SIMULATING SURFACE AND GROUND WATER FLUXES IN COMPLEX KARST SYSTEMS – ANALYSIS OF MIDDLE AND EAST TENNESSEE'S SEQUATCHIE VALLEY USING THE ANNAGNPS WATERSHED POLLUTANT LOADING MODEL

By

Benjamyn Savard

APPROVED:

Graduate Committee:

Dr. Henrique G. Momm (Geosciences)

Dr. Racha El Kadiri (Geosciences)

Dr. Mark Abolins (Geosciences)

Dr. Henrique G. Momm, Chairperson Department of Geosciences

Dr. David L. Butler, Dean of the College of Graduate Studies

SIMULATING SURFACE AND GROUND WATER FLUXES IN COMPLEX KARST SYSTEMS – ANALYSIS OF MIDDLE AND EAST TENNESSEE'S SEQUATCHIE VALLEY USING THE ANNAGNPS WATERSHED POLLUTANT LOADING MODEL

By

Benjamyn Savard

A thesis submitted in partial fulfillment

of the requirements for the degree of

MASTER OF SCIENCE

in

Geosciences

Middle Tennessee State University

August 2023

DEDICATION

I would like to dedicate this thesis to my wife Jennifer and family & friends who have always encouraged my decision to pursue a degree following a career in the United States Navy. Your unwavering support then and now will always be cherished.

ACKNOWLEDGEMENTS

There are many individuals who in their invaluable support led to the completion of this work, for that I give many thanks. First and foremost, I want to say thank you to my advisor, Dr. Henrique G. Momm, whose enthusiastic support, mild manner, and countless hours of providing thoughtful insight and edits to this thesis is admirable. Your "open door policy" and commitment to help those around you is an infectious trait, even though you recent office move will not allow you stay in the background as your desire. To my committee members, Dr. Racha El Kadiri and Dr. Mark Abolins, I appreciate your devoted time and support to this research and thank you for your indispensable knowledge, many inputs, and review of this manuscript. Thanks to Ben Miller for sharing his insight on the Sequatchie Valley and vast experience in hydrology and dye trace investigations, your assistance provided a clear path towards completion of this work. I would like to thank Dr. Albert Ogden for his warm hospitality and availability to sit and chat regarding his past work and tips for moving forward with my research. Much appreciation to the faculty and staff of Middle Tennessee State University, especially the numerous meetings with Katy Moore to help explain and provide guidance for the hydrological modeling processes used in this study. Many thanks to Karen Wolfe who always offers a warm shoulder and a good ear; you are the glue that holds the department together. To Chuck Sutherland and Emily Davis of the Tennessee Cave Survey for providing the cave data that became the foundation of this thesis. Special thanks to the helpful staff of the James E. Walker campus library and Tennessee State Library & Archives in Nashville who allowed access to their many references and digital equipment which proved instrumental to the completion of my thesis. Finally, I would like to express my gratitude to the land owners who allowed open access to their property, especially

Jimmy Kemmer who took the time to walk his property with me and show the relevant karst features.

ABSTRACT

Topographic feature evaluation is a critical component for accurate interpretation and quantification of hydrologic flow in watersheds. In a typical drainage basin, topographic features identify and define surface drainage paths as flow traditionally travels through the most direct pathway. Alternatively, hydrology in karst terrain is often heterogeneous to surface topography and water rapidly redirects underground only to emerge elsewhere as discharge. This study evaluated the predictability of the Annualized Agricultural Non–Point Source (AnnAGNPS) model to simulate runoff over a six–year period (2016 to 2021) in the Sequatchie Valley and adjacent karst valleys located in the Cumberland Plateau of Middle and East Tennessee. After runoff calibration and statistical analysis, it was determined that utilizing the automated/semi–automated processes of AnnAGNPS produced results that indicate the model to be appropriate in simulating surface and surface–ground flow at annual–scales, although future work is recommended to improve monthly–scale results.

TABLE OF CONTENTS

LIST OF TABLES vii

LIST OF FIGURES ix

LIST OF SYMBOLS AND ABBREVIATIONS xii

Chapter

I. INTRODUCTION 1

II. SEQUATCHIE VALLEY 5

III. MATERIALS AND METHODS 18

IV. RESULTS AND DISCUSSION 65

V. CONCLUSIONS 72

BIBLIOGRAPHY 74

Appendix

A – TOPAGNPS & ANNAGNPS TECHNICAL PROCESSES 79

B – STUDY AREA PHOTOGRAPHS 95

C – IMAGES OF SURROUNDING KARST VALLEYS 101

LIST OF TABLES

Table 1. Study area relevant karst feature coordinates referenced from the Tennessee Cave Survey (TCS) in comparison to ground verified GPS coordinates of features. 19

Table 2. List of output files provided by TopAGNPS reduced–mode execution that are utilized in this study, including brief descriptions of their function. 30

Table 3. Summary of reach and cell characteristics following CSA and MSCL threshold assignment and TopAGNPS watershed segmentation and sub–catchment parametrization. 33

Table 4. Example of soil layer data section output displaying the number of soil layers, layer depth of soils, and the content of clay, silt, sand, and very fine sand for each unique soil ID and AnnAGNPS cell. 36

Table 5. Example of soil data section output displaying the unique soil ID and AnnAGNPS cell utilizing the NITA program and Python script, including the hydrologic soil group, soil erodibility factor (K factor) value, and soil texture description. 36

Table 6. Conversion of master category number (zonal majority) of land cover/land use into a representative letter to help determine the management sequence for each CDL year. 38

Table 7. Example of four sequences of operation illustrating the annual land use/land cover majority assigned to each grid code (sub–catchment) in the Sequatchie Subbasin. 41

Table 8. Example of annual farming operations associated with distinctive managements for the row crops of corn and soybeans in the Sequatchie Subbasin during 2016. 42

Table 9. Spatial location and elevation of the climate stations used in the study area. 44

Table 10. Summary of the drainage basin hydrologic surface areas for each 12–digit HUC within the Sequatchie Subbasin. 47

Table 11. Example of observed total streamflow measured in Mg and the corresponding annual time–series variation of alphas values. * Denotes removal of baseflow. 48

Table 12. Summary of the template folders and sub–content file inputs required for the master control file and initial simulation of AnnAGNPS. 50

Table 13. Incremental increase of Curve Numbers from initial values for each hydrological soil group and corresponding majority land use/land cover type in the Sequatchie Subbasin. 53

Table 14. Simulated annual time–series scenarios of total streamflow summaries created from the incremental increase of CN values during calibration of the Sequatchie Subbasin................ 54

Table 15. Summary of statistical metrics used to determine model performance for simulated and observed total streamflow (runoff) comparisons at annual and monthly time–scales. 57

Table 16. Initial and final CN values following calibration for each land use/land cover majority in the Sequatchie Subbasin. 61

Table 17. Statistical parameters of model performance for the final calibrated Sequatchie Subbasin summary file when compared to .800 gage file. 62

Table 18. Mean daily, monthly, annual, and total rainfall in mm at each drainage basin respective outlet during the study period (2016 to 2021). 66

Table 19. Output summaries of AnnAGNPS gaging station files for surface hydrological modeling. 68

Table 20. Comparison of statistical parameters in the calibrated surface and integrated (ground–surface) Sequatchie Subbasin and observed .800 gage files. 69

Table 21. Comparison of surface and ground–surface hydrology models mean total streamflow (Mg) and their percent change. 69

LIST OF FIGURES

Figure 1. Prominent structural features of the Cumberland Plateau located in the southern Appalachian Plateau Province of Middle Tennessee, from Wilson and Stearns (1958). 7

Figure 2. Block diagram illustrating development of the Sequatchie Valley from undisturbed sediments to present configuration due to karstification and headward valley erosion, modified from Miller (1979). 8

Figure 3. Northeastern section in Sequatchie Valley and key topographic divides of the surrounding karst valleys (coves), compiled from 1:24,000–scale USGS topographic map. 9

Figure 4. Illustrative relationship between subterranean stream invasion and slope retreat of geologic units along the western escarpment of the Cumberland Plateau, adapted from Crawford (1987). 11

Figure 5. Illustration of perceived subsurface drainage routes from swallets and resurgence at springs with overland flow of the Sequatchie River to the USGS watershed outlet. Refer to Figure 3 for key topographic divide locations. 12

Figure 6. Injection and detection point locations and their corresponding subsurface drainage referenced from the Tennessee Cave Survey (TCS) digital data subset. 14

Figure 7. Study area and climate station locations within the surrounding counties of Middle and East Tennessee, USA. See Table 9 for names and locations of climate stations. 16

Figure 8. Photographs of relevant karst features taken by author during December 2022 field investigations. (A) Head of Sequatchie Spring in Cumberland Plateau State Park and Mill Cave entrance in Grassy Cove (B), Cumberland County, Tennessee (Photo B used with permission). 19

Figure 9. Conceptual watershed model of surface flow and subsurface drainage routes in the Sequatchie Subbasin and interconnected karst valleys. 21

Figure 10. Internal representation of surface flow connections (A) and watershed characterization (B) between the basic modeling unit process of sub–catchment (cell) and concentrated flow (reach) simulated by AnnAGNPS. 25

Figure 11. Schematic diagram of system components required for watershed modeling represented by input and output spatial datasets (light– and dark–green), input (light–yellow) and output (dark–yellow) tabular datasets, external databases (white), AGNPS tools (dark–blue), and required primary inputs for AnnAGNPS (light blue). 28

Figure 12. Identification of the Sequatchie Subbasin outlet cell along the flow vector (white line) using the UPAREA raster grid and location coordinates in GIS (A) and the no–data pixel value indicating the outside of the drainage basin boundary (B). 32

Figure 13. Identification of the Grassy Cove outlet location by editing the ASCII DEM raster elevation from the original meter value to a -9999 no–data value (A) and pixel result represented on the raster layer in GIS (B). 32

Figure 14. GIS layers for the Sequatchie Subbasin 9–m resolution DEM (A), channel network of drainage basin with prominent spring and outlet locations (B), and sub–catchments representing the watershed division into AnnAGNPS cells (C). 34

Figure 15. GIS layers of the Sequatchie Subbasin original soil type layer displaying state and county codes (A), the soil type and soil layer merged into unique soil IDs (B), and soil zonal statistic majorities discretized to the AnnAGNPS cells (C). 37

Figure 16. Assignment of the 2016 (A), 2019 (B), and 2021 (C) Crop Data Layers (CDLs) to the Sequatchie Subbasin raster boundary. 39

Figure 17. Crop majorities discretized to the Sequatchie Subbasin sub–catchments displaying progression of crop majorities from 2016 (A), 2019 (B), and 2021 (C). A majority value of 0 represents background. 40

Figure 18. Assignment of climate station zones of influence to the Sequatchie Subbasin and surrounding karst valleys based on Thiessen polygon analysis. 45

Figure 19. Map of the delineated Sequatchie Subbasin in order to capture the Sequatchie River tributary systems. 46

Figure 20. Hydrograph of observed total streamflow measured in Mg over the six–year study period (2016 to 2021). ***** Denotes removal of baseflow. 48

Figure 21. Comparison of monthly simulated and observed (gage) total streamflow (Mg) with varying alpha values during the study period (2016 to 2021). Simulated scenarios of initial CN value (dark–blue line), CN +3 (red line), CN +5 (green line), CN +7 (orange line), and observed (light–blue line) .300 gage (A), simulated scenarios and observed (.500 gage) (B), simulated scenarios and observed (.800 gage) (C), and simulated scenarios and observed (.975 gage) (D).55

Figure 22. Comparison between the CN +5 simulated and .800 gage observed total streamflow (Mg) values at annual (A) and monthly (B) time–scales. The obtained R^2 values were .86 and .65 respectively, explaining 86% and 65% of data variance between observed and the model. 59

Figure 23. Time–series of annual (A) and monthly (B) CN +5 compared with observed .800 gage summary files of total streamflow (Mg) from January 2016 through December 2021. 62

Figure 24. Conversion of Grassy Cove (A) and Sequatchie Subbasin (B) reach raster grids into GIS layers in order to identify the stream code IDs that correspond to each input reach. 64

Figure 25. Monthly time–scale simulated rainfall (mm) summaries of the Sequatchie Subbasin (A), Grassy Cove, Little Cove, McClough Hollow, and Bat Town Cove (B), and 67

Figure 26. Comparison between the integrated and .800 gage observed total streamflow (Mg) values at annual (A) and monthly (B) time–scales. The obtained R^2 values again explain 86% and 65% of data variance between observed and the model. 70

Figure 27. Time–series of annual (A) and monthly (B) calibrated surface and integrated Sequatchie Subbasin compared with observed .800 gage summary files of total streamflow (Mg) from January 2016 through December 2021. 70

LIST OF SYMBOLS AND ABBREVIATIONS

% – percent

º – degrees

~ – approximate value

∞ – infinity

-- – no–data

∑ – sum

$\bar{O}$ – observed mean

$\bar{P}$ – predicted mean

AgGEM – Generation of Weather Elements for Multiple Applications

AGNPS – Agricultural Non–Point Source

AMC – antecedent moisture condition

AnnAGNPS – Annualized Agricultural Non–Point Source

ASCII – American Standard Code for Information Interchange

BMP – Best Management Practice

C – Celsius

CDL – Crop Data Layer

CN – Curve Number

CSA – critical source area

csv – comma separated value

D8 – deterministic eight

DEDNM – Digital Elevation Drainage Network Model

DEM – Digital Elevation Model

ft – feet

ft^3/s – cubic feet per second

GIS – Geographic Information System

GPS – Global Positioning System

ha – hectare

HUC – Hydrologic Unit Code

HUSLE – Hydro–geomorphic Universal Soil Loss Equation

i – incremental duration of time

KHI – Karst Hydrological Inventory

km – kilometer

km^2 – kilometer squared

LiDAR – light detection and radar

m – meter

m^3/d – cubic meter per day

Mg – megagram

mm – millimeter

MSCL – minimum source channel length

N – nitrogen

n – number of observations

NITA – NASIS Import to AnnAGNPS

NOAA – National Oceanic and Atmospheric Association

NPS – Non–Point Source

NRCS – Natural Resources Conservation Service

NSE – Nash–Sutcliffe efficiency

O – observed data

P – phosphorous

P – predicted data

PBIAS – percent bias

PL – Pollutant Loading

r – Pearson’s product–moment correlation coefficient

R^2 – coefficient of determination

RASFOR – Raster Formatting

RASPRO – Raster Properties

RITA – RUSLE2 Import to AnnAGNPS

RMSE – root mean square error

RSR – root mean square error–observations standard deviation ratio

RUSLE – Revised Universal Soil Loss Equation

SCS – Soil Conservation Service

SD – standard deviation

SSURGO – Soil Survey Geographic

TCS – Tennessee Cave Survey

TMDL – Total Maximum Daily Loads

TN – Tennessee

TopAGNPS – Topographic Parameterization for Agricultural Non–Point Source

TOPAZ – Topographic Parameterization

USDA–ARS – United States Department of Agriculture–Agricultural Research Service

USGS – United States Geological Survey

UTM – Universal Transverse Mercator

CHAPTER I

INTRODUCTION

Topographic feature evaluation is a critical component for accurate interpretation and quantification of hydrologic flow in a watershed. In a typical drainage basin, topographic features identify the basin boundary while defining surface drainage paths and flow traditionally travels through the most direct pathway. Alternatively, hydrology in karst terrain is often heterogeneous to surface topography and water is rapidly redirected downward into a pipeline network of subsurface cracks, conduits, or caves that emerges elsewhere as discharge from springs, seeps, and wells.

Karst landscapes are classically characterized by temperate, subtropical, or tropical climates where rainfall varies from moderate to heavy, topography often lacks surface water with thin to patchy regolith, and thinly–bedded strata of non–porous, highly–fractured carbonate rocks are subject to the excavating effects of efficient groundwater flow. Development of karst occurs in less–soluble bedrock, such as marble, limestone, and dolomite, as slightly acidic rainwater interacts with soils and assumes additional carbon dioxide to form carbonic acid, a naturally occurring mild acid commonly found in groundwater (Palmer, 1990; Veni et al., 2001). As acidic groundwater circulates within rock passageways it dissolves calcite, the principal mineral of carbonates, to gradually alter small cracks and fractures into larger, unique features in downstream directions such as conduits, caves, losing streams, and sinkholes that interconnect the surface with the subsurface groundwater environment.

Since groundwater stores within rock pores and occupies bedrock voids in the phreatic zone, most subsurface flow and dissolution enlargements occur just below or at the water table, where fracture connectivity, dissolved openings, and permeability are greatest (Veni et al., 2001). Water entering a karst hydrological system as inputs (e.g., infiltration, losing streams, and sinkholes) is characterized as zones of recharge whereas water leaving the system through output features identify as zones of discharge (Miller, 2010). Current field methods like geochemical and discharge measurements, cave surveys, and dye trace investigations in karst environments greatly aid the interpretation and quantification of karst hydrologic processes, although karst research remains difficult due to complex flow paths, flow regime variety, and unpredictable groundwater behavior (Wolfe et al., 1997; Miller, 2010). Furthermore, there remains a lack of hydrological modeling tools that integrate the surface and subsurface connection of karst.

Some of the most extensive karst and cave systems in the United States are displayed in the Sequatchie Valley, a prominent 105 kilometer (km) linear gorge breaking the surface trace of the Cumberland Plateau just north of the Alabama–Georgia border in Middle Tennessee. The Sequatchie Valley developed along the axis of a thrust–faulted anticline in association with Valley and Ridge rock deformation during the Paleozoic uplift of the Southern Appalachians (Wilson and Stearns, 1958; Miller, 1979). This physiographic region was historically an anticlinal mountain that currently transitions northeast as an anticlinal valley due to karst landform development. At one time Middle and parts of East Tennessee were entirely capped by thick sequences of Pennsylvanian strata, now a geomorphic transition between valley and mountain topography occurs at the Head of Sequatchie Valley and northernmost portion of the Sequatchie Anticline (Milici, 1960). This transition is due to extensive downward and outward erosion resulting in a series of coves and closed depressions in earlier–stage karst valley retreat (Crawford, 1984; Miller and Ham, 2021). Today, the geomorphic agent of water continues to promote headward migration due to siliciclastic caprock removal and conduit cavern expansion caused by subsurface stream invasion along the western Cumberland Plateau escarpment, unlike the eroded fault escarpment along the eastern flank (Crawford, 1984; 1989).

The Cumberland Plateau is an expansive topographic feature extending 725 km from southern West Virginia to northern Alabama, varying from 65 to 80 km wide with general elevations of ~300 to 500 meters (m). Extraction of topographic information by traditional, manual techniques from such a large land surface would be impracticable for the scope of this research. In addition, manual techniques to obtain topographic information from large or complex watersheds like those in the Sequatchie Valley are time–consuming and tend to introduce subjectivity and/or possible human error. To mitigate this, the Topographic Parameterization for Agricultural Non–Point Source (TopAGNPS) software package will be utilized. TopAGNPS is an automated digital landscape analysis tool that analyzes high–quality Digital Elevation Models (DEM) to evaluate topography features and define drainage network and sub–watershed properties. TOPAGNPS interfaced with a Geographic Information System (GIS) can rapidly evaluate digital landscape elevations by performing topographic parametrization to generate the basic modeling units of cells and reaches required for automated watershed hydrologic analysis in the Annualized Agricultural Non–Point Source (AnnAGNPS) pollution model system.

AnnAGNPS is designed to predict long–term sediment–transported pollutant loadings and water quality responses at the agricultural watershed–scale. The analyzed watershed is subdivided into homogenous hydrologically–derived cells, or *sub–catchments*, in terms of user–generated data inputs of land use and farming management, soil hydrological properties, and continuous time–series of climate and weather (primarily temperature and precipitation). Analysis of land surfaces (cells) at any point within the watershed is capable since sub–

catchments have relatively uniform hydrological and physical characteristics. AnnAGNPS then simulates daily time–step constituent quantities of runoff, sediment, and pollutant loadings that are routed from each cell to the channel network over a given simulation period.

The purpose of this study is to conduct hydrological modeling of the Sequatchie Valley from the United States Geological Survey (USGS) Sequatchie–Whitwell streamgage to the Head of Sequatchie Spring and adjacent karst valley hydrologic systems of the Cumberland Plateau in Middle and East Tennessee. The hydrological modeling processes within the AnnAGNPS model are used to simulate surface water flow in the Sequatchie Valley subbasin combined with the influence of surrounding cove surface and subsurface flow contributions (Crawford, 1984; 1989; Miller and Ham, 2021). Ultimately, the objective of this research is to evaluate the AnnAGNPS model's capability to predict runoff from a series of interconnected karst systems that contribute to the total streamflow volume of the Sequatchie River over a six–year period (2016 to 2021).

<u>Hypothesis</u>*: The AnnAGNPS model can accurately predict runoff from surface and subsurface flow from the adjacent karst valleys that contribute to the total streamflow volume of the Sequatchie River.*

CHAPTER II

SEQUATCHIE VALLEY

Literature Review

Previous Investigations

Early geologic investigations in or adjacent to the Sequatchie Valley region prior to 1926 were primarily concerned with the general geology and economic & mineral resource (e.g., coal, iron ore, hematite, oil, and gas) possibilities (Milici, 1960). After 1926, investigations transitioned from general geology and economic concerns to stratigraphy and structural geology of the region. The first generalized section of formation exposures in the Sequatchie Valley was published by R. S. Bassler in 1932. Rich (1934) suggested that the Sequatchie Valley anticline resulted from the mechanics of overthrust faulting and low–angle thrusting. This was supported by new information provided by Rodgers (1950) and his consideration that outlying structures of the Cumberland overthrust block are analogous to the Sequatchie Valley structure. Stearns (1954) discussed the geology of the Crab Orchard Mountain area and demonstrated a northwestward movement of Cumberland Plateau's most southern portion along the overthrust, and Wilson and Stearns (1958) investigated the structural features of the Cumberland Plateau.

Milici (1960) extensively studied, mapped, and measured rock sections of the Sequatchie Valley overthrust block while Milici's (1963) additional interpretations of the Sequatchie anticline–Cumberland Plateau structural relationship differ from Wilson and Stearns (1958). Milici (1974) described strata of Upper Mississippian and Lower Pennsylvanian depositional environments in the southern Cumberland Plateau and Milici and Wedow (1977) measured sections and strata distribution of Upper Ordovician and Silurian depositional environments in the Sequatchie Valley. Milici and others (1979) discussed Carboniferous strata systems and coal–bearing strata in the Sequatchie Valley. Crawford (1984) investigated karst landform development processes and Pennsylvanian to Late Mississippian caprock removal due to groundwater along the retreating escarpments of the Cumberland Plateau and Crawford (1989) studied surface and subsurface hydrogeology of karst valleys in the Grassy Cove area using dye trace investigations. Most recently, Steinmann (2018) determined controls on speleogenesis of caves and karst processes in Upper Mississippian strata utilizing dye tracing, cave geomorphology, and GIS while Miller and Ham (2021) discussed Mississippian to Pennsylvanian unit stratigraphy and hydrogeology of major karst systems in the Cumberland Plateau and Head of Sequatchie Valley.

Geological History

The Sequatchie Valley is a linear deep gorge that cuts into the Cumberland Plateau physiographic region in Middle Tennessee. It lies in the southern portion of the Appalachian Plateau province and displays some of the most extensive karst and cave systems in the United States. The Sequatchie Valley developed along the axis of a thrust–faulted anticline (Rich, 1934) and presents as a continuation of mountain–building events associated with the deformation of the Valley and Ridge province (Wilson and Stearns, 1958; Milici, 1960; 1963) during the Allegheny orogeny of the Paleozioic Era (Miller, 1979). According to Wilson and Stearns (1958), three subprovinces divide the structure of the Cumberland Plateau: 1. a province area absent of faulting just northeast of the Sequatchie Valley anticline; 2. the thinly sheeted Pine Mountain overthrust extending northeast of the undisturbed area; and 3. the nearly–horizontal beds of the Cumberland Plateau overthrust sheet which is incised and folded by the Sequatchie Valley anticline (Figure 1), although Milici (1963) interprets the fault relationships of the Cumberland Plateau and Sequatchie Valley as one.

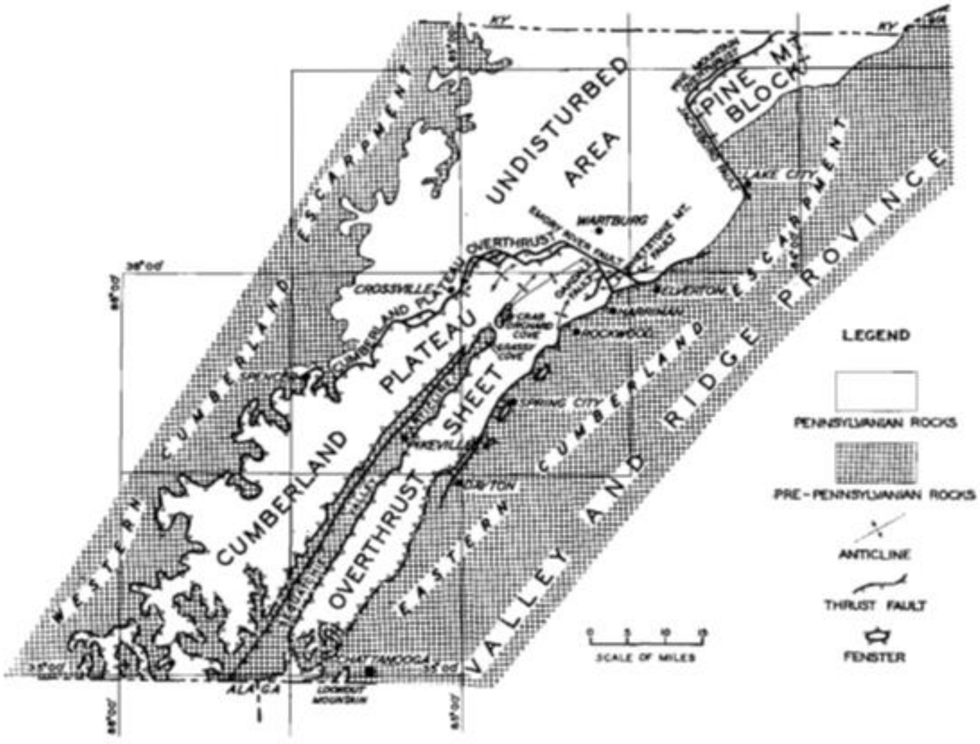


Figure 1. Prominent structural features of the Cumberland Plateau located in the southern Appalachian Plateau Province of Middle Tennessee, from Wilson and Stearns (1958).

The physiographic feature of the Sequatchie Valley was historically an anticlinal mountain that currently transitions northeast as an anticlinal valley due to karst landform development (Figure 2). Today, the geomorphic agent of groundwater continues to create and mature karst in this region due to Pennsylvanian–aged siliciclastic caprock removal, subsurface stream invasion of underlying carbonates causing gradual conduit cavern enlargement and expansion (Crawford, 1984; 1989), and retreating slopes of the Cumberland Plateau escarpment due to subsurface erosion (Milici, 1960; 1963; Miller, 1979; Crawford, 1984). Prominent Pennsylvanian caprock beds underlain by less resistant Mississippian carbonates mark the eastern and western escarpment limits that rim the Cumberland Plateau (Milici, 1963; Crawford,

1984) with Cambro–Ordovician aged dolomite bed exposure (Wilson and Stearns, 1958) and outcrops of the Knox Group (Brahana and Bradley, 1985) in the broad Sequatchie Valley.

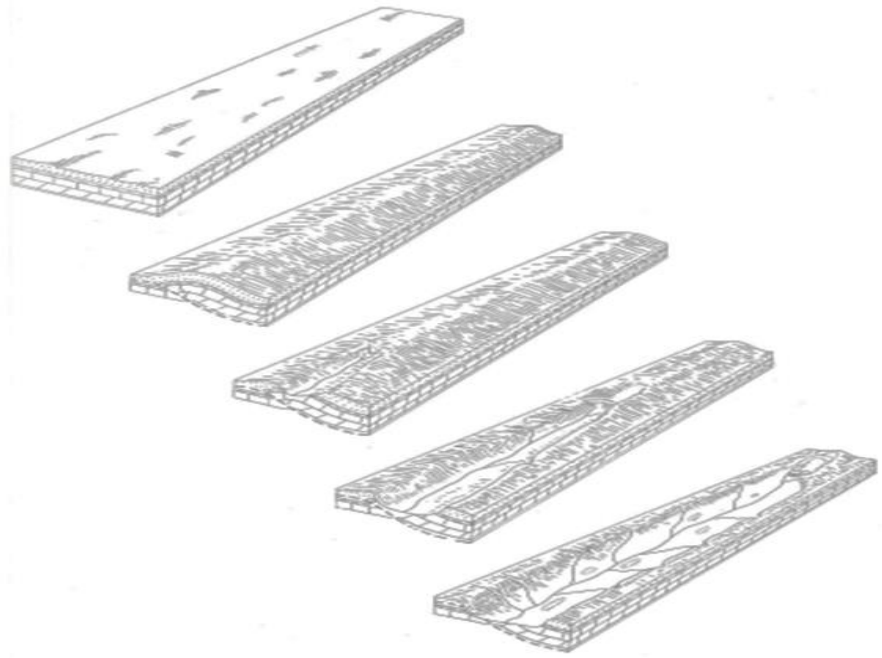


Figure 2. Block diagram illustrating development of the Sequatchie Valley from undisturbed sediments to present configuration due to karstification and headward valley erosion, modified from Miller (1979).

