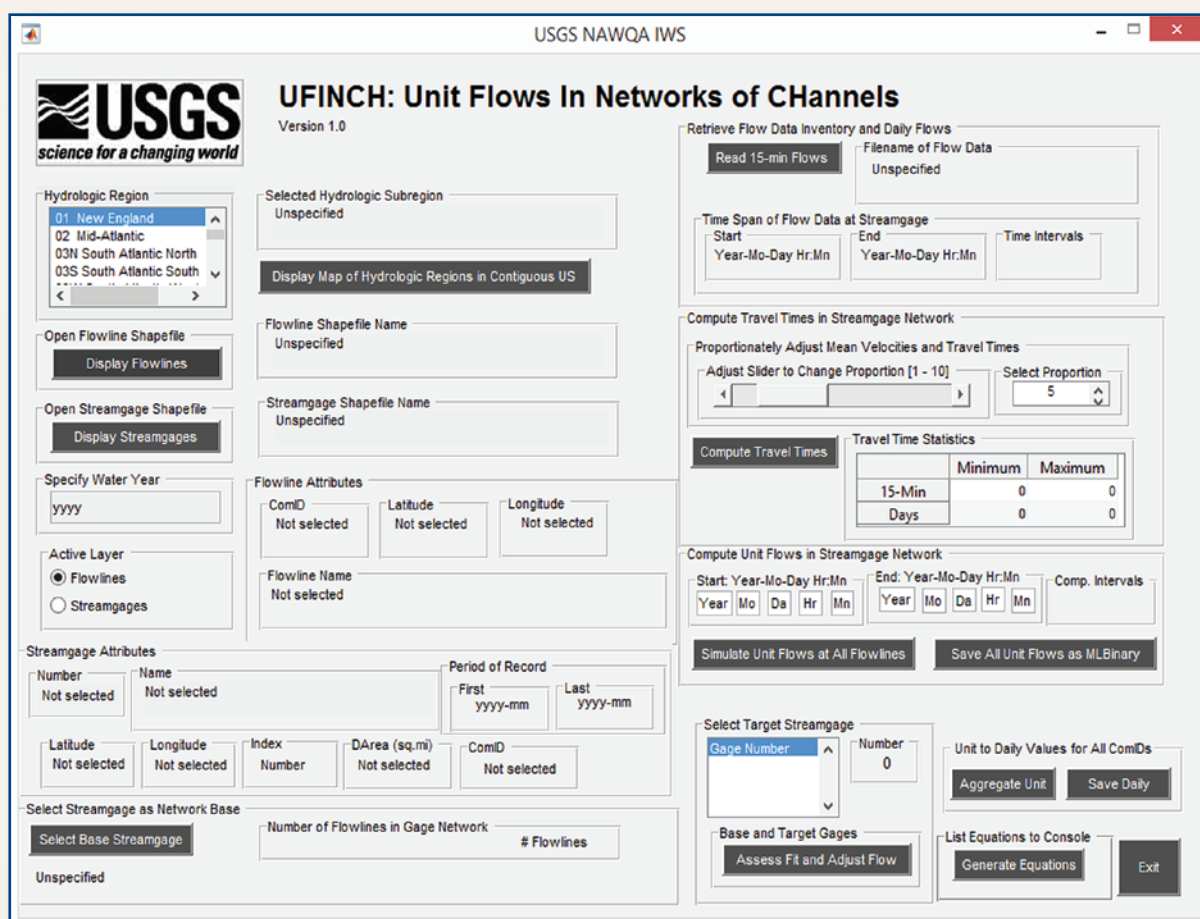# UFINCH—A Method for Simulating Unit and Daily Flows in Networks of Channels Described by NHDPlus Using Continuous Flow Data at USGS Streamgage

Scientific Investigations Report 2016–5074

UFINCH—A Method for Simulating Unit and Daily Flows in Networks of Channels Described by NHDPlus Using Continuous Flow Data at USGS Streamgages

By David J. Holtschlag

National Water Quality Program

Scientific Investigations Report 2016–5074

U.S. Department of the Interior
RYAN K. ZINKE, Secretary

U.S. Geological Survey
James F. Reilly II, Director

U.S. Geological Survey, Reston, Virginia: 2018

Contents

Figures

Conversion Factors

U.S. customary units to International System of Units

Multiply	By	To obtain
Length		
mile (mi)	1.609	kilometer (km)
Area		
square mile (mi^2)	259.0	hectare (ha)
square mile (mi^2)	2.590	square kilometer (km^2)
Volume		
cubic mile (mi^3)	4.168	cubic kilometer (km^3)
Flow rate		
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as follows:

$$°F = (1.8 \times °C) + 32.$$

Temperature in degrees Fahrenheit (°F) may be converted to degrees Celsius (°C) as follows:

$$°C = (°F - 32) / 1.8.$$

Abbreviations

AFINCH   Analysis of Flows in Networks of Channels

ARMA   autoregressive-moving average

ECDF   Empirical Cumulative Distribution Function

HUC   hydrologic unit code

LCLU   land-cover land-use

MCR   Matlab Compiler Runtime

MOVE   Maintenance Of Variance Extension

NAWQA   National Water-Quality Assessment Project

NHD   National Hydrography Dataset

NHDPlus   National Hydrography Dataset with value-added attributes

OLS   ordinary least square

PRISM   Parameter-elevation Regressions on Independent Slopes Model

R^2 coefficient of determination

REG   regression

RMSE   root mean square error

RPN   regression plus noise

UFINCH   Unit Flows in Networks of CHannels

USGS   U.S. Geological Survey

UFINCH—A Method for Simulating Unit and Daily Flows in Networks of Channels Described by NHDPlus Using Continuous Flow Data at USGS Streamgages

By David J. Holtschlag

Abstract

The Unit Flows In Networks of Channels (UFINCH) computer application can be used to simulate daily and unit flows in networks of streams based on geospatial data in the National Hydrography Dataset NHDPlus (with value added attributes), and U.S. Geological Survey daily streamflow data from a downstream (or base) streamgage. Among streamflow augmentation methods, UFINCH has the unique capability to estimate time series of flows from a single base (downstream) streamgage to many upstream reaches, while conserving flows within the basin. UFINCH also provides a simple statistical model to adjust simulated flows to better match continuous flows from data at an upstream streamgage. Parameters of the statistical model are estimated using overlapping periods of record at the two streamgages, but the adjustment can be applied to all years of record available at the base streamgage. This report describes the main features of UFINCH and presents results from a sample application. Interactive graphical user interfaces and automated geographical information processing facilitate flow-data retrievals and provide an intuitive environment for efficient and effective generation of flow information in a network. UFINCH is coded in the Matlab programming language and can be run in the Matlab programming environment, with supporting statistical, optimization, and mapping toolboxes, or from compiled code on a Microsoft Windows computer.

Introduction

The U.S. Geological Survey (USGS) provides publically accessible streamflow information at numerous streamgages in the Nation. Each streamgaged site occurs within a drainage network of interconnected stream segments that receive surface runoff, interflow, and groundwater flow from upstream catchments. This system of catchments and stream segments transport water throughout the basin to the outlet. Without developing a continuous hydrologic process model, there is currently (2016) no simple method for estimating flows in a stream network at a unit or daily time step that conserves flow in a basin, is consistent with measured flows at the outlet, and provides incremental increases in flow between stream segments that is proportional to drainage area increments. Unit Flows In Networks of Channels (UFINCH) provides this capability to support hydrologic and ecologic studies requiring coherent, continuous flow information in a detailed stream network.

In the medium resolution version (2) of the National Hydrography Dataset with value-added attributes (NHDPlus) system (McKay and others, 2015), drainage networks in the Nation have been discretized into approximately 4 million georeferenced catchments and stream segments; thus, each catchment has an average drainage area of about 1 square mile (mi^2). Data in NHDPlus can be used to describe the connectivity between catchment areas and stream segments and among stream segments, specifies flow directions for stream segments, and estimates lengths and velocities in stream segments. Collectively, NHDPlus data provide a geospatial framework for simulating flows in all stream segments of a basin. This framework can be used to help compare or extend streamflow records at an individual upstream streamgage based on measured flows at a downstream streamgage.

Hirsch (1982) compared four streamflow record extension techniques by use of Monte Carlo simulations. The four methods were ordinary least-squares regression (REG), regression plus noise (RPN), and two methods for Maintenance Of Variance Extension (MOVE) techniques (MOVE.1 and MOVE.2), where MOVE.1 is equivalent to a method widely used in psychology and geomorphology that is sometimes referred to as the line of organic correlation (Hirsch, 1982). Hirsch (1982) showed that MOVE.1 estimates were less biased and more accurate than REG and RPN, and that MOVE.2 estimates had marginally lower bias and higher accuracy than the MOVE.1 technique. MOVE.2 uses the same four parameters estimated in MOVE.1, but their estimation is based on more information (Hirsch, 1982). Where outliers may be present in the series, Khalil and Adamowski (2014) indicate that a robust version of the line of organic correlation may be preferable. In this report, MOVE refers generally to the family of Maintenance of Variance Estimators.

Streamflow record extension methods discussed by Hirsch (1982) have been used (Nielsen, 1999; Emerson, 2005) to estimate daily flows at a streamgage having a shorter period of record on the basis of daily values from a streamgage with a longer period of record. Unlike UFINCH, a MOVE estimator cannot be used to estimate flows for ungaged stream segments or provide consistent estimates of flow increments associated with increments in drainage area along a stream or flow network.

In contrast to MOVE as a one-to-one streamgage estimator, UFINCH is a one-to-many estimator, including ungaged sites, which is based on a drainage area ratio approach and on travel times described by NHDPlus attributes. The UFINCH approach provides flow information for all NHDPlus stream segments in a basin and conserves flow within a drainage network. In addition, simulated flows monotonically increase in proportion to increasing downstream drainage area. Where streamflow records are available for upstream segments, a simple statistical corrector can be applied to improve the fit between measured and simulated flows. This corrector can be applied to simulated flows outside the interval of measured flows at the target site. The extended series (measured plus UFINCH simulated and [or] corrected flows) at an upstream target site can be used in subsequent UFINCH applications to propagate flows further upstream in the network. Although UFINCH can be used to estimate flows in an upstream network of stream segments based on flows at a single base streamgage, upstream flow information can minimally be used to confirm basic assumptions underlying the UFINCH model and understand the uncertainty of simulated flows.

