Attachment 4

CRITERIA FOR DESIGNATION OF EQUIVALENCE METHODS FOR CONTINUOUS SURVEILLANCE OF $PM_{2.5}$ AMBIENT AIR QUALITY

by

Office of Air Quality Planning and Standards
Office of Research and Development
U.S. ENVIRONMENTAL PROTECTION AGENCY
Research Triangle Park, North Carolina 27711

BATTELLE
505 King Avenue
Columbus, Ohio 43201-2693

DISCLAIMER

This report is a work prepared by the United States Government by Battelle. In no event shall either the United States Government or Battelle have any responsibility or liability for any consequences of any use, misuse, inability to use, or reliance upon the information contained herein, nor does either warrant or otherwise represent in any way the accuracy, adequacy, efficacy, or applicability of the contents hereof.

TABLE OF CONTENTS

List of Tables

Draft Guidance Report
on
Criteria for Designation of Equivalence Methods for Continuous Surveillance of $PM_{2.5}$ Ambient Air Quality

Introduction

This report describes the proposed methodology and criteria for labeling a continuous particulate matter (PM) sampler as being *equivalent* to Federal Reference Method (FRM) samplers for sampling $PM_{2.5}$. The methodology employed adheres to EPA regulatory requirements for designation of federal reference and equivalent methods for $PM_{2.5}$ sampling as defined in 40 CFR Part 53. Under 40 CFR Part 53, continuous samplers are defined as a Class III candidate method of sampling:

> *Class III equivalent methods include any candidate $PM_{2.5}$ methods that cannot qualify as either Class I or Class II. This class includes any filter-based integrated sampling method having other than a 24-hour $PM_{2.5}$ sample collection interval followed by moisture equilibrium and gravimetric mass. More importantly, Class III also includes filter-based continuous or semi-continuous methods, such as beta attenuation instruments, harmonic oscillating element instruments, and other complete in situ monitor types. Non-filter-based methods such as nephelometry or other optical instruments will also fall into the Class III category.*

Because a wide variety of possible candidate sampling methods fall into the Class III designation, testing procedures and performance requirements must be individually designed and adapted for the specific sampler method of interest. In 40 CFR Part 53, U.S. EPA regulations state that the specific test procedures and performance requirements for each Class III candidate method should be determined on a case-by-case basis upon request, in connection with each proposed or anticipated application for a Class III equivalent method determination. This report provides details and guidance on the proposed criteria for establishing equivalence of a continuous $PM_{2.5}$ sampler with FRM samplers. More detailed information on the derivations of the equivalency criteria are provided in Appendix A.

Additional related information on the National Ambient Air Quality Standard (NAAQS) and reference method for $PM_{2.5}$ are defined in 40 CFR Part 50 and the network requirements for surveillance of ambient air quality at State and Local Air Monitoring Stations (SLAMS) are detailed in 40 CFR Part 58.

Sampling Requirements

To determine the equivalency of a candidate continuous sampler (uniquely specified by manufacturer, brand, and model number) relative to the FRM sampler, daily concentration data need to be obtained from $PM_{2.5}$ samples collected from co-located candidate and FRM samplers at

multiple sites. To support an equivalence evaluation, the data collection process involving co-located samplers needs to adhere to the following requirements:

- Three (3) to five (5) candidate samplers will be co-located with three (3) FRM samplers. This number of FRM samplers is consistent with existing requirements and improves on the ability to identify statistical outliers in daily concentrations.

- Within a given season of the year, each sampler will be run daily for a target of 30 days (with at least one site having samples collected in multiple seasons).

- On a given day, the required sample collection period for each sampler will be a minimum of 22 hours. Although the recommended sampling period for FRM samplers is 23 to 25 hours, EPA has lowered the minimum sampling period to 22 hours for purposes of this equivalency evaluation in order to allow sufficient time to change out samplers and to perform necessary maintenance between sample runs.

- On a given day, valid data must be available for at least two (2) FRM samplers and at least two (2) candidate samplers in order for any data associated with the day's sample collection to be used in the equivalency evaluation.

- Each sampler at a given site will produce valid measurements on at least 75 percent of the sampling days in a given season. For a 30-day sampling period, this corresponds to a minimum of 23 days per season.

- The acceptable concentration range of sample data is 3 $\mu g/m^3$ to 200 $\mu g/m^3$. Although previous $PM_{2.5}$ method designations had a minimum concentration requirement of 10 $\mu g/m^3$ due to concerns about large variations in measurements as one approached zero as well as a desire to test at higher concentrations, recent experience has shown that testing at the higher range is not as representative as it used to be, and fairly repeatable concentrations can be obtained at low concentrations. Thus, the minimum of 3 $\mu g/m^3$ is consistent with proposed edits of 40 CFR Part 58, Appendix A.

Data collection will be replicated at multiple sites to ensure that the sampling is representative of different aerosol types. Furthermore, for at least one site, sampling will occur in at least two distinct seasons of the year. The above sampling requirements will hold across seasons for each site.

Appendix A outlines the development of the specific criteria in this report. Appendix B discusses concerns regarding the potential for a seasonal component to the variability of the concentrations, the expected minimum concentration CV, and related assumptions.

Notation and Intermediate Formulas

In the presentation that follows, the following notation is used:

D	=	number of sampling days. (As noted above, D must be at least 23 within a given season.)
t_i	=	number of FRM samplers utilized in the evaluation that have valid sample concentrations reported on the i^{th} sampling day (i=1, …, D). Possible values of t_i are 2 or 3.
k_i	=	number of candidate samplers utilized in the evaluation that have valid sample concentrations reported on the i^{th} sampling day (i=1, …, D). Possible values of k_i are 2 through 5.
$FRM_{i,j}$	=	concentration reported on the i^{th} sampling day (i=1, …, D) from the j^{th} FRM sampler (j=1, …, t_i) that had a valid concentration reported on that day.
$Cand_{i,j}$	=	concentration reported on the i^{th} sampling day (i=1, …, D) from the j^{th} candidate sampler (j=1, …, k_i) that had a valid concentration reported on that day.

Using this notation, intermediate formulas used in determining equivalency of a candidate sampler are as follows. These formulas are presented in a form that can be easily applied within standard spreadsheet packages.

1. The arithmetic mean of the valid sample concentrations originating from the FRM samplers on the i^{th} sampling day (i=1, …, D) is calculated as follows:

$$FRM\_daily\_mean_i = \frac{1}{t_i}\sum_{j=1}^{t_i} FRM_{i,j} \quad . \qquad \text{(Eq. 1)}$$

2. The arithmetic mean of the valid sample concentrations originating from the candidate samplers on the i^{th} sampling day (i=1, …, D) is calculated as follows:

$$Cand\_daily\_mean_i = \frac{1}{k_i}\sum_{j=1}^{k_i} Cand_{i,j} \quad . \qquad \text{(Eq. 2)}$$

3. The overall mean of the D daily means associated with the FRM sampler is computed as follows:

$$Overall\_mean = \frac{1}{D}\cdot\sum_{i=1}^{D} FRM\_daily\_mean_i \quad . \qquad \text{(Eq. 3)}$$

4. The *root mean square* of the D daily means associated with the FRM samplers is computed as follows:

$$FRM\_RMS = \sqrt{\frac{D \cdot \sum_{i=1}^{D} (FRM\_daily\_mean_i)^2 - \left(\sum_{i=1}^{D} FRM\_daily\_mean_i \right)^2}{D^2}} \quad \text{(Eq. 4)}$$

This formula is equivalent to the square root of the mean of the D squared deviations between the daily means and the overall mean for the FRM sampler.

5. The root mean square of the D daily means associated with the candidate sampler is computed as follows:

$$Cand\_RMS = \sqrt{\frac{D \cdot \sum_{i=1}^{D} (Cand\_daily\_mean_i)^2 - \left(\sum_{i=1}^{D} Cand\_daily\_mean_i \right)^2}{D^2}} \quad \text{(Eq. 5)}$$

6. The *coefficient of variation* (CV) of a set of reported measurements is equal to the standard deviation divided by the mean. The CV of the valid sample concentrations associated with the FRM samplers on the i^{th} sampling day (i=1, …, D) is calculated as follows:

$$FRM\_daily\_CV_i = \left(\frac{1}{FRM\_daily\_mean_i} \right) \bullet \sqrt{\frac{t_i \sum_{j=1}^{t_i} (FRM_{i,j})^2 - \left(\sum_{j=1}^{t_i} FRM_{i,j} \right)^2}{t_i (t_i - 1)}} \quad \text{(Eq. 6)}$$

7. The CV of the valid sample concentrations associated with the candidate samples on the i^{th} sampling day (i=1, …, D) is calculated as follows:

$$Cand\_daily\_CV_i = \left(\frac{1}{Cand\_daily\_mean_i} \right) \bullet \sqrt{\frac{k_i \sum_{j-1}^{k_i} (Cand_{i,j})^2 - \left(\sum_{j-1}^{k_i} Cand_{i,j} \right)^2}{k_i (k_i - 1)}} \quad \text{(Eq. 7)}$$

Requirements and Guidelines on the Collected FRM Concentration Data

Requirement on Precision: In this context, precision is calculated as the square root of the mean of the squared daily CVs associated with a given site. The precision associated with the FRM sampler data is calculated as:

$$FRM\_prec = \sqrt{\frac{\sum_{i=1}^{D} (FRM\_daily\_CV_i)^2}{D}} \quad . \qquad \text{(Eq. 8)}$$

The precision of the FRM sampler data is required to be **no greater than 7 percent** in order to allow any data collected at that site (including data associated with the candidate sampler) to be used in the equivalency evaluation.

***Guideline on Concentration Coefficient of Variation* (CCV):** The CCV is defined as the CV of the D daily means associated with the FRM samplers at a given site. It is computed as follows:

$$CCV = \left(\frac{1}{Overall\_mean}\right) \bullet \sqrt{\frac{D \cdot \sum_{i=1}^{D} (FRM\_daily\_mean_i)^2 - \left(\sum_{i-1}^{D} FRM\_daily\_mean_i\right)^2}{D \cdot (D-1)}} \quad . \qquad \text{(Eq. 9)}$$

For the equivalence evaluation, sampling sites and/or seasons should be chosen to maximize the likelihood that the value of CCV within each site will be **at least 0.35**. Note that this lower limit on the value of CCV represents a target rather than a requirement.

