Calibration and Validation of the AVGWLF Model for New England and New York State

Prepared for:

U.S. Environmental Protection Agency

Prepared by:

New England Interstate Water Pollution Control Commission
116 John Street
Lowell, MA 01852

Questions concerning this document should be directed to:

Rebecca Weidman
New England Interstate Water Pollution Control Commission
(978) 323-7929
rweidman@neiwpcc.org

Section A – Project Management

A1 Approval Page

NEIWPCC PROJECT OFFICER

______________________________________ ____________

Rebecca Weidman DATE
NEIWPCC
116 John Street, Lowell, MA 01852; (978) 323 - 7929

NEIWPCC QUALITY ASSURANCE OFFICER

______________________________________ ____________

Mike Jennings DATE
NEIWPCC
116 John Street, Lowell, MA 01852; (978) 323 - 7929

AVGWLF DEVELOPER

______________________________________ ____________

Barry M. Evans, Ph.D. DATE
Penn State Institutes of the Environment
128 Land and Water Research, University Park, PA 16802; (814) 865-3357

USEPA PROJECT OFFICER

______________________________________ ____________

Stuart Lehman DATE
USEPA
4503T Ariel Rios Building, 1200 Pennsylvania Avenue, N.W., Washington, DC 20460; (202) 566-1205

USEPA QUALITY ASSURANCE OFFICER

______________________________________ ____________

Alan Peterson DATE
USEPA
11 Technology Drive, North Chelmsford, MA 01863; (617) 918- 8322

A2 Table of Contents

List of Tables

Table 1. Acronyms and Abbreviations

AVGWLF	ArcView Generalized Watershed Loading Function
BMP	Best Management Practice
DP	Dissolved Phosphorus
GWLF	Generalized Watershed Loading Function
MEDEP	Maine Department of Environmental Protection
MRLC	Multi-Resource Land Characterization
NEIWPCC	New England Interstate Water Pollution Control Commission
NASS	National Agricultural Statistics Service
NAWQA	National Water-Quality Assessment
NH_3	Ammonia
NHDES	New Hampshire Department of Environmental Services
NLCD	National Land Cover Data
NO_2	Nitrite
NO_3	Nitrate
NOAA	National Oceanic & Atmospheric Administration
NPS	Nonpoint Source
NRCS	Natural Resource Conservation Service
NYSDEC	New York State Department of Environmental Conservation
QA	Quality Assurance
QA/QC	Quality Assurance/Quality Control
QAPP	Quality Assurance Project Plan
STATSGO	State Soil Geographic
TAC	Technical Advisory Committee
TDS	Total Dissolved Solids
TKN	Total Kjehldahl Nitrogen
TMDL	Total Maximum Daily Load
TN	Total Nitrogen
TP	Total Phosphorus
TS	Total Solids
TSS	Total Suspended Solids
USDA	United States Department of Agriculture
USGS	United States Geological Survey
VTDEC	Vermont Department of Environmental Conservation

A3 Distribution List

Rebecca Weidman, Environmental Analyst, Project Officer
New England Interstate Water Pollution Control Commission
116 John Street
Lowell, MA 01852
(978) 323-7929

Laura Blake, Environmental Analyst
New England Interstate Water Pollution Control Commission
116 John Street
Lowell, MA 01852
(978) 323-7929

Mike Jennings, Environmental Analyst, QA Officer
New England Interstate Water Pollution Control Commission
116 John Street
Lowell, MA 01852
(978) 323-7929

Barry M. Evans, Ph.D., Research Associate
Penn State Institutes of the Environment
128 Land and Water Research
University Park, PA 16802
(814) 865-3357

Stuart Lehman, Project Officer
US EPA Headquarters
4503T Ariel Rios Building
1200 Pennsylvania Avenue, N.W.
Washington, DC 20460
(202) 566-1205

Alan Peterson, QA Officer
US EPA Region 1
11 Technology Drive
North Chelmsford, MA 01863
(617) 918- 8322

Eric Perkins, TAC Member
US EPA Region 1
1 Congress Street
Suite 1100
Boston, MA 02114-2023
(617) 918-1602

Tim Clear, TAC Member
Vermont Department of Environmental Conservation
103 South Main Street
Waterbury, VT 05671
(802) 241-3770

Melissa Evers, TAC Member
Maine Department of Environmental Protection
SHS #17
Augusta, ME 04333
(207) 287-2838

Russell Isaac, TAC Member
Massachusetts Department of Environmental Protection
627 Main Street, 2nd Floor
Worcester, MA 01608
(508) 767-2876

Shohreh Karimipour, TAC Member
New York State Department of Environmental Conservation
Division of Water, Bureau of Water Permits
625 Broadway
Albany, New York 12233
(518) 402-8102

Kelly L. Streich, TAC Member
Connecticut Department of Environmental Protection
79 Elm Street
Hartford, CT 06106
(860) 424-3864

Eric Williams, TAC Member
New Hampshire Department of Environmental Services
P.O. Box 95
Concord, NH 03302-0095
(603) 271-2358

A4 Project/Task Organization

New England Interstate Water Pollution Control Commission – Lead Organization

Rebecca Weidman, Project Manager and Co-Chair
This person is responsible for the overall coordination, management, and implementation of the Northeast AVGWLF Model Project in accordance with the cooperative agreement and workplan approved by the EPA Project Officer. Responsibilities include QAPP development, TAC

coordination, data collection and compilation, model calibration/verification, and report development.

Laura Blake, Project Co-Chair
This person is responsible for assisting with the implementation of the Northeast AVGWLF Model Project, including QAPP development, TAC coordination, data collection and compilation, model calibration/verification, and report development.

Mike Jennings, Quality Assurance Officer
This person is responsible for verifying that a QAPP is generated for projects in which environmental data are collected for decision making and that the QAPP meets the requirements of NEIWPCC's Quality Management Plan. He is also responsible for verifying that the QAPP is followed.

Water Quality Modeling Research Assistant
This person is responsible for assisting with the implementation of the Northeast AVGWLF Model Project, including QAPP development and collecting and formatting the necessary input data from federal, regional, and state agencies.

Pennsylvania State University

Barry M. Evans, Ph.D.
This person is responsible for advising the formatting and calculation of input data for AVGWLF. He will be in charge of model calibration and validation.

Environmental Protection Agency

Stuart Lehman, Project Officer
This person is responsible for oversight of tasks performed by NEIWPCC. This includes reviewing and approving the project workplan and QAPP, as well as reviewing progress and deliverables, including a final report.

Alan Peterson, Quality Assurance Officer
This person is responsible for reviewing and approving the QAPP, its attachments, and subsequent amendments.

Project Technical Advisory Committee
Eric Perkins, EPA Region 1
Tim Clear, VT DEC
Melissa Evers, ME DEP
Russ Isaac, MA DEP
Shohreh Karimipour, NYS DEC
Kelly Streich, CT DEP
Eric Williams, NH DES

A Technical Advisory Committee (TAC) has been created by NEIWPCC to assist with the development of the Northeast AVGWLF Model Project. Responsibilities of the TAC include: provide technical assistance, advice, and recommendations on model development; review and provide input on project reports and other written documents; act as a liaison between TAC and individual state programs; assist in the compilation of model input data; and participate in all teleconferences and meetings.

