Quality Assurance Project Plan

for

Total Maximum Daily Load (TMDL) and Model Development for Floyd's Fork

Contract Number EP-C-08-004, Task Order 83

Prepared for:

United States Environmental Protection Agency Region 4
61 Forsyth Street, SW
Atlanta, GA 30303-8960

Prepared by:

Tetra Tech, Inc.
2110 Powers Ferry Road SE, Suite 202
Atlanta, GA 30339

July 15, 2011
QAPP 296, Draft Revision 0

This quality assurance project plan (QAPP) has been prepared according to guidance provided in the following documents to ensure that environmental and related data collected, compiled, or generated for this project are complete, accurate, and of the type, quantity, and quality required for their intended use:

- *EPA Requirements for Quality Assurance Project Plans.* (EPA QA/R-5, EPA/240/B-01/003, U.S. Environmental Protection Agency, Office of Environmental Information, Washington DC, March 2001 [Reissued May 2006]) (USEPA 2001)
- *Guidance for Quality Assurance Project Plans for Modeling.* (U.S. EPA QA/G-5M, EPA/240/R-02/007, U.S. Environmental Protection Agency, Office of Environmental Information, Washington, DC, December 2002) (USEPA 2002a)
- *EPA New England Quality Assurance Project Plan (QAPP). Checklist for Model Applications.* (U.S Environmental Protection Agency, Region 1, Boston, MA, 2009) (USEPA 2009a)
- *Guidance on the Development, Evaluation, and Application of Environmental Models* (EPA/100/K-09/003, U.S. Environmental Protection Agency, Office of the Science Advisor, Council for Regulatory Environmental Modeling, Washington DC, March 2009) (USEPA 2009b)
- *EPA Office of Water Quality Management Plan.* (EPA/821/R-09/001, U.S. Environmental Protection Agency, Office of Water, Washington DC, February 2009).

Tetra Tech, Inc., will conduct work in conformance with the procedures detailed in this QAPP.

Approvals:

Brian Watson, P.E., P.H.
Task Order Leader
Tetra Tech, Inc.
7/21/2011
Date

Steven Davie, P.E.
Modeling QC Officer
Tetra Tech, Inc.
7/21/2011
Date

John O'Donnell
Quality Assurance Officer
Tetra Tech, Inc.
7/22/11
Date

Approvals, continued

Tim Wool	7/21/11	William Melville	7/21/11
Date		Date	
Task Order Manager		Quality Assurance Coordinator	
U.S. Environmental Protection Agency, Region 4		U.S. Environmental Protection Agency, Region 4	

Contents

Tables iv
Figures iv
Acryonyms and Abbreviations v
Distribution vi

Project Management

1.0 Project/Task Organization 1
2.0 Problem Definition/Background 3
3.0 Project/Task Description 4
4.0 Quality Objectives and Criteria 10
5.0 Special Training Requirements/Certification 18
6.0 Documentation and Records 18

Data Acquisition

7.0 Sampling Process Design 19
8.0 Sampling Methods 20
9.0 Sample Handling and Custody 20
10.0 Analytical Methods 20
11.0 Quality Control 20
12.0 Instrument/Equipment Testing, Inspection, and Maintenance 21
13.0 Instrument/Equipment (Model) Calibration and Frequency 22
14.0 Inspection/Acceptance of Supplies and Consumables 23
15.0 Nondirect Measurements 23
16.0 Data Management 23

Assessment/Oversight

17.0 Assessment and Response Actions 23
18.0 Reports to Management 27

Data Validation and Usability

19.0 Data Review, Verification, and Validation 27
20.0 Verification and Validation Methods 28
21.0 Reconciliation with User Requirements 29
References 29

Tables

Table 1 Project Schedule 9
Table 2 Statistical Measures for Model Comparisons (USEPA 2000) 13
Table 3 Relative errors and statistical targets for hydrologic calibration 13
Table 4 Secondary Environmental Data to Be Collected for the Floyds Fork TMDL Development .. 17

Figures

Figure 1. Project organization 2
Figure 2. Problem assessment and correction options 22
Figure 3. Corrective Action Request and Response Verification form 23

Appendices

A. Data Quality Indicator Definitions

B. Data Quality Requirements for Secondary and Calculated Data

Acronyms and Abbreviations

BMP	Best management practice
BOD	Biochemical oxygen demand
DO	Dissolved oxygen
DQO	Data quality objective
EFDC	Environmental Fluid Dynamic Code
EPA	U.S. Environmental Protection Agency
GIS	Geographic information system
LSPC	Loading Simulation Program in C++
N	Nitrogen
NCDC	National Climatic Data Center
NED	National Elevation Dataset
NH_3	Ammonia
NHD+	National Hydrography Dataset
NLCD	National Land Cover Data
NO_2^-	Nitrite
NO_3^-	Nitrate
NOAA	National Oceanic and Atmospheric Administration
NPDES	National Pollutant Discharge Elimination System
NRCS	Natural Resources Conservation Service
P	Phosphorus
Phosphate	phosphate, PO_4^{3-}
QA	Quality assurance
QAPP	Quality assurance project plan
QC	Quality control
SOD	Sediment oxygen demand
TMDL	Total maximum daily load
TN	Total nitrogen
TOL	Task Order Leader
TOM	Task Order Manager
TP	Total phosphorus
USDA	U.S. Department of Agriculture
USGS	U.S. Geological Survey
WASP	Water Quality Analysis and Simulation Program
WQS	Water quality standard
WRDB	Water Resources Database

Distribution

This document will be distributed to the following U.S. Environmental Protection Agency and Tetra Tech, Inc., personnel and any subcontractor staff involved in the project.

Name	Phone, fax, e-mail	Mailing address
U.S. Environmental Protection Agency		
Elizabeth Belk Alternate Task Order Manager	404-562-9236 (phone) 404-562-9224 (fax) decker.ed@epa.gov	USEPA Region 4 61 Forsyth Street, S.W. Mail Code: 9T25 Atlanta, GA 30303-8960
Tim Wool Task Order Manager	404-562-9260 (phone) 404-562-9224 (fax) wool.tim@epa.gov	
William Melville Quality Assurance Coordinator	404-562-9266 (phone) Melville.william@epa.gov	
Tetra Tech, Inc.		
Steven Davie, P.E. Technical Monitor, QC Officer	770-850-0949 (phone) 770-850-0950 (fax) steven.davie@tetratech.com	Tetra Tech, Inc. 2110 Powers Ferry Road, Suite 202 Atlanta, GA 30339
Brian Watson, P.E., P.H. Task Order Leader	770-850-0949 (phone) 770-850-0950 (fax) brian.watson@tetratech.com	
Hugo Rodriguez, Ph.D. Modeling Lead	770-850-0949 (phone) 770-850-0950 (fax) hugo.rodriguez@tetratech.com	
John O'Donnell Quality Assurance Officer	703-385-6000 (phone) 703-385-6007 (fax) susan.lanberg@tetratech.com	Tetra Tech, Inc. 10306 Eaton Place, Suite 340 Fairfax, VA 22030

(This page is intentionally blank.)

1.0 Project/Task Organization

This document presents the quality assurance project plan (QAPP) for developing total maximum daily loads (TMDLs) in Floyd's Fork, Kentucky. U.S. Environmental Protection Agency (EPA) Region 4 has tasked Tetra Tech with support to this TMDL development under Contract No. EP-C-08-004, Task Order (TO) 83. Primary technical support for this effort is being conducted by the Atlanta, Georgia, office of Tetra Tech in conformance with the quality assurance (QA) program described in this QAPP.

The organizational aspects of the program provide the framework for planning and conducting tasks. They can also facilitate project performance and adherence to quality control (QC) procedures and QA requirements. Key project roles are filled by those persons responsible for ensuring the gathering of valid data and the routine assessment of the data for precision and accuracy, as well as the data users and the person(s) responsible for approving and accepting final products and deliverables. The program organization chart, presented as Figure 1, includes relationships and lines of communication among all participants and data users. The responsibilities of these persons are described below.

The EPA Region 4 Task Order Manager (TOM) is Tim Wool, and the alternate TOM is Elizabeth Belk. They will provide overall project and program oversight for the task order. They will review and approve project work and QA plans, as well as interim and final deliverables. The EPA TOM and alternate TOM will also coordinate with contractors, reviewers, and others to ensure technical quality in all deliverables and adherence to the contract, as appropriate.

The EPA Region 4 QA Coordinator (QAC) for this TO is William Melville, and his responsibilities include reviewing and approving the QAPP and participating in any EPA reviews of work performed, as appropriate.

Tetra Tech's Task Order Leader (TOL) is Brian Watson, and the technical monitor is Steven Davie. They will oversee all work conducted under the TO. Specific responsibilities of the Tetra Tech TOL include the following:

- Coordinating project assignments, establishing priorities, and scheduling
- Ensuring completion of high-quality projects within established budgets and time schedules
- Acting as primary point of contact for the EPA TOM and alternate TOM
- Providing guidance, technical advice, and performance evaluations to those assigned to the project
- Implementing corrective actions and providing professional advice to staff
- Preparing or reviewing preparation of project deliverables, including the QAPP and other materials developed to support the project
- Distributing the approved QAPP and any updates to the approved QAPP to staff on the distribution list
- Providing guidance on development of new site-specific models and review of developed models
- Providing support to EPA in interacting with the project team, technical reviewers, and others to ensure that technical quality requirements of the study design objectives are met in accordance with EPA's objectives

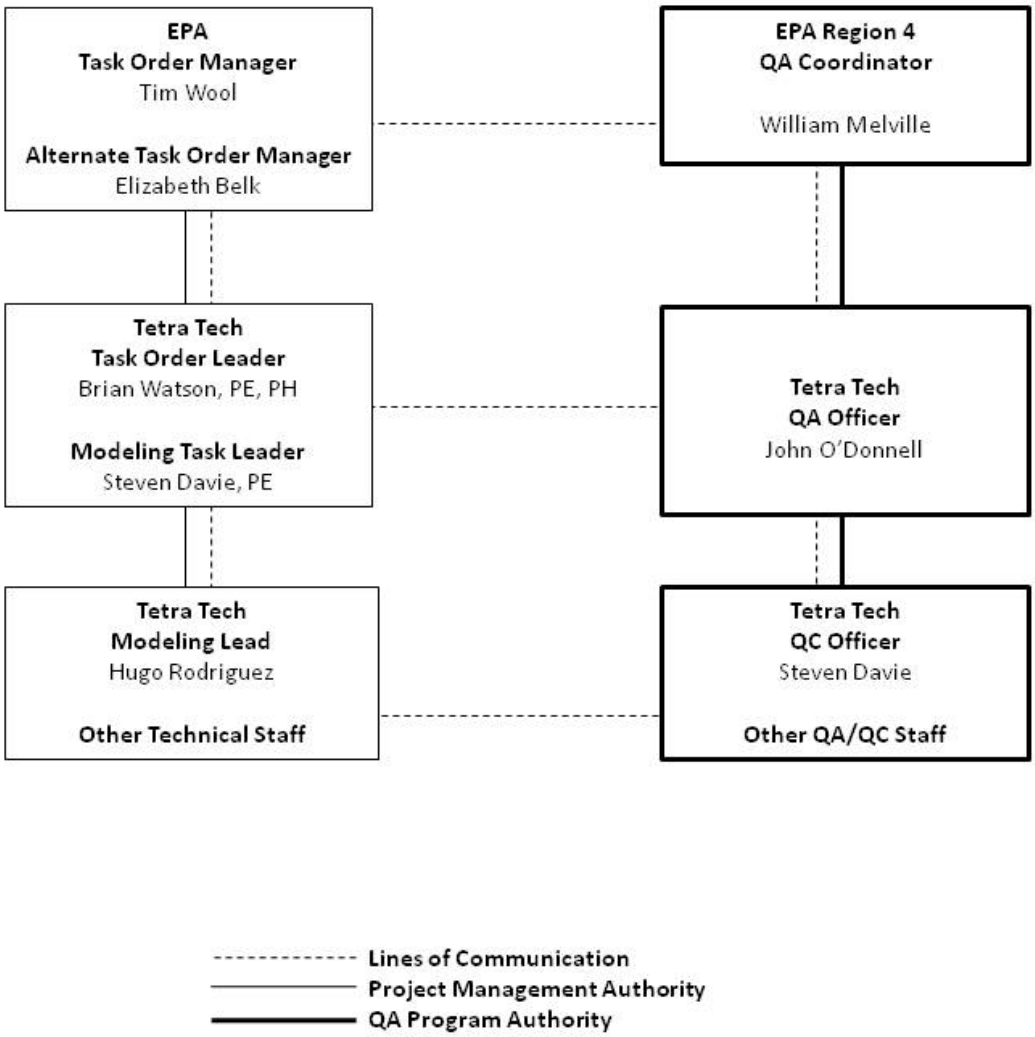


Figure 1 Project Organization

Tetra Tech's QA Officer is John O'Donnell. His primary responsibilities include providing support to the Tetra Tech TOL in the development of the QAPP, reviewing and approving the QAPP, and, with the assistance of the QC Officers, monitoring QC activities to determine conformance with QA/QC requirements.

