TETRA TECH, INC.
4401 Building, Suite 200
79 T.W. Alexander Drive
P.O. Box 14409
Research Triangle Park, NC 27709
Telephone: (919) 485-8278 Telefax: (919) 485-8280

MEMORANDUM

To: Ed Garvey (TAMS/NJ) **Date:** February 17, 1998

From: J. B. Butcher **Project:** Hudson

Subject: Low Res vs CC DN-50/Mean Phi stats **Pjn:** 1182-06

NUMBER OF PAGES: 5 Original will follow as e-mail attachment

Purpose

This memo examines the relationship between DN50 and Mean Phi in the Low Resolution and CC data sets, and examines the data for consistency, or lack of consistency in the relationships. There are two related types of questions we can ask: (1) Are the low resolution results consistent with the high resolution results, and do the low resolution results confirm the statistical models derived from high resolution results?; and (2) Are the two sets of data drawn from the same population?

The analyses were completed using the spreadsheet supplied by you. Only minor changes were necessary, plus some additions. I noted that the subscript bounds were too large (by one) for calculating the regression coefficients for the LR and LR+CC data sets. Also, a minor correction was made in the calculation of $s_{y.x}$.

DN-50 is predicted as a linear model of Mean Phi. Coefficient estimates are given below

	CC Data	Low Resolution	Pooled Data
Intercept	74.13	84.26	76.23
Slope	-6.741	-10.01	-7.804
R^2 (percent)	53.7	60.8	58.7

The R^2 coefficients indicate that the models are not particularly strong predictors. It is instructive to examine a plot of DN50 versus Mean Phi for both sets of data, shown in Figure 1 along with regression lines. There is considerable scatter in both distributions (thus the low R^2), but the relationships appear similar for Mean Phi greater than about 3. Below Phi of 3 the two data sets diverge, with the center of mass of the DN50s from the LR data above that for the CC data. Most notably, there are many more data points for Mean Phi less than 2 in the CC data than in the LR data.