A5 Project Description and Background

Introduction

Our nation's waterbodies continue to receive significant amounts of pollutant loading, much of which originates from nonpoint sources. Section 303(d) of the Clean Water Act and EPA guidance require states to identify waters that fail to meet (or are not expected to meet) water quality standards. Such waters are considered water quality-limited and require the development of Total Maximum Daily Loads (TMDLs). Methods for TMDL development and/or determining the extent of nonpoint source pollutant loads typically include long-term surface water monitoring and computer-based simulation modeling. As resources for monitoring have declined, reliance on modeling (for making necessary determinations) has increased.

Nonpoint source pollution (NPS) is considered a primary threat to the quality of waters in the country. Section 319 of the Clean Water Act presents guidelines for the implementation of state NPS management programs; specifically, the guidance documents urge state NPS programs to implement a watershed approach. This entails the development of watershed-based plans that should identify sources of pollutants, describe management measures necessary to achieve pollutant (nitrogen, phosphorus, sediment) load reductions, and estimate these resulting pollutant load reductions.

The Environmental Protection Agency's Nonpoint Source Program Grants guidelines and the TMDL Regulations and Guidance both advocate a watershed approach to better address water quality problems. Both of these guidelines and regulations require the development of pollutant load reduction estimates to a watershed. Modeling has become an essential tool for evaluating the sources and controls of sediment and nutrient loading to surface waters. For the NPS program however, there is concern over the reporting inconsistencies of load reduction estimates (LRE). Such inconsistencies may arise through the use of more than one model, since different models have different purposes and levels of accuracy. In addition, there are huge variations in estimated pollutant load reductions being reported by different states. The states have therefore expressed a desire to use one model that is neither too complicated nor oversimplified. Using a regional approach to develop LREs will help eliminate data reporting inconsistencies and give a better overall picture of the status of regional water quality. The states therefore recognize the tremendous benefits provided by a model that is regional in scope.

Background

The New England Interstate Water Pollution Control Commission (NEIWPCC) is calibrating and validating a Northeast AVGWLF Model in response to its member states' need for regionally-consistent tools for estimating nutrient and sediment pollutant load reductions. Dr. Barry Evans of the Pennsylvania State University's Environmental Resources Research Institute developed the comprehensive, GIS-based modeling approach for use in predicting nutrient and sediment loads for the state of Pennsylvania. The state of Pennsylvania was interested in developing a model that would not need to be calibrated prior to each use, but that could accurately estimate nutrient and sediment loading for every watershed in the state, including those for which there were minimum water quality data available.

The use of AVGWLF in Pennsylvania showed that the model provides reasonably good estimates for watersheds that exhibit a wide range of landscape characteristics. Based on 32 calibration and verification watersheds in the state, Nash-Sutcliffe coefficients of correlation were derived with high values ranging from 0.92 to 0.97 for mean annual nutrient loads. AVGWLF was successful at simulating nutrient load variations for monthly, seasonal, and yearly time periods. The success of AVGWLF applications in Pennsylvania and its applicability to a variety of water programs (e.g., NPS, TMDL, monitoring, etc.) made it a highly-desirable model for development and for calibration in the Northeast region.

The core model within AVGWLF is the Generalized Watershed Loading Function (GWLF) model. The US EPA has endorsed GWLF as a good "mid-level" model that employs the algorithms needed to simulate the mechanisms that guide nutrient fluxes within a watershed. GWLF provides a simulation of sediment and nutrient (N and P) loads from a watershed. It integrates loadings from sources such as septic systems and point source discharges. It is a continuous simulation model that uses daily time steps for weather data and water balance calculations. Monthly calculations are derived from the daily water balance in order to obtain sediment and nutrient loads.

AVGWLF depends on GIS-based input data generation and model execution. Pennsylvania State University developed a customized interface for AVGWLF that parameterizes input data for the GWLF model. Using the interface, AVGWLF automatically loads the statewide GIS data sets. The user then provides additional information related to non-spatial parameters, which includes the start and end times for growing season and the months when manure is spread on land. The model uses this information to derive values for required input parameters and stores them in input files used to execute the model.

PRedICT (Pollution Reduction Impact Comparison Tool) is an additional tool to assist AVGWLF model users to develop more accurate sediment and nutrient loads. PRedICT was developed in order to evaluate pollution reduction strategies at the watershed level. It allows the user to create scenarios in which current landscape conditions and pollutant loads can be compared against future conditions that take into account the use of best management practices, for example, the conversion of septic system to centralized wastewater treatment, and the upgrade of treatment plants. It also includes cost information for various pollution reduction techniques. Presently, the tool provides different approaches to cost accounting so that the user can identify the strategy that will most efficiently reduce pollution and keep costs low.

The Northeast AVGWLF Model will incorporate the specifications of the model used in Pennsylvania while adding properties provided for a model developed by Penn State for Canada (known as CANWET). These additional properties include a provision for the inclusion of values for water extraction and the utilization of tile drainage spatial information.

Objectives

NEIWPCC is collaborating with Penn State to calibrate and validate the AVGWLF Model for the New England states and New York ("the Northeast"). The specialization of the AVGWLF Model for the Northeast will provide the states and their partners with an enhanced technical "tool kit" for use in the development of nonpoint source pollutant load reduction estimates and TMDLs.

The Northeast AVGWLF will be calibrated and then validated using a different sub-set of twenty-seven representative watersheds throughout New England and New York State. The calibration and validation of this model will provide the Northeast states with a tool to estimate load reduction and TMDLs more consistently for the entire region.

The project will help the states to more efficiently implement the NPS and TMDL programs by building the capacity of all levels of government to develop effective, comprehensive programs for watershed protection and management. States will be able to make more informed decisions regarding such issues as choosing BMPs for specific areas, deciding on feasibility of centralized wastewater treatment, and determining the need for treatment upgrades. This capacity-building effort will also encourage the implementation of these programs on a regional scale.

A6 Project/Task Description and Schedule

NEIWPCC will work with Dr. Barry Evans at Penn State to validate and calibrate AVGWLF. AVGWLF model development for the region will be guided by the project's technical advisory committee (TAC), made up of representatives from NEIWPCC, Penn State, EPA, the six New England states, and New York.. The project is comprised of four primary tasks: 1) preparation of a Quality Assurance Project Plan; 2) selection of the calibration watersheds; 3) compilation and preparation of the model input data; and 4) calibration/validation of the model for the Northeast.

Task 1 – Preparation of a Quality Assurance Project Plan

The QAPP will be prepared and submitted to EPA for approval in accordance with EPA Requirements for Quality Assurance Project Plans (QA/R-5, 3/01) and EPA NE Compendium of Quality Assurance Project Plan Requirements and Guidance (10/99). The QAPP must be fully approved before any data compilation and/or generation activities begin. NEIWPCC is responsible for developing the QAPP. The TAC will act as a review committee.

Task 2 – Selection of Calibration and Verification Watersheds

Twenty-seven watersheds, located in the seven participating states, will be utilized for calibrating and verifying the model output. The watersheds, selected by the TAC, range in size from 15 to over 2000 square miles. The watersheds are selected on the basis of size, characteristics, quality of available data, and location. Watersheds that span multiple state boundaries will be favored. (See Section A7 for details on calibration watershed criteria.)

Of the 27 watersheds used, about half (13 or 14) will be used for calibration purposes, and the remainder will be used for verification. An attempt will be made to have at least one calibration watershed and one verification watershed in each member state. However, since there are 7 states included in the project, most states will have 2 calibration and 2 verification watersheds. These watersheds have been selected both on the basis of available data (both GIS and monitoring data), as well as their inherent diversity in landscape and sources of pollution. A list of the watersheds selected for the calibration of the model and the validation of the model; and a description of why certain watersheds were used for a specific task will be included in the final report.

