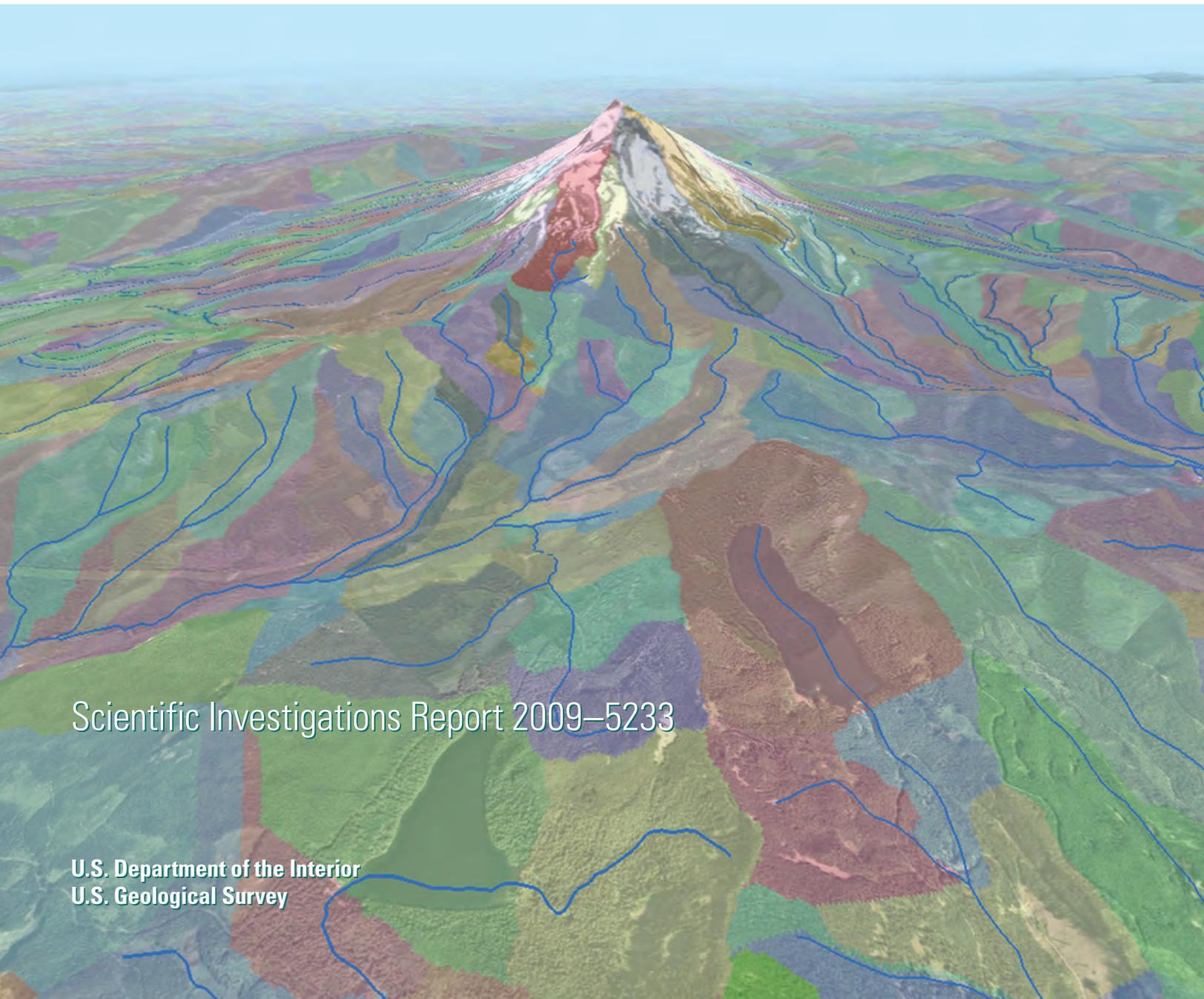


Prepared in cooperation with the U.S. Environmental Protection Agency

# Evaluation of Catchment Delineation Methods for the Medium-Resolution National Hydrography Dataset



Scientific Investigations Report 2009–5233

U.S. Department of the Interior  
U.S. Geological Survey

**Cover.** National Hydrography Dataset (NHD) streams and associated catchments overlaid on a 3-dimensional perspective view (facing southeast) of Mount Hood, Oregon. Imagery provided by Environmental Systems Research Institute (ESRI) ArcGIS Online Standard Map Services.

# Evaluation of Catchment Delineation Methods for the Medium-Resolution National Hydrography Dataset

By Craig M. Johnston<sup>1</sup>, Thomas G. Dewald<sup>2</sup>, Timothy R. Bondelid<sup>3</sup>, Bruce B. Worstell<sup>4</sup>, Lucinda D. McKay<sup>5</sup>, Alan Rea<sup>1</sup>, Richard B. Moore<sup>1</sup>, and Jonathan L. Goodall<sup>6</sup>

<sup>1</sup>U.S. Geological Survey.

<sup>2</sup>U.S. Environmental Protection Agency.

<sup>3</sup>Consultant, Woodville, VA.

<sup>4</sup>Hughes Corporation, Sioux Falls, SD.

<sup>5</sup>Horizon Systems Corporation, Herndon, VA.

<sup>6</sup>University of Texas at Austin.

Prepared in cooperation with the U.S. Environmental Protection Agency

Scientific Investigations Report 2009–5233

**U.S. Department of the Interior**  
**U.S. Geological Survey**

**U.S. Department of the Interior**  
KEN SALAZAR, Secretary

**U.S. Geological Survey**  
Marcia K. McNutt, Director

U.S. Geological Survey, Reston, Virginia: 2009

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment, visit <http://www.usgs.gov> or call 1-888-ASK-USGS

For an overview of USGS information products, including maps, imagery, and publications, visit <http://www.usgs.gov/pubprod>

To order this and other USGS information products, visit <http://store.usgs.gov>

Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this report is in the public domain, permission must be secured from the individual copyright owners to reproduce any copyrighted materials contained within this report.

Suggested citation:

Johnston, C.M., Dewald, T.G., Bondelid, T.R., Worstell, B.B., McKay, L.D., Rea, Alan, Moore, R.B., and Goodall, J.L., 2009, Evaluation of catchment delineation methods for the medium-resolution National Hydrography Dataset: U.S. Geological Survey Scientific Investigations Report 2009–5233, 88 p., available at <http://pubs.usgs.gov/sir/2009/5233>.

## Foreword

The U.S. Geological Survey (USGS) and the U.S. Environmental Protection Agency (USEPA) have long shared the goal of creating a nationally consistent digital database of important environmental and natural resource information, termed a national geospatial surface-water framework. For water resources, this means the creation of a dataset that links mapped streams to the landscape, and that essentially allows a drop of water to be followed, from where it falls on the land surface, to the stream, and all the way to the ocean. This requires an accurate geospatial data model that includes both the necessary digital data and the processes to manipulate and interpret the data.

Towards this goal, the National Hydrography Dataset (NHD) was created by the USEPA and USGS to provide the necessary stream network for this geospatial model. The NHD is a comprehensive set of digital spatial data representing the surface water of the United States and includes common water features such as lakes, ponds, streams, rivers, canals, and oceans. The NHD contains a flow-direction network that traces the water downstream or upstream. It also uses a reach addressing system that permits the linking of specific information such as wastewater discharges or water-quality monitoring sites to specific locations along the stream network.

This report represents the next step towards the national goal of linking the mapped stream network to the landscape—the identification of catchments or contributing drainage areas to individual stream segments. The catchments were used in computing streamflow and velocity estimates for NHD streams. The catchment creation effort was coordinated by the USEPA involving a team of experts from the USGS, USEPA, the private-consulting community, and academia. This report documents the collective efforts to determine the most accurate and appropriate means to delineate catchments, which establish the linkage between the NHD stream network and the landscape.

Since the conclusion of the study described in this report, the catchment delineation process that provides the most accurate catchments has been completed for the medium resolution 1:100,000-scale NHD in the conterminous United States. The results are the first national dataset that links the mapped stream network to the landscape. The data, called NHDPlus, are distributed for all States, except Alaska, through the website <http://www.epa.gov/waters>. The dataset is a milestone for geospatial data infrastructure for the Nation and we at the USGS are pleased to be a part of the team that developed it.

Matthew C. Larsen  
Associate Director for Water

## Acknowledgments

The authors wish to thank staff at contributing U.S. Geological Survey (USGS) Water Science Centers who provided streamgage and Watershed Boundary Dataset boundary data for verification purposes: William Ellis, Karen Hanson, David Litke, Mark Lowery, Theresa Olsen, and Chris Sanocki.

Thanks also goes to those who provided technical review for the project, including those at USGS: Susan Greenlee, Curtis Price, Gregory Schwarz, Peter Steeves, and David Wolock, as well as Kristine Verdin (Hughes Corporation, Sioux Falls, SD), David Maidment (University of Texas at Austin, Austin, TX), and Tracey R. Szajgin (consultant, Reston, VA).

# Contents

Foreword .....	iii
Acknowledgments .....	iv
Terminology.....	x
Abstract.....	1
Introduction.....	1
Background.....	2
National Hydrography Dataset.....	2
Concept of National Hydrography Dataset Catchments .....	5
National Elevation Dataset.....	5
Elevation Derivatives for National Applications .....	8
Watershed Boundary Dataset.....	8
Catchment Delineation Methods .....	9
Raster Seeding Method .....	9
New England Method with Watershed Boundary Dataset .....	10
New England Method without Watershed Boundary Dataset.....	15
Outlet Matching Method .....	15
Thiessen Polygon Method.....	18
Testing the Catchment Delineation Methods.....	20
Study Subbasins.....	20
Basis for Comparison.....	20
Datasets Obtained for Verification in the Comparison Analysis.....	20
Streamgage Data .....	22
Boundaries from the Watershed Boundary Dataset .....	23
Evaluation Metrics.....	23
Coefficient of Areal Correspondence.....	24
Normalizing Coefficient of Areal Correspondence .....	27
Polygon-Area Percent Error .....	30
Cumulative Drainage-Area Analysis .....	30
Convergence Analysis .....	31
Evaluation of Catchment Delineation Methods .....	32
Summary.....	35
References Cited.....	36
Glossary.....	37
Appendix. Detailed Discussion for each Study Subbasin .....	45

## Figures

1. Map showing the medium-resolution National Hydrography Dataset surface-water drainage system, central New Hampshire.....	3
2. Map showing the United States with the source map scales of the medium-resolution National Hydrography Dataset.....	4
3. Map showing an example of isolated networks in the medium-resolution National Hydrography Dataset (NHD), east-central Tennessee.....	6
4. Map showing the National Hydrography Dataset flowlines and their associated catchments .....	7
5. Schematic design showing three vector National Hydrography Dataset flowline features (blue lines) overlaid onto their raster representations, along with the corresponding catchments generated by the Raster Seeding Method.....	9
6. Schematic design showing multiple polygon features that define the catchment boundary for a National Hydrography Dataset (NHD) flowline .....	10
7. Map showing the spatial detail of the stream network and 42,000 catchments used for the New England <u>SP</u> ATIALLY <u>R</u> EFERENCED <u>R</u> EGRESSIONS <u>O</u> N <u>W</u> ATERSHED attributes (SPARROW) model.....	11
8. Maps showing (A) horizontal displacement between the National Hydrography Dataset flowlines and the Elevation Derivatives for National Applications (EDNA) synthetic streams, (B) catchment delineations using the Raster Seeding Method with the EDNA flow direction grid, and (C) catchment delineations using the New England Method with a new flow-direction grid created using the AGREE computer program with the National Elevation Dataset.....	12
9. Schematic design showing the manner in which the AGREE computer program modifies the digital elevation model (DEM) of the National Elevation Dataset (NED) to align with the National Hydrography Dataset (NHD) flowlines.....	13
10. Schematic design showing a three-dimensional perspective view of a modified digital elevation model with walling of the Watershed Boundary Dataset and stream burning, using the National Hydrography Dataset.....	14
11. Schematic design showing the Junction Intersect technique of the Outlet Matching Method .....	16
12. Schematic design showing the Drainage Point Proximity technique of the Outlet Matching Method .....	16
13. Schematic design showing the Next Downstream Area technique of the Outlet Matching Method .....	17
14. Map showing Thiessen Polygon Method catchments for National Hydrography Dataset flowlines .....	19
15. Map showing the location of the six study Subbasins in the continental United States .....	21
16. Schematic design showing how Coefficient of Areal Correspondence (CAC) score is computed .....	25
17. Map showing boundaries from the Watershed Boundary Dataset (WBD) Subwatersheds and the corresponding boundaries from the New England Method without WBD (NEM-without-WBD), National Hydrography Dataset (NHD) flowlines including those most closely corresponding to the WBD outlets, and Coefficient of Areal Correspondence scores, Pine River Subbasin, MN.....	26
18. Schematic design showing a mismatch between a Watershed Boundary Dataset (WBD) Subwatershed boundary and the corresponding New England Method with WBD boundary caused by the WBD outlet not coinciding with the downstream end of the corresponding NHD flowline .....	27

19.	Schematic design showing a Watershed Boundary Dataset Subwatershed and corresponding boundaries from the New England Methods and the Raster Seeding Method, with Coefficient of Areal Correspondence (CAC) scores before and after normalization .....	28
20.	Graph showing the non-normalized and normalized Coefficient of Areal Correspondence (CAC) scores for the New England Method without the Watershed Boundary Dataset (NEM-without-WBD) in relation to the corresponding boundaries from the Watershed Boundary Dataset labeled by National Hydrography Dataset flowline outlet ComID in the Great Egg Harbor Subbasin, NJ .....	29
21.	Graph showing the Watershed Boundary Dataset cumulative drainage-area measurements (verification) in relation to the cumulative drainage-area measurements of the methods tested, Great Egg Harbor Subbasin, NJ .....	31
22.	Graph showing percent error of convergence for drainage-area measurements for all methods tested in relation to the Watershed Boundary Dataset verification cumulative drainage-area measurements, in the Seneca River Subbasin, SC .....	32

## Tables

1.	Hierarchical structure of the Watershed Boundary Dataset .....	8
2.	Subbasin name, Hydrologic Unit Code, State, and description for six study Subbasins in the United States .....	20
3.	Description of National Water Information System streamgages, boundaries, and associated drainage-area measurements used for verification datasets.....	22
4.	Summary of Watershed Boundary Dataset Hydrologic Units used for comparative analysis of National Hydrography Dataset catchment boundary delineation methods .....	23
5.	Mean normalized Coefficient of Areal Correspondence scores for the four methods used in the Watershed Boundary Dataset comparison analysis, by study Subbasin ....	33
6.	Mean normalized polygon-area percent errors for catchment delineation methods used in the Watershed Boundary Dataset comparison analysis .....	33
7.	Mean cumulative drainage-area errors, by drainage-area size category, for catchment delineation methods for all study Subbasins, except Willapa Bay.....	34

## Conversion Factors, Datum, and Abbreviations

<b>Multiply</b>	<b>By</b>	<b>To obtain</b>
	Length	
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
	Area	
square foot (ft <sup>2</sup> )	0.09290	square meter (m <sup>2</sup> )
square mile (mi <sup>2</sup> )	2.590	square kilometer (km <sup>2</sup> )
<b>Multiply</b>	<b>By</b>	<b>To obtain</b>
	Length	
meter (m)	3.281	foot (ft)
kilometer (km)	0.6214	mile (mi)
	Area	
square meter (m <sup>2</sup> )	10.76	square foot (ft <sup>2</sup> )
square kilometer (km <sup>2</sup> )	0.3861	square mile (mi <sup>2</sup> )

Vertical coordinate information is referenced to the North American Vertical Datum of 1988 (NAVD 88).

Horizontal coordinate information is referenced to the North American Datum of 1983 (NAD 83).

### ABBREVIATIONS AND ACRONYMS USED IN THIS REPORT

AML	Arc Macro Language
CAC	Coefficient of Areal Correspondence
ComID	Common Identifier (a unique identifier for an NHD feature)
DEM	digital elevation model
DLG	Digital Line Graph
DOQ	Digital Orthophoto Quadrangle
DRG	Digital Raster Graphic
EDNA	Elevation Derivatives for National Applications
ESRI	Environmental Systems Research Institute
GIS	geographic information system
GNIS	Geographic Names Information System
HUC	Hydrologic Unit Code
ID	identifier
LIDAR	Light Detection and Ranging
NED	National Elevation Dataset
NEM	New England Method

NHD	National Hydrography Dataset
NRCS	Natural Resources Conservation Service
NWIS	National Water Information System (USGS)
RF1	USEPA Reach File Version 1.0
RF3	USEPA Reach File Version 3.0
SPARROW	<u>SP</u> ATIALLY <u>R</u> EFERENCED <u>R</u> EGRESSIONS <u>O</u> N <u>W</u> ATERSHED attributes
U.S.	United States
USEPA	United States Environmental Protection Agency
USGS	U.S. Geological Survey
WBD	Watershed Boundary Dataset

## Terminology

The Glossary at the end of this report provides a comprehensive list of key and specialized terms used in this report along with their definitions. The following information provides a brief overview of certain key terms and their relations. The initial use of a Glossary term is italicized in this report. Many of the terms used in this report are common hydrologic terms, but they may be used in a different context in this report.

Surface-water *drainage* refers to the process by which water flows downhill following the lay of the land. A *drainage area* defines the geographic extent of the land surface that drains to a specific location. A *drainage-area measurement* is the number that defines a drainage area in terms of square miles or kilometers. A *drainage-area boundary* is the polygon that defines the perimeter of a drainage area. The *incremental drainage area* for a specific hydrologic feature, such as a *stream segment*, defines the local area that drains directly into that feature. In figure 4, for example, the incremental drainage-area boundaries are shown for each hydrologic feature in the pictured drainage system. The incremental drainage area for a hydrologic feature found in the National Hydrography Dataset (NHD) is known as a *catchment*. Lastly, the *cumulative drainage area* for a specific hydrologic feature includes its incremental drainage area as well as the drainage area upstream from that feature, that is, all the way to the headwaters (farthest upstream). In figure 4, the entire yellow area shown is the cumulative drainage area for the hydrologic feature at the downstream end of the pictured drainage system. Figure 4 also illustrates that when a particular hydrologic feature is the farthest upstream feature, such as a headwater stream segment, the incremental drainage area and the cumulative drainage area for that feature are the same. When the term drainage area is used unqualified in this report, it means the same as cumulative drainage area.

This report reserves the terms *Region*, *Subregion*, *Basin*, *Subbasin*, *Watershed*, and *Subwatershed* for use when discussing the Watershed Boundary Dataset as it relates to the project. These terms are the proposed names for the six levels of *Hydrologic Units* defined within the Watershed Boundary Dataset.

# Evaluation of Catchment Delineation Methods for the Medium-Resolution National Hydrography Dataset

By Craig M. Johnston, Thomas G. Dewald, Timothy R. Bondelid, Bruce B. Worstell, Lucinda D. McKay, Alan Rea, Richard B. Moore, and Jonathan L. Goodall

## Abstract

Different methods for determining catchments (incremental drainage areas) for stream segments of the medium-resolution (1:100,000-scale) National Hydrography Dataset (NHD) were evaluated by the U.S. Geological Survey (USGS), in cooperation with the U.S. Environmental Protection Agency (USEPA). The NHD is a comprehensive set of digital spatial data that contains information about surface-water features (such as lakes, ponds, streams, and rivers) of the United States. The need for NHD catchments was driven primarily by the goal to estimate NHD streamflow and velocity to support water-quality modeling. The application of catchments for this purpose also demonstrates the broader value of NHD catchments for supporting landscape characterization and analysis.

Five catchment delineation methods were evaluated. Four of the methods use topographic information for the delineation of the NHD catchments. These methods include the Raster Seeding Method; two variants of a method first used in a USGS New England study—one used the Watershed Boundary Dataset (WBD) and the other did not—termed the “New England Methods”; and the Outlet Matching Method. For these topographically based methods, the elevation data source was the 30-meter (m) resolution National Elevation Dataset (NED), as this was the highest resolution available for the conterminous United States and Hawaii. The fifth method evaluated, the Thiessen Polygon Method, uses distance to the nearest NHD stream segments to determine catchment boundaries.

Catchments were generated using each method for NHD stream segments within six hydrologically and geographically distinct Subbasins to evaluate the applicability of the method across the United States. The five methods were evaluated by comparing the resulting catchments with the boundaries and the computed area measurements available from several verification datasets that were developed independently using manual methods.

The results of the evaluation indicated that the two New England Methods provided the most accurate catchment boundaries. The New England Method with the WBD

provided the most accurate results. The time and cost to implement and apply these automated methods were also considered in ultimately selecting the methods used to produce NHD catchments for the conterminous United States and Hawaii.

This study was conducted by a joint USGS–USEPA team during the 2-year period that ended in September 2004. During the following 2-year period ending in the fall of 2006, the New England Methods were used to produce NHD catchments as part of a multiagency effort to generate the NHD streamflow and velocity estimates for a suite of integrated geospatial products known as “NHDPlus.”

## Introduction

The medium-resolution (1:100,000-scale) *National Hydrography Dataset* (NHD) (U.S. Geological Survey, 1999a) is a comprehensive set of digital spatial data that contains information about surface-water features such as lakes, ponds, streams, and rivers. The NHD was produced by and is maintained as a collaborative effort between the U.S. Geological Survey (USGS), the U.S. Environmental Protection Agency (USEPA), and State cooperators. These organizations have linked sites from many different water-related databases, including streamgages, water-quality monitoring sites, point-source discharge sites, and impaired waters, to the underlying NHD by assigning an NHD stream address to each site. These linkages enable upstream/downstream query and analysis of the sites and their associated data.

To enhance this basic NHD capability, the USGS and USEPA established the goal of developing estimates of *streamflow* and velocity for each NHD stream segment in the United States. The streamflow and velocity estimates support pollutant dilution modeling and other water-resource modeling applications. Streamflow combined with pollutant concentration measurements collected through water-quality monitoring activities provide a means to estimate the total amount of pollutants (pollutant load) in a stream. Stream velocity and measured distance provide a means to estimate average travel times between contamination sources and other locations in

## 2 Evaluation of Catchment Delineation Methods for the Medium-Resolution National Hydrography Dataset

the stream network. To develop these estimates of streamflow and velocity for NHD features using regression techniques, *drainage-area measurements* and attributes are needed. From a practical standpoint, given the large number of NHD features, this is best accomplished by delineating a *catchment* for each relevant feature. A catchment is the *incremental drainage area* for a single NHD hydrologic feature, such as a stream segment. The catchment boundaries enable simple computation of the incremental drainage-area measurement and may also be overlaid on other ingredient data for computing streamflow and velocity (such as temperature and precipitation) to associate these attributes with the NHD features.

The resulting NHD catchments are expected to support a variety of applications beyond their initial use in estimating streamflow and velocity. Catchments are essential in determining the relations between any land-surface information (such as soil types, land use, land cover, and crop maps) and streams and waterbodies. In addition, catchments can be an aid in understanding the relations between atmospheric information (such as precipitation, temperature, and air-pollutant deposition), and streams and waterbodies. Thus, by relating these land-surface and atmospheric characteristics to streams, users have the ability to build comprehensive, integrated models of the land, atmosphere, and surface waters.

The goal of estimating streamflow and velocity for the NHD led the USGS and USEPA to integrate the NHD (1:100,000-scale) with the *National Elevation Dataset* (30-m) (U.S. Geological Survey, 1999b) and the *Watershed Boundary Dataset* (WBD) (where it existed) (U.S. Department of Agriculture, 2004) for purposes of determining NHD catchment boundaries. This integration made possible many additional geospatial data products found in what is now known as *NHDPlus* (U.S. Environmental Protection Agency, 2006).

The purpose of this report is to present the results of different NHD catchment boundary delineation methods that were tested for six study *Subbasins* from across the United States, each having distinct hydrologic characteristics. The delineation methods consisted of the New England Method with the WBD (NEM-with-WBD), the New England Method without WBD (NEM-without-WBD), the Raster Seeding Method, the Outlet Matching Method, and the Thiessen Polygon Method.

The catchment boundary delineations produced by the different methods were compared with verification datasets that included existing *drainage-area boundaries* in *geographic information system* (GIS) format and computed drainage-area measurements. The evaluation of the catchment drainage-area measurements and boundaries was conducted using analysis metrics, including the *Coefficient of Areal Correspondence* (CAC), polygon-area percent error, *cumulative drainage area*, and drainage-area convergence analysis. Detailed findings for each study Subbasin and method are presented in the appendix.

The evaluation described in this report was used to determine the most accurate catchment boundary delineation method for producing NHD catchments. Though not presented

in this report, the time and cost to implement and apply these automated methods also were considered in ultimately selecting the method used to produce NHD catchments for the conterminous United States and Hawaii.

## Background

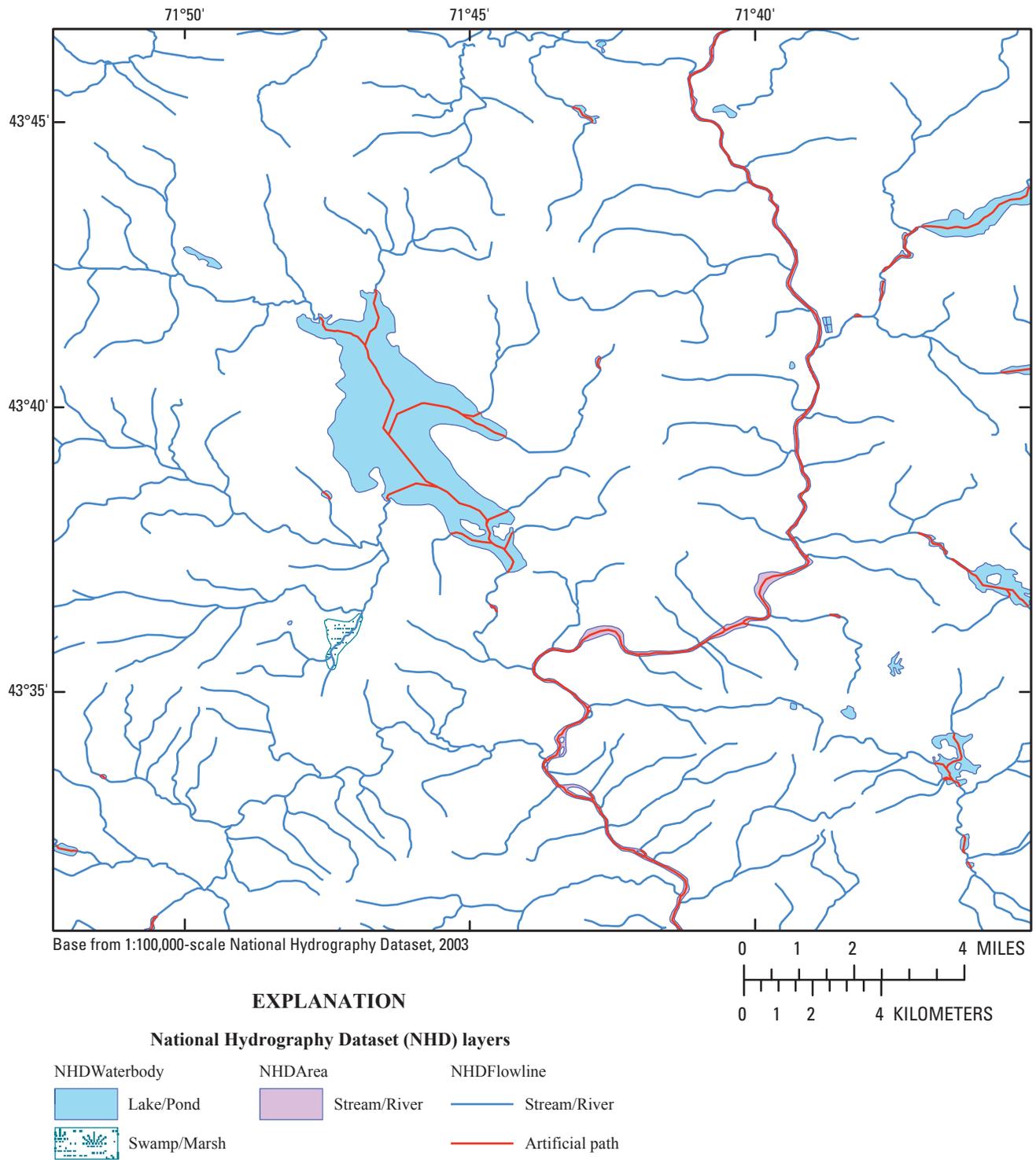
The NHD stream network is discussed in this section along with the concept of NHD catchments. Also described in this section are the other national geospatial datasets used by the methods evaluated.

### National Hydrography Dataset

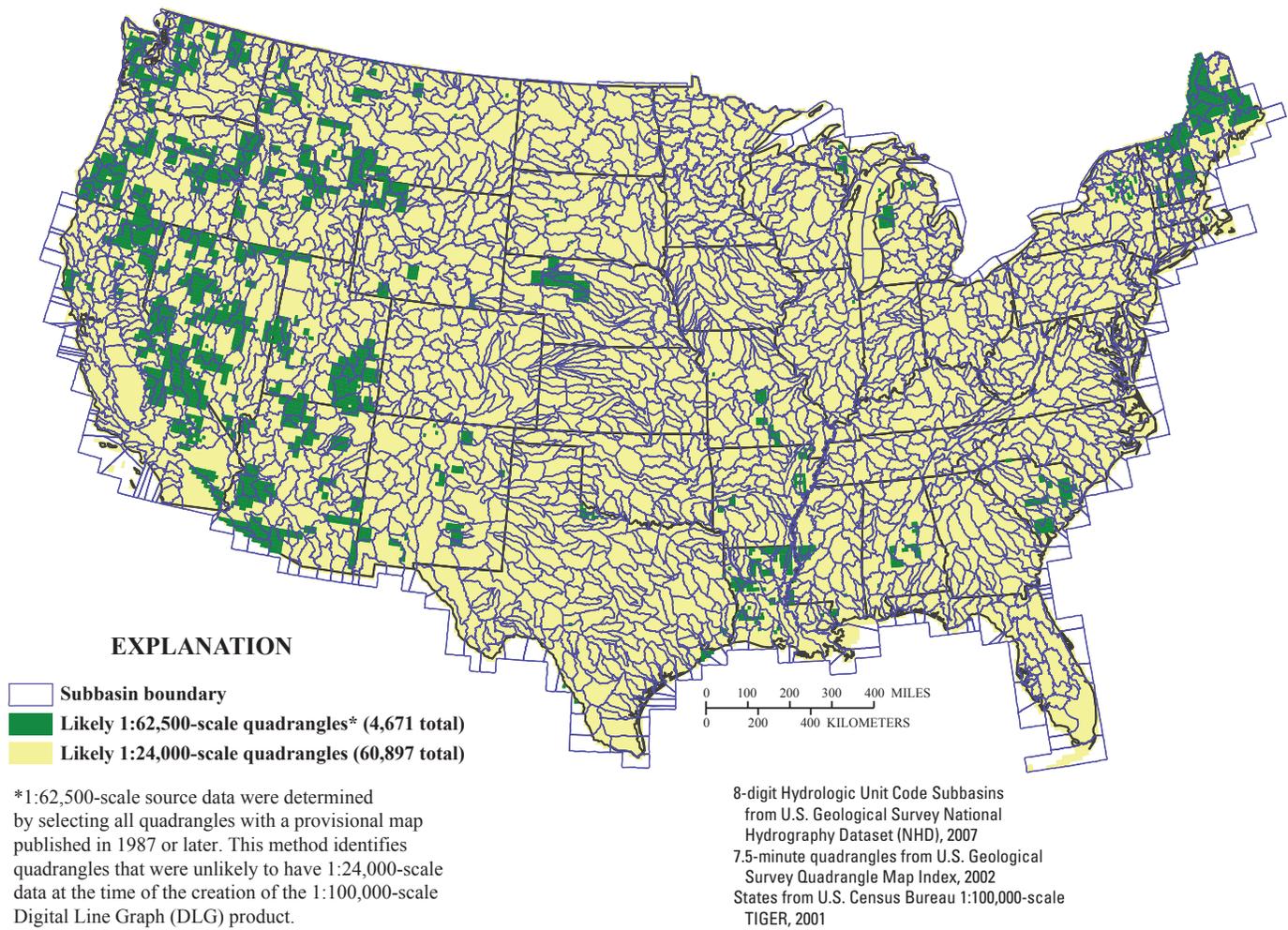
The linear surface-water network in the NHD is represented by linear features (such as streams and canals) along with *artificial path* features that provide flow paths through areal features (lakes, ponds, and other waterbodies) in the network. An example of the linear and areal features in the NHD is shown in figure 1.

In the medium-resolution NHD, the content of the USGS 1:100,000-scale *Digital Line Graph* (DLG) hydrography data (U.S. Geological Survey, 1996) is integrated with the reach-related information from *USEPA's Reach File Version 3.0* (RF3) (U.S. Environmental Protection Agency, 1994). DLG hydrography data are digital representations of the surface-water features shown on the USGS 1:100,000-scale topographic maps. As shown in figure 2, most of the 1:100,000-scale DLGs for the United States were developed from 1:24,000-scale USGS 7.5-minute topographic quadrangle maps. At the time the 1:100,000-scale DLGs were produced, 1:24,000-scale maps did not exist for the entire country and, therefore, USGS relied on older USGS 1:62,500-scale 15-minute quadrangle maps for these areas (fig. 2). The data were compiled by 30 × 60-minute quadrangles. The source hydrography was photo-reduced, edge-matched, and digitized at the 1:100,000 scale.

The NHD contains unique permanent identifiers, known as *reach codes*, for most linear and areal features. This includes coastline features that represent the shorelines of the Great Lakes, the Atlantic and Pacific Oceans, and the Gulf of Mexico. Reach codes provide the mechanism for linking many types of water-resource information to NHD hydrologic features in the same manner that street addresses are used to associate houses with specific locations on streets. The NHD provides a national framework of reach addresses for linking water-related entities, such as point-source pollution discharges, impaired water designations, drinking-water intakes, water-quality monitoring sites, wild and scenic river designations, and streamgages. The USGS and USEPA have linked numerous water-related databases to the NHD by assigning NHD stream (*reach*) addresses in this manner (<http://www.epa.gov/waters>). The NHD provides the analytical framework to support hydrologic modeling,



**Figure 1.** The medium-resolution National Hydrography Dataset surface-water drainage system, central New Hampshire.



**Figure 2.** The United States with the source map scales of the medium-resolution National Hydrography Dataset.

water-quality studies, characterization of fish and other aquatic organism habitat, and many other applications. Some major characteristics of the NHD are listed below.

- It is a feature-based dataset that interconnects and uniquely defines the stream segments that comprise the surface-water drainage system of the United States.
- Permanent identifiers known as “reach codes” are assigned to NHD stream network and waterbody features.
- Reach codes on linear features, along with a measurement system, form “reach addresses” that establish the locations of surface-water entities on the NHD surface-water drainage network.
- Once the upstream/downstream relations of water-related entities and associated information are linked to the NHD by their reach addresses, they can be analyzed using various software tools.
- The medium-resolution NHD is based on the content of the USGS 1:100,000-scale DLG data, giving it spatial resolution similar to that found on the 1:100,000-scale USGS topographic maps.
- The high-resolution NHD for the conterminous United States, which is based on a 1:24,000 scale or better source, was finished in 2007 after this study had been completed.
- NHD datasets are maintained using one consistent spatial reference—geographic (latitude/longitude coordinates in decimal degrees) referenced to the North American Datum of 1983.
- Names with *Geographic Names Information Systems* (GNIS) identification numbers are included for lakes, other water bodies, and many stream courses (<http://nhd.usgs.gov/gnis.html>).

- The data contain many feature attributes such as flow directions, names, and stream levels.
- The NHD is maintained through a national stewardship program led by USGS and implemented by Federal and State stewards.

The *flowline* is the basic unit of the NHD linear surface-water network. An NHD reach is a uniquely identified linear feature that consists of one or more flowlines. Flowline features are segmented at stream confluences and at changes of NHD feature types and attributes along the network. Flowline features include streams, rivers, intermittent streams, canals, ditches, pipelines, coastlines, and artificial paths (through areal water bodies). The set of connected flowlines, with known flow direction, forms the drainage network.

Isolated networks are those networks with known flow direction that stop at a dead-end or terminus (that is, without connecting to a downstream flowline). In the NHD, stream features may be present on one source quadrangle, but not continued on the adjacent source map because of inconsistently applied production procedures across quadrangles. In these situations, a stream mapped as flowing downstream toward an adjacent quadrangle may seem to end at the quadrangle boundary. Actually, the stream does continue onward to connect to the downstream network, but this is not reflected in the NHD. This is an example of an erroneous isolated network associated with quadrangle boundaries. Isolated networks are not always found at quadrangle boundaries, and they may or may not be legitimate. Legitimate isolated networks are those streams that percolate into the ground or deliver water to playas such as the evaporative salt flats found in the western part of the United States. An example of two isolated NHD networks in a karst terrain setting is shown in figure 3.

The NHD *flow table* was of particular importance during this study. The flow table is a database table that contains the interconnections between the flowlines. Only flowlines with known flow directions are included in the flow table. Where two flowlines exchange water, a flow relationship is established within the flow table. In cases where two coastline features touch, a nonflowing relationship is included in the flow table. Where a stream flows into the Pacific or Atlantic Ocean, the Great Lakes, or the Gulf of Mexico, a nonflowing relationship in the flow table connects the stream to the coastline. These nonflowing relations enable network navigation along the coasts to identify water-related information that has been linked to the coastlines.

For all the methods evaluated, the NHD flowlines were first projected into the National Albers Equal-Area coordinate system and converted to an *ArcInfo* coverage format. Catchments were delineated only for NHD flowlines having known flow direction and, for that reason, are included in the flow table. All features with unknown flow direction were deleted from the ArcInfo coverage prior to catchment boundary delineation.

## Concept of National Hydrography Dataset Catchments

The goal of this study was to determine the most accurate method for delineating catchment boundaries for NHD flowlines. An NHD catchment is the land surface that drains directly into an NHD flowline without first flowing into an upstream flowline (fig. 4). A catchment boundary is a polygon that extends upstream from the terminus (downstream end) of the flowline and defines the incremental drainage area for that flowline. Combining the catchment boundary for a specific flowline with the catchments associated with all upstream flowlines defines the cumulative drainage area for that flowline (fig. 4).

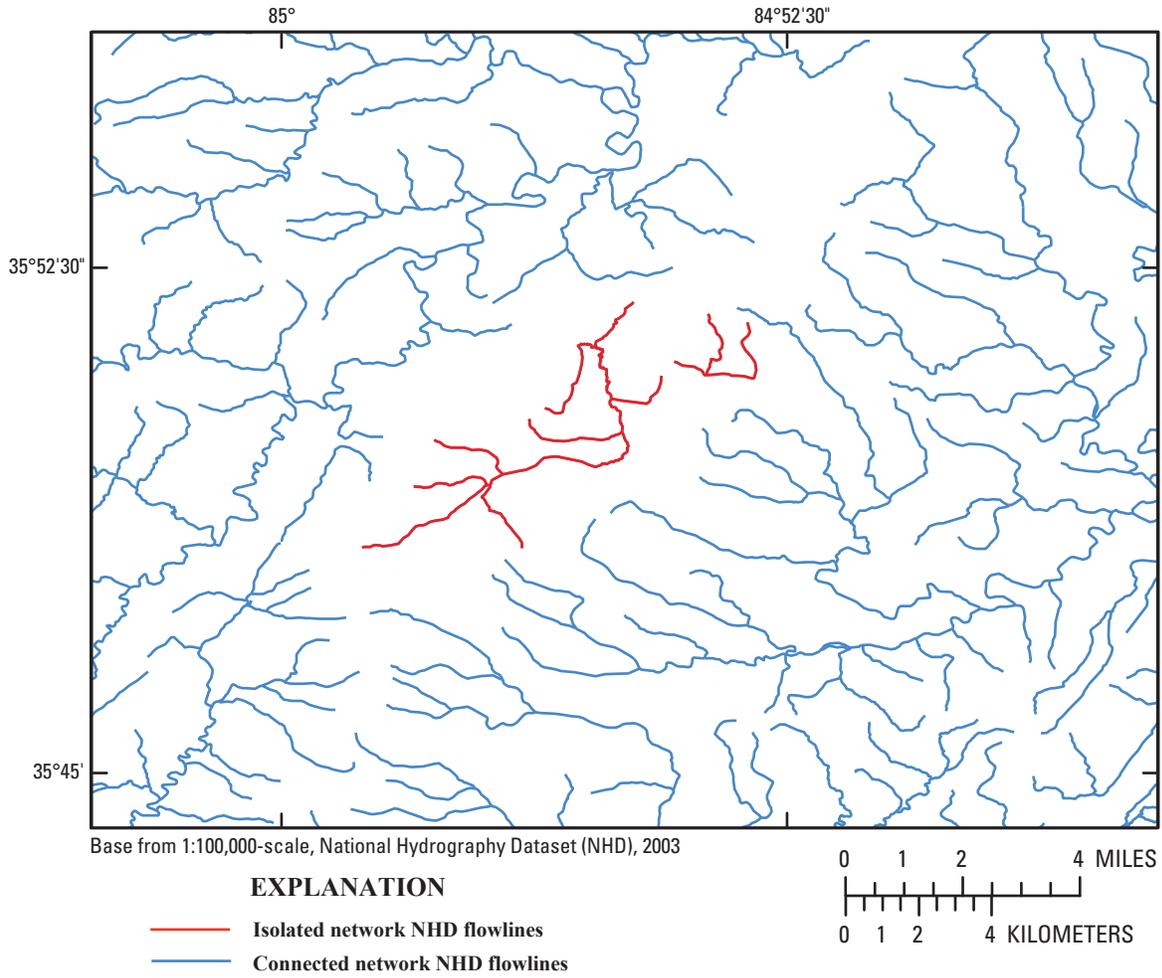
## National Elevation Dataset

The *National Elevation Dataset* (NED) (U.S. Geological Survey, 1999b) is a *raster* (grid) product assembled by the USGS (<http://ned.usgs.gov>) that provides nationwide elevation data in a seamless form referenced to a common coordinate system. The NED contains data corrections made during the assembly process to minimize production artifacts, perform edge matching, and fill sliver areas of missing data.

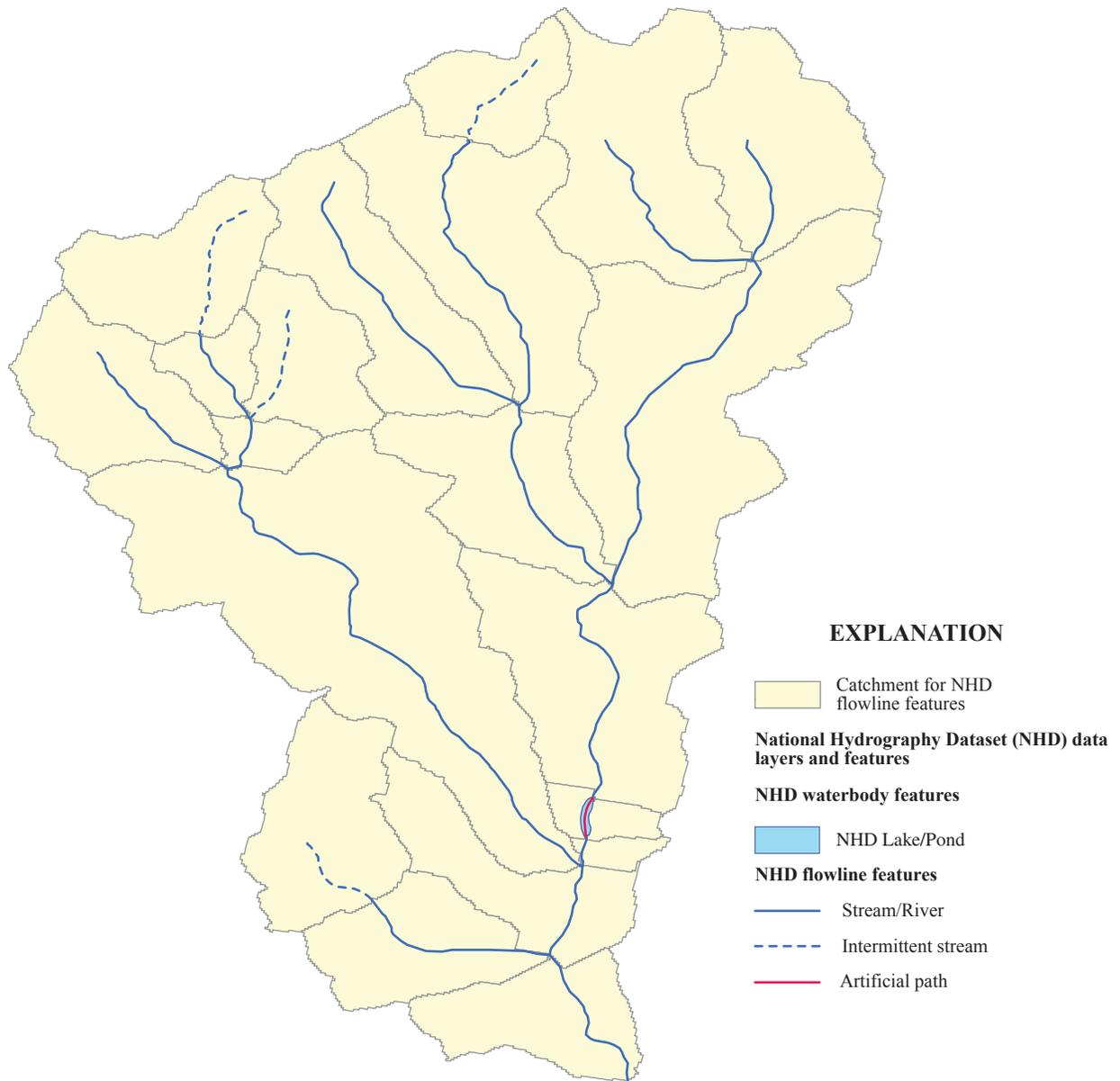
NED is available in multiple resolutions. A one arc-second resolution NED (approximately 30-m cell size) is available for the conterminous United States, Hawaii, Puerto Rico, and island territories. The one arc-second resolution NED currently provides the most extensive coverage for the United States, with the exception of Alaska (Alaska has a two-arc-second resolution, or approximately a 60-m cell size). A NED of one-third arc-second resolution (approximately 10-m cell size) is available for Hawaii and about two-thirds of the conterminous United States. The highest resolution NED of one-ninth arc-second (approximately 3-m cell size) is available for selected areas across the United States and is derived from large-scale planimetric contours or newer digital topographic capture methods, such as *Light Detection and Ranging* (LIDAR).

As higher resolution or higher quality elevation data become available, the NED is updated bimonthly to incorporate best-available coverage. The NED data dictionary, release notes, and update information can be found on the NED Web site at <http://ned.usgs.gov/ned/downloads.asp>. A visual index documenting the individual source elevation datasets is found in the Data Source Index for NED on the Web site. The *digital elevation models* (DEMs) used to construct the NED have a wide range of source-data characteristics. The NED is designed to incorporate the most accurate source data for the one-ninth, one-third, and one arc-second grid spacing. For instance, if a new one-ninth arc-second NED is produced, the coarser resolutions of one-third and one arc-second NED are updated by resampling the one-ninth arc-second data. Seamless NED data are distributed through the Seamless Data Distribution System (<http://seamless.usgs.gov/>).

