A method for estimating daily loads from coastal watersheds with adjustments for dilution and concentrating effects, utilizing the National Hydrography Dataset Plus Estimated Runoff Model. By Karen R. Worcester and David M. Paradies. Central Coast Ambient Monitoring Program. 2015 draft.

Abstract

We describe an approach for estimating daily flows and loads from small coastal streams in central coastal California. Estimations are derived from monthly measured flow and concentration data, coupled with daily flow and dilution models. Daily flow estimation makes use of average annual stream flow estimates from the National Hydrography Dataset Plus Estimated Runoff Model (EROM). From locally available stream gage data, daily ratios are developed between daily flow averages and the EROM average annual flow estimates for each gage location. Where multiple gages are used, the ratios are averaged. These ratios are then applied to the EROM average annual stream flow estimates at nearby stream sampling locations to estimate daily flows. Measured monthly flows are used as benchmarks to assess performance.

Loading estimates derived from linear interpolation of monthly flow and concentration data are likely to have a high degree of error because of intervening storm events which may dramatically effect loading and may cause dilution effects. In some cases high flow may reduce concentrations due to dilution and in other cases, pollutants may increase in concentration with flow because of overland transport. We characterize this relationship at the scale of individual watersheds and describe a method for improving the accuracy of loading estimates by application of watershed specific relationships between pollutant concentration and flow.

Introduction

The California central coast is characterized by relatively small coastal watersheds that in some cases transport relatively high loads of pollutants to the Pacific Ocean. Few of these watersheds contain flow gages and very few contain any water quality instrumentation, but monthly water quality and flow data is available from the Central Coast Ambient Monitoring Program (CCAMP) and related efforts. An approach was needed to better characterize flows and loads leaving these watersheds; there is a continuing need to understand pollutant loading to the ocean in support of epidemiological studies of marine mammal disease (M.A. Miller et al., 2002; M.A. Miller et al., 2005; W.A. Miller et al., 2006; Stoddard et al, 2008), studies on the role of anthropogenic nutrient inputs in plankton bloom initiation (Lane et al., 2009, Kudela et al., 2008; Ryan, et al., 2008), and other purposes. Our approach is adaptable to other areas where local gage data and stream flow and water quality measurements are available.

CCAMP is the California Central Coast Water Board's regionally-scaled water quality monitoring program. It operates along a 300-mile stretch of California coastline from southern San Mateo County through northern Ventura County. CCAMP's "Coastal Confluences" monitoring program provides a long-term dataset of monthly stream flow and water quality measurements at the ocean discharge point of thirty-three streams and rivers. Associated watersheds range in size from X to Y square kilometers.

The Central Coast Daily Flow Estimation (CCDFE) approach described here utilizes annual average flow estimates available within the National Hydrography Dataset Plus Version 2 (NHDPlus V2) (Horizon Systems, 2005) and improves upon these estimates for more locally-scaled purposes, using locally available stream gage and flow measurement data. NHDPlus V2 is a suite of nationally scaled geo-spatial datasets that provide a hydrologic framework for analysis, and that integrate features from the the National Hydrography Dataset, National Elevation Dataset, and the Watershed Boundary Dataset (McKay, L. et al., 2012).

Included in the NHDPlus V2 geospatial framework is the Estimated Runoff Model (EROM), which estimates average annual flows for each NHD stream catchment. Runoff is derived from a "flow balance model" (Wolock and McCabe, 1999), which incorporates precipitation, potential evapotranspiration, evapotranspiration, and soil moisture storage (McCay, et al., 2012). Incremental flows are calculated by routing and accumulating catchment level flows using NHDPlus flow lines. This estimate of "natural" runoff is further refined when necessary by accounting for instream losses due to natural hydrologic processes, an important correction for flow estimates in the arid West. The estimate is also corrected for a negative bias in base flow estimates through a log-log regression-based adjustment using “reference” gage data from largely unaltered river and stream systems (McCay, et al., 2012). The model can account for flow additions, removals and transfers where data is available. Flow estimates are adjusted where on-stream flow gage data is available, pro-rating for differences in drainage area. Only gages that meet specific screening criteria (accuracy of reported drainage area, period of record, number of years of record) are used (McCay, et al., 2012).

