Dominguez Channel and Greater Los Angeles and Long Beach Harbor Waters Toxic Pollutants Total Maximum Daily Loads Draft

Prepared by California Regional Water Quality Control Board, Los Angeles Region
And the U.S. Environmental Protection Agency, Region 9

External Peer Review

by

Patrick L. Brezonik[1]
Apple Valley, Minnesota

Overview

My initial impression upon starting to read the TMDL document was favorable. It was clear that a very large effort went into the development of the document and its associated appendices. The Introduction and Problem Statement sections are well written, and the analysis of impairments identified in 303(d) lists, as well as the assessment of findings for each water body is thorough. Unfortunately, as I continued to read the report, my opinion became less positive. The writing in key sections on numeric targets and the TMDL development (sections 3 and 6) was unclear, and I had difficulty understanding the scientific basis for some numeric targets and TMDLs.

My opinion further declined as I read the two appendices related to the critical modeling components. Although the models that the authors used are widely used and represent the state-of-the-art in watershed and hydrodynamic modeling, the calibrations were poor to mediocre. Similarly, although an attempt was made at model validation for some of the contaminants, it was not successful. As a result, to the extent that the models were used to generate the TMDLs, WLAs and LAs, I do not think that much confidence can be placed in the numbers.

A broad framework is provided in the TMDL document for the implementation plan, which includes a monitoring program. Actual details of the implementation plan and monitoring program are left to the responsible parties to develop. Additional monitoring of water and sediment quality is critically important, not only to gather information on the extent to which compliance with the TMDL objectives is achieved, but equally important to provide more and better data to calibrate and validate the models on which the TMDLs were based.

An analysis of costs to implement the TMDL is provided at the end of the report. The authors indicated that such an analysis was not a requirement of the TMDL process but presented it anyway. I found the analysis to be largely superficial, but if one accepts the numbers generated in that analysis to be even roughly correct, it is clear that the implementation will impose a large economic burden on the region. It is not within my role as a reviewer of the scientific merits of the TMDL report to make judgments on the economic impacts relative to the need or desirability of various components of the implementation program. In my opinion, however, it is within my purview to state that given the high projected costs, the science behind the analyses leading to the TMDLs (and thus the necessity for

---

[1] Dr. Brezonik is Professor Emeritus, University of Minnesota, Minneapolis, MN. This review was not performed as part of his (former) employment by the University of Minnesota, and no endorsement by the University of this report should be inferred.

implementing BMPs and sediment remediation) needs to be sound and the results need to be reliable. I conclude that unfortunately the current TMDL document does not meet this standard.

Responses to Major Issues

1. ***Appropriateness of selected sediment, fish tissue, and water numeric targets for OC pesticides, PCBs, PAHs, and metals.***

The numeric targets were based largely on state and federal water quality standards and criteria. These standards and criteria were developed over many years based on the best scientific information available, and I do not have any basis for criticizing them. Even if I did, I think that the authors of the report were constrained legally to use these values. The TMDL document notes that there are no numeric standards for sediments (called sediment quality objectives) in the California Toxics Rule (CTR), but the TMDL document relied on guidelines in a 2006 study on the development of California's 303d (impaired waters) list to develop the sediment quality guidelines (Table 2-4 of the document) that were used to assess whether sediments were impaired or not. This approach seems reasonable.

Nonetheless, I found Section 3 "Numeric Targets," (pp. 43-52) very difficult to follow and understand (see my detailed remarks regarding these pages in the section of this review titled "Other Comments"). The section on numeric criteria for chronic toxicity (pp. 44-45) lacks clarity. For example, I don't understand what the authors mean when they say "sample concentration was expressed as a percentage" (p. 45, below equation 1). Percentage of what? In the end, I was unable to make a firm conclusion regarding the scientific validity of the specific numeric targets because of the lack of clarity and details in the section.

