

Chapter

Performance Measures for Evaluating and Communicating Data Quality in Aquatic Environmental Monitoring in the U.S.

James Stribling, Susan Lanberg, Blaine Snyder, Jerry Diamond, Benjamin Jessup and Hugh Sullivan

Abstract

Quantitative data quality descriptors are important for evaluating and communicating acceptability of information used in environmental monitoring and assessment. In this chapter, we present (1) the rationale for establishing and using performance measures and MQOs in routine quality control planning and analysis, (2) field and laboratory methods for capturing input data required for performance calculations, and (3) approaches for setting data acceptability thresholds and determining the need for corrective actions. Relevant examples are available from local, regional, and national programs in the U.S. charged with monitoring and assessing aquatic biological condition, physical habitat, contaminants, and toxicity testing. We will describe techniques for calculating and determining acceptability of performance measures, such as, among other data quality indicators, precision, accuracy, sensitivity, representativeness, and completeness of field sampling, laboratory processing, and data management and analysis. Data types on which these examples will be based include benthic macroinvertebrates, fish assemblage, tissue body burden, laboratory analytical and toxicity testing, physical habitat, selected geomorphic characteristics, and algal toxins.

Keywords: precision, bias, indicators, error, corrective actions, acceptability

1. Introduction

Science is recognized as treating uncertainty and variability as information that serves as a key component of decision-making, helping formulate new questions, experimental designs, and testing and measurement procedures [1–3]. This is the essence of the scientific method; knowledge itself increases through the process of trial and error. Perhaps the most well-known effort to begin quantifying data variability as part of the decision-making process in technical endeavors led to development

of the concept of statistical process control [4]. With a focus on manufacturing, Shewhart's ideas largely originated with the desire to better understand causes of anomalous or unwanted output, and to provide clarity on what might be necessary to improve outcomes. He stated [4] "*Through the use of the scientific method...it has been found possible to set up limits within which the results of routine efforts must lie if they are to be economical. Deviations in the results of a routine process outside such limits indicate that the routine has broken down and will no longer be economical until the cause of the trouble has been removed*". Shewhart [5] further developed statistical techniques and demonstrated their application, helping to broaden the appeal of using control charts to document and track the quality of various data characteristics. The quality of data, and especially environmental data, is tracked through various quality control (QC) processes as discussed in this chapter. As Woodall [6] and others have noted, QC analyses and their interpretation are best handled by those who are knowledgeable about the field of practice.

The 1993 passing of the Government Performance and Results Act (GPRA) in the United States (US) elevated attention to documenting effectiveness, efficiency, and accountability of programs, and resulted in Federal agencies setting goals for program performance [7]. The GPRA focused the need for and use of performance characteristics and quantitative measurement quality objectives (MQO) to strengthen programs at any scale, not just Federal. The Information Quality Act (IQA) of 2000, sometimes referred to as the Data Quality Act, required Federal programs to ensure the "quality, objectivity, utility, and integrity" of publicly available information they produced [8]. It also required agencies to develop techniques for acquiring, reporting, and acting on results, where necessary.

The concept of process or QC is adaptable to any measurement system, requiring only that key points of the process are identified as providing opportunity for taking measurements, and that there is some standard or criterion for comparison. When anomalous or extreme results are detected via comparison to MQO, they would be investigated to determine what might be causing performance deviations.

Routine environmental monitoring requires consistent collection of data and information such that they are of known and acceptable quality. The purpose of this chapter is to describe data requirements, numeric structure, and interpretation thresholds for MQO related to several diverse and important indicators used in aquatic environmental monitoring throughout the US. These include biological, physical, chemical, and toxicological indicators.

2. Quality assurance and control for environmental monitoring

In the US, environmental data are collected by many federal, state, tribal, and local agencies, including the US Environmental Protection Agency (EPA), as well as non-EPA organizations supporting environmental programs on behalf of EPA in accordance with the EPA agency-wide quality system. Other federal agencies such as the US Geological Survey (USGS), the National Oceanographic and Atmospheric Administration (NOAA), and the US Fish and Wildlife Service (USFWS) collect data under different data quality frameworks but have similar requirements for known and acceptable quality. The quality assurance (QA) planning processes established by the EPA are recognized as a high standard that should be *attempted* even in non-EPA projects, such as state-, industry-, or non-profit-funded special projects. The EPA quality system is based on ANSI/ASQC E4-1994, *Specifications and Guidelines*

for *Quality Systems for Environmental Data Collection and Environmental Technology Programs*, a national standard for quality management practices for environmental programs involving the collection and evaluation of environmental data and the design, construction, and operation of environmental technologies [9, 10]. Quality planning documentation prepared for collection of environmental data (by or for EPA) includes descriptions of project-specific data quality objectives (DQO), quality assurance project plans (QAPP), and standard operating procedures (SOP). DQO are integral to the QA planning process. The DQO process includes identifying the decisions to be made based on the information collected, as well as the data quality and quantity acceptance criteria required to make those project decisions [11]. QAPPs are developed and implemented to ensure that data collected for a project are complete and of a quality sufficient for their intended purpose [12]. A QAPP includes a section on DQO, and SOP for relevant field collection and laboratory analysis procedures are often included as QAPP attachments. SOP are developed and followed to ensure that procedures for data collection and analysis are performed consistently within boundaries defined by MQO, thus meeting acceptance criteria.

There are different sources of error, some of which may yield uncertainty and all of which can affect variability observed in data and outcomes. This chapter discusses several commonly used indicators of aquatic environmental condition and the types of performance measures and QC processes used to ensure that data are acceptable to use in a particular environmental program.

3. Indicators of environmental condition

3.1 Biological

Field sampling, laboratory processing, and data analysis procedures for biological indicators are relatively well-established for many monitoring programs. For programs focused on community level indicators of biological integrity, such as the Index of Biological Integrity (IBI) or River Invertebrate Prediction and Classification System (RIVPACS) of observed to expected (O/E) conditions, based on consistent sampling and interpreting of taxa and individual counts. Field sampling for these indicators typically gathers composite samples from multiple habitats distributed throughout some defined area of the stream, river, lake, or estuarine/near-coastal waters. Depending on the program, the sampling area for rivers and streams can be a defined channel length, such as 100 m, or some multiple of the wetted width. Organism groups targeted by this kind of sampling includes, for example, benthic macroinvertebrates (BMI), fish, and algae/diatoms. Laboratory processing for BMI and diatoms includes sorting, subsampling, and taxonomic identifications. Estuarine and near-coastal programs sample benthic invertebrates from a surface area defined by gear type. Example methods documents are [13, 14], and several field and laboratory operations manuals from EPA National Aquatic Resource Surveys (NARS) [15–24].

