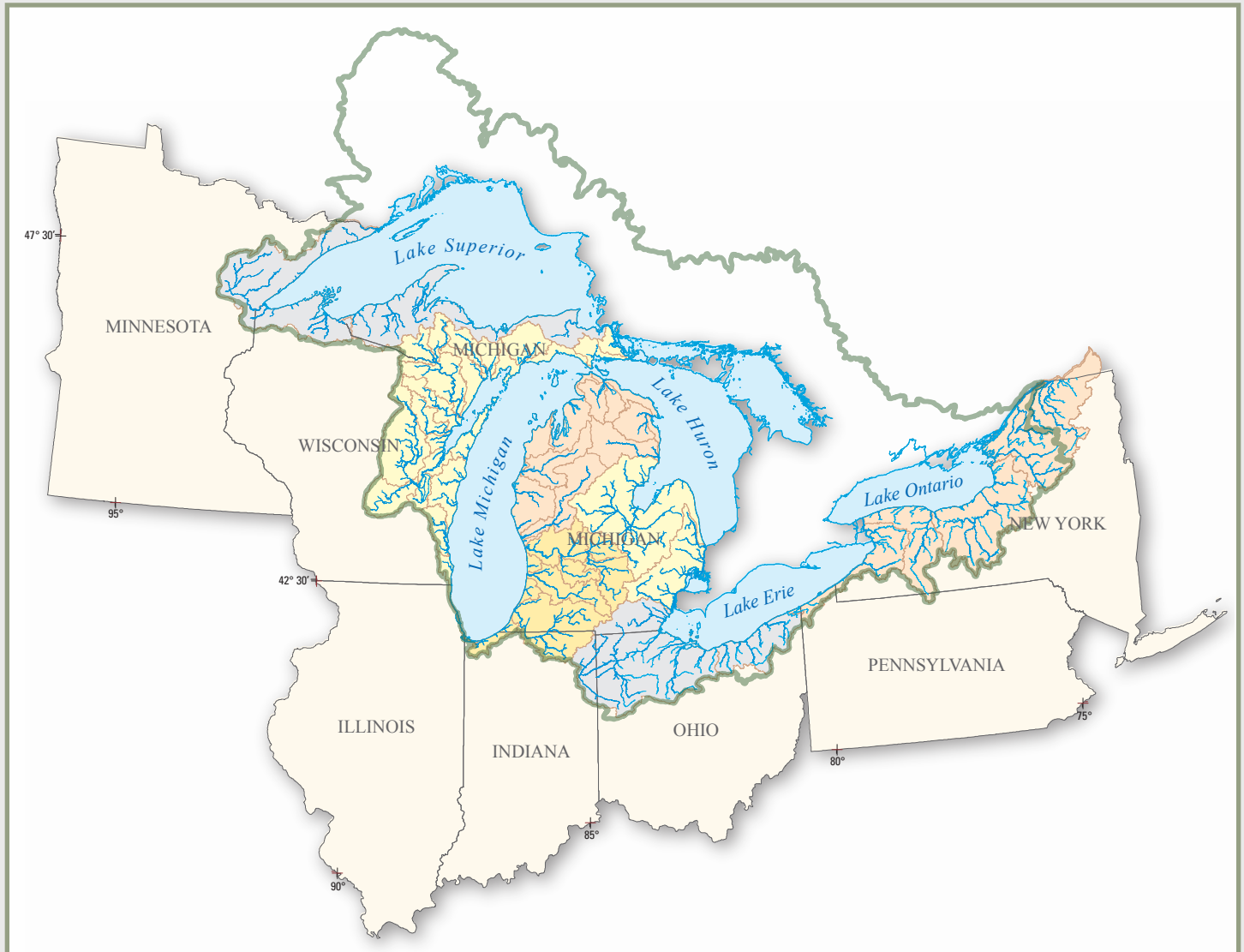


Estimation of Monthly Water Yields and Flows for 1951–2012 for the United States Portion of the Great Lakes Basin with AFINCH



Scientific Investigations Report 2014–5192

Cover image. The United States portion of the Great Lakes Basin.

Estimation of Monthly Water Yields and Flows for 1951–2012 for the United States Portion of the Great Lakes Basin with AFINCH

By Carol L. Luukkonen, David J. Holtschlag, Howard W. Reeves, C.J. Hoard, and Lori M. Fuller

Scientific Investigations Report 2014–5192

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

U.S. Department of the Interior
SALLY JEWELL, Secretary

U.S. Geological Survey
Suzette M. Kimball, Acting Director

U.S. Geological Survey, Reston, Virginia: 2015

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

Although this information product, for the most part, is in the public domain, it also may contain copyrighted materials as noted in the text. Permission to reproduce copyrighted items must be secured from the copyright owner.

Suggested citation:

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, <http://dx.doi.org/10.3133/sir20145192>.

ISSN 2328-0328 (online)

Contents

Abstract.....	1
Introduction.....	1
Purpose and Scope	2
Description of the Great Lakes Basin.....	2
Data Integration within AFINCH and NHDPlus	2
AFINCH Modeling Approach	2
Predictor Step	2
Corrector Step.....	8
Model Limitations.....	9
NHDPlus Geospatial Data Frame	9
Response Data	9
Monthly Streamflow Data	9
Monthly Water-Use Data.....	10
Explanatory Data.....	11
PRISM Monthly Climatic Data	11
National Land Cover Data	11
NHDPlus National Catchment Characteristics Data	11
User-Supplied Catchment Characteristics Data	11
SSURGO Soil Data	11
GAP Vegetation and Land-Use Pattern Data	11
Growing Degree Day Data	12
Output Data	12
Methods of Flow and Water-Yield Estimation for Study Areas.....	14
Study Area 1.....	15
Study Area 2.....	23
Study Area 3.....	33
Study Area 4.....	41
Study Area 5.....	47
Study Area 6.....	55
Study Area 7.....	63
Summary.....	75
References Cited.....	75
U.S. Geological Survey Water-Data Reports for Michigan and the St. Lawrence River Basin.....	77
Regional compilations (listed chronologically).....	77
Michigan data only (listed chronologically).....	77
Appendix 1	78

Figures

1. Map showing the hydrologic subregions and cataloguing units in the U.S. Great Lakes Basin	3
2. Map showing the distribution of estimated water yields for August averaged over 1951–2012	12
3. Time-series plot of monthly flows at flowline 11959440, streamgaging station 04058000, Middle Branch Escanaba River near Ishpeming, Michigan	13
4. Map showing the location of study area 1 (hydrologic subregions 0401 and 0402) in the U.S. Great Lakes Basin	15
5. Graph showing the t-value-indexed colors corresponding to each explanatory variable in the monthly regression equation for estimating water yield for study area 1 (hydrologic subregions 0401 and 0402) in the U.S. Great Lakes Basin	21
6. Map showing the location of study area 2 (hydrologic subregion 0403 and cataloguing units 04040002, 04040003, 04060106, 04060107, 04070001, and 04070002) in the U.S. Great Lakes Basin	23
7. Graph showing the t-value-indexed colors corresponding to each explanatory variable in the monthly regression equation for estimating water yield for study area 2 (hydrologic subregion 0403 and cataloguing units 04040002, 04040003, 04060106, 04060107, 04070001, and 04070002) in the U.S. Great Lakes Basin	32
8. Map showing the location of study area 3 (hydrologic subregion 0405 and cataloguing unit 04040001) in the U.S. Great Lakes Basin	33
9. Graph showing the t-value-indexed colors corresponding to each explanatory variable in the monthly regression equation for estimating water yield for study area 3 (hydrologic subregion 0405 and cataloguing unit 04040001) in the U.S. Great Lakes Basin	40
10. Map showing the location of study area 4 (cataloguing units 04060101, 04060102, 04060103, 04060104, 04060105, 04070003, 04070004, 04070005, 04070006, and 04070007) in the U.S. Great Lakes Basin	41
11. Graph showing the t-value-indexed colors corresponding to each explanatory variable in the monthly regression equation for estimating water yield for study area 4 (cataloguing units 04060101, 04060102, 04060103, 04060104, 04060105, 04070003, 04070004, 04070005, 04070006, and 04070007) in the U.S. Great Lakes Basin	46
12. Map showing the location of study area 5 (hydrologic subregions 0408 and 0409) in the U.S. Great Lakes Basin	47
13. Graph showing the t-value-indexed colors corresponding to each explanatory variable in the monthly regression equation for estimating water yield for study area 5 (hydrologic subregions 0408 and 0409) in the U.S. Great Lakes Basin	54
14. Map showing the location of study area 6 (hydrologic subregions 0410 and 0411) in the U.S. Great Lakes Basin	55
15. Graph showing the t-value-indexed colors corresponding to each explanatory variable in the monthly regression equation for estimating water yield for study area 6 (hydrologic subregions 0410 and 0411) in the U.S. Great Lakes Basin	62

Figures (continued)

16. Map showing the location of study area 7 (hydrologic subregions 0412, 0413, 0414, and 0415) in the U.S. Great Lakes Basin	63
17. Map showing the t-value-indexed colors corresponding to each explanatory variable in the monthly regression equation for estimating water yield for study area 7 (hydrologic subregions 0412, 0413, 0414, and 0415) in the U.S. Great Lakes Basin	74

Tables

1. Summary of study area information	14
2. Streamgaging stations considered for analysis of study area 1 (hydrologic subregions 0401 and 0402) in the U.S. Great Lakes Basin, by water year.....	16
3. Selected explanatory variables for each study area.....	20
4. Regression statistics and explanatory variables used in the AFINCH analysis of flows in study area 1 (hydrologic subregions 0401 and 0402), in the U.S. Great Lakes Basin, by month.....	21
5. Streamgaging stations considered for analysis of study area 2 (hydrologic subregion 0403 and cataloguing units 04040002, 04040003, 04060106, 04060107, 04070001, and 04070002) in the U.S. Great Lakes Basin, by water year.....	24
6. Regression statistics and explanatory variables used in the AFINCH analysis for study area 2 (hydrologic subregion 0403 and cataloguing units 04040002, 04040003, 04060106, 04060107, 04070001, and 04070002) in the U.S. Great Lakes Basin, by month.....	32
7. Streamgaging stations considered for analysis of study area 3 (hydrologic subregion 0405 and cataloguing unit 04040001) in the U.S. Great Lakes Basin, by water year.....	34
8. Regression statistics and explanatory variables used in the AFINCH analysis for study area 3 (hydrologic subregion 0405 and cataloguing unit 04040001) in the U.S. Great Lakes Basin, by month.....	40
9. Streamgaging stations considered for analysis of study area 4 (cataloguing units 04060101, 04060102, 04060103, 04060104, 04060105, 04070003, 04070004, 04070005, 04070006, and 04070007) in the U.S. Great Lakes Basin, by water year.....	42
10. Regression statistics and explanatory variables used in the AFINCH analysis for study area 4 (cataloguing units 04060101, 04060102, 04060103, 04060104, 04060105, 04070003, 04070004, 04070005, 04070006, and 04070007) in the U.S. Great Lakes Basin, by month.....	46
11. Streamgaging stations considered for analysis of study area 5 (hydrologic subregions 0408 and 0409) in the U.S. Great Lakes Basin, by water year.....	48
12. Regression statistics and explanatory variables used in the AFINCH analysis for study area 5 (hydrologic subregions 0408 and 0409) in the U.S. Great Lakes Basin, by month.....	54

Tables (continued)

13. Streamgaging stations considered for analysis of study area 6 (hydrologic subregions 0410 and 0411) in the U.S. Great Lakes Basin, by water year.....	56
14. Regression statistics and explanatory variables used in the AFINCH analysis for study area 6 (hydrologic subregions 0410 and 0411) in the U.S. Great Lakes Basin, by month.....	62
15. Streamgaging stations considered for analysis of study area 7 (hydrologic subregions 0412, 0413, 0414, and 0415) in the U.S. Great Lakes Basin, by water year.....	64
16. Regression statistics used in the AFINCH analysis for study area 7 (hydrologic subregions 0412, 0413, 0414, and 0415) in the U.S. Great Lakes Basin, by month.....	74

Appendix Tables

1-1. Revisions to NHDPlus flowline value-added attribute (VAA) information	78
1-2. Description of classes in the 2006 National Land Cover Database.	82
1-3. Description of soil groups in the Soil Survey Geography (SSURGO) database.	83
1-4. Description of variables from Aquatic Gap Analysis Program (GAP) used in the analysis.....	83

Conversion Factors and Datums

Multiply	By	To obtain
	Length	
mile (mi)	1.609	kilometer (km)
kilometer (km)	0.6214	mile (mi)
	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:

$$^{\circ}\text{F}=(1.8\times^{\circ}\text{C})+32$$

Vertical coordinate information is referenced to the North Geodetic Vertical Datum of 1929 (NGVD 29).

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

Altitude, as used in this report, refers to distance above the vertical datum.

A water year is the 12-month period October 1 through September 30. The water year is designated by the calendar year in which it ends and which includes 9 of the 12 months. Thus, the year ending September 30, 2002, is called the “2002 water year.”

Abbreviations

AFINCH	Analysis of Flows In Networks of CHannels
GAP	USGS National Gap Analysis Program
HUC	hydrologic unit code
NED	National Elevation Dataset
NHD	National Hydrography Dataset
NLCD	National Land Cover Data
NPDES	National Pollutant Discharge Elimination System
OLS	ordinary least square
POA	period of analysis
PRISM	Parameter-elevation Regressions on Independent Slopes Model
R ²	coefficient of determination
RMSE	root mean square error
SSURGO	Soil Survey Geography
USGS	U.S. Geological Survey
VAA	value-added attributes

Estimation of Monthly Water Yields and Flows for 1951–2012 for the United States Portion of the Great Lakes Basin with AFINCH

By Carol L. Luukkonen, David J. Holtschlag, Howard W. Reeves, C.J. Hoard, and Lori M. Fuller

Abstract

Monthly water yields from 105,829 catchments and corresponding flows in 107,691 stream segments were estimated for water years 1951–2012 in the Great Lakes Basin in the United States. Both sets of estimates were computed by using the Analysis of Flows In Networks of CHannels (AFINCH) application within the NHDPlus geospatial data framework. AFINCH provides an environment to develop constrained regression models to integrate monthly streamflow and water-use data with monthly climatic data and fixed basin characteristics data available within NHDPlus or supplied by the user. For this study, the U.S. Great Lakes Basin was partitioned into seven study areas by grouping selected hydrologic subregions and adjoining cataloguing units. This report documents the regression models and data used to estimate monthly water yields and flows in each study area. Estimates of monthly water yields and flows are presented in a Web-based mapper application. Monthly flow time series for individual stream segments can be retrieved from the Web application, The Great Lakes Restoration Initiative (GLRI) Mapper for Monthly Streamflow and Yield of Catchments using Analysis of Flows in Networks of CHannels (AFINCH), <http://cida.usgs.gov/glri/afinch/>, and used to approximate monthly flow-duration characteristics and to identify possible trends.

Introduction

A regionally consistent estimate of streamflow provides unified information across the U.S. Great Lakes Basin for restoration, assessment, management, and conservation of stream ecosystems. These estimated flows provide information for natural resource professionals and the general public to assess stream ecosystems across State jurisdictional boundaries and

provide a framework that can be used to help enhance existing data networks. In addition, this information can be used to assess the status of stream ecosystems, to identify and prioritize locations to focus stream restoration efforts, to plan and manage game and non-game species, to protect and enhance aquatic systems for recreational use, to provide tools for assessing disturbance of aquatic ecosystems, to plan adaptive management strategies, and to aid in implementation of the Great Lakes-St. Lawrence River Basin Water Resources Compact (Council of Great Lakes Governors, 2005).

The Analysis of Flows In Networks of CHannels (AFINCH) application (Holtschlag, 2009) was applied to estimate monthly time series of yields and flows for water years 1951–2012 within the U.S. Great Lakes Basin (Hydrologic Region 04) (Seaber and others, 1987). The resulting long-term flow time series can be used to describe monthly flow-duration characteristics and to investigate trends at 107,691 stream segments in the U.S. Great Lakes Basin. AFINCH is an interactive program used to facilitate the estimation of time series of monthly water yields from catchments and flows in stream reaches defined within the 1:100,000-scale NHDPlus geospatial data framework (McKay and others, 2012). NHDPlus is an integrated suite of application-ready geospatial-data products that includes features from the National Hydrography Dataset (NHD; U.S. Geological Survey, 2009a), the National Elevation Dataset (NED; U.S. Geological Survey, 2009b), and the National Watershed Boundary Dataset (U.S. Department of Agriculture-Natural Resources Conservation Service, 2009; U.S. Geological Survey, 2008), which describes the geometry and attributes of the hydrologic units. Monthly streamflow data at included streamgages, adjusted for upstream water use, along with monthly climatic data and land-cover or basin characteristics provide a basis for developing regression equations to estimate natural monthly water yields at individual

catchments. Streamflows are estimated from the catchment yields and flow is conserved through the NHDPlus network. Choropleth maps of monthly water yield and flow can be generated and analyzed. Estimated monthly flows can be displayed within AFINCH, examined for nonstationarity, and tested for monotonic trends. Monthly flows also can be used to estimate flow-duration characteristics at stream segments within AFINCH.

