THE BRAKE PAD PARTNERSHIP
Compilation of Technical Reviewers' and Stakeholders' Comments on Watershed Modeling Draft Report
July 22, 2007

Background

On behalf of the Brake Pad Partnership, AQUA TERRA in conjunction with the US EPA has conducted watershed modeling of the environmental fate and transport of copper from vehicle brake pad wear debris and other copper sources in order to better understand how copper travels through the environment and what the relative significance is of copper from brake pad wear debris on copper levels in the San Francisco Bay. The watershed modeling effort is one of three interlinked modeling components at the core of the BPP's effort. An air deposition model was used to estimate the amount of copper from brake wear debris and other air releases of copper that is deposited in the watershed. The results of that model provided some of the inputs to the watershed model. The watershed model, a draft of which you are being asked to review, estimated the relative amount of copper from brake wear debris that is discharged from the watershed in runoff to the San Francisco Bay. That information will provide inputs to the bay model, which will estimate short- and long-term concentrations of copper in the South San Francisco Bay.

The work plan for the watershed modeling effort was reviewed and approved and is available at http://www.suscon.org/brakepad/documentArchive.asp. Final results for the air deposition modeling and for the source release inventory effort can be found at http://www.suscon.org/brakepad/documents.asp.

The Brake Pad Partnership Steering Committee is seeking an independent expert review of watershed modeling results for the environmental fate and transport of copper to ensure that the approach and results of this element of the Partnership's work are technically sound, to determine if there are feasible opportunities to strengthen the presentation of the results, and to help build in-depth understanding of and confidence in the technical studies on the part of the Steering Committee and the stakeholder communities.

Charge to Reviewers

With the aim of meeting these objectives, the Steering Committee is seeking comments that specifically address the following questions:

1. In your assessment, do the modeled results adequately estimate the relative amount of copper from vehicle brake pad wear debris that enters the San Francisco Bay in runoff from the watershed? What improvements, if any, do you recommend for the presentation of results?
2. The original intent of the technical studies was to develop information on the relative impacts of copper from brake pad wear debris in the bay. In your assessment, what factors would need to be taken into account when using the results presented in this

report for assessing the relative contribution of brake pad wear debris copper in the creeks?

3. In your assessment, do the modeled results adequately present the level of uncertainty involved in the amount of copper from vehicle brake pad wear debris that enters the bay in runoff from the watershed? Do the modeled results adequately represent the uncertainty in the relative contributions of copper in runoff from brake pad wear debris and from non-brake pad copper sources?
4. During your review, did you identify anything that could render the modeled results inappropriate for subsequent use in the Bay model as described above? Did you identify any limitations on the use of these results in the Bay model?
5. In your assessment, what factors would be important when extrapolating the modeled results to other geographic areas?

===

COMMENTS RECEIVED

===

From Jim Pendergast, US EPA (July 17, 2007)

P1, second paragraph: This text seems better placed in section 1.2. It doesn't provide background but rather describes what model was used, which seems to be better placed in the discussion of the work.

P2, Table: Suggest deleting the Total Loads column because it is confusing without adding discussion text. We know that we matched the high brakes with low non- BP contribution, etc., which resulted in no real change in total loading. But without saying this, readers will wonder why the total loads did not change. Since the table is designed to show that the BP contribution ranges from 10% to 35% depending on one's assumption on the variability in sources, that is all the table should show here.

P18, Last paragraph: Did we decide not to pay for a BP high – non-BP high and BP low – non-BP low scenario? How much would it cost to run these scenarios? Given the insensitivity of the total watershed load to the bay to the three scenarios run, I'll guess that reviewers may want to see the high-high and low-low scenarios to check the sensitivity on the total load calculation.

P26, L2: Suggest not saying "fall back", but rather say "secondary".

P26, Table: Need to fix the word wrapping.

P26, Last sentence: Ought to give the reason why the Alameda Creek simulation was run.

P26, Last paragraph: Fix formatting to left justify.

Figures 3.2 and 3.3: Very good simulation of higher and middle flow ranges, but why the departure between observed and predicted at the lower flow ranges of Napa River near Napa and Alameda Creek near Niles? The Napa River is more troubling because the departure is significant for the lower half of the flows. Suggest additional text in the report describing this.

P30, L3: Don't need to emphasize NO. Lower case will do.

P31, Third paragraph, Last sentence: What reason do we give to consider these 6 points as outliers? An alternate hypothesis is that the model is missing a significant loading, because the observations are 4 to 5 times higher than the predicted values. I do suggest more text on this topic.

P31, Last paragraph, L3: Suggest replacing "NOT calibrated" with "not generated by further calibrating the model"

P36, Third paragraph: Is there a reason why the high end of BP loads is offset by the low end of non-BP loads? Since the loads were developed independently, this is only a coincidence, right? We ought to say this so that a reader doesn't wonder if the non-BP loads were generated by subtracting the BP loads from a total watershed load. Remember that not all readers will read all the reports.

P37, Table: Check the total row. The sums for "Watershed Load" and "Load to Bay" should differ because not all watersheds modeled drain to the Bay.

P40, Figure 3.9: Check the word wrap on the figure title.

==

From Bob Peters, Akebono-USA (July 18, 2007)

Page 3, Paragraph 3. I'm not familiar with the report cited but it seems to me that those estimates would be based on roughly the same thinking and same modelling that this report is based on. Additionally, it had major involvement from the same individual, so it's no real surprise that the estimates are similar. I don't think this adds much to the credibility of the model.

Page 6, Section 2.1 - when the number of paths from a watershed to a bay have been reduced from many to 1 for the purposes of modeling - how is the background copper concentration (I think in the sediment) of the many small creeks comprehended, or maybe more appropriately, compensated for?

Page 9, Table 2.1 - This isn't the only example - but it illustrates my point. Rain gage measurements are reported to 0.001". I doubt any rain gage is that accurate. I think significant digits need to be considered throughout the report.

Page 17, Paragraph 2. Minor point - just wondering if the assumption about irrigation efficiency is a good one. I'm pretty sure my lawn watering isn't that efficient.

Page 18, Section c. "BPW" is used as an acronym for Brake Pad Wear - I think it would be more accurate to call it BPWD - since it really is the debris we're concerned with.

Page 26, Table 3.1. Text and table are interfering with each other.

