Review of “Delta Mercury Control Program Phase 1 Review of the Sacramento-San Joaquin Delta Estuary Total Maximum Daily Load for Methylmercury: Staff Report for Scientific Peer Review (March 2024)

Dr. Laurel Larsen
Associate Professor, Departments of Geography and Civil & Environmental Engineering, University of California-Berkeley
January 6, 2025

Scope of Review

Based on my expertise and experience, I am reviewing the findings, assumptions, or conclusions I agreed I could review with confidence:

- Conclusion 3: The DMCP Review’s proposed water balance and methylmercury mass balance reasonably quantify and account for all water and methylmercury source and loss types in the Delta.
- Conclusion 4a: DMCP Review’s proposed methylmercury load allocations and waste load allocations are achievable considering current technology, feasibility of controlling the sources, and recommended methylmercury allocation compliance calculations
- Conclusion 4c: Achieving load allocations and waste load allocations for Delta regulated entities (e.g., MS4s, WWTPs, irrigated agriculture) will result in a measurable reduction in Delta aqueous methylmercury concentrations. This conclusion should be considered apart from whether other loads are achieved.
- Conclusion 5: The DMCP Review’s proposed methylmercury source analysis, allocations, and compliance calculation methods reasonably account for climatic variability.

I structure this review by first providing a set of overall comments and then detailed comments associated with each of the four conclusions above.

General Comments

Board staff are to be commended for the comprehensive set of updates made to the 2010 TMDL staff report. For the most part, they have done an excellent job at summarizing and incorporating results from the Delta Mercury Control Program studies, the independent scientific peer review panel findings from those studies, and in greatly expanding the range of data and water-year types included in their analysis. The comprehensive evaluation of over 400 candidate linkage models is likewise commendable. The processes affecting mercury transformation and transport in the Delta are incredibly complex, and the science highly uncertain, and in general, staff have taken a reasonable approach to developing quantitative estimates of linkages, loads, and needed TMDL allocations in the face of this uncertainty. My detailed critical comments often suggest alternative approaches that may be regarded as more scientifically robust, but given the high degree of uncertainty and variability in mercury fluxes and transformations, whether to undertake these analyses is a decision that should be weighed against staff time involved and likely sensitivity of the outcome to changes.

For future public release of this staff report, it would likely be helpful to include more general information about the approach to establishing load allocations up front, as a sort of road map. Specifically, there should be a statement that, as in the 2010 staff report, the net 110 kg/yr reduction in mercury load required by the San Francisco Bay TMDL is proportionally allocated to tributary inflows, and that the analyses covered in the staff report are primarily oriented toward achieving goal concentrations in fish tissue (which are linked to corresponding aqueous goal concentrations). Forefronting this information (currently, the information that the SF Bay TMDL load reduction is to be met by reductions to tributary inflows is in the final chapter of the report) provides important context to the decision to use median flow and concentration values and could reduce concerns about the implications of this decision for calculating annual loads. Secondly, doing so highlights important details about the timescale of the calculations: large fish like black bass accumulate mercury over their 10-15-year lifespan, so goal concentrations in fish tissue are much more sensitive to longer-term averages than to seasonal or interannual variability.

Assessment of Conclusion 3: The DMCP Review’s proposed water balance and methylmercury mass balance reasonably quantify and account for all water and methylmercury source and loss types in the Delta

In my professional opinion, conclusion 3 is *somewhat supported.* The staff report does an excellent job in identifying and attempting to quantify all significant sources and sinks of water and methylmercury in the Delta. It is commendable that in this report, tidal and nontidal sources are separated. Further, comparisons of some individual estimated fluxes to relevant fluxes computed elsewhere (e.g., the 2003 and 2008 CALFED

Reports) provide confidence in estimates. However, there are several salient areas of concern:

Method of medians for load calculations:

While the use of median methylmercury concentrations is appropriate for the linkage model given the integrative nature of bioaccumulation, the use of median monthly flow and concentration values is less appropriate for developing a mass balance that elucidates relative concentrations of different sources and sinks, particularly for highly variable time series. Unlike mean values, medians may be completely insensitive to storm events or other “hot moments” that transport biologically relevant loads of methylmercury into or out of the Delta and should likely be considered in understanding relative contributions of dischargers. For this reason, widely used software packages for developing load contributions from discrete data, such as the USGS’s LOADEST (Load Estimation; Runkel et al., 2004) or RLOADEST packages (the implementation of LOADEST in R; Lorenz et al., 2015) are grounded in the computation of representative averages, and there is widespread precedent in the use of average loads for mass balance calculations in systems such as the Great Lakes (e.g. Robertson et al., 2018). I suspect that the use of medians may be a big component of the noted mass balance closure errors.[1]

