

Wastewater Management Validation and Design Committee
Date Originated 19 November 2008

Subject: Summary Chapter VI (Water Quality Modeling)

Key Issues:

1. The adequacy of water quality sampling and the data determined from the samples.
2. How the data has been used. Specifically as to:
 - a. The use of annual averages and the unavailability of some of the raw data. For instance, the use of annual averages rather than actual sample data tends to reduce the variability and, thus, enhances the impression that the data fit the model.
 - b. How have outliers been handled? Have they been excluded? Why?
 - c. The raw data should be examined.
 - d. No confidence levels are provided for the data or the model's ability to predict nitrogen levels in Pleasant Bay.
3. There are large standard deviations for the nitrogen data. Large standard deviations lead to large uncertainties in the outcomes that would be predicted by any model.
4. It is not clear whether or not the model fits the water quality data. In other words, it is not clear that the model works as a predictor of nitrogen concentrations in Pleasant Bay.
5. Dispersion coefficients of the model do not appear to account for gyric circulation in Pleasant Bay.

Tasks to Be Undertaken to Address the Key Issues:

Task 1. Obtain copies of all raw field and analytical water quality data, and have statistical analysis performed.

Task 2. Statistically analyze the impact of the uncertainties on the prediction of nitrogen concentrations in Pleasant Bay and the sub-embayments and the prediction for compliance with improved water quality by the proposed sewerage program.

Task 3. Engage a consultant with expertise in finite difference modeling to evaluate the SMAST model described in the Pleasant Bay Report.

Expertise and Experience Required to Complete the Tasks:

1. Expertise in statistical analysis of experimental data.
2. Expertise in construction and use of finite difference models.

APPENDIX

Commentary:

Description of Model Development

Water quality modeling as described in Chapter VI brings together the following:

1. the hydrodynamic model (Chapter V) to describe the movement of the water in the Pleasant Bay system;
2. anthropogenic and natural nitrogen loadings as estimated for the watersheds in Chapter IV;
3. nitrogen loadings from rainfall on the surface of the Pleasant Bay system waters;
4. nitrogen loadings from "benthic flux" (nitrogen exchange taking place between the embayment sediments and the water column above it);
5. background nitrogen concentration in the Atlantic Ocean water entering Pleasant Bay;
6. field nitrogen concentration measurements from a sampling program during the 1995-2004 period.

Based on the data input sources identified above, the water quality model was "calibrated" by modifying the dispersion coefficients used in the model until the model output of nitrogen concentrations matched the measured nitrogen concentrations.

The purpose of the water quality modeling is to realistically simulate the behavior of bioactive nitrogen (DIN + PON) in the Pleasant Bay system in response to the inputs (nitrogen loads going into the system) and the flushing action of the tidal flow from the Atlantic (taking nitrogen out of the system).

Hydrodynamic Model as a Basis

Reference to the Chapter V summary document developed by the Orleans Wastewater Management Validation and Design Committee should be made for a complete evaluation of the issues raised regarding the hydrodynamic model. The two primary issues identified in the Chapter V summary document were: (1) a question as to the adequacy of the hydrodynamic model to describe mixing taking place in the Pleasant Bay system and directly relating to the "selection" of dispersion coefficients used in the model; (2) the hydrodynamic modeling in the Pleasant Bay report does not account for the improved flushing provided by the 2007 North Breach.

It seems unlikely that the water quality model could provide a realistic simulation of bioactive nitrogen behavior in the Pleasant Bay system without adequately accounting for the mixing taking place due to the currents and flows and without considering the impact of the significantly increased dilution and flushing provided by the North Breach.

Pleasant Bay System – Measured Nitrogen Concentrations

Data Reliability

Questions Raised Concerning Data Reliability:

1. Were all of the field measurements and collection of samples conducted in accordance with protocols established in a DEP approved Quality Assurance Project Plan (QAPP)? A site-specific DEP approved QAPP for Orleans was not in place during the years covered in this report.
2. In a committee discussion of the sampling program, a citizen participant in the sampling program stated that sampling locations were not consistent from year

to year and sometimes sampling buoys were moved during the year. Is it possible that inconsistencies in results may not be due to changes in nitrogen concentrations but may, instead, be due to changes in where the samples were collected?

3. Orleans is not in possession of some of the raw field and analytical data for some of the monitoring years covered in the Pleasant Bay Report, so it could not be reviewed. In addition, the model calibration was done with multi-year averages of nitrogen concentrations and many of the standard deviations are quite large. Large standard deviations in the data employed would lead to large uncertainties in the results.

