ELSEVIER

The Science of the Total Environment 260 (2000) 1–9

the Science of the Total Environment
An International Journal for Scientific Research into the Environment and its Relationship with Man

www.elsevier.com/locate/scitotenv

Assessing water quality impacts and cleanup effectiveness in streams dominated by episodic mercury discharges

Dyan C. Whyte[a,*], James W. Kirchner[b]

[a] *California Regional Water Quality Control Board, 1515 Clay St., Suite 1400, Oakland, CA 94612, USA*
[b] *Department of Geology and Geophysics, University of California, Berkeley, CA 94720-4767, USA*

Received 27 September 1999; accepted 4 February 2000

Abstract

Accurate pollutant mass budgets are needed for identifying contaminant sources and establishing cleanup goals. We monitored mercury discharges from an abandoned mine site in northern California with the objectives of: (1) estimating the mass loading of mercury from the site; (2) evaluating the factors that control the mercury discharges; (3) assessing the significance of peak flows in transporting contaminants; and (4) developing methods for measuring the effectiveness of cleanup efforts. We sampled water downstream from the mine site over a wide range of streamflows. Mercury concentrations varied over 2000-fold, from 485 to 1 040 000 ng/l, grossly exceeding the regulatory water quality objective of 12 ng/l at all times. Particulate mercury represented over 99.97% of the total mercury, and mercury concentrations were closely correlated to suspended sediment concentrations ($r = 0.98$). Thus, we can use suspended sediment concentrations as a proxy for mercury concentrations, and calculate a continuous record of mercury flux from continuous monitoring of streamflow (using a small flume) and turbidity (using an optical backscatter sensor). Mercury fluxes inferred in this way are consistent with fluxes estimated from field samples. In January and February of 1998, our small abandoned mine site released approximately 82 kg of mercury to downstream waters. Most of the mercury was released during brief intense rainstorms. For example, in one 200-min period we recorded 3.4 cm of rain, a 2.6-fold increase in streamflow (460–1120 l/s), and an 82-fold increase in mercury flux (1.2–99 g/min). Over 75% of the total mercury flux during this 2-month period occurred in less than 10% of the total time. In systems such as this one, where contaminant transport is highly episodic, sampling programs that miss the high-flow episodes may greatly underestimate the actual water quality threat. In addition, changes in pollutant fluxes or concentrations in receiving waters may not reflect changes in pollutant sources (such as an environmental cleanup) if the stochastic forcing (e.g. intense rainstorms) varies through time. We propose that water

* Corresponding author. Tel.: +1-510-622-2441; fax: +1-510-622-2460.
E-mail address: dcw@rb2.swrcb.ca.gov (D.C. Whyte).

PII: S0048-9697(00)00537-4

quality trends can be more accurately measured by changes in the relationship between contaminant flux and stochastic driving factors, as expressed by contaminant rating curves.

Keywords: Mercury; Contaminant flux; Episodic discharge; Mine remediation; Rating curve

1. Introduction

Mercury mines, and gold mines that used mercury as amalgamate, are widespread sources of mercury contamination in lakes and estuaries (Lacerda and Salomons, 1998). For persistent bioaccumulative pollutants such as mercury, estimates of pollutant mass loadings are needed for identifying and quantifying sources of contamination, and evaluating the threat that they pose.

Pollutant loads are typically estimated by multiplying contaminant concentrations in water by discharge, and integrating over time. The accuracy of this method depends on how the contaminant flux varies between the moments when the samples are taken. In streams in which contaminant concentrations and/or discharge are highly variable, fluxes inferred in this manner may be misleading. For example, Mason and Sullivan (1998) documented a three to fivefold increase in mercury concentrations in the Anacostia River during storm events, and concluded that mercury fluxes could be significantly underestimated by not sampling during storms.

Thomas and Lewis (1995) recommend using a time-stratified sampling regime to estimate suspended sediment loads, in which the sampling frequency is varied according to discharge (see also Balogh et al., 1998). In small episodic streams where fluctuations in discharge are rapid and unpredictable, high-frequency sampling during storms may be difficult. This is often the case at mines located in steep, remote mountainous areas. Here we show that, where mercury fluxes are dominated by particulate mercury, loadings can be accurately estimated from continuous records of stream discharge and suspended sediment concentrations. Relatively few mercury concentrations need to be measured, and thus sampling and analysis costs are minimized. Our results show that mercury fluxes in small streams can be extremely episodic, and thus mass loadings derived by conventional methods can vary greatly, depending on the particular sampling times that are chosen.

