pennsylvania
DEPARTMENT OF ENVIRONMENTAL PROTECTION

OFFICE OF WATER PROGRAMS
BUREAU OF CLEAN WATER

DISCRETE PHYSICOCHEMICAL ASSESSMENT METHOD

2021

Chapter 3 Chemical Assessment Methods

Prepared by:

Brian Chalfant
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2013

Edited by:

Heidi Biggs, Dustin Shull, and Rebecca Whiteash
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2021

INTRODUCTION

Water quality criteria, located in 25 Pa. Code Chapter 93 (relating to water quality standards), are assessed considering magnitude (e.g., concentration), duration (i.e., averaging period), and frequency (i.e., how often excursions from criteria are allowed). The purpose of this document is to create standardized procedures for DEP to follow to make water use assessment decisions, as required under the Clean Water Act (CWA) Sections 303(d) and 305(b), based on discrete physical and chemical (physicochemical) data in relation to water quality standards (WQS). This document relies on contextual and conceptual discussions found in *Water Use Assessment Decision-making Based on Physicochemical and Bacteriological Sampling* (Chalfant 2013). It contains relatively little discussion of the planning and execution phases of physicochemical water quality sampling, such as outlining study objectives, choosing sampling plan designs, and setting data quality objectives, which are described in more detail within the *Water Quality Monitoring Protocols for Streams and Rivers* (Shull and Lookenbill 2018). This document aims to describe how the inherent variation and sampling error of physicochemical data are addressed by DEP in the use assessment process for discrete physicochemical water quality sampling data and to expand upon the general use assessment determination methods in Chapter 5, where appropriate.

Sampling Error

The inferential process of using discrete, spatiotemporally-limited observations (i.e., samples) to estimate a larger set of unobserved, continuously dynamic conditions can introduce uncertainty – called sampling error – into the use assessment determination process. Uncertainty attributable to analytical measurement techniques, known as measurement error, is discussed in the *Quality Assurance Manual for the Pennsylvania Department of Environmental Protection Bureau of Laboratories* (DEP 2016).

Ideally, physicochemical data will inform use assessment determinations by sampling frequently enough to minimize sampling error and to accurately characterize the conditions for each parameter of concern. These conditions should be characterized over a long enough time frame to account for variations in concentrations attributable to changes in all relevant factors. For many water quality parameters, including toxic substances listed in § 93.8a, this can only be achieved through very intense discrete sampling efforts, as they require physical site visits to collect data with handheld field meters and chemistry samples for laboratory analysis. When intense discrete sampling efforts are not possible, the resulting physicochemical data provides limited windows into the dynamic continuum of water quality conditions. However, more frequent sampling is possible for specific water quality criteria listed in § 93.7, through the deployment of automated, continuous instream monitoring (CIM) devices (see the Continuous Physicochemical Assessment Method).

Continuous instream monitoring devices can measure conditions frequently and can be deployed in remote locations. Monitoring water quality conditions at frequencies as high as every 15 minutes minimizes the amount of time sample results must be extrapolated into unobserved time, thereby minimizing the potential for sampling error (Hoger 2018). Continuous instream monitoring devices can be set up to report observations via telemetry or through occasional retrievals and downloads. While CIM devices can provide extremely detailed, temporally-dense observational records, many such devices can only measure a few water quality parameters (e.g., dissolved oxygen, temperature, conductivity, pH) for which WQS exist.

Another approach to reduce sampling error in the absence of temporally-dense observations, is to collect enough information on other relevant variables (e.g., stream flow, precipitation, water temperature) to allow for confident extrapolation from observed conditions to unobserved conditions based on an empirical understanding of variability. A wide variety of interrelated factors can contribute to spatial and temporal variation in the concentrations of water quality parameters, such as precipitation, stream flow, geology, watershed drainage patterns, and anthropogenic influences. Different water quality parameters often vary in unique ways relating to these and other factors. For example, dissolved oxygen concentrations in streams often exhibit strong annual and diel patterns attributable to interrelated patterns of solar flux, stream temperature, and photorespiratory activity. Meanwhile, concentrations of total dissolved solids often vary much less with diel or annual patterns of solar flux, but instead vary primarily with stream flow and related patterns of surface runoff, geology, and groundwater flow patterns. Knowledge and understanding of such patterns can strengthen inferences about unobserved conditions (USEPA 2005).

