AUXILIARY PROCEDURES FOR THE AGNPS MODEL IN URBAN FRINGE WATERSHEDS

by

Eugene R. Yagow

Dissertation Submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Biological Systems Engineering

Vernon Shanholtz, Chair
James Bohland
Theo Dillaha
Saied Mostaghimi
John Perumpral
Mary Leigh Wolfe

February 28, 1997
Blacksburg, Virginia

Keywords: geographic information systems, index, model interface, nonpoint source pollution

AUXILIARY PROCEDURES FOR THE AGNPS MODEL IN URBAN FRINGE WATERSHEDS

Eugene R. Yagow

(ABSTRACT)

The Agricultural Nonpoint Source model (AGNPS) is a single-event grid-based model used for simulating runoff, sediment and nutrients from agricultural areas. This study involved using geographic information system (GIS) spatial data and functionality to improve the spatial and temporal assignment of parameter values for the AGNPS 5.0 model and incorporated methods for representing urban fringe land uses and their nonpoint source (NPS) pollution contributions in model inputs.

Auxiliary procedures for modeling with AGNPS were developed both for enhancing input into the model and for enhancing modeled output. On an event basis, one procedure automated the creation of complex-formatted AGNPS 5.0 model input files using GIS as a spatial data manager. One pair of alternative procedures were developed to automate the assignment of parameter values on an event basis. One procedure used typical average annual parameter values, and the second assigned parameter values using adaptations of existing time-dependent relationships. On a monthly basis, a sequencing procedure was created to perform multiple runs with the model for a list of storms while updating parameters for each event and aggregating monthly modeled spatial output. Another pair of alternative procedures were developed to facilitate the simulation of monthly output from AGNPS modeled events. The first of these aggregated event output for all storms in each month, while the second supplemented the aggregated output with baseflow and septic system loads.

The study area was the 6,500 ha urbanizing Bull Run watershed in northern Virginia, which was modeled as 14,621 cells. Databases were assembled and 109 selected storm events within a 16-year period were modeled using the above procedures. Event data were added together, where necessary, to correspond with observed data from composite-sampled intervals. Output from the two event parameterization procedures were compared with monitored loads calculated for 89 composite periods, while output from the two monthly simulation procedures were compared with monthly monitored data for 23 complete months.

The monitored-modeled comparisons were considered inconclusive. Evidence strongly suggested that the rainfall records from a rain gauge outside the watershed did not correspond well with monitored runoff. The average runoff produced with the AGNPS model from the 109 selected storms amounted to 40.7% of rainfall, consistent with the calculated long-term average of 38% for the Bull Run watershed.

A nonpoint source pollution index was developed to utilize monthly modeled total nitrogen, total phosphorus, and suspended sediment. Individual rating curves were developed to

separately transform loads and concentrations of each pollutant into sub-index values. The maximum sub-index from each parameter was added together and averaged for the index. The index was calculated at the watershed outlet from monitored data, and in a spatially-distributed fashion along all streams from simulated output.

Support for this work was provided in part by funding from the Virginia Division of Soil and Conservation for Project Nos. 437116, 437134, and 438132.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dr. Vernon Shanholtz, my major professor and colleague, for his support, trust and guidance throughout my graduate program. He provided me with the freedom to explore my area of research, while helping me to sharpen the focus of my work. He also has provided and/or helped to arrange funding support for me, both for many years prior to, and during, my graduate years.

The other members of my graduate committee - Jim Bohland, Theo Dillaha, Saied Mostaghimi, John Perumpral and Mary Leigh Wolfe - have graciously helped me to distill my topic from a multitude of vaguely connected ideas, to develop relevant hypotheses, and to realistically re-assess the focus and objectives of my research. Their time and assistance have been greatly appreciated.

This project, furthermore, would not have been possible without the cooperation from many individuals and agencies who graciously provided data for use with this study: Tom Grizzard from the Occoquan Water Monitoring Lab, Kurt Neumiller from the Environmental Protection Agency's (EPA) Chesapeake Bay Liaison Office, Patrick Callahan from Loudoun County, Kathy Prescott and Charles McNoldy from Prince William County, and Norman Goulet and Harry Post with the Northern Virginia Planning District Commission. Spatial data transfer and translation were greatly facilitated by Randy Marchany at Virginia Tech's Computing Center.

Site specific land use and cropping data were gathered with the help of Misty Jones in the Warrenton ASCS office and Jim Bonar from the Prince William Soil and Water Conservation District. Barry Harris and Larry Wilkinson, both with the Natural Resources Conservation Service (NRCS), provided information on predominant cropping rotations in their respective counties. Paul Welle, with the NRCS Technical Center, provided curve number table revisions.

Closer to home, in the Biological Systems Engineering Department, Barbara Crawford and Linda Altizer were crucial in helping to arrange my transitions from one project to another. And, our Department Head, John Perumpral, has provided support over the years through various tuition waivers and scholarships, which have all been greatly appreciated.

My colleagues at work - Chetan Desai, Oscar Aquilar, Fu Youtong, and Phil McClellan - have helped me in many ways with programming details, computer functionality and network access.

Dr. Walter Pirie, with the Statistical Consulting Center, was invaluable in helping me think through the analysis of my research, and in further refining my statistical analysis.

And last, but not least, I want to thank my family - Froggie, C.T. and Patrick - for their patience and support throughout the long journey to my dissertation. The list of undone projects around the house has most considerately not been shown in quite some time, but most assuredly, has grown long.

Table of Contents

List of Tables

List of Figures

1. INTRODUCTION

All models attempt to simulate real world conditions, and are simplifications of real world processes. As such, models require input data to describe the various physical parameters used by the models. Almost all parameter values are averages over space and time. By reducing the spatial and temporal increments used for averaging, a closer approximation to the complexity of the physical system should be obtained, resulting in more realistically modeled output.

Nonpoint source (NPS) models are spatial models which simulate physical processes over a user-specified area of land. By and large, historical development of most of the NPS models occurred prior to the popularization of geographic information systems (GIS). NPS models, therefore, were not developed to directly utilize GIS's efficient spatial management capabilities. Furthermore, models were generally created for research applications with little forethought about user-friendliness for other potential users of the models. Today, more and more researchers are discovering the advantages of linking models with GIS to efficiently and consistently assign parameter values to all parts of a study area, and to facilitate modeling with increased spatial resolution. GIS functionality also provides an avenue for creating user interfaces that increase the ease of model use by people other than that model's developers.

Research has produced a wide variety of relationships that help us to understand and to describe the many and various physical processes as they vary with time. Linking these existing procedures within modeling interfaces can provide the user with an in-house expert system for assigning time-variable parameters. Increases in time variability through the use of existing research algorithms and increases in spatial variability made possible with the use of GIS are both means of improving the assignment of model parameter values for modeling NPS pollution.

1.1 NPS Pollution in the Urban Fringe

NPS pollution has only been widely recognized as a source of water quality impairment since the passage of the Clean Water Act in 1972. From the end of the last century until the early 1960s, all pollution control efforts were aimed at controlling point sources - sewage and wastewater discharges. Soil loss, one of the principal components of surface runoff, was recognized as being harmful to agricultural productivity, but was not associated with water quality problems. Only after many of the municipal and industrial discharges had been regulated, with minimal improvement in water quality, was NPS pollution recognized as a major contributor to water quality problems. Advances have been made in NPS pollution modeling, but improvements are still needed for site-specific targeting of NPS pollution sources and for evaluating the impact of various land use management strategies on resultant NPS pollutant loads (Humenik et al., 1987). Although much emphasis has been placed on agricultural NPS pollution, implementation of NPS pollution controls and land use planning is also critical in urban fringe areas, where pollution prevention, the best method of future pollution reduction, can be realized at least cost (Myers et al., 1985).

1.1.1 Urban Fringe NPS Sources

Sources of NPS pollution vary among regions and states, but generally include agriculture, urban, mining, silvicultural and construction activities. For the purposes of this research, NPS pollution sources in the urban fringe focused on agricultural and rural/suburban residential activities. NPS pollution from agriculture is a function of tillage, land management and nutrient management activities. Agricultural cropland is a source of pollution from sediment, nutrients, and pesticides, with a variety of salts and minerals coming from irrigated farmland. Runoff from barnyards and feedlots is an additional source of nutrients, and also contributes organic matter, ammonia, fecal bacteria, and other microorganisms to receiving water bodies (EPA, 1984).

Rural/suburban residential NPS pollution results from increased amounts of runoff due to hydrologic modifications, fertilization of lawns, pet populations, street litter accumulation, combustion by-products, and street salting. NPS pollutants from transportation corridors include heavy metals, such as copper, lead and zinc, asbestos, PCBs, petroleum products from spills, nutrients, suspended solids, BOD, sediment and debris (EPA, 1984; WPCF, 1990). Although NPS pollution from construction sites represents only 4 to 5% of the nationwide sediment load, erosion rates from construction sites can be 10 to 20 times that of agricultural land, with runoff rates up to 100 times greater, causing severe localized problems (EPA, 1984). In rural residential and suburban areas, failing septic systems can also be a major source of NPS pollution, contributing both nutrients and bacteria (WPCF, 1990).

Rural commercial/industrial land uses contribute to NPS pollution through increased surface runoff and associated pollutants from impervious areas. Increases in imperviousness associated with commercial and industrial development produce increased runoff volumes, shorter detention times, and increases in total solids, BOD, nutrients and metals from atmospheric and transportation sources and downstream channel erosion.

In EPA's 1984 Report to Congress, some of our nation's major water pollutants which showed a worsening trend since the original Clean Water Act (CWA) in 1977 were contributed primarily by nonpoint rather than point sources. At that time, agricultural sources were identified as the most pervasive nonpoint source in every EPA region in the U.S. (EPA, 1984). Despite progress in implementation of NPS control measures, the most recent Report to Congress (1992) still cites agriculture as the leading source of water quality impairment in rivers and lakes, and urban runoff as the second leading source in lakes and estuaries (EPA, 1994).

1.1.2 NPS Characteristics

Nonpoint and point sources contribute many of the same kinds of pollutants, although these pollutants are generated in different volumes, combinations, and concentrations during different flow conditions. Pollutants from nonpoint sources are mobilized primarily during storm events, which are intermittent in nature. Surface runoff, which originates predominantly from hydrologically active land areas, is the catalyst for NPS pollution (Novotny and Chesters, 1981). Pollution episodes, therefore, will occur with lower frequency and for shorter duration from nonpoint sources than from point sources (EPA, 1984). Point source pollution is often generated on a continuous basis, and is of most critical concern during low flow conditions,

when reduced stream volumes result in increased in-stream concentrations of influent pollutants (Chesters and Schierow, 1985).

