APPENDIX D

SEWRPC Memorandum – February 5, 2004 – Draft memo of CSO & SSO Pollutant Concentrations for Purposes of Water Quality Modeling – Dated October 29, 2003 [Addresses nitrogen species concentration development in MMSD's CSOs and SSOs]

SEWRPC STAFF MEMORANDUM

TO: Pat Marchese and Mary Recktenwalt

FROM: Bob Biebel, Ron Printz, Thomas Slawski, and Joseph Boxhorn

DATE: February 5, 2004

SUBJECT: DRAFT MEMO OF CSO & SSO POLLUTANT CONCENTRATIONS FOR PURPOSES OF WATER QUALITY MODELLING DATED OCTOBER 29, 2003

With regard to the subject memorandum and a subsequent meeting with Triad Engineering staff on January 28, 2004, we have developed several recommendations for your consideration. These recommendations pertain to issues first raised in a Memorandum to Mr. Pat Marchese dated December 8, 2003, attached hereto as Exhibit A and summarized below.

Use of Analysis of Variance (ANOVA)

We agree with the appropriateness of performing an ANOVA for each parameter among sites as well as among watersheds. Based upon further reanalysis and *post-hoc* testing of the CSO data (see Watershed and Collector Sites Analysis section below and section I of Exhibit B), we agree that the Menomonee River collector CT56 is statistically different from all other sites in terms of biological oxygen demand (BOD), total phosphorus (P), and total suspended solids (TSS), which justifies using a separate mean concentration for each of these constituents as recommended in the draft memo. In addition, we also recommend removal of this collector site prior to performing ANOVA by watershed for BOD, P, and TSS constituents. In contrast, fecal coliform counts at collector CT5/6 were not found to be significantly different and we suggest that collector CT5/6 not be removed prior to performing ANOVAs by watershed for this constituent (see section II of Exhibit B).

Watershed and Collector Sites Analysis

For identification of which watersheds and sites differ in CSO chemistry parameters, we suggest using *post hoc* pairwise comparisons of means in those instances where ANOVA has identified the existence of differences among means. This methodology is preferred over conducting a large series of *t* tests. At a significance level of 0.05, a series of 80 *t* tests would be expected to produce four significant results by chance alone. Performing a large numbers of *t* tests can result in numerous spurious findings of statistically significant differences. The pairwise comparisons that we suggest using were designed to eliminate this problem. Several of these tests exist and are commonly supported by statistics software. The different tests do have different properties. In this case, our suggestion is to use Bonferroni's test. With small numbers of means to be tested, it is more powerful than Tukey's test. Scheffé's test was designed to test differences among all possible linear combinations of group means and is generally less sensitive to differences than Bonferroni's. Fisher's LSD test is to be avoided as it produces the same problems as multiple *t* tests.

Our analyses do suggest that for some parameters there are statistically significant differences between the mean CSO chemistry from site CT5/6 and other sites. ANOVAs conducted by site showed statistically significant differences among means for BOD, TSS, and P. For BOD, Bonferroni's test showed that the mean BOD at CT5/6 was different from each of the other sites. No other differences among sites were detected. For TSS, Bonferroni's test showed that mean TSS at CT5/6 was different from that at LMS, NS10, and NS11. No other differences among sites were detected. For P, Bonferroni's test showed that mean P at CT5/6 was different from that at LMN, LMS, NS6, NS8, NS9, and NS12. Again, no other differences among sites were detected.

Based upon additional information provided by Triad Engineering as shown in Exhibit C, collector CT5/6 does contain one of the largest drainage areas and serves the highest number of persons compared to the other collector basins, which may explain why this collector contains higher levels of pollutants. However, this increased pollutant loading may also be a function of a variety of additional factors that include but not limited to: the nature, amount, and number of pollutants generated in this sewershed area; type and proportions of land use in the sewershed area; as well as some inherent difference in the physical structure in the system of CSO pipes themselves. Despite these unknown factors, we still recommend that separate geometric means for the recommended CSO modeling concentrations for BOD, TSS, and Total Phosphorus be utilized for collector CT5/6 as shown in Table 2 below.

CSO Trend Analysis

Based upon additional time-series analyses performed by Triad Engineering staff, we suggest that the subject memorandum include the linear regression analyses of CSO BOD, P, TSS, and fecal coliform counts as a function of time. In the cases of BOD and P, the p values in the regression statistics are greater than 0.05, showing that there were no statistically significant trends in these factors over time during the period sampled (Exhibit D). With respect to these two variables, the regressions indicate that the average composition of the CSOs has not changed over the period sampled. The situation is more complicated with respect to TSS and fecal coliform counts. For TSS, the p value in the regression statistics is less than 0.05. This shows that there was a statistically significant trend over the period sampled. The R^2 value in the regression statistics was 0.04, indicating that this trend accounted for only about 4 percent of the variation in data set. The situation is similar for fecal coliform counts. While the p value in the regression statistics does indicate a statistically significant trend over the sample period, the R^2 value shows that the trend accounted for only about 14 percent of the variation. In both of these cases, the effects of the trends were small compared to the effects caused by other sources of variability. Overall, these analyses suggest that change over the sampled period in the composition of the CSOs, with respect to BOD, P, TSS, and fecal coliform counts, is not a major consideration in the selection of CSO values for the model and supports the existing approach in the original draft memo.

Nitrogen Species Concentrations for SSO and CSO

Based upon additional analyses, we suggest that the subject memorandum include the linear regression analysis of SSO Jones Island influent data and the CSO data from the Ohio River Valley Water Sanitation Commission (ORVWSC) in Louisville of ammonia concentrations as a function of BOD as shown in Figure 1 below (see sections III and IV of Exhibit B). The p values in the regression statistics are less than 0.001 for each analysis and shows that there are statistically significant trends between these constituents for each facility. The R^2 values in the regression statistics were 0.91 for the SSO and 0.61 in the CSO analysis, indicating that these trends accounted for about 91 and 61 percent of the variation in each data set, respectively.

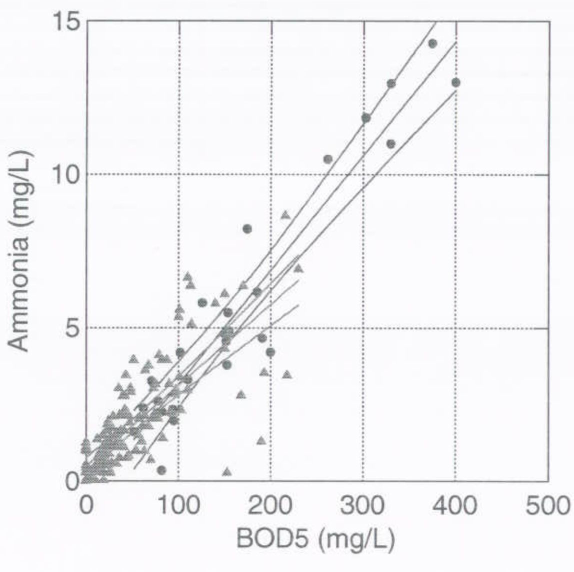


Figure 1.

Regression analysis between BOD and Ammonia concentrations from the MMSD Jones Island SSO influent and the Ohio River Valley Sanitation Commission's CSO, which include 95 percent confidence intervals.

Results from the Jones Island influent data indicate that the concentration of ammonia changes linearly as a function of BOD concentration. The regression shows that this relationship has a significant y-intercept. Because of this, using ratios of average ammonia concentration to BOD concentration will introduce bias into the estimates of ammonia concentration. For instance, the linear regression model indicates that the recommended SSO mean concentration of 3.5 mg/L for Ammonia is a significant overestimate based upon an average SSO mean concentration of 31 mg/L BOD. Unfortunately, the Jones Island influent BOD concentration data, which range from 52 to 400 mg/L, are beyond the concentration range of the recommended SSO mean concentration of 31 mg/L BOD. Hence, it is not recommended that the Jones Island relationship between BOD and ammonia be utilized to estimate the actual concentration of ammonia, either based upon the linear regression model or as a percentage of the BOD concentration. Given that the Jones Island data are not dissimilar to the reported data in the Ohio River Valley Sanitation Commission study, and that the ORVWSC data do cover the range of BOD concentrations indicated, we suggest that the recommended SSO and CSO nitrogen species mean concentrations for modeling be based upon data from the ORVWSC. It is the only real data we are aware of that contains a high number of actual measurements of BOD, Total Suspended Solids, Total Phosphorus, nitrate/nitrate, Ammonia, and organic nitrogen concentrations within the range of reported values in SSO and CSO within MMSD.

