U.S. ENVIRONMENTAL PROTECTION AGENCY

THE PRECISION AND ACCURACY OF ENVIRONMENTAL MEASUREMENTS FOR THE BUILDING ASSESSMENT SURVEY AND EVALUATION PROGRAM

Previously submitted date: March 31, 1999
September 29, 2000

Prepared For:

Ms. Laureen Burton
U.S. Environmental Protection Agency
Indoor Environments Division
501 3rd Street N.W.
Washington, DC 20001

Prepared By:

Environmental Health & Engineering, Inc.
60 Wells Avenue
Newton, MA 02459-3210

EH&E Report #11995
September 28, 2001

P:\11995\Precision and Accuracy\REPORT7.DOC

©2001 by Environmental Health & Engineering, Inc.
All rights reserved

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS (Continued)

LIST OF ABBREVIATIONS & ACRONYMS

CO	carbon monoxide
CO_2	carbon dioxide
CFU/m^3	colony forming units per cubic meter
CV	coefficient of variation
DNPH	2,4-dinitrophenyl hydrazine
EH&E	Environmental Health & Engineering, Inc.
EPA	U.S. Environmental Protection Agency
GC/MS	gas chromatography/mass spectrometry
GM	geometric mean
GSD	geometric mean standard deviation
HPLC	high performance liquid chromatography
IAQ	indoor air quality
IEQ	indoor environmental quality
LOD	limit of detection
LOQ	limit of quantitation
MEA	malt extract agar
MEM	microenvironmental exposure monitors

TABLE OF CONTENTS (Continued)

NDIR	non-dispersive infrared radiation
NIST	National Institute for Standards and Technology
PE/PD	performance evaluation and performance demonstration
p.i.	prediction interval
ppm	parts of vapor or gas per million parts of air by volume
QA	Quality Assurance
QAPP	Quality Assurance Project Plan
QC	Quality control
TLV	threshold limit value
TSA	trypticase soy agar
TWA	time-weighted average
$\mu g/m^3$	micrograms per cubic meter
VOC	volatile organic compound

1.0 OBJECTIVE

The objective of this report is to evaluate the precision and accuracy of the environmental measurements collected during the Building Assessment Survey and Evaluation (BASE) Study. Moreover, this report will provide available indicators of precision and accuracy for this data set.

This report evaluates the precision and accuracy of ten time-integrated sampling methods and four continuous sampling methods used in the BASE study. These sampling methods are summarized in Table 1.1, along with the name used to refer to them throughout the report.

Table 1.1 Environmental Parameters Discussed within this Report

Parameter	Designation within Report
Time-integrated	
VOCs by canister method	Canister VOCs
VOCs by multisorbent method	Multisorbent VOCs
Fungal spores by Burkard Sampler	Burkard
Culturable fungi by Andersen Sampler	Fungi
Culturable mesophilic bacteria by Andersen Sampler	Mesophilic
Culturable thermophilic bacteria by Andersen Sampler	Thermophilic
Formaldehyde	Formaldehyde
Acetaldehyde	Acetaldehyde
Particulate matter less than 10 μm aerodynamic diameter	PM_{10}
Particulate matter less than 2.5 μm aerodynamic diameter	$PM_{2.5}$
Continuous	
Carbon dioxide	CO_2
Carbon monoxide	CO
Temperature	Temperature
Relative Humidity	RH

VOC = volatile organic compound
μm = micron

For the time-integrated samples, precision is evaluated by calculating residual differences of duplicate samples and accuracy is evaluated by reviewing results of spiked and blank samples. For the continuous data, precision and accuracy are evaluated by comparison of differential results between instrument readings and gases of known concentration or reference instruments. These sampling and validation procedures are described in greater detail in other documents (i.e., EPA *BASE Study Protocol* and the *BASE Quality Assurance Project Plan*).

2.0 INTRODUCTION

The BASE Program was a cross-sectional information-gathering study sponsored by the U.S. Environmental Protection Agency (EPA) from 1993 through 1998. It was designed to address a significant information gap in indoor environmental information by collecting baseline data from one hundred public and commercial office buildings in the United States with respect to key characteristics of indoor air quality (IAQ) and occupant perceptions (EPA 1994).

One of the primary goals of the BASE program was to provide a foundation for researchers to understand the role that environmental variables played in causing health symptom reports and how these variables affected the perception of IAQ. However, before these relationships can be developed, it is essential to have an understanding of how data quality was impacted by measurement error. For all practical purposes, assessing error in a data set is as important as the data itself (ACGIH 1995).

Two central themes in the discussion of error analysis are the concepts of precision and accuracy. For application to environmental sampling, the definitions of accuracy and precision can be defined as follows:

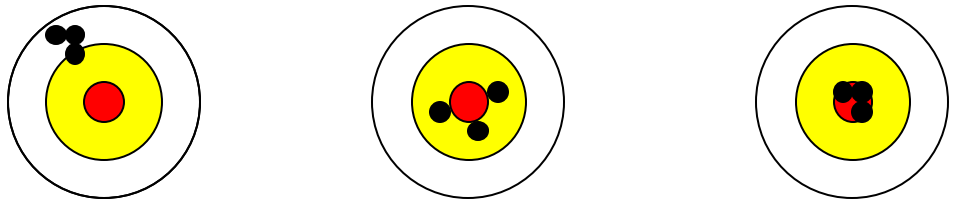
accuracy - the degree of correctness with which a measurement reflects the true value of the parameter being assessed

precision- the degree of variation in repeated measurements of the same quantity of a parameter

For example, if ten measurements for a given parameter are taken at the same time at the same location by the same method, the accuracy would be indicated by how well the average of the ten measurement results reflects the actual concentration present and the precision would be indicated by the variation in the results of the ten measurements.

Using a classic example of marksmanship, Figure 2.1 distinguishes the concepts of precision and accuracy.

Figure 2.1 Precision and Accuracy



Precise but Not Accurate | Accurate but Not Precise | Precise and Accurate

A high degree of precision and accuracy do not necessarily occur simultaneously in a process, as illustrated in the previous figure. Measurements may have a high degree of precision, while not being very accurate. Conversely, a set of data may have high accuracy but lack precision. When results are both precise and accurate, confidence in data quality is maximized.

One factor that may affect precision and accuracy is bias in the sampling methods, such as sample media contamination or continuous data instruments reading higher than actual concentrations. Poor precision and accuracy are seen in high variability of sampling results.

In order to address those issues that would affect confidence in the BASE data, a Quality Assurance Project Plan (QAPP) was developed in conjunction with the BASE protocol to ensure the collection of reproducible and accurate data and to provide guidelines that would allow the investigators to recognize and control factors that would compromise the data quality. Table 2.1 describes the terms and quality assurance measures that will be addressed in the body of this report to measure precision and accuracy as they relate to the BASE environmental measurements.

Table 2.1 Quality Assurance and Quality Control Measures of Precision and Accuracy Included in this Report

Sample Type	Definition
Integrated Samples	
Duplicate Sample	A sample run concurrently with a field sample to assess repeatability of methods as well as a redundant safeguard in case a sample is voided.
Field Blank	A sample prepared by the field team using the procedure for preparing integrated sample, but which is not run as a regular sample. This is sent blindly to the laboratory. The results of the field blanks can be used to determine whether there was any contamination in the preparation or shipping process of the other samples or during the analysis of the samples by the laboratory.
QC Spike	A sample that is spiked by the analytical laboratory, sent to the field team, then sent back to the analytical laboratory within a regular sample shipment. This sample is sent blindly to the laboratory. The results of a QC spike can be used to determine if laboratory analytical procedures are accurate.
Continuous Data	
Verification	Verification in the field includes: weekly side-by side comparisons of similar field instruments and daily comparisons of sensors' response to known standards (*e.g.*, zeros and spans). These decisions are based upon EPA's Large Buildings QAPP Table 8.4, Data Acceptability Criteria for Validation.

Table 2.2 illustrates the parameters and applicable quality assurance measures under which precision and accuracy could be evaluated. Note that continuous data requires slightly different performance measurements, which will be discussed later in the report. Radon is not discussed in this report because it was rarely present above its limit of quantitation. Bulk and dust samples were also omitted because no meaningful methods of estimating precision or accuracy exist for these types of samples. For the continuous data, light and noise data were not evaluated because no standards or methods for estimating precision or accuracy exist.

Table 2.2 Available Measures of Precision and Accuracy with Respect to the BASE Environmental Samples

Integrated Data			
Method	**Precision**	**Accuracy**	
	Duplicate	Blank	QC Spike
VOC: Canister	A	A	A
VOC: Multisorbent	A	A	A
Aldehydes	A	A	A
PM_{10} and $PM_{2.5}$	A	A	NA
Radon	A	A	NA
Fungi, mesophilic bacteria, and thermophilic bacteria	A	A	NA
Burkard	A	NA	NA
Dust	NA	NA	NA
Bulk	NA	NA	NA
Continuous Data			
	Precision	Accuracy	
Carbon Dioxide	A	A	
Carbon Monoxide	A	A	
Temperature	A	A	
Relative Humidity	A	A	
Light	NA	NA	
Sound	NA	NA	

VOC = volatile organic compound
A = Available
NA = Not available

3.0 INTEGRATED DATA ERROR ANALYSIS

While the sampling and analytical methods for the integrated parameters vary widely, the analyses of precision and accuracy do not. As reported in Table 2.2, duplicates, lab samples, field blanks, and spikes were the tools used for measuring confidence with the integrated samples.

It is important to understand the manner in which the samples were collected and analyzed. This will affect the limitations of evaluating precision and accuracy for some of the samples. Appendix A contains sections describing environmental measurements with respect to the sample collection and analytical methods used in the study.

The following sections describe methods used to estimate the precision and accuracy of integrated samples collected as part of the BASE study.

For the purposes of comparison of the data analyzed in this report, Table 3.1 presents the nominal limits of quantitation of the integrated sample methods discussed in this report, expressed as quantity per sample and as quantity per unit volume of air sampled. The volume used was the median sample volume. For the volatile organic compounds (VOCs) by both canister and multisorbent methods, the median limit of quantitation (LOQ) across all analytes was used. The LOQs presented are as reported by the respective analytical laboratory. EH&E did not review the methods used by the laboratories to derive these values. The nominal LOQ per cubic meter of air sampled assumes the sample volume specified in the BASE protocol was collected. The median volume was chosen to provide a simple indicator of the LOQ across all BASE samples. Most sample volumes varied by only ±20% across the study.

Table 3.1 Summary of Nominal Limits of Quantitation for BASE Sampling Methods

Analyte	Units	Nominal LOQ per Sample	Nominal LOQ per m^3 of Air Sampled
VOC canister	µg	0.0041	1.8
VOC multisorbent	µg	0.0010	0.34
Formaldehyde	µg	0.040	0.40
Acetaldehyde	µg	0.050	0.50
PM_{10} and $PM_{2.5}$	µg	10	1.0
Fungi	CFU	1.0	18
Mesophilic bacteria	CFU	1.0	18
Thermophilic bacteria	CFU	1.0	18
Burkard	spores	1.0	25

Throughout this report, box plots are used to summarize populations of data. Figure 3.1 presents a sample box plot. This figure summarizes the distribution of values from a population. Each box plot in this report includes five points from the data population: the 5^{th}, 25^{th}, 50^{th} (median), 75^{th} and 95^{th} percentiles. Each of these percentiles will always occupy the same position in the box plot. The lowest point is always the 5^{th} percentile; the second from the lowest is always the 25^{th}, the middle is always the 50^{th}; the second from the top is the 75^{th}; and the top is the 95^{th}. In some cases, the lower percentiles may have the same value and, therefore, these points may not appear distinct on the box plot.

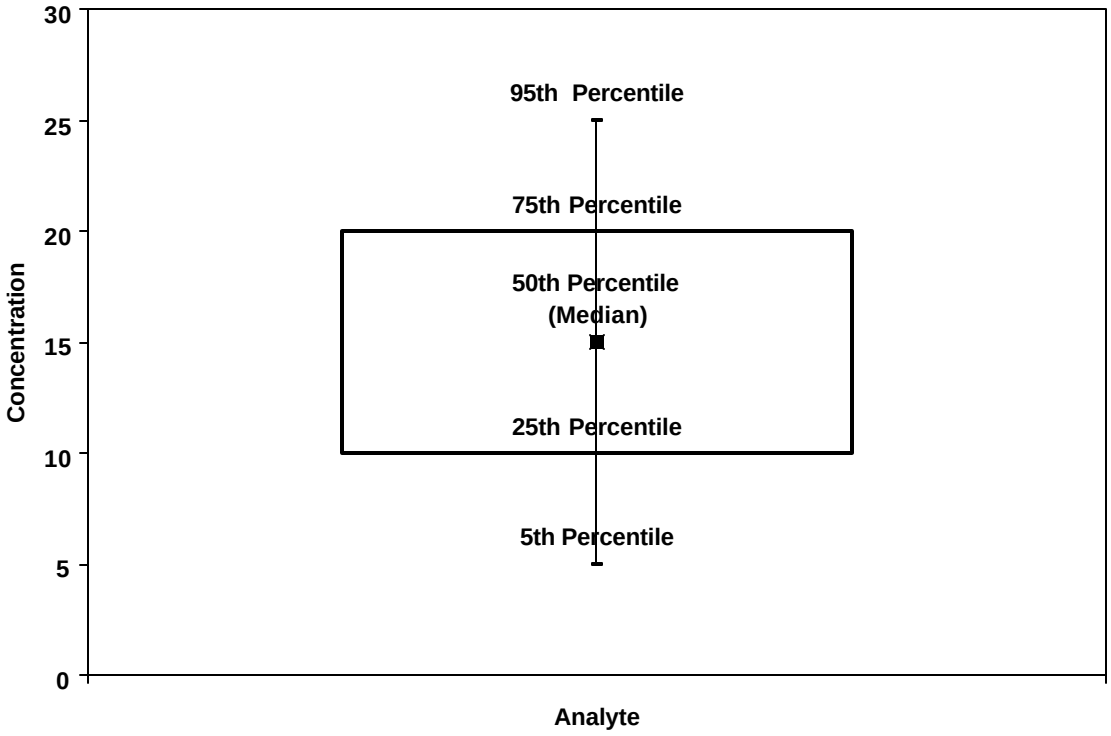


Figure 3.1 Sample box plot indicating the values represented by each point in the box . This chart includes the 5th, 25th, 50th (median), 75th and 95th percentiles.

Percentiles reflect the cumulative frequency at which a value is found in a data set. For example, if the 95th percentile of a population is 16, that means that 95% of the values in the population are equal to or less than 16. In addition to box plots, some of the charts in this report plot cumulative frequency (percentiles) against duplicate residual values. These charts indicate the percent of the population that has duplicate residuals less than or equal to given concentrations. All figures of integrated data error analysis are presented in Appendix C of this report. Summary statistics for much of the data in this section is also presented in Tables 5.1, 5.2, and 5.3 found in Section 5.0 of this report.

3.1 PRECISION

Previous analyses of BASE data relied on the coefficient of variation (CV) between duplicate samples as an indicator of agreement ([standard deviation/mean]*100). However, it was determined that this may not be the most representative or appropriate method for describing precision. When the sample is limited to a pair, the CV is related only to the ratio of the pair. For example, duplicate pairs of 0.1 and 1.0 parts per billion (ppb) and 10 and 100 ppb both have the same CV, but in practical terms, the differences

at the low end are less significant than those of the greater concentrations. Overall, this statistic tends to be misleading by overemphasizing differences at low-end concentrations.

As a result, the duplicate residuals were used as an indicator of the precision of the duplicate pairs. The duplicate residual is simply the absolute difference between the concentrations of the duplicate and its co-located sample. Using this method, the difference between 1.0 and 10 ppb is much more significant than that of 0.1 and 1.0 ppb.

The duplicate residual method of precision analysis is also subject to a type of bias. For example duplicates of 10 and 20 ppb and 290 and 300 ppb both have residuals of 10 ppb. However, the significance of the difference on the higher concentrations is exaggerated because, in practical terms, a 10 ppb difference is negligible at higher concentrations. As a result, the duplicate residual method tends to overstate results at high concentrations.

Because the majority of environmental data collected in the BASE study appear to be lognormal, the duplicate residual method provides a more representative measure of precision than the CV.

The analysis of duplicate precision included only duplicate pairs where at least one of the concentrations was greater than the LOQ. In addition, pairs were excluded where one value was below the LOQ and the second value was less than twice the LOQ. This was done because these differences cannot be quantified accurately and generally are much smaller and, therefore, less significant than the pairs included. The effect of this exclusion is likely to bias the results toward slightly higher values since many of the smaller residuals have been excluded.

For VOCs and microbial organisms, duplicate residuals are presented as an aggregate across all individual compounds, species, or types. VOC samples, for example, were analyzed for between 29 and 56 target compounds over the course of the study. For each pair of duplicate samples, the results of each analyte were compared. If at least one

of the analyte's results was greater than twice its LOQ, then that analyte-duplicate pair would be included in the analysis. For VOC canisters, 198 pairs of duplicate samples and 2,818 analyte-duplicate pairs were available for analysis. A similar logic was applied to analysis of the precision of the multisorbent VOCs and the microbiological samples.

For each method, cumulative frequencies of duplicate residuals were derived for each season of the BASE study from Summer 1994 through Summer 1998. The samples were segregated in this manner to assess changes in precision over time.

The distribution of duplicate residuals appeared to be lognormal. Hence, the geometric mean (GM) and geometric standard deviation (GSD) are used to characterize these distributions. The formulas used to compute the GM and GSD are presented below.

$$GM = 10^{\sum \log(xi)/n}$$

$$GSD = 10^{\sum [\log(xi) - \log (GM)] / (n - 1)}$$

Where x_i is the value of the individual observations and n is the number of observations.

Figure C1 presents the cumulative frequencies of duplicate residuals for VOCs sampled by SUMMA® canisters. This includes results from comparison of 2,818 analyte-duplicate pairs. The geometric mean across all seasons was 1.1 micrograms per cubic meter ($\mu g/m^3$) with a geometric standard deviation of 5.0 $\mu g/m^3$. As is evident from Figure C1, precision for SUMMA® canister samples was worse in the earlier summer seasons. Based on these data, the upper limit of the 95% prediction interval for individual duplicate residuals is 16 $\mu g/m^3$.

Figure C2 presents the cumulative frequencies of duplicate residuals for VOCs sampled by multisorbent methods. This data comprises 3,360 analyte-duplicate pairs. This number is greater than the number of pairs by the SUMMA® canister method, despite the fact that multisorbent samples were collected in only seventy of the one hundred buildings where SUMMA® canister samples were collected. This is due in large part to the substantially lower limits of detection for multisorbent samples. Duplicate residuals during the summer seasons are slightly greater than during winter seasons. The

geometric mean (GM) residual was 0.25 $\mu g/m^3$, and the geometric standard deviation (GSD) was 5.0 $\mu g/m^3$. The GM residual from multisorbent sampling was statistically significantly lower than that from canister sampling. Based on these data, the upper limit of the 95% prediction interval for individual duplicate residuals is 3.5 $\mu g/m^3$.

Figure C3 presents the cumulative frequency of duplicate residuals for formaldehyde sampling. The analysis included 191 duplicate pairs. The GM duplicate residual was 0.61 $\mu g/m^3$, and the GSD was 4.3 $\mu g/m^3$. The variation between seasons appears greater when compared to VOC results. However, the greater variation may be attributable to the substantially smaller set of data points, less than 200 for formaldehyde versus approximately 3,000 for the VOCs. When considered across all seasons, however, the formaldehyde GSD was less than the VOC GSDs. Based on these data, the upper limit of the 95% prediction interval for individual duplicate residuals is 6.8 $\mu g/m^3$.

Figure C4 presents the duplicate residuals for acetaldehyde sampling. These data include 167 duplicate pairs. The GM residual was 0.28 $\mu g/m^3$, and the GSD was 3.4 $\mu g/m^3$. As with formaldehyde, acetaldehyde appears to have greater variation across the seasons, but this is likely a function of the fewer available data points. Based on these data, the upper limit of the 95% prediction interval for individual duplicate residuals is 2.1 $\mu g/m^3$.

Figure C5 presents cumulative frequency of duplicate samples for particulate matter samples. This analysis includes a total of 394 duplicate pairs. The GM for PM_{10} was 0.98 $\mu g/m^3$ and the GSD was 2.9 $\mu g/m^3$. For $PM_{2.5}$, the GM was 0.99 $\mu g/m^3$ and the GSD was 3.2 $\mu g/m^3$. Based on these data, the upper limit of the 95% prediction interval for individual duplicate residuals is 5.7 $\mu g/m^3$ for PM_{10} and 6.6 for $PM_{2.5.}$

Figure C6 presents cumulative frequency of duplicate residuals for culturable fungal types from the 5-minute samples by Andersen sampler. The GM was 15 colony forming units per cubic meter (CFU/m^3) and the GSD was 3.2 CFU/m^3. This analysis included 1,311 duplicate pairs. Based on these data, the upper limit of the 95% prediction interval for individual duplicate residuals is 100 CFU/m^3.