2. ***Development of the sources and linkage analysis to show how sources of contaminant loading to the harbors are lined to the sediment and water quality.***

The authors of the TMDL document clearly expended considerable efforts in gathering background data for their modeling efforts. This included extensive historical information on water and sediment quality in the subject water bodies, as well as data on fish tissue levels of contaminants. The analysis of existing conditions appears to be thorough and credible, and the remaining uncertainties regarding the degree of impairment in the water bodies and their sediments reflect the absence or inadequacies of past monitoring programs rather than insufficient efforts on the part of the authors.

To run the contaminant loading models (LSPC and EFDC), the authors obtained detailed information on point and nonpoint pollution sources in the watershed, detailed watershed information needed to configure the models to the complicated set of watersheds in the study, and a variety of meteorological and water data needed to calibrate the basic hydrodynamic components of the model and the water quality (pollutant transport and fate) components of the model. As the authors note, the models they used are widely used in the environmental engineering community for surface water modeling in complicated systems, and they are accepted and supported by the U.S. EPA.

I have no criticism of the models per se except to note that such models do much better at simulating the movement of water itself than they do in modeling/predicting the transport and fate of non-conservative substances (e.g., pollutants) in the water. This is because the physics of water movement is well understood and can be described quantitatively by mathematical equations with physical coefficients that can be determined with fair accuracy. In contrast we simply do not understand how organic pollutants or metals behave sufficiently to write analytical equations with coefficients that are truly fundamental. In spite of their apparently "analytical nature," when models like EFDC are used to simulate the environmental behavior of non-conservative chemicals or biological components, they become inherently empirical, meaning that the accuracy of their simulations depends strongly on the availability of a robust set of calibration data.

The calibration exercises conducted by the authors for the TMDL study showed that the model generally did a good job in simulating water flows (at least insofar as water surface elevations at a NOAA tide gauge appear to be close to observed water levels. Results were not quite as good for modeled versus measured salinity, but part of the problem here is that many of the stations do not show substantial variations over time in salinity. In contrast, modeled trends generally did not accurately fit observed values for concentrations or loads of the three heavy metals (Cu, Pb, Zn) either in the subwatersheds used to calibrate the model or (even more strongly) in the subwatersheds used for model validation. The authors state several times (e.g., Appendix II, p. 15) that the differences between observed and modeled results were small and well within acceptable modeling ranges, but I simply do not agree with this statement. Furthermore, the validation results that are presented in the TMDL document and Appendix II really do not "validate" the accuracy of the model nor do they demonstrate that it is able to predict the behavior of the metals in the system with sufficient accuracy for the purposes needed in the TMDL analysis. (Just because one conducts a validation exercise does not mean that a model has been validated.)

There are at least two reasons why the calibration/validation exercise failed. First, as the authors point out, there was a paucity of data that could be used for calibration and validation purposes. This was especially the case for the organic pollutants, for which within-event calibration data were almost completely lacking. Perhaps this can be rectified by establishing a monitoring program (which is part of the implementation phase). Second, the model itself simply may not be sufficiently defined and refined to simulate the behavior of the pollutants in this system. The equations describing the behavior of metals in the model are not described in any detail in the TMDL document or modeling appendices, but my impression from the latter documents is that metal behavior is modeled primarily in terms of a partition coefficient, K_p, that quantifies the amount of metal in the dissolved state and that sorbed onto suspended particles. The behavior of the former presumably is modeled by water transport and the latter is modeled by equations intended to predict the settling and scouring of suspended particles. This certainly is a simplification of the complicated chemical and biological processes that affect behavior of the metals in aquatic systems, but it may be adequate if two conditions are met: (1) sorption/desorption to/from suspended sediments is the dominant process, and (2) this process can be quantified in terms of a single value for K_p. The results presented in Appendix I, Figure 31 (p. 47) clearly show that the latter is not the case. Values of K_p exhibit a wide range for all three metals, and they do not show a predictable relationship with the concentration of suspended solids. Consequently, the use of a single (average) value of K_p in the modeling effort is inappropriate and may account for much of discrepancy between modeled and observed concentrations and loads.

Use of the complicated hydrodynamic model may have been intended to give the impression that the authors used a sophisticated modeling approach, but given the lack of fit and inadequacy of calibration data, the results are no more reliable than if the authors had used simpler, more empirical approaches (e.g., plug-flow and completely-mixed reactor models) to conduct their loading and transport studies.