Sampling and Criteria Frequency

Within the regulatory framework outlined above, DEP must determine if waterbodies meet WQS. In 25 Pa. Code Chapter 16 (relating to water quality toxics management strategy – statement of policy), § 16.21 states that aquatic life criteria for toxic substances are developed such that the frequency of occurrence is accounted for through the specification of factors appropriate to the criteria in Chapter 96 (relating to water quality standards implementation), but also, that the basis for the magnitude, duration, and frequency is described in criteria development rationale or other appropriate supporting documentation. Section 16.22 states that DEP looks to National guidelines (USEPA 1985) in establishing aquatic life criteria for toxic substances. In 25 Pa. Code Chapter 96 (relating to water quality standards implementation), § 96.3

(relating to water quality protection requirements) provides that – to protect existing and designated surface water uses – the water quality criteria described in Chapter 93 shall be achieved in all surface waters at least 99% of the time. For fish and aquatic life criteria for toxic substances, the National guidelines most often state that criteria excursions are to occur no more than once in three years on average (USEPA 1985). For human health criteria for toxic substances, because EPA's derivation of the WQC recommendations "involves the calculation of the maximum water concentration for a pollutant that ensures drinking water and/or fish ingestion exposures will not result in human intake of that pollutant in amounts that exceed a specified level based upon the toxicological endpoint of concern" (USEPA 2000), it is understood that any human health criteria excursions are unacceptable. A number of interrelated considerations – discussed in more detail in Chalfant (2013) – must be addressed when assessing if waterbodies meet WQS "at least 99% of the time" or to determine if excursions occur more than "once in three years on average" based on physicochemical samples.

The frequency of "at least 99% of the time" addresses the temporal aspect of criteria for which this consideration is not otherwise explicitly specified in Chapter 93 or addressed in the criteria development documents for Nationally recommended criteria. Note that some water quality criteria in § 93.7 have frequency components explicitly specified as part of the criteria, including the ammonia nitrogen criterion for aquatic life and the bacteria criterion for water contact sports. The underlying concept in the phrase "at least 99% of the time" is straightforward: there is some acceptable frequency – albeit relatively low (i.e., ≤ 1% of the time) – at which water quality criteria excursions are allowed without constituting criteria exceedances.

The fish and aquatic life criteria for toxic substances are different with regard to temporal aspects because these criteria are based on USEPA's recommended frequencies, which most often state that excursions from the criteria are not to occur more than once in three years on average (USEPA 1985, USEPA 1991, USEPA 2002b). This concept is discussed further in the *Criteria Duration Considerations* section, below. As noted previously, the human health criteria for toxic substances consider any criteria excursions to be unacceptable (USEPA 2000). Determining if WQS are met at the appropriate frequency requires context-specific considerations that take into account the particular standard(s) being evaluated and the expected site-specific patterns of variability.

MAKING DISCRETE PHYSICOCHEMICAL ASSESSMENTS

DEP will follow the determination framework (Chapter 5) when making use assessments using physicochemical data. In order to have sufficient data for assessment determinations, physicochemical data should be collected according to monitoring protocols in *Water Quality Monitoring Protocols for Streams and Rivers* (Shull and Lookenbill 2018), and data collections should consider the duration component of the water quality criterion, including the durations for toxic substance criteria as specified in Table 1 below. In the use assessment process, DEP must also consider the sampling design employed – including critical sampling periods, when applicable – sample size, and quality assurance methods (discussed below).

Judgement-Based Sampling

Due to interrelated considerations of reasonable decision error rates, sample sizes, and extreme percentiles of frequency distributions, it will often be impractical to employ a probability-based sample design – discussed further in *Water Use Assessment Decision-making Based on Physicochemical and Bacteriological Sampling* (Chalfant 2013) – to assess against meeting WQS "at least 99% of the time" or allowing for one excursion in a three year period without collecting large numbers of samples. Especially when accounting for monitoring costs, the most resource-effective approach will often be to focus monitoring at times when excursions are most likely to occur, hereafter referred to as *critical sampling periods*. Collecting samples during these critical sampling periods will require an understanding of the variables at play. Some of the variables that should be considered are discussed further in Chalfant (2013). According to USEPA (2002a), critical sampling periods can be thought of as temporal "hot spots," and sampling that is targeted to observe these "hot spots" based on an understanding of context-specific variations is a targeted sampling approach referred to as "*judgment-based sampling*" (as contrasted with probability-based sampling). Since judgment-based sampling is not as suited to some forms of quantitative statistical analyses as probability-based sampling, assessment processes based on judgement-based sampling may involve a different analytical toolset than assessment processes based on probability-based sampling.