The report does not discuss confidence levels for the original data or for the multi-year average data used to calibrate the model. Knowing the confidence levels for the data and determining the confidence level for the model results are critical to understanding how well the model describes nitrogen concentrations in Pleasant Bay.

There are questions concerning confidence levels and data reliability in the Pleasant Bay report. Data, estimates of parameters and results of models are presented as absolute numbers without any analysis of their reliability, margins of error or confidence levels.

Nitrogen Concentrations at Build Out

In the last section of Chapter VI, the model is used to forecast the impact of “build out” on nitrogen concentrations.

However, the report provides no description or assumptions for the “build-out” scenario. What is the basis for “build-out”?

Using the build-out scenario data from Figure VI-7 and the no anthropogenic load data from Figure VI-8 plus the nitrogen threshold concentration of 0.21 mg/L, shows that the no anthropogenic load nitrogen concentrations for Meetinghouse, Areys and Pah Wah, for example, are nearly equal to or are greater than the nitrogen threshold concentration established in Chapter VII. The no anthropogenic load nitrogen concentration for Meetinghouse and Areys Ponds exceeds the nitrogen threshold concentration. If the sub-embayment exceeds the nitrogen threshold concentration targeted, no nitrogen remediation program will achieve the target. This could mean that either the nitrogen threshold concentration is not appropriate (as it appears to be unachievable) or that there is something wrong with the nitrogen sample data or the model itself.

Reviewer Questions:

VI.1.3 : p 129

Measured Nitrogen Concentrations in Embayments. “In order to create a model that realistically simulates bioactive nitrogen (DIN+PON) concentrations in a system in response to the existing flushing conditions and loadings, it is necessary to calibrate the model to actual measurements of water column nitrogen concentrations. The refined and approved data for each monitoring station used in the water quality modeling effort are presented in Table VI-1. Station locations are indicated in the area map presented in Figure VI-1. The multi-year averages present the “best” comparison to the water quality model output, since factors of tide, temperature and rainfall may exert short-term influences on the individual sampling dates and even cause inter-annual differences. Three years of baseline field data are the minimum required to provide a baseline for

MEP analysis. Typically, ten years of data (collected between 1995 and 2004) were available for stations monitored by in Pleasant Bay."

Not all earlier data was not collected under a DEP approved QAPP. In addition, some analyses may have not been performed by a certified laboratory, or analyses may have been performed with different analytical methods or detection limits, and therefore may not have been accepted by DEP for the purpose of TMDL development. These facts call into question the reliability of the data. How can samples collected and analyzed under circumstances be relied upon as a basis for the water quality model?

Why do the multi-year averages present the "best" comparison? What happens if all values are included in the comparison?

Table VI-1 (pg. 130):

The table lists Total and Bioactive Nitrogen means and Standard deviations. Many nitrogen standard deviations are very large. This would lead to large uncertainties in the predicted outcomes. This should be part of the analysis of the tolerance of the prediction of the model.

The DIN species may include ammonia, urea, and nitrates, and possibly others. Are these species individually measured? Of the DIN and DON, are there measurements that show the relative proportion of inorganic and organic nitrogen? Would it show the relationship between sources from (1) people and (2) from natural decay from vegetation and marine life?

VI.2. Model Description (pg. 132)

Para. 2. Why is total nitrogen used as a "non-conservative constituent, where bottom sediments act as a source or sink of nitrogen..."? Does the ratio of DIN to [DON + PON] vary according to the characteristics of the bottom surface?

Para. 2. Has the model been used for winter conditions? If no, why not? Has SMAST considered year-around material balances in its calculations?

Para. 3. Is it always correct that there is vertical mixing, with no vertical composition gradients?

What about periods of light or no winds? Are there areas in the Pleasant Bay System where vertical composition gradients may exist for significant periods of time (days)?

Para. 4. Is there only one set of dispersion coefficients that provide a solution to the inputs and boundary conditions, or are there multiple combinations? In other words, is there a unique solution to the mathematical computations?

VI.2.1. Model Formulation.

The Differential Equation needs to be understood if the model is to be believed. Many terms (i.e. dispersion coefficients) that are selected by the user are derived from the calibration process. It states that the model dispersion coefficients are adjusted to fit the data. What if the nitrogen inputs are incorrect? Then the calculations and results could be misleading.

