Response to Comments on Draft Watershed Model Calibration Report

Stephen Carter
Tetra Tech, Inc.
September 18, 2006

1. Los Angeles Regional Water Quality Control Board

1.1 Comment:

The model seems to assume that metals in rainwater are uniformly deposited throughout the watershed. It further assumes that the metal concentration of rainwater and the dry deposition mass that occurs between storms will be added to that rainfall concentration. There are two issues that are of concern regarding this. The first is the implicit assumption that metal concentration in runoff is the same for all land uses. The second is that there is nothing shown in the report or in any discussions testing that assumption. If stormwater (from atmospheric dry/wet deposition) is a significant contributor to the metal mass loading, the contributions from rainfall to runoff need to be verified. One of two things should be done to help with this potentially significant shortcoming. The best option would be to monitor runoff (flow and metal) from various land uses in the watershed to ensure that the model is predicting land use runoff correctly. The second would be to rely on historic regional data collection for comparison. The Los Angeles Department of Public Works has historic data that includes metal monitoring at both land use and watershed scales. Those data could be compiled and compared with model predictions.

1.1 Response

The model does not include specific parameterization of wet- and dry-weather deposition of metals. Also, modeling parameters associated with metals are specific to each land use. A description of the regional modeling approach, including a list of previous studies and references that has led to its development, is discussed on page 25 and Section 3.3.1 of the memo. This approach is consistent with options described in the comment, which outline a methodology for monitoring and modeling individual land uses.

1.2 Comment

Air deposition in wet weather condition should be included in the model. Also, if urbanized areas are a significant source of metal, how will the runoff from those areas be independently calibrated?

1.2 Response

As much as atmospheric deposition contributes to land use loadings, they can be considered implicitly included in the loading assumption. The model does not include explicit assignment of atmospheric deposition of pollutants. However, atmospheric deposition can be considered a general source to pervious and impervious areas contributing to surface loads and influencing calibration of land use modeling parameters.

1.3 Comment

There should be a discussion of how the model will account for overland flow in areas that are built up. Will a percent impervious be assumed by land use? If so, how will those numbers be determined?

1.3 Response

Percent impervious is assumed in the model for each land use. As stated on page 8 of the memo, assumptions for imperviousness were based on TR-55 percentages.

1.4 Comment

The validation results for Maritime Museum station did not match very well with the measured flow. The hydrology parameters in this area need to be adjusted to capture the measured peak value.

1.4 Response

As discussed on page 16 of the modeling report, development of modeling parameters were based on calibrations performed in other watersheds where long-term datasets were available for robust analyses and comparison of model results with observed data. The Maritime Museum station is located downstream of a mix of land uses, with each land use represented in the model with unique hydrologic modeling parameters. The single storm monitored for Maritime Museum is not a compelling argument for re-calibration of all land use modeling parameters, nor did it provide information to guide which land uses were not calibrated appropriately. Also, based upon our experience of calibration and validation to regional hydrographs, not all storms are captured effectively by the model at such short time scales and small sites. In some cases, the model would be unable to predict a specific storm at a site based on locally measured rainfall data, but effectively modeled all other storms at the same site. Therefore, before modification of the regionally calibrated modeling parameters, we suggest more data collection at the Maritime Museum site to verify that other storms are also not modeled effectively, and that this is not an anomaly.

2. Exxon Mobil (Lial Tischler of Tischler/Kocurek)

2.1 Comment

Difficult to evaluate the calibration and validation results in the context of the predictive capability of the model

2.1 Response

EPA is open to suggestions to improve reporting of calibration and validation results to improve the ability to review the model's predictive capability. Limited watershed data collected in the vicinity of the Ports provides little opportunity for comparison of model output with observed data. We also encourage a thorough review of other regional modeling efforts referenced in the modeling report that further document the calibration of regional hydrologic, sediment, and metals modeling parameters. The current study is considered a continuation of that effort, and relies heavily on the associated previous documentation of model predictive capability.

2.2 Comment

Graphical comparisons give no information on the quantitative errors in the model flow estimates, which are necessary to assess the uncertainty in the predictions. A quantitative comparison can be done by calculating the total runoff event volume from observed data and comparing it to the total event volume predicted by the model.

