Mapping Headwater Streams:

Intermittent and Perennial Headwater Stream Model Development and Spatial Application

North Carolina Division of Water Quality

Final Report for Federal Highway Administration

Contract: Feasibility Study WBS: 36486.4.2

DWQ Contact: Periann Russell

January 29, 2008

EXECUTIVE SUMMARY

The North Carolina Department of Environment and Natural Resources (DENR) has numerous Divisions where maps depicting streams are used for regulatory and non-regulatory purposes. Additionally, a large number of other federal, state and local agencies use these maps for a wide variety of purposes. For instance, DOT uses available stream maps to compare alternatives, select potential road alignments, and to approximate the amount of stream mitigation required for a particular project. However, these maps and their depiction of streams were not originally designed for many of the technical and regulatory applications used today do not always accurately represent streams on the landscape. Accurate representation of streams, or at least, "known accuracy" of streams on maps, is a basic requirement for use in regulatory, non-regulatory and research programs across North Carolina. Improved stream maps will benefit DWQ and DOT as well as other state, federal and local government programs across the state.

In response to the need for accurate stream maps, the North Carolina Division of Water Quality, North Carolina Department of Transportation, and North Carolina State University partnered in early 2004 to initiate a pilot project to examine the feasibility of developing a stream map for the state that more accurately depicts 1st and 2nd order (headwater) streams. The products developed from DWQ headwater stream mapping will be integrated with the CGIA Stream Map (NC OneMap) as it is developed. For the pilot phase of the project, field data were collected from 23 watersheds in five ecoregions in the mountains, piedmont and coastal plain. Using the field data, terrain variables derived from a digital elevation model were used to fit a logistic regression model for the Carolina Slate Belt Ecoregion to predict headwater stream locations. This model was applied spatially to the Silk Hope site in Chatham County to test model performance.

While additional work is needed, each objective was completed with varying, though positive, results based on the results for the Silk Hope watershed. The use and application of logistic regression is effective in predicting the correct locations of streams (83% accuracy) and stream length (77% accuracy). The current version of the model poorly predicts flow duration, i.e., intermittent verses perennial flow, but further exploration of the model with additional

variables may provide more favorable results. The effort by DWQ and NCSU documented in this report demonstrates the feasibility of mapping headwater streams in North Carolina using field-collected data and logistic regression analysis.

The remaining short-term work for this project includes expanded logistic regression modeling with additional variables, and then application of the expanded model to the remaining ecoregions for which field data have been collected. We also plan to use these data with help from DOT staff to test the method on several selected DOT road projects to explore the applicability of the method. Long-term work includes collection of field data for the remaining 26 ecoregions in the state and subsequent modeling of stream locations for the remaining portion of the state, continued coordination with CGIA for the statewide NC OneMap, and systematic use of stream origin data as collected by federal, state and local agency staff, and consultants. The ultimate result of this work will be a stream map for the entire state, based on field data, GIS and logistical modeling with known levels of accuracy and will be useful to a variety of agencies, industry and the public.

Mapping Headwater Streams:
Intermittent and Perennial Headwater Stream Model
Development and Spatial Application

North Carolina Division of Water Quality
Final Report for Federal Highway Administration
Contract: Feasibility Study WBS: 36486.4.2
DWQ Contact: Periann Russell
January 29, 2008

Contract Description: The overall goal of the Headwater Stream Mapping project is to improve the accuracy and representation of headwater streams and their flow duration on stream maps in North Carolina. The primary objective of the Federal Highway Administration (FHWA) contract is to conduct a pilot study to evaluate the feasibility of developing a methodology for predicting intermittent and perennial stream origins within selected study areas in the coastal plain, the piedmont and the mountains.

Project Cooperators: North Carolina Department of Environment and Natural Resources (DENR) Division of Water Quality (DWQ), North Carolina Department of Transportation (DOT), North Carolina State University (NCSU)

Contract Deliverables (TO DOT): Final Report, Stream Maps and all GIS layers, electronic copies of all programs/macros/software written and used for mapping.

INTRODUCTION

The North Carolina Department of Environment and Natural Resources (DENR) has multiple divisions where maps depicting streams are used daily to complete program business and environmental goals. The Division of Water Quality (DWQ) is a DENR division where staff are dependent on United States Geological Survey (USGS) and Natural Resource Conservation Service (NRCS) soil maps to facilitate water quality monitoring and modeling programs, 401 certification, compliance and enforcement, water supply watershed protection, buffer and NPDES permitting. However, these maps and their depiction of streams were not originally intended for many of the technical and regulatory applications used today and consequently do not always accurately represent streams on the landscape. This fact is particularly true of 1st and 2nd order streams commonly referred to as headwater streams. Many states, including North Carolina, regulate activity in and near headwater streams for protection of water quality and aquatic resources as defined by the state administrative code. Accurate representation of streams, or at least, "known accuracy" of streams on presently used maps, is a basic requirement for use in research and regulatory programs across North Carolina. Improved stream maps benefit other state, federal and local government programs as well. The Ecosystem Enhancement Program and the North Carolina Department of Transportation need accurate stream maps in

- planning for mitigation needs,
- evaluating impact alternatives,
- identifying potential mitigation sites,
- and, facilitating hydraulic issues encountered in road construction and maintenance.

Additionally, timesavings will be realized by these agencies and the consultants that contract for them since more accurate maps will result in fewer stream delineations.

In response to the need for accurate headwater stream maps, the North Carolina Division of Water Quality, The North Carolina Department of Transportation, and North Carolina State University partnered in early 2004 to initiate a pilot project to develop a stream map for the state that more accurately depicts 1st and 2nd order streams. Additionally, DENR Center of Geographic Information and Analysis (CGIA) was tasked to produce a more accurate state-wide stream map in response to the Hurricane Recovery Act (Senate Bill 1152) passed in late 2004. Due to monetary, time and program constraints, neither program mapping effort alone will produce the multiple-use stream map needed by the state. For this reason, an agreement between DWQ and CGIA was made to integrate the headwater stream mapping products with CGIA to produce one North Carolina stream map that meets multiple use requirements.

Funding from FHWA provided many of the necessary components for the overall DWQ headwater mapping efforts completed to date. The work reported in this document is the result of several cooperators in addition to DWQ, and details progress, feasibility and future needs that encompass the total DWQ Headwater Stream Mapping Project. Particular acknowledgement is given to Dr. Tom Colson and Dr. James Gregory of NC State University who provided a substantial portion of the research and data used in project.

BACKGROUND

<u>Headwater Streams</u>

The hydrologic and ecologic importance of headwater streams in watersheds has been thoroughly documented (Coats 1972; Vannote et al., 1980; Kiffney et al., 2000; Peterson et al., 2001; McGlynn and Seibert, 2002; and others). More recently, The Journal of America Water Resources dedicated a portion of the February 2006 issue to headwater stream research. Headwater streams are the primary sources of water in a drainage network (Stanford 1996) and serve as a critical hydrologic link between the surrounding landscape and larger, downstream surface waters. The progressive downstream connection between small watersheds results in a continuous hydrologic network consisting of streams, rivers, ponds, lakes, and wetlands (Colson 2006). Due to their location and prevalence in the landscape, headwater streams are the primary transport mechanism for nonpoint source pollution since they convey stormwater and associated pollutants to downstream surface waters. Research suggests that small first order streams cumulatively drain up to 85% of a watershed area (McGlynn and Seibert 2002; Peterson et al. 2001). Additionally, headwater streams are an important component of the aquatic habitat, as they transport water, sediments, nutrients, organic matter, and woody debris to downstream reaches where they influence aquatic productivity (Kiffney et al. 2000; Vannote et al. 1980).

Geographical Information Systems (GIS) and Stream Mapping

Research over the past several decades has shown that many of the maps on which agencies rely greatly underestimate the number and length of headwater streams, misrepresent flow duration and are generally unsuitable for identifying streams (Morisawa 1957; Hansen 2001; Firman and Jacobs 2002; Heine et al. 2004; Paybins 2002; Colson 2006). With increased use and accessibility of Geographic Information Systems (GIS) and Digital Elevation Models (DEM), methods for improving the accuracy of stream representation have been and still are actively researched and analyzed (Dietrich et al., 1987; Montgomery and Dietrich 1989; Tarboton et al., 1992; Tucker et al., 2001; Garbrecht et al., 2001; Vogt et al., 2003; Heine et al. 2004). Recent research by Colson (2006) compared and analyzed multiple methods for creating and minimizing errors in digital elevation models (DEM), and examined the best DEM resolution for use in stream extraction. Colson (2006) also tested the use and accuracy of several existing stream extraction tools for use in predicting headwater streams. Based on the results of this research, existing stream extraction models do not perform at a level acceptable for meeting the objectives of the DWQ headwater stream mapping project.

The extensive research conducted by Colson (2006) with the guidance of NCSU professor Dr. James Gregory provides the basis and essential information for the continuation and success of the headwater stream project. For a complete literature review of GIS methods, stream map accuracy and the function of headwater streams, please refer to Colson's (2006)

Stream Network Delineation from High-Resolution Digital Elevation Models accessible by the link http://www.lib.ncsu.edu/theses/available/etd-10302006-122024/ and Colson and Gregory (2007 unpublished), Appendix E. Details of recommended procedures for DEM creation, refinement and resolution considerations are listed in Appendix A.

APPROACH AND OBJECTIVES

Based on the findings of Colson (2006), a plan for accurately mapping headwater streams was developed. The basic plan elements are:

- Use LiDAR data acquired from the NC Floodplain Mapping Program to create new topographic maps at a resolution of 5 meters (16.4 feet).
- Conduct field identification and mapping of 1st and 2nd order intermittent and perennial headwater streams and their origins in selected watersheds.
- Build a geodatabase of study watersheds to include field data, DEMs and DEM-derived data and other GIS data, e.g., slope, contributing drainage area, curvature, soils, precipitation and Ecoregion.
- Conduct spatial analysis using GIS data and field data.
- Use field data and other GIS data to create predictive models of streams and origins.

Accurate mapping of headwater streams must address stream representation as part of the larger concept of a stream network. The basic questions asked when referencing streams on current maps are 1) does the stream truly exist on the ground and, is there a stream on the ground that is not depicted on the map? Subsequently, if a stream is depicted on the map, is the length of the stream correct, and is that stream length correctly labeled as ephemeral, intermittent, or perennial? Each question refers to a specific element of a mapped stream within the stream network. Model performance with respect to these questions can be evaluated by performing a stream network accuracy assessment that specifically addresses different components of the stream network. The term “stream network accuracy” encompasses topography, geometry and hydrology, all of which are components of stream networks that influence geomorphic, hydrologic and biologic processes. Stream network accuracy considers multiple scales and is most important for DWQ, USACOE and EPA applications. Therefore, the objectives of this study are designed to capture several aspects of stream network accuracy of the predictive model.

The objectives, in order of priority, are to develop a spatially-based model that most accurately predicts:

- Presence or absence of a stream in the correct valley,
- Stream length, and
- Flow Duration – i.e., ephemeral, intermittent or perennial.

Development of predictive models requires the evaluation of landscape characteristics and processes to determine influencing factors on headwater streams, the processes these factors represent and how these factors and processes vary spatially. Understanding the strength and contribution of these factors is critical in developing a meaningful model and for evaluating its success in predicting the presence and extent of headwater streams.

STUDY SITES

Study sites for field data collection were selected to represent several EPA Level IV ecoregions (Griffith et al., 2002) with the number of sites per ecoregion based on the size the ecoregion (Figure 1). Since EPA Level IV Ecoregions are intended to represent landscapes with similar geology, topography and vegetation, sites were stratified by ecoregion given the assumption that ecoregions will control for landscape variability. Study watershed sizes ranged from 150 to 5500 acres in size and an average of 40 stream origins (intermittent and perennial) were surveyed within each study site. (Table 1).

