SWEM QUESTIONS FOR MODEL WORKSHOP

A. Management Overview

What is the purpose (i.e., heuristic or predictive?) of the model as a management tool?

In principle, SWEM is meant to be a predictive tool, although it has some utility as a heuristic or learning tool as well. Although considerable efforts have been expended to calibrate both the hydrodynamic and water quality submodels of SWEM against observed data, the model can only reproduce the general features of circulation, primary production, nutrient cycling, and dissolved oxygen dynamics within the SWEM domain. In particular, representing the myriad of phytoplankton species present within the Sound over an annual cycle with two functional phytoplankton group (although that is pretty much the state-of-the-science) limits the model in its ability to "truly" predict the future. Rather the model is best suited to answer the relative differences and benefits that might result in Long Island Sound water quality in response to various management actions. The model, however, can also be viewed as a learning tool – often to identify areas of the model, which are weak and could benefit from additional research and/or field monitoring programs.

B. Hydrodynamic Submodel

I. East River volume flux has been subject to considerable scrutiny and discussion.

1. **How representative are the transport estimates with respect to other models and empirical evaluations?**
2. **What has been used as the "standard" for evaluating accuracy?**

The transport of water through the East River has been the subject of many studies over the past 100 years. The most definitive result is that of Blumberg and Pritchard (1997) who found for a 144 day period in 1989 using high quality current measurements, a transport of 360 m^3/s directed from Long Island Sound to NY Harbor. A transport in the upper layers of 120 m^3/s was directed towards Long Island Sound while the lower layers were directed in the opposite direction with a transport of 480 m^3/s. Their detailed 3D hydrodynamic modeling extended the analysis period to cover 18 months with the transport estimates not changing more than 20% from the values mentioned previously. Blumberg and Pritchard (1997) also show that the transports exhibit considerable fluctuation on 34 hr to 5 day time scales. Peak fluctuations of about 1500 m^3/s towards New York Harbor were evident. The SWEM computed values of the transport have been compared to the transport data-derived estimates using a goodness of fit assessed via acceptability criteria developed by Pritchard (undated). That criteria essentially stated that the monthly averaged transport for the net, upper and lower layers should be within +- 100 m^3/s of the Blumberg and Pritchard (1989) estimates. This criterion was the result of extensive negotiations between the late Prof Donald Pritchard and the NOAA Long Island Sound project managers and was considered to be a strict assessment of model skill. The SWEM fluxes are within this acceptance criterion 58% of the time. The only other 3D hydrodynamic model of the Long Island Sound - New York Harbor system, a model developed by NOAA, met the criteria 50% of the time.

II. East River and Hydrodynamics in general

1. **Can some perspective be put on the model calibration year as to its uniqueness or is it judged to represent an average condition?**
2. **If the calibration period is unique, what adjustments have been or can be made to the model to improve results for more common conditions or more extreme events?**

The model as been run for 6 different years. In each year there have been an extensive set of model result to data comparisons. In every case the model has performed well. The 6 years represent a range of conditions. It appears that the main calibration years, 1988-1989 and 1994-1995, represent average and low river flow periods, respectively. This conclusion is based on a 22 year record of Connecticut River flow as measured at Thompsonville, CT. The other four years were used only for hydrodynamic model validation and were not involved with the water quality SWEM runs. These additional years tend to have higher than average river flows.

The same set of model coefficients/parameters have been used in the 6 year hydrodynamic model run. No adjustments were necessary to reproduce conditions ranging from normal to extreme. This is a testament to the model physics and numerics.

III. Model Calibration and Verification

1. **Can X-Y scatter plots be prepared to compare model current velocity (perhaps net, maximum and minimum) outputs with measured data for a few locations on the East River and in the Sound?**

There are many ways to compare model results with data. Creating X-Y scatter plots is one method. These plots could be made for the locations where current measurements were taken in the East River but they have not been. The effort is a very large one. The X-Y plots have been made for the salinity and temperature comparisons. Time series comparisons have been made instead of X-Y plots for the currents. Visual and statistical measures were used to quantify the goodness of fit. Near College Point, the RMS errors between the modeled currents and those observed at three different depths were 7% at the surface, 8% at mid-depth and 12% near the bottom. In future model skill assessments the X-Y plot will be part of the statistical comparison package.

2. **If so, r-squared values will help describe model fit as an easily recognizable statistic.**

The r-squared metric is a good one. It has been used in the analysis of salinity and temperature. When X-Y plots are created for currents, this metric will be used.

