United States Department of the Interior

U.S. GEOLOGICAL SURVEY
Reston, VA 20192
http://www.usgs.gov/

In Reply Refer To:
Mail Stop 412

May 20, 2016

OFFICE OF WATER QUALITY TECHNICAL MEMORANDUM 2016.05

SUBJECT: Guidance on Annotating Results Affected by Contamination Bias, with Examples for Water Samples Affected by Co and Mn Contamination from High-Capacity Capsule Filters

This memorandum provides guidance on annotating results affected by contamination bias with examples for water samples that may have been affected by cobalt (Co) or manganese (Mn) from high-capacity capsule filters. Contamination attributable to these filters (National Field Supply Service item #Q398FLD) was noted from about October 1, 2008 to about September 30, 2014, as reported in WaQI Note 2015.04. Subsequently, a nationwide retrieval of NWIS data for filtered water samples from the national aggregated database for this period found Co detected in approximately 42 percent of field blanks and Mn detected in approximately 33 percent of field blanks.

Contamination with Co and (or) Mn was noted in water samples processed with 0.45-µm high-capacity capsule filters made by Pall Corporation. These filters are sold as National Field Supply Service (NFSS) item #Q398FLD through USGS One-Stop Shopping. Although the NFSS routinely tests filters, not all filters are found to be contaminant-free (e.g., WaQI Note 2009.11 describes a contaminant issue). To mitigate the risk of contamination, the Office of Water Quality (OWQ) provides guidance for rinsing filters (e.g., WaQI Note 2014.07). However, in the recent case of Co and Mn, the prescribed field rinsing did not always remove all of the contamination.

Whenever contamination bias is significant relative to the concentrations in the environmental samples, action is needed to communicate the degree of the bias to ensure that results reported to the National Water Information System (NWIS) and other products are of known quality. These actions can include applying a result-level comment ("e") or the "V" remark code, applying a raised reporting level, or otherwise censoring the affected data.

Information in this memorandum, summarized from a nationwide dataset, may be useful in guiding local decisions about annotating affected results for Co and Mn; however, because projects or networks may differ in data-collection practices, it is important for each project or network to consider their local context when formulating decisions. The attached guidance includes three steps:

1. Identifying which data may be affected by contamination bias,

2. Determining the frequency and magnitude of the contamination bias on the basis of field-blank data, and
3. Selecting and applying an approach for annotating the affected results for environmental samples in NWIS and other products.

Explicit steps for identifying and annotating environmental sample results that are affected by field sources of contamination are contained in the attached instructions. Although this guidance provides a specific example of how to address contamination bias based on an evaluation of Co and Mn detections in field-blank data gathered from NWIS, the techniques discussed can be adapted for other forms of contamination bias. A spreadsheet tool called the "Binomial Calculator" that was used for this example is attached to this memorandum and can be accessed through the OWQ internal Web page at http://water.usgs.gov/usgs/owq/software.html. Please send questions or comments regarding this memorandum to Lisa Olsen (ldolsen@usgs.gov) of the Water Science Field Team.

REFERENCES

Office of Water Quality Water-Quality Information Note 2015.04, Low-level detections in high capacity filters, accessed January 14, 2016, at http://water.usgs.gov/usgs/owq/WaQI/WaQI2015.04.pdf.

Office of Water Quality Water-Quality Information Note 2014.07, Inspection and rinse protocol for high-capacity capsule filters, accessed January 14, 2016, at http://water.usgs.gov/usgs/owq/WaQI/WaQI2014.07-Filter-inspection-rinse.pdf.

Office of Water Quality Water-Quality Information Note 2009.11, Field methods—Cobalt detected in Pall Versapor® capsule filters, lot #FQ8545, accessed January 14, 2016, at http://water.usgs.gov/usgs/owq/WaQI/WaQI09.11.pdf.

