CLEAN WATER / RMP

NUMBER 788
MARCH 2016

Regional Watershed Spreadsheet Model (RWSM) Year 5 Progress Report

Prepared by:
Jing Wu, Alicia Gilbreath, and Lester McKee

For
Regional Monitoring Program for Water Quality in San Francisco Bay (RMP)
Sources Pathways and Loadings Workgroup (SPLWG)
Small Tributaries Loading Strategy (STLS)

SAN FRANCISCO ESTUARY INSTITUTE | AQUATIC SCIENCE CENTER
4911 Central Avenue, Richmond, CA 94804 • p: 510-746-7334 (SFEI) • f: 510-746-7300 • www.sfei.org

Regional Watershed Spreadsheet Model (RWSM)

Year 5 Progress Report

Prepared by

Jing Wu, Alicia Gilbreath, and Lester McKee

San Francisco Estuary Institute, Richmond, California

On

March 20th, 2016

For

Regional Monitoring Program for Water Quality in San Francisco Bay (RMP)
Sources Pathways and Loadings Workgroup (SPLWG)
Small Tributaries Loading Strategy (STLS)

Preface

This continuation of the development of the Regional Watershed Spreadsheet Models (RWSM) was completed with funding provided by the Regional Monitoring Program for Water Quality in San Francisco Bay (RMP). This report is the fourth on this project and will be succeeded by a final report and model publication sometime in later 2016 or 2017. This report provides a short update to facilitate review of model development.

Acknowledgements

We appreciate the support and guidance from members of the Sources, Pathways and Loadings Workgroup (SPLWG) of the RMP. The detailed work plan behind this work was developed through the Small Tributaries Loading Strategy (STLS) team and updated each year with SPLWG input. Local members on the STLS at this time are Arleen Feng (for the Alameda Countywide Clean Water Program), Bonnie de Berry (for the San Mateo Countywide Water Pollution Prevention Program), Lucile Paquette (for the Contra Costa Clean Water Program) and Chris Sommers (for the Santa Clara Valley Urban Runoff Pollution Prevention Program); and Richard Looker, and Jan O'Hara (for the Regional Water Board). We received workplan review and technical guidance over the past five years from Michael Stenstrom (UCLA) and more recently in this past year from Peter Mangarella (Retired from GeoSyntec). Review comments on the report and the scientific work behind it were provided by Arleen Feng, Lisa Austin, Bonnie de Berry, Nick Zigler, Richard Looker, Kelly Moran, and Barbara Mahler.

Suggested citation:

Wu, J., Gilbreath, A.N., McKee, L.J., 2016. Regional Watershed Spreadsheet Model (RWSM): Year 5 Progress Report. A technical report prepared for the Regional Monitoring Program for Water Quality in San Francisco Bay (RMP), Sources, Pathways and Loadings Workgroup (SPLWG), Small Tributaries Loading Strategy (STLS). Contribution No. 788. San Francisco Estuary Institute, Richmond, California.

Executive Summary

The Regional Watershed Spreadsheet Model (RWSM) is a regional-scale planning tool developed primarily to estimate long-term average annual loads from the small tributaries surrounding San Francisco Bay, and secondarily to provide supporting information for prioritizing watersheds or areas within watersheds for management actions. The RWSM has been in development since 2010. In previous years, obstacles to calibrating the PCB and Hg models included issues associated with GIS data underlying the Model, a simplistic model structure and parameterization, lack of available empirical data for calibrating concentrations associated with specific land uses and source areas, and the high uncertainty with the empirical calibration dataset. The modeling effort in Years 2014 and 2015 (Year 4 and 5 of the model development) focused on improving the model calibration to attempt to resolve these issues. The major changes included switching from a sediment-based model to a water-based model, elimination of double counting of source areas on top of general land uses, changes in the model calibration approach, and changes to the land use grouping.

The PCB and Hg models were calibrated after making these modifications, and the calibration results were evaluated using a two-step method: (1) examining the calibrated PCB and Hg concentrations for each land use group; and (2) comparing the modeled and observed concentrations and loads for the calibration watersheds. The assessment of the calibration results indicates that the PCB model calibration appears improved over previous models and reasonable enough to move on to the next step of estimating regional loads, while the Hg model calibration remains uncertain as the model overestimates concentrations in cleaner areas but fails to capture the high concentrations in the most polluted watersheds.

The regional PCB loads estimated from the Model appear to be consistent with previously reported estimates, while Hg loads appear lower than the previous estimates and need to be interpreted within the context of a less satisfactory model calibration. Although the similarity between the modeled PCB loads and the TMDL estimate should not be used as a success indicator and there is still room for improvement in the calibration, the PCB model that has emerged from efforts during 2014 and 2015 may now be ready to be used for planning purposes.

Contents

1. Introduction

The San Francisco Bay Regional Water Quality Control Board (SFBRWQCB) has identified a number of Pollutants of Concern in San Francisco Bay including mercury (Hg) and polychlorinated biphenyls (PCBs). To help address information needs for these POCs and provide a coordinated approach for stormwater programs and the RMP to monitor and model POC loads to the Bay, a Small Tributaries Loading Strategy (STLS) was developed that outlined four key management questions about loadings and a general plan to address these questions (SFEI, 2009). These questions were developed to be consistent with Provision C.8.e of the first Municipal Regional Stormwater Permit (MRP) for MS4 phase I stormwater agencies ("MRP 1.0", SFBRWQCB, 2009; 2011 (update)).

MQ1. Which Bay tributaries (including stormwater conveyances) contribute most to Bay impairment from POCs?

MQ2. What are the annual loads or concentrations of POCs from tributaries to the Bay?

MQ3. What are the decadal-scale loading or concentration trends of POCs from small tributaries to the Bay?

MQ4. What are the projected impacts of management actions (including control measures) on tributaries and where should these management actions be implemented to have the greatest beneficial impact?

To help determine the magnitude of regional-scale loads for POCs as well as to support the estimation of regional loads in the future for other "yet to be determined pollutants of interest", the Strategy called for the development of a Regional Watershed Spreadsheet Model (RWSM or the "Model"). The RWSM was envisioned as a regional-scale planning tool primarily to estimate average annual loads from the small tributaries, but secondarily to provide supporting information for prioritizing watersheds or sub-watershed areas for management. A spreadsheet model was chosen primarily because it is easy to construct and use at a regional scale. The RWSM is structured with three stand-alone empirical models: hydrology model, sediment model, and pollutant models. The hydrology model uses runoff coefficients based on land use-soil-slope combinations to estimate annual runoff from a watershed and can serve as the basis for any pollutant model (Figure 1). The sediment model uses a function of geology, slope and land-use to simulate suspended sediment transport in landscape while adjusting for watershed storage factors. The pollutant model is essentially a "concentration map" that can be driven by either the hydrology model (for pollutant concentrations in water, Figure 1) or the

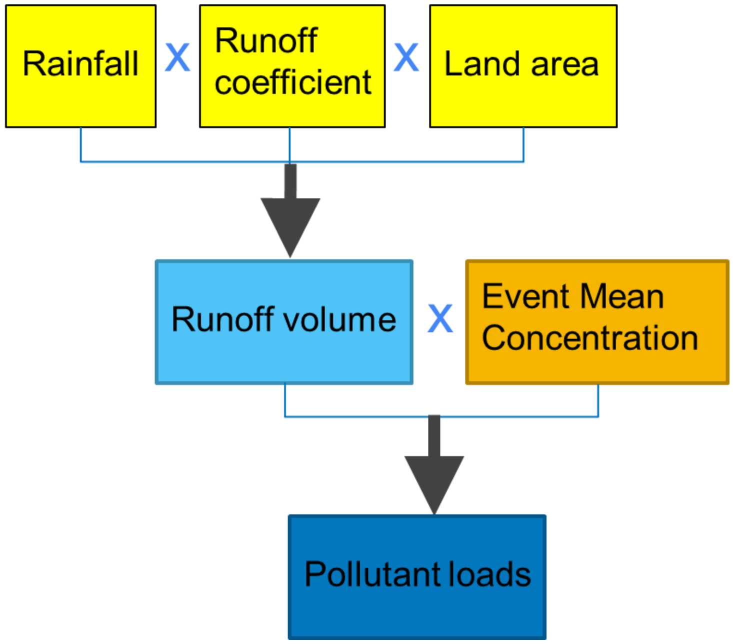

Figure 1. General structure of a hydrology-based pollutant model.

sediment model (for pollutant concentrations on fine sediment particles as particle ratios[1] for specific land use or source areas). The choice of modeling approach is pollutant-specific and depends on whether a pollutant is mainly sediment-associated and the type of concentration data that is available.

Starting in 2010, a multi-year effort was undertaken to systematically develop and calibrate the Model. The development process has been documented through three previous progress reports (Lent and McKee, 2011; Lent et al., 2012, McKee et al., 2014). The model development plan was to structure the pollutant models of the RWSM with the option of being able to use either a hydrology model or suspended sediment (SS) model as the basis. The modeling plan also included linkages to other efforts by Bay Area Stormwater Management Agencies Association (BASMAA) and the RMP. Studies to improve GIS data about the sources of PCBs and Hg in the urban landscape of the Bay Area and to characterize stormwater concentrations during wet weather provided input data for the Model (SFEI, 2009; McKee et al., 2012; McKee et al., 2013; Gilbreath et al., 2014; McKee et al, 2015 in review; McKee et al., 2016 in review). Functionally, the PCB and Hg models relate physical characteristics in each watershed (flow or suspended sediment production, land uses, and source areas) to the average annual PCB or Hg loads at a watershed scale. The outputs of the PCB and Hg models are the Event Mean Concentrations (EMC) or particle ratios for specific land use and source area groups that can be

[1] Particle ratios = pollutant concentration in water (ng/L) / suspended sediment concentration (mg/L) equivalent to mg/kg.

used to estimate the pollutant loads for unmonitored watersheds and for the region as a whole. Issues associated with the accuracy and specificity of the GIS data in relation to PCB and Hg source areas, the simplistic model structure and parameterization (annual time step, land-use based structure), the lack of available EMC and particle ratio data for PCB and Hg for specific land uses and source areas, as well as the high uncertainty with the calibration dataset and representativeness of the calibration watersheds, made it difficult in the past to achieve a reasonable calibration for all three models. The PCB and Hg model calibration was further complicated by the issues associated with the SS model, including an unstable model calibration. This report documents how many of these challenges have been largely resolved.

In light of these challenges, the modeling effort in Years 4 and 5 (2014 and 2015) was focused on improving the calibration of the PCB and Hg models. This Progress Report details the improvements made on GIS data, model structure, and calibration approach, and presents model calibration results and updated regional load estimates. Although there are still some calibration challenges, the models that have emerged from efforts during 2014 and 2015 produce more reasonable results (especially for PCB) and may be ready for planning level applications. The findings from this year's effort and future steps for improvement are summarized at the end of this progress report.

2. Methodological Improvements

2.1. Quality Checking and Improvement of GIS Data

The GIS layers for land use and source areas are the basis of the PCB and Hg models, and the quality of these GIS layers is critical to ensure the models are structured properly and calibrated reasonably. The term land use (LU) was used to describe the set of standard urban land use categories commonly used for urban planning, zoning, and scientific investigations: industrial, commercial, residential, agriculture, open space, and transportation (Park et al., 2009). The term source area (SA) was used to refer to specific locations where PCB or Hg may have been transported, used, or spilled, potentially leaving a legacy of contaminated equipment and/or higher concentrations in soils and in or near surface water. Source areas can and often do cut across land use boundaries (for example railway lines) and one SA type can be embedded in all land use classes (e.g., electrical transmission facilities). However the predictive value of SA mapping is inherently limited by high variability in site-specific mechanisms of actual release and dispersal of pollutants, particularly for PCBs. The present components of the SA category (see Table 1 below) are based on general correlation of known uses or activities with PCB occurrence and do not account for several major source categories including: PCB remediation sites with cleanup targets above those now of interest for TMDL purposes, illegal dumping, and past releases of PCB-containing caulks, sealants or other building materials.

During 2014 and 2015, the improvement of GIS data was focused on determining the degree of inconsistencies in general LU categories across the region. The general LU dataset (ABAG 2005) used in the 2013 version of the RWSM was compared to the recently published National Land Cover Database (NLCD2011) (Homer et al., 2015) to check for significant differences in urban land uses. In general, urban LUs in NLCD2011 are consistent with the ABAG (2005) dataset (<4% absolute difference). Some discrepancies (>10%) were noted but were associated with the land use sub-groups we chose for coalescing the more than 200 categories within the ABAG database into the six basic land use categories (industrial, commercial, residential, agriculture, open space, and transportation). Based on this analysis, the data that were used to support the 2013 RWSM update were deemed appropriate and continued to be used as the basis for 2015 RWSM update.