6 Evaluation of Catchment Delineation Methods for the Medium-Resolution National Hydrography Dataset



**Figure 3.** An example of isolated networks in the medium-resolution National Hydrography Dataset (NHD), east-central Tennessee.



**Figure 4.** The National Hydrography Dataset flowlines and their associated catchments.

For the purpose of this study, the one arc-second NED was selected as the elevation data source for delineating NHD catchment boundaries, as it was the highest resolution NED available for the conterminous United States and Hawaii. National coverage is essential in meeting the goal of estimating streamflow and velocity for the flowlines in the medium-resolution NHD.

## Elevation Derivatives for National Applications

The *Elevation Derivatives for National Applications* (EDNA) is a suite of data layers that are commonly used for topographic analyses. These layers include aspect, shaded relief, slope, and synthetic streamlines (<http://edna.usgs.gov/>). Several of these layers have importance for hydrologic applications. The EDNA data are based on the one arc-second NED. In EDNA, elevation data have been projected to a coordinate system appropriate for direction and area analysis and processed to remove spurious sinks so that flow routing across the entire elevation surface can be achieved.

The EDNA flow direction and flow accumulation grids provide a digital representation of land-surface flow. For each cell in the grid, the EDNA flow direction grid determines the direction of flow in relation to its neighboring cells. The flow direction grid makes it possible to delineate drainage-area boundaries from EDNA data. For each cell in the grid, the EDNA flow accumulation grid provides the structure to determine the number of upstream cells that flows into that cell. Using the flow accumulation grid, EDNA “synthetic” stream lines were generated to an upstream threshold of 1.74 square miles (mi<sup>2</sup>); that is, the synthetic lines extend upstream until the total area of the upstream cells reached 1.74 mi<sup>2</sup>. EDNA catchment boundaries are then produced for each synthetic stream. These synthetic stream lines do not always follow the same flow path as the mapped NHD flowlines, and in some cases differ considerably. In addition, all EDNA synthetic stream lines begin and end at confluences, whereas NHD flowlines begin and end at both confluence and NHD feature-type changes.

The development of the EDNA dataset is envisioned as a three-stage process. Stage 1, the current stage of the dataset, is mostly automated. The main component of this step is a blind pass filling of spurious sinks to create the depressionless DEM. Stage 2 focuses on identifying potential errors in the data with the assistance of local agencies that use EDNA stage 1 data and who are familiar with the hydrology in their geographic area. In stage 3, the errors identified in stage 2 will be addressed to create an improved dataset (Kost and others, 2003).

## Watershed Boundary Dataset

The Watershed Boundary Dataset (WBD) is a national seamless framework for the compilation of digital hydrologic area features delineated using manual interpretive methods

from 1:24,000-scale USGS topographic maps (U.S. Department of Agriculture, 2004). These WBD hydrologic area features, also known as *Hydrologic Units*, support a variety of water-quality reporting uses. The term *Subwatershed* refers to the subdivisions of a *Watershed* (table 1) in a hierarchal system that subdivides the landscape into successively smaller Hydrologic Units (Seaber and others, 1987). Hydrologic Units are spatially nested within one another and each unit is assigned a unique *Hydrologic Unit Code* (HUC). HUCs are 2 to 12 digits in length, based on the hierarchy of geographic areas defined within the WBD. The largest Hydrologic Units subdivide the United States into 22 major geographic areas, called *Regions*, and are identified by unique 2-digit HUCs. A Subwatershed, the smallest Hydrologic Unit, is identified by a unique 12-digit HUC and is generally 20 to 60 mi<sup>2</sup> in size. The hierarchal structure of the Hydrologic Units contained in the WBD is given in table 1. These hierarchal relations enable the aggregation of smaller Hydrologic Units into larger Hydrologic Units based upon the HUCs. Many individual WBD hydrologic units are not complete drainage areas since they do not extend all of the way upstream to the headwaters.

The WBD is the culmination of collaborative efforts by Federal, State, and local agencies. The WBD delineations go through various stages of review before final certification and entry into the national WBD database. During the review process, delineations are checked for accuracy against 1:24,000-scale *Digital Raster Graphics* (DRG) to ensure that subdivisions meet defined hydrologic and cartographic criteria. DRGs are digital versions of the same USGS topographic maps that served as the source for the medium-resolution NHD and, as such, align well with NHD features. WBD interstate edge-matching issues are resolved during the review process. Once all Hydrologic Units in a State are certified, the WBD for that State becomes available to the public. The WBD is maintained by the Natural Resources Conservation Service (NRCS) in partnership with USGS and State cooperators. For more information, visit the WBD Web site at (<http://www.ncgc.nrcs.usda.gov/products/datasets/watershed>).

**Table 1.** Hierarchal structure of the Watershed Boundary Dataset.

Hydrologic Unit name	Hydrologic Unit Code (HUC) length in digits	Approximate number nationwide
Region	2	22
Subregion	4	222
Basin	6	352
Subbasin	8	2,149
Watershed	10	22,000
Subwatershed	12	160,000

## Catchment Delineation Methods

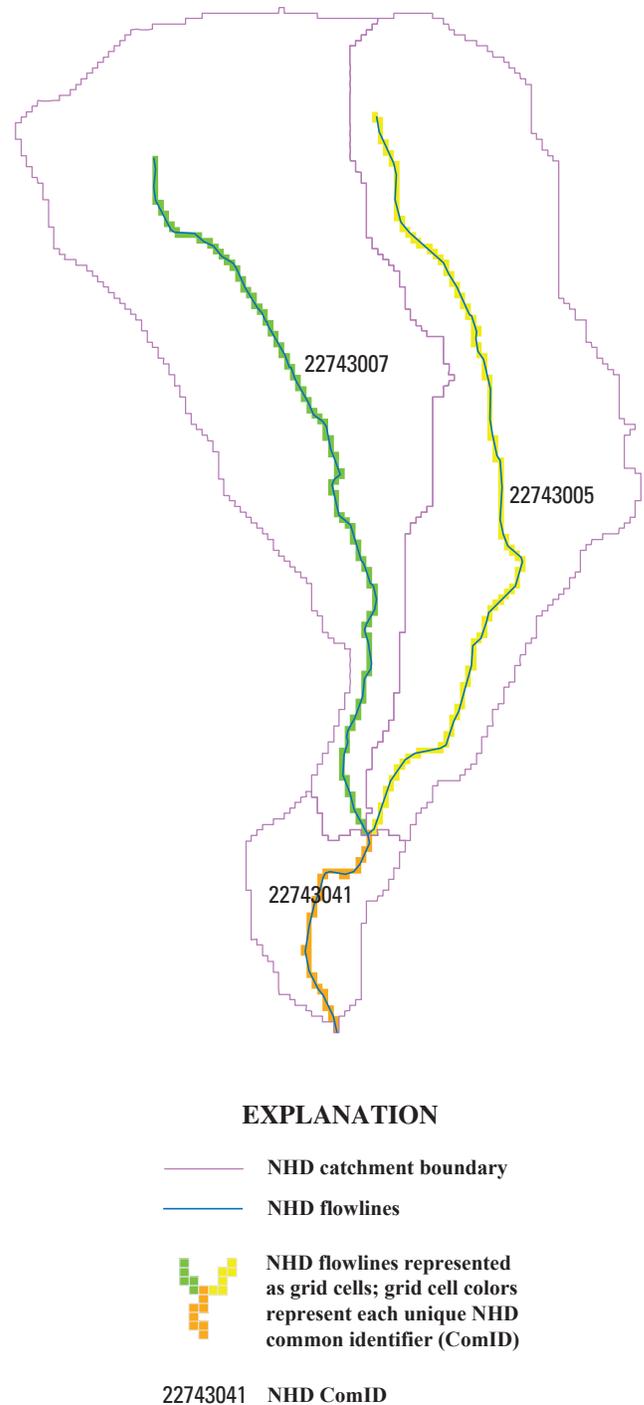
Five NHD catchment delineation methods were tested. Four methods used elevation information as the basis for catchment delineation: the Raster Seeding Method; two variants of an approach first used in a USGS New England study (termed the “New England Method”)—one variant with WBD and one variant without WBD—and the Outlet Matching Method. A fifth method, the Thiessen Polygon Method, determines catchment boundaries by evenly splitting the distance between nearby NHD flowlines. As previously noted, catchments are only generated for NHD flowlines with known flow direction (that is, flowlines with flow relations defined in the NHD flow table).

### Raster Seeding Method

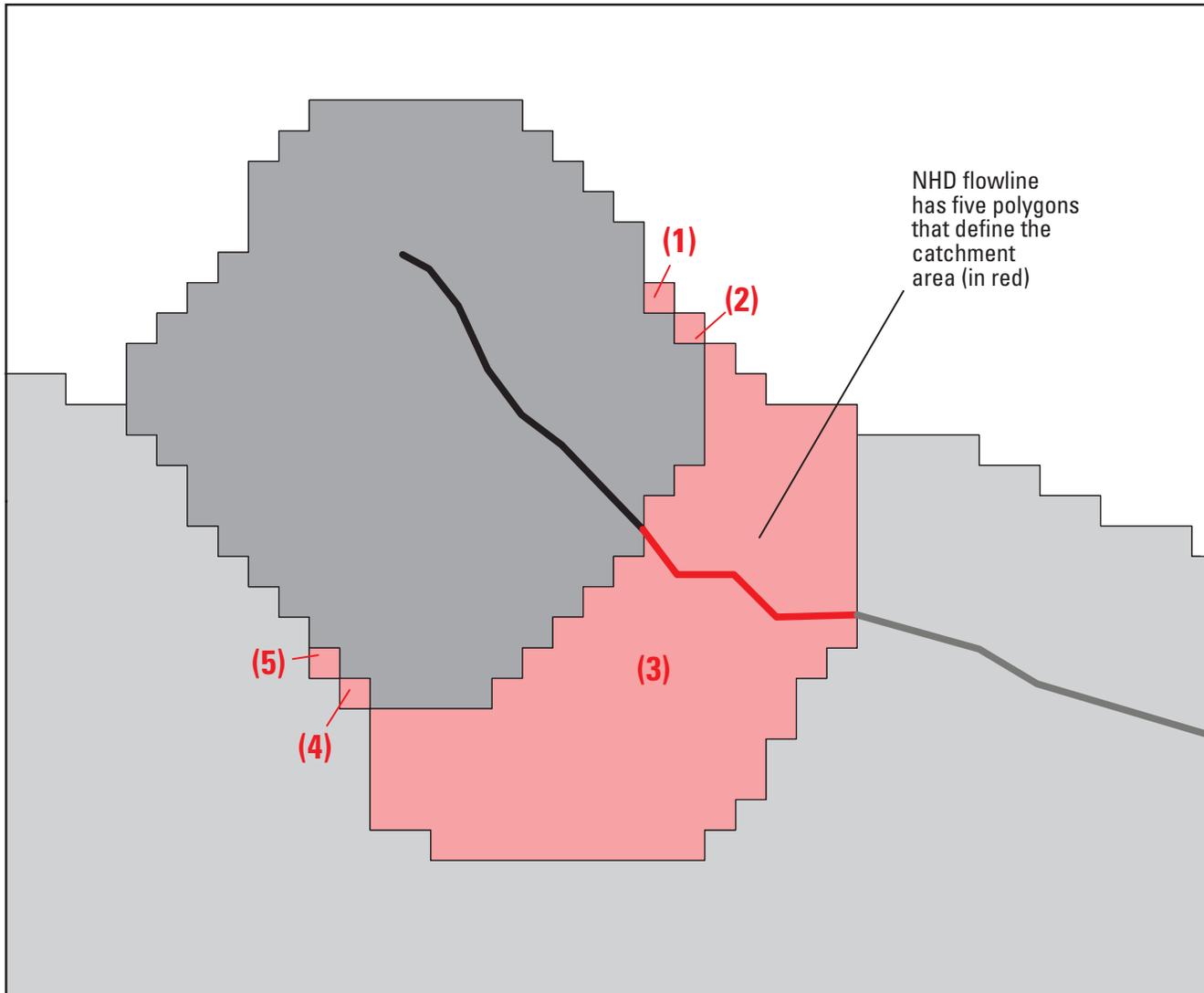
The Raster Seeding Method uses the NHD and EDNA datasets. The method converts the vector NHD flowlines into a raster ArcInfo Grid (Environmental Systems Research Institute, Inc., 1999) data model to be used as “seeds” for elevation-based drainage-area delineation (fig. 5). This approach has been used in other studies to delineate catchments for cartographic streams (Nolan and others, 2003).

In the Raster Seeding Method, NHD flowline features are rasterized to a 30-m × 30-m grid cell resolution and aligned to the EDNA flow-direction grid. The NHD flowline common identifier (ComID) codes are used as the grid values in the raster representation of the flowlines. The rasterized NHD flowlines and the EDNA flow direction grid are the required inputs to the Environmental Systems Research Institute (ESRI) ArcInfo Grid WATERSHED command (Environmental Systems Research Institute, Inc., 1999), which is used to generate catchments for each unique NHD flowline. The WATERSHED command delineates catchments upstream from seed cells (in this case, those cells that make up the rasterized NHD flowlines) using an algorithm developed by Jensen and Dominique (1988).

To facilitate the spatial analyses used to compare results of the different catchment boundary delineation methods, the boundaries were converted from raster format to vector format (ArcInfo coverage data model). Vector catchment boundaries were used for many of the evaluation metrics discussed later in this report. The coverage format associates all noncontiguous areas for any multipart catchments by assigning them the same ComID. They are recognized as one vector feature in the ArcInfo region subclass data model. These multiple-part polygon features are created during the raster-to-vector conversion process, when one or more cells with directional flow traveling diagonally into an adjacent cell along a catchment boundary result in separate polygons in the vector data model (fig. 6). In addition, catchment boundaries are not generated by the WATERSHED command for very short NHD flowlines (less than about 42 m) because of the 30-m resolution of the ingredient EDNA dataset.



**Figure 5.** Schematic design showing three vector National Hydrography Dataset (NHD) flowline features (blue lines) overlaid onto their raster representations, along with the corresponding catchments generated by the Raster Seeding Method.



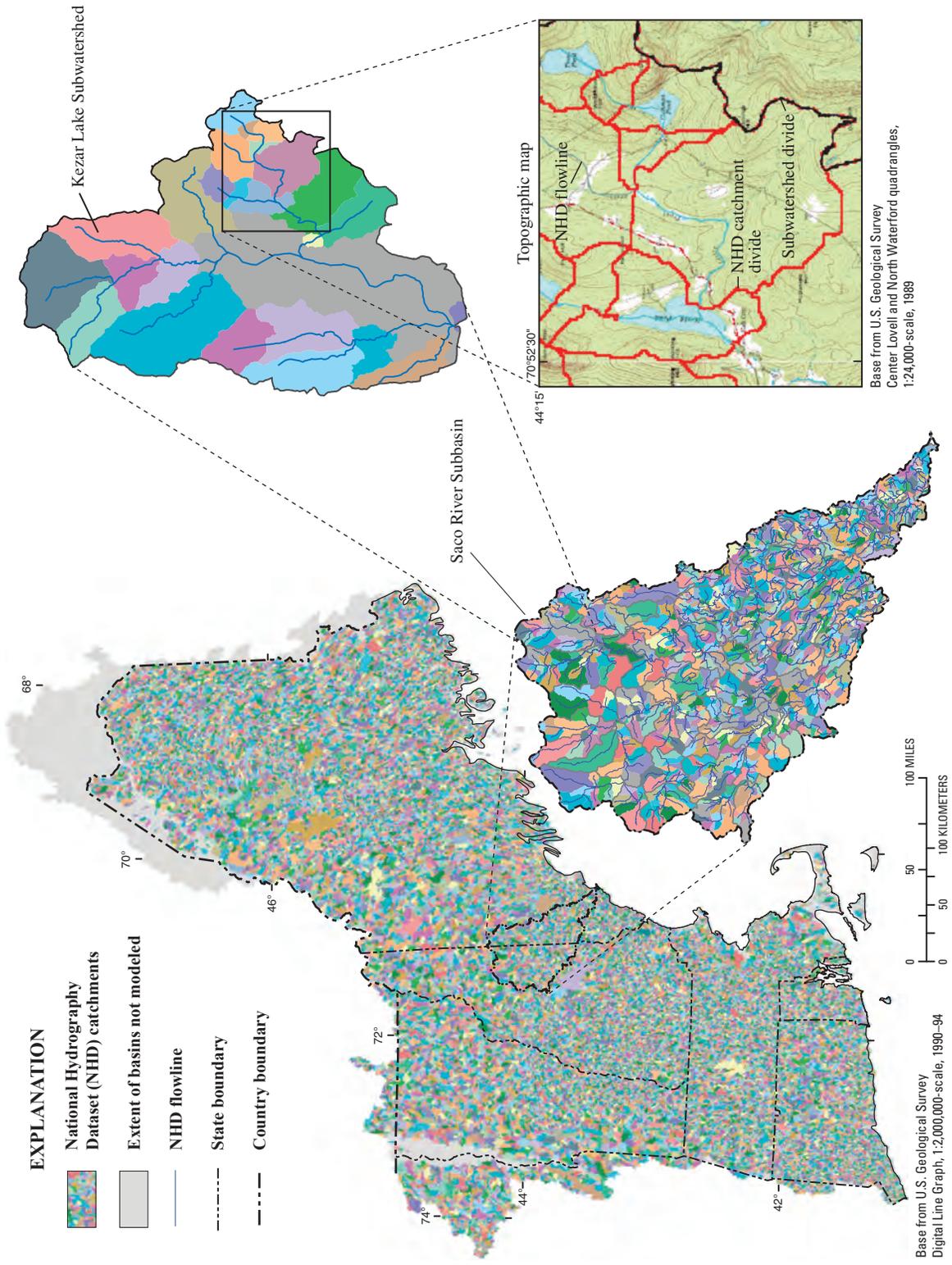
**Figure 6.** Schematic design showing multiple polygon features that define the catchment boundary for a National Hydrography Dataset (NHD) flowline.

### New England Method with Watershed Boundary Dataset

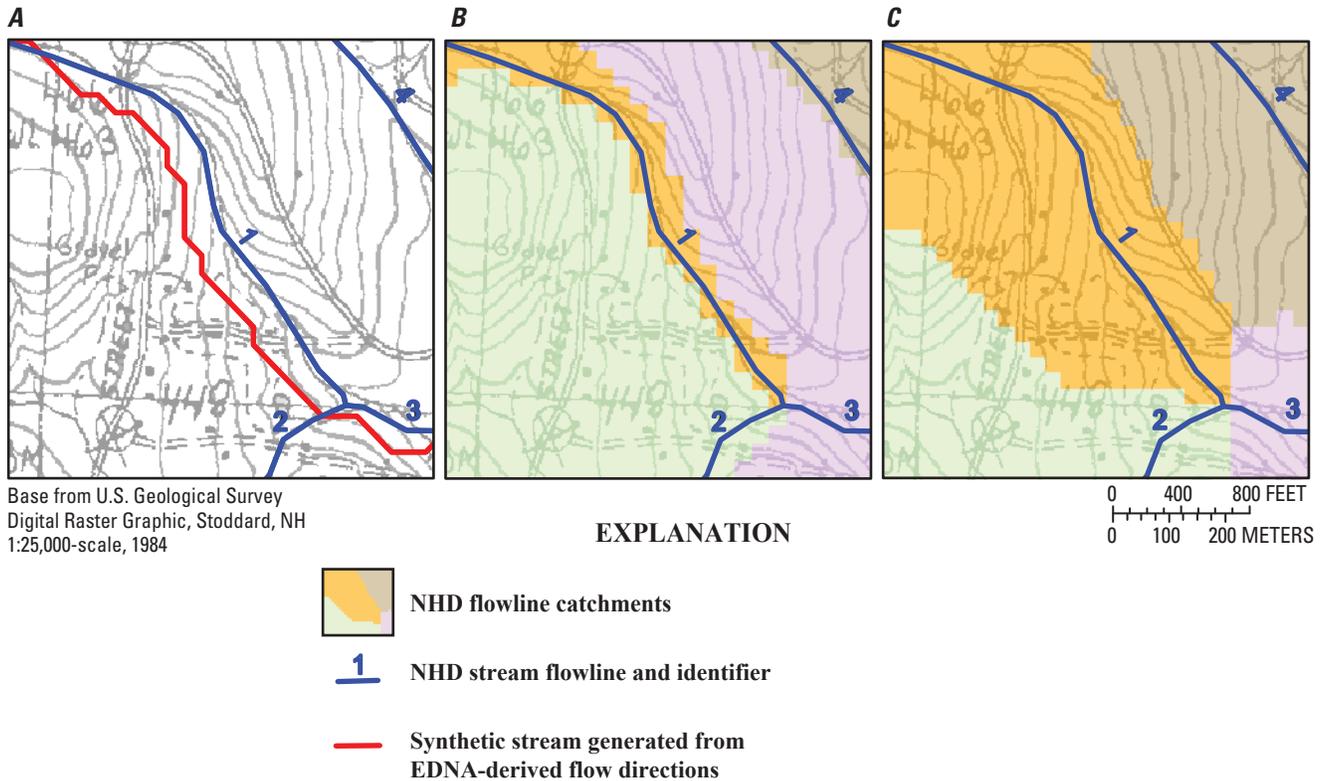
The New England Method with Watershed Boundary Dataset (NEM-with-WBD) was developed for the New England SPATIALLY REFERENCED REGRESSIONS ON WATERSHED attributes (SPARROW) water-quality model (Moore and others, 2004). This method was also employed without the WBD as described in the section “New England Method without Watershed Boundary Dataset.”

The SPARROW model requires a hydrologically connected representation of a stream network and associated

catchments through which stream nutrient loads are transported downstream, thus relating upstream sources to downstream monitored nutrient data. Previous SPARROW models in the United States were based on the *USEPA River Reach File Version 1.0* (RF1) (Dewald and others, 1985). The need for improved resolution in the New England models resulted in the selection of the medium-resolution NHD as the base hydrography for the model. For comparison, in New England, there are 2,462 RF1 catchments (mean size 29.2 m<sup>2</sup>) and 42,000 NHD catchments (mean size 1.7 m<sup>2</sup>). The spatial detail of the NHD in New England is shown in figure 7.



**Figure 7.** The spatial detail of the stream network and 42,000 catchments used for the New England SPATIALLY Referenced Regressions On Watershed attributes (SPARROW) model (modified from Moore and others, 2004).



**Figure 8.** (A) Horizontal displacement between the National Hydrography Dataset (NHD) flowlines and the Elevation Derivatives for National Applications (EDNA) synthetic streams, (B) catchment delineations using the Raster Seeding Method with the EDNA flow direction grid, and (C) catchment delineations using the New England Method with a new flow-direction grid created using the AGREE computer program with the National Elevation Dataset.

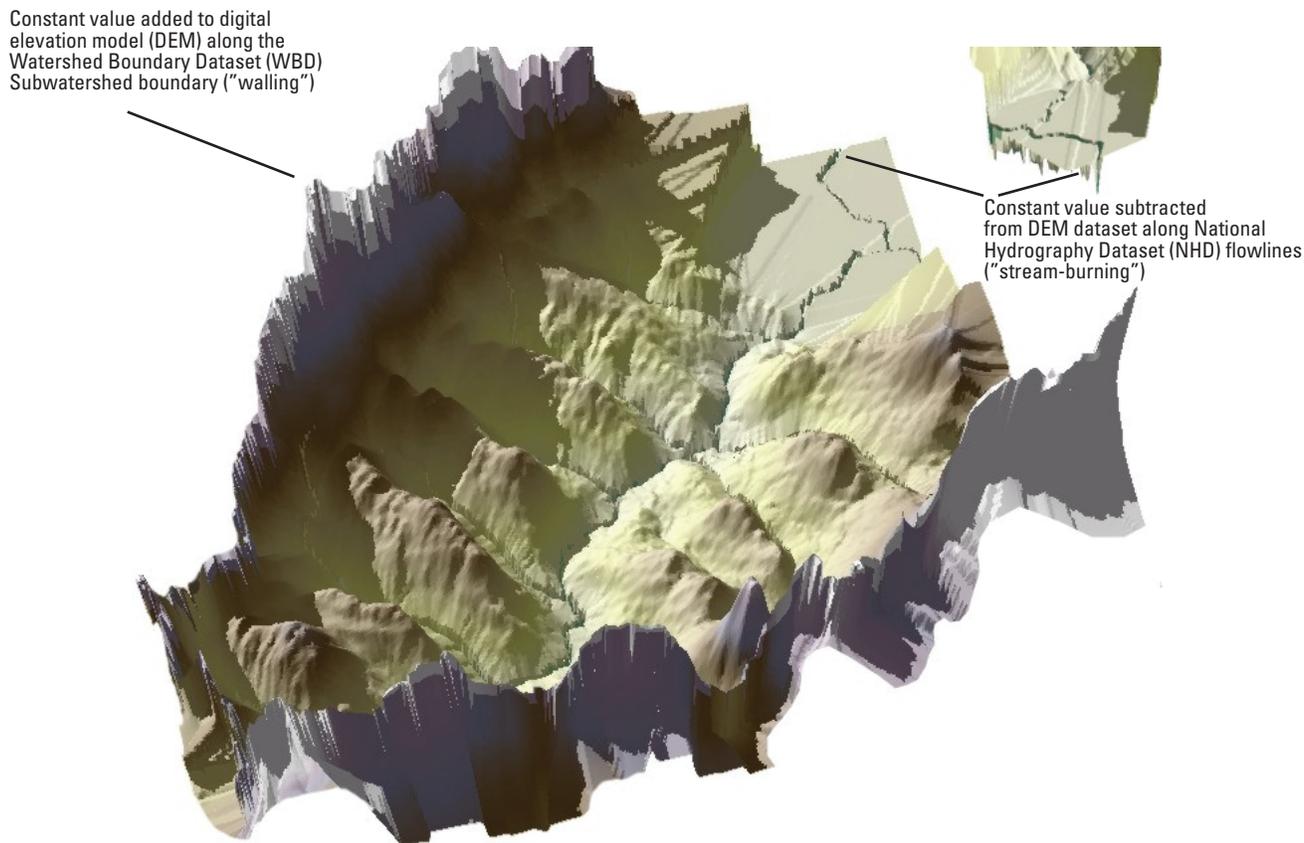
Similar to the Raster Seeding Method, catchments that are generated using the NEM-with-WBD were delineated using DEM-derived flow directions. The only difference between the methods is that the Raster Seeding Method used the EDNA flow direction grid and the NEM-with-WBD used a new flow direction grid. In the NEM-with-WBD, NED data were first modified to enforce drainage to the NHD flowlines, and a new flow direction grid was created as the source for catchment boundary delineations.

The NED modifications were considered necessary because the flow path defined by the EDNA flow direction grid can at times differ substantially from the mapped 1:100,000-scale NHD flowlines. This difference, referred to as the *DEM flow-path displacement*, is defined as the horizontal positional offset between the mapped NHD streams and the inherent drainage pattern of the DEM-based flow direction data. A common example of the DEM flow-path displacement between NHD flowlines and EDNA-derived streams is shown in figure 8A. In many cases, the NHD flowlines and EDNA-derived synthetic stream lines run parallel to each other at

some offset distance. If this offset distance is greater than one grid cell width, then part of the catchment generated using the Raster Seeding Method is not identified as being uphill from the NHD stream flowline. As a result, this part of the catchment is erroneously excluded from the delineated catchment boundary (fig. 8B).

To alleviate displacement problems between NHD flowlines and EDNA flow paths (fig. 8A), a new flow direction grid was developed by modifying the NED data in a process, often referred to as *stream burning* (Saunders, 2000), which integrates the NHD vector flowlines into the raster elevation data. A stream burning technique developed by Hellweger and Maidment (1997), called *DEM surface reconditioning*, was used with modifications. A series of computer algorithms implemented in the computer program called AGREE were used. AGREE is based on the *Arc Macro Language* (AML) developed by Hellweger and Maidment (1997). The AGREE computer program “burns a canyon” into the DEM by using a specified vertical distance, which is subtracted from the DEM elevation grid cells that coincide with rasterized NHD





**Figure 10.** Schematic design showing a three-dimensional perspective view of a modified digital elevation model with walling of the Watershed Boundary Dataset and stream burning, using the National Hydrography Dataset.

and create errors in catchment delineations. To minimize this headwater stream buffer intrusion for the study Subbasins, the TRIMSTREAMS AML program was developed to remove the most upstream 90 m of all headwater vector flowlines. The TRIMSTREAMS program was included in the New England Method processing steps before the AGREE processing.

To generate NHD catchments for the New England SPARROW model, the hydrologic conditioning of the NED data was taken a step further by creating a series of computer algorithms to force the DEM to be consistent with manually interpreted *drainage-area divides* of the WBD. The process of conditioning DEM data to WBD drainage areas is called *walling*. The New England Method process of walling was implemented using an AML program that increases the elevation values of DEM cells that coincide with the location of vector divides in the WBD. The vertical distance used to increase the drainage divide elevation cells was set arbitrarily high at 10,000 m. A three-dimensional representation of a hydrologically conditioned DEM showing the effects of walling and burning is presented in figure 10. This walling technique was used in the five study Subbasins where WBD boundaries were available.

The resulting hydrologically conditioned DEM, with stream burning of the NHD and walling of the WBD, was filled to remove spurious sinks on the DEM surface and used to derive the new flow direction grid. The final step of the NEM process is identical to that of the Raster Seeding Method where catchments were produced by rasterizing the NHD flowlines as input with the new flow direction grid to the ArcInfo WATERSHED command. The hydrologically conditioned DEM was discarded after this process, as measures of elevation, slope, and aspect would be meaningless for use in any analysis.

As with the Raster Seeding Method, flowlines less than 42 m in length may not have a catchment delineated by the New England Method, owing to the 30-m raster cell size. The resulting catchments from the NEM were converted to a vector region ArcInfo coverage data model.

In the application of the New England Method, the modified walling technique ensured that catchment delineations conformed to manually delineated Subwatershed boundaries of the WBD. NEM-with-WBD can be used for areas where the certified WBD is available. In general, digital drainage-area boundaries should not be used in the New England Method

walling process if the boundary delineations were made using elevation contours of a substantially different scale than the DEM data used for the New England Method. For this study, WBD boundary delineations (based upon 1:24,000-scale maps) were used with comparable one arc-second resolution DEM data (approximately 30-m cell size). In addition, the New England Method can be used if only part of the landscape has WBD boundaries. The use of the WBD also was helpful in locating inaccuracies in the NHD spatial and tabular relations. For the study Subbasins, no attempts were made to correct errors found in either the NHD or the WBD.

## New England Method without Watershed Boundary Dataset

At the time of this study, many States in the United States did not have a certified WBD. Some States possessed digital WBD data that were still in the development phase and, as a result, were deemed not acceptable for New England Method walling. Where the WBD is not available, the New England Method without Watershed Boundary Dataset (NEM-without-WBD) can be applied by simply omitting the algorithms used in the walling process. For completeness of the methods evaluations, the NEM-without-WBD was applied to all study Subbasins, including those with available WBD boundaries.

## Outlet Matching Method

The Center for Research in Water Resources at the University of Texas at Austin developed the Outlet Matching Method to relate a catchment to the point along a river network that most accurately approximates the catchment *outlet*. The algorithm is included in the *ArcHydro Tools* (Olivera and others, 2002) as Store Area Outlets. The Outlet Matching tool in ArcHydro Tools links existing catchments to the river network when the catchments are developed from a DEM that was not conditioned with the river network. For this study, the Outlet Matching tool was used to link existing EDNA catchments, derived from a synthetic stream network, with the flowline features that compose the NHD stream network. Using the existing DEM-derived EDNA catchments offered the potential of producing NHD catchments for less time and cost.

The Outlet Matching Method supports three different techniques for selecting the proper outlet for each EDNA catchment. Only one of these techniques may be specified for each execution of the Outlet Matching routine, thus requiring a GIS analyst to execute the routine several times, applying each technique in turn until an outlet has been identified for all EDNA catchments.

In the first technique, Junction Intersect (fig. 11), junction points were identified on the NHD network at the intersection of the EDNA catchment boundaries and the NHD flowlines. It is assumed that the outlet of the catchment will be at or near a junction of an NHD flowline and a EDNA catchment boundary. When there is more than one junction at the EDNA

catchment boundary, the appropriate outlet junction will have the greatest length of stream network upstream of the junction and inside the catchment boundary.

Figure 11 illustrates that the Junction Intersect technique selects the outlet that has the longest upstream network length inside the EDNA catchment boundary. Junction Intersect of the Outlet Matching technique typically determines outlets for more than 75 percent of the EDNA catchments.

In the second technique, Drainage Point Proximity (fig. 12), it is assumed that the outlet of an EDNA catchment is near the bottom of an NHD flowline. This technique is applied only to EDNA catchments that did not have an outlet assigned using the Junction Intersect technique. In the Drainage Point Proximity technique, a set of candidate outlets was defined as the downstream ends of the NHD flowlines. The drainage point for each EDNA catchment was defined as the lowest elevation point in each catchment. On the basis of an analyst-supplied tolerance, the Drainage Point Proximity technique finds the NHD outlet that is within the tolerance and nearest to the catchment drainage point and assigns that NHD outlet as the outlet of the catchment. The Drainage Point Proximity technique cannot be used with catchments that are not derived from elevation data.

The third technique, Next Downstream Area (fig. 13), was used when EDNA catchments were not assigned an outlet by either of the previous two techniques. In this technique, it is assumed that the topology of the EDNA catchments mimics the topology of the NHD stream network. This technique locates an adjacent catchment, which is downstream of the subject catchment that already has an assigned outlet, and then assigns that outlet to the subject catchment as well. The EDNA catchment topology is based on unique catchment identifiers, known as *Pfafstetter codes* (Verdin and Verdin, 1999), and their associated relations, shown as HydroID and NexDownID, respectively, in figure 13.

These three techniques were applied in sequence. The second technique was used only on those EDNA catchments that were not assigned an outlet using the first technique. The second technique was applied several times while increasing the analyst-specified tolerance and operating only on catchments that had not yet received an outlet. Likewise, the third technique was used only for those catchments that did not receive an outlet assignment using the first two techniques. After running all three techniques, all EDNA catchments were assigned an outlet.

The ultimate goal of this method was to have each EDNA catchment linked to a single NHD flowline feature rather than to an outlet point. A custom computer program was written to perform this final process. For a given EDNA catchment, the program relates the assigned outlet for the catchment to the flowline underlying the outlet. If an outlet was located at the intersection of multiple NHD flowline features, the flowline with the largest length within the catchment boundary was selected.

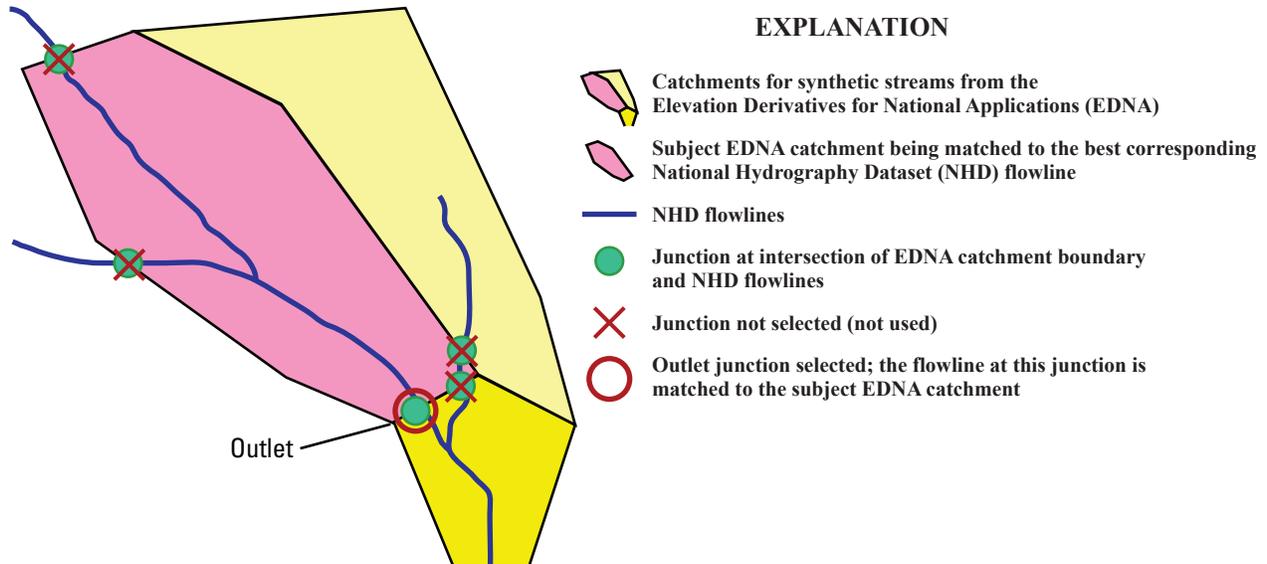


Figure 11. Schematic design showing the Junction Intersect technique of the Outlet Matching Method.

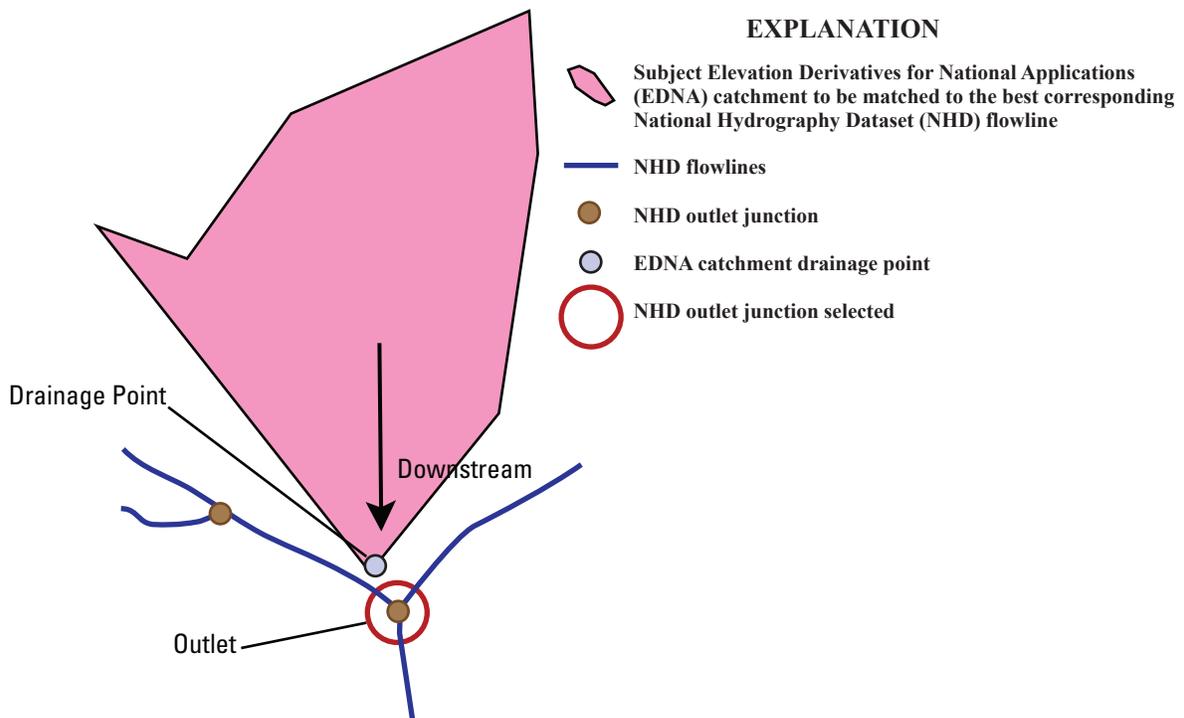
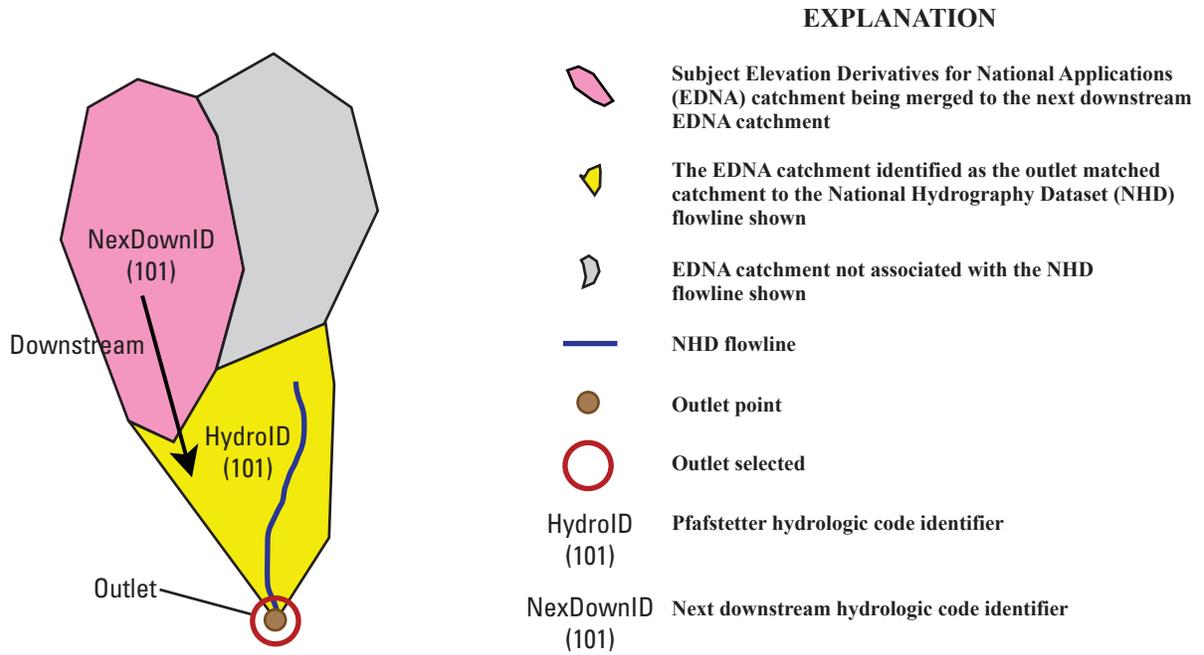


Figure 12. Schematic design showing the Drainage Point Proximity technique of the Outlet Matching Method.



**Figure 13.** Schematic design showing the Next Downstream Area technique of the Outlet Matching Method.

## Thiessen Polygon Method

The name “Thiessen” refers to work by Thiessen and Alter (1911) that defined areas (polygons) around rain gages whose boundaries were computed as half way between the nearest rain gages. The *Thiessen polygons* are constructed using a raster or grid-based approach to produce the boundary lines halfway between the nearest NHD flowlines. Because the Thiessen Method did not require the time-consuming processing of topographic information like the other methods, it offered the potential of producing NHD catchments for less time and cost.

To define the catchment boundary for each NHD flowline with the Thiessen Polygon Method, the vector NHD flowlines were first converted to a grid (raster) format having 30-m grid cells. Then a new grid of 30-m cells, referred to as the catchment grid, was created using the ArcInfo Grid EUCALLOCATION function (ESRI, 2003). In this step, each cell in the new catchment grid is assigned the grid code of the nearest NHD flowline grid cell measured using the straight-line euclidean distance. The collection of catchment grid cells with the grid code of a particular NHD flowline represents an approximation of the catchment area for that flowline. The catchment grid then is vectorized and assigned attributes

to create a vector polygon dataset approximating Thiessen polygons. The NHD flowline ComID is entered into the resulting polygon attribute table using the ArcInfo JOINITEM command (ESRI, 2003).

Thiessen polygon approximations of catchment boundaries are called “Thiessen catchments” (fig. 14). Because Thiessen catchments are based on purely geometric proximity, and terrain elevation information is not taken into consideration, the polygons represent a less accurate approximation of the true catchment boundaries. Conversely, the Thiessen Polygon Method has the advantage of being computationally efficient and does not require any dataset other than the NHD.

In addition to the flowlines in a study Subbasin, the NHD grid included the flowlines for all surrounding Subbasins. The surrounding flowlines were assigned grid codes of zero, and their associated areas were subsequently removed from the resulting datasets. This technique ensured that the resulting catchment boundary for a flowline on the perimeter of a study Subbasin was computed in the same manner as a catchment boundary for a flowline in the interior of the study Subbasin, that is, the boundary is halfway between the nearest NHD flowlines, including those in adjacent Subbasins.

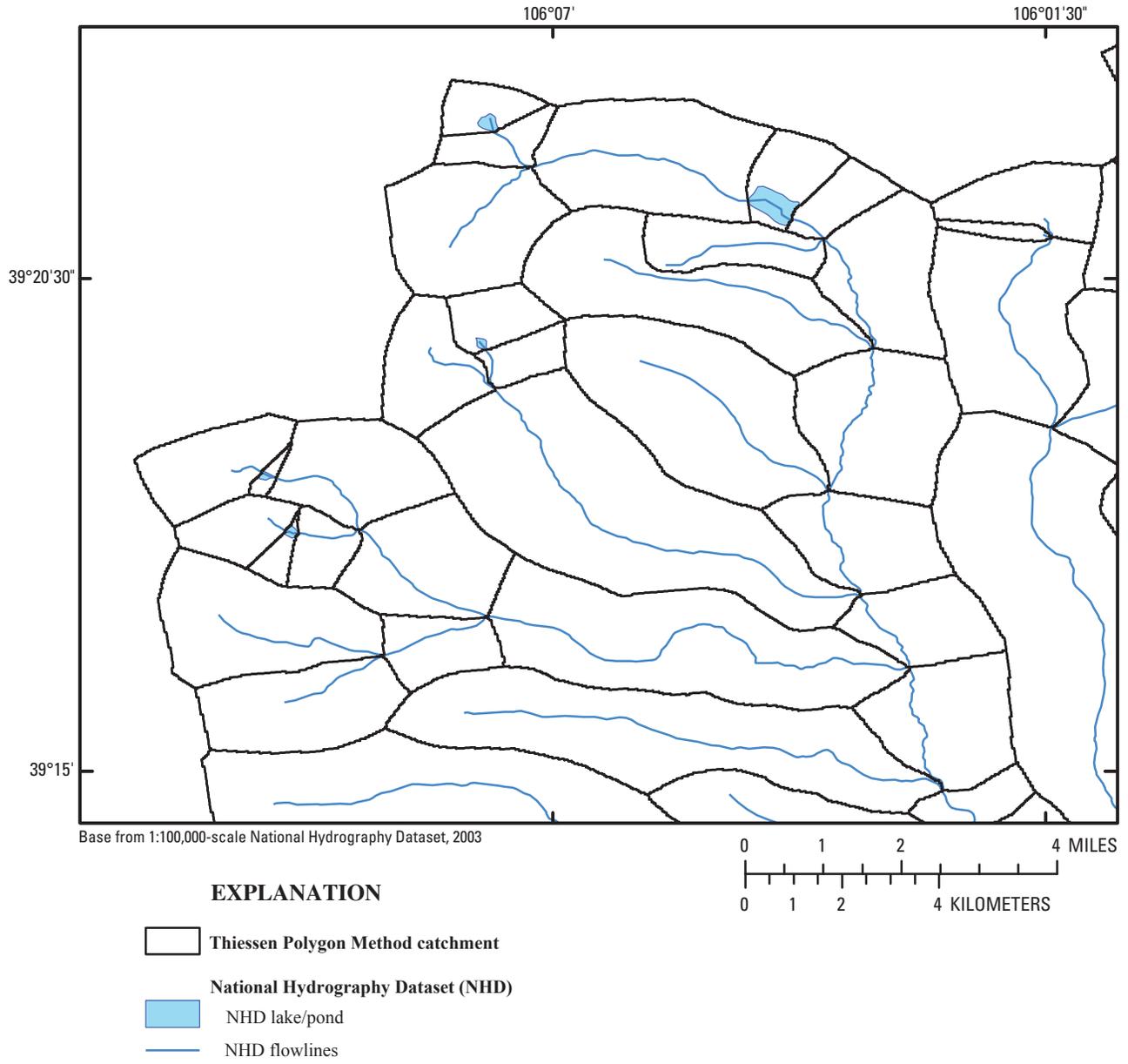


Figure 14. Thiessen Polygon Method catchments for National Hydrography Dataset flowlines.