Page 32, Figure 3.4. I assume the line is a correlation line and my read is that correlation is pretty poor based on that graph. What is the correlation factor r^2? How can this be considered "good" correlation?

Page 32, Figure 3.4 The plots (not just on this page) are very difficult to decrypt. For the benefit of the folks that aren't used to viewing the plots - I suggest the legends and axes are labeled in English, avoiding technical abbreviation wherever possible.

==

From Wayne C. Huber, Oregon State University (July 18, 2007)

Donigian, A.S., Jr. and B.R. Bicknell, *Watershed Modeling of Copper Loads to San Francisco Bay*, Aqua Terra Consultants, Mountain View, CA, June 21, 2007.

I am pleased to provide this review of the report. Overall, it looks like a good job to me. I will make a few general and specific comments about the report and then address the specific questions to the reviewers. The majority of my general comments are really questions for clarification and do not reflect generally upon the good quality of the modeling.

General Comments

1. I like continuous simulation for the purpose of defining loads to the Bay. HSPF is an appropriate model, run by good modelers. The tools of this study are appropriate.

2. I agree with the "weight of the evidence" means for model credibility in the absence of more data and greater study resources. Once in a while I might quibble with what is "good agreement," but I generally accept the opinion of these qualified modelers.

3. Loads developed from HSPF are driven by the loadings supplied by others. HSPF associates copper mass with runoff events and develops hydrographs and pollutographs for Cu. But I expect that the loads are driven most strongly in the model by the deposition modeling of Pun (2006) and copper release estimates of Rosselot (2006a,b). (I have not reviewed those reports.) In other words, if the loading estimates that drive the model are on target then so will be the HSPF estimates. The fact that HSPF Cu loads are in the same ballpark as spreadsheet estimates (Davis et al., 2000) is good. The good agreement of HSPF TSS loads with estimates of McKee et al.(2003) is also good.

4. Having said this, I do not see a good discussion of sensitivity analysis in the report. What really drives the copper loads in the opinion of the authors?

5. Since load (mass or mass/time) = concentration x flow (volume or flow rate), the long-term hydrographs are also very important. The quality of the observed vs. predicted hydrographs is difficult to judge from Figures 3.8 – 3.11, but looks generally good. The flow-duration curve comparisons (Figures 3.2 and 3.3) look good to me, and a maximum volume error of 9.0% over a 24-year simulation (Table 3.2) is very good.

6. I like the use of best available topographic, land use and other spatial data, e.g., National Hydrography Dataset (NHD), coupled with GIS. I would appreciate a few more references for datasets such as NHD with which I have not worked personally.

7. P. 8, re. volume-stage-discharge functions: Were actual rating curves available for the various streams that would serve the same function? And couldn't stream width be measured directly or estimated from photos or GIS coverage? For instance, some of the "streams" are rectangular concrete channels. Just asking.

8. Regarding precipitation input (p. 8), the report indicates hourly values were used. Were 15-min data available and if so, why not use them? However, with large PERLND and IMPLND subareas, maybe HSPF does not respond much differently to hourly vs. 15-min rainfall. But in most urban modeling scenarios, the shorter time increment rainfall data, the better.

9. I wonder if 15-min rainfall vs. hourly would also affect sediment washoff? My recollection is that sediment is generated in HSPF on the basis of rainfall energy. Would higher intensity, 15-min rainfall generate more sediment than lower intensity, hourly rainfall?

10. I really don't think this matters much here, but Thiessen weighting of nearby gages loses temporal patterns as weighted gages are combined. Another option is to use the closest gage so that temporal resolution is retained.

11. I like the use of the California Irrigation Management Information System (CIMIS) data for evaporation estimates. In the Pacific NW, we have something similar called AgriMet (U.S. Bureau of Reclamation). I am not clear how the CIMIS data were combined with pan evaporation data for modeling purposes. If there were a way to use the data directly in the model, I'd choose CIMIS ET estimates based on the Penman equation and good local meteorological data over pan data in San Jose.

12. Pp. 16-17, re. irrigation. Great idea to include this in the model. How is irrigation introduced into HSPF? As additional rainfall? If so, does this show up on impervious as well as pervious areas? Probably not, but I'm interested in how this works.

13. P. 18. paragraph c: Does the "impervious" land use category include both roadways and non-roadways? Is imperviousness "directly connected impervious area" (DCIA)? It is not clear to me whether loadings differ for roadways and, say, rooftops, and whether such a difference is accounted for, or whether it matters at all in this modeling. I probably just don't understand exactly how the loadings, land uses, and model land covers are coupled. The model is likely calibrated to whatever assumptions were made in the loadings. But I'd assume that roadway surfaces need to be separated from other imperviousness if brake pads themselves impact mainly

roadways?

14. Page 30, paragraph b: Another reason for deviations in flow-duration curves at low flows might be baseflow resulting from irrigation and the many different kinds of urban discharges that occur during dry weather.

15. Is there a reason for dropping the six high observed Cu concentrations in Figure 3.4? Agreement between simulated and observed concentrations is not so good when they are retained.

16. Finally, what is the rationale for the three scenarios (Table 1.1), in which non-BP loads are increased/reduced as BP loads are reduced/increased? Yes, this results in three copper loadings of similar magnitude (56,000 kg/yr), but I don't understand why the scenarios assume compensating loads.

Specific Questions to Reviewers

1. In your assessment, do the modeled results adequately estimate the relative amount of copper from vehicle brake pad wear debris that enters the San Francisco Bay in runoff from the watershed? What improvements, if any, do you recommend for the presentation of results?

I believe the relative amount of copper from brake pads vs. other sources is strongly a function of the loadings applied to the different land uses in the model (see comments 3 and 13). Hence, I believe the HSPF modeling itself is secondary to the loading estimates, at least as far as producing total loads to the Bay is concerned. As noted in the report, the earlier spreadsheet modeling (66,000 kg) comes close to the HSPF annual loads (56,000 kg). The HSPF modeling increases in utility as control alternatives are considered and spatial and temporal patterns might affect them. Regarding the spatial distribution of loads, this is also strongly a function of the land use and loading data entered into the model and also, presumably, the spreadsheet. I don't see any special improvements for presentation of the result, other than, perhaps, some higher resolution (shorter time span) presentation of some of the observed vs. simulated hydrographs.

2. The original intent of the technical studies was to develop information on the relative impacts of copper from brake pad wear debris in the bay. In your assessment, what factors would need to be taken into account when using the results presented in this report for assessing the relative contribution of brake pad wear debris copper in the creeks?