Task 3 – Compilation and Preparation of Model Input Data

NEIWPCC and the TAC will work closely together to compile input data for the region and the calibration watersheds. Critical watershed-related data such as hydrology, land cover, soils, topography, weather, and pollutant discharges will comprise the core database for the model. NEIWPCC will be responsible for working with the TAC to locate necessary data and inserting it into the model. Specific data inputs for any given watershed may include:

- weather data (average temperature and precipitation), a minimum of 3 weather stations per watershed is ideal;
- mean monthly and/or annual nutrient loads and flows from point source discharges;
- information about water extractions;
- tile drainage systems;
- information on unpaved roads;
- septic system usage;
- animal density;
- soils information;
- land use;
- elevation data (100m resolution);
- groundwater nitrogen (based on background nitrogen estimates)
- physiographic province;
- soil phosphorus levels;
- 5+ years of consecutive water quality data (for calibration watersheds).

Dr. Evans will monitor data formatting to ensure that the data are consistent and appropriate for the model. He will also assist with the calculation of some input data and review the database to ensure that information was properly entered and formatted.

Task 4 – Calibration and Validation of the AVGWLF Model for the Northeast

Once the input database is completed, NEIWPCC and Dr. Evans will calibrate and validate the model's algorithms and verify model output (estimated nutrient and sediment loading) for the watersheds throughout the region, using historic water chemistry data. Dr. Evans is responsible for the calibration and validation process, with input from NEIWPCC and the TAC.

Task Schedule

The schedule for completion of the project is as follows:

2005

January	Project commences.
	TAC meeting for overview of AVGWLF and discussion of input data needs and calibration watersheds.
March-April	TAC teleconferences to discuss and select calibration and verification watersheds.
August	Begin compilation of input data for model. NEIWPCC to work with TAC to collect necessary data.
October	Complete compilation of input data for model. Begin model calibration and validation with Dr. Evans.

2006

January	TAC Teleconference to provide project update and discuss status of model calibration.
February	Complete beta version of the Northeast Regional AVGWLF Model.
	The draft model report is completed.
March	Draft model report is distributed to TAC members.
April	TAC Meeting to review and test beta model and discuss draft model report.
May	Comments form TAC on draft model report to NEIWPCC.
	TAC teleconference, if necessary, to discuss draft model report comments.
June	Submit compiled comments on draft model report to Dr. Evans.
July	Northeast Regional AVGWLF Model and final model report complete.
	AVGWLF Model Overview Workshop for NEIWPCC member states.

A7 Quality Objectives and Criteria for Model Inputs/Outputs

Watershed Selection Criteria

The TAC held three conference calls to discuss the eligibility of watersheds proposed for use in the calibration or verification of the model, and selected a total of twenty-seven watersheds (Table 2). The watersheds are collectively representative of the variations in landscape characteristics found in the region. They were chosen for best meeting the criteria listed below. Any watersheds that do not meet certain criteria were chosen because they have special characteristics that make them worthwhile to the model. In such cases, deviations from the criteria will be acknowledged as likely causes of deviations from expected results. The criteria are as follows:

- Watershed size should range from 20 to 500 square miles. Larger watersheds provide less accurate nutrient estimates since the model does not account for N or P attenuation from the point source. (Improvements in AVGWLF undertaken by Penn State may result in an accounting for in-stream attenuation by the time that the model is ready for calibration.)

Smaller watersheds typically aren't detailed enough to provide the model with sufficient input. To capture regional variability, the calibration watersheds should vary in size.

Table 2. Selected Watersheds for AVGWLF Calibration

State	Watershed	Sampling Location	Area (mi^2)	Comments
CT	Broad Brook	Broad Brook	15	This is a Connecticut River tributary, located upstate. Land use is primarily forest and agriculture. Significant groundwater nutrient load.
CT	Farmington River	Tariffville	577	Land use is primarily forest.
CT	Norwalk River	Winnipauk	33	A significant change of gradient exists from rural to highly developed land use as river proceeds downstream. Land use is primarily forest and development.
CT/RI	Pawcatuck River	Westerly, RI	295	Interstate watershed. Land use is primarily forest.
CT	Salmon River	East Hampton	100	Good reference watershed. This is a Connecticut River tributary, located near the coast.
MA	Squannacook River	Groton/Shirley	66	Land use is primarily forest.
MA	Sudbury River	Framingham	106	Land use is primarily developed and forest. It is the most developed of the Massachusetts calibration watersheds.
MA	Swift River	Ware/Belchertown	189	This watershed is below the Windsor Dam and the Quabbin Reservoir.
MA	Ware River	Barre	96	Land use is primarily forest. Data from both downstream and upstream gages will be utilized.
ME	Narraguagus River	Cherryfield	227	Geology includes sandy soils. Land use is primarily forest and agriculture. There are relatively few historical land use changes.
ME/NH	Saco River	Cornish, ME	1293	Land use is primarily forest. Despite its size, this watershed was chosen because it crosses state boundaries.
ME	Sheepscot River	North Whitefield	145	The watershed is located midcoast. It has a mix of forest, agriculture, and developed land. It is more developed than the other Maine watersheds.
ME	Wild River	Gilead	70	Watershed is almost completely forested. It would serve as a good reference watershed.
NH	Lamprey River	Newmarket	183	This is a coastal watershed. Land use is primarily forest, with a significant level of urbanization.
NH	Souhegan River	Merrimack	170	Land use is primarily forest, with a significant level of urbanization and some point sources.
NH	Contoocook River	West Hopkinton	150	Land use is primarily forest. The watershed area is predominantly rural. The headwater is regulated by the US Army Corps of Engineers. 85% of the watershed area is accounted for by the two streamflow gages.

NH	Ashuelot River	Keene	420	Land use is primarily forest. The watershed area is predominantly rural. The headwater is regulated by the US Army Corps of Engineers.
NY	Chemung River	Chemung	2506	This watershed is a tributary to the Susquehanna River. Despite its size, this watershed was chosen since Penn State already has the data needed to run the model.
NY	West Branch	Beerston	351	This watershed is a tributary to the Delaware River and an input to the Cannonsville Reservoir, which provides drinking water to New York City. Penn State already has the necessary input data.
NY	Dutch Hollow Brook	Owasco	30	This river flows into Owasco Lake from the east, providing 20% of the lake's water. Land use is predominantly agriculture and forest.
NY	Lake Lonely	Saratoga Springs	23	Land use is primarily forest and developed.
NY	Owasco Lake	Southeast Owasco	115	The primary inlet to this lake is at the southern end. Land use is a mixture of forest, agriculture, and development.
NY	Little Chazy River	Chazy	55	This river flows into Lake Champlain. Land use is primarily forest and agriculture.
VT	Lewis Creek	North Ferrisburg	77	This watershed originates in the Green Mountains and flows into Lake Champlain. It has undergone little change in last ten years. Land use is primarily forest and agriculture.
VT	LaPlatte River	Shelburne Falls	44	This river flows into Lake Champlain. The watershed is the most agriculture-based of the Vermont watersheds. There is also a significant level of forested land, with some recent shift toward residential use.
VT	Little Otter Creek	Ferrisburg	57	This river flows into Lake Champlain. Land use is primarily agriculture and forest, and very little residential use.
VT/NY	Poultney River	Fair Haven, VT	187	This river flows into Lake Champlain. Land use is primarily forest and agriculture.