## Testing the Catchment Delineation Methods

This section describes the areas where the catchment delineation methods were applied, the basis for comparing each method, and the evaluation metrics used for the comparisons.

### Study Subbasins

To test the catchment boundary delineation methods, six study areas within the continental United States were chosen (fig. 15). Each study area consists of a single headwater Subbasin, identified by its 8-digit HUC as defined within the WBD. These study Subbasins, listed in table 2, were chosen to represent the typical surface-water network morphologies found within the conterminous United States. The decision to use only headwater Subbasins was made to simplify the comparisons since it limits the upstream extent of the study Subbasins.

### Basis for Comparison

Evaluation of the accuracy of the catchment boundaries produced by the five delineation methods is based on conformance to real-world topographic relief and known drainage areas (presumed to be definitive). Catchment boundary delineation methods using topographic information (the Raster Seeding Method, the New England Method

variants, and the Outlet Matching Method) are bound to the resolution of the nationally available one arc-second (30-m) NED. In the conterminous United States, nearly all data in the one arc-second NED, at the time of this study, were derived from the largest scale topographic data available nationally, which is the USGS 1:24,000-scale topographic maps. Also, at the time of the study, nearly all one-third arc-second NED were derived from the 1:24,000-scale maps. Therefore, until the time when the one-ninth arc-second NED is completed nationally, the accuracy of the delineation methods can only refer to their conformance to the 1:24,000-scale topographic map program.

### Datasets Obtained for Verification in the Comparison Analysis

A variety of data was used to evaluate the accuracy of the catchment boundary delineation methods. These data are listed below.

- Published drainage-area measurements for USGS streamgages. These published measurements were used for comparisons with the drainage-area measurements computed during the study for NHD flowlines associated with streamgages.
- Drainage-area boundaries for USGS streamgages and draft WBD boundaries. These GIS datasets enable spatial drainage-area comparisons and analysis techniques.

**Table 2.** Subbasin name, Hydrologic Unit Code, State, and description for six study Subbasins in the United States.

Subbasin name	Hydrologic Unit Code (HUC)	State(s)	Description
Great Egg Harbor	02040302	New Jersey	A headwater Subbasin along the Atlantic Coast in New Jersey, just north of Delaware Bay. This area was selected as representative of low-relief areas along the east coast.
Seneca River	03060101	North Carolina, South Carolina	Headwaters of the Savannah River. This area was selected because of the large impoundment, Hartwell Lake, found at the outlet of the Subbasin.
Pine River	07010105	Minnesota	The headwater of the Mississippi River in Minnesota. This Subbasin contains a complex network of lakes and wetlands.
South Platte Headwaters	10190001	Colorado	This Subbasin contains braided streams and divergences that flow into the ground.
Pine Valley	16020302	Utah	Part of the Great Basin in Utah. This Subbasin is a closed drainage system where many isolated stream networks drain into the ground near the center of the Subbasin in a salt flat called the Pine Valley Hardpan.
Willapa Bay	17100106	Washington	This Subbasin is typical of the rugged terrain along the west coast.

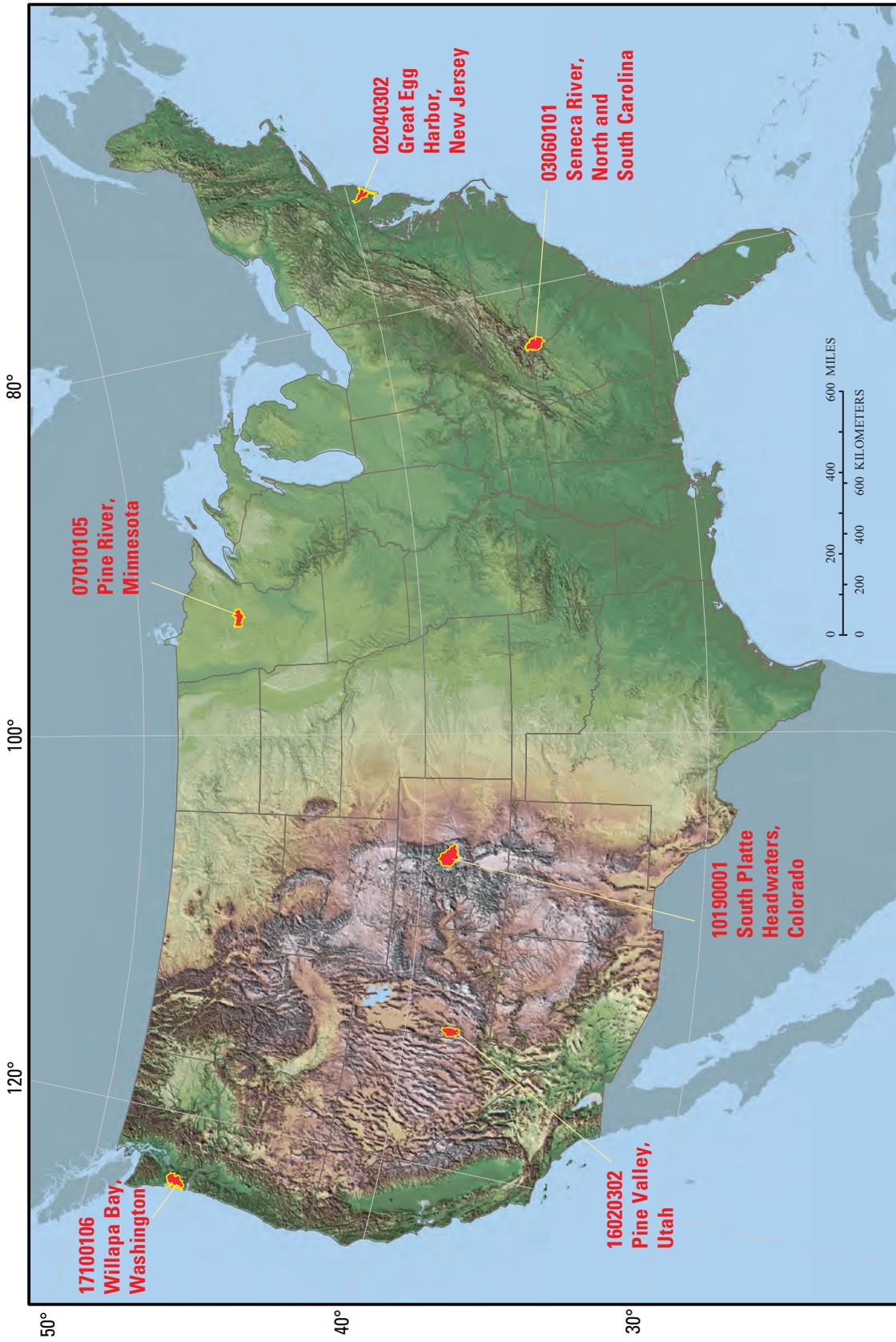


Figure 15. The location of the six study Subbasins in the continental United States.

- Digitally scanned images of the USGS 1:24,000-scale topographic maps called Digital Raster Graphics (DRGs) (U.S. Geological Survey, 2001a). DRGs were used to evaluate the delineations only for site-specific investigations where results indicated that detailed visual review was warranted.
- Aerial photographs called *Digital Orthophoto Quadrangles* (DOQ) (U.S. Geological Survey, 2001b) were used along with the interpretation of DRGs, especially for areas of flat relief with limited contour detail. DRGs and DOQs were used where warranted to evaluate differences between those areas where the largest errors occurred.

## Streamgage Data

The national set of currently operating and discontinued USGS streamgages provides the most readily available information on reported drainage-area measurements for evaluation purposes. Streamgage information is stored in the USGS's *National Water Information System* (NWIS) database and can be accessed through the NWIS Web site at <http://waterdata.usgs.gov/nwis>. Streamgage information in NWIS includes latitude and longitude coordinates, along with the reported drainage-area measurement for each streamgage. The streamgage data were also valuable for purposes of comparing methods because they provided a wide range of streamgage drainage-area sizes (1.1 to 1,084 mi<sup>2</sup>) across the study Subbasins.

The drainage-area measurements for the streamgages typically were manually computed by trained hydrologists using paper topographic maps and a mechanical planimeter. Drainage area measurements for some of the historical streamgages may be based on computations from USGS 15-minute 1:62,500-scale maps or other topographic maps. Topography and other features on the 15-minute map series are generalized and can produce different drainage-area measurements than those produced using the larger scale 1:24,000-scale series maps. Drainage-area measurements computed manually using the 1:24,000-scale series maps are assumed to be the most accurate. No documentation exists within NWIS that describes the source map scale used to determine the drainage-area measurement for streamgages.

A point dataset of the streamgages was created using the latitude and longitude coordinates from the NWIS database. Each streamgage point location was visually displayed along with the NHD flowlines and, through manual verification by USGS state offices, was linked to the appropriate location along the appropriate NHD flowline. Streamgages that were successfully linked with NHD flowlines were selected for use in the verification process. Drainage-area boundaries for these streamgages, in GIS polygon format, were also obtained from USGS Water Science Centers. The number of NWIS streamgages obtained for each study Subbasin with drainage-area measurements and boundaries that were selected for verification is shown in table 3. NWIS streamgage drainage-area measurements were not available for the Pine Valley Subbasin, and streamgage drainage-area boundaries were not available for the South Platte Headwaters and the Pine Valley Subbasins.

**Table 3.** Description of National Water Information System streamgages, boundaries, and associated drainage-area measurements used for verification datasets.

[NWIS, National Water Information System; GIS, geographic information systems; mi<sup>2</sup>, square miles; n/a, not available]

Subbasin Hydrologic Unit Code and name	Number of selected NWIS streamgages with drainage-area measurements	Number of selected streamgages with GIS drainage-area boundaries	Minimum drainage area (mi <sup>2</sup> )	Maximum drainage area (mi <sup>2</sup> )	Mean drainage area (mi <sup>2</sup> )
02040302 Great Egg Harbor	2	2	30.77	57.10	43.90
03060101 Seneca River	4	4	46.76	106.00	72.38
07010105 Pine River	1	2	1.03	269.54	177.19
10190001 South Platte Headwaters	15	0	11.80	1,084.00	286.41
16020302 Pine Valley	0	0	n/a	n/a	n/a
17100106 Willapa Bay	12	12	3.97	219.00	49.07

The minimum, maximum, and mean drainage-area measurements are included in table 3 to show the range and variation in streamgage drainage areas used in the study. The values shown in table 3 were computed using both published streamgage drainage-area measurements and drainage-areas measurements calculated from streamgage drainage-area boundaries where available.

## Boundaries from the Watershed Boundary Dataset

A number of USGS Water Science Centers provided draft WBD Subwatershed boundaries for use as verification datasets for this study. Many of these Subwatershed boundaries have been subsequently incorporated into the WBD. At the time of the study (2003), only limited coverage of the WBD was available, so the study team used draft WBD boundaries for the six study Subbasins. These draft WBD boundaries were digitized using manual methods from 1:24,000-scale sources and were validated by state water managers and staff familiar with the hydrology of the areas. The draft WBD boundaries were also validated through visual spot checking against contours depicted on DRGs. Where both datasets were available, the streamgage drainage-area boundaries were observed to be in good spatial agreement with corresponding draft WBD boundaries. For the purposes of this report, the draft WBD boundaries are simply referred to as WBD.

Four of the six study Subbasins were delineated to the WBD 12-digit Hydrologic Unit Subwatershed level, which generally ranges in size from 20 to 60 mi<sup>2</sup>. The Great Egg

Harbor Subbasin was delineated to the 14-digit hydrologic unit level, which is typically less than 20 mi<sup>2</sup>. Currently, the WBD does not incorporate boundaries for the 14-digit hydrologic unit level. The Willapa Bay Subbasin had no WBD boundaries available. A summary of the minimum, maximum, and mean WBD area measurements for each study Subbasin is presented in table 4.

## Evaluation Metrics

Evaluation metrics addressed two distinct characteristics of the catchments: physical-boundary accuracy and drainage-area measurement. Shape analysis techniques were used to assess the accuracy of the physical boundary delineations, and area-size analysis techniques were used to compare the drainage-area measurements. The evaluation metrics used in this study were the Coefficient of Areal Correspondence (CAC), polygon-area percent error, cumulative drainage-area analysis, and convergence analysis. These metrics and associated analyses helped identify errors in ingredient data and shortcomings in the methods themselves. Collectively, they enabled reasonable comparisons of the accuracy of the NHD catchment delineations resulting from each method.

The evaluation metrics were determined for all methods for all study Subbasins except for: (1) metrics requiring WBD for the Willapa Subbasin, where WBD was absent; and (2) the Outlet Matching Method, which was only determined for the Great Egg Harbor Subbasin. Outlet Matching Method metrics for other Subbasins are not presented because of poor performance in the Great Egg Harbor Subbasin.

**Table 4.** Summary of Watershed Boundary Dataset Hydrologic Units used for comparative analysis of National Hydrography Dataset catchment boundary delineation methods.

[WBD, Watershed Boundary Dataset; mi<sup>2</sup>, square miles; n/a, not available]

Study Subbasin Hydrologic Unit Code and name	Category of WBD Hydrologic Units	Number of WBD Hydrologic Units	Minimum area (mi <sup>2</sup> )	Maximum area (mi <sup>2</sup> )	Mean area (mi <sup>2</sup> )
02040302 Great Egg Harbor	14 digit	46	4.37	51.34	10.80
03060101 Seneca River	12 digit	33	11.63	57.27	31.15
07010105 Pine River	12 digit	28	6.28	53.98	27.85
10190001 South Platte Headwaters	12 digit	55	0.65	81.80	29.17
16020302 Pine Valley	12 digit	9	10.36	50.48	28.73
17100106 Willapa Bay	n/a	n/a	n/a	n/a	n/a

## Coefficient of Areal Correspondence

The CAC is a metric used to quantify the overlap of two areas (Taylor, 1977) and is computed by dividing the intersection of two areas by the union of the two areas (fig. 16). For catchment boundary delineation comparisons, the CAC conveys how well a boundary, based on one of the five methods, compares with an independent boundary, such as one for a WBD Subwatershed or a streamgage drainage area. The CAC quantifies the spatial correspondence between two boundaries where a CAC value of 100 indicates a perfect spatial correspondence. For the catchment delineation method comparisons, these CAC values can be thought of as a score, where higher scores indicate better performance, and with 100 being the highest score.

In figure 16, two squares measuring 30 units per side—one with a solid black border and one with a dashed red border—intersect. The area of the intersection between the squares is shown in orange and measures 600 square units (20 units wide by 30 units tall). The area of the union between the squares includes all colored areas and measures 1,200 square units (40 units wide by 30 units tall). Dividing the intersection area by the union area and multiplying the result by 100 produces a CAC score of 50 indicating that half of the squares overlap. A CAC score of 100 would indicate that the two squares overlap completely since the intersection and the union areas would be equal. Both the WBD boundaries and the streamgage drainage-area boundaries were used as GIS verification datasets for computing CAC scores.

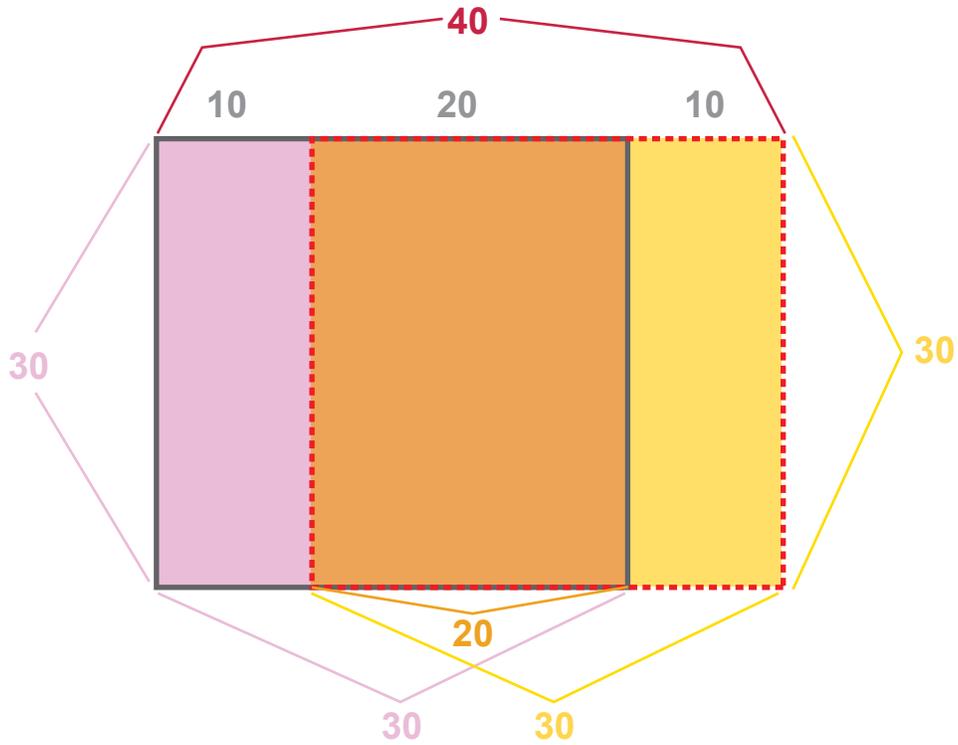
In the comparative analysis with the WBD data, CAC scores were computed using each individual WBD boundary polygon and the polygon for the same area delineated using each of the five catchment boundary delineation methods. To perform the comparison, first the NHD flowline that most closely corresponds to the outlet of each WBD boundary was identified. Using the NHD flow table, the NHD flowlines within the WBD boundary were then selected by navigating upstream until another WBD outlet flowline was encountered or the headwaters were reached. Next, the catchments boundaries for those flowlines were combined to create a single boundary corresponding to the actual WBD boundary,

and a CAC score was computed from the two. CAC scores were computed in this manner for all study Subbasins having WBD boundaries (the Willapa Subbasin did not have WBD boundaries and could not be tested in this manner). Figure 17 shows the WBD boundaries, the corresponding catchment-based boundaries produced using the NEM-without-WBD, and the CAC scores for the Pine River Subbasin.

CAC scores were also computed for the streamgage drainage-area boundaries using the same process just described for the WBD boundaries. CAC scores were computed in this manner for all study Subbasins having streamgage drainage-area boundaries (the South Platte Headwaters and the Pine Valley Subbasins did not have streamgage drainage-area boundaries and could not be tested in this manner).

Given that the NEM-with-WBD uses the WBD to hydrologically condition the elevation when producing NHD catchments, one would assume that the CAC scores calculated when comparing NEM-with-WBD based WBD boundaries with the actual WBD would most often be 100. However, this was often not the case. When comparing the NEM-with-WBD and WBD boundaries, CAC scores of less than 100 are usually attributable to the WBD boundary outlet not matching the location of the outlet (downstream end) of the corresponding NHD flowline. This is most often the result of differences between the 1:100,000-scale NHD flowlines and how the WBD boundaries were delineated using 1:24,000-scale hydrography on the topographic maps, where either the same confluences are in slightly different locations or the WBD outlet is located at a 1:24,000-scale stream confluence not included in the 1:100,000-scale NHD. An example of a WBD/NHD-flowline outlet mismatch that results in a NEM-with-WBD CAC score of 88.6 is shown in figure 18. When these WBD/NHD-flowline outlet mismatches occurred, all of the methods were affected.

CAC scores for all methods are also affected by NHD isolated networks. As shown in figure 17, isolated networks are not included when aggregating NHD catchments to correspond to the actual WBD boundary and, as a result, will lower the CAC scores.

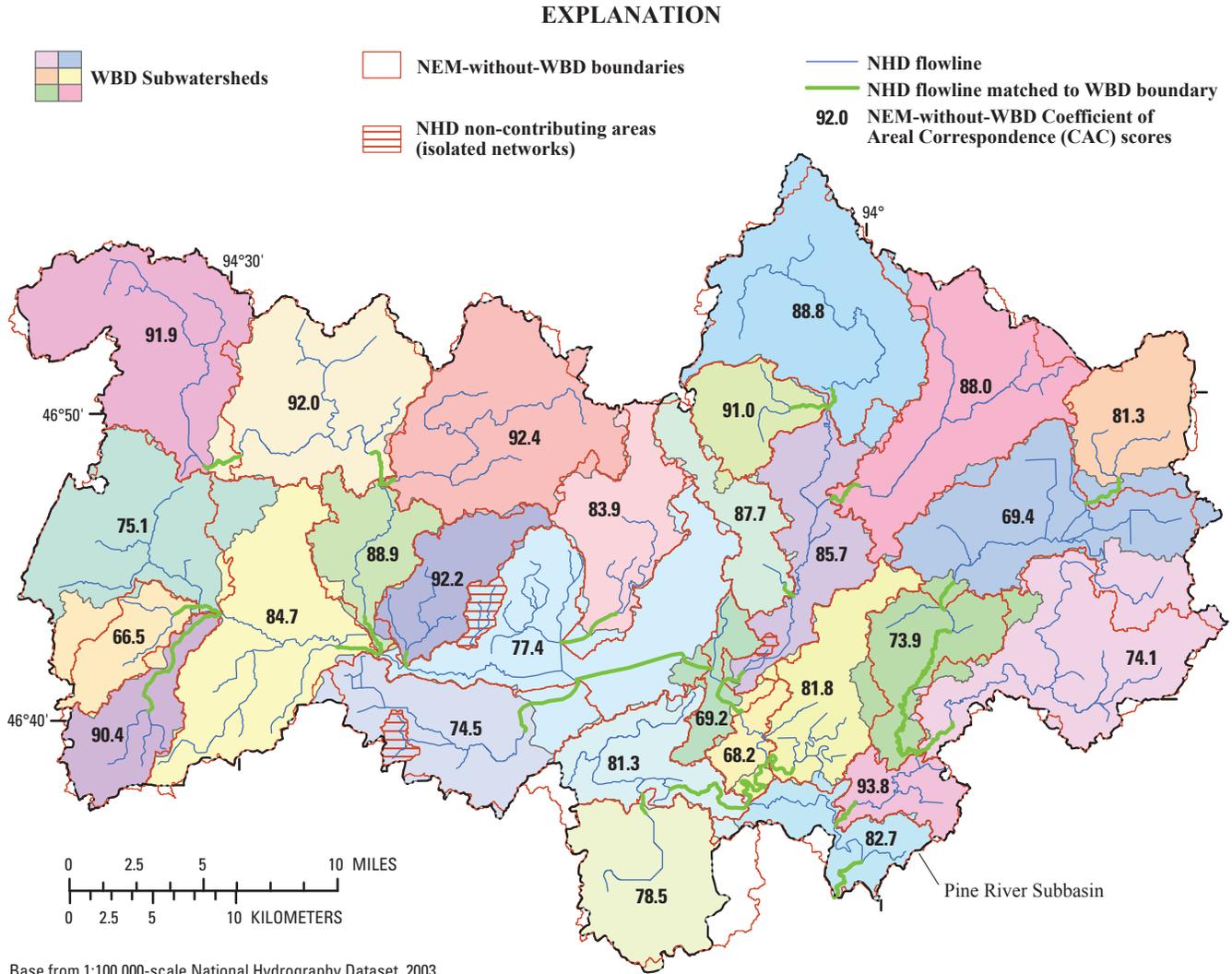


**EXPLANATION**

-  Area of 900
-  Area of 900
-  Intersection area of 600
-  Union area of 1,200

$$\frac{600}{1,200} \times 100 = 50 \text{ (CAC)}$$

**Figure 16.** Schematic design showing how Coefficient of Areal Correspondence (CAC) score is computed.



**Figure 17.** Boundaries from the Watershed Boundary Dataset (WBD) Subwatersheds and the corresponding boundaries from the New England Method without WBD (NEM-without-WBD), National Hydrography Dataset (NHD) flowlines including those most closely corresponding to the WBD outlets, and Coefficient of Areal Correspondence scores, Pine River Subbasin, MN. (The location of this Subbasin is shown in figure 15.)

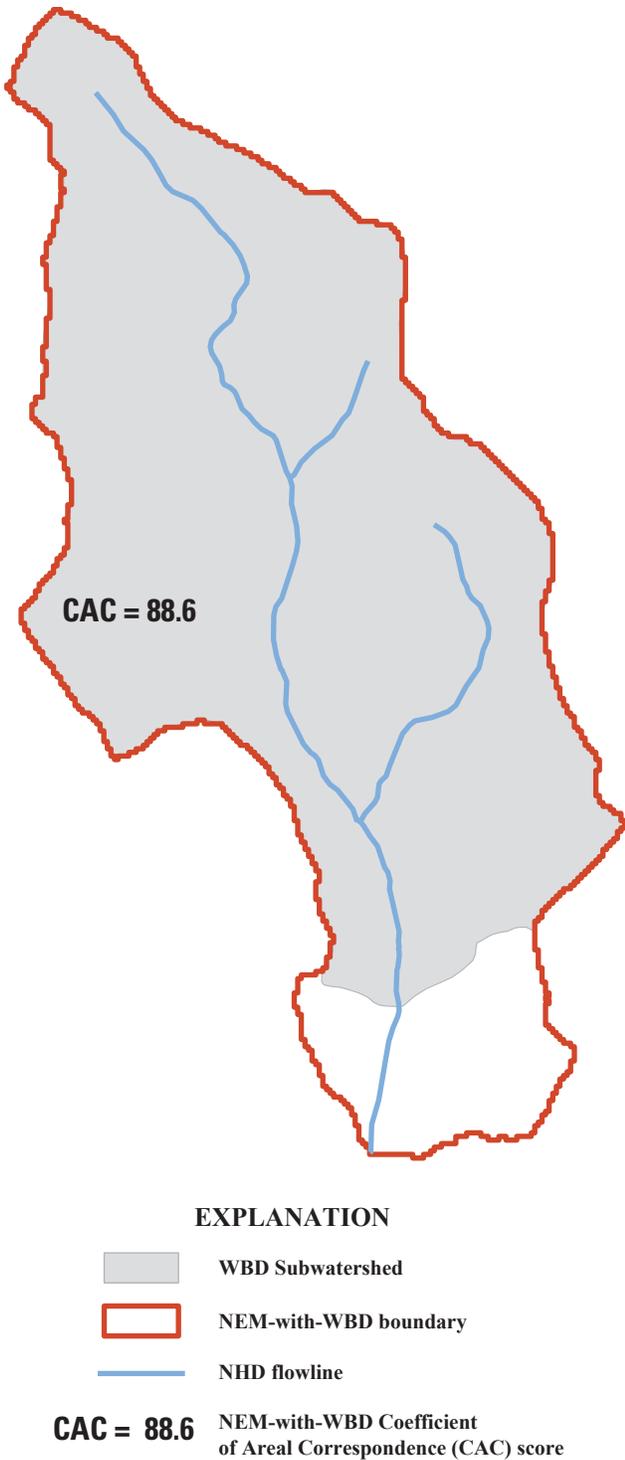
### Normalizing Coefficient of Areal Correspondence

The WBD/NHD-flowline outlet mismatches and NHD isolated networks just described can limit the effectiveness of the CAC as a descriptive tool to assess the methods. Normalizing the CAC scores by removing these known sources of error would bring more focus on the differences between the methods, rather than the errors associated with these known problems. The NEM-with-WBD was expected to be the most accurate method when compared with the WBD, since it benefited from hydrologically conditioning the DEM with both the NHD network and the WBD. As a result, the effect of WBD/NHD-flowline outlet mismatches and NHD isolated networks on CAC scores can be removed by using the NEM-with-WBD values to normalize the CAC scores of all other delineation methods, except the Outlet Matching Method (because of the prevalent outlet mismatching that occurred between the EDNA catchments and the NHD flowlines). The normalization formula shown below and in figure 19 is applied to adjust the CAC scores.

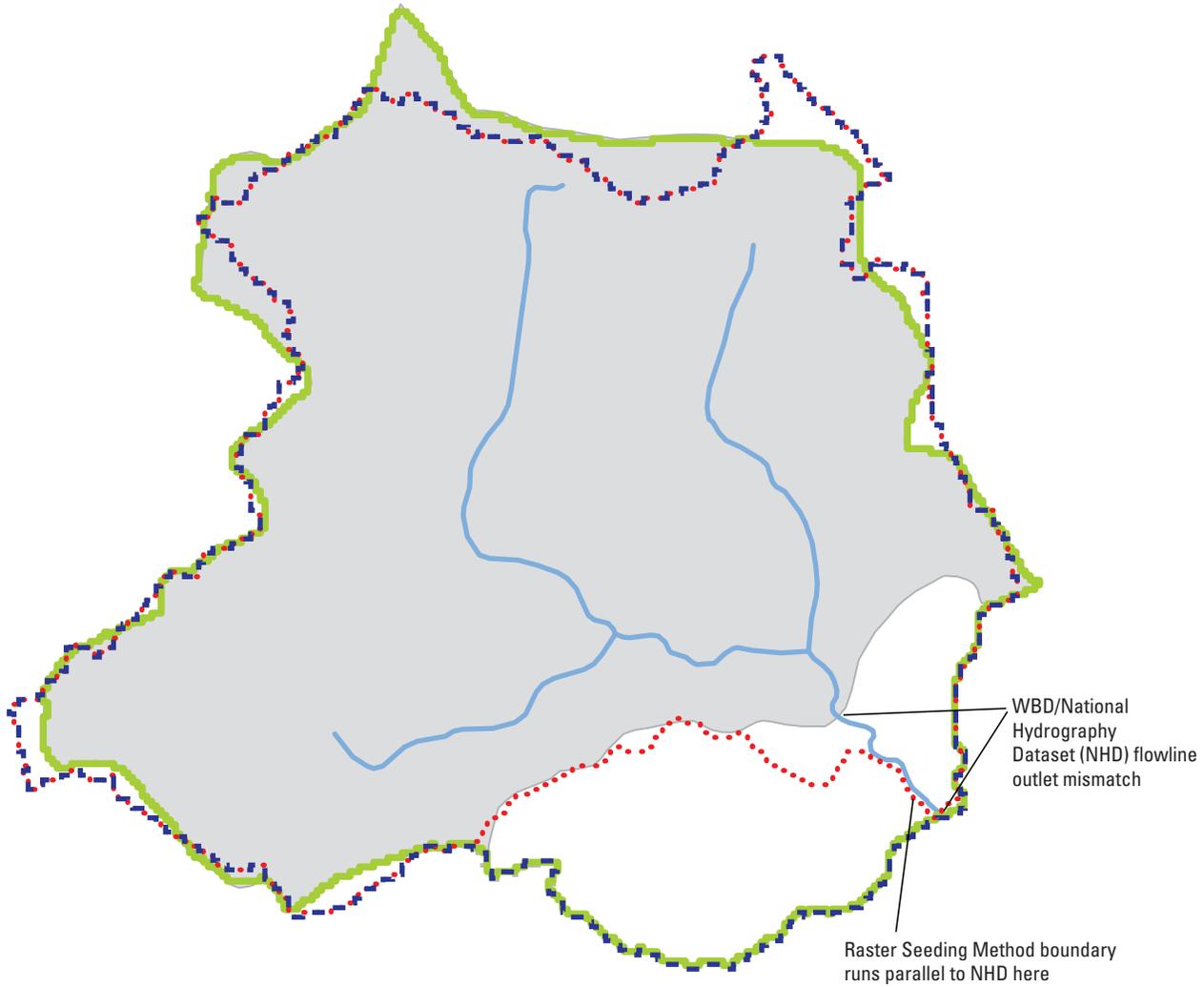
$$\text{Normalized CAC} = (\text{Delineation Method CAC} / \text{NEM-with-WBD CAC}) * 100 \quad (1)$$

The WBD boundary shown in figure 19 has a WBD/NHD-flowline outlet mismatch. The NEM-with-WBD catchment-based boundary follows the WBD boundary, except at the mismatched flowline outlet, resulting in a CAC score of 83.2 for the NEM-with-WBD boundary. CAC scores for each catchment delineation method are normalized such that NEM-with-WBD CAC score reflects a 100 percent spatial correspondence. The graph in figure 20 shows how the CAC scores for the NEM-with-WBD were adjusted by normalization in the Great Egg Harbor Subbasin. Each WBD unit is identified on the x-axis by the ComID of the NHD flowline that corresponds to the WBD outlet.

Normalizing CAC scores in this manner highlighted a potential issue when analyzing the Raster Seeding Method as shown in figure 19. The catchment-based boundaries from the Raster Seeding Method and NEM-without-WBD are identical, except for the area associated with the NHD-flowline outlet. At the NHD-flowline outlet, the two NEMs share the same boundary. For the Raster Seeding Method, the boundary delineation associated with the flowline outlet is smaller than the areas delineated using the NEM methods because of the affect of DEM flow-path displacement. In some cases, such as the example shown in figure 19, the normalized CAC score for the Raster Seeding Method (103.8) is erroneously higher than the NEM-with-WBD score (100). The highest a CAC score can be is 100, therefore, any CAC scores for the Raster Seeding Method greater than 100 were reviewed to determine what led to such a high score. These high scores happen in situations where there is both a WBD/NHD-flowline outlet mismatch and a Raster Seeding Method error associated with DEM flow-path displacement along the outlet flowline. This situation occurs because the Raster Seeding boundary runs parallel



**Figure 18.** Schematic design showing a mismatch between a Watershed Boundary Dataset (WBD) Subwatershed boundary and the corresponding New England Method with WBD (NEM-with-WBD) boundary caused by the WBD outlet not coinciding with the downstream end of the corresponding NHD flowline.



**EXPLANATION**

Watershed Boundary Dataset (WBD) Subwatershed

**Delineation Method, Coefficient of Areal Correspondence (CAC), and Normalized CAC**

- New England Method with WBD boundary (NEM-with-WBD)  
 CAC = 83.2    Normalized CAC = 100.0
- ⋯ Raster Seeding Method boundary  
 CAC = 86.4    Normalized CAC = 103.8
- - - New England Method without WBD boundary (NEM-without-WBD)  
 CAC = 76.6    Normalized CAC = 92.1

$$\text{Normalized CAC} = \frac{\text{Delineation Method CAC}}{\text{NEM-with-WBD CAC}} \times 100$$

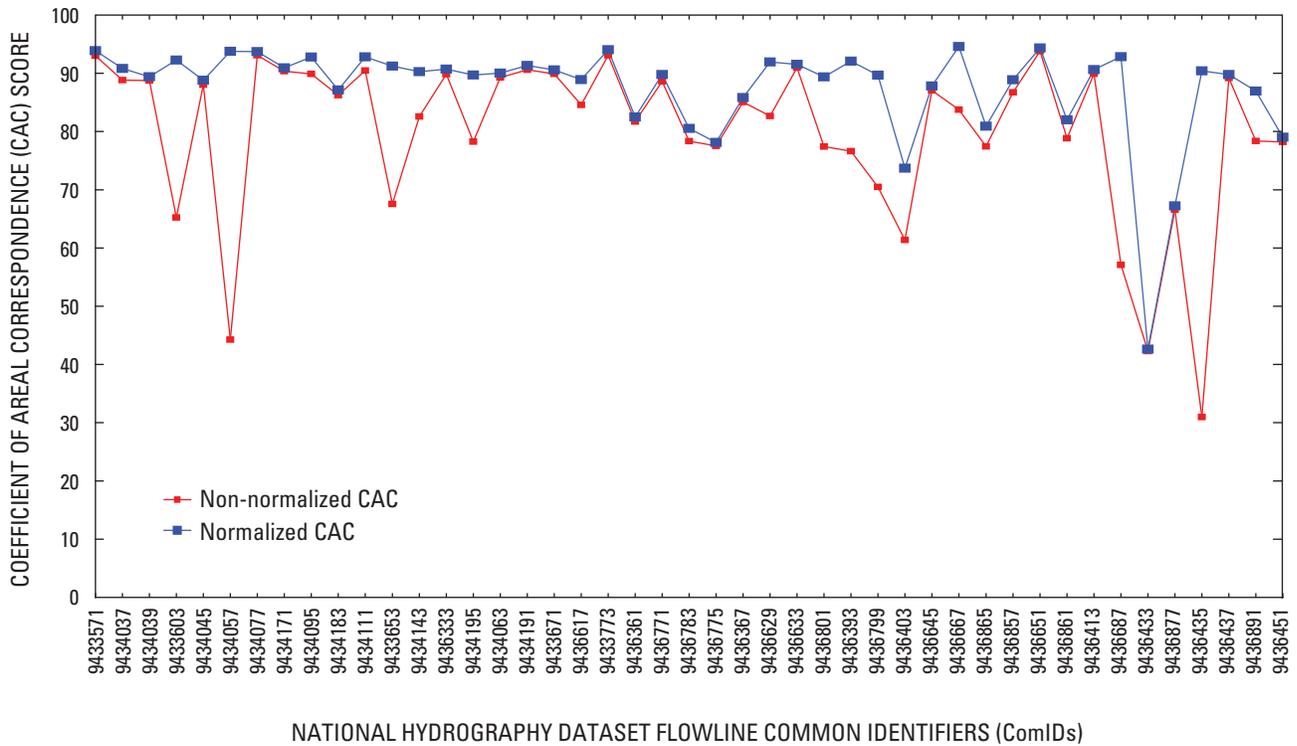
**Figure 19.** Schematic design showing a Watershed Boundary Dataset Subwatershed and corresponding boundaries from the New England Methods and the Raster Seeding Method, with Coefficient of Areal Correspondence (CAC) scores before and after normalization. The formula used to normalize the CAC scores is also shown.

to the outlet NHD flowline for some distance causing this Raster Seeding area outside the WBD boundary to be smaller (in error) than the NEM methods. Less obvious to find, but still prevalent, many Raster Seeding Method normalized CAC scores less than 100 can be artificially high where the combination of WBD/NHD-flowline outlet mismatch and DEM flow-path displacement occur at the outlet NHD flowline.

Outlet mismatch errors also occur for the streamgage boundaries because a streamgage can be located on any point along an NHD flowline. The CAC scores for the streamgage *drainage-area boundary* analysis were normalized in the same manner, since both the WBD and most streamgage drainage-area boundaries were manually reviewed against, or manually delineated from, 1:24,000-scale sources. As previously noted, the streamgage boundaries were also observed to be in good spatial agreement with corresponding WBD boundaries. The CAC scores for the NEM-with-WBD were used to normalize

the CAC scores for the other methods (except the Outlet Matching Method) in the comparison of the streamgage drainage-area boundaries. This CAC normalization was performed for all study Subbasins having WBD boundaries (the Willapa Bay Subbasin did not).

In the CAC evaluation, special attention was paid to those Subwatersheds that had a CAC (NEM-with-WBD) or normalized CAC (NEM-without-WBD and the Raster Seeding Method) score less than 80 to determine the cause of the specific error. Differences in delineations were examined closely using the information from the DRGs and DOQs to determine the correct delineation and to consider the effects on applications of the delineated boundaries. Low normalized CAC scores (less than 80) for the delineations from the Thiessen Polygon Method were not evaluated in detail because the Thiessen Polygon Method does not incorporate topographic information.



**Figure 20.** The non-normalized and normalized Coefficient of Areal Correspondence (CAC) scores for the New England Method without the Watershed Boundary Dataset (NEM-without-WBD) in relation to the corresponding boundaries from the Watershed Boundary Dataset labeled by National Hydrography Dataset flowline outlet ComID in the Great Egg Harbor Subbasin, NJ.

## Polygon-Area Percent Error

Polygon-area percent errors were computed between the area in the WBD Subwatershed polygons and the area of the corresponding catchment-based polygons produced by each method that was used in the CAC analysis. Polygon-area percent errors, as defined in the following equation, further quantify the magnitude of the delineation errors observed in the CAC analysis:

$$\text{Polygon-Area Percent Error} = \left[ \frac{|\text{Delineation Method Area} - \text{WBD Area}|}{\text{WBD Area}} \right] \times 100 \quad (2)$$

Polygon-area percent errors were computed for all study Subbasins having WBD boundaries (the Willapa Bay Subbasin did not have WBD boundaries to compute this metric). The polygon-area percent errors were also normalized using the NEM-with-WBD polygon-area percent errors to remove the effects of mismatches between the WBD boundary outlets and NHD-flowline outlets. This metric became even more relevant when applied to the cumulative drainage areas required for streamflow estimates.

## Cumulative Drainage-Area Analysis

The WBD cumulative drainage-area measurements were compared with the catchment-based cumulative drainage-area measurements from each delineation method. This analysis was applied to all study Subbasins having WBD boundaries (the Willapa Bay Subbasin did not have WBD boundaries).

To determine the WBD cumulative drainage-area measurements, the outlet flowlines identified for each WBD Subwatershed in the CAC analysis were used with the NHD flow table to determine all flowlines upstream of each WBD Subwatershed. Using the Subwatershed outlet flowlines and upstream flowlines, the upstream Subwatersheds for each Subwatershed can be determined and used to compute the WBD cumulative drainage-area measurement. For each of the catchment delineation methods, the flowlines for each WBD cumulative drainage area were linked to the catchments generated by each method to compute their respective cumulative drainage-area measurements.

For this analysis, X–Y graphs of WBD cumulative drainage-area measurements (x-axis) in relation to the cumulative drainage-area measurements for the delineation methods (y-axis) were produced. The closer the points are to the one-to-one line beginning at the origin, the closer the agreement. For example, the graph for the Great Egg Harbor Subbasin cumulative drainage-area measurements in figure 21 shows that the results of the two New England Methods and the Raster

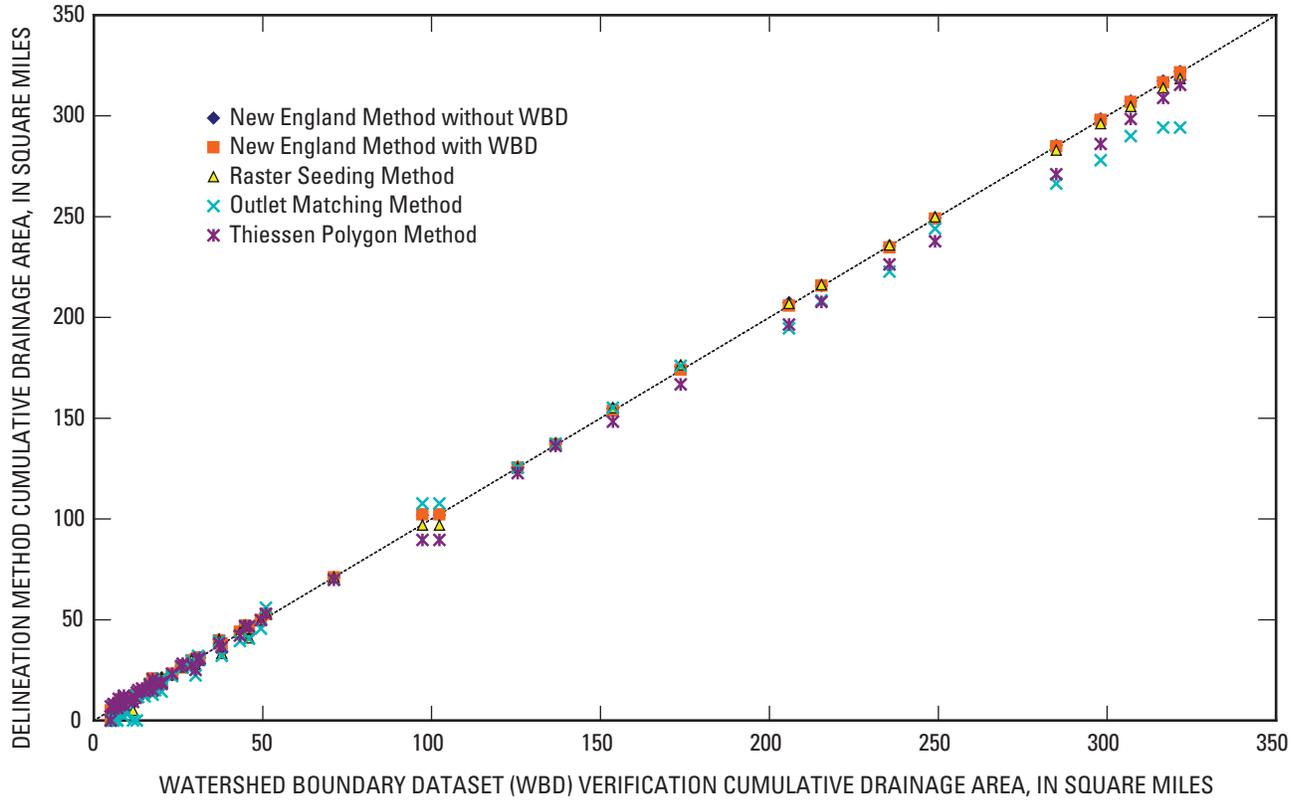
Seeding Method plot close to the one-to-one line, whereas the results of the Outlet Matching and Thiessen Polygon Methods plot fairly close to the one-to-one line, but tend to underestimate the cumulative drainage-area measurements.

The cumulative drainage-area analysis was also employed using drainage-area measurements from the streamgage verification data, which are cumulative by definition in the headwater Subbasins. The analysis required the accumulation of the catchment areas generated by each method for the NHD outlet flowlines on which the streamgages were located. For the Pine Valley Subbasin, streamgage data were not available. For the South Platte Headwaters Subbasin, streamgage drainage-area measurements were available from NWIS, but streamgage boundaries were not available. For the other study Subbasins, drainage-area data were available from both sources, and data from both were used in the analyses.

Better agreement between the streamgage verification drainage-area measurements and the drainage-area measurements from the catchment delineation methods is obtained when a streamgage is located at or near the corresponding NHD-flowline outlet. This agreement is true because, at the outlet of the flowline, the boundary of the streamgage drainage-area is a close match to the corresponding catchment-based boundary for each method. Greater differences in drainage areas occur the farther upstream a streamgage is located from a corresponding NHD-flowline outlet. This situation is similar to the WBD/NHD-flowline outlet mismatching previously described.

Drainage-area differences attributed to the streamgage location in relation to the corresponding NHD-flowline outlet occur only within the catchment where the streamgage is located. All upstream catchments included in the drainage area are unaffected by the location of the streamgage relative to the NHD-flowline outlet. For streamgages with large drainage areas, the percentage error contributed by disagreements between the streamgage location and the flowline outlet is expected to be less than 1 percent, because the outlet catchment area is relatively small compared with the upstream accumulated catchment-based drainage area. Conversely, for streamgages with small drainage areas, the errors contributed by disagreements between the streamgage location and the flowline outlet are expected to be larger because the outlet catchment area is larger relative to the upstream accumulated catchment-based drainage area.

