Regional PMF Analysis -- Blank Influence Analysis
Conceptual Scope of Work

This scope of work is the first phase of a project to utilize Positive Matrix Factorization to identify sources of PCBs to the Spokane River. Upon completion of Summer 2018 sampling activities and receipt of the analytical data a scope of work for the second phase will be developed. The second phase scope of work will identify the data sets to be used in the PMF analysis and detailed approach to PCB source characterization in the Spokane River.

In the first phase of the project described in the following scope of work we propose to study the effect of blank contamination on source apportionment of PCBs in the ambient surface water of the Spokane River by conducting factor analysis using Positive Matrix Factorization Model (PMF2) software on a number of permutations of the SRRTTF Spokane River water column data set in order to identify which approach (or combination of approaches) best addresses the impact of blank contamination on the analysis of low levels of PCB measured in the Spokane River water column.

What is Positive Matrix Factorization?

Positive Matrix Factorization (PMF) is a mathematical receptor model developed by Paatero and Tapper (1994) who developed the PMF2 software. EPA later adopted the technology, issuing PMF versions 3.0 and most recently 5.0. PMF is used to quantify the contribution of sources to samples based on the composition or fingerprint of the sources. The PMF model can analyze a wide range of environmental sample data on sediments, wet deposition, surface water, ambient air, and indoor air. It reduces the large number of variables in complex analytical data sets to combinations of species called source types and source contributions. The source types are identified by comparing them to measured profiles. Source contributions are used to determine how much each source contributed to a sample. Algorithms used in the various PMF versions have been peer reviewed by leading air and water quality management scientists.

Project Goals

The goals of this project are:

- to evaluate whether the PMF source apportionment techniques used in other watersheds can be applied to the Spokane River ambient water data despite the blank issues and
- to determine how the blank contamination issue should be handled in the second phase of this project.

Proposed Approach

Data sets will be analyzed in four main ways:

- No blank correction.
- Censor (exclude) concentrations of peaks that were present in the blanks.
- Subtract blank masses from sample masses.
- Exclude specific congeners that are often present in the blanks.

The resulting data sets will be analyzed via PMF2 software and the results compared to identify which approach(es) provide the most useful information about PCB sources to the Spokane River.

Background

It has been proposed to conduct factor analysis on PCB data from the Spokane River watershed using PMF2 in order to characterize the sources of PCBs to the watershed. This approach involving the PMF2 software has been used successfully in many watersheds, including the Green-Duwamish River (Rodenburg and Leidos 2017, Rodenburg and Leidos 2017), the Delaware River (Du, Belton et al. 2008, Du, Belton et al. 2008, Praipipat, Rodenburg et al. 2013), the Portland Harbor Superfund Site (Rodenburg, Krumins et al. 2015), and the NY-NJ Harbor (Rodenburg, Du et al. 2011, Rodenburg and Ralston 2017). In all of these previous cases, the concentrations of PCBs in the affected water bodies were generally greater than 1,000 pg/L and blank masses were therefore negligible. In the Spokane River, by contrast, the sum of 209 PCB congeners is often on the order of 100 pg/L with roughly 30 pg/L found in the blanks. Thus the blank may constitute as much as 30% of the PCB mass in the samples

EPA guidelines suggest that blank subtraction is not recommended. For example, EPA method 8270D (EPA 2017) for measurement of semivolatiles by gas chromatography/mass spectrometry states that "The laboratory should not subtract the results of the [method blank] from those of any associated samples. Such "blank subtraction" may lead to negative sample results." However, method 1668 (USEPA 2003) notes that "The recommended procedure for blank correction (Reference 20) is that a result is significantly above the blank level, and the level in the blank may be subtracted, if the result is greater than the mean plus 2 standard deviations of results of analyses of 10 or more blanks for a sample medium." Reference 20 is a peer-reviewed paper (Ferrario, Byrne et al. 1997) that specifically discusses the difficulties associated with obtaining low or zero concentrations of PCBs in blanks when using high resolution mass spectrometry. These authors note that blank contamination frequently consists of Aroclor-type congeners. However, in our experience, non-Aroclor congeners are frequently abundant in blanks. It is important to note that most of the guidance provided by EPA on this issue concerns the use of data for assessing absolute concentrations, i.e. whether they are above detection or above an applicable limit, such as a water quality standard. To our knowledge, there is no specific guidance from EPA concerning blank correction in scenarios involving fingerprinting or source apportionment.

