Prepared in cooperation with the Pennsylvania Department of Conservation and Natural Resources, Bureau of Topographic and Geologic Survey

# Regression Method for Estimating Long-Term Mean Annual Ground-Water Recharge Rates from Base Flow in Pennsylvania

Scientific Investigations Report 2008-5185

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

# Regression Method for Estimating Long-Term Mean Annual Ground-Water Recharge Rates from Base Flow in Pennsylvania

By Dennis W. Risser, Ronald E. Thompson, and Marla H. Stuckey

Prepared in cooperation with the Pennsylvania Department of Conservation and Natural Resources, Bureau of Topographic and Geologic Survey

Scientific Investigations Report 2008-5185

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

**U.S. Department of the Interior**
DIRK KEMPTHORNE, Secretary

**U.S. Geological Survey**
Mark D. Myers, Director

U.S. Geological Survey, Reston, Virginia: 2008

For product and ordering information:
World Wide Web: http://www.usgs.gov/pubprod
Telephone: 1-888-ASK-USGS

For more information on the USGS--the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment:
World Wide Web: http://www.usgs.gov
Telephone: 1-888-ASK-USGS



Suggested citation:
Risser, D.W., Thompson, R.E., and Stuckey, M.H., 2008, Regression method for estimating long-term mean annual ground-water recharge rates from base flow in Pennsylvania: U.S. Geological Survey Scientific Investigations Report 2008-5185, 23 p.

# Contents

## Figures



## Tables

# Conversion Factors

| Multiply | By | To obtain |
|---|---|---|
| | Length | |
| inch (in.) | 2.54 | centimeter (cm) |
| inch (in.) | 25.4 | millimeter (mm) |
| foot (ft) | 0.3048 | meter (m) |
| mile (mi) | 1.609 | kilometer (km) |
| | Area | |
| square mile ($mi^2$) | 259.0 | hectare (ha) |
| square mile ($mi^2$) | 2.590 | square kilometer ($km^2$) |
| | Flow rate | |
| cubic foot per second ($ft^3/s$) | 0.02832 | cubic meter per second ($m^3/s$) |
| inch per year (in/yr) | 25.4 | millimeter per year (mm/yr) |
| | Hydraulic gradient | |
| foot per mile (ft/mi) | 0.1894 | meter per kilometer (m/km) |

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

°C=(°F-32)/1.8

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

# Regression Method for Estimating Long-Term Mean Annual Ground-Water Recharge Rates from Base Flow in Pennsylvania

By Dennis W. Risser, Ronald E. Thompson, and Marla H. Stuckey

## Abstract

A method was developed for making estimates of long-term, mean annual ground-water recharge from streamflow data at 80 streamflow-gaging stations in Pennsylvania. The method relates mean annual base-flow yield derived from the streamflow data (as a proxy for recharge) to the climatic, geologic, hydrologic, and physiographic characteristics of the basins (basin characteristics) by use of a regression equation.

Base-flow yield is the base flow of a stream divided by the drainage area of the basin, expressed in inches of water basinwide. Mean annual base-flow yield was computed for the period of available streamflow record at continuous streamflow-gaging stations by use of the computer program PART, which separates base flow from direct runoff on the streamflow hydrograph. Base flow can provide a reasonable estimate of recharge for basins where losses and interbasin transfers of ground water are minimal.

Twenty-eight basin characteristics were included in the exploratory regression analysis as possible predictors of base-flow yield. Basin characteristics found to be statistically significant predictors of mean annual base-flow yield during 1971–2000 at the 95-percent confidence level were (1) mean annual precipitation, (2) average maximum daily temperature, (3) percentage of sand in the soil, (4) percentage of carbonate bedrock in the basin, and (5) stream channel slope. The equation for predicting recharge was developed using ordinary least-squares regression. The standard error of prediction for the equation on log-transformed data was 9.7 percent, and the coefficient of determination was 0.80.

The equation can be used to predict long-term, mean annual recharge rates for ungaged basins, providing that the explanatory basin characteristics can be determined and that the underlying assumption is accepted that base-flow yield derived from PART is a reasonable estimate of ground-water recharge rates. For example, application of the equation for 370 hydrologic units in Pennsylvania predicted a range of ground-water recharge from about 6.0 to 22 inches per year. A map of the predicted recharge illustrates the general magnitude and variability of recharge throughout Pennsylvania.

## Introduction

Ground-water recharge is a major component of the water budget, but because it is difficult to measure directly, the magnitude and variability of recharge in Pennsylvania is not well known. The droughts of 1998–2002 heightened public awareness about the sustainable use of ground water and increased interest in quantifying water budgets and in obtaining better estimates of ground-water recharge rates. In 2002, the Commonwealth of Pennsylvania enacted legislation directing the Pennsylvania Department of Environmental Protection (PaDEP) to update the State Water Plan and to include an assessment of "prime recharge areas" and "recharge capacity" as part of an inventory of ground-water resources.

Many approaches for estimating ground-water recharge rates are available (Scanlon and others, 2002), but most methods provide results at a local scale of 10s to 1,000s of square feet (Delin and Risser, 2007, table 1). To map recharge rates on a statewide scale, one widely used approach is to estimate recharge from available streamflow records at gaged basins and develop a regression equation relating those recharge estimates to physical and climatic characteristics of the gaged basins. Provided the basin characteristics are known, regression equations can then be used to make estimates of recharge for basins across the state. This approach has been used to estimate recharge in other states (Holtschlag, 1997; Flynn and Tasker, 2004; Delin and others, 2007; Gebert and others, 2007). The availability of streamflow-gaging stations with long-term streamflow records throughout Pennsylvania made the regression approach used in this study feasible.

### Base-Flow Yield and Recharge

In this study, long-term average base-flow yield was used as a proxy for the long-term average ground-water recharge rate. Base-flow yield is the base flow of a stream divided by the drainage area of the basin upstream from the location where base flow is determined. Base flow is the part of streamflow not caused by direct runoff from precipitation or melting snow and usually is attributed to ground-water discharge (U.S. Geological Survey, 1989). Mean annual base-flow yield for the period of record is assumed to equal the mean annual recharge rate. For the assumption to be reasonable, base flow needs to be a good representation of ground-water discharge to the stream, and additions (other than from recharge) or losses of ground water from the basin need to be minimal.

Mean annual base-flow yield was computed for the period of available continuous streamflow record and was expressed in inches of water over the area of the basin upstream from each streamflow-gaging station. Converting mean annual base flow from units of cubic feet per second to inches basinwide normalizes the value by its basin area and makes it easier to visualize as an estimate of recharge rate.

Separation of the streamflow hydrograph is subjective and can be determined by different methods (for example, Pettyjohn and Henning, 1979; Rutledge, 1993; Arnold and others, 1995; Wahl and Wahl, 1988) causing uncertainty about the base-flow value. However, numerous studies have used base flow as a proxy for long-term average recharge because other methods for estimating regional recharge rates also have inherent difficulties (Scanlon and others, 2002).

### Previous Investigations

Previous studies have used base flow from streamflow-gaging stations to estimate ground-water recharge. These include studies in North Carolina (Heath, 1994; Daniel, 1996), Michigan (Holtschlag, 1997), the Great Lakes Basin (Neff and others, 2006), Wisconsin (Gebert and others, 2007), and the entire United States (Wolock, 2003). Regression equations have been used to estimate recharge for ungaged basins by Holtschlag (1997), Flynn and Tasker (2004), Delin and others (2007), and Gebert and others (2007).

In Pennsylvania, base flow for various recurrence intervals was determined by White and Sloto (1990), but mean annual base flow was not included. Risser and others (2005) estimated mean annual base flow in Pennsylvania by use of the PART computer program (Rutledge, 1993) and recharge by use of the RORA computer program (Rutledge, 1998) at 197 gaged basins but did not develop equations for making estimates at ungaged basins. Regression was used by Stuckey (2006) to develop equations for predicting the 10-, 25-, and 50-year recurrence-interval base flows for Pennsylvania streams, but mean annual base flow was not included in the analysis. In Monroe County, Pa., a regression equation to predict mean annual base flow in cubic feet per second from basin characteristics was developed by Thompson and Hoffman (2006).

### Purpose and Scope

This report presents the results of a study by the U.S. Geological Survey (USGS) in cooperation with the Pennsylvania Department of Conservation and Natural Resources, Bureau of Topographic and Geologic Survey, to develop a method for estimating ground-water recharge rates for basins in Pennsylvania. The method assumes that base-flow yield, determined from streamflow at 80 continuous-record streamflow-gaging stations, can be used as a proxy for ground-water recharge. The report describes the development of a regression equation relating base-flow yield to the significant climatic, geologic, hydrologic, and physiographic characteristics of a basin. To illustrate the use of the equation, mean annual recharge for 1971–2000 was estimated for 370 hydrologic units in Pennsylvania. The results of the equation were mapped to show the spatial variability of ground-water recharge rates across Pennsylvania.

## Regression Method for Estimating Recharge Rates

The method for estimating long-term, mean annual ground-water recharge involved three steps: (1) values of mean annual base-flow yield were determined from streamflow records at gaging stations with flow relatively unaffected by regulation, diversion, or mining; (2) basin characteristics hypothesized to affect base-flow yield, such as bedrock type, land cover, and precipitation, were determined for the drainage basin upstream from the gaging stations; and (3) the base-flow yield (dependent variable) was then related to the basin characteristics (independent or explanatory variables) using ordinary least-squares regression techniques to obtain a predictive equation for recharge. This equation can be used to compute estimates of recharge rates for basins where gaging-station data are not available. The values of base-flow yield determined from streamflow-gaging-station records are referred to as “observed” in this report, and those computed from regression equations are referred to as “predicted.”