DEP does not discount any data or information from consideration, so no strict guidelines are set with regard to what sampling designs are acceptable for assessments; however, of the various sampling plan designs discussed by USEPA (2002a), DEP believes that the judgment-based sampling design is the most suited method to assess extreme, infrequent ends of water quality parameter distributions. A judgment-based sampling design offers the benefit of more resource-efficient sampling (i.e., needing fewer observations to achieve a given level of precision) than other sampling designs when a sound understanding of the sites and systems being sampled is incorporated (USEPA 2002a).

Sample Size and Representativeness

Depending on the data available, the criterion being assessed, or the ambient conditions of the waterbody, a single observation (sample) can represent different periods of time. These factors may suggest that more than one sample needs to be collected to confidently make an assessment determination. USEPA guidance discourages rigid minimum sample size requirements and requires States to evaluate all existing, readily available, and appropriate water quality-related data for determining WQS attainment determinations (USEPA 2006, USEPA 2013). As a result of these factors and federal requirements, DEP recommends multiple sampling events for assessing any criterion but will evaluate all existing and readily available data when making assessments. More specifically, DEP encourages at least three sampling events within the criterion duration period. For example, the total iron criterion for aquatic life is written as a 30-day average, so DEP encourages at least three samples be collected within a 30-day period to compare conditions to the criterion.

DEP generally considers discrete samples to be representative of one day unless convincing evidence exists to suggest otherwise (e.g., a documented spill, influence of a known biological process, supporting high-frequency monitoring data). For most criteria, this strikes an acceptable balance between resource expenditure and sampling error. For some criteria, however, literal interpretation of criteria would result in sample collection that is very resource intensive and not feasible. For example, literal interpretation of the USEPA's national recommended toxic criteria would require samples to be collected at least once each day within a consecutive four-day period. Consequently, USEPA recommends that a single sample may represent periods longer than one day if conditions are stable (USEPA 1997). DEP conforms to this guidance and accepts that a sample collected for the assessment of time-averaged criteria can represent more than one day unless convincing evidence exists to suggest otherwise.

Criteria Duration, Frequency Considerations

<u>Instantaneous Criteria</u>

The allowable frequency of excursions depends on whether the criteria duration is instantaneous or time-averaged. For criteria expressed as instantaneous maxima or minima, there is no averaging period required to compare measured values to the criteria (i.e., the criteria magnitude duration is instantaneous). For these instantaneous criteria in § 93.7 (relating to specific water quality criteria), Table 3, the allowable frequency of excursions – unless otherwise specified as part of the criteria – follows Chapter 96, meaning that criteria must be met at least 99% of the time. For the human health criteria in § 93.8c (relating to human health and aquatic life criteria for toxic substances), Table 5, no excursions are allowable, as previously discussed above based on EPA's *Methodology for Deriving Ambient Water Quality Criteria for Human*

Health (USEPA 2000). As such, excursions of instantaneous criteria in § 93.7 occurring on four or more separate days within twelve months typically constitute an exceedance, as excursions from the criteria magnitudes occur more than 1% of the time (i.e., 4 days / 365 days ≈ 1.1%) and therefore the criteria are met less than 99% of the time.

Time-Averaged Criteria

For the aquatic life criteria at § 93.8c (relating to human health and aquatic life criteria for toxic substances), Table 5, adopted based on the National recommended criteria (see Table 1) for which the criteria durations are expressed as averages (e.g., seven-day, 30-day, or monthly averages), an excursion from the criteria is not to occur more than once in three years on average.

