SFEI AQUATIC SCIENCE CENTER
4911 Central Ave, Richmond, CA 94804, p: (510) 746-7334, f: (510) 746-7300 www.sfei.org

To: Melissa Foley, RMP Manager
Jay Davis, RMP Lead Scientist
From: Don Yee, Quality Assurance Officer
Date: September 18, 2019
Re: 2018 RMP Tissue Data Quality Assurance Report

Introduction

In 2018, bivalve tissue samples were collected from six Bay/Delta stations and a reference site for the Regional Monitoring Program for Water Quality in San Francisco Bay. Bird egg tissue samples were collected from two sites for cormorants, and four sites for terns. General descriptions of the sample collection methods are provided in the RMP Quality Assurance Program Plan, cruise plans, cruise reports, and sampling reports. These documents are available from the SFEI website (http://www.sfei.org/content/status-and-trends-monitoring-documents).

Bivalve samples were analyzed for the following parameters by the laboratories indicated:

- BAL – selenium
- SGS-AXYS – PAHs

Bird egg samples were analyzed for the following parameters by the laboratories indicated:

- MPSL-DFW – selenium in cormorant and tern eggs, mercury in cormorant eggs
- USGS-WERC – mercury in tern eggs
- SGS-AXYS – PCBs and perfluoronates in cormorant eggs, and PBDEs in tern and cormorant eggs

The SFEI Data Services Team checked the laboratory results using the methods and data quality objectives in the RMP Quality Assurance Project Plan (QAPP). For bivalves, 100% of the BAL field sample results, and 67% of SGS-Axys results were determined to be acceptable for use in RMP reports and calculations. The main cause of data rejection for the PAH data was blank contamination at concentrations close to (>33% of) concentrations found in field samples. For bird eggs, 100% of the selenium and mercury data form MPSL-DFW and USGS-WERC were reportable. Of the organic compounds reported by SGS-AXYS, 100% of PFAS results, 95% of PCBs, and 76% of PBDEs were reportable, with the main cause of censoring being blank concentrations close to those in field samples for some of the analytes.

This memo provides a high-level summary of the quality assurance assessment for each dataset. Non-conformances with the QAPP and possible indicators of variability and uncertainty in reported values, with corrective actions needed for the next round of monitoring are highlighted in gray shading. The details of the quality assurance assessment of each dataset are provided in Appendix A for bivalves, and Appendix B for bird eggs.

The data have been approved by the RMP Manager and Lead Scientist, and all results have been uploaded to the San Francisco Regional Data Center and CEDEN.

Quality Assurance Summary for 2018 RMP Bivalve Samples

Selenium and PAHs in bivalves were reported for 6 RMP historical sites (4 transplant sites, and 2 with resident bivalves), and one reference/source site for the transplanted bivalves. The selenium data had no issues, with detections in all samples, no detections in method blanks, and recovery errors in matrix spikes and CRMs averaging <5%, and likewise precision <5% RPD on the lab replicate. For PAHs, results were less consistent, with about 1/3 of the PAH analytes having non-detects in >50% of the samples. About half of the analytes were also detected in the method blank, which combined with the non-detects or low concentrations for many of the PAHs in field samples, resulted in about 1/3 of all the PAH data being censored (for being <3x over the blanks). C4-Fluoranthenes/Pyrenes (an alkylated PAH group) was also censored for very poor precision, with an RSD on lab replicates of 97%. Fluorene and 2-Methylanthracene had marginal recovery (around 50-55% error), over the target 35%, but not enough to censor their results. *The extensive non-detects in field samples, and detections of many target analytes in the lab blank limit the usability of the PAH data; although the data may still be useful for assurance that concentrations are below those likely to cause impacts (e.g., if the concentrations, even including blank contributions, are below effects thresholds), they may be less useful or usable for detecting differences among locations or in seeing trends over time. A reduction in the number of target analytes to a subset that is usually reliably quantified (i.e., usually detected in field samples, and with no or negligible blank contamination) might be beneficial, as the poorly quantified analytes may generally not provide useful information. Raising detection or reporting limits to levels above typical random blank noise may make the data appear somewhat cleaner and more consistent. However, the outcome would be similar; a larger percentage of the data would simply be reported as non-detects (as opposed to being censored for blank contamination), but in either case, the results are not useful for quantitative comparison.*

Quality Assurance Summary for 2018 RMP Bird Egg Samples

Selenium, PFASs, PBDEs, and PCBs were reported for composites of cormorant and tern eggs, and mercury for individual tern eggs or composited cormorants. No issues were encountered in the analysis of selenium or mercury in cormorant or tern eggs analyzed by MPSL-DFW. There were no non-detects for selenium or mercury in any egg composites, and they were not detected in the method blanks. Recovery errors for selenium and mercury in a CRM (NRC DORM-4) or in matrix spikes, averaged < 10%, well within the target 35%, and required no flags. Precision in lab replicates and matrix spike replicates averaged 3% RPD for selenium and 5% for mercury, also well within the target of < 35% and needing no flags. Average concentrations are somewhat lower than in past years (about 55-65% of past averages for selenium, 87% for mercury), but not enough different to raise immediate concerns about comparability of the analytical methods over time.

Mercury analyses were done for 84 individual tern eggs by USGS-WERC. There were no non-detects for mercury in field samples. Mercury was detected in the blank for one batch, but concentrations in field samples were well over 3x higher, so results were flagged but none censored. Recovery of mercury in CRMs averaged within 2% of the target values, well within the 35% requirement. Precision on lab control samples replicates was also <1% RSD, within the 35% target and not requiring any flags. Average concentrations were similar to the past, about 93% of prior years' (2002-2016) averages.

Organic compounds reported in bird egg composites by SGS-Axys had some issues with quality for some analyte groups, but were mostly reportable. Non-detects (NDs) were prevalent for these organic groups,

with 6 of 13 PFAS analytes, 36 of the 209 PCBs, and 11 of 48 reported PBDEs 100% ND. None of the PFASs were detected in the method blanks. In contrast, about 1/3 of the PCB congeners were detected in one or more method blanks. 8 PCBs, mostly lighter ones (PCB 56 and lower) were 100% censored (VRIP flag) for field sample results being <3x the blank. Some PBDE congeners were detected in one or more blanks, with some sample results censored (VRIP flag) for being <3x higher than the blank average in their batches for PBDEs 37, 138, and 208. Recoveries in CRM or matrix spike samples were acceptable with <35% erro for most target analytes, >35% for PCBs 49, 87, and 138, which were flagged but not censored, and PBDE 17 with 29% recovery (71% error) censored (VRIU flag) for poor recovery. Lab replicate RPDs were 6% or better for the PFASs detected in lab replicates, requiring no flags. A few PCBs (1, 3, 7) had marginal precision (RSDs 35-70%), and only PCB 42 was censored (RSD 200%). Ten of the lower abundance PBDE congeners similarly had RSDs >70% and thus were censored (VRIL flagged) for poor precision. Average concentration for PFAS compounds were generally lower than (20-70% of) their previous averages (for 2006, 2009, and 2016 combined) with only Perfluorododecanoate higher at 106%. PCBs averaged 1.1-1.5x their past averages (2009-2016). For the most common congeners (47, 49, 99, 100, 153, 154), PBDEs in cormorants were within 70-130% of their averages in the prior 2 sampling events, but for terns average concentrations were only about 40-50% of the prior 2 events. *Similar to the case for PAHs is bivalves, the most issues were found with less abundant compounds, where the quantitation is less certain. However, raising detection and reporting limits would mask but not fix the underlying uncertainty in quantitative measurement in the low concentration range, as the flagged censored values would largely then just be non-detects.*

Most of these organic compounds appear to be on a downward trajectory in concentrations. It may be difficult to drop measurement of PCBs due to continued bioaccumulation in fish exceeding levels of concern and TMDL information needs, but as concentrations of other compound groups drop below toxicity thresholds of concern, a reduction in the frequency of measurement or in the number of target analytes (e.g. reducing to a few most abundant compounds for use as a general index of the group) may be warranted.