2.2 Response

Although the comparison of storm volumes can provide indication of the uncertainty of model predictions for the "Port Activities" land use, this information will not provide guidance regarding the overall uncertainty of the model including all land uses. Also, typical quantitative analyses of model uncertainty are based on long-term data for multiple years so that statistical analysis of model accuracy can be performed. A single storm monitored at three small watersheds does not provide the necessary amount of data for quantitative evaluation of model uncertainty. For this reason, we encourage a thorough review of other regional modeling efforts referenced in the modeling report that further document hydrologic calibration and validation based on large hydrologic monitoring datasets.

2.3 Comment

Use a similar approach as described above to calculate the total event sediment mass from observed data and compare it to the predicted sediment mass per event from the model.

2.3 Response

As stated above, data from a single storm monitored at three small watersheds prevents a robust quantitative or statistical analysis of model uncertainty. EPA will evaluate the recommended approach to determine if results are helpful in review of model uncertainty. However, there is concern that results may be misinterpreted as indicative of overall model uncertainty, which cannot be evaluated without also including review of previous modeling studies that led to development of the regional modeling approach.

2.4 Comment

Use similar approach for a quantitative comparison as described above for flow and sediment. EMCs in Table 6 show acceptable predictive capabilities. Compare storm water concentrations with water column concentrations to determine compliance with acute water quality criteria.

2.4 Response

See response to comment 2.3.

2.5 Comment

Estimated concentrations and loadings are questionable. These chemicals present a bigger problem than Paths did with respect to model predictions and uncertainty. Back-calculating sediment concentrations in runoff from concentrations in the harbor and bay sediments implicitly assumes that there are continuing sources of these pollutants from the nearshore watersheds; however, none of the runoff measurements by the County show any measureable concentrations of the pollutants. The data described in the report from Ackerman and Schiff show only DDT in only the agricultural land use category. Therefore, the assumption that the nearshore watersheds contribute ANY DDT Chlordane, or PCBs to the harbors and bay is unproven and may be inappropriate. There is an equal or greater probability that the DDT, Chlordane, and PCBs in the harbor sediments is legacy material that is not being added by the current runoff - it may simply be moving around due to sediment transport. There should not be an assumption that the chemicals are present in the runoff, including the assumption that that they are present at 1/2 the detection limit. Modeling with the assumption that they are not present in the nearshore runoff may show a better fit to the existing data.

2.5 Response

Lack of detects for these pollutants in runoff measured by the County may be due to several factors, including historic high detection limits, which do not necessarily indicate lack of presence. Based on the back-calculated methods, predicted concentrations were mostly below those same detection limits. Data reported in Ackerman and Schiff were also influenced by high detection limits. Additional water quality calibration of the receiving water model, based on predicted watershed loadings, will provide further evaluation of the accuracy of assumptions for watershed loads.

2.6 Comment

PAHs, DDT, chlordane, and PCBs - Assumptions are appropriate and consistent with measured data.

2.9 Response

Comment noted.

3. Port of Los Angeles

3.1 Comment

For DDT, PCB, and Chlordane the method utilized to calibrate the TT model is based on the assumption that the pollutant concentration from the watershed during wet weather is similar to the pollutant concentration in the harbor bottom sediments. This assumption is likely to over predict the watershed pollutant loading because the pollutants in the bottom sediments are likely to be more concentrated than in the water column. In addition, the pollutants in the bottom sediments are likely to be coming from multiple locations throughout the harbor but have become concentrated in areas that undergo increased sedimentation. In general, for a complex water body like the LA/LB Harbor complex, it is difficult to identify meaningful direct "cause and effect" correlations between the receiving water pollutant concentrations and pollutant loadings from the immediate adjacent nearshore subwatersheds. For example, as pointed out by many other TAC members, the pollutant sediment concentrations at the Bight 03 East Basin Station are more likely associated with pollutants from the DCW rather than the seven, small nearshore subwatersheds located nearby.

3.1 Response

Additional water quality calibration of the receiving water model, based on predicted watershed loadings, will provide further evaluation of the accuracy of assumptions for watershed loads.