Figure C7 presents the cumulative frequency of duplicate residuals from the Burkard samples for fungi. Only 151 pairs were available for inclusion in this analysis. The GM was 130 spores per cubic meter (spores/m^3) and the GSD was 4.7 spores/m^3. It should be noted that the value of these residuals is somewhat inflated because the samples were not true duplicates, but rather consecutive samples taken at the same location. The 95% prediction interval for individual duplicate residuals was 1,400 spores/m^3.

Figure C8 presents the cumulative frequency of duplicate residuals for mesophilic bacteria from 5-minute Andersen samples. The GM was 16 CFU/m^3, and the GSD was 2.8 CFU/m^3. There appears to be little variation across the seasons. Based on these data, the upper limit of the 95% prediction interval for individual duplicate residuals is 87 CFU/m^3.

Figure C9 presents the cumulative frequency of duplicate residuals for thermophilic bacteria from 5-minute samples. This analysis was comprised of only 104 points because thermophilic bacteria were rarely detected. The GM was 11 CFU/m^3 and the GSD was 3.7 CFU/m^3. Based on these data, the upper limit of the 95% prediction interval for individual samples is 47 CFU/m^3.

3.2 ACCURACY

Compared to estimates of precision, there is far less certainty in available estimates of accuracy with respect to the environmental integrated data collected from the BASE study. As mentioned in Table 2.1, field blanks and spikes were the only quality assurance devices employed for assessing how close the data points conformed to truth. Spiked samples were only available for the canister, multisorbent, and aldehyde sampling methods. No spiked samples were available for the other sampling methods because of the difficulty in actually creating a meaningful spike. Also, most samples were not available for the Burkard method.

Most measures of accuracy appear to be normally distributed. Hence, the arithmetic mean and standard deviation are used when discussing these results. It should also be noted that many of the estimates of accuracy have considerable variation in their values. However, this variation reflects the precision, which is best estimated by the field

duplicate samples discussed in Section 3.1. For estimates of accuracy, the mean is by far the most important value.

3.2.1. Field Blanks

Field blanks, while vital for identifying contamination issues in preparation and shipping, are limited with respect to their indication of accuracy. Field blanks are a measure of how close the data points in exposed samples conformed to a zero value; however, with environmental data, the determination of accuracy is best measured at values greater than zero; hence, spiked samples are ultimately more useful for measuring accuracy. As noted earlier, blank samples were not available for the Burkard method. A limited number of blanks were analyzed for the canister method for VOCs, but their use was discontinued early in the study, and they are not included in this analysis.

Figures C10 through C13 demonstrate a measure of accuracy using field blanks. Mean values and percentile concentrations are presented. Figure C10 presents the results of analyses of multisorbent field blank samples. Only sixteen of the VOC analytes are presented because only those sixteen VOCs had more than one quantifiable result on a field blank. Most of those sixteen VOCs were detected infrequently. The mean field blank results of several compounds were greater than their LOQ from Table 3.1. The compounds most frequently detected above the detection limit were acetone, benzene, trichlorofluoromethane, and nonanal.

Figure C11 summarizes the results of field blank analyses for formaldehyde and acetaldehyde. The mean blank concentration was 0.022 μg/sample for formaldehyde and 0.054 μg/sample for acetaldehyde. The acetaldehyde mean was slightly greater than the LOQ for this method of 0.05 μg/sample.

Figure C12 summarizes the results of analysis of field blanks from particulate samples (both PM_{10} and $PM_{2.5}$ sampling). The mean value was 2.2 μg/sample, and the standard deviation was 4.0 μg/sample. Less than 25% of values exceeded the method LOQ of 10 μg/sample.

Figure C13 presents the results of analyses of field blanks for thermophilic bacteria, mesophilic bacteria, and fungi. The mean was 0.57 for fungi, 1.02 for mesophilics, and 0.55 for thermophilics. The greater contamination of the mesophilic bacteria plates likely

originates from the field personnel handling the plates, since humans are common sources of mesophilic bacteria.

3.2.2 Quality Control Spikes

Spiked samples can be used to estimate accuracy. Most of these spikes were generated by injecting known masses of the analyte of interest into a container and then drawing the air from the container into the sample media.

Quality control spikes were available for only three integrated sampling methods (SUMMA® canisters, multisorbent samplers, and aldehyde samplers). There are several limitations to how well the spiked samples can estimate accuracy. First, for the VOCs, only a subset of the target analytes is represented in the spikes. For the VOC samples, over fifty compounds were included on the target analyte list, but only fifteen of those fifty were among the compounds spiked for the canisters, and only six of these fifty were used for the multisorbent samples. Second, almost all spike quantities were the same, 200 μg/sample. The spiked quantities tend to be much higher than quantities typically present on the BASE samples. Third, the spike was not applied to the sample using the same collection methods as field samples. Actual field analyte recovery from the samplers may have differed from those of spiked samples. The ideal spiked sample would have been to collect samples in a chamber with known concentrations of target analytes, thus replicating field collection methods.

The charts below present the percentile distribution of quantities of spikes recovered. The recovered quantity is expressed as percentage recovered. For each analyte, the 5th, 25th, 50th, 75th, and 95th percentiles are presented.

Figure C14 presents the results of spiked SUMMA® canister samples. Of the fifteen analytes, most analytes have a mean value near 100%, and 90% of the spikes have values within ±20%. The significant exceptions are acetone, 1,1-dichloroethene, and o-xylene, which all had a substantial number of spike recoveries greater than ±20%.

Figure C15 summarizes the results of analysis of spiked multisorbent tubes. For the six spiked analytes, all mean values were near 100% and variation in over 90% of spikes was less than ±20%. Note that only three of the target analytes overlap with the compounds spiked on the SUMMA canisters. The three compounds that showed the

greatest variation in spike recovery from the canister samples were not included in the compounds spiked on multisorbent media.

Figure C16 presents the results of spiked aldehyde tubes. Mean recoveries were 97% for formaldehyde and 99% for acetaldehyde. Over 90% of the data fell with ±20%.

4.0 CONTINUOUS DATA ERROR ANALYSIS METHODS

Evaluating the precision and accuracy of continuous monitoring parameters is based upon predictable and repeatable instrument response as well as on documented comparisons to National Institute of Standards & Technology (NIST) traceable standards. As per the BASE protocol and QAPP, each sensor not only underwent seasonal multipoint calibrations, but weekly calibration checks and daily verification assessments. This resulted in approximately 100 calibrations and 300 verifications on each continuous monitoring sensor throughout the duration of the study. All figures of continuous data error analysis are presented in Appendix D of this report. Summary statistics are presented in Table 5.4 found in Section 5.0 of this report.

4.1 PRECISION AND ACCURACY

The concepts of precision and accuracy are more inherent to continuous monitoring than time-integrated sampling, as sensor performance can be more directly evaluated and/or manipulated. While integrated data requires laboratory intervention and subsequent analysis, continuous data records can be adjusted and tracked in the field.

At the beginning of each field week, instrument response was set or calibrated to a primary standard device, such as a zero or span gas, or mercury thermometers and hygrometers. Each day during the field week, the performance of each sensor was measured or verified against these primary standards. This method allows both the repeatability (precision) and the instrument accuracy to be recorded. The following analyses summarize all recorded validation checks.

These daily validation performance verifications were completed to determine if sensors were performing within established limits in accordance with the QAPP. These primary and secondary performance ranges were essential in establishing consistent sensor response throughout the study. The primary and secondary criteria used in the BASE study for continuous measurements are summarized in Table 4.1.

Table 4.1 Primary and Secondary Ranges for Each Parameter

Parameter	Primary Range[1]	Secondary Range[2]
Temperature	± 1.0°C	± 2.0°C
Relative humidity	± 5% RH	± 7% RH
Carbon dioxide	zero ± 50 ppm span ± 75 ppm	zero ± 75 ppm span ± 150 ppm
Carbon monoxide	zero ± 2 ppm span ± 3 ppm	zero ± 3 ppm span ± 5 ppm

[1] Instrument is "in spec" if within this range.
[2] Values within this range require justification for continued use; outside this range renders data unusable.

4.1.1 Carbon Dioxide

The accuracy of these sensors was measured by a comparison of instrument response to a zero and to a span gas (usually 0 parts per million [ppm] CO_2 and 1,000 ppm CO_2), which were introduced to the sensors repeatedly throughout the week. Precision is measured using repeated verification records, which track the degree of deviation.

Using the measured difference between expected and actual instrument response, Figures D1-D4 demonstrate overall carbon dioxide in-field sensor performance. Primary and secondary performance acceptability standards for zero gas were ± 50 and ±75 ppm, respectively. Span gas acceptability ranges were ±150 and ±175 ppm. Note that the primary acceptance range is marked at the y-axis with a solid line, while the secondary range is marked with a dotted line. Statistical analysis was performed across all sensors despite changes in instrumentation, and data records include all points recorded from log sheets.

The 5th, 25th, 50th, 75th, and 95th percentiles are represented on the box and whisker plots. The accuracy and precision of these sensors were measured by comparison zero and span gas (usually 0 ppm and 350 ppm for outdoors or 1,000 ppm for indoors).

Figures D1 and D2 present the indoor and outdoor sensor response to zero gas concentrations. The response across all years is fairly consistent. With a few exceptions in 1994, over 90% of sensor responses fell within the primary criteria of ±50 ppm.

Figures D3 and D4 summarize the indoor and outdoor CO_2 responses to span gas. In general, the indoor span gas was approximately 1,000 ppm CO_2, and the outdoor span was 350 ppm. However, on some occasions, such as during a shortage of available span gas, a different span concentration may have been used. All span results are summarized as the difference between the instrument response and the span concentration. For both indoor and outdoor sampling, nearly 90% of all samples were within the primary criteria of ±75 ppm.

4.1.2 Carbon Monoxide

The accuracy of these sensors was measured by a comparison of instrument response to a zero and to a span gas (usually 0 ppm CO and 10 ppm CO), which were introduced to the sensors repeatedly throughout the week. Precision is measured using these repeated verification records, which track the degree of deviation.

Using the measured difference between expected and actual instrument response, Figures D5-D8 demonstrate overall carbon monoxide in-field sensor performance. Primary and secondary performance acceptability standards for zero gas were ±2 and ±3 ppm, respectively. Span gas acceptability ranges were ±3 and ±5 ppm. Note that the primary acceptance range is marked at the y-axis with a solid line, while the secondary range is marked with a dotted line. Statistical analysis was performed across all sensors despite changes in instrumentation, and data records include all points recorded from log sheets.

The 5th, 25th, 50th, 75th, and 95th percentiles are represented on the box and whisker plots. The accuracy and precision of these sensors were measured by a comparison to a zero and span gas (usually 0 ppm and 10 ppm), which were introduced at least daily throughout the week.

Figures D5 and D6 summarize the response of indoor CO sensors to zero and span gas respectively. In both cases, the mean differential response was near zero, and nearly 95% of the responses were within the primary response standard of ±2 ppm of the calibration gas for the zero and ±3 ppm for the span gas.

Figures D7 and D8 present the response of outdoor CO sensors to zero and span gas, respectively. Although different sensors were used indoors and outdoors, the response

was similar. The mean differential response was -0.022 ppm for zero gas and 0.50 ppm for span gas.

4.1.3. Temperature

The accuracy of these sensors was measured by comparisons with NIST-traceable devices, either mercury thermometers or digital thermo-hygrometers. These actions were repeated, at least daily, throughout the week. Precision and accuracy were evaluated by these repeated comparisons, which track the degree of deviation.

Figures D9 and D10 summarize the difference between BASE temperature sensors and reference thermometers. For the indoor sensors, nearly 90% of all results fell within the primary criteria of a difference of less than ±1 °C. For outdoors sensors, more responses failed to meet the acceptability criteria.

4.1.4. Relative Humidity

The accuracy of these sensors was measured by comparisons with NIST traceable digital thermo-hygrometers. These actions were repeated throughout the week. Precision was measured by these repeated verification records, which track the degree of deviation.

Figures D11 and D12 summarize humidity in-field sensor performance. Indoors, over 90% of all measurements fell with ±5% of the reference measurement. Outdoors, however, in many seasons over 5% of the sensors failed to meet the primary and secondary criteria for relative humidity. This may have been associated with the more extreme humidities seen outdoors, which may have affected both the measurement device and QA reference units.

5.0 SUMMARY AND CONCLUSIONS

Error is inherent in measurement because it embodies such things as the precision of both measuring and analytical tools, their proper adjustment, and competent application. Analyses of the magnitude of errors is necessary in examining the suitability and limitations of methods or equipment used to obtain, portray, and utilize, an acceptable result.

The environmental parameters sampled as part of the BASE study, while collected and analyzed using standardized methods and quality assurance protocols, can not be meaningfully interpreted without assessing confidence in the quality of the result. Two such measures of confidence are the concepts of precision and accuracy.

Table 5.1 summarizes the results of duplicate residual analyses for all integrated sampling methods. Most GM duplicate residuals were low, near the LOQ for the given sampling methods. The notable exception was the GM for the Burkard sampling method, which was several times its LOQ. It should be cautioned that of all the samples analyzed, only the Burkard samples were not true duplicates. Due to limited availability of one Burkard sampler per building for this study, the Burkard samples were not collected simultaneously, but instead consecutively. This likely inflated the estimates of precision for the Burkard samples. It should also be cautioned that when interpreting the upper 95% prediction interval for all parameters, many of these values may be associated with higher environmental concentrations and thus may be less significant than they appear. For example, the upper 95% prediction interval for canister VOC residuals is 16 $\mu g/m^3$. If that 16 $\mu g/m^3$ is associated with a pair of 100 and 116 $\mu g/m^3$, it is less significant than if it had been associated with a pair of 1 and 17 $\mu g/m^3$, and it is more likely to be associated with a pair with higher concentrations.

Table 5.1 Summary of Duplicate Residuals

Analyte	Units	# Pairs Analyzed	Geometric Mean	Geometric Standard Deviation	Upper 95% Prediction Interval of Individual Residuals
Canister VOCs	µg/m^3	2,818	1.1	5.0	16
Multisorbent VOCs	µg/m^3	3,360	0.25	5.0	3.5
Formaldehyde	µg/m^3	191	0.61	4.3	6.8
Acetaldehyde	µg/m^3	167	0.28	3.4	2.1
PM_{10}	µg/m^3	198	0.98	2.9	5.7
$PM_{2.5}$	µg/m^3	196	0.99	3.2	6.6
Bacteria (Meso 5 min)	CFU/m^3	959	16	2.8	87
Bacteria (Therm 5 min)	CFU/m^3	104	11	2.4	47
Fungi (5 min)	CFU/m^3	1,311	15	3.2	100
Burkard	spores/m^3	151	108	4.7	1,400

CFU/m^3 = colony forming unit per cubic meter
µg/m^3 = micrograms per cubic meter
spores/m^3 = spores per cubic meter

Table 5.2 summarizes the results of analyses of field blank samples. Only sixteen of the more than fifty VOCs analyzed for on multisorbent samples were detected on the field blank samples. Of those sixteen, most were infrequently detected and were at low levels. Likewise, the other sampling methods showed little contamination of the blanks.

Table 5.2 Summary of Field Blank Analyses

Analyte	Units	# Samples Analyzed	Mean	Standard Deviation
Multisorbent VOCs				
1,1,1-Trichloroethane	ng/sample	67	1.2	0.10
3-Methyl pentane	ng/sample	67	2.2	0.20
Acetone	ng/sample	67	7.1	0.61
Benzene	ng/sample	67	3.4	0.34
Ethyl acetate	ng/sample	67	0.78	0.10
Hexanal	ng/sample	41	1.4	0.25
Methylene chloride	ng/sample	67	2.4	0.04
Nonanal	ng/sample	41	1.9	0.15
Phenol	ng/sample	41	0.67	0.06
Styrene	ng/sample	67	0.57	0.05
Toluene	ng/sample	67	0.73	0.06
Trichloroethene	ng/sample	67	0.62	0.05
Trichlorofluoromethane	ng/sample	67	5.7	0.29
m- & p-Xylenes	ng/sample	67	0.53	0.01
n-Dodecane	ng/sample	67	0.87	0.03
n-Hexane	ng/sample	41	2.0	0.44
Other Analytes				
Formaldehyde	ng/sample	100	0.022	0.0070
Acetaldehyde	ng/sample	86	0.054	0.021
PM	ng/sample	100	2.2	4.0
Mesophilic bacteria	CFU/sample	108	1.0	0.94
Thermophilic bacteria	CFU/sample	112	0.55	0.11
Fungi	CFU/sample	112	0.57	0.18

ng = nanogram
CFU = colony forming unit

Table 5.3 summarizes the results of analyses of the spiked samples. Spiked samples were available for only canister, multisorbent, and aldehyde samples. Across all spiked analytes, the mean recovery differed by less than 12% from the quantity spiked. Most of the standard deviations fall below 11%, with the exception of three of the canister VOCs. However, as noted before, the mean reflects accuracy, and the standard deviation reflects precision of the method.

Table 5.3 Summary of Spiked Sample Results

Analyte	# Spikes Analyzed	Mean	Standard Deviation
		% of Spike Recovered	
Canister VOCs			
Acetone	79	109	27
Benzene	99	107	8.3
2-Butanone	87	108	11
Chlorobenzene	87	100	11
Chloroform	62	110	7.3
1,1-Dichloroethene	46	112	19
cis-1,2-Dichloroethene	46	107	10
4-Methyl-2-pentanone	87	104	10
Tetrachloroethene	99	97	11
Toluene	99	106	10
1,1,1-Trichloroethane	99	107	9.3
Trichloroethene	99	103	10
Vinyl chloride	53	106	12
m- & p-Xylenes	99	102	14
o-Xylene	12	99	29
Multisorbent VOCs			
Benzene	25	98	10
n-Hexane	13	103	13
Styrene	25	97	9.1
Tetrachloroethene	25	96	8.9
Trichloroethene	25	94	7.8
p-Xylene	25	99	8.3
Aldehydes			
Formaldehyde	59	97	3.6
Acetaldehyde	41	99	2.8

Table 5.4 summarizes the responses of continuously monitored parameters to zero and span checks. These results indicate that, in most cases, the results were both precise and accurate. In most cases, the mean response was near zero, indicating that results were accurate. Almost all data met or were near the primary target criteria set for BASE continuous monitoring over 90% of the time. The greatest exception was the outdoor

temperature sensors, where the 90% prediction interval for individual measurements ranged between –3.8 and 4.1 ^{o}C. The 90% prediction interval of outdoor relative humidity was slightly greater than the secondary criteria. Other parameters that were slightly outside the primary criteria but within the secondary criteria 90% of the time were indoor CO zero, outdoor CO span, indoor CO_2 span, and indoor temperature.

Table 5.4 Summary of Differential Responses of Continuous Measurement Sensors to Known Concentrations or to Reference Instruments

Analyte	Location	Comparison	Units	n	mean	SD	Lower 90 % of p.i. Individual Measurements	Upper 90% of p.i. Individual Measurements
CO	Indoor	Zero	ppm	1,573	0.068	1.4	-2.2	2.4
CO	Indoor	Span	ppm	1,576	0.30	1.5	-2.2	2.8
CO	Out	Zero	ppm	476	-0.02	0.99	-1.6	1.6
CO	Out	Span	ppm	475	0.50	1.6	-2.2	3.2
CO_2	Indoor	Zero	ppm	1,672	-7.34	25	-49	34
CO_2	Indoor	Span	ppm	1,704	-5.11	49	-86	76
CO_2	Out	Zero	ppm	466	0.81	20	-32	33
CO_2	Out	Span	ppm	467	1.49	26	-41	44
Temp.	In	Reference	deg C	1,288	0.01	0.89	-1.5	1.5
Temp.	Out	Reference	deg C	253	0.35	2.4	-3.8	4.1
RH	In	Reference	%	1,289	0.14	1.4	-1.9	2.6
RH	Out	Reference	%	254	-0.37	4.5	-7.7	7.0

p.i. = prediction interval
SD = standard deviation

These data provide useful estimates of the precision and accuracy of environmental measurements collected in the BASE study. They can be used by both researchers analyzing the BASE data and by researchers designing studies that will use these sampling methods.

It must also be noted, however, that these analyses of precision and accuracy have limitations as well. For the time-integrated data, the estimates of precision are extensive but could be explored further to make them more useful. It may be useful to characterize changes in precision as a function of changes in concentration. These estimates of precision may not apply to samples where actual concentrations are substantially different from the BASE results. For the VOCs and the microbiological organisms, it may also be valuable to characterize the precision of individual compounds or type of

microbial. The estimates of the accuracy of integrated data are limited because the methods used do not replicate field collection methods, considering the manner in which the sensors were spiked. In addition, spiked samples were not available for all sampling methods or, in the case of the VOCs, for all analytes. It should also be noted that these estimates of precision and accuracy can not be applied to different sampling methods or to analyses by different laboratories.

The estimates of precision and accuracy for the continuous data have fewer limitations. One potential limitation is that the zero and span gases used were dry, but measurements were generally conducted in environments with varying levels of humidity. This may have influenced the measurements to an unknown degree. Review of the multipoint calibration records for continuous sensors would be useful to assess whether responses were linear between the zero and span concentrations. To ensure that the standards used were reliable, these records could also be examined for the NIST traceable reference instruments.