Erosional processes exposing underlying strata, regional structural features, and Paleozoic headward migration of the anticlinal axis have resulted in a local breach of Pennsylvanian strata to the northeast of the Sequatchie Valley and formation of two additional valleys, Crab Orchard Cove and Grassy Cove (Wilson and Stearns, 1958). Grassy Cove is one of North America’s largest closed depressions and represents the next significant portion of the Sequatchie Anticline. The natural drainage divide of Hinch Mountain separates the Head of Sequatchie Valley and Grassy Cove, where at the base of Dorton Knob discharges the Head of

Sequatchie Spring. North of Crab Orchard Cove the anticlinal mountains still exist along the northernmost part of the Sequatchie Anticline. In the transitional area between the anticlinal mountains of Crab Orchard and anticlinal valley at the Head of Sequatchie lies a series of karst valleys, locally known as *coves*, upgradient from Grassy Cove in earlier–stage valley retreat (Crawford, 1984). These include Little Cove, McClough Hollow, and Bat Town Cove, in addition to Swagerty Cove along the eastern flank of the Sequatchie Valley anticline axis, just south of Bear Den Mountain (Crawford, 1989; Miller and Ham, 2021; Figure 3).

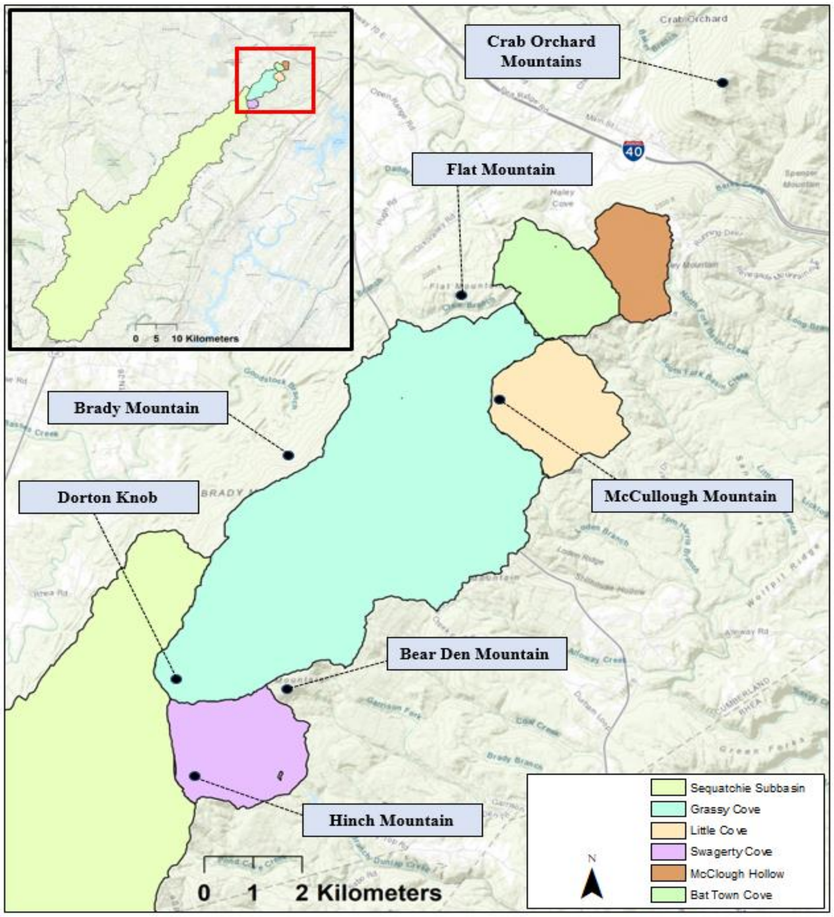


Figure 3. Northeastern section in Sequatchie Valley and key topographic divides of the surrounding karst valleys (coves), compiled from 1:24,000–scale USGS topographic map.

Karst Hydrogeology and Previous Dye Trace Investigations

In general, exposed strata within and adjacent to the Sequatchie Valley range in age from Cambro–Ordovician dolomites to Quaternary alluvium. The oldest exposure in the area occurs along the Sequatchie Anticline crest where portions of the Knox Group crop out (Milici, 1960). The Knox Group is a relatively undeformed thick sequence of limestone and dolomite where the upper 150 m is a deep aquifer under confined conditions (Brahana and Bradley, 1985). Nearly all groundwater in this aquifer system occurs in a dynamic flow system of small tubular voids with a potentiometric surface that gently slopes west towards the Tennessee River. Vertical leakage out or recharge in through the overlying Ordovician rocks is based on driller logs of the 1940's which indicate the aquifer responds to regional discharge and recharge but groundwater flow is not well defined (Brahana and Bradley, 1985).

At the Head of Sequatchie and Grassy Cove area the stratigraphy sequence ranges from Mississippian clastics and carbonates to Pennsylvanian caprock plateau clastics. Near the top of Brady Mountain is the Vandever Formation, transitioning down sequence along the cove flanks and plateau escarpments to the St. Louis Limestone. Uniquely, Grassy Cove differs from the rest of the nearly–horizontal geologic units of the Cumberland Plateau as a result of extensive local structural deformation due to faulting (Miller and Ham, 2021). Portions of the Sequatchie Valley sinkhole plain and floor of Grassy Cove partially extend down to the Fort Payne Formation but consist mainly of the Warsaw Formation. Along this geologic sequence there is a relationship between surface–flowing streams sinking into swallets then subsurface stream resurgence at springs, with caves being the primary conduits of discharge.

Both the Hartselle Formation and Warsaw Formation act as major confining layers in this sequence and control the structural elevation of stream caves in their respective overlying limestones, including the Bangor Limestone and St. Louis Limestone (Crawford, 1984). As surface streams flow from the caprock plateau they interact with carbonates through bedding planes and joints, gradually enlarging the subsurface conduit by abrasion, causing a downcut of the bed or erosion of a lower passage along dip (Crawford, 1984; 1987; Figure 4). It is believed the product of caprock plateau removal is left behind at the retreating escarpment base, extending outward for several kilometers as a sinkhole plain.

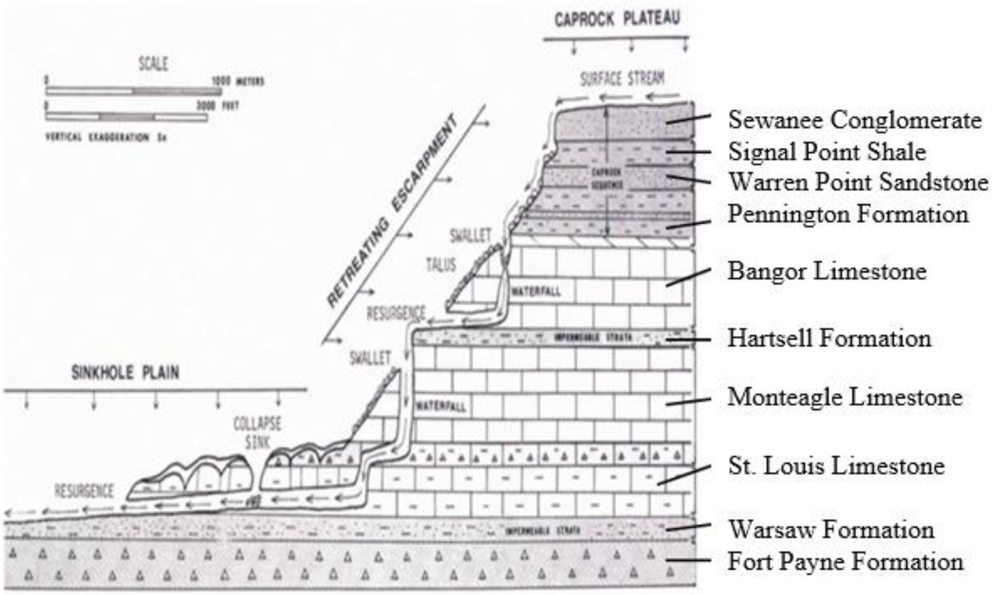


Figure 4. Illustrative relationship between subterranean stream invasion and slope retreat of geologic units along the western escarpment of the Cumberland Plateau, adapted from Crawford (1987).

Along Grassy Cove's interior circumference lie multiple small–scale karst systems with streams that rapidly sink into conduits to resurge from springs at lower elevations. This discharge accumulates on the surface and flows northwest across Grassy Cove towards Mill Cave, where it sinks underground. This cave stream then travels south under Brady Mountain, resurging in a large trunk passage inside Run to the Mill Cave, continues subsurface to reappear at Devilstep Hollow Cave Window, and then rises at Head of Sequatchie Spring. Three other coves contribute to the flow system of Mill Cave: Little Cove drainage flows underneath McCullough Mountain to Bristow Spring in Grassy Cove; and McClough Hollow and Bat Town Cove drainage, both travelling under Flat Mountain before joining the subsurface flow in Mill Cave under Brady Mountain (Miller and Ham, 2021; Figure 5).

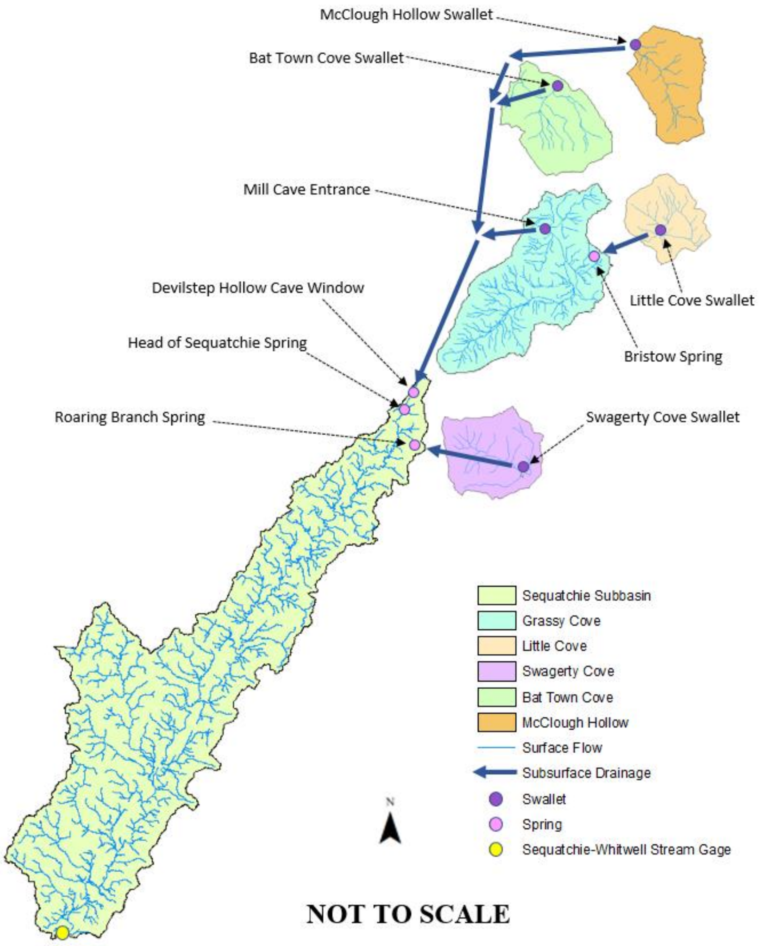


Figure 5. Illustration of perceived subsurface drainage routes from swallets and resurgence at springs with overland flow of the Sequatchie River to the USGS watershed outlet. Refer to Figure 3 for key topographic divide locations.

The dye trace, geochemical analysis, and hydrological investigations conducted by Nicholas Crawford during the 1970's established the recharge connections between drainage and resurgence in this hydrological system, including the subsurface drainage under Hinch Mountain from Swagerty Cove swallet to Roaring Branch Spring, revealing the interrelationship between these complex karst valleys. When combined, the discharge from Grassy Cove, Devilstep Hollow Overflow Spring, and Head of Sequatchie Spring become the headwaters of the Sequatchie River. Injection point (swallet) and detection point (spring) locations and subsurface drainage flow directions identified by Crawford's (1984; 1989) dye investigations and Tennessee Cave Survey (TCS) digital data subsets (Figure 6) provided general locations of relevant karst features during the field investigation of this research. The TCS database did not contain subsurface drainage routes or injection point locations from previous dye traces performed in McClough Hollow or Bat Town Cove, although the swallet locations, subsurface streams, and resurgence location at the Head of Sequatchie Spring are depicted on pages 312 & 313 of Crawford (1984).

In Tennessee, dye injection into subsurface waters requires permit application and acceptance notification from the Tennessee Division of Water Supply prior to dye trace investigations. These investigations study the movement of groundwater to help determine their origin, subsurface route, destination, and velocity of flow in karst aquifers. After key karst features are identified in a Karst Hydrological Inventory (KHI) (Crawford Hydrology Laboratory, 2021), dye receptors (e.g., activated charcoal packets) are deployed in groundwater resurgence locations, such as springs or karst windows. Once placed, a dye (or dyes) is directly injected into wells, sinking streams, or swallets and dye receptors are monitored, exchanged, and inventoried over a determined interval based on study objectives. Following the dye trace investigation, typically lasting weeks to months, collected dye receptors are processed and analyzed in a laboratory for fluorescence (CHL, 2021).

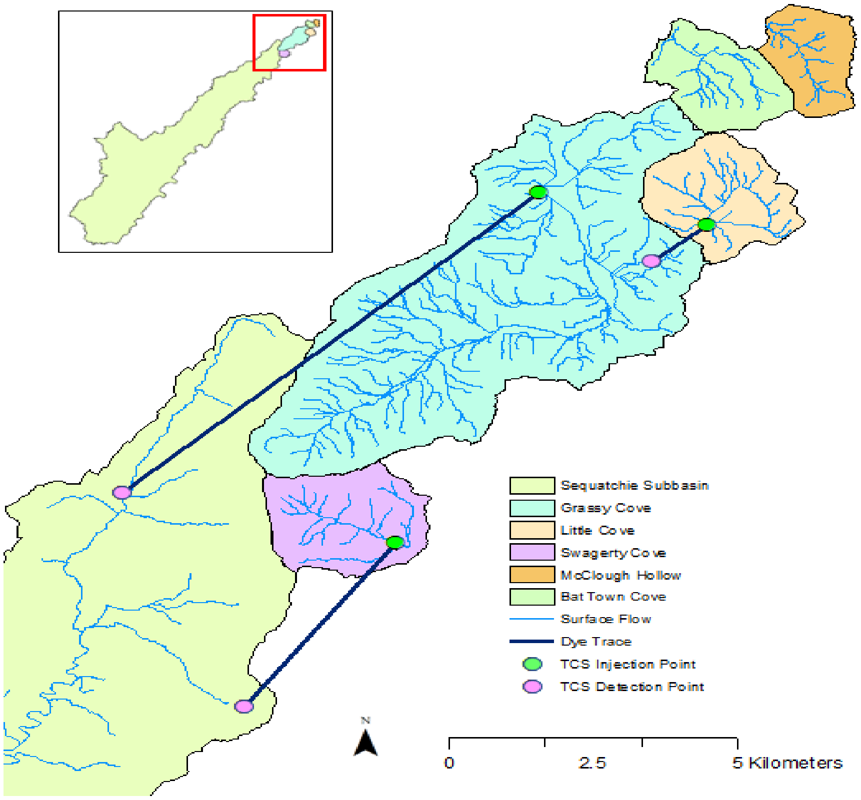


Figure 6. Injection and detection point locations and their corresponding subsurface drainage referenced from the Tennessee Cave Survey (TCS) digital data subset.

<u>Study Area</u>

Sequatchie Subbasin and Surrounding Coves

The study area (Figure 7) includes the Sequatchie Subbasin and surrounding karst valleys of Grassy Cove, Little Cove, Swagerty Cove, McClough Hollow, and Bat Town Cove. This expansive area extends across five counties between the northeastern section of the Sequatchie

Valley and the transitional area southwest of the Crab Orchard Mountains in Middle Tennessee. Elevation here ranges from 193.9 to 926.7 m above sea level with the perennial headwaters of the Sequatchie River originating from a cove–like setting of 336.5 m and the outlet (USGS Sequatchie–Whitwell streamgage) at an elevation of 192.82 m. Flanking the valley are the eastern and western escarpments, averaging nearly 300–500 m in height and 6–8 km wide.

The northeastern section of the Sequatchie Valley is characterized by a humid subtropical climate averaging 1,349 millimeters (mm) of rainfall per year and mean daily minimum and maximum air temperatures of 49.7 degrees Celsius (°C) to 71.5 °C (2016 to 2021). Daily precipitation and air temperature information is monitored from ten nearby climate (weather) stations. Soils in the study area are comprised of 267 unique types and land use/land cover in the karst valleys are dominated by forest and pasture whereas the Sequatchie Valley includes forest, pastures, and row crop mainly of corn and soybeans. The total drainage area is 105,420.1 hectare (ha) or 1,054.2 km^2 which is distributed as follows: Sequatchie Subbasin – 100,107 ha; Grassy Cove – 3,466.4 ha; Swagerty Cove – 574.9 ha; Little Cove – 537.4 ha; Bat Town Cove – 421.7 ha; and McClough Hollow – 312.6 ha.

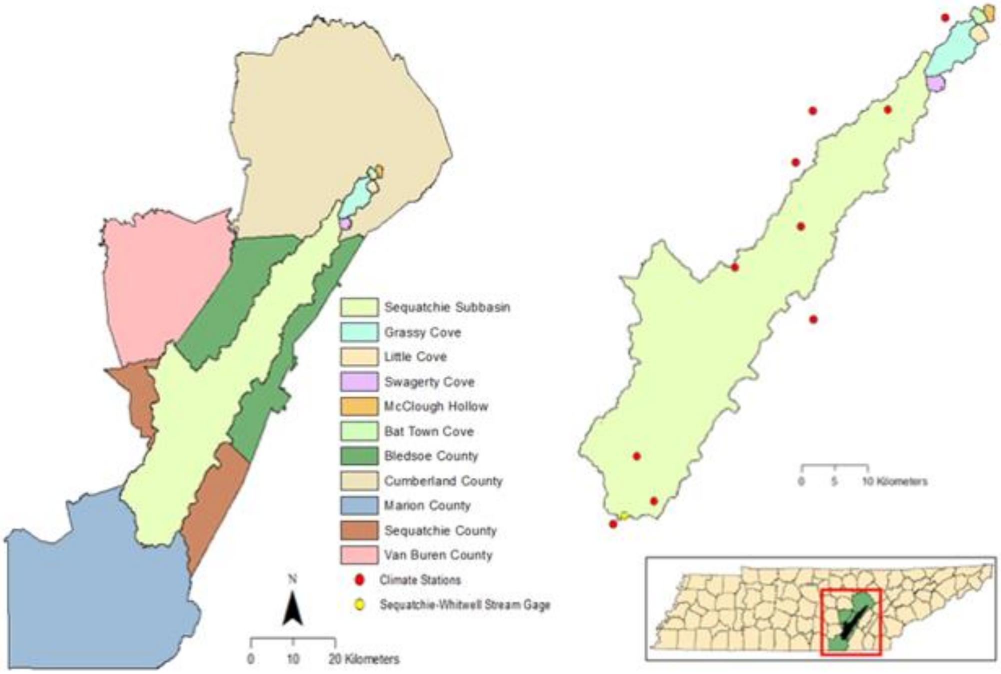


Figure 7. Study area and climate station locations within the surrounding counties of Middle and East Tennessee, USA. See Table 9 for names and locations of climate stations.

Scope of Research

The study period was from 2016 to 2021. The scope of this research covers the hydrological modeling of integrated surface and subsurface runoff for the northeastern portion of the Sequatchie Valley and five interconnected coves. High resolution DEMs are used to construct drainage basins for each location and their sub–division into homogeneous sub–catchments (AnnAGNPS cells) and channel network utilizing the TopAGNPS–GIS software package. Daily temporal weather variations were represented using ten climate stations. Raster grid layers of annual variations in land–cover and soil type were processed to define the dominant land–management and soil type for each sub–catchment. Required input parameters of topography, climate, soil, and farming management are then assembled for use in the AnnAGNPS watershed pollutant model.

Eight total AnnAGNPS simulations were used to estimate surface and subsurface runoff and include: 1. Sequatchie Subbasin; 2. Grassy Cove; 3. Swagerty Cove; 4. Little Cove; 5. Bat Town Cove; 6. McClough Hollow; 7. integrated simulation of Little Cove and Grassy Cove; and 8. integrated simulation of Sequatchie Valley subbasin and all surrounding coves. Output data is then calibrated in the Sequatchie Subbasin by pairwise comparison of simulated AnnAGNPS total streamflow data against historical (observed) streamgage data after baseflow removal. The final AnnAGNPS simulation integrates all five coves and the Sequatchie Subbasin. Efficiency of the hydrologic model is assessed through multiple statistical analyses to visualize and investigate patterns, relationships, and trends of qualitative and quantitative data in order to accept or reject the hypothesis.

Limitations in this study are as follows: (1) dye tracing information regarding subsurface flow were restricted to published studies and new dye trace investigations are either unpublished or not performed, (2) only the AnnAGNPS watershed pollutant model was used to capture surface flow and no ground water model was considered, and (3) analyses were considered at the monthly time step to overcome uncertainties in subsurface latency time.

CHAPTER III

MATERIALS AND METHODS

Field Investigation

A modified Karst Hydrological Inventory (KHI), as outlined in Crawford Hydrology Laboratory (2021), was conducted in December of 2022 under high flow conditions to visually identify key karst features (springs, sinkholes, sinking streams, swallets, karst windows, and caves) while dominant resurgence points were active. Features on private properties were accessed after consent from the land owner and Global Positioning System (GPS) coordinates of the Tennessee Cave Survey (TCS) database and Crawford (1984; 1989) served as reference to general locations of prior dye injection (swallet) and detection (spring) points. Relevant karst features were ground verified then inventoried by name and location, photographed by author (Figure 8; Appendix B), and updated coordinates captured using a hand–held GPS receiver (see Table 1). The USGS Sequatchie River near Whitwell, Tennessee (TN) streamgage (stream site: 03571000; elevation: 192.82 m; latitude 35°12'23.42", longitude 85°29'49.68") just east of Whitwell, TN in Marion County was inventoried and identified as the study area’s watershed outlet.

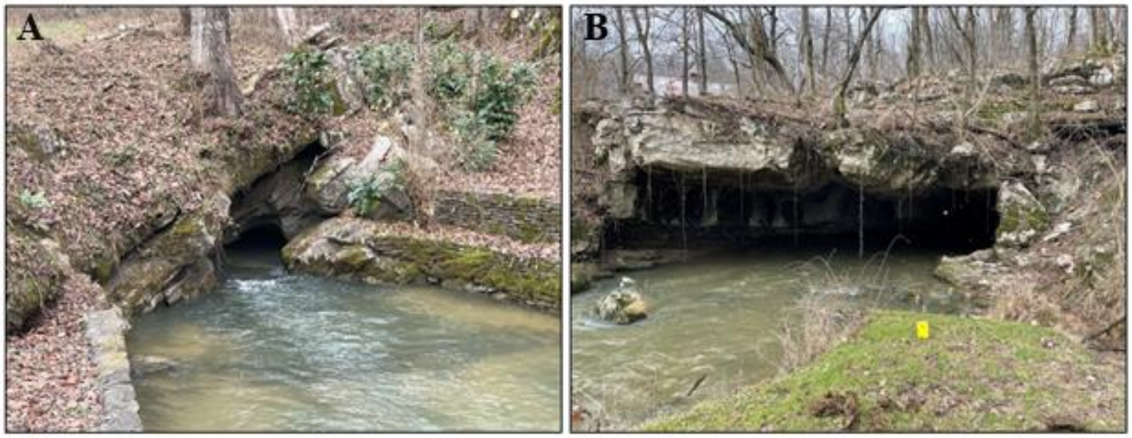


Figure 8. Photographs of relevant karst features taken by author during December 2022 field investigations. (A) Head of Sequatchie Spring in Cumberland Plateau State Park and Mill Cave entrance in Grassy Cove (B), Cumberland County, Tennessee (Photo B used with permission).

Table 1. Study area relevant karst feature coordinates referenced from the Tennessee Cave Survey (TCS) in comparison to ground verified GPS coordinates of features.

	TCS Database		Ground Verified		
Karst Feature	**Latitude**	**Longitude**	**Latitude**	**Longitude**	**Elevation (m)**
Head of Sequatchie Spring	35.79527	-85.00883	35.79305	-85.00760	336.49
Devilstep Cave Window	35.81964	-84.98877	35.79603	-85.00867	369.42
Devilstep Hollow Spring	35.82372	-84.98973	35.79550	-85.00872	352.35
Swagerty Cove Swallet	35.78416	-84.95636	35.78500	-84.95666	576.38
McClough Hollow Swallet	--	--	35.89067	-84.88242	523.95
Bat Town Cove Swallet	--	--	35.88575	-84.89157	525.47
Little Cove Swallet	35.84874	-84.89480	35.84992	-84.89542	491.34
Bristow Spring	35.84147	-84.90528	35.84620	-84.90381	533.70
Roaring Branch Spring	35.75076	-84.98636	35.75390	-84.98189	740.66
Mill Cave Entrance	35.85604	-84.92697	35.85675	-84.92709	464.82

Note: (--) implies no data.

Conceptual Watershed Model

The conceptual watershed model (Figure 9) illustrates the flow of surface water in each drainage basin (subbasin and karst valley or cove) to their respective outlet (primary swallet, cave entrance, or streamgage), the perceived subsurface drainage flow direction and resurgence at springs, and their contribution to the Head of Sequatchie River. Streamflow then travels south–southwest to the watershed outlet at the USGS streamgage.

Summary of flow: The surrounding coves contribute to the flow system of Mill Cave as Little Cove drainage flows underground to Bristow Spring in Grassy Cove and McClough Hollow and Bat Town Cove drainage both travel underground before joining the subsurface flow in Mill Cave. Stream flow accumulates on the surface and flows northwest across Grassy Cove towards Mill Cave entrance where it sinks underground. The cave stream then travels south to reappear at Devilstep Hollow Cave Window, Devilstep Hollow Overflow Spring, and then rises at Head of Sequatchie Spring. The combined discharge from Grassy Cove, Devilstep Hollow Overflow Spring (if under wet conditions), and Head of Sequatchie Spring become the headwaters of the Sequatchie River. The Sequatchie River then continues south–southwest, combines with the subsurface drainage from Swagerty Cove swallet to Roaring Branch Spring, and flow proceeds to the outlet of the drainage basin.

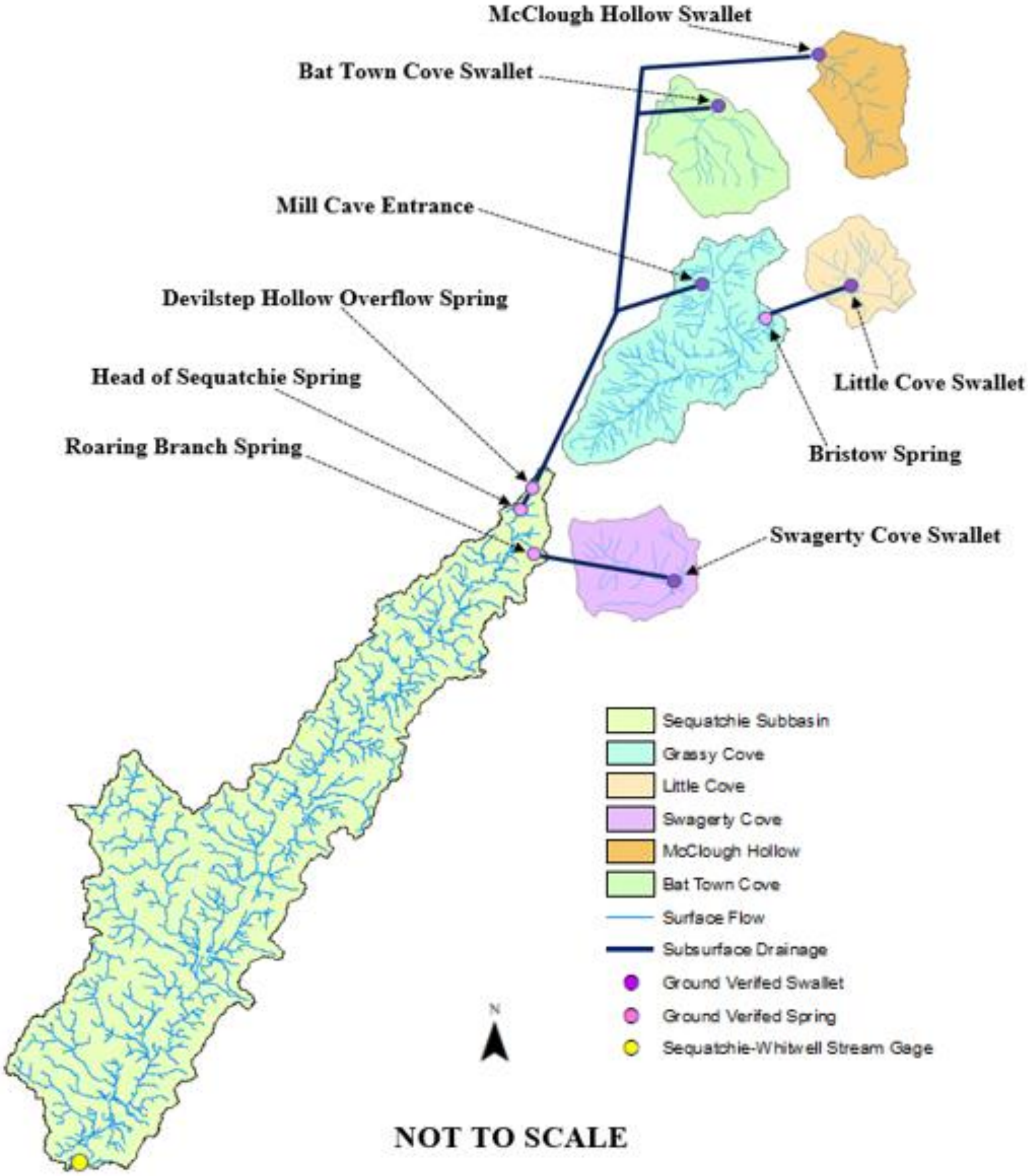


Figure 9. Conceptual watershed model of surface flow and subsurface drainage routes in the Sequatchie Subbasin and interconnected karst valleys.