- Watersheds should have at least one monitoring station, and a minimum of 5 years of consecutive water quality and flow data. Watersheds with year-round data are preferred; however, it is possible to work with a watershed that only has seasonal data. Watersheds should contain at least one USGS (or other long-term, continuous) flow gage that calculates mean daily flow. Daily flow data will better reflect the variability of the weather data. Table 3 indicates what flow data are available for each calibration watershed.

Table 3. Flow Data for Calibration Watersheds

State	Watershed	Flow Data Record	Data Source
CT	Broad Brook	1961-1976, 1982-2003	USGS
CT	Farmington River	1977-2003	USGS
CT	Farmington River	1928-1939, 1971-2003	USGS
CT	Norwalk River	1962-2004	USGS
CT/RI	Pawcatuck River	1940-2003	USGS
CT	Salmon River	1928-2003	USGS
MA	Squannacook River	1949-present; every other month	USGS
MA	Sudbury River	1979-present; every other month	USGS
MA	Swift River	1912-present; every other month	USGS
MA	Ware River	1928-present	USGS
ME	Narraguagus River	1948-present	USGS
ME/NH	Saco River	1916-present	USGS
ME	Sheepscot River	1938-present	USGS
ME	Wild River	1964-present	USGS
NH	Lamprey River	1934-present	USGS
NH	Souhegan River	1909-present	USGS
NH	Contoocook River	Peakflow, 1921-present. Daily 1929-1989. Daily, 1945-present.	USGS
NH	Ashuelot River	1907-present	USGS
NY	Chemung River	1903-2003	USGS
NY	West Branch	1991-present (daily)	USGS
NY	Dutch Hollow Brook	1997-present	local gaging station
NY	Lake Lonely	2003-2004	local gaging station
NY	Owasco Lake	1988-present	USGS
NY	Little Chazy River	1992-2003	USGS
VT	Lewis Creek	1992-2003	USGS
VT	LaPlatte River	1992-2003	USGS
VT	Little Otter Creek	1992-2003	USGS
VT/NY	Poultney River	1992-2003	USGS

- Watersheds should have water quality data on total nitrogen, total phosphorus, and sediments (preferably total suspended solids). Watersheds that lack data for nitrogen, phosphorus or sediments will not be discounted, because the model can be run based on any one of the parameters. Speciation of total nitrogen or total phosphorus is acceptable, as totals can be calculated. Watershed monitoring stations with a good nutrient data record, but which lack sediment data will not be discounted as potential model candidates. The species of nutrients for which data are collected will not matter, because through relationships we can calculate total N and P which is what the model actually uses. For example, data for such parameters such as TKN, NO_x, or NH_3 will suffice when calculating loads. To calculate sediment loads, total suspended solids are the preferred data type. However, turbidity can also be used, since sediment load data can be derived from it. At least 5 or 6 years of data should be available to reflect variability in weather. Table 4 indicates the type of nutrient and sediment data available and their period of record for each calibration watershed.

Table 4. Nitrogen, Phosphorus, and Sediment Data for Calibration Watersheds

State	Watershed	Data Type	Data Record	Source
CT	Broad Brook	TN, TKN, TP, TSS, TDS	1993-2003 (TSS 1997-2003)	USGS
CT	Farmington River	TN, TKN, TP, TSS, TDS	1991-2003	USGS
CT	Farmington River	TN, TKN, TP, TSS, TDS	1967-1968, 1971-2003	USGS
CT	Norwalk River	TN, TKN, TP, TSS, TDS	1980-2003 (TN 1980-1986, other N species sampled beyond 1986)	USGS
CT/RI	Pawcatuck River	TN, TKN, TP, TSS, TDS	1976-2002 (TDS 1952-2002, TN 1976-1986)	USGS
CT	Salmon River	TN, TKN, TP, TSS, TDS	TDS 1953-2003, TSS 1974-2003, TP 1971-2003, TKN 1973-2003, TN 1972-1986	USGS
MA	Squannacook River	NH3, NO2/NO3, TKN (recently), TP	1998-present	MADEP
MA	Sudbury River	NH3, NO2/NO3, TKN (recently), TP	1998-present	MADEP
MA	Swift River	NH3, NO2/NO3, TKN (recently), TP	1998-present	MADEP
MA	Ware River	NH3, NO2/NO3, TKN (recently), TP	1998-present	MADEP
ME	Narraguagus River	TN, TP, TSS	1978-1986, 1999-2001, 2004	USGS, MEDEP
ME/NH	Saco River	TN, TP, turbidity	2001-2004	Saco River Corridor Commission
ME	Sheepscot River	TN, TP, TSS	1999-2001, 2004	MEDEP
ME	Wild River	TN, TP, TSS	1968-1996	USGS
NH	Lamprey River	NH3, NO2/NO3, TKN, TP, TSS, TS	2002-present, March-December	NHDES
NH	Souhegan River	NH3, NO2/NO3, TKN, TP, TSS, TS	~1990-present, Summer	NHDES
NH	Contoocook River	NH3, NO2/NO3, TKN, TP, TSS, TS	~1990-present, Summer	NHDES
NH	Ashuelot River	NH3, NO2/NO3, TKN, TP, TSS, TS	~1990-present, Summer	NHDES
NY	Chemung River	TN, TP, TSS	1989-1998	USEPA
NY	West Branch	NH3, NO3+NO2, TKN, TP, TS, TSS	1959-1999	USEPA
NY	Dutch Hollow Brook	TP, TSS	1997-present	Cayuga County Planning Commission
NY	Lake Lonely	TP, TSS	2003-2004	NYSDEC
NY	Owasco Lake	TP, TSS	1995-present	Cayuga County Planning

				Commission
NY	Little Chazy River	TN, TP, DP, OP, TSS	1992-2003	NYSDEC
VT	Lewis Creek	TN, TP, DP, OP, TSS	1992-2003	VTDEC
VT	LaPlatte River	TN, TP, DP, OP, TSS	1992-2003	VTDEC
VT	Little Otter Creek	TN, TP, DP, OP, TSS	1992-2003	VTDEC
VT/NY	Poultney River	TN, TP, DP, OP, TSS	1992-2003	VTDEC

- Watersheds should have data regarding instantaneous flow (measured at time of sample collection). If flow data were not collected at the time of sample collection, it may be possible to calculate flow based on measurements from a nearby stream gage. Flow gage data at the outlet and in-stream water quality data are needed. At least 5 or 6 years of data should be collected to reflect variability in weather.

- Watersheds with major point source inputs are acceptable as long as they are not dominated by point source inputs.

- Watersheds should have a variety of the characteristics that are of regional interest, since the goal for selection is to best capture and represent the variable characteristics of the region. For example, the five watersheds chosen in Connecticut include two coastal watersheds and three inland watersheds, while the watersheds in New Hampshire include a high enough level of urbanization to balance the more rural watersheds chosen for Maine.

- Watersheds with few best management practices (BMPs) should be selected. If all of the chosen watersheds have many BMPs, then the selected calibration period should end just prior to BMP implementation.

- The most recent watershed data should be utilized for any calibration watershed that has experienced significant changes over the past 10 years.

Input Data Criteria

All input data acquired for the model must conform to QA/QC procedures established by the source agency (e.g., states, EPA, USGS, etc.). If necessary, the data will be converted (by NEIWPCC or Penn State) to the units and projection needed to run the model, as a single projection must be utilized for all data sets in order to run the model.

