APPENDIX II-J

APPLICABILITY OF SSAP CORES FOR ESTIMATION OF PCB MASS IN UNEXPECTED INVENTORY:

THE THOMPSON ISLAND DAM 2009

Appendix II-J

APPLICABILITY OF SSAP CORES FOR ESTIMATION OF PCB MASS IN UNEXPECTED INVENTORY:

THE THOMPSON ISLAND DAM 2009

March 5, 2010

Prepared for:

United States Environmental Protection Agency
Region 2
290 Broadway
New York, NY 10007-1866

United States Army Core of Engineers
Kansas City District
Kansas City, Missouri

Under Contract to:

Louis Berger Group, Inc.
412 Mt. Kemble Avenue,
Morristown, NJ 07960

Introduction

GE planned to remove 144,439 cubic yards (cy) of sediment from dredged certification units (CUs) in the Phase I project (Table D-11). Design dredging cut lines did not in general capture actual depths of contamination (DoC), due to:

1. Uncertainty in measured DoC due to incomplete penetration of the contaminated layer by SSAP cores.
2. Failure to follow EPA recommendations to validate DoC extrapolation models.
3. Failure to "hedge" the design cut lines (through incorporation of a dredging overcut) to compensate for spatial variability and uncertainty in the DoC interpolation models.

As a result, all 10 CUs required at least 3 dredging passes with up to 5 passes in CU-1. The total volume of contaminated sediment and mass of total and Tri-plus PCBs removed were reported by GE and EPA.

GE and EPA reported removal of similar volumes of contaminated sediments: 286,354 cy and 267,804 cy, respectively, with a relative percent difference (RPD; 2xDifference/Sum) of just 7 percent. Conversely, GE and EPA estimated the mass of PCB removed to be 16,320 kg and 20,020 kg, respectively, resulting in an RPD of over 20 percent and an absolute difference of 3,700 kg.

Although seemingly small, determination/achievement of compliance with the Residuals and Resuspension Standards is sensitive to this difference. Therefore it is important to resolve discrepancies in GE and EPA mass estimates. Understanding the root cause of differences between these mass estimates is important in order to interpret loading data to the Lower Hudson River and compliance with the Resuspension Standard. The following is an analysis of a likely source of bias in the mass estimates.

Potential Root Causes

1) Low bias of PCB concentration in SSAP samples near and below the design cut lines.
2) Differences in handling of bulk density.
3) Order of operations in mass calculations—product of averages vs. sum of products.
4) Weighted vs. un-weighted averaging.

Bias in SSAP Samples to Characterize Unexpected Inventory

EPA and GE base their mass calculations on different sets of PCB concentration data. GE uses a combination of post-dredging core samples and SSAP cores collected prior to dredging in forming the basis for setting cut lines. Regardless of calculation methods, a difference in the distribution of PCB concentrations among the two data sets would necessarily cause problems with reconciliation of any other steps in the process. Therefore this report focuses on an investigation of potential biases associated with SSAP PCB data.

Residuals Samples are Unbiased

Because the post-dredging samples were collected from the nodes of a regularly spaced grid, and because the post-dredging samples fully penetrate the 6-inch contaminated layer below the design cut lines, the un-weighted arithmetic average of PCB concentration in post-dredging core samples is an unbiased estimator of the concentration within the 6-inch interval below the design cut lines. These samples are sufficient to estimate mass of PCBs in unexpected inventory below the design cut lines. Because post-dredging core data are based on a relatively large (N=40) unbiased systematic sample, inclusion of other sources of unbiased data should result in little or no change in the estimated mean concentration. The primary benefit of inclusion of other sample data would be to improve precision of the estimated sample mean. If the SSAP data are also unbiased to the PCB concentrations in unexpected inventory their inclusion should not change estimates of concentration substantively.

Because of the prevalence of up to several feet of unexpected PCB inventory found below the design cut lines it is clear that DoC, as inferred from the SSAP cores, was frequently understated. Because the SSAP cores frequently do not fully penetrate the PCB-contaminated layer, one should expect that SSAP samples would be biased low.

