FUZZY CLASSIFICATION AND POST-PROCESSING OF SATELLITE IMAGERY

TO DERIVE WATERSHED MODEL PARAMETER VALUES

by

ROBERT CLARK FULLER

(Under Direction the of E. Lynn Usery)

ABSTRACT

Nonpoint source water pollution, because it is diffuse in its origin, is more difficult to locate and control than point source water pollution. Modeling of the hydrologic and water quality characteristics of watersheds is a valuable tool in the effort to control nonpoint source pollution. Remotely sensed imagery, including satellite imagery and aerial photography, are increasingly being used as sources of data in the formulation of watershed models. Image processing software and geographic information systems (GIS) have been used to classify remotely sensed data, organize those data, and prepare them for use in watershed models. In prior studies of this type, remotely sensed data have been classified into land use/land cover classes, then linked to tables containing watershed model parameter values associated with the classes. In this study, field estimates of selected watershed model parameter values were used as training set data for fuzzy classification of Landsat Enhanced Thematic Mapper Plus (ETM+) satellite imagery into three sets (image layers) of possible values of the parameters. The associated spectral distance files, based on Mahalanobis distances, were used in a post-classification model to assign weighting values to the three fuzzy layers and calculate composite parameter values for subsequent use in the AGNPS watershed model. A second spatial model was developed to perform accuracy assessment on the classification and modeling used to generate watershed model parameter values. The technique was tested on three watersheds in Georgia, one in the Coastal Plain and two in the Piedmont. The technique produced reasonable results but it was not possible to judge its merits relative to existing techniques based on the specific application that was tested. Several other, more promising applications for the technique are suggested.

INDEX WORDS: Nonpoint source, Water pollution, Watershed, Hydrologic modeling, Water quality, Satellite imagery, Image processing, Geographic information system, GIS, Fuzzy classification, Mahalanobis distance, Spatial model, Accuracy assessment, Georgia

FUZZY CLASSIFICATION AND POST-PROCESSING OF SATELLITE IMAGERY TO DERIVE WATERSHED MODEL PARAMETER VALUES

by

ROBERT CLARK FULLER

B.S.E., University of South Florida, 1975

M.S.E, University of South Florida, 1975

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2003

FUZZY CLASSIFICATION AND POST-PROCESSING OF SATELLITE IMAGERY TO DERIVE WATERSHED MODEL PARAMETER VALUES

by

ROBERT CLARK FULLER

Major Professor:	E. Lynn Usery
Committee:	David D. Bosch Thomas W. Hodler David S. Leigh Chor-pang Lo

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2003

DEDICATION

This is dedicated to my wife of 32 years and best friend, Kathy C. Fuller.

ACKNOWLEDGEMENTS

Many people have helped me in this project, some directly and some indirectly. I am grateful to all, though I cannot hope to give all the credit that is due. Dr. E. Lynn Usery has provided advice, access to data, and the freedom to explore, learn, and grow. His work ethic and intellectual honesty continue to be an inspiration and his mentoring has been valuable and appreciated. Dr. Tom Hodler taught me about maps, showed me by example how to run a class, served as one of my most valuable mentors, and pushed when I needed pushing. Dr. C. P. Lo taught me things I thought I already knew about aerial photography but did not and stimulated my interest in both population studies and China. Dr. David Leigh showed me that there is a lot more to moving water than just the equations describing its velocity and energy. He also showed me how to develop a story in the classroom, the story of how the world works, and how to use that story to stimulate students' interest. Dr. David Bosch has made multiple lengthy trips on my behalf without protest and has generously shared his knowledge of soils, hydrology, and AGNPS. Dr. Roy Welch showed me the power of employing sophisticated satellite imagery and aerial photography to solve environmental problems, introduced me to serious environmental field work, and showed me the positive impact on students of bringing a true command of the subject matter to the classroom.

Mr. Mike Finn and Mr. Gregory Jaromack, with USGS in Rolla, Missouri, were generous with their time, expertise, and the product of their labor, AGNPS Data Generator. Mr. Rob Rivers, Director of Public Works for Hall County, Georgia, provided

much needed aerial photographs along with support and encouragement. My colleagues at North Georgia College and State University have encouraged me, cheered me on, forgiven me when I didn't carry my fair share of the load, and offered valuable advice on how to get it all done and still enjoy the doing. I thank you all.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Research background

Nonpoint source pollution is the contamination of surface waters by pollutant-laden runoff from distributed areas, as opposed to point discharges, such as pipes discharging directly into streams. Nonpoint source pollution is increasingly being seen as a major danger to the health of the surface and subsurface waters of the United States. Omernick (1976) performed pioneering work in developing empirical relationships connecting the percentages of various land uses in a watershed with the quality of the water discharging from the watershed. A 1986 report to Congress prepared by the U. S. Environmental Protection Agency (USEPA) reported that 60% of the surface water pollution problems in the United States were the result of routine agricultural activities (Tim and Jolly 1994). This estimate was reaffirmed two years later in another report to Congress (USEPA 1990).

If these problems are to be solved or at least partially ameliorated, it is essential that the mechanisms responsible for generating nonpoint source pollution be understood. Osmond et al. (1997, p. 328) stated, "In order to improve or protect water quality, it is necessary to first determine its status and then devise a protection or remediation plan." Water quality models are tools that allow for the relatively rapid simulation and

comparison of existing and proposed sets of watershed characteristics in order to predict the impact of proposed watershed management strategies.

The large quantities of data that must be employed to adequately describe a watershed for the purpose of modeling have led to the use of geographic information systems (GIS), linked to water quality models, as tools to aid in the organization, management, and preprocessing of the watershed data. In addition, GIS and image processing software have been used to extract information from remotely sensed images that can be used to develop certain water quality model parameters**.** To date, this has been done by classifying the remotely sensed imagery into traditional land use/land cover classes. Water quality model input parameter values have then been estimated for each class in the image and linked to locations (pixels) in the image by the use of lookup tables. While this method retains a linkage between the familiar land use/land cover classes and water quality model parameters that is useful to some in visualizing the processes taking place within the watershed, it is also a potential source of error that could be eliminated if the imagery were classified directly into values of the water quality input parameters. That is the focus of this research.

Research objectives

A new method is proposed for generating certain nonpoint source water quality model input parameters from satellite imagery. It is proposed that the satellite imagery be classified into fuzzy classes corresponding directly to the possible range of values of the selected nonpoint source water quality model input parameters. The fuzzy classifier uses a maximum likelihood classifier to derive a user-selected number of layers, where the first layer contains the class values with the smallest spectral distances from the class

signature, the second layer contains the class values with the second smallest spectral distances from the class signature, etc. The spectral distance used by the Imagine fuzzy classifier is the Mahalanobis distance, which includes the covariance matrix in its calculations (ERDAS 1997). In this scheme, one classification would be required for each distinct nonpoint source water quality model input parameter. Thus, if it is intended to extract values from the remotely-sensed imagery for six nonpoint source water quality model input parameters, then six separate classifications of the imagery would be required. Each classification would require that training sets be identified within the area to be modeled and that the appropriate value of the nonpoint source water quality model input parameter for that classification be field-determined for the subject area.

The objectives of the research project are:

1. to demonstrate the feasibility of classifying a satellite image into fuzzy classes representing nonpoint source water quality model input parameters rather than the traditional land use/land cover classes,

2. to develop a method of using the spectral distance associated with pixels in a fuzzy-classified image to develop weighted averages of the fuzzy-classified layers, to more closely approximate the continuity of environmental variables than is possible without using this method,

3. to compare the accuracy of the nonpoint source water quality model input parameters obtained directly from a satellite image classified in terms of nonpoint source water quality model input parameters vs. similar parameters derived from look-up tables linked to land use/land cover classes derived from the same satellite image,

4. to compare the accuracy, ease of construction, and efficacy of the nonpoint source water quality model derived from the proposed new classification method with that of the model derived from traditional land use/land cover classification of the same satellite image.

Hypotheses

This project is intended to test the following hypotheses:

1) Classifying satellite imagery directly into water quality model parameters, eliminating the use of land use/land cover classes and associated lookup tables, will improve the accuracy of the resultant water quality model by eliminating the forced fit of water quality model parameters into classes defined by land use/land cover classes.

2) The creation of weighted averages of fuzzy-classified layers of environmental variables will produce a better approximation of the continuity of those variables in the environment than can be achieved by more traditional methods.

The first hypothesis is tested by classifying each satellite image two ways: 1) into conventional land use/land cover classes and 2) into fuzzy classes covering the possible ranges of values of water quality model parameters. The results of these two classification schemes are used, with other data, to develop two different water quality models for each of two test areas. The models are run to simulate recorded rainfall events and the results are compared with each other and with water quality data collected at the respective watershed outlets for the observed rainfall events. The second hypothesis is tested by comparison of pixel values resulting from the implementation of the weighted averaging methodology with more conventional classification methods.

CHAPTER 2

LITERATURE REVIEW

Nonpoint source water quality models

Batchelor (1994) defined a model as "a deliberately simplified construct of nature erected for the purpose of understanding a phenomenon." A nonpoint source water quality model, then, is a simplification of the natural process of rainfall, runoff, erosion, sedimentation, and the solution or suspension and transport of other pollutants from the land surface, designed to simulate only those elements of the natural process that are essential to the understanding of the process and its results. Nonpoint source water quality models are variously referred to as hydrologic/water quality models and sometimes simply watershed models or water quality models. A watershed model may also refer to a hydrologic or a hydrologic/hydraulic watershed model, with no water quality elements included. Similarly, a water quality model may be a nonpoint source water quality model or it may be a model of body of water subjected only to point sources of contaminants.

Based on the work of Chow et al. (1988), models can be classed first as physical or abstract, with abstract models further divided into theoretical or empirical. Theoretical models are developed on the basis of one or more underlying theories, believed to describe the modeled system's behavior. Empirical models, on the other hand, are abstractions of compiled data representing the system inputs and responses. Empirical models are frequently developed through the use of regression techniques but also may

be refined by calibration of one or more model parameters in an attempt to force the model to respond in a manner similar to the physical system. Abstract models can also be described as being one of the four possible combinations of either lumped parameter or distributed parameter and event-based or continuous.

A lumped parameter model treats an entire watershed as a homogeneous unit, ignoring the spatial variation of factors such as slope, soil type, ground cover, and soil moisture. A distributed parameter model subdivides a watershed into smaller units in an attempt to capture the heterogeneity of attributes affecting runoff and pollutant motility and discharge from the watershed.

In the context of nonpoint source water quality modeling, the primary difference between single event and continuous models is in the way rainfall is represented. Single event models treat rainfall as a uniform phenomenon, while continuous models reflect the temporal variation of rainfall during the course of a storm. Some continuous models also account for changing soil moisture during the course of a storm and the resultant changes in infiltration rates and soil cohesiveness.

A watershed, also known as a drainage basin, is any geographic area from which all rainfall runoff drains to and through a single point or a single cross-section of a stream or river (Ritter et al. 1995, Parker 1994). A watershed model, then, is a model of a watershed and may include simulation of the processes of rainfall runoff as well as the detachment and transport of sediment, nutrients (typically nitrogen and phosphorus), pesticides, organic material, and any other material impacted by fluvial action.

Nonpoint source water pollution refers to the pollution of surface waters by contaminated runoff from distributed geographic areas. An agricultural field that might

yield runoff polluted with excess fertilizer and pesticides would be such a distributed area, as opposed to a concentrated pollution source, such as a wastewater treatment plant discharge pipe. The most common interpretation of the term includes pollutants generated in urban areas, suburban neighborhoods, and in agricultural areas. In an urban environment, the pollutants of concern include sediment, nutrients (primarily nitrogen and phosphorus), oxygen demanding organic litter, bacteria, pesticides, gasoline, oil, tire residue, and various industrial wastes (Ventura and Kim 1993). In both suburban and agricultural areas, sediment, nutrients, oxygen demanding organic litter, bacteria, pesticides, and herbicides are major pollutants of concern, with sediment being more prevalent in agricultural areas.

Because agricultural nonpoint source pollution has come to be seen as the largest water pollution problem in the U.S., major efforts have been undertaken to understand and mitigate the adverse impacts of agricultural practices. These efforts have included attempts to better understand the processes that influence surface water runoff and the movement of pollutants by that runoff, efforts to evaluate the nature and extent of agricultural pollution of surface and ground water, and the application of the knowledge gained in improving agricultural land management practices. One of the methods employed to accomplish this new level of understanding and land management control has been the development of computer models of hydrologic and nonpoint source pollutant runoff processes (Tim and Jolly 1994). According to Ciesiolka et al. (1995), the first mention of an erosion model in the literature was a description of a summary of an extensive series of field runoff plot experiments on agricultural land in the United States.

The model that eventually resulted became known as the Universal Soil Loss Equation (USLE).

The United States Environmental Protection Agency (EPA) has developed a model called BASINS (U.S. EPA 2003) which stands for "Better Assessment Science Integrating point and Nonpoint Source." This model actually links other point and nonpoint-source models though ESRI's Arcview GIS package. BASINS also includes databases of river reaches and cataloged point sources.

One example of a single event, distributed parameter nonpoint source pollution model is the Agricultural Nonpoint Source (AGNPS) pollution model. This model was developed by the United States Department of Agriculture, Agricultural Research Service specifically to estimate the runoff and sediment and nutrient yields from agricultural lands (Young et al. 1989). While the model provides estimates of the runoff and water quality from watersheds, it is not a predictive model. Its intended purpose is to provide a means of comparing the impact on runoff and water quality from watersheds containing varying agricultural practices within the watersheds. AGNPS has been selected as the nonpoint source water quality model to be used in a five-year study intended to lead to the development of management guidelines for animal-based agriculture that take into account the effect of animal-based agriculture on landscape-scale environmental quality, particularly water quality (Vellidis et al. 1995). There is some disagreement regarding the appropriate classification of AGNPS. In one publication (Baker et al. 1995), the U.S. Department of Agriculture (USDA) makes a point of saying that AGNPS is a "lump-sum parameter" model. Other researchers have different views. Mitchell et al. (1993) call AGNPS a "spatially distributed grid model." Tim and Jolly (1994) specifically call

AGNPS a "distributed parameter model." Whatever the model is called, it does allow for spatially varied input parameters to represent the geographic variation of the watershed characteristics more realistically than could be done with a lumped parameter model.

AGNPS has three major components: hydrology, soil erosion, and nutrient pollution (CIESIN 1996). Early versions of the model required that the watershed be divided into equal-area cells. These cells are created by Dirichlet tesselation (USDA 1997), which, in the case of regular gridded data points, results in a set of equal, adjacent squares. The latest release, version 5.00, still requires that the watershed be divided into cells, but it now allows for subdivision of selected cells into subcells to facilitate the capture of more detailed spatial variation of watershed characteristics (USDA 1995).

The current version of AGNPS, version 5.00, requires twelve data values common to the entire watershed and specific to the current simulation. These parameters are listed in Table 2.1.

Table 2.1. AGNPS initial watershed information

	Parameter
1.	Watershed identification
2.	Description
3.	Area of each cell
4.	Number of cells
5.	Precipitation
6.	Nitrogen concentration in rainfall
7.	Energy intensity value or Duration
8.	Storm Type
9.	Peak flow calculations
10.	Geomorphic calculations
11.	Hydrograph shape factor toggle
12.	K coefficient or % runoff prior to peak

Source: USDA (1995)

In addition to the general watershed and simulation data, specific data are required to describe the individual cells and any subdivided cells within the watershed. The AGNPS program requires that the watershed to be simulated be divided into a series of square grid cells which must be uniform in size and shape. However, individual cells may be subdivided into fourths. The resultant subdivided cells may themselves be subdivided on an individual basis, i.e. the subdivision may be applied to none, one, two, three, or all of the subdivided cells. This process of subdivision, which results in a quadtree structure (Burrough 1986) may be completed in this manner to one, two, or three levels, resulting in cell divisions that are 1/4, 1/16, or 1/64 the size of the original cell.

The major cell parameters used in Version 5.00 of AGNPS and its precursor, version 4.03, their default values if they are supplied, and suggested sources of the parameter values are listed in Table 2.2 (USDA 1997). The first and second cell parameters are simply identifying numbers, which are created automatically by the AGNPS spreadsheet editor. Parameter1 values are created on the basis of the number of cells identified for the watershed. Values for parameter 2 are assigned on an individual cell basis only in response to invoking the cell division option in the spreadsheet editor. While it is most likely possible to generate the array values for cell numbers automatically within Imagine or other GIS software, it is unlikely that any meaningful cell subdivision could be accomplished without the intervention of human judgment. The automatic generation of a cell array and cell numbers within Imagine could be accomplished regardless of whether the watershed boundaries were human-produced or generated by a watershed boundary routine operating on a digital elevation model (DEM)

Table 2.2. AGNPS cell information

Parameter	Data values or possible source
1. Cell number	AGNPS generates west to east, north to south
2. Cell division	AGNPS generates if divided
3. Receiving cell number	DEM: Aspect - Recipient cell of most flow
4. Receiving cell division	DEM: Aspect - Recipient division of most flow
5. Flow direction	DEM: Aspect - 8 classes
6. SCS curve number	Soils and Land cover class: Look-up table
7. Land slope (%)	DEM: Slope
8. Slope shape	DEM: 1st derivative of slope; Reclass
9. Slope length	DEM + hydrography
10. Overland Manning's coefficient	Land cover class: Look-up table
11. K-factor (soil erodibility)	Soils: Look-up table - USDA Ag Handbook 537
12. C-factor (cropping management)	Land cover class: Look-up table
13. P-factor (conservation practice)	Land cover class: Look-up table
14. Surface condition constant	Land cover class: Look-up table
15. COD factor	Land cover class: Look-up table
16. Soil texture #	Soils: Look-up table
17. Fertilizer indicator	Select 0 or 1
18. Pesticide indicator	Select 0 or 1
19. Point source indicator	Select 0 or 1
20. Additional erosion indicator	Select 0 or 1
21. Impoundment indicator	Select 0 or 1
22. Channel indicator	Select or from Hydrography

Note: DEM = digital elevation model.
Source: USDA (1997)

of the area of interest. Parameters 3, 4, 5, 7, 8, and 9 can all be generated by ERDAS Imagine from a DEM of the watershed, except that parameter 9 will also require an overlay of a hydrography layer in the GIS. In their work to integrate AGNPS with Imagine, Olivieri et al. (1991) found it beneficial to work with a DEM grid spacing that was six times as dense as the grid spacing used in AGNPS, i.e. each AGNPS cell was broken into 36 subcells within Imagine for greater accuracy in generating slopes, aspects, and slope shapes. Slopes and aspect values can be directly determined from the DEM by

ERDAS Imagine. Aspects will require recoding to one of eight directions, determined clockwise from 1 (north) through 8 (northwest). Application of the slope function to the slope layer previously created should result in a second derivative of the elevation or a slope shape (curvature). Receiving cell number and receiving cell division can be automatically determined from the aspect with reference to the AGNPS grid structure. However, some judgment may be appropriate, with reference to an overlay of hydrography. The determination of slope length will require the overlay of hydrography with the finer grid ERDAS Imagine aspect file (e.g. 36 subcells) in order to determine the representative length of overland flow to the nearest identified channel. Srinivasan and Engel (1994) reported using both aspect and elevation, in conjunction with unit stream power theory, to derive average field slope length. A restriction within AGNPS that this value should not exceed 300 feet may result in assignment of a value of 300 feet to this parameter for most cells if cell sizes are relatively large.

Parameters 6, 11, and 16 relate to the dominant soils in each cell. These values can all be determined using ERDAS Imagine to reclass a soils layer in the GIS with reference to look-up tables. Guidelines for the selection of values for these parameters and suggested values for a variety of soil types for use in creating the look-up tables can be found in Young et al. (1994).

Parameters 10, 12, 13, 14, and 15 have reportedly been generated through classification of Landsat Thematic Mapper (TM) images and aerial photographs (Olivieri et al. 1995; Olivieri et al. 1991; Tim and Jolly 1994). Olivieri et al. (1991) assumed worst case conservation practices (parameter 13) but used supervised and unsupervised classification of TM images to estimate Manning's coefficients (parameter 10), cropping

factors (parameter 12), and surface conditions (parameter 14). Although they attributed some of problems to accuracy assessment methodology, Olivieri et al. (1991) reported supervised classification accuracy of 59% and unsupervised classification accuracy of 56%. These numbers were not reflective of the additional errors inevitably introduced by the translation of land use/land cover classifications to the AGNPS parameters previously identified through the use of look-up tables.

AGNPS is designed to process an ASCII text file in which the principal data values are arranged in columns of eight characters. The first, third, and fourth lines of the data file are reserved for descriptive text values. The remainder of the file comprises a mixture of text keywords and numerical values, all arranged in eight-character per datum format.

Role of GIS in nonpoint source modeling

One of the major limitations of nonpoint source water quality models has been the very large volumes of data required to accurately describe the spatial variation of the numerous factors that affect the runoff of storm water, the solution or suspension of pollutants, and the transport of those pollutants into surface waters. GIS are tools that show great promise for the organization, preprocessing, and passing of spatially varied attributes of the land surface to nonpoint source water quality models. GIS also have been shown to be useful for the display and interpretation of the results of nonpoint source water quality modeling (Srinivasan and Engel 1994).

Table 2.1 and Table 2.2, in describing the AGNPS model, present a typical range of parameters required for a nonpoint source water quality model to reasonably describe those watershed characteristics that cause variance in rainfall runoff and the

concentrations of selected pollutants in that runoff from a watershed. Other models may require more or less input data but the data required for AGNPS is not untypical (Oliver and Solomon 1990; Rose et al. 1990). If the model is a lumped parameter model, then only single values for these descriptive parameters need be supplied for the entire watershed. If, however, the model is a distributed parameter model, then the number of parameters required to describe the watershed must be multiplied by the number of divisions of the watershed for modeling purposes.

The requirement for large amounts of input data required to describe the heterogeneities of the natural system is a major limitation in the use of hydrologic/water quality (H/WQ) models (nonpoint source water quality models). Tim and Jolly (1994, p. 25) stated that for quite some time, "researchers have recognized that the spatio-temporal variability in landscape characteristics including soil, land use, topography, and climate, affects the hydrologic response of the physical system and severely limits the applicability of models." They further pointed out that the "extreme complexity of manipulating large volumes of spatial and nonspatial (or attribute) input data, for example, severely limits the use of distributed H/WQ models."

The recent development and proliferation of GIS, however, has provided a new set of tools for capturing, storing, manipulating, organizing, and preprocessing spatially varied watershed data for use in distributed parameter nonpoint source water quality models. Tim (1996) evaluated 35 hydrologic and water quality models, describing them as lumped or distributed parameter and as event-based or continuous. He then rated each model based on its potential for integration with GIS. These descriptions and ratings are summarized in Table 2.3. The potentials for integration with GIS were determined on the

Table 2.3. Classification of Hydrologic/Water Quality Models

Model Acronym	Space Domain		Time Domain		Potential for Integration with GIS
	Lumped	Distributed	Event-based	Continuous	
ACTMO	X			X	Low
AGNPS		X	X	X	High
ANSWERS		X	X		High
ARM II	X			X	Very low
CNS	X			X	Very low
CPM	X			X	Very low
CPS	X			X	Very low
CREAMS	X			X	Low
EPIC	X		X	X	Low
FESHM		X	X	X	High
GAMES		X	X		Moderate
GLEAMS	X			X	Low
HSPF	X	X		X	High
LEACHM	X		X	X	Moderate
NPS	X			X	Very low
NTRM	X			X	Low
NLEAP	X			X	High
OPUS	X			X	Low
PLIERS	X		X	X	Low
PRMS	X			X	High
PRZMII	X			X	Low
ROTO	X			X	Moderate
RUSLE	X		X		High
RZWQM	X			X	Low
SHE		X	X	X	High
SPAW	X			X	Low
SPUR	X			X	Moderate
STORM	X			X	Moderate
SWAT		X		X	High
SWM	X			X	Low
SWMM	X			X	Low
SWRRB	X		X	X	Moderate
TOPMODEL		X	X	X	High
USDAHL	X		X	X	Low
WEPP	X	X	X	X	Moderate

Adapted from Tim (1996)

basis of the structure of the models and the perceived difficulty likely to be encountered in the process of linking the models with GIS. Very low and low integration potential implies that loose coupling might be feasible. Moderate integration potential corresponds to a potential for close coupling of the model and GIS; high integration potential corresponds to a potential for tight coupling of the model and GIS.

Tim (1996) defined loose coupling of GIS with hydrologic/water quality models as the use of GIS to generate and organize input data and display the model output. In this type of coupling, an interface program is employed to edit and reformat the data required by the hydrologic/water quality model from the GIS and pass those data to the model. Similarly, an interface program converts the model results into a form usable by the GIS and passes the data to the GIS, where any desired post-processing takes place and the results are displayed. Close coupling is defined as the use of macros, macro languages, user-callable routines, and/or analytical functions to cause the GIS and the hydrologic/water quality model to execute together. Program modifications may be made in either or both the GIS and the hydrologic/water quality model. Tight coupling is defined as the addition of GIS functionality directly into the hydrologic/water quality model.

In addition to these three linkage strategies, some efforts have been made to implement hydrologic/water quality modeling functions directly in a GIS. Robinson and Ragan (1993) wrote a GIS specifically to calculate nonpoint source loading rates for sediment, BOD, lead, zinc, nitrogen, and phosphorus based solely on empirical functions derived for eleven land use types and four soil types. These functions were developed by

the Northern Virginia Planning District Commission as average annual rates for the region. As such, they do not reflect variation in hydrology and, thus, do not represent a hydrologic/water quality model. A procedure for calculating rainfall erosivity from rainfall data, suggested by Reyes and Gayle (1995) may have potential for integration with the work by Robinson and Ragan (1993). Wilson and Gallant (1996) used the grid-based terrain analysis programs, TAPES-G and EROS, developed by Panuska and Moore (1991) to derive erosion indices for areas and converted them for further manipulation and display within Arc/Info. This procedure also failed to use hydrology as an independent variable and, therefore, is not a true hydrologic/water quality model. Wang and Yin (1997) used a very different approach, employing the statistical analysis program SAS to perform correlation analysis on USGS water quality data and land use data contained in an Arc/Info GIS. This approach produced interesting relationships between land use and a large number (199) of water quality parameters but still falls short of being a hydrologic/water quality model.

Integration of GIS and hydrologic/water quality models

The first two levels of coupling GIS with hydrologic/water quality models, as defined by Tim (1996), have been accomplished. Tim and Jolly (1994) developed a closely coupled linkage between AGNPS and Arc/Info GIS software using special-purpose computer programs and interfaces. Digital data representing the spatially varied landscape and hydrologic parameters were assembled in Arc/Info vector format, converted to raster format, and passed to AGNPS using the special-purpose computer programs and interfaces. The process was reversed to pass AGNPS results back to Arc/Info for further processing and display. While this effort was successful and resulted

in modeling at a finer resolution than was feasible using manual input of AGNPS data, the authors pointed out that the entry of data into the GIS that was linked to AGNPS was still very time-consuming and costly. They maintain that significant time and cost savings will only be realized with the use of existing GIS databases for hydrologic/water quality modeling or, conversely, using the GIS developed for a hydrologic/water quality modeling project for other purposes. Morse et al. (1994) reported the development of an integrated Arc/Info - AGNPS system, with the additional linkage of ORACLE data files. All three were linked with special-purpose FORTRAN programs. Other researchers have reported successful integration of SWAT (Soil and Water Assessment Tool) and the GRASS GIS (Mitchell et al. 1993; Srinivasan and Engel 1994) as well as AGNPS and GRASS (Srinivasan and Arnold 1994).

Olivieri et al. (1991) took this linkage a step further by using Imagine GIS and image processing capabilities to generate eighteen of the 22 spatially varied input parameters required by AGNPS (excluding fertilization level, fertilizer availability, point source, and gully source indicator values) from classified Landsat thematic mapper (TM) images or aerial photographs, digitized soil map units, and digitized USGS topographic maps. Olivieri et al. (1991) were working with an earlier release of AGNPS than that which is currently available, so the input parameters they used differed somewhat from those described herein for versions 4.03 and 5.00. The soils and topographic data were sometimes obtained from other sources and converted from other GIS formats, including GRASS and Arc/Info (Olivieri et al. 1995). Values for the eighteen AGNPS input parameters were calculated partially within Erdas Imagine and partially using programs developed specifically for this purpose.

Finn et al. (2002) developed a program, Generator.exe, and related models to generate AGNPS parameter values from a digital elevation model, soils image, and land cover image. This program executes from within Imagine, using file types and procedures that are familiar to Imagine software users. The program also has Version 5.0 of AGNPS embedded within it and has additional routines that generate Imagine images from the AGNPS output file. This suite of programs and models, known collectively as Data Generator, was used in this study as one method of creating input data files for processing by AGNPS 5.0.

CHAPTER 3

STUDY SITES

Three watersheds were selected for this study, Subwatershed K (including its tributary, Subwatershed M) of Little River near Ashburn, Georgia; West Fork Little River, in Hall, White, and Lumpkin Counties, Georgia; and East Fork Little River in Hall and White Counties, Georgia. The fact that the two forks of Little River in North Georgia share the name of Little River in South Georgia is purely coincidental. The Little River in South Georgia has been under intensive study by the U.S. Department of Agriculture (USDA) and the University of Georgia (Vellidis et al. 1995) and in the course of that work, has been divided into several subwatersheds for convenience of nomenclature. They were assigned letters "A" through "N", though some subwatersheds drain into others. To avoid unnecessary confusion, Subwatershed K of Little River, including its tributary, Subwatershed M, will hereinafter be referred to simply as Subwatershed K. West Fork Little River and East Fork Little River were originally tributaries to Little River in Hall County, Georgia. With the construction of Buford Dam and the filling of Lake Sidney Lanier behind it, these two rivers now discharge into the lake. These rivers were selected because their watersheds are similar in size and land use characteristics and because it was desired to have test watersheds in both the Coastal Plain and the Piedmont physiographic provinces of Georgia. Subwatershed K is located in the Coastal Plain physiographic province of Georgia. West Fork Little River and East Fork Little River are located in the Piedmont physiographic province of Georgia. The

three test watersheds are located on a map of Georgia in Figure 3.1. The watersheds are shown to scale and county boundaries are shown to aid in locating the watersheds. Also shown are approximate 13X enlargements of the watersheds, to illustrate their sizes relative to one another. The visualization and digitizing of the watershed boundaries are described in Chapter 4.

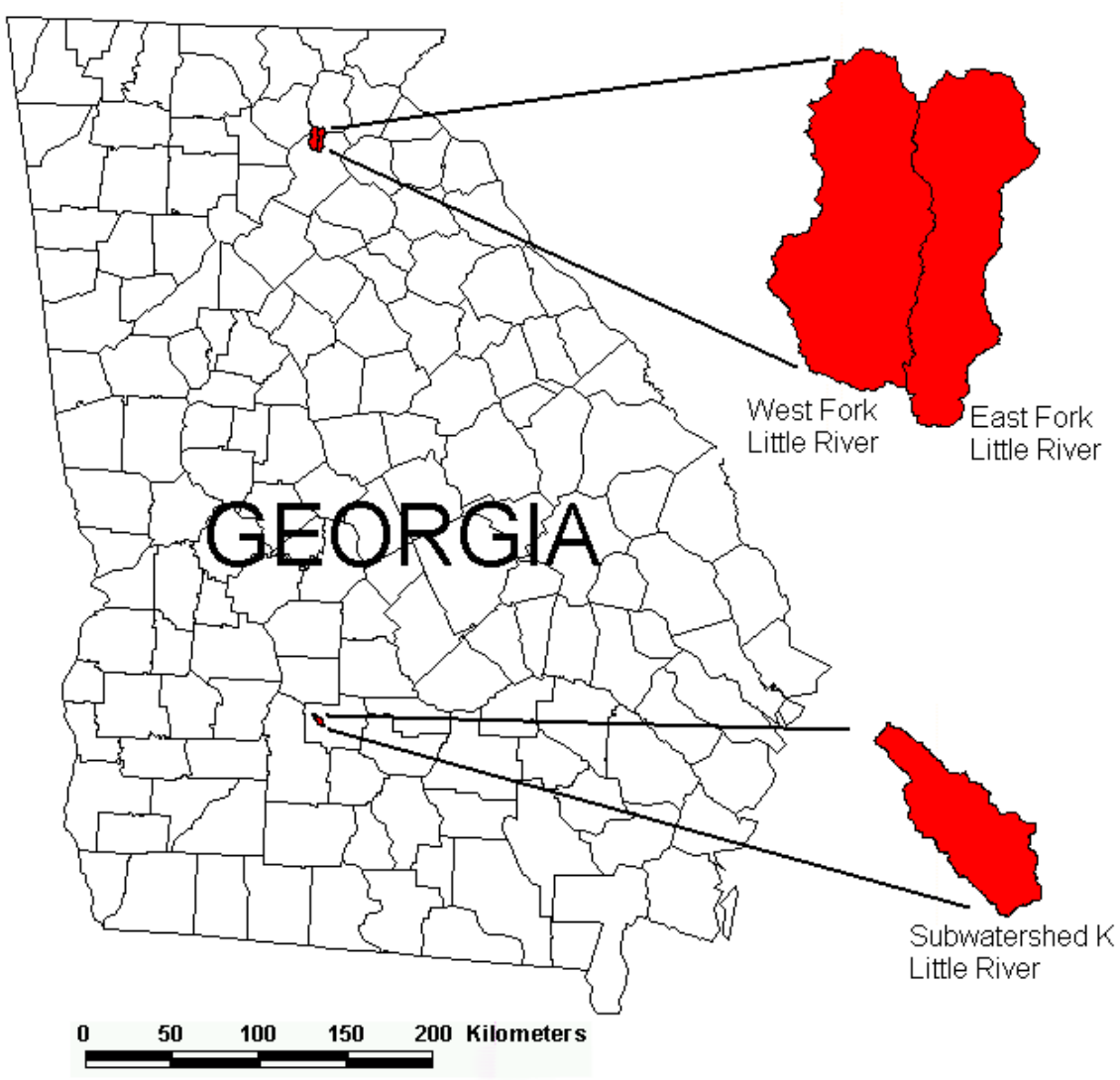


Figure 3.1. Location of the three test watersheds. East Fork Little River and West Fork Little River are tributaries to a different river than that to which Subwatershed K Little River drains. The name similarity is strictly coincidental.

Some of the characteristics of the three watersheds are listed in Tables 3.1 and 3.2. The Shreve Magnitude and Strahler stream order are described in Ritter et al. (1995). The sum of all stream lengths for each basin was, as the term implies, calculated by adding the lengths of all perennial stream segments shown on the 7.5-minute topographic maps. This was accomplished by digitizing the stream segments from the digital raster graphics (DRGs) with ERDAS Imagine, creating line topology in the resultant vector layer, and reading the stream lengths from the associated attribute tables created by Imagine. Reading the highest and lowest elevations from the DRG and taking the arithmetic difference yielded the maximum relief for each basin. The bifurcation ratios given in Table 3.2 are, in each case, means of the first to second-order ratio and the second to third-order ratio, given by

$$\overline{R}_b = \frac{\left(\frac{N_1}{N_2} + \frac{N_2}{N_3}\right)}{2}, \text{ where}$$

$\overline{R}_b =$ *the mean bifurcation ratio,*

$N_1 =$ *the number of Strahler first order stream segments,*

$N_2 =$ *the number of Strahler second order stream segments,* and

$N_3 =$ *the number of Strahler third order stream segments.*

Drainage density figures, which are a way of quantifying the average length of stream channel per unit area, were calculated by dividing the sum of the lengths of all perennial stream segments in each basin by the drainage basin area. Relief ratios were calculated by

$R_h = \frac{H}{L}$, where

R_h = *the relief ratio,*

H = *the basin maximum relief,* and

L = *the longest horizontal length of the basin parallel to the main channel.*

Table 3.1. Selected characteristics of test watersheds.

Test watershed	Location, county(ies)	Drainage area, km^2	Drainage pattern	Shreve Magnitude	Strahler stream order
Subwatershed K Little River	Turner	16.95	Dendritic	12	3
West Fork Little River	Hall, White, Lumpkin	47.34	Dendritic	29	3
East Fork Little River	Hall, White	36.37	Dendritic	18	3

Table 3.2. Selected basin morphometric characteristics.

Test watershed	Bifurcation ratio (mean)	Sum of all stream lengths, km	Drainage density, km/km^2	Maximum relief, m	Relief ratio
Subwatershed K Little River	3.50	26.47	1.56	42.4	.00517
West Fork Little River	5.57	65.84	1.39	458.1	.03704
East Fork Little River	4.50	49.30	1.36	449.6	.03455

Georgia Coastal Plain watershed

Vellidis et al. (1995) proposed to study the Little River watershed in South Georgia, including its hydrology, water quality, and the impact of various agricultural practices on its water quality. In a personal communication, Richard Lowrance (1999), one of the study participants, suggested using Subwatershed K, a small segment at the north end of the Little River watershed in the Georgia Coastal Plain, as a test watershed for this study. He said that more water quality data had been gathered for that watershed than for any other in the Little River watershed. Figure 3.2 depicts Subwatershed K with each perennial stream segment labeled with both the Shreve Magnitude and the Strahler stream order.

Subwatershed K is located entirely in Turner County, Georgia. The centroid of the watershed is approximately five and a half kilometers west-northwest of the center of Ashburn, Georgia. The area ranges from nearly flat to gently rolling. It is nearly equally divided between row crops and forested. The most abundant row crop is cotton, though a significant portion of the cropped area is used to grow peanuts. Much of the forested land is covered with planted pine trees, though the wetlands immediate adjacent to most creeks are largely covered with hardwood or mixed hardwood and coniferous forest.

Subwatershed K drains generally from the north-northwest to the south-southeast. Its elongated shape would be expected to cause the runoff hydrograph to peak a little later and have a somewhat lower peak than would the hydrograph for a watershed that was proportionally wider but identical in all other respects. The watershed drains an area of 16.95 square kilometers (4,188 acres).

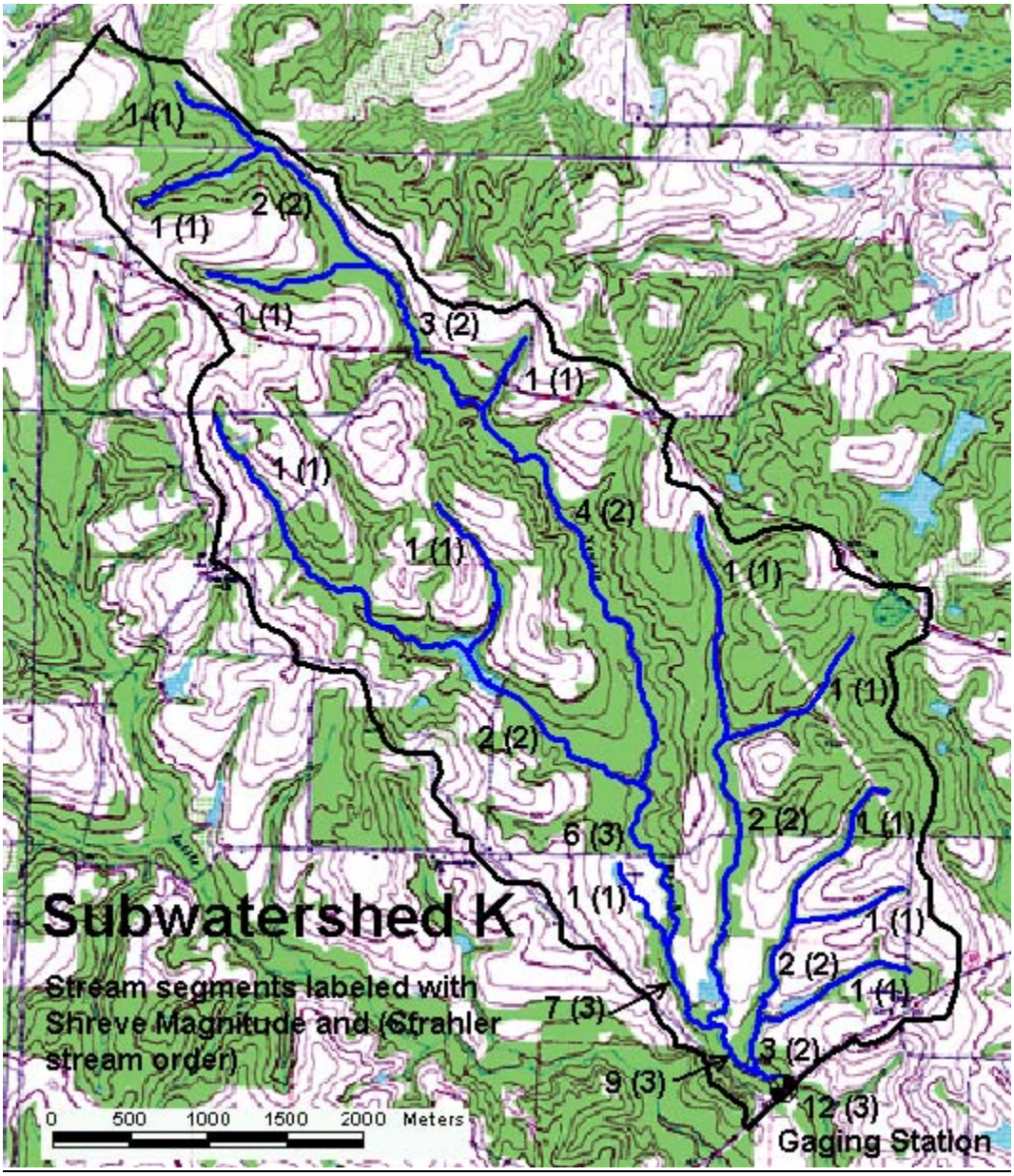


Figure 3.2. Subwatershed K Little River with stream segments digitized from the DRGs and labeled with the Shreve Magnitude and the Strahler stream order in parentheses. A small segment of the watershed boundary with topographic detail is depicted in Figure 4.1.

Georgia Piedmont watersheds

The primary goal in seeking test watersheds in the Piedmont region of Georgia was to find two similar watersheds that were also somewhat similar in size and configuration to Subwatershed K in the Coastal Plain. Two watersheds were sought in the Piedmont so that one could be used to compare and contrast the methodology of this study between it and the Coastal Plain watershed and the other could be used to test the results of the new classification process on a highly similar watershed. The West Fork Little River watershed was chosen as the Piedmont watershed to have ground truth data used to generate the watershed parameter values in a manner similar to that used on the Coastal Plain watershed. The East Fork Little River watershed was selected to test the results of the classification process used on West Fork Little River. The two Piedmont watersheds are shown in Figure 3.3 with each perennial stream segment labeled with both the Shreve Magnitude and the Strahler stream order.

The portion of West Fork Little River used as a test watershed lies mostly in Hall County, Georgia (86%), though small portions of it lie in White (12%) and Lumpkin (2%) Counties. The counties in which this watershed and East Fork Little River watershed lie became important in the acquisition and use of digital elevation, land cover, and soils data, as discussed in Chapter 4. The watershed is aligned north to south, with the general direction of drainage being to the south. Its shape is somewhat atypical, being wider at the downstream end than at the upstream end. The watershed, as defined in Figure 3.3, drains an area of 47.34 square kilometers (11,698 acres).

Eighty-eight percent of the East Fork Little River watershed is in Hall County and 12% is in White County. This watershed is also aligned north to south, with drainage

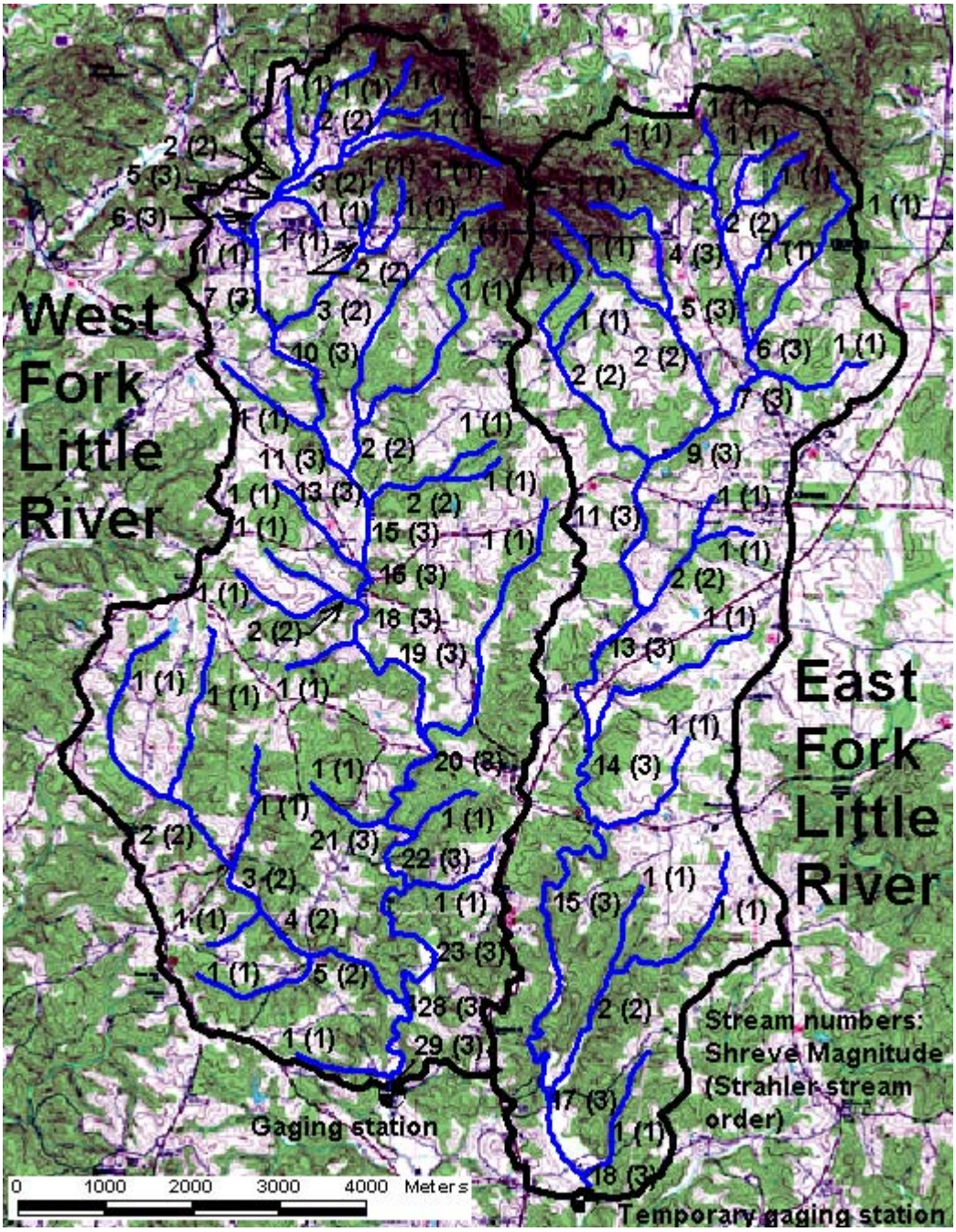


Figure 3.3. West Fork Little River and East Fork Little River watersheds with stream segments digitized from the DRGs and labeled with the Shreve Magnitude and the Strahler stream order in parentheses. A small segment of the two watershed boundaries with topographic detail is depicted in Figure 4.2.

being toward the south. The shape of this watershed is a bit more typical, exhibiting a slightly greater width near the upstream end than near the downstream end. This watershed drains an area of 36.37 square kilometers (8,987 acres), approximately 77% the size of West Fork Little River.

CHAPTER 4

METHODOLOGY

Overview

The overall approach of this research was to gather information on three test watersheds and use it to develop models of two of the watersheds in two different ways and then compare the results of the two approaches. The third watershed was then modeled using parameters developed in the modeling of one of the first two watersheds and the results of that modeling effort were compared to field-measured rainfall, runoff quantity, and runoff quality to assess the accuracy of the total modeling effort.

Topographic maps – paper and digital

Paper copies of USGS 7.5-minute topographic maps (quadrangles) were obtained covering Subwatershed K, West Fork Little River, and East Fork Little River. Subwatershed K covers an area that falls within four of these maps. West Fork Little River and East Fork Little River watersheds are covered by only two 7.5-minute quadrangles. Table 4.1 provides a breakdown of the distribution of Subwatershed K within the four quadrangles that cover the area. Table 4.2 provides similar information for the West Fork Little River and East Fork Little River watersheds.

Table 4.1. Distribution of Subwatershed K in four 7.5-minute quadrangles.

Portion of Subwatershed K in 7.5 min. quad:	Area, m^2	% of Subwatershed K	% of 7.5-minute quadrangle
Ashburn	16,238,080	95.8%	9.9%
Arabi	688,012	4.1%	0.4%
Pateville	21,039	0.1%	0.0%
Cuffietown	489	0.0%	0.0%
Total watershed	16,947,620	100.0%	10.3%

Table 4.2. Distribution of West Fork Little River and East Fork Little River in two 7.5-minute quadrangles.

Portion of each watershed in indicated 7.5-minute quad:	Area, m^2	% of watershed	% of 7.5-minute quadrangle
West Fork in Cleveland	8,190,064	17.3%	5.1%
West Fork in Clermont	39,145,080	82.7%	24.6%
West Fork Total	47,335,144	100.0%	29.7%
East Fork in Cleveland	7,435,821	20.4%	4.7%
East Fork in Clermont	28,934,963	79.6%	18.2%
East Fork Total	36,370,784	100.0%	22.8%

Digital raster graphic topographic maps (DRGs) served a variety of purposes in this project. The most important of these was to provide the topographic information

needed to delineate watershed boundaries for the three watersheds. At the outset of the project, it was anticipated that watershed boundaries would only be delineated from DRGs for the two Georgia Piedmont watersheds, West Fork Little River and East Fork Little River, since other researchers had already delineated the boundaries of Subwatershed K. Upon close examination, however, this watershed boundary was found to be much coarser (i.e. with the boundary defined by very long line segments) than could be achieved digitizing over the 7.5-minute quadrangle DRGs, so this watershed boundary was also digitized from the DRGs. A secondary purpose was to serve as a background for maps illustrating the features of the watersheds (see Figures 3.2, 3.3, 4.1, and 4.2).

The original DRGs used in the project for West Fork Little River and East Fork Little River were obtained from CD-ROMs distributed by the U.S. Geological Survey (USGS) (1996). Most of the drainage areas of both West Fork Little River and East Fork Little River are contained within the Clermont, Georgia 7.5-minute quadrangle, though the northern-most sections of both are mapped on the Cleveland, Georgia 7.5-minute quadrangle.

Since the USGS DRG CD-ROMs are no longer sold by USGS (USGS 2002), this option was not available for obtaining DRGs covering Subwatershed K. The Georgia GIS Data Clearinghouse (2002) was found to have most of the DRGs covering the state of Georgia, including those needed for this study, available for downloading free of charge. DRGs from this source had the added advantage of already being trimmed along the border of the map area, which makes creating a mosaic of the DRGs much simpler than if they first need to be trimmed. This option was adopted as the source for all DRGs

used in the study, including those covering West Fork Little River and East Fork Little River.

Metadata for the Georgia GIS Data Clearinghouse files, listed under "Data Preview", states that they were scanned from paper maps and projected to the Universal Transverse Mercator (UTM) coordinate system, majority zone (Georgia GIS Data Clearinghouse 2002). The metadata states that the datum of the DRGs match the datum of the source map, which is NAD 27 for all maps used in this study. The majority zone was taken to mean the UTM zone covering all or most of the DRG. All six DRGs used in the project, four covering Subwatershed K and two covering West Fork Little River and East Fork Little River, fell within UTM zone 17. Since the DRGs obtained from the USGS CD-ROM were not trimmed (i.e. borders had not been removed) and those obtained from the Georgia GIS Data Clearinghouse were trimmed, it was decided that it would be much simpler and as effective to use the Georgia GIS Data Clearinghouse DRGs for the Piedmont watersheds as well as for the Coastal Plain watersheds. It was not apparent at the time, but this decision led to a great deal of effort being expended on dealing with DRGs but it had the positive result of leading to the uncovering of a widespread error in Georgia GIS Data Clearinghouse DRGs. This is discussed further in the section, Georeferencing issues, later in this chapter. All DRGs obtained were in Tagged Interchange File Format (TIFF).

Subwatershed K is located almost entirely in the Ashburn, Georgia 7.5-minute quadrangle, specifically in the northeast quadrant, occupying approximately 10.3% of the quadrangle. The figures presented for the percent of each 7.5-minute quadrangle covered by Subwatershed K (last column) were based on the areas of each quadrangle. Because

they are nearer the equator and, therefore, slightly wider, the areas of the Ashburn and Cuffietown quadrangles are approximately fourteen-hundredths of one percent larger than the areas of the Arabi and Pateville quadrangles.

Because the DRGs have been projected from geographic coordinates (i.e. latitude and longitude) onto the rectangular UTM grid, the edges of the maps, which are aligned north-south-east-west, do not line up with the UTM grid. Furthermore, since the DRG pixels representing the scanned maps are stored in a rectangular digital array that coincides with the coordinates of the UTM grid, where x in the array represents easting in the UTM grid and y in the array represents northing in the UTM grid, the edges of the maps do not line up with the DRG array. Since trimming away the map borders involves removing rows and columns of the digital array, this presents a problem: dealing with a single map, one is faced with the choice of not trimming all the map margin if it is desired to retain all of the map area or with trimming away part of the map, if it is desired to remove all the map margin area. The DRGs distributed by the Georgia GIS Data Clearinghouse were produced by adding a small triangular strip from each of the four adjoining DRGs to expand the map into a rectangle aligned with the UTM grid and, therefore, aligned with the DRG's rectangular array data structure. This enlargement of each DRG had the added benefit in this project of eliminating one DRG from those needed to cover the area of Subwatershed K. The Cuffietown 7.5-minute quadrangle contains slightly less than three tenths of one percent of the Subwatershed K drainage area. This small area happened to be covered by the portion of the Cuffietown DRG that was added to the west side of the Georgia GIS Data Clearinghouse's Ashburn DRG to fill out the rectangular array aligned with the UTM grid. Because of this, the Ashburn,

Arabi, and Pateville DRGs were the only ones needed to create a single mosaic image to serve as a map background and as the source of topographic information for delineating the Subwatershed K boundary and perennial streams. Figure 4.1 depicts the boundaries of the 7.5-minute quadrangles covering the area of Subwatershed K.