Tetra Tech's Modeling Lead is Dr. Hugo Rodriguez, and he will assist in the evaluation of currently developed watershed and water quality models identified by the TOM. As appropriate and as directed by the TOM, he will assist the TOL in the oversight and development of additional model input data sets, model application, comparison of model results to observed data, evaluation of model calibration, and document preparation. He will also oversee and supervise the details of any additional modeling efforts and provide guidance on revising and debugging existing, EPA-approved models, as appropriate. Dr. Rodriguez will implement the QA/QC program, and ensure that assigned work is completed on schedule, with strict adherence to the established procedures, and that all procedures and analyses are fully documented, as appropriate. Other technical staff will perform literature searches; assist in secondary data collection, compilation, and QA review; and help complete draft and final modeling reports, as appropriate.

Tetra Tech modeling staff will be responsible for developing model input data sets, calibrating and validating the model, applying the model results, and writing a final report. They will implement the QA/QC program, complete assigned work on schedule and with strict adherence to the established procedures, and complete required documentation.

In his role as the Technical Monitor and QC Officer, Mr. Steven Davie, will provide additional oversight. He is a member of the project staff and is familiar with the models to be used. Mr. Davie will not participate in the application of the models. He will be responsible for performing evaluations to ensure that QC is maintained throughout the data collection and analysis process. QC evaluations will include reviewing site-specific model equations and codes (when necessary), double-checking work as it is completed, and providing written documentation of these reviews to ensure that the standards set forth in the QAPP and in other planning documents are met or exceeded. Other QA/QC staff, including technical reviewers and technical editors selected, as needed, will provide review oversight of the content of the work products and ensure that the work products comply with EPA's specifications.

2.0 Problem Definition/Background

Section 303(d)(1)(c) of the Clean Water Act (CWA) and its associated policy and program requirements for water quality planning, management, and implementation (at Title 40 of the *Code of Federal Regulations* [CFR] Part 130) require the establishment of a TMDL for the achievement of state water quality standards when a waterbody is water quality-limited. A TMDL identifies the pollutant/waterbody-specific assimilative capacity, which includes an appropriate margin of safety. The focus of the TMDL is reduction of pollutant inputs to a level (or "load") that fully supports the designated uses of a given waterbody. The mechanisms used to address water quality problems after the TMDL is developed can include a combination of best management practices or effluent limits and monitoring required through National Pollutant Discharge Elimination System (NPDES) permits.

The Kentucky 303(d) list identifies streams within the state that do not meet their designated use or partially fail to meet their designated uses. EPA Region 4 has contracted with Tetra Tech to assist Kentucky Department of Environmental Protection's, Division of Water (KDOW) in the development of a watershed and riverine model of the Floyds Fork watershed and river to the confluence with the Salt River, and for development of TMDLs for waters within watershed not

meeting prevailing water quality criteria for their attainment of their designated uses. Segments of Floyds Fork fail to meet state water quality criteria for warm water habitat and nutrient-related impairments. In preparation for the modeling and TMDL development effort, EPA funded the U.S. Geological Survey (USGS) to collect nutrient and organic enrichment data to assist KDOW in evaluating the current condition of the watershed. Tetra Tech will develop and calibrate a watershed and riverine model using all available data including those information sources identified in Section 3 of this plan to address nutrient loadings and meeting water quality standards for chlorophyll *a* and dissolved oxygen (DO). Concurrent with the initial modeling phase of the project KDOW will be developing numeric nutrient criteria for total nitrogen and phosphorus (TN and TP) for review by EPA and subsequent use as TMDL load allocation targets.

TMDL development includes running any required model applications; development and calibration of models, if appropriate; and development of draft TMDL and modeling reports and preparation of the associated administrative records.

This QAPP describes the quality system that Tetra Tech will implement to effectively plan throughout this project and provides general descriptions of the work to be performed to support the development of required TMDLs and modeling reports, the standards to be met, and the procedures that will be used to ensure that the results are scientifically valid and defensible and that uncertainty has been reduced to a known and practical minimum. This project does not require the collection of primary data. During the initial data collection and inventory operations, it might be determined that primary data should be collected to better characterize the waterbodies. Should additional data collection be required in consultation with the TOM and stakeholders, Tetra Tech will, under the TOM's technical direction or in response to an amendment to this TO, develop a separate field sampling QAPP or QAPPs to detail the locations and frequency of sampling, parameters to be measured, and the data quality objectives to be met, and to document the rationale for the monitoring plan and the procedures to be used to collect those data.

3.0 Project/Task Description

Task 1. Quality Assurance Project Plan

This QAPP has been developed in accordance with the requirements in to guidance provided in *EPA Requirements for Quality Assurance Project Plans.* (EPA QA/R-5, EPA/240/B-01/003, U.S. Environmental Protection Agency, Office of Environmental Information, Washington DC, March 2001 [Reissued May 2006]) (USEPA 2001); EPA's *Guidance for Quality Assurance Project Plans for Modeling.* (U.S. EPA QA/G-5M, EPA/240/R-02/007, U.S. Environmental Protection Agency, Office of Environmental Information, Washington, DC, December 2002) (USEPA 2002a) and *Guidance on the Development, Evaluation, and Application of Environmental Models* (EPA/100/K-09/003, U.S. Environmental Protection Agency, Office of the Science Advisor, Council for Regulatory Environmental Modeling [CREM], Washington DC, March 2009) (USEPA 2009b); and on the basic requirements of the *EPA Office of Water Quality Management Plan*. (EPA/821/R-09/001, U.S. Environmental Protection Agency, Office of Water, Washington DC, February 2009). The draft QAPP has been developed for EPA review and approval and the final plan will be developed in consideration of EPA comments and requested revisions for implementation by the project team.

Task 2. Floyds Fork Watershed Model

Tetra Tech will develop a watershed model for the entire Floyds Fork watershed to the confluence with the Salt River. The watershed model will be calibrated for: flow, nutrients (nitrogen and phosphorus), and biochemical oxygen demand. Outputs for temperature, dissolved oxygen and chlorophyll-a will also be evaluated. Tetra Tech will utilize the Loading Simulation Program C++ (LSPC) as the watershed model. The simulation period will be determined at the first meeting with EPA and KDOW, but is expected to be from January 1, 2000 through at least December 31, 2010.

The watershed model will be divided into modeling sub-basins based on hydrologic criteria to be represented as a series of hydraulically connected sub-watersheds in which the watershed model will calculate surface water runoff and the advective transport of constituents using historic precipitation data. The following data and other modeling requirements will be required to perform the continuous watershed model simulations:

- Meteorological Data: Hourly meteorological data from weather stations within, or in close proximity to, the sub-watershed will be used in the watershed model. Precipitation data for the watershed will be gathered from several sources and the watershed will be subdivided into Thiessen polygons with precipitation stations as centers, in order to select the station for the watershed. The potential evapotranspiration will be calculated from the maximum and minimum daily temperatures obtained from either NCDC or other local network stations. The Hamon PET method will be used to calculate hourly potential evapotranspiration using air temperature, a monthly variable coefficient, the number of hours of sunshine (based on latitude), and absolute humidity (computed from air temperature).

- Land Use/Land Cover: The watershed model uses land cover data as the basis for representing hydrology and nonpoint source loading. Tetra Tech will discuss with EPA and KDOW what source of land use/land cover that shall be utilized for this modeling effort. If KDOW does not have any local data, then the 2011 NLCD land use coverage will be used. Land cover categories for modeling are expected to include open water, urban, barren or mining, cropland, pasture, forest, and grassland. Coverages of imperviousness will also be utilized to develop the typical imperviousness percentages for each land use category. The percent imperviousness of a given land category will be calculated as an area-weighted average of land use classes encompassing the modeling land category.

- Soils Data: Soils data for the watershed will be obtained from the STATSGO or SSURGO Database. There are four main hydrologic soil groups. The different soil groups range from soils that have a low runoff potential to soils that have a high runoff potential. The total area that each hydrologic soil group covers within each sub-watershed will be determined. The hydrologic soil group that has the highest percent of coverage within each sub-watershed will be used to represent the sub-watershed.

- Digital Elevation Model: Digital elevation model (DEM) data will be obtained for the watersheds modeled and shall have at least a 10-meter grid resolution. These data will be used to determine the channel and watershed slopes for use in the watershed model.

- Point Source Discharge Data: The watershed model will include point source discharge data. This data will be obtained from either KDOW or EPA's PCS database. At the first meeting with KDOW and EPA, it will be discussed if there should be a flow cutoff limit for point sources. For example, should point sources with a permitted flow below 0.1 MGD be excluded from the model. In addition, it will be determined to what level of detail that will be included for each point source. This will determine whether monthly DMR data or daily data will be needed for each point source.

- Water Withdrawal Data: The watershed model will include water withdrawal data. This data will be obtained from either KDOW or EPA's PCS database. Similar to the point source discussion, at the first meeting with KDOW and EPA, it will be discussed if there should be a flow cutoff limit for water withdrawals.

- Septic Systems and Land Application System Data: The watershed model will include septic systems and any land application systems data. This data will be obtained from either KDOW or county health departments..

The watershed model will be calibrated to daily flows and discrete instream water quality data measured by KDOW, USGS, local municipalities, counties, and other sources of data identified by KDOW. The watershed models will simulate the rainfall runoff process for both flow and water quality and the results of these models will be used as tributary inputs to the river model.

This task will also include the following:

- Identify existing data and informational sources, including where and how to retrieve these data and information.
  - USGS Flows
  - USGS Water Quality Data (Collected under IAG for EPA)
  - NPDES Data (Kentucky Department of Water, EPA)
- Data will be processed into the Water Resources Database (WRDB), this should include all of the observed data (meteorological, water quality, flow, environmental), watershed predictions for flow, nutrients, and BOD.
- Develop model and calibrate to all available data. The watershed will be delineated to allow for the investigated of each individual tributary to Floyds Fork.
- Develop a watershed modeling report (Draft & Final) detailing sources of data, assumptions made in the calibration, calibration time series graphs and any sensitivity analysis that is done.
- Deliver Calibrated Model (Models and Input Files)
- Provide onsite training to KDOW modeling staff in the use of the watershed model. This will be scheduled at the end of the project and will coincide with the river water quality model.

LSPC

LSPC represents receiving waters as one-dimensional, completely mixed stream reaches or reservoirs. It is a dynamic watershed model driven by time-variable weather input data. It produces

time series of flow and pollutants simulating transport in overland flow, the vadose and saturated zones, and in-stream components of the system, using an area-weighted or lumped methodology. LSPC can simulate loadings from multiple land uses and represent in-stream processes that affect the fate of nutrients within the stream network. Model documentation is available from EPA (http://www.epa.gov/athens/wwqtsc/).

Input data for LSPC includes three main categories of information: (1) landscape data, including topography, point source locations, locations and connections among streams, and the like; (2) meteorological data, including precipitation, air temperature, humidity; and (3) land use and pollutant-specific data (land use areas, monitoring data, and such). The watershed loading component of the model divides all land uses into pervious and impervious segments, which are further grouped by land use and subbasin. Loads from subbasins are routed to receiving waters (representative stream reaches or reservoirs).