Other pollutants for which the criteria duration is expressed as averages – which includes several of the criteria in § 93.7 (relating to specific water quality criteria), Table 3, and the fish and aquatic life criteria at § 93.8c (relating to human health and aquatic life criteria for toxic substances) Table 5 that are not described in Table 1 below – any single averaging period showing an excursion from the mean concentration of the criterion will be considered an exceedance, and thus an impairment of the relevant protected use. For instance, a single, seven-day average dissolved oxygen observation below the criterion – which is considered a single excursion from the criterion – indicates an exceedance of the criterion, as there are 52 seven-day cycles per year, and one seven-day cycle is more than one percent of a year [100 * (1 seven-day cycle / 52 cycles / year) = 1.9% of the year]. Or, a single 30-day average ammonia concentration above the criterion magnitude indicates an of the criterion, as there are 12 30-day cycles per year, and one 30-day cycle is more than one percent of a year [100 * (1 30-day cycle / 12 cycles / year) ≈ 8.3% of the year]).

It is important to note that the criteria for toxic substances in § 93.8c do not specify durations. However, many of these fish and aquatic life criteria were adopted based on USEPA's *National Recommended Water Quality Criteria – Human Health Criteria* ("the national recommended criteria") (USEPA 2020), for which the aquatic life criterion maximum concentration (CMC, or acute criterion) and criterion continuous concentration (CCC, or chronic criterion) durations are provided in Table 1 below.

DEP will assess data based on the corresponding duration component each parameter's water quality criterion. The frequency components of the national recommended fish and aquatic life criteria for toxic substances indicate that more than one excursion in three years would result in a criterion exceedance and use impairment (USEPA 1985, USEPA 1991, USEPA 2002b).

Table 1. Criteria durations and other information for the fish and aquatic life criteria for toxic substances from § 93.8c (relating to human health and aquatic life criteria for toxic substances) including the Chemical Association System (CAS) number, Chemical Name, and durations for the Criterion Maximum Concentration (CMC, acute criterion) and Criterion Continuous Concentration (CCC, chronic criterion) as adopted from USEPA's national recommended criteria, identified by the Rationale Document.

CAS	Chemical Name	CMC Duration	CCC Duration	Rationale Document
333415	DIAZINON	1-hour average	4-day average	EPA-822-R-05-006
104405	NONYLPHENOL	1-hour average	4-day average	EPA-822-R-05-005
87865	PENTACHLORO-PHENOL	1-hour average	4-day average	EPA-820-B-96-001
7440382	ARSENIC (As3+)	1-hour average	4-day average	EPA-820-B-96-001
7440439	CADMIUM	1-hour average	4-day average	EPA-820-B-96-001
16065831	CHROMIUM III	1-hour average	4-day average	EPA-820-B-96-001
18540299	CHROMIUM VI	1-hour average	4-day average	EPA-820-B-96-001
7440508	COPPER	1-hour average	4-day average	EPA-820-B-96-001
7439921	LEAD	1-hour average	4-day average	EPA 440/5-84-027
7439976	MERCURY (Hg2+)	1-hour average	4-day average	EPA-820-B-96-001
7440020	NICKEL	1-hour average	4-day average	EPA-820-B-96-001
7782492	SELENIUM	1-hour average	4-day average	EPA-820-B-96-001
7440224	SILVER	Instantaneous	N/A	EPA 440/5-80-071
7440666	ZINC	1-hour average	1-hour average	EPA-820-B-96-001
57125	CYANIDE, FREE	1-hour average	4-day average	EPA-820-B-96-001
58899	gamma-BHC (LINDANE)	1-hour average	N/A	EPA-820-B-96-001
57749	CHLORDANE	Instantaneous	24-hour average	EPA 440/5-80-027
50293	4,4-DDT	Instantaneous	24-hour average	EPA 440/5-80-038
60571	DIELDRIN	1-hour average	4-day average	EPA-820-B-96-001
959988	alpha-ENDOSUL-FAN	Instantaneous	24-hour average	EPA 440/5-80-046
33213659	beta-ENDOSULFAN	Instantaneous	24-hour average	EPA 440/5-80-046
72208	ENDRIN	1-hour average	4-day average	EPA-820-B-96-001
76448	HEPTACHLOR	Instantaneous	24-hour average	EPA 440/5-80-052
1024573	HEPTACHLOR EPOXIDE	Instantaneous	24-hour average	EPA 440/5-80-052
--	PCB (Polychlorinated Biphenyls)	24-hour average	24-hour average	EPA 440/5-80-068
8001352	TOXAPHENE	1-hour average	4-day average	EPA 440/5-86-006
107028	ACROLEIN	1-hour average	1-hour average	Not Numbered
108883	TOLUENE	Not specified	Not specified	EPA 440/5-80-075

Fish and aquatic life criteria for toxic substances not described in Table 1 below, were determined by the Commonwealth rather than directly adopting USEPA's National recommendations (USEPA 1985). As such, the allowable frequency of occurrence follows Chapter 96 and criteria must be met at least 99% of the time.