2.2. Changes in Model Structure

PCB and Hg Models can conceptually be based on either water concentrations (ng/L) or particle based concentrations (mg/kg) (McKee et al., 2014). Initially (in 2011) no a priori decision was made on either basis before running early versions of the model; rather the intent was to use the Model to explore and justify the basis. However, the initial preference was to use sediment as the basis for the PCB and Hg Models in watersheds or landscape components that have higher sediment production rates such as agricultural and open space areas. This was seen as a means for preserving the variability of pollutant supply associated with the erosion of clean sediments in the model structure. Since the Model calibration procedures require a reasonable range of concentrations or particle ratio data for each parameter as a starting point, considerable effort was taken to generate this information, but despite these effort, information on concentrations in flowing stormwater in relation to land uses or source areas remained sparse. Therefore, even though the sediment model calibration was unstable, the decision was made continue to base the PCB and mercury models on sediment and use particle ratios as the calibration parameters.

However, upon a more thorough examination, it became clear that the unstable nature of the sediment Model and the elevated SS loads it produced were the main reason why the PCB and Hg Models failed to calibrate. Therefore, the use of the hydrological Model as the basis for the PCB and Hg Models was explored in 2015. The models were then calibrated to water concentrations instead of particle ratios. Where initially there had been concerns about the lack of enough water based concentration data in land uses or source areas to support the water based Models, lessons learned during the trials with the sediment Model were used to develop the constraints for each parameter in the water based Models.

Table 1. Land use and source area categories used in model development.

2013 data structure		land area in the 22 calibration watersheds
Land uses	Null	0.42%
	Agriculture	2.2%
	Industrial	2.5%
	Commercial	8.4%
	Transportation	13%
	Residential	33%
	Open	40%
		100%
Source areas	Cement	Overlay "double counting" with land use categories to facilitate variable mass production in relation to land use context. For example, a rail line that crosses through differing land uses would have a unique unit area mass production for each land use.
	Crematoria	
	ElectricPower	
	ElectricTransf	
	Highways	
	ManufMetals	
	Military	
	OilRefineries	
	OldUrbanAndIndustrial	
	RecycAuto	
	RecycDrums	
	RecycMetals	
	RecycWaste	
	Streets	
	TranspAir	
	TranspRail	
	TranspShip	

2015 option 1		land area in the 22 calibration watersheds
Land uses	Null	0.42%
	Open	40%
	Agriculture	2.2%
	New Residential	7.1%
	New Commercial	1.8%
	New Industrial	1.7%
	New Transportation	2.4%
	Old Industrial	1.3%
	Old Residential	26%
	Old Commercial	6.3%
	Old Transportation	10%
		100%
Source areas	Cement	Overlay "double counting" with land use categories to facilitate variable mass production in relation to land use context. For example, a rail line that crosses through differing land uses would have a unique unit area mass productions for each land use.
	Crematoria	
	ElectricPower	
	ElectricTransf	
	Highways	
	ManufMetals	
	Military	
	RecycAuto	
	RecycDrums	
	RecycMetals	
	RecycWaste	
	Streets	
	TranspAir	
	TranspRail	

2015 option 2		land area in the 22 calibration watersheds
Land uses	Null	0.42%
	Open	40%
	Agriculture	2.2%
	New Residential	7.1%
	New Commercial	1.7%
	New Industrial	1.6%
	New Transportation	2.3%
	Old Industrial	0.91%
	Old Residential	26%
	Old Commercial	6.2%
	Old Transportation	10%
	All source areas combined	1.7%
		100%
Source areas		No “double counting”.

2.3. Improvements to the Model Calibration Approach

A weakness in the original model structures was the output of single estimates of calibration coefficients and singles estimates for regional loads causing the need to use best professional judgment to estimate potential model uncertainty. The calibration approach for 2015 modeling was improved to provide a model structure that generates an estimate of uncertainty during the process of model calibration. Instead of calibrating the Model to a single average or median concentration, the Model code was rewritten to calibrate to a randomly selected calibration point drawn from the distribution of observed data at each calibration site. The goal of this approach was to provide confidence intervals for the calibrated concentrations and resulting loads estimates to quantify the uncertainty associated with the data and the calibration process. The detailed calibration procedure was as follows:

1) Construct a distribution of the observed data for each calibration watershed. The log-normal distribution was deemed appropriate based on data analysis and the typical pattern of stormwater data.
2) Randomly generate observed data points based on the distribution, one for each calibration watershed.
3) Calibrate the Model for each watershed simultaneously in the same manner previously applied using the Box method (optimization process), and save the Model parameters when the calibration is deemed completed.
4) Repeat the process (steps 2-3) for a number of iterations (currently set as 100) and save the Model parameters for each iteration.
5) Establish the distribution of model parameters from the 100 points that were produced by all iterations.

Based on the recommendation from our advisors (Stenstrom and Mangarella, personal communication, October, 2015), the 25th percentile, median, and 75th percentile of calibrated concentrations were then used to estimate a range of PCB and Hg loads for the whole region.

2.4. Land Use and Source Area Grouping

In addition to the new calibration approach, considerable effort was also spent on exploring how to group LUs and SAs in the models to best describe PCB and Hg observed in Bay Area stormwater runoff. In the 2013 model, the best six category model structure was proposed for PCB and the best five category model for Hg after a series of trials of different grouping options (McKee et al., 2013). However, questions remained as to whether these model structures were reasonable or if the groups were well represented in the calibration watersheds. It was recognized that the selection of one group over another had a large influence on the resulting regional loads as well as the relative load estimates between watersheds. Further, the model estimates based on the previous groups resulted in very high yields in certain watersheds for which stakeholders' anecdotal knowledge did not support the model results.

During the completion of the 2013 model development, one of the emerging concerns was the potential for excessive double counting. During the early phases of model development (Lent and McKee, 2011), it was decided that a limited amount of double counting would be beneficial allowing, for example, a SA such as a railway line to have unique coefficients as it passes through multiple LU types. As such, the 2013 model structure was purposely designed to compile estimates of load per unit area for each LU and SA and add those together (therefore leading to a double-counting of load wherever there was a SA given that the LU coverage was complete for the region). However, as the model development and testing evolved, it was discovered that the overlapping of SAs (for example, older industrial, railway lines, electrical facilities, and waste recyclers) was inadvertently causing an excessive level of "double counting" of pollutant loads in certain watersheds that led to a model complexity potentially beyond the simplistic nature of the model structure and parameterization. In addition, the stakeholders, in parallel, had been working on a multiple linear regression approach (ACCWP, 2014; CCCWP, 2014; SMCWPPP, 2014; SCVURPPP, 2014) that supported a land use grouping that was slightly different from the choices in the 2013 model (McKee et al. 2013). Therefore, SPLWG reviewers recommended that a new set of categories should be explored to be more inclusive of the SAs and more aligned with the LU groupings chosen by the stakeholders.

Two options were explored during 2014 and 2015 to address concerns over the land use and source area groupings and to improve model calibration: 1) treat the SAs the same way as with the previous models and double count their loads except remove Old Industrial and Old Urban as separate SAs and instead improve the LU designations to include "old" and "new" for all urban land use categories; and 2) remove double counting altogether by lumping all source areas as a separate land category and integrating it into the LU coverage.

To support the exploration of these options, the GIS definitions of the main LU categories were revised. The SA categories of Old Urban and Old Industrial as defined in the previous model were eliminated to reduce double counting. All urban LUs (Residential, Commercial, Industrial, and Transportation) were split into New– and Old– (breaking at 1968 for Industrial and 1974 for other three LUs) to make it easier and more flexible for grouping them during the calibration process. To explore Option 2, all source areas (now excluding Old Industrial and Old Urban) were lumped together and burned into the general LU layer such that Source Area became a new general LU and each Source Area was identified only as Source Area and maintained no other general LU category.

The model calibration results using Option 2 were more desirable than Option 1 (Appendix) and therefore the Option 2 grouping was selected to construct the model. This option has the following advantages: all SAs are included within the model calibration thus avoiding the trap of selective inclusion or exclusion of SAs, and the total area of the revised LU groups adds up to the total area for each watershed (Table 1). Adding this Source Area group, a four-category model was constructed for both PCB and Hg, with three land use groups as defined and shown in Table 2.

3. Calibration Results

After making the improvements described in the previous section, the Model was calibrated for PCB and Hg. The calibration results were evaluated using a two-step method: (1) examining the calibrated PCB and Hg concentrations for each group and (2) comparing the modeled and observed concentrations and loads for the calibration watersheds. For the first step, a positive evaluation was made if the values for each LU matched our conceptual understanding of unit load production based on the PCB and Hg contaminant profiles (McKee and Lent, 2011). If this first criterion was met, the modeled concentrations were compared across all the calibration watersheds to determine how closely the modeled concentrations matched the observed data. Also, the measured loads from a fewer number of watersheds (11 watersheds for PCB and 9 watersheds for Hg) (McKee et al., 2015) were compared to both modeled loads and estimated loads to further verify model performance. The modeled (or simulated) loads were calculated by multiplying the calibrated concentrations with modeled runoff, while the estimated loads were calculated by multiplying the calibrated concentrations with measured runoff. Comparing these two loads helped identify whether the mismatch for each watershed was due to the modeled concentrations or the modeled flow.

The calibration for the PCB model appears to be reasonable based on the two-step evaluation process. The PCB model calibration results (Table 3) were highest for Old Industrial, followed by the SA group, Old Urban, and lowest in the clean (ag-open-new urban) LU group (Figure 2), generally matching our conceptual understanding of unit load production from these land uses. The modeled mean concentrations show an overall pattern of over-simulation of the cleaner watersheds and under-simulation of the dirtiest watersheds (Figure 3). This pattern is consistent with all previous calibration attempts (McKee et al., 2013) and reflects the inherent limitations of a regional, one-size-fit-all model. With the calibration dataset exhibiting variation as high as two orders of magnitude, the calibration that was aimed to minimize the sum of errors between simulated and observed data at all stations was bound to over-simulate some and under-simulate others (Figure 3).

The measured, simulated, and estimated PCB loads for the subset of calibration watersheds where load measurements are available are shown in Figure 4. The simulated loads match the measured data well at four of the watersheds, including the highly polluted Sunnyvale East

Table 2. Land use category groupings for the PCB and Hg models.

Land use or source area	PCB Model LU Group		Land use or source area	Hg Model LU Group
LU Old Industrial	Old Industrial		LU Old Industrial	Industrial
LU New Industrial	Old urban / new industrial		LU New Industrial	
LU Old Residential			LU Old Residential	Old urban
LU Old Commercial			LU Old Commercial	
LU Old Transportation			LU Old Transportation	
LU Agriculture	Clean		LU Agriculture	Clean
LU Open			LU Open	
LU New Residential			LU New Residential	
LU New Commercial			LU New Commercial	
LU New Transportation			LU New Transportation	
SA manufMetals	All Source Areas		SA manufMetals	All Source Areas
SA recycAuto			SA recycAuto	
SA recycMetals			SA recycMetals	
SA recycWaste			SA recycWaste	
SA recycDrums			SA recycDrums	
Marine repair scrap yards			Marine repair scrap yards	
SA electricPower			SA electricPower	
SA electricTransf			SA electricTransf	
SA transpRail			SA transpRail	
SA transpAir			SA transpAir	
SA military			SA military	
			SA crematoria	
			SA cement	
			SA Refinery and petrochem	

Table 3. Low, median and high concentrations estimated by the calibrated PCB model. Note, the model calibration process tended towards the lower boundary for the clean (0.5 ng/L) and allSA group (50 ng/L).