The model input dataset for weather is automatically prepared using daily climate data from nearby weather stations. As stated by Haith et al., 1992, “The GWLF model assumes an April-March ‘weather year’ similar in concept to the ‘hydrologic year’ used by the U.S. Geological Survey that begins on October 1 and ends on September 30. In this case, it is assumed that runoff events have ‘flushed out’ the previous year’s accumulated sediment by the beginning of early spring of each year. In the file, a line is required to specify the number of days in each month, and subsequent lines are used to record the average daily temperature (in degrees C) and the total amount of precipitation (in centimeters).” If one or more weather stations are within the watershed polygon, the mean daily values for temperature and precipitation are used; otherwise, the daily mean values of the two stations nearest to the perimeter are used (Evans et al. 2003).

Model Limitations

AVGWLF is meant to provide estimates of nutrient and sediment loadings from the landscape into a given watershed or waterbody. AVGWLF is not an in-stream model and therefore does not account for in-stream attenuation and sedimentation. The model can be, however, linked to other in-stream models to account for these factors. Basically, the output from AVGWLF can be used as the input for an in-stream model. For the calibration and validation phase of the model development, watersheds with significant point source inputs are not selected due to the limitations of this model.

A8 Documentation and Records

All documentation, including the QAPP, progress reports, records of monthly meetings, and final reports will be developed by NEIWPCC. Dr. Evans will also document relevant steps and decisions made with respect to the calibration process. NEIWPCC will be provided with a copy of the documentation compiled during the course of the project. A User's Guide (Evans et al., 2003) for AVGWLF is presently available and will be modified for the current AVGWLF Model. All records will be stored electronically, and available for review, on the NEIWPCC server. Paper copies of all documents will also be kept on file in the NEIWPCC office. At its discretion, NEIWPCC may seek review and comment from the TAC on draft deliverables. Copies of all documents will also be provided to the EPA Project Officer.

Section B – Measurement and Data Acquisition

B1 AVGWLF Calibration

For the purpose of testing and adapting AVGWLF to conditions in New York and New England, the AVGWLF watershed modeling application will be used to predict hydrology, sediment loads and nutrient loads for a number of watersheds within the region. In the calibration process, flow data will be calibrated first, followed by nutrient data/loads, and then the sediment data/loads. The hydrology representation must be correct before continuing with the model development. Penn State will be responsible for all of the calibration activities for this model under the guidance of NEIWPCC. This calibration process should ensure that the model will only need to be calibrated once rather than before each use.

Hydrologic simulation will be evaluated via a comparison of simulated and observed monthly flow data. To assess the correlation, or "goodness-of-fit", between observed and predicted values, the Nash-Sutcliffe statistical measure recommended by ASCE (1993) for hydrological studies will be used. With the Nash-Sutcliffe measure, an R^2 coefficient is calculated using the equation

$$R^2 = 1 - \frac{\sum (Q_o - Q_p)^2}{\sum (Q_o - Q_a)^2}$$

where: Q_o is the observed value
Q_p is the predicted value
Q_a is the average of the observed values.

Coefficient (R^2) values equal to 1 indicate a perfect fit between observed and predicted data, and R^2 values equal to 0 indicate that the model is predicting no better than using the average of the observed data. Therefore, any positive value above 0 suggests that the model has some utility, with higher values indicating better model performance.

During the calibration process, an attempt will be made to achieve the best Nash-Sutcliffe coefficient in each calibration watershed by adjusting key model parameters such as ET and groundwater recession coefficients. For the purposes of this project, watershed in which Nash-Sutcliffe values greater than 0.1 for monthly flow comparisons are obtained will be assumed to be calibrated. In these cases, additional calibrations for sediment and nutrients will subsequently be performed.

During the initial calibration step, available in-stream flow and water quality data will be used to derive observed flows and loads for each watershed against which model-simulated results will be compared. In each case, simulations will be performed for the same period in which historical stream flow and water quality sample data were compiled. Model input files for all areas will be created using the GIS-based AVGWLF modeling application that automatically assigns parameter values using the GIS data layers and default values.

In recognition of the fact that various AVGWLF routines are based on default values and algorithms developed in Pennsylvania, some effort will be expended during the calibration process to fine-tune selected default values and algorithms used to better reflect conditions in New York and New England. The primary parameters and routines to be adjusted during calibration will primarily include those that affect stream flow, nutrient and sediment loads due to upland erosion, sediment loads from streambank erosion, and background concentration of nitrogen and phosphorus in groundwater. During the calibration process, an attempt will be made to adjust these parameter values (or algorithms used to estimate these values) in a way that will achieve an overall “best fit” between the simulated and observed nutrient loads in all of the test watersheds.

Subsequent to the calibration step, AVGWLF will be run in other watersheds to validate the accuracy of the refined modeling algorithms. In this case, observed flows and loads in the validation watersheds will be compared against “un-calibrated” results.

To assess the correlation, or “goodness-of-fit,” between observed and predicted values in the calibration and validation watersheds, the Nash-Sutcliffe statistical measure recommended by ASCE (1993) for hydrological studies will be used. With the Nash-Sutcliffe measure, an R^2 coefficient will be calculated using the equation $R^2 = 1 - [\sum(Q_o - Q_p)^2 \ / \ \sum(Q_o - Q_a)^2]$, where Q_o is the observed value, Q_p is the predicted value, and Q_a is the average of the observed values.

Coefficient (R^2) values equal to 1 indicate a perfect fit between observed and predicted data, and R^2 values equal to 0 indicate that the model is predicting no better than using the average of the observed data (Evans et al., 2002). The model will be calibrated to achieve the best possible R^2 value. There is no minimum value set to determine that the model has been calibrated successfully.

Subsequent to performing individual watershed calibrations, the algorithms used within AVGWLF for estimating various GWLF model parameters will be adjusted for the purpose of obtaining "optimized" results among the calibration watersheds. AVGWLF will then be re-run on each of these watersheds to produce new results based on the adjusted (i.e., "optimized") algorithms (i.e., no additional adjustments will be made for each individual watershed). These results will again be compared statistically against observed data as described above and will be assumed to represent the "calibration" results. Subsequent to the final calibration runs, AVGWLF will then be run on the verification watersheds. Again, no attempt will be made to adjust model input parameters after the model runs. The results of the verification runs will be compared against observed data, and Nash-Sutcliffe values will be calculated.

The mean monthly, seasonal and annual results for both calibration and verification watersheds will be entered into a table such as that shown in Table 1, and the median values for each period will be calculated. Median values less than 0.1 will be assumed to represent "poor" model performance; median values between 0.1 and 0.5 will be assumed to represent "acceptable" model performance, and median values greater than 0.5 will be assumed to represent "good" model performance. In the calibration report prepared as part of the project, descriptive information and commentary will be provided to aid model users in interpreting model results.

Table 5. Summary of calculated Nash-Sutcliffe coefficients.