4. Port of Long Beach (Weston Solutions)

4.1 Comment

The model assumes that land use dependent hydrologic runoff and pollutant loading model coefficients can be considered "transferable" to neighboring watersheds with similar land use characteristics. However, the assessment of land use categories and runoff coefficients appear very coarse. Adjustment of these coefficients may be a better way to calibrate the model.

4.1 Response

The approach presented is based on application of the regional modeling approach. We encourage a full review of all such supporting documentation to provide the reviewer an expanded understanding of the development, calibration, and validation of modeling parameters, and the proven transferability among watersheds in the LA Region. The model does not simulate hydrology and pollutant loads based on model coefficients, but rather on specific parameters included within HSPF and LSPC.

4.2 Comment

Preliminary model results are shown in Figure 4 of the cited memo for "Forest" flow and TSS for a single storm event during February 2003. The memo notes that the "… This small discrepancy in flow is well within acceptable modeling ranges." This judgment call is technically indefensible since no data quality objectives (DQOs) for modeling results have been identified. U.S. EPA's Council on Regulatory Environmental Modeling identifies best modeling practices. These practices include DQOs that quantify the acceptable level of total uncertainty that will still enable model results to be used for the intended purpose.

4.2 Response

Monitoring data in the Port area were limited to a single storm observed at three small watersheds. These single storms do not provide the necessary amount of data for full evaluation of model uncertainty based on quantitative and statistical methods, and instead must rely on regional analyses that have been performed in multiple LA watersheds that have furthered the science and development of the watershed models though continued testing, validation, and sensitivity analysis. The present study is a continuation of this effort, so previous documentation referenced in the memo should be included in review of model uncertainty. Data quality objectives were not set for comparison to the limited storms monitored in the port area, as these results are confined to calibration of modeling parameters associated with the "port activities" land use based on refinement of previously calibrated industrial land use parameters calibrated regionally, and do not reflect overall model accuracy.

4.3 Comment

The acceptability of the preliminary calibration is solely based on visual observation (Figure 4 in the cited memo). The predicted modeled flow and more importantly, the total volume (area under the flow curve) do not adequately represent the measured flow. Accurate calibration of the hydrology is essential or the modeling will lead to inaccuracies. It is obvious that the model is predicting significantly more volume of runoff than was measured. Also based on visual observation (Figure 4 in the cited memo), the discrepancies in modeled TSS vs. observed TSS are not within an "acceptable" range. This figure shows that most predictions are off by an order of magnitude. Since the model over-predicts flow and under-predicts TSS, the total load (product of the two factors) may be relatively accurate. Although this is a convenient result, it is not an indication of an accurate model.

4.3 Response

It is important to point out the ranges of flow we are considering (+/- 1 cfs in Figure 4), which are difficult to pinpoint with any watershed model. Also, through the calibration and validation process, a balance needed to be reached to provide the best fit to both storms. In other words, the model could provide a close match to the Forest site during calibration, but modeling parameters needed to be balanced to also provide a good match during the validation at the Pier A site. It is important to also re-state that this calibration/validation was focused on the “Port Activities” land use modeling parameters. For refinement of “Port Activities” sediment parameters (Figures 7 and 8), a balance was also reached between the calibration performed at the Forest site, and the validation performed at the Pier A site.

4.4 Comment

Identify specific modeling DQOs such as: What summary statistics for model performance have been computed from model results vs. observed flow data to support the above statement that the hydrologic model of runoff for the new nearshore sub-watershed drainage areas can be considered calibrated?

4.4 Response

Not enough monitoring data are available for meaningful statistical calculations, therefore statistical DQOs were not set. However, such summary statistics and DQOs have been reported and applied in other studies cited in the memo that have contributed to development of the regional modeling approach.

4.5 Comment

Identify specific modeling DQOs such as: What are summary statistics relative errors (as %) for wet weather indicators such as 10% highest flows; 25% highest flows; relative errors for total storm volume and average storm peak flow error for each stormwater monitoring site location?

4.5 Response

See response to comment 4.4.

4.6 Comment

Identify specific modeling DQOs such as: What are the calibration targets (as % relative error) assigned for each of the summary statistic to quantify hydrologic model performance?

4.6 Response

See response to comment 4.4.