APPENDIX A

Sample Collection
and
Analysis Methods

SAMPLE COLLECTION AND ANALYSIS METHODS

Although the environmental samples collected during the BASE study were based on established sampling protocols and considered "state of the art," the scope and magnitude of evaluating precision and accuracy differed for each measured parameter, and was predicated by the methods used to collect and analyze the data.

The majority of data collected was accompanied by quality assurance and quality control (QA/QC) samples, from which precision and/or accuracy could be measured. However, parameters including dust and bulk samples, and continuous light and noise monitoring were limited with respect to the demonstration of precision and accuracy. The aforementioned continuous parameters lacked the applicable technology for in-field performance evaluation; therefore, annual manufacturer calibrations and certifications were the only methods available for ensuring data quality. The methods used to collect dust and bulk samples were not applicable to collecting duplicate or blank samples, and the contractor relied on established laboratory standards to ensure data quality.

In order to better understand the limitations and conditions under which error can be evaluated, the following sections will outline the various sample collection methods, analytical procedures, and quality assurance measures employed during the BASE study.

VOLATILE ORGANIC COMPOUNDS: SUMMA® CANISTER

The objective of this procedure is to collect a representative sample of air containing volatile organic compound (VOC) contaminants present in an indoor environment using an evacuated canister. This sample is subsequently analyzed for the concentrations of 52 VOCs, as selected by EPA. The procedure involves several steps, including canister preparation, sampling of the indoor and outdoor air, and the analyses of the samples collected.

The sampling apparatus consists of a SUMMA® canister with an attached precalibrated low volume flow controller for time integrated sampling. The flow controller contains a critical orifice that ensures a constant flow rate during the sampling period. The flow

controllers are set to fill the sample canister approximately three quarters full (4.0 liters) over an eight- to ten-hour sampling period.

The fixed site sampling convention is as follows, provided the study area can accommodate the configuration.

- Outdoor Site: One sample, one duplicate
- Fixed Site 1 (indoors): One sample
- Fixed Site 3 (indoors): One sample, one spiked sample
- Fixed Site 5 (indoors): One sample, one duplicate

Analytical Method

VOC samples are analyzed using gas chromatography/mass spectrometry (GC/MS). The analyses are performed according to the methodology outlined in EPA Method TO-14 from EPA's *Compendium of Methods for the Determination of Toxic Organic Compounds in Ambient Air*. The analyses are performed by GC/MS utilizing thermal desorption/cryogenic concentration.

Quality Control and Assurance

There is no blank sampling associated with VOC sampling with evacuated canisters; however, the laboratory does analyze a method blank sample in compliance with in-house standards. For ensuring quality control with this sampling method, there are different sampling and analysis procedures employed. Specifically, these include the submission and analysis of a spiked sample, the submission of field duplicates, in-house duplicate analysis of one field sample as a lab duplicate.

To test the accuracy of the analytical procedures, some canisters are filled at the laboratory with a spiked sample and are delivered to the field team each week as part of the normal canister shipment. These spiked canisters are issued an IADCS-generated ID label, making them indistinguishable from all the other VOC samples, and they are returned to the laboratory for analyses as part of the complete shipment. No additional air

is sampled in the field for the spiked samples. In the early stages of the BASE study, additional spiked samples were generated by an independent laboratory. These were referred to as performance evaluation and performance demonstration (PE/PD) samples. Their use was discontinued early in the study.

The repeatability (precision) of sampling and analysis is assessed by performing sample duplicates. Each week one set of duplicate samples is collected at a specified indoor location and another set of duplicate samples is collected at the outdoor location.

As Environmental Health & Engineering, Inc. is not in possession of all PE/PD results, no data for these samples will be presented in this report.

VOLATILE ORGANIC COMPOUNDS: MULTISORBENT SAMPLERS

The objective of this procedure is to collect a representative sample of volatile organic compound (VOC) contaminants present in indoor and outdoor environments using multisorbent samplers. This sample is subsequently analyzed for the concentrations of 46 VOCs, as selected by EPA. The procedure involves several steps, including the assembly of the sampling apparatus, the sampling of the indoor and outdoor air, and the analyses of the samples collected.

The multisorbent sampler consists of a 200 mm long by 6 mm O.D. glass tube packed with several sorbent materials, as detailed in Figure 1 of appendix A. There is a glass frit at the sample inlet end, followed by sections of sorbent materials consisting of: 1) glass beads; 2) Tenax TA; 3) Ambersorb XE-340; and 4) activated carbon. Each section is separated by glass-wool plugs.

The multisorbent sampling apparatus, which consists of a diaphragm pump and rotameters, is configured to deliver a nominal flow rate of 5.0 ml/min over an eight- to ten-hour period.

The fixed site sampling convention is as follows, provided the study area can accommodate the configuration.

- Outdoor Site: One sample, one duplicate
- Fixed Site 1 (indoors): One sample, one field blank
- Fixed Site 3 (indoors): One sample, one spiked sample
- Fixed Site 5 (indoors): One sample, one duplicate

Analytical Method

VOC samples are analyzed using gas chromatography/mass spectrometry (GC/MS). The analyses are performed according to a modified version of the methodology outlined in EPA Method TO-1 from EPA's *Compendium of Methods for the Determination of Toxic Organic Compounds in Ambient Air.* The samplers are thermally desorbed, analyzed by gas chromatography and mass spectrometry

Quality Control and Assurance

One field blank is submitted for laboratory analysis on a per building basis as a blinded sample, to ensure the quality of contaminant-free samplers. Additionally, duplicate samples are also submitted to assess the repeatability (precision) of this analytical method. Each week, one set of duplicate samples is collected at a specified indoor location and another set of duplicate samples is collected at the outdoor location.

To test the accuracy of the analytical procedures, samplers are spiked for subsequent analysis. Some samplers are spiked at the laboratory and delivered to the field team for each city as part of the normal sample shipment. These spiked samplers are issued an IADCS-generated ID label, making them indistinguishable from all the other VOC samples, and they are returned to the laboratory for analyses as part of the complete shipment.

ALDEHYDES

The objective of this procedure is to collect representative samples of formaldehyde and other carbonyl compounds present in an indoor environment and in the outdoor air supplied to the space. These samples are subsequently quantitated by desorption and

chromatographic analyses. The test is intended to characterize the indoor level of aldehydes and to determine whether the likely source is indoors or outdoors.

The sampling train consists of an ozone scrubber (potassium iodide-based denuder) at the front end, followed by the Sep-Pak® cartridge. Placed immediately behind the cartridge is a trap filled with activated charcoal to remove acetonitrile liberated from the DNPH during sampling. The pump (low noise personal pump with a low flow adapter) is located downstream of the charcoal trap and is configured to deliver a nominal flow rate of 200 ml/min over an eight- to ten-hour period.

The fixed site sampling convention is as follows, provided the study area can accommodate the configuration.

- Outdoor Site: One sample, one duplicate
- Fixed Site 1: One sample, one field blank
- Fixed Site 3: One sample, one spiked sample
- Fixed Site 5: One sample, one duplicate

Analytical Method

As described above, formaldehyde and acetaldehyde are sampled using a sorbent tube containing silica gel coated with 2,4-dinitrophenyl hydrazine (DNPH). Formaldehyde and other aldehydes react with DNPH to form stable hydrazones, which are extracted from the silica gel and analyzed by high performance liquid chromatography (HPLC). The procedures follow the Ambient Air Compendium, Method TO-11 (1987). Results are reported for only formaldehyde and acetaldehyde. Formaldehyde was analyzed for in all 100 buildings, but acetaldehyde was only included in the last 86 buildings.

Quality Control and Assurance

The quality control samples for the adsorbent cartridges for the BASE study consist of lab blanks, lab spikes, a field blank, and a field spike.

The field blank is prepared by going through the process of placing a cartridge in line for 30 seconds and then removing it without having drawn air through the sample. The field blank is then placed in a foil envelope to sit at the blank fixed site for the duration of the sampling period.

Spiked sample cartridges are prepared by the analytical laboratory and delivered to the field team for each city as part of the normal sample shipment. These spiked samplers are issued an IADCS-generated ID label, making them indistinguishable from all the other samples, and they are returned to the laboratory for analysis with the complete shipment.

PARTICULATES

Particles are collected in this procedure from indoor air and from the outdoor air supplied to the indoor space tested. The particle samples are collected by inertial impaction onto pre-weighed filters. The collected material is weighed and quantitated in terms of weight per unit volume of air sampled. The sampling is conducted on two size ranges: particles less than 10 microns in diameter (PM_{10} "inhalable" fraction) and particles less than 2.5 microns in diameter ($PM_{2.5}$ "fine respirable" fraction).

Each sampling apparatus consists of a single-stage impactor with a classifier (i.e., orifices) for one of the size ranges of interest (2.5 microns or 10 microns). The impactor classifier selected defines the size (aerodynamic diameter) of the particles collected and has a nominal “cut-off size.” The cut-off size is defined as the aerodynamic diameter of particles that are collected with 50% efficiency. Sampling is conducted by drawing air through the impactor at a rate of 20.0±1 lpm. The samples are collected on 37 mm Teflon air sampling (PTFE, Gelman) membrane filters.

These filters are weighed before use after equilibrating in a temperature- and humidity-controlled weighing room. After sample collection, the filters with the collected particles are again equilibrated in the weighing room, and then weighed to determine the increase in mass.

The fixed site sampling convention is as follows provided the study area can accommodate the configuration.

- Outdoor Site: One sample, one duplicate
- Fixed Site 1: One sample, one field blank
- Fixed Site 3: One sample
- Fixed Site 5: One sample, one duplicate

Analytical Method

The 37mm Teflon filters are weighed prior to use using a microbalance in a temperature- and humidity-controlled room. The filters are then loaded in dichot filter holders and shipped out to the field teams. Ten per cent of the filters prepared for field use are kept as lab blanks at the analyzing laboratory. The samples received from the field teams are then re-weighed again after equilibration with controlled temperature and humidity. Ten percent of the filters are re-weighed for QC purposes.

Quality Control and Assurance

The required precision of the balance is ±0.004 mg.

A field blank is included in each sample batch. The field blank is prepared by inserting it into an impactor and then removing it without running air through the instrument. The blank is issued an IADCS generated label and sent blinded to the laboratory.

VIABLE AEROBIOLOGICALS

The objective of this procedure is to determine concentrations of viable airborne microbiological organisms (bioaerosols) that may be present in indoor air and in the outdoor air supplied to the space tested. The collected samples are subsequently cultured to speciate and quantify the organisms collected. The colony forming units (CFUs) of each organism are quantitated in terms of their number per unit volume (STP) of air sampled. The organisms of interest are mesophilic bacteria, thermophilic bacteria, and fungi.

Bioaerosol samples are collected using single stage (N6) Andersen inertial impactors with isokinetic sampling orifices. Samples are collected on two different types of culture plates. Trypticase soy agar (TSA) culture plates are employed for thermophilic and mesophilic bacteria, and malt extract agar (MEA) culture plates are used for fungi.

Air is sampled at a flow rate of 28.3±1.4 liters per minute, for time intervals of 2 minutes and 5 minutes as specified in the BASE Protocol.

Samples are collected during the morning and during the afternoon on one day during the field week at the following locations, provided the test space can accommodate this configuration.

- **Outdoor Site (near the outdoor air intake for the study area).** Two and five minute samples and duplicates for mesophilic and thermophilic bacteria and for fungi.

- **Fixed Site 1 (indoors)**. Two and five minute samples for mesophilic and thermophilic bacteria and for fungi.

- **Fixed Site 3 (indoors).** Two and five minute samples for mesophilic and thermophilic bacteria and for fungi, three field blanks collected in the morning round, and three shipping blanks collected in the afternoon round for both TSA and MEA plates.

- **Fixed Site 5 (indoors).** Two and five minute samples and duplicates for mesophilic and thermophilic bacteria and for fungi.

Analytical Method

Aerobiological samples are collected on 100mm x 15mm polystyrene petri plates containing either MEA or TSA. Fungi samples are incubated at 25° with 12-hour light and dark cycles. Colonies are examined seven days after incubation using standard light microscopy. Bacteria samples are incubated at 30° and 55° to promote the growth of

mesophilic and thermophilic bacteria respectively (mesophilics, 2-4 day incubation; thermophilics, 4-6 day incubation). Colonies are subsequently gram-stained, analyzed and counted using standard light microscopy techniques.

Quality Assurance and Control

QC samples consist of field blanks, shipping blanks and duplicate samples. Handling and sampling procedures are assessed using the field blanks, while the sterility and quality of the plates are evaluated using shipping blanks. Field and shipping blanks of each plate (mesophilic, thermophilic, fungi) are forwarded the laboratory.

The repeatability (i.e., precision) of sampling and analysis is assessed by sample duplicates. One set of duplicate samples is collected at a specified indoor location and another set of duplicate samples is collected at the outdoor location.

VIABLE AND NON-VIABLE AEROBIOLOGICALS

The objective of the procedure described is to determine airborne concentrations of viable and non-viable fungal spores combined. A Burkard air sampler was used to direct air through a slit orifice onto coated microscopic slides. Each slide is carefully prepared by applying a small amount of stop-cock grease to a finger and spreading the grease thinly, covering only the central portion of the slide. Air is sampled for nominally four (4) minutes at a flow rate of approximately 14.5 liters/minute, and for a total volume of approximately 58 liters. Prior to sampling both indoors and outdoors, the top inlet of each sampler is wiped with isopropyl alcohol.

The samples from the Outdoor Site and Fixed Sites 1, 3, and 5 and the duplicate sample are collected simultaneously with the Andersen sampling.

Analytical Method

Samples were analyzed to identify and quantify fungal types present by light microscopy by a qualified microbiology laboratory. Results were reported in spores per cubic meter of air sampled (spores/m^3).

Quality Assurance and Control

Meaningful spike and blank samples could not be produced for this method. Only duplicate samples could be used to assess precision. However, due to equipment limitations, only one Burkard sampler was available per field team. Consequently, the duplicate samples were actually collected consecutively rather than simultaneously. This would likely increase differences between duplicate samples.

CARBON DIOXIDE AND CARBON MONOXIDE: COLLECTION METHOD AND ANALYSIS

The objectives of these methods are to measure representative levels of carbon dioxide (CO_2) and carbon monoxide (CO) from the indoor environment, and the outdoor air serving the test space.

Both the CO and CO_2 sensors employed are part of the aq-502 Indoor Environmental Monitor manufactured by Metrosonics. The active CO_2 sensors employ non-dispersive infrared radiation (NDIR) adsorption to quantify the concentration of CO_2 present in a gas stream. The instrument output is a voltage that relates linearly to CO_2 concentration in the 0 to 5,000 parts per million (ppm) range. The precision (repeatability) of the instrument is, according to the manufacturer's specifications, 3% of full scale. Calibration records at EH&E show a repeatability for different instruments and at different times in the range of 0 to 40 ppm, based on measurements with gases of known concentrations. For comparison, the BASE Method Performance Requirement is ±50 ppm. The manufacturer's specified accuracy of the instrument is 3% of the reading or ±50 ppm, whichever is greater. Calibration records at EH&E show average deviations of -13 to 28 ppm for the 0 to 1,000 ppm concentration range. These data are presented here to facilitate the evaluation of the results of span checks and calibrations conducted under field conditions.

The CO sensors are CiTiceL®, 7E/F aqueous electrolyte-based gas detectors (model gs-7701, manufactured by Metrosonics Inc.). The gas detected penetrates through a diffusion barrier and, upon equilibrating with the electrolyte solution, changes its

conductivity in proportion to the equilibrium concentration of the gas species of interest (in this case CO). The instrument output is a voltage that relates linearly to the partial pressure (i.e., concentration) of CO. The conductivity of the electrolyte is also affected by temperature and by the presence of some other gases in the atmosphere. Ethylene and hydrogen are the principal gases affecting CO concentration measurements. With regard to temperature effects, the cells are partly temperature compensated. According to the manufacturer (City Technology, Inc.), the mean square deviation from the true concentration in the temperature range of 10 to 40°C is less than 5%. EH&E's experience with CiTiceL® sensors is that, at low CO concentrations (< 5 ppm), a 5°C temperature variation can produce a variation in instrument response equivalent to 1 to 2 ppm of CO concentration. Repeated detector calibrations show mean square deviations (precision) ranging between 0.3 and 1.0 ppm (compared to the BASE method performance requirement of ±1 ppm). Similarly determined accuracy shows deviations between 1.5 and 2.0 ppm (compared to the BASE method performance requirement of ±2 ppm).

Sensors and dataloggers are configured to collect samples at five-minute intervals over a 56- to 72-hour period. The fixed site sampling convention is as follows, provided the study area can accommodate the configuration.

- Outdoor Site: One sensor
- Fixed Site 1: One sensor
- Fixed Site 3: One sensor
- Fixed Site 5: One sensor

Fixed Site 5 previously housed a duplicate sensor array. However, it was later deemed redundant, given the employed quality assurance and control measures. Toward the latter half of the study, the duplicate sensor was moved to Fixed Site 2 to acquire more meaningful data (i.e., intraspace variability).

Quality Assurance and Control

Before and after each season, a multipoint calibration is performed on each sensor to ensure consistent and linear response over the range of concentrations. Prior to in-field continuous sampling, all CO_2 sensors undergo zero checks against a zero-air gas and span checks (against 1,000 ppm for the indoor system, and 350 ppm for the outdoor setup) using calibration gases with certificates of traceability to NIST. Investigators are responsible for consulting the validation table (BASE Protocol) to determine the performance level of the sensors. During the sampling period, each CO_2 sensor is checked daily (validation) using the same zero and span procedure followed on Monday. All zero and span checks are recorded on field log sheets, so that drift of sensor response can be tracked through the course of the week and through the course of the study. Finally, informal sensor performance checks are performed throughout the day to see if the sensor reading seems reasonable for each particular site.

As with the CO_2 sensors, all indoor and outdoor CO sensors are subject to zero checks against a zero air gas, and span checks against a 10 ppm span gas on the Monday before sampling begins. Once sampling begins, each fixed site CO sensor is checked daily using the same zero and span procedure followed on Monday. All zero and span checks are recorded on field log sheets, so that drift of sensor response can be tracked through the course of the week and through the course of the study.

TEMPERATURE AND RELATIVE HUMIDITY: COLLECTION METHOD AND ANALYSIS

Thermal stratification is measured employing Telaire, Inc., Model 1058T thermistor-based sensors with a range of 2 °C to 98 °C. EH&E's in-house cumulative calibration records for these sensors show that, generally, the precision of the Telaire model 1058-T is better than ±0.1°C, and the accuracy better than 0.2 °C. The performance requirement specified in the BASE Protocol is ± 2 °C accuracy.

The relative humidity detectors employed use an electronic thin film capacitor. Increased humidity raises the dielectric constant of the capacitor proportional to the relative humidity in the air.

The output of the sensors is factory set to read zero volts at 0% RH and 1 V at 100% RH. The instrument response is nearly linear (maximum deviation from linearity is 3.0% RH). The instruments are rated to measure the entire relative humidity range at operating temperatures between 0°C and 55°C. This meets the BASE method performance requirement of 20% to 90% RH.

Sensors and dataloggers are configured to collect samples at five minute intervals over a 56- to 72-hour period. The fixed site sampling convention is as follows, provided the study area can accommodate the configuration.

- Outdoor Site: One sensor
- Fixed Site 1: One sensor
- Fixed Site 3: One sensor
- Fixed Site 5: One sensor

Fixed Site 5 previously housed a duplicate sensor array. However, it was later deemed redundant, given the employed quality assurance and control measures. Toward the latter half of the study, the duplicate sensor was moved to Fixed Site 2 to acquire more meaningful data (i.e., intraspace variability).

Quality Assurance and Control

Before and after each season, a multipoint calibration is performed on the sensors to assess consistency and linearity of response. Prior to in-field sampling, all instruments used to take temperature measurements are gathered and compared to NIST-traceable glass thermometers or a digital thermo-hydrometer, such as Model PA-1 (Rotronic Instrument Company, Inc.) with a range of 0°C to 50°C. The accuracy of the digital thermo-hygrometer is better than ±0.5°C.

This “cross check” procedure is important for the early identification of malfunctioning instrumentation, and as a means of understanding the range of variation within the instrumentation. Once continuous monitoring has begun, all temperature sensors are compared to thermometers/digital thermo-hygrometer positioned at each fixed site each

sampling day. Again, the sensors are monitored to determine their performance level with respect to EPA validation ranges. If sensor performance falls outside the secondary range of 2 °C, it is replaced. Finally, informal sensor performance checks are conducted throughout the day to see if the sensor readings are reasonable.

As with the air temperature monitoring equipment, before any RH equipment is deployed for sampling it is compared to NIST traceable digital thermo-hygrometers, which are calibrated using known humidities from the head spaces of different salt solutions. Once again, the “cross check” procedure is important for the early identification of malfunctioning instrumentation, and is important as a means to understand the range of variation within the instrumentation being used. Relative humidity sensors in the field are also checked daily throughout the sampling week against the thermo-hygrometer. If sensor performance falls outside the secondary range of 7%, it is replaced. Furthermore, informal sensor performance checks are made throughout the day to see if the sensor readings seem reasonable.