LSPC can evaluate both point and nonpoint sources and simulate both land-based (e.g., rainfall-runoff) and in-stream water movement and processes. Nitrogen/phosphorus pollution loading can originate from a wide variety of nonpoint sources, such as agriculture, on-site wastewater treatment systems (e.g., septic systems), urban runoff, timber production areas, and atmospheric deposition. Only runoff processes are simulated for impervious land units, whereas interflow and baseflow are simulated in addition to runoff for pervious land units.

For each pervious and impervious land segment, LSPC will estimate the water budget, sediment transport, and general water quality constituents, which will represent total nitrogen (TN), total phosphorus (TP) and biochemical oxygen demand (BOD). In each waterbody segment, LSPC will simulate hydraulic behavior, water temperature, transport of TN, TP, and sediments, and transport of BOD. LSPC simulates nitrogen/phosphorus pollution loading from watersheds using a buildup/washoff approach wherein nutrients accumulate between rain events and are mobilized and transported into streams during rain events. When using that method, nitrogen, phosphorus and other constituents can be applied to the land surface over time so that a mass of the pollutant accumulates and is then removed at a rate correlated to a corresponding quantity of sheet flow on the land surface.

Task 3. Development of Floyds Fork Water Quality Model

Tetra Tech will develop a water quality model for the Floyds Fork River to address nutrients loadings and meeting a water quality standard for chlorophyll a and dissolved oxygen. Tetra Tech will utilize the Water Quality Analysis Simulation Program (WASP 7.x) as the water quality model. The simulation period will be determined at the first meeting with EPA and KDOW, and will coincide with the watershed model simulation period.

This task will include:

- Identify existing data and informational sources, including where and how to retrieve these data and information.
- These data will be put into a form that will be easily transmitted to KDOW (WRDB)
- Develop model and calibrate to all available data
- Develop a water quality modeling report detailing sources of data, assumptions made in the calibration, calibration time series graphs and any sensitivity analysis that is done.
- Develop load reduction scenarios in consultation with the TOM to meet water quality criterion.

- Provide onsite training to KDOW modeling staff in the use of the water quality model. This training will be coordinated with the training of the watershed model and will be conducted at the end of the task.

WASP

WASP is an EPA-developed and supported water quality model that is routinely applied throughout the United States and worldwide to investigate water quality issues. WASP7 is the newest version of WASP (released June 7, 2010) and has many upgrades to the user interface and to the model capabilities. WASP can be downloaded at http://www.epa.gov/athens/wwqtsc/html/wasp.html.

The LSPC will be linked to the WASP model by providing flows and loads at tributaries and local drainage areas. Three organic carbon variables play an equivalent role as BOD, representing organic matter that is relatively refractive, of an intermediate reactivity, or labile. Nitrogen is divided into organic and inorganic fractions. Organic nitrogen state variables are dissolved organic nitrogen, labile particulate organic nitrogen, and refractory particulate organic nitrogen. Inorganic nitrogen forms are ammonia (NH3) and nitrate (nitrite is implicitly represented). Both nitrate (NO3-) and ammonia (NH3) are used to satisfy algal nitrogen requirements, with ammonia (NH3) being preferred. The primary reason for distinguishing the two is that ammonia is oxidized by nitrifying bacteria into nitrate. Nitrification can be a significant sink of oxygen in the water column. Sediment nitrification is represented implicitly via the boundary condition value for sediment oxygen demand (SOD). As with carbon and nitrogen, organic phosphorus is considered in three states: dissolved organic phosphorus, labile particulate phosphorus, and refractory particulate phosphorus. Only a single inorganic form, orthophosphate, is considered. Orthophosphate exists as several states in the model ecosystem: dissolved phosphate, phosphate adsorbed to inorganic solids, and phosphate incorporated in algal cells. Equilibrium partition coefficients are used to distribute the total among the three states.

WASP is capable of simulating four classes of algae, each targeting a specific ecological "niche" defined by distinctive characteristics of the class and the role those characteristics play in ecosystem function. Cyanobacteria, commonly called blue-green algae, are characterized by their abundance as picoplankton in saline water and by their bloom-forming characteristics in fresh water. Being very small, picoplankton do not sink at appreciable rates. Another key feature of cyanobacteria is that some species, generally in lakes, fix atmospheric nitrogen and can form harmful blooms. Diatoms are distinguished by their need for silica to form their siliceous tests. Diatoms can grow quickly if given sufficient nutrients but also sink relatively quickly, being large cells and lacking flagella to actively maintain position in the water column. Planktonic algae that do not fall into the two groups are lumped into the heading of green algae. Green algae settle at a rate intermediate between cyanobacteria and diatoms and are subject to greater grazing pressure than are cyanobacteria. The fourth category is macroalgae, which could be expected to increase in biomass when nutrients are available but high flushing rates prevent accumulations of planktonic algae.

WASP is able to simulate sediment-water oxygen and nutrient exchanges by simulating sediment processes using a sediment diagenesis model. Such an approach entails substantial data requirements, and a need for adequate data to calibrate the model. The approach that Tetra Tech is considering using for Floyds Fork is to specify those oxygen and nutrients fluxes as boundary conditions. That could improve model performance relative to a poorly specified and calibrated simulation of the processes. A limitation of the approach is that if changes in benthic processes such as SOD or nutrient fluxes are an important aspect of the response of the ecosystem to changes in nutrient loading, the model will not be able to reflect those responses. For example, if estimates of SOD are

based on recent observations (e.g., Murrell et al. 2009) and SOD would decrease if nitrogen/phosphorus pollution loading rates decreased, the model would tend to underestimate the improvement in bottom water oxygen condition that could be associated with reduced nitrogen/phosphorus pollution loading. EPA could address that by modifying the boundary condition used in pollution reduction scenarios.

Task 4. Floyds Fork TMDL Development

Tetra Tech will develop a TMDL document for the Floyds Fork watershed. The TMDL will have initial allocations for all point sources (continuous point and MS4) and potential reductions in nonpoint sources. The TMDL reduction scenarios will be done in consultation with the TOM and reviewed by KDOW personnel. The draft TMDL will be submitted in electronic format with the accompanying watershed and water quality modeling reports.

Task 5. Stakeholder Meetings

Tetra Tech will conduct and facilitate at least 6 public outreach meetings in the Louisville, KY area. The purpose of these meetings are to bring all interested parties (Municipal Sewer District - Louisville, Environmental Interest Groups, EPA and KDOW) together to form a consensus on the modeling and TMDL development. EPA and KDOW will provide a list of potential invitees. The first meeting will be used to introduce the technical approach that will be used in model development and TMDL determination. This meeting will also be used to solicit additional information and approaches that could be used in the TMDL determination. Other stakeholder meetings will be held at the completion of the watershed modeling, water quality modeling, the draft of the TMDL, and the finalization of the TMDL.

Task 6. Training

Tetra Tech will provide a one week onsite training that will cover the development and use of the watershed and water quality modeling. The training will take place at KDOW's offices. Training will include:

- Overview of Watershed and Water Quality Model
- Delineation of Watershed and Water Quality Model
- Overview of all forcing functions for the Watershed and Water Quality Model
- Calibrations process
- TMDL Determination and overview of reduction scenarios.

Schedule

The following schedule outlines key deliverables dates:

Table 1 Project Schedule

Task	Item	Deliverable	Draft Date	Final Date
Task 1	Quality Assurance Plan	Signed QAPP Document	07/15/2011	07/22/2011
Task 2	Draft Watershed Modeling Report	MS Word Document	12/30/2011	04/01/2012
Task 2	Final Watershed Modeling Report	MS Word Document	03/01/2012	04/01/2012
Task 3	Draft Water Quality Modeling Report	MS Word Document	05/01/2012	07/18/2012
Task 3	Final Water Quality Modeling Report	MS Word Document	06/01/2012	07/18/2012

Task 4	Draft TMDL Document	MS Word Document	8/17/2012	09/16/2012
Task 4	Final TMDL Document	MS Word Document and Final Models	8/17/2012	09/16/2012
Task 6	Training	Training Materials	TBD	11/15/2012

4.0 Quality Objectives and Criteria

Data quality objectives (DQOs) are qualitative and quantitative statements that are used in the project planning and implementation to clarify the intended use of the data, define the type of data needed to support the decision, identify the conditions under which the data should be collected, and specify tolerable limits on the probability of making a decision error because of uncertainty in the data (if applicable). Data users develop DQOs to specify the data quality needed to support specific decisions.

Data of known and documented quality are essential to the success of any water quality modeling study, which in turn generates data for use in various evaluations and to make decisions. Model setup, calibration, and validation for the project under this QAPP will be accomplished using currently available data. The QA process for this study consists of using data of acceptable quality, data analysis procedures, modeling methodology and technology, administrative procedures, and auditing. Project quality objectives and criteria for measurement data will be addressed in the context of the two tasks discussed above: (1) evaluating the quality of the data used, and (2) assessing the results of the model application.

The quality of an environmental monitoring program can be evaluated in three steps: (1) establishing scientific assessment quality objectives, (2) evaluating program design for whether the objectives can be met, and (3) establishing assessment and measurement quality objectives that can be used to evaluate the appropriateness of the methods being used in the program. The quality of a data set is some measure of the types and amount of error associated with the data.

Sources of error or uncertainty in statistical inference are commonly grouped into two categories:

- *Sampling error:* The difference between sample values and in situ *true* values from unknown biases due to sampling design. Sampling error includes natural variability (spatial heterogeneity and temporal variability in population abundance and distribution) not specifically accounted for in a design (for design-based inference), and variability associated with model parameters or incorrect model specification (for model-based inference).
- *Measurement error:* The difference between sample values and in situ *true* values associated with the measurement process. Measurement error includes bias and imprecision associated with sampling methodology; specification of the sampling unit; sample handling, storage, preservation, and identification; and instrumentation.

Sections 4.1 through 4.7 describe DQOs and criteria for model inputs and outputs for this project, written in accordance with the seven steps described in EPA's *Guidance on Systematic Planning Using the Data Quality Objectives Process* (EPA QA/G-4) (USEPA 2006).

4.1 State the Problem

Excess inputs of nutrients (nitrogen/phosphorus pollution) and pathogens in surface waters can be harmful in aquatic ecosystems by directly producing excess plant and algal growth, and indirectly leading to reduced clarity, reduced oxygen levels as the algae and plants decompose, and decreased

biodiversity. Primary sources of nutrient and pathogen loading to aquatic ecosystems may include illegal dumps or other inappropriate waste disposal; urban runoff/storm sewers; package plant or other permitted small flow discharges; municipal point source discharges; highway/road/bridge runoff (non-construction related); grazing in riparian or shoreline zones; agriculture; and site clearance (land development or redevelopment; construction related).

4.2 Identify the Study Question

The study's objective is to support development of TMDLs and associated waste load allocations which will meet prevailing water quality standards for chlorophyll *a* and DO, and potentially those numeric criteria for TP and TN under development by KDOW which will result in attainment of designated uses.

4.3 Identify Information Needs

The objective of this modeling project is to determine the allowable loads of nutrients that the waters within the Floyds Fork watershed can receive while still attaining water quality standards (WQS) for chlorophyll a and DO, and possibly nutrients, pending the final development and approval of KDOW's numeric criteria. Wasteload allocations for point sources and load allocations for nonpoint sources need to be developed to support attainment of standards and designated uses.

The load allocations will be used to develop point and nonpoint source reduction plans on the basis of meeting relevant ambient water quality criteria. In general, ambient water quality criteria have incorporated a margin of safety such that concentrations at or just less than the criterion indicate a potential for unacceptable risks to aquatic life; exceedances are anticipated to produce impairment.

Input data for LSPC includes three main categories of information: (1) landscape data, including topography, point source locations, locations and connection among streams; (2) meteorological data, including precipitation, air temperature, and humidity; and (3) land use and pollutant-specific data (e.g., land use areas, monitoring data). The watershed loading component of the model divides all land uses into pervious and impervious segments, which are further grouped by land use and subbasin. Loads from subbasins are routed to receiving waters (representative stream reaches or reservoirs).