I gather this question deals with the relative impact of brake pad wear on the Bay vs. creeks. Factors that might need to be considered include: point sources of Cu (e.g., industries) that might be located along a creek (greater Cu downstream); sorption and speciation of Cu along the creek; relative runoff from a subbasin directly out of a creek into the Bay vs. runoff distributed elsewhere along the shoreline (discussed in the report a little bit on p. 8).

3. In your assessment, do the modeled results adequately present the level of uncertainty involved in the amount of copper from vehicle brake pad wear debris that enters the bay in runoff from the watershed? Do the modeled results adequately represent the uncertainty in the relative contributions of copper in runoff from brake pad wear debris and from non-brake pad copper sources?

How to assess uncertainty is a bugaboo in most modeling studies. Presumably this was considered in this modeling study through the stated uncertainty in the dry and wet deposition fluxes of copper given in Table 2.8. But the scenarios seem to consider only compensating loading variations (see my comment 16), so I don't really understand how uncertainty is considered here. And I don't know how this would be done other than arbitrarily to vary model parameters, which is often done to produce a range of estimates. On the other hand, I'm not sure what this would prove. A better sense of model sensitivity would help here (see my comment 4), to see where additional focus might be needed.

4. During your review, did you identify anything that could render the modeled results inappropriate for subsequent use in the Bay model as described above? Did you identify any limitations on the use of these results in the Bay model?
Most of my general comments reflect more of a need for clarification than any negative criticism of the HSPF modeling. I think the modeling reflects the spatial and temporal patterns of copper and sediment load to San Francisco Bay to the extent that such estimates can be supported by the data. This is to say, I believe these results are credible. I also believe annual loads might be produced just as well by the spreadsheet model. But HSPF provides the opportunity for good temporal resolution and many more options for controls.

If brake pad contributions of copper are on the order of 10-35% of the total loads to the Bay, it will be interesting to see whether reductions in brake pad copper make much difference in the health of the Bay's sediments. For instance, it isn't possible (I assume) to get rid of 100% of brake pad copper, given the prevalence of older vehicles and long life of brake pads in general. What if BP contributions were reduced by 50%? Could the impact of this be seen in the Bay?

Control options will also depend on the particle size distribution and speciation of copper entering the Bay. If it's all attached to sediment, it will be more easily removable than if more is in a dissolved form.

5. In your assessment, what factors would be important when extrapolating the modeled results to other geographic areas?
Loadings (atmospheric and other) have to be the key since they are the driving parameters in determining loads to the Bay or wherever a similar model is run. Good land use data will also be necessary since loadings are determined as a function of land use. Point sources and upstream sources of Cu will also matter. HSPF will do a good job of modeling the hydrology most anywhere. It's the parameters that drive the quality processes that are most important. And calibration/verification data must always be a part of the study as well.

==

Kirsten Rosselot, Process Profiles (July 19, 2007)

General comments:

I found this to be a well-written report that was a pleasure to read.

The portion of the copper load to the Bay that is due to brake pads, non-brake pad anthropogenic sources, and sediment is going to be of central interest to many of the Partnership's stakeholders. Could you include a discussion of this in your final report? I tinkered around with Tables 3.3 and 3.4 because I wanted to see what portion of the anthropogenic copper loads were due to releases from brake pads. I used a concentration of 25 mg Cu/kg sediment to eliminate the nonanthropogenic load so I could get the total estimated anthropogenic loads to the Bay and from there get the the portion of the anthropogenic loads that is due to brake pads. Is this a reasonable way to determine the anthropogenic fraction? When I do this, I get that the total contribution from brake pad wear debris towards total anthropogenic loads of copper to the Bay for the mid-brakes case varies from 15% (for the Sonoma sub-watershed) to 57% (for the Upper Colma sub-watershed). For the rural sub-watersheds, the brake pad contribution is much lower than for the heavily urbanized sub-watersheds. There are six sub-watersheds whose total copper load to the bay is larger than 4,000 kg/y (for the mid-brakes case). They are Contra Costa Central, East Bay Central, Napa, Petaluma, Santa Clara Valley West, and Sonoma. These six sub-watersheds contribute about 60% of the total copper load to the Bay. It's interesting that some of these sub-watersheds have their largest contribution from sediment (Napa, Petaluma, Sonoma)), some have their largest contribution from non-brake pad anthropogenic sources (Contra Costa Central, East Bay Central), and one has its largest contribution from brake pad sources (Santa Clara Valley West).

Another thing that's of interest to folks who might want to use this study as a springboard for understanding copper in areas other than the Bay area is the fraction of releases that winds up in runoff for various types of land use. Could you discuss average runoff factors for the various land use types in the Bay area?

I would like to see more mass balance-related discussion in the report. For example, I think the runoff factors for releases to directly-connected impervious surfaces depend heavily on the build up and washoff parameters that were developed using monitoring data in the Castro Valley sub-watershed. What happens to the copper released to these surfaces after the buildup function value is satisfied? Physically we know that this copper blows off or is blown off the impervious surface, but is it modeled as if it blows off the impervious surface and lands on an adjacent pervious surface where it is treated like a release to pervious surface, or is it modeled as if it goes to a sink and becomes unavailable? Could you prepare a mass balance of copper in the watersheds based on the source term inputs and the modeled outputs?

Specific comments:

First two paragraphs, page 3: Does this agreement occur because the model was calibrated to produce these results? Or does this agreement demonstrate that using parameters that were developed from calibration exercises in one watershed throughout the greater watershed produced results that are in agreement with other research? This issue returns in the 3rd and 4th full paragraphs of page 36.

Table 2.1 page 9: There's an asterisk in the second column heading that doesn't have a footnote. What does the "(S)" in the third column heading mean? Please define "MFACT" and "CIMIS" in footnotes. What is the Zone 8 MFACT (the last column)? I see that it looks like Gage

MFACT normalized to Zone 8 but I don't know what it was used for, unless it's the multiplier instead of Gage MFACT.

Paragraph just below Table 2.1, p. 9: define "ET." "ET" shows up again on page 16 and I think it should be defined again there.

Third paragraph, p. 12: Could you discuss whether the choice of land use dataset would be expected to influence the results in any particular direction?