Conditions for Equivalency of a Candidate Sampler

From the daily sample concentration data to be collected from the co-located samplers at a given site, the following four essential measures will be calculated:

- Precision
- Correlation
- Multiplicative bias
- Additive bias.

A candidate sampler needs to achieve specified criteria placed on each of these four measures in order to be classified as equivalent to the FRM sampler. Values for these four measures are calculated separately for each site, and the candidate sampler needs to achieve the specified criteria at each site.

Using the notation and intermediate formulas above, formulas are presented below that specify how each of these four measures is calculated, and the criteria that the calculated measures need to satisfy are presented with the formulas. Detail is provided in Appendix A on the derivation of these formulas and the determination of the equivalence criteria. In calculating the four equivalency measures, true daily $PM_{2.5}$ concentrations at a given site are estimated from the daily means associated with the FRM samplers.

Precision: Similar to the definition given in Equation 8 above, the precision associated with the candidate sampler data is calculated as:

$$Cand\_prec = \sqrt{\frac{\sum_{i-1}^{D}\left(Cand\_daily\_CV_i\right)^2}{D}} \quad . \qquad \text{(Eq. 10)}$$

The precision of the candidate sampler data most be **no greater than 15 percent** in order for the candidate sampler to qualify for equivalency classification.

Correlation: Correlation in the daily means between the FRM and candidate samplers is calculated as follows:

$$r = \frac{D\sum_{i=1}^{D}\left(FRM\_daily\_mean_i \cdot Cand\_daily\_mean_i\right) - \left(\sum_{i-1}^{D} FRM\_daily\_mean_i\right)\cdot\left(\sum_{i=1}^{D} Cand\_daily\_mean_i\right)}{D^2 \cdot Cand\_RMS \cdot FRM\_RMS} \qquad \text{(Eq. 11)}$$

The value of this correlation **must exceed the following lower bound**, determined by the value of the CCV for the FRM sampler (Equation 9), in order to qualify for equivalency classification.

$$Correlation\ lower\ bound(CCV) = \begin{cases} 0.93 & if \quad CCV < 0.3 \\ 0.87 + 0.2 \cdot CCV & if \quad 0.3 \le CCV < 0.4 \\ 0.95 & if \quad 0.4 \le CCV \end{cases} \qquad \text{(Eq. 12)}$$

Thus, the value of this lower bound is determined by the value of CCV. It equals 0.93 when CCV is no higher than 0.3, increases linearly from 0.93 to 0.95 as the value of CCV increases from 0.3 to 0.4, then equals 0.95 for all values of CCV above 0.4. Details on the derivation of this lower bound are given in Appendix A.

Multiplicative bias: The multiplicative bias is calculated as the correlation (Equation 11) multiplied by the ratio of the root mean squares between the candidate and FRM samplers (Equations 6 and 7):

$$b = r \cdot \frac{Cand\_RMS}{FRM\_RMS} \quad . \qquad \text{(Eq. 13)}$$

The multiplicative bias associated with the candidate sampler **must fall between 0.90 and 1.10** in order for the candidate sampler to qualify for equivalency classification.

Additive bias: The additive bias is dependent on the daily means associated with the candidate sampler (Equation 2), the overall mean associated with the FRM sampler (Equation 3), and the calculated value for multiplicative bias (Equation 12). The formula for additive bias is as follows:

$$a = \left(\frac{1}{D} \cdot \sum_{i=1}^{D} Cand\_daily\_mean_i \right) - b \cdot Overall\_mean \qquad \text{(Eq. 14)}$$

The additive bias associated with the candidate sampler **must fall between $a_1(b)$ and $a_2(b)$** in order to qualify for equivalency classification. These lower and upper bounds, which are linear functions of the multiplicative bias (b), are as follows:

$$a_1(b) = 15.05 - 0.92(18.8)b = 15.05 - 17.31b \qquad \text{(Eq. 15)}$$

$$a_2(b) = 15.05 - 1.08(12.2)b = 15.05 - 13.20b \qquad \text{(Eq. 16)}$$

Details on the derivation of this acceptance region are given in Appendix A.

Example

As an illustrative example, the equivalence evaluation process described above has been applied to the set of simulated $PM_{2.5}$ concentration data presented in Table 1. In this example, which is a simulation and not based upon actual data, three co-located samplers of each type (candidate and FRM) were placed at a single site. Sampling occurred within two distinct seasons of the year. Within each season, valid data from 23 runs were available for a total of D=46 days. As seen within Table 1, all reported daily concentrations fell within the required range of 3 $\mu g/m^3$ to 200 $\mu g/m^3$ except for the candidate samplers on Day 4. All data were deemed valid for use in the evaluation.

For both the candidate and FRM sampler types within each site, daily means and CVs were calculated across co-located samplers from the data reported in the first six columns of Table 1, using Equations 1, 2, 6, and 7 above, and these statistics are reported in the last four columns of Table 1. The overall FRM mean (Equation 3) is the mean of the daily means for the FRM samplers across all sampling days, and the value of this statistic are given in the second row of Table 2. Within the last three rows of Table 2, the calculated root mean square (Equation 4), the precision calculation (Equation 8), and the CCV (Equation 9) for the FRM samplers are presented. Note from the third row of Table 2 that FRM precision achieves the requirement of being below 7 percent. In addition, the value of CCV (last row of Table 2) is above the target of 0.35.

Table 3 reports the results of statistics calculated on the candidate sampler data. Specifically, rows 3 through 6 of this table represent the four measures used to determine equivalency of the candidate sampler relative to the FRM sampler. The findings reported in Table 3 are as follows:

- The precision value is 15 percent, and therefore, the precision criterion is achieved.
- Because the value of CCV (Table 2) is above 0.4, the lower bound on the correlation is 0.95. The correlation value exceeds 0.95 and, therefore, the correlation criterion is achieved.
- Multiplicative bias is within the range of 0.90 to 1.10 and, therefore, the criterion for multiplicative bias is achieved.
- Given the value of multiplicative bias, additive bias falls within the range of (-3.3, 1.1) and, therefore, the criterion for additive bias is achieved.

By applying the data reported in this example to the evaluation criteria presented earlier, the candidate sampler would have been declared equivalent to the FRM sampler.

Table 1. Listing of Simulated $PM_{2.5}$ Concentration Data ($\mu g/m^3$) Used in Example

Run	Daily Concentrations						Daily Means and CVs			
	FRM Samplers			Candidate Samplers			FRM Samplers		Candidate Samplers	
	#1	#2	#3	#1	#2	#3	Mean	CV	Mean	CV
Season 1										
1	6.4	5.8	6.6	5.5	5.0	3.7	6.3	0.066	4.7	0.196
2	6.9	6.4	7.2	4.4	6.5	4.3	6.8	0.059	5.1	0.245
3	5.6	5.9	5.6	4.5	3.5	4.3	5.7	0.030	4.1	0.129
4	4.4	4.4	4.5	2.7	2.4	2.2	4.4	0.013	2.4	0.103
5	7.9	8.9	8.4	7.5	6.0	7.6	8.4	0.060	7.0	0.127
6	8.0	7.7	8.2	7.2	5.8	6.3	8.0	0.032	6.4	0.110
7	9.0	10.3	9.1	5.4	10.4	8.3	9.5	0.076	8.0	0.313
8	16.9	18.6	15.8	16.2	17.8	15.6	17.1	0.082	16.5	0.069
9	10.3	10.2	11.5	9.6	10.8	7.6	10.7	0.068	9.3	0.173
10	11.1	11.7	11.6	10.9	8.6	12.5	11.5	0.028	10.7	0.184
11	10.9	10.7	10.5	11.9	8.6	10.1	10.7	0.019	10.2	0.162
12	11.2	10.9	11.0	9.1	9.5	12.0	11.0	0.014	10.2	0.154
13	10.2	10.2	10.2	9.2	8.5	8.4	10.2	0.000	8.7	0.050
14	5.3	5.6	5.5	3.6	4.2	2.8	5.5	0.028	3.5	0.199
15	7.8	8.8	8.5	5.0	6.7	5.5	8.4	0.061	5.7	0.152
16	5.6	6.4	6.2	5.3	4.5	5.3	6.1	0.069	5.0	0.092
17	11.3	10.2	10.8	7.8	8.7	9.3	10.8	0.051	8.6	0.088
18	7.0	7.9	7.1	6.3	4.5	6.9	7.3	0.067	5.9	0.212
19	6.6	6.0	6.0	4.7	5.4	4.0	6.2	0.056	4.7	0.149
20	6.1	6.3	6.4	5.7	5.4	4.3	6.3	0.024	5.1	0.144
21	6.7	6.3	7.3	4.3	5.6	7.6	6.8	0.074	5.8	0.285
22	9.8	9.1	9.4	7.4	7.7	10.3	9.4	0.037	8.5	0.188
23	12.5	13.5	14.1	15.4	12.4	10.3	13.4	0.060	12.7	0.202
Season 2										
24	16.4	16.1	15.0	15.0	18.1	19.4	15.8	0.047	17.5	0.129
25	26.7	22.2	24.4	21.2	20.8	20.0	24.4	0.092	20.7	0.030
26	12.5	10.5	10.0	10.8	10.1	11.0	11.0	0.120	10.6	0.044
27	15.6	15.0	14.9	11.5	16.1	13.9	15.2	0.025	13.8	0.166
28	20.8	22.2	20.4	20.8	17.3	19.9	21.1	0.045	19.3	0.094
29	19.7	20.3	20.0	22.8	17.6	22.4	20.0	0.015	20.9	0.138
30	5.4	5.2	5.3	3.6	3.1	4.3	5.3	0.019	3.7	0.164
31	7.0	7.3	6.9	4.8	7.2	5.9	7.1	0.029	6.0	0.201
32	17.1	14.9	16.1	15.9	17.4	15.5	16.0	0.069	16.3	0.062
33	12.5	12.6	11.3	10.9	8.4	11.5	12.1	0.060	10.3	0.160
34	9.7	10.1	10.1	9.1	9.6	7.9	10.0	0.023	8.9	0.099
35	14.8	15.4	16.3	13.4	14.6	14.1	15.5	0.049	14.0	0.043
36	19.4	19.7	19.8	18.5	15.9	18.5	19.6	0.011	17.6	0.085
37	17.1	15.7	17.0	14.0	18.1	18.2	16.6	0.047	16.8	0.143
38	14.1	14.2	14.0	17.2	16.9	14.0	14.1	0.007	16.0	0.110
39	11.6	10.8	11.0	10.2	8.1	10.6	11.1	0.037	9.6	0.139
40	12.5	12.7	13.6	11.3	12.7	10.6	12.9	0.045	11.5	0.093
41	11.1	11.6	12.0	10.8	8.5	11.9	11.6	0.039	10.4	0.167
42	14.2	14.3	14.9	13.4	13.6	14.6	14.5	0.026	13.9	0.046
43	22.8	23.5	20.3	24.3	17.9	22.0	22.2	0.076	21.4	0.151
44	14.4	17.3	16.6	17.1	19.7	15.9	16.1	0.094	17.6	0.111
45	18.2	18.9	17.1	19.7	22.2	17.0	18.1	0.050	19.6	0.132
46	17.2	20.6	19.5	19.5	17.5	17.0	19.1	0.091	18.0	0.073