Efforts to customize performance measures to biological monitoring programs have sought to use the process to isolate potential sources of variability, or error, and determine the need for and nature of corrective actions [14, 25]. A biological assessment protocol is a series of methods encompassing field sampling, laboratory processing (if necessary, and including sorting/subsampling and taxonomic identification), enumeration, data analysis, and assessment endpoints such as a regionally calibrated multimetric IBI. Community-level fish indicators typically do not involve laboratory

work as identification and counting is done in the field while on site. [14, 25] propose performance measures and MQO to cover the sequential phases of biological assessments. Key components considered are field sampling precision, and for BMI, sorting/subsampling and taxonomic identification. In the framework they propose, the ability to detect or highlight errors in these phases requires specific activities that provide data to calculate performance measures, the results of which are then compared to the MQO (**Table 1**). Descriptions below are examples of performance measures and how data are acquired for their calculation.

Field sampling precision (requires duplicate samples). Biological samples are taken from duplicate 100 m channel reaches that are immediately adjacent to each other. Laboratory processing and indicator calculation proceeds for each as separate samples. Comparison of results using specific performance measures (relative percent difference [RPD], coefficient of variability [CV], and confidence intervals [CIs]) (**Table 1**) reveals the precision and repeatability of the sampling method and its application.

Sorting/subsampling bias (requires sort residue rechecks). The objective of primary sorting of BMI samples is to remove all organisms from nontarget sample material, such as leaf litter, twigs, sand/silt, and other organic and inorganic detritus. The remaining sample material (sort residue) is checked for specimens missed by the primary sorter, and the performance measure, percent sorting efficiency (PSE) (**Table 1**) calculated as indicative of bias in the process.

Taxonomic precision (requires sample re-identification). Biological samples undergo identification by a primary taxonomist, then reidentification by a separate, independent taxonomist. Identification and count results are directly compared, and differences or error rates are quantified as a measure of taxonomic performance, specifically, precision. Terms calculated are percent taxonomic disagreement (PTD), percent difference in enumeration (PDE), and percent taxonomic completeness (PTC) (**Table 1**). All three terms quantify distinctly different aspects of the taxonomic identification process and relate directly to overall sample characteristics. Further, PTC indicates the proportion of the sample identified to the target hierarchical level (species, genus, tribe, family, or higher), where the absolute value of the difference between primary and QC taxonomist ($|PTC|$) indicates precision and consistency. Results from QC analyses can be presented in reports or associated with datasets in a straightforward manner (**Table 2**) that allows the data user to understand and move ahead with subsequent analyses.

The sites and samples for which these analyses are done use a *randomly selected* subset of sites, sort residue samples, and samples, respectively. As a rule of thumb, approximately 10% would be selected from the sample lot. The outcomes of these calculations and comparison to MQO can and should be used to (1) help detect potential problems in how the specific activity was implemented, (2) help inform the nature and need for corrective actions; and (3) summarize the overall quality of the full dataset. Subsequent values exceeding the MQO are not automatically taken to be unacceptable data points; rather, such values should receive closer scrutiny to determine reasons for the exceedance and might indicate a need for corrective actions.

The rationale for determining numeric values to be used as MQO should be based on observable data which are relevant to the monitoring program and the indicators that are being tracked as a part of it [25]. As an example, the MQO for PTD is 15 [26] and was arrived at through recognizing that taxonomic comparison (TAXCOMP) results for many samples were <20 and that there were very few <10. The 15% simply splits the difference.

Indicator category	Indicator/group	Data	Data origin	Performance term	Source
Biological	BMI	Assemblage-level, taxonomy, count	Laboratory processing	Sample sorting: PSE Taxonomic identification: PTD, PDE, PTC, PTC	[14, 18, 19, 25–27]
Biological	BMI	Individual metrics, MMI	Field sampling; MMI	Precision (among sites): CV, CI90 Precision (within sites): RPD	[14, 17, 21, 22, 25–27]
Biological	Fish	Individual metrics, MMI	Field sampling; field processing; MMI	Precision (among sites): CV Precision (within sites): RPD Taxonomic identification: PTD Percent completeness: % comp.	[28, 29]
Physical	Physical habitat		Field observations	Precision (among sites): CV, CI90 Precision (within sites): RPD	[13, 14, 25]
Physical	Sediment	Sediment grain size and total organic carbon	Laboratory processing	Precision and accuracy: recovery of spikes in blanks and matrices; MDLs (calculated for lab)	[23]
Physical	Water clarity	Photosynthetically active radiation transmittance at 1 m	field measurements, calculation	Slope of least squares regression [$-\ln(\text{light UW}/\text{light AMB})$ vs depth]; $R^2 > 0.75$	[24, 30]
Physical	Water clarity	Mean Secchi depth	Field measurements, calculation	Precision: all disappear and reappear values (3 of each) within 0.5 m	[23, 30]
Chemical	Fish	Tissue contaminant load	Laboratory processing	Sample preparation: RPD for duplicate homogenized tissue sample pairs; sample analysis: RSD for initial precision recovery, matrix spike, and matrix spike duplicate samples	[31–34, 51]
Chemical	Residuals and water quality	PFAS—16 analytes	Laboratory analysis of samples collected by facilities	Accuracy of measurements: % recovery for internal standards, LCS % recovery, MS % recovery; precision: RPD for MS/MSD and FDs	[35]
Harmful algal blooms	Algal toxins	Cylindrospermopsin, microcystins	Laboratory analysis of proficiency test (PT) samples	Accuracy of measurement: % recovery; precision among analytical laboratories: RPD	[18–20, 23, 36–40]

Indicator category	Indicator/group	Data	Data origin	Performance term	Source
Harmful algal blooms	Algal toxins	Cylindrospermopsin, microcystins	Generally field sampling	False positive rate; false negative rate; sensitivity (detection limit); CV for precision	[41, 42]
Ecotoxicity testing	Sediment toxicity	Acute toxicity of whole sediment sample	Laboratory processing	Minimum mean control corrected % survival	[43, 44]
Ecotoxicity testing	Aquatic toxicity	Counts, weight, % survival, % fertilization	Lab testing; field exposures	Within-test variability; sensitivity to specific contaminants; control precision (CV); PMSD	[45, 46]

BMI, benthic macroinvertebrates; PSE, percent sorting efficiency; PTD, percent taxonomic disagreement; PDE, percent difference in enumeration; PTC, percent taxonomic completeness; |PTC|, absolute value of PTC difference; MMI, multimetric index; CV, coefficient of variability; RPD, relative percent difference; RSD, relative standard deviation; CI90, 90% confidence interval; UW, under water; AMB, ambient; PFAS, per- and polyfluoroalkyl substances; LCS, laboratory control sample; MS, matrix spike; MS/MSD, matrix spike/matrix spike duplicate; FD, field duplicate; MDL, method detection limit; PMSD, percent minimum significant difference.

