Equivalency of a personal dust monitor to the current United States coal mine respirable dust sampler

Steven J. Page, Jon C. Volkwein, Robert P. Vinson, Gerald J. Joy, Steven E. Mischler, Donald P. Tuchman, Linda J. McWilliams

US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Pittsburgh Research Laboratory, 626 Cochrans Mill Road, Pittsburgh, PA 15236, USA

Appendix A

Mathematical representation

Comparison of the PDM to the personal sampler was made using regression analysis. The data indicate an error term increasing with the independent variable, or multiplicative error, in addition to the required constant additive error term. As a result, the total sum of squares will largely be influenced by the large dependent variable values and lead to an analysis bias. This situation is typical of data collected with dust sampling instrumentation[1] and there are several different remedial data transformations to eliminate, or at least minimize, the non-constant variance problem.

The general equation used by Eagleson and Muller[2] to represent only multiplicative errors can be written as:

Eq. A-1 $$Y = g(X)*(\varepsilon_1).$$

In the present analysis,

Y = a personal sampler response variable,
$g(X)$ = some function of the PDM predictor variable (X),
ε_1 = a normally distributed random multiplicative error term, mean = 1 and variance resulting from in-canister spatial variation in the concentration, coupled with sampling and analytical error of the personal sampler.

The only requirement is that g(x) be a smooth function. Eq. A-1 can also be expressed in terms of the usual error term ε_0 with mean = 0, with inclusion of an additive error term, as

Eq. A-2 $$Y = g(X)*(1+\varepsilon_0) + {}_0\varepsilon_0 = g(X) + g(X)*(\varepsilon_0) + {}_0\varepsilon_0,$$

where

${}_0\varepsilon_0$ = a normally distributed random additive error term with mean = 0 and constant variance $({}_0\sigma)^2$, resulting from weighing imprecision as the true concentration approaches zero.

A decision for g(X) representing the true underlying model for the data must be made. The model should agree with similar published data, previous experience, and be based on sound statistical arguments. Intuitively, one would expect, in the absence of measurement bias, a linear and monotonic relationship (and ideally with zero intercept, unity slope) between different instruments designed and developed to measure the same true but unknown quantity. In this case, that quantity is the airborne respirable coal mine dust concentration.

Weight variable estimation

Weighted regression can directly stabilize the variance if the variance function can be estimated. There are numerous weighting factors that can be used in regression analysis, the more common of which are $(1/X)$ and $(1/X^2)$.[3,4] The data of this investigation were used to internally estimate the variance relationship of the personal sampler with the independent PDM variable. Typical $1/X^2$ weighting assumes that dependent variable variance increases proportionally with X^2 over the entire range of independent variable. However, at low concentration values there is the limiting error term $({}_0\varepsilon_0)$ due to weighing imprecision. The constant variance $({}_0\sigma)^2$ of this error term is known quite accurately for the personal sampler samples and is presented in the Results section. It is readily seen that the proper weight variable is the reciprocal of the true total variance σ_T^2, given by

Eq. A-3 $$\sigma_T^2 \approx ({}_0\sigma)^2 + (RSD)^2 * X^2,$$

where RSD can be considered to be the variation of the dependent variable about the regression.

The process for estimating the proper weight variable is iterative, using the following procedure for the personal sampler data:

Step 1: An initial regression of Eq. A-2 using $1/X^2$ weighting is performed to establish initial weight variables, where $g(X) = Y_0 + a*X$.
Step 2: Using the definition of variance, the values $(Y_i - Y_{ip})^2$, representing the variance between the measured Y_i and predicted Y_{ip} from the initial regression of step 1, are calculated.
Step 3: The plot of $(Y_i - Y_{ip})^2$ vs X_i is fit with the function of Eq. A-3. The second weight estimation is then approximated point-by-point as $1/\sigma_T^2$.
Step 4: Perform a weighted regression with the new weight variable.
Steps 2-4 are then repeated with each new estimate of weight variable and $(Y_i - Y_{ip})^2$ until convergence to a solution.