First paragraph, p. 17: 0.8 doesn't seem small enough, observation indicates that more overwatering than that occurs in urban landscaping. How sensitive are the modeled results to this number? IIf this number is too high, what effect would that have on the modeled results? Would it make the runoff from developed land too low? Is this factor the same for crops as it is for landscaping on developed land?

Item d(e) on p. 18: Sediment associated losses don't actually show up in Tables 2.6 to 2.7. I would suggest you switch to plain bullets or something besides lowercase letters for the items in this list.

Item e on p. 18: I want to make sure these releases are being treated as a non-constant point source load direct to the streams and that there is no buildup function involved.

Items 1, 2, and 3 at bottom of p. 18: There needs to be a paragraph or so explaining the rationale for choosing to run the cases this way.

Second full paragraph p. 31/Figure 3.4 p. 32: I think there should be some discussion of the uncertainty in the measured copper concentrations. Also, the modeled results appear to be much flatter than measured concentrations. The slope of the modeled results does not appear to be near unity, either – based on an eyeball assessment that slope looks greater than one. I think there may be source term reasons why the modeled results are flatter than measured and I would like to see some discussion about these reasons, which include 1) the releases were applied evenly throughout the year when in reality releases can be seasonal, episodic, or vary by time of day, and 2) releases were modeled as if they were applied evenly throughout the land use types for each sub-watershed when in reality they may have occurred more in one part of the subwatershed for that land use area than another. The effect of these temporal and geographic variations in releases might be captured in grab samples, but they wouldn't be in the modeling. That doesn't mean the model didn't do a good job of estimating annual copper loads to the bay, but it may be important to understand that the way things were modeled may mean that modeled peak concentrations would not be expected to match measured.

==

From Kelly D. Moran, TDC Environmental (July 19, 2007)

This memorandum contains my own comments as well as comments from a BASMAA reviewer on the draft watershed modeling report prepared for the Brake Pad Partnership (BPP).

Overall, the way the watershed modeling was implemented appears to provide a reasonable estimate of copper loads conveyed from San Francisco Bay Area watersheds to San Francisco Bay. When supplemented with additional information to address our comments below and any substantive comments from the BPP Scientific Advisory Team, we believe that the watershed modeling report should be suitable to serve the needs of the BPP.

Our substantive comments are listed below. During our review, we identified a few minor editorial items that we expect AQUA TERRA will wish to address while completing its report. These are listed in an attachment.

- Anthropogenic copper contributions. We would appreciate it if the report were amended to include additional information to clarify the portion of the copper loads attributable to anthropogenic sources. Copper control strategies necessarily focus on anthropogenic copper. Clarification of the anthropogenic contribution will help BPP stakeholders understand the relative importance of the brake pad copper contribution. (We appreciate that the BPP project manager has already initiated efforts to address this comment.)
- Approach to modeling of brake pad releases in relationship to other copper sources. The approach to addressing the range of the uncertainty of the copper source estimates is based on an assumption that does not seem reasonable—that the errors in the release estimates for copper sources other than brakes are directly linked—and inversely proportional to—the errors in brake pad copper release estimates (called "anti-variance" below). Despite our concerns about this approach, we hope that it is possible to obtain the information necessary for the BPP from the existing modeling runs. To address this issue, we suggest the following actions:
    - Consult with the Bay modeling contractor (URS) to determine whether an approach to handling the Bay modeling scenarios without assuming anti-variance in release estimates can be developed with existing information.
    - Clarify the reason for the selection of the "anti-variance" assumption in the text.
    - Revise or remove the text in the third full paragraph on Page 36, which relies on the anti-variance assumption.
    - Discuss the additional potential release estimate variance options (e.g., high-brakes and high-non-brake sources, variance of individual non-brake sources) qualitatively.
    - Consider additional description of the error estimates for the primary copper load estimate comparison value, which is from Davis et al. 2000. In particular, we believe that it would be helpful to provide the estimated range of the load estimate ("low" and "high" estimates can be obtained mathematically from information in Table II-7 of Davis et al. 2000). We also suggest further clarification of the preliminary nature of the SFEI copper load estimates, e.g., the SFEI report states, "[t]he simple model used in this report was intended only to provide preliminary estimates of emissions….(Davis et al. 2000, page 69).
- Inherent shortcomings of modeling approach. The report should list and discuss the implications of the inherent shortcomings of the modeling approach. The most important shortcomings to list are those that are likely to be relevant to (a) the Bay modeling or (b) interpretation of the report by stakeholders. For example:
    - The modeling systematically underestimates peak copper concentrations/loads. This result is an expected shortcoming of the modeling approach, which entails averaging out of sources that may occur as "events" (e.g., swimming pool

discharges). In conjunction with the Bay modelers (URS), the BPP should explore what this means for the Bay modeling and for comparison to Bay water quality criteria, and the BPP should ask AQUA TERRA to include appropriate recommendations in its report (e.g., explain how to adjust for this shortcoming and/or or to understand its implications).

  - Similarly, the modeling design was focused on providing information that will allow Bay modelers to assess annual or seasonal changes in the Bay—it is not appropriate for evaluating short-term copper concentrations in creeks. Because the modeling was approached on a Bay-wide basis, the values for individual modeling watersheds are necessarily approximate. (More locally-specific release estimates and/or modeling parameters could significantly affect results for higher-resolution modeling at smaller geographical scales or time steps).

We believe that the following changes would improve the clarity of the report:

- <u>Table 1.1, Summary of copper loads for modeled scenarios</u>. This table does not connect well to the text, which refers to 6 model runs. This table would be more useful and clearer for the reader if it specifically showed the three non-brake runs. While it is possible for a clever reader to tease this information out of the table, the summary would be clearer to stakeholders if it were formatted such that it is obvious that there were 6 model runs. Table 3.3 is a good example of how the six runs can be clearly presented.
- <u>"Best Estimate</u>." It would be helpful to label the "mid-brakes"/mid-other copper sources scenario in a manner indicating that it is the central estimate of loads from all copper sources (e.g., the "Best Estimate" or "Central Estimate" scenario).
- <u>Explaining the value of BPP-conducted studies</u>. In Section 1.3, it would be helpful to provide a short explanation of the framework of the BPP studies (i.e., the boxes in Figure 1.2) that identifies the studies that were specifically prepared to provide input data for this watershed modeling report. We believe that is important to explain that the BPP studies were completed with the cooperative oversight of the BPP Steering Committee and were peer reviewed by the BPP Scientific Advisory Team. The cooperative approach to preparation of these input data distinguishes the BPP-supplied studies from other information sources that happen to have relevance to this work. The cooperative approach and the peer review lend credibility to the BPP-supplied inputs.
  When BPP-funded studies are mentioned for the first time in the body of the report, it would be helpful to identify them as BPP studies. For example, where the report draws on the BPP studies in the first sentence of Section 2.5, it would be helpful to link back to the explanation of the credibility of the source (i.e., by adding the phrase "…as described in Section 1.3").
- <u>Clarification that runoff from the delta is not part of the modeling</u> (Page 6, Section 2.1, first paragraph). It would be helpful to clarify that the model does not address runoff flowing into the study area from the Sacramento/San Joaquin River delta.
- <u>Clarify references to Bay Area Hydrology Model</u> (Page 25, letters "a" and "d"). We believe that this would be more clear for Bay Area readers if these two items were combined into one and included the context (i.e., "HSPF parameter development for the Bay Area Hydrology Model design tool through calibration studies in Castro Valley and Upper Alameda Creeks (AQUA TERRA 2006) and in two Santa Clara County creeks (Clear Creek Solutions in preparation; parameters listed in Clear Creek Solutions 2007)."), with appropriate editing of the first sentence in the following paragraph to

reflect this change. To be completely accurate, it would be appropriate to add a footnote after this mention of “Upper Alameda Creek” clarifying that the subwatershed referenced is a subset of the BPP modeled watershed with the same name.

- Clarify data source (Page 31, first paragraph, last sentence). Please clarify the source(s) of the “limited available instream copper concentrations” and add appropriate citations to the reference list.
- Clarify references to watersheds. BASMAA asks that text, figure captions and/or tables use the term "BPP modeled watershed(s)" (as was done in Dufour and Cooke 2005) to minimize potential confusion with different watershed areas delineated by other organizations that have identical or similar names.

ATTACHMENT

The following minor editorial items were identified during our review:

- Page 1, first paragraph, second sentence. Delete “San Francisco Bay Area”. The project is the “Brake Pad Partnership,” not the “San Francisco Bay Area Brake Pad Partnership.”
- Page 1, first paragraph, last sentence. It would be preferable to use a word other than “deposition” to refer to the quantity of copper transported to San Francisco Bay, because “deposition” is the term used for the air pathway in this work.
- Section 1.2. Consider using this section to create an Executive Summary. The summary is quite helpful, but we found its placement confusing. We found the text in this summary a bit confusing to follow (it seemed to jump back and forth between assumptions, results and validation/calibration comparisons with outside datasets); perhaps more paragraph breaks or subheadings could be added to assist the reader.
- Page 2, Section 1.2, second paragraph, third sentence (and throughout the report). Please replace “Alameda study” with “study prepared for the Alameda Countywide Clean Water Program (ACCWP).” Similarly, please clarify other ambiguous uses of the term "Alameda" (e.g., page 26, paragraph 1, which could be interpreted as referring to Alameda Creek, another geographical area in the East Bay, or to several possible agencies).
- Page 3, last paragraph, second sentence. The final verb should be “are” rather than “is”. (“…while the values for the East Bay watershed are 48% to 53%.”)
- Page 4, Section 1.3, last sentence. This sentence contains a “he” that should be a “the.”
- Page 5, Figure 1.2. Recommend addition of a source for more context for this figure (at a minimum, a link to the BPP web site).
- Page 8, Section 2.2., first paragraph, first sentence. The verb should be plural (i.e., “include” rather than “includes.”)
- Page 9, Table 2.1. This table has a few formatting issues. It may need footnotes. There is a “*” in the header of the Rainfall column that doesn’t connect to a note below the table. There are two columns for “Gage Annual Rainfall” and “Gage MFACT,” with two values for Upper Alameda watershed, and one value for the other watersheds. It appears that 2 rain gauges were associated with Upper Alameda watershed. A short footnote to the table would be helpful to clarify whether this was to fill gaps in individual gauge records, or for other purposes.
- Page 26, last two paragraphs. Paragraph formats need fixing (one wraps around table; the other is centered, rather than left justified.)

- Page 31, first line. Replace “Cleanwater” with “Clean Water” (two words).
- Pages 43-44, references. At least one reference is missing (ACFCWCD 2005).
- Page 43, references, final reference on page. Correct spelling of “Leatherbarrow”.

===

From Lester McKee, San Francisco Estuary Institute (July 19, 2007)

P = Page number, Par. = paragraph number, L. = line number

P2, section 1.2. Please make it clear in the opening paragraphs if you are modeling total copper or dissolved copper.

P2, Par4. Given the model objective if to estimate loads of Copper from urban areas, I am hopeful you included the gauges that represent greater urbanization (only two come to mind: Guadalupe River at San Jose/Hwy 101, San Lorenzo Creek at San Lorenzo)

P6, Par5. …small streams and creeks (add and storm drains).

P8, Par1. Define major. For example there are 14 “major reservoirs in the Bay Area that capture >20 sqmi area (see SFEI report Davis, McKee et al., 2000).

P9, Par1. …with similar rainfall total. Do you mean daily rainfall totals. Did you check the results of this methodology by trying to predict the hourly rainfall of a station with hourly rainfall data? This would provide you with an understanding of your %error both in timing and in magnitude.

P9, Par3. Pan evaporation. This section needs better description so that the reader is not left feeling disconnected from your modeling procedure. How exactly did you estimate evaporation on an hourly basis? What about impervious surfaces? How did you cope with roofs, roads and parking lots? Did you check the outcomes? Please demonstrate the methods and assumptions gave reasonable input data for the model. Evapotranspiration is likely 70% of the total water budget so I suspect it is a really important term in the model. Was the range of the estimated data reasonable over a single day and over the longer term? If not, what are the potential consequences of the assumptions and errors?

P12, Par3. …and associated characteristics. What characteristics? Please be specific. …and associated activities. Please add some references to the end of the paragraph so the reader can follow from where you are deriving your assumptions.

P12, Par4. I agree – impervious surfaces along with evapotranspiration are very important parameters in your model. Again, I am left having to just trust that you mode reasonable decisions. Please add some more detail in this very important section. How did you actually estimate total imperviousness and disconnected impervious surfaces? Did you use land use data, if so which data and how was it classified? You reference Defour and Cooke, 2006. Please add their table 10 in this report so the reader can see what assumptions where made and how all the calculations were done. Please discuss the potential impacts of these assumptions to the model outcomes.