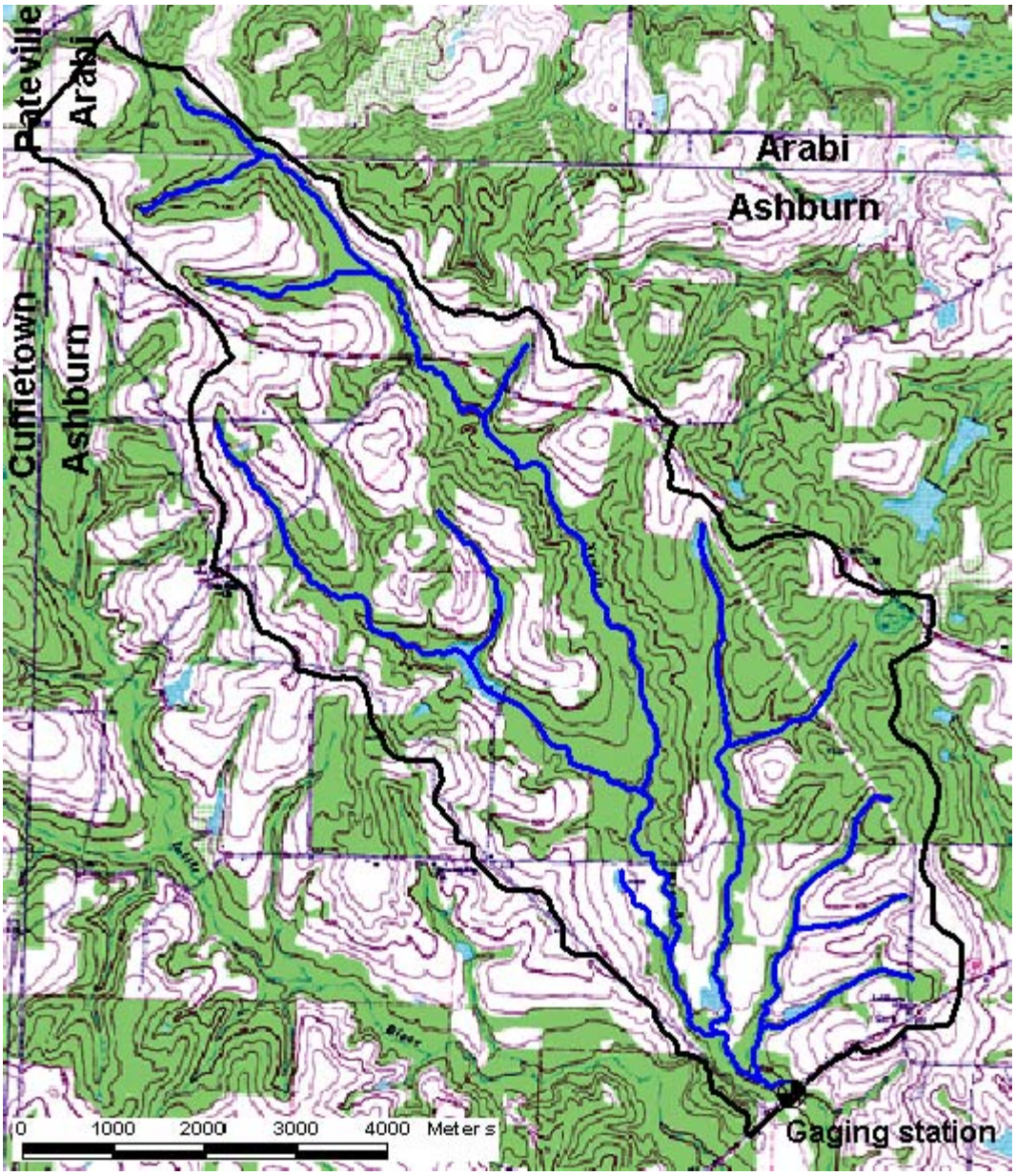


Figure 4.1. 7.5-minute quadrangles covering Subwatershed K area.

Since the DRGs were provided by the Georgia GIS Data Clearinghouse in TIFF format and ERDAS Imagine 8.5 was used to create a mosaic from them, it was necessary to convert the files to Imagine's ".img" file format. The resultant files displayed coordinates that appeared to be meters (consistent with the UTM norm) but indicated no projection information and "other" units when they were displayed in an Imagine Viewer. A mosaic of the three images was created using the mosaic function in Imagine. This image was then subset to just the area needed to cover Subwatershed K. Based on the metadata previously described, projection, datum, and units information were supplied for each file. With this option, it was possible to specify meters and UTM only, not the UTM zone or the datum, NAD 27. Once this was done, it was then necessary to specify the Clarke 1866 spheroid, NAD27 datum, and UTM zone 17. The same process was repeated to produce a subset of a mosaic of the Clermont and Cleveland DRGs for use with the West Fork Little River and East Fork Little River watersheds. The 7.5-minute quadrangles used to create the mosaic for mapping West Fork Little River and East Fork Little River watersheds are depicted in Figure 4.2. This process was accomplished without great difficulty. However, when the subset and mosaiced DRGs were put to use, significant problems arose. These problems and their solution are discussed in a following section, Georeferencing issues.

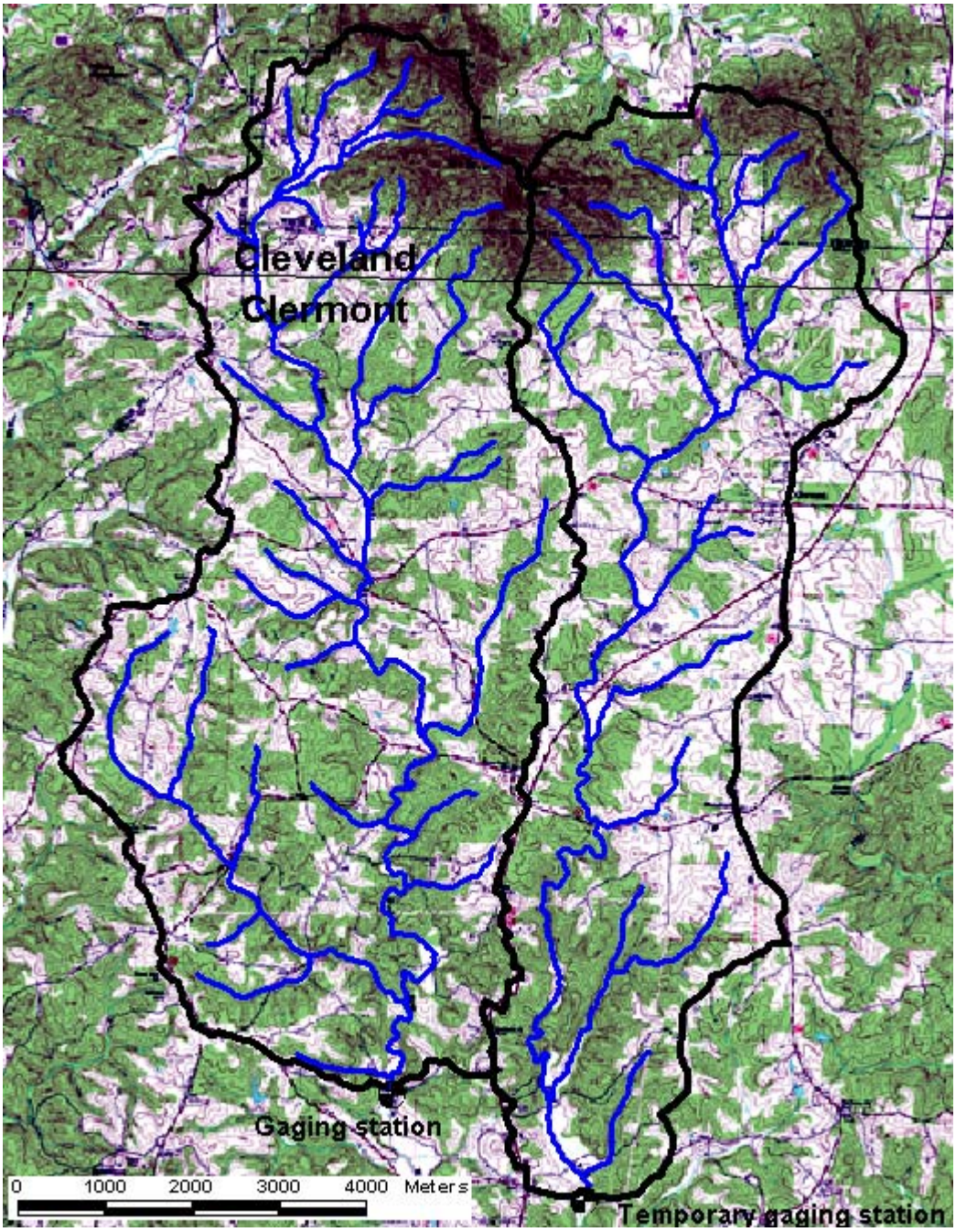


Figure 4.2. 7.5-minute quadrangles covering West Fork Little River and East Fork Little River watersheds. The black line at the quadrangle boundary was added for emphasis.

Delineation of watershed boundaries

As discussed in the section of this chapter on topographic maps – paper and digital, a digital file representing the boundary of Subwatershed K was acquired from the National Environmentally Sound Production Agriculture Laboratory (NESPAL) in Tifton, Georgia. This file was digitized at NESPAL at a scale of 1:125,000. When this file was opened in an Imagine Viewer superimposed upon the 1:24,000-scale DRG mosaic that was created for the Subwatershed K area, it was found to comprise line segments that were significantly longer than one would expect when digitizing at a scale of 1:24,000. Because of this, the 1:125,000-scale watershed boundary file failed to capture the fine detail of topography that should be achievable when digitizing over a 1:24,000-scale DRG. It was also apparent that the boundary deviated by approximately 200 meters, mostly north but also slightly east, from where the topography mapped in the 1:24,000-scale DRG mosaic indicated the boundary should be. This fact is discussed in more detail later in this chapter.

The watershed boundaries that were ultimately used in the project were digitized as polygons in Imagine. The subset and mosaiced DRGs for the area of the watershed were opened in a viewer, then a new vector file was created. Within the vector file, a polygon was digitized to represent the watershed boundary by analyzing the topography depicted in the underlying DRG. With experience, the interfluves (i.e. drainage divides) defining the watershed boundary can be visualized from the DRG. In so doing, the watershed boundaries for all three test watersheds were digitized. Because several problems arose in digitizing with Imagine, the details of performing these digitizing tasks are discussed in Appendix C. Figure 4.3 is an enlargement of the southern portion of

Subwatershed K. The details of the topography available in the DRG that was used to digitize the watershed boundary is clearly depicted in this figure. The ambiguity in the drainage that is apparent near the southernmost corner of the watershed was resolved by field observation of the drainage patterns in the area, which in this case were altered by human activity (road construction). Similar field observations were used to resolve questionable segments of the watershed boundaries for West Fork Little River and East Fork Little River. Figure 4.4 is an enlargement of the northeastern section of West Fork Little River watershed and the northwestern section of East Fork Little River watershed. This figure shows a portion of the topographic detail used to digitize the watershed boundaries.

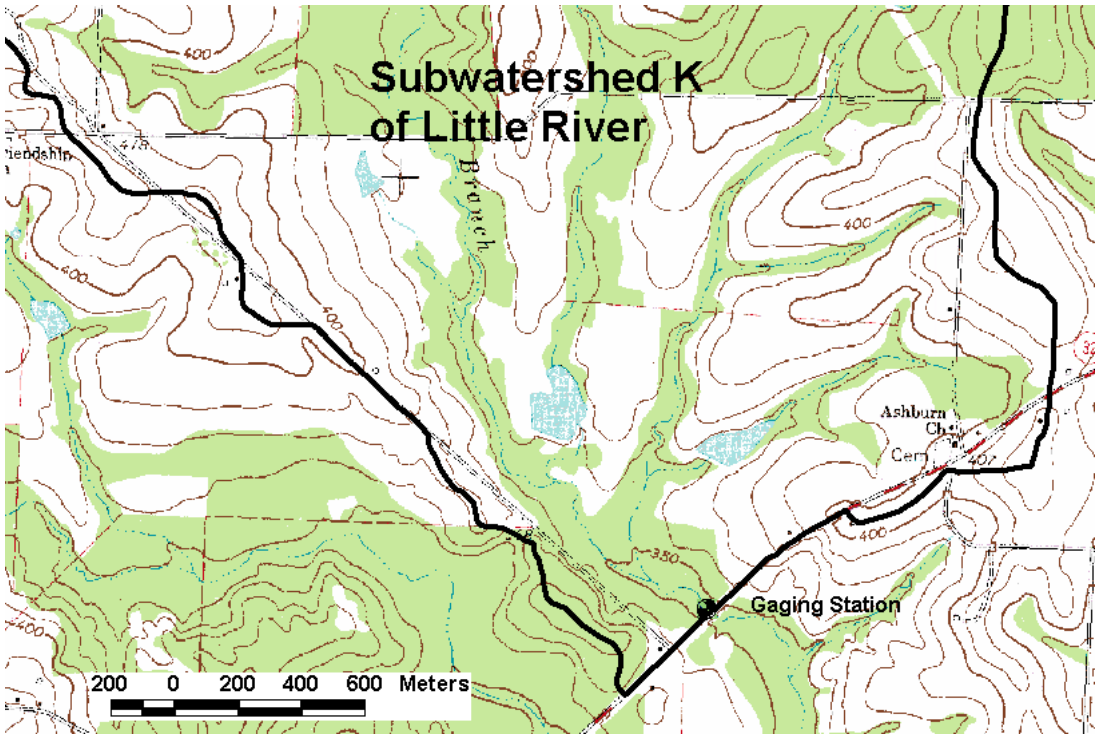


Figure 4.3. Enlarged detail of approximately the southern fourth of Subwatershed K, showing the topographic information from the DRG used to digitize the watershed boundary.

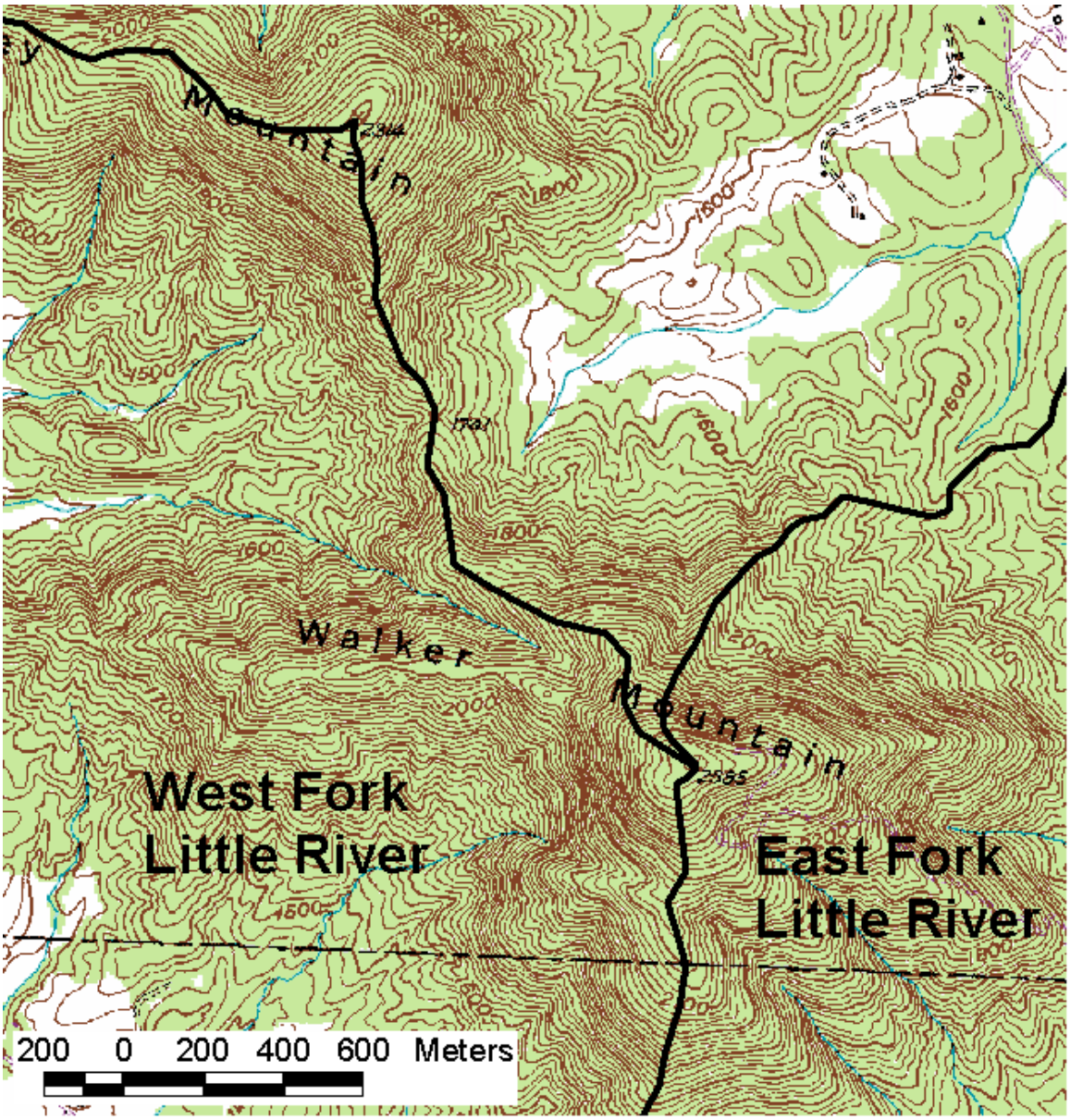


Figure 4.4. Enlarged detail of the northeastern section of West Fork Little River watershed and the northwestern section of East Fork Little River watershed, showing the detailed topographic information from the DRG used to digitize the watershed boundaries. "Walker Mountain", shown here, is known locally as "Wauka Mountain".

Recognizing that the actual flow path in the streams cannot be precisely depicted on a 1:24,000 scale map, care was taken to digitize all streams at the same display scale on the computer screen. Digitizing of streams was accomplished by displaying both the

watershed boundary of the stream under consideration and the DRGs upon which the watershed boundary was digitized. The screen resolution was maintained at a scale of 1:12,000 for all digitizing operations so that the departure of the digitized lines from the actual position of stream flow paths should average out to be approximately the same in all instances. Unfortunately, all digitizing had to be redone because it turned out that the DRGs that were used as the basis for digitizing perennial streams and watershed boundaries were not properly georeferenced.

Georeferencing issues

In order to check the differences between the newly digitized Subwatershed K boundary that was created with reference to the 1:24,000-scale DRGs and the one previously supplied by NESPAL, the new vector file was opened in a Viewer along with the DRG mosaic for the area and the watershed boundary vector file that had been received from NESPAL. While there were numerous differences in the details of the two watershed boundaries, presumably due to the difference in the scale of the DRGs that were employed as the source of topographic information used to visualize the watershed boundary location, the striking difference was the generally consistent north-south shift of approximately 200 meters between the two boundaries. After several attempts to find errors in the digitized vector files, the coordinates of points in the DRGs were checked against coordinates in copies of the paper maps from which the DRGs were originally scanned.

The paper copies of 7.5-minute topographic maps used in this study, Cleveland, Clermont, Ashburn, Arabi, Pateville, and Cuffietown, all have geographic coordinates (latitude and longitude) as well as UTM grid coordinates based on NAD27 marked along

the margins of the maps. To check UTM coordinates in the DRGs against coordinates on the paper maps, it was first necessary to accurately determine the UTM coordinates of well-defined points on the paper maps. To facilitate this, a grid was carefully drawn on each paper map connecting the UTM coordinate tic marks along the borders of the maps. Then, for each map, separate scales were determined for the easting and northing (x and y) coordinates, on the assumption that paper maps sometimes stretch slightly differently in different directions. The differences were quite small. For example, using a particular metric ruler that was judged to be quite uniform in its graduations, the northing scale on the paper Clermont 7.5-minute topographic map was found to be *1 cm represents 240.82 m*, while the easting scale on the same map was found to be *1 cm represents 240.24 m*. With the scale that was used, the best estimate of the precision to which these measurements could be made was +/- 0.2 mm, which equates to +/- 4.8 m on the ground.

The upper left corner of the Clermont 7.5-minute topographic map was checked first. The UTM Zone 17, NAD27 coordinates of this point were determined using five different methods: 1) by measurement relative to the UTM scale on the paper map, 2) by reading the coordinates of the cursor when it was placed over upper left corner of the Georgia GIS Data Clearinghouse DRG loaded in Imagine, 3) by reading the coordinates of the cursor when it was placed over upper left corner of the USGS DRG (from CD-ROM) loaded in Imagine, 4) by converting the latitude and longitude to UTM coordinates with Blue Marble Geographics Geographic Calculator 5.1, and 5) by converting the latitude and longitude to UTM coordinates with the Imagine "Tools – Coordinate Calculator..." functions but with the input datum reset to NAD27, as discussed below. The results of these measurements and calculations are summarized in Table 4.3.

Table 4.3. UTM coordinates of the upper left corner of the Clermont quadrangle.

Coordinate determination method	Easting, m	Northing, m	Easting error, m	Northing error, m
Measured on paper map	236020	3821147	0	0
GA GIS Data Clearinghouse DRG	236040	3821350	+20	+203
USGS DRG	236024	3821147	+4	0
Blue Marble Geographics Geographic Calculator 5.1	236020	3821154	0	+7
Imagine 8.5 Coordinate Calculator with datum reset to NAD27	236020	3821154	0	+7

Initially, the coordinates that were calculated with Imagine were done using the default spheroid and datum values supplied by Imagine's Coordinate Calculator. However, the use of default values was inappropriate in this case and it was necessary to reset the datum to NAD27. Once this was done, the coordinates agreed very closely with the coordinates measured from the paper map.

It should be noted that the first three coordinate determination methods listed in Table 4.3, measurement on a paper map and placing a cursor in the upper left corner of two different versions of the DRG, involved manual operations that were prone to human error. For that reason, none of the first three sets of coordinates in Table 4.3 should be considered precise. In fact, because the measurements on a paper map were taken to be the "true" values against which the others were judged and error values were calculated, the small errors associated with the USGS DRG, the Blue Marble Geographics

Geographic Calculator 5.1 calculations, and the ERDAS Imagine Coordinate Calculator calculations, are not significant.

The source of the errors in the UTM coordinates of the Georgia GIS Data Clearinghouse DRGs remain a mystery. All six of the Georgia GIS Data Clearinghouse DRGs were found to have similar errors. The corners of these DRGs were checked, as were several other points within the DRGs and all were found to have similar translation errors. The coordinates of the upper left corners of the six Georgia GIS Data Clearinghouse DRGs and the paper maps from which they came are shown in Table 4.4. The fact that the absolute values of the errors in these DRGs are very close to the absolute values of the errors in the coordinate conversions produced by the author's first, incorrect attempt to use ERDAS Imagine Coordinate Calculator may be significant. However, the fact that the signs of the errors are opposite argues against that implied relationship. This issue warrants further study but, since the underlying problem with Georgia GIS Data Clearinghouse DRGs did not bear directly on the outcome of this project once corrections to the DRGs used in the project were made, that investigation was set aside for later study.

Table 4.4. UTM coordinates of the upper left corners of the quadrangles used in this study. "GA GISDCH" refers to the Georgia GIS Data Clearinghouse.

Quadrangle	Paper quad easting, m	Paper quad northing, m	GA GISDCH DRG Easting, m	GA GISDCH DRG Northing, m	Easting error, m	Northing error, m
Ashburn	239492	3515831	239515	3516015	+23	+184
Arabi	239849	3529690	239862	3529878	+13	+188
Pateville	228022	3529988	228041	3530185	+19	+197
Cuffietown	227653	3516137	227667	3516322	+14	+185
Cleveland	236419	2835019	236437	3835218	+18	+199
Clermont	236020	3821147	236040	3821350	+20	+203

Landsat 7 Enhanced Thematic Mapper + data

Datasets from the Enhanced Thematic Mapper Plus (ETM+) sensor onboard the Landsat 7 satellite were acquired for both the Subwatershed K area and the two forks of the Little River north of Lake Lanier. Both datasets were from the passes of the satellite closest in time to the period of field data acquisition that produced cloud free images.

The image covering Subwatershed K, identified by USGS as Scene ID L71018038_03819991224, is an image of Path 018, Row 038, acquired December 24, 1999. This image comprises 7231 rows and 8051 columns of data at the nominal pixel size of 30 meters. The original image was projected onto the Universal Transverse Mercator (UTM) Zone 17 grid, based on the World Geodetic System of 1984 (WGS84) spheroid and the WGS84 datum. It was reprojected onto UTM Zone 17 based on the Clarke 1866 spheroid and the North American Datum of 1927 (NAD27) in order to be

consistent with maps produced by USGS, which are typically projected onto UTM, NAD27, Clarke 1866.

The image used in the mapping and analysis of West Fork Little River and East Fork Little River was Scene ID LE 7019036000108250 covering Path 019, Row 036. This scene of 6315 rows and 6864 columns of data was acquired March 23, 2001. It was projected onto UTM Zone 17, WGS84. However, when the projection was changed to UTM Zone 17, NAD27, Clarke 1866, it was found, by checking against paper maps, that the coordinates were in serious error. A representative point located within the watersheds of interest was found to be off by approximately 37.7 kilometers. The coordinates for the point in the ETM+ image were too far east by 23,958 meters and too far north by 29,089 meters. Because of these large errors, this image was rectified rather than reprojected. Ground control points were taken from USGS 7.5-minute topographic maps.

Tax maps and aerial photographs

Tax maps and aerial photographs were used to augment ground truth data, to assist in locating ground truth points in the field, to identify and verify apparent changes that occurred in the watersheds since the most recent topographic maps were made, and to help locate areas of interest on the Landsat ETM+ images. A set of four black and white paper prints of 1:7,920-scale tax maps were purchased from Turner County for use in Subwatershed K. The principal value of these maps was the aerial photographic base of the maps. They also proved valuable in locating landowners for the purpose of obtaining permission to enter private property. These four maps covered all of Subwatershed K and were used extensively during the fieldwork in that watershed.

A set of 32, 1:2,400-scale photographic prints, made from 1:12,000-scale nine-inch by nine-inch aerial negatives, and covering most of West Fork Little River and East Fork Little River watersheds, was generously provided by Hall County. These aerial photographs covered all of the two Piedmont watersheds that are in Hall County, the small segment of West Fork Little River watershed located in Lumpkin County, and most of the two watersheds in White County. A set of four, 1:4,800-scale tax maps were purchased from White County to cover the small area not covered by the aerial photographs. Like the Turner County tax maps, these were printed on aerial photographic bases.

Soils data

Soils data, in the form of digital vector files, were downloaded from the Georgia GIS Data Clearinghouse (2002b). Soils maps created by the U.S. Department of Agriculture (USDA) in the SSURGO program were downloaded for Turner, Hall, and White counties. The files downloaded were in Arc Interchange format (.e00) and were of necessity imported in the form of ArcInfo coverages, using Imagine's Import command. In addition, two soil survey books were obtained from the USDA, Natural Resource Conservation Service (NRCS) in Hall County. One book covers Barrow, Hall, and Jackson Counties (USDA 1977). The other covers Dawson, Lumpkin, and White Counties (USDA 1972). These books were used to correlate properties of interest in hydrologic modeling with soil types found in the SSURGO data.

Since digital soils data in the SSURGO dataset were not available for Lumpkin County, and a small portion of the West Fork Little River watershed falls in Lumpkin County, it was necessary to digitize portions of soils maps found in the soils survey

(USDA 1972, sheets 57 and 68). Lacking a digitizing tablet, a field-expedient digitizing method was devised for this purpose. Clear sheets of acetate were taped over the soils maps and soils polygons, roads, and other landmarks were carefully traced onto these sheets. The DRG mosaic that had been prepared for this area was then opened in an Imagine Viewer. Using road segments that appeared in both the DRG and the soils maps as references, the ratio of the scale in the paper soils maps (1:15,840) to the screen display scale of the DRG was calculated. Imagine's Scale Tool was then used to adjust the screen scale to that of the soils maps. The clear tracings were then overlaid on the computer screen, aligned with landmarks that were in both the tracings and the DRGs, and taped in place. Polygons were digitized on-screen to represent the soils mapping units depicted in the soils survey.

One disadvantage to this process was evident when the overlays were aligned with the on-screen DRG: the maps did not align precisely. Both the DRG used to create the overlays and the soils maps were projected using the UTM projection, based on NAD27, yet they quite obviously did not align correctly. While the reasons for this remain unknown, aligning the overlays for best fit in the center of the region of interest, then taking note of any discrepancies dealt with the situation. The majority of the polygons digitized appeared to be within one millimeter of their correct positions, based on the alignment of nearby landmarks. At a scale of 1:15,840, this corresponds to a ground distance error of 16 meters or just over half a pixel in the 30-meter resolution raster images used for this study. At the edges of the digitized areas, errors were estimated at up to three millimeters, corresponding to ground error of 48 meters or just over one and one half pixels.

Three soils properties were needed for use in the AGNPS model. These were hydrologic soil group, soil erodibility, and soil texture. These properties had already been determined by the Southeast Watershed Research Lab in Tifton, Georgia for use in the Little River watershed, of which Subwatershed K is a part, and were provide by Dr. David Bosch for use in this study. For the Piedmont watersheds, West Fork Little River and East Fork Little River, these properties were found in published sources or were calculated based on published properties of the soils. Hydrologic soil group designations of A, B, C, or D were found for the soils mapping units, based on their associated soils series, in Appendix B-1 of an erosion and sediment control publication of the Georgia Soil and Water Conservation Commission (1992). Soil texture values were assigned based on the descriptions of the soils mapping units in soils surveys, with integer values of 0, 1, 2, 3, or 4 corresponding to water, sand, silt, clay, or peat, respectively. Where the soils descriptions did not make it clear which texture dominated in a particular mapping unit, the soil triangle (Georgia Soil and Water Conservation Commission 1992, Appendix B-2) was used for guidance. Values of soil erodibility were determined using a nomograph produced for that purpose (Georgia Soil and Water Conservation Commission 1992, Appendix B-2; Wischmeier and Smith 1978). Additional soils properties required for use of the nomograph were obtained from the soils mapping unit descriptions in the soils surveys, with reference to the soil triangle. These were percent silt, percent sand, percent organic matter, soil structure, and permeability.

The soil mapping units were represented as polygons, in a vector data structure, in the digital SSURGO soils files obtained for the study areas. The polygons had attributes associated with them naming the soil mapping units in the source soils surveys from

which the digital files were created. Dr. E. Lynn Usery converted these vector files to raster images for use in this project. Both the soil mapping unit names and the polygon numbers from which they were derived became raster attributes in the soils raster images. Subsequent operations on the raster images using Imagine, including reprojecting into UTM NAD27, subsetting, and mosaicing, deleted the soil mapping unit designations from the raster attribute tables. The fact that the positions of all attributes in the raster attribute tables corresponded to the polygon numbers from which the attributes were derived was used to reestablish the soil mapping unit labels to the raster attribute tables for the soils images. In addition, columns were added to the attribute tables representing the polygon number, hydrologic soil group, soil erodibility, and soil texture. All these data were copied from the attribute tables and pasted into Microsoft Excel spreadsheets. In the spreadsheets, the data were sorted on the basis of soil mapping units, in alphabetical order. Excel's "if, then, else" function was used to select all soil mapping units from the list that had non-zero histograms, i.e. they were actually represented in the subset images immediately surrounding the study watersheds, and then all duplicates were eliminated. In the region surrounding Subwatershed K, 5,427 entries in the raster attribute table were reduced to 21 unique soil mapping units. In the West Fork Little River and East Fork Little River watershed area, 65 unique soil mapping units were extracted from 8,141 raster attribute table entries. The hydrologic soil group, soil erodibility, and soil texture properties were added in rows in the Excel spreadsheets associated with their appropriate soil mapping units. These values were then copied and pasted back into the full, sorted spreadsheet containing all the attribute table entries. Once this was accomplished, the spreadsheets were re-sorted on the basis of soil polygon

number, to get the data back in the same order as it was in the raster attribute table. Soil property values were transferred into the raster attribute tables using copy and paste.

Digital elevation data

Digital elevation data, in the form of compressed, 30-meter resolution raster files, (digital elevation models, DEMs) were downloaded from the Georgia GIS Data Clearinghouse (2002c). Separate files were obtained for Turner, Hall, Lumpkin, and White Counties. Although these data were produced by USGS and were available from that agency, it was simpler to obtain the data from the Georgia GIS Data Clearinghouse because datasets were readily available for individual counties.

Rainfall and runoff data

At the beginning of this project, it was intended that existing rainfall and runoff data would be supplied for Subwatershed K for comparison with AGNPS modeling results and that similar data would be gathered in the field for West Fork Little River and East Fork Little River. However, runoff data were not collected from Subwatershed K during the period when field data and Landsat ETM+ imagery were collected for use in this study. Lacking such data, a representative rainfall was assumed and used to compare the results of running AGNPS based on the two methods of generating AGNPS parameters for the watershed.

In order to collect usable data for West Fork Little River and East Fork Little River watersheds, it was necessary to monitor rainfall for a storm that could be expected to produce measurable response in the watersheds and also to monitor flow and selected water quality parameters exiting both watersheds. Six non-recording rain gages were installed in the two watersheds, one near the outlet of each watershed, one each near the

centroids of the watersheds, and one each near the headwaters of the watersheds. The locations of the six rain gages are shown in Figure 4.5. A recording rain gage was also installed next to the non-recording rain gage at the outlet of East Fork Little River. The purpose of these gages was to determine the distribution of rainfall over the two watersheds for the entire storm, with the intention of adjusting the recorded rainfall amounts if the storm proved to be significantly non-uniform across the watersheds.

Since a USGS monitoring station had been in existence and had been monitored for flow and served as a water quality sampling point near the most downstream point of West Fork Little River, this site was chosen as the monitoring site and the defining point of the West Fork Little River watershed. The staff gage and stilling well at this station are depicted in Figure 4.6. Unfortunately, it was discovered well into this project that USGS had ceased monitoring this station. It also proved impossible to obtain the rating curve linking flow to stream stage for the station, since the rating curve applied to the stage monitoring equipment as it had been installed and the equipment had been removed. Therefore, it was necessary to develop a new rating curve for the site, though the existing staff gage, shown in Figure 4.7, was used for that purpose.

Beginning in the late afternoon of March 28, 2001, just before an anticipated rain event began, and lasting into the afternoon of March 31, 2001, a day and a half after the rain ended, the stages of both West Fork Little River and East Fork Little River were monitored. This monitoring period coincided with a rise and subsequent fall and leveling of the stages in both rivers. The flow in both rivers was measured three times during this period over a wide range of stage and flow conditions. In both rivers, areas of the channels had previously been selected where it appeared that flow would be reasonably

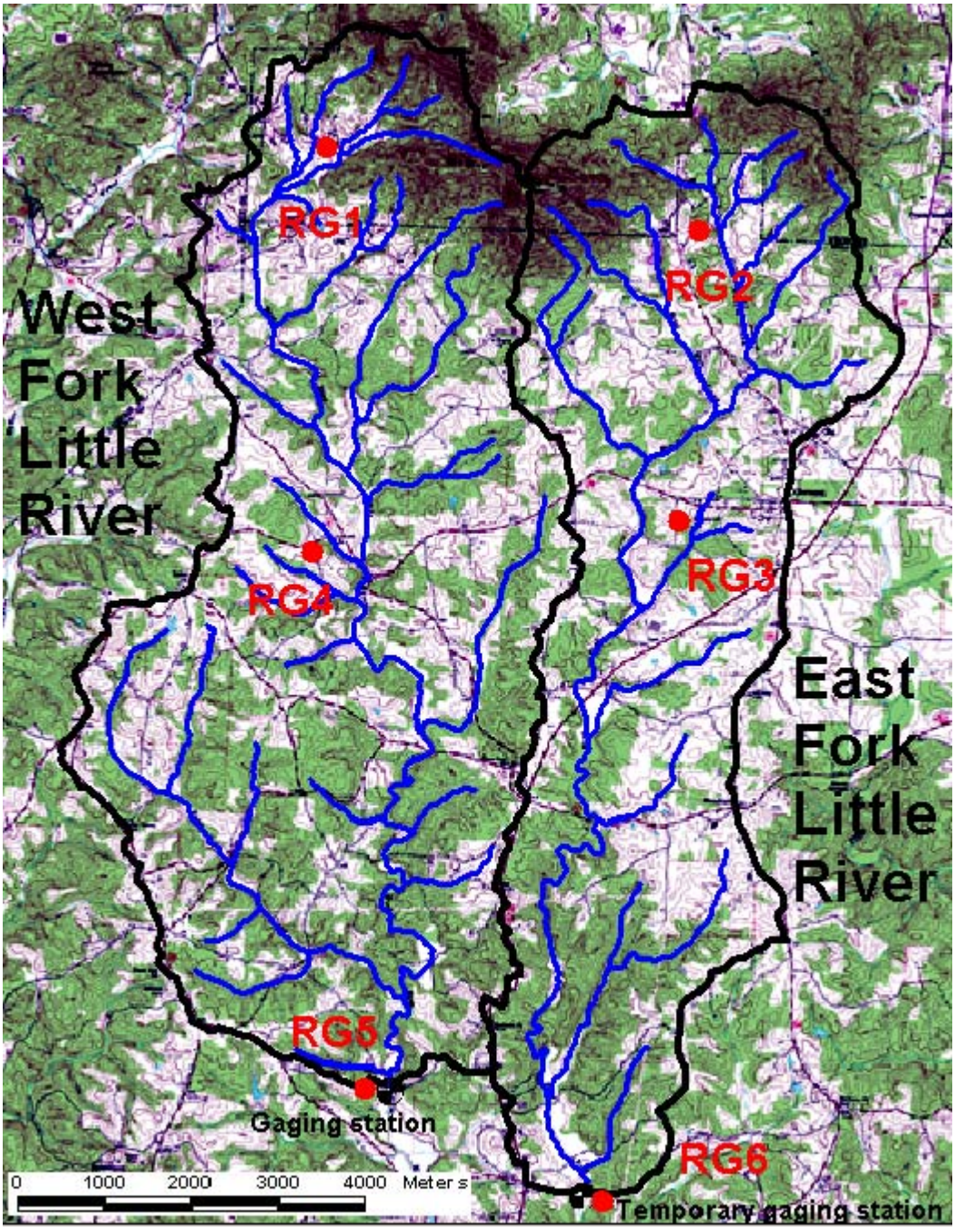


Figure 4.5. Locations of the six rain gages in West Fork Little River and East Fork Little River watersheds.

Figure 4.6. Staff gage used for determining flow in West Fork Little River at Jess Helton Road, attached to the southeast culvert wing wall by USGS.

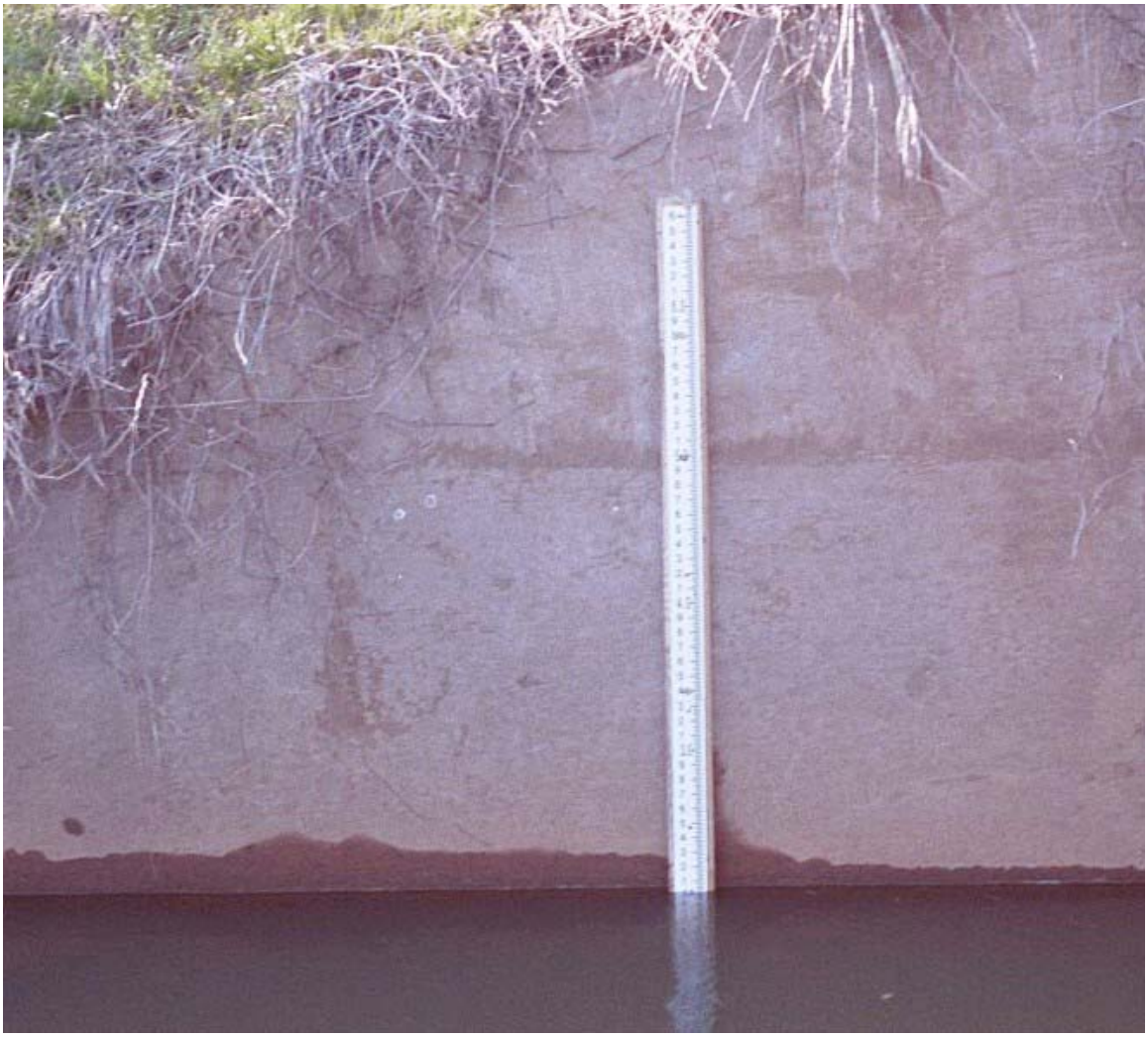


Figure 4.7. Staff gage used for determining flow in East Fork Little River at No Pone Road, attached to the southwest culvert wing wall by the author.

uniform, unaccelerated, and free of significant eddy currents. In both rivers, the selected cross sections were located three to four channel widths downstream of the nearby culverts. The procedure followed for each flow measurement at each sampling point was to divide the channel into small, relatively uniform subsections and record the distance from the bank, depth of water, and water velocity at the boundaries of each subsection.

This actually only required a number of velocity and depth measurements equal to the number of subsections plus one, since the end of one subsection would be the beginning of the next. Boundaries of subsections were selected where the channel bottom cross-slope changed noticeably or where it appeared that the water velocity changed significantly. The was to identify approximately trapezoidal submerged areas for which the average depth could be easily calculated from the depths at both sides and for which the average of the velocities at either side of the subsection reasonably represented the average velocity for the subsection. For each subsection, then, the flow was determined by

$$Flow = (W_2 - W_1)\left(\frac{d_1 + d_2}{2}\right)\left(\frac{v_1 + v_2}{2}\right)$$, where

W_1 and W_2 are the widths of the stream from the bank to the subsection boundaries, i.e. the points at which measurements were taken,

d_1 and d_2 are the depths of water at the subsection boundaries, and

v_1 and v_2 are the water velocities at the subsection boundaries.

Water velocity was determined by holding a propeller meter at a point approximately six tenths of the water depth above the river bottom, recording propeller turns for one minute, and reading the associated velocity from a rating curve provided by the propeller meter manufacturer. The flow for each river was determined by summing the individual flows calculated for all subsections at the monitoring site. Figure 4.8 is a photograph of the author collecting channel geometry and flow velocity data in East Fork

Little River during the rain event of March 28 and 29, 2001. The long pole leaning against a tree in the background was used to lower a one liter sample bottle to collect water samples and allowed moving the sample bottle up and down within the flow column. Measurements were taken at each site of the controlling channel area, which was the culvert under the nearby road in each case, and this was used to determine the stage at which flow would be zero. The three sets of stage and flow measurements for each monitoring site were then graphed using Microsoft Excel. These graphs were used in a linear regression to determine coefficients for equations of the form

$Flow = a(b - b_0)^n + c$, where

Flow is in cubic feet per second,

b is the stage in feet,

b_0 is the stage corresponding to zero flow,

n is an exponent adjusted by trial and error until the best regression fit was achieved and *c* is a constant determined in the regression.

For West Fork Little River, the rating equation was determined to be

$Flow = 81.01(b - 4.0)^{1.70} + 14.76$

For East Fork Little River, the rating equation was determined to be

$Flow = 39.89(b - 1.1)^{1.65} + 1.56$

Figure 4.8. The author collecting channel geometry and flow velocity data on March 29, 2001.

During the monitoring period 6:28 p.m. March 28, 2001 through 11:45 a.m. April 1, 2001, stage was monitored at each site and samples were collected approximately every hour as stage and or the apparent sediment load in the rivers changed rapidly. Monitoring was less frequent as changes in the rivers became more gradual.

Cumulative rainfall was also monitored at each river sampling point throughout the monitoring period and at the other four rain gages installed in the watersheds. A recording rain gage was also operated at the West Fork Little River monitoring site during the initial six hours of the rain event but the recording gage failed at that time.

Until its failure, the recording rain gage and the non-recording rain gage at the West Fork Little River monitoring site agreed to within 0.01 inch each time the gages were checked. As Table 4.5 indicates, the storm was actually quite uniform across the two watersheds, so no attempt was made to adjust the rainfall recorded at the two stream sampling points. The average of the rainfall recorded at the six gages was 1.44 inches. This uniformity was further confirmed by the fact that the rain gage installed at the Gainesville Airport, ID 093621, (latitude N34º 18', longitude W83º 52') recorded 1.39 inches of rain during the same period (NCDC 2002).

Table 4.5. Total rainfall recorded for the storm of March 29-March 31, 2001. Coordinates are UTM Zone 17, NAD27, Clarke 1866.

Gage	Watershed	Location, easting (m) and northing (m)	Rainfall, in.
RG1	WFLR	Poverty Acres west of Wauka Mtn. 239927, 3822590	1.33
RG2	EFLR	Whelchel Rd. approx 200m E of GA 284 244177, 3821467	1.44
RG3	EFLR	Hulsey Rd. approx. 280 m S of GA 283 243986, 3818145	1.49
RG4	WFLR	Elmer Truelove Rd. 90m SE of GA 283 239704, 3817800	1.43
RG5	WFLR	WFLR @ Jess Helton Rd. 240600, 3811595	1.48
RG6	EFLR	EFLR @ No Pone Rd. 242937, 3810329	1.48

Measurable rain began at about 5:30 a.m. March 29, 2001. About an hour later, the first discernible increase in the stage of the two rivers was recorded. Rainfall intensities were calculated at both Rain Gage 5 and Rain Gage 6 by dividing each incremental change in recorded rain by the time it took that rainfall amount to

accumulate. A graph of the rainfall intensity at Rain Gage 5, which is also representative of the pattern at Rain Gage 6, is presented in Figure 4.9. Graphs of the stages of West Fork Little River and East Fork Little River, as they responded to the rainfall, are presented in Figures 4.10 and 4.11, respectively. The data from which these graphs were made and other data reflecting the changes in flow and water quality at the two monitoring sites are presented in Tables 4.6 and 4.7.

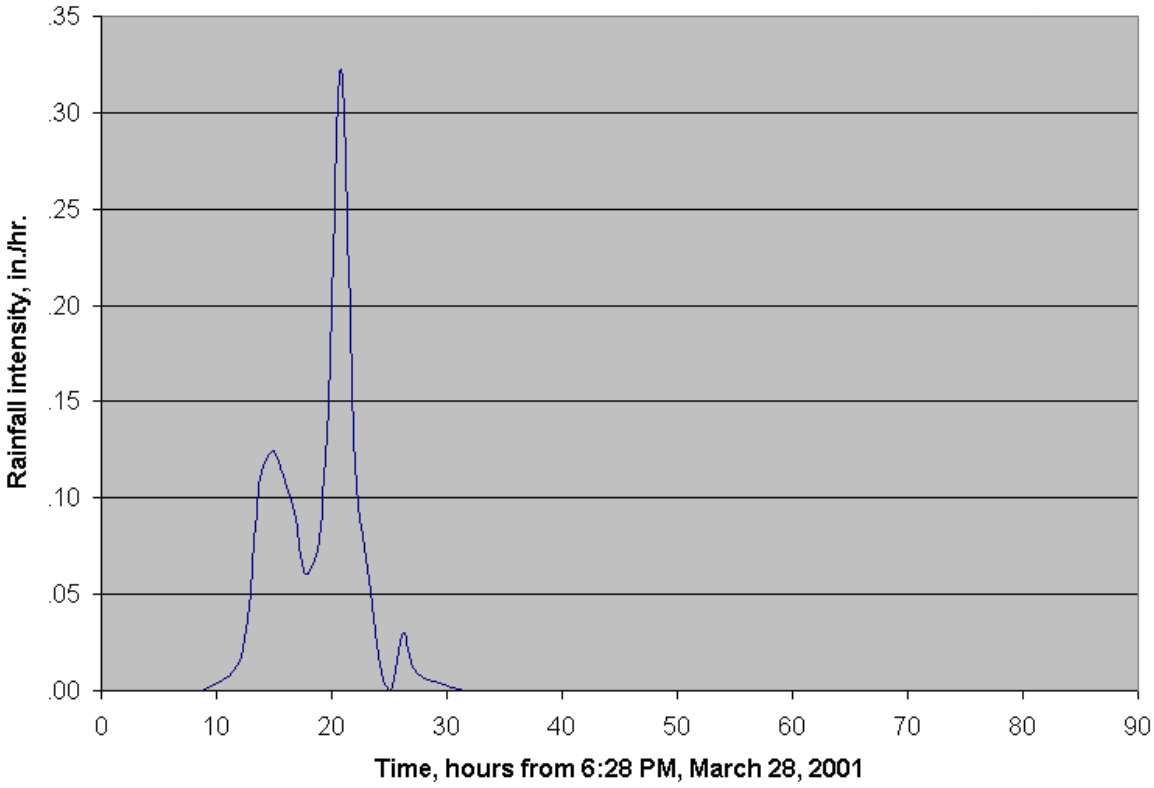


Figure 4.9. Graph of rainfall intensity vs. time at Rain Gage 5 near West Fork Little River at Jess Helton Road.

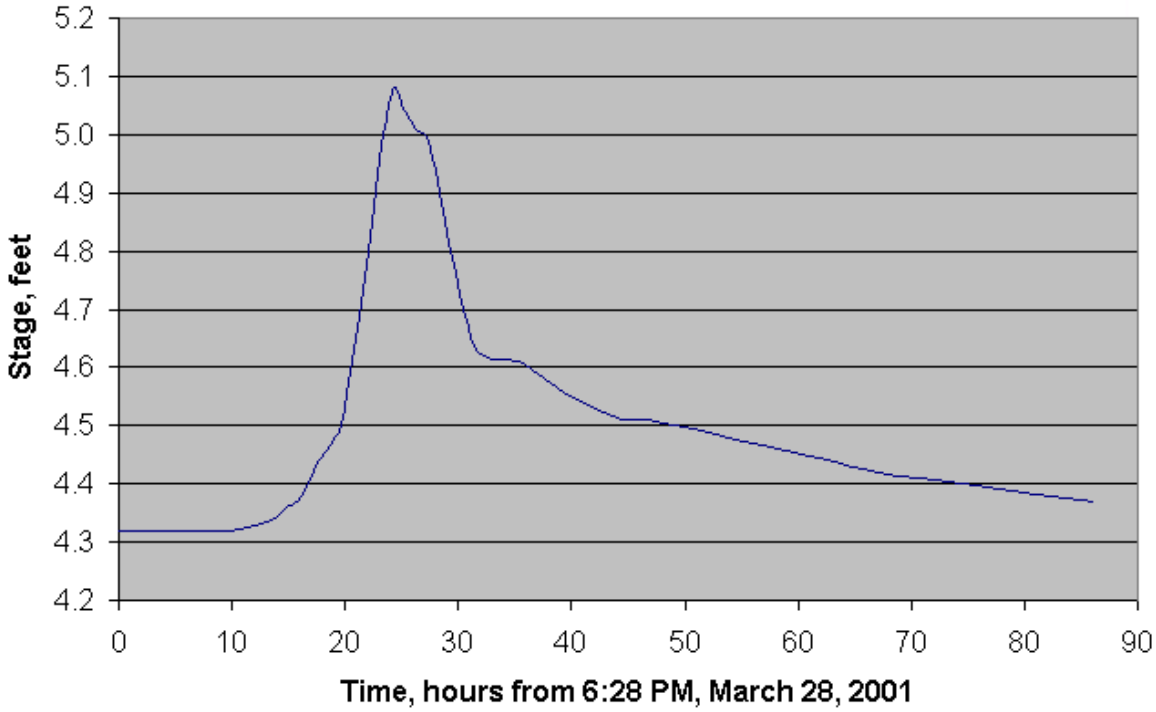


Figure 4.10. Graph of the West Fork Little River at Jess Helton Road water surface elevation (stage) response to the storm of March 29-30, 2001.

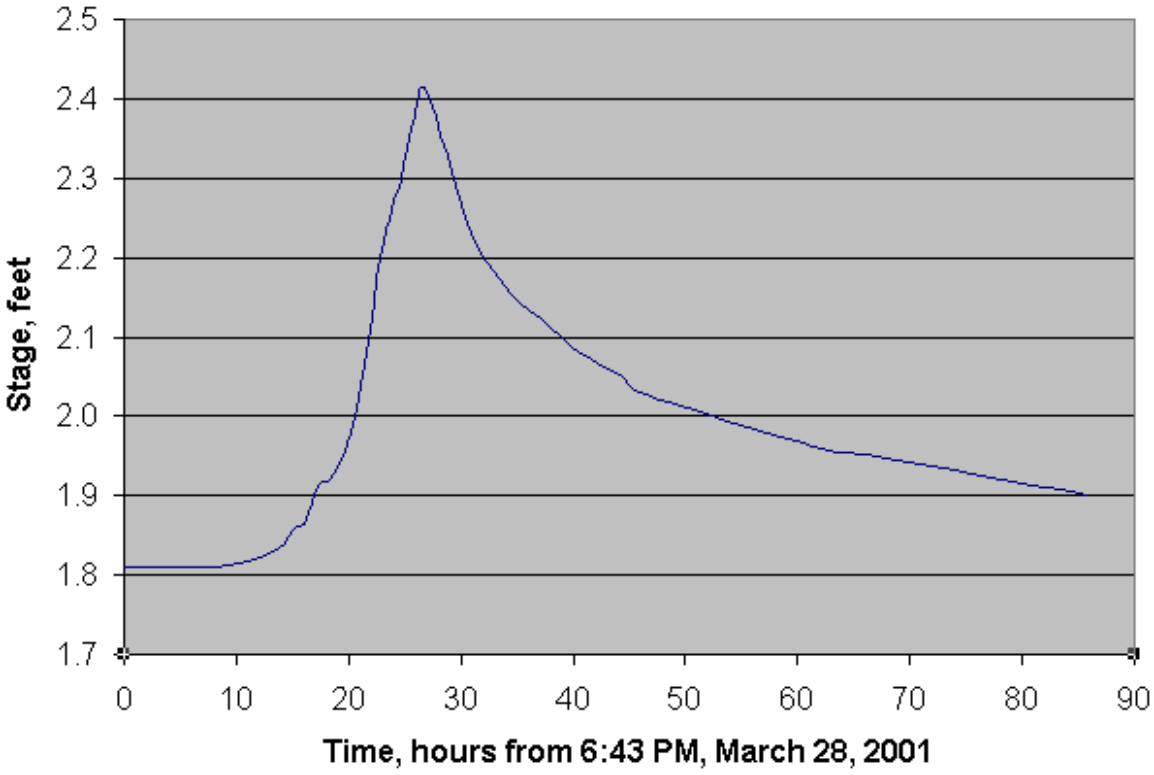


Figure 4.11. Graph of the East Fork Little River at No Pone Road water surface elevation (stage) response to the storm of March 29-30, 2001.