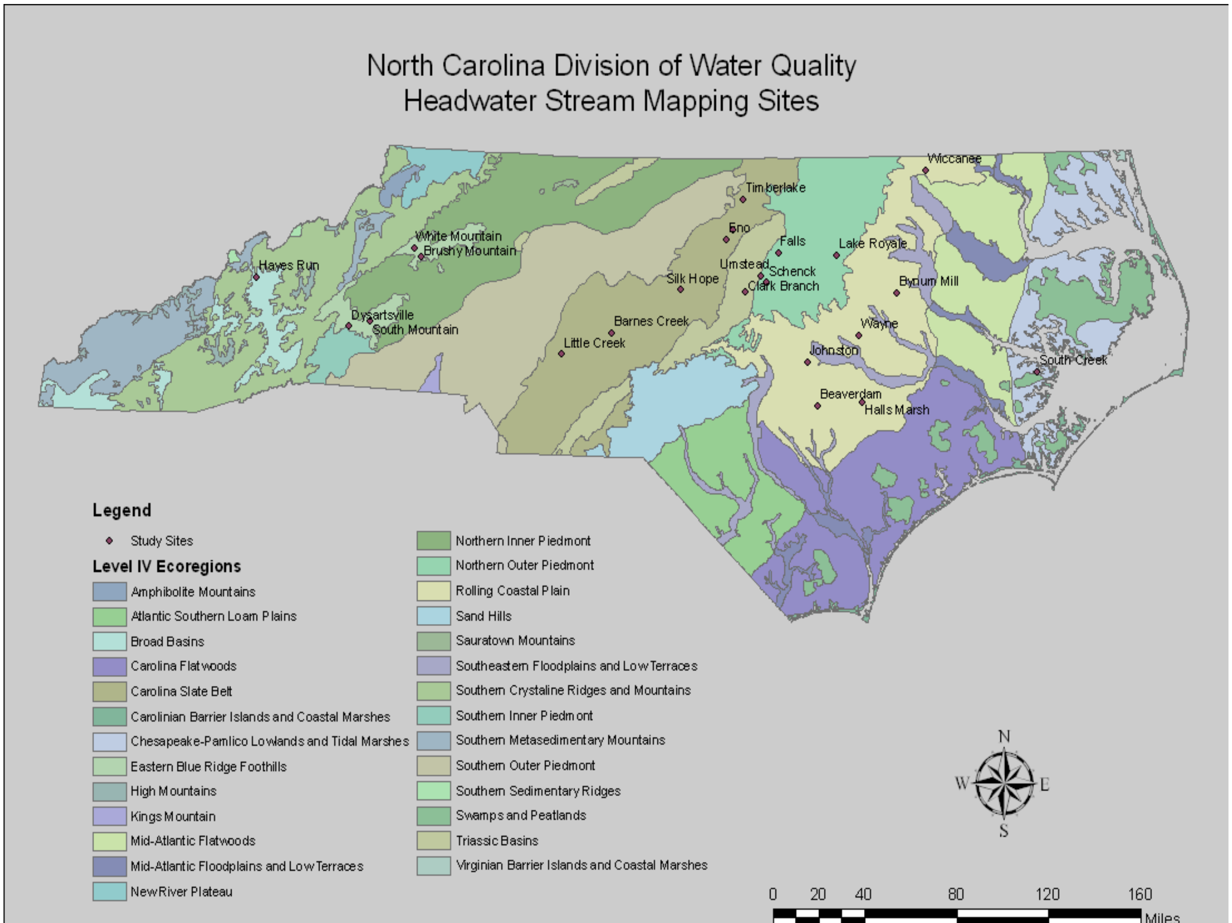

Figure 1: Map of Field Sites

Table 1: Study Sites (* Sites used in Model ** Site used for validation)

Site Name	Area (acres)	County	Level IV Ecoregion	River Basin
Hayes Run	1600	Madison	Broad Basin	Catawba
Barnes Creek*	1385	Randolph	Carolina Slate Belt	Yadkin
Eno*	1708	Orange	Carolina Slate Belt	Neuse
Little Creek*	1377	Stanly	Carolina Slate Belt	Yadkin
Northwest Durham*	1053	Orange	Carolina Slate Belt	Neuse
Silk Hope**	3919	Chatham	Carolina Slate Belt	Cape Fear

Site Name	Area (acres)	County	Level IV Ecoregion	River Basin
Timberlake*	1262	Person	Carolina Slate Belt	Neuse
South Creek	5438	Beaufort	Chesapeake-Pamlico	Tar-Pam
Brushy Mountain	1642	Caldwell	Eastern Blue Ridge	Catawba
Dysartville	1286	Rutherford	Eastern Blue Ridge	Broad
South Mountain	1312	Rutherford	Eastern Blue Ridge	Broad
White Mountain	727	Caldwell	Eastern Blue Ridge	Yadkin
Falls	150	Wake	Northern Outer Piedmont	Neuse
Lake Royale	215	Franklin	Northern Outer Piedmont	Neuse
Schenck	113	Wake	Northern Outer Piedmont	Neuse
Umstead	443	Wake	Northern Outer Piedmont	Neuse
Beaverdam	3921	Sampson	Rolling Coastal Plain	Cape Fear
Bynum Mill	1967	Edgecomb	Rolling Coastal Plain	Tar-Pam
Halls Marsh	3681	Duplin	Rolling Coastal Plain	Cape Fear
Johnston	1435	Johnston	Rolling Coastal Plain	Neuse
Wayne	1882	Wayne	Rolling Coastal Plain	Neuse
Wiccanee	2002	North Hampton	Rolling Coastal Plain	Chowan
Clark Branch	720	Durham	Triassic	Neuse
Jack Branch	684	Durham	Triassic	Neuse
Total	39,927			

METHODS

Field Delineation of Streams

Field data were collected by starting at the most downstream point of a study watershed and walking upstream until a tributary was observed. The tributary was subsequently followed upstream until another tributary was encountered, or until an origin (or origins) was reached. Both intermittent and perennial origins were identified on each stream where they existed. Origins were surveyed with a GPS capable of sub-meter accuracy according to standards established by DWQ with input from NC Geodetic Survey (Appendix A.5). Intermittent and/or perennial flow was determined by staff certified to use the North Carolina Stream Identification Methodology (NCDWQ 2005). Once the origin was identified and surveyed,, the field investigator returned to the mainstem to continue the walk upstream. The process was repeated until the entire stream network within the study watershed boundary was observed, all surface drainage features were investigated, and all origins and flow duration data were collected. FHWA funding financed a portion of the field mapping conducted by EcoScience Corp., Soil and Environmental Consultants, Inc. and Environmental Services, Inc. The reports submitted from each consulting group are included in Appendix B. The remaining field mapping was conducted by Dr. Tom Colson and Dr. James Gregory of NCSU and several DWQ staff.

Geographic Information Systems

Field Data

Stream mapping field data were transferred and post-processed with Leica and Trimble proprietary GPS software. ARCMap GIS shape files were created for use in the geo-spatial phase of work. Stream maps were created from field data by application of the free ESRI extention, TAUDEM, to a 5-meter resolution DEM to extract streamlines. The GIS stream network output was then manually edited to "trim" stream segments that extended upstream of GPSed stream origins. Methods for interpolation and pre-processing of DEMs and methods for extracting streams from DEMs were thoroughly evaluated by Colson (Chapter 6, Stream Network Delineation from High-Resolution Digital Elevation Models, 2007). In his research, Colson (2007) found that the greatest source of error in the positional accuracy of extracted streams is attributed to the DEM source. He concluded that interpolation of bare-earth LiDAR with the ArcMap tool ANUDEM, and the Impact Reduction Approach (IRA) (Lindsay and Creed 2005) for depression removal and drainage enforcement produced the DEM that, when used in the TAUDEM stream network delineation tool, resulted in the most accurate positioning of stream channels within the watershed.

Horizontal accuracy assessment was based on GPS points collected on the ground (overall GPS accuracy = 0.6 meters) in sequence along the centerline of the stream for each of his sites. Horizontal accuracy assessment was based on GPS points collected on the ground (overall GPS accuracy = 0.6 meters) in sequence along the centerline of the stream for each of his sites. Horizontal accuracy varied by ecoregion (topography), but the Carolina Slate Belt analysis indicated a range of 48.75% - 65.85% of GPS points fell within 3.06 meters of the centerline of the stream (6.10 meter buffer, 3.06 meters on each side of the stream). Colson's (2007) results provided the methods for producing baseline streams used in this study. These streams represent field-mapped streams and are referenced as such in the remainder of the document. Field-mapped streams are considered the standard to which modeled streams will be compared. These streams were also used to generate all spatial variables used in modeling.

Additional Data Sources

Ecoregion, soils and geology maps were available in digital format and easily acquired. LiDAR bare earth points were downloaded from NC Flood Mapping Program website (NCFMP 2003). As stated above, digital elevation models (DEM) were generated from LiDAR bare earth points using the combination of ARC TopoGrid (ANUDEM) and Terrain Analysis Software (IRA) tools evaluated by Colson (2007). All data were imported to an ARC geodatabase.

DEM-Derived Data (GIS)

Terrain characteristics were derived from the 5-meter (16.4 feet) resolution DEMs for testing their influence on headwater streams and flow duration. Terrain derivatives include local and averaged topographic slope curvature, gradient, and contributing drainage area for each origin. To date (November 2007), data from 21 watershed sites have been GIS-processed for analysis. Complete coverage of LiDAR bare earth points for two of the mountain sites has not been released to DWQ and delayed DEM processing. Tom Colson (NCSU) generated all terrain derivatives and detailed methods for obtaining derivatives can be found in Appendix C.

Statistical analysis and modeling were conducted using SAS and JMP software.

DATA DESCRIPTION

Field-Mapped Stream Origin Data

Stream origins were mapped in twenty-three watersheds resulting in over 500 mapped origins. GPS accuracy after differential correction varied across the state, but generally, GPS accuracy in the Rolling Coastal Plain averaged equal to or less than 2 meters, the Carolina Slate Belt average less than 1 meter, and in the Blue Ridge Foothills, averaged 1 meter. Potential error associated with individual field investigators using the DWQ stream identification method is unknown. However, it is a DWQ requirement that field investigators be trained in the stream identification method, and have indicated competency by passing a written and field test. Additionally, at the start of and during field data collection, several "calibration" days are conducted between DWQ certified staff and stream mapping field investigators to ensure consistent observation. For more information regarding the use and limitations of stream identification, please refer to the DWQ Stream Identification Manual (2005).

All completed stream maps created from field data are included in Appendix D.

Origin Transitions

The most common origin transitions are transitions from ephemeral flow to intermittent flow (46%) and intermittent flow to perennial flow (36%) (Table 2). These transitions reflect the current perception of headwater stream behavior where discharge and flow duration increase as drainage area increases. Variation in first order landscape influences is revealed in the distribution of transition types between geologic province and selected ecoregions. In the low-relief, sedimentary deposits of the Triassic Basin, ephemeral to intermittent transitions account for over 71% of all types within that ecoregion. In contrast, the majority of origins in the high relief, metamorphic rocks of the mountain sites are ephemeral to perennial and are typically springs. Few mountain stream origins transition from ephemeral to intermittent. Rolling Coastal Plain stream flow transitions are also predominately ephemeral to intermittent, but more transition types are present, including the 5% that are wetland to intermittent transitions.

Table 2: Headwater Stream Origin Transitions as Percentage of Totals

Type of Origin	Carolina Slate Belt	Northern Outer Piedmont	Triassic Basin	Rolling Coastal Plain	Eastern Blue Ridge Foothills	% Total Origin Type
	% Origin Type and Ecoregion					
Ephemeral to Intermittent	52	45	71	55	10	46
Intermittent to Perennial	30	51	17	36	43	36
Ephemeral to Perennial	11	0	10	1	47	14
Intermittent Modified	1	1	0	1	0	1
Perennial Modified	0	0	0	1	0	0
Intermittent Ditch	1	1	0	2	0	1
Perennial Ditch	0	1	0	0	0	0
Wetland to Intermittent Transition	1	0	1	5	0	1
Wetland to Perennial Transition	1	0	0	0	0	0
Intermittent to Pond	1	0	1	0	0	1
Pond to Intermittent	1	0	0	0	0	0
Perennial to Pond	1	0	0	0	0	0
Intermittent to Ephemeral	1	0	0	0	0	0
% Total by Ecoregion	30	15	15	20	20	100

Stream Length

The length of first order intermittent streams in the mountains, the coast and in the Triassic Basin is greater relative to first order perennial stream length, and is roughly equal in the

Carolina Slate Belt (Table 3). Although most streams begin as perennial springs (ephemeral to perennial transitions in Table 2) in the Eastern Blue Ridge Foothills, the length of the first order intermittent streams in this ecoregion is greater. First order perennial streams originating from springs may occur at lower elevations in the landscape, and subsequently have a shorter distance to travel to reach the next ordered stream.

Table 3: Percent Length of Headwater Streams by Level IV Ecoregion

		Length of Intermittent and Perennial Streams by Stream Order (%miles)							
Geologic Province	Ecoregion	1		2		3		Total	
		Int	Per	Int	Per	Int	Per	Int	Per
Mountains	Eastern Blue Ridge	33	22	8	20	0	16	42	58
		55		28		16		100	
Piedmont	Carolina Slate Belt	22	24	3	33	0	18	25	75
		46		36		18		100	
	Triassic Basin	26	18	2	38	0	16	28	72
		44		40		16		100	
	Northern Outer Piedmont	20	43	1	30	0	6	21	79
		63		31		6		100	
	Total	22	29	2	33	0	14	24	76
		51		35		14		100	
Coast	Rolling Coast Plain	32	18	5	29	0	16	37	63
		50		34		16		100	
	Chesapeake-Pamlico Lowlands and Tidal Marshes	56	44	0	0	0	0	56	44
		100		0	0	0	0	100	
	Total	34	21	4	26	0	14	39	61
		55		31		14		100	
All Total		28	25	4	29	0	14	31	69
		53		33		14		100	

The intermittent streams likely originate higher in the landscape, and thus, a longer distance is required to reach a 2nd order stream, or to accumulate sufficient watershed area for perennial flow. As the overall relief decreases east of the mountains, the Piedmont first order intermittent lengths also decrease about 10%. But, first order intermittent stream lengths tend to increase in response to declining slopes in the coastal areas. Not unexpectedly, these data suggest that stream processes are strongly associated with landscape characteristics (e.g., soils, precipitation, geology, and elevation) and the relationships need further examination.

