Proposal to Change $PM_{2.5}$ and PM_{10} Collocation Sampling Frequency Requirements

Mike Papp and Louise Camalier

November, 2005

Proposal to Change $PM_{2.5}$ and PM_{10} Collocation Sampling Frequency Requirements

In order to provide both adequate and cost effective data quality information, the OAQPS Ambient Air Quality Assurance Team reviewed the $PM_{2.5}$ and PM_{10} precision data to ensure that this information was providing representative and adequate precision estimates for reporting organizations. Unlike the gaseous criteria pollutants, where one can use a standard of known concentration to estimate precision and perform this at every site; for the particulate matter pollutants, one must rely on duplicate measurements at a representative sample of sites for estimates of precision.

Precision is estimated using collocated sampling. Since only a portion of the monitoring sites are represented, precision is assessed at the reporting organization level. In order to provide an adequate level of confidence in our estimates of precision, an adequate number of collocation samples must be collected. The pertinent regulation for collocated precision is found in 40 CFR Part 58, Appendix A. The strategy is to collocate an identical FRM or FEM $PM_{2.5}$ or PM_{10} air sampling instrument with an established primary sampler at a routine air monitoring site, operate both samplers over the same time period, and then compare the results. The current regulation for the number of sites that need to be collocated are different for $PM_{2.5}$ versus PM_{10}, however they are both based on the percentage of routine monitoring sites in the reporting organizations' network. However, the frequency at which the collocated sites are sampled is the same (once every 6 days).

Recently, OAQPS was asked to look at whether another QA activity, the $PM_{2.5}$ Performance Evaluation Program (PEP), could be reduced without adversely affecting the confidence in our bias estimates. As we started to develop techniques to review this data, we realized that similar techniques could be used to assess the collocated precision data. Since three years of routine data are used for comparisons to the National Ambient Quality Standards (NAAQS) comparisons, EPA uses 3 years of precision data to determine if the data quality objectives (DQOs) are being achieved. Our assessments suggested that we could reduce the number of sites that required collocation to 15% within each reporting organization resulting in a reduction in sampling frequency from every six days to every twelve days without significantly affecting the precision estimates.

We set about reviewing the data with the following assumptions.

1) The precision data is aggregated at the reporting organization level over a 3 year period.
2) We are using the new proposed precision statistic which is the 90% confidence limit upper bound for the mean coefficient of variation.
3) Even with a reporting organization with only 1 collocated site, at a 1-in-12 day sampling frequency and 75% completeness (CFR requirement), we would expect to receive **66** collocated values (22/year) over 3 years. Our evaluations show that this provides estimates of precision with acceptable levels of confidence.

$PM_{2.5}$ and PM_{10} data was assessed for the calendars years 2002-2004. However, since the $PM_{2.5}$ precision, in general, is of higher quality than the $PM_{10,}$ the examples and findings that follow are related to PM_{10} data yet are also applicable to the $PM_{2.5}$ data.

Collocated Sample Frequency Reduction to 1 in 12 Day Sampling

Table 1 reports the 3-year precision estimates (CV-upper bounds) using 1-in-6 day (columns with "_6") and 1-in-12 day data (columns with "_12"). Figure 1 illustrates this information. Since we would expect 66 samples over 3 years as a minimum, we separated the CV_UB estimates where we had at least 66 1-in-12 day samples (right-most portion of graph, red/dark blue lines) from those where we had < 66 sample pairs (samples starting on the left, yellow/light blue lines). The data demonstrate that the CV_UB estimates at the two sampling frequencies are not significantly different. We also looked at the absolute value difference between the two CV_UB estimates and graphed that on the chart (bottom line). As illustrated, there is less variability in the CV estimates when one has > 66 pairs. Table one shows that the mean difference for sites with less than 66 pairs is 1.71% while the difference with >66 samples was 0.21%. However there are two things that are also important to notice:

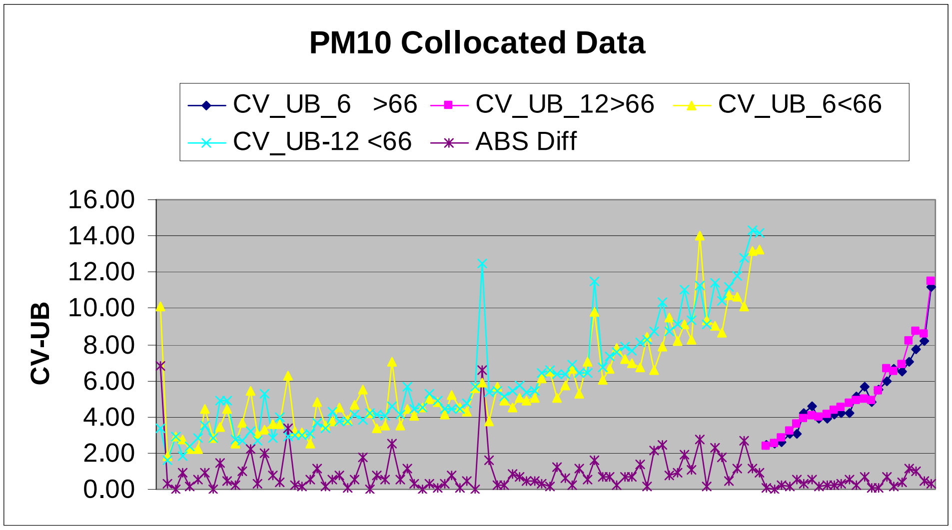

Figure 1. 3-Year PM10 data CV-UB estimates at 1-in-6 & 1-in-12 sampling frequency

1. The collocated precision estimates using the CV_UB are good. Out of 104 reporting organizations, there are only a few reporting organization precision estimates greater than 10%, which is quite assuring knowing that there are various types of high-Vol and low-Vol instruments in the field with varying degrees of precision.
2. In general, the differences between precision estimates using 1-in-6 day and 1-in-12 data are reasonable, and are minimized with more data. There are only a few sites where the difference between the precision estimates using 1-in-6 day and 1-in-12 day frequencies is > 2%; such sites had sites with < 20 collocated pairs over a 3-year period.

As we looked at the graph and tables, we wondered why there were so many sites that had less than 66 sample pairs in a 3-year period. We realized this was due to incomplete data submittal, however it also had to do with the 20 ug/m^3 cut off value that precludes any concentration < 20 ug/m^3 for use in precision estimates (this is a CFR requirement). Similar to our 2001 evaluation[1] of the $PM_{2.5}$ 6 ug/m^3 cut off, which led us to adjust this value down to 3 ug/m^3, we performed a similar analysis of the PM_{10} data. Table 2 provides the 2002-2004 PM_{10} precision estimates at cutoff values of 20 ug/m^3 (current requirement) and 15 ug/m^3 for reporting organizations that had at least 20 sample pairs.

[1] 3-Year Quality Assurance Report for Calendar Years 1999, 200 and 2001. March,2004 http://www.epa.gov/ttn/amtic/cy9901qa.html

Lowering our cut off value from 20 to 15 will not inflate the CV_ub values that would generally be calculated for a reporting organization. Two things happen when we lower our cut off value:

1. **Variability increases slightly.** Based on the statistics used, our estimates of variability increase as concentration lowers. When we lower our cut-off value to 15ug/m_3, the CV estimate that is calculated will in general be slightly higher than the CV estimate containing only values at concentrations at or above 20ug/m3. The regression line (red line) in Figure 2 illustrates the magnitude of the variability increase from 20ug/m_3 to 15ug/m_3 .

2. **Sample size increases because less data is excluded.** When we lower our cut off value to 15ug/m3, the upper bound confidence limit estimate that is calculated will be, in general, slightly lower than the upper-bound estimate containing only values at concentrations at or above 20ug/m^3. Having more data increases our sample size, which causes our upper bound confidence limit to decrease.