Hydrological Modeling Technology

TopAGNPS

Accurate interpretation and evaluation of topography is a critical aspect in the hydrologic processes of surface runoff, soil erosion, and water storage. Digital landscape extraction tools such as Topographic Parameterization for Agricultural Non–Point Source (TopAGNPS) use automated programs that directly analyze and evaluate high–quality Digital Elevation Models (DEM), the basic program input requirement used to rapidly parameterize a watershed (Garbrecht and Martz, 1999). TopAGNPS, a subset module of the Topographic Parameterization (TOPAZ) software package developed by the United States Department of Agriculture - Agricultural Research Service (USDA–ARS), is primarily utilized to analyze and support hydrologic modeling. The fundamental concepts of the program are to pre–process DEM data (i.e., rectify localized sink and flat surface artifacts), hydrographically segment (i.e., generation of the channel network, sub–catchments, and drainage divides), and topographically parameterize (i.e., measuring a variety of properties and hydrologic variables) the raster DEM, channel network, and sub-catchments (Garbrecht and Martz, 1999; Armstrong and Martz, 2003). TopAGNPS v5.5 is the most recent version and provides interface with a user–selected Geographic Information System (GIS) from generated attribute tables and raster data outputs produced during topographic processing. Output files interfaced with GIS allow the user to display raster files, perform layer algebra, and interact with the basic modeling units of cells and reaches required for automated hydrologic analysis and watershed modeling in AnnAGNPS.

Organization of TopAGNPS consists of three interdependent programs that operate in specific order based on the requirement of one module ‘output’ being another modules ‘input’ for proper sequence operation (Armstrong and Martz, 2003). These include program DEDNM

(Digital Elevation Drainage Network Model), program RASPRO (Raster Properties), and program RASFOR (Raster Formatting). DEDNM is the primary component of TopAGNPS and always applied first to perform data quality control followed by either RASPRO or RASFOR based on user selection (Garbrecht at el., 2000). DEM raster, sub–catchment, and channel network processing is based on the deterministic eight (D8) method and the concepts of downslope flow routing and accumulation to define cell landscape properties, flow direction, and drainage boundaries (Garbrecht and Martz, 1999). Sub–catchment drainage areas and the topographic and spatial configuration of the drainage network are controlled by user–specified value thresholds of the critical source area (CSA) and minimum source channel length (MSCL) parameters (Garbrecht and Martz, 1999; Armstrong and Martz, 2003).

The D8 flow method defines the surface drainage area over the DEM landscape by first determining the steepest downslope of each raster cell. It then models the flow vector for each of those cells to one of the eight adjacent cells with the steepest downslope which identifies the upstream drainage area of that sub–catchment (Garbrecht and Martz, 1999). A watershed boundary (i.e., the extent of the flow vector grid) is then defined by all cells contributing flow to the user–defined outlet cell (Armstrong and Martz, 2003). Identification of the channel network is based upon the contributing area of the watershed and user–specified thresholds of CSA and MSCL parameters. The drainage area at the end of flow vectors (where the smallest or first–order channel begins) defines the CSA, and removal of exterior channel links (source channels) shorter than the specified threshold (i.e., retaining source channels based on the meter length of the flow path) defines the MSCL (Armstrong and Martz, 2003; Momm at el., 2017). Definition of the sub–catchments and channel network following the execution sequence of TopAGNPS allows the assignment of spatial landscape attributes (e.g., soil layers, soil majorities, management and land use, climate stations, and crop layers) to the required basic modeling units (cells and reaches) used in AnnAGNPS watershed model simulation.

AnnAGNPS

The Annualized Agricultural Non–Point Source (AnnAGNPS) watershed pollution model system uses a continuous simulation, event–based Pollutant Loading (PL) surface runoff model designed to predict long–term sediment–transported PLs and water quality responses. This extensively used computer simulation model is capable of analyzing and estimating surface runoff based on the Natural Resources Conservation Service (NRCS) Curve Number (CN) method (USDA–SCS, 1986; TDOT, 2021), sediment transport, and continuous pollutant (e.g., nutrient and pesticide) load simulations (Momm et al., 2019), predominately at the agricultural watershed–scale. Operation of AnnAGNPS requires extensive data processing competence, sufficient training, and use of large amounts of input data. System components and primary inputs of AnnAGNPS include user–provided continuous climate data, land use and management, soils, stream network and topographic information from DEMs, and point sources (e.g., construction sites, animal feedlots, and springs) (Young and Shepherd, 1995; Yuan, 2020).

AnnAGNPS is the enhanced version of the Agricultural Non–Point Source (AGNPS) pollution modeling system, a single event–based non–point source (NPS) computer model jointly developed by the Agricultural Research Service (ARS) and NRCS system during the 1980s (Bingner and Theurer, 2001). The latest version, AnnAGNPS v6.1, is designed to subdivide the watershed into basic modeling units (Figure 10) of homogenous hydrologically–derived cells (sub–catchments) in terms of topography and hydrologic soil properties (Yuan, et al., 2020). Utilizing additional automated analysis tools, such as the GIS assisted flow–routing algorithms of TopAGNPS (Momm et al., 2017), one can estimate the properties of sub–catchment and channel network layers to provide quantitative and spatial information on slopes, drainage areas, reach lengths and patterns, and topographic elevations (Garbrecht and Martz, 1999; Legates and McCabe, 1999; Momm et al., 2019). Analysis of land surfaces (cells) at any point within the watershed is viable since sub–catchments have relatively uniform hydrological and physical characteristics. AnnAGNPS utilizes the Revised Universal Soil Loss Equation (RUSLE) method (Renard et al., 1997) to estimate sheet and rill erosion, and the Hydro–geomorphic Universal Soil Loss Equation (HUSLE) (Theurer and Clarke, 1991; Luo et al., 2015) to simulate sediment delivery. Detached sediments, referred to as *erosion*, are either deposited or routed from each cell to the channel network, referred to as *yield*, and evaluated for transport to the watershed outlet, referred to as *load* (Momm et al., 2017; 2019).

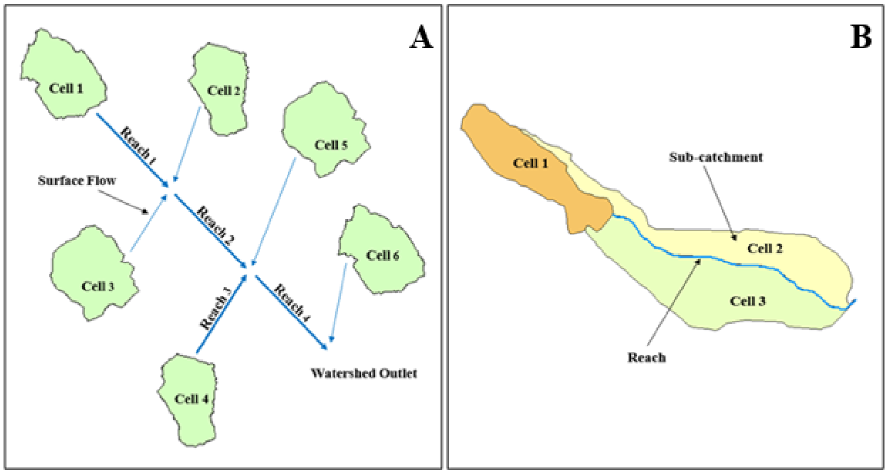


Figure 10. Internal representation of surface flow connections (A) and watershed characterization (B) between the basic modeling unit process of sub–catchment (cell) and concentrated flow (reach) simulated by AnnAGNPS.

Pollutants, particularly those emerging in agricultural areas, are considered NPS pollutants as their generation from unidentifiable, or not easily traced sources, originates through agricultural runoff, adulterated lands, or informal settlements (Adu and Kumarasamy, 2018). NPS pollutants typically discharge in uncertain quantities through multiple outlets within agricultural watersheds as their chemical and physical constituents transport from their perceived origin and either deposit in the confines of the stream channel or route to the watershed outlet (Mulla, et al., 2019). The AnnAGNPS model can be utilized to predict these PLs at their potential source and then simulate their movement in a stepwise manner through the watershed (Parsons, et al., 2007; Mulla, et al., 2019). The behavior of nutrients like nitrogen (N) and phosphorous (P), both major contributors to water pollution but essential to agriculture, in addition to pesticides are estimated and considered in the AnnAGNPS model (Young and Shepherd, 1995).

AnnAGNPS is widely used around the world to describe, address, and identify the potential source of problematic NPS pollutants in agricultural watersheds and proves to be an effective tool to prioritize remedial practices. Recent applications of AnnAGNPS include the evaluation of sediment loads and landscape "hot spot" identification within a Hawaiian tropical watershed (Polyakov et al., 2007), identification of conservation practices due to sediment losses from agricultural fields in the Mississippi Delta (Yuan, et al., 2008), evaluation of runoff, peak flow, and sediment yield in a Belgian agricultural watershed (Zema et al., 2012), and the applicability of the AnnAGNPS model in karst agricultural areas of Southwest China (Liying, et al., 2021). Model simulation components and various user–selected output parameters include watershed evaluation, sediment and nutrient transport, erosivity, simulated peak flow and runoff volume at the watershed outlet, and estimations of upland erosion (Yuan et al., 2008; Mulla, et al., 2019). Additionally, AnnAGNPS helps determine Best Management Practices (BMPs) for the surface runoff model, set Total Maximum Daily Loads (TMDLs) of waterbody pollutants, and aids in quantification of farming conservation practices (Polyakov et al., 2007; Yuan et al., 2008; Lou et al., 2015).

Schematic Diagram of Hydrological Modeling

The schematic diagram of hydrological modeling provides a visual representation of the system components (spatial and tabular datasets, external programs, AGNPS tools, and primary inputs for AnnAGNPS) required for watershed modeling (Figure 11). Input and output *spatial datasets* include the DEM, Crop Data Layer (CDL), soil layer, climate stations, and TopAGNPS channel network and sub–catchment rasters. Input and output *tabular datasets* include the soil, climate, statistical seed (synthetic climate), gage (simulated streamflow), and source output tables. AGNPS tools utilized are TopAGNPS, custom Python scripts, RUSLE2 Import to AnnAGNPS (RITA), NASIS Import to AnnAGNPS (NITA), and the Generation of Weather Elements for Multiple Applications (AgGEM) program. Integrated external databases and tools include a user–specified GIS, RStudio scripts, RUSLE2 templates, and the primary inputs for AnnAGNPS simulation include reach, cell, management schedule and operation, soil and soil layer, climate, and other (e.g., point source pollutants or riparian buffers) data section.

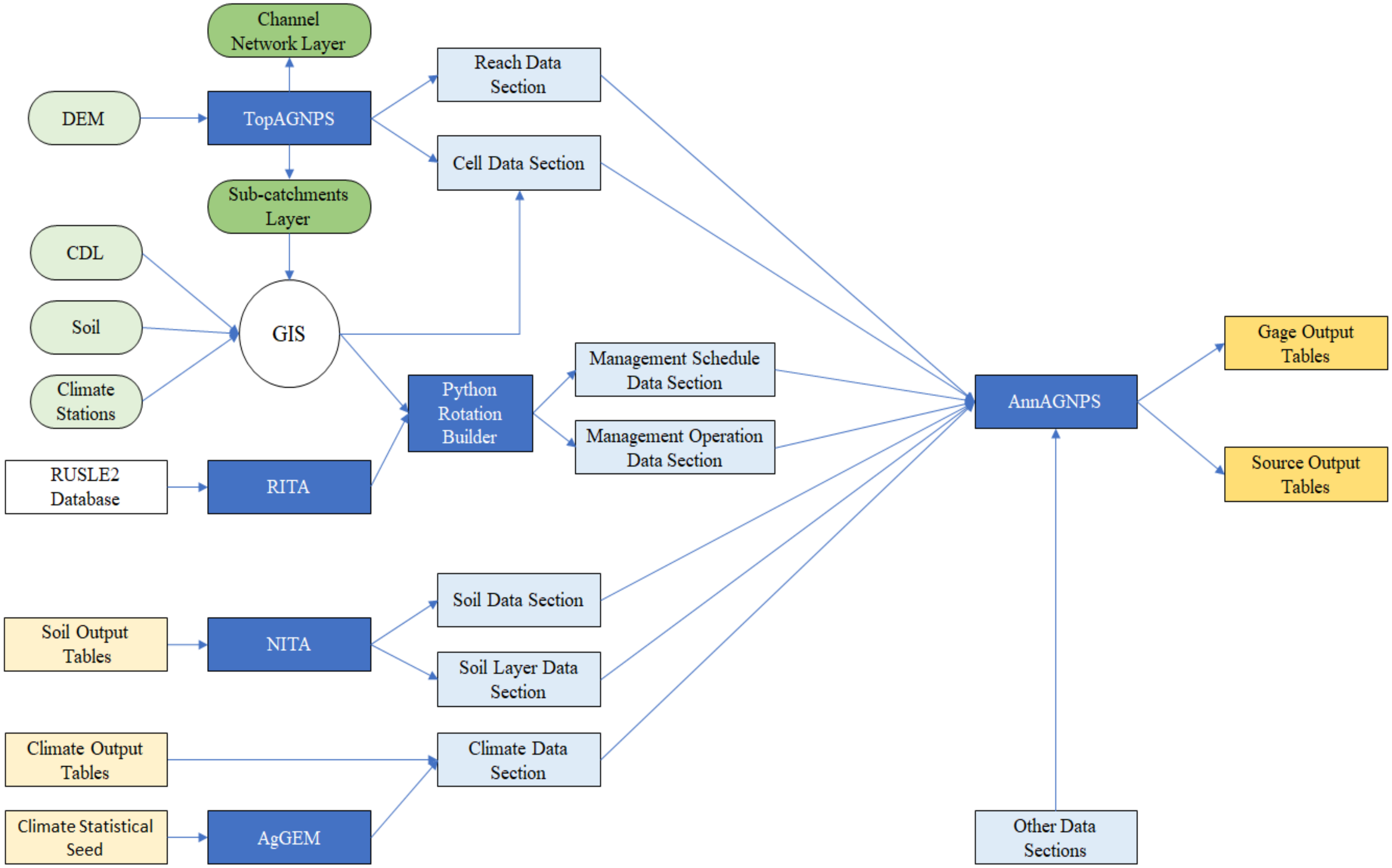


Figure 11. Schematic diagram of system components required for watershed modeling represented by input and output spatial datasets (light– and dark–green), input (light–yellow) and output (dark–yellow) tabular datasets, external databases (white), AGNPS tools (dark–blue), and required primary inputs for AnnAGNPS (light–blue).

GIS Analysis

* Further clarification into the technical aspects of modeling applications and syntax performed for TopAGNPS, AnnAGNPS, and other tools & external databases utilized in this study should refer to *TopAGNPS & AnnAGNPS Technical Processes* (Appendix A).

Digital Elevation Model

A Digital Elevation Model (DEM) is a raster grid layer representing a matrix composed of numeric values that continuously describe terrain elevation. Topography determines the formation and evolution of land features and surface water flow in watersheds. DEMs for the Sequatchie Subbasin and five surrounding karst valleys were generated from high–quality (0.762 m) light detection and ranging (LiDAR) surveys. LiDAR–generated DEMs are mosaicked and converted from feet (ft) to m and spatially referenced to Universal Transverse Mercator (UTM) Zone 16 North. Each karst valley is then resampled to 3–m spatial resolution and the Sequatchie Subbasin to 9–m due to its large area and file size.

Topographic Characterization Using the TopAGNPS Tool

The TopAGNPS software package interfaced with a Geographic Information System (GIS) can rapidly evaluate digital landscape elevations by performing watershed segmentation and sub–catchment parametrization. The DEM in American Standard Code for Information Interchange (ASCII) file format is the primary input required for TopAGNPS (Garbrecht and Martz, 2000). Processing of each karst valley is performed using CSA and MSCL values of 5 ha and 250 m and 50 ha and 500 m for the Sequatchie Subbasin, respectively. Key raster grid files generated using the TopAGNPS program in reduced–mode are listed in Table 2 with a brief description of each. These raster grids were used to analyze the catchments generated and define the catchment outlet location. The outlet location for the Sequatchie Subbasin was defined at the

USGS gage station while the outlet for each karst valley was selected to match the swallet information described in dye trace studies combined with the field investigation. Once the outlet is identified, full–processing of TopAGNPS produces outputs comprised of spatial and tabulated attribute data. The organization of spatial data is ASCII format and provides the ability to visualize raster grids depicting topographic attributes, channel network, and watershed subdivision into sub–catchments while tabular data contains attribute tables of channel network structure and sub–catchment properties.

Table 2. List of output files provided by TopAGNPS reduced–mode execution that are utilized in this study, including brief descriptions of their function.

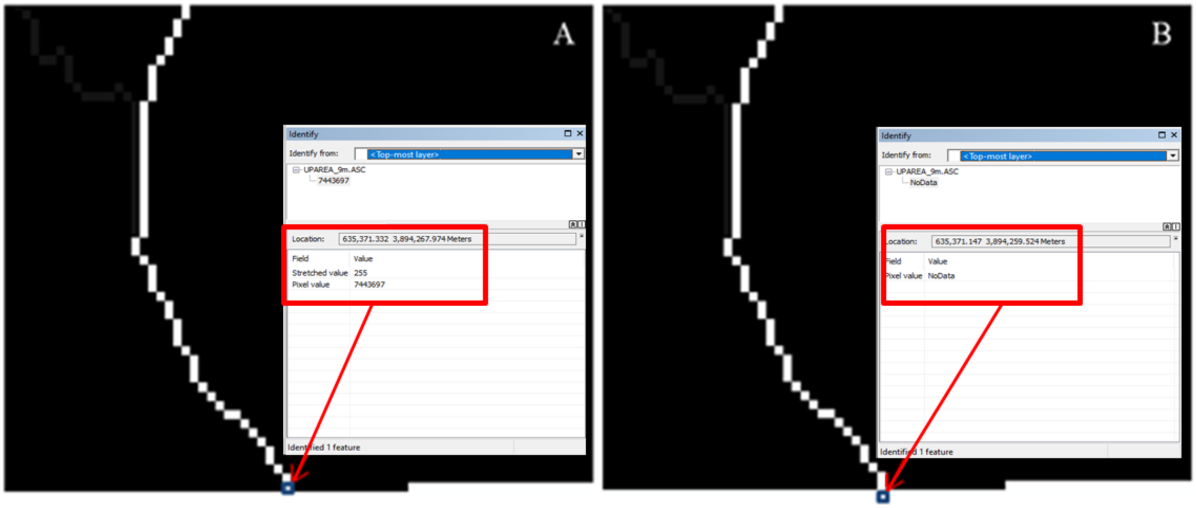
TopANGPS Output File	**Description**
BOUND.OUT	Drainage basin boundary raster
NETW.OUT	Channel network (reach) within drainage basin raster
UPAREA.OUT	Flow Accumulation raster
SUBWTA.OUT	Sub–catchment (cell) index raster

Outlet Identification in Bowl–Like Topography

Typical processing of topographic data by flow routing algorithms assumes that all the flow exits the raster grid layer at its edges (edge of the raster grid) or to raster grid cells with no–data (-9999) elevation values indicating the area is outside of the study area. The identification of a typical drainage basin outlet is at the extent of the flow vector grid defined by all cells contributing flow to the user–selected outlet cell and out of the boundary to a no–data value elevation. Commonly, the outlet raster grid cell is determined by visual analysis of the flow accumulation raster grid in TopAGNPS (referred as the UPAREA raster grid file in GIS) and using the *identify* tool. The location coordinates in meters (Figure 12) are copied from the outlet cell and added to the TopAGNPS control file X (outlet column) and Y (outlet row) columns, then full processing is executed to finalize the location.

The karst valleys in the study area are bowl–shaped closed depressions with swallow holes (swallets) along the bottom. Surface water of the vector grid flows from the higher–elevation rims into low–spots of the sinkhole and disappear into the subsurface. Swallet locations ground–verified during the field investigation are used to identify each cove outlet. Traditionally, flow routing algorithms artificially fill depressions so that all flow leaves the DEM through its edges. However, such modification in bowl–shaped closed depressions would be an incorrect representation of the surface–subsurface water flow. An alternative approach was used to force the flow routing algorithm within the TopAGNPS toolset not to fill the entire depression.

After identifying the outlet location coordinates, the outlet elevation value must be altered. This procedure consisted of opening the respective DEM ASCII file in a text editor, such as *notepad* ++, using the *find* tab to locate the cells surrounding the outlet, and changing the given outlet cell elevation in meters to the no–data value of *-9999* (Figure 13). Again, the location X and Y meter coordinates for the altered outlet are copied and added to the TopAGNPS control file and full–processing is executed to finalize the location. Similar to the typical drainage basin outlet, TopAGNPS fills the depressions within the DEM but assumes the raster grid cell with a no–data value to be outside of the study area and flow is routed downslope along the flow vector grid to the -9999 outlet cell and ‘out of the boundary’.



A: UPAREA_9m.ASC — 7443697

Location: 635,371.332 3,894,267.974 Meters

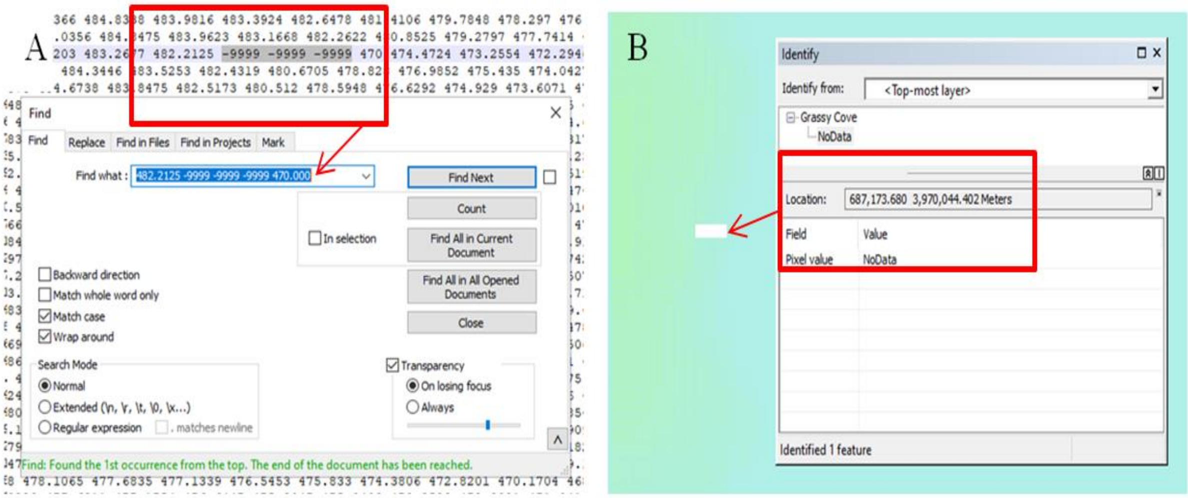
Field	Value
Stretched value	255
Pixel value	7443697

B: UPAREA_9m.ASC — NoData

Location: 635,371.147 3,894,259.524 Meters

Field	Value
Pixel value	NoData

Figure 12. Identification of the Sequatchie Subbasin outlet cell along the flow vector (white line) using the UPAREA raster grid and location coordinates in GIS (A) and the no–data pixel value indicating the outside of the drainage basin boundary (B).



A:

Find what : 482.2125 -9999 -9999 -9999 470.000

☐ In selection

☐ Backward direction
☐ Match whole word only
☑ Match case
☑ Wrap around

Search Mode
● Normal
○ Extended (\n, \r, \t, \0, \x...)
○ Regular expression ☐ . matches newline

☑ Transparency
● On losing focus
○ Always

Find: Found the 1st occurrence from the top. The end of the document has been reached.

B: Grassy Cove — NoData

Location: 687,173.680 3,970,044.402 Meters

Field	Value
Pixel value	NoData

Figure 13. Identification of the Grassy Cove outlet location by editing the ASCII DEM raster elevation from the original meter value to a *-9999* no–data value (A) and pixel result represented on the raster layer in GIS (B).

Files created from TopAGNPS full processing include the AnnAGNPS reach and cell data sections characterizing concentrated flow (reaches) and fields (AnnAGNPS cells) summarized in Table 3. Additional raster grid layers include BOUND (watershed boundaries), NETW (channel network), and AnnAGNPS cell ID (SUBWTA). These raster grids are then converted into GIS vector layers for improved visualization and analysis as illustrated for the Sequatchie Subbasin in Figure 14. Furthermore, the shapefile images for the karst valleys are displayed in Appendix C.

Table 3. Summary of reach and cell characteristics following CSA and MSCL threshold assignment and TopAGNPS watershed segmentation and sub–catchment parametrization.

	Reach Data		**Cell Data**		
Drainage Basin	**Number of Reaches**	**Reach Length Mean (m)**	**Number of Cells**	**Cell Area Mean (ha)**	**Average Slope Grade (%)**
Sequatchie Subbasin	799	1,397.9	1,996	50.2	0.17
Grassy Cove	253	502.7	626	5.54	0.21
Little Cove	38	463.9	94	5.72	0.24
Swagerty Cove	33	470.6	83	6.93	0.29
McClough Hollow	27	330.5	68	4.59	0.27
Bat Town Cove	27	489.0	67	6.29	0.22

Note: mean = $(\sum n)/n$

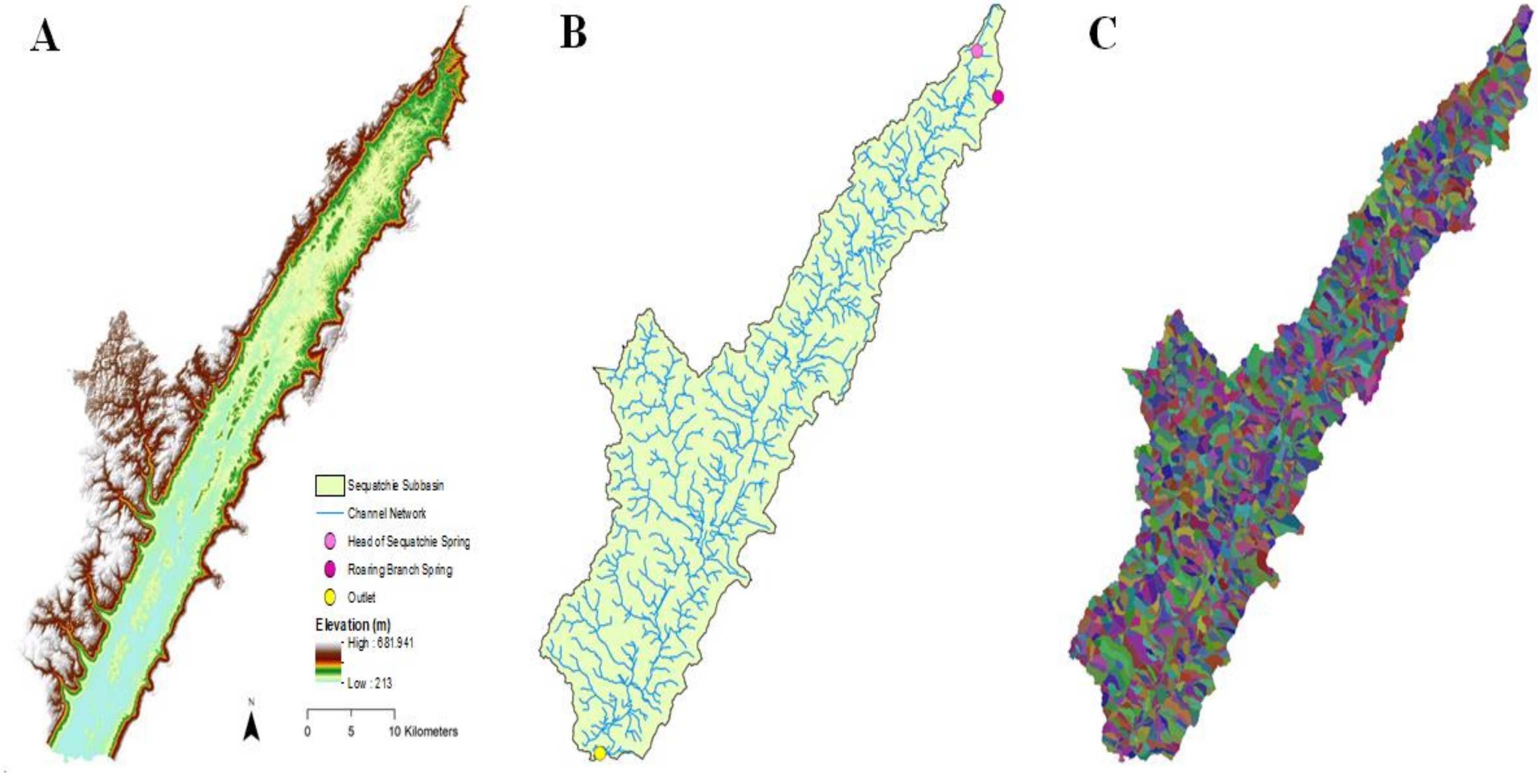


Figure 14. GIS layers for the Sequatchie Subbasin 9–m resolution DEM (A), channel network of drainage basin with prominent spring and outlet locations (B), and sub–catchments representing the watershed division into AnnAGNPS cells (C).