The character of streams in the Chesapeake-Pamlico Lowlands is more indicative of outer coast plain streams that flow directly into estuaries and sounds. The South Creek area where the data were collected has a parallel drainage pattern with few incoming tributaries (http://h2o.enr.state.nc.us/ncwetlands/documents/pcsdocfinal.pdf). Additional data from this Ecoregion are needed to determine if the drainage pattern and stream flow duration is typical for similar areas.

Generally, the ratio of intermittent to perennial streams in the Piedmont is 1:3, that is, total intermittent length is approximately 1/3 of perennial length, whereas the approximate ratio in the mountains and coast is closer to 1:2. Although intermittent and perennial stream lengths vary with landscape, according to the current stream mapping data, about half the length of all first order streams across the state are intermittent streams.

Terrain Derivatives

Variables derived from 5-meter resolution DEMs for each study site are described in Table 4. These variables are indicative of water flow and flow paths on the landscape, and therefore provide information regarding surface (and near surface) water accumulation. The accuracy of terrain derivatives is highly dependent on the accuracy of the DEM source. DEMs are artificial representations of topography, and although they are usually based on groundtruthed elevation data, they are subject to the methods used to create, smooth and remove artifacts (Colson 2007). Kienzle (2004) found that the optimum grid cell size for creating derivatives is between 5 and 20 m, depending on terrain complexity and terrain derivative. Local and averaged upslope values were derived for the hillslope gradient and all curvatures. The curvature derivatives are numerical representations of shape, convex, concave or planar, and include plan and profile curvature (Figure 2).

Table 4: Terrain Derivatives

Independent Variable	Description
Local Slope	Slope at the point
Average Upslope Flow –Weighted Slope	Average slope above point
Contributing Drainage Area	Area draining to point
Local Profile Curvature	Profile curvature at point
Average Upslope Flow –Weighted Profile Curvature	Average profile curvature at the point
Local Plan Curvature	Plan curvature at point
Average Upslope Flow –Weighted Plan Curvature	Average plan curvature at the point
Local Curvature (local plan – local profile)	Curvature at point
Average Upslope Flow –Weighted Curvature (avg plan – avg profile)	Average curvature at the point

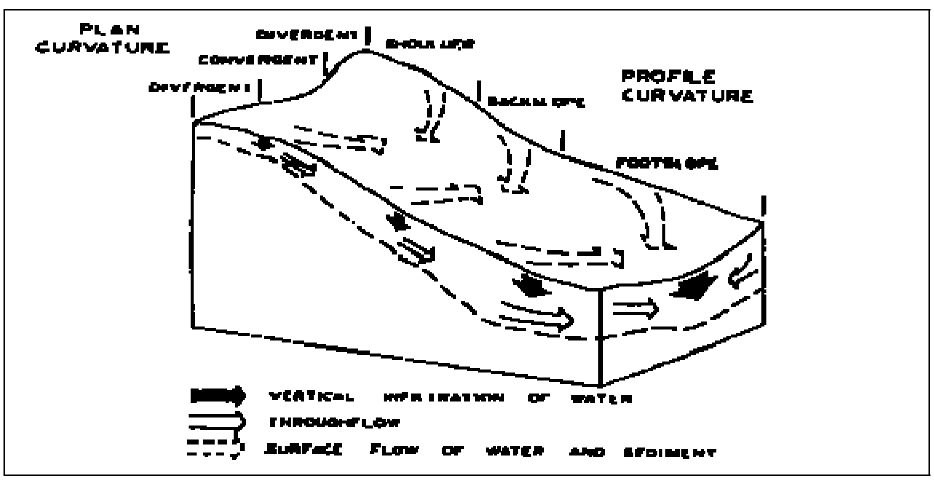

Figure 2: Profile and plan curvature. Source: Garg and Harrison 1990

Plan curvature is computed perpendicular to the slope and represents convergent or divergent flow. Profile curvature is computed in the direction of the slope and represents the change in flow velocity (Mitasova and Hoerka, 1993). The general curvature derivative is a measure of net curvature calculated as plan curvature-profile curvature.

MODEL DEVELOPMENT

Assessment of published literature indicates the use of logistic regression produced the most accurate results for modeling stream networks using GIS (Appendix C). Heine et al. (2004) modeled the channel network from the true channel head using logistic regression. The channel head in most research applications is the upper most point of water accumulation on the landscape, usually the ephemeral origin. The Heine et al. (2004) approach involved determination of the ephemeral origins in a small watershed by use of aerial photographs and then deriving terrain variables from a 10-meter resolution DEM for the same watershed. Average slope, plan and profile curvature and drainage area were included in a logistic regression model. The model was used to predict the probability of an ephemeral origin for each cell of the DEM. The authors concluded that a probability equal to and greater than 0.5 accounted for 95% of the stream network.

The basic principles described by Heine et al. (2004) were used to create a stream network. The DWQ approach departed from most of the stream network research, including Heine, et al. (2004), in that the intermittent or perennial origin is modeled rather than the ephemeral origin (true channel head). Additionally, the DWQ headwater stream model incorporates field-identified stream origin locations that may result in more accurate prediction of headwater streams. Due to time and resource constraints, variables used to develop the logistic regression model were limited to the terrain derivatives that could be measured from a DEM. The addition of other landscape variables is currently underway, but could not be included in this initial feasibility phase of the project.

Logistic Regression

The goal of logistic regression is to determine the "best fit" (yet reasonable) model to predict the probability of the success of a dichotomous or binomially distributed dependent variable as a function of a set of independent (predictor or explanatory) variables. The logistic regression model is expressed as

$$\text{logit}(p) = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \ldots + b_k X_k$$

where p is the probability of success, i.e., binary response = 1 and

$$\text{logit}(p) = \ln\left[\frac{p}{1-p}\right]$$

. For more detail about logistic regression, see http://www2.chass.ncsu.edu/garson/PA765/logistic.htm or http://www.medcalc.be/manual/logistic_regression.php. In this case, the dependent variable is either 0 (failure) - not a stream or 1 (success) – is a stream. One of the primary advantages of logistic regression is its flexibility in its assumptions. Unlike linear regression or discriminant analysis, logistic regression does not require the independent variables to be normally distributed, linearly related, or equal variance within each group (Tabachnick and Fidell, 1996, p575). However logistic regression assumes a linear relationship between the independent variables and the log odds (logit) of the dependent. The flexibility in logistic regression assumptions (opposed to linear regression) is analytically advantageous in obtaining a model where spatial data are used and the variability is poorly understood. While logistic regression does not have many assumptions, it does require a large sample size. It is recommended that at least 50 cases per independent variable be required for accurate hypothesis testing, especially when the dependant variable has many groups (Grimm and Yarnold, p. 221). http://www.kmentor.com/socio-tech-info/archives/000480.html. Although the field data collected for each stream includes delineation of flow duration, that is, intermittent and perennial origins, a dichotomous variable was defined as 0 (not a stream) or 1 (is a stream), regardless of flow duration. Hypothetically, the probability of a stream would increase in the downstream direction and the probability values would indicate the intermittent/perennial break.

Carolina Slate Belt Logistic Regression Model

Terrain variables derived from the DEM were used to fit a logistic regression model for the Carolina Slate Belt Ecoregion. Insufficient sample size for the Blue Ridge Foothills (LiDAR data gaps) and Triassic Basin (2 sites) Ecoregions prevented model development for those areas. Sufficient data was collected in the coastal plain region, but the Beaverdam dataset proved too large to process due to computer limitations. The sites used to model the Carolina Slate Belt include Little Creek, Barnes Creek, Timberlake, Eno and Northwest Durham. In a first step of the analysis, only independent variables correlated with the response and found to be statistically significant (p<=.05) were kept in the model (Schoonjans 2007). The full model/reduced model approach was used to determine the final model. The final model equation is

$$P = \frac{1}{1 + e^{\substack{(-21.2544 + (18.3974 * [\text{average flow-weighted slope}]) + (1.7469 * \mathrm{Ln}([\text{Contributing Drainage Area}])) + \\ (-9.3164 * [\text{local slope}]) + (-7.9802 * [\text{average flow-weighted plan curvature}]))}}}$$

where P = probability of a grid cell not being a stream, and 1-P = probability of a grid cell being stream.

MODEL VALIDATION

The logistic regression model was applied to the Silk Hope site in Chatham County to test its predictive capability. The two watersheds at the sites collectively drain 3919 acres of moderately sloped topography of 88 meters (279 feet) of relief. The Silk Hope site is located in the Carolina Slate Belt Ecoregion and has similar landscape characteristics as the sites used to develop the model. A probability threshold of 0.98 (i.e. cells with values <= 0.98 are considered part of the stream network) was selected to best represent the field-mapped stream network (Figure 3). Probability thresholds for intermittent (P=0.00375) and perennial (P=0.004255) stream delineation were selected independently to achieve the best fit with the field-mapped streams. The model was evaluated by performing accuracy assessments specific to each objective of the DWQ headwater stream project.

Stream Network Accuracy

Accuracy is defined by “the degree of conformity of a measure to a standard or a true value” (Merriam-Webster Online Dictionary, accessed January 4, 2008). For this study, the standard or true value is represented by the field-mapped streams created using the field-identified stream origins (GPS’d) and by editing the streamlines extracted with TAUDEM to end at the stream origins. Modeled streams are assessed for stream network accuracy against field-mapped streams and include three accuracy measures that describe the modeled stream network. The three measures are

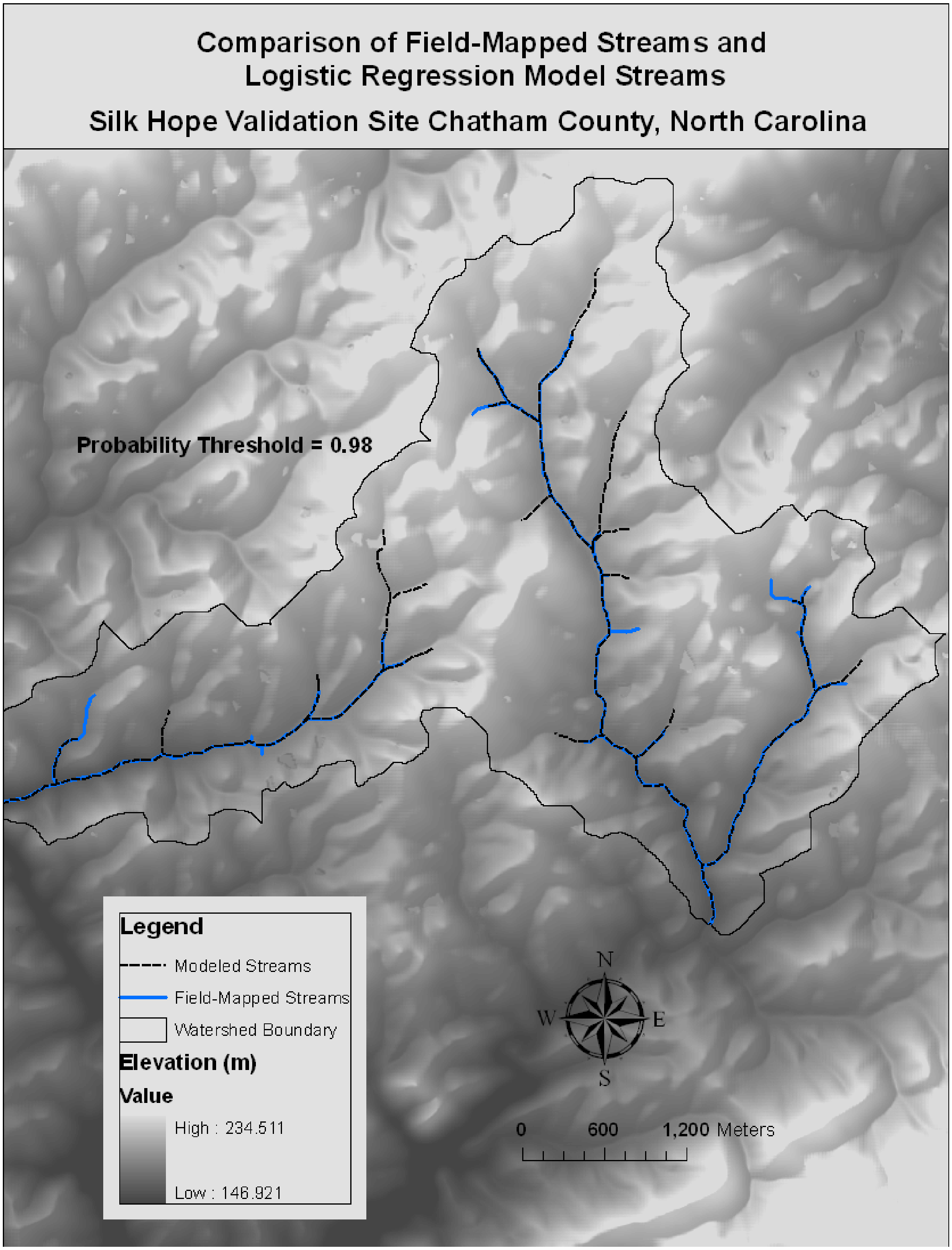

Figure 3: Comparison of Modeled Streams and Field-mapped Streams for Silk Hope Site.