For further comments on this topic, see comments for pp. 69-80 of the TMDL document and all the comments for Appendices I and II in the section of this review title "Other Comments."

3. ***Calculating loading capacity (TMDLs)***.

In many cases explanations in the section on TMDLs are given considerably after the results are presented or are not given at all, making it very difficult for readers to understand what was done and what the basis for the TMDL really was. Overall, this section of the report was difficult to follow and understand. As a result, I am not able to provide a firm conclusion about the validity of the final results. One example regarding the lack of clarity involves Table 6-1, which provides WLAs and LAs based on toxicity criteria. It would seem that the various loads would be additive to the overall toxicity of the receiving water and thus the TUc values should be distributed fractionally among the dischargers. Perhaps I just don't understand what was done and how the calculations were made, but I do not think the report provides an adequate description for me to develop this understanding. Similarly, I was not able to figure out how the wet-weather loading capacities in Table 6-2 were obtained.

It is clear from previous sections of the TMDL document that large uncertainties exist in the modeling and analyses and that the available data is not sufficient in many respects. Given this situation, it seems to me that the small margin of safety (10%) provided in Table 6-4 is unrealistic. The values reported in Table 6-8 presumably represent 95 percentile values of historical data, but the text is not clear regarding how they "translate" to either a WLA. Similarly, the meaning of the TMDL values and allocations for bed sediments in Table 6-10 is not clear, and with regard to the first note at the bottom of this table, it is not obvious why no reductions in atmospheric deposition of Cu, Zn and PAHs should be anticipated. If atmospheric sources are contributing to the problem, they should be subject to regulation just as much as land-based point and non-point sources.

It also is not obvious why an implicit margin of safety exists in the final allocations to Dominguez Channel estuary and the greater Harbor waters (Section 6.5.3) just because multiple numeric targets were selected. They all could be "unprotective."

Finally, I wonder whether the tiny values listed in Table 6-12 for DDT and PCB WLAs are meaningful. Could one actually make measurements to show that a discharge was in compliance with a WLA of 0.35 g/yr? In general, the numbers in the table seem unreasonably low.

4. ***Development of a proposed monitoring program to assess effectiveness of the TMDL and attainment of water quality goals.***

The proposed monitoring program is an essential component of the TMDL implementation. The data that will be obtained will be critically important not only for compliance purposes but also for improving the database available for calibrating transport and fate models. In an adaptive management context, this will allow improvement of the analyses conducted originally as part of the TMDL study, thus likely allowing modification and improvement of the implementation plan, as well as the TMDL targets themselves.

The water parts of the monitoring program appear generally to be sound. In particular, the requirement to monitor two wet-weather and one dry-weather events each year, including the first major wet-weather event of the season, is reasonable. The monitoring plan described in section 7.6, starting on p. 116 of the TMDL document, does not provide sufficient information, however, on the nature of the sampling frequency within the wet and dry events. This may be spelled out in the SWAMP protocol and various MRPs and QAPPs, but it would be appropriate for the document at least to specify that sufficient samples should be taken within events to define the "pollutograph"—that is, the concentration and load versus time over the period of the event. In addition, the report does not provide specific information on the number and location of storm drain sites that will need to be monitored. I believe the report easily could be modified to present this information, which would make it much easier to evaluate the adequacy of the monitoring program. Finally, it is not clear what is meant by a dry-weather "event." It would be useful for the report to clarify this terminology and also the timing and duration of a dry-weather sampling program.