Unmodified, the NHDPlus EROM model provides annual average flow estimates for medium resolution (100 K) hydrographic stream reaches defined by the National Hydrography Dataset. CCDFE uses the EROM annual flow estimates as the basis for more spatially and temporally explicit estimates of flow, adjusted with locally available flow gages and verified with measured stream flows. CCDFE describes flow at CCAMP Coastal Confluence monitoring sites at daily resolution. and reconciles the following issues inherent in the NHDPlus EROM model because of its national scale: (1) CCDFE allows use of flow data from more locally available stream gages that may have a shorter period of record or that may otherwise not be employed in the EROM calculations , (2) CCDFE allows for flow adjustments using gages from nearby watersheds with similar hydrologic, climatic, and land use characteristics , (3) CCDFE makes use of monthly measured flow data at coastal confluence sites to optimize choice of gages and to assess “fit” of the model.

Richards (1997) addresses the challenges associated with load estimates developed from recording flow gages paired with infrequent interval analyte grab samples. He offers several methods for addressing these challenges, some of which are employed here. Others have compared the performance of various estimation approaches and generally note that bias increases and precision decreases with sampling interval (Stelzer et al., 2006; Birgand, et al., 2011; Zamyadi et al., 2007). Birgand et al. (2011) reported uncertainties of +- 40% for monthly sampling of most nutrients using either straight linear interpolation or a ratio approach where annual flow volume is multiplied by annual flow-weighted concentration (generated from monthly measurements). Stelzer et al. (2006) addresses the issue o f "flashiness", a measure of day-to-day flow variation, where streams that are highly flashy are more influenced by sampling frequency than less flashy flow systems. This can result in higher bias and/or lower precision in modeled load estimates.

Studies have shown that for some analytes there is a "diluting" effect with increased flow, whereas with other analytes the opposite is true. Birgand et al. (2011) report a concentrating effect with increased flow for analytes associated with particles (including phosphorus and total suspended solids), a diluting effect for soluble analytes such as dissolved organic carbon, and mixed effects for nitrogen analytes. Stelzer et al. (2006) reported an exponentially declining effect associated with dissolved silica. These effects can result in loads to either be over-estimated or under-estimated by straight linear interpolation (Stelzer et al., 2006; Birgand et al.).

Methods

Field Data Collection **-** Monthly flow and nutrient data are collected by CCAMP in accordance with California State Board's Surface Water Ambient Monitoring Program Quality Assurance Program Plan (www.waterboards.ca.gov/water_issues/programs/swamp/). Depth-integrated samples are collected into 1 L plastic bottles from the center of the stream flow or thalweg and immediately placed in cold ice chests (4 °C) for transport to the laboratory for nutrient analysis. Flow measurements used for validation of the CCAMP stream flow estimation model are typically collected monthly along a ten-point cross section using a Marsh-McBirney conductive probe flow meter and setting rod. In locations where this is not feasible (for example at culverts) other methods of estimating cross-sectional area are used.

Central Coast Daily Flow Estimation (CCDFE) - CCAMP enhancements to the NHDPlus EROM model include selection of one to three USGS gages that more directly represent localized flow conditions at the site of interest. Gages are also selected to ensure that flow data throughout the time period of interest is available from at least one gage at any given time. In some cases a gage may reside on the same stream system as the site of interest. These gages may not be represented in the existing EROM model because of its required period of record and other criteria. Ratios are developed between gaged daily flow measurements and the EROM annual mean flow at each gage location. If more than one gage is used these ratios are averaged. Mean daily flow ratios are then multiplied by the NHDPlus-derived annual mean daily flow at the discharge location of interest to estimate flow at that location and point in time. Gage choice is optimized by evaluating performance against CCAMP measured monthly stream flows, collected since 2005. Measured flows provide benchmark data for flow model validation, through simple linear regression. In some cases, flow has not been collected at the site of interest. In two of these cases, the San Lorenzo and Pajaro rivers, modeled flow was benchmarked against measured flow collected at upstream sites, after adjusting for difference in watershed area.

<u>Load Assessment Including Adjustments for Dilution effects</u> - The bulk of the concentration and measured flow data available for the coastal loading assessment is collected at monthly intervals via the CCAMP coastal confluences monitoring program. We initially explored using linear interpolation to impute concentration values for days between two consecutive measured concentrations (Chescheir, et al., 2010; Birgand et al., 2010; Zamyadi et al. 2007). Because of the monthly sampling interval, these estimates are likely to have a relatively high degree of error; Birgand et al. (2011) estimated 30 to 80%.

