Optimizing sampling methods for pollutant loads and trends in San Francisco Bay urban stormwater monitoring

Technical Report

Melwani, A.R., Lent, M., Greenfield, B.K., and McKee, L.

July 2010.
REPORT SUMMARY

The purpose of this document is to summarize efforts to evaluate the optimization of sampling methods for pollutant loads and trend monitoring at Guadalupe River (GR) and Zone 4 Line A (Z4LA). This report presents a technical evaluation of sampling methods, load estimators, and strategies for storm selection. The sampling optimization focused on Hg, PCBs, and suspended sediments (SS) since these are the high management priorities in San Francisco Bay. The information summarized here will facilitate further discussion to develop appropriate study designs to address MRP questions and priorities at these and future sites. The focus of this study was to evaluate sampling designs for obtaining annual loads estimates. The study included two components:

- Comparison of the accuracy and precision of a variety of stormwater monitoring designs and mathematical equations (estimators) for determining annual pollutant loads; and
- Determining the power and sample size needed to detect declining trends in Hg and PCBs in the next 10 - 40 years.

The MRP default design is the automated sampling of four random storms using a composite sample method. The estimated range in bias (-50 – 13%) and standard error (4.3 – 6.5%) for the default MRP method was among the highest of the designs evaluated. Alternatives were explored such as increasing the number of samples and storms to six or 10 storms using a composite sampling method. Although sampling of 10 storms would provide better precision than four or six storms, a design with 10 storms would likely exceed budgetary limits. A six storm sampling strategy was simulated to include the first flush and largest storm. This design produced a similar range in bias (-16 – 31%) and standard error (1.4 – 3.6%) to the sampling of four storms (-13 – 57% and 2.2 – 5.0%, respectively). It is likely that the small improvement in precision with six storms would not warrant the extra on-going cost for this design, but inclusion of first flush and largest storms may warrant consideration. Automated sampling of two, four, or six storms using a discrete sampling method was also explored. The total number of samples was assumed to remain the same in each scenario, thus the range in bias (-7 – 4%) and standard error (0.1 – 1.4%) of these designs did not change. The best configuration was four storms (3 samples per storm).

The addition of turbidity was also explored using the turbidity surrogate regression estimator for the loads calculation method. This method produced the highest accuracy and least bias of all the alternative designs. To use regression on the turbidity surrogate records for estimating annual loads, at least 10 but ideally 16 samples per year should be collected at each site. Given results from the discrete among-storm evaluations, it is likely that scenarios that include first flush and one of the largest storms of the year would provide more robust loads estimates than random sampling alone when applying the turbidity surrogate method.

Power for detecting trends appeared to be possible with just 10 samples collected per year, based on a preliminary scenario in which the samples were randomly selected and did not confirm to any of the tested sampling designs.
ACKNOWLEDGEMENTS

The authors are grateful to The Regional Monitoring Program for Water Quality in San Francisco Bay (RMP), through its Sources, Pathways, and Loadings Workgroup (SPLWG) for funding this study. We thank Jim Griswold from Western Ecosystems Technology, Inc. (www.west-inc.com) for statistical guidance. The project work plan and draft report were guided by regular and timely input received from Arleen Feng (ACPWA / BASMAA), Richard Looker (Water Board), and Chris Sommers (EOA / BASMAA).

This report can be cited as:
TABLE OF CONTENTS

Summary

Acknowledgements

Introduction

Methods
   “Best estimate” loads calculations
   Loads analysis
   Trend analysis

Results
   Optimal sampling designs for estimating annual loads
      - Guadalupe River
      - Zone 4 Line A
   Optimal turbidity surrogate regression design for estimating annual loads
      - Guadalupe River
      - Zone 4 Line A
   Trend analysis
      - Guadalupe River
      - Zone 4 Line A

References

Tables

Figure Captions

Figures

Appendices

Appendix A – SSC: Hg regressions at Z4LA
Appendix B – Literature review of mercury variation by McKee et al. (2004)
Appendix C – Within-storm sampling designs (complete results for flow and turbidity-based storm selection criteria)
Appendix D – Among-storm sampling designs (turbidity-based storm selection)
Appendix E – Among-storm loads optimization (PCB and SS loads for flow-based storm selection criteria at Guadalupe River)
INTRODUCTION

Worldwide, coastal ecosystems adjacent to large urban, industrial and agricultural centers are subject to contamination, toxicity, and subsequent demise of wildlife and fisheries (Lauenstein and Daskalakis 1998, Linkov et al. 2002, Trimble 2003, Newton and Mudge 2005). In response, environmental laws in many countries are being developed and implemented to slowdown or reverse the process of contamination and even restore lost ecosystem attributes. In many cases, estimation of ecosystem-scale mass loads emanating from sources is one of the first data requirements needed to develop a plan of action (e.g., Godfrey et al. 1995, Schiff and Bay 2003, Balcom et al. 2004). San Francisco Bay is one such ecosystem that has been highly impacted by a history of urban, industrial, agriculture and mining land uses spanning about 150 years (Flegal et al. 2007). Approximately seven million people currently live in the nine counties bordering the Bay, and runoff and contaminants from mining legacies, urban areas, and agriculture drain to the Bay from about 37% of California (McKee et al. 2006a, David et al. 2009). Today, mercury (Hg) and polychlorinated biphenyls (PCBs) are considered the greatest threat to human and wildlife uses of the Bay (Conaway et al. 2007, Davis et al. 2007, Flegal et al. 2007, Yee et al. in review). However, there are also concerns about a number of emerging contaminants (Oros et al. 2003, Hoenicke et al. 2007, She et al. 2008).

In San Francisco Bay, urban runoff is considered one of the largest controllable sources of pollutant discharge. Total maximum daily loads (TMDL) reports written by the local Regional Water Quality Control Board (Water Board), summarize current estimates of loads from the main sources and pathways (urban and industrial wastewater, urban stormwater, Central Valley rivers, atmospheric deposition). The TMDL reports also argue for studies linking loads and toxic effects to beneficial uses, and provide loads allocations for each source and pathway (SFRWQCB 2006, 2008). The allocations are particularly stringent for urban stormwater and allow for 82 kg of Hg and 2 kg of PCBs with the objective of improving water quality in the Bay to desirable standards in 20 years (2028 for Hg and 2030 for PCBs). These represent estimated reductions of 50% and 90% over the present load estimates of 160 kg of Hg and 20 kg of PCBs, respectively. However, these load estimates remain uncertain, since measurements have only been made in a few of over two hundred tributaries (SFEI 2010).

The Regional Monitoring Program for Water Quality in San Francisco Bay (RMP), through its Sources, Pathways, and Loadings Workgroup (SPLWG), has been conducting tributary loading studies for nine years. The focus has been to provide information on sediment and pollutant transport processes in urban watersheds around the Bay (McKee et al. 2004, McKee et al. 2005, McKee et al. 2006b, Davis et al. 2007, Oram et al. 2008, David et al. 2009, McKee and Gilbreath 2009). The primary objective of these studies has been to achieve precise and unbiased estimates of loads of particle-associated pollutants-of-concern (particularly, Hg, PCBs, and suspended sediments). Most of the sampling effort has been focused on three locations: Mallard Island on the Sacramento River; Guadalupe River in San Jose; and the Zone 4 Line A flood control channel in Hayward. At all three study locations, a turbidity surrogate methodology has been employed, as it has been reported to be an appropriate and cost-effective method for accurate and unbiased particulate loads calculation (Grayson 1996, Wall et al. 2005). The tributary loading studies have provided valuable information for the development of the
San Francisco Bay and Guadalupe River Hg TMDLs (Austin 2006, SFRWQCB 2006), and the Municipal Regional Stormwater NPDES Permit (MRP) (SFRWQCB 2009).

The TMDLs and the MRP call for the Bay Area Stormwater Management Agencies Association (BASMAA) to improve loads information. In response, the RMP developed a small tributaries loading strategy (STLS) to guide the development of loads information over the next five years and to ensure coordination between the RMP and BASMAA. The STLS and Provision C.8.e. of the MRP aim to answer the following management questions:

1. Identify which Bay tributaries (including stormwater conveyances) contribute most to Bay impairment from pollutants of concern,
2. Quantify annual loads or concentrations of pollutants of concern from tributaries to the Bay,
3. Quantify the decadal-scale loading or concentration trends of pollutants of concern from small tributaries to the Bay, and
4. Quantify the projected impacts of management actions (including control measures) on tributaries and identify where these management actions should be implemented to have the greatest beneficial impact.

All of these questions require some level of information on the concentrations and mass loads in tributaries but the focus here is on optimization of study design for questions 2 and 3.

There are a number of sampling methods and corresponding mathematical calculation methods available for developing mass loading information. The optimal balance of sampling frequency within and among storm events is important to achieve precise, accurate, and cost effective loads measurements. Several methods, such as random and time-interval based sampling designs, have already been evaluated in other studies and essentially rejected as ineffective methods for evaluating tools towards our management questions (Walling 1985, 1988, Leecaster et al. 2002, Ma et al. 2009), and thus need not be evaluated further. Other calculation methods, including flow-weighted means, have been tested previously in southern California (Leecaster et al. 2002, Ma et al. 2009), and additional methods (turbidity surrogate, simple means, and linear interpolation) were examined in this study.

The focus of this study was to evaluate sampling designs for obtaining annual loads estimates. The study included two components:

1. Comparison of the accuracy, precision, and cost of a variety of stormwater monitoring designs and mathematical equations (estimators) for determining annual pollutant loads; and
2. Determining the power and sample size needed to detect declining trends in Hg and PCBs in the next 10 - 40 years.