4.7 Comment

The extraction of single storm event results for one or more specific stormwater station sites are very meaningful to show the onset and decline of flow and pollutant loading from a storm. However, it is also necessary to show time series plots of observed measurements and model results for the entire calibration time period to show model performance against all available data used to build the model.

4.7 Response

The only available data at the three monitoring sites were for the storm events shown in the memo. Additional comparison to other monitoring data collected throughout the LA region are documented in additional studies cited in the memo that have contributed to development of the regional modeling approach.

4.8 Comment

The modeling assumptions state that TSS is the primary carrier of contaminants and the best indicator to be tracked. The dissolved phase of these contaminants is not considered important. Sediments are not the only carriers of contaminants.

4.8 Response

In stormwater, metals are primarily observed in the particulate phase. Other organic pollutants are also known to have an affinity to sediment during transport.

4.9 Comment

The order and magnitude of storms also plays a large role in TSS and contaminant transport. The same annual rainfall will produce vastly different flows and contaminant transport depending on whether they come in one large storm or several smaller storms. This variability needs to be addressed in the model through sensitivity analysis.

4.10 Response

Flows and contaminant transport are simulated in the model at an hourly time step based on hourly rainfall data. This ability to simulate rainfall variability was one of the major factors that led to model selection.

4.11 Comment

The same comments described above apply to the quantification of calibration DQOs for metals loading models. Predicted vs. observed metals data should also be presented for the entire model calibration period. In addition, based on visual observation (Figure 5 in the memo), the model does not adequately represent the first flush, but its impact is evident in the measured values.

4.11 Response

All historic data and comparisons to model predictions were reported. The two peaks were the result of two separate peaks in rainfall intensity or the convergence of two hydrographs with staggered peak flows.

4.12 Comment

Using simulated continuous flow and an Event Mean Concentration (EMC) for estimating pollutant loading of PAHs is reasonable methodology for the development of watershed runoff and pollutant loading models for the study area. The simplicity of the method is its chief advantage. The inherent variability of the chemical load generated during a storm event can be adequately evaluated using the sensitivity analysis as outlined memorandums to Peter Kozelka and L.B. Nye dated May 5, 2006 from Tetra Tech.

4.12 Response

Comment noted.

4.13 Comment

The EMC approach as defined for PAHs estimation of pollutant loads will be adequate for organic chemicals if such data becomes available for the study. In the absence of EMC, the cited memo indicates that time series of sediment loading data will be used to estimate organic chemical concentrations and loads. The methodology to perform such a sediment-based estimate is not detailed in the cited memo. Will chemical-specific partition coefficients be assigned to estimate the particulate form of the organic chemical loads? What is the source of data that will be used to define chemical-specific partition coefficients?

4.13 Response

The sediment-based approach for estimating organic pollutant loads was detailed in Section 3.3.3 of the memo.

4.14 Comment

The assumption of zero concentration and load for PAHs and organic chemicals (DDT, chlordane and PCBs) is based on results of dry weather monitoring studies in the LA region conducted by LADPW between 1988 and 2005. With the much higher number of dry weather days than wet weather events, dry weather flows contribute a larger volume of water than wet weather flows discharged to the Harbors from the watersheds. Even very low concentrations of PAHs this larger water volume can amount to significant loading. Additional sampling should be conducted of dry weather flows with improved analytical detection limits or modified sampling protocols.

4.14 Response

We agree that additional sampling would be beneficial to validate this assumption; however, we are not aware of plans for this type of monitoring.

4.15 Comment

The watershed loading report needs documentation of what criteria is actually used to define wet weather conditions compared to dry weather conditions. Will a listing be prepared to identify a sequence of wet weather time intervals with dates not included on the list presumed to be dry weather conditions? Presumably the model calibration period will include a mix of wet and dry weather conditions so that loads provided as linkage input to the EFDC water quality model will be calculated on any given day will be estimated using either dry or wet weather methodology. The calibration report should present time series plots of flows and pollutant loads for (a) the entire simulation period and (b) selected wet weather event station conditions.

4.15 Response

Wet weather conditions are defined as all flows predicted by the LSPC model exceeding those calculated for each model sub watershed based on equation (2) of the memo. The use of these flows by the receiving water model will be further discussed in future reports following development and calibration of the model.