We observed that, depending on the watershed, some analytes display concentrating effects with additional flow, whereas others show diluting effects; this effect has been noted by others in other locations (Chescheir et al., 2010; Birgand et al, 2011; Stetzer, et al.). We also noted that this effect could vary in different watersheds depending on source of pollutant; point sources tended to dilute with flow increases, whereas in some instances non-point sources tended to concentrate. In order to improve our estimates we determined that we should account for these effects where they are statistically significant.

We developed an algorithm to scan site-level analyte - flow relationships to identify possible concentrating or dilution effects. The scanning algorithm employs flow interval, ratio estimation, and regression approaches (EPA, 2003; Richards,1997). We created 30 logarithmically scaled bins for characterizing flow ranges. For all sites and analytes of interest, we produced tables of average. maximum and minimum concentrations within each of the 30 flow bins. For each site – analyte combination, we used linear regression to calculate coefficients of determination on average concentration and the lower boundary of the respective flow bin. We successively removed low-end flow bin values from the regression, until either the p-value was less than 0.05 or fewer than six flow bins remained. The intent in doing this was to eliminate lower flow relationships with concentrations that may be dominated by factors other than runoff (such as groundwater contributions, local effects, etc.). Only sites with significant positive or negative relationships between flow and concentration were adjusted for diluting or concentrating effects.

After we determined which relationships were significant, values were imputed for empty bins using linear interpolation between bin values on either side of the empty bin(s). Values were calculated for both bin maximums and minimums. Appendix A shows analyte relationships for all sites, using nitrate as an example (John – not yet developed…we may or may not want this in the report).

For site - analyte combinations where linear regression showed a significant increasing or decreasing slope, adjustments were made to incorporate concentrating and diluting effects. Adjustments were made using a hybrid algorithm involving both flow interval and ratio estimator approaches (described by Richards, 1997 and EPA, 3002). The maximum and minimum concentration/flow ratios were calculated for each flow bin. In the case of diluting effects, the maximum concentration in the bin was divided by the lower flow bin boundary. This provides the most conservative estimate of maximum dilution effect for that flow regime. In the case of concentrating effects, the minimum concentration in the bin was divided by the lower flow bin boundary. This provides the most conservative estimate of minimum concentrating effect for that flow regime.

The hybrid algorithm utilizes linear interpolation between two monthly measured concentrations to establish a trial concentration for each day that lacks a measured concentration. Where no significant diluting or concentrating relationship is detected, the trial concentration is used as the imputed concentration. For site - analyte combinations that show significant dilution effects, the maximum daily flux is multiplied by the modeled flow to derive a maximum concentration. When a trial concentration exceeds the calculated maximum concentration, then the maximum concentration is used as the imputed value. If the trial concentration does not exceed the maximum concentration, the trial concentration is used as the imputed value. In the case of site - analyte combinations with significant concentrating effects, when a trial concentration is less than the calculated minimum concentration, then the minimum concentration is used as the imputed value. If the trial concentration is not less than the minimum concentration, the trial concentration is used as the imputed value.

Results

Linear correlations between model results and measured stream flows for most coastal confluence sites are relatively high. R-square values for the twenty-eight benchmarked sites range from 0.307 to 0.988 and average 0.756. Slopes range from 0.877 to 1.663 and average 1.077 (Table 1). Some watersheds present more modeling challenges, depending on gage location, anthropogenic disturbance, and watershed behavior. For example, the Santa Maria River is highly influenced by extractions, irrigation runoff, and by an upstream reservoir. The Old Salinas River has flow influence from the Salinas River via a tide gate during higher flow events, and is heavily influenced by irrigation runoff and tile drain water. Field measurements at some sites are confounded by tidal influence and/or access problems. The coastal confluence sites which are not routinely measured for flow by CCAMP crews include Scott Creek (tidal influence), San Lorenzo River (not wadable), Pajaro River (not wadable), Old Salinas River and Tembladero Slough (tidal influence), Little Sur River (no access), and Mission Creek (limited access). Of these, Pajaro and San Lorenzo rivers were benchmarked using upstream site data, and Mission Creek was benchmarked using a relatively limited dataset.

