Chesapeake Bay Point Source Data Project: Common Themes in Jurisdictional Experience and Recommendations for Streamlining

Introduction

This document highlights common themes among the jurisdictions in gathering, reviewing, and submitting point source data to the Chesapeake Bay Program for the progress model run each year. More information about specific experiences, and who attended each meeting, is available in the meeting notes from each conversation, as well as the Stakeholder Register for this project. The newly finalized NPDES E-Reporting rule brings additional changes and requirements, which will impact the way jurisdictions submit data to the Bay program. This document concludes with a recommended approach to streamlining the process for submitting and reviewing data to be used in the Chesapeake Bay progress model.

Summary of Input

What are the barriers you see to submitting uniform data across the jurisdictions?

- Jurisdictions see differences in each other's QA/QC processes that would make developing a single tool or portal difficult.
- Differences exist among jurisdictions in the treatment of nonsignificant facilities, in terms of the data available and the extent to which default values are used.
- A tool would have to accomm33odate entry of different formats of data (from a state system, from an Excel spreadsheet), or the jurisdictions would have to agree to a common entry format.
- Jurisdictions also differ in the proportion of facilities within their jurisdiction are included in the Chesapeake Bay Watershed and, therefore, are of concern to the Bay Program.

What does your current data gathering process look like?

- Some jurisdictions start their data gathering process with a pull from ICIS-NPDES, while others take data directly from the state database that feeds ICIS-NPDES.
- The effort required to gather data for significant facilities is substantial, but the true pain point is combining this data with data (or lack thereof) for nonsignificant facilities.
- Jurisdictions are in various states of transition to an electronic reporting system, through either the EPA system (net DMR) or a state-built or state-customized e-DMR system. When reissuing permits, most jurisdictions add a new requirement to report electronically.
- Region 3 is in the process of analyzing how the NPDES e-Reporting rule will affect its jurisdictions.
    - Preliminary results of this analysis show that several states will have to complete CROMERR approval of their system (or use the federal system, net-DMR) to comply with the new rule.

  - Rule will require state systems to interface using net or e-DMR with the EPA NPDES system (ICIS-NPDES).
- One of the challenges of data collection is the required format for some systems. For example, net DMR does not accept raw data, only calculated statistics (e.g., average flow) that could contain errors made when performing these calculations.
- Use of default values:
  - Many jurisdictions use default values per Attachment 6 of the grant guidance for nonsignificant facilities, either because they don't have the data (not required to report) or, in some cases, the data set of nonsignificant facilities is too large to assess.
    - PA raised the issue of data quality in relation to default values. In PA's experience, default values can fail to represent the progress that is actually occurring.
    - Once PA began to monitor those facilities for which it had used default values, it found the actual data to be lower than the default values.
    - Default values help other jurisdictions fill in the blanks for facilities with irregular or infrequent reporting requirements.
    - The workgroup recognizes a balance between the value of real data and the payoff received from the high level of effort associated with gathering and QA/QCing the data.
  - Moving away from using default values places more burden on jurisdictions to QA/QC data for nonsignificant facilities in addition to data for significant facilities.
- Jurisdictions are in agreement that there is room for streamlining in the data gathering, review, and submission process.

What does your current QA/QC of gathered data look like?

- States have invested in different QA/QC tools to save time and effort, including:
  - macros built into spreadsheets to pull data from ICIS-NPDES
  - run programs (e.g., SAS) to check data
  - series of database queries
  - investing in new e-DMR systems with validation and error catching built into the system
- An additonal QA/QC tool may be useful, if it provided benefits in addition to those provided by existing jurisdiction tools.
  - Jurisdiction data managers can easily see missing data in the current process.
  - A tool will help most if it is able to perform more sophisticated checks, like identifying outliers and checking for expected trends in the data.

Can ICIS-NPDES help report the parameters the Bay program is interested in?

- There was general agreement among jurisdictions that building a separate system for submitting data would be redundant. It would be better to create the QA/QC tools to make better use of existing systems.
- Many jurisdictions do not require facilities to report all parameters the Bay Program is interested in.
  - For example, many report simply total nitrogen and total phosphorus.