Since the cumulative drainage-area measurements are the direct input into streamflow and velocity estimates, they were not normalized for either the cumulative drainage-area analysis or the convergence analysis; thus, all errors in the measurements would be present in these final metrics.



**Figure 21.** The Watershed Boundary Dataset cumulative drainage-area measurements (verification) in relation to the cumulative drainage-area measurements of the methods tested, Great Egg Harbor Subbasin, NJ.

### Convergence Analysis

Convergence analysis was used to better discern the results of the cumulative drainage-area analysis by determining a drainage-area percent error. Of particular interest was whether the errors for each method tended to decrease as the point of accumulation moved downstream. Determining a percentage error for comparison of each cumulative drainage area and graphing the results were used in the convergence analysis. Graphs of the data tend to show a decrease in drainage-area errors as cumulative drainage-area measurements increase. The errors tend to converge closer to the 0 percent error line as the drainage areas increase; hence, the term “convergence analysis” was developed for this study.

The drainage-area errors are expressed as percent error in the following equation that refines the polygon-area percent error by using areas that are cumulative drainage-areas.

$$\text{Percent error} = \frac{(|MV - V|)}{V} \times 100 \quad (3)$$

where

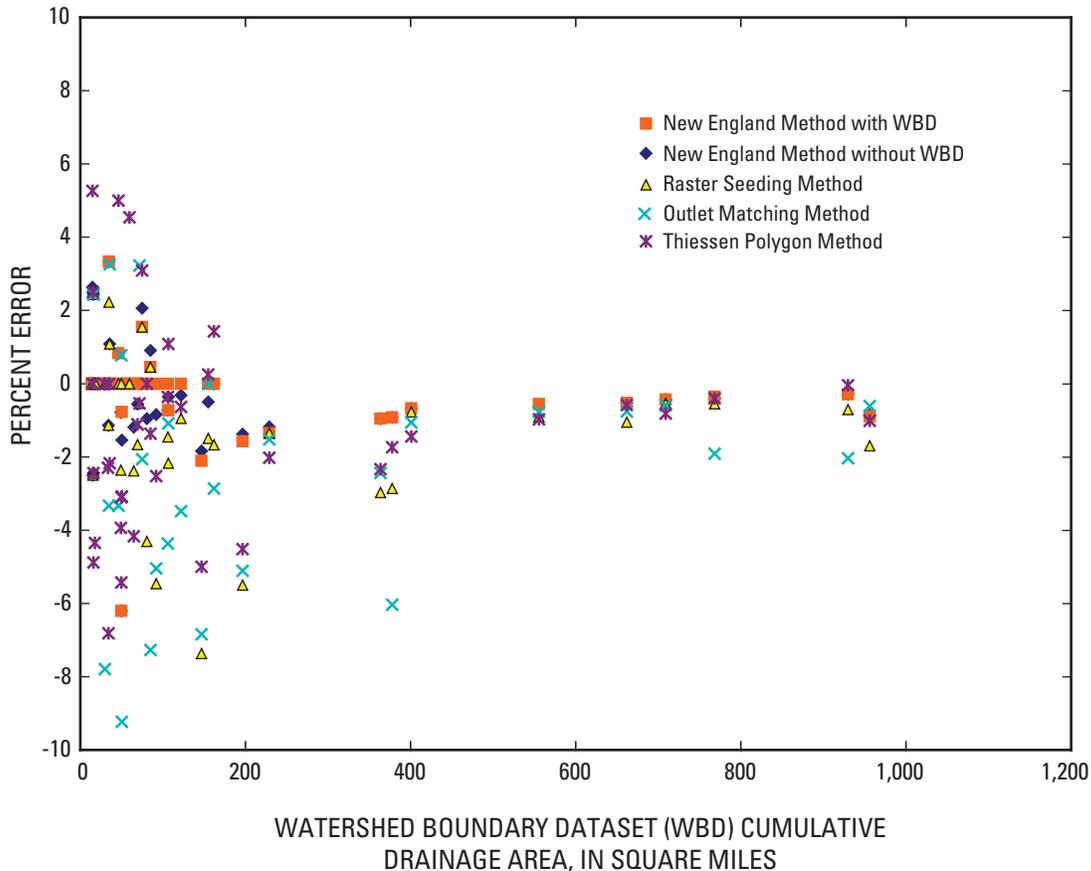
- $V$  = Cumulative WBD drainage-area measurement, and
- $MV$  = Cumulative delineation-area method value.

An example of the graph used in the convergence analysis for the Seneca River Subbasin is shown in figure 22. Graphs of convergence analysis are shown in the appendix for each study Subbasin that has WBD boundaries. (The Willapa Bay Subbasin did not have WBD boundaries.)

To assist further in the evaluation of the convergence analysis, cumulative drainage-area measurement errors for each study Subbasin are expressed as a mean error. The mean drainage-area measurement errors were computed by classifying the WBD cumulative drainage-area measurements into the following drainage-area size categories, which for convenience are based on the WBD Subbasin, Watershed, and Subwatershed size ranges:

1. Less than 20 mi<sup>2</sup>
2. 20 to 60 mi<sup>2</sup>
3. Greater than 60 to 390 mi<sup>2</sup>
4. Greater than 390 mi<sup>2</sup>

Classifying the results into size categories helped to identify the strengths or weaknesses of a particular method for producing drainage-area measurements.



**Figure 22.** Percent error of convergence for drainage-area measurements for all methods tested in relation to the Watershed Boundary Dataset verification cumulative drainage-area measurements, in the Seneca River Subbasin, SC.

## Evaluation of Catchment Delineation Methods

Results from all the analyses for all the study Subbasins indicate that the two New England Methods (with and without WBD) produced the most accurate catchment boundaries. Detailed results for each of the study Subbasins are presented in the appendix of this report. The NEM-with-WBD produced the most accurate catchment boundaries overall because this method incorporates the WBD in the process. For all study Subbasins, the mean normalized CAC scores from the WBD comparative analysis for each method are listed in table 5. Normalized CAC scores for the NEM-with-WBD indicate perfect spatial correspondence because this method was the basis for normalization. The overall average normalized CAC score using the NEM-without-WBD is 92.4 and using the Raster Seeding Method is 89.9. The Thiessen Polygon Method produced the least accurate spatial correspondence with an average normalized CAC score of 74.5.

The Outlet Matching method was found to have numerous limitations in its use as a catchment delineation

method. The fundamental differences between the NHD flowline network and the EDNA synthetic stream network used to generate EDNA catchments made it difficult to associate the flowlines and catchments using the Outlet Matching Method. NHD flowlines begin and end on the basis of cartographic collection standards that include minimum stream length, confluences, changes in feature types and characteristics. EDNA synthetic streams begin and end at confluences and are generated using a DEM-derived flow accumulation grid and a constant drainage-area threshold of 1.74 mi<sup>2</sup>, which result in catchments that often do not align favorably with NHD flowlines.

Horizontal displacement between the synthetic streams derived from the 30-m EDNA dataset and the NHD flowlines can also make the matching of the EDNA catchments to the NHD flowlines difficult. In areas of flat-relief, the horizontal displacement can be so great that the Outlet Matching Method process fails to identify a proper assignment of EDNA catchments to the NHD flowlines. In many cases, no match of the EDNA catchments to the flowlines could be made. For the Great Egg Harbor Subbasin, in 16 percent of the WBD comparisons, EDNA catchments could not be outlet-matched

**Table 5.** Mean normalized Coefficient of Areal Correspondence scores for the four methods used in the Watershed Boundary Dataset comparison analysis, by study Subbasin.

[CAC, Coefficient of Areal Correspondence; WBD, Watershed Boundary Dataset; NEM, New England Method]

Catchment delineation method	Mean normalized CACs from the WBD comparison analysis for study Subbasins					
	Great Egg Harbor 02040302	Seneca River Basin 03060101	Pine River Basin 07010105	South Platte Headwaters 10190001	Pine Valley 16020302	All study Subbasins (except Willapa Bay)
NEM-with-WBD	100.0	100.0	100.0	100.0	100.0	100.0
NEM-without-WBD	87.5	97.4	89.1	94.7	95.9	92.4
Raster Seeding Method	85.8	91.7	81.7	95.6	96.6	89.9
Thiessen Polygon Method	70.9	85.1	66.3	74.7	77.7	74.5

to the corresponding NHD flowlines assigned to the WBD Subwatersheds. Similarly, 11 percent of the EDNA catchments in the Pine River Subbasin, 7 percent in the South Platte Subbasin, and 3 percent in the Seneca Subbasin could not be outlet-matched to NHD flowlines. Appendix figure 1–6 shows a major error in the Outlet Matching Method within the Great Egg Harbor Subbasin, where an outlet-matched boundary for a streamgage erroneously composes of two separate polygons.

These significant difficulties encountered in achieving even an approximate coupling of the two datasets through database relations make the Outlet Matching Method problematic for producing catchments for NHD flowlines. Results for the Outlet Matching Method are provided only for the Great Egg Harbor Subbasin in the appendix of this report because it was representative of the poor performance of the Outlet Matching Method. Because of these identified problems, Outlet Matching Method analysis results are not described further in this report.

The Raster Seeding Method appears to have performed slightly more accurately than the NEM-without-WBD for the South Platte Headwaters and the Pine Valley Subbasins

(table 5), but further analysis of the data shows that the error for the Raster Seeding Method, resulting from DEM flow-path displacement at the outlet flowline and WBD/NHD-flowline outlet mismatching, artificially increased the CAC score. This is described in detail for the individual study Subbasins in the appendix.

The mean normalized polygon-area percent errors are summarized in table 6. Because the NEM-with-WBD was used to normalize the results for the other methods, the NEM-with-WBD shows 0 percent error for all Subbasins tested. The NEM-without-WBD was more accurate than the Raster Seeding and Thiessen Polygon Methods for all study Subbasins. The differences in normalized CAC scores between the NEM-without-WBD and the Raster Seeding Method shown in table 5 are further emphasized by the normalized polygon-area percent errors shown in table 6. In the South Platte Headwaters and Pine Valley Subbasins where the Raster Seeding Method appears to have been more accurate according to the normalized CAC score results (table 5), the normalized polygon-area percent errors for the Raster Seeding Method (table 6) are greater than that of NEM-without-WBD for the two Subbasins.

**Table 6.** Mean normalized polygon-area percent errors for catchment delineation methods used in the Watershed Boundary Dataset comparison analysis.

[All values are in percent. WBD, Watershed Boundary Dataset; NEM, New England Method]

Catchment delineation method	Mean normalized polygon-area percent errors for the WBD comparison analysis for study Subbasins					
	Great Egg Harbor 02040302	Seneca River Basin 03060101	Pine River Basin 07010105	South Platte Headwaters 10190001	Pine Valley 16020302	All study Subbasins (except Willapa Bay)
NEM-with-WBD	0.00	0.00	0.00	0.00	0.00	0.00
NEM-without-WBD	6.89	0.56	5.70	2.74	1.11	3.83
Raster Seeding Method	9.00	5.75	15.44	3.70	2.61	7.42
Thiessen Polygon Method	17.17	5.47	21.95	12.87	10.04	13.93

The Raster Seeding Method rated lower than the two New England Methods in the CAC and polygon-area percent error analyses because of the errors introduced by DEM flow-path displacement. DEM flow-path displacement errors were consistently observed in the Raster Seeding Method, causing catchment boundary delineations to erroneously parallel the NHD flowlines. As presented in the study Subbasin analyses in the appendix, the Raster Seeding Method errors were sometimes very large, especially in areas of flat relief or for artificial path flowlines within water bodies. The Pine River Subbasin has extreme examples of DEM flow-path displacement related errors within waterbodies and flat relief.

The Thiessen Polygon Method produced the least accurate catchment boundary delineations using the CAC analysis. This result was expected because the method does not take into account any topographic information. The Thiessen Polygon Method polygon-area percent error analysis confirmed these findings.

In the comparison of catchment-based delineations from the different methods with the streamgage drainage-area boundaries, the CAC scores produced are similar to those from the WBD comparison. Again, the New England Method without WBD was slightly more accurate than the Raster Seeding Method, and the Thiessen Polygon Method had the least accurate catchment boundaries.

Values obtained from the cumulative drainage-area analysis indicate, for each method, that drainage-area estimates increase in relative accuracy as the drainage-area size increases. The mean cumulative drainage-area errors for each method are listed in table 7. Those results are weighted averages, based on the number of observations from five of the six study Subbasins (except Willapa Bay, because this Subbasin did not have WBD data to compare with).

Drainage-area percent errors for the NEM-with-WBD are attributed to WBD/NHD-flowline outlet mismatching and NHD isolated networks. These errors were also present for the other methods. The second most accurate method, which

is based on the mean cumulative drainage-area percent error, was the NEM-without-WBD. The NEM-without-WBD was more accurate than the Raster Seeding and Thiessen Polygon Methods in all drainage-area size categories (table 7).

The Thiessen Polygon Method provided reasonable drainage-area estimates for the larger size drainage areas, but was progressively less accurate as the drainage-area size decreased. The mean cumulative drainage-area error for the Thiessen Polygon Method is 16.56 percent for drainage areas less than 20 mi<sup>2</sup>; for the Raster Seeding Method, 6.39 percent; for the NEM-without-WBD, 6.01 percent; and for the NEM-with-WBD, 1.88 percent.

Performance of the methods did not appear to vary across the landscape and hydrologic settings of the study Subbasins. For all study Subbasins, the walling of the WBD in the DEM conditioning process greatly enhanced the accuracy of the delineations. Without the use of walls from the WBD, the DEMs were generally less accurate, especially for areas of flat relief. In certain cases, a proper comparison with the verification data could not be made because of the limited contour detail of the DRGs in flat relief settings. Also, because the WBD boundaries were in draft stage, occasional delineations of the WBD were observed to provide less accurate boundaries compared with the NEM-without-WBD and the Raster Seeding Method.

The NEM-without-WBD and the Raster Seeding Method produced similar results, but the Raster Seeding Method showed significant problems in flat-relief settings, as observed in the Pine River Subbasin. Consistently across study Subbasins, differences between the Raster Seeding Method and the NEM-without-WBD were primarily due to the horizontal displacement between the DEM flow-path and the NHD flowlines associated with the Raster Seeding Method.

The relative differences in accuracy between the different catchment delineation methods decreases (they converge) as the drainage-area size increases (table 7). Figure 22 and appendix figures 1–3, 1–12, 1–20, and 1–26 support

**Table 7.** Mean cumulative drainage-area errors, by drainage-area size category, for catchment delineation methods for all study Subbasins, except Willapa Bay.

[mi<sup>2</sup>, square miles; NEM, New England Method; WBD, Watershed Boundary Dataset; All size categories were computed as a weighted average using the number of observations as a weights]

Method	Mean cumulative drainage area error, in percent				
	Drainage-area size category				
	Less than 20 mi <sup>2</sup>	20–60 mi <sup>2</sup>	Greater than 60–390 mi <sup>2</sup>	Greater than 390 mi <sup>2</sup>	All drainage areas
NEM-with-WBD	1.88	2.76	0.89	0.64	1.65
NEM-without-WBD	6.01	3.98	1.28	0.56	3.21
Raster Seeding Method	6.39	7.42	2.05	1.48	4.65
Thiessen Polygon Method	16.56	9.31	4.94	1.46	8.85

this observation as well. This convergence implies that the accuracy of the catchment delineation methods is most important in smaller drainage areas. Collectively, these results show that the NEM-with-WBD and NEM-without-WBD are the most accurate for smaller drainage areas. This finding is important because most of the flowlines in the United States have drainage areas less than 20 mi<sup>2</sup> (U.S. Environmental Protection Agency, 2006). The results of convergence analyses imply that the NEM methods will not improve the flow estimates for large drainage areas, but these methods will improve flow estimates for flowlines with smaller drainage areas.

## Summary

The need for catchment delineations originated from the common goal of the U.S. Environmental Protection Agency and U.S. Geological Survey (USGS) to estimate streamflow and velocity for all flowlines (stream segments) of the medium-resolution (1:100,000-scale) National Hydrography Dataset (NHD) in the conterminous United States and Hawaii. Accurate NHD catchment delineations are needed because streamflow and velocity estimates are based on drainage areas and characteristics of the landscape within the flowline catchment areas. This study was conducted to evaluate the accuracy of various methods for determining NHD catchments.

Five catchment delineation methods were evaluated, four of which used topographic information. These four were (1) the Raster Seeding Method; two variants of a method used in a USGS New England study, (2) one using drainage divides from the Watershed Boundary Dataset (WBD) and (3) the other without, termed the “New England Methods;” and (4) the Outlet Matching Method. For each of these topographic delineation methods, the source of elevation data was the 30-m resolution National Elevation Dataset (NED), since this was the highest resolution available for the conterminous United States and Hawaii. (5) The fifth method evaluated, the Thiessen Polygon Method, uses distance to the nearest NHD flowlines alone to determine a catchment boundary.

The New England Methods and the Raster Seeding Method applied the same GIS computer algorithm to determine catchments using the NHD flowlines and a digital elevation model (DEM)-derived flow direction grid. The input flow direction grid, however, differed between these methods in that the New England Methods modified the source DEM data, and the Raster Seeding Method did not.

The New England Methods modified the DEM data to ensure that drainage flow paths represented in the flow direction grid matched the drainage network of the NHD. For the New England Method that incorporated the WBD, the DEM was further modified to ensure that the WBD drainage divides also were represented in the DEM.

The Outlet Matching Method attempted to relate existing catchments from the Elevation Derivatives for National Applications (EDNA) database to the NHD flowlines. The EDNA catchments, which were based on unmodified DEM

flow paths, represented the areas draining to each EDNA synthetic stream.

The methods were applied to six study areas within the conterminous United States. These study areas, called “Subbasins,” are designated with 8-digit Hydrologic Unit Codes, as defined by the WBD. The six Subbasins chosen for this study represent different surface-water morphologies found within the United States. The study Subbasins and their locations are: Great Egg Harbor, New Jersey; Seneca River, North Carolina and South Carolina; Pine River, Minnesota; the South Platte Headwaters, Colorado; Pine Valley, Utah; and Willapa Bay, Washington.

The study used verification datasets, including USGS streamgages and the WBD, to compare the results of the methods. Evaluation metrics included drainage-area measurement and shape analysis techniques, including the Coefficient of Areal Correspondence (CAC), polygon-area percent error, cumulative drainage-area analysis, and convergence analysis.

The results of this study show that the New England Methods provided the most accurate results overall, especially the method that incorporated the WBD drainage divides in the DEM. Performance for all methods did not appear to vary across the various Subbasins. In general, for topographic-based methods using 30-m NED data, greater accuracy was attained in areas of high relief contrasted with flatter areas. The accuracy for flatter areas was improved by the integration of the WBD into the New England Method.

The Outlet Matching Method was determined to be the least effective method because of the fundamental differences between the NHD streams and the EDNA synthetic stream network. These differences led to the inability to match many EDNA catchments to NHD flowlines. Because of this, the Outlet Matching Method was deemed unsuitable for determining catchments for the NHD flowlines.

The Thiessen Polygon Method also was problematic because it did not include topographic information in the delineation of catchments. The Thiessen Polygon Method did not yield accurate drainage area measurements for flowlines with relatively small drainage areas. This deficiency is especially important because most of the NHD flowlines have a cumulative drainage area of less than 20 square miles.

On the basis of summary statistics of the various evaluation metrics, the Raster Seeding Method results were slightly less accurate than the New England Method without WBD. A detailed inspection of the larger boundary differences between the verification data and the Raster Seeding Method revealed problems that were caused by the horizontal displacement of the DEM flow paths and the NHD streams. These boundary errors in the Raster Seeding Method caused erroneous parallel-to-stream delineations that were sometimes extensive, particularly in areas of low relief and within waterbodies.

Catchments now have been produced nationally using this process, integrating the NHD with the NED and WBD. These catchments and the related suite of geospatial datasets have become known as the “NHDPlus” and currently are being used in a wide variety of hydrologic applications.

## References Cited

- Dewald, T.G., Horn, R., Greenspun, R., Taylor, P., Manning, L., and Montalbano, A., 1985, STORET reach retrieval documentation: Washington, DC, U.S. Environmental Protection Agency.
- Environmental Systems Research Institute, Inc. (ESRI), 1999, Using ARC GRID with Arc/Info: Redlands, CA, ESRI, v. 2, 436 p.
- Environmental Systems Research Institute, Inc. (ESRI), 2003, ArcDoc: Redlands, CA, ArcInfo Command Reference, ArcInfo v. 8.3 software documentation.
- Hellweger, Ferdi, and Maidment, David, 1997, AGREE-DEM surface reconditioning system: Austin, TX, University of Texas, accessed November 2005, at <http://www.crwr.utexas.edu/gis/gishyd98/quality/agree/agree.htm>
- Jenson, S.K., and Dominique, J.O., 1988, Extracting topographic structure from digital elevation data for Geographic Information System analysis: Photogrammetric Engineering and Remote Sensing, v. 54, no. 11, p. 1593–1600.
- Jobson, H.E., 1996, Prediction of traveltime and longitudinal dispersion in rivers and streams: U.S. Geological Survey Water-Resources Investigations Report 96–4013, 69 p.
- Kost, Jay, Verdin, K., Worstell, B., and Kelly, G., 2003, Methods and tools for the development of hydrologically conditioned elevation data and derivatives for national applications: U.S. Geological Survey, accessed October 2007, at <http://edna.usgs.gov/Edna/pubs/KostEDNA.pdf>
- Moore, R.B., Johnston, C.M., Robinson, K.W., and Deacon, J.R., 2004, Application of spatially referenced regression models to evaluate total nitrogen and phosphorus in New England streams: U.S. Geological Survey Water-Resources Investigations Report 2004–5012, 42 p., available at <http://pubs.water.usgs.gov/sir2004-5012>
- Nolan, J.V., Brakebill, J.W., Alexander, R.B., and Schwarz, G.E., 2003, ERF1\_2—Enhanced River Reach File 2.0: U.S. Geological Survey Open-File Report 2002–40, accessed August 2006, at [http://water.usgs.gov/lookup/getspatial?erf1\\_2](http://water.usgs.gov/lookup/getspatial?erf1_2)
- Olivera, F., Furnans, J., Maidment, D., Djokic, D., and Ye, Z., 2002, Drainage systems, in Maidment, D.R., ed., Arc Hydro—GIS for water resources: Redlands, CA, ESRI Press, chap. 4, 78 p.
- Saunders, W., 2000, Preparation of DEMs for use in environmental modeling analysis, in Maidment, D., Djokic, D., eds., 2000, Hydrologic and hydraulic modeling support: Redlands, CA, ESRI Press, p. 29–51.
- Seaber, P.R., Kapinos, P.F., and Knapp, G.L., 1987, Hydrologic Unit maps: U.S. Geological Survey Water-Supply Paper 2294, 64 p.
- Taylor, P.J., 1977, Quantitative methods in geography, an introduction to spatial analysis: Prospect Heights, IL, Waveland Press, chap. 5, 375 p.
- Thiessen, A.H., and Alter, J.C., 1911, Climatological data for July 1911: District No. 10, Great Basin: Monthly Weather Review, July 1911, p. 1082–1089.
- U.S. Department of Agriculture, Natural Resources Conservation Service, 2004, Watershed Boundary Dataset, accessed November 2006, at <http://www.ncgc.nrcs.usda.gov/products/datasets/watershed/>.
- U.S. Environmental Protection Agency, 1994, The U.S. EPA Reach File Version 3.0 alpha release technical reference, accessed December 2006, at <http://www.epa.gov/waters/doc/techref.html>
- U.S. Environmental Protection Agency, 2006, National Hydrography Dataset Plus, accessed March 13, 2009, at <http://www.epa.gov/waters>
- U.S. Geological Survey, 1996, US GeoData Digital Line Graphs: U.S. Geological Survey Fact Sheet 078–96, accessed September 2007, at <http://erg.usgs.gov/isb/pubs/factsheets/fs07896t.pdf>
- U.S. Geological Survey, 1999a, The National Hydrography Dataset: U.S. Geological Survey Fact Sheet 106–99, accessed December 2006, at <http://erg.usgs.gov/isb/pubs/factsheets/fs10699.html>
- U.S. Geological Survey, 1999b, National Elevation Dataset: U.S. Geological Survey Fact Sheet 148–99, accessed September 2007, at <http://erg.usgs.gov/isb/pubs/factsheets/fs14899.html>
- U.S. Geological Survey, 2001a, USGS GeoData Digital Raster Graphics: U.S. Geological Survey Fact Sheet 088–01, accessed October 2007, at <http://erg.usgs.gov/isb/pubs/factsheets/fs08801.html>
- U.S. Geological Survey, 2001b, Digital Orthophoto Quadrangles: U.S. Geological Survey Fact Sheet 057–01, accessed October 2007, at <http://erg.usgs.gov/isb/pubs/factsheets/fs05701.html>
- Verdin, K.L., and Verdin, J.P., 1999, A topological system for delineation and codification of the Earth's river basins: Journal of Hydrology, v. 218, p. 1–12.
- Vogel, R.M., Wilson, I.W., and Daly, C., 1999, Regional regression models of annual streamflow for the United States: Journal of Irrigation and Drainage Engineering, v. 125, no. 3, p. 148–157.

# Glossary

## A

**ArcHydro Tools** A set of public domain utilities developed jointly by the Center for Research in Water Resources (<http://www.crrw.utexas.edu>) of the University of Texas at Austin, and the Environmental Systems Research Institute, Inc. These tools support terrain processing, watershed delineation and attribute management.

**ArcInfo** Geographic information systems (GIS) software developed by Environmental Systems Research Institute.

**Arc Macro Language (AML)** A computer program scripting language used by ArcInfo.

**Artificial path** A National Hydrography Dataset (NHD) flowline feature type that represents a flow-path through a waterbody in the surface water network of the NHD.

## B

**Basin** The third largest category (level 3) of Hydrologic Units defined within the Watershed Boundary Dataset (WBD). *See* table 1.

## C

**Catchment** The incremental drainage area for a linear hydrologic feature found in the National Hydrography Dataset (NHD).

**Coefficient of Areal Correspondence (CAC)** A simple measure of the extent to which two polygons correspond to one another.

**Cumulative drainage area** The catchment area for a specific flowline combined with the catchment areas for all upstream flowlines.

## D

**DEM flow-path displacement** The horizontal positional offset between a mapped stream in the NHD and that of a synthetic stream derived from a digital elevation model.

**DEM surface reconditioning** *See* stream burning and walling.

**Digital elevation model (DEM)** Digital raster representations of terrain elevations for ground positions at regularly spaced horizontal intervals.

**Digital Line Graphic (DLG)** Digital vector representations of cartographic information derived from USGS maps and related sources.

### Digital Orthophoto Quadrangle (DOQ)

A computer-generated image of an aerial photograph in which the image displacement caused by terrain relief and camera tilt has been removed.

**Digital Raster Graphic (DRG)** A georeferenced digital image resulting from scanning a paper USGS topographic map for use on a computer.

**Drainage area** An area that defines the geographic extent of the land surface that drains to a specific location.

**Drainage-area boundary** The polygon that defines the perimeter of a drainage area.

**Drainage-area divide** The boundary line between two different drainage areas along a topographic ridge.

**Drainage-area measurement** The number that defines a drainage area in terms of square miles or square kilometers.

## E

**Elevation Derivatives for National Applications (EDNA)** A set of geospatial layers derived from the National Elevation Dataset.

## F

**Flowline** A mapped stream segment or a path through a waterbody in the surface-water network of the National Hydrography Dataset. The basic unit of the NHD linear surface-water network.

**Flow table (also NHD flow table)** A database table that contains the interconnections between the NHD flowlines.

**G****Geographic Names Information Systems**

**(GNIS)** A database that contains name and location information about more than two million physical and cultural features located throughout the United States and its territories.

**H**

**Hydrologic Unit** The Hydrologic Unit system is a standardized watershed classification system developed by USGS in the mid 1970s. Hydrologic units are area boundaries organized in a nested hierarchy by size. They range in size from Region, Subregion, Basin, Subbasin, Watershed, Subwatershed. *See* table 1.

**Hydrologic Unit Code (HUC)** Codes assigned to a series of nested hydrologic divisions of the landscape. *See* table 1.

**I**

**Incremental drainage area** For a specific hydrologic feature, such as a stream segment, it defines the local area that drains directly into that feature. Synonymous with catchment.

**L**

**Light Detection and Ranging (LIDAR)** An optical remote sensing technology that measures properties of scattered light to find range and (or) other information of a distant target.

**N**

**National Elevation Dataset (NED)** A seamless elevation coverage of the conterminous United States, Hawaii, Alaska, and the island territories.

**National Hydrography Dataset (NHD)**

A comprehensive set of digital spatial data representing the surface water of the United States using common features such as lakes, ponds, streams, rivers, canals, and oceans.

**National Water Information System**

**(NWIS)** A principal repository of national water resources data.

**NHD flow table (also Flow table)** A database table that contains the interconnections between the NHD flowlines.

**NHDPlus** An integrated suite of application-ready geospatial datasets that incorporate many of the best features of the National Hydrography Dataset (NHD), the National Elevation Dataset (NED), the National Land Cover Dataset (NLCD), and the Watershed Boundary Dataset (WBD).

**NHD Reach (also Reach)** A uniquely identified linear feature that consists of one or more flowlines.

**O**

**Outlet** The most downstream location of a surface-water feature, such as a flowline or catchment.

**P**

**Pfafstetter code** A hierarchical watershed identifier based upon topology of the land surface assigned by a methodology developed by Otto Pfafstetter in 1989.

**R**

**Raster** A digital computer format for representing an image or mapped data information. The structure consists of data stored in a rectangular grid array of pixels. Synonymous with grid.

**Reach (also NHD Reach)** A uniquely identified linear feature that consists of one or more flowlines.

**Reach code** A unique, permanent identifier in the National Hydrography Dataset.

**Region** The largest category (level 1) of Hydrologic Units defined within the Watershed Boundary Dataset (WBD). *See* table 1.

**S**

**Stream burning** Using a mapped stream network to trench a DEM to improve how accurately the DEM flow paths match the streams, to ensure DEM-derived catchment boundaries fit the stream network.

**Streamflow** The volume of water flowing past a fixed point in a fixed unit of time.

**Stream segment (also see flowline)** Part of a stream, often extending between tributary confluences.

**Subbasin** The fourth largest (level 4) category of Hydrologic Units defined within the Watershed Boundary Dataset (WBD).  
*See table 1.*

**Subregion** The second largest category (level 2) of Hydrologic Units defined within the Watershed Boundary Dataset (WBD).  
*See table 1.*

**Subwatershed** The sixth largest category (level 6) of Hydrologic Units defined within the Watershed Boundary Dataset (WBD).  
*See table 1.*

## T

**Thiessen polygon** A polygon surrounding an individual feature whose boundary is halfway between that feature and the nearest other features.

## W

**Walling** Using a representation of the known drainage boundaries to build up or mathematically warp a DEM to improve how accurately catchment boundaries replicate known drainage boundaries.

**Watershed** The fifth largest (level 5) category of Hydrologic Units defined within the Watershed Boundary Dataset (WBD).  
*See table 1.*

### **Watershed Boundary Dataset (WBD)**

A base-line hydrologic drainage boundary framework, accounting for all land and surface areas of the United States.

## U

### **USEPA Reach File Version 1.0 (RF1)**

A national hydrologic database that uniquely identifies and interconnects the stream segments or “reaches” that compose the Country’s surface water drainage system at a 1:500,000 scale.

### **USEPA Reach File Version 3.0 (RF3)**

A national hydrologic database that uniquely identifies and interconnects the stream segments or “reaches” that compose the Country’s surface water drainage system at a 1:100,000 scale.

THIS PAGE INTENTIONALLY LEFT BLANK

## Appendix. Detailed Discussion for each Study Subbasin

Great Egg Harbor, NJ (02040302).....	45
Comparison of Method Boundaries with WBD Boundaries.....	45
Comparison of Methods with Streamgage Data .....	51
Summary of Results for Great Egg Harbor .....	51
Seneca River, NC and SC (03060101).....	54
Comparison of Method Boundaries with WBD Boundaries.....	54
Comparison of Methods with Streamgage Data .....	57
Summary of Results for Seneca River.....	59
Pine River, MN (07010105) .....	59
Comparison of Method Boundaries with WBD Boundaries.....	59
Comparison of Methods with Streamgage Data .....	69
Summary of Results for Pine River .....	71
South Platte Headwaters, CO (10190001).....	71
Comparison of Method Boundaries with WBD Boundaries.....	71
Comparison of Methods with Streamgage Data .....	77
Summary of Results for South Platte Headwaters.....	80
Pine Valley, Utah (16020302).....	81
Willapa Bay, WA (17100106).....	85

## Figures

1-1. Map showing the draft Watershed Boundary Dataset (WBD) boundaries, and National Hydrography Dataset (NHD) flowlines, Great Egg Harbor Subbasin, NJ .....	46
1-2. Graph showing normalized Coefficient of Areal Correspondence (CAC) scores, by National Hydrography Dataset (NHD) flowline common identifier, for all methods tested (except the Outlet Matching Method), Great Egg Harbor Subbasin, NJ .....	47
1-3. Graph showing drainage-area percent error in relation to Watershed Boundary Dataset cumulative drainage areas, for all methods tested, Great Egg Harbor Subbasin, NJ .....	48
1-4. Map showing (A) Watershed Boundary Dataset (WBD) and the New England Method without WBD (NEM-without-WBD) boundaries, (B) area of detail where boundaries differ overlaid on a Digital Raster Graphic, and (C) area in detail with boundary error overlaid on a Digital Orthophoto Quadrangle, Great Egg Harbor Subbasin, NJ ....	49
1-5. Map showing location of streamgages and associated drainage areas used as verification datasets, Great Egg Harbor Subbasin, NJ .....	52
1-6. Map showing (A) the Outlet Matching Method and verification drainage-area boundaries for a streamgage (01411000) and (B) the unusual boundary from the Outlet Matching Method for a streamgage (01411300), Great Egg Harbor Subbasin, NJ.....	53
1-7. Map showing the Watershed Boundary Dataset (WBD) Subwatershed boundaries, and National Hydrography Dataset (NHD) flowlines, Seneca River Subbasin, NC and SC.....	55
1-8. Graph showing normalized Coefficient of Areal Correspondence (CAC) scores, by National Hydrography Dataset (NHD) flowline common identifier, for all methods tested, Seneca River Subbasin, NC and SC.....	56

1–9.	Map showing location of streamgages and associated drainage areas used as verification datasets, Seneca River Subbasin, NC and SC.....	58
1–10.	Map showing the draft Watershed Boundary Dataset (WBD) Subwatershed boundaries, and National Hydrography Dataset (NHD) flowlines, Pine River Subbasin, MN.....	60
1–11.	Graph showing normalized Coefficient of Areal Correspondence (CAC) scores, by National Hydrography Dataset (NHD) flowline common identifier, for all methods tested, Pine River Subbasin, MN.....	61
1–12.	Graph showing drainage-area percent error in relation to Watershed Boundary Dataset cumulative drainage areas, for all methods tested, Pine River Subbasin, MN.....	62
1–13.	Map showing (A) Watershed Boundary Dataset (WBD) and New England Method without WBD (NEM-without-WBD) boundaries, and (B) and (C) areas of detail showing boundary differences at wetlands, Pine River Subbasin, MN.....	64
1–14.	Map showing (A) Watershed Boundary Dataset (WBD) and New England Method without WBD (NEM-without-WBD) boundaries, (B) area of detail where the boundaries differ overlaid on a Digital Raster Graphic, and (C) area in detail with boundary error overlaid on a Digital Orthophoto Quadrangle, Pine River Subbasin, MN.....	65
1–15.	Map showing (A) Watershed Boundary Dataset (WBD) and New England Method without WBD (NEM-without-WBD) boundaries, (B) area of detail, showing boundary differences at Downey Lake, and (C) area of detail, showing boundary differences at a wetland, Pine River Subbasin, MN.....	66
1–16.	Map showing Watershed Boundary Dataset (WBD) and the Raster Seeding Method boundaries for the two Subwatersheds with the lowest normalized Coefficient of Areal Correspondence (CAC) scores, Pine River Subbasin, MN.....	67
1–17.	Map showing location of streamgages and associated drainage areas used as verification datasets, Pine River Subbasin, MN.....	70
1–18.	Map showing the draft Watershed Boundary Dataset (WBD) Subwatershed boundaries, and National Hydrography Dataset (NHD) flowlines, South Platte Headwaters Subbasin, CO.....	72
1–19.	Graph showing normalized Coefficient of Areal Correspondence (CAC) scores, by National Hydrography Dataset (NHD) flowline common identifier, for all methods tested, South Platte Headwaters Subbasin, CO.....	73
1–20.	Graph showing drainage-area percent error in relation to Watershed Boundary Dataset cumulative drainage areas, for all methods tested, South Platte Headwaters Subbasin, CO.....	74
1–21.	Map showing (A) the Watershed Boundary Dataset (WBD) and the New England Method without WBD (NEM-without-WBD) boundaries for the Antero Reservoir Subwatershed and (B) area of detail showing boundary differences overlaid on the Digital Orthophoto Quadrangle (DOQ), South Platte Headwaters Subbasin, CO.....	76
1–22.	Map showing the Watershed Boundary Dataset (WBD) Subwatersheds and Raster Seeding Method boundaries with normalized Coefficient of Areal Correspondence (CAC) scores and digital elevation model (DEM) flow-path displacement boundary errors, South Platte Headwaters Subbasin, CO.....	78
1–23.	Map showing location of streamgages used as verification data, South Platte Headwaters, CO.....	79
1–24.	Map showing the draft Watershed Boundary Dataset (WBD) Subwatershed boundaries, and National Hydrography Dataset (NHD) flowlines, Pine Valley, Utah.....	82
1–25.	Graph showing normalized Coefficient of Areal Correspondence (CAC) scores, by National Hydrography Dataset (NHD) flowline common identifier, for all methods tested, Pine Valley Subbasin, Utah.....	83

1-26.	Graph showing drainage-area percent error in relation to Watershed Boundary Dataset cumulative drainage areas, for all methods tested, Pine Valley Subbasin, Utah .....	84
1-27.	Map showing location of streamgages and associated drainage areas used as verification datasets, Willapa Bay, WA .....	86
1-28.	Graph showing the computed Coefficient of Areal Correspondence (CAC) scores (not normalized), by streamgage, for the New England Method without Watershed Boundary Dataset (NEM-without-WBD), Raster Seeding Method, and Thiessen Polygon Method, Willapa Bay Subbasin, WA.....	87
1-29.	Graph showing drainage-area percent error in relation to streamgage drainage areas of the verification data for the New England Method without Watershed Boundary Dataset (NEM-without-WBD), Raster Seeding Method and Thiessen Polygon Method, Willapa Bay Subbasin, WA.....	88

## Tables

1-1.	Mean normalized Coefficient of Areal Correspondence score by delineation methods for Watershed Boundary Dataset boundaries in the Great Egg Harbor Subbasin, NJ ..	47
1-2.	Mean cumulative drainage-area errors for all delineation methods for various drainage-area sizes, Great Egg Harbor Subbasin, NJ .....	48
1-3.	Mean normalized Coefficient of Areal Correspondence score by delineation methods for Watershed Boundary Dataset Subwatersheds in the Seneca River Subbasin, NC and SC .....	56
1-4.	Mean cumulative drainage-area errors for all delineation methods for various drainage-area sizes, Seneca River Subbasin, NC and SC.....	57
1-5.	Mean normalized Coefficient of Areal Correspondence score by delineation methods for Watershed Boundary Dataset Subwatersheds in the Pine River Subbasin, MN .....	61
1-6.	Mean cumulative drainage-area errors for all delineation methods for various drainage-area sizes, Pine River Subbasin, MN .....	62
1-7.	Mean normalized Coefficient of Areal Correspondence score by delineation methods for Watershed Boundary Dataset Subwatersheds in the South Platte Headwaters Subbasin, CO .....	73
1-8.	Mean cumulative drainage-area errors for all delineation methods for various drainage-area sizes, South Platte Headwaters Subbasin, CO.....	74
1-9.	Mean drainage-area errors for all delineation methods (except the Outlet Matching Method), for the National Water Information System (NWIS) streamgages, grouped by various drainage-area sizes, South Platte Headwaters, CO .....	80
1-10.	Mean normalized Coefficient of Areal Correspondence scores by delineation method for Watershed Boundary Dataset Subwatersheds in the Pine Valley Subbasin, Utah..	83
1-11.	Mean cumulative drainage-area errors for all delineation methods for various drainage-area sizes, Pine Valley Subbasin, Utah.....	84
1-12.	Coefficient of Areal Correspondence scores for the New England Method without Watershed Boundary Dataset (NEM-without-WBD), Raster Seeding Method, and Thiessen Polygon Method, for streamgage drainage-areas, Willapa Bay Subbasin, WA .....	87
1-13.	Streamgage drainage-areas of verification data, drainage-area percent errors, and mean errors for the New England Method without Watershed Boundary Dataset, Raster Seeding Method, and Thiessen Polygon Method, Willapa Bay, WA .....	88

ABBREVIATIONS USED IN THIS APPENDIX

CAC	Coefficient of Areal Correspondence
ComID	Common Identifier, a unique identifier for an NHD feature
DEM	digital elevation model
DOQ	Digital Orthophoto Quadrangle
DRG	Digital Raster Graphic
EDNA	Elevation Derivatives for National Applications
GIS	geographic information system
HUC	Hydrologic Unit Code
NEM	New England Method
NHD	National Hydrography Dataset
NWIS	National Water Information System
USGS	United States Geological Survey
WBD	Watershed Boundary Dataset

## Appendix. Detailed Discussion for each Study Subbasin

### Great Egg Harbor, NJ (02040302)

#### Comparison of Method Boundaries with WBD Boundaries

Draft Watershed Boundary Dataset (WBD) boundaries delineated to the 14-digit hydrologic unit level, as shown in figure 1–1, were obtained for the Great Egg Harbor Subbasin in New Jersey. Forty-five WBD boundaries were used in the Coefficient of Areal Correspondence (CAC) analysis—the results of which are shown in figure 1–2 and table 1–1. A graph, developed as part of the convergence analysis, displaying the distribution of drainage-area percent errors is shown in figure 1–3. Convergence analysis of cumulative drainage-area percent errors are also summarized by drainage-area size categories, as shown in table 1–2. The Outlet Matching Method produced fewer comparisons with the WBD because the Outlet Matching Method catchments could not always be linked to the National Hydrography Dataset (NHD) flowlines. For the Outlet Matching Method, 39 WBD boundaries were compared in the analysis.

#### New England Method with WBD (NEM-with-WBD)

Fourteen of the 45 WBD boundaries have CAC scores (before normalization to 100) for the NEM-with-WBD of less than 95. For all 14, the lower CAC scores were attributed to mismatches between the outlet of the WBD boundaries and the NHD flowlines assigned as best representing the WBD outlets (WBD/NHD-flowline outlet mismatch). The Great Egg Harbor Subbasin had a higher incidence of WBD/NHD-flowline outlet mismatches than the other Subbasins owing to the greater detail in the divisions of the WBD (14-digit Hydrologic Unit Code (HUC) level). WBD data for the Great Egg Harbor Subbasin include subdivisions along flowlines where no confluences occur in the medium-resolution NHD. These subdivisions may represent the location of a 1:24,000-scale stream confluence or water-quality monitoring site, dam, or other feature.

The mean CAC score (before normalization to 100) for the NEM-with-WBD is 91.6, the highest of all the methods. The mean cumulative drainage-area error for NEM-with-WBD is 1.65 percent (table 1–2) and is attributed to the WBD/NHD-flowline outlet mismatch errors. No cumulative drainage areas from the NHD flowlines in this Subbasin fit within the 8-digit HUC designation.

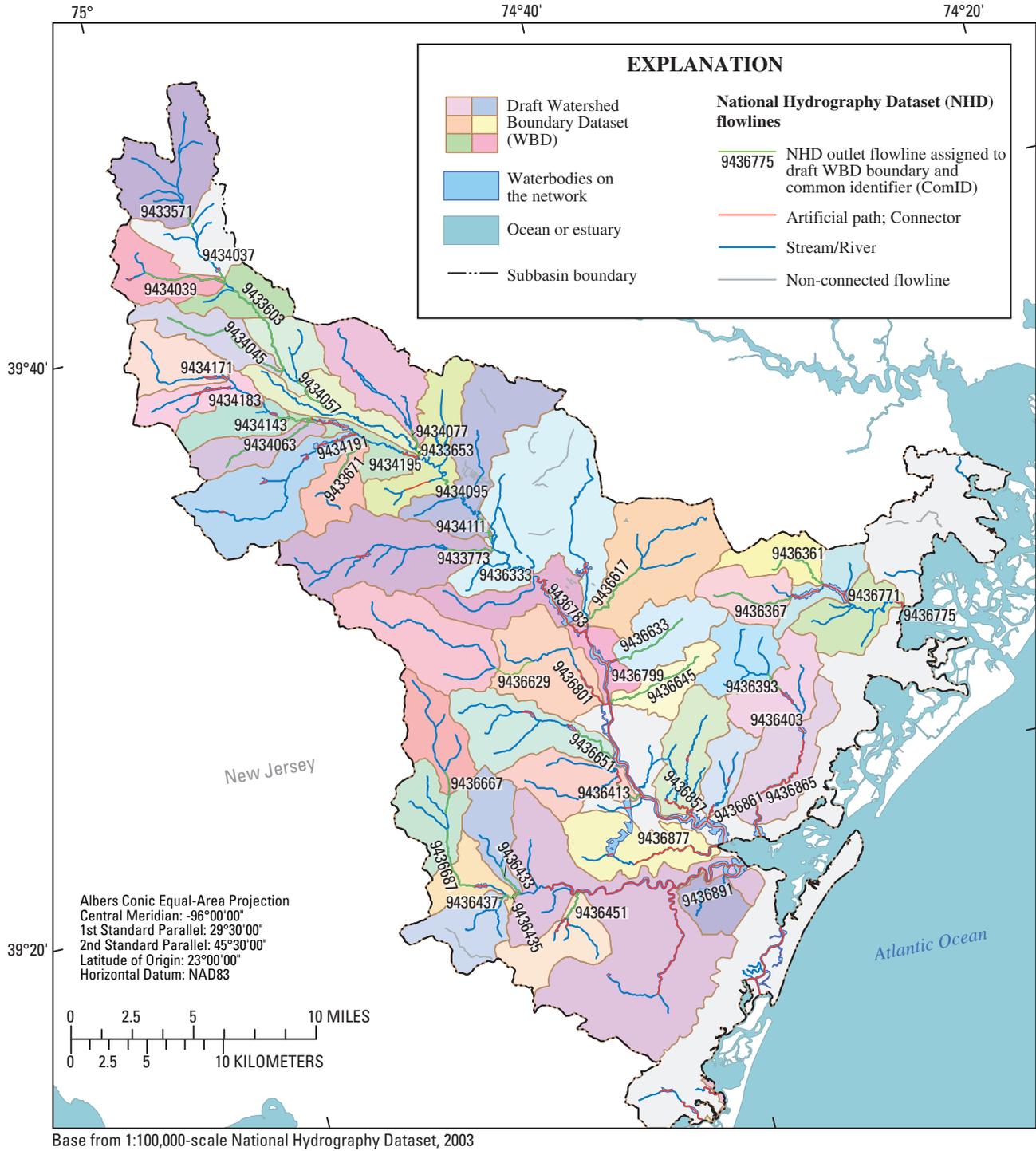
#### New England Method without WBD (NEM-without-WBD)

The mean normalized CAC score for the NEM-without-WBD is 87.5 (table 1–1), the highest score after NEM-with-WBD. Twenty-six of the 45 WBD boundaries (58 percent) have normalized CAC scores greater than 90. Five WBD boundaries (11 percent) have normalized CAC scores less than 80 (table 1–1). The five lowest individual normalized CAC scores are 42.6, 67.2, 73.7, 78.1, and 79.0.