High quality loading data from local watersheds collected by the SPLWG and RMP provided a resource to evaluate potential future monitoring approaches. A variety of sampling and mathematical loads calculation methods were simulated by statistically subsampling the existing high temporal resolution empirical data sets. Combining empirical data with simulation methods to test and optimize loading measurements has been carried out in numerous studies before (e.g., Walling 1985, Walling 1988, Leecaster et al. 2002, Ma et al. 2009). This study focused on sampling optimization for Hg, PCBs, and suspended sediments (SS) since these are the high management priorities in San
Francisco Bay, and suspended sediment concentration and load is an important vector for transport of sediment-associated pollutants. However, the findings are likely relevant for other particulate substances in similar sized watersheds.

METHODS

Three years of urban runoff data from the Guadalupe River (GR) and Zone 4 Line A (Z4LA) monitoring stations in San Francisco Bay, California were statistically analyzed in this study. GR is located near San Jose, the largest city in the San Francisco Bay Area. Its watershed is the fourth largest in the Bay Area (about 500 km$^2$) and is a network of mostly natural channels that have been modified by impoundments and flood control engineering. The monitoring station operated by the SPLWG developed through collaboration with the US Geological Survey (USGS) (station number 11169025) is upstream from tidal influence, but resides downstream from five main reservoirs, the City of San Jose, and the majority of flood control channels. The typical flood hydrograph produced by heavy rainfall passes through this watershed over a period of about 12-24 hours but larger and later season floods may last for several days. The free flowing area downstream from reservoirs is 236 km$^2$, of which approximately 80% is urbanized landscape. In addition, this area also drains the Quicksilver County Park, formerly the New Almaden Mining District, which, since 1849, has produced 6% of the total Hg worldwide (McKee, L., unpublished data) and is a known source of Hg to San Francisco Bay.

In contrast, the monitoring station in Z4LA, located in Hayward, drains a relatively small 4.47 km$^2$ watershed of completely urbanized landscape with over 38% industrial land use. Historically, there was no creek draining this area of the Bay margin. The flood channel of Zone 4 Line A is entirely engineered with approximately one-third open to the air and two-thirds underground culverts and storm drains. The monitoring station resides approximately 1.7 km from the Bay and upstream from tidal influence. There are no reservoirs in this watershed and rain passes largely unabated through the network of flood channels in minutes to hours.

Though an excellent data set is also available for the Mallard Island on the Sacramento River for a range of trace contaminants (David et al. 2009), these data were not included in the present study. There were two primary reasons for this: 1. the system is not representative of typical small tributaries to be monitored in the MRP, due to extreme size of the watershed (covering two-thirds of the land area of California); 2. time and resource limitations precluded the detailed examination necessary to evaluate optimal sampling design for this system. Since the system is much larger than GR, Z4LA, and others watersheds previously studied (e.g., Leecaster et al. 2002, May and Sivakumar 2009), and has considerably higher water volume and longer flood wave travel times, an optimal sampling strategy for this system is likely quite different. Performing a similar exercise for larger watersheds such as represented by our sampling station on the Sacramento River at Mallard Island remains a data gap in the published literature.

Three water years (WY) of data that spanned a range in climatic conditions (and thus a range in wet weather discharge and pollutant concentrations) were chosen for each watershed. A water year begins October 1st each year and ends September 30th and is
designated by the end date. At GR, during WYs 2003 – 2005, the peak discharge was 172 m³/s on December 16th, 2002 in WY 2003 (Table 1). In comparison, WYs 2004 and 2005 had relatively lower maximum discharge and pollutant concentrations. At Z4LA, three water years of data were also available, although not all years have complete records due to permitting issues and upstream construction. During WYs 2007 - 2009, Z4LA wet season discharge varied from 4.7 – 6.7 m³/s. WY 2008 exhibited relatively lower peak wet season discharge than the other two years, but this was predominated by many small storms.

“Best Estimate” Loads Calculations

Statistical simulations of sampling designs were compared to existing “best estimates”. The “best estimate” of annual loads (based on wet season data) was determined for each watershed based on the mathematical combination of estimated pollutant concentration and discharge volume. Hg, PCBs, and suspended sediment (SS) loads were examined. At both study locations, a turbidity surrogate regression (TSR) methodology has been used. Specifically, turbidity was monitored at short time intervals (15 minutes or less) and a statistical regression developed with a subset of water samples analyzed for suspended sediment concentration. This turbidity: SSC regression was combined with the continuous turbidity measurements to generate a time-continuous SSC record.

Additional depth integrated water samples (10-40 samples per year) were collected manually during high-flow events (storms), and analyzed for trace contaminants. Clean hands protocols were used. All analytical results were certified by the RMP data management and quality assurance plan (Lowe et al. 1999). Subsequently, during well sampled floods, linear interpolation was used to estimate concentrations between data points which were then combined with short interval flow measurements to determine loads. During storm periods when no sampling was conducted or during dry weather flows, regression relationships were determined between turbidity and each pollutant and used to calculate time-continuous estimates of contaminant concentration (turbidity surrogate regression or TSR). These estimates were then combined with discharge measurements to calculate loads. These combined methods were used to determine the “best estimate” of annual pollutant concentrations and loads to these watersheds over the years. It has been acknowledged in previous reporting (McKee et al. 2006b) that under complex conditions (e.g., Guadalupe River in 2005) professional judgment was used to guide these calculations. For example, Hg loads in GR were often stratified based on the predominant source of runoff indicated on hydrographs, resulting in separate regression relationships for urban vs. non-urban signals. These professional-judgment-based turbidity surrogate load estimates were used as the best available load estimates, against which all sampling design scenarios were compared.

Loads Analysis

Sampling programs for watershed loads estimates are designed with two attributes in mind; the number of samples taken within a storm and the number of storms sampled during a year or wet season (Leecaster et al. 2002). Our analysis of the optimal sampling
method was performed in two steps. First, within-storm load estimates were compared among sampling designs and calculation methods. Secondly, using the results of the optimal within-storm sampling strategy, scenarios among storms were examined. A number of prospective designs were considered, including variations on sample collection and loads calculation within individual storms (i.e., within-storm designs), and sampling designs across the wet season (i.e., among-storm designs). All designs were evaluated by Monte Carlo simulation. Bias and precision were calculated to evaluate all design scenarios. Bias was calculated as the median percentage difference between the expected estimate and the actual results. Precision was calculated as the variability in bias, measured as the standard deviation. Results were calculated and compiled using the statistical software package R (v. 2.10.1).

The within-storm design analysis considered four aspects of sampling design: 1. storm sampling basis (i.e., flow based vs. turbidity based); 2. sampling emphasis; 3. sample size; and 4. loads estimation method (Table 2). To examine the first aspect of the analysis, each year of data for the two watersheds was analyzed for the presence of sampled storms using a) flow thresholds and b) turbidity thresholds (Table 3a and 3b). The use of a pre-set threshold simulates protocols for an automated sampler, which will collect water samples when preprogrammed thresholds are surpassed that characterize flow and concentration during each storm event hydrograph. Since both flow and turbidity increase during high-flow events, both were evaluated for use as primary drivers for sampling design. To define and select a storm, thresholds were statistically established for the start, peak, and end of each storm hydrograph. Storm selection criteria differed between the two watersheds. For GR, flow events greater than 200 cfs, with peak flow greater than 736 cfs, were flagged as storms (Table 3a). For Z4LA, flow events greater than 5 cfs, with peak flow greater than 26 cfs, were flagged as storms. The storm selection criteria were chosen to achieve thorough coverage of storm flow events, without including baseflow events. For the purposes of this analysis, flow that did not meet these criteria was deemed baseflow. In contrast to flow thresholds, turbidity-based storm selection thresholds were similar between GR and Z4LA (Table 3b). For GR, turbidity measurements greater than 30 NTU, during storms with peak turbidity greater than 84 NTU, were flagged as storms. For Z4LA, the thresholds were 30 NTU and 89 NTU, respectively.

The second aspect of the within-storm design was sampling emphasis, which refers to relative frequency of sample collection in the rising vs. falling stage. Two approaches were considered: a) equal spacing of the samples across the rising and falling stages (i.e., 1:1 sampling emphasis) or b) rising-stage emphasis, where twice as many samples were spaced on the rising stage of the hydrograph relative to the falling stage (i.e., 2:1 sampling emphasis). The rationale for considering a rising stage emphasis is that suspended sediment pollutant loads are typically greater and more variable during the rising stage (McKee et al. 2006b).

The third aspect of within-storm design evaluated was the number of water samples collected per storm (i.e., sample size). For Hg and SS, four sample sizes were considered: 6, 12, 18, and 24 collections per storm. For PCBs, 6 and 12 collections per storm were considered; larger numbers of collection would be unfeasible due to the large sample volumes required for PCB lab analysis. The actual number of samples that could be evaluated for each scenario varied based on the size of each storm sampled.
The final aspect of within-storm design was the loads calculation method (Table 4, Equations 1 - 3). Methods for loads calculation will depend on the method used to integrate individual water collection events (Leecaster et al. 2002). Specifically, auto samplers may obtain discrete or composite samples. Discrete samples are small samples (referred to as “sips”) taken by the auto sampler throughout the storm. Composite samples are the combined collection of many discrete sips that are used to represent conditions over an entire storm. The data generated from composite samples collected in this manner are often referred to as event mean concentrations (USEPA 2002, Ma et al. 2009). For the discrete sampling method, two loads calculation methods were examined: 1) linear interpolation; and 2) flow weighted mean (Table 2). For the composite sampling, only a simple mean method was used to estimate loads because the other estimators require discrete data. These three loads calculation methods were tested for all combinations of sampling basis (flow or turbidity), number of samples collected per storm, and sampling stage emphasis (Table 2). The resulting within-storm load estimates were compared to the “best estimate” loads to assess performance of the sampling design and loads calculation methods.