Purpose and Scope

The purpose of this report is to document regression models and estimates of monthly water yields and flows developed for seven study areas spanning the U.S. Great Lakes Basin for the period 1951–2012. These models were developed within the AFINCH modeling environment by using the NHDPlus geospatial data framework to estimate monthly water yields and flows for 105,829 catchments and 107,691 flowlines in the region. The estimated flows are constrained to be consistent with measured flows at 823 U.S. Geological Survey (USGS) streamgages. These constraints also adjust regression estimates upstream from streamgages to generally improve flow estimates from regression estimates. Monthly water-use data were applied, where available, to mitigate effects of water use on the regionalization process.

Description of the Great Lakes Basin

The Great Lakes Basin in the northeastern United States and southeastern Canada contains the five Laurentian Great Lakes of Superior, Michigan, Huron, Erie, and Ontario. These Great Lakes make up the largest surface freshwater system on Earth (U.S. Environmental Protection Agency, 2014), and contain approximately 5,500 cubic miles (mi³) of water (Waples and others, 2008). The approximate land drainage area of the basin is 201,460 square miles (mi²), water area is 94,250 mi², and the total area of the basin is 296,000 mi² (Government of Canada and U.S. Environmental Protection Agency, 1995). Parts of the States of Michigan, New York, Ohio, Wisconsin, Indiana, Minnesota, Illinois, and Pennsylvania, and the provinces of Ontario and Quebec drain into the Great Lakes.

Hydrologic Region 04 subdivides the U.S. Great Lakes Basin into 15 hydrologic subregions based on surface-water divides (watersheds) (U.S. Geological Survey, 2008) (fig. 1). The hierarchical hydrologic unit code (HUC) system describes increasingly refined levels of subdivision by appending a 2-digit code onto the end of the code identifying

the previous level. In particular, the coding system divides surface-water drainage basins in the U.S. and the Caribbean into 21 hydrologic regions (2 digits), 221 subregions (4 digits), 378 accounting units (6 digits), and 2,264 cataloguing units (8 digits). The subregions and cataloguing units in the U.S. Great Lakes Basin were formed into seven study areas for analysis with AFINCH. These study areas facilitated the analysis by (1) combining hydrologically similar areas, (2) utilizing a larger pool of streamgaging stations than would have been possible by analyzing subregions individually, and (3) improving consistency of estimates between study areas.

Data Integration within AFINCH and NHDPlus

The AFINCH modeling approach and method of estimating flows along with descriptions of potential explanatory variables are described in the following sections.

AFINCH Modeling Approach

AFINCH implements a two-step predictor-corrector process. The predictor step is based on a user-specified multiple-linear-regression equation, which estimates monthly yields using selected monthly climatic and basin characteristics data and flows at streamgages. The corrector step adjusts the regression estimates to match measured flows at included streamgages.

Predictor Step

The predictor step in an AFINCH analysis is based on the statistical relation between a response variable that is a function of monthly flows and one or more explanatory variables that may be formed from time-varying monthly climatic data or fixed basin characteristics data. In particular, a user-specified multiple-linear-regression model is used to predict the response variable for each NHDPlus catchment in the study area. As discussed later in more detail, catchment yield (the response variable in this study) is defined as the streamflow contribution from the catchment divided by the catchment area. The estimated square root of yield is transformed and multiplied by catchment area to produce a runoff or catchment flow. The catchment flows are accumulated downstream by using the stream network topology described in NHDPlus. The following paragraphs develop the notational conventions underlying the AFINCH model.



Figure 1. Hydrologic subregions and cataloguing units in the U.S. Great Lakes Basin.

A time series of monthly flows for the period of analysis at each NHDPlus flowline in the study area is the primary output from an AFINCH analysis. Monthly measured flows $^{mea}Q_{Gage(i,j)}$, for water year index i and streamgage index j , are critical to developing these estimates. Stream-flow data are typically reported as mean daily flows at the streamgage; monthly flows used in this study are the mean of the daily flows for each month. This set of monthly measured flows at included streamgages in water year i can be formed into a column vector with length n_i , written $^{mea}Q_{n_i \times 1}$.

In some stream reaches, withdrawals, diversions, or augmentations associated with water use can augment or diminish flows measured at downstream streamgages. If monthly water-use data are available to document these uses, measured flows at streamgages can be adjusted to more accurately represent natural flow conditions, that is, without water use. For example, if a consumptive water use of 1 cubic foot per second (ft^3/s) is withdrawn from a reach upstream of one or more streamgages, the natural flow at the affected streamgages would be expected to have $1 \text{ ft}^3/\text{s}$ more than the measured flow. Understanding natural flow conditions helps regionalize flow information with AFINCH. Measured flows that have been adjusted with available water use are designated as $^{mea}\tilde{Q}_{Gage(i,j)}$.

Streamflow generally occurs in dendritic drainage patterns in which headwater streams converge to form tributaries that combine flow from upstream branches. This process progresses downstream as tributaries continue to merge until only a main stem remains. A streamgaging network follows this pattern, so some streamflow is commonly measured by multiple streamgages. To avoid this redundancy in streamflow information, AFINCH computes an incremental flow, ${}_{\Delta}^{mea}Q_{Gage(i,j)}$, and a corresponding incremental drainage area, ${}_{\Delta}^{basin}A_{Gage(i,j)}$, for each streamgage based on the stream network configuration. Consider, for example, a simple two-gage network in which the water measured at the upstream gage is combined with additional flow as it moves downstream and is re-measured at the downstream gage. In this network, the incremental flow and area for the upstream streamgage is the same as the measured flow and total area; at the downstream streamgage, however, the incremental flow is the measured flow minus the flow measured at the upstream gage. Similarly, the incremental area is the contributing drainage area at the downstream gage minus the contributing drainage area at the upstream gage.

Over a significant range of flows and drainage areas, (incremental) streamflow is proportional to the (incremental) drainage area. For example, all other factors being equal in a local flow system, the mean streamflow at a stream site with twice the drainage area of another stream site is likely to have twice the mean flow. The dominant effect of drainage areas on flows can mask more subtle effects statistically associated with the impact that land use and land cover, monthly precipitation and temperature, and other factors might have on streamflow. To increase the sensitivity of the analysis to these more subtle basin and climate effects, AFINCH analysis is based on water yield, which is obtained by dividing incremental flows by corresponding incremental areas as ${}_{\Delta}^{mea}Y_{Gage(i,j)} = {}_{\Delta}^{mea}Q_{Gage(i,j)} / {}_{\Delta}^{basin}A_{Gage(i,j)}$ or in vector notation ${}_{\Delta}Y_{n_i \times 1} = {}_{\Delta}Q_{n_i \times 1} \circ {}_{\Delta}^{basin}A_{n_i \times 1}^{-1}$. This approach removes drainage area as a possible explanatory variable, which also mitigates some concerns about developing parameter estimates from analysis of (incremental) streamgage basins, made up of hundreds of individual NHDPlus catchments, for prediction at individual catchments.

A multiple regression model provides a basis for prediction and for assessing the uncertainty of the prediction. This uncertainty is based on simplifying assumptions concerning the characteristics and distribution of model residuals ε formed by the differences between measured and predicted values of the response. In particular, model residuals are assumed to be independently and identically distributed with a normal (Gaussian) distribution having a mean of 0 and a variance of σ_{ε}^2 (Draper and Smith, 1998). So that this assumption is more likely satisfied, a square-root transformation was applied to the incremental yield to reduce its skewness. Skewness is a measure of the departure of a distribution from a symmetric distribution like the normal distribution. The transformed yield data are denoted by ${}_{\Delta}^{mea}Y_{Gage(i,j)}^{0.5} = \sqrt{{}_{\Delta}^{mea}Y_{Gage(i,j)}}$ and the residuals are formed as $\varepsilon_{n_i \times 1} = {}_{\Delta}^{mea}Y_{n_i \times 1}^{0.5} - {}_{\Delta}^{est}Y_{n_i \times 1}^{0.5}$.

In the multiple regression model underlying AFINCH, the response variable of monthly water yields—expressed generically for any individual water year i with n_i streamgages in the study area as ${}_{\Delta}^{mea}Y_{n_i}^{0.5}$ —is linearly related to a design matrix containing a column of 1's, corresponding to an intercept term, augmented on the right with a set of explanatory variables. Explanatory variables are user specified, but generally include monthly climatic and basin characteristics. All columns have the same number of elements (lengths of n_i) and the rows must be independent as

$$\frac{mea}{\Delta} \tilde{Y}_{n_i \times 1}^{0.5} = \begin{bmatrix} \mathbf{1}_{n_i \times 1} & X_{n_i \times k} \end{bmatrix} \cdot \beta_{k+1} + \varepsilon_{n_i}, \quad (1)$$

where

$$\frac{mea}{\Delta} \tilde{Y}_{n_i \times 1}^{0.5}$$
 is an n_i -column vector of the square-roots of measured incremental water yields, adjusted for water use;

 $\begin{bmatrix} \mathbf{1}_{n_i \times 1} & X_{n_i \times k} \end{bmatrix}$ is a design matrix of explanatory variables in which the first column is a n_i -vector of ones for the intercept term, which is augmented on the right by an $n_i \times k$ matrix of explanatory variables indexed over water years by i ;

 β_{k+1} is a $k+1$ column vector of parameters relating the explanatory variables to the response variable; and

 ε_{n_i} is an n_i -column vector of regression residuals, which are assumed to be independent and identically distributed random variables following a normal (Gaussian) distribution with mean 0 and variance σ_ε^2 .

In cases where the k explanatory variables can be partitioned into a set of k_v time-varying characteristics such as monthly precipitation, and a set of k_f time-invariant (fixed) characteristics such as soil types, the design matrix for a selected month in the i^{th} year can be written as

$$X_{n_i \times (1+k_v+k_f)}^i = \begin{bmatrix} \mathbf{1}_{n_i} & X_{n_i \times k_v}^i & X_{n_i \times k_f} \end{bmatrix} \quad (2)$$

Ordinary least-square (OLS) estimates of the parameter vector β^i for monthly water yields in the i^{th} water year can be computed (Draper and Smith, 1998) as

$$\begin{bmatrix} \hat{\beta}_0^i \\ \hat{\beta}_{k_v}^i \\ \hat{\beta}_{k_f}^i \end{bmatrix} = \left[\begin{bmatrix} X_{n_i \times (1+k_v+k_f)}^i \end{bmatrix}' \cdot X_{n_i \times (1+k_v+k_f)}^i \right]^{-1} \cdot \begin{bmatrix} X_{n_i \times (1+k_v+k_f)}^i \end{bmatrix}' \cdot \frac{mea}{\Delta} \tilde{Y}_{n_i \times 1}^{0.5} \quad (3)$$

Annual estimates of parameters may be useful in cases where the relation between the explanatory and response variables is suspected of being time varying. For example, a trend in parameter estimates might occur if an explanatory variable, such as land use, is treated as fixed when it is actually time varying. More commonly, the design matrix for the user-specified period of analysis (POA) for N years can be developed by vertically stacking design matrixes for individual water years and corresponding water yield vectors $\frac{mea}{\Delta} \tilde{Y}_{n_i \times 1}^{0.5}$ as

$$\underline{X} = X_{\sum_i n_i, 1+k_v+k_f} = \begin{bmatrix} X_{n_1 \times (1+k_v+k_f)}^1 \\ X_{n_2 \times (1+k_v+k_f)}^2 \\ \vdots \\ X_{n_N \times (1+k_v+k_f)}^N \end{bmatrix} = \begin{bmatrix} \mathbf{1}_{n_1} & X_{n_1 \times k_v}^1 & X_{n_1 \times k_f} \\ \mathbf{1}_{n_2} & X_{n_2 \times k_v}^2 & X_{n_2 \times k_f} \\ \vdots & \vdots & \vdots \\ \mathbf{1}_{n_N} & X_{n_N \times k_v}^N & X_{n_N \times k_f} \end{bmatrix} \quad (4)$$

The corresponding OLS estimate of the regression parameter vector for the period of analysis is

$$\begin{bmatrix} \hat{\beta}_0^{poa} \\ \hat{\beta}_{k_v}^{poa} \\ \hat{\beta}_{k_f}^{poa} \end{bmatrix} = \left[\underline{X}' \cdot \underline{X} \right]^{-1} \cdot \underline{X}' \cdot \overset{mea}{\Delta} \underline{\tilde{Y}}_{\sum_i n_i \times 1}^{0.5} \quad (5)$$

In AFINCH, 12 equations are individually estimated for monthly yields (either annually or for the POA) providing the flexibility to include different explanatory variables, β , in different monthly equations. A stepwise-selection process is used to automate the selection process in which the user specifies an α – level (the probability of a type I error, which in this example, is incorrectly rejecting the null hypothesis that individual parameters are equal to zero) to control the apparent statistical significance of included explanatory variables. Stepwise selection is an iterative process in which p -values (the apparent probability that the magnitude of a parameter estimate, relative to its uncertainty, would occur by chance under the null hypothesis) are repeatedly computed and compared to the specified α – level. Explanatory variables with p -values smaller than the α – level are maintained in the equation. The user can override the stepwise-selection process by specifying which explanatory variables to include in the monthly regression equations regardless of their apparent statistical significance. Finally, the user can specify whether OLS or robust parameter estimates are to be used in the estimation of catchment water yields. Robust parameter estimates use a bisquare weight function with a tuning constant of 4.685 by default (The MathWorks, Inc., 2012). Robust estimates may be preferred to OLS estimates to reduce the sensitivity of parameter estimates to one or more irresolvable outliers in the water yield series; however, for the estimates presented in this report, OLS parameter estimates were used.

Once the monthly regression equations are developed from streamflow information at gaged basins, the results are applied to estimate water yields for all gaged and ungaged NHDPlus catchments. For the case where estimation equations are developed by using POA parameter estimates, monthly estimates of water yield for individual water years are computed as

$$\overset{est}{\underline{\tilde{Y}}}_{M \times 1}^{0.5} = \underline{x}_{M \times (k+1)} \cdot \hat{\beta}_{k+1}^{poa}, \quad (6)$$

where

$\overset{est}{\underline{\tilde{Y}}}_{M \times 1}^{0.5}$ is the estimated vector of natural water yields (adjusted for water use), in the square-root metric, for all M NHDPlus catchments-water years in the period of analysis;

$\underline{x}_{M \times (k+1)}$ are catchment characteristics that correspond to basin characteristics for all the water-years in equation [4]; and

$\hat{\beta}_{k+1}^{poa}$ is the vector of parameter estimates.

Estimated runoff (flows) from individual catchments are computed from estimated yields and catchment areas as

$$\underline{\tilde{q}}_{M \times 1}^{est} = \left(\underline{\tilde{y}}_{M \times 1}^{est, 0.5} \right)^2 \circ^{catch} \underline{a}_{M \times 1}, \quad (7)$$

where

- $\underline{\tilde{q}}_{M \times 1}^{est}$ is the vector of natural flows (runoff) for all catchments-water years in the period of analysis,
- $\left(\underline{\tilde{y}}_{M \times 1}^{est, 0.5} \right)^2$ is the estimated natural catchment yield, and
- $\underline{a}_{M \times 1}^{catch}$ is a vector of *NHDPlus* catchment areas.

For notational simplicity, let $\underline{\tilde{q}}_{M \times 1}^{est} \rightarrow \underline{\tilde{q}}_{m \times 1}^{est} \forall wy(i) \in poa$. Then accumulate monthly flows from catchments (runoff) forming the incremental areas upstream of included streamgages in water year i as

$$\underline{\tilde{Q}}_{n_i \times 1}^{est} = \sum_{catch_{Gage(i,j)} = catch_{StudyArea} \cap^{nBasin} A_{\Delta}^{i, n_i \times 1}} \underline{\tilde{q}}_{catch}^{est}, \quad (8)$$

where

- $\underline{\tilde{Q}}_{n_i \times 1}^{est}$ is the vector of accumulated AFINCH estimates of incremental flows at the n_i active streamgages in water year i ,
- $catch$ is the set of all study area catchments in the incremental area of individual streamgages monitored in water year i , and
- $\underline{\tilde{q}}_{catch}^{est}$ is the natural monthly runoff from the set of catchments forming the incremental streamgage areas in the i^{th} water year.

and compare with measured flows.

