Evaluation of SWAC vs. RAL from Alternate COPC Mapping Groups and Stratum

October 11, 2016

Problem Statement

The United States Environmental Protection Agency (EPA) has been working collaboratively with the Cooperating Parties Group (CPG) to refine proposed mapping methods for contaminants of potential concern (COPC) at the Lower Passaic River portion of the Diamond Alkali Superfund Site. The CPG initially proposed a method that incorporated stratification of the river based on proximity to left or right shoals, channel, presence or absence of silt deposits upstream of River Mile 8 and based on erosion and deposition patterns downstream of River Mile 8. The USEPA had concerns with the mapping methods because some relationships are supported by relatively small numbers of samples, and because the CPG had not characterized uncertainty in the mapping procedures.

As a result of the working group meetings, the CPG developed new mapping methods based on geostatistical procedures which provided estimates of uncertainty in mapped values and also provided a means to propagate mapping error through estimation of the surface weighted average concentrations (SWAC) as well as to evaluate potential bias in the SWAC vs remedial action limit (RAL) relationship. These efforts also lead to a reduction in the number of analysis groups (strata) used for estimating the semivariogram and process variance, but retained the original stratification for estimating the stratum means, and simulated values were developed separately for each group. This constraint had the effect of holding stratum boundaries fixed as if the edges of silt deposits and erosional and depositional areas were knowns. EPA observed that these stratum boundaries are actually estimated from data which in some areas are relatively sparse, so this raised concern that the CPG mapping and simulation procedure might have understated uncertainty surrounding mapped values and also may not have fully mitigated potential bias in the SWAC vs RAL relationship. As a result EPA tested an alternative COPC mapping model which was used to develop additional simulations to study the degree to which EPA concerns might be substantively important to the RI/FS process.

Methods

EPA developed a geostatistical model of 2,3,7,8 TCDD (TCDD) concentrations in surface sediments for the lower Passaic River from River Mile 8 through 14. This portion of the river was selected for analysis because EPA has issued a Record of Decision (ROD) for the lower 8.3 miles. The selected remedy for the lower 8.3 miles of the river is spatially extensive and unlikely to be sensitive to the COPC mapping methods. Conversely, the remedial options under consideration for river miles 8 through 14 may be more focused, and as such, evaluation of the feasibility of such options may be more sensitive to mapping methods.

The geostatistical model developed by EPA is a nonparametric method combining indicator interpolation using natural neighbor interpolation with P-field simulation using the fast Fourier Transform method. Indicator interpolation involves selection of a range of concentration threshold values covering the percentiles of the TCDD distribution followed by coding locations as zero when concentration exceeds threshold values and as one when concentration does not exceed threshold values. For each threshold, these binary values are interpolated forming a map of the probability that TCDD concentration is less than the threshold value. Because the values being interpolated are binary, there are no distributional assumptions and because the data were interpolated using natural neighbor interpolation, there are no stationarity assumptions for the interpolation step. This use of natural neighbor interpolation is not a standard approach, but has the added advantage that development of the interpolated surfaces does not require development of semivariograms for each threshold value, as would be necessary for the indicator kriging method usually applied.

The indicator interpolation analysis described above results in an estimated cumulative distribution of concentrations for each grid cell. The cumulative distribution provides a correspondence between uniformly distributed probabilities, ranging from 0 to 1, and the range of TCDD values for each cell in the study grid.

Simulated concentration surfaces are constructed by selecting a random uniform value at each grid cell, and identifying the corresponding TCDD concentration at each location. To insure that the randomly selected values honor the spatial correlation of the sample concentration data, the random uniform values are simulated as a spatially correlated random uniform value bounded between 0 and 1.0.

The spatial continuity of this uniform value is obtained from the semivariogram of the uniform transform of the TCDD data. With this approach, the probability field is simulated as a stationary random field, but nonetheless the resultant concentration simulations are non-stationary because the cumulative distributions have spatially varying means, and variances.

EPA's primary focus of evaluation was a comparison of two simulation procedures based on "narrow stratification" with over 20 groups as defined by EPA and "broad stratification" based on a smaller number of general groups—left shoal, right shoal, navigation channel and silt deposits. The narrow stratification approach errs on the side of maximizing accuracy while potentially resulting in poor precision, whereas the broad stratification trades accuracy for increased precision. The SWAC vs RAL relationship was simulated based on both stratification approaches to determine how sensitive the relationship is to the assumed underlying stratification. Contrary to the approach developed by CPG, both methods used by EPA allowed for uncertainty in the stratum boundaries by simulating a single mean-zero residual process, independently of stratum boundaries which was then added to the estimated stratum means. In this way local high or low concentration deposits were allowed to extend across imperfectly known stratum boundaries.