The among-storm design evaluations focused on number and type of storms sampled. Using the results from the optimum within-storm design, three strategies for sampling among storms were considered for their ability to estimate annual loads (Table 5). The first among-storm design (Design A in Table 5) sampled the first flush (i.e., the first storm of a wet season) plus a variable number of random storms. The second among-storm design (Design B in Table 5) sampled the first flush plus one of the three largest storms of the wet season and a variable number of random storms (Note we chose one of the three largest because although it is easy to define and then respond to a weather forecast for a large storm, we may also miss a storm that ends up larger than the forecast predicted - we can never know until the end of the season if we sampled the largest storm of the season or one of the three largest). The third among-storm design (Design C in Table 5) is the design written in the MRP and was evaluated using either 2 or 4 random storms. Designs A and B were evaluated for 2, 4, 6, and 10 storms (actual number depends on available data). To correspond with the MRP requirements, Design C evaluated 2 and 4 storms only. All results were extrapolated to an estimated annual load by dividing by the ratio of sampled storm flow volume vs. total wet season flow volume (Table 4, Equation 4). Note that WY 2008 at Z4LA was deemed inappropriate for this analysis since the sampling began later in the season, and thus, an assessment of first flush and largest storms was not possible.

The accuracy and precision for annual loads calculation using each sampling strategy was compared. Monte Carlo simulations were employed to obtain random storm subsamples under each design method. Each design was run 50 times for each number of storms (to allow for random selection of storms), and an annual load was calculated for each run. The optimum strategy was identified as the design with the median closest to the annual best estimate of load, and the lowest variability in estimated annual loads. Each year of data was analyzed separately to demonstrate performance under a variety of climatic conditions.

A parallel analysis was performed to examine performance of the turbidity surrogate regression method for annual pollutant loads estimation. First, Monte Carlo simulation was employed to examine the sample size requirements for developing a
relationship of turbidity to Hg, PCBs, and SS. This was performed for each design within each year of data. The optimal sample size was examined by varying the number of samples in the TSR from 4 – 40 using Monte Carlo simulation (1000 simulations per design/year combination). Again, the actual sample sizes that could be simulated varied based on the empirical data sets. Each regression generated from a sub-sampled data set was converted into a continuous estimated pollutant concentration record by applying the regression to the continuous turbidity record. The continuous pollutant concentrations were then extrapolated to loads using the same methods as for the mass emission estimator. These annual loads based on sub-sampled data were compared against loads calculated using each water years’ complete grab sample data set to determine how many grab samples were necessary to obtain precise and unbiased load estimates. Once the optimum number of TSR samples was identified for each pollutant, the average regression slope and intercept (from the 1000 runs at the optimal sample number) were applied to the continuous flow and turbidity records to calculate per-storm loads. Finally, the TSR-based storm loads were extrapolated to annual loads using the same three among storms sampling strategies (Table 5), employing 50 runs per design. The performance of the TSR in the among-storm designs was compared to the TSR loads using all samples collected in a year, and the “best estimate” of annual loads.

*Trend Analysis*

Provision C.8.e. of the MRP calls for testing for trends towards compliance with loads allocations (SFRWQCB 2009). To support that provision, the objective of the trends analysis performed here was to determine the power to observe declining trends in the ratio of SSC to Hg concentration or SSC: PCB concentration given the current mean slope and variability. This is consistent with the presentation of TMDL targets on SSC normalized basis (SFRWQCB 2006, 2008). Trends were examined for reductions in the estimated particle concentration [mass/unit mass] from its current value to a value of 0.2 mg Hg / kg suspended sediment (i.e., the SSC: Hg target) and 0.002 mg PCB / kg suspended sediment. These targets assume that urban suspended sediment loads in the Bay Area average 400 million kg annually (following Lewicke and McKee 2009). Note that for the Z4LA Hg trend analysis, initial regression results demonstrated that the current SSC: Hg slope estimates were below the 0.2 target (Appendix A). Therefore, in this analysis, the trend was examined for a target value of 0.05 mg Hg / kg suspended sediment (75% below 0.2 mg/kg).

Power to observe trends were evaluated at time periods of 10, 20, 25, and 40 years. The analysis examined the power to detect a decline in SSC: Hg and SSC: PCBs (at alpha = 0.05) based on the coefficient of variation (CV = s.e. / mean). Sample sizes in future years were assumed to be the same as current (approximately 12 to 20 PCB samples per year and 15 – 50 Hg samples per year) or reduced to 7 or 10 samples per year. The CV was adjusted for the n = 7 and n = 10 scenarios. Although intuitively, one might expect CV to diminish over time, since a trait of cleaner systems is lower concentration variation (Appendix B), in the absence of information to quantitatively predict the shape of such a trend for Bay Area watersheds and pollutants of interest, power was evaluated assuming that the CV would show a linear decline over the time period evaluated.
RESULTS

Optimal Sampling Designs for Estimating Annual Loads

Guadalupe River

Within-storm sampling design scenarios for Hg, PCBs, and suspended sediment (SS) generally indicated linear interpolation to be the most accurate estimator of loads per storm (Table 6). Complete results are tabulated in Appendix C.1. For all three pollutants, using either the flow-based or turbidity-based storm sampling methods, the accuracy (median bias) and precision (variability of bias) was higher at n = 12 than at n = 6. However, accuracy and precision for Hg did not notably improve from n = 12 to n = 18 or n = 24. Variability of bias generally decreased with increasing sample size. No improvement in accuracy or precision was evident using a rising stage (2:1) emphasis compared to samples evenly spaced over a storm hydrograph (1:1). Based on these findings, linear interpolation was used to characterize annual loads in the among-storm scenarios. To obtain a cohesive analysis, a single storm sampling method and sample size was used for each pollutant and site combination (Table 7).

Evaluation of the three among-storm sampling designs at GR (Table 5) indicated that the number of sampled storms strongly affected accuracy of estimated pollutant loads (Hg results in Figure 1; PCBs and SS results in Appendix E). Scenario results were generally similar for Hg, PCBs and SS loads, and for flow based vs. turbidity based storm sampling. The Hg flow based selection results are described in further detail here (Figure 1), while PCB and SS results, as well as all turbidity-based selection results, may be found in Appendices D.1 and E. The highest sample size of storms evaluated generally resulted in the lowest variability and bias in loads estimates (Figure 1). In WY 2004 and 2005, Design A (including first flush) and Design B (including first flush plus largest storm) demonstrated pronounced increases in accuracy and precision with each increase in storm sampling frequency (Figure 1b, 1c). Depending on the available data for simulations, either 6 or 10 storms were optimal for reducing bias. Design C (random storms only) consistently exhibited the least precision of the three designs at GR. However, Design C also exhibited less bias for 2 and 4 storms than the other designs.

Zone 4 Line A

Consistent with GR results, linear interpolation was the most accurate estimator of Hg, PCBs, and suspended sediment (SS) loads within individual storms (Table 6; Appendix C.2). Flow weighted mean performed particularly poorly for Z4LA. This may suggest that flow and concentrations were not as closely related in the storms analyzed (strong hysteresis). Accuracy and standard deviation of Z4LA load estimates were improved at the higher sample sizes, using either the flow-based or turbidity-based storm selection methods. For Hg and SS, the bias and precision were similar at n = 18 and n =
For PCBs, the highest sample size \((n = 12)\) was optimal. Larger sample sizes were not evaluated due to the limitations on PCB sample volume in auto-samplers. Both the magnitude of bias and variability in bias generally decreased with increasing sample size, particularly for linear interpolation. Similar to GR, no change in accuracy or precision was evident using a rising stage \((2:1)\) emphasis versus evenly spaced \((1:1)\) emphasis. Based on these findings for Z4LA, the flow-based design with 18 samples per storm for Hg and SS, 12 samples per storm for PCBs, and linear interpolation was used to characterize annual loads in the among-storm scenarios (Table 7).

Simulation of the three strategies for sampling among storms indicated that 10 storms (the highest sample size) generally resulted in the lowest variability and bias in loads estimates (Figures 2, 3, and 4). Turbidity-based results are summarized in Appendix D.2. Designs A and B generally had the lowest variability for both Hg and PCB loads (Figures 2 and 3). In WY 2009, there was little difference between Designs A and B. In contrast, simulation of WY 2007 indicated more variable results among designs. For Hg, PCBs, and SS (Figures 2, 3, and 4), Design C with 4 storms sampled approximated the best estimate load as well as the other designs (i.e., similar accuracy) but was associated with much higher variability (i.e., lower precision).

### Optimal Turbidity Surrogate Regression Designs for Estimating Annual Loads

**Guadalupe River**

Similar loads estimates could be obtained by the turbidity surrogate regression (TSR) method with significantly fewer samples than the full available sample size. Generally, all simulations indicated median loads that were similar to the best estimate load, reflecting the close relationships of these pollutants to turbidity. Simulations of the TSR showed that variability in the load estimates was markedly reduced at sample sizes of 7 or more (Figures 5, 6, and 7). For example, the median Hg load in WY 2004 at \(n = 7\) was \(12.8 \pm 2\) kg, compared to the best estimate load of \(13.0\) kg. Monte Carlo simulations of the TSR also indicated that 7 samples were needed to accurately estimate for PCBs and SS loads. For example, median PCB load estimates in WY 2003 and 2004 were \(1.7 \pm 0.4\) kg and \(1.1 \pm 0.4\) kg, respectively, compared to loads determined using all samples of \(0.9\) kg and \(0.5\) kg, respectively. In WY 2005, a limited pool of samples \((n = 12)\) was available for PCB simulation, and thus annual loads exhibited wider variability. However, SS loads were well sampled in all years and thus were generally consistent. Based on the finding that 7 samples provided an adequate basis for TSR in most GR scenarios, among-storm sampling strategies were examined using this sample size.

Simulation of three among-storm sampling (Table 5) designs for annual loads estimation using TSR indicated that sampling 10 storms per year was optimal to approximate the best estimate loads (Figures 8 – 10). The error bounds for annual loads generally narrowed as the sample size increased, but there was considerable variability among years and pollutants. Simulation of Designs A and B most consistently produced the least bias estimates, but not in all cases. For example, WY 2004 results indicated similar Hg load estimates using either Design A or B, and wider variability for Design C. In contrast, estimated loads with Design C were more consistent in WY 2005 than either
of the other designs. PCB and SS loads were less variable than Hg loads, but still indicated that 10 storms were required for minimum bias in loads.