Table 2. Statistics Calculated from the FRM Sampler Data

Overall Mean (Equation 3)	11.95 $\mu g/m^3$
RMS (Equation 4)	5.05
Precision (Equation 8)	0.055
CCV (Equation 9)	0.427

Table 3. Statistics Calculated from the Candidate Sampler Data

RMS (Equation 5)	5.44
Precision (Equation 10)	0.150
Correlation (Equation 11)	0.981
Multiplicative bias (Equation 12)	1.058
Additive bias (Equation 13)	-1.692
Lower and upper bounds on additive bias (Equations 15 and 16)	(-3.3, 1.1)

APPENDIX A:

DERIVATIONS OF THE EQUIVALENCY CRITERIA

Appendix A: Derivations and Justifications for the Equivalency Criteria

This appendix outlines the main calculations performed in developing the specific equivalency criteria. It provides supplemental information that is not needed for applying the criteria, but should be useful in finalizing the criteria.

Recall that four measures will be calculated when determining whether a continuous $PM_{2.5}$ sampler can be classified as equivalent to FRM samplers:

- Precision
- Correlation
- Multiplicative bias
- Additive bias.

Data to be used in calculating these measures will be generated by collecting samples from three co-located FRM samplers and three to five co-located candidate samplers (of the same type, brand, and model) at multiple sites over approximate 30-day periods in a given season. For at least one site, sampling will occur in at least two distinct seasons of the year.

Expected Population Variability

The quality of the criteria estimates and, in the case of the correlation, the magnitude of the estimate, is affected by the variability associated with the true $PM_{2.5}$ concentrations at the given site. For this reason, variability has played a key role in the DQO development process for $PM_{2.5}$ and has become a widely-studied topic of interest. Historical estimates can be used for selecting sites and setting reasonable expectations for the equivalency criteria development.

The DQO model for $PM_{2.5}$ features three factors for characterizing variability in the underlying population:

- population coefficient of variation (CV)
- autocorrelation
- seasonality ratio.

Data for a site are modeled with autocorrelated, log-normally distributed random deviations with a constant CV from a sinusoidal curve having a period of one year. The parameters for the sinusoidal curve, the autocorrelation, and population CV were all evaluated at the site level using data collected from 1999 though 2001 (Coutant and Holloman, 2003). This allowed for locally relevant evaluations of the DQOs and national evaluations based on the extremes in the parameter estimates. To set expectations for equivalency testing, the ranges and median of these parameter estimates should be considered.

Figures A-1 through A-3 show the distributions of the parameter estimates. Figure A-1 shows that approximately 95 percent of all the sites in the network had monthly means that differed by at least a factor of 1.5. Figure A-2 shows that approximately 95 percent of all the

sites had a population CV estimate greater than 40 percent. The sites that do not sample daily were automatically given an autocorrelation estimate of zero. Figure A-3 shows the distribution of the non-zero autocorrelation estimates. The autocorrelation acts to reduce the observed CV for time intervals that are short compared to the annual period. Hence, for any given site and month, a CV of 30 percent should be close to the minimum (Appendix B). However, by choosing samples from different seasons, the seasonality effect can be used to obtain a larger CV across the data from a specific site.

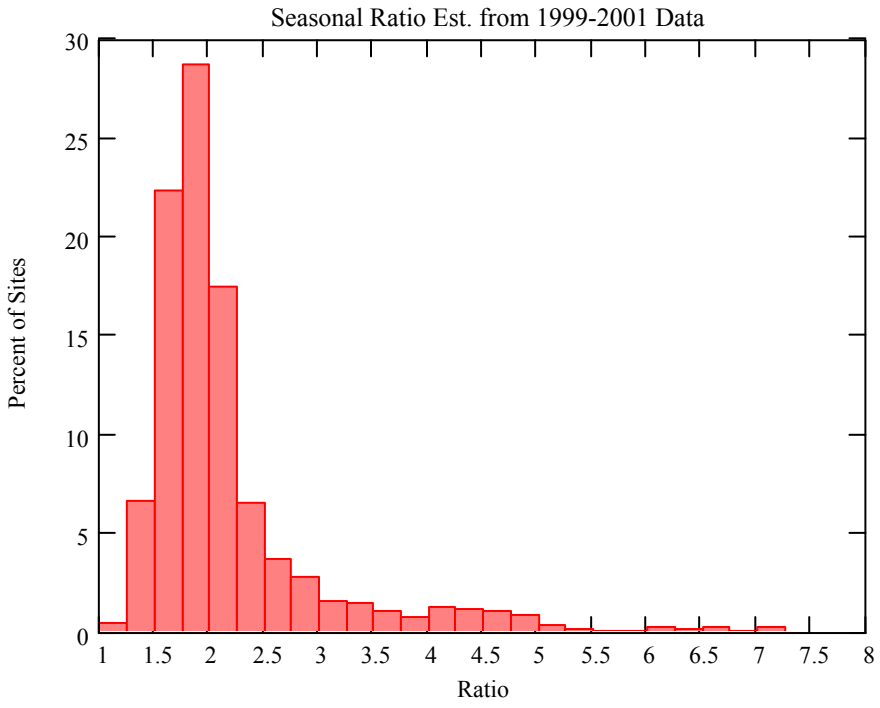


Figure A-1. The distribution of the estimated $PM_{2.5}$ seasonality ratios from 1999 to 2001 data.

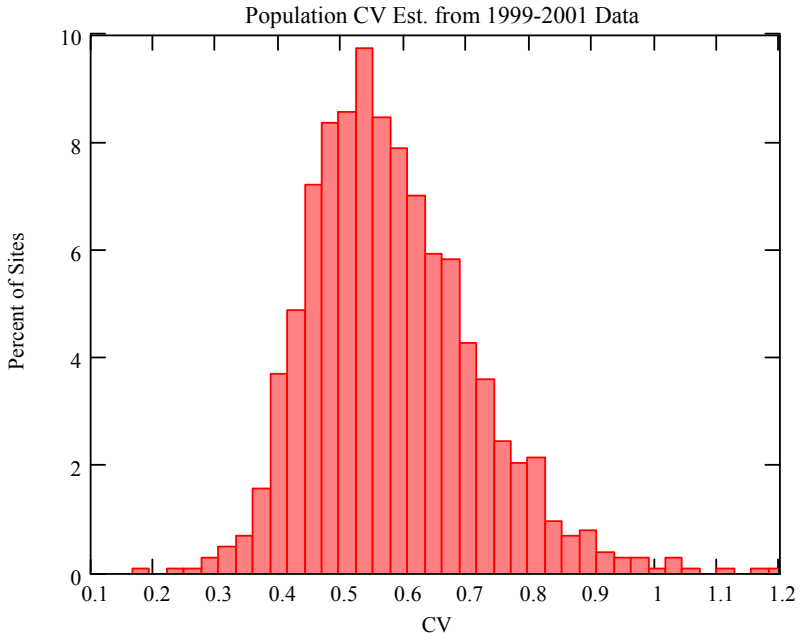


Figure A-2. The distribution of the estimated $PM_{2.5}$ population CV (after removing seasonality) from 1999 to 2001 data.

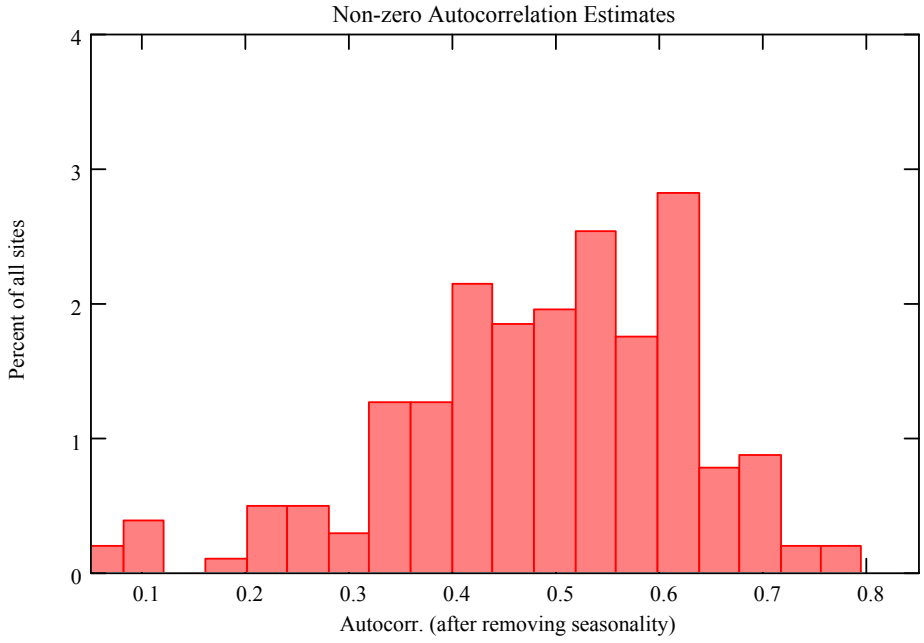


Figure A-3. The distribution of the non-zero estimates of $PM_{2.5}$ autocorrelation (after removing seasonality) from 1999 to 2001 data.