Because the post-dredging core data are known to be representative (unbiased) of the concentration in the 6-inch layer of unexpected inventory, SSAP data should not be included in the mass estimation procedure without first demonstrating that they are equally unbiased. This can be investigated statistically to provide evidence of the nature of potential bias in SSAP cores particularly when used to estimate concentration of PCBs below the design cut lines.

If both the SSAP and the post-dredging core samples are unbiased to the true mean of PCB concentration in sediments below the design cut lines there should be no systematic differences between concentration in the SSAP and post-dredging sample data within the same CUs. To investigate this hypothesis, SSAP core segments with average depth (*i.e.*, centroid of the core section) within the 6-inch horizon below the first pass design elevation were compared with corresponding post-dredging cores from the same depth interval.

These subsets of data were grouped by CU and summarized as boxplots in Figure 1. In all 10 CUs dredged in 2009 the median (horizontal red line) PCB concentration for SSAP cores is less than that for the corresponding post-dredging core distribution. Under the null hypothesis of equal median concentrations the probability of observing fully 100 percent of the medians from the SSAP population below that of the post-dredging population is $0.5^{10} = 1/1000$. This strongly suggests that the SSAP data are not representative of PCB concentrations in the 6-inch layer of unexpected inventory directly below the design cut lines.

Acceptability of Complete Cores

One might conjecture that this bias is primarily due to the incomplete subset of the cores in hopes that the complete SSAP cores (*i.e.*, high confidence cores) might be suitable for application to estimation of mass of PCBs removed. To investigate this question incomplete cores were removed from the data and the distributions were again compared (Figure 2). Removal of the incomplete cores actually increases the magnitude of the bias, so calculations that preferentially incorporate complete cores would be expected to accentuate the degree to which mass may be understated when the SSAP cores are incorporated into the analysis.

Magnitude of the Bias

To quantify the magnitude of the bias, geometric means (appropriate for right skewed data) were calculated for the post-dredging and SSAP samples and the ratio of the geometric means was calculated for each CU based on all SSAP data as well as the complete core subset. For complete cores, ratios varied from 4:1 in CU-7 to 55:1 in CU-4, with an overall ratio of geometric means of 15:1. For complete and incomplete cores combined, the ratios ranged from approximately 1:1 at CU-1 to 20:1 at CU-18 with an overall ratio of approximately 7:1 for all CUs combined. These ratios demonstrate that use of SSAP cores would create a statistically significant and materially substantive low bias in estimates of PCB concentration and by extension PCB mass in unexpected inventory below design cut lines.

Source of the Bias

A high proportion of SSAP cores did not fully penetrate the PCB-contaminated layer. This is a form of right censoring of the PCB concentration distribution. At depths below the design cut lines, low concentration samples (complete cores) are over represented in the sample population because higher concentration PCB values in the population are unobservable due to incompleteness of cores—by definition a core is incomplete if the bottom sample exceeds 1 mg/kg.

The likelihood that an individual location would be incomplete is a function of the thickness of the sediment deposit—the deeper the deposit the greater the likelihood that the core does not penetrate the PCB-contaminated layer. Combining this with the fact that deeper deposits represent the depositional areas and contain more highly contaminated sediments than thinner deposits, the net effect is that unobserved core sections are likely to have higher concentrations than those that were observed in the bottoms of nearby complete cores. Figure 3 shows a hypothetical group of 7 cores and how observable core sections preferentially sample the lower concentration fraction of the unexpected PCB inventory.

In Figure 3 there are 7 cores, 4 of which are complete and three of which are incomplete. The data are laid out horizontally as if the core was on a table with surface elevations at the left and sediment depth increasing to the right. The designed dredging cut line is shown as a pink-shaded column, observable core sections below the dredge cut line are shaded green, and unobservable (*i.e.*, censored) observations are shaded gray. Because the complete cores, by definition, have observed clean sections below the DoC elevation, they are observable. In contrast, incomplete cores have concentrations greater than 1 mg/kg below their deepest recovered sections that are unobservable. Therefore the low concentration fraction of the population is over represented by the observable complete cores retained in the mass estimation analysis. The bottom two rows of the table compare the “observed” average concentration with the true (observed and unobserved sections) average of PCB in the sediment layer. This example illustrates the bias in estimated concentration that is likely. This is consistent with results seen in practice comparing SSAP and post-dredging core samples above.

