

National Water Quality Program

Quality of Data From the U.S. Geological Survey National Water Quality Network for Water Years 2013–17

Scientific Investigations Report 2020–5116

Quality of Data From the U.S. Geological Survey National Water Quality Network for Water Years 2013–17

By Laura Medalie and Laura M. Bexfield

National Water Quality Program

Scientific Investigations Report 2020–5116

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
DAVID BERNHARDT, Secretary

U.S. Geological Survey
James F. Reilly II, Director

U.S. Geological Survey, Reston, Virginia: 2020

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment—visit <https://www.usgs.gov> or call 1–888–ASK–USGS.

For an overview of USGS information products, including maps, imagery, and publications, visit <https://store.usgs.gov/>.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this information product, for the most part, is in the public domain, it also may contain copyrighted materials as noted in the text. Permission to reproduce copyrighted items must be secured from the copyright owner.

Suggested citation:

Medalie, L., and Bexfield, L.M., 2020, Quality of data from the U.S. Geological Survey National Water Quality Network for water years 2013–17: U.S. Geological Survey Scientific Investigations Report 2020–5116, 21 p., <https://doi.org/10.3133/sir20205116>.

Associated data for this publication:

Bexfield, L.M., Sandstrom, M.W., and Beaty, D., 2020, Field, laboratory, and third-party data for assessment of the quality of pesticide results reported by the National Water Quality Laboratory for groundwater samples collected by the National Water-Quality Assessment Project, 2013–18: U.S. Geological Survey data release, <https://doi.org/10.5066/P90BFKA4>.

Medalie, L., and Shoda, M.E., 2020, Field blank and field replicate datasets for inorganic and organic compounds collected for the National Water Quality Network, water years 2013–17: U.S. Geological Survey data release, <https://doi.org/10.5066/P96VY980>.

Contents

Abstract.....	1
Introduction.....	1
Purpose and Scope	2
Types and Uses of QC Information.....	3
Methods.....	4
Field Quality-Control Samples	4
Laboratory Quality-Control Samples	8
Integrated Assessment for Pesticides.....	9
Quality of Data for Nutrients, Carbon, and Sediment	10
Quality of Data for Major Ions and Trace Elements	13
Quality of Data for Pesticides	14
Summary.....	18
Acknowledgments	19
References Cited.....	19

Figures

1. Map of surface-water sites in the U.S. Geological Survey National Water Quality Network in the United States; mapped from Lee and Reutter (2019)2

Tables

1. Information about chemical compounds in water samples from the U.S. Geological Survey (USGS) National Water Quality Network analyzed at the USGS National Water Quality Laboratory for water years 2013–17.....3
2. Quality-control datasets and associated data-quality objectives for water-quality sampling in the U.S. Geological Survey National Water Quality Network and repositories for interpreted quality-control results for water years 2013–17
3. Summary information for field blanks and surface-water samples of the U.S. Geological Survey National Water Quality Network and comparisons to benchmarks for water years 2013–17
4. The 90-percent upper confidence limit for the 95th-percentile concentrations of nutrient, carbon, major ion, trace element, and select pesticides in field blanks of the U.S. Geological Survey National Water Quality Network and suggestions for censoring based on detections in field blanks for water years 2013–17
5. Estimated variability in detection frequency of field replicates for samples in the U.S. Geological Survey National Water Quality Network for water years 2013–17
6. Estimated variability in concentrations of field replicates for samples in the U.S. Geological Survey National Water Quality Network for water years 2013–17
7. Summary of results of data-quality assessment for inorganic compounds analyzed in surface-water samples of the U.S. Geological Survey National Water Quality Network based on all field quality-control sample types for water years 2013–17

demonstrates contamination with a detectable concentration of an analyte; if this contamination is indicated, then analysis of the source-solution blank may help determine whether the contamination is from the source-solution water (if the analyte is detected in the source-solution blank) or from the field (if the analyte is not detected in the source-solution blank).

Field replicates are used to estimate the potential variability in detection frequency and concentrations of environmental samples. Variability is typically expressed as uncertainty with a specified level of confidence; project protocols might define acceptable levels of uncertainty for their data. Uncertainty might be expressed as two different metrics for low- and high-concentration ranges (as defined in the section “Field replicate samples”). For example, variability for water-quality data can be estimated as the average standard deviation of field replicates within the low range of concentrations as defined by the project and as the relative standard deviation of field replicates in the high range of concentrations for datasets where there is a change in slope for standard deviations as a function of mean concentration for replicate pairs (Mueller and others, 2015).

Field matrix spikes are environmental samples with spike solutions (described in the section “Collection and processing of samples”) added in the field to measure the effect of the sample water matrix on the recovery efficiency (recovery bias) of the analytical method. Recovery is the measured amount of the compound in a spiked sample expressed as a percentage of the amount spiked, ideally 100 percent (Martin and others, 2009; Shoda and others, 2018). For NWQN samples, field matrix spikes are collected only for organic compounds (pesticides); they are not collected for inorganic constituents, which generally do not degrade and are not affected by matrix interference (Mueller and others, 2015). Field matrix spikes, as well as laboratory spikes from the NWQL and blind spikes from the QSB, assess recovery bias and variability in recovery (table 2). The combination of these three different types of spike samples could help to determine if recoveries in field matrix spikes outside of a target range are caused by analyte degradation, matrix interference, or analytical performance (Mueller and others, 2015).

Methods

Field-blank, source-solution blank, and field replicate data for all compound groups sampled by the NWQN are presented in Medalie and Shoda (2020). To provide a single source for all field replicate results, published field replicate results for pesticides for 2013–15 (Shoda and others, 2017a) are combined with previously unpublished field replicate results for pesticides for 2016–17 and all field replicate results for 2013–17 for inorganic compounds sampled by the NWQN (Medalie and Shoda, 2020). Data for field matrix spikes have been published in Shoda and others (2017b) and Wieben (2019) and are not reproduced with this work. Results for all

laboratory QC samples that are summarized in this report have been published in Shoda and others (2017b), Wieben (2019), and Bexfield and others (2020a).

Field Quality-Control Samples

Collection and processing of samples.—Protocols for cleaning equipment and collecting and processing QC and environmental samples are described in chapters A4 and A5 of the USGS National Field Manual (USGS, 2019) and in Riskin and others (2018) and are briefly described in this section. Depending on the analyte, sample handling protocols for the nutrient, carbon, major ion, and trace element schedules used by the NWQN specify whether or not samples are filtered, acidified, and chilled. All pesticide compounds in schedule 2437 are filtered and chilled.

Field blanks are collected for all constituents analyzed by the NWQN except for sediment. The rarity of detections in field blanks for sediment led to the decision to discontinue their collection because enough data had been collected to be confident that samples were free from contamination bias (Riskin and others, 2018). Field blanks are prepared using water from the NWQL that is certified to be free of constituents of concern. Inorganic blank water (IBW) is used for field blanks for inorganic constituents. Pesticide-grade organic blank water (OBW) is used for field blanks for pesticides and most other organic constituents and for ultraviolet-absorbance measurement blanks. Considerations for determining the timing and sites for collection of NWQN field blanks include characteristics of the constituents, expected constituent concentrations at the site, and results of previous QC analyses (Riskin and others, 2018). For NWQN sites, the percentages of environmental samples with associated field blanks range from about 4 to 9 for nutrients and carbon, from 4 to 5 for major ions and trace elements, and from 5 to 6 for pesticides during the 5-year study timespan.

Field blanks are prepared in the field immediately before collection of the environmental sample using the same equipment and protocols as the environmental sample (Riskin and others, 2018; USGS, 2019). Approximately 2–3 gallons of either IBW or OBW (depending on the type of analysis) is used to rinse and process precleaned equipment for field blanks. The basic procedure for field blanks is to first rinse the sample collection bottle and assembly, including the compositing container and filter apparatus, with IBW or OBW to simulate as closely as possible the field rinsing that occurs prior to collection of the environmental sample. After rinsing, the remaining blank water is processed through the sample collection bottle into the compositing container and then either directly into sample bottles (for unfiltered samples) or through the filter apparatus into sample bottles (for filtered samples) that are sent to the NWQL for analysis. Filtration steps specific to the types of analytes are documented in the National Field Manual: section 5.2.1 describes procedures for inorganic

constituents, section 5.2.2.B describes procedures for pesticides, and section 5.2.2.C describes procedures for carbon analysis (USGS, 2019).

Collected less frequently than field blanks but still routinely at NWQN sites, source-solution blanks are prepared at the field site before collecting the field blank by rinsing and then pouring IBW or OBW into a sample bottle or vial (Riskin and others, 2018). The IBW or OBW used for source-solution blanks is the same source-solution water that is used to prepare the field blank and tests whether the blank water itself is free from contamination. Often, source-solution blanks are stored and only analyzed if there is a detection in the associated field blank.

Most field replicates for the NWQN are collected as split replicates, where the environmental and field replicate samples are decanted from the same compositing container (Riskin and others, 2018). The environmental sample is processed first, followed by the field replicate. If filtering is required, the filter apparatus is changed before processing the replicate sample.

The procedure to collect field replicates for sediment analysis, which is a concurrent replicate, is different from that for environmental samples. To sample a stream for sediment analysis, individual bottles are collected at multiple vertical points over a cross section of the stream. For field replicates, the replicate sample collection bottle is filled immediately after the environmental collection bottle at each vertical (Riskin and others, 2018).

Collection of a field matrix spike for pesticides begins with an environmental sample and a split replicate sample, using procedures described in the previous paragraph. Using a micropipette, 0.1 milliliter of the spike solution warmed to room temperature is added to the split replicate sample. The spike solution used for field matrix spikes is a mixture prepared by the NWQL that contains known aliquots of target analytes and is the same spike solution that the NWQL uses for laboratory spikes. An important part of the protocol is to record the lot number and volume of the spike solution and volume of the spiked water sample, without which recovery cannot be computed.

Compilation and screening of data.—Analytical results for field blank and field replicate samples summarized in this report were retrieved during February 2019 from the internal-use portal of the USGS National Water Information System (NWIS) database and are presented in Medalie and Shoda (2020). Environmental data only are available through the public NWIS portal (USGS, 2020); field and laboratory QC data are not available through the public portal. From 2013 through 2017, the NWQN had about 1.7 million analytical results for environmental water-quality samples, 94,000 for field blanks, 107,000 for field replicates, and 56,000 for field matrix spikes.

Before publication in NWIS, USGS water-quality data undergo several review steps. After initial review by primary and secondary analysts at the NWQL, every preliminary analytical result for environmental and field QC samples from the NWQL is provided to a data analyst at the water science

center that collected the samples and is assessed for quality, completeness, and reasonableness. Some large USGS projects have additional review steps for data and metadata that might happen before or after publication in NWIS, depending on project protocols. As an example, analysts involved with the NWQN periodically evaluate long-term QC datasets that are in a collective “inference space” after results are published in NWIS (Medalie and others, 2019). An inference space for QC is the spatial and temporal conditions for which QC samples can be related to environmental samples because the QC results reflect the same conditions, in terms of potential bias and variability, under which the environmental samples were collected (Mueller and others, 2015). This latter type of review is the primary topic of this report.