2. Site history and description

The Gambonini mine, located approximately 50 miles north of San Francisco in the steep headwaters of the Tomales Bay drainage, is one of 51 major mercury deposits located within the California Coast Range mercury mineral belt (Rytuba and Enderlin, 1999). The ore body was a high-grade, localized cinnabar deposit that produced approximately 5000 flasks of mercury from 1964 to 1970 using open pit mining techniques. The environmental conditions at this mine illustrate the consequences of poor mining waste disposal practices. Over 300 000 m^3 of mine wastes (overburden and calcines), with an average mercury concentration of 320 mg/kg, were dumped on the slope of an adjacent ravine. In an attempt to protect water quality, the mining company dammed the ravine roughly 400 m downstream from the mining waste pile. In 1982 the dam failed, inundating the channel with mine waste and creating an alluvial fan of poorly sorted mercury-laden sediments downstream from the former impoundment. By 1990, the steep ephemeral creek draining the ravine was incising through the toe of the waste pile and causing mass failure, as revealed by a 5-m landslide scarp in the top surface of the waste pile. The waste pile also showed abundant evidence of surface erosion, including numerous rills, large gullies (some over 3 m deep), and debris flow scars. Our primary objective was to quantify the environmental threat posed by the ongoing erosion of the waste pile and alluvial fan.

3. Water column sampling and analysis

We measured discharge, turbidity, suspended solids concentrations, and mercury concentrations in the stream draining the mine site, at a point roughly 800 m downstream from the waste pile, near the base of the alluvial fan comprised of mine wastes. We sampled the stream 18 times during the 1998 rainy season, aiming to characterize a large range of flows by sampling before, during, and after storm events. Frontier Geoscience performed all our analytical laboratory work. Stream samples were taken in triplicate and composited to minimize the effects of short-term fluctuations in concentrations. Samples were immediately placed on ice, and within 24 h, samples were composited, filtered, and oxidized with BrCl in the lab. Stream samples were analyzed for total, dissolved (filter size 0.45 μm), and particulate mercury concentrations using $SnCl_2$ reduction, dual gold amalgamation, and cold vapor atomic fluorescence detection. Nine matrix spikes were analyzed, and the percent recovery results ranged from 96.8 to 102.1%. The percent recovery for certified standard references was 99.4, 101.1, and 114.1%. The minimum detection limits for total mercury, reported as three times the standard deviation of three sets of procedural blanks, were 0.06, 0.09, and 0.39 ng/l, more than three orders of magnitude below sample analytical results.

4. Water column monitoring results

Total mercury concentrations ranged from 485 to 1 040 000 ng/l, varying by more than three orders of magnitude and grossly exceeding the water quality objective of 12 ng/l (California Regional Water Quality Control Board, 1986) at all times. Despite the large fluctuations in total mercury concentrations, dissolved concentrations remained relatively constant, ranging from 31 to 139 ng/l. Dissolved concentrations had little effect on total mercury concentrations, as essentially all of the mercury (92.78–99.99%) was in the solid phase. There are strong positive correlations between total mercury and total suspended solids ($r = 0.98$, Fig. 1a), total mercury and discharge ($r = 0.94$), and total suspended solids and