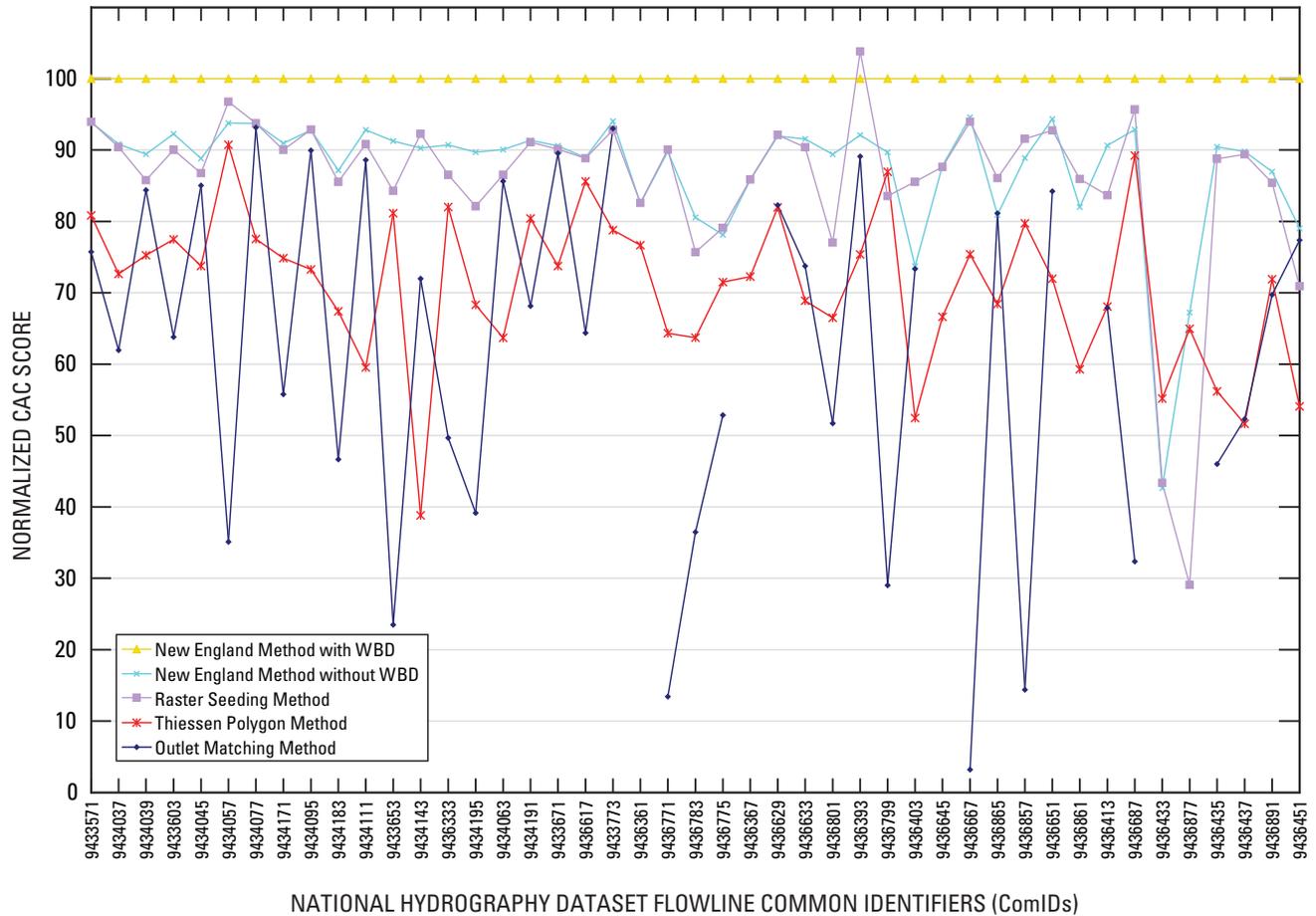
The lowest normalized CAC score of the NEM-without-WBD is 42.6. The difference in area between the WBD boundary and the corresponding NEM-without-WBD boundary (fig. 1–4A) is 3.25 mi<sup>2</sup>, which is the highest normalized polygon-area error for this method at 46.4 percent. The WBD and the NEM-without-WBD boundaries were overlaid on the 1:24,000-scale Digital Raster Graphic (DRG) (fig. 1–4B) to determine visually which boundary was more accurate. The differences in the delineations occur at a wetland and the area of difference is attributed, in most part, to the drainage area of this wetland. In the WBD dataset, a road grade was used as a divide to delineate the WBD boundary through the wetland area. On the basis of the DRG, it was difficult to determine which boundary was correct (fig. 1–4B). However, on the basis of the surface waters depicted on the U.S. Geological Survey (USGS) Digital Orthophoto Quadrangle (DOQ), it does appear that the WBD data are correct (fig. 1–4C).

The second lowest normalized CAC score of the NEM-without-WBD is 67.2. The area of difference between the NEM-without-WBD and the WBD boundary is 3.02 mi<sup>2</sup>, a 25.5 percent normalized polygon-area error. The WBD area drains into the tidal portion of the Great Egg Harbor River, and large tidal wetlands with complex drainage patterns and flood levees are present within the WBD boundary. Both delineations appear to be acceptable here on the basis of on the complex drainages.

The NEM-without-WBD boundary with the third lowest normalized CAC score at 73.7 has a normalized polygon-area error of 34 percent. There are several areas of difference between the WBD and NEM-without-WBD boundaries. On the basis of DRG information, it was difficult to determine which method produced the more accurate boundary at the largest area of difference (approximately 2 mi<sup>2</sup>). Data from the WBD provided a more accurate delineation at another area with a difference of 0.5 mi<sup>2</sup>. Other differences along the boundaries involved areas of less substantial sizes. In one case, neither the NEM-without-WBD boundary nor the WBD boundary is correct.



**Figure 1-1.** The draft Watershed Boundary Dataset (WBD) boundaries, and National Hydrography Dataset (NHD) flowlines, Great Egg Harbor Subbasin, NJ. (The location of this Subbasin is shown in figure 15.)



**Figure 1–2.** Normalized Coefficient of Areal Correspondence (CAC) scores, by National Hydrography Dataset (NHD) flowline common identifier, for all methods tested (except the Outlet Matching Method), Great Egg Harbor Subbasin, NJ.

**Table 1–1.** Mean normalized Coefficient of Areal Correspondence score by delineation methods for Watershed Boundary Dataset boundaries in the Great Egg Harbor Subbasin, NJ.

[CAC, Coefficient of Areal Correspondence; NEM, New England Method; WBD, Watershed Boundary Dataset]

Method	Mean normalized CAC	Normalized CACs greater than 90		Normalized CACs less than 80	
		Number of boundaries	Percentage of boundaries	Number of boundaries	Percentage of boundaries
NEM-with-WBD	100	45	100	0	0
NEM-without-WBD	87.5	26	58	5	11
Raster Seeding Method	85.8	20	44	6	13
Thiessen Polygon Method	70.9	1	2	36	80

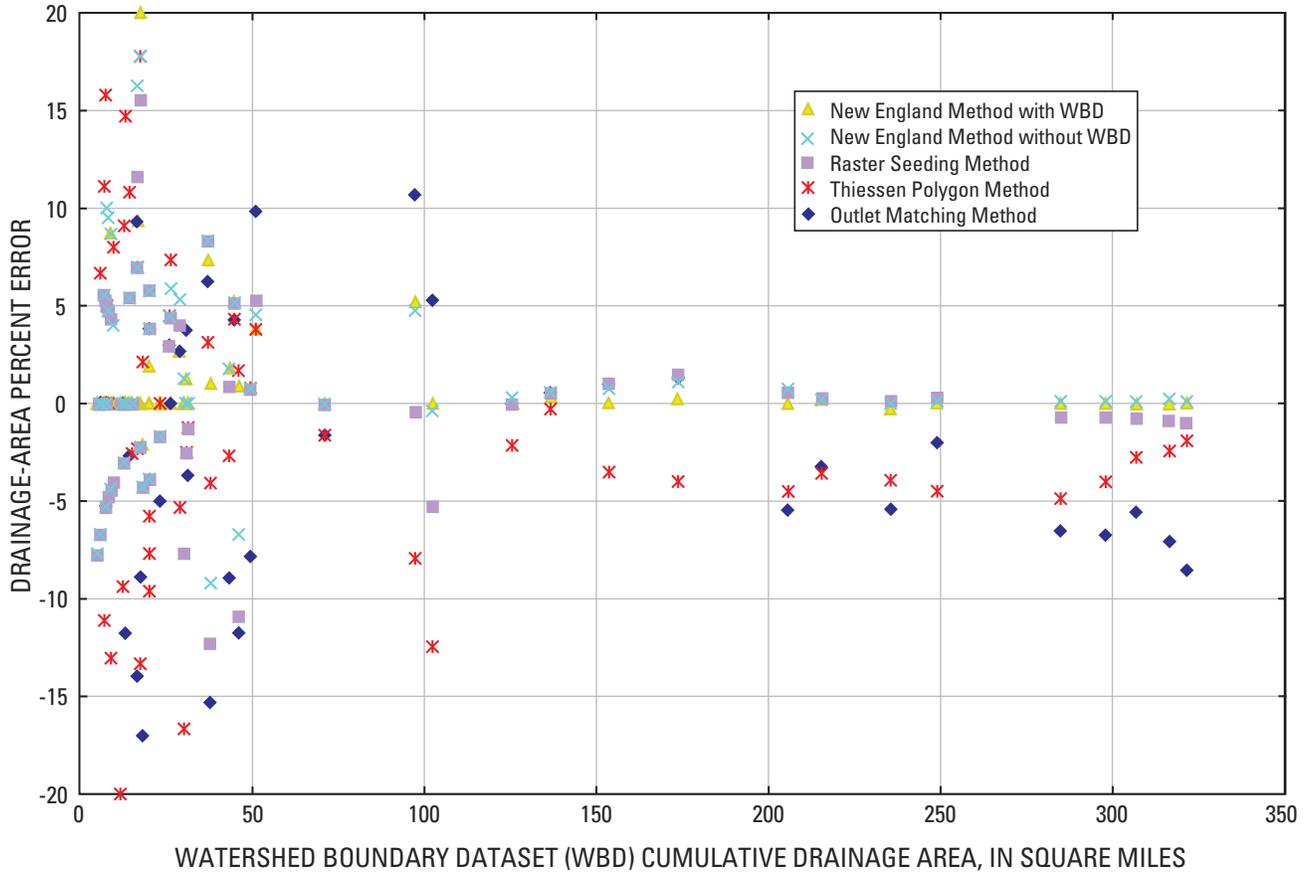
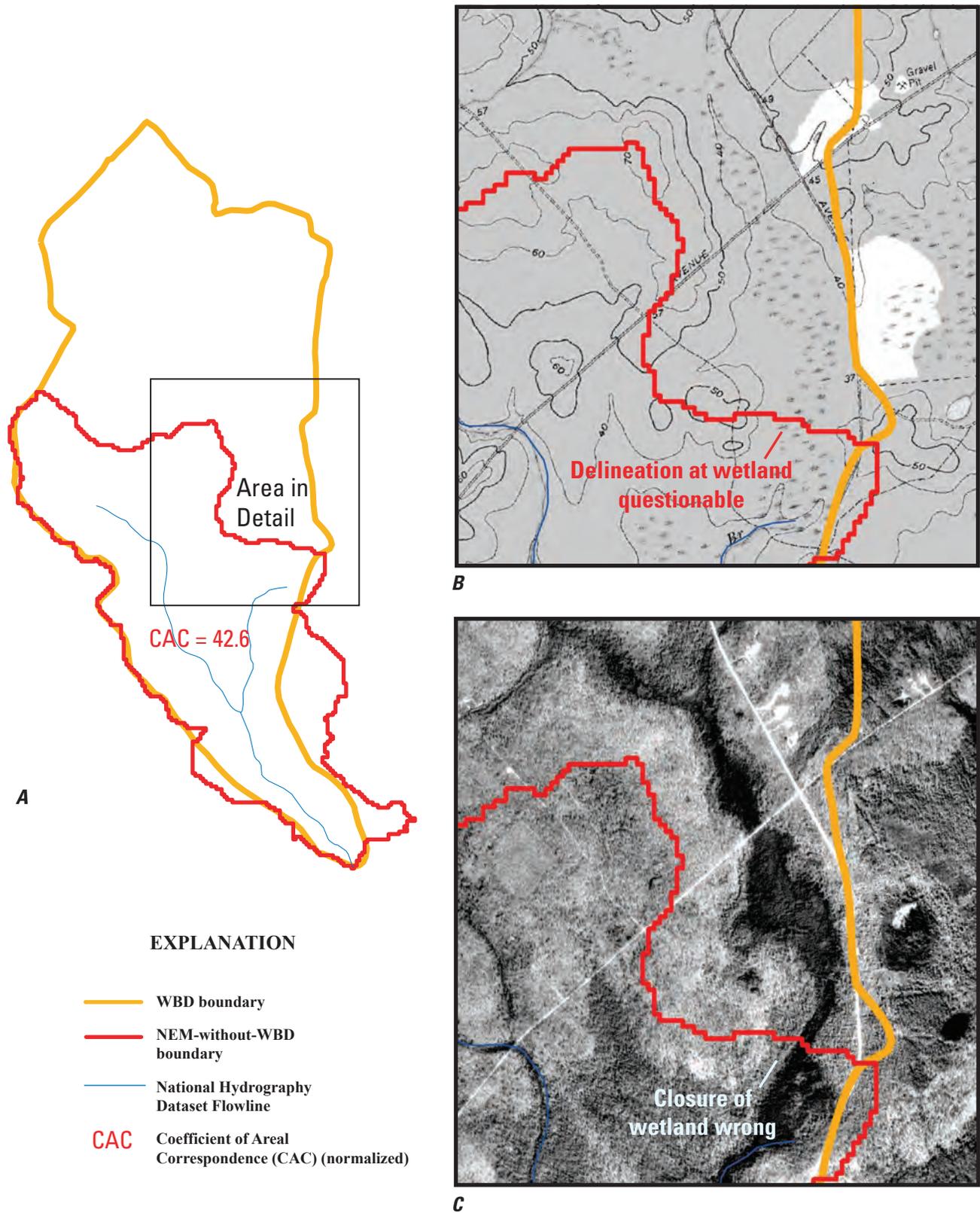


Figure 1-3. Drainage-area percent error in relation to Watershed Boundary Dataset cumulative drainage areas, for all methods tested, Great Egg Harbor Subbasin, NJ.

Table 1-2. Mean cumulative drainage-area errors for all delineation methods for various drainage-area sizes, Great Egg Harbor Subbasin, NJ.

[OM, Outlet Matching; mi<sup>2</sup>, square miles; NEM, New England Method; WBD, Watershed Boundary Dataset; na, not applicable]

Method (61 units total; 54 units OM)	Mean cumulative drainage-area error, in percent				
	Drainage-area size category				
	All drainage areas	Less than 20 mi <sup>2</sup>	20–60 mi <sup>2</sup>	Greater than 60–390 mi <sup>2</sup>	Greater than 390 mi <sup>2</sup>
NEM-with-WBD	1.65	2.18	1.96	0.39	na
NEM-without-WBD	4.98	8.06	4.04	0.61	na
Raster Seeding Method	5.25	8.01	4.80	0.88	na
Thiessen Polygon Method	9.86	16.35	4.77	4.03	na
Outlet Matching Method	11.30	18.81	8.50	4.44	na



**Figure 1-4.** (A) Watershed Boundary Dataset (WBD) and the New England Method without WBD (NEM-without-WBD) boundaries, (B) area of detail where boundaries differ overlaid on a Digital Raster Graphic, and (C) area in detail with boundary error overlaid on a Digital Orthophoto Quadrangle, Great Egg Harbor Subbasin, NJ. (The location of this WBD boundary is shown in figure 1-1.)

It was difficult to determine which dataset was correct for the fourth lowest normalized CAC score, 78.1, for the NEM-without-WBD. Cultural features in a heavy urban setting depicted on the DRGs made interpreting the underlying topographic contours difficult. The digital elevation model (DEM) delineation of the NEM-without-WBD appears to be correct, but this is not verifiable. The difference in area between the dataset boundaries is 0.97 mi<sup>2</sup> with a normalized polygon-area error of 20.7 percent.

The fifth lowest CAC score is 79.0. Either the WBD or the NEM-without-WBD delineation seems to be possible. The WBD includes road grades as drainage-area divides in certain areas, which are presumed to be correct. Because the NEM-without-WBD delineations rely solely on topographic data that do not take into account road grades, differences occur between the dataset boundaries. In these road-related differences, both boundaries can be considered acceptable on the basis of the limited information available. However, from the DOQ imagery it does appear that the use of the roads as divides in certain areas enhanced the accuracy of the boundaries. The difference in area for this WBD boundary is 0.16 mi<sup>2</sup>, which represents a normalized polygon-area error of 2.2 percent.

In the cumulative drainage-area analysis, the NEM-without-WBD results closely tracked with the Raster Seeding Method results, but were slightly more accurate. The mean cumulative drainage-area error is 4.98 percent, compared to 5.25 percent for the Raster Seeding Method (table 1–2). The NEM-without-WBD mean cumulative drainage-area error for drainage areas less than 20 mi<sup>2</sup> is 8.06 percent, which is similar to the Raster Seeding Method error of 8.01 percent. An 8 percent error for computed drainage areas is somewhat high; however, about 2 percent of the error contribution is attributed to WBD/NHD-flowline outlet mismatch.

### Raster Seeding Method

The mean normalized CAC score for the Raster Seeding Method is 85.8, a close second to the NEM-without-WBD of 87.5. Twenty of the 45 Raster Seeding Method boundaries (44 percent) have normalized CAC scores greater than 90, and just 6 (13 percent) have a normalized CAC score less than 80. The six normalized CAC scores less than 80 are 29.1, 43.4, 70.9, 75.7, 77.0, and 79.1. One normalized Raster Seeding Method CAC score of greater than 100 (103.8) is explained by a combination of DEM flow-path displacement and WBD/NHD-flowline outlet mismatch, shown in figure 19 of this report.

Parallel stream artifacts from DEM flow-path displacement also contributed to the lowering of normalized CAC scores, affecting three of the six lowest Raster Seeding Method scores. Overall, the catchment delineations of the Raster Seeding Method and the NEM-without-WBD are very similar. Of the Raster Seeding Method catchment delineation lines, 81 percent are in perfect correspondence to the drainage divides of the NEM-without-WBD. Most of the delineation

differences are due to errors associated with DEM flow-path displacement observed with the Raster Seeding Method.

The lowest normalized CAC score of 29.1 is attributed mostly to the DEM flow-path displacement. This compared boundary is also where the NEM-without-WBD scored 67.2, the second lowest for that method; the area contains tidal wetlands with complex drainage patterns. The Raster Seeding Method boundary runs parallel to the NHD, leading to a less accurate delineation. The area of difference between the Raster Seeding Method and WBD boundaries is 6.54 mi<sup>2</sup> (a 55.8-percent normalized polygon-area error), and the difference between the NEM-without-WBD and the WBD boundaries is 3.02 mi<sup>2</sup> difference (a 25.5-percent error).

The second lowest Raster Seeding Method normalized CAC score of 43.4 is for the WBD boundary shown in figure 1–4, which also had the lowest normalized CAC score for the NEM-without-WBD. Delineations from both the Raster Seeding Method and NEM-without-WBD are the same at the wetland area identified in figure 1–4B. Once again, the WBD provided the more accurate delineation. The area of difference is 3.61 mi<sup>2</sup> between the Raster Seeding Method and the WBD boundaries, representing a normalized polygon-area error of 51.6 percent.

DEM flow-path displacement led to a less accurate delineation and contributed to the third lowest Raster Seeding Method normalized CAC score of 70.9. This is the same WBD area discussed earlier as the NEM-without-WBD with the fifth lowest CAC score (79.0) for that method. As for the NEM-without-WBD boundary, the use of roads as divides in the WBD contributed to some of the differences between the Raster Seeding Method and the WBD boundaries. There is a stretch in the Raster Seeding Method boundary where the delineation improperly parallels the NHD stream, contributing to additional differences from the NEM-without-WBD boundary. The area of difference between the Raster Seeding Method and corresponding WBD boundary is 0.63 mi<sup>2</sup> (7.3-percent normalized polygon-area error), of which about 0.47 mi<sup>2</sup> is directly attributable to DEM flow-path displacement.

Both the WBD and the Raster Seeding Method data produced boundary delineations with weaknesses, resulting in the fourth lowest normalized CAC score for the Raster Seeding Method (75.7). The delineations for this WBD boundary from both the Raster Seeding Method and NEM-without-WBD are identical in many areas. The effects of DEM flow-path displacement contributed to additional errors in the Raster Seeding Method boundary.

A determination could not be made as to which delineation was more accurate for a Raster Seeding boundary with the fifth lowest normalized CAC score of 77.0. It was also difficult to determine which delineation was the most accurate for the boundary with the sixth lowest normalized CAC score (79.1) for the Raster Seeding Method.

## Thiessen Polygon Method

The mean normalized CAC score for the Thiessen Polygon Method is 70.9. Only one normalized CAC score was determined to be greater than 90 (at 90.36). In contrast, 80 percent of all Thiessen Polygon Method boundaries have normalized CAC scores less than 80, compared to just 11 and 13 percent for the NEM-without-WBD and Raster Seeding Method, respectively. The Thiessen Polygon Method has the highest mean normalized polygon-area error of 17.17 percent, whereas the Raster Seeding Method has a 9 percent mean error, and NEM-without-WBD has a 6.89 percent mean error.

In the cumulative drainage-area analysis, the Thiessen Polygon Method has the fourth highest error with a mean drainage-area error of 9.86 percent (table 1–2). The mean Thiessen Polygon Method cumulative drainage-area error for drainage areas less than 20 mi<sup>2</sup> is roughly twice that of the mean errors for the Raster Seeding Method and the NEM-without-WBD. The mean Thiessen Polygon Method cumulative drainage-area error (4.77 percent) for drainage areas of 20 to 60 mi<sup>2</sup> tracks closely with the Raster Seeding Method (4.80 percent) and the NEM-without-WBD (4.04 percent). For drainage areas greater than 60 to 390 mi<sup>2</sup> in size, the Thiessen Polygon Method has a higher mean error at 4.03 percent, compared to 0.88 percent for the Raster Seeding Method and 0.61 percent for NEM-without-WBD.

## Outlet Matching Method

The Outlet Matching Method has the lowest CAC scores of all the methods with a mean CAC of 61.4. Thirty-nine of the 45 WBD boundaries were used in the CAC analysis because six NHD flowlines could not be assigned a corresponding Elevation Derivatives for National Applications (EDNA) catchment area. The Outlet Matching Method also has the highest polygon-area errors (not normalized) with a mean error of 40.17 percent; the Thiessen Polygon Method, the Raster Seeding Method, and NEM-without-WBD have mean errors of 19.54, 11.71, and 10.46 percent, respectively. In the cumulative drainage-area analysis, the Outlet Matching Method has the highest mean error of 11.30 percent. The Outlet Matching Method also has the highest mean error in each drainage-area size category; the highest is 18.81 percent for drainage areas less than 20 mi<sup>2</sup> (table 1–2).

Results for the Outlet Matching Method are provided only for the Great Egg Harbor Subbasin in this Appendix because it was representative of the poor performance of the Outlet Matching Method.

## Comparison of Methods with Streamgage Data

Geographic information system (GIS) boundary data were obtained for two streamgage drainage areas (fig. 1–5) to be used as verification datasets for the Great Egg Harbor Subbasin. National Water Information System (NWIS) drainage-area measurements also were obtained for the two streamgages. The NWIS drainage-area measurement for streamgage 01411000 is 57.1 mi<sup>2</sup>, and the GIS measurement

is 56.94 mi<sup>2</sup>. For streamgage 01411300, the NWIS drainage area and GIS drainage-area measurements are the same, 30.77 mi<sup>2</sup>. The streamgage drainage-area boundaries are the same as the WBD boundaries for both streamgages, except at the areas near the streamgages.

Imperfect NEM-with-WBD CAC scores (less than 100) were obtained for the comparison of the drainage-area boundaries for the two streamgages. In both cases, the CAC scores can be linked to mismatches between the streamgage drainage-area outlet and the NHD-flowline outlet. Streamgage 01411000 is located at a point that is 62 percent along the length of the flowline, resulting in a CAC score of 96.63. The location of streamgage 01411300 on its matched NHD flowline is 29 percent along the length of the flowline, resulting in a higher CAC score of 98.01. The mean CAC score is 97.32.

The mean normalized CAC score for the NEM-without-WBD is 95.88. For both streamgages, the delineations of NEM-without-WBD boundaries are acceptable on the basis of the limited detail of the contour intervals depicted on the DRGs. Differences in the delineations were observed, but no major errors could be identified in either dataset.

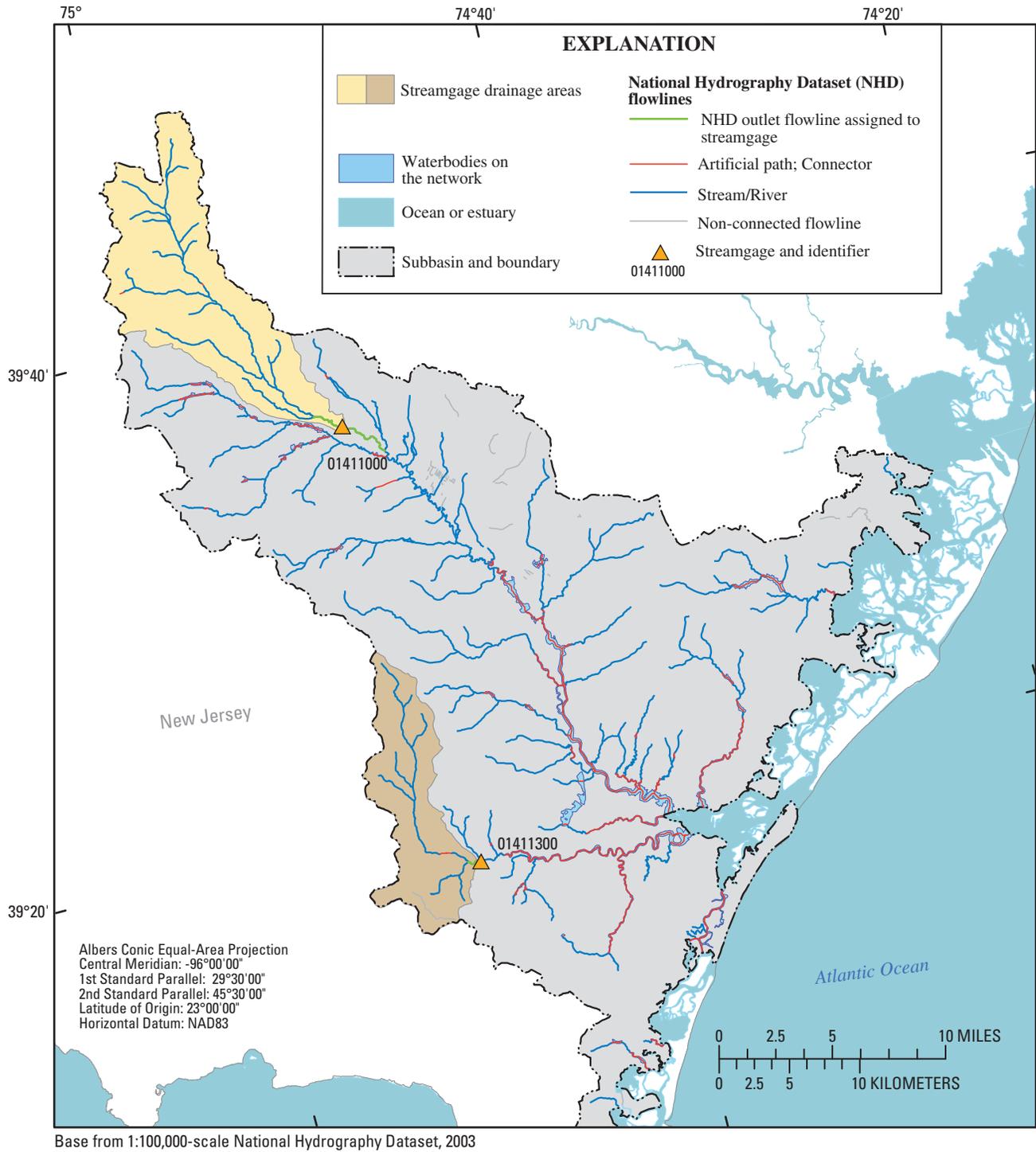
The Raster Seeding Method boundaries for the streamgages are nearly identical with the NEM-without-WBD boundaries in most areas. Boundaries running parallel to streams at the matched flowline for both streamgages were observed in the Raster Seeding Method. The mean normalized CAC score for the Raster Seeding Method of 95.66 is close to the NEM-without-WBD score (95.88).

The mean CAC score for the Outlet Matching Method is 90.88. An unusual Outlet Matching Method assignment for streamgage 01411000 is shown in figure 1–6A. In figure 1–6A, two separate drainage areas were assigned as contributing areas to the streamgage. The improper assignment of the contributing area for streamgage 01411000 reflects the errors in the DEM-derived stream network and shows the limitations to the Outlet Matching Method. Errors unique to the Outlet Matching Method were observed, such as the example presented in figure 1–6B, where the delineation error occurs at the drainage-area outlet for streamgage 01411300. The correct delineation for the streamgage in figure 1–6B is depicted in the streamgage drainage-area boundary used as verification.

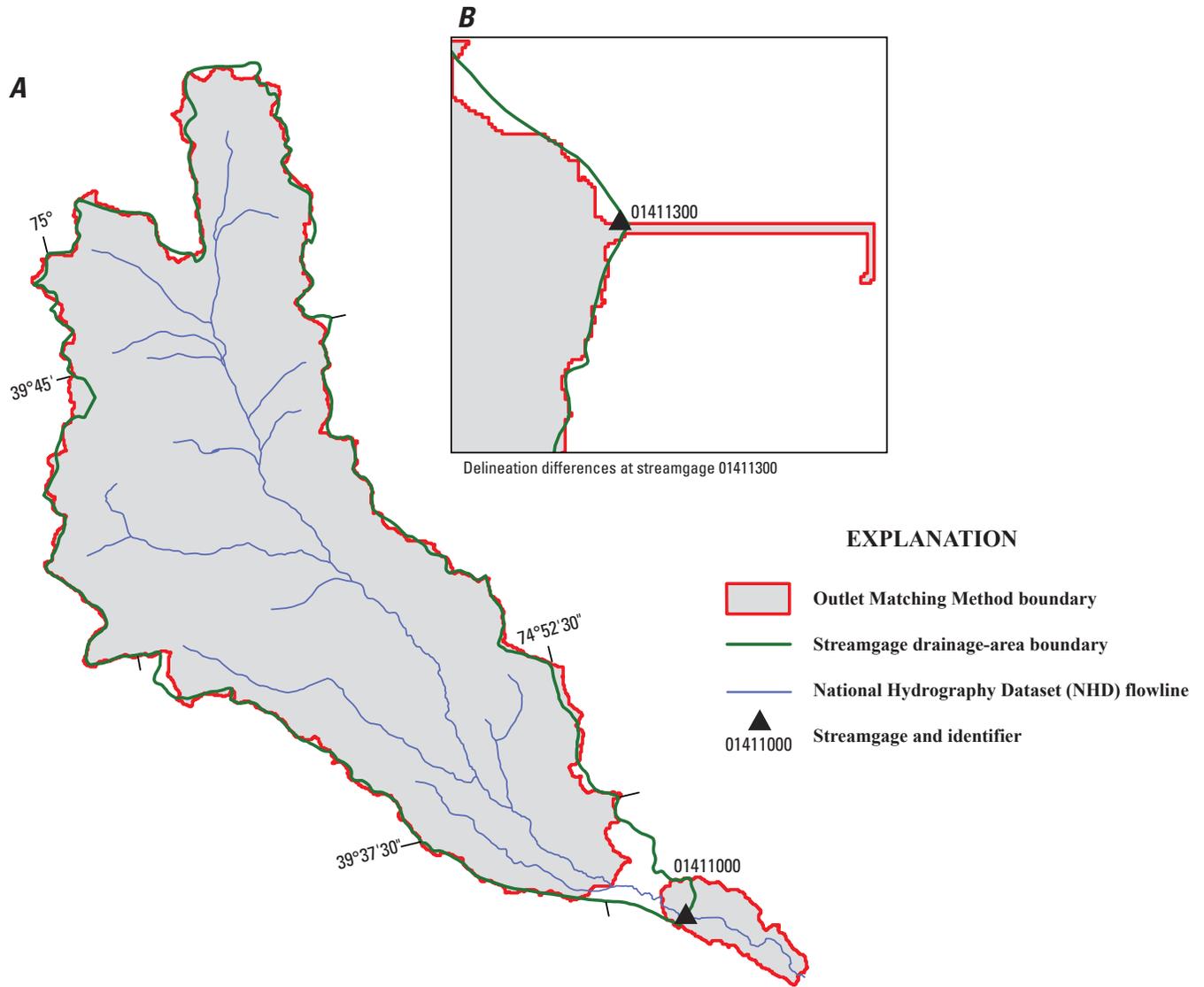
The Thiessen Polygon Method produced the lowest normalized CAC score (79.89) in the GIS comparison analysis.

## Summary of Results for Great Egg Harbor

The normalized CAC scores show that the NEM-with-WBD provided the most accurate results, followed by the NEM-without-WBD. The lowest CAC score of the NEM-without-WBD shows that the WBD boundary is correct and the DEM delineation of the NEM-without-WBD is in error. Mean cumulative drainage-area errors are as high as 8.06 percent (table 1–2) for small drainages of less than 20 mi<sup>2</sup> in size and as low as 0.61 percent for drainage areas greater than 60 to 390 mi<sup>2</sup> in size.



**Figure 1-5.** Location of streamgages and associated drainage areas used as verification datasets, Great Egg Harbor Subbasin, NJ. (The location of this Subbasin is shown in figure 15.)



**Figure 1-6.** (A) The Outlet Matching Method and verification drainage-area boundaries for a streamgage (01411000) and (B) the unusual boundary from the Outlet Matching Method for a streamgage (01411300), Great Egg Harbor Subbasin, NJ. (The location of these streamgages are shown in figure 1-5.)

The mean normalized CAC score of the Raster Seeding Method is 85.8, a close second to the NEM-without-WBD Method, 87.5 (table 1–1). Eighty-one percent of the boundary lines from the Raster Seeding Method correspond to the boundaries from the NEM-without-WBD. Many of the delineation differences of the Raster Seeding Method are due to errors associated with parallel-to-stream boundaries observed from DEM flow-path displacement. For three of the six Raster Seeding Method boundaries with the lowest CAC scores, DEM flow-path displacement had negligible effects. Similar results were obtained for the Raster Seeding Method and the NEM-without-WBD in the cumulative drainage-area analysis.

The Thiessen Polygon Method boundaries are not topographically derived and, therefore, do not share boundaries with the WBD. The Thiessen Polygon Method, according to CAC scores, provided inaccurate results in terms of spatial correspondence to the WBD data, with a mean normalized CAC score of 70.9. The Thiessen Polygon Method also had the highest mean normalized polygon-area error. The Thiessen Polygon Method produced the fourth highest mean error in the cumulative drainage-area analysis, with a mean error as high as 16 percent for the smallest drainage-area size category.

The Outlet Matching Method provided the least accurate results of all the methods, with a mean CAC score of 61.4. The Outlet Matching Method had the highest cumulative drainage-area errors for all drainage-area size categories, with mean errors as high as 18.81 percent for drainage areas less than 20 mi<sup>2</sup> in size (table 1–2). The Outlet Matching Method failed to provide a one-to-one correspondence with the NHD flowlines used in the analysis.

In the streamgage comparison, all the methods except the Thiessen Polygon Method had normalized CAC scores greater than 90. The Thiessen Polygon Method mean normalized CAC score is 70.9. The Outlet Matching Method mean CAC score was not normalized and has an uncorrected CAC of 90.88. The Outlet Matching Method provided erroneous and unusual results (fig. 1–6) and, thus, provided the overall least accurate results in the comparison.

## Seneca River, NC and SC (03060101)

### Comparison of Method Boundaries with WBD Boundaries

Previously determined draft WBD Subwatershed boundaries for the Seneca River Subbasin (fig. 1–7), delineated to the 12-digit hydrologic unit level, were used in the CAC and cumulative drainage-area and convergence analyses. Thirty-three WBD Subwatersheds were used in the CAC analysis, the results of which are shown in figure 1–8 and table 1–3. Results of the cumulative drainage-area convergence analysis are summarized in table 1–4.

### New England Method with WBD (NEM-with-WBD)

WBD/NHD-flowline outlet mismatches and an isolated NHD stream network contributed to the CAC scores (before normalizing to 100) of the NEM-with-WBD. The Seneca River Subbasin's outlet Subwatershed has the lowest CAC score at 76.4 caused by the WBD/NHD-flowline outlet mismatch. The Seneca River flows into a man-made reservoir, Lake Hartwell. The WBD boundary subdivides the reservoir at the pre-flooded main flow path of the Seneca River at the convergence with the Tugaloo River, which together comprise the headwaters of the Savannah River. The NHD artificial path outlet flowline does not follow the original flow path of the Seneca River and drains into Lake Hartwell at another location upstream. WBD/NHD-flowline outlet mismatches also occurred at other Subwatersheds.

An isolated NHD stream network in a headwater Subwatershed lowered the CAC score for the NEM-with-WBD boundary to 93.0 with a polygon-area error of 6.77 percent. At the Subbasin flowline outlet, the effect of the WBD/NHD-flowline outlet mismatch and the isolated stream network was minimal in the cumulative drainage-area analysis, with a 0.89 percent error for this 956-mi<sup>2</sup> Subbasin.

### New England Method without WBD (NEM-without-WBD)

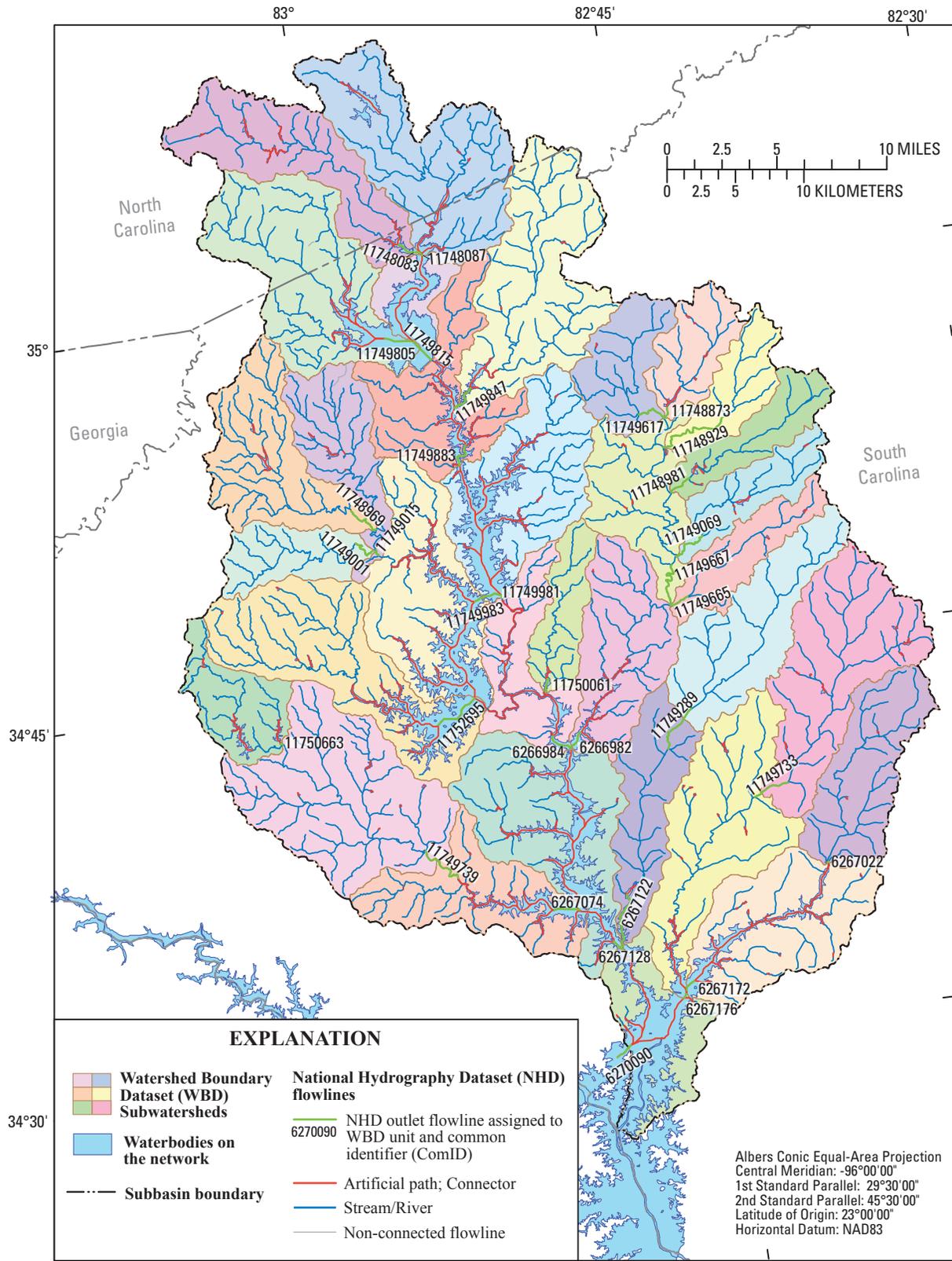
The mean normalized CAC for the NEM-without-WBD is 97.4, the highest score after NEM-with-WBD (see table 1–3). There were no normalized CAC scores less than 90 or normalized polygon-area errors greater than 2.01 percent. The mean normalized polygon-area error for the NEM-without-WBD is 0.56 percent. No major boundary differences were observed between the NEM-without-WBD boundaries and the WBD Subwatersheds. Where small differences occurred, the WBD provided a more accurate boundary than the NEM-without-WBD.

The cumulative drainage-area convergence analysis results summarized in table 1–4 show that the NEM-without-WBD provided the most accurate results (after NEM-with-WBD) for all drainage-area size categories. NEM-without-WBD was the only method besides NEM-with-WBD that produced drainage-area errors of less than 1 percent for drainage areas greater than 60 mi<sup>2</sup> in size. No cumulative drainage-area errors are greater than 6.20 percent. The highest cumulative drainage-area error of 6.20 percent is due mostly to an isolated stream network.

### Raster Seeding Method

The mean normalized CAC score for the Raster Seeding Method is 91.7. Twenty-six of the 33 boundaries compared (79 percent) have normalized CAC scores greater than 90, and 4 boundaries (12 percent) have CAC scores less than 80. The four CAC scores less than 80 are 60.6, 60.8, 72.9, and 79.1.

The Raster Seeding Method has the highest mean normalized polygon-area error at 5.75 percent (compared to: Thiessen Polygon Method, 5.47 percent; NEM-without-WBD,



**Figure 1-7.** The Watershed Boundary Dataset (WBD) Subwatershed boundaries, and National Hydrography Dataset (NHD) flowlines, Seneca River Subbasin, NC and SC. (The location of this Subbasin is shown in figure 15.)

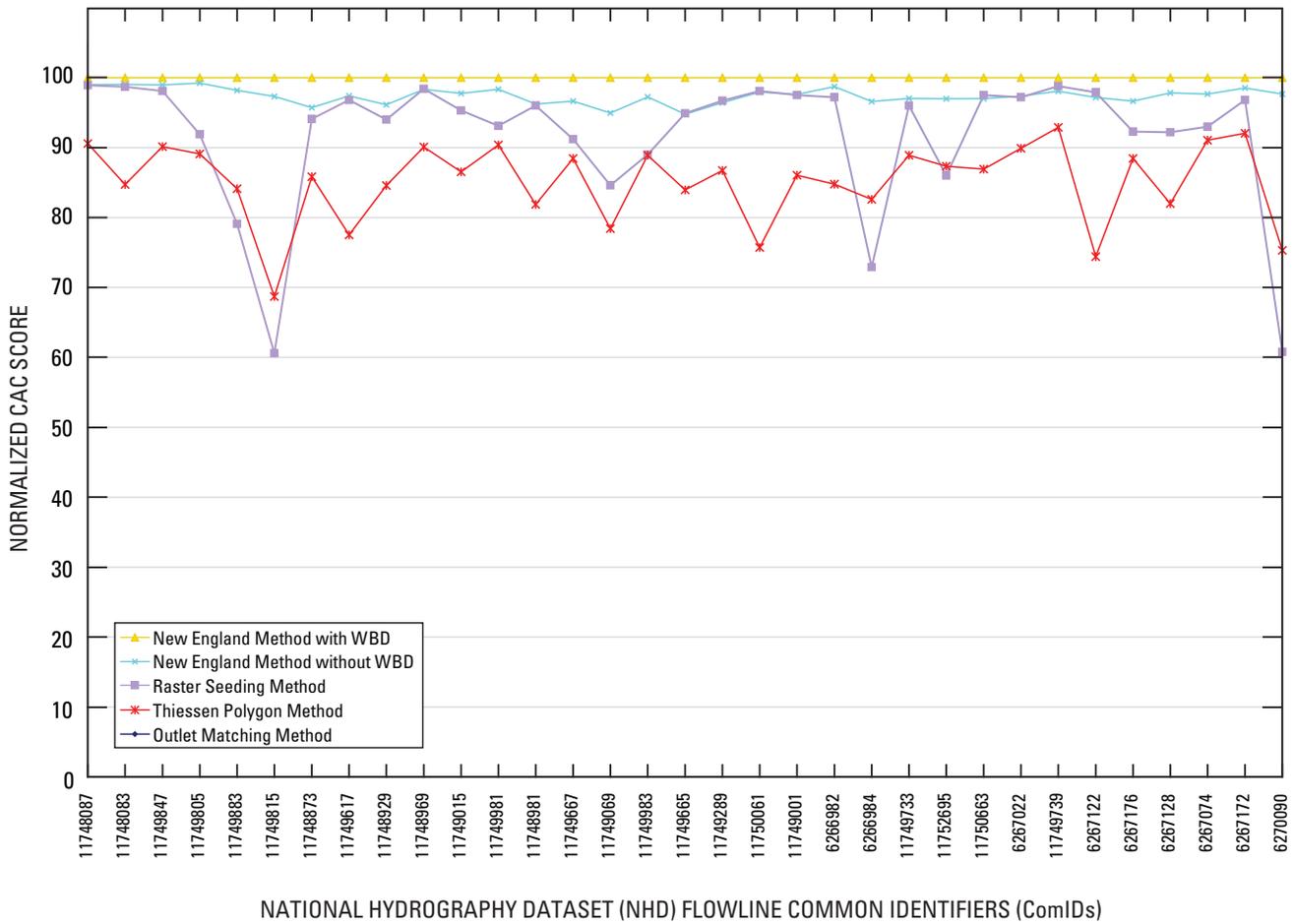


Figure 1–8. Normalized Coefficient of Areal Correspondence (CAC) scores, by National Hydrography Dataset (NHD) flowline common identifier, for all methods tested, Seneca River Subbasin, NC and SC.