SiteTag	Waterbody	R-squared	Slope
304APW	Aptos Creek	0.8815	1.0008
304GAZ	Gazos Creek	0.9692	0.9890
304LOR	San Lorenzo River*	0.9596	0.9898
304SOK	Soquel Creek*	0.9392	1.0102
304WAD	Waddell Creek	0.7543	1.0631
305THU	Pajaro River*	0.9879	0.9979
307CML	Carmel River*	0.9555	1.0842
308BGC	Big Creek	0.7586	0.9895
308BSR	Big Sur River*	0.9712	1.0883

308WLO	Willow Creek	0.6703	1.0419
309DAV	Salinas River*	0.9667	1.0309
310ADC	Arroyo de la Cruz	0.5681	1.0532
310ARG	Arroyo Grande Creek*	0.5278	0.9972
310PIS	Pismo Creek	0.7040	1.0045
310SLB	San Luis Obispo Creek	0.8163	1.0114
310SRO	Santa Rosa Creek	0.8590	0.9987
310SSC	San Simeon Creek	0.9014	1.1805
310TWB	Chorro Creek	0.5932	0.8765
312SMA	Santa Maria River*	0.3963	1.4918
313SAC	San Antonio Creek	0.7309	0.9965
314SYN	Santa Ynez River*	0.6485	1.0792
315ABU	Arroyo Burro	0.9317	1.1794
315ATA	Atascadero Creek (S Barbara Co.)		
315CRP	Carpinteria Creek	0.3272	1.1748
315FRC	Franklin Creek	0.3068	1.0643
315GAV	Gaviota Creek	0.8636	1.6632
315JAL	Jalama Creek	0.5773	1.0009
315RIN	Rincon Creek	0.8565	1.0281

Table 1. R-square values and slopes of regressions relationships for benchmarked coastal confluences. * denotes at least one of the gages is present in the watershed being modeled.

Notes for future analysis – We need to look at how the loading values differ between this approach and a straight linear interpolation. It would be good to be able to say, % difference in total loading varied from x% to y% in estimates that were adjusted based on a significant diluting or concentrating

We should develop a Table of site – analytes with significant relationships and note direction of change.

Birgand, F. T.W. Appelboom, G. M. Chescheir, R. W. Skaggs. 2011. Estimating Nitrogen, Phosphorus, and Carbon Fluxes in Forested and Mixed-Use Watersheds of the Lower Coastal Plain of North

Carolina: Uncertainties Associated with Infrequent Sampling. Amercian Society of Agricultural and Biological Engineers ISSN 2151-0032. Vol. 54(6): 2099 – 2110.

Chescheir, G. M., F. Birgand, S. Tian, M. A. Youssef, D. M. Amatya. 2010. The Effect of Sampling Frequency on the Accuracy of Nitrogen Load Estimates from a Drained Loblolly Pine Plantation in Eastern North Carolina. XVIIth World Congress of the International Commission of Agriculture and Biosystems Engineering. Quebec City, Canada, June 13 – 17, 2010.

Lane, J. Q., P. T. Raimondi, R.M. Kudela. 2009. Development of a logistic regression model for the prediction of toxigenic *Pseudo-nitzschia* blooms in Monterey Bay, California. Marine Ecology Progress Series. Vol 383: 37 – 51.

Kudela, R. M., Lane, J. Q., Cochlan, W. P. (2008a). The potential role of anthropogenically derived nitrogen in the growth of harmful algae in California, USA. Harmful Algae 8: 103–110.

McKay, L., Bondelid, T., Dewald, T., et al. 2012. "NHDPlus Version 2: User Guide". U.S. Environmental Protection Agency.

Wolock, D. M. and G. J. McCabe. 1999. Estimates of Runoff Using Water-Balance and Atmospheric General Circulation Models. Journal of the American Water Resources Association 35(6):1341 – 1350, December 1999.

Miller, M.A., I.A. Gardner, C. Kreuder, D.M. Paradies, K. R. Worcester, D.A. Jessup, E. Dodd, M.D. Harris, J.A. Ames, A.E. Packham, P.A. Conrad. 2002. Coastal freshwater runoff is a risk factor for Toxoplasma gondii infection of southern sea otters (Enhydra lutris nereis). International Journal for Parasitology 32: 997 - 1006.

Miller, W.A., M.A. Miller, I.A. Gardner, E.R. Atwill, M. Harris, J. Ames, D. Jessup, A. Melli, D. Paradies, K. Worcester, P. Olin, N. Barnes, P.A. Conrad. 2005. New genotypes and factors associated with Cryptosporidium detection in mussels (Mytilus spp.) along the California coast. International Journal for Parasitology 35:1103–1113.

Miller, W.A., M.A. Miller, I.A. Gardner, E.R. Atwill, M. Harris, J. Ames, D. Jessup, A. Melli, D. Paradies, K. Worcester, P. Olin, N. Barnes, P.A. Conrad. 2005. New genotypes and factors associated with Cryptosporidium detection in mussels (Mytilus spp.) along the California coast. International Journal for Parasitology 35:1103–1113.