Purpose and Scope

The purpose of UFINCH is to help extend unit and daily flow information in stream networks defined within the NHDPlus geospatial framework (McKay and others, 2015) by use of continuous flow information at USGS streamgages. Historically, the USGS has operated about 18,500 streamgages nationwide (Whal and others, 1995). Preliminary testing indicates that UFINCH may be most applicable for simulating unregulated, non-estuarial, flows in basins as large as 3,000 mi^2 in area. Simulated flows can be compared with and adjusted to better match measured flows at upstream streamgages within UFINCH. Adjusted flows have shown close comparisons to measured flows in subbasins that are as much as two orders of magnitude smaller than the basin being simulated. This report documents the capabilities and features of the UFINCH model for simulating unit and daily flows in networks of channels described by NHDPlus based on continuous flows at USGS streamgages.

Methodology

Streamflow monitored at streamgaged sites represents accumulated catchment runoff, including overland flow, interflow, or groundwater outflow, which has flowed downstream through a network of channels. As a first approximation, water yields from upstream catchments are assumed to be uniformly distributed across the streamgaged basin in a UFINCH analysis. Basin water yields, in feet per day, are computed as (runoff) streamgaged flow divided by drainage area. Basin water yields are distributed backward in time and applied to individual upstream catchments so that differences in travel times from catchments to the outlet are accounted for in the analysis. Travel times are proportional to average stream segment velocities and cumulative stream segment lengths (as defined in NHDPlus) from individual catchments to the outlet.

To initiate a UFINCH analysis, a streamgage with continuous flow data is selected as the downstream terminus for a stream network of interest. This downstream streamgage at the basin outlet is referred to as the base streamgage, and is used to identify the upstream catchments and interconnected flowlines. Additional streamgages may be upstream in the network, and data from those streamgages provide a basis for comparing and adjusting simulated flows. Any streamgages upstream from the base streamgage may be referred to as target streamgages. Variations in flows at the base streamgage are expected to reflect variations in average water yields throughout the upstream basin. Particular exceptions to plausible selections of base streamgages are streamgages located on estuaries, those affected by significant water regulation, or on large drainage basins.

A UFINCH analysis proceeds by retrieving daily streamflow values for the base streamgage. For USGS streamgages, an automated process has been developed for retrieving a data inventory from the water data website (waterdata.usgs.gov/nwis/inventory/) and daily flows data from the water services website (waterservices.usgs.gov). The daily flow series are automatically interpolated to 15-minute (unit) value series and divided by drainage area at the base streamgage to form a time-series of basin-wide-average water yields at unit (15-minute) time steps. Interpolated daily flows are used as the default for UFINCH simulations because unit flow variations often reflect effects of local factors that are not representative of fluctuations in water yields throughout the contributing basin.

Historically daily flows, rather than unit flows, have been considered as the primary flow record; thus, unit flows have not been the focus of quality control to the extent that daily flows have been. In addition, unit flows are commonly unavailable for periods of missing (stage) record, or for periods when flows are ice affected. Flows are interpolated using "shape-preserving piecewise cubic interpolation" (http://www.mathworks.com/help/matlab/ref/interp1.html, accessed July 29, 2015) referred to as "pchip" interpolation in Matlab (The MathWorks, Inc., 2015a). This type of interpolation maintains

the appearance of differentiability in daily intervals while constraining interpolated values to the range of bounding daily values. Alternatively, the user may provide a data file of measured unit flows for UFINCH simulations or develop unit flow estimates that are consistent with stochastic variations in unit flow data.

Unit intervals were chosen for UFINCH simulation time steps because the minimum travel time in most stream segments, based on NHDPlus flowline lengths and average flow velocities, is about 15 minutes. UFINCH computes the travel times from the base streamgage to all upstream catchments and applies the unit water yields to these catchments with appropriate lead times so that flow through the network arrives synchronously at the base streamgage.

Spatial variations in water yields commonly occur in natural basins because of persistent differences in the spatial distribution of rainfall patterns, land-cover land-use characteristics, or other factors. Thus, simulation of catchment flows based on average basin water yields may produce simulated flows that systematically differ from measured flows at some target streamgages. UFINCH has two methods for adjusting simulated flows to better match measured flows at upstream streamgages. The static method approximates the discrepancy between ($\log_{10}$ transformed) measured and simulated flows as a regression equation with possible linear and quadratic terms of $\log_{10}$ simulated flows. Parameters for these regression equations can be estimated by use of ordinary least squares or robust techniques (The MathWorks, 2015b). When the parameters of the regression equation are statistically significant, predicted values from this regression can be added to simulated flows to improve the match with measured flows. The regression method is parsimonious, which may facilitate regionalization, and can readily be applied to simulated values during unmeasured periods at the target streamgages to synthetically extend streamflow records.

Alternatively, the Empirical Cumulative Distribution Function (ECDF) method provides a more flexible, but less parsimonious, approach than static regression for improving the match between measured and simulated flow characteristics. The ECDF method computes sample cumulative probabilities and corresponding quantiles for measured and simulated flows. These sample cumulative probabilities have the same length as the number of 15-minute time steps in the simulation. To reduce the size of this set, measured and simulated quantiles are interpolated at 0.001 intervals of probability from 0 to 1 using the sample estimates. A set of 1,000 flow ratios is computed by dividing the interpolated measured quantiles by the interpolated simulated quantiles.

To reduce the irregularities in consecutive ECDF ratios, a double exponential smoothing filter (Kalekar, 2004) is applied to smooth the sequence. A smoothing parameter, which is applied to level components and trend components of the filter and is user specified through a linked edit-text field and slider control on the AssessAdjustGUI graphical user interface in UFINCH. The paired interpolated quantiles of simulated flows and smoothed ratios can be exported to an ASCII file and applied to future simulated flows to provide a flow adjustment during non-overlapping periods of record. Matching the measured and simulated ECDFs may not improve the match between measured and simulated flows, as two white noise sequences can have the same ECDF but remain uncorrelated.

UFINCH can simulate zero flows measured at the base streamgage by applying zero water yields to upstream catchments. In geographic areas where loss of water with downstream distance in a stream reach is common, UFINCH flows would generally underestimate upstream flows. Simulation of flows in stream networks with losing reaches, however, is inconsistent with the assumption implicit in UFINCH simulations that flows at the base streamgage are proportional to, or through adjustments for water use can be made proportional to, water yields throughout the basin. In UFINCH, both the static and ECDF methods for error correction involve simulated flows in the denominator of the corrector; thus, application of these correctors requires positive (no-zero) simulated streamflow values.

UFINCH is designed for application anywhere within the United States based on medium-resolution NHDPlus geospatial data and daily streamflow data available from the USGS. The extensive geospatial dataset supporting the application includes both catchment and streamflow topology, and information on the USGS streamgage network available from the NHDPlus website (http://www.horizon-systems.com/nhdplus/). Streamflow data can be accessed on demand in UFINCH, provided an internet connection is available. Software installation, features, and operation of the UFINCH program and the supporting data are described in the following paragraphs.

Model Assumptions and Limitations

UFINCH is applicable to basins where continuous daily streamflow data are available for an extended period at the basin outlet. A streamgage at the basin outlet is referred to as the base streamgage in this report. Temporal variations in flow at the streamgage are assumed to be generally consistent with temporal variations in water yields throughout the basin. Specifically, basins where daily flows at the base streamgage are commonly affected by artificial controls, substantial (in stream) water uses, diversions or flow augmentations, losing stream reaches, or estuary conditions are not consistent with assumptions underlying the UFINCH model, unless the effect of these variations can be abstracted from the flow record at the base streamgage before application of UFINCH. Basins larger than 3,000 mi^2 have not been extensively evaluated but may act similarly to large impoundments that would substantially degrade the relation between outlet flows and catchment water yields. The period of record extension for the upstream flow network is limited to the period of record, either measured or extended by other means, at the outlet.

Software Installation

UFINCH is written in the Matlab programming language (The MathWorks, 2015a) and uses functions from the Statistics and Machine Learning Toolbox (The MathWorks, 2015a), the Optimization Toolbox (The MathWorks, 2015c), the Mapping Toolbox (The MathWorks, 2015d), as well as the base Matlab functions. Computers equipped with this suite of software can run UFINCH from the source code. Alternatively, UFINCH can be run on 64-bit Windows-based computers without the Matlab programming environment by use of an executable version compiled under Matlab R2015a or later.

For distribution within the U.S. Department of the Interior, the file UFinch_Installer_web.exe, which will be shared through the USGS Google Drive, is designed to facilitate this process, but requires administrative privileges for software installation. Once initiated, the installer should automatically download the 2015a version of the Matlab Compiler Runtime (MCR) program from the MathWorks Web site, unless it already is installed. The MCR translates Matlab code to binary instructions that Windows can execute. Once the MCR file is installed, the executable UFINCH program, UFinch.exe, and updates compiled for Matlab 2015a (also shared through the Google Drive), can be run from the Windows command line or desktop icon. Hopple and Holtschlag (2018) provide code, data, executables, and other guidance for using UFINCH on Microsoft Windows computers.

Data Structure

In addition to the UFINCH software, geographic information and flow data are needed to support the application. Given the extensive geographic information system (GIS) data involved in UFINCH applications, the GIS data has been partitioned by hydrologic regions, as defined in NHDPlus. These regional units generally correspond to 21 USGS Hydrologic Regions (fig. 1; U.S. Geological Survey, 2008), although some of the regions have been subdivided in NHDPlus to form 23 regions (Horizon Systems Corporation, 2014). Within the contiguous United States, region 03 South Atlantic-Gulf has been partitioned into 03 North (03N), 03 South (03S), and 03 West (03W) divisions, and Region 10 Missouri has been partitioned into 10 Upper (10U) and 10 Lower (10L) divisions. Also, region 13 Rio Grande, as defined by NHDPlus, has been extended into Mexico. Outside the contiguous United States, region 20 covers Hawaii, and region 21 describes Puerto Rico and the U.S. Virgin Islands. Preparation of NHDPlus data for Alaska (region 19) is in progress. UFINCH GIS data are partitioned into folders consistent with this NHDPlus GIS data structure.