APPENDIX B

References

REFERENCES

ACGIH. 1995. *Air Sampling Instruments for evaluation of atmospheric contaminants.* 8th edition. Cohen BS and Hering SV, editors. Cincinnati: American Conference of Governmental and Industrial Hygienists.

U.S. Environmental Protection Agency. 1994. "*A Standardized EPA Protocol for Characterizing Indoor Air Quality in Large Office Buildings*," Washington DC: Office of Research and Development and Office of Air and Radiation, U.S. Environmental Protection.

Environmental Heath and Engineering, Inc., Newton, MA and U.S. Environmental Protection Agency, Washington DC, December 1996. *BASE Quality Assurance Project Plan*.

APPENDIX C

Charts for Integrated Data Error Analysis

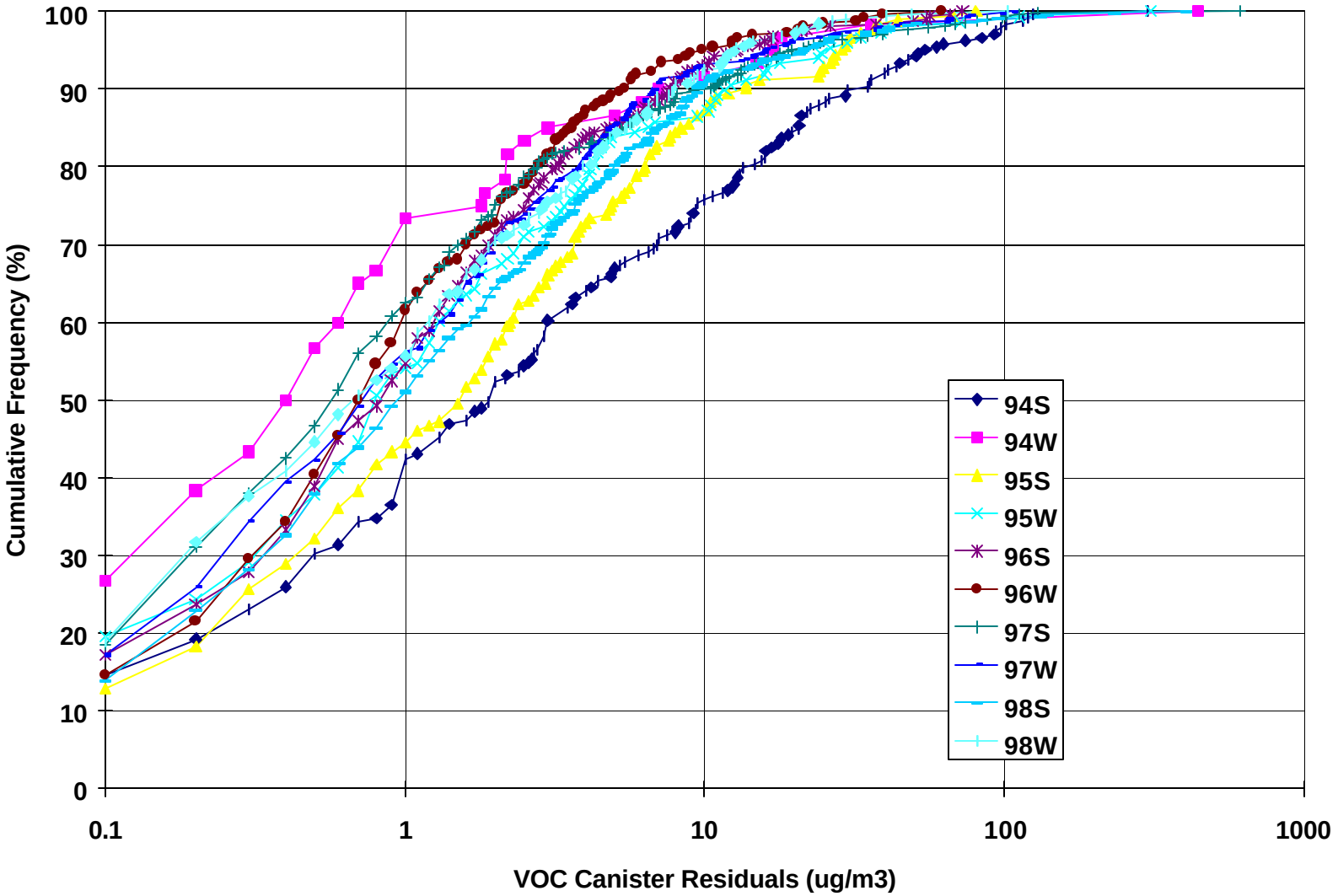


Figure C1 SUMMA® Canister: Cumulative Frequency of Residuals

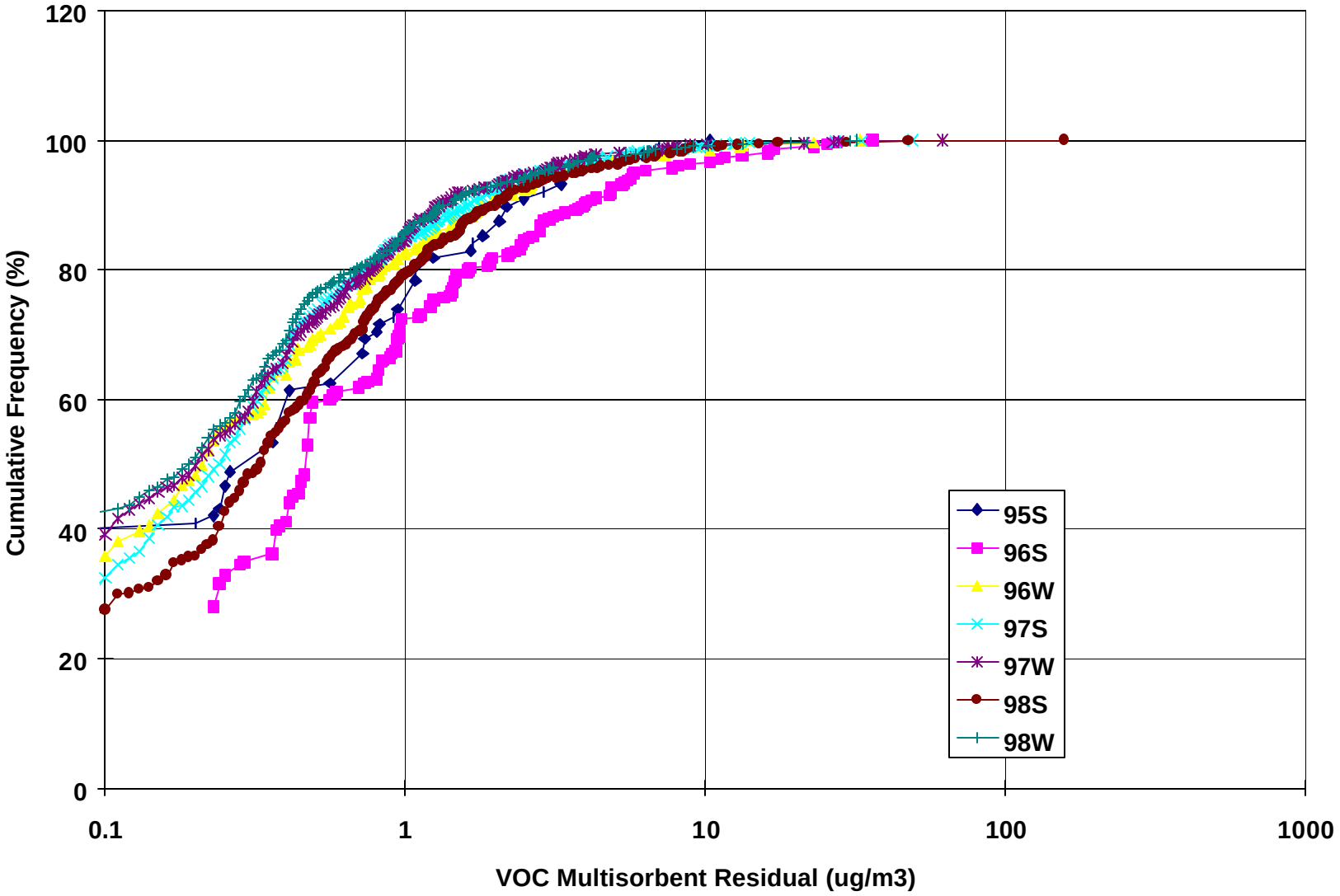


Figure C2 Multisorbent Samplers: Cumulative Frequency of Residuals

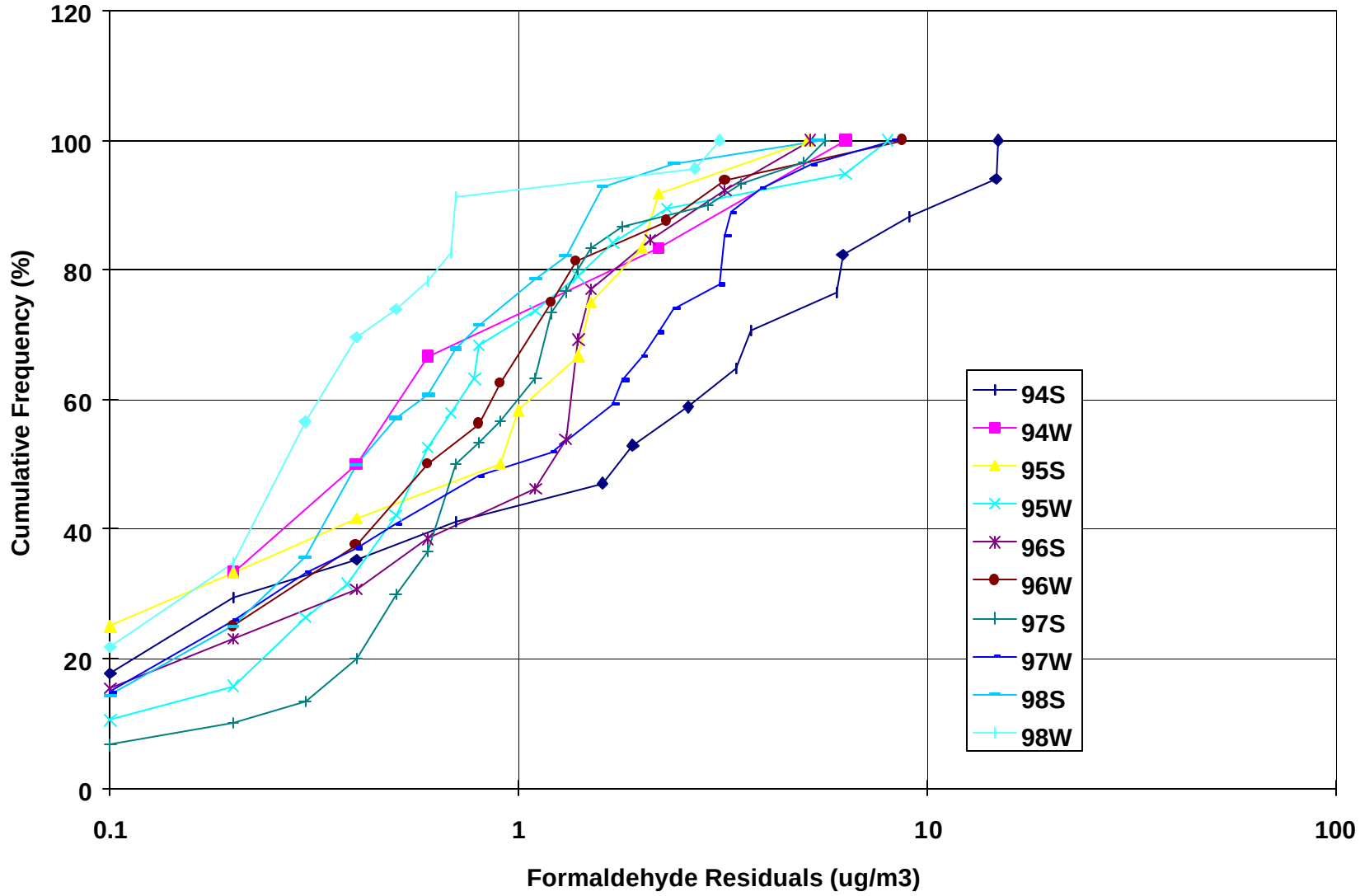


Figure C3 Formaldehyde: Cumulative Frequency of Residuals

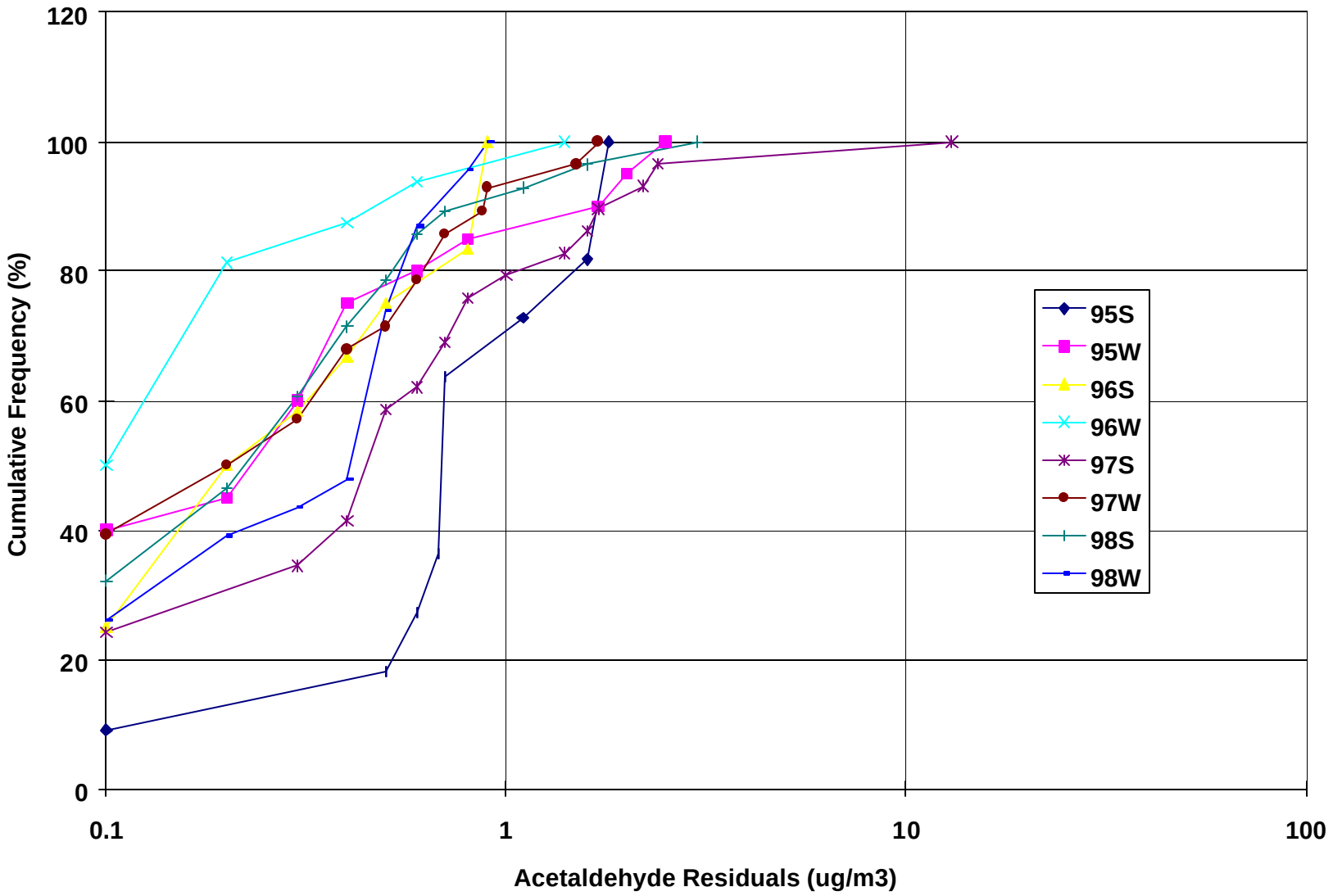


Figure C4 Acetaldehyde: Cumulative Frequency of Residuals

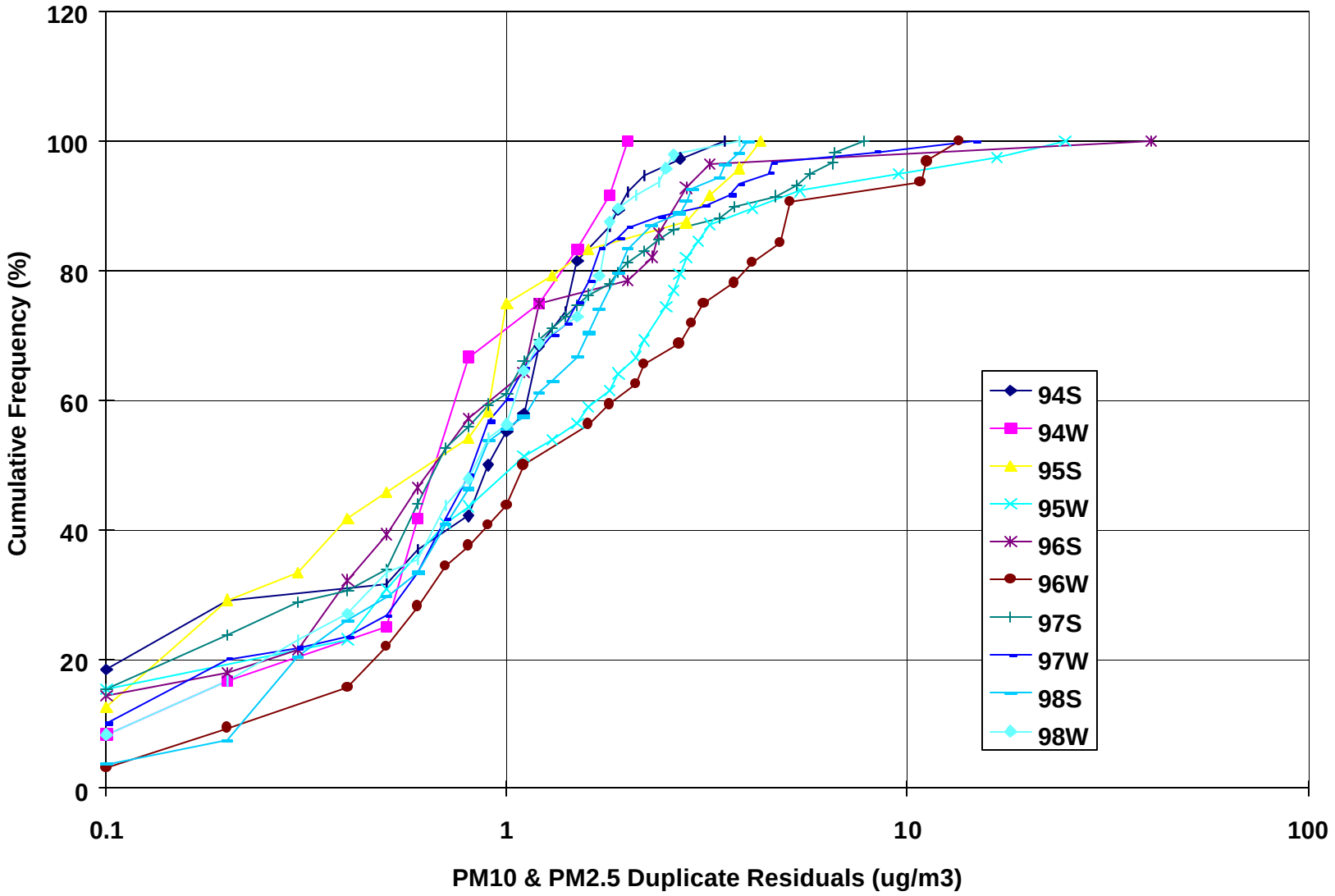


Figure C5 Particulates: Cumulative Frequency of Residuals (PM_{10} and $_{2.5}$)