Second, the step in the data processing pipeline at which the relevant median is computed is important, and I believe it has been handled inappropriately. In general, to ensure that the median value is most representative of long-term conditions, the median should be taken as late in the data processing pipeline as possible. Since the goal for the mass balance calculations is to develop an annual mass balance, this means the most relevant median would be to take the median of the summed monthly totals (i.e., to take the median across years, rather than the median of the summed monthly median, as is done now). When flow and methylmercury concentrations are not statistically independent, it is also very important to take the median of the instantaneous product of concentration and flow (rather than the product of median flow and median concentration, which can be substantially different from the former and not representative of an “average” load.) Furthermore, wherever possible, measured daily data should be used to develop estimates aggregated to longer timescales. For example, to estimate flow from the Cache Creek Settling Basin (CCSB), staff calculated

---

[1] For certain components of the mass balance that are not expected to be highly variable within a month, the median may be an appropriate choice. Board staff are commended for conducting some comparisons of load calculations using the mean vs. median, such as the comparison of methylmercury loading from MS4 dischargers discussed on p. 148, for which the use of the median instead of the mean was not found to be the main cause of why the present estimate is less than that of the 2010 TMDL staff report.

monthly averages (medians) from the dataset, disaggregated those values to daily values, and then summed the disaggregated values to compute representative monthly volumes that were summed to develop annual volumes (section 6.3.7.2). This approach may hide variability in the annual volumes that would otherwise be computed from gap-filled daily measurements and may skew the estimation of a representative average or median.

In some places, choices for why the median annual load was computed in a certain way were not made clear. For example, on p. 85, it is explained that on several main tributaries, annual loads were determined as the sum of monthly loads, whereas on other tributaries, the estimated annual median flow was multiplied by the annual median methylmercury concentration. Presumably this difference is due to differences in data frequency/availability, but this should be stated.

Representation of interannual variability

Presently, mass balance computations represent a long-term central tendency, masking the interannual variability that characterizes California's hydrologic dynamics. For purposes of the linkage model, this decision makes sense, as fish tissue integrates long-term variability in aqueous concentrations. However, for purposes of determining relative contributions of different sources and sinks to long-term aqueous concentrations in order to distribute allocations, this approach may be problematic in that it can underestimate contributions of sources that are nonlinearly activated during unusually wet or unusually dry periods.

I encourage Board staff to stratify mass balance results by water year type to better account for this climatic variability. Presenting this data graphically could help document uncertainty in the estimates of source and sink contributions by bracketing their variability. It would also highlight whether certain interventions would need to be targeted to abnormally wet or dry years and whether the calculus of relative source contributions changes substantially with climate. Finally, as climate change is expected to enhance extremes, stratifying the annual load data in this way would help refine the identification of how these loads are likely to change in the future.

Gap filling

For sources such as certain tributaries or the Cache Creek Settling Basin that have gaps in the gage data, staff set the gaps to zero with the justification that these gaps tended to appear during the dry season and likely reflect periods of intermittent flow. While I agree that this often may be the case, my own experience with gage data interpretation suggests that short-timescale gaps due to sensor malfunctioning or reasons other than zero-flow conditions are also common. There are many well documented approaches (e.g., USGS standard practices) for diagnosing the potential cause of gaps and filling them for annualized analyses, using methods that depend on the duration of the gap (e.g., regression-based, interpolation-based). I suspect that several of the gaps staff encountered may not actually be due to zero-flow conditions and that setting their values to zero may bias flow estimates to be low. At a minimum, additional information about the characteristics of gaps (e.g., percentage of the available data, range of durations) would be helpful information to include in the staff report, as well as information about how the data may have been pre-processed by the collector prior to Board staff's additional processing.