Two water samples were collected periodically at each of the monitoring sites, one for measurement of total suspended solids (TSS) and one for measurement of chemical oxygen demand (COD). The water samples were preserved in ice until they could be transported back to the author's laboratory for analysis. Samples collected for COD were also brought to a pH below 2.0 by adding 0.5 ml concentrated sulfuric acid to each 125 ml of sample.

The rating curves previously described were used to calculate flow from the stages measured at the two sampling points. The flow in each river prior to the storm was considered base flow, which is flow contributed by groundwater rather than surface runoff. In the eight days prior to the monitored storm, only 0.12 inches of rainfall was

recorded at the Gainesville Airport, approximately 21 kilometers south of the two Piedmont study areas (NCDC, 2002). For each time increment between samplings at each sampling point, the total flow was calculated and then the base flow was subtracted from this to yield the flow resulting from rainfall runoff (storm flow). The incremental storm flow numbers were then multiplied by the time during the increment to which they applied to get a time-weighted storm flow value for the increment. The applicable time increment in each case except the first and last was taken to be the period from midway between the time of the prior sampling and the current sampling to midway between the current sampling and the next sampling. For the first sampling at each site, the applicable time increment was taken to be half of the time between the first sampling and the second; for the last sampling at each site, the applicable time increment was taken to be half of the time between the next-to-last sampling and the last. Each of these numbers was ultimately divided by the total storm flow for the appropriate watershed to get a percentage contribution of the samples to be used to create composite samples for each watershed for TSS and COD measurement. Thus, the composites were composed of portions of the samples collected during the monitoring period proportional to the total volume of storm flow estimated to have flowed from the basin during the period represented by each water sample. Tables 4.6 and 4.7 summarize the data collected in the two Piedmont watersheds during the monitored storm.

Table 4.6. Field data for West Fork Little River.

Date	Time	Δt, min.	Stage, ft.	ΔStage, ft.	Rain, in.	ΔRain, in.	Sample #
10328	1828	0	4.32	0.00	0.00	0.00	36
10329	0312	524	4.32	0.00	0.00	0.00	72
10329	0636	204	4.33	0.01	0.06	0.06	78
10329	0819	103	4.34	0.01	0.25	0.19	69
10329	0917	58	4.36	0.02	0.37	0.12	77
10329	1016	59	4.37	0.01	0.48	0.11	71
10329	1114	58	4.40	0.03	0.57	0.09	66
10329	1213	59	4.44	0.04	0.63	0.06	76
10329	1319	66	4.47	0.03	0.71	0.08	83
10329	1408	49	4.50	0.03	0.83	0.12	82
10329	1515	67	4.63	0.13	1.19	0.36	73
10329	1632	77	4.80	0.17	1.34	0.15	35
10329	1744	72	5.00	0.20	1.41	0.07	27
10329	1845	61	5.08	0.08	1.42	0.01	28
10329	1940	55	5.04	-0.04	1.42	0.00	32
10329	2040	60	5.01	-0.03	1.45	0.03	25
10329	2150	70	4.99	-0.02	1.46	0.01	38
10330	0155	245	4.64	-0.35	1.46	0.00	44
10330	0540	225	4.61	-0.03	1.47	0.01	49
10330	0811	151	4.58	-0.03	1.47	0.00	42
10330	1133	202	4.54	-0.04	1.47	0.00	51
10330	1455	202	4.51	-0.03	1.47	0.00	39
10330	1705	130	4.51	0.00	1.47	0.00	54
10331	0855	950	4.44	-0.07	1.48	0.01	55
10331	1335	280	4.42	-0.02	1.48	0.00	47
10401	0840	1145	4.37	-0.05	1.48	0.00	45

Table 4.6, continued

Date	Time	Flow, cfs	Storm flow, cfs	Time-weighted total flow	Composite volume, ml	Comments
10328	1828	26.44	0.00	6926.23	35.36	Base flow
10329	0312	26.44	0.00	9622.70	49.13	Drizzle
10329	0636	27.06	0.63	4154.18	21.21	Light rain
10329	0819	27.70	1.27	2230.13	11.39	Medium rain
10329	0917	29.02	2.59	1697.93	8.67	Light rain
10329	1016	29.70	3.27	1737.71	8.87	Light rain
10329	1114	31.82	5.39	1861.61	9.50	Light rain
10329	1213	34.82	8.39	2176.47	11.11	Drizzle
10329	1319	37.20	10.77	2139.24	10.92	Medium rain
10329	1408	39.69	13.26	2302.24	11.75	Medium rain
10329	1515	51.69	25.26	3721.91	19.00	Medium rain
10329	1632	70.20	43.76	5229.60	26.70	Medium rain
10329	1744	95.77	69.33	6368.71	32.51	Drizzle
10329	1845	107.09	80.66	6211.42	31.71	No rain
10329	1940	101.36	74.92	5827.94	29.75	No rain
10329	2040	97.15	70.72	6314.88	32.24	Drizzle
10329	2150	94.40	67.96	14867.63	75.90	No rain
10330	0155	52.70	26.26	12383.42	63.22	No rain
10330	0540	49.72	23.29	9347.78	47.72	No rain
10330	0811	46.85	20.41	8268.98	42.21	No rain
10330	1133	43.18	16.74	8722.15	44.53	No rain
10330	1455	40.55	14.11	6730.87	34.36	No rain
10330	1705	40.55	14.11	21895.61	111.78	No rain
10331	0855	34.82	8.39	21416.49	109.33	No rain
10331	1335	33.30	6.86	23724.81	121.12	No rain
10401	0840	29.70	3.27	17005.83	86.82	No rain
				195880.65	1000.00	

Table 4.7. Field data for East Fork Little River.

Date	Time	Δt, min.	Stage, ft.	ΔStage, ft.	Rain, in.	ΔRain, in.	Sample
10328	1843	0	1.81	0.00	0.00	0.00	37
10329	0300	497	1.81	0.00	0.00	0.00	70
10329	0618	198	1.82	0.01	0.07	0.07	64
10329	0800	102	1.83	0.01	0.20	0.13	63
10329	0900	60	1.84	0.01	0.33	0.13	65
10329	1000	60	1.86	0.02	0.49	0.16	62
10329	1100	60	1.87	0.01	0.57	0.08	61
10329	1200	60	1.91	0.04	0.63	0.06	68
10329	1300	60	1.92	0.01	0.70	0.07	81
10329	1355	55	1.94	0.02	0.83	0.13	74
10329	1458	63	1.98	0.04	1.11	0.28	67
10329	1616	78	2.08	0.10	1.32	0.21	34
10329	1729	73	2.20	0.12	1.45	0.13	33
10329	1831	62	2.26	0.06	1.46	0.01	26
10329	1925	54	2.30	0.04	1.46	0.00	22
10329	2027	62	2.37	0.07	1.47	0.01	23
10329	2135	68	2.41	0.04	1.47	0.00	21
10330	0138	243	2.23	-0.18	1.47	0.00	43
10330	0525	227	2.15	-0.08	1.47	0.00	56
10330	0800	155	2.12	-0.03	1.47	0.00	41
10330	1119	199	2.08	-0.04	1.47	0.00	50
10330	1503	224	2.05	-0.03	1.47	0.00	30
10330	1630	87	2.03	-0.02	1.47	0.00	53
10331	0840	970	1.96	-0.07	1.48	0.01	29
10331	1348	308	1.95	-0.01	1.48	0.00	31
10401	0830	1122	1.90	-0.05	1.48	0.00	40

Table 4.7, continued

Date	Time	Flow, cfs	Storm flow, cfs	Time-weighted total flow	Composite volume, ml	Comments
10328	1843	21.11	0.00	5245.69	31.75	Base flow
10329	0300	21.11	0.00	7335.52	44.39	Drizzle
10329	0618	21.64	0.53	3245.80	19.64	Light rain
10329	0800	22.17	1.06	1795.99	10.87	Light rain
10329	0900	22.71	1.60	1362.69	8.25	Light rain
10329	1000	23.80	2.69	1428.20	8.64	Light rain
10329	1100	24.36	3.25	1461.38	8.84	Light rain
10329	1200	26.62	5.51	1596.90	9.66	Drizzle
10329	1300	27.19	6.08	1563.50	9.46	Light rain
10329	1355	28.36	7.25	1673.09	10.13	Light rain
10329	1458	30.74	9.63	2167.47	13.12	Medium rain
10329	1616	37.02	15.91	2795.18	16.92	Medium rain
10329	1729	45.12	24.01	3045.83	18.43	Light rain
10329	1831	49.40	28.29	2865.13	17.34	No rain
10329	1925	52.33	31.22	3035.18	18.37	No rain
10329	2027	57.62	36.51	3744.99	22.66	Drizzle
10329	2135	60.72	39.61	9442.25	57.14	No rain
10330	0138	47.24	26.13	11102.02	67.19	No rain
10330	0525	41.67	20.56	7959.75	48.17	No rain
10330	0800	39.65	18.55	7018.92	42.48	No rain
10330	1119	37.02	15.91	7830.20	47.39	No rain
10330	1503	35.09	13.98	5456.94	33.03	No rain
10330	1630	33.83	12.72	17878.31	108.20	No rain
10331	0840	29.54	8.43	18877.25	114.24	No rain
10331	1348	28.95	7.84	20697.39	125.26	No rain
10401	0830	26.04	4.93	14610.35	88.42	No rain
				165235.92	1000.00	

Flow-weighted composite water samples were created in the laboratory from the samples taken from West Fork Little River and East Fork Little River during the monitoring period. These samples were then analyzed for total suspended solids (TSS)

and chemical oxygen demand (COD). The total volume of flow in each river during the monitoring period was then multiplied by the TSS concentrations, with appropriate unit conversions, to obtain the total weight of suspended solid exiting each watershed during the monitoring period. These results are summarized in Table 4.8.

Table 4.8. Laboratory analyses of water samples

	Water volume exiting watershed, ft^3	COD, mg/l Flow-weighted average	TSS, mg/l Flow-weighted average	TSS weight, tons
West Fork Little River	11,753,000	11.2	37.2	11.5
East Fork Little River	9,914,000	7.5	25.0	7.7

Ground truth data acquisition

Ground truth data were gathered at all three study sites during field investigations of the study areas. This was accomplished during January 5-7, 2000 in Subwatershed K and on April 17, 2001 in West Fork Little River and East Fork Little River. The field-gathered data were later supplemented by ground truth data acquired from aerial photographs and ground-based photographs.

In the Subwatershed K field study, considerable effort was expended locating property owners and securing their permission to walk around their land, gathering information and taking photographs. The identity and contact information of landowners were obtained from the Turner County Courthouse, in the form of tax maps and printouts of tax records referenced to parcel numbers on the tax maps. The process of obtaining entry permissions was made more difficult by the fact that there had been a great deal of

discussion in the community in the months immediately prior to the field study regarding increased regulation of water use by the Georgia Department of Natural Resources Environmental Protection Division. This issue came up often in discussions with landowners, yet permission was secured to enter all private land within the watershed, except in the case of land owned by two absentee owners who could not be contacted.

Once access permissions had been obtained in Subwatershed K, ground truth data were gathered in essentially a semi-random, stratified sampling. The sampling was semi-random, in that the entire watershed was driven and/or walked generally by whim. However, the sampling was more closely stratified in nature because the maps of the watershed were consulted occasionally to ensure that representative data had been collected in all regions of the watershed and that, as far as could be ascertained, all land use/land cover classes were represented in the sampling.

In the field, an attempt was made to collect data that were, at each point, representative of an area larger than the area of a single Landsat 7 ETM+ pixel, in this case, thirty by thirty meters or 900 square meters. At each selected location, a one-square meter grid was placed on the ground and a stake was pushed into the ground near the center of the grid to provide scale for subsequent photographs. Typically, one or two photographs were taken at each point. These became valuable reference data later, when training sets were defined for supervised classification of the ETM+ imagery. Example photographs are shown in Figures 4.12 and 4.13.

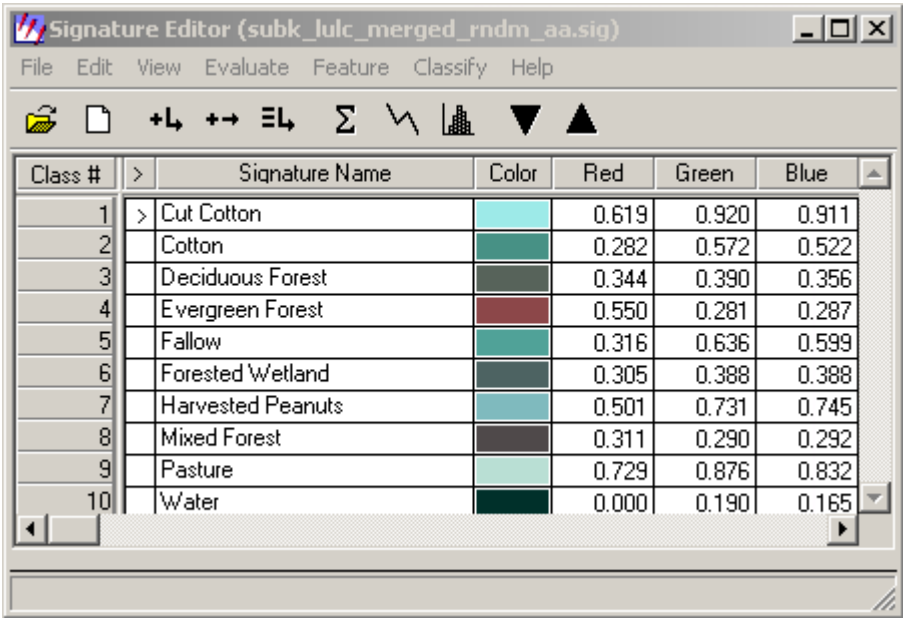
Figure 4.12. Vertical photograph at Subwatershed K ground truth point 4 with meter sticks placed in an approximate square for scale.

Figure 4.13. Annotated oblique photograph at Subwatershed K ground truth point 4.

A Magellan ProMARK X™ Global Positioning System (GPS) receiver was placed in the grid to collect data while other data were collected. The point identification was entered into the GPS receiver for later reference and, when the receiver indicated it was receiving signals from enough satellites to ensure a good positional fix, the data point location and the precise time were stored in the receiver's memory by the simple pressing of a key on the receiver. The time associated with each position was later used to apply differential GPS correction to the entire dataset. This process was performed with the gracious assistance of Mr. Chris Semerjian at Gainesville College in Oakwood, Georgia, who provided differential correction data collected at Gainesville College during the field sampling period. While the GPS receiver was collecting data, the point number and time were recorded and estimates were made and recorded of the SCS curve number (CN), Manning's n (n), the practice factor (P-fact), and the surface condition constant (SCC). These factors are described in detail in Chapter 2. In addition, a landuse/landcover (LU/LC) class was recorded for each ground truth point from among the classes listed in Table 4.9. The classes were selected to reflect the specific conditions found during the field investigation and, as much as was deemed reasonable, to reflect the nomenclature used by Young et al. (1994). The signatures associated with the landuse/landcover classes are shown in Figure 4.14. The LU/LC classes were used in an alternative method of classifying the Landsat 7 ETM+ images to generate input data for the AGNPS model. The Subwatershed K field data are presented in Table 4.10.

Table 4.9. LU/LC classes used in the Subwatershed K field study.

Symbol	LU/LC class
Ps	pasture
CC	cotton cut but not harrowed, stems still in ground
Cot	cotton not harvested, left in field
HP	harvested peanuts, not harrowed
EF	evergreen forest
DF	deciduous forest
MF	mixed forest
FW	forested wetland
F	fallow
W	water



Class #	Signature Name	Color	Red	Green	Blue
1	Cut Cotton		0.619	0.920	0.911
2	Cotton		0.282	0.572	0.522
3	Deciduous Forest		0.344	0.390	0.356
4	Evergreen Forest		0.550	0.281	0.287
5	Fallow		0.316	0.636	0.599
6	Forested Wetland		0.305	0.388	0.388
7	Harvested Peanuts		0.501	0.731	0.745
8	Mixed Forest		0.311	0.290	0.292
9	Pasture		0.729	0.876	0.832
10	Water		0.000	0.190	0.165

Figure 4.14. Landuse/landcover classes and their signatures derived from and used in Subwatershed K.

Table 4.10. Subwatershed K ground truth data.

Point	Date	Time	Easting, m	Northing, m	CN	n	P-fact	SCC	LU/LC
1	1-5-00	1340	244985	3510699	39	.13	1.00	.22	Ps
2	1-5-00	1413	244867	3510724	39	.13	1.00	.22	Ps
3	1-5-00	1437	244829	3510657	30	.20	1.00	.29	Ps
4	1-5-00	1505	244591	3510604	67	.06	1.00	.05	CC
5	1-5-00	1538	244447	3510454	67	.06	1.00	.05	CC
6	1-5-00	1557	244211	3510630	36	.25	1.00	.40	FW
7	1-5-00	1622	244166	3510801	67	.06	1.00	.05	CC
8	1-5-00	1646	244160	3511313	67	.06	1.00	.05	CC
9	1-5-00	1708	244638	3511116	67	.06	1.00	.05	CC
10	1-6-00	0848	244945	3510302	39	.13	1.00	.22	Ps
11	1-6-00	0913	244671	3510265	49	.06	1.00	.15	Ps
12	1-6-00	0925	244392	3510266	100	.99	0	0	W
13	1-6-00	0930	244583	3510382	36	.25	1.00	.40	FW
14	1-6-00	0950	244202	3510110	39	.08	1.00	.22	Ps
15	1-6-00	1005	244202	3510175	36	.25	1.00	.40	FW
16	1-6-00	1015	244208	3509977	36	.15	1.00	.29	EF
17	1-6-00	1028	244369	3509897	39	.08	1.00	.22	Ps
18	1-6-00	1035	244292	3509873	25	.20	1.00	.59	MF
19	1-6-00	1318	242532	3511300	67	.05	1.00	.05	CC
20	1-6-00	1335	242675	3511349	25	.20	1.00	.59	MF
21	1-6-00	1406	242178	3511420	67	.06	1.00	.15	CC
22	1-6-00	1430	242302	3511691	65	.06	1.00	.29	CC
23	1-6-00	1445	242432	3511721	67	.06	1.00	.05	HP
24	1-6-00	1458	242486	3511707	67	.05	1.00	.05	HP
25	1-6-00	1510	242643	3511618	67	.06	1.00	.05	HP
26	1-6-00	1525	242470	3511475	25	.20	1.00	.59	MF
27	1-6-00	1530	242481	3511387	25	.20	1.00	.59	DF
28	1-6-00	1552	241297	3512451	67	.06	1.00	.15	CC
29	1-6-00	1600	241219	3512689	67	.06	1.00	.15	CC
30	1-6-00	1615	239928	3515627	67	.05	1.00	.05	HP
31	1-6-00	1630	239697	3515743	67	.05	1.00	.05	HP
32	1-6-00	1635	239618	3515787	67	.06	1.00	.15	CC
33	1-6-00	1638	239588	3515879	67	.05	1.00	.05	HP
34	1-6-00	1647	239777	3516093	77	.06	.83	.22	F
35	1-6-00	1700	239737	3515942	25	.20	1.00	.59	MF
36	1-6-00	1717	239985	3515774	67	.05	1.00	.05	HP
37	1-6-00	1725	239984	3515843	36	.15	1.00	.29	EF
38	1-6-00	1740	240132	3515766	36	.15	1.00	.29	EF
39	1-7-00	0911	244745	3513008	36	.05	1.00	.29	EF
40	1-7-00	0930	244757	3512793	25	.30	1.00	.59	EF

41	1-7-00	1020	244548	3513057	67	.05	1.00	.05	CC
42	1-7-00	1030	244565	3512910	25	.15	1.00	.59	EF
43	1-7-00	1050	244418	3512913	25	.15	1.00	.59	EF
44	1-7-00	1103	244378	3512445	67	.05	1.00	.05	Ps
45	1-7-00	1135	244819	3511967	25	.30	1.00	.59	EF
46	1-7-00	1153	244220	3511785	67	.05	1.00	.05	CC
47	1-7-00	1230	244600	3513190	67	.04	1.00	.05	CC
48	1-7-00	1240	244672	3513163	36	.05	1.00	.29	EF
49	1-7-00	1312	243455	3513720	67	.04	1.00	.05	CC
50	1-7-00	1325	243380	3513746	36	.05	1.00	.29	EF
51	1-7-00	1335	243289	3513725	25	.30	1.00	.59	EF
52	1-7-00	1344	243167	3514232	67	.04	1.00	.05	CC
53	1-7-00	1358	243092	3514059	67	.05	1.00	.05	CC
54	1-7-00	1410	242994	3514030	67	.05	1.00	.05	CC
55	1-7-00	1434	242844	3514307	67	.04	.13	.05	CC
56	1-7-00	1450	242845	3514308	67	.04	.13	.05	CC
57	1-7-00	1510	242527	3513269	67	.04	1.00	.05	CC
58	1-7-00	1530	242626	3512868	77	.04	1.00	.22	F
59	1-7-00	1545	242792	3512530	67	.05	1.00	.05	CC
60	1-7-00	1550	242852	3512416	77	.04	1.00	.22	F
61	1-7-00	1604	242253	3513055	67	.05	1.00	.05	CC
62	1-7-00	1616	242142	3512928	67	.04	1.00	.05	Cot
63	1-7-00	1647	242625	3514164	67	.04	1.00	.05	F
64	1-7-00	1706	242047	3514655	67	.04	1.00	.05	HP
65	1-7-00	1720	241842	3514633	36	.25	1.00	.40	FW
66	1-7-00	1725	241601	3515214	67	.05	1.00	.05	CC
67	1-7-00	1728	241597	3515291	67	.05	1.00	.05	CC
68	1-7-00	1733	341329	3515519	67	.05	1.00	.05	CC
69	1-7-00	1740	241613	3514590	67	.05	1.00	.05	CC
70	1-7-00	1743	241000	3514600	67	.05	1.00	.05	HP
71	1-7-00	1746	240598	3514832	67	.05	1.00	.05	HP
72	1-7-00	1750	240166	3515283	67	.04	1.00	.05	HP
73	1-7-00	1758	241658	3514331	67	.04	1.00	.05	CC
74	1-7-00	1800	241713	3514199	67	.05	1.00	.05	CC
75	10-6-02	Map	243828	3510337	100	.99	0	0	W
76	10-6-02	Map	242345	3512343	100	.99	0	0	W
77	10-6-02	Map	242195	3512523	100	.99	0	0	W
78	10-6-02	Map	243861	3512915	100	.99	0	0	W
79	10-6-02	Map	244972	3512809	100	.99	0	0	W
80	10-6-02	Map	242090	3512841	67	.04	1.00	.05	Cot
81	10-6-02	Map	242090	3512780	67	.04	1.00	.05	Cot
82	10-6-02	Map	242841	3511445	25	.20	1.00	.59	DF
83	10-6-02	Map	242916	3511460	25	.20	1.00	.59	DF

Essentially the same process was employed in the West Fork Little River and East Fork Little River watersheds, though there were some differences. Primary among these was that many of the ground truth points were not directly visited but were, instead, observed from a distance, normally from a public road. This was because there were many times the number of landowners associated with the West Fork Little River and East Fork Little River watersheds than there were for Subwatershed K and many of these could not be contacted. Where landowners were contacted, most responded favorably to the request for permission to walk their property. Fortunately, the rolling terrain of the Piedmont watersheds made it rather easy to observe most of the watershed from a public road. Where this was necessary, the locations of ground truth points were recorded as azimuth and estimated offset distances from the GPS points collected and were also visually located on paper 7.5-minute topographic maps, from which the locations were later checked. The offset azimuth and distance values were used to calculate coordinates for the ground truth points by first converting the latitude and longitude values based on WGS 84 to rectangular UTM coordinates based on NAD 27 with Imagine's Coordinate Calculator, then using trigonometric functions to calculate the shifted coordinates. The trigonometric calculations were performed with Excel, using the following functions:

$X = X_0 + R\sin\theta$ and

$Y = Y_0 + R\cos\theta$, where

$X = ground\ truth\ point\ easting$,

$X_0 = base\ point\ (before\ offset)\ easting$,

$Y = ground\ truth\ point\ northing$,

$Y_0 = base\ point\ (before\ offset)\ northing$,

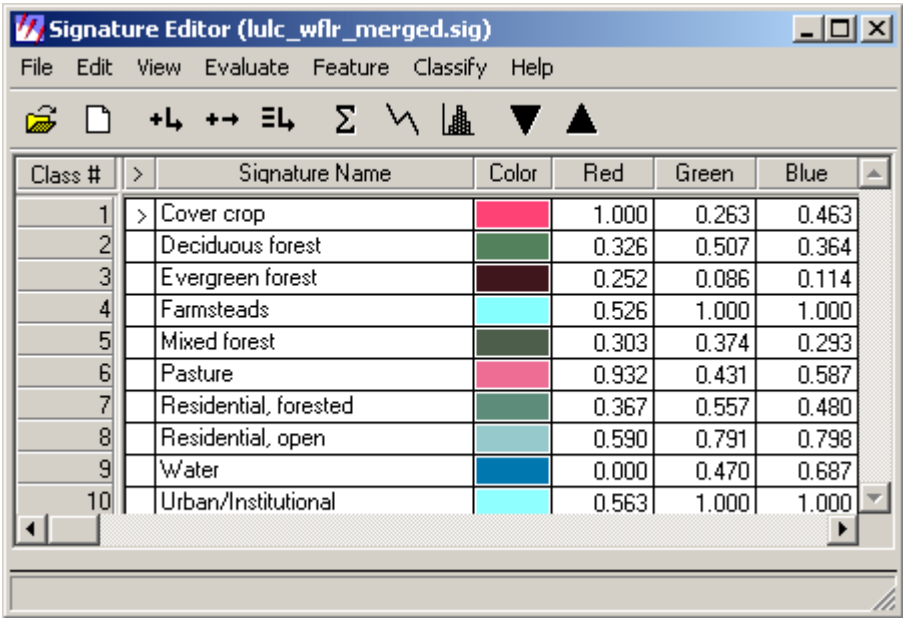
$R = offset\ distance$, and

$\theta = offset\ azimuth$

In addition, low altitude, large scale (1:2,500) aerial photographs were used to supplement the field observations in the West Fork Little River watershed. Both the coordinates and the sketches of ground truth areas on the maps were ultimately used in the satellite image classification process. The landuse/landcover classes and their signatures, derived from West Fork Little River watershed and used to describe West Fork Little River and East Fork Little River watersheds, are shown in Figure 4.15. Ground truth data collected in these watersheds are presented in Table 4.11. The last three columns in this table are discussed in the following section.

Another difference between handling of the field data associated with both West Fork Little River and the field data associated with Subwatershed K was that the GPS points collected in the Piedmont region were not differentially corrected, as they were for the Coastal Plain watershed. During the time between the field work in these two regions, selective availability, the mechanism that had been used by the U.S. Department of Defense to intentionally degrade the GPS signals available to civilians, was turned off. This removed the major source of error in positions determined using the GPS system, making differential correction unnecessary in the Piedmont. Also, since the raw GPS points were used for West Fork Little River and East Fork Little River watersheds and these points were based on the WGS 84 spheroid, it was necessary to use a different

transformation of the points to convert them to UTM coordinates. In contrast, the GPS points collected in Subwatershed K were referenced to NAD 27 during the differential correction process.



Class #	Signature Name	Color	Red	Green	Blue
1	Cover crop		1.000	0.263	0.463
2	Deciduous forest		0.326	0.507	0.364
3	Evergreen forest		0.252	0.086	0.114
4	Farmsteads		0.526	1.000	1.000
5	Mixed forest		0.303	0.374	0.293
6	Pasture		0.932	0.431	0.587
7	Residential, forested		0.367	0.557	0.480
8	Residential, open		0.590	0.791	0.798
9	Water		0.000	0.470	0.687
10	Urban/Institutional		0.563	1.000	1.000

Figure 4.15. Landuse/landcover classes and their signatures derived from West Fork Little River and used in both Piedmont watersheds.

Classification into land use/land cover classes

All three test watersheds were classified into land use/land cover classes with Imagine. Bands 1, 2, 3, 4, 5, and 7 were used with Imagine's maximum likelihood classifier. The results of these classifications are presented in Chapter 5.

Table 4.11. Ground truth sampling times, pixel locations, AGNPS parameter values, LU/LC classes, and cluster sizes for Subwatershed K.

Point	Date	Time	Easting, m	Northing, m	CN	n	P-fact	SCC	LU/LC	Use as	Pixels	Area, m^2
1	1/5/02	1340	244985	3510699	39	0.13	1.00	0.22	Ps	AA	10	9,000
2	1/5/02	1413	244867	3510724	39	0.13	1.00	0.22	Ps	Training	10	9,000
3	1/5/02	1437	244829	3510657	30	0.20	1.00	0.29	Ps	Training	8	7,200
4	1/5/02	1505	244591	3510604	67	0.06	1.00	0.05	CC	Training	50	45,000
5	1/5/02	1538	244447	3510454	67	0.06	1.00	0.05	CC	Training	21	18,900
6	1/5/02	1557	244211	3510630	36	0.25	1.00	0.40	Fw	Training	18	16,200
7	1/5/02	1622	244166	3510801	67	0.06	1.00	0.05	CC	Training	61	54,900
8	1/5/02	1646	244160	3511313	67	0.06	1.00	0.05	CC	Training	31	27,900
9	1/5/02	1708	244638	3511116	67	0.06	1.00	0.05	CC	Training	35	31,500
10	1/6/02	0848	244945	3510302	39	0.13	1.00	0.22	Ps	Training	68	61,200
11	1/6/02	0913	244671	3510265	49	0.06	1.00	0.15	Ps	Training	92	82,800
12	1/6/02	0925	244392	3510266	100	0.99	0.00	0.00	W	AA	26	23,400
13	1/6/02	0930	244583	3510382	36	0.25	1.00	0.40	Fw	AA	15	13,500
14	1/6/02	0950	244202	3510110	39	0.08	1.00	0.22	Ps	Training	13	11,700
15	1/6/02	1005	244202	3510175	36	0.25	1.00	0.40	Fw	Training	19	17,100
16	1/6/02	1015	244208	3509977	36	0.15	1.00	0.29	Ef	Training	10	9,000
17	1/6/02	1028	244369	3509897	39	0.08	1.00	0.22	Ps	AA	24	21,600
18	1/6/02	1035	244292	3509873	25	0.20	1.00	0.59	Mf	Training	14	12,600
19	1/6/02	1318	242532	3511300	67	0.05	1.00	0.05	CC	Training	9	8,100
20	1/6/02	1335	242675	3511349	25	0.20	1.00	0.59	Mf	Training	25	22,500
21	1/6/02	1406	242178	3511420	67	0.06	1.00	0.15	CC	Training	14	12,600
22	1/6/02	1430	242302	3511691	65	0.06	1.00	0.29	CC	Training	18	16,200
23	1/6/02	1445	242432	3511721	67	0.06	1.00	0.05	HP	Training	20	18,000
24	1/6/02	1458	242486	3511707	67	0.05	1.00	0.05	HP	AA	15	13,500
25	1/6/02	1510	242643	3511618	67	0.06	1.00	0.05	HP	Training	15	13,500
26	1/6/02	1525	242470	3511475	25	0.20	1.00	0.59	Mf	Training	8	7,200
27	1/6/02	1530	242481	3511387	25	0.20	1.00	0.59	Df	Training	5	4,500
28	1/6/02	1552	241297	3512451	67	0.06	1.00	0.15	CC	Training	13	11,700
29	1/6/02	1600	241219	3512689	67	0.06	1.00	0.15	CC	Training	24	21,600
30	1/6/02	1615	239928	3515627	67	0.05	1.00	0.05	HP	AA	24	21,600
31	1/6/02	1630	239697	3515743	67	0.05	1.00	0.05	HP	Training	12	10,800
32	1/6/02	1635	239618	3515787	67	0.06	1.00	0.15	CC	AA	12	10,800
33	1/6/02	1638	239588	3515879	67	0.05	1.00	0.05	HP	AA	14	12,600
34	1/6/02	1647	239777	3516093	77	0.06	0.83	0.22	F	Training	53	47,700
35	1/6/02	1700	239737	3515942	25	0.20	1.00	0.59	Mf	Training	40	36,000
36	1/6/02	1717	239985	3515774	67	0.05	1.00	0.05	HP	Training	12	10,800
37	1/6/02	1725	239984	3515843	36	0.15	1.00	0.29	Ef	AA	8	7,200
38	1/6/02	1740	240132	3515766	36	0.15	1.00	0.29	Ef	Training	20	18,000
39	1/7/02	0911	244745	3513008	36	0.05	1.00	0.29	Ef	Training	11	9,900
40	1/7/02	0930	244757	3512793	25	0.30	1.00	0.59	Ef	Training	20	18,000
41	1/7/02	1020	244548	3513057	67	0.05	1.00	0.05	CC	Training	15	13,500
42	1/7/02	1030	244565	3512910	25	0.15	1.00	0.59	Ef	Training	15	13,500
43	1/7/02	1050	244418	3512913	25	0.15	1.00	0.59	Ef	Training	18	16,200
44	1/7/02	1103	244378	3512445	67	0.05	1.00	0.05	Ps	Training	8	7,200
45	1/7/02	1135	244819	3511967	25	0.30	1.00	0.59	Ef	Training	114	102,600

46	1/7/02	1153	244220	3511785	67	0.05	1.00	0.05	CC	Training	31	27,900
47	1/7/02	1230	244600	3513190	67	0.04	1.00	0.05	CC	Training	16	14,400
48	1/7/02	1240	244672	3513163	36	0.05	1.00	0.29	Ef	Training	6	5,400
49	1/7/02	1312	243455	3513720	67	0.04	1.00	0.05	CC	Training	22	19,800
50	1/7/02	1325	243380	3513746	36	0.05	1.00	0.29	Ef	Training	14	12,600
51	1/7/02	1335	243289	3513725	25	0.30	1.00	0.59	Ef	Training	21	18,900
52	1/7/02	1344	243167	3514232	67	0.04	1.00	0.05	CC	Training	16	14,400
53	1/7/02	1358	243092	3514059	67	0.05	1.00	0.05	CC	Training	12	10,800
54	1/7/02	1410	242994	3514030	67	0.05	1.00	0.05	CC	Training	83	74,700
55	1/7/02	1434	242844	3514307	67	0.04	0.13	0.05	CC	Training	46	41,400
56	1/7/02	1450	242845	3514308	67	0.04	0.13	0.05	CC	AA	35	31,500
57	1/7/02	1510	242527	3513269	67	0.04	1.00	0.05	CC	Training	61	54,900
58	1/7/02	1530	242626	3512868	77	0.04	1.00	0.22	F	Training	34	30,600
59	1/7/02	1545	242792	3512530	67	0.05	1.00	0.05	CC	AA	10	9,000
60	1/7/02	1550	242852	3512416	77	0.04	1.00	0.22	F	Training	30	27,000
61	1/7/02	1604	242253	3513055	67	0.05	1.00	0.05	CC	Training	30	27,000
62	1/7/02	1616	242142	3512928	67	0.04	1.00	0.05	Cot	Training	8	7,200
63	1/7/02	1647	242625	3514164	67	0.04	1.00	0.05	F	Training	14	12,600
64	1/7/02	1706	242047	3514655	67	0.04	1.00	0.05	HP	Training	67	60,300
65	1/7/02	1720	241842	3514633	36	0.25	1.00	0.40	Fw	Training	14	12,600
66	1/7/02	1725	241601	3515214	67	0.05	1.00	0.05	CC	Training	31	27,900
67	1/7/02	1728	241597	3515291	67	0.05	1.00	0.05	CC	Training	18	16,200
68	1/7/02	1733	341329	3515519	67	0.05	1.00	0.05	CC	Training	15	13,500
69	1/7/02	1740	241613	3514590	67	0.05	1.00	0.05	CC	AA	33	29,700
70	1/7/02	1743	241000	3514600	67	0.05	1.00	0.05	HP	Training	55	49,500
71	1/7/02	1746	240598	3514832	67	0.05	1.00	0.05	HP	Training	34	30,600
72	1/7/02	1750	240166	3515283	67	0.04	1.00	0.05	HP	Training	27	24,300
73	1/7/02	1758	241658	3514331	67	0.04	1.00	0.05	CC	Training	11	9,900
74	1/7/02	1800	241713	3514199	67	0.05	1.00	0.05	CC	Training	61	54,900
75	10/6/02	Map	243828	3510337	100	.99	0	0	W	Training	9	8,100
76	10/6/02	Map	242345	3512343	100	.99	0	0	W	AA	29	26,100
77	10/6/02	Map	242195	3512523	100	.99	0	0	W	Training	33	29,700
78	10/6/02	Map	243861	3512915	100	.99	0	0	W	Training	33	29,700
79	10/6/02	Map	244972	3512809	100	.99	0	0	W	Training	19	17,100
80	10/6/02	Map	242090	3512841	67	.04	1.00	.05	Cot	Training	8	7,200
81	10/6/02	Map	242090	3512780	67	.04	1.00	.05	Cot	AA	8	7,200
82	10/6/02	Map	242841	3511445	25	.20	1.00	.59	DF	Training	5	4,500
83	10/6/02	Map	242916	3511460	25	.20	1.00	.59	DF	AA	8	7,200

Classification into AGNPS model parameter values

Classification of the ETM+ image into AGNPS parameter values was accomplished using the Imagine fuzzy classifier. The Coastal Plain watershed,

Subwatershed K, was used as the test bed for the entire classification and model development process. The process began with the identification of training sets to establish relationships between reflectance in the six bands of the ETM+ image that were being used and the values of AGNPS parameters that were estimated in the field. In Imagine, training sets are identified either as polygons or as area of interest (AOI) files. It was found that this could best be accomplished by simultaneously opening two viewers in Imagine. In one viewer, the DRG covering the Subwatershed K area was opened, along with the watershed boundary vector file. The Landsat 7 ETM+ image file was opened in the other viewer, also with the watershed boundary file. The two viewers were then linked on the basis of geographical coordinates allowing query of attributes at common locations in the two images. This resulted in the opening of a third window displaying the coordinates of the common point in the two viewers. This was possible because both images were georeferenced to the same coordinate system, based on the same spheroid and datum. Information provided by the linked query included the precise coordinates, in this case in UTM, Zone 17 coordinates, based on the Clarke 1866 spheroid and the NAD27 datum, of the map location of the cursors in the two viewers. A fourth window was opened to display a spreadsheet containing the coordinates of the ground truth data points that had previously been gathered in the study area. Figure 4.16 is a screen capture image of the four windows that were typically opened on the computer in order to locate field points and digitize AOIs representing each of the field locations. The UTM coordinates in the Inquire Cursor window at upper right match those for ground truth point PT1 in the upper left window.

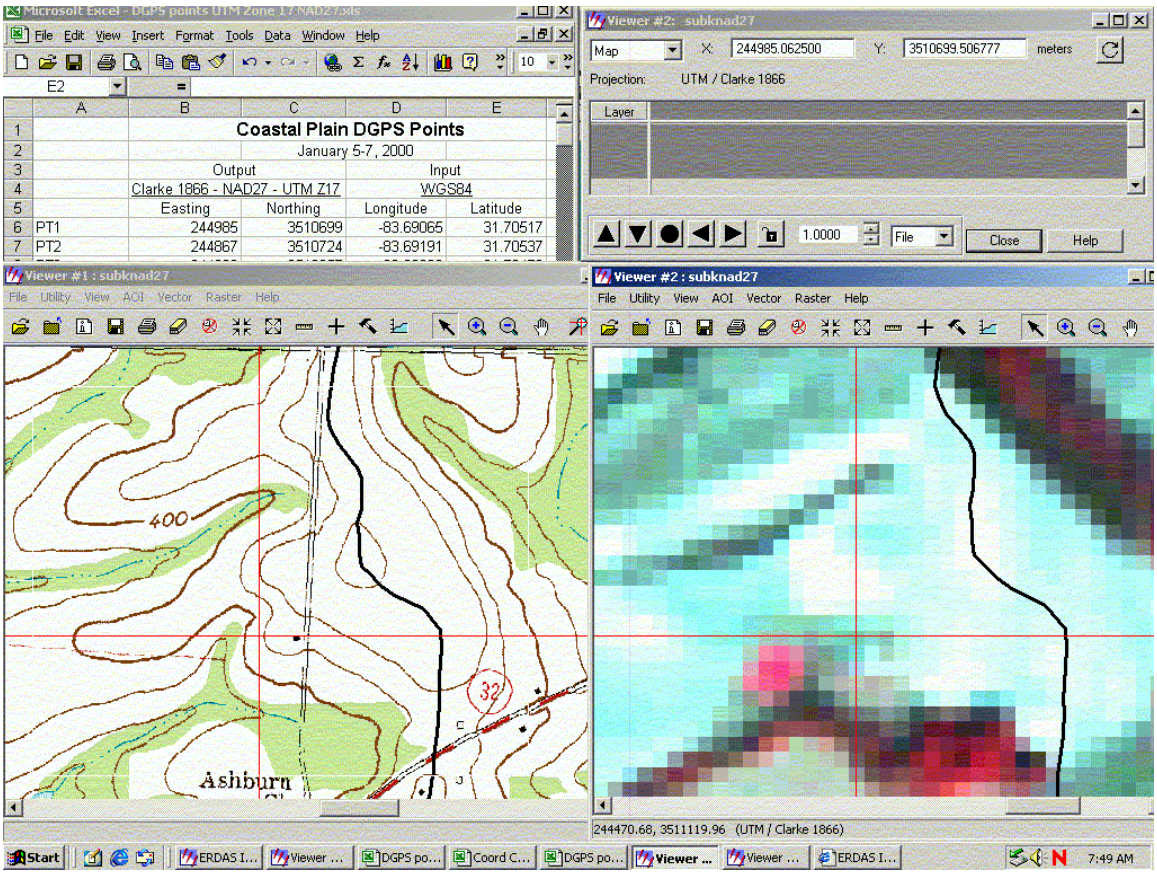


Figure 4.16. Linked Imagine Viewers, 7.5-minute DRG on the left and ETM+ imagery on the right, with Inquire Cursor active (upper right) to match image coordinates with the field-gathered, differentially corrected, GPS coordinates of ground truth data (upper left). The black line in both viewers is part of the Subwatershed K boundary.

AOIs were created for all ground truth data points, regardless of whether they were to be used as training sets or for later accuracy assessment work. The geographical extent and orientation of each AOI created to represent each ground truth point was determined through a process of evaluating DGPS data, field notes, photographs taken in the field, the Landsat ETM+ image, the DRG topographic map image, and a set of tax map aerial photographs that were purchased from Turner County. Using all of the clues available in the data that had been collected, AOIs were digitized for each ground truth point in the watershed. The site photographs, such as the examples shown in Figure 4.12

and 4.13, were very useful in this process. The sizes of the AOIs ranged from 4,500 square meters (five pixels) to 102,600 square meters (114 pixels). The extent of each AOI was determined on the basis of the best judgment that could be made based on the information collected in the field plus the data described above, of the area sharing the same characteristics. Table 4.10 lists pertinent characteristics of the Subwatershed K ground truth data points.

The designated usages of ground truth pixels in training sets or for accuracy assessment (AA) shown in Table 4.10 were determined using the random number generator function in Microsoft Excel. Imagine also has a random selection routine for accuracy assessment pixels but this routine selects from the entire dataset rather than from just those pixels that have been identified as ground truth pixels, so it was not suitable for this study. According to Congalton (1991), a minimum of 250 pixels in a class is needed to assess the mean accuracy of that class to within five percent. Since the numbers of pixels in classes were variable in this study and in many cases the number of pixels in certain classes was considerably less than 250, it was decided to randomly select a minimum of 250 pixels for accuracy assessment from the population of ground truth pixels and use this set for all accuracy assessment work. Since only 2094 total pixels were selected as ground truth pixels in Subwatershed K, 250 pixels would represent twelve percent of all ground truth pixels.

The accuracy assessment pixels were selected by first determining the number of pixels associated with each group of ground truth pixels that had been selected as representative of a ground truth point visited in the field or selected from maps, aerial photographs, and other ancillary data. Since there were 83 groups of ground truth pixels

in Subwatershed K, Microsoft Excel's random number generator function was used to generate a list of random numbers between one and 83. Duplicates were eliminated from the list and the numbers of pixels associated with ground truth points coinciding with the numbers in the randomly generated list were summed until a number greater than 250 was reached. In two cases, ground truth points were designated by the list of random numbers that would have eliminated all pixels of a certain class from the group available for use as training sets. These were eliminated from the pixels designated for accuracy assessment in order to preserve the classes associated with them in the training sets and, subsequently, in the classification. It turned out that 269 pixels were selected in order to exceed the chosen threshold of 250 pixels. The ground truth points (groups of pixels) selected in Subwatershed K for accuracy assessment were 31, 24, 81, 1, 69, 32, 37, 13, 76, 17, 56, 83, 59, and 12.

Once AOIs had been digitized, these were used to create signature files in Imagine representing the training sets that would be used for classification. Separate signature files were created for each parameter that was to be determined by classification of the ETM+ imagery, i.e. four files were used to create and store the training signatures for SCS curve number (CN), Manning's n (n), the practice factor (P-fact), and the surface condition constant (SCC).

For this operation, a viewer was opened in Imagine and the ETM+ imagery file was opened in the viewer. A Microsoft Excel spreadsheet was opened in a separate window. A spreadsheet was created displaying the ground truth points, their easting and northing coordinates, the values of the parameter for which a signature file was being created, and an indicator as to whether a particular point was to be included in the

signature file or to be reserved for later accuracy assessment of the classification procedure. To make the process easier, the data in the spreadsheet were sorted, first in ascending order on the basis of the parameter values, secondly on the basis of their intended use for accuracy assessment or training set data (AA or Training), and finally on the basis of point number. This facilitated convenient grouping of the signatures in the signature file. The Imagine Signature Editor was started, opening a window in which pertinent information about the signatures associated with the ground truth areas were stored and edited.

The AOIs representing the ground truth locations were then opened one at a time in the viewer. The AOI was selected in the viewer and the spectral information associated with that AOI in the ETM+ image was then added to the signature file as a signature entry. Each signature entry was renamed for its associated ground truth point number and the value associated with the point was reset to the parameter value for that signature. As subsequent AOIs were opened in the viewer, selected, and added as signatures in the signature editor, the viewer often became crowded with enough AOIs that it was difficult to keep track of which one was currently under consideration. Having the easting and northing coordinates of each ground truth point displayed in the spreadsheet significantly reduced this confusion. This allowed the location of each AOI, as determined by placing the cursor within the boundary of the AOI, to be checked against the values in the spreadsheet. Occasionally, confusion was still difficult to avoid, so all AOIs were cleared from the viewer before any additional AOIs were opened. Figure 4.17 is a screen capture image of the windows that were open during the process of creating a signature file for CN values in Subwatershed K.

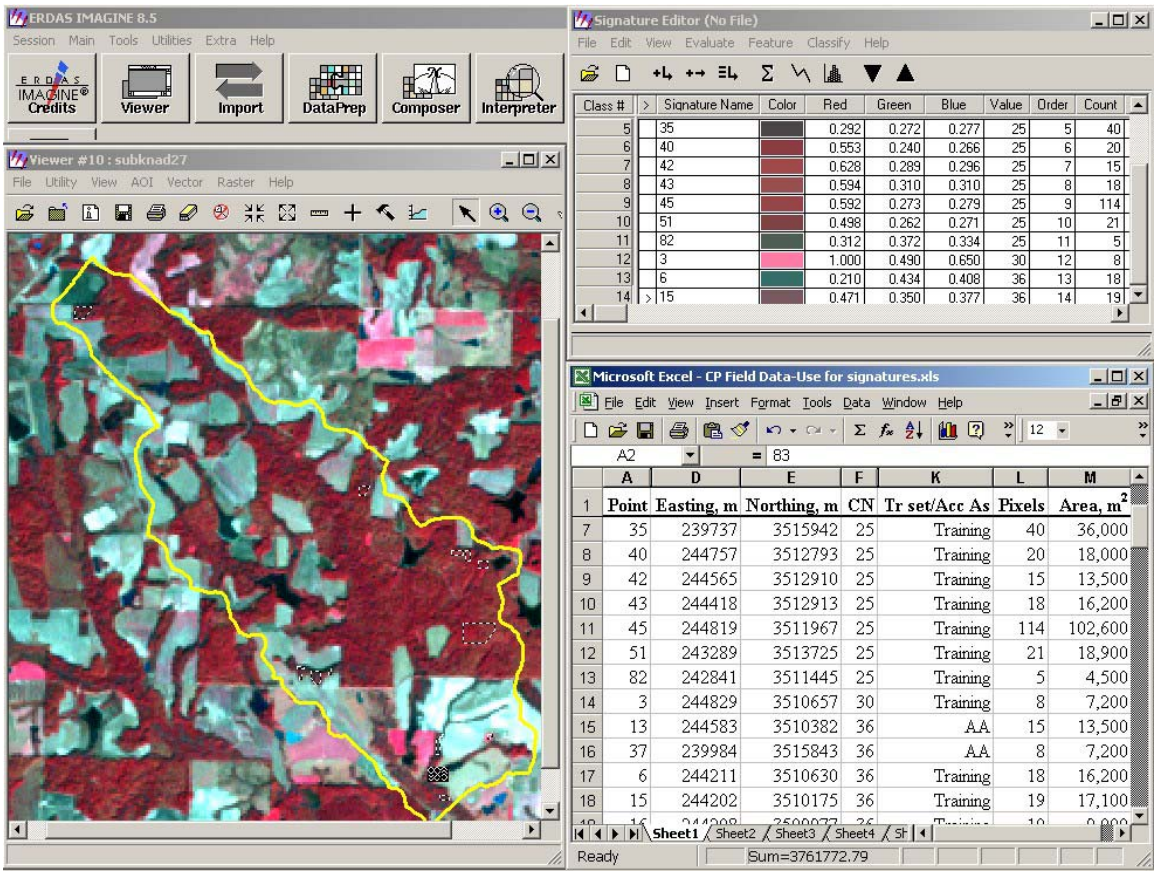


Figure 4.17. Screen capture of the three windows open during the collection of training signatures in Subwatershed K. Training set AOIs can be seen in the left window as small black and white outlined objects. "AA" in the lower right window identifies training sets reserved for accuracy assessment.

An alternative method of organizing signatures in the Signature Editor was tried, which proved to be more convenient and more representative of the watershed as a whole than saving the signatures of each training set individually. This involved opening all AOIs associated with a single value of the parameter under consideration in the viewer, then rapidly adding these signatures to the signature file, without regard to keeping track of which AOI corresponded to a particular signature. This was acceptable because only those AOIs with identical parameter values would be open at any one time. The

signatures thus added to the Signature Editor window were then selected and combined into a composite signature. Figure 4.18 is a side-by-side comparison of a portion of the Signature Editor window displaying the individual signatures associated with all AOIs that were individually associated with ground truth points and the Signature Editor window displaying the merged signatures for all assigned values of CN. The signatures, i.e. the displayed values for red, green, and blue, for each merged signature are the weighted averages of all the signatures for the individual AOIs having the same parameter value. That is, the signatures for CN 25 on the right side of Figure 4.18 are, in each case (red, green, and blue), the sum of the products of the signature values times the count (associated number of pixels) for all signatures having a value of 25, divided by the sum of the counts of all the signatures having a value of 25. For example, if

$CN25_{red}$ = the merged red band signature for all ground truth pixels for which CN equals 25, then

$$CN25_{red} = \frac{(.427)(14) + (.190)(8) + (.377)(5) + (.292)(40) + (.553)(20) + (.628)(15) + (.592)(114) + .498)(21)}{(14 + 8 + 5 + 40 + 20 + 15 + 114 + 21)}$$

or

$$CN25_{red} = .504$$

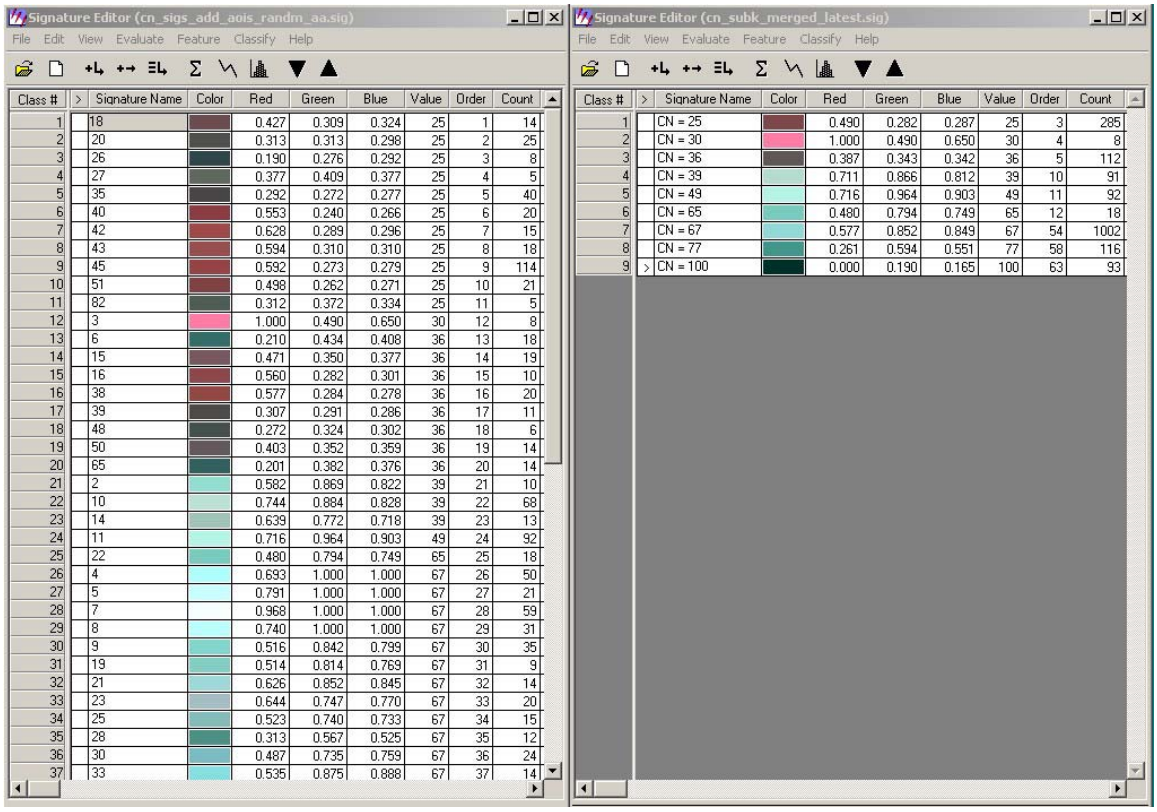


Figure 4.18. Comparison of individual training signatures with merged training signatures.

Although this is discussed in more detail later in this chapter, it should be mentioned here that the results of fuzzy classification using these two strategies for organizing ground truth signatures in the Signature Editor were quite similar but not identical. It was decided, more on philosophical terms than on the basis of rigorous analysis, that classification based on the merged signatures was more representative of the watershed as a whole than classification based on individual ground truth signatures. Since one of the major goals of this project was to develop a means of accurately representing the behavior of a watershed, taken as a whole, the method of classification based on merged signatures was adopted as the preferred method.