UFINCH uses a relative folder structure for locating and storing files. The UFINCH program is expected to be run from the folder ..\UFinch\UWork, where ..\ can be any folder on a computer system. An additional folder, ..\UFinch\HR00\GIS, is needed to provide supporting GIS data. This data has two basic types: GIS data specific to individual hydrologic units and a national streamgaging dataset. Any or all of the two-digit hydrologic region [xx] datasets can be loaded in subfolders under ..\UFinch\HR00\GIS\. These subfolders are named (HR01, HR02, HR03N, HR03S, HR03W,...,HR21) corresponding to the NHDPlus convention, and can be downloaded from the designated Google Drive. In addition, the streamgage shapefile set also should be downloaded and stored in the GIS folder ..\UFinch\HR00\GIS\Streamgage\.

In the first analysis of any two-digit hydrologic region [xx], UFINCH will create a new subfolder named ..\UFinch\HR[xx]\GIS\. Region-specific GIS data will be copied to this subfolder from the ..\UFinch\HR00\GIS\HR[xx]\ folder. In addition, a FlowData subfolder will be created along with subfolders for storing unit and daily simulated and measured flows. Subsequent analysis will use the region-specific folder to access GIS and flow data. The region-specific data include (1) binary and supporting files of the boundary of the hydrologic region (two-digit) along with subregional (four-digit) boundaries in the shapefile set HUC02wHUC04.*, where the '*' refers to a set of common filename extensions commonly associated with shapefiles (shp, sbn, sbx, shx, prj, dbf); (2) binary and supporting files in flowline shapefile set NHDFlowline.*; (3) text files NHDFlowlineVAA.txt and PlusFlow.txt containing value added attributes describing flowline connectivity, flowline lengths, and corresponding catchment drainage areas; and (4) text file EROM_MA0001.txt of average flow velocity for individual flowlines identified uniquely by ComID. Nationwide GIS data for the USGS streamgage locations and attributes are stored in the binary and supporting shapefile set StreamGageEvent.*. These GIS shapefile sets for all U.S. streamgages are stored in folder ..\HR00\GIS\Streamgage\.

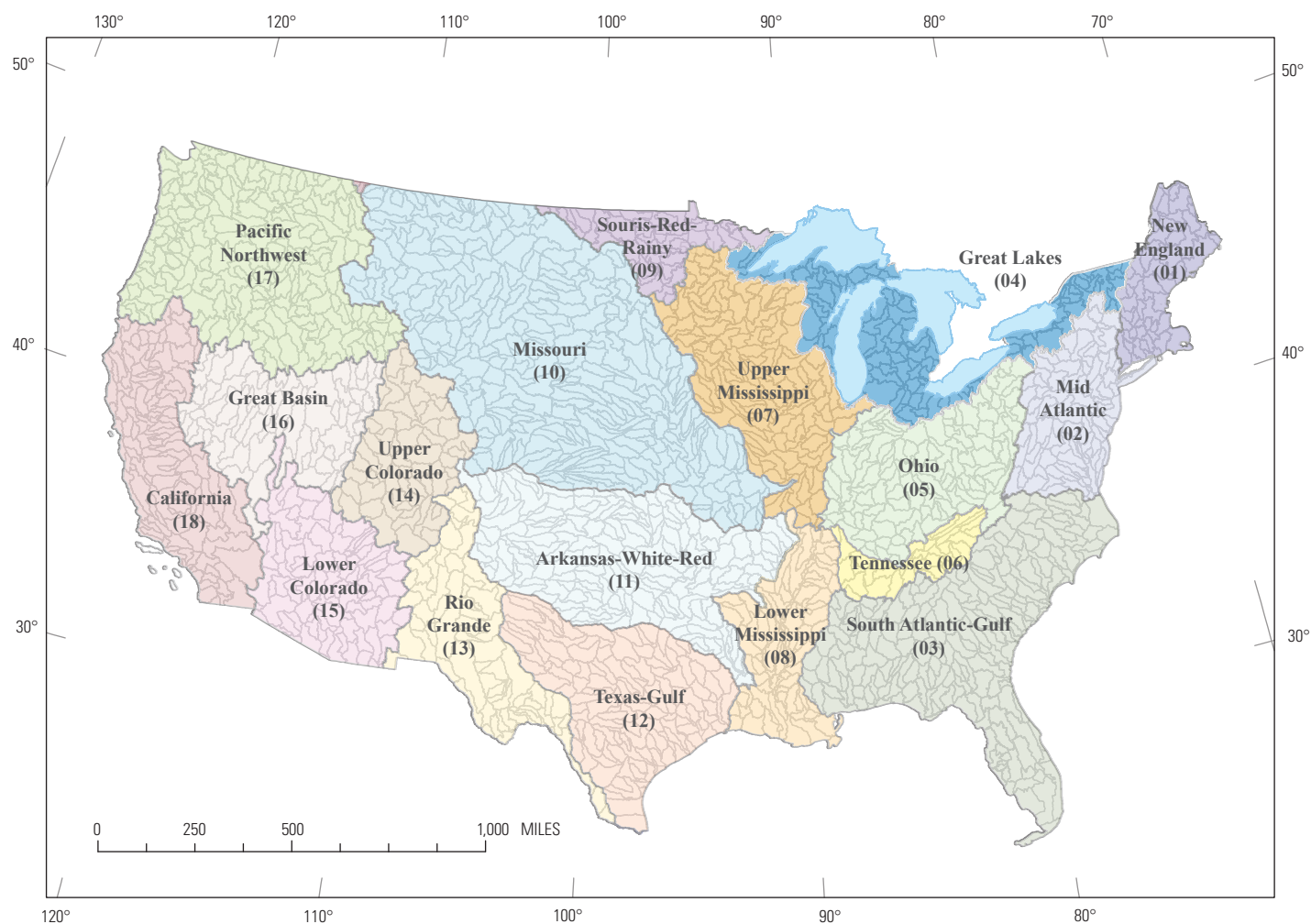**Figure 1.**   Map of two-digit NHDPlus Hydrologic Regions showing two-digit NHDPlus regions for the conterminous United States.

UFINCH Processing and Results

If the UFINCH software is properly installed and the folder HR[xx], for the region of interest, and the Streamgage folder and expected contents are in place at ..\UFinch\HR00\GIS\, the user can initiate a UFINCH analysis. After launching UFINCH, the main GUI is displayed (fig. 2). The main GUI provides access to much of the functionality of UFINCH, including selection of the two-digit hydrologic region for study, selection of the four-digit hydrologic study area, selection of base and target streamgages, and specification of the simulation period. There is a button on the main GUI labeled "Display Map of Hydrologic Regions in Contiguous US" that displays a (nonclickable) map of NHDPlus hydrologic regions for reference (fig. 3).

Once a two-digit [xx] named region is selected from a drop-down list in the upper left-hand corner of the MainGUI, UFINCH will determine if the corresponding folder ..\UFinch\HR[xx] exists. If the folder does not exist, UFINCH will prompt for permission to create the folder structure and copy regional GIS data from ..\UFinch\HR00\GIS\ HR[xx] to populate ..\UFinch\HR[xx]. Otherwise, data in the existing regional folder will be used.

Following selection of the two-digit region, the Region-MapGUI will be displayed showing the names of all hydrologic subregions (fig. 3). The user then selects a four-digit subregion within this region by left clicking on the Region-MapGUI, which causes the corresponding subregion map component to be highlighted in blue. A right click ends the selection process. The RegionMapGUI then closes, and a new SubRegionMapGUI is displayed (fig. 4) showing the outline of the selected subregion. The user then can click the buttons labeled "Display Flowlines" and "Display Streamgages" to show NHDPlus flowlines and streamgages within the selected subregion on the SubRegionMapGUI. To facilitate flowline and streamgage file selection, UFINCH will launch an open file dialog box and prompt for the expected file name of the corresponding coverage in the expected directory location. For standard applications, the user can merely accept the default file location by clicking the "Open" button on the dialog box. Otherwise, the user can navigate the open file dialog box to the location of the custom user-created shapefile.