I doubt that it makes sense to require analyses of filtered water samples for dissolved DDT, chlordane, PAHs and PCBs at all sites. It is known from many studies in the literature that these highly hydrophobic substances occur on particulate phases rather than in the dissolved phase, and prior work in these watersheds (described in the document) has shown that levels generally are undetectable in the water itself. It may be appropriate, however, to require collection of dissolved natural organic matter (NOM) and analysis of this material for the above mentioned pollutants if dissolved organic carbon (DOC) concentrations in the stormwater are known to be high; this usually is done by passing water samples through columns of resins like DAX-8 and extracting the sorbed NOM. It is well known that organic pollutants sorb onto macromolecular NOM, which operationally is a part of the "dissolved" fraction when water samples are filtered using conventional filters. Given the geological and climatic conditions in the Los Angeles region, I doubt that DOC (and dissolved NOM) is high enough in surface waters of region to represent a significant transport medium for the pollutants, but aquatic chemists in the region should be able to evaluate this.

Sampling of sediments and fish within the various units of Dominguez Channel and the Greater Harbor also is a component of the monitoring program. Although the proposed sampling frequency of every five years may be sufficient for compliance purposes, in my opinion, it is not sufficient to improve the database needed for better calibration and validation of the transport and fate models. Therefore, I recommend that sampling and analyses of sediment and fish should be undertaken at least every two years for an initial period—until sufficient data are obtained to improve the models. It may be possible for this sampling to be done at fewer sites than needed for the five-year compliance monitoring, but sampling will need to be based on the requirements to achieve the goal of improved scientific understanding of pollutant distributions and dynamics in sediments and fish of the system rather than on compliance issues. The TMDL document does not necessarily need to include details on the exact sites to be included in this more frequent sampling, but it should be modified to address the need for more and better data to achieve the aforementioned goal.

5. *Evaluation of implementation plan and allocations.*

Insofar as I lack confidence in the results of the EFDC model used to generate the proposed implementation plan and allocations, I must conclude that the TMDL report does not provide a sufficient scientific basis for the proposed plan and allocations. That said, the report does provide a sound general approach to implementation that involves five broad processes: 1) implement and evaluate the effectiveness of BMPs and source control in conjunction with remediation to remove contaminated sediments; 2) evaluate the effectiveness of controlling sediment loading from major river sources (Los Angeles and San Gabriel Rivers and Machado Lake) through implementing effective TMDLs; 3) conduct compliance monitoring; 4) determine whether reductions in loadings from controllable sources in the Los Angeles and San Gabriel Rivers will be required and addressed through revision of the TMDL; and 5) re-evaluate the WLAs and LAs, as necessary.

Overall, the implementation plan provides a general framework for implementation rather than specific details, which are left to the "responsible parties" (local agencies and governmental units in the affected area) to develop. The implementation plan also is not prescriptive in stating specific activities, including BMPs, that should be undertaken to achieve the WLAs. In one sense, this approach is good in that it allows for local decisions to be made based on local knowledge. On the other hand, the approach adds uncertainty and vagueness to the implementation phase.

The implementation plan is also described as consisting of three phases. Phase I includes incorporating interim limits into NPDES permits and waste discharge requirements, implementing BMPs in the watersheds, implementing TMDLs for the Los Angeles and San Gabriel Rivers and Machado Lake, and developing and initiating a monitoring program. Phase II extends the implementation to clean-up of high priority areas, including sediment removal in harbor areas, implementation of additional BMPs, and other targeted source reduction activities identified in Phase I. Plans for Phase III are very sketchy and simply state that secondary and additional remediation actions as necessary will be implemented to insure compliance with final load allocations by the end of the implementation period. Table 7-2 (p. 122) indicates that Phase I should last five years, Phase II ten years, and Phase III an additional five years of the total 20-year implementation plan. Overall, the idea of a phased approach makes sense, and although the report does not use the term "adaptive management," the implementation plan does have many elements of adaptive management. Considering the very large costs associated with implementation of this TMDL, I agree that a phased approach is appropriate, and I also recommend that the implementing agencies develop an implementation approach that specifically follows the principles of adaptive management.

Clearly, the implementing agencies will need to develop more detailed plans for the three phases than are presented in the TMDL document. Although it is not feasible at the outset to provide as much specificity for Phase II as for Phase I, the plan at least should describe the mechanism and timing for formulating a detailed Phase II plan, and a similar requirement should exist regarding Phase III plans.