Corrector Step

A corrector step is used to adjust AFINCH estimated monthly water yields upstream from an included gage so that resulting AFINCH constrained monthly flows match measured flows at every streamgage for each water year in the period of analysis as ${}^{con}_{\Delta} \tilde{Q}_{Gage_j}^i \equiv {}^{mea}_{\Delta} \tilde{Q}_{Gage_j}^i$. The adjustment is applied to the catchment corresponding to the NHDPlus flowline where the streamgage is located and to estimated yields on all catchments upstream of the streamgage. The flow-based correction ratio is computed as

$$\tilde{qRatio}_{Gage_j}^i = {}^{mea}_{\Delta} \tilde{Q}_{Gage_j \times 1} \circ {}^{est}_{\Delta} \tilde{Q}_{Gage_j \times 1}^{-1} \quad (9)$$

$${}^{con} \tilde{y}_{usGage_j}^i = {}^{est} \tilde{y}_{usGage_j}^i \cdot \tilde{qRatio}_{Gage_j}^i, \quad (10)$$

where

- ${}^{con} \tilde{y}_{usGage_j}^i$ is the constrained estimate of water yields in *NHDPlus* catchments forming the incremental drainage area upstream of the j^{th} streamgage in the i^{th} water year,
- $usGage_j$ is the set of catchments contributing to flowlines upstream from $Gage_j$, and
- $\tilde{qRatio}_{Gage_j}^i$ is a constant for the j^{th} streamgage and i^{th} water year. In ungaged areas and areas downstream from active gages, estimated water yields are unconstrained.

Constrained monthly water yields for each catchment and water year are used with corresponding catchment areas to compute constrained runoff ${}^{con} \tilde{q}_{catch}^i = {}^{con} \tilde{y}_{catch}^i \circ a_{catch}$. The constrained runoffs are accumulated downstream by using the NHDPlus flowline topology to compute constrained accumulated AFINCH flows ${}^{con} \tilde{Q}_{flowline}^i$. Lastly, water-use information is added to affected flowlines and combined with ${}^{con} \tilde{Q}_{flowline}^i$ so that ${}^{con} Q_{Gage_j}^i \equiv {}^{mea} Q_{Gage_j}^i \forall j \in \text{Streamgage Network in water year } i$.

Model Limitations

The regression model structure underlying AFINCH has some inconsistencies with the data structure that may affect the reliability of reported statistics (p -values) used for assessing model suitability. In particular, model residuals are assumed to be independent in regression models, but the data structure provides repeated (annual) measures from a network of streamgages. Thus, residual series from individual streamgages may not be independent in time, and residual series from different streamgages may not be identically distributed. These limitations are likely to be exacerbated when there are few streamgages with relatively long periods of record. Lack of temporal independence and non-identically distributed residuals series could inflate the apparent significance (reducing p -values) of individual explanatory variables, thus leading to greater type I errors than what would be consistent with the specified α – level. To mitigate this potential problem, the number of explanatory variables potentially included in the regression model might be constrained, say to a number less than the square root of the number of included streamgages, or the α – level might be otherwise reduced.

NHDPlus Geospatial Data Frame

NHDPlus describes the geometry and the attributes of the hydrologic units and links flowlines to catchments and to other flowlines. Notably, NHDPlus includes the flowline topology, which is the routing information providing the upstream and downstream connections for each flowline. AFINCH estimates of water yields are based on incremental flows and areas within streamgage networks. These incremental flows are summed down the network from the headwaters to the downstream-most segment in the watershed under analysis. There are 105,829 catchments making up Hydrologic Region 04 (U.S. Great Lakes Basin), which have an average area of 1.16 mi². Each catchment, which is uniquely identified by a *Gridcode*, is associated with a flowline (stream segment), which is uniquely identified by a *ComID*, that drains the catchment.

The 107,691 flowlines characterizing the drainage network in the U.S. Great Lakes Basin include streams and rivers, canals and ditches, pipelines, artificial paths, coastlines, and connectors. Flowlines in the U.S. Great Lakes Basin average about 1 mile (mi) in length; some flowlines have no associated catchments. The topology of flowlines defined within NHDPlus provides a mechanism to accumulate flows from

catchments and route flows through a drainage network. Initial inspection of some of the network design matrixes and the drainage area plots produced by AFINCH indicated the need for revisions to the flowline routing. In addition, early estimated flows also indicated the need for changes to the specified diversions for some flowlines. These revisions to flowline routing and diversion information are outlined in appendix 1, table 1–1.

Response Data

AFINCH uses multiple regression analysis to regionalize and spatially distribute water yields across catchments spanning (stream) gaged and ungaged basins in the study area. Monthly streamflow and water-use data are used together to describe the response variable.

Removing effects of artificial augmentations and diversions from directly impacted streamflow records helps describe natural streamflow conditions and associated yields, which can be regionalized. From these natural streamflow conditions, regression analysis describes how monthly climatic and basin characteristics affect the distribution of natural water yields across catchments. Catchment yields with corresponding drainage areas are used to compute runoff (flow), which can be accumulated downstream by using the NHDPlus streamflow topology. In gaged basins, the regression yield estimates are adjusted (constrained) so that resulting flows, with water-use data re-applied, match measured monthly flows. In ungaged basins, the natural water yields are used (without constraints) to estimate flows, with available water-use data re-applied to provide the final estimates of flow.

Monthly Streamflow Data

The NHDPlus GageLoc is a shapefile that locates USGS streamgages along flowlines. A USGS streamgage is uniquely identified by its station number; other information included in the analysis is the drainage area at the gage. In the U.S. Great Lakes Basin (Hydrologic Region 04), 1,339 streamgages are identified; however, stations that (1) represented lake sites, (2) had no specified drainage area, and (or) (3) did not have complete data for at least 1 water year within the period 1951–2012 were excluded from the analysis. Therefore, data from 823 streamgages were compiled for the AFINCH analysis. Time series of monthly streamflow values were computed from data obtained through the USGS cooperative streamgaging program (Wahl and others, 1995).

Occasionally, within NHDPlus, one or more streamgages may be located at different points along a single flowline. The streamgage that best represents flow conditions, for example the gage with the longer period of record, was selected for the analysis. In each subregion, flow at sites that were regulated or influenced by water withdrawals and diversions, as indicated by station descriptions or analysis of the water yield graphs, was not included in the analysis if water-use or routing data were not readily available to adjust the flow to more natural conditions. NHDPlus routing tables were used to route flow upstream to headwaters or downstream to discharge points. For each streamgaging station, all upstream flowlines (and associated catchments) were identified by using the NHDPlusV2 Flow Table Navigator Toolbar in ArcMap (Horizon Systems Corporation, 2014).

Monthly Water-Use Data

In AFINCH, water use (including withdrawals, augmentations, and diversions) can be associated with flowlines and accounted for along with catchment flows. Monthly flows can be adjusted for specified water uses by using a companion (currently (2014) unpublished) program to AFINCH, which aids with water-use accounting and tracks withdrawals or returns from their point of origin to the downstream-most segment in the analysis. The Water Use Data Tool is available upon request from the USGS Michigan Water Science Center. Water withdrawals are indicated by negative water-use values, and flow augmentations are indicated by positive water-use values. Monthly water-use data for each water year within the period of analysis were specified for those flowlines with readily available data; however, determination of monthly water use for the period of analysis at each possible location was beyond the scope of this study. In addition, much of these data are unavailable, incomplete, or are not in a format that is readily amenable for inclusion in this analysis. Collection of additional data would likely improve the estimates and permit inclusion of additional streamgage data in the analysis.

Water withdrawals, diversions, and augmentations were determined from available records for selected flowlines in hydrologic subregions 0402, 0403, and 0409 using data from the National Pollutant Discharge Elimination System (NPDES) permit program, USGS Annual Data Reports, and (or) other USGS studies. The NPDES permit program requires industrial, municipal, and other facilities to report point-source discharges to surface-water bodies. Data were retrieved using the Permit Compliance System (U.S. Environmental Protection Agency, 2001). Monthly discharges that could be linked

to an adjacent flowline were included in the AFINCH analysis each year the discharge occurred for subregions 0403 and 0409. Other withdrawal and diversion data were available from USGS Annual Data Reports (U.S. Geological Survey Water-Data Reports for Michigan and the St. Lawrence River Basin, 1958–2012). For example, reservoir stage data are collected that can be related to a flow or change in storage in the reservoir or to a withdrawal and later release to a different flowline. Data from these sources were included in the AFINCH analysis for subregions 0402 and 0403. In addition, local data used in the Kalamazoo area study (Luukkonen and others, 2004), describing water diversions to Portage Creek for selected years during 1966–2002, were used to estimate water-use data for hydrologic subregion 0405 for this time period.

In hydrologic subregion 0402, at streamgaging station 04034000 Bond Falls Reservoir near Paulding, Michigan, water is diverted to South Branch Ontonagon River through Bond Falls Canal (streamgaging station 04033500), and water is used for power production at Victoria Dam near Rockland. Gage height data provided by Upper Peninsula Power Company are converted to acre-feet by USGS (Neal Craig, U.S. Geological Survey, oral commun., 2011). Water diverted from Schweitzer Reservoir to this subregion as described below also was included in the analysis. Data for 1950–2010 were used in the AFINCH analysis for subregion 0402.

In hydrologic subregion 0403, at streamgaging station 04057811 Greenwood Reservoir near Greenwood, Mich., water can flow over a concrete spillway into the Middle Branch Escanaba River below streamgaging station 04057814 Greenwood Release near Greenwood, Mich., with some downstream diversions by local industries above streamgaging station 04058000. An outlet structure also permits flow from the reservoir into the afterbay (conservation pool) where water can be diverted to Green Creek gaged at Greenwood Diversion (streamgaging station 04057813) or released to the Middle Branch Escanaba River below streamgaging station 04057814. Beginning in October 1979, some water also was released after use to the East Branch Escanaba River via Goose Lake Outlet. Since January 1973, Greenwood Reservoir diverts some impounded water to Schweitzer reservoir (streamgaging station 04058190 Schweitzer Reservoir near Palmer, Mich.) for use in iron ore processing. Some flow from this reservoir is diverted to the city of Ishpeming, Mich., for municipal supply and discharged to the Carp River Basin (in hydrologic subregion 0402). After processing, water is returned to Green Creek or, beginning in 1980, to Goose Lake Outlet. Data for 1963–2009 were used in the AFINCH analysis for hydrologic subregion 0403.

Explanatory Data

Multiple regression analyses relate explanatory variables consisting of time varying (monthly) climatic characteristics, and time invariant (fixed) catchment characteristics to monthly varying water yields. In this study, Parameter-elevation Regressions on Independent Slopes Model (PRISM) monthly precipitation and temperature data were used, and catchment characteristics in the National Land Cover Data (NLCD) set were supplemented or replaced with user-specified catchment characteristics. The source and characteristics of explanatory variables used in this analysis are described in the following sections.

PRISM Monthly Climatic Data

For this report, average monthly air temperature and total precipitation data were computed for each year during the period 1950–2012 for all catchments in the U.S. portion of the Great Lakes Hydrologic Region (04) (fig. 1). The climatic data are based on 2.5 arc-minute (approximately 4 kilometers) grids of PRISM data by Daly and Taylor (1998a, b). Monthly minimum and maximum air temperature and precipitation datasets for 1895–2012 from the PRISM Climate Group (2013) were overlain with polygons representing catchment boundaries to estimate spatially averaged precipitation and temperature values. The average of the minimum and maximum temperature values was used to estimate the average monthly temperature.

National Land Cover Data

Land-cover information for the conterminous United States for 2006 is available through the NLCD from the Multi-Resolution Land Characteristics Consortium (Fry and others, 2011). The 16-class land-cover classification scheme includes more broadly defined land-cover classes including water, developed, barren, forest, shrub land, herbaceous, planted or cultivated, and wetlands (appendix 1, table 1–2). Percentages of these categories were determined for each catchment in Hydrologic Region 04.

NHDPlus National Catchment Characteristics Data

NHDPlus includes a stream network based on the medium resolution (1:100,000 scale) NHD, with connectivity within the stream network and to associated catchments. Stream order is a method of assigning a numeric order to links

within a stream network and in NHDPlus is a modified version of stream order as defined by Strahler (1957). The Strahler stream order algorithm does not account for flow splits in the network. The NHDPlus algorithm for stream order does take flow splits into consideration. These value-added attributes (VAA) are computed from upstream to downstream (McKay and others, 2012).

The NHDPlus suite of data includes information from the NED (Gesch and others, 2009). For each NHDPlus catchment, average catchment elevations were determined using the zonal statistics ArcMap toolbox command (ESRI, 2012). In a similar way, average and minimum catchment slope were determined for each catchment using the ArcMap toolbox commands for slope and zonal statistics (ESRI, 2012).

User-Supplied Catchment Characteristics Data

The user may provide one or more sets of shapefiles containing catchment attributes that can be used as explanatory variables in the subsequent regression analysis. User-specified variables for this analysis included soil information from the Soil Survey Geography (SSURGO) database (U.S. Department of Agriculture-Natural Resources Conservation Service, 2011), surficial and bedrock geology from the USGS National Gap Analysis Program (GAP), and growing degree day data (computed using PRISM air temperature data).

SSURGO Soil Data

The SSURGO database contains information about soil components and properties as collected by the National Cooperative Soil Survey and is available for most areas in the United States. The hydrologic group is based on estimates of runoff potential and are assigned according to the rate of water infiltration when the soils are not protected by vegetation, are thoroughly wet, and receive precipitation from long-duration storms (appendix 1, table 1–3). If a soil is assigned to a dual hydrologic group (A/D, B/D, or C/D), the first letter is for drained areas and the second is for undrained areas. Only the soils that in their natural condition are in group D are assigned to dual classes.

GAP Vegetation and Land-Use Pattern Data

The USGS National GAP Land Cover Data Set includes detailed vegetation and land-use patterns for the continental United States to aid in State or regional assessments of natural resources and to provide for a regionally consistent database with uniformity across State boundaries (U.S. Geological Survey, 2011). GAP is a nationwide effort under the direction

of the USGS and provides information on the distribution of native vegetation types, modified and introduced vegetation, developed areas, and agricultural areas of the United States. Surficial and bedrock geology variables used in the AFINCH analysis are shown in appendix 1, table 1–4.

Growing Degree Day Data

Growing degree days provide a measure of heat accumulation because plant and animal development is related to the accumulation of heat or temperature units above a threshold or base temperature below which little growth occurs. The average annual growing degree days for the continental United States were calculated for 1971–2000 from downloaded PRISM air temperature datasets (Mike Slattery, U.S. Geological Survey, written commun., 2013). The base temperature (10 degrees Celsius (°C)) was subtracted from the average of the maximum and minimum monthly air temperatures. Only values equal to or greater than 0 °C were used and the resulting grid was then multiplied by the number of days in the month; the total from the previous month was added in a cumulative method to arrive at an approximation of the average annual growing degree days.

Output Data

The AFINCH program writes comma-delimited ASCII files with constrained and unconstrained monthly yields and flows for all the stream segments in the part of the NHDPlus area being analyzed and for each month and water year of the time period of the analysis. These flows and yields are indexed by the *ComID* of the stream segment or the *Gridcode* of the catchment. Headers in these files help document the analyses by describing explanatory variables, estimated parameters, and other user specifications. A representative choropleth map showing the distribution of water yields for August averaged over 1951–2012 is shown in figure 2. An example time-series plot of monthly flows for a streamgaging station in hydrologic subregion 0403 showing the square root of measured flow with the square root of estimated incremental flow is shown in figure 3. Estimated flows and yields have been incorporated into a mapper application by the USGS Center for Integrated Data Analytics, <http://cida.usgs.gov/glri/afinch/>, for ease of use, display, and retrieval. These long-term flow series can be used to investigate trends or to characterize monthly flow durations.

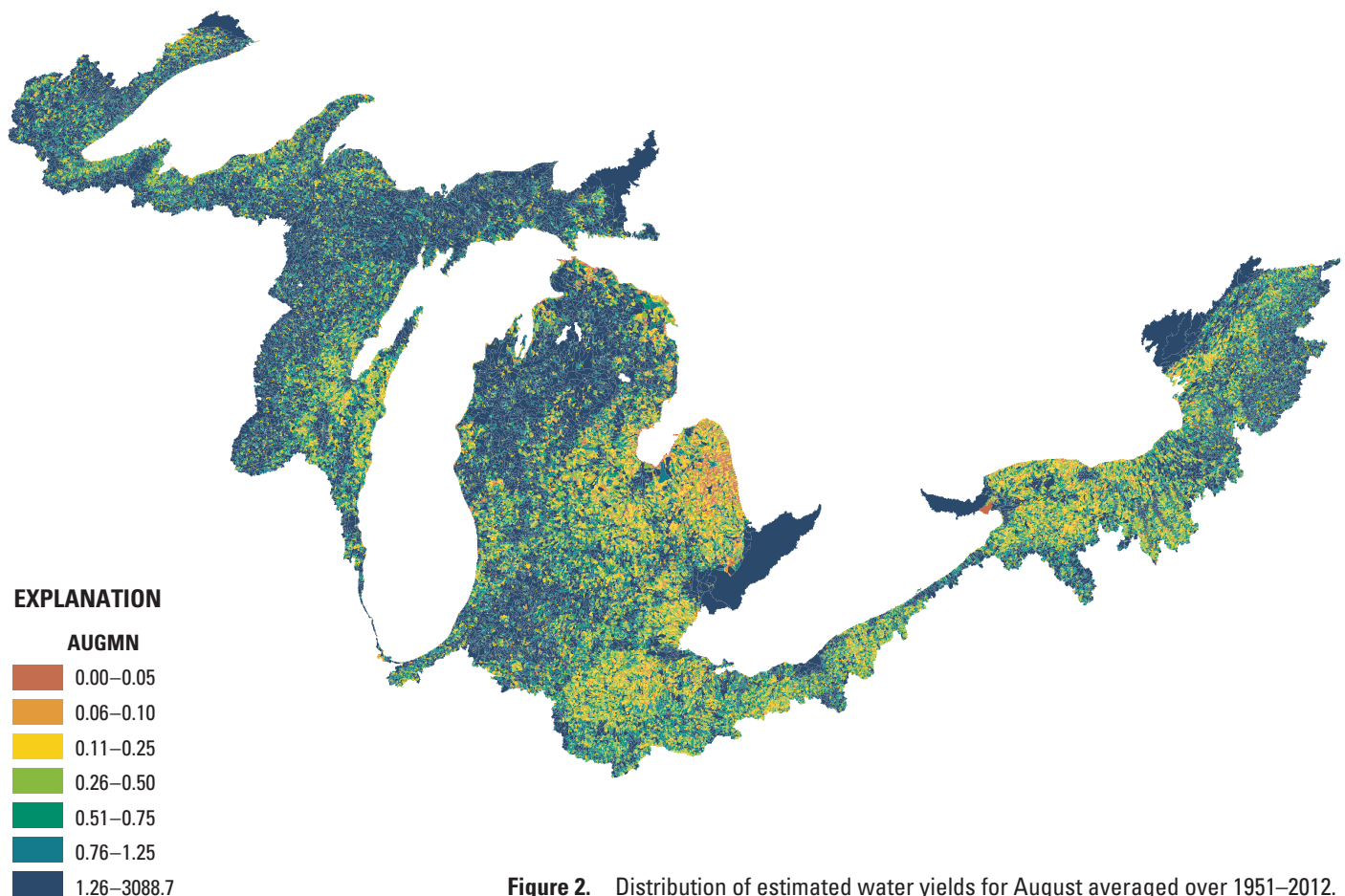


Figure 2. Distribution of estimated water yields for August averaged over 1951–2012.