In our previous experience, we have faced two situations in which blank contamination was problematic. The first case concerned PCBs in the effluent of the Spokane County Regional Water Reclamation Facility (SCRWRF). This data set included measurements of PCBs in both the influent and effluent of the plant. PCB concentrations in the influent were high enough that blank masses (median 130 pg/L) were negligible, but median PCB concentrations in the effluent were about 200 pg/L, such that blank contamination was a significant issue for the effluent. Notably, the non-Aroclor congener PCB 11 was the most abundant congener in the majority of blanks. This is problematic because one of the main issues to be investigated in the Spokane River is the extent to which PCB sources are associated with Aroclors versus non-Aroclor sources. It will be difficult to determine the true impact of PCB 11 if is it abundant in the blanks.

In consultation with the SCRWRF, we decided to blank correct the data by subtracting the average concentration of each congener across all blanks (field, lab, rinsate) collected for each sampling event. The results of this blank subtraction were not noticeable when the influent and effluent were analyzed together in a combined data set, probably because the concentrations in the influent were so much higher and 'swamped' the effluent, dominating the resolved source profiles. However, when the effluent was analyzed separately via PMF2, a factor was generated that contained high proportions of PCBs 44+47+65, 45+51, and 68, and it became clear that one sample of effluent was dominated by these three peaks, which are known to be associated with silicone rubber. This sample was then discarded from further analysis, and the final PMF2 solution for the effluent contained four factors that resembled the four main Aroclor formulations (1016, 1248, 1254, and 1260). It is not clear whether the discarded sample reflected a real PCB source in the sewage of Spokane County, or if it became contaminated during sampling, handling, or analysis.

The second scenario in which we have faced the issue of blank contamination concerned the ambient water data from the Green River, which flows into the Duwamish River in Washington State (Rodenburg and Leidos 2017). The Green River is relatively remote and therefore has relatively low PCB concentrations. The sum of 209 PCB concentrations in these samples ranged from 5 to 450 pg/L, but the concentrations of PCBs 44+47+65 plus 45+51 plus 68 made up between 0% and 91% of the PCBs in the samples. It was subsequently confirmed that silicone rubber tubing had been used for sample collection (Greyell and Williston 2018). Blank data was not available, but the concentrations of these three peaks in the samples ranged from non-detect to 270 pg/L. Taking an average of this wide range of values and subtracting it from each sample would not have solved the contamination problem. Instead, it would have resulted in roughly half of the samples continuing to display high concentrations of congeners associated with contamination. Therefore it was decided to exclude these three peaks from the PMF2 analysis. Due to the large number of non-detects, only 42 PCB peaks representing 69 congeners were included in the final PMF2 model runs. These peaks contained only about 60% of the total PCB mass detected across all 209 congeners. Of the 'missing' 40% of mass, about 15% was explained by the three peaks that were excluded due to silicone rubber contamination. Thus the final model explained about 75% of the PCB mass detected in the samples. The results yielded four factors which resembled the four main Aroclors, although for the factor that was most similar to Aroclor 1248, the correlation coefficient between the congener patterns of the Aroclor and the factor was just 0.44.

Taken together, EPA guidance and our experience lead to several conclusions regarding fingerprinting of PCB data via PMF2 analysis:

- It is important to have blank data available for examination.
- Blank subtraction has in the past resulted in data sets in which factor analysis identified Aroclors.
- Absence of blank correction can sometimes lead to factors that are presumed to resemble the blank contamination.
- It is difficult to perform factor analysis on data sets with low concentrations not only because of blank contamination issues, but also because many of the congeners have to be excluded from the PMF model because they are not detected in enough samples (even when they are not detected in the blanks).
- Non-Aroclor congeners are often prevalent in blanks, making a determination of their true levels in the sample difficult.

Based on these observations, we propose the following approaches to deal with blank contamination issues in the Spokane River ambient water data. We propose to analyze several permutations of the ambient water data set with the following modifications:

- No blank correction. Interpret the output of the model with the assumption that one or more of the resolved factors may represent blank contamination.
- Censor (exclude) concentrations of peaks that were present in the blanks. Censoring may be done at several levels, such as censoring concentrations that are within 3x the blank level, 5x the blank level, and 10x the blank level. Censored concentrations would be treated as non-detects and therefore assigned a higher uncertainty in the PMF2 model. This approach may result in a very high number of congeners excluded from the PMF2 model due to lack of detected concentrations. Dr. Rodenburg will work with SRRTTF to determine which blanks should be used to determine average blank masses.
- Subtract blank masses from sample masses. The blank masses to be subtracted may be calculated as the average of all blanks in a sampling campaign, all blanks in a sampling event, only field blanks, only equipment blanks, etc. We will consult with the SRRTTF for guidance on how to perform blank subtraction. When such a subtraction results in a zero or negative concentration, the data point will be treated as a non-detect. Congeners that are non-detect in a majority of samples are typically excluded from the PMF2 analysis, so this might result in the exclusion of a large number of congeners from the PMF2 model.
- Exclude from the PMF2 analysis specific congeners that are often present in the blanks, as, for example, in the Green River ambient water data in which PCBs 68, 44+47+65, and 45+51 were excluded. This may underestimate the importance of non-Aroclor sources in the Spokane River.