We also suggest that the subject memorandum include the linear regression analysis of CSO data from the ORVWSC of organic nitrogen concentrations as a function of BOD, as shown in Figure 2 below (see section IV of Exhibit B). The p value in the regression statistics is less than 0.0001 for this analysis and shows that there is a statistically significant trend between these constituents. The R^2 value in the regression statistic was 0.64, indicating that this trend accounted for about 64 percent of the variation in the data set. For these reasons, we suggest that the recommended SSO and CSO mean concentrations for organic nitrogen be based upon this linear regression model with BOD concentrations as shown in Tables 1 and 2 below.

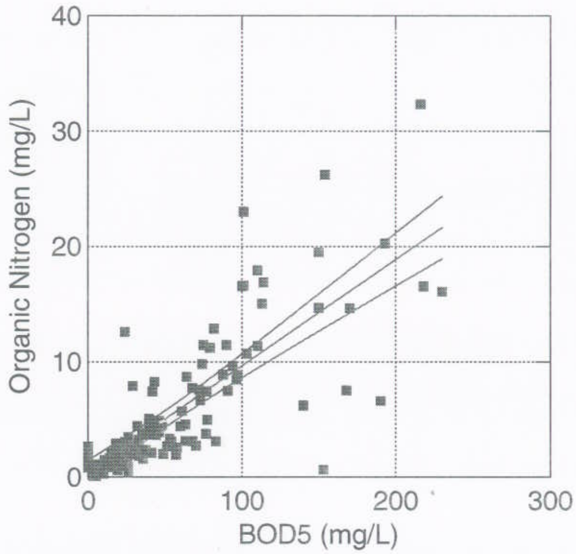


Figure 2

Regression analysis between BOD and Organic Nitrogen concentrations from the Ohio River Valley Sanitation Commission's CSO, which include 95 percent confidence intervals.

Table 1
Recommended SSO Mean Concentrations For Modeling (bolded values indicate recommended changes from the original draft memo)

Parameter	BOD_5 (mg/L)	Total Suspended Solids (mg/L)	Fecal Coliform (#/100 mL)	Phosphorus (mg/L)	Organic Nitrogen as N (mg/L)	Ammonia as N (mg/L)
Source	MMSD sampling	MMSD sampling	MMSD sampling	MMSD sampling	**ORVWSC sampling***	**ORVWSC sampling***
All Watersheds	31	126	530,000	2.2	**3.3**	**1.3**

Assume: 1) Nitrate, nitrite, chlorophyll a, and dissolved oxygen to be negligible. 2) Treat temperature similar to how it is treated for storm water runoff. Note: Fecal coliform concentration was rounded to two significant figures.
*Kim Mays from the Ohio River Valley Water Sanitation Commission (ORVWSC) provided sampling data August 19, 2003.

Table 2
Recommended CSO Mean Concentrations For Modeling (bolded values indicate recommended changes from the original draft memo)

Parameter	BOD_5 (mg/L)	Total Suspended Solids (mg/L)	Fecal Coliform (#/100 mL)	Phosphorus (mg/L)	Copper ICP (mg/L)	Zinc ICP (mg/L)	Organic Nitrogen-as N (mg/L)	Ammonia-as N (mg/L)	Nitrate/Nitrite-as N (mg/L)
Source	MMSD sampling	MMSD sampling	MMSD sampling	MMSD sampling	MMSD sampling	MMSD sampling	**ORVWSC sampling***	**ORVWSC sampling***	**ORVWSC sampling***
Menomonee River (all but CT 5/6)	9	56	160,000	0.64	0.02	0.09	**1.3**	**0.70**	1.0
Menomonee River (only CT 5/6)	54	116	**160,000**	1.07	0.02	0.12	**5.4**	**2.0**	1.0
Kinnickinnic River	9	56	160,000	0.64	0.02	0.09	**1.3**	**0.70**	1.0
Milwaukee River	9	56	160,000	0.48	0.02	0.09	**1.3**	**0.70**	1.0

*Kim Mays from the Ohio River Valley Water Sanitation Commission (ORVWSC) provided sampling data August 19, 2003.

* * *

TMS/JEB
#91155 V1 - DRAFT CSO & SSO POLLUTANT CONC-MEMO 012804

Exhibit A to Appendix D

MEMORANDUM

TO: Pat Marchese

FROM: Bob Biebel, Ron Printz, and Thomas Slawski

DATE: December 8, 2003

SUBJECT: DRAFT MEMO OF CSO & SSO POLLUTANT CONCENTRATIONS FOR PURPOSES OF WATER QUALITY MODELLING DATED OCTOBER 29, 2003

With regard to the subject memorandum, we suggest consideration be given to carrying out additional analyses to demonstrate the appropriateness of performing an Analysis of Variance (Anova) for each parameter among watersheds as well as a potential alternative approach to better understand variance of each parameter among outfall collectors.

We understand that Anova is a fairly robust statistical model that can tolerate a fair amount of departure from normality, however, many large outliers or extreme differences in variability among groups are not as easily tolerated. Based upon the reported mean and variance values for BOD, TSS, fecal coliform, and total phosphorus parameters as shown in Appendix A, it seems likely that this condition is met. Nonetheless, we recommend that a test for the homogeneity of variance be performed among watersheds to be certain that the variances within each group are roughly equal. SYSTAT™ recommends developing a box plot by watershed for each parameter to visually inspect whether or not distributions differ (SYSTAT™ 10.2 Statistics I, Copyright © 2002 by SYSTAT Software Inc.). If few differences are observed in the spread of the boxes, Levene's test for unequal variances is unlikely to be significant. If large differences exist then performance of Levene's test is warranted.

If the homogeneity of variance assumption is not met, then we shouldn't really carry out the Anova as the variance within groups is different for different groups. Such a case might reflect something more systematic. For example, Anova would be unable to distinguish any difference among watersheds, if the variance within each watershed is higher than the variance between the watersheds: the variances within each group should be roughly equal in order to use Anova. Specifically, this is what appears to be happening, based upon an inspection of the data set forth in Table 4 and Appendix A, in the case of fecal coliform concentrations. If there is a lack of homogeneity among watersheds for a particular parameter, then we recommend each watershed be treated/analyzed separately.

Assuming that the variances are equal, then, we further recommend that an Anova be modified to test for effect of watershed (Menomonee, Kinnickinnic, Lake MI, and Milwaukee) and the interaction between watershed and sampling sites (20 outfall collection sites). This analysis may offer the added ability to simultaneously test to see if there are similarities or differences among watersheds, as well as differences among the CSO collector sites themselves. Hence, the results of this analysis will allow us to either (a) statistically identify CSO sites that are loading significantly higher or lower compared to all of the sites combined, or (b) conclude that a system-wide value for the particular parameter can be used. Because the existing analysis suggests that there are differences between CSO collector sites, this approach would form a more valid approach to identifying high pollutant outlier sites, such as may be the case for site CT5/6, as well as low pollutant outlier sites, which have yet to be identified, for each parameter, but may be equally as important from a modeling simulation point of view.

If the results of this analysis indicate significant differences among watersheds or among sites within watersheds, then each of the watersheds should be treated separately. It is not recommended that a site be removed from the data set and the remaining data reanalyzed. In addition, if there is significant difference between sites within a watershed, then each site would have to be treated separately based upon this statistical difference.

Ancillary issues:

Identify phosphorus as "total phosphorus" in the recommended mean concentrations for modeling in all of the tables.

On page 4 of 10, an assumption is made that the SSO concentration for ammonia is similar to that measured for the Jones Island inflow. What was the basis for this? Were any further analyses made to verify this? A comparison of other constituents measured at both SSO's and Jones Island could be made to see if those concentrations are similar.

In recommendation #3 on page 6 of 10, what pollutant loadings report that used arithmetic means is being referred to?