3. **How accurate are the atmospheric heat and momentum fluxes used to force the hydrodynamic model?**

The atmospheric heat and momentum fluxes are calculated from meteorological data using the bulk aerodynamic formulae of Large and Pond (1982). This method of computing fluxes is subject to error because both the data and formulae have uncertainty. It is difficult to quantify the error as no flux measurements exist in the area and time periods of interest. One way to quantify the heat flux error is to see how well the model can reproduce temperature. If the model's heat fluxes are wrong, then it wouldn't be expected that the model surface temperatures would be right. For temperature data collected in Long Island Sound at 60 locations at approximately two week intervals for three different years, the model is able to reproduce the surface data with an r-squared value of 0.98 . Bottom temperatures are also well reproduced. The r-squared is a little less, however, about 0.97 .

4. **What components have the greatest uncertainty?**

The terms in the heat flux balance are all important. The greatest uncertainty lies with the data that is used in their computation.

5. **How representative are the fluxes of a typical year? Is there significant interannual variability in these quantities?**

We have not looked into the details of the interannual variability of the atmospheric forcing. The modeling has spanned 6 complete annual cycles so we expect to have covered much variability.

6. **How much does the year-to-year variability and/or uncertainty in the atmospheric momentum and heat fluxes contribute to uncertainty in the hydrodynamic outputs (e.g., stratification, circulation)?**

The error in the temperature computations does not seem to be a function of year to year variability in the atmospheric forcing. The r-squared values are pretty much constant. But, it can be expected that the model results will be worse for years when there are big differences in atmospheric properties between those measured over land and those measured over water.

7. What sorts of measurements are missing that would better constrain the above uncertainities?

The biggest improvement to the computation of heat and momentum fluxes would come from more representative measurements of the data needed for the computations. Now meteorological data is based primarily on information gathered from seven airports in the study region. Only one offshore station is used. The best results would be those that involved data measured over the water as opposed to land based stations. This alone would go a long way in further improving the model results.

C. Eutrophication Submodel

I. Calibration and Verification

1. Were the hydrodynamics and physical components of the model adjusted to improve calibration? How? What was the rationale?

For the hydrodynamic calibration, the SWEM model grid at the Race was adjusted to refine a preliminary calibration to better account for land masses (islands) not originally included. Bottom roughness coefficients in the East River were also adjusted to improve the transport calibration using total nitrogen (in the water quality submodel) as the tracer. In the water quality model, vertical mixing coefficients were adjusted in a portion of Long Island Sound and certain other geographical regions of SWEM to better reproduce observed DO stratification. The experience is that vertical salinity is relatively insensitive to the vertical mixing coefficient adjustments in the water quality model and cannot discriminate to the degree necessary to account for vertical DO variation.

2. Can X-Y scatter plots be prepared to compare model nutrient and dissolved oxygen outputs with measured data for a few critical cells in the Sound and perhaps New York Harbor?

Yes, some examples for DO, total nitrogen and particulate organic carbon were shown at the SWEM Workshop.

II. Loading

1. **How important is the timing of nutrient delivery with respect to severity of hypoxia? (e.g., wet spring vs. wet summer, vs. dry winter, etc.)**

   This assessment has not yet been made with SWEM. Additional years of model/observed water quality data comparisons would be needed.

2. **Riverine (fall line) loads are large with respect to other sources, and fluctuate widely based on rainfall. How do the model calibration and verification years compare to long-term averages and what are the implications for model predictive capability?**

   Water years 1988-89 and 1998-95 were relatively wet and relatively dry, respectively, on the basis of long term rainfall statistics. Model validation with additional years of data would be required to better assess the implications regarding predictive capability.

3. **How accurate are boundary loads delivered into LIS and can they be compared to a pre-Colonial load?**

   The water quality boundary conditions used in SWEM which are the basis of the boundary loads are limited: ten stations along the SWEM open boundary sampled on four occasions during 1994-95. Due to this limited data, the accuracy is difficult to quantify. Pre-Colonial conditions are unknown.

4. **Is nearshore (embayment) attenuation of nitrogen accounted for?**

   No. The model grid is not sufficiently refined in most embayments.

III. *Biological components*

1. What biological add-ons have been included in the model?

Traditionally, the various versions of HydroQual's eutrophication models represented zooplankton grazing using a simple (temperature-corrected) first-order