Appendix A: Dataset QA Summaries
Bay RMP 2018 Bivalves

BA

Bivalves

Selenium, Total Solids, and Moisture

QA Issues for Project Manager to Review

Overall acceptability

Overall the dataset is acceptable. 100% of the results are reportable.

Accuracy

The accuracy for mercury was flagged following the SFEI RMP Status and Trends convention of using the average percent error of the certified material samples (CRMs), when present, in preference to the percent error of the matrix spike/matrix spike replicates or the percent error of the laboratory control sample, as the CRMs are externally validated values.

The selenium certified reference material (CRM) result meet the method quality objective listed in the 2018 RMP QAPP of "expected value ± 35%", and the average percent error of about 2% was less than a RPD of 35%; no accuracy flags are indicated.

The selenium matrix spike/matrix spike duplicate results meet the method quality objective listed in the 2018 RMP QAPP of "expected value ± 35%, and the average percent error of about 4% was < 35%.

The selenium laboratory control sample result meet the method quality objective listed in the 2018 RMP QAPP of "expected value ± 35%, and the average percent error of about 3% was < 35%; accuracy was sufficient based on the CRM results.

The accuracy of the moisture and total solids data could not be determined as no spiked samples of any kind were reported/analyzed.

Precision

The precision of field samples in the database is flagged following the SFEI RMP Status and Trends convention of using lab replicates in preference to using field replicates, although both are reviewed and described narratively when provided. No field replicates were included in this data submission.

The single selenium lab replicate met the method quality objective listed in the 2018 RMP QAPP of an RPD < 35% so the set of field samples was not flagged in the database.

The single moisture and total solids replicate met the method quality objective listed in the 2018 RMP QAPP of an RPD < 20%, therefore, no field samples were flagged in the database.

Reporting Issues for Lab to Review

Formatting Issues for Data Manager to Review

Hold time review (especially desired by stormwater programs)
Selenium samples were analyzed between 148 and 256 days after collection. This is well within the 1 year holding time specified in the 2018 RMP QAPP. No holding time requirements listed for either Moisture or Total Solids.

QA Review
Selenium results were reported for 7 bivalve tissue samples (5 *Mytilus californianus* and 2 *Corbicula fluminea*) analyzed in 1 lab batch. One lab replicate (*Mytilus californianus*) and 1 matrix spike/matrix spike replicate (*Mytilus californianus*) were analyzed for the 7 bivalve tissue samples meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for those sample types.

Four method blanks were analyzed meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for those sample types. One certified reference material samples (NRC DORM-4: Fish protein certified reference material for trace metals), and one laboratory control sample was also analyzed meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for those sample types. Data were reported not blank corrected.

Moisture results were reported for 7 bivalve tissue samples analyzed in 1 lab batch. One lab replicate and two method blanks were also reported meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for those sample types. However, the method blanks reported were not analyzed, but instead determined by means of a calculation (100% - Total Solids% = Moisture%). No moisture was reported for the matrix spike/matrix spike replicate or the certified reference material and laboratory control sample as listed in the QAPP.

Total Solids results were reported for 7 bivalve tissue samples analyzed in 1 lab batch. One lab replicate and two method blanks were also reported meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for those sample types.

Overall acceptability
Overall the data submission is acceptable. 100% of the results are reportable.

MDLs sensitivity
The lab reported results above the detection limit for all bivalve composite tissue samples for selenium, moisture, and total solids. This indicates that the analysis methods used were of sufficient sensitivity to detect concentrations found in the bivalve composites.

QB averages (procedural, field blank)

Selenium was not measured in the method blanks at concentrations equal to, or above the method detection limit (MDL), meeting the method objective of the 2018 RMP QAPP of being "<MDL". No qualifiers were added.

Total Solids was found in the method blanks at concentrations >MDL. All total solids results were flagged with the non-censoring qualifier "VIP" (Analyte detected in field or lab generated blank, flagged by QAO).

Calculated method blank moisture results were above the reported MDL value so all moisture results were flagged with the "VIP" qualifier like the Total Solids results.

Accuracy (using a variety of SRMs or Matrix spike QRECs)
For selenium samples with a known concentration, consisting of certified reference material (CRM), were run at a minimum frequency of one per analytical batch (for analytical batches consisting of up to 20 field samples) or per 20 (field) samples for larger analytical batches. Analysis of CRMs allows us to evaluate measurement accuracy, or how close our measurement comes to a consensus/expected value. Matrix spikes, where an environmental sample is "spiked" with a known amount of mercury, provide an alternative determination of method accuracy that can account for matrix interferences or other analytical problems. Laboratory control samples are an aliquot of a reference matrix fortified (spiked) with known quantities of specific compounds and subjected to the entire analytical procedure to determine the accuracy of the method by measuring recovery; they are not externally validated values.

The average percent error for the certified reference material samples for selenium of 2.1% (average recovery 102.1%) was well below the target MQO of 35%. No qualifiers were added. The average percent error examined for the selenium matrix spikes was 3.9% (average recovery 103.9%), and for the laboratory control sample it was 2.5% (average recovery 102.5), both within the target 35% error.

The accuracy of the moisture and total solids data could not be determined as no spiked samples of any kind were reported/analyzed.

Average precision from replicate field sample
The precision of analysis methods (ability to consistently obtain the same result) is determined by analyzing replicate or duplicate samples. The analysis of lab replicates (split and analyzed in the laboratory) allows us to assess the repeatability of lab measurements.

Lab replicates were used to decide whether precision flags were needed. The average RPD for the single selenium lab replicate was 3%, well below the MQO target of 35%. No field replicates were analyzed. No qualifiers were added.

Matrix spike replicates were examined, but not used for flagging the dataset. The average RPD for the selenium matrix spike replicate was 1%, below the MQO target of 35%.

The precision of the moisture and total solids results was evaluated using the single laboratory replicate. The average RPD of 0.02% for moisture and 0.08% for total solids were both below the MQO target of 20%.

Comparison of dissolved and total phases
Not Applicable

Comparison to previous years
Average concentration for total selenium in bivalves was 104% of the previous average of the 1993-2016 RMP Status and Trends samples (in units ug/g dw). Average concentration for total selenium in the *Mytilus californianus* was 98% and for the *Corbicula fluminea* it was 122% of the 1993-2016 RMP Status and Trends samples.

Ratio checking summary (if applicable)
Not applicable.

SGS AXYS

Bivalves

PAH

QA Issues for Project Manager to Review

Overall acceptability

Overall the dataset is acceptable. 67% of the results are reportable.

Accuracy

The accuracy for PAH and alkylated PAH was flagged following the SFEI RMP Status and Trends convention of using the average percent error of the certified material samples (CRMs), when present, in preference to the percent error of the matrix spike/matrix spike replicates or the percent error of the laboratory control sample, as the CRMs are externally validated values.

The majority of PAH and alkylated PAH certified reference material (CRM) results did meet the method quality objective listed in the 2018 RMP QAPP of “expected value ± 35%”. The average percent errors were generally less than a RPD of 35%, except for Fluorene and 2-Methylanthracene with average percent errors of 48.05% and 54.78%, respectively. All Fluorene and 2-Methylanthracene results were flagged with the non-censoring qualifier “VIU”.