WATERSHED/WQN No.	N-mo	N-seas	N-ann	P-mo	P-seas	P-ann
Beech Creek (V)	0.52	0.38	0.31	0.54	0.62	0.95
Blacklick Creek (V)	-1.03	-2.36	-11.85	-0.81	-1.36	-7.08
Brodhead Creek (V)	0.21	0.26	-0.06	0.44	0.70	0.59
Casselman Creek (V)	0.80	0.92	0.94	0.59	0.91	0.92
Chartiers Creek (V)	0.40	0.46	0.55	0.69	0.75	0.44
Clarion River (V)	0.79	0.85	0.87	0.72	0.82	0.73
Clearfield Creek (C)	0.84	0.87	0.93	0.73	0.78	0.90
Codorus Creek (C)	0.77	0.80	0.89	-0.49	-0.15	0.89
Conestoga Creek (C)	0.75	0.75	0.66	0.10	0.52	0.80
Conewago Creek (V)	0.59	0.72	0.77	0.22	0.28	-0.01
Conodoguinet Creek (V)	0.81	0.83	0.90	0.66	0.82	0.77
Driftwood Branch (C)	0.73	0.81	0.66	0.74	0.84	0.62
Fishing Creek (V)	0.55	0.34	0.86	-0.99	-1.22	0.46
Juniata/Raystown Br. (C)	0.71	0.72	0.66	0.65	0.80	0.90
Kettle Creek (V)	0.50	0.54	0.26	0.64	0.73	0.74
Loyalsock Creek (C)	0.52	0.25	-1.14	0.66	0.58	0.19
Lycoming Creek (C)	0.69	0.62	0.08	0.62	0.69	0.82
Neshaminy Creek (V)	0.47	0.41	0.65	0.57	0.59	0.40

Oil Creek (V)	0.57	0.37	0.01	0.54	0.67	0.85
Penns Creek (C)	0.65	0.67	0.79	0.64	0.66	0.85
Pine Creek (C)	0.72	0.71	0.14	0.68	0.65	0.08
Redbank Creek (V)	-0.89	-2.35	-9.46	0.00	-0.30	-2.36
Schuylkill River (V)	0.76	0.70	-0.02	0.67	0.78	0.74
Sherman Creek (C)	0.71	0.65	0.68	0.64	0.50	0.62
Slippery Rock Creek (C)	-1.10	-1.72	-4.49	-0.46	-0.49	-2.13
Spring Creek (C)	0.14	-0.07	0.06	0.07	0.31	0.74
Swatara Creek (C)	0.77	0.78	0.85	0.55	0.69	0.61
Tioga River (C)	0.82	0.86	0.95	0.64	0.76	0.92
Towanda Creek (V)	0.40	0.39	0.41	0.47	0.69	0.78
Tunkhannock Creek (C)	0.82	0.88	0.63	0.46	0.40	-0.49
Yellow Breeches Creek (C)	0.70	0.52	0.83	0.71	0.72	0.67
Young Woman Creek (V)	0.70	0.60	-0.13	0.82	0.81	0.70
Median Value	0.70	0.64	0.64	0.61	0.68	0.72

C = calibration watershed
V = verification watershed

B2 Non-Direct Measurements (Data Acquisition Requirements)

To provide for a high-quality input database, data will be acquired from a variety of qualified sources, including federal and state agencies, universities, and watershed groups. Input data for the region will include: hydrology, land cover, soils, topography, weather, pollutant discharges, and other critical watershed-related characteristics that will act as the core database for the model.

Specific data inputs for any given watershed may include:

- weather data (average temperature and precipitation), a minimum of 3 weather stations per watershed is ideal;
- mean monthly and/or annual nutrient loads and flows from point source discharges;
- information about water extractions;
- tile drainage systems;
- information on unpaved roads;
- septic system usage;
- animal density;
- soils information;
- land use;
- elevation data (100m resolution);
- groundwater nitrogen (based on background nitrogen estimates)
- physiographic province;
- soil phosphorus levels;

- 10+ years of consecutive water quality data (for calibration watersheds).

Weather data – The weather data originates from the National Oceanic & Atmospheric Administration (NOAA). The data, however, may be acquired from EarthInfo (a company based in Colorado) because they package data on a CD that has an easy-to-use interface for extracting data by location. Weather data may also be obtained from the National Weather Service website, at www.ncdc.noaa.gov/oa/climate/climatedata.html.

Nitrogen, Phosphorus, and Sediment Loads – more than 5 years of load measurement data are needed for the calibration watersheds. Obtaining total nitrogen data may require obtaining constituent load data, such as for TKN, NO_2, NO_3, and NH_3. Much of the nutrient data will be obtained from USEPA and USGS (via STORET), although some will come from other sources, including state agencies. Sediment data are available from several sources, including USGS, US EPA, state agencies, as well as watershed groups. TSS should be used to derive sediment loads, but turbidity data can also be used if the former is unavailable. NEIWPCC will ensure that all data conforms to the state government level quality assurance standards.

For the following categories of input data, data will be used when available. All data will follow quality assurance standards established by the source agency.

Water Extraction Information – Each state may have these data available in a GIS database. New Hampshire, for example, has it on the Department of Environmental Services website for water extractions that are equal to or greater than 20,000 gallons per day.

Tile Drainage – Tile drainage data will be used where available. Data will originate from the Natural Resource Inventory, prepared by the Natural Resource Conservation Service (NRCS).

Unpaved Road Information – These data will be included in road datalayers obtained from each state, as available. Data have already been confirmed as available for Massachusetts, New Hampshire, and Vermont.

Septic System Usage – Information will be derived using a statewide census tract layer, which contains attribute data for the number of people served by septic systems as recorded in U.S. census of 1990, the last time the census included this data. For modeling purposes, this number is estimated based on the proportion of one or more tracts that fall within a watershed. Data may also be obtained by obtaining sewer data from states, and assuming that residential areas without sewer access use septic systems.
Data link regarding data accuracy: http://factfinder.census.gov/metadoc/stf3appc.pdf

Animal Density –County-level data are obtained from National Agricultural Statistics Service (NASS). NASS data are the official data of the USDA. Data also exists at the zip code level, although data are held for all zip code districts that have one to four farms. These data will conform to the USDA's quality assurance standards. Values are in units of AEU_Acres where 1 AEU = 1000 lbs.
Data link: http://www.nass.usda.gov/census/

Soils Information – Soils data will be obtained from a generalized statewide data layer called STATSGO. The source is the STATSGO soil mapping products developed by NRCS (Bliss and Reynold 1989). The data layer is scaled at 1:250,000. The map data are projected into an Albers Equal Area projection in meters.
Data link: ftp://ftp-fc.sc.egov.usda.gov/NCGC/products/statsgo/statsgo-user-guide.pdf

Physiographic Province – These data will be obtained from Natural Resource Conservation Service. These data are used to derive values of rainfall intensity (i.e., rainfall erosivity factor) with categories of Rain_cool and Rain_warm; and groundwater recession (typically ranging between 0.01 and 1, usually averaged to 0.1).

Land Use – The multi-resource land characterization (MRLC) data will be obtained from the states. The 2001 data layer is expected to be released later in 2005. MRLC data exists for all of the New England states and New York. Land use data used in AVGWLF will need to fit with MRLC categories. MRLC data are obtained from the EPA and conforms to the agency's quality assurance standards. 1992 NLCD data may also be used.

Elevation Data – Data should be available as part of the National Elevation Dataset. 100-meter data may be used, and should conform to USGS quality assurance standards.

Groundwater Nitrogen – These input data are derived from land use data, soils data, and information in the USGS National Water-Quality Assessment (NAWQA) report. Depending on the availability of data, the nitrogen grid for most of the study areas will likely be created using spatial relationships between land use/cover and rock type (surface geology) described in various National Water Quality Assessment (NAWQA) reports prepared by the U.S. Geological Survey in the Northeast (www.water.usgs.gov/nawqa). This process will involve an evaluation of land use/cover GIS data layers with available surface geology, physiography, and/or soils data layers. In the absence of such data for any given watershed, existing water quality data (specifically, low-flow nitrate concentration sample data) will be used to derive these grids.