### Determining Base-Flow Yield for Basins

Base flow was determined from streamflow records at gaging stations in Pennsylvania by use of the computer program PART, from a previous study by Risser and others (2005). The computer program PART was selected for this study because it had been widely used to compute base flow in the eastern United States (Holtschlag, 1997; Nelms and others, 1997; Neff and others, 2006) and the software is supported by the USGS.

The PART program computes base flow from the streamflow hydrograph by first identifying days of negligible surface runoff and assigning base flow equal to streamflow on those days; the program then interpolates between those days. PART locates periods of negligible surface runoff after a storm by identifying the days meeting a requirement of antecedent-recession length and rate of recession. It uses linear interpolation between the log values of base flow to connect across periods that do not meet those tests. A detailed description of the algorithm used by PART is described in Rutledge (1998, p. 33-38). An example illustrating the separation of the base-flow component from a streamflow hydrograph is shown in figure 1.

Values of mean annual base-flow yield were derived from 183 active and inactive continuous-record streamflow-gaging stations in Pennsylvania. The stations were selected from 197 stations in Risser and others (2005) that had at least 10 years of recorded streamflow from watersheds of less than 550 $mi^2$ and were relatively unaffected by upstream regulation from reservoirs, withdrawals, wastewater return flow, and mining. The extent that those and other human activities affected the streamflow were determined from USGS annual data reports for Pennsylvania (*http://pa.water.usgs.gov/ar/index.html*) and from reports by Ehlke and Reed (1999) and Stuckey and Reed (2000). The mean period of record for the 183 streamflow-gaging stations was 27 years, and the mean drainage-basin area was 157 $mi^2$.

Mean annual base-flow yield was computed in Risser and others (2005) for the period of available streamflow records collected through 2001. For this study, base-flow yield also was computed for the period 1971–2000 so that the base-flow record would be on a time period coincident with 1971–2000 climatic normal data sets for temperature and precipitation.

## Basin Characteristics

A list of 28 climatic, geologic, hydrologic, and physiographic basin characteristics that might possibly affect base-flow yield was developed from a variety of geographic information system (GIS) sources (table 1). Twenty-three of the basin characteristics were used directly from the work of Stuckey (2006) and five additional basin characteristics (percentage of sandstone bedrock, percentage of shale bedrock, percentage of crystalline bedrock, percentage of sand in the soil, and average maximum daily temperature) were added for this study because they were postulated as possibly affecting base-flow yield or they had been shown to be statistically significant variables in other recharge studies.

## Regression Techniques

Mean annual base-flow yield (expressed in inches over the basin area) at the 183 candidate streamflow-gaging stations was related to basin characteristics using ordinary least-squares regression techniques. The regression analysis was conducted using the statistical software package S-PLUS (MathSoft, Inc., 1997) to determine the explanatory variables significant at the 95-percent confidence level.

Exploratory regression analysis was conducted by initially examining the relation between mean annual base-flow yield for the period of record from 183 continuous-record

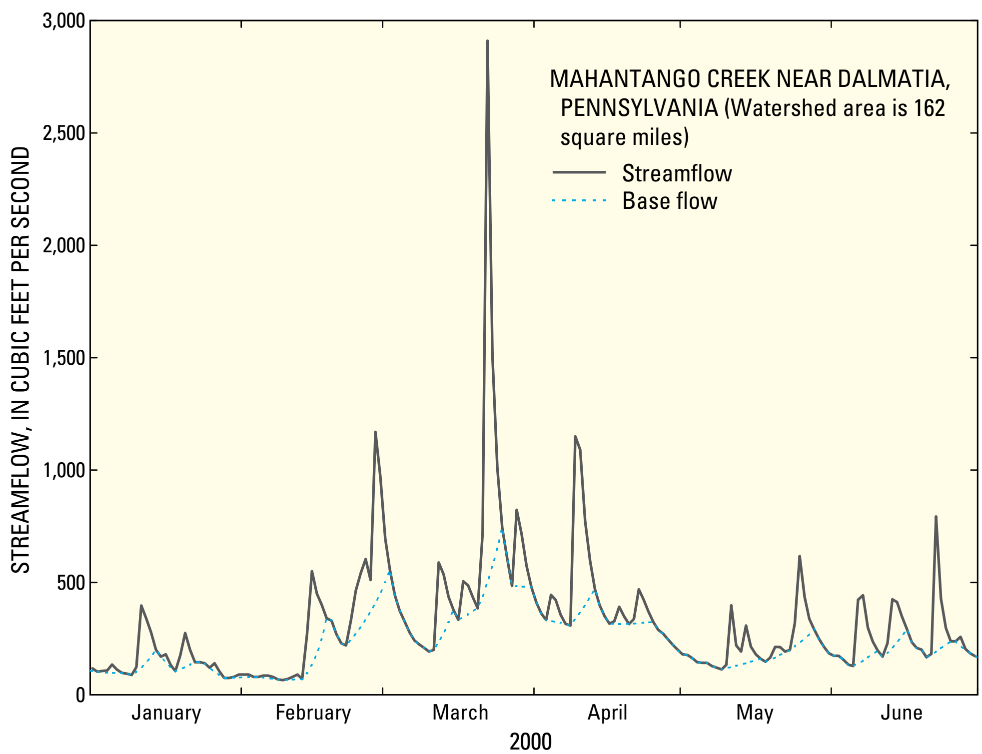


**Figure 1.** Streamflow hydrograph showing separation of the base-flow component with the PART computer program for a basin in Pennsylvania (From Risser and others, 2005, fig. 2).

**Table 1.** Basin characteristics selected for use in the development of regression equations for recharge in Pennsylvania.

| Regression variable (alternate variable name) | Definition | Source | Reference |
|---|---|---|---|
| **Dependent Variables [used as proxy for ground-water recharge]** | | | |
| UBC195 (BFY) | Mean annual base-flow yield for years of available streamflow record during 1971–2000, concurrent with the Parameter-elevation Regressions on Independent Slopes Model (PRISM) precipitation record (expressed as inches of water over basin area) | Derived by application of PART base-flow separation computer program on daily values of streamflow recorded at 80 U.S. Geological Survey gaging stations | Risser and others (2005) *http://pa.water.usgs.gov/recharge/* |
| UBC200 | Mean annual base-flow yield for period of streamflow record (expressed as inches of water over basin area) | Derived by application of PART base-flow separation computer program on daily values of streamflow recorded at 197 U.S. Geological Survey gaging stations | Risser and others (2005) *http://pa.water.usgs.gov/recharge/* |
| **Independent Variables [basin characteristics determined from geographic information system (GIS) data sets]** | | | |
| bslope | Basin slope (in degrees) | Digital Elevation Model (DEM) | U.S. Geological Survey (2000a) |
| elev | Mean basin elevation (feet) | Digital Elevation Model (DEM) | U.S. Geological Survey (2000a) |
| forest | Forested (percentage of basin area) | National Land Cover Dataset enhanced version (NLCDe) | Homer and others (2004) |
| glacer | Glaciated (percentage of basin area) | Pennsylvania Department of Conservation and Natural Resources Glacial data set 1:100,000 scale | Pennsylvania Department of Conservation and Natural Resources (1995) |
| lake | Lakes (percentage of basin area) | Digitized from USGS 1:24,000 topographic quadrangle maps | U.S. Geological Survey (*http://topomaps.usgs.gov/*) |
| length | Stream density (length in miles per basin area) | National Hydrography Dataset (NHD), 1:24,000 scale | U.S. Geological Survey (2000b) |
| precip (Ppt) | Mean annual precipitation, 1971–2000 (inches) | Parameter-elevation Regressions on Independent Slopes Model (PRISM) | Daly (1996) |
| slope (Sl) | Channel slope (average from 10th to 85th percent of stream length measured from headwaters) (feet per mile) | Digital Elevation Model (DEM) | U.S. Geological Survey (2000a) |
| soilin | Soil infiltration index (unitless, 1=well to 4=poor) | State Soil Geographic (STATSGO) database | U.S. Department of Agriculture (1994) |
| vallgh | Longest drainage path (mile) | National Hydrography Dataset (NHD), 1:24,000 scale | U.S. Geological Survey (2000b) |
| ubc024 | Sinkhole density (number of sinkholes per square mile) | Sinkhole inventory and online database | Pennsylvania Department of Conservation and Natural Resources (2008) |
| ubc025 | Depth to bedrock (inches) | State Soil Geographic (STATSGO) database | U.S. Department of Agriculture (1994) |
| ubc026 | Drainage run-off curve (unitless, 1=well to 7=poor) | State Soil Geographic (STATSGO) database | U.S. Department of Agriculture (1994) |
| ubc027 | Soil available water content (percentage) | State Soil Geographic (STATSGO) database | U.S. Department of Agriculture (1994) |
| ubc028 | Soil permeability (inches per hour) | State Soil Geographic (STATSGO) database | U.S. Department of Agriculture (1994) |