Spatial Heterogeneity

One might suggest that differences between averages based on SSAP and post-dredging core samples could be due to spatial heterogeneity induced by the lack of collocation of samples. Both SSAP and post-dredging core sampling plans are based on regular systematic grids and therefore should both be representative of the concentrations within the CU. Any biases introduced by spatial variation of PCBs within CUs would require that the high concentration values were preferentially located at the nodes of one design while the low concentrations would necessarily be located at the spatially-intermingled nodes of the other design—in effect PCB concentrations distributed in an ‘egg carton’ pattern. This is really not a plausible situation. Additionally it is also implausible that lack of collocation might produce a bias between SSAP and post-dredging core samples that is consistently negative across all CUs.

Material Importance of the Bias

Because GE has understated concentration by combining SSAP cores with post-dredging core samples, the mass of PCBs in unexpected inventory is understated. Because aspects of the Resuspension Standard, as well as estimates of remedial efficiency depend on these estimates, this mistake in mass estimation is propagated into calculations intended to evaluate efficiency of removal of deeper layers of PCB-contaminated sediment as well as compliance with the Resuspension Standard.

GE stated that the percentage of mass removed declines rapidly with successive dredging passes (*i.e.*, with depth). Because the bias in the SSAP cores increases with depth (*i.e.*, likelihood of incomplete cores increases with thickness of the sediment deposit) it is fully expected that the difference between SSAP and post-dredging core PCB concentrations would increase with depth. This suggests that the apparent reduction in percentage mass removed identified by GE may actually be in part a spurious consequence that the bias in SSAP and post-dredging core samples increases with depth.

This can be investigated by comparing the RPD between EPA and GE estimates as they relate to successive dredging passes. Figure 4 shows that RPDs between EPA and GE mass estimates increase with each dredging pass, ranging from around 10 percent in the first dredge pass to nearly 60 percent by the fourth dredge pass. This may be due to the bias in the SSAP data described above.

GE argues that dredging beyond the first or perhaps second pass is inefficient based on these mass estimates, which unlike EPAs estimates, decline substantially on a per unit volume basis with increasing depth.   This understatement of PCB mass is likely due to:

1. Biased estimates of the efficiency of moving deeper sediment deposits.
2. Cloud issues related to evaluation of compliance with the Resuspension Standard.
3. Understating potential benefits of the active remedy.
4. Understating the extent to which DoC delineations failed to accurately target the DoC.