Land use or source area	Grouping	25th percentile (ng/L)	Median (ng/L)	75th percentile (ng/L)
Old Industrial	Old Industrial	50.0	98.0	312.4
Old Commercial	Old urban	14.7	24.5	41.4
Old Transportation				
Old Residential				
New Industrial				
New Commercial	Clean	0.5	0.5	0.5
Agriculture				
New Residential				
New Transportation				
Open				
All Source Areas combined	allSA	50.0	50.0	65.8

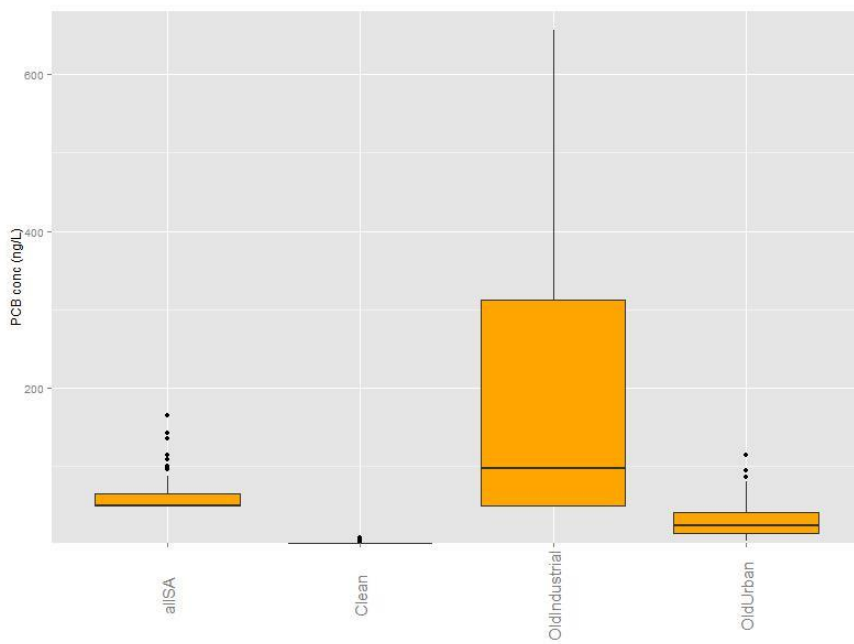

Figure 2. Boxplot showing the range of simulated PCB concentrations for each land use group in the calibrated Model. The box is bounded by 25th and 75th quartiles of the calibrated concentrations. The upper and lower whisker extend from the hinge to 1.5 * IQR of the hinge, where IQR is the distance between 25th and 75th quartiles. Data beyond the end of the whiskers are outliers and plotted as points.

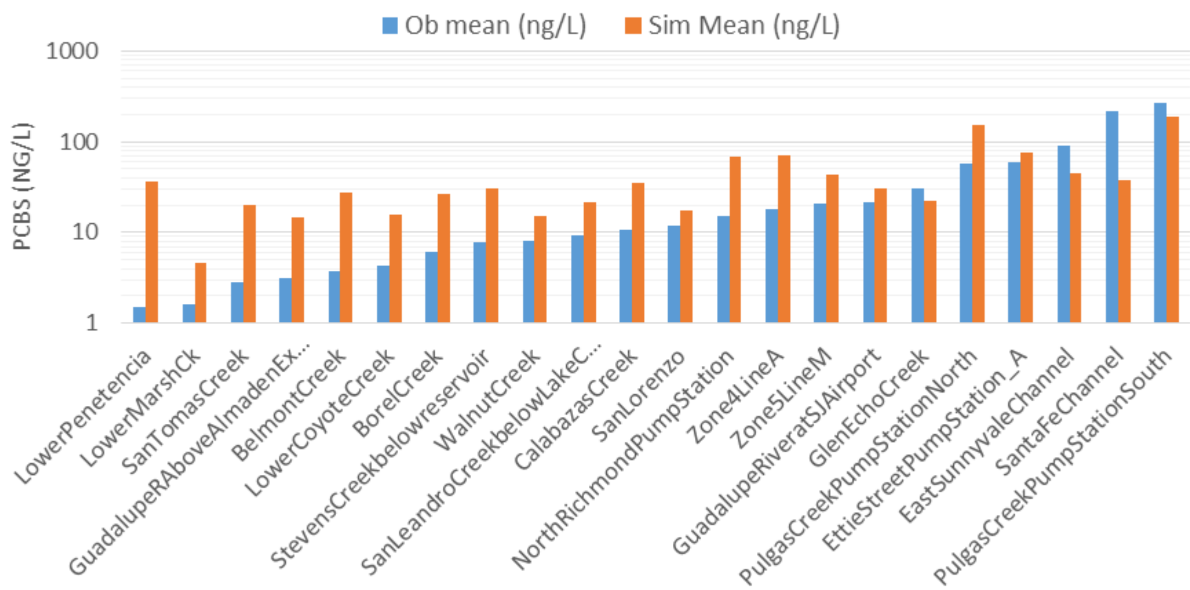

Figure 3. Simulated and observed mean PCB concentrations in water at calibration watersheds. “Ob” = observed and “Sim” = simulated (note the log scale on the y-axis). Note, there are less watersheds included in Figure 4 below because measured loads are available in just a small subset of watersheds.

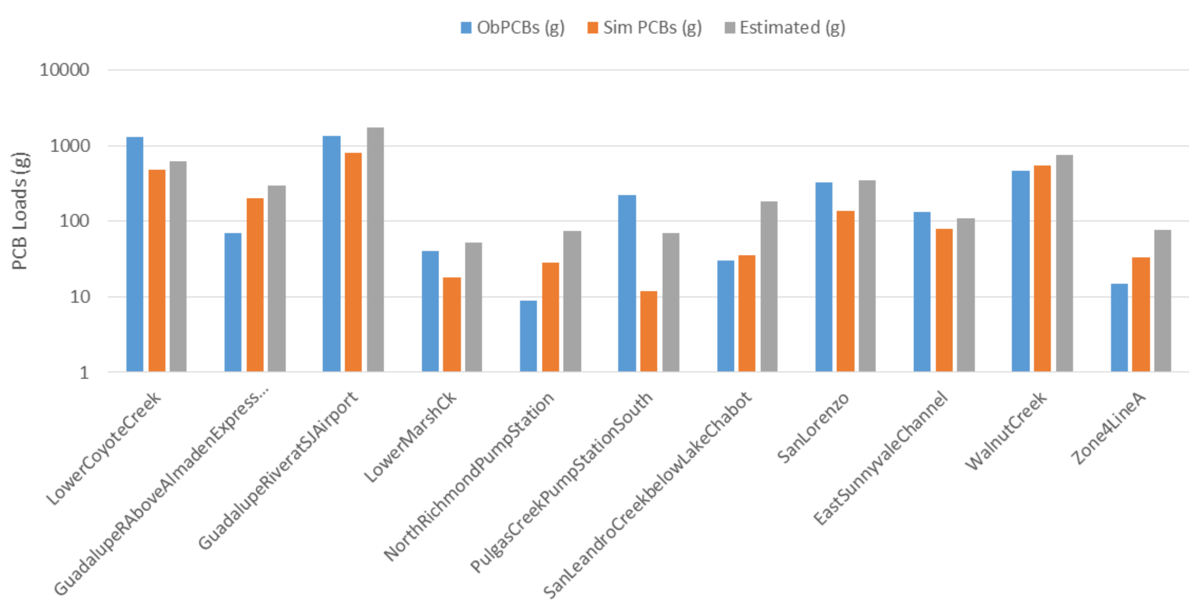

Figure 4. Measured, simulated, and estimated PCB loads in water at calibration watersheds (note the log scale on the y-axis). Note, there are less watersheds included in this figure compared to Figure 3 above because measured loads are available in just a small subset of watersheds.

Channel and the cleaner watersheds of Walnut Creek, in addition to San Leandro creek, and Guadalupe River at SJ Airport. For Lower Coyote Creek and Guadalupe River above Almaden, the simulated loads vary more greatly from the measured loads, yet the simulated loads are very similar to the estimated loads, suggesting that the runoff simulation is reasonable for these watersheds and the discrepancy between measured and simulated loads is largely due to the calibrated PCB concentrations[2]. The largest discrepancy between measured and simulated loads was for North Richmond Pump Station where the model over-predicted PCB load by ~200%. Further investigation is needed on both model results and measured loads to better understand the load results.

In comparison, the calibration of the Hg model appears less satisfactory. The calibration results show a relatively even distribution of Hg concentrations among LU and SA groups, with the highest from the clean LU group (ag-open-new urban) and the lowest from the combined SA group (Figure 5, Table 4). This relatively even distribution of concentrations generally matches our conceptual understanding of the diffuse nature of Hg sources in the landscape and the influence of atmospheric deposition. The simulated Hg mean concentrations tended toward the middle range of the observed data (Figure 6), resulting in over-prediction of 15 watersheds and under-prediction of the three dirtiest watersheds (significantly so for the most polluted watershed, Zone5LineM). While the simulated mean concentrations for all the watersheds were relatively equal, around 100 ng/L, the observed mean concentrations showed a small slope in the first 18 watersheds from 14 ng/L up to approximately 100 ng/L, and then a sharp increase in the final three watersheds. This again highlights the limited ability of a regional model to explain the large variations in the monitoring data, especially at extremities, but it also suggests that the calibration for the Hg model is not yet complete or satisfactory. There is clearly something unique about the Zone5LineM dataset that neither the model nor our anecdotal knowledge of the watershed can explain; further investigation into the potential Hg sources within the watershed should be made and given the impact on the model calibration, resampling the watershed could be considered to verify these high concentrations.

The Hg load comparison for individual watersheds shows an overall reasonable match between the simulated loads and measured loads at the majority of the watersheds but with clear discrepancies between measured and simulated loads at three watersheds (East Sunnyvale Channel, San Leandro Ck and Zone 4 Line A; Figure 7). The largest discrepancy was for East Sunnyvale Channel where the model over-predicted Hg loads by ~ 130%. But since the estimated loads matches the simulated loads well (the only difference between these estimates

[2] Other outcomes are possible too. For example, for some watersheds the difference between concentration and load results can be explained by flow. If a watershed is calibrated poorly for concentration but shows a better estimates load, this tends to indicate a likely over-simulation of runoff.

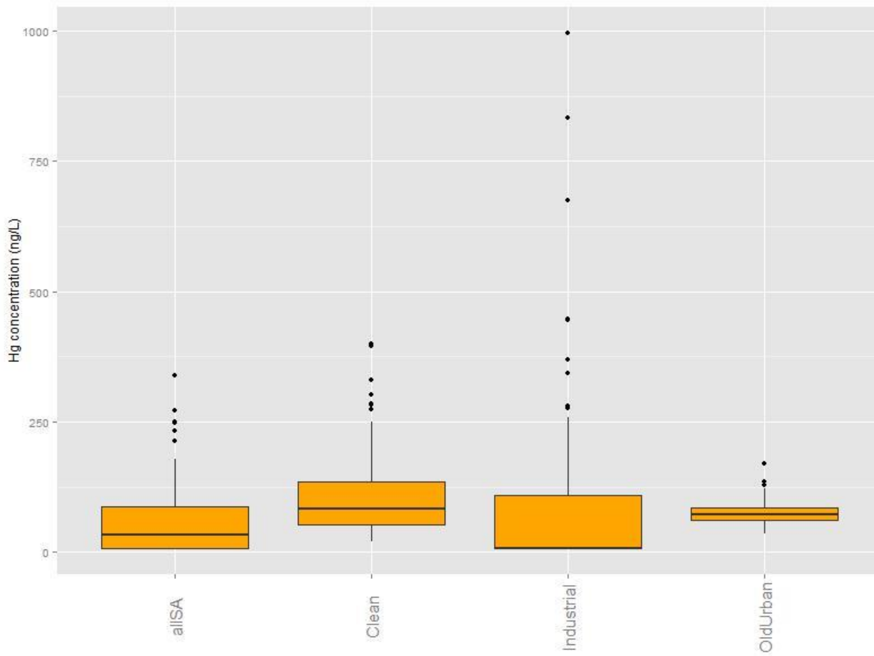

Figure 5. Boxplot showing the range of Hg concentrations for each land use group in the calibrated model. The box is bounded by 25th and 75th quartiles of the calibrated concentrations. The upper and lower whisker extend from the hinge to 1.5 * IQR of the hinge, where IQR is the distance between 25th and 75th quartiles. Data beyond the end of the whiskers are outliers and plotted as points.

Table 4. Low, median and high concentrations estimated by the calibrated Hg model. Note, the model calibration process tended towards the lower boundary (7 ng/L) for the industrial parameter.