With NPS pollution, cause and effect relationships are difficult to establish between individual sources and a specific water quality problem. The monitoring data gathered since enactment of the CWA legislation have come during a time of increasing industrial and population growth. Therefore, increased contributions from these sources tend to mask improvements resulting from the implementation of NPS controls (Humenik et al., 1987). Cause and effect is further clouded by the fact that reductions in sediment loads may result in increases in naturally generated sediment and attached pollutants stored within the system, resulting in barely noticeable changes for years (EPA, 1984).

1.2 Goal and Objectives

This research utilized the single event, agricultural model, AGNPS version 5.0 (USDA-ARS-MWA, 1995), to model agricultural and urban fringe NPS pollutant loads. The overall goal was to improve the spatial and temporal distribution of parameter values within, or supplemental to, the AGNPS model, so that model output will more closely approximate actual surface runoff and pollutant loads. The following are a list of objectives for this study:

- to identify distributed methods for assigning spatial and temporal parameter values within the model,
- to incorporate methods for representing urban fringe land uses and their NPS contributions,
- to create supporting spatial data layers and attribute files,
- to develop procedures for automating the creation of AGNPS input files both for individual events and for a series of storm events,
- to evaluate output from auxiliary procedures for alternative AGNPS event parameterization and monthly simulation, and
- to outline a protocol for indexing NPS pollutants on a monthly basis.

2. LITERATURE REVIEW

2.1 The AGNPS Model

The Agricultural Nonpoint Source (AGNPS) model was developed at USDA-ARS (Young et al., 1987) as a single-event watershed model for evaluation of alternative agricultural management scenarios. AGNPS is a grid-based model, whose spatial variability is a function of cell size. AGNPS simulates loads of total and soluble N and P, soluble COD, erosion, sediment, and runoff. The model can be used to identify critical source areas within a watershed and then to allow users to view results, either graphically or in tabular form, for any cell in the watershed. The model is designed to handle smaller watersheds, where overland flow dominates, since channel processes are not considered. Version 4.03, released in September 1994, also models a variety of soluble pesticides (Young et al., 1994). Version 5.00, released in May 1995, corrected most of the errors uncovered in version 4.03, but contains no other enhancements. Version 5.00 is intended to be the final version of the event model, as development efforts are now being diverted towards the development of the continuous simulation version, ANN-AGNPS.

The model divides the watershed area into square grid cells both for parameter assignment and for subsequent model computations. Routing is performed on a per cell and per particle-size basis, for each of five particle-size classes. Computations and routing from cell to watershed outlet are performed in three loops. The first loop calculates erosion, soluble nutrient yields, overland runoff volumes, concentrations of substances in overland flow, and contributions from point sources. The second loop calculates yields from impoundments and primary cells. The third loop accumulates output from the first two loops along each flow path, and routes them to the watershed outlet, as channel transport capacity permits.

AGNPS uses many of the same sediment and nutrient algorithms as the CREAMS model (Knisel, 1980). Upland erosion is modeled with a modified USLE for single storm events (Wischmeier and Smith, 1978), transport with Bagnold's stream power equation (Bagnold, 1962), deposition with equations by Foster et al. (1981) and Lane (1982), and channel flow velocity with Manning's equation. AGNPS uses the SCS curve number method to estimate runoff, which is routed using a unit hydrograph procedure. Nutrients are modeled in both sediment and adsorbed phases, through the use of extraction coefficients and potency factors. AGNPS can accommodate both feedlot and non-feedlot point source inputs and user-defined streambank and gully erosion.

An urban component for AGNPS was developed conceptually (Lucord and Young, 1989), which models urban NPS pollutants using accumulation and washoff procedures. Urban loads are to be simulated as point sources, which are a function of land use, length of roads, % impervious surfaces, and maximum depression storage. This component will be executed in a continuous simulation mode, since, between storms, accumulation of pollutant loads occurs as a function of

days since the last runoff event. This option, therefore, was intended for incorporation into the continuous simulation version of the model when it becomes available.

Initially developed for single event simulation, AGNPS is in the process of being converted to a continuous simulation model, ANN-AGNPS (Needham and Young, 1993). ANN-AGNPS will incorporate many enhancements over the current AGNPS model, at the expense of greater input data requirements. ANN-AGNPS will operate on a daily time-step, and include provision for temporally variable parameters for many parameters and processes. Curve numbers will be evaluated as a daily variable based on soil type, crop type, management practice, tillage type and crop residue, rather than as temporal constants. Daily rainfall will be input into ANN-AGNPS either manually by the user, or generated by the stochastic weather generator, CLIGEN, currently used in several other USDA-ARS models. Evaporation, plant growth, and decay processes are all planned for the new version. RUSLE routines are included for erosion modeling, most of whose factors are updated for each runoff event. While ANN-AGNPS will require greater amounts of input than AGNPS, it will also include procedures to simplify preparation of the data input file, such as provision for initialization of baseline soil moisture parameters using WEPP procedures (Lane and Nearing, 1989). The overland flow and channel procedures also use WEPP procedures for linear storage routing. ANN-AGNPS will still have a 1900 cell limit as does version 3.65, currently in wide usage, but AGNPS version 5.00 has this limit removed by compiling the code using an extended memory compiler.

2.2 Modeling Urban Fringe Land Uses

Land use representation for agricultural fields, pervious urban areas, and construction sites are modeled in the AGNPS versions 5.00 and ANN-AGNPS models through the use of USLE or RUSLE factors, land use parameters, CN, fertilization level and availability, and gully erosion. Feedlots are modeled as point sources, with specified acreage, roofed area, animal number and type, and CN by sub-area. Septic systems can be modeled as non-feedlot point sources by specifying daily flow rates and N, P, and COD concentrations. Impervious areas can be modeled using a daily accumulation-washoff model as proposed by Lucord and Young, also under the point source option. BMPs can be simulated using most of the parameters mentioned above, as well as surface residue and timing of operations on a cell/time-step basis in ANN-AGNPS.

The ANN-AGNPS model output will include choices of daily, event, monthly, and annual summaries of water yield, peak flow rates, erosion, sediment, sediment and soluble N and P, soluble COD, heavy metals (urban areas only), and pesticides. AGNPS 5.00 includes event summaries of all of the above parameters with the exception of the heavy metals.

One of the major strengths of the AGNPS 5.00 model is its flexibility in cell division and parameterization, especially in defining the desired urban fringe land uses, and the history of successful linkage of previous versions with GIS for input file creation. This latest version, however, has a more complex data input structure than the previous cell matrix format, thus increasing the level of programming needed for linkage. Among AGNPS 5.00's many

enhancements over the 3.65 version are improved sediment detachment and transport, and a choice of variable flow rates throughout a storm by using a triangular hydrograph rather than a constant rate. Its use of extended memory compilers eliminates the previous 1900 cell limit and makes the increased computational effort required by the annualized version possible. Additionally, many soil nutrient parameters which were treated as constants in the 3.65 version, are now made available as optional parameters. AGNPS still lacks tabular links between fields and individual cells, and between land use and parameters needed to simulate various BMPs, making assignments and changes at the field level difficult to verify. The impact of these limitations, however, should be minimal with the linkage to GIS for data management of both input and output files. The AGNPS spreadsheet editor/user interface is attractive, but is limited in its analysis capabilities, and is not recommended for use with large watersheds (Baker et al., 1995). The runoff generation process is still dependent on the CN, but the improved sediment detachment and transport functions make this model more appealing. ANN-AGNPS is currently unvalidated and in a beta testing stage, but the continuous simulation capabilities should make it a good match for the type of analyses proposed with this research.

2.3 Using Models in a GIS Environment

GIS are tools to collect, store, manage and display spatially varying data. GIS are useful in reducing manual data input requirements when linked with distributed parameter models. Use of GIS data further simplifies debugging of model input files, interpretation of output files, and creation of alternative scenarios.

The Map Analysis Package (MAP; Tomlin, 1980) along with supplemental computer programs, were used to create a geographic data layer of potential sediment yield for targeting agricultural NPS pollution within Virginia's portion of the Chesapeake Bay drainage area (Hession and Shanholtz, 1988). MAP was used to create base data layers from existing maps or aerial photography. Support software was developed for derivation of additional data layers. The combined set of base and derived data layers comprised the factors used for calculating USLE soil loss and a delivery ratio. Potential sediment yield was calculated on a cell-by-cell basis through straight-forward overlay techniques using the GIS data layers. Additional data layers were created by categorizing and ranking output data into a water quality index on a cell-by-cell basis. A second prioritizing data layer, called the pollution density index, was created through normalization of the agricultural load by watershed (total potential sediment yield divided by the agricultural acreage within each watershed). All of these data layers became known collectively as the VirGIS (Virginia Geographic Information System) database (Shanholtz et al., 1987b), and the developed software was eventually incorporated into the PC-VirGIS software package (MapTech, Inc., 1994).

Automated procedures were created using Arc/INFO GIS (Environmental Systems Research Institute, Inc.; Redlands, California) to reduce manual data input for the AGNPS model (Hession et al., 1989). This study was conducted utilizing the existing VirGIS geographical database for

the State of Virginia (Shanholtz et al,. 1988a). Raster data from this database was vectorized and overlaid with a number grid according to AGNPS specifications. Relational tables were created in INFO to facilitate data input: a soils attribute table linked to soil type; a land use attribute table; a curve number look-up table based on SCS specified combinations of soil and land use attributes; and a field table relating annual historical land use to individual fields and AGNPS cell numbers. These inputs were manipulated, area-weighted for the user-specified cell size, and combined to create a partial input file for AGNPS. This partial input file contained all soil and land use related parameters, curve numbers and slopes for individual cells. Additional AGNPS inputs were obtained manually or based on specified assumptions.