Expected Correlation and Candidate Sampler Precision

Correlation is a measure of how well a linear error model fits the data. While unlikely, samplers with seasonal differences in how well they predict the FRM measurements could fail to meet the correlation requirement even though they may have total square errors less than the amounts implied by the other requirements. This is, in fact, a desirable characteristic of the criteria. The reason is that significant seasonal effects would require extensive testing (at least a year per site) to assure that the sampler would always be within a reasonable range of an FRM. On the other hand, consistent biases are much easier to estimate and control. Given that the aim is to have equivalency testing criteria that can be done with much less than a year of data, this property is desirable.

The expected correlation between the daily means of n_1 reference samplers with precision σ_r, n_2 candidate samplers with precision σ_c and over a period with a population CV (estimated by CCV, Equation 9) of τ_u can be written as follows:

$$\frac{1}{\sqrt{\left[1+\sigma_r^2\frac{1}{n_1}\cdot\left(1+\tau_u^{-2}\right)\right]\cdot\left[\left[1+\sigma_c^2\frac{1}{n_2}\cdot\left(1+\tau_u^{-2}\right)\right]\right]}} \qquad \text{(Eq. A-1)}$$

This formula is derived in Mosquin and McElroy (2004). Notice that this formula is dependent on the population CV.

Figure A-4 shows plots of the expected correlation for three reference samplers with a precision of 5 percent and three candidate samplers with a precision of 15 percent over a range of population CVs. This shows the sharp dependency on the population CV below 40 percent and the relative insensitivity to the number of FRM samplers. Figure A-5 shows the case for three reference samplers with a precision of 5 percent and three candidate samplers with a precision of 10 percent. In this case, the sensitivity to the population CV is much reduced. This is one of the reasons for setting the limit on the candidate precision to, at most, 15 percent. The reason for the limit is not driven by how well the NAAQS decision can be made, but rather how well one can estimate the parameters that do effect the NAAQS decision.

Figure A-4 also shows the proposed lower limit on the correlation, calculated as follows:

$$Correlation\ lower\ bound(CCV)=\begin{cases}0.93 & if \quad CCV<0.3\\ 0.87+0.2\cdot CCV & if \quad 0.3\le CCV<0.4\\ 0.95 & if \quad 0.4\le CCV\end{cases} \qquad \text{(Eq. A-2)}$$

where CCV is the population CV estimate for the site. For a fixed sampler precision it will be both to EPA's and the manufacture's benefit to choose sites for equivalency testing that are expected to have a large population CV (above 40 percent). This would provide for some

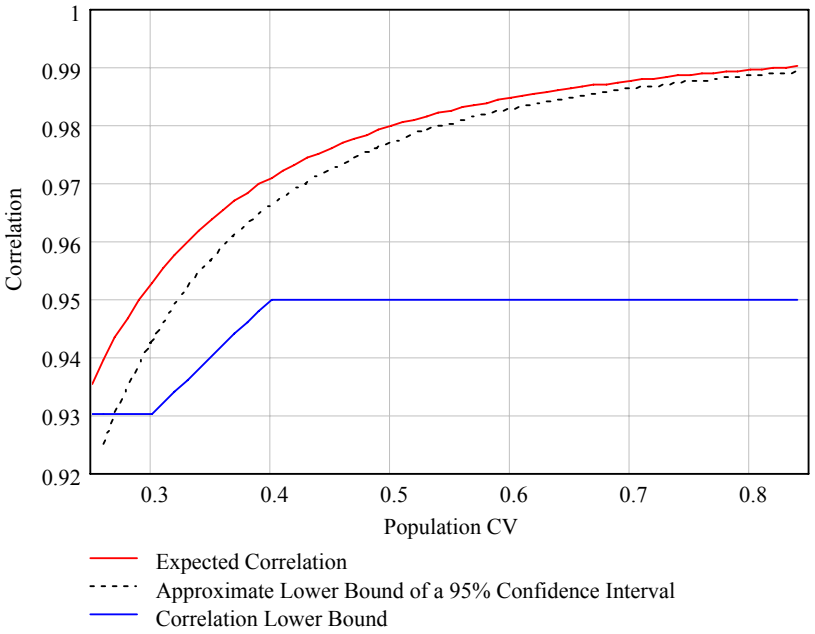


Figure A-4. Expected correlation between 46 daily means of three FRM samplers and three candidate samplers with a precision of 15 percent.

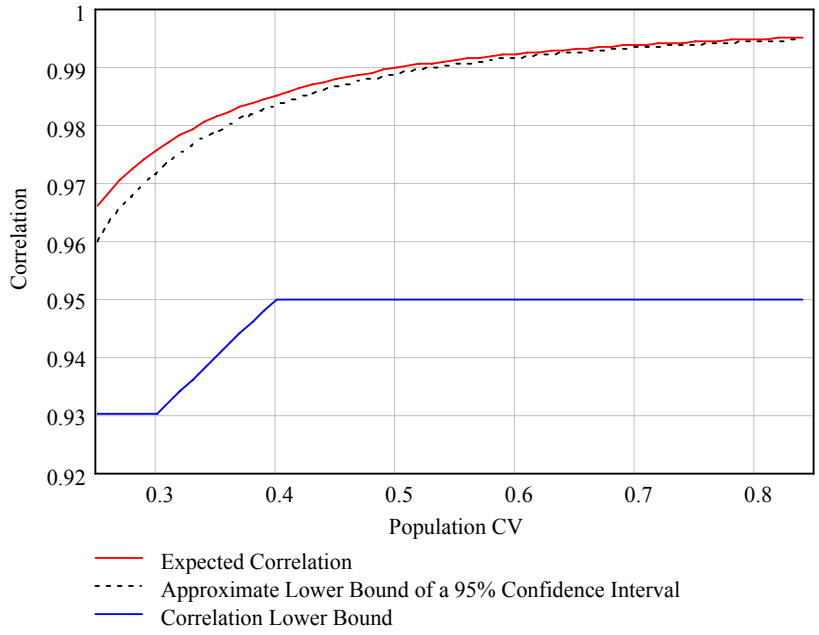


Figure A-5. Expected correlation between 46 daily means of three FRM samplers and three candidate samplers with a precision of 10 percent.

sampling error in the correlation estimate for the manufacture and as will be seen, will improve the bias estimates for EPA.

The approximate lower bound of a 95 percent confidence interval shown in Figure A-4 is based on the following transformation of the correlation, r:

$$transform\ (r, n) = \frac{r \cdot \sqrt{n-2}}{\sqrt{1-r^2}} \qquad \text{(Eq. A-3)}$$

where n is the number of sample pairs. This transformation is approximately distributed as a Student t-distribution with n-2 degrees of freedom. (Note that both the correlation estimate and the above transformation ignore the fact that the FRM samplers have a small amount of measurement error.)

The Acceptance Region for the Additive and Multiplicative Bias

In Mosquin and McElroy (2004), a proposed acceptance region was derived for the additive and multiplicative bias components. The proposed region was based on 1-in-6 day sampling in order to account for increased error rate by local agencies. Because this was felt to be too conservative, a corresponding region for 1-in-3 day sampling was derived and is presented below.

The objective was to find combinations of additive and multiplicative bias that would yield the same gray zone that had been established for the 1-in-6 day FRM samplers (i.e., 12.2 $\mu g/m^3$ to 18.8 $\mu g/m^3$), but with a higher sampling rate. First, the DQO Companion software tool (Battelle, 2003) was used to derive a gray zone for 1-in-3 day sampling assuming a 10 percent measurement CV and absolute bias. (All parameters were left at the default values except for the sampling rate.) This yielded a gray zone that ranged from approximately 12.64 $\mu g/m^3$ to 18.16 $\mu g/m^3$. (Because the software tool uses random simulations, results will vary among different runs.) From these results, the 5th and 95th percentiles were determined by solving the following two equations for x and y:

$$15.05 - 18.16(0.9)x = 0 \qquad \text{(Eq. A-4)}$$

$$15.05 - 12.64(1.1)y = 0 \qquad \text{(Eq. A-5)}$$

Here, x = 0.92 and y = 1.08.

Consequently, the same simulations can be used to derive the following acceptable range of additive bias:

$$a_1(b) = 15.05 - 0.92(18.8)b = 15.05 - 17.31b \qquad \text{(Eq. A-6)}$$

$$a_2(b) = 15.05 - 1.08(12.2)b = 15.05 - 13.20b \qquad \text{(Eq. A-7)}$$

This resulted in a gray zone of 12.2 μg/m^3 to 18.8 μg/m^3 for a given multiplicative bias b (Mosquin and McElroy, 2004)[1].

EPA has set the bounds on the multiplicative bias, b, to be ±10 percent from 100 percent (i.e., $0.9 \leq b \leq 1.1$). Under this range, the box in Figure A-6 represents the region of values for multiplicative and additive biases that will correspond to gray zones that range from 12.2 μg/m^3 to 18.8 μg/m^3 for 1-in-3 day sampling. The gray zones for the daily sampling with the continuous instrument will be tighter.

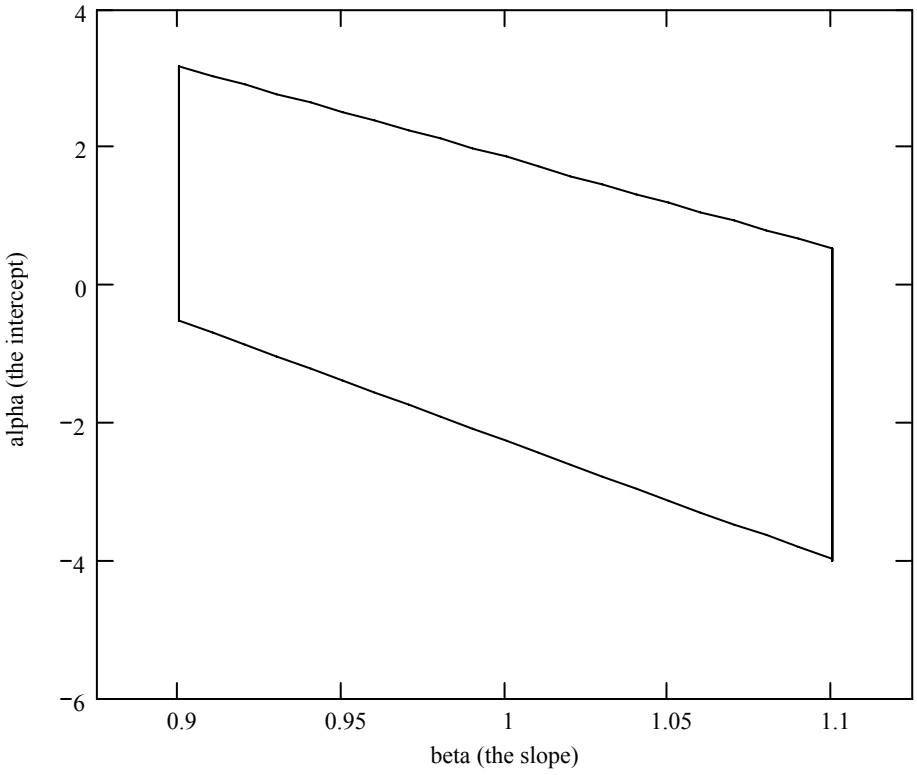


Figure A-6. Acceptance range for the additive and multiplicative bias components.