Tetra Tech is considering using the following model inputs to support development of numeric nutrient criteria for Florida estuarine waters:

- USGS NHD+ catchments
- U.S. Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS) 12-digit HUCs
- Subwatershed information (e.g., elevations, slopes, reach lengths)
    - National Elevation Dataset (NED) 1/3 arc-second (10 meter by 10 meter)
- Land use
    - National Land Use/Land Cover Dataset (2011).
- Meteorological data
    - NOAA NCDC rainfall data
    - Kentucky State Climatological Office
- NPDES-permitted point sources and water withdrawals from KDOW

4.4 Specify the Characteristics that Define the Population of Interest

The tasks in this project must support the goal of quantifying the amount of nutrient and sediment loading that Floyds Fork watershed can assimilate while attaining relevant water quality standards. The statistical criteria for the designations/allocations are detailed with the error discussion in Section 4.6.

In most cases, the statistical criteria for loads and concentrations are detailed in the error discussion in Section 4.6.

4.5 Develop the Strategy for Information Synthesis

Tetra Tech will use a systematic planning process to apply LSPC and WASP models for developing the watershed model and the subsequent TMDLs in Floyds Fork watershed. That process takes into account the following elements:

- The accuracy and precision needed for the models to predict a given quantity at the application site of interest to satisfy regulatory objectives
- The appropriate criteria for making a determination of whether the models are accurate and precise based on past general experience combined with site-specific knowledge and completeness of the conceptual models
- How the appropriate criteria would be used to determine whether model outputs achieve the needed quality

Acceptance criteria that result from systematic planning address the following types of components for modeling projects. Criteria used in selecting the appropriate model will be documented in the modeling reports and typically include the following:

- Technical criteria (concerning the requirements for the model's simulation of the physical system)
- Regulatory criteria (concerning constraints imposed by regulations, such as water quality standards)
- User criteria (concerning operational or economical constraints, such as hardware/software compatibility)

The Tetra Tech TOL compared available models to select the most ones to use for this study. In addition, existing model programming language can be converted into a different programming language to enhance software compatibility. The models that will be used are

- LSPC
- WASP
- WRDB

4.6 Specify Performance and Acceptance Criteria

Quantitative measures, sometimes referred to as calibration criteria, include the *relative error* between model predictions and observations as defined below.

$$E_{rel} = \frac{\Sigma |O - P|}{\Sigma O} \times 100$$

where E_{rel}= relative error in percent. The relative error is the ratio of the absolute mean error to the mean of the observations and is expressed as a percent.

Models will be deemed acceptable when they are able to simulate field data within predetermined statistical measures. Examples of these statistical measures are described in Table 2. Future variables can be added to a future appendix if deemed necessary. Those statistical criteria will vary depending on the focus of the model study. When applying watershed hydrologic models, for example, Tetra Tech will use a hydrologic calibration spreadsheet to determine the acceptability of modeling results. The spreadsheet computes the relative error for various aspects of the hydrologic system. Statistical targets that have been developed and implemented in previous studies (Lumb et al. 1994) are defined and met for each aspect of the system before accepting the model (**Error! Reference source not found.** 3). Similar comparisons are made for other modeling components (e.g., watershed pollutant loads and receiving water quality).

Table 2 Statistical Measures for Model Comparisons (USEPA 2000)

State Variable	Percent Difference between Simulated and Observed Values		
	Excellent	Good	Fair
Water Temperature	<7	8-12	13-18
Water Quality / Dissolved Oxygen	<15	15-25	25-35
Nutrients / Chlorophyll a	<30	30-45	45-60

Table 3 Relative errors and statistical targets for hydrologic calibration

Relative errors (simulated-observed)	Statistical target
Error in total volume	±10%
Error in 50% lowest flows:	±10%
Error in 10% highest flows:	±15%
Seasonal volume error - Summer:	±30%
Seasonal volume error - Fall:	±30%
Seasonal volume error - Winter:	±30%
Seasonal volume error - Spring:	±30%
Error in storm volumes:	±20%
Error in summer storm volumes:	±50%

It is important to clarify that these targets are intended to be applied to mean values, and that individual events or observations may show larger differences and still be acceptable.

Calibration for water quality will proceed sequentially after hydrologic calibration. Unlike flow, water quality parameters are not observed continuously. The calibration must therefore rely on comparison of continuous model output to point-in-time-and-space observations. This creates a situation in which it is not possible to fully separate error in the model from variability inherent in the observations. For example, a model could provide an accurate representation of an event mean or

daily average concentration in a reach, but an individual observation at one time and one point in a reach itself may differ significantly from the average. In addition, any uncertainty present in the hydrologic calibration will also propagate into the water quality simulation.

For these reasons, specific acceptance criteria are not pre-specified for the water quality simulation. Instead, attempts will be made to minimize relative error and bias. For model uses to establish allocations it is particularly important to control bias in the simulation. Bias relative to flow regime will be evaluated through the use of rating curve regressions and logistic regressions on the load-duration curves.

4.7 Optimize the Design for Obtaining and Generating Adequate Data or Information

The data requirements of this project encompass aspects of both laboratory analytical results obtained as secondary data and database management to reduce sources of errors and uncertainty in the use of the data. Data commonly required for populating a database to supply data for calibrating a model are listed in

DATA TYPE	EXAMPLE MEASUREMENT ENDPOINT(S) OR UNITS
Geographic or Location Information (Typically in GIS Format)	
Land use	acres
Soils (including soil characteristics)	hydrologic group
Topography (stream networks, watershed boundaries, contours, or digital elevation)	elevation in feet and meters; percent slope
Water quality and biological monitoring station locations	latitude and longitude, decimal degrees (North American Datum 1983; NAD83)
Meteorological station locations	latitude and longitude, decimal degrees
Permitted facility discharge locations	latitude and longitude, decimal degrees
Impaired waterbodies (georeferenced 303(d)-listed AUs)	latitude and longitude, decimal degrees
Dam locations	latitude and longitude, decimal degrees
Municipal separate storm sewer discharge locations	latitude and longitude, decimal degrees
Mining locations	latitude and longitude, decimal degrees
Flow	
Historical record (daily, hourly, 15-minute interval)	cubic feet per second (cfs)
Dam release flow records	cfs
Peak flows	cfs
Meteorological Data	
Rainfall	inches
Temperature	degrees C
Wind speed	miles per hour
Dew point	degrees C
Humidity	percent or grams per cubic meter
Potential evapotranspiration	Inches
Cloud cover	percent
Solar radiation	watts per square meter
Water Quality (Surface Water, Ground Water)	
Chemical monitoring data	milligrams per liter (mg/L)
Field parameters (turbidity, pH, water temperature)	Nephelometric turbidity units (NTU), hydrogen ion activity (pH), temperature (degrees C).
Discharge Monitoring Report	discharge characteristics (flow and chemical loading)

DATA TYPE	EXAMPLE MEASUREMENT ENDPOINT(S) OR UNITS
Permit Limits	mg/L
Regulatory or Policy Information	
Applicable state water quality standards	mg/L
EPA water quality standards	mg/L
On-site Waste Disposal	
Septic systems	number of systems, locations, failure rates
Illicit discharges	straight pipes
Land Management Information	
Agricultural practices (major crops, rotation; manure management practices, fertilization practices, pesticide use)	description of crop rotations; pounds manure applied per acre
Best Management Practices	length and width of buffer strips
Additional Anecdotal Information as Appropriate	
Stream networks, watershed boundaries, contours or digital elevation, storm water permits, storm characteristics, reservoir characteristics, facility type, permit status, applicable permits, best management practices, major crops, crop rotation, manure management and application practices, livestock population estimates, fertilization application practices, pesticide use, wildlife population estimates, citizen complaints, relevant reports, existing watershed and receiving water models	specific descriptive codes

.

Secondary data will be downloaded electronically from various sources to reduce manual data entry whenever possible. Secondary data will be organized into a standard model application database. A screening process will be used to scan through the database and flag data that are outside typical ranges for a given parameter; values outside typical ranges will not be used to develop model calibration data sets or model kinetic parameters. The data used in the model, the time period from which the data were collected, and the quality requirements of the data will be described in the TMDL report. If no quality requirements exist or if the quality of the secondary data cannot be determined, a disclaimer that indicates that the quality of the secondary data is unknown will be added. The wording of this disclaimer will be as follows:

The quality of the secondary data used in developing the TMDL could not be determined.

The goal of the modeling effort is to calculate water or sediment contaminant levels resulting from one or more point and nonpoint sources. The results of the modeling effort could be used to establish NPDES permit limits or nonpoint source reduction plans on the basis of meeting relevant ambient water or sediment quality criteria. In general, ambient water and sediment quality criteria have incorporated a margin of safety such that concentrations at or just less than the criterion indicate a potential for unacceptable risks to human health or aquatic life, and exceedances are anticipated to produce impairment. If the calculated point source permit limit for the contaminant is exceeded, water or sediment quality will be reduced possibly presenting a hazard.

Uncertainty in the data due to sampling and measurement errors or errors introduced during data manipulation could result in identifying a hazard when one does not actually exist or in not identifying a hazard when one does exist. The overall assumption being made during this process is that the results of the assessment should be conservative, i.e., errors made by identifying a hazard

when one does not actually exist are more acceptable than errors made by not identifying a hazard when one does exist. Reducing data uncertainty is of the highest priority. Because these data will be used to develop control measures, including NPDES permits and actions taken by state, territorial,

Table 4 Secondary Environmental Data to Be Collected for the Floyds Fork TMDL Development

DATA TYPE	EXAMPLE MEASUREMENT ENDPOINT(S) OR UNITS
Geographic or Location Information (Typically in GIS Format)	
Land use	acres
Soils (including soil characteristics)	hydrologic group
Topography (stream networks, watershed boundaries, contours, or digital elevation)	elevation in feet and meters; percent slope
Water quality and biological monitoring station locations	latitude and longitude, decimal degrees (North American Datum 1983; NAD83)
Meteorological station locations	latitude and longitude, decimal degrees
Permitted facility discharge locations	latitude and longitude, decimal degrees
Impaired waterbodies (georeferenced 303(d)-listed AUs)	latitude and longitude, decimal degrees
Dam locations	latitude and longitude, decimal degrees
Municipal separate storm sewer discharge locations	latitude and longitude, decimal degrees
Mining locations	latitude and longitude, decimal degrees
Flow	
Historical record (daily, hourly, 15-minute interval)	cubic feet per second (cfs)
Dam release flow records	cfs
Peak flows	cfs
Meteorological Data	
Rainfall	inches
Temperature	degrees C
Wind speed	miles per hour
Dew point	degrees C
Humidity	percent or grams per cubic meter
Potential evapotranspiration	Inches
Cloud cover	percent
Solar radiation	watts per square meter
Water Quality (Surface Water, Ground Water)	
Chemical monitoring data	milligrams per liter (mg/L)
Field parameters (turbidity, pH, water temperature)	Nephelometric turbidity units (NTU), hydrogen ion activity (pH), temperature (degrees C).
Discharge Monitoring Report	discharge characteristics (flow and chemical loading)
Permit Limits	mg/L
Regulatory or Policy Information	
Applicable state water quality standards	mg/L
EPA water quality standards	mg/L
On-site Waste Disposal	
Septic systems	number of systems, locations, failure rates
Illicit discharges	straight pipes
Land Management Information	
Agricultural practices (major crops, rotation; manure management practices, fertilization practices, pesticide use)	description of crop rotations; pounds manure applied per acre
Best Management Practices	length and width of buffer strips
Additional Anecdotal Information as Appropriate	

DATA TYPE	EXAMPLE MEASUREMENT ENDPOINT(S) OR UNITS
Stream networks, watershed boundaries, contours or digital elevation, storm water permits, storm characteristics, reservoir characteristics, facility type, permit status, applicable permits, best management practices, major crops, crop rotation, manure management and application practices, livestock population estimates, fertilization application practices, pesticide use, wildlife population estimates, citizen complaints, relevant reports, existing watershed and receiving water models	specific descriptive codes

tribal, or local authorities, to implement TMDLs to reduce pollution, it is important to reduce uncertainty by using appropriate QC protocols. Discussion of conventional data quality indicators (precision, accuracy, representativeness, completeness, and comparability) appear in the Appendix.

Tetra Tech documents all data sources, including full reference citations in a bibliography and parenthetical references in report text. Tetra Tech also maintains paper and electronic copies of all references. Documentation for all data sources (i.e., full bibliographical information and metadata where appropriate) will be collected and recorded.