**Table 1.** Basin characteristics selected for use in the development of regression equations for recharge in Pennsylvania.—Continued

| Regression variable (alternate variable name) | Definition | Source | Reference |
|---|---|---|---|
| ubc029 | Urbanized area (percentage of basin area) | National Land Cover Dataset enhanced version (NLCDe) | Homer and others (2004) |
| ubc030 | Residential area (percentage of basin area) | National Land Cover Dataset enhanced version (NLCDe) | Homer and others (2004) |
| ubc031 | Mined area (percentage of basin area) | National Land Cover Dataset enhanced version (NLCDe) | Homer and others (2004) |
| ubc038 | Commercial and industrial area (percentage of basin area) | National Land Cover Dataset enhanced version (NLCDe) | Homer and others (2004) |
| ubc039 | Wetlands (percentage of basin area) | National Land Cover Dataset enhanced version (NLCDe) | Homer and others (2004) |
| ubc066 | Ground-water head, defined as mean basin elevation minus minimum elevation (feet) | Digital Elevation Model (DEM) | U.S. Geological Survey (2000a) |
| ubc067 | Shape factor (unitless 1 to ~25) | Digital Elevation Model (DEM) | U.S. Geological Survey (2000a) |
| ubc069 (C) | Carbonate bedrock (percentage of basin area) | Bedrock geology, with formations aggregated by U.S. Geological Survey for this study | Miles and others (2001) |
| UBC166 | Sandstone bedrock (percentage of basin area) | Bedrock geology, with formations aggregated by U.S. Geological Survey for this study | Miles and others (2001) |
| UBC167 | Shale bedrock (percentage of basin area) | Bedrock geology, with formations aggregated by U.S. Geological Survey for this study | Miles and others (2001) |
| UBC169 | Crystalline bedrock (percentage of basin area) | Bedrock geology, with formations aggregated by U.S. Geological Survey for this study | Miles and others (2001) |
| UBC170 (T) | Average daily maximum temperature (degrees Fahrenheit) | Parameter-elevation Regressions on Independent Slopes Model (PRISM) | Daly (1996) |
| UBC182 (S) | Percentage of sand in soil | State Soil Geographic (STATSGO) database | U.S. Department of Agriculture (1994) |

stations and 28 basin characteristics. Exploratory regression analysis indicated that mean annual precipitation was the most significant basin characteristic in the equation on the basis of the *p*-value and the sensitivity analysis for predicting base-flow yield. Because base-flow yield was computed for various periods during 1884–2000, the variability in precipitation for differing periods of record could introduce a potential source of error into the equation. For example, mean annual precipitation at 12 long-term meteorological stations across Pennsylvania was about 6 in. greater during the 1990s than during the 1960s. In the equation, a 6-in. difference in precipitation could cause predicted base-flow yield (or recharge) to differ by more than 3 in. between the two periods.

Because of the potential source of error introduced by climatic variability, base-flow-yield data were used only from the period 1971–2000, which is the normal climatic period used in the regression for the basin characteristics *mean annual precipitation* and *average daily maximum temperature* (table 1). Of the 183 streamflow-gaging stations, 90 stations having at least 15 years of record available during 1971–2000 were selected for further analysis to lessen the potential error caused by climatic variability.

Of the 90 stations, 10 stations were eliminated during analysis of diagnostic statistics. Regression diagnostics, including colinearity, leverage, Cooks D, and DFFITS (Helsel and Hirsch, 1993), in conjunction with location and period of record were used to further decrease the number of streamflow-gaging stations used in the analysis. Eighty of the 90 stations used in the analysis had values below the critical values for leverage and Cooks D and were near or less than the critical value for DFFITS. These 80 streamflow-gaging stations were retained and used to develop the final regression equation. The stations are listed in table 3 (back of report) and locations are shown in figure 2.

In an attempt to improve the power of the regression equation to predict base-flow yield, Pennsylvania was divided into four different regions, and separate equations were computed for each region. However, because regionalization did not substantially improve the residuals between observed and predicted mean annual base-flow yield, a single statewide equation for base-flow yield was developed. The variables selected for the final regression equation were all statistically significant well below the 95-percent confidence level (*p*-values all less than 0.000003). The following five basin characteristics were significant for predicting mean annual base-flow yield: mean annual precipitation, average daily maximum temperature, percentage of sand in the soil, percentage of carbonate bedrock within the basin, and channel slope. Values of the basin characteristics for each of the 80 stations used in the equation are listed in table 4 (back of report). The distributions of mean annual precipitation, average daily maximum temperature, percentage of sand in the soil, and percentage of carbonate bedrock within the basin are shown in figures 3 and 4.

All variables were log-transformed before regression analysis to form a near-linear relation between the base-flow yield and the basin characteristics. Exploratory regression using non-transformed variables showed a non-constant variance in the residuals, indicating a need to log-transform the variables. Because percentages of sand in the soil and of carbonate rock in a basin can have values of zero, 1.0 was added to the decimal form of the percentages before transformation. The regression model took the form in log units:

$$LogBFY = 2.4343 + 1.7218 logPpt - 2.4602 logT + 1.2297 log(1 + 0.01S) + 0.5673 log(1 + 0.01C) + 0.0828 log(Sl) \tag{1}$$

Or, in arithmetic space:

$$BFY = 10^{2.4343} (Ppt)^{1.7218} (T)^{-2.4602} (1 + 0.01S)^{1.2297} (1 + 0.01C)^{0.5673} (Sl)^{0.0828} \tag{2}$$

where:

| | |
|---|---|
| *Log* | is log to base 10; |
| *BFY* | is the predicted mean annual base-flow yield during 1971–2000, in inches per year over the drainage basin; |
| *Ppt* | is mean annual precipitation during 1971–2000, in inches; |
| *T* | is average daily maximum temperature during 1971–2000, in degrees Fahrenheit; |
| *S* | is percentage of sand in the soil, by weight; |
| *C* | is percentage of basin underlain by carbonate bedrock, in percent; and |
| *Sl* | is the channel slope, in feet per mile. |

Observed and predicted mean annual base-flow yields are listed along with residuals in table 5 (back of report). Predicted mean annual base-flow yield was generally within 2.5 in. of the observed value (fig. 5). The coefficient of determination ($r^2$) for the equation was 0.80, and the standard error of prediction was 0.043 log units (9.7 percent). Residuals (observed minus predicted base-flow yield) ranged from -4.39 to 4.18 in., and the average absolute percent difference of the residuals was 8 percent. These statistics are measures of unexplained variation between the observed and predicted values.

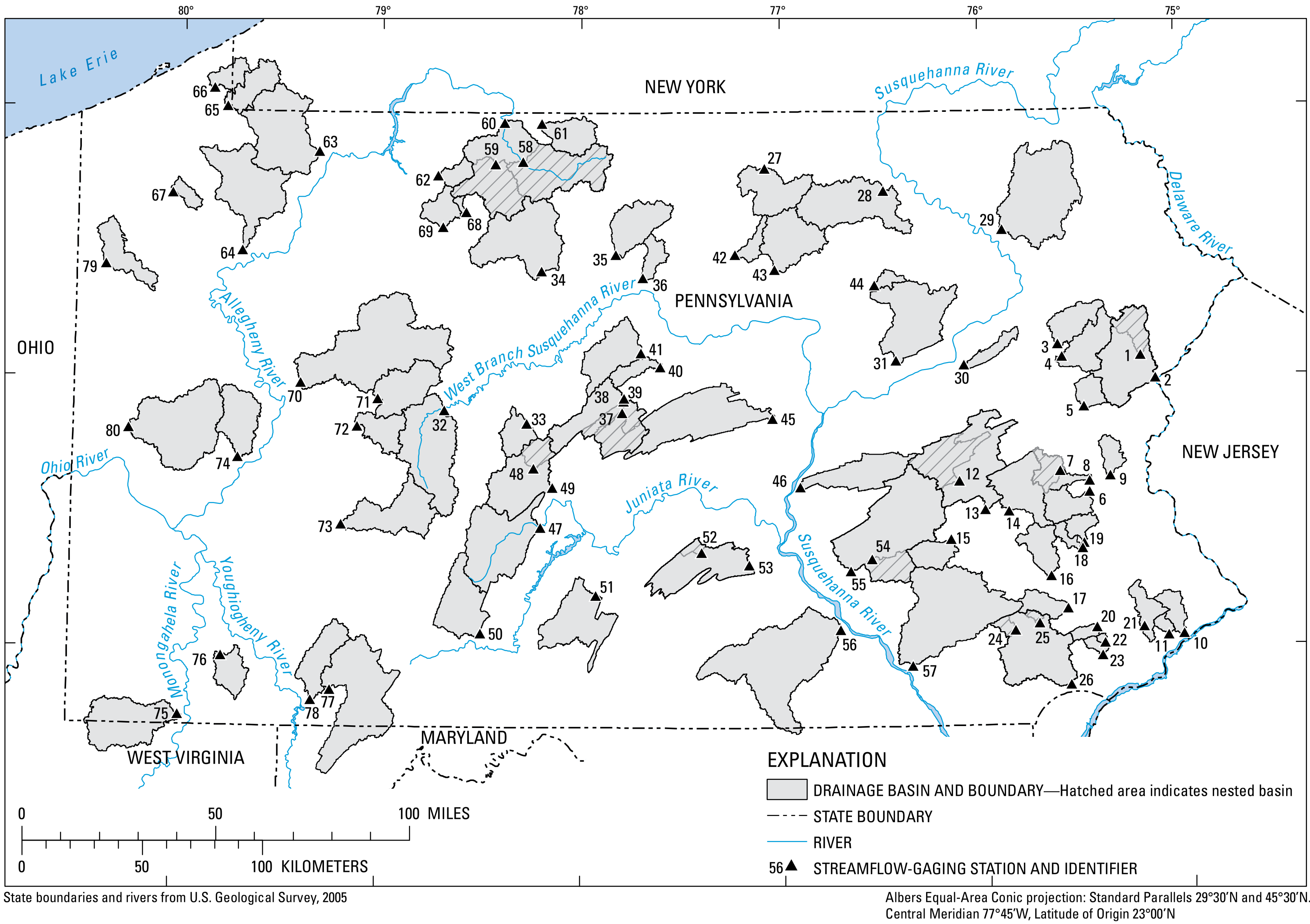


**Figure 2.** Locations of 80 streamflow-gaging stations and basins used to develop regression equations for estimating ground-water recharge in Pennsylvania. (Identifiers and descriptions of streamflow-gaging stations are presented in table 3.)