* * *

TMS/
#89489 V1 - DRAFT CSO & SSO POLLUTANT CONC-MEMO 120503

Exhibit B to Appendix D

Exhibit B

The following analyses outputs were completed using Systat™ (SYSTAT™ 10.2 Statistics I, Copyright © 2002 by SYSTAT Software Inc.) and separated in the following sections:

Section I
MMSD CSO Analysis of variance (ANOVA) by all collector sites for each of the following constituents:

- Biological Oxygen Demand (BOD)
- Total suspended solids (TSS)
- Fecal coliform, and
- Total Phosphorus (TP)

Section II
MMSD CSO ANOVA by watershed without collector CT5/6 for each of the following constituents:

- Biological Oxygen Demand (BOD)
- Total suspended solids (TSS)
- Total Phosphorus (TP)

MMSD CSO ANOVA by watershed with all collector sites for the following constituent:

- Fecal coliform

Section III
MMSD SSO Jones Island influent regression analysis between BOD and Ammonia

Section IV
Ohio River Valley Water Sanitation Commission (ORVWSC) regression analysis between BOD versus Ammonia and BOD versus Organic Nitrogen

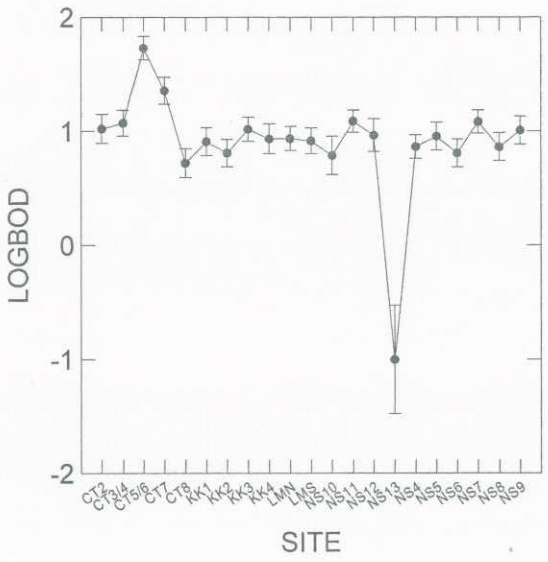
Section I

MMSD CSO ANOVAS analyzed by site
These analyses include all sites on the Menomonee River

```
Effects coding used for categorical variables in model.

Categorical values encountered during processing are:
SITE$ (21 levels)
   CT2, CT3/4, CT5/6, CT7, CT8, KK1, KK2, KK3, KK4, LMN, LMS, NS10, NS11, NS12,
   NS13, NS4, NS5, NS6, NS7, NS8, NS9

Dep Var: LOGBOD   N: 332   Multiple R: 0.4825   Squared multiple R: 0.2328

Analysis of Variance
```

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
SITE$	21.2940	20	1.0647	4.7184	0.0000
Error	70.1760	311	0.2256		

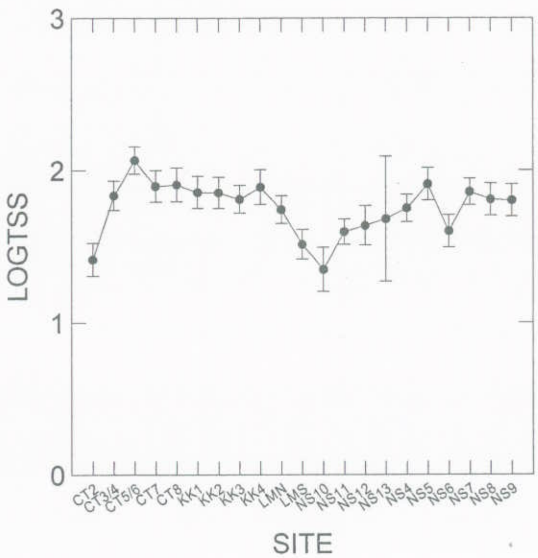
Least Squares Means




```
*** WARNING ***
Case         74 is an outlier        (Studentized Residual =      -3.8407)
Case         85 is an outlier        (Studentized Residual =      -4.2728)
Case        100 is an outlier        (Studentized Residual =      -4.0259)
Case        117 is an outlier        (Studentized Residual =      -4.4870)
Case        137 is an outlier        (Studentized Residual =      -4.3569)
Case        150 is an outlier        (Studentized Residual =      -4.2950)
Case        189 is an outlier        (Studentized Residual =      -4.1228)
Case        261 is an outlier        (Studentized Residual =      -4.1633)
Case        276 is an outlier        (Studentized Residual =      -4.5097)

Durbin-Watson D Statistic         2.1668
First Order Autocorrelation      -0.0835
Effects coding used for categorical variables in model.

Categorical values encountered during processing are:
SITE$ (21 levels)
   CT2, CT3/4, CT5/6, CT7, CT8, KK1, KK2, KK3, KK4, LMN, LMS, NS10, NS11, NS12,
   NS13, NS4, NS5, NS6, NS7, NS8, NS9
1 case(s) deleted due to missing data.

Dep Var: LOGTSS   N: 331   Multiple R: 0.3853   Squared multiple R: 0.1485

Analysis of Variance
```

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
SITE$	9.0656	20	0.4533	2.7024	0.0001
Error	51.9972	310	0.1677		

Least Squares Means




```
Durbin-Watson D Statistic          1.7067
First Order Autocorrelation        0.1427
```

Effects coding used for categorical variables in model.

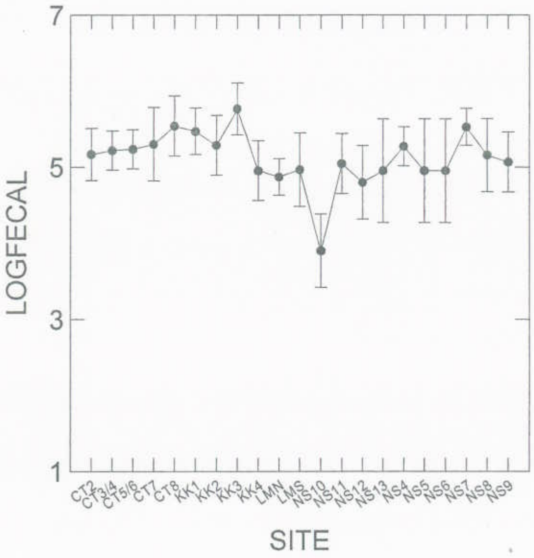
Categorical values encountered during processing are:
SITE$ (21 levels)
   CT2, CT3/4, CT5/6, CT7, CT8, KK1, KK2, KK3, KK4, LMN, LMS, NS10, NS11, NS12,
   NS13, NS4, NS5, NS6, NS7, NS8, NS9
254 case(s) deleted due to missing data.

Dep Var: LOGFECAL   N: 78   Multiple R: 0.4837   Squared multiple R: 0.2340

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
SITE$	8.0722	20	0.4036	0.8706	0.6221
Error	26.4264	57	0.4636		

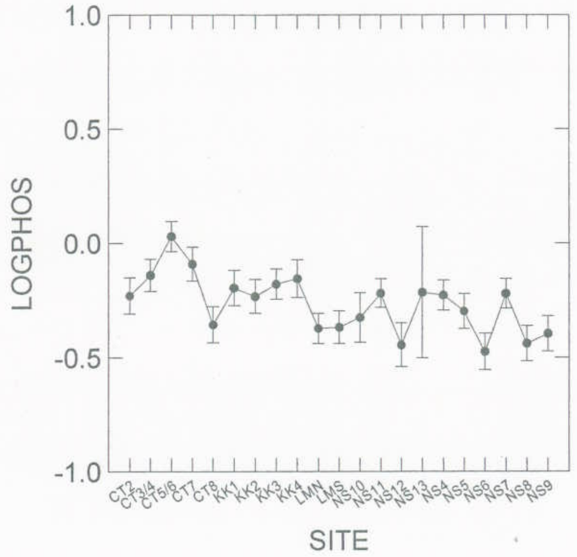
Least Squares Means




```
*** WARNING ***
Case          151 is an outlier        (Studentized Residual =       -4.0080)

Durbin-Watson D Statistic         2.3527
First Order Autocorrelation      -0.1777


Effects coding used for categorical variables in model.

Categorical values encountered during processing are:
SITE$ (21 levels)
   CT2, CT3/4, CT5/6, CT7, CT8, KK1, KK2, KK3, KK4, LMN, LMS, NS10, NS11, NS12,
   NS13, NS4, NS5, NS6, NS7, NS8, NS9
28 case(s) deleted due to missing data.

Dep Var: LOGPHOS   N: 304   Multiple R: 0.4166   Squared multiple R: 0.1736

Analysis of Variance
```

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
SITE$	4.9114	20	0.2456	2.9722	0.0000
Error	23.3821	283	0.0826		

Least Squares Means



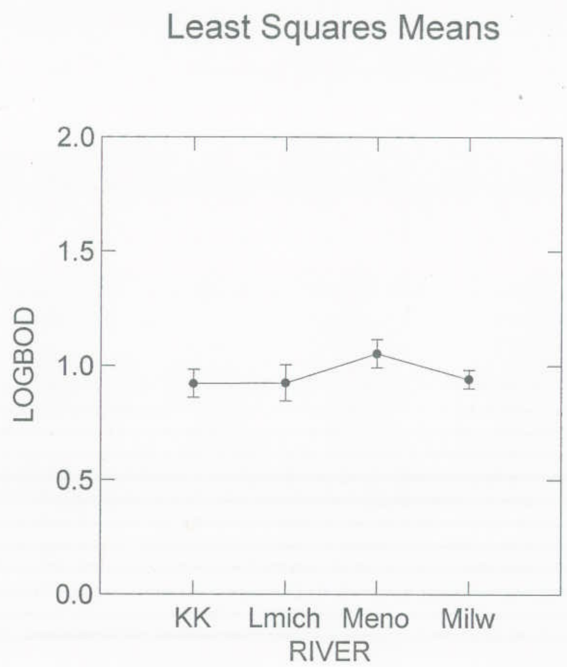

```
*** WARNING ***
Case           77 is an outlier        (Studentized Residual =       -5.0727)

Durbin-Watson D Statistic         2.0290
First Order Autocorrelation      -0.0159
```