P16, Par1. 23% seems low. Do you mean the developed area that received irrigation? Overall based on ABAG land use data from 1995, the area in urban and ag use in the BA was more like 44% (see SFEI report Davis, McKee et al., 2000). Perhaps the differences in the model area are the cause of the apparent discrepancy?

P17, Par4. Please add the units from Pun and the units you converted to for HSPF in to the text and better explain how you did the conversion. Please comment of the errors and potential effects on the model output of the assumptions that you had to make.

P19. Figure 2.6. Given that the model objective is to understand urban copper runoff, more effort needs to be put towards disaggregating the urban source and transport characteristics and less effort on the 4 non urban classes (Ag, Grass, Scrub and For.)

P23, Table 2.8. These are lower that Tsai et al, 2000 (RMP study). Please comment on how did your estimates of loads compare with the work of Pam Tsai et al. (see http://www.sfei.org/rmp/reports/air_dep/tracemetals/AD_TMFinalReport.pdf). I realize Pam's work was primarily focused on deposition to the Bay not to watershed surfaces but she did estimate deposition to watershed surfaces as well. And her concentrations in rainfall are certainly valid.

P24, Table 2.9. Your concentrations in wet deposition here appear lower than Pam Tsai's work by a factor of 2 or 3 (see page 8 in http://www.sfei.org/rmp/reports/air_dep/tracemetals/AD_TMFinalReport.pdf). Please add a comment of why that is so and the potential impact to the model outcomes.

P28, Figure 3.2. I was hoping to see a better calibration of the urbanized San Lorenzo system. Please comment on the potential impacts of the over estimation in the two urbanized systems (San Lorenzo and Guadalupe).

P30, Par b, c, d. Please define "good" in this report. I see the model results for hydrology being out be 30 – 50%. I see both over and under estimation that is apparently based on degree of urbanization. I realize that +/- 50% is not considered bad, but I am left to ponder what causes the over or under estimation and therefore I am not sure I agree with "very good" without some further commentary.

P31, Par1. It would be more helpful to see scatter graph of observed versus modeled data for TSS so that the reader could see if the model is performing consistently over the range of observed concentrations. Please add the scatter plot. P31, Par 2 come part way to achieving this.

P31, Par3. Do you mean San Lorenzo Creek? San Lorenzo River is not within you model boundaries (also Figure 3.6). I have collected three years of high flow data collection to total Cu at Hwy 101. JC used this data to verify his model. Is that factored in to your model as well? If so please reference SFEI and my reports and provide more detail about what was done.

P31, Par4. …will provide sound technical basis for performing the alternate scenario runs. You may be right, but I am not convinced. I think adding some more scatter plots would help. However the main addition that would help would be a sensitivity analysis of errors. I suggest that such an analysis should be completed to provide validity to you statement.

P37, Table 3.3. The loads for Santa Clara Valley Central and West represent an area of about 335 sq mi. You loads of sediment and Cu are 75554 tonnes and 7037 kg. The loads of sediment and Cu I have measured during WY 2003 (about average flow) in Guadalupe River at Hwy 101 were 10,806 tonnes and 820 kg for an area of 160 sqmi about 91 sqmi of which is downstream from reservoirs. No matter which way I scale up the sediment or copper loads, it appears your model over estimates Cu load for that part of the Bay Area. Sediment is over estimated by at least 1.9 times and Cu is over estimated by at least 2.3 times (see McKee et al., 2005 for the Cu loads http://www.sfei.org/watersheds/reports/409_GuadalupeRiverLoadsYear2.pdf). Please comment the cause of the discrepancy and on how this might impact the usefulness of the model.

P38, Table 3.4. Please clarify that each column represents the low best and high scenarios. Please comment on any other literature that estimates BP contribution to urban runoff. For example, perhaps the modeling done in Stockholm made these estimates. Is 35% reasonable? Given you have potentially underestimated atmospheric contributions, perhaps 35% may be the upper estimate?

P43. The report lack a discussion of potential errors associated with the assumptions. It would benefit from a sensitivity analysis, a conclusion, and section describing data gaps that would improve future modeling efforts for the Bay Area.

===

From Robert Ambrose, US EPA (July 19, 2007)

Reviewer Response

I'll start with some general comments and questions about the report and the modeling work before addressing the specific charge questions. First, I want to express my great confidence in the technical expertise of the report's authors, Tony Donigian and Brian Bicknell. They are certainly among the best in the world in applying HSPF to analyze watershed hydrological and pollutant response. Second, I acknowledge the coordinator's observation of the strong effort to get this draft report together in time for review, and that certain items would be added and corrected in the final report.

I have two or three main technical questions about this application.

Modeling of Land Use
The report makes it clear that individual land uses were considered, but it does not make clear if they were modeled explicitly or implicitly in the assignment of watershed parameter values. I have always thought of HSPF as a lumped parameter model as applied to individual

subwatersheds, but I understand that there are ways to model individual land uses. Were land uses lumped and modeled explicitly within each watershed? That would make sense to me, but it is never stated that this is how the modeling was conducted. If land uses are not modeled this way, how is the information in Table 2.4 used?

<u>Assignment of Hydrological Parameter Values</u>

On page 13, the last paragraph states that model parameters for each watershed and land use category were assigned based on the soil classification (one of 4 HSGs). In the last paragraph on p. 25, the report states that hydrology model parameters were based on soils and slopes. These two paragraphs imply to me that the model would use the same parameters for forested, grassland, developed, and agricultural land uses if the HSG and slope is the same. That wouldn't seem right to me. The report should state clearly the methods used to account for individual land uses within each watershed. A table summarizing model parameters by land use and watershed would be helpful.

<u>Modeling Copper Fate in Watersheds</u>

There is no clear description of how the model handles the fate of copper loads deposited to the watershed. A section should be added summarizing how this is modeled. This could mostly be copied from the HSPF technical manual. The report should also list the sediment and copper parameters used in this application, and describe how their values were derived. Text on p. 31 mentions washoff parameters, but no equations or values are given.

On p. 18, item "e" states that "rain-dependent releases are represented in the model as originating from impervious surfaces based on a calculated concentration and the runoff volume from those surfaces for each watershed." Though it's not entirely clear, I assume that these releases are not meant to represent runoff and erosion of atmospherically-deposited copper from the watershed (such as wet and dry deposition and BPW roadway releases), but rather the "architectural and industrial 'rain-dependent' releases" mentioned in item "d.d." This should be clarified.