[1] The calculations to obtain the functions in Equations A-6 and A-7 were done with the full precision and rounded at the end, not at the intermediate steps shown.

Approximate Standard Errors for the Additive and Multiplicative Bias Components

Table A-1 displays approximate standard errors for the bias components. The formulas used are variations on the usual formulas for the standard errors of linear regression slope and intercept. (Casella and Berger, 1990).

Table A-1. Approximate standard errors for slope and intercept estimates.

Number of Candidate Samplers	Candidate Sampler Precision	Number of Days Sampled	Conc. Mean	Overall Population CV	Intercept ($\mu g/m^3$)		Slope	
					Approx. Standard Error	Approx. Radius of 95% Conf. Interval	Approx. Standard Error	Approx. Radius of 95% Conf. Interval
3	10.0%	23	10	30%	0.42	0.82	4.0%	7.9%
3	10.0%	23	10	50%	0.27	0.53	2.4%	4.7%
3	10.0%	23	20	30%	0.84	1.64	4.0%	7.9%
3	10.0%	23	20	50%	0.54	1.06	2.4%	4.7%
3	10.0%	46	10	30%	0.30	0.58	2.8%	5.6%
3	10.0%	46	10	50%	0.19	0.37	1.7%	3.3%
3	10.0%	46	20	30%	0.59	1.16	2.8%	5.6%
3	10.0%	46	20	50%	0.38	0.75	1.7%	3.3%
3	12.5%	23	10	30%	0.52	1.03	5.0%	9.8%
3	12.5%	23	10	50%	0.34	0.66	3.0%	5.9%
3	12.5%	23	20	30%	1.05	2.05	5.0%	9.8%
3	12.5%	23	20	50%	0.67	1.32	3.0%	5.9%
3	12.5%	46	10	30%	0.37	0.73	3.5%	7.0%
3	12.5%	46	10	50%	0.24	0.47	2.1%	4.2%
3	12.5%	46	20	30%	0.74	1.45	3.5%	7.0%
3	12.5%	46	20	50%	0.48	0.93	2.1%	4.2%
3	15.0%	23	10	30%	0.63	1.23	6.0%	11.8%
3	15.0%	23	10	50%	0.40	0.79	3.6%	7.1%
3	15.0%	23	20	30%	1.26	2.46	6.0%	11.8%
3	15.0%	23	20	50%	0.81	1.58	3.6%	7.1%
3	15.0%	46	10	30%	0.44	0.87	4.3%	8.3%
3	15.0%	46	10	50%	0.29	0.56	2.6%	5.0%
3	15.0%	46	20	30%	0.89	1.74	4.3%	8.3%
3	15.0%	46	20	50%	0.57	1.12	2.6%	5.0%

If $n_{samplers}$ is the number of candidate samplers and n_{days} is the number of days with valid data, the standard error of the slopes and intercepts given in Table A-1 are calculated by:

$$SE(b) = \sqrt{\frac{(Overall\ sample\ mean)^2 \cdot (Candidate\ \ sampler\ \ measurement\ \ CV)^2}{(Standard\ deviation\ of\ the\ daily\ concentrations)^2 \cdot n_{samplers} \cdot n_{days}}} \qquad \text{(Eq. A-8)}$$

$$SE(a) = SE(b) \cdot \sqrt{(Overall\ sample\ mean)^2 + (Standard\ deviation\ of\ the\ daily\ concentrations)^2} \qquad \text{(Eq. A-9)}$$

The usual least squares assumption that the independent variable is measured without error does not strictly hold, but should not be grossly incorrect. Also, these calculations assume that the multiplicative bias is approximately 1 and that the intercept is approximately 0 and, therefore, should be interpreted only qualitatively. However, assuming a goal of a 5 percent radius for a 95 percent confidence interval for the slope, then a combination of a 10 percent measurement precision and a site with a 50 percent concentration CV during a single month is sufficient. Otherwise, the two-season sampling (which is needed for other reasons) is needed to achieve that goal.

References

Battelle, (2003). "DQO Companion for PMcoarse Software Tool." Manual and software program for U.S. EPA, Contract No. 68-D-02-061, May 2.

Casella, G., and Berger, R.L. (1990). Statistical Inference. Belmont, California: Duxbury Press.

Coutant, B.W., and Holloman, C.H. (2003). "Statistics for Criteria Pollutant Data Quality Objectives." Battelle technical report to U.S. EPA, Contract No. 68-D-02-061, September 15.

Mosquin, P., and McElroy, F. (2004). "Data Quality Objectives for PM Continuous Methods II." ManTech technical report to the U.S. EPA, report number TR-CAN-04-02, June.

APPENDIX B:

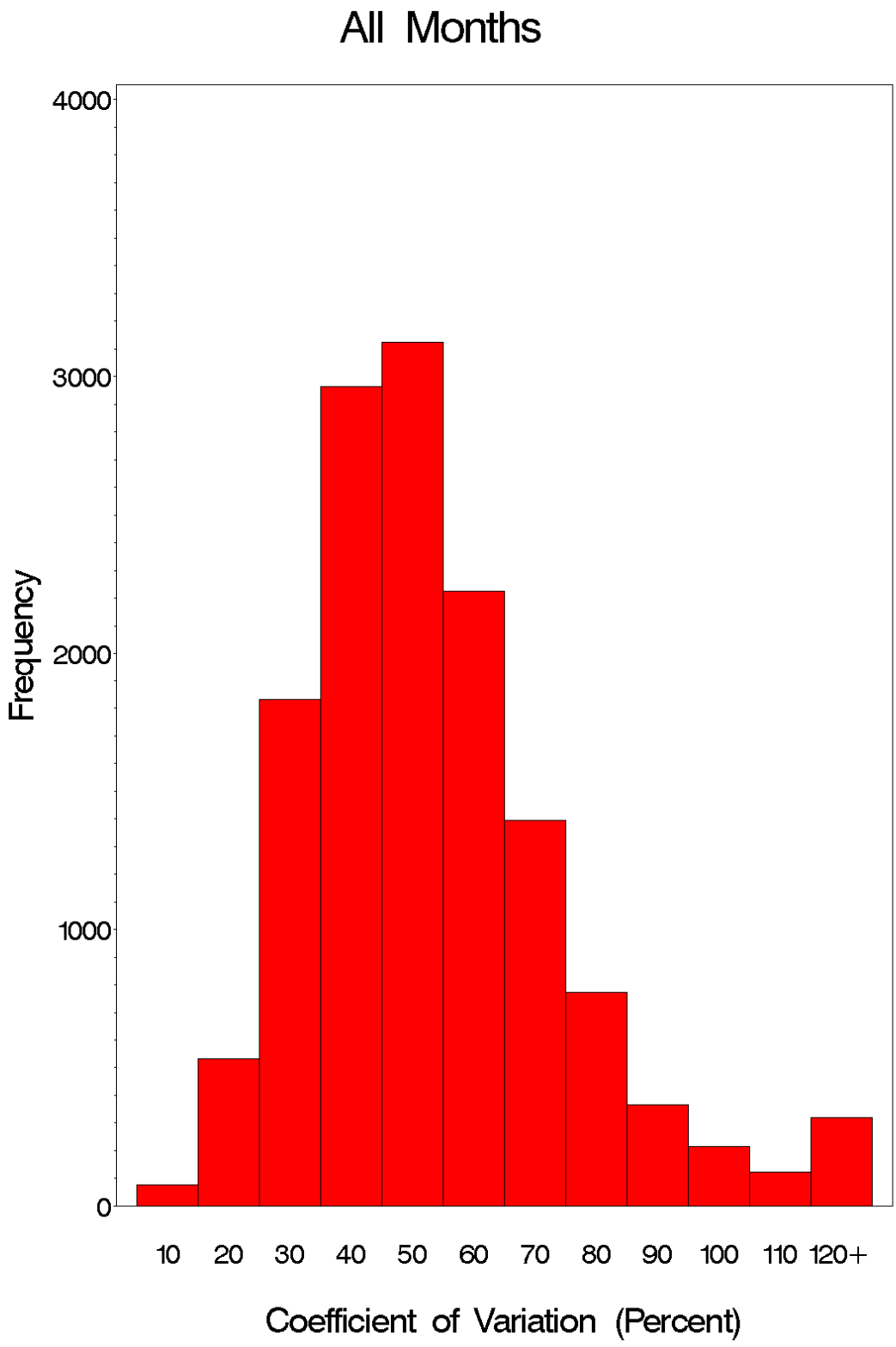
THE DISTRIBUTION OF MONTHLY CONCENTRATION CVS

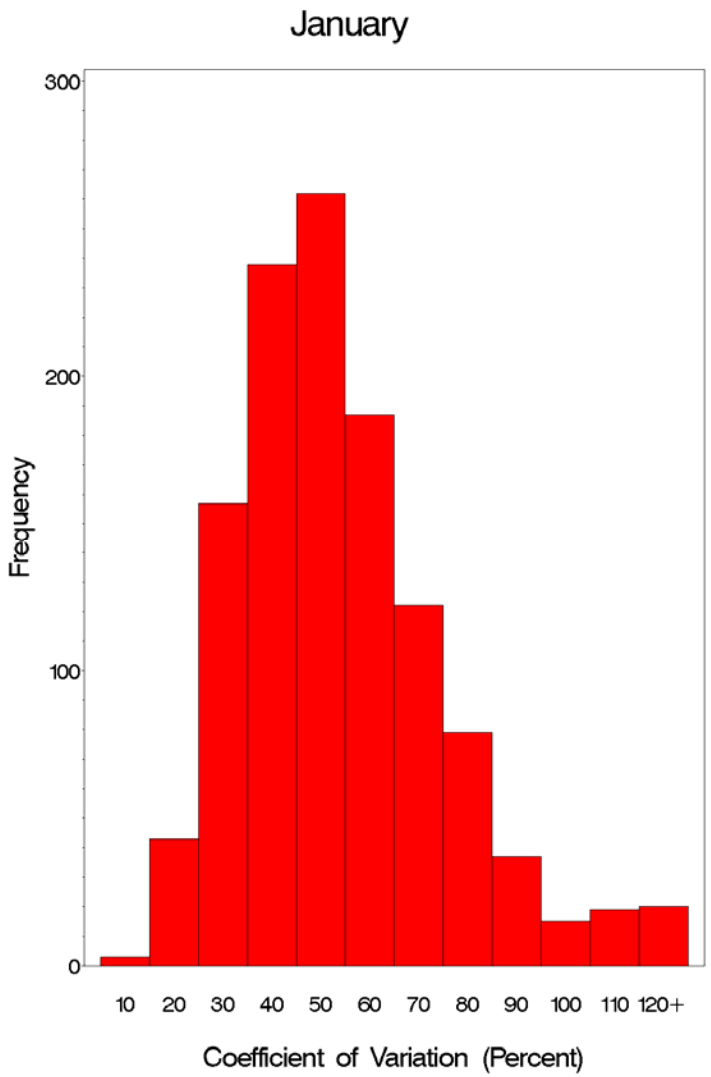
Appendix B: The Distribution of Monthly Concentration CVs

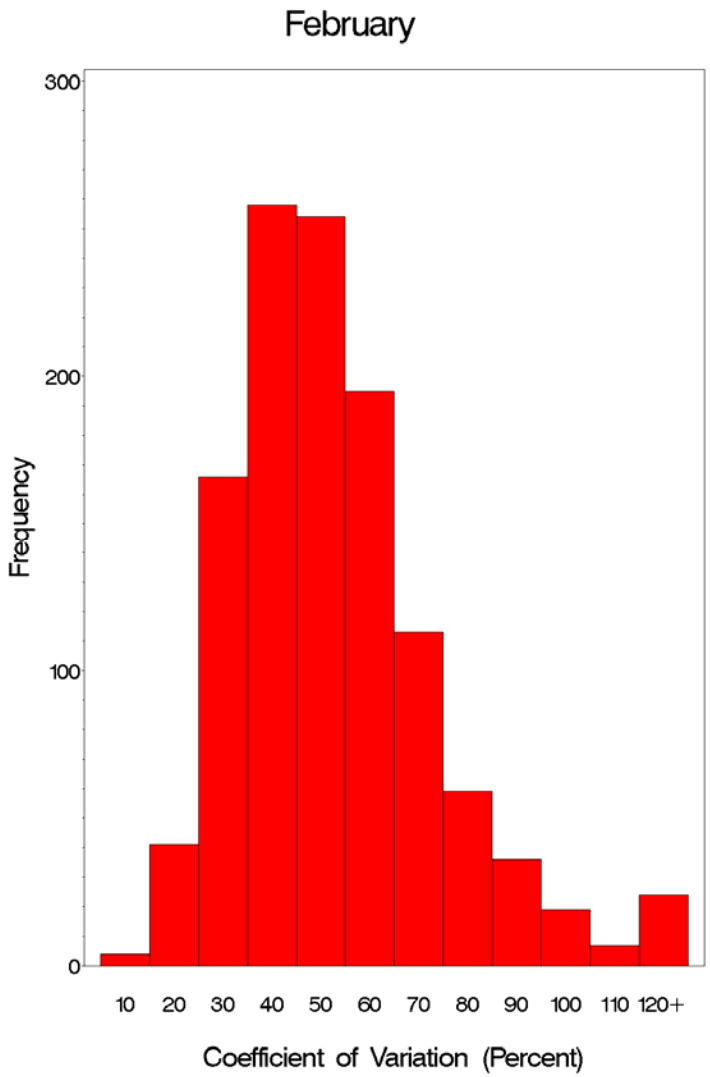
In the course of developing the equivalency criteria, there was concern that due to the known seasonality effect on $PM_{2.5}$ concentrations, there would be a seasonal component to the variability of the concentrations. There was also concern about the minimum concentration CV that could be expected. While the possibilities were not extensively investigated, the plots in this appendix and previous work in verifying the assumptions made in the DQO development should help alleviate some of these concerns.

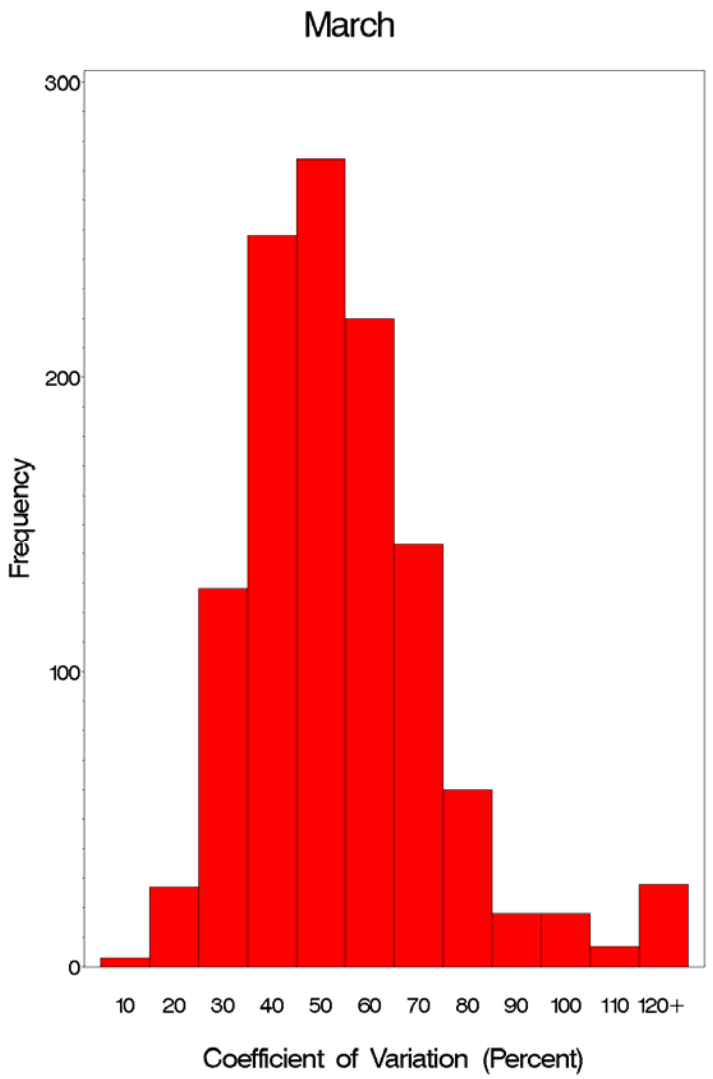
It should be noted that the DQO development for $PM_{2.5}$ assumes that there is a very specific seasonal component to the population, namely that for fixed sites, the coefficient of variation about the seasonal curve is constant. Over periods relative to the annual cycle (e.g., one month), this should translate into a constant CV. To simultaneously investigate both issues, the monthly CVs were calculated from the collected 2003 data that was downloaded from AQS the first week of July 2004. The first graph that follows shows the distribution of all of these monthly CVs. From this illustration, one can see that 30 percent is a reasonable lower bound. The remaining graphs show the distribution of the CVs restricted to each calendar month. These show that the overall distribution did not change much from month to month during the year.

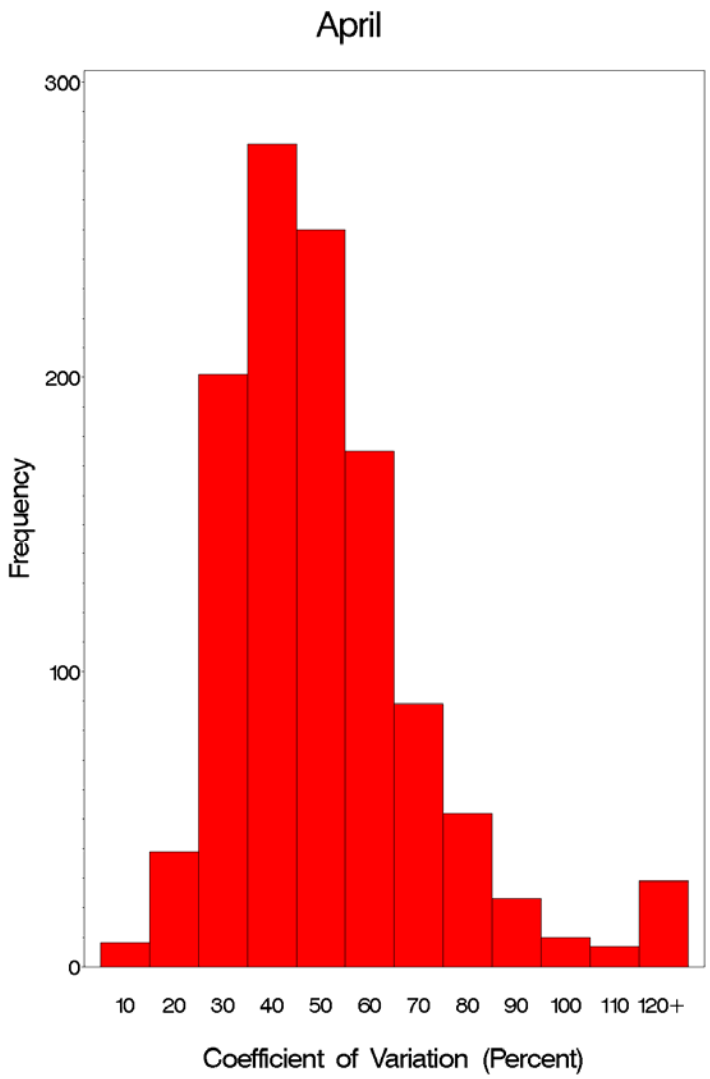
Note that because many of these sites only sample once every six days, the bin classifications for any given site and month are not certain. These data should only be reviewed in aggregate, and not at site level. Site selection should be based on several years' worth of data, preferably restricted to sites with higher sampling frequencies.