A grid cell-based GIS, ERDAS (Earth Resources Data Analysis System), was used to derive model input data for the AGNPS model (Evans and Miller, 1988). Since the cell size used for the AGNPS model (1500m x 1500m) was considerably larger than the cell size used to gather data within the GIS (30m x 30m), a cross-tabulation was performed to calculate average values within each AGNPS cell using existing GIS analysis capabilities. A user interface was developed to facilitate input file creation using the output of the various GIS analyses. For channel slope calculations, the stream drainage GIS file was created by digitizing distinct stream segments, and assigning all cells in each segment the same code. This layer was then overlaid with the corresponding elevation data layer (from DEM) to compute distance and change in elevation for each stream segment, from which channel slope was computed. While most AGNPS input parameters were created directly from the digital data layers, a few, such as field slope length, practice factor and fertilization level, still required manual entry for each cell.

Needham and Vieux (1989) described the development of procedures used to operate and link PC-Arc/INFO GIS and the AGNPS model. A 5000 acre watershed in the northern lower peninsula of Michigan was used as the test site for development. Existing spatial data in the Michigan Information and Resource Inventory System (MIRIS; scale 1:24,000) and related relational databases were utilized. Soil mapping units, land use cover, and elevation were the basic spatial data coverages used. SCS curve numbers were determined from corresponding combinations of land cover and soil hydrologic groups in lookup tables. Slope and aspect were computed using the TIN module. Channel slope was calculated as a weighted average of all grid cells predominantly contained within an overlaid vector stream segment. Default values were used for channel side slopes and field slope shape. Model output was read into a relational database by cell number for display and further analysis within the GIS.

Prato et al. (1989) used p-MAP (Professional Map Analysis Package) to assemble the spatial data layers needed to calculate the Universal Soil Loss Equation (USLE) using various alternative management systems. Boolean algebraic functions were used within the GIS to overlay the data layers corresponding to each USLE factor, resulting in a soil erosion data layer for each management alternative. An optimization program was used to select the management alternative which maximized net farm income subject to specified erosion tolerance levels. Water quality was then evaluated for select storms using the AGNPS model using data layers and parameters which defined the optimal management scenario.

Interface programs for GIS were written to quantify input parameters for the ANSWERS model (Areal Nonpoint Source Watershed Environment Response System; Beasley and Huggins, 1980) for studying surface runoff and soil erosion in a European setting (de Roo et al., 1989). The Map Analysis Package, Autometric MOSS and Deltamap GIS were all used in various stages of creating the data layers for the study watersheds. Maps of input variables, stored within the geographic database, were converted to input parameters for the model. Maps of altitude, slope and aspect were computed directly from DEM within the GIS. Data layers of concavity / convexity and potential stream channels were also derived from DEM. Geostatistical interpolation techniques, such as block kriging, were used with point data to create grid based data, where digitized data for all cells was either not available, or impractical to define. This technique was used with point measurements of many of the infiltration and surface cover parameters needed by the ANSWERS model. Several management alternative scenarios were modeled and subtraction of output layers was performed within the GIS to look at areas affected by changes in management.

Rewerts and Engel (1991) describe a process to develop the ANSWERS on GRASS Project Manager, a user interface between ANSWERS and GRASS (Geographical Resources Analysis Support System; Construction Engineering Research Laboratory, Champaign, Illinois). The Manager is being designed as an interactive tool to simplify preparation of model input files. The Manager is under development and currently functions only as a means of organizing data. Data for illustrations in the report were entered manually, with links to relational databases and spatial data layers not yet implemented.

Olivieri et al. (1991) developed a set of procedures for creating AGNPS model input files from digital topographic, soils and real-time classification of Landsat TM imagery for delineating crop residue and amount. Slope and aspect were generated by ERDAS using the digital topographic data for 16.7 m grid cells. Values from individual grid cells were aggregated to a 100 m cell size, using average slope and predominant aspect from a 6 x 6 cell matrix of the smaller cells. ERDAS GIS was used to generate data layers needed for data extraction, conversion of raster data to tabular model input data, and conversion of AGNPS output files into relational files. These output relational files are then accessible for display and post-processing within the GIS. Though problems were encountered in the TM classification process, this technology will increasingly be used to quantify differences in land use, residue cover and tillage type as influenced by seasons for NPS modeling, as with AGNPS.

GRASS GIS routines were written in "C" language to integrate input/output from three cell-based NPS models: AGNPS, ANSWERS and SWAT (Engel et al., 1993). Routines were developed to predict flow directions for grid cells from digital elevation maps (DEMs), eliminate flow direction problems from DEM data, display cell flow direction, edit cell values, estimate SCS curve numbers for each cell, and develop soil property data layers from soil series data layers by accessing the SOILS-5 database (Srinivasan and Engel, 1991; Rewerts and Engel, 1991; and Engel et al., 1993). After the required data input layers were created, programmed routines were used to extract parameters and create the input data files required by any of several

different models. Once the models were run, programmed routines created relational database files linked to cell numbers for display and visualization within the GIS. This type of modular development provides flexibility in model choices, by combining selected parameters required by individual models. Comparisons were made between modeled output and 4 years of observed data, without calibration. Runoff correlated well for all three models with the observed data, though calibration, and perhaps, more resolute data, appear to be essential for sediment and nutrients.

A calibration and validation study of AGNPS was performed by Mitchell et al. (1992) on two watersheds and five nested sub-watersheds in East Central Illinois using selected events from four growing seasons. Twenty-nine rainfall events provided 94 monitored station-runoff events. After screening for events with minimal runoff, 50 station events remained. These events were rank ordered by total sediment yield. Odd-numbered events were used for calibration and even-numbered events for validation. The GIS interface developed by Srinivasan and Engel (1991) was used to develop the AGNPS input data file. Calibration runs for one of the watersheds (18.2 ha) showed that 20 x 20 m cells provided better simulation than 80 x 80 m cells. Smaller cells were not tested. Seventy-five calibration runs were conducted, using from one to six of the drainage areas and one of seven different events, along with variations in cell size, C-factor, channel and channel length, slope length, slope estimation technique, and 5-day antecedent moisture content (AMC). The most sensitive parameter from their calibration runs was shown to be AMC. The AMC boundaries suggested by their study were: AMC 1 < .12 mm, and AMC 3 > 41 mm. The validation results indicated AGNPS simulated real events poorly compared to the observed data. The study was conducted with watersheds which were small in size (1.6 - 30.$ ha), with very mild slopes (average basin slopes: 1.0 - 2.1%), indicating conditions for which the model may not be appropriate.

Digital data in both raster and vector format were used in conjunction with the USGS Precipitation Runoff Modeling System (PRMS) to partition a watershed into spatially non-contiguous hydrologic response units (HRUs) based on categories or patterns of altitude, slope, aspect, vegetation and soil (Jeton and Smith, 1993). Digital data were also used to assist in defining parameters and in calibrating the model. Cells are linked to HRU categories using relational database tables. After creation of the HRU data layer, this layer was intersected with the five original data layers to determine the frequency distribution of each attribute within each HRU, which was added to the relational database table. This procedure provided the modeler a means of disaggregating the watershed, while at the same time, being computationally efficient at the basin scale.

GRASS-GIS procedures were written in the "C" language for a variety of hydrologic, database access and generic aggregation routines, to automate data input to the continuous simulation SWAT model (Srinivasan and Arnold, 1993). The SCS curve number data layer was generated based on four data layers (hydrologic soil group, hydrologic condition, management practice and land use) using rules stipulated by the SCS Hydrology Handbook (USDA, 1972), and based on equations from Arnold et al. (1990) for estimating antecedent moisture conditions. Another

routine calculated the following stream attributes: stream length, stream slope, and cross-sectional dimensions. Length and slope are calculated directly from the DEM, while cross-sectional dimensions are calculated using an empirical exponential function of drainage area and/or interpolating between known stream dimensions at various points along the stream. Overland slope and slope length were also calculated from DEM data using a neighborhood algorithm. The unit stream power theory was used to estimate overland flow length. The programs were designed to be portable, for use in other hydrologic models with similar inputs, or for use as stand alone modules. A user interface was also created using a knowledge-based approach to create the model input file for SWAT in a step-by-step fashion. The interface uses the developed GIS routines, relational database extraction procedures, and rules appropriate to each step, to guide the user through the input file creation process. The developed interface is believed to reduce data collection and manipulation by several orders of magnitude.

GRASS WATERWORKS was developed to derive parameters needed for a regional scale watershed in Michigan to be modeled with AGNPS (He et al., 1993). WATERWORKS functions are categorized into four main types: watershed derivation, parameter interface, model interface, and output. The parameter and model interface functions guide the user in deriving all 22 parameters needed on a cell basis, using user input, database extraction or GIS spatial analysis. Three GISs were used for data layer creation; Arc/INFO, C-Map and GRASS. Twelve alternative agricultural management scenarios were analyzed on the cropland land use segments of the watershed. All other land uses were held constant. Larger cells were used for analysis, in order to accommodate the watershed area within the 1900 cell limit in version 3.65 of the model.

The TYDAC SPANS GIS was used to provide spatial data handling capabilities for the Florida Institute of Phosphate Research hydrologic model (FHM), which integrates surface water, evapo-transpiration, and groundwater components (Ross and Tara, 1993). The HSPF (Hydrologic Simulation Program - Fortran) model was used as the surface water component. FHM is a model interface to integrate the various components, and to provide users with a choice of procedures for defining individual model parameters, either by GIS data layers, or by constant default, previously used, or user-specified values. One of the main functions of GIS in FHM was to provide a standardized format for data conversion from digital data with varying projections and scales. GIS also provided a means of assimilating the vast amount of model output from a spatially distributed continuous simulation model, for qualitative and quantitative comparisons with previous conditions and other areas.

The GIS-based Watershed Simulation Model (GSWM) combines continuous time-series data with geo-referenced watershed and land use data (Rodstrom et al., 1994). GWSM links output from a continuous simulation model such as SWMM (Storm Water Management Model) or HSPF with GIS analytic routines. GWSM uses the Arc/INFO GIS, requires vector coverages of input data, and continuous simulation output for each land use type modeled. Output from these models are single land use water quality and quantity time series for each land use. Programs in the Fortran programming language have been written to combine the land use time-series along with land use and drainage area files to generate watershed loads and summary statistics. The

GIS component was used for initial characterization and selection of watersheds, and for manipulating output data for visualization, analysis and editing for alternative management scenarios.

A methodology is proposed for predicting rural municipal NPS pollution using GIS and the AGNPS model (Ramalingam and Farrell-Poe, 1995). The model is essentially a model that works parallel to AGNPS for cells designated as rural municipal. This model uses an unreported method for calculating surface runoff based on soils data and hydrological parameters extracted from a DEM, and the Yalin equation to model sediment transport. Nutrient and biological components are also to be developed and incorporated into the model.