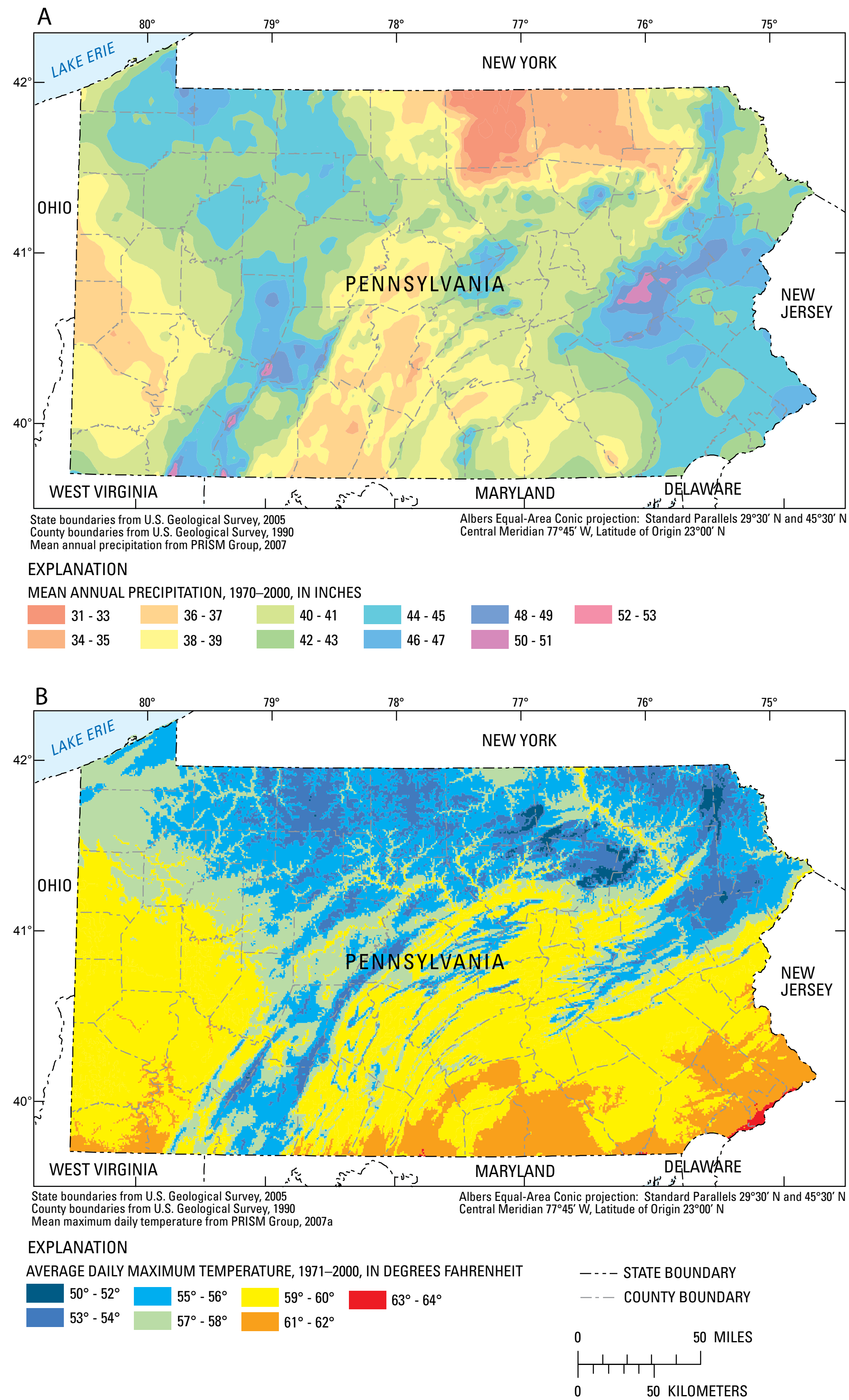


**Figure 3.** Climatic variables that were significant basin characteristics in the regression equation for predicting mean annual base-flow yield in Pennsylvania – (*A*) mean annual precipitation and (*B*) average daily maximum temperature.

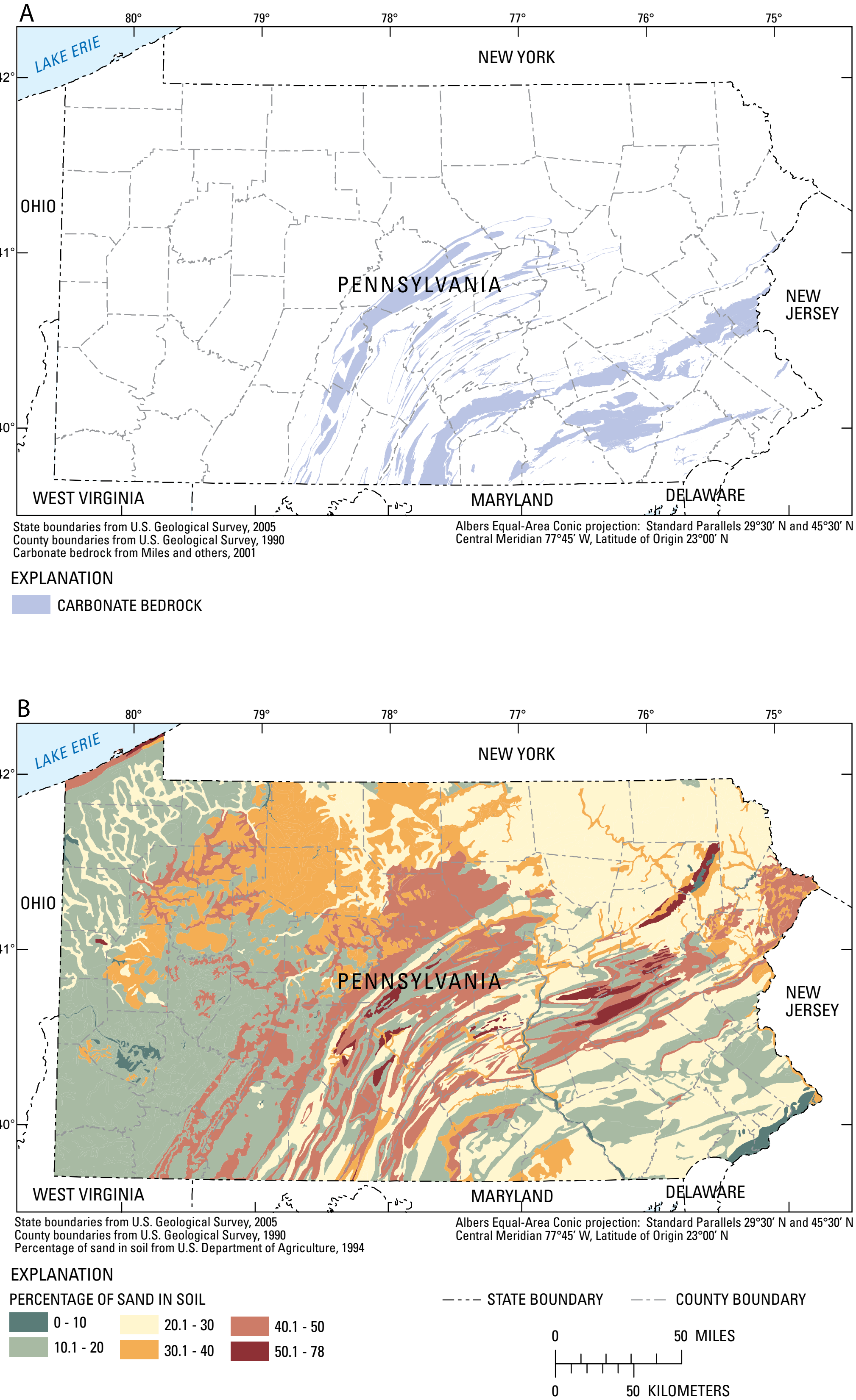


**Figure 4.** Geologic and soil variables that were significant basin characteristics in the regression equation for predicting mean annual base-flow yield in Pennsylvania – (*A*) locations of carbonate bedrock and (*B*) percentage of sand in the soil.