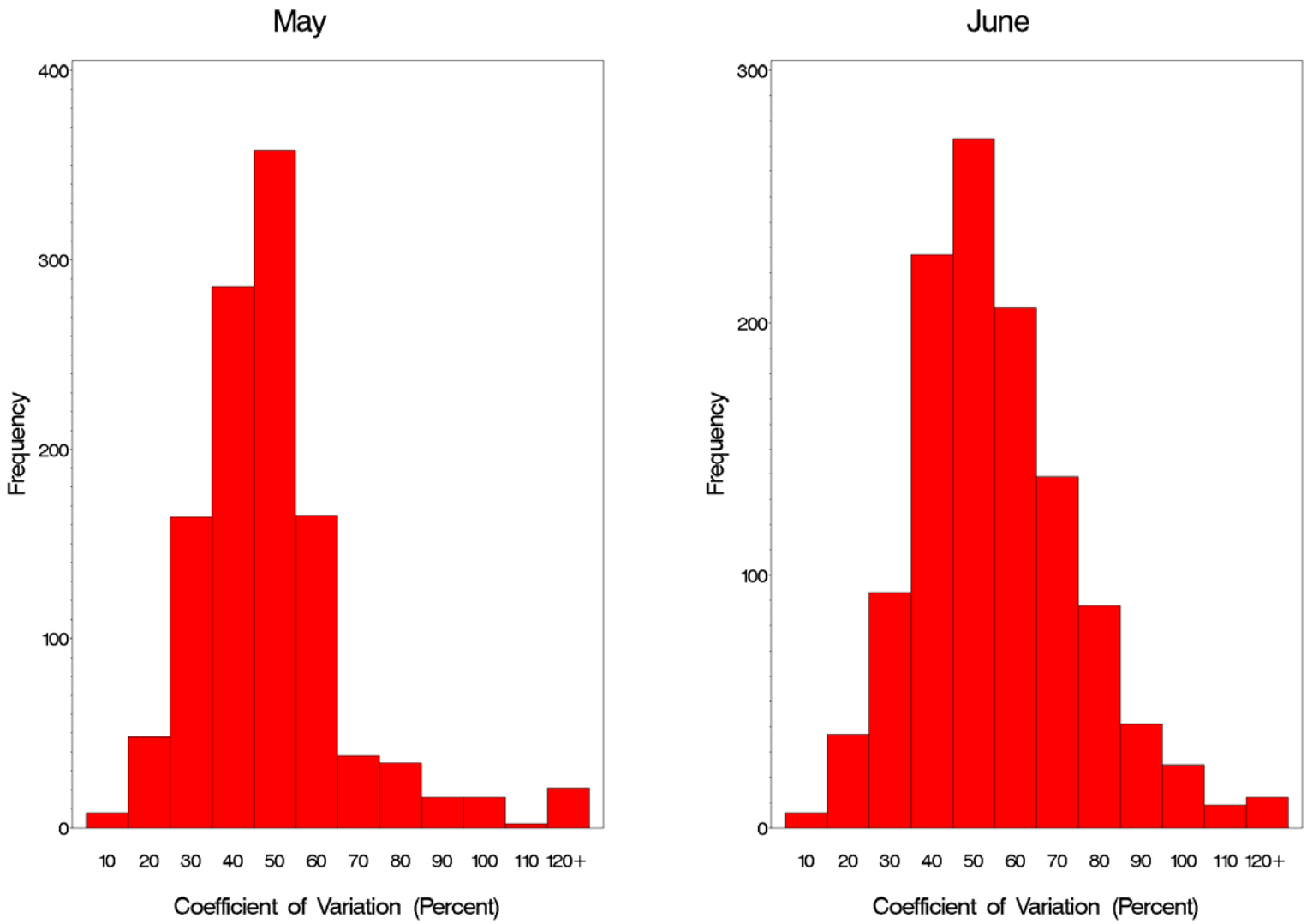

All Months
4000
3000
2000
1000
0
Frequency
10
20
30
40
50
60
70
80
90
100
110
120+
Coefficient of Variation (Percent)

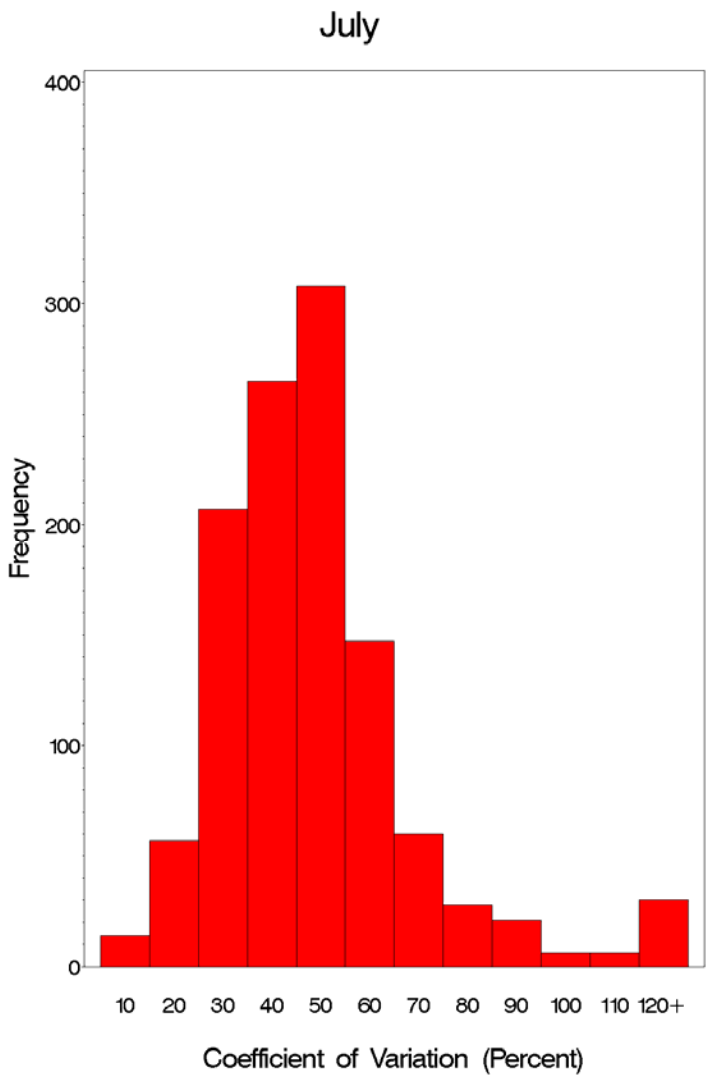

January
300
200
100
0
Frequency
10 20 30 40 50 60 70 80 90 100 110 120+
Coefficient of Variation (Percent)

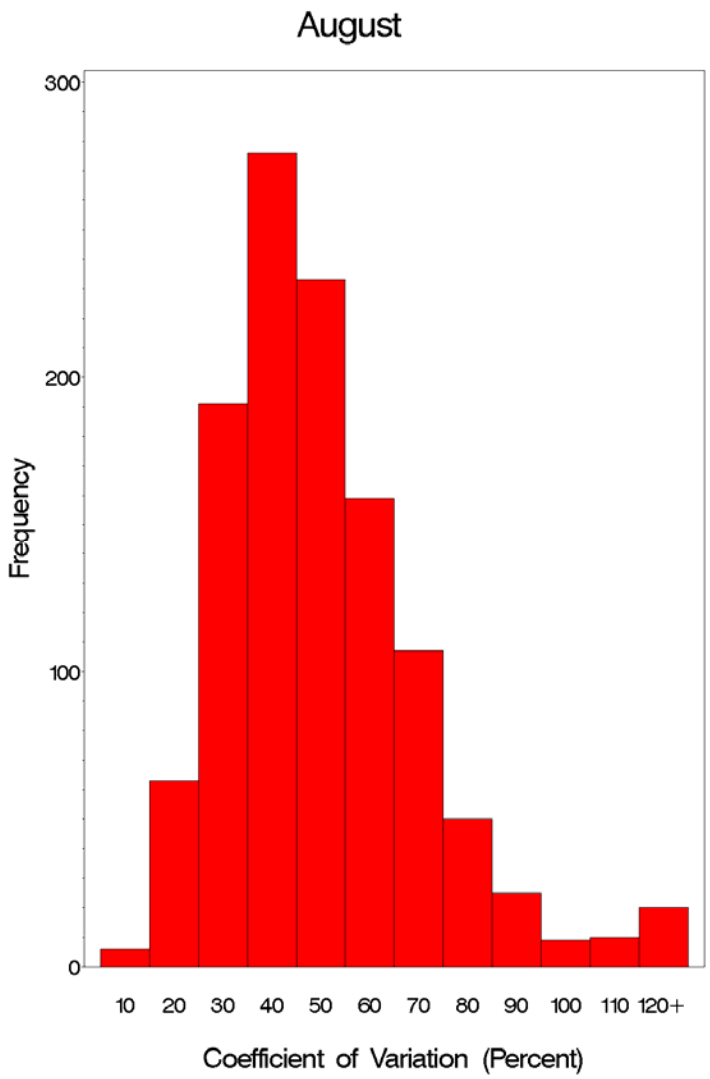

February
300
200
100
0
Frequency
10 20 30 40 50 60 70 80 90 100 110 120+
Coefficient of Variation (Percent)