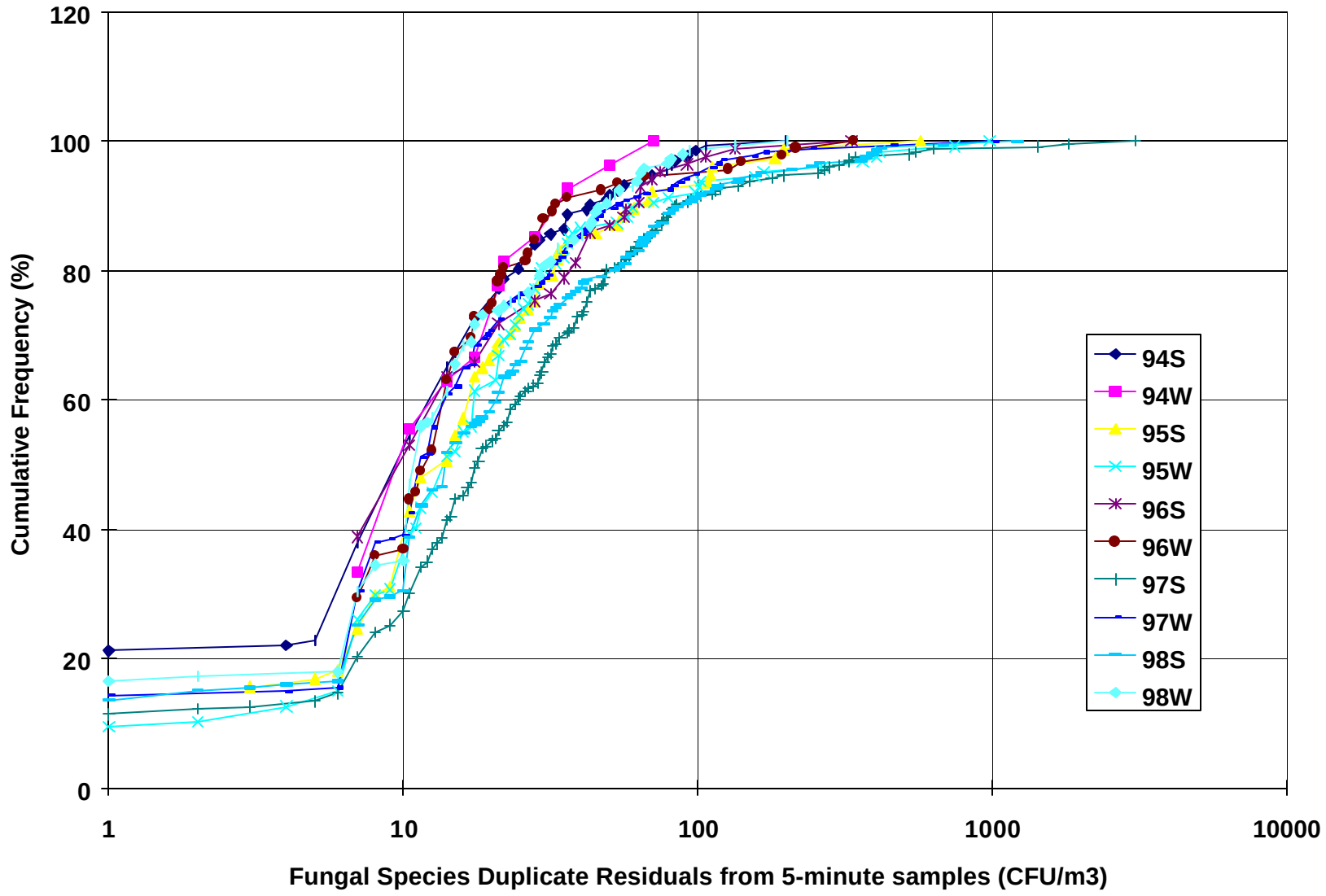


Figure C6 Fungi: Cumulative Frequency of Residuals

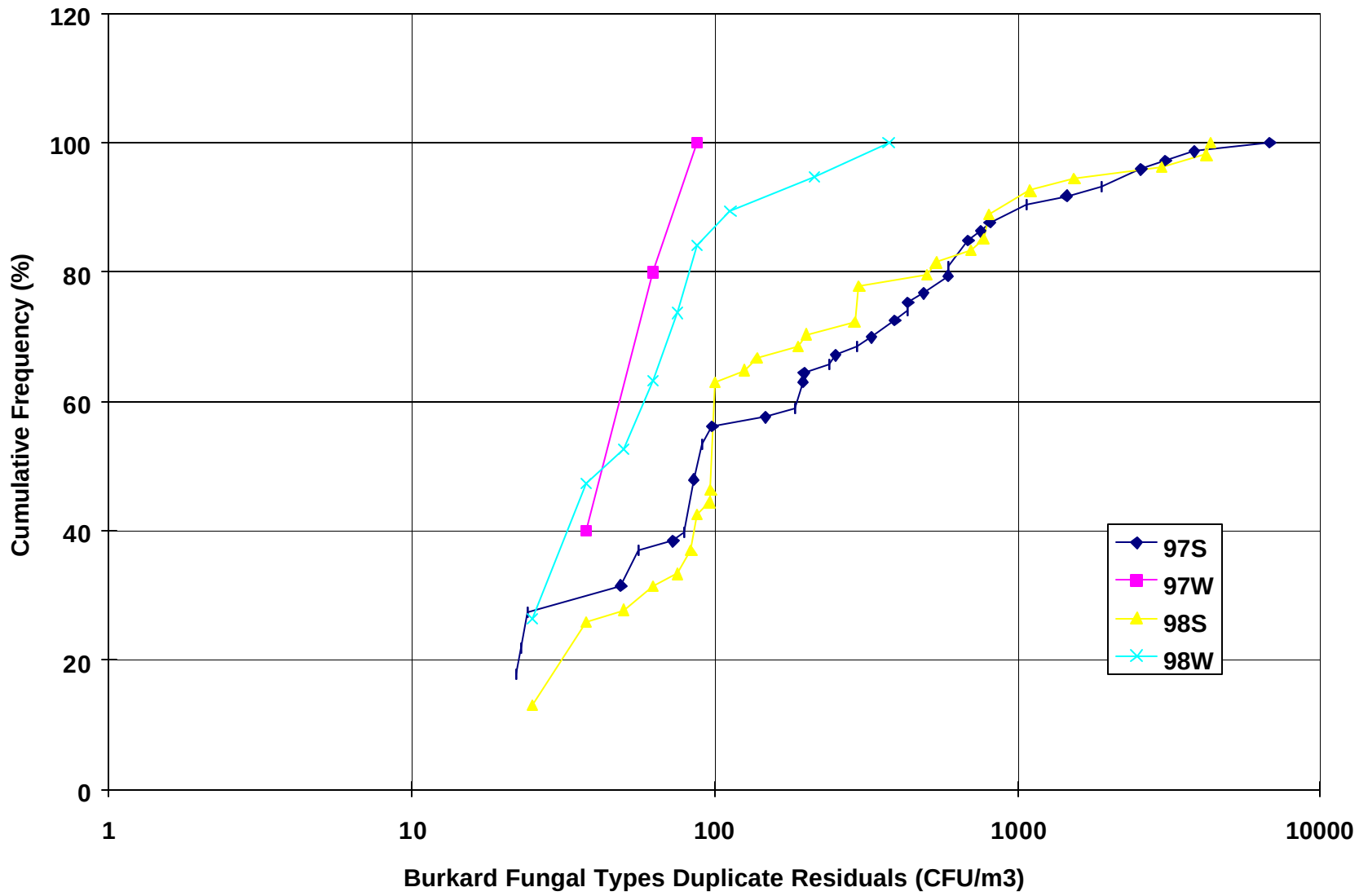


Figure C7 Burkard Fungal Types: Cumulative Frequency of Residuals

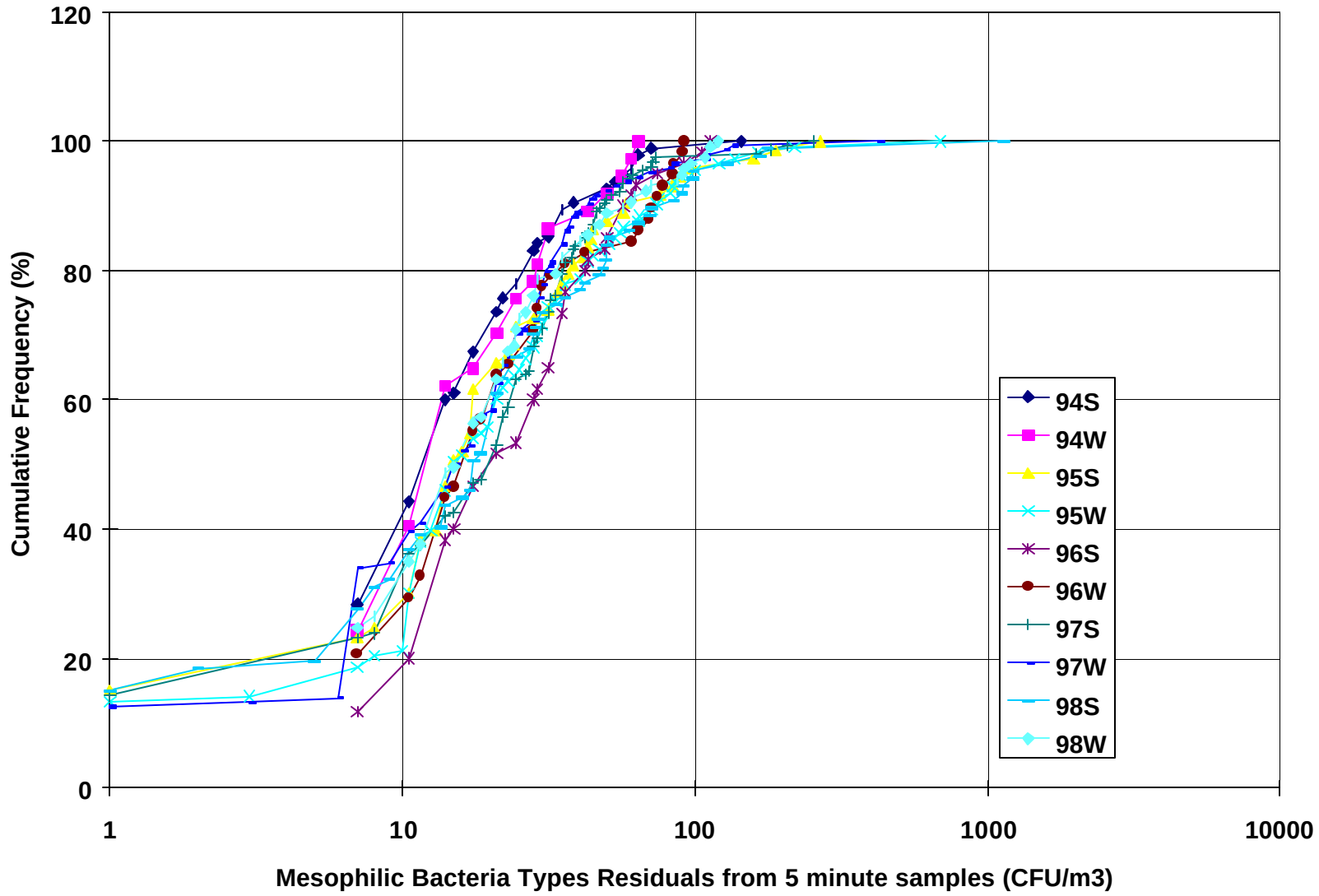


Figure C8 Mesophilic Bacteria: Cumulative Frequency of Residuals

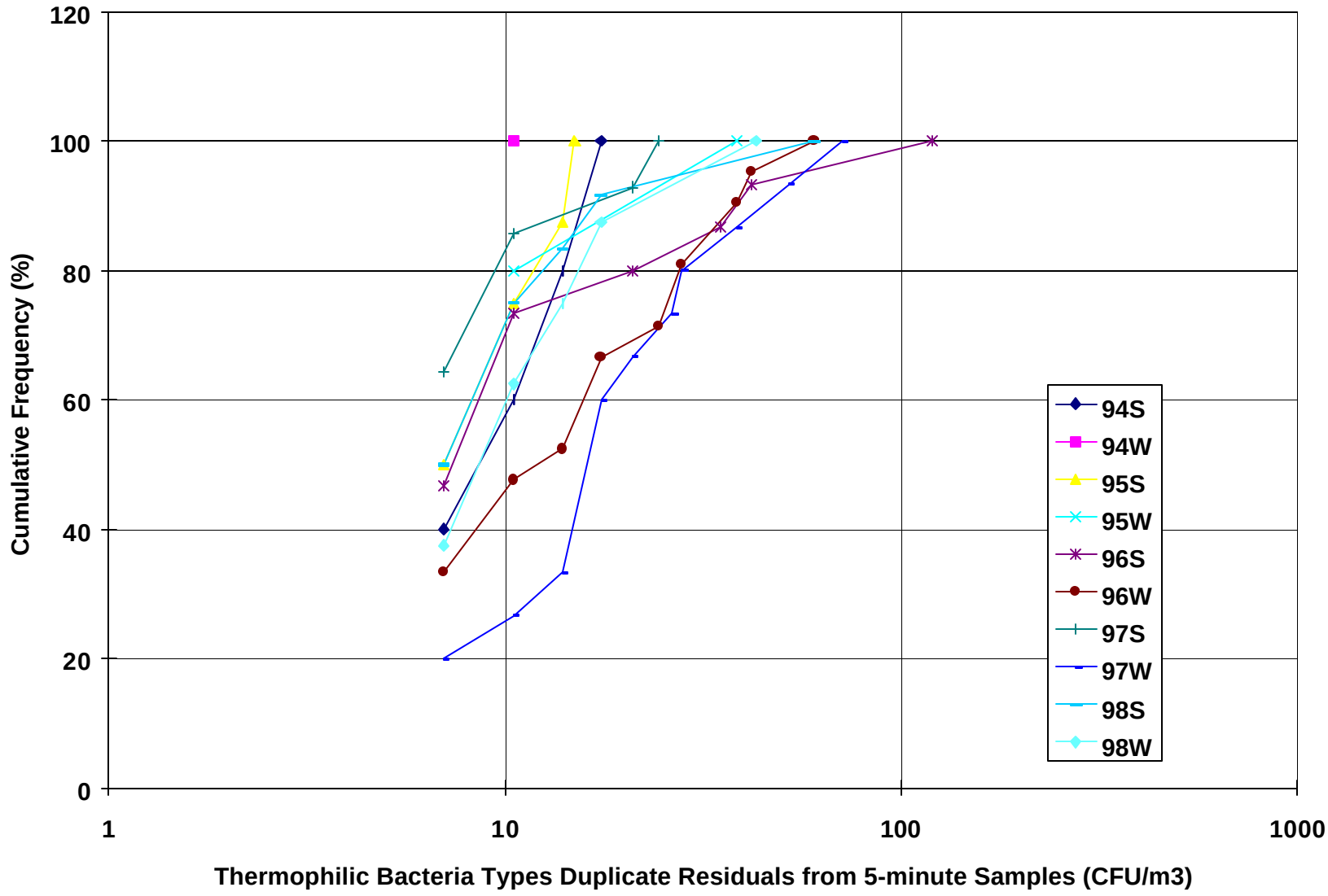


Figure C9 Thermophilic Bacteria: Cumulative Frequency of Residuals

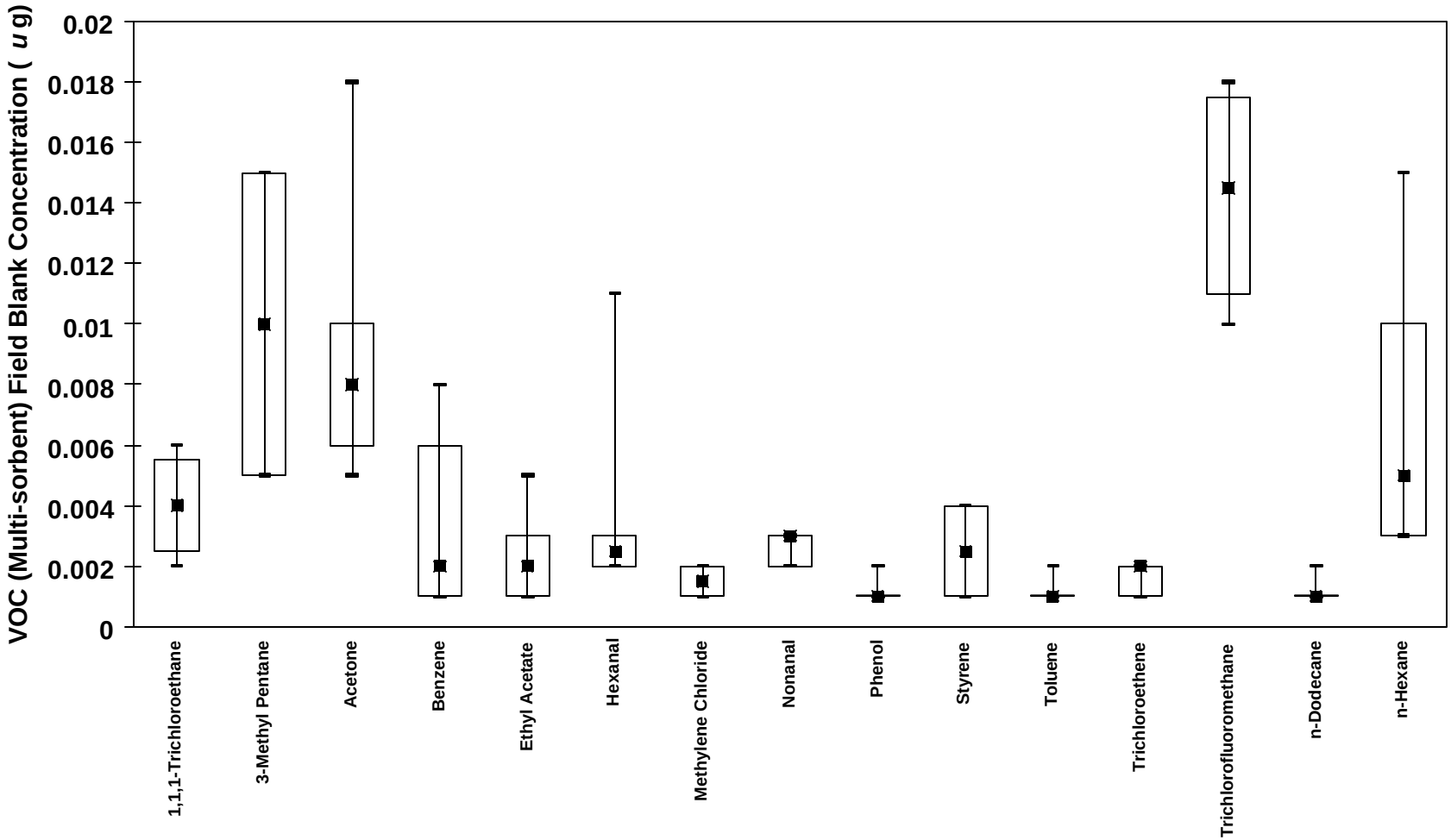


Figure C10 Summary of Field Blanks from Multisorbent Samples

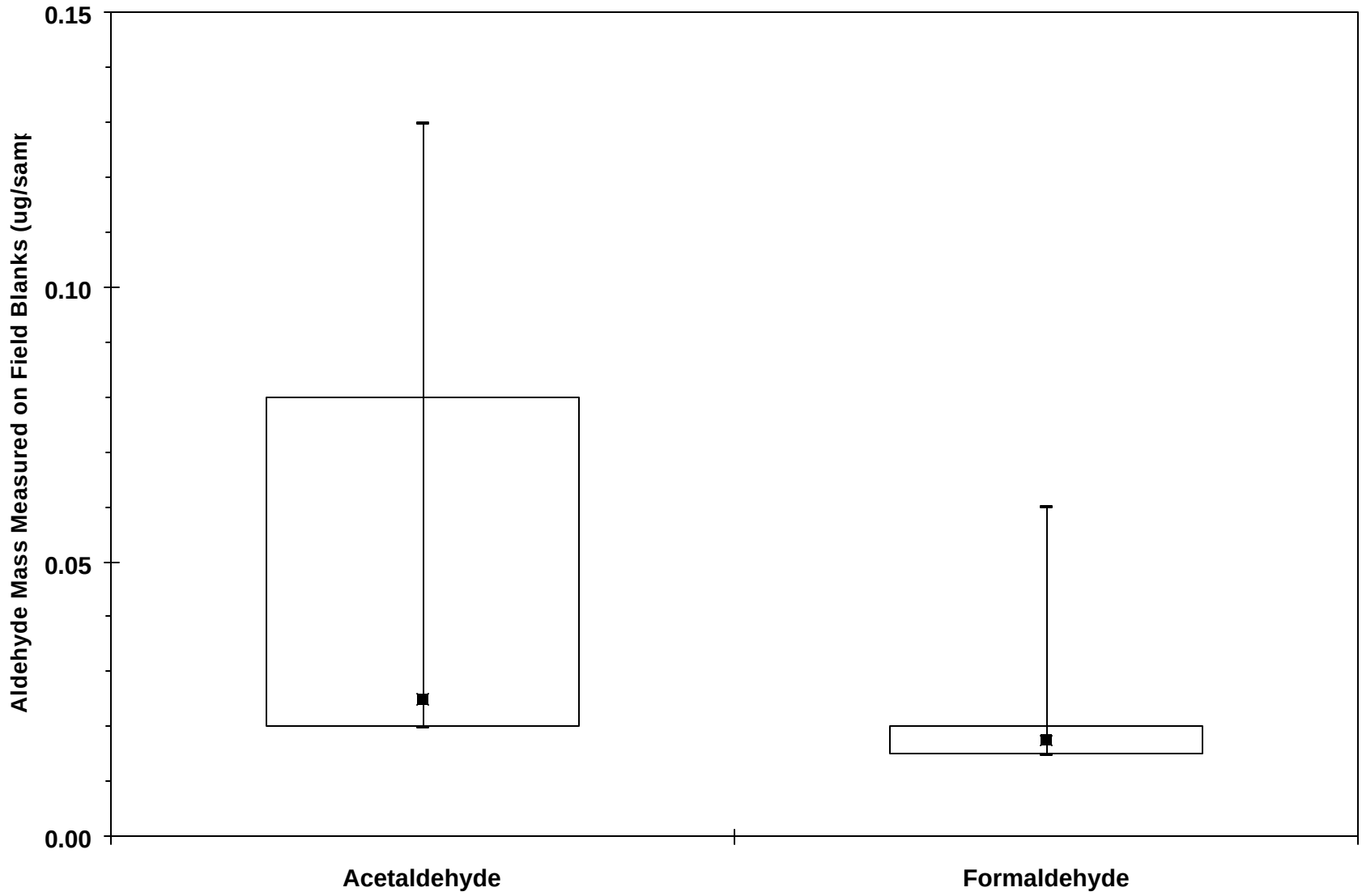


Figure C11 Summary of Aldehyde Field Blanks

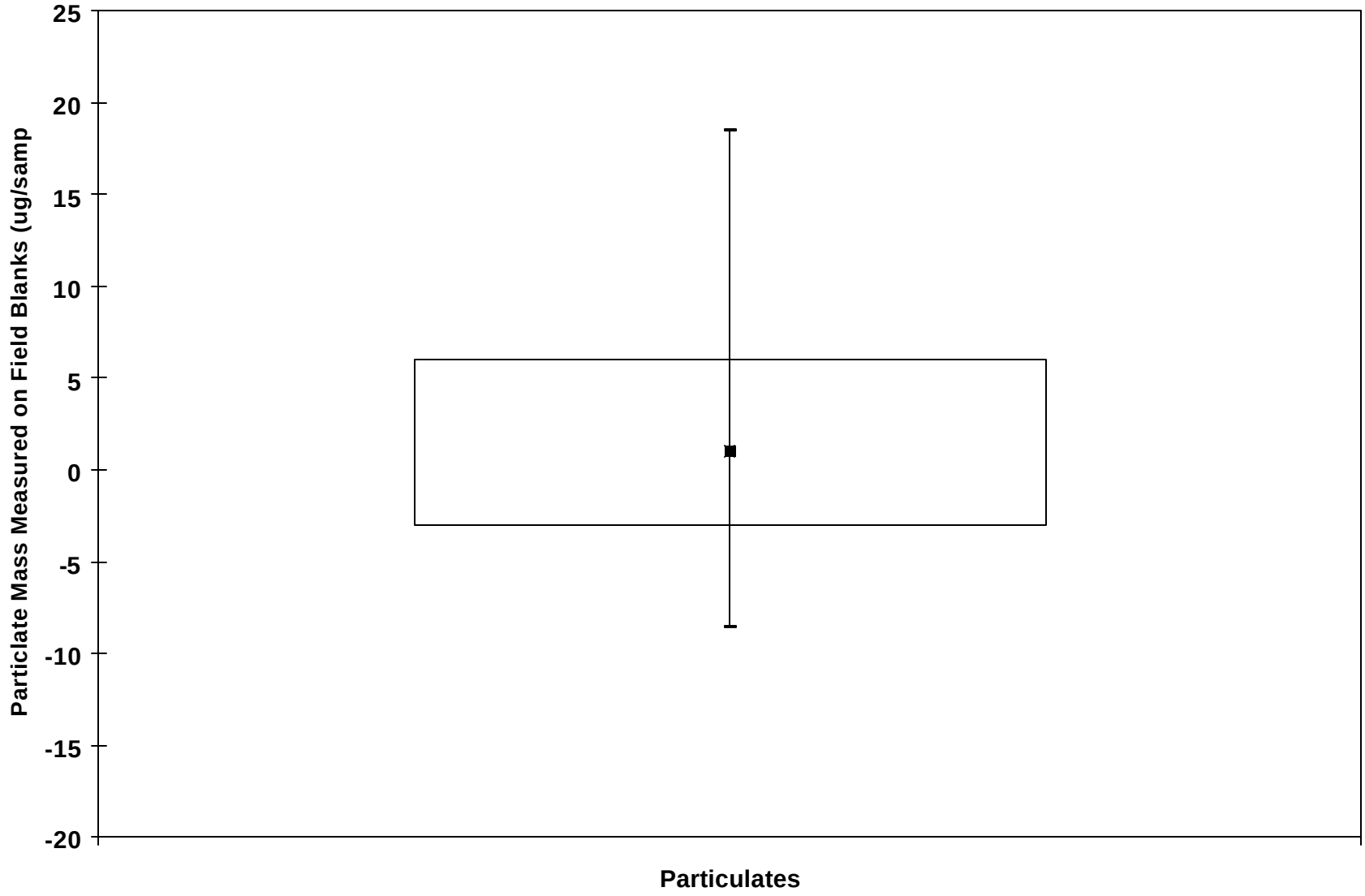


Figure C12 Summary of Particulate Field Blanks

