



IN REPLY REFER TO:

# United States Department of the Interior

U.S. GEOLOGICAL SURVEY

Box 25046 M.S. 407

Denver Federal Center

Denver, Colorado 80225

## **NATIONAL WATER QUALITY LABORATORY TECHNICAL MEMORANDUM 2000-01S**

January 6, 2000

To: Gary Cottrell

From: Jeff Pritt

Subject: Analysis and interpretation of reporting environmental data with a non-proper calibration curve for Kjeldahl nitrogen and phosphorus methods

### **Summary**

For the time period between April 20, and October 4, 1999, samples were analyzed by a modified calibration model. An additional low concentration standard was added that improperly eliminated the blank standard from the calibration model. It was initially thought that the blank was critical in the calibration model, by introducing potentially a 0.01 and 0.05 mg/L high bias in results for phosphorus and nitrogen, respectively. Additional data analysis described here, shows that the non-standard calibration model was still a good model to determine the phosphorus and nitrogen concentrations in environmental samples. The correct 8-point calibration model with the blank standard was developed and all original sample data were processed through the calculation, giving paired results. The variances observed in the data set would be caused entirely by the original non-standard calibration model and the corrected 8-point calibration model, since the original instrument signal for the samples was used. The larger issue is "should the NWIS and LIMS databases be updated"? In conclusion, the differences between the calibration models is well within the routinely observed method variability with a median bias of 0 mg/L for phosphorus and 0.002 mg/L for nitrogen. A majority of differences calculated in the original result and the 8-point calibration result is within the LT-MDL and LRL for each method. Results in this range are already in a low confidence for quantitation region and data are reported with the estimated "E" attached to each result, reflecting the higher relative variability for measurement.