Table 1–3. Mean normalized Coefficient of Areal Correspondence score by delineation methods for Watershed Boundary Dataset Subwatersheds in the Seneca River Subbasin, NC and SC.

[CAC, Coefficient of Areal Correspondence; NEM, New England Method; WBD, Watershed Boundary Dataset]

Method	Mean normalized CAC	Normalized CACs greater than 90		Normalized CACs less than 80	
		Number of Subwatersheds	Percentage of Subwatersheds	Number of Subwatersheds	Percentage of Subwatersheds
NEM-with-WBD	100.0	33	100	0	0
NEM-without-WBD	97.4	33	100	0	0
Raster Seeding Method	91.7	26	79	4	12
Thiessen Polygon Method	85.1	7	21	6	18

**Table 1–4.** Mean cumulative drainage-area errors for all delineation methods for various drainage-area sizes, Seneca River Subbasin, NC and SC.[mi<sup>2</sup>, square miles; NEM, New England Method; WBD, Watershed Boundary Dataset]

Method (44 units total)	Mean cumulative drainage-area error, in percent					
	Drainage-area size category					
	All drainage areas	Less than 20 mi <sup>2</sup>	20–60 mi <sup>2</sup>	Greater than 60–390 mi <sup>2</sup>	Greater than 390 mi <sup>2</sup>	At Subbasin outlet 956 mi <sup>2</sup>
NEM-with-WBD	0.56	0.00	1.01	0.57	0.53	0.89
NEM-without-WBD	1.04	1.39	1.36	0.86	0.54	0.85
Raster Seeding Method	2.35	2.45	2.80	2.60	0.89	1.70
Thiessen Polygon Method	2.94	6.18	3.31	1.89	0.75	1.01

0.56 percent; and NEM-with-WBD, 0 percent). The normalized polygon-area errors for the lowest CAC scores in the Raster Seeding Method boundaries range from 19.5 to 39.9 percent. The Raster Seeding Method has the highest normalized polygon-area error of 39.9 percent compared to 29.5 percent for the Thiessen Polygon Method and 2.01 percent for the NEM-without-WBD.

There was a 69-percent spatial correspondence between the drainage-area divides of the Raster Seeding Method catchment delineations and those of the NEM-without-WBD, as determined by GIS overlay analysis. Differences in the spatial delineations between the Raster Seeding Method and NEM-with-WBD are directly attributed to DEM flow-path displacement errors that occurred in the Raster Seeding Method delineations. In all cases where DEM flow-path displacement errors occur, the effects of such errors had negligible effects. DEM flow-path displacement was most problematic for flowlines in reservoirs; this was observed for the boundaries with the four lowest CAC scores for the Raster Seeding Method.

The highest Raster Seeding Method cumulative drainage-area errors were observed in headwater Subwatersheds. A 12.4 percent cumulative drainage-area error was observed for a 43.6-mi<sup>2</sup> Subwatershed. In this Subwatershed, about half of the error is linked to an isolated NHD network, whereas the other half of the error is attributed to DEM flow-path displacement. DEM flow-path displacement was the most significant contributor for a 12.2 percent error in a 13.9-mi<sup>2</sup> Subwatershed, and an 11.54 percent error in a 44.4-mi<sup>2</sup> Subwatershed.

The Raster Seeding Method has the third lowest mean error in the cumulative drainage-area analysis at 2.35 percent (table 1–4). No Raster Seeding Method cumulative drainage-area mean errors greater than 2.80 percent were observed for any drainage-area size categories shown in table 1–4.

### Thiessen Polygon Method

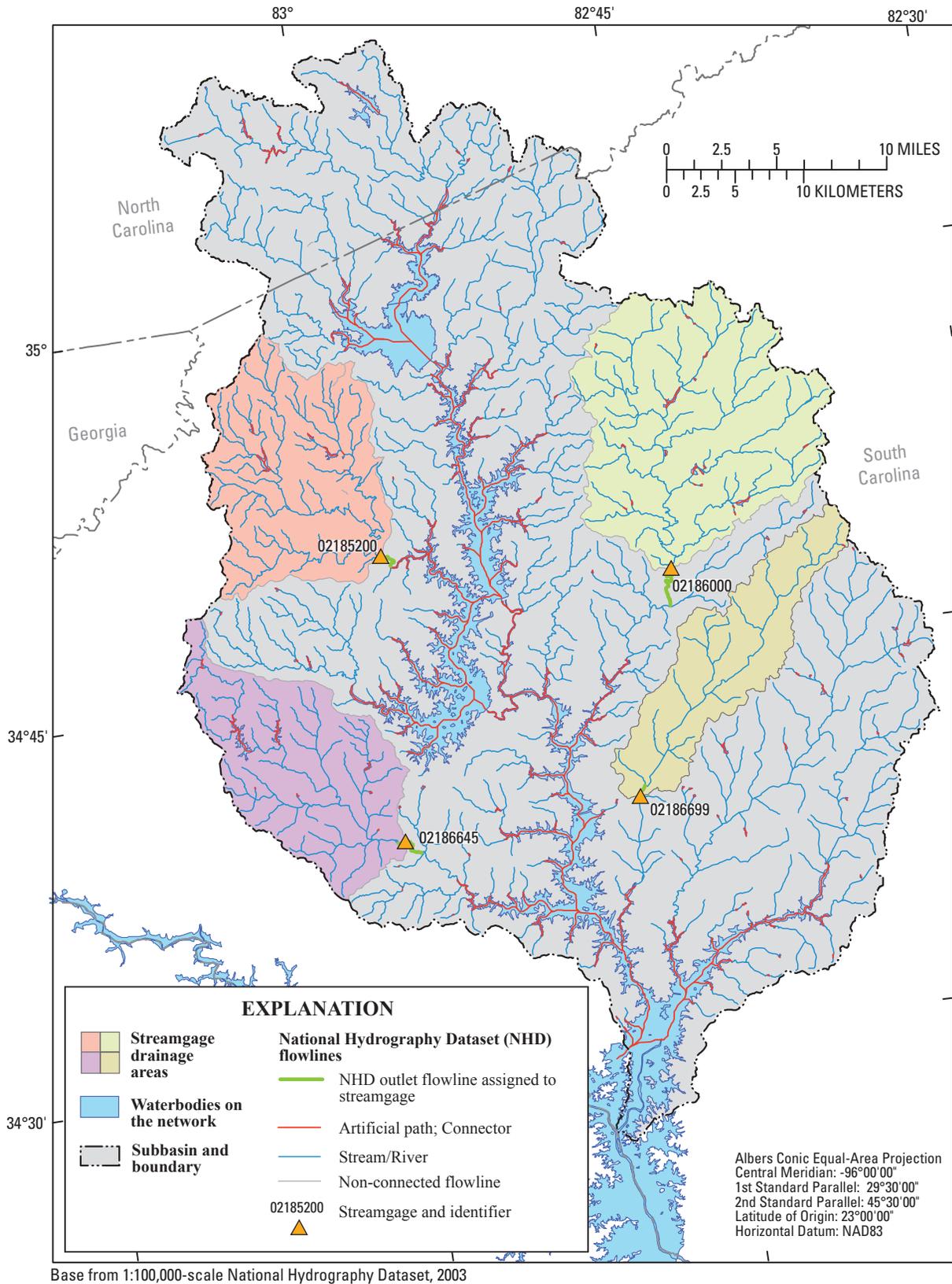
The mean normalized CAC score for the Thiessen Polygon Method is 85.1. Seven of the 33 Thiessen Polygon Method boundaries (21 percent) have normalized CAC scores greater than 90. Six Thiessen Polygon Method boundaries (18 percent) have normalized CAC scores less than 80.

In the cumulative drainage-area analysis, the Thiessen Polygon Method has the highest mean drainage-area error of just 2.94 percent (table 1–4). The Thiessen Polygon Method has the highest mean error of all the methods for drainage areas less than 20 mi<sup>2</sup> in size, 6.18 percent, and for drainage areas 20 to 60 mi<sup>2</sup> in size, 3.31 percent. For drainage areas greater than 60 mi<sup>2</sup>, the Thiessen Polygon Method mean cumulative drainage-area errors are not substantial; the method produced slightly lower mean drainage-area errors than the Raster Seeding Method.

### Comparison of Methods with Streamgage Data

Four streamgage drainage-area boundaries, shown in figure 1–9, were obtained as verification datasets in the Seneca River Subbasin. Drainage-area measurements for the streamgages range from 47 to 104 mi<sup>2</sup>. The boundaries from the streamgage drainage-area data align with those from the WBD Subwatershed boundaries, except at the areas near the streamgages. NWIS drainage-area measurements also were obtained for these four streamgages.

CAC scores less than 100 for the NEM-with-WBD are directly related to the location of the streamgage along the NHD flowline. The farther upstream the streamgage is located along the flowline, the lower the CAC score. The lowest CAC score (not normalized) of the NEM-with-WBD is 97.75, and the mean CAC score of the NEM-with-WBD is 98.39.



**Figure 1–9.** Location of streamgages and associated drainage areas used as verification datasets, Seneca River Subbasin, NC and SC. (The location of this Subbasin is shown in figure 15.)

The mean normalized CAC score for the NEM-without-WBD is 98.32. No major errors were observed in the NEM-without-WBD boundaries in the comparison with the streamgage drainage-area boundaries. The streamgage drainage-area boundaries provided more accurate boundaries than the NEM-without-WBD for many areas where differences were observed. On the basis of the limited contour information of the DRGs, it was difficult to determine which boundary was more accurate for some areas where small differences were observed.

The results of the CAC analysis of the Raster Seeding Method boundaries were nearly the same as those for the NEM-without-WBD boundaries. The mean normalized CAC score for the Raster Seeding Method is 98.14, and for NEM-without-WBD, 98.32. The streamgage drainage-area boundaries from the Raster Seeding Method have a 92 percent spatial correspondence to the boundaries from the NEM-without-WBD. DEM flow-path displacement artifacts in the Raster Seeding Method delineations contributed to the differences between the Raster Seeding Method and the NEM-without-WBD boundaries.

The Thiessen Polygon Method achieved reasonable CAC results for the boundaries of the four streamgage drainage areas. The mean Thiessen Polygon Method normalized CAC score is 91.23. The highest normalized drainage-area error of 3.68 percent occurred for streamgage 02186699; it is the highest error obtained for all the methods.

## Summary of Results for Seneca River

The results of the WBD Subwatershed boundary and streamgage drainage-area boundary comparisons for the NEM-with-WBD were useful in identifying an isolated network, WBD/NHD-flowline outlet mismatches, and NHD-flowline outlet and streamgage location mismatches. The effects of the isolated network in the cumulative drainage-area analysis for the NEM-with-WBD were minimal at the outlet flowline and within the headwater Subwatershed in which the isolated network is located.

The NEM-without-WBD provided the highest normalized CAC scores in the comparisons with WBD Subwatershed and streamgage drainage-area boundaries and the most accurate areas in the cumulative drainage-area analysis (after the NEM-with-WBD). NEM-without-WBD is the only other method besides NEM-with-WBD where no normalized CAC scores are less than 90. No major delineation errors were identified in the comparison between the NEM-without-WBD boundaries and the WBD boundaries.

DEM flow-path displacement artifacts occurring in the Raster Seeding Method contributed to the delineation differences between the Raster Seeding Method and the NEM-without-WBD boundaries. The Raster Seeding Method is the third best method, using CAC analysis with the WBD Subwatersheds and streamgage drainage areas. The Raster Seeding Method has the highest mean cumulative drainage-area errors for the larger drainage-area size categories; most

errors are attributed to the DEM flow-path displacement artifacts. The Raster Seeding Method also has the highest mean normalized Subwatershed polygon-area error and the highest polygon-area error of all the methods. DEM flow-path displacement related errors were most problematic in waterbodies. No major errors were identified in the streamgage drainage-area comparisons; however, DEM flow-path displacement errors were observed.

The Thiessen Polygon Method was more accurate for this Subbasin than for the other study Subbasins. The mean normalized CAC score is 85.1 for the WBD comparison and 91.23 for the streamgage analysis. The Thiessen Polygon Method has the highest mean cumulative drainage-area error at just 2.94 percent. Thiessen Polygon cumulative drainage-area errors are highest in the two smaller drainage-area size categories shown in table 1–4. The highest mean error, 6.18 percent, was observed for the Thiessen Polygon Method for drainage areas less than 20 mi<sup>2</sup> in size. In the streamgage analysis, the Thiessen Polygon Method has the highest normalized drainage-area error of all the methods at 3.68 percent.

## Pine River, MN (07010105)

### Comparison of Method Boundaries with WBD Boundaries

Draft WBD Subwatershed boundaries for the Pine River Subbasin, MN (fig. 1–10), which were delineated to the 12-digit hydrologic unit level, were used in the CAC and cumulative drainage-area analyses. Twenty-eight WBD boundaries were used in the CAC analysis; results are shown in figure 1–11 and table 1–5. Cumulative drainage-area convergence analysis results are summarized in table 1–6. A graph developed as part of the convergence analysis, displaying the distribution of drainage-area errors is shown in figure 1–12.

### New England Method with WBD (NEM-with-WBD)

WBD/NHD-flowline outlet mismatches and other NHD issues contributed to the lowering of the CAC scores (before normalizing to 100) of the NEM-with-WBD. NHD issues include three isolated stream networks, an improper NHD flow direction, and an NHD stream incorrectly crossing a WBD divide. WBD/NHD-flowline outlet mismatches, along with the other NHD issues, also led to errors in the cumulative drainage-area convergence analysis results shown in table 1–6.

Three isolated NHD stream networks account for the majority of the cumulative drainage-area error computed for the outlet flowline of the Subbasin. The three isolated stream network drainage areas were not included in the catchment drainage-area accumulation for the outlet flowline of the Subbasin, thereby introducing an error of 0.61 percent. The drainage areas of the three isolated networks are 2.6, 1.8, and 0.4 mi<sup>2</sup> in size, totaling 4.8 mi<sup>2</sup>. One isolated network,

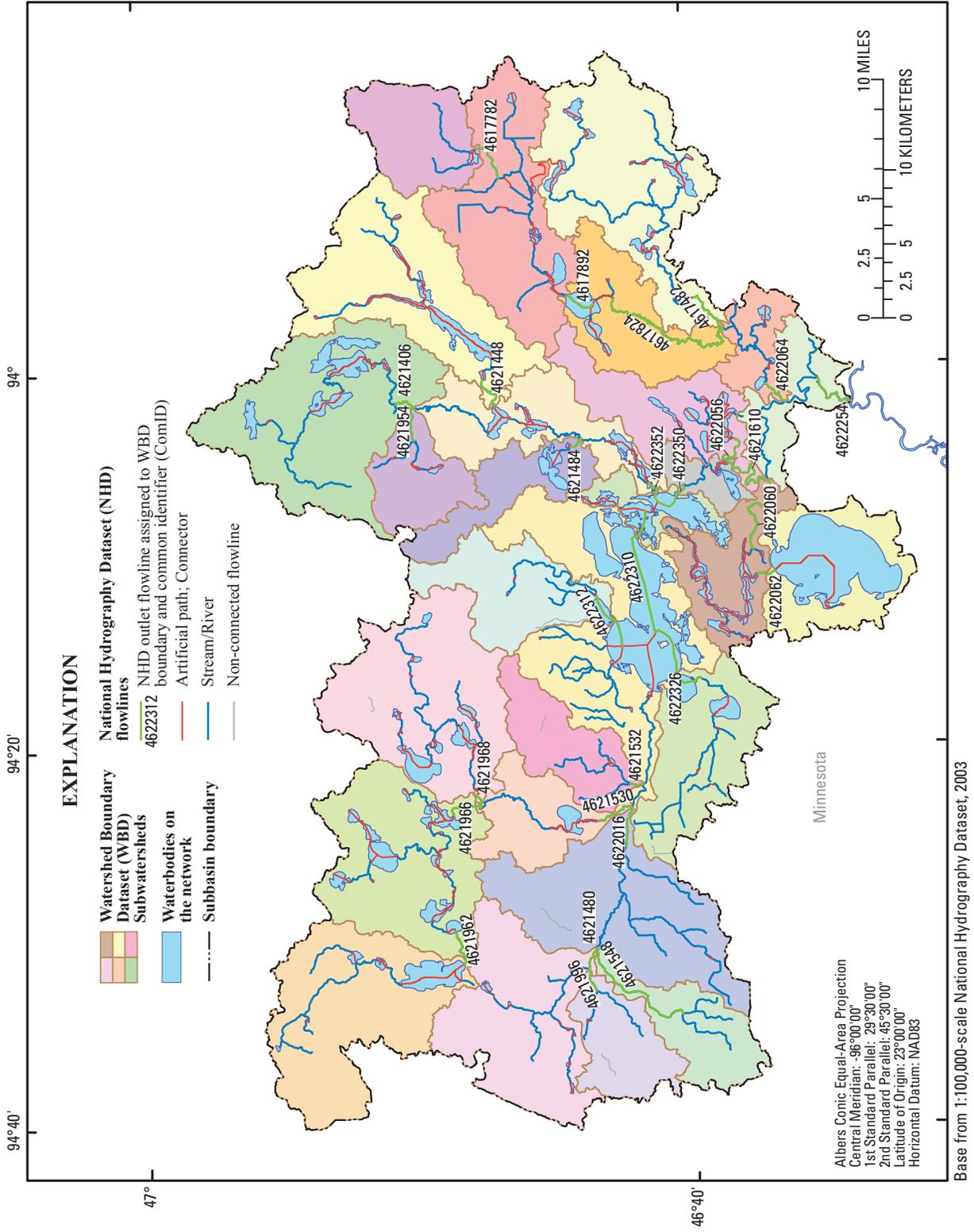
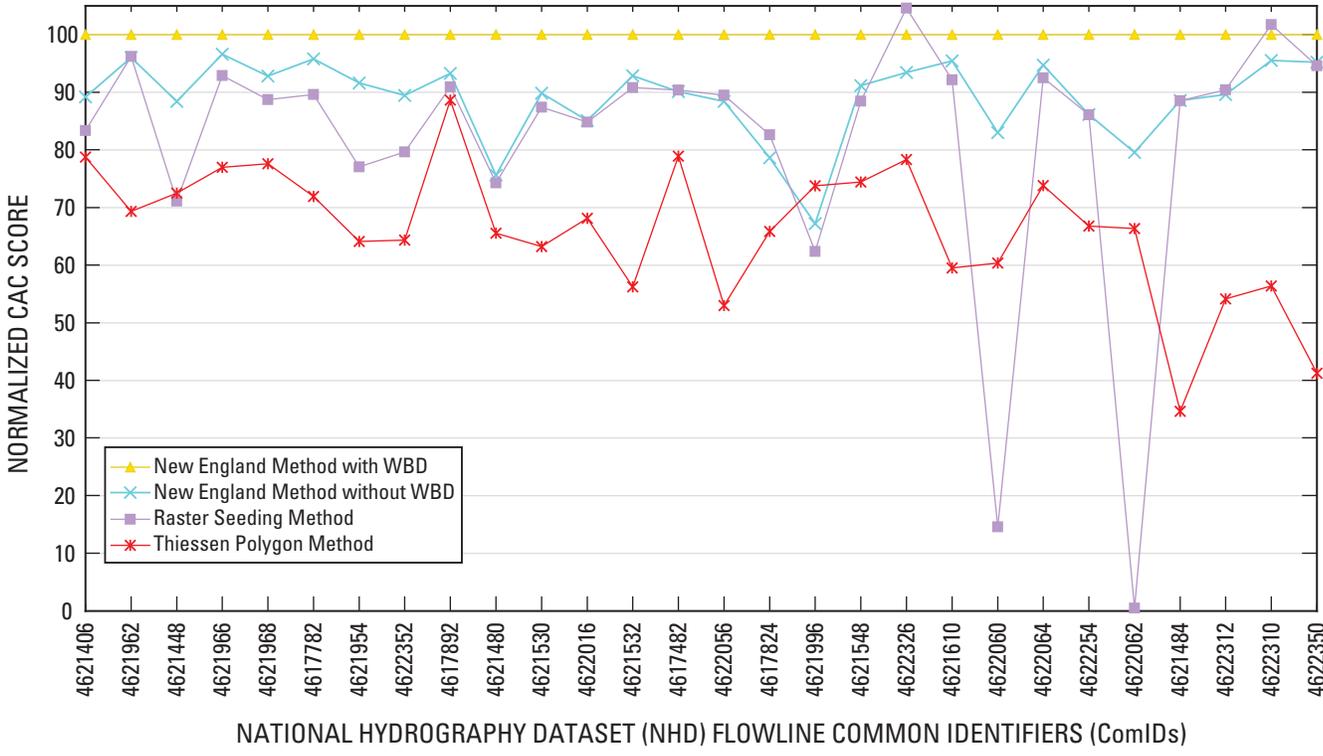


Figure 1-10. The draft Watershed Boundary Dataset (WBD) Subwatershed boundaries, and National Hydrography Dataset (NHD) flowlines, Pine River Subbasin, MN. (The location of this Subbasin is shown in figure 15.)



**Figure 1-11.** Normalized Coefficient of Areal Correspondence (CAC) scores, by National Hydrography Dataset (NHD) flowline common identifier, for all methods tested, Pine River Subbasin, MN.

**Table 1-5.** Mean normalized Coefficient of Areal Correspondence score by delineation methods for Watershed Boundary Dataset Subwatersheds in the Pine River Subbasin, MN.

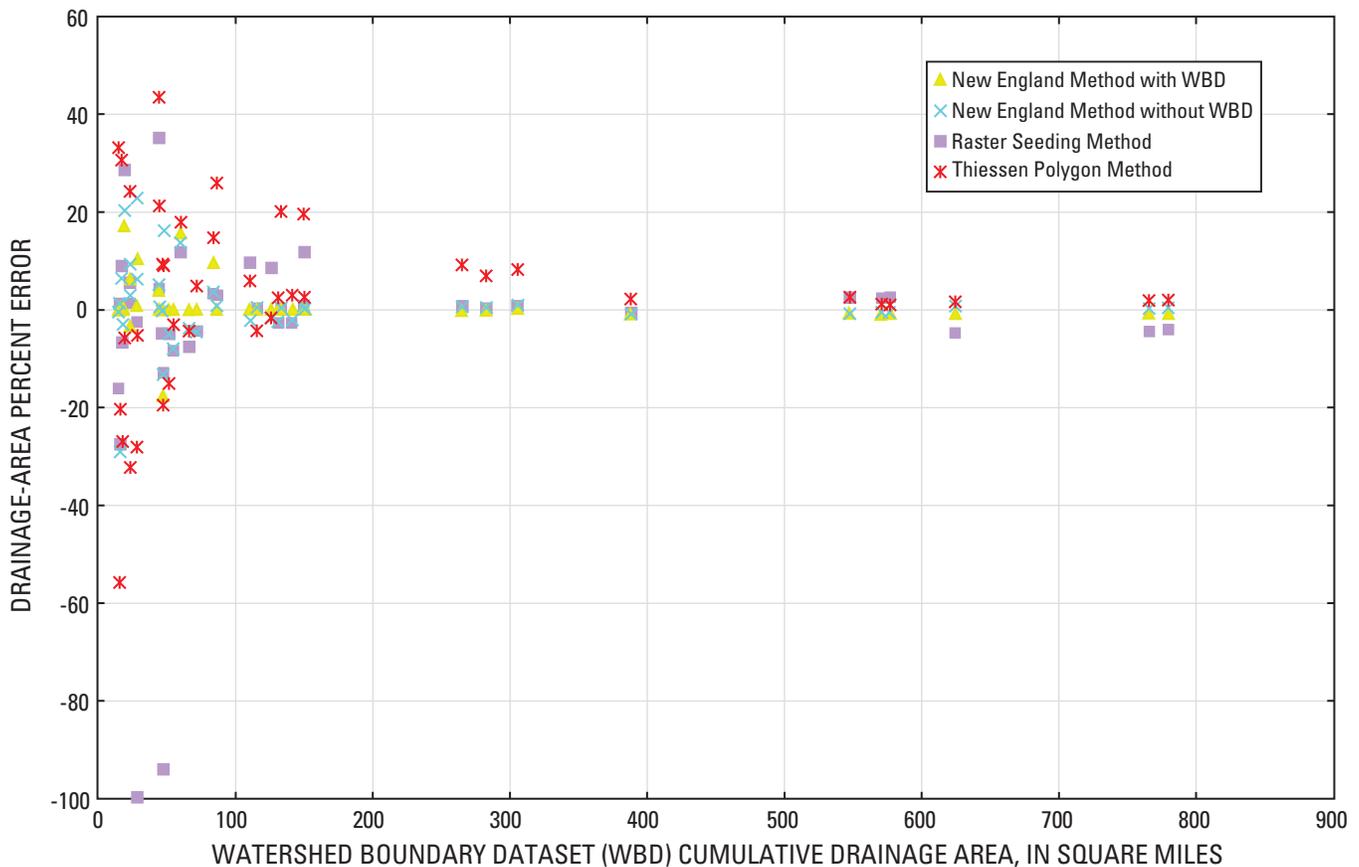
[CAC, Coefficient of Areal Correspondence; NEM, New England Method; WBD, Watershed Boundary Dataset]

Method	Mean normalized CAC	Normalized CACs greater than 90		Normalized CACs less than 80	
		Number of Subwatersheds	Percentage of Subwatersheds	Number of Subwatersheds	Percentage of Subwatersheds
NEM-with-WBD	100.0	28	100	0	0
NEM-without-WBD	89.1	14	50	4	14
Raster Seeding Method	81.7	11	39	7	25
Thiessen Polygon Method	66.3	0	0	27	96

**Table 1-6.** Mean cumulative drainage-area errors for all delineation methods for various drainage-area sizes, Pine River Subbasin, MN.

[mi<sup>2</sup>, square miles; NEM, New England Method; WBD, Watershed Boundary Dataset]

Method (40 units total)	Mean drainage-area error, in percent					
	Drainage-area size category					
	All drainage areas	Less than 20 mi <sup>2</sup>	20–60 mi <sup>2</sup>	Greater than 60–390 mi <sup>2</sup>	Greater than 390 mi <sup>2</sup>	At Subbasin outlet 780 mi <sup>2</sup>
NEM-with-WBD	2.29	2.99	3.86	1.59	0.69	0.61
NEM-without-WBD	4.84	10.10	8.13	2.37	0.52	0.40
Raster Seeding Method	11.32	14.90	24.79	4.13	3.46	3.91
Thiessen Polygon Method	13.69	28.76	19.14	9.07	1.73	2.02



**Figure 1-12.** Drainage-area percent error in relation to Watershed Boundary Dataset cumulative drainage areas, for all methods tested, Pine River Subbasin, MN.

located within a headwater WBD Subwatershed, caused a 10.6 percent cumulative drainage-area error for that Subwatershed. An improper NHD flow relationship in the flowline network caused cumulative drainage-area errors of as much as 17.38 percent.

### New England Method without WBD (NEM-without-WBD)

The mean normalized CAC score for the NEM-without-WBD was 89.1, the highest score after NEM-with-WBD (table 1–5). Data show that 14 of the 28 NEM-without-WBD boundaries (50 percent) have normalized CAC scores greater than 90 (fig. 1–11 and table 5), the highest normalized CAC scores after the NEM-with-WBD. Just four boundaries, or 14 percent, have normalized CAC scores less than 80 (the lowest after NEM-with-WBD). The four lowest normalized CAC scores for the NEM-without-WBD are 67.2, 75.6, 78.7, and 79.6. The NEM-without-WBD had the smallest mean normalized polygon-area error, after the NEM-with-WBD, of 5.7 percent, compared to 15.44 and 21.95 percent for the Raster Seeding and Thiessen Polygon Methods.

Results of the cumulative drainage-area analysis summarized in table 1–6 show that the NEM-without-WBD provided the most accurate drainage areas (after NEM-with-WBD) for all drainage-area size categories. The NEM-without-WBD was the only method besides NEM-with-WBD that produced drainage-area errors of less than 1 percent for drainage areas greater than 390 mi<sup>2</sup>.

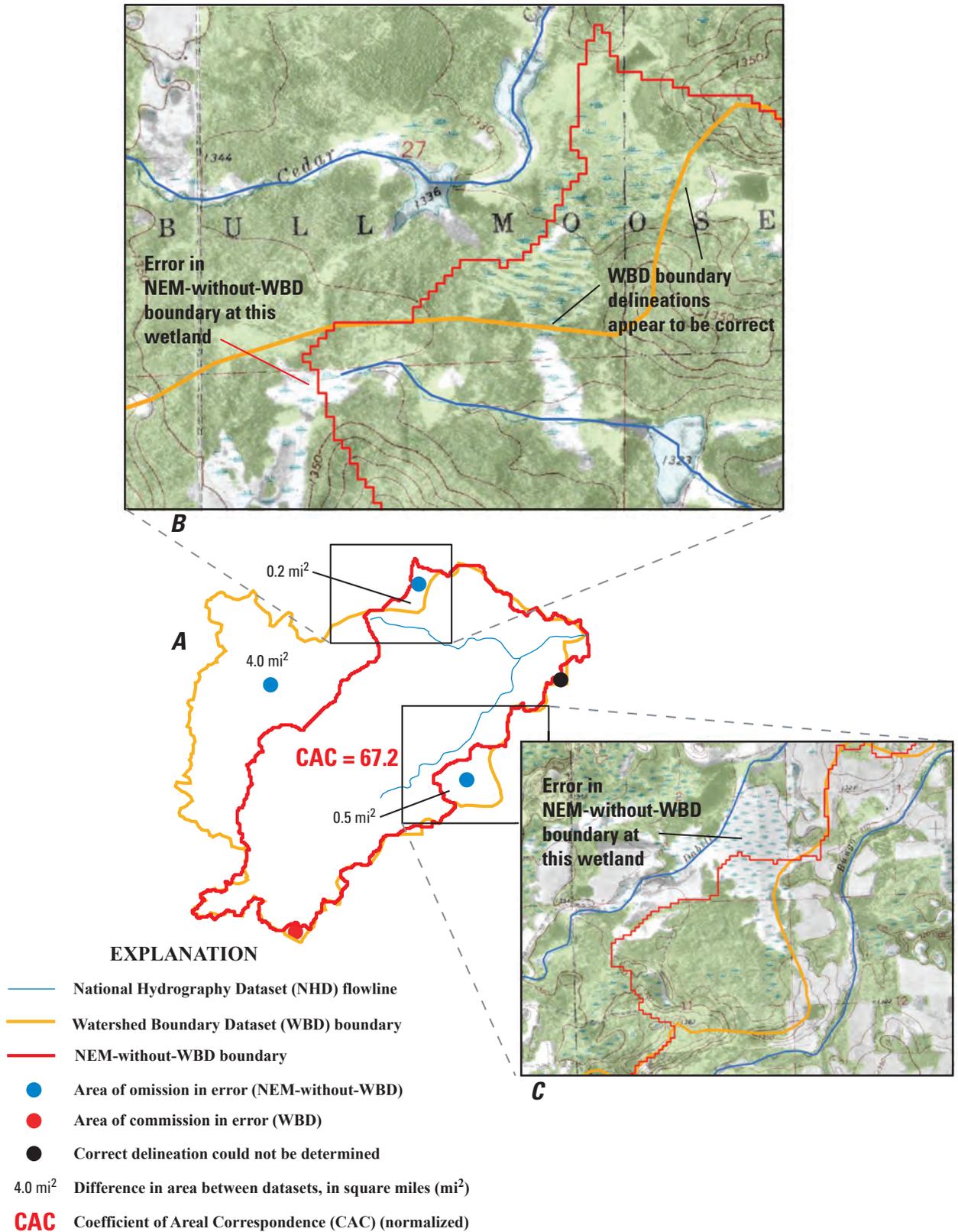
The WBD Subwatershed data provided a more accurate boundary for the compared boundary of the NEM-without-WBD, which has the lowest normalized CAC score for its method, at 67.2. The NEM-without-WBD boundary and the corresponding WBD boundary are shown in figure 1–13. The most noticeable areas of difference are shown in figure 1–13A, indicated by the colored point symbols inside each area in error. The colors designate which delineation method provided more accurate results. A 4-mi<sup>2</sup> area draining into a wetland was the area with the most obvious difference. Information from the DRG and DOQ for this area confirms that the WBD inclusion of this wetland and its associated drainage area is correct (fig. 1–13B). The DEM-based delineation of the NEM-without-WBD improperly assigns the wetland-drainage area to an adjacent flowline. A smaller wetland-drainage area (0.2 mi<sup>2</sup>) was incorrectly included in the NEM-without-WBD boundary (fig. 1–13B). The NEM-without-WBD delineation also was wrong for another wetland drainage area identified in figure 1–13C. Several additional boundary discrepancies affected much smaller areas; however, using the DRG and DOQ, it was not possible to determine which method provided the most accurate boundaries. This Subwatershed also has the highest cumulative drainage-area error of the NEM-without-WBD at 29.02 percent.

The second lowest normalized CAC score for the NEM-without-WBD is 75.6. The WBD Subwatershed boundary and the NEM-without-WBD boundary are shown in figure 1–14A. The contour information depicted on the DRGs was difficult to

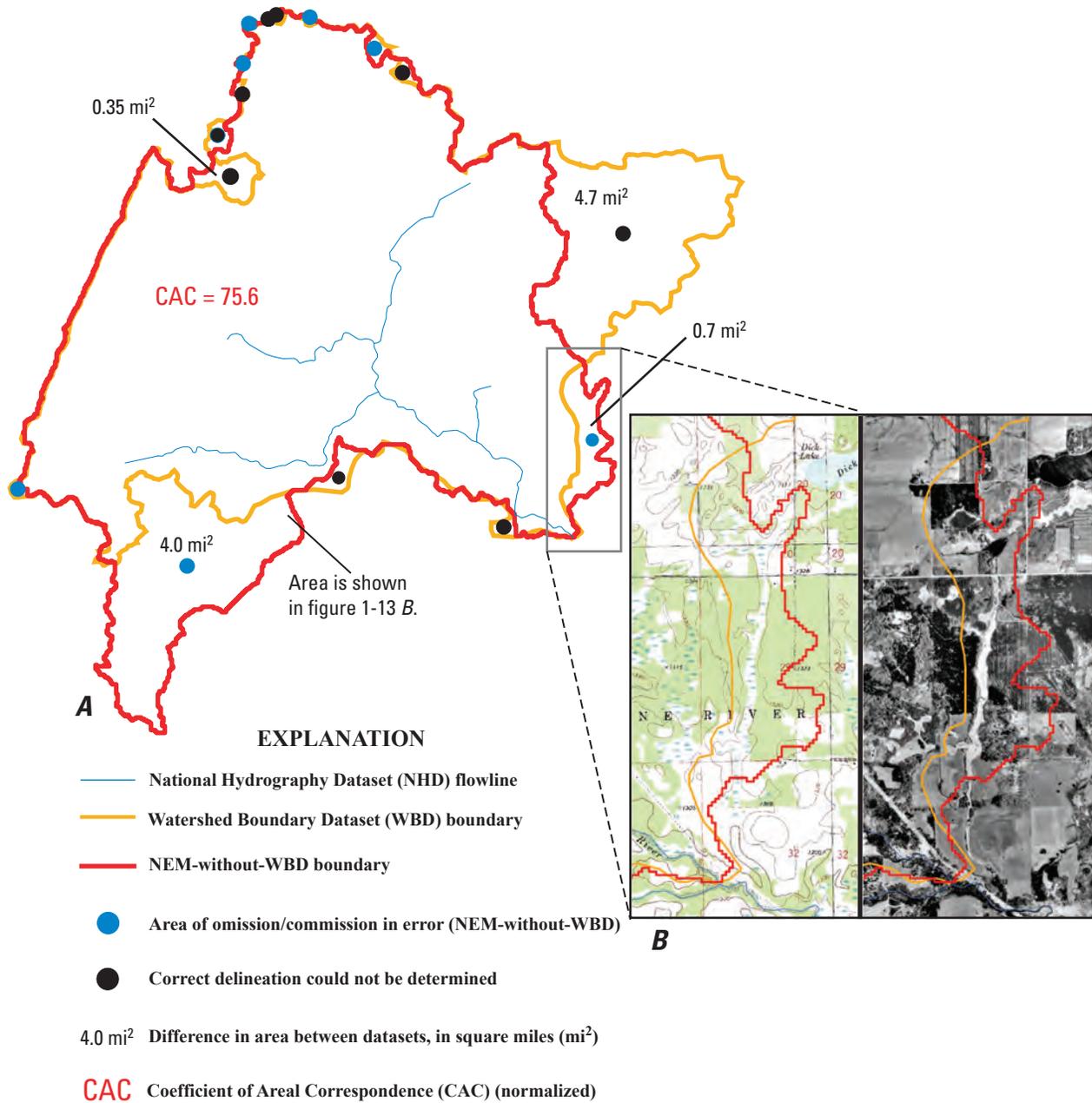
interpret when determining which delineation was correct for the largest area of difference, 4.7 mi<sup>2</sup>. The second largest area of difference is the same 4-mi<sup>2</sup> area identified in figure 1–13 of the adjacent Subwatershed. Again, the WBD delineation provided the more accurate boundary delineation. The third largest area of difference is a 0.7-mi<sup>2</sup> wetland area that was incorrectly included in the NEM-without-WBD delineation. Information from the DRG and DOQ depicted in figure 1–14B shows that the WBD boundary is correct. The wetland drainage area included in the NEM-without-WBD boundary clearly drains to the downstream adjacent Subwatershed. Several other delineation differences between the WBD and NEM-without-WBD affect much smaller areas. In many cases, it was not possible on the basis of the limited detail of the contours to determine which method is more accurate. In other areas of less significance, the WBD Subwatershed provided a more accurate boundary, based on the contour information of the DRGs. In one case, the areas of difference canceled each other out, resulting in a normalized polygon-area error of only 0.95 percent; however, the effects of these delineation differences are reflected in the cumulative drainage-area analysis with an error of 7.88 percent for this 55-mi<sup>2</sup> drainage area.

The NEM-without-WBD area with the third lowest normalized CAC score of 78.7 (fig. 1–15) has a normalized polygon-area error of 17.5 percent. The largest area of difference is a 1.94-mi<sup>2</sup> drainage area for Downey Lake (fig. 1–15B). The NEM-without-WBD boundary does not include the lake or its drainage area, whereas the WBD boundary does. On the basis of the limited information available for the Downey Lake site, a determination of which boundary is more accurate could not be made. Information from the DRGs indicates that no streams are connected to the lake, which would provide an indication of outlet flow direction. Two wetlands abut the lake, but no determination could be made of the primary drainage for these wetlands. The second largest area of difference is a 1.45-mi<sup>2</sup> wetland area (shown in figure 1–15C), which was omitted from the NEM-without-WBD boundary. The WBD boundary for this wetland is correct on the basis of the DRG and DOQ information. The most accurate boundaries for other areas with differences could not be established because of limited or complex contour detail. For other areas where a determination could be made, the WBD provided the most accurate delineation.

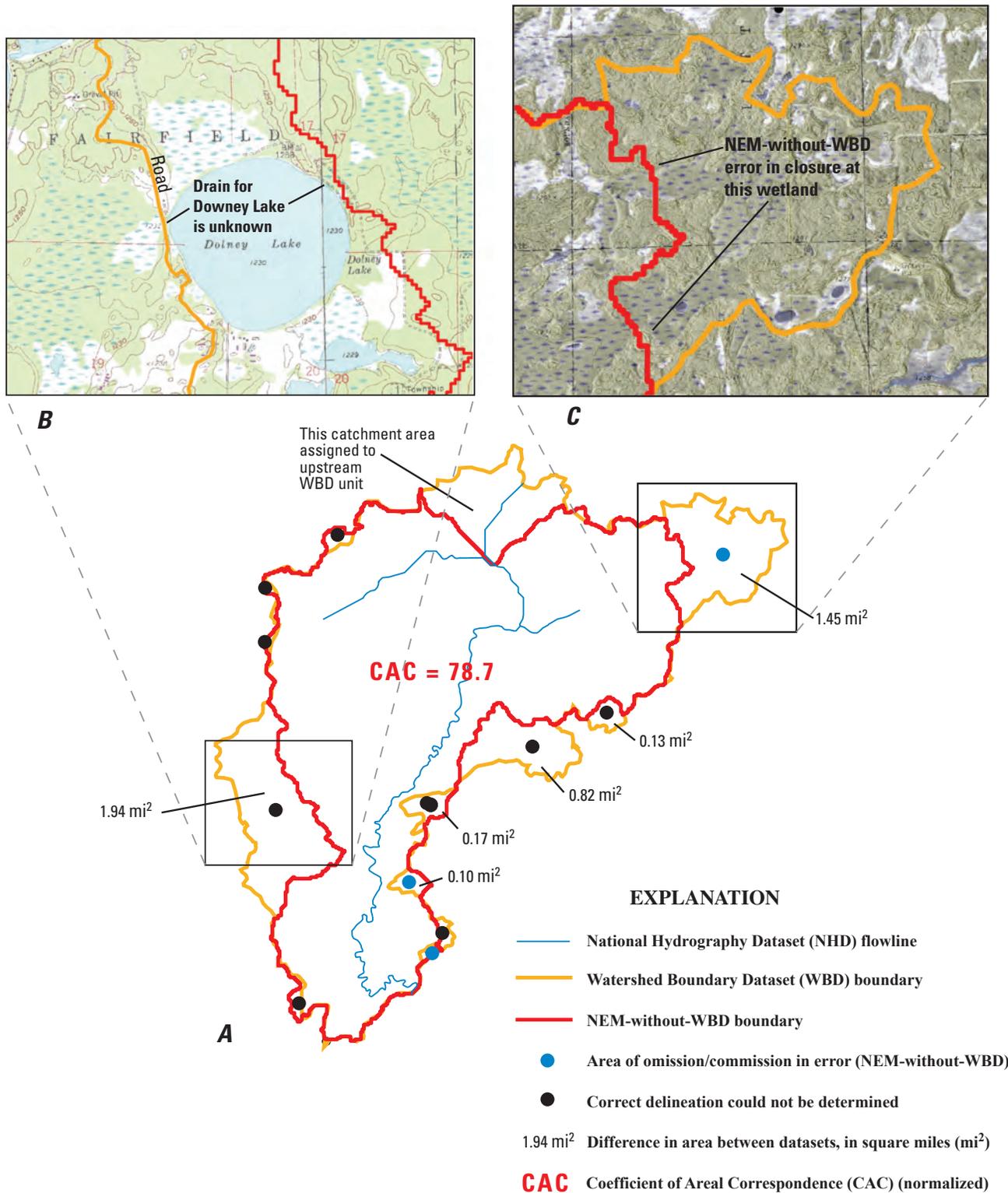
The NEM-without-WBD boundary with the fourth lowest score has a normalized CAC score of 79.6. The largest area of difference (approximately 6 mi<sup>2</sup>) between the NEM-without-WBD and WBD boundaries occurred for the Subwatershed that represents the drainage area for Pelican Lake (fig. 1–16). Drainage areas for three adjacent lakes to the east are in question. The WBD boundary does not include the lakes, whereas the NEM-without-WBD boundary does. The 1:24,000-scale DRGs do not show any outlet drainage for these lakes, and therefore, no definitive conclusion about which delineation is more accurate could be made. It does appear, however, that the manual delineation of the WBD uses a road grade as a



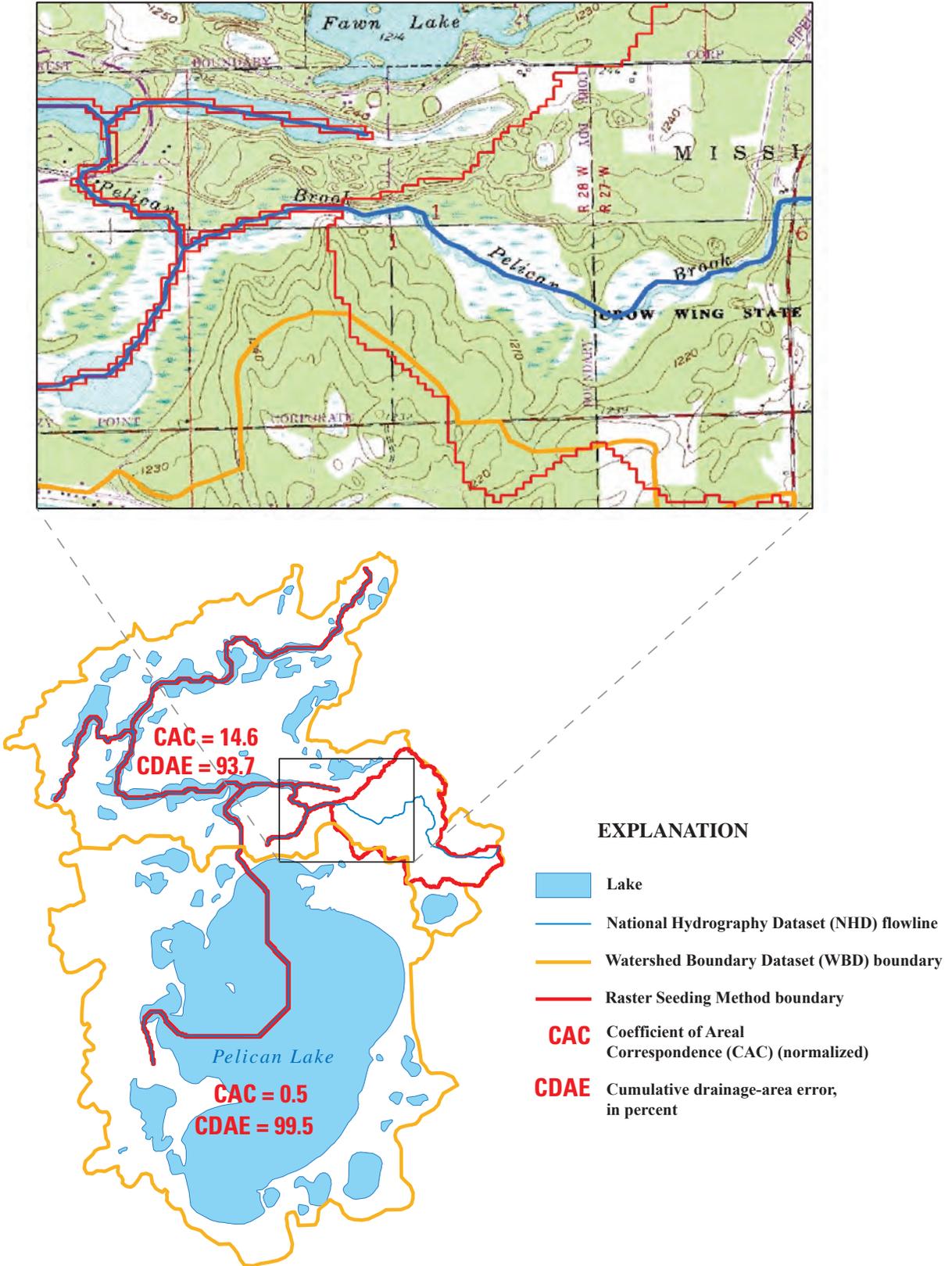
**Figure 1-13.** (A) Watershed Boundary Dataset (WBD) and New England Method without WBD (NEM-without-WBD) boundaries, and (B) and (C) areas of detail showing boundary differences at wetlands, Pine River Subbasin, MN. (The location of this WBD boundary is shown in figure 1-10.)