Section II

MMSD CSO ANOVAS by watershed
CT5/6 had been removed from these analyses

```
Data for the following results were selected according to:
      (SITECODE= 1)

Effects coding used for categorical variables in model.

Categorical values encountered during processing are:
RIVER$ (4 levels)
   KK, Lmich, Meno, Milw

Dep Var: LOGBOD   N: 311   Multiple R: 0.0993   Squared multiple R: 0.0099

Analysis of Variance
```

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
RIVER$	0.7302	3	0.2434	1.0183	0.3848
Error	73.3801	307	0.2390		




```
*** WARNING ***
Case          74 is an outlier        (Studentized Residual =      -4.3574)
Case          85 is an outlier        (Studentized Residual =      -4.0614)
Case         100 is an outlier        (Studentized Residual =      -4.0614)
Case         117 is an outlier        (Studentized Residual =      -4.0614)
Case         137 is an outlier        (Studentized Residual =      -4.0614)
Case         150 is an outlier        (Studentized Residual =      -4.0920)
Case         189 is an outlier        (Studentized Residual =      -4.0859)
Case         261 is an outlier        (Studentized Residual =      -4.0859)
Case         276 is an outlier        (Studentized Residual =      -4.0859)
Case         324 is an outlier        (Studentized Residual =      -4.0859)

Durbin-Watson D Statistic          2.0499
First Order Autocorrelation       -0.0251
COL/
ROW RIVER$
  1  KK
  2  Lmich
  3  Meno
  4  Milw
Using least squares means.
Post Hoc test of LOGBOD
Using model MSE of 0.239 with 307 df.
```

Matrix of pairwise mean differences:

	1	2	3	4
1	0.0000			
2	0.0032	0.0000		
3	0.1316	0.1285	0.0000	
4	0.0196	0.0165	-0.1120	0.0000

Bonferroni Adjustment.

Matrix of pairwise comparison probabilities:

	1	2	3	4
1	1.0000			
2	1.0000	1.0000		
3	0.7913	1.0000	1.0000	
4	1.0000	1.0000	0.7884	1.0000

Data for the following results were selected according to:
(SITECODE= 1)

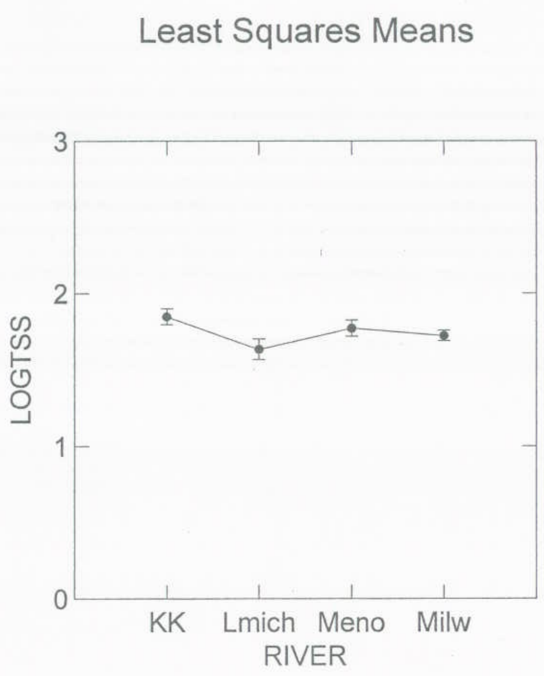
Effects coding used for categorical variables in model.

Categorical values encountered during processing are:
RIVER$ (4 levels)
KK, Lmich, Meno, Milw
1 case(s) deleted due to missing data.

Dep Var: LOGTSS N: 310 Multiple R: 0.1499 Squared multiple R: 0.0225

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
RIVER$	1.2532	3	0.4177	2.3434	0.0731
Error	54.5449	306	0.1783		



Durbin-Watson D Statistic 1.5157
First Order Autocorrelation 0.2406
COL/
ROW RIVER$
1 KK
2 Lmich
3 Meno

4 Milw
Using least squares means.
Post Hoc test of LOGTSS
Using model MSE of 0.178 with 306 df.

Matrix of pairwise mean differences:

	1	2	3	4
1	0.0000			
2	-0.2138	0.0000		
3	-0.0761	0.1377	0.0000	
4	-0.1242	0.0895	-0.0481	0.0000

Bonferroni Adjustment.

Matrix of pairwise comparison probabilities:

	1	2	3	4
1	1.0000			
2	0.0837	1.0000		
3	1.0000	0.6871	1.0000	
4	0.3034	1.0000	1.0000	1.0000

Data for the following results were selected according to:
(SITECODE= 1)

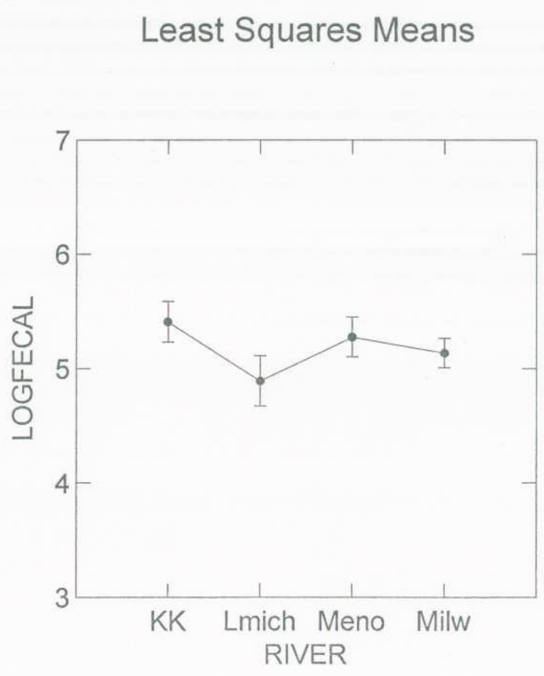
Effects coding used for categorical variables in model.

Categorical values encountered during processing are:
RIVER$ (4 levels)
KK, Lmich, Meno, Milw
240 case(s) deleted due to missing data.

Dep Var: LOGFECAL N: 71 Multiple R: 0.2303 Squared multiple R: 0.0531

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
RIVER$	1.8115	3	0.6038	1.2512	0.2983
Error	32.3345	67	0.4826		



*** WARNING ***
Case 151 is an outlier (Studentized Residual = -3.8080)

Durbin-Watson D Statistic 1.9775

First Order Autocorrelation 0.0092
COL/
ROW RIVER$
1 KK
2 Lmich
3 Meno
4 Milw
Using least squares means.
Post Hoc test of LOGFECAL
Using model MSE of 0.483 with 67 df.

Matrix of pairwise mean differences:

	1	2	3	4
1	0.0000			
2	-0.5168	0.0000		
3	-0.1327	0.3841	0.0000	
4	-0.2726	0.2442	-0.1399	0.0000

Bonferroni Adjustment.

Matrix of pairwise comparison probabilities:

	1	2	3	4
1	1.0000			
2	0.4373	1.0000		
3	1.0000	1.0000	1.0000	
4	1.0000	1.0000	1.0000	1.0000

Data for the following results were selected according to:
(SITECODE= 1)

Effects coding used for categorical variables in model.

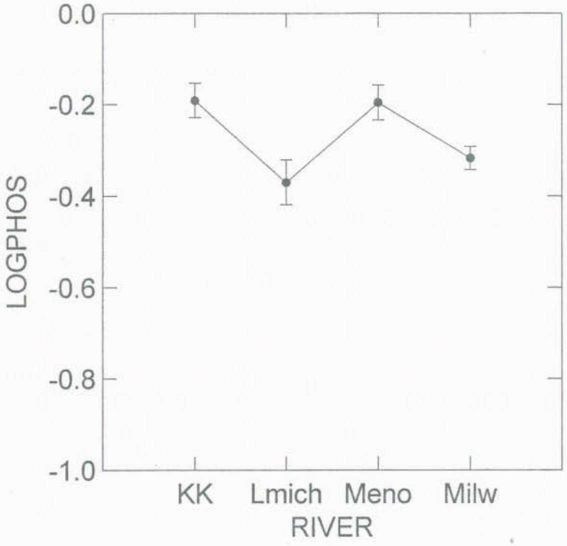
Categorical values encountered during processing are:
RIVER$ (4 levels)
KK, Lmich, Meno, Milw
26 case(s) deleted due to missing data.