The initial screening for field-blank and field-replicate data retrieved from NWIS for this project was to remove missing, rejected, or duplicated results. Sample coding (metadata) was checked for accuracy and completeness. For example, a field QC sample with a concentration similar to the associated environmental sample may have been erroneously coded as a field blank when it should have been coded as a field replicate.

Detections in field blanks were plotted over time and overlain with plots of detections in laboratory blanks and in NWQN environmental samples. Unusual patterns, for example clusters of field blanks with detections, were investigated by checking for non-field-related sources of contamination such as the source-solution water from which the field blank was produced. The certificates of analysis of IBW and OBW and of analytical results from all available source-solution blanks were examined for compounds that were detected in field blanks to ensure that contamination indicated by the field blank was not from the source solution itself. Where an investigation revealed that the contamination in a field blank might have stemmed from a nonfield source (such as the source solution) and thus had no bearing on contamination of environmental samples, a comment was added to the data in Medalie and Shoda (2020), and the result was not included in the summary of field blanks or the determination of suggested censoring because this result would have no effect on interpretation of the associated environmental sample.

Because there were very few detections in field blanks for pesticides, all detections were individually investigated to assess whether the detection was related to something pertaining only to the particular sample or might be signaling something more generally applicable to other samples. For example, individual field personnel, equipment, or source solutions may introduce errors, which in turn may introduce biases. In some cases, this investigation into detected pesticide field blanks prompted requests to the NWQL for verification of results. Some requests for verification resulted in confirmation of the detections, and other requests where the original detection was not verified resulted in a change of the original detection to a nondetection. Changes from these types of verifications were made to results in the NWQL Laboratory Information Management System and in NWIS.

The benefit of reviewing long-term datasets within an inference space is illustrated with the following example. Individual sample collectors review analytical results for samples that they submit to the NWQL. Suppose that a field-blank sample submitted by hydrologist A had a detected result for a specified analyte. Hydrologist A would likely surmise that contamination was introduced during a field-related process and that the associated environmental sample was subject to the same source of field-related contamination. Based on that assumption, hydrologist A might censor the environmental result. Assume, however, that the source of contamination was not from the field but from a laboratory-related process that affected a random subset of environmental or QC samples in each of several consecutive sets. In this hypothetical scenario, the random laboratory contamination only affected the field-blank sample submitted by hydrologist A and did not affect any environmental samples in the sample set. For an environmental sample submitted by hydrologist B and processed in the same analytical set as hydrologist A's samples, hydrologist B would have no knowledge of the detection in hydrologist A's field blank and would assume that the environmental sample is free of contamination. Thus, hydrologist A would correctly censor the environmental sample but for the wrong reason (assuming contamination was from a field-related rather than a laboratory-related process) and hydrologist B would incorrectly neglect to censor the environmental sample.

This type of nonpervasive contamination also might not get noticed in reviews by NWQL analysts. This is because the primary approach used by the NWQL to address laboratory-derived contamination is set censoring (censoring of samples within a given preparation set), which addresses contamination that occurs within sets of samples prepared and analyzed together. Although effective at addressing contamination within sets, set censoring does not address contamination that might be randomly present but is not necessarily present in every set of samples (Medalie and others, 2019). Retrospective review of long-term QC datasets might be the only way to detect a pattern of low-level contamination that does not surface during routine reviews by NWQL analysts or data-reviewing hydrologists.

As described in Shoda and others (2018), screening steps for field matrix spikes include investigation and analysis of value qualifier codes and comparison of spiked concentrations with reporting levels and with environmental concentrations, raised reporting levels, and effects of spike lot and spike lot age on recovery.

Temporal patterns (2013–17) of QC information within the inference space of samples evaluated for this report that could be used to refine interpretations of environmental results were investigated. For example, field replicates for a compound might indicate high variability in results in 2013 and 2014 and minimal or negligible variability in results in 2016.

Two pesticide compounds, phosmet and methyl parathion (parameter codes 65101 and 65089, respectively) were dropped from schedule 2437 for samples from 2015 because poor analytical performance rendered them unsuitable for

interpretive use (USGS, 2015); these compounds are not included in summaries of this report or in Medalie and Shoda (2020). Nineteen pesticide compounds have a permanent “E” (estimated) remark code for quantified results (table 1) because data-quality objectives of the NWQL were not met during method validation (Sandstrom and others, 2015). Effective January 1, 2016, 15 additional pesticide compounds were given a permanent “E” remark code after routine review of performance by the NWQL (Duane Wydoski, USGS, written commun., December 3, 2015). An “E” remark code indicates some uncertainty of the concentration value but does not indicate uncertainty in the identification of the analyte because all qualitative identification criteria must be met before a concentration can be reported (Sandstrom and others, 2015).

The QSB provides independent information about data quality for results produced by the NWQL. In addition to QC results from the QSB that are discussed in the “Laboratory Quality-Control Samples” section, the Inorganic Blind Sample Project, which is the inorganic branch of the QSB, provides internal bimonthly data-quality assessments, and the Organic Blind Sample Project, which is the organic branch of the QSB, has provided five comprehensive data-quality assessments for compounds in schedule 2437 between 2013 and 2017. These assessments for inorganic and organic compounds describe performance-related biases or variability and were consulted to help interpret any unusual results in field QC samples.

Statistical and graphical data analysis.—Field blank, field replicate, and field matrix spike results are presented with results of the associated environmental sample so that the data user can make side-by-side comparisons (Shoda and others, 2017a, b; Wieben, 2019; Medalie and Shoda, 2020).

Field blank samples.—Field-blank data for inorganic and organic compounds are summarized as percent detections, maximum concentration, and the 90-percent upper confidence limit (UCL) for the 95th percentile concentration. The nonparametric binomial probability approach was used to determine the latter concentration, whereby the ranked (ordered) blank value corresponding to the 90-percent UCL, calculated as Jeffreys interval (Brown and others, 2001; Fram and others, 2012), is the maximum contamination expected, with 90-percent confidence, for 95 percent of all samples. The 90-percent UCL for the 95th percentile concentration was calculated using all data for each compound as well as separately for each year in the study period.

For the six pesticide compounds in schedule 2437 with at least two detections in field blanks during the study period, the 90-percent UCL for the 95th percentile concentration was calculated by year. This calculation required an adaptation to determine the ranked position because these compounds, unlike the inorganic compounds, could have detections that are reported below the reporting level (RL) or the detection level (DL). The adaptation was necessary because calculation of the 90-percent UCL for the 95th percentile concentration relies on order statistics, which would erroneously rank nondetections (reported with a “<” remark and the RL concentration) above

detections that are less than the RL or the DL. The adaptation for the ranked position was to rank all nondetections below all detections (Bexfield and others, 2020b).

Raised reporting levels (RRLs), which are nondetections reported at a nonstandard level, introduce another complication for ranking needed for the 90-percent UCL for the 95th percentile concentration calculation and were handled as in the USGS studies of Shoda and others (2018) and Bexfield and others (2020b). Results reported with an RRL at or below the first calibration standard above the maximum standard RL were included in data analysis as nondetections at the RRL. Results reported with an RRL greater than this were excluded from analysis. RRL results with no indication of the reason for the RRL were excluded from analysis for the 90-percent UCL for the 95th percentile concentration calculation. Field blank results excluded from this analysis are tracked through use of the comment field in the Table 1_datarelease_c3fieldblanks.csv table in Medalie and Shoda (2020).

Environmental relevancy of potential contamination bias is established in several ways. The 90-percent UCL for the 95th percentile concentration is a proxy for the potential bias that could be introduced to samples as a result of contamination (Bender and others, 2011; Mueller and others, 2015). Providing this metric for individual years allows the data user to focus on the time period and environmental results of interest and assess whether there might be concerns about contamination bias in a smaller time period than the entire 5-year study period.

Relevancy of potential contamination also is determined by comparing the maximum concentration of the field-blank results with surface-water results and with benchmark concentrations. Because low concentrations in surface water potentially are more sensitive to low-level contamination, the minimum and the 1st percentile of detected concentrations in surface-water samples are provided; the 1st percentile buffers against potentially anomalous minimum results. Maximum concentrations of field blanks are presented as fractions of benchmarks to enable a worst-case assessment of whether potential contamination in NWQN samples could be affecting concentrations near benchmark values. Time-series plots of field-blank and surface-water results for all inorganic and organic compounds are in Medalie and Shoda (2020, plotfiles 1, 3, and 5). The minimum benchmark concentration is shown on these time-series plots if it falls within the range of concentrations of environmental samples.

Types and availability of benchmarks are specific to individual compounds. Human-health benchmarks (HHBs) for inorganic compounds in this report include U.S. Environmental Protection Agency (EPA) maximum contaminant levels (MCLs) for nitrite, nitrate, fluoride, arsenic, and selenium and health-based screening levels (HBSLs) for boron and strontium (EPA, 2018a; Norman and others, 2018). HHBs for pesticides include MCLs for alachlor, atrazine, carbofuran, simazine, 2,4-D, and oxamyl; cancer HBSLs for carbaryl, cyanazine, diuron, fentin, and fluometuron;

noncancer HBSLs for 34 pesticides; carcinogenic HHBs for 11 pesticides; and chronic noncancer HHBs for 65 pesticides (Norman and others, 2018).

Other types of benchmarks for inorganic compounds include the minimum concentration in the range of EPA (2019a) nutrient criteria for ecoregions for phosphorus and total nitrogen and an acute aquatic-life criterion for ammonia (EPA, 2019b). About half of the pesticides in this report have at least one of several possible aquatic-life benchmarks (ALBs; EPA, 2018b). For compounds with more than one ALB, only the lowest benchmark concentration is presented to provide a worst-case comparison with concentrations in blank samples. Although acute ALBs are the most directly comparable type of benchmark to instantaneous concentrations of discrete sample results, chronic ALBs (for invertebrates or fish) are often lower than acute and are presented in this report as conservative (protective) thresholds.

The conventions used for censoring suggestions for NWQN samples based on field-blank detections, as described in this section, also could apply to other samples collected and processed using the same procedures as NWQN samples and analyzed at the NWQL using the same methods during the same time frame as the QC samples described in this report. If the 90-percent UCL for the 95th percentile concentration field-blank value for any year is a detected concentration that is less than the RL, then a suitable censor level for that year is the RL. If the 90-percent UCL for the 95th percentile concentration for any year is a detected concentration greater than the RL, then a suitable censor level for that year is a concentration between the RL and the 90-percent UCL for the 95th percentile concentration. The appearance of one or two detections in field blanks likely reflects potential contamination from sample processing or a unique condition at a specific site (or an anomalous value) and is unlikely to be indicative of a pervasive contamination problem that is generalizable to samples collected across the NWQN. Furthermore, with generally less than 100 field blanks in any given year, detections indicated at the 90-percent UCL for the 95th percentile concentration value of the binomial calculation usually have a higher UCL than 90 because the convention of the calculation is to round up to the next highest ranked sample (Mueller and others, 2015, p. 18). For these reasons and to guard against censoring that might be overly conservative, the general convention is refined such that censoring is not suggested based on a single detection in a field blank for a given year, regardless of the 90-percent UCL for the 95th percentile concentration calculation. Similarly, if one or two field-blank detections are greater than the RL, the suggestion is to censor at the RL and not at the potentially higher 90-percent UCL for the 95th percentile concentration calculated value.