The signature file that was developed for CN values in the watershed was then used to classify the Landsat ETM+ image, using bands 1, 2, 3, 4, 5, and 7. Only the portion of the ETM+ image contained within the watershed boundary, as defined by the AOI that was digitized for that purpose, was classified. Classification was performed using the maximum likelihood fuzzy classification technique. This technique returns two or more bands in the classified image, where the first band contains pixel values calculated to be the most likely value for the pixel, based on the signatures used in the classification, the second band contains the next most likely of the available values, the third band contains the third most likely values, etc. The number of bands in the fuzzy-classified image is selected by the software user and was set at three for this study.

Figure 4.19 depicts side-by-side screen capture images of Imagine Viewers displaying a false color infrared image (bands 2, 3, and 4 displayed as blue, green, and red, respectively) of the Landsat ETM+ image, with the watershed boundary shown as an AOI, and the fuzzy classified image. In the fuzzy classified image, the pixel values in each layer were set equal to the CN value associated with the signature values assigned to that layer by the classifier. For instance, the values in layer 2 were set to the CN values second most likely to be associated with each pixel by the classifier. Figure 4.20 shows pixel values for the three layers in a small segment near the upper left corner of the fuzzy classified image of CN values. This small segment includes some pixels outside the Subwatershed K AOI, which became important in the development of the post-classification model that was used for subsequent processing. The same segment of the image, comprising 66 pixels from row 10 to row 20 and from column 3 to column 8, inclusive, was used for comparison of the inputs and outputs of the six operations in the

model that was developed for post-classification processing of the classified satellite image.

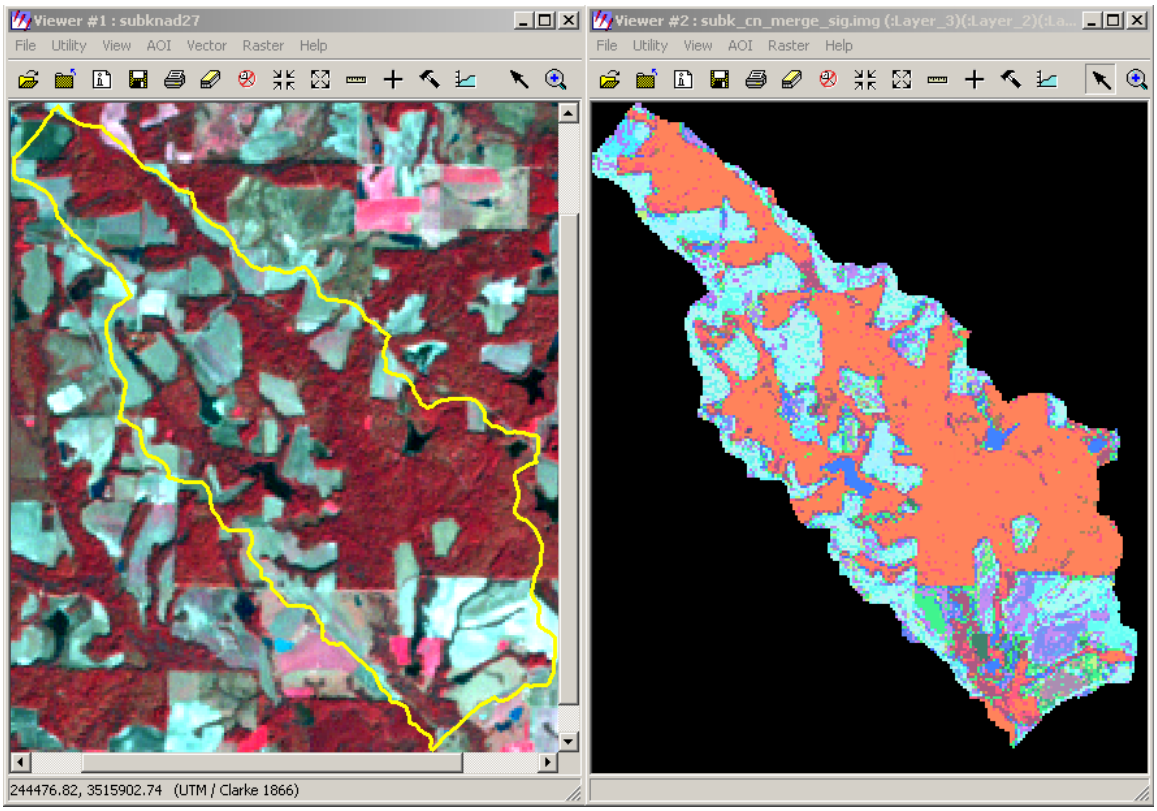


Figure 4.19. Subwatershed K depicted in the ETM+ image (left) and the fuzzy-classified image (right). Bands 2, 3, and 4 of the ETM+ image are displayed as blue, green, and red, respectively, while the same color order depict layers 1, 2, and 3 of the fuzzy-classified image classified for CN.

ImageInfo (subk_cn_merge_sig.img)

File Edit View Help

1 Layer_1

General | Projection | Histogram | Pixel data

Row	3	4	5	6	7	8
10	0	0	0	0	0	77
11	0	0	0	0	0	77
12	0	0	0	36	77	77
13	0	0	77	77	77	77
14	0	0	77	77	77	77
15	0	77	77	77	77	77
16	77	77	77	77	77	77
17	77	77	77	77	77	77
18	77	77	77	77	77	77
19	77	77	77	77	77	77
20	67	67	67	67	67	67

ImageInfo (subk_cn_merge_sig.img)

File Edit View Help

2 Layer_2

General | Projection | Histogram | Pixel data

Row	3	4	5	6	7	8
10	0	0	0	0	0	36
11	0	0	0	0	0	36
12	0	0	0	77	36	67
13	0	0	36	67	36	67
14	0	0	67	36	67	67
15	0	67	36	67	67	67
16	67	67	67	67	67	67
17	36	67	67	67	67	67
18	67	67	67	67	67	67
19	67	67	36	67	67	67
20	77	77	77	77	77	77

ImageInfo (subk_cn_merge_sig.img)

File Edit View Help

3 Layer_3

General | Projection | Histogram | Pixel data

Row	3	4	5	6	7	8
10	0	0	0	0	0	67
11	0	0	0	0	0	67
12	0	0	0	67	67	25
13	0	0	67	25	67	25
14	0	0	25	67	25	25
15	0	25	67	36	36	36
16	65	36	36	36	25	36
17	67	36	25	36	36	36
18	65	36	36	36	25	36
19	65	65	67	36	36	36
20	65	65	30	65	65	65

Figure 4.20. Pixel values for portions of the three layers of the ETM+ image classified for CN.

The Imagine fuzzy classifier provides the option of creating and writing a distance file along with the classified image file. In the case of a three-layer fuzzy classification, as was used in this study, the distance file contained three layers. Each value in the distance file contained the Mahalanobis distance (ERDAS 1997, 252) between the associated pixel in the classified image file and the mean vector for the class (value) to which the pixel was assigned in that particular layer of the fuzzy classified image. Segments of the three layers of the distance file corresponding to the segments of the classified image shown in Figure 4.20 are shown in Figure 4.21. Note that the Mahalanobis distance values for each pixel increase from layer 1 to layer 2 and from layer 2 to layer 3. This is because the maximum likelihood fuzzy classifier assigns pixels to classes by minimizing the Mahalanobis distance for layer 1, finding the second smallest value of the Mahalanobis distance for layer 2, etc.

A spatial model was created with Imagine to create weighted averages for the pixel values (e.g. CN value) associated with each of the three layers of the fuzzy classified image, with the weighting based on the Mahalanobis distance of each pixel value in each layer. The model used the three layers of the distance file, inverted and normalized (i.e. scaled so that the sum of the three layer values associated with each pixel was equal to one) to weight the values of the classified image, which represented parameter values for later use in the AGNPS model. Figure 4.22 is a graphical representation of the spatial model. The objects with uneven edges represent files used by the model. The circular objects in the graphical model represent operations performed on files. The directions of arrows indicate whether a file is an input or an output of a particular operation. Two of the files are strictly inputs to model operations. Five of the

ImageInfo (subk_cn_merge_sig_dist.img)

File Edit View Help

1 Layer_1

General | Projection | Histogram | Pixel data

Row	3	4	5	6	7	8
10	000.000	000.000	000.000	000.000	000.000	014.898
11	000.000	000.000	000.000	000.000	000.000	013.155
12	000.000	000.000	000.000	023.599	019.098	016.872
13	000.000	000.000	016.380	018.563	016.917	023.879
14	000.000	000.000	017.043	024.906	012.591	013.565
15	000.000	010.466	013.202	011.403	010.983	013.453
16	014.616	015.933	013.220	011.279	010.464	011.417
17	020.536	010.434	010.519	016.286	018.616	012.277
18	015.407	015.009	011.218	017.503	016.254	010.732
19	010.681	026.033	013.358	018.754	011.353	009.848
20	017.212	020.945	020.264	018.690	018.911	020.625

ImageInfo (subk_cn_merge_sig_dist.img)

File Edit View Help

2 Layer_2

General | Projection | Histogram | Pixel data

Row	3	4	5	6	7	8
10	000.000	000.000	000.000	000.000	000.000	029.058
11	000.000	000.000	000.000	000.000	000.000	030.189
12	000.000	000.000	000.000	023.979	027.593	028.225
13	000.000	000.000	018.906	028.747	021.248	035.416
14	000.000	000.000	036.190	031.676	031.224	031.350
15	000.000	033.546	029.615	030.646	030.735	035.134
16	027.866	035.710	033.995	031.853	029.304	033.002
17	029.039	030.400	032.570	034.354	035.062	031.983
18	024.666	035.492	033.284	037.176	033.282	028.477
19	020.799	032.898	025.593	031.238	026.884	027.094
20	020.165	028.368	025.701	032.041	036.776	021.110

ImageInfo (subk_cn_merge_sig_dist.img)

File Edit View Help

3 Layer_3

General | Projection | Histogram | Pixel data

Row	3	4	5	6	7	8
10	000.000	000.000	000.000	000.000	000.000	030.781
11	000.000	000.000	000.000	000.000	000.000	031.744
12	000.000	000.000	000.000	025.555	028.706	059.482
13	000.000	000.000	025.511	063.121	027.442	093.129
14	000.000	000.000	087.230	032.996	064.531	070.335
15	000.000	070.249	031.726	046.795	037.964	039.835
16	043.962	037.976	036.331	040.250	074.178	036.795
17	029.261	034.086	070.188	041.815	048.643	033.405
18	030.436	059.217	038.400	048.440	082.940	033.101
19	031.399	049.428	025.937	036.259	036.769	032.794
20	020.856	040.053	027.203	039.829	042.428	028.344

Figure 4.21. Values for portions of the three layers of the distance file produced for the ETM+ image classified for CN.

files are outputs of one operation and inputs to one or more operations. Only the final file in the model is solely an output file. The functions used in the graphical model are listed in Table 4.12.

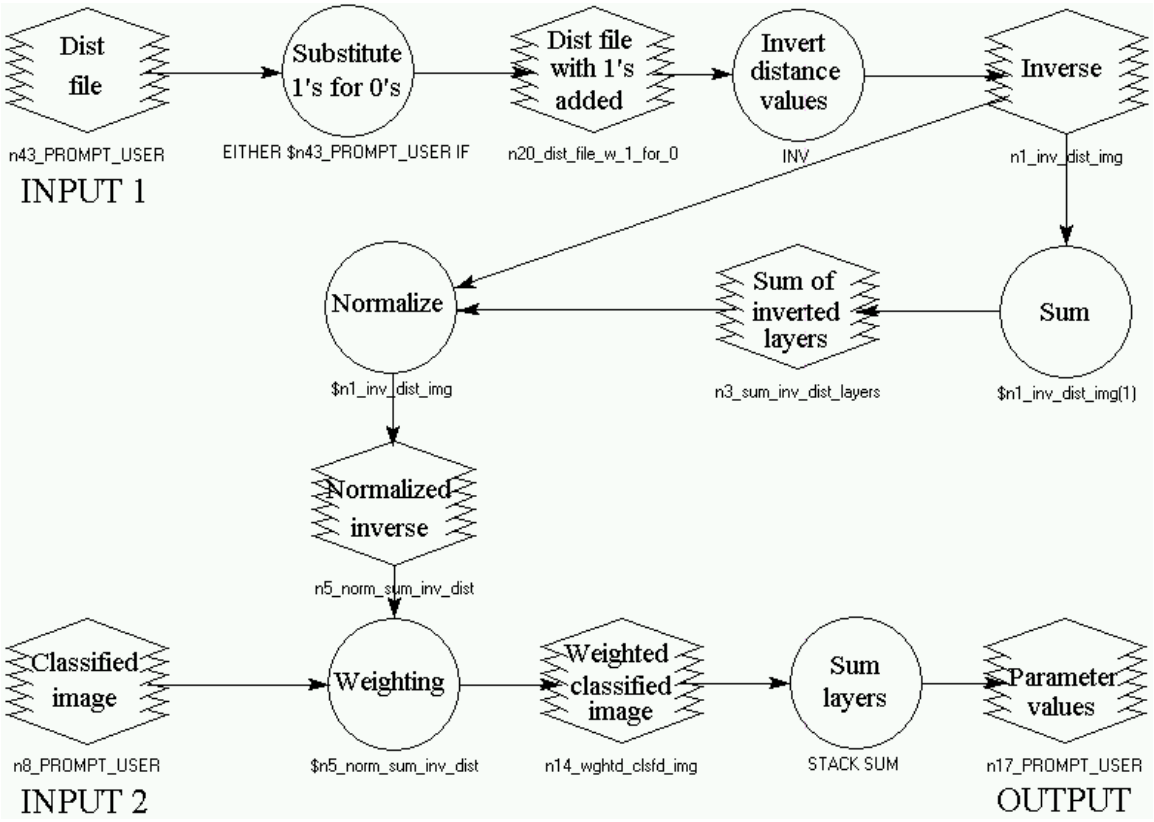


Figure 4.22. Spatial model for generating parameter values using fuzzy classified image and its distance file.

Figures 4.23 through 4.28 depict the sequential steps performed by the spatial model, the input files used in each step, the output file of each step, and representative pixel values from layer 1 of each output image for the same segment of the watershed. All images used or created by the model had three layers except for the sum of the inverse distance values image and the parameter values image, each of which contained

only one layer. Where the output of one operation is shown in a figure, it is not repeated in the subsequent figure where it serves as an input to an operation.

Table 4.12. Functions used in the post-classification spatial model

Model function	Function definition
Substitute 1's for 0's	EITHER $n43_PROMPT_USER IF ($n43_PROMPT_USER > 0) OR 1 OTHERWISE
Invert distance values	INV ($n20_dist_file_w_1_for_0)
Sum	$n1_inv_dist_img(1) + $n1_inv_dist_img(2) + $n1_inv_dist_img(3)
Normalize	$n1_inv_dist_img / $n3_sum_inv_dist_layers
Weighting	$n5_norm_sum_inv_dist * $n55_clsfd_img_dvdd_by_scalar
Sum layers	STACK SUM ($n14_wghtd_clsfd_img)

Figure 4.23 depicts the operation of substituting values of one for values of zero in the distance file. This operation was necessary to avoid division by zero when the distance file was inverted. The substitution of ones for zeros did not impact the final model output because all zero distance values were outside the watershed boundary. Although both input and output images for this step appear identical, the small sample of pixel values clearly shows that zeros have been replaced by ones.

Figure 4.24 depicts the operation of inverting the distance file after ones had been substituted for zeros. The image appears black in the display because a linear bin function was used and 28,852 of the pixel values (60%) were equal to one, while the other 18,986 pixel values (40%) were less than 0.190876. This highly bipolar histogram led to the display of the file as nearly black and white.

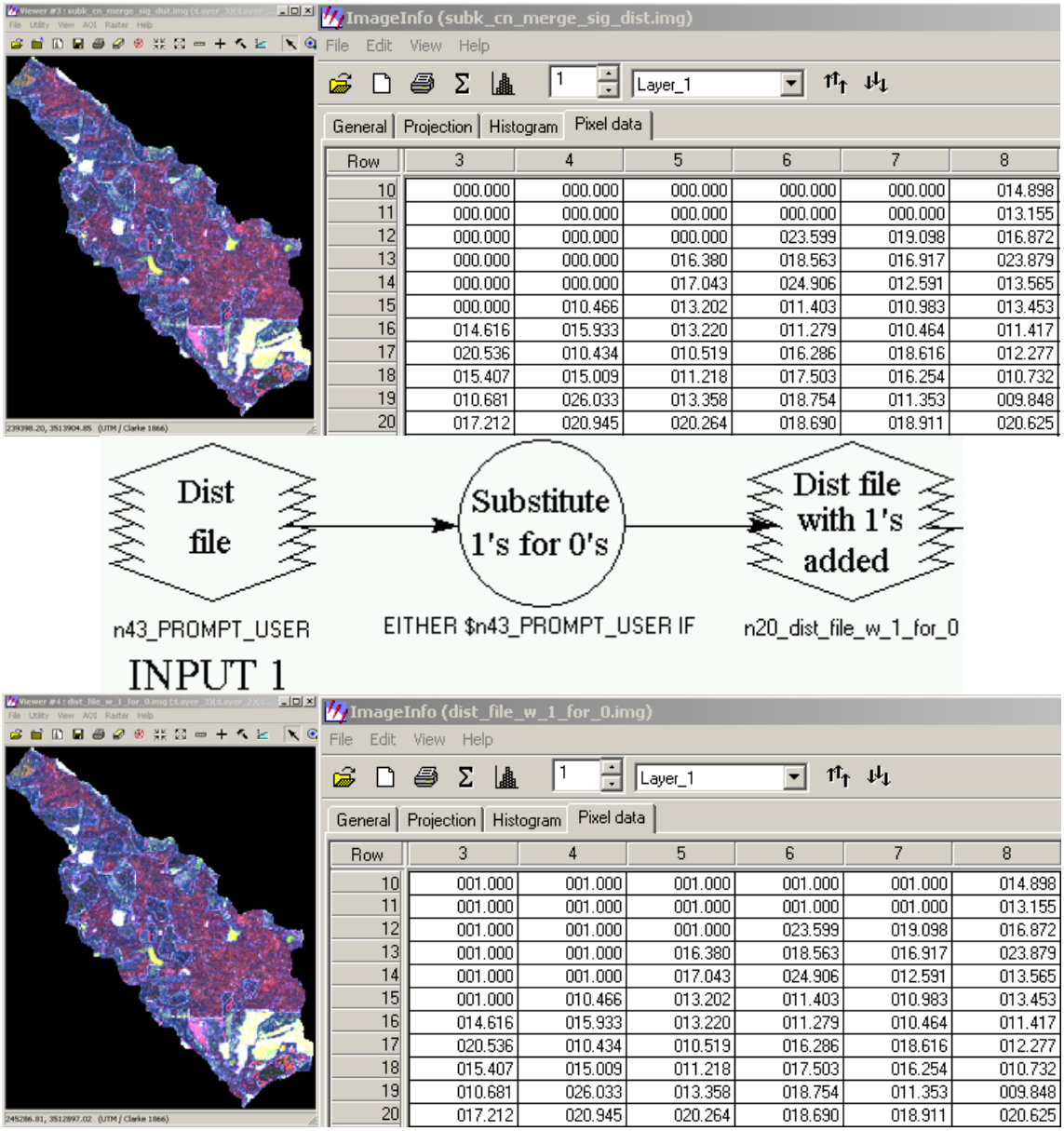
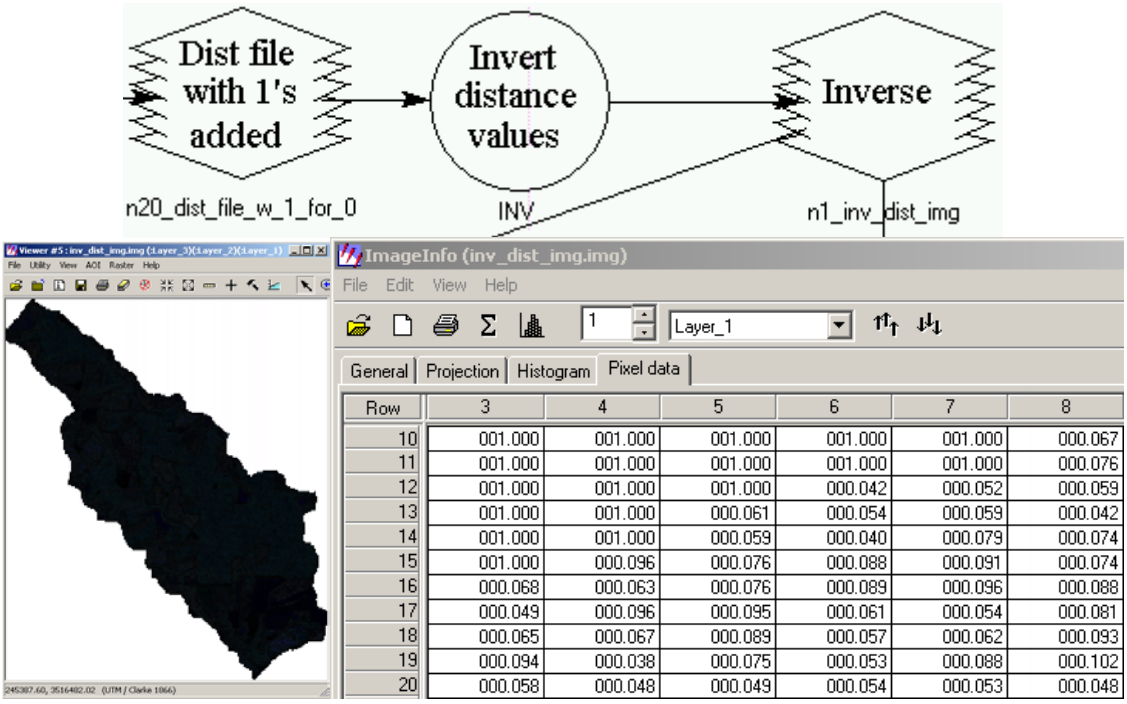


Figure 4.23. Substituting 1's for 0's so that the matrix can be inverted.

Figure 4.25 depicts the operation of summing all three layers of the inverted distance file image for subsequent use in normalizing the inverted distance file. This single-layer image appears to be black and white for the same reason as the inverted distance file does. The pixels outside the watershed boundary all contain values of 3.00

because the three layers of the inverted distance file all contained values of 1.00 in these same locations.

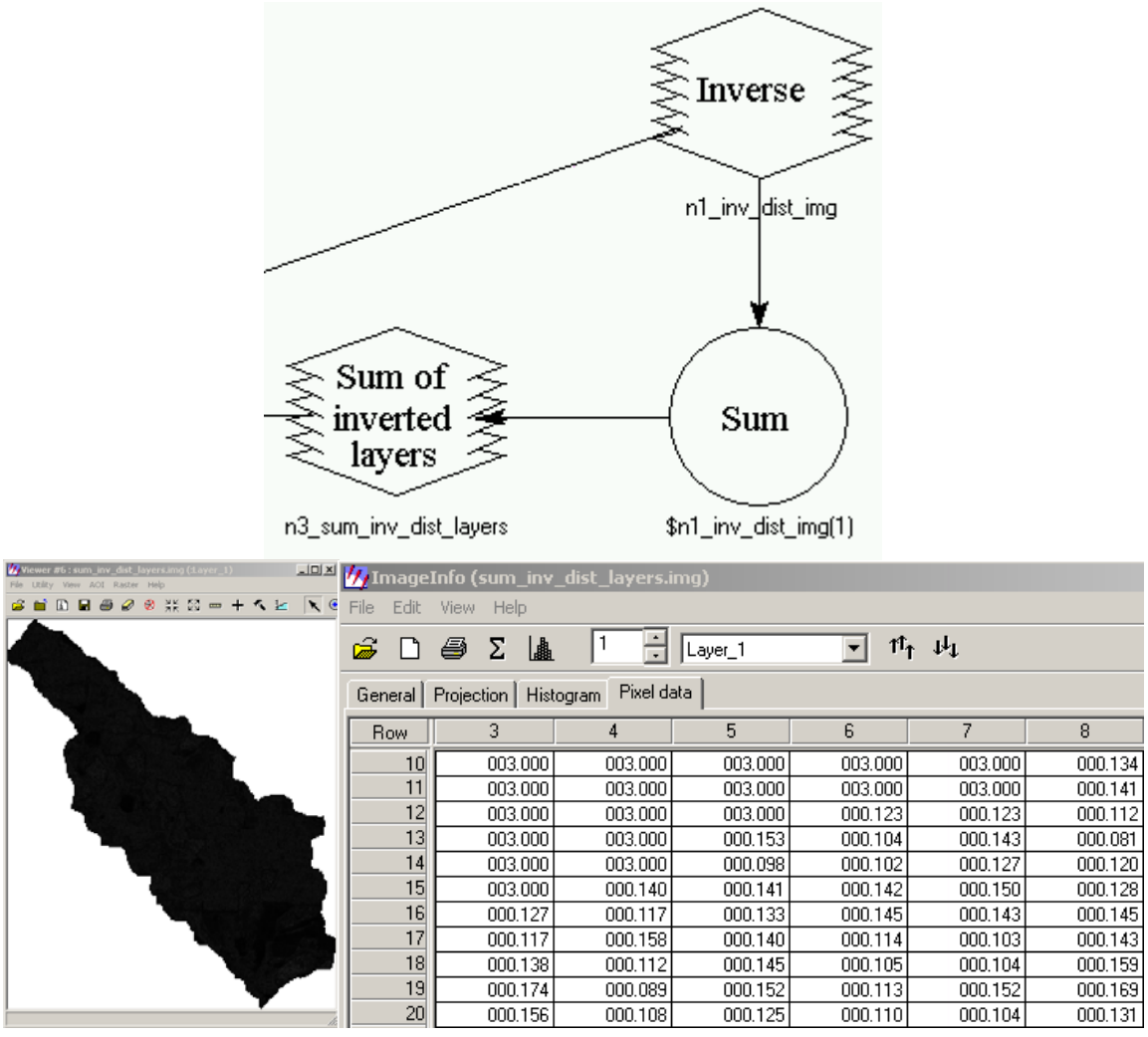


Row	3	4	5	6	7	8
10	001.000	001.000	001.000	001.000	001.000	000.067
11	001.000	001.000	001.000	001.000	001.000	000.076
12	001.000	001.000	001.000	000.042	000.052	000.059
13	001.000	001.000	000.061	000.054	000.059	000.042
14	001.000	001.000	000.059	000.040	000.079	000.074
15	001.000	000.096	000.076	000.088	000.091	000.074
16	000.068	000.063	000.076	000.089	000.096	000.088
17	000.049	000.096	000.095	000.061	000.054	000.081
18	000.065	000.067	000.089	000.057	000.062	000.093
19	000.094	000.038	000.075	000.053	000.088	000.102
20	000.058	000.048	000.049	000.054	000.053	000.048

Figure 4.24. Inverting the distance file after 1's have been added.

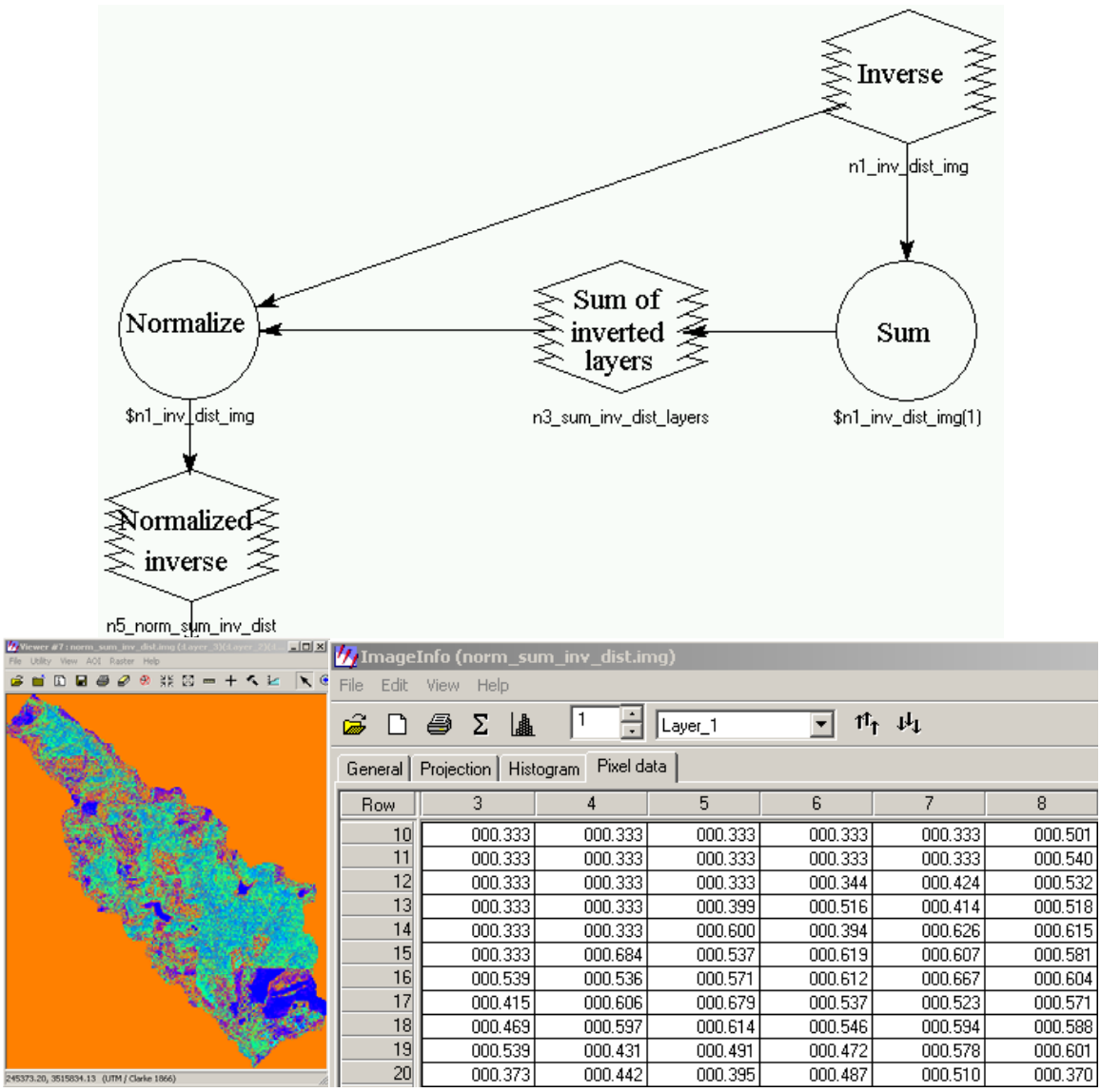
Figure 4.26 depicts the operation of normalizing the values in the three layers of the inverted distance file so that for every pixel, the sum of the values in the three layers would equal one. This was accomplished by dividing each value in each layer of the inverted distance file by the corresponding pixel value in the sum of inverted layers file. While only layer one of the three layer output image for this operation is shown, it is clear from examining the values outside the watershed boundary, i.e. those in the upper left corner of the pixel values image, what happened in this operation. Representative samples of values inside the watershed boundary were also checked to ensure that the

operation had the intended effect. The output image, normalized by the sum of the inverted distance values, was treated as containing pseudo-probability values, each of which was considered a surrogate for the probability that the classified image file value in the same location and layer was the correct value for the parameter for which the ETM+ image had been classified.



Row	3	4	5	6	7	8
10	003.000	003.000	003.000	003.000	003.000	000.134
11	003.000	003.000	003.000	003.000	003.000	000.141
12	003.000	003.000	003.000	000.123	000.123	000.112
13	003.000	003.000	000.153	000.104	000.143	000.081
14	003.000	003.000	000.098	000.102	000.127	000.120
15	003.000	000.140	000.141	000.142	000.150	000.128
16	000.127	000.117	000.133	000.145	000.143	000.145
17	000.117	000.158	000.140	000.114	000.103	000.143
18	000.138	000.112	000.145	000.105	000.104	000.159
19	000.174	000.089	000.152	000.113	000.152	000.169
20	000.156	000.108	000.125	000.110	000.104	000.131

Figure 4.25. Summing the three layers of the inverted distance file.

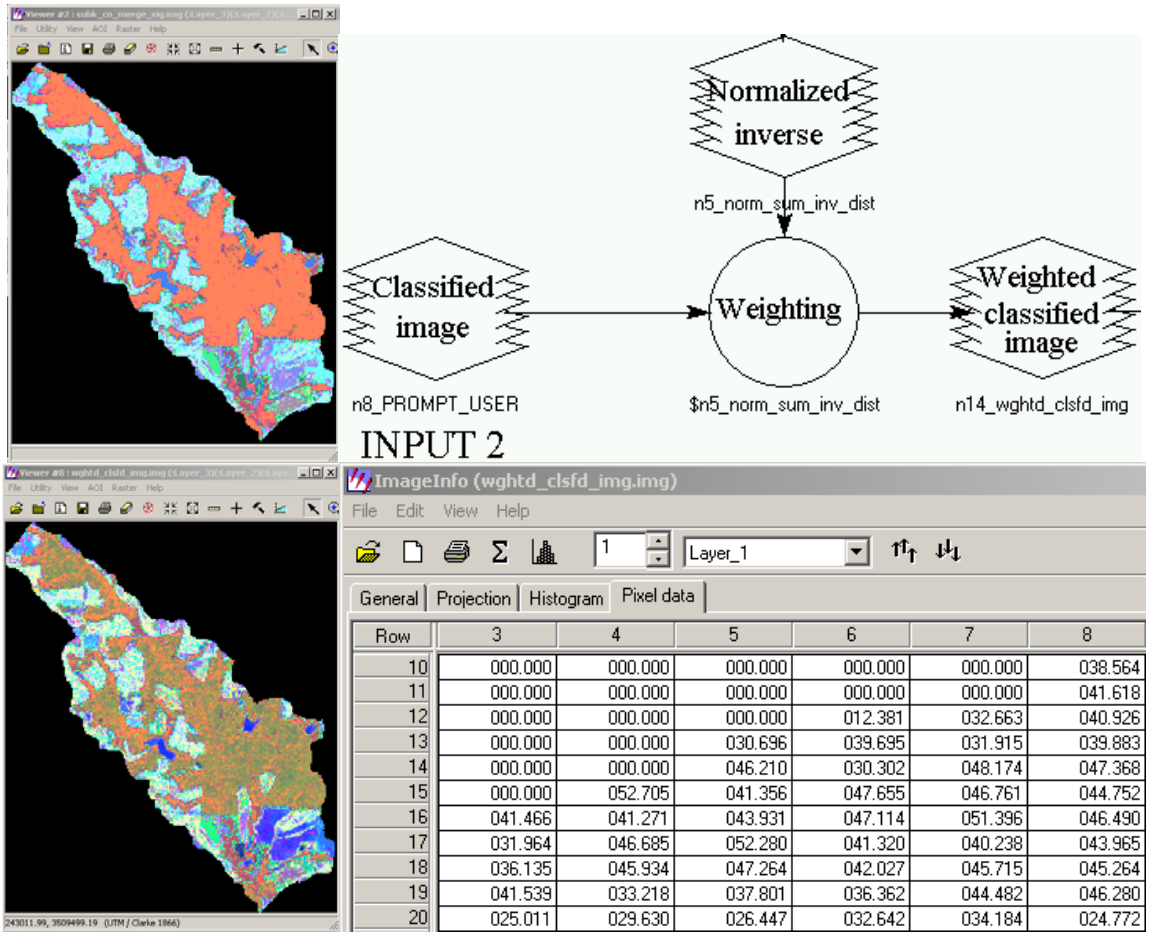


Row	3	4	5	6	7	8
10	000.333	000.333	000.333	000.333	000.333	000.501
11	000.333	000.333	000.333	000.333	000.333	000.540
12	000.333	000.333	000.333	000.344	000.424	000.532
13	000.333	000.333	000.399	000.516	000.414	000.518
14	000.333	000.333	000.600	000.394	000.626	000.615
15	000.333	000.684	000.537	000.619	000.607	000.581
16	000.539	000.536	000.571	000.612	000.667	000.604
17	000.415	000.606	000.679	000.537	000.523	000.571
18	000.469	000.597	000.614	000.546	000.594	000.588
19	000.539	000.431	000.491	000.472	000.578	000.601
20	000.373	000.442	000.395	000.487	000.510	000.370

Figure 4.26. Normalizing the inverted distance file by dividing by the sum of inverted distance values.

The operation of weighting the pixel values in the classified image with the pseudo-probability values calculated in the previous step is depicted in Figure 4.27. This was a simple multiplication operation, which restored the zero values outside the watershed boundary because the classified image had zero values there. The Viewer in the upper left corner of Figure 4.27 depicts the classified ETM+ image, in this case

classified for CN values. The output image of the operation, the weighted, classified CN image, is shown at the lower left of the figure.

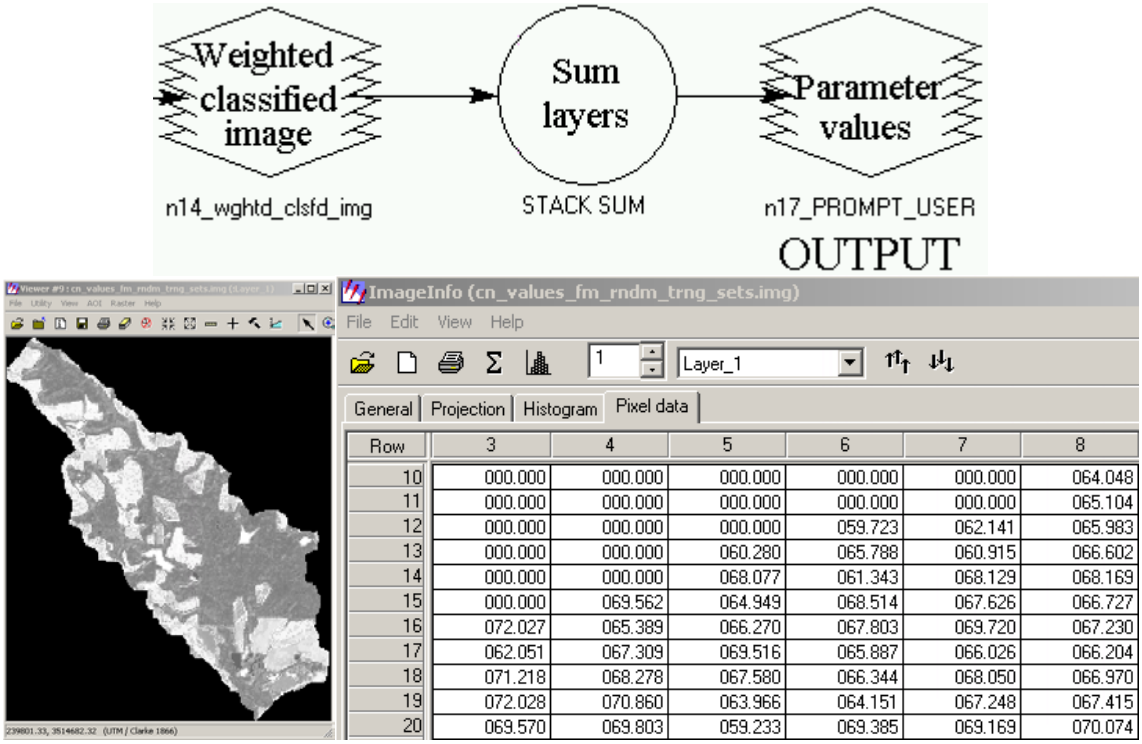


Row	3	4	5	6	7	8
10	000.000	000.000	000.000	000.000	000.000	038.564
11	000.000	000.000	000.000	000.000	000.000	041.618
12	000.000	000.000	000.000	012.381	032.663	040.926
13	000.000	000.000	030.696	039.695	031.915	039.883
14	000.000	000.000	046.210	030.302	048.174	047.368
15	000.000	052.705	041.356	047.655	046.761	044.752
16	041.466	041.271	043.931	047.114	051.396	046.490
17	031.964	046.685	052.280	041.320	040.238	043.965
18	036.135	045.934	047.264	042.027	045.715	045.264
19	041.539	033.218	037.801	036.362	044.482	046.280
20	025.011	029.630	026.447	032.642	034.184	024.772

Figure 4.27. Weighting the classified image with normalized inverted distance values. The input classified image is above and the output weighted classified image is below.

The final step of the model is depicted in Figure 4.28. This step simply adds the values of the three layers for each pixel to create a single layer output file. Each pixel value in the output file is the sum of the three possible values of the parameter being modeled that were determined to be the first, second, and third most probable values for

the parameter for that pixel, each of which has been multiplied by the weighting factor created from the pseudo-probability of the associated value being the correct value.



Row	3	4	5	6	7	8
10	000.000	000.000	000.000	000.000	000.000	064.048
11	000.000	000.000	000.000	000.000	000.000	065.104
12	000.000	000.000	000.000	059.723	062.141	065.983
13	000.000	000.000	060.280	065.788	060.915	066.602
14	000.000	000.000	068.077	061.343	068.129	068.169
15	000.000	069.562	064.949	068.514	067.626	066.727
16	072.027	065.389	066.270	067.803	069.720	067.230
17	062.051	067.309	069.516	065.887	066.026	066.204
18	071.218	068.278	067.580	066.344	068.050	066.970
19	072.028	070.860	063.966	064.151	067.248	067.415
20	069.570	069.803	059.233	069.385	069.169	070.074

Figure 4.28. Summing the weighted classified image values to produce CN values.

A slightly modified version of this process was used to generate signatures and classify the ETM+ image to produce matrices of n, P-fact, and SCC values for Subwatershed K. The difference was due to the fact that the values for n, P-fact, and SCC were all decimals, estimated to two decimal places. In contrast, the CN values were all integers. The signature editor for ERDAS Imagine requires that values associated with the signatures be integer values. All values for n, P-fact, and SCC, therefore, were multiplied by 100 to convert them to integers. These values were used in the signature

files and were the pixel values in the three layers of the classified images input into the post-classification model. In order to recover the correct values of these parameters, the input image files were declared as floating point files, then the pixel values in all three layers were divided by 100. Figure 4.29 illustrates the additional step added to the model to process the integer values and return them to decimal values. This step was accomplished by adding the function, "$n53_PROMPT_USER / 100" to the model, where "$n53_PROMPT_USER" represents the classified image that was created from signature values that had been multiplied by 100. The process of calculating P-fact values was nearly trivial but it was done anyway to ensure that the process was valid even for the case of the nearly constant values for P-fact across the watershed.

The entire process of classifying and modeling the Landsat ETM+ image for AGNPS parameter values CN, n, P-factor, and SCC described in this chapter was repeated for West Fork Little River and for East Fork Little River, with only slight modifications. The first difference between the two was that the ground truth data collected in the West Fork Little River watershed were used for East Fork Little River as well. No ground truth data were gathered in the East Fork Little River watershed except rainfall data, stream flow, and water quality data. Another difference was that values for n were estimated to three decimal places in the West Fork Little River rather than two, as was done in Subwatershed K. This required a slight modification to the post-classification model. Instead of multiplying the ground truth values by 100 prior to using them in the Signature Editor, they were multiplied by 1000. This required that the classified image pixel values then be multiplied by 1000 rather than 100. The final difference in the process was that the ETM+ image was not classified for P-factor nor

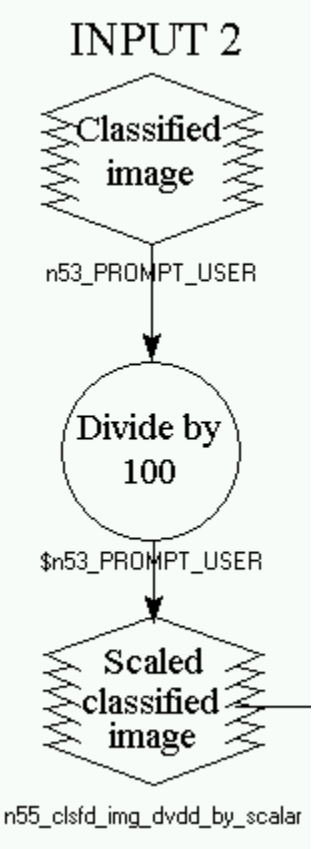


Figure 4.29. Model function added to correct classified image pixel values after their associated signature values were multiplied by 100.

was the post-classification model run for P-factor in West Fork Little River. This is because all ground truth estimates of P-factor in West Fork Little River were 1.00. Thus, classifying the satellite image and running the model for this parameter would have been superfluous in West Fork Little River and East Fork Little River.

The Landsat ETM+ images were also classified within the three watersheds, Subwatershed K, West Fork Little River, and East Fork Little River, in terms of land use and land cover (LU/LC). The results of these classifications were used, in conjunction with tabular data linking CN, n, P-factor, and SCC, to generate values for those

parameters for use in separate runs of the AGNPS model, for comparison with the runs using the parameters obtained by classifying and post-processing the ETM+ images directly for those constituent values.

Accuracy assessment

Accuracy assessment of the results of the image classification process fell into two distinct categories: 1) accuracy assessment of the classification of Landsat ETM+ imagery into LU/LC classes and 2) accuracy assessment of the fuzzy classification and post-classification modeling of Landsat ETM+ imagery into numerical values of AGNPS parameters, CN, n, P-factor, and SCC.

The number of pixels used in the accuracy assessment process significantly impacts the statistical validity of the process. As has been discussed earlier in the chapter, only 2094 pixels were selected for ground truth in Subwatershed K. This placed a significant limitation on the number that could be reserved for accuracy assessment and still leave enough for training sets for the supervised classification process. The lessons learned from the limitations this imposed on the analysis of Subwatershed K were applied to the analysis of the two Piedmont watersheds. Significantly more ground truth information was obtained in West Fork Little River (4567 pixels), which allowed the reservation of far more data (1220 pixels) for accuracy assessment use.

Accuracy assessment for classifications of the ETM+ imagery into LU/LC classes was more straightforward than the process involving AGNPS parameter values. Imagine's Accuracy Assessment module was used, though with modification to the most commonly used method. This module normally generates a random sampling of pixels in a classified image and the user then supplies the ground truth values for the selected

pixels. Accuracy assessment statistics are then generated based on these data. In this study, however, ground truth points were selected prior to analysis and from the ground truth points thus selected, clusters of pixels were randomly selected for accuracy assessment. Because of this, accuracy assessment points had to be converted from pixels to ASCII text files. Although an Imagine utility exported pixel locations and file values, only the file values were retained by Imagine when these files were imported into the Accuracy Assessment module. Therefore, it was necessary to link the Accuracy Assessment module to the classified LU/LC image being assessed to extract pixel values. This process was found to have a feature which was deemed undesirable for this study: the pixel value extracted for each accuracy assessment pixel was not necessarily the value for that pixel; rather, it was the majority value for a three by three window of pixels centered on the subject pixel. It was discovered, however, that an option found in the Accuracy Assessment module allowed the selection of the center pixel value from the three by three window in the event that a majority threshold was not obtained.

With either method, using randomly selected pixels or user-supplied pixels, it was necessary to manually add the ground truth pixel values to the accuracy assessment table for comparison to the classified image pixel values at the same locations. Once this was done, it was a simple matter to have Imagine generate an accuracy assessment report giving user's accuracy, producer's accuracy, overall accuracy, and Kappa statistics for the classified image. These results of the accuracy assessment effort are presented in Chapter 5 for Subwatershed K and West Fork Little River.

To assess the accuracy of the process of classifying Landsat ETM+ imagery into three fuzzy classes of each of the targeted AGNPS parameter values and calculating

weighted averages of those fuzzy-classified images, two models were developed to calculate statistics that were then used to judge the accuracy of the classification and modeling process. The first model calculated the absolute value of residuals of the process, where the residuals were defined as the differences between the ground truth values of the AGNPS parameters and the modeled values. The residuals were calculated for each cell within the regions randomly selected from the ground truth data for accuracy assessment use. The second model was identical to the first except that it included a step to divide pixel values by 1000. This was necessary because the Imagine Recode function would only assign integer values to the pixels, even though the data type was declared as floating point. This was essentially the same problem that was encountered with the Imagine Signature Editor, so the work-around procedure had already been devised and tested. Figure 4.30 is a graphical representation of the model that includes the step of dividing pixel values in the image containing accuracy assessment parameter values by 1000 to reverse the multiplication of the original values by 1000, which was required for parameters n, P-factor, and SCC. Functions used in the model are listed in Table 4.13. The other model is identical to this one, except for the division by 1000 step.

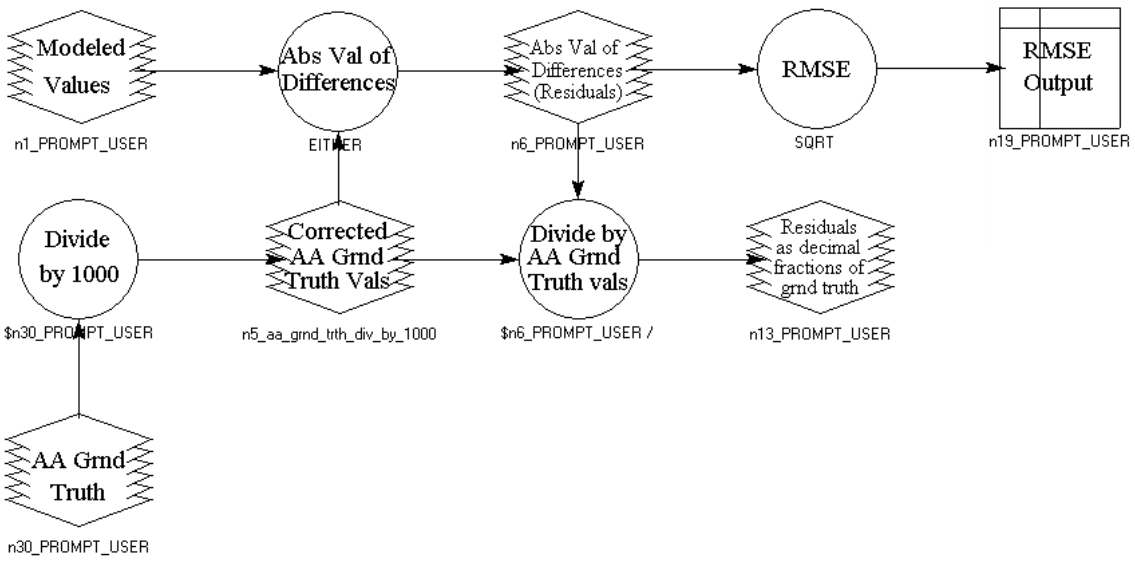


Figure 4.30. Imagine accuracy assessment model used for imagery with floating point pixel values representing AGNPS parameter values that were previously multiplied by 1000.

Table 4.13. Functions used in the accuracy assessment of AGNPS parameter values model.

Model function	Function definition
Divide by 1000	$n30_PROMPT_USER / 1000 (Note: Not used with CN)
Abs Val of Differences	EITHER (ABS ($n1_PROMPT_USER-$n5_aa_grnd_trth_div_by_1000)) IF (($n5_aa_grnd_trth_div_by_1000>0) AND ($n1_PROMPT_USER>0)) OR 0 OTHERWISE
Divide by AA Grnd Trth vals	$n6_PROMPT_USER/(EITHER 1 IF (($n5_aa_grnd_trth_div_by_1000 EQ 0) OR ($n5_aa_grnd_trth_div_by_1000 OTHERWISE
RMSE	SQRT (GLOBAL MEAN (($n6_PROMPT_USER*$n6_PROMPT_USER), IGNORE 0))

The following steps were used in the accuracy assessment of the process of classifying and modeling AGNPS parameter values: 1) A blank image was created for

each of the two watersheds for which accuracy assessment was conducted, Subwatershed K and West Fork Little River, 2) spreadsheets were created with Microsoft Excel containing all ground truth AOIs for Subwatershed K and West Fork Little River and the values of all AGNPS parameter values estimated for each AOI (Tables 4.12 and 4.13), 3) the spreadsheets were successively sorted for each AGNPS parameter value and for the "AA" indicator of AOIs designated for accuracy assessment use, 4) AOIs were created for all accuracy assessment pixels sharing the same value of each parameter in each watershed by opening all accuracy assessment AOIs sharing the same parameter value, then saving the resultant multi-element AOI under a new name (e.g., cn67aa.aoi), 5) the Recode function was used in Imagine to assign either the parameter value (CN) or 1000 times the parameter value (n, P-factor, and SCC) to all accuracy assessment pixels sharing the same parameter value, as delineated by the AOI described in step 4, 6) the Imagine Stack function was used *n-1* times, where *n* represents the number of discrete values for a given AGNPS parameter in each watershed, to create a composite image for each parameter in each watershed containing all pixel values intended for accuracy assessment use, 7) the appropriate one of the two accuracy assessment models was used to calculate the absolute values of the residuals, i.e. the differences between the estimated accuracy assessment values and the values calculated by fuzzy classification and post-classification modeling, and the root mean square error (RMSE) for each AGNPS parameter in each watershed.

In the early stages of developing and testing the accuracy assessment models, it was discovered that, where accuracy assessment ground truth data fell outside the boundaries of the test watersheds, the accuracy assessment model calculated a value of 1

for the residuals. Because of this, a conditional test was added to the model to eliminate those accuracy assessment values for which there was no corresponding modeled parameter value, i.e. all pixels outside the watershed boundaries. The results of accuracy assessment are presented in the following chapter.

GIS-AGNPS linkages

The U.S. Geological Survey has been working on the development of a set of models working inside Imagine designed to estimate AGNPS parameters on the basis of information associated with soil mapping unit, land cover, and elevation. As of this writing, the work is essentially complete (Finn et al. 2002) but has undergone only limited testing. This project served as an additional test for the models, known collectively as AGNPS Data Generator.

Data Generator was used as the foundation for generating 22 cell-based AGNPS parameters. Fifteen of these, parameters 1, 3, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, and 22, were generated on the basis of elevation from the DEMs; SCS curve number (CN) based on hydrologic soil group, soil erodibility, and soil texture associated with soil mapping units in the soils images; and Manning's n (n), cropping factor, surface condition constant (SCC), COD factor, a fertilizer indicator, a pesticide indicator, and the hydrologic soil group associated with land cover classes. Many of these factors had previously been associated with land cover classes by the Southeast Watershed Research Lab in Tifton Georgia for use in the Little River watershed and were provide by Dr. David Bosch for use in this study. Data Generator set the other seven cell-based AGNPS parameters internally.

For each of the three watersheds, a digital elevation model (DEM) and a soils image were obtained as described in prior sections of this document. The Turner County DEM and soils image were prepared as described in the "Soils Data" section of this chapter, adding appropriate attributes, and successfully used with Data Generator. The Subwatershed K land cover file generated by classifying the Landsat ETM+ image was also used for generating AGNPS parameters with Data Generator. The results are presented in the following chapter.

In order to run Data Generator on the West Fork Little River and East Fork Little River watersheds, it was necessary to create mosaics of the Hall, White, and Lumpkin County DEMs and soils images. The mosaic of DEMs was created without great difficulty, though several empty (zero-value) pixels resulted along the boundaries of the three files. These pixels were assigned values using Imagine. The elevation values were assigned by performing bilinear interpolation based on the four surrounding pixel values. The DEM mosaic file was subset to minimize storage requirements and processing time required by the file. For reasons that were never determined, the original subset image and subsequent, slightly different, subset DEM images failed to work with Data Generator. Ultimately, the full, 14-megabyte (14 MB) DEM mosaic was successfully used with Data Generator.