5.0 Special Training Requirements/Certification

Tetra Tech staff involved in developing TMDLs, model input data sets and model applications have experience in all phases of TMDL development and support. Staff from Tetra Tech's water resources practice have developed more than 6,000 approved TMDLs for states, territories, local agencies, and tribes for all water body types and numerous pollutants, including more than 600 completed in EPA Region 4. Much of this work has focused on developing and applying tools and models in TMDL and management decision-making, and training EPA and state staff to develop TMDLs and to appropriately select and use models.

6.0 Documentation and Records

Thorough documentation of all modeling activities is necessary for the interpretation of study results. As directed by the EPA TOM, Tetra Tech will prepare progress reports and other deliverables, which will be distributed to project participants as indicated by the EPA TOM. Data and assumptions used to develop the TMDL models will be recorded and documented in the TMDL report.

The format of the raw data to be used for TMDL model parameters, model input, model calibration, and model output will be converted to the appropriate units, as necessary, for use in TMDL development.

An administrative record will be developed concurrently with TMDL development. EPA requires that all information used in TMDL calculations for a specific project is included in the administrative record for that project. The administrative record is a complete repository of all the data and information used to develop a TMDL. The administrative record will serve as the project files for TMDL development and will contain copies of all records and documents, including electronic versions of the data and model input data sets. The administrative record will be delivered to EPA at the end of the TO. A copy of the administrative record will be

maintained at the Tetra Tech, Atlanta, GA, office for at least 3 years (unless otherwise directed by EPA's TOM). The TOM and TOL will maintain files, as appropriate, as repositories for information and data used in models and for the preparation of any reports and documents during the project. Electronic project files are maintained on network computers and are backed up periodically. The Tetra Tech TOL will supervise the use of materials in the administrative record. The following information will be included in the hard copy or electronic project files in the administrative record:

- Any reports and documents prepared.
- Contract and TO information.
- Electronic copies of model input/output (for model calibration and allocation scenarios).
- Results of technical reviews, model tests, data quality assessments of output data, and audits.
- Documentation of response actions during the project to correct model development or implementation problems.
- Assessment reports for acquired data.
- Statistical goodness of fit methods and other rationale used to decide which statistical distributions should be used to characterize the uncertainty or variability of model input parameters.
- Communications (electronic mail, memoranda, internal notes, telephone conversation records, letters, meeting minutes, and all written correspondence among the project team personnel, subcontractors, suppliers, or others).
- Maps, photographs, and drawings.
- Studies, reports, documents, and newspaper articles pertaining to the project.
- Spreadsheet data files: physical measurements, analytical chemistry data, and microbiological data (hard copy and on diskette).

The model application will include complete recordkeeping of each step of the modeling process. The documentation will consist of reports and files addressing the following items:

- Assumptions.
- Parameter values and sources.
- Nature of grid, network design, or subwatershed delineation.
- Changes and verification of changes made in code.
- Actual input used.
- Output of model runs and interpretation.
- Calibration and validation of the model(s).

Formal reports submitted to EPA that are generated from the data will be maintained in the central file (electronic and hard copy) at Tetra Tech's Atlanta office. The data reports will include a summary of the types of data used, sampling dates, and any problems or anomalies observed during data analysis and review.

7.0 Sampling Process Design

Not applicable.

8.0 Sampling Methods

Not applicable.

9.0 Sample Handling and Custody

Not applicable.

10.0 Analytical Methods

Not applicable.

11.0 Quality Control

The project team will follow the policies and procedures detailed in this QAPP. In general, training programs, materials, manuals, and reports prepared by Tetra Tech staff will be subjected to internal or external technical and editorial reviews before the final versions are submitted. Specific QC procedures for the development, application, and calibration of the models used for this project are described in this section.

The data quality of model input and output is addressed, in part, by the training and experience of project staff (Section 5.0) and documentation of project activities (Section 6.0). This QAPP and other supporting materials will be distributed to all personnel involved in nutrient criteria model development. The QC Officer will ensure that all tasks described in the project action plan for developing the required analyses and TMDLs are carried out in accordance with the QAPP. Staff performance will be reviewed throughout each of the TMDL development phases to ensure adherence to project protocols.

QC is defined as the process by which QA is implemented in a modeling project. All project modelers will conform to the following guidelines:

- All modeling activities including data interpretation, load calculations, or other related computational activities are subject to audit or internal review. Thus, the modelers are instructed to maintain careful written and electronic records for all aspects of model development.
- A written record of where the data used in the models were obtained will be kept, and any information on data quality will be documented in the final report. A written record on where this information is located on a computer or backup media will be maintained in the project files.

The QC Officer or their designee will periodically conduct surveillance of each modeler's work. Modelers will be asked to provide verbal status reports of their work at periodic modeling subgroup teleconferences. Detailed modeling documentation will be made available to members of the modeling subgroup as necessary.

The ability of computer code to represent model theory accurately will be ensured by following rigorous programming protocols, including documentation within the source code. Specific tests will be required of all model revisions to ensure that fundamental operations are verified to the extent possible. Those tests include testing numerical stability and convergence properties of the model code algorithms, if appropriate. Model results will be generally checked by comparing results to those obtained by other models or by comparison to hand calculations. Visualization of model results

will assist in determining whether model simulations are realistic. Model calculations will be compared to field data. If adjustments to model parameters are made to obtain a *fit* to the data, the modelers will provide an explanation and justification that must agree with scientific knowledge and with process rates within reasonable ranges as found in the literature.

Both project-generated and non-project-generated data will be used for model development and calibration. The QA procedures for project-generated data and database development have been discussed elsewhere in this QAPP. All analytical data for the model's target parameters and most supporting data will have been verified through field QAPP processes before release to the modelers.

The DQOs for this project are discussed in Section 4.0 of this QAPP. Rigorous examination of precision, accuracy, completeness, representativeness, detectability, and comparability will be conducted on project-generated data by the Modeling QC Officer during model calibration and application. Project-generated data will be verified and validated using a process that controls measurement uncertainty, evaluates data, and flags or codes data against various criteria. That portion of the QA process is also associated with the final database construction. Modelers will cross-check data for bias, outliers, normality, completeness, precision, accuracy, and other potential problems. These data and processes will be documented in the TMDL Modeling Report.

Non-project-generated data might be obtained from either published or unpublished sources, and the modelers will examine those data as part of a data quality assessment. Databases that have not been published are also examined in light of a data quality assessment. Data provided by EPA or other sources will be assumed to meet precision objectives established by those entities. For example, we anticipate using IWR run 40 for the base water quality dataset. But we know additional data will become available such as collected by Tampa Bay Estuary Program and Florida's Water Management Districts. The acceptance criteria for individual data values generally address the issues described in the appendix.

12.0 Instrument/Equipment Testing, Inspection, and Maintenance

No instruments that need testing, inspection, or maintenance for collection and analysis of field samples will be used in this study. If sampling is required for the project, the instrument and equipment testing, inspection, and maintenance for the field sampling program conducted for the modeling study will be addressed in a separate field sampling QAPP.

The majority of work conducted by Tetra Tech for TMDL development in the Floyds Fork watershed will involve the acquisition or processing of data and the generation of reports and documents, both of which require the maintenance of computer resources. Tetra Tech's computers are either covered by on-site service agreements or serviced by in-house specialists. When a problem with a microcomputer occurs, in-house computer specialists diagnose the trouble and correct it if possible. When outside assistance is necessary, the computer specialists call the appropriate vendor. For other computer equipment requiring outside repair services and not covered by a service contract, local computer service companies are used on a time-and-materials basis. Routine maintenance on microcomputers is performed by in-house computer specialists. Electric power to each microcomputer flows through a surge suppressor to protect electronic components from potentially damaging voltage spikes. All computer users have been instructed on the importance of routinely archiving project data files from hard drive to compact disc or floppy disk storage. The Atlanta and Fairfax office network servers are backed up on tape nightly during the week. Screening for viruses on electronic files loaded on microcomputers or the network is standard company policy. Automated

screening systems have been placed on Tetra Tech's computer systems and are updated regularly to ensure that viruses are identified and destroyed promptly.

13.0 Instrument/Equipment (Model) Calibration and Frequency

If sampling is required for this project, the calibration and frequency of calibration for instruments and equipment used to collect new data will be addressed in a separate field sampling QAPP.

A model calibration is a measure of how well the model results represent field data. Because the TMDL program requires the implementation of load reduction scenarios that may, in many cases, require enormous capital expenditures, the use of a calibrated model, the scientific veracity of which is well defined, is of paramount importance..

The Tetra Tech TOL and Modeling Task Leader will direct the model calibration efforts. Some model parameters will need to be estimated using site-specific field data for the model's application. Some example parameters follow:

- Kinetic coefficients and parameters (e.g., partition coefficients, decay coefficients)
- Forcing terms (e.g., sources and sinks for state variables)
- Boundary conditions (specified concentrations, flows)

Models are often calibrated through a subjective trial-and-error adjustment of model input data because a large number of interrelated factors influence model output. The model calibration *goodness of fit* measure can be either qualitative or quantitative. Qualitative measures of calibration progress are commonly based on the following:

- Graphical time-series plots of observed and predicted data
- Graphical transect plots of observed and predicted data at a given time interval
- Comparison between contour maps of observed and predicted data, providing information on the spatial distribution of the error
- Scatter plots of observed versus predicted values in which the deviation of points from a 45-degree straight line gives a sense of fit
- Tabulation of measured and predicted values and their deviations

The TMDL models will be calibrated to the best available data, including literature values and interpolated or extrapolated existing field data. If multiple data sets are available, an appropriate period and corresponding data set will be chosen on the basis of factors characterizing the data set, such as corresponding weather conditions, amount of data, and temporal and spatial variability of data. The model will be considered calibrated when it reproduces data within an acceptable level of accuracy or approved by the EPA modeling subgroup. Sensitivity analysis (refer to Section 17.0) is used to identify the most influential parameters in determining the accuracy and precision of model predictions. Sensitivity analysis can be used to improve the efficiency of the calibration process.

Quantitative calibration measures include time-series error measures, and other statistic-based, dimensionless performance indices. Quantitative measures allow comparison of the level of calibration and validation between modeling studies of different water bodies and different modeling studies of a specific water body. Time series error measures, particularly root mean square errors, are typically used to evaluate model performance with respect to predicting water surface elevation, temperature and salinity. There are not quantifiable limits because the Tetra Tech modelers and the

EPA TOM may decide for a particular station that the statistics (quantitative) are more or less important than the graphical plots (qualitative). This will be a site-specific discussion for each estuary; the limits used will be documented in the watershed and estuary modeling reports.

14.0 Inspection/Acceptance of Supplies and Consumables

Not applicable.

15.0 Nondirect Measurements

Nondirect measurements (also referred to as non-project-generated data) are data that were previously collected under a different effort outside this contract. Nondirect data can come from a number of sources, but the nondirect data most often used in developing numeric nutrient criteria modeling projects are typically obtained from EPA, NOAA NCDC, USGS NHD+ and NED, USDA NRCS, and databases maintained by state agencies.

Non-project-generated data could be obtained from published or unpublished sources. The published data will have been previously peer reviewed. Those data are generally examined by modelers as part of a data quality assessment. Databases that have not been published are also examined in light of a data quality assessment. Data provided by EPA or other sources will be assumed to meet precision objectives established by those entities, as described in the acceptance criteria issues described in the appendix. If historical data are used, a written record of where the data were obtained and any information on their quality will be documented in the final report (see Section 11.0).