An interface was built between the UNIX version 3.0 of the AGNPS model and the public domain GIS software, XGRASS (Park et al., 1995). The interface consists of three sets of tools for processing inputs, displaying outputs, and for creating alternative management scenarios. The management tool can be used to identify critical NPS source areas, to evaluate alternative BMPs, and to explore the impacts of land development on NPS loads. The interface requires four basic GIS map layers - elevation, soils, streams, and land use - to define the cell parameters for this earlier version of AGNPS. After comparing the model output with monitored data, parameter calibration was performed on the curve number, AMC and the peak runoff relationships.

The HU/WQ model GIS interface is designed to rapidly and consistently prepare model input files to four ARS water quality models for use by the Natural Resources Conservation Service (NRCS) as a screening tool, and in developing water quality project plans for impaired watersheds and to evaluate the impacts of alternative land management plans (Drungil et al., 1995; Geter et al., 1995). The hydrologic unit water quality (HU/WQ) project was initiated by the Soil Conservation Service (SCS) in 1991 to develop a common interface for various models used by the agency. The intended use of the resulting interface was for running models that would be used to assist SCS (now NRCS) in its emerging role in water quality planning for watersheds. Four models were selected from an extensive evaluation of available models for incorporation into this interface: two watershed-scale models - AGNPS and SWRRB-WQ, and two field-scale models. The resulting interface, HU/WQ Tools uses GIS spatial data and related attribute data to develop model input files, greatly simplifying user input of parametric data. This interface also allows compositing of field-scale data to a watershed scale, where watershed size demands that cell size be increased to accommodate the entire watershed within the allowable number of cells; and also temporal compositing to determine average dates and associated parameters for an event. The HU/WQ interface uses the GRASS GIS for spatial data management. The interface incorporates five different types of data: spatial GIS data layers, collected data on physical properties and land management practices, internal reference databases, data derived from the previously mentioned data, and run-time data, e.g. event-specific information such as rainfall amount, initial conditions, and output options. Once data are entered through the interface, it can be used to run any of the four water quality models. This allows the planner to analyze different applications and different water quality factors,

taking into account the strengths and weaknesses of the individual models. The interface also provides consistent methods for deriving parameters which are not based on direct field observations.

2.4 Validation Procedures

2.4.1 Validation in the Context of NPS Loading Models

A comparison was made between model evaluations using the established groundwater quality protocol, and those in the field of agricultural NPS pollution modeling (Shepherd and Geter, 1995). Published model evaluations of agricultural NPS models tended to be non-uniform in content, inconsistent in statistical measures used, and generally inconclusive. Another apparent difference between the two protocols was in semantics, where some terms, such as verification and validation, have very different meanings.

A NPS simulation model is considered validated for a given watershed when the comparison between simulated and observed values meets some chosen criteria over a specified range of hydrologic conditions (Thomann, 1982). Reckhow et al. (1990) refer to verification as the process of testing a model to see if it can adequately predict system response within an acceptable level of error. In an earlier publication, Reckhow and Chapra (1986) discuss the process of model testing as one of checking the degree to which a model performs under extreme conditions, and thus proving its adaptability and reliability. Validation is not seen, therefore, as a one-time process, but as an on-going evaluation to check a model's compatability with current evidence and spatial, climatological and hydrologic variations not previously tested.

NPS models consist of several different components: primarily hydrology, sediment and nutrients. As these models have developed over time, hydrologic relationships were developed first. The next generation of models added sediment generation and transport, with many relationships based on those already established in the hydrology component. The last of these primary components was that of nutrients, whose processes built on both the hydrologic and sediment relationships already tested and in common use. If modeled output is dependent on more basic components, e.g. nutrients based on sediment or runoff volumes, the basic components may need to be calibrated in ascending order from runoff to sediment to nutrients. As has been noted by many different authors, for models whose components build incrementally on other components, no component can be predicted more accurately than those components on which they are based. If they purport to, they end up simulating part of the random error inherent in the system. Better fits are expected, therefore, with the hydrology parameters, than with sediment and nutrients. If the model lends itself to calibration, components must be calibrated sequentially, first hydrology, then sediment and finally nutrients, if the model is to maintain its intended degree of physical significance.

NPS loading models attempt to model natural processes, where broad variations in spatial and temporal conditions are expected. Compared to point source loading models, where pollution

inputs are easier to quantify and locate, NPS models should be expected to predict loads with a lesser degree of accuracy. Validation of point source models is often performed by testing a hypothesis of equality between a time-series of simulated and observed concentrations. Parrish and Smith (1989) consider the hypothesis of equality too rigorous for the objective of determining NPS model validity. They reason that, in reality, predicted values can never equal the true values, so the hypothesis of equality can not hold. Validity in this sense is inconsistent with the inexact nature of modeling. If a model is to be judged valid, it must be considered so only to the extent that it is capable of producing values approximately equal or sufficiently close to true values. The essence of model validity then is defining what is meant by "reasonably close", and then testing the model according to that criterion.

In evaluating and "validating" the AGNPS and SWRRB-WQ models for a northern great lake drainage area, both quantitative and qualitative measures were used (MacAlpine et al., 1995). The quantitative measures included average error, relative error, standard error and coefficient of variation. For the low flow conditions in this watershed, AGNPS performed poorly. It was noted that AGNPS predictions of total phosphorus (TP) would be more reasonable for dry-average AMC conditions, as shown in Table 2-1. This would require modifications to the AMC boundaries, as suggested by Mitchell et al. (1992), to "calibrate" the model to better represent observed conditions. The results for one sub-basin are shown in Table 2-1 for daily mean TP. Qualitatively, the model was able to identify the same critical areas pinpointed by an earlier diagnostic study of the area. These results reinforce AGNPS's main utility as a management tool, not as a predictive one.

Table 2-1. Daily Mean Total Phosphorus Comparison (Mitchell et al., 1992)

AMC	No. of Events	Observed Load (kg)	Predicted Load (kg)
Dry	102	0.54	0.00
Average	102	0.54	1.41
Wet	102	0.54	14.55

2.4.2 Suggested Measures of Validation

Currently, there is no standard criterion, let alone standardized statistical tests, which are prescribed for judging the adequacy of a NPS loading model. Common practice with validation studies of distributed parameter models is to show bi-variate plots of observed and simulated data values, usually with regression lines included, a variety of statistical measures that vary from reviewer to reviewer, and a qualitative assessment of the results, based on the experience of the modeler, usually without a firm statement of validation confirmation or rejection (Young et al., 1989; Dillaha and Beasley, 1983; Storm et al., 1988; Mitchell et al., 1992).

Visual comparison of modeled and observed data, and experience-based judgment on the part of the modeler have been deemed important by researchers for assessing model validity and applicability for use in the decision making process. These same researchers cite the need to

strike a balance between objective and subjective methods of evaluation, while noting that no single method of statistical evaluation alone is sufficient to judge model adequacy (Pennell et al., 1990; Thomann, 1982; Reckhow et al., 1990; James and Burges, 1982; Loague and Green, 1991). Statistical measures are used to provide objectivity for evaluating aspects of model adequacy, though, they too may be based on somewhat subjective assumptions of data and modeling conditions. Reckhow et al. (1990) state that good model verification should combine the conventional judgment-based approach with one or more statistical goodness-of-fit tests.

Pennell et al. (1990) state that graphical analysis allows for identification of general trends in the data, systematic errors, and other potential sources of error, such as outliers. Objective functions provide a numerical measure of the agreement between measured and predicted data.

Thomann (1982) suggests the following statistical comparisons as criteria in judging the adequacy of a model: hypothesis testing of the slope and intercept of the regression line, mean relative error, root mean square (RMS) error, and a comparison of the means of the observed and simulated data. He suggests using regression analysis in conjunction with tests of slope and intercept of the regression line, along with the standard error of estimate to properly interpret correlation results.

Reckhow and Chapra (1986) list measures of error, the t-test, the non-parametric Mann-Whitney-Wilcoxon test, regression analysis, cross-correlation and box plots as appropriate statistical methods for confirmation of deterministic models. Measures of error include mean relative error and mean squared error. Reckhow et al. (1990) also recommend various combinations of graphic and statistical procedures based on the proposed analysis and intended use of the modeling results. Their graphical recommendations include bi-variate plots, histograms, and box plots.

Loague and Green (1991) likewise, suggest the use of both statistical and graphical measures for evaluation. Model performance can be compared using either summary statistics (mean, range, standard deviation) or using individual observed vs. predicted pairs of data, which can also be displayed in both statistical and graphical forms. Evaluation of data pairs usually proceeds with an analysis of the residual errors in the forms of maximum error, root mean square error, modeling efficiency and the coefficients of determination and residual mass (James and Burges, 1982; Green and Stephenson, 1986). These tests can be performed with either sorted or unsorted data, analyses with unsorted data being more rigorous. Graphical displays suggested include 1) comparison of observed and predicted values; 2) comparison of ranges, medians and means of integrated values; 3) comparison of matched predicted and observed time-series values and/or residuals; 4) comparison of cumulative integrated values; and 5) cumulative frequency distributions. Choice of validation techniques should also be based on the modeled output parameter, e.g. error in cumulative loads may be more appropriate for seasonal or annual loads than on an event basis.

Zacharias and Coakley (1993) categorize validation techniques into three main categories: summary statistics, hypothesis testing, and measures of goodness-of-fit. They list examples of summary statistics as the mean, standard deviation, and those statistics commonly used with box or whisker plots: range, interquartile range, and median. Examples of goodness-of-fit include maximum error, the normalized root mean square error, the coefficient of determination, modeling efficiency and the coefficient of residual mass. Hypothesis testing is a more formal approach to validation where either summary statistics or goodness-of-fit are tested against some prescribed criteria. Where the data distribution is either unknown or does not conform with a normal distribution, additional robust measures are suggested: the normalized median absolute error, the robust coefficient of determination, and the robust modeling efficiency, which substitute median values for mean values in their formulation.

Although regression analysis is commonly used in NPS model evaluation, discussions with a local statistician indicate this may be an inappropriate use. Regression analysis was designed to test the impact of an input parameter on an output parameter (Pirie, 1996).