Soil Layers and Soil Data

The physical properties and cover conditions of soil influence the rate of water infiltration and runoff potential following a storm. Digital datasets of soil maps help determine soil types, hydrologic soil groups, texture description, and the spatial frequency for each drainage basin and individual sub–catchment in their respective counties. The datasets for soil type and soil table used in this study are downloaded from the USDA–Soil Survey Geographic (SSURGO) database, which contain spatial soil layers and tabular data. Spatial datasets of each county are merged and clipped to the boundary of each drainage basin based on a GIS overlay. The *map unit symbol* (soil type) and *state and county code* (area symbol) in the attribute table are then combined to create a signifier (i.e., the soil type plus county code) for each unique soil type which is merged and clipped to the study area drainage basins. A total of 267 unique soil types across five counties are used in the study area. Unique soil types are then combined to each downloaded county soil table using the NITA program to format a new value field of unique numbers shared between soil type and soil table. This zonal statistic value is created utilizing a Python script to write the soil layer and soil data comma separated value (csv) output files required for AnnAGNPS simulation (see examples in Table 4 and Table 5). The soil majority values (zonal statistic spatial layers) are then allocated to each AnnAGNPS Cell ID shapefile (Figure 15).

Table 4. Example of soil layer data section output displaying the number of soil layers, layer depth of soils, and the content of clay, silt, sand, and very fine sand for each unique soil ID and AnnAGNPS cell.

AnnAGNPS Cell	Unique Soil ID	Layer Number	Layer Depth (mm)	Content % Clay	Silt	Sand	Very Fine Sand
224	Aa TN115	1	5.91	0.29	0.37	0.34	0.1
234	Ab TN115	1	7.87	0.16	0.16	0.68	0.17
211	Ac TN115	2	27.17	0.3	0.37	0.34	0.1
210	Ad TN115	3	29.13	0.3	0.15	0.56	0.1
183	Ae TN115	2	12.99	0.28	0.38	0.34	0.1
53	AeC TN153	2	3.15	0.17	0.4	0.43	0.07

Table 5. Example of soil data section output displaying the unique soil ID and AnnAGNPS cell utilizing the NITA program and Python script, including the hydrologic soil group, soil erodibility factor (K factor) value, and soil texture description.

AnnAGNPS Cell	Unique Soil ID	Hydrologic Group	K Factor	Soil Texture
224	Aa TN115	B	0.24	Clay loam
234	Ab TN115	B	0.2	Fine sandy loam
211	Ac TN115	B	0.2	Fine sandy loam
210	Ad TN115	B	0.1	Stony fine sandy loam
183	Ae TN115	B	0.1	Stony fine sandy loam
53	AeC TN153	A	0.15	Flaggy loam

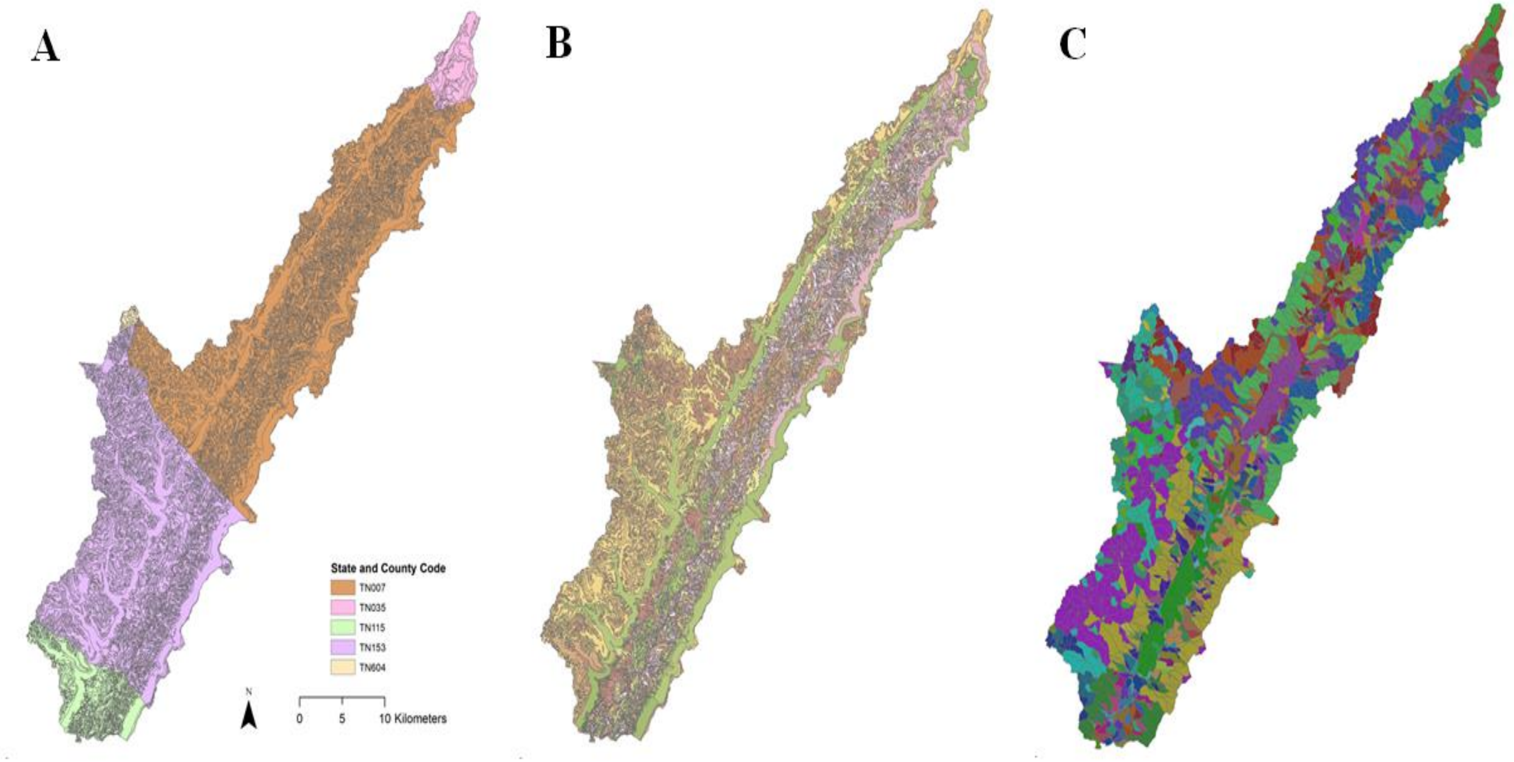


Figure 15. GIS layers of the Sequatchie Subbasin original soil type layer displaying state and county codes (A), the soil type and soil layer merged into unique soil IDs (B), and soil zonal statistic majorities discretized to the AnnAGNPS cells (C).

Land Cover and Farming Management

Land use information on dominant crop types and statistical majorities are determined from the USDA–NRCS *Crop Data Layers* (CDL) annually. Farming management is composed by a sequence of farming operations for individual crop combined into annual crop rotations. Farming management operations and schedules were created by integrating temporal sequences of crop type with sequences of typical crop operations (Momm at el., 2017). After downloading the individual CDL rater grids for the study area, all CDL years (2016 to 2021) are mosaicked and clipped to study area boundaries (Figure 16). A custom Python script is utilized to calculate majority (i.e., a master category number for each grid code) zonal statistics and land use/land cover percentages for each sub–catchment (Figure 17). CDL majorities are then converted from a master category number into a letter representative of the majority land use/land cover type (Table 6). The objective is to reduce the total number of land use types while representing the dominant ones.

Table 6. Conversion of master category number (zonal majority) of land cover/land use into a representative letter to help determine the management sequence for each CDL year.

Master Category Number	Letter Representation	Description
1	S	Soybeans (Row Crop)
5	C	Corn (Row Crop)
26, 29, 37, 152, 176, or 190	G	Grassland, Pasture, or Meadow
83 or 111	W	Water Body
121, 122, 123, or 124	U	Urban Development
141, 142, or 143	F	Deciduous Forest

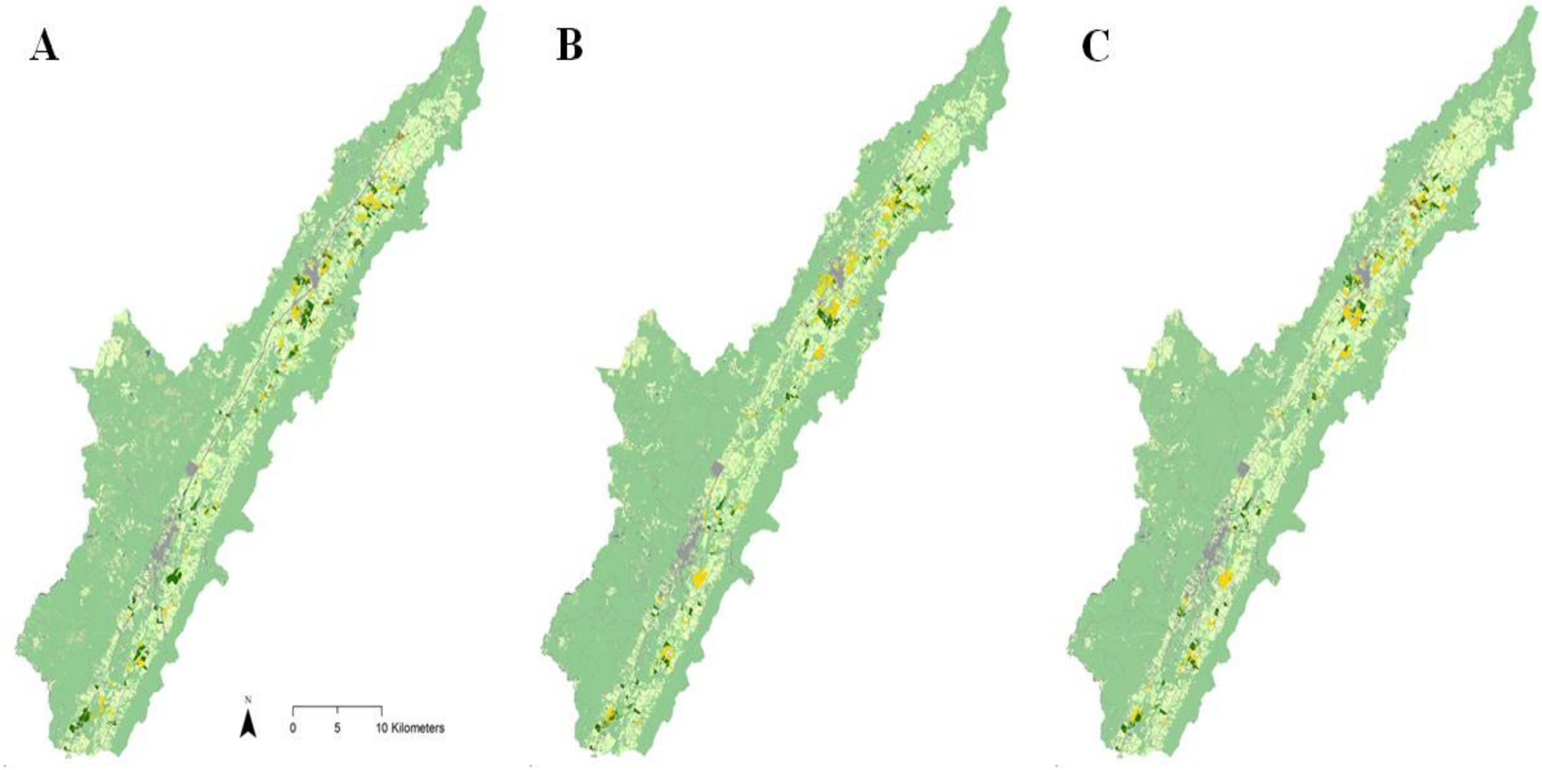


Figure 16. Assignment of the 2016 (A), 2019 (B), and 2021 (C) Crop Data Layers (CDLs) to the Sequatchie Subbasin raster boundary.

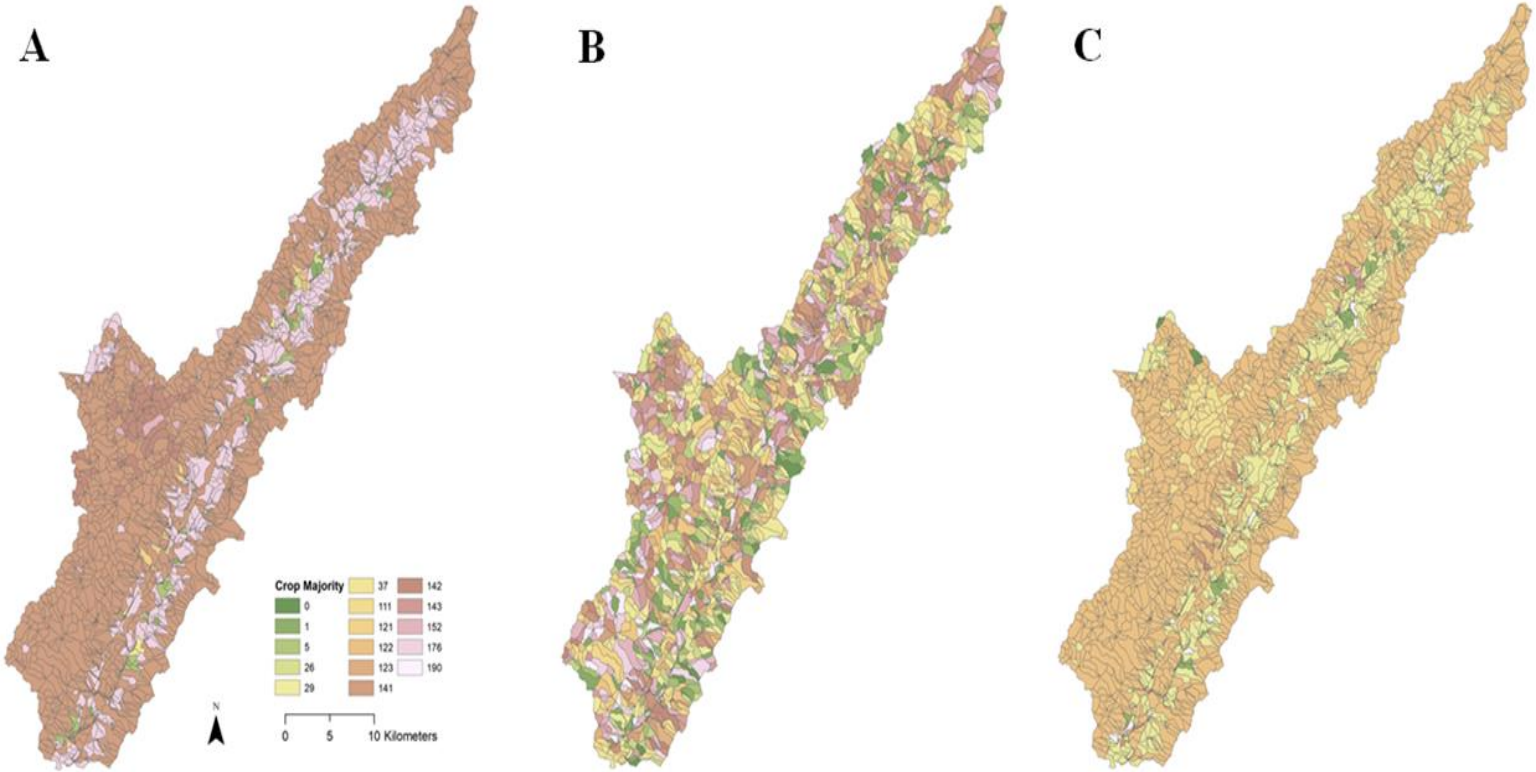


Figure 17. Crop majorities discretized to the Sequatchie Subbasin sub–catchments displaying progression of crop majorities from 2016 (A), 2019 (B), and 2021 (C). A majority value of 0 represents background.

In the study area, a management practice considers the temporal rotations of grassland/meadow, forest, urban, water, and the row crops of soybean and corn to develop a sequence of operations based on the statistical majority for that year. For example, a sequence of S; S; S; S; C; S refers to soybean being the majority from 2016 to 2019, corn for 2020, and soybean for 2021 (see example in Table 7). The other aspect of management practices are the associated farming operations, these include activities such as fertilization, spraying, and harvesting crop (see example in Table 8).

Table 7. Example of four sequences of operation illustrating the annual land use/land cover majority assigned to each grid code (sub–catchment) in the Sequatchie Subbasin.

AnnAGNPS Cell ID	Majority	Sequence of Operation
1003	5; 1; 5; 1; 1; 1	C; S; C; S; S; S
1023	190; 190; 190; 141; 141; 141	G; G; G; F; F; F
1073	141; 176; 176; 176; 176; 176	F; G; G; G; G; G
1732	141; 141; 122; 122; 141; 122	F; F; U; U; F; U

Table 8. Example of annual farming operations associated with distinctive managements for the row crops of corn and soybeans in the Sequatchie Subbasin during 2016.

Corn	**Soybeans**
4/30: Sprayer	5/25: Sprayer
5/1: Fertilizer Application	5/26: Fertilizer Application
5/10: Plant	6/5: Plant
5/11: Sprayer	6/25: Sprayer
6/10: Fertilizer Application	10/27: Harvest
9/27: Harvest	

The conversion of land use/land cover layers for each sub–catchment into unique management schedules with associated farming operations requires RUSLE2 templates, the RITA AGNPS tool, and a custom Python script to develop the management schedule and operation data sections required for AnnAGNPS. A total of 72 distinctive management schedules and respective farming operations were created and assigned to the sub–catchments of the Sequatchie Subbasin. No management schedules or operations were considered for the five surrounding karst valleys since land cover majorities were forest and grassland/meadows.

Climate Stations

The primary controlling parameters for hydrological models are precipitation and temperature. This information is provided, along with other climatic parameters, into two types of input files: primary and secondary. The primary climate file consists of a complete file without any gaps (i.e., a backup file). Secondary files can contain gaps and are often created using observed information while primary files are created using a combination of observed and synthetic climate data. Any informational gaps within the climate/weather measurements of the secondary files are substituted from the primary file, which contain additional observed data such as dew point, wind speed, and solar radiation. Daily summaries of precipitation and temperature information are downloaded as tabular datasets from the National Oceanic and Atmospheric Association (NOAA) agency. Synthetic weather/climate data used in the primary climate file is created from the Generation of Weather Elements for Multiple Applications (agGEM) simulation program.

AgGEM uses existing long–term–statistical parameter files of synthetic meteorological data (i.e., statistical seeds of the closest city in relation to the study area) to simulate 10 years of observed data, generating the climate station template and no–gap primary station climate daily output in the correct AnnAGNPS format. For this research, the statistical seed input of Chattanooga, Tennessee is utilized for simulation purposes and used to determine the study area's annual average precipitation of 1,349 mm with 119.8 wet days per year. Subsequently, the primary and secondary climate station files are combined, organized, and quality controlled to create a database of daily climate coverages (e.g., date, elevation, precipitation, maximum temperature, and minimum temperature) over the entire study period (January 1st, 2016 to December 31st, 2021) that correspond to one of the ten study area climate stations summarized in Table 9. The XY (longitude and latitude) data of each climate station file is then added to GIS, creating a new point layer displaying the spatial assignment of the climate stations (Figure 18). These zones of influence allocate space around the nearest station to assign a secondary climate file to AnnAGNPS cells.

Table 9. Spatial location and elevation of the climate stations used in the study area.

Climate Station	County Location and Station Name	Latitude	Longitude	Elevation (m)
1	Cumberland Co., Crossville 7.2 SE	35.8774	-84.9464	165.750
2	Bledsoe Co., Pikeville 7.5 SW	35.5442	-85.3046	194.158
3	Bledsoe Co., Pikeville 13.2 NE	35.7539	-85.0452	106.101
4	Bledsoe Co., Pikeville 5.4 N	35.6851	-85.2013	158.130
5	Bledsoe Co., Pikeville 10.3 N	35.7544	-85.1711	165.354
6	Bledsoe Co., Graysville 5.5 WNW	35.4711	-85.1761	180.503
7	Marion Co., Whitwell 0.3 SE	35.1980	-85.5143	68.001
8	Marion Co., Whitwell 4.5 ENE	35.2290	-85.4457	71.354
9	Sequatchie Co., Sequatchie–Whitwell	35.2897	-85.4736	94.762
10	Bledsoe Co., Pikeville (Primary Station)	35.5982	-85.1939	80.254

Note: Climate stations 1 through 9 are secondary stations.

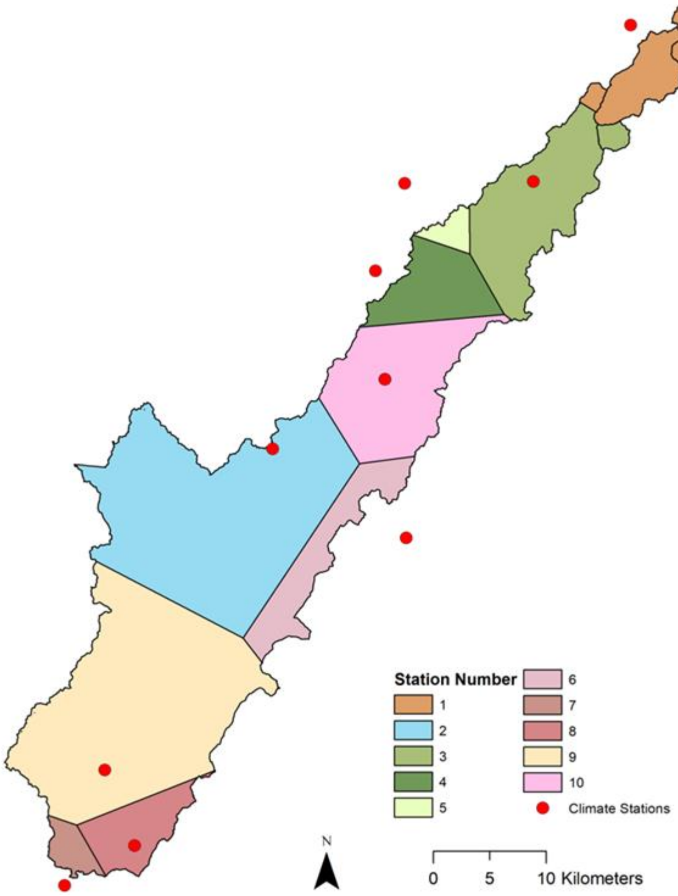


Figure 18. Assignment of climate station zones of influence to the Sequatchie Subbasin and surrounding karst valleys based on Thiessen polygon analysis.

Observed Data

Observed daily streamflow conditions at the outlet of the Sequatchie Subbasin is monitored by United States Geologic Survey (USGS) streamgage site 03571000 on the Sequatchie River near Whitwell, TN (see photo in Appendix B). Contribution of discharge or surface runoff from precipitation plus baseflow (groundwater) at this gage location includes the upstream flowline of the Sequatchie River to the headwaters at Head of Sequatchie Spring and contributory flow from the closed basin of Grassy Cove and surrounding karst valleys. The Sequatchie Subbasin, an 8–digit Hydrologic Unit Code (HUC) cataloging unit, is composed of seventeen 12–digit sub–watersheds that map and capture the tributary systems of the Sequatchie River. The study area consists of ten of the seventeen sub–watersheds located throughout Grassy Cove, the Upper Sequatchie River, and portions of the Lower Sequatchie River (Figure 19), with their total area of observed streamflow summarized in Table 10. Additionally, the karst valleys of Bat Town Cove and McClough Hollow are positioned in the Upper Daddy's Creek sub–watershed and both Swagerty Cove and Little Cove located in Upper Whites Creek sub–watershed (not shown in Figure 19 or Table 10).

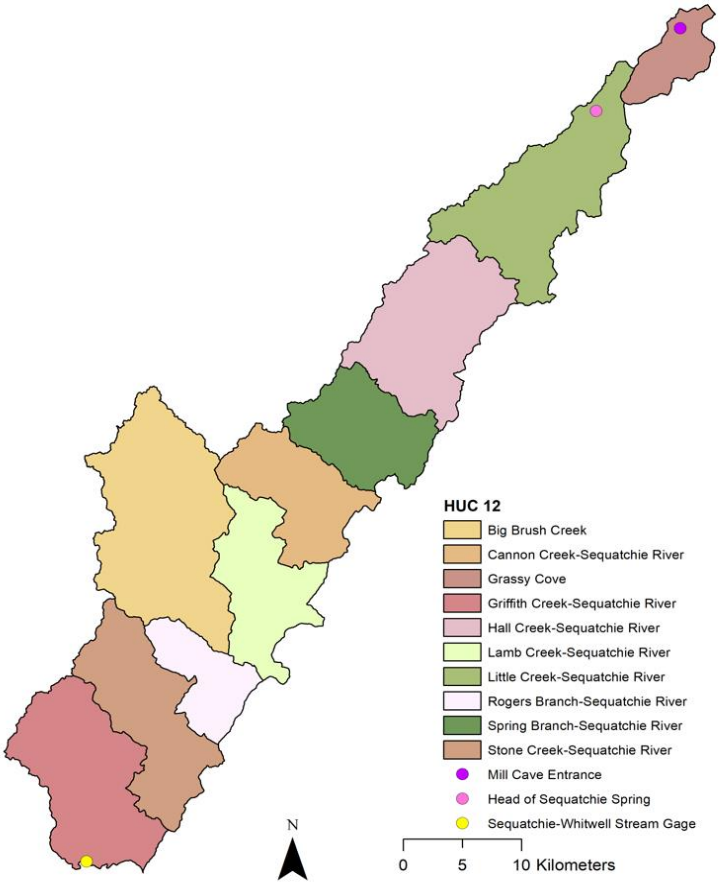


Figure 19. Map of the delineated Sequatchie Subbasin in order to capture the Sequatchie River tributary systems.

Table 10. Summary of the drainage basin hydrologic surface areas for each 12–digit HUC within the Sequatchie Subbasin.

Name	**12–Digit HUC**	**Surface Area (km²)**
Griffith Creek–Sequatchie River	060200040303	130.86
Rogers Branch–Sequatchie River	060200040301	52.42
Stone Creek–Sequatchie River	060200040302	110.41
Hall Creek–Sequatchie River	060200040103	130.53
Grassy Cove	060200040101	34.67
Little Creek–Sequatchie River	060200040102	138.02
Spring Branch–Sequatchie River	060200040104	85.06
Cannon Creek–Sequatchie River	060200040105	83.83
Big Brush Creek	060200040106	181.45
Lamb Creek–Sequatchie River	060200040107	89.17
		Total: 1,036.42

Baseflow is generally groundwater discharge and accounts for the portion of streamflow contained in subsurface flow during the dry season. Water below the soil profile that interacts with stream channels is not simulated in the AnnAGNPS model. Accurate correlation between simulated streamflow and observed USGS streamgage data requires the conversion of observed discharge from cubic feet per second (ft^3/s) to cubic meters per day (m^3/d) and separation of digital baseflow using modified RStudio and Python scripts. The observed total streamflow output files include monthly and annual time–series variations of the significance level (alpha) between the values of .100 and .999 (see example in Table 11), where an alpha of .975 is considered a default value. The observed total streamflow measured in megagrams (Mg) versus time is then graphed (Figure 20) and the formatted observed time–series variations are ready for statistical analyses using pairwise comparison against AnnAGNPS simulated streamflow.

Table 11. Example of observed total streamflow measured in Mg and the corresponding annual time–series variation of alpha values. * Denotes removal of baseflow.

	Gage Annual Time–Series Variations of Alpha				
	.100	**.300**	**.500**	**.800**	**.975**
Date	**Observed Streamflow* (Mg)**				
12/31/2016	103,113,690.4	128,097,518.9	156,452,685.3	218,413,717.8	324,889,201.6
12/31/2017	142,061,873.4	181,704,694.9	229,582,160.5	331,549,172.4	509,064,085.5
12/31/2018	226,090,930.5	287,832,729.1	364,483,039.9	524,207,257.4	739,044,915.4
12/31/2019	211,518,808.6	264,372,967.4	326,478,089.5	475,985,513.5	707,026,232
12/31/2020	223,096,308.1	281,587,201.1	347,303,582.6	484,788,493.5	714,003,588.8
12/31/2021	151,017,448.1	191,480,527.1	243,069,944.1	355,516,803.2	516,588,044.3

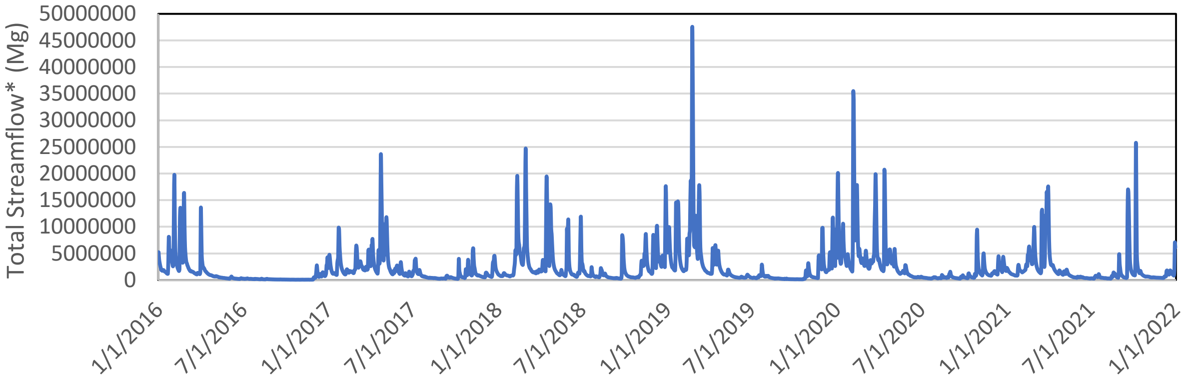


Figure 20. Hydrograph of observed total streamflow measured in Mg over the six–year study period (2016 to 2021). * Denotes removal of baseflow.