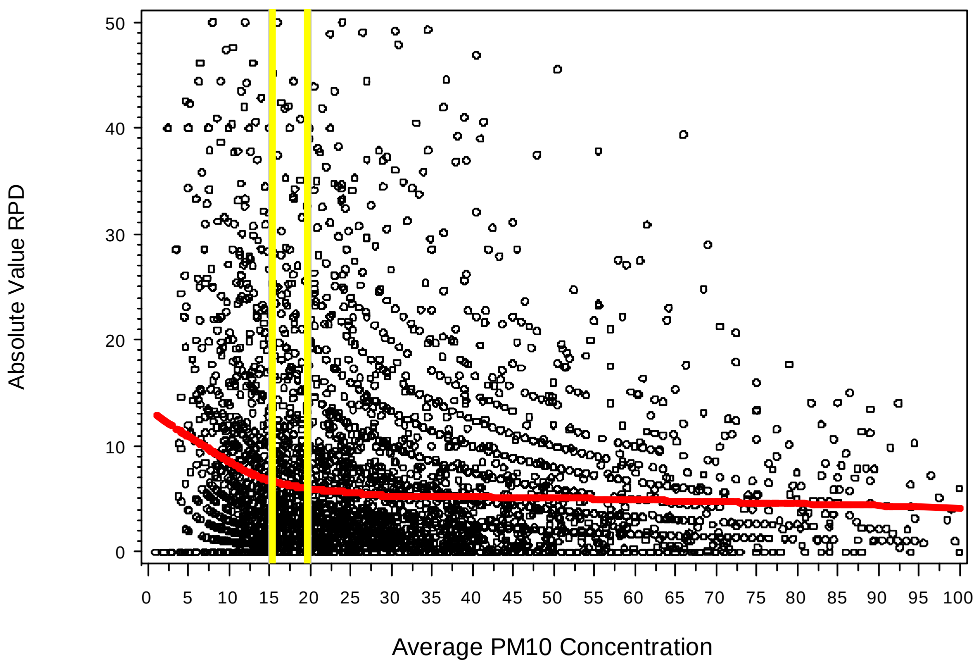

Figure2. 3-Year PM10 data RPD of collocated sample pairs.

CV_UB is calculated by multiplying CV by its corresponding upper bound value. When both events 1 and 2 occur simultaneously, they cancel each other out causing any major effect from the lowered cut-off value to be minimized. Figure 3 demonstrates this. While the CV (CV_15) for a reporting organization with a cut-off value of 15ug/m^3 may be higher (in some cases) the multiplier (ub_15) used to calculate the upper bound is lower than the corresponding multiplier (ub_20) for the CV with the cut off value of 20 ug/m^3. Therefore, the CV_ub that results from the 15ug/m^3 cut off value will not be too different from its original CV_ub value calculated using the original 20 ug/m^3 restriction.