Table 1.
Selected example performance measures for QC planning and analysis.

Performance characteristic	MQO	Observed
1. Field sampling precision (MMI)	CV < 15%	10.6
	CI90 ≤ 1.0	0.8
2. Sorting/subsampling bias	PSE ≥ 90	96.7
3. Taxonomic precision	Median PTD ≤ 15%	5.4
	Median PDE ≤ 5%	0.5
4. Taxonomic completeness	Median PTC ≥ 90%	91
	Median PTC ≤ 5%	1.5

Table 2.
Summary results from QC analyses BMI samples (n = 9) from the Prince George’s County (Maryland, USA) biological monitoring program, 2010–2013.

Subsequent TAXCOMP results support using 15%, whether at broad national scales or smaller programs of anywhere from 10 to 50 samples. MQO are also not necessarily intended to be permanently fixed. As a monitoring program or testing procedure matures and more experience is gained, subsequent values often are observed as being consistently lower; a program may determine it would be beneficial to lower the MQO. Among all programs, PTD values are increasingly more commonly observed <10. It is advisable to use improved understanding of variability and its causes to adjust thresholds.

3.2 Physical habitat

3.2.1 Wadeable streams

One approach for characterizing the quality of stream physical habitat is a visual-based procedure [13] that assesses channel conditions in terms of stability, complexity,

and availability of habitat for stream biota. There are 10 parameters, seven of which are rated for all streams, and 3 each for low and high gradient streams (Table OS-1¹). Parameters are graded along a continuum of conditions from the perspective that as a stream becomes physically degraded, it loses physical complexity. Each parameter is rated on a 20-point scale while the observer is on site, then the values are summed for an overall site score. The range for the overall score is 0–200, with low values indicating poor quality habitat incapable of supporting stream biota and high indicating optimal conditions.

Data for input to QC calculations are from assessments done on adjacent 100 m channel reaches, identical to those discussed above for biological sampling. Reaches for which duplicate assessments are performed are randomly selected from the full site load, and pairs of habitat assessment results are used to calculate different performance measures (Table 1). As an example of results from such a QC analysis, consider a project that assessed 87 wadeable stream locations in Prince George's County, Maryland USA, and thus had nine (9) pairs of habitat assessment scores (Table OS-2 cdn.intechopen.com/public/259766_osi.zip).

Even though the field technique is qualitative, these numbers demonstrate the consistency of the results, particularly the median relative percent difference (mRPD) and CV. The values of RPD range from 1.4 to 35.3, with the substantial difference at the high end of the range suggesting that either the two reaches are dramatically different in quality, or potentially a data recording error occurred. These numbers characterize quality of the physical habitat data, as well as provide a roadmap for investigating potential anomalous results.

3.2.2 Estuarine/near coastal

Environmental monitoring programs assess abiotic indicators to understand how stressors may impact organisms, as well as how the habitat may be impacted by human disturbance. For example, because light underwater diminishes with depth [47] programs such as the U.S. EPA NARS National Coastal Condition Assessment (NCCA) survey and the Chesapeake Bay Program collect *in situ* water clarity measurements to estimate the impact of cultural eutrophication on light attenuation through the water column [24]. The EPA measures water clarity as Secchi depth at Great Lakes nearshore sites (the average depth of disappearance and reappearance of a 20 cm black and white disk lowered and retrieved through the water column three times), or transmission of photosynthetically active radiation (PAR) by comparing simultaneous ambient and underwater light measurements at incremental depths for estuarine sites. Performance measures for water clarity are intended to ensure accuracy and precision, as well as repeatability and consistency across the wide array of sites encountered in the survey. Secchi depth performance checks are implemented in the field and reviewed by analysts before use. They require that all six measurements are within 0.5 m of each other. When the difference between the maximum and minimum Secchi measurements at a site exceeds 0.5 m, the field crew repeats the entire set of measurements [24]. Data analysts again check Secchi data; values exceeding the maximum difference of 0.5 m among measurements at a site are reviewed and obvious transcription errors are corrected. Final values that do not meet the quality requirement are excluded from analysis. Table OS-3 cdn.intechopen.com/public/259766_osi.zip.

¹ Due to space limitations, Tables OS-1 through OS-11 are provided as Online Supporting Information cdn.intechopen.com/public/259766_osi.zip.

com/public/259766_osi.zip illustrates the decisions made when reviewing Secchi data collected at 20 sites during the NCCA 2010 field season. For PAR, light sensors and data loggers are required to have been calibrated within 2 years prior to use and NCCA analysts conduct post measurement data checks to verify data quality. To ensure that the underwater light measurements decrease with depth (that is, light attenuation increases with depth), the PAR attenuation coefficient (K_d) is first calculated as the negative of the natural log of the ratio of underwater light to ambient light [$-\ln(UW/AMB)$]. K_d is then plotted on the Y axis against the measurement depth on the X axis. If there is a negative slope of the resulting least squares regression line, or the coefficient of determination (R^2) $\leq 0.75^2$, measurements are investigated further. When specific measurements are found to be incorrect, they are excluded from regression [30]. Figure OS-1 cdn.intechopen.com/public/259766_osi.zip illustrates an example of erroneous UW PAR measurements that were excluded from analysis at a site sampled during the 2010 NCCA field season.

3.3 Chemical

3.3.1 Algal toxins

Recent NARS, including the National Lakes Assessment (NLA 2017), National Rivers and Streams Assessment (NRSA 2018/2019), and the NCCA (2020), sampled assessment locations (sites) from across the US. Locations were selected using a probability-based approach to provide representative results to estimate conditions at broad spatial scales. For purposes of discussion in this section, we will focus on water grab samples that were collected from a subset of sites representing lakes, streams and rivers, and coastal areas for analysis of cyanobacteria-produced algal toxins (microcystins and cylindrospermopsin).