Donna N. Myers
Chief, Office of Water Quality

Distribution: All WMA Employees

Attachments:

- INSTRUCTIONS: Evaluating and annotating environmental sample results affected by contamination bias for cobalt (Co) and manganese (Mn)
- Binomial_Calculator_and_Test_Datasets (Microsoft Excel® spreadsheet)

INSTRUCTIONS: Evaluating and annotating environmental sample results affected by contamination bias for cobalt (Co) and manganese (Mn)

1. Identify which data are affected by contamination bias:
Identify which data are affected by field sources of contamination bias by one or more of the following means:

- Review results from field blanks (FBs) and environmental samples (ENVs) collected for your project or network. If contaminants are detected in the FBs, first determine whether the contamination could have originated from the laboratory or the blank water used to collect the FBs by checking the results of laboratory blanks, certificates of analysis for blank water used for the field blanks (see http://wwwnwql.cr.usgs.gov/qas.shtml?ibw), and any source-solution blanks that were collected. In the case of the Co and Mn in filtered water samples, the contamination bias in the nationwide dataset was not attributed to laboratory blanks or source water.
- Look for patterns associated with specific equipment or processes. Time-series plots (concentration *vs.* time) may be helpful in illustrating the timing of the contamination in the FBs and ENVs. Take note of the NWIS filter lot numbers, typically recorded under NWIS parameter code 99206. With sufficient data, it may be possible to identify subsets of samples (affected vs. unaffected), based on filter lot numbers. Review WaQI Note 2015.04 as a basis for comparison to your data.
- Consider pooling data among projects that used the same sampling procedures, in order to increase the confidence in determining the distribution of potential contamination. For example, a minimum of 22 FBs is needed to have at least 90-percent confidence that the maximum concentration in the FBs represents the 90th percentile of the contamination in the ENVs (Mueller & others, 2015).
- The guidance in this document presumes that each project or network collected FBs; thus a comparison of local data to the nationwide dataset is possible. If no FBs were collected, please take steps to ensure that future data-collection efforts include appropriate numbers and types of QC samples.

2. Identify the frequency and magnitude of the contamination on the basis of data from field blanks:
If contamination is identified in the FBs, the ENVs may also be affected. The results from the FBs can be used to infer how the ENVs were affected. The premise in collecting QC samples is that the FBs represent the same conditions affecting the ENVs. Thus, a set of FBs can be compared to a set of ENVs, rather than matching the FBs to ENVs on a 1-to-1 basis. The detection frequency in field blanks, *DF*, is calculated by dividing the number of FBs with detections by the total number of FBs:

$$DF = \frac{FBs\ with\ detections}{total\ number\ of\ FBs} \times 100\ percent$$

If the contamination was infrequent (detection frequency < 5 percent) or if the number of FBs was inadequate, statistical approaches for determining the magnitude of contamination for populations of FBs and ENVs may not be practical. In these cases, the FBs and ENVs can be evaluated as follows:

(1) Were the FB concentrations less than (<) one-tenth of the ENV concentrations? If so, no censoring is warranted (see Office of Water Quality Technical Memorandum 97.08). For

products outside of NWIS, one would mention the infrequent contamination occurrence for the FBs and explain that no censoring was applied to the ENVs.

(2) If FB concentrations were greater than (>) one-tenth the ENV concentrations, determine if any patterns exist in the FB detections; for example, did the FB detections occur during a limited period or with filters of a specific lot number? If specific detections in the ENVs can be inferred to be affected by the same source as affected the FBs (e.g., same lot number, similar concentrations), those ENV results can be annotated by assigning the "**e**" value-qualifier code and including a field comment, such as "*Analyte detected in __ percent of corresponding field blanks at concentrations of up to ___ µg/L.*" For products outside of NWIS, explain how the affected ENVs were identified and treated in any interpretations.