At each time step the model computes Nitrogen over the entire grid and utilizes a continuity of mass equation to check these results. Question: Is this a Mass Balance calculation?

Pg. 133. Para. 1. If the model was operated with "total nitrogen" inputs, how did the RMA-4 model predict the "bio-active" nitrogen concentrations?

VI.2.2. Water Quality Model Setup (pg. 133).

Describes that an initial N concentration is applied to the entire model domain. Model was then run for a simulated 28 day spin-up period. Then the model was run for a five day tidal period. Question: Why and for what reason?

Para. 3. What is a "spin-up period"?

VI. 2.3 Boundary Condition Specification.

Para. 5. In the watershed loads, why were the contributions from "Natural Background Watershed Loads" [Pleasant Bay Report, Table ES-1a, Page ES-11.] not included? See Table VI-2, page 134. Or were they included?

Para. 5. Why would the sediment in some areas of Pleasant Bay "take up", or absorb" nitrogen during the summer weeks? What happens to this nitrogen long term?

Para. 6. The report states that the concentration of "bio-active" nitrogen in the Atlantic Ocean was 0.094 mg/liter in 2005. Were any other measurements made at any other time? Shouldn't measurements have been made in at least one or two other years?

Para. 6. Were measurements made of "total nitrogen" in the Atlantic Ocean, since this was the basis for the input data? In more than just 2005?

Does the model allow for the counterclockwise gyric in Pleasant Bay and Little Pleasant Bay described by Dr. Gregory Horne? How would this motion affect the dispersion coefficients? And the computational results?

Nitrogen loads for each sub-embayment watershed were evenly distributed at grid cells that form the perimeter of the sub-embayment. Question: Was attenuation by the fringing marsh applied in the model?

Nitrogen loads were distributed across the embaymentbased on core samples and applied to the entire surface of the embayment. Question : Where were the cores taken? Is it reasonable to apply the flux for the sample core to the entire embayment??

Does the benthic flux vary with distance from the center of the pond or embayment?

VI.2.4 Model Calibration p134 ¶ 1

"Generally, the relatively quiescent main basin of Pleasant Bay and its tributary sub-embayment systems required values of E that are lower compared to the riverine estuary systems evaluated by Fischer, *et al.*, (1979). Observed values of E in these calmer areas typically range between order 10 and order 0.001 m²/sec (USACE, 2001)."

Re: footnoted reference: "Water quality indicators are usually numerically modeled using mass balance equation in steady or unsteady conditions. The mass balance equation contains accumulation, loading by streams, rivers, outfalls, outflow, reaction and settling. Dynamic mass balance equation in the form of Advective-Diffusion Equation (ADE) for 3D, 2D and 1D framework in unsteady condition is the basic equation in modeling water quality indicators. The accuracy of the solution to the numerical scheme, which is usually used to solve the ADE, is highly dependent upon the accuracy of longitudinal Dispersion coefficients (DI). It has been established that many hydrodynamic parameters affect the longitudinal dispersion coefficient, particularly for natural streams where the dispersion characteristics may vary greatly from one stream to another."¹

Note that the dispersion coefficient for plug flow is zero ($E=0$). This indicates (low E values used) that the model for the main basin of Pleasant Bay has it in a very quiet state (lacking significant current flow and mixing). As noted in reference 1, "The accuracy of the solution to the numerical scheme, which is usually used to solve the Advective-Diffusion Equation (ADE), is highly dependent upon the accuracy of longitudinal Dispersion coefficients (E values)."

Figures VI-2 (p137) and VIII-1 (p216) seem to confirm this.

Is Pleasant Bay quiescent? Are there not significant currents producing mixing, at least sometimes? If there were, higher dispersion coefficients would be needed to reflect the currents in Pleasant Bay.

VI.2.4 Model Calibration p134

Para. 1. "Calibration of the bio-active nitrogen model of Pleasant Bay proceeded by changing model dispersion coefficients so that the model output of nitrogen concentrations matched measured data." What does this mean?

Para. 2. Figure VI-2. How is the concentration for "bio-active" nitrogen, DIN+PON, derived from the total nitrogen input data? What is the conversion? Are the concentrations of oxygen in the model elements considered in the calculations?