**Figure 1–14.** (A) Watershed Boundary Dataset (WBD) and New England Method without WBD (NEM-without-WBD) boundaries, (B) area of detail where the boundaries differ overlaid on a Digital Raster Graphic, and (C) area in detail with boundary error overlaid on a Digital Orthophoto Quadrangle, Pine River Subbasin, MN. (The location of this WBD boundary is shown in figure 1–10.)



**Figure 1-15.** (A) Watershed Boundary Dataset (WBD) and New England Method without WBD (NEM-without-WBD) boundaries, (B) area of detail, showing boundary differences at Downey Lake, and (C) area of detail, showing boundary differences at a wetland, Pine River Subbasin, MN. (The location of this WBD boundary is shown in figure 1-10.)



**Figure 1-16.** Watershed Boundary Dataset (WBD) and the Raster Seeding Method boundaries for the two Subwatersheds with the lowest normalized Coefficient of Areal Correspondence (CAC) scores, Pine River Subbasin, MN. (The locations of these WBD boundaries are shown in figure 1-10.)

divide to separate these lakes in this area of flat relief. The cumulative drainage-area error for this 28-mi<sup>2</sup> drainage area is 22.92 percent, the second highest error for this method in the cumulative drainage-area analysis.

Overall, it appears that the WBD data enhanced the delineations when used with the NEM-with-WBD. The Subbasin is mostly an area of low relief with many lakes and wetland areas where delineations based on USGS 30-m DEM data tend to be less accurate.

### Raster Seeding Method

The mean normalized CAC score for the Raster Seeding Method was 81.7. Two normalized CAC scores greater than 100 are linked to DEM flow-path displacement at the outlet flowline and WBD/NHD-flowline outlet mismatch. Eleven of the 28 boundaries compared (39 percent) have normalized CAC scores greater than 90, and 7 boundaries (25 percent) have a normalized CAC score less than 80. The seven normalized CAC scores less than 80 are 0.5, 14.6, 62.4, 71.14, 74.28, 77.09, and 79.65.

In the WBD comparison, the mean normalized polygon-area error for the Raster Seeding Method is 15.44 percent, a distant third in percentage following the NEM-without-WBD error of 5.70 percent. The Raster Seeding Method has the highest error, 99.5 percent, compared with the errors from the New England and Thiessen Polygon Methods.

The two boundaries with the lowest Raster Seeding Method normalized CAC scores of 0.5 and 14.6 also have the highest errors in the cumulative drainage-area analysis. Both of these Raster Seeding Method boundaries (shown in figure 1–16) had parallel-to-stream boundaries from DEM flow-path displacement that led to extreme errors in the Raster Seeding Method results. Most of the flowlines within these two boundaries are artificial flow paths or streams located in relatively flat terrain. The parallel stream effect is greatly increased because of the flat surface of the lake or terrain in the DEM. The Raster Seeding Method boundaries for the NHD flowlines in these flat areas resemble a 30-meter buffer surrounding the NHD flowlines, as opposed to correct catchment delineations. For the Pelican Lake area, the true drainage area is 27.96 mi<sup>2</sup>, not the 0.14 mi<sup>2</sup> area produced by the Raster Seeding Method, a cumulative drainage-area error of 99.5 percent. Pelican Lake drains to Pelican Brook. The Pelican Brook Subwatershed has the second lowest normalized Raster Seeding Method CAC score, 14.6, and displays similar parallel stream effects for many of its flowlines. The true accumulated drainage area for Pelican Brook, which includes the Pelican Lake drainage area, is 47.7 mi<sup>2</sup>, not the 3 mi<sup>2</sup> area produced by the Raster Seeding Method, a drainage-area error of 93.65 percent.

One WBD Subwatershed downstream from Pelican Brook contains the outlet flowline for the entire Pine River Subbasin. At the outlet, the accumulation of upstream contributing drainage area totals 780 mi<sup>2</sup>. The Raster Seeding Method drainage area for the same outlet flowline is 749 mi<sup>2</sup>,

yielding a 31-mi<sup>2</sup> difference and a cumulative drainage-area error of 3.9 percent.

The Subwatershed boundary with the third lowest normalized CAC score for the Raster Seeding Method, 62.4, also has the lowest NEM-without-WBD score, 67.2. The Raster Seeding Method area shares the same delineation differences as the NEM-without-WBD area (see figure 1–13) with additional errors introduced by DEM flow-path displacement. Compared with the NEM-without-WBD, the WBD boundary provided the more accurate delineation. This boundary has the fifth highest cumulative drainage-area error for the Raster Seeding Method at 27.36 percent.

The primary delineation error observed for the boundary with the fourth lowest normalized CAC score (71.14) for the Raster Seeding Method could have been avoided in the Raster Seeding Method production process. The error involves an area of approximately 13.6 mi<sup>2</sup> outside the Pine River Subbasin that was assigned as an area contributing to a headwater flowline in the Pine River Subbasin. If the flowlines for the adjacent Subbasin had been used in the Raster Seeding Method process as boundary constraints, then the error for this area would have been reduced substantially. The NEM-without-WBD and the Thiessen Polygon Method incorporated adjacent flowlines in their catchment delineation process and, therefore, did not yield this type of error. Smaller errors from DEM flow-path displacement also were observed for this boundary.

The area with the fifth lowest normalized CAC score for the Raster Seeding Method, 74.28, is also the area with the second lowest normalized CAC score for the NEM-without-WBD, 75.6. The Raster Seeding Method boundary has the same delineation differences as the NEM-without-WBD (see figure 1–14) with additional errors introduced by DEM flow-path displacement.

DEM flow-path displacement contributed substantially to the sixth lowest normalized CAC score for the Raster Seeding Method, 77.09. The NEM-without-WBD score for this same boundary is 91.0. The boundary is a headwater Subwatershed with a true drainage area of 14.7 mi<sup>2</sup>. The drainage area of the Raster Seeding Method boundary is 12.3 mi<sup>2</sup> with a 15.9 percent cumulative drainage-area error. The drainage-area error for the NEM-without-WBD for this area is just 0.49 percent and for the Thiessen Polygon Method, 33.20 percent. DEM flow-path displacement also affected the seventh lowest Raster Seeding Method normalized CAC score, 79.65 (85.69 for NEM-without-WBD).

Raster Seeding Method catchment boundaries in the Pine River Subbasin have a 38 percent correspondence to the catchment boundaries of the NEM-without-WBD, the lowest correspondence of all the study Subbasins. Differences between the catchment boundaries from the NEM-without-WBD and Raster Seeding Method are mostly the result of DEM flow-path displacement errors that occur with the Raster Seeding Method.

DEM flow-path displacement errors in the Raster Seeding Method also caused inaccurate results in the cumulative drainage-area convergence analysis (table 1–6). If the parallel stream artifact errors of DEM flow-path displacement had not occurred in the two boundaries shown in figure 1–16, results of the cumulative drainage-area analysis would have been improved. By removing the two boundaries with the highest cumulative drainage-area errors, the mean error for all drainage-area sizes would be 7 percent instead of 11.32 percent; the mean error for drainage areas 20 to 60 mi<sup>2</sup> in size would be 8.84 percent instead of 24.79 percent. The Raster Seeding Method had the highest mean error in this category. DEM-flow-path displacement errors, though not as great a problem as seen in figure 1–16, contributed 32 percent to the mean error shown for the Raster Seeding Method in the smallest drainage-area size category of the cumulative drainage-area results (table 1–6).

### Thiessen Polygon Method

The mean normalized CAC score for the Thiessen Polygon Method is 66.3. No normalized CAC scores are greater than 90. Twenty-seven of the 28 Thiessen Polygon Method boundaries (96 percent), have normalized CAC scores less than 80. The Thiessen Polygon Method produced the highest mean normalized polygon-area error, 21.95 percent, in the WBD Subwatershed comparison, compared to 15.4, 5.7, and 0 percent for the Raster Seeding Method, NEM-without-WBD, and NEM-with-WBD, respectively.

Thiessen Polygon Method boundaries tested in the cumulative drainage-area analysis have a mean drainage-area error for all drainage-area sizes of 13.69 percent, the highest mean error of all the methods. The Thiessen Polygon Method also has the highest mean error of all the methods for drainage areas less than 20 mi<sup>2</sup> in size, 28.76 percent, and for drainage areas greater than 60 to 390 mi<sup>2</sup> in size, 9.70 percent. A complete summary of the results of the cumulative drainage-area analysis for the Thiessen Polygon Method is provided in table 1–6. The Thiessen Polygon Method produced reasonably small errors, about 2 percent or less, for drainage areas greater than 390 mi<sup>2</sup>.

### Comparison of Methods with Streamgage Data

Verification data for two streamgage drainage-areas were obtained for the Pine River Subbasin (fig. 1–17). The delineations from the WBD data were compared to the GIS verification drainage-area boundaries for both streamgages. The delineations are the same, except for the areas near the streamgage locations.

The streamgage drainage-area verification measurements differed between the NWIS and GIS data. The NWIS drainage-area measurement for the streamgage on Pine River near Pine River, MN (05229450), is 261 mi<sup>2</sup>, and the GIS measurement is 269.54 mi<sup>2</sup>, a 3 percent difference. Because the NWIS

drainage-area measurement probably was based upon older delineation methods and source maps, the NWIS drainage area is most likely less accurate than the GIS measurement. NWIS did not contain a drainage-area measurement for streamgage 05229430 on Hoblin Creek near Pine River, MN; the GIS measurement is 1.03 mi<sup>2</sup>.

Imperfect CAC scores (less than 100) were obtained for the comparison between the two streamgage drainage-area boundaries and the NEM-with-WBD boundaries. For each drainage area, the CAC score can be linked to the location of the streamgage in relation to the matched NHD-flowline outlet. The streamgage on Pine River is located 8 percent of the way up the length of the flowline, resulting in a CAC score of 98.18. The streamgage on Hoblin Creek is 85 percent of the way up the length of the flowline, resulting in a CAC score of only 15.65. The CAC scores of the NEM-with-WBD were used to normalize CACs from the other methods in order to adjust for the streamgage location/NHD-flowline outlet mismatch error.

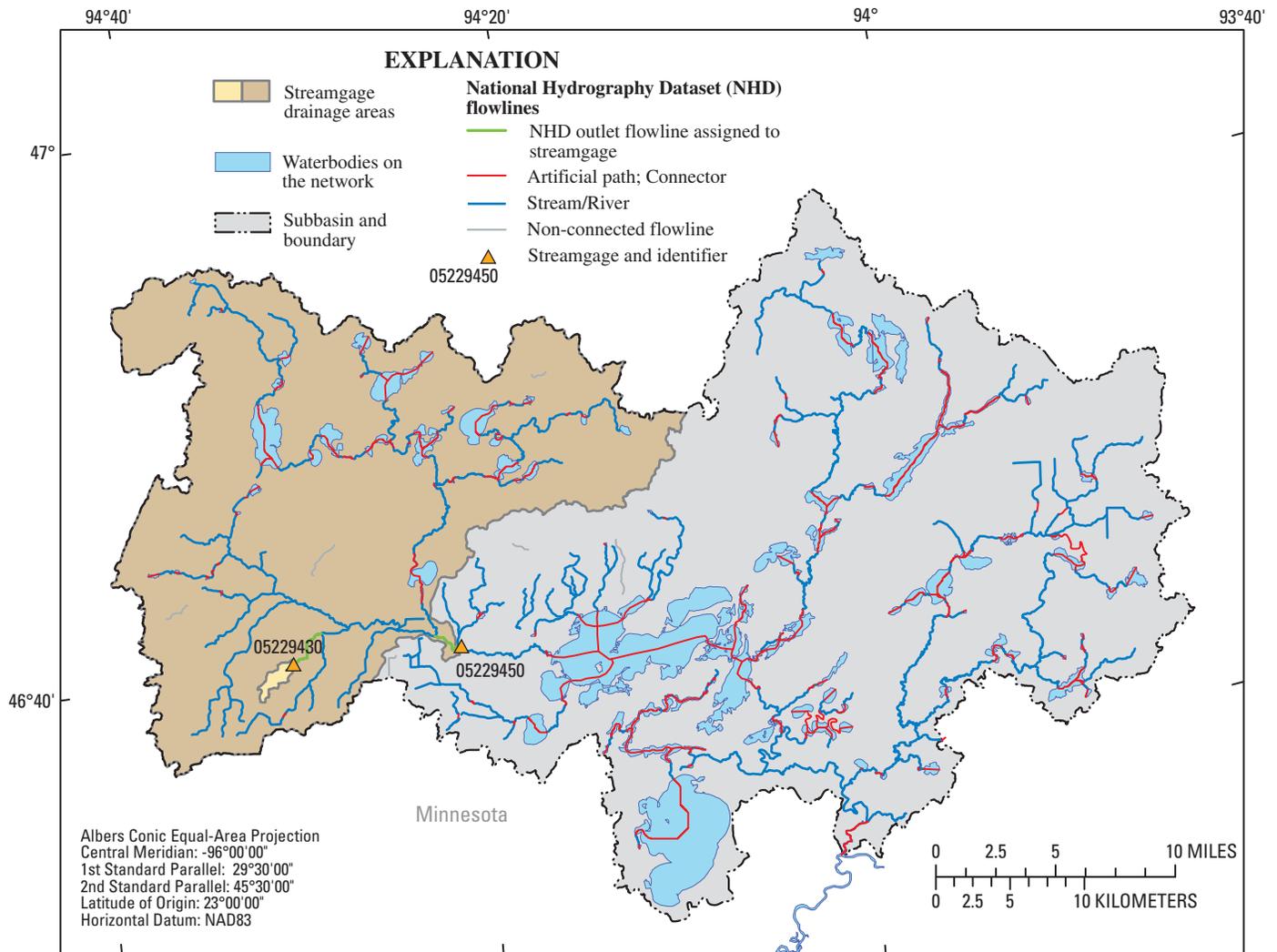
The mean normalized CAC score for the NEM-without-WBD is 98.69. For the streamgage on Pine River near Pine River, MN, the CAC score is 97.63. No major differences were observed between the NEM-without-WBD boundary and the streamgage drainage-area boundary. The verification boundary provided the more accurate delineation for the streamgage on Hoblin Creek.

The Raster Seeding Method delineation for the Pine River streamgage is nearly identical to the NEM-without-WBD delineation in most areas. A boundary running parallel to the flowline that the streamgage is located on was observed in the Raster Seeding Method. At 97.51, the normalized Raster Seeding Method CAC score for the drainage area for the streamgage on Pine River is a close second to NEM-without-WBD (97.63).

The normalized Raster Seeding Method CAC score for the streamgage on Hoblin Creek is 207.03. This high CAC score resulted from an error introduced by DEM flow-path displacement and streamgage location mismatch with the assigned flowline (see figure 1–17). The verification boundary provided the more accurate delineation for the area of discrepancy (same as in the NEM-without-WBD).

The Thiessen Polygon Method produced the lowest normalized CAC score for the Pine River streamgage (05229450), 87.91. The Thiessen Polygon Method also has the highest cumulative drainage-area error, compared to the other methods, for the streamgage at Pine River, 7.72 percent. The drainage-area error for this streamgage for NEM-without-WBD is 0.82 percent and the Raster Seeding Method, 0.58 percent.

The streamgage location/NHD-flowline outlet mismatch made it difficult to quantify the Thiessen Polygon Method results for the small drainage area associated with the streamgage at Hoblin Creek (05229430). Visual review of the Thiessen Polygon Method delineation for this streamgage determined that the other methods provided more accurate boundary delineations and associated drainage-area statistics.



Base from 1:100,000-scale National Hydrography Dataset, 2003

**Figure 1–17.** Location of streamgages and associated drainage areas used as verification datasets, Pine River Subbasin, MN. (The location of this Subbasin is shown in figure 15.)

## Summary of Results for Pine River

The NEM-with-WBD CAC comparative analysis revealed errors related to WBD/NHD-flowline outlet mismatches and NHD issues, which had similar effects on all the methods. Three isolated NHD stream networks were identified. The effects of the isolated stream networks on the total drainage area for the Pine River Subbasin outlet flowline was considered minimal with a 0.61 percent cumulative drainage-area error. An isolated network within a headwater WBD Subwatershed boundary caused a 10.6 percent drainage-area error. An improper NHD flow relationship affected cumulative drainage-area analysis results for three Subwatersheds with errors as great as 17.38 percent.

The normalized CAC scores show that the NEM-without-WBD provided the most accurate results after the NEM-with-WBD. The mean normalized CAC score for the NEM-without-WBD is 89.1 in the WBD comparison analysis and 98.7 in the streamgage analysis. Four boundaries have normalized CAC scores less than 80. The lowest normalized CAC score for this method indicates that the WBD boundary is correct and that the DEM-derived boundary from the NEM-without-WBD is in error. Many noticeable differences in delineations favored the WBD boundary, yet other differences were undeterminable because of the limited or complex contour detail depicted on the DRGs. The NEM-without-WBD also provided the second lowest error in the cumulative drainage-area analysis with a mean error of 4.84 percent for all drainage-area sizes tested.

For the Raster Seeding Method, the mean normalized CAC score in the WBD analysis was 81.7. Only 38 percent of the catchment delineations from the Raster Seeding Method correspond to the catchments from the NEM-without-WBD. Most of the delineation differences for the Raster Seeding Method are the result of errors associated with DEM flow-path displacement. For all seven boundaries with the lowest normalized Raster Seeding Method CAC scores, parallel stream boundary artifacts from DEM flow-path displacement had negligible effects. Parallel stream artifacts were especially problematic for two WBD Subwatershed boundaries encompassing lakes and flat terrain; in these cases, the delineations resembled 30-m buffer zones around the NHD flowlines. The normalized Raster Seeding Method CAC scores for the two Subwatersheds are 0.15 and 14.3. The cumulative drainage-area errors for these two Subwatersheds are 99 and 93 percent. These values skew the results of the Raster Seeding Method and make it the least accurate method in the cumulative drainage-area analysis for drainage areas of 20 to 60 mi<sup>2</sup> and greater than 390 mi<sup>2</sup>. Parallel stream artifacts, though not causing as great a problem as those mentioned above, also contributed to higher errors for other boundaries in the results of the cumulative drainage-area analysis. For the streamgage analysis, one streamgage has a normalized CAC score of 97.5,

whereas the normalized CAC score for the other streamgage is 207, which is explained by DEM flow-path displacement and the streamgage location mismatch with the flowline outlet.

In the CAC analysis with WBD, the Thiessen Polygon Method provided the least accurate results with a mean normalized CAC score of 66.3. No normalized CAC scores greater than 90 resulted from this method; 96 percent of the Thiessen Polygon boundaries resulted in a CAC score of less than 80. The Thiessen Polygon Method had the highest mean normalized polygon-area error at 21.95 percent with the WBD comparisons. The Thiessen Polygon Method also produced the highest mean cumulative drainage-area error of 13.69 percent and the highest mean drainage-area error of 28.76 percent for drainage areas less than 20 mi<sup>2</sup>. In this method, the largest drainage areas (greater than 390 mi<sup>2</sup>) had fairly small errors, 2.02 percent or less.

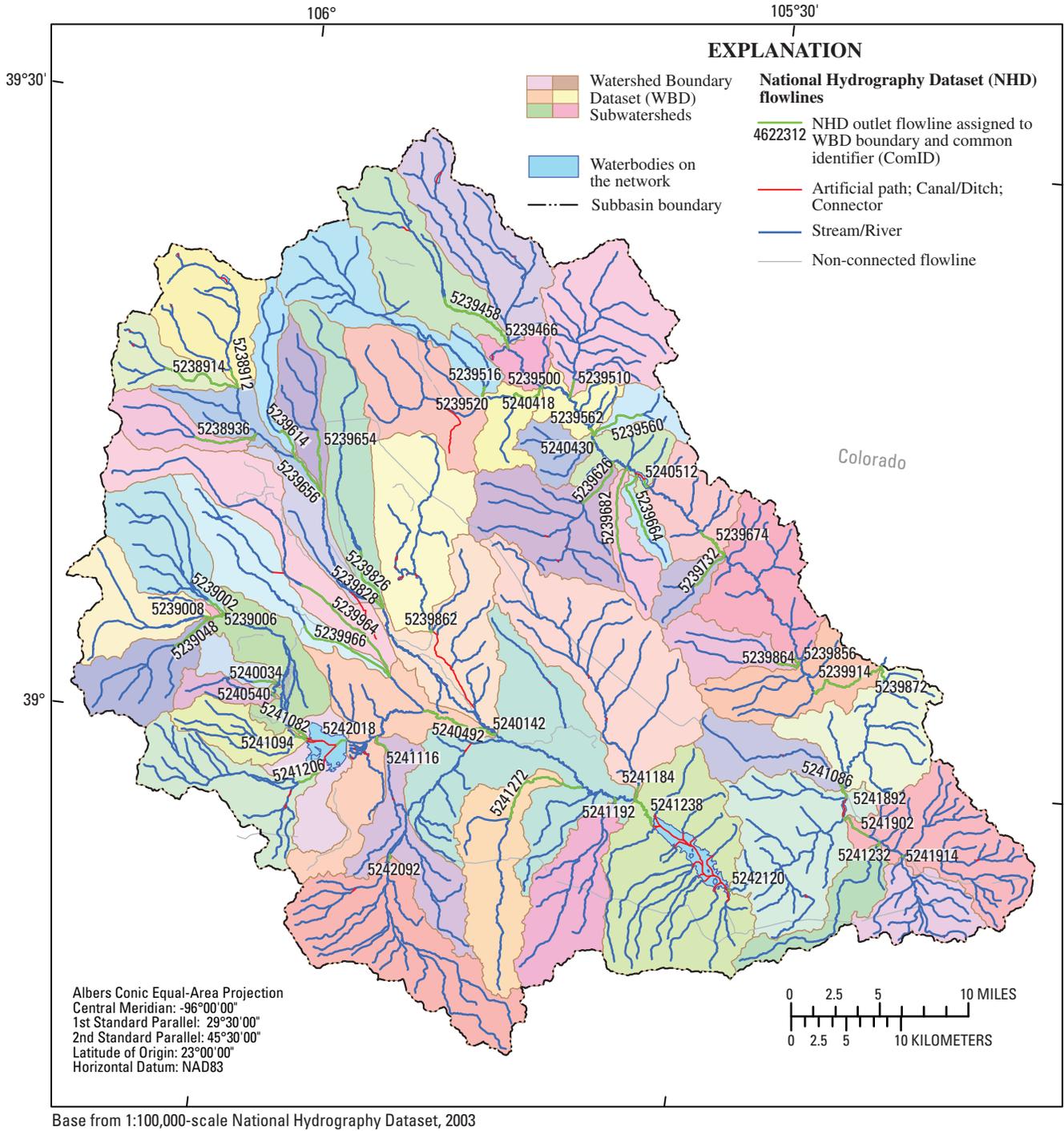
## South Platte Headwaters, CO (10190001)

### Comparison of Method Boundaries with WBD Boundaries

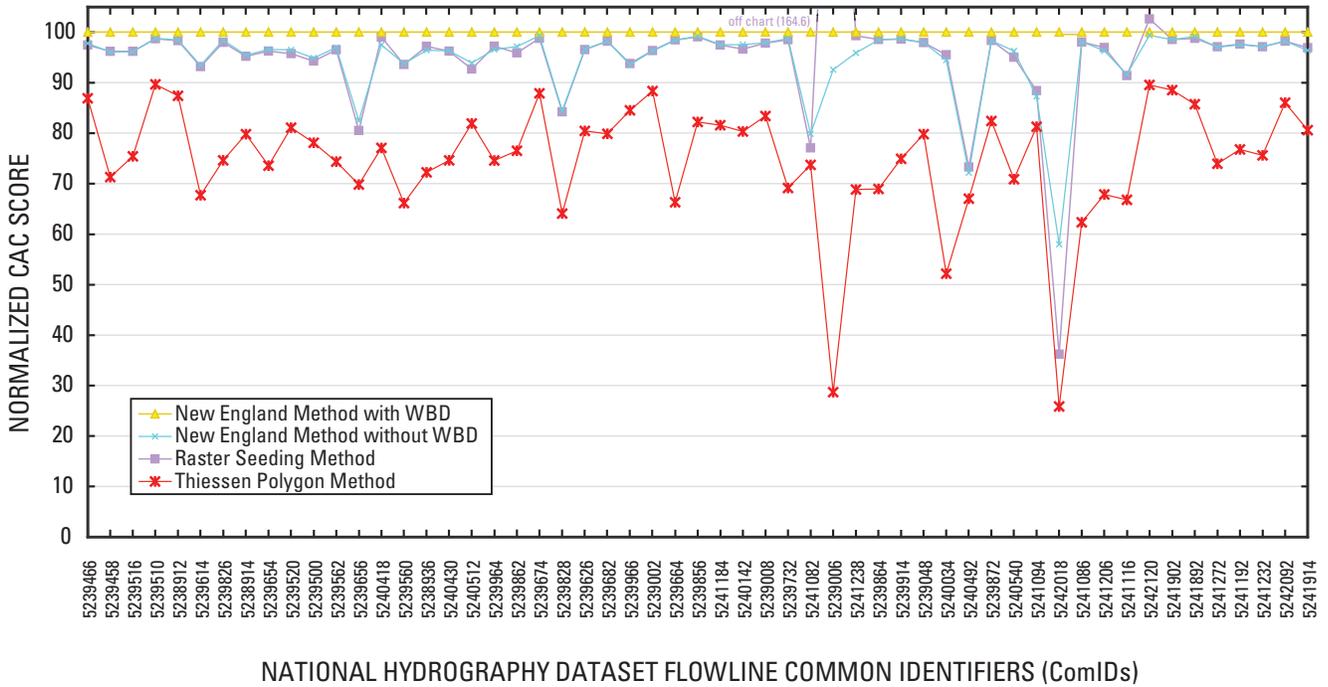
Draft WBD Subwatershed boundaries in the South Platte Headwaters Subbasin (fig. 1–18), delineated to the 12-digit hydrologic unit level, were obtained and used in the CAC and cumulative drainage-area analyses. Fifty-five WBD Subwatershed boundaries were used in the CAC analysis; results are shown in figure 1–19 and table 1–7. Results of the cumulative drainage-area convergence analyses are summarized in table 1–8. A graph developed as part of the convergence analysis displaying the drainage-area errors for each method is shown in figure 1–20.

### New England Method with WBD (NEM-with-WBD)

WBD/NHD-flowline outlet mismatches, two isolated NHD stream networks, and NHD flowlines crossing WBD boundaries contributed to the lowering of the CAC scores (before normalizing to 100) of the NEM-with-WBD. The lowest CAC score of the NEM-with-WBD is 50.3 for a catchment-sized boundary in the draft WBD with an area of 0.65 mi<sup>2</sup>. (Certified WBD data would not include such a small area as found in this draft WBD data.) For this area, the NHD stream crosses the WBD boundary at a point where the WBD boundary runs parallel and close to the hydrography depicted on the 1:24,000-scale DRGs. The 1:100,000-scale NHD flowline is shifted slightly from the hydrography on the DRGs and actually meanders across the WBD divide, causing the delineation to extend beyond the divide. Both New England Methods and the Thiessen Polygon Method included areas extending beyond the WBD divide.



**Figure 1–18.** The draft Watershed Boundary Dataset (WBD) Subwatershed boundaries, and National Hydrography Dataset (NHD) flowlines, South Platte Headwaters Subbasin, CO. (The location of this Subbasin is shown in figure 15.)



**Figure 1–19.** Normalized Coefficient of Areal Correspondence (CAC) scores, by National Hydrography Dataset (NHD) flowline common identifier, for all methods tested, South Platte Headwaters Subbasin, CO.

**Table 1–7.** Mean normalized Coefficient of Areal Correspondence score by delineation methods for Watershed Boundary Dataset Subwatersheds in the South Platte Headwaters Subbasin, CO.

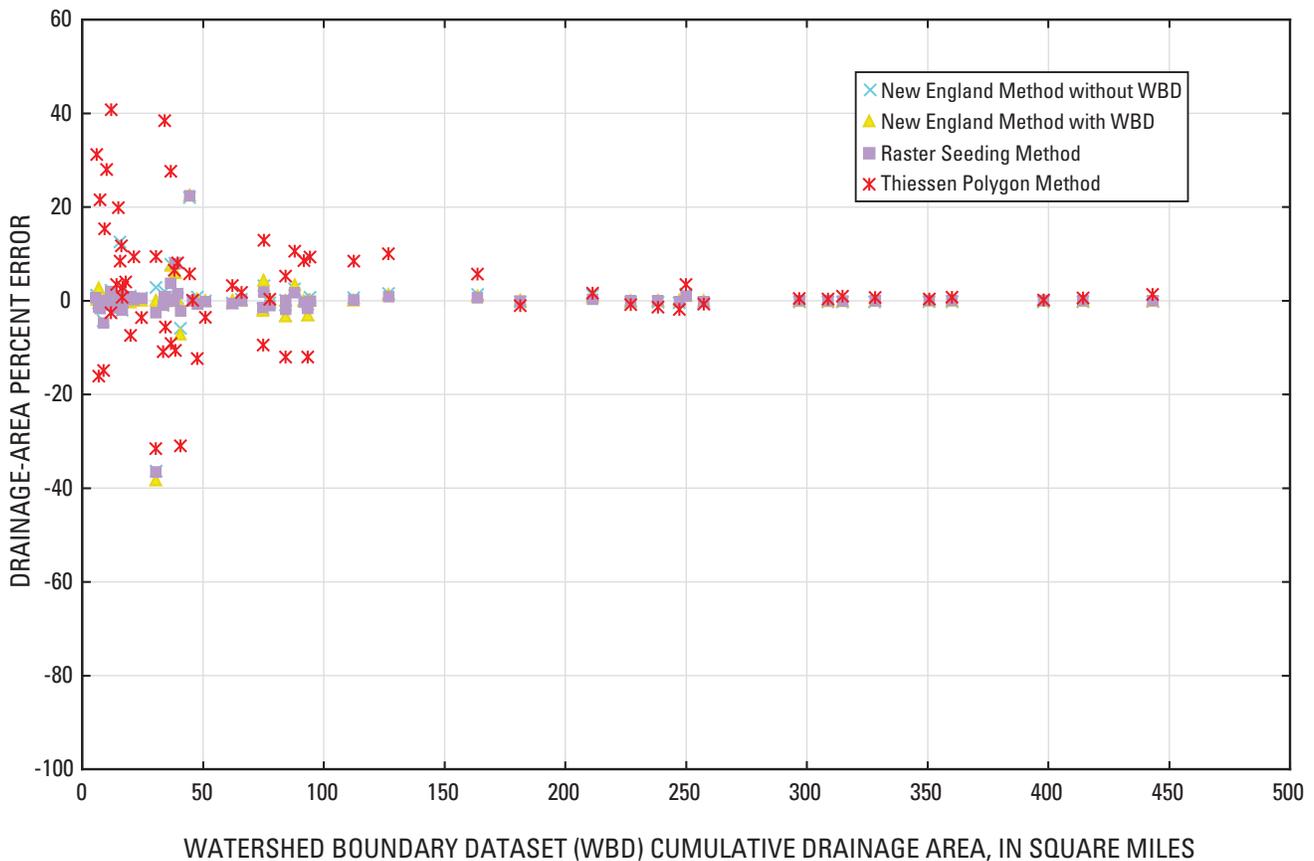
[CAC, Coefficient of Areal Correspondence; NEM, New England Method; WBD, Watershed Boundary Dataset]

Method	Mean normalized CAC	Normalized CACs greater than 90		Normalized CACs less than 80	
		Number of Subwatersheds	Percentage of Subwatersheds	Number of Subwatersheds	Percentage of Subwatersheds
NEM-with-WBD	100.0	55	100	0	0
NEM-without-WBD	94.7	49	89	3	5
Raster Seeding Method	95.6	49	89	3	5
Thiessen Polygon Method	74.7	0	0	35	64

**Table 1-8.** Mean cumulative drainage-area errors for all delineation methods for various drainage-area sizes, South Platte Headwaters Subbasin, CO.

[mi<sup>2</sup>, square miles; NEM, New England Method; WBD, Watershed Boundary Dataset]

Method (69 units total)	Mean cumulative drainage-area error, in percent					
	Drainage-area size category					
	All drainage areas	Less than 20 mi <sup>2</sup>	20–60 mi <sup>2</sup>	Greater than 60–390 mi <sup>2</sup>	Greater than 390 mi <sup>2</sup>	At Subbasin outlet 1,605 mi <sup>2</sup>
NEM-with-WBD	1.71	0.16	3.54	1.07	0.69	0.72
NEM-without-WBD	1.97	1.57	3.99	1.43	0.60	0.63
Raster Seeding Method	1.99	1.04	4.02	1.31	0.78	0.72
Thiessen Polygon Method	8.66	15.64	11.05	5.53	1.76	1.32



**Figure 1-20.** Drainage-area percent error in relation to Watershed Boundary Dataset cumulative drainage areas, for all methods tested, South Platte Headwaters Subbasin, CO.

There were numerous places where NHD flowlines crossed the WBD divides. Based on a review of the DRGs, in many cases, the NHD is correct and the draft WBD divides are in error. Some of the lowest CAC scores of the NEM-with-WBD occurred where these flowlines cross the divides. In one case, a connector flowline in the NHD crossed a WBD divide and connected with a stream located in the adjacent WBD Subwatershed. The affected boundaries have CAC scores of 76.0 and 61.3 with cumulative drainage-area errors of 22.71 and 38.39 percent.

The Subbasin had two isolated NHD stream networks with delineated areas totaling 11.39 mi<sup>2</sup>. The NEM-with-WBD CAC scores of the two affected WBD boundaries are 83.6 and 82.1 with normalized polygon-area errors of 7.06 and 14.58 percent. The effects of the isolated networks on cumulative drainage area were minimal with an error of 0.72 percent for the outlet of this 1,605-mi<sup>2</sup> Subbasin. The effects of WBD/NHD-flowline outlet mismatches were minimal as most WBD Subwatershed boundary outlets matched reasonably well with the confluences in the NHD.

### New England Method without WBD (NEM-without-WBD)

The mean normalized CAC score for the NEM-without-WBD is 94.7. Forty-nine of the 55 WBD boundaries (89 percent) have normalized CAC scores greater than 90 (table 1–7). Only three normalized CAC scores fall below 90 (5 percent) with values of 58.0, 72.2, and 79.9. NEM-without-WBD produced the second smallest mean normalized polygon-area error of 2.74 percent; the Raster Seeding Method error is 3.7 percent; and the Thiessen Polygon Method error is 12.9 percent.

The NEM-without-WBD produced mean cumulative drainage-area results similar to those of the NEM-with-WBD and Raster Seeding Method. The mean drainage-area error for the NEM-without-WBD is 1.97 percent; for the NEM-with-WBD, 1.71 percent; and for the Raster Seeding Method, 1.99 percent (table 1–8). The mean cumulative drainage-area errors for the various drainage-area size categories are similar among the two New England Methods and the Raster Seeding Method. The mean cumulative drainage-area errors for the NEM-without-WBD were not substantial, especially if the errors introduced by isolated networks and other NHD-related issues were removed from the analysis.

The lowest normalized CAC score for the NEM-without-WBD, 58.0, was calculated for the WBD Subwatershed for the Antero Reservoir, which had the highest normalized polygon-area error of 49.96 percent. Several delineated areas that differ between the Antero Reservoir WBD boundary and the corresponding boundary of the NEM-without-WBD are shown in figure 1–21. Because of limited information on the DRGs or DOQs, the correct delineation for the largest area of difference, approximately 9.3 mi<sup>2</sup>, could not be determined. Most of the 9.3-mi<sup>2</sup> area is associated with an NHD flowline that is not connected to the network, and

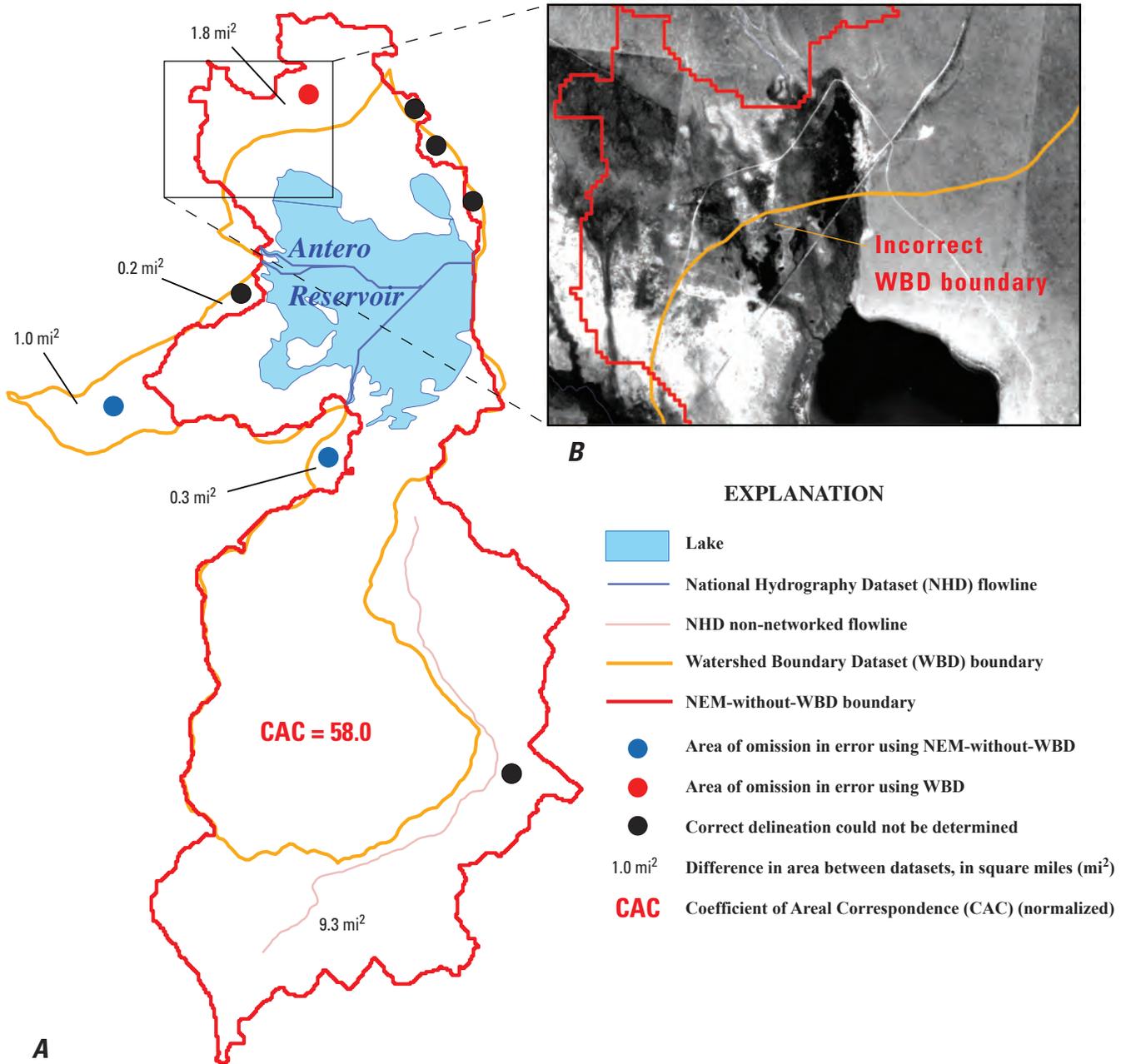
therefore, the associated area has actually been assigned to another NHD flowline. Like the NHD, the non-networked stream segment, as depicted on the DRG, ends abruptly. At the stream termination ending point, the land surface appears flat on the DRGs, making it difficult to determine the direction of flow past the downstream end of the flowline. The WBD delineation for Antero Reservoir does not include the drainage area associated with the non-networked flowline, whereas the NEM-without-WBD delineation does.

The next greatest difference in an area for the WBD boundary shown in figure 1–21A is 1.8 mi<sup>2</sup>. Information shown on the DOQ for this area indicates that the draft WBD divide shown in figure 1–21B is incorrect. For other area differences of 1.0 and 0.3 mi<sup>2</sup>, the WBD data provide a more accurate delineation over the NEM-without-WBD. For other smaller areas of difference, a determination could not be made on the basis of limited contour detail depicted on the DRGs as to which method was more accurate.

The cumulative drainage-area error is 3.3 percent for the Subwatershed boundary produced by the NEM-without-WBD, which has a WBD cumulative drainage area of 191 mi<sup>2</sup> (fig. 1–21A). The 9.3-mi<sup>2</sup> area of difference (shown in figure 1–21A) was the major contributor of the cumulative drainage-area error. An isolated stream network upstream from the WBD Subwatershed canceled out some of the error. If the 2.7-mi<sup>2</sup> area associated with the isolated network had been included in the cumulative drainage area, the error would have been 4.4 percent.

The 9.3-mi<sup>2</sup> delineation area of difference shown in figure 1–21A is largely responsible for lowering the CAC score for the adjacent boundary. The adjacent boundary has the second lowest normalized CAC score of 72.2 and the second highest normalized polygon-area error of 20.17 percent. Other delineation differences were observed elsewhere for this WBD boundary, but none was as significant as the 9.3-mi<sup>2</sup> area. For many areas, the WBD was correct, whereas the NEM-without-WBD was more accurate in other areas. The NEM-without-WBD cumulative drainage-area error for this WBD Subwatershed is 1.25 percent. If the effects of an isolated stream network were removed, the cumulative drainage-area error would be 0.58 percent.

The third lowest normalized CAC score, 79.9, was determined for a NEM-without-WBD boundary located north of, and adjacent to, the WBD boundary shown in figure 1–21A. This boundary also has the third highest NEM-without-WBD normalized polygon-area error at 18.27 percent. Five distinct areas with delineation differences were observed between the WBD and NEM-without-WBD boundaries. For the largest area of difference, the 1.8-mi<sup>2</sup> area shown in figures 1–21A and B, the NEM-without-WBD provided the more accurate boundary. For this Subwatershed, with an actual drainage area of 122 mi<sup>2</sup>, the NEM-without-WBD cumulative drainage-area error is 5.08 percent. An isolated stream network contributed 2.2 percent of the error.



**Figure 1–21.** (A) The Watershed Boundary Dataset (WBD) and the New England Method without WBD (NEM-without-WBD) boundaries for the Antero Reservoir Subwatershed and (B) area of detail showing boundary differences overlaid on the Digital Orthophoto Quadrangle (DOQ), South Platte Headwaters Subbasin, CO. (The location of this WBD boundary is shown in figure 1–18.)

## Raster Seeding Method

The Raster Seeding Method has a mean normalized CAC score of 95.6. As observed in the NEM-without-WBD, 89 percent of the Raster Seeding Method boundaries tested had normalized CAC scores greater than 90, with just three boundaries less than 80. Two of the boundaries had normalized CAC scores greater than 100, 164.6 and 102.6. The three lowest normalized CAC scores for the Raster Seeding Method are 36.2, 73.3, and 77.1. The cumulative drainage-area results for the Raster Seeding Method compared well with the results of the two New England Methods. The Raster Seeding Method mean cumulative drainage-area errors for all drainage-area size categories were not substantial. Catchment boundaries from the Raster Seeding Method have a 75-percent spatial correspondence with those from the NEM-without-WBD.

The 0.65-mi<sup>2</sup> catchment-sized WBD boundary, discussed previously in the NEM-with-WBD Method section, has a normalized Raster Seeding Method CAC score of 164.6. The error from DEM flow-path displacement in the Raster Seeding Method, by chance, actually provided the more accurate delineation for this WBD boundary.

The boundary with the lowest normalized CAC score for the Raster Seeding Method, 36.2, is the same corresponding boundary of the NEM-without-WBD shown in figure 1–21A. The 1.8-mi<sup>2</sup> area of difference shown in figure 1–21 occurs for the Raster Seeding Method boundary as well. Once again, for this 1.8-mi<sup>2</sup> area, the DEM delineation provided the more accurate boundary. The area with the most substantial area of difference between the Raster Seeding Method and WBD boundaries was a direct result of DEM flow-path displacement. The catchment boundary delineation errors due to DEM flow-path displacement are shown in figure 1–22.

The Raster Seeding normalized polygon-area percent error for the Antero Reservoir Subwatershed is 43.32 percent. For this Subwatershed, the Raster Seeding Method has the highest cumulative drainage-area error at 8.48 percent, compared to 1.62 percent for the NEM-with-WBD, 3.03 percent for the NEM-without-WBD, and 2.63 percent for the Thiessen Polygon Method. If a 4.2-mi<sup>2</sup> area associated with an isolated NHD flowline had been included in the cumulative drainage area for the reservoir, the Raster Seeding Method cumulative drainage-area error would have been 6.26 percent.

A Subwatershed adjacent to the Antero Reservoir Subwatershed (fig. 1–22) has the second lowest normalized Raster Seeding Method CAC score of 73.3. This boundary also has the second lowest normalized NEM-without-WBD CAC score of 72.2. In the Raster Seeding Method boundary, the same parallel-to-stream boundary error, caused by DEM flow-path displacement, for an artificial path flowline within the Antero Reservoir (fig. 1–22) served as a major error. Other delineation errors due to DEM flow-path displacement were observed throughout the Subwatershed. Many differences between the Raster Seeding Method and WBD boundaries were identical to those observed in the NEM-without-WBD boundary. The Raster Seeding Method cumulative drainage-area error for the

boundary is 1.63 percent. If the effects of an isolated stream network were removed, the cumulative area in error would have been 0.58 percent.

The boundary with the third lowest normalized CAC score (77.1) for the Raster Seeding Method also is adjacent to the Antero Reservoir Subwatershed (fig. 1–22). This boundary also has the third lowest normalized CAC score for the NEM-without-WBD (79.9). Some of the differences between the Raster Seeding Method boundary and the WBD boundary were the same as those observed for the NEM-without-WBD boundary, including the 1.8-mi<sup>2</sup> area identified in figure 1–21. Additional Raster Seeding Method errors due to DEM flow-path displacement were observed here.