Processes 2 and 4 in the implementation plan involve actions outside the domain of this TMDL, specifically, the development of separate TMDLs for the Los Angeles and San Gabriel Rivers and Machado Lake. The latter is a small and apparently impaired water body between Wilmington and Harbor City and west of I-110, just north of Los Angeles Harbor. Given the proximity of Machado Lake to the harbor and the fact that it drains into the harbor, it is difficult to understand why this water body was not part of the present TMDL or at least why it was not described in more detail in the TMDL document. Its location is not even noted in Figure 2-1, although I believe it is present on the map as an unnamed water body just northwest of the Los Angeles Inner Harbor. Overall, this situation (i.e., three additional TMDLs being required to fully implement the TMDL for Dominguez Channel and the Greater Harbors) represents an unfortunate complication, but I understand that this may reflect legal requirements and is not necessarily an issue relevant to the scientific review of the TMDL document.

Responses to overarching questions

(a) ***Are there additional scientific issues that are part of the scientific basis of the proposed rule not described above?***

There may be other issues, but I believe that my major concerns with the proposed rule and its scientific basis have been addressed in responding to the above five issues and in the comments included below, which were developed as I observed issues and problems while reading the report and associated appendices.

(b) ***Taken as a whole, is the scientific portion of the proposed rule based on sound scientific knowledge, methods and practices?***

The authors of the report show clear evidence of detailed familiarity with scientific knowledge about the environmental problems in Dominguez Channel and Los Angeles and Long Beach Harbors and about the scientific bases for addressing these issues. In addition, the scientific portion of the proposed rule relied on generally accepted and sound scientific methods. For example, the models used in the study are generally accepted as “state-of-the-art” and are widely used by both government agencies and scientists and engineers in the private sector. The application of sound scientific practices was not always followed, however. Examples of instances where there was a lapse of sound scientific practices range from small statistical issues, such as using regression analysis when the basic assumptions inherent in the method were not present in the data (e.g., see comments on pages 52 and 53 of Appendix II below), to much larger issues like the continued use of the EFDC model to determine transport and fate of pollutants in the system in spite of the fact that the calibrations and validations showed that the model did not come close to matching the observed values.

Other Comments

The following comments were developed during my reading of the report and associated appendices. Many of the comments served as the basis for my remarks on the five major issues and two overarching questions in Attachment 2 “Description of scientific issues to be addressed by peer reviewer for proposed TMDL for Toxic Pollutants in Dominguez Channel and Greater Los Angeles and Long Beach Harbor Waters,” which were addressed in the two preceding sections. Many other comments relate to the clarity of the document (or lack thereof). The list is not an exhaustive compendium of this reviewer’s concerns and issues with the report and associated appendices, but it complements the responses to the five major issues and two over-arching questions raised in the attachment.