The PAH and alkylated PAH matrix spike results meet the method quality objective listed in the 2018 RMP QAPP of “expected value ± 35%, and the average percent error ranged from ranged from 0% to 23.63% all < 35%.

The PAH and alkylated PAH laboratory control sample results meet the method quality objective listed in the 2018 RMP QAPP of “expected value ± 35%, and the average percent error ranged from 0% to 31.5% all < 35%.

The accuracy of the lipid and moisture and data could not be determined.

Precision

The precision of field samples in the database is flagged following the SFEI RMP Status and Trends convention of using lab replicates in preference to using field replicates, although both are reviewed and described narratively when provided. No field replicates were included in this data submission.

The PAH and alkylated PAH lab replicate average RSD's ranged from 0% to 96.99%, but was below the MQO target of 35%, except for Dibenzothiophenes, C2- (51.3%), and Fluoranthenes/Pyrenes, C4- (96.99%). All the Dibenzothiophenes, C2- results were flagged with the non-censoring qualifier "VIL". While all the Fluoranthenes/Pyrenes, C4- results were flagged with the censoring qualifier "VRIL".

The moisture replicate met the method quality objective listed in the 2018 RMP QAPP of an RSD < 20%, therefore, no field samples were flagged in the database. The lipid result could not be evaluated.

Reporting Issues for Lab to Review

Formatting Issues for Data Manager to Review

Hold time review (especially desired by stormwater programs)

PAH and alkylated PAH samples were analyzed between 133 and 242 days after collection well within the 1 year holding time specified in the 2018 RMP QAPP. No holding time requirements specified for lipid and moisture.

QA Review

PAH and alkylated PAH results were reported for 7 bivalve tissue samples (5 *Mytilus californianus* and 2 *Corbicula fluminea*) analyzed in 1 lab batch.

One lab replicate (*Mytilus californianus*) and 1 matrix spike (*Mytilus californianus*) were analyzed for the 7 bivalve tissue samples meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for those sample types.

One method blank was analyzed meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for those sample types. One certified reference material samples (NIST 1974c: Organics in Mussel Tissue (*Mytilus edulis*)), and one laboratory control sample was also analyzed for a subset of the analytes meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for those sample types. Data were reported not blank corrected.

Lipid results were reported for 7 bivalve tissue samples analyzed in 1 lab batch. No lipid was reported for the lab replicate, matrix spike, certified reference material and laboratory control

sample, therefore, failing to meet the minimum requirement in the RMP QAPP of 1 per 20 samples, or 1 per batch for those sample types.

Moisture results were reported for 7 bivalve tissue samples analyzed in 1 lab batch. One lab replicate, and matrix spike sample moisture were reported meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for those sample types. Moisture was not reported for the certified reference material, and laboratory control sample failing to meet the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for those sample types.

Overall acceptability
Overall the data submission is acceptable. 67% of the results are reportable

MDLs sensitivity
The lab reported results above the detection limit in the *Corbicula fluminea* composite tissue samples for 82% (64 out of 78) of the lipid, moisture, PAH and alkylated PAH analytes. Extensive non-detects (> 50% NDs) were reported for 64% (9 out of 14) of the PAH and alkylated PAH analytes (Acenaphthenes, C1-, Benz(a)anthracene, Dimethylfluorene, 1,7-, Dimethylnaphthalene, 1,2-, Dimethylphenanthrene, 3,6-, Methylanthracene, 2-, Methylphenanthrene, 9/4-, Tetramethylnaphthalene, 1,4,6,7-, and Trimethylphenanthrene, 1,2,6-). Five PAH and alkylated PAH analytes had 50% non-detects (Benz(a)anthracenes/Chrysenes, C4-, Dimethyldibenzothiophene, 2,4-, Dimethylphenanthrene, 1,8-, Dimethylphenanthrene, 2,6-, and Methylfluorene, 2-).

The lab reported results above the detection limit in the *Mytilus californianus* composite tissue samples for 51% (40 out of 78) of the lipid, moisture, PAH and alkylated PAH analytes. Extensive non-detects (> 50% NDs) were reported for 37% (14 out of 38) of the PAH and alkylated PAH analytes (Benz(a)anthracene, Benz(a)anthracenes/Chrysenes, C3-, Benz(a)anthracenes/Chrysenes, C4-, Benzo(e)pyrene, Dimethyldibenzothiophene, 2,4-, Dimethylfluorene, 1,7-, Dimethylphenanthrene, 1,7-, Dimethylphenanthrene, 1,8-, Dimethylphenanthrene, 3,6-, Methylanthracene, 2-, Methylbenzo(a)pyrene, 7-, Methylphenanthrene, 9/4-, Tetramethylnaphthalene, 1,4,6,7-, and Trimethylphenanthrene, 1,2,6-). Twenty-four PAH and alkylated PAH analytes had 17% to 50% non-detects.

QB averages (procedural, field blank)
A majority of the lipid, moisture, PAH and alkylated PAH analytes (54%) were measured in the method blank at concentrations equal to, or above the method detection limit (MDL), failing to meet the method objective of the 2018 RMP QAPP of being "<MDL". Results for 46 of the PAH and alkylated PAH analytes were flagged for some degree of blank contamination; some of the results for 38 of the analytes were qualified with the censoring "VRIP" (Data rejected - Analyte detected in field or lab generated blank, flagged by QAO).

Lipid and moisture were not reported/analyzed in the method blank.

Accuracy (using a variety of SRMs or Matrix spike QRECs)

For samples with a known concentration, consisting of certified reference material (CRM), were run at a minimum frequency of one per analytical batch (for analytical batches consisting of up to 20 field samples) or per 20 (field) samples for larger analytical batches. Analysis of CRMs allows us to evaluate measurement accuracy, or how close our measurement comes to a consensus/expected value. Matrix spikes, where an environmental sample is "spiked" with a known amount of mercury, provide an alternative determination of method accuracy that can account for matrix interferences or other analytical problems. Laboratory control samples are an aliquot of a reference matrix fortified (spiked) with known quantities of specific compounds and subjected to the entire analytical procedure to determine the accuracy of the method by measuring recovery; they are not externally validated values.

Only a subset of the PAH and alkylated PAH analytes were analyzed in the certified reference material (NIST 1974c). The percent errors for the analytes in the CRM with certified values were generally below the target MQO of 35%, the exceptions were Fluorene and Methylanthracene, 2-, with average percent errors of 48.05% (average recovery 51.95%) and 54.78% (average recovery 45.22%), respectively. All the Fluorene and Methylanthracene, 2- results were flagged with the non-censoring qualifier "VIU" (Percent Recovery exceeds laboratory control limit, flagged by QAO).

The average percent error examined for the subset of PAH and alkylated PAH analytes in the matrix spike ranged from 0% (average recovery 100%) to 23.63% (average recovery 123.63%), and for the laboratory control sample it ranged from 0% (average recovery 100%) to 31.5% (average recovery 68.5%), both within the target 35% error.

The accuracy of the lipid and moisture and data could not be determined.

Average precision from replicate field sample
The precision of analysis methods (ability to consistently obtain the same result) is determined by analyzing replicate or duplicate samples. The analysis of lab replicates (split and analyzed in the laboratory) allows us to assess the repeatability of lab measurements.

Lab replicates were used to decide whether precision flags were needed. The average RSD for the subset of PAH and alkylated PAH analytes in the lab replicate ranged from 0% to 96.99%, but was below the MQO target of 35%, except for Dibenzothiophenes, C2- (51.3%), and Fluoranthenes/Pyrenes, C4- (96.99%).