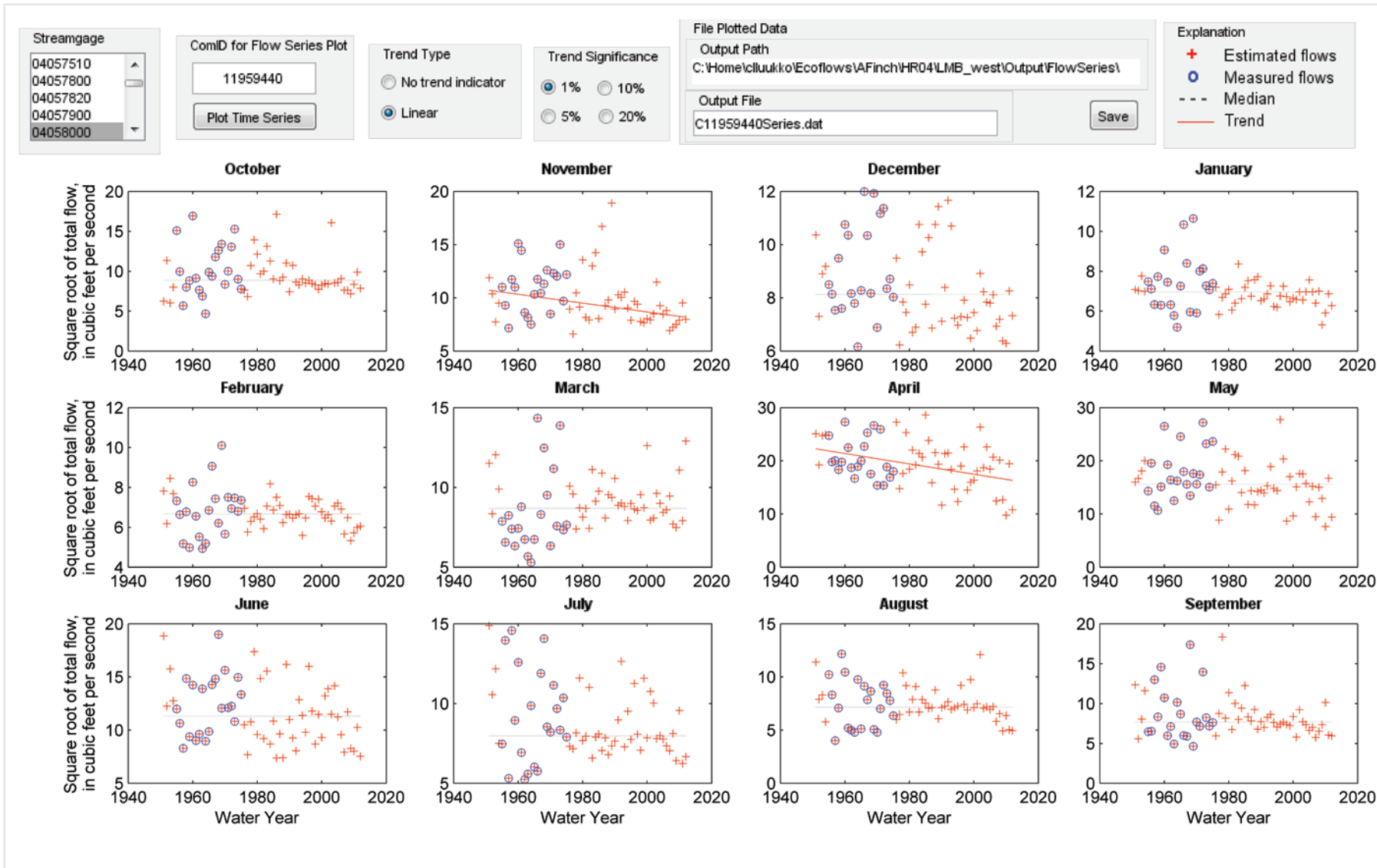


Figure 3. Time-series plot of monthly flows at flowline 11959440, streamgaging station 04058000, Middle Branch Escanaba River near Ishpeming, Michigan.

Methods of Flow and Water-Yield Estimation for Study Areas

Flows and water yields for water years 1951–2012 were estimated for 107,691 flowlines and 105,829 catchments in Hydrologic Region 04 representing the portion of the Great Lakes Basin in the U.S. (fig. 1). The 15 hydrologic subregions were combined to form 7 study areas consisting of multiple subregions and (or) cataloguing units for analysis with AFINCH (table 1). Details on the analysis and estimation of flows and water yields are presented for each study area. Each section includes (1) location map, (2) table of the streamgaging stations in the study area indicating each year each streamgauge was active during the analysis period and whether or not the streamgauge data were included in the analysis, (3) the regression coefficients and fit statistics for each variable, and (4) graph showing the t -values for each explanatory variable by month.

The regression coefficients indicate the size and direction (positive or negative) of the effect that each variable has on the yield estimate. However, some of the differences in coefficient magnitudes are owing to the fact that the potential explanatory variables are in different units, for example, variables with larger values would likely have smaller coefficients. To aid in the interpretation of the effect of each variable, graphs of t -values for all variables in the multiple-linear-regression equation are shown for each study area. The magnitudes of the t -values are test statistics that are used to interpret the likelihood that individual parameters in the regression equation are equal to zero. The sign of the t -values indicate whether the corresponding parameter is positive or negative. The cells in the image are white if the magnitude of the corresponding test statistic did not indicate significance at the level prescribed by the user. Cells are graded with increasing tones

of red for increasingly significant parameters with negative signs. Similarly, cells are graded with increasing tones of blue for increasingly significant parameters with positive signs. The color scale at the bottom indicates the magnitude of the t -statistics. This visual representation permits interpretation of the information used to estimate water yields by each variable and for each month. For example, parameters (represented in each horizontal band) with most months tinted in shades of blue and (or) red likely provide more information in estimating the water-yield response than would a variable where most months are white indicating that the test statistics were not significant. Likewise, months (represented in each vertical band) with most cells tinted blue and (or) red would be estimated using multiple parameters, whereas months where most cells are white would not be estimated using much parameter information. The regression-fit statistics include the root-mean-square error (RMSE), F statistic, and the coefficient of determination (R^2) for the fit to the square-root of water yield. As noted previously, drainage area is so strongly correlated to streamflow that this relation is built into the analysis by writing the regression equation in terms of yield (streamflow divided by drainage area). Examination of the tables reveals that the reported R^2 values for the fit to the square root of yield can be fairly low (0.2–0.7); however, the corresponding R^2 for the fit of estimated to measured streamflows can be quite high (0.8–0.9) because of the strong correlation between area and streamflow.

Selection of explanatory variables was guided by knowledge of each study area, identification of variables likely to affect water yields, and inspection of the images showing monthly parameter t -values. Furthermore, an additional consideration was to improve the estimation of yields and flows during April (higher flow month) and August (lower flow month) in order for the estimates to be more useful for

Table 1. Summary of study area information.

Study area	Hydrologic subregions (4-digits) and (or) cataloguing units (8-digits)	Drainage area (square miles)	Number of streamgages	Number of flowlines	Number of catchments
1	0401 and 0402	16,632	76	15,442	15,221
2	0403, 04040002, 04040003, 04060106, 04060107, 04070001, and 04070002	23,133	177	16,529	16,153
3	0405 and 04040001	13,620	105	11,523	11,300
4	04060101, 04060102, 04060103, 04060104, 04060105, 04070003, 04070004, 04070005, 04070006, and 04070007	14,844	61	7,946	7,808
5	0408 and 0409	12,881	116	13,303	13,151
6	0410 and 0411	14,942	120	16,935	16,713
7	0412, 0413, 0414, and 0415	21,655	168	26,013	25,483

ecological stream-classification work. The variables selected for each study area represent those with the most significance of the combinations that were considered during the regression analysis; however, consideration of all possible variables and combinations was beyond the scope of this report. It is possible that there are areas where local conditions differ greatly from the variable characteristics that were selected for the study area and that in these areas flow and yield estimates are not as representative of actual conditions than they could be with a different variable set. It also is possible that a regionally important variable was not included in the analysis. Equally likely is the potential that major water uses (withdrawals, augmentations, or diversions) were not accounted for during the analysis. However, regionally these estimates present a reasonable, consistent, and defensible set of flow and yield estimates for the U.S. Great Lakes Basin.

Study Area 1

Study area 1 is formed by hydrologic subregions 0401 and 0402 in the western part of the U.S. Great Lakes Basin and encompasses parts of Minnesota, Wisconsin, and the Upper Peninsula of Michigan (fig. 4). The number of stations included in the analysis ranged from 17 (1986, 1998, 1999, 2001, 2007, 2008, and 2009) to 30 (1954 and 1955) (table 2). Nine explanatory variables were selected during the regression analysis with the number of variables specified for each monthly equation ranging from 6 (June) to 8 (October, November, December, January, April, May, and August). Selected explanatory variables are summarized in table 3. Variables and regression statistics based on the stepwise regression for each month are indicated in table 4 and figure 5.



Figure 4. Location of study area 1 (hydrologic subregions 0401 and 0402) in the U.S. Great Lakes Basin.

Table 3. Selected explanatory variables for each study area.

[PRISM, Parameter-elevation Regressions on Independent Slopes Model; NLCD, National Land Cover Data; SSURGO, Soil Survey Geography; GAP, USGS National Gap Analysis Program; GDD, growing degree days]

Explanatory variables	Study area 1	Study area 2	Study area 3	Study area 4	Study area 5	Study area 6	Study area 7
PRISM							
Current precip, in hundredths of millimeters	X	X	X	X	X	X	X
Preced precip, in hundredths of millimeters	X	X	X	X	X	X	X
Current temp, in degrees Celsius * 100	X	X	X	X	X	X	X
NLCD							
Open water, in percent			X				X
Developed, in percent							X
Deciduous and evergreen forest, in percent				X			
Forest, in percent	X	X	X			X	X
Grass herbaceous, in percent					X		
Pasture hay, in percent			X		X		
Row crops, in percent			X	X	X		
Pasture hay and row crops, in percent		X				X	X
Wetland woody, in percent		X					
Wetland herbaceous, in percent		X	X				
NHDPlus catchment characteristics							
Stream order, dimensionless	X		X	X	X	X	
Mean elevation, in centimeters	X		X			X	
Mean slope, in centimeters/meter			X		X	X	X
SSURGO							
Hydric group A and AD, in percent				X			
Hydric group B and BD, in percent					X		
Hydric group C and CD, in percent		X					
Hydric groups C, CD, and D, in percent	X						X
GAP							
Coarse outwash, in percent			X		X		
Fine end and ground moraines, in percent						X	
Medium end and ground moraines, in percent						X	
Coarse end and ground moraines, in percent	X						X
Coarse outwash and ice contact, in percent						X	X
Coarse stratified sediment, in percent		X					
Till, in percent		X					
GDD							
Average GDD, in days			X		X		X

Table 4. Regression statistics and explanatory variables used in the AFINCH (Analysis of Flows In Networks of Channels) analysis of flows in study area 1 (hydrologic subregions 0401 and 0402), in the U.S. Great Lakes Basin, by month.

[–, indicates variable not included in monthly regression equation; RMSE, root mean square error; R², coefficient of determination. See appendix tables for variable descriptions]

Month	Regression coefficients									Regression fit statistics		
	Intercept	Current precipitation	Preceding precipitation	Current temperature	Forest	Coarse end and ground moraines	Mean elevation	Hydric groups C, CD, and D	Stream order	RMSE	F-statistic	R ²
October	0.797	0.116	0.080	–	–	–	–	–0.326	–0.203	0.236	391.3	0.537
November	1.133	0.120	0.068	–0.011	–	0.073	–6.0E-06	–0.379	–0.180	0.221	144.4	0.428
December	1.314	0.060	0.062	0.015	0.113	0.145	–7.0E-06	–0.496	–0.158	0.184	144.4	0.462
January	1.279	0.027	0.054	0.008	0.225	0.193	–9.0E-06	–0.547	–0.132	0.192	124.3	0.425
February	1.261	–	0.049	0.015	0.307	0.196	–9.0E-06	–0.498	–0.111	0.193	125.7	0.395
March	0.926	0.091	0.047	0.068	0.681	0.074	–7.0E-06	–	–	0.254	248.5	0.525
April	0.395	0.101	0.089	–0.036	0.887	–0.285	–	0.709	0.278	0.465	62.1	0.244
May	1.064	0.146	0.066	–0.048	–	–0.262	–	–0.121	0.169	0.343	137.1	0.379
June	0.818	0.133	0.071	–0.016	–	–0.134	–4.0E-06	–0.241	–	0.246	184	0.450
July	1.296	0.123	0.061	–0.037	0.150	–	–1.0E-05	–0.385	–	0.225	263.7	0.540
August	0.044	0.086	0.060	0.028	0.290	–	–9.0E-06	–0.470	–	0.217	213	0.486
September	0.781	0.100	0.064	–	0.156	–	–6.0E-06	–0.458	–0.158	0.222	229.5	0.505

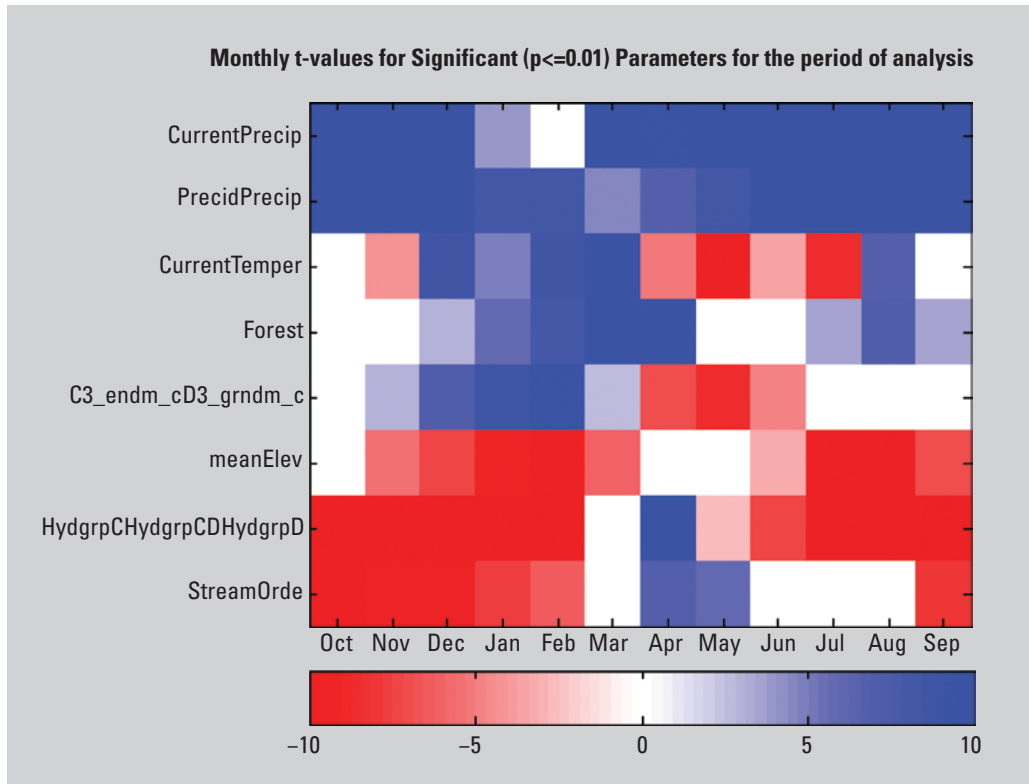


Figure 5. Image showing the t-value-indexed colors corresponding to each explanatory variable in the monthly regression equation for estimating water yield for study area 1 (hydrologic subregions 0401 and 0402) in the U.S. Great Lakes Basin.