If the contamination was frequent (detection frequency > 5 percent), as was observed for Co and Mn in filtered water samples, statistical approaches may be useful in determining the magnitude of the contamination in the FBs relative to the ENVs, as shown in Fig. 1.
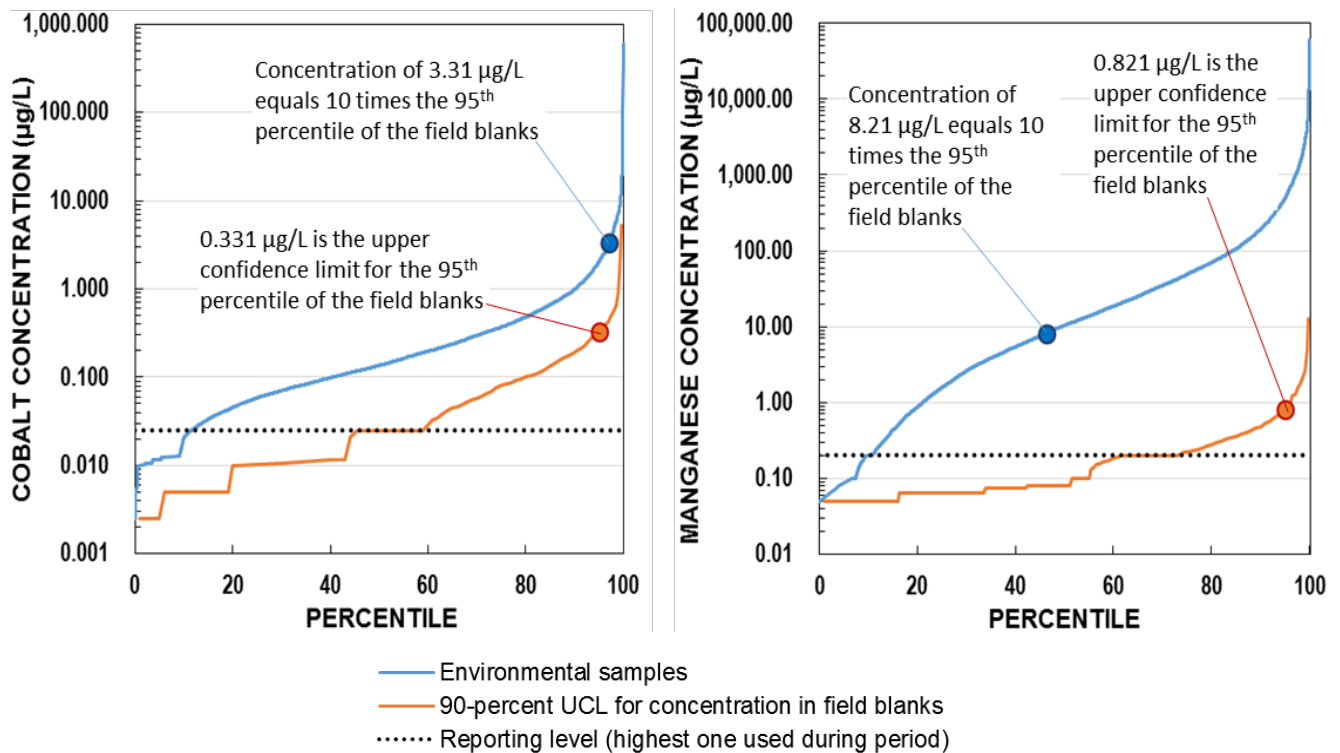

Figure 1. Distribution of cobalt (Co) and manganese (Mn) concentrations in environmental samples and 90-percent upper confidence limit (UCL) for percentiles of concentrations in field blanks collected from October 1, 2008, through September 30, 2014.