Field replicate samples.—Field replicate samples (replicate samples) are used to evaluate variability in detection frequency and in concentration of water samples. Water is intrinsically variable in time, space, and matrix (composition); variability can also arise during sample collection and processing. Variability in detection frequency is evaluated in

terms of inconsistent detections of the replicate pairs (Mueller and others, 2015) and mean detection rate. Analytical results for compounds analyzed at the NWQL are reported either as a quantified detection or as a censored (nondetected) result, which is expressed as less than the reporting level. An inconsistent detection is defined as one result of the replicate pair being a detection and the other result being censored.

By convention, pairs in which one sample is reported as a detection (for example, 7 nanograms per liter [ng/L]) and the second sample is censored but at a higher reporting level (for example, <10 ng/L) than the concentration reported in the first sample, this sample pair is still recorded as inconsistent, even though the first (detected) concentration is within the range defined by the second, uncensored value. This may result in a positive bias of the number of inconsistent detections, especially for pesticides, because schedule 2437 has a fairly high rate of raised reported levels (11 percent of all data for 2013–15), which represents greater uncertainty in the reported result for a pesticide than detections (Shoda and others, 2017a). Inconsistent detections are presented in this report relative to the total number of replicate pairs and also relative to the number of replicate pairs with at least one detection (that is, excluding replicate pairs where both samples are censored). The mean detection rate is presented for replicate pairs with at least one detection. Compounds that exhibit low variability in detection frequency for the purposes of this report are defined as those with (1) 25 percent or fewer inconsistent detections or (2) a mean detection rate of at least 90 percent of replicate pairs with at least one detection (Martin, 2002). Compounds that exhibit high variability in detection frequency are defined as those with (1) at least 50 percent inconsistent detections of replicate pairs with at least one detection or (2) a mean detection rate of 75 percent or less (Martin, 2002). Compounds with neither low nor high variability in detection frequency are assumed to have moderate variability.

For compounds with a sufficient number of detected replicate results at low and high concentrations to characterize two distinct slopes (as described in this paragraph), variability in concentrations is described in this report using the two-range model—because for many analytes, variability at low concentrations is best described differently from variability at high concentrations (Mueller and Titus, 2005; Mueller and others, 2015). At low concentrations, the standard deviation of replicate pairs is approximately constant and is the most appropriate description of variability, whereas at high concentrations the relative standard deviation is approximately constant and is the most appropriate description of variability. The approximate concentration forming the boundary between low and high concentrations is determined to be the breakpoint in slope of the scatterplot of mean concentration against the standard deviation of replicate pairs (scatterplots are not shown; breakpoints in slope are provided in the “Potential variability bias of nutrient, carbon, and sediment data based on field replicates” part of the “Quality of Data for Nutrients, Carbon, and Sediment” section). Variability for analytes with no breakpoint in slope is determined as the mean standard deviation. Using

standard deviation or relative standard deviation as an estimate of variability, data users may construct confidence intervals around critical values to help interpret environmental data (Mueller and others, 2015).

The mean relative standard deviation of replicate pairs for each given year is used to calculate variability in concentrations on an annual basis because, for most combinations of pesticides and years, there are too few replicate pairs to separate into low and high ranges of concentration for the two-range model. It is important to note that mean relative standard deviation overestimates variability for low concentrations and thus is a conservatively high estimate of variability. For interpretations in this report, a threshold of 0.2 mean relative standard deviation for replicate pairs was used to indicate high variability. However, if a data user has strict requirements for interpretations, such as performing statistical comparisons among groups of data before and after an event (such as the Deepwater Horizon oil spill; Nowell and others, 2013), a lower threshold might be used.

Field matrix spike samples for pesticides.—All spike results (field matrix, laboratory, and third-party) are provided in terms of percentage of the spiked aliquot (of known volume and concentration) that is recovered during analysis, known as “percent recovery” or “recovery.” Reference documents for analytical methods typically provide target percent recoveries to indicate acceptable method performance. For example, beginning in 2013, the analytical method used by the NWQL to analyze samples collected for the NWQN program for pesticides was a liquid chromatography-tandem mass spectrometry method that set target recoveries of 100±30 percent (70–130 percent; Sandstrom and others, 2015).

Summary statistics of field matrix spike recovery percentages include minimum, 25th percentile, median, 75th percentile, maximum, mean, standard deviation, relative standard deviation, and F-pseudostigma (Bexfield and others, 2020b). F-pseudostigma is a nonparametric measure of variability that is less sensitive to outliers than the relative standard deviation and is calculated as the absolute value of the difference between the 25th and 75th percentiles divided by 1.349 (Hoaglin and others, 1983). Compounds with median recoveries outside the target range of 70–130 percent and those with F-pseudostigmas less than or equal to 30 percent are flagged because these metrics indicate that environmental data could be exhibiting a high or low bias.

Laboratory Quality-Control Samples

Two separate USGS entities, the NWQL and the QSB, provided the three types of laboratory QC samples that are discussed in this report. A fourth type of QC sample from the QSB, blind blanks, are discussed for pesticides in Bexfield and others (2020b) but were not available for the time frame of this study. The NWQL provided laboratory blanks and laboratory spikes, and the QSB provided blind-spike samples (table 2). Laboratory blanks from the NWQL (also called

method, reagent, or preparation blanks) are used to assess possible contamination for a set of samples during preparation, processing, and instrument analysis. Laboratory spikes from the NWQL are a synthetic matrix fortified with known concentrations of some or all of the method analytes and are used to assess method performance bias that in theory can be used to adjust concentrations for method-related recovery issues.

Blind spikes from the QSB are samples made up of a matrix fortified with known concentrations of all or a representative selection of the method analytes. The synthetic matrix of blind spikes for pesticides from the QSB is intended to be similar to blank water, which precludes evaluation of matrix effects from environmental waters. The matrix for blind spikes for inorganics may be synthetic or may be environmental water. Nondetected analytes in blind spikes known to include the given analyte in the spike mixture are considered to be false-negative results. Detected analytes in blind spikes that were not included in the spike mixture are false-positive results and can be used to assess the false-positive risk in environmental samples—except when the detected compound may be a degradate of a parent analyte (Bexfield and others, 2020b). Because spike solutions used for blind spikes are made by the QSB and not by the NWQL and concentrations of analytes in spike solutions of QSB and NWQL are different, a direct comparison of blind-spike results with field matrix spikes and laboratory spikes may reflect spiking errors or variability between the two laboratories. This caveat only applies to the comparison of field spikes to blind spikes because the NWQL makes spike solutions for laboratory spikes and field matrix spikes.

Laboratory blank, laboratory spike, and blind-spike datasets for the organic pesticide compounds were obtained from the NWQL or the QSB for a period that overlaps this report; results are available in Bexfield and others (2020a) and discussed in Bexfield and others (2020b). The discussion and interpretations in Bexfield and others (2020b) about these datasets also are relevant to the NWQN data and therefore are incorporated into this report, although those data (Bexfield and others, 2020a) are not duplicated in this report. The only part of Bexfield and others (2020a, b) not directly relevant to the NWQN data is the field QC component for groundwater samples. For those interested, comparable laboratory QC datasets for the inorganic compounds included in this report are available by request from the NWQL (labhelp@usgs.gov). Comparable QSB datasets for blind blanks are available in USGS (2018) and for blind spikes are available by request from the Inorganic Blind Sample Project (<https://qsb.usgs.gov/ibsp/>).

Integrated Assessment for Pesticides

Evaluations from assessments of field QC data and laboratory QC data from the NWQL and the QSB are combined to provide an integrated approach to identification of pesticides

with potential for contamination bias, recovery bias, or high variability. This information is provided for individual years. Bexfield and others (2020a, b) integrates information from laboratory QC as well as field QC information for groundwater samples, providing a summary assessment of potential contamination and recovery biases, bias from false-negative risk, and overall high variability. This report combines the laboratory QC summary information from Bexfield and others (2020b, table 11) with field QC information related to surface-water samples in the NWQN. The assessment of bias from false-negative risk where the false-negative occurrence in blind spikes exceeds 10 percent is presented in Bexfield and others (2020b) but is not included in this report because only four compounds (disulfoton oxon, naled, orthosulfamuron, and terbufos oxon) were determined to have that risk where corroboration was provided by median recoveries that were less than 70 percent.

Evidence of the potential for substantial contamination bias of pesticide results is determined by laboratory blanks, blind spikes (false-positive occurrences for compounds that are not part of spike mixtures), and field blanks. The basis of evidence for laboratory blanks and blind spikes is detailed in Bexfield and others (2020b, table 11). For laboratory blanks and field blanks, evidence of bias is indicated if the 90-percent UCL for the 95th percentile concentration is a detection. Evidence of bias is indicated for individual years based on the 90-percent UCL for the 95th percentile concentration using data for the given year. Blind spikes provide evidence of substantial contamination bias if the false-positive occurrence at concentrations greater than the maximum detection level exceeds 1 percent across all years (Bexfield and others, 2020b).

Evidence of the potential for substantial recovery bias is determined by laboratory spikes, blind spikes, and field matrix spikes. For each of these types of samples, evidence of recovery bias is determined from median recoveries less than 70 percent (low bias) or greater than 130 percent (high bias). The basis of evidence for laboratory spikes and blind spikes is detailed in Bexfield and others (2020b, table 11).

Evidence of the potential for high variability is determined using laboratory spikes, blind spikes, and field replicates. Field replicate pairs with variable detection frequencies or analyte concentrations are determined to be “highly variable.” Laboratory spikes and blind spikes with variable recoveries also are determined to be “highly variable.” The basis of evidence for laboratory spikes and blind spikes is Bexfield and others (2020b, table 11). For laboratory spikes and blind spikes, evidence is indicated if the F-pseudosigma (described in the “Field matrix spike samples for pesticides” section of this report) exceeds 30 percent. Criteria that describe the potential for high variability based on field replicates, either for inconsistent detections or for concentrations, are described in the “Field replicate samples” section of this report.

Quality of Data for Nutrients, Carbon, and Sediment

Potential contamination bias of nutrient and carbon data.—Between 5 and 6 percent of surface-water samples collected for nutrient and carbon analysis in the NWQN were sampled with a concurrent field blank. All nutrient and carbon results for the NWQN had at least one detection in a field-blank result; however, field-blank detection frequencies for 8 of 13 types of nutrient results were less than 5 percent (table 3). The percentage of field blanks with detections for nutrient results ranged from less than 1 (nitrate plus nitrite [analyzed with USGS laboratory method RED01], phosphorus [analyzed with methods CL020 and KJ005], and orthophosphate) to 30 percent (ammonia). Compounds that have maximum and 90-percent UCL for the 95th percentile concentrations in field blanks (ammonia, ammonia plus organic nitrogen, low-level phosphorus [analyzed with method CL021], total particulate nitrogen, and total nitrogen) greater than both the minimum and the 1st percentile concentrations of surface-water samples (table 3; Medalie and Shoda, 2020, plotfile 1) are indicative of some potential for contamination that could affect the lowest concentrations seen in environmental samples.

Censoring suggestions are based on the approach described in the “Statistical and graphical data analysis” section of this report. It is important to note that, even when censoring is suggested, uncensored data are appropriate to use for many purposes such as making general statements about occurrence of compounds in the environment or comparing data to a benchmark concentration that is much higher than the suggested censoring level. Censoring suggestions are most important if there is a benchmark concentration or the need to evaluate individual environmental results that are close to or within an order of magnitude of the censoring level.