Dep Var: LOGPHOS N: 285 Multiple R: 0.2310 Squared multiple R: 0.0534

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
RIVER$	1.3358	3	0.4453	5.2790	0.0015
Error	23.7023	281	0.0843		

Least Squares Means




```
*** WARNING ***
Case           77 is an outlier        (Studentized Residual =       -5.4860)

Durbin-Watson D Statistic          1.8735
First Order Autocorrelation        0.0602
COL/
ROW RIVER$
  1  KK
  2  Lmich
  3  Meno
  4  Milw
Using least squares means.
Post Hoc test of LOGPHOS
Using model MSE of 0.084 with 281 df.
```

Matrix of pairwise mean differences:

	1	2	3	4
1	0.0000			
2	-0.1793	0.0000		
3	-0.0046	0.1746	0.0000	
4	-0.1259	0.0534	-0.1212	0.0000

Bonferroni Adjustment.

Matrix of pairwise comparison probabilities:

	1	2	3	4
1	1.0000			
2	0.0240	1.0000		
3	1.0000	0.0319	1.0000	
4	0.0344	1.0000	0.0511	1.0000

Section III

MMSD SSO Jones Island influent regression analysis between BOD versus Ammonia

>IMPORT "I:\ENVA\WORK\Tom's Folder\RWQMP-Update\Data\JI bod & nh3.XLS" / TYPE=EXCEL,SHEET=1
IMPORT successfully completed.

>REGRESS
>MODEL INJI_AMM = CONSTANT+INJI_BOD

>ESTIMATE

```
Dep Var: INJI_AMM    N: 27    Multiple R: 0.951    Squared multiple R: 0.905

Adjusted squared multiple R: 0.901    Standard error of estimate: 1.245
```

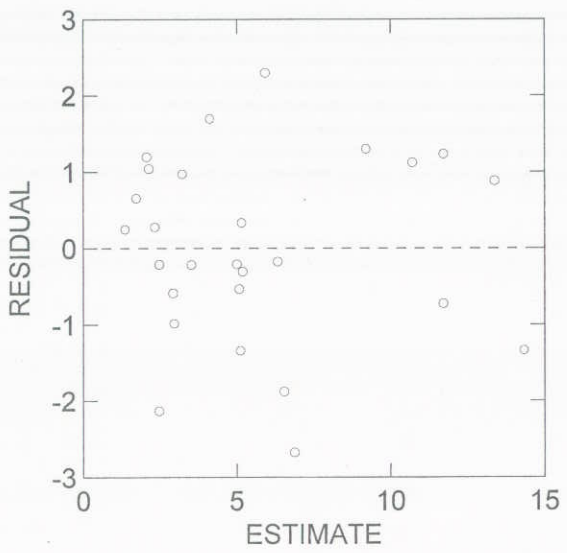
Effect	Coefficient	Std Error	Std Coef	Tolerance	t	P(2 Tail)
CONSTANT	-0.589	0.472	0.000	.	-1.249	0.223
INJI_BOD	0.037	0.002	0.951	1.000	15.435	0.000

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
Regression	369.447	1	369.447	238.242	0.000
Residual	38.768	25	1.551		

```
Durbin-Watson D Statistic      1.332
First Order Autocorrelation    0.220
```

Plot of Residuals against Predicted Values



Section IV

Ohio River Valley Water Sanitation Commission (ORVWSC) regression analysis between BOD versus Ammonia and BOD versus Organic Nitrogen.

```
Dep Var: NH3   N: 147   Multiple R: 0.7842   Squared multiple R: 0.6149

Adjusted squared multiple R: 0.6123   Standard error of estimate: 1.0454
```

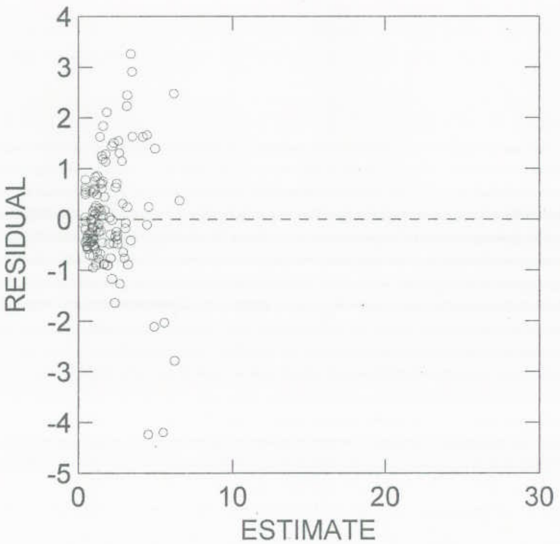
Effect	Coefficient	Std Error	Std Coef	Tolerance	t	P(2 Tail)
CONSTANT	0.4728	0.1196	0.0000	.	3.9543	0.0001
CBOD5	0.0264	0.0017	0.7842	1.0000	15.2163	0.0000

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
Regression	253.0547	1	253.0547	231.5346	0.0000
Residual	158.4771	145	1.0929		

```
*** WARNING ***
Case        116 has large leverage   (Leverage =        0.0985)
Case        157 is an outlier        (Studentized Residual =     -4.3802)
Case        208 is an outlier        (Studentized Residual =     -4.3962)
Case        209 has large leverage   (Leverage =        3.0612)

Durbin-Watson D Statistic        1.1458
First Order Autocorrelation      0.4241
```

Plot of Residuals against Predicted Values



Dep Var: TOTALORGANI   N: 138   Multiple R: 0.8023   Squared multiple R: 0.6437

Adjusted squared multiple R: 0.6410   Standard error of estimate: 3.4446

Effect	Coefficient	Std Error	Std Coef	Tolerance	t	P(2 Tail)
CONSTANT	0.4313	0.4179	0.0000	.	1.0320	0.3039
CBOD5	0.0922	0.0059	0.8023	1.0000	15.6733	0.0000

Analysis of Variance

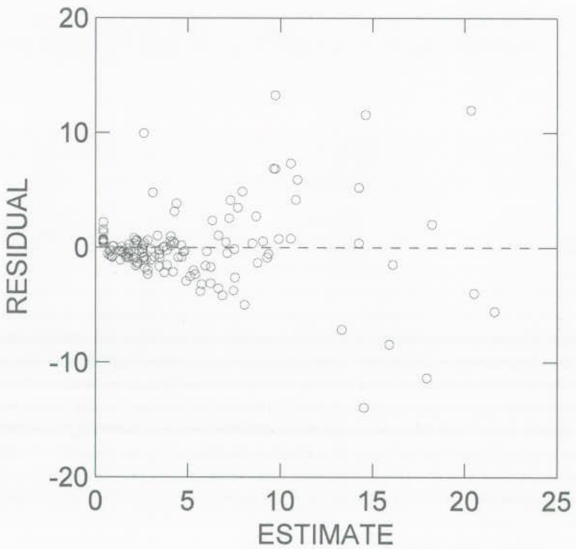
Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
Regression	2914.6684	1	2914.6684	245.6531	0.0000
Residual	1613.6369	136	11.8650		

*** WARNING ***

```
Case       90 is an outlier       (Studentized Residual =      3.8232)
Case       92 is an outlier       (Studentized Residual =      4.0998)
Case      116 has large leverage  (Leverage =        0.1011)
Case      157 is an outlier       (Studentized Residual =     -4.3858)
Case      168 is an outlier       (Studentized Residual =      3.5744)
Case      208 is an outlier       (Studentized Residual =     -3.5464)
```

Durbin-Watson D Statistic   1.2893
First Order Autocorrelation   0.3551

Plot of Residuals against Predicted Values



MMSD CSO ANOVAS BY SITE
These analyses include all sites

SYSTAT Rectangular file C:\Data\MMSD CSO Sampling Data 1994-2002 Stat Data.SYD,
created Thu Jan 29, 2004 at 13:56:01, contains variables:

SITE$	SITECODE	RIVER$	DATE	BOD	LOGBOD
TSS	LOGTSS	FECAL	LOGFECAL	PHOS	LOGPHOS

Effects coding used for categorical variables in model.