2.4.3 Criteria Used with Validation Measures

Traditional hypothesis testing, as with point source models, has focused on evaluating the significance of the mean and standard deviation at the 0.05 probability level. Criteria suggested as being practical by Anderson (1942, In: James and Burges, 1982) include getting means to within 5%, standard deviation between 5 and 10%, and the lag-one correlation tendency to 0. As a good calibration is approached, lag-one is normally distributed with mean 0, and variance, (N-2)/(N-1) squared. James and Burges suggest that overall model quality is best reflected by the mean and standard deviation of the relative error, model efficiency and the coefficient of the residual mass. Since models tend to perform unevenly over the entire range of flow conditions, they suggest calculating several coefficients at low, average and high (or dry, average and wet) conditions. Criteria used in a sample calculation included coefficients of determination, efficiency and residual mass > 0.97, mean daily flow/yr within 5%, standard deviation of daily flows within 15% and individual monthly volumes within 10% of observed values.

A qualitative evaluation of expected relative errors from various NPS modeling scenarios, created by CH2M Hill, Inc. (1990), listed the following magnitude of errors for the given deterministic modeling components:

	------Relative Error (%)------		
	Event	Continuous	Measurement
Annual Runoff Volume	40	20	10
Pollutant Runoff	100	25	
Management Effects	100	50	
Stream Processes	50	25	
Pollutant Load	250	100	

Dillaha (1990) states that good hydrologic assessment models should predict observed values within a factor of 2, where parameters are measured on site, or where the model is calibrated; and within a factor of 10 otherwise. Zacharias and Heatwole (1993) used a criterion that model predictions be within a factor of 2 of observed pesticide concentrations at sampled dates and depths. They further evaluated model performance over the entire study period, by looking at the percentage of date-depth pairs meeting their criterion. These positive indicators could also be broken down by individual depths and, on a temporal basis, either monthly or seasonally. Bouraoui (1994) used a factor-of-2 validation criteria with long-term cumulative measured values in testing his continuous version of the ANSWERS model. Fu (1994) used the following goodness-of-fit statistics: graphic comparison, relative error, model efficiency, linear regression, hypothesis testing, and comparison of hydrograph shapes. His validation criteria included peak discharge rates within 15%, total runoff volumes within 20%, and time to peak within 15 minutes of observed data.

2.4.4 Validation Studies with the AGNPS Model

AGNPS was developed for agricultural watersheds from a few hectares up to 20,000 ha. The model has undergone preliminary testing for the runoff components with data from 20 different watersheds in the north central U.S. (Young et al., 1989). The study showed relatively good correlation of peak flows with a coefficient of determination of 0.81. Results were shown as plots of observed vs. simulated data, overlaid with linear regression lines. Qualitative assessments were also used, e.g. “sediment yield compared favorably”, “AGNPS provided realistic estimations of nutrient concentrations in runoff”. Additional testing was conducted on the sediment component with data from 2 experimental watersheds near Treynor, Iowa and another near Hastings, Nebraska . The observed vs. simulated plot showing this data was noted as a favorable comparison, but was dominated by two outliers. Additional testing showed a favorable comparison with several other models on three different types of watersheds in Mississippi (Bingner et al., 1989). Chemical components of AGNPS received basic testing during development of the CREAMS model, for which they were originally developed. Chemical data from watersheds is not widely available, but data from several measurable rainfall-runoff events on seven different watersheds in Minnesota over a three year period were used to test the model. The storms during this period were small, however, with relatively little runoff. The model provided realistic results compared with these smaller events (Young et al., 1989). AGNPS has been used to prioritize watersheds and to pinpoint critical areas within a watershed for potential severity of water quality problems in several states, and for evaluation of alternative management plans. AGNPS was used in the Garvin Brook and the Salmonson Creek watersheds in Minnesota to pinpoint critical sediment source areas and to prioritize critical areas for BMP treatment, using a 25-year, 24-hour design storm.

Hession et al. (1989) evaluated the AGNPS model (v. 3.51) for its reliability in assessing BMP effectiveness on a monitored watershed, by comparing pre-BMP, post-BMP and 100% forested conditions. VirGIS data were converted to Arc/INFO coverages and used to facilitate the creation of AGNPS input files. Evaluation was not performed on specific storm events, but using design storms ranging from 1 to 6 inches for 1, 2, 5 and 10 year events. Input parameters

were, therefore, selected to reflect average conditions. The 1157 ha watershed contained five livestock producers, with only one using an animal waste storage facility, and excessive field applications of manure for pre-BMP conditions. Post-BMP included waste facilities for all animal operations and nutrient management plans to reduce fertilization to recommended application rates. No comparison was made with monitored data. Simulation showed that the state program's 40% nutrient reduction goal could be met with full implementation of BMPs.

Mitchell et al. (1992) used the AGNPS model to evaluate runoff and sediment delivery predictions from small watersheds with mild topography. Fifty sediment yield events were monitored from two predominantly row-cropped watersheds and five nested sub-watersheds in East Central Illinois over a period of four growing seasons. A series of calibration runs was performed before attempting to validate AGNPS for these small watersheds with mild topography. Grid cell size, C-factors (calculated with RUSLE), inclusion or exclusion of channel segments, slope length, slope estimation algorithms and antecedent moisture content (AMC) were the parameters used for calibration. Half of the events were used to calibrate and half to test the AGNPS model. Average calibrated input parameters were used for all events. Input data were facilitated through the use of GRASS GIS. Management practices, nutrient levels, fertilizer incorporation levels and the USLE C-factor were estimated for each storm date. No validation criteria were defined, but observed and simulated data plots were presented with regression lines. The simulation results were reported as "poor", with annual runoff varying from 65% to 151% of observed, and total annual sediment yield varying from 29% to 557% of observed values. The applicability of AGNPS on small watersheds (1-30 ha) with mild slopes (1-2%) was deemed questionable.

A post-audit verification of the AGNPS model (v.2.52) was performed by Clausen (1993) on two watersheds in northwestern Vermont. Both watersheds were used predominantly for dairy agriculture and had received extensive implementation of BMPs accompanied by comprehensive water quality monitoring. The LaPlatte River watershed was 400 acres in size, and the Jewett Brook watershed was 1,384 ha in size. For both watersheds, AGNPS predictions were compared to monitored values of storm flow depth, peak flow, and the concentrations and mass loadings of sediment, nitrogen and phosphorus. A total of 15 storms were selected between late March and late November with 3 or more consecutive 8-hr composite runoff samples during the 5 years of record. Winter storms were ignored since AGNPS does not predict snowmelt runoff. Calibration was performed, since AGNPS was developed for Minnesota conditions, which varied considerably from those in Vermont. Calibration parameters were unspecified. Statistics used in the verification process included comparison of observed and predicted means, relative error, use of the t-test, linear regression, the coefficient of determination, RMS error, and the F-statistic, some of which are reported in Table 2-2. The hypothesis that the mean observed and predicted values were not different was rejected. Peak flow and phosphorus concentration were the only variables with low relative errors, although regression found no significant relationships for any of the variables. For the smaller watershed, discharge was generally lower than observed, while both sediment and nitrogen concentrations were greatly overpredicted along with their mass export. No real trend was observed with phosphorus variables, except for

grossly overpredicted mass export for storms greater than 1.00" on the smaller watershed. Since this study was the first validation study of AGNPS utilizing a wide range of statistical tests as recommended by Thomann (1982), the utility of each selected test was unknown. Based on the analysis of this study, comparisons of load means and RMS errors were found to be inappropriate measures, since the discrepancies between predicted and observed discharges were so large. The planned use of the two-sample Kolmogorov-Smirnov test was omitted since no information could really be gained over that obtained from the poor results of the previous statistics. Clausen's recommendations for NPS model validation statistics included linear regression in conjunction with an analysis of variance (ANOVA) to check regression significance, the coefficient of determination (R^2), and tests of the slope and intercept of the regression line.

Table 2-2. Validation Statistics: Mean Observed vs. AGNPS Predicted Values for 15 Storms (Claussen, 1993)

Parameter	Means		Relative Error	t-value	Prob > t	R^2
	Observed	Predicted				
runoff volume (in)	0.31	0.04	87	4.86	0.003	0.03
sediment load (lb/ac)	0.53	124.6	23,409	-1.11	0.290	0.20
nitrogen load (lb/ac)	0.06	1.09	1,717	-1.65	0.128	0.14
phosphorus load (lb/ac)	0.01	0.54	5,300	-1.64	0.125	0.03

The current team of model developers at USDA-ARS in Morris, Minnesota, who are responsible for developing the continuous simulation model, ANN-AGNPS, was chosen to conduct a verification of the AGNPS 4.03 model, in response to reported errors by model users, and since the original model developers were no longer available (Baker et al., 1995; Witte et al., 1995). A verification protocol was defined for the AGNPS 4.03 model, and used to find and correct coding errors. These corrections were incorporated into AGNPS version 5.00, as were software modifications to simplify debugging and for tracking specific parameters. No changes were made to the intended algorithms between these two versions, only corrections. During the verification procedure, a limited series of tests were run on a four cell (2x2) watershed with 100 acre cells, so that model calculations could be checked against hand calculations. These tests were conducted on most of the hydrologic, sediment and chemical calculations. Additionally, a partial verification was conducted on the sediment trap feature by the verification team. All computer-calculated values agreed with corresponding hand-calculated values. Several additional errors were noted, but were left uncorrected in AGNPS version 5.00, because of time constraints. One such flaw was in the algorithm used to create the composite hydrograph. This results in a consistent overprediction of about 10% of sediment transport capacity and sediment yield , and is most pronounced for primary cells. A second flaw noted was the assumption of no deposition of sheet and rill erosion in the fields. These flaws are to be corrected in the continuous simulation version of AGNPS.

2.5 The Use of Indexes

Attempts to categorize water according to its degree of purity date back to the mid-nineteenth century (Landwehr, 1974). The first numerical scale used to quantify the measure of water quality in an index was developed by Horton in 1965. Since then, more than 20 different water quality indexes have been published. The indexes have served as valuable aids in communicating information and in evolving water resource policy. A useful index synthesizes data such as analytical results by means of a simple quality vector. It is easily and readily interpretable. It communicates some relative quality of interest, and instills confidence in the reliability of the parameters used in the index (Couillard and Lefebvre, 1985). Water quality indexes allow meaningful spatial and temporal comparisons to be made, and integrate effects of a variety of pollutants in a simple, objective and reproducible manner (House and Newsome, 1989). A water quality index can be a unique symbol or a combination of numerical or alphanumerical variables, and may include secondary symbols which provide qualifying information such as parameter reliability, exceedance of standards (Steinhart et al., 1982) or availability of correlating monitoring data (DSWC, 1993).