Six nutrients or nutrient and method combinations were found to be completely free of contamination bias: nitrite, nitrate plus nitrite (analyzed with method RED01), phosphorus (analyzed with methods KJ009, CL020, and KJ005), and orthophosphate (table 4). Two nutrient or nutrient and method

combinations (nitrate plus nitrite [analyzed with method RED02] and phosphorus [analyzed with method CL021]) demonstrated potential contamination bias for 1 year for samples at low concentrations near the RL. Five nutrient or nutrient and method combinations demonstrated potential contamination bias for 4 or 5 years (ammonia, ammonia plus organic nitrogen, total Kjeldahl nitrogen, total particulate nitrogen, and total nitrogen); however, total nitrogen is the only one of these with a benchmark low enough to be within range of environmental data. Corroboration for the suitability of censoring suggestions offered in this report is provided visually by the patterns in time and concentrations of field blanks (Medalie and Shoda, 2020, plotfile 1). To illustrate censoring suggestions using ammonia (parameter code 00608) as an example, between years 2013 and 2015, there is no more than one detected field blank with concentrations greater than the RL of 0.02 milligram per liter (mg/L) each year, and many detections are between the RL and the DL, rendering censoring at the RL of 0.02 mg/L an appropriate choice. Because there are several detected field blanks with concentrations greater than the RL for 2016 and 2017, a censoring level between the RL and each year’s 90-percent UCL for the 95th percentile concentration would be appropriate.

An internal investigation of ammonia contamination in NWQN samples (Janice Fulford, USGS, written commun., July 23, 2018) found that laboratory method processes, rather than field procedures, were likely the contributing sources in low-level detections of ammonia in field blanks and thus of low-level contamination in environmental samples. Low-level ammonia contamination in surface-water samples from NWQN sites also has been observed in historical data (Mueller and Titus, 2005; Medalie and Martin, 2016).

One context for consideration of ammonia contamination is to compare detections in field blanks to the ammonia ALB, or more specifically, the acute ALB criterion of 17 mg/L for ammonia. The maximum concentration of ammonia in a field blank (0.465 mg/L) is about one thirty-sixth of the ALB, and the 90-percent UCL for the 95th percentile concentration for ammonia is one-thousandth of the ALB, meaning that comparisons of surface-water samples in the NWQN relative to the

Table 3. Summary information for field blanks and surface-water samples of the U.S. Geological Survey National Water Quality Network and comparisons to benchmarks for water years 2013–17.

[This table is available for download at <https://doi.org/10.3133/sir20205116>]

Table 4. The 90-percent upper confidence limit for the 95th-percentile concentrations of nutrient, carbon, major ion, trace element, and select pesticides in field blanks of the U.S. Geological Survey National Water Quality Network and suggestions for censoring based on detections in field blanks for water years 2013–17.

[This table is available for download at <https://doi.org/10.3133/sir20205116>]

ALB are unlikely to be affected by contamination. Regardless of comparisons to the ALB for ammonia, many surface-water results overlap the ranges of time and concentrations of field blanks (Medalie and Shoda, 2020, plotfile 1); therefore, to minimize contamination bias, censoring suggestions in table 4 would benefit from the use of low-concentration ammonia data in occurrence or trend studies.

Ammonia contamination also might affect samples analyzed for other nitrogen compounds that include ammonia, such as ammonia plus organic nitrogen and total Kjeldahl nitrogen (parameter codes 00623 and 00625, respectively). Additional sources of data on the extent, timing, and magnitude of potential laboratory sources of low-level ammonia include (1) concentrations of blind-blank samples from the QSB (available upon request from <https://qsb.usgs.gov/ibsp/>), (2) concentrations in laboratory blanks from NWQL (available upon request from labhelp@usgs.gov), and (3) concentrations in source-solution blanks (Medalie and Shoda, 2020).

The 90-percent UCL for the 95th percentile and maximum concentrations in field blanks for nitrite and for nitrate plus nitrite, the two nutrient compounds with HHBs (MCLs of 1 and 10 mg/L, respectively), are at least 100 times less than the respective MCL. Thus, contamination from field-related sources is unlikely to interfere with interpretations of concentrations of environmental samples near the concentrations of the MCLs. Censoring to the RL concentration of 0.02 mg/L is suggested for NWQN results in 2013 for nitrate plus nitrite analyzed with method RED02 (table 4).

With one exception (affecting low-level results analyzed with one method for 1 year), phosphorus and orthophosphate environmental results are free of contamination bias. Although the maximum concentrations of phosphorus in field blanks analyzed using three of the four methods (method CL020 being the exception) were greater than the minimum value of the range of nutrient criteria as set by the EPA (2019a), detections in field blanks for phosphorus were rare: the 90-percent UCL for the 95th percentile concentration values were all nondetections or less than the RL (table 3). The only suggested censoring for phosphorus is to censor samples collected during 2016 analyzed using the low-level method CL021 at the RL concentration of 0.008 mg/L. This timing coincides with an internal investigation of detections in field blanks and in false-positive results from QSB blind samples, which found two lots of sulfuric acid sample preservative contaminated with phosphorus that could have affected field and environmental samples collected during 2016 (Teresa Williams, USGS, written commun., February 23, 2017). The investigation found that phosphorus was detected in up to 50 percent of samples and at concentrations up to 0.005 mg/L when analyzed using one of the contaminated lots of preservative and using low-level method CL021.

Samples analyzed for total particulate nitrogen and total nitrogen had enough detections in field blanks at high enough concentrations to suggest additional censoring for each year during the study period at either their respective RLs or at the 90-percent UCL for the 95th percentile concentration when

greater than the RL (table 4). Although total nitrogen (parameter code 62854; Medalie and Shoda, 2020, plotfile 1) is the only type of nutrient result that has a benchmark in the range of low concentrations of environmental data, the 0.12-mg/L minimal EPA benchmark (table 3) is the lowest value out of 13 ecoregions; the other 12 ecoregions had higher benchmarks, ranging from 0.31 to 2.18 mg/L (EPA, 2019a). Where potential contamination bias of total nitrogen when comparing with EPA nutrient criteria concentrations may be a concern, an examination of geographic overlap of ecoregions (geographic location, concentration range, and time) with sites that have field blanks with detections may be helpful (Medalie and Shoda, 2020). Potential contamination bias of total particulate nitrogen in 2013 (parameter code 49570; Medalie and Shoda, 2020, plotfile 1) may be a reflection of variability found in filters used for blind blanks and in laboratory reagent blanks, which was identified by the NWQL and resulted in an increase in the RL in 2014 (Tedmund Struzeski, USGS, written commun. [from data quality assessment summary], April 2013; Douglas Stevenson, USGS, written commun. [from Rapi-Note 14-008], March 31, 2014).

Despite 11 detections in field blanks for each of the two types of ultraviolet-absorbance measurement (parameter codes 50624 and 61726) during the study period, potential concerns about contamination bias are negligible because for individual years, the 90-percent UCL for the 95th percentile concentration is either less than the RL or is not detected (table 4). Nearly all detections in environmental surface-water samples are greater than the RL and are at least an order of magnitude higher than the detections in field blanks (Medalie and Shoda, 2020, plotfile 1). There are no benchmarks for ultraviolet-absorbance measurements.

Three of the five carbon compounds in this study (particulate organic carbon, dissolved inorganic carbon, and total particulate carbon; parameter codes 00689, 00691, and 00694, respectively) have detection frequencies in field-blank samples greater than 40 percent (table 3). Eleven source-solution blanks for these three carbon compounds during the study period had detections between 0.1 and 0.5 mg/L (Medalie and Shoda, 2020, plotfile 1), which are similar concentrations to field blanks, presenting evidence that at least some of the carbon contamination in field blanks is from the source-solution IBW used to create the field blanks. The implication is that detections of carbon compounds in field blanks are unrelated to field activities and do not reflect potential field-related contamination of environmental samples, mitigating evidence of contamination bias in environmental samples based on field blank data.

A slightly high bias from laboratory carryover and instrument gas-flow issues may have affected environmental samples analyzed from about April through October 2014 for particulate inorganic carbon and total particulate carbon (parameter codes 00688 and 00694, respectively) and possibly other carbon compounds (Tedmund Struzeski, USGS, written commun. [from data quality assessment summary], October, 2014; this information is available upon request from the

USGS Inorganic Blind Sample Project at <https://qsb.usgs.gov/ibsp/>). This potential bias, identified from blind blanks with concentration greater than the DLs of 0.03 mg/L for particulate inorganic carbon or 0.05 mg/L for total particulate carbon, would only have affected low-level results and is not discernable in higher concentrations during 2014 compared with other years for these compounds.

Potential variability bias of nutrient, carbon, and sediment data based on field replicates.—Variability from field replicates is discussed in terms of variability in detection frequency and variability in concentration. Twenty out of 23 compounds in the nutrient, carbon, and sediment group had low variability in detection frequency, and 2 compounds had too few replicate pairs to assess variability (table 5). Total particulate inorganic carbon was the only compound in this group found to have high variability in detection frequency (table 5).

Two compounds (total particulate inorganic carbon and suspended-sediment concentration) in this group had high variability in concentrations for some part of the study period based on an annual mean relative standard deviation greater than 0.2 (table 6). These compounds are measures of particulate matter in water and thus are subject to natural variability from instream, sampling, or processing procedures (Gray and others, 2000). Total particulate inorganic carbon had high variability in concentrations for each year of the study period. The illustration of replicate pairs for this compound shows points plotting off the 1:1 line for all years at low and high concentrations (Medalie and Shoda, 2020, plotfile 2). High variability based on relative standard deviation in concentrations for suspended-sediment concentration analyzed with method SED10 during 2015 is likely an overestimate given that concentrations of all three replicate pairs in 2015 were in the low

range (less than 150 mg/L; table 6) where variability would have been more suitably assessed using standard deviation rather than relative standard deviation. Although variability in concentrations for suspended-sediment concentration appears to be high in 2014 (red circles in Medalie and Shoda, 2020, plotfile 2), points for these years that plot off the 1:1 line are at relatively high concentrations (greater than 150 mg/L) where relative standard deviation is the appropriate measure.

Integrated assessment of quality-control information for nutrients, carbon, and sediment.—Information from field blanks and field replicates is combined to present a snapshot of years and compounds with potential QC issues (table 7). Although some potential QC issues are identified, it is important to keep in mind the context of those issues on interpretations of environmental data. For example, none of the nutrients have highly variable data; potential contamination bias, which is present for five types of nitrogen results (ammonia, ammonia plus organic nitrogen as nitrogen, total Kjeldahl nitrogen, total particulate nitrogen, and total nitrogen), affects only the lowest concentrations. Total nitrogen is the only type of nutrient result with a benchmark (for 1 of 13 ecoregions) within the range of detections in field blanks and some environmental samples. Nitrate plus nitrite (analyzed with method RED02) and phosphorus (analyzed with method CL021) each have potential bias that is limited to a single year and low ranges of concentration. All other phosphorus and orthophosphate results are free of bias and high variability. None of the carbon compounds or ultraviolet-absorbance measurements have biased results. High variability was seen for total particulate inorganic carbon and suspended-sediment concentrations for 1 year, which might be partly an artifact of sampling for particulate substances.