1) Accuracy of presence/absence of a stream in the correct valley indicates the degree to which the model correctly identifies a valley that contains a stream. Presence/absence is measured as the percentage of streams occupying the same valley as the field-mapped streams.

2) Accuracy of stream length indicates the ability of the model to predict the same stream length as the field-mapped stream length. Length is measured as the percentage of modeled stream length by stream order that is the same as the field-mapped stream length by stream order. Accuracy is also measured by the evaluating the flowpath distance of the modeled stream endpoint to the field-mapped origin. The flowpath distance may be positive, 0, or negative representing over-prediction, correct prediction or under-prediction, respectively.

3) Accuracy of flow duration indicates the degree to which the model correctly identifies intermittent and perennial stream segments. Flow duration accuracy is measured as the percentage of intermittent and perennial stream cells that are coincident with the same intermittent and perennial stream cells as the field-mapped streams.

<u>Horizontal Accuracy</u>

Stream network accuracy (and its components) is the first order requirement for headwater stream mapping since correct network representation has the greatest influence on routing water, sediment, nutrients and organic matter through the watershed. NC One stream mapping reports horizontal accuracy, so horizontal accuracy is also included in this assessment. Accuracy of horizontal placement indicates the degree to which the modeled streamlines are placed along the same path as the field-mapped streams. Horizontal accuracy is measured as the percentage of modeled streamlines that are coincident, or that lie in the same 5-meter cell, as the field-mapped streams. Raster streamlines are located in the center of the cell allowing for a 2.5-

meter distance to the left and right of the centerline. Presently, horizontal accuracy is not considered a critical component of headwater stream mapping since it does not have an important influence on DWQ, USACOE and EPA's stream regulatory and non-regulatory programs. But, implications of the degree of horizontal accuracy required for DWQ and NC One should be re-visited in the future.

Assessment

The true skill statistic (TSS) (Allouche et al., 2006) is used to test the performance of the model for presence/absence and horizontal accuracy. TSS is an alternative to the traditional kappa statistic and evaluates predictive accuracy by use of omission and commission errors in the model. Allouche et al. (2006) developed the true skill statistic to eliminate statistical artifacts introduced by the kappa statistic due to its dependence on prevalence of true values. TSS corrects for dependence while maintaining all the advantages of kappa (Allouche et al. 2006). The kappa statistic was used to measure flow duration accuracy since it handles multiple cases of omission and commission errors.

Presence/Absence

Predictive accuracy of presence/absence of a stream in the correct valley was measured by constructing a confusion matrix that records the number of true positive, false positive, false negative and true negative cases predicted by the model (Figure 4). Each case represents the number of valleys correctly predicted to have streams (a-true positive), the number of valleys predicted to have streams that do not actually have streams (b-false positive), the number of valleys that have streams, but were not predicted by the model (c-false negative), the number of valleys correctly predicted to not have streams (d-true negative). TSS is calculated from a 2X2 matrix (Table 5) with the formulas listed in Table 6.

Table 5: 2X2 Data Matrix

		Field-Mapped	
		True	False
Silk Hope Model	True	(a) 18	(b) 4
	False	(c) 3	(d) 153

Table 6: TSS Formulas, Source:Allouche et al., 2007

Measure	Formula
Overall Accuracy	$\frac{a+d}{n}$
Sensitivity	$\frac{a}{a+c}$
Specificity	$\frac{d}{b+d}$
TSS	sensitivity + specificity - 1
n = a + b + c+ d	

Overall accuracy denotes the rate of correctly classified valleys (with and without streams). Sensitivity is the probability that the model will correctly identify a true positive and specificity is the probability that the model will correctly identify a true negative. TSS normalizes the overall accuracy by the accuracy that may occur by chance (Allouche et al., 2006). The results of the calculations indicate the true skill statistic is equal to 0.83 where 'perfect' accuracy would be equal to 1.0 (Table 7).

Table 7: Results

Measure			
Overall Accuracy	Sensitivity	Specificity	TSS
0.96	0.86	0.97	0.83

The results show the TSS equal to 0.83 indicating the high probability of the model to correctly identify valleys that actually have streams and valleys that do not actually have streams. The most relevant aspect of the results is that the model performed well at predicting the correct valleys that have streamflow thereby excluding valleys that only convey stormwater.

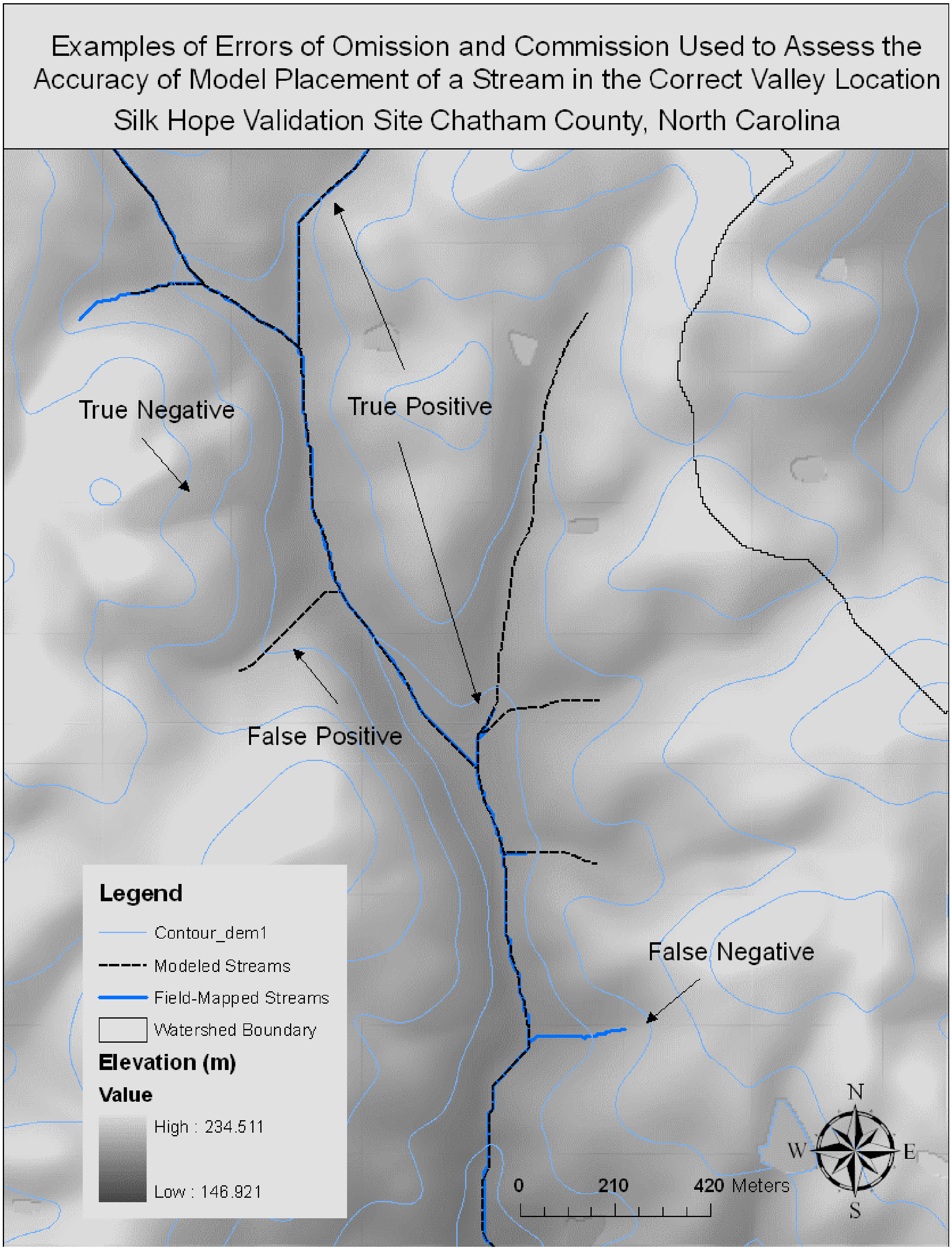

Figure 4: Omission and Commission Examples

Stream Length

Stream length accuracy was assessed by stream order and by first order stream flowpath distance from the modeled stream endpoint to the field-mapped stream origin (Figure 5). The overall accuracy of the model in predicting total stream length is 76.6% (Table 8). The model over-predicted total stream length by 23.4%, most of which occurs in 1st order catchments.

Table 8: Accuracy of Total Stream Length

	Total Length of Field-Mapped Streams (meters)	Total Length of Modeled Streams (meters)	Total Model Length Error (% over/under)	Overall Model Length Accuracy* (%)
Stream Order				
1st	4770	7943	66.5	33.5
2nd	9858	10228	3.8	96.3
3rd	462	451	-2.4	97.6
Total	15090	18622	23.4	76.6
*Overall model accuracy = 100% - model error				

Flowpath distance calculations indicate that approximately 87% of the model stream endpoints are less than 400 meters and approximately 23% are less than 100 meters from the field-mapped stream origins (Figure 6). The majority of modeled stream endpoints is less than 300 meters from the true stream origin and cumulatively represents 77% of the modeled streams. Approximately 928 meters of false-positive and 398 meters of false negative stream length contributed to the length of the 1st order streams. By simply improving the predictive capability of the model to correctly represent stream presence/absence (objective 1), the first order length accuracy increases by 28% thereby improving accuracy in the down network direction. Additionally, the maximum distance from a single modeled stream endpoint to the true origin is +918 meters and lies at the outlet of a large beaver impoundment. Since the model does not incorporate the influence of beaver activity, the stream was modeled as if the pond

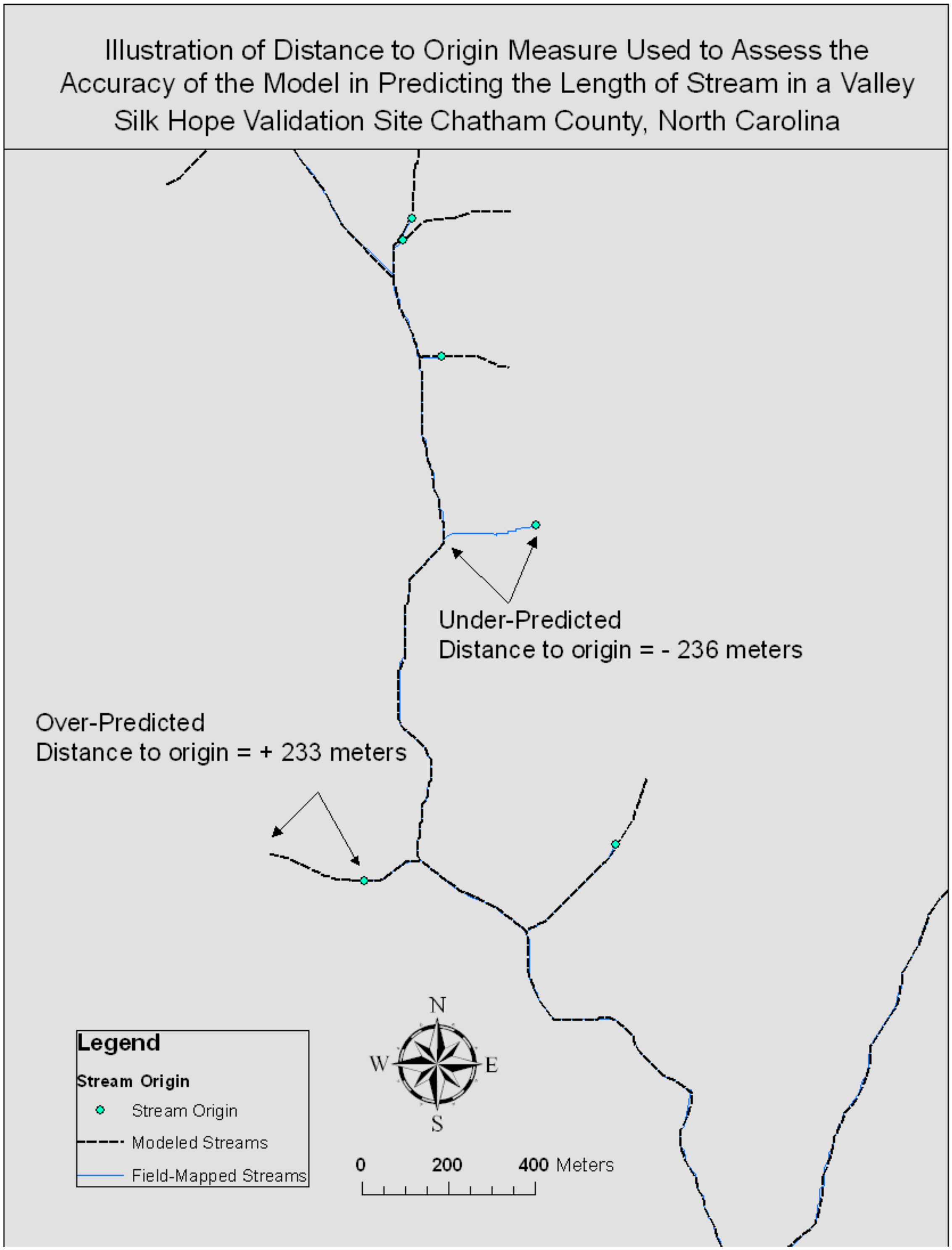

Figure 5: Example of Valley Distance from Stream Origin Measurement

were absent. The drainage enforcement algorithm (TAS) suggested by Colson forces flow from the lowest "upstream" side of the beaver pond to the lowest "downstream" side of the pond, and does not consider the possibility of sinuous flow in absence of the pond. This is the procedure used by the USGS when mapping a Flowline as an "Artificial Path" through water bodies (pers. Comm., Colson 2008). By removing the beaver pond stream outlier, the total 1st order stream length accuracy increases by 19%. Managing natural and manmade impoundments is one of the issues slated for future analysis.