Appendix B

Table B-1 Valid area sample raw data without PDM bias corrections

		mg m^{-3}				
MSHA District	**Field Office**	**PDM**	**Void**	**CMDPSU**	**Void**	**Notes**
6	Whitesburg, KY	0.041	-	0.047	-	
4	Pineville, WV	0.050	-	0.055	-	
7	Barbourville, KY	0.050	-	0.048	-	
6	Whitesburg, KY	0.073	-	0.063	-	
6	Elkhorn City, KY	0.076	-	0.097	-	
5	Norton, VA	0.080	-	0.078	-	
6	Elkhorn City, KY	0.080	-	0.114	-	
6	Pikeville, KY	0.080	-	0.088	-	
4	Logan, WV	0.095	-	0.088	-	
5	Norton, VA	0.100	-	0.124	-	
5	Norton, VA	0.100	-	0.065	-	
2	Indiana, PA	0.115	-	0.109	-	
6	Martin, KY	0.119	-	0.098	-	
2	Kittanning, PA	0.126	-	0.104	-	
9	Delta, CO	0.129	-	0.129	-	
5	Norton, VA	0.130	-	-	0.004	(a)
2	Ruff Creek, PA	0.134	-	0.132	-	
5	Norton, VA	0.140	-	0.112	-	
2	Kittanning, PA	0.143	-	0.113	-	
9	Craig, CO	0.155	-	0.220	-	
9	Castle Dale, UT	0.158	-	0.205	-	
4	Princeton, W VA	0.180	-	0.160	-	
11	Hueytown, AL	0.186	-	0.208	-	
5	Norton, VA	0.190	-	0.149	-	
4	Logan, WV	0.204	-	0.224	-	
2	Ruff Creek, PA	0.213	-	0.257	-	
8	Benton, IL	0.220	-	0.282	-	
7	Jacksboro, TN	0.222	-	0.195	-	
2	Kittanning, PA	0.240	-	0.180	-	
8	Hillsboro, IL	0.240	-	0.276	-	
9	Castle Dale, UT	0.248	-	0.301	-	
11	Hueytown, AL	0.253	-	0.331	-	
2	Ruff Creek, PA	0.254	-	0.220	-	
10	Beaver Dam, KY	0.254	-	0.308	-	
9	Craig, CO	0.265	-	0.246	-	
4	Logan, WV	0.265	-	0.262	-	
4	Madison, WV	0.272	-	0.244	-	
4	Pineville, WV	0.280	-	0.233	-	
6	Whitesburg, KY	0.284	-	0.295	-	
2	Johnstown, PA	0.292	-	0.246	-	
3	Morgantown, WV	0.323	-	0.387	-	
8	Benton, IL	0.330	-	0.441	-	
11	Hueytown, AL	0.348	-	0.408	-	
3	Morgantown, WV	0.355	-	0.303	-	
5	Norton, VA	0.360	-	0.415	-	
6	Pikeville, KY	0.360	-	0.352	-	
8	Hillsboro, IL	0.360	-	0.346	-	
8	Vincennes, IN	0.360	-	0.390	-	
4	Madison, WV	0.369	-	0.404	-	
4	Logan, WV	0.376	-	0.368	-	
4	Madison, WV	0.379	-	0.369	-	

7	Hindman, KY	0.380	-	0.354	-	
3	St. Clairsville, OH	0.383	-	0.429	-	
7	Harlan, KY	0.440	-	0.392	-	
10	Madisonville, KY	0.449	-	0.546	-	
6	Martin, KY	0.451	-	-	0.395	(b)
7	Hazard, KY	0.452	-	0.459	-	
6	Martin, KY	0.455	-	0.482	-	
4	Madison, WV	0.475	-	0.558	-	
5	Vansant, VA	0.480	-	0.612	-	
7	Hindman, KY	0.485	-	0.543	-	
6	Phelps, KY	0.490	-	0.520	-	
5	Vansant, VA	0.500	-	0.503	-	
8	Benton, IL	0.500	-	0.587	-	
4	Logan, WV	0.513	-	0.434	-	
3	Morgantown, WV	0.540	-	0.599	-	
9	Delta, CO	0.563	-	0.614	-	(c)
4	Madison, WV	0.570	-	0.595	-	
2	Ruff Creek, PA	0.579	-	0.588	-	
7	Barbourville, KY	0.588	-	0.644	-	
10	Beaver Dam, KY	0.596	-	0.565	-	
6	Pikeville, KY	0.610	-	0.601	-	
5	Norton, VA	0.620	-	0.612	-	
2	Kittanning, PA	0.628	-	0.600	-	
8	Hillsboro, IL	0.630	-	0.819	-	
4	Pineville, WV	0.640	-	0.689	-	
2	Johnstown, PA	0.644	-	0.533	-	
6	Phelps, KY	0.660	-	0.650	-	
11	Hueytown, AL	0.686	-	0.875	-	
3	Bridgeport, WV	0.689	-	0.888	-	
9	Delta, CO	0.741	-	0.965	-	
5	Norton, VA	0.760	-	0.717	-	
6	Phelps, KY	0.760	-	0.703	-	
8	Vincennes, IN	0.820	-	0.841	-	
9	Craig, CO	0.842	-	0.805	-	
10	Beaver Dam, KY	0.852	-	1.046	-	
9	Price, UT	0.888	-	1.156	-	
4	Mt. Carbon, WV	0.890	-	1.070	-	
6	Whitesburg, KY	0.914	-	0.822	-	
3	Morgantown, WV	0.921	-	0.872	-	
4	Mt. Hope, WV	0.960	-	1.076	-	
9	Price, UT	0.979	-	1.059	-	
6	Pikeville, KY	1.020	-	1.035	-	
7	Barbourville, KY	1.041	-	1.203	-	
7	Jacksboro, TN	1.058	-	1.100	-	
7	Harlan, KY	1.070	-	1.280	-	
7	Jacksboro, TN	1.103	-	1.255	-	
3	Bridgeport, WV	1.103	-	1.271	-	
10	Madisonville, KY	1.171	-	1.528	-	
10	Madisonville, KY	1.244	-	1.435	-	
4	Logan, WV	1.285	-	1.473	-	
4	Madison, WV	1.297	-	1.607	-	
7	Harlan, KY	1.330	-	1.491	-	
6	Whitesburg, KY	1.362	-	1.301	-	
6	Martin, KY	1.401	-	1.428	-	
3	Morgantown, WV	1.419	-	1.420	-	
4	Madison, WV	1.482	-	1.680	-	
5	Norton, VA	1.520	-	1.291	-	