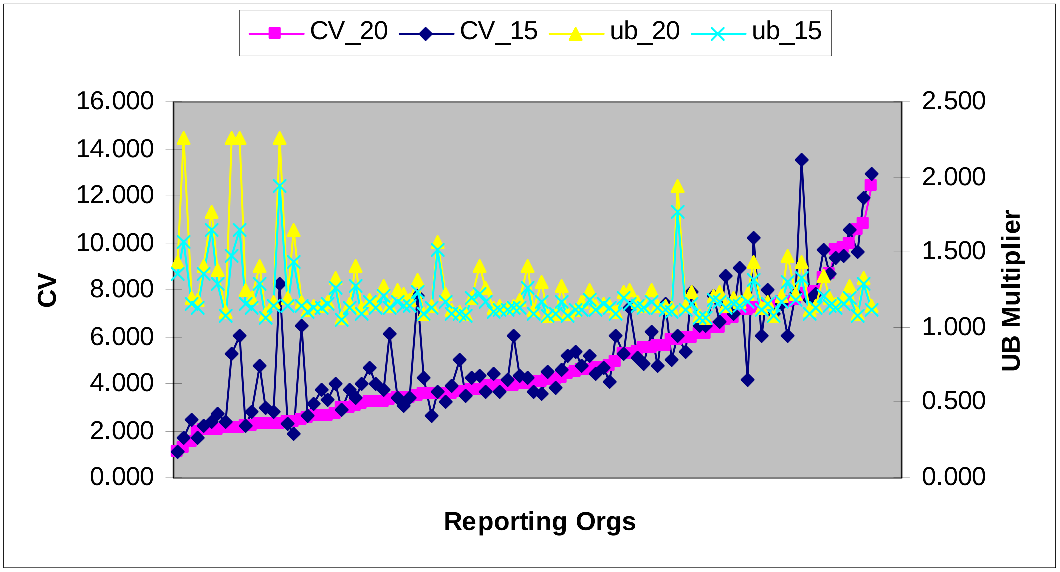

Figure 3. The "canceling-out" effect of gathering more data while slightly increasing variability

The data in Fig. 2 represent the absolute value of the relative percent differences (RPD) of each collocated pair in the 3-year PM_{10} data base. The x-axis represents the mean concentration estimate of the collocated pair. A regression line (in red) was calculated using this data via the GAM procedure in SAS. The GAM procedure is used to fit generalized additive models using nonparametric regression and smoothing techniques defined by degrees of freedom. Nonparametric regression is beneficial as it can bypass standard assumptions of linearity, which may in turn reveal an underlying relationship between the independent and dependent variables.

The regression line shows how the variability increases only marginally when PM_{10} concentration is 15ug/m^3 versus 20ug/m^3. This graphic sufficiently justifies our movement toward lowering the 15ug/m^3 cut off from our current cut point of 20ug/m^3.

Reducing this cut-off value will in turn increase the sample size to compensate for initial data loss resulting from our plan to reduce the collocated sample frequency burden from 1-in-6 day sampling to 1-in-12 day sampling.

The 1-in-12 day sampling frequency is only a minimum that will be proposed in CFR. Monitoring organizations are encouraged to collect PM_{10} samples at higher frequencies. Monitoring organizations with only one collocated monitoring site representing their primary quality assurance organization may want to perform additional collocated sampling. Also, the collocated site should be selected where there is the greatest chance for values > 15 ug/m^3.

Conclusions

Our evaluations of the $PM_{2.5}$ and PM_{10} data lead us to conclude that reducing the sample frequency and the PM_{10} cutoff value will not have a significant, adverse effect on our ability to provide precision estimates with adequate confidence. However, an important aspect in this conclusion is that the monitoring organizations select collocated sites that, for the majority of the time, are sampling concentrations above the cut-off value, and at a minimum, meet the 75% completeness criteria in order to collect a minimum of 22 samples/collocated site/year.

Table 1 PM10 Collocated Data 2002-2004

n= number of sample pairs (6=1-in-6 day sampling, 12= 1-in12-day sampling)

CV= coefficient of variation

CV_UB= coefficient of variation upper 90 confidence limit (> 66= >66 sample pairs, <66= < 66 sample pairs)

ABS Diff = absolute value of the difference between the 1-12 CV_UB from the 1-in-6 CV_UB