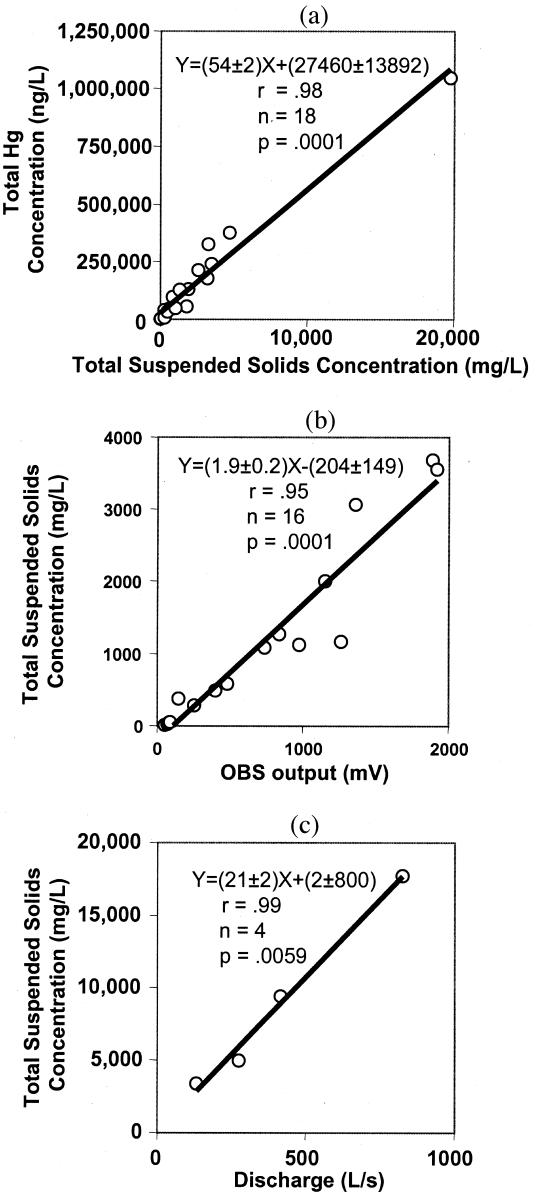

Fig. 1. Regressions used to calibrate monitoring system. (a) Mercury concentrations are inferred from total suspended solids concentrations. (b) During low flow periods, total suspended solids concentrations are inferred from turbidity readings (OBS output). (c) During high flow periods, total suspended solids concentrations are inferred from discharge.

discharge ($r = 0.87$). A single peak flow sample contained very high values for all three of these variables. When this sample is excluded, each pair of variables still exhibits significant positive correlation ($r = 0.94$, $r = 0.88$, and $r = 0.75$, respectively), confirming that the high flow sample is neither an outlier nor driving the correlation. Particulate mercury concentrations (ng Hg/g suspended sediment) were normally distributed and were not affected by changes in discharge, suspended sediment concentrations, or total mercury concentrations, suggesting that upstream mercury sources were unchanged throughout the monitoring period. These observations indicate that at this site, mercury and sediment transport processes are closely linked, and suspended solids can be used as a proxy for mercury.

5. Flux calculations

Here we detail our method for estimating mercury fluxes from turbidity and discharge measurements. Because total mercury and total suspended solids (TSS) concentrations are tightly correlated at our site (Fig. 1a), we can construct a continuous record of mercury concentrations from continuous TSS measurements. We estimated TSS concentrations from turbidity, which we measured using an optical backscatter sensor (OBS). The OBS is a small submersible nephelometric turbidimeter, which emits an infrared pulse of light that is scattered by particles in the water column. The OBS records the intensity of backscattered light and converts this to a voltage output, which is a relative measure of turbidity.

The relationship between OBS readings and TSS concentrations is strongly influenced by the size distribution and optical properties of the suspended particles, and by water color (Lewis, 1996; Buchanan and Schoellhamer, 1998), so site-specific calibration is required. We calibrated our OBS readings with TSS measurements obtained using an isokinetic depth-integrated sampler. Isokinetic samplers are designed so that water and suspended solids enter the collection nozzle at the same velocity they are traveling in the water column, thereby minimizing sampling bias caused by water accelerating around the nozzle (Martin et al., 1992). Isokinetic samplers also can be used to withdraw an integrated sample from an entire stream cross-section, or in our case, the area from which the OBS is recording backscatter measurements. We used the OBS calibration curve (Fig. 1b, $r = 0.95$, $n = 16$) to predict TSS concentrations from in-stream turbidity measurements. At this site the OBS was only capable of measuring turbidity when TSS concentrations were less than 3500 mg/l. During storm events (and thus periods of high turbidity and high flow), we measured TSS concentrations up to 17 700 mg/l. These high TSS concentrations are very closely correlated to discharge (Fig. 1c), so whenever turbidity levels in the stream exceeded the OBS's measuring range, we inferred total suspended solids concentrations from discharge measurements, and subsequently inferred mercury concentrations from TSS.

We constructed a flume to continuously record stream discharge. The flume was designed to accommodate a large range of flows and to pass bedload, both essential requirements for measuring discharge in flashy, steep, headwater streams like ours. In the flume, we recorded water levels in two stilling wells every 5 min. We used a form of the Bernoulli equation to estimate discharge from stage during low flows, when flow in the control section of the flume was critical. We developed a rating curve to estimate discharge from stage at higher flows.