Stelzer, R.S., and G. E. Likens. 2006. Effects of sampling frequency on estimates of dissolved silica export by streams: The role of hydrological variability and concentration-discharge relationships. Water Resources Research, Vol. 42, W07415, doi:10.1029/2005WR004615.

Stoddard, R. A., E. R. Atwill, F. M.D. Gulland, M.A. Miller, H. A. Dabritz, D.M. Paradies, K.R. Worcester, S. Jang, J. Lawrence, B. A. Byrne, P. A. Conrad. 2008. Risk Factors for Infection with Pathogenic and Antimicrobial-Resistant Fecal Bacteria in Northern Elephant Seals in California. Public Health Reports 123:360 - 370.

Richards, P. 1997. Estimation of Pollutant Loads in Rivers and Streams: A Guidance Document for NPS Programs. Water Quality Laboratory, Heidelberg College, 310 E. Market St., Tiffin, OH 44883. U.S. EPA Grant X998397-01-0

Ryan, J. P., J. F. R. Gower, S. A. King, W. P. Bissett, A. M. Fischer, R. M. Kudela, Z. Kolber, F. Mazzillo, E. V. Rienecker, and F. P. Chavez. 2008. Geophysical Research Letters. Vol. 35, L12602, doi:10.1029/2008GL034081.

U.S. Environmental Protection Agency, 2003. National Management Measures to Control Nonpoint Source Pollution from Agriculture. EPA 841-B-03-004.

Zamyadi, A., J. Gallichand, and M. Duchemin. 2007. Comparison of methods for estimating sediment and nitrogen loads from a small agricultural watershed. Canadian Biosystems Engineering, Vol. 49(1): 27-36.

Appendex A. Flow – analyte relationships, using Nitrate + Nitrite as an example. Graphs in column A show all concentration and flow data used in the analysis. Graphs in column B show average concentrations in each flow bin, and the associated regression lines. The longer line is for the entire set of flow bins; the shorter red line is for higher end flow bins where R-2 exceeds 0.5. Graphs in column C show bin values after a smoothly algorithm is employed and missing values are imputed. Relationship for other analytes can be found at http://www.ccamp.us/k/???

Methods for Estimating Pollutant Loads (Richards, 1997)

Numeric Integration — Total load is calculated as the sum of the individual loads calculated for each sample.

Worked Record Procedure — Chemical observations are plotted onto a detailed hydrograph, and smooth curves are drawn through chemical data points based upon analyst's experience with the relationship of concentration and flow.

Averaging Approaches — Calculation that uses averaging of concentration and/or flow to estimate loads. For example, analyst might multiply average weekly suspended solids concentration by daily flow to estimate daily loads for the week.

Flow Interval Technique — Semi-graphical technique that calculates "interval loads" as the product of average flux for a range of daily flow values times the number of days in which flows were within the particular flow range.

Ratio Estimators — Total loads are estimated using a known relationship between the less-frequently sampled parameter of interest and a more-frequently sampled parameter (e.g., discharge) to fill gaps in the data record for the parameter of interest.

Regression Approaches — Relationship is established between concentration and flow based on samples taken, and then applied to estimate concentration for days not sampled.

Flow-Proportional Sampling — Mechanical approach in which representative samples are taken to determine concentration for a known discharge. Pollutant load is calculated as the sum of the sample concentrations multiplied by the measured discharge.

Figure From usepa 2003

Notes below:

Estimating Loads with available data

Given constraints of the available dataset:

Analyte concentrations were measured monthly

Flows were measured monthly when conditions permitted such measurements

We pursued an approach guided by tenets set forth in Richards,1997

"

1. Find a way to estimate "missing" concentrations: i.e. concentrations to go with the flows observed at times when chemical samples were not taken.
2. Abandon most of the flow data and calculate the load using the concentration data and just those flows which were observed at the same time the samples were taken.
3. Do something in between - find some way to use the more detailed knowledge of flow to adjust the load estimated from matched pairs of concentration and flow.

The second approach is usually totally unsatisfactory because the frequency of chemical observations is inadequate to lead to a reliable load estimate when simple summation is used. Thus almost all of the load estimation approaches which have been shown to give good results are variants of approaches 1 or 3. Total load vs. unit load

If all of the flow and concentration observations are available for all n intervals in formula (3) or (4) above,

the summation is an easy task. The summation is the total load for the time period of interest; each individual product, ciqiDt or ciqiti, could be called the unit load. If the total load is an annual load, the unit

load might be the daily load. If the total load is a weekly load, the unit load might be the hourly load. This distinction is important, because we tend to focus on the total load as our goal. The considerable complexities of the various methods for calculating loads, however, are almost always related to trying to accurately characterize the unit loads. Once this is done, the total load is easily obtained.