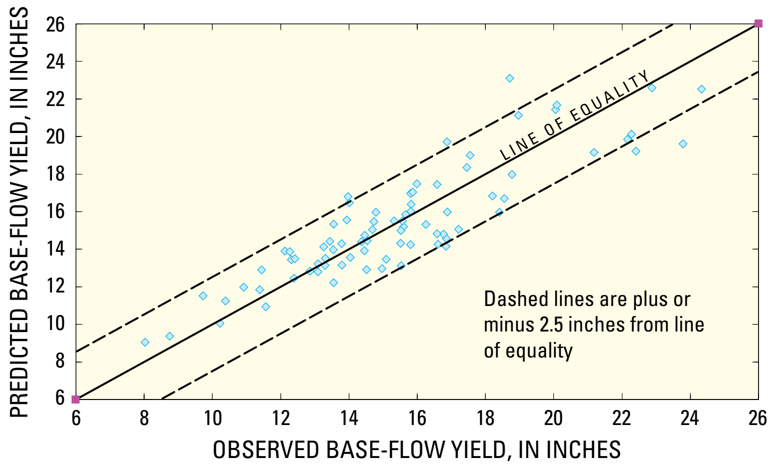


**Figure 5.** Relation between observed mean annual base-flow yield from streamflow records and predicted mean annual base-flow yield from the regression equation.

A measure of error associated with a predicted value of base-flow yield is the prediction interval, which gives the range of values of predicted base-flow yield for a given confidence level. For a 90-percent prediction interval, the probability that a predicted base-flow value falls within the prediction interval is 90 percent. The 90-percent prediction interval (*T*) for the equation on the log-transformed values is 0.0746. To compute the 90-percent prediction interval for a predicted base-flow yield, the following equation can be used:

$$10^{|qo-T|} \leq BFY \leq 10^{qo+T} \quad (3)$$

Where:

*qo* is the predicted base-flow yield at a site expressed as inches over the basin, in log units;

*T* is the 90-percent prediction interval; and

*BFY* is the predicted base-flow yield at a site expressed in inches over the basin.

## Sensitivity of Predictions

The sensitivity of base-flow yield values predicted from the equation was examined by changing each of the five independent variables individually by its full range in the 80 basins used to develop the equation; the remaining four variables in the equation were assigned the mean value from table 4. On the basis of the sensitivity analysis, a change in mean annual precipitation of 14.6 in. (the range of precipitation in the 80 basins) caused the greatest change in predicted base-flow yield (8.84 in.). Base-flow yield increased with an increase of each of the explanatory variables except average daily maximum temperature, which produced a decrease (table 2).

The most influential basin characteristics in the equation were climatic variables (table 2). Predicted base-flow yield was positively related to mean annual precipitation and negatively related to average daily maximum temperature. An increase in precipitation on a basin logically produces more ground-water recharge and base flow. Temperature affects ground-water recharge and base flow through its association with growing season and evapotranspiration–warmer temperatures allow a longer growing season and increased evapotranspiration, which decreases the amount of water available for recharge. Percentage of sand in the soil and percentage of carbonate rock in the basin are positively related to base-flow yield. This relation is probably caused by the relatively large vertical permeability of sandy soils and sinkholes and closed depressions developed on carbonate bedrock–both factors that promote increased infiltration within a basin. The positive relation between channel slope and predicted base-flow yield is difficult to explain. One might expect more direct runoff (and less recharge) on basins with steep slopes. Basins having steep slopes are most likely correlated with other variables not represented in the equation that cause greater base-flow yield relative to other areas. Possibilities include the occurrence of permeable colluvium on slopes, greater accumulation of snowpack on steep slopes in mountainous areas, or variability in precipitation amount or land cover not included in the generalized data sets used in the regression analysis for those explanatory variables.

**Table 2.** Change in predicted base-flow yield for an increase in the explanatory variable over the range of its values used to develop the regression equation for estimating ground-water recharge rates in Pennsylvania.

| Independent explanatory variable | Range of explanatory variable | Change in predicted base-flow yield, in inches | Change in predicted base-flow yield, in percent |
|---|---|---|---|
| Average daily maximum temperature (degrees Fahrenheit) | 9.0 degrees Fahrenheit | -6.08 | -32 |
| Mean annual precipitation (inches) | 14.6 inches | 8.84 | 85 |
| Percentage of sand in the soil | 27 percent | 4.09 | 29 |
| Channel slope (foot per mile) | 108 feet per mile | 4.62 | 36 |
| Carbonate bedrock (percentage of basin area) | 83 percent of basin area | 6.02 | 41 |

## Limitations of Regression Equations

The regression equation for estimating recharge was developed from base-flow yields at streams in Pennsylvania; thus, results are limited to basins in Pennsylvania. The largest source of uncertainty in the estimates of recharge is most likely from the determination of base flow and from the assumption that base flow is a good approximation of ground-water recharge. The observed base-flow yield for the continuous-record stations was computed from the computer program PART; the values would be different if another base-flow separation method was used. For example, estimates of mean annual base-flow yield of East Mahantango Creek near Dalmatia (site 46 in tables 3–5), determined by four different hydrograph-separation routines, ranged from 10.8 to 12.9 in/yr for 1968–2001 (Risser and others, 2005, table 5). Neff and others (2006) compared base-flow yields from six different separation methods for streams in the Great Lakes Basin and showed an average difference of 41 percent in base-flow index (ratio of base flow to streamflow) among the methods.

Base flow of streams in Pennsylvania can be affected by factors other than ground-water discharge. Slow release of water from wetlands and from snowmelt to streams is sometimes included as base flow by the hydrograph-separation routines, which might cause ground-water recharge to be overestimated in the northern and glaciated areas of Pennsylvania. Although an attempt was made to select streamflow-gaging stations in basins without major reservoirs or discharges, some of those impacts may be incorporated into the base-flow estimates.

The determinations of basin characteristics used in the regression analyses are limited by the accuracy of the information in the analyzed data sets and the GIS-based automated procedures used to process the data sets. The results of the regression analysis explain as much variance in the data as possible at the $p < 0.05$ level of significance but should not be thought of as providing cause and effect relations. The standard error of prediction of 9.7 percent for the regression of log-normalized base-flow-yield data may not be acceptable for some purposes.

The equation is useful for illustrating regional variability of ground-water recharge rates across Pennsylvania but should not be used for estimating site-specific values. Within a basin, recharge rates could vary considerably from the average rate estimated by the equation because the characteristics of the basin are not homogeneous throughout. The regression equation was derived from base-flow yields determined at drainage basins ranging from about 5 to 550 $mi^2$ and from five explanatory variables with ranges given in table 4. Caution is advised in estimating recharge for basins with properties that fall substantially outside of those ranges.

## Example Calculation

Calculate the base-flow yield (recharge) and 90-percent prediction interval for a site at the confluence of Mountain Creek with the Yellow Breeches Creek in Cumberland County in south-central Pennsylvania at latitude 40.1448° N and longitude 77.1775° W. Mean annual precipitation in the basin is 41.3 in., average daily maximum temperature is 59.6°F, percentage of sand in soil is 31.1 percent, percentage of carbonate bedrock is 16.3 percent, and channel slope is 37.7 ft/mi. The watershed is unaffected by regulation, diversion, or mining.

1. Substituting the coefficient and basin characteristics for the site into equation 1 produces:

$$logBFY = 2.4343 + 1.7218 log(41.3) - 2.4602 log(59.6) + 1.2297 log(1 + 0.01(31.1)) + 0.5673 log(1 + 0.01(16.3)) + 0.0828 log(37.7)$$

$$logBFY = 1.1615$$

$$BFY = 14.5 \text{ in/yr}$$

2. To compute the prediction interval, the predicted base-flow yield and the $T$ value of 0.0746 are substituted into equation 3:

$$10^{(1.1615 - 0.0746)} \leq 14.5 \leq 10^{(1.1615 + 0.0746)}$$

$$10^{(1.0869)} \leq 14.5 \leq 10^{(1.2361)}$$

$$12.2 \leq 14.5 \leq 17.2$$

3. The predicted base-flow yield for the site as an estimate of ground-water recharge is 14.5 in/yr with a 90-percent probability that the base-flow yield is between 12.2 and 17.2 in/yr.

# Recharge Estimates for Basins in Pennsylvania

To illustrate the use of the equation, mean annual ground-water recharge for 1971–2000 was estimated for 370 hydrologic units (HUC-11 basins) in Pennsylvania. Recharge was assumed to equal the mean annual base-flow yield predicted by the equation. For each basin, the values for the five significant basin characteristics were determined, recharge was predicted, and the results were mapped (fig. 6). Mean annual recharge rates ranged from about 6 to 22 in/yr. About 80 percent of the basins (10th–90th percentiles) had predicted recharge rates between 10 and 18 in/yr.

# Summary

This report, in cooperation with the Pennsylvania Department of Conservation and Natural Resources, Bureau of Topographic and Geologic Survey, presents a regression equation for estimating long-term, mean annual ground-water recharge rates for basins in Pennsylvania and describes the methodology used to develop the equation. The equation should be useful for mapping regional variability in the magnitude of ground-water recharge. The equation relates base-flow yield (base flow divided by basin area expressed as inches basin-wide) to the climatic, geologic, hydrologic, and physiographic characteristics of a basin. In this study, base-flow yield is used as a proxy for ground-water recharge.