AnnAGNPS

Primary Inputs and Simulation Preparation

AnnAGNPS simulation requires the following primary data section inputs in comma separated value (csv) file format: TopAGNPS reach and cell data sections; GIS and RUSLE2 management schedule and management operation data sections; NITA soil data and soil layer data sections; and the climate data section file. These input files and other csv files are updated and become the sub–contents within the template folders (Table 12) of climate, general, simulation, watershed, and the AnnAGNPS master control file. The initial Curve Number (CN) values used in the management schedule and rocurve sub–content folders are obtained from TDOT (2021). The master control file utilizes links to the template folders enabling content pulls from multiple inputs to create the AnnAGNPS simulated gaging station output files.

Table 12. Summary of the template folders and sub–content file inputs required for the master control file and initial simulation of AnnAGNPS.

Template Folder	**Sub–Content File Inputs**	**Required Updates**
Climate	Primary Daily Climate Primary Climate Station Secondary Daily Climate Secondary Climate Station	- Assignment of drainage basin *Primary* and *Secondary Daily Climate* and *Climate Stations* based on spatial relationship between AnnAGNPS cells and station location using Thiessen polygons
General	Management Field	- Requires RUSLE2 inputs to *Land Use Type ID* and *Management Schedule ID* column
	Management Schedule	- Requires RUSLE2 inputs to *Management Schedule ID* and initial CN values to *CN ID* column
	Soil Layers Soil Data	- Assignment of NITA *Unique Soil ID* for AnnAGNPS cells column
	Rocurve	- Requires initial CN values
	Management Operation	- No required updates
	Crop Data	- No required updates
	Crop Growth Data	- No required updates
	Cell ID	- No required updates
	Non Crop Data	- No required updates
Simulation	Simulation Period	- Requires inputs to *Simulation Begin* and *End* columns in month–date–year format
Watershed	Watershed Data	- Requires inputs to *latitude*, *longitude*, and *watershed location* and *name*
	Cell Data	- Assignment of NITA *Unique Soil ID* and *Management Field ID* (CDL majority) column
	Reach Data	- No required updates

Initial AnnAGNPS Simulation

Initial simulation of the Sequatchie Subbasin creates the AnnAGNPS gaging station file containing requested mass and area unit data of reach information, watershed outlet total streamflow, drainage area, precipitation, and pollutant and sediment loadings of the simulation period. These files are used to create time–series summaries of monthly and annual simulated total streamflow values using Python script, similar to those for observed streamgage data. Once the summary files are complete, a pairwise correlation between simulated and observed data is performed to define the relationship between the two variables, calibrate the datasets, and determine the final CN values.

Calibration of Data and Adjusting Curve Numbers

Calibration of AnnAGNPS output tables in watershed modeling is an important step to improve model performance and ensure confidence in simulation quality. This iterative process compares the model results of runoff with observed (actual) runoff (i.e., total streamflow in Mg) by adjusting input parameters of land use categories to obtain results that best represent reality. This approach of estimating surface runoff from agricultural watersheds is based on the empirical method developed by the Natural Resources Conservation Service (NRCS) CN model. CN's represent the estimation of runoff producing potential as a function of cumulative precipitation, hydrological soil group, land use, and the antecedent moisture condition (AMC) of the watershed (TDOT, 2021). AMC groups range from I to III, with I being dry, II as normal or average, and III representing wet conditions. Runoff CN values range from 0 to 100 and characterize the runoff properties of a particular ground cover type, its hydrologic condition, and the soil group classification.

CN values in this study were incrementally increased by 3, 5, and 7 (i.e., CN +3, CN +5, and CN +7) for all land use/land cover majority types (G, F, U, S, and C) from their initial values (Table 13), for a total of 4 scenarios in the Sequatchie Subbasin (Table 14). Each iteration of CN value increase are succeeded by a new AnnAGNPS simulation to create an updated gaging station file, the conversion of simulated total streamflow to monthly and annual time–series summary files, and subsequent pairwise correlation with observed values of gage .100 through .999 summary files. These results are used to visualize hydrographs for each scenario (see example in Figure 21) and evaluated using various statistical metrics to determine the final CN values and alpha variation for AnnAGNPS modeling. No alternative scenarios or CN value iterations were considered for the five surrounding karst valleys.

Table 13. Incremental increase of Curve Numbers from initial values for each hydrological soil group and corresponding majority land use/land cover type in the Sequatchie Subbasin.

Hydrological Soil Group	**Land Use/Land Cover Majority**				
	Grassland	**Forest**	**Urban**	**Soybeans**	**Corn**
	Initial CN Value				
A	30	36	57	60	56
B	58	60	72	68	66
C	71	73	81	74	71
D	78	79	86	76	75
	CN +3 Value				
A	33	39	60	63	59
B	61	63	75	71	69
C	74	76	84	77	74
D	81	82	89	79	78
	CN +5 Value				
A	35	41	62	65	61
B	63	65	77	73	71
C	76	78	86	79	76
D	83	84	91	81	80
	CN +7 Value				
A	37	43	64	67	63
B	65	67	79	75	73
C	78	80	88	81	78
D	85	86	93	83	82

Table 14. Simulated annual time–series scenarios of total streamflow summaries created from the incremental increase of CN values during calibration of the Sequatchie Subbasin.

	Simulated Annual Time–Series Scenarios			
	Initial CN Value	**CN +3 Value**	**CN +5 Value**	**CN +7 Value**
Date	**Simulated Total Streamflow (Mg)**			
12/31/2016	209,537,556.1	246,660,362	273,965,768.6	303,246,140.1
12/31/2017	255,589,280.9	307,991,486.8	347,362,351.2	390,537,911.1
12/31/2018	369,795,301.2	430,006,009.6	474,104,023.4	521,561,891.7
12/31/2019	394,879,403	461,485,212.4	510,610,894.6	563,745,747
12/31/2020	386,710,816.2	449,846,080.6	496,419,532.8	546,935,447.7
12/31/2021	347,410,057.1	402,375,149.7	442,897,808.4	486,762,461.7

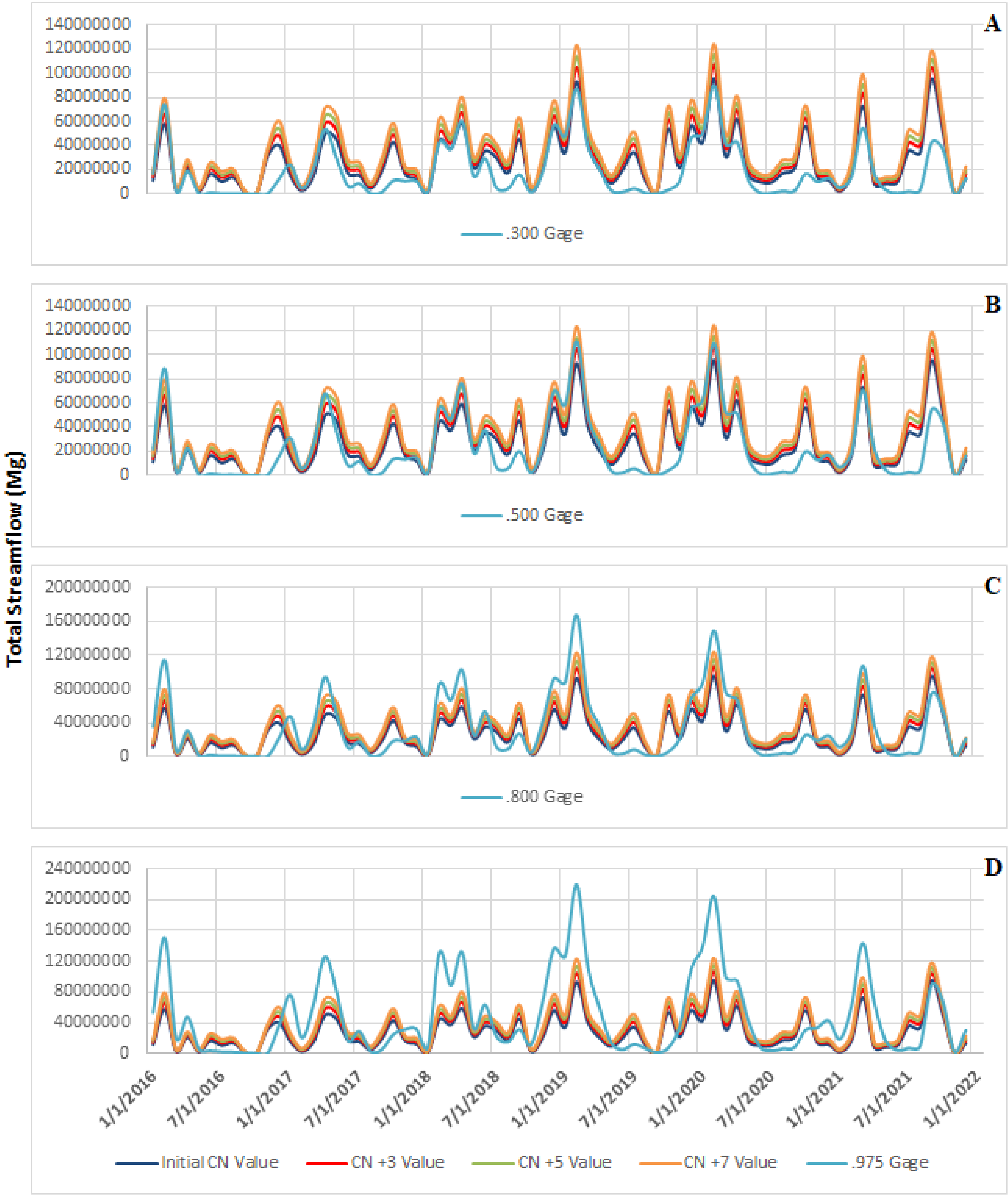


Figure 21. Comparison of monthly simulated and observed (gage) total streamflow (Mg) with varying alpha values during the study period (2016 to 2021). Simulated scenarios of initial CN value (dark–blue line), CN +3 (red line), CN +5 (green line), CN +7 (orange line), and observed (light–blue line) .300 gage (A), simulated scenarios and observed (.500 gage) (B), simulated scenarios and observed (.800 gage) (C), and simulated scenarios and observed (.975 gage) (D).

Statistical Metrics

Evaluations of hydrological model performance often utilize multiple statistical analysis methods to visualize and investigate patterns, relationships, and trends of quantitative data. Quantitative data typically measures numerical values between model–simulated/predicted data (P) and observed data (O) under identical conditions, where the total number of observations (n) during a given incremental duration of time (i) evaluates the model's performance. Pairwise comparisons are statistical techniques often used to evaluate the spatial and/or temporal relationship of these variables from a set of inputs, although correlation–based measures like Pearson's product–moment correlation coefficient (r) and the coefficient of determination (R^2) are overly sensitive to extreme outliers (Legates and McCabe, 1999).

In model evaluation both graphical techniques (e.g., hydrographs and scatterplots) and quantitative statistical measures are recommended to determine model efficacy (Moriasi et al., 2007; Daggupati et al., 2015). These measures should include the central tendency concerning the mean of values ($\bar{P}$ and $\bar{O}$) and their frequency distribution, r, R^2, the Nash–Sutcliffe efficiency (NSE) model evaluation statistic (Nash and Sutcliffe, 1970), average tendency of P and O expressed as a percentage (PBIAS), and the root mean square error– (RMSE) observations standard deviation ratio (RSR). Model evaluation in this study included comprehensive quantitative comparisons of simulated streamflow (runoff) against observed values at annual and monthly time–scales, with an example of results summarized in Table 15 utilizing statistical measurements of r, R^2, NSE, PBIAS, and RSR.

Table 15. Summary of statistical metrics used to determine model performance for simulated and observed total streamflow (runoff) comparisons at annual and monthly time–scales.

Statistical Metric	**Time–Scale**					
	Annual			**Monthly**		
	Simulated Mean (Mg)					
	CN +3	**CN +5**	**CN +7**	**CN +3**	**CN +5**	**CN +7**
	383,060,717	424,226,730	468,798,267	31,921,726	35,352,227	39,066,522
	.500 Gage Observed Mean (Mg)					
	277,894,917			23,157,910		
r	0.91	0.91	0.92	0.81	0.82	0.82
R²	0.83	0.84	0.84	0.66	0.67	0.67
NSE	-0.94	-2.09	-3.40	0.49	0.44	0.37
PBIAS	15.91	20.84	25.57	27.45	34.49	40.72
RSR	1.39	1.76	2.10	0.72	0.75	0.80
	.800 Gage Observed Mean (Mg)					
	398,410,160			33,200,847		
r	0.92	0.92	0.93	0.80	0.80	0.81
R²	0.85	0.86	0.86	0.64	0.65	0.65
NSE	0.63	0.66	0.24	0.18	0.31	0.38
PBIAS	-1.96	3.14	8.12	-4.01	6.09	15.01
RSR	0.61	0.58	0.87	0.91	0.83	0.79
	.975 Gage Observed Mean (Mg)					
	585,102,678			48,758,557		
r	0.92	0.93	0.93	0.76	0.77	0.77
R²	0.85	0.86	0.87	0.58	0.59	0.59
NSE	-6.62	-3.32	-1.15	-1.42	-0.88	-0.48
PBIAS	-20.87	-15.94	-11.04	-52.74	-37.92	-24.81
RSR	2.76	2.08	1.47	1.56	1.37	1.22

Pearson’s product–moment correlation coefficient (r) is a benchmark measure of linear dependence between predicted and observed data, where the quality of fit is based on summarizing the plot of the two variables into a single number between -1 and 1. The direction and strength of correlation is represented by the degree of index with values of -1 to 0 associated with a negative slope, 0 to 1 having a positive slope, and r equal to 0 having no linear relationship. Correlation values close to 1 or -1 constitute two quantitative variables that are strongly related (Moriasi et al., 2007). Pearson’s r is given as:

$$r = \frac{\sum_{i=1}^{n} ((O_i - \bar{O})(P_i - \bar{P}))}{\sqrt{\sum_{i=1}^{n} (O_i - \bar{O})^2 \sqrt{\sum_{i=1}^{n} (P_i - \bar{P})^2}}}$$

The coefficient of determination (R^2) provides a measure of how well the observed data is replicated or predicted by the model (Figure 22). R^2 values range from 0 to 1 and describe the percentage of total variation explained by the relationship between two variables (Legates and McCabe, 1999), where the mean of *y*–axis values varies from the total squared area between points of a line, or *x*. A small percentage of error variance indicates the line is a good fit and R^2 values will be close to 1, although large error percentiles are poor linear fits with values closer to 0, meaning very little variation in *y* is represented in *x*. Typically an R^2 value above 0.60 is acceptable (Moriasi et al., 2015). R^2 is given as:

$$R^2 = \left\{ \frac{\sum_{i=1}^{n} ((O_i - \bar{O})(P_i - \bar{P}))}{\sqrt{\sum_{i=1}^{n} (O_i - \bar{O})^2 \sqrt{\sum_{i=1}^{n} (P_i - \bar{P})^2}}} \right\}^2$$

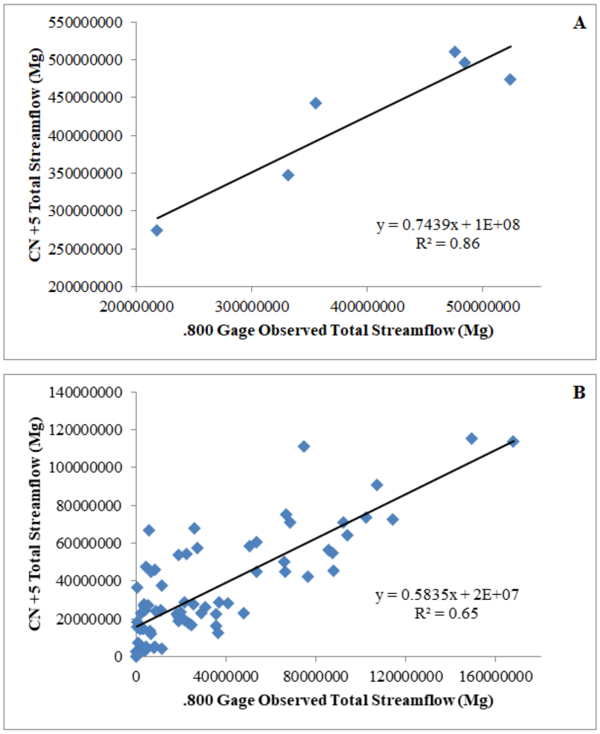


Figure 22. Comparison between the CN +5 simulated and .800 gage observed total streamflow (Mg) values at annual (A) and monthly (B) time–scales. The obtained R^2 values were .86 and .65 respectively, explaining 86% and 65% of data variance between observed and the model.

The Nash-Sutcliffe efficiency (NSE) hydrologic model evaluation statistic is defined as an objective function that best reflects the overall fit of hydrographs (Momm et al., 2019). NSE, or the coefficient of efficiency (Nash and Sutcliffe, 1970), is a dimensionless normalized statistic ranging in value from negative infinity (∞) to 1, with 1 being most optimal since higher values indicate greater agreement between observations and simulations (Legates and McCabe, 1999). This statistical measure determines the predictability of the model and how well a simulated versus observed data plot fits the line of equality (1:1 line). NSE values of 0 indicate model simulations and the observed mean have similar explanatory power but values less than 0 signify the observations are more accurate predictors than the hydrological model. Generally, NSE values > 0.50 are acceptable levels of model performance (Moriasi et al., 2015) and given as:

$$\text{NSE} = 1\text{-}\left[\frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2}\right]$$

The percent bias (PBIAS) is a measurement of the average tendency of simulated data to be a greater or lesser value than their observed data counterpart. Bias is introduced to the statistical result by overestimation of the model resulting in positive values or model underestimation indicated by negative values. PBIAS results are given as a percentage (%), ranging in values from -∞ to ∞, with optimal values being 0 since lower–weighted values indicate higher accuracy in model simulation (Moriasi et al., 2007; 2015). PBIAS is given as:

$$\text{PBIAS} = \left(\frac{\sum_{i=1}^{n} O_i - P_i}{\sum_{i=1}^{n} O_i}\right) \times 100$$

The root mean square error– (RMSE) observations standard deviation ratio (RSR) statistic is a relatively new hydrological modeling performance measure developed by Moriasi and others (2007). RMSE is the standard deviation (SD) measurement of residual distance from the 1:1 line (i.e., the concentration of observed data points around the regression line). RSR results range in value from 0 to ∞ with an optimal value recognized at 0, as it is commonly accepted that low RMSE and RSR values are indicative of better model performance (Moriasi et al., 2015). RSR is given as:

$$\text{RSR} = \frac{\sqrt{\sum_{i=1}^{n}(O_i - P_i)^2}}{\sqrt{\sum_{i=1}^{n}(O_i - \bar{P})^2}}$$

Final Integrated Simulation

Hydrological model evaluation utilizing the previous statistical metrics determined the appropriate alpha variation for baseflow removal and final CN adjustment values were .800 and CN +5 respectively, at both annual and monthly temporal scales. After adequate calibration, the curve number data section is updated with the final CN values (Table 16) and a new simulation is executed to create the final Sequatchie Subbasin gaging station file. Results are converted from daily into monthly and annual time–series and model efficiency is statistically evaluated both quantitatively (Table 17) and graphically (Figure 23) with the corresponding observed gage .800 temporal summary files in preparation for the final integrated simulation.

Table 16. Initial and final CN values following calibration for each land use/land cover majority in the Sequatchie Subbasin.

Land Use/Land Cover Majority	Hydrologic Soil Groups							
	Initial CN Values				Final CN Values			
	A	B	C	D	A	B	C	D
Grassland	30	58	71	78	35	63	76	83
Forest	36	60	73	79	41	65	78	84
Urban	57	72	81	86	62	77	86	91
Soybeans	60	68	74	76	65	73	79	81
Corn	56	66	71	75	61	71	76	80

Table 17. Statistical parameters of model performance for the final calibrated Sequatchie Subbasin summary file when compared to .800 gage file.

	Statistical Metric						
Time–Scale	**CN +5 Mean (Mg)**	**.800 Gage Mean (Mg)**	**r**	**R^2**	**NSE**	**PBIAS**	**RSR**
Annual	424,226,730	398,410,160	0.92	0.86	0.66	3.14	0.58
Monthly	35,352,227	33,200,847	0.80	0.65	0.31	6.09	0.83

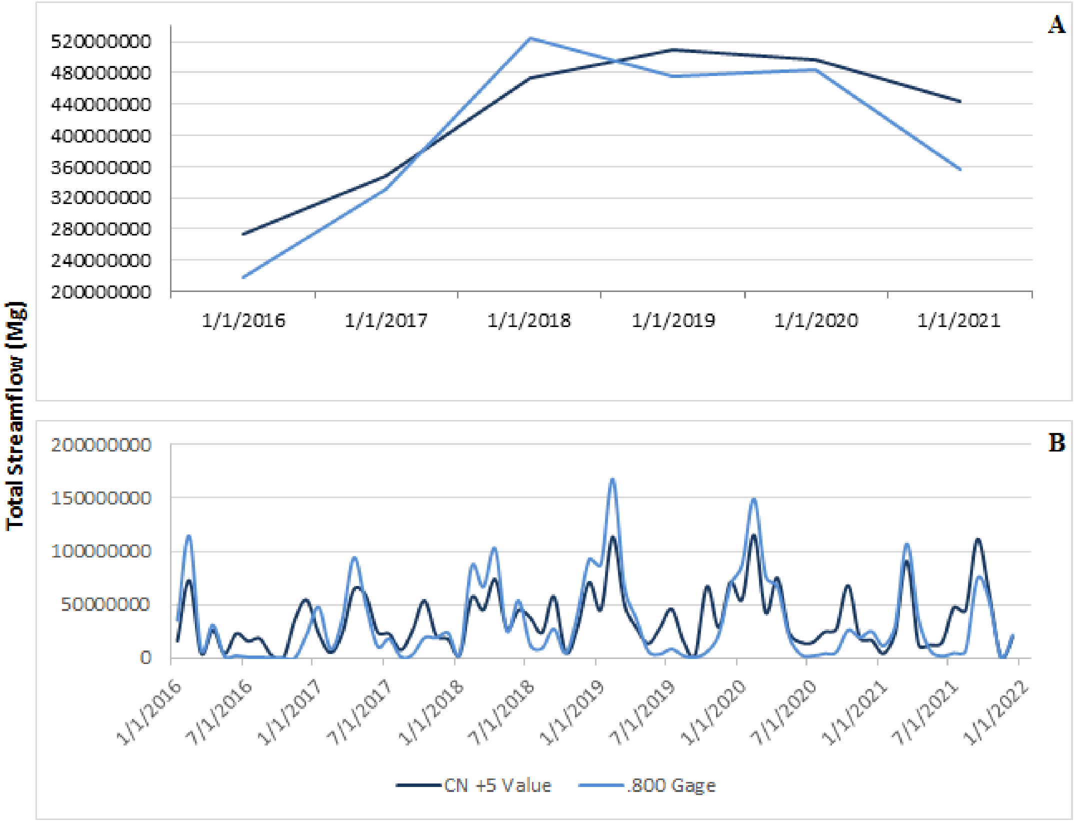
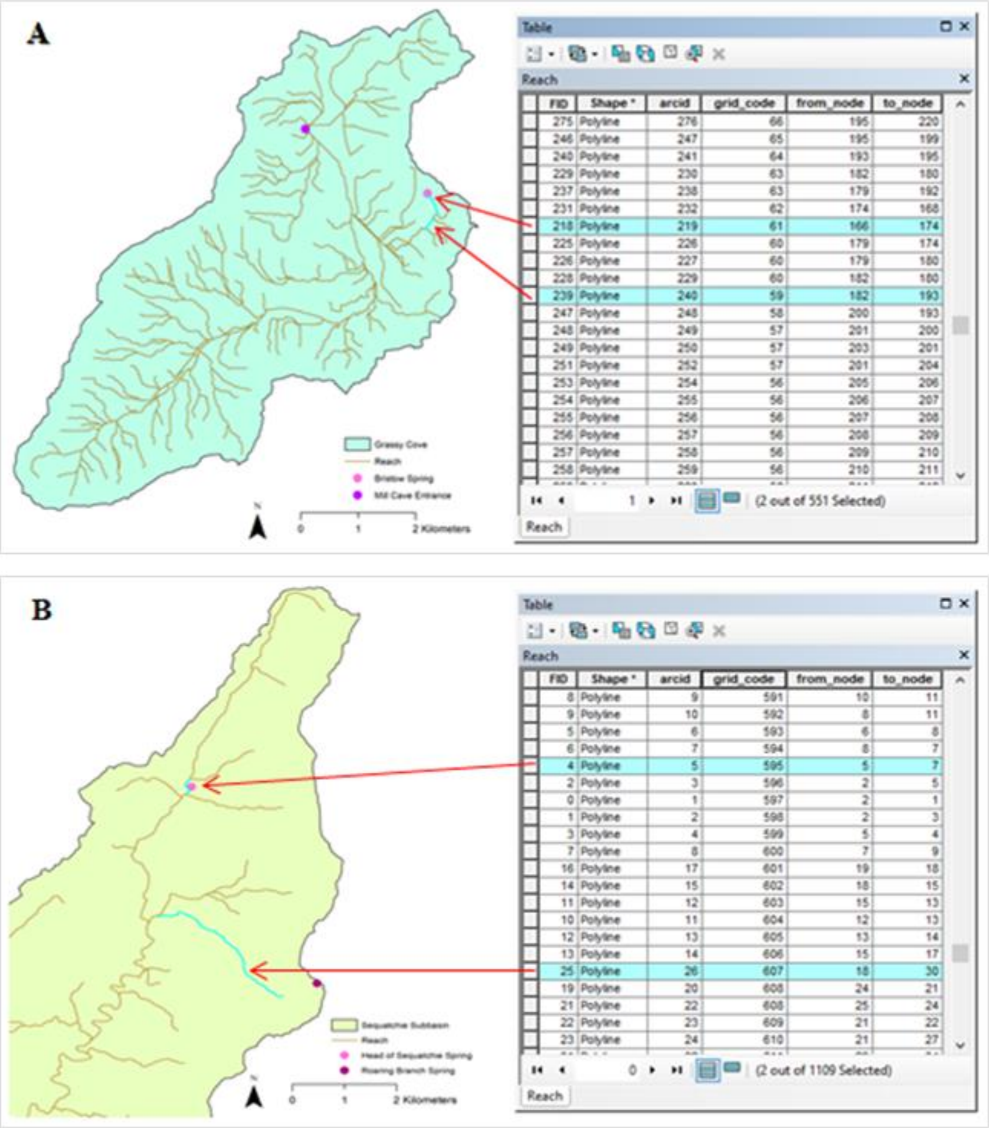


Figure 23. Time–series of annual (A) and monthly (B) CN +5 compared with observed .800 gage summary files of total streamflow (Mg) from January 2016 through December 2021.

The next step is integration of Little Cove with Grassy Cove since the contributing subsurface flow in these two karst valleys originates in Little Cove, travels underground from Little Cove swallet to resurge at Bristow Spring in Grassy Cove, then flows to the Mill Cave entrance and rises again at Head of Sequatchie Spring. The second integration combines the contributing subsurface flow originating in Swagerty Cove, traveling underground from Swagerty Cove swallet and resurges at Roaring Branch Spring to join as a tributary stream to the overland flow of the Sequatchie River. Subsurface flow integration is accomplished by converting the Grassy Cove and Sequatchie Subbasin *reach* raster grids into GIS layers to identify the grid code ID that spatially corresponds to each reach input (Figure 24). Use of the new AnnAGNPS feature of integrating separated simulations allows outputs from upstream simulations to be entered as inputs in downstream simulations.

Once identified, these inputs (AnnAGNPS stream code IDs) are added to their respective AnnAGNPS watershed data sections and the integrated Grassy Cove–Little Cove AnnAGNPS simulation is executed. The goal is to account and simulate the contribution of these coves to the Sequatchie Subbasin. After completion, the new integrated Grassy Cove–Little Cove and remaining karst valley (Bat Town Cove, McClough Hollow, and Swagerty Cove) previously simulated gaging station files are added as inputs to the Sequatchie Subbasin AnnAGNPS watershed and executed. The final step is to compare the total streamflow against the corresponding observed gage to statistically re–evaluate the hydrological efficiency of the model.