Table 5. Estimated variability in detection frequency of field replicates for samples in the U.S. Geological Survey National Water Quality Network for water years 2013–17.

[This table is available for download at <https://doi.org/10.3133/sir20205116>]

Table 6. Estimated variability in concentrations of field replicates for samples in the U.S. Geological Survey National Water Quality Network for water years 2013–17.

[This table is available for download at <https://doi.org/10.3133/sir20205116>]

Table 7. Summary of results of data-quality assessment for inorganic compounds analyzed in surface-water samples of the U.S. Geological Survey National Water Quality Network based on all field quality-control sample types for water years 2013–17.

[This table is available for download at <https://doi.org/10.3133/sir20205116>]

Quality of Data for Major Ions and Trace Elements

Potential contamination bias of major ions and trace elements.—There is no environmentally relevant contamination bias from field-related activities for surface-water samples for several major ions and trace elements (calcium, magnesium, sodium, potassium, sulfate, strontium, and lithium; [table 3](#); Medalie and Shoda, 2020, plotfile 3). Additional censoring suggested for calcium, magnesium, sodium, sulfate, and strontium for all or most of the study period and for fluoride in 2014 (based on the censoring approach described in the “Statistical and graphical data analysis” section; [table 4](#)) is a hypothetical exercise because there are no environmental samples with concentrations less than the censoring threshold. Similarly, the additional censoring suggested for silica, at censoring thresholds ranging from 0.051 to 0.187 mg/L, is expected to affect less than 1 percent of surface-water samples (the 1st percentile of silica concentration in surface-water samples is 0.686 mg/L; [table 3](#)).

The relatively large number of random chloride detections in field blanks (Medalie and Shoda, 2020, plotfile 3) indicates that some contamination bias appears to warrant additional censoring of surface-water samples at a threshold between the RL of 0.12 mg/L and the annual 90-percent UCL for the 95th percentile concentration value for most of the study period ([table 4](#)); this suggests that censoring has the potential to affect more than the 1st percentile concentration (0.139 mg/L) of environmental samples collected during 2015 and 2017 ([table 3](#)). Potential contamination bias for fluoride, arsenic, iron, and vanadium is not pervasive throughout the indicated years ([table 4](#)) but, rather, occurs in clusters or discrete periods (Medalie and Shoda, 2020, plotfile 3) and affects very low environmental concentrations (except for vanadium in 2014).

Mixed evidence for whether the source of detections seen in the cluster of 16 detected field blanks for vanadium in 2014 is from source-solution water used for blank samples (which is not relevant to environmental samples) or from factors that would also affect environmental samples suggests that censoring should be considered in order to ensure that environmental results are not biased by contamination. In contrast, there is strong evidence that two clusters of boron contamination in field blanks (Medalie and Shoda, 2020; not reflected in [table 4](#) because these detections have no bearing on environmental samples) reflect two lots of contaminated source-solution water (NWIS lot 81202 tested June 1, 2012, with average boron concentration of 27.8 micrograms per liter [$\mu\text{g/L}$] and NWIS lot 81401 tested February 1, 2014, with average boron concentration of 5.4 $\mu\text{g/L}$; NWQL, USGS, written commun.,

January 10, 2020), meaning that environmental samples are unlikely to be contaminated. Source-solution water test results for vanadium for lot 81402 had concentrations that ranged from 0.07 to 0.12 $\mu\text{g/L}$ (NWQL, USGS, written commun., January 10, 2020); also in 2014, vanadium was detected in some samples of groundwater source-solution blanks, groundwater equipment blanks, and groundwater field blanks (Arnold and others, 2017) at concentrations similar to detections in NWQN surface-water field blanks. These results provide evidence that vanadium contamination affects field blanks but might not affect environmental samples. However, there were also three detections in QSB blind blanks in 2014 at similar concentrations as those seen in NWQN surface-water field blanks. Because the QSB uses different source solutions than that used for field blanks and by the NWQL, some vanadium contamination could also be related to laboratory processes, which could affect environmental sample results.

Censoring suggestions presented in this report as censoring data for an entire year at a time might be overly conservative if only part of the year is susceptible to the contamination bias (see for example, arsenic in 2016 in Medalie and Shoda, 2020, plotfile 3). Although beyond the scope of this report, an alternate approach is to censor for the part of the year that is susceptible to the contamination bias, for example, by applying a moving average of detections in blank samples (Fram and Belitz, 2011; Medalie and others, 2019).

Because the 90-percent UCLs for the 95th percentile concentration values for the five major ions and trace elements (fluoride, arsenic, boron, strontium, and selenium) with an HHB are all nondetections, there is negligible risk that contamination bias would affect concentrations near HHB values for these compounds in surface-water samples ([table 3](#)). No major ions have an ALB.

Potential variability bias of major ions and trace elements based on field replicates.—With variability assessed from pairs of field replicate samples, all major ions and trace elements had low variability in detection frequency according to criteria for inconsistent detections and mean detection rate ([table 5](#)). Similarly, no compound of this type had high variability in concentrations for any of the years of the study period ([table 6](#); Medalie and Shoda, 2020, plotfile 4).

Integrated assessment of quality-control information for major ions and trace elements.—All major ions and trace elements had low variability in detection frequency, and there were no indications of high variability in concentrations ([table 7](#)). All major ions and trace elements except potassium and lithium had some contamination bias for at least 1 year, although bias other than for chloride (and possibly for vanadium in 2016) was not at environmentally relevant concentrations. Most bias or variability for trace elements occurred in 2014 and 2016.

Quality of Data for Pesticides

Potential contamination bias of pesticides based on field blanks.—The low number of detections in field blanks is evidence that the risk of contamination bias for pesticides in NWQN samples is low. Twenty-three pesticides had at least 1 detection in a field blank, and 6 pesticides (metolachlor, hydroxymetolachlor, propoxur, siduron, tebuthiuron, and *trans*-permethrin) had 2 or more detections (table 3). The calculated 90-percent UCLs for the 95th percentile concentration results for these six compounds were nondetections. Five of the six compounds (all except the degradate hydroxymetolachlor) have at least one HHB or ALB. Maximum field-blank concentrations for metolachlor, propoxur, siduron, and tebuthiuron were more than 10 times less than the respective HHB or ALB. This was not the case for *trans*-permethrin, which had a maximum field-blank concentration that was double the concentration of the chronic invertebrate ALB. This apparent overlap in concentrations for *trans*-permethrin, however, needs to be qualified with the perspective that potential contamination bias is minimal, because detections were seen in only 0.8 percent of field blanks (3 of 358) during the study period. For all pesticides except propoxur, detections in field blanks were spread out in time whereby no more than one or two occurred in any year (Medalie and Shoda, 2020, plotfile 5). Whereas metolachlor had no more than two detections in field blanks for any year, seven of the nine detections in field blanks for propoxur occurred in 2017 (Medalie and Shoda, 2020, plotfile 5).

The timing and concentrations of detections of these six pesticide compounds indicate that only propoxur in 2017 might warrant additional censoring (up to 3.5 ng/L) to address potential contamination bias as indicated by field blanks in surface-water samples (to put this into perspective, propoxur had detections in 9, or 2.5 percent, of 358 field blanks during the study period; tables 3 and 4). Internal testing by the NWQL of the 20-milliliter vials used to collect samples for pesticide analysis revealed that propoxur was present at low concentrations in some tested vials. Detections in field blanks in 2017 indicate that some contaminated vials were likely used for field blanks and, by inference, also for environmental samples. Consequently, additional censoring of surface-water results at a threshold between the 90-percent UCL for the 95th percentile concentration value of 2.9 ng/L and the 2017 RL of 3.2 ng/L is suggested (table 4). Although the minimum concentration of surface-water samples (0.241 ng/L) was less than both these values, the 1st percentile concentration (0.877 ng/L) was greater than the RL, indicating that more than 99 percent of surface-water results are greater than either suggested censoring threshold and would be minimally affected by contamination bias.

Potential variability bias of pesticides based on field replicates.—Typically, 381 replicate pairs (the number of pairs ranged from 306 to 398) were available for the determination of variability bias for each pesticide between 2013

and 2017. Variability for each compound is assessed in two ways: as variability in detection frequency and as variability in concentration.

Forty-eight compounds (21 percent) met one or both criteria for low variability in pesticide detection frequency in replicate pairs with at least 1 detection; 169 pesticides (75 percent) had neither low nor high variability (table 5); and 8 pesticides had high variability in detections. These last eight pesticides (chlorsulfuron, alachlor oxanilic acid, deisopropyl prometryn, demethyl hexazinone B, hydroxyalachlor, 2-hydroxy-6-ethylamino-4-amino-*s*-triazine, sulfosulfuron, and alachlor sulfonic acid) had values that were greater than the criteria for both inconsistent detections and mean detection rate (at least 50 percent inconsistent detections or a mean detection rate of less than or equal to 75 percent). Where variability of detection frequency for these eight compounds (or for any other compound) is of concern, confidence limits can be calculated for the percentage of inconsistent detections (Martin, 2002). Because detections at concentrations near the reporting level might be prone to false-negative errors, estimates of detection frequencies of pesticides in environmental water samples are likely to be biased low (Martin, 2002).

There were 139 pesticides with at least 1 pair of detected results. Fifty-eight of these had enough replicate pairs with detected results to show a different pattern in variability between low and high concentrations, so that variability could be assessed using mean standard deviation at low concentrations and mean relative standard deviation at high concentrations (table 6). Variability for the other 81 pesticides with at least 1 pair of detected results was characterized by the mean relative standard deviation for all concentrations, which was typically based on a small number of replicate pairs and probably overestimated variability at lower concentrations (Mueller and others, 2015).

Variability in concentrations for most pesticides (170 of 225, or 76 percent) was moderate (table 6). The remaining 24 percent of pesticides (55 compounds) had high variability in concentrations for at least 1 year, assessed as relative standard deviation greater than 0.2 for pairs of field replicates. None of these compounds were highly variable in all 5 years, chlorpyrifos was the only compound that was highly variable during 4 years of the study period, and 39 of the 55 compounds were highly variable for just 1 year. These assessments of high variability based on relative standard deviation for all concentrations are likely to overestimate variability, especially when the dataset includes replicate pairs with low concentrations. For example, high variability is indicated for chlorsulfuron and dichlorvos in 2014 based on relative standard deviation greater than 0.2; however, each of these assessments is based on one data point (one replicate pair) in 2014 at a relatively low concentration, known to be overestimated by relative standard deviation. Therefore, these estimates of high variability (and other similar estimates) should be seen as very conservative. Another example is with the compound 1*H*-1,2,4-triazole (parameter 68498): where the data (table 6) show variability is high for 2014, a scatter plot of replicate

concentrations (Medalie and Shoda, 2020, plotfile 6) shows that the replicate concentrations for 2014 were in the low range of the two-range model (less than 500 ng/L), where high variability assessed by relative standard deviation is likely to be overestimated.

With replicate pairs from only 8 out of 225 compounds demonstrating high variability in detection frequency, overall variability in detection frequency is low. Variability in replicate pair concentrations is high in 78 year-and-compound combinations out of a possible 1,125 combinations (225×5), or about 7 percent of the time; however, as with the example of 1*H*-1,2,4-triazole, this is likely an overestimate. Information about variability from field replicates can be used to help interpret environmental results. Example applications of information about variability are (1) calculating the uncertainty in the concentration of an individual result from the standard deviation or the relative standard deviation, depending on whether the concentration falls below or above the breakpoint in slope; (2) using variability to estimate the probability that the true concentration in the sample exceeded a benchmark; and (3) constructing confidence intervals around concentrations, such as benchmarks, that are considered to be critical values (Mueller and Titus, 2005; Mueller and others, 2015).