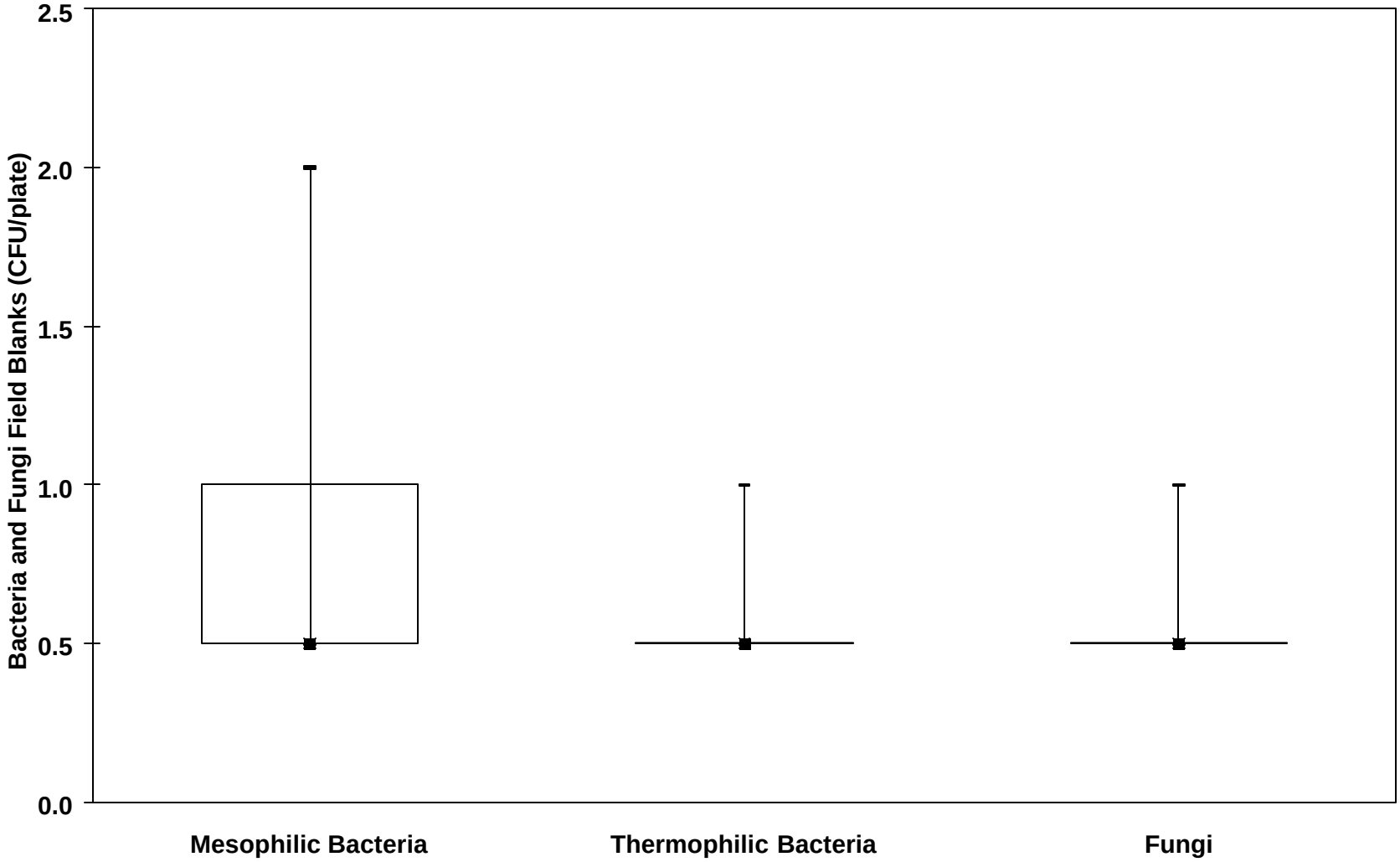


Figure C13 Summary of Bacteria and Fungi Field Blanks from the Andersen Sampler

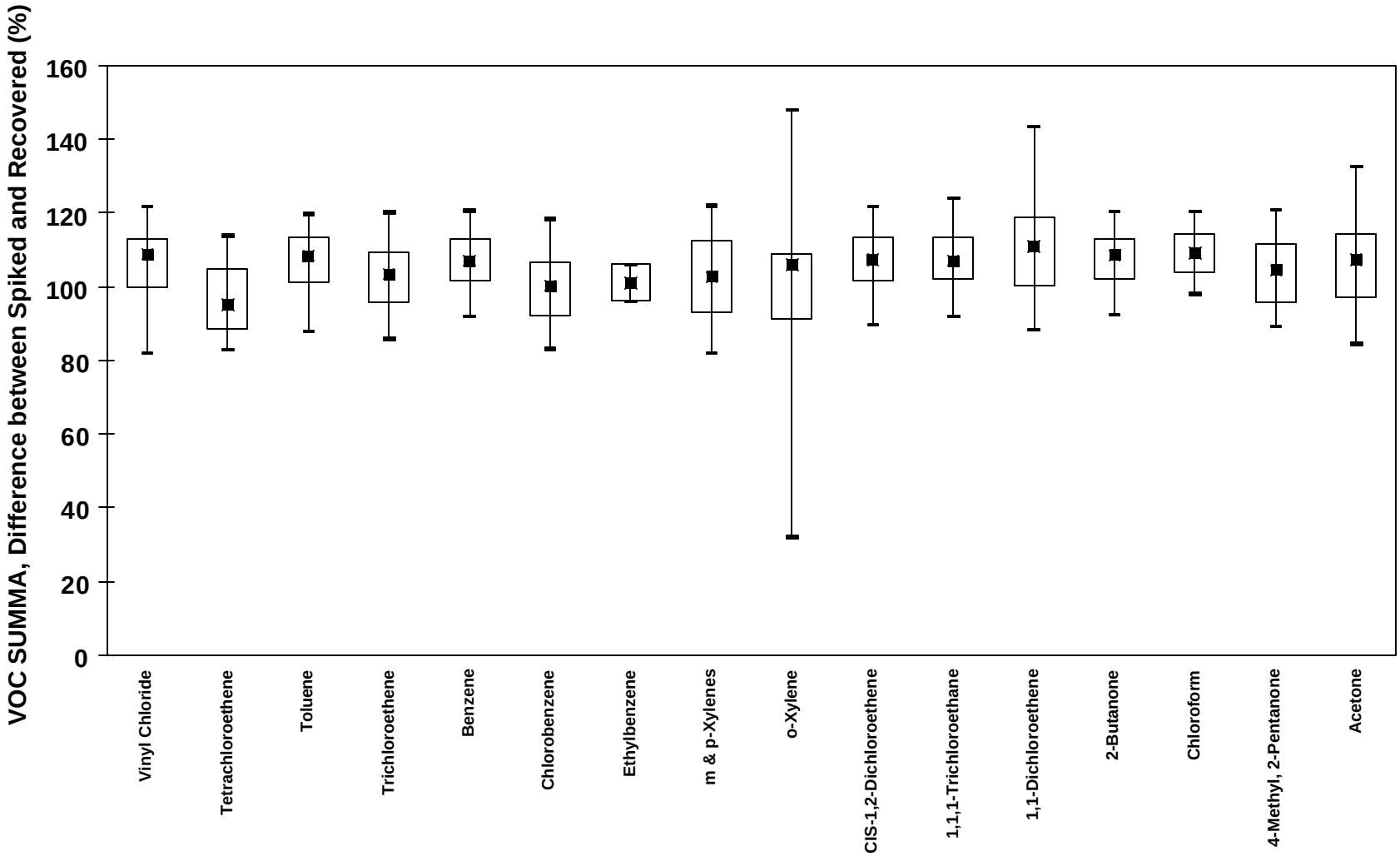


Figure C14 Percent Recovered from Spiked Canister Samples

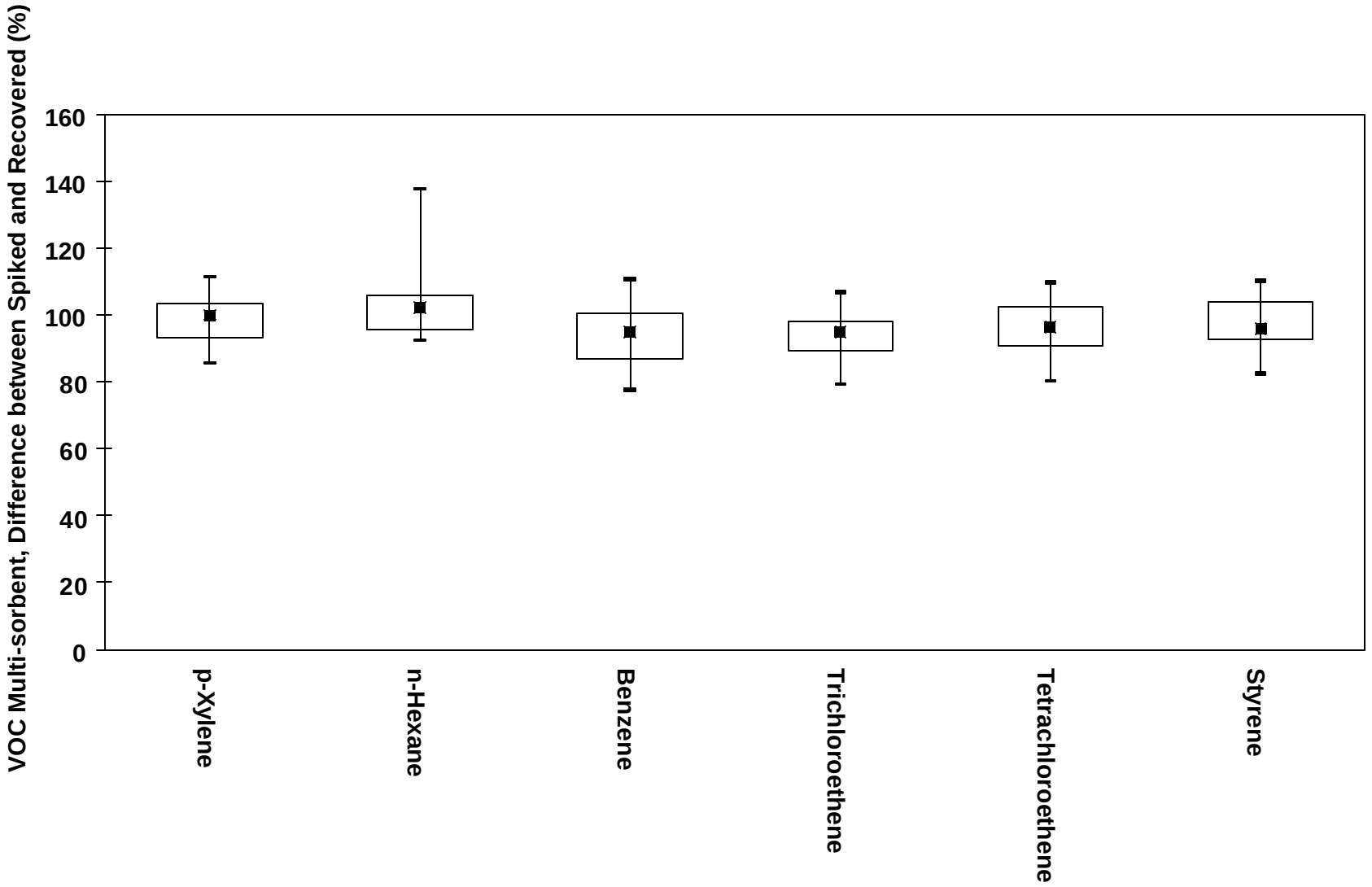


Figure C15 Percent Recovered from Spiked Multisorbent Samples

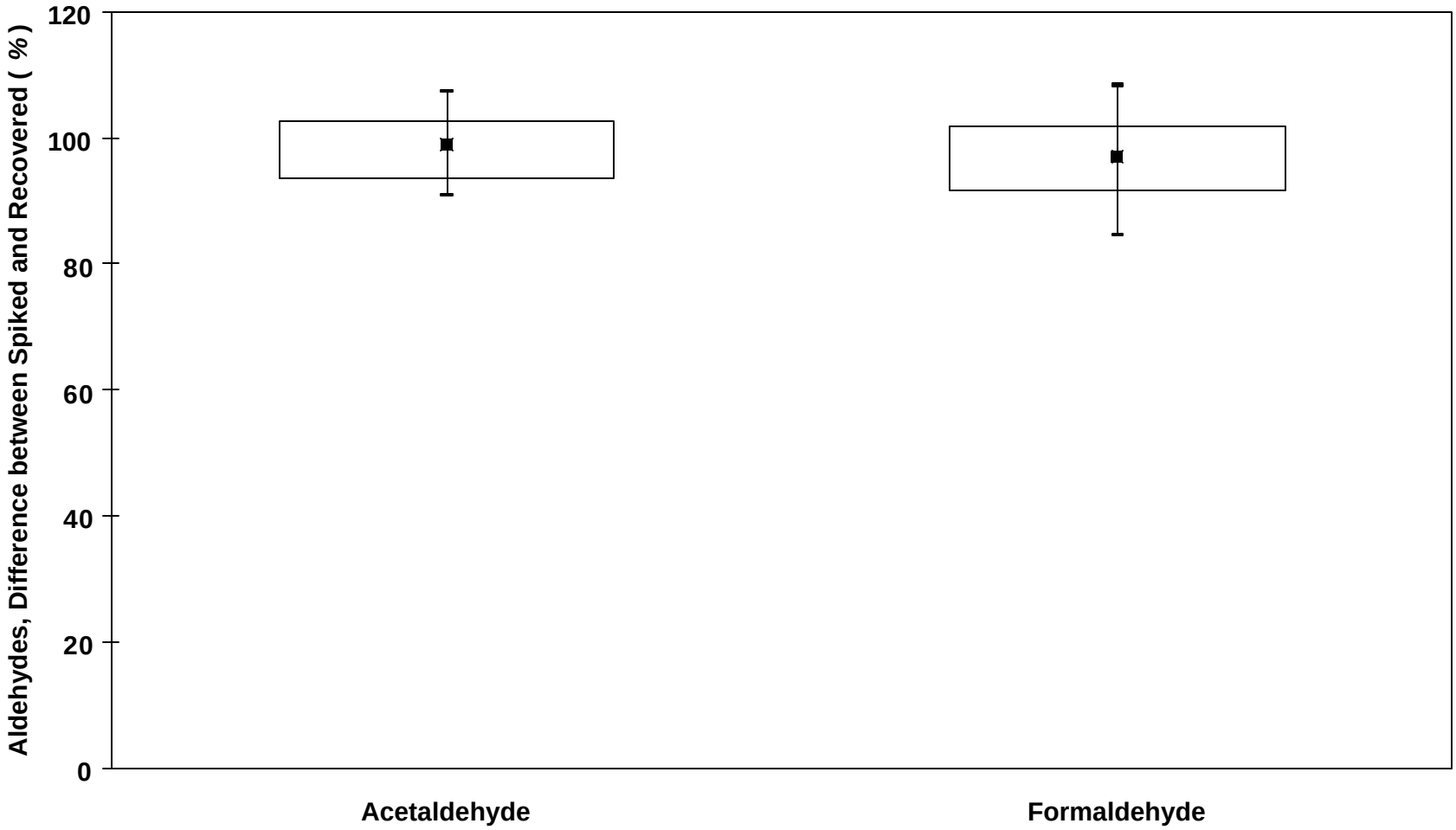


Figure C16 Percent Recovered from Spiked Aldehyde Samples

APPENDIX D

Charts for Continuous Data Error Analysis

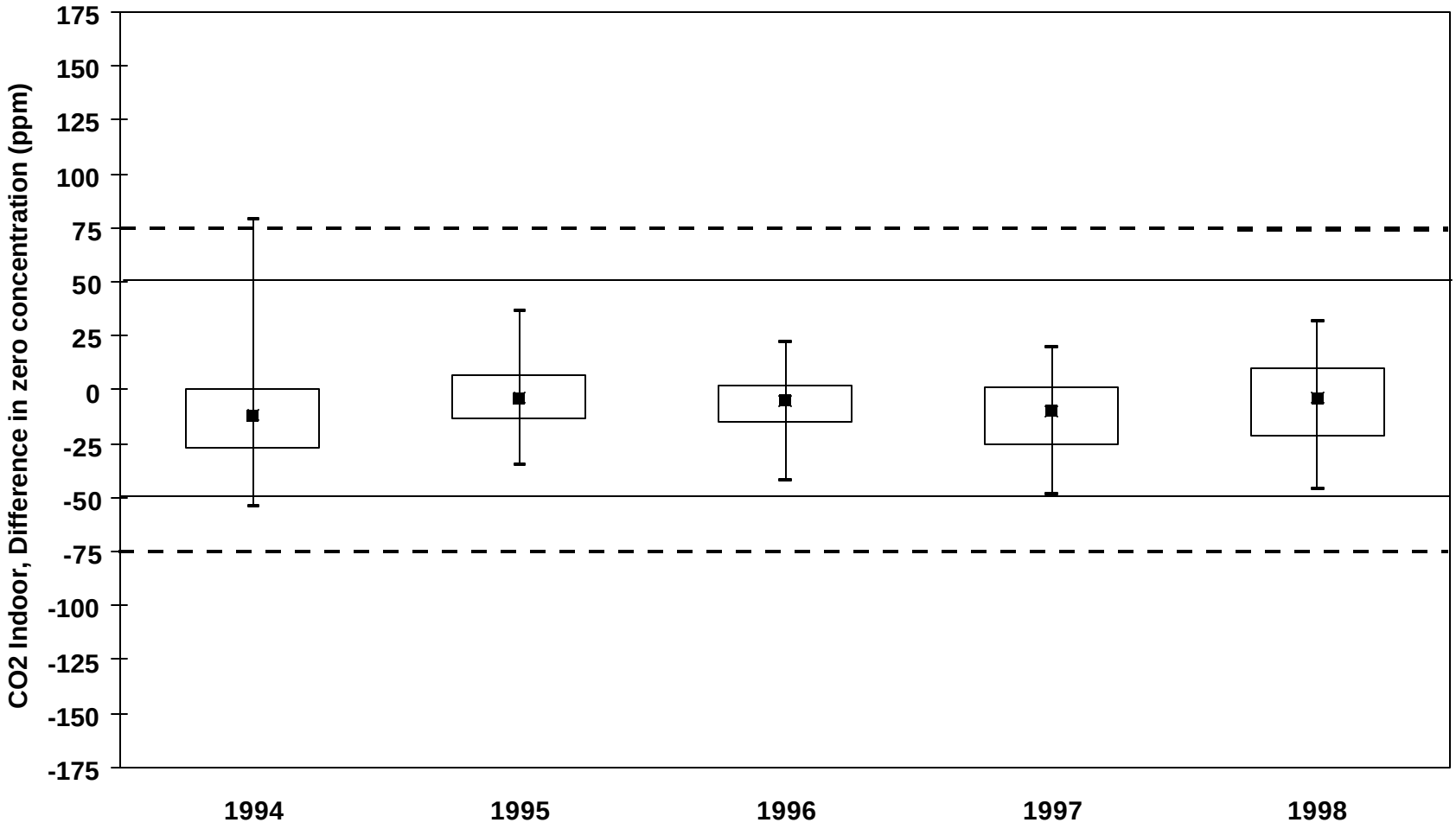


Figure D1 Differential Indoor CO_2 Sensor Response to Zero Gas, 1994 through 1998

Figure D2 Differential Outdoor CO_2 Sensor Response to Zero Gas, 1994 through 1998