On p. 18, item "c" states that roadway releases of copper were adjusted to account for losses due to street sweeping. The adjustment method (equation, algorithm, parameter values, etc.) should be described. The amount of copper reduced due to street sweeping must go somewhere in the watershed. How were those loads handled? Were they added to wastewater loads somewhere?

<u>Modeling Results</u>

As noted, the flow calibration seems good to very good. The general pattern TSS and copper loading in Castro Valley Creek, Guadalupe River, and San Lorenzo River seem good, except that high loadings in Castro Valley Creek are not captured. I don't believe these are "outliers," as stated on p. 31, as there are 6 of these values. The implications of this apparent bias should be addressed.

<u>Charge Questions</u>

<u>1. In your assessment, do the modeled results adequately estimate the relative amount of copper from vehicle brake pad wear debris that enters the San Francisco Bay in runoff from the watershed? What improvements, if any, do you recommend for the presentation of results?</u>

I would judge the modeled estimates of the relative loadings to be adequate. Previous estimates were that brake pad depris constitutes about half of the copper loading to the Bay, whereas this study centers on about a quarter. I do have questions about how copper lost in street sweeping was accounted for, and whether it might enter the Bay in other waste streams.

<u>2. The original intent of the technical studies was to develop information on the relative impacts of copper from brake pad wear debris in the bay. In your assessment, what factors would need to be taken into account when using the results presented in this report for assessing the relative contribution of brake pad wear debris copper in the creeks?</u>

I would like more information on the differences in loading estimates in this report, which centered on 23%, versus earlier work by Carlton and Cocca that centered on about 50%. I also would need to know more about how street sweeping was handled.

<u>3. In your assessment, do the modeled results adequately present the level of uncertainty involved in the amount of copper from vehicle brake pad wear debris that enters the bay in runoff from the watershed? Do the modeled results adequately represent the uncertainty in the relative contributions of copper in runoff from brake pad wear debris and from non-brake pad copper sources?</u>

I believe that the results showing high, medium, and low estimates give an adequate representation of the range of uncertainty in the relative contributions of copper from brake pad wear (10% to 35%). I'm not sure about the accuracy of the overall loading estimates, since the modeling took overall loading estimates as given. I'd be a little concerned about the under-simulation of high copper loads from Castro Valley Creek.

<u>4. During your review, did you identify anything that could render the modeled results inappropriate for subsequent use in the Bay model as described above? Did you identify any limitations on the use of these results in the Bay model?</u>

One of my main concerns is how land use patterns affect results. These concerns might be alleviated by the authors with some further explanation of procedures. At present, I'd have to caution use of this model to explore scenarios with different land use patterns.

<u>5. In your assessment, what factors would be important when extrapolating the modeled results to other geographic areas?</u>

At present I couldn't recommend extrapolating modeled results to other geographic areas. I would need to understand more about how the model was applied to individual land uses, and how parameter values were derived. If important (sensitive) model parameters represent lumped processes and are calibrated, then the model must be considered descriptive of present conditions in the modeled watersheds. This is still useful, of course.

==

Ken Schiff, Southern California Coastal Water Research Project (July 19, 2007)

The stated goal of this paper is to develop a watershed model of flow, sediment, and copper loads to San Francisco Bay. The uniqueness of this project is the linkage with atmospheric deposition, and more specifically, the atmospheric deposition associated with brake pad debris wear. A wonderful question and fun challenge indeed! Below are a few of the major items I saw that could improve the report.

I assume that this watershed modeling is an extension of the original watershed modeling conducted by Carleton for Castro Valley Creek. If so, a quick comparison of the two models would be helpful. Overall, the model setup is very straightforward and the typical model parameters are from the regular sources (i.e., NHD+, DEM, local rain gages with extrapolation, NLCD, etc.). What's unique are the BPDW source terms. To the person who has not read the atmospheric deposition component by Pun and Rosselot, a good summary here would be very helpful, especially since these terms are not specifically defined. Roadway release, although variable, appears to be empirical data. Atmospheric terms are modeled. Two issues that are not addressed in the report:

1) Are there non-BPDW deposition rates for both wet and dry conditions?
2) Of the non-BPDW releases, a very large portion may be from historically deposited BPDW that was either resuspended (for atmospheric terms) or washed off (for nonatmospheric terms). This is an important element as we have found a large fraction of deposition is from nonspecific dusts.

The modeling calibration and validation deserve a bit more effort. The authors begin the most important section (3.0 application and results) with a paragraph of excuses. This does not give the reader confidence in the results. The fact that such a high profile project would not provide adequate time and resources (in the authors opinion) for thorough calibration and validation is startling. In fact, looking at the "QA checks" leads me to believe that some calibration and validation would definitely help. Some of the larger items I noticed:

1) Flow duration curves match well (fig 3.2, 3.3), but little other hydrologic validation is provided that could significantly influence sediment or copper concentrations (i.e., daily mean flow, daily max flow, time to peaks, etc.)
2) Units in table 3.2 need to be examined
3) The TSS and copper plots do not show a "remarkable degree of consistency with observations" to me. I do not see much relationship in the simulated concentration and observed data points in the long-term simulations. In fact, the correlations of observed vs. simulated data for copper in Fig 3.4 looks doubtful, even with the outliers removed. I would have the authors double-check the slope and r^2.
4) The observed copper data vs. long-term copper simulations in figs 3.5-3.7 are insufficient for me to evaluate validation. I recommend more scatter plots and/or summary statistics of accuracy, precision and bias.

Although I am somewhat harsh in some of my criticisms, I do think this is a worthwhile and important effort. I encourage the authors to spend some more time performing the necessary calibration and validation steps. If this extra effort requires additional resources and time from the sponsors, it will be a valuable investment. Even if the numbers change little, the extra confidence it provides the reader will be important.

One last issue, albeit nontechnical. Who is it that will decide that the BPDW fraction is large enough to warrant management action? Is 35% too much? Is the action threshold lower (say 10%) or does BPDW have to be the majority (>50%) of the load? These policy decisions are often times equally as important as the scientific information.   Sometimes agreeing to action levels beforehand can help avoid miscommunication at the end of the study.