Land use or source area	Grouping	25th percentile (ng/L)	Median (ng/L)	75th percentile (ng/L)
New Industrial	Industrial	7.0	7.0	110.2
Old Industrial				
Old Residential	Old urban	62.2	73.3	86.2
Old Commercial				
Old Transportation				
Open	Clean	52.1	83.8	134.6
Agriculture				
New Residential				
New Commercial				
New Transportation				
All Source Areas combined	allSA	7.0	32.6	88.1

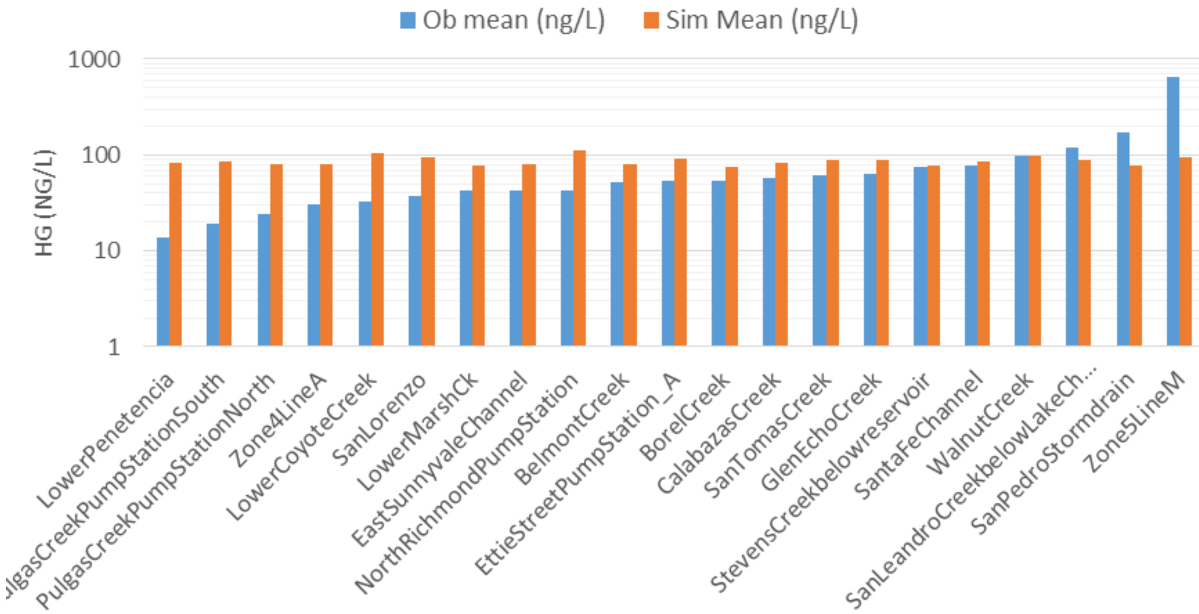

Figure 6. Simulated and observed mean Hg concentrations in water at calibration watersheds "Ob" = observed and "Sim" = simulated (note the log scale on the y-axis).

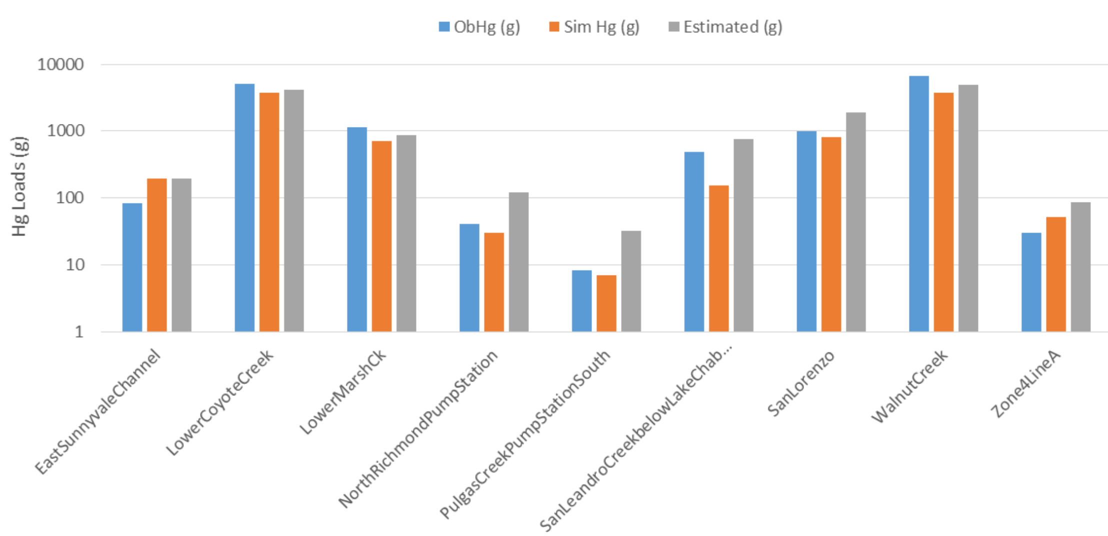

Figure 7. Measured, Simulated and estimated Hg loads in water at calibration watersheds.

is the flow volume), suggesting that the calibrated Hg concentrations may be biased high for the watershed. For North Richmond Pump Station and Pulgas Pump Station South, the estimated loads are significantly higher than both simulated and measured loads, indicating the flow volume is significantly underestimated and Hg concentrations are too high[3]. Revisiting the Hydrology Model and recalibrating using the monitoring data that has been collected since the last Hydrology Model calibration may be warranted based on these findings. Additionally, these loads results need further investigation in the context of uncertainty in both the calibrated model and measured data.

In summary, the PCB model calibration appears reasonable enough to move onto the next step of estimating loads, but the Hg model calibration remains uncertain. Therefore the regional Hg loads described in next section should be interpreted with this caveat.

4. Pollutant Load Estimates and Discussion

Total Maximum Daily Load reports (TMDLs) (SFBRWQCB, 2006; 2007) call for the development of improved information about PCB and Hg sources and loads and a reduction of stormwater PCB loads from 20 kg to 2 kg by 2030 and Hg loads from 160 kg to 80 kg by 2028 with an interim milestone of 120 kg by 2018. These needs were reflected in the first Municipal Regional Stormwater Permit (MRP) for MS4 phase I stormwater agencies (SFBRWQCB, 2009; 2011 (update)). MRP 1.0, as it came to be known, contained provisions aimed at improving information on storm water loads (Provision C.8.) and piloting a number of management techniques to reduce PCB and Hg loading entering the Bay from smaller urbanized tributaries (Provisions C.11. and C.12.).

Therefore, in relation to provision C.8, after the calibration step, the Model was used to estimate regional loads of PCB and Hg. The loads were calculated four different ways. First, the load of pollutants from the different land use types was calculated. Second, all the results of the Model were summed to calculate the total load for the whole region. Third, the load from individual watersheds was estimated. And, finally, the yield or the load normalized by watershed area was calculated for each watershed. Presenting the model results in these four ways provides useful demonstration of information for management decisions about prioritizing watersheds and land use types for load reduction actions and demonstrating progress toward achieving the goals of the TMDLs.

[3] It is sometimes very difficult to determine the relative sensitivity of the model to estimated flow or land use. Although there are considerable challenges with the quality of the land use and source area data that are used as inputs to the model, in this case, it appears that flow volume affected both the PCB and Hg model calibrations. The seemingly different impacts on each model come largely from the calibrated concentration coefficients. The Hg model remains more challenged at this time by the choice of parameters and the resulting calibrated coefficients.

4.1. Loads by Land Use Type

The simulated regional loads were summed by land use type to help understand the relative contributions of PCB and Hg loads from different land use or source area categories. At the regional level, 57% of PCB loads were contributed by Old urban, 24% by the SA group, 16% by Old Industrial, and only 2.7% by Clean LUs (Table 2, Figure 8). This pattern generally fits our conceptual model - PCBs were primarily used in urban and industrial applications and atmospheric redistribution is assumed to be a small component of the PCB environmental cycle[4]. Therefore, the majority of loads come from urban areas, and the Clean LU contributes the least. However, it is also possible that PCB load estimates from the "Clean" LU are biased low. This may require further investigation, or alternatively, the nonurban land-use could be dropped altogether with little impact on the regional load estimation.

In contrast, for mercury, an estimated 82% of the regional load is coming from Clean LUs, and the other three LU groups together contributed 18%, with the SA group contributing only 0.6% (Figure 9). The large load from Clean LUs is the result of both a high concentration coefficient for the Clean LU category, a large total area for this land use group, and finally disproportionately greater rainfall in the northern third of the region which is also disproportionately low in urban land uses. As discussed in Section 3, the current Hg model calibration is not robust and needs further verification, and the interpretation of the regional loads needs to take this into account. Unfortunately, the model calibration experience to date suggests only modest improvement could be expected, unless the proposed improvements in parameterization[5] and the increased size of the calibration data set end up having a larger than expected influence on model performance. With this caveat noted, the relative contribution from each land use or source area contrasts considerably between PCB and Hg, consistent with the expectation of a more ubiquitous dispersion of Hg relative to PCB.

4.2. Regional PCB and Hg Loads

One of the primary objectives of the spreadsheet model was to generate improved estimates of regional scale loads. Functionally this is done by summing up the simulated PCB and Hg loads for each of the individual watersheds in the region or for a sub-region. As discussed above in Section 2, in contrast to the previous modeling phases (McKee et al., 2013), the regional and sub-regional loads of PCB and Hg were computed by the model for the median (best estimate),

[4] Concrete and other construction materials recycling is presently being explored by BASMAA agencies as a possible local atmospheric source in the areas where it is occurring.

[5] Presently, the parameterization of the Hg model is a little weaker than for the PCB model due to a greater focus on completing a better PCB calibration. There are indications that the "clean parameter" needs breaking up and overall the lack of a large variation between the parameter coefficients suggests overall parameterization of the Hg model needs considerable improvement.

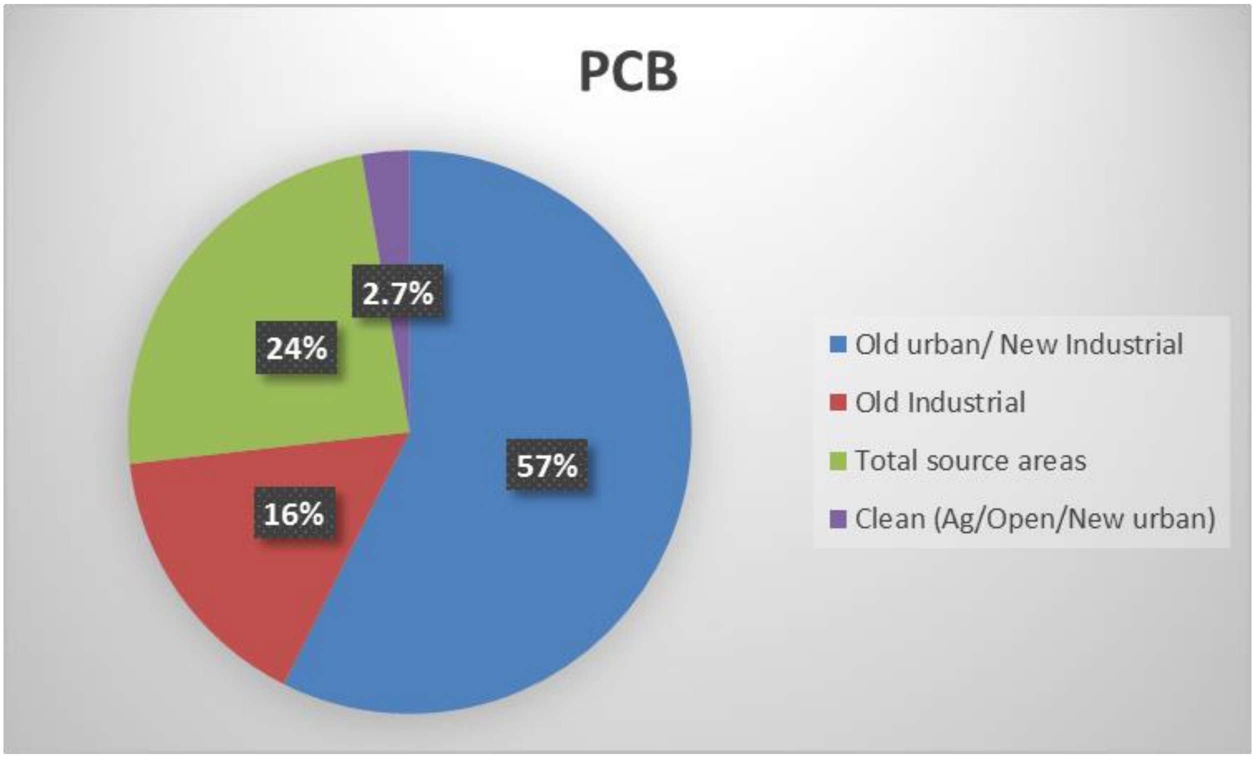

Figure 8. Relative contribution of mass from each LU group in the PCB model.