Results

The simulated TCDD maps derived from the EPA models were used to estimate the SWAC vs RAL relationship and were overlaid to compare each stratification approach. The SWAC vs RAL relationship was plotted for the broad stratification approach shown as a solid blue line in Figure 1. Approximate 95% confidence bands were also plotted as black dashed lines, and the relationship based on the narrow stratification approach (e.g. CPG's smaller stratum sizes and therefore larger number of data groups) was overlaid and shown as a red dashed line. It can be seen that the best estimates, shown as blue solid line and a red dashed line are very similar and that the narrow stratification result is within the 95% confidence band for most actions limits. The narrow stratification result deviates slightly outside the 95% confidence bands for action limits below 200 ng/kg, but this is primarily due to narrowing of the confidence bands at low action limits, as opposed to actual larger deviations between the two estimates. While there is a wide range of evaluations that could be investigated by further post processing of simulation results, these summaries show that the SWAC vs RAL relationship, which is central to development of the FS, is insensitive to assumptions regarding how the concentration distribution is stratified, in the section of the river from River Mile 8 to 14. This analysis does not necessarily imply that the spatial distribution of contaminants is as highly resolved locally as may be necessary to develop a remedial design. It is anticipated that the FS can be supported by the current mapping procedure, but that, depending on the remedy selected following completion of the 17-mile RI/FS, the remedial design may require additional pre-design sampling.

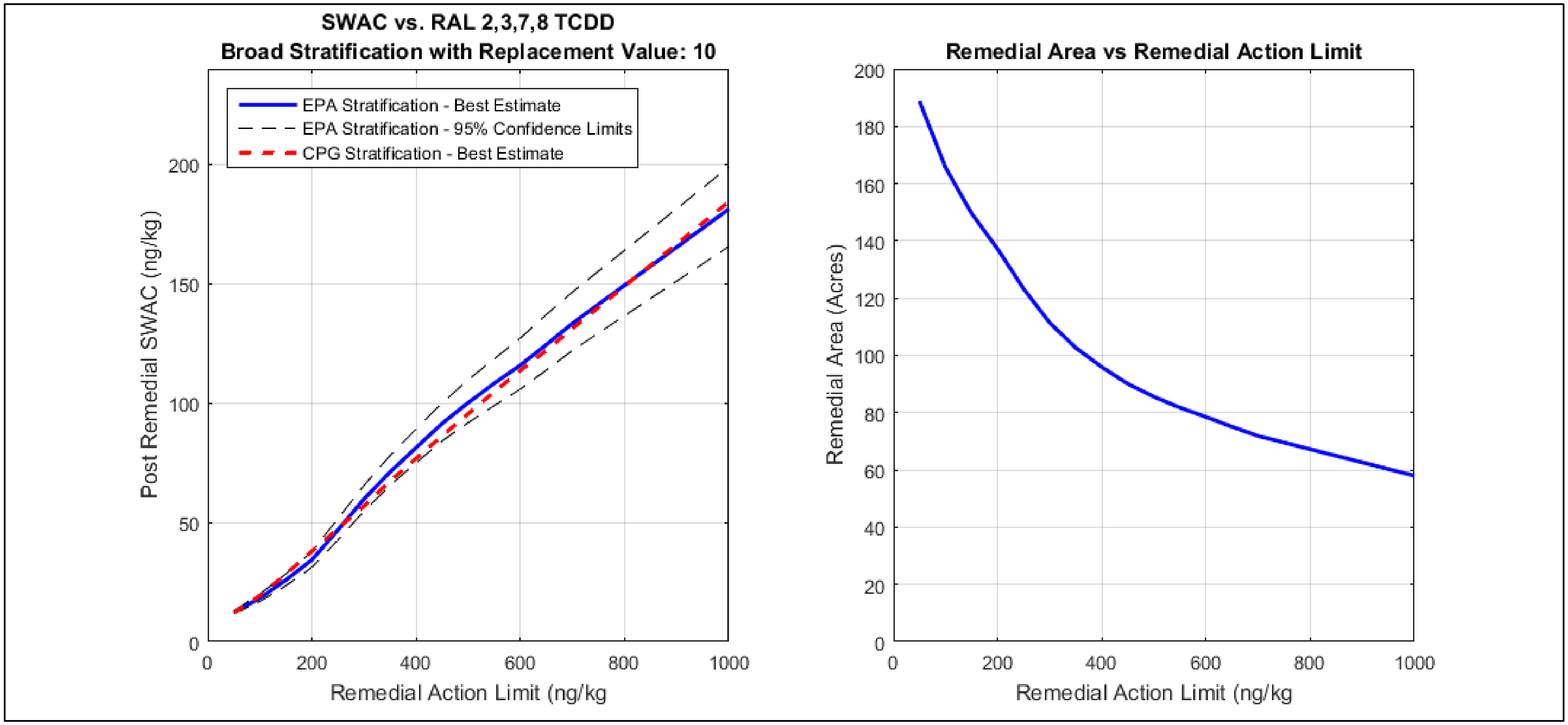

Figure 1. Surface weighted average concentration (SWAC) for 2,3,7,8 TCDD and remedial area versus remedial action limit (RAL) in sediments from River Mile 8 through 14 of the Lower Passaic River.