Potential recovery bias of pesticides based on field matrix spikes.—The number of field matrix spikes per compound in the NWQN for the study period ranged from 42 to 274; most compounds had more than 200 results (table 8). Median percent recoveries ranged from 20 to 158. Comparing median recoveries with the method target range of between 70 and 130 percent, 22 compounds had median recoveries below 70 percent (table 8), which could indicate that environmental results are biased low, and three compounds (didealkylatrazine, diketonitrile isoxaflutole, and 2-hydroxy-6-ethylamino-4-amino-*s*-triazine) had recoveries greater than 130 percent (table 8), which could indicate that environmental results are biased high.

The NWQN has set a data-quality objective for F-pseudosigma, another measure of recovery bias, to be less than 30 percent. Eight compounds (bifenthrin, didealkylatrazine, chlorosulfonamide acid, diketonitrile isoxaflutole, fenbutatin oxide, hydroxyfluometuron, novaluron, and 2-hydroxy-6-ethylamino-4-amino-*s*-triazine) had an F-pseudosigma value

for field matrix spikes greater than 30 percent, indicating the potential for high variability in environmental concentrations (table 8). Five of these eight compounds (bifenthrin, didealkylatrazine, diketonitrile isoxaflutole, fenbutatin oxide, and novaluron) have at least one benchmark for which the comparison with environmental concentrations might be affected by high replicate-sample variability; one compound (2-hydroxy-6-ethylamino-4-amino-*s*-triazine) has no benchmark; and the remaining two compounds (chlorosulfonamide acid and hydroxyfluometuron) had no detections in environmental samples during the study period.

Integrated assessment of quality-control information for pesticides.—In this section, QC information from the NWQL and the QSB is combined with information from field QC samples for pesticide results. This discussion emphasizes the use of corroboration from multiple lines of evidence to note compounds with potential QC issues. For example, some blind-spike entries in the “Evidence of potential for contamination bias” column of table 9 are qualified with a note to indicate that potential contamination bias from third-party blind spikes (false-positive results for compounds that are not included in the spike mixture) is inconclusive evidence of bias because there is no corroboration from laboratory or field blanks. Explanations for these false-positive results in blind spikes could be that spiking issues or postspiking degradation occurred, both of which could create the appearance of contamination bias. Similarly, some blind-spike entries in the “Evidence of potential for overall high variability” column of table 9 are qualified with a note to indicate that although F-pseudosigma of these blind spikes is greater than 30 percent, this is deemed inconclusive evidence of overall high variability because there is no corroboration from laboratory spikes.

Potential for contamination bias based on integrated assessment.—In addition to field blanks, information on potential contamination bias is available from third-party blind spikes and from detections in laboratory blanks. The detailed interpretation from these types of samples done by Bexfield and others (2020b) is reproduced in table 9. Although 34 compounds had potential contamination bias based on blind spikes, 18 of these were qualified by Bexfield and others (2020b) and are likewise qualified with a note in table 9 because of lack of corroborating evidence of false positives

Table 8. Summary statistics for pesticide field matrix spikes in samples of the U.S. Geological Survey National Water Quality Network for water years 2013–17.

[This table is available for download at <https://doi.org/10.3133/sir20205116>]

Table 9. Summary of results of data-quality assessment for pesticides analyzed in samples of the U.S. Geological Survey National Water Quality Network based on all quality-control sample types for water years 2013–17.

[This table is available for download at <https://doi.org/10.3133/sir20205116>]

that could come from other QC samples or from NWQL indicators of interference. The remaining 16 compounds, as well as the 47 additional compounds indicated in table 9 as having potential contamination bias from laboratory blanks, might warrant further consideration for contamination bias using guidance described in this section.

Two considerations are offered to help data users sort out applicability of potential contamination bias to environmental data. First, the timing and relative magnitudes of detections in blank samples and environmental samples is an important part of this evaluation. Tools for the comparison are the first set of year-by-year columns in table 9 and time-series plots of detections of laboratory blanks, field blanks, and surface-water samples (Medalie and Shoda, 2020, plotfile 5). For example, there might be several years with potential concerns of contamination bias for compounds like 2,4-D, pyridaben, and etoxazole, as indicated by detections in laboratory blanks (table 9); however, reviewing the plotting of the data (Medalie and Shoda, 2020, plotfile 5) can alleviate most concerns because there is very little overlap of detections in blank and environmental samples. In some cases, such as for 2,4-D in 2014, the absence of reported detections in environmental samples (including all environmental and field QC results) demonstrates that the NWQL took a proactive censoring approach beyond set censoring to address widespread detections in laboratory blanks to reduce the occurrence of false positives.

In typical cases where detections in laboratory blanks are not as widespread as with 2,4-D and did not engender a proactive censoring approach by the NWQL, the NWQL censoring protocol is to censor environmental and field QC samples that would be reported as detections but are less than 3 times the concentration of detections in laboratory blanks (USGS, 2011). When environmental and QC samples need to be censored because concentrations are less than 3 times the concentration of detections in laboratory blanks, the standard RL is used if the concentration in the environmental sample is not greater than the standard reporting level, and an RRL is used if the concentration in the environmental sample is greater than the standard RL (USGS, 2011; Shoda and others, 2018). In cases where an RRL is used, the RRL is the concentration measured in the environmental or QC sample, and the result is reported as less than the RRL. Censored results reported with RRLs are conservative because they are designed to reduce the occurrence of false-positive results. This censoring protocol applies only to environmental samples in the same set as detections in laboratory blanks and does not address random laboratory contamination that may occur in proximate sets (Medalie and others, 2019). For compounds such as fipronil in 2013–14 and triclopyr in 2014, which had detections of laboratory blanks and surface-water results in proximate sets (Medalie and Shoda, 2020, plotfile 5), data users may choose to apply an additional censoring schema, such as a fixed censoring threshold or a moving average detection frequency in a specified number of laboratory blanks (Fram and Belitz, 2011; Bexfield and others, 2019; Fram and Stork, 2019; Medalie and others, 2019) for the period affected based on project-specific data-quality objectives.

The second consideration addresses relevancy by comparing concentrations of QC and environmental results to ALBs and HHBs. For example, six pesticides with very low ALBs (bifenthrin, dichlorvos, diflubenzuron, fipronil, and *cis*- and *trans*-permethrin), all with minimum surface-water concentrations less than the ALB, also have maximum laboratory-blank concentrations greater than the respective ALB. Although the number of environmental results that are detections less than or about equal to these ALBs is very small (fewer than 1 percent of results are detected in surface water [table 3; the 1st percentile is a nondetection] for all these compounds except fipronil), these results should be assessed for relevancy (in timing and magnitude) using information available in table 9 and Medalie and Shoda (2020, plotfile 5).

A caveat to the comparison of concentrations in surface water to ALBs for these compounds relates to the lowest ALB for these six compounds being the chronic invertebrate benchmark. This measure of an exposure duration is typically calculated as a 21-day moving average pesticide concentration, which would be a lower value than any instantaneous environmental or QC result presented as a comparison in this report when calculated (EPA, 2017). However, the relative value of the low chronic invertebrate ALB and the environmental or laboratory blank concentration being compared with the ALB is not as important for purposes of this report as the general idea that some environmental concentrations near relevant benchmarks might have some contamination bias.

The only pesticide and associated year potentially affected by contamination bias based on field blanks was propoxur in 2017 (“Potential contamination bias of pesticides based on field blanks” section of this report). An analysis adding laboratory blank and blind spike data to the evaluation (Bexfield and others, 2020b) shows that propoxur is also susceptible to contamination bias in 2013 and 2016 (table 9). Because no surface-water field blanks had detections of propoxur in 2013, we might infer that the NWQL sufficiently addressed potential laboratory contamination identified in the 2013 data (field QC samples are subject to the same censoring protocols by the NWQL as environmental samples) and that surface-water results for 2013 are likely free of contamination bias. This conclusion is generally supported by Medalie and Shoda (2020, plotfile 5), which shows few surface-water detections that coincide with the cluster of laboratory blank detections in 2013. Thus, the integrated picture of contamination for propoxur shows that the only contamination biases of concern to NWQN data users occurred during 2016 and 2017.

Potential for recovery bias based on integrated assessment.—Information about recovery of compounds from spike mixtures added to QC samples is used primarily to assess biases stemming from method performance and matrix or degradation effects. Recovery information from each of the three different types of samples (field matrix spikes, laboratory spikes, and third-party spikes) tells a different story. Put together, this information can either present multiple lines of consistent evidence to identify QC issues or biases or it can present inconsistent evidence used to qualify or discount identification of QC issues. Field matrix spikes can assess

biases from the water matrix, degradation, and the laboratory environment or analysis. Laboratory spikes can assess biases from the laboratory environment or analysis. Third-party blind spikes (blind spikes) can assess biases from degradation and the laboratory environment or analysis.

Twenty-seven of 225 pesticide compounds in schedule 2437 had at least one issue with recovery bias (table 9; unqualified [none of these 27 pesticide compounds included a note to indicate that there might be a problem with the low or high recovery value itself]). None of the laboratory spikes had spike recoveries outside of acceptable limits. Seven compounds (malaoxon, lactofen, naled, phorate oxon sulfone, terbufos oxon, and *cis*- and *trans*-permethrin) had recoveries of less than 70 percent for field matrix spikes and for blind spikes, but not for laboratory spikes, indicating that the method performance is likely good but there might be matrix effects (which could affect field matrix spikes but not blind or laboratory spikes) or degradation (which could affect field matrix or blind spikes; these types of spikes have a relatively long lag time between sample collection and preparation compared with laboratory spikes, which are prepared and analyzed in a shorter time frame) affecting recoveries. Such instances of corroborating evidence where the same conclusion is drawn from two different types of QC samples present the strongest cases for the presence of QC issues. Other reasons that there might be differences between blind spikes and laboratory spikes is that the spike mixtures are different, both in composition and concentration, whereas the spike mixtures are the same for field matrix spikes and laboratory spikes.

This evaluation places less emphasis on recoveries outside of limits for blind spikes than for other types of spikes because of the difficulty in determining whether recoveries for blind spikes are from degradation or from a process specific to the spike mixture used exclusively for blind spikes (the QSB prepares spike mixtures for blind spikes, and the NWQL prepares spike mixtures for field matrix and laboratory spikes). To resolve this difficulty, analyses should provide corroborating evidence where two or more types of QC samples point to the same conclusion. Most of the compounds that have evidence for substantial recovery bias (table 9) have that bias consistently for 2014–17. Bifenthrin, naled, novaluron, lactofen, 2-(1-hydroxyethyl)-6-methylaniline (HEMA), fenbutatin oxide, asulam, and 1*H*-1,2,4-triazole have been permanently labeled with an “E” code, meaning these compounds are known to have high variability in recovery. All except HEMA have at least one type of benchmark.