Soil Phosphorus – These data are derived from background concentrations of soil phosphate depicted in a national map available in the GWLF User's Manual (Haith et al., 1992).

Point Source Information – Data will probably come from National Permit Discharge Elimination System (NPDES) permits. Permitted, rather than actual, values of nitrogen and phosphorus concentrations will be used. Data may also come from the Permit Compliance System (PCS) inventory. Information regarding total nitrogen and total phosphorus levels are available by conducting a Water Discharge Permits query. These data are published by the EPA and conform to the agency's quality assurance standards. Actual observed data, when available, may also be used.

Streams Data – Data should be available from the National Hydography Dataset. It is used to calculate stream density, which will in turn be used to derive slope length and stream bank erosion.
Data link: http://nhdgeo.usgs.gov/metadata/nhd_local.htm

Flow Data – Flow data will be obtained from USGS flow gages. Five years or more are needed for model input.
Data link: http://waterdata.usgs.gov/nwis/about

Quality Assurance/Quality Control

The success of the model will depend greatly on the quality of data collected for model input. These data will be obtained from organizations that allow for a wide range of QA/AC standards. The input data obtained for the model will be subject to the quality control standards of the organization's host-state government agency.

Some of the data are derived from QAPP-approved sampling projects, including for the Vermont and New York rivers that flow into Lake Champlain. All Massachusetts data obtained, since they span no earlier than 1998, are also QAPP-approved. Data for the selected watersheds from New Hampshire are collected using procedures specified in the state's recent QAPP-approved Ambient River Monitoring Program. Sampling data from the New York State DEC are gathered pursuant to EPA-approved protocols. All data collected by local agencies have met with the approval of their respective states.

Table 5 summarizes the input data types, their data sources, and web links that discuss the accuracy of the data.

Table 6. Data Input Types and Sources

Data	**Source**	**Links**
Weather	NOAA	http://www.ncdc.noaa.gov/oa/climate/climateinventories.html
Land Use	MRLC	Re NLCD 1992 accuracy: http://www.epa.gov/mrlc/accuracy.html Re NLCD data: http://landcover.usgs.gov/natllandcover.asp 1992 NLCD accuracy assessment: http://landcover.usgs.gov/accuracy/index.asp Bibliography: http://www.epa.gov/mrlc/pubs.html
Elevation	States	Re methodology: http://gisdata.usgs.gov/Ned/methodology.asp Re accuracy: http://gisdata.usgs.gov/Ned/accuracy.asp Re standards: http://gisdata.usgs.gov/Ned/standards.asp
Soils	NRCS	http://www.nrcs.usda.gov/technical/techtools/statsgo_db.pdf
Flow	USGS	http://waterdata.usgs.gov/nwis/about
Streams	USGS	http://nhdgeo.usgs.gov/metadata/nhd_local.htm
Groundwater Nitrogen	USGS	http://water.usgs.gov/GIS/metadata/usgswrd/XML/nawqacyc1.xml
Soil Phosphorus	States	http://water.usgs.gov/GIS/metadata/usgswrd/XML/nawqacyc1.xml
Physiographic Province	NRCS	http://www.nrcs.usda.gov/intranet/rad/data.html#nalcc
Septic System Usage	US Census	http://factfinder.census.gov/metadoc/stf3appc.pdf
Unpaved Roads	States	MA: http://www.mass.gov/mgis/mrd.htm
Point Sources	States	N/A; NPDES permit values - not actual ones - are used

Animal Density	NASS	http://www.nass.usda.gov/census/census02/censusfaqs.htm
Water Extraction	States	
Tile Drainage	NRCS	
Nitrogen Loads	USGS, USEPA, state agencies	http://waterdata.usgs.gov/nwis/about
Phosphorus Loads	USGS, USEPA, state agencies	http://waterdata.usgs.gov/nwis/about
Sediment Loads	USGS, USEPA, state agencies, volunteer groups	http://waterdata.usgs.gov/nwis/about

B3 Data Management

Most data will be obtained from state agencies, typically through the state representatives on the Technical Advisory Committee. Additional datasets will be obtained from national agencies (e.g., weather data will be obtained from NOAA). Data will be obtained by CD or email, whichever method is most convenient for NEIWPCC and the states.

A copy of all datasets will be stored in the original form in which it was received. Whenever data are reformatted, the new files will be saved in a separate folder to ensure that the original formats of the data are not lost. To ensure that "pre-formatted" data are not used as model input, the filenames of all modified datasets will include the date they were modified, along with a written description of the file and any modifications that were made. All emails regarding data information and transfer will also be saved. All data will be stored on a secure server at NEIWPCC.

To ensure correct values and units, data will be reviewed by NEIWPCC's Research Assistant for 100% of the data that are transcribed from paper and 10% of data that are obtained electronically and converted to the needed dataset format. The formulas used to perform unit conversions will be checked and reviewed by the project officer. When creating new fields for values that are converted using mathematical formulas to the units required by AVGWLF, 10% of all the data will be checked by calculator. Any errors in the data will be reported to the project officer. Upon assessment of the any reported errors, the Project Officer may require the Research Assistant to go back to the "pre-formatted data" and begin the development of a given dataset from scratch. In this scenario, the Project Officer will oversee the review of the formatted datasets to ensure their accuracy.

With the exception of data sets for each individual weather station, all final versions of datasets will be placed in a data directory. Within this directory, state-specific data will be kept in subdirectories for their respective states, and region-wide data will be kept in a regional subdirectory. Data for individual weather stations will be stored in the AVGWLF directory. Because of the large amount of data that will be collected to create the input for this model, a word document will be maintained that gives, for each dataset, the file name and a description, including any modifications made to the dataset.

AVGWLF requires daily weather datasets to be put in Excel comma-delimited format. The input data sets are automatically loaded when using the AVGWLF interface developed by Penn State. Once they are loaded, the user provides other information related to "non-spatial" model parameters. The GIS and non-spatial data are subsequently used to derive values for required model input parameters that are written to one of three different files: weather.dat, transport.dat, and nutrient.dat. A summary of categories for different data and the sources used to derive the data are given in Table 6.

Table 7. Information Sources for GWLF Model Parameterization (Evans et al. 2003)

File	Information Source
WEATHER.DAT file	Historical weather data from National Weather Service monitoring stations
TRANSPORT.DAT file	
Basin size	GIS/derived from basin boundaries
Land use/cover distribution	GIS/derived from land use/cover map
Curve numbers by source area	GIS/derived from land cover and soil maps
USLE (KLSCP) factors by source area	GIS/derived from soil, DEM, and land cover
ET cover coefficients	GIS/derived from land cover
Erosivity coefficients	GIS/ derived from physiography map
Daylight hrs. by month	Computed automatically for state
Growing season months	Input by user
Initial saturated storage	Default value of 10 cm
Initial unsaturated storage	Default value of 0 cm
Recession coefficient	Default value of 0.1
Seepage coefficient	Default value of 0
Initial snow amount (cm water)	Default value of 0
Sediment delivery ratio	GIS/based on basin size
Soil water (available water capacity)	GIS/derived from soil map
NUTRIENT.DAT file	
Dissolved N in runoff by land cover type	Default values/adjusted using AEU density
Dissolved P in runoff by land cover type	Default values/adjusted using AEU density
N/P concentrations in manure runoff	Default values/adjusted using AEU density
N/P buildup in urban areas	Default values (from GWLF Manual)
N and P point source loads	GIS/derived from NPDES point coverage
Background N/P concentrations in GW	GIS/derived from new background N map
Background P concentrations in soil	GIS/derived from soil P loading map
Background N concentrations in soil	Based on map in GWLF Manual
Months of manure spreading	Input by user
Population on septic systems	GIS/derived from census tract map
Per capita septic system loads (N/P)	Default values (from GWLF Manual)

Section C – Assessment and Oversight

C1 Assessment and Response Actions

The Project Officer and the Water Quality Modeling Research Assistant at NEIWPCC shall meet on a weekly or biweekly basis to discuss progress of the model, including the collection and format of input data. All correspondence regarding the gathering of input data and model development shall be copied to the Project Officer. Any errors found in the dataset will be documented and reported to the Project Officer. In case any backtracking is needed to locate the source of errors, all original files will be saved separately, along with each modified variant of the dataset will be saved. All data files shall be named and documented so that all project participants will understand the file contents.