Study Area 2

Study area 2 is formed by hydrologic subregion 0403 and cataloguing units 04040002, 04040003, 04060106, 04060107, 04070001, and 04070002 on the western and northern side of Lake Michigan and encompasses parts of Wisconsin, northern Illinois, and the Upper Peninsula of Michigan (fig. 6). The number of stations included in the analysis ranged from

21 (1951 and 1953) to 47 (2005) (table 5). Ten explanatory variables were selected during the regression analysis with the number of variables specified for each monthly equation ranging from 6 (March and June) to 10 (February) (table 3). Variables and regression statistics based on the stepwise regression for each month are indicated in table 6 and figure 7.



Figure 6. Location of study area 2 (hydrologic subregion 0403 and cataloguing units 04040002, 04040003, 04060106, 04060107, 04070001, and 04070002) in the U.S. Great Lakes Basin.

Table 6. Regression statistics and explanatory variables used in the AFINCH (Analysis of Flows In Networks of Channels) analysis for study area 2 (hydrologic subregion 0403 and cataloguing units 04040002, 04040003, 04060106, 04060107, 04070001, and 04070002) in the U.S. Great Lakes Basin, by month.

[–, indicates variable not included in monthly regression equation; RMSE, root mean squared error; R², coefficient of determination. See appendix tables for variable descriptions]

Month	Regression coefficients											Regression fit statistics		
	Intercept	Current precipitation	Preceding precipitation	Current temperature	Pasture hay and row crops	Forest	Wetland woody	Wetland herbaceous	Hydric group C and CD	Coarse stratified sediment	Till	RMSE	F-statistic	R ²
November	0.336	0.139	0.076	-0.006	-0.167	0.131	–	–	-0.100	0.125	–	0.194	408.2	0.579
January	0.411	0.091	0.086	0.020	-0.077	0.510	-0.585	0.525	–	0.309	0.111	0.194	150.4	0.395
March	0.653	0.119	0.097	0.029	0.225	–	–	0.839	0.293	–	–	0.285	282.8	0.450
May	1.238	0.134	0.068	-0.033	-0.221	–	0.537	–	–	-0.430	-0.390	0.271	345.3	0.538
July	1.057	0.103	0.066	-0.028	-0.518	-0.232	-0.800	–	-0.176	0.231	–	0.197	404.7	0.609
September	0.411	0.106	0.050	–	-0.410	–	-0.600	-0.341	-0.189	0.159	–	0.203	446	0.601

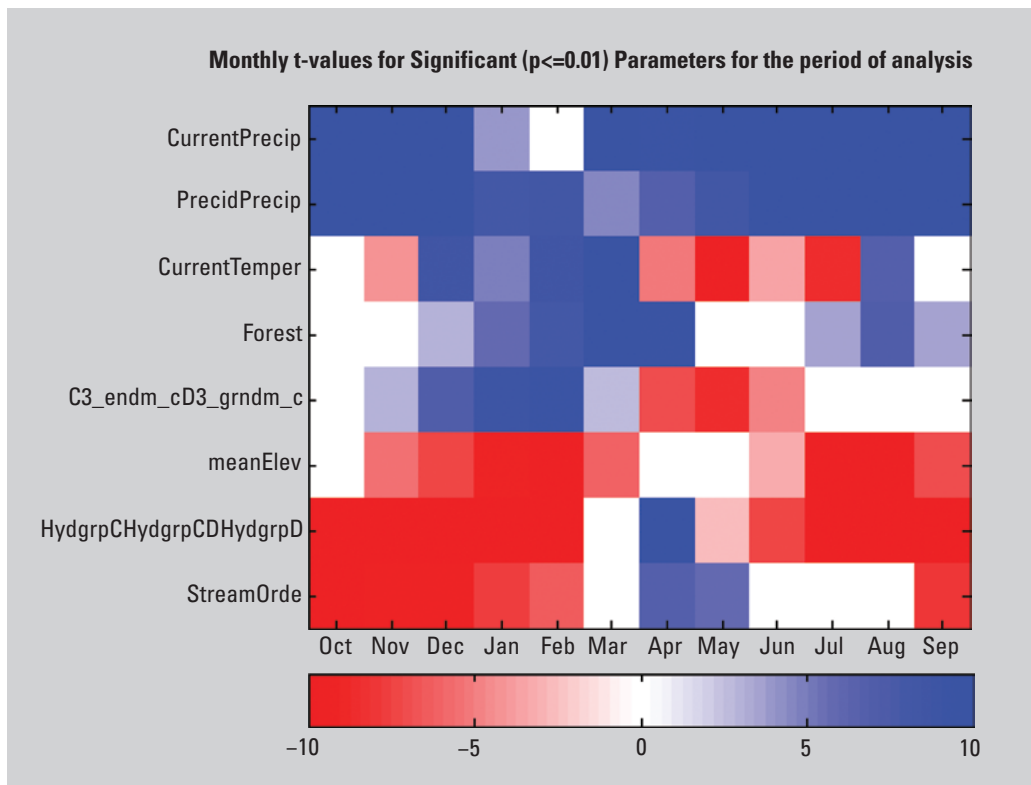


Figure 7. Image showing the t-value-indexed colors corresponding to each explanatory variable in the monthly regression equation for estimating water yield for study area 2 (hydrologic subregion 0403 and cataloguing units 04040002, 04040003, 04060106, 04060107, 04070001, and 04070002) in the U.S. Great Lakes Basin.

Study Area 3

Study area 3 is formed by hydrologic subregion 0405 and cataloguing unit 04040001 in the south-central part of the U.S. Great Lakes Basin along the southeastern side of Lake Michigan and encompasses parts of Michigan, Indiana, and a small part of Illinois (fig. 8). The number of stations included in the analysis ranged from 27 (1951) to 54 (1976)

(table 7). Thirteen explanatory variables were selected during the regression analysis with the number of variables specified for each monthly equation ranging from 5 (April) to 12 (September) (table 3). Variables and regression statistics based on the stepwise regression for each month are indicated in table 8 and figure 9.

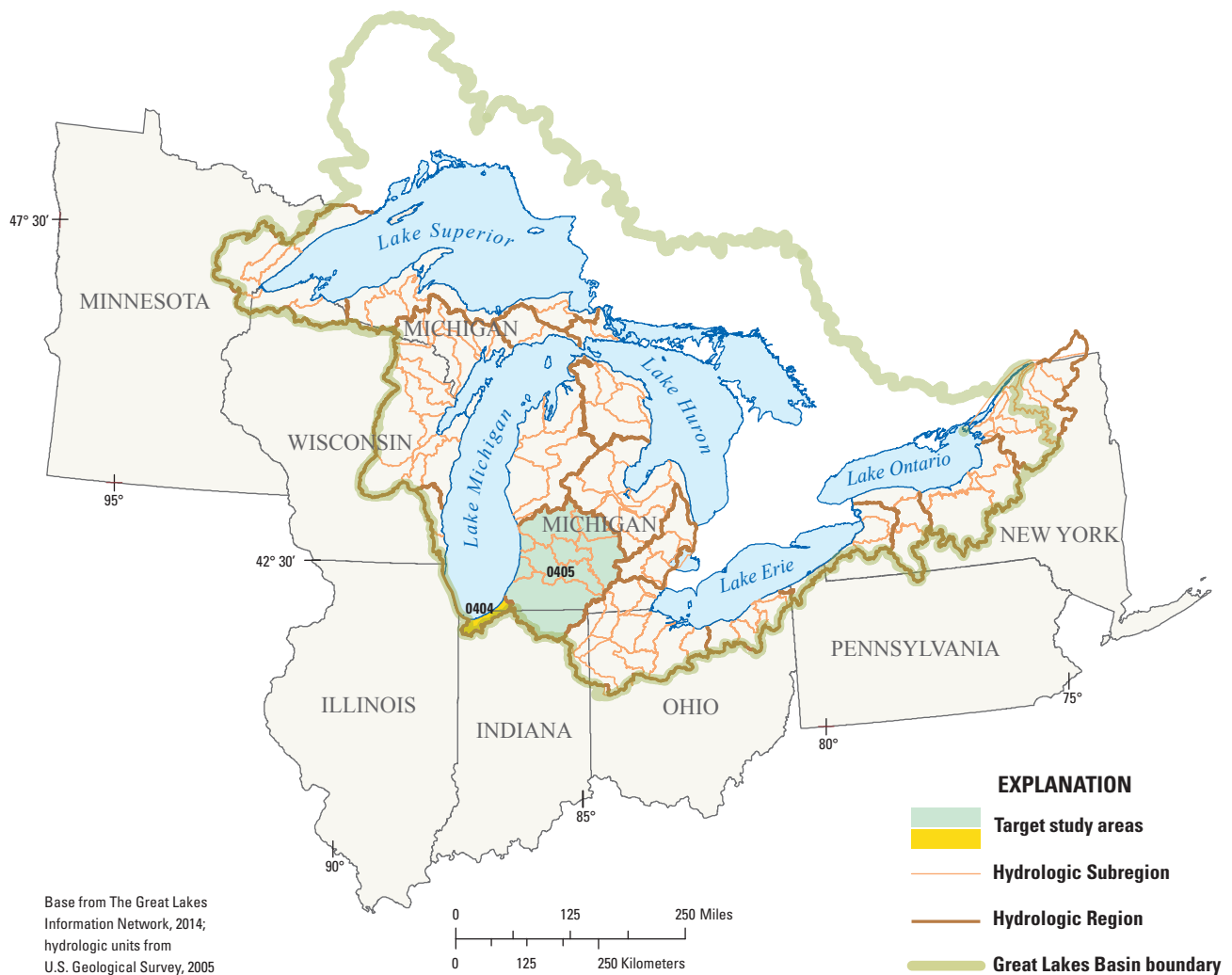


Figure 8. Location of study area 3 (hydrologic subregion 0405 and cataloguing unit 04040001) in the U.S. Great Lakes Basin.

Table 7. Streamgaging stations considered for analysis of study area 3 (hydrologic subregion 0405 and cataloguing unit 04040001) in the U.S. Great Lakes Basin, by water year.—Continued

[●, station active and included in the analysis; ○, station active but not included in the analysis]

Station number	Water year																																	
	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982		
04112850																											●	●	●	●	●			
04112904																											●							
04113000	●	●	●	●	●	●	●	●	○	●	●	●	●	●	●	●	○	●	●	●	●	●	●	●	●	●	●	●	○	●	●	●	●	
04113097																											●	●	●	●	●			
04114000			○	○	●	●	●	●	●	●	●	●	●	●	●	●	○	●	●	●	●	●	●	●	●	●	●	●	●	○	●	●	●	
04114498	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
04114500	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	
04115000	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	
04115265																																		
04116000		●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
04116500	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
04117000					●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
04117500	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
04118000		●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
04118500			●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
04119000	●	●	●	●	●	●	○	●	●	○	●	●	●	●	●	●	○	○	●	●	○	●	●	●	●	●	●	●	●	●	●	○	●	
04119300																											●							

Table 8. Regression statistics and explanatory variables used in the AFINCH (Analysis of Flows In Networks of CHannels) analysis for study area 3 (hydrologic subregion 0405 and cataloguing unit 04040001) in the U.S. Great Lakes Basin, by month.

[–, indicates variable not included in monthly regression equation; GDD, growing degree days; RMSE, root mean square error; R², coefficient of determination. See appendix tables for variable descriptions]

Month	Regression coefficients														Regression fit statistics		
	Intercept	Current precipitation	Preceding precipitation	Current temperature	Pasture hay	Row crops	Forest	Open water	Wetland herbaceous	Stream order	Mean elevation	Mean slope	Coarse outwash	Average GDD	RMSE	F-statistic	R ²
October	0.529	0.091	0.055	–	–0.289	–	0.277	–2.622	–	–2.2E-05	0.110	0.077	0.501	–	0.181	496.8	0.641
November	0.648	0.103	0.065	–0.011	–0.334	–	0.490	–1.802	–2.646	–2.0E-05	0.065	–	0.359	4.8E-05	0.181	355.4	0.609
December	0.710	0.110	0.074	0.016	–0.337	–	0.545	–1.131	–3.075	–1.4E-05	–	–	0.236	4.5E-05	0.199	296.8	0.542
January	0.805	0.120	0.104	0.032	–0.286	–	0.305	–	–4.094	–1.2E-05	–	–	0.183	4.9E-05	0.223	315.5	0.531
February	0.972	0.155	0.049	0.021	–0.586	–	–	–1.331	–3.384	–1.3E-05	0.042	–	0.145	5.0E-05	0.228	202.8	0.447
March	1.161	0.118	0.094	–	–0.331	–	–	–2.131	–	–7.0E-06	–	–	–0.057	–	0.235	227.4	0.352
April	1.039	0.080	0.070	–0.013	–0.568	–	–	–1.655	–	–	–	–	–	–	0.214	209.4	0.294
May	0.818	0.094	0.074	–0.015	–0.439	–	–	–2.226	–	–9.0E-06	0.043	0.090	0.202	–	0.184	354.7	0.560
June	0.412	0.103	0.072	–	–0.593	–	–	–2.306	–2.858	–1.2E-05	0.070	0.082	0.355	–	0.201	362.8	0.566
July	0.372	0.069	0.066	–	–0.661	–0.115	–	–2.336	–4.678	–1.5E-05	0.099	0.127	0.472	–	0.17	335.5	0.572
August	0.449	0.057	0.058	–	–0.595	–0.187	–	–2.584	–4.794	–1.9E-05	0.120	0.139	0.509	–	0.171	305.6	0.549
September	0.711	0.076	0.047	–0.016	–0.571	–0.172	–	–2.538	–5.036	–2.2E-05	0.114	0.129	0.523	2.9E-05	0.171	355.1	0.630

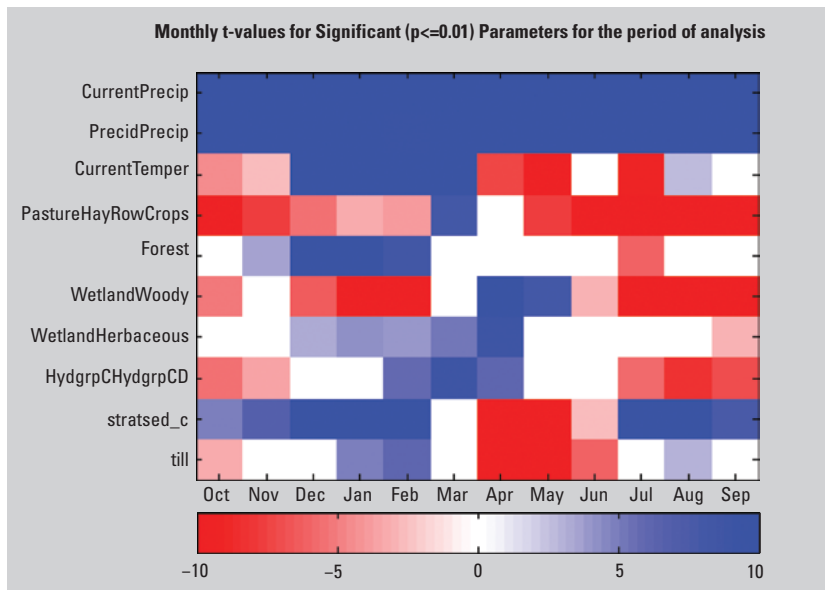


Figure 9. Image showing the t-value-indexed colors corresponding to each explanatory variable in the monthly regression equation for estimating water yield for study area 3 (hydrologic subregion 0405 and cataloguing unit 04040001) in the U.S. Great Lakes Basin.

Study Area 4

Study area 4 is formed by cataloguing units 04060101, 04060102, 04060103, 04060104, 04060105, 04070003, 04070004, 04070005, 04070006, and 04070007 near the central part of the U.S. Great Lakes Basin in the northern part of the Lower Peninsula of Michigan (fig. 10). The number of stations included in the analysis ranged from 15 (1994) to

30 (1970) (table 9). Seven explanatory variables were selected during the regression analysis with the number of variables specified for each monthly equation ranging from 5 (May) to 7 (December, January, February, and July) (table 3). Variables and regression statistics based on the stepwise regression for each month are indicated in table 10 and figure 11.



Figure 10. Location of study area 4 (cataloguing units 04060101, 04060102, 04060103, 04060104, 04060105, 04070003, 04070004, 04070005, 04070006, and 04070007) in the U.S. Great Lakes Basin.