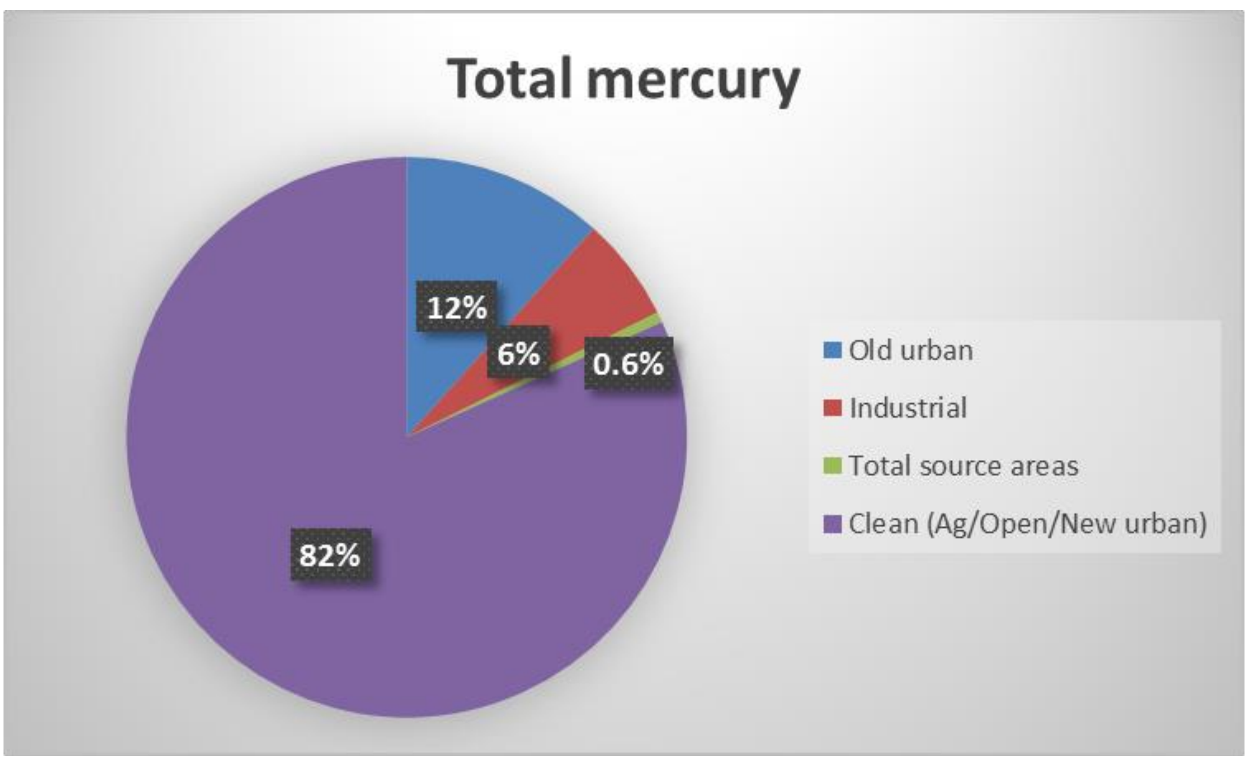

Figure 9. Relative contribution of mass from each LU group in the Hg model.

25th percentile (low estimate), and 75th percentile (high estimate) of the calibrated concentrations (Table 3 and 4).

The new estimate of regional PCB loads ranged from 11.6 kg to 30.1 kg, with a best estimate (median value) of 16.8 kg (Table 5)[6]. This is very similar to the estimated regional load in the San Francisco Bay PCB TMDL (20 kg: Water Board, 2008) that was based on extrapolation of loading data from Guadalupe River and Coyote Creek to the area of urban land use in the Bay Area. It is also similar to the 18 or 19 kg load (using two slightly differing scaling methods) recently presented by McKee et al., 2015 (in review).

In the case of mercury, the best estimate (median load) from the model simulation was 95 kg, about half as big as the load estimated in the San Francisco Bay mercury TMDL (160 kg urban; 25 kg non-urban: Water Board, 2006), but the high estimate of 170 kg is more consistent with TDML load. Given the interim milestone of 120 kg written into the Hg TMDL for 2018, it is somewhat important to generate a reasonably accurate regional load. Recently McKee et al., 2015 (in review) presented a new regional estimate of 113 kg for urban loads. Thus, although the new modeled loads are very similar to the other recent regional estimates, the main contrast is in the relative proportions of urban to non-urban; the RWSM predicts a high percentage of total loads from non-urban land uses, whereas previous estimates were based on the assumption that urban areas contributed the majority of loads[7]. This is yet another indication that Hg model needs further examination or that the assumptions underlying our conceptual model of Hg loads need to be reassessed; again important issues given the need for clarity around the regional loads estimates for Hg that would be timely in 2018 in relation to the TMDL interim milestone.

Although overall the similarities in loadings estimates for PCB, and to a lesser degree for Hg, are generally encouraging, that loads are similar to previous estimates does not provide evidence that the estimates are correct. In fact, both the current estimates and the previous estimates tend to bias towards the central tendency of the data, failing to properly address the weaknesses of the data and our general lack of knowledge about highly polluted areas: how many there are, where they are in the Bay Area, and their pollution characteristics. Work is ongoing to identify more of these areas and further exploration is needed to improve the way the calibration of the model accounts for the extremes of the calibration data set. These

[6] Although it is encouraging that the model is getting some convergence in a narrow range, that there is agreement with loads estimates generated using other methods is not an indication of performance. Improvements are still possible that will potentially lead to regional loads estimates that are greater than those suggested by the present model calibration.
[7] This needed to be resolved since it has large implications for the benefits of treatment.

cautions in mind, the model results are closer to other regional load estimates than any previous calibration attempts. It appears that the substantial changes that were made in 2014 and 2015 to the GIS layers, model structure, and the calibration procedures have greatly improved model performance.

Table 5. Estimated sub-regional and regional loads (kg) of PCB and Hg in RMP Bay segments.

Bay segment	Watershed area (km2)	Total PCB load (kg)			Total mercury load (kg)		
		Low estimate	High estimate	Best estimate	Low estimate	High estimate	Best estimate
Central (East)	87	0.62	1.83	0.92	0.51	2.67	1.08
Central (West)	155	0.53	1.26	0.80	1.21	5.10	2.48
Lower South Bay	1,313	2.50	7.01	3.87	7.17	24.81	13.09
San Pablo Bay (Southeast)	174	0.56	2.11	0.90	1.61	4.60	2.61
San Pablo Bay (Northwest)	1,851	1.81	3.56	2.35	24.22	65.34	39.30
South Bay (East)	1,360	2.25	5.71	3.20	8.40	26.73	14.47
South Bay (West)	258	1.07	3.13	1.70	1.14	6.41	2.77
Suisun Bay (East)	621	1.52	3.94	2.15	4.19	14.05	7.30
Suisun Bay (West)	908	0.78	1.53	0.93	7.39	20.11	11.87
Total Regional load	**6,726**	**11.6**	**30.1**	**16.8**	**55.8**	**169.8**	**95.0**

4.3. Loads by Watershed

The regional loads can be viewed on a map to illustrate their spatial distribution. For the sake of simplicity, the best estimates of loads (based on the median concentration coefficients) were used in subsequent figures and tables. Although uncertainty remains for these load estimates due to uncertainty with model input data, structure, and calibration, individual watershed loads are presented to provide an overall picture of load production across the region. This overall pattern can be further examined based on our conceptual understanding of the regional load distribution and local knowledge on individual watersheds to verify model performance.

As expected, the model simulations predicted that the larger watersheds in the Bay Area contribute generally larger PCB loads (Figure 10). This was generally true for mercury as well (Figure 11) although simulated Hg loads were a little more influenced by runoff volume and PCB loads were more influenced by urban and industrial land uses. As such, there is another group of smaller watersheds in the size range between a few square kilometers up to a few tens of square kilometers where the presence of more polluting land uses also resulted in high simulated PCB loads from individual watersheds. Overall, the 25 watersheds with the largest loads account for 45% of the total regional PCB load, while for mercury, the 25 watersheds that are estimated to produce the largest loads account for 42% of the total regional load (Table 6).

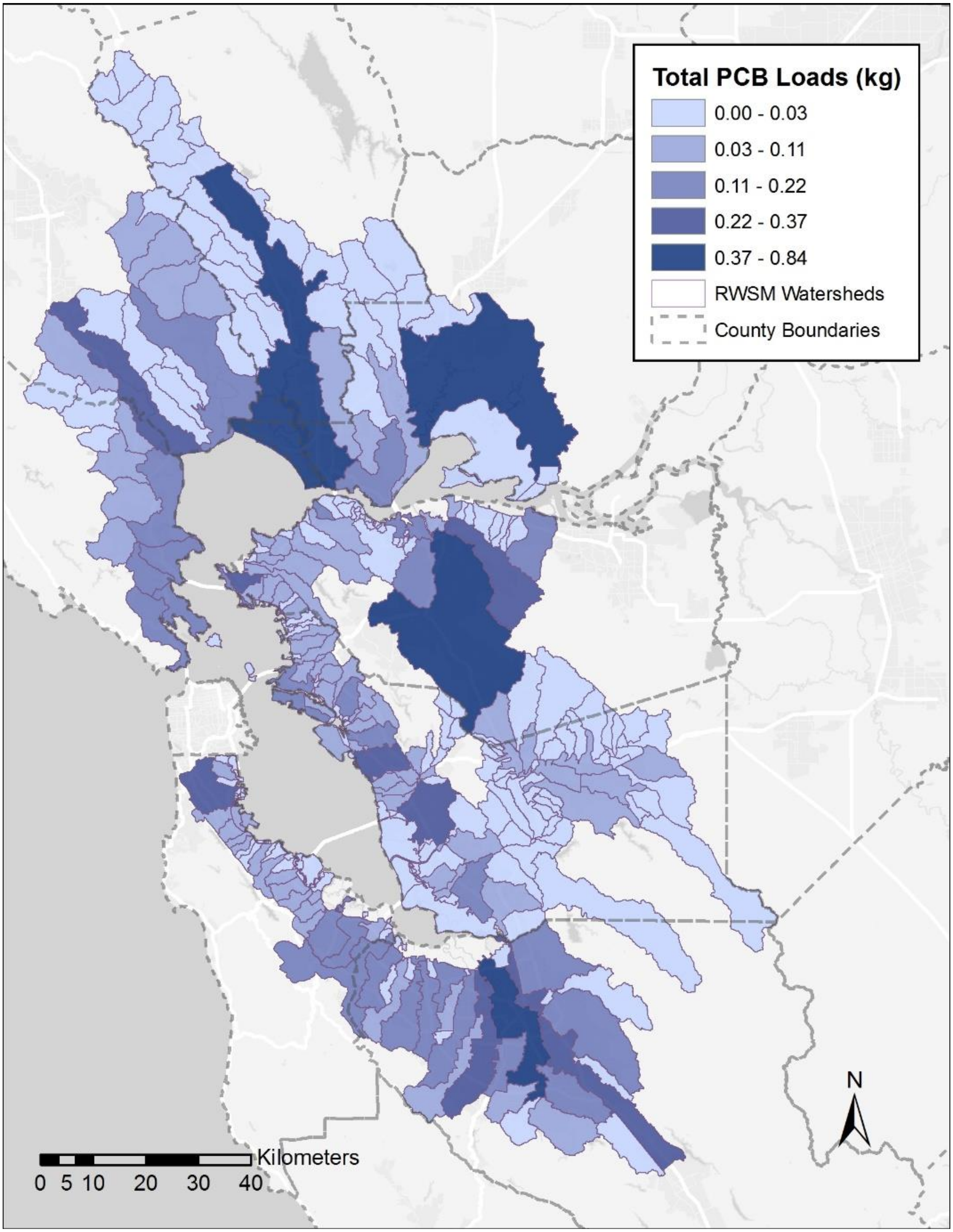

Figure 10. The distribution of estimated PCB loads in watersheds of the Bay Area.