Table VI-2 (pg. 134) shows that the total contribution of nitrogen (total N) from the direct atmospheric deposition (85.69 kg/day) and from benthic flux (184.519 kg/day) is 2.1 times the contribution from the watershed load of 127.203 kg/day. Aren't the contributions from septic systems, fertilizer, etc., relatively small?

Para. 3. Figure VI-3 (pg. 138). What does "yearly data" mean? Were measurements made throughout a year? Or does it mean the "average of several years"?

Why is the one sigma measured data variance so large, especially for: Meetinghouse Pond, Upper Muddy River, and Frost Fish [CM-14]?

The larger variance tends to occur with the larger average values of "DIN+PON".

If the data represent the average of measurements made over several years, they would not describe trends in the concentrations, either increases or decreases. Aren't trends, if they occur, important?

¹ Zedah, A. A. T. and Kashefipour, S. M., Department of Hydraulic Structures, Shahid Chamran University, Ahwaz, Iran, "Effect of Dispersion Coefficient in Surface Water Quality Modeling"

VI.2.5 Model Salinity Verification (pg. 139)

Para. 1. Table VI-4 (pg. 140). How were the fresh water flows determined? Are they assumed to be the same for every year? How do they vary with annual rainfall and with municipal well production?

Para. 1. The average annual rainfall is 27.25 inches net infiltration into undeveloped land [42 inches rainfall – 14.75 inches evapo-transpiration; or (44.0 – 16.75)?]². This should be the same for water bodies since there is evaporation from these surfaces as well.

VI.2.6.1 Build-Out p 143-144

The first paragraph jumps right into conclusions of significant increases in nitrogen loads due to build-out. The impact is all about the assumptions used and there is no hint of how this was generated. What assumptions were used? Are there decision choices that could be made to lessen the impact?

Pg. 144 bottom of page: Again “nitrogen recycling”, also called “benthic flux net” appears...and it is a very significant part of the total nitrogen loading. We need to understand this: the calculations (it is not clear what some of the terms mean); the analytical methods used and the mechanism of benthic flux. It is key to develop a nitrogen mass balance (at least around selected sub-embayments as examples).

Taken from Tables: VI-7 and VI-9

Data for bioactive N (DIN+PON) mg/L

Sub-embayment	No Anthropogenic Load	DEP Target	Septic	Present	Buildout
Meetinghse PBA16	0.233	0.21		0.380	0.441
Meetinghse WMO 10	0.184	0.21		0.261	0.296
Areys Pond	0.214	0.21		0.297	0.345
Pah Wah Pond	0.181	0.21		0.257	0.304

Forgetting for a moment that we need to ground this work with explanations for how the above numbers are arrived at; assumptions going into the model, the mechanism and analysis related to benthic nitrogen flux, etc, it is not clear from the above as to what approach needs to taken to the problem. In the Meetinghouse Pond case it seems eliminating all man made nitrogen inputs (all septage, all fertilizer, etc) will not get us to the target. For the Arey's and Pah Wah, the targets are very close to the “no load” scenario.

Para. 1. Figure VI-7 (pg. 143). The model shows very poor correlation with field measurements for Upper and Lower Muddy Creek. Why?

Para. 1. Figure VI-7. The variance on the yearly measurements is significantly large, especially at: Upper Pochet, Upper and Lower Muddy Creek, and lower Frost Fish. Why?

² Pleasant Bay Report, pg 36, table IV-4, suggests recharge rate from impervious surfaces at 40 inches per year. But only 90 percent of precipitation reaches groundwater, so precipitation equals 44 inches per year.

Para. 2. "Loads are presented in kilograms per day (kg/day) in this Section, since it is inappropriate to show benthic flux loads in kilograms per year due to seasonal variability." This needs to be explained. Does this mean that the calculations and results are useful for only a few days per year (in the summer)?

Para. 2. Table VI-5. Is the purpose of the "no anthropogenic load" scenario to represent an ideal situation of complete sewers in all watersheds contributing to the Pleasant Bay system? Or alternatively, no human presence in the watersheds?

Buildout shown in Table VI-5 (pg. 144), e.g., Meetinghouse Pond +33% . Question, is this correct?

Comment : "Benthic flux is assumed to vary proportionally to the watershed load." Seems reasonable on the surface but assumes that the benthic flux is caused by anthropogenic loading. This assumption seems to ignore flux caused by vegetative and marine decay.