As a final comparison, the bias and precision in sampling 10 storms using either Design A or B (Table 5) were compared between linear interpolation and TSR. For all three pollutants, linear interpolation provided more accurate and precise estimation of the best estimate load in WY 2005 (Table 8). A sampling strategy employing first flush and 10 total storms (Design A) with linear interpolation suggested relatively high accuracy for Hg and PCB loads of approximately 10%. Using TSR, PCB loads had very low accuracy, suggesting variability on the order of 50%. For SS, linear interpolation with Design B suggested the best design, which was estimated to have accuracy of approximately 1% under the WY 2005 scenario.

Zone 4 Line A

Simulations of the TSR for Z4LA supported the GR results that similar estimates of loads could be obtained with significantly less samples than the full available sample size (Figures 11, 12, 13). Simulations for Hg and PCBs indicated that 7 samples per year were needed to accurately estimate loads each year. Although the median load estimate did not vary greatly with increasing sample size, variability was significantly reduced with 7 or more samples. Simulation of these datasets indicated that the TSR was robust at all sample sizes in WY 2007 and 2008. Due to lower sample size of PCBs in WY 2009, there was greater variability in the load estimates. The accuracy of SS loads was relatively high at all sample sized evaluated due to the larger number of available sample points for simulation. However, the variability in loads spanned more than two orders of magnitude at sample sizes less than 10 (particularly in 2007). In summary, storm sampling strategies based on TSR as described below, were examined using 7 samples for Hg and PCBs, and 10 samples for SS.

Simulation of three among-storm sampling (Table 5) designs for annual loads estimation using TSR indicated that sampling the maximum number of storms each year was optimal for minimum bias and precise load estimates. For Hg loads, sampling of 10 storms per year using Design A or B achieved the least amount of variability and most accurate loads in 2007 (Figure 14). Simulations using 2009 data, indicated 4 – 6 storms using Design A would be sufficient, as the median load and variability did not vary greatly at greater sampling intensity or when one of the largest storms was included. Design C under predicted the best estimate loads in WY 2007, but attained reasonably close estimation of the best estimate load in WY 2009. PCB and SS loads were similar to Hg and best approximated loads at Z4LA by sampling of 10 storms (Figures 15 and 16).

Finally, TSR was compared to linear interpolation to evaluate bias in loads using the first flush designs when sampling 10 storms in WY 2009 (Table 5). Using either TSR or linear interpolation, Design A indicated better accuracy and precision relative to Design B. PCB loads were the most variable of the three pollutants in both methods, with an estimated bias of around 30% relative to the best estimate. However, estimated loads were very accurate for Hg (~ 1%) using either method and represented similar levels of precision (5%). Interestingly, SS loads were generally more accurate using TSR, but exhibited less precision than linear interpolation.
Trend Analysis

Guadalupe River

Trend analysis indicated that the power to detect trends in SSC: Hg using the current sampling intensity (Table 9) was generally greater than 90%. In WY 2004 and 2005, future sample sizes could be reduced to 7 samples per year from current sample sizes of 37 and 52, respectively, without loss of power to detect trends in the next 10 – 40 years (Table 10). Due to the lower SSC: Hg slope estimate (1.14) and weak regression relationship in 2003 (CV = 2.35; R² = 0.30), there was very low power to detect trends in that year.

Estimates of power to detect trends in the SSC: PCB relationship were also generally high (> 90%). Based on WY 2003 and 2004 data, future sample sizes could be reduced to 10 samples per year from current sample sizes of 21 and 19, respectively (Table 10). However, trend analysis performed with the lower SSC: PCB slope estimate measured in WY 2005 (0.12) indicated that relatively high power would not be achieved for a 10 year trend. Overall, the power analysis suggested that fewer sample sizes at GR would not inhibit the ability to detect declines in Hg and PCB concentrations.

Zone 4 Line A

The trend analysis for Z4LA indicated the power to detect trends in SSC: Hg using the current sampling intensity (Table 11) was generally greater than 90%. Scenarios run with the SSC: Hg slope estimate from 2009, indicated that future sample sizes could be reduced to 10 samples per year from a current sample sizes of 30 without loss of power for trends in 20, 25, and 40 years (Table 12).

Estimates of power to detect trends in the SSC: PCB relationship were very high (> 95%). In WY 2007 and 2008, future sample sizes could be reduced to 7 samples per year from current sample sizes of 15 and 14, respectively, for all trend scenarios evaluated (Table 12). However, at a sample size of 7 per year, 95% power would only be achieved in WY 2007 for 25 and 40 year trends. Overall, the power analysis suggested that lower sizes would also not inhibit the ability to detect declines in Hg and PCB concentrations at Z4LA.

SUMMARY

- The optimal within-storm design in GR and Z4LA evaluations was an equal-spacing, flow or turbidity-based sampling method, with the linear interpolation estimator.

- The optimal among-storm design was highly dependant on sample size. When small numbers of storms were simulated per year, sampling strategies that included first flush or largest storms per year (i.e., Designs A and B) exhibited substantial upward bias in estimated annual load. The first flush and large storm events generally have
greater suspended sediment and pollutant concentrations than other storms; as a result, overemphasizing these events would result in overestimates of annual loads. Not surprisingly, the best estimates of annual loads were achieved in the largest sample sizes examined (10 storms per year).

- Designs that randomly sample storms throughout the year (i.e., Design C) without emphasizing first flush or large events have better accuracy at small sample sizes. However, these designs exhibit poor precision, with highly variable estimated loads.

- Evaluation of the turbidity-surrogate regression methods suggested that sampling frequency could be significantly reduced. For example, 10 storms sampled per year with one or two samples per storm were indicated.

- Results of the trend analysis indicated that power to detect long term trends in SSC: Hg concentrations and SSC: PCB concentrations should be high using a variety of sampling designs.

References


Lewicke, M., and L. J. McKee. 2009. Watershed specific and regional scale suspended sediment loads for Bay Area small tributaries. SFEI Contribution #566, San Francisco Estuary Institute, Oakland, CA.


McKee, L., J. Leatherbarrow, and R. Eads. 2004. Concentrations and loads of mercury, PCBs, and OC pesticides associated with suspended sediments in the Lower Guadalupe River, San Jose, California. 86, San Francisco Estuary Institute, Oakland, CA.


Table 1. Summary of Guadalupe River and Zone 4 Line A data examined in this study.

<table>
<thead>
<tr>
<th>Location</th>
<th>Water Years Examined</th>
<th>Peak Discharge (m^3/s)</th>
<th>Suspended Sediment FWMC (mg/L)</th>
<th>THg FWMC (ng/L)</th>
<th>TPCBs FWMC (ng/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guadalupe River</td>
<td>2003</td>
<td>172</td>
<td>204</td>
<td>2190</td>
<td>55</td>
</tr>
<tr>
<td>(GR)</td>
<td>2004</td>
<td>124</td>
<td>191</td>
<td>329</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>2005</td>
<td>112</td>
<td>79</td>
<td>140</td>
<td>45</td>
</tr>
<tr>
<td>Zone 4</td>
<td>2007</td>
<td>6.3</td>
<td>212</td>
<td>48</td>
<td>27</td>
</tr>
<tr>
<td>Line A (Z4LA)</td>
<td>2008*</td>
<td>6.7</td>
<td>350</td>
<td>60</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>4.7</td>
<td>109</td>
<td>23</td>
<td>11</td>
</tr>
</tbody>
</table>

* Data for partial year that included 2 dry season months.

Table 2. Design options examined for sampling within storms.

<table>
<thead>
<tr>
<th>Design Criteria</th>
<th>Design Options for Sampling Within Storms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Storm Sampling Basis</td>
<td>Flow-based or turbidity-based thresholds</td>
</tr>
<tr>
<td>2. Emphasis</td>
<td>Rising stage (2:1) or evenly spaced (1:1)</td>
</tr>
<tr>
<td>3. Max Sample Size (n) per Storm</td>
<td>24*, 18*, 12, 6</td>
</tr>
<tr>
<td>(actual n depends on storm size)</td>
<td></td>
</tr>
<tr>
<td>4. Load Calculation Methods**</td>
<td>LI – linear interpolation; FWM – flow-weighted mean</td>
</tr>
<tr>
<td>Composite Designs</td>
<td>SM – simple mean **</td>
</tr>
</tbody>
</table>

* Evaluated for Hg and suspended sediments only due to limitation on volumes required for PCB lab analysis
** Loads calculation methods differed for discrete vs. composite designs.
*** The other methods require discrete measurements

Table 3a. Flow-Based storm selection criteria by watershed.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Flow Thresholds for Storm Selection (cfs)</th>
<th>Minimum Peak Flow for Storm Selection (cfs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guadalupe River</td>
<td>200</td>
<td>736</td>
</tr>
<tr>
<td>Zone 4 Line A</td>
<td>5</td>
<td>26</td>
</tr>
</tbody>
</table>

Table 3b. Turbidity-Based storm selection criteria by watershed.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Turbidity Thresholds for Storm Selection (NTU)</th>
<th>Minimum Turbidity Peak for Storm Selection (NTU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guadalupe River</td>
<td>30</td>
<td>84</td>
</tr>
<tr>
<td>Zone 4 Line A</td>
<td>30</td>
<td>89</td>
</tr>
</tbody>
</table>
Table 4. Equations used to evaluate pollutant loads within- and among-storms.