The plots in fig. 1 were constructed using data from NWIS for filtered water samples analyzed by the NWQL, following procedures described by Mueller and others (2015), following the conceptual example in fig. 9 of that report. Results below the reporting level were assigned a concentration of one-half the reporting level (in this case, the LT-MDL) for plotting. For Co, data were limited to NWQL method PLM10 (elements in filtered water using collision/reaction cell inductively coupled plasma-mass spectrometry) in order to reduce variation in reporting levels. For Mn, data from

PLA11 (metals, filtered water, inductively coupled plasma-atomic emission spectrometry) and PLM43 (metals, filtered water, inductively coupled plasma-mass spectrometry) were used. Data were excluded if the analyzing entity or laboratory method could not be identified, or if fundamental metadata such as the sample type or medium code were missing or inconsistent.

Information from the FBs (orange lines) can be used to identify which range of ENV concentrations (blue lines) are likely to be affected by contamination bias (fig.1). At the low end of the range, ENVs that are non-detections are considered to be unaffected by bias because the contaminant is absent. At the upper end, ENVs with concentrations greater than 10 times the 90-percent upper confidence limit (UCL) for the 95th percentile of the FBs also are unlikely to be affected by contamination bias, although there is still a ≤ 5 percent chance of a contamination effect.

Results between the detection limit and 10 times the 90-percent UCL for the 95th percentile of the FBs may be influenced by contamination bias of a magnitude that affects the first or second significant figures in concentrations reported for the ENVs, up to 10 times the estimate made using FBs (0.321 µg/L for Co and 0.821 µg/L for Mn). **Thus, on the basis of the nationwide dataset, detected concentrations of up to 3.31 µg/L for Co and 8.21 µg/L for Mn would warrant annotation** as being potentially affected by contamination bias attributable to field processes, in excess of any actual contamination present in the sample. Annotating ENVs within these ranges would be a conservative approach to evaluating data quality. ENVs of higher concentration would have a lower relative effect of the bias than ENVs of lower concentration, although all ENVs have the same likelihood of being affected or unaffected, regardless of concentration.

The concentrations for the 95th percentile of the FBs for Co and Mn (at the 90-percent UCL) are plotted along with the results for FBs and ENVs in NWIS to illustrate the relative magnitude of the contamination bias (figs. 2-3). These results from the nationwide dataset are provided as a basis for comparison for local projects and networks. If one's local context provides a reason to expect the local results to differ from the nationwide dataset (for example, if a local project used a different source of filters), an independent statistical analysis of that project's data would be warranted.

3. Select and apply an approach for annotating the affected results for ENVs:
Once the range of affected concentrations in the ENVs has been identified through the use of time-series plots and statistical approaches described in the previous section, the next step is to annotate the affected results in NWIS and other products. Approaches for results in NWIS are limited to the result-level remark codes and value-qualifier codes permitted by the system, as described at http://help.waterdata.usgs.gov/codes-and-parameters/codes#WQ and as follows:

- **Apply the "V" remark code to the affected data.** The "V" remark code identifies which results are affected by contamination, but does not convey the magnitude of contamination. This solution is based on guidance in Office of Water Quality Technical Memorandum 97.08, which stipulates that the following conditions are met:
    1. There is direct evidence of contamination.
    2. The concentration of the contamination is significant relative to the environmental concentration.
    3. The contamination is sufficiently understood to allow some interpretation of the environmental data. Generally, this will require that the contamination can be shown to be systematic and the source and magnitude can be identified.

- **Use the "e" value-qualifier code to add a result-level comment to the affected data.** Because the "**V**" remark code alone does not tell data users about the source and magnitude of the contamination, additional annotation is warranted. The "**e**" value-qualifier code can be used to add an informative comment, such as "*Analyte detected in __% of associated field blanks at concentrations of up to ___µg/L (__th percentile with __% confidence), attributed to contamination in capsule filters.*"
- **The combined use of the "V" remark code and "e" value-qualifier code** serve the dual functions of alerting data users to the issue and describing the issue.