TMDL Report

Page	**Comment**
31	*Line 4 from bottom*: The text makes reference to summary tables for all the data but does not indicate where these tables are located.
39	*Paragraph 2.6.3 and subsequent ones*: No summary statement is provided regarding conclusions on what is impaired, as was done in paragraphs 2.6.1 and 2.6.2.
39	*Paragraph 2.6.5 and subsequent paragraphs*: I don't understand what the authors mean by "*certain* DDT and PCBs…" As far as I am aware there is only one kind of DDT, although there are several DDT degradation products.
43	*Line 2 above Table 3-1*: I don't understand what the authors mean by "…the CTR vice…"
43	*Table 3-1*: The relevance of including a water quality criterion for mercury in water based on protection of human health is not obvious. Previous text did not establish that there was any problem with mercury concentrations in the water column of any of the water bodies.
44	*Table 3-2*: It is not obvious to this reviewer how the freshwater wet weather metal targets in this table were obtained, nor is it clear what is meant by "translators," or why this was done.
47	*Paragraph 2 of Section 3.2.2*: Insufficient information is provided on the benthic invertebrate indices, including their nature and references to literature on them.
47	*Second last paragraph*: The text states that the combination of the four benthic invertebrate indices provides more information than any single index. I am not convinced that this is the case if all one uses is the median value of the four indices. If anything, use of the median value will decrease information on extreme conditions that the individual indices may provide. I do not think that this approach yields results that are helpful in deciding whether the sediment benthos is impaired or not.
48	*Last paragraph*: Proper names of organisms should be italicized. The second last sentence is not clear and needs further elaboration.
49	*Last paragraph*: At best this paragraph is unclear, but it seems to me to represent circular reasoning.
57	*Second paragraph under Section B*: If the analytical methods were not sufficiently sensitive to detect the pesticides and PCBs, how can the authors know that the discharge is a minimal source of these contaminants to Dominguez Channel and the harbor waters?
58	*Third last paragraph* (**and many other places in the report**): The report is sloppy with regard to citing references. Including the date of the Stenstrom et al. report in the text would tell the reader that the authors are citing a reference that can be found in the bibliography or reference section.
58	*Third paragraph*: the mean values given for copper, lead and zinc are very high (> 1 mg/L), and I wonder whether these are correct.
62	*Table 4-3*: It is difficult to evaluate the significance of the numbers in this table. Reporting the results as areal based loads ($g\ m^{-2}\ yr^{-1}$) would be more useful.
66	*Table 4-6*: Same comment applies.
68	*Third last paragraph*: By this point in the analysis, the authors should not have to resort to weak statements like "…atmospheric deposition may be a potential nonpoint source of metals, DDT and PAHs to the watershed…." Is it or isn't it? Data sources were cited earlier that should have allowed a more conclusive statement than this.
69	*Section 5.1*: The terms LSPC, LAR, and SGR were not defined previously and are not in the list of acronyms. Authors should define these terms and describe how the models work.
73	Mention of the three appendices much earlier in the section would have been helpful to the reader in understanding where to look for more information about the modeling approach.

77 *Figure 5-2*: It is impossible to distinguish the modeled results from the actual data in the black and white printed version of the report sent to me to review. The reader must accept on faith that the figure actually shows both.

77 *The second last "sentence"* actually is not a sentence and does not express a complete thought.

78 *Figure 5-3*: One cannot distinguish which data point and line represent the bottom water and which represent the surface.

78 *First paragraph under Figure 5-3* states "As can be seen from the comparisons indicated in the above figures, the hydrodynamic model provides a good foundation …." This is not really the case. Without presenting any statistics, the authors cannot make such a conclusive statement.

78 *Section 5.2.2, first paragraph*: The first sentence is not clear. What is meant by "only a calibration effort"?

80 *Second paragraph*: No data are presented here or cited to support this statement.

82 *Paragraph below Eq. 1*: Sample concentration is expressed as a percentage, but it is not clear or obvious what this means (percentage of "what"?).

Table 6.1: are not the various loads additive? If so, shouldn't the final TUc values be allocated fractionally among the permittees?

84 *Table 6-2*: It is not at all clear to this reviewer how the numbers in this table were obtained.

85 *Paragraphs 6.2.2.1 and 6.2.2.2*: The term "MOS" in the equations is not defined.

86 *Table 6-4*: Given the large uncertainties in the data, modeling and analyses leading to the allocations listed in this table, the margin of safety (10%) seems unrealistically small. I note that MOS finally is defined, after the fact, in Table 6-4.

87 *Section 6.3.2*: The wet-weather allocations given here seem reasonable given the lack of data, but one wonders why there are no data.

89 *Table 6-8*: It is not clear what the numbers in the table mean. Based on the text at the bottom of page 88, I assume that they are 95 percentile values of historical data, but the text is not clear regarding how they "translate" to either a TMDL value or a WLA.

91 *Third paragraph from bottom*: The paragraph, particularly the last sentence, strikes me as a bit of "hand-waving."

92 *Table 6-10*: The document is not clear on what TMDL values mean for bed sediments.

94 *Note under Table 6-10*: it is not obvious why no reductions in atmospheric deposition of Cu, Zn and PAHs should be anticipated. If atmospheric sources are contributing to the problem, they should be subject to regulation just as much as land-based pint and non-point sources.