Figure 2. Main graphical user interface (MainGUI) for UFINCH.
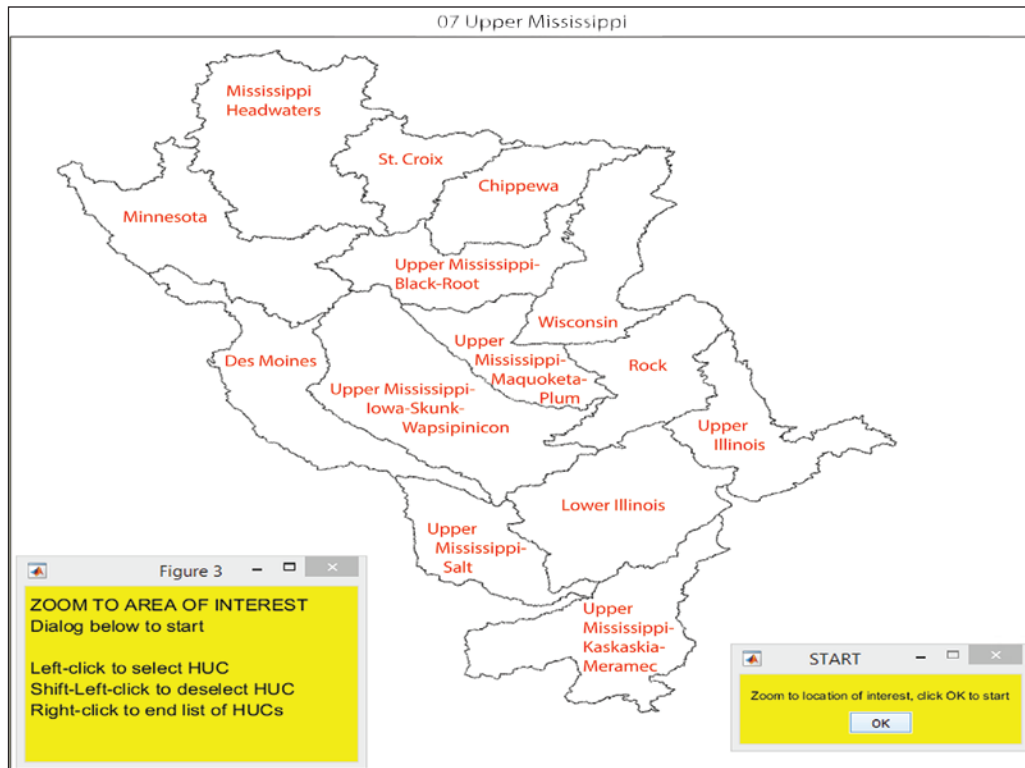
Figure 3. RegionMapGUI for Hydrologic Region 07 Upper Mississippi showing names of hydrologic subregions, with subregion St. Croix selected.
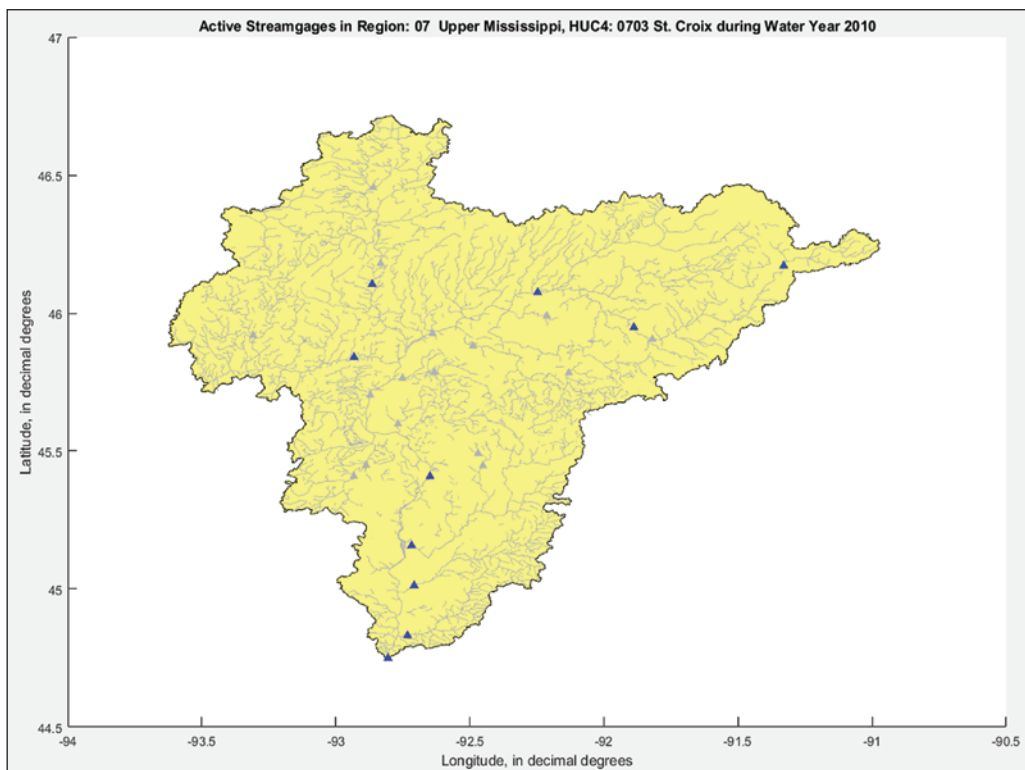

Figure 4. SubRegionMapGUI for Region 07 Upper Mississippi, Subregion 03 St. Croix, showing flowlines and streamgages. (Streamgages active in water year 2010 are highlighted in dark blue).

Flowline shapefiles provide detailed geometry for tens of thousands of flowlines in a region. These files tend to be large, greater than 50 megabytes, and require about a minute to read and display, depending on the performance characteristics of the computer. A message dialog box is displayed indicating this expected time requirement and is closed when the process is complete. Before selecting a base streamgage, the user may use the "Specify Water Year" edit text box to enter a year. Those triangular streamgage symbols displayed on the SubRegionMapGUI are greyed out if the streamgage was not active in the specified year (fig. 4). This feature may be helpful in visually identifying streamgages that were active during a particular time periods.

The SubRegionMapGUI can be used to identify streamgages and record characteristics by left clicking on the triangular symbols representing streamgages. Both blue highlighted and greyed-out symbols are clickable. Attributes of clicked streamgage symbols are displayed on the MainGUI in the "Identify Streamgage Attributes" panel. Attributes include streamgage number and name, first and last date of data collection, latitude, longitude, drainage area, and the ComID, which uniquely identifies the flowline (stream segment) on which the streamgage is located. Once the desired base streamgage is identified, the base streamgage can be selected for analysis by use of the "Select Base Streamgage" button in

the "Select Streamgage as Network Base" panel. Selection of the base streamgage highlights the streamgage and flowline network upstream of the streamgage on the SubRegionMap-GUI (fig. 5). Fields on the MainGUI in the "Select Streamgage as Network Base" panel also are populated, including those identifying the streamgage and the number of flowlines in the network. The drop down list in the mainGUI "Select Target Streamgage" panel also is populated with streamgage numbers for all streamgages at or upstream from the selected base streamgage. The static text field labeled "Number" displays the number of streamgages at or upstream from the base streamgage in the network.

The user may then select the "Read 15-min Flows" button on the MainGUI to retrieve or access unit flow data. Once this button is selected, UFINCH will prompt for the expected name of the flow data file in the expected directory. For example, if base streamgage 05333500 is selected in the St. Croix subregion of the Upper Mississippi region (HR07), UFINCH will prompt for a data file 05333500*.uFloMea in the folder ..\ UFinch\HR07\FlowData\ Measured\Unit.

If no data file has been previously retrieved or created for this streamgage, the user can select the "Cancel" button in the get file dialog box. This action launches a question dialog box with buttons labeled "Retrieve Data," "Reselect Gage," and "Cancel." Selecting the "Retrieve Data" button will initiate an
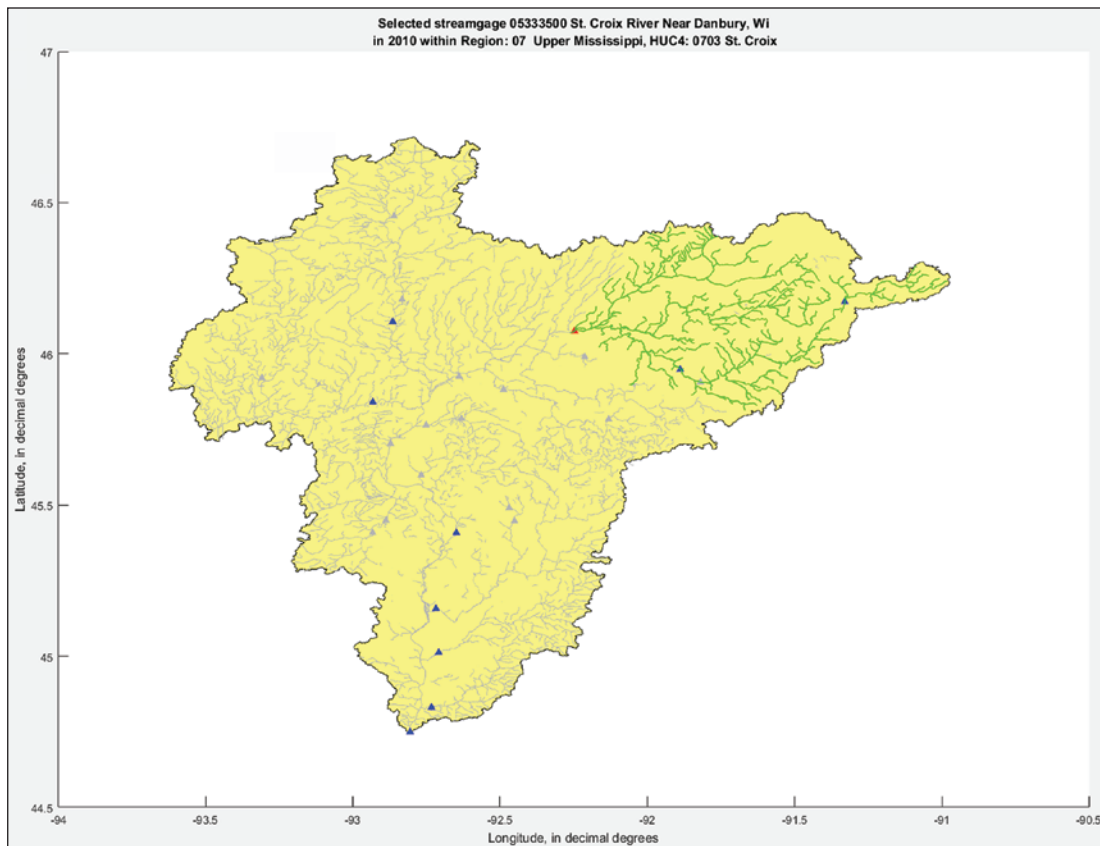

Figure 5.   Hydrologic subregion 0703 with selected streamgage (05333500) St. Croix River near Danbury, Wisconsin, and upstream flow network identified.

input dialog box for the station number (that is populated with the base streamgage number) and two boxes for the start and end date of data retrieval. This date range can be based on an inventory of available data that is initiated from UFINCH and automatically displayed in a browser window returned from an NWIS query (fig. 6). Also, the beginning and ending dates for data retrieval should extend a few days before and after the period for simulation to allow for travel time from the upper catchments of the target basin to the outlet.

The specified streamgage and date range are used in a water services data (http://waterservices.usgs.gov/) retrieval to obtain daily flows. The daily flows are interpolated to 15-minute increments for simulation and are overlain on one or more hydrographs of measured flows in a new figure window. Multiple hydrographs, and corresponding data files, are sometimes returned to span all the continuous periods of record at the streamgage. Multiple periods of continuous record within the streamgaged period are not evident from information displayed in the data inventory. Separate data files are created for each period of continuous record and are labeled to identify the beginning and ending dates of the period. For example, data retrieval for 05331833 from October 1, 1997, to September 30, 2014, created two data files named 05331833_1997Oct01_2001Sep30.uFloMea and 05331833_2005May02_2014Sep30.uFloMea that were stored in folder ..\UFinch\HR07\FlowData\Measured\Unit\.

Once a unit flow file is generated, the user can select the data file for the period of interest using the "Open" button on the file dialog box. This selection causes UFINCH to read in the unit flow data. A hydrograph of this data is displayed in a new figure window (fig. 7), and attributes of the flow data are shown in the panel labeled "Retrieve Flow Data Inventory and Daily Flows" panel in the upper right corner of the MainGUI.