As part of the effort to meet programmatic data quality requirements [18, 20, 23], EPA designed a performance analysis to document the reliability and consistency with which analytical laboratories detected the presence and concentration of the algal toxins cylindrospermopsin and microcystins. With a focus on accuracy (percent recovery), the design provided performance test (PT) samples to state and national laboratories analyzing field samples for which the nominal concentrations were known to the NARS QC administrators. The objective of the PT analysis is to allow use of the results to evaluate the quality of the analytical procedures, specifically through use of enzyme-linked immunosorbent assay (ELISA) test kits, and potentially develop recommendations for improvement in sample handling, preparation, and analytical techniques.

Sets or “waves” of PT samples were prepared and delivered to the target laboratories during the same period that primary project samples were undergoing analysis. Two waves were analyzed for the NLA (2017), and three waves of PT samples each were analyzed for the NRSA (2018/2019) and the NCCA (2020). The procedures for analyzing microcystins and cylindrospermopsin included necessary cleanup steps for samples with salinity >3.5 parts per thousand, as well as dilution steps for samples with concentrations $>$ upper detection limit (UDL) of the ELISA test kits. The PT

² The protocol in [30] calls for a minimum R of 0.95; the minimum R for the NCCA is relaxed to 0.75 to allow for variability in measurement due to factors such as differing sun angles throughout the day or underwater light reflection at shallower estuarine sites.

samples were subjected to multi-temperature stability studies before shipment, and then shipped on ice packs overnight to the laboratories analyzing NARS field samples.

PT samples were prepared to specified concentrations of cyanotoxins (Table OS-4 cdn.intechopen.com/public/259766_osi.zip) and distributed to the target laboratories. We used two performance measures in evaluating the PT results. First, percent recovery was used for accuracy, and RPD or relative standard deviation (RSD) [40, 48–50] for precision. Although all PT concentrations are shown (**Table 1**), for reasons of space limitations we have selected example results to illustrate results for one round of analyses for which the most accurate % recovery results were obtained and another for the least accurate from the most recent NARS, including NLA2017, NRSA2018/2019, and NCCA2020.

Both Lab A and Lab B met the % recovery goal of 70–130% [38] for the freshwater microcystins 2018/2019 NRSA Round 1 PT samples (Table OS-5 cdn.intechopen.com/public/259766_osi.zip). In comparison, Lab A did not meet the % recovery goal for the two of the freshwater microcystins 2017 NLA Round 1 PT samples. It should be noted that the results for sample M-7 were only slightly outside the % recovery range. In addition, although the results for M-10 were lower than 70% recovery, the PT sample concentration was much higher than the test kit range and required several dilutions for analysis.

Lab A met or nearly met the % recovery goal of 70–130% [38] for the estuarine microcystins 2020 NCCA Round 3 PT samples (Table OS-6 cdn.intechopen.com/public/259766_osi.zip). In contrast, Lab D did not meet the % recovery goal for 2 of the estuarine microcystins 2020 NCCA Wave 1 PT samples. The 2020 NCCA Wave 1 estuarine microcystins % recovery results ranged from 63.0 to 131.1, excluding the two non-detect results from Lab D. The 63.0% recovery value was a calculated PT sample concentration above the upper limit of detection (20MC-9) and the 131.1 % recovery value was calculated for the lowest microcystins concentration (20MC-8). The non-detect results reported by Lab D were for concentrations at the lower end of detection (20MC-8 and 20MC-10).

Lab A met the % recovery goal of 70–130% [39] for the freshwater cylindrospermopsin 2020 NCCA Wave 3 PT sample (Table OS-7 cdn.intechopen.com/public/259766_osi.zip). In comparison, Lab A did not meet the % recovery goal for four of the freshwater cylindrospermopsin 2017 NLA Wave 1 PT samples. It should be noted that of the 2017 NLA Wave 1 PT sample concentrations with % recovery value outside the % recovery goal, only C-4 had a concentration within the detection range of the test kit.

Lab A met the % recovery goal of 70–130% [39] for the estuarine cylindrospermopsin 2020 NCCA Wave 3 PT sample (Table OS-8 cdn.intechopen.com/public/259766_osi.zip). In contrast, Lab A did not meet the % recovery goal for all five of the estuarine cylindrospermopsin 2020 NCCA Wave 1 PT samples and Lab D did not meet the % recovery goal for one of the estuarine cylindrospermopsin 2020 NCCA Round 1 PT samples. The vendor laboratory noted that the salts used to prepare the estuarine PT samples might have caused the elevated % recovery values for the lower concentrations (<1 µg/L) due to background interference. The vendor laboratory indicated that the salts would not lead to false positive results if there were no cylindrospermopsin in the sample.

The analyses and comparisons of analytical results highlighted potential issues that allowed the QC coordinators to inquire for additional information. Although these particular instances did not result in anomalous results, the evaluations did help improve understanding of the sample handling and analysis process.

3.3.2 Per- and polyfluoroalkyl substances in residuals

Entities permitted to sell or distribute wastewater residuals for land application in Massachusetts were required by the Massachusetts Department of Environmental Protection (MDEP)³ to collect and submit quarterly samples in 2020–2021 for analyses of 16 PFAS (Table OS-9 cdn.intechopen.com/public/259766_osi.zip). In 2020–2021, no EPA-approved methods were available for testing residuals for PFAS. Laboratories used “modified” EPA Method 533 (*Determination of Per- and Polyfluoroalkyl Substances in Drinking Water by Isotope Dilution Anion Exchange Solid Phase Extraction and Liquid Chromatography/Tandem Mass Spectrometry*) [35] to analyze samples. The laboratory SOP were reviewed and approved by the MassDEP before they were used to analyze the residuals samples. In addition, a standardized data quality evaluation checklist was developed and used to consistently perform reviews of the quality of results reported in laboratory data packages. Implementing these steps allowed for evaluation of whether the analytical results met the quality requirements outlined in EPA Method 533 “modified” [35], as well as the overall analytical quality requirements in 40 CFR Part 136.7 (*Guidelines Establishing Test Procedures for the Analysis of Pollutants, Quality Assurance and Quality Control*).