As noted in the *Sample Size and Representativeness* section above, while DEP will evaluate all existing and readily available data, DEP encourages at least three samples be collected within the time-averaged criterion duration.

Quality Assurance

DEP makes every effort to verify the accuracy of all data used in the use assessment determination process. DEP strongly encourages anyone submitting data to familiarize themselves with DEP Bureau of Laboratories quality assurance and quality control procedures (DEP 2016) regarding record keeping, methods documentation, sampling techniques, selection of analytic laboratories, chain of custody concerns, and so forth. DEP will not exclude extreme values (outliers) from a dataset unless there is reason to believe the extreme value is invalid. For example, a dissolved oxygen concentration of 100 mg/L is physically impossible at tropospheric temperatures and pressures; it is likely that such a record is a typographical error actually meant to be 1 mg/L or 10 mg/L. Similarly, in a water temperature dataset submitted in degrees Celsius where one value is recorded at 72, it is highly unlikely this is a valid reading and may be recorded in degrees Fahrenheit.

Sample Precedence

In DEP's assessments of WQS, more recent data take precedence over older data, especially in situations where conditions have recently changed (e.g., installation of pollution remediation projects, alteration of permit limits in the watershed, changing land use patterns, discontinuation of combined sewer overflows). In some instances, older and newer data may be considered together to document temporal trends.

CONCLUSION

This document outlines much of what should be considered when conducting an assessment of water quality physicochemical data for making CWA 303(d) and 305(b) decisions. It should be read in its entirety and followed in combination with the *Water Quality Monitoring Protocols* (Shull and Lookenbill 2018). Additional technical information and rationale on physicochemical data collections and assessment decisions can be obtained in *Water Use Assessment Decision-making Based on Physicochemical and Bacteriological Sampling* (Chalfant 2013).

LITERATURE CITED

Chalfant, B.A. 2013. Water Use Assessment Decision-making Based on Physicochemical and Bacteriological Sampling. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

DEP. 2016. Quality Assurance Manual for the Pennsylvania Department of Environmental Protection Bureau of Laboratories. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

Hoger, M.S. 2018. Continuous Physicochemical Assessment Method. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

Shull, D. R., and M. J. Lookenbill (editors). 2018. Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

USEPA. 1985. Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses. U.S. Environmental Protection Agency, Office of Research and Development. PB85-227049

USEPA. 1991. Technical Support Document for Water Quality-Based Toxics Control. EPA/505/2-90-001. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

USEPA. 1997. Guidelines for Preparation of the Comprehensive State Water Quality Assessments (305(b) Reports) and Electronic Updates: Supplement. EPA-841-B-97-002B. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

USEPA. 2000. Methodology for Deriving Ambient Water Quality Criteria for the Protection of Human Health. EPA 822-B-00-004. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

USEPA. 2002a. Guidance on Choosing a Sampling Design for Environmental Data Collection for Use in Developing a Quality Assurance Project Plan. EPA QA/G-5S. U.S. Environmental Protection Agency, Office of Environmental Information, Washington, D.C.

USEPA. 2002b. Consolidated Assessment and Listing Methodology – Toward a Compendium of Best Practices. First edition. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

USEPA. 2005. Guidance for 2006 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d), 305(b) and 314 of the Clean Water Act. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

USEPA. 2006. Guidance on Systematic Planning Using the Data Quality Objectives Process. EPA QA/G-4. U.S. Environmental Protection Agency, Office of Environmental Information, Washington, D.C.

USEPA. 2013. Information Concerning 2014 Clean Water Act Sections 303(d), 35(b), and 314 Integrated Reporting and Listing Decisions. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

USEPA. 2020. National Recommended Water Quality Criteria – Human Health Criteria Table. U.S. Environmental Protection Agency, Office of Water, Washington, D.C., 2020 Feb 20 [accessed 2021 Jan 8].

Walters, G., and M. Pulket. 2015. Assessment and Delisting Methods. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.