Travel times through the network are computed as the product of flowline length and average flow velocity defined in NHDPlus. Flow velocities in NHDPlus reflect particle velocities in streams. The user can proportionally adjust NHDPlus flow velocities to better reflect the speed of a wave in an open channel (wave celerity), which potentially improves the match between measured and simulated hydrographs. In particular, the user may specify a multiplier to adjust mean flow velocities in each flowline by use of controls in the panel named "Compute Travel Times in Streamgage Network." The default of 1 can be changed by adjusting the linked slider/edit-text box to a preferred value. Based on limited testing, a value of 2.5 often produces a closer match between simulated and measured flows at the base streamgage than a value of 1.0. The default value of 1 may produce simulated hydrographs that are dampened with respect to measured (interpolated) values. The adjusted velocities in UFINCH are not intended to reflect information on particle velocities in streams.
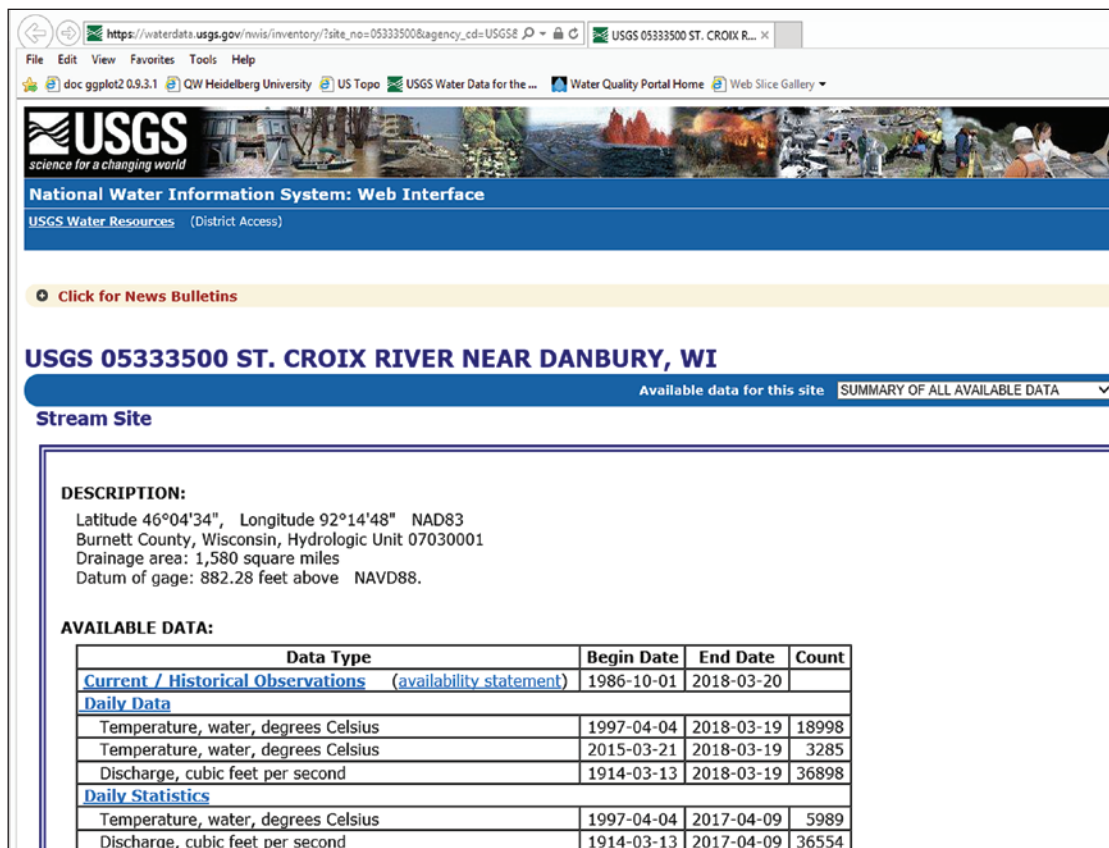

Figure 6. Inventory of data at the selected streamgage (05333500) St. Croix River near Danbury, Wisconsin.
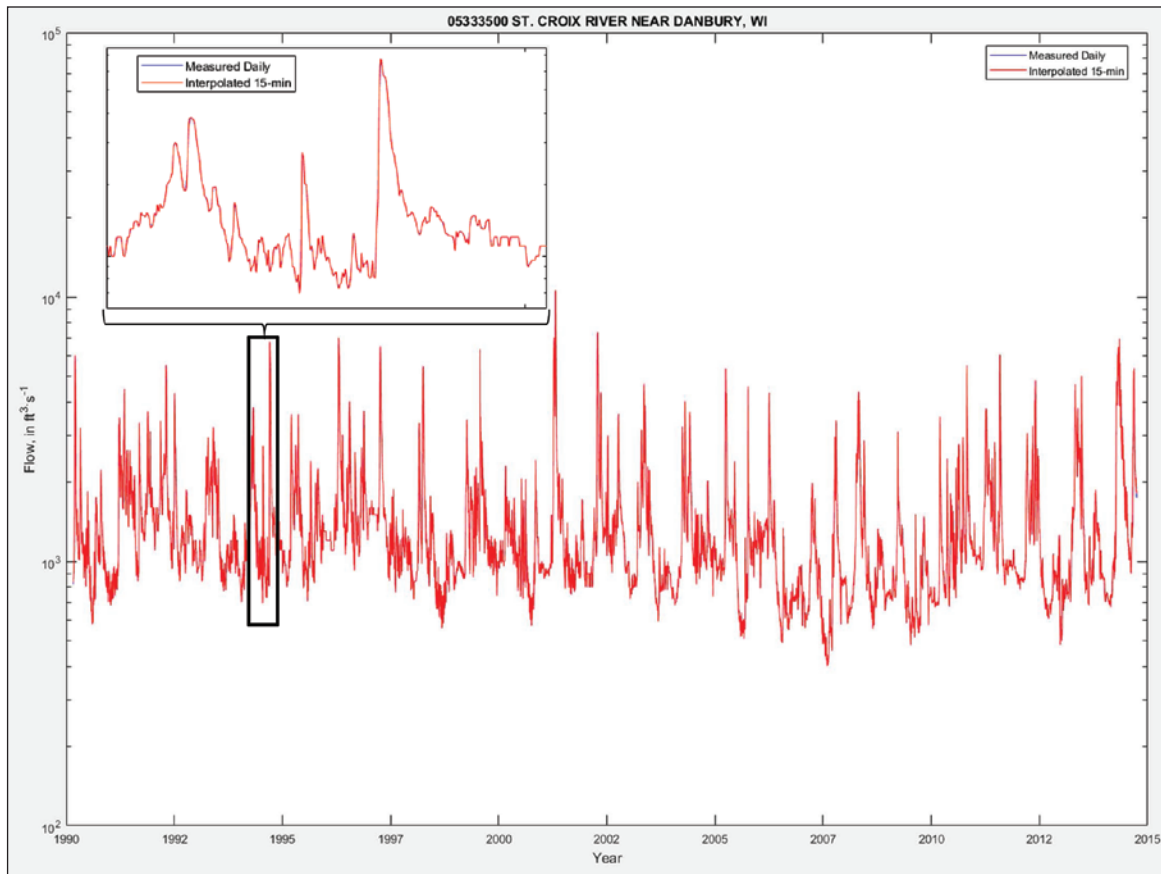
Figure 7.   Daily and interpolated 15-minute flows at streamgage (05333500) St. Croix River near Danbury, Wis.

Once the multiplier is specified, travel times in the network can be computed by selecting the "Compute Travel Times" button. A summary of minimum and maximum travel times in the network in 15-minute time increments, and days are tabulated in the "Travel Time Statistics" table to the right of this button. Travel times determine the apparent lead time needed for water to travel from a catchment to the base streamgage. Basin water yields are computed for each time step on the basis of interpolated flows at the base streamgage. Basin yields are applied to upstream catchments with time leads designed so that corresponding flows appear to arrive synchronously at the base streamgage.

The simulation period can be specified in the "Compute Unit Flows in Streamgage Network" panel. Start- and end-time values are prespecified in the editable text boxes based on period of record in the flow file. The number of computational time steps for the simulation is displayed in a static text box to the right of the simulation period. Computational requirements are proportional to the product of the number of time steps and the number of flowlines, which is displayed in the panel labeled "Number of Flowlines in the Gage Network." For example, the simulation of the highlighted network in figure 5, which has 781 flowlines, 140,161 time steps from October 1, 2010, to September 30, 2014, or about 10 million flowline-time steps, was

completed in 6.5 minutes on a conventional workstation using a single processor. A wait bar is displayed during the simulation to indicate the temporal progress of the computations.

Finally, the user may select a target streamgage to compare measured and simulated flows. The streamgage numbers in the drop down list shown in the "Select Target Streamgage" panel include all streamgages at or upstream from the base streamgage. Selecting a target streamgage from this list highlights the corresponding streamgage on the SubRegionMap-GUI. To confirm the connectivity of the catchments and the flow network for the base streamgage, the user may wish to initially specify the target streamgage as the base streamgage. Connectivity would be confirmed if the average simulated flow is consistent with the average measured flow. Alternatively, if simulated flows are consistently less than measured flows, drainage from some catchment areas may not be represented at the base streamgage. One possible reason for this difficulty is that UFINCH may not have correctly interpreted the network topology.