In 2021, an evaluation of the analytical results was performed for quarterly residuals samples collected during the last quarter of 2020 through the third quarter of 2021 using the standardized data evaluation checklists. The method quality objectives (e.g., holding times, minimum reporting limits, RPD for laboratory or field duplicates) presented (**Table 1**) were evaluated and documented for each sample using a standardized data quality evaluation checklist. Additional issues that the laboratories encountered during analysis were also documented in these checklists. Results from these standard evaluations were used to qualify the data to enable end users to interpret the quality of results. We provide a summary of the qualifiers used (and frequency of use) for each of the reported 16 analytes from a total of 164 samples (Table OS-10 cdn.intechopen.com/public/259766_osi.zip).

Elevated reporting limits (>1 ng/g) were the most frequently used qualifier (Table OS-10 cdn.intechopen.com/public/259766_osi.zip). The R qualifier was used for at least one analyte for 79% of the samples analyzed. These elevated reporting limits were less frequently observed in samples with low moisture content, with all samples with less than 28.3% solids having elevated reporting limits for at least one analyte. It should be noted that the remaining qualifiers used for the results were only applied when the results were greater than the detection limit. The J1- qualifier, indicating that the isotopically labeled analogue recovery was below the lower acceptance limit and that the residual result is estimated (could be biased low) for the corresponding target PFAS, was used for at least one analyte for 37% of samples analyzed. The J6+ qualifier, indicating that the ratio of the quantifier ion response to qualifier ion response (i.e., primary mass transition) falls outside of the laboratory established criteria (i.e., outside ratio limits) and that results are estimated maximum concentrations, was used for at least one analyte for 37% of the samples analyzed. The J5+/- qualifier was used for at least one analyte for 34% of the samples, commonly indicating that the RPD for the field sample duplicate (or less commonly the MSD)

³ 310 CMR 32.00: Land Application of Sludge and Septage, which states “any additional substance for which sampling and analysis is required by the Department, before or after the sludge or septage is approved by the Department pursuant to 310 CMR 32.11.” Also, see URL: <https://www.mass.gov/doc/required-laboratory-procedures-for-testing-pfas-in-residuals/download>.

was above the upper acceptance limit or not analyzed with the residual extraction batch; this indicated that the residual PFAS results above the RL were estimated (could be biased high or low).

Results of the 2020–2021 QC evaluations were used to inform ongoing residual analyses in Massachusetts. MassDEP communicated results for individual data packages and for the overall analysis to the laboratories, contributing facilities, and their management to refine protocols and execution of the residual PFAS monitoring program. Additional analyses of the magnitudes of PFAS concentrations over time and of duplicate precision were used to recommend field sampling and duplication frequency and is a technical issue many states and other entities are beginning to address.

3.3.3 Tissue contaminants

As with biological monitoring and bioassessments, performance measures and MQOs are essential for both the field and laboratory aspects of tissue contaminant monitoring studies of aquatic biota (e.g., fish, mollusk, or crustacean tissue studies for human health or ecological risk management and communication). QA planning and implementation should focus on defining DQOs, designing a QC system to measure data quality, and assessing data quality to determine its suitability to support management decisions regarding future monitoring, risk assessment, or issuance of consumption advisories [31, 51].

Field QC procedures need to be detailed in SOPs and as noted previously, sampling practitioners need to be trained in those program-specific procedures. A primary QA concern for the field collection, handling, preservation, and shipping stages of tissue contaminant studies is the preservation of tissue sample integrity. The accuracy of analytical results depends in part on the immediate preservation (i.e., freezing) of tissues and the prevention of exposure to extraneous sources of contamination. Those sources need to be identified and avoided or eliminated. Field blanks, or rinsates of empty field sample containers have been used by some investigators to evaluate field sample packaging materials as sources of contamination, with a control limit of less than the MDL as determined for the particular analytical method or monitoring program [51]; however, immediate freezing of whole organisms in the field (and preparation of tissue in the laboratory) and the use of food-grade packaging materials reduces or even eliminates the need for field blanks. Some studies may require tissue resection in the field, but sample processing (including resections) conducted under controlled laboratory conditions reduces the potential for sample contamination. One means of evaluating the efficacy of tissue preparation cleaning and decontamination procedures is the preparation and analysis of processing blanks or rinsates of the equipment used for dissecting and homogenizing tissues. As with field contamination QC measures, the control limit for processing blanks would also be <MDL for the particular analytical method or monitoring program. Control limit exceedances require suspension of sample preparation and specific corrective action by the preparation laboratory before resection or homogenization may resume.

Overall completeness is the number of valid sample measurements relative to the number of samples planned for collection, and it may be impacted by a variety of circumstances, e.g., storm events, samples lost during shipment, etc. Completeness objectives vary by study administrators and can range from 80% to 99%, with levels <80% generally requiring corrective action such as resampling or reanalysis [33, 34, 51]. Sampling precision (or the degree of agreement among replicate measurements caused by random error) can be estimated by comparing field replicates using RSD; however,

acceptable field replicate samples require the collection of target organisms of the same species and the same sizes collected from the same location which may not always be possible. Rather than establishing acceptance limits for sampling precision, some researchers have instead used field replicate results to aid in the evaluation of study results and characterize the variability of the sampled population [32, 34]. Variability arising from tissue preparation (e.g., homogenization, compositing, and aliquoting), shipping, and laboratory analysis processes can be estimated by having the sample preparation laboratory prepare duplicate tissue homogenate or processed composite sample pairs to be analyzed as blind duplicates. [32] applied a MQO specifying that the RPD for these duplicate tissue composite pairs should be <50% for values greater than 5× the minimum level of quantification (ML) for each target contaminant and <100% for values <5× the ML.

In addition to the use of duplicate homogenate or composite sample pairs, a standard suite of laboratory QC measures including initial precision and recovery (IPR) samples, and matrix spike and matrix spike duplicate samples provides information about the precision associated with various components of the analytical process. IPR samples are used to demonstrate that a laboratory can achieve precision and accuracy using a particular analytical method prior to the analysis of any tissue study samples. They consist of a reference matrix (i.e., one that matches the study tissue matrix) that is spiked to a known level with the target contaminant. Accuracy is measured by the average recovery of the target chemical in replicate IPR samples. Precision is assessed by calculating RSD of the measured concentrations of the target chemical in the IPR samples. Matrix spike samples are field sample tissue homogenates with known amounts of a target chemical spiked into the sample to assess the effect of matrix interferences on compound identification and quantitation (measured as percent recovery of the chemical). Duplicate matrix spike samples consist of additional aliquots of matrix spike samples that are analyzed to assess the effect of tissue matrix interferences and are routinely used to assess method precision. Summarizing measurement QC limits for tissue studies is not as straightforward as identifying measurement quality indicators. Analytical QC limits vary with target chemicals and analytical methods. [51] provides general control limit recommendations and associated corrective actions for fish and shellfish tissue studies.