All the Dibenzothiophenes, C2- results were flagged with the non-censoring qualifier "VIL" (RPD exceeds control limit, flagged by QAO). While all the Fluoranthenes/Pyrenes, C4- results were flagged with the censoring qualifier "VRIL" (Data rejected - RPD exceeds control limit, flagged by QAO).

No field replicates were analyzed.

The precision of the moisture results was evaluated using the single laboratory replicate. The average RSD of 0.33% was below the MQO target of 20%. The lipid result could not be evaluated as it was not reported/analyzed in the lab replicate.

Comparison of dissolved and total phases
Not Applicable

Comparison to previous years
Average concentration for the PAH and alkylated PAHs in bivalves ranged from 0% to 148% of the previous average of the 1993-2016 RMP Status and Trends samples (in units ug/g dw), where they could be compared. The new analyte list contains 31 out of 76 PAH and alkylated PAH analytes not previously analyzed/reported.

Ratio checking summary (if applicable)

2018 bivalve PAH and alkyl PAH data appear reasonable overall. No changes suggested.

PAH Detailed Narrative:
Replicate samples (BC10) showed excellent agreement, with minor exceptions for methylchrysene, 1- and 5/6-.

As noted in the 2016 dataset, there were unusually high method detection limits (MDLs) for benz(a)anthracene (avg MDL 8) and somewhat high MDLs for benzo(e)pyrene (avg MDL 2.7). For benz(a)anthracene, the highest average ratio observed in previous datasets is 6%, suggesting this should not invalidate ratios or sums significantly. For benzo(e)pyrene, the ratios where detected were also ~6%, again suggesting no major impact to sums.

Means and standard deviations of ratios were calculated excluding T-0 (reference site, patterns not typical of Bay). Apart from expected ratio deviations in T-0, the only values that (just) exceeded two standard deviations above the mean were Dimethyldibenzothiophene, 2,4- and Dimethylphenanthrene, 1,8- in BG30, both with ratios <1%.

HPAHs were proportionally higher in BA10 and BA40 (Lower South and South Bay) and BG30 (San Joaquin River). This may suggest a difference in PAH sources. Higher proportions of LPAHs, along with higher proportions of alkyl relative to parent compounds, is indicative of a petrogenic PAH source; higher proportions of HPAHs is indicative of a pyrogenic source.

In contrast to 2016 data, samples from BC10 did not show unusually high levels of contamination. The T-0 reference site showed very low levels of contamination, as expected.

Alkyl PAH Detailed Narrative:
Replicate samples (T-0) showed good agreement. There is one major exception, for C2-naphthalenes, where one of the sample values was rejected due to blank contamination. (The other value was 9.8 ng/g, and flagged with GN and JA.)

T-0 alkyl PAH values, where available, were generally low, as would be expected for this reference site. C2- and C3-fluorenes at this site were exceptions, comparable to other sites.

Patterns in alkyl vs parent compound proportions are generally unclear as to potential sources of contamination (pyrogenic vs. petrogenic).

Appendix B: Dataset QA Summaries
Bay RMP 2018 Bird Eggs

USGS-WERC

Tern Eggs

Hg and Moisture

QA Issues for Project Manager to Review

Results for the mercury composites were in units of ug/g fww while the units for the equipment and method blanks were in ug/g ww. The difference between fww and ww is usually small, consequently the blank contamination check will be slightly too conservative (ww blank concentration > fww equivalent).

Reporting Issues for Lab to Review

Formatting Issues for Data Manager to Review

Hold time review (especially desired by stormwater programs)

Total mercury samples were analyzed between 65 and 128 days after collection within the 1 year holding time specified in the 2018 RMP QAPP. No holding time issues were flagged.

QA Review

Dataset completeness

Total Mercury and Moisture results were reported for 84 composite tissue samples analyzed in 1 lab batch.

One non-project sample and its lab replicate, 9 project lab replicates, 5 matrix spike/matrix spike duplicates, 10 method blanks, and 13 certified reference materials (CRMs) were analyzed for the 84 composite samples which meets the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples for those sample types. Ten continuing calibration verification samples and 10 equipment blanks were also reported. Data were reported not blank corrected.

Overall acceptability

Overall the data submission is acceptable. 100% of the results are reportable.

MDLs sensitivity

No non-detects (NDs) were reported for moisture or mercury results.

QB averages (procedural, field blank)

Mercury was found in the method blanks of the single lab batch at concentrations above the method detection limit. As a result all mercury results were flagged with the non-censoring qualifier “VIP”.

Mercury was also found in the equipment blanks at concentrations above the method detection limit.

Accuracy (using a variety of SRMs or Matrix spike QRECs)
Accuracy of the total mercury results was evaluated using the certified reference material samples. The average %error of ~2% (average recovery 100.95%) was well below the 35% target MQO. No qualifiers were needed.

Average precision from replicate field sample
Precision was evaluated using the laboratory control sample replicates. The average RSD for total mercury was ~1% well below the 35% MQO target. No qualifiers were added.

Comparison of dissolved and total phases
Not Applicable

Comparison to previous years
Average concentration for total mercury in *Forster's Terns* was 93% of (~0.9x) the previous average of the 2002-2016 RMP EEPS samples (in units ug/g fww).

Ratio Checking Summary
Not Applicable

Sums Summary
Not Applicable

AXYS

Cormorant Eggs

PCBs

QA Issues for Project Manager to Review

Reporting Issues for Lab to Review

Formatting Issues for Data Manager to Review
Changed CRM query to look at SFEISampleID rather than SampleID since the latter do not contain SRM or other relevant matching information. Unique SFEISampleID is likely the case needed for other non-field matrix samples, e.g. field or equipment blanks, LCSs, LRMs where there may not be sufficient info in StationCode, SampleDate, etc., to distinguish sets of samples that should have shared vs different expected results. Adam indicated SFEISampleID is inconsistently populated (e.g. often empty), which is fine for regular field matrix samples, but we should establish an SOP for CRM/LCS/LRM, where CRMs get

their CRM name, and LRM/LCS should get a unique (within dataset) ID for every different ExpectedValue case. There are likely cases where field or equipment or other blanks are slightly different, and this field should be populated to distinguish those cases (e.g. tubing vs bottle equipment blanks, travel (unopened) vs opened field blanks).

Hold time review (especially desired by stormwater programs)

All samples were analyzed within 365 days, as required in the 2018 RMP QAPP.
No holding time is listed in the 2018 RMP QAPP for lipid or moisture.

QA Review

Overall acceptability

Overall the dataset is acceptable. 94% of the results are reportable.
Samples were analyzed within the recommended 1 year hold time for PCBs in EPA Method 1668. 6 field samples with one lab replicate were analyzed for lipid, moisture, and PCBs in 2 batches. A lab blank and lab control sample were reported for each PCB batch. An MS/MSD pair, and 1 CRM (NIST 1946) were also reported for the overall effort for PCBs.

There were non-detects (NDs) for 68 of the PCBs, with 36 of them 100% ND . These are generally congeners with very low abundance, rather than being due to very high MDLs. PCBs were detected in the method blank at or above the MDL in at least one batch for about a third of the congeners. 8 PCBs, mostly lighter ones (PCB 56 and lower) were 100% censored (VRIP flag, for sample results <3x the blank), and 3 PCBs had 1 to 6 of their 7 sample results censored. Lipids and moisture were not reported/analyzed for the method blank.

Percent error for the various PCBs in the NIST 1946 CRM ranged up to 88% if uncertified values were included, but 49% was the worst error for certified values, with only 3 congeners (PCBs 49, 87, and 138) over the 35% target. The average percent error for the matrix spike samples for the PCBs were all 23% or better for all congeners that were spiked at least to double the original concentration. Recovery errors were 19% or better for congeners spiked at >3xMDL in LCSs. The lipid and moisture samples are seldom analyzed as spiked or certified values, and none were included in this reporting.