Selecting the "Assess Fit and Adjust Flow" button in the "Base and Target Gages" panel on the MainGUI will initiate an input file dialog box defaulting to a file name associated with the target streamgage number, commonly an eight-digit value [yyyyyyyy], with an extension for unit flows measured
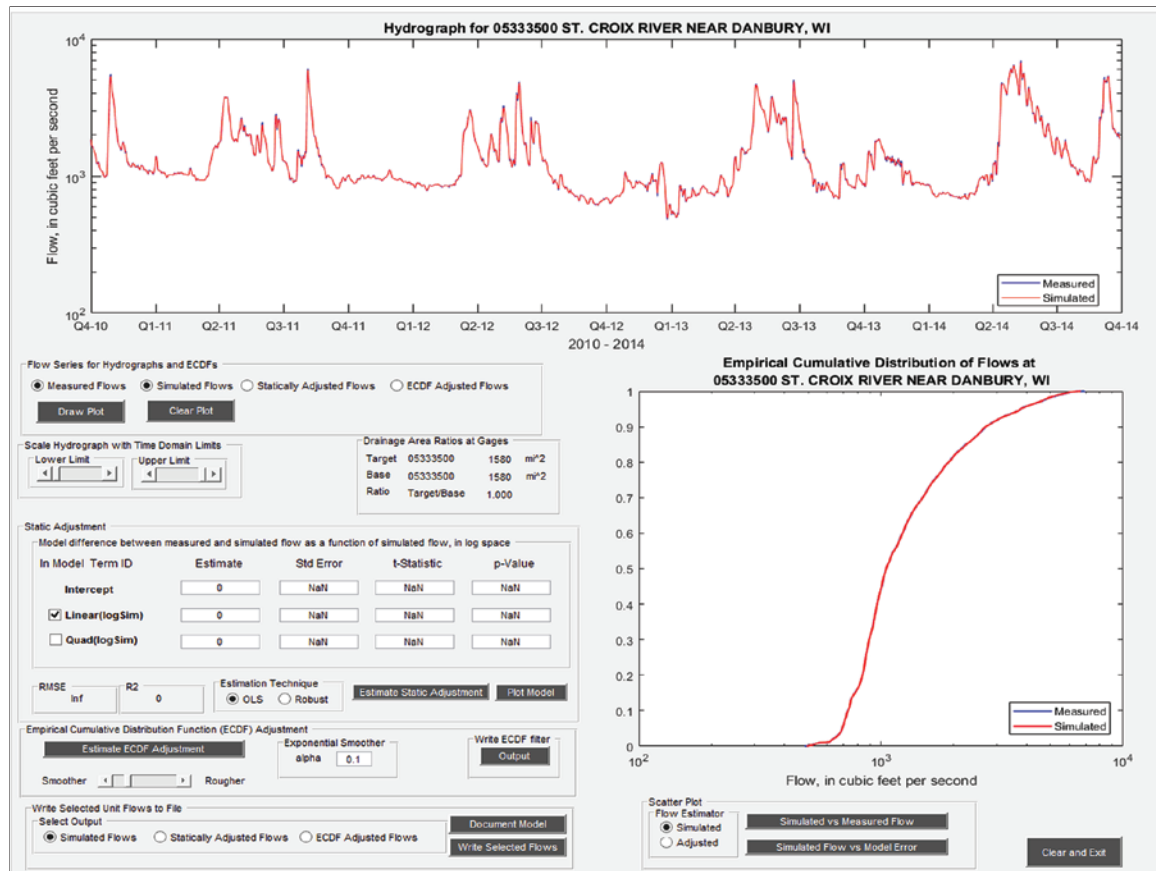
Figure 8. Hydrograph and plot of the empirical cumulative distribution function comparing measured and simulated flows at streamgage (0533500) St. Croix River near Danbury, Wis.

at the target streamgage in folder ..\UFinch\HR[xx]\FlowData\Measured\Unit\. For example, if target streamgage 05333500 St. Croix River near Danbury, Wisconsin, were selected for the target site, the default file name would be 05333500*.uFloMea. The asterisk embedded in the file name would enable display of files with unit flows at the selected site with continuous flows during multiple time periods. Again, if no data files had been previously retrieved for the target site, the user could press "Cancel" in the input file dialog box. Cancelling would again initiate a question dialog box to facilitate retrieving daily flows and interpolating them to unit values. A data inventory would be retrieved to help the user specify the period of record.

Results for the example scenario involving streamgage 05333500 are displayed on the AssessAdjustGUI (fig. 8). Measured and simulated flows are plotted on hydrographs and ECDFs in this figure. In this scenario, careful inspection is required to confirm that both measured and simulated flows are displayed because of their close correspondence. These results appear to confirm the connectivity of the flow network and that flow is conserved in UFINCH. Controls in the "Hydrographs and ECDF Series" panel can be used to clear or select sets of flow series for display. Slider controls in the "Scale Hydrograph with Time Domain Limits" can be used to zoom in on the time axis of the hydrograph. In this case, there

is little potential for improving the match between simulated and measured flows by applying adjustments. Selecting the "Clear and Exit" button in the lower right hand corner of the AssessAdjustGUI will close the figure and return the user to the MainGUI.

Once the connectivity and directionality of the flow network is confirmed, the user may wish to compare simulated and measured flows at one or more upstream streamgages. This process will be illustrated with streamgage (05332500) Namekagon River near Trego, Wis., and streamgage (05331833) Namekagon River at Leonards, Wisconsin. Selecting streamgage number 05332500 in the "Select Target Streamgage" panel displays characteristics of the streamgage and flow record in the "Streamgage Attributes" panel on the MainGUI. The period of record at 05332500 extends from 1928 to 2014, thus providing a temporal overlap with the simulated period. The drainage area at the streamgage is 488 mi^2, or about 31 percent of the area at the base streamgage. Selecting the "Assess Fit and Adjust Flow" button will again provide the opportunity to use an existing unit flow data file or retrieve daily flows. In this scenario, a data file was created from data retrieved for the period October 1, 1990, to September 30, 2014 (05332500_1990Oct_2014Sep30.uFloMea). Selecting this file initiates the AssessAdjustGUI shown in figure 8.
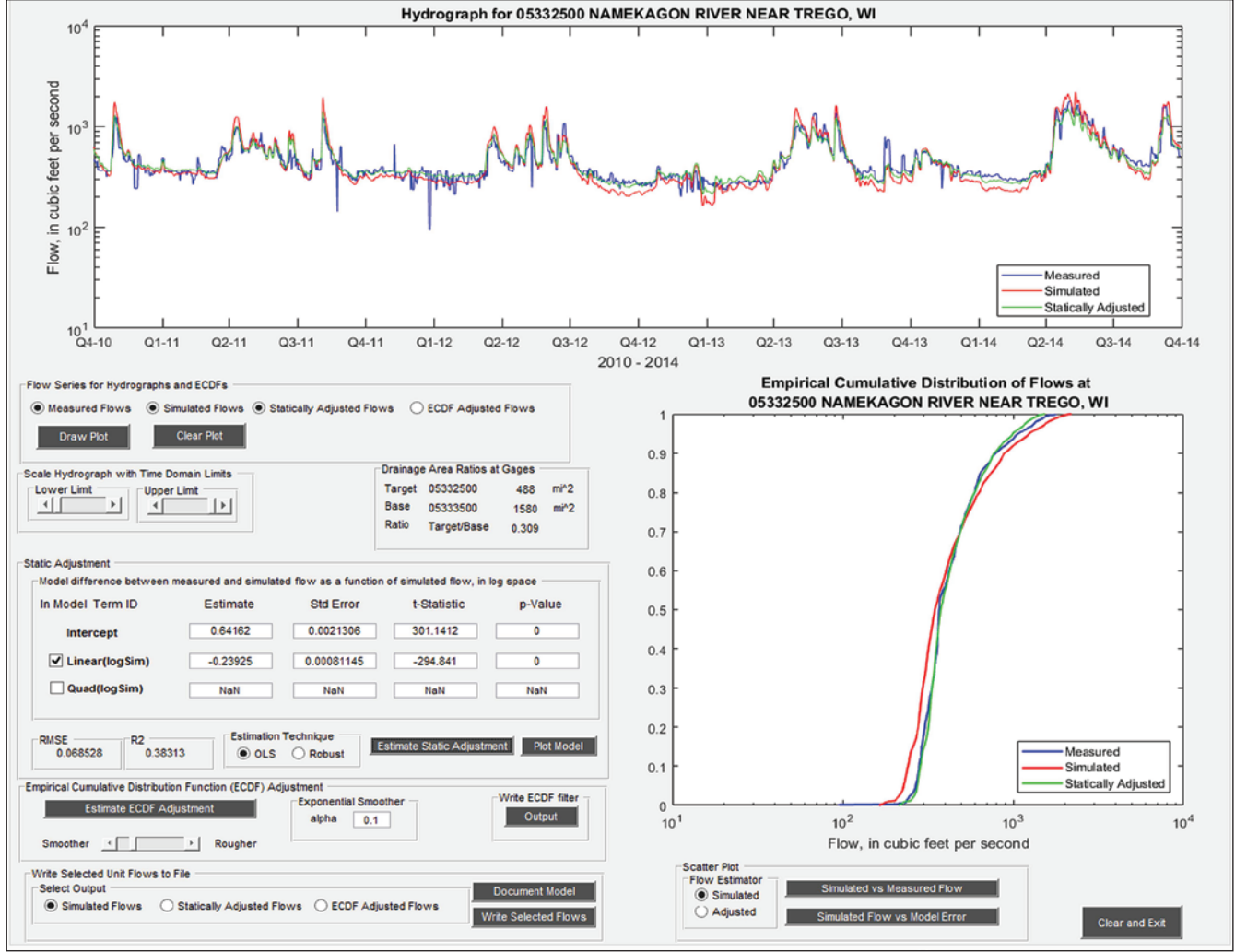
Figure 9. Hydrograph and plot of the empirical cumulative distribution function comparing measured with UFINCH simulated and statically adjusted flows at streamgage (05332500) Namekagon River near Trego, Wisconsin, based on measured flows at streamgage (0533500) St. Croix River near Danbury, Wis.

In this comparison, measured and simulated flows at streamgage 05332500 are quite distinct. The hydrograph indicates that measured flows tend to be more variable than simulated flows. The ECDF indicates that at lower flows, simulated values commonly underestimate measured flows, but that at higher flows simulated values overestimate measured flows. Based on this comparison, a static regression model was developed to adjust simulated flows to better match measured flows. In particular, a constant and linear term was selected from the "Static Adjustment" panel to estimate this adjustment (fig. 9). A quadratic term also can be added to improve estimation in some cases. The form of the error correction is

$$\log_{10} Q_{\varepsilon(t)} = \log_{10}\left(Q_{mea(t)}\right) - \log_{10}\left(Q_{sim(t)}\right),$$

where $Q_{\varepsilon(t)}$ is the modelled error component, and

$$\log_{10} Q_{\varepsilon(t)} = \hat{\beta}_0 + \hat{\beta}_1 \cdot \log_{10}\left(Q_{sim(t)}\right) + \hat{\beta}_2 \cdot \log_{10}\left(Q_{sim(t)}\right)^2 + \delta_t,$$

where $\hat{\beta}_0 + \hat{\beta}_1$, and $\hat{\beta}_2$ are estimated parameters for constant, linear, and quadratic models, respectively.

$$\hat{Q}_{stt(t)} = 10^{\log_{10}\left(Q_{sim(t)}\right) + \log_{10} \hat{Q}_{\varepsilon(t)}},$$

where $\hat{Q}_{stt(t)}$ is the statistically corrected flow estimate.