<table>
<thead>
<tr>
<th>Within-storm Estimators</th>
<th>Among-storm Ratio Estimator</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Simple Mean</strong></td>
<td></td>
</tr>
<tr>
<td>Equation 1:</td>
<td></td>
</tr>
<tr>
<td>[ L_{S.M.} = \left( \sum_{j=1}^{N} [x_j]/N \right) \left( \sum_{i=1}^{n} Q_i \Delta t \right) ]</td>
<td></td>
</tr>
<tr>
<td><strong>Linear Interpolation</strong></td>
<td></td>
</tr>
<tr>
<td>Equation 2:</td>
<td></td>
</tr>
<tr>
<td>[ L_{L.I.} = \sum_{i=1}^{n} \left[ x_{i,\text{int}} \right] Q_i \Delta t ]</td>
<td>[ WY_{L,L,I.} = \sum_{k=1}^{M} V_k \times \sum_{k=1}^{m} L_k / \sum_{k=1}^{m} V_k ]</td>
</tr>
<tr>
<td><strong>Flow-weighted Mean</strong></td>
<td></td>
</tr>
<tr>
<td>Equation 3:</td>
<td></td>
</tr>
<tr>
<td>[ L_{F.W.} = \left( \sum_{j=1}^{N} [x_j] Q_j \right) \left( \sum_{i=1}^{n} Q_i \Delta t \right) / \left( \sum_{j=1}^{N} Q_j \right) ]</td>
<td></td>
</tr>
</tbody>
</table>

1 = Ratio Estimator used to calculate annual loads using optimal within-storm estimation method (i.e. Equations 1, 2 or 3).

Where, \( L \) = estimate of mass loading for a storm; \( WY_L \) = estimate of annual mass emissions; \( \Delta t \) = time interval between discharge measurements; \( N \) = number of samples taken during storm; \( n \) = number of time intervals in storm (based on frequency of discharge measurements); \([x_j]\) = concentration of sample \( j \); \([x_{i,\text{int}}]\) = \([x_j]\) interpolated to all \( n \) time intervals in storm; \( Q_i \) = discharge at time step \( i \); \( Q_j \) = discharge at sampling event \( j \); \( V_k \) = discharge volume for storm \( k \); \( m \) = number of storms sampled; and \( M \) = number of storms.

Table 5. Design options examined for sampling among storms.

<table>
<thead>
<tr>
<th>Design Criteria</th>
<th>Design Options for Sampling Among Storms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>Which storms</td>
<td>First flush, and</td>
</tr>
<tr>
<td>Total Number</td>
<td>random N</td>
</tr>
<tr>
<td>Storms (N)</td>
<td>10, 6, 4, 2</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* To account for selection uncertainty, the largest storm was selected randomly from the three highest total volume discharges per water year
** MRP design and sample sizes
Table 6. Comparison of bias in within-storm loads estimates for sample designs at Guadalupe River and Zone 4 Line A (sample emphasis = 1:1, flow-based criteria).

<table>
<thead>
<tr>
<th>Pollutant*</th>
<th>Simple Mean</th>
<th>Linear Interpolation</th>
<th>Flow-weighted Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hg (n = 12)</td>
<td>-0.21 ± 0.18</td>
<td>-0.05 ± 0.25</td>
<td>-0.02 ± 0.33</td>
</tr>
<tr>
<td>PCBs (n = 12)</td>
<td>-0.04 ± 0.30</td>
<td>0.02 ± 0.28</td>
<td>0.15 ± 0.35</td>
</tr>
<tr>
<td>SS (n = 12)</td>
<td>-0.20 ± 0.24</td>
<td>-0.01 ± 0.24</td>
<td>0.07 ± 0.42</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pollutant*</th>
<th>Simple Mean</th>
<th>Linear Interpolation</th>
<th>Flow-weighted Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hg (n = 18)</td>
<td>-0.14 ± 0.23</td>
<td>0.01 ± 0.23</td>
<td>0.17 ± 0.38</td>
</tr>
<tr>
<td>PCBs (n = 12)</td>
<td>-0.26 ± 0.26</td>
<td>-0.11 ± 0.30</td>
<td>-0.09 ± 0.45</td>
</tr>
<tr>
<td>SS (n = 18)</td>
<td>-0.17 ± 0.27</td>
<td>0.003 ± 0.26</td>
<td>0.15 ± 0.44</td>
</tr>
</tbody>
</table>

* Number in parentheses designates number of samples evaluated per storm

Table 7. Within-storm design strategies used for among-storm analyses.

<table>
<thead>
<tr>
<th>Site</th>
<th>Pollutant</th>
<th>Loads Estimation Method</th>
<th>Sampling Emphasis</th>
<th>Sample Size per Storm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guadalupe River</td>
<td>Hg</td>
<td>Linear Interpolation</td>
<td>Evenly Spaced</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>PCBs</td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>SSC</td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>Zone 4 Line A</td>
<td>Hg</td>
<td>Linear Interpolation</td>
<td>(1:1)</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>PCBs</td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>SSC</td>
<td></td>
<td></td>
<td>18</td>
</tr>
</tbody>
</table>
Table 8. Comparison of bias in median annual loads (+/- st. dev) resulting from turbidity-surrogate and linear interpolation in WY 2005 at GR and WY 2009 at Z4LA using two among storm sampling strategies (Design A and B in Table 5; N = 10 storms).

<table>
<thead>
<tr>
<th>Watershed</th>
<th>Pollutant</th>
<th>Water Year</th>
<th>Turbidity-surrogate</th>
<th>Linear Interpolation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Design A</td>
<td>Design B</td>
</tr>
<tr>
<td>Guadalupe River (GR)</td>
<td>Hg</td>
<td>2005</td>
<td>-0.20 ± 0.25</td>
<td>0.09 ± 0.22</td>
</tr>
<tr>
<td>Guadalupe River (GR)</td>
<td>PCBs</td>
<td></td>
<td>0.54 ± 0.22</td>
<td>0.57 ± 0.20</td>
</tr>
<tr>
<td>Guadalupe River (GR)</td>
<td>SS</td>
<td></td>
<td>0.09 ± 0.23</td>
<td>0.17 ± 0.22</td>
</tr>
<tr>
<td>Zone 4 Line A (Z4LA)</td>
<td>Hg</td>
<td>2009</td>
<td>-0.02 ± 0.05</td>
<td>-0.01 ± 0.04</td>
</tr>
<tr>
<td>Zone 4 Line A (Z4LA)</td>
<td>PCBs</td>
<td>2009</td>
<td>0.26 ± 0.09</td>
<td>0.33 ± 0.10</td>
</tr>
<tr>
<td>Zone 4 Line A (Z4LA)</td>
<td>SS</td>
<td></td>
<td>0.13 ± 0.08</td>
<td>0.19 ± 0.09</td>
</tr>
</tbody>
</table>

Table 9. Data used to examine power for trend analysis at Guadalupe River.

<table>
<thead>
<tr>
<th>Pollutant</th>
<th>Year</th>
<th>Current Sample Size</th>
<th>Mean Slope</th>
<th>S.D. Slope</th>
<th>95% C.I. Slope (lower, upper)</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hg</td>
<td>2003</td>
<td>25</td>
<td>1.14</td>
<td>13.4</td>
<td>-4.42, 6.70</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>37</td>
<td>1.44</td>
<td>0.73</td>
<td>1.20, 1.69</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>2005</td>
<td>52</td>
<td>2.23</td>
<td>1.34</td>
<td>1.86, 2.61</td>
<td>0.93</td>
</tr>
<tr>
<td>PCBs</td>
<td>2003</td>
<td>21</td>
<td>0.06</td>
<td>0.05</td>
<td>0.03, 0.08</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>19</td>
<td>0.11</td>
<td>0.11</td>
<td>0.06, 0.16</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>2005</td>
<td>12</td>
<td>0.12</td>
<td>0.18</td>
<td>0.00, 0.23</td>
<td>0.75</td>
</tr>
</tbody>
</table>
### Table 10. Estimates of power to detect trends in the slope of SSC: Hg and SSC: PCBs at Guadalupe River.

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Years to Reach Target</th>
<th>Power for Current Sample Size*</th>
<th>Power for Current Sample Size**</th>
<th>Power for n = 7</th>
<th>Power for n = 10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n = 7</td>
<td>n = 10</td>
<td>n = 7</td>
<td>n = 10</td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td>10</td>
<td>12</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>13</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>14</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>17</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td>10</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2005</td>
<td></td>
<td>10</td>
<td>100</td>
<td>98</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

* For 2003, n = 25; For 2004, n = 37; For 2005, n = 52
** For 2003, n = 21; For 2004, n = 19; For 2005, n = 12

### Table 11. Data used to examine power for trend analysis at Zone 4 Line A.

<table>
<thead>
<tr>
<th>Pollutant</th>
<th>Year</th>
<th>Current Sample Size</th>
<th>Mean Slope</th>
<th>S.D. Slope</th>
<th>95% C.I. Slope (lower, upper)</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hg</td>
<td>2007</td>
<td>30</td>
<td>0.13</td>
<td>0.11</td>
<td>0.09, 0.17</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>15</td>
<td>0.38</td>
<td>0.61</td>
<td>0.04, 0.72</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>21</td>
<td>0.08</td>
<td>0.05</td>
<td>0.05, 0.10</td>
<td>0.71</td>
</tr>
<tr>
<td>PCBs</td>
<td>2007</td>
<td>18</td>
<td>0.06</td>
<td>0.07</td>
<td>0.03, 0.10</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>15</td>
<td>0.16</td>
<td>0.13</td>
<td>0.09, 0.23</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>14</td>
<td>0.08</td>
<td>0.02</td>
<td>0.07, 0.10</td>
<td>0.98</td>
</tr>
</tbody>
</table>
Table 12. Estimates of power to detect trends in the slope of SSC: Hg and SSC: PCBs at Zone 4 Line A. Note that the SSC: Hg are currently below the target of 0.2 mg/kg, therefore the trend was examined for a target of 0.05 mg/kg, which is 75% below the original target of 0.2 mg/kg for urban stormwater (see Methods).