We propose to analyze Spokane River ambient water column data collected by the SRRTTF by the above approaches and compare the various model outputs to determine which approach (or combination of approaches) yields the most useful information about PCB sources to the river. We propose to conduct this study as soon as possible, with the intent of determining the best approach to address blank contamination for the second phase of this project.

For this scope of work blanks include laboratory method blanks, travel blanks, and field blanks. The Spokane River ambient water data set includes Spokane River water column samples collected in conjunction with the SRRTTF 2014 synoptic sampling, 2015 synoptic sampling and the 2016 monthly sampling. The water column data and blank data will be provided to Dr. Rodenburg in an Access Database by the SRRTTF. The specified blank correction and other associated data processing and manipulation will be done by Dr. Rodenburg. Dr. Rodenburg will submit the input data sets to the SRRTTF for approval before conducting the PMF2 analysis.

Deliverables

The following data and reports will be provided:

- Spreadsheets of PMF input and output for all data sets analyzed.
- A report interpreting the results and making a recommendation regarding the best approach to dealing with blank contamination in future factor analysis work to be conducted in the Spokane River.

Budget

Dr. Rodenburg will bill her time at a rate of $200 per hour.

PMF analysis: 50 hours

Write final report: 15 hours

Total = 65 hours @ $200 per hour = $13,000

References

Du, S., T. J. Belton and L. A. Rodenburg (2008). "Source Apportionment of PCBs in the Tidal Delaware River." Environ. Sci. Technol. **42**: 4044–4051.

Du, S., T. J. Belton and L. A. Rodenburg (2008). "Source apportionment of polychlorinated biphenyls in the tidal Delaware River." Environmental Science & Technology **42**: 4044-4051.

EPA, U. S. (2017). Method 8270e: Semivolatile Organic Compounds By Gas Chromatography/Mass Spectrometry.

Ferrario, J., C. Byrne and A. E. Dupuy (1997). "Background contamination by coplanar polychlorinated biphenyls (PCBs) in trace level high resolution gas chromatography high resolution mass spectrometry (HRGC/HRMS) analytical procedures." Chemosphere **34**(11): 2451-2465.

Greyell, C. and D. Williston (2018). Green River PCB Equipment Blank Study Data Rep. Seattle, WA, King County Water and Land Resources Division.

Paatero, P. and U. Tapper (1994). "Positive Matrix Factorization - a Nonnegative Factor Model with Optimal Utilization of Error-Estimates of Data Values." Environmetrics **5**: 111-126.

Praipipat, P., L. A. Rodenburg and G. J. Cavallo (2013). "Source Apportionment of Polychlorinated Biphenyls in the Sediments of the Delaware River." Environ. Sci. Technol. **47**: 4277−4283.

Rodenburg, L. A., S. Du, B. Xiao and D. E. Fennell (2011). "Source Apportionment of Polychlorinated Biphenyls in the New York/New Jersey Harbor." Chemosphere **83**: 792–798.

Rodenburg, L. A., V. Krumins and J. C. Curran (2015). "Microbial Dechlorination of Polychlorinated Biphenyls, Dibenzo-p-dioxins, and -furans at the Portland Harbor Superfund Site, Oregon, USA." Environmental Science & Technology **49**(12): 7227-7235.

Rodenburg, L. A. and Leidos (2017). Green-Duwamish River Watershed PCB Congener Study: Phase 2 Source Evaluation. Seattle, WA.

Rodenburg, L. A. and Leidos (2017). Green-Duwamish River Watershed PCB Congener Study: Phase 2. Initial Data Assessment. Bothell, Washington, Department of Ecology, State of Washington,.

Rodenburg, L. A. and D. K. Ralston (2017). "Historical sources of polychlorinated biphenyls to the sediment of the New York/New Jersey Harbor." Chemosphere **169**: 450-459.

USEPA (2003). Method 1668, Revision A. Chlorinated Biphenyl Congeners in Water, Soil, Sediment, and Tissue by HRGC/HRMS (EPA 821-R-00-002), United States Environmental Protection Agency.