Evaporation

The estimation of evaporation (technically evapotranspiration) is based on scaled-up discrete measurements from four Delta islands collected over various periods from 1948-1977 and from a more modern estimate of open-water evaporation. This generates a very coarse estimate at best and could potentially be one of the sources of error in the mass balance. Board staff included a question in the staff report about whether the mass balance might be double-counting evapotranspiration by including channel depletions computed by the DICU model, and I can confidently confirm that it is; all channel depletions in the DICU model are ultimately due to evapotranspiration, as groundwater seepage is from channels into the islands. Alternative and more relevant and reliable estimates of evapotranspiration for the time period of interest may be derived from DWR models such as Bay-Delta SCHISM, which provides this information in a spatially explicit manner that does not double-count Delta island channel depletions. A discussion between Board staff and DWR Bay-Delta modeling staff would be informative. Another more robust alternative would be to calibrate an evapotranspiration model to the available data and apply that model over the years of the study based on available climatological data.

Need for additional clarification:

Little information is provided about the precipitation gage data that were used in the analyses. Appendix A references an Excel spreadsheet of precipitation that was not provided and indicates that "several" gages were used in the analysis. When several gages are available, there are multiple methods available to apply the data from the gages to spatial areas to ensure appropriate areal weighting, including through kriging or the use of Thiessen polygons. It would be helpful for the update to the staff report to specify the exact number of gages used and how the data were distributed over space.

Given the known spatial variability in precipitation, reliance on discrete gages and assumptions about how their data apply geographically can be another source of error in mass balance computations.

Methods of scaling up estimates:

In general, the staff report is robust in highlighting limitations inherent in the estimates, based on available data. One of the biggest limitations in the development of the mass balance that is perhaps not emphasized enough is that of scaling up estimates from just a few experiments, such as the BREW study control ponds and Twitchell Island experimental ponds. Although the flux rates for nontidal wetlands adopted for the current staff report somewhat address a previous concern about constant flow-through conditions, they are still limited in applying rates measured in these studies to all nontidal wetlands. An alternative to consider would be to determine representative inflow and outflow rates from spatially explicit models, such as the Bay-Delta SCHISM model or RMA2D/3D models.

A related concern—involving scaling up in time—pertains to the method of computing methylmercury outflux to the San Francisco Bay. Methylmercury data were available only for a selection of water years in the 2000-2019 time frame. Staff used the available data to compute monthly medians, which were then multiplied by median monthly flows for the entire time period of interest. However, this approach for scaling up might not be robust if the water year classifications for the years with available methylmercury data are not representative of the time period of interest. For example, the years with available methylmercury data contain no critically dry years. A better approach would be to undertake a similar analysis subset by water year type, or to develop regressions between flow and methylmercury concentration for the monthly or annual fluxes and use this approach to fill gaps. The use of load computation packages such as the RLOADEST package mentioned earlier would be helpful in this type of analysis.

Representation of the Yolo Bypass

I am especially concerned that the Yolo Bypass may not have been appropriately represented in the methylmercury source accounting, which may have contributed substantially to the mass-balance discrepancy that methylmercury losses exceeded source contributions. As mentioned in section 6.2.8.1, it is assumed that methylmercury fluxes from the inundated Yolo Bypass is assumed to be indirectly accounted for in source estimations such as agriculture, wetland, and atmospheric deposition. Nevertheless, these estimates do not account for the higher methylmercury loads in the soil of farmed regions of the Yolo Bypass cited in Windham-Myers (2010), and contributions from wetlands are represented as sinks in the present analysis, as they fall under the nontidal wetlands category. The result is that the Yolo Bypass has a relatively small contribution to total methylmercury loads in the DMCP analysis, which contrasts with the recent findings of the DWR Mercury Open Water Final Report (2020), which suggested that the Yolo Bypass is an important internal source of methylmercury, contributing up to 1/3 the load of the tributaries to the Delta. These contributions were attributed primarily to flooded agricultural soils. I contend that a separate method of accounting for contributions of flooded soils in the Yolo Bypass is needed for the DMCP analysis, given the substantially different hydrologic management characteristics for this region compared to farmed islands elsewhere in the Delta.

Double counting in particle settling (sink) estimates

On p. 173 of the staff report, Board staff note that the settling estimate based on the 2008 CALFED report may be an overestimate, as settling is also inherently included in load estimations for wetlands and agricultural lands. This staff concern seems relatively straightforward to address by subtracting out the settling that occurs over these areas, presuming that the areal coverage is known. However, this is likely a small source of error within the mass balance.