Table 10. Regression statistics and explanatory variables used in the AFINCH (Analysis of Flows In Networks of CHannels) analysis for study area 4 (cataloguing units 04060101, 04060102, 04060103, 04060104, 04060105, 04070003, 04070004, 04070005, 04070006, and 04070007) in the U.S. Great Lakes Basin, by month.

[–, indicates variable not included in monthly regression equation; RMSE, root mean square error; R², coefficient of determination. See appendix tables for variable descriptions]

Month	Regression coefficients								Regression fit statistics		
	Intercept	Current precipitation	Preceding precipitation	Current temperature	Deciduous and evergreen forest	Row crops	Hydric group A and AD	Stream order	RMSE	F-statistic	R ²
November	–0.086	0.053	0.043	–	0.839	0.763	0.200	0.122	0.147	208.1	0.481
January	0.093	0.040	0.054	0.021	0.766	0.979	0.243	0.139	0.152	166.6	0.464
March	0.497	0.072	0.057	0.028	0.366	0.830	–	0.138	0.177	144	0.390
May	0.332	0.062	0.033	–	0.565	0.285	–	0.122	0.154	154.8	0.364
July	–0.379	0.039	0.035	–0.008	0.945	0.724	0.468	0.207	0.141	268.4	0.582
September	–0.642	0.043	0.030	–	1.127	0.958	0.457	0.186	0.144	359.4	0.615

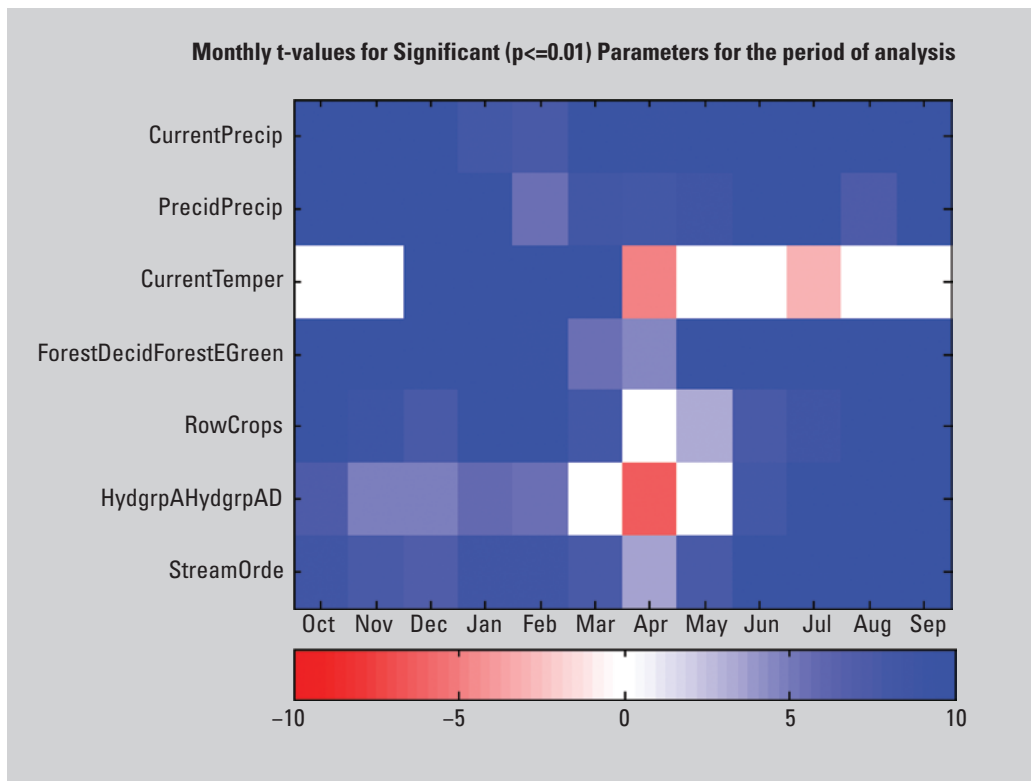


Figure 11. Image showing the t-value-indexed colors corresponding to each explanatory variable in the monthly regression equation for estimating water yield for study area 4 (cataloguing units 04060101, 04060102, 04060103, 04060104, 04060105, 04070003, 04070004, 04070005, 04070006, and 04070007) in the U.S. Great Lakes Basin.

Study Area 5

Study area 5 is formed by hydrologic subregions 0408 and 0409 near the central part of the U.S. Great Lakes Basin along the southwestern side of Lake Huron in the Lower Peninsula of Michigan (fig. 12). The number of stations included in the analysis ranged from 21 (2008) to 52 (1966) (table 11). Eleven explanatory variables were selected during

the regression analysis with the number of variables specified for each monthly equation ranging from 6 (March) to 10 (August) (table 3). Variables and regression statistics based on the stepwise regression for each month are indicated in table 12 and figure 13.



Figure 12. Location of study area 5 (hydrologic subregions 0408 and 0409) in the U.S. Great Lakes Basin.

Table 11. Streamgaging stations considered for analysis of study area 5 (hydrologic subregions 0408 and 0409) in the U.S. Great Lakes Basin, by water year.—Continued

[●, station active and included in the analysis; ○, station active but not included in the analysis]

Station number	Water year																																
	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	
04164600															●	●	●	●	●														
04164800												●	●	●	●	○	○		●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
04164980																																	
04165200											●	●	●	●	●																		
04165500	●	●	●	●	●	●	●	●	●	●	●	○	○	○	○	○	○	○	○	●	○	●	●	●	●	●	●	●	●	●	●	●	
04165557																																	
04165559																																	
04166000	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
04166040																																	
04166100									●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
04166200									●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
04166300									●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
04166470																																	
04166500	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	
04166750																																	
04167000	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	
04167150																																	
04167625																																	
04168000	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	○	●	●	●	○	●	●	○	●	●	●	●	●	●	●	●	
04168400																																	
04168530																																	
04168580																																	
04169500	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
04170000	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	○	●	●	●
04170500	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	
04171500		●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
04172000		●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	○	●	●	●
04172500	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
04173000	●	●	●	●	●	○	●	●	●	●	●	●	○	○	●	○	●	○	●	○	●	○	●				●	●					
04173500			●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	○	●	●	●
04174500	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
04174518																																	
04174800																											○	○	○	○	○	○	○

Table 12. Regression statistics and explanatory variables used in the AFINCH (Analysis of Flows In Networks of CHannels) analysis for study area 5 (hydrologic subregions 0408 and 0409) in the U.S. Great Lakes Basin, by month.

[–, indicates variable not included in monthly regression equation; GDD, growing degree days; RMSE, root mean square error; R², coefficient of determination. See appendix tables for variable descriptions]

Month	Regression coefficients												Regression fit statistics		
	Intercept	Current precipitation	Preceding precipitation	Current temperature	Pasture hay	Row crops	Grass herbaceous	Stream order	Mean slope	Hydric group B and BD	Coarse outwash	Average GDD	RMSE	F-statistic	R ²
November	0.0815	0.1212	0.0795	-0.0137	-0.5761	-0.1675	–	0.1342	0.1100	-0.1647	–	–	0.203	292.3	0.512
January	0.1645	0.1383	0.1078	0.0443	-0.5149	0.1856	1.5386	0.2060	0.1442	-0.2622	–	–	0.232	280.1	0.530
March	1.0156	0.1500	0.1153	–	-0.2659	0.3633	–	–	–	–	-0.1671	-9.5E-05	0.284	174.3	0.319
May	0.2966	0.1606	0.0782	-0.0242	-0.7910	–	2.5436	0.1041	0.0765	–	-0.1402	4.0E-05	0.184	515.9	0.675
July	-0.3059	0.0998	0.0745	–	-1.0679	–	2.3501	0.1598	0.1173	-0.2079	-0.1211	4.8E-05	0.186	286	0.536
September	0.1771	0.1088	0.0530	-0.0180	-0.7754	-0.2785	–	0.1527	0.1148	-0.1551	-0.1159	–	0.200	328.7	0.57

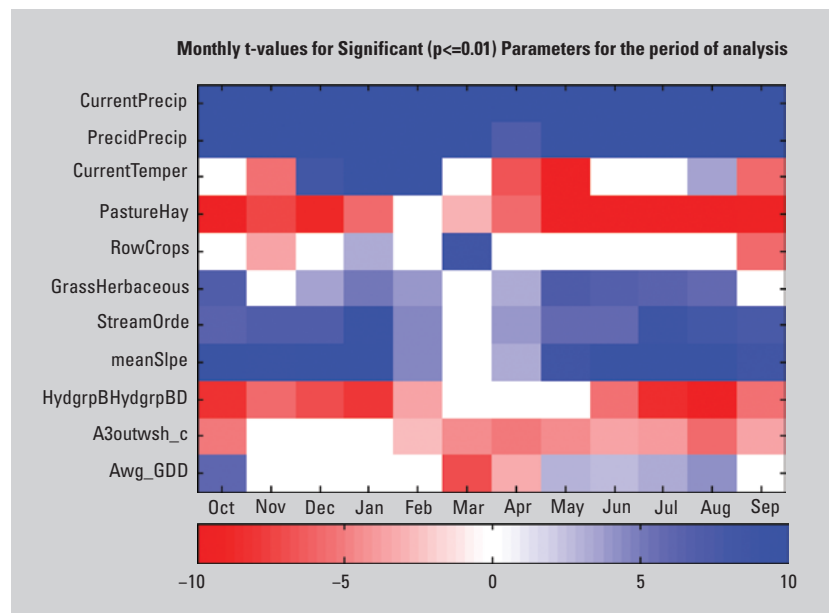


Figure 13. Image showing the t-value-indexed colors corresponding to each explanatory variable in the monthly regression equation for estimating water yield for study area 5 (hydrologic subregions 0408 and 0409) in the U.S. Great Lakes Basin.

Study Area 6

Study area 6 is formed by hydrologic subregions 0410 and 0411 in the southeastern part of the U.S. Great Lakes Basin and encompasses parts of Michigan, Ohio, Indiana, and a small part of Pennsylvania (fig. 14). The number of stations included in the analysis ranged from 26 (1953, 1963, 1982, and 1983) to 48 (2002) (table 13). Eleven explanatory

variables were selected during the regression analysis with the number of variables specified for each monthly equation ranging from 7 (July) to 11 (April and September) (table 3). Variables and regression statistics based on the stepwise regression for each month are indicated in table 14 and figure 15.

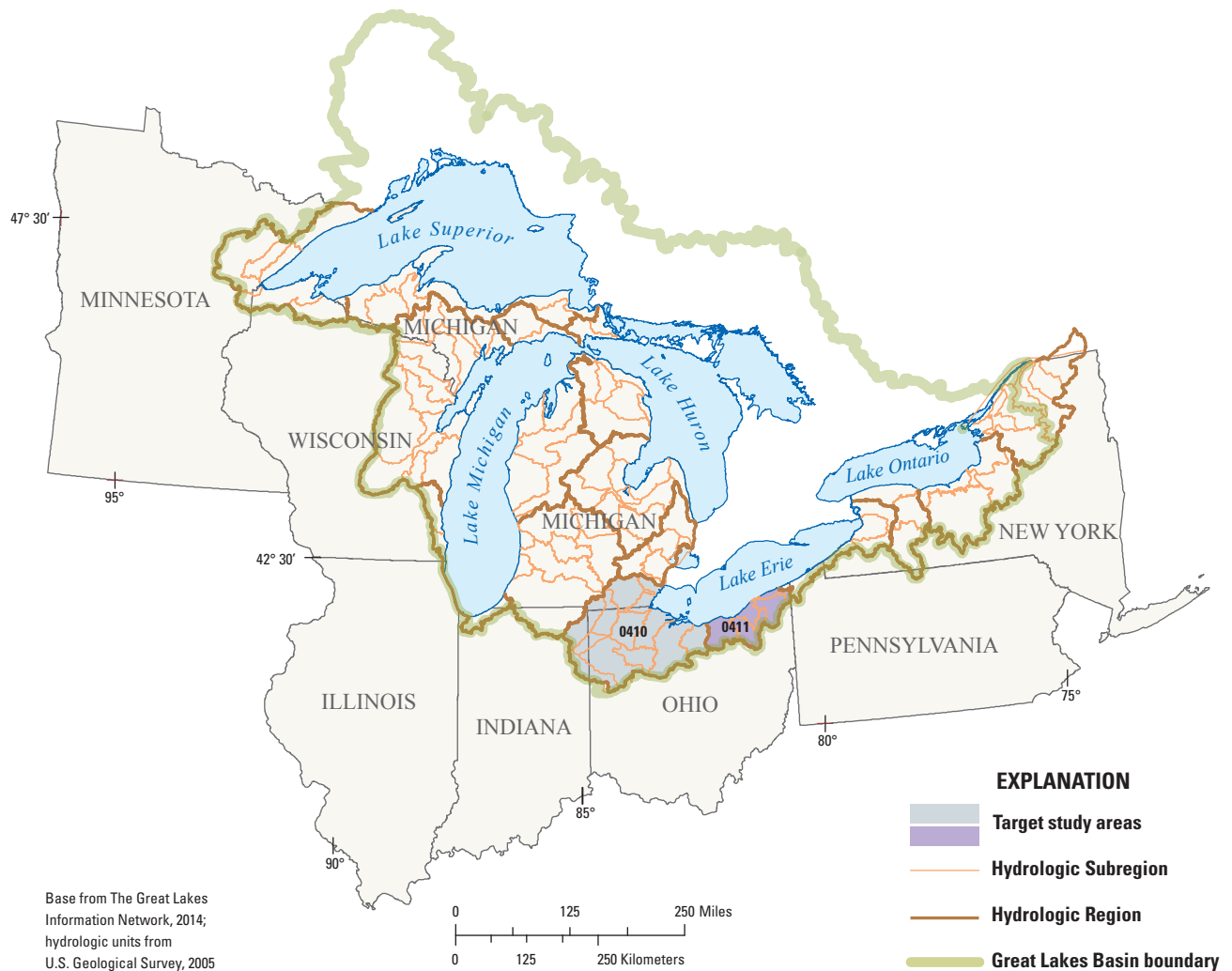


Figure 14. Location of study area 6 (hydrologic subregions 0410 and 0411) in the U.S. Great Lakes Basin.

Table 13. Streamgaging stations considered for analysis of study area 6 (hydrologic subregions 0410 and 0411) in the U.S. Great Lakes Basin, by water year.—Continued

[•, station active and included in the analysis; ○, station active but not included in the analysis]

Station number	Water year																														
	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	
04191500	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	
04192000																															
04192500	•	○	•	•	○	○	○	○	○	•	○	○	•	•	•	•	•	○	○	•	•	•	•	•	•	○	•	•	○	○	○
04193500	○	○	○	○	○	○	•	•	•	•	•	○	○	•	○	○	○	•	○	○	○	○	○	○	○	○	•	○	•	○	•
04194085																										○	○	○	○	○	○
04195500	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
04195600																															
04195820																		•	•	•	•	•	•	•	•	•	•	○	•	○	•
04195825							•				•																				
04195830							•				•																				
04196000														•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
04196200																															
04196500																						•	•	•	•	•	•	○	•	•	•
04196800	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
04197000																															
04197020	•	•	•	•	•	•	•																								
04197100	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
04197170		•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
04197300																															
04197450																															
04198000	•	•	•	•	•	•	•	•	•	•	•	○	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	○	•	○
04199000							•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
04199155							•	•	•	•	•	•		•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
04199287							•	•		•	•	•																			
04199500																						•	•	•	•	•	•	•	•	•	•
04200430	○	○																													
04200500	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
04201300																															
04201404																							•								
04201495																								•							
04201500	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
04202000	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
04203900																															
04204000																															
04204500																															
04205000																															
04205700																															
04206000	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
04206014																							•	•	•	•					
04206021																							•	•	○	•					
04206029																							•	•	•	•					
04206038																							•	○	○	•					
04206043																							•	○	○	○					

Table 14. Regression statistics and explanatory variables used in the AFINCH (Analysis of Flows In Networks of CHannels) analysis for study area 6 (hydrologic subregions 0410 and 0411) in the U.S. Great Lakes Basin, by month.