Lab replicates were primarily used to assign whether precision flags. If the 1 sample with a lab replicate was not >3xMDL, the average RSD within each of the sample sites could also be used to estimate an effective field replicate RSD (i.e., each of the composites considered as a field replicate of the other composite from that site). Average RSDs were within the target MQO of 35% for nearly all PCBs detected >3xMDL, with only PCBs 1, 3, 7, having RSDs between 35-70% (non-censoring VIL flag) and only PCB 42 censored for 200% (all flagged VRIL, for being over 70% RSD). The RSDs of the lipid and moisture results in the lab replicates were 6% and 1% respectively, and averaged 11% and 1% for all the composites within each site, so no added flags were needed .

Average total PCB concentrations in the cormorant egg samples were similar to previous years, with results averaging about 1.5x the 2016 results reported in common, but only about 1.1x the

combined average of 2009, 2012, and 2015 results, (for the reported values in 2012 and 2009 including somewhat fewer analytes)

****

Detailed notes below:

Dataset completeness
Results were reported for 209 PCBs (with coeluters) measured in 6 cormorant (*Phalacrocorax auritus*) egg composite samples (from three nests at two stations) analyzed in 1 lab batch, plus a lab replicate.

The 1 lab replicate and 1 matrix spike and matrix spike replicate, and 1 certified reference material samples (NIST 1946: Lake Superior Fish Tissue) were analyzed for PCBs, meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples for those sample types.

One method blank and 1 lab control sample per batch were analyzed meeting the minimum requirement in the 2018 RMP QAPP of 1 per batch (or per 20 samples for larger batches) for these sample types.

Lipid and moisture results were reported for the 6 egg composite samples, with one lab replicate analyzed for each. No other QC sample types were reported for these parameters, and no minimum requirements are listed in the 2018 RMP QAPP.

MDLs sensitivity
Non-detects (NDs) were reported for 68 of the PCBs, with 48 of those over 50% NDs, and 36 of them 100% ND . These are generally congeners with very low abundance, rather than being due to very high MDLs.

QB averages (procedural, field blank)
PCBs were detected in the method blank at or above the MDL in at least one batch for about a third of the congeners. 8 PCBs, mostly lighter ones (PCB 56 and lower) were 100% censored (VRIP flag) for field sample results being <3x the blank, with 3 more congeners having between 1 and 6 of their 7 sample results censored.

Lipids and moisture were not reported/analyzed for the method blank.

Accuracy (using a variety of SRMs or Matrix spike QRECs)
A certified reference material (CRM), was run at a minimum frequency of one per per 20 (field) samples. Analysis of CRMs allows us to evaluate measurement accuracy, or how close our measurement comes to an externally referenced consensus/expected value.

The NIST 1946 CRM has certified values for a subset of PCBs congeners, with others reported as reference values. Percent error for the various PCBs ranged up to 88% (12% recovery) when

uncertified values were included, but the worst case was 49% error for certified values, with only 3 congeners (PCBs 49, 87, and 138) over the 35% target.

Matrix spikes, where an environmental sample is "spiked" with a known amount of mercury, provide an alternative determination of method accuracy that can account for matrix interferences or other analytical problems. Laboratory control samples are an aliquot of a reference matrix fortified (spiked) with known quantities of specific compounds and subjected to the entire analytical procedure to determine the accuracy of the method by measuring recovery; they are not externally validated values.

The average percent error for the matrix spike sample for the PCBs were all 23% or better for all congeners that were spiked at least to double the original concentration. Recovery rrrors were 19% or better for congeners spike at >3xMDL in LCSs.

The lipid and moisture samples are seldom analyzed as spiked or certified values, and none were included in this reporting.

Average precision from replicate field sample
The precision of analysis methods (ability to consistently obtain the same result) is determined by analyzing replicate or duplicate samples. The analysis of lab replicates (split and analyzed in the laboratory) allows us to assess the repeatability of lab measurements.

Lab replicates were primarily used to assign whether precision flags, if results were sufficiently above MDLs (averaging >3xMDL). If the 1 sample with a lab replicate was not sufficiently above detection limits, the average RSD within each of the sample sites could also be used to estimate an effective field replicate RSD (i.e., each of the composites considered as a field replicate of the other composite from that site). Average RSDs were within the target MQO of 35% for nearly all PCBs detected >3xMDL, with only PCBs 1, 3, 7, having RSDs between 35-70% (non-censoring VIL flag) and only PCB 42 censored at 200% (flagged VRIL, being over 70%).

The RSDs of the lipid and moisture results in the lab replicates were 6% and 1% respectively, and averaged 11% and 1% for all the composites within each site, so no added flags were needed .

Comparison of dissolved and total phases
Not Applicable

Comparison to previous years
Average total PCB concentrations in the cormorant egg samples were similar to previous years, with results averaging about 1.5x the 2016 results reported in common, but only about 1.1x the combined average of 2009, 2012, and 2015 results, (for the reported values in 2012 and 2009 including somewhat fewer analytes)

Ratio Checking Summary
The data look good.
The congener ratios look fine, and the total concentrations are consistent with past data.
They're good to go.

Sums Summary
Not Applicable

AXYS

Cormorant Eggs

Perfluoronates

QA Issues for Project Manager to Review

Overall acceptability
Overall the dataset is acceptable. 100% of the results are reportable.

Accuracy
No certified reference material (CRM) sample was required per the 2018 RMP QAPP, therefore, the accuracy for the perfluoronates was flagged following the SFEI RMP Status and Trends convention of using the average percent error of the percent error of the matrix spike/matrix spike replicates in preference to the percent error of the laboratory control sample.

The perfluorooctanesulfonate (PFOS) certified reference material (CRM) result was not used as the expected value is not certified, instead it is a reference value. The result, however, still meet the method quality objective listed in the 2018 RMP QAPP of "expected value ± 35%", and the average percent error of about 15% was less than a RPD of 35%.

The perfluoronate matrix spike results meet the method quality objective listed in the 2018 RMP QAPP of "expected value ± 35%, and the average percent error for the individual perfluoronates were all less than 18% and less than a RPD of 35%; no accuracy flags are indicated. One missing

The perfluoronates laboratory control sample result meet the method quality objective listed in the 2018 RMP QAPP of "expected value ± 35%, and the average percent error for the individual perfluronates were all less than 24% and less than a RPD of 35%; accuracy was sufficient based on the matrix spike results.

The accuracy of the lipid and moisture data could not be determined as no spiked samples of any kind were reported/analyzed.

Precision
The precision of field samples in the database is flagged following the SFEI RMP Status and Trends convention of using lab replicates in preference to using field replicates, although both are reviewed and described narratively when provided. No field replicates were included in this data submission.

The single lab replicate perfluoronate results that could be examined, six perfluoronate results could not be used due to non-detects (NDs), met the method quality objective listed in the 2018 RMP QAPP of an RPD < 35% so the set of field samples was not flagged in the database.

The precision of the lipid and moisture data could not be determined as no replicates of any kind were reported/analyzed.

Reporting Issues for Lab to Review

Formatting Issues for Data Manager to Review

Hold time review (especially desired by stormwater programs)

PFC samples were analyzed between 84 and 86 days after collection. No holding time is available in the 2018 RMP QAPP.

No holding time is listed in the 2018 RMP QAPP for lipid or moisture.

QA Review____

Dataset completeness

Results were reported for 13 perfluoronates (PFCs) measured in 6 cormorant (*Phalacrocorax auritus*) egg composite samples (from three nests at two stations) analyzed in 1 lab batch. The target was 7 homogenized eggs.