Approaches for results in USGS publications and other products include this additional option:
- **Apply a raised reporting level (RRL) to affected ENV results.** This solution entails identifying a concentration (RRL) at or above which the ENV results are considered representative of the sampled matrix, with minimal influence from contamination bias. ENV results that are greater than the detection limit but less than the RRL are reported as <Q, where Q is the quantified result. The RRL is selected so that ENVs within the same concentration range as the majority of the contaminated FBs are censored. The selection of an RRL depends on the degree of protection against false-positive risk and desired confidence level. Nonparametric statistics are advised for data that are not normally distributed, as was the case with Co and Mn in field blanks. A method derived from Hahn and Meeker (1991) uses order statistics and the binomial distribution to determine upper confidence levels (UCLs) for specified percentiles of the contaminant concentrations. This method is described by Mueller & others (2015) and is taught in the USGS Quality-Control Sample Design and Interpretation class (QW 2034).
- **Based on the examples in fig. 1 for cobalt and manganese,** a set of RRLs of 0.331 µg/L for Co and 0.821 µg/L for Mn could be applied in publications that cover the affected time period, with a citation of this memorandum, provided that the results in the publication represent filtered water samples collected using the Pall capsule filters and were analyzed using the NWQL methods PLM10 for Co and PLA11 or PLM43 for Mn.

Summary and Implications to Data in NWIS
Deciding how or whether to annotate data in NWIS and other products in response to contamination bias is the responsibility of each program or network that collected the data. Evaluation of environmental data relative to field (and laboratory) blank data is essential to ensure that data reported to the public is not a result of false positives or extrinsic contamination bias due to field- or laboratory-derived contamination. Failure to address known or suspected contamination bias can lead to erroneous scientific conclusions about the resources that were sampled and can give an impression that inadequate consideration was given to assuring the quality of the data.

With respect to cobalt (Co) and manganese (Mn) results from water samples that could be biased from contamination attributable to high-capacity capsule filters, it is apparent from figs. 1-3 that large numbers of environmental samples have concentrations within the same ranges as were observed in field blanks. These pooled results cannot take into account the specific circumstances that pertain to each data-collection effort. Therefore, it is imperative that the staff responsible for data quality perform the necessary steps to evaluate their data within their local context, and to annotate the affected data as described in this document.

Additional resources:

An Excel spreadsheet tool, that facilitates use of the binomial distribution method is attached to this document. The use of the binomial distribution for trace-element contamination observed in field blanks is described in several publications. Davis and others (2014) and Olsen and others (2010) explicitly address the use of RRLs (called "SRLs" in those reports).

- Mueller and others, 2015: http://pubs.usgs.gov/tm/04/c04/
- Davis and others, 2014: http://dx.doi.org/10.3133/sir20145105
- Olsen and others, 2010: http://pubs.usgs.gov/sir/2009/5220/
- Apodaca and others, 2006: http://pubs.usgs.gov/sir/2006/5093/

