

Quality Assurance Project Plan

For

Reporting of Pennsylvania NPDES Point Source Data to EPA's Chesapeake Bay Program

PADEP QA Manager

(signature)

Approved:

(date)

PADEP Chief of Data Systems Section

(signature)

(date)

EPA Project Officer (Nita Sylvester)

(signature)

(date)

EPA QA Officer (Richard Batiuk)

(signature)

(date)

Distribution List

PADEP

Robert Kachonik, Environmental Group Manager, Data Systems Section, BPNPSM
Sean Furjanic, Environmental Program Manager, BPNPSM
Lori Mitzel, Water Program Specialist, Data Systems Section, BPNPSM
Michele Leota, Business Analyst 2, Office of Program Integration

EPA

Nita Sylvester, EPA Project Officer
Richard Batiuk, EPA QA Officer

Table of Contents

Acronym List iii

Introduction 1

Reporting Procedures..... 1

QA/QC Procedures..... 2

 Facilities 2

 Effluent Monitoring Data..... 3

 Data Collection 3

 Identifying Data Gaps 3

 Estimates and Defaults..... 4

 QA/QC Protocol 5

Organization Chart..... 7

Acronym List

BOD₅ – Five-day biochemical oxygen demand

CBPO - U.S. Environmental Protection Agency’s Chesapeake Bay Program Office

DMR – Discharge Monitoring Report

eDMR - electronic Discharge Monitoring Report system

EPA - U.S. Environmental Protection Agency

ICIS – EPA’s Integrated Compliance Information System

MDL - Method Detection Limit

NH₃-N – Ammonia nitrogen

NO₂/NO₃-N – Nitrite and nitrate as nitrogen

NPDES – National Pollutant Discharge Elimination System

PADEP - Pennsylvania Department of Environmental Protection

PO₄ - Phosphate

MGD – Million Gallons per Day

QAPP - Quality Assurance Project Plan

QA/QC - Quality Assurance / Quality Control

RL - Reporting Limit

SIS - PADEP Sample Information System

TKN – Total Kjeldahl Nitrogen

Total N – Total Nitrogen

Total P – Total Phosphorus

BPNSM – PADEP Bureau of Point and Non-Point Source Management

**PA Point Source Quality Assurance Project Plan
Revised December 2014**

INTRODUCTION

This Quality Assurance Project Plan (QAPP) was developed for documenting quality assurance / quality control (QA/QC) activities that the Pennsylvania Department of Environmental Protection (PADEP) will perform prior to submission of National Pollutant Discharge Elimination System (NPDES) facility effluent monitoring data (point source data) to U.S. Environmental Protection Agency’s (EPA’s) Chesapeake Bay Program Office (CBPO). The QAPP is a required deliverable, from PADEP to EPA, under the Chesapeake Bay Implementation Grant. This QAPP will be reviewed periodically and revised, as necessary, as data reporting requests or requirements change. All updates to this document will be prepared in draft form and sent to EPA CBPO for concurrence prior to implementation. The Data Systems Section within PADEP’s Bureau of Point and Non-Point Source Management (BPNPSM) will be the contact for Chesapeake Bay point source data reporting.

REPORTING PROCEDURES

PADEP will submit its point source data report to EPA CBPO on an annual basis by December 1, which will include point source data for the previous July 1 – June 30 period.

Listed below are the parameters that will be reported to EPA in the point source data report, at a minimum:

Parameter Name	ICIS Parameter Code	Statistical Base Code(s) Reported
Dissolved Oxygen	300	Minimum*
Total Nitrogen	00600, 51445	Average Monthly (mg/l), Total Monthly (lbs)
Ammonia Nitrogen	00610, 51446	Average Monthly (mg/l), Total Monthly (lbs)
Total Kjeldahl Nitrogen	00625, 51449	Average Monthly (mg/l), Total Monthly (lbs)
Nitrate and Nitrite Nitrogen	00630, 51450	Average Monthly (mg/l), Total Monthly (lbs)
Total Phosphorus**	00665, 51451	Average Monthly (mg/l), Total Monthly (lbs)
Flow	50050	Average Monthly (MGD)
CBOD ₅	80082	Average Monthly (mg/l and lbs/day)
Total Suspended Solids	00530	Average Monthly (mg/l and lbs/day)

* Dissolved Oxygen concentrations are generally only reported on DMRs in PA as minimum values obtained during the month.