Dr. Evans will be responsible for calibrating the model once the data are collected and formatted. During calibration, he will adjust the algorithms used to estimate various input parameters, including those that affect stream flow, upland erosion, sediment loads from stream bank erosion, background concentrations of nitrogen and phosphorous in groundwater, and average soil phosphorous concentration. Dr. Evans will aim to achieve an overall "best fit" between the simulated and observed nutrient loads for the calibration watersheds (Evans et al., 2002). NEIWPCC will oversee Dr. Evans during the calibration period. All model adjustments will be documented and reviewed by NEIWPCC's Project Officer.

The NEIWPCC staff shall refer to state TAC representatives and Dr. Barry Evans at Penn State concerning questions about data gathering and formatting.

NEIWPCC may implement, at their discretion, various audits or reviews of this project to assess conformance and compliance to the quality assurance project plan in accordance with NEIWPCC Quality Management Plan.

AVGWLF was adopted and approved by the EPA for use in Pennsylvania. Penn State conducts internal assessments of the model only when there are changes to the model. This project alone will not trigger any further such assessments.

The model will be calibrated to achieve the best possible R^2 value. Members of NEIWPCC and the TAC, along with researchers at Penn State, will test the model by running model simulations for the verification watersheds, in order to evaluate how well the model works in watersheds not considered in the original calibration effort.

C2 Reports to Management

Deliverables to be submitted for quality assurance purposes includes:

- QAPP, submitted by NEIWPCC
- Quarterly progress reports
- Final report, submitted by Penn State and NEIWPCC – should include input data sources and a synthesis of the calibration and validation results
- CD and/or DVD of the regional AVGWLF, submitted by Penn State
- Presentation to the TAC by Penn State – should include discussion of model development and calibration

Frequency and distribution of reports

Research Assistant to Management – will report daily on activity performed. Drafts of QAPP will be submitted daily to the project officers for review, revisions, and comments concerning issues that require further investigation.

NEIWPCC to TAC – will provide a project update and report on progress of input dataset by holding teleconferences as needed until the dataset has been fully compiled and prepared. A draft model report will be submitted to the TAC in March 2006. In April 2006, the TAC will meet at NEIWPCC to test and review the beta Northeast Regional AVGWLF model. The TAC will provide written comments regarding the beta model and submit them to NEIWPCC. A teleconference will subsequently be held in May 2006 to discuss the draft model report and to review the written comments.

Section D – Data Validation and Usability

Validation
All acquired data will be reviewed and verified for conformance to project requirements, and validated against the data quality objectives which are listed in Section A7. Only those data which are supported by appropriate quality control data and meet the data quality objectives defined for this project will be considered acceptable for use in developing the Northeast AVGWLF Model. These objectives require the best available quality control standards of the source agency and as approved by their respective state environmental agencies.

NEIWPCC, Penn State, and EPA Project Managers are responsible for ensuring that data are properly reviewed and verified for integrity, and are suitable for use in model development. Verification, validation and integrity review of data will be performed using self-assessments by the NEIWPCC and Penn State Project Managers, and TAC reviews (as appropriate).

Prior to inclusion in the input database, the data will be reviewed for accuracy, representativeness, sufficiency, and analytical quality, as discussed in section A7. Data will be checked for errors, especially errors in transcription, calculations, and data input, as discussed in section B3. Issues and/or errors that can be corrected will be corrected and documented. If an issue cannot be corrected, the research assistant will consult with the project management to establish the appropriate course of action.

Usability
The success of this project, the creation of the Northeast AVGWLF model, will be assessed in two ways: the results from the calibration and verification process and the ability of regional state agencies to utilize this tool in their NPS and TMDL programs. All input data for the model will undergo a review to ensure the quality of the data, as outlined in Section B2. Data that does not meet these standards will not be utilized in this model; this will insure that the model output will met the necessary QA/QC standards for use in both the TMDL and NPS programs.

The Northeast AVGWLF model does not spatially distribute pollutant loading areas, it aggregates these loads at a watershed level. As a result, this would not be an ideal model for looking at the pollutant loads to a small segment of a body of water. In addition, AVGWLF is not an in-stream model and therefore not does not take into account the loading and losses of nutrients and sediments in transit in a specific body of water; it only considers pollutant loading (with minimal losses) from the landscape. The output from the AVGWLF model can be used as the input data for an in-stream model, thus taking nutrient and sediment in-stream loading and losses into account.

A benefit of AVGWLF is that once the model has been calibrated, model users can “swap” coarser-scale data for finer-scale data. For example, if finer-scale landuse information is available, that information could be inserted into the model in place of the coarser-scale landuse information. The “swapping” of data sets should not effect the model calibration. AVGWLF also provides users with a tool called PRedICT which provides model users with financial and efficiency information concerning various water quality management measures.

References

ASCE Task Committee on Definition of Criteria for Evaluation of Watershed Models of the Watershed Management Committee, Irrigation and Drainage Division, 1993. Criteria for Evaluation of Watershed models. Journal of Irrigation and Drainage Engineering, Vol. 199, No. 3.

Bliss, N.B. and W.U. Reybold, 1989. Small-scale Digital Soil Maps for Interpreting Natural Resources, Journal of Soil and Water Conservation, Vol. 44, pp. 30-34.

Evans, Barry M., et al., 2003. AVGWLF Version 5.0 User's Guide, Environmental Resources Research Institute, Pennsylvania State University, University Park, PA.

Evans, Barry M., et al., 2002. A Comprehensive GIS-Based Modeling Approach for Predicting Nutrient Loads in Watersheds, Journal of Spatial Hydrology, Vol. 2, No. 2.

Evans, Barry M., 2002. Development of an Automated GIS-Based Modeling Approach to Support Regional Watershed Assessments, Ph.D. Dissertation in Soil Science, The Pennsylvania State University, University Park, PA.

Haith, D.R., R. Mandel, and R.S. Wu, 1992. GWLF: Generalized Watershed Loading Functions User's Manual, Vers. 2.0, Cornell University, Ithaca, NY.

Lee, K.Y., et al., 2000. Modeling the Hydrochemistry of the Choptank River Basin Using GWLF and Arc/INFO, Biogeochemistry, Vol. 49, No. 2, pp. 143-173.

Schneiderman, E., et al., 1998. Application of GWLF to Non-Point Source Loading Model to the NYC Catskill and Delaware System Watersheds, Study Report to EPA.