One lab replicate and 1 matrix spike were analyzed for the 6 cormorant (*Phalacrocorax auritus*) egg composite samples meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for those sample types.

One method blank was analyzed meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for this sample type. One laboratory control sample was also analyzed meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for this sample type.

Even though it was not required one certified reference material samples (NIST 1947: Lake Michigan Fish Tissue), was analyzed. The result reported in the certified reference material for Perfluorooctanesulfonate (PFOS), however, is not a certified result, but a reference value. All data were reported not blank corrected.

Lipid and moisture results were reported for 6 cormorant (*Phalacrocorax auritus*) egg composite samples (from three nests at two stations) analyzed in 1 lab batch. No replicate or spiked samples of any kind were reported/analyzed. No minimum requirements are listed in the 2018 RMP QAPP.

Overall acceptability

Overall the data submission is acceptable. 100% of the results are reportable.

MDLs sensitivity

Non-detects (NDs) were reported for eight out of 13 of the Perfluoronates (62%): six out of those eight Perfluoronates were 100% NDs (Perfluorobutanesulfonate, Perfluorobutanoate, Perfluoroheptanoate, Perfluorohexanoate, Perfluorooctanesulfonamide, and Perfluoropentanoate). Forty-three percent of Perfluorooctanoate and twenty-nine percent of Perfluorohexanesulfonate results were NDs.

QB averages (procedural, field blank)
None of the 13 Perfluoronates were measured in the method blank at concentrations equal to, or above the method detection limit (MDL), meeting the method objective of the 2018 RMP QAPP of being "<MDL". No qualifiers were added.

Lipids and moisture were not reported/analyzed for the method blank.

Accuracy (using a variety of SRMs or Matrix spike QRECs)
For samples with a known concentration, consisting of certified reference material (CRM), were run at a minimum frequency of one per analytical batch (for analytical batches consisting of up to 20 field samples) or per 20 (field) samples for larger analytical batches. Analysis of CRMs allows us to evaluate measurement accuracy, or how close our measurement comes to a consensus/expected value.

The CRM analyzed for the perfluoronates (NIST 1947), however, contains only a reference value for Perfluorooctanesulfonate (PFOS). The average percent error for this analyte in the CRM was 14.92% (average recovery 85.08%).

Matrix spikes, where an environmental sample is "spiked" with a known amount of mercury, provide an alternative determination of method accuracy that can account for matrix interferences or other analytical problems. Laboratory control samples are an aliquot of a reference matrix fortified (spiked) with known quantities of specific compounds and subjected to the entire analytical procedure to determine the accuracy of the method by measuring recovery; they are not externally validated values.

The average percent error for the matrix spike sample for the perfluoronates ranged from 2.01% (average recovery 97.99%) for Perfluoroundecanoate to 17.59% (average recovery 82.41%) for Perfluoropentanoate well below the target MQO of 35%. No qualifiers were added.

The average percent error examined for the perfluoronates in the laboratory control sample ranged from 0.5% (average recovery 100.5%) for Perfluorobutanesulfonate to 23.5% (average recovery 76.5%) for Perfluoropentanoate well below the 35% target MQO.

The accuracy of the lipid and moisture data could not be determined as no spiked samples of any kind were reported/analyzed.

Average precision from replicate field sample

The precision of analysis methods (ability to consistently obtain the same result) is determined by analyzing replicate or duplicate samples. The analysis of lab replicates (split and analyzed in the laboratory) allows us to assess the repeatability of lab measurements.

Lab replicates were used to decide whether precision flags were needed. RPDs could not be calculated for six of the individual perfluoronates as at least one of results for the pair was a non-detect (ND). The average RPD for the replicate perfluoronates ranged from 0.41% for Perfluorononanoate to 5.75% for Perfluorohexanesulfonate well below the target MQO of 35%. No qualifiers were added. No field replicates were analyzed. No matrix spike replicates were analyzed.

The precision of the lipid and moisture data could not be determined as no replicate samples of any kind were reported/analyzed.

Comparison of dissolved and total phases
Not Applicable

Comparison to previous years
Average concentration for Perfluorodecanoate in the cormorant egg samples was 70% of the previous average of the combined 2006, 2009, and 2016 RMP EEPS samples (in units of ng/g ww). Average concentration for Perfluorododecanoate was 106%, for Perfluorohexanesulfonate it was 69%, for Perfluorononanoate it was 32%, for Perfluorooctanesulfonate it was 34%, for Perfluorooctanoate it was 21%, and for Perfluoroundecanoate it was 66%. Other perfluoronates could not be compared as they were non-detects (NDs).

Ratio Checking Summary
Overall 2018 shows continued declines for PFOS, PFOA as well as most the long-chained carboxlates (C8>) which is to be expected as the market shifts away from these chemicals (due to bans and voluntary restrictions). We do not see the shorter chained substitutes in biota (<C7) because they do not bioaccumulate. There is a market shift towards polyfluorinated alkyl substances some of which can degrade to the carboxylates but it doesn't seem like we see this in the data (i.e. no increase in carboxylates). In South Bay, perfluorodecanoate (C10) and perfluorododecanoate (C12) concentrations appear to be holding relatively stable.

The duplicate data are astonishing close, so really good RPD. And no detects in the blank above MDL.

Sums Summary

AXYS

Bird Eggs (Tern 2016, Tern 2018, Cormorant 2018)

PBDE

QA Issues for Project Manager to Review
Tentative identified compounds in target list (184), and outside target list (202).
2016 tern average concentrations are about half those for the prior 2 rounds, possibly due to different sites being included?

Reporting Issues for Lab to Review
Resub via email is not ideal.

Formatting Issues for Data Manager to Review
Protocol code varied among samples, some RMP, some RMP_EEPS, all changed to RMP for now to stop queries from breaking data up into different groups. Should be changed to all match whatever the final choice for upload.

Looks like all the Cormorant samples are labeled as replicate 2? Not sure the need for this since we have species to distinguish, my only guess is that the lab does not keep species as a groupable field, so needed to separate the tern dupes and MS from cormorant dupes and MS? Could be left as is unless someone thinks it may cause confusion with people looking for/expecting a Replicate 1 for cormorants.

Hold time review (especially desired by stormwater programs)
All the Cormorant eggs were analyzed within the QAPP listed 1 year, but all the Tern results were analyzed much later (~3 years). All tern results were therefore flagged with VH for hold time exceedance, but the losses are likely minimal with cold/frozen storage.

QA Review
Dataset completeness
Results were reported for 48 PBDEs (including coeluters), for 6 cormorant composites and 1 lab rep, and for 24 tern composites plus 1 lab rep, meeting the overall QAPP target of 1 lab dupe per 20 field samples (here 2 per 30 overall). Lipid and moisture were also reported for all field samples and their lab dupes (with slight variation in the lab dupes, so independent results rather than results copied over from a single analysis). 5 lab blanks and 5 LCSs, 1 MS and 1 MSD, and 4 CRM results were also reported for PBDEs.

Percent usable (non-reject) field data
About 73% of the results were reportable, with most of the censored results coming from poor precision on lab replicates for less abundant congeners.

Overall acceptability
Overall the data are acceptable, with the dominant/most abundant congeners were for the most part reportable.

MDLs sensitivity
A bit over half the 48 PBDEs reported had NDs in at least some samples, with 11 of them ND in 100% of cormorants, and 6 100% ND in terns. The detections being concentrated in a smaller group of analytes is common for the PBDE sets however.

QB averages (procedural, field blank)
PBDE 015, 100, 138, 153, 154, 208 were found in one or more blanks with average concentrations >MDL, although the lab flagged additional analytes with the blank flag (IP). We did not remove any of the lab applied blank flags, but censored some sample results (VRIP flag) for being <3x higher than the blank average in their batches for PBDEs 37, 138, and 208.