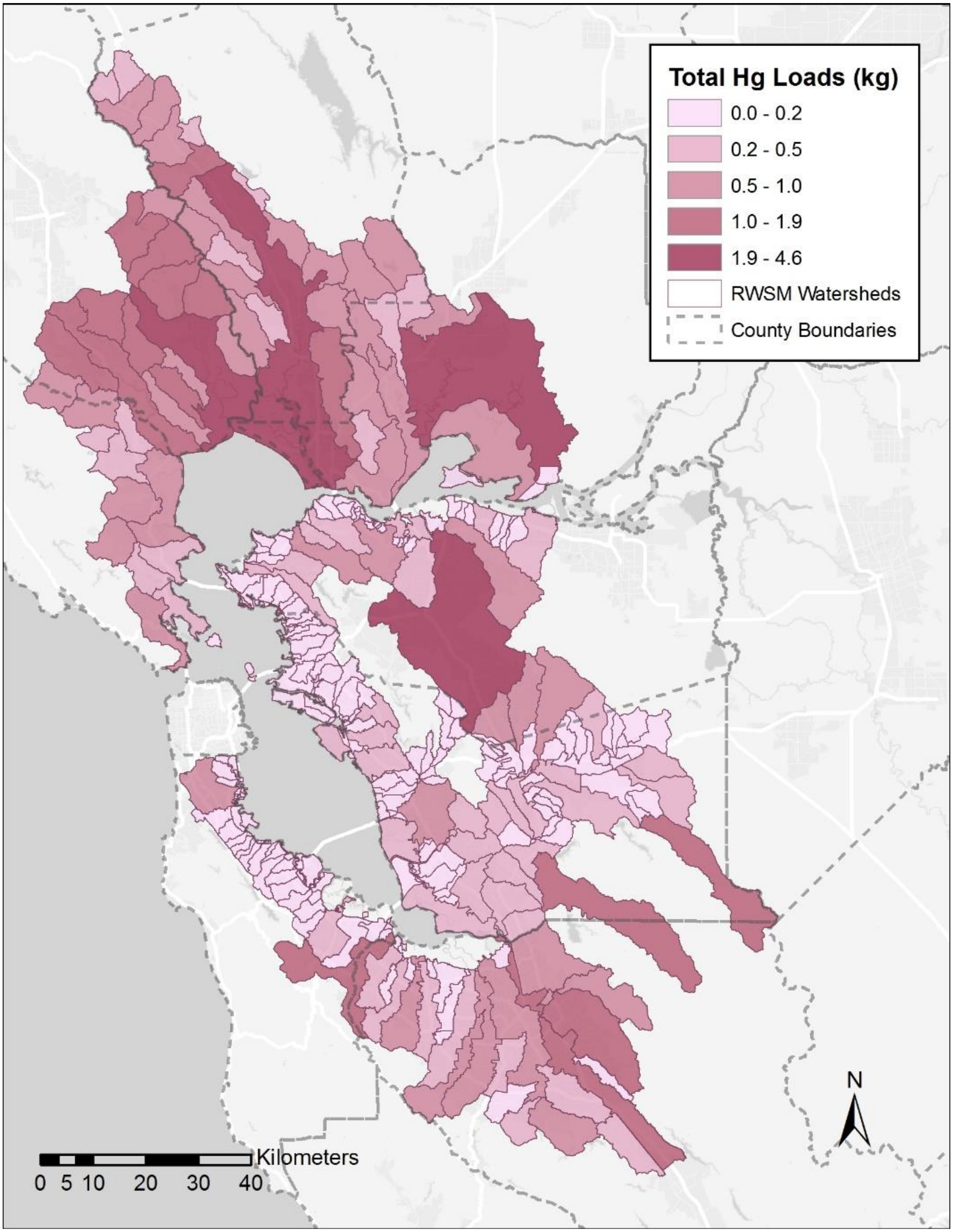

Figure 11. The distribution of estimated Hg loads in watersheds of the Bay Area.

4.4. Unit Loads (Yields) by Watershed

Watersheds with high area-normalized loads (yields) indicate concentrated sources of pollutants. Discharges from watersheds with high pollutant yields may contribute disproportionate loads to smaller or semi-enclosed areas on the Bay margin. In areas where water circulation and mixing and water and sediment dispersion may be reduced, localized impacts may be more prevalent. Areas with this type of disproportional impact and their watersheds have been referred to as “high leverage”; management actions focused in these areas may be relatively cost-effective and have a greater chance of improving water quality.

The simulated results from the RWSM indicate that area-normalized loads for PCB (yields: µg/m2 per year) range from 14-32 µg/m^2 in the top 25 simulated watersheds (Figure 12), while Hg yields range between 24-36 µg/m^2 (Figure 13; Table 7), which are within the range of reported values observed in the Bay Area. For example, PCB yields from multiple years of monitoring at Zone 4 Line A and North Richmond Pump Station are 3.5 µg/m^2 and 4.7 µg/m^2, respectively, and yield of 85 µg/m^2 have been observed at Pulgas Creek South (McKee et al., 2015 in review). The model appeared unable to predict the high yield at Pulgas Creek Pump Station South when using the best estimates of calibrated concentrations, in large part due to the regional model calibration toward the central tendency of the data rather than the extremes, but also because the hydrology model under-simulates flow in the watershed and because of limitations in the GIS data set (the lack of source areas identified in Pulgas within the current GIS data sets meant that the model did not assign mass there). However, when using the high estimate of calibrated concentrations, the model could produce the PCB yield of 57 µg/m^2 for this watershed, which compared more closely with the measured loads. This example highlights the possible need for recalibration of the hydrology model now that more calibration datasets are available, the importance of improving the GIS datasets to identify source areas within high leverage watersheds, and the need for monitoring watersheds with those source areas such that they are included in the calibration of the concentration coefficients.

From a management perspective, it is easier to manage and reduce loads when they emanate from concentrated sources and smaller watershed areas. The Model results indicate that PCBs tend to have concentrated sources while mercury sources are widespread. For PCBs, the top 25 highest yielding watersheds generate 1.0 kg of PCB from just 49.6 km^2 of area, or 6% of the simulated annual average PCB load from just 0.7% of the regional area. In contrast, the Model predicts that the top 25 highest yielding Hg watersheds generate 10% of the simulated annual average Hg load (or 18.7 kg) from an equal proportion (10%) of the regional area (or 662 km^2) (Table 7). Comparisons like these are useful for comparing and contrasting loads between differing land use types within a single Model and for exploring the potential for multiple management benefits across multiple pollutants.

Table 6. Watersheds in the Bay Area with the 25 largest estimated PCB and Hg loads[8].

Watershed	Total Area (km2)	PCBs load (kg)	Watershed	Total Area (km2)	Hg loads (kg)
MouthofNapaRiver	339	0.84	MouthofNapaRiver	339	4.6
MallardReservoir	317	0.81	SuisunSlough	344	4.2
SuisunSlough	344	0.59	MallardReservoir	317	3.7
GuadalupeRiver	93	0.53	MouthofSonomaCreek	149	3.1
DonnerCreek	80	0.37	PetalumaRiver	108	1.9
ColmaCreek	41	0.36	AlamedaCreek5	141	1.5
WardandZeileCreeks	55	0.31	UpperSonomaCreek	49	1.5
LowerCoyoteCreekbelowAndersonDam	106	0.28	LynchCreek	42	1.41
SanTomas	70	0.26	AdobeCreekLakeville	36	1.31
EstudilloCanal	29	0.25	UpperCalabazas	47	1.27
PetalumaRiver	108	0.25	FaganCreek	74	1.24
HermanSloughandCastroCreek	10	0.25	LowerCalabazas	49	1.22
CalabazasCreek	53	0.22	LowerCoyoteCreekbelowAndersonDam	106	1.20
AC_unk09	14	0.22	Petaluma	60	1.10
SanFrancisquitoCreek	82	0.21	TolayCreek	30	1.08
SanRafaelCreek	31	0.21	NathansonCreek	36	1.07
OldMillCreek	39	0.20	LowerSilverThompsonCreek	112	1.05
MartinezCreek4	16	0.19	SanFrancisquitoCreek	82	1.05
PineLake	62	0.19	ArroyoMocho6	97	1.04
KirkerCreek	45	0.18	HeathCanyon	41	1.04
GraysonCreek	45	0.18	GalinasCreek	68	0.98
LowerPenitenciaCreek	76	0.18	PineLake	62	0.96
GalinasCreek	68	0.18	YorkCreek	34	0.95
CorteMaderaCreek	24	0.17	BearCanyon	38	0.94
SunnyvaleWestChannel	19	0.17	CamerosCreek	38	0.90

[8] Note, there are 324 watersheds in the model. It is not practical to create an appendix that includes all watersheds in the model sorted by county that would include watershed area, annual load, annual yield for every watershed for flow, copper, PCB and Hg. Such an appendix would be very long and span many pages given the numbers of watersheds in the model. Such data can be made available upon request but is probably best provided as a GIS data file given that the watershed names and locations are often not self-explanatory.

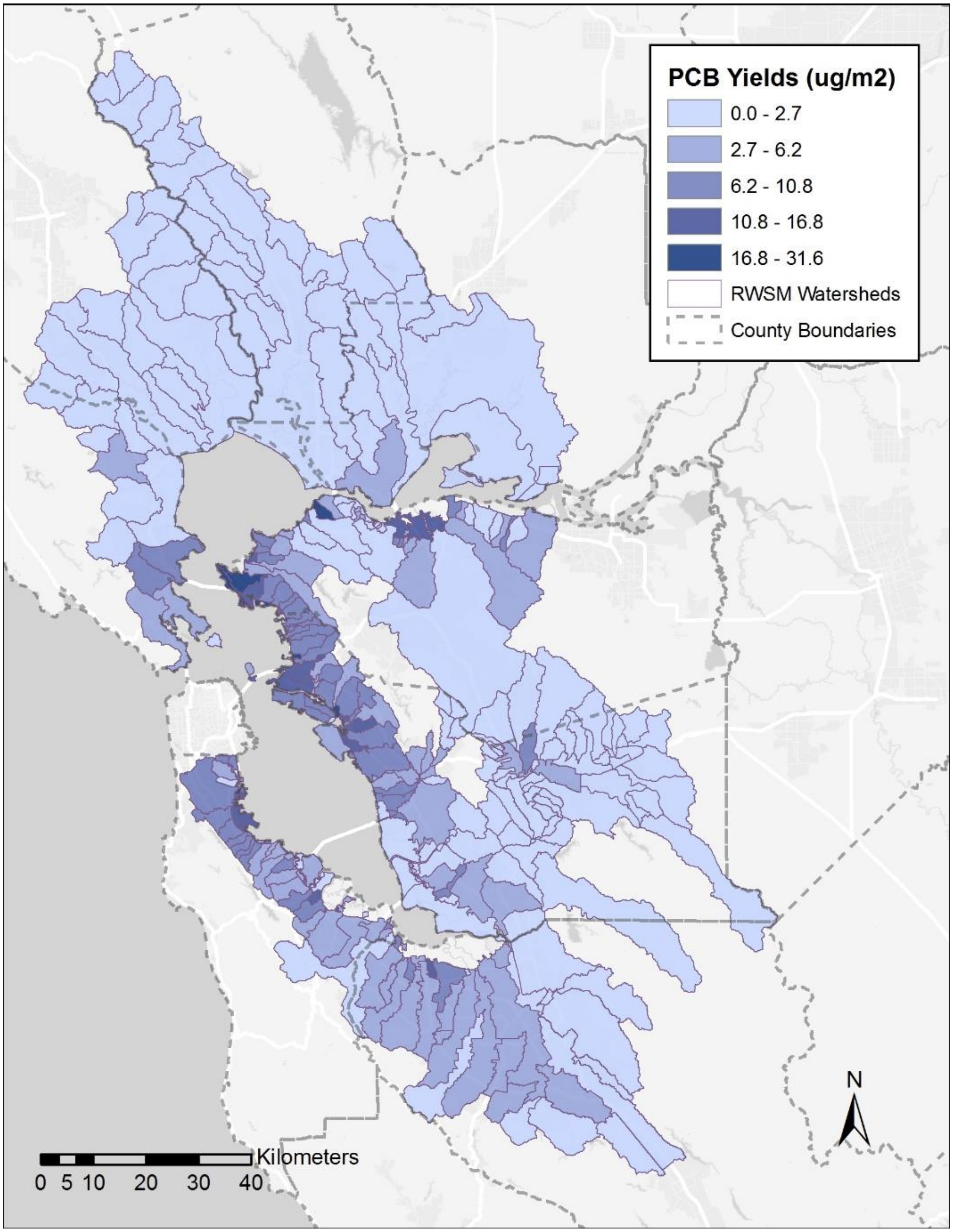

Figure 12. The distribution of PCB yield (area normalized PCB loads) in watersheds of the Bay Area estimated from the RWSM. Although there are still improvements to be made, the distribution shown by this map seems generally reasonable.