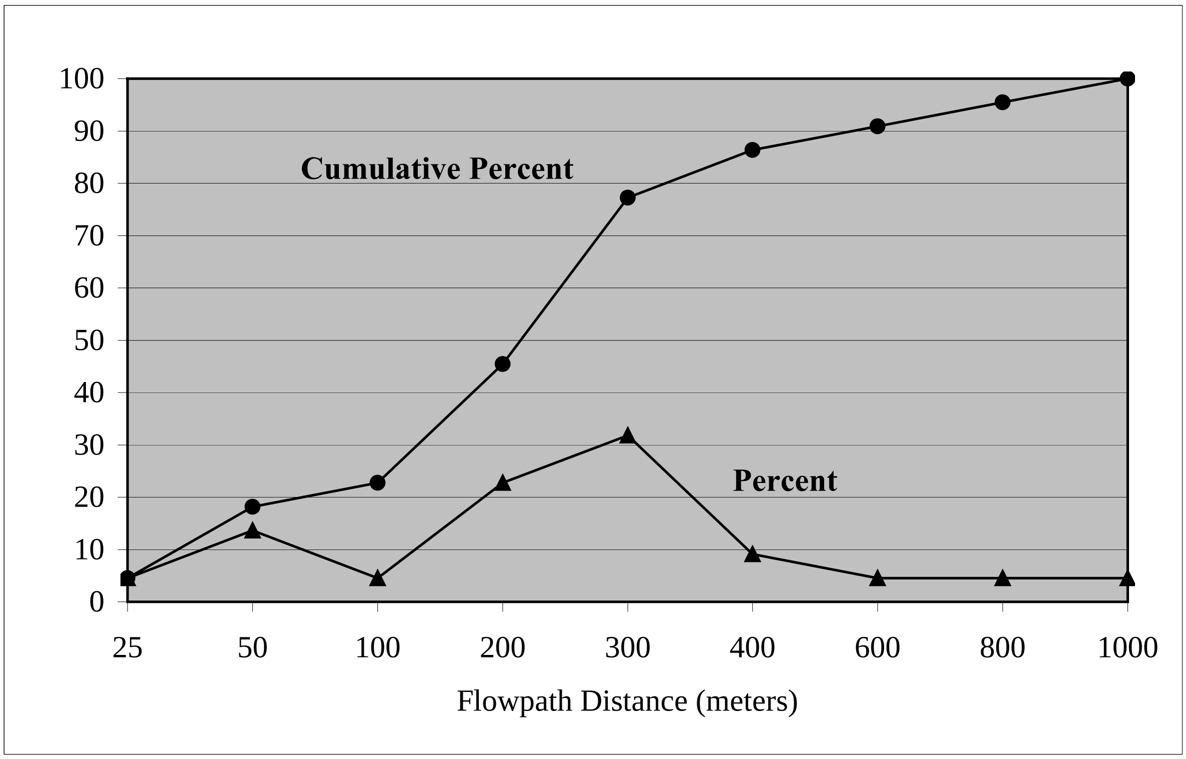

Figure 6: Percentage of Modeled Stream Endpoints within Flowpath Distance categories of Field-mapped Stream Origins.

Flow Duration

Flow duration analysis for the Silk Hope modeled streams was conducted independently of presence/absence and length analysis by NCSU (Colson and Gregory 2007, Unpublished report, Appendix E) using the kappa statistic previously mentioned. Colson & Gregory produced a 4X4 confusion matrix that incorporates omission and commission errors of correct and incorrect identification of intermittent, perennial and "not" streams. Their analysis resulted in a kappa statistic of 20% accuracy for intermittent streams and 47% accuracy for perennial streams.

Horizontal Accuracy

Horizontal accuracy was evaluated with the true skill statistic similar to the presence/absence assessment previously described, but was conducted on a cell-by-cell basis. To evaluate alignment only, analysis was conducted only on the streams that share the same valley and stream length from each grid (Figure 7). The exclusion of the "non-matching" streamlines allows testing alignment where alignment is possible. True and false positive and negative values for each cell in the grid were used to construct a 2X2 matrix and the same formulas applied to the data. The use of this method provides omission and commission accuracy of cells, and therefore, stream alignment. However, since the 5-meter resolution field-mapped streams provide the standard for accuracy, five meters is the minimum threshold of precision that can be assessed (2.5 m either side of centerline). Prior research determined that 45% - 67% of the GPS'd stream points were within 3.05 meters of the TAUDEM streamlines (Colson 2007), resulting in an inferred accuracy of greater than or equal to 45%- 67%. The TSS analysis for the modeled streams resulted in an overall horizontal accuracy of 0.988 representing the model's

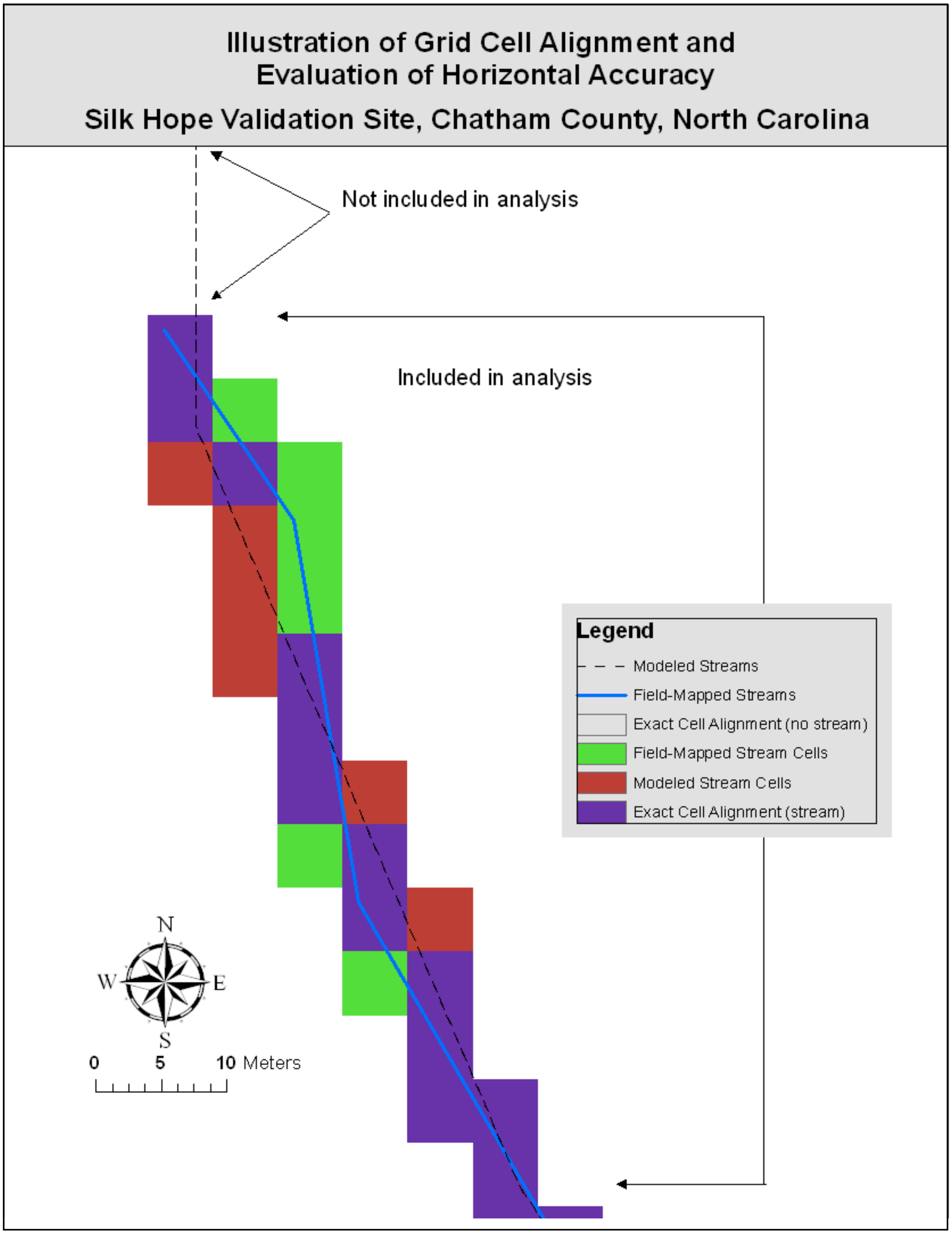

Figure 7: Cells Included and Excluded in Horizontal Analysis

total rate of correctly classified cells. The true skill statistic normalizes the overall accuracy by incorporating sensitivity (cell alignment) and specificity (correctly identify a cell as not a stream) resulting in a TSS value of 0.694.

Table 9: Horizontal Accuracy

Measure			
Overall Accuracy	Sensitivity	Specificity	TSS
0.988	0.699	0.995	0.694

DISCUSSION

The application of the logistic regression model to the Silk Hope site in the Carolina Slate Belt ecoregion performed well at predicting the stream network, particularly since the variables were limited to those measured from a DEM. Most notable is the ability of the model to correctly identify valleys where streams exist and do not exist. A large body of research addresses locating the channel head usually associated with ephemeral streams, but few have considered locating valleys that convey only intermittent or perennial surface water. Correctly predicting the presence/absence of a stream in a valley is the first step in mapping stream networks. Further evaluation of landscape characteristics and processes that influence headwater streams and their origins, and how they vary spatially, will help identify additional variables needed for model improvement.

Predicting accurate stream length is dependent on the ability of the model to determine the correct location of a stream origin.   The same landscape forms and processes that influence the ephemeral stream origin become more complicated in their influence on intermittent and perennial stream origins, particularly due to groundwater interactions. Stream length accuracy will likely improve simply by addressing stream presence/absence. However, further analysis of influencing landscape characteristics may reveal other factors that specifically apply to stream length. In the accuracy assessment for the current model, the total stream length predicted by the model accounted for 76.6% of the field-mapped stream length. Second and third order streams were more accurately represented at approximately 96% and 97%, respectively. Stream order is a numerical representation of a stream's location or contribution within a network and is based on a counting method (Strahler 1957). The number and configuration of first order streams affects the determination of second order streams, so errors in first order stream presence/absence and length are carried through to higher stream orders. Since most of the error in the modeled stream length occurs in first order streams, increased accuracy would be expected given model improvement targeted at first order streams. With that said, the current version of the model as applied to the Silk Hope site still provides total stream length and second and third order stream length at an accuracy acceptable for use in planning, such as estimating stream impact lengths, stream mitigation needs (linear feet), and buffer square footage.

Further analysis and evaluation of the model, along with the addition of independent variables as described above will likely contribute to increased model prediction of flow duration. Correctly identifying transition points or zones in flow duration along a stream is a critical part of modeling headwater streams. However, the terrain variables used to build the logistic regression model do not adequately represent the factors influencing flow duration and so this aspect of stream mapping was not explored in depth Modeling flow duration is strongly dependent on correctly identifying variables that describe, index or act as surrogates for ground water. The additional data acquisition and analysis mentioned previously will provide information as to the feasibility and potential for predicting flow duration.

ADDITIONAL DATA ANALYSIS

Variable Distribution

While all of the terrain derivatives likely have an influence on headwater streams, their relative contribution likely varies across different landscapes. Descriptive statistics for average slope and the contributing drainage area above stream origins indicate variation within and between Ecoregions. Statistical analysis of the remaining terrain derivatives are not complete and so are not included in this section. While normal distribution and equal variance are not requirements for logistic regression, statistics are used here to quantitatively describe, and potentially gain more understanding, of the independent variables.

Comparisons Between Ecoregions

Distributions of average slope and contributing drainage area above intermittent and perennial stream origins illustrate the range of values between Ecoregions (Figure 8 and 9, Tables 10 & 11). During the analysis process, closer examination of the Carolina Slate Belt data indicated higher than expected variability between sites. Additional analysis revealed similarities and differences in rock strength between sites and so the Slate Belt was divided into 2 groups, A-strong rocks (Silk Hope, Little Creek, Timberlake) and B-weak rocks (Eno, Northwest Durham, Barnes Creek) (pers. comm., Phil Bradley, Professional Geologist, North Carolina Geologic Survey, 2007). Geology may be one of the explanatory variables due for further analysis in the near future and therefore, its use to sub-divide Slate Belt Ecoregion sites is preliminary.
As expected, the Eastern Blue Ridge Foothills and Rolling Coastal Plain Ecoregion average slopes represent the upper and lower extremes of the data collected thus far. The average slopes above intermittent origins in the foothills and coastal plain are the only slopes that are

significantly different from all other Ecoregions. Slopes in remaining areas are similar to at least one other area. Slopes above intermittent stream origins tend to have less overlap between piedmont and coast Ecoregions than perennial stream origins. The differences between contributing drainage areas above both intermittent and perennial origins are coincident with the less competent rocks in the Slate Belt-B and coastal plain origins and the harder, more competent rocks in the other regions. Triassic Basin stream origin drainage areas are not statistically different than those in the competent piedmont Ecoregions even though Triassic rocks are considered relatively weak. But, since Triassic rocks and soils have very low infiltration rates, appx. 4% (Weaver 1998), surface flows behave similarly to harder, impervious rocks.