The soils images for Hall and White Counties and the small soils image created from digitized soil mapping unit polygons for Lumpkin county were successfully mosaiced and subset for use with Data Generator. As with the DEM mosaic, some empty pixels resulted along the boundaries of the subset files. These values were filled with values associated with the appropriate soil mapping units, as determined from the paper

soils survey maps. Figure 4.31 is a screen capture of the subset soils image mosaic, with a small selection of the attributes assigned to the soil mapping units represented in the image.

The land cover images used in West Fork Little River and East Fork Little River watersheds were the images created by classifying the Landsat ETM+ image into land use/land cover classes, as previously described. Attributes required by Data Generator were assigned based on those used in Subwatershed K or, where the classes were significantly different from those used in Subwatershed K, based on published values in AGNPS references (Young et al. 1995, USDA 1997). The expanded attributes of the land cover image for West Fork Little River and East Fork Little River watersheds are shown in Figure 4.32.

AGNPS model

Version 5.0 of the AGNPS model has been incorporated into the Data Generator models developed by USGS (Finn et al. 2002). Once image files had been created by Data Generator containing values for the cell-based AGNPS parameters, a single image was created with Data Generator containing all the generated values as image layers. A separate utility within Data Generator was then used to extract cell-based parameter values from the image and format them into a text tile in ".dat" format. The AGNPS program was called from within Data Generator with the text file as input.

Output from the AGNPS model is produced in another text file with a ".nps" file extension. Data Generator has routines built into it that convert this text file into raster images that can be displayed by Imagine. The individual pixel values within these images can be examined with Imagine's Image Information utility.

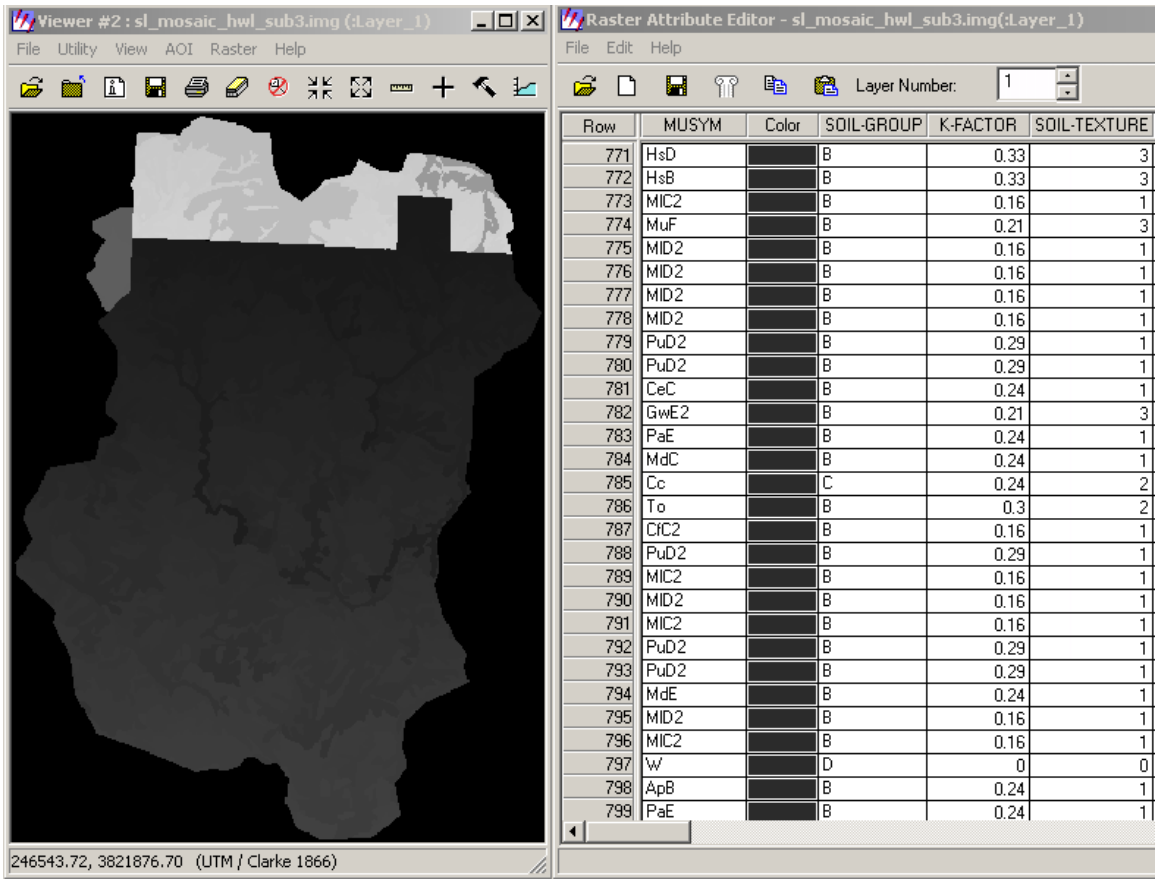


Figure 4.31. Soils image mosaic of Hall, White, and Lumpkin Counties, subset to a region immediately surrounding West Fork Little River and East Fork Little River watersheds, with representative soils attributes.

Data Generator, as used in this study, did not include any provision for changing basin parameters, i.e. parameters that apply to the entire watershed. The parameters that were of particular interest in this study were total precipitation for a storm event, rainfall duration, storm type, a geomorphic indicator that determines whether basin channels will be modeled as triangular or rectangular, and the K-coefficient, a measure of the distribution of rainfall intensity within a storm event. The locations of these parameters in the AGNPS 5.0 input file are described by USDA (1998). Since there were no rainfall

Raster Attribute Editor - lulc_wflr_fm_merged_sig.img(:Layer_1)

File Edit Help

Layer Number: 1

Row	Class Names	Color	MANNING	C-FACTOR	SCC	COD	FERT	PEST	A	B	C	D
0	Unclassified		0	0	0	0	0	0	0	0	0	0
1	Cover crop		0.13	0.1	0.29	80	2	1	63	74	82	85
2	Deciduous forest		0.3	0.02	0.59	65	0	0	25	55	70	77
3	Evergreen forest		0.3	0.02	0.59	65	0	0	25	55	70	77
4	Farmsteads		0.03	0.2	0.01	80	1	1	59	74	82	86
5	Mixed forest		0.3	0.02	0.59	65	0	0	25	55	70	77
6	Pasture		0.06	0.04	0.15	60	1	1	49	69	79	84
7	Residential, forested		0	0.02	0.4	70	1	1	42	65	76	82
8	Residential, open		0.08	0.03	0.1	80	3	1	59	74	82	86
9	Water		0.99	0	0	0	0	0	100	100	100	100
10	Urban/Institutional		0.03	0.2	0.01	80	0	0	98	98	98	98

Figure 4.32. Expanded attribute table for the land use/land cover images of West Fork Little River and East Fork Little River watersheds.

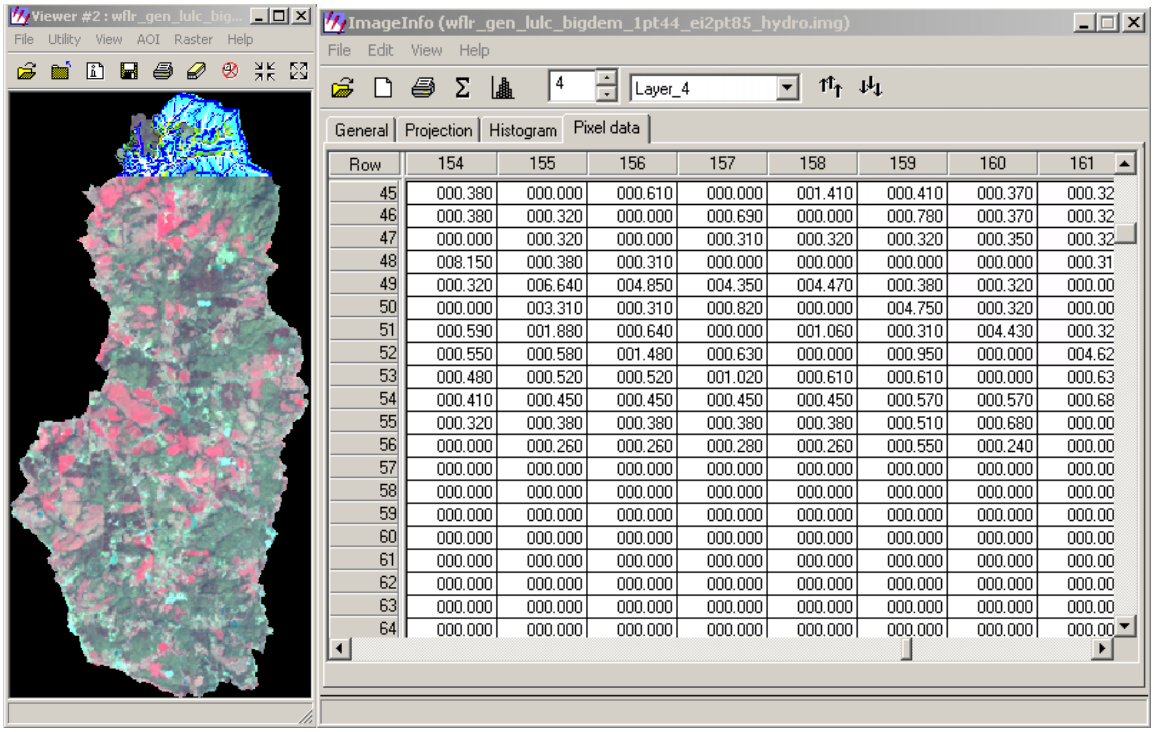
or runoff data for Subwatershed K to compare with modeled values, this issue was not considered in implementing AGNPS to model that watershed. With the collection of rainfall and runoff quantity and quality data for West Fork Little River and East Fork Little river watersheds, it was important to be able to adjust the basin parameters to try to simulate the conditions monitored in the watersheds. This turned out to be not possible, however.

The storm that was monitored in West Fork Little River and East Fork Little river watersheds produced an average of 1.44 inches of rainfall over the watersheds. After AGNPS parameters had been generated for West Fork Little River watershed, the AGNPS input file was opened in a text editor and selected basin parameters were changed from their default values. The geomorphic indicator was changed to 1 to indicate that the model should adjust the channel geometry to reflect the basin characteristics. The K-coefficient was changed to 140 from its default value of 366 to

better represent a less intense storm (Merritt 1983, p. 21-90). The storm type was changed from 3 to II, the most common type in North Georgia (Georgia Soil and Water Conservation Commission 1992, p. A-1-104). The energy intensity value (EI index) was changed from the default value of 160 to 0 because storm duration was to be specified and this would trigger the internal calculation of EI index by AGNPS. Rainfall duration was changed from the default of 0 to 12.1 to reflect the effective duration of the monitored storm and precipitation was changed from the default of 7.3 to 1.44. These data were used, along with cell data developed with Data Generator, as input to AGNPS and the model was run. The results were surprising and remain unexplained. As Figure 4.33 illustrates, the model calculated plausible values approximately 13% of the way through the watershed raster grid, then inserted what appeared to be background values of 000.000 in the case of flow in all the remaining cells of the grid within the watershed boundary. The image shown in Figure 4.33 is for the AGNPS hydro output image overlaid on the Landsat ETM+ image of West Fork Little River watershed. The pixel values shown are the concentrated flow values, in cubic feet per second (cfs) entering the cells. All other images generated for this simulation, such as for clay, silt, and sand yields, showed the same pattern of having only calculated results part way through the watershed.

Experiments with adjusting the storm duration, K-coefficient, EI index, and storm type produced the same results. Adjusting the precipitation amount, however, produced change. Increasing the precipitation amount was the only method found that caused AGNPS 5.0 to generate plausible results all the way through the watershed. Increasing precipitation invariably resulted in the propagation of results farther down through the

watershed. A precipitation amount of 3.0 inches produced a seemingly complete simulation. Because of this, the one-year 24-hour storm for the area, which is 3.2 inches (Georgia Soil and Water Conservation Commission 1992, p. A-2-10), was used for AGNPS simulations in West Fork Little River and East Fork Little River Watersheds. This was the closest rainfall amount to that which was monitored that had some statistical significance and that appeared to produce a complete AGNPS simulation. To date, the reason for this problem in AGNPS 5.0 is not known.



Row	154	155	156	157	158	159	160	161
45	000.380	000.000	000.610	000.000	001.410	000.410	000.370	000.32
46	000.380	000.320	000.000	000.690	000.000	000.780	000.370	000.32
47	000.000	000.320	000.000	000.310	000.320	000.320	000.350	000.32
48	008.150	000.380	000.310	000.000	000.000	000.000	000.000	000.31
49	000.320	006.640	004.850	004.350	004.470	000.380	000.320	000.00
50	000.000	003.310	000.310	000.820	000.000	004.750	000.320	000.00
51	000.590	001.880	000.640	000.000	001.060	000.310	004.430	000.32
52	000.550	000.580	001.480	000.630	000.000	000.950	000.000	004.62
53	000.480	000.520	000.520	001.020	000.610	000.610	000.000	000.63
54	000.410	000.450	000.450	000.450	000.450	000.570	000.570	000.68
55	000.320	000.380	000.380	000.380	000.380	000.510	000.680	000.00
56	000.000	000.260	000.260	000.280	000.260	000.550	000.240	000.00
57	000.000	000.000	000.000	000.000	000.000	000.000	000.000	000.00
58	000.000	000.000	000.000	000.000	000.000	000.000	000.000	000.00
59	000.000	000.000	000.000	000.000	000.000	000.000	000.000	000.00
60	000.000	000.000	000.000	000.000	000.000	000.000	000.000	000.00
61	000.000	000.000	000.000	000.000	000.000	000.000	000.000	000.00
62	000.000	000.000	000.000	000.000	000.000	000.000	000.000	000.00
63	000.000	000.000	000.000	000.000	000.000	000.000	000.000	000.00
64	000.000	000.000	000.000	000.000	000.000	000.000	000.000	000.00

Figure 4.33. An image of the partial output of AGNPS flow calculations superimposed on the West Fork Little River Landsat ETM+ image with sample pixel values shown in the area where values ceased being calculated.

Comparison of AGNPS parameters acquisition methodologies

Once AGNPS had been run with parameters generated by Data Generator, it was a relatively simple task to modify the process to run AGNPS with parameters created in the Landsat ETM+ classification and modeling process. Three images created by classifying and post-classification modeling of the Landsat ETM+ image were substituted for images that had been created by Data Generator. The classified and modeled image of curve number (CN) was substituted for the Data Generator parameter 6 image. The classified and modeled image of Manning's n (n) was substituted for the Data Generator parameter 10 image. Finally, the classified and modeled image of surface condition constant (SCC) was substituted for the Data Generator parameter 14 image. Data Generator was then used to create a new stacked image, format that image into a ".dat" text file, run AGNPS with that file, and create images from the AGNPS output. This was done for both West Fork Little River and East Fork Little River watersheds.

CHAPTER 5

RESULTS

Coastal Plain watershed

The coastal plain watershed was used largely as a test bed for the methodology developed in this study. As a result, many of the methods used were in development and the results reflected the author's inexperience in field estimation of AGNPS parameters and LU/LC classes. The results in the Coastal Plain watershed were generally not as good as the results achieved in the Piedmont watersheds.

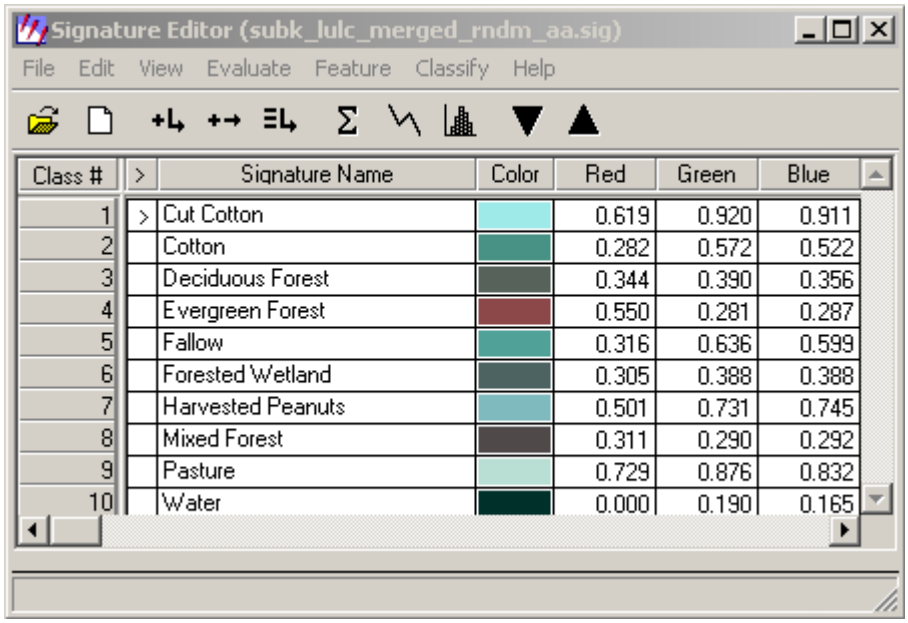
Using LU/LC classes

A portion of the Imagine Signature Editor window, depicting the classes and signatures used for classifying Subwatershed K into land use/land cover classes, is presented in Figures 5.1. A subset of the Landsat ETM+ image and the results of the classification of the Landsat ETM+ image within Subwatershed K, in terms of land use and land cover (LU/LC) are presented in Figure 5.2. Comparison of the classified LU/LC image in Figure 5.2 with the adjacent ETM+ image subset gives the impression of a better classification result than the LU/LC accuracy assessment indicated, as discussed in a subsequent section of this chapter.

Classifying imagery to AGNPS parameter values

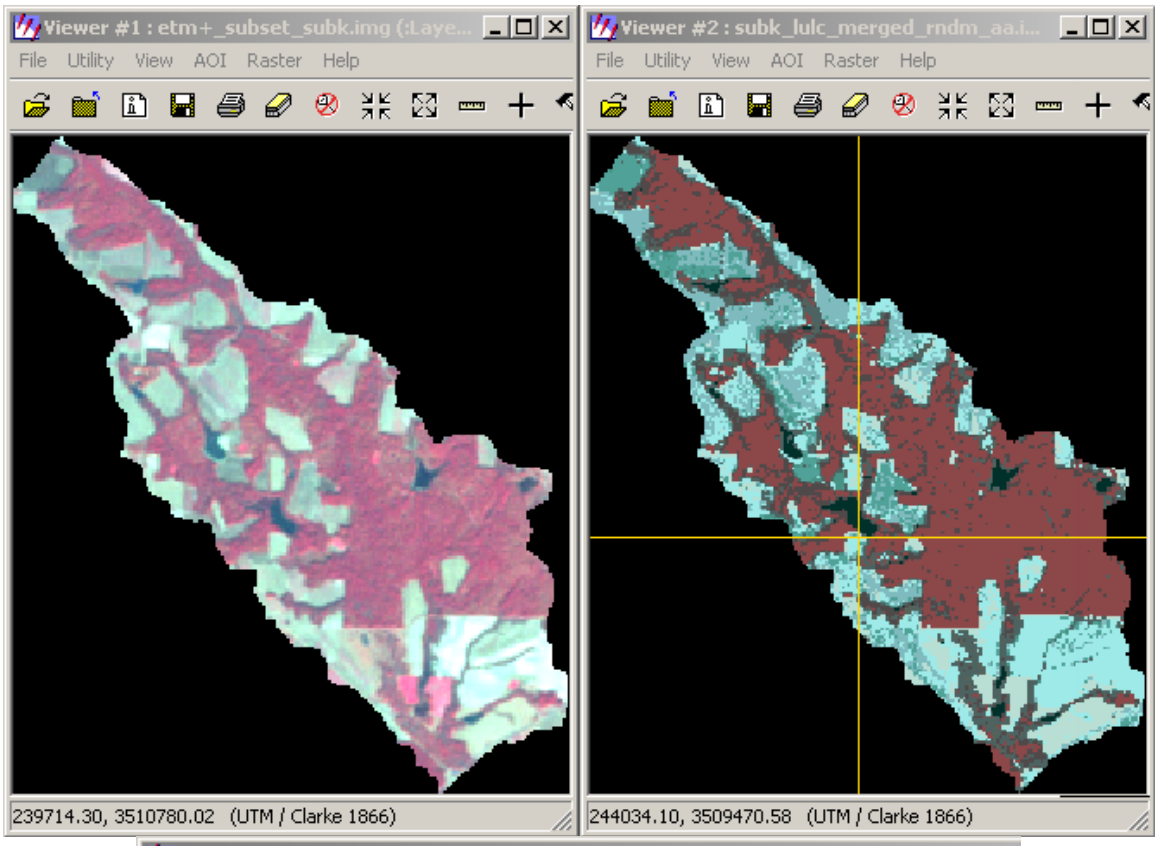
Figures 5.3 through 5.6 depict visualizations of the fuzzy classified images, distance files, and output files and samples of the resultant parameter values as shown in the pixel values of the output files for CN, n, P-fact, and SCC, respectively for

Subwatershed K. The samples of output pixel values for all four parameters were selected from the same area of the images, near the northwest (upper left) corner of the images where zero values for pixels outside the watershed give way to positive values for the parameters within the watershed.



Class #	Signature Name	Color	Red	Green	Blue
1	Cut Cotton		0.619	0.920	0.911
2	Cotton		0.282	0.572	0.522
3	Deciduous Forest		0.344	0.390	0.356
4	Evergreen Forest		0.550	0.281	0.287
5	Fallow		0.316	0.636	0.599
6	Forested Wetland		0.305	0.388	0.388
7	Harvested Peanuts		0.501	0.731	0.745
8	Mixed Forest		0.311	0.290	0.292
9	Pasture		0.729	0.876	0.832
10	Water		0.000	0.190	0.165

Figure 5.1. Imagine Signature Editor window depicting the classes and signatures used in classifying the Landsat ETM+ image in terms of LU/LC for Subwatershed K.



ImageInfo (subk_lulc_merged_rndm_aa.img)

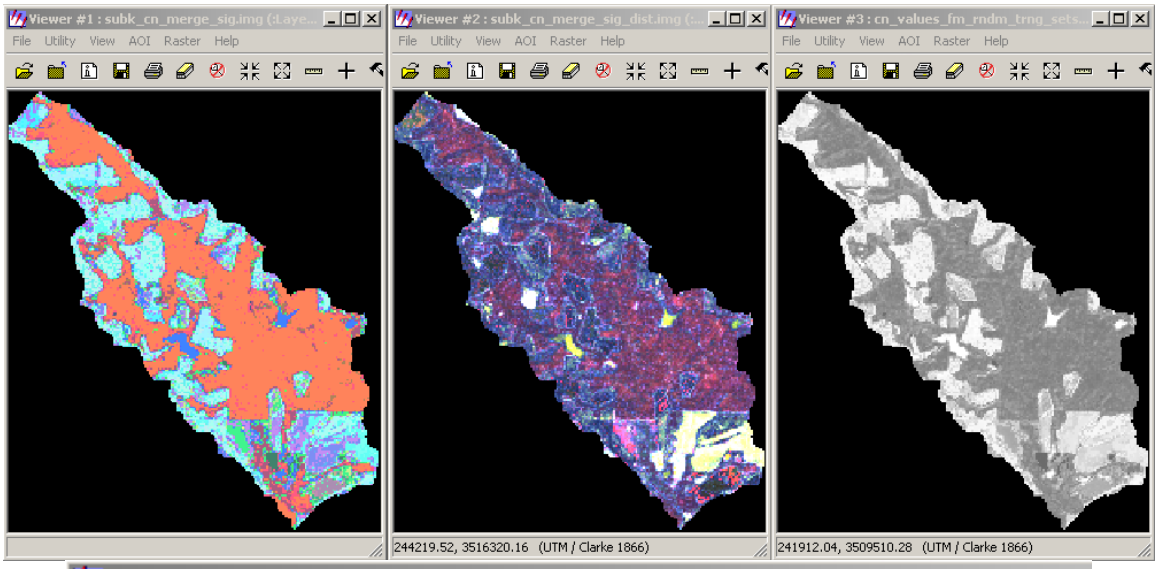
File Edit View Help

1 Layer_1

General | Projection | Histogram | Pixel data

Row	94	95	96	97	98	99
140	10	10	10	10	10	10
141	4	10	10	10	10	10
142	4	4	8	10	10	10
143	9	9	9	6	10	10
144	7	5	7	9	6	6
145	1	9	1	1	1	6
146	1	1	1	1	1	9
147	1	1	1	5	5	5
148	1	1	1	9	1	3
149	1	1	1	1	5	6
150	1	1	5	1	5	6
151	1	7	5	1	7	6

Figure 5.2. Clockwise from upper left, subset of the Landsat ETM+ image, classified image with crosshairs indicating the center of the area for which pixel values are shown, and a sample of output pixel values for LU/LC in Subwatershed K.



ImageInfo (cn_values_fm_rndm_trng_sets.img)

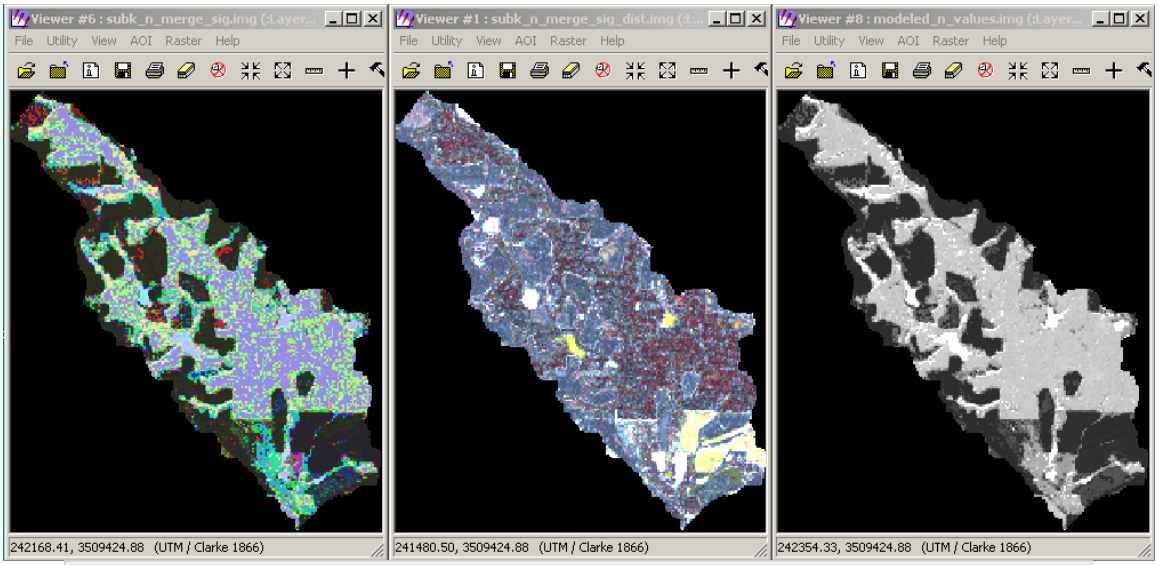
File Edit View Help

1 Layer_1

General | Projection | Histogram | Pixel data

Row	3	4	5	6	7	8
10	000.000	000.000	000.000	000.000	000.000	064.048
11	000.000	000.000	000.000	000.000	000.000	065.104
12	000.000	000.000	000.000	059.723	062.141	065.983
13	000.000	000.000	060.280	065.788	060.915	066.602
14	000.000	000.000	068.077	061.343	068.129	068.169
15	000.000	069.562	064.949	068.514	067.626	066.727
16	072.027	065.389	066.270	067.803	069.720	067.230
17	062.051	067.309	069.516	065.887	066.026	066.204
18	071.218	068.278	067.580	066.344	068.050	066.970
19	072.028	070.860	063.966	064.151	067.248	067.415
20	069.570	069.803	059.233	069.385	069.169	070.074

Figure 5.3. Clockwise from upper left: fuzzy classified image, distance file, output file, and a sample of output pixel values for CN in Subwatershed K.



ImageInfo (modeled_n_values.img)

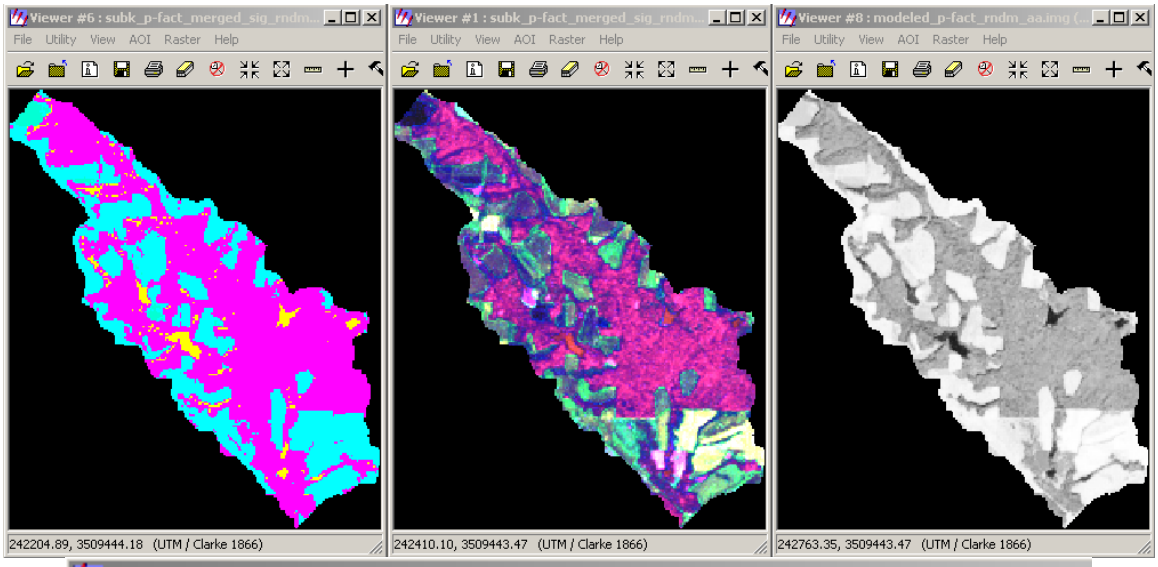
File Edit View Help

1 Layer_1

General | Projection | Histogram | Pixel data

Row	3	4	5	6	7	8
10	000.000	000.000	000.000	000.000	000.000	000.111
11	000.000	000.000	000.000	000.000	000.000	000.050
12	000.000	000.000	000.000	000.096	000.050	000.051
13	000.000	000.000	000.112	000.051	000.108	000.050
14	000.000	000.000	000.051	000.050	000.051	000.051
15	000.000	000.051	000.109	000.051	000.051	000.051
16	000.050	000.051	000.102	000.051	000.051	000.051
17	000.116	000.051	000.051	000.051	000.050	000.101
18	000.110	000.051	000.051	000.051	000.051	000.050
19	000.050	000.051	000.116	000.051	000.050	000.051
20	000.049	000.050	000.049	000.050	000.050	000.049

Figure 5.4. Clockwise from upper left: fuzzy classified image, distance file, output file, and a sample of output pixel values for n in Subwatershed K.



ImageInfo (modeled_p-fact_rndm_aa.img)

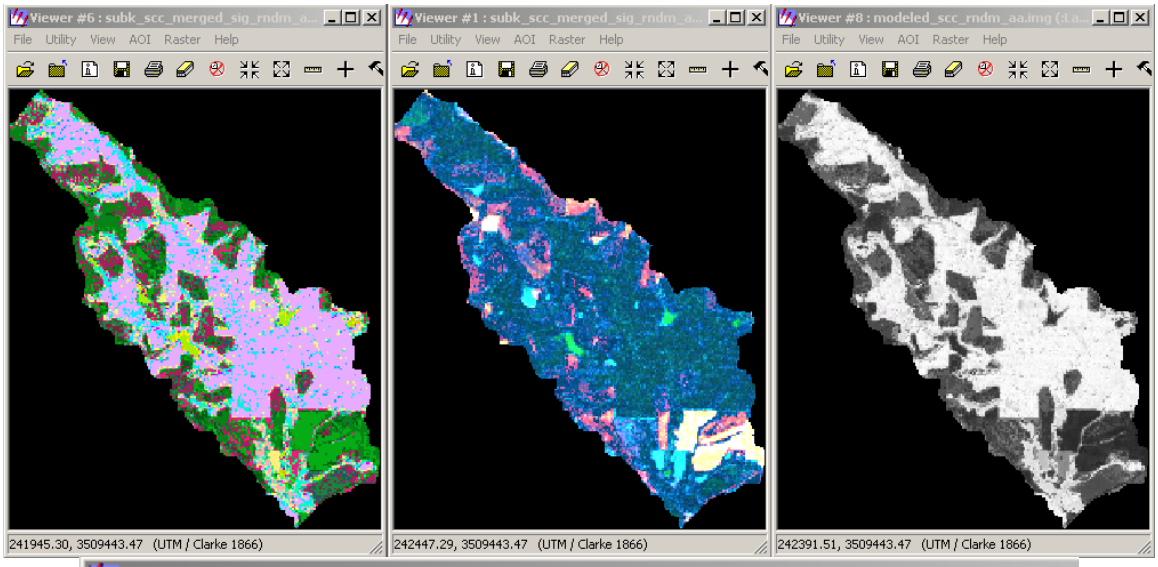
File Edit View Help

1 Layer_1

General | Projection | Histogram | Pixel data

Row	3	4	5	6	7	8
10	000.000	000.000	000.000	000.000	000.000	000.827
11	000.000	000.000	000.000	000.000	000.000	000.835
12	000.000	000.000	000.000	000.815	000.826	000.824
13	000.000	000.000	000.817	000.818	000.833	000.824
14	000.000	000.000	000.832	000.825	000.829	000.832
15	000.000	000.835	000.838	000.841	000.841	000.839
16	000.859	000.849	000.840	000.841	000.839	000.840
17	000.863	000.840	000.837	000.849	000.840	000.843
18	000.867	000.843	000.840	000.838	000.842	000.848
19	000.896	000.889	000.854	000.853	000.850	000.839
20	000.948	000.943	000.941	000.954	000.957	000.921

Figure 5.5. Clockwise from upper left: fuzzy classified image, distance file, output file, and a sample of output pixel values for P-factor in Subwatershed K.



ImageInfo (modeled_scc_rndm_aa.img)

File Edit View Help

1 Layer_1

General | Projection | Histogram | Pixel data

Row	3	4	5	6	7
10	000.000	000.000	000.000	000.000	000.000
11	000.000	000.000	000.000	000.000	000.000
12	000.000	000.000	000.000	000.192	000.195
13	000.000	000.000	000.295	000.192	000.295
14	000.000	000.000	000.189	000.195	000.195
15	000.000	000.197	000.284	000.190	000.194
16	000.189	000.196	000.196	000.194	000.194
17	000.299	000.197	000.197	000.192	000.190
18	000.293	000.190	000.198	000.191	000.187
19	000.186	000.182	000.295	000.187	000.193
20	000.123	000.122	000.126	000.112	000.106

Figure 5.6. Clockwise from upper left: fuzzy classified image, distance file, output file, and a sample of output pixel values for SCC in Subwatershed K.

The statistics presented in Table 5.1 provide summary information of the results of the classification and modeling process for Subwatershed K. It is important to note here that, while the image files containing the classified and modeled values of these constituents and from which these statistics were obtained were all rectangular raster

images, all values outside the boundaries of the subject watershed in each image were equal to zero and these values were not included in the calculation of the statistics in Table 5.1.

Table 5.1. Statistical summary of the Subwatershed K AGNPS model parameter values generated by ETM+ image classification and modeling.

Parameter	Minimum	Mean	Maximum	Dominant bin	% of pixels
CN	30.20389	50.2653	88.65240	40-50	33.6%
n	.04799	.16454	.79948	0.2-0.3	39.8%
P-fact	.11066	.77238	.98572	0.9-1.0	34.3%
SCC	.06822	.30777	.47464	0.4-0.5	46.3%

Accuracy Assessment

The Classification Accuracy Assessment Report for classification of Subwatershed K into LU/LC classes is presented in Table 5.2. The results for Subwatershed K, although lacking in statistical rigor due to scarcity of pixels used in the accuracy assessment, do clearly indicate some trends in the entire process of class definition, ground truth data acquisition, and classification. There was considerable doubt in the author's mind that cut cotton and harvested peanuts could be clearly distinguished in a Landsat ETM+ image. Yet, because it appeared on the ground that there were distinguishable differences in the appropriate values of AGNPS parameters for these two types of land cover, the decision was made to separate the classes. As the accuracy assessment report clearly indicates, though, the classes were not well distinguished in the supervised classification process into LU/LC values. The low producers accuracy for cut cotton was largely due to misclassification of cut cotton pixels

as harvested peanuts. Of the 114 cut cotton ground truth pixels, only 34 (29.82% of 114) were correctly classified as cut cotton. This, in turn, led to the low, 21.59%, users accuracy for harvested peanuts. Of the 88 ground truth pixels classified as harvested peanuts, only 19 (21.59% of 88) were actually harvested peanuts, while 69 were cut cotton. The methodology used to select pixels for accuracy assessment led to there being no pixels representing either fallow or mixed forest used in the accuracy assessment for Subwatershed K. Cotton, deciduous forest, evergreen forest, and forested wetlands were so underrepresented in the accuracy assessment that it is difficult to generalize based on the results, although the good results obtained with the small samples for cotton, deciduous forest, and evergreen forest are not surprising, based on observations in the field and the apparent separability of these classes in the ETM+ imagery. The confusion between water and forested wetlands is also not surprising, considering the small size of water bodies in the watershed and the inevitable mixed pixels around the edges of water bodies. Generally good results were obtained with both pasture and water because these classes were quite distinctive in the imagery. The overall classification accuracy of 50.57% and the overall Kappa statistics of 0.3986 were disappointing and were partially responsible for improvements in the entire process used in West Fork Little River.

Table 5.2. Accuracy assessment of LU/LC classification of Subwatershed K produced by Imagine Accuracy Assessment module.

CLASSIFICATION ACCURACY ASSESSMENT REPORT

Image File : c:/documents and settings/rfuller/my documents/dissertation/coastal plain/imagine-working-cp/landsat etm+ working/lu-lc/subk_lulc_merged_rndm_aa.img
User Name : rcfuller
Date : Sun Jan 19 15:27:16 2003

ERROR MATRIX

	Reference Data			
Classified Data	Unclassified	Cut Cotton	Cotton	Deciduous
Unclassified	0	0	0	0
Cut Cotton	0	34	2	1
Cotton	0	0	5	0
Deciduous Forest	0	0	0	3
Evergreen Forest	0	0	0	0
Fallow	0	10	1	0
Forested Wetland	0	1	0	3
Harvested Peanuts	0	69	0	0
Mixed Forest	0	0	0	1
Pasture	0	0	0	0
Water	0	0	0	0
Column Total	0	114	8	8

	Reference Data			
Classified Data	Evergreen	Fallow	Forested W	Harvested Peanuts
Unclassified	0	0	0	0
Cut Cotton	0	0	2	6
Cotton	0	0	0	0
Deciduous Forest	0	0	0	0
Evergreen Forest	6	0	1	0
Fallow	0	0	0	0
Forested Wetland	0	0	7	0
Harvested Peanuts	0	0	0	19
Mixed Forest	0	0	1	0
Pasture	2	0	2	2
Water	0	0	0	0
Column Total	8	0	13	27

	Reference Data			
Classified Data	Mixed Forest	Pasture	Water	Row Total
Unclassified	0	0	0	0
Cut Cotton	0	13	1	59
Cotton	0	0	0	5
Deciduous Forest	0	0	0	3
Evergreen Forest	0	0	0	7
Fallow	0	0	0	11
Forested Wetland	0	0	12	23
Harvested Peanuts	0	0	0	88
Mixed Forest	0	0	1	3
Pasture	0	21	0	27
Water	0	0	39	39
Column Total	0	34	53	265

----- End of Error Matrix -----

ACCURACY TOTALS

Class Name	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy
Unclassified	0	0	0	---	---
Cut Cotton	114	59	34	29.82%	57.63%
Cotton	8	5	5	62.50%	100.00%
Deciduous Forest	8	3	3	37.50%	100.00%
Evergreen Forest	8	7	6	75.00%	85.71%
Fallow	0	11	0	---	---
Forested Wetland	13	23	7	53.85%	30.43%
Harvested Peanuts	27	88	19	70.37%	21.59%
Mixed Forest	0	3	0	---	---
Pasture	34	27	21	61.76%	77.78%
Water	53	39	39	73.58%	100.00%
Totals	265	265	134		

Overall Classification Accuracy = 50.57%

----- End of Accuracy Totals -----

KAPPA (K^) STATISTICS

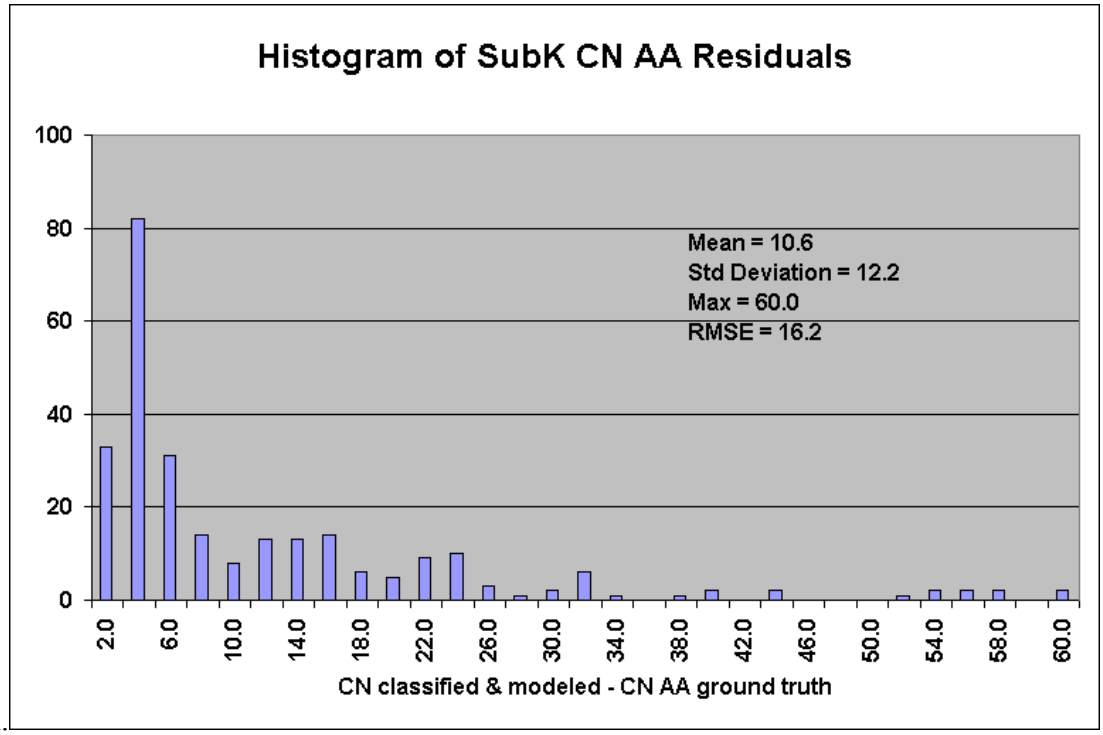
Overall Kappa Statistics = 0.3986

Conditional Kappa for each Category.

Class Name	Kappa
Unclassified	0.0000
Cut Cotton	0.2564
Cotton	1.0000
Deciduous Forest	1.0000
Evergreen Forest	0.8527
Fallow	0.0000
Forested Wetland	0.2685
Harvested Peanuts	0.1270
Mixed Forest	0.0000
Pasture	0.7451
Water	1.0000

The results of accuracy assessment of the fuzzy classification and modeling process into AGNPS parameter values in Subwatershed K are depicted in Figures 5.7 through 5.10. The mean value of 0.17 for the histogram of CN residuals divided by CN ground truth data values is interpreted to mean that the average error in the combined classification and modeling process for CN in Subwatershed K was seventeen percent of the ground truth values selected for accuracy assessment. Likewise, the mean value of 0.29 for the histogram of n residuals divided by n ground truth data values indicates that the average error in the combined classification and modeling process for n in Subwatershed K was twenty-nine percent.

A.



B.

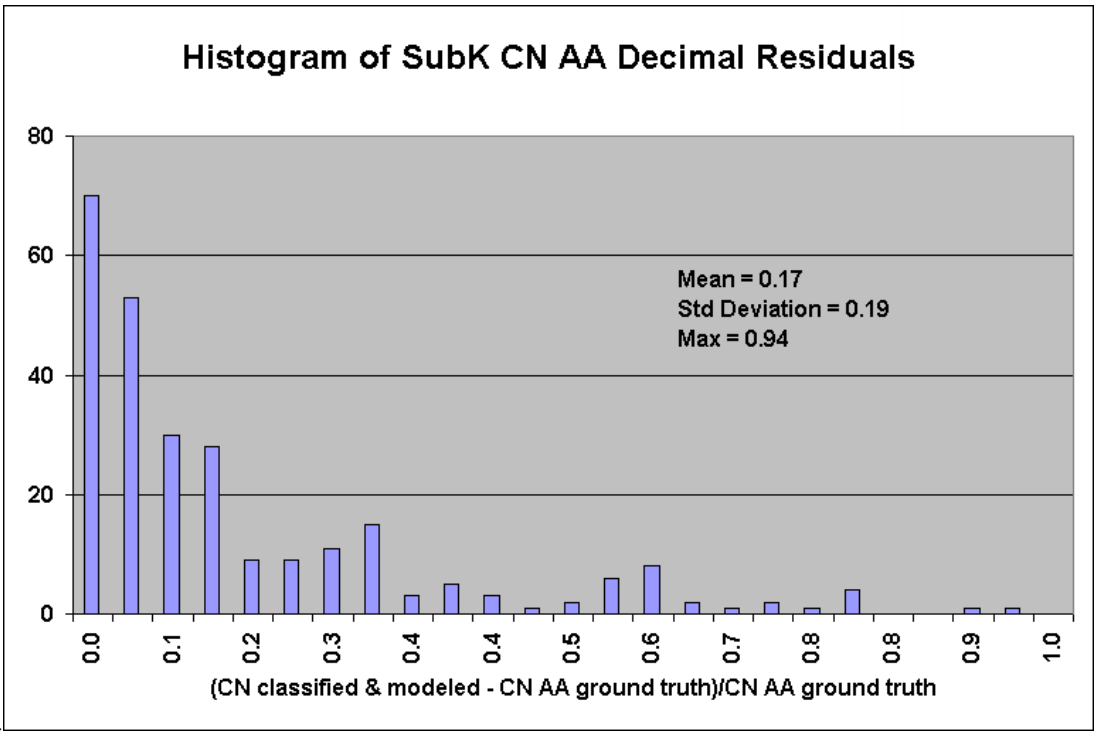
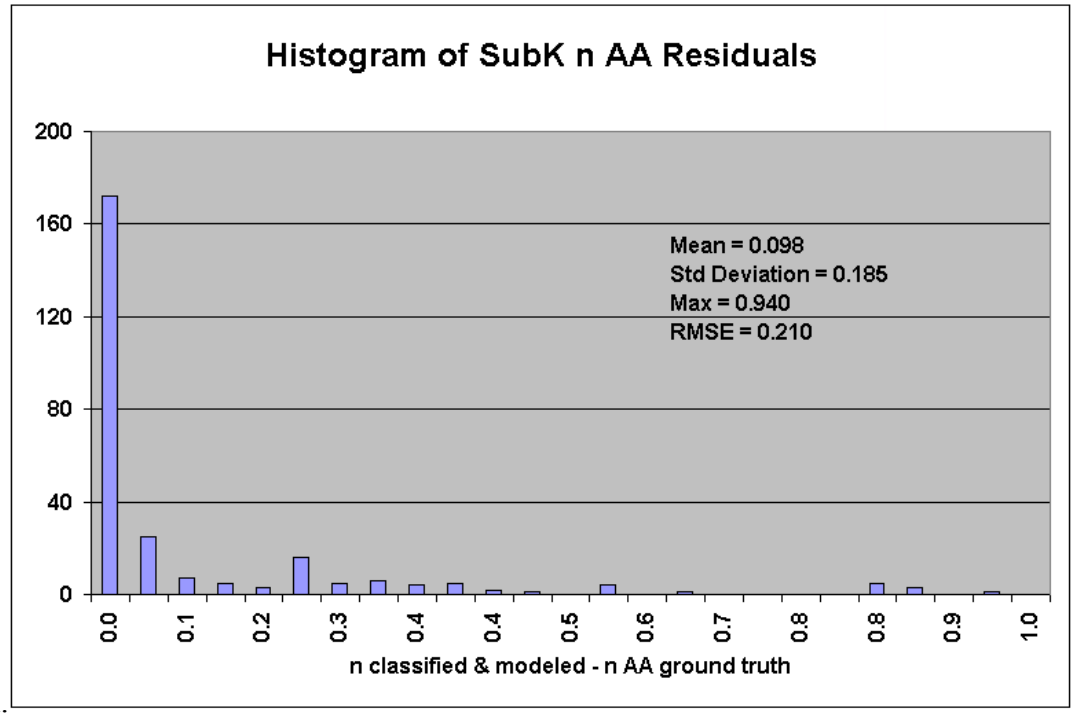


Figure 5.7. Residuals (predicted – observed) of the classification and modeling process for CN in Subwatershed K (A) and the decimal fraction of the residuals divided by observed values (B).

A.



B.

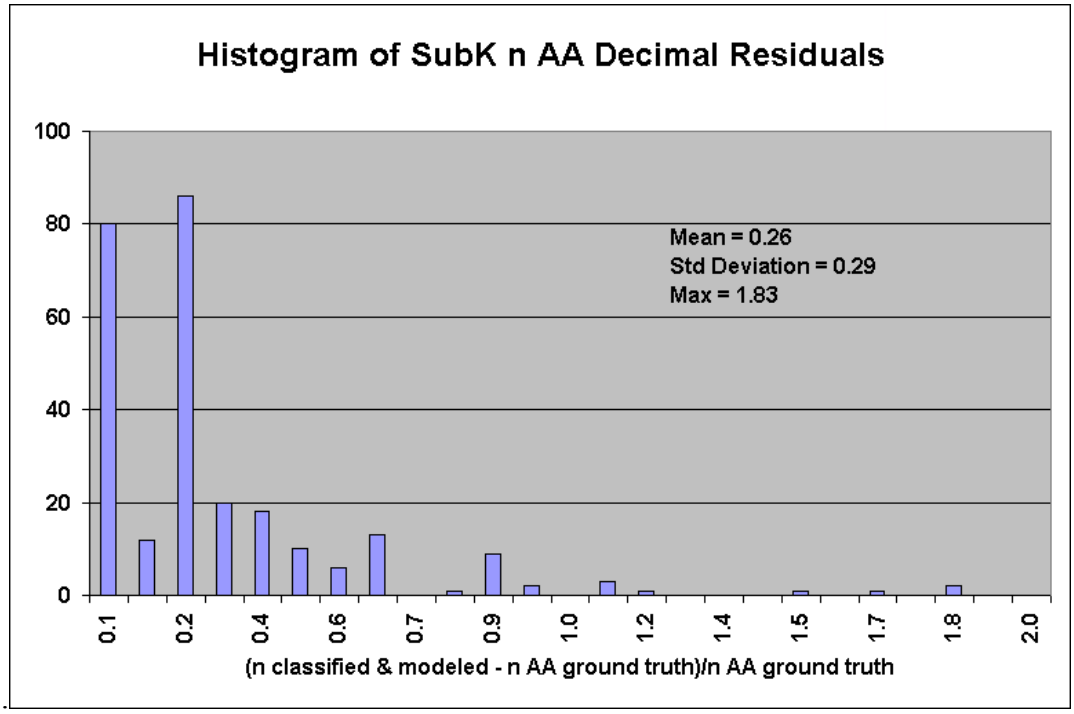
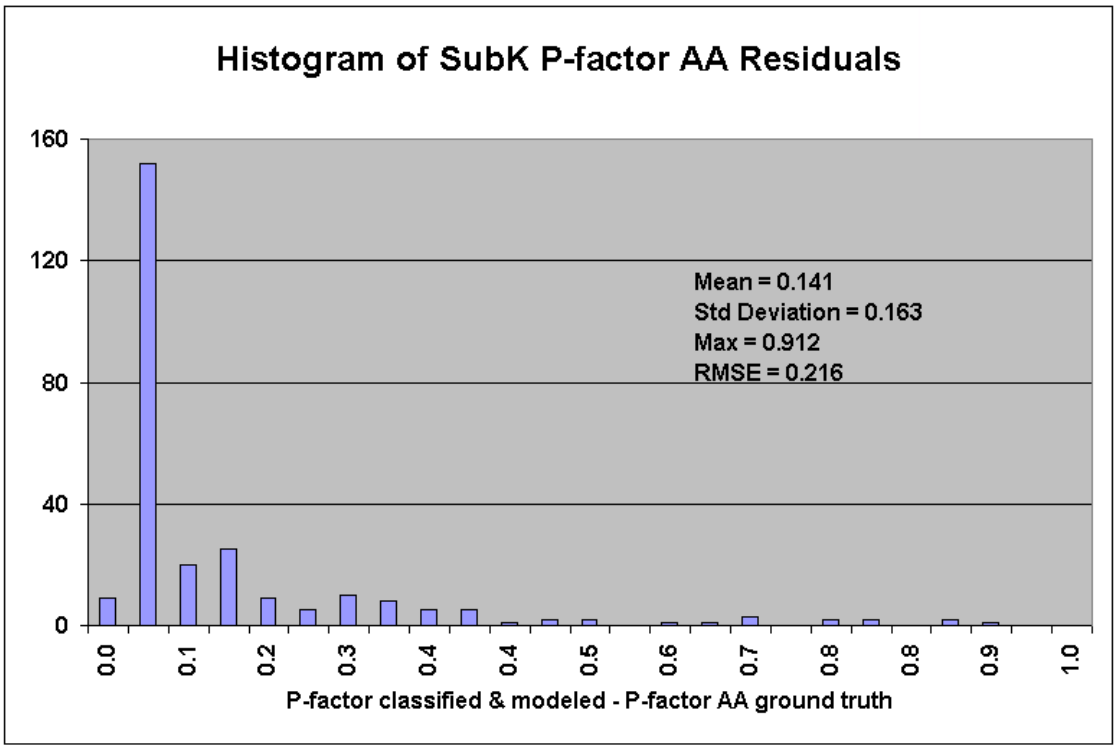


Figure 5.8. Residuals (predicted – observed) of the classification and modeling process for n in Subwatershed K (A) and the decimal fraction of the residuals divided by observed values (B).

Figures 5.9 and 5.10, showing the residuals for P-factor and SCC, respectively, are not as easy to interpret. The first graph in each figure is, in fact, fairly straightforward, showing the numerical values and selected statistics concerning the residuals for accuracy assessment pixels for the two parameters. In Figure 5.9, it is shown that the mean residual was 0.141 and the maximum was 0.942. When one considers that all values of P-factor in the ground truth data had values of either 0.00 or 1.00 and that the fuzzy classified and modeled clustered very close to these two values, the residuals make more sense than upon first inspection. The mean value of 0.141 actually indicates most modeled values were quite close to their respective ground truth values. However, the decimal values require a bit more explanation to understand. It should be noted that examination of the pixel values in the modeled P-factor image and the accuracy assessment ground truth image revealed that many of the accuracy assessment ground truth values were zero. Due to the restrictions against division by zero in the model, these values were set to 0.001 prior to their use in the model. The result was that the associated residuals were divided by a very small number, yielding a very large quotient. Much the same thing occurred with SCC residual values. As can be seen in the top half of Figure 5.10, the mean residual for SCC was 0.109 and the maximum was 0.422. Compared to the range of values, 0.000 to 0.590, the mean represents less than nineteen percent of the range, which is not unacceptably high. As was the case with P-factor, the decimal residuals for SCC presented in the bottom half of Figure 5.10 defies meaningful interpretation.