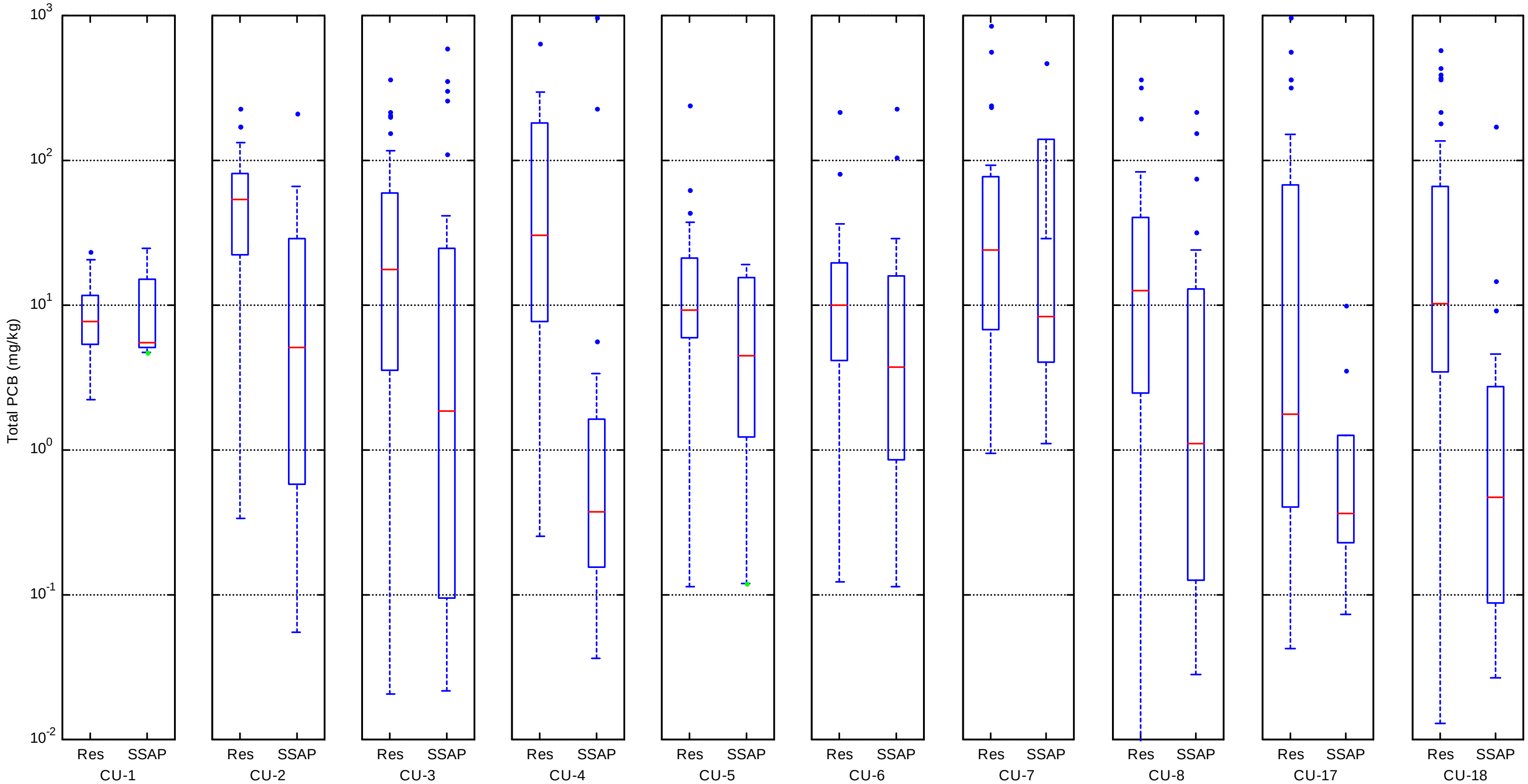

Figure 1. Boxplots of total PCB comparing residual and SSAP cores with centroids within the first six inch interval below the design elevation in Phase-I dredging units, upper Hudson River, NY. Red lines represent the median concentration, the boxes represent the 25^{th} and 75^{th} percentiles and the "whiskers" are the lesser of 1.5 times the box length (inter-quartile range) and the maximum PCB value. Median PCB concentration in SSAP cores was lower than in Residual cores in all 10 certification units.