Average precision from replicate field sample
Average RSDs were within the target 35% for most PBDEs, except for 10 mostly lower abundance congeners which all had RSDs >70% and thus were censored (VRIL flagged) for poor precision.

Composites within a given field site were often highly variable, especially for terns, with average RSDs over 100% (stdev > mean) in many cases. This is likely due to inter-individual variation, so results were not flagged for variable field results within sites.

Accuracy (using a variety of SRMs or Matrix spike QRECs)
Recoveries were within the target of <35% error from the certified or expected value for all but one of the analytes with certified results in CRMs and concentrations >3xMDL. PBDE 17 had 29% recovery (71% error) so was censored (VRIU flag) for poor recovery. MS recoveries were all within 35% of target for analytes without certified CRM values, so no additional flags were needed.

Comparison of dissolved and total phases
Not applicable

Comparison to previous years
For the most common congeners (47, 49, 99, 100, 153, 154) average concentrations in 2018 cormorants were between 70-130% of prior 2 events averages, so reasonable. Tern concentrations on these common congeners generally biased low however, mostly around 40-50% of the prior 2 events.

Ratio Checking Summary
Thanks to a far lower BDE-209 MDL (averaging 0.064 ng/g ww, instead of 2016's 1.976 ng/g ww), this congener was frequently detected in bird eggs, unlike previous analyses. The low levels are to be expected for aquatic birds; terrestrial birds often have higher levels (and ratios) of this congener.

Ratio review reveals few significant outliers exceeding two standard deviations greater/less than the mean for each bird group (cormorant 2018, tern 2016, tern 2018). Those deviations that do occur all involve congeners found in the Penta commercial mixture (BDE-47, 99, 100, 153, 154), and do not suggest any analytical issues. A brief list of affected samples includes:
L29696-3, cormorant 2018, high BDE-47 likely explained by lower levels of BDE-99 and -100
HS-2-PBDE, tern 2016, lower levels of BDE-99 and -100 are reported, though they are quite similar to the duplicate, which was just inside the cutoff
HS-3-PBDE, tern 2016, higher levels of BDE-154 reported

18-A3W-2-PBDE, tern 2018, higher levels of BDE-99 and -154 reported

Temporal and regional trends:
As noted in QA review, 2018 cormorant levels are similar to those reported in 2016. Sums of PBDEs in samples taken from the Cargill Salt Pond A7 site (2EEPSPA7) have a mean value (53 ng/g ww) nearly double that of those taken from the Richmond Bridge (2EEPSRB; 28 ng/g ww). The Cargill Salt Pond site appears to be new (for cormorants), and potentially more contaminated than the previous South Bay site, Don Edwards (2EEPSDEP9/10C). Meanwhile, the Richmond Bridge samples have lower levels of contamination than reported in 2012 (53) and 2016 (72), suggestive of continuing declines over the past decade+.

As noted in the QA review, tern levels are about half those reported in 2012. This may be due to selection of different sites, or to long-term declines anticipated following the phase-out of Penta and Octa BDE commercial mixtures.

Sums Summary

MPSL-DFW

Cormorant and Tern Eggs

Selenium

QA Issues for Project Manager to Review

Overall acceptability

Overall the dataset is acceptable. 100% of the results are reportable.

Forster'stern moisture values (in units of % ww) are very low averaging 1.24% (compared to 84% for moisture in the Double-crested cormorant egg samples). This is because the Forster's tern moisture results are the result of an analysis of an already dried sample sent to the lab by USGS.

Adam has suggested that there may be a need to discuss either putting a note about the values, or combining the moisture results from DFW and USGS so that people can convert the dw Selenium value to ww (currently it would be difficult).

Accuracy

The accuracy for selenium was flagged following the SFEI RMP Status and Trends convention of using the average percent error of the certified reference material (CRM), if available, in preference to the percent error of the matrix spike/matrix spike replicates and any laboratory control samples.

The selenium certified reference material (CRM) result meet the method quality objective listed in the 2018 RMP QAPP of "expected value ± 35%", and the average percent error of about 10% was less than a RPD of 35%; no accuracy flags are indicated.

The selenium matrix spike results meet the method quality objective listed in the 2018 RMP QAPP of "expected value ± 35%, and the average percent error of 7% was less than a RPD of 35%; accuracy was sufficient based on the certified reference material results.

The accuracy of moisture data could not be determined as there are no certified or reference values for moisture in the CRM and no spiked samples of any other kind were reported/analyzed.

Precision

The precision of field samples in the database is flagged following the SFEI RMP Status and Trends convention of using lab replicates in preference to using field replicates, although both are reviewed and described narratively when provided. No field replicates were included in this data submission.

The selenium lab replicate results that could be evaluated, one RPD was not used as the result was less than three times the method detection limit (MDL), met the method quality objective listed in the 2018 RMP QAPP of an RPD < 35% so the set of field samples was not flagged in the database.

Precision of the moisture data could not be determined as no replicate samples were reported/analyzed.

Reporting Issues for Lab to Review

Formatting Issues for Data Manager to Review

Hold time review (especially desired by stormwater programs)

Forster's tern (*Sterna forsteri*) egg composite selenium samples were analyzed between 256 and 270 days after collection within the 1 year holding time specified in the 2018 RMPQAPP.

Double-crested cormorant (*Phalacrocorax auritus*) selenium samples were analyzed between 268 and 269 days after collection within the 1 year holding time specified in the 2018 RMPQAPP.

QA Review

Dataset completeness

Selenium results were reported for 12 Forster's tern (*Sterna forsteri*) egg composite samples (three eggs at four stations) analyzed in 1 lab batch. The target was 7 homogenized eggs. One lab replicate and 1 matrix spike/matrix spike duplicate were analyzed for 12 Forster's tern (*Sterna forsteri*) egg composite samples meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for those sample types. Two method blanks and 1 certified reference material sample (NRC DORM-4: Fish protein certified reference material for trace metals) were analyzed meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for these sample types. All data were reported not blank corrected.

Moisture results were reported for 12 Forster's tern (*Sterna forsteri*) egg composite samples (three eggs at four stations) analyzed in 1 lab batch. No minimum requirements are listed in the 2018 RMP QAPP.

Selenium results were reported for 6 double-crested cormorant (*Phalacrocorax auritus*) egg composite samples (three eggs at two stations) analyzed in 1 lab batch. The target was 7 homogenized eggs. One lab replicate and 1 matrix spike/matrix spike duplicate were analyzed for the 6 double-crested cormorant (*Phalacrocorax auritus*) egg composite samples meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for those sample types. Two method blanks and 1 certified reference material sample (NRC DORM-4: Fish protein certified reference material for trace metals) were analyzed meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for these sample types. All data were reported not blank corrected.

Moisture results were reported for 6 double-crested cormorant (*Phalacrocorax auritus*) egg composite samples (three eggs at two stations) analyzed in 1 lab batch. No minimum requirements are listed in the 2018 RMP QAPP.

Overall acceptability
Overall the data submission is acceptable. 100% of the results are reportable.

MDLs sensitivity
Non-detects (NDs) were not reported for moisture or selenium.

QB averages (procedural, field blank)
Selenium was not measured in the method blanks at concentrations equal to, or above the method detection limit (MDL), meeting the method objective of the 2018 RMP QAPP of being "<MDL". No qualifiers were added.

Moisture was not reported/analyzed for the method blanks.