[–, indicates variable not included in monthly regression equation; RMSE, root mean square error; R², coefficient of determination. See appendix tables for variable descriptions]

Month	Regression coefficients												Regression fit statistics		
	Intercept	Current precipitation	Preceding precipitation	Current temperature	Pasture hay and row crops	Forest	Fine end and ground moraines	Medium end and ground moraines	Coarse outwash and ice contract	Stream order	Mean elevation	Mean slope	RMSE	F-statistic	R ²
November	0.421	0.226	0.090	-0.015	-0.355	-0.063	0.076	0.259	–	–	-1.5E-05	0.074	0.267	516.5	0.661
January	0.495	0.288	0.081	0.037	–	0.406	0.128	0.448	–	0.084	-1.9E-05	0.075	0.310	442.9	0.653
March	0.930	0.254	0.047	-0.035	–	0.677	0.177	0.527	–	0.095	-2.3E-05	0.047	0.335	224.9	0.488
May	0.985	0.200	0.072	-0.018	-0.218	–	–	0.263	0.105	–	-2.4E-05	0.104	0.244	528.7	0.666
July	0.029	0.160	0.079	–	-0.406	-0.668	–	–	–	0.077	-7.0E-06	0.106	0.246	478.4	0.612
September	0.338	0.141	0.077	-0.011	-0.521	-0.601	0.113	0.156	0.339	0.099	-1.2E-05	0.080	0.218	364	0.654

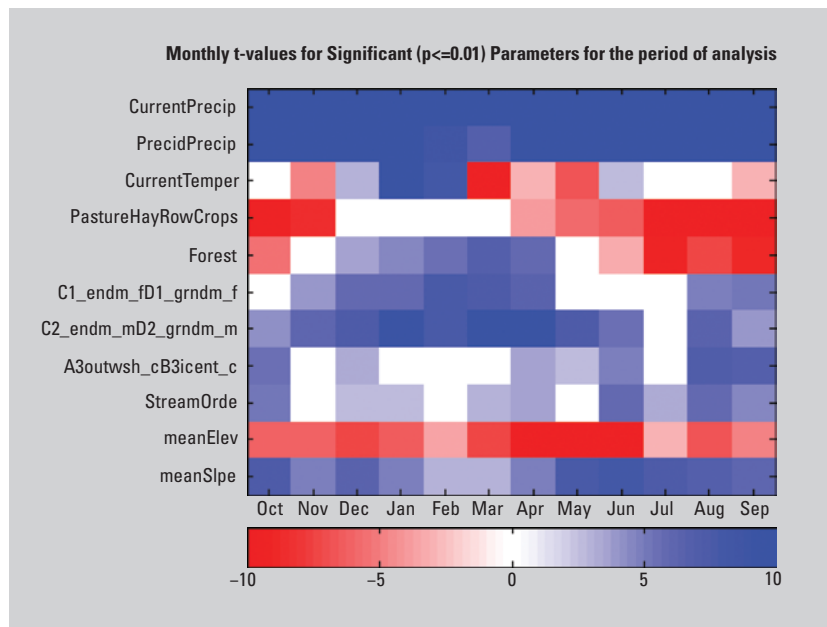


Figure 15. Image showing the t-value-indexed colors corresponding to each explanatory variable in the monthly regression equation for estimating water yield for study area 6 (hydrologic subregions 0410 and 0411) in the U.S. Great Lakes Basin.

Study Area 7

Study area 7 is formed by hydrologic subregions 0412, 0413, 0414, and 0415 in the eastern part of the U.S. Great Lakes Basin and encompasses parts of Pennsylvania, New York, and a small part of Ohio (fig. 16). The number of stations included in the analysis ranged from 31 (1951 and 1953) to 65 (1967) (table 15). Twelve explanatory variables

were selected during the regression analysis with the number of variables specified for each monthly equation ranging from 6 (February) to 11 (March) (table 3). Variables and regression statistics based on the stepwise regression for each month are indicated in table 16 and figure 17.



Base from The Great Lakes Information Network, 2014; hydrologic units from U.S. Geological Survey, 2005

Figure 16. Location of study area 7 (hydrologic subregions 0412, 0413, 0414, and 0415) in the U.S. Great Lakes Basin.

Table 15. Streamgaging stations considered for analysis of study area 7 (hydrologic subregions 0412, 0413, 0414, and 0415) in the U.S. Great Lakes Basin, by water year.—Continued

[●, station active and included in the analysis; ○, station active but not included in the analysis]

Station number	Water year																																					
	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982						
0423204920																																						
0423205010																																			○			
0423205023																																			○			
0423205025																																						
0423368620																																						
0424014980																																						
0424016205																																						
0424016825																																						
0424016975																																						
0426545290																																			●	●		
0426545295																																			○	○	○	○

Table 16. Regression statistics used in the AFINCH (Analysis of Flows In Networks of CHannels) analysis for study area 7 (hydrologic subregions 0412, 0413, 0414, and 0415) in the U.S. Great Lakes Basin, by month.

[–, indicates variable not included in monthly regression equation; GDD, growing degree days; RMSE, root mean square error; R², coefficient of determination. See appendix tables for variable descriptions]

Month	Regression coefficients													Regression fit statistics		
	Intercept	Current precipitation	Preceding precipitation	Current temperature	Open water	Forest	Developed	Mean slope	Hydric group C, CD, and D	Coarse outwash and ice contact	Coarse end and ground moraines	Average GDD	Pasture hay and row crops	RMSE	F-statistic	R ²
November	0.462	0.179	0.107	–0.011	–	–	–0.221	–0.726	–0.030	–	–	–	5.1E-05	0.266	709.8	0.658
January	1.061	0.170	0.085	0.045	–	–	–0.480	–0.403	–0.017	0.071	–	–	–	0.285	398.1	0.519
March	2.051	0.155	0.046	0.039	–3.962	–0.520	–1.088	–0.817	–	0.253	–0.203	–0.263	–1.1E-04	0.349	101.1	0.302
May	1.208	0.165	0.067	–0.024	1.809	–0.383	–0.543	–0.778	–	–	0.178	0.140	–	0.222	847.6	0.748
July	0.772	0.119	0.077	–0.018	1.089	–0.459	–	–0.562	–	–0.115	0.359	0.082	–	0.182	732.2	0.719
September	0.061	0.152	0.094	–	1.058	–0.305	–	–0.671	–	–0.133	0.304	–	4.0E-05	0.224	761.5	0.703

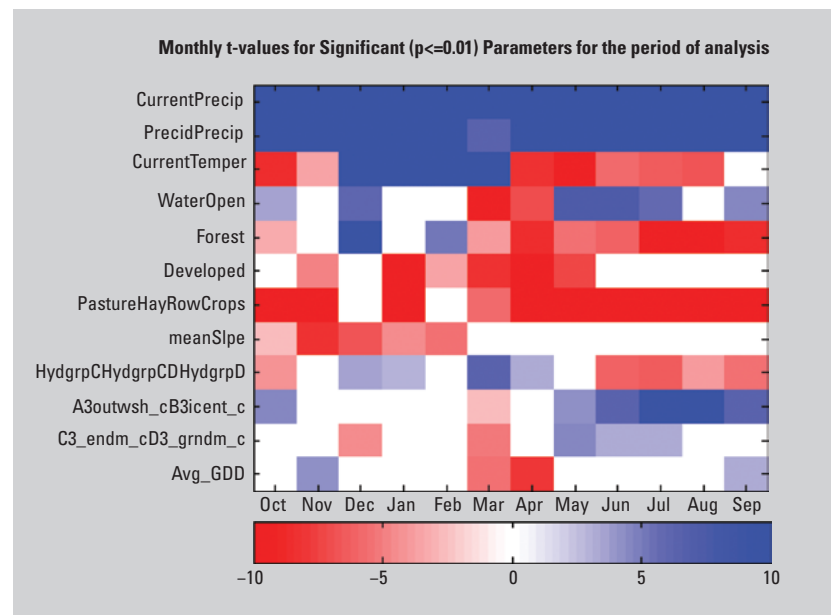


Figure 17. Image showing the t-value-indexed colors corresponding to each explanatory variable in the monthly regression equation for estimating water yield for study area 7 (hydrologic subregions 0412, 0413, 0414, and 0415) in the U.S. Great Lakes Basin.

Summary

A regionally consistent estimate of streamflow provides unified information across the United States portion of the Great Lakes Basin for restoration, assessment, management, and conservation of stream ecosystems. Monthly water yields from catchments and corresponding flows for stream segments (flowlines) were estimated for water years 1951–2012 for the U.S. Great Lakes Basin by using the computer application Analysis of Flows In Networks of CHannels (AFINCH). The AFINCH approach to estimate ungaged flows is a two-step process. The predictor step is based on a user-specified multiple-linear-regression equation to estimate monthly yields using selected monthly climatic and basin characteristics data and flows at streamgages. The corrector step adjusts the regression estimates to match measured flows at the streamgages.

The 15 hydrologic subregions comprising the U.S. Great Lakes Basin were combined to form 7 study areas consisting of multiple subregions and (or) cataloguing units for analysis with AFINCH. Input data, explanatory variables, and regression details used in the estimation of water yields and flows are described for each study area. Data from 823 streamgages were compiled for the AFINCH analysis and were used to estimate monthly flows and water yields for 107,691 stream segments (flowlines) and 105,829 catchments in Hydrologic Region 04 for 1951–2012. These long-term flow series can be used to describe monthly flow-duration characteristics and to investigate trends for the stream segments in the U.S. Great Lakes Basin.

AFINCH utilizes the NHDPlus geospatial framework to provide a basis for integrating monthly flow measured at U.S. Geological Survey (USGS) streamgages, available water-use data, monthly climatic data, land-cover characteristics, and catchment attributes described in user-defined shapefiles to develop sets of multiple-linear-regression equations for estimating monthly water yields. The explanatory variables describing monthly precipitation for the current month, monthly precipitation for the preceding month, and average monthly temperatures were included in the multiple-linear-regression equations for all seven study areas. National land-cover characteristics defining forest, pasture, and cultivated areas were significant explanatory variables in most of the study areas. NHDPlus variables defining stream order, as well as catchment elevation and (or) slope, and Soil Survey Geography (SSURGO) soil data also were significant explanatory variables in most of the study areas. Surficial areas of coarse outwash tended to be included in the regression equations

for study areas in the southern and eastern parts of the U.S. Great Lakes Basin. Growing degree day data and areas with end and ground moraines also were important in several study areas. In general, coefficient of determination (R^2) values and F-statistics for the study areas were lowest in the late spring and early summer months (February–April) and higher in the late summer and early winter months (July–October). Considering a potential additional variable, such as an index relating to snow pack, would likely be significant and might improve the estimates during months of lower flow conditions. Regionally, however, these estimates present a reasonable, consistent, and defensible set of flow and yield estimates for the U.S. Great Lakes Basin.

References Cited

- Council of Great Lakes Governors, 2005, Great Lakes–St. Lawrence River Basin Water Resources Compact, accessed May 22, 2014, at http://www.cglg.org/projects/water/docs/12-13-05/Great_Lakes-St_Lawrence_River_Basin_Water_Resources_Compact.pdf.
- Daly, Chris, and Taylor, G.H., 1998a, United States average monthly or annual precipitation, 1961–90: Corvallis, Ore., Spatial Climate Analysis Service at Oregon State University, accessed May 20, 2009, at http://www.climate-source.com/cd1/ppt_met_us.html.
- Daly, Chris, and Taylor, G.H., 1998b, United States average monthly or annual mean temperature, 1961–90: Corvallis, Ore., Spatial Climate Analysis Service at Oregon State University, accessed May 20, 2009, at http://www.climate-source.com/cd1/ppt_met_us.html.
- Draper, N.R., and Smith, Harry, 1998, Applied regression analysis (3d ed.): New York, John Wiley & Sons, Inc., 736 p.
- ESRI (Environmental Systems Research Institute), 2012, ArcMap 10.1: Redlands, Calif., available at <http://www.esri.com/software/arcgis/extensions/spatialanalyst/key-features/statistical>.
- Fry, J.A., Xian, G., Jin, S., Dewitz, J.A., Homer, C.G., Yang, L., Barnes, C.A., Herold, N.D., and Wickham, J.D., 2011, Completion of the 2006 National Land Cover Database for the Conterminous United States: Photogrammetric Engineering and Remote Sensing, v. 77, no. 9, p. 858–864.
- Gesch, Dean; Evans, Gayla; Mauck, James; Hutchinson, John; and Carswell, W.J., Jr., 2009, *The National Map—Elevation: U.S. Geological Survey Fact Sheet 2009–3053*, 4 p., (Also available at <http://pubs.usgs.gov/fs/2009/3053/>.)

- Government of Canada and U.S. Environmental Protection Agency, 1995, *The Great Lakes—An environmental atlas and resource book* (3d ed.): U.S. Environmental Protection Agency EPA-905-B-95-00 1 and Environment Canada EN40-349/1995E, 46 p., map scales differ. (Also available at <http://www.epa.gov/glnpo/atlas/>.)
- 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/>.)
- Horizon Systems Corporation, 2012, NHDPlus version 2.1, accessed June 2012, at http://www.horizon-systems.com/NHDPlus/NHDPlusV2_home.php.
- Horizon Systems Corporation, 2014, NHDPlusV2 flow table navigator toolbar, accessed May 27, 2014, at http://www.horizon-systems.com/NHDPlus/NHDPlusV2_tools.php#NHDPlusV2.
- Luukkonen, C.L., Blumer, S.P., Weaver, T.L., and Jean, Julie, 2004, Simulation of the ground-water-flow system in the Kalamazoo County area, Michigan: U.S. Geological Survey Scientific Investigations Report 2004-5054, 65 p.
- McKay, Lucinda; Bondelid, Timothy; Dewald, Tommy; and others, 2012, NHDPlus version 2—User guide, 170 p., accessed August 28, 2014, at http://www.horizon-systems.com/NHDPlus/NHDPlusV2_documentation.php.
- 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 August 1, 2013, at <http://www.ocs.orst.edu/prism>.
- Seaber, P.R., Kapinos, F.P., and Knapp, G.L., 1987, Hydrologic unit maps (rev. 2007): U.S. Geological Survey Water-Supply Paper 2294, 63 p. (Also available at <http://pubs.usgs.gov/wsp/wsp2294/>.)
- Strahler, A.N., 1957, Quantitative analysis of watershed geomorphology: *Eos, Transactions American Geophysical Union*, v. 38, no. 6, p. 913–920.
- The MathWorks, Inc., 2010, MATLAB release 2010b: Natick, Mass., accessed August 28, 2014, at http://www.mathworks.com/products/new_products/release2010b.html.
- The MathWorks, Inc., 2012, Matlab R2012a release highlights: Natick, Mass., accessed August 28, 2014, at http://www.mathworks.com/products/new_products/release2012a.html.
- U.S. Department of Agriculture-Natural Resources Conservation Service, 2009, National Watershed Boundary Dataset, accessed August 28, 2014, at <http://www.nrcs.usda.gov/wps/portal/nrcs/main/national/water/watersheds/dataset/>.
- U.S. Department of Agriculture-Natural Resources Conservation Service, 2011, Web Soil Survey, accessed November 8, 2011, at <http://websoilsurvey.nrcs.usda.gov/>.
- U.S. Environmental Protection Agency, 2001, EPA/OW Permit Compliance System for BASINS Version 3 in CONUS, accessed August 28, 2014, at <http://water.epa.gov/scitech/datatit/models/basins/pcs3.cfm>.
- U.S. Environmental Protection Agency, 2014, Great Lakes, accessed May 27, 2014, at <http://www.epa.gov/greatlakes/>.
- U.S. Geological Survey, 2008, Hydrologic unit maps—What are hydrologic units?, accessed November 24, 2008, at <http://water.usgs.gov/GIS/huc.html>.
- U.S. Geological Survey, 2009a, National Hydrography Dataset, available at <http://nhd.usgs.gov/>.
- U.S. Geological Survey, 2009b, National Elevation Dataset, available at <http://ned.usgs.gov/>.
- U.S. Geological Survey, 2011, National Gap Analysis Program (GAP)—Land cover data and modeling, accessed May 2011, at <http://gapanalysis.usgs.gov/gaplandcover/data>.
- Wahl, K.L., Thomas, W.O., Jr., and Hirsch, R.M., 1995, Stream-gaging program of the U.S. Geological Survey: U.S. Geological Survey Circular 1123, accessed May 27, 2014, at <http://pubs.usgs.gov/circ/circ1123/index.html>.
- Waples, J.T.; Eadie, Brian; Klump, J.V.; Squires, Margaret; Cotner, James; and McKinley, Galen, 2008, Chapter 7—The Laurentian Great Lakes, in Hales, Burke; Cai, Wei-Jun; Mitchell, B.G., Sabine, C.L., and Schofield, Oscar, eds., *North American Continental Margins—A Synthesis and Planning Workshop—Report of the North American Continental Margins Working Group for the U.S. Carbon Cycle Scientific Steering Group and Interagency Working Group*: Washington, D.C., U.S. Carbon Cycle Science Program, p. 73–81, accessed May 27, 2014, at <http://www.glerl.noaa.gov/pubs/fulltext/2008/20080024.pdf>.

U.S. Geological Survey Water-Data Reports for Michigan and the St. Lawrence River Basin

Regional compilations (listed chronologically)

- Wells, J.V.B., 1958, Compilation of records of surface waters of the United States through September 1950—Part 4, St. Lawrence River Basin: U.S. Geological Survey Water-Supply Paper 1307, 397 p.
- Hendricks, E.L., 1964, Compilation of records of surface waters of the United States, October 1950 to September 1960—Part 4, St. Lawrence River Basin: U.S. Geological Survey Water-Supply Paper 1727, 379 p.
- U.S. Geological Survey, 1971, Surface water supply of the United States, 1961–65—Part 4, St. Lawrence River Basin—Volume 1, Basins of streams tributary to Lakes Superior, Michigan, and Huron: U.S. Geological Survey Water-Supply Paper 1911, 651 p.
- U.S. Geological Survey, 1974, Surface water supply of the United States, 1966–70—Part 4, St. Lawrence River Basin—Volume 1, Basins of streams tributary to Lakes Superior, Michigan and Huron: U.S. Geological Survey Water-Supply Paper 2111, 754 p.