3.4 Ecotoxicology

Ecotoxicology tests are used in many countries and environmental programs as one of several approaches to assess environmental condition of soils, sediments, and water, toxicity of chemicals (including pesticides), and compliance with environmental regulatory statutes (e.g., the Clean Water Act in the U.S.). Toxicity testing for these various programs is largely conducted in a controlled laboratory setting according to specific test method protocols, e.g., [46, 52, 53], although mesocosm and in situ toxicity testing is also used in some cases in aquatic testing of chemicals, for example, (e.g., [54, 55]). Toxicity test results consist of two types of information: biological measurements and statistical interpretation of the observed biological data. Biological measurements are the raw data recorded when conducting toxicity tests (e.g., survival, weight, number of young produced). The statistical interpretation of a toxicity test is derived from the observed biological data.

Like other types of methods that rely on biological data, results of a toxicity test depend on the method used. Ecotoxicological testing relies on several QC procedures

and analyses to help document that the test method performs acceptably given program DQO [46, 52]. Two key QA procedures used in all ecotoxicology testing are: (1) results from testing with a reference toxicant and (2) meeting minimum test acceptability criteria.

In reference toxicant testing, test organisms are exposed to a range of concentrations of a known toxicant or positive control (e.g., a metal such as copper or a salt such as potassium chloride for aquatic testing, e.g., [56–58]). Organism response to that toxicant is compared against an acceptable range of response previously established by the laboratory for the test organism and test method. Control charts are developed based on several reference toxicant tests for a given test species and test method to document an acceptable range of response to the toxicant [46]. In practice, statistical point estimate endpoints rather than the raw data are used to document results of each test and establish an acceptable range of response for a test method and reference toxicant. Often, a series of performance measures is used with corresponding MQO to address a range of relevant concerns (Table OS-11 cdn.intechopen.com/public/259766_osi.zip). Examples of point estimate endpoints include the lethal concentration to 50% of the test organisms (LC50) and the concentration resulting in a 25% inhibition in response compared to the control organisms (IC25). Point estimate endpoints have the advantage of generating 95% CIs around the mean value so that within test variability as well as between test variability can be established. These endpoints can be compared across tests and laboratories for a given chemical because the endpoint is not dependent on the concentration series used.

The second key QA requirement is that each test method has minimum test acceptance criteria (TAC) for control organisms that should (must for some programs such as the NPDES program in the U.S.) be met in a test for the results to be considered of acceptable quality. Examples of TACs include metrics such as minimum acceptable percent survival for organisms in a clean control matrix, minimum growth, and minimum number of offspring per female that must be achieved in the controls in a test [46].

A key performance measure in ecotoxicological testing is within-test variability or precision, both in the controls alone and for the entire test. Laboratories track performance metrics for the control over time to assess within-test variability. This is accomplished by calculating the mean, standard deviation (SD), and coefficient of variation (CV) of the control replicate data for each test conducted by the laboratory for a given test method. A statistical metric that is used to calculate within-test variability for the test as a whole is percent minimum significant difference (PMSD) [59, 60], which is derived from an Analysis of Variance (ANOVA) and Dunnetts Multiple Comparison Analysis. The PMSD documents the percent effect that can be statistically distinguished as compared to the control in the test based on the within-test variability observed.

Allowable ranges of PMSD values were derived by EPA using multiple tests for a given test method [59]. Controlling both minimum as well as maximum intra-test variability in whole effluent toxicity (WET) tests is seen as an important test acceptance factor. Too much variability among control replicates reduces the ability to distinguish statistical difference in organism response among treatments. Too little variability among control replicates, on the other hand, can yield statistically significant differences among-test concentrations and the control that are *biologically meaningless*. Controlling within-test precision is key to achieving the optimal sensitivity possible using a particular test species and ecotoxicology test method.

4. Conclusions

It should be noted that there has been vigorous debate on the appropriateness of actions that should result from interpreting statistical deviation in terms of process or QC [6], including that practitioners should avoid over-interpretation. This includes suggestions that unnecessary adjustments in processes could actually increase frequency of anomalous results. The implication here is that someone interpreting and developing recommendations from QC analysis who is not knowledgeable about the field of practice or study risks having a program just working toward a number, rather than truly trying to improve a process or determine the quality of environmental data for use in assessing ecological outcomes.

Recognition of the causes, magnitude, and effects of variability and error is attained through consistent observation and measurement and can simultaneously provide direction on the need for and types of corrective actions. Appropriately developed and implemented MQO, as part of consistent and routine measurement and monitoring programs, not only function to keep them on-track, but in the long run can also lead to more cost- and time-efficient processes.

Acknowledgements

The authors here express gratitude to Dr. Richard Mitchell (US EPA/Office of Water) for participation in and contribution to discussions related to the structure and content of this chapter. We would also like to recognize Lealdon Langley, Susannah King, and Jennifer Wood of the Massachusetts Department of Environmental Protection for their contributions during reviews of the Massachusetts PFAS data packages.

Author details

James Stribling^{1*}, Susan Lanberg², Blaine Snyder¹, Jerry Diamond¹, Benjamin Jessup³
and Hugh Sullivan⁴