n_6	**n_12**	**CV_6**	**CV_12**	**CV_UB_6 >66**	**CV_UB_12 >66**	**CV_UB_6 <66**	**CV_UB-12 <66**	**ABS Diff**
6	3	5.75	1.09			10.13	3.35	6.78
7	3	3.59	4.05			5.93	12.48	6.55
7	4	1.75	1.29			2.89	2.91	0.02
8	4	2.19	2.15			3.44	4.87	1.44
9	4	2.92	2.15			4.43	4.87	0.45
7	4	1.98	2.33			3.27	5.28	2.00
11	5	6.86	5.91			9.84	11.46	1.63
12	6	3.16	2.00			4.44	3.53	0.91
14	7	2.62	2.41			3.56	3.98	0.42
16	8	3.37	3.59			4.47	5.64	1.17
21	10	7.21	7.51			9.14	11.04	1.90
23	11	1.49	1.09			1.86	1.57	0.29
22	11	6.28	7.21			7.91	10.34	2.43
22	11	11.15	7.85			14.04	11.26	2.78
24	12	1.80	2.00			2.24	2.81	0.57
25	12	4.38	2.28			5.42	3.21	2.21
24	12	3.04	3.08			3.78	4.32	0.54
24	12	3.97	3.76			4.94	5.28	0.34
24	12	4.51	4.01			5.61	5.63	0.02
26	13	2.33	2.08			2.87	2.87	0.00
30	15	7.48	8.50			9.06	11.39	2.33
33	16	4.02	2.75			4.81	3.64	1.17
32	16	10.94	10.79			13.16	14.29	1.13
33	17	5.90	3.48			7.08	4.57	2.51
36	18	3.14	4.10			3.73	5.33	1.59
41	20	3.99	3.26			4.68	4.16	0.51
40	20	4.81	4.27			5.65	5.46	0.20
41	20	8.59	10.00			10.08	12.77	2.69
44	22	4.16	3.90			4.86	4.91	0.05
46	23	2.19	2.20			2.54	2.75	0.21
46	23	4.33	4.60			5.03	5.76	0.72
46	23	5.59	5.28			6.49	6.61	0.12
48	24	3.62	3.37			4.19	4.19	0.01
48	24	5.77	5.55			6.68	6.90	0.22
49	25	5.74	5.94			6.63	7.35	0.72
52	26	5.32	5.24			6.12	6.46	0.33
52	26	6.27	6.42			7.21	7.91	0.70
54	27	2.96	3.40			3.39	4.17	0.78
53	27	7.18	7.61			8.24	9.33	1.09
56	28	3.58	3.61			4.09	4.40	0.31
56	28	5.75	7.17			6.57	8.75	2.18
58	29	7.17	7.50			8.18	9.12	0.93
64	32	2.44	1.55			2.76	1.87	0.90
65	32	6.89	6.30			7.80	7.58	0.22
63	32	9.38	9.79			10.64	11.77	1.13

n_6	n_12	CV_6	CV_12	CV_UB_6 >66	CV_UB_12 >66	CV_UB_6 <66	CV_UB-12 <66	ABS Diff	
66	33	3.16	2.36			3.57	2.82	0.74	
65	33	3.99	3.75			4.51	4.50	0.01	