Base-flow yield was determined at 183 candidate stream-flow-gaging stations by use of the hydrograph-separation com-

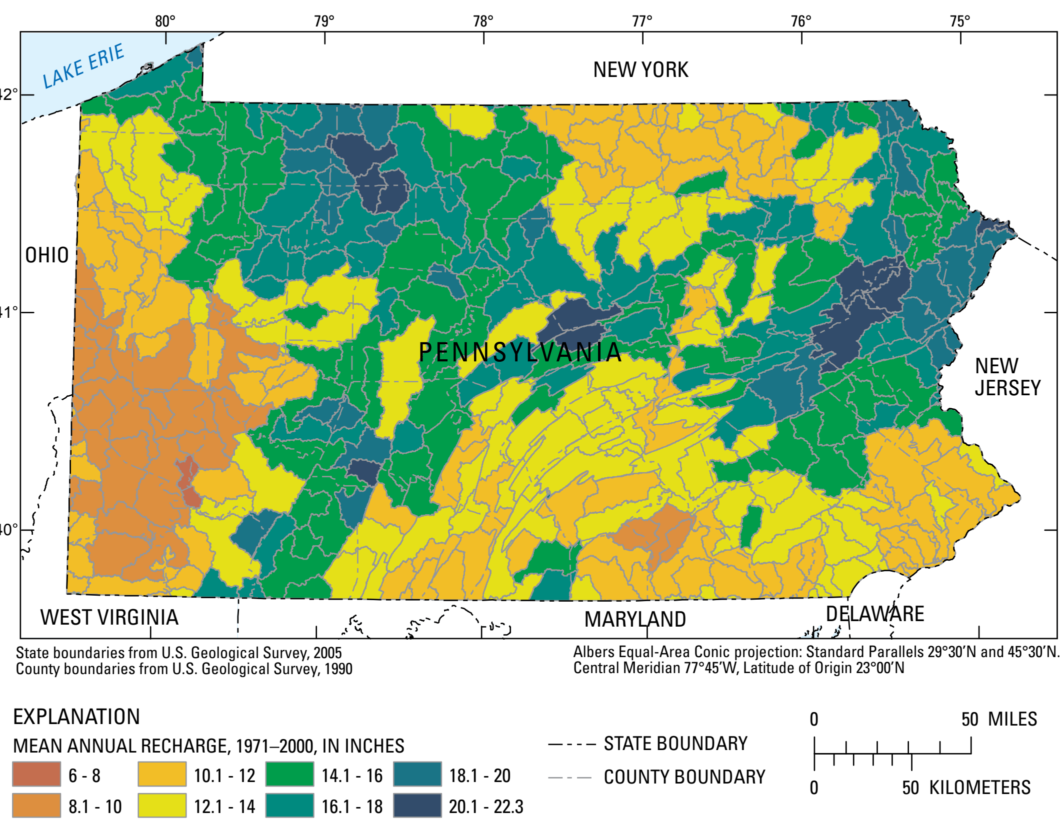


**Figure 6.** Ground-water recharge rates predicted from the regression equation for 370 hydrologic units (HUC-11 basins) in Pennsylvania.

puter program PART. Twenty-eight basin characteristics that might possibly affect base-flow yield were quantified for each basin from a variety of GIS sources. Base-flow yield was then related to the basin characteristics using ordinary least-squares (OLS) regression techniques to determine the explanatory variables (basin characteristics) significant at the 95-percent confidence level. Exploratory regression analysis indicated that mean annual precipitation was one of the most significant basin characteristics in the equation for predicting base-flow yield. Of the 183 streamflow-gaging stations, only 90 had at least 15 years of record available during 1971–2000, so those stations were selected for further analysis. Of the 90 stations, 10 stations were eliminated during evaluation of diagnostic statistics and period of record, to determine the 80 stations used in the final equation.

Basin characteristics significant at the 95-percent confidence level for predicting mean annual base-flow yield were mean annual precipitation, average daily maximum temperature, percentage of sand in the soil, percentage of carbonate bedrock within the basin, and channel slope. These five variables were used to develop a regression equation that could be used to estimate recharge for ungaged basins. The coefficient of determination ($r^2$) for the equation was 0.80, and the standard error of prediction was 0.043 log units (9.7 percent). To illustrate the use of the equation, mean annual recharge for 1971–2000 was estimated for each of 370 hydrologic units in Pennsylvania. The results were mapped and showed variability of ground-water recharge rates from about 6 to 22 in/yr across Pennsylvania.

## References Cited


Arnold, J.G., Allen, P.M., Muttiah, R., and Bernhardt, G., 1995, Automated base flow separation and recession analysis techniques: Ground Water, v. 33, no. 6, p. 1010-1018.

Daly, Christopher, 1996, Overview of the PRISM Model [online]: accessed May 27, 2008, at *http://www.prism.oregonstate.edu/docs/tech_desc.html*

Daniel, C.C., III, 1996, Ground-water recharge to the regolith-fractured crystalline rock aquifer system, Orange County, North Carolina: U.S. Geological Survey Water-Resources Investigations Report 96-4220, 59 p.

Delin, G.N., Healy, R.W., Lorenz, D.L., and Nimmo, J.R., 2007, Comparison of local- to regional-scale estimates of ground-water recharge in Minnesota, USA: Journal of Hydrology, v. 334, no. 1-2, p. 231-249.

Delin, G.N., and Risser, D.W., 2007, Ground-water recharge in humid areas of the United States—A summary of Ground-Water Resources Program studies, 2003-2006: U.S. Geological Survey Fact Sheet, 4 p.

Ehlke, M.H., and Reed, L.A., 1999, Comparison of methods for computing streamflow statistics for Pennsylvania: U.S. Geological Survey Water-Resources Investigations Report 99-4068, 80 p.

Flynn, R.H., and Tasker, G.D., 2004, Generalized estimates from streamflow data of annual and seasonal ground-water-recharge rates for drainage basins in New Hampshire: U.S. Geological Survey Scientific Investigations Report 2004-5019, 27 p.

Gebert, W.A., Radloff, M.J., Considine, E.J., and Kennedy, J.L., 2007, Use of streamflow data to estimate base flow/ground-water recharge for Wisconsin: Journal of the American Water Resources Association, v. 43, no. 1, p. 220-236.

Heath, R.C., 1994, Ground-water recharge in North Carolina: North Carolina State University, report to the N.C. Dept. of Environment, Health, and Natural Resources, Division of Environmental Management, Groundwater Section, 52 p.

Helsel, D.R., and Hirsch, R.M., 1993, Statistical methods in water resources: Amsterdam, The Netherlands, Elsevier, Sarah Burgerhartstraat 25, 1000 AE Studies in Environmental Science 49, 529 p.

Holtschlag, D.J., 1997, A generalized estimate of ground-water-recharge rates in the lower peninsula of Michigan: U.S. Geological Survey Water-Supply Paper 2437, 37 p.

Homer, C., Huang, C., Yang, L., Wylie, B., and Coan, M., 2004, Development of a 2001 National Landcover Database for the United States: Photogrammetric Engineering and Remote Sensing, v. 70, no. 7, July 2004, p. 829-840, accessed May 27, 2008, at *http://www.mrlc.gov/pdf/July_PERS.pdf*

MathSoft, Inc., 1997, S-PLUS user's guide: Seattle, Wash., Data Analysis Products Division, MathSoft, Inc., 620 p.

Miles, C.E., Whitfield, G.T., and other staff and interns of the Pennsylvania Geological Survey, 2001, Bedrock geologic units of Pennsylvania, based on Berg, T.M., Edmunds, W.E., Geyer, A.R., Glover, A.D., Hoskins, D.M., MacLachlan, D.B., Root, S.I., Sevon, W.D., and Socolow, A.A., 1980, Geologic map of Pennsylvania: Pennsylvania Geological Survey, 4th ser., Map 1, scale 1:250,000. (*http://www.dcnr.state.pa.us/topogeo/map1/bedmap.aspx*)

Neff, B.P., Piggott, A.R., and Sheets, R.A., 2006, Estimation of shallow ground-water recharge in the Great Lakes Basin: U.S. Geological Survey Scientific Investigations Report 2005-5284, 20 p.

Nelms, D.L., Harlow, G.E., Jr., and Hayes, D.C., 1997, Base-flow characteristics of streams in the Valley and Ridge, the Blue Ridge, and the Piedmont physiographic provinces of Virginia: U.S. Geological Survey Water-Supply Paper 2457, 48 p.

Pennsylvania Department of Conservation and Natural Resources, 1995, DCNR Glacier 100k 1995: Pennsylvania Bureau of Topographic and Geologic Survey, accessed July 13, 2006, at *http://www.pasda.psu.edu/uci/MetadataDisplay.aspx?entry=PASDA&file=pags_glacier1k.xml&dataset=293*

Pennsylvania Department of Conservation and Natural Resources, 2008, Sinkhole inventory and online database: Pennsylvania Geological Survey, Sinkholes in Pennsylvania, accessed April 4, 2008, at *http://www.dcnr.state.pa.us/topogeo/hazards/sinkholes.aspx*

Pettyjohn, W.A., and Henning, Roger, 1979, Preliminary estimate of ground-water recharge rates, related streamflow and water quality in Ohio: Ohio State University Water Resources Center Project Completion Report Number 552, 323 p.

PRISM Group, 2007, Average annual precipitation GIS data 1971–2000: accessed November 8, 2007, at *http://www.prism.oregonstate.edu/products/matrix.phtml?vartype=ppt*

PRISM Group, 2007a, Average maximum temperature GIS data 1971–2000: accessed November 8, 2007, at *http://www.prism.oregonstate.edu/products/matrix.phtml?vartype=tmax*

Risser, D.W., Conger, R.W., Ulrich, J.E., and Asmussen, M.P., 2005, Estimates of ground-water recharge based on streamflow-hydrograph methods—Pennsylvania: U.S. Geological Survey Open-File Report 2005-1333, 30 p.