6. Results and discussion

We constructed a 59-day record (January and February 1998) of mercury fluxes from streamflow and turbidity monitoring data, using site-specific field-based relationships between mercury concentrations, TSS concentrations, turbidity, and discharge. This method yielded the following equations (coefficients are reported as means $\pm$ 1 S.E.):

Low flow periods

$$\text{Hg flux (ng/s)} = \text{Discharge (l/s)} \times [(100 \pm 10) \times \text{Turbidity (mV)} + (16\,557 \pm 16\,040)]$$

High flow periods

$$\text{Hg flux (ng/s)} = \text{Discharge (l/s)} \times [(1145 \pm 106) \times \text{Discharge (l/s)} + (27\,591 \pm 45\,166)]$$

These calibration equations are necessarily specific to this site and to this sampling period. New field measurements and calibration equations would be required if site conditions changed (as a result of remediation efforts, for example).

The mercury fluxes we inferred using the relationships described above are generally consistent with field measured fluxes, although there is some tendency to over-predict low fluxes and under-predict high fluxes (Fig. 2). In both the low- and high-flow equations, the second term has large uncertainty, but contributes little to the total mercury flux estimate. Using first order, second moment error propagation techniques, we calculate that the total uncertainty associated with our 59-day cumulative mercury flux estimate is approximately 11%. The simple linear regressions used in our calibrations can predict some unrealistic values, such as negative TSS concentrations at very low flows. This problem can be avoided by using power functions for the regressions. We used linear regressions because they simplify the error propagation calculations; using power functions yields the same cumulative flux estimate as the linear regressions do, within standard error. Because extrapolating beyond the calibration data can be problematic, it is important to include peak flows in the calibration samples. In this study, the highest inferred mercury concentration was 1 390 000 ng/l, only 33% larger than the highest measured concentration in our calibration data (1 040 000 ng/l).

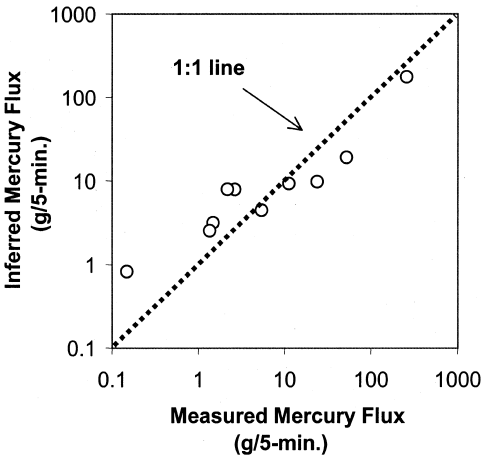

Fig. 2. Comparison of mercury fluxes from direct measurements, and mercury fluxes inferred from turbidity and discharge monitoring (1:1 line indicates perfect agreement).

6.1. Episodicity of contaminant flux

We recorded discharge and OBS readings every 5 min, yielding a practically continuous record of flow, TSS, and inferred mercury flux. The resulting time series (Fig. 3) illustrate the processes driving mercury transport, and highlight the extreme episodicity of flows and mercury fluxes in our small headwater stream. Plots a and b in Fig. 3 show that streamflow responded very quickly to precipitation, with a lag time of only 5–15 min between peak rainfall and channel response. In such a flashy stream, it is almost impossible to synchronize grab sampling with peak flows, making continuous monitoring essential for accurate flux estimates. The TSS time series (Fig. 3c) indicates that sediment loads also responded very quickly to changes in precipitation and streamflow. Intense storms transport mercury-laden sediment from the waste pile to the stream by surface erosion and mass wasting. High flows also increase the stream's capacity to transport sediment, as well as its ability to erode its bed and banks (which also contain mining waste). The highest mercury fluxes occurred during brief intense rainstorms, and mercury discharges were more episodic than streamflow (Fig. 3d). For example, during a single 200-min period we recorded 3.4 cm of rain, a 2.6-fold increase in streamflow