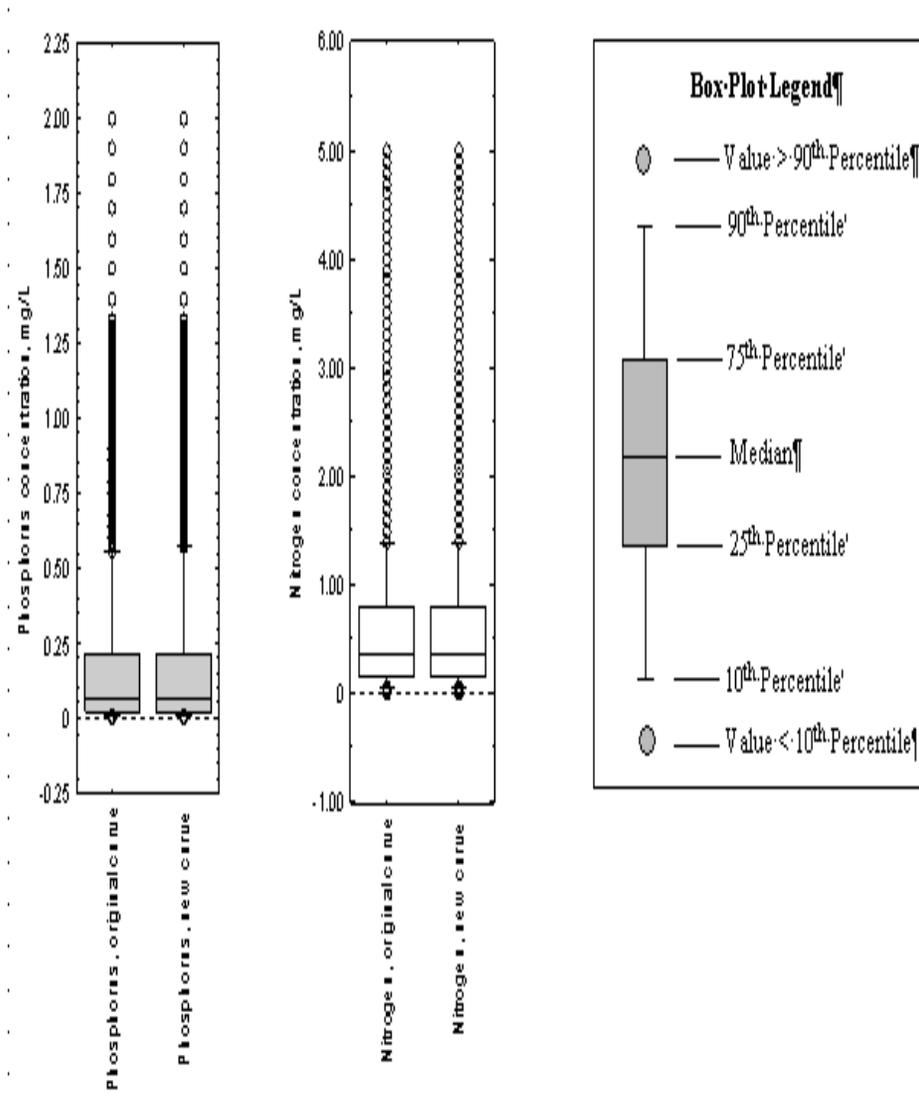
### **Description of the Data Set**

Gary Cottrell prepared the data set for analysis from all the analytical runs associated with the time period of the nonstandard calibration curve. The data supplied were environmental data determined from the original curve and the correct 8-point calibration curve. For this analysis the data within the calibration range were used. All environmental results greater than 2 mg/L for phosphorus were not

used for the phosphorus analysis. Similarly, all environmental results greater than 5 mg/L for nitrogen were not used for the nitrogen analysis. The data set size for phosphorus is 14,582 pairs, and for nitrogen is 15,660 pairs. The data were further adjusted from unrounded, unprocessed to correctly rounded using the ASTM standard guide. Generally, 2 or 3 significant figures were retained in the processed results. Negative concentrations were not used, since that data would be reported as less than the LRL.

## Data Analysis

Box plots of the original and recalculated data are shown below. Ninety percent of the environmental data are below approximately 0.06 mg/L for phosphorus and approximately 1.5 mg/L for nitrogen. Since the data are not normally distributed or symmetric, a paired sign test was used to test for significance. This sign test null hypothesis states the median is 0 and the alternative hypothesis states the median is not 0.



The sign test shows significance for both sets of data. The conclusion that can be drawn is there is enough evidence to show that the medians are different between the original and the recalculated results for the data. The sign test results are tabulated below.

**Paired Sign Test for Phosphorus**

# Differences > 0	6692
# Differences < 0	2677
# Differences = 0	5213
P-Value	<.0001

208 cases were omitted due to missing values.