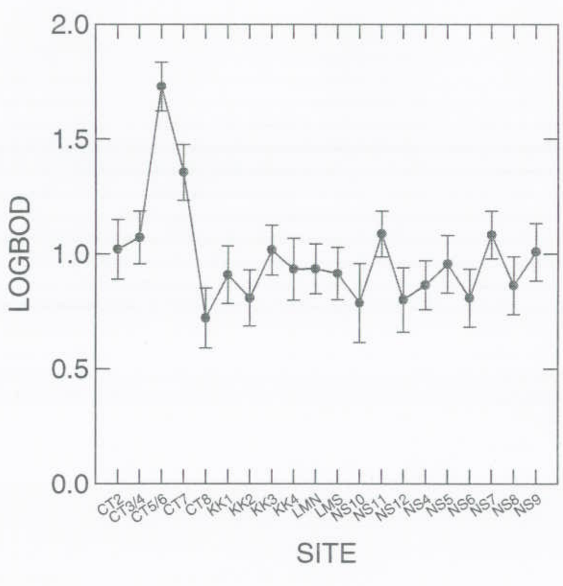
Categorical values encountered during processing are:
SITE$ (20 levels)
CT2, CT3/4, CT5/6, CT7, CT8, KK1, KK2, KK3, KK4, LMN, LMS, NS10, NS11, NS12,
NS4, NS5, NS6, NS7, NS8, NS9

Dep Var: LOGBOD   N: 332   Multiple R: 0.4406   Squared multiple R: 0.1942

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
SITE$	17.7589	19	0.9347	3.9563	0.0000
Error	73.7110	312	0.2363		

Least Squares Means




```
*** WARNING ***
Case         85 is an outlier       (Studentized Residual =     -4.1699)
Case        100 is an outlier       (Studentized Residual =     -3.9296)
Case        117 is an outlier       (Studentized Residual =     -4.3784)
Case        137 is an outlier       (Studentized Residual =     -4.2518)
Case        150 is an outlier       (Studentized Residual =     -4.1915)
Case        189 is an outlier       (Studentized Residual =     -4.0239)
Case        261 is an outlier       (Studentized Residual =     -4.0633)
Case        276 is an outlier       (Studentized Residual =     -4.4004)
Case        324 is an outlier       (Studentized Residual =     -3.9581)

Durbin-Watson D Statistic          2.1855
```

First Order Autocorrelation -0.0933

Pairwise comparisons identify site CT5/6 as being different from all other sites except CT7.

Effects coding used for categorical variables in model.

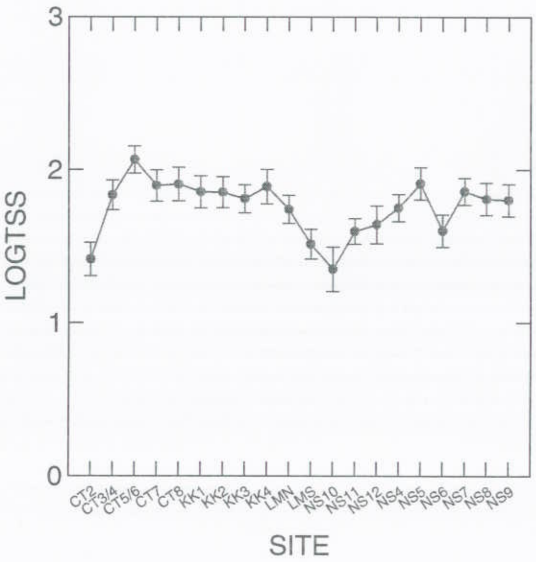
Categorical values encountered during processing are:
SITE$ (20 levels)
CT2, CT3/4, CT5/6, CT7, CT8, KK1, KK2, KK3, KK4, LMN, LMS, NS10, NS11, NS12,
NS4, NS5, NS6, NS7, NS8, NS9
1 case(s) deleted due to missing data.

Dep Var: LOGTSS N: 331 Multiple R: 0.3853 Squared multiple R: 0.1484

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
SITE$	9.0639	19	0.4770	2.8532	0.0001
Error	51.9989	311	0.1672		

Least Squares Means



Durbin-Watson D Statistic 1.7069
First Order Autocorrelation 0.1426
Effects coding used for categorical variables in model.

Pairwise comparisons identify site CT5/6 as being different from sites LMS, NS10, and NS11.

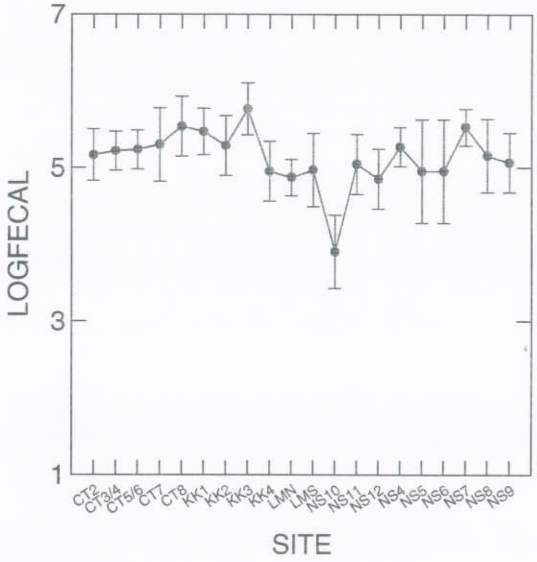
Categorical values encountered during processing are:
SITE$ (20 levels)
CT2, CT3/4, CT5/6, CT7, CT8, KK1, KK2, KK3, KK4, LMN, LMS, NS10, NS11, NS12,
NS4, NS5, NS6, NS7, NS8, NS9
254 case(s) deleted due to missing data.

Dep Var: LOGFECAL N: 78 Multiple R: 0.4833 Squared multiple R: 0.2335

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
SITE$	8.0566	19	0.4240	0.9301	0.5509
Error	26.4420	58	0.4559		

Least Squares Means




```
*** WARNING ***
Case          151 is an outlier        (Studentized Residual =       -4.0421)
```

Pairwise comparisons are irrelevant as the ANOVA detected no significant differences among sites.

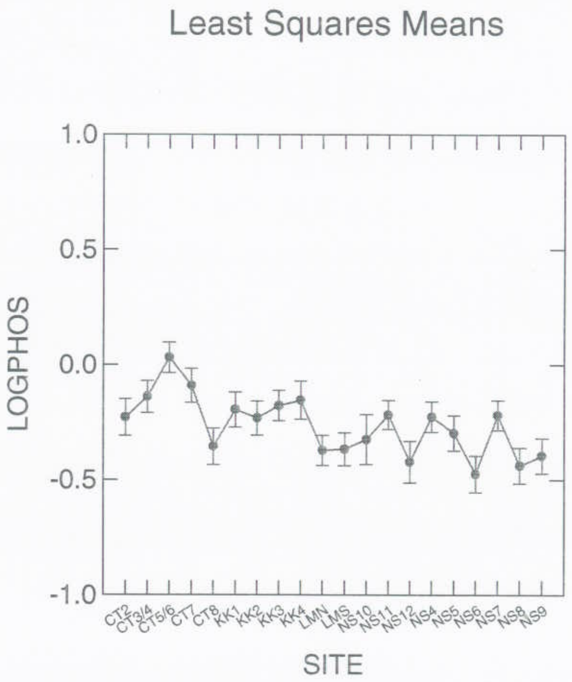
Effects coding used for categorical variables in model.

Categorical values encountered during processing are:
SITE$ (20 levels)
CT2, CT3/4, CT5/6, CT7, CT8, KK1, KK2, KK3, KK4, LMN, LMS, NS10, NS11, NS12,
NS4, NS5, NS6, NS7, NS8, NS9
28 case(s) deleted due to missing data.

Dep Var: LOGPHOS N: 304 Multiple R: 0.4146 Squared multiple R: 0.1719

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
SITE$	4.8636	19	0.2560	3.1028	0.0000
Error	23.4299	284	0.0825		




```
*** WARNING ***
Case            77 is an outlier        (Studentized Residual =      -5.0760)

Durbin-Watson D Statistic          2.0316
First Order Autocorrelation       -0.0175
```

Pairwise comparisons identify site CT5/6 as being different from sites LMN, LMS, NS12, NS6, NS8, and NS9.

Exhibit C to Appendix D

#91452 V1 - EXHIBIT C
[300-4003]
[JEB]
[2-05-2004]

Exhibit C

Table C-1
MMSD Collectors Sewershed Areas and Population

Watershed	Collector	Area (acres)	Population
Menomonee River	CT 2	339	5,345
	CT ¾	2,547	46,159
	CT 5/6	2,324	46,383
	CT 7	818	18,918
	CT 8	550	5,847
Kinnickinnic River	KK 1	618	17,105
	KK 2	126	1,677
	KK 3	1,460	33,205
	KK 4	66	873
Lake Michigan	LMN	480	5,849
	LMS	396	6058
Milwaukee River	NS 4	930	15,622
	NS 5	359	5,482
	NS 6	832	12,177
	NS 7	1,670	30,857
	NS 8	692	8,482
	NS 9	690	9,638
	NS 10	358	616
	NS 11	497	5,653
	NS 12	167	1,794
Total		15,919	277,740
Mean		795.95	13,887

Source: Triad Engineering and SEWRPC.