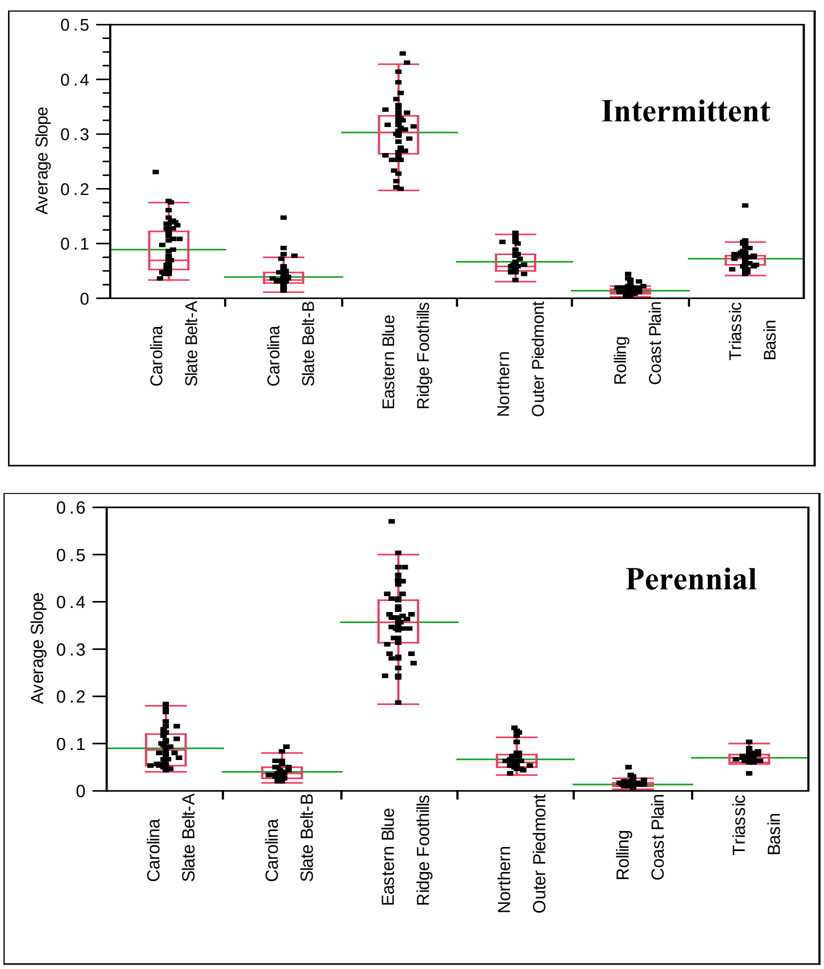

Figure 8: Distribution of Average Slope Above Intermittent and Perennial Stream Origins

Table 10: Significance Between Ecoregions – Avg Slope

	Intermittent							Perennial				
Ecoregion						Mean	Ecoregion					Mean
Eastern Blue Ridge Foothills	A					0.3032	Eastern Blue Ridge Foothills	A				0.3581
Carolina Slate Belt-A		B				0.0877	Carolina Slate Belt-A		B			0.0906
Triassic Basin		B	C			0.0718	Triassic Basin		B	C		0.0698
Northern Outer Piedmont			C			0.0666	Northern Outer Piedmont		B	C		0.0682
Carolina Slate Belt-B				D		0.0394	Carolina Slate Belt-B			C	D	0.0398
Rolling Coast Plain					E	0.0129	Rolling Coast Plain				D	0.0133

Levels not connected by same letter are significantly different at Alpha=0.05.

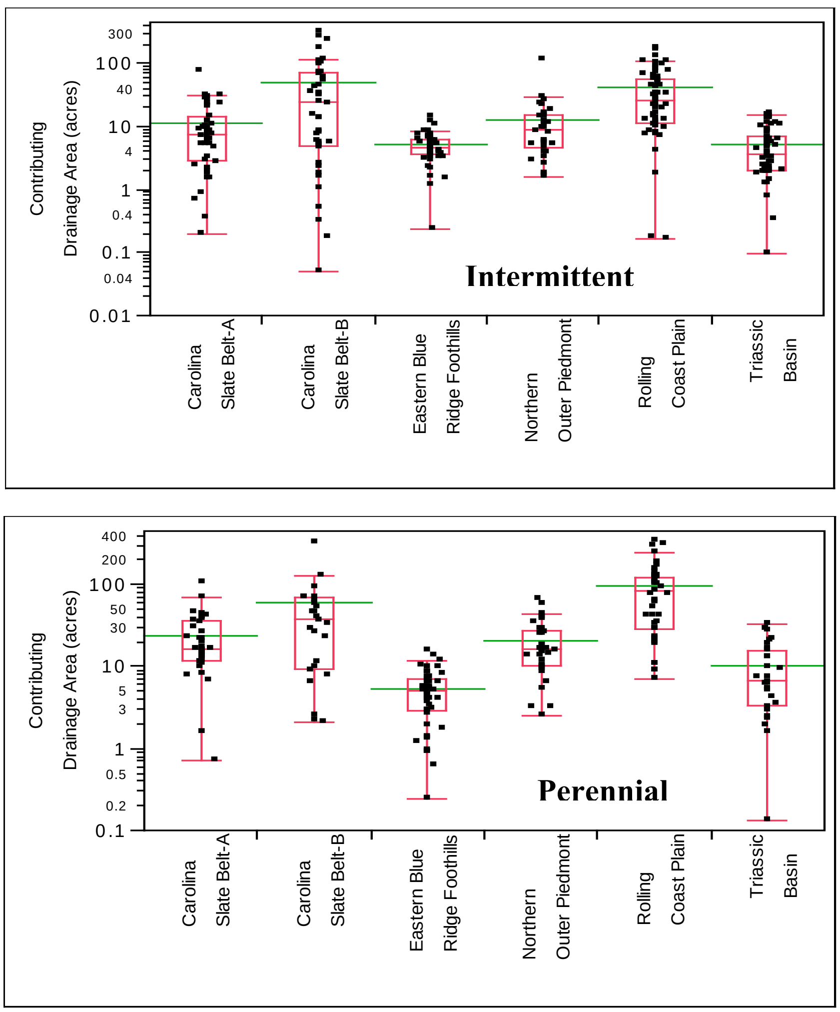

Figure 9: Distribution of Contributing Drainage Area Above Intermittent and Perennial Stream Origins

Table 11: Significance Between Ecoregions -Area

	Intermittent				Perennial			
Ecoregion			Mean	Ecoregion				Mean
Carolina Slate Belt-B	A		50.8631	Rolling Coast Plain	A			95.5882
Rolling Coast Plain	A		40.6575	Carolina Slate Belt-B		B		60.8503
Northern Outer Piedmont		B	12.7221	Carolina Slate Belt-A			C	23.7380
Carolina SlateBelt-A		B	11.1957	Northern Outer Piedmont			C	20.5212
Eastern Blue Ridge Foothills		B	5.1570	Triassic Basin			C	10.4031
Triassic Basin		B	5.1138	Eastern Blue Ridge Foothills			C	5.2683

Levels not connected by same letter are significantly different at Alpha=0.05.

The range in average slope tends to correspond with topographic relief. Low relief areas, such as the Coast Plain and Triassic Basin, have narrow slope ranges, but in high relief terrains, the range broadens to reflect the overall availability of varying slopes. Generally, the opposite is true for contributing drainage area where lower relief corresponds to higher drainage area and visa versa (Figure 10).

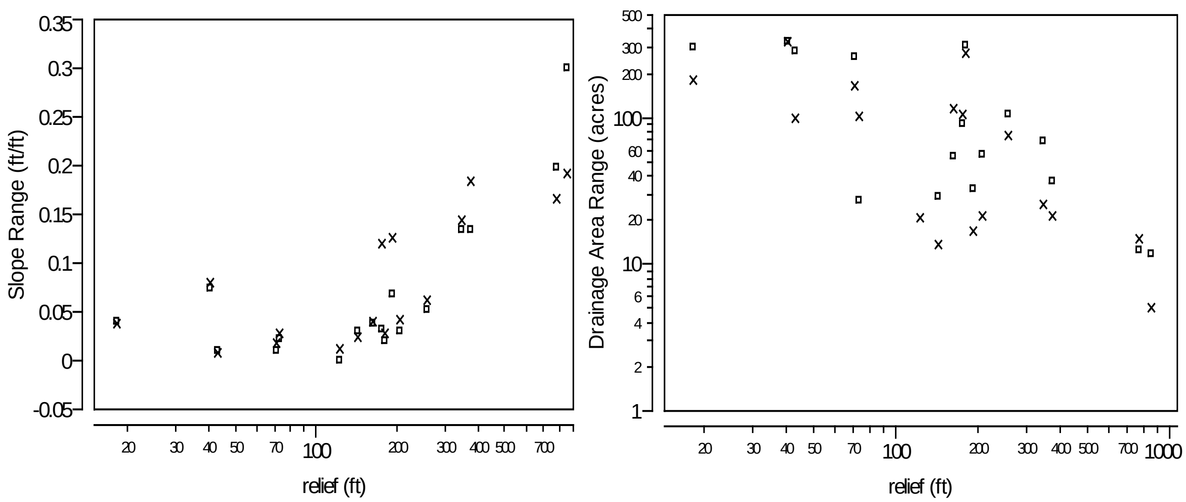

Figure 10: Trend Plots of Average Slope and Contributing Drainage Area (acres) Range (max-min) and Topographic Relief (ft) by Site. x=intermittent, □=perennial

Comparisons Within Ecoregions

With the exception of the Eastern Blue Ridge Foothill Ecoregion, the average slope above an intermittent origin is not statistically different than the average slope above a perennial origin (Table 12). Contributing drainage area distribution between intermittent and perennial origins indicated all were significantly different (Lower Confidence only for Northern Outer Piedmont) except the Eastern Blue Ridge Foothills and the group B-weak rocks in the Carolina Slate Belt (Tables 13 & 14).

Table 12: Average Slope Means Between Intermittent and Perennial Origins

Ecoregion	Difference	Upper CL Difference	Prob > |t|	Lower CL Difference	Prob > t
Carolina Slate Belt-A	0.00292	0.01989	0.7334	-0.01405	0.3667
Carolina Slate Belt-B	0.00047	0.01054	0.9264	-0.00961	0.4632
Northern Outer Piedmont	0.00164	0.01377	0.7878	-0.01049	0.3939
Triassic Basin	-0.00196	0.00519	0.5870	-0.00911	0.7065
Rolling Coastal Plain	0.00046	0.00341	0.7570	-0.00249	0.3785
Eastern Blue Ridge Foothills	0.05502	0.08007	<.0001	0.02998	<.0001

Table 13: Drainage Area Means Between Intermittent and Perennial Origins

Ecoregion	Difference	Upper CL Difference	Prob > |t|	Lower CL Difference	Prob > t
Carolina Slate Belt-A	12.54230	20.00650	0.0014	5.07820	0.0007
Carolina Slate Belt-B	9.98700	49.87300	0.6168	-29.89900	0.3084
Northern Outer Piedmont	7.79900	15.73000	0.0538	-0.13200	0.0269
Triassic Basin	5.28924	8.93516	0.0058	1.64331	0.0029
Rolling Coastal Plain	54.93070	81.90140	0.0001	27.95990	<.0001
Eastern Blue Ridge Foothills	0.11130	1.26820	0.8490	-1.04560	0.4245

The average slope may not vary enough between intermittent and perennial stream origins to be effective in predicting location, except for the mountain sites. Conversely, contributing drainage area may help in determining origin locations in all Ecoregions except for mountain sites, and geologically weak areas in the Carolina Slate Belt. This finding provides additional insight into landscape processes that influence intermittent and perennial stream origins.

Table 14: Distribution of Intermittent and Perennial Origin Contributing Drainage Area (acres)

	Carolina Slate Belt-A		Carolina Slate Belt-B		Eastern Blue Ridge Foothills		Northern Outer Piedmont		Rolling Coast Plain		Triassic Basin	
	int	per	int	per	int	per	int	per	int	per	int	per
Min	0.20	0.72	0.05	2.04	0.23	0.24	1.55	2.54	0.16	7.16	0.10	0.13
10%	1.47	7.53	0.77	2.39	2.17	1.02	1.80	4.07	7.52	10.76	1.24	1.89
25%	2.85	11.58	4.89	9.52	3.72	2.91	4.48	10.05	11.15	28.82	1.95	3.27
50%	7.36	15.99	23.80	37.50	4.60	4.98	8.82	16.18	25.67	84.00	3.70	6.85
Mean	11.20	23.74	50.86	60.85	5.16	5.27	12.72	20.52	40.66	95.59	5.11	10.40
75%	14.47	35.40	69.96	68.16	6.34	7.04	15.06	27.11	55.15	122.00	7.16	15.79
90%	27.39	43.33	142.41	187.26	8.16	9.81	22.99	41.31	101.33	217.34	11.87	27.80
Max	74.63	107.00	322.27	328.28	14.60	15.85	115.95	64.81	173.65	343.66	16.51	32.49

The product of average slope and contributing drainage area (Figure 11 & Table 15) is an index that serves as a surrogate for erosion or energy potential of overland flow or stream flow. In the context of stream origins, it represents the interaction of drainage area and slope and the potential to sufficiently incise through the soil profile to intersect with the groundwater. The index was examined for its predictive strength in delineating intermittent and perennial origins.