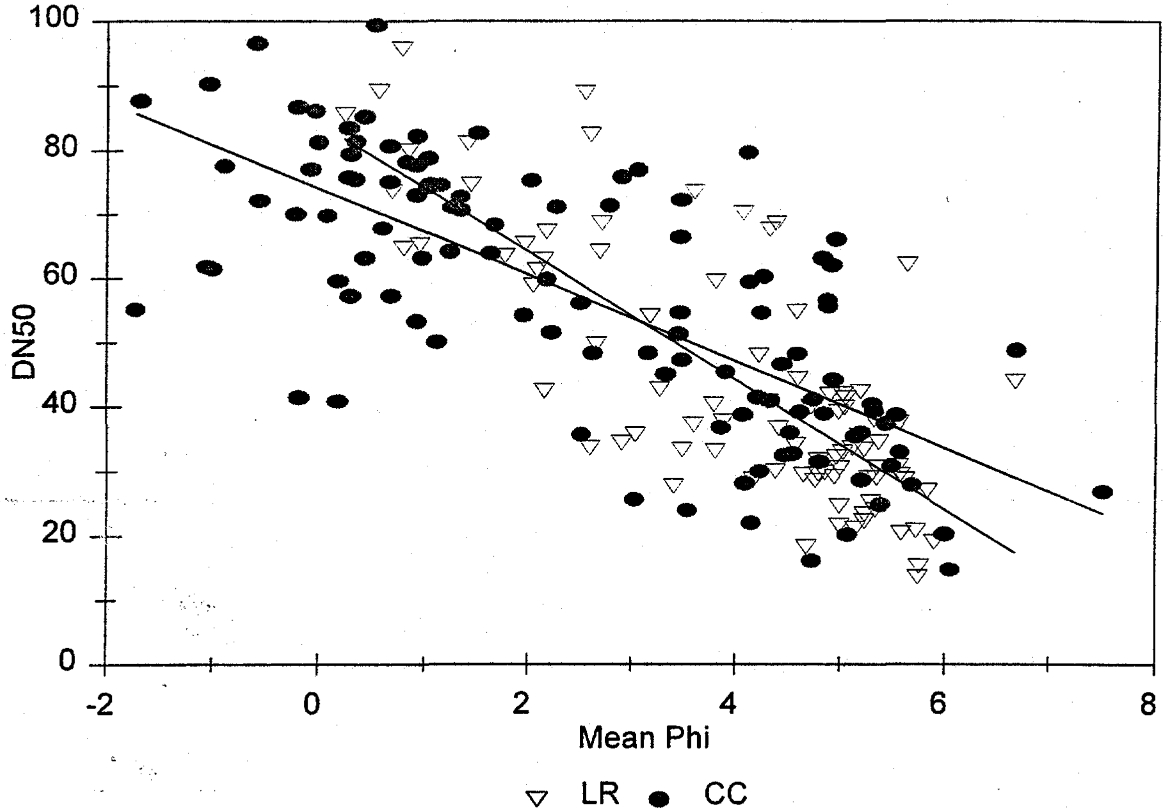

Figure 1. DN50 versus Mean Phi for LR and CC Data Sets

Does this represent a systematic difference in the sampling process? The difference in the lower end of the distribution results in a noticeable difference in the regression lines for the two data sets.

1. Calculate an R^2 predicting results from one data set with the model from the other

This test applies one data set's fitted model coefficients to predict the other data set's results. Quality of fit is summarized by the R^2 value, which represents the percent of total variation explained by the regression model. To test performance of using one model on the other set of data, we manually compute an R^2 as 1 - SSE/SST. If this R^2 is not much below the R^2 obtained in the original "native" regression for the data set, the two models are functionally similar.

This comparison was run in both directions. When the LR data are predicted using the CC model, the calculated R^2 is 52.0, which can be compared to the native LR model R^2 of 60.8. When the CC data are predicted with the LR model, the calculated R^2 is 40.4, which can be compared to the native CC model R^2 of 53.7. In either case, switching models results in a moderate decline in predictive ability.

2. Proportion within confidence bounds

This is an informal, qualitative test, in which we examine the proportion of Low Resolution data which fall within the 95% confidence bounds on the CC model. The confidence bounds used here are those for the prediction interval, the error in predicting specific realizations, from the CC regression. The position of the low resolution data points within the CC model prediction limits is shown below. 88 our of 89 points (98.9%) fall within the prediction limits. The Low Resolution data are thus not inconsistent with the CC model, even though the estimates of coefficients may differ. The 95% prediction limits are broad because of the relatively low predictive power of the CC model.