Exhibit D to Appendix D

Test	**Linear regression**
	Log SSO Data by Parameter and Sample Data
Fit	Log BOD v Date
Performed by	Jeremy Nitka

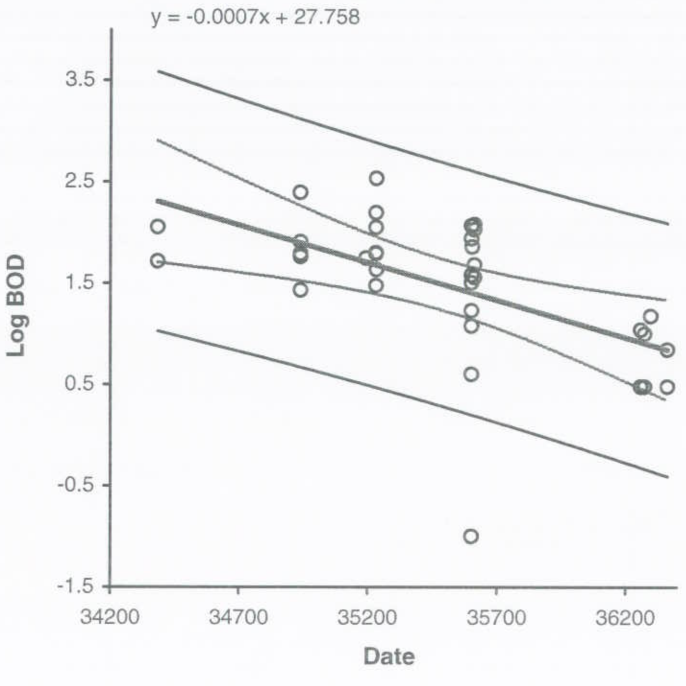
Date 30 January 2004

n	35 (cases excluded: 1 due to missing values)

R^2	0.31
Adjusted R^2	0.29
SE	0.5834

Term	Coefficient	SE	p	95% CI of Coefficient
Intercept	27.7584	6.8042	0.0003	13.9152 to 41.6015
Slope	-0.0007	0.0002	0.0005	-0.0011 to -0.0004

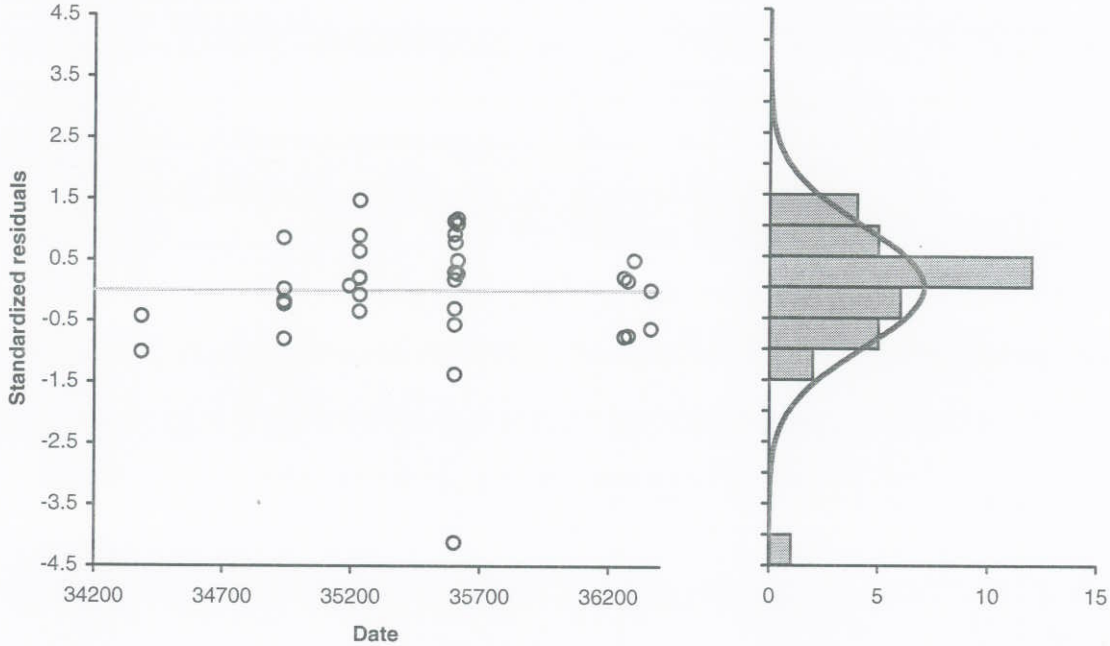
Source of variation	SSq	DF	MSq	F	p
Due to regression	5.076	1	5.076	14.91	0.0005
About regression	11.233	33	0.340		
Total	16.310	34			



Test	**Linear regression**
	Log SSO Data by Parameter and Sample Data
Fit	Log BOD v Date
Performed by	Jeremy Nitka
Date	30 January 2004

Test	**Linear regression**
	Log SSO Data by Parameter and Sample Data
Fit	Log BOD v Date
Performed by	Jeremy Nitka

Date 30 January 2004



analysed with: Analyse-it + General 1.68

Test	**Linear regression**
	Log SSO Data by Parameter and Sample Data
Fit	Log TSS v Date
Performed by	Jeremy Nitka

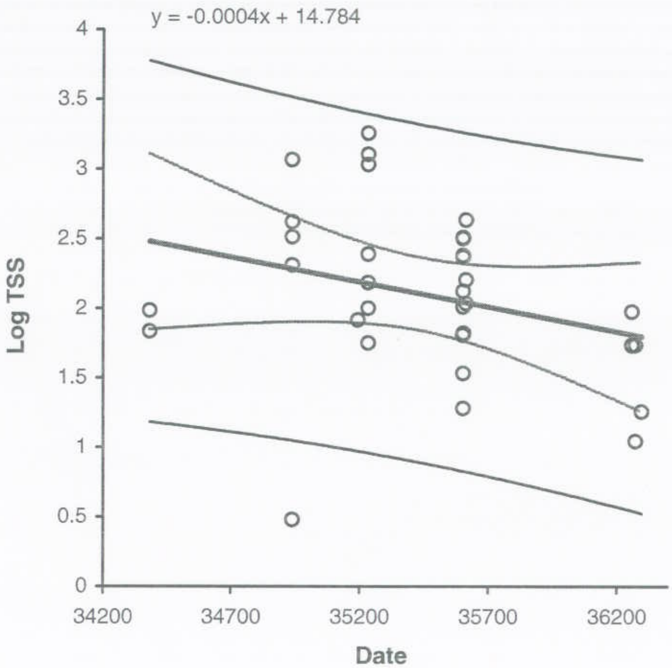
Date 30 January 2004

n	33 (cases excluded: 3 due to missing values)

R^2	0.08
Adjusted R^2	0.05
SE	0.5866

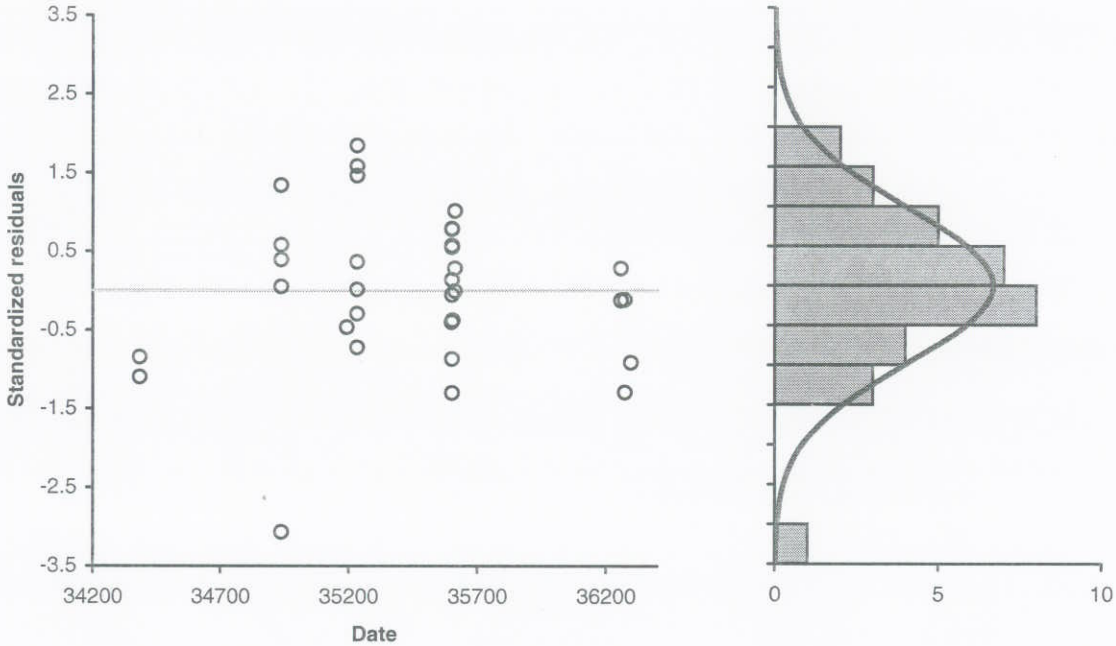
Term	Coefficient	SE	p	95% CI of Coefficient
Intercept	14.7842	7.5135	0.0581	-0.5397 to 30.1080
Slope	-0.0004	0.0002	0.1014	-0.0008 to 0.0001