===

John Sansalone, University of Florida (July 22, 2007)

The reviewer examined the report in the context that the draft report will be a component step towards building a comprehensive calibrated/validated model of copper loads to San Francisco Bay. In the report the authors clearly indicate the constraints they worked with; specifically lack of data, lack of resources and lack of funds. As a result, the HSPF model developed in the study was not calibrated and validated. It is assumed that all parties accept that the results and conclusions are developed from an uncalibrated/unvalidated model due to the reality of the constraints identified. The report title should reflect these limitations accordingly. As a draft report, the report appears to be a reasonable draft summary, although the original scope and constraints are not directly known by this reviewer.

Since the study and report represents an important evolutionary step that builds a foundation from the previous literature and data collection, it is of critical importance to fully details the input parameter values and ranges and their specific sources in a series of tables within the report. This is an organizational exercise since all of these input parameters were required to run the model. Large databases such as rainfall should have their specific source identified, and the actual digital data should be provided to the sponsor in an organized and clear format on a CD or similar. (This may be part of the deliverable package to the client, but is not clear from the report). Input parameters that were assumed or estimated should also be identified in an organized format of tables based on model algorithms for watershed or catchments and references and rationale identified. The use of input data and parameters is discussed throughout the report but without the ability of the reviewer or client to assess the validity and basis of input data and parameters.

In general the reviewer finds the report language replete with lack of precision and lack of clarity for an engineering document intended to build the knowledge base. There are a number of grammar and punctuation corrections that are required; and can be identified through careful editing. Again, this observation is important if the report is to have lasting value in the evolutionary process. As with the language, the graphical results could also need improved clarity. The graphics are not report quality.

There are a number of specific comments. These comments are generally organized from the beginning to the end of the report.

The authors indicate that HSPF is considered a premier model. Can the authors provide an independent reference to demonstrate “premier”. The authors should concisely describe the other models they considered and why these models were not selected in favor of HSPF. This rationale is important as others undertake parallel and future modeling efforts.
The authors have properly included an expanded database with 20 precipitation gages given the micro-climate variability of the watersheds. However the watershed components do not appear

to reflect a database of equal quality and detail in order to translate rainfall to runoff. It would be helpful if the authors create input parameter tables with quantitative values and units, parameters ranges, reference sources so that the client and future investigators can recognize input parameter strengths and weaknesses. Future work can rely on the defensible strengths and work towards improving weakness in the input parameters.

It is unclear if the authors used GIS data to generate a carefully disaggregated set of hydrologic functional units (HFU) for each land use. Each HFU can then have hydrologic properties identified such as depressional storage, infiltration parameters based on physical soil properties, etc. There should be tables of HFUs and a drawing delineating and quantifying HFUs for each watershed.

The authors indicate the lack of data and resources. The authors should identify the specific data needs and resource needs for the benefit of future investigations. Again this can be best carried out in tables and with a description of these needs.

The authors should quantitatively describe their methods of limited calibration and the quantitative results of limited calibration.

The authors should quantitatively describe their "consistency" and QA checks in terms of both methods utilized and quantitative results that illustrate the consistency and QA checks. Some of these are illustrated in Figure 3.2 and 3.3.

It is noted that while two modeling methods provide similar load results, that does not necessarily imply that this represents reality without an objective yardstick.

The text and figures are inconsistent regarding suspended sediment and TSS. What is being measured is particulate matter in aqueous solution. Suspended sediment may be measured as TSS (a method of measurement) but TSS is not an accurate representation of suspended sediment. The proper terminology is suspended sediment measured as TSS. In addition suspended sediment should not be confused with the SSC methodology, which I suspect was not utilized in most of the data collection.

There should be a list of symbols and abbreviations at the front of the report with definitions and units provided.

Nearly all of the figures need graphical and clarity improvement in order to provide clear, report quality figures. This is especially true for the area maps, which in a number of cases are missing scales and north arrows. In some cases map boundaries and sub-delineations should be shown more clearly. This should not be a problem given the GIS capabilities. Each map caption should provide references for the sources of information.

How do model results respond to various levels of watershed and loading parameter discretization? This parametric evaluation is important not only for copper loadings, but for hydrologic parameters and watershed hydrologic functional unit parameters.

The study talks about data types – define data types.

The authors should indicate why it is necessary for the model setup, “by necessity” to aggregate all watershed drainage.

The use of indefinites and personal pronouns, at least in this reviewer’s opinion, are not appropriate for an engineering report.

The use of a single level value for multiple streams should be justified with a reference. This is the case for many, many assumptions made in the report and model.

The storage characteristics of the reservoirs should be summarized in a table.

Identify how many of the study stations were updated with respect to the total number of stations utilized.

The HSPF hydraulic parameters should be provided in table format with references.

Briefly describe methodology for “actual simulated” ET. Also briefly describe what procedures were used from the Alemeda County study and why they are applicable to this study.

Describe infiltration algorithms and how these were related to soil properties and moisture conditions. For example the use of physical soil properties to determine gravitational and soil suction components of soil hydraulic conductivity. How were soil properties delineated across catchments?

Is the mean the representative statistic for slope – are slopes normally distributed; if not, use the median.

Provide a basis and reference for the irrigation efficiency factor of 0.8, this appears high. Identify the copper loads from the irrigation water.

Are annual historical precipitation values normally distribution, so that a mean annual value is representative of the distribution?

How is the overall long-term mean annual precipitation shown to demonstrate how the mean varies?

Describe and quantify how losses due to street sweeping were determined.

How was a “best estimate” of copper contributions determined?

Figure 2.6 could use more work; I cannot put my hand on it, but it is not particularly useful in the present form.

Tables should be report quality.

Describe and illustrate statistical comparisons.

While weight-of-evidence approaches are described in the Guide, demonstrate how the approach was used in this study and why it is deemed valid in lieu of calibration/validation.

Be specific, what are the sediment and copper parameter values?

Figure 3.2 and 3.3 are important but the graphical quality is weak. How do authors reconcile the large differences are the far more frequent low flows?

Consistency and QA checks are very unclear in the report and not detailed or quantified.

Demonstrate the decision to throw out 6 of the data points is statistically valid; there are specific test to demonstrate this. Why 6, why not 5 or 7? If the data cannot be thrown out statistically they should be included.

Describe all calibrations quantitatively.

What is a “remarkable” degree of consistency? This is probably inappropriate for a quantitative engineering report.