1 Tetra Tech, Inc., Center for Ecological Sciences, Owings Mills, Maryland, USA


2 Tetra Tech, Inc., Environmental Sciences, Fairfax, Virginia, USA

3 Tetra Tech, Inc., Center for Ecological Sciences, Montpelier, Vermont, USA

4 US Environmental Protection Agency, Office of Water, Washington, DC, USA

*Address all correspondence to: james.stribling@tetratech.com

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] IOM (Institute of Medicine). *Environmental Decisions in the Face of Uncertainty*. Institute of Medicine. Washington, DC: The National Academies Press; 2013. p. 281. DOI: 10.17226/12568
- [2] Bradshaw G and Borchers J. Uncertainty as information: Narrowing the science-policy gap. *Conservation Ecology*. 2000;4(1):7. Available from: <https://www.ecologyandsociety.org/vol4/iss1/art7/>
- [3] Lindley D. *Uncertainty Einstein, Heisenberg, Bohr, and the Struggle for the Soul of Science*. New York, NY: Anchor Books, a Division of Random House; 2007. p. 257. ISBN: 978-1-4000-7996-4
- [4] Shewhart W. *Economic Control of Quality of Manufactured Product*. 7th Printing. D. Van Nostrand and Company; 1923 (Note, contemporary references to this book use 1931 as the publication date; 1923 is the date at the bottom of the Preface)
- [5] Shewhart W. *Statistical Methods from the Viewpoint of Quality Control*. Washington, D.C.: The Graduate School, U.S. Department of Agriculture; 1939, 1939. Republished 1986, with a new Foreword by W.E. Deming. Mineola, NY: Dover Publications, Inc. p. 105
- [6] Woodall W. Controversies and contradictions in statistical process control. *Journal of Quality Technology*. 2000;32(4):341-350. DOI: 10.1080/00224065.2000.11980013
- [7] Divorski S, Scheirer M. Improving data quality for performance measures: Results from a GAO study of verification and validation. *Evaluation and Program Planning*. 2001;24:83-94. DOI: 10.1016/S0149-7189(00)00049-5
- [8] Copeland C, Simpson M. *The Information Quality Act: OMB's Guidance and Initial Implementation*. Congressional Research Bulletin Report for Congress. Order Code RL32532. Washington, D.C.: The Library of Congress, Congressional Research Service; 2004
- [9] ANSI/ASQC. *American National Standard Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs*. E4-1994. American National Standards Institute/American Society for Quality Control, Milwaukee, WI: ASQ Quality Press; 1995
- [10] U.S. EPA. *Overview of the EPA Quality System for Environmental Data and Technology*. EPA/240/R-02/003. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2002a
- [11] U.S. EPA. *Guidance on Systematic Planning Using the Data Quality Objectives Process*. EPA QA/G-4. EPA/240/B-06/001. Washington, D.C.: U.S. Environmental Protection Agency, Office of Environmental Information; 2006
- [12] U.S. EPA. *EPA Requirements for Quality Assurance Project Plans*. EPA QA/R-5. EPA/240/B-01/003. 2001; Reissued May 2006. U.S. Environmental Protection Agency, Washington, D.C.: U.S. Environmental Protection Agency, Office of Environmental Information
- [13] Barbour M, Gerritsen J, Snyder B, Stribling J. *Revision to the Rapid Bioassessment Protocols for Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish*. 2nd ed. Washington, D.C.: U.S. Environmental

- Protection Agency, Office of Water; 1999
EPA/841-D-97-002
- [14] Flotemersch J, Stribling J, Paul M. Concepts and Approaches for the Bioassessment of Non-Wadeable Streams and Rivers. EPA/600/R-06/127. Cincinnati, OH: U.S. Environmental Protection Agency, Office of Research and Development, National Exposure Research Laboratory; 2006
- [15] U.S. EPA. National Lakes Assessment. Field Operations Manual. EPA 841-B-11-003. 2011. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2012
- [16] U.S. EPA. National Lakes Assessment. Laboratory Operations Manual. Version 1.1. EPA-841-B-11-004. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2012
- [17] U.S. EPA. National Lakes Assessment 2017 Field Operations Manual Version 1.1. EPA 841-B-16-002. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2017a
- [18] U.S. EPA. National Lakes Assessment 2017 Laboratory Operations Manual Version 1.1. EPA 841-B-16-004. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2017b
- [19] U.S. EPA. National Rivers and Streams Assessment 2018-19: Laboratory Operations Manual Version 1.1. EPA-841-B-17-004. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2018
- [20] U.S. EPA. National Rivers and Streams Assessment 2018/19 Quality Assurance Project Plan, Version 1.2. EPA 841-B-17-001. Washington, D.C.: United States Environmental Protection Agency, Office of Water; 2019b
- [21] U.S. EPA. National Rivers and Streams Assessment 2018/19: Field Operations Manual – NonWadeable. EPA-841-B-17-003b. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2019c
- [22] U.S. EPA. National Rivers and Streams Assessment 2018/19 Field Operations Manual Wadeable Version 1.2. EPA-841-B-17-003a. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2019d
- [23] U.S. EPA. National Coastal Condition Assessment 2020 Quality Assurance Project Plan, Version 1.2. EPA 841-F-19-003. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2021a
- [24] U.S. EPA. National Coastal Condition Assessment 2020 Field Operations Manual. EPA 841-F-19-005. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2021b
- [25] Stribling J. Partitioning error sources for quality control and comparability analysis in biological monitoring and assessment. Chapter 4. In: Eldin A, editor. *Modern Approaches to Quality Control*. London, UK: INTECH Open Access Publisher; 2011. pp. 59-84. ISBN 978-953-307-971-4. DOI: 10.5772/22388
- [26] Stribling J, Jessup B, Feldman D. Precision of benthic macroinvertebrate indicators of stream condition in Montana. *Journal of the North American Benthological Society*. 2008a;27(1):58-67. DOI: 10.1899/07-037R.1
- [27] Stribling J, Pavlik K, Holdsworth S, Leppo E. Data quality, performance, and uncertainty in taxonomic identification for biological assessments. *Journal of the North American Benthological Society*. 2008b;27(4):906-919. DOI: 10.1899/07-175.1