Estimates of the magnitudes and uncertainties of this model are displayed in the table to the right of the selected parameters. Because of the high autocorrelation expected for unit flows in residual series, the conventional t-Statistics and p-Values shown for the corresponding parameters in the model have only marginal utility. The R^2 (coefficient of determination) value displayed in the panel indicates that the estimated linear function of $\log_{10}$ simulated flows explains about 38 percent of the variability in the residuals between $\log_{10}$ measured and $\log_{10}$ simulated flows. Adjusting the simulated flows by the regression model appears to improve the match in hydrographs and reduce the bias in the ECDFs.

Given the observed improvement in fit, the user may wish to document the model and save the adjusted flow information. To document the model, the user can apply controls in the "Write Selected Unit Flows to File" panel. In this case, the "Statically Adjusted Flows" radio button might be selected. The "Document Model" button launches an input dialog box with editable fields prepopulated with base and target streamgage information (fig. 10). Selecting the "Write Selected Flows" button launches an input file dialog box with a default file name in an expected folder location. In this example, the default folder is ..\UFinch\HR07\FlowData\Simulated\Unit\ and the default file is 05332500_05333500_2010Oct01_2014 Sep27.uFloStt to indicate that this is unit flows (uFlo) that have been statically (Stt) adjusted for target streamgage 05332500 conditioned on flow at base streamgage 05333500 from October 1, 2010, to September 27, 2014.

The contents of the output flow file are shown in figure 11. The file begins with header information that is delimited by an opening /* and closing */, which provides documentation for the analysis but is ignored by UFINCH when the file is read in for possible future simulations. The "Clear and Exit" button on the lower right hand corner of the AssessAdjustGUI returns control to the MainGUI.

The second upstream target streamgage selected for analysis was streamgage (05331833) Namekagon River at Leonards, Wis. The streamgage measures flow from a basin of 126 mi^2, or about 8 percent of the flow at the base streamgage (fig. 12). The hydrograph shows that simulated and measured flows are generally consistent. The ECDF, however, shows that simulated flows are generally biased below measured flows, with this tendency decreasing with higher measured flows. A quadratic term in the regression was not significant in reducing this local bias. In general, the hydrograph shows that the adjustment generally improved the match between statically adjusted simulated flows and measured flows. The robust parameter estimation technique, rather than ordinary least squares, was selected for this analysis by use of controls in the radio button panel labeled "Estimation Technique." In particular, the robust parameters may provide an adjustment that improves the match between measured and statically adjusted flows at higher flow magnitudes. Robust techniques are less sensitive to outliers than ordinary least square techniques. Model document and adjusted simulated unit flow may output as discussed previously before selecting the "Clear and

Figure 10.   Example input dialog box for model documentation.

```
/*¶
Target·gage:·05332500·NAMEKAGON·RIVER·NEAR·TREGO,·WI·at·ComID·1854633.¶
···Drainage·area·of·target·gage·is··488.0·mi^2.¶
Base·gage·for·simulation:·05333500·St.·Croix·River·Near·Danbury,·Wi·at·ComID·2263699¶
···Drainage·area·of·the·base·gage·is·1580.0·mi^2¶
Date·and·time·of·analysis:·2018-Mar-20·09:56·¶
Static·Error·Model·Equation·¶
Model·used·in·Static·Adjustment:¶
¶
Linear·regression·model:¶
····logSimErr·~·1·+·logSimFlow¶
¶
Estimated·Coefficients:¶
···················<strong>Estimate</strong><strong>·······SE····</strong><strong>·····tStat·</strong><strong>····pValue</strong>¶
···················<strong>_____</strong>····<strong>_____</strong>····<strong>_____</strong>····<strong>_____</strong>¶
¶
····<strong>(Intercept)</strong>·····0.64162····0.0021306····301.14····0····¶
····<strong>logSimFlow·</strong>····-0.23925····0.00081145····-294.84····0····¶
¶
¶
Number·of·observations:·139963,·Error·degrees·of·freedom:·139961¶
Root·Mean·Squared·Error:·0.0685¶
R-squared:·0.383,··Adjusted·R-Squared·0.383¶
F-statistic·vs.·constant·model:·8.69e+04,·p-value·=·0¶
Unit·flows·saved:·20-Mar-2018·09:56:27·¶
*/¶
2010,10,1,0,0,594.37¶
2010,10,1,0,15,593.28¶
2010,10,1,0,30,592.19¶
2010,10,1,0,45,591.11¶
2010,10,1,1,0,590.03¶
2010,10,1,1,15,588.96¶
2010,10,1,1,30,587.9¶
2010,10,1,1,45,586.84¶
2010,10,1,2,0,585.79¶
2010,10,1,2,15,584.75¶
2010,10,1,2,30,583.72¶
2010,10,1,2,45,582.7¶
2010,10,1,3,0,581.69¶
2010,10,1,3,15,580.69¶
2010,10,1,3,30,579.71¶
2010,10,1,3,45,578.73¶
2010,10,1,4,0,577.77¶
.¶
.¶
.¶
2014,9,27,22,15,585.25¶
```

Figure 11.   Example unit flow output file following static adjustment.

Exit" button on the AssessAdjustGUI to return to the Main-GUI. Parts of the adjusted simulated unit flow series can be replaced with available measured flows at the target site to form a hybrid series that spans the time interval at both the base and target streamgages. This hybrid series can be used as a base flow series in subsequent UFINCH analysis to propagate measured or adjusted flows further upstream.

The MainGUI provides several minor features that may be of interest. The button labeled "Generate Equations" lists the time-indexed equations that UFINCH uses to route flow from the upper catchments downstream through the flow network to the outlet at the base streamgage. The output is listed to the computer console. This listing may have utility in identifying UFINCH coding limitations for unusual network configurations or NHDPlus network anomalies when flow is apparently not conserved.

The "Aggregate" button in the "Unit to Daily Values for All ComIDs" panel aggregates unit values to daily flows for all flowlines in the network. These simulated daily flows can be saved to an ASCII output file by use of the labeled "Save" button for further analysis. The aggregation may require several moments to complete the computation for large networks and extended simulation periods. Unit flows for all flowlines also can be saved in Matlab binary format by selecting the "Save All Unit Flows as MLBinary," although the large size of this binary tends limit its perceived utility.

The "Exit" button on the MainGUI will end the application and exit to the operating system. Before exiting, the Microsoft notepad application will be launched to display the contents of an automatically generated, time-stamped log file entitled "uFinchSummary[yyyymmmdd_HHMM].txt. This file will contain model documentation for any models documented during the session. Finally, UFINCH will display a question dialog box requesting to close any open figures.
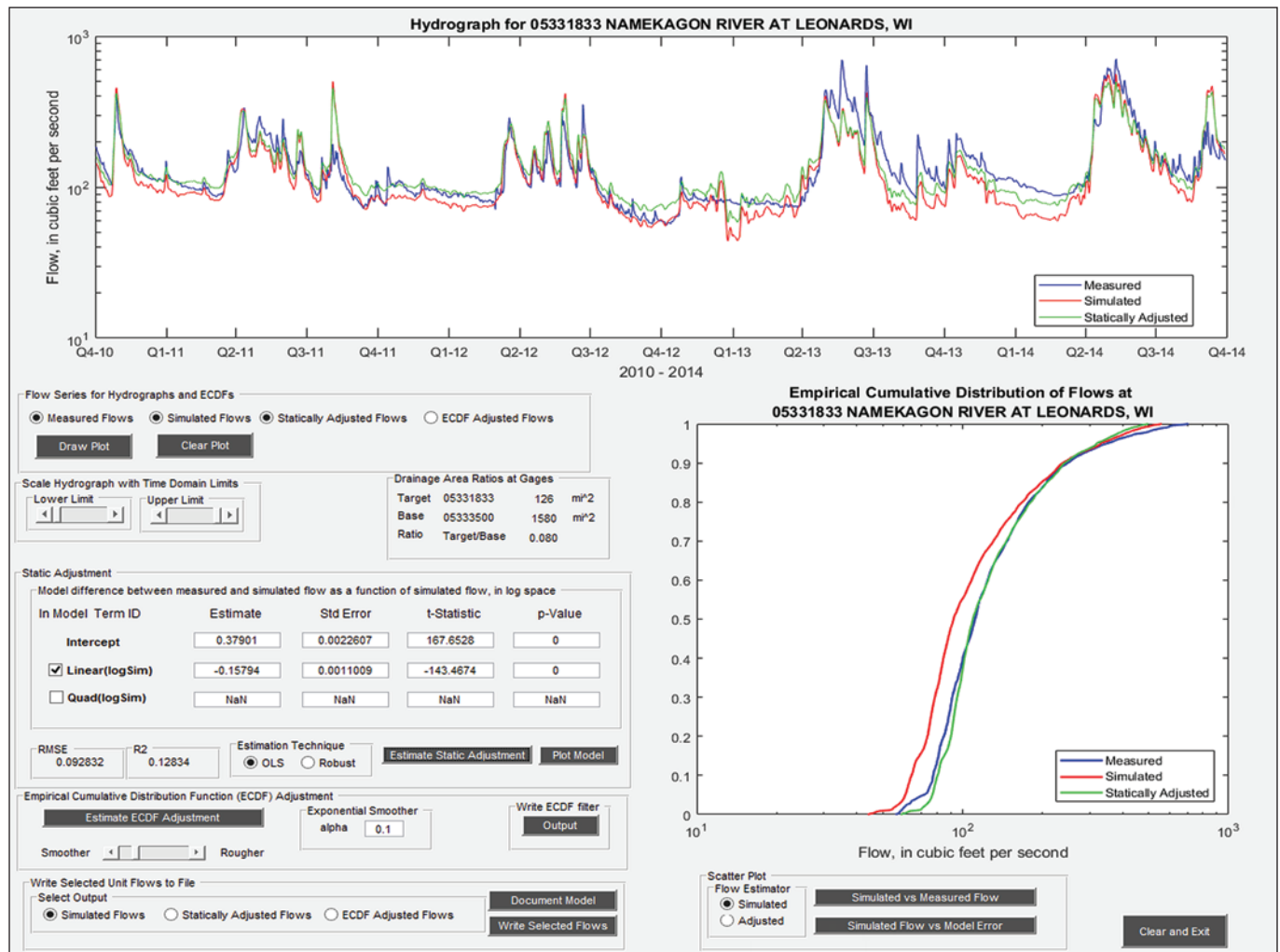
Figure 12. Hydrograph and plot of the empirical cumulative distribution function comparing measured with UFINCH simulated and statically adjusted flows at streamgage (05331833) Namekagon River at Leonards, Wisconsin, based on measured flows at streamgage (0533500) St. Croix River near Danbury, Wis.