FID	Shape *	arcid	grid_code	from_node	to_node
275	Polyline	276	66	195	220
246	Polyline	247	65	195	199
240	Polyline	241	64	193	195
229	Polyline	230	63	182	180
237	Polyline	238	63	179	192
231	Polyline	232	62	174	168
218	Polyline	219	61	166	174
225	Polyline	226	60	179	174
226	Polyline	227	60	179	180
228	Polyline	229	60	182	180
239	Polyline	240	59	182	193
247	Polyline	248	58	200	193
248	Polyline	249	57	201	200
249	Polyline	250	57	203	201
251	Polyline	252	57	201	204
253	Polyline	254	56	205	206
254	Polyline	255	56	206	207
255	Polyline	256	56	207	208
256	Polyline	257	56	208	209
257	Polyline	258	56	209	210
258	Polyline	259	56	210	211

(2 out of 551 Selected)

FID	Shape *	arcid	grid_code	from_node	to_node
8	Polyline	9	591	10	11
9	Polyline	10	592	8	11
5	Polyline	6	593	6	8
6	Polyline	7	594	8	7
4	Polyline	5	595	5	7
2	Polyline	3	596	2	5
0	Polyline	1	597	2	1
1	Polyline	2	598	2	3
3	Polyline	4	599	5	4
7	Polyline	8	600	7	9
16	Polyline	17	601	19	18
14	Polyline	15	602	18	15
11	Polyline	12	603	15	13
10	Polyline	11	604	12	13
12	Polyline	13	605	13	14
13	Polyline	14	606	15	17
25	Polyline	26	607	18	30
19	Polyline	20	608	24	21
21	Polyline	22	608	25	24
22	Polyline	23	609	21	22
23	Polyline	24	610	21	27

(2 out of 1109 Selected)

Figure 24. Conversion of Grassy Cove (A) and Sequatchie Subbasin (B) reach raster grids into GIS layers in order to identify the stream code IDs that correspond to each input reach.

CHAPTER IV

RESULTS AND DISCUSSION

Statistical Analysis and Surface Hydrology Modeling

Complete assessment of hydrological model performance should utilize comprehensive techniques and numerous statistical measures to determine the predictive ability of the model to simulate reality. After the calibration process, acceptable statistical performance measures were attained for the Sequatchie Subbasin at the annual time–scale, where: $R^2 = 0.86$, NSE = 0.66, PBIAS = 3.14%, and RSR = 0.58. Although, combinations of very good to unsatisfactory performance measures were obtained at the monthly time–scale, showing that: $R^2 = 0.65$, NSE = 0.31, PBIAS = 6.09%, and RSR = 0.83. Interpretation of these statistical results are based on Table 4 and Table 9 recommended final performance evaluation criteria of Moriasi and others (2007, 2015) respectively, as both monthly temporal–scale statistics for NSE >0.50 and RSR > 0.70 are considered unsatisfactory performance ratings for flow output responses. This could possibly be explained by uncertainties in characterization of the contribution of baseflow and subsurface flow to streamflow. Additionally, potential latency in karst flow combined with rainfall variability adds to this uncertainty.

During the 6–year simulation period, the mean daily, monthly, and annual rainfalls for all drainage basins ranged from 2.05 to 5.07 mm, 110.22 to 133.87 mm, and 1,322.58 to 1,606.38 mm, respectively. Table 18 summarizes the mean rainfalls at each temporal series and total rainfall at the outlet for the Sequatchie Subbasin and 5 surrounding karst valleys. Simulated rainfall summaries at monthly time–scales are illustrated in Figure 25 to help visualize rainfall data and determine which months are rainy versus dry periods for each drainage basin. The climate stations assigned to the Sequatchie Subbasin (A) are stations 1 through 10, while Grassy

Cove, Little Cove, McClough Hollow, and Bat Town Cove (B) all utilize climate station 1, and Swagerty Cove (C) is simulated from climate station 3.

Table 18. Mean daily, monthly, annual, and total rainfall in mm at each drainage basin respective outlet during the study period (2016 to 2021).

	Mean Daily Rainfall (mm)						Mean Monthly Rainfall (mm)	Mean Annual Rainfall (mm)	Total Rainfall at Outlet (mm)
Drainage Basin	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**			
Sequatchie Subbasin	2.96	4.35	4.70	4.99	5.07	4.14	133.87	1,606.38	9,638.30
Grassy Cove	2.80	3.92	4.28	4.12	4.56	3.30	117.13	1,405.57	8,433.43
Little Cove	2.78	3.89	4.26	4.10	4.56	3.30	116.61	1,399.37	8,396.24
Swagerty Cove	2.05	4.36	3.72	4.42	4.26	2.83	110.22	1,322.58	7,935.48
McClough Hollow	2.73	3.87	4.19	4.03	4.55	3.30	115.50	1,405.57	8,433.43
Bat Town Cove	2.75	3.89	4.23	4.08	4.56	3.30	116.15	1,393.74	8,362.46
								Total:	**51,199**

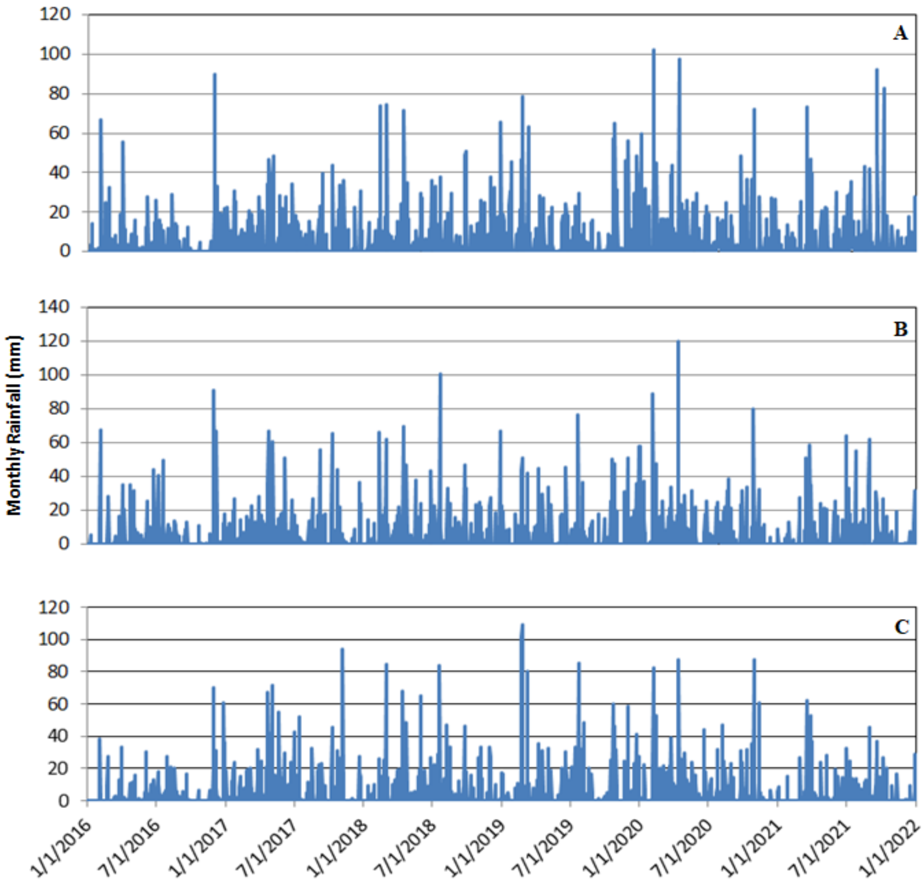


Figure 25. Monthly time–scale simulated rainfall (mm) summaries of the Sequatchie Subbasin (A), Grassy Cove, Little Cove, McClough Hollow, and Bat Town Cove (B), and Swagerty Cove (C).

AnnAGNPS gaging station output files for surface hydrological modeling are summarized in Table 19. The total drainage area of the Sequatchie Subbasin and surrounding karst valleys is 105,420 ha with an overall reach length of 1,299,425 m (1,299.4 km). McClough Hollow and Bat Town Cove contribute the least amount of total streamflow to the surface hydrological model outlet as expected, as these are the least mature and northeastern–most karst systems with the smallest surface area.

Table 19. Output summaries of AnnAGNPS gaging station files for surface hydrological modeling.

Drainage Basin	**Average Annual Total Streamflow (Mg)**	**Reach Length (m)**	**Drainage Area (ha)**
Sequatchie Subbasin CN +5	424,226,730	1,116,942.12	100,107.01
Grassy Cove	5,292,742.6	127,194.72	3,466.36
Little Cove	784,441.2	17,631.11	537.36
Swagerty Cove	549,894.7	15,528.91	574.93
McClough Hollow	82,973.9	8,923.69	312.66
Bat Town Cove	306,283.3	13,204.07	421.75
Total:	**431,243,066**	**1,299,425**	**105,420**

Ground–Surface Hydrology Model

The ground–surface hydrology AnnAGNPS simulation integrates the contributory subsurface and surface flow from the Sequatchie Subbasin and surrounding karst valleys. We then compared the simulated (predicted) to observed streamflow values at the USGS outlet. Statistical measures similar to the surface hydrology model are incorporated to determine the predictive ability of the model and measure how well the observed data is replicated. Acceptable statistical performance measures were again attained at the annual time–scale for the integrated ground–surface model, where: $R^2 = 0.86$, NSE = 0.63, PBIAS = 3.95%, and RSR = 0.61. This showed some moderate changes to the overall model though, with NSE values decreasing by 4.54% and PBIAS and RSR values increasing by 25.8% and 5.17%, respectively.

Combinations of very good to unsatisfactory performance measures were repeated at the monthly time–scale as well, showing that: $R^2 = 0.65$, NSE = 0.34, PBIAS = 7.61%, and RSR = 0.81. These results display a value increase in NSE and PBIAS by 9.67% and 24.95%, respectively, and RSR values decreasing by 2.40%. A summary of these results are shown in Table 20 and Figure 26. The value difference of annual and monthly time–scales between the surface calibrated (424,226,730 and 35,352,227 Mg) and ground–surface integrated (431,207,948 and 35,933,996 Mg) Sequatchie Subbasin summary files are 6,981,218 and 581,769 Mg respectively, or only an increase of 1.64% total streamflow both annually and monthly as summarized in Table 21 and graphed in Figure 27.

Table 20. Comparison of statistical parameters in the calibrated surface and integrated (ground–surface) Sequatchie Subbasin and observed .800 gage files.

	Statistical Metric						
Time–Scale	**Integrated Mean (Mg)**	**.800 Gage Mean (Mg)**	**r**	**R^2**	**NSE**	**PBIAS**	**RSR**
Annual	431,207,948	398,410,160	(0.92) 0.93	(0.86) 0.86	(0.66) 0.63	(3.14) 3.95	(0.58) 0.61
Monthly	35,933,996	33,200,847	(0.80) 0.80	(0.65) 0.65	(0.31) 0.34	(6.09) 7.61	(0.83) 0.81
		Percent Change:	**+1.08%** **--**	**--** **--**	**-4.54%** **+9.67%**	**+25.47%** **+24.95%**	**+5.17%** **-2.40%**

Table 21. Comparison of surface and ground–surface hydrology models mean total streamflow (Mg) and their percent change.

	Total Streamflow (Mg)			
Time–Scale	**Ground–Surface Mean (Mg)**	**Surface Mean (Mg)**	**Difference (Mg)**	**Percent Change**
Annual	431,207,948	424,226,730	+6,981,218	**+1.64%**
Monthly	35,933,996	35,352,227	+581,769	**+1.64%**

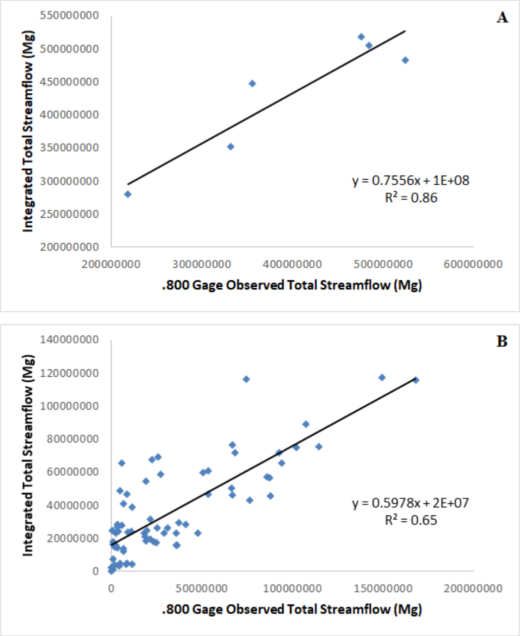


Figure 26. Comparison between the integrated and .800 gage observed total streamflow (Mg) values at annual (A) and monthly (B) time–scales. The obtained R² values again explain 86% and 65% of data variance between observed and the model.

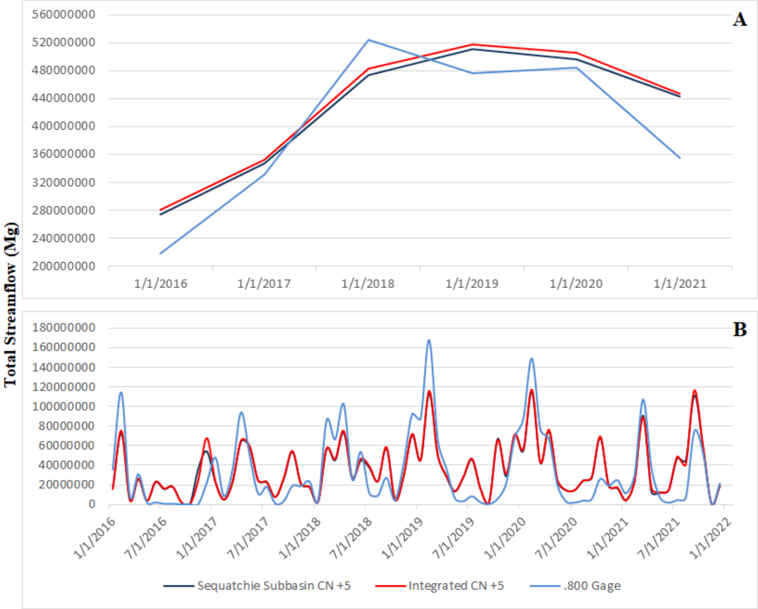


Figure 27. Time–series of annual (A) and monthly (B) calibrated surface and integrated Sequatchie Subbasin compared with observed .800 gage summary files of total streamflow (Mg) from January 2016 through December 2021.

Changes in statistical parameters describing comparisons between observed and simulated indicate the contribution of the coves to the overall simulated streamflow. Although, it is not possible to know with certainty the real contribution without field measured information at sinking streams and springs, as integrated simulation only provides an estimation of their contribution. It is important to recognize the uncertainties involved in this study.

One source of uncertainty is the AnnAGNPS model does not simulate baseflow, and when compared with measured streamgage values it must be removed. Baseflow removal is appropriate in most systems due to small amounts of variability over time, however, in karst systems this can be significantly impacted by direct sub–surface flow (other than common groundwater flow), adding to the uncertainty. Another potential source is the conceptual model used assumes all subsurface flow is accounted for. However, it is possible that additional unknown zones of recharge and discharge exists and/or are not documented. Finally, the latency (time between water entering the ground and resurgence) for each cove was not accounted for. This study used monthly and yearly time–scales to reduce this unpredictability, however, variations in flow latencies added to the uncertainty, especially at the monthly time–scale.

CHAPTER V

CONCLUSIONS

The purpose of this study was to apply and evaluate hydrological modeling technology to the northeastern portion of the Sequatchie Valley in Middle and East Tennessee in order to determine the model capacity of AnnAGNPS to predict runoff over a six–year study period (2016 to 2021). This topographically–complex area of the Cumberland Plateau is comprised of a series of interconnected karst systems that contribute subsurface and surface flow to the Sequatchie River. Using automated/semi–automated processes significantly reduced time and effort in developing AnnGNPS simulations in this large and intricate system.

After the required data inputs were sourced and integrated with GIS, the Sequatchie Subbasin and five surrounding cove watersheds were subdivided and discretized in order to route predicted streamflow and compare it with observed data to determine if correlation exists. To optimally achieve runoff prediction, the model was calibrated by adjusting Curve Number (CN) values and evaluated using comprehensive statistical metrics. Once the best–fit model produced satisfactory calibration results, an integrated simulation combining ground and surface flow was conducted in order to accept or reject the research hypothesis that the AnnAGNPS model can accurately predict contributing runoff from the adjacent karst valleys of the Sequatchie River.

Changes in statistical results between a scenario considering only the Sequatchie Subbasin and one considering additional subsurface flow to the Sequatchie Subbasin document the AnnAGNPS model capability of capturing cove contribution in runoff. Differences in statistical outcomes depict the level of uncertainties involved when studying surface and subsurface flows, as seen with monthly time–scale results ranging from unsatisfactory to good in comparison to acceptable annual results. Nonetheless, findings indicate the AnnAGNPS model to be appropriate in simulating karst systems with integrated surface and subsurface hydrology. Future work should focus on additional field measurements to enhance model calibration for each cove with the objective of improved characterization of volumes and latencies into the Sequatchie Subbasin.

BIBLIOGRAPHY

Adu, J.T., and Kumarasamy, M.V. (2018). Assessing Non-Point Source Pollution Models: A Review, *Polish Journal of Environmental Studies, 27:5*, pg 1-2.

Armstrong, R.N. and Martz, L.W. (2003). Topographic Parameterization in Continental Hydrology: A Study in Scale, *Hydrological Processes, 17*. https://doi.org/10.1002/hyp.1352, pg 3763-3767.

Bingner, R.L. and Theurer, F.D. (2001). AGNPS 98: A Suite of Water Quality Models for Watershed Use, In: *Proceedings of the Sedimentation: Monitoring, Modeling, and Managing*, 7th Federal Interagency Sedimentation Conference, United States Department of the Interior, Bureau of Land Management, p. VII 1-8.

Brahana, J.V., and Bradley, M.W. (1985). Delineation and Description of the Regional Aquifers of Tennessee - The Knox Aquifer in Central and West Tennessee, U.S. Geological Survey, *Water-Resources Investigations Report 83-4012,* pg 2-5.

Crawford Hydrology Laboratory (CHL). (2021). Karst Groundwater Investigation Research Procedures, pg 3-7.

Crawford, N.C. (1984). Karst Landform Development along the Cumberland Plateau Escarpment of Tennessee, Groundwater as a Geomorphic Agent, Binghamton Geomorphology Symposium 13. https://doi.org/10.4324/9781003028833, pg 3-19.

Crawford, N.C. (1987). The Karst Hydrogeology of the Cumberland Plateau Escarpment of Tennessee, (Part I), Center for Cave and Karst Studies, Dept. of Geography and Geology, Western Kentucky University, Chapter 1, pg 1-20.

Crawford, N.C. (1989). The Karst Hydrogeology of the Cumberland Plateau Escarpment of Tennessee, (Part II), Center for Cave and Karst Studies, Dept. of Geography and Geology, Western Kentucky University, Chapter 1, pg 1-40.

Daggupati, P., Pai, N., Ale, S.., Mankin, K., Zeckoski, R., Jeong, J., Parajuli, P., Saraswat, D., & Youssef, M. (2015). A Recommended Calibration and Validation Strategy for Hydrologic and Water Quality Models.*Transactions of the ASABE. 58(6),* doi: 10.13031/trans.58.10712, pg 1705-1719.

Garbrecht, J., and Martz, L.W. (1999). TopAGNPS Overview, an Automated Digital Landscape Analysis Tool for Topographic Evaluation, Drainage Identification, Watershed Segmentation, and Subcatchment Parameterization, pg 5-23.

Garbrecht, J., Martz, L.W., & Bingner, R.L. (2000). TOPAGNPS Version 3.1: An Automated Digital Landscape Analysis Tool for Topographic Evaluation, Drainage Identification, Watershed Segmentation, and Subcatchment Parameterization Integrated with AGNPS 2001 Watershed Modeling Technology. *TOPAGNPS User Manual*, USDA, Agricultural Research Service, pg 6-24.

Legates, D.R., and McCabe, G.J. (1999). Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation, *Water Resources Research*, *Vol. 35, No. 1*, pg 233-241.

Liying, L., Litang, Q., Guangsheng, P., Honghu, Z., Zheng, L., & Jianwen, Y. (2021). Non-point source pollution and long-term effects of best management measures simulated in the Qifeng River Basin in the karst area of Southwest China, *Water Supply, 21:1*, pg 262-275.

Luo, C., Li, Z., Li, H., & Chen, X. (2015). Evaluation of the AnnAGNPS Model for Predicting Runoff and Nutrient Export in a Typical Small Watershed in the Hilly Region of Taihu Lake. *International journal of environmental research and public health*, *12*(9), 10955–10973. https://doi.org/10.3390/ijerph120910955, pg 1-15.

Milici, R.C. (1960). The Geology of the Sequatchie Valley Overthrust Block, Sequatchie Valley, Tennessee. Ph.D. diss., University of Tennessee. http://trace.tennessee.edu/utk_graddiss/3232, pg 51-63.

Milici, R.C. (1963). Geology of the Sequatchie Valley Overthrust Block and its Relationships to the Cumberland Plateau Overthrust, Tennessee, Guide Book for the 1963 Field Trip, pg 33-51.

Milici, R.C. (1974). Stratigraphy and Depositional Environments of Upper Mississippian and Lower Pennsylvanian Rocks in the Southern Cumberland Plateau of Tennessee. *Geological Society of America Special Paper 148*, pg 115-133.

Milici, R.C., and Wedow, H. JR. (1977). Upper Ordovician and Silurian Stratigraphy in the Sequatchie Valley and Parts of the Adjacent Valley and Ridge, Tennessee, *Geologic Survey Profession Paper 996*, pg 7-17.

Milici, R.C., Briggs, G., Knox, L.M., Sitterly, P.D., and Statler, A.T. (1979). The Mississippian and Pennsylvanian (Carboniferous) Systems in the United States – Tennessee. *United States Geological Survey Professional Paper 1110-G*, pg G5-G18.

Miller, B.V. (2010). The Hydrology of the Carroll Cave-Toronto Springs System: Identifying and Examining Source Mixing through Dye Tracing, Geochemical Monitoring, Seepage Runs, and Statistical Methods. Masters Theses & Specialist Projects, Paper 216. http://digitalcommons.wku.edu/theses/216, pg 3-15.

Miller, B.V., and Ham, B. (2021). Karst Interest Group Field Trip Guide to the Cumberland Plateau of Tennessee, *U.S. Geological Survey Karst Interest Group Proceedings*, *October 19-20*, pg 126-147.

Miller, R.A. (1979). The Geologic History of Tennessee, *Department of Environment and Conservation, Bulletin 74*, pg 3-44.

Momm, H.G., Bingner, R.L., Emilaire, R., Garbrecht, J., Wells, R.R., & Kuhnle, R.A. (2017). Automated watershed subdivision for simulations using multi-objective optimization, *Hydrological Sciences Journal, 62:10*, https://doi.org/10.1080/02626667.2017.1346794, pg 1564-1582.

Momm, H.G,, Bingner, R.L., Wells, R.R., Porter, W., Yasarer, L., & Dabney, S. (2019). Enhanced field-scale characterization for watershed erosion assessments, *Environmental Modelling & Software, 117*, https://doi.org/10.1016/j.envsoft.2019.03.025, pg 3-13.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D. & Veith, T.L. (2007) . Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations, *Transactions of the ASABE*, 50(3). doi:10.13031/2013.23153, pg 1-16.

Moriasi, D.N., Gitau, M.W., Pai, N., & Daggupati, P. (2015). Hydrologic and water quality models: performance measures and criteria. *Trans. ASABE* (Am. Soc. Agric. Biol. Eng.) 58(6). https://doi.org/10.13031/TRANS.58.10715, pg 1771–1777.

Mulla, N.H., Krishna, B.M., & Manoj, K.B. (2019). A Review on Water Quality Models: QUAL, WASP, BASINS, SWAT, and AGNPS, *International Journal of Scientific Research in Civil Engineering*, *3:10*, pg 65-66.

Nash, J.E., and Sutcliffe, J.V. (1970). River Flow Forecasting through Conceptual Model. Part 1—A Discussion of Principles, *Journal of Hydrology*, *10*. https://doi.org/10.1016/0022- 1694(70)90255-6, pg 282-290.

Palmer, A.N. (1990). Groundwater Processes in Karst Terranes. Groundwater Geomorphology: The role of subsurface water in earth surface processes and landforms, edited by Higgins, CG and Coates, DR, *Geol. Soc. Am., spec. paper, 252*, pp.177-200.

Parsons, J.E., Thomas, D.L., & Huffman, R.L. (2004). Agricultural Non-Point Source Water Quality Models, their use and application. *Southern Cooperative Series Bulletin*, *398*, pg 45-54.

Polyakov, V., Fares, A., Kubo, D., Jacobi, J., & Smith, C. (2007). Evaluation of a non-point source pollution model, AnnAGNPS, in a tropical watershed. *Environmental Modelling & Software*, *22:11*, pg 1617-1627.

Rich, J.L. (1934). Mechanics of Low-Angle overthrust Faulting as Illustrated by Cumberland Thrust Block, Va., Ky., and Tenn.: *A.A.P.G. Bull.*, *vol. 18*, pg 1584-1596.

Renard, K.G., Foster, G.R., Weesies, G.A., Mccool, D.K., & Yoder, D.C. (1997). Predicting soil erosion by water: A guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE), *United States Department of Agriculture*, pg 259-324.

Rodgers, J. (1950). Mechanics of Appalachian Folding as Illustrated by Sequatchie Anticline, Tennessee and Alabama: *A.A.P.G. Bull., vol. 34*, pg 672-681.

Stearns, R.G. (1954). The Cumberland Plateau Overthrust and Geology of the Crab Orchard Mountains Area, Tennessee: Ten. Div. Geol. Bull. 60.

Steinmann, H. (2018). Controls on Speleogenesis in the Upper-Mississippian Pennington Formation on the Western Cumberland Plateau Escarpment: Bowling Green, Western Kentucky University, Master's thesis, https://digitalcommons.wku.edu/theses/3088.

Tennessee Department of Transportation (TDOT). (2021). Tennessee Department of Transportation, *Design Division Drainage Manual*, *Chapter 4 – Hydrology, Section 4.06,* pg 37-53.

Theurer, F.D., and Clarke, C.D. (1991). Wash Load Component for Sediment Yield Modeling, In: *Proceedings of the Fifth Federal Interagency Sedimentation Conference* (5th FISC), pg 7: 1-8.

United States Department of Agriculture - Soil Conservation Service (USDA-SCS). (1986). Engineering Division, Urban Hydrology for Small Watersheds - *Technical Release 55*. http://www.wcc.nrcs.usda.gov/water/quality/common/tr55/tr55.html, accessed 07 June 2023.

Veni, G., DuChene, H., Crawford, N.C., Groves, C.G., Huppert, G.N., Kastning, E.H., Olson, R., Wheeler, B.J., & LaMoreaux, P.E. (2001). Living with Karst: A Fragile Foundation. American Geological Institute, *AGI Environmental Awareness Series*, *4*, pg 11-17.

Wilson, C.W., and Stearns, R.G. (1958). Structure of the Cumberland Plateau, Tennessee, *Bulletin of the Geological Society of America, Vol. 69*, pg 1284-1295.

Wolfe, W.J. (1997). Preliminary Conceptual Models of the Occurrence, Fate, and Transport of Chlorinated Solvents in Karst Regions of Tennessee-*WRIR 97-4097*, pg 1-6.

Young, R.A., and Shepherd, R.G. (1995). AGNPS: An agricultural nonpoint source model, *Workshop on Computer Applications in Water Management (Vol. 33)*, pg 33-34.

Yuan, Y.O.N.G.P.I.N.G., Locke, M.A., & Bingner, R.L. (2008). Annualized agricultural non-point source model application for Mississippi Delta Beasley Lake watershed conservation practices assessment, *Journal of Soil and Water Conserv.*, *63*:*6*, pg 542-551.

Yuan, L., Sinshaw, T., & Forshay, K.J. (2020). Review of watershed-scale water quality and nonpoint source pollution models, *Geosciences*, *10*:*1*, pg 11-12.

Zema, D.A., Bingner, R.L., Denisi, P., Govers, G., Licciardello, F., & Zimbone, S.M. (2012). Evaluation of Runoff, Peak Flow and Sediment Yield for Events Simulated by the AnnAGNPS Model in a Belgian Agricultural Watershed, *Land Degradation & Development*, *23*, pg 205-215.

APPENDICES

A – TOPAGNPS & ANNAGNPS TECHNICAL PROCESSES

Topography

Topography determines the arrangement of land features and surface water flow in the watershed from high-quality [LIDAR (≤3m)] Digital Elevation Models (DEM). DEMs are a value raster containing a cell matrix composed of numeric values that represent terrain elevation.

DEM Data Collection

- Download LIDAR DEMs for watershed into new folder and add *.img* files
- Convert DEM *.img* files to *.tif* using *Raster to Other Format* (*Multiple*) *Tool* in ArcMap, inputs are all *.img* files for study area/watershed, select folder for output workspace, raster format is *TIFF*, then *ok*
- Add new *.tif* files to ArcMap, mosaic all *.tif* files using *Mosaic to New Raster* (*Data Management*) *Tool* and drag *.tif* files from ArcMap contents to input, output location is new folder, output is new mosaic with *.tif* extension, spatial reference for raster is *NAD_1983_UTM_Zone_16N*, change pixel type to *32_bit_float*, number of bands is *1*, mosaic operator is *blend*, mosaic colormap mode is *match*, then *ok*
- Convert *mosaic.tif* from *feet to meters* by searching *Raster Calculator (Spatial Analysis) Tool*, input is mosiac.*tif*, then multiply (*) by *.3048*, and save as mosaic_m.*tif*
- Change cell size of *mosaic_m.tif* to 3m x 3m resolution using *Resample* (*Data Management*) *Tool,* input is *mosaic_m.tif*, output is resampled *.tif*, change X to *3*, Y to *3*, resampling technique is *NEAREST*, then *ok*
- Search *Raster to ASCII (Conversion)* tool, input is resampled *.tif*, output is final *DEM.asc* shapefile
- If needed, change projection with *Define Projection (Management) Tool*, input is *DEM.asc*, output is correct UTM Zone

TOPAGNPS

TopAGNPS

Automated digital landscape analysis tools such as Topographic Parameterization for Agricultural Non-Point Source (TopAGNPS) identifies and analyzes high-quality Digital Elevation Models (DEMs) to evaluate topography, calculate drainage network information, and finalize the watershed outlet location. TopAGNPS interfaced with a Geographic Information System (GIS) can evaluate digital landscape elevations by performing topographic parametrization to generate the basic modeling units of cells and reaches required for automated hydrologic analysis and watershed modeling in AnnAGNPS.