Table VI-6 (pg. 145). Why are the "net benthic fluxes" for Lower Muddy Creek, Pleasant Bay/Chatham Harbor Channel, Bassing Harbor and Chatham Harbor negative? What were the results for the cores taken from these areas? Were these negative values derived from the model calculations?

The benthic flux values in Table VI-6 are slightly higher than in Table VI-2. What was the methodology for determining the higher values in Table VI-6?

VI.2.6.2 No Anthropogenic Load (pg. 148)

Para. 1. Table VI-8 (pg. 148). The benthic flux for Meetinghouse Pond is shown as 7.40 kg/day, compared to 14.365 kg/day in Table VI-2. Does that mean that the difference, 6.695 kg/day of total nitrogen (46.6%), derives from vegetative and marine matter decay? How was this determined? Is this calculation satisfied by an annual material balance?

For Arey's Pond, the number is 3.70 kg/day in the "no-load" case, compared to 5.669 kg/day in Table VI-2, or more than half. This means that in the current day scenario, in Arey's Pond, the contribution from the sediments during a summer month is 3.70 kg/day (61.7%) from natural matter and 2.30 kg/day from (38.3%) anthropogenic sources. How was this determined?

"The benthic flux load was reduced toward zero based on the reduction in the watershed load discussed I VI.2.6.1" Question : How was the benthic flux calculated?

Pg. 149. Table VI-9 (pg. 149).

The table presents data as DIN+PON. How were the data obtained from "total nitrogen"?

The following is statistical work by Mark Fiegel

Review of Chapter VI
Massachusetts Estuaries Project
Final Report – May 2006

By Mark Fiegel

"The overall result of the sensitivity analysis is that the Linked Model predictions of embayment nitrogen level and distribution appear to be relatively robust. The Model is most sensitive to (in the following order of most to least sensitive):

- Water column dispersion
- source water nitrogen concentration
- benthic regeneration, septic load
- attenuation, fertilizer, impermeable surfaces."³

Page 132 VI.2.1 Model Formulation.

"The formulation of the model is for two-dimensional depth-averaged systems in which concentration in the vertical direction is assumed uniform. The depth-averaged assumption is justified since vertical mixing by wind and tidal processes prevent stratification in the modeled sub-embayment."

Is it reasonable to make this assumption given the model is most sensitive to water column dispersion? Should this assumption not have been tested?

Page 129 VI.1 Data Sources for the Model:

"Several different data types and calculation are required to support the water quality modeling effort for the Pleasant Bay system. These include the output from the hydrodynamics model, calculations of external nitrogen loads from the watersheds, measurements of internal nitrogen loads from the sediment (benthic flux), and measurements of nitrogen in the water column."

How good is this data, do we have any confidence intervals (levels) for the data? Models are only as good as the input. In the discussion I can find no reference to confidence intervals. Should we have some idea regarding the confidence levels?

Page 133 VI.2.3 Boundary Condition Specification:

"The boundary concentration Atlantic Ocean region offshore Pleasant Bay was set at 0.094 mg/L, based on Chatham data collected in the summer of 2005 and analyzed by SMAST." This is the mean data for PBA-17A for the period summers of 2000 through 2005 in Table VI-1. It is also the same as the model min for PBA-17 in Table VI-1. Is this really reasonable? I ask this question because it seems really coincidental that the two measurements are the same.

Regarding Boundary Condition Specification, there are no boundary conditions set for the East side of Pleasant Bay. As I understand boundary conditions for dynamic models, they are a set of values for state variables and their rates along problem domain boundaries, sufficient to determine the state of the system within the problem domain. Given what Dr. Horne has stated, is this reasonable?

Reference is to Figure VI-4 on page 139:

³ Page V of Executive Summary of "Nitrogen Modeling to Support Watershed Management: Comparison of Approaches and Sensitivity Analysis" By Brian C. Howes, and John S. Ramsey from October 2000 – October 2001

This figure is replicated (MiniTab) with data in the report (specifically Table VI-1):