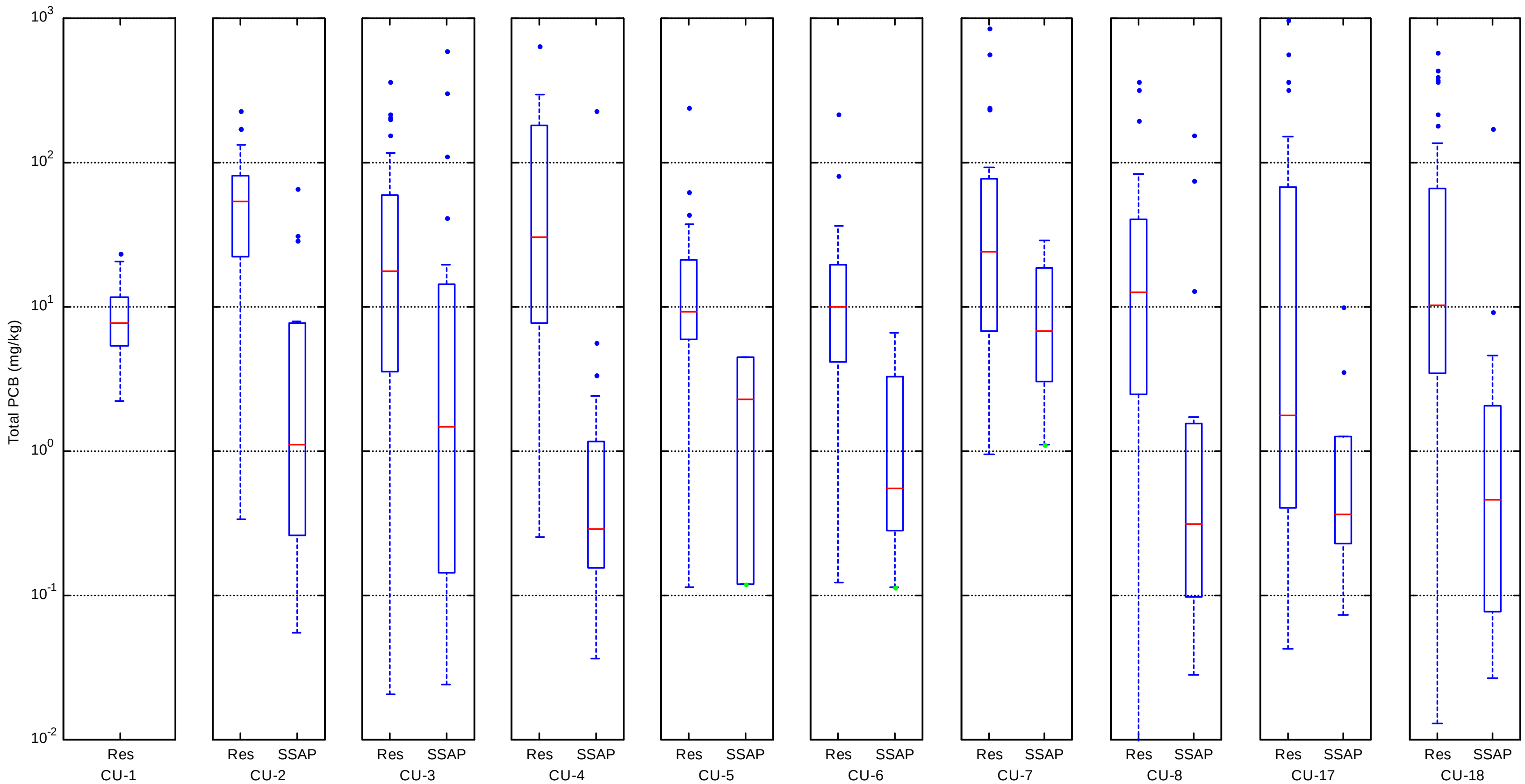

Figure 2. Boxplots of total PCB comparing residual and complete SSAP cores with centroids within the first six inch interval below the design elevation in Phase-I dredging units, upper Hudson River, NY. Red lines represent the median concentration, the boxes represent the 25th and 75th percentiles and the "whiskers" are the lesser of 1.5 times the box length (inter-quartile range) and the maximum PCB value. Median PCB concentration in SSAP cores was lower than in Residual cores in all 10 certification units.

Bias Associated with SSAP Cores for Estimating Post Dredging Mass								
						First Pass	Second Pass	Third Pass
Core Type	**0 to 6**	**6 to 12**	**12-18**	**18-24**		**24 to 30**	**30 to 36**	**36 to 42**
1--Complete	30	10	5	1	**D**	0.5	0.5	0.1
2—Incomplete	20	12	10	1	**R**	**10**	**5**	**1**
3—Complete	10	5	2	1	**E**	0.2	0.75	0.1
4—Complete	20	10	4	1	**D**	0.3	0.5	0.1
5--Increasing Profile	5	10	Double Depth		**G**	**15**	**10**	**5**
6--Increasing Profile	10	20	Double Depth		**E**	**10**	**5**	**1**
7—Complete	10	5	6	1		0.6	0.1	0.1
True Average						**5.2**	**3.1**	**1.1**
Apparent Estimate						**0.4**	**0.5**	**0.1**

Only complete SSAP cores are available to inform average

Excluded from estimate because the core is incomplete and sample is unobservable

Figure 3. Illustration of bias resulting from hypothetical group of complete and incomplete cores.