** PA facilities typically do not monitor for phosphates, but rather Total Phosphorus, although the data reporting guidance requested phosphate data. PADEP will use the general species relationship for PO₄/TP presented in the guidance to report phosphate concentrations and loads. [Note: EPA Default: Calculated as 85% of TP by CBP species ratio.]

Data quantity will conform to EPA CBPO’s latest grant guidance. All concentrations will be presented as gross values for sewage facilities, i.e., subtraction to account for ambient or background concentrations (net concentrations and loads) will not be conducted unless specifically requested. Concentrations for industrial facilities will be presented as net values per the grant guidance; however, until NPDES permits require collection of influent samples and analysis for Chesapeake Bay parameters, influent sample concentrations will be assumed to be zero (i.e., in the absence of any information on ambient concentrations, it should be assumed, without further EPA guidance, that the gross values reported to EPA are net values). If influent sample data are available for particular facilities, the values will be subtracted from effluent values so that accurate net values can be

**PA Point Source Quality Assurance Project Plan
Revised December 2014**

reported, accompanied by appropriate documentation. The report will be transmitted to the CBPO point of contact as an Excel spreadsheet file. All rationale for assumptions or estimates will be explained in the key for the data source abbreviations.

QA/QC PROCEDURES

PADEP will perform a series of procedures to assure consistency and integrity in the list of NPDES facilities that are reported and point source data. These procedures are discussed sequentially below.

Facilities

PADEP has segregated sewage (municipal and non-municipal) discharges in the Chesapeake Bay watershed into five groups or phases, based on discharge flow rates, corresponding to the timing of NPDES permitting requirements for attainment of annual Total Nitrogen (Total N) and Total Phosphorus (Total P) mass load limitations. Sewage discharges with design flows of at least 0.4 MGD, are, for reporting purposes, assigned to Phases 1 through 3 ("significant"). There are 190 significant sewage dischargers at this time.

PADEP has delineated industrial waste discharges in the Chesapeake Bay watershed into two groups, based on the estimated Total N and Total P loads. An industrial waste discharge with an estimated Total N and Total P loads of 75 lbs/day and 25 lbs/day, respectively, is considered "significant". There are, at this time, 23 industrial waste discharges in PA that are considered significant.

The total number of PA "Significant Bay Dischargers" is now 213. BPNPSM monitors NPDES permits issued by its regional offices, and updates the list of Significant Bay Dischargers as necessary in its Phase 2 Watershed Implementation Plan (WIP) Wastewater Supplement document, available on DEP's website (www.depweb.state.pa.us/npdes-bay) and provided to EPA in Section 106 Grant Semi-Annual Status Reports.

Newly identified facilities will be communicated to EPA via email. The location of each facility outfall will be reported by county and by latitude/longitude coordinates.

At EPA's request, PADEP continues to include effluent data on the point source data reports for facilities that are no longer considered significant Chesapeake Bay dischargers. These facilities are clearly identified in the submissions.

Effluent Monitoring Data

Data Collection

At this time, most of the 213 Significant Bay Dischargers have monitoring requirements in their NPDES permits for the full suite of parameters that must be reported to EPA. The effluent limits and DMRs for all past and current significant Bay Dischargers are coded in ICIS. Therefore, with few exceptions, all of the reported data needed to construct the annual Bay report should be available in ICIS.

As a first step in the data collection process, PADEP will query ICIS to gather all available monitoring data for the period of record. The resulting report is presented with data in rows; this format is not amenable to QA/QC activities, and therefore PADEP uses Access to convert data in rows to columns (with parameter names at the top of the worksheet). The information is retained in an Excel spreadsheet while additional data collection steps are completed.

The final source of monitoring data is PADEP's eDMR system. Beginning with the June 2009 monitoring period, all eDMR data for Significant Bay Facilities is automatically uploaded to and available from ICIS.

There are still a few permits in PA that do not have the full suite of monitoring requirements. When no data are available due to lack of monitoring, PADEP will use approaches identified below to develop concentration estimates for reporting purposes.

Identifying Data Gaps

Although total loadings are not part of the data requested in the grant guidance, most PA Bay permits require the reporting of total monthly loads for Total Nitrogen, Ammonia Nitrogen, TKN, Nitrate/Nitrite and Total Phosphorus. These values for total monthly loads are critical in evaluating the reported monthly average concentration values for accuracy. Consequently, PADEP initially builds a spreadsheet, the "Reported Data Spreadsheet", which includes the total monthly loads. Missing total monthly loads, as well as missing concentration values, will be considered data gaps.