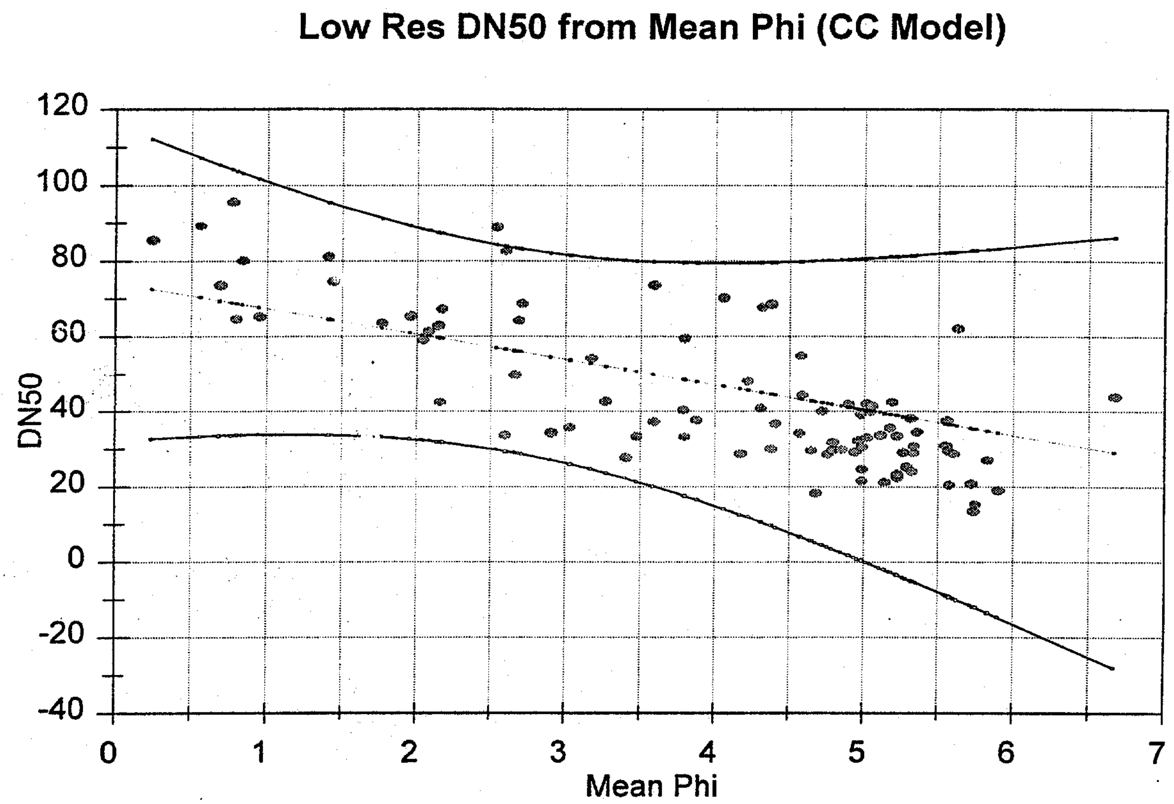

Figure 2. Low Resolution Data and 95% Prediction Intervals from CC Model

3. Regress predicted on observed Low Resolution values and test coefficients

For this test we want Low Resolution values predicted with the CC regression coefficients. We can then test whether the intercept is significantly different from 0 and the slope significantly different from 1. The graph presented above suggests there is a slight bias in the relationship of the Low Resolution data to the CC regression line. A regression of Low Resolution predictions (with the CC model) on Low Resolution observations provides a fit of about the same quality as the native models, with R^2 of 60.8. The intercept of -25.81 is statistically different from 0 at the 95% confidence level (confidence interval -38.03 to -13.59) and the slope of 1.485 is statistically different from 1 (confidence interval 1.231 to 1.739). This suggests a systematic change in the relationship between Phi and DN50 between the two data sets, consistent with the distribution of low Phi observations seen in Figure 1.

4. Theil's U Statistic

Theil's U statistic gives a measure of the consistency between forecasts (e.g., Low Resolution predictions using the High Resolution model) and the data used to develop the forecasts. It ranges from 0 to 1, with 0 indicating perfect prediction, and it's variance can be approximated (for U less than 0.3) as U^2/T, where T is the number of samples in the "forecast".

The U statistic may also be decomposed into portions attributed to bias (U^m), variance (U^s), and covariance (U^c). When U is non-zero, we would ideally like the decomposition to show that the difference is entirely attributable to the covariance component, which represents non-controllable random variability. Weight on the bias component indicates that the linear relationship differs between the two data sets. Weight on the variance component indicates that the difference is attributable primarily to differing variances between the two data sets.

U	0.14
Var (U)	0.00235
Lower 95% Confidence	0.112
U^m	0.05
U^s	0.47
U^c	0.49

For the prediction of Low Resolution DN50 from Mean Phi using the CC model, results are shown in the box. The U statistic is significantly different from zero, indicating the presence of some difference between the data sets. The decomposition shows that there is a small bias. Most of the difference, however, is attributed, in approximately equal amounts, to the variance and covariance portions. The large weight on the variance portion indicates that the decline in fit in applying the CC model to the Low Resolution data is partly due to differing variance between the Low Resolution and CC data. In fact, the variances of DN50 observations in the two data sets are nearly equal (407.8 for CC data, 392.1 for Low Res data); however, the variance in the phi data is more than twice as large for the CC data (4.82) than for the Low Res data (2.37). Thus the Low Resolution data achieves the same amount of variance in DN50 over a smaller range of Phi values compared to the CC model.

5. *Chow's F Test*

Chow's F test addresses the hypotheses that the parameters have or have not changed between the two data sets. It is developed by calculating SSEs for regressions on each of the data sets individually and an SSE for a regression on the pooled data. The comparison is made by forming an F statistic with k and t_1+t_2-2k degrees of freedom, formed as

$$F = \frac{[SSE(constrained) - SSE(unconstrained)]/k}{SSE(unconstrained)/(t_1 + t_2 - 2k)}$$

in which *SSE(unconstrained)* is the sum of the *SSE*s from the two separate regressions, *SSE(constrained)* is the *SSE* from the regression on the pooled data, t_1 is the number of observations in the first sample set, t_2 is the number of observations inn the second sample set, and k is the number of parameters. The resulting statistic can then be compared to a tabulation of the F distribution.

In this case, $k = 2$, because both a slope and an intercept are calculated. The resulting F statistic is 4.722, which has a probability value under the null hypothesis of no change in parameters of 0.0093. Thus, the F test indicates that the change in parameter values between the High Resolution and Low

Resolution regressions is statistically significant at the 95% level, matching the results seen in item 3.

Summary

Individual DN50 versus Mean Phi observations in the Low Resolution data set are not inconsistent with the CC data set. However, the two data sets produce significantly different model coefficient estimates. The difference in models may arise from a truncation of low Phi values in the Low Resolution data.