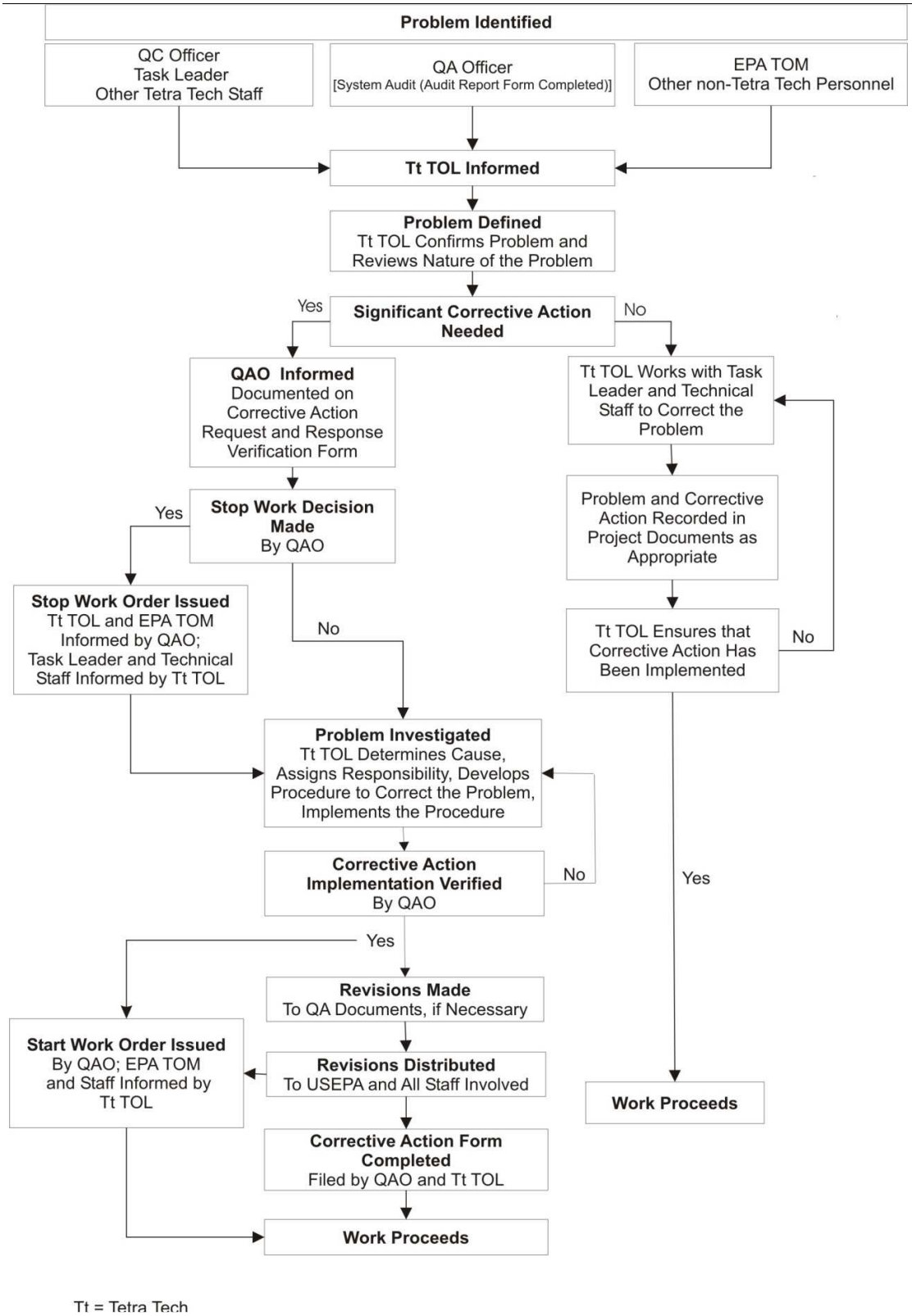
16.0 Data Management

The data management process and the computer hardware and software configuration requirements will be developed and submitted to the EPA TOM for review before model equations and related algorithms are coded into an integrated, efficient computer code. Modeling staff members will work closely with the TL and will consult with experts as necessary to ensure the theory is accurately represented in the code. The modeling code is continually checked by the developers and compared to bench test runs to ensure the accuracy of the mechanistic equations and solution techniques. The QC Officer will conduct internal reviews of the computer code, as appropriate.

17.0 Assessment and Response Actions

The QA program under which this project will operate includes surveillance, with independent checks of the data obtained from sampling, analysis, and data-gathering activities. This process is illustrated in Figure 4. The essential steps in the QA program are as follows:

- Identify and define the problem
- Assign responsibility for investigating the problem
- Investigate and determine the cause of the problem
- Assign and accept responsibility for implementing appropriate corrective action
- Establish the effectiveness of and implement the corrective action
- Verify that the corrective action has eliminated the problem



Tt = Tetra Tech

Figure 2. Problem assessment and correction options.

Many of the technical problems that might occur can be solved on the spot by the staff members involved, for example, by modifying the Initial Technical Approach or correcting errors or deficiencies in documentation. Immediate corrective actions form part of normal operating procedures and are noted in records for the project. Problems that cannot be solved in that way require more formalized, long-term corrective action.

If quality problems that require attention are identified, Tetra Tech will determine whether attaining acceptable quality requires either short- or long-term actions. If a failure in an analytical system occurs (e.g., performance requirements are not met), the Modeling QC Officer will be responsible for corrective action and will immediately inform the Tetra Tech TOL or the QA Officer, as appropriate. Subsequent steps taken will depend on the nature and significance of the problem, as illustrated in Figure 4.

CORRECTIVE ACTION REQUEST AND RESPONSE VERIFICATION

Contract (name) _____

Date of Assessment _____ Request No. _____

Title (of project or other) _____

Project Leader _____ TC# _____

Other Responsible Personnel _____

Auditor or Initiator of This Corrective Action Request _____

Problem Description:

Recommended Action: **Date to Be Completed:**

_____ _____
Quality Assurance Officer Date

_____ _____
Principal-in-Charge or Program Manager Date

Action Taken: **Date:**

Verification of Completion of Corrective Action:

_____ _____
Quality Assurance Officer Date

_____ _____
Principal-in-Charge or Program Manager Date

Original form to be filed in QAO File; one copy to be filed in Project File and one copy in Contract File (if corrective action pertains to a project), or one copy to be filed in Contract File (if corrective action pertains to a contract).

Figure 3. Corrective Action Request and Response Verification form.

The Tetra Tech TOL and QC Officer have primary responsibility for monitoring the activities of this project and identifying or confirming any quality problems. These problems will also be brought to the attention of the Tetra Tech QA Officer, who will initiate the corrective action system described above, document the nature of the problem (using a form such as that shown in Figure 5), and ensure that the recommended corrective action is carried out. The Tetra Tech QA Officer has the authority to stop work on the project if problems affecting data quality that will require extensive effort to resolve are identified.

The EPA TOM and Tetra Tech TOL will be notified of major corrective actions and stop work orders. Corrective actions could include the following:

- Reemphasizing to staff the project objectives, the limitations in scope, the need to adhere to the agreed-upon schedule and procedures, and the need to document QC and QA activities.
- Securing additional commitment of staff time to devote to the project.
- Retaining outside consultants to review problems in specialized technical areas.
- Changing procedures. The Tetra Tech TOL can replace a staff member, if appropriate, if it is the best interest of the project to do so.

Performance audits are quantitative checks on different segments of project activities; they are most appropriate for sampling, analysis, and data-processing activities. The QC Officer is responsible for overseeing work as it is performed and periodically conducting internal assessments during the data entry and analysis phases of the project. As data entries, model codes, calculations, or other activities are checked, the QC Officer will sign and date a hard copy of the material or complete Tetra Tech's

standard Technical/Editorial Review Form, as appropriate, and provide it to the Tetra Tech TOL and TL to include in the administrative record. Performance audits will consist of comparisons of model results with observed historical data. Performing control calculations and post-simulation validation of predictions are major components of the QA framework.

The Tetra Tech TOL will periodically perform or oversee the following qualitative and quantitative assessments of model performance to ensure that the model is performing the required task while meeting the quality objectives:

- Data acquisition assessments
- Model calibration studies
- Sensitivity analyses
- Uncertainty analyses
- Data quality assessments
- Model evaluations
- Internal reviews

Sensitivity to variations, or uncertainty in input parameters, is an important characteristic of a model. Sensitivity analysis is used to identify the most influential parameters in determining the accuracy and precision of model predictions. That information is important to the user who must establish required accuracy and precision in model application as a function of data quantity and quality. Sensitivity analysis quantitatively or semi-quantitatively defines the dependence of the model's performance assessment measure on a specific parameter or set of parameters. Sensitivity analysis can also be used to decide how to simplify the model simulation and to improve the efficiency of the calibration process.

Model sensitivity can be expressed as the relative rate of change of selected output caused by a unit change in the input. If the change in the input causes a large change in the output, the model is considered to be sensitive to that input parameter. Sensitivity analysis methods are mostly nonstatistical or even intuitive by nature. Sensitivity analysis is typically performed by changing one input parameter at a time and evaluating the effects on the distribution of the dependent variable. Nominal, minimum, and maximum values are specified for the selected input parameter.

Initially, sensitivity analysis is performed at the beginning of the calibration process to design a calibration strategy. After the calibration is completed, a more elaborate sensitivity analysis is performed to quantify the uncertainty in the calibrated model caused by uncertainty in the estimates of the model input parameters.

Informal sensitivity analyses (iterative parameter adjustments) are generally performed during model calibrations to ensure that reasonable values for model parameters will be obtained, resulting in acceptable model results. The degree of allowable adjustment of any parameter is usually directly proportional to the uncertainty of its value and is limited to its expected range of values. Formal sensitivity analyses will be performed in accordance with technical direction from the EPA TOM when a certain aspect of the system requires further investigation. For example, formal sensitivity analyses are often performed on the effects of loadings from different sources on instream water quality to allow the development of more feasible and reasonable allocations and load reductions on the basis of the dominant sources.

Internal reviews, as well as results of EPA modeling subgroup reviews provided to Tetra Tech, will be documented in the project files. Documentation will include the names, titles, and positions of the reviewers; their report findings; and the project management's documented responses to their findings.

The Tetra Tech TOL will perform surveillance activities throughout the duration of the project to ensure that management and technical aspects are being properly implemented according to the schedule and quality requirements specified in this QAPP. The surveillance activities will include assessing how project milestones are achieved and documented, corrective actions are implemented, budgets are adhered to, reviews are performed, and data are managed and whether computers, software, and data are acquired in a timely manner.

System audits are qualitative reviews of project activity to check that the overall quality program is functioning and that the appropriate QC measures identified in the QAPP are being implemented. If requested by the EPA TOM, and additional funding is provided by EPA, the Tetra Tech QA Officer or designee will conduct an internal system audit of the project and report results to the EPA TOM and Tetra Tech TOL.

18.0 Reports to Management

Throughout the TMDL development process, , the Tetra Tech TOL will submit to the EPA TOM technical memoranda describing each major step in TMDL development for review and approval before taking further steps.

On a monthly basis, the Tetra Tech will provide the EPA TOM with a report describing the status of the project and the results of any intermediate assessments. The results of the project will be provided to the EPA TOM in TMDL and modeling reports submitted for inclusion in the overall administrative record. In addition, Tetra Tech will deliver the project files for TMDL development that will contain copies of all records and documents, including soft copy versions of the data and model input data sets. Tetra Tech will deliver the files to EPA at the end of the project.

The draft and final TMDL modeling reports will include a separate section titled *Data Quality* to relate the results of the study back to this QAPP and the project action plan. This section will contain, at a minimum, (1) where all the data were obtained, (2) the reason the data were originally collected, and (3) all readily available information related to QA associated with the collection and handling of the data. This section will also present a brief summary of the QA procedures used when the data were processed. The report will also include results of technical reviews, model tests, data quality assessments of output data and audits, actual input and databases used, response actions to correct model development of implementation problems, and if applicable, pre- and post-software development.

19.0 Data Review, Verification, and Validation

Data review and validation services provide a method for determining the usability and limitations of data and provide a standardized data quality assessment. Verification of new model components or parameters (when applicable) improves the predictive capabilities of new models or modified existing models. Experienced professionals will be used in the data review, compilation, and evaluation phases of the study. Tetra Tech will be responsible for reviewing data entries, transmittals, and analyses for completeness and adherence to QA requirements. The data will be organized in a

standard database on a microcomputer. A screening process that scans through the database and flags data that are outside typical ranges for a given parameter will be used. Values outside typical ranges will not be used to develop model calibration data sets or model kinetic parameters.

20.0 Verification and Validation Methods

The QC Officer will review or oversee review of all data related to the project for completeness and correctness. The Tetra Tech modeling staff will make all data available to the QC Officer within 2 weeks of receiving data. The QC Officer will identify any issues of concern to the Tetra Tech TOL, and he will resolve those issues with the modeling team.