67	33	7.67	8.71			8.66	10.45	1.79	
68	34	5.97	6.78			6.73	8.10	1.37	
70	35	3.26	2.25			3.67	2.68	0.98	
73	37	1.96	2.00			2.20	2.37	0.16	
78	39	4.07	3.19			4.55	3.76	0.79	
77	39	3.13	3.42			3.51	4.04	0.53	
79	39	3.83	4.00			4.28	4.72	0.43	
80	40	4.07	4.59			4.55	5.40	0.85	
81	40	4.56	5.39			5.08	6.34	1.26	
83	41	5.60	2.46			6.24	2.89	3.35	
86	43	2.72	2.31			3.03	2.69	0.33	
86	43	7.57	7.09			8.41	8.28	0.13	
87	44	2.28	2.62			2.54	3.05	0.52	
90	45	8.56	7.49			9.50	8.72	0.78	
92	46	3.47	3.25			3.85	3.77	0.08	
92	46	4.53	4.70			5.02	5.46	0.44	
99	49	3.21	2.94			3.54	3.39	0.15	
102	51	9.71	9.71			10.69	11.18	0.49	
103	52	2.93	2.60			3.22	2.99	0.23	
103	52	5.21	5.51			5.74	6.34	0.60	
104	52	8.47	7.94			9.32	9.13	0.19	
106	53	3.22	3.58			3.54	4.12	0.57	
106	53	4.81	5.61			5.29	6.45	1.16	
107	54	2.87	2.61			3.15	2.99	0.16	
107	54	5.04	3.32			5.53	3.81	1.73	
109	54	4.44	4.68			4.88	5.37	0.49	
108	54	12.06	12.32			13.24	14.13	0.89	
111	55	3.79	3.91			4.15	4.48	0.33	
109	55	6.38	5.63			7.01	6.45	0.56	
114	57	6.36	6.71			6.96	7.66	0.70	
118	59	4.75	3.92			5.19	4.46	0.73	
118	59	4.50	4.52			4.92	5.15	0.23	
120	60	5.52	5.93			6.04	6.74	0.71	
123	61	4.16	3.92			4.54	4.46	0.08	**1.71**
130	65	5.47	5.89	5.96	6.67			0.71	
134	67	6.45	7.25	7.01	8.19			1.18	
141	71	3.83	3.90	4.15	4.39			0.24	
143	72	2.79	3.18	3.02	3.57			0.55	
154	77	2.40	2.55	2.60	2.85			0.26	
164	82	5.25	4.48	5.66	4.99			0.66	
165	83	3.88	4.05	4.19	4.51			0.33	
170	85	3.62	3.71	3.89	4.13			0.23	
174	87	7.22	7.87	7.76	8.74			0.98	
203	101	4.29	3.67	4.59	4.04			0.54	
203	102	5.14	4.92	5.49	5.42			0.08	
213	107	2.31	2.15	2.46	2.36			0.11	
235	117	4.81	4.47	5.12	4.89			0.23	
235	118	10.53	10.53	11.20	11.51			0.31	
246	123	3.96	4.33	4.21	4.72			0.51	