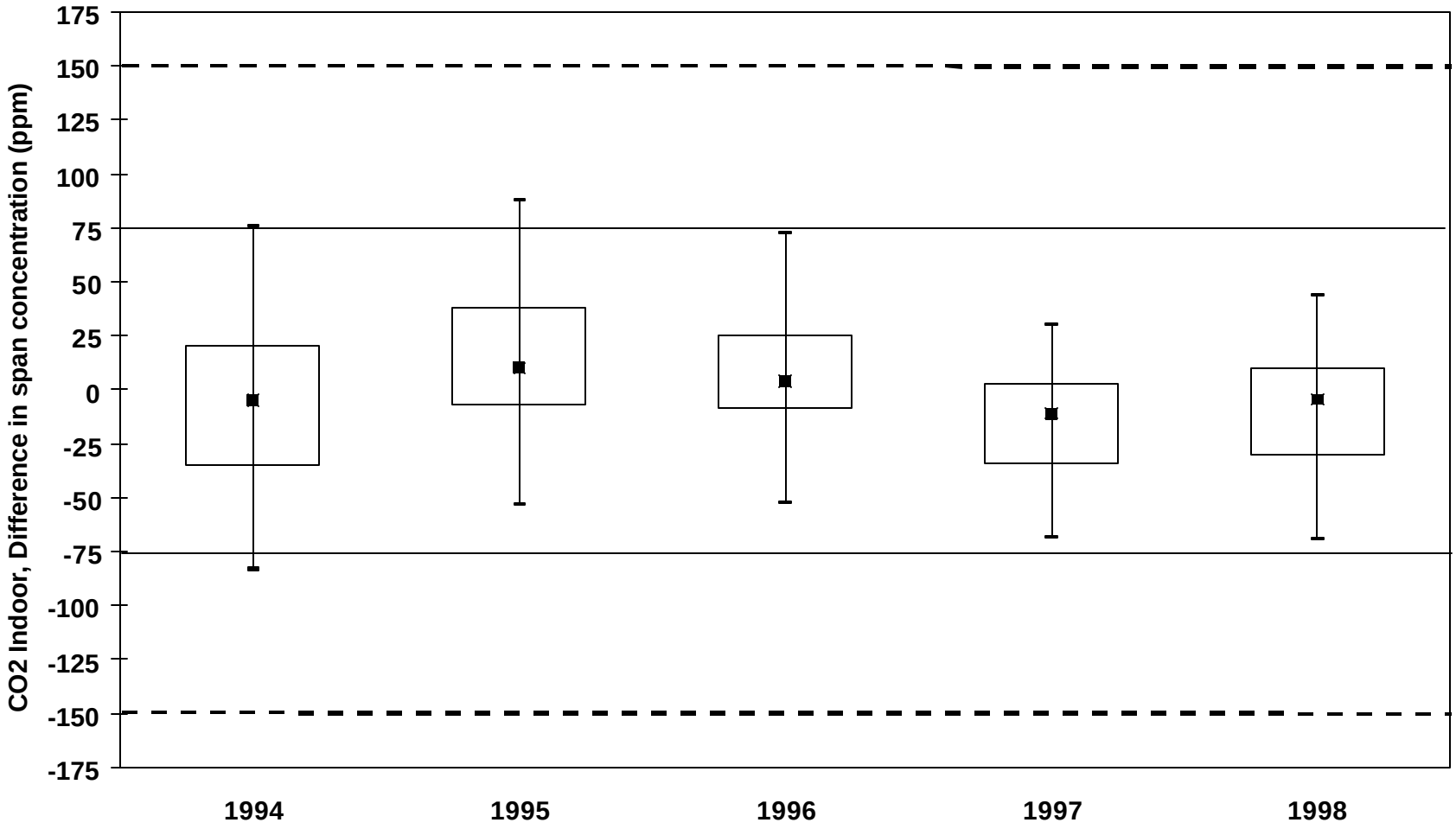


Figure D3 Differential Indoor CO_2 Sensor Response to Span Gas

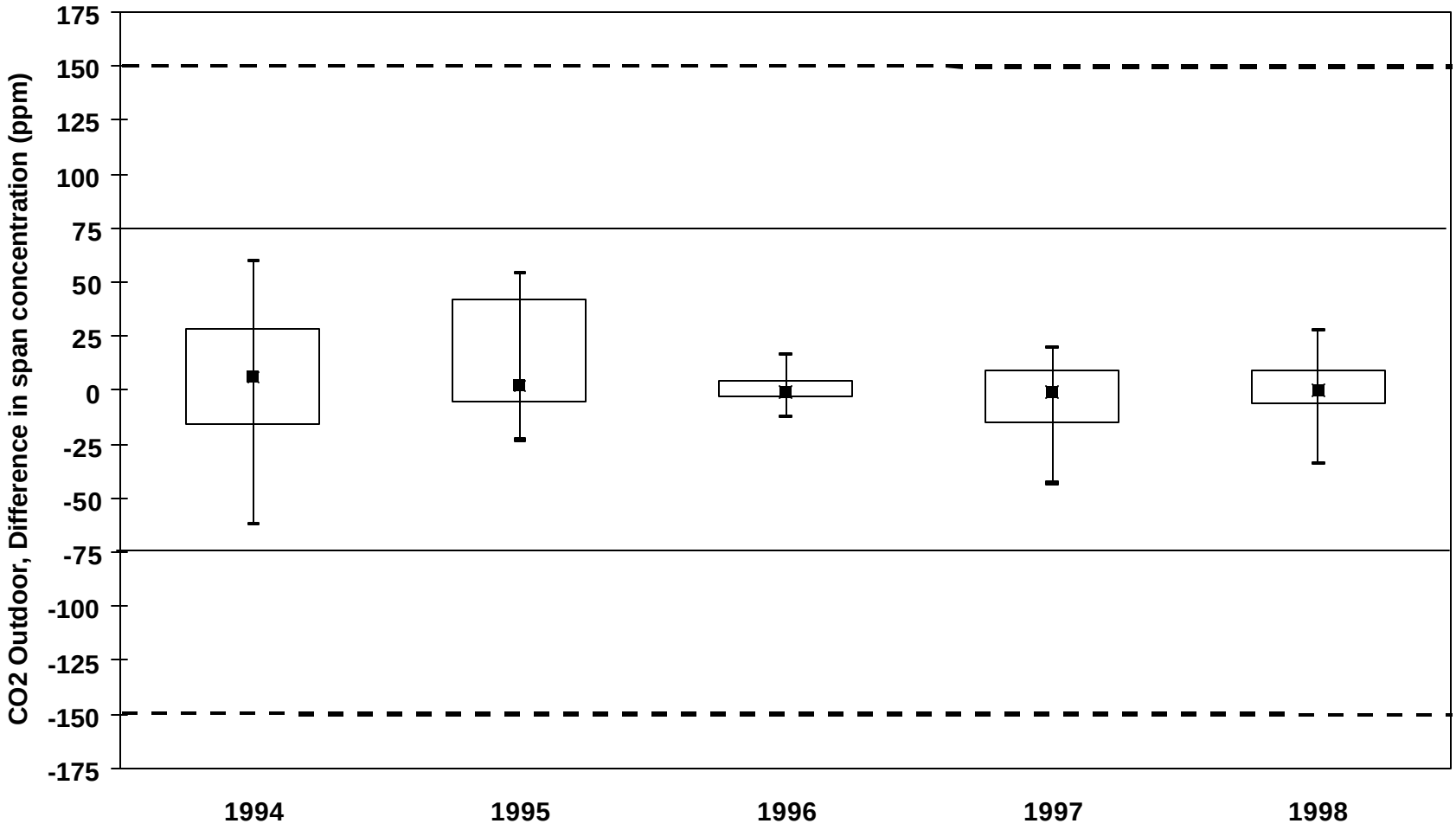


Figure D4 Differential Outdoor CO_2 Sensor Response to Span Gas

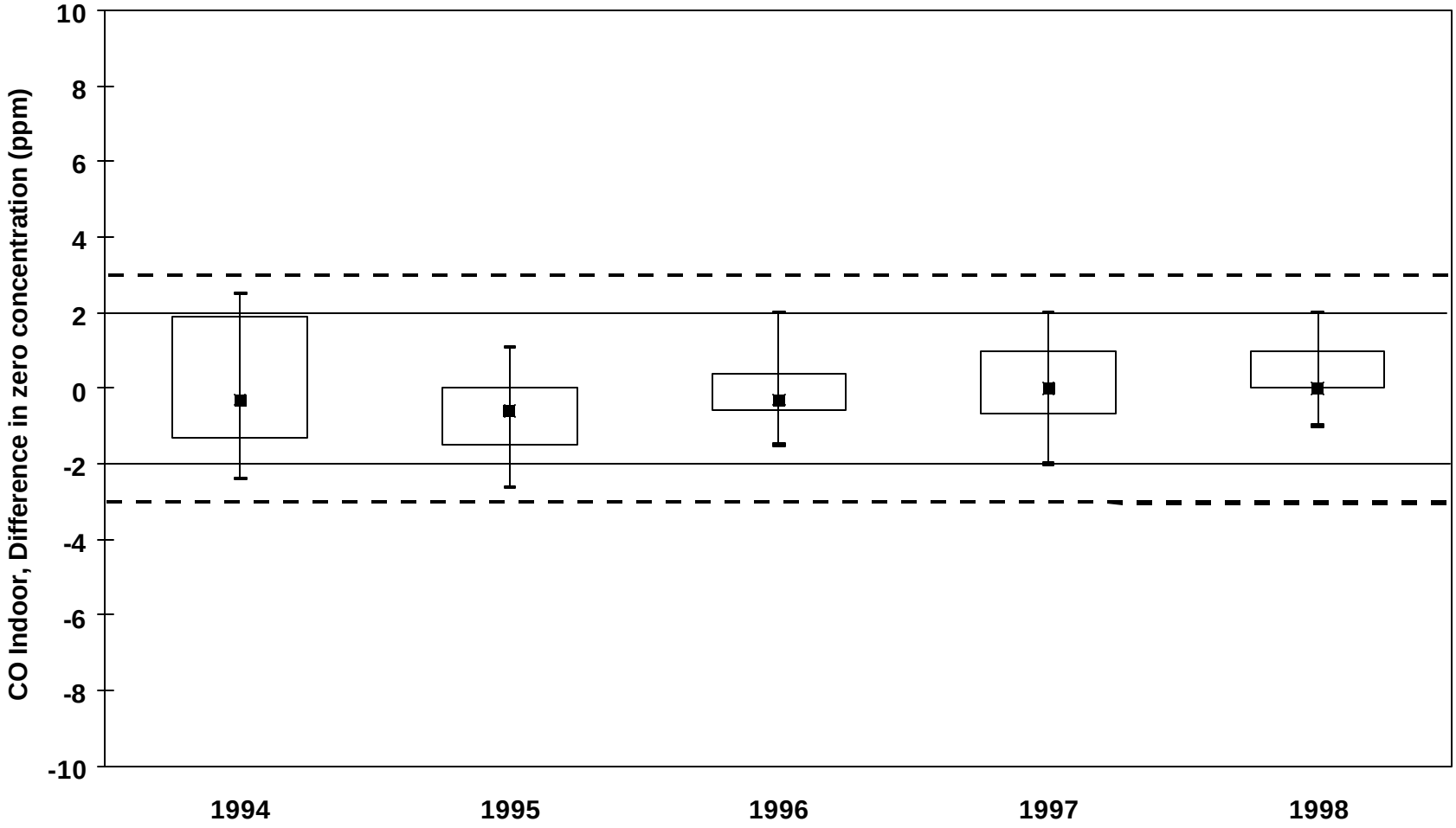


Figure D5 Differential Responses of Indoor CO Sensors to Zero Gas

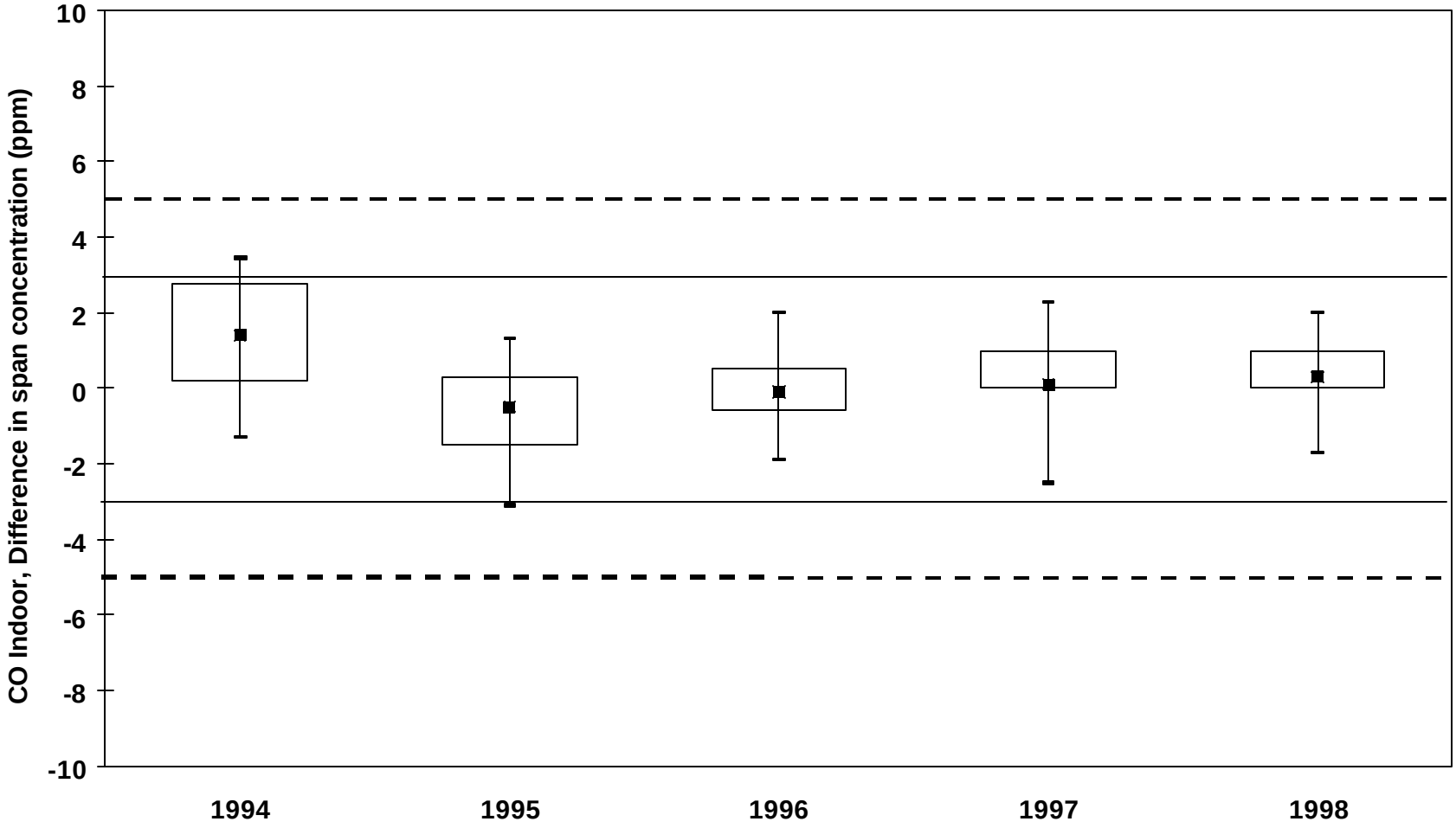


Figure D6 Differential Responses of Indoor CO Sensors to Span Gas

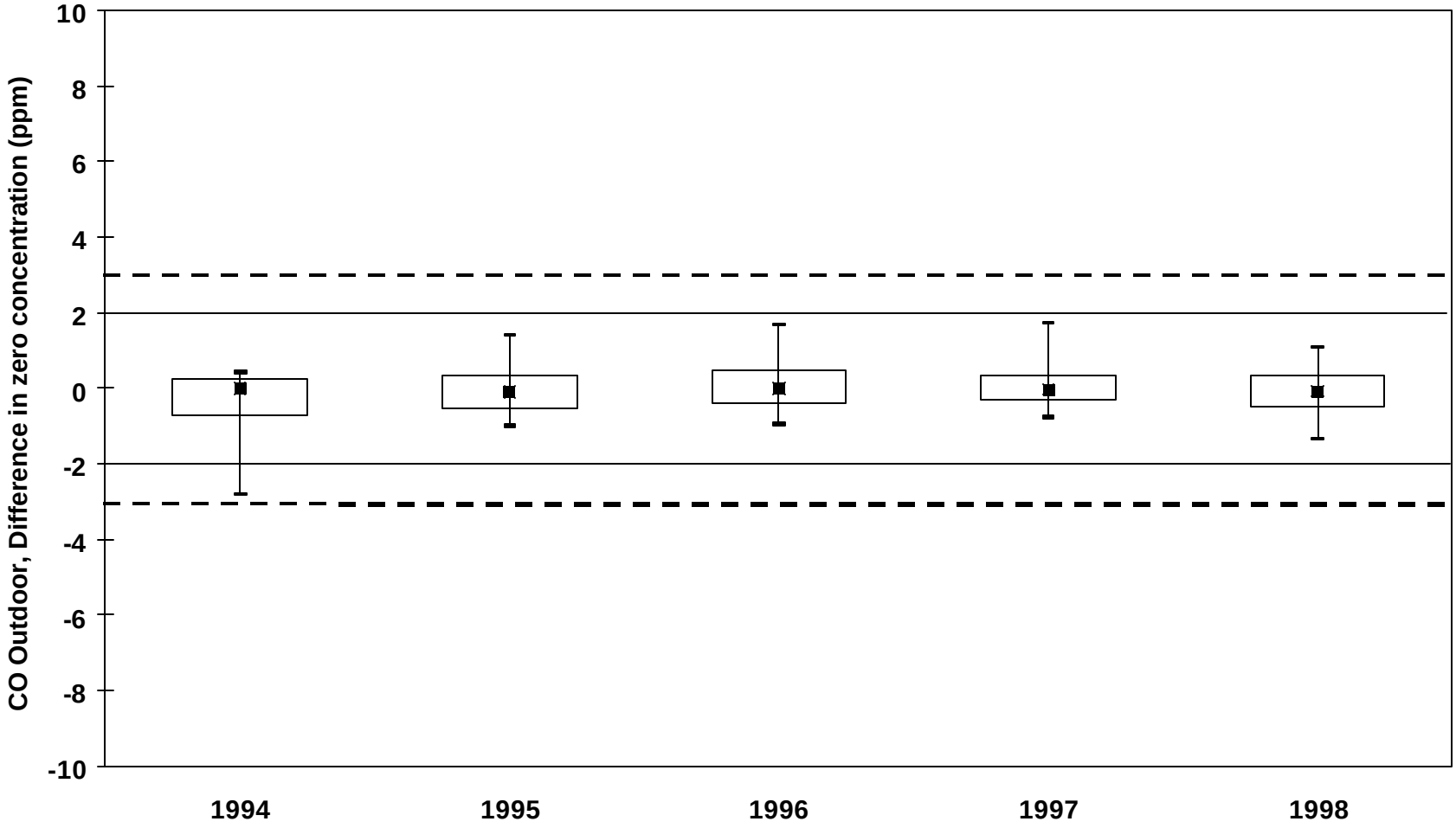


Figure D7 Differential Responses of Outdoor CO Sensors to Zero Gas

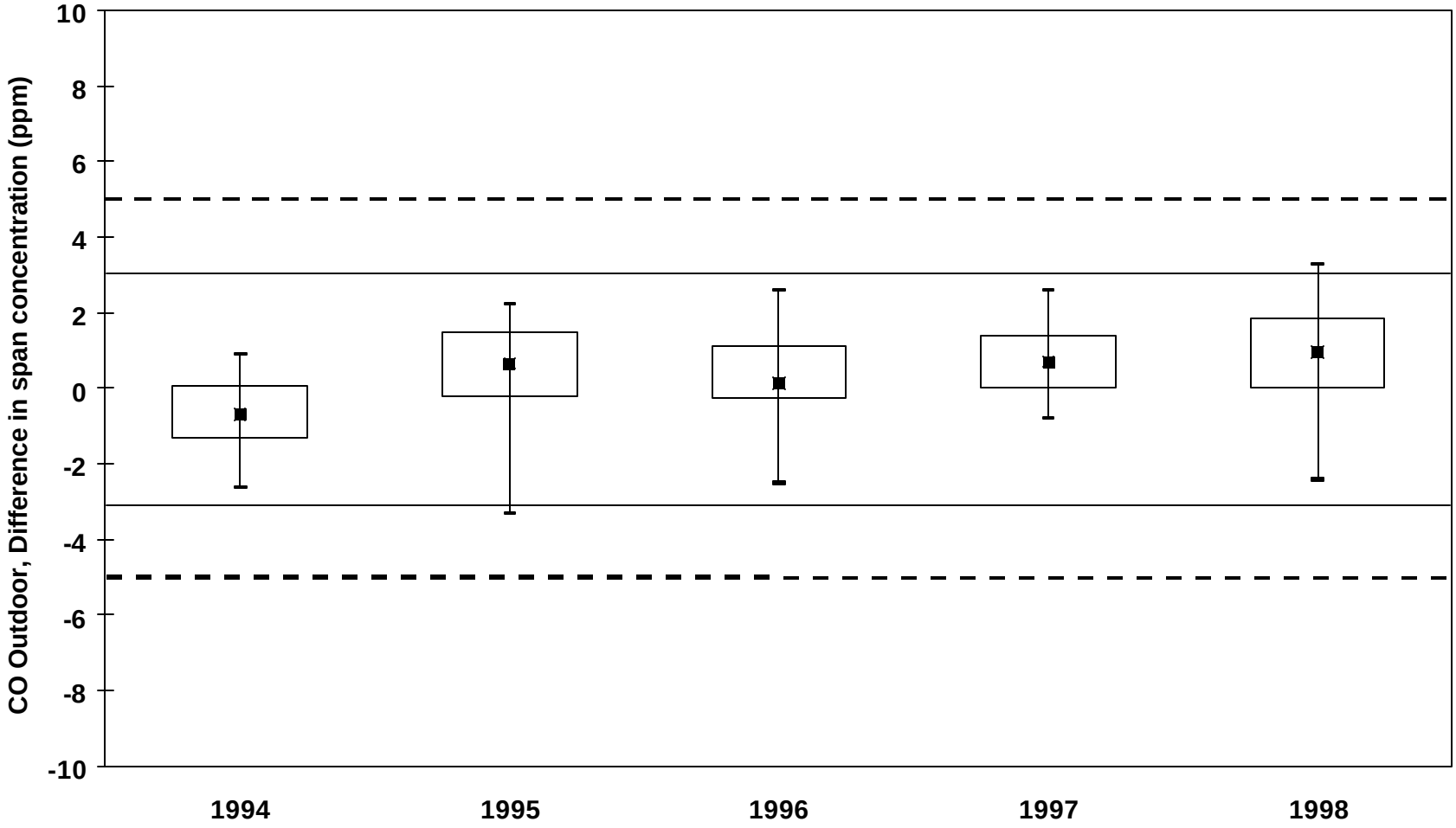