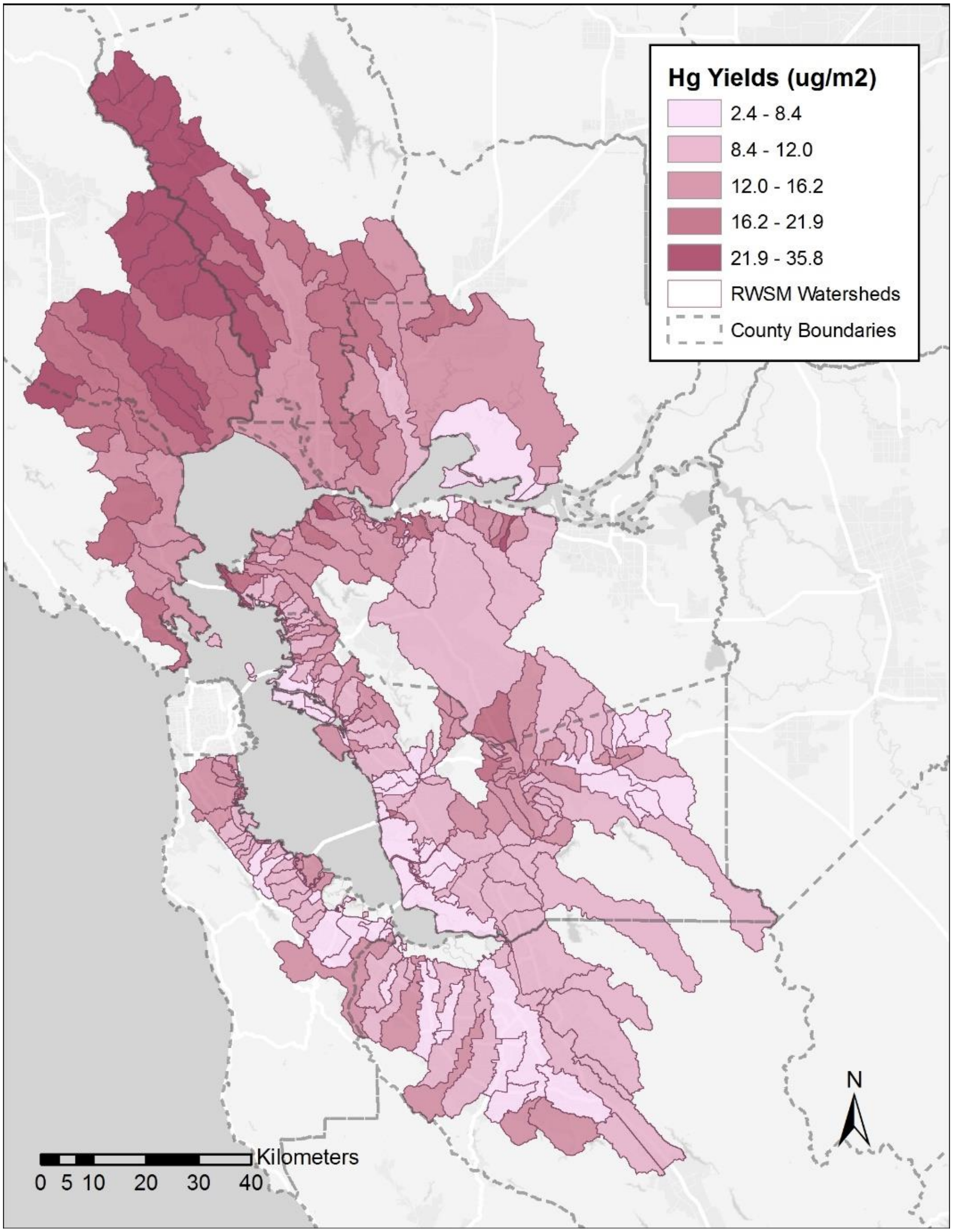

Figure 13. The distribution of Hg yields (area normalized Hg loads) in watersheds of the Bay Area estimated from the RWSM. The distribution shown by this map does not follow the conceptual model that urban areas should produce higher yields of Hg; improvements could include improving parametrization with a focus on subdividing the land uses and source areas more carefully in relation to sediment production (appears this may matter more for Hg than for PCBs) and recalibration with some unexplainable outliers in the Hg data set removed.

Table 7. Watersheds in the Bay Area with the 25 greatest estimated PCB and Hg yields.

Watershed	Area (km2)	PCB loads (kg)	PCB yields (ug/m2)	Watershed	Total Area (km2)	Hg loads (kg)	Hg yields (ug/m2)
PointSanPabloPeninsulaWest	4.0	0.127	32	AdobeCreekLakeville	36.5	1.31	36
DavisPoint	4.9	0.152	31	BearCreek	21.2	0.76	36
SMC_unk03	0.6	0.015	25	TolayCreek	30.4	1.08	35
HermanSloughandCastroCreek	9.7	0.247	25	ChamplinCreek	19.0	0.65	34
HerculesCreekandRefugioCreek	0.2	0.005	24	LynchCreek	42.4	1.41	33
PointRichmond	0.7	0.017	24	UpperSonomaCreek	49.1	1.49	30
SMC_unk09	0.1	0.003	23	NathansonCreek	36.5	1.07	29
OysterPoint	0.2	0.004	21	StageGulch	30.3	0.85	28
AC_unk04	0.9	0.018	21	YorkCreek	34.0	0.95	28
SMC_unk07	0.2	0.003	20	UpperCalabazas	46.6	1.27	27
AC_unk15	1.3	0.026	20	RedwoodCreek	28.2	0.76	27
PointSanPabloPeninsulaNorth	0.9	0.017	20	PointSanPabloPeninsulaNorth	0.9	0.02	27
RichmondInnerHarbor	1.4	0.026	19	DavisPoint	4.9	0.13	27
AC_unk12	0.1	0.003	19	UpperNapaRiver	15.9	0.41	26
AC_unk14B	0.3	0.005	19	HeathCanyon	40.9	1.04	25
AC_unk17	0.5	0.008	17	LowerCalabazas	48.6	1.22	25
SMC_unk06	0.4	0.007	16	PointSanPabloPeninsulaWest	4.0	0.10	25
AC_unk09	13.8	0.215	16	RectorReservation	9.8	0.25	25
SanPablo	1.6	0.024	15	SimmonsCanyon	34.1	0.84	25
AC_unk06	0.1	0.002	15	BearCanyon	37.9	0.94	25
WalnutCreek	6.4	0.094	15	UpperDryCreek	24.4	0.60	24
AC_unk16	0.4	0.005	14	BellCanyonReservoir	10.9	0.26	24
SMC_unk05	0.4	0.006	14	RitchieCreek	35.4	0.85	24
YerbaBuenaIsland	0.5	0.008	14	GarnettCreek	20.6	0.49	24
SantaFeChannel	7.8	0.105	14	LawlorRavine1	3.4	0.08	24

5. Summary and Recommendations

The modeling effort in Years 4 and 5 was focused on improving the calibration of the PCB and Hg models. The major changes included switching from a sediment-based model to a water-based model, elimination of double counting of source areas on top of general land uses, changes in the model calibration approach, and changes in land use grouping. As a result of these modifications, the Model calibration improved, and the regional PCB and Hg load estimates simulated from the models are conceptually reasonable and more consistent with previously reported loads estimates. Although there is still room for improvement in the calibration, the Models that have emerged from efforts during 2014 and 2015, especially the PCB model, may be ready to be used for planning purposes. The following steps are recommended to further improve model calibration in 2016:

- **Explore LU and SA grouping.** With the new calibration datasets and as QA of GIS data improves, there might be an opportunity to revisit the LU and SA grouping. Although the use of a global source area category contributed to encouraging results, provided more monitoring data is generated that will increase the number of calibration watersheds (and improved representation of some source areas), it might be possible to split the global SA group into a few source area sub-groups such as a global recycling group for PCB or a global atmospheric deposition group for Hg. In addition, there is the potential for improving the "clean parameter" for Hg by reconsidering the variation in sediment yield. Doing this may resolve concern over the relative production of Hg between urban and non-urban areas.

- **Calibrate the models with expanded datasets.** WY 2015 data collected at 19 additional watersheds will soon be available and added to the calibration dataset to further improve model calibration. With 41 watersheds, a subset can be used for model verification.
    - Exploring creating a fuzzy scale of fitness to gauge model calibration. To identify what factors or any patterns (i.e. close to a source area) that may have big influence on model calibration, it may be helpful to create a fuzzy scale of "good fit," "average fit," "bad fit" for model calibration and map it like colored pins with GIS layers of land uses and source areas. Any patterns discerned from this exercise will be useful for interpreting and validating model results.
    - Explore adding aerial deposition into Hg model. Mercury has a large input from air deposition, which is not a function of land use and therefore currently not included in the RWSM. Adding an aerial deposition (mass/time-area) into Hg model might help its calibration. The same deposition would be added to all the land uses, and the model would then only be simulating the extra Hg from land uses.

- **Get the RWSM ready for stakeholder use.** Once the model calibration is deemed reasonable and acceptable to stakeholders, the final step of model development is to get the model ready for stakeholder use. To the extent possible and as budget allows, the effort will begin with identification of target users and may also include: QA and clean up the GIS data; redesign of the model structure to make it more flexible; and creation of an improved model interface to make it more user-friendly.
    - Ensure that the model output tables and graphics are designed to feed directly into identified uses of the information:
        - A regional map of flow production
        - A regional concentration map
        - A regional yield map
        - Land use/source area parameter based event mean concentrations (ng/L)
        - Land use/source area parameter based yields (g/km^2)

- Use the model to explore basic concepts of land use change in the form of redevelopment. For example, the conversion of older industrial areas to mixed zone commercial and residential land use should cause a reduction in both PCB and Hg loads.
- Could also use the model to explore the likely increase in runoff and loads associated with conversion of open space or agricultural land to mix urban as a means for providing a rationale to developers to implement BMPs to manage that potential impact.
- Could model other pollutants. For example pyrethriod pesticides. Given the regional nature of the model, could compare the regional pyrethriod usage data that is documented on a county by county basis to estimate the percentage of pyrethriod use that gets into our stormwater system. Use the model to answer the simple question: Is the stormwater load 0.1%, 1%, or 10% of the annual average use?

6. References

ABAG 2005. Land use map for the region including a description of land use classifications categories. Association of Bay Area Governments (ABAG), Oakland, CA.

ACCWP, 2014. Alameda Countywide Clean Water Program Integrated Monitoring Report. http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/stormwater/Municipal/IMR/Alameda_Countywide_2014.pdf

CCCWP, 2014. Contra Costa Clean Water Program Integrated Monitoring Report – Part C: Pollutants of Concern Implementation Plan. http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/stormwater/Municipal/IMR/CC_Countywide_C_2014.pdf

Gilbreath, A.N., Gluchowski, D.C., Wu, J., Hunt, J.A., and McKee, L.J., 2014. Pollutants of concern (POC) loads monitoring data progress report, water year (WYs) 2012 and 2013. A technical report prepared for the Regional Monitoring Program for Water Quality in San Francisco Bay (RMP), Sources, Pathways and Loadings Workgroup (SPLWG), Small Tributaries Loading Strategy (STLS). Contribution No. 708. San Francisco Estuary Institute, Richmond, California. http://www.sfei.org/documents/pollutants-concern-poc-loads-monitoring-data-progress-report-water-years-wys-2012-and-2013

Gilbreath, A.N., Hunt, J.A., Wu, J., Kim, P.S., and McKee, L.J., 2015. Pollutants of concern (POC) loads monitoring progress report, water years (WYs) 2012, 2013, and 2014. A technical report prepared for the Regional Monitoring Program for Water Quality in San Francisco Bay (RMP), Sources, Pathways and Loadings Workgroup (SPLWG), Small Tributaries Loading Strategy (STLS). Contribution No. 741. San Francisco Estuary Institute, Richmond, California.

Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J.D., and Megown, K., 2015, Completion of the 2011 National Land Cover Database (NLCD) for the conterminous United States-Representing a decade of land cover change information. Photogrammetric Engineering and Remote Sensing, v. 81, no. 5, p. 345-354

Lent, M.A. and McKee, L.J., 2011. Development of regional suspended sediment and pollutant load estimates for San Francisco Bay Area tributaries using the regional watershed spreadsheet model (RWSM): Year 1 progress report. A technical report for the Regional Monitoring Program for Water Quality, Small Tributaries Loading Strategy (STLS). Contribution No. 666. San Francisco Estuary Institute, Richmond, CA. http://www.sfei.org/sites/default/files/RWSM_EMC_Year1_report_FINAL.pdf

Lent, M.A., Gilbreath, A.N., and McKee, L.J., 2012. Development of regional suspended sediment and pollutant load estimates for San Francisco Bay Area tributaries using the regional watershed spreadsheet model (RWSM): Year 2 progress report. A technical progress report prepared for the Regional Monitoring Program for Water Quality in San Francisco Bay (RMP), Small Tributaries Loading Strategy (STLS). Contribution No. 667. San Francisco Estuary Institute, Richmond, California. http://www.sfei.org/sites/default/files/RWSM_EMC_Year2_report_FINAL.pdf

McKee, L.J., Gilbreath, A.N., Hunt, J.A., and Greenfield, B.K., 2012. Pollutants of concern (POC) loads monitoring data, Water Year (WY) 2011. A technical report prepared for the Regional Monitoring Program for Water Quality in San Francisco Bay (RMP), Small Tributaries Loading Strategy (STLS). Contribution No. 680. San Francisco Estuary Institute, Richmond, California. http://www.sfei.org/documents/pollutants-concern-poc-loads-monitoring-data-water-year-wy-2011

McKee, L.J., Gluchowski, D.C., Gilbreath, A.N., and Hunt, J.A., 2013. Pollutants of concern (POC) loads monitoring data progress report, water year (WY) 2012. A technical report prepared for the Regional Monitoring Program for Water Quality in San Francisco Bay (RMP), Sources, Pathways and Loadings Workgroup (SPLWG), Small Tributaries Loading Strategy (STLS). Contribution No. 690. San Francisco Estuary Institute, Richmond, California. http://www.swrcb.ca.gov/rwqcb2/water_issues/programs/stormwater/UC_Monitoring_Report_2012.pdf

McKee, L.J., Gilbreath, A.N., Wu, J., Kunze, M.S., Hunt, J.A., 2014. Estimating Regional Pollutant Loads for San Francisco Bay Area Tributaries using the Regional Watershed Spreadsheet Model (RWSM): Year's 3 and 4 Progress Report. A technical report prepared for the Regional Monitoring Program for Water Quality in San Francisco Bay (RMP), Sources, Pathways and Loadings Workgroup (SPLWG), Small Tributaries Loading Strategy (STLS). Contribution No. 737. San Francisco Estuary Institute, Richmond, California. http://www.sfei.org/documents/estimating-regional-pollutant-loads-san-francisco-bay-area-tributaries-using-regional-wate

McKee, L.J. Gilbreath, N., Hunt, J.A., Wu, J., and Yee, D., 2015. Sources, Pathways and Loadings: Multi-Year Synthesis with a focus on PCBs and Hg. A technical report prepared for the Regional Monitoring Program for Water Quality in San Francisco Bay (RMP), Sources, Pathways and Loadings Workgroup

(SPLWG), Small Tributaries Loading Strategy (STLS). Contribution No. 773. San Francisco Estuary Institute, Richmond, California.

McKee, L.J., Gilbreath, A.N., Yee, D., and Hunt, J.A., 2016. Pollutants of concern (POC) reconnaissance monitoring draft progress report, water year (WY) 2015. A technical report prepared for the Regional Monitoring Program for Water Quality in San Francisco Bay (RMP), Sources, Pathways and Loadings Workgroup (SPLWG), Small Tributaries Loading Strategy (STLS). Contribution No. 787. San Francisco Estuary Institute, Richmond, California.

Park, M.-H., Swamikannu X., and Stenstrom, M.K., 2009. Accuracy and precision of the volume concentration method for urban stormwater modeling. Water Research 43, 2773-86.

SCVURPPP, 2014. Santa Clara Valley Urban Runoff Pollution Prevention Program Integrated Monitoring Report – Part C: Pollutants of Concern Load Reduction Opportunities. http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/stormwater/Municipal/IMR/SCV_C_2014.pdf

SFEI, 2009. RMP Small Tributaries Loading Strategy. A report prepared by the strategy team (**L McKee**, A Feng, C Sommers, R Looker) for the Regional Monitoring Program for Water Quality. SFEI Contribution #585. San Francisco Estuary Institute, Oakland, CA. http://www.sfei.org/sites/default/files/biblio_files/Small_Tributary_Loading_Strategy_FINAL.pdf

SFRWQCB, 2006. Mercury in San Francisco Bay: Proposed Basin Plan Amendment and Staff Report for Revised Total Maximum Daily Load (TMDL) and Proposed Mercury Water Quality Objectives. California Regional Water Quality Control Board San Francisco Bay Region, August 1st, 2006. 116pp. http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/TMDLs/sfbaymercury/sr080906.pdf

SFRWQCB, 2007. Total Maximum Daily Load for PCBs in San Francisco Bay Proposed Basin Plan Amendment and Staff Report. San Francisco Bay Regional Water Quality Control Board. Oakland, CA. December 4th, 2007. 178pp. http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/TMDLs/sfbaypcbs/Staff_Report.pdf

SFRWQCB, 2008. Guadalupe River Watershed Mercury Total Maximum Daily Load (TMDL) Project BASIN PLAN AMENDMENT. California Regional Water Quality Control Board San Francisco Bay Region October 8, 2008. http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/TMDLs/guadalupe_river_mercury/Guad_Hg_TMDL_BPA_final_EOcorrSB_clean.pdf

SFRWQCB, 2009. California Regional Water Quality Control Board San Francisco Bay Region Municipal Regional Stormwater NPDES Permit, Order R2-2009-0074, NPDES Permit No. CAS612008. Adopted October 14, 2009. 279pp.

http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/stormwater/Municipal/index.shtml

SFRWQCB, 2011. California Regional Water Quality Control Board San Francisco Bay Region Municipal Regional Stormwater NPDES Permit, Order R2-2009-0074, NPDES Permit No. CAS612008. Adopted October 14, 2009. Revised November 28, 2011 http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/stormwater/Municipal/R2-2009-0074_Revised.pdf

SMCWPPP, 2014. San Mateo Countywide Water Pollution Prevention Program Integrated Monitoring Report – Part C: Pollutants of Concern Load Reduction Opportunities. http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/stormwater/Municipal/IMR/San_Mateo_C_2014.pdf

7. Appendix - Double counting source areas

With 2015 option 1 (defined in Table 1), a range of land use and source area groupings were explored that included groupings focusing on the SAs most prevalent in the calibration watersheds and a series of "global" SAs that included all SAs, or logical subsets (e.g., electrical user based SAs, a waste recycling SA group, and an air emissions SA group). To aid the selection, Spearman Rank correlations between various LUs and SAs and mean concentrations and particle ratios of PCB and Hg were calculated, and a set of decision criteria first proposed by McKee et al. (2013) were formalized.

1) All LUs must be in the model (thus covering all land area at least once).
2) LUs cannot be grouped with SA categories in order to avoid double counting.
3) The LUs and SA categories with similar Spearman Rank correlations can and likely should be grouped together.
4) The grouping should generally follow our conceptual understanding of yield from various LUs and SA (see the pollutant profiles: McKee and Lent, 2011; McKee et al., 2013).
5) LU or SA groups should be presented at sufficient number of watersheds and runoff volume contributions from them should be sufficient as well.

Based on these criteria, five possible options for land use grouping were explored for the PCB model (Table 3) and five options for Hg model (Table 4). The logical subsets (e.g., electrical user based SAs, a waste recycling SA group, and an air emissions SA group) were dropped for further exploration at this time due to failure in relation to criterion 5 (sufficient representation). The model calibration was attempted for each of the groupings.

Assessment of both the PCB and Hg model calibration results suggests that a three LU group model (option 1 in Tables A1 and A2) met all the calibration criteria and performed the best. Although the model calibration with this grouping showed some promise, this model structure was deemed not desirable as it did not include any SAs, which are major pollution sources based on empirical data. Therefore, further steps were taken to explore the possibility of adding a global SA group to this LU-based model.

Table A1. Possible land use and source area groupings for the PCB model.

Land use or source area	Spearman Rank *PCB (pg/L) median*	Conceptual relative pollution	% volume contribution	Number of watersheds	Conceptual largest influence (Combined rank)	Option 1 Base model	Option 2 Base model (+)	Option 3 Base model (+)	Option 4 Base model (+)	Option 5 Base model (+)
LU Old Industrial	0.6	M	2.1%	20	41	Old Industrial	Old Industrial	Old Industrial	Old Industrial	Old Industrial
LU Old Commercial	0.3	-	8.9%	22	195	Old urban / new industrial	Old urban / new industrial	Old urban / new industrial	Old urban / new industrial	Old urban / new industrial
Transportation	0.3	-	20%	22	447					
LU Old Residential	0.1	-	16%	22	347					
LU New Industrial	0.1	M/L	3.1%	22	69					
LU New Commercial	-0.4	L	2.8%	13	36	Ag / Open / New Urban	Ag / Open / New Urban	Ag / Open / New Urban	Ag / Open / New Urban	Ag / Open / New Urban
LU Agriculture	-0.5	VL	1.7%	9	16					
LU New Residential	-0.5	L	4.5%	12	54					
LU New Transportation	-0.6	-	4.9%	14	68					
LU Open	-0.6	VL	36%	22	781					
SA manufMetals	0.7	M	0.2%	14	3		all SA	Metal / TranspRail	Metal / TranspRail	Metal / TranspRail
SA transpRail	0.5	M	1.4%	15	22					
SA transpAir	Insufficient data	?	0.3%	2	0.6				Other SA	Recycle
SA recycMetals	0.2	M/L	0.1%	5	0.3					
SA recycAuto	0.2	M/L	0.2%	10	1.7					
SA recycWaste	0.0	M/L	1.0%	9	9					
SA recycDrums	Insufficient data	H	0.0%	2	0.0					
Marine repair scrap yards	Insufficient data	M/L	-	-	-					
SA military	Insufficient data	H	0.0%	1	0.0					
SA electricTransf	0.0	VH	0.1%	13	0.7					
SA electricPower	Insufficient data	VH	0.0%	1	0.0					

Table A2. Possible land use and source area groupings for the Hg model.

Land use or source area	Spearman Rank HgT (ng/L) median	Conceptual relative pollution	% volume contribution	Number of watersheds	Conceptual largest influence (Combined rank)	Option 1 Base model	Option 2 Base model (+)	Option 3 Base model (+)	Option 4 Base model (+)	Option 5 Base model (+)
LU New Industrial	-0.18	H	3.1%	23	72	Industrial	Industrial	Industrial	Industrial	Industrial
LU Old Industrial	-0.27	H	2.1%	21	44					
LU Old Residential	0.28	M	16%	23	363	Old urban	Old urban	Old urban	Old urban	Old urban
LU Old Commercial	-0.04	-	8.9%	23	204					
LU Old Transportation	0.29	-	20%	23	467					
LU Open	0.33	VL	35%	23	816	Ag / Open / New Urban	Ag / Open / New Urban	Ag / Open / New Urban	Ag / Open / New Urban	Ag / Open / New Urban
LU Agriculture	0.23	VL	1.7%	9	15					
LU New Residential	0.32	-	4.5%	12	54					
LU New Commercial	0.29	-	2.8%	13	36					
LU New Transportation	0.20	L	4.9%	14	68					
SA manufMetals	-0.46	MH	0.2%	14	3		All Source Areas	ManufMetals		
SA recycAuto	-0.24	VH	0.2%	10	2				Recycle	
SA recycMetals	-0.17	VH	0.1%	5	0					
SA recycWaste	-0.06	VH	1.0%	10	10					
SA recycDrums	Insufficient data	VH	0.0%	2	0					
Marine repair scrap yards	Insufficient data	VH	-	-	-					
SA crematoria	0.22	MH	6.9%	9	62					
SA cement	Insufficient data	MH	0.3%	1	0					
SA Refinaries and petroche	Insufficient data	MH	-	-	-					
SA electricPower	Insufficient data	H	0.0%	1	0					
SA transpRail	-0.01	M	1.5%	16	23					TranspRail
SA electricTransf	0.21	-	0.1%	13	1					
SA transpAir	Insufficient data	?	0.3%	2	1					
SA military	Insufficient data	-	0.0%	1	0					