

Project Planning and Documentation for Projects using Existing Data
Illustration of the checklist, *Using Data from Other Sources - A Checklist for Quality Concerns*

Project planning for modeling projects is important in order to ensure that the model is scientifically sound, robust and defensible. EPA's Quality Staff has developed a draft checklist for use in planning modeling and other projects using existing data, *Using Data from Other Sources - A Checklist for Quality Concerns*. A past GLNPO modeling project, Lake Erie Total Phosphorus Loads, illustrates how the checklist can be used for planning and documenting a project as detailed below.

1. *Identify the decision you are making or the project objectives:* The overall objective of the project was to revive phosphorus load estimation efforts for the Great Lakes, using Lake Erie (1996-2000) as an example. The information required by GLNPO was the estimate of total phosphorus load for Lake Erie for the years 1996-2000. The over-arching decision was whether changes in environmental management were needed to address phosphorus loads to Lake Erie.
2. *Identify the data and information from outside sources proposed for the project:* A list of government agency databases was identified and used to provide data for input to study models.
3. *Determine whether the data have any constraints affecting their use in the new project:* The government agency databases that were used were verified to be available and accessible for use in the project.
4. *Determine where the acquired data will be used in the decision making process:* The data were used as inputs to the phosphorus loading model. Use of the data and model calculations were detailed.
5. *Scrutinize data for quality concerns pertinent to the intended use:* The Principal Investigator (PI) and GLNPO noted that the responsibility for basic data review, validation and verification was with the agency that collected the data. However, additional data screening for the purpose of load estimation was conducted prior to input to the model. The PI used statistical programs to identify outliers by checking for internal consistency and comparing to historical information. A plan was developed to investigate outliers and determine when data points would not be used in the model. Another quality concern involved the critical assumption that the quality objectives and criteria associated with the targeted agency databases would be adequate for the purposes of this project. GLNPO and the PI noted that numerous studies have been conducted in the past to ensure that this was the case. Data comparability also was a concern. A main project requirement was that the monitoring data should be obtained in the same way as previous load estimates to ensure comparability to historical data. The PI determined that the flow and total phosphorus data currently being generated by the agencies responsible for the existing databases are of comparable quality to data reported by these agencies previously. The PI noted one exception and developed a plan to address this data issue.
6. *Document your analysis plan in a Quality Assurance Project Plan (QAPP):* GLNPO and the PI developed a QAPP according to EPA QA/R-5, *EPA Requirements for Quality Assurance Project Plans*. In accordance with the graded approach, the QAPP noted when components discussed in R-5 were not applicable to the project. For example, Section B.4 Sampling Methods, was cited as "not applicable" because the project used existing data and environmental sampling was not a component of the project (see more on this at the end of this section). The plan also included a detailed description and citations for peer reviewed equations used in the model.
7. *Execute your analyses and document the outcome appropriately:* The QAPP included a schedule for model completion and development of a final report. As noted under Step 5, a main project requirement was data comparability to historical data. To address this, the PI used standard methods for load estimation and documented any deviations in the final report.