98 *First sentence in section 6.5.3*: It is not obvious why an implicit margin of safety exists in the final allocations to Dominguez Channel estuary and the greater Harbor waters just because multiple numeric targets were selected. They all could be "unprotective."

Table 6-12: One wonders whether the tiny values listed in the table for DDT and PCB WLAs are meaningful. Could one actually measure a WLA of 0.35 g/yr? In general, the numbers in the table seem unreasonably low.

101 In many cases explanations in the section on TMDLs are given considerably after the results are presented, making it very difficult for readers to understand what was done and what the basis for the TMDL really was. Overall, this section of the report was difficult to follow and understand.

106 *Third paragraph*: The sentence that forms this paragraph is garbled and difficult to understand.

124 Overall, the cost analysis is very superficial and inadequate.

Second paragraph: It does not seem appropriate to simply average the two widely disparate estimates of dredging costs.

126 The cost analyses for sand/organic filters and vegetated swales also are superficial and inadequate.

128 *Table 7-7*: Even by today's standards, these are huge cost estimates. Although it is readily apparent that a large effort was expended in developing the TMDL document and associated appendices, the large uncertainties associated with the modeling analyses lead me to be very

skeptical that the work provides a sufficient scientific basis for the expenditure of such large amounts of money.

Appendix I

Page	Comment
14, 15	*Figures 5 and 6*: One cannot distinguish the "observed" and "predicted" lines in the black and white versions of these figures in the printed document. Authors of reports need to avoid using color for lines unless they are certain that the report will be printed in color.
20-22	*Figures 8-11*: The same comment applies to these figures.
26-27	*Figures 14 and 15*: The two different sets of data in both figures cannot be distinguished by the symbols used in the figures, nor is it obvious which line refers to the "bottom fit" and the "surface fit."
28	*Last paragraph*: Although there may be an entirely reasonable explanation for not including physical bed data from inside the breakwater for years prior to 1997, no explanation is provided, leading me to be concerned about whether this was an arbitrary decision.
29	*Second paragraph*: Similarly, no explanation is included for eliminating sediment metals data from inside the breakwater prior to 2000, leading to concerns about arbitrariness. In addition, the text is not clear on how initial concentrations of metals and organic contaminants in the sediments (displayed in the maps in Figures 23-28) actually were estimated.
32	*Figure 18*: No r^2 values are given in the plots to demonstrate the level of precision of the predictive equations, nor is it clear whether the outlier value in the upper figure was included in the regression analysis.
47	*Figure 31*: The data shown in the three plots of Figure 31 are all "over the map," leading to two conclusions: (i) there is no predictive relationship between the partition coefficient (K_p) for heavy metals and total solids concentrations, and (ii) use of a mean value of K_p for modeling purposes would result in large uncertainties in predicted results because of the large range in K_p values.
50-51	*Figures 34-36*: The same comment applies to K_p values for organic contaminants.
64-65	*Figures 41 and 42*: There is virtually no relationship in the scatter plots for copper in the figures. All that one can conclude is that the predicted numbers are in the same order of magnitude as the observed values. I suspect that the latter fact reflects "tweaking" associated with the calibration effort. I conclude from the figures that the model cannot be used to predict effects of changes in external loading on sediment concentrations with any degree of accuracy or reliability and that it would be even worse in predictions of the effects of other environmental/management variables on sediment levels of copper.
66-67	*Figures 43 and 44*: The results for lead and zinc are even worse in that the predicted sediment concentrations exhibit a much large range than the observed values for two of the three lead plots and all three zinc plots. As such the results suggest that the model may produce differences or trends in concentrations of Pb and Zn in runs where environmental or management-related parameters are varied even though such differences or trends may not occur in reality.
68	*Figure 46*: The same comment as above applies to the PAH plot.
69	*Paragraph 2*: This is an honest appraisal of the adequacy of the data for modeling purposes, but I am not convinced that the statement in the third paragraph "…it has been demonstrated to respond appropriately to load reductions and is therefore considered useful for load reduction scenarios…" is true or accurate. I certainly would not be surprised if the model produced simulations in the right direction—i.e., a reduction in load produces a reduction in concentrations; the model would have to be seriously flawed not to do that, and I do not think that the model itself is that flawed. Nonetheless, one cannot conclude that the model is adequate for the proposed purposes just because it gets the direction of change correct. The results presented in preceding pages do not lead me to think that it can do more than that.