Rutledge, A.T., 1993, Computer programs for describing the recession of ground-water discharge and for estimating mean ground-water recharge and discharge from streamflow records: U.S. Geological Survey Water-Resources Investigations Report 93-4121, 45 p.

Rutledge, A.T., 1998, Computer programs for describing the recession of ground-water discharge and for estimating mean ground-water recharge and discharge from streamflow records—Update: U.S. Geological Survey Water-Resources Investigations Report 98-4148, 43 p.

Scanlon, B.R., Healy, R.W., and Cook, P.G., 2002, Choosing appropriate techniques for quantifying groundwater recharge: Hydrogeology Journal, v. 10, p. 18-39.

Stuckey, M.H., 2006, Low-flow, base-flow, and mean-flow regression equations for Pennsylvania streams: U.S. Geological Survey Scientific Investigations Report 2006-5130, 84 p.

Stuckey, M.H., and Reed, L.A., 2000, Techniques for estimating magnitude and frequency of peak flows for Pennsylvania streams: U.S. Geological Survey Water-Resources Investigations Report 00-4189, 43 p.

Thompson, R.E., and Hoffman, S.A., 2006, Selected streamflow statistics and regression equations for predicting statistics at stream locations in Monroe County, Pennsylvania: U.S. Geological Survey Scientific Investigations Report 2006-5244, 45 p.

U.S. Department of Agriculture, 1994, State Soil Geographic (STATSGO) data base for Pennsylvania [online]: accessed January 17, 2006, at *http://www.ncgc.nrcs.usda.gov/products/datasets/statsgo*

U.S. Geological Survey, 1989, Federal Glossary of selected terms – subsurface-water flow and solute transport: U.S. Geological Survey, Office of Water Data Coordination, Ground Water Subcommittee of the Federal Interagency Advisory Committee on Water Data, 38 p.

U.S. Geological Survey, 1990, County boundary GIS data: accessed March 3, 2008, at *http://www.nationalatlas.gov/atlasftp.html*

U.S. Geological Survey, 2000a, US GeoData Digital Elevation Models: U.S. Geological Survey Fact Sheet 040-00, 3 p., accessed January 30, 2006, at *http://erg.usgs.gov/isb/pubs/factsheets/fs04000.html*

U.S. Geological Survey, 2000b, The National Hydrography Dataset [online]: accessed January 3, 2006, at *http://nhd.usgs.gov/chapter1/index.html*

U.S. Geological Survey, 2005, State boundary and river GIS data: accessed March 3, 2008, at *http://www.nationalatlas.gov/atlasftp.html*

Wahl, K.L., and Wahl, T.L., 1988, BFI—A computer program for determining an index to base flow: U.S. Bureau of Reclamation Water Resources Research Laboratory, accessed August 25, 2004, at *http://www.usbr.gov.pmts.hydraulics_lab/twahl/bfi/.*

White, K.E., and Sloto, R.A., 1990, Base-flow-frequency characteristics of selected Pennsylvania streams: U.S. Geological Survey Water-Resources Investigations Report 90-4160, 67 p.

Wolock, D.M., 2003, Estimated mean annual natural ground-water recharge in the conterminous United States: U.S. Geological Survey Open-File Report 2003-311, digital data set, accessed July 12, 2008, at *http://water.usgs.gov/GIS/metadata/usgswrd/XML/rech48grd.xml*

**Table 3.** Description of streamflow-gaging stations and period of record used to develop a regression equation for estimating ground-water recharge in Pennsylvania.

| Identifier on figure 2 | U.S. Geological Survey streamflow-gaging-station number | Station name | Latitude (decimal degrees) | Longitude (decimal degrees) | Period of record used in analysis (calendar years) | Number of years used in analysis | Drainage area (square miles) |
|---|---|---|---|---|---|---|---|
| 1 | 01440400 | Broadhead Creek near Analomink | 41.08472 | -75.21500 | 1971-2000 | 30 | 67.5 |
| 2 | 01442500 | Broadhead Creek near Minisink Hills | 40.99861 | -75.14306 | 1971-2000 | 30 | 261 |
| 3 | 01447500 | Lehigh River at Stoddartsville | 41.13028 | -75.62583 | 1971-2000 | 30 | 91.8 |
| 4 | 01447720 | Tobyhanna Creek near Blakeslee | 41.08472 | -75.60583 | 1971-1985 | 15 | 119 |
| 5 | 01449360 | Pohopoco Creek at Kresgeville | 40.89750 | -75.50278 | 1971-2000 | 30 | 49.8 |
| 6 | 01451500 | Little Lehigh Creek near Allentown | 40.58222 | -75.48333 | 1971-2000 | 30 | 81.9 |
| 7 | 01451800 | Jordan Creek near Schnecksville | 40.66167 | -75.62722 | 1971-2000 | 30 | 53.0 |
| 8 | 01452000 | Jordan Creek at Allentown | 40.62306 | -75.48278 | 1971-2000 | 30 | 76.3 |
| 9 | 01452500 | Monocacy Creek at Bethlehem | 40.64111 | -75.37972 | 1971-2000 | 30 | 43.3 |
| 10 | 01467048 | PennyPack Creek at Lower Rhawn St Bdg, Phila. | 40.05000 | -75.03306 | 1971-2000 | 30 | 49.8 |
| 11 | 01467086 | Tacony Creek at County Line, Philadelphia | 40.04639 | -75.11111 | 1971-1985 | 16 | 16.2 |
| 12 | 01468500 | Schuylkill River at Landingville | 40.62917 | -76.12500 | 1974-2000 | 26 | 137 |
| 13 | 01470500 | Schuylkill River at Berne | 40.52250 | -75.99861 | 1971-2000 | 30 | 358 |
| 14 | 01470756 | Maiden Creek at Virginville | 40.51417 | -75.88333 | 1974-1994 | 21 | 159 |
| 15 | 01470779 | Tulpehocken Creek near Bernville | 40.41333 | -76.17194 | 1975-2000 | 25 | 70.5 |
| 16 | 01471980 | Manatawny Creek near Pottstown | 40.27278 | -75.68028 | 1975-2000 | 24 | 85.5 |
| 17 | 01472157 | French Creek near Phoenixville | 40.15139 | -75.60167 | 1971-2000 | 30 | 59.0 |
| 18 | 01472198 | Perkiomen Creek at East Greenville | 40.39389 | -75.51583 | 1982-2000 | 18 | 37.7 |
| 19 | 01472199 | West Branch Perkiomen Creek at Hillegrass | 40.37389 | -75.52278 | 1982-2000 | 18 | 23.1 |
| 20 | 01473169 | Valley Creek at Pa Turnpike Bridge near Valley Forge | 40.07917 | -75.46111 | 1983-2000 | 17 | 20.8 |
| 21 | 01474000 | Wissahickon Creek at Mouth | 40.07917 | -75.22861 | 1971-2000 | 30 | 63.9 |
| 22 | 01475300 | Darby Creek at Waterloo Mills near Devon | 40.02250 | -75.42222 | 1973-1993 | 21 | 5.19 |
| 23 | 01475850 | Crum Creek near Newtown Square | 39.97639 | -75.43694 | 1982-2000 | 19 | 15.8 |
| 24 | 01480300 | West Branch Brandywine Creek near Honey Brook | 40.07278 | -75.86111 | 1971-2000 | 30 | 18.5 |
| 25 | 01480675 | Marsh Creek near Glenmoore, | 40.09778 | -75.74194 | 1971-2000 | 30 | 8.54 |
| 26 | 01481000 | Brandywine Creek at Chadds Ford | 39.86972 | -75.59361 | 1971-2000 | 30 | 288 |
| 27 | 01516350 | Tioga River near Mansfield | 41.79417 | -77.08056 | 1977-2000 | 23 | 153 |
| 28 | 01532000 | Towanda Creek near Monroeton | 41.70694 | -76.48500 | 1971-2000 | 30 | 215 |
| 29 | 01534000 | Tunkhannock Creek near Tunkhannock | 41.55833 | -75.89500 | 1971-2000 | 30 | 393 |
| 30 | 01538000 | Wapwallopen Creek near Wapwallopen | 41.05917 | -76.09389 | 1971-2000 | 30 | 42.9 |