Statistically, the slope-area product is significantly different between intermittent and perennial origins in all Ecoregions except Slate Belt – B and Blue Ridge Foothills. The t-test result is similar to the t-test result for drainage suggesting that drainage area is the strongest individual predictive variable. However, the slope-area product yielded higher statistical confidence and lower p-values for significant differences as well as boosting the confidence for Northern Outer Piedmont intermittent and perennial origins.

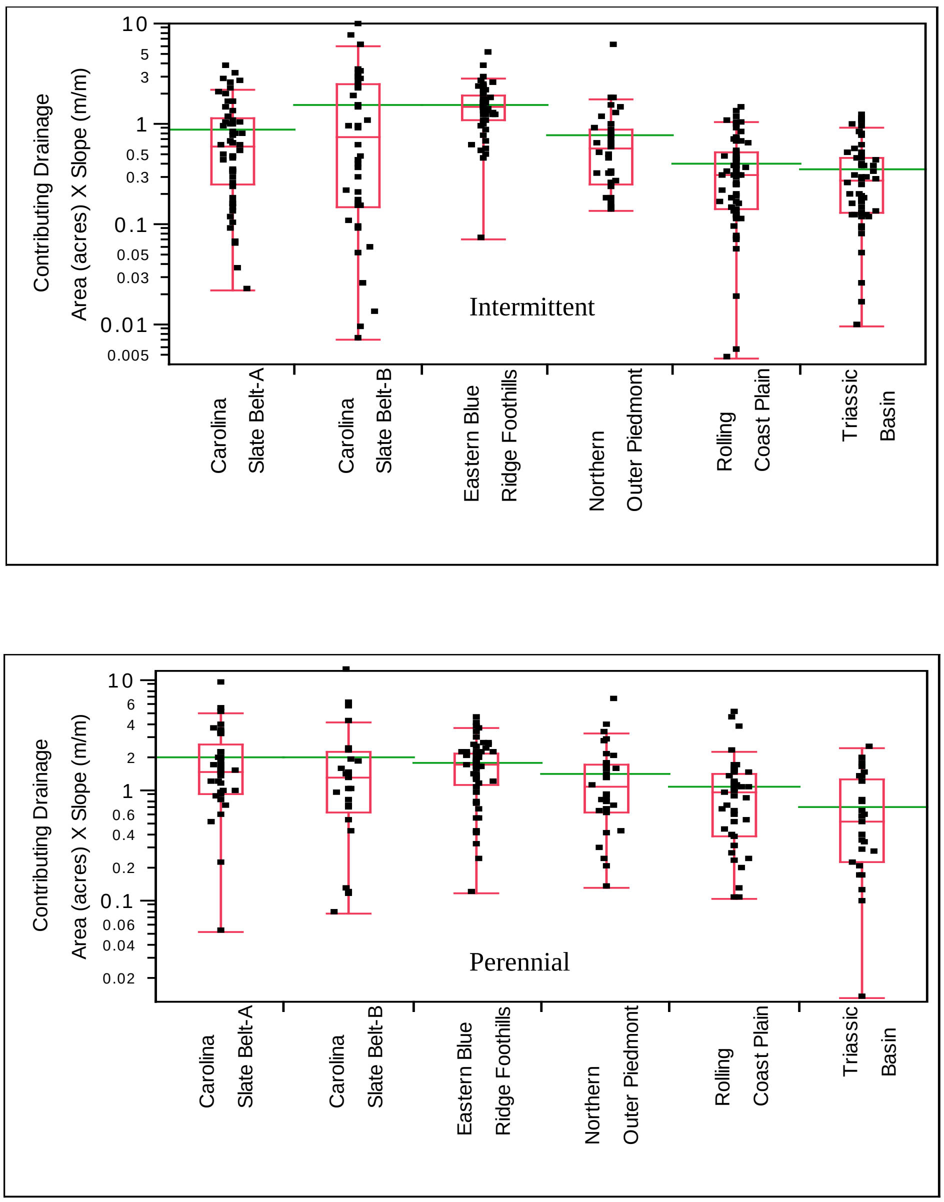

Figure 11: Distribution of Slope-Area Index

Table 15: Avg Slope X Contributing Drainage Area Means Between Intermittent and Perennial Origins

Ecoregion	Difference	Upper CL Difference	Prob > |t|	Lower CL Difference	Prob > t
Carolina Slate Belt-A	1.12232	1.74708	0.0007	0.49757	0.0004
Carolina Slate Belt-B	0.39840	1.58350	0.5020	-0.78680	0.2510
Northern Outer Piedmont	0.64797	1.19217	0.0204	0.10377	0.0102
Triassic Basin	0.36205	0.61687	0.0067	0.10723	0.0034
Rolling Coastal Plain	0.68537	1.01258	0.0001	0.35816	<.0001
Eastern Blue Ridge Foothills	0.19554	0.54370	0.2678	-0.15262	0.1339

The descriptions and analyses of stream origin data provide information regarding influences on stream origin and flow duration within selected regions of North Carolina. Additional data collection is planned and analysis of the stream origin data are far from complete. Although more work in needed, information derived to date illustrates the influence of landscape on the complex processes that govern the flow duration of streams and their origins. While drainage area and slope are commonly used surrogates for landscape process, they represent only a fraction of determinants that require investigation.

CONCLUSIONS and FUTURE WORK

The effort by NCSU and DWQ documented in this report demonstrates the feasibility of accurately mapping headwater stream networks in North Carolina. While additional work is needed, each objective was met with varying results. The use and application of logistic regression is effective in predicting presence/absence valleys with streams and stream length. The current version of the logit model is inefficient in predicting flow duration, i.e., intermittent verses perennial flow, but further exploration of the model with additional variables may provide more favorable results. Headwater stream mapping can be improved with greater understanding of landscape characteristics and processes, additional independent variable analysis and integration into the model.

One unexpected, but positive outcome of the headwater stream mapping effort has been the opportunity to transfer knowledge and technology between agencies. GPS and GIS methods and protocols developed during the different phases of this project have been shared with various government entities tasked with regulation of surface water features. GPS training was coordinated and a stream mapping protocol was developed to facilitate DWQ and local governments in mapping stream origins. DWQ transportation permitting, 401 certification and program development staff were trained in the use of the GPS protocol thereby facilitating the collection of stream origin data and its rapid dissemination via digital means and storage in a medium that allows geostatistical analysis by interested users.

Additionally, GPS and GIS applications used for stream delineation have been adopted by several consultants, and federal and state (in and out of state) regulatory staff, including use for several jurisdictional determinations. The "mobile" GIS application developed by Tom Colson for use in this study is available for download from ESRI web page and is voluntarily maintained by Colson under the supervision of DWQ. The application has been updated on a regular basis in response to "lessons learned" and knowledge gained by working on headwater stream mapping. To date over 700 users have downloaded this application. Additionally, an

ArcGIS tool developed by Colson for interpolating the LiDAR DEMs and for preparing the DEMs for statistical models is also available for download at ESRI.

Finally, analysis of headwater streams and origins characteristics (based on the field data) has provided guidance in North Carolina coastal stream restoration policy, and in state and nationwide Rapanos-Carabell decision procedure and policy adjustments by EPA and USACOE.

Additional Data and Analytical Needs

Incorporation of soil attributes, specifically, hydraulic and groundwater properties, has the greatest potential for accurately predicting flow duration. Precipitation, watershed shape, wetness and roughness indices may also reduce error and provide more accurate results. Expertise from the soil, hydrogeologic and hydrologic sciences is being sought to provide specific information regarding the most applicable attributes of each discipline to test. Also, current field data cover only a small portion of the state, so more data are needed for an adequate statewide sample.

Further statistical and spatial analysis of terrain derivatives, as well as analysis of additional data are required to determine the dominant influences on stream origin location, stream length and flow duration. Sample size, landscape division or grouping, and the intricacies of logistic regression also need further exploration.

GIS/staff/resources

The timely success of further stream mapping work is subject to additional dedicated staff resources. Building geodatabases, processing DEMs, and populating databases with external data requires considerable time and skill. Assistance in GIS data production by DOT staff will allow for greater progress in a shorter timeframe.

Long-Term And Short-Term Plans

Results of the pilot study are reported herein and indicate the effectiveness of the modeling approach. Given the acquisition of more data and support for processing that data, short-term and long-term plans may be initiated.

<u>Short-Term</u>

In the next six months to one year (with GIS support from DOT staff), the following activities will be completed to expand the DWQ stream mapping project.

- Develop a comprehensive inventory of additional landscape factors, indices and surrogates with the potential for influencing headwater stream origin locations, length and flow duration.
- Populate stream mapping geodatabase with those additional data.
- Conduct statistical and other quantitative analyses of ‘new’ data to determine the relative contribution and applicability of each variable.
- Improve Slate Belt logistic regression model based on the results of the above analysis.
- Develop models for remaining Ecoregions where data were collected during the pilot phase (Rolling Costal Pain, Eastern Blue Ridge Foothills, Northern Outer Piedmont).

<u>Long-Term (1year +)</u>

Future direction and success of the headwater stream mapping program requires the development of complimentary procedures and standards to fully integrate headwater streams with the CGIA stream map. With additional funding, this work will expand the pilot study into a statewide mapping product.

Long-term goals include the following:

- Establish acceptable horizontal accuracy standards for modeled streams and origins. This effort will be coordinated with DOT, DWQ and CGIA personnel.
- Develop a web-based spatial database to track and maintain in-field stream identification calls.
- Develop business and technical methodologies and accuracy standards for updating modeled headwater stream data with field verified or field identified data.
- Develop business and technical methodologies for updating CGIA NC stream map based on CGIA data maintenance plan.
- Collect data to represent the remaining areas of the state not included in the pilot project.
- Apply model (s) to the remaining Ecoregions in North Carolina.

<u>Establish acceptable horizontal accuracy standards for modeled streams and origins. This effort has been coordinated with DWQ and CGIA personnel.</u>

Horizontal accuracy of stream lines and origin points analyzed using field data and modeled data will be determined. Based on the accuracy analysis, all participating agencies must establish a minimum horizontal distance error for stream lines and a minimum radial distance error for origin points that will be acceptable for all future mapping. These standards must be established in accordance with CGIA/USGS accuracy standards to conform to state and federal GIS accuracy policy.

<u>Continue to collect data and develop models for the remaining areas in North Carolina.</u>

Improvement of initial model and potential development of additional models will continue. DWQ Priority areas (Table 13) for headwater mapping need to be coordinated with the CGIA steam mapping areas.

Table 13: DWQ Priority Areas for Headwater Stream Mapping

Ecoregion Name Level III	Ecoregion Name Level IV	Ecoregion Number	Priority
Outer Coastal Plain			
	Pamlico Lowlands	63b	I
	Nonriverine Swamps	63c	I
	Barrier Islands	63d, 63g	III
	Mid-Atlantic Flatwoods	63e	I
	Carolina Flatwoods	63h	Scheduled
	Floodplains and Terraces	63n	I
Inner Coastal Plain			
	Sandhills	65c	III
	Atlantic Loamy Plains	65l	III
	Rolling Coastal Plain	65m	Complete
	Floodplains and Terraces	65p	I
Piedmont			
	Southern Inner	45a	I
	Southern Outer	45b	I
	Slate Belt	45c	Complete
	Northern Inner	45e	Complete
	Northern Outer	45f	I
	Triassic	45g	Partially Complete
	Kings Mountain	45i	Scheduled
Mountains			
	New River	66c	II
	Southern Crystalline	66d	II
	Southern Sedimentary Ridges	66e	II
	Southern Metasedimentary	66g	II
	High Mountains	66i	II
	Broad Basins	66j	Need more
	Foothills	66l	Complete
	Amphibolite Mountains	66k	II
	Sauratown Mountains	66m	II

Develop a web-based spatial database to track and maintain in-field stream identification calls.

North Carolina Division of Water Quality proposes to develop and implement a WEB-Based Spatial Database for Perennial and Intermittent Stream Origins. A detailed proposal has been written and funding is being sought for this effort. This system can be used by public agencies and private industry for wetland permit applications, buffer rule implementation, basinwide studies, water quality [305(b)] reporting, stream restoration site search and planning, analysis of ecosystem resources and impacts and development project planning and design.

The implementation of a web-based perennial and intermittent stream origin spatial database will provide a centrally located and maintained system for use by public agencies and private industry in North Carolina as well as outside the state. A long-term benefit includes improved performance in permitting throughout DENR, DOT and other state, federal, local and private entities by streamlining current processes as well as by providing universal access to stream data.