Following the data collection process, PADEP will make an inventory of the missing data, which would fall into one of the following categories: 1) there was no discharge from the facility, resulting in no data, 2) there was a discharge and analyses were completed, but data are not available, and 3) there was a discharge and analyses were not completed.

The following efforts will be made to identify and fill data gaps:

- Identification of "No Discharge" Data Gaps – PADEP will run an ICIS report to identify "No Discharge" reported on DMRs. Where the information appears suspect, for instance where only one parameter of an outfall is missing, or where there is no flow from a municipal STP, PADEP will request review by regional staff for verification of "No Discharge" reporting. If data are missing as a result of there not being a discharge during the period of interest, fields will remain blank and will be annotated with a footnote.
- Identification of Data Gaps Where Analyses Were and Were Not Completed – PADEP will inspect actual permits to verify that monitoring was required by the permit. If permits did require monitoring but no data are available, PADEP will contact regional staff and facilities to resolve the issue and complete blank fields in the database to the extent possible. If permits did not require monitoring, PADEP will implement measures described below.

Estimates and Defaults

Upon completion of the identification of data gaps, PADEP will assign values to each field (except for the "No Discharge" scenario), using the following approaches, presented in order of preference:

- In the event that zero (0) or "non-detect" values are reported, PADEP will research and identify an appropriate quantitation limit (QL) for the parameter, and replace the non-detect with the chosen QL. Currently, zero or non-detect values are replaced with the PADEP Bureau of Laboratories' QLs. Current replacement values are (mg/l): Ammonia-Nitrogen = 0.02; Total Kjeldahl Nitrogen = 1.0; Nitrate+Nitrite-Nitrogen = 0.04; Total Phosphorus = 0.01; BOD = 0.2; CBOD = 0.2; Total Suspended Solids = 2.0. If data are reported as "< QL", the "<" sign is dropped, and the value is used. For example, for a reported value of "< 0.5", the less than sign is dropped and "0.5" is used.
- If only partial monthly data is reported, calculations will be performed where appropriate to determine or estimate the remaining data. For instance: 1) if Total Nitrogen and NO₂₋₃ concentrations are reported, TKN will be calculated from the reported data [Note: EPA Default - Calculated as NO₂₋₃ = TN – TKN]; 2) If monthly average flow and Total Nitrogen concentration is reported, the estimated monthly Total Nitrogen load will be calculated.
- Where parameter value data has been required on a quarterly basis rather than monthly or more frequently, the quarterly value will be applied to the other months in that monitoring period.
- If parameter value data exists for a facility for 8 or more months during the annual reporting period but not others, average data for available months for that facility will be used to populate blank fields. It is expected that this rare situation would occur only where sample collection or analysis was overlooked, or where there are newly instituted nutrient monitoring requirements. [Note: EPA Default - Average of reported monthly data.]

PA Point Source Quality Assurance Project Plan
Revised December 2014

- If parameter value data does not exist for 4 months or more during the reporting period, prior reports will be consulted to see if data (either default, reported or a combination) are available for the reporting period. If available, average data for that previous 12 month period will be used, with possible adjustments made due to seasonal factors. For example, if a facility produced no Total Nitrogen data during September 1 – December 31, 2009 but Total Nitrogen data are available for the period of July 1 – December 31, 2008, average data for the second half of 2008 would be reported for each blank month for the 2009 period. A situation where data was reported in the past but not currently could occur if a permit allows a facility to discontinue monitoring of parameters during the permit term or a permit is reissued without the same level of monitoring. [Note: EPA default - Default value based on state specific information.]
- If parameter value data does not exist for any months during the reporting period, and no historical data are available for the parameter at the facility, PADEP will apply EPA default data based on the level of treatment or, in the event default data are not available for parameters, PADEP will use best professional judgment with documentation. For example, PADEP will consider utilizing its Sample Information System (SIS) database to determine appropriate concentration values. [Note: EPA default: Default value based on SIS database.] SIS is a compilation of grab samples collected by PADEP inspectors during wastewater facility inspections.

If EPA defaults are chosen, the following default concentrations will be used, as supplied by EPA CBPO: [Note: EPA Default - Default values agreed to by the workgroup.]