A.



B.

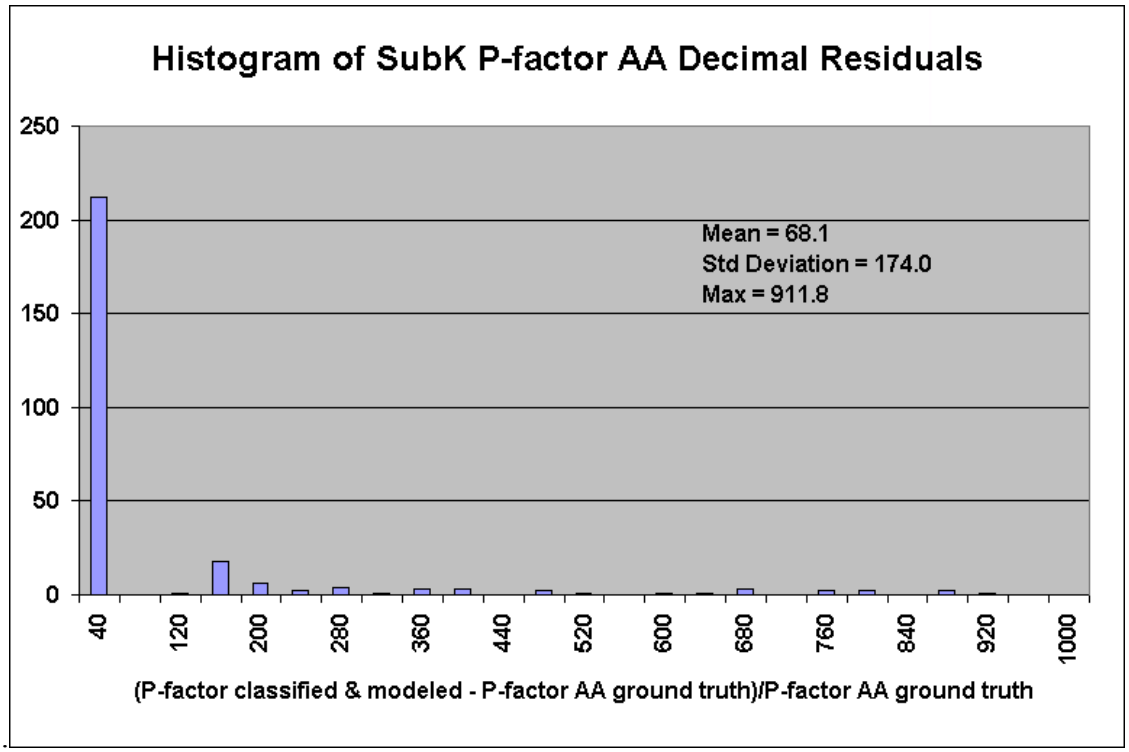
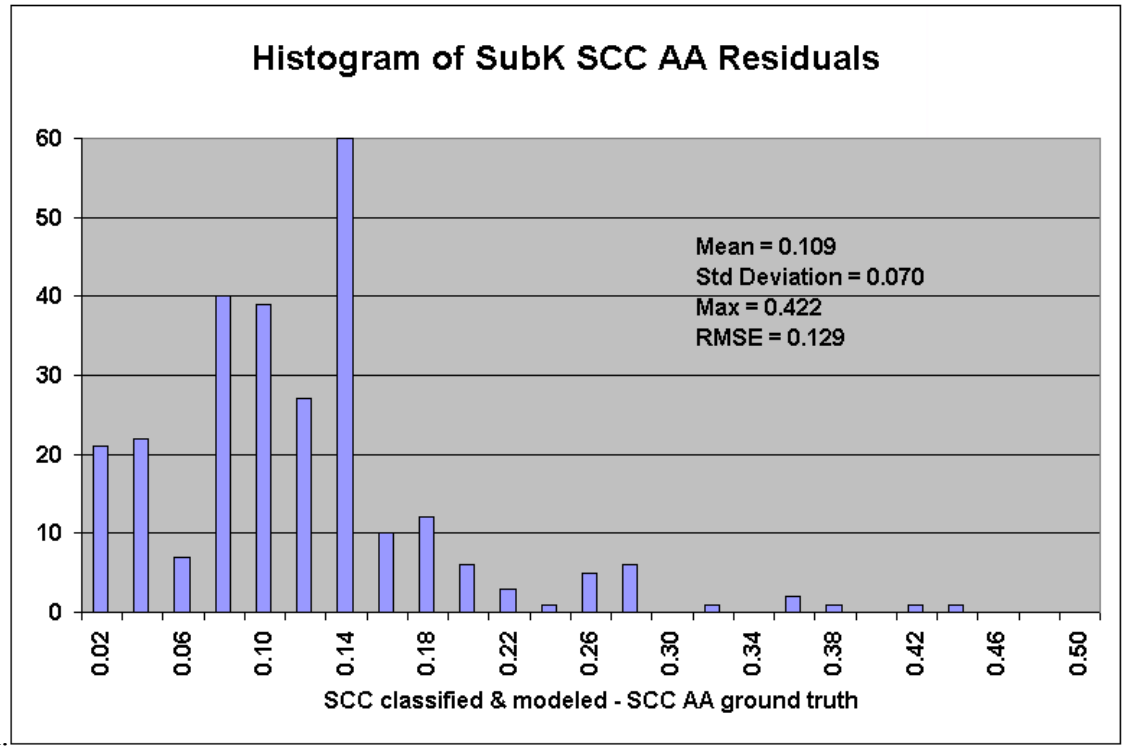


Figure 5.9. Residuals (predicted – observed) of the classification and modeling process for P-factor in Subwatershed K (A) and the decimal fraction of the residuals divided by observed values (B).

A.



B.

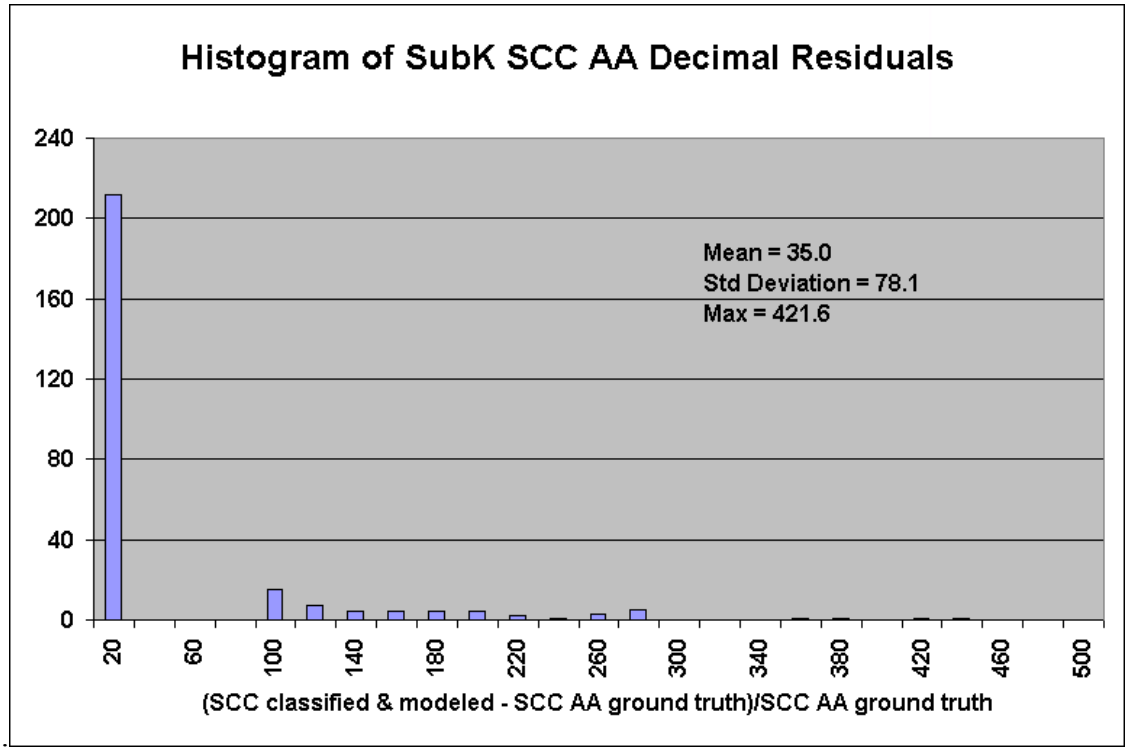


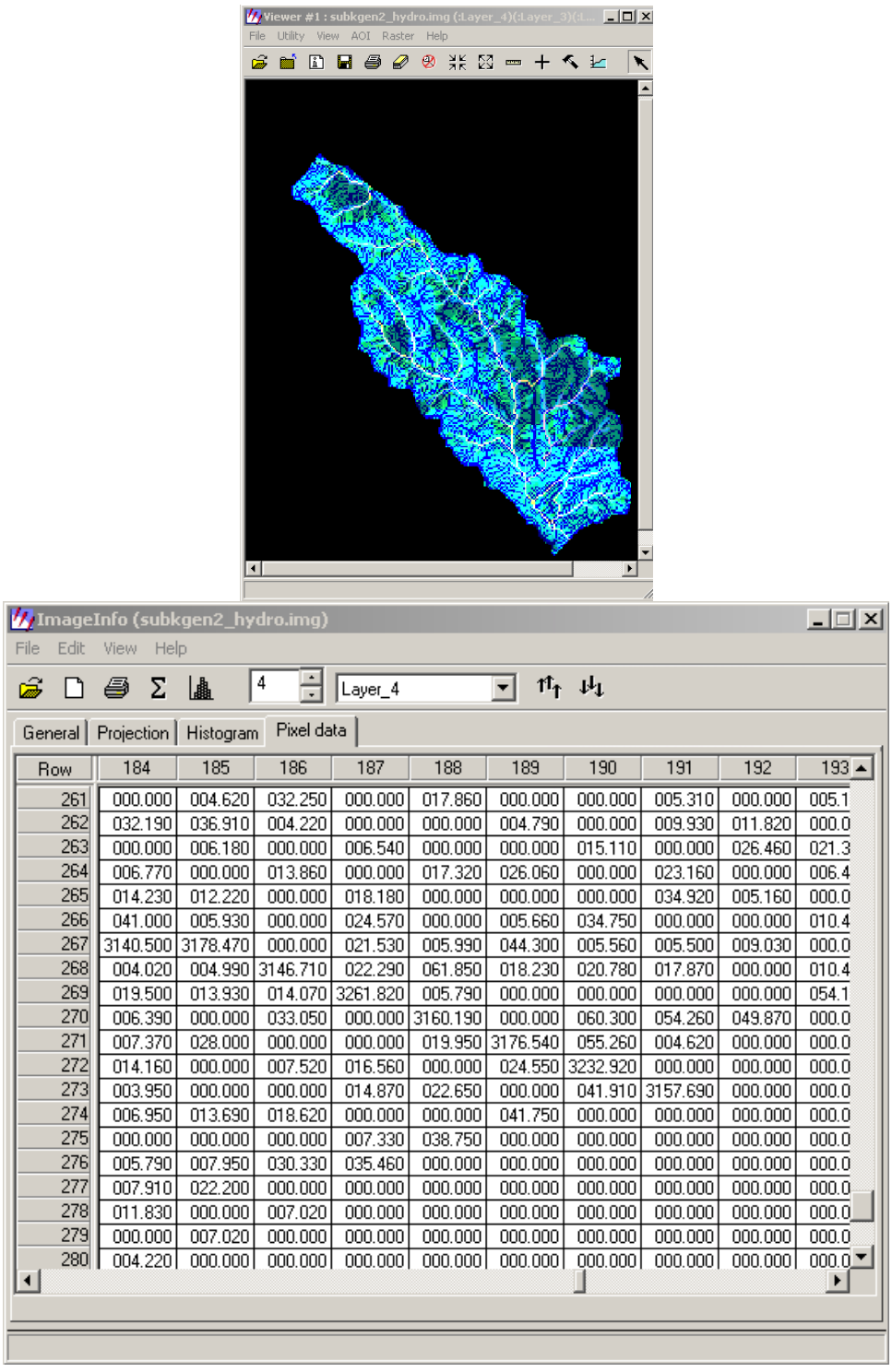
Figure 5.10. Residuals (predicted – observed) of the classification and modeling process for SCC in Subwatershed K (A) and the decimal fraction of the residuals divided by observed values (B).

AGNPS modeling results

Parameters for AGNPS were created and processed with AGNPS 5.0 in two ways. The first was with the use of Data Generator, extracting parameter values from a DEM, a soils image, and a land cover image. The second involved replacing images containing CN, n, and SCC values with those created by fuzzy classifying a Landsat ETM+ image and performing post-classification processing. Figure 5.11 is the Subwatershed K hydro image produced by Data Generator with representative flow values, in cfs. The stream channels can clearly be seen to be represented by the locations of the high flow values in the Image Information window. Blue in this image represents layer 2, which contains values of cell overland runoff in inches. Green represents layer 3, the accumulated runoff volume into each cell in inches. Red represents layer 4, the peak flow entering the cell in cfs. Figure 5.12 presents the same information for the AGNPS modeling using images of CN, n, and SCC developed by classifying and post-classification modeling of the Landsat ETM+ image as presented in Figure 5.11. The results of these two methods of AGNPS modeling of Subwatershed K are compared in Table 5.3.

Table 5.3. Selected results of AGNPS modeling of Subwatershed K using two methods of generating input data.

	All input created with Data Generator	Parameters CN, n, and SCC generated from Landsat ETM+ image	Ratios of the results of the two methods
Peak flow, cfs	3,158	1,475	0.47
Clay yield, tons	77,539	48,122	0.62
Silt yield, tons	186	52	0.28
Sand yield, tons	11	6	0.54
Total sediment yield, tons	77,805	48,219	0.62
Chemical oxygen demand (COD)	0	0	not applicable



Row	184	185	186	187	188	189	190	191	192	193
261	000.000	004.620	032.250	000.000	017.860	000.000	000.000	005.310	000.000	005.1
262	032.190	036.910	004.220	000.000	000.000	004.790	000.000	009.930	011.820	000.0
263	000.000	006.180	000.000	006.540	000.000	000.000	015.110	000.000	026.460	021.3
264	006.770	000.000	013.860	000.000	017.320	026.060	000.000	023.160	000.000	006.4
265	014.230	012.220	000.000	018.180	000.000	000.000	000.000	034.920	005.160	000.0
266	041.000	005.930	000.000	024.570	000.000	005.660	034.750	000.000	000.000	010.4
267	3140.500	3178.470	000.000	021.530	005.990	044.300	005.560	005.500	009.030	000.0
268	004.020	004.990	3146.710	022.290	061.850	018.230	020.780	017.870	000.000	010.4
269	019.500	013.930	014.070	3261.820	005.790	000.000	000.000	000.000	000.000	054.1
270	006.390	000.000	033.050	000.000	3160.190	000.000	060.300	054.260	049.870	000.0
271	007.370	028.000	000.000	000.000	019.950	3176.540	055.260	004.620	000.000	000.0
272	014.160	000.000	007.520	016.560	000.000	024.550	3232.920	000.000	000.000	000.0
273	003.950	000.000	000.000	014.870	022.650	000.000	041.910	3157.690	000.000	000.0
274	006.950	013.690	018.620	000.000	000.000	041.750	000.000	000.000	000.000	000.0
275	000.000	000.000	000.000	007.330	038.750	000.000	000.000	000.000	000.000	000.0
276	005.790	007.950	030.330	035.460	000.000	000.000	000.000	000.000	000.000	000.0
277	007.910	022.200	000.000	000.000	000.000	000.000	000.000	000.000	000.000	000.0
278	011.830	000.000	007.020	000.000	000.000	000.000	000.000	000.000	000.000	000.0
279	000.000	007.020	000.000	000.000	000.000	000.000	000.000	000.000	000.000	000.0
280	004.220	000.000	000.000	000.000	000.000	000.000	000.000	000.000	000.000	000.0

Figure 5.11. Subwatershed K AGNPS hydro image and sample pixel values using input data generated with Data Generator.

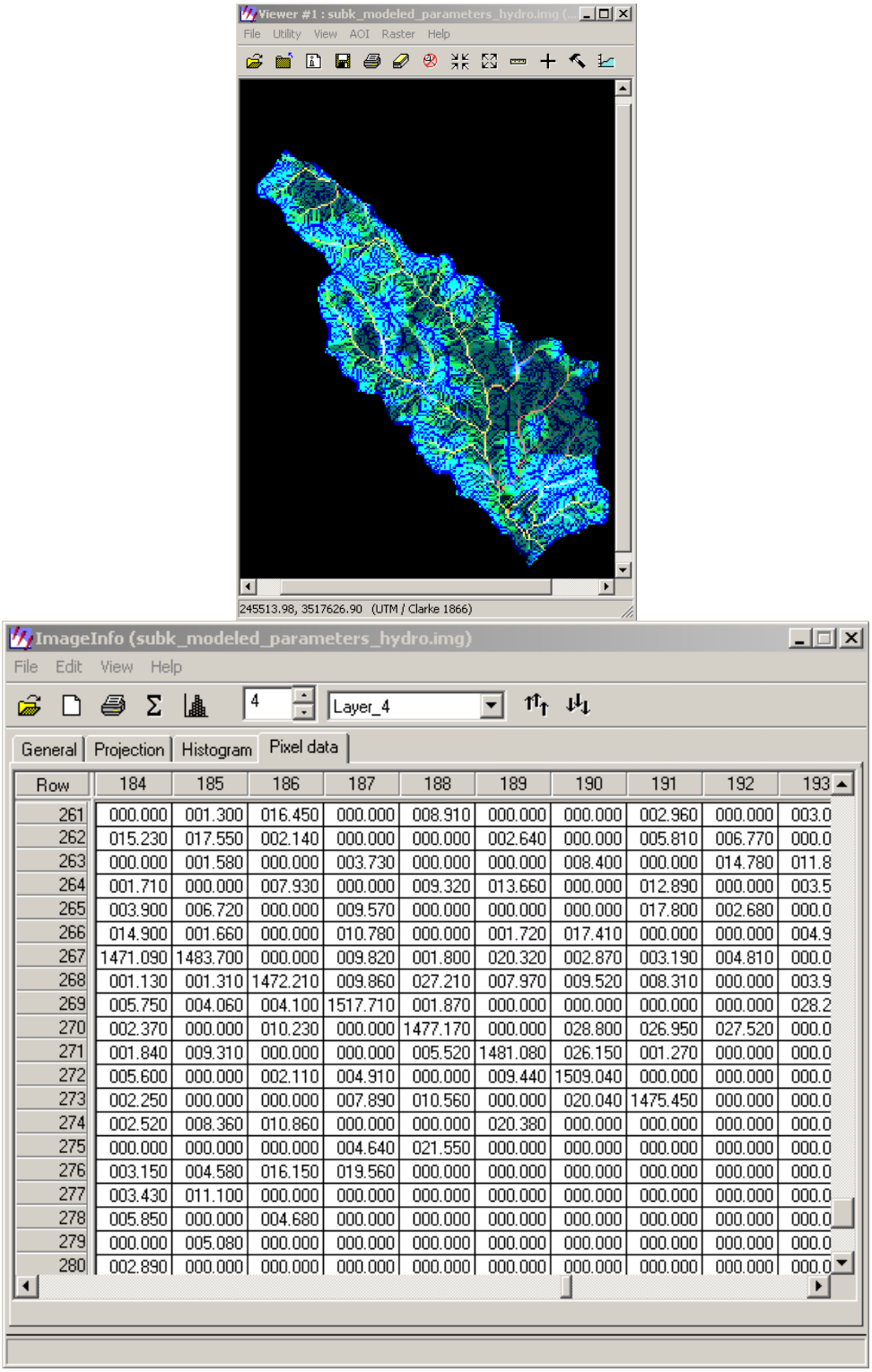
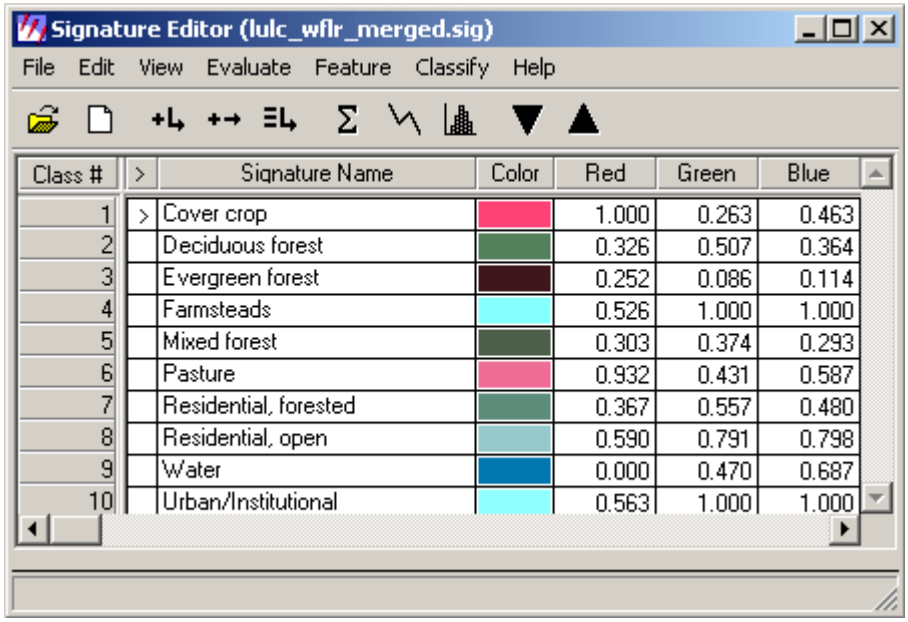


Figure 5.12. Subwatershed K AGNPS hydro image and sample pixel values substituting images for CN, n, and SCC created in the classification of a Landsat ETM+ image and post-classification modeling of those images.

Piedmont watersheds

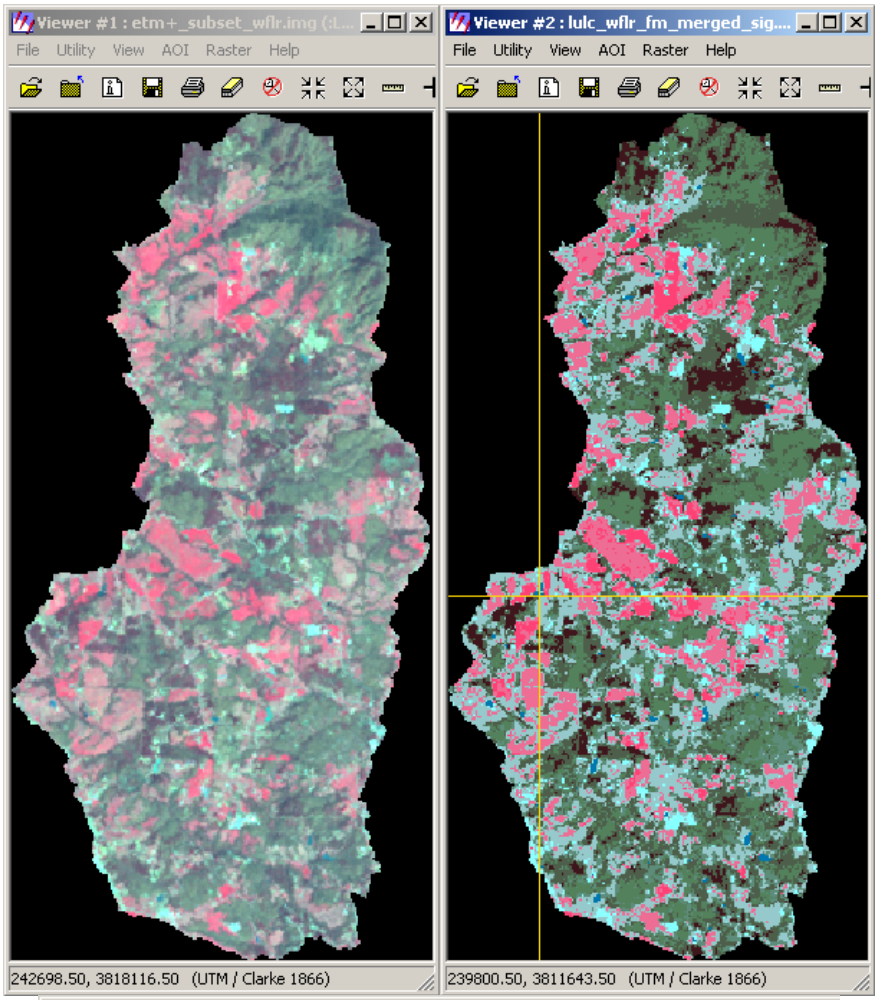
Using LU/LC classes

Figure 5.13 depicts a portion of the Imagine Signature Editor window, showing the classes and signatures used for classifying West Fork Little River and East Fork Little River into land use/land cover classes. A subset of the Landsat ETM+ image over northern Georgia and the results of the LU/LC classifications of that image within West Fork Little River and East Fork Little River are presented in Figures 5.14 and 5.15.



Class #	Signature Name	Color	Red	Green	Blue
1	Cover crop		1.000	0.263	0.463
2	Deciduous forest		0.326	0.507	0.364
3	Evergreen forest		0.252	0.086	0.114
4	Farmsteads		0.526	1.000	1.000
5	Mixed forest		0.303	0.374	0.293
6	Pasture		0.932	0.431	0.587
7	Residential, forested		0.367	0.557	0.480
8	Residential, open		0.590	0.791	0.798
9	Water		0.000	0.470	0.687
10	Urban/Institutional		0.563	1.000	1.000

Figure 5.13. Imagine Signature Editor window depicting the classes and signatures derived from West Fork Little River and used in classifying the Landsat ETM+ image in terms of LU/LC for West Fork Little River and East Fork Little River.



ImageInfo (lulc_wflr_fm_merged_sig.img)

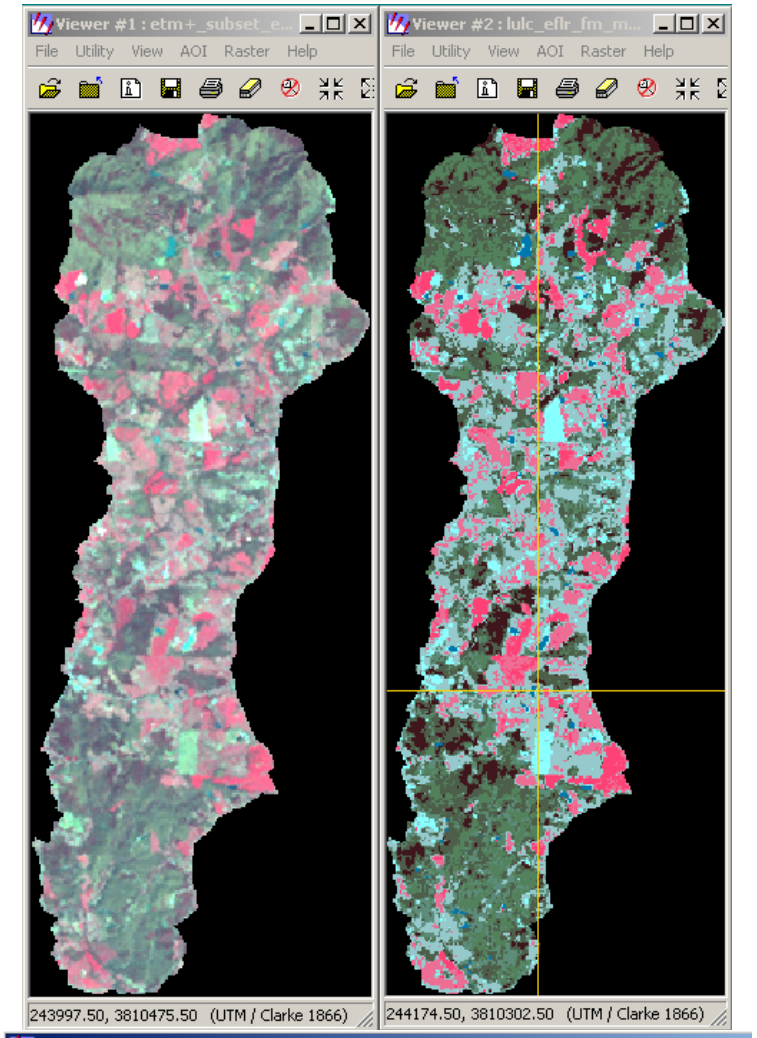
File Edit View Help

1 Layer_1

General | Projection | Histogram | Pixel data

Row	40	41	42	43	44	45
230	1	1	8	6	9	9
231	1	1	1	8	9	9
232	1	1	1	8	9	9
233	1	1	9	9	9	10
234	8	5	9	9	9	9
235	3	3	9	9	5	5
236	3	5	5	5	5	8
237	5	5	5	8	5	8
238	7	5	8	8	8	8
239	5	7	7	8	8	8
240	7	7	8	8	8	8

Figure 5.14. Clockwise from upper left, subset of the Landsat ETM+ image, classified image with crosshairs indicating the center of the area for which pixel values are shown, and a sample of output pixel values for LU/LC in West Fork Little River watershed.



ImageInfo (lulc_eflr_fm_merged_sig2.img)

File Edit View Help

1 Layer_1

General | Projection | Histogram | Pixel data

Row	71	72	73	74	75	76
278	5	3	3	3	3	3
279	8	8	8	8	5	3
280	8	8	6	8	8	3
281	9	9	7	8	7	8
282	9	9	10	7	8	1
283	9	10	8	7	8	6
284	9	10	8	7	5	8
285	9	8	8	8	8	8
286	8	7	8	8	8	8
287	8	8	8	8	6	8
288	10	8	8	6	8	8

Figure 5.15. Clockwise from upper left, subset of the Landsat ETM+ image, classified image with crosshairs indicating the center of the area for which pixel values are shown, and a sample of output pixel values for LU/LC in East Fork Little River watershed.

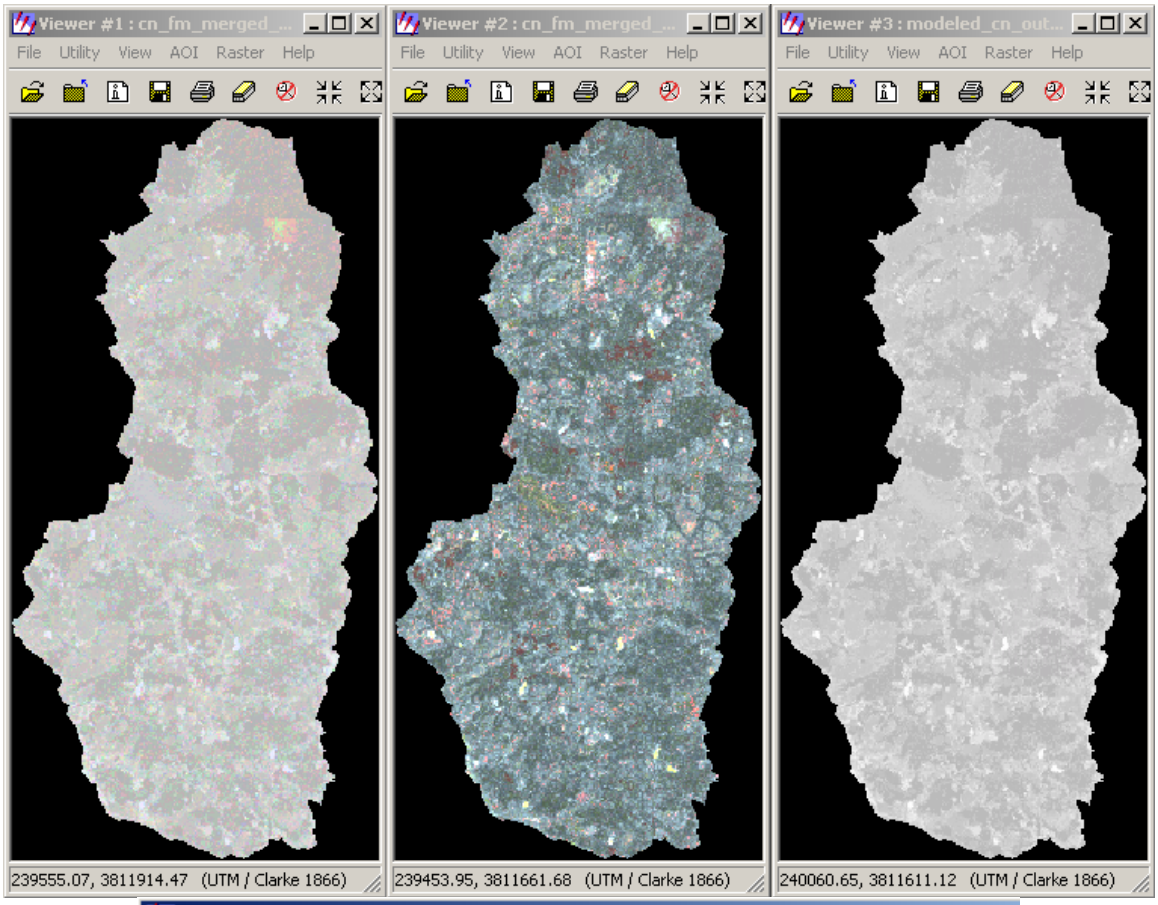
Classifying imagery to AGNPS parameter values

Figures 5.16 through 5.18 depict visualizations of the fuzzy classified images, distance files, and output files and samples of the resultant parameter values as shown in the pixel values of the output files for CN, n, and SCC, respectively for West Fork Little River. Comparable images and parameter values for East Fork Little River are presented in Figures 5.19 through 5.20. The Landsat ETM+ image was not classified for P-fact in the Piedmont watersheds because all ground truth values for P-fact were 1.00. The samples of output pixel values for all three parameters in both Piedmont watersheds were selected from the same areas of the images as were used in the LU/LC classification, in order to facilitate comparisons.

The statistics presented in Tables 5.4 and 5.5 provide summary information about the results of the classification and modeling process for West Fork Little River and East Fork Little River. As was the case with Subwatershed K, the image files containing the classified and modeled values of these constituents and to which these statistics apply were all rectangular raster images and all values outside the boundaries of the subject watershed in each image were equal to zero. These values were not included in the calculation of the statistics in Tables 5.4 and 5.5.

Table 5.4. Statistical summary of the West Fork Little River AGNPS model parameter values generated by ETM+ image classification and modeling.

Parameter	Minimum	Mean	Maximum	Dominant bin	% of pixels
CN	71.26852	77.66381	98.42751	70-80	70.8%
n	.01263	.22237	.73726	0.3-0.4	26.0%
SCC	.00537	.24611	.52392	0.1-0.2	27.2%



ImageInfo (modeled_cn_output_wflr_merged_sig.img)

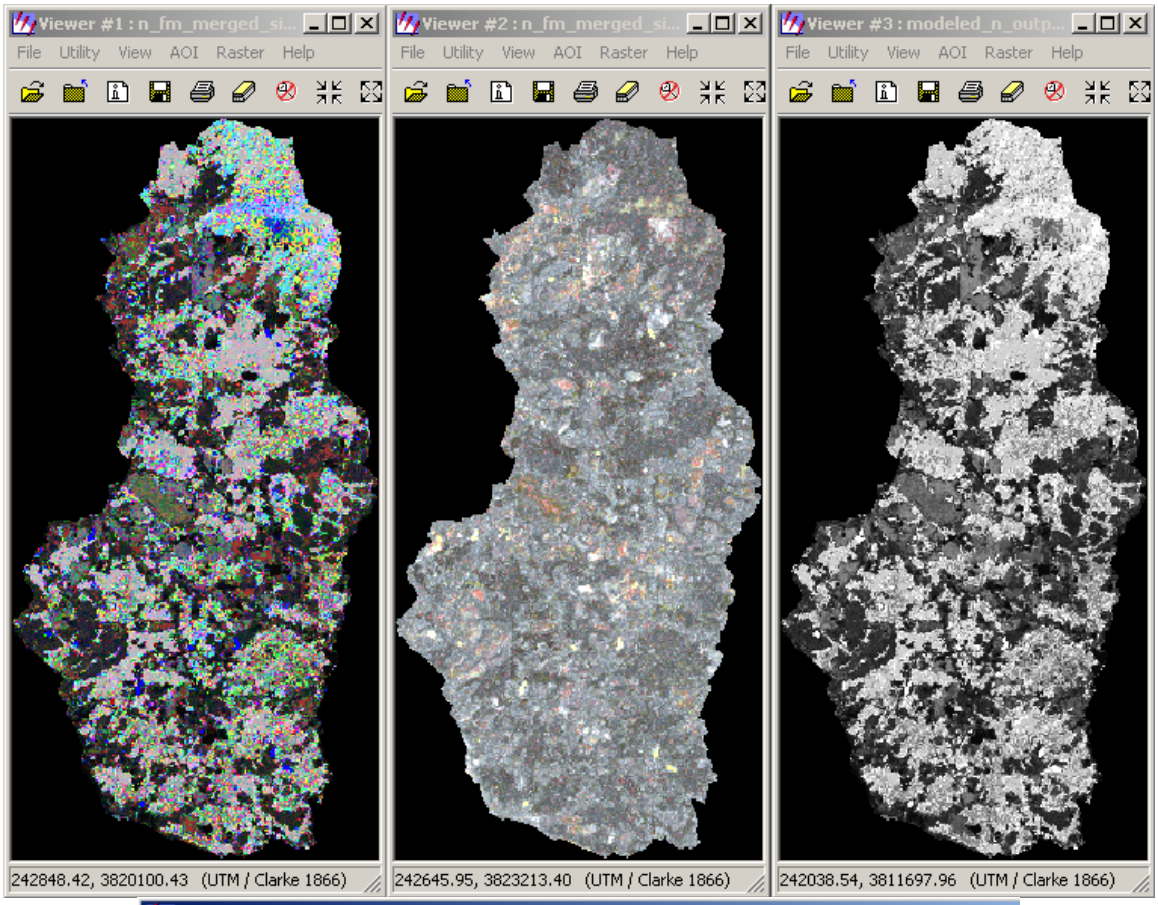
File Edit View Help

1 Layer_1

General | Projection | Histogram | Pixel data

Row	40	41	42	43	44	45
230	078.634	079.038	077.660	078.263	094.158	095.884
231	077.981	078.751	079.022	082.720	094.229	095.710
232	078.091	078.314	078.321	081.206	091.532	095.886
233	078.517	078.202	095.062	096.421	095.749	086.455
234	078.929	078.845	095.820	095.698	095.360	094.834
235	072.122	071.873	094.457	095.977	077.792	072.489
236	072.122	072.320	074.386	074.880	074.893	078.090
237	071.458	074.439	076.923	082.303	080.580	080.311
238	076.333	072.341	081.450	077.999	077.936	079.554
239	080.194	076.594	077.343	078.047	079.596	079.518
240	081.143	076.701	077.851	079.706	079.726	079.649

Figure 5.16. Clockwise from upper left: fuzzy classified image, distance file, output file, and a sample of output pixel values for CN in West Fork Little River.



ImageInfo (modeled_n_output_wflr_merged_sig.img)

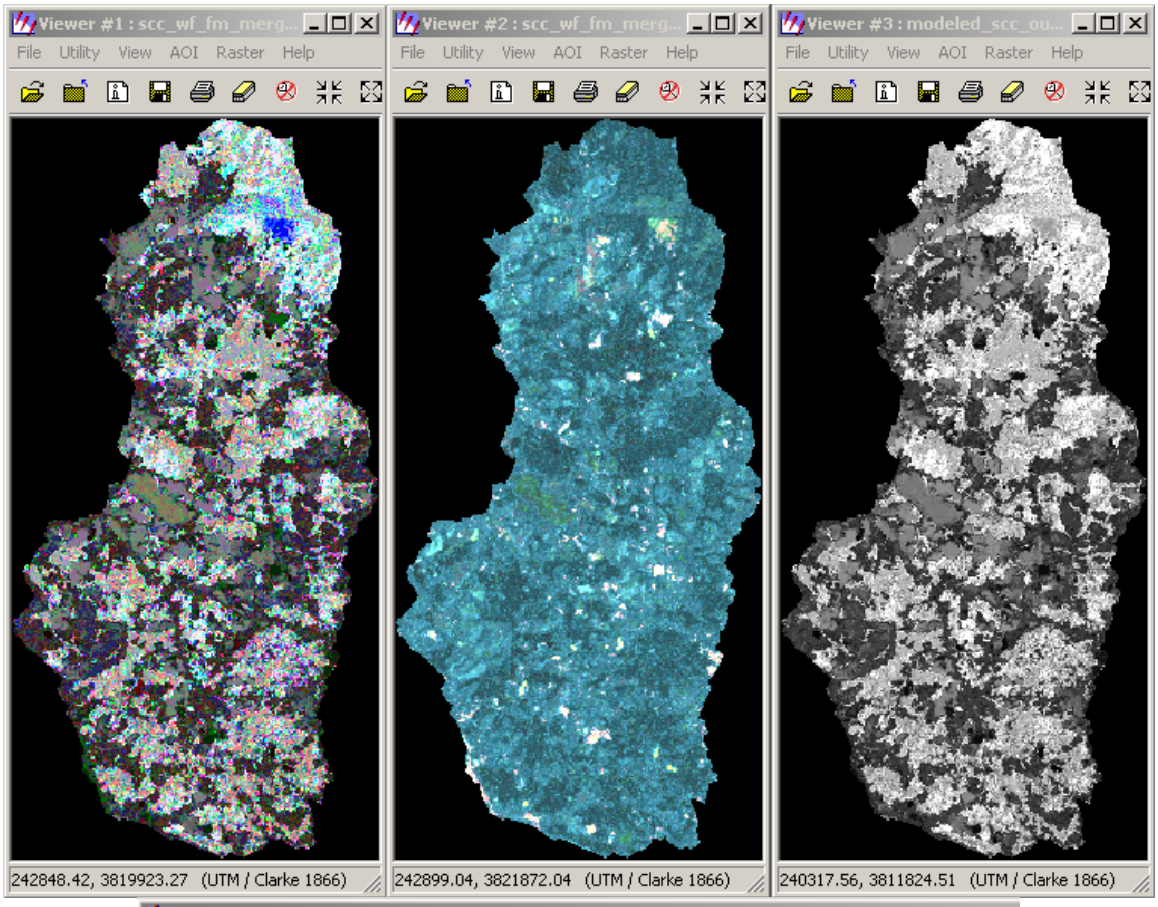
File Edit View Help

1 Layer_1

General | Projection | Histogram | Pixel data

Row	40	41	42	43	44	45
230	000.249	000.211	000.080	000.081	000.343	000.535
231	000.216	000.194	000.184	000.077	000.418	000.426
232	000.215	000.253	000.215	000.065	000.361	000.429
233	000.191	000.216	000.476	000.573	000.530	000.068
234	000.125	000.416	000.534	000.514	000.488	000.433
235	000.385	000.343	000.400	000.550	000.454	000.466
236	000.343	000.348	000.356	000.480	000.466	000.051
237	000.333	000.393	000.374	000.179	000.240	000.051
238	000.378	000.343	000.051	000.076	000.076	000.078
239	000.257	000.377	000.155	000.077	000.077	000.078
240	000.087	000.285	000.077	000.077	000.077	000.077

Figure 5.17. Clockwise from upper left: fuzzy classified image, distance file, output file, and a sample of output pixel values for n in West Fork Little River.



ImageInfo (modeled_scc_output_wflr_merged_sig.img)

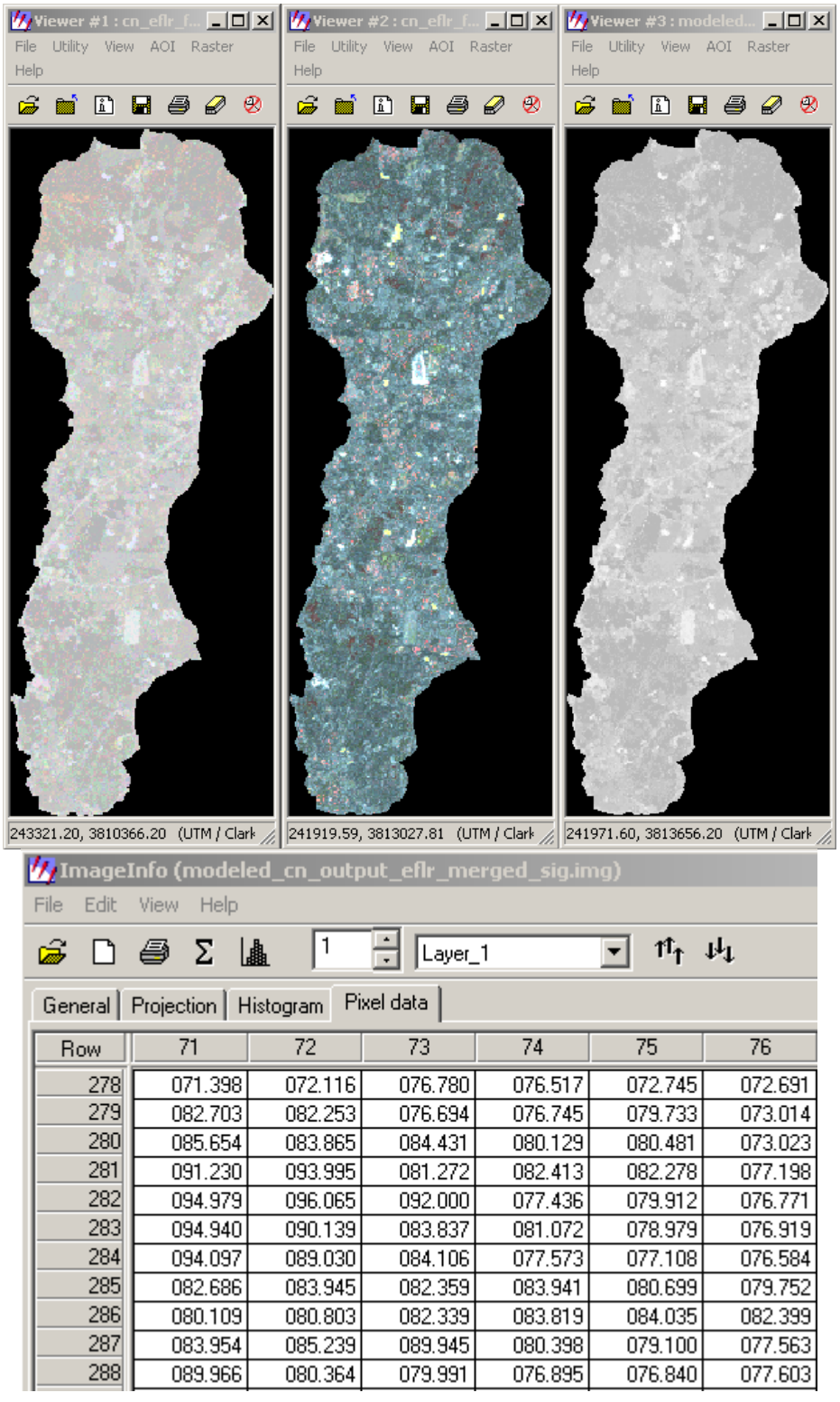
File Edit View Help

1 Layer_1

General | Projection | Histogram | Pixel data

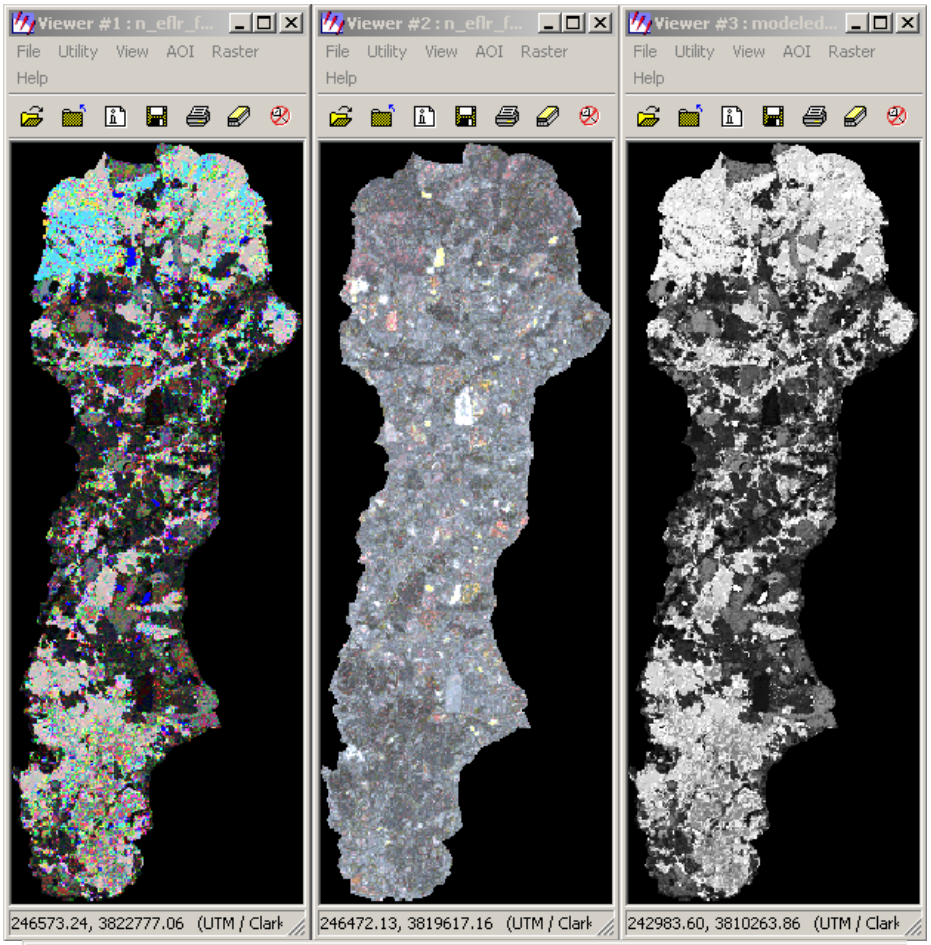
Row	40	41	42	43	44	45
230	000.271	000.267	000.116	000.157	000.079	000.012
231	000.268	000.272	000.195	000.076	000.014	000.017
232	000.268	000.269	000.233	000.041	000.068	000.016
233	000.270	000.269	000.014	000.010	000.013	000.054
234	000.264	000.244	000.012	000.012	000.014	000.099
235	000.275	000.317	000.184	000.012	000.399	000.258
236	000.352	000.308	000.352	000.434	000.402	000.113
237	000.385	000.403	000.373	000.077	000.077	000.111
238	000.389	000.358	000.077	000.115	000.118	000.172
239	000.077	000.135	000.135	000.115	000.118	000.120
240	000.077	000.078	000.130	000.117	000.076	000.121

Figure 5.18. Clockwise from upper left: fuzzy classified image, distance file, output file, and a sample of output pixel values for SCC in West Fork Little River.



Row	71	72	73	74	75	76
278	071.398	072.116	076.780	076.517	072.745	072.691
279	082.703	082.253	076.694	076.745	079.733	073.014
280	085.654	083.865	084.431	080.129	080.481	073.023
281	091.230	093.995	081.272	082.413	082.278	077.198
282	094.979	096.065	092.000	077.436	079.912	076.771
283	094.940	090.139	083.837	081.072	078.979	076.919
284	094.097	089.030	084.106	077.573	077.108	076.584
285	082.686	083.945	082.359	083.941	080.699	079.752
286	080.109	080.803	082.339	083.819	084.035	082.399
287	083.954	085.239	089.945	080.398	079.100	077.563
288	089.966	080.364	079.991	076.895	076.840	077.603

Figure 5.19. Clockwise from upper left: fuzzy classified image, distance file, output file, and a sample of output pixel values for CN in East Fork Little River.



ImageInfo (modeled_n_output_eflr_merged_sig.img)

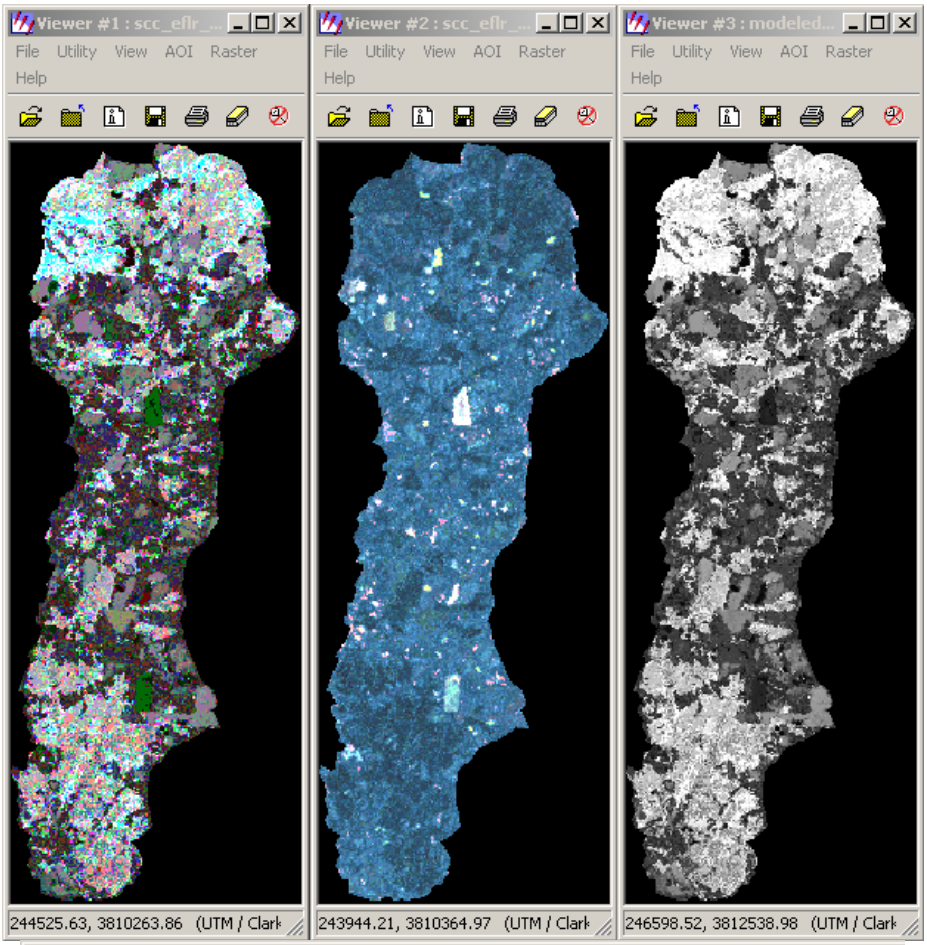
File Edit View Help

1 Layer_1

General | Projection | Histogram | Pixel data

Row	71	72	73	74	75	76
278	000.377	000.375	000.164	000.328	000.440	000.343
279	000.079	000.098	000.112	000.127	000.238	000.298
280	000.051	000.077	000.076	000.050	000.180	000.315
281	000.369	000.347	000.082	000.076	000.146	000.108
282	000.436	000.557	000.045	000.149	000.051	000.184
283	000.441	000.034	000.087	000.168	000.050	000.141
284	000.362	000.016	000.051	000.117	000.279	000.160
285	000.346	000.076	000.058	000.076	000.057	000.077
286	000.088	000.104	000.088	000.067	000.076	000.077
287	000.077	000.048	000.037	000.050	000.126	000.091
288	000.016	000.049	000.077	000.159	000.090	000.090

Figure 5.20. Clockwise from upper left: fuzzy classified image, distance file, output file, and a sample of output pixel values for n in East Fork Little River.