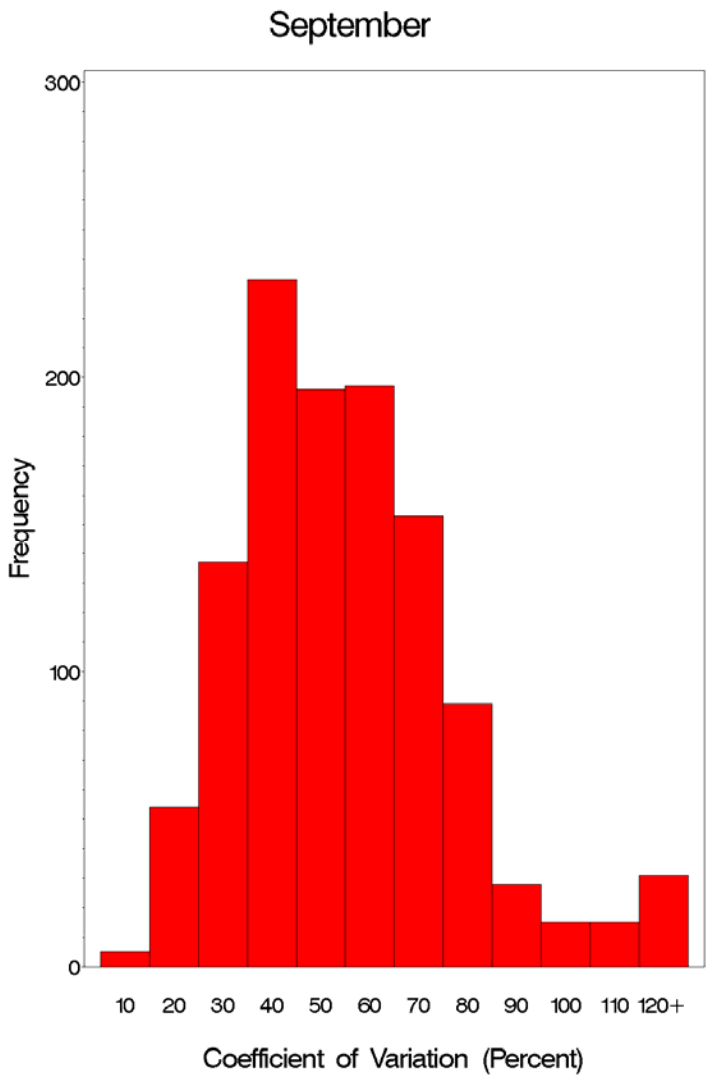

March
300
200
100
0
Frequency
10 20 30 40 50 60 70 80 90 100 110 120+
Coefficient of Variation (Percent)

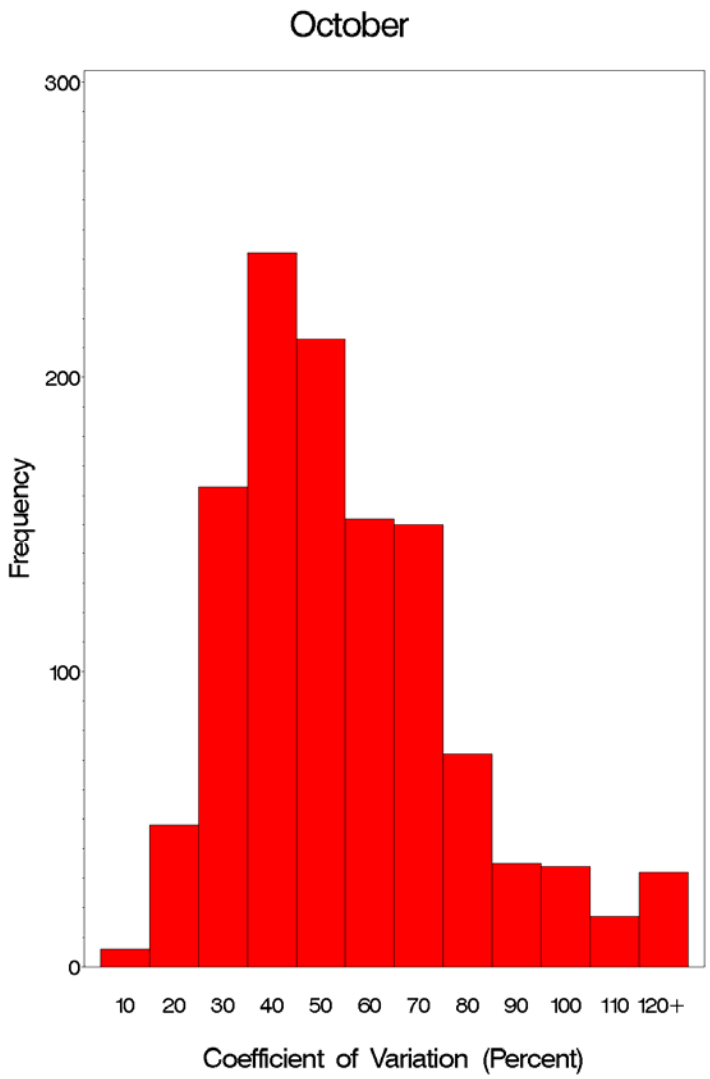

April
300
200
100
0
Frequency
10 20 30 40 50 60 70 80 90 100 110 120+
Coefficient of Variation (Percent)

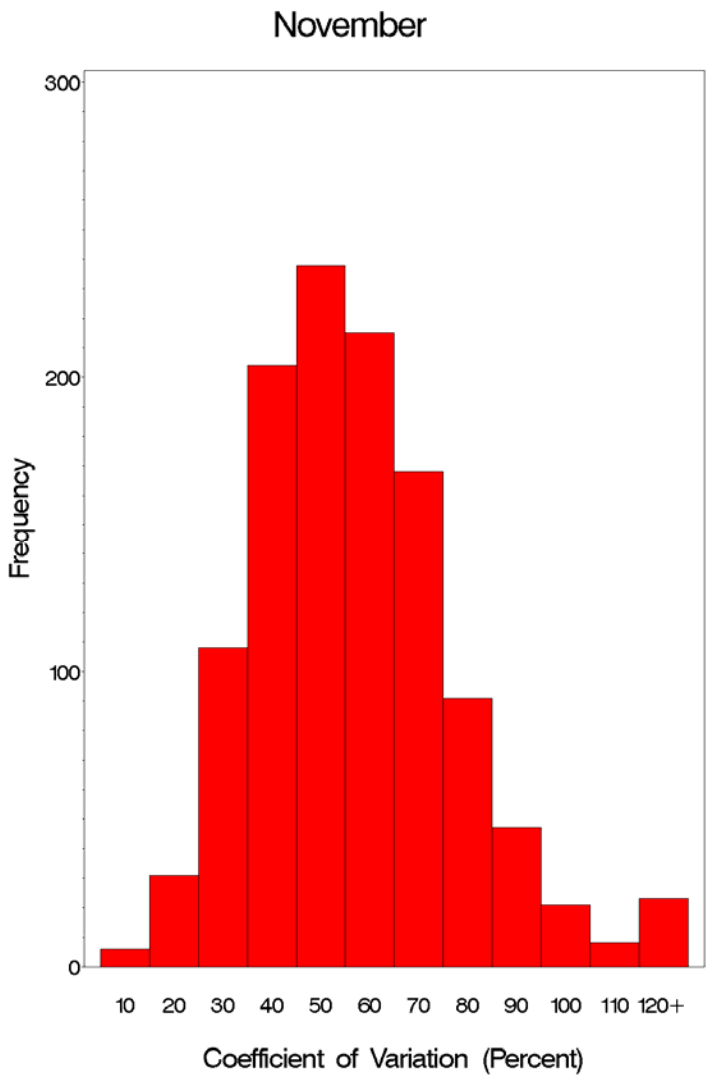

May
Frequency
400
300
200
100
0
10 20 30 40 50 60 70 80 90 100 110 120+
Coefficient of Variation (Percent)
June
Frequency
300
200
100
0
10 20 30 40 50 60 70 80 90 100 110 120+
Coefficient of Variation (Percent)

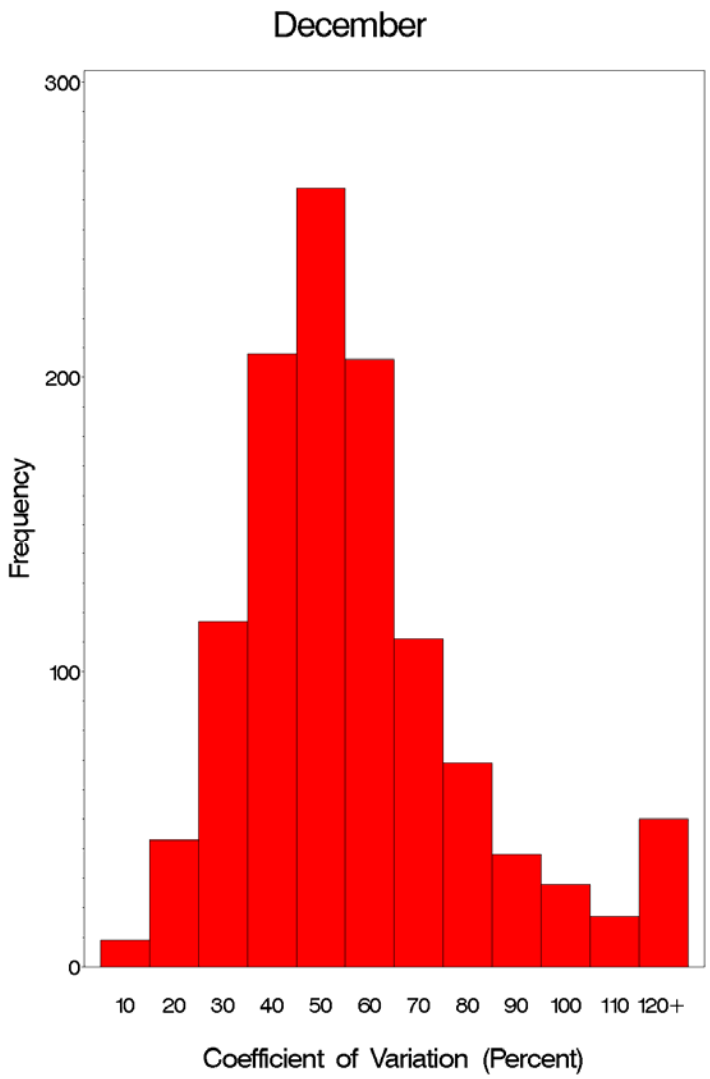

July
400
300
200
100
0
Frequency
10 20 30 40 50 60 70 80 90 100 110 120+
Coefficient of Variation (Percent)


August
300
200
100
0
Frequency
10 20 30 40 50 60 70 80 90 100 110 120+
Coefficient of Variation (Percent)


September
300
200
100
0
Frequency
10 20 30 40 50 60 70 80 90 100 110 120+
Coefficient of Variation (Percent)


October
300
200
100
0
Frequency
10 20 30 40 50 60 70 80 90 100 110 120+
Coefficient of Variation (Percent)


November
300
200
100
0
Frequency
10 20 30 40 50 60 70 80 90 100 110 120+
Coefficient of Variation (Percent)


December
300
200
100
0
Frequency
10 20 30 40 50 60 70 80 90 100 110 120+
Coefficient of Variation (Percent)