Summary of Major Recommendations:

- In a future update, Board staff may wish to consider using peer-reviewed load computation packages (such as the USGS RLOADEST package) for data pre-processing (e.g., addressing issues of missing values or the need to scale up measurements) and computing loads from sources with discrete discharge and concentration time series over various desired timescales of aggregation.
- If it is infeasible to update the analysis with a package such as RLOADEST, it is recommended that staff undertake a more rigorous (e.g., regression-based, interpolation-based, etc.) gap-filling procedure to compute daily or monthly-scale loads (concentration x discharge) and then aggregate those loads to the annual scale before taking the median.
- In a future update, Board staff may wish to use a combination of the DICU model and either DSM2 or Bay-Delta SCHISM to estimate evapotranspiration, applying their current approach only to those areas of the Delta Mercury TMDL boundary that are outside the domain of the DICU.
- Board staff should consider a different method for estimating potential methylmercury contributions from flooded agricultural soils in the Yolo Bypass, given substantial differences in hydrologic regime between this region and farmed islands elsewhere in the Delta and large discrepancies between the

DMCP source estimates for the Yolo Bypass and the DWR Mercury Open Water Final Report (2020).

- Interannual variability in estimates should be better accounted for by stratifying load and mass balance computations by water year type (or, if more feasible, aggregated water year types, such as dry + critical).

Assessment of Conclusion 4a: DMCP Review's proposed methylmercury load allocations and waste load allocations are achievable considering current technology, feasibility of controlling the sources, and recommended methylmercury allocation compliance calculations

In my professional opinion, Conclusion 4a is *potentially supported, with caveats*. In making this assessment, I focus on the new proposed allocations of the DMCP Review, excluding the allocations assigned to Delta tributaries to meet the SF Bay TMDL load reduction for the Delta, as there is little scientific evidence to suggest whether these reductions are feasible. For the source allocations within the Delta assigned to meet the aquatic methylmercury goal level, the DMCP Review suggests that these allocations are achievable in many subregions given implementation of current technology such as LID controls, as discussed in summary of the characterization and control studies (Appendix E). Additionally, the expansion of the data analysis to include additional years has shown that WWTFs, and MS4s have already achieved measurable reductions in methylmercury loads (Section 8). However, for some subregions or sources, control and characterization studies did not propose or address the feasibility of methylmercury control options (e.g., the control and characterization study on irrigated agriculture). Further, with respect to the Yolo Bypass, the Independent Scientific Peer Review Report on the Delta characterization and control studies concluded that it is unlikely that the TMDL target would be met.

As Appendix E highlights, while the early characterization and control studies are often promising as to whether proposed allocations are achievable, supporting data are strictly limited. For this reason, it will be important to adopt an adaptive management strategy to document progress toward achieving targets and make adjustments to rebalance allocations if certain types of allocations are infeasible.

Measuring progress toward achieving allocation targets, however, requires adequate monitoring. Recommended compliance calculations are based on a minimum of two data points per year for load: one during a wet period, and one during a dry period. This frequency is likely to be inadequate for estimating annual load contributions. I would recommend that at least one sample per month is acquired, though for more variable fluxes (or fluxes that might occur in threshold fashion, such as fluxes from flooded agriculture or managed wetlands), this frequency also may not be adequate. I recommend that Board staff determine a more representative frequency of measurement for compliance calculations by resampling existing datasets representative of the entities to which allocations are assigned. For example, to test the adequacy of the currently recommended two sample points, staff could randomly sample from among the data representative of wet and dry parts of the year and compare the resulting two-point-estimated annual flux to the more robustly estimated value.

Compliance is also based on a five-year rolling window, which I support, given the high interannual variability of flows and the fact that allocations are currently based on a central tendency. The choice to adopt a five-year window for compliance determinations is also in line with requests and findings that resulted from the subregional DMCP control studies and is consistent with the basis for compliance with the San Francisco Bay TMDL load allocation for the Delta.

Summary of Major Recommendations

- As the Board moves forward with implementing the TMDL allocations, an adaptive management strategy should be put into place to evaluate achievability of scaled-up allocations and to flexibly adapt the distribution of allocations if scaled-up implementation is infeasible. Monitoring associated with adaptive management will also refine understanding of load reductions achievable with particular technologies in particular places.
- Two data points per year for compliance is unlikely to be sufficiently informative of how annual loads are impacted by mercury control measures. It is recommended that Board staff engage in a data-guided selection process to identify the minimum sufficient number of control points that can robustly estimate annual loads.