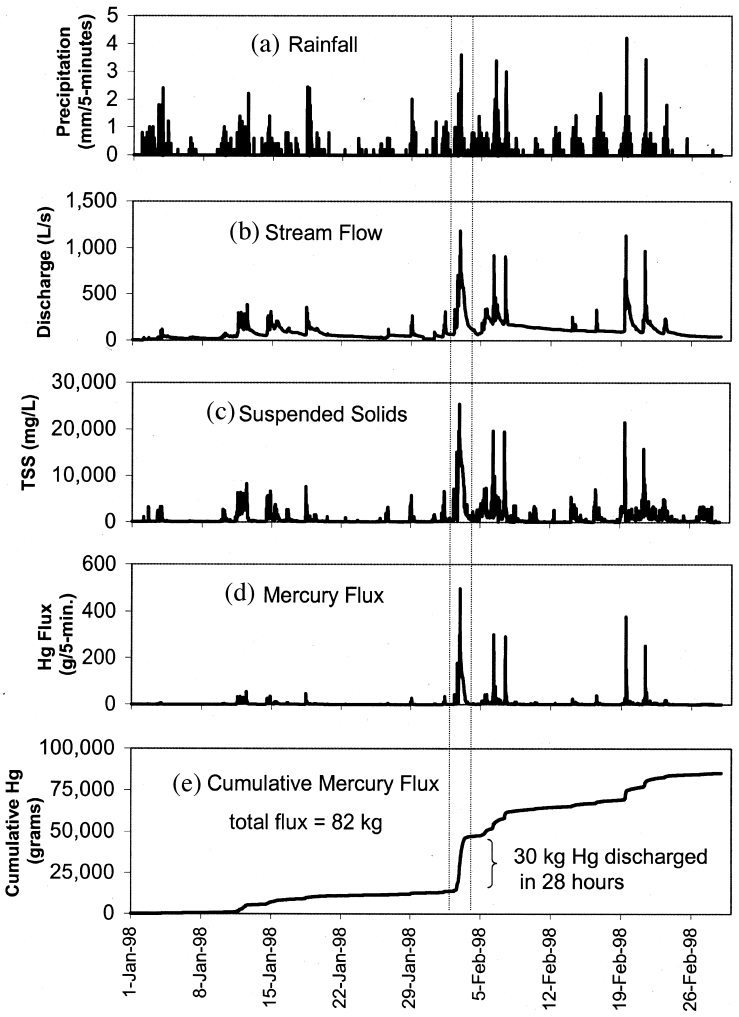

Fig. 3. Effects of rainfall intensity (a) on stream discharge (b), total suspended solids concentrations (c), mercury fluxes (d), and cumulative mercury flux (e) during January and February 1998. Precipitation was measured by a tipping bucket rain gauge, discharge was measured by a flume, total suspended solids concentrations were estimated from turbidity and discharge monitoring, mercury fluxes were inferred from total suspended solids concentrations and discharge (see text), and cumulative fluxes were determined by summing 5-min mercury loads.

(460–1200 l/s), and an 82-fold increase in mercury flux (1.2–99 g/min).

Fig. 3e shows the cumulative mercury flux through time, calculated by summing our continuous record of 5-min mercury loads. We estimate that the mine site released 82 kg (S.E. 9 kg) of mercury to downstream waters in this 2-month period. Our calculations indicate that the site released 1 300 000 kg of suspended sediment over the same period, with an average mercury concentration of 65 mg/kg.

The brief, intense storm events that account for most of the mercury flux comprise only a small fraction of the period of record (Fig. 3d and Fig. 4). For example, 75% of the total mercury flux over the 2-month period was released in less than 10% of the total time, or roughly 5 days (Fig. 4). A single 28-h storm discharged 30 kg of mercury,

or nearly 40% of the total (Fig. 3e). This highlights the need for continuous monitoring, because mercury fluxes could be significantly underestimated if one or two storm events are missed.

6.2. Implications for sampling

To illustrate how sample timing can affect flux estimates, we used our continuous 2-month record to simulate the fluxes that we would have inferred from weekly sampling on a fixed schedule (Fig. 5). These cumulative flux estimates varied up to 24-fold, depending on the day of the week sampled. For example, if we had sampled only on Wednesdays we would have concluded that the mercury load for our 2-month study period was only 14 kg, with a S.E. of only 4 kg. Thus, sampling on Wednesdays would lead to a well-constrained estimate, but one that understates the true flux by more than fivefold. By contrast, if we had sampled on Thursdays our estimate would have been 380 kg (S.E. = 277, $n = 9$), overstating the true flux by more than fourfold. Continuous monitoring avoids under- or over-sampling storm events, and thus under- or over-estimating the cumulative flux.