Figure D8 Differential Responses of Outdoor CO Sensors to Span Gas

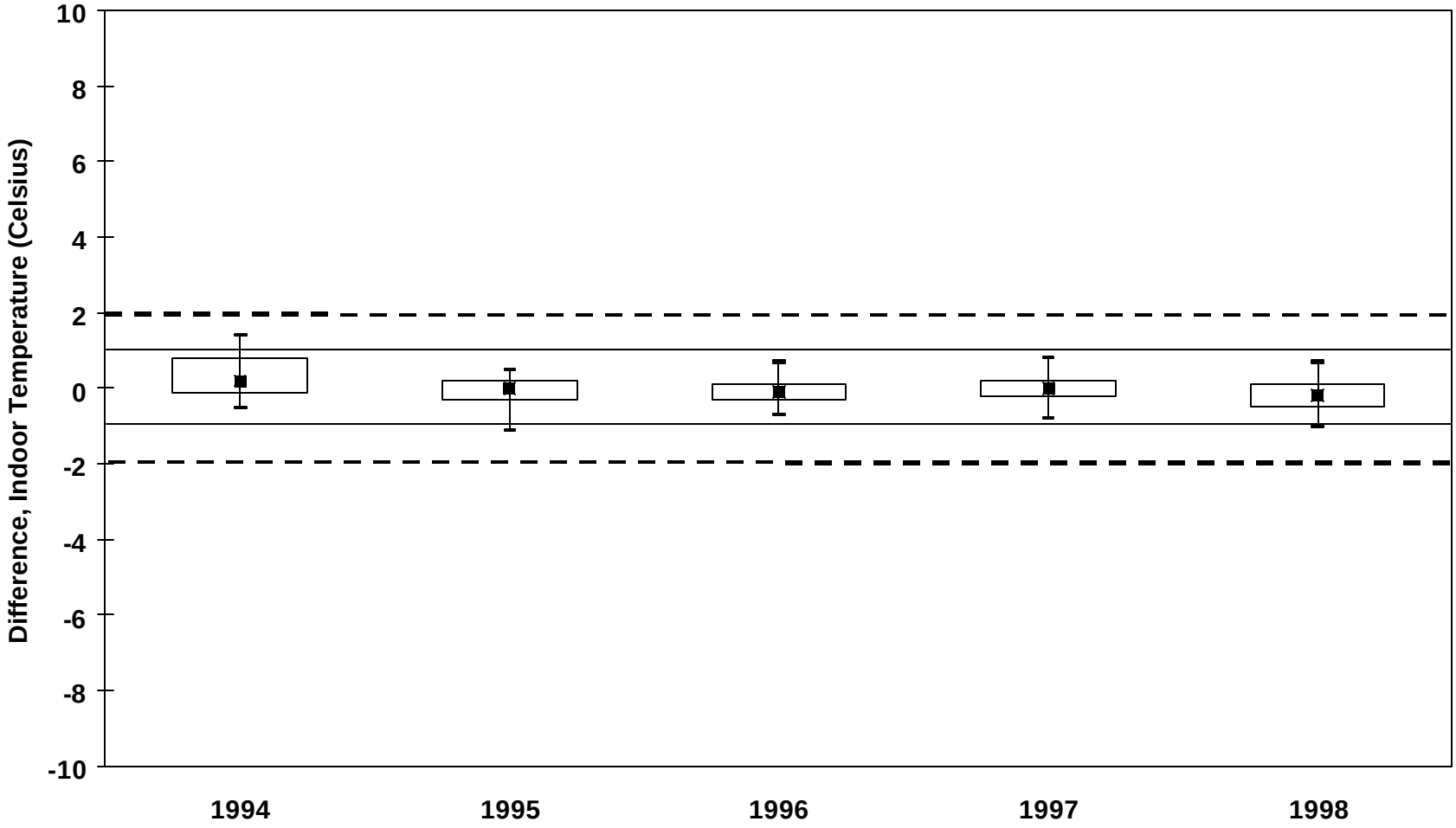


Figure D9 Differential Temperatures from Indoor Sensors Relative to Reference Thermometer

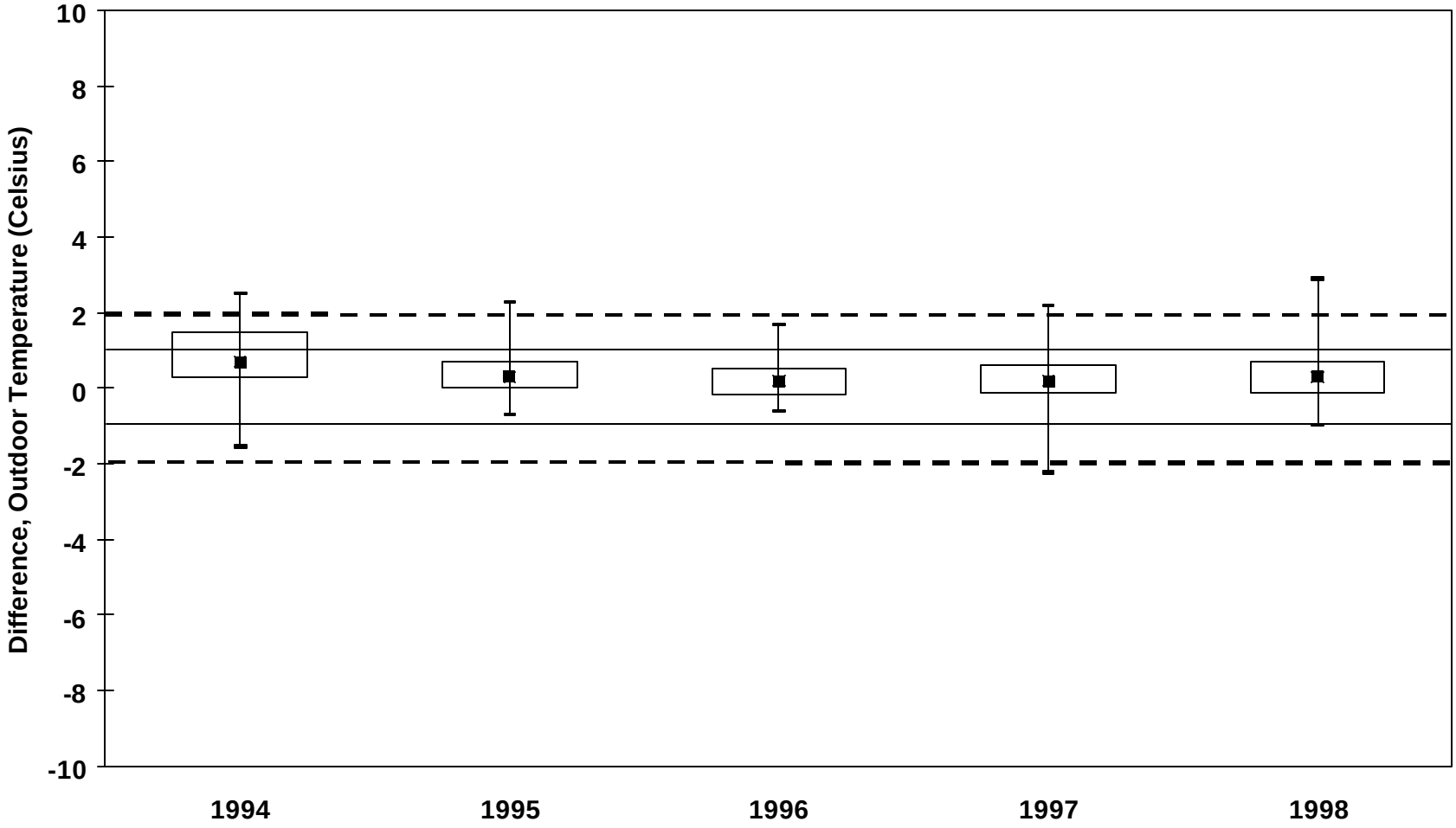


Figure D10 Differential Temperatures from Outdoor Sensors Relative to Reference Thermometer

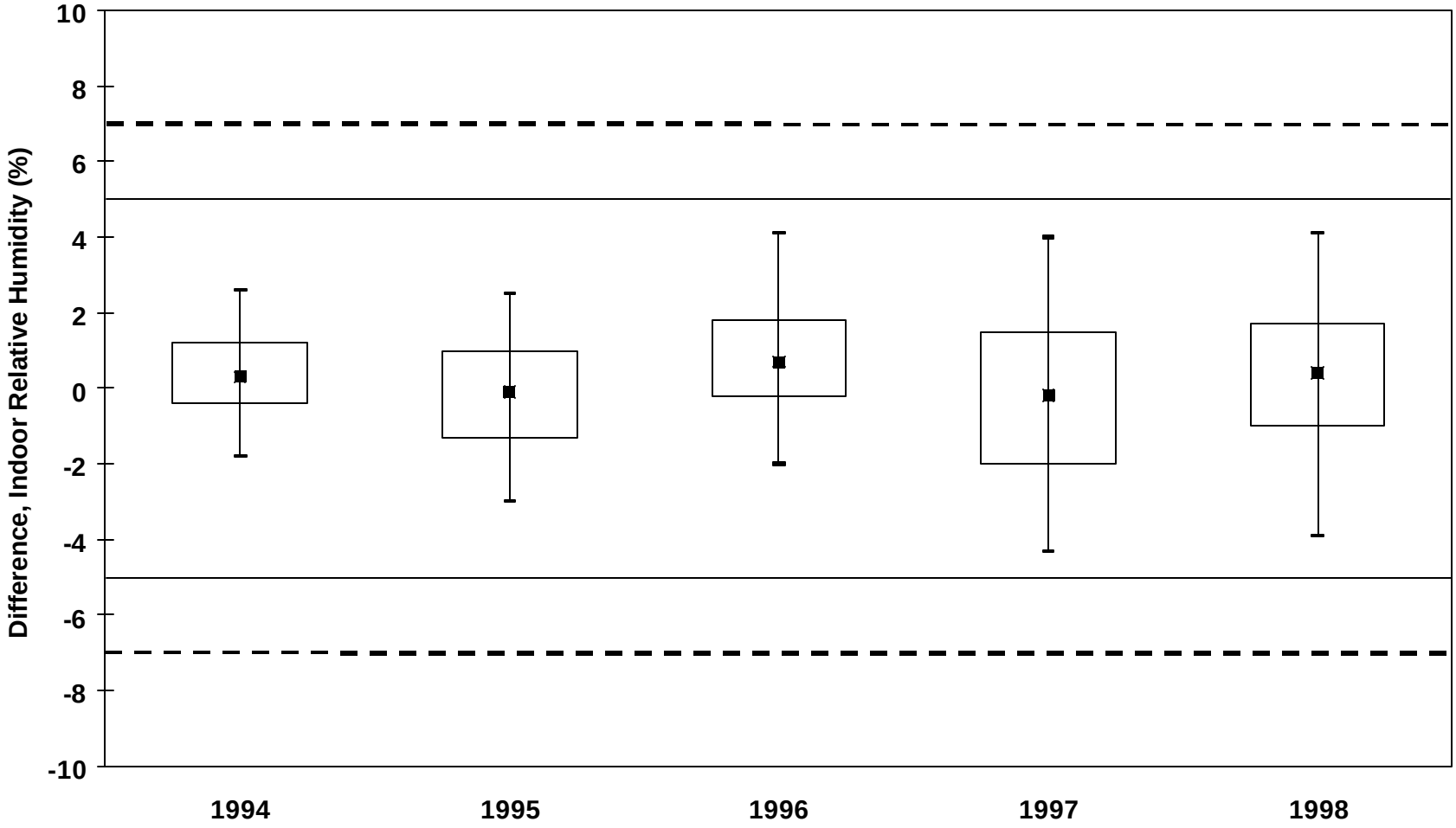


Figure D11 Differential Relative Humidities from Indoor Sensors Compared to Reference Sensor

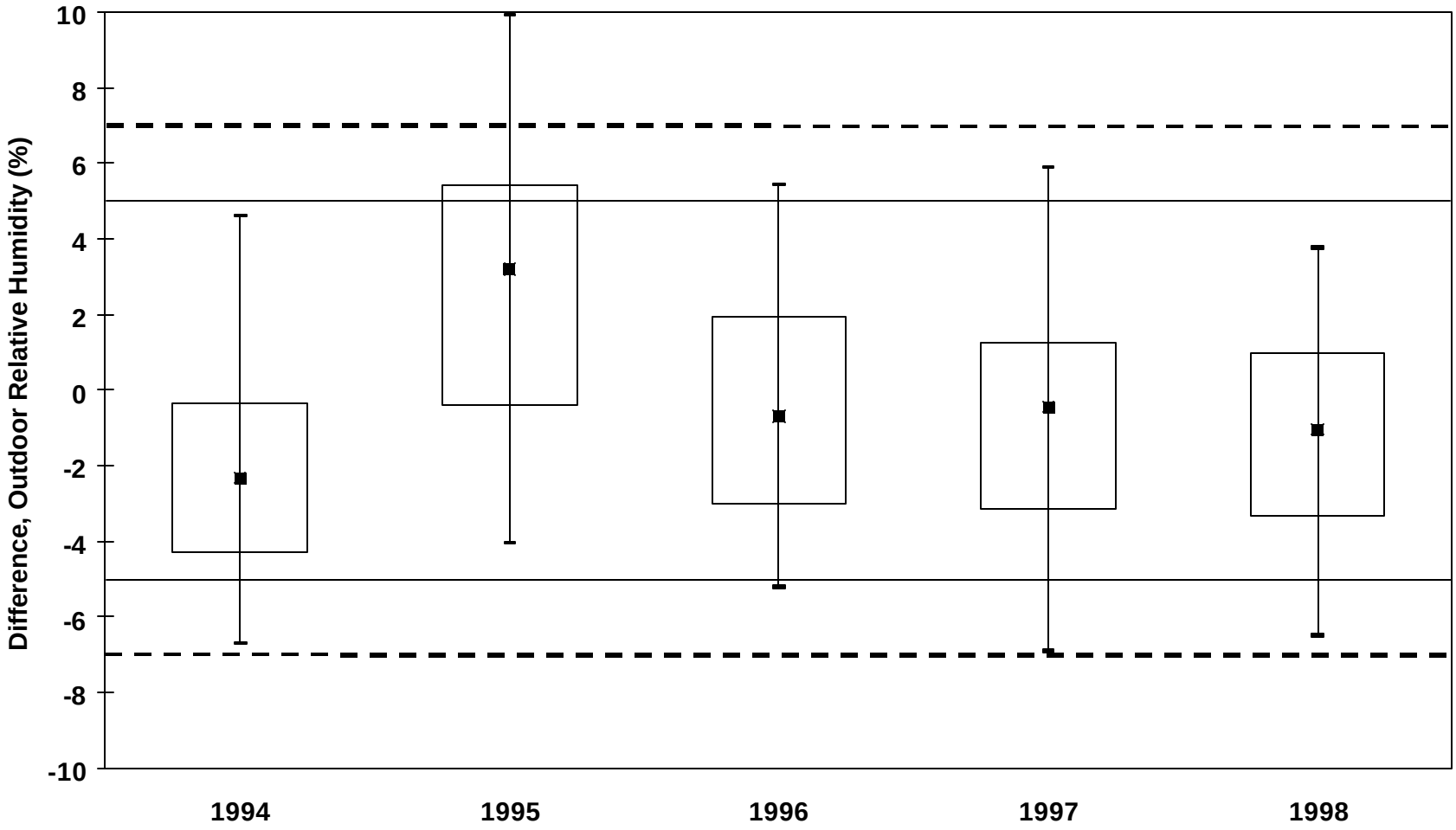


Figure D12 Differential Relative Humidities from Outdoor Sensors Compared to Reference Sensor