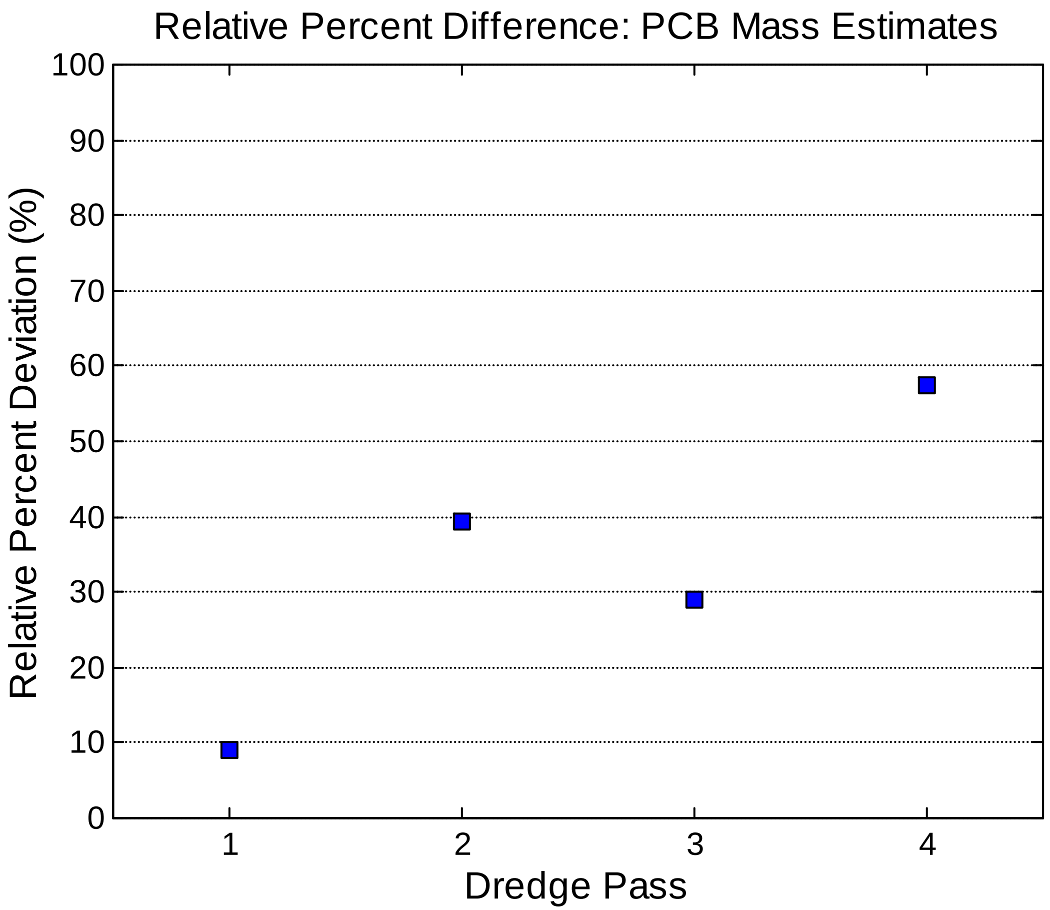

Figure 4. Relative percent difference between GE and EPA estimates as a function of dredge pass.

Table 1. Geometric mean concentrations in residual samples and SSAP samples within the unexpected inventory below the first pass design elevations in Phase I certification units dredged in 2009 in the Upper Hudson River.

	Residual Cores		Complete SSAP Cores		
Certification Unit	Count	Geometric Mean (mg/kg)	Count	Geometric Mean (mg/kg)	Ratio: Residual: SSAP
1	43	2.05	0	ND	ND
2	40	41.76	15	1.70	24.62
3	47	12.26	23	1.61	7.62
4	42	27.02	20	0.49	55.02
5	28	9.92	2	0.73	13.50
6	30	7.38	5	0.79	9.30
7	41	21.89	4	6.08	3.60
8	52	8.00	15	0.60	13.37
17	39	4.08	10	0.57	7.20
18	43	11.61	27	0.51	22.98
Overall	405	12.19	121	0.84	14.59