In summary, where pollutants are primarily released during episodes of high stream flow, sample timing is critical and flux estimates generated from intermittent sampling may misrepresent the true water quality threat. In such cases, accurately quantifying contaminant fluxes requires continuous monitoring. Where contaminant concentrations are strongly correlated to suspended solids concentrations, turbidity and discharge monitoring can yield robust contaminant flux estimates.

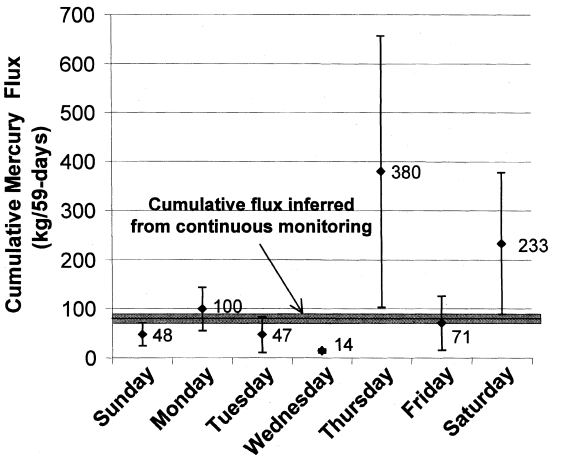

Fig. 5. Effects of fixed-schedule (once per week) sampling on flux estimates. Fluxes were calculated from instantaneous concentration and discharge at 12.00 h on each Mon., Tues., etc., and averaged for the 2 months of record. Error bars denote one standard error. Cumulative flux from entire continuous record (82 kg) is presented for comparison; gray band denotes 1 S.E. (9 kg).

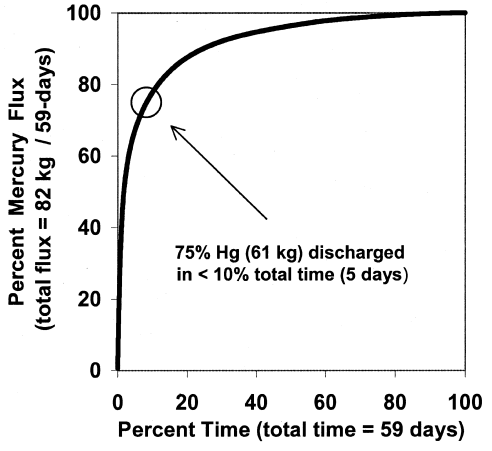

Fig. 4. Episodicity curve, expressing relationship between percentile of cumulative flux and percentile of period of record, obtained by sorting 5-min mercury fluxes in descending order.

6.3. Measuring cleanup effectiveness

Our continuous monitoring data demonstrate that a relatively small mine can deliver large quantities of mercury to downstream waters, if mercury-laden sediments are readily available for off-site fluvial transport. The water quality threat posed by mercury is a function of its bioavailability. Although the particulate mercury that dominates our measured fluxes is not directly bioavailable, our flux measurements were high enough to warrant an investigation of mercury transformation and bioaccumulation downstream from the mine. Our flux estimates, in conjunction with measurements documenting elevated mercury concentrations in waterfowl (Hoffman et al., 1998)

and clams (Whyte, 1998), prompted the US EPA and the California Regional Water Quality Control Board to implement Superfund emergency response cleanup measures at the mine (Whyte, 1998). Our monitoring program will be continued for several years to evaluate the success of the remediation effort.

The net environmental benefits of cleanup projects are typically assessed by comparing pre- and post-remediation water concentrations and contaminant fluxes. In episodic systems such as ours, comparing average concentrations or fluxes between pre- and post-remediation years may not accurately reflect the effectiveness of remediation, if the episodic driving factors (such as storm frequency or intensity) also differ between the two periods. We propose that cleanup effectiveness can be more accurately measured by changes in the contaminant rating curve, which expresses the relationship between contaminant fluxes and stochastic driving factors. The solid line in Fig. 6, fitted to our calibration samples, is the contaminant rating curve for baseline conditions at the mine site. The dashed line in Fig. 6 is a hypothetical post-remediation rating curve that illustrates an approximate 80% reduction in mercury fluxes. Because mercury concentrations can overlap between the two rating curves, post-remediation sampling may not accurately reflect the impact of remediation, depending on when samples are taken and on the frequency and intensity of storm events. For example, if high-flow conditions (and thus high mercury concentrations) were undersampled before remediation and over-sampled afterwards, average mercury concentrations for the two periods could be statistically indistinguishable. However, the respective contaminant rating curves for the two different periods would be much less affected by this kind of bias. A comparison between the pre- and post-remediation period rating curves would show a downward shift in concentrations for a given streamflow. Other shifts in a rating curve, such as a change in slope, may reflect changing contaminant transport processes at the site. Contaminant rating curves permit an 'all else equal' comparison of pre- and post-remediation conditions, by distinguishing changes in stochastic driving factors from changes in contaminant sources. In streams with highly episodic contaminant fluxes, they may be particularly useful.