Raw data received in hard copy format will be entered into the standard database. All entries will be compared to the original hard copy data sheets by the team personnel. Screening methods will be used to scan through the database and flag data that are outside typical ranges for a given parameter. Data will also be manipulated using specialized programs and Microsoft Excel 2007. Ten percent of the calculations will be recalculated to ensure that correct formula commands were entered into the program. If 5 percent of the data calculations are incorrect, all calculations will be rechecked after the correction is made to the database. Data quality will be assessed by comparing entered data to original data; performing the data and model evaluations described in Sections 4.0, 11.0, and 17.0; and comparing results with the measurement performance or acceptance criteria summarized in the data review and technical approach documentation to determine whether to accept, reject, or qualify the data. Results of the review and validation processes will be reported to the EPA TOM.

General guidelines and procedures for model data validation and calibration are listed in Section 13.0. Verification will be performed by comparing new model parameters or components to theory. Tetra Tech will specify which methods will be used to ensure that the model results will ultimately be valid for EPA use in each project-specific data review and technical approach document. Model validation evaluates the model's ability to appropriately simulate conditions under a data set or time period that is independent from those used in the calibration. The calibration and validation process will be documented in the TMDL report.

Because the goal is to be able to predict when point and nonpoint source loads produce water quality impairment on the basis of the ambient water quality criteria, model calibration and validation should strive to reduce errors (deviations between model predictions and observed measurement data) to zero.

A set of parameters used in the calibrated model might not accurately represent field values, and the calibrated parameters might not represent the system under a different set of boundary conditions or hydrologic stresses. Therefore, a second model validation period helps establish greater confidence in the calibration and the predictive capabilities of the model. A site-specific model is considered *validated* if its accuracy and predictive capability have been proven to be within acceptable limits of error independently of the calibration data. In general, model validation is performed using a data set that differs from the calibration data set (i.e., low-flow data set for calibration versus higher-flow data set for verification). If only a single time series is available, the series can be split into two sub-series, one for calibration and another for validation. If the model parameters are changed during the validation, the exercise becomes a second calibration, and the first calibration needs to be repeated to account for any changes. Acceptable limits are those defined by the combined process of quantitative and qualitative examination of the model versus the data. There are not quantifiable limits because the Tetra Tech modelers and the EPA modeling subgroup may decide for a particular

station that the statistics (quantitative) are more or less important than the graphical plots (qualitative). This will be a site-specific discussion for each estuary; the limits used will be documented in the modeling report.

Model validation will be accomplished by calibration. A model calibration is the process of adjusting model inputs within acceptable limits until the resulting predictions give good correlation with observed data. Commonly, the calibration begins with the best estimates for model input on the basis of measurements and subsequent data analyses. Results from initial simulations are then used to improve the concepts of the system or to modify the values of the model input parameters. The success of a model calibration is largely dependent on the validity of the underlying model formulation.

21.0 Reconciliation with User Requirements

All data quality indicators will be calculated at the completion of the data analysis phase. Measurement quality requirements will be met and compared with the DQOs to confirm that the correct type, quality, and quantity of data are being used for model development in support of numeric nutrient criteria in Florida estuaries. The interpretation and presentation stage includes inspection of the form of the results, and the meaning and reasonableness of the computation results and post-simulation analysis.

The Tetra Tech TOL, Modeling Task Leader, or their designees will perform internal reviews to assess departures from assumptions established in the planning phase of the TMDL modeling process. Tetra Tech, in consultation with the EPA TOM and alternate TOM, will determine how anomalies will be resolved.

If requested by the EPA TOM, Tetra Tech will perform a post-audit for the project. A post-audit is an evaluation of the correctness of the initial model predictions conducted several years after the original modeling study is completed. If the models' predictions were accurate, the model can be considered valid for the specific site and the actual stresses. A post-audit requires new field observations for the predicted variables, which are to be collected at a time after the system has had a chance to adjust to the management changes. Uncertainties and limitations in the use of such data and interpretation of results will be provided to EPA.

References

Anscombe, P.J. 1967. Topics in the investigation of linear relations fitted by the method of least squares. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)* 29:1–52.

Arhonditsis, G.B., and M.T. Brett. 2004. Evaluation of the current state of mechanistic aquatic biogeochemical modeling. *Marine Ecology Progress Series* 271:13–26.

Chapra, S.C. 2003. Engineering water quality models and TMDLs. *Journal of Water Resources Planning and Management*. 129(4): 247–256.

Di Toro, D.M., and J.P. Connolly. 1980. *Mathematical Models of Water Quality in Large Lakes, Part 2: Lake Erie*. EPA-600-3-80-065. U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC.

Di Toro, D.M., J.J. Fitzpatrick, and R.V. Thomann. 1983. *Water Quality Analysis Simulation Program (WASP) and Model Verification Program (MVP) Documentation*. Hydroscience, Inc., Westwood, NY.

Hamrick, J.M. 1992. *A Three-Dimensional Environmental Fluid Dynamics Computer Code: Theoretical and Computational Aspects*. Special Report 317. The College of William and Mary, Virginia Institute of Marine Science, Williamsburg, VA.

James, R.T., J. Martin, T. Wool, and P.F. Wang. 1997. A sediment resuspension and water quality model for Lake Okeechobee. *Journal of the American Water Resources Association* 33:661–680.

Lumb, A.M., R.B. McCammon, and J.L. Kittle Jr. 1994. *Users Manual for an Expert System (HSPEXP) for Calibration of the Hydrologic Simulation Program–Fortran*. U.S. Geological Survey Water-Resources Investigations Report 94-4168. U.S. Geological Survey, Reston, VA.

Multi-Resolution Land Characteristics Consortium (MRLC). 2007. National Land Cover Data 2005. U.S. Environmental Protection Agency, National Exposure Research Laboratory, Multi-Resolution Land Characteristics Consortium, Research Triangle Park, NC. <http://www.epa.gov/mrlc/>. Accessed February 2011.

Murrell, M.C., J.G. Campbell, J.D. Hagy, and J.M. Caffrey. 2009. Effects of irradiance on benthic and water column processes in a Gulf of Mexico estuary: Pensacola Bay, Florida, USA. *Estuarine Coastal and Shelf Science* 81(4):501–512.

National Climatic Data Center (NCDC). 2011. *National Climatic Data Center Online Climate Data Directory*. National Oceanic and Atmospheric Administration, National Climatic Data Center, Asheville, NC. <http://www.ncdc.noaa.gov/oa/climate/climatedata.html>. Accessed February 2011.

Thomann, R.V. 1975. *Mathematical Modeling of Phytoplankton in Lake Ontario, 1. Model Development and Verification*. EPA-600/3-75-005. U.S. Environmental Protection Agency, Corvallis, OR.

Thomann, R.V., and J.J. Fitzpatrick. 1982. *Calibration and Verification of a Mathematical Model of the Eutrophication of the Potomac Estuary*. Washington, District of Columbia, Department of Environmental Services, Washington, DC.

Thomann, R.V., R.P. Winfield, D.M. Di Toro, and D.J. O'Connor. 1976. *Mathematical Modeling of Phytoplankton in Lake Ontario, 2. Simulations Using LAKE 1 Model*. EPA-600/3-76-065. U.S. Environmental Protection Agency, Grosse Ile, MI.

Thomann, R.V., R.P. Winfield, and J.J. Segna. 1979. *Verification Analysis of Lake Ontario and Rochester Embayment Three Dimensional Eutrophication Models*. EPA-600/3-79-094. U.S. Environmental Protection Agency, Grosse Ile, MI.

USEPA (U.S. Environmental Protection Agency). 2001 (Reissued May 2006). *EPA Requirements for Quality Assurance Project Plans, QA/R-5*. EPA 240-B-01-003. U.S. Environmental Protection Agency, Office of Environmental Information, Washington, DC.

———. 2002a. *Guidance for Quality Assurance Project Plans for Modeling, QA/G-5M.* EPA 240-R-02-007. U.S. Environmental Protection Agency, Office of Environmental Information, Washington, DC.

———. 2006. *Guidance on Systematic Planning Using the Data Quality Objectives Process, EPA QA/G-4*. EPA 240-B-06-001. U.S. Environmental Protection Agency, Office of Environmental Information, Washington, DC.

———. 2009a. *Office of Water Quality Management Plan*. EPA 821/R-09/001. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

———. 2009b. *Guidance on the Development, Evaluation, and Application of Environmental Models.* EPA/100/K-09/003. U.S. Environmental Protection Agency, Office of the Science Advisor, Council for Regulatory Environmental Modeling, Washington, DC.

Wool, T.A., S.R. Davie, and H.R. Rodriguez. 2003. Development of a Three Dimensional Hydrodynamic and Water Quality Model to Support the TMDL Decision Process for the Neuse River Estuary, North Carolina. *American Society of Engineers Journal of Water Resources Planning and Management* 129(129):295-306.

APPENDIX A

DATA QUALITY INDICATOR DEFINITIONS

This page left intentionally blank

DATA QUALITY INDICATOR DEFINITIONS

Measurement acceptance or performance criteria are quantitative statistics used to interpret the degree of acceptability or utility of the data to the user. The quality of existing environmental monitoring data and generated data is some measure of the types and amount of error associated with the data. Those criteria, also known as *data quality indicators*, are the following:

- Precision
- Accuracy
- Representativeness
- Comparability
- Completeness

Data used in model development are generally data in federal and state government water quality databases. Data obtained from government agency databases should have already been screened and met specified measurement performance criteria. Those criteria might not be reported for the parameters of interest in the databases. In consultation with the EPA TOM, it will be determined how much effort should be expended to find reports or metadata that might contain that information. Measurement performance or acceptance criteria for various parameters will be documented in the modeling report. Parameters for which measurement performance or acceptance criteria could be set are the following:

- Software run time
- Software processing capabilities
- Model prediction results relative to decision error
- Data used in model(s)

Precision is a measure of internal method consistency. It is demonstrated by the degree of mutual agreement between individual measurements or enumerated values of the same property of a sample, usually under demonstrated similar conditions. Precision of field sampling methods is estimated by taking duplicate samples for analysis. This QC calculation also addresses uncertainty due to natural variation and sampling error.

Precision of available data used for this project will be noted if available. Precision of generated data produced by the model will be examined by performing replicate runs.

Accuracy is defined as the degree of agreement between an observed value and an accepted reference or true value. Accuracy is a combination of random error (precision) and systematic error (bias), which are due to sampling and analytical operations. Bias is the systematic distortion of a measurement process that causes errors in one direction so that the expected sample measurement is always greater or lesser to the same degree than the sample's true value. Because accuracy is the measurement of a parameter and comparison with a *truth*, and the true values of environmental physicochemical characteristics cannot be known, use of a surrogate is required.

Accuracy of non-direct data obtained from government agency databases and entered into the project database can be expressed as the percentage of values, by field, not included as valid values in their associated system reference tables. For example, a code entered incorrectly or in the wrong field would constitute inaccurate data. The accuracy of non-direct data will be controlled by

double-checking all automatically mapped data. Accuracy of the model will be determined by comparing the contaminant concentrations calculated for a given area with actually measured contaminant concentrations reported in the database under conditions used in the model simulation. Accuracy of data entry into the project database will be controlled by double-checking all manual data entries.

Data representativeness is defined as the degree to which data accurately and precisely represent a characteristic of a population, a parameter, variations at a sampling point, a process condition, or an environmental condition. It therefore addresses the natural variability or the spatial and temporal heterogeneity of a population. Comparisons of the loadings data and measured environmental concentrations will be made to examine sources and sinks of materials. Preliminary knowledge of the area will be used to select appropriate sites and stations in the vicinity of point source discharges for the initial and later modeling phases.

Two data sets are considered to be comparable when there is confidence that the two sets can be considered equivalent with respect to the measurement of a specific variable or group of variables. Measurement data used in the model will follow protocols established by the appropriate government agency to permit comparisons of water quality data at different sites on the study site. Data sets will be examined with respect to variables of interest, commonality of units of measurement, and similarity in analytical and QA procedures. Additional comparability of data can be ensured by similarity in geographic, seasonal, and sampling method characteristics.