Most indexes are created from a number of different measurable or observable parameters, which are scored against a rating curve. Rating curves are different for each parameter, and establish a relationship between a range of expected parameter values and unitless sub-index values. The shape of the rating curve may also be influenced by applicable environmental standards, criteria or objectives. Individual parameter or sub-index scores are then usually weighted and aggregated, using one of a variety of different methods, into a final index value. Differential weighting of sub-index scores is used to assign relative importance among the parameters. Aggregation methods consolidate and scale the weighted individual sub-index scores into an overall index value (Couillard and Lefebvre, 1985). Rating curves and parameter weighting schemes have been developed using subjective techniques, such as the Delphi consensus approach (Brown et al., 1970), statistical techniques, such as factor analysis (Joung et al., 1979), and nonparametric classification procedures (Harkins, 1974).

Indexes have been used as a method both to communicate technical consensus and to simplify technical information for policy-makers and the general public (Ball and Church, 1980). In the U.S., approximately 20% of all water pollution control agencies work with some type of water quality index (Couillard and Lefebvre, 1985).

2.6 Water Quality Indexes

Horton (1965) developed one of the first indexes for use in describing water quality and changes in water quality in comparative terms. Horton included ten parameters in his index, which included one demographic factor, eight measurements of physical, chemical and biological parameters, and one visual assessment. Each parameter was assigned a rating between 0 and 100 based on a comparison between a measurement or observation and suggested rating scales for each parameter, except for temperature and "obvious pollution". Horton excluded toxicity

from his index, since water containing any toxics was deemed to be impaired, and the index invalid under these conditions. The sources of data for the ranges and breakpoints used in the rating curves were not given. The values used were for illustrative purposes only, though spatial or temporal comparisons are still valid. Arbitrary weighting factors were also used to show relative importance of the various parameters. Horton's index is a simple additive formula, summing the weighted scores of each parameter, then dividing by the sum of the weights to constrain the final score within the range of 0 to 100. The formula included two factors which were used to adjust for temperature and "obvious pollution" conditions.

Brown et al. (1970, 1972) developed a water quality index through the use of systematic opinion survey techniques with a large group of water quality management personnel. Through a series of mailed questionnaires, the group's collective judgment was used in ranking important water quality parameters, from which the surveyors arbitrarily chose the top eleven. The group then defined rating curves to relate individual parameter measurements to water quality and to assign relative weights to each parameter. Special procedures were suggested for pesticides and toxics, such that cumulative concentrations of pesticides greater than 100 ppb, or toxic levels exceeding maximum permissible levels in the Drinking Water Standards of 1962 would automatically equate the index with zero, the lowest value on the water quality scale. An arithmetic weighted mean index, WQIA, was then defined summing the products of the individual parameter scores, q_i , times their weights, w_i ,

$$\text{WQIA} = \sum w_i * q_i \qquad (2.1)$$

where WQIA and q_i range from 0 to 100, and $\sum w_i = 1$. The higher the WQIA score, the better the water quality. Some concerns were raised by survey respondents that the index should address a specific water use rather than overall water quality. The selection of parameters was considered the most crucial aspect of model development, both in reducing redundancy between factors, and in minimizing the number of parameters to those measuring the most common denominators of water quality. Minimization of parameters was effected by choosing parameters which were regularly used in water quality assessment, and for which data were readily available. These restrictions were necessary to minimize expense in data collection and to ensure widespread applicability. The WQIA is composed of nine parameters: dissolved oxygen, fecal coliform, pH, BOD_5, nitrate, phosphate, temperature, turbidity, and total solids. In 1974, work by Landwehr altered the index formulation to a multiplicative one using the original parameters with the following form:

$$\text{WQIM} = \prod T_i * p_i * a_i \qquad (2.2)$$

where *T* is the transform function, *p* is the measured value of a parameter, and *a* is the individual parameter weight.

An index developed by Prati et al. (1971) indicated the relative level of water pollution and was used for inventorying the quality of water resources in a region or country. This index incorporated 12 commonly measured parameters: pH, DO, BOD_5, COD, suspended solids, NH_3, NO_3, CL, iron, manganese, ABS, and CCE. The index was constructed so that a higher index indicated a higher polluting effect, in contrast to most water quality indices, where a higher value indicates higher water quality. Standards for comparison of each parameter were based on water quality classifications from various published world-wide sources. Transformation algorithms were constructed so that parameter values associated with each of the five classes increased in an approximately geometrical progression. An arithmetic mean with equal weighting was used to arrive at an overall index of pollution.

Harkins (1974) used a distribution-free statistical procedure in the development of his objective index. This procedure is an application of Kendall's non-parametric multivariate ranking scheme, which can be used with any number of parameters for comparison between sampling points either spatially or temporally. A critical minimum or maximum value, such as a water quality standard, is chosen for each parameter. Then all values for each parameter, along with its control are ranked, a rank variance calculated, and an index, S, is calculated for each observation by summing the normalized rankings for each parameter, represented by:

$$S_j = \frac{\sum_{i=1}^{n} \left(R_{ij} - C_i\right)^2}{V_i} \qquad (2.3)$$

where *R* is the parameter ranking, *C* is the parameter control value, *V* is the rank variance, *n* is the number of parameters, and *j* is the specific observation. Comparisons using this method are only valid between observations within the given data set. When additional data are added to the set, rankings, rank variances and indexes must all be recalculated. The method is intended for summarizing large quantities of data for trend assessment. While this index will allow comparison between relative rank, it has the drawback that when the data set changes and recalculation is performed, the rank score of any individual observation is subject to change.

A joint undertaking by the MITRE Corporation and the EPA Office of Water Programs Operations led to the development of three indexes (Truett et al., 1975): a water quality index, a needs index for pollution abatement, and a management planning index. The index of concern here is the water quality index, termed the PDI index, which stands for prevalence, duration and intensity of pollution. Prevalence (P) refers to the number of stream miles within a planning area not in compliance with some established criteria of water quality. Duration (D) reflects the number of quarter-year periods when violations have occurred. Intensity (I) is an indicator of the severity of pollution as assessed in terms of ecological, utilitarian and aesthetic effects. The PDI is calculated as the product of P, D and I ratings divided by the total number of stream miles (M) within a planning area, as follows:

$$\text{PDI} = \frac{P * D * I}{M} \qquad (2.4)$$

where PDI, P/M and I range between 0 and 1, and D ranges between 0.4 and 1. This index was first computed on a widespread basis by various regional EPA offices in 1971. Parts of this index are currently used by the Commonwealth of Virginia's Department of Environmental Quality (DEQ) as part of its biennial 305(b) reports.

Dunnette (1979) developed a water quality index similar to Landwehr's arithmetic formulation, with a summation of transformed parameter values times weights, with the WQI ranging between 10 and 100. Parameters were selected using a variety of criteria. Initially, a wide variety of parameters used in previously proposed water quality indices were considered and then limited based on various rejection criteria. This list was then further reduced based on an importance ranking by state DEQ personnel. The number of parameters was then further reduced to six, representing the common water quality impairment categories: oxygen status, eutrophication, health aspects, physical characteristics, and dissolved substances. Weights were calculated based on a statistical analysis of relative importance values assigned by state DEQ personnel. The six parameters selected were percent saturation of DO, BOD, nitrate and ammonia nitrogen, total solids, pH and fecal coliform count. Dunnette (1979) used a logarithmic transform for each parameter with the slope of each relationship determined by assigning index values based on historical arithmetic means and no load or impact conditions, e.g. %DO Sat. =100, pH=7, TS=background load, BOD=0. Dunnette (1979) used an approach similar to Brown et al. (1972) in his treatment of toxicity and radioactivity. Samples including any indication of these substances were not given an index score, but were flagged and footnoted separately. Dunnette (1979) states "as long as appropriate parameters are selected from each of the major impairment categories and are widely accepted in water quality monitoring, then parameter selection and weighting or transformation to common index units are not critical".

A generalized water quality index was developed by Joung et al. (1979) based on multivariate factor analysis. Factor analysis is a technique which can be used to identify and quantify underlying patterns of variation in a data set. Ten common water quality parameters were considered in development of the index. Factor analysis was used both to select the most significant parameters, and as a basis for calculation of relative weights. The analysis resulted in a set of five parameters selected for the index: dissolved oxygen deficit percentage (DODP), temperature, BOD, total phosphorus and electrical conductivity. Each of these parameters was not significantly correlated with each other. Rating equations were statistically developed for each parameter using polynomial regression analysis between original parameter concentrations and literature importance ratings. The assigned weight for each selected parameter was its corresponding correlation coefficient. The guidelines used for index construction included: that the index be conceptually simple and easy to compute, that it be significantly influenced by each of the selected parameters, and that it be applicable over varying geographic areas. Factor analysis indicated that only 70% of the water quality index could be explained by the selected parameters. The authors acknowledged the omission of certain unspecified water quality parameters on the basis that they were currently not widely used. The sensitivity of individual parameters also needs further research to explicitly define minimum and maximum levels of acceptable water quality.

Steinhart et al. (1982) developed an environmental quality index for the near shore waters of the Great Lakes. This index was based on nine physical (P), chemical (C), biological (B) and toxic (T) substance variables. Raw monitored data were converted to sub-index values by mathematically defined functions based on national or international objectives. Sub-index values were multiplied by weighting factors, summed, and scaled to a final index score ranging from 0 (worst quality) to 100 (best quality). Letters with subscripts following the index score indicated the types and numbers of variables whose values exceeded some objective-specified limit, e.g. $70C_1P_1$ indicated moderately good water quality with one chemical and one physical variable exceeding specified limits.