n_6	n_12	CV_6	CV_12	CV_UB_6 >66	CV_UB_12 >66	CV_UB_6 <66	CV_UB-12 <66	ABS Diff	
255	128	3.96	3.55	4.21	3.87			0.34	
270	135	2.36	2.30	2.50	2.50			0.01	
284	142	3.66	3.69	3.88	4.00			0.13	
304	152	4.59	4.52	4.84	4.89			0.04	
346	173	6.17	6.38	6.49	6.86			0.37	
378	189	7.77	8.02	8.16	8.59			0.44	
471	235	6.40	6.11	6.68	6.50			0.18	
578	289	2.94	3.01	3.06	3.19			0.13	**0.21**
		4.86	**4.68**	**5.09**	**5.28**			**0.85**	

Table 2. Collocated PM_{10} Precision Estimates using 20 ug/m^3 and 15 ug/m^3 cutoff values

n= number of sample pairs (6=1-in-6 day sampling, 12= 1-in12-day sampling)

CV= coefficient of variation

CV_UB= coefficient of variation upper 90 confidence limit (_20 = 20 ug/m^3 cutoff, 15 = 15 ugm^3 cutoff)

Diff 15-20 = CV_UB_20 – CV_UB_15

n_20	n_15	CV_20	CV_15	CV_UB_20	CV_UB_15	Diff 15-20
20	31	3.26	3.73	4.16	4.51	0.34
33	32	3.75	4.27	4.50	5.14	0.64
20	33	10.00	10.55	12.77	12.65	0.12
26	34	5.24	5.23	6.46	6.26	0.20
31	36	6.37	7.74	7.69	9.20	1.51
23	37	4.60	5.22	5.76	6.18	0.42
27	39	7.61	7.79	9.33	9.19	0.14
24	40	3.37	3.40	4.19	3.99	0.20
22	40	3.90	3.69	4.91	4.34	0.57
20	40	4.27	4.63	5.46	5.45	0.01
29	40	7.50	7.38	9.12	8.68	0.44
28	41	7.17	4.20	8.75	4.93	3.82
24	44	5.55	6.25	6.90	7.29	0.39
28	45	3.61	3.23	4.40	3.76	0.65
39	45	3.19	4.65	3.76	5.42	1.65
26	45	6.42	6.60	7.91	7.68	0.23
23	46	2.20	2.19	2.75	2.55	0.21
46	46	4.70	4.70	5.46	5.46	0.00
40	47	7.32	7.97	8.61	9.24	0.63
34	48	6.78	6.99	8.10	8.10	0.01
25	48	5.94	7.90	7.35	9.14	1.79
44	49	2.62	3.36	3.05	3.88	0.83
23	49	5.28	7.24	6.61	8.36	1.75
32	51	1.55	2.48	1.87	2.86	0.99
32	51	9.79	9.46	11.77	10.89	0.88
27	53	3.40	3.08	4.17	3.53	0.64
41	53	2.46	6.43	2.89	7.38	4.49
33	53	8.71	8.72	10.45	10.02	0.43
33	54	2.36	2.29	2.82	2.63	0.20
44	54	7.12	8.98	8.31	10.29	1.98
43	57	2.31	2.78	2.69	3.18	0.48
39	58	3.42	3.44	4.04	3.93	0.11
40	59	5.39	5.07	6.34	5.77	0.57
53	62	4.73	4.11	5.43	4.67	0.76
54	62	3.32	6.09	3.81	6.91	3.11
52	62	7.94	7.83	9.13	8.89	0.24
37	63	2.00	1.69	2.37	1.92	0.45
39	63	4.00	4.37	4.72	4.95	0.24
35	64	2.25	2.77	2.68	3.13	0.45
57	65	6.71	8.62	7.66	9.75	2.09
53	67	5.61	4.76	6.45	5.37	1.08
52	68	2.60	3.11	2.99	3.51	0.51
51	68	9.71	9.38	11.18	10.58	0.60
49	69	2.94	3.73	3.39	4.20	0.81
53	71	3.58	2.62	4.12	2.94	1.17

n_20	n_15	CV_20	CV_15	CV_UB_20	CV_UB_15	Diff 15-20
54	72	2.61	3.74	2.99	4.20	1.21
46	72	3.25	3.96	3.77	4.46	0.68
52	72	5.51	4.84	6.34	5.44	0.90
58	72	3.95	6.06	4.50	6.81	2.31
40	73	4.59	4.79	5.40	5.38	0.02
53	73	12.43	12.97	14.27	14.56	0.29
54	74	4.68	4.46	5.37	5.00	0.37
67	74	7.25	6.04	8.19	6.77	1.42
61	75	3.92	4.20	4.46	4.71	0.25
55	80	5.63	7.42	6.45	8.29	1.84
55	82	3.91	3.68	4.48	4.10	0.37
60	85	5.93	5.32	6.74	5.92	0.82
82	86	4.48	5.35	4.99	5.95	0.96
65	90	5.89	5.06	6.67	5.61	1.05
71	97	3.90	4.43	4.39	4.89	0.51
77	104	2.55	2.61	2.85	2.88	0.02
84	111	3.62	3.93	4.03	4.31	0.28
72	111	3.18	4.01	3.57	4.39	0.82
87	117	7.87	7.60	8.74	8.32	0.42
83	119	4.05	3.65	4.51	3.99	0.52
101	127	3.67	5.05	4.04	5.50	1.46
102	130	4.92	6.08	5.42	6.62	1.21
121	136	4.17	3.80	4.56	4.13	0.42
123	141	3.55	4.26	3.88	4.62	0.75
107	145	2.15	2.37	2.36	2.57	0.21
142	148	3.69	3.53	4.00	3.82	0.19
115	152	4.46	5.20	4.88	5.62	0.74
148	156	4.14	4.52	4.48	4.88	0.41
118	159	10.53	9.64	11.51	10.41	1.10
152	164	7.32	7.00	7.92	7.54	0.37
135	198	2.30	2.96	2.50	3.17	0.67
172	215	6.11	6.43	6.57	6.86	0.29
226	237	6.11	6.50	6.51	6.91	0.40
272	292	2.94	2.89	3.11	3.05	0.06
Average		**4.91**	**5.25**	**5.72**	**5.95**	**0.79**