Other short-term and long-term benefits are:

- Quality control for data
- Allow prompt and complete data entry
- Prevention of duplicate stream identification data
- Immediate dispersal and access of data
- Streamlining of permit processes where streams are involved
- Provide more accurate information for permit applications, e.g., stream length and buffer impact estimates
- Facilitate automated update and maintenance of stream data for the CGIA stream map
- Improvement to stream ID methodology by adding to perennial and intermittent stream data.
- Additional data will allow consistent evaluation of perennial and intermittent streams and origin thresholds and determination parameters.

- Link to Basinwide Information Management System (BIMS)
- Link to the Corps of Engineers and EPA for reporting and tracking purposes

The development of a web-based stream identification database requires the acquisition of an individual with expertise in SQL Server, ARCGIS, and web-based applications to develop the stream identification application. DWQ proposes to establish a technical advisory group of federal, state and local agencies to assist in the development of this database. At present, a data dictionary and geodatabase for stream delineations is being maintained by DWQ stream mapping staff. This geodatabase will serve as a starting point for program development and will be revised according to web-based field, form and function requirements. The system will be designed to allow expansion of utility for future needs, such as wetland and stream functional assessment data and possibly wetland delineation data.

<u>Establish business and technical methodologies and accuracy standards for updating modeled headwater stream data with field verified or field identified data.</u>

The continued field identification of intermittent and perennial streams and origin by certified individuals will be entered and tracked with the web-based spatial stream database. These data are inherently more accurate than modeled data since stream identification and origin is determined on the ground. For this reason, these field data should be used to update the modeled data whenever possible. Standard update methods will be developed for both technical and business procedures based on the accuracy of the field and modeled data. All procedures will be developed in conjunction with DWQ and CGIA standards.

Develop business and technical methodologies for updating CGIA NC stream map based on CGIA data maintenance plan.

A critical piece of the CGIA NC stream mapping project is the creation of a maintenance plan to effectively manage the continued update and maintenance of the CGIA stream map. It is within the CGIA maintenance plan that headwater streams will be addressed and included in the process. A conceptual flow of processes was developed by DWQ (Figure 13) to illustrate the basic components required for successful transfer of stream data. Critical processes include the acceptance criteria and approval of data based on CGIA guidelines prior to updating the CGIA stream map. Core business and technical issues should be addressed within these two processes to minimize transfer and delivery problems and end-product updates to the CGIA stream map. The methodologies will be developed within the existing CGIA technical advisory committee task framework.

Business and Logistical Requirements

Additional requirements for successful completion and implementation of the DWQ headwater stream project include training in GPS data acquisition and data transfer, training for web-based spatial database, assignment of work, determination of data storage, determination of system administration, continued coordination with CGIA and acquisition of additional funding. The overall goal for DWQ is to have a fully integrated, accessible and accurate stream mapping product for use by multiple agencies, industry and the public.

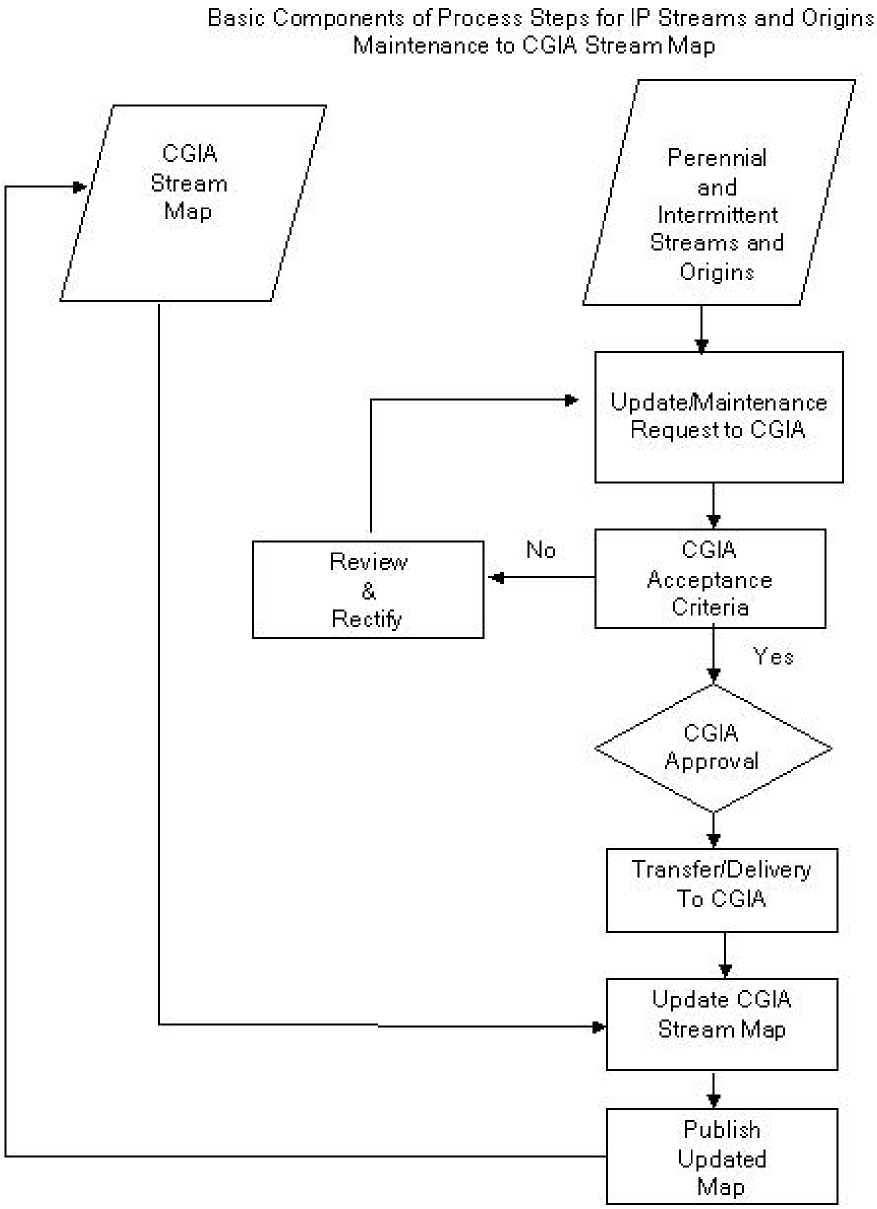

Figure 13: Conceptual Process Flow for Update and Maintenance of headwater streams to NC CGIA

REFERENCES

"Accuracy." *Merriam-Webster Online Dictionary*. 2007. http://www.m-w.com (4 Jan. 2008).

Coats, D. R., 1972. Quantitative geomorphology of small drainage basins in Southern Indiana. Tech. Rep. 10, Office of Naval Research, Department of Geography, Columbia University.

Colson, T. P. 2006. Stream Network Delineation from High Resolution Digital Elevation Models. Ph.D. Dissertation, Dept. of Forestry and Environmental Resources, NC State University, Raleigh, NC.

Colson, T.P. and Gregory, James D., unpublished 2007. Predicting Locations and Origins of First Order Streams in LIDAR DEMs Using a Logistic Regression Modeling Approach: A Preliminary Study in the Carolina Slate Belt. Department of Forestry and Environmental Resources, North Carolina State University. Contract Research Report to: Wetlands and Stormwater Branch North Carolina Division of Water Quality.

Dietrich, W., Reneau, S., Reneau, C., 1987. Overview: 'zero-order basins' and problems of drainage density, sediment transport and hillslope morphology. In: Erosion and Sedimentation in the Pacific Rim. International Association of Hydrological Sciences, Corvallis, Or, pp. 27–37.

Firman, J. C., Jacobs, S. E., 2002. Comparison of stream reach lengths measured in the field and from maps. North American Journal of Fisheries Management 22 (4), 1325–1328.

Garbrecht, J., Ogden, F., DeBarry, P., Maidment, D., 2001. GIS and distributed watershed models. I: Data coverages and sources. Journal of Hydrologic Engineering 6 (6), 506–514.

Garg, P.K and Harrison,A.R. 1990. Quantitative Representation of Land Surface Morphology from Digital Elevation Models, *Proc. 4th Int. Symp. on Spatial Data Handling* (1990), pp. 273-284.

Griffith, G.E., J.M. Omernik, J.A. Comstock, M.P. Shafale, W.H. McNab, D.R. Lenat, J.B. Glover, and V.B. Shelburne. 2002. Ecoregions of North Carolina and South Carolina. (2 sided color poster with map, descriptive text, summary tables, and photographs). U.S. Geological Survey, Reston, VA. Scale 1:1,500,000.

Grimm, L.G. & Yarnold, P.R. eds. (1995). Reading and Understanding Multivariate Statistics. Washington D.C.: American Psychological Association.

Hansen, W. F., 2001. Identifying stream types and management implications. Forest Ecology and Management 143 (1-3), 39–46.

Heine, R. A., Lant, C. L., Sengupta, R. R., 2004. Development and comparison of approaches for 71 automated mapping of stream channel networks. Annals of the Association of American Geographers 94 (3), 477–490.

Kienzle, Stefan (2004). The Effect of DEM Raster Resolution on First Order, Second Order and Compound Terrain Derivatives. Transactions in GIS 8 (1), 83–111.

Kiffney, P. M., Richardson, J. S., Feller, M. C., 2000. Fluvial and epilithic organic matter dynamics in headwater streams of Southwestern British Columbia, Canada. Archiv Fur Hydrobiologie 149 (1), 109–129.

Lindsay, J. B., Creed, I. F., 2005. Removal of artifact depressions from digital elevation models: towards a minimum impact approach. Hydrological Processes 19 (16), 3113–3126.

Morisawa, M., 1957. Accuracy of determination of stream lengths from topographic maps. Transactions, American Geophysical Union 38, 86–88.

McGlynn, B. L., Seibert, J., 2002. Distributed assessment of contributing area and riparian buffering along stream networks. Water Resources Research 39 (4), 2(1)–2(7).

Montgomery, D. R., Dietrich, W. E., 1989. Source areas, drainage density, and channel initiation. Water Resources Research 25 (8), 1907–1918.

North Carolina Division of Water Quality(NCDWQ), 2005. Identification methods for the origins of intermittent and perennial streams, version 3.1[Online]. Available at http://h2o.enr.state.nc.us/ncwetlands/documents/NC Stream ID Manual.pdf. Raleigh, NC.

North Carolina Floodplain Mapping (NCFMP), 2003. LiDAR and digital elevation data[Online], N.C. Division of Emergency Management, Available at http://www.ncfloodmaps.com/pubdocs/lidar final jan03.pdf. Raleigh, NC.

Mitasova, H., and Hoerka, Jaroslav.,1993. Interpolation by Regularized Spline with Tension: II. Application to Terrain Modeling and Surface Geometry Analysis: Math. Geol., v. 25, p. 657-669.

Morisawa, M., 1957. Accuracy of determination of stream lengths from topographics maps. Transactions, American Geophysical Union 38, 86–88.

Paybins, K., 2002. Flow origin, drainage area, and hydrologic characteristics for headwater streams in the mountaintop coal-mining region of southwest Virginia, 2000-01. Water-Resources Investigations Report 02-4300, U.S. Geological Survey.

Peterson, B. J., Wollheim, W. M., Mulholland, P. J., Webster, J. R., Meyer, J. L., Tank, J. L., Marti, E., Bowden, W. B., Valett, H. M., Hershey, A. E., McDowell, W. H., Dodds, W. K., Hamilton, S. K., Gregory, S.,Morrall, D. D., 2001. Control of nitrogen export from watersheds by headwater streams. Science 292 (5514), 86–90.

Schoonjans, Frank 2007. MedCalc Software Manual, Broekstraat 52, 9030 Mariakerke, Belgium. Version 9.3.8.0 - © - Last modified: 25 September 2007 http://www.medcalc.be/manual/logistic_regression.php

Stanford, J., 1996. Landscapes and catchment basins. In: Hauer, F.R. and Lamberti., G.A. (Eds.) Methods in Stream Ecology. Academic Press, San Diego, pp. 3–22.

Tabachnick, B.G. and Fidell, L.S. (1996). Using Multivariate Statistics. NY: HarperCollins.

Tarboton, D. G., Bras, R. L., Rodriguez-Iturbe, I., 1992. A physical basis for drainage density. Geomorphology 5 (1-2), 59.

Vannote, R. L., Minshall, G.W., Cummins, K.W., Sedell, J. R., Cushing, C. E., 1980. River continuum concept. Canadian Journal Of Fisheries And Aquatic Sciences 37 (1), 130–137.

Tucker, G. E., Catani, F., Rinaldo, A., Bras, R. L., 2001. Statistical analysis of drainage density from digital terrain data. Geomorphology 36 (3-4), 187.

Vogt, J., Colombo, R., Bertolo, F., 2003. Deriving drainage networks and catchment boundaries: a new methodology combining digital elevation data and environmental characteristics. Geomorphology 53 (3), 281–298.

Weaver, Curtis J. 1998. Low Flow Characteristics and Discharge Profiles for Selected Streams in the Neuse River Basin, North Carolina. USGS Water Resources Investigation Report 98-4135.