Appendix II

Page	**Comment**
1	*Figure 1*: The map does not "work" in gray scale. The watersheds simply cannot be distinguished in the b/w printed version of the report.
15	*Last paragraph*: The authors are disingenuous in stating that "the predicted flow for the Forest subwatershed has a similar pattern, but *slightly* (italics added) higher peaks than the observed flow at the POLA/POLB stormwater sampling station." The second simulated peak is twice as high as the observed peak; I do not consider that to be a "slight" or "small" difference, and I don't consider that to be "well within acceptable modeling ranges."
16	*Figure 5*: The comparison of modeled and measured flows in this figure also is not impressive. The modeled results completely miss the two-peak nature of the observations.
17	*Figure 6*: Modeled versus observed peak flow for the subwatershed in the figure differ by a factor of five. At least, the text (bottom of p. 16) acknowledges the lack of fit, but the results certainly to not provide validation for the model. *Statement in the first sentence*: "Once the model was calibrated and validated…." This statement makes it seem that everything worked, but as the previous comments indicate, the model really was not validated. I don't think one can say that a model was validated simply because one ran a validation exercise. If the simulation didn't fit the observed data in the validation exercise, one cannot conclude that the model was validated.
19-20	*Figures 7 and 8*: The modeled trends in TSS and measured data are not even close in Figures 7 and 8, and I do not consider these results to be "well within acceptable modeling ranges" as the report states at the end of the first paragraph. The same comment applies to the "validation" in Figure 9 (p. 21).
21	*First paragraph*: Plots like those in Figures A-2 to A-15 in the appendix are almost useless in evaluating the validity of the model. The range of the TSS data is so large that it would be amazing if the model didn't predict TSS concentrations "generally within the range of the observed data."
24	*Figure 10*: Overall, the modeled versus observed concentrations of copper, lead, and zinc in Figure 10 do not represent acceptable fits of the data. The modeled results largely under-predict the initial high concentrations and the double peaks of the modeled results are not to be found in the observed data.
25	*Figure 11*: Similarly, the second modeled peak is not found in the observed data in Figure 11, but the first peak roughly captures the initial observed data.
26	*Table 6*: Are the EMCs flow-weighted or simple averages?
27	*Figure 12*: The comparisons of modeled and measured EMCs are actually quite good for two of the three sites shown in Figure 12 (and awful for the third site), but it is difficult to understand how the EMCs can be as close as shown given the poor match of modeled and measured results in the preceding Figures 10 and 11, from which the bars for the Forest Industries site in Figure 12 were based.
29	*Figure 13*: Same comment applies here as for Figure 10 (p. 24).
31-32	Given that the authors showed previously that they were not able to simulate flows for the Maritime Museum subwatershed, one wonders why they even bothered to model the metal concentrations and loads. Clearly, they were unsuccessful in doing those as well.
52	*Figure 27*: The one data point at the right side of Figure 27 is the "tail wagging the dog." That is, this one datum is driving the regression and is largely responsible for the high r^2. The distribution of the data does not fit one of the basic assumptions of regression analysis—that data are distributed roughly equally across the range of the independent variable.
53	*Tables 13 and 14*: From the magnitude of the standard deviations relative to the means in Table 13, it is clear that the data are highly skewed and not normally distributed. Mean values are not appropriate in such cases. The authors should have log-transformed the data, which likely would

have yielded at least close to a normal distribution. As a result, the values calculated for the "low range" and "high range" in Table 14 are not correct.

54 No basis is presented for the statement "Trace metals were bound to a particle during wet-weather wash off until they dissociated upon reaching the receiving water body." This may or may not be true, depending on dissolved metal concentrations in the receiving water body, the kinetics of desorption, and the mode by which the metals are bound to particles. Not all metals are bound by reversible (ion-exchange-like) processes.