Michigan data only (listed chronologically)

- U.S. Geological Survey, 1961–4, Surface water records of Michigan [annual data reports, water years 1961–4].
- U.S. Geological Survey, 1965–75, Water resources data for Michigan, Part 1, Surface water records [annual data reports, water years 1965–74].
- U.S. Geological Survey, 1976–82, Water resources data, Michigan, water years 1975–81: U.S. Geological Survey Water-Data Reports MI-75-1 through MI-81-1.
- Miller, J.B., Oberg, J.L., and Sieger, Theodore, Jr., 1983–85, Water resources data, Michigan, water years 1982–84: U.S. Geological Survey Water-Data Reports MI-82-1 through MI-84-1.
- Miller, J.B., Oberg, J.L., and Failing, J.C., 1986, Water resources data, Michigan, water year 1985: U.S. Geological Survey Water-Data Report MI-85-1, 297 p.

- Miller, J.B., Failing, J.C., and Larson, W.W., 1987, Water resources data, Michigan, water year 1986: U.S. Geological Survey Water-Data Report MI-86-1, 353 p.
- Blumer, S.P., Failing, J.C., Larson, W.W., Whited, C.R., and LeuVoy, R.L., 1988–90, Water resources data, Michigan, water years 1987–89: U.S. Geological Survey Water-Data Reports MI-87-1 through MI-89-1.
- Blumer, S.P., Larson, W.W., Minnerick, R.J., Whited, C.R., and LeuVoy, R.L., 1991–92, Water resources data, Michigan, water years 1990–91: U.S. Geological Survey Water-Data Reports MI-90-1 through MI-91-1.
- Blumer, S.P., Behrendt, T.E., Larson, W.W., Minnerick, R.J., LeuVoy, R.L., and Whited, C.R., 1993–94, Water resources data, Michigan, water years 1992–93: U.S. Geological Survey Water-Data Reports MI-92-1 through MI-93-1.
- Blumer, S.P., Behrendt, T.E., Ellis, J.M., Minnerick, R.J., LeuVoy, R.L., and Whited, C.R., 1995–2004, Water resources data, Michigan, water years 1994–2003: U.S. Geological Survey Water-Data Reports MI-94-1 through MI-03-1.
- Blumer, S.P., Behrendt, T.E., Whited, C.R., T.E., Ellis, J.M., Minnerick, R.J., and LeuVoy, R.L., 2005, Water resources data, Michigan, water year 2004: U.S. Geological Survey Water-Data Report MI-04-1, 496 p.
- Blumer, S.P., Whited, C.R., Ellis, J.M., Minnerick, R.J., and LeuVoy, R.L., 2006, Water resources data, Michigan, water year 2005: U.S. Geological Survey Water-Data Report MI-05-1, 572 p.
- U.S. Geological Survey, 2007–13, Water-resources data for the United States, water years 2006–12 [annual data compilations for stations 04034000 Bond Falls Reservoir near Paulding, Mich.; 04033500 Bond Falls Canal near Paulding, Mich.; 04057811 Greenwood Reservoir near Greenwood, Mich.; 04057814 Greenwood Release near Greenwood, Mich.; 04057813 Greenwood Diversion near Greenwood, Mich.; 04058000 Middle Branch Escanaba River near Ishpeming, Mich.; 04058190 Schweitzer Reservoir near Palmer, Mich.; and 04058400 Goose Lake Outlet near Sands Station, Mich.; accessed July 5, 2011, at <http://wdr.water.usgs.gov/>.

Table 1–1. Revisions to NHDPlus flowline value-added attribute (VAA) information.

[--, not changed]

Subregion	ComID	Original values					Modified values				
		FromNode	ToNode	Hydroseq	Divergence	Startflag	FromNode	ToNode	Hydroseq	Divergence	Startflag
0401	1798775	--	--	90072931	1	--	--	--	90072920	0	--
	1798903	90003017	90003052	90072920	2	0	9003052	9003017	90072931	0	1
	1771252	--	--	--	1	--	--	--	--	0	--
	1772034	--	--	--	2	0	--	--	--	0	1
0402	11951467	90021466	90021629	90046042	1	--	90021629	90021466	90046053	0	--
	11952189	90021629	90021470	90040541	--	--	90021470	90021629	90046185	--	--
	11951469	90021470	90021630	90036575	--	--	90021630	90021470	90046186	--	--
	11952191	90021630	90021472	90033483	--	--	90021472	90021630	90046187	--	--
	11951473	90021472	90021631	90031008	--	--	90021631	90021472	90046188	--	--
	11952193	90021631	90021628	90028979	--	0	90021628	90021631	90046189	--	1
	11951459	--	--	--	1	--	--	--	--	0	--
	11930450	--	--	--	1	0	--	--	--	0	1
	11930458	--	--	--	2	--	--	--	--	0	--
	11929798	--	--	--	1	--	--	--	--	2	--
	11929802	--	--	--	2	--	--	--	--	1	--
	11930458	--	--	--	2	--	--	--	--	0	--
	11930450	90019836	90023688	--	1	--	90023688	90019836	--	0	--
	12027444	--	--	--	--	0	--	--	--	--	1
0403	14444692	--	--	--	1	--	--	--	--	0	--
	6820848	--	--	--	1	0	--	--	--	0	1
	6849469	--	--	--	2	--	--	--	--	0	--
	11959514	--	--	--	--	0	--	--	--	--	1
	14444674	--	--	--	--	--	--	--	--	--	--
	9027663	--	--	--	2	0	--	--	--	0	1
	6809298	--	--	--	1	--	--	--	--	0	--
	14444312	--	--	--	--	0	--	--	--	--	1
	12006809	--	--	--	2	--	--	--	--	1	--
	12006813	--	--	--	1	--	--	--	--	2	--
	6865517	--	--	--	1	0	--	--	--	0	1
	6866125	--	--	--	2	--	--	--	--	0	--
	6821620	--	--	--	2	--	--	--	--	1	--
6821578	--	--	--	1	--	--	--	--	2	--	

Table 1–1. Revisions to NHDPlus flowline value-added attribute (VAA) information.—Continued

[--, not changed]

Subregion	ComID	Original values					Modified values				
		FromNode	ToNode	Hydroseq	Divergence	Startflag	FromNode	ToNode	Hydroseq	Divergence	Startflag
0403 (continued)	12174040	--	--	--	2	0	--	--	--	0	1
	12174028	--	--	--	1	--	--	--	--	0	--
	9036437	--	--	--	2	0	--	--	--	0	1
	9036449	--	--	--	1	--	--	--	--	0	--
	11959396	--	--	90098082	--	1	--	--	90017082	--	0
	11959398	--	--	--	--	0	--	--	--	--	1
	11959926	--	90021774	--	--	--	--	90082911	--	--	--
0404	3397118	--	--	--	2	--	--	--	--	0	--
	3396784	--	--	--	1	0	--	--	--	0	1
0405	12261958	--	--	--	1	--	--	--	--	0	--
	12260476	--	--	--	2	0	--	--	--	0	1
	13032275	90030086	90040059	90051985	2	--	90040059	90030086	90044497	0	--
	13032045	90040059	90040060	90044497	--	0	90040060	90040059	90051985	--	1
	12233796	--	--	--	1	--	--	--	--	0	--
	12260238	--	--	--	1	--	--	--	--	0	--
	12263234	--	--	--	2	0	--	--	--	0	1
	12258396	--	--	--	1	0	--	--	--	0	1
	12258498	--	--	--	2	--	--	--	--	0	--
	12263190	--	--	--	2	0	--	--	--	0	1
12260112	--	--	--	1	--	--	--	--	0	--	
0406	12222264	90029225	90029214	90067021	2	0	90029214	90029225	90067041	0	1
	12221900	--	--	--	1	--	--	--	--	0	--
	12222160	--	--	--	2	--	--	--	--	0	--
	12221662	--	--	--	1	--	--	--	--	0	--
0407	12502899	--	--	--	1	--	--	--	--	2	--
	12502901	--	--	--	0	--	--	--	--	1	--
	12962413	--	--	--	2	0	--	--	--	0	1
	12962487	--	--	--	1	--	--	--	--	0	--

Table 1–1. Revisions to NHDPlus flowline value-added attribute (VAA) information.—Continued

[--, not changed]

Subregion	ComID	Original values					Modified values				
		FromNode	ToNode	Hydroseq	Divergence	Startflag	FromNode	ToNode	Hydroseq	Divergence	Startflag
0408	12944858	--	--	--	2	--	--	--	--	0	--
	12944468	90036462	90036538		1	0	90036538	90036462		0	1
	12944860	--	--	--	1	0	--	--	--	0	1
	13046031	--	--	--	2	0	--	--	--	0	1
	13016527	--	--	90044601	--	--	--	--	90018387	--	--
	13016547	--	--	90052133	1	--	--	--	90018934	0	--
	13016565	90038914	90039104	90052132	2	--	90039104	90038914	90019518	0	--
	13019475	90039104	90039109	90044580	--	--	90039109	90039104	90020159	--	--
	13019493	90039109	90039117	90035767	--	--	90039117	90039109	90020830	--	--
	13019523	90039117	90039292	90032845	--	--	90039292	90039117	90021565	--	--
	13020129	90039292	90039294	90030478	--	--	90039294	90039292	90022385	--	--
	13020133	90039294	90039129	--	--	--	90039129	90039294	--	--	--
	13019559	90039129	90039298	90022385	--	--	90039298	90039129	90030478	--	--
	13020141	90039298	90039135	90021565	--	--	90039135	90039298	90032845	--	--
	13019577	90039135	90039138	90020830	--	--	90039138	90039135	90035500	--	--
	13019583	90039138	90039142	90020159	--	--	90039142	90039138	90035550	--	--
	13019591	90039142	90039300	90019518	--	--	90039300	90039142	90044601	--	--
	13020145	90039300	90039299	90018934	--	0	90039299	90039300	90052132	--	1
	13020143	--	--	90018387	--	0	--	--	90052133	--	1
	13019637	--	--	90052102	--	--	--	--	90064410	--	--
	13020157	--	--	90044561	--	--	--	--	90064405	--	--
	13019631	--	--	90039494	--	--	--	--	90064400	--	--
	13020155	--	--	90035756	--	--	--	--	90064395	--	--
	13019605	--	--	90032841	--	--	--	--	90064390	--	--
	13020147	--	--	90030477	--	--	--	--	90064385	--	--
	13007650	--	--	--	1	0	--	--	--	0	1
	13007796	--	--	--	1	0	--	--	--	0	1
	13007622	--	--	--	1	0	--	--	--	0	1
	13020539	--	--	--	1	0	--	--	--	0	1
	13031999	--	--	--	2	--	--	--	--	0	--
	13017263	--	--	--	1	--	--	--	--	0	--
	13028773	--	--	--	2	0	--	--	--	0	1

Table 1–1. Revisions to NHDPlus flowline value-added attribute (VAA) information.—Continued

[--, not changed]

Subregion	ComID	Original values					Modified values				
		FromNode	ToNode	Hydroseq	Divergence	Startflag	FromNode	ToNode	Hydroseq	Divergence	Startflag
0408 (continued)	13029155	--	--	--	2	0	--	--	--	0	1
	13028795	--	--	--	1	--	--	--	--	0	--
	13006714	--	--	--	1	--	--	--	--	0	--
	13009450	--	--	--	2	0	--	--	--	0	1
	13008924	--	--	--	1	--	--	--	--	0	--
0409	13192736	--	--	--	2	--	--	--	--	0	--
	13192794	--	--	--	2	--	--	--	--	0	--
	13192762	--	--	--	2	--	--	--	--	0	--
	13192474	--	--	--	2	0	--	--	--	0	1
	13194146	--	--	--	2	0	--	--	--	0	1
	13194090	--	--	--	1	--	--	--	--	0	--
	13193728	--	--	--	1	0	--	--	--	0	1
	13193748	--	--	--	2	--	--	--	--	0	--
0410	15612436	90060257	90060500		1	0	90060500	90060257		0	1
	15613610	--	--	--	--	0	--	--	--	--	1
	15644452	--	--	--	2	--	--	--	--	0	--
	15627609	--	--	--	1	0	--	--	--	0	1
	15627613	--	--	--	2	0	--	--	--	0	1
	15636243	--	--	--	2	0	--	--	--	0	1
	15627291	--	--	--	1	--	--	--	--	0	--
	15610538	--	--	--	2	0	--	--	--	0	1
	15605172	--	--	--	1	--	--	--	--	0	--

Table 1–2. Description of classes in the 2006 National Land Cover Database.

Class	Code	Description
Open water	11	areas of open water, generally with less than 25 percent cover of vegetation or soil
Perennial ice/snow	12	areas characterized by a perennial cover of ice and (or) snow, generally greater than 25 percent of total cover
Developed-open space	21	areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20 percent of total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes.
Developed-low intensity	22	areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 20 to 49 percent of total cover. These areas most commonly include single-family housing units.
Developed-medium intensity	23	areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 50 to 79 percent of the total cover. These areas most commonly include single-family housing units.
Developed-high intensity	24	highly developed areas where people reside or work in high numbers. Examples include apartment complexes, row houses, and commercial/industrial. Impervious surfaces account for 80 to 100 percent of the total cover.
Barren land	31	areas of bedrock, desert pavement, scarps, talus, slides, volcanic material, glacial debris, sand dunes, strip mines, gravel pits, and other accumulations of earthen material. Generally, vegetation accounts for less than 15 percent of total cover.
Deciduous forest	41	areas dominated by trees generally greater than 5 meters tall, and greater than 20 percent of total vegetation cover. More than 75 percent of the tree species shed foliage simultaneously in response to seasonal change.
Evergreen forest	42	areas dominated by trees generally greater than 5 meters tall, and greater than 20 percent of total vegetation cover. More than 75 percent of the tree species maintain their leaves all year. Canopy is never without green foliage.
Mixed forest	43	areas dominated by trees generally greater than 5 meters tall, and greater than 20 percent of total vegetation cover. Neither deciduous nor evergreen species are greater than 75 percent of total tree cover.
Shrub/scrub	52	areas dominated by shrubs less than 5 meters tall with shrub canopy typically greater than 20 percent of total vegetation. This class includes true shrubs, young trees in an early successional stage, or trees stunted from environmental conditions.
Grassland/herbaceous	71	areas dominated by graminoid or herbaceous vegetation, generally greater than 80 percent of total vegetation. These areas are not subject to intensive management such as tilling, but can be utilized for grazing.
Pasture/hay	81	areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay crops, typically on a perennial cycle. Pasture/hay vegetation accounts for greater than 20 percent of total vegetation.
Cultivated cropland	82	areas used for the production of annual crops, such as corn, soybeans, vegetables, tobacco, and cotton, and also perennial woody crops such as orchards and vineyards. Crop vegetation accounts for greater than 20 percent of total vegetation. This class also includes all land being actively tilled.
Woody wetlands	90	areas where forest or shrubland vegetation accounts for greater than 20 percent of vegetative cover and the soil or substrate is periodically saturated with or covered with water
Herbaceous wetlands	95	areas where perennial herbaceous vegetation accounts for greater than 80 percent of vegetative cover and the soil or substrate is periodically saturated with or covered with water

Table 1–3. Description of soil groups in the Soil Survey Geography (SSURGO) database.

Group	Description
A	Soils having a high infiltration rate (low runoff potential) when thoroughly wet. These consist mainly of deep, well drained to excessively drained sands or gravelly sands. These soils have a high rate of water transmission.
B	Soils having a moderate infiltration rate when thoroughly wet. These consist chiefly of moderately deep or deep, moderately well drained or well drained soils that have moderately fine texture to moderately coarse texture. These soils have a moderate rate of water transmission.
C	Soils having a slow infiltration rate when thoroughly wet. These consist chiefly of soils having a layer that impedes the downward movement of water or soils of moderately fine texture or fine texture. These soils have a slow rate of water transmission.
D	Soils having a very slow infiltration rate (high runoff potential) when thoroughly wet. These consist chiefly of clays that have a high shrink-swell potential, soils that have a high water table, soils that have a claypan or clay layer at or near the surface, and soils that are shallow over nearly impervious material. These soils have a very slow rate of water transmission.

Table 1–4. Description of variables from Aquatic Gap Analysis Program (GAP) used in the analysis.

Code	Group	Landform (texture)
Surficial geology		
1	A3	outwash (coarse)
2	B3	ice-contact (coarse)
3	C1	end-moraine (fine)
4	C2	end-moraine (medium)
5	C3	end-moraine (coarse)
6	D1	ground-moraine (fine)
7	D2	ground-moraine (medium)
8	D3	ground-moraine (coarse)
Bedrock geology		
		Coarse-grained stratified sediment
		Till

Publication services provided by the U.S. Geological Survey
Science Publishing Network
Columbus Publishing Service Center

For more information concerning the research in this report
contact the

Director, Michigan Water Science Center

U.S. Geological Survey

6520 Mercantile Way, Suite 5

Lansing, MI 48911-5991

<http://mi.water.usgs.gov/>