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Years to Reach Target</th>
<th>Hg</th>
<th>PCBs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Power for Current Sample Size*</td>
<td>n = 7</td>
<td>n = 10</td>
</tr>
<tr>
<td>2007</td>
<td>10</td>
<td>98</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>100</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>100</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>100</td>
<td>97</td>
</tr>
<tr>
<td>2008</td>
<td>10</td>
<td>62</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>87</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>93</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>99</td>
<td>85</td>
</tr>
<tr>
<td>2009</td>
<td>10</td>
<td>74</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>94</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>98</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>100</td>
<td>84</td>
</tr>
</tbody>
</table>

* For 2007, n = 30; For 2008, n = 15; For 2009, n = 21
** For 2007, n = 18; For 2008, n = 15; For 2009, n = 14
Figure Captions

Figure 1. Comparison of annual Hg loads at Guadalupe River in 2003-2005 based on three designs for sampling among storms (Table 5). Loads were calculated using linear interpolation with flow-based storm selection criteria. Design A simulated sampling of the first flush only and a variable number of random storms. Design B simulated the first flush plus one of the three largest storms and a variable number of random storms. Design C only tested the random storm component. -------- = best estimate Hg load for year (96 kg, 13 kg, and 7 kg, respectively).

Figure 2. Comparison of annual Hg loads at Zone 4 Line A in 2007 and 2009 based on three designs for sampling among storms (Table 5). See Figure 1 caption and text for further information. -------- = best estimate Hg load for year (17 g and 11 g, respectively).

Figure 3. Comparison of annual PCB loads at Zone 4 Line A in 2007 and 2009 based on three designs for sampling among storms (Table 5). See also Figure 1 caption and text. -------- = best estimate PCB load for year (8 g and 5 g, respectively).

Figure 4. Comparison of annual SS loads at Zone 4 Line A in 2007 and 2009 based on three designs for sampling among storms (Table 5). See also Figure 1 caption and text. -------- = best estimate SS load for year (0.10 10^-6 kg and 0.05 10^-6 kg, respectively).

Figure 5. Results of Monte Carlo simulations to determine the optimum number of samples required to estimate Hg loads at Guadalupe River using the turbidity surrogate regression method. -------- = load determined using all samples collected in each year.

Figure 6. Results of Monte Carlo simulations to determine the optimum number of samples required to estimate PCB loads at Guadalupe River using the turbidity surrogate regression method. -------- = load determined using all samples collected in each year.

Figure 7. Results of Monte Carlo simulations to determine the optimum number of samples required to estimate suspended sediment (SS) loads at Guadalupe River using the turbidity surrogate regression method. -------- = load determined using all samples collected in each year.

Figure 8. Comparison of annual Hg loads at Guadalupe River in 2003-2005 based on three designs for sampling among storms (Table 5) using turbidity surrogate regression. -------- = Hg load from all storms sampled each year.

Figure 9. Comparison of annual PCB loads at Guadalupe River in 2003-2005 based on three designs for sampling among storms (Table 5) using turbidity surrogate regression. -------- = PCB load from all storms sampled each year.
Figure 10. Comparison of annual SS loads at Guadalupe River in 2003-2005 for three designs for sampling among storms (Table 5). -------- = SS load from all storms sampled each year.

Figure 11. Results of Monte Carlo simulations to determine the optimum number of samples required to estimate Hg loads at Zone 4 Line A using the turbidity surrogate regression method. -------- = load determined using all samples collected in each year.

Figure 12. Results of Monte Carlo simulations to determine the optimum number of samples required to estimate PCB loads at Zone 4 Line A using the turbidity surrogate regression method. -------- = load determined using all samples collected in each year.

Figure 13. Results of Monte Carlo simulations to determine the optimum number of samples required to estimate suspended sediment (SS) loads at Zone 4 Line A using turbidity surrogate regression. -------- = load determined using all samples collected in each year.

Figure 14. Comparison of annual Hg loads at Zone 4 Line A in 2007-2009 based on three designs for sampling among storms (Table 5) using turbidity surrogate regression. -------- = Hg load from all storms sampled each year.

Figure 15. Comparison of annual PCB loads at Zone 4 Line A in 2007-2009 based on three designs for sampling among storms (Table 5) using turbidity surrogate regression. -------- = PCB load from all storms sampled each year.

Figure 16. Comparison of annual SS loads at Zone 4 Line A in 2007-2009 based on three designs for sampling among storms (Table 5) using turbidity surrogate regression. -------- = SS load from all storms sampled each year.
Figures

a) 2003

![Figures 2003](image1)

b) 2004

![Figures 2004](image2)

c) 2005

![Figures 2005](image3)

Figure 1a-c
a) 2007

Figure 2a-b

b) 2009
DRAFT – FINAL

a) 2007

b) 2009

Figure 3a-b
a) 2007

b) 2009

Figure 4a-b
Figure 5
Figure 6
Figure 7
a) 2003

b) 2004

c) 2005

Figure 8a-c
Figure 9a-c
Figure 10a-c
Figure 11
Figure 12
Figure 13
a) 2007

b) 2009

Figure 14a-b
a) 2007

![Graphs showing PCBs levels for Design A, B, and C in 2007.]

b) 2009

![Graphs showing PCBs levels for Design A, B, and C in 2009.]

Figure 15a-b
a) 2007

![Graphs showing data for 2007 across different designs.](image)

b) 2009

![Graphs showing data for 2009 across different designs.](image)

Figure 16a-b
Appendix A – SSC: Hg relationships at Zone 4 Line A

The mean slopes of SSC: Hg at Z4LA were below the target slopes for the trend analysis of 0.2 mg Hg / kg suspended sediment in two of the three years. Due to this situation, the Hg trend analysis examined power for trends assuming a target of 0.05 mg Hg / kg suspended sediment. The revised target was 75% below 0.2 mg/kg and was selected to ensure the trend could be examined for all years.

Figure A.1. SSC: Hg relationships at Zone 4 Line A. The mean slope estimate for 2007 – 2009 were 0.13, 0.38, and 0.08, respectively. The mean slope of all three years was 0.19.
Appendix B. Literature review of mercury variation by McKee et al. (2004)
<table>
<thead>
<tr>
<th>River (Source, see McKee et al. 2004)</th>
<th>Description</th>
<th>Min THg</th>
<th>Max THg</th>
<th>Max/min</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 Rappahannock R., Chesapeake (Lawson et al., 2001)</td>
<td>Mostly undisturbed, 12% Agriculture (5)</td>
<td>10.3</td>
<td>24.9</td>
<td>2.4</td>
</tr>
<tr>
<td>2 Site B1 Sudbury R., Massachusetts (Waldron et al., 2000)</td>
<td>Undisturbed rural (2)</td>
<td>1.9</td>
<td>5.4</td>
<td>2.8</td>
</tr>
<tr>
<td>1 Site B2 Sudbury R., Massachusetts (Waldron et al., 2000)</td>
<td>Undisturbed rural (1)</td>
<td>0.99</td>
<td>3.6</td>
<td>3.6</td>
</tr>
<tr>
<td>6 Choptank R., Chesapeake (Lawson et al., 2001)</td>
<td>83% Agriculture (6)</td>
<td>6.8</td>
<td>26.2</td>
<td>3.9</td>
</tr>
<tr>
<td>10 Anacostia R. NE. Branch (Mason and Sullivan, 1998)</td>
<td>Mostly Urban (10)</td>
<td>8.72</td>
<td>39.5</td>
<td>4.5</td>
</tr>
<tr>
<td>9 Susquehanna R., Chesapeake (Lawson et al., 2001)</td>
<td>61% Agriculture (9)</td>
<td>7</td>
<td>32.8</td>
<td>4.7</td>
</tr>
<tr>
<td>11 Herring Run R., Chesapeake (Lawson et al., 2001)</td>
<td>100% urban (11)</td>
<td>12.2</td>
<td>62.8</td>
<td>5.1</td>
</tr>
<tr>
<td>7 At Freeport, Sacramento Basin (Domagalski and Dileanis, 2000; Roth et al., 2001)</td>
<td>2% urban, 35% Agriculture (7)</td>
<td>4.2</td>
<td>29</td>
<td>6.9</td>
</tr>
<tr>
<td>8 Anacostia R. NW. Branch (Mason and Sullivan, 1998)</td>
<td>Mostly Urban (8)</td>
<td>4.45</td>
<td>30.8</td>
<td>6.9</td>
</tr>
<tr>
<td>3 Below Keswick Dam, Sacramento Basin (Domagalski and Dileanis, 2000; Roth et al., 2001)</td>
<td>Dam (3)</td>
<td>1.1</td>
<td>7.9</td>
<td>7.2</td>
</tr>
<tr>
<td>14 Potomac R., Chesapeake (Lawson et al., 2001)</td>
<td>28% urban, 20% Agriculture (14)</td>
<td>12.1</td>
<td>93.1</td>
<td>7.7</td>
</tr>
<tr>
<td>12 At Colusa, Sacramento Basin (Domagalski and Dileanis, 2000; Roth et al., 2001)</td>
<td>1% urban, 31% Agriculture (12)</td>
<td>6.5</td>
<td>81</td>
<td>12</td>
</tr>
<tr>
<td>4 Above Bend Bridge, Sacramento Basin (Domagalski and Dileanis, 2000; Roth et al., 2001)</td>
<td>1% urban, 27% Agriculture (4)</td>
<td>1.2</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td>13 Site M1 Sudbury R., Massachusetts (Waldron et al., 2000)</td>
<td>Nyanza chemical dump super fund (13)</td>
<td>5.2</td>
<td>92</td>
<td>18</td>
</tr>
<tr>
<td>15 Guadalupe R., Bay Area (Leatherbarrow et al., 2002)</td>
<td>70% urban, New Almaden mining district (15)</td>
<td>18</td>
<td>730</td>
<td>41</td>
</tr>
<tr>
<td>16 Kuskakwim R. Basin, SW Alaska (Gray et al., 2000)</td>
<td>Abandoned mercury mines (17)</td>
<td>10</td>
<td>2500</td>
<td>250</td>
</tr>
<tr>
<td>17 Walker Ck. Marin County, California (Whyte and Kirchner, 2000)</td>
<td>Gambonini mine (18)</td>
<td>485</td>
<td>104000</td>
<td>2144</td>
</tr>
<tr>
<td>15 Cache Ck., Sacramento Basin (Domagalski and Dileanis, 2000)</td>
<td>Cache Ck. Mining (16)</td>
<td>1</td>
<td>2250</td>
<td>2250</td>
</tr>
</tbody>
</table>
Appendix C – Within-storm Sampling Designs

Results are shown here for evaluation of within-storm sampling designs using flow and turbidity-based selection criteria. Bias here refers to the best estimate of loads per storm. Both flow and turbidity-based sampling criteria identified similar levels of accuracy (median load bias) and precision (standard error in load bias) in estimation of loads for the three pollutants (Hg, PCBs, suspended sediment).
### C.1. Guadalupe River

Table C.1a. Summary of within-storm Hg loads (g) at Guadalupe River determined using three mass emission estimators. Flow-based storm selection criteria.