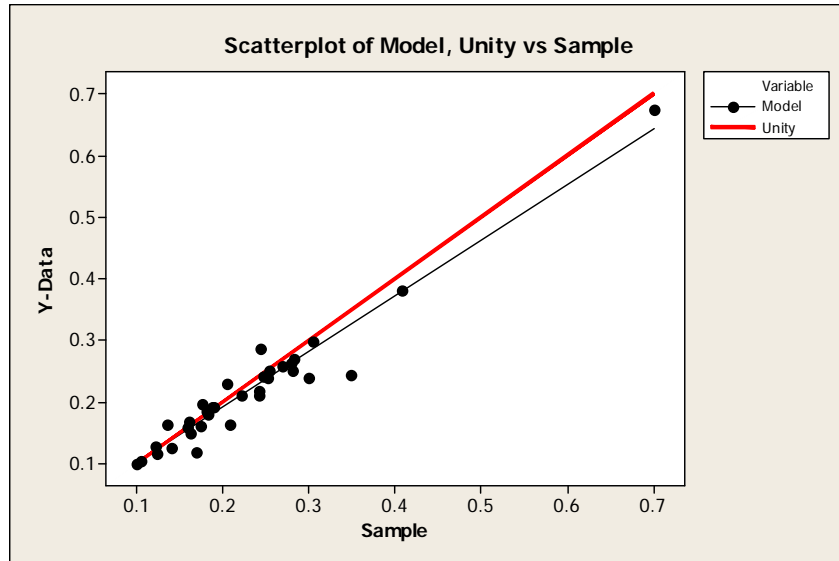


Figure 1

Two observations, first the black line is the linear regression and the R^2 : 0.962 (note the correlation coefficient should be r not R^2) refers to the strength of the linear fit of the data, and not how well the model simulates reality. Also, the outlier (in the upper right corner) of the graph is the mean of Muddy Creek Upper (PBA-05A) which has an extremely large standard deviation. The question: what is the effect of this sampling site on the model, and how do the modelers handle this outlier? Does the dynamic model just accept this outlier and accept that it may be an influential point. I didn't find any discussion in the report regarding this data point.

Why wasn't all the data plotted or used, not just the mean of the data? The measured concentration is an average (mean) of the sampled data; thus, the statistic (mean) is less variable than the data on which they are based, so it tends to inflate the impression of the strength of the relationship. In other words, the high correlation coefficient may be overstated.

A 95% confidence interval from the data provided in the report is delineated in the graph below:

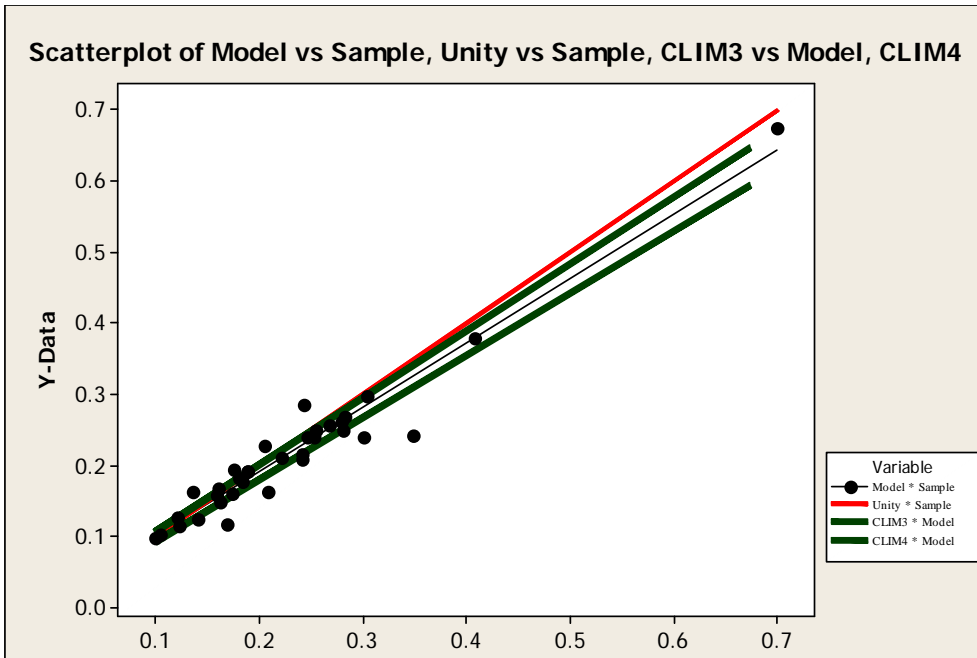


Figure 2

There are a number of data points outside the 95% confidence interval. What impact does this imply about the model? Further what does it mean in terms our solutions suggested at the end of the report?

I generated 25 random samples using the mean and standard deviation for Muddy Creek Upper (PBA-05A). I assumed the distribution was approximately normal. I used the new mean to generate a graph like the one above (Figure 1).