Potential for high variability based on integrated assessment.—High variability based on this integrated assessment was determined from field replicates (variability of detection frequencies or variability of concentrations), laboratory spikes, and blind spikes (variability for both types of spike samples refers to variability in recovery as assessed by F-pseudostigma). Based on laboratory spike analysis, fenbutatin oxide and naled samples showed high variability of recovery, and based on blind spike analysis, samples for 24 pesticides showed high variability of recovery. This

dichotomy in variability between the two types of spikes when consistency might have been expected was investigated by Bexfield and others (2020b), who determined that substantially higher variability in recovery for blind spikes than for laboratory spikes might be an effect either from spike mixture degradation or from a discrepancy in concentrations in spike solutions (blind spikes are spiked at a range of concentrations, whereas laboratory spikes are always spiked at the same concentration). Because higher variability in blind spikes likely reflects either of these two effects, which are not directly related to environmental sample variability, all cases of high variability in blind spikes in table 9 that do not show concomitant high variability in laboratory spikes are qualified with a note. For compounds with this inconsistent metric for variability, blind spikes qualified with a note are not counted as evidence of potential for overall high variability in the summation of QC issues in this section; however, they are left in table 9 in order to be consistent with information in Bexfield and others (2020b).

Disregarding high variability indicated only by footnoted blind spikes, 61 pesticide compounds had a high variability indication in table 9 based on field replicate information (same as tables 4 and 5), 1 was indicated from laboratory spike information (fenbutatin oxide), and 1 was indicated from laboratory spike, blind spike, and field replicate information (naled). Although fenbutatin oxide and naled have indications of high variability and both have a chronic ALB and an HHB, both compounds are rarely detected (less than 1 percent of results are detections) in surface water at NWQN sites (table 3; Medalie and Shoda, 2020, plotfile 5). Most of the high variability from field replicates (from inconsistent detections and concentrations) was in 2013, 2014, and 2015 (table 9). All but 15 of these 58 compounds have some kind of benchmark.

Overall integrated assessment.—Eighty-seven pesticide or degradate compounds have at least one QC issue (table 9); all except 33 of these compounds have at least 1 benchmark. The least common QC issue, affecting 36 compounds, was recovery bias. Thirty-three compounds were affected by 2 potential QC issues, with the combination of contamination bias and overall high variability being the most common; all but 2 of these compounds have at least 1 benchmark. Eight compounds affected by two QC issues (chlorpyrifos, linuron, naled, novaluron, pyridaben, tebufos, and *cis*- and *trans*-permethrin) have a chronic invertebrate benchmark of less than 100 ng/L and thus need particularly careful consideration of timing and magnitude of QC results in relation to surface-water results before interpretive use.

As part of this evaluation of using QC information to help interpret environmental data, it is important to determine the intended use of the environmental data. The presence of one or more biases or of high variability and the magnitude of the bias or variability could come into play for determining acceptability of data for use. For example, to determine time trends of environmental data, the variability of the dataset needs to be small enough so that any trend signal is not obscured by noise and any changes in recovery or other biases

cannot account for the changes in environmental data over time. Surveys of data that are casting a wide net for occurrence of pesticides in environmental water need data with sufficient recoveries and low false-negative rates to detect the presence of pesticides that are truly in the water. Water-quality results that are meant for comparisons to benchmarks, especially low benchmarks, need thorough assessments of potential contamination bias at concentrations near concentrations of benchmarks, need to have sufficient recoveries, and need low false-negative and low false-positive rates to accomplish the intended purpose.

Summary

The quality of water-quality data at 122 sites in the U.S. Geological Survey National Water Quality Network for water years 2013–17 was assessed in an integrated process. All compounds routinely sampled by the National Water Quality Network and analyzed at the National Water Quality Laboratory were included. Summaries of field quality-control (QC) results (field blanks, field replicates, and field matrix spikes) for all compounds were produced.

QC evaluations for inorganic compounds used field blanks to assess contamination bias. Six of the 13 types of nutrient results—nitrite, nitrate plus nitrite (method RED01), phosphorus (methods KJ009, CL020, and KJ005), and orthophosphate—had no contamination bias. Nitrate plus nitrite (method RED02) and phosphorus (method CL021) demonstrated potential contamination bias for 1 year for samples at low concentrations near the reporting level. Five types of nutrient results—ammonia, ammonia plus organic nitrogen, total Kjeldahl nitrogen, total particulate nitrogen, and total nitrogen—demonstrated potential contamination bias potentially affecting low-level concentrations for 4 or 5 years; however, total nitrogen was the only one of these nutrients with a benchmark low enough to be within range of environmental data. None of the carbon compounds had biased results because detections in field blanks were traced to contaminated source solutions. Contamination bias was not found for ultraviolet-absorbance measurements and was not assessed for sediment. All major ions and trace elements except potassium and lithium had some contamination bias for at least 1 year, although those biases, except for chloride and possibly for vanadium in 2016, were not at environmentally relevant concentrations.

QC evaluations for inorganic compounds used field replicates paired with environmental samples to assess variability in detection frequency and variability in concentrations. With two exceptions, all compounds with a sufficient number of field replicate samples in the nutrient, carbon, and sediment group and in the major ions and trace elements group had low variability both in detection frequency and in concentration.

Exceptions were for total particulate inorganic carbon and for sediment for 2015, both of which are particulate substances with intrinsically high sampling variability.

Very few detections in field blanks is evidence that the risk of contamination bias for National Water Quality Network samples is low. No contamination bias was found for 219 of 225 pesticides, and limited potential contamination bias was found for 6 pesticides (metolachlor, hydroxymetolachlor, propoxur, siduron, tebuthiuron, and *trans*-permethrin). The maximum field-blank concentration for *trans*-permethrin was double the concentration of the chronic invertebrate aquatic-life benchmark. Seven of the nine detections of propoxur in field blanks were in 2017, likely from contaminated vials, which also could have affected environmental samples.

An additional assessment of pesticide contamination bias was based on the integration of information from field blanks, third-party blind spikes, and laboratory blanks. Sixteen compounds showed potential contamination bias based on blind spikes, and 47 different compounds showed potential contamination bias from laboratory blanks. Timing and relative magnitudes of detections in blank samples and environmental samples as well as environmental relevancy are important considerations when evaluating potential contamination bias.

Overall variability in detection frequency for pesticides from field replicates was low or moderate: 48 pesticides (21 percent) were characterized with low variability, 169 pesticides (75 percent) were moderate with neither low nor high variability, and 8 out of 225 compounds (4 percent) had high variability in detection frequency. Also based on field replicates, 55 pesticides had overall high variability in concentrations for at least 1 year, with 39 of the 55 being highly variable for just 1 year, although these assessments likely overestimate high variability.

High variability based on the integrated assessment was determined from field replicates, laboratory spikes, and blind spikes. Disregarding high variability indicated only by blind spikes (which might be reflecting a degradation or concentration effect), 61 out of 69 pesticide compounds with high variability were indicated as having high variability from field replicates, 1 from laboratory spikes (fenbutatin oxide), and 1 from laboratory spikes and blind spikes (naled). Although fenbutatin oxide and naled had indications of high variability and both had a chronic aquatic-life benchmark and a human-health benchmark, both compounds were rarely detected in surface water at National Water Quality Network sites.

Median percent recoveries in field matrix spikes ranged from 20 to 158 for pesticides. Median recoveries for 22 compounds were less than 70 percent, an indication that environmental results might be biased low, and were greater than 130 percent for 3 compounds, an indication that environmental results might be biased high. Eight compounds—bifenthrin, didealkylatrazine, chlorosulfonamide acid, diketonitrile isoxaflutole, fenbutatin oxide, hydroxyfluometuron, novaluron, and 2-Hydroxy-6-ethylamino-4-amino-*s*-triazine—had an F-pseudosigma value for field matrix spikes above 30 percent, indicating the potential for high variability in

recovery of environmental concentrations. Five of these eight compounds—bifenthrin, didealkylatrazine, diketonitrile isoxaflutole, fenbutatin oxide, and novaluron—had at least one aquatic-life benchmark or human-health benchmark.

Information from laboratory spikes and third-party blind spikes was used to augment information about recovery bias from field matrix spikes. Although none of the laboratory spikes had pesticide recoveries outside of acceptable limits, 30 pesticides had recoveries outside of acceptable limits for either field matrix or blind spikes. Seven compounds—malaoxon, lactofen, naled, phorate oxon sulfone, terbufos oxon, *cis*- and *trans*-permethrin—had recoveries of less than 70 percent for both of these types of spikes, indicating that the method performance is likely good but there might be degradation or a matrix effect impacting recoveries. All results for bifenthrin, naled, novaluron, lactofen, 2-(1-hydroxyethyl)-6-methylaniline (HEMA), fenbutatin oxide, asulam, and 1*H*-1,2,4-triazole were flagged by the NWQL because they are known to have high variability in recovery.

Overall, 87 pesticides had at least 1 QC issue, although most of these issues had little or no effect on the interpretation of environmental results because (1) the National Water Quality Laboratory addressed the QC issue before publishing the environmental results (like 2,4-D in 2014 and propoxur in 2013), (2) environmental results were almost entirely nondetections (like novaluron), (3) concentrations of environmental results were higher than potential contamination bias (like etoxazole), or (4) benchmark concentrations were orders of magnitude higher than all environmental results (such as for metolachlor and siduron). Eight compounds—chlorpyrifos, linuron, naled, novaluron, pyridaben, tebufirimfos, *cis*- and *trans*-permethrin—affected by two QC issues had an aquatic-life benchmark less than 100 ng/L and thus warrant particularly careful consideration of timing and magnitude of QC results in relation to surface-water results before interpretive use.

Acknowledgments

The authors would like to thank the numerous U.S. Geological Survey (USGS) personnel who provided the environmental and field quality-control data summarized and described in this report—they took meticulous care to follow standard protocols for collection, processing, shipping, and coding water samples; and for reviewing water-quality results. We also thank analysts at the USGS National Water Quality Laboratory and the Quality Systems Branch who provide high-quality data and are always available to help hydrologists understand what goes on “behind the scenes” in the laboratory during production of water-quality results. Appreciation is extended to Mark Sandstrom and Lisa Nowell, both with the USGS, for help with refining the interpretation of quality-control data.