- [28] Klemm D, Stober Q, Lazorchak J. Fish Field and Laboratory Methods for Evaluating the Biological Integrity of Surface Waters. EPA/600/R-92/111. Cincinnati, OH: U.S. Environmental Protection Agency, Office of Research and Development; 1993
- [29] IDEM. Quality Assurance Project Plan for Biological Community and Habitat Measurements. B-003-OWQ-WAP-XXX-20-Q-R0. Indianapolis, IN: Indiana Department of Environmental Management, Office of Water Quality, Watershed Assessment and Planning Branch; 2020
- [30] Kopp B, Neckles H. A Protocol for Monitoring Estuarine Nutrient Enrichment in Coastal Parks of the National Park Service Northeast Region. Natural Resource Report NPS/NCBN/NRR—2009/110, Augusta. ME: U.S. Geological Survey Patuxent Wildlife Research Center; 2009
- [31] U.S. EPA. Quality Assurance Project Plan for Analytical Control and Assessment Activities in the National Study of Chemical Residues in Lake Fish Tissue. EPA 823-R-02-006. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2000
- [32] U.S. EPA. The National Study of Chemical Residues in Lake Fish Tissue. EPA-823-R-09-006. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2009
- [33] Neebling T. Standard Operating Procedure for Fish Tissue Contaminant Sampling – Field and Laboratory Methods, Casper. WY: Wyoming Game and Fish Department, Fish Division; 2012
- [34] Seiders K, Sandvik P. Freshwater Fish Contaminant Monitoring Program. Publication 20-03-106, Olympia. WA: Washington State Department of Ecology; 2020
- [35] U.S. EPA. Method 533, Determination of Per- and Polyfluoroalkyl Substances in Drinking Water by Isotope Dilution Anion Exchange Solid Phase Extraction and Liquid Chromatography/Tandem Mass Spectrometry. EPA 815-B-19-020. Cincinnati, OH: U.S. Environmental Protection Agency, Office of Ground Water and Drinking Water; 2019a
- [36] U.S. EPA. National Lakes Assessment 2017 Quality Assurance Project Plan Version 1.1. EPA-841-B-16-003. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2017c
- [37] U.S. EPA. National Coastal Condition Assessment 2020 Laboratory Operations Manual, Version 1.2. EPA 841-F-19-004. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2021c
- [38] U.S. EPA. Method 546: Determination of Total Microcystins and Nodularins in Drinking Water and Ambient Water by Adda Enzyme-Linked Immunosorbent Assay. EPA 815-B-16-011. Cincinnati, OH: U.S. Environmental Protection Agency, Office of Ground Water and Drinking Water; 2016
- [39] U.S. EPA. Method 545: Determination of Cylindrospermopsin and Anatoxin-a in Drinking Water by Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometry (LC/ESI-MS/MS). EPA 815-R-15-009. Cincinnati, OH: U.S. Environmental Protection Agency, Office of Ground Water and Drinking Water; 2015
- [40] Berger W, McCarty H, and Smith R. Environmental Laboratory Data Evaluation. Schenectady, NY: Genium Publishing Company. 1996. ISBN 1-931690-91-9

- [41] U.S. EPA. Single Laboratory Validated Method for Determination of Cylindrospermopsin and Anatoxin-a in Ambient Freshwater by Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometry (LC/ESI-MS/MS). EPA/600/R-17/130. 2017d. Cincinnati, OH: U.S. Environmental Protection Agency, Office of Research and Development;
- [42] U.S. EPA. Single Laboratory Validated Method for Determination of Microcystins and Nodularin in Ambient Freshwater by Solid Phase Extraction and Liquid Chromatography/Tandem Mass Spectrometry (LC/MS/MS). EPA/600/R-17/344. Washington, D.C.: U.S. Environmental Protection Agency, Office of Research and Development; 2017e
- [43] ASTM. Standard Test Method for Measuring the Toxicity of Sediment-Associated Contaminants with Freshwater Invertebrates. ASTM E1706-19, West Conshohocken. PA: American Society for Testing and Materials; 2019
- [44] U.S. EPA. Methods for Measuring the Toxicity and Bioaccumulation of Sediment-associated Contaminants with Freshwater Invertebrates. EPA/600/R-99/064. Washington, D.C.: U.S. Environmental Protection Agency, Office of Research and Development and Office of Water; 2000b
- [45] U.S. EPA. Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms. EPA-821-R-02-012. 2002b. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water;
- [46] U.S. EPA. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms. EPA-821-R-02-013. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2002c
- [47] Kirk J. Light and Photosynthesis in Aquatic Ecosystems. Cambridge University Press; 1994. DOI: 10.1017/CBO9780511623370
- [48] APHA. Standard Methods for the Examination of Water and Wastewater. 21st ed. Washington, D.C.: American Public Health Association, American Water Works Association, and Water Environment Federation; 2005
- [49] Keith L. Environmental Sampling and Analysis. A Practical Guide. 1991. Chelsea, MI: Lewis Publishers. p. 160. ISBN: 9780873713818
- [50] Smith R. Interpretation of Organic Data. Amsterdam, New York: Genium Publishing Corporation, Genium Group, Inc., 2000. ISBN: 1-890911-19-4
- [51] EPA US. Guidance for Assessing Chemical Contaminant Data for Use in Fish Advisories, Volume 1: Fish Sampling and Analysis. 3rd ed. EPA-823-B-00-007. Washington, D.C.: U.S. Environmental Protection Agency, Office of Water; 2000
- [52] ASTM. Standard Test Method for Measuring the Toxicity of Sediment-Associated Contaminants with Freshwater Invertebrates. ASTM E1706-19, Philadelphia. PA: American Society for Testing and Materials; 2016
- [53] OECD (Organization for Economic Cooperation and Development). OECD Test Guidelines for Chemicals [Internet]. 2022. Available from: <https://www.oecd.org/chemicalsafety/testing/oecdguidelinesforthetestingofchemicals.htm> [Accessed: 08 February 2022]
- [54] U.S. EPA. Series 850 - Ecological Effects Test Guidelines. Washington,

D.C.: U.S. Environmental Protection Agency, Office of Pesticides and Toxic Substances; 2022

[55] Pereira A, Soares A, Alves F, Ribeiro R. Test chambers and test procedures for in situ toxicity testing with zooplankton. *Environmental Toxicology and Chemistry*. 1999;**18**: 1956-1964. DOI: 10.1002/etc.5620180914

[56] ECCC. Guidance Document on Control of Toxicity Test Precision Using Reference Toxicants. Environmental Protection Series. [Reference methods] EPS 1/RM/12. Ottawa, Ontario: Environment and Climate Change Canada, Environment Canada; August 1990

[57] Lazorchak J, Smith M. Rainbow trout (*Oncorhynchus mykiss*) and brook trout (*Salvelinus fontinalis*) 7-day survival and growth test method. *Archives of Environmental Contamination and Toxicology*. 2007;**2007**(53):397-405. DOI: 10.1007/s00244-006-0227-8

[58] Environmental Health Standing Committee (enHealth). The Role of Toxicity Testing in Identifying Toxic Substances in Water: A Framework for Identification of Suspected Toxic Compounds in Water. Canberra: Environmental Health Standing Committee, Australian Health Protection Principal Committee; 2012

[59] EPA US. Understanding and Accounting for Method Variability in Whole Effluent Toxicity Applications Under the National Pollutant Discharge Elimination System Program. EPA 833-R-00-003. Washington, D.C.: U.S. Environmental Protection Agency, Office of Wastewater Management; 2000d

[60] Denton D, Fox J, Fulk F. Enhancing toxicity test performance by using a statistical criterion. *Environmental Toxicology and Chemistry*. 2003;**22**: 2323-2328. DOI: 10.1897/02-384