Parameter	Secondary Treatment	Tertiary Treatment
NH ₃ -N	2.6	0.9
BOD ₅	13	7.3
Total N	11.2	5.6
Total P	2	0.8
TKN	6.7	2.8
TSS	13.85	7.05

The Reported Data Spreadsheet on which PADEP collects reported data will flag (qualify) all data that is modified based on estimates and assumptions. Where a modified value extends into the spreadsheet for submission to EPA CBPO, the qualifier will be shown on the submitted spreadsheet as well. Data reported to EPA CBPO will include qualifier codes that document all estimates and assumptions. PADEP will attempt to develop standardized qualifier codes that are consistent between submissions. If CBPO adopts its own qualifier codes through guidance, it is anticipated that PADEP will adopt these codes.

QA/QC Protocol

Once all data fields are populated, PADEP will conduct quality assurance on the data as follows:

- As the procedures listed above are conducted, any data that appears to be outside the predominant data pattern for an individual facility will be highlighted for further evaluation. Expected seasonal variations and across the board high and low flow trends will be taken into account.
- PADEP will use judgment in identifying data entry errors, and make appropriate corrections. For example, if average monthly Total Phosphorus concentrations are 1.2, 1.5, 1.8, 1.8, 1.4, and 0.18 mg/l, PADEP will assume that the facility or regional staff meant to enter 1.8 mg/l and will make the change accordingly. However, PADEP will apply caution to the rejection of data that could be lower than historical statistics, as it is possible a facility has implemented nutrient removal technologies. PADEP will contact regional staff and/or facilities as necessary to verify such improvements.

PA Point Source Quality Assurance Project Plan
Revised December 2014

- PADEP will also use technical judgment for replacing data that appear to be grossly miscalculated. For example, if the reported Total Nitrogen monthly load is significantly lower (less than half) than its expected level based on flows and Total Nitrogen concentrations, a DEP-calculated Total Nitrogen load may replace the reported value. For facilities where these types of gross miscalculations are prevalent, PADEP will contact the permittee to discuss the correct calculation procedures.
- Other QA/QC Procedures – PADEP will apply the following rules in validating data in constructing the Reported Data Spreadsheet:
 - TKN concentration and load must be greater than the NH₃-N concentration and load. In the event that NH₃-N values exceed TKN values, the TKN values will be matched to the NH₃-N values.
 - Total N values must equal TKN values plus NO₂/NO₃-N values. If not, Total N values will be adjusted to equal TKN plus NO₂/NO₃-N. [Note: EPA Default - Calculated as NO₂₋₃ = TN - TKN.]
 - Total P values must exceed PO₄ values (since facilities typically monitor for only Total P, it is rare that this verification will need to be performed).
 - Monthly Average flows are less than or equal to the Maximum Daily Flow.
 - Verify that there are no missing or negative values in the report.
 - As time allows, where data discrepancies are discovered for users of PA's eDMR system, the eDMR submission will be examined for inclusion of DMR supplemental forms where raw sampling data is recorded. The permittee's calculation method and their reporting accuracy will be evaluated. The statistic in question maybe recalculated using the raw data. Data errors will be corrected and flagged as qualified data on the DEP Reported Data Spreadsheet.
 - As resources allow, the following additional QA/QC procedures may be performed:
 - Evaluation of Facility-Specific Data Trends and Variability – PADEP may compile all historical non-default monitoring data for a facility (as submitted to EPA CBPO and as available) that have undergone quality assurance review and generate mean and standard deviation values for concentrations. The period of record would then be compared with the historical data to determine outliers. A reported value in the period of record would be rejected if the value is greater than or equal to three standard deviations from the historical mean value. For example, if Total Nitrogen was monitored during 2007 and 2008, mean and standard deviation values will be determined for those years and serve as baseline for comparison to 2009 values. If an outlier is identified, the value will be replaced with the historical mean value. This replacement will not apply if site-specific information is available to support the apparent outlier, for example, newly completed construction of facility upgrades.
 - For situations where there are no historical data, PADEP may conduct a “sensitivity analysis” to evaluate whether one or more outliers exist by removing values that appear to be outside normal variability for the period of interest and determining the mean and standard deviation of the remaining data. If the inclusion of data suspected to be outliers results in an increase in the standard deviation of 300% or more, the data will be removed and replaced with the mean value for the remaining data, or a calculated value based on the species relationships, or actual flow and concentrations values, if applicable.

ORGANIZATION CHART