House and Newsome (1989) report on the development of four indexes developed for use by operational managers in the U.K., both for planning and for day-to-day management of surface water quality. These indexes were designed to be used either independently or in combination. The first index reflects the water quality in general, relates nine routinely monitored physical, chemical, and biological parameters to a variety of potential uses, and is called the general water quality index (WQI). Its parameters include DO, NH_4-N, BOD, suspended solids, NO_3-N, pH, temperature, chlorides, and total coliforms. The second index, the potable water supply index (PWSI) relates water quality to suitability for use as a potable water supply using the WQI parameters plus sulfates, fluorides, color and dissolved iron. The last two indexes, the aquatic toxicity index (ATI) and the potable sapidity index (PSI), are based upon toxic parameters and are use-related; the ATI for protection of fish and wildlife, and the PSI for public water supplies. ATI and PSI are based upon nine and twelve parameters, respectively, which consist of heavy metals, pesticides and hydrocarbons. Rating curves were based upon published European water quality standards. Weightings were determined by expert consensus for the WQI and PWSI, but were considered inappropriate for use with the toxic indexes. The aggregation method used was a modified arithmetic formulation. Lower ratings of either of the toxic indexes would over-ride values of the WQI or PWSI, when used in combination. In a comparison with a national rivers classification scheme, the authors found that their indexes allowed for more precise location of a water body within a class, and allowed for better quantification of water quality deterioration or improvement. Interpretation of the indexes by experienced personnel was still necessary to put scores in perspective, and to identify exceptions and requirements not included in the logic of the indexes. The indexes, by and large, were found to be very useful for reducing large quantities of data to a single index in a rapid, reproducible and objective fashion.

Growth limiting factors for algae have been developed using several different aggregation methods which may be applicable to the index development process. The four major approaches include: a multiplicative approach where each of a variety of factors ranging from 0-1 are multiplied together, a minimum formulation which chooses the minimum value from a range of various factor values, a harmonic mean formulation which combines the reciprocal of each limiting factor, and an arithmetic mean formulation which takes the average of all of the factors. The multiplicative approach assumes that several nutrients in short supply will more severely limit growth than a single nutrient in short supply. However, unless the number of nutrients or factors is kept constant, the severity of reduction will increase with the number of

factors chosen. The minimum approach assumes that the nutrient in shortest supply controls growth. The harmonic mean is based on an electronic analogy of resistors in series. It includes some interaction between factors but is not as severely limiting as the multiplicative approach. The arithmetic mean is rarely used since it reduces the influence of a critical nutrient, if all other factors are relatively higher (EPA, 1985).

Two basic problems have been identified with the use of indices: ambiguity and eclipsing (Ott, 1978). Ambiguity is the situation where none of the individual sub-indexes exhibit a problem, but collectively, the index indicates a problem. Eclipsing, on the other hand, occurs when one sub-index indicates a problem, while the overall index does not indicate a problem. To further identify situations where these conditions occur, index structures can be broken down into two basic types:

- those in which index numbers increase with the degree of pollution (increasing scale indexes), and
- those in which index numbers decrease with the degree of pollution (decreasing scale indexes).

An analysis of aggregation functions and index structures show that additive index structures tend to exhibit ambiguity in increasing scale indices and eclipsing in decreasing scale indices. Some combinations of aggregation function and index structure are not applicable, such as the multiplication and minimum operator aggregation functions with increasing scale indices, and the maximum operator aggregation function with decreasing scale indices. Those index forms which avoid the problems of both eclipsing and ambiguity are the maximum operator with increasing scale indices, and the weighted-product multiplicative and minimum operator aggregation functions with decreasing scale indices.

van Vuuren et al. (1994) present an argument to show that reductions in NPS pollution do not always correspond to water quality improvements, contrary to popular assumptions. The authors create scenarios where water quality impairment results from standards exceedances from a differing number of NPS pollutants. The illustration analyzes water quality using a multi-pollutant water quality index based on the minimum operator aggregation function. An index is suggested as a reasonable approach to use, since water quality represents a degree of fitness or unfitness, rather than an either-or proposition. The water quality index is conceived of as a step function, broken down into several classes over the range of 0-100. The use of the minimum operator implies that the pollutant with the highest relative loading is the most immediate limiting factor for improving water quality. If reduction in a different pollutant occurs, no improvement will be seen in water quality. By the same reasoning, improvement in water quality can only occur when the load of the limiting pollutant is first reduced. The improvement, however, may not be in direct proportion, or over the whole range of limiting pollutant reduction. As the loading of one pollutant is reduced, another pollutant may then become the limiting factor.

2.7 NPS Pollutant Indexes

Nonpoint source (NPS) pollutant indexes are relatively new in comparison with water quality indexes, mainly due to the comparatively recent origin of the whole concept of NPS pollution. NPS indexes are distinctly different from water quality (WQ) indexes in several ways. First, measurements included in the rating procedures tend to be in terms of pollutant loading rather than in terms of concentrations. Second, the procedures used to arrive at a final NPS index are not as straightforward as with WQ indexes, which tend to have more precise mathematical formulations. NPS indexes also appear to be more concerned with rank order than with comparison to some established threshold, and weighting procedures are rarely used explicitly.

Though not explicitly an index, the Universal Soil Loss Equation (USLE; Wischmeier and Smith, 1978) has the form of an unweighted multiplicative index, with each factor a sub-index, representative of some characteristic which strongly influences soil loss rates. The USLE is an empirical equation, calibrated from many years of plot studies for estimation of average annual soil loss (A) usually expressed in tons/acre. The USLE is given as:

$$A = R * K * LS * C * P \tag{2.5}$$

where R is the rainfall erosivity factor, K is the soil erodibility factor, LS is the slope-length factor, C is the vegetative cover factor, and P is the supporting practices factor. A variation on the USLE is the unitless erodibility index (EI) based on the work of Heimlich and Bills (1984). The EI was used as the basis for determining highly erodible land in conjunction with the 1985 Farm Bill. EI was calculated as:

$$EI = \frac{R * K * LS}{T} \tag{2.6}$$

where T is allowable soil loss, a characteristic of soil type, and the RKLS quotient represents the non-management factors from the USLE. Under guidelines established by USDA, land was classified as highly erodible where $EI \geq 8$.

In Virginia, the Virginia Geographic Information System (VirGIS) was developed as a means of targeting NPS pollution potential from cropland in the Chesapeake Bay drainage area of Virginia (Shanholtz et al., 1987a; Hession and Shanholtz, 1988). VirGIS was built around a modified USLE with a delivery factor for the prediction of potential sediment loading (PSL; Hellmund et al., 1986), where

$$PSL = R * K * LS * C * P * DR \tag{2.7}$$

The R, K and LS factors are as in the USLE. The C-factor was assigned a value of 0.35 for cropland, 0.08 for pasture, and 0.01 for non-agricultural land. The P-factor was a constant value of 1, since management practices were not evaluated. DR was a delivery ratio based on the distance and slope to the nearest stream, the current cell's land use, and the intervening land use

(Heatwole and Shanholtz, 1991; Shanholtz et al., 1987b; Shanholtz and Kleene, Draft). The VirGIS database contains a series of raster data layers at a 1/9 ha cell size, on which all calculations are based. A water quality index (WQI) was defined by ranking all cells based on calculated PSL values and assigning index values based on rank categories. The WQI has a maximum value of 9 for the highest agricultural NPS pollution potential and a value of 3 for the lowest. Non-agricultural areas have been assigned an index of 1. The WQI is a relative index based on percentages of the total agricultural area being ranked. The WQI values and their percentage categories are: 3(1-60%); 5(61-80%); 6(81-85%); 7(86-90%); 8(91-95%); and 9(96-100%). Cells with a WQI of 9 represent 5% of all agricultural land with the highest PSL values, and, therefore, the highest pollution potential. With the current category definitions, WQI comparisons are only valid between sub-areas. The procedure could readily be adapted for more widespread use by standardizing the category threshold values. To further assist in watershed assessment, the watershed pollution density index (PDI) was created to rank watersheds within a Soil and Water Conservation District (Shanholtz et al., 1987). The PDI value represents the average potential stream sediment loading from all agricultural land within each watershed, and was calculated as:

$$PDI_j = \frac{\sum PSL_{ij}}{n_j} \tag{2.8}$$

where n is the number of agricultural cells in watershed j, and i is an agricultural cell counter. Later in 1988, the above index was revised to account for relative total potential sediment loading per watershed, based on the ratio of cropland within any watershed to total cropland in the study area (Shanholtz et al., 1988b). This watershed pollution density index (WPDI) was defined as:

$$WPDI_j = \frac{10}{n_j} * \frac{A_j}{A_{TOT}} * PDI_j \tag{2.9}$$

where A_j is the cropland acreage in watershed *j*, A_{TOT} is the total cropland acreage in the study area, 10 is a scale factor, n is the number of agricultural cells in watershed j, and i is an agricultural cell counter.

In a parallel application of the VirGIS database for targeting NPS pollution potential due to livestock, the animal waste pollution index (AWPI) was created (Shanholtz et al., 1988a; Heatwole and Shanholtz, 1991). The AWPI is an index of potential nutrient load delivered to the nearest stream both from a concentrated livestock facility and from the application of manure to agricultural land within a specified radius of the facility. The same basic formula was used for calculating potential loads either from the facility or from the area where manure was spread. Potential stream nutrient loading (R) was calculated as a product of: the manure application rate (L) based on land use, nutrient type and animal type; an availability factor (A), each for the confinement site and for the fields used for spreading; and a delivery ratio (DR)

based on distance and slope to the nearest stream. The equation for potential stream nutrient loading was:

$$R = L_s * A_s * DR_s + \sum_{f=1}^{n} \left(L_f * A_f * DR_f \right) \quad (2.10)$$

where subscripts s and f refer to confinement site and field, respectively, and n represents the number of agricultural cells within a 48 cell region of the confinement site. The AWPI was calculated as a normalized potential NPS stream load of nitrogen (N) and phosphorus (P) as:

$$AWPI = \frac{R}{(n+1)} \quad (2.11)$$

Maryland developed two NPS pollution indexes - the potential phosphorus release index (PPRI) and the potential nitrogen release index (PNRI) - in the process of prioritizing watersheds within their state's NPS programs (Spickler, 1984). Watersheds within the state were ranked based on their potential risk for releasing N and P. Potential risk was assessed based on soil and land characteristics, management factors, general cropping patterns and animal waste inputs within each watershed. The phosphorus index was calculated as:

$$PPRI_i = \frac{PP_i * AP_i * WA_i}{I_i} * DR_i \quad (2.12)$$

where PP, the specific potential P release, was defined as:

$$PP_i = \frac{C_i + P_i}{WA_i} + \frac{C_i}{WA_i} + \frac{CONV_i}{C_i} + 2 * \frac{CR_i}{C_i} \quad (2.13)$$

where C_i = cropland acreage in watershed i,
P_i = pasture acreage in watershed i,
WA_i = watershed i total acreage,
$CONV_i$ = cropland acreage in conventional tillage, and
CR_i = acreage of cropland on critical (steep or erodible) soils.