Source of variation	SSq	DF	MSq	F	p
Due to regression	0.981	1	0.981	2.85	0.1014
About regression	10.668	31	0.344		
Total	11.649	32			



Test	**Linear regression**
	Log SSO Data by Parameter and Sample Data
Fit	Log TSS v Date
Performed by	Jeremy Nitka
Date	30 January 2004

Test	**Linear regression**
	Log SSO Data by Parameter and Sample Data
Fit	Log TSS v Date
Performed by	Jeremy Nitka
Date	30 January 2004



Test	**Linear regression**
	Log SSO Data by Parameter and Sample Data
Fit	Log Fecal v Date
Performed by	Jeremy Nitka

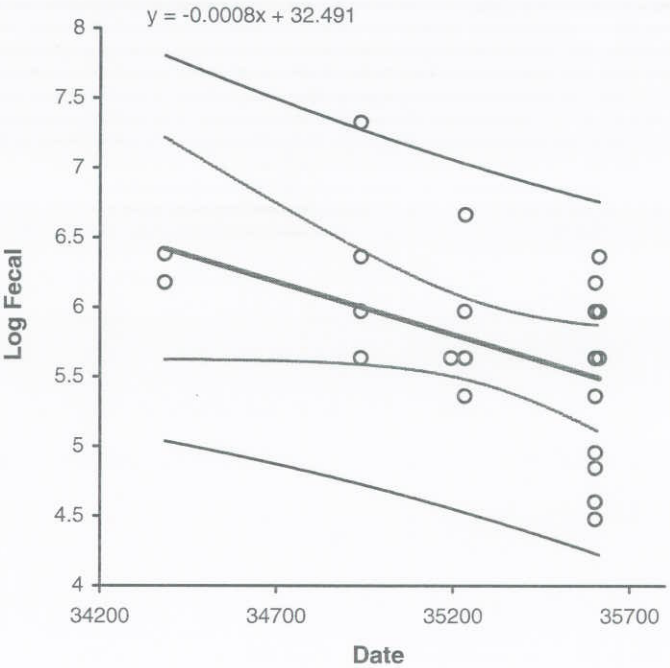
Date 30 January 2004

n	29 (cases excluded: 7 due to missing values)

R^2	0.18
Adjusted R^2	0.15
SE	0.5994

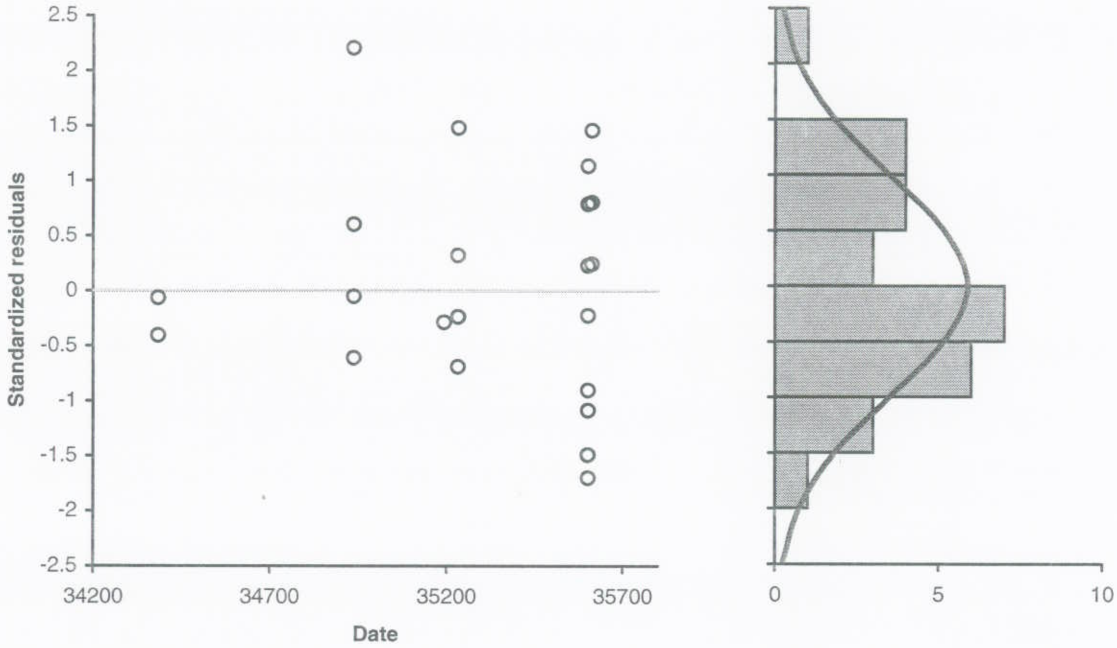
Term	Coefficient	SE	p	95% CI of Coefficient
Intercept	32.4907	11.0413	0.0066	9.8358 to 55.1456
Slope	-0.0008	0.0003	0.0223	-0.0014 to -0.0001

Source of variation	SSq	DF	MSq	F	p
Due to regression	2.112	1	2.112	5.88	0.0223
About regression	9.702	27	0.359		
Total	11.813	28			



Test	**Linear regression**
	Log SSO Data by Parameter and Sample Data
Fit	Log Fecal v Date
Performed by	Jeremy Nitka
Date	30 January 2004

Test	**Linear regression**
	Log SSO Data by Parameter and Sample Data
Fit	Log Fecal v Date
Performed by	Jeremy Nitka
Date	30 January 2004



analysed with: Analyse-it + General 1.68

Test	**Linear regression**
	Log SSO Data by Parameter and Sample Data
Fit	Log Total P v Date
Performed by	Jeremy Nitka

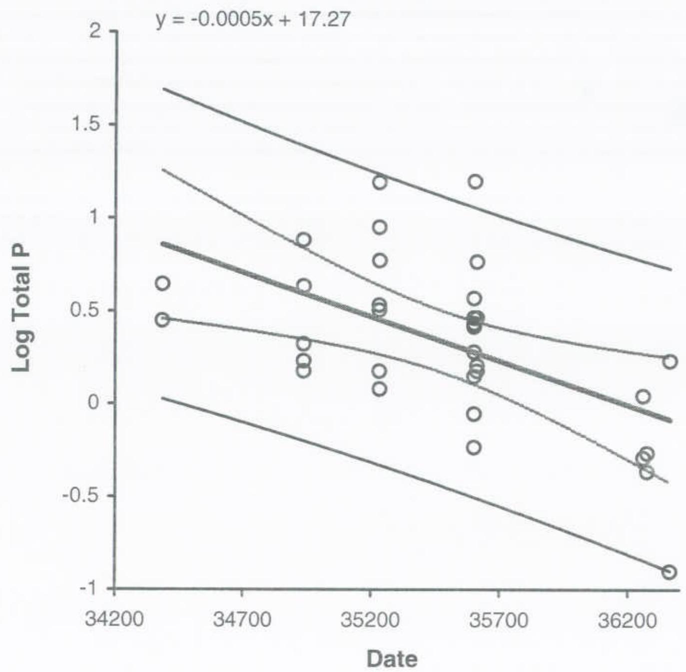
Date 30 January 2004

n	34 (cases excluded: 2 due to missing values)

R^2	0.30
Adjusted R^2	0.28
SE	0.3777

Term	Coefficient	SE	p	95% CI of Coefficient
Intercept	17.2702	4.5872	0.0007	7.9265 to 26.6140
Slope	-0.0005	0.0001	0.0008	-0.0007 to -0.0002

Source of variation	SSq	DF	MSq	F	p
Due to regression	1.945	1	1.945	13.63	0.0008
About regression	4.565	32	0.143		
Total	6.509	33			



analysed with: Analyse-it + General 1.68

Test	**Linear regression**
	Log SSO Data by Parameter and Sample Data
Fit	Log Total P v Date
Performed by	Jeremy Nitka

Date 30 January 2004