References Cited

- Arnold, T.L., Bexfield, L.M., Musgrove, M.L., Lindsey, B.D., Stackelberg, P.E., Barlow, J.R., DeSimone, L.A., Kulongoski, J.T., Kingsbury, J.A., Ayotte, J.D., Fleming, B.J., and Belitz, K., 2017, Datasets from groundwater-quality data from the national water-quality assessment project, January through December 2014 and select quality-control data from May 2012 through December 2014: U.S. Geological Survey data release, accessed May 23, 2020, at <https://doi.org/10.5066/F7W0942N>.
- Bender, D.A., Zogorski, J.S., Mueller, D.K., Rose, D.L., Martin, J.D., and Brenner, C.K., 2011, Quality of volatile organic compound data from groundwater and surface water for the national water-quality assessment program, October 1996–December 2008: U.S. Geological Survey Scientific Investigations Report 2011–5204, 128 p. [Also available at <https://doi.org/10.3133/sir20115204>.]
- Bexfield, L.M., Belitz, K., Sandstrom, M.W., Beaty, D., Medalie, L., Lindsey, B.D., and Nowell, L.H., 2020b, Quality of pesticide data for groundwater analyzed for the National Water-Quality Assessment Project, 2013–18: U.S. Geological Survey Scientific Investigations Report 2020–5072, accessed July 27, 2020, at 35 p., <https://doi.org/10.3133/sir20205072>.
- Bexfield, L.M., Sandstrom, M.W., and Beaty, D., 2020a, Field, laboratory, and third-party data for assessment of the quality of pesticide results reported by the National Water Quality Laboratory for groundwater samples collected by the National Water-Quality Assessment Project, 2013–18: U.S. Geological Survey data release, <https://doi.org/10.5066/P90BFKA4>.
- Bexfield, L.M., Toccalino, P.L., Belitz, K., Foreman, W.T., and Furlong, E.T., 2019, Hormones and pharmaceuticals in groundwater used as a source of drinking water across the United States: *Environmental Science & Technology*, v. 53, p. 2950–2960. [Also available at <https://doi.org/10.1021/acs.est.8b05592>.]
- Brown, L.D., Cai, T.T., and DasGupta, A., 2001, Interval estimation for a binomial proportion: *Statistical Science*, v. 16, no. 2, p. 101–133. [Also available at <https://doi.org/10.1214/ss/1009213286>.]
- Deacon, J.R., Lee, C.J., Norman, J.E., and Reutter, D.C., 2017, Nutrient and pesticide data collected from the USGS National Water Quality Network and previous networks, 1963–2016: U.S. Geological Survey data release, accessed May 23, 2020, at <https://doi.org/10.5066/F73777K4>.

- Fram, M.S., and Belitz, K., 2011, Occurrence and concentrations of pharmaceutical compounds in groundwater used for public drinking-water supply in California: *Science of the Total Environment*, v. 409, no. 18, p. 3409–3417. [Also available at <https://doi.org/10.1016/j.scitotenv.2011.05.053>.]
- Fram, M.S., Olsen, L.D., and Belitz, K., 2012, Evaluation of volatile organic compound (VOC) blank data and application of study reporting levels to groundwater data collected for the California GAMA priority basin project, May 2004 through September 2010: U.S. Geological Survey Scientific Investigations Report 2012–5139, 94 p. [Also available at <https://doi.org/10.3133/sir20125139>.]
- Fram, M.S., and Stork, S.V., 2019, Determination of study reporting levels for pesticide data for the California Groundwater Ambient Monitoring and Program priority basin project, 2004–2018—Part 1—National water quality schedules 2003, 2032, or 2033, and 2060: U.S. Geological Survey Scientific Investigations Report 2019–5107, 129 p., accessed May 23, 2020, at <https://doi.org/10.3133/sir20195107>.
- Gray, J.R., Glysson, G.D., Turcios, L.M., and Schwarz, G.E., 2000, Comparability of suspended-sediment concentration and total suspended solids data: U.S. Geological Survey Water-Resources Investigations Report 00–4191, 20 p. [Also available at <https://doi.org/10.3133/wri004191>.]
- Hoaglin, D.C., Mosteller, F., and Tukey, J.W., eds., 1983, *Understanding robust and exploratory data analysis*: New York, John Wiley, 447 p.
- Lee, C., and Reutter, D., 2019, Nutrient and pesticide data collected from the USGS National Water Quality network and previous networks, 1963–2018: U.S. Geological Survey data release, accessed June 11, 2020, at <https://doi.org/10.5066/P94F31R8>.
- Martin, J.D., 2002, Variability of pesticide detections and concentrations in field replicate water samples collected for the National Water-Quality Assessment Program, 1992–97: U.S. Geological Water-Resources Investigations Report 01–4178, 102 p., accessed September 3, 2019, at <https://doi.org/10.3133/wri20014178>.
- Martin, J.D., Stone, W.W., Wydoski, D.S., and Sandstrom, M.W., 2009, Adjustment of pesticide concentrations for temporal changes in analytical recovery, 1992–2006: U.S. Geological Survey Scientific Investigations Report 2009–5189, 23 p., apps., accessed May 23, 2020, at <https://doi.org/10.3133/sir20095189>.
- Medalie, L., and Martin, J.D., 2016, Nutrient and pesticide contamination bias estimated from field blanks collected at surface-water sites in U.S. Geological Survey water-quality networks, 2002–12: U.S. Geological Survey Scientific Investigations Report 2016–5129, 40 p., accessed September 3, 2019, at <https://doi.org/10.3133/sir20165129>.
- Medalie, L., Sandstrom, M.W., Toccalino, P.L., Foreman, W.T., ReVello, R.C., Bexfield, L.M., and Riskin, M.L., 2019, Use of set blanks in reporting pesticide results at the U.S. Geological Survey National Water Quality Laboratory, 2001–15: U.S. Geological Survey Scientific Investigations Report 2019–5055, 147 p., accessed September 3, 2019, at <https://doi.org/10.3133/sir20195055>.
- Medalie, L., and Shoda, M.E., 2020, Field blank and field replicate datasets for inorganic and organic compounds collected for the National Water Quality Network, water years 2013–17: U.S. Geological Survey data release, <https://doi.org/10.5066/P96VY980>.
- 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., accessed September 3, 2019, at <https://doi.org/10.3133/tm4C4>.
- Mueller, D.K., and Titus, C.J., 2005, Quality of nutrient data from streams and ground water sampled during water years 1992–2001: U.S. Geological Survey Scientific Investigations Report 2005–5106, 27 p., accessed September 3, 2019, at <https://doi.org/10.3133/sir20055106>.
- Norman, J.E., Toccalino, P.L., and Norman, S.A., 2018, Health-based screening levels for evaluating water-quality data (updated May 2018): U.S. Geological Survey data release, accessed September 3, 2019, at <https://doi.org/10.5066/F71C1TWP>.
- Nowell, L.H., Ludtke, A.S., Mueller, D.K., and Scott, J.C., 2013, Organic contaminants, trace and major elements, and nutrients in water and sediment sampled in response to the Deepwater Horizon oil spill: U.S. Geological Survey Scientific Investigations Report 2012–5228, 96 p., apps., accessed September 3, 2019, at <https://doi.org/10.3133/sir20125228>.
- Riskin, M.L., Reutter, D.C., Martin, J.D., and Mueller, D.K., 2018, Quality-control design for surface-water sampling in the National Water-Quality Network: U.S. Geological Survey Open-File Report 2018–1018, 15 p., accessed September 3, 2019, at <https://doi.org/10.3133/ofr20181018>.
- Sandstrom, M.W., Kanagy, L.K., Anderson, C.A., and Kanagy, C.J., 2015, Determination of pesticides and pesticide degradates in filtered water by direct aqueous-injection liquid chromatography-tandem mass spectrometry: U.S. Geological Survey Techniques and Methods, book 5, chap. B11, 54 p., accessed September 3, 2019, at <https://doi.org/10.3133/tm5B11>.

- Shoda, M.E., Nowell, L.H., and Bexfield, L.M., 2017a, National water-quality assessment project replicate surface water and groundwater pesticide data analyzed by the USGS National Water Quality Laboratory schedule 2437, water years 2013–15: U.S. Geological Survey data release, accessed September 3, 2019, at <https://doi.org/10.5066/F75H7DS8>.
- Shoda, M.E., Nowell, L.H., Bexfield, L.M., Sandstrom, M.W., and Stone, W.W., 2017b, Recovery data for surface water, groundwater and lab reagent samples analyzed by the USGS National Water Quality Laboratory schedule 2437, water years 2013–15: U.S. Geological Survey data release, accessed September 3, 2019, at <https://doi.org/10.5066/F7QZ28G4>.
- Shoda, M.E., Nowell, L.H., Stone, W.W., Sandstrom, M.W., and Bexfield, L.M., 2018, Data analysis considerations for pesticides determined by National Water Quality Laboratory schedule 2437: U.S. Geological Survey Scientific Investigations Report 2018–5007, 458 p., accessed September 3, 2019, at <https://doi.org/10.3133/sir20185007>.
- U.S. Environmental Protection Agency, [EPA], 2017, Overview of risk assessment in the pesticide program: U.S. Environmental Protection Agency web page, accessed December 17, 2019, at <https://www.epa.gov/pesticide-science-and-assessing-pesticide-risks/overview-risk-assessment-pesticide-program>.
- U.S. Environmental Protection Agency, [EPA], 2018a, 2018 edition of the drinking water standards and health advisories: U.S. Environmental Protection Agency report EPA 822–F–18–001, 20 p., accessed October 24, 2019, at <https://www.epa.gov/sites/production/files/2018-03/documents/dwtable2018.pdf>.
- U.S. Environmental Protection Agency, [EPA], 2018b, Aquatic life benchmarks and ecological risk assessments for registered pesticides (updated October 24, 2018): U.S. Environmental Protection Agency web page, accessed December 10, 2018, at <https://www.epa.gov/pesticide-science-and-assessing-pesticide-risks/aquatic-life-benchmarks-and-ecological-risk>.
- U.S. Environmental Protection Agency, [EPA], 2019a, Ecoregional criteria: U.S. Environmental Protection Agency web page, accessed October 24, 2019, at <https://www.epa.gov/nutrient-policy-data/ecoregional-criteria>.
- U.S. Environmental Protection Agency, [EPA], 2019b, Aquatic life criteria—Ammonia: U.S. Environmental Protection Agency web page, accessed October 24, 2019, at <https://www.epa.gov/wqc/aquatic-life-criteria-ammonia>.
- U.S. Geological Survey, [USGS], 2011, Application of the result-level ‘v’ value qualifier code and ‘E’ remark code to selected organic results reported by the National Water Quality Laboratory (NWQL): U.S. Geological Survey Office of Water Quality Laboratory Technical Memorandum 2012.01, 4 p., accessed December 8, 2017, at <https://water.usgs.gov/admin/memo/QW/qw12.01.pdf>.
- U.S. Geological Survey, [USGS], 2015, Retrospective review and guidance for interpretation of results from National Water Quality Laboratory custom method (LC2437) “Determination of pesticides and pesticide degradates by direct aqueous injection liquid chromatography-mass spectrometry (DAI LC–MS/MS)”: U.S. Geological Survey Office of Water Quality Technical Memorandum 2015.04, 3 p., accessed November 1, 2019, <https://water.usgs.gov/admin/memo/QW/qw2015.04.pdf>.
- U.S. Geological Survey, [USGS], 2018, Inorganic blind sample project: U.S. Geological Survey web page, accessed September 3, 2019, at https://qsb.usgs.gov/ibsp/archive/2018/Blind%20Blank%20Charts/BBPCat_an.html.
- U.S. Geological Survey, [USGS], 2019, National field manual for the collection of water-quality data: U.S. Geological Survey Techniques and Methods, book 9, chaps. A1–A10, [variously paged], accessed September 3, 2019, at <https://water.usgs.gov/owq/FieldManual/index.html>.
- U.S. Geological Survey [USGS], 2020, USGS water data for the nation: U.S. Geological Survey National Water Information System database, accessed June 24, 2020, at <https://doi.org/10.5066/F7P55KJN>.
- Wieben, C.M., 2019, Pesticide recovery data for surface-water and lab reagent samples analyzed by the USGS National Water Quality Laboratory schedule 2437, water years 2016–17: U.S. Geological Survey data release, accessed September 3, 2019, at <https://doi.org/10.5066/P93MWMVF>.