Accuracy (using a variety of SRMs or Matrix spike QRECs)
For samples with a known concentration, consisting of certified reference material (CRM), were run at a minimum frequency of one per analytical batch (for analytical batches consisting of up to 20 field samples) or per 20 (field) samples for larger analytical batches. Analysis of CRMs allows us to evaluate measurement accuracy, or how close our measurement comes to a consensus/expected value.

The average percent error for selenium in the CRM (NRC DORM-4: Fish protein certified reference material for trace metals) was 10.43% (average recovery 89.57%) below the MQO target of 35%. No qualifiers were added.

Matrix spikes, where an environmental sample is "spiked" with a known amount of mercury, provide an alternative determination of method accuracy that can account for matrix interferences or other analytical problems. Laboratory control samples are an aliquot of a

reference matrix fortified (spiked) with known quantities of specific compounds and subjected to the entire analytical procedure to determine the accuracy of the method by measuring recovery; they are not externally validated values.

The average percent error for selenium in the matrix spike samples was 7.23% (average recovery 92.77%) well below the target MQO of 35%.

The accuracy of moisture data could not be determined as there are no certified or reference values for moisture in the CRM and no spiked samples of any other kind were reported/analyzed.

Average precision from replicate field sample
The precision of analysis methods (ability to consistently obtain the same result) is determined by analyzing replicate or duplicate samples. The analysis of lab replicates (split and analyzed in the laboratory) allows us to assess the repeatability of lab measurements.

Lab replicates were used to decide whether precision flags were needed. One RPD was not used in the evaluation as the result was less than three times the method detection limit (MDL). The RPD for the other selenium replicate was 2.9% well below the target MQO of 35%. The average RPD for the matrix spike samples was 2.81%. No field replicates were analyzed. No qualifiers were added.

Precision of the moisture data could not be determined as no replicate samples were reported/analyzed.

Comparison of dissolved and total phases
Not Applicable

Comparison to previous years
Average concentration for selenium in the cormorant egg samples was 66% of the average of the combined 2002, 2004, 2006, 2009, and 2016 RMP EEPS samples (in units of ng/g dw). Average concentration for selenium in the Forster's tern egg samples was 54% of the average of the combined 2009 and 2016 RMP EEPS samples (in units of ng/g dw).

Forster'stern moisture values (in units of % ww) are very low averaging 1.24% (compared to 84% for moisture in the Double-crested cormorant egg samples). This is because the Forster's tern moisture results are the result of an analysis of an already dried sample sent to the lab by USGS.

Ratio Checking Summary
Not Applicable

Sums Summary

MPSL-DFW

Cormorant Eggs

Mercury

QA Issues for Project Manager to Review

Overall acceptability

Overall the dataset is acceptable. 100% of the results are reportable.

Accuracy

The accuracy for mercury was flagged following the SFEI RMP Status and Trends convention of using the average percent error of the certified reference material (CRM), if available, in preference to the percent error of the matrix spike/matrix spike replicates and the laboratory control samples.

The mercury certified reference material (CRM) result meet the method quality objective listed in the 2018 RMP QAPP of "expected value ± 35%", and the average percent error of about 3% was less than a RPD of 35%; no accuracy flags are indicated.

The mercury matrix spike results meet the method quality objective listed in the 2018 RMP QAPP of "expected value ± 35%, and the average percent error of 4.67% was less than a RPD of 35%; accuracy was sufficient based on the certified reference material results.

The accuracy of moisture data could not be determined as there are no certified or reference values for moisture in the CRM and no spiked samples of any other kind were reported/analyzed.

Precision

The precision of field samples in the database is flagged following the SFEI RMP Status and Trends convention of using lab replicates in preference to using field replicates, although both are reviewed and described narratively when provided. No field replicates were included in this data submission.

The mercury lab replicate results met the method quality objective listed in the 2018 RMP QAPP of an RPD < 35% so the set of field samples was not flagged in the database.

Precision of the moisture data was evaluated using the lab replicate samples with an average RPD of 1.24%.

Reporting Issues for Lab to Review

Formatting Issues for Data Manager to Review

Hold time review (especially desired by stormwater programs)

Mercury samples were analyzed between 344 and 349 days after collection within the 1 year holding time specified in the 2018 RMPQAPP. No holding time issues were flagged.

QA Review

Dataset completeness

Mercury results were reported for 34 double-crested cormorant (*Phalacrocorax auritus*) egg composite samples (19 eggs at one station and 15 eggs at a second) analyzed in 3 lab batches. Three lab replicate and 3 matrix spike/matrix spike duplicate samples were analyzed for the 34 double-crested cormorant (*Phalacrocorax auritus*) egg composite samples meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for those sample types. Nine method blanks and 3 certified reference material sample (NRC DORM-4: Fish protein certified reference material for trace metals) were analyzed meeting the minimum requirement in the 2018 RMP QAPP of 1 per 20 samples, or 1 per batch for these sample types. All data were reported not blank corrected.

Moisture results were reported for 34 double-crested cormorant (*Phalacrocorax auritus*) egg composite samples (19 eggs at one station and 15 eggs at a second), 3 lab replicates, and 3 certified reference material samples (NRC DORM-4: Fish protein certified reference material for trace metals) analyzed in 3 lab batches. No minimum requirements are listed in the 2018 RMP QAPP.

Overall acceptability

Overall the data submission is acceptable. 100% of the results are reportable.

MDLs sensitivity

Non-detects (NDs) were not reported for moisture or mercury.

QB averages (procedural, field blank)

Mercury was not measured in the method blanks at concentrations equal to, or above the method detection limit (MDL), meeting the method objective of the 2018 RMP QAPP of being "<MDL". No qualifiers were added. Moisture was not reported/analyzed for the method blank.

Accuracy (using a variety of SRMs or Matrix spike QRECs)

For samples with a known concentration, consisting of certified reference material (CRM), were run at a minimum frequency of one per analytical batch (for analytical batches consisting of up to 20 field samples) or per 20 (field) samples for larger analytical batches. Analysis of CRMs allows us to evaluate measurement accuracy, or how close our measurement comes to a consensus/expected value.

The average percent error for mercury in the CRM (NRC DORM-4: Fish protein certified reference material for trace metals) was 2.94% (average recovery 101.46%) below the MQO target of 35%. No qualifiers were added.

Matrix spikes, where an environmental sample is “spiked” with a known amount of mercury, provide an alternative determination of method accuracy that can account for matrix interferences or other analytical problems. Laboratory control samples are an aliquot of a reference matrix fortified (spiked) with known quantities of specific compounds and subjected to the entire analytical procedure to determine the accuracy of the method by measuring recovery; they are not externally validated values.

The average percent error for mercury in the matrix spike samples was 4.67% (average recovery 104.47%) well below the target MQO of 35%.

The accuracy of moisture data could not be determined as there are no certified or reference values for moisture in the CRM and no spiked samples of any other kind were reported/analyzed.

Average precision from replicate field sample
The precision of analysis methods (ability to consistently obtain the same result) is determined by analyzing replicate or duplicate samples. The analysis of lab replicates (split and analyzed in the laboratory) allows us to assess the repeatability of lab measurements.

Lab replicates were used to decide whether precision flags were needed. The average RPD for the mercury lab replicates was 4.8% well below the target MQO of 35%. The average RPD for the matrix spike samples was 4.75%. No field replicates were analyzed. No qualifiers were added.

Precision of the moisture data was evaluated using the lab replicate samples with an average RPD of 1.24%.

Comparison of dissolved and total phases
Not Applicable

Comparison to previous years
Average concentration for mercury in the cormorant egg samples was 87% of the average of the combined 2002, 2004, 2006, and 2016 RMP EEPS samples (in units of ng/g ww).

Ratio Checking Summary
Not Applicable

Sums Summary