"

We addressed flow measurement issues through use of the ccamp flow model (Worcester, Paradies, initally presented at the 6th National Water Quality Monitoring Conference, Atlantic City NJ 2008)

Daily flows were estimated using the ccamp flow model

We addressed issues associated with flow dependent concentration variability in the following fashion:

We initially explored using linear interpolation to impute concentration values for days between measured concentrations.(as in Chescheir etal and other studies)

Exploratory data analysis revealed that site specific relationships existed between concentrations and flows. (as in Birgand etal, Stelzer etal,Zamyadi etal and other studies)

In some instances concentrations were reduced as flows increased.(as in Birgand etal, Stelzer etal,Zamyadi etal and other studies)

In some instances concentrations were increased as flows increased.(as in Birgand etal, Stelzer etal,Zamyadi etal and other studies)

We were concerned about the subjectivity associated with graphical identification of concentration and dilution effects (e.g. the K. Worcester pers comm and the USGS 'worked record' methodology).

We explored various methods of detecting the presence of these relationships.(we found no literature which provided an automated scanning methodology for detecting possible concentration or dilution effects)

We adopted a hydrid approach for scanning site - analyte combinations to identify possible concentration or dilution effects.

The scanning algorithm employs flow interval, ratio estimation, and regression (as described in EPA 841-B-03-004, July 2003 and Richards,1997).

We calculated average concentrations at each site in each of 30 flow interval-regime bins to produce a table of of limits for each flow interval regime

We successively calculated the regression based coefficient of determination (R squared) for average concentration in each flow regime for each site analyte combination in stepwise fashion begining with all flow regimes and successively eliminating low flow regime values one bin at a time until an R squared value exceeds .5 (or not)

The slope of the regression line is used to distinguish between concentration and dilution effects.

Notes:

The use of regression here is used only to detect site-analyte situations where concentration - dilution adjustments are needed.

The results of the regression are not used in the process of 'load estimation'.

We explored various methods of improving load estimates which address detected site specific relationships.

We adopted a hybrid approach based on a framework set forth in EPA 841-B-03-004, July 2003 and Richards, P 1997.

The hybrid approach includes:

A 'worked record' approach (as described in EPA 841-B-03-004, July 2003 and Richards, P 1997) modified to utilize the flow regime algorithm described above to mitigate the subjectivity normally associated with the 'worked record' approach with respect to dilution and concentration effects.

A 'flow interval' approach (as described in EPA 841-B-03-004, July 2003 and Richards, P 1997) which utilizes the flow interval regimes described above to impute daily flux, but not to calculate loads for more than a single day.

A 'ratio estimator' approach (as described in EPA 841-B-03-004, July 2003 and Richards, P 1997) which utilizes the flow interval regimes described above to impute daily flux, but not to calculate loads for more than a single day.

The hybrid algorithm initally utilizes linear interpolation to establish a trial value for concentration for each day which is missing a concentration measurement.

In the case of dilution effects; When a trial value exceeds the average concentration in a given flow regime, the average concentration for the flow regime is used as the imputed value.

In the case of concentration effects; When a trial value is less than the average concentration in a given flow regime, the average concentration for the flow regime is used as the imputed value.

Error associated with linear interpolation was examined in more detail using high frequency nitrate concentration data from an in-situ ultraviolet spectrophotometer sensor, deployed by Monterey Bay Aquarium Research Institute through its Land Ocean Biogeochemical Observatory (LOBO) network. The L03 sensor is located at the lower end of the Old Salinas River in Moss Landing Harbor. This instrument collects nitrate readings continuously at an hourly interval and has been in operation since 1994. Its location in a tidal area presents additional sources of variability that would not be encountered were the sensor located in a non-tidal riverine environment. We adapted this data for use by selecting daily measurements collected at salinity low points. We extracted monthly interval measurements from this dataset and used the subset to create a daily linear interpolation of nitrate concentrations. Linear regression between the measured and interpolated daily concentrations produced a significant relationship ($y = 0.7228x + 4.7881$; $P=0.00$, $r^2 = 0.53$). When averaged daily interpolated concentrations were compared to averaged measured concentrations for the period of record, the interpolated values underestimated average concentrations by 3%.