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Simple Mean</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>177</td>
<td>-50%</td>
<td>6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>330</td>
<td>-21%</td>
<td>3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>402</td>
<td>-14%</td>
<td>3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>393</td>
<td>-10%</td>
<td>3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>177</td>
<td>-50%</td>
<td>5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>304</td>
<td>-25%</td>
<td>3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>365</td>
<td>-15%</td>
<td>3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>419</td>
<td>-10%</td>
<td>3%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Linear Interpolation</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>205</td>
<td>-40%</td>
<td>7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>411</td>
<td>-5%</td>
<td>5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>510</td>
<td>0%</td>
<td>4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>429</td>
<td>3%</td>
<td>3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>205</td>
<td>-40%</td>
<td>7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>454</td>
<td>-3%</td>
<td>6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>460</td>
<td>3%</td>
<td>4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>520</td>
<td>2%</td>
<td>4%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Flow-weighted Mean</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>183</td>
<td>-47%</td>
<td>7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>510</td>
<td>-2%</td>
<td>6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>623</td>
<td>4%</td>
<td>5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>533</td>
<td>10%</td>
<td>5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>183</td>
<td>-45%</td>
<td>7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>518</td>
<td>0%</td>
<td>7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>479</td>
<td>9%</td>
<td>6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>598</td>
<td>9%</td>
<td>6%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table C.1b. Summary of within-storm Hg loads (g) at Guadalupe River determined using three mass emission estimators. Turbidity-based storm selection criteria.

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Mean</td>
<td></td>
<td>285</td>
<td>149</td>
<td>-41%</td>
<td>4%</td>
</tr>
<tr>
<td>Even</td>
<td>6</td>
<td>149</td>
<td>241</td>
<td>-14%</td>
<td>3%</td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>251</td>
<td>272</td>
<td>-8%</td>
<td>-1%</td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>272</td>
<td>149</td>
<td>-41%</td>
<td>4%</td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>289</td>
<td>237</td>
<td>-12%</td>
<td>4%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>274</td>
<td>277</td>
<td>-4%</td>
<td>4%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>277</td>
<td>274</td>
<td>4%</td>
<td>3%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>277</td>
<td>274</td>
<td>4%</td>
<td>3%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>278</td>
<td>274</td>
<td>4%</td>
<td>3%</td>
</tr>
<tr>
<td>Linear Interpolation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>6</td>
<td>175</td>
<td>275</td>
<td>-28%</td>
<td>6%</td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>284</td>
<td>277</td>
<td>0%</td>
<td>2%</td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>277</td>
<td>274</td>
<td>4%</td>
<td>2%</td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>277</td>
<td>274</td>
<td>4%</td>
<td>2%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>175</td>
<td>270</td>
<td>3%</td>
<td>2%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>270</td>
<td>274</td>
<td>4%</td>
<td>3%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>270</td>
<td>274</td>
<td>4%</td>
<td>3%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>270</td>
<td>274</td>
<td>4%</td>
<td>3%</td>
</tr>
<tr>
<td>Flow-weighted Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>6</td>
<td>177</td>
<td>178</td>
<td>-30%</td>
<td>6%</td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>321</td>
<td>321</td>
<td>12%</td>
<td>3%</td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>331</td>
<td>331</td>
<td>9%</td>
<td>3%</td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>331</td>
<td>331</td>
<td>9%</td>
<td>3%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>178</td>
<td>307</td>
<td>13%</td>
<td>7%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>307</td>
<td>307</td>
<td>13%</td>
<td>5%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>307</td>
<td>307</td>
<td>13%</td>
<td>5%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>307</td>
<td>307</td>
<td>13%</td>
<td>5%</td>
</tr>
</tbody>
</table>
Table C.1c. Summary of within-storm PCB loads (g) at Guadalupe River determined using three mass emission estimators. Flow-based storm selection criteria.

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>37</td>
<td>-18%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>33</td>
<td>-4%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>37</td>
<td>-15%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>42</td>
<td>-1%</td>
<td>6%</td>
<td></td>
</tr>
</tbody>
</table>

**Simple Mean**

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>40</td>
<td>-9%</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>34</td>
<td>2%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>43</td>
<td>-1%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>52</td>
<td>5%</td>
<td>6%</td>
<td></td>
</tr>
</tbody>
</table>

**Linear Interpolation**

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>38</td>
<td>-13%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>33</td>
<td>15%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>39</td>
<td>1%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>53</td>
<td>14%</td>
<td>7%</td>
<td></td>
</tr>
</tbody>
</table>

**Flow-weighted Mean**
Table C.1d. Summary of within-storm PCB loads (g) at Guadalupe River determined using three mass emission estimators. Turbidity-based storm selection criteria.

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Even 6</td>
<td>9.1</td>
<td>-10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Even 12</td>
<td>11.3</td>
<td>-2%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td></td>
<td>9.1</td>
<td>-8%</td>
<td>5%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td></td>
<td>11.5</td>
<td>3%</td>
<td>6%</td>
</tr>
</tbody>
</table>

**Simple Mean**

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Even 6</td>
<td>9.3</td>
<td>-3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Even 12</td>
<td>11.3</td>
<td>-1%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td></td>
<td>9.1</td>
<td>-1%</td>
<td>7%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td></td>
<td>13.5</td>
<td>12%</td>
<td>6%</td>
</tr>
</tbody>
</table>

**Linear Interpolation**

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Even 6</td>
<td>9.1</td>
<td>-3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Even 12</td>
<td>13.4</td>
<td>10%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td></td>
<td>9.1</td>
<td>-1%</td>
<td>8%</td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td></td>
<td>14.9</td>
<td>19%</td>
<td>8%</td>
</tr>
</tbody>
</table>

**Flow-weighted Mean**

---

50
Table C.1e. Summary of within-storm suspended sediment loads (kg) at Guadalupe River determined using three mass emission estimators. Flow-based storm selection criteria.

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>111</td>
<td>-50%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>199</td>
<td>-20%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>197</td>
<td>-10%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>219</td>
<td>-10%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>111</td>
<td>-46%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>182</td>
<td>-17%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>218</td>
<td>-8%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>219</td>
<td>-1%</td>
<td>4%</td>
<td></td>
</tr>
</tbody>
</table>

Simple Mean

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>124</td>
<td>-40%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>226</td>
<td>-1%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>254</td>
<td>2%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>245</td>
<td>2%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>124</td>
<td>-34%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>244</td>
<td>3%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>258</td>
<td>6%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>278</td>
<td>6%</td>
<td>4%</td>
<td></td>
</tr>
</tbody>
</table>

Linear Interpolation

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>114</td>
<td>-47%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>270</td>
<td>7%</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>286</td>
<td>16%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>280</td>
<td>19%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>114</td>
<td>-38%</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>282</td>
<td>17%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>281</td>
<td>22%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>316</td>
<td>20%</td>
<td>5%</td>
<td></td>
</tr>
</tbody>
</table>

Flow-weighted Mean

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>114</td>
<td>-47%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>270</td>
<td>7%</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>286</td>
<td>16%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>280</td>
<td>19%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>114</td>
<td>-38%</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>282</td>
<td>17%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>281</td>
<td>22%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>316</td>
<td>20%</td>
<td>5%</td>
<td></td>
</tr>
</tbody>
</table>
Table C.1f. Summary of within-storm suspended sediment loads (kg) at Guadalupe River determined using three mass emission estimators. Turbidity-based storm selection criteria.

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>47</td>
<td>-37%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>57</td>
<td>-19%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>63</td>
<td>-2%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>64</td>
<td>-3%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>47</td>
<td>-37%</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>68</td>
<td>-4%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>68</td>
<td>3%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>72</td>
<td>4%</td>
<td>4%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>50</td>
<td>-24%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>68</td>
<td>-2%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>72</td>
<td>2%</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>68</td>
<td>1%</td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>50</td>
<td>-20%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>78</td>
<td>18%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>76</td>
<td>9%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>77</td>
<td>5%</td>
<td>2%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>48</td>
<td>-28%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>74</td>
<td>13%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>78</td>
<td>22%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>80</td>
<td>19%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>48</td>
<td>-19%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>86</td>
<td>31%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>86</td>
<td>30%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>89</td>
<td>27%</td>
<td>3%</td>
<td></td>
</tr>
</tbody>
</table>
**C.2. Zone 4 Line A**

Table C.2a. Summary of within-storm Hg loads (mg) at Zone 4 Line A determined using three mass emission estimators. Flow-based storm selection criteria.