Time-series plots (figs. 2-3) generated from the nationwide dataset used to generate fig. 1 may be useful for illustrating the timing of the contamination bias as well as for evaluating the effects of applying an RRL to the environmental samples; for example, at the 95th percentile at $\geq$ 90 percent UCL (the solid red line). Detections were plotted as filled symbols; orange for FBs and blue for ENVs. Nondetections were plotted as open symbols at their respective reporting levels.
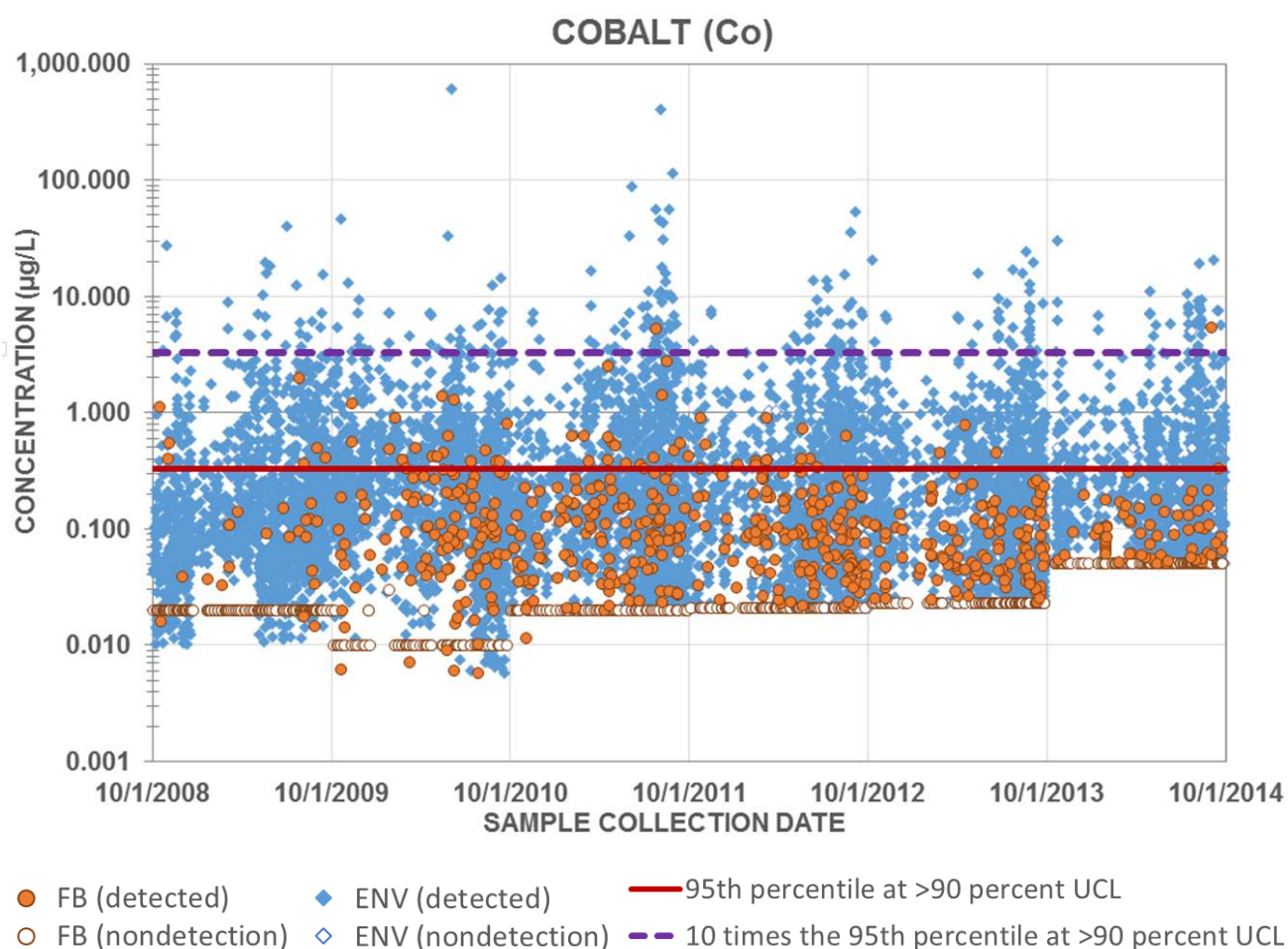