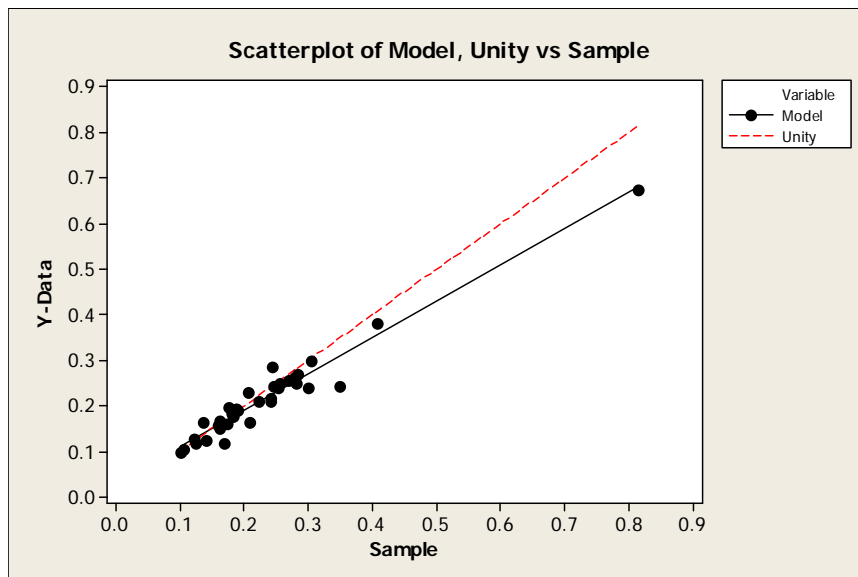


Figure 3

The correlation coefficient is higher using this data ($r = 0.973$), but obviously the model isn't as "good" as the first run. You can see this by the way the regression line falls away from the unity line.

The confidence interval was also produced in Figure 4.

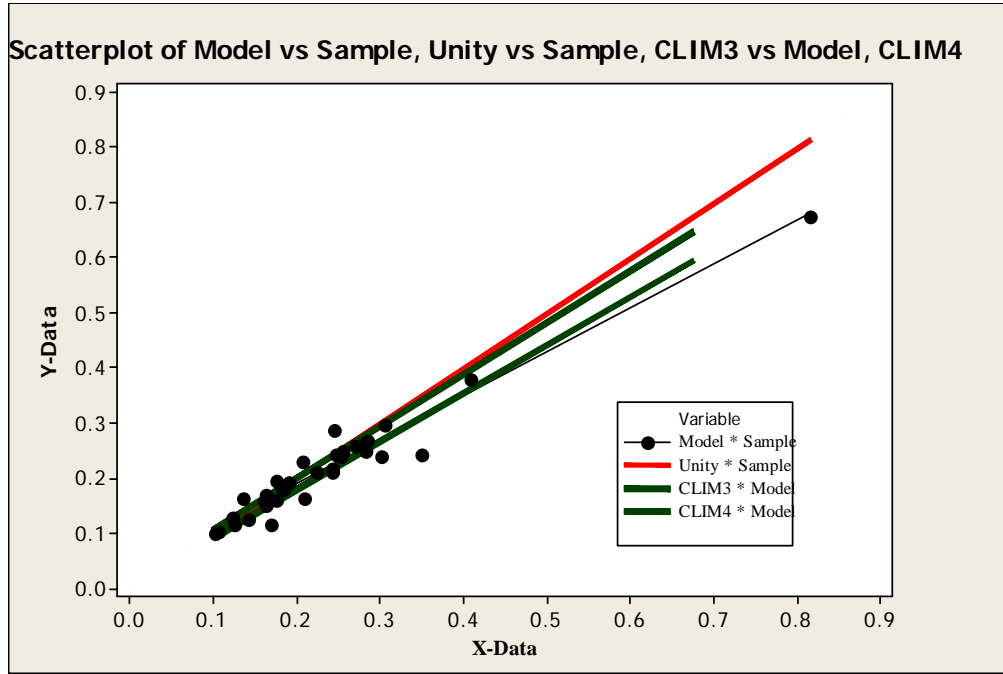


Figure 4.

Notice that the confidence interval doesn't even include the unity line when we get past 0.25 mg/L of bioactive nitrogen.

The calculations suggest that the model appears to underestimate the concentration of nitrogen in the Pleasant Bay system. The model has a bias!