ImageInfo (modeled_scc_output_eflr_merged_sig.img)

File Edit View Help

1 Layer_1

General | Projection | Histogram | Pixel data

Row	71	72	73	74	75	76
278	000.356	000.277	000.198	000.181	000.354	000.359
279	000.109	000.310	000.204	000.266	000.270	000.207
280	000.116	000.077	000.076	000.109	000.112	000.349
281	000.121	000.020	000.076	000.077	000.077	000.113
282	000.015	000.011	000.054	000.166	000.117	000.290
283	000.015	000.036	000.076	000.166	000.111	000.207
284	000.020	000.037	000.076	000.077	000.077	000.237
285	000.050	000.076	000.116	000.077	000.124	000.156
286	000.077	000.115	000.077	000.110	000.109	000.077
287	000.109	000.113	000.109	000.107	000.118	000.206
288	000.094	000.108	000.116	000.237	000.238	000.121

Figure 5.21. Clockwise from upper left: fuzzy classified image, distance file, output file, and a sample of output pixel values for SCC in East Fork Little River.

Table 5.5. Statistical summary of the East Fork Little River AGNPS model parameter values generated by ETM+ image classification and modeling.

Parameter	Minimum	Mean	Maximum	Dominant bin	% of pixels
CN	71.25420	78.29622	98.47377	70-80	65.6%
n	.01254	.21652	.73868	0.3-0.4	24.5%
SCC	.00524	.23666	.52459	0.1-0.2	28.3%

Accuracy Assessment

The results of classification into LU/LC classes in West Fork Little River were both better and more statistically valid than that for Subwatershed K, as shown in Table 5.6. The overall classification accuracy of 61.62% and the overall Kappa statistics of 0.5590 are both less than desirable (Foody, 2002) but these low numbers are largely due to discrimination problems between classes that had only very small differences in their respective associated AGNPS parameter values, so the confusion between these classes should not have significantly affected the final outcome of the AGNPS modeling process based on this classification. A close examination of the error matrix shows that the largest errors occurred between mixed forest, evergreen forest and deciduous forest; between deciduous forest and forested residential; and between pasture and open residential. The first of these three groupings was of little consequence, since the AGNPS parameters were nearly identical within these land uses. In the second grouping, an attempt was made to distinguish the differences in response to rainfall between the two land uses, though the anticipated differences were still somewhat small, due to the large lot sizes in most wooded residential areas in the watershed. Areas identified as open residential tended to be more densely developed than the forested residential areas, so the

difference in response to rainfall of the two land uses was expected to be more significant. The fact that many of the open residential ground truth pixels were classified as mixed forest or forested residential is particularly troubling, since these land uses were generally assigned considerably different AGNPS parameter values than were open residential pixels. The confusion between open residential and forested residential is not terribly surprising because of the contribution of reflectance from roof and pavement areas. However, the confusion between open residential and mixed forest is difficult to understand.

The results of accuracy assessment of the fuzzy classification and modeling process into AGNPS parameter values in West Fork Little River are depicted in Figures 5.22 through 5.24. The mean value of 0.044 for the histogram of CN residuals divided by CN ground truth data values was taken to mean that the average error in the combined classification and modeling process for CN in West Fork Little River was, at 4.4% of the ground truth values selected for accuracy assessment, significantly better than the 17% achieved in Subwatershed K. However, the mean value of 1.00 for the histogram of n residuals divided by n ground truth data values, representing an average discrepancy between predicted (classified and modeled) values of n and ground truth values reserved for accuracy assessment of 100%, indicates that the average error in the combined classification and modeling process for n in West Fork Little River was substantially worse than the 29% result in Subwatershed K. This result bears further consideration, though. The range of classified and modeled values for n in West Fork Little River was from 0.0126 to 0.737 (Table 5.4), with a mean value of 0.222. From Figure 5.23, it can be seen that the mean residual value for n was 0.095. This represents 13% of the total

Table 5.6. Accuracy Assessment of LU/LC classification of West Fork Little River produced by Imagine Accuracy Assessment module.

CLASSIFICATION ACCURACY ASSESSMENT REPORT

Image File : c:/documents and settings/rfuller/my documents/dissertation/piedmont/imagine-working-pi/landsat etm+ working/lulc/wflr/lulc_wflr_fm_merged_sig.img
User Name : rcfuller
Date : Sun Jan 19 18:06:03 2003

ERROR MATRIX

	Reference Data			
Classified Data	Unclassified	Cover crop	Deciduous	Evergreen
Unclassified	0	0	0	0
Cover crop	0	119	0	0
Deciduous forest	0	0	146	0
Evergreen forest	0	0	0	69
Farmsteads	0	0	2	0
Mixed forest	0	1	24	8
Pasture	0	13	0	0
Residential, forested	0	0	6	0
Residential, open	0	2	0	0
Water	0	0	0	0
Urban/Institutional	0	0	0	0
Column Total	0	135	178	77

	Reference Data			
Classified Data	Farmsteads	Mixed forest	Pasture	Res, forested
Unclassified	0	0	0	0
Cover crop	0	0	14	0
Deciduous forest	0	66	1	78
Evergreen forest	0	45	0	0
Farmsteads	30	0	1	7
Mixed forest	2	55	8	9
Pasture	0	0	119	0
Residential, forested	1	3	2	57
Residential, open	8	6	51	9
Water	0	0	0	0
Urban/Institutional	3	0	0	1
Column Total	44	175	196	161

	Reference Data			
Classified Data	Res, open	Water	Urban/Inst	Row Total
Unclassified	0	0	0	0
Cover crop	1	0	0	134
Deciduous forest	1	0	0	292
Evergreen forest	3	0	0	117
Farmsteads	2	0	0	42
Mixed forest	22	0	0	129
Pasture	8	0	0	140
Residential, forested	23	0	0	92
Residential, open	74	0	0	150
Water	0	13	0	13
Urban/Institutional	5	0	18	27
Column Total	139	13	18	1136

----- End of Error Matrix -----

ACCURACY TOTALS

Class Name	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy
Unclassified	0	0	0	---	---
Cover crop	135	134	119	88.15%	88.81%
Deciduous forest	178	292	146	82.02%	50.00%
Evergreen forest	77	117	69	89.61%	58.97%
Farmsteads	44	42	30	68.18%	71.43%
Mixed forest	175	129	55	31.43%	42.64%
Pasture	196	140	119	60.71%	85.00%
Residential, forstd	161	92	57	35.40%	61.96%
Residential, open	139	150	74	53.24%	49.33%
Water	13	13	13	100.00%	100.00%
Urban/Institutional	18	27	18	100.00%	66.67%
Totals	1136	1136	700		

Overall Classification Accuracy = 61.62%

----- End of Accuracy Totals -----

KAPPA (K^) STATISTICS

Overall Kappa Statistics = 0.5590

Conditional Kappa for each Category.

Class Name	Kappa
Unclassified	0.0000
Cover crop	0.8730
Deciduous forest	0.4071
Evergreen forest	0.5599
Farmsteads	0.7028
Mixed forest	0.3219
Pasture	0.8187
Residential, forested	0.5567
Residential, open	0.4227
Water	1.0000
Urban/Institutional	0.6613

range of values and 43% of the mean value of n. This is still less than ideal but not nearly as bad as the 100% mean decimal residual value indicates. This value, necessarily, then, indicates a preponderance of the n residual values occurring with small ground truth values, which led to the large decimal residual values.

Interpretation of Figure 5.24, showing the residuals for SCC is similar to the interpretation of SCC residual values for Subwatershed K except that the results in West Fork Little River were slightly better. In the top half of Figure 5.24, the mean residual for SCC is seen to be 0.098 and the maximum is 0.480. Compared to the range of values, 0.000 to 0.590 in the accuracy assessment ground truth data and 0.005 to 0.524 in the modeled values, the mean represents less than 17% of the of the range of accuracy

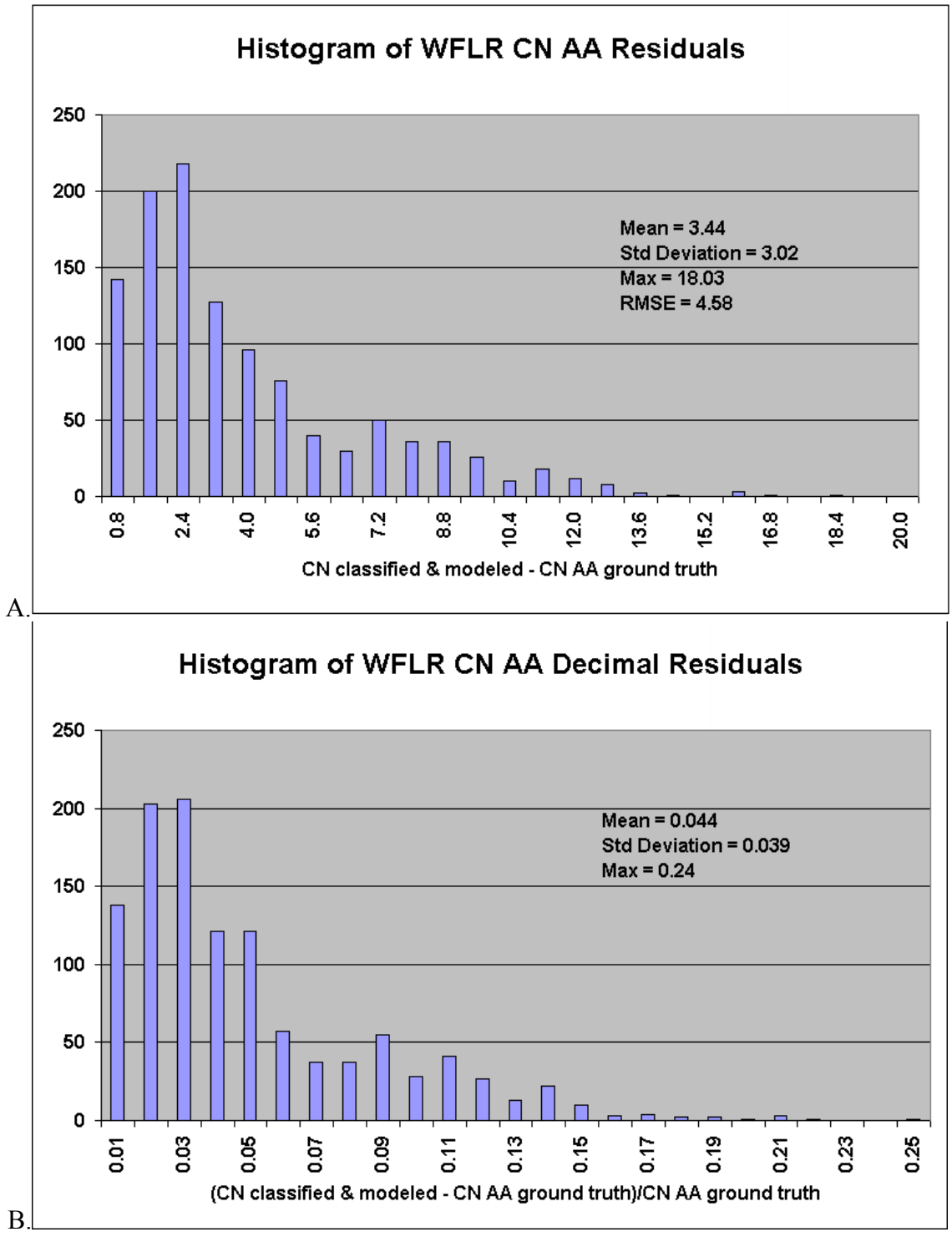
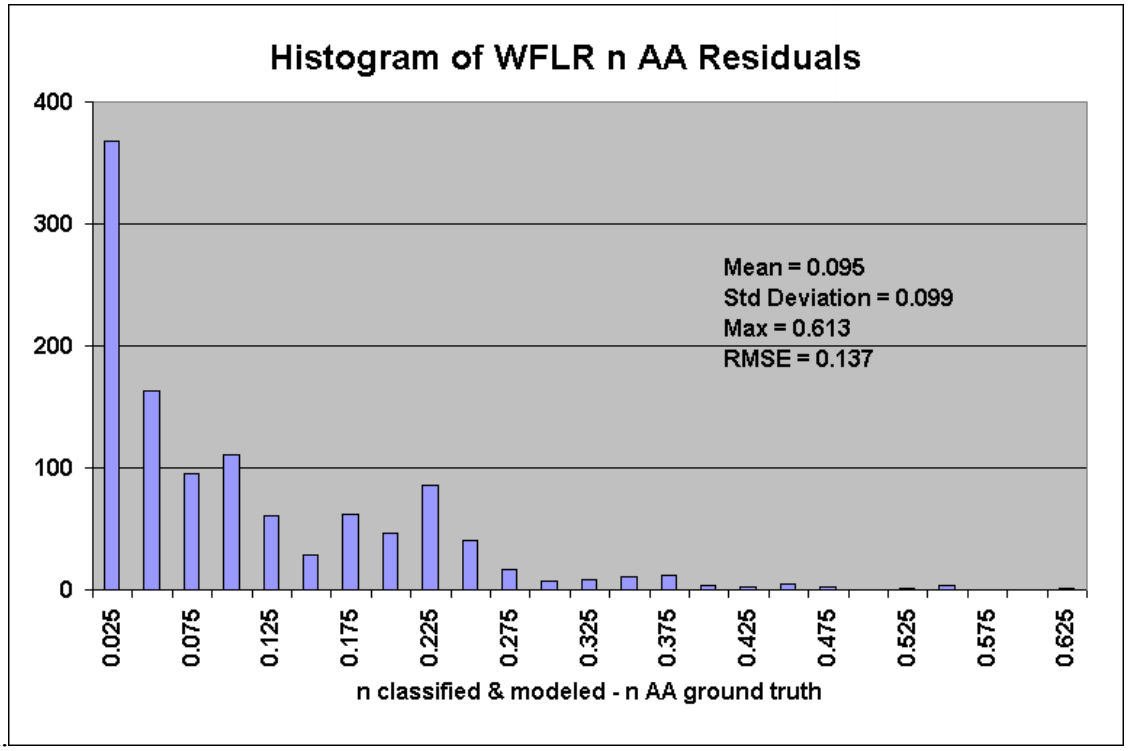


Figure 5.22. Residuals (predicted – observed) of the classification and modeling process for CN in West Fork Little River (A) and the decimal fraction of the residuals divided by observed values (B).

A.



B.

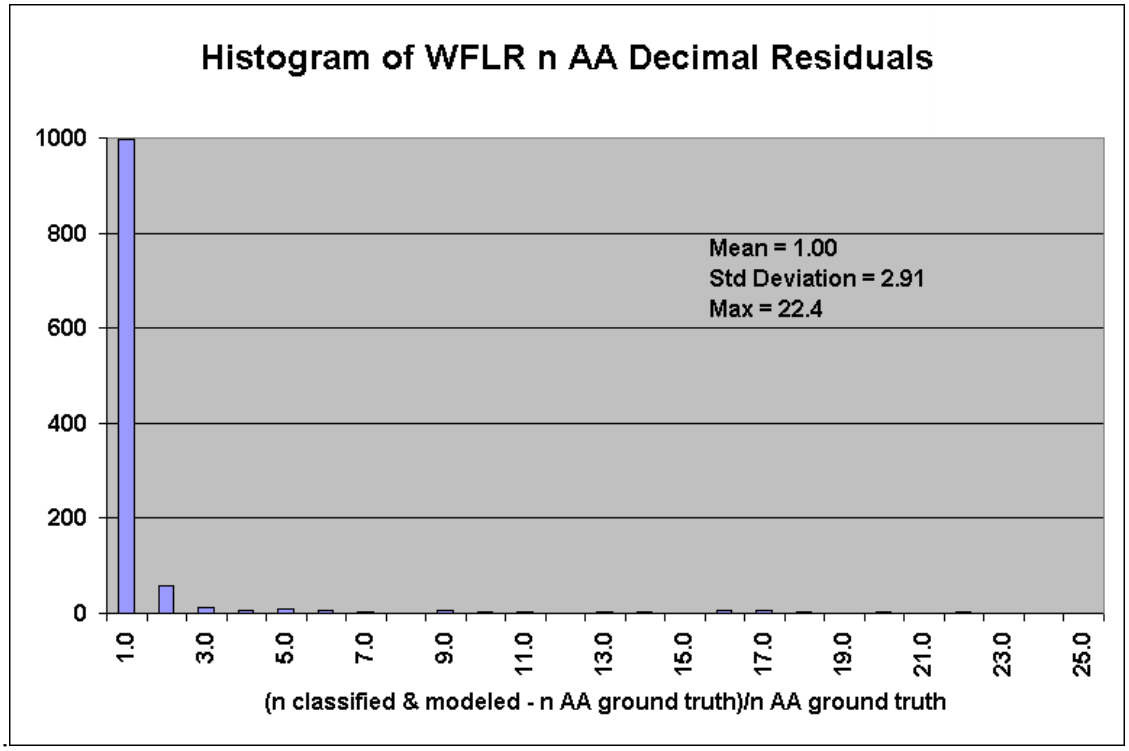
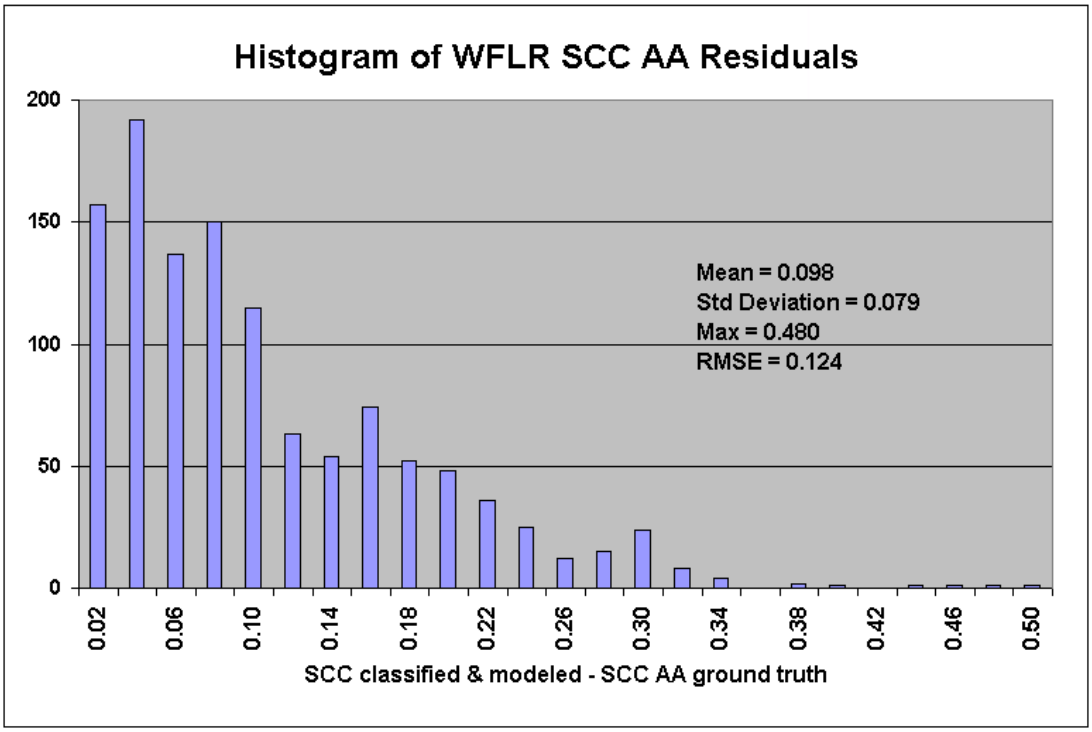


Figure 5.23. Residuals (predicted – observed) of the classification and modeling process for n in West Fork Little River (A) and the decimal fraction of the residuals divided by observed values (B).

A.



B.

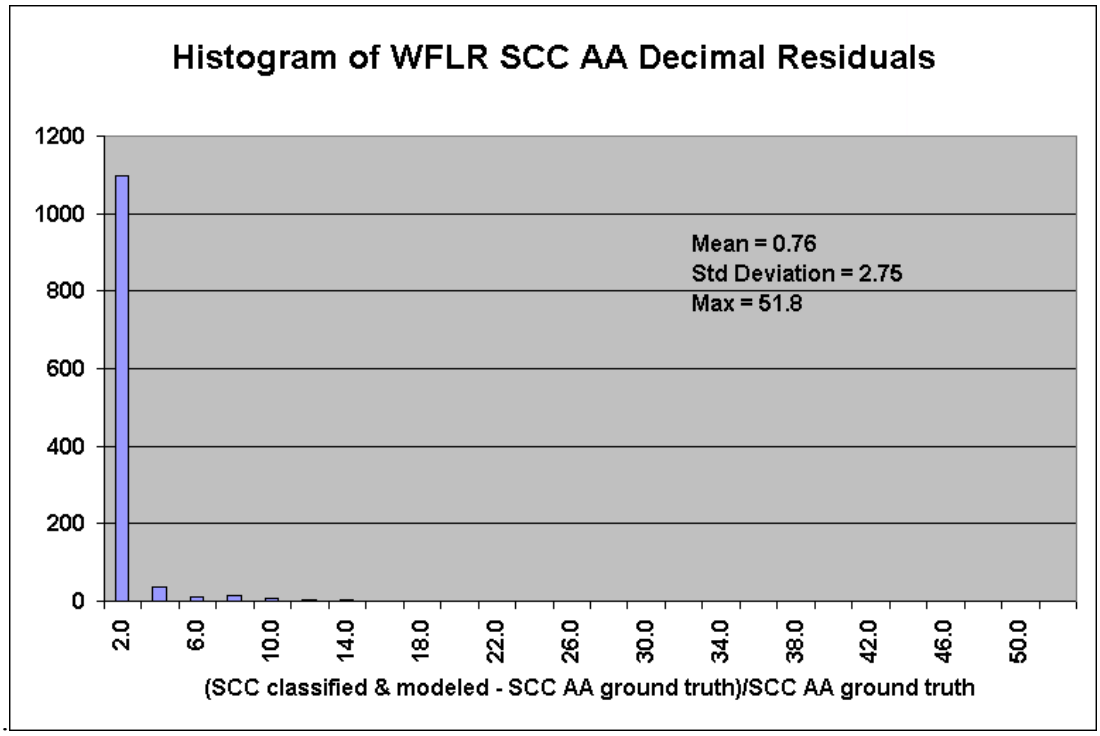


Figure 5.24. Residuals (predicted – observed) of the classification and modeling process for SCC in West Fork Little River (A) and the decimal fraction of the residuals divided by observed values (B).

assessment ground truth data and slightly under 19% of the modeled values. These values are not unacceptably high.

AGNPS modeling results

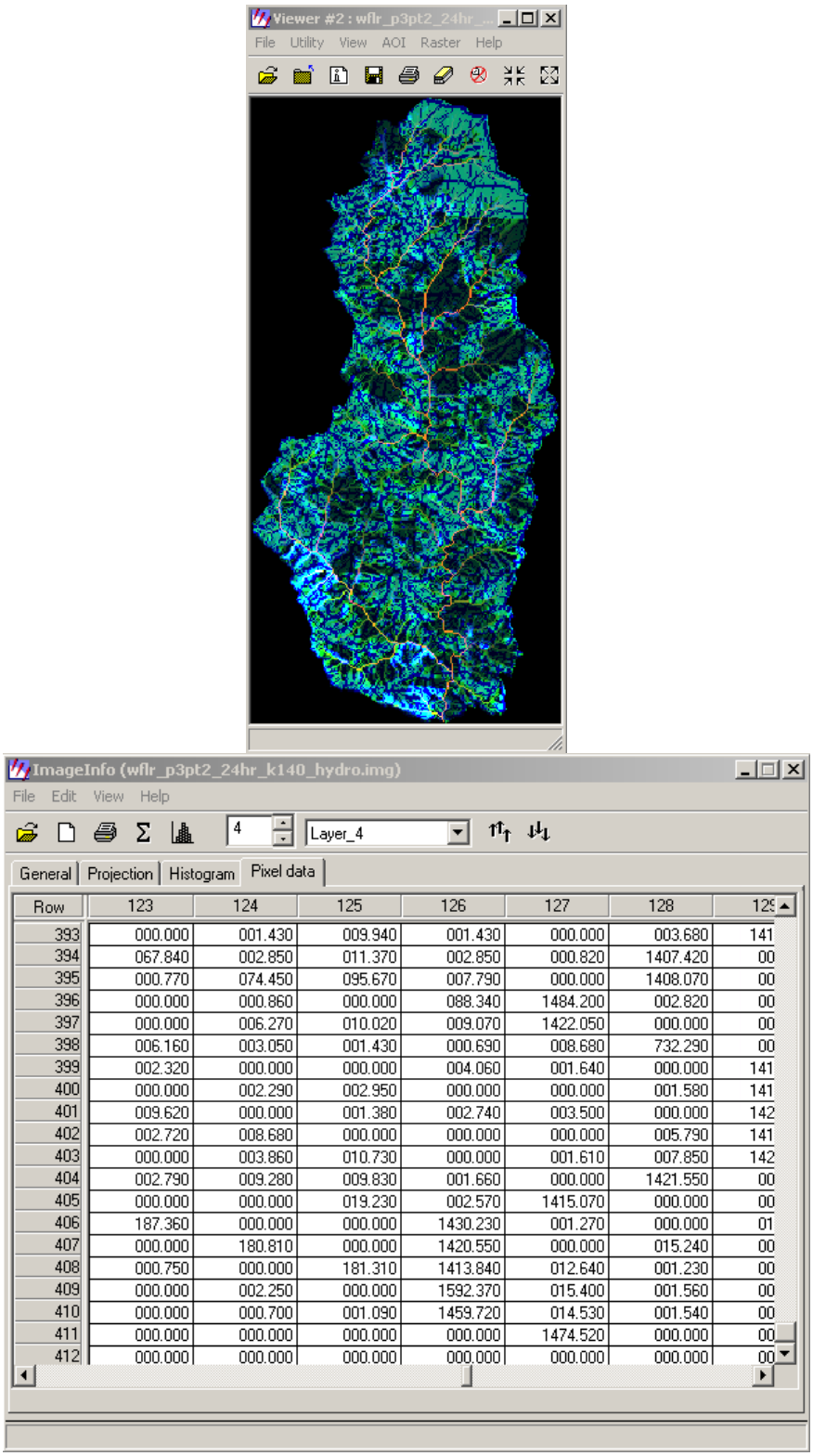
Table 5.7 presents the total runoff volume, with base flow removed, for six rainfall events in West Fork Little River. These values and associated peak flow values were calculated from recorded data in order to provide a comparison with AGNPS modeling results. Linear regression was used on these values to derive total runoff volume and peak flows for the modeled 3.2-inch rainfall for both West Fork Little River and East Fork Little River. The runoff per unit area derived for West Fork Little River was applied to East Fork Little River, then total volume was calculated based on the watershed area. Peak flows for the 3.2-inch storm were calculated for both watersheds using the SCS triangular hydrograph method (Georgia Soil and Water Conservation Commission 1992, A-2-1 – A-2-5; Northcott 2003). The peak flow for a 3.2-inch rain event in West Fork Little River, based on the regression of recorded data and SCS peaking, was calculated as 5,607 cubic feet per second (cfs), based on a time to peak of 20.6 hours. Using the same method, the peak flow for a 3.2-inch rain in East Fork Little River, based on a time to peak of 21.3 hours, was calculated as 4,454 cfs. These values are presented in Tables 5.8 and 5.9 for comparison with AGNPS-modeled values.

Figure 5.25 presents an image and sample pixel values of the West Fork Little River AGNPS results based on Data Generator input. Figure 5.26 presents similar information for AGNPS results produced with images of CN, n, and SCC extracted from the Landsat ETM+ image. The results of these two methods of AGNPS modeling of West Fork Little River are compared in Table 5.8 with peak flow for a 3.2-inch rainfall

Table 5.7. Total runoff volume, with base flow removed, for West Fork Little River, USGS Station 02332830, and associated rainfall at Clermont, NCDC Station 91998.

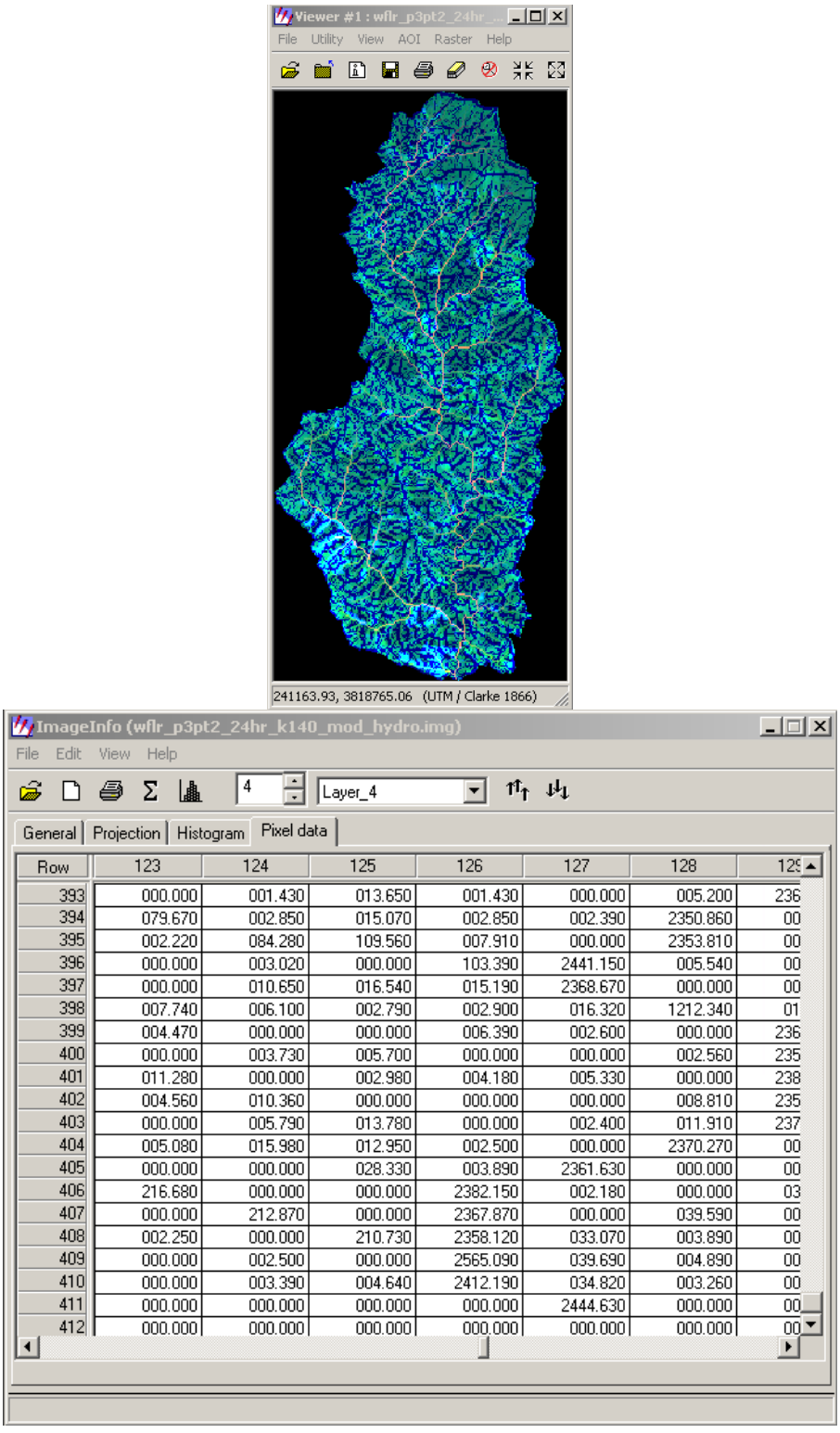
Date of start of rainfall	Rainfall, in.	Recorded runoff volume, base flow removed, $ft.^3$	Runoff volume from regression line, $ft.^3$
1/22/99	1.7	5,875,200	8,156,271
3/12/93	1.8	7,516,800	10,663,453
3/25/93	2.2	29,289,600	20,692,178
6/1/95	4.0	30,628,800	65,821,441
1/25/96	4.4	131,630,400	75,850,166
8/14/94	4.8	62,121,600	85,878,891

based on regression of recorded flow and rainfall for similar storms and with field monitoring data for the 1.44-inch storm of March 29-30, 2001. The total sediment yield figure for the monitored storm in this table was calculated by multiplying the sum of all incremental flows by the length storm in seconds, then by the average total suspended solids concentration, expressed as parts per million, which is equivalent to milligrams per liter, then performing appropriate units conversion. The resultant figure was entered in the table as total sediment yield, though it might more closely represent the sum of the clay and silt yield from the watershed. For a larger, more energetic storm, more sand-size particles would be in suspension in the flow and would be included in the measured total suspended solids. Since the overwhelming majority of the sediment yield calculated by AGNPS was clay-size, the lack of exact correspondence in the measured vs. modeled sediment yield figures is not particularly significant. Comparable information for East Fork Little River is presented in Figures 5.27 and 5.28 and Table 5.9.



Row	123	124	125	126	127	128	129
393	000.000	001.430	009.940	001.430	000.000	003.680	141
394	067.840	002.850	011.370	002.850	000.820	1407.420	00
395	000.770	074.450	095.670	007.790	000.000	1408.070	00
396	000.000	000.860	000.000	088.340	1484.200	002.820	00
397	000.000	006.270	010.020	009.070	1422.050	000.000	00
398	006.160	003.050	001.430	000.690	008.680	732.290	00
399	002.320	000.000	000.000	004.060	001.640	000.000	141
400	000.000	002.290	002.950	000.000	000.000	001.580	141
401	009.620	000.000	001.380	002.740	003.500	000.000	142
402	002.720	008.680	000.000	000.000	000.000	005.790	141
403	000.000	003.860	010.730	000.000	001.610	007.850	142
404	002.790	009.280	009.830	001.660	000.000	1421.550	00
405	000.000	000.000	019.230	002.570	1415.070	000.000	00
406	187.360	000.000	000.000	1430.230	001.270	000.000	01
407	000.000	180.810	000.000	1420.550	000.000	015.240	00
408	000.750	000.000	181.310	1413.840	012.640	001.230	00
409	000.000	002.250	000.000	1592.370	015.400	001.560	00
410	000.000	000.700	001.090	1459.720	014.530	001.540	00
411	000.000	000.000	000.000	000.000	1474.520	000.000	00
412	000.000	000.000	000.000	000.000	000.000	000.000	00

Figure 5.25. AGNPS hydro image and sample pixel values for West Fork Little River using input data generated with Data Generator.

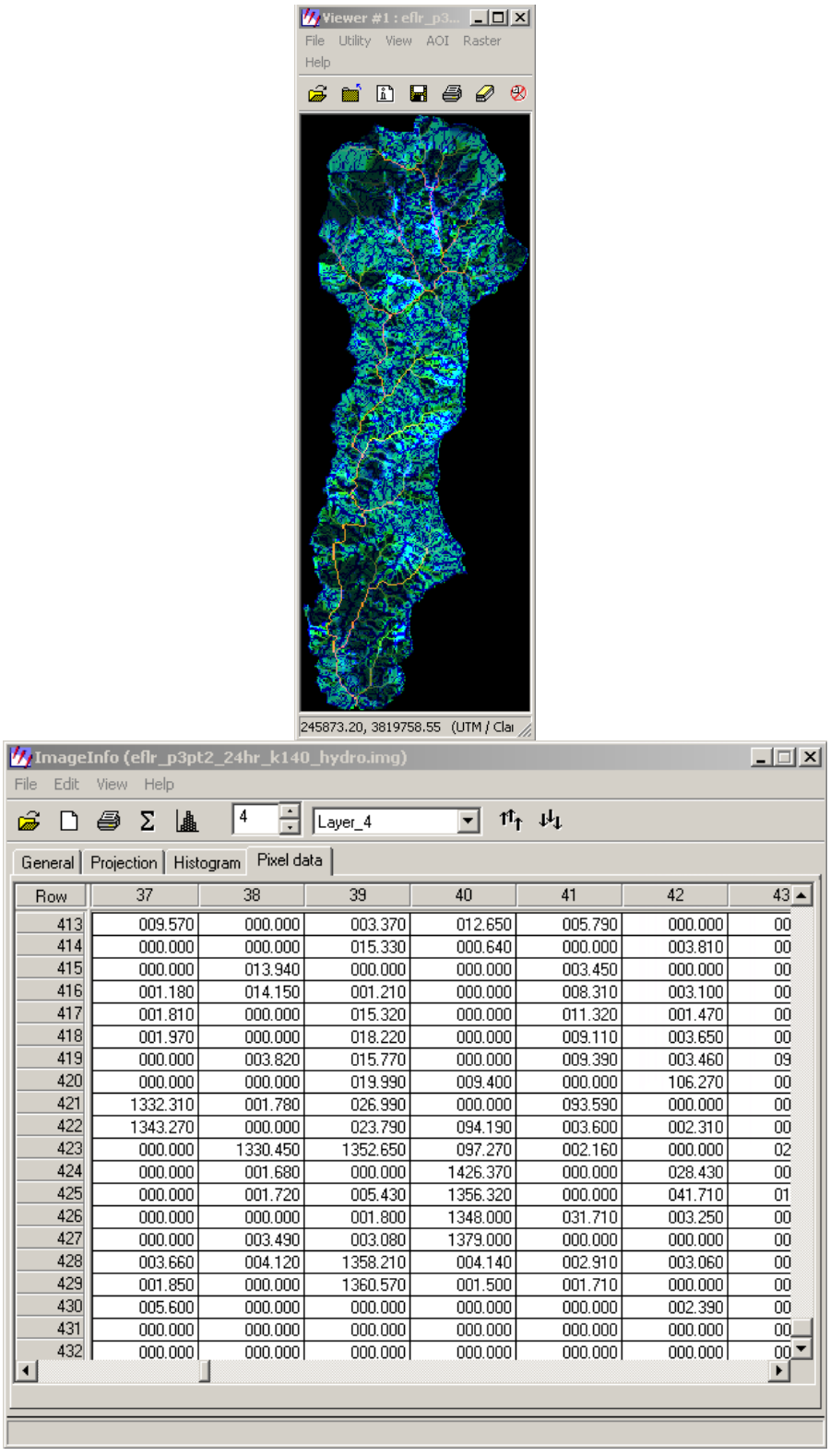


Row	123	124	125	126	127	128	129
393	000.000	001.430	013.650	001.430	000.000	005.200	236
394	079.670	002.850	015.070	002.850	002.390	2350.860	00
395	002.220	084.280	109.560	007.910	000.000	2353.810	00
396	000.000	003.020	000.000	103.390	2441.150	005.540	00
397	000.000	010.650	016.540	015.190	2368.670	000.000	00
398	007.740	006.100	002.790	002.900	016.320	1212.340	01
399	004.470	000.000	000.000	006.390	002.600	000.000	236
400	000.000	003.730	005.700	000.000	000.000	002.560	235
401	011.280	000.000	002.980	004.180	005.330	000.000	238
402	004.560	010.360	000.000	000.000	000.000	008.810	235
403	000.000	005.790	013.780	000.000	002.400	011.910	237
404	005.080	015.980	012.950	002.500	000.000	2370.270	00
405	000.000	000.000	028.330	003.890	2361.630	000.000	00
406	216.680	000.000	000.000	2382.150	002.180	000.000	03
407	000.000	212.870	000.000	2367.870	000.000	039.590	00
408	002.250	000.000	210.730	2358.120	033.070	003.890	00
409	000.000	002.500	000.000	2565.090	039.690	004.890	00
410	000.000	003.390	004.640	2412.190	034.820	003.260	00
411	000.000	000.000	000.000	000.000	2444.630	000.000	00
412	000.000	000.000	000.000	000.000	000.000	000.000	00

Figure 5.26. AGNPS hydro image and sample pixel values for West Fork Little River substituting images for CN, n, and SCC created in the classification of a Landsat ETM+ image and post-classification modeling of those images.

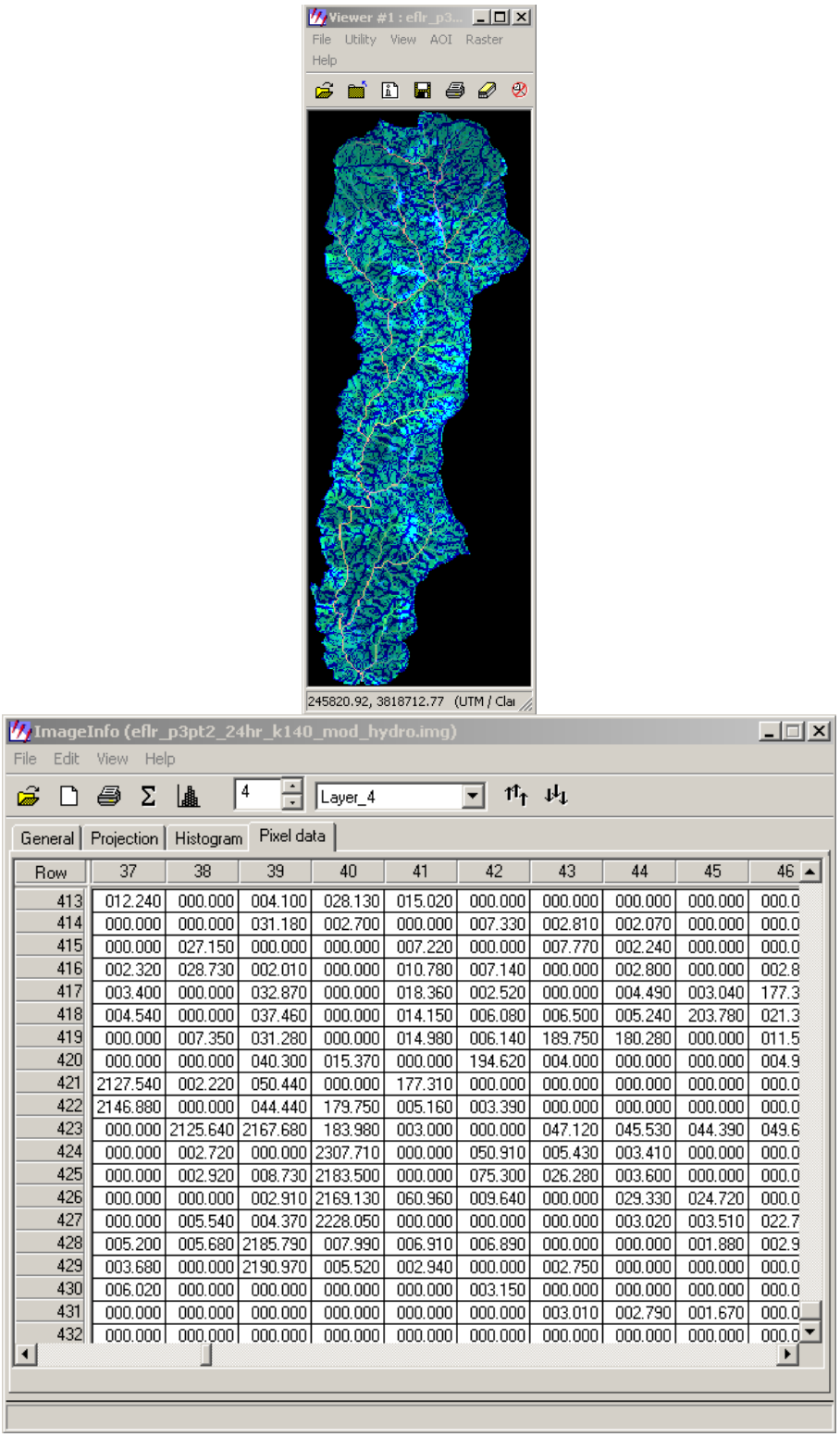
Table 5.8. Selected results of AGNPS modeling of West Fork Little River watershed using two methods of generating input data and field monitoring data.

	All input created with Data Generator	Parameters CN, n, and SCC generated from Landsat ETM+ image	SCS peaking based on regression from recorded WFLR data	Field/lab monitored data
Rainfall, inches	3.2	3.2	3.2	1.44
Peak flow, cfs	1,475	2,445	5,607	107
Clay yield, tons	277	473	no data	no data
Silt yield, tons	32	48	no data	no data
Sand yield, tons	9	12	no data	no data
Total sediment yield, tons	372	613	no data	11.5
Chemical oxygen demand (COD), mg/l (ppm)	0	0	no data	11.2



Row	37	38	39	40	41	42	43
413	009.570	000.000	003.370	012.650	005.790	000.000	00
414	000.000	000.000	015.330	000.640	000.000	003.810	00
415	000.000	013.940	000.000	000.000	003.450	000.000	00
416	001.180	014.150	001.210	000.000	008.310	003.100	00
417	001.810	000.000	015.320	000.000	011.320	001.470	00
418	001.970	000.000	018.220	000.000	009.110	003.650	00
419	000.000	003.820	015.770	000.000	009.390	003.460	09
420	000.000	000.000	019.990	009.400	000.000	106.270	00
421	1332.310	001.780	026.990	000.000	093.590	000.000	00
422	1343.270	000.000	023.790	094.190	003.600	002.310	00
423	000.000	1330.450	1352.650	097.270	002.160	000.000	02
424	000.000	001.680	000.000	1426.370	000.000	028.430	00
425	000.000	001.720	005.430	1356.320	000.000	041.710	01
426	000.000	000.000	001.800	1348.000	031.710	003.250	00
427	000.000	003.490	003.080	1379.000	000.000	000.000	00
428	003.660	004.120	1358.210	004.140	002.910	003.060	00
429	001.850	000.000	1360.570	001.500	001.710	000.000	00
430	005.600	000.000	000.000	000.000	000.000	002.390	00
431	000.000	000.000	000.000	000.000	000.000	000.000	00
432	000.000	000.000	000.000	000.000	000.000	000.000	00

Figure 5.27. AGNPS hydro image and sample pixel values for East Fork Little River using input data generated with Data Generator.



Row	37	38	39	40	41	42	43	44	45	46
413	012.240	000.000	004.100	028.130	015.020	000.000	000.000	000.000	000.000	000.0
414	000.000	000.000	031.180	002.700	000.000	007.330	002.810	002.070	000.000	000.0
415	000.000	027.150	000.000	000.000	007.220	000.000	007.770	002.240	000.000	000.0
416	002.320	028.730	002.010	000.000	010.780	007.140	000.000	002.800	000.000	002.8
417	003.400	000.000	032.870	000.000	018.360	002.520	000.000	004.490	003.040	177.3
418	004.540	000.000	037.460	000.000	014.150	006.080	006.500	005.240	203.780	021.3
419	000.000	007.350	031.280	000.000	014.980	006.140	189.750	180.280	000.000	011.5
420	000.000	000.000	040.300	015.370	000.000	194.620	004.000	000.000	000.000	004.9
421	2127.540	002.220	050.440	000.000	177.310	000.000	000.000	000.000	000.000	000.0
422	2146.880	000.000	044.440	179.750	005.160	003.390	000.000	000.000	000.000	000.0
423	000.000	2125.640	2167.680	183.980	003.000	000.000	047.120	045.530	044.390	049.6
424	000.000	002.720	000.000	2307.710	000.000	050.910	005.430	003.410	000.000	000.0
425	000.000	002.920	008.730	2183.500	000.000	075.300	026.280	003.600	000.000	000.0
426	000.000	000.000	002.910	2169.130	060.960	009.640	000.000	029.330	024.720	000.0
427	000.000	005.540	004.370	2228.050	000.000	000.000	000.000	003.020	003.510	022.7
428	005.200	005.680	2185.790	007.990	006.910	006.890	000.000	000.000	001.880	002.9
429	003.680	000.000	2190.970	005.520	002.940	000.000	002.750	000.000	000.000	000.0
430	006.020	000.000	000.000	000.000	000.000	003.150	000.000	000.000	000.000	000.0
431	000.000	000.000	000.000	000.000	000.000	000.000	003.010	002.790	001.670	000.0
432	000.000	000.000	000.000	000.000	000.000	000.000	000.000	000.000	000.000	000.0

Figure 5.28. AGNPS hydro image and sample pixel values for East Fork Little River substituting images for CN, n, and SCC created in the classification of a Landsat ETM+ image and post-classification modeling of those images.

Table 5.9. Selected results of AGNPS modeling of East Fork Little River watershed using two methods of generating input data and field monitoring data.

	All input created with Data Generator	Parameters CN, n, and SCC generated from Landsat ETM+ image	SCS peaking based on regression from recorded WFLR data	Field/lab monitored data
Rainfall, inches	3.2	3.2	3.2	1.44
Peak flow, cfs	1,361	2,191	4,454	61
Clay yield, tons	4,296	8,216	no data	no data
Silt yield, tons	38	83	no data	no data
Sand yield, tons	7	10	no data	no data
Total sediment yield, tons	4,385	8,371	no data	7.7
Chemical oxygen demand (COD), mg/l (ppm)	0	0	no data	7.5

CHAPTER 6

SUMMARY AND CONCLUSIONS

A process has been developed to use the reflectances contained within a digital satellite image as means of extrapolating discrete field data to better approximate the continuum that most environmental variables exhibit in reality. The process used a three-level fuzzy classification and Mahalanobis distances associated with each pixel to create weighted averages of the environmental variables under consideration. Weighting was based on the reciprocals of the Mahalanobis distances associated with values in each layer of the fuzzy-classified image.

The example used in the development of the process was the estimation of selected parameter values for the AGNPS watershed model. A watershed in the Georgia Coastal Plain was used as the developmental test-bed for the entire process. An early goal of comparing the results of the modeling of this watershed to existing field data was not realized because the field data could not be acquired for this watershed. However, the results of modeling the watershed with AGNPS using a classified land use/land cover image as one source of input data was compared with the results produced using the fuzzy classification and post-classification modeling methodology. Two watersheds in the Georgia Piedmont were selected for modeling and the results of modeling these watersheds using the two distinct methods of generating input parameters were compared. In addition, these results were compared with theoretical peak runoff, based on the SCS triangular hydrograph method and with data gathered in the field and

laboratory reflecting the responses of the watersheds to a rain event. In the Coastal Plain watershed, the results of modeling the watershed with AGNPS using the fuzzy classification and post-classification modeling methodology were consistently lower than the results obtained using a classified land use/land cover image as one source of input data. In the two Piedmont watersheds, the relationship was reversed. The peak flows predicted by the AGNPS modeling using fuzzy classification and post-classification modeling in the Piedmont watersheds were about half those based on the SCS triangular hydrograph method. However, this was a significantly better result than that achieved using only the AGNPS parameter values generated with AGNPS Data Generator. The discrepancy between AGNPS modeling results and SCS triangular hydrograph peak runoff predictions is not surprising, considering the steep terrain in both West Fork Little River and East Fork Little River. AGNPS was developed for relatively flat, agricultural watersheds (Young et al. 1989), so it should not be expected to perform well in steep, forested watersheds.

Few meaningful conclusions can be drawn from comparing the results of monitoring the responses of the two Piedmont watersheds to the modeling results because software limitations precluded modeling a storm event of the same magnitude as was monitored. The fact that the modeled watershed responses to the larger modeled storm event were significantly larger than the measured responses of the watersheds to the smaller monitored event is certainly correct but the appropriateness of the magnitudes of the differences is unknown. Based on these results, it is apparent that the first hypothesis of this study, "Classifying satellite imagery directly into water quality model parameters, eliminating the use of land use/land cover classes and associated lookup tables, will

improve the accuracy of the resultant water quality model by eliminating the forced fit of water quality model parameters into classes defined by land use/land cover classes," was not proven. On the other hand, these first trials of the methodology did not confirm the null hypothesis, "Classifying satellite imagery directly into water quality model parameters, eliminating the use of land use/land cover classes and associated lookup tables, will not improve the accuracy of the resultant water quality model by eliminating the forced fit of water quality model parameters into classes defined by land use/land cover classes." Observation of the pixel values resulting from the different methods employed to extract AGNPS parameter values from satellite imagery did support the second hypothesis, "The creation of weighted averages of fuzzy-classified layers of environmental variables will produce a better approximation of the continuity of those variables in the environment than can be achieved by more traditional methods."

The process that has been developed would be better suited for applications in which environmental variables were actually measured in the field instead of being estimated on the basis of published values, as was done in this study. In this study, environmental variables estimated in the field from published values inevitably turned out to be categorical data, even though the categories were numerical. One such possible application is in extrapolating field observations of plant communities. It is anticipated that the techniques developed in this study would be well-suited to the creation of vegetation maps depicting the expected densities of various plant species that might be found, based on a fuzzy classification followed by the post-classification techniques developed in this study. Soils classification is another possible application.

With these types of applications, though, it is anticipated that difficulties might arise due to the nature of the way the field data would likely be gathered and the mathematical requirements for classifying multi-layered imagery. A multi-layered satellite image contains a wealth of information about the reflectance of the subject area in a variety of wavelengths of electromagnetic radiation. The eight bands of the Landsat ETM+, for instance, contain information that is very useful in discriminating between a wide variety of vegetation types, rocks, soil types, and other environmental variables. In order to use the depth of information contained in the multiple layers of this imagery, however, it is necessary to associate with each field value a number of image pixels equal to the number of bands being used plus one (ERDAS 2001). This is not normally a problem if the field data are nominal, i.e. representing a finite number of categories. However, if the field data are interval or ratio type, then problems arise. For example, if three bands of the imagery were being used to develop a map depicting the cation exchange capacity, then a minimum of four pixels in the image would need to have the same numerical value associated with them in order to classify the image using a maximum likelihood fuzzy classification strategy. While this may limit the applicability of this strategy somewhat, it does not preclude its use in developing continuous maps of a variety of environmental variables. It simply limits the procedure to the collection of categorical rather than continuous field data for the variable under consideration. For instance, if soil moisture were the environmental variable under consideration and only one band of an image were being used in the analysis, it would still be necessary to have a minimum of two pixels in the imagery associated with each discrete value of soil moisture measured in the field. Unless a very large number of field measurements were

made, it is likely that quite a few values would be measured only once within a given study area. It would, therefore, be necessary to group similar field-measured values together in order to get a minimum of two pixels (in this example) associated with each discrete soil moisture value.

In practical application, field measurements of environmental variables are not truly continuous. Continuity is limited by the ability to measure. In the case of digital measurement instruments, continuity is limited by the number of digits displayed by the measurement device, whatever it may be. Continuity should, in fact, be limited not by the instrument display but by the precision (i.e. the repeatability) of the instrument as well as by its accuracy (i.e. its closeness to the correct value). This may well be less than the limits imposed by the number of digits displayed. The limiting case of the need for discrete values to be reflected in the field data, therefore, would be where all possible values of the environmental variable under consideration that could precisely and accurately be measured and displayed were included in the training set data a minimum number of times equal to one plus the number of imagery bands used in the classification. Therefore, the limitation of being able to gather and use only categorical environmental data is, strictly speaking, absolute. However, the ideal of gathering continuous field data can be approached by using a clustering sampling scheme, with the clusters designed to cover at least one more pixel in the imagery to be classified than the number of bands to be used.