  - Different species behave differently; thus the Chesapeake Bay model requires speciation of nitrogen and phosphorus.
  - Some jurisdictions are moving towards including speciation parameters as reporting requirements. With this increase in data comes an increase in the effort to gather and, in particular, to review the data with a robust QA/QC process prior to submission to the Bay Program.
  - A few jurisdictions are moving away from including more reporting requirements.
    - Having different species that don't add up to the total makes the QA/QC lengthier.
    - These jurisdictions use default values or design flow to fill in missing parameters.
  - Most jurisdictions do not have regular reporting requirements for nonsignificant facilities, for a variety of reasons, including resources, level of effort, quality of estimated data, etc.
  - Without standardization of reporting requirements, ICIS-NPDES may be difficult to use as a data source for Bay Program model runs.
- Jurisdictions see challenges associated with getting minor permitees to report data into ICIS-NPDES (i.e., changing and enforcing additional reporting requirements).
- ICIS-NPDES does not allow the flexibility needed for entry of actual data.
  - The system doesn't allow data entry for intermittent flow; it assumes the entered data is a daily flow.
  - System does not accept negative values, which may occur for non-contact cooling waters.
    - This means the net value is calculated over time, resulting in a manipulation of the data.
    - Eventually, the negative and positive values cancel each other out over time.
  - Any tool built to enhance ICIS-NPDES (data quality tool) or upload data to the Bay Program would have to be flexible enough to not prevent entering of accurate data (e.g., flag something as an outlier, but do not prohibit its entry).
- There were varied views on the quality of data in ICIS-NPDES.
  - For some jurisdictions, data quality of ICIS-NPDES is not an important issue because it is a reflection or duplication of the data in their state database, until they do the QA/QC of the data prior to submission to the Bay Program.
  - Others cited data quality issues that led to unnecessary work investigating possible compliance issues based on data errors.
- Some jurisdictions expressed support for an automatic data flow from ICIS-NPDES to the Bay Program, but this was not widely expressed. The Bay Program would need to bring this suggestion back to the larger group to get a better read on how it would affect the process of all jurisdictions.

How do your progress run data and your compliance data sets differ?

- Jurisdictions agree that the submitted DMR (electronic or in paper) is legally definitive, regardless of what system it is entered into.

- Data errors should be corrected by requesting and accepting a revised DMR from the facility.
- For electronic DMRs, this corrected data can easily flow into the state database and/or ICIS-NPDES.

Major Challenges and Themes

- **Speciation** makes ICIS-NPDES a less useful tool and requires more QA/QC to make totals add up. It requires some data manipulation—changing raw data to fit reported total.
- **Data quality in ICIS-NPDES** is still lacking but not necessarily prohibitive to using ICIS-NPDES as a starting point for a way to collect Bay Program data.
- There is no legal incentive to **consolidate places where data is collected** (ICIS-NPDES vs Bay Program), as the submitted DMR is the basis for legal action. There is, however, a **process incentive to improve the quality of the data**, as jurisdictions often find they can address flagged compliance issues with a correction to a data error.
- The Bay Program needs to better understand the **existing QA/QC tools** that jurisdictions have to be able to create a tool that adds more value.
- Jurisdictions **vary on reporting requirements** in permits—and whether the trend is to include more or fewer parameter requirements, and whether more or fewer nonsignificant facilities will have reporting requirements.
- The group varies in opinion concerning the use of **default values**.
- Shared challenges of issues and data—reiterate verbally as well.

Recommendation

Based on this feedback, the Bay Program offers a recommendation for consideration that incorporates (1) separating the significant facilities from nonsignificant facilities, and (2) creating a pre-processor to assist in QA/QC and preparing the data for the Chesapeake Bay Model.

The Chesapeake Bay Program will build a pre-processor tool that will review and accept the submissions of the jurisdictions for the progress model run each year. The pre-processor will have **two components**: one for **significant facilities**, and one for **nonsignificant facilities**. The functionality of each component is discussed below; the details and nuances of these functionalities will be discussed in greater detail with the jurisdictions during the development of the tool. The outputs of each component would be reviewed by the jurisdiction prior to submitting the data sets within the pre-processor tool.

Focus on significants—pilot this first, iteratively, revisit nonsignificants

The **Significant Facilities Component** will enable the Bay Program to pull data directly from ICIS-NPDES for significant facilities. The component would be programmed to complete tasks such as:

- Reviewing the data for outliers
- Checking submitted data against historical data to review the data trend
- Flag missing data (and possibly add last year's values, with permission)
- Automate speciation—either estimate speciation from a given total or adjust reported numbers to add to total of species based on agreed-upon ratios

The **Nonsignificant Facilities Component** will empower the jurisdictions to gather data for nonsignificant facilities faster and with more accuracy. The component would be programmed to enable jurisdictions to:

- Reviewing the data for outliers
- Checking submitted data against historical data to review the data trend
- Flag missing data (and possibly add last year's values, with permission)
- Add in default values, if desired
- Automate speciation—either estimate speciation from a given total or adjust reported numbers to add to total of species based on agreed-upon ratios

The intent of this tool is to ease the burden of preparing already-reported numbers for the Bay Program Model, based on the needs of the model. The tool would provide a path for submission as well.

In the long-term, the Bay Program would like to explore further the **reporting requirements jurisdictions assign to nonsignificant facilities**. If jurisdictions trend towards increasing reporting requirements or gathering more data for nonsignificant facilities, this data could be incorporated into the "significant facilities component" of this tool, relieving jurisdictions of the burden of preparing this data separately.

If there is agreement to move forward, come back to the group with a demo of existing QA/QC tool (Duet).

Language on resources available: Opportunities

Exchange Network grants priority for complying with e-Reporting requirements

WIP assistance funds