Figure 2. Concentrations of cobalt (Co) in filtered water samples collected from October 1, 2008, through September 30, 2014, and analyzed at the NWQL
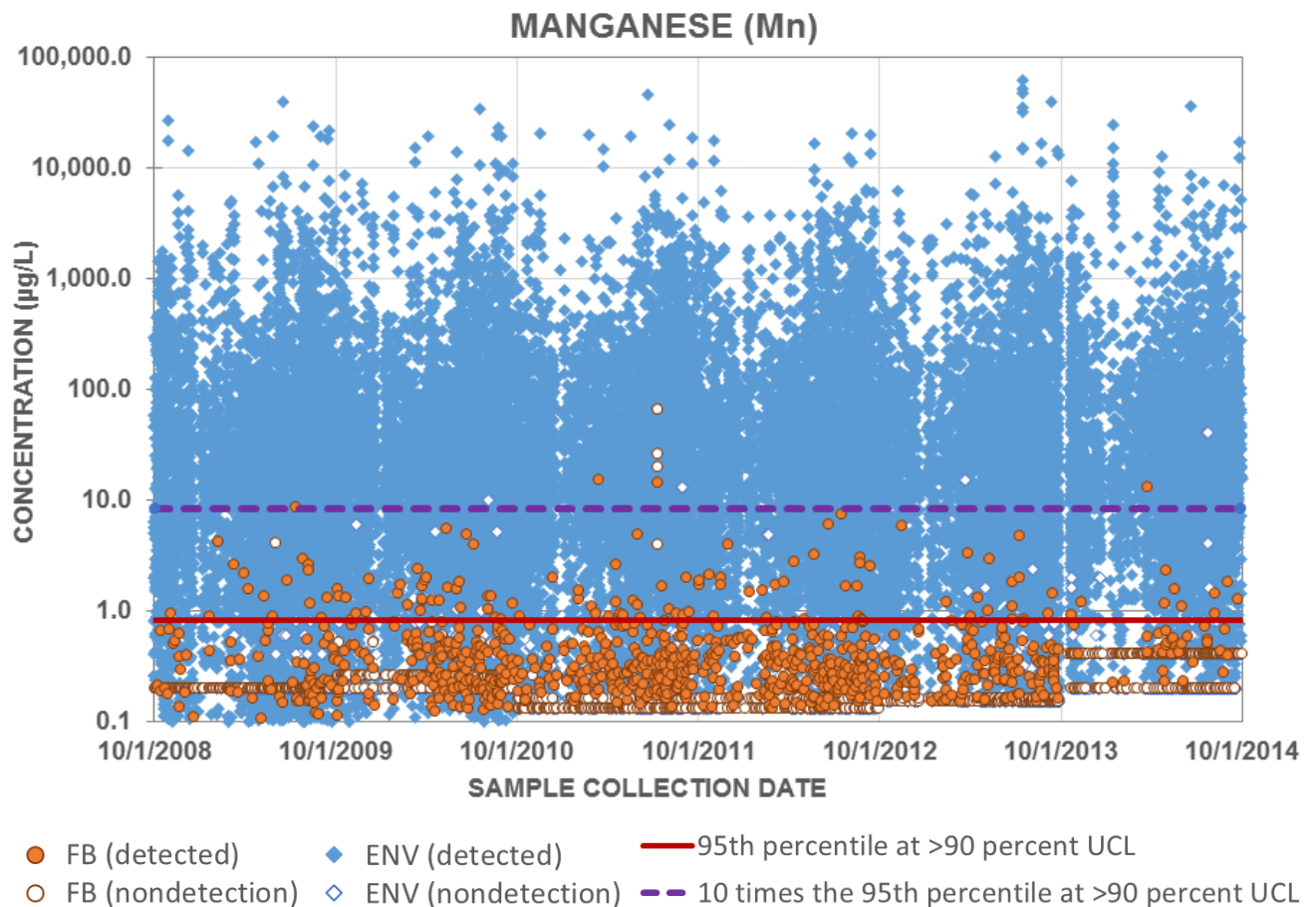
Figure 3. Concentrations of manganese (Mn) in filtered water samples collected from October 1, 2008, through September 30, 2014, and analyzed at the NWQL

The solid red lines on figs. 2-3 represent the 90-percent UCL for the 95th percentile of the FBs, which was of 0.331 µg/L for Co (fig. 2) and 0.821 µg/L for Mn (fig. 3). The dashed purple line represents the upper extent of the influence of the contamination bias based on a threshold of 10 times the 90-percent UCL of the 95th percentile of the FBs.