REFERENCES

Anderson, J. R., E. E. Hardy, J. T. Roach, and R. E. Witmer. 1976. *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*, U. S. Geological Service Professional Paper 964, U.S.G.S Washington, D.C.

Baker, K. D., F. D. Theurer, and J. Witte.1995. "AGNPS version 5.00 verification: science". *Soils.mrsars.usda.gov*. File Name: vrfscien.doc. Revision date: 30 May 1995.

Batchelor, P. 1994. "Models as metaphors: the role of modeling in pollution prevention". *Waste Management*. 14(3): pp. 243-251.

Burrough, P. A. 1986. *Principles of Geographic Information Systems for Land Resources Assessment*. Oxford: Clarendon Press.

Chow, V. T., D. R. Maidment, and L. W. Mays, eds. 1988. *Applied Hydrology*. New York: McGraw-Hill.

CIESIN (Consortium for Intern Science Information Network). 1996. "AGNPS: Agricultural Non-point Source Pollution Model". *http://dino.wiz.uni-kassel.de/model_db/mdb/agnps.html*. Maintained by Joachim Benz, University of Kassel. Updated 5-30-96.

Ciesiolka, C. A., K. J. Coughlan, C. W. Rose, M. C. Escalante, G. Mohd. Hashim, E. P. Paningbatan Jr., and S. Sombatpanit. 1995. "Methodology for a multi-country study of soil erosion management". *Soil Technology*. 8: pp. 179-192.

Congalton, R. 1991. "A review of assessing the accuracy of classifications of remotely sensed data". *Remote Sensing of Environment*, 37: pp. 35-46.

ERDAS, Inc. 2001. "Classification On-Line Manual, Signature Editor" *ERDAS Imagine 8.5 On-Line Help Manual*. Atlanta: ERDAS, Inc.

ERDAS, Inc. 1997. *ERDAS® Field Guide™, 4th Ed.* Atlanta: ERDAS, Inc., 656 p.

Finn, M. P., E. L. Usery, D. J. Scheidt, T. Beard, S. Ruhl, and M. Bearden. 2002. "AGNPS watershed modeling with GIS databases". *Proceedings Second Federal Interagency Hydrologic Modeling Conference*. Las Vegas, Nevada.

Georgia GIS Data Clearinghouse. 2002. Data Preview, 7.5-minute Digital Raster Graphics(tiff), Cuffietown. URL (varies for other DRGs): http://gis1.state.ga.us/preview.asp?dataId=7060&data_type=1, accessed June 4, 2002.

Georgia GIS Data Clearinghouse. 2002b. SSURGO. URL (varies for other soils files): http://gis1.state.ga.us/.

Georgia GIS Data Clearinghouse. 2002c. Turner County ned287. URL (varies for other dem files): http://gis1.state.ga.us/.

Georgia Soil and Water Conservation Commission. 1992. *Manual for Erosion and Sediment Control in Georgia, 3rd Ed.* Athens: Georgia Soil and Water Conservation Commission.

Merritt, F. S. 1983. *Standard Handbook for Civil Engineers, 3rd Ed.* New York: McGraw-Hill Book Company.

Mitchell, J. K., B. A. Engel, R. Srinivasan, and S. S. Y. Wang. 1993. "Validation of AGNPS for small watersheds using an integrated AGNPS/GIS system". *Water Resources Bulletin*. 29(5): pp.833-842.

Morse, G., A. Eatherall, and A. Jenkins. 1994. "Managing agricultural pollution using a linked geographical information system and non-point source pollution model". *Journal of the Institute of Water and Environment Management*. 1994(8): pp. 277-286.

NCDC. 2002. "Gainesville, GA, Record of Climatological Observations". http://www.ncdc.noaa.gov/dailyform/dlyFORMv2, accessed 07:57:05 EST Nov. 27, 2002.

Northcott, W. 2003. “Estimating peak discharge” in *Biosystems Engineering 481 Agricultural and Small Watershed Hydrology*. http://www.egr.msu.edu/~northco2/BE481/Peakdischarge.htm, accessed 16:34:21 EST Apr. 24, 2003.

Oliver, G. and R. Solomon. 1990. “The use of surface runoff models for water quality decisions - user’s perspective”. In D. G. Decoursey (ed.) *Proceedings of the International Symposium on Water Quality Modeling of Agricultural Nonpoint Sources, Logan, UT*. 19-23 June 1988. ARS-81. USDA-ARS, Washington, DC: pp. 197-204.

Olivieri, L. J., G. M. Schaal, T. J. Logan, W. J. Elliot, and B. Motsch. 1991. “Generating AGNPS input using remote sensing and GIS”. American Society of Agricultural Engineers Paper no. 91-2622. St. Joseph, MI.

Olivieri, L. J., G. M. Schaal, B. Motsch, and W. J. Elliot. 1995. “Linking remote sensing and GIS to soil conservation: generating AGNPS input using remote sensing and GIS” *http://www.dnr.ohio.gov/odnr/relm/remsen/agnps/cleanh2o.html*. Posted 7-6-95.

Omernick, J. M. 1976. *The influence of land us on stream nutrient levels.* Environmental Reseach Laboratory, Corvallis, OR. EPA Report No. EPA-600/3-76-014.

Osmond, D. L., R. W. Gannon, J. A. Gale, D. E. Line, C. B. Knott, K. A. Phillips, M. H. Turner, M. A. Foster, D. E. Lehning, S. W. Coffey, and J. Spooner. 1997. “WATERSHEDSS: a decision support system for watershed-scale nonpoint source water quality problems”. *Journal of the American Water Resources Association*. 33(2): pp. 327-341.

Panuska, J. C. and I. D. Moore. 1991. *Water Quality Modeling: Terrain Analysis and the Agricultural Non-Point Source Pollution (AGNPS) Model*. TR-132. Water Resources Research Center, Minnesota University, St. Paul. NTIS Springfield, VA.

Parker, S. P. ed. 1994. *McGraw-Hill Dictionary of Scientific and Technical Terms*, 5th ed. New York: McGraw-Hill.

Reyes, M. R. and G. A. Gayle. 1995. "Automatic calculation of rainfall erosivity". *Transactions of the American Society of Agricultural Engineers*. 38(2): pp. 551-553.

Ritter, D. F., R. C. Kochel, and J. R. Miller. 1995. *Process geomorphology*, 3rd ed. Dubuque. IA: Wm. C. Brown Publishers.

Robinson, K. J. and R. M. Ragan. 1993. "Geographic information system based nonpoint pollution modeling". *Water Resources Bulletin*. 29(6): pp. 1003-1008.

Rose, C. W., W. T. Dickinson, S. E. Jorgensen, and H. Ghadiri. 1990. "Agricultural nonpoint source runoff and sediment yield water quality (NPSSWQ) models: modeler's perspective". In D. G. Decoursey (ed.) *Proceedings of the International Symposium on Water Quality Modeling of Agricultural Nonpoint Sources, Logan, UT*. 19-23 June 1988. ARS-81. USDA-ARS, Washington, DC: pp. 145-170.

Srinivasan, R. and J. G. Arnold. 1994. "Integration of a basin-scale water quality model with GIS". *Water Resources Bulletin*. 30(3): pp. 453-462.

Srinivasan, R. and B. A. Engel. 1994. "A spatial decision support system for assessing agricultural nonpoint source pollution". *Water Resources Bulletin*. 30(3): pp. 441-452.

Tim, U. S. 1996. "Emerging technologies for hydrologic and water quality modeling research". *Transactions of the American Society of Agricultural Engineers*. 39(2): pp 465-476.

Tim, U. S. and R. Jolly. 1994. "Evaluating agricultural nonpoint-source pollution using integrated geographic information systems and hydrologic/water quality model". *Journal of Environmental Quality*. 23(1): pp. 25-35.

USDA. 1972. *Soil Survey of Dawson, Lumpkin, and White Counties, Georgia.* U.S. Government Printing Office, Washington, D.C.

USDA. 1977. *Soil Survey of Barrow, Hall, and Jackson Counties, Georgia.* U.S. Government Printing Office, Washington, D.C.

USDA. 1997. "About AGNPS - from USDA-ARS". *http://www.cee.odu.edu/cee/model/agnps_desc.html*. Updated 2-5-97.

USDA. 1998. “AGNPS Version 5.0 Input File (*.DAT) Format Guide”. *http://grass.itc.it/gdp/erosion/agnps/usda-ars/500in.txt*. Accessed March 31, 2003.

USEPA. 2003. “BASINS – Better Assessment Science Integrating point and Nonpoint Sources”. *http://www.epa.gov/OST/BASINS/*.

USEPA. 1990. *National Water Quality Inventory: 1988 Report to Congress*. U. S. Environmental Protection Agency, EPA 440-4-90-003, Washington, D. C.

U.S. Geological Survey, U.S. Department of the Interior. 1996. *Georgia, SC, NC 34083, Digital Raster Graphic Data, 1-Degree Cell*. Edition 1. CD-ROM. Mastered 8/96.

USGS (unidentified employee). 2002. Personal communication by telephone, June 4.

Ventura, S. J. and K. Kim. 1993. “Modeling urban nonpoint source pollution with a geographic information system”. *Water Resources Bulletin*. 29(2): pp. 189-198.

Vellidis, G., C. K. Kvien, and D. Thomas. 1995. *A Landscape Approach to Protecting Water Quality in the Southeastern Coastal Plain*. Proposal submitted to the Cooperative State Research, Education, and Extension Service. 80 pp.

Wang, X. and Z. Yin. 1997. Using GIS to assess the relationship between land use and water quality at a watershed level. *Environment International*. 23(1): pp. 103-114.

Wischmeier, W. H. and D. D. Smith. 1978. *Predicting Rainfall Erosion Losses: a Guide to Conservation Planning*. U.S. Department of Agriculture, Agricultural Handbook Number 537.

Wilson, J. P. and J. C. Gallant. 1996. EROS: a grid-based program for estimating spatially-distributed erosion indices. *Computers and Geosciences*. 22(7): pp. 707-712.

Witte, J., F. D. Theurer, and K. D. Baker. 1995. *AGNPS version 5.00 verification report: software*. Soils.mrsars.usda.gov. File name: vrfsftwr.doc. Revision date: 30 May 1995.

Young, R. A., C. A. Onstad, D. D. Bosch, and W. P. Anderson. 1994. *Agricultural Non-Point Source Pollution Model, Version 4.03 AGNPS USER’S GUIDE. ftp://ftp:mrsars.usda.gov/pub/ars/agnps.* Note: USDA states, “The 4.03 User’s Manual is valid for version 5.00.”

Young, R. A., C. A. Onstad, D. D. Bosch, and W. P. Anderson. 1989. “AGNPS: a nonpoint source pollution model for evaluating agricultural watersheds”. *Journal of Soil and Water Conservation*. 44(2): pp. 168-173.

APPENDIX A

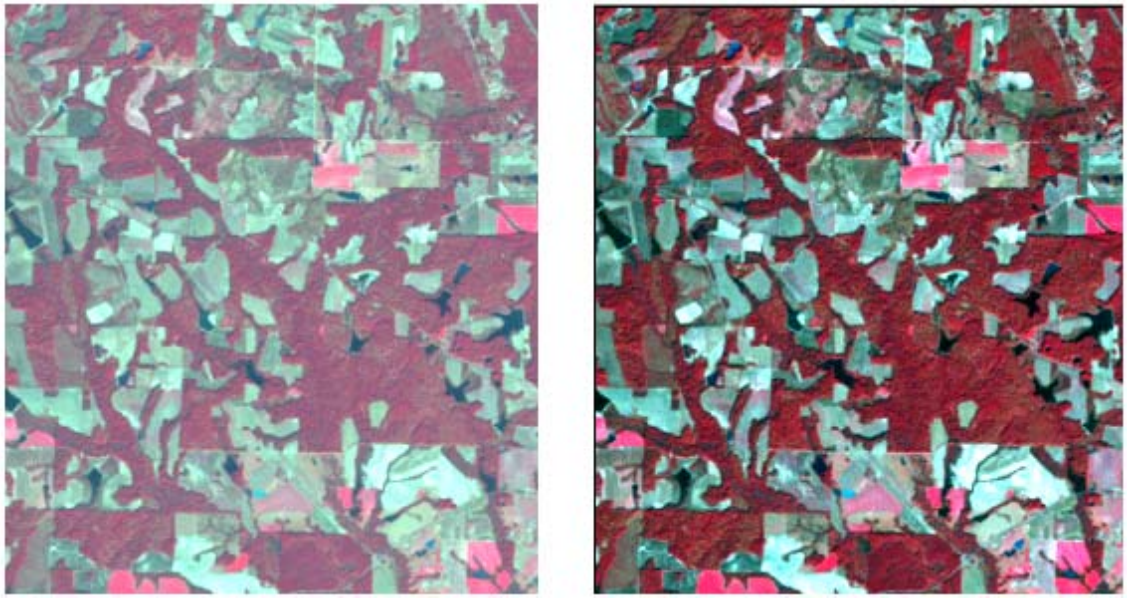
IMAGINE 8.5 PIXEL SHIFT WITH SUBSET OPERATION

Subsetting a raster image in Erdas Imagine is a useful technique for minimizing computer random access memory (RAM) requirements and for speeding image processing. However, the subsetting process can lead to subtle geometric changes that can have an adverse impact on subsequent operations on the subset image.

In the development of a model of Subwatershed K of the Little River watershed, a Landsat 7 (LS7), Enhanced Thematic Mapper Plus (ETM+) scene was used as the source of remotely sensed imagery for the subwatershed. Scene L71018038_03819991224, which was acquired December 24, 1999 and contained the entire area of interest, was obtained from the U.S. Geological Survey (USGS). Bands 1, 2, 3, 4, 5, and 7 were then combined, using the Imagine Interpreter – Utilities – Layer Stacker command, to form a single, multi-layer file. In order to minimize the amount of computer RAM needed during the many image processing steps that would be required and to speed up computer processing, a subset of this multi-layer image was selected and saved as a separate file. This operation was performed with the Imagine Interpreter – Utilities – Subset command, though it can also be done with Imagine's DataPrep – Subset Image option. The area to be included in the subset image was selected on-screen by specifying "From Inquire Box" in the Subset dialog box, then dragging a box around the desired area of the larger

source image. The subset image contained all of the layers contained in the source image, i.e., bands 1, 2, 3, 4, 5, and 7 were represented as layers in the subset image.

One of the first things that was noticed after the subset image was created and displayed with an Imagine Viewer was that with the same bands displayed and at the same scale, Subwatershed K looked distinctly different in the subset image than it did in the original, source image. The patterns discernable in the two images were identical but the colors were significantly different in the two images. Figure 1 is a side by side comparison of portions of the source image and the subset image, illustrating the displayed color differences between the two. The images displayed in Figure 1 are print versions of Windows® bitmap (.bmp) images captured from Imagine Viewer screen displays using the Parsons Screen Shot® screen capture software. The colors displayed in Figure 1 do not precisely match those displayed on the author's computer monitor because of the differences in the processes of displaying colors on the printed page and those on a computer monitor but they are indicative of the nature of the color display problem.



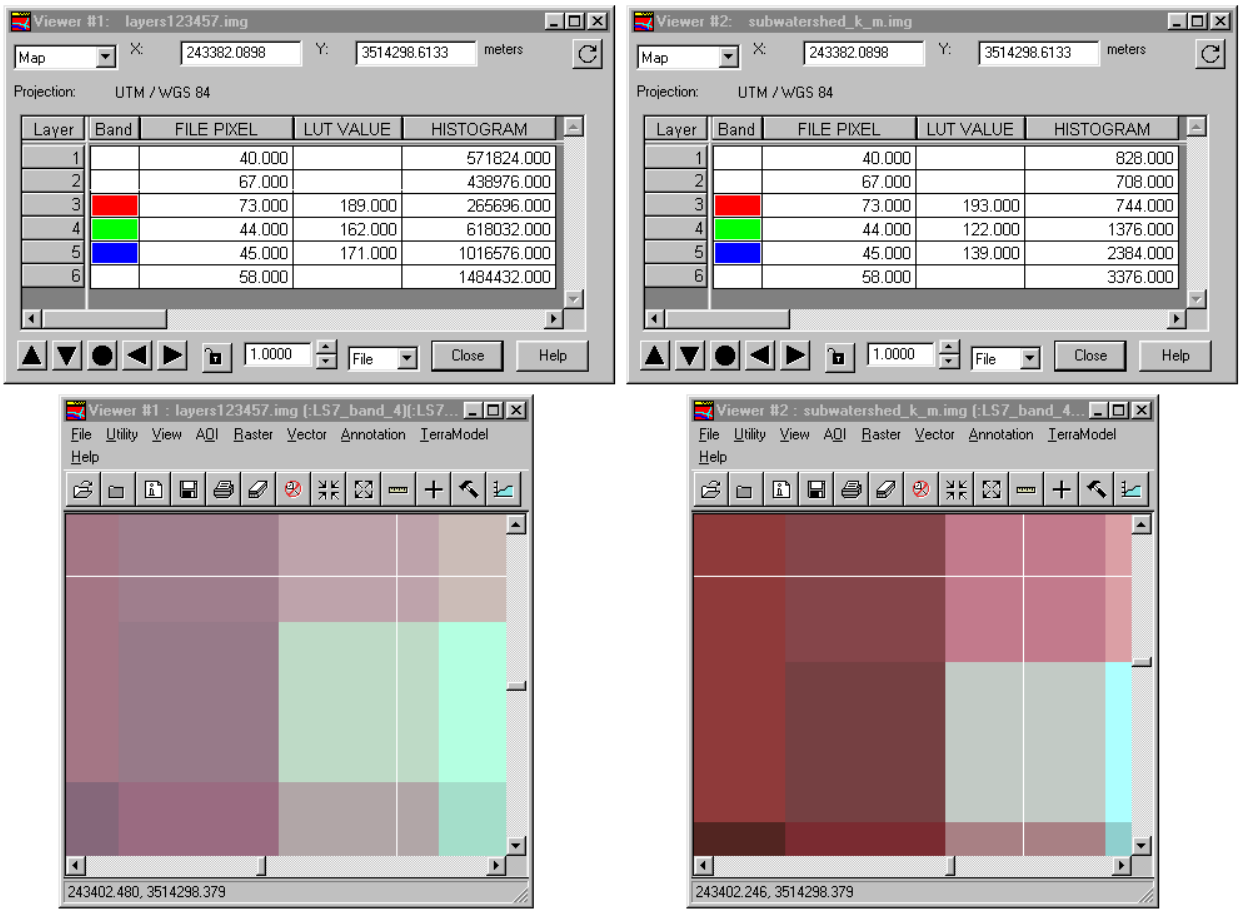
(a) A clipped portion of the source image with default Imagine display

(b) The Subwatershed K subset image with default Imagine display

Figure A1. Comparison of the default Imagine displays of (a) a portion of the source image that coincides in area and scale to (b) the Subwatershed K subset image.

In order to further investigate this problem, the two images displayed in Figure A1 were magnified using the Imagine zoom tool until individual pixels were clearly visible. The same geographical area was enlarged in the two images so that the enlarged pixels in the two images corresponded to the same geographical area. The Inquire Cursor function was used in each Viewer to display information about the data associated with individual pixels. This function caused crosshairs to appear in the Viewers. The crosshairs were carefully moved in one image until they were located at the same geographical coordinates as in the other image. Although the crosshairs were obviously not located in exactly the same portion of a pixel in both images, they were located in corresponding pixels, based on the fact that the zooming operation was carried out incrementally in both Viewers, increasing magnification by a factor of only two each

time, first in one Viewer, then in the other. In this way, a single pixel corresponding to a specific



(a) A portion of the source image (b) The Subwatershed K subset image

Figure A2. Lookup table values (LUT) for identical points with identical file pixel values in (a) a pixel in the full original scene that coincides with (b) a pixel in the Subwatershed K subset image.

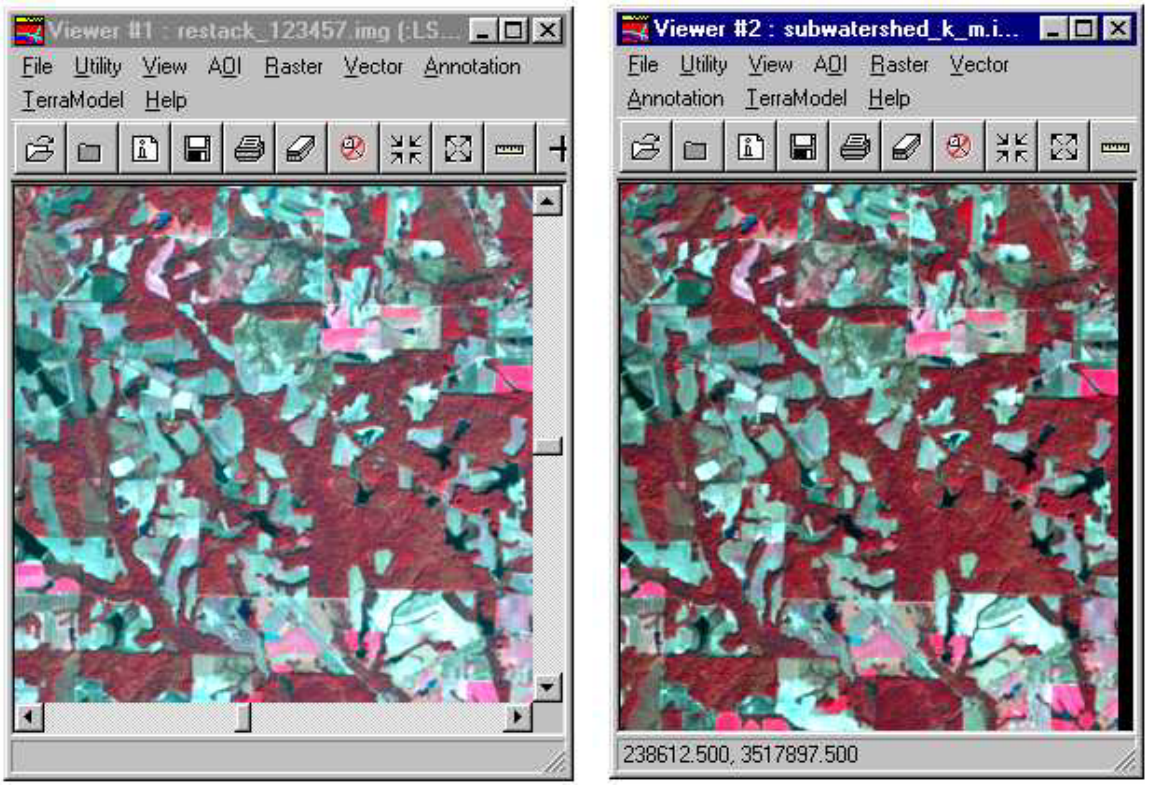
portion of the image was identified and zoomed into in each Viewer. The apparent shifting of corresponding pixels in the subset image is discussed further in a later part of this section.

Figure A2 depicts the enlarged areas and the listings of file pixel values, lookup table (LUT) values, and histogram values for the six layers making up both images. The

file pixel values are the reflectance values contained in the enhanced ETM+ dataset received from USGS. The file pixel values are identical for these two corresponding pixels in the two images. The LUT values, however, differ between the two images. LUT values are the numbers that determine the brightness, in this case on a 0 to 255 scale, for each of the three color guns (red, green, and blue) in the computer display monitor. The histogram values tabulated for the two images in Figure A2 represent the numbers of pixels in each image with exactly the same file pixel values as those listed in the associated lines in the tables. The histogram values associated with the source image are much larger than the histogram values for the subset image because there are just under 742 times the number of pixels in the source image (58,201,500) as there are in the subset image (78,470).

The problem of visible color differences between the source image and subset images can be dealt with in two ways. The first is to recompute the image statistics for the source image, instructing Imagine to ignore zeros in computing statistics. The image statistics include the histogram of pixel values for each band in the image, which is used to apply the linear stretch that is the default display algorithm used by Imagine in displaying an image in a Viewer. Figure A3 depicts a screen capture of a side by side comparison of a portion of the source image displayed with a linear stretch (the Imagine default) based on statistics computed ignoring zero values and the subset image previously discussed. The display colors depicted for the two images in Figure A3 are much closer in visual appearance than those displayed in Figure A1. The Inquire cursor tool was used on these two images in the same way that it had been on both the source image in which zeros had not been ignored in computing statistics and the subset image.

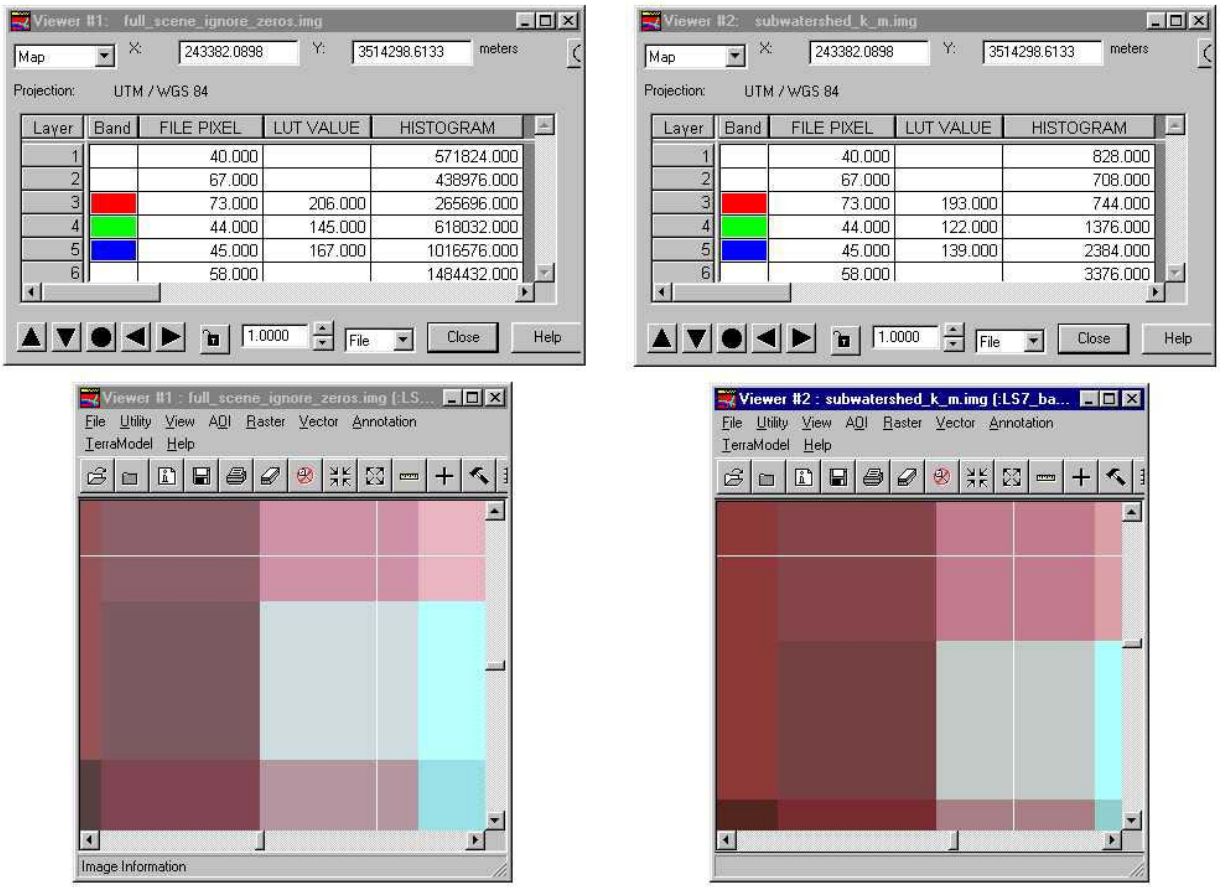
The results are depicted as a screen capture in Figure A4. Although the red LUT values are slightly farther apart in Figure A4 than in Figure A2, the blue and green LUT values are much closer together.



(a) A clipped portion of the source image with zeros ignored and default Imagine display

(b) The Subwatershed K subset image with default Imagine display

Figure A3. Comparison of the default Imagine displays of (a) a portion of the source image with image statistics recomputed ignoring zero values that coincides in area and scale to (b) the Subwatershed K subset image.



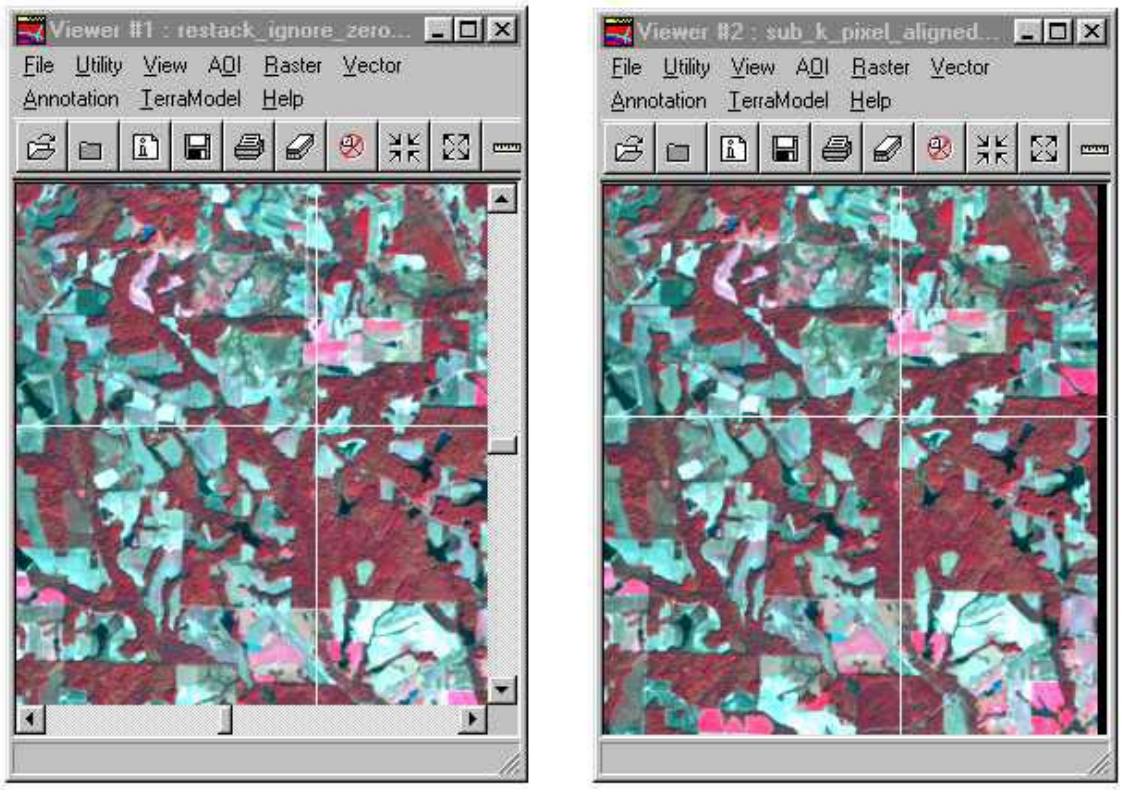
(a) A portion of the source image

(b) The Subwatershed K subset image

Figure A4 Lookup table values (LUT) for identical points with identical file pixel values in (a) a pixel in the source image with statistics recomputed ignoring zero values that coincides with (b) a pixel in the Subwatershed K subset image.

The second way to overcome the problem of color differences between a subset image and its source image is to use the same or nearly the same contrast histogram in displaying the two images. This can be done to a close approximation by copying the breakpoints of one image's display histogram and applying them to the other image. In Imagine, breakpoints refer to the transition points between linear sections of the histogram used to map file pixel values to one of 256 possible brightness values assigned to each of the computer display's three color guns (red, green, and blue). With the source

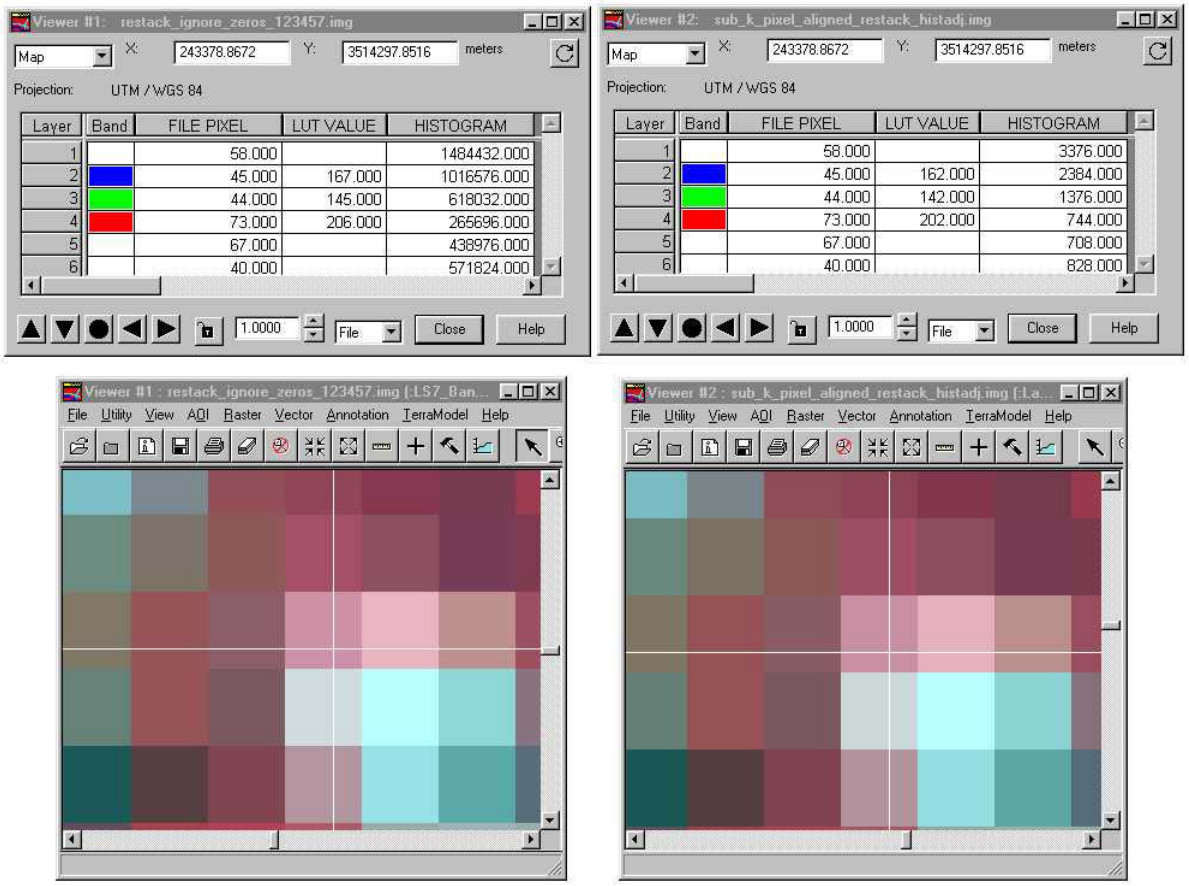
image displayed in a Viewer, the Raster - Contrast – Save Breakpoints command was used to save that image's breakpoints to a file. The subset image was then displayed in another Viewer and the Raster - Contrast – Load Breakpoints command was used to load the source image's breakpoints into the subset image. The results are depicted in Figures A5 and A6. Figure A5 indicates that the two images are visually quite similar, while Figure A6 demonstrates that there are still subtle differences in the LUT values of the two images. This is because the two images have different histograms of file pixel values and when Imagine maps the breakpoints to create display histograms for the two images, the results are still slightly different. Since all three LUT values are slightly higher in the source image than in the subset image, the result is that the intensity of the source image is slightly greater than the intensity of the subset image, while the hues of the two images are nearly identical. However, since the file pixel values rather than the LUT values are used for most image processing tasks, this is not a serious handicap. Similarity in appearance, achieved by creating similar LUT values for corresponding areas of a subset image and its source image, is useful for purposes of feature identification and location of identical features in both images. The level of display similarity demonstrated in Figure A5 is adequate for that purpose.



(a) a portion of the source image

(b) subset image contrast adjusted with breakpoints from source image

Figure A5. Comparison of (a) a portion of the source image with statistics computed ignoring zeros and (b) the subset image with contrast adjusted using the breakpoints from the source image (a).



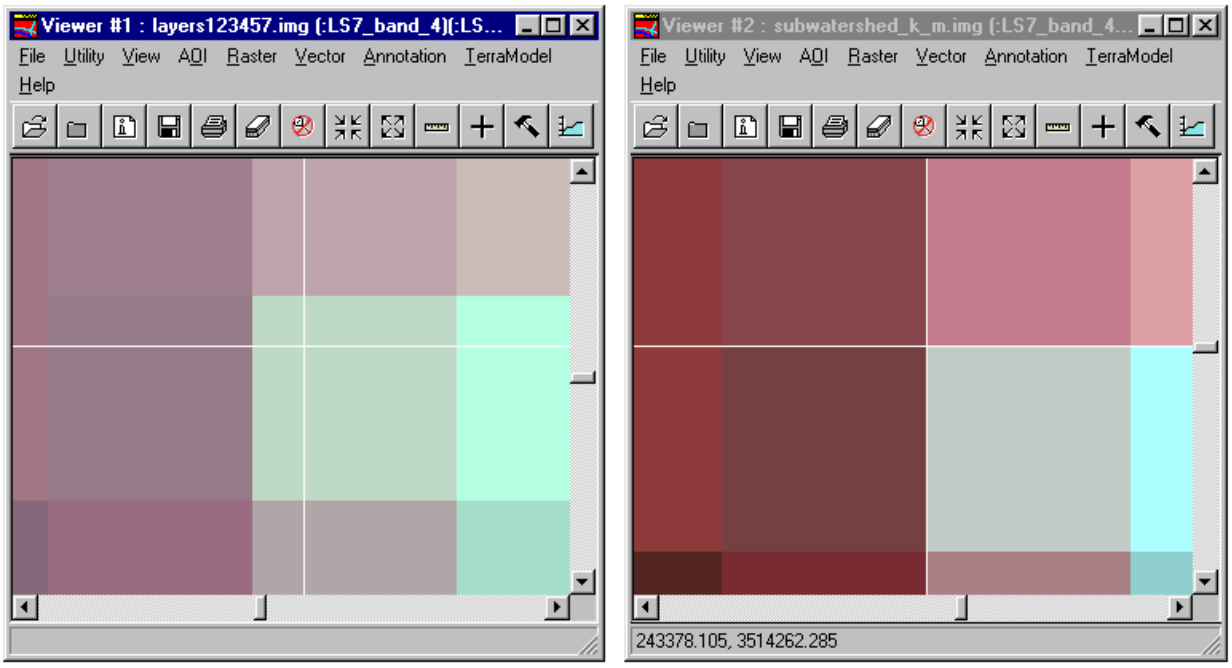
(a) LUT for a pixel in the source image

(b) LUT for the corresponding pixel in the subset image

Figure A6. Comparison of LUT values for (a) a portion of the source image with statistics computed ignoring zeros and (b) the subset image with contrast adjusted using the breakpoints from the source image (a).

One of the checks that was made to determine what had been done by the Imagine software to cause the display differences between the full scene and the subset image was to examine individual pixels in each image. This was done to check that the reflectance values, i.e., the data values in the corrected ETM+ dataset received from USGS, stored in the image files were identical for corresponding pixels in the two images. If this was the case, then the only logical remaining conclusion would be that the differences observed on the computer screen were the result of the manner in which the Imagine was

displaying the two images. The pixel to pixel comparison was made by opening two Viewers and displaying the subset image in one and the corresponding area of the full scene in the other. Both Viewers were displayed at the same on-screen scales. The two Viewers were then linked, using the Geographical Link feature of the Imagine Viewer. Each image was zoomed by the same factor until the individual 30 meter by 30 meter pixels were clearly visible. At this point, it became evident that the pixels did not precisely line up in the two Viewers. Figure A7 is a screen capture of the two linked Viewers with the enlarged pixels shown. The original full scene is on the left and the subset image is on the right. The crosshairs that appear in both Viewers are geographically linked and, therefore, represent the same geographical location in each Viewer. With the Viewers zoomed as indicated in Figure A7, it was possible to determine the geographical coordinates of the corners of corresponding pixels in the two Viewers to a precision of plus or minus 0.1 meter. Thus, it was determined that the pixels in the subset image had been shifted down 7.5 meters and right 7.5 meters, a total distance of 10.6 meters.



(a) A portion of the full scene linked to (b) Viewer 2

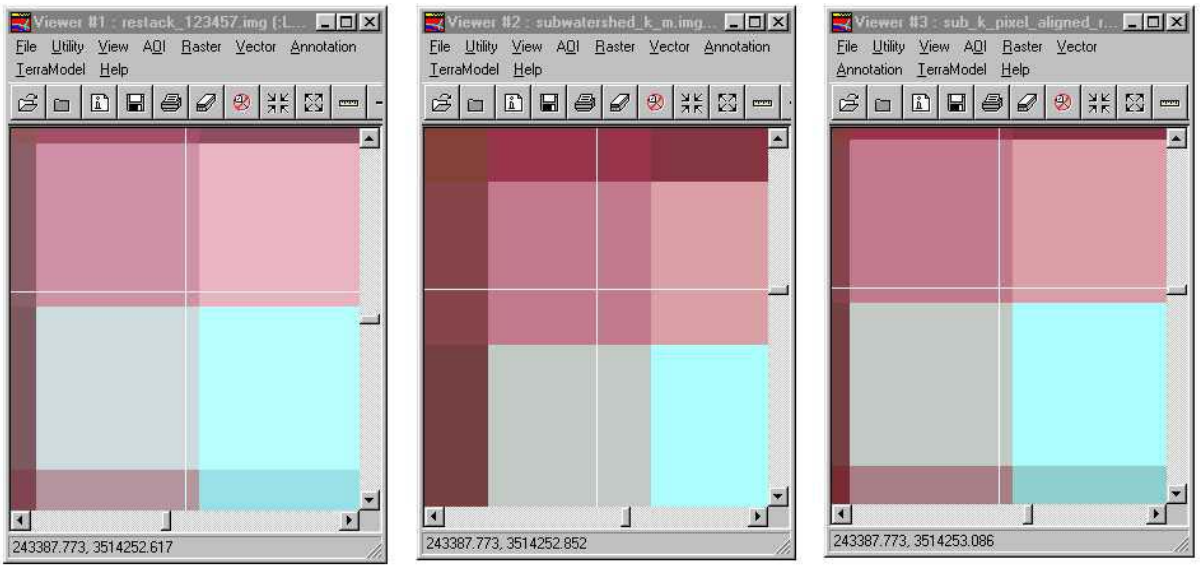
(b) The Subwatershed K subset image linked to (a) Viewer 1

Figure A7. Comparison of the location of linked crosshairs generated by the Inquire Cursor command in linked Viewers.

The key to understanding how this happened is to note that the subset image was defined by an on-screen Inquire Box procedure, whereby a box is dragged around the area to be subset visually in the Imagine Viewer. This operation then generates the coordinates of the upper left and lower right corners of the desired subset image. Apparently, because the on-screen box is defined by cursor movements that are not linked to pixel coordinates in the underlying image, this causes the upper left corner to be defined based on the cursor location rather that the upper left corner of a pixel in the image from which the subset is to be defined. The Imagine software then selects the whole number of rows and columns that most closely matches the Inquire Box and resets the coordinates of the upper left corner of the pixel in the upper left corner of the subset

image to the upper left coordinates of the Inquire Box. The result is a subset image that is shifted some fraction of a pixel in the X and Y directions but whose file pixel values remain unchanged. The result is the same as if a new image had been defined with pixels shifted slightly and with pixel values defined by a nearest neighbor resampling operation.

To avoid the offset pixels described above, it is simply necessary to define the upper left and lower right coordinates of the subset image to coincide with coordinates of pixel corners in the original image from which the subset is to be selected. In this study, the appropriate coordinates were determined by determining the upper left and lower right coordinates of both the original larger image and the pixel-shifted subset image. By subtracting the upper left X and Y coordinate values of the original larger image from the upper left coordinates of the pixel-shifted subset image and dividing the resultant two numbers by the pixel size (30 meters on a side), the X and Y components of the distance between the upper left corners of the original larger image and the pixel-shifted subset image were determined in pixels. In this case, the resultants were not whole numbers of pixels. These resultants were then rounded to the nearest whole numbers of pixels, multiplied by the pixel size in meters, and added back to the coordinates of the upper left corner of the original larger image to obtain coordinates of the upper left corner of a pixel in the original larger image nearest to the upper left corner of the first attempt at creating a subset image. The coordinates of the lower right corner of a pixel in the original larger image closest to the lower right corner of the first subset image were obtained in a similar manner. The coordinates thus obtained were then used to define the upper left and lower right corners of a new subset image that would have pixels precisely aligned with the pixels of the original larger image. The results are illustrated in Figure A8.



(a) Source image

(b) Subset defined on-screen with Inquire Box

(c) Subset defined by coordinates aligned with source image pixels

Figure A8. Inquire cursor crosshairs in three linked Viewers illustrating the alignment of pixels in (c) a subset image defined by coordinates that coincide with coordinates of pixel corners in the (a) source image but not aligned with (b) a subset image defined on-screen with Inquire Box.

The significance of pixel shifting in subset images is that it adds an element of geometric error to any attempts at correlating the image reflectance values with ground truth data. In the example described above, the subset image pixels were shifted 7.5 meters in the X direction and 7.5 meters in the Y direction. This is illustrated schematically in Figure A9.

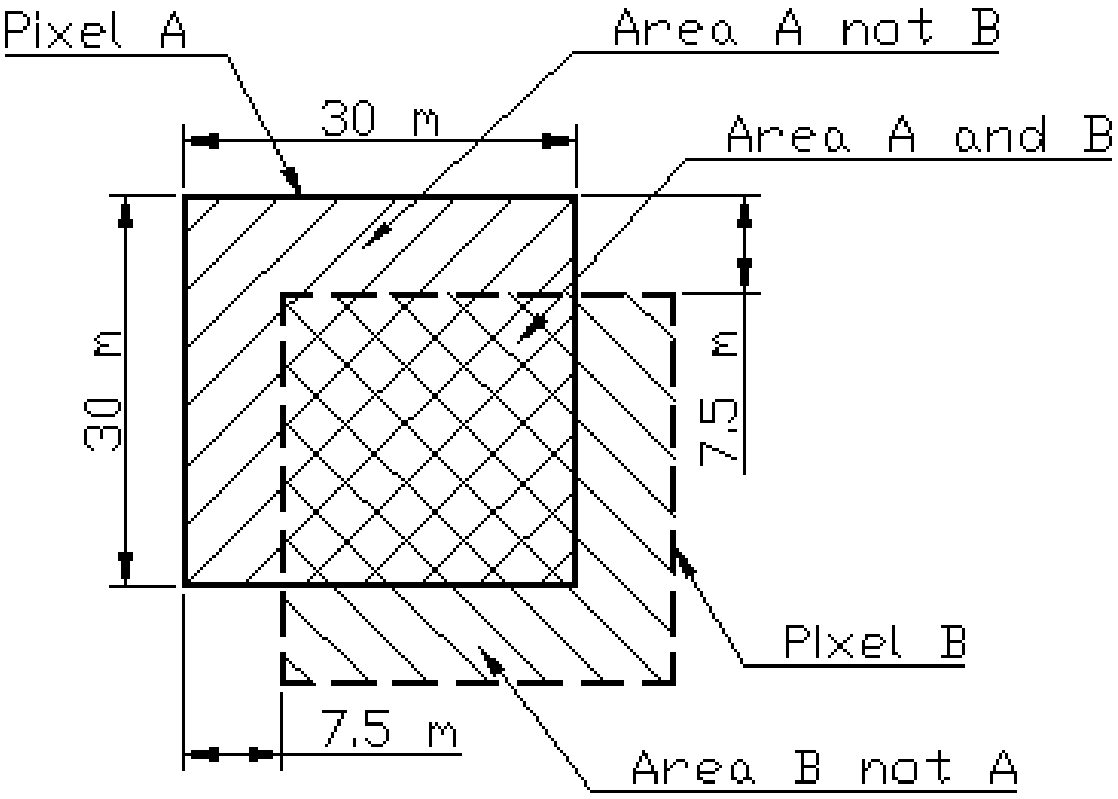


Figure A9. Geometric relationship of Pixel A (solid boundary) in the source image and the corresponding Pixel B (dashed boundary) in the subset image.

From Figure A9, it can be seen that the area identified as Area A and B is the only area in which objects that are contained within this particular pixel in the source image would also be contained within the corresponding pixel in the subset image. It follows that the probability of an object randomly located within pixel A also being located within pixel B is represented by the ratio of the area shared by the two pixels divided by the area of pixel A. That is, the probability, P, of a ground truth data point randomly located in pixel A being correctly located in pixel B, the subset pixel corresponding to pixel A, is found by

$$P = \frac{Area\ A\ and\ B}{Area\ A} = \frac{(30m - 7.5m)^2}{(30m)^2} = \frac{(22.5m)^2}{(30m)^2} = \frac{506.25m^2}{900m^2} = .5625$$

Thus, for a shift of only one fourth the linear dimension of a pixel in both the X and Y directions, the probability of a ground truth data point being assigned to the correct pixel is only slightly better than 56%. If ground truth data are geometrically well-defined and selected near the edge of an area represented by the data, then defining the boundaries of a subset image with coordinates matching coordinates of pixel corners in the source image becomes critically important.

APPENDIX B

SATELLITE IMAGE RECTIFICATION PROBLEMS

Rectifying the Landsat Enhanced Thematic Mapper Plus (ETM+) data involved selecting ground control points (GCPs) from digital raster graphic (DRG) versions of USGS 1:24,000 topographic maps, and pairing these points with control points (CPs) selected from the ETM+ image. This latter element turned out to be quite difficult when ArcView was used to display and extract GCP coordinates from the DRG and ERDAS IMAGINE 8.5 was used to select CPs paired with the GCPs. Once the DRG was opened in ArcView, it was rather easy to place the cursor on candidate GCPs in the DRG and read off the coordinates displayed by ArcView. However, the 30-meter pixel size of the ETM+ dataset led to great difficulty in precisely locating intersections of smaller, rural roads and other CPs in the ETM+ image that had appeared to be prominent points in the DRG. In order to aid in dealing with this problem, IMAGINE uses a point prediction function to place the cursor in the satellite image where the next CP is predicted to lie, based on the coordinates that had been entered for the associated GCP.

After entering several sets of GCP coordinates and locating their respective CPs in the ETM+ image, it was noticed that several previously entered CPs had been moved in the ETM+ image. After entering several more points, keeping close track of the locations of previously set CPs and making a record of their image coordinates (x- and y-

coordinates in the ETM+ image), it was discovered that point prediction was causing the problem. When "Predict Next" in "Edit – Point Prediction" was turned off, the shift problem went away. All previously entered CPs had to be relocated in the image, but none of the subsequently entered CPs were shifted after this was done.

APPENDIX C

DIGITIZING WITH IMAGINE 8.5

Imagine 8.5 incorporates some vector functionality within the main body of the program. However, for extensive vector operations, a user must install the Vector Module. This module uses the ARC/INFO data model, developed by ESRI, Inc. of Redlands, California (ERDAS 1997). Entering vector data (points, lines, or polygons) is accomplished by either digitizing from hardcopy originals using a digitizing tablet or digitizing from digital originals on-screen using a mouse pointing device. Digitizing for this project was accomplished with a mouse, with both ArcView and Imagine software.

In digitizing watershed boundaries and stream segments with ERDAS Imagine, a digital raster graphics (DRGs) image of a 7.5-minute topographic map was displayed on the screen as a raster layer. A new vector layer was then created to contain either the boundary or stream segments being digitized. A separate layer was created for each type of information, since different topology was defined for the two data types: polygons for watersheds and lines for stream segments. Watersheds were entered as polygons since this option caused the ERDAS Imagine software to automatically calculate the area within the polygons, which was useful for the comparison and analysis of watersheds. The following information concerning the mechanics of digitizing is presented because the process was not well documented in either the help files available with the software or with the print references available (ERDAS 1997, 49).

Digitizing with reference to an existing digital image, such as a DRG, is accomplished by displaying the image in an ERDAS Imagine viewer at a screen resolution of approximately 1:12,000. When a new vector layer is created in the same viewer, editing is automatically enabled. If the vector data are to be added to an existing vector layer (file), then it is necessary to first open the layer in the viewer and then enable editing, using the "Vector – Enable Editing" menu commands. Clicking "Vector – Tools..." opens a set of tools in a small window, which allow the creation and modification of vector data. In the case of the watershed boundaries, the "Place a simple closed polygon" tool was used to insert vertices defining a polygon into the vector layer in locations determined with reference to the raster DRG. The interfluve defining the watershed boundary was defined visually by evaluating the topography depicted in the DRG (points along hilltops, ridges, and saddles) and by referring to the drainage network depicted in the DGR. The Subwatershed K and West Fork Little River watershed boundaries were successfully digitized in this manner.

Unfortunately, when the polygon defining the East Fork Little River watershed boundary was completed and the mouse was double clicked to close it, the polygon disappeared. This polygon was created to overlap the West Fork Little River watershed boundary, with the intention of then using the "Clean" command to create polygon topology out of the digitized points and, simultaneously, insert nodes at the points where the two polygons intersected. The creation of these nodes would then define new endpoints to line segments making up the boundaries of the polygons and allow for the unneeded line segments to be deleted. The East Fork Little River watershed boundary was digitized again and the results disappeared again when the polygon was closed. At

this point, several smaller trial polygons were digitized overlapping the West Fork Little River watershed polygon and these were all completed, topology was created with "Clean", they were trimmed, and topology was revised with the "Build" command successfully. The full East Fork Little River watershed boundary was digitized again and the process failed again. At this point, the "Place a line feature in the coverage" vector tool was used to digitize the watershed boundary as a line but this process also failed. In two instances, when the mouse was double clicked to end the line, the 500+ points that were digitized disappeared and a single line segment appeared connecting the first and last points that were digitized. This process was tried twice more, with the same results one time and with a fatal error that shut down the viewer the other time. Digitizing was tried several more times after various combinations of closing all running programs, restarting ERDAS Imagine, and restarting the computer. None of these attempts were successful. Finally, success was achieved digitizing the watershed as a polygon after turning off all snapping options associated with digitizing (accessed through "Vector – Options..."). It is not known if turning snapping off directly led to successful digitizing or if this was coincidental.

After successfully completing the digitizing of the East Fork Little River watershed boundary, multiple backup copies were made of the results to minimize the risk of having to repeat this arduous task. The "Clean" command was then used to create topology and insert nodes at intersections. Unnecessary lines were remove by highlighting them and selecting "Vector – Delete" and the topology was corrected with the "Build" command.

Three separate vector layers were created and used to store digitized lines representing the stream networks depicted in the DRGs for the three watersheds. Digitizing of stream segments was carried out with the DRGs displayed at a screen resolution varying from 1:6,000 to 1:12,000, depending upon the complexity of the lines being digitized. Only perennial streams were digitized. Streams were digitized in segments corresponding to the segments used in determining Shreve Magnitudes and Strahler stream orders (Ritter, Kochel, and Miller 1995). Where an impoundment (man-made lake) was shown on a stream, the stream segment line was terminated at the downstream end of the impoundment unless the stream was further depicted in the DRG upstream of the impoundment. In this case, the stream segment line was digitized through the impoundment and was only terminated at the end of the perennial stream symbol in the DRG or at the next junction of stream segments upstream of the impoundment, whichever came first.