TopAGNPS Simulation

Calculates final outlet location(s) of catchment(s) in *DEM.asc*.

- Running TopAGNPS requires the *TOPAGNPS.exe (executable)*, *TOPAGNPS.csv control file*, and newly *converted ASCII (DEM.asc) Raster file*
- Update first 2 rows (keywords and values) of *excel control file*
- Update filename with *DEM.asc*, values for *DEMPROC* is *2* (network generation), *OUTFORMAT* is *1* (latitude/longitude coordinates) or *0* (row/column), *CSA* is *5*, *MSCL* is *250*, and input *ncols*, *nrows*, and *cellsize* by editing *DEM.asc* file in *notepad++* then enter values
- Click *TOPAGNPS.exe* or use *cmd window* to run model (reduced mode)
- Outlet location for catchment is determined by opening *UPAREA.asc* file in ArcMap
- White line in *UPAREA.asc* file is path of water
- Click *UPAREA.asc* file and go to properties: Symbology; *Stretch,* Type; *Standard Deviations,* and Statistics; *From Current Display Extent*
- Identify and finalize *X (outcol)* & *Y (outrow)* of outlet in *meters* by using *Identify*, click outlet location, copy *outcol* and *outrow* from *location* bar, paste in *TOPAGNPS.csv control file,* and change to *General* format cells
- Outlet locations for coves (sinking streams/caves) vary from typical watershed outlets and must be altered by opening the watershed *DEM.asc* in text editor (*notepad ++)* and change the given outlet *XY* elevations from meters to *-9999* (*no data value*)
- Run full processing of TopAGNPS by updating the *TOPAGNPS.csv control file DEMPROC* value to *0* and execute TopAGNPS to create updated *.asc* files of *BOUND, RELIEF, UPAREA, NETW,* and *AnnAGNPS cell ID*
- Open *BOUND.asc* in GIS to compare to each DEM/Study Area catchment boundary

Hydrological Correction of DEM

Attempts to recalculate catchment flow patterns as close to real world scenarios by changing elevations within the DEM due to TopAGNPS identifying flow channels using the lowest raster value, but miscalculates infrastructure (bridges, culverts, ect.).

* *Not utilized in this research*

Converting Final TopAGNPS Files for AnnAGNPS

Converts *AnnAGNPS_Cell_ID.asc*, *NETW.asc*, and *BOUND.asc* to *.shp* (shapefiles).

- Converting *AnnAGNPS_Cell_ID.asc* and *BOUND.asc* to *.shp (polygon):*
  - Update *.asc* (ASCII) files by first opening in text editor (*Notepad ++*)
  - The first 4 lines of header remain unchanged, change *cellsize* and *nodata value* lines from *floats* (decimals) to *integers* (no decimals)
  - Use *Raster to Polygon (Conversion) tool* in *ArcMap* to convert *.asc* to *polygon*

- Input is *AnnAGNPS_Cell_ID.asc* and/or *BOUND.asc* files, field is *VALUE*, and check *simplify polygons* box
- The outputs are the new *polygon* shapefile containing the watershed boundary and the AnnAGNPS cell values / sub-catchments
- Repeating cell number values (cells split in half) in the new *AnnAGNPS polygon.shp* file are changed to only used once by searching *Dissolve (Data Management) tool*
- Input is new *polygon* file, check *gridcode box* in Dissolve Field(s), and check *create multipart features*
- Output is *dissolved* AnnAGNPS *.shp* containing non-repeating cell values
- Use *Clip (Data Management) tool* to clip *DEM.asc* of each catchment to its *BOUND.shp* to represent each catchment within the correct watershed boundary
- Input raster is *DEM.asc*, output extent is *BOUND.shp (polygon)*, check *use input features for clipping geometry*, save the output as *.tif*, check *maintain clipping extent*, then *ok*
- Search *Raster to ASCII (Conversion)*, input is *clipped DEM.tif*, output is new *clipped DEM.shp*, save type is *.ASC* file, then *ok*

- Converting *NETW.asc* to *.shp:*
    - Update *.asc* (ASCII) files by opening in text editor (Notepad ++) to change from floats to integers
    - Use *Raster to Polyline tool* to convert *.asc* to .shp (*vector)* in ArcMap
    - Input is *NETW.asc*, field is VALUE, and uncheck *simplify polyline* box
    - Output is *NETW.shp* (concentrated flow of water / reaches throughout watershed)

SOIL

Soil

The physical properties and cover conditions of soil influence the rate of water infiltration and runoff potential following a storm. Digital datasets of soil help determine soil types, hydrologic soil groups, and the spatial frequency in each sub-catchment (per county) required to run AnnAGNPS.

Soil Type Download

- Download spatial datasets from *USDA* for each county within study area
- Zip file contains spatial, tabular, metadata, and state soil shapefiles
- Soil shapefile dataset(s) are added to ArcMap and projected to what coordinate system is used in watershed using *search* tab, type *project*, and select Projection (Data Management) Tool
- Combine *state and county code* (*area symbol*) & *map unit symbol (MUSYM)* in the *attribute table* of each soil shapefile to identify unique signifier for each soil type as a new field (*MUSYM_NEW*), type *text*, length *50*
- Right click *MUSYM_NEW* field, select *Field Calculator*, change parser to *python*

- Type formula *!MUSYM!+’_’+!AREASYMBOL!* in the *MUSYM_NEW* field then *ok* to create unique soil signifier
- Merge datasets of new field for all soil polygon shapefiles with *Merge (Data Management) Tool*, save as *soil_merge.shp*
- *Soil_merge* is then clipped to catchment boundary (*BOUND.shp)* polygon
- Search *Clip (Analysis) tool*, input is *soil_merge.shp*, clip feature is *BOUND.shp*, save output as *soil_clip.shp*

Soil Table Download

- Modify the *soil_table_script.txt* document to download the soil data table from the USDA website using *area symbol code* for each county in study area
- Select ‘*Queued/Text*’ for results format, check box ‘*First row contains column names’*, set options as vertical bar ‘ **|** ’, and ‘ **“** ’ for text delimiter
- Submit query and download *table.txt* soil zip file(s) from link

NITA

Program that formats the *table.txt* soil files to correct AnnAGNPS format.

- Combine all *table.txt* soil files into one txt file/record of data (*soil_table_ALL*)
- Copy soil table to blank Excel document and separate by columns (*text to columns tab*)
- Create *MUSYM_NEW* column and update existing *MUSYM* column by concatenating with *Area Symbol Code* to match polygon shapefile
- Delete old *MUSYM* column, rename, and save Excel document as *.csv file*
- Open *NITA.csv* file and update FILENAME using new *.csv file*
- Run *NITA.exe* to create final soil files of *log*, *soil_data*, and *soil_layers_data.csv*

Creating Unique Values for Soil Type

Changes the *musym_areasymbol code* to *numeric value* (*soil_values.csv*) in *Python* so polygon shapefile and soil table data efficiently runs in AnnAGNPS.

- Open *soil_value.py* script and change *F:* to *Table_ALL_FINAL_soil.data.csv* read location and *K:* to location *soil_values.csv* to write
- Run module to read soil data and write *soil_values.csv* with no soil type duplicates
- Create second column of sequenced numbers starting at 1 in *soil_values.csv*

Combining Soil Polygon Shapefile and Soil Table

Creates new VALUE field of unique numbers so *soil table data* and *spatial soil data* have same numbering (value).

- Create separate *county_names.csv* file with *area symbol codes* of each county
- Open *attribute table* of soil polygon shapefile (*soil_merge*) in ArcMap and add new *VALUE* field, *Type* is *short integer* with *Precision* of *5*

- Open *soil_value_update.py* script and update line 5 with file location of *soil_values.csv*, update line 9 with file location of *county_names.csv*, and update polygon shapefile name (*soil_merge*) in line 19
- With *soil_merge* open in ArcMap, click *Python* module and copy updated *soil_value_update.py script* into *Python* box and press enter to update new *VALUE* field with *unique values* of soil table data and soil spatial layers

Updating the Soil Number in the Soil Data and Soil Layers Data Files from NITA

- Add both NITA *.csv* output files (*Table_ALL_FINAL_soil_data.csv* and *Table_ALL_FINAL_soil_layers_data.csv*) to ArcMap
- For each *.csv* file, go to *joins and relates*, *join*, then *soil ID* field, add *soil_values.csv* table, and *musym_new* for next field, then *ok*
- Once each *.csv* file is joined with the unique soil number, *export* data and change the save type as *.txt* file using the *.csv extension*
- Open each *.csv* file, scroll to end of file and *delete* the *Field 1* (*MUSYM_NEW*) column, then *copy* the *Field_2* column contents and *paste* into the *Soil ID* column so it only contains *numbers*, once updated the *.csv* files are ready for AnnAGNPS

Adding the Soil Unique Value to the AnnAGNPS Cells

- Open the *soil_clip.shp* file from previous section (*Soil Type Download*) and convert soil polygon to raster using *Polygon to Raster (Conversion)* Tool, the input is *soil_clip.shp*, the Value field is *VALUE*, output raster is saved as *.tif* file (*soil_raster.tif*), set cell assignment to *Cell_Center*, priority field is *none*, and change cellsize to *3*
- Before running conversion tool click *Environments* tab and open *processing extent* tab, change the extent and snap raster to *AnnAGNPS_Cell_ID.asc* file then click *run*
- Search for *Zonal Statistics as Table (Spatial Analysis) Tool*, the input raster or feature is the dissolved *AnnAGNPS_Cell_ID.shp*, zone field is *gridcode*, the input value raster is newly created soil raster from previous step (*soil_raster.tif*), the output table contains *majority voting* with *.dbf* extension, check *ignore NoData calculations*, change statistics type to *Majority*, then click *ok*
- The tool creates a *.dbf* table (*Soil_Zonal_Statistics.dbf*) containing *gridcode*, *count*, *area*, and *majority*

Adding the Zonal Statistics Back to the AnnAGNPS Cells

Creates new *.shp* file joining *zonal statistics table* with *dissolved AnnAGNPS_Cell_ID.shp* file.

- Right click *dissolved AnnAGNPS_Cell_ID.shp* file and click *joins and relates*, then *join*
- Select *gridcode*, choose the newly created zonal statistic table (*Soil_Zonal_Statistics.dbf*), select *gridcode* field for 3rd box, then *ok*
- Check to see if values joined correctly by opening *attribute table* and sort the *Majority* field as *descending*

- Once the *.shp* file is verified export the data as a *.txt* file by opening the *attribute table*, click *table options*, then *export*
- Save the *.txt* file to a new *Excel* sheet by copying data, *separate by columns (text to columns)*, and use *comma* as delimiter
- Create new *.shp* file of joined datasets by opening the attribute table of the newly joined *dissolved AnnAGNPS_Cell.ID.shp* file, *export*, then save as *.shp* file
- Once displayed in ArcMap, change symbology to *categories* and correct color ramp, use *majority* value field, and add *All Values*

MANAGEMENT

Management

Management information on dominant crop types and majorities are determined from the USDA-NRS *Crop Data Layers* (CDL) unique county codes for annual land use. CDL years are joined and clipped to study area sub-catchments to calculate majority zonal statistics, crop type percentages, and develop management schedule for AnnAGNPS.

Management Download and Boundary Clip

- Download CDL files by defining time period (years) for counties within study area
- Spatially join the <u>same year</u> of each county *CDL.tif* to the study area in ArcMap
- Raw data files will have gaps between datasets, to remove gaps use *Mosaic to New Raster* tool if catchment/study area contains <u>separate</u> counties
- Run *mosaic* tool for all years during time period of study
- *Clip (Data Management)* <u>all years</u> of *CDL Mosaic.tif* to each catchment boundary shapefile in study area to mask out areas not in catchment(s) using *BOUND.shp* (polygon) created in conversion of final TOPAGNPS files

Dominant Crop Type

Determines dominant crop type for specified years in each sub-catchment.

- Create attribute tables for the annual CDL raster datasets clipped to each catchment boundary using *Clip (Data Management)* in GIS using the first, middle, and last years of time period
- Join CDL type to raster by selecting CDL raster then *Joins and Relates* and *join* by <u>value</u> of *CDL code csv file*
- Copy newly joined attribute table to *excel document* and sort by count (largest to smallest) and *sum* Count then calculate *percentage* of each crop type (each type count / Count sum)
- Repeat process for specified years in time period

Zonal Statistics (Management)

- Open the dissolved AnnAGNPS shapefile (*AnnAGNPS_Cell_ID_dissolve.shp)* containing non-repeating cell values and search for *Zonal Statistics as Table (Spatial Analysis)* Tool
- Input raster/feature is *ANNAGNPS_Cell_ID_dissolve.shp*, zone field is *gridcode,* input value raster is each *CDL* for all years of study, output *majority table* is *.dbf* file of each *CDL year*, check *ignore NoData in calculations* box, select *majority* statistics then *ok*
- Create separate *.shp* file for each *CDL majority table* by right clicking *AnnAGNPS_Cell_ID_dissolve.shp*, select *Joins and Relates*, click *Joins*, enter *gridcode* for part 1, *CDL year* for part 2, and *gridcode* for part 3 then ok
- Right click *AnnAGNPS_Cell_ID_dissolve.shp* after CDL is joined then *data*, *export data*, save as *.shp* file, remove join from *AnnAGNPS_Cell_ID_dissolve.shp* then repeat for each year in study
- Join ALL years of CDL *majority table.dbf* files created in previous step in chronological order to the first year of study (2016) *CDL_join_year .shp* file
- Once all years are joined to (*CDL_join_2016.shp*) right click file in contents, *data*, *export data*, and save as *AnnAGNPS_Cell_ID_CDL_ALLYEARS.shp*
- Open attribute table of new *.shp* and click on *majority* field then select *field calculator*
- Update each *majority* field by changing parser to *python*, check *show codeblock* box, then enter the following *Python* script in the *Pre Logic Script Code* box:

  ```
  def updatevalue(value): #no indention
    if value == 0: #indented 2 spaces
      return 176 #indented 4 spaces
    else: #indented 2 spaces
      return value #indented 4 spaces
  ```
- Next enter the following *Python* script in the *Majority=* box:

  ```
  updatevalue(!MAJORITY!)
  ```
- Change the *Majority=* box to match each MAJORITY field, (*ex: MAJORITY_1)*
- Once all majority fields are updated save the attribute table of *AnnAGNPS_Cell_ID_ CDL_ALLYEARS.shp* by clicking *table options*, *export*, then save as *.txt* file
- Open the newly created *.txt* file and copy to new *Excel* file then separate the data by columns under *Data* tab, select *text to columns,* check delimited by *comma* only, then click *finish*
- Create separate *.csv* files for each CDL year but keep only *Gridcode* and *Majority* columns, include *year* in the column header and save a copy as *.xlsx* file
- Open *AnnAGNPS_Cell_ID_CDL_ALLYEARS.xlsx* and combine each CDL majority into a new *excel* file
- Convert the CDL majority Master Category *number* for each grid code into a *letter* representative of the crop type: (176 = *G*: *Grassland/Meadow*; 141-143 = *F*: *Forest*; 121-124 = *U*: *Urban*;1 = *S*: *Soybean*; 5 = *C*: *Corn*; and 83 or 111 = *W*: *Water*)
- *Management sequence* for each year is determined by the highest percentage Master Category *letter* per *grid code* (F;G;F;F;G;F) then save as *Management_Sequence.csv*

- *Management schedule* is created from *Management_Sequence.csv* and *management_code_template.csv* files of G, F, U, S, C, and W, the output file after running *Python* script *mngt_sch_v08.py* is *mgnt_sched.csv*

Management Practices (RUSLE2)

The Revised Universal Soil Loss Equation version 2 (*RUSLE2*) program predicts the long-term average of rill and sheet flow sediment loss due to erosion by rainfall and runoff. This simple empirical and process-based computer model is maintained by the Natural Resources Conservation Service (NRCS). Their primary focus are agricultural lands but contribute to collecting and sharing information on natural resources while providing information on soil surveys and classification, agricultural operations, water quality, and technical assistance to farmers through partnerships with local and state agencies.

- Go to *RUSLE2* official database website to download management practices for the study area, *https://fargo.nserl.purdue.edu/rusle2_dataweb/RUSLE2_Index.htm*
- Click *Crop Management Zone Maps* under *Crop Management Templates*
- Select the appropriate map box of the *Crop Management Zone* (*CMZ*) based on watershed location and notate the unique 2 digit number of the CMZ (*CMZ 63*)
- Click *Data Files* under *Crop Management Templates* and select the correct *CMZ.zip* file to download from the parent directory
- Unzip the CMZ file to extract the *.gdb* file then open the *RUSLE2* program
- Click *Database*, then *Import database*, and select *CMZ 63.gdb* file
- Check the folder box under *Import Database*, check *Import to new folder*, and click *Import*
- Click *File*, *Open*, *Management*, then select *CMZ 63* to navigate folders
- *AnnAGNPS_Management_Operation_Data.csv*, *AnnAGNPS_Crop_Growth_Data.csv*, and *AnnAGNPS_Crop_Data.csv* files are populated from *CMZ 63.gdb* in *RUSLE2* and prior AnnAGNPS simulations for *urban*, *grass*, and *forest* in similar Tennessee watersheds

CLIMATE DATA

Climate Data

The primary input for the AnnAGNPS watershed model consisting of two types of climate/weather files: *Primary* (observed and/or synthetic data without gaps) and *Secondary* (assigned to AnnAGNPS cells as main source of climate information). Information gaps of missing measurements in *Secondary* files are read from *Primary* files.

Secondary Climate Stations Download

- Download climate station data from *NOAA* website based on county and catchment relation/location of study area

- Climate data on *NOAA* has four inputs: Dataset type is *daily summaries*; Data range is *time period* of study; Search *by counties*; and Search Term is *name of county*
- Station selection is based on *station name* and *geographic location* with selection of data custom outputs *precipitation* and *air temperature*
- NOAA data file download is *Excel* format and becomes *secondary* climate files (stations) after organization and quality control of data

Organizing and Quality Control of Downloaded Climate Information

- Each county dataset may contain multiple climate station per file (use *filter* tool)
- Keep all original *Excel* climate data files then separate *Excel* files by station
- Keep *station name, lat, long, elevation, date* (start and end of study period), *prcp, tmax*, and *tmin* columns, delete rest of columns
- Create separate database containing *station name, lat, long, start date, end date*, and *elevation* of each climate station and save as *.csv* file
- Once organized modify each *Excel* climate data file by moving *prcp, tmax*, and *tmin* <u>5</u> columns to right
- Add new date (*actual date*) column to the right of original date (*observed date*)
- Use formula to read both date columns to check if they match, dates that do not match will display *#N/A* meaning climate data is missing
- Calculate coverage score in empty column to return *percentage* of missing data
- Add calculated percentage to *climate station database.csv* file
- Create *climate_daily.csv* file for each *Excel* climate station after completing quality control

Generating Secondary Station Files

Station files are associated to each *climate_daily.csv* file and provide information about each climate station used in AnnAGNPS simulation.

- Each station will have its own climate station file
- Columns include: *Version* is *6* (AnnAGNPS version 6.1), *Input Unit Code* is *1* for standard international and *0* for not standard international, *Begin* and *End date, Lat, Long*, and *Elevation*
- Stations are saved as *climate_station_n.csv* files based on spatial assignment and amount of secondary climate stations in study area

Generating Synthetic Weather Climate Files (AgGEM)

The <u>G</u>eneration of Weather <u>E</u>lements for <u>M</u>ultiple Application (AgGEM) program uses existing long-term *statistical parameter files* of synthetic meteorological data to generate a *no gap* dataset of *primary* statistic climate files in the correct AnnAGNPS format.

- Open the USDA AgGEM database to select the file (*statistical seed*) of the closest city in relation to the study area and download

- Add the city *seed* and *agGEM.exe* program to a folder and access the command line shortcut by typing *cmd* in the folder location bar
- Click *tab* until *agGEM.exe* is in the command line and hit *enter*
- Type the input file name (*TN_Chattanooga*) without the *gm6* extension then *enter*, say *no* to entering a new value for average annual precipitation, *no* to English units, then type the desired number of simulated data years, no selection defaults to producing *10 years* of data
- Hit *enter* to complete AgGEM simulation and produce the gem output *TN_Chattanooga_gem6.out*, AGNPS output *TN_Chattanooga_climate.inp*, *climate_daily.csv*, and *climate_station.csv* files.
- Update the *year* column of the *climate_daily.csv* from 1-10 to study period years (including leap years)
- Change the version to *6* and update the *beginning* and *ending climate date* columns of the *climate_station.csv* file

Spatial Assignment of Climate Stations to AnnAGNPS Cells

Creates point shapefile(s) to display spatial assignment of climate stations in ArcMap.

- Add the *Climate_Stations_Final.csv* file to ArcMap
- *Right click* the *.csv* file and navigate to *Display XY Data*, the *X* (longitude) and *Y* (latitude) columns should automatically be selected
- Select *edit* then choose a general GCS (Geographic Coordinate System) like *GCS_North_American_1983* and click *ok*
- The output is a new layer called *Climate_Stations_Final.csv Events*, *right click* the layer and select *Data*, *Export Data*, and save as a permanent shapefile in ArcMap
- Open the *search tab*, type *project* and select *project (Data Management) tool*, the input is the new *.shp*, save the *output feature class* then change the *output coordinate system* to match the projection of the watershed study area (*NAD_1983_UTM_Zone_16N*) and then *ok*

Creating Thiessen Polygons

Allocates space around the nearest station/point shapefile to create a boundary where everything inside has the attributes of that point feature.

- Open the *search tab* in ArcMap, type *Thiessen polygon* and select *Create Thiessen Polygons (Analysis) tool*
- Input feature is *Climate_Stations_Final_Projected.shp*, save output, select *ALL* in output fields (optional), then *ok*
- Output file is *Thiessen polygon shapefile*

Zonal Statistics (Secondary Climate Stations)

Creates table from Thiessen polygon containing majority voting of the climate station number(s).

- Click the Thiessen polygon *.shp* file, open the *search tab* and type *polygon to raster* and select *Polygon to Raster (Conversion) Tool*
- Input is new Thiessen polygon *.shp* file, the *Value* field contains the climate station numbering system (*Input_FID*), save output as a new Thiessen raster file with *.tif* extension
- Set the cell assignment to *Cell_Center*, priority field is *none*, and change cellsize to *9*
- Click the *Environments* tab then open the *Processing Extent* tab, change *extent* and *snap* raster to *AnnAGNPS_Cell_ID.asc* file of watershed then *ok* to create raster
- Open the search tab, type Zonal Statistics as Table and select *Zonal Statistics as Table (Spatial Analysis) tool*, input raster is *AnnAGNPS_Cell_ID_dissolve.shp*, Zone field is *gridcode*, input value raster is new *Thiessen raster.tif*
- Save output table with *.dbf* extension (*Thiessen_Majority_table.dbf*), <u>check</u> *Ignore NoData* box, select *Majority* for Statistics type, then *ok*

Adding Majority Zonal Statistics back to AnnAGNPS cells

- Right click the *AnnAGNPS_Cell_ID_dissolve.shp* file in contents, go to *Joins and Relates*, then *Joins*
- Choose *gridcode* for field 1, select *Thiessen_Majority_table.dbf* as table to join, *gridcode* for field 3, then *ok*
- Verify that *Majority* values joined correctly by right clicking the header then *sort as ascending* (Null values will show if table did not join correctly and must be manually updated)
- Once verified right click the newly joined *AnnAGNPS_Cell_ID_dissolve.shp* file, click *Attribute Table*, select *Table Options*, *Export*, and save as *text* file
- Open the *Secondary_Climate_Station_Majority.txt* file, copy the text then add to new excel sheet, click *Data*, *Text to Columns*, select *comma* as delimiter, then save as *.xlsx* file
- Save joined *AnnAGNPS_Cell_ID_dissolve.shp* file by right clicking file in contents, *Data*, *Export Data*, then save as *AnnAGNPS_Secondary_Climate_Majority.shp*

Discharge Conversion and Separation of Baseflow from Observed Data

Baseflow is generally groundwater discharge and accounts for the portion of streamflow contained in subsurface flow during the dry season. Water below the soil profile that interacts with stream channels is not simulated in AnnAGNPS and therefore ignored in the system and treated as a loss. Accurate correlation between simulated and observed data requires the conversion of discharge (Q) from cubic feet per second (f^3/s) to cubic meters per day (m^3/d) and separation of baseflow from Q of United States Geologic Survey (USGS) site/gage water data using modified *RStudio* and *Python* scripts.

- Open the website *waterdata.usgs.gov* and locate the individual site(s) link using the *Map* or *Site Information tab*
- Display stream gage water data as *tab-separated* daily time series based on *begin* and *end date* of study period
- Copy the statistics under *Description* displayed as *Discharge, cubic feet per second (Mean)* and save study period data as *gage.txt* file
- Copy *gage.txt* file into new *.csv* file using *text to columns* then multiply the f^3/s value of *Q* by *2447* to convert into m^3/d and save as *gage.csv*
- Open *gage.csv* and add *Date* in *month/day/year* format to *column 1* and converted *Q* in *column 2* to correctly format file for *RStudio* and save as *SequatchieGage.csv*
- Read *SequatchieGage.csv* in *RStudio* using *Hydrostats package* and modified *baseflows* script to vary the significance level (*alpha*) between .100 and .999 to separate the digital baseflow from *Q*
- Write new *SequatchieGage_bf_removed_a100.csv* file in *RStudio* then save in correct *Python* script format using *time_series_template.csv*
- *Time_series_template.csv* uses *year* in column 1, *month* for column 2, *day* in column 3, and newly calculated *Q* in column 4. Update the columns with study period dates then open a *cmd* window and run *timeSeries.py script* to create *Gage.csv* summaries of *weekly, monthly, quarterly,* and *annual* removed baseflow values of observed *Q*

ANNAGNPS

AnnAGNPS

The Annualized Agricultural Non-Point Source (AnnAGNPS) model contains a continuous simulation, single event-based Pollutant Loading (PL) surface runoff model designed to predict long-term sediment-transported pollutant loadings and water quality responses at the agricultural watershed-scale. The analyzed watershed is subdivided into homogenous hydrologically-derived cells, or *sub-catchments*, in terms of land use and soil properties. AnnAGNPS, integrated with additional modeling interfaces such as TopAGNPS and GIS, can then estimate the topographic properties in the basic modeling units of cells and concentrated flow paths, or *reaches*, to provide parameter information on slopes, drainage areas, reach lengths, and elevations. Analysis of land surfaces (cells) at any point within the watershed is viable since sub-catchments have relatively uniform hydrological and physical characteristics. AnnAGNPS then simulates constituent quantities of runoff, sediment, and pollutant loadings that are routed from each cell and either detach and deposit within the channel network or transport to the watershed outlet, respectively referred as *erosion*, *yield*, and *load*.

AnnAGNPS Input Preparation

- AnnAGNPS requires the *AnnAGNPS version 6.01.exe (executable)*, *AnnAGNPS.fil* file, *AnnAGNPS_master.csv* control file, and template folders of *Climate*, *General*,

Simulation, and *Watershed* containing sub-content *.csv* inputs from GIS, RUSLE2, TopAGNPS, and other AGNPS tools
- *AnnAGNPS.fil*: FIL file required for simulations
- *AnnAGNPS_master.csv*: Master control file containing links to other *.csv* files enabling content pulls from multiple input files to create AnnAGNPS gage output files
- *Climate folder*: Folder sub-contents include *Primary* (*climate_daily.csv* and *climate_station.csv*) and *Secondary* (*climate_daily_#.csv* and *climate_station_#.csv*) climate/weather stations based on spatial relationship between watershed and station location
- *General folder*: Folder sub-contents include, *manfield_field.csv* file requiring inputs for *Landuse Type ID* and *Management Schedule ID* columns, *Mgnt_Schedule.csv* file requiring inputs for *Management Schedule ID* and *Curve Number ID* columns, and *Table_All_FINAL_soil_layers_withNUM.csv* & *Table_All_FINAL_soil_data_withNUM.csv* files containing *NITA* unique soil numbers (*Soil IDs*) for AnnAGNPS cells
- *Simulation folder*: Folder sub-contents include *sim_period.csv* file requiring inputs for *Simulation Begin Month-Date-Year* and *Simulation End Month-Date-Year* columns based on time period of study
- *Watershed folder*: Folder sub-contents include *watershed_data.csv* requiring update of *Lat, Long,* & *watershed name/location*, *AnnAGNPS_Cell_Data_Section.csv* requiring inputs for *Soil_ID* (referenced from *Majority* column of *Soil_Zonal_Statistics.xlsx* file within *Soil_Tables* folder) & Management_Field (referenced from *#year_Majority* column of *AnnAGNPS_Cell_ID_CDL_ALLYEARS.xlsx* file located in *Annual_CDL_Clip* folder within *Management* folder)

Initial AnnAGNPS Simulation

Creates the file *AnnAGNPS_TBL_Gaging_Station_Data_Hyd.csv* containing requested mass and area unit data of reach information, watershed outlet streamflow, drainage area, precipitation, and pollutant and sediment loadings of the simulation period.