Completeness is defined as the percentage of measurements made that are judged to be valid according to specific criteria and entered into the data management system. To achieve that objective, reasonable effort is made to avoid accidental or inadvertent sample or data loss. Lack of data entered into the databases will reduce the ability of the project to calibrate and verify the model. Although some fields in the project database should never contain blanks (e.g., facility name), other fields could be impossible to fill or might not be filled until later (e.g., completion date of an activity). Completeness is thus also defined as the percentage of data available to cover all aspects of model development. In any complex model study, it is inevitable that some data gaps will exist. The data gaps and the assumptions used in filling the gaps will be documented in the modeling report. Percent completeness (%*C*) for measurement parameters can be defined as follows:

$$\%C = \frac{v}{T} \times 100$$

where v = the number of measurements judged valid and T = the total number of measurements. For this project, the model application will be considered complete when no less than 85 percent of the measurement data, parameter variables, and output values are judged valid; however, other considerations must be taken into account as well, depending on the use of the data.

Acceptance criteria will be obtained from any existing QAPPs, sampling and analysis plans, standard operating procedures, laboratory reports, and other correspondence for a given source of non-direct measurement data, if available. The data assessment and quality guidelines associated with a given type of measurement will be developed from those sources and documented. The secondary data will be reviewed and compared with the guidelines in this QAPP. Data not meeting the acceptance criteria requirements will be rejected or their status documented, as deemed appropriate by the EPA TOM and Tetra Tech TOL and TL.

Model Sensitivity Analysis

The sensitivity to variations or uncertainty in input parameters is an important characteristic of a model. Sensitivity analysis is used to identify the most influential parameters in determining the accuracy and precision of model predictions. That information is of importance to the user who must establish required accuracy and precision in model application as a function of data quantity and quality. Sensitivity analysis quantitatively or semi-quantitatively defines the dependence of the model's performance assessment measure on a specific parameter or set of parameters. Sensitivity analysis can also be used to decide how to simplify the model simulation and to improve the efficiency of the calibration process.

Model sensitivity can be expressed as the relative rate of change of selected output caused by a unit change in the input. If the change in the input causes a large change in the output, the model is then considered to be sensitive to that input parameter. Sensitivity analysis methods are mostly nonstatistical, or even intuitive by nature. Sensitivity analysis is typically performed by changing one input parameter at a time and evaluating the effects on the distribution of the dependent variable. Nominal, minimum, and maximum values are specified for the selected input parameter.

Initially, sensitivity analysis is performed at the beginning of the calibration process to design a calibration strategy. After a calibration is completed, a more elaborate sensitivity analysis is performed to quantify the uncertainty in the calibrated model caused by uncertainty in the estimates of the model input parameters.

Informal sensitivity analyses (iterative parameter adjustments) are generally performed during model calibrations to ensure that reasonable values for model parameters will be obtained, resulting in acceptable model results. The degree of allowable adjustment of any parameter is usually directly proportional to the uncertainty of its value and is limited to its expected range of values. Formal sensitivity analyses will be performed in accordance with technical direction from the EPA TOM when a certain aspect of the system requires further investigation.

APPENDIX B

Data Quality Requirements for Secondary and Calculated Data

This page left intentionally blank

DATA QUALITY REQUIREMENTS
SECONDARY DATA

Data Element	Source(s)	Acceptance Criteria
Technical information or specifications	Government publications (e.g., technical support documents, Reports to Congress, fact sheets) Scientific journal Trade journal Book or handbook Compliance data Maps, GIS layers, plots, land surveys, photographs NPDES permit data from the permittee, EPA, and/or the State	Received from the client and stated by the client to be acceptable Received from the client and verified accurate against its cited original data source Published in peer-reviewed journal Published in reliable source Statements of data quality are documented (QC checks implemented, level of oversight)
Measurements obtained during a primary data collection activity	Government databases (e.g., EPA Environmental Information Management System: STORET, PCS, or TRI; USGS water quality; Census Bureau statistics) Government catalog (e.g., EPA's Environmental Data Registry) Government report Scientific journal Unpublished research or pilot studies NPDES permit data from the permittee, EPA, and/or the State	Metadata are available that explain the study objectives and conditions under which the data were obtained QC implementation is described Statements of data quality are provided (e.g., data quality indicators: accuracy, bias, precision, completeness, comparability, representativeness, sensitivity) Latitude/longitude of sample is correct Sample type is clearly identified (e.g., sediment, water, air) Sample collection data is provided Measurement or analytical method used is stated Data are reasonable (e.g., correct units, within accepted ranges)

Data Element	Source(s)	Acceptance Criteria
Data generated from calculations by using spreadsheet or modeling tools	Published reports NPDES permit data from the permittee, EPA, and/or the State	Published in peer-reviewed journal or report Data sources used are cited Limitations and uncertainties in the data used to generate the results are discussed Confidence in the estimated values from the calculations performed is provided
Results of surveys and their interpretation	Published reports	Survey objectives and applicability of results are clearly stated Respondent selection criteria are provided Participation rate is stated Dates survey conducted provided
Site investigation/audit findings	EPA and state agencies	Collected by EPA, a state agency, or an approved contractor Findings based on actual observations or facility records Findings released by the agency
Comments	Comments submitted by the public on draft/final policies, documents, or rules	All records accepted

Procedure for Achieving Secondary Data Quality Requirements

The following approach will be used to evaluate the quality of secondary data and ensure accuracy in its use in deliverables:

1. Identify specific data needs for the particular deliverable. The data elements listed above are general categories of data that might be useful in developing responses to comments, syntheses, evaluations, or analyses. Specific data (parameters, information, model results) must be determined for each need (e.g., water chemistry values taken during the last five years for total phosphorus or dissolved organic carbon for a stream segment in a particular county or modeled results of phosphorus loadings for that stream segment). Because data needs could change during the course of deliverable preparation, communicate with the client frequently.

2. Identify source(s) from which to obtain the data. When possible, prioritize the sources based on confidence in the data, for example, the original database or published report where the data first appeared is preferable to its citation in a review chapter in a book.

3. Obtain the database or publication. Be sure the database is marked “Read Only.” Copyright

restrictions on publications may not permit copying a page or document for the project file. Flag page in book or journal where data are found that will be extracted to the deliverable. Whenever possible, secondary data will be downloaded from various electronic sources to reduce manual data entry.

4. Prepare a list of specific acceptance criteria for the project, including spatial, temporal, and data quality requirements to provide consistency in the evaluation. For example, measurement data may only be acceptable only if collected from a single station during the past five years, but technical information may be acceptable if prepared during the past ten years and regionally relevant. Determine what is important to know about the data or information to be able to judge its utility to the project and confidence in its accuracy.

5. Read the publication or metadata and evaluate the data or information against the acceptance criteria. Decide whether it can be used without concern or whether its use must be justified or limited or whether it should not be used in the preparation of the deliverable. Consult with the client if there are any questions about acceptability. The client must provide written technical direction if additional work should be performed to verify values and extract statements of data quality from the raw data, metadata, or original final report. If, in consultation with the client (who has considered the use of the data and importance of the decision to be made), it is determined that such searches are not necessary or no quality requirements exist or can be established but the data must be used in the project, a disclaimer is added to the deliverable indicating that the quality of the secondary data is unknown:

   The quality of the secondary data used [in this project] could not be determined.

6. Document the outcome of the evaluation and decision (e.g., use a spreadsheet tracking system, checklist, rubber stamp, or written justification). Note if you have to use the data because they are the only data available and you need to proceed with the project or whether background information is not available.

7. As a database is populated (data entry) or deliverable is prepared (word processing), enter the value or information and then review the transfer for accuracy.

8. Write the complete citation for the source of the data, following the appropriate style guide format. Enter the short citation in the database, text, or footnote, as appropriate (e.g., Author, Year). All data needs to have its source identified during deliverable preparation.

9. A second person (e.g., QC Officer) manually checks the data transfer from hard copy materials to deliverables or databases for accuracy and proper citation of the data source. The degree to which data are checked could vary based on the amount of data involved and resources available. For example, checking hundreds of values against the original sources could be impractical because of time and costs; however, checking 10 percent of those values may be appropriate. A customized user input interface to perform certain appropriate checks on data as they are being manually entered can also reduce incorrect data entry. Additional checks that might be used, if appropriate, include displaying data graphically to visually inspect for potential errors, analyzing data to detect invalid data, and checking for duplicate data entries. If data extraction and transfer reveal more than 1 percent incorrectly entered data, more data could be checked and corrected.

10. The second reviewer marks and reports data discrepancies to the technical staff member(s) who originally extracted the data or information. The reviewer initials and dates a copy of the marked data or notes the review in a spreadsheet or on a review form to document the review.

11. Correct data in the database or deliverable and recheck to ensure that corrections are made.

DATA QUALITY REQUIREMENTS
CALCULATED DATA

Data Element	Source(s)	Acceptance Criteria
Equation representing a mathematical concept that will provide a quantitative result (value)	Government publications (e.g., technical support documents, Reports to Congress, fact sheets)	Received from the client and stated by the client to be acceptable Received from the client and verified accurate against its cited original data source
	Scientific journal	Published in peer-reviewed journal
	Trade journal	Published in reliable source
	Book or handbook	Operations are correct
	Spreadsheet	Result units are correct
	Model	Result is reasonable
	NPDES permit data from the permittee, EPA, and/or the state	

Procedure for Achieving Calculated Data Quality Requirements

The following approach will be used to evaluate the quality of calculated data and ensure accuracy in its use in deliverables:

1. Obtain the equation and document the source of the equation (e.g., textbook, journal article, spreadsheet). Transfer the equation into a spreadsheet and heed any warnings generated by the spreadsheet concerning the format or content of the equation.

2. A second person reviews the equation transfer and makes any necessary corrections.

3. Obtain the data to be analyzed using the equation or series of equations. Be sure input data have been checked for accuracy and completeness, outliers, or other issues. Document data transfers and QC checks.

4. A second person reviews the dataset format and content.

5. Perform any necessary data reductions. These are processes that change either the values or numbers of data items; the original data set from which reduced data are generated cannot be recovered, i.e., there is a loss of information. For example, averages of replicate measurements may be required for statistical comparisons. For manual calculations, actual raw data are transformed to the reduced data. When a programmable calculator is used, a copy of the program used is provided to the client. When a computer is used to process large quantities of data, references to the specific program and database documentation are provided. Before using statistical techniques for data analysis or when reviewing a report where statistics have been used, the assumptions required by the methods should be evaluated (e.g., data distribution, how missing data are handled). In some cases, relaxation of the assumptions

might have a minimal effect on the results; in others, the results might be rendered invalid.

6. Perform the calculation using the equation and data set.

7. A second person checks the calculation result for reasonableness and correct units. The second reviewer marks and reports calculation errors to the technical staff member(s) who originally performed the calculation. The reviewer initials and dates a copy of the calculation or notes the review in a spreadsheet or on a review form to document the review.

8. Correct the calculation result in the database or deliverable and recheck to ensure that the correction has been made.

9. At every stage of data processing at which a permanent collection of data is stored, maintain a separate copy for purposes of integrity and security. Data are securely archived in a suitable manner. Address aspects such as storage media, conditions, location, access by authorized personnel, and retention time in consultation with the client. Before archiving, ensure that all data sets are complete, with all of the client-required data standards honored, including, for example, the use of Chemical Abstract Service registry numbers for identifying chemical substances, descriptions to define sampling locations, standard QA qualification codes, and facility identification numbers, and that appropriate metadata accompany every data set, in accordance with Federal Geographic Data Committee standards (Standards for Digital Geospatial Metadata, Washington, DC, June 1994), including product identification, data quality information, spatial data organization, spatial reference information, entity and attribute information, and metadata reference information.

USE OF SECONDARY DATA DISCLAIMER

Prior to the development of any final deliverable under this Work Assignment where secondary data has been utilized, that is intended for pubic distribution or access, Tetra Tech will consult with the EPA WAM to determine the need for an appropriate disclaimer on the document. The disclaimer will identify that the document has been produced utilizing data from secondary sources. The agency can not guarantee the accuracy or validity of such data. ***[Note: EPA will need to work with OGC to develop specific language to be used as a standard disclaimer in such circumstances.]***