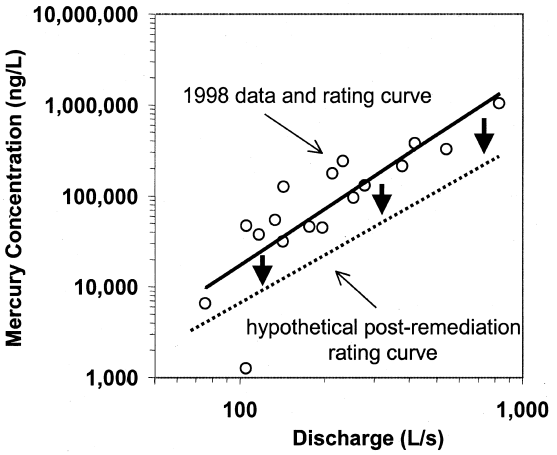

Fig. 6. Use of contaminant rating curves to measure cleanup effectiveness. The 1998 monitoring data and contaminant rating curve (solid line) reflect baseline conditions at the mine site. Hypothetical post-remediation rating curve (dashed line) corresponds to an 80% reduction in mercury fluxes. The downward shift between the two rating curves shows how much remediation hypothetically would reduce contaminant flux at any given discharge. It thus reflects cleanup effectiveness, free of confounding differences in stochastic forcing (frequency and intensity of storms) between the pre- and post-remediation periods.

Acknowledgements

Special thanks to Jill Marshall for her assistance in the field, especially during storms, and John Replogle for designing and helping install the flume (twice!). We also wish to thank the Regional Board staff who assisted with the flume installation, Ron Gervason for his support, and Mathias Kolehmainen for his perl programming. This work was partially supported by the California Water Quality Control Board cleanup and abatement fund, and by NSF grant EAR-9357931 to J.W.K.

References

Balogh SJ, Meyer ML, Johnson DK. Transport of mercury in three contrasting river basins. Environ Sci Technol 1998; 32(4):456–462.

Buchanan PA, Schoellhamer DH. Summary of suspended-solids concentration data, San Francisco Bay, California, Water Year 1996. USGS Open-File Report 98-175, 1998:59.

California Regional Water Quality Control Board. Water Quality Control Plan, San Francisco Bay Basin Region (2). Oakland, California, 1986.

Hoffman DJ, Ohlendorf HM, Marn CM, Pendleton GW. Association of mercury and selenium with altered glutathione metabolism and oxidative stress in diving ducks from San Francisco Bay Region, USA. Environ Toxicol Chem 1998;17(2):167–172.

Lacerda LD, Salomons W. Mercury from gold and silver mining: a chemical time bomb? Germany: Springer-Verlag, 1998:141.

Lewis J. Turbidity-controlled suspended sediment sampling for runoff-event load estimation. Water Resourc Res 1996;32(7):2299–2310.

Martin GR, Smoot JL, White KD. A comparison of surface-grab and cross sectionally integrated stream water quality sampling methods. Water Environ Res 1992;64:866–876.

Mason RP, Sullivan KA. Mercury and methylmercury transport through an urban watershed. Water Res 1998;31:321–330.

Rytuba JJ, Enderlin DA. Geology and environmental geochemistry of mercury and gold deposits in the northern part of the California Coast Range mercury belt. In: Wagner DL, Graham SA, editors. Geologic field trips in northern California Centennial Meeting of the Cordilleran Section of the Geological Society of America. Special Publication 119, 1999:214–235.

Thomas RB, Lewis J. An evaluation of flow-stratified sampling for estimating suspended sediment loads. J Hydrol 1995; 170:27–45.

Whyte DC. The Gambonini Mercury Mine — water quality threats and remediation alternatives. California Regional Water Quality Control Board, San Francisco Bay Region, Staff Summary Report. Oakland, California, 1998:15.