4.16 Comment

In the full model calibration report, a pollutant load budget for each contaminant for the model calibration period should be compiled to identify what proportion of the total simulated load is contributed by dry weather conditions and what proportion is contributed by wet weather conditions. The contributions of wastewater reclamation plants (WRPs) should be quantified as a separate line item of as a component of the total dry weather load.

4.15 Response

Comment noted.

5. Heal the Bay

5.1 Comment

The fact that the model under-predicts the flow and parameter concentrations for the Maritime Museum site is concerning. If this model is used in the development of the TMDLs, this could be an issue. For instance if the loadings are underestimated, the final waste load allocations may not get us to the goal of water quality standards attainment. Also, there are numerous model assumptions provided on Page 54 of the Report. For instance, the modelers assumed that the LAR and SGR models were representative of the loadings from their respective watersheds without further validation. We do not know that this is the case. Therefore, several of these assumptions compound the uncertainty with the model. A very large margin of safety would be necessary to off-set the model uncertainty.

5.1 Response

Results of comparison of model predictions with data from the Maritime Museum site are not representative of model uncertainty. It is important to note that rather than calibration to a single storm monitored at a single location, calibration and validation must be performed at several location/storm combinations to eliminate variables that can unfairly impact the calibration/validation process. Such variables can include error in rainfall, flow, or water quality measurements, spatial variability of a storm not captured by the rainfall gage, etc. Several monitored storms and locations, in addition to long-term datasets collected at County mass emissions stations, have been used in development of the regional modeling approach that has led to models developed for Ballona Creek, LAR, and SGR, which were cited in the memo and should be reviewed for a full understanding of model development. These models and their land-use-specific modeling parameters have been fully calibrated and validated, with results documented in the reports cited in the memo. Application of these modeling parameters to the nearshore areas of the harbors is therefore an expansion of the regional modeling approach, with additional parameters added specific to "Port Activities," calibrated and validated using data from the Forest and Pier A sites.

5.2 Comment

Does the model fall apart based on the size of the drainage? (i.e. for hydrology the Forest Sub watershed (the smallest area) predictions appear to be the best). The model also appears to under-predict the flow at Pier A at the beginning. Can there be some sort of correction factor?

5.2 Response

The size of the Pier A and Forest watersheds were unlikely a factor in model accuracy. Regardless of the attempts to calibrate or "correct" some of the modeling parameters, the model consistently predicted a general shape of the hydrograph with a smaller initial storm peak followed by a higher peak (Figure 5). The model could have been modified to capture the initial peak, but then would have greatly over-predicted the magnitude of the second peak. For this reason, we opted to minimize the overall error in stormflow prediction, resulting in the model results reported.

5.3 Comment

Are the land uses of the Forest, Pier A and Maritime sites very similar? If so, how well will the model predict loadings from other types of land uses?

5.3 Response

The Forest and Pier A watersheds were representative of the "Port Activities" land use, which provided calibration of a unique set of modeling parameters specific to this land use. Modeling parameters associated with other land uses were based on regionally calibrated parameters reported in studies cited in the memo, and previously applied successfully to models of Ballona Creek, LAR, and SGR. Accuracy of the model for the other land uses are fully documented in the other regional modeling reports cited in the memo. The Maritime Museum watershed is a mix of these land uses.

5.4 Comment

What average EMCs were used for the metals modeling?

5.4 Response

EMCs were not used to model metals. The approach used in the LSPC model for metals simulation is discussed in Section 3.3.1 of the memo.

5.5 Comment

Why would the EMCs for PAHs be greater at agricultural sites than transportation sties?

5.5 Response

The differences can be further defined by the large standard deviation associated with agriculture (1.0 E3) reflecting the large variability. Therefore, there is also more uncertainty associated with the agricultural EMC (order of magnitude greater than transportation at 2.8 E2) that may explain how the average can be greater. Without more data describing EMCs for transportation or agricultural sites, assumptions will be based on these averages, with sensitivity analysis performed to reflect the variability associated with the standard deviations. Regardless, little or no agricultural land use is relevant except for the LAR and SGR watersheds, where PAH concentrations are not estimated based on land use EMCs but rather on stormwater monitoring performed at the bottom of the respective watersheds (see p. 36).