## Thiessen Polygon Method

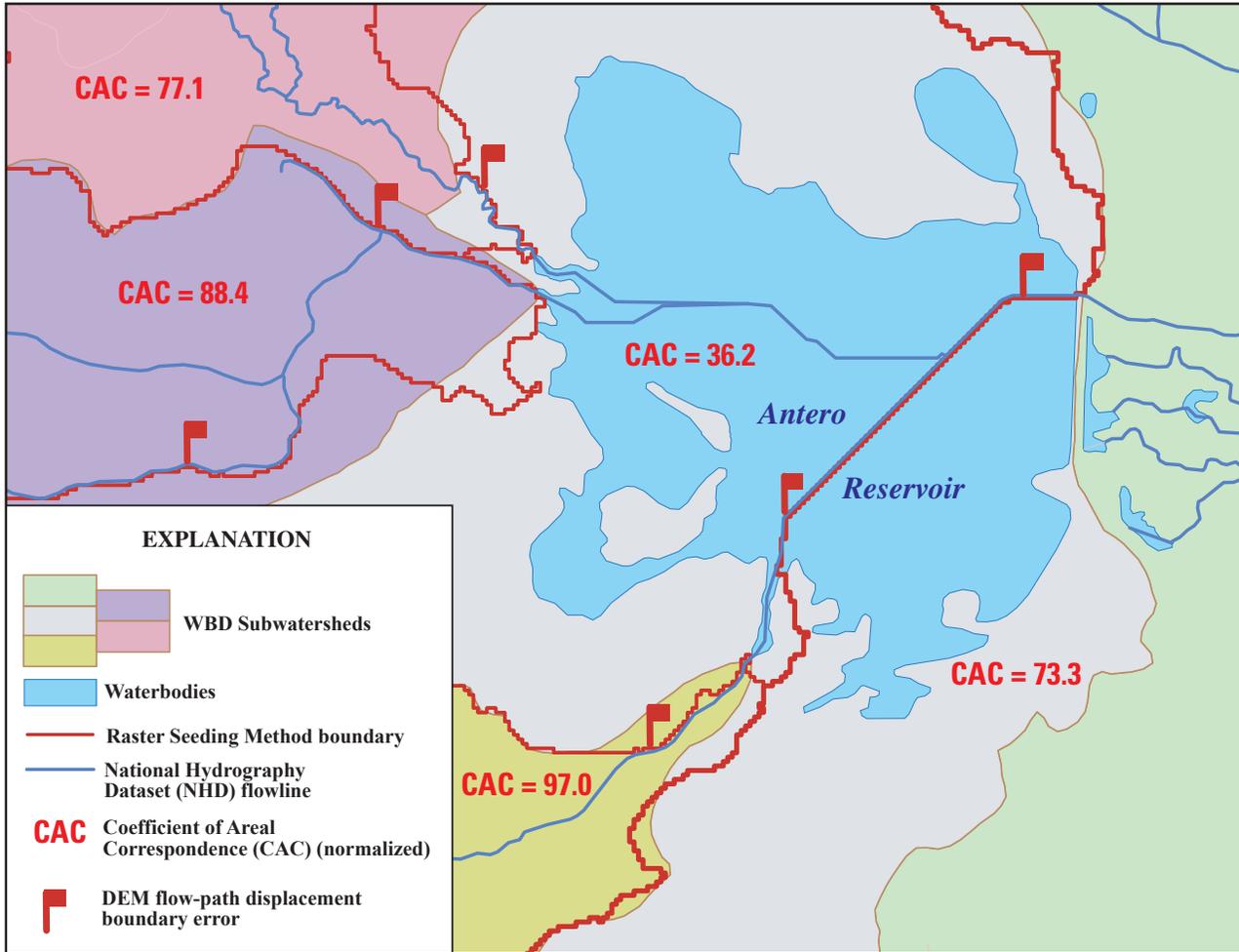
The mean normalized CAC for the Thiessen Polygon Method is 74.7, the lowest score of all the methods. Thiessen Polygon was the only method for which none (0 percent) of the boundaries has a normalized CAC score greater than 90. The Thiessen Polygon Method also has the highest percentage of boundaries (64 percent) with normalized CAC scores less than 80. For NEM-without-WBD and the Raster Seeding Method, just 5 percent of the boundaries have scores less than 80. The Thiessen Polygon Method also has the highest normalized polygon-area error of 12.87 percent, compared with 3.70 percent for the Raster Seeding Method and 2.74 percent for NEM-without-WBD.

The Thiessen Polygon Method has the highest mean cumulative drainage-area error at 8.66 percent, and the highest mean errors for all the various drainage-area size categories shown in table 1–8. Cumulative drainage-area errors are especially high for the smaller drainage-area size categories. For drainage areas less than 20 mi<sup>2</sup> in size, the Thiessen Polygon Method mean error is 15.64 percent, greatly surpassing the mean errors of the other methods. Errors also are high for drainage areas 20 to 60 mi<sup>2</sup> in size, with a mean error at 11.05 percent.

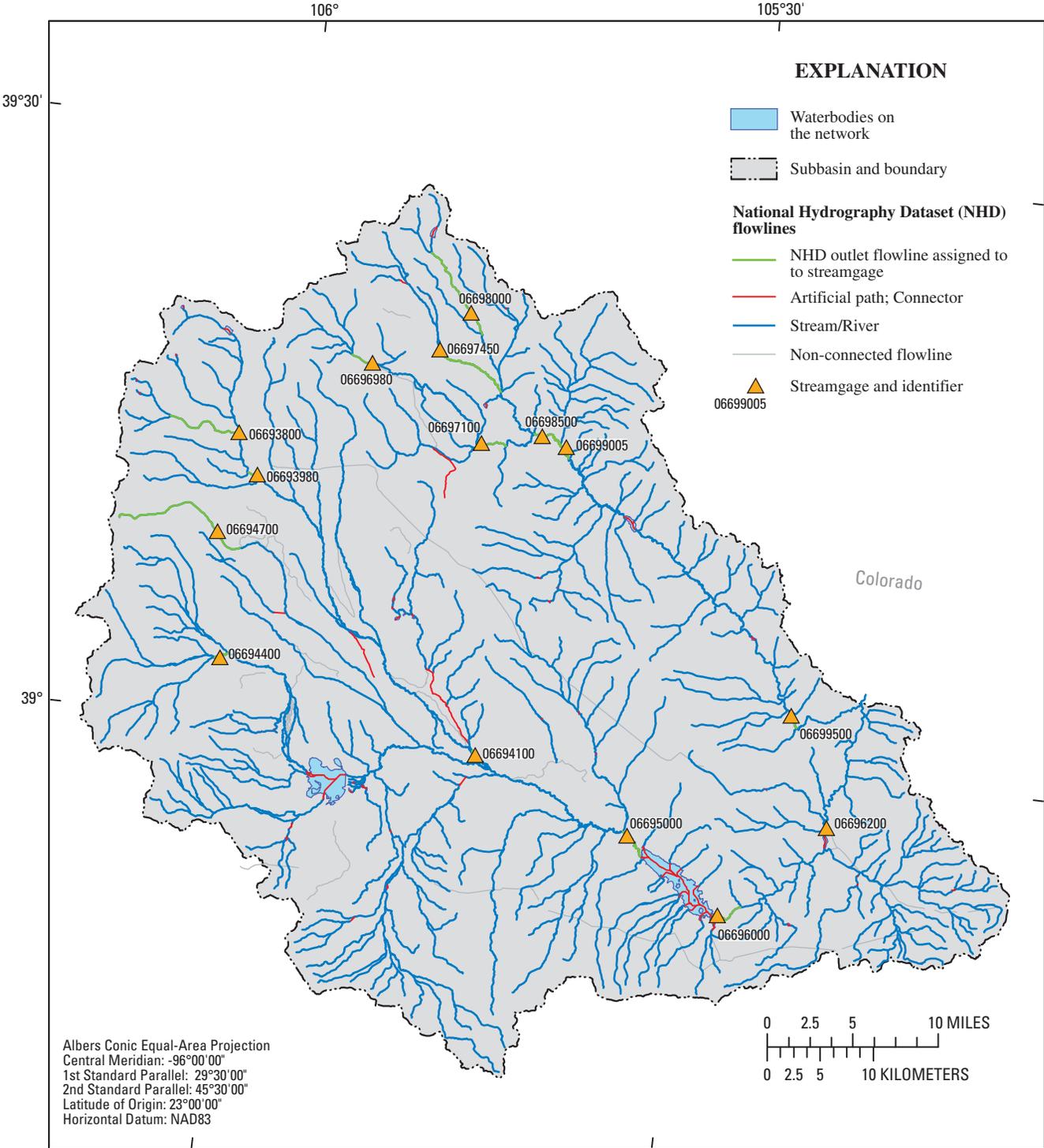
## Comparison of Methods with Streamgage Data

For the streamgage comparative analysis, only published drainage-area measurements from the USGS NWIS database were used as verification. No streamgage drainage-area boundaries were available at the time of the study. From NWIS, 15 streamgages were selected for the analysis (fig. 1–23). Drainage-area measurements for the streamgages range from 12 to 1,084 mi<sup>2</sup>. Mean cumulative drainage-area errors, in percent, are summarized in table 1–9.

NEM-with-WBD, the presumed most accurate delineation method, had drainage-area error contributions from NHD-related artifacts. NHD artifacts include isolated networks and NHD flowlines crossing the streamgage drainage divides. Many of the streamgages were located upstream to the outlet of the assigned NHD flowline, introducing some errors from these artifacts. Cumulative drainage-area errors also could be



**Figure 1–22.** The Watershed Boundary Dataset (WBD) Subwatersheds and Raster Seeding Method boundaries with normalized Coefficient of Areal Correspondence (CAC) scores and digital elevation model (DEM) flow-path displacement boundary errors, South Platte Headwaters Subbasin, CO. (The location of this area is shown in figure 1–18.)



**Figure 1-23.** Location of streamgages used as verification data, South Platte Headwaters, CO. (The location of this Subbasin is shown in figure 15.)

**Table 1–9.** Mean drainage-area errors for all delineation methods (except the Outlet Matching Method), for the National Water Information System (NWIS) streamgages, grouped by various drainage-area sizes, South Platte Headwaters, CO.[mi<sup>2</sup>, square miles; NEM, New England Method; WBD, Watershed Boundary Dataset]

Method (15 streamgages total)	Mean drainage-area error, in percent				
	Drainage-area size category				
	All drainage areas	Less than 20 mi <sup>2</sup>	20–60 mi <sup>2</sup>	Greater than 60–390 mi <sup>2</sup>	Greater than 390 mi <sup>2</sup>
NEM-with-WBD	2.62	4.12	3.93	4.17	0.66
NEM-without-WBD	2.68	4.93	3.94	2.27	0.55
Raster Seeding Method	2.48	4.46	3.45	2.09	0.76
Thiessen Polygon Method	6.41	18.50	2.28	5.25	1.88

the result of differences in the way the drainage areas were computed for NWIS. As discussed in the “Testing the Catchment Delineation Methods” section of the report, the sources of the NWIS drainage-area estimates are unknown. The three topographic-based delineation methods (New England Methods and Raster Seeding Method) have similar mean NWIS drainage-area error results for most of the drainage-area size categories shown in table 1–9.

The Thiessen Polygon Method cumulative drainage areas have the highest mean drainage-area error of all the methods, 6.41 percent. The mean Thiessen Polygon Method drainage-area error for small drainages (less than 20 mi<sup>2</sup>) is 18.5 percent; other methods have errors of less than 5 percent.

## Summary of Results for South Platte Headwaters

For the evaluation of the South Platte Headwaters Subbasin, 55 draft WBD Subwatersheds were used for verification in the CAC analysis. Cumulative drainage-area errors also were derived using published drainage-area values for 15 streamgages obtained from the NWIS database.

Several source data factors affected the analysis. Two isolated stream networks with a total associated drainage area of 11.39 mi<sup>2</sup> had minimal effects on the drainage area at many downstream Subwatersheds and at the outlet of the Subbasin. NHD flowlines crossing divides in the WBD Subwatershed data also were observed. In some cases, the NHD flowlines meandered across the WBD divides in error, whereas in other cases the draft WBD delineations at these intersections were in error.

The New England Methods and the Raster Seeding Method provided similar results—having high spatial correspondence with the WBD and acceptable drainage-area errors for various drainage-area size categories. For the NEM-without-WBD and Raster Seeding Method, only three boundaries for each method (5 percent) have normalized CAC scores less than 80, and they are the same units in each of the two methods. The catchment delineations from the Raster Seeding Method have a 75 percent spatial correspondence with the catchments from the NEM-with-WBD.

The Subwatershed for the Antero Reservoir had the lowest NEM-without-WBD and Raster Seeding Method CAC scores. A determination could not be made as to which method (WBD or NEM-without-WBD) provided the more accurate boundary for the largest area (9.3 mi<sup>2</sup>) of difference between the datasets. The Raster Seeding Method boundary is definitely in error for this same location, as severe parallel-to-stream boundary delineations occurred at artificial path NHD flowlines located within the Antero Reservoir. Other parallel-to-stream boundary delineation errors from DEM flow-path displacement were observed throughout the Subbasin in the Raster Seeding Method boundaries.

The Thiessen Polygon Method, as expected, had the lowest spatial correspondence with the WBD Subwatershed boundaries as indicated by the CAC analysis. The Thiessen Polygon Method is the only method in which no individual boundaries have a normalized CAC score greater than 90. The Thiessen Polygon Method has the highest number of boundaries (64 percent) with normalized CAC scores less than 80. The method also produced the highest normalized polygon-area error in the WBD comparison analysis. In

the WBD and NWIS streamgage cumulative drainage-area analysis, the Thiessen Polygon Method boundaries had the highest mean drainage-area errors. The Thiessen Polygon Method produced especially high errors for drainage areas less than 20 mi<sup>2</sup> in size.

## Pine Valley, Utah (16020302)

No streamgage data were available for the Pine Valley Subbasin; therefore, the comparative analysis for this Subbasin entails the use of the draft WBD only. For this Subbasin, available 1:24,000-scale draft WBD Subwatershed boundaries, delineated to the 12-digit hydrologic unit level (fig. 1–24), were used in the CAC and cumulative drainage-area analyses. Nine WBD boundaries were used in the CAC analysis, the results of which are shown in figure 1–25 and table 1–10. Mean cumulative drainage-area errors derived from the WBD data are summarized in table 1–11. The results of the convergence analysis, in which drainage-area errors were plotted for all methods tested, are shown in figure 1–26. No cumulative drainage areas met the largest drainage size criteria in the summary results. No outlet for the Subbasin exists; it is a closed system with all streams in the Subbasin draining toward a salt flat called the Pine Valley Hardpan.

### New England Method with WBD (NEM-with-WBD)

WBD/NHD-flowline outlet mismatches and an isolated NHD stream contributed to the lowering of the CAC scores (before normalizing to 100) of the NEM-with-WBD. Two NHD flowlines were assigned incorrectly to the WBD Subwatershed boundaries, causing some of the mismatch errors.

The effects of an isolated NHD stream on the results of the CAC and drainage-area analyses were minimal. The omitted catchment area (0.17 mi<sup>2</sup>) associated with the isolated flowline contributed to a cumulative drainage-area error of just 0.26 percent. The CAC score of 96.8 for the Subwatershed with the isolated stream was also affected by errors from a WBD/NHD-flowline outlet mismatch. A WBD/NHD-flowline outlet mismatch in another Subwatershed is responsible for a CAC score of 76.7, and a cumulative drainage-area error of 28.98 percent.

### New England Method without WBD (NEM-without-WBD)

The mean normalized CAC score for the NEM-without-WBD is 95.9, indicating a high spatial correspondence to the WBD. All boundaries compared have normalized CAC scores greater than 90. The NEM-without-WBD has the second lowest (after NEM-with-WBD) mean normalized polygon-area error of 1.1 percent. No normalized polygon-area errors are greater than 2 percent. For the Raster Seeding Method, the highest normalized area in error is 7.8 percent, and for the Thiessen Polygon Method, 32.07 percent.

The differences in delineations were minor among the WBD Subwatersheds and the corresponding boundaries of the NEM-without-WBD. On the basis of topographic information from the DRGs, the WBD provided the more accurate delineations in the places where there were differences.

In the cumulative drainage-area analysis, the NEM-without-WBD has a mean error of 3.95 percent. The mean cumulative drainage-area error for drainage areas less than 20 mi<sup>2</sup> in size is 16.85; this is attributed mostly to a WBD/NHD-flowline outlet mismatch at one Subwatershed.

### Raster Seeding Method

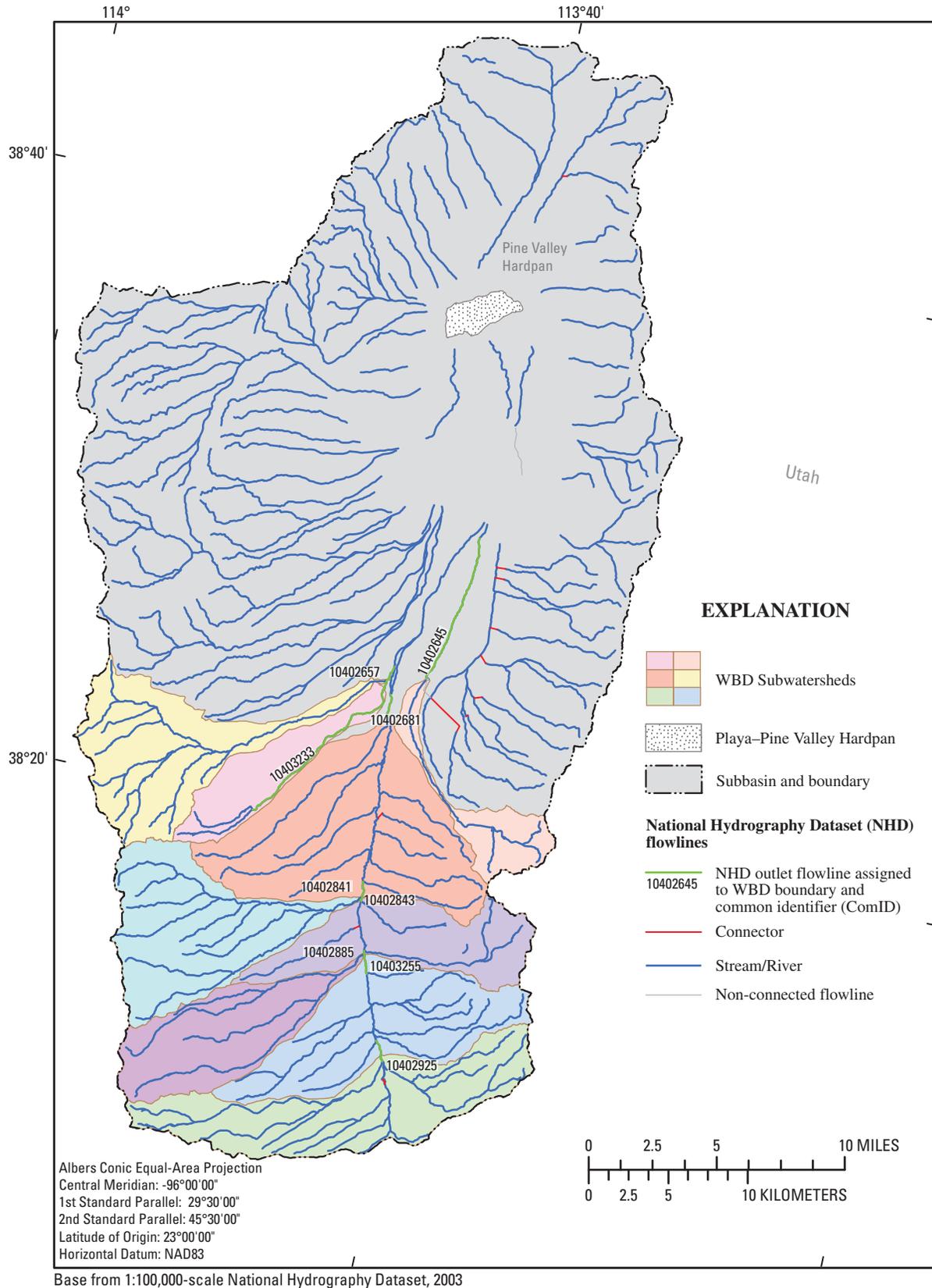
The Raster Seeding Method also had a high spatial correspondence with the WBD boundaries, with a mean normalized CAC score of 96.9. As observed for the two New England Methods, all the Raster Seeding Method boundaries have normalized CAC scores greater than 90. The mean normalized polygon-area error for the Raster Seeding Method is 2.61 percent, the third highest mean error after the two New England Methods at 0 percent (NEM-with-WBD) and 1.1 percent (NEM-without-WBD).

The mean Raster Seeding Method drainage-area error is 2.98 percent, the NEM-with-WBD is 3.38 percent, and NEM-without-WBD is 3.95 percent. For the Raster Seeding Method, some of the drainage-area errors were artificially skewed in a positive direction as a result of DEM flow-path displacement errors at the outlet flowline where WBD/NHD-flowline outlet mismatches occurred. Overall, the Raster Seeding Method produced good drainage-area delineations tracking closely to those of the New England Methods.

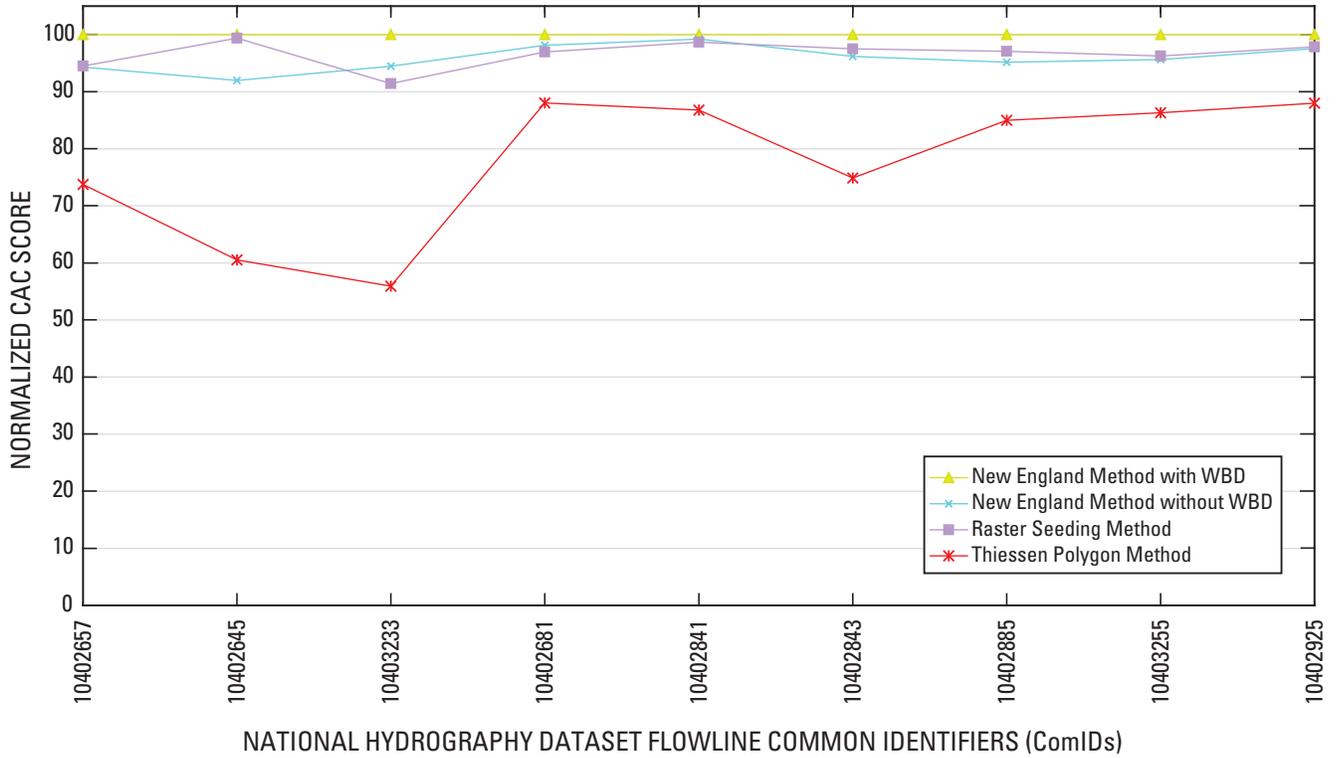
### Thiessen Polygon Method

The Thiessen Polygon Method produced boundaries that had, as expected, the least spatial correspondence to the areas of the WBD Subwatershed boundaries with a mean normalized CAC score of 77.7. Unlike the other methods, for which 100 percent of the boundaries have normalized CAC scores greater than 90, the Thiessen Polygon Method has none. Forty-four percent of the Thiessen Polygon boundaries have normalized CAC scores of less than 80. The Thiessen Polygon Method has the highest mean normalized polygon-area error of 10.04 percent and the highest normalized polygon-area error of 32.07 percent.

The Thiessen Polygon Method also gave the least accurate results in terms of delineated drainage area. In the cumulative drainage-area analysis, the Thiessen Polygon Method has the highest mean error at 10.49 percent (about 3 percent of that is due to WBD/NHD-flowline outlet mismatches) and the highest mean errors for all drainage-area size categories tested. Errors are especially high for drainage areas less than 60 mi<sup>2</sup> in size (table 1–11).



**Figure 1–24.** The draft Watershed Boundary Dataset (WBD) Subwatershed boundaries, and National Hydrography Dataset (NHD) flowlines, Pine Valley, Utah. (The location of this Subbasin is shown in figure 15.)



**Figure 1–25.** Normalized Coefficient of Areal Correspondence (CAC) scores, by National Hydrography Dataset (NHD) flowline common identifier, for all methods tested, Pine Valley Subbasin, Utah.

**Table 1–10.** Mean normalized Coefficient of Areal Correspondence scores by delineation method for Watershed Boundary Dataset Subwatersheds in the Pine Valley Subbasin, Utah.

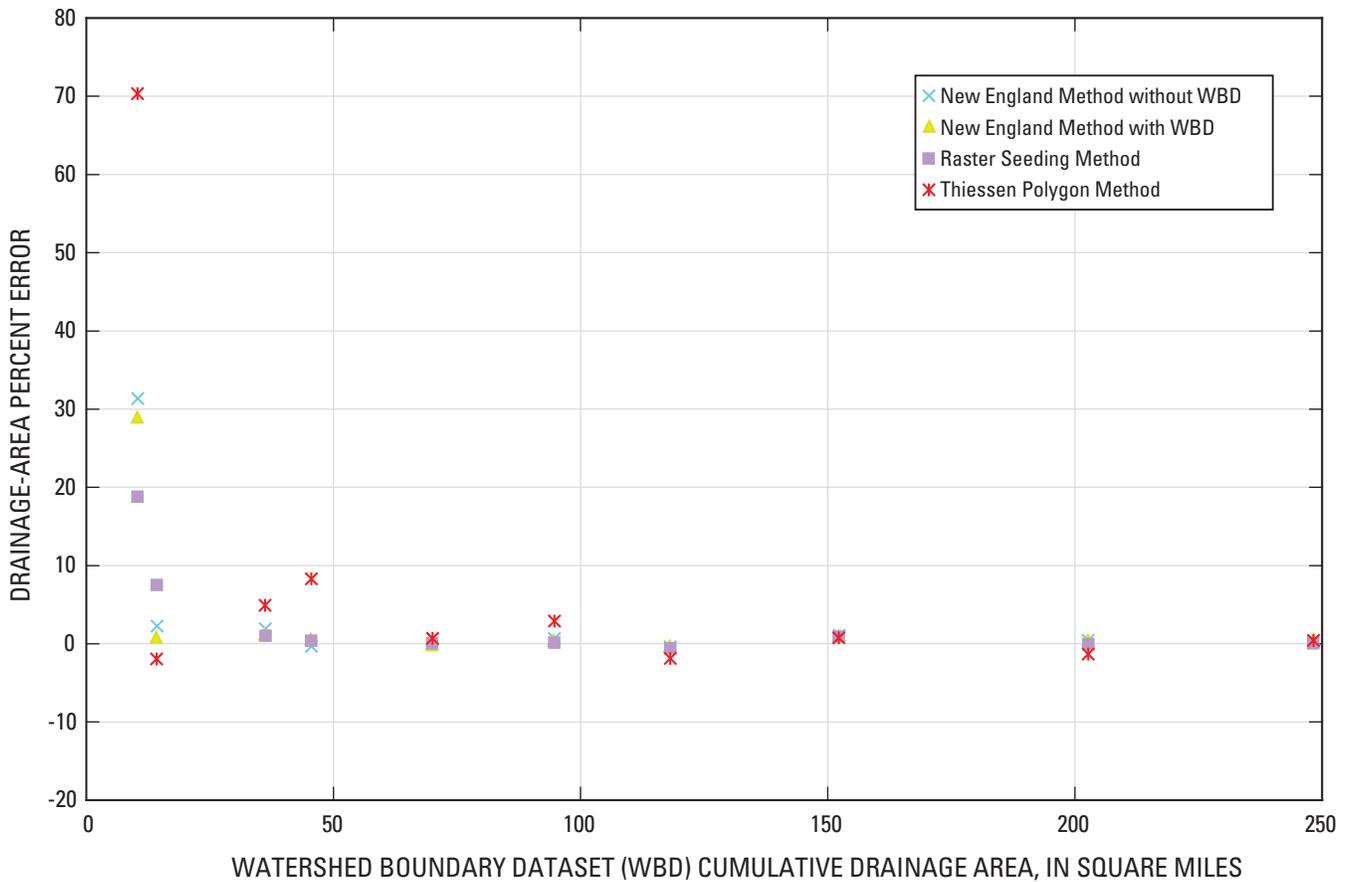
[CAC, Coefficient of Areal Correspondence; NEM, New England Method; WBD, Watershed Boundary Dataset]

Method	Mean normalized CAC	Normalized CACs greater than 90		Normalized CACs less than 80	
		Number of Subwatersheds	Percentage of Subwatersheds	Number of Subwatersheds	Percentage of Subwatersheds
NEM-with-WBD	100.0	9	100	0	0
NEM-without-WBD	95.9	9	100	0	0
Raster Seeding Method	96.9	9	100	0	0
Thiessen Polygon Method	77.7	0	0	4	44

**Table 1–11.** Mean cumulative drainage-area errors for all delineation methods for various drainage-area sizes, Pine Valley Subbasin, Utah.

[mi<sup>2</sup>, square miles; NEM, New England Method; WBD, Watershed Boundary Dataset; na, not applicable]

Method (10 units total)	Mean cumulative drainage-area error, in percent					
	Drainage-area size category					
	All drainage areas	Less than 20 mi <sup>2</sup>	20–60 mi <sup>2</sup>	Greater than 60–390 mi <sup>2</sup>	Greater than 390 mi <sup>2</sup>	At Subbasin outlet (closed system)
NEM-with-WBD	3.38	14.90	1.46	0.31	na	na
NEM-without-WBD	3.95	16.85	1.81	0.49	na	na
Raster Seeding Method	2.98	13.28	1.27	0.22	na	na
Thiessen Polygon Method	10.49	36.16	8.94	1.46	na	na



**Figure 1–26.** Drainage-area percent error in relation to Watershed Boundary Dataset cumulative drainage areas, for all methods tested, Pine Valley Subbasin, Utah.

## Willapa Bay, WA (171001106)

No draft WBD data were available for the Willapa Bay Subbasin; therefore, the comparative analysis for this Subbasin entails the use of the streamgage drainage-area boundaries and NWIS drainage-areas measurements. Twelve streamgage drainage-area boundaries were obtained for the evaluation. Published drainage-area measurements from the NWIS database also were obtained for the 12 streamgages. The streamgage drainage areas, shown in figure 1–27, range in size from 3.97 to 219 mi<sup>2</sup>.

After detailed review of the verification streamgage drainage-area boundaries with the DEM-based delineations of the New England Methods and the Raster Seeding Method, it became apparent that the verification data were not completely accurate. The boundaries from the verification dataset, confirmed with the DRGs, provided less accurate delineations where differences were observed than were obtained for the other study areas. The delineation differences were not substantial, but where differences were noticeable, the DEM based-delineations were more accurate. Also, the verification boundaries from the streamgage data did not edgemark correctly with each other, further calling their accuracy into question.

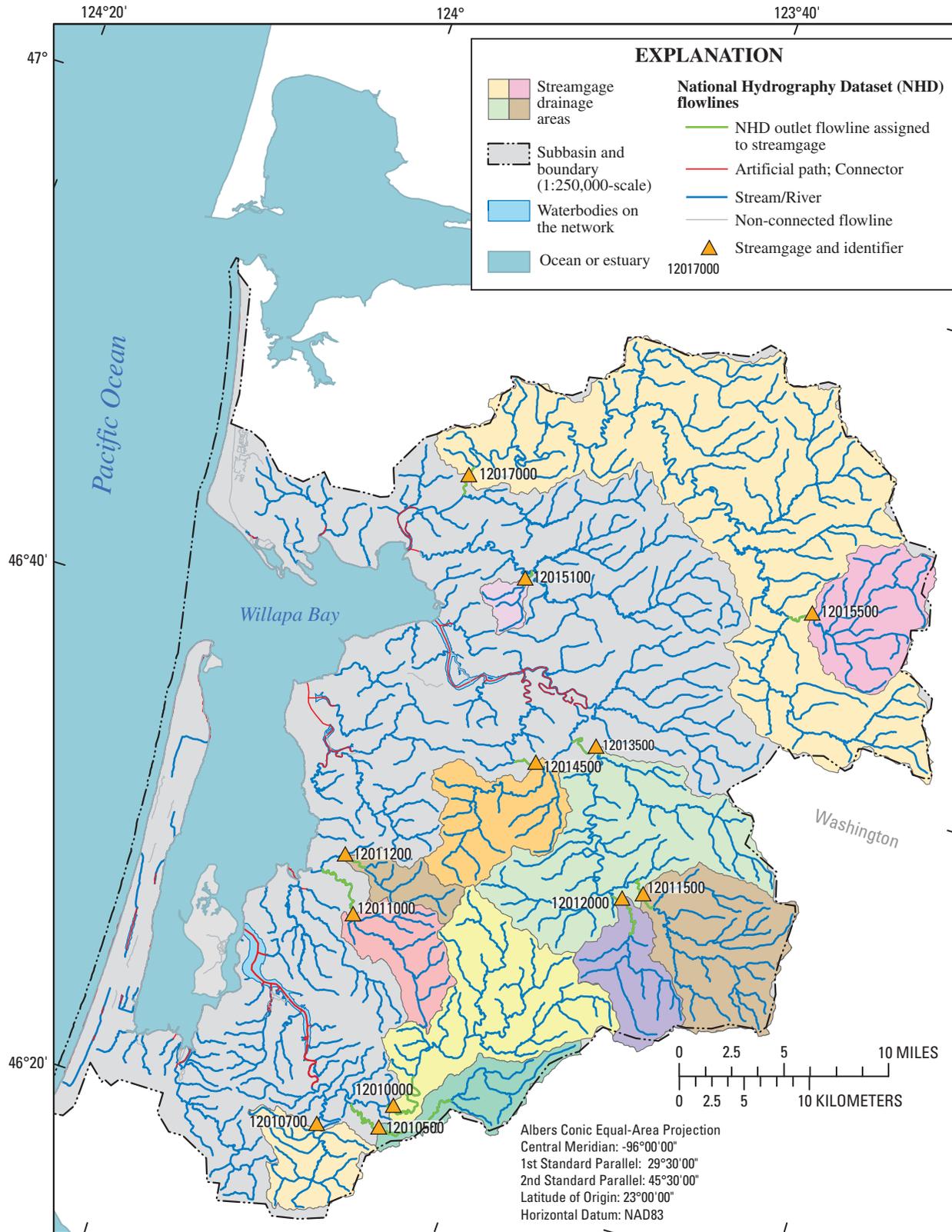
The CAC scores could not be normalized for this Subbasin because no versions of the WBD were available to produce NEM-with-WBD. The CAC scores for the streamgages computed for the NEM-without-WBD, Raster Seeding Method, and Thiessen Polygon Method, are listed in table 1–12 and plotted in figure 1–28. Drainage-area percent

errors for the methods are listed in table 1–13 and plotted in figure 1–29.

The NEM-without-WBD and Raster Seeding Method provided the most accurate boundaries. The mean CAC score of 94.41 for NEM-without-WBD is slightly higher than that for the Raster Seeding Method, 94.37. Where delineation differences were observed, the DEM-based delineations of the NEM-without-WBD surpassed the quality of the verification data. CAC scores less than 90 and high polygon-area errors for the NEM-without-WBD and Raster Seeding Method are directly related to NHD flowline and streamgage location mismatches.

The Raster Seeding Method also provided excellent delineations, which have a 94-percent spatial correspondence with the streamgage drainage boundaries of the NEM-without-WBD. DEM flow-path displacement errors, inherent with the Raster Seeding Method, are the direct cause for the delineation differences with the NEM-without-WBD. The areas in error from DEM flow-path displacement are not substantial in this study Subbasin, as indicated by the results in table 1–13.

The Thiessen Polygon Method delineations provided the least accurate results of the methods, with many CAC scores less than 90 (see table 1–12). The Thiessen Polygon Method has the highest mean drainage-area in error at 8.82 percent, compared to NEM-without-WBD with 4.45 percent and Raster Seeding Method with 4.44 percent. Excluding streamgages with substantial locational mismatches with the flowline outlet, the Thiessen Polygon Method has the highest single drainage-area percent error at 22.19 percent for streamgage 12010700, compared to 0.33 percent and 0.08 percent for the NEM-without-WBD and Raster Seeding Method, respectively.

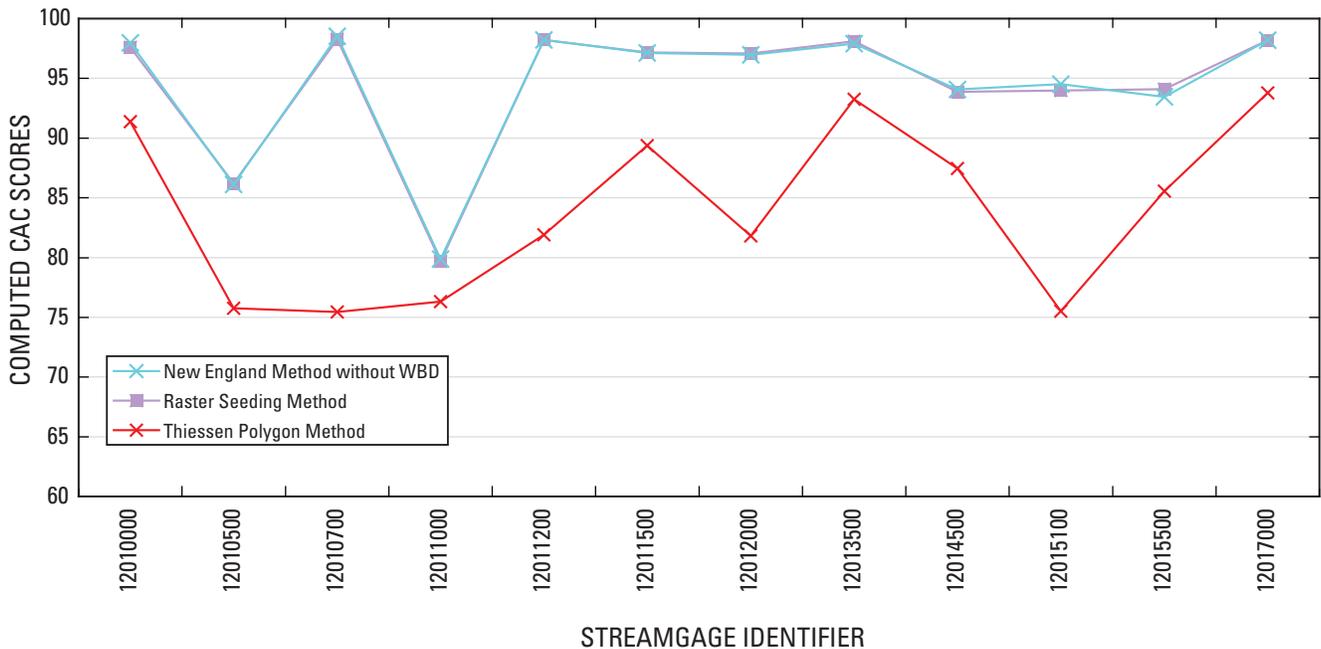


**Figure 1–27.** Location of streamgages and associated drainage areas used as verification datasets, Willapa Bay, WA. (The location of this Subbasin is shown in figure 15.)

**Table 1–12.** Coefficient of Areal Correspondence scores for the New England Method without Watershed Boundary Dataset (NEM-without-WBD), Raster Seeding Method, and Thiessen Polygon Method, for streamgage drainage-areas, Willapa Bay Subbasin, WA.

[ID, Identifier; NHD, National Hydrography Dataset; NEM, New England Method; WBD, Watershed Boundary Dataset; ComID, Common identifier; CAC, Coefficient of Areal Correspondence]

Station ID	NHD flowline ComID	NEM-without-WBD CAC	Raster Seeding Method CAC	Thiessen Polygon Method CAC
12010000	23864404	97.98	97.59	91.37
12010500	23864390	86.12	86.21	75.76
12010700	23865092	98.53	98.29	75.45
12011000	23864488	79.87	79.67	76.31
12011200	23864504	98.22	98.22	81.90
12011500	23864670	97.12	97.17	89.36
12012000	23864656	96.98	97.07	81.82
12013500	23864616	97.90	98.10	93.25
12014500	23864588	94.07	93.87	87.46
12015100	23865380	94.51	93.98	75.50
12015500	23864888	93.44	94.10	85.57
12017000	23864810	98.20	98.19	93.78
	<b>Mean CAC</b>	<b>94.41</b>	<b>94.37</b>	<b>83.96</b>

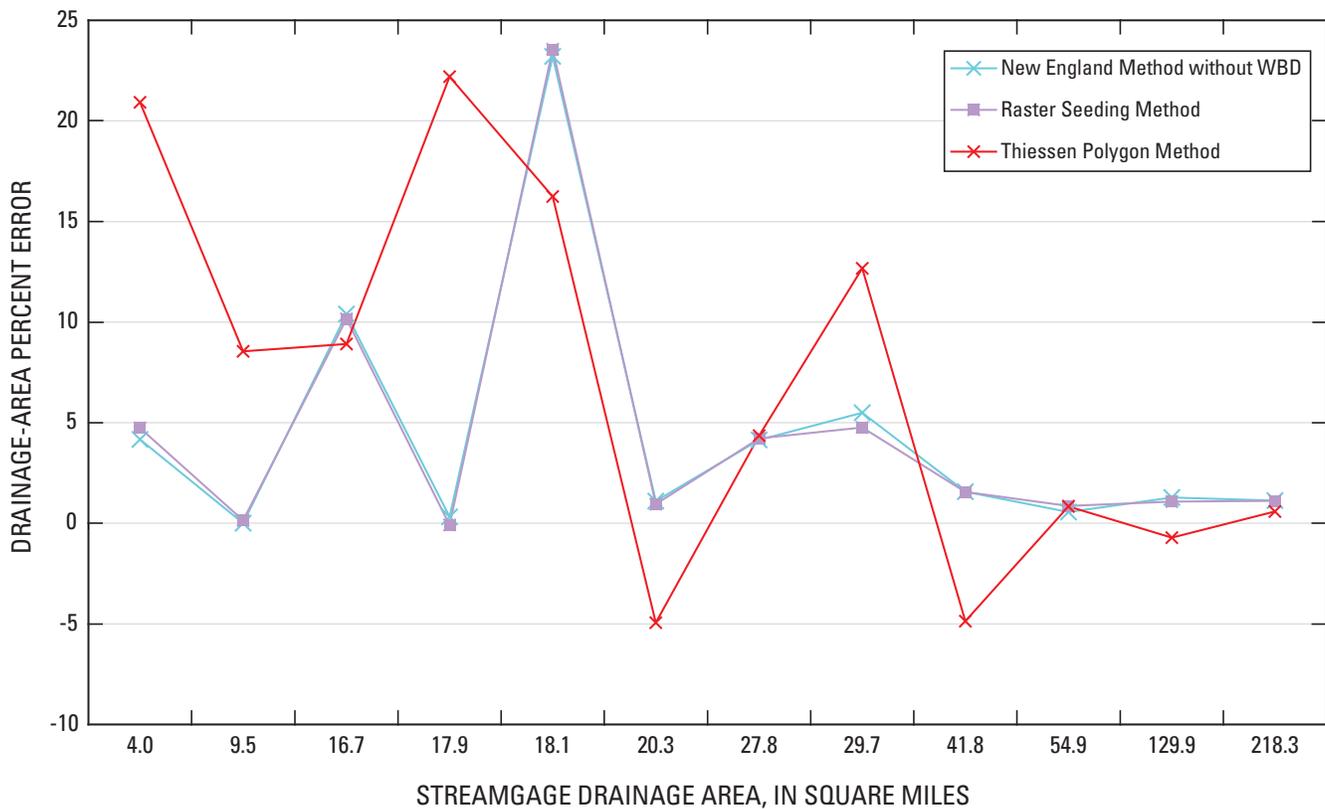


**Figure 1–28.** The computed Coefficient of Areal Correspondence (CAC) scores (not normalized), by streamgage, for the New England Method without Watershed Boundary Dataset (NEM-without-WBD), Raster Seeding Method, and Thiessen Polygon Method, Willapa Bay Subbasin, WA.

**Table 1–13.** Streamgauge drainage-areas of verification data, drainage-area percent errors, and mean errors for the New England Method without Watershed Boundary Dataset, Raster Seeding Method, and Thiessen Polygon Method, Willapa Bay, WA.

[ID, identifier; mi<sup>2</sup>, square miles; GIS, geographic information systems; NEM, New England Method; WBD, Watershed Boundary Dataset]

Station ID	Streamgaging station drainage area (mi <sup>2</sup> ) (from GIS data)	Drainage-area percent error		
		NEM-without-WBD	Raster Seeding Method	Thiessen Polygon Method
12015100	3.97	4.17	4.74	20.92
12011200	9.45	0.01	0.14	8.55
12010500	16.73	10.40	10.16	8.91
12010700	17.89	0.33	0.08	22.19
12011000	18.11	23.21	23.57	16.24
12012000	20.28	1.10	0.95	4.94
12014500	27.80	4.14	4.22	4.36
12015500	29.75	5.49	4.77	12.68
12011500	41.76	1.57	1.56	4.86
12010000	54.90	0.57	0.86	0.84
12013500	129.88	1.28	1.08	0.72
12017000	218.29	1.12	1.12	0.59
	<b>Mean error</b>	<b>4.45</b>	<b>4.44</b>	<b>8.82</b>



**Figure 1–29.** Drainage-area percent error in relation to streamgauge drainage areas of the verification data for the New England Method without Watershed Boundary Dataset (NEM-without-WBD), Raster Seeding Method and Thiessen Polygon Method, Willapa Bay Subbasin, WA.

Prepared by the Pembroke Publishing Service Center.

For more information concerning this report, contact:

Director  
U.S. Geological Survey  
New Hampshire-Vermont Water Science Center  
331 Commerce Way, Suite 2  
Pembroke, NH 03275  
dc\_nh@usgs.gov

or visit our Web site at:  
<http://nh.water.usgs.gov>