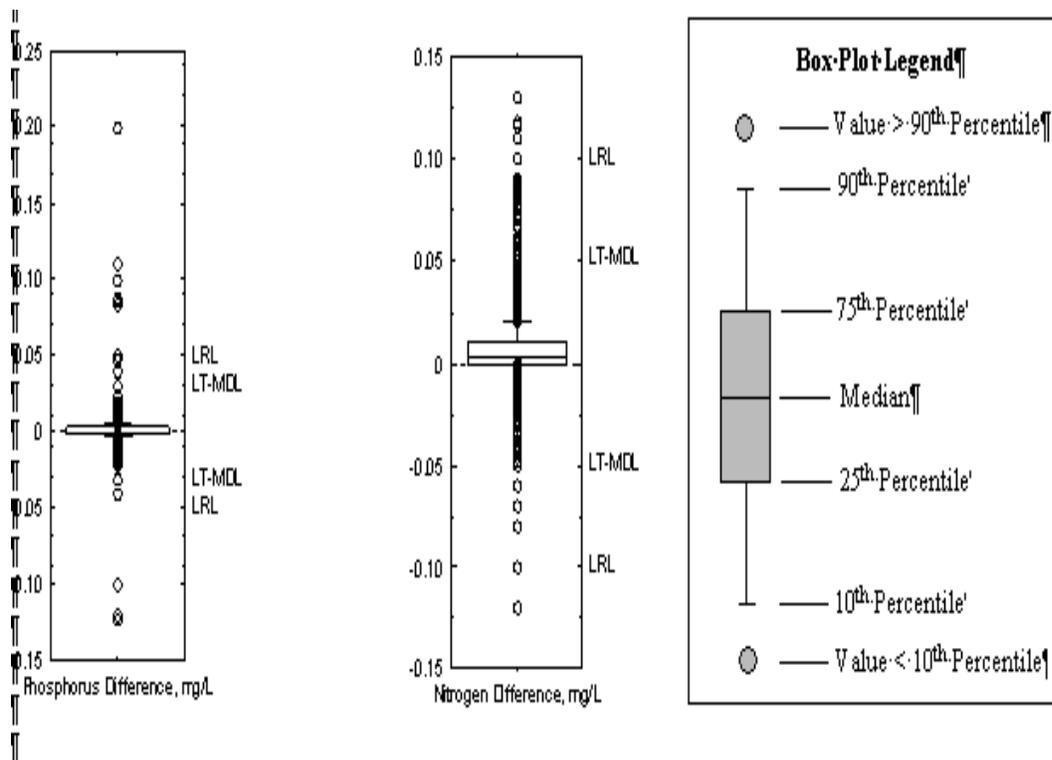
**Paired Sign Test for Nitrogen**

# Differences > 0	7985
# Differences < 0	1367
# Differences = 0	6308
P-Value	<.0001

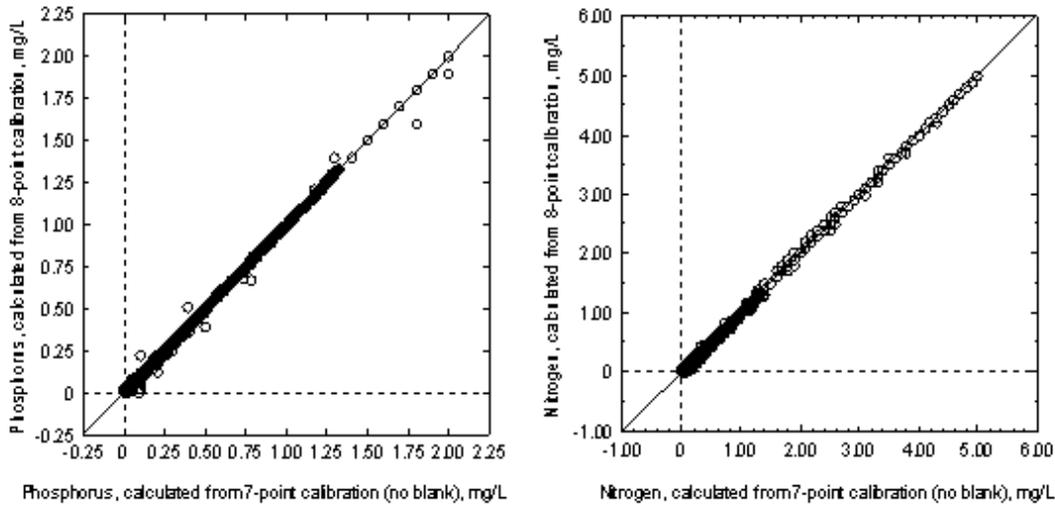
77 cases were omitted due to missing values.

In large data sets it is not uncommon to see very small but significant differences. Considering the expected variability of the method, a box plot of the differences is shown to determine if the statistical significance is valid. In both cases, the differences are biased high, confirming the omission of the blank in the original calibration curve did have a slight impact, but nowhere near the estimated 0.01 mg/L for phosphorus or 0.05 mg/L for nitrogen.

The box plots of the differences show the majority of differences between results are less than the LT-MDL (long-term method detection level) and the LRL (laboratory reporting level). The LT-MDL is statistically determined from many measurements of a low concentration standard submitted blindly for analysis. The LT-MDL is strictly a measure of the variability of the method at very low concentrations and confirmed by blank data. The policy of the NWQL is to report all data below the LRL with and "E" indicating the presence of large variability relative to the concentration reported.



As a final comparison, scatter plots, comparing the non-standard calibration curve and the 8-point calibration curve, and regression analysis show just how small the error is associated with the incorrect calibration curve.



The regression summary for these plots are shown below.

### Phosphorus Regression Summary

Regression Summary

Count	14582
Num. Missing	208
R	.999859319
R Squared	.999718658
Adjusted R Squared	.999718639
RMS Residual	.005242848

Regression Coefficients

	Coefficient	Std. Error	Std. Coeff.	t-Value	P-Value
Intercept	-.0017688647	.000051244	-.0017688647	-34.514232629	<.0001
Phosphorus, original curve	1.001604557	.000139154	.999859319	7197.809162723	<.0001

ANOVA Table

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	1424.082368489	1424.082368489	5.180845674E7	<.0001
Residual	14580	.400767022	.000027487		
Total	14581	1424.483135511			

## Nitrogen Regression Summary

1

### Regression Summary

Count	15660
Num. Missing	77
R	.999849565
R Squared	.999699153
Adjusted R Squared	.999699134
RMS Residual	.011946840

### Regression Coefficients

	Coefficient	Std. Error	Std. Coeff.	t-Value	P-Value
Intercept	-.008600572	.000127718	-.008600572	-67.340101997	<.0001
Nitrogen, original curve	1.004050678	.000139196	.999849565	7213.237750895	<.0001

### ANOVA Table

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	7426.199707028	7426.199707028	5.203079885E7	<.0001
Residual	15658	2.234819330	.000142727		
Total	15659	7428.434526357			