Discussion

This report documents the capabilities and features of the UFINCH model for simulating unit and daily flows in networks of channels described by NHDPlus based on continuous flows at USGS streamgages. UFINCH provides a basis for comparing simulated flows with measured flows at upstream streamgages, and designing a filter to improve the match between overlapping periods of record at the base and target streamgages. This filter can be applied to simulated flows during periods when the target streamgage was not operating to synthetically extend the period of record at the target streamgage. By integrating periods of measured and simulated flows at the target streamgage, UFINCH can be reinitialized to use the integrated flows as the base streamgage and progressively extend streamflow information further upstream in the basin.

Because flow is conserved in a basin, a filter designed to improve the match between simulated flows at a base streamgage and measured flows at a target streamgage may have transfer value to the remaining parts of the basin. If a filter systematically increases simulated flows at a target streamgage during low flow periods, then simulated low flows elsewhere in the network may need to be systematically reduced to conserve flows within the entire basin during these periods. In basins with multiple upstream target streamgages, a corresponding set of filters may be able to further improve flow simulations in intervening areas. A set of such filters may provide a basis for assessing how spatial variations in precipitation patterns or land-use, land-cover characteristics within a basin effect flow characteristics.

Future UFINCH enhancements could provide access to simulated unit flows at all flowlines in the network. Currently (2016), individual flowlines on the SubRegionMapGUI are clickable if flowlines are the active layer, rather than the streamgage layer. The "Active Layer" radio button panel on the MainGUI is used to select the active layer, which defaults to the streamgage layer. Characteristics of the selected flowlines currently (2016) can be displayed in the "Flowline Attributes" panel on the MainGUI. A feature for accessing, plotting, and exporting the simulated unit flows may be helpful, although additional feature layers, such as major roads, may be needed to facilitate flowline selection.

Another feature that may have utility in future versions of UFINCH would be the ability to use discrete flow measurements at miscellaneous sites to assess and adjust simulated flows. In particular, the static regression model for adjusting simulated flows is parsimonious, which would allow the 1 to 4 needed model parameters to be estimated from a relatively small number of discrete measurements. Such relatively limited data requirements could be helpful in confirming and improving model simulations for ecological investigations on small streams and other applications.

Integration of UFINCH with AFINCH (Analysis of Flows in Networks of Channels; Holtschlag, 2009) would eliminate the assumption of uniform water yields underlying UFINCH. In particular, AFINCH uses the same NHDPlus network of catchments and flowlines and streamgage network as UFINCH to compute monthly flows. The AFINCH application conditions the spatial distribution of monthly water-yields on monthly climatic data (PRISM Climate Group, 2013), including precipitation and temperature, and selected land-use, land-cover characteristics using a constrained regression approach. Within the U.S. Great Lakes area, monthly AFINCH water yields are precomputed for all NHDPlus catchments from 1951 to 2012 (Luukkonen and others, 2015). Finally, adjustment of measured flows at base or target streamgages with water use data may facilitate the application of UFINCH to regulated streams.

Summary

The UFINCH (Unit Flows In Networks of Channels) computer application can be used to simulate daily and unit flows in networks of streams based on geospatial data in the National Hydrography Dataset NHDPlus (with value added attributes), and U.S. Geological Survey daily streamflow data from a downstream (or base) streamgage. Among streamflow augmentation methods, UFINCH has the unique capability to estimate time series of flows from a single base streamgage to many upstream reaches, while conserving flows within the basin. UFINCH also provides a simple statistical model to adjust simulated flows to better match continuous flows from data at an upstream streamgage. Parameters of the statistical model are estimated using overlapping periods of record at the two streamgages, but the model can be applied to all years of record available at the base streamgage. This report describes the main features of UFINCH and presents results from a sample application.

To simulate unit (15-minute) flows, measured or interpolated unit flows at an outlet (base) streamgage are converted to basin average unit water yields by dividing flows by the area drained at the outlet. Unit yields are then shifted backwards in time based on their expected travel times from catchments to the outlet, using NHDPlus flow velocities and reach lengths for intervening channel segments (NHDPlus flowlines). Time delayed yields are applied to upstream catchments to produce catchment flows, in a process that conserves flows throughout the basin. Catchment flows are then translated downstream through intervening flowlines. Simulated flows can be compared with measured upstream flows and adjusted by use of tools in the UFINCH application to improve the match. Simulations using long-term flow records at the outlet can be used to extend unit flows at upstream flowlines, with adjustments derived from overlapping periods of record.

Interactive graphical user interfaces and automated geographical information processing facilitate flow-data retrievals provide an intuitive environment for efficient and effective generation of flow information in a network. UFINCH is coded in the Matlab programming language and can be run in the Matlab programming environment, with supporting statistical, optimization, and mapping toolboxes, or from compiled code on a Microsoft Windows computer.

Acknowledgments

References Cited

Box, G.E.P., and Jenkins, G.M, 1976, Time series analysis forecasting and control: Oakland, Calif., Holden-Day, 575 p.

Emerson, D.G., 2005, Historic and naturalized monthly streamflow for selected sites in the Red River of the North Basin in North Dakota, Minnesota, and South Dakota, 1931–2001, U.S. Geological Survey Scientific Investigations Report 2005–5092, 228 p., accessed on August 14, 2015, at https://www.usbr.gov/gp/dkao/redriver/rrvwsp/Support/final_usgs_naturalized_flow.pdf.

Holtschlag, D.J., 2009, Application guide for AFINCH (Analysis of Flows In Networks of CHannels) described by NHDPlus: U.S. Geological Survey Scientific Investigations Report 2009–5188, 106 p. [Also available at http://pubs.usgs.gov/sir/2009/5188/.]

Hopple, J.A., and Holtschlag, D.J., 2018, Code, data, executables, and other information used to run Unit Flows in Networks of Channels (UFINCH) —A method for simulating unit and daily flows in networks of channels described by NHDPlus using continuous flow data at U.S. Geological Survey streamgages: U.S. Geological Survey data release, https://doi.org/10.5066/F7319TC5.

Hirsch, R.M., 1982, A comparison of four streamflow record extension techniques: Water Resources Research, v. 18, no. 4, p. 1081–1088

Horizon Systems Corporation, 2014, NHDPlusV2 flow table navigator toolbar: accessed May 27, 2014, at http://www.horizon-systems.com/NHDPlus/NHDPlusV2_tools.php.

Kalekar, P.S., 2004, Time series forecasting using Holt-Winters exponential smoothing: Bombay, India, Kanwal Rekhi School of Information Technology, 13 p., accessed January 6, 2017, at https://c.forex-tsd.com/forum/69/exponentialsmoothing.pdf.

Khalil, Bahaa, and Adamowski, Jan, 2014, Comparison of OLS, ANN, KTRL, KTRL2, RLOC, and MOVE as record-extension techniques for water quality variables: Water Air Soil Pollution, v. 225, 1966, p. 1–19, accessed on January 6, 2017, at http://dx.doi.org/10.1007/s11270-014-1966-1.]

Luukkonen, C.L., Holtschlag, D.J., Reeves, H.W., Hoard, C.J., and Fuller, L.M., 2015, Estimation of monthly water yields and flows for 1951–2012 for the United States portion of the Great Lakes Basin with AFINCH: U.S. Geological Survey Scientific Investigations Report 2014–5192, 83 p. [Also available at http://dx.doi.org/10.3133/sir20145192.]

MaKay, L., Bondelid, T., Dewals, T., Johnston, J., Moore, R., and Rea, A., 2016, NHDPlus Version 2—User Guide (Data Model Version 2.1), Horizon Systems, Inc., 181 p., accessed on January 6, 2017, at ftp://ftp.horizon-systems.com/NHDPlus/NHDPlusV21/Documentation/NHDPlusV2_User_Guide.pdf.

Nielsen, J.P., 1999, Record extension and streamflow statistics for the Pleasant River, Maine: U.S. Geological Survey Water-Resources Investigations Report 99–4078, 22 p. [Also available at http://pubs.usgs.gov/wri/1999/4078/report.pdf.]

The MathWorks, 2015a, Matlab Programming Fundamentals, R2015a: Natick, Maine, 1064 p., accessed June 11, 2015, at http://in.mathworks.com/help/pdf_doc/matlab/matlab_prog.pdf.

The MathWorks, 2015b, Statistics and Machine Learning Toolbox User's Guide: R2015a, Natick, Maine, 6682 p., accessed June 11, 2015, at http://www.mathworks.com/help/pdf_doc/stats/stats.pdf.

The MathWorks, 2015c, Optimization Toolbox User's Guide, R2015a: Natick, Maine, 780 p., accessed June 11, 2015, at http://www.mathworks.com/help/pdf_doc/optim/optim_tb.pdf.

The MathWorks, 2015d, Mapping Toolbox User's Guide, R2015a: Natick, Maine, 836 p., accessed June 11, 2015, at http://www.mathworks.com/help/pdf_doc/map/map_ug.pdf.

PRISM Climate Group, 2013, United States average mean annual precipitation and air temperature: Northwest Alliance for Computational Science and Engineering at Oregon State University, accessed June 10, 2015, at http://www.ocs.orst.edu/prism.

U.S. Geological Survey, 2008, Hydrologic unit maps—What are hydrologic units?, Web site accessed November 24, 2008, at http://water.usgs.gov/GIS/huc.html.

U.S. Geological Survey, 2014, Water webserver team—Hydrologic unit map (Based on data from USGS Water-Supply Paper 2294): Web site accessed June 11, 2015, at http://water.usgs.gov/GIS/regions.html.

Wahl, K.L., Thomas, Jr., W.O., and Hirsch, R.M., 1995, Stream-Gaging Program of the U.S. Geological Survey, U.S. Geological Survey Circular 1123, Reston, Va., 22 p. Web site accessed January 6, 2017, at https://pubs.usgs.gov/circ/1995/1123/report.pdf.

USGS