6	Phelps, KY	1.520	-	1.445	-	
8	Vincennes, IN	1.520	-	1.930	-	
4	Logan, WV	1.522	-	1.705	-	
3	Oakland, MD	1.529	-	1.956	-	
5	Vansant, VA	1.530	-	1.746	-	
6	Elkhorn City, KY	1.570	-	1.681	-	
6	Pikeville, KY	1.590	-	1.455	-	
4	Mt. Hope, WV	1.610	-	1.543	-	
7	Harlan, KY	1.620	-	1.586	-	
5	Norton, VA	1.630	-	1.680	-	
8	Vincennes, IN	1.650	-	-	0.481	(d)
7	Jacksboro, TN	1.669	-	1.700	-	
7	Hazard, KY	1.670	-	2.074	-	(c)
7	Harlan, KY	1.680	-	1.496	-	
7	Barbourville, KY	1.720	-	1.697	-	
4	Mt. Hope, WV	1.740	-	1.993	-	
7	Barbourville, KY	1.742	-	1.616	-	
7	Harlan, KY	1.840	-	2.012	-	
7	Hindman, KY	1.934	-	2.702	-	
7	Hindman, KY	1.972	-	-	1.515	(b)
6	Martin, KY	2.042	-	2.250	-	
4	Mt. Carbon, WV	2.060	-	2.020	-	
6	Pikeville, KY	2.100	-	2.666	-	
4	Pineville, WV	2.320	-	2.715	-	
4	Logan, WV	2.415	-	2.550	-	
4	Pineville, WV	-	0.610	1.370	-	(e)
4	Pineville, WV	-	0.120	0.209	-	(e)

(a) Cyclone hose off when opened can
(b) Cyclone pump out of calibration
(c) light rockdusting
(d) possible pre-weigh error on 2.0 L min^{-1} filter--outlying data point
(e) PDM flow restriction

Table B-2 Excluded area sample raw data without PDM bias corrections

		mg m^{-3}		
MSHA District	**Field Office**	**PDM**	**CMDPSU**	**Notes**
2	Ruff Creek, PA	-	-	(a)
3	Bridgeport, WV	16.465	14.361	(b)
3	Bridgeport, WV	0.885	0.778	(c)
3	Bridgeport, WV	1.842	1.911	(a)
4	Mt. Carbon, WV	-	-	(d)
4	Mt. Hope, WV	3.930	4.714	(e)
5	Norton, VA	1.340	1.406	(a)
6	Martin, KY	0.174	0.228	(f)
6	Whitesburg, KY	0.520	0.411	(g)

(a) PDM failed
(b) PDM filter overload error after 4 hr 42 min
(c) PDM TE fail/remove error after 7 hr 10 min
(d) PDM did not start.
(e) PDM greater than twice the protocol limit.
(f) Sample terminated early, mine shut down by inspector.
(g) heavy rockdusting

References

[1] NIOSH, *The Precision of Coal Mine Dust Sampling*. US Department of Health and Human Services, Public Health Services, Centers for Disease Control, National Institute for Occupational Safety and Health, NTIS Pub. No. PB-85-220721, 1984.

2. G. K. Eagleson and H. G. Muller, *J. Royal Stat. Soc.* 1997, **59**, 173-189.

3. J. Neter, W. Wasserman and M. H. Kutner, in *Applied Linear Statistical Models,* Richard D. Irwin, Inc. 3rd edn., 1990, ch.11, p. 420.

4. SPSS, Inc. SPSS v.14.0, Chicago, IL.