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>142</td>
<td>-46%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>180</td>
<td>-26%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>248</td>
<td>-14%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>258</td>
<td>-8%</td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>143</td>
<td>-41%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>275</td>
<td>-15%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>272</td>
<td>-5%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>295</td>
<td>1%</td>
<td>3%</td>
<td></td>
</tr>
</tbody>
</table>

**Linear Interpolation**

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>144</td>
<td>-44%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>203</td>
<td>-11%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>279</td>
<td>1%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>278</td>
<td>3%</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>144</td>
<td>-38%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>351</td>
<td>13%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>311</td>
<td>9%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>310</td>
<td>7%</td>
<td>2%</td>
<td></td>
</tr>
</tbody>
</table>

**Flow-weighted Mean**

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>146</td>
<td>-44%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>225</td>
<td>-9%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>351</td>
<td>17%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>358</td>
<td>17%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>153</td>
<td>-38%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>401</td>
<td>28%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>395</td>
<td>28%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>388</td>
<td>28%</td>
<td>4%</td>
<td></td>
</tr>
</tbody>
</table>
Table C.2b. Summary of within-storm Hg loads (mg) at Zone 4 Line A determined using three mass emission estimators. Turbidity-based storm selection criteria.

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Simple Mean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>6</td>
<td>93</td>
<td>-41%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>113</td>
<td>-12%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>138</td>
<td>-5%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>136</td>
<td>-1%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>99</td>
<td>-34%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>127</td>
<td>-1%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>147</td>
<td>5%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>152</td>
<td>8%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td><strong>Linear Interpolation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>6</td>
<td>108</td>
<td>-30%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>143</td>
<td>3%</td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>166</td>
<td>3%</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>158</td>
<td>3%</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>109</td>
<td>-22%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>184</td>
<td>22%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>177</td>
<td>12%</td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>166</td>
<td>8%</td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td><strong>Flow-weighted Mean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>6</td>
<td>95</td>
<td>-29%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>157</td>
<td>13%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>192</td>
<td>17%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>187</td>
<td>18%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>109</td>
<td>-22%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>202</td>
<td>34%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>191</td>
<td>33%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>196</td>
<td>31%</td>
<td>4%</td>
<td></td>
</tr>
</tbody>
</table>
Table C.2c. Summary of within-storm PCB loads (mg) at Zone 4 Line A determined using three mass emission estimators. Flow-based storm selection criteria.

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>59 -47% 3%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>89 -26% 3%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>61 -42% 3%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>119 -15% 4%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>62 -44% 4%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>104 -11% 4%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>65 -40% 5%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>145 13% 4%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>61 -44% 4%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>113 -9% 6%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>63 -38% 6%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>180 27% 7%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table C.2d. Summary of within-storm PCB loads (mg) at Zone 4 Line A determined using three mass emission estimators. Turbidity-based storm selection criteria.

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>6</td>
<td>39</td>
<td>-45%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>54</td>
<td>-14%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>44</td>
<td>-40%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>58</td>
<td>-5%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Linear Interpolation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>6</td>
<td>44</td>
<td>-35%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>63</td>
<td>3%</td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>48</td>
<td>-22%</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>93</td>
<td>18%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Flow-weighted Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>6</td>
<td>47</td>
<td>-35%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>68</td>
<td>14%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>47</td>
<td>-22%</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>90</td>
<td>34%</td>
<td>6%</td>
<td></td>
</tr>
</tbody>
</table>
Table C.2e. Summary of within-storm suspended sediment loads (g) at Zone 4 Line A determined using three mass emission estimators. Flow-based storm selection criteria.

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>536</td>
<td>-54%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>705</td>
<td>-29%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>1055</td>
<td>-17%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>1171</td>
<td>-7%</td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>536</td>
<td>-48%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>1105</td>
<td>-15%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>1163</td>
<td>-3%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>1407</td>
<td>3%</td>
<td>3%</td>
<td></td>
</tr>
</tbody>
</table>

**Simple Mean**

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>549</td>
<td>-51%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>947</td>
<td>-11%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>1150</td>
<td>0%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>1272</td>
<td>3%</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>585</td>
<td>-46%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>1381</td>
<td>15%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>1430</td>
<td>11%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>1443</td>
<td>8%</td>
<td>3%</td>
<td></td>
</tr>
</tbody>
</table>

**Linear Interpolation**

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Sample Size</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>6</td>
<td>551</td>
<td>-51%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>12</td>
<td>1185</td>
<td>-10%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>18</td>
<td>1337</td>
<td>15%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Even</td>
<td>24</td>
<td>1592</td>
<td>18%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>6</td>
<td>565</td>
<td>-43%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>12</td>
<td>1739</td>
<td>33%</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>18</td>
<td>1594</td>
<td>35%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage</td>
<td>24</td>
<td>1836</td>
<td>32%</td>
<td>5%</td>
<td></td>
</tr>
</tbody>
</table>

**Flow-weighted Mean**
Table C.2f. Summary of within-storm suspended sediment loads (g) at Zone 4 Line A determined using three mass emission estimators. Turbidity-based storm selection criteria.

<table>
<thead>
<tr>
<th>Simple Mean</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even 6</td>
<td>328</td>
<td>-43%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even 12</td>
<td>590</td>
<td>-11%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Even 18</td>
<td>643</td>
<td>-6%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Even 24</td>
<td>662</td>
<td>2%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage 6</td>
<td>342</td>
<td>-38%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage 12</td>
<td>673</td>
<td>-1%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage 18</td>
<td>791</td>
<td>8%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage 24</td>
<td>741</td>
<td>10%</td>
<td>7%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Linear Interpolation</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even 6</td>
<td>351</td>
<td>-35%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even 12</td>
<td>682</td>
<td>4%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Even 18</td>
<td>653</td>
<td>4%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Even 24</td>
<td>640</td>
<td>4%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage 6</td>
<td>351</td>
<td>-29%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage 12</td>
<td>956</td>
<td>19%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage 18</td>
<td>745</td>
<td>15%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage 24</td>
<td>655</td>
<td>9%</td>
<td>3%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Flow-weighted Mean</th>
<th>Median &quot;Best Estimate&quot; Load</th>
<th>Median Load</th>
<th>Median Bias</th>
<th>St. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even 6</td>
<td>353</td>
<td>-34%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Even 12</td>
<td>709</td>
<td>13%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Even 18</td>
<td>768</td>
<td>22%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Even 24</td>
<td>820</td>
<td>21%</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage 6</td>
<td>353</td>
<td>-27%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage 12</td>
<td>918</td>
<td>37%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage 18</td>
<td>890</td>
<td>36%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Rising Stage 24</td>
<td>811</td>
<td>34%</td>
<td>7%</td>
<td></td>
</tr>
</tbody>
</table>
Appendix D. Among-storm sampling designs (turbidity-based sampling)

Results are presented here for among-storm sampling designs with turbidity-based storm sampling. These results were generally more variable than the flow-based results presented in the main text of the report.

Figure Captions

Figure D.1a-c. Comparison of annual Hg loads at Guadalupe River in 2003-2005 based on three designs for sampling among storms (Table 5). Loads were calculated using linear interpolation with turbidity-based storm selection criteria. Design A simulated sampling of the first flush only and a variable number of random storms. Design B simulated the first flush plus one of the three largest storms and a variable number of random storms. Design C only tested the random storm component. --------- = best estimate Hg load for year

Figure D.2a-c. Comparison of annual PCB loads at Guadalupe River in 2003-2005 based on three designs for sampling among storms (Table 5). See Figure D.1 caption and text for further information. --------- = best estimate PCB load for year.

Figure D.3a-c. Comparison of annual SS loads at Guadalupe River in 2003-2005 based on three designs for sampling among storms (Table 5). See Figure D.1 caption and text for further information. --------- = best estimate SS load for year.

Figure D.4a-c. Comparison of annual Hg loads at Zone 4 Line A in 2007 and 2009 based on three designs for sampling among storms (Table 5). See Figure D.1 caption and text for further information. --------- = best estimate Hg load for year.

Figure D.5a-c. Comparison of annual PCB loads at Zone 4 Line A in 2007 and 2009 based on three designs for sampling among storms (Table 5). See Figure D.1 caption and text for further information. --------- = best estimate PCB load for year.

Figure D.6a-c. Comparison of annual SS loads at Zone 4 Line A in 2007 and 2009 based on three designs for sampling among storms (Table 5). See Figure D.1 caption and text for further information. --------- = best estimate SS load for year.
a) 2003

![Graphs for 2003](image)

b) 2004

![Graphs for 2004](image)

c) 2005

![Graphs for 2005](image)

Figure D.1a-c
a) 2003

b) 2004

c) 2005

Figure D.2a-c
a) 2003

b) 2004

c) 2005

Figure D.3a-c
a) 2007

Figure D.4a-b

b) 2009

Figure D.4a-b
a) 2007

b) 2009

Figure D.5a-b
Figure D.6a-b

a) 2007

b) 2009
Appendix E. Comparison of annual PCB and SS loads at Guadalupe River in 2003 – 2005 based on three designs for sampling among storms.

Results for estimation of annual PCB and SS loads at Guadalupe River mirrored that of Hg loads. Accuracy and precision were optimal at 6 or 10 storms samples per water year. Designs A and B performed the best and similarly well, with Design C exhibiting good accuracy, but poor precision.

Figure Captions

Figure E.1a-c. Comparison of annual PCB loads at Guadalupe River in 2003-2005 based on three designs for sampling among storms (Table 5). Loads were calculated using linear interpolation with flow-based storm selection criteria. In addition to the random component, Designs A and B simulated sampling of the first flush (A) and first flush plus one of the three largest storms (B). Design C only tested the random storm component. -------- = best estimate PCB load for year (0.9 kg, 0.5 kg, and 0.5 kg, respectively).

Figure E.2a-c. Comparison of annual SS loads at Guadalupe River in 2003-2005 based on three designs for sampling among storms (Table 5). Loads were calculated using linear interpolation with flow-based storm selection criteria. In addition to the random component, Designs A and B simulated sampling of the first flush (A) and first flush plus one of the three largest storms (B). Design C only tested the random storm component. -------- = best estimate SS load for year (10 x 10^6 kg, 8 x 10^6 kg, and 4 x 10^6 kg, respectively)
a) 2003

b) 2004

c) 2005

Figure 1a-c.
Figure 2a-c.