AP, the annual available P per acre is:

$$AP = AW + BG + EX \quad (2.14)$$

where AW = phosphorus from animal waste, BG = background mineralizable P, and EX = P extracted from plant tissue and residue at soil surface.

I is the internal impedance factor (defined as WA divided by the total stream length in the watershed), and DR is the delivery ratio based on estimated travel time from the watershed outlet to the fall line.

The potential nitrogen release index (PNRI) was calculated similarly with the following modifications: 1) since leaching is assumed to be the main pathway for nitrogen loss, critical soils were defined as highly permeable soils instead of steep or erodible; 2) mineralization of N occurs from animal manure; 3) available N accounts for excess fertilization; and 4) delivery ratio is 100% for leaching. Neither of the indexes were normalized, and so were not confined within a specific range. The watersheds were ranked according to the most critical 10, 25 and 50% areas of the state in terms of potential release and delivery of phosphorus and nitrogen to the Chesapeake Bay.

DRASTIC is a standardized ranking scheme which produces a relative index of ground water contamination vulnerability based on the product of seven hydrologic factors represented by the letters in its acronym (NWWA, 1985). These factors, representing measurable parameters, are depth to water table (D), net recharge (R), aquifer media (A), soil media (S), topography (T), impact of the vadose zone (I), and hydraulic conductivity of the aquifer (C). The values for each parameter are transformed through the use of standardized rating curves into sub-index values. The rating curves are either nonlinear or step functions, based on the professional experience of the researchers and a technical advisory committee. Weights were determined by the committee using the Delphi consensus approach. Each sub-index value is multiplied by the assigned weights and summed for the overall DRASTIC index. The higher the index, the greater the potential for ground water contamination.

Pennsylvania developed a GIS screening model to rank watersheds statewide for agricultural pollution potential based on an NPS index (Hamlett et al., 1992; Peterson et al., 1991). The agricultural pollution potential index (APPI) was composed of four sub-indexes: RI the runoff index, SPI - a sediment production index, ALI - an animal loading index, and CUI - a chemical use index. Raster data were gathered at the 100 meter cell size from 1:250,000 scale source maps, and analyzed at the watershed level. The index was calculated as follows:

$$APPI_i = RI_i * WF_1 + SPI_i * WF_2 + ALI_i * WF_3 + CUI_i * WF_4 \qquad (2.15)$$

where i represents individual watersheds, and WF is the weighting factor for each sub-index. The RI uses the SCS curve number (CN) model to predict runoff volume, based on maximum monthly precipitation values (five-yr maximum monthly 24-hr totals), but excluding management considerations in determining CN. The SPI uses the USLE model, with interpolations of local station EI values and an assumed slope length of 60m for all cells, and a non-spatial delivery ratio based on the inverse of the drainage area to the 0.2 power. ALI was calculated as the potential manure produced from all livestock sources divided by the agricultural area within a given ZIP code area. A matrix of relative loadings of N, P and pesticides for various land uses was used to assign chemical use potential values. These values were then summed and divided by the watershed area for the CUI value. Sub-index development was performed by researchers, while relative weights were assigned by state DER personnel. The final rank order of watersheds was from the highest pollution potential (APPI = 1), to the lowest pollution potential (APPI = 104). Kasi (1994) reported on progress made with

Pennsylvania's APPI. After considerations of various options, the APPI was revised to calculate unit area loads for the various indexes using the agricultural acreage only rather than the total watershed acreage. A technical work group further revised the index to the following form:

$$APPI_i = SPI_i * 0.2 + ALI_i * 0.8 \qquad (2.16)$$

Plans are underway to expand the GIS database so that a DRASTIC ground water index can also be calculated and possibly incorporated into the APPI.

Virginia has expanded its original agricultural NPS program to consider other sources of NPS pollution (DSWC, 1993). They developed a statewide NPS pollution assessment procedure for agricultural, urban and forestry areas to broaden the basis for their NPS priority watersheds. Inventory data were used to rank watersheds for NPS pollution potential based on land use, animal densities, estimated loading factors and other related data. Inventory data were initially collected at the county level and then disaggregated to the watershed level through the use of questionnaires to county and field agricultural services personnel. This assessment did not account for any NPS management controls which may have existed to limit the pollution potential. Three agricultural, two urban and two forestry indexes were created in arriving at the statewide rankings. Separate agricultural and urban nutrient loads were estimated by multiplying individual land uses by nutrient loading factors representing fertilizer sales and sludge distribution. Fertilizer nutrients were distributed to both cropland and residential land, whereas sludge was assumed to apply only to agricultural land. Animal nutrient loads were calculated by multiplying the number of each animal type by an appropriate waste generation factor. Erosion loads were calculated for agricultural land based on the 1987 NRI erosion ranges and acreages; for urban areas, erosion loads were calculated using erosion rates of 45 t/ac for disturbed land and 0.6 t/ac for undisturbed urban areas. Forestry harvesting erosion loads were calculated as total acres harvested times logging erosion rates, and site-preparation rates were calculated in a similar fashion. The individual indexes were then created by dividing loads by watershed area and normalizing by dividing by the corresponding statewide average watershed unit load. The overall NPS pollution (NPSP) ratings were calculated as:

$$NPSP = 0.475*AGTOT + 0.475*UTOT + 0.05*FTOT \qquad (2.17)$$

where AGTOT is the sum of the three normalized agricultural indexes, UTOT is the sum of the two normalized urban indexes, and FTOT is the sum of the two normalized forestry indexes. The higher the NPSP value, the higher the pollution potential. Rather than using the index values *per se*, the watersheds were broken into three categories based on rank order; the top 20% were ranked high, the next 30% ranked medium priority, and the remaining 50% as low. These rankings were further qualified with a '+', ' ', or '-' to indicate, respectively, the presence of monitoring data indicating water quality problems, the absence of monitoring data, and the presence of monitoring data showing no water quality problems.

The watershed index of pollution potential (WIPP) was developed by the Tennessee Valley Authority (TVA) for use in assessing agricultural NPS pollution potential in areas where land use databases were either not available or were subject to rapid changes (Sagona and Phillips, 1994). TVA uses low altitude color infrared photography to identify and characterize land uses and nonpoint sources of pollution as part of its water resources management activities. Sagona and Phillips (1994) utilized these survey results along with physical watershed characteristics from topographic maps and published stream flow data in the development of their index. Two distinct sets of scenarios were defined for evaluation within the WIPP: land use activities significant during runoff events, and those significant during non-runoff conditions. This index used the classic sum-of-weighted-factors approach, combined with a surrogate set of metrics to represent NPS pollution potential under both runoff and non-runoff conditions. They first identified separate categories of agricultural NPS pollution which they considered most important during both types of scenarios. The non-runoff scenario metrics are nutrients/pathogens, sediment/turbidity, and streambank condition. The runoff scenario metrics are nutrients/pathogens, sediment, and transport. They then identified a number of measurable features which they would use to score each metric. For instance, the feature, *areal extent of cropland*, would be rated as one of the measures of the *sediment* metric for runoff conditions. Each feature would be given either a high (1), medium (3), or low (5) pollution potential score based on a set of rules for each feature. The feature scores within each metric were then averaged to ensure equal representation of all metrics, since the number of features used to evaluate any metric may vary. Scores were then summed for metrics within each scenario. The WIPP index is the average of the runoff and non-runoff scenario scores, and can range from 12 - 60. The lower the WIPP score, the greater the potential for agricultural NPS pollution. The index was constructed to be flexible in the number and type of features used to evaluate each pollution category. All WIPP scores are relative only within the area surveyed.

Shanholtz and Kleene (Draft) define a procedure for indexing multiple NPS pollutant parameters based on computer modeled output. Their watershed non-point source management system (WATNPS) model considers sediment, nitrogen, phosphorus and the N/P ratio. The modeled output for each parameter is assigned an individual integer index from 1 (lowest pollution potential) to 5 (highest), based on user-defined range criteria. The overall index is a concatenation of the four indexes in the order: sediment, nitrogen, phosphorus, and N/P ratio. For example, an index of 5243 ranks sediment in the highest pollution potential category, nitrogen near the lowest, phosphorus near the highest, and the N/P ratio intermediate. This procedure is being developed for use both at the field and watershed level. An alternative index, the Stream Drainage Quality Index, is proposed from a regression equation between selected sub-metrics of the Index of Biotic Integrity (IBI), and sediment, the N/P ratio, disturbed streambank and the ratio of good pasture to agricultural land. This index is a predictor of pseudo-IBI scores, which integrate biological effects, along with physical and chemical effects, of NPS pollution on stream water quality.

The Chesapeake Bay Decision Support System (CBDSS) is an evolving, GIS-based, multi-layered and hierarchical system which is being built in response to various policy and

programmatic needs within the Chesapeake Bay watershed. Initial focus is on the agricultural and forestry resources. CBDSS is being cooperatively developed through a publicly-funded, multi-organizational, multi-institutional effort. Plans are underway to adopt, modify or create a whole range of indexes for assessing the NPS pollution potential on ground water, surface water and wildlife habitat, and the NPS potential of urban runoff and potentially highly erodible croplands. This system will include indexes for potential sources and transport of soluble and adsorbed pollutants from both agricultural and urban areas (CBDSS Working Group, 1994).

Other ranking schemes do not explicitly calculate indexes, but some of the factors considered may be useful for future indexes. EPA (1987) described prioritizing schemes for additional states, not already mentioned. Illinois approached their watershed prioritization from the grassroots level, where problems were identified and then reviewed and screened at county, regional and state levels. The emphasis within its rating scheme was on soil erosion. Ohio's approach was more data intensive and included wastewater treatment, waste disposal and other forms of NPS in addition to agricultural NPS, and incorporated the intended uses of the various water resources. Wisconsin ranked each of the state's 330 watersheds based on both agricultural and urban NPS problems. An initial technical screening identified the top 25% watersheds with land management and water quality problems, based on the extent of severe soil erosion, the extent of urban land, the concentration of animals and the acreage of lakes and streams in each watershed. Additional review of the top watersheds was more subjective, and aimed at involving local and regional interests. Its selection criteria included: 1) the severity of water quality use impairments; 2) the practicability of alleviating the impairments; and 3) the threat to high quality, recreationally valuable waters. EPA also provided suggested criteria and example scoring schemes for both phosphorus-specific and pesticide-specific problems.