The nationwide data retrieval that was used to generate figs. 2-3 also included data for NWQL method PLA11 for cobalt, and data for two other laboratories, the University of Southern Mississippi (MS-USM) and Huffman Laboratories in Golden, Co (CO-HFMAN). Table 1 summarizes these results in terms of the frequency and magnitude of contamination in FBs. Differences in results might be due to differences in laboratory reporting levels or the types or lot numbers of filters used to process samples. Results in table 1 can be used as a basis for comparison; however, because projects or networks may differ in data-collection practices, it is important for each project or network to evaluate their own data for Co and Mn in filtered water samples with consideration of their local field practices and laboratory detection limits.

Table 1. Summary statistics for cobalt and manganese in field blanks collected from October 1, 2008, through September 30, 2014, from the National Water Information System (NWIS)

[UCL, upper confidence level; µg/L, micrograms per liter; NWQL, USGS National Water Quality Laboratory; MS-USM, University of Southern Mississippi; CO-HFMAN, Huffman Laboratories, Golden, CO; LT-MDL, long-term method reporting level; LRL, lower reporting level; nc, not calculated because number of field blanks was insufficient for the specified UCL.]

Laboratory	Method codes[1]	Highest reporting level used (µg/L)	Reporting level type	Results for field blanks			
				Number	Detection frequency (percent)	Maximum concentration (µg/L)	95th Percentile at UCL ≥ 90 percent (µg/L)
Cobalt (Co)							
NWQL	PLM10	0.05	LT-MDL	1,453	42.1	5.448	0.3264
NWQL	PLA11	1.6	LRL	33	3.0	1.1	nc
MS-USM	PLM58	0.037	LRL	34	52.9	0.1475	nc
Manganese (Mn)							
NWQL	PLA11, PLM43	0.40	LT-MDL	2,738	32.6	14.93	0.821
MS-USM	PLM58	5.00	LRL	34	32.4	7.153	nc
CO-HFMAN	PLM43	0.20	LRL	59	91.5	0.52	0.52

[1] Laboratory method codes are defined in NWIS as follows:

PLA11: Metals, filtered water, Inductively coupled plasma-atomic emmission spectrometry

PLM10: Elements in filtered water using collision/reaction cell inductively coupled plasma-mass spectrometry

PLM43: Metals, filtered water, Inductively coupled plasma-mass spectrometry

PLM58: Elements in filtered water by inductively coupled plasma-mass spectrometry

References Cited

Apodaca, L.E., Mueller, D.K., and Koterba, M.T., 2006, Review of trace element blank and replicate data collected in ground and surface water for the National Water-Quality Assessment Program, 1991–2002: U.S. Geological Survey Scientific Investigations Report 2006-5093, 32 p., accessed June 30, 2015, at http://pubs.usgs.gov/sir/2006/5093/.

Davis, T.A., Olsen, L.D., Fram, M.S., and Belitz, Kenneth, 2014, Updated study reporting levels (SRLs) for trace-element data collected for the California Groundwater Ambient Monitoring and Assessment (GAMA) Priority Basin Project, October 2009–March 2013: U.S. Geological Survey Scientific Investigations Report 2014–5105, 52 p., http://dx.doi.org/10.3133/sir20145105.

Hahn, G.J., and Meeker, W.Q., 1991, Statistical intervals—a guide to practitioners: New York, John Wiley and Sons, 392 p.

Mueller, D.K., Schertz, T.L., Martin, J.D., and Sandstrom, M.W., 2015, Design, analysis, and interpretation of field quality-control data for water-sampling projects: U.S. Geological Survey Techniques and Methods book 4, chap. C4, 54 p., http://dx.doi.org/10.3133/tm4C4.

Olsen, L.D., Fram, M.S., and Belitz, Kenneth, 2009, Review of trace-element field-blank data collected for the California Groundwater Ambient Monitoring and Assessment (GAMA) Program, May 2004–January 2008: U.S. Geological Survey Scientific Investigations Report 2009-5220, 47 p., accessed July 29, 2015, at http://pubs.usgs.gov/sir/2009/5220/.