Assessment of Conclusion 4c: Achieving load allocations and waste load allocations for Delta regulated entities (e.g., MS4s, WWTPs, irrigated agriculture) will result in a measurable reduction in Delta aqueous methylmercury concentrations. This conclusion should be considered apart from whether other loads are achieved.

In my professional opinion, this conclusion is *supported*, given that the allocations were computed to achieve the aqueous mercury goal concentration, which is presently well below actual mercury concentrations. It is commendable that Board staff have evaluated multiple scenarios for quantification of present subregion-specific methylmercury concentrations and selected the one that is most protective of human and aquatic system health. Staff have also used robust calculations to estimate the load reduction needed in order to achieve the goal concentration. I have verified that the subregion-specific needed percent reduction values reported in table 8.1 for scenario B correspond with the total estimated reduction values reported in the region-specific tables.

Strong support for this conclusion is predicated on the assumption that annualized load calculations for sources are robust, and that the represented subregional calculations of concentrations are robust. Although more monitoring is always helpful, scenario B does seem to be the best choice for calculating concentrations (though Board staff should be aware that pilot studies to estimate methylmercury concentrations from remote sensing are underway but that they are not currently accurate enough for application). Please see my previous comments on the computation of source input loads (Assessment of Conclusion 3), which I believe could be made sufficiently more robust through the use of means or through improved calculation of medians.

Assessment of Conclusion 5: The DMCP Review's proposed methylmercury source analysis, allocations, and compliance calculation methods reasonably account for climatic variability

In my professional opinion, this conclusion is *poorly supported* overall. However, support for the conclusion is contextually dependent. For the context of development of the linkage model and achievement of desired concentrations of methylmercury in fish tissue, the conclusion is fairly supported, given that the central tendencies used integrate data over a period of time that is likely representative with respect to the proportion of water year types. However, I will also note that the fish concentration data used were sampled from the years immediately following the 2012-2015 drought and hence may be skewed to be more representative of long-term dry conditions.

While data from representative water year types were used to develop the annual medians used in the load calculations and mass balances, reporting of central tendencies only masks interannual variability in loads that may ultimately be important in understanding relative contributions of sources and subregions to Delta methylmercury concentrations. As argued in my assessment of Conclusion 3 (p. 4), it would be helpful to report a range of loads derived from analyses stratified by water-year types. This reporting may be as simple as reporting two numbers (representing wet and dry conditions) rather than the single number reported now. Based on this reporting, staff may decide that adjustments may be needed for the allocation strategy. At a minimum, regulated entities would have improved information for anticipating/planning their load reductions across wet and dry years. Reporting how loads vary in extreme wet and dry years will also help better anticipate the impacts of future climate change.

As discussed in the Assessment of Conclusion 3, estimates of aspects of the water balance (particularly evapotranspiration) that are based on scaled-up measurements inadequately account for the effects of climate variability. However, it is challenging to speculate on whether these simplifications are likely to be consequential for the present analysis. A sensitivity test in which different flux variables are perturbed within their likely range would be informative about which aspects of the TMDL analysis are likely to be worth investing time in to improve.

Last, as discussed in the assessment of conclusion 4a, the two annual samples currently required for compliance are unlikely to adequately represent potential intra-annual variability in load due to climate variability. However, the use of a rolling window of 5 years for calculation of the annual median is likely to capture interannual variability.

Summary of major recommendations:

- Load and mass-balance computations should be stratified by water year type to bracket the effects of interannual hydrologic variability. At a minimum, two values, representative of wet and dry conditions, should be reported.
- More than two annual measurements should be required for compliance monitoring in order to adequately capture intra-annual variability.

Additional References Cited

References cited beyond those in the staff report are detailed below:

Lorenz, D., Runkel, R. and De Cicco, L., 2015. River load estimation, rloadest package. *US Geol. Surv. Software*. https://github.com/USGS-R/rloadest

Robertson, D.M., Hubbard, L.E., Lorenz, D.L. and Sullivan, D.J., 2018. A surrogate regression approach for computing continuous loads for the tributary nutrient and sediment monitoring program on the Great Lakes. *Journal of Great Lakes Research*, *44*(1), pp.26-42.

Runkel, R.L., Crawford, C.G. and Cohn, T.A., 2004. *Load Estimator (LOADEST): A FORTRAN program for estimating constituent loads in streams and rivers* (USGS Techniques and Methods No. 4-A5). https://doi.org/10.3133/tm4A5.