- Open the watershed folder containing the *AnnAGNPS version 6.01.exe*, *AnnAGNPS.fil* file, *AnnAGNPS_master.csv* control file, and template folders containing sub-content .csv inputs
- Once template folders and *AnnAGNPS_master.csv* are verified click the *AnnAGNPS version 6.01.exe* to begin simulation
- Confirm the status of the AnnAGNPS simulation and the occurrence of no errors
- Output files include *AnnAGNPS_AA.csv*, *AnnAGNPS_dbg.csv*, *AnnAGNPS_LOG.csv*, and *AnnAGNPS_TBL_Gaging_Station_Data_Hyd.csv*

Summary Files of Simulated Data

Uses *Python* script to create *.csv* summaries of weekly, monthly, quarterly, and annual total streamflow values from *AnnAGNPS_TBL_Gaging_Station_Data_Hyd.csv* files following initial AnnAGNPS simulation.

- Open *time_series_template.csv*, update study period dates, and copy *Total Streamflow* (*Mg*) values (*Column P*) for entire study period from each *AnnAGNPS_TBL_Gaging_Station_Data_Hyd.csv* and paste into template then save as *AnnAGNPS.csv* files
- Type *cmd* in command window and run *timeSeries.py script* to create *AnnAGNPS.csv* summaries of *weekly*, *monthly*, *quarterly*, and *annual* simulated AnnAGNPS data

AnnAGNPS Model Calibration and Adjusting Curve Numbers

Calibration of AnnAGNPS output tables in watershed modeling is an important step to improve model performance and ensure confidence in simulation quality. This iterative process compares the model results of runoff with observed (actual) runoff by adjusting input parameters of land use categories to obtain results that best represent reality. This approach of estimating surface runoff from agricultural watersheds is based on the empirical method developed by the NRCS Curve Number (CN) model. CN's represent the estimation of runoff producing potential as a function of cumulative precipitation, hydrological condition, soil group classification, land use, and the antecedent moisture condition (AMC) of the watershed. Runoff CN values range from 0 to 100 and characterize the runoff properties of a particular ground cover type, hydrologic condition, and soil group.

- Open *AnnAGNPS.csv* and *Gage.csv monthly* and *annual* summary files and copy results into *Statistics_Template.xlsx* then update *Date* (columns *A* and *H*)
- Compare stream flow of *initial* CN values for *AnnAGPNS.csv monthly* and *annual* summary files to observed values of *Gage_100_M_sum.csv* through *Gage_999_M_sum.csv* and *Gage_100_A_sum.csv* through *Gage_.999_A_sum.csv*
- Increase the *initial* CN values of *TN_rocurve.csv* in theAnnAGNPS *General Folder* by *3* for all land use types (G, F, U, S, and C) to *increase* runoff and re-run AnnAGNPS
- Open the new *AnnAGNPS_TBL_Gaging_Station_Data_Hyd.csv* and copy *Total Streamflow* (*Mg*) values (*Column P*) and paste to *time_series_template.csv*
- Rename the new *time_series_template.csv* as *AnnAGNPS_withCurve_3.csv*
- Type *cmd* in command window and run *timeSeries.py script* to create *AnnAGNPS_withCurve_3.csv* summaries of *weekly*, *monthly*, *quarterly*, and *annual* simulated AnnAGNPS data
- Compare stream flow of *AnnAGPNS_withCurve_3.csv monthly* and *annual* summary files to observed stream flow values of *Gage_100_M_sum.csv* through *Gage_999_M_sum.csv* and *Gage_100_A_sum.csv* through *Gage_.999_A_sum.csv*
- Repeat the processes of increasing CN values of *TN_rocurve.csv* by *5* and *7* for all land use types in AnnAGNPS, pasting *Total Streamflow* values into *time_series_template.csv*, and creating *monthly* and *annual* summaries by running *timeSeries.py script* then save *Statistics_Template.xlsx* as new file to execute pairwise comparisons and evaluate runoff relationships

Evaluation of Model Performance

Evaluations of hydrological model performance often utilize multiple statistical analysis methods to visualize and investigate patterns, relationships, and trends of quantitative data. Quantitative data typically measures numerical values between model-simulated data and observed data under identical conditions, where the total number of observations during a given incremental duration of time evaluates the model's performance. Pairwise comparisons are statistical techniques often used to evaluate the spatial and/or temporal relationship of these variables from a set of inputs, although correlation-based measures like Pearson's product-moment correlation coefficient (r) and the coefficient of determination (R^2) are overly sensitive to extreme outliers. In model evaluation, both graphical techniques and quantitative statistical measures are recommended to determine model efficacy and should include central tendency concerning the mean of values and their frequency distribution, r, R^2, the Nash-Sutcliffe efficiency (*NSE*) model evaluation statistic, the average tendency of simulated and observed data expressed as a percentage (*PBIAS*), and root mean square error (RMSE) -observations standard deviation ratio (*RSR*).

- Open revised *Statistics_Template.xlsx* after adjusting Curve Numbers and creating all *monthly* and *annual* summaries for *AnnAGNPS.csv* (*initial* through *withCurve_7*) and *Gage.csv* (*.100* through *.999*)
- Perform pairwise comparison and statistical analysis of *AnnAGPNS.csv initial* CN summary *and Gage_100* through *Gage_.999_sum.csv* from formulas for *annual* (column *E* and *F*) and *monthly* (column *L* and *M*) from *simulated streamflow* (column *B* and *C*) and *observed/gage* (column *J* and *K*) summary values:

    - *mean of simulated:* `=AVERAGE($B$3:$B$8)`
    - *mean of observed*: `=AVERAGE(C3:C8)`
    - *coefficient of determination*: `=RSQ($B$3:$B$8,C3:C8)`
    - *correlation coefficient*: `=CORREL($B$3:$B$8,$C3:$C8)`
    - *NSE*: `=1-(SUMPRODUCT(($B$3:$B$8-C3:C8)^2)/SUMPRODUCT(($B$3:$B$8-F2)^2))`
    - *PBIAS*: `=(SUMPRODUCT(($B$3:$B$8-C3:C8)*100))/(SUMPRODUCT($B$3:E8))`
    - *RSR*: `=(SQRT(SUMPRODUCT(($B$3:$B$8C3:C8)^2)))/(SQRT(SUMPRODUCT(($B$3:$B$8-F2)^2)))`

- Repeat process for *AnnAGNPS_withCurve_3* through *AnnAGNPS_withCurve_7* and *Gage.csv* (*.100* through *.999*) to determine the adequate calibration based on model evaluation statistics and their satisfactory corresponding performance ratings
- Once model is adequately calibrated update *TN_rocurve.csv* with the *final* CN values in the AnnAGNPS *General Folder* and run final AnnAGNPS simulation

Integrated Simulation of Study Area

Creates the final integrated *AnnAGNPS_TBL_Gaging_Station_Data_Hyd.csv* file containing requested mass and area unit data of reach information, watershed outlet streamflow, drainage area, precipitation, and pollutant and sediment loadings for the entire study area.

- Add the *Reach_GC.shp* and *Reach_SR.shp* raster grids to GIS and use the *identify* tool to determine the contributing reach inputs (*grid code IDs*) for each drainage basin

  - Grassy Cove inputs: (*59 = Grassy Cove reach*; *61 = Bristow Spring*)
  - Sequatchie inputs: (*595 = Head of Sequatchie*; *607 = Roaring Branch Spring*)
- Once identified, grid code IDs are added as *sub-content files* to the respective *watershed* template folders and integrated simulation of *Grassy Cove-Little Cove* is executed
- Open the new integrated *LC-GC AnnAGNPS_TBL_Gaging_Station_Data_Hyd.csv* and add as a *sub-content file* (along with *Bat Town Cove*, *McClough Hollow, Swagerty Cove*) to the Sequatchie Subbasin *watershed* template folder and execute AnnAGNPS
- Open the <u>*final*</u> integrated *Sequatchie Subbasin AnnAGNPS_TBL_Gaging_Station_Data_Hyd.csv* and copy *Total Streamflow* (*Mg*) values (*Column P*) and paste to *time_series_template.csv*
- Type *cmd* in command window and run *timeSeries.py script* to create *AnnAGNPS.csv* summaries of *weekly*, *monthly*, *quarterly*, and *annual* simulated AnnAGNPS data
- Open revised *Statistics_Template.xlsx* after creating *monthly* and *annual* summaries for *Integrated_AnnAGNPS_withCurve_5.csv*
- Perform statistical analysis of *Integrated_withCurve_5.csv* and *Gage .800* from formulas for *annual* (column *E* and *F*) and *monthly* (column *L* and *M*) from *simulated streamflow* (column *B* and *C*) and *observed/gage* (column *J* and *K*) summary values

B – STUDY AREA PHOTOGRAPHS

View toward east of Sequatchie-Whitwell USGS Stream Gage on Sequatchie River in Marion County.

View toward the northeast of Head of Sequatchie Spring in Cumberland Trail State Park, Cumberland County. Note the 10 to 25 degree northeast dipping Monteagle Limestone beds.

View toward the north of Devilstep Hollow Overflow Spring runoff and Head of Sequatchie River (towards building) junction in Cumberland County State Park, Cumberland County.

View toward the northwest of Devilstep Hollow Overflow Spring in Cumberland County State Park, Cumberland County.

View toward the southwest of Devilstep Hollow Cave karst window in Cumberland County State Park, Cumberland County.

View toward the northwest of Mill Cave entrance on Grassy Cove Creek, Grassy Cove in Cumberland County. (*Photo used with permission from land owner*)

View toward the southeast of Little Cove swallet in Cumberland County.

View toward the northeast of Bristow Spring in Grassy Cove, Cumberland County.

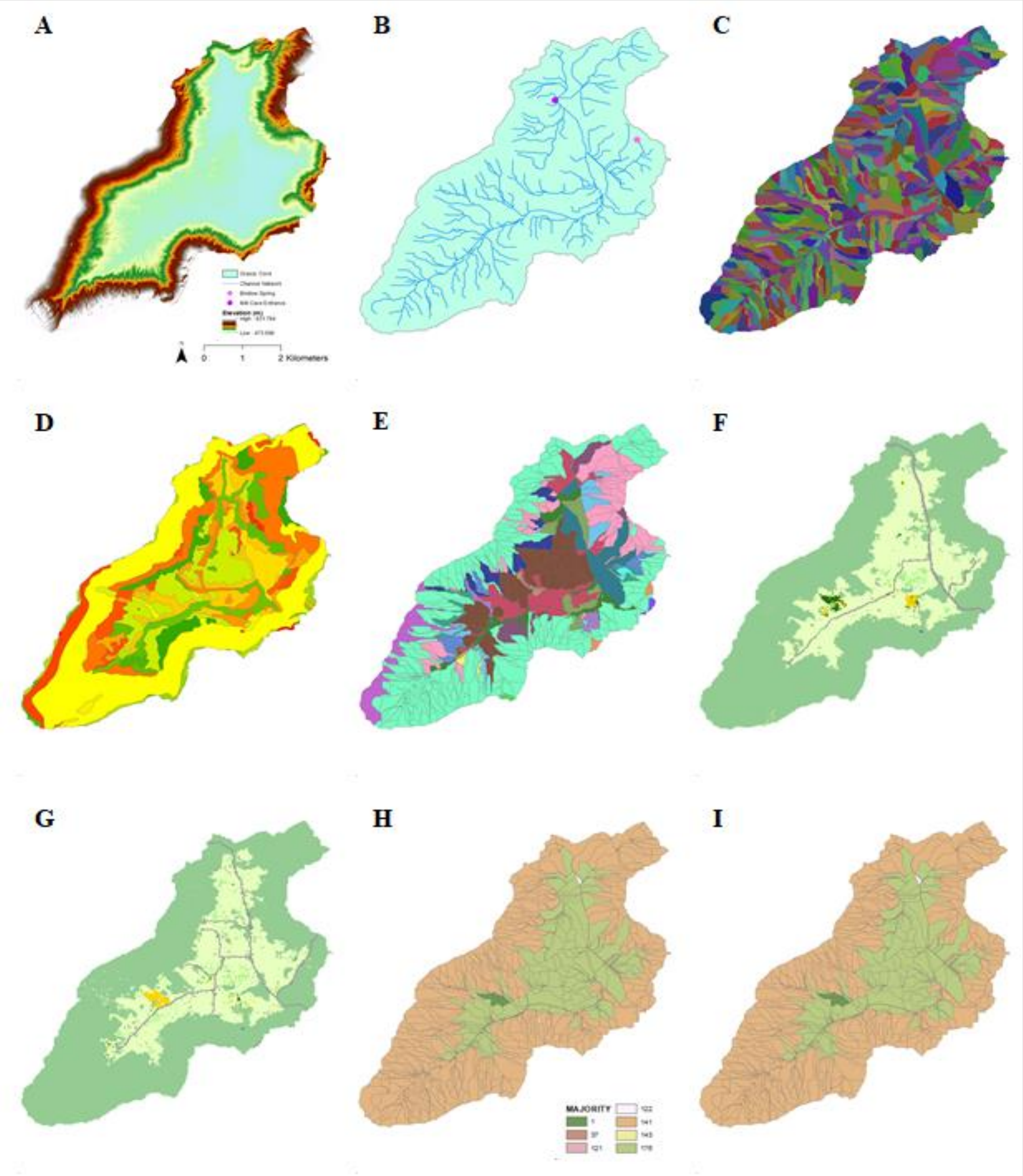
View toward the northeast of Swagerty Cove swallet in Cumberland County.

View toward the southwest of Roaring Branch Spring in Sequatchie River Subbasin, Cumberland County.

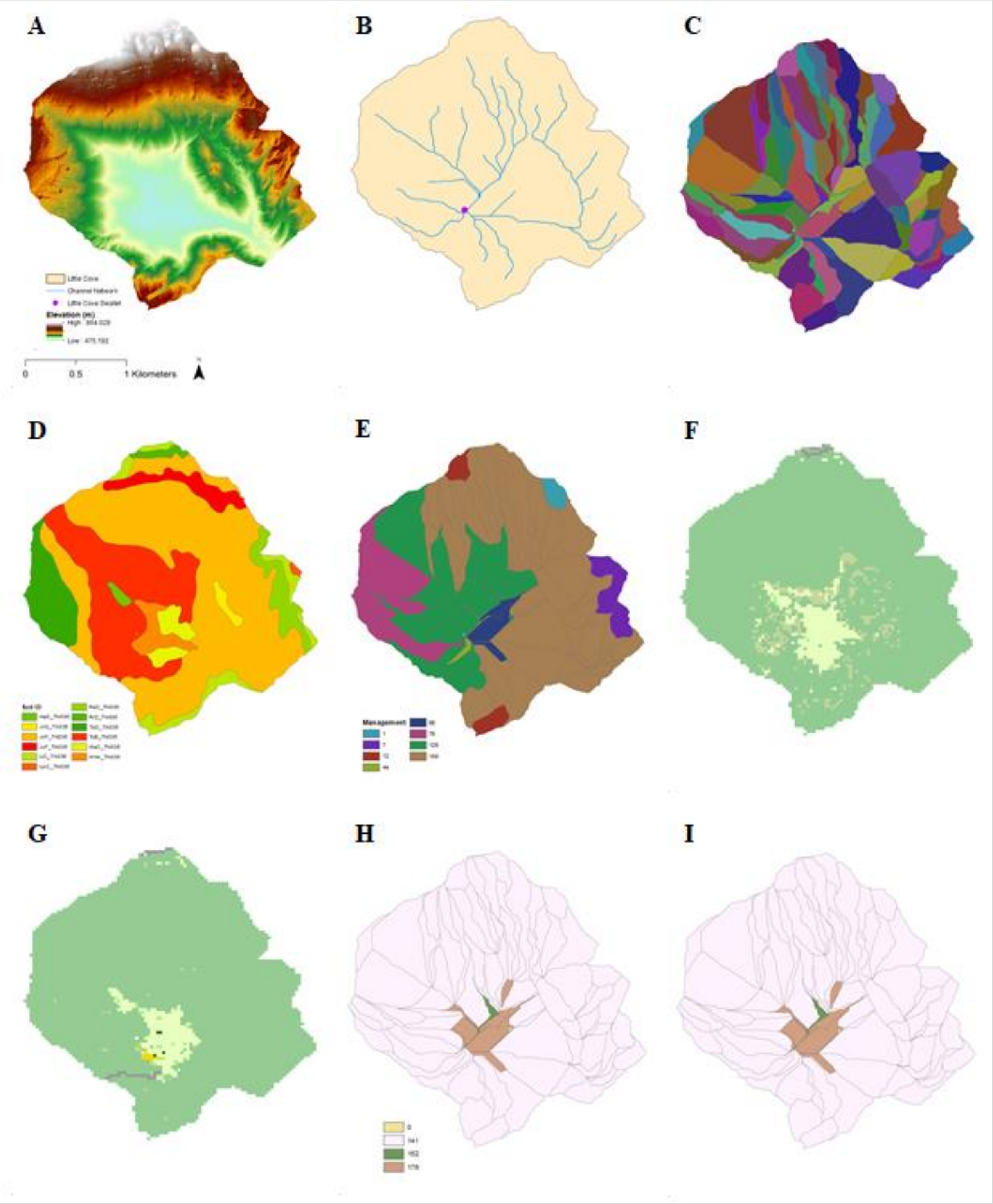
View toward the northwest of Bat Town Cove swallet in Cumberland County.

View toward the northeast of McClough Hollow swallet in Cumberland County.

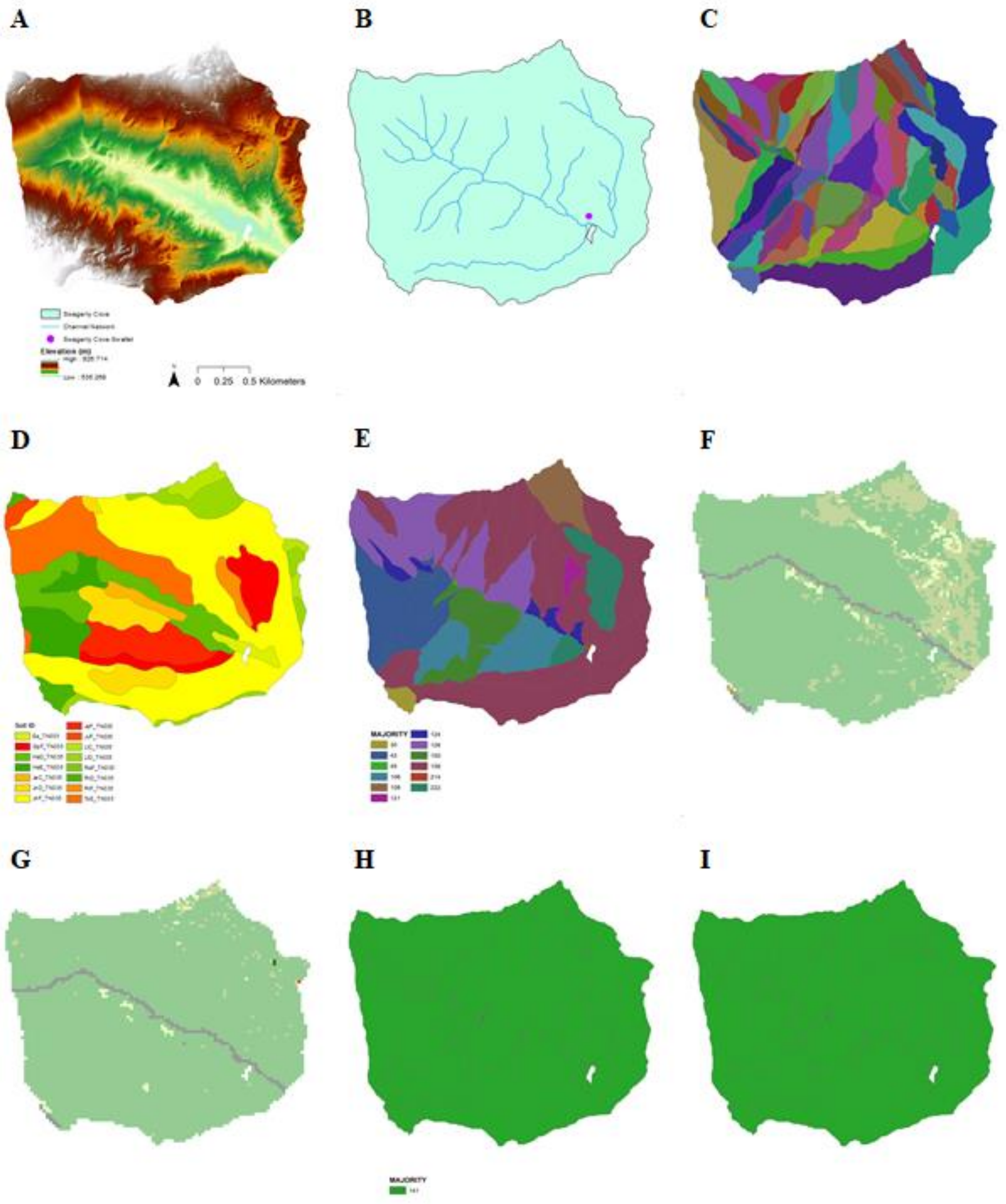
C – IMAGES OF SURROUNDING KARST VALLEYS



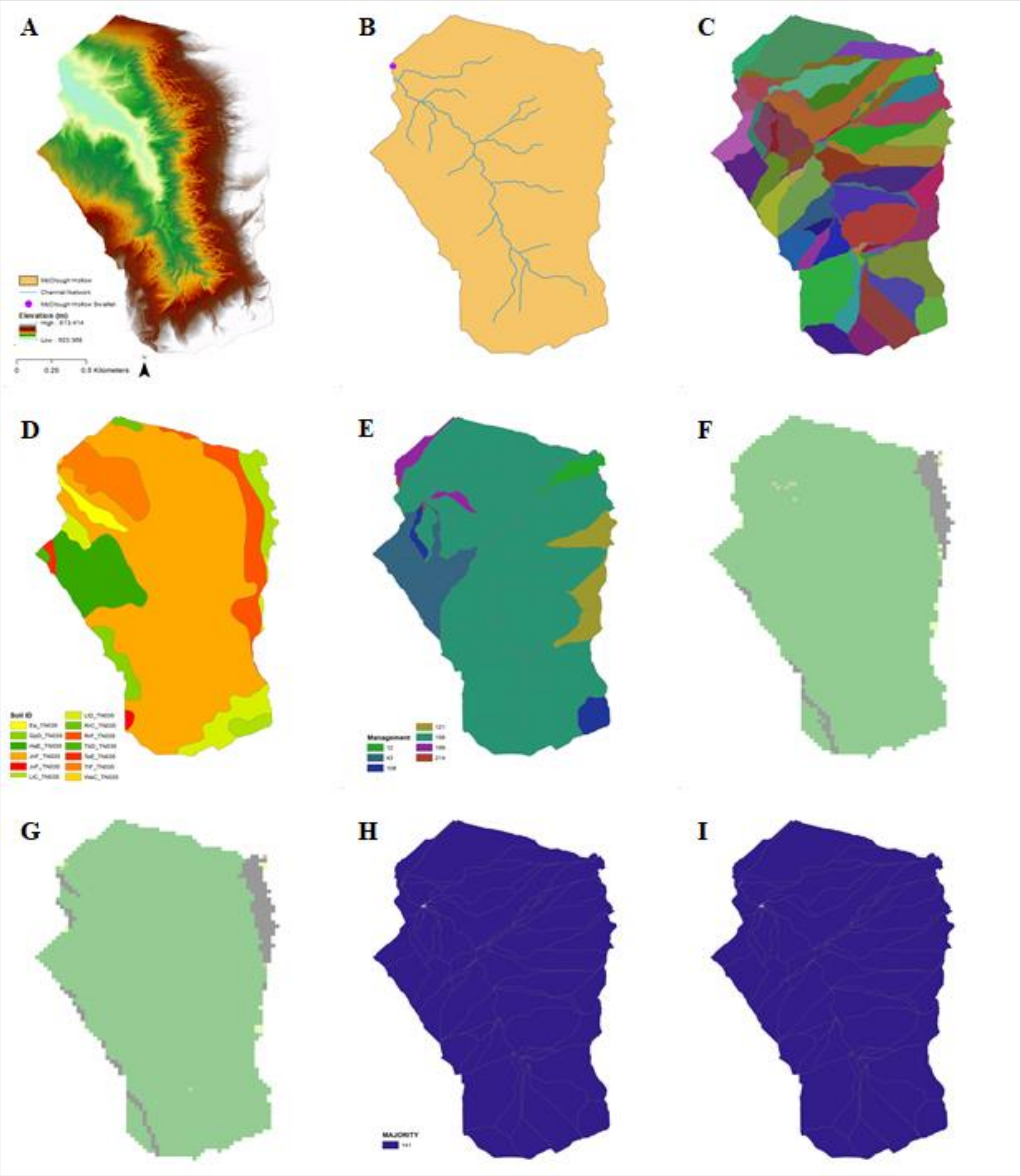
Grassy Cove: DEM (A), Channel Network (B), AnnAGNPS Cells (C), Soil Types (D), Soil Majority (E), CDL-2016 (F), CDL-2021 (G), Management-2016 (H), and Management-2021 (I).



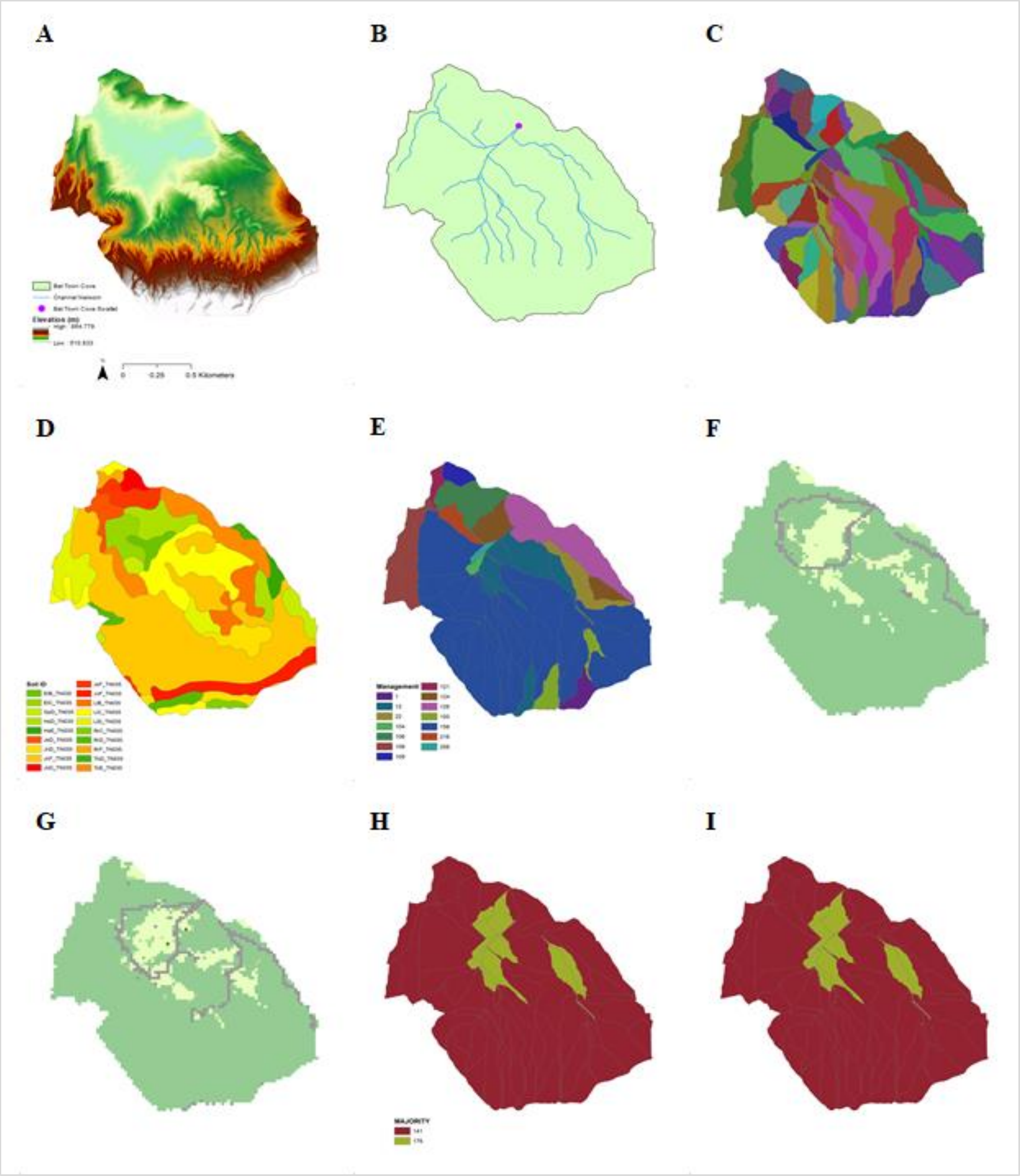
Little Cove: DEM (A), Channel Network (B), AnnAGNPS Cells (C), Soil Types (D), Soil Majority (E), CDL-2016 (F), CDL-2021 (G), Management-2016 (H), and Management-2021 (I).



Swagerty Cove: DEM (A), Channel Network (B), AnnAGNPS Cells (C), Soil Types (D), Soil Majority (E), CDL-2016 (F), CDL-2021 (G), Management-2016 (H), and Management-2021 (I).



McClough Hollow: DEM (A), Channel Network (B), AnnAGNPS Cells (C), Soil Types (D), Soil Majority (E), CDL-2016 (F), CDL-2021 (G), Management-2016 (H), and Management-2021 (I).



Bat Town Cove: DEM (A), Channel Network (B), AnnAGNPS Cells (C), Soil Types (D), Soil Majority (E), CDL-2016 (F), CDL-2021 (G), Management-2016 (H), and Management-2021 (I).