**Table 3.** Description of streamflow-gaging stations and period of record used to develop a regression equation for estimating ground-water recharge in Pennsylvania.—Continued

| Identifier on figure 2 | U.S. Geological Survey streamflow-gaging-station number | Station name | Latitude (decimal degrees) | Longitude (decimal degrees) | Period of record used in analysis (calendar years) | Number of years used in analysis | Drainage area (square miles) |
|---|---|---|---|---|---|---|---|
| 31 | 01539000 | Fishing Creek near Bloomsburg, | 41.07806 | -76.43139 | 1971-2000 | 30 | 272 |
| 32 | 01541000 | West Branch Susquehanna River at Bower | 40.89694 | -78.67722 | 1971-2000 | 30 | 315 |
| 33 | 01542000 | Moshannon Creek at Osceola Mills | 40.84944 | -78.26806 | 1971-1992 | 23 | 68.7 |
| 34 | 01543000 | Driftwood Br Sinnemahoning Cr at Sterling Run | 41.41333 | -78.19722 | 1971-2000 | 30 | 297 |
| 35 | 01544500 | Kettle Creek at Cross Fork | 41.47583 | -77.82611 | 1971-2000 | 30 | 137 |
| 36 | 01545600 | Young Womans Creek near Renovo | 41.38944 | -77.69111 | 1971-2000 | 30 | 46.3 |
| 37 | 01546500 | Spring Creek near Axemann | 40.88972 | -77.79444 | 1971-2000 | 30 | 85.9 |
| 38 | 01547100 | Spring Creek at Milesburg | 40.93167 | -77.78694 | 1971-2000 | 30 | 145 |
| 39 | 01547200 | Bald Eagle Creek below Spring Creek at Milesburg | 40.94306 | -77.78667 | 1971-2000 | 30 | 267 |
| 40 | 01547700 | Marsh Creek at Blanchard | 41.05944 | -77.60611 | 1971-2000 | 30 | 44.6 |
| 41 | 01547950 | Beech Creek at Monument | 41.11167 | -77.70250 | 1971-2000 | 30 | 153 |
| 42 | 01549500 | Blockhouse Creek near English Center | 41.47361 | -77.23111 | 1971-2000 | 30 | 37.9 |
| 43 | 01550000 | Lycoming Creek near Trout Run | 41.41833 | -77.03306 | 1971-2000 | 30 | 173 |
| 44 | 01552500 | Muncy Creek near Sonestown | 41.35694 | -76.53500 | 1971-2000 | 30 | 23.4 |
| 45 | 01555000 | Penns Creek at Penns Creek | 40.86667 | -77.04861 | 1971-2000 | 30 | 306 |
| 46 | 01555500 | East Mahantango Creek near Dalmatia | 40.61111 | -76.91222 | 1971-2000 | 30 | 162 |
| 47 | 01556000 | Frankstown Br Juniata River at Williamsburg | 40.46306 | -78.20000 | 1971-2000 | 30 | 289 |
| 48 | 01557500 | Bald Eagle Creek at Tyrone | 40.68361 | -78.23389 | 1971-2000 | 30 | 44.6 |
| 49 | 01558000 | Little Juniata River at Spruce Creek | 40.61250 | -78.14083 | 1971-2000 | 30 | 221 |
| 50 | 01560000 | Dunning Creek at Belden | 40.07167 | -78.49278 | 1971-2000 | 30 | 172 |
| 51 | 01564500 | Aughwick Creek near Three Springs | 40.21250 | -77.92556 | 1971-2000 | 30 | 204 |
| 52 | 01567500 | Bixler Run near Loysville | 40.37083 | -77.40250 | 1971-2000 | 30 | 15.0 |
| 53 | 01568000 | Sherman Creek at Shermans Dale | 40.32333 | -77.16917 | 1971-2000 | 30 | 207 |
| 54 | 01573160 | Quittapahilla Creek near Belle-grove | 40.34278 | -76.56278 | 1976-1993 | 18 | 73.4 |
| 55 | 01573560 | Swatara Creek near Hershey | 40.29833 | -76.66806 | 1976-2000 | 23 | 483 |
| 56 | 01574000 | West Conewago Creek near Man-chester | 40.08222 | -76.72028 | 1971-2000 | 30 | 512 |
| 57 | 01576754 | Conestoga River at Conestoga | 39.94639 | -76.36806 | 1985-2000 | 15 | 468 |
| 58 | 03007800 | Allegheny River at Port Allegany | 41.81861 | -78.29306 | 1975-2000 | 24 | 251 |
| 59 | 03009680 | Potato Creek at Smethport | 41.80972 | -78.43056 | 1975-1990 | 16 | 162 |
| 60 | 03010500 | Allegheny River at Eldred | 41.96333 | -78.38639 | 1971-2000 | 30 | 549 |

**Table 3.** Description of streamflow-gaging stations and period of record used to develop a regression equation for estimating ground-water recharge in Pennsylvania.—Continued

| Identifier on figure 2 | U.S. Geological Survey streamflow-gaging-station number | Station name | Latitude (decimal degrees) | Longitude (decimal degrees) | Period of record used in analysis (calendar years) | Number of years used in analysis | Drainage area (square miles) |
|---|---|---|---|---|---|---|---|
| 61 | 03010655 | Oswayo Creek at Shinglehouse | 41.96167 | -78.19833 | 1975-2000 | 23 | 98.2 |
| 62 | 03011800 | Kinzua Creek near Guffey | 41.76639 | -78.71889 | 1971-2000 | 30 | 38.7 |
| 63 | 03015500 | Brokenstraw Creek at Youngsville | 41.85250 | -79.31750 | 1971-2000 | 30 | 304 |
| 64 | 03020500 | Oil Creek at Rouseville | 41.48167 | -79.69556 | 1971-2000 | 30 | 283 |
| 65 | 03021350 | French Creek near Wattsburg | 42.01528 | -79.78278 | 1975-2000 | 24 | 94.4 |
| 66 | 03021410 | West Branch French Creek near Lowville | 42.08167 | -79.85056 | 1975-1992 | 18 | 52.7 |
| 67 | 03022540 | Woodcock Creek at Blooming Valley | 41.69056 | -80.04833 | 1975-1994 | 20 | 30.2 |
| 68 | 03026500 | Sevenmile Run near Rasselas | 41.63111 | -78.57694 | 1971-2000 | 30 | 7.85 |
| 69 | 03028000 | West Branch Clarion River at Wilcox | 41.57528 | -78.69250 | 1971-2000 | 30 | 63.3 |
| 70 | 03032500 | Redbank Creek at St. Charles | 40.99444 | -79.39444 | 1971-2000 | 30 | 533 |
| 71 | 03034000 | Mahoning Creek at Punxsutawnery | 40.93917 | -79.00861 | 1971-2000 | 30 | 157 |
| 72 | 03034500 | Little Mahoning Creek at McCor-mick | 40.83611 | -79.11028 | 1971-2000 | 30 | 87.7 |
| 73 | 03042000 | Blacklick Creek at Josephine | 40.47333 | -79.18361 | 1971-2000 | 30 | 193 |
| 74 | 03049000 | Buffalo Creek near Freeport | 40.71583 | -79.69972 | 1971-2000 | 30 | 138 |
| 75 | 03072000 | Dunkard Creek at Shannopin | 39.75917 | -79.97083 | 1971-2000 | 30 | 227 |
| 76 | 03074500 | Redstone Creek at Waltersburg | | -79.76444 | 1971-2000 | 30 | 73.9 |
| 77 | 03079000 | Casselman River at Markleton | 39.85972 | -79.22778 | 1971-2000 | 30 | 382 |
| 78 | 03080000 | Laurel Hill Creek at Ursina | 39.82028 | -79.32167 | 1971-2000 | 30 | 121 |
| 79 | 03102500 | Litte Shenango River at Greenville | 41.42194 | -80.37639 | 1971-2000 | 30 | 104 |
| 80 | 03106000 | Connoquenessing Creek near Zelienople | 40.81694 | -80.24250 | 1971-2000 | 30 | 356 |

**Table 4.** Basin characteristics for streamflow-gaging stations used to develop a regression equation for estimating ground-water recharge in Pennsylvania.

[Sources used to derive basin characteristics are referenced in table 1. ft/mi, feet per mile; in., inches; %, percent; °F, degrees Fahrenheit]

| Identifier on figure 2 | U.S. Geological Survey streamflow-gaging-station number | Channel slope (ft/mi) | Mean annual precipitation (in.) | Carbonate rock in basin (%) | Sand in the soil (%) | Average daily maximum temperature (°F) |
|---|---|---|---|---|---|---|
| 1 | 01440400 | 78.3 | 44.46 | 0 | 29.25 | 54.3 |
| 2 | 01442500 | 56.2 | 46.29 | 0 | 30.33 | 56.2 |
| 3 | 01447500 | 17.5 | 46.02 | 0 | 31.63 | 53.08 |
| 4 | 01447720 | 20.6 | 47.64 | 0 | 35.29 | 52.95 |
| 5 | 01449360 | 67.4 | 47.43 | 0 | 31.04 | 57.20 |
| 6 | 01451500 | 11.5 | 45.20 | 64.0 | 22.67 | 59.93 |
| 7 | 01451800 | 17.9 | 45.64 | 0 | 27.76 | 59.76 |
| 8 | 01452000 | 11.8 | 45.41 | 11.0 | 26.30 | 59.85 |
| 9 | 01452500 | 23.7 | 44.56 | 63.0 | 21.70 | 59.74 |
| 10 | 01467048 | 11.9 | 45.62 | 1.0 | 23.62 | 61.95 |
| 11 | 01467086 | 27.4 | 44.99 | 0 | 21.73 | 61.99 |
| 12 | 01468500 | 51.4 | 48.72 | 0 | 44.28 | 57.21 |
| 13 | 01470500 | 19.9 | 48.71 | 0 | 36.65 | 57.66 |
| 14 | 01470756 | 8.1 | 46.77 | 11.0 | 29.34 | 59.77 |
| 15 | 01470779 | 10.2 | 43.28 | 83.0 | 19.16 | 60.01 |
| 16 | 01471980 | 25.6 | 45.95 | 26.0 | 21.05 | 60.30 |
| 17 | 01472157 | 25.0 | 44.93 | 1.0 | 23.59 | 60.23 |
| 18 | 01472198 | 47.7 | 45.31 | 3.0 | 22.16 | 60.34 |
| 19 | 01472199 | 39.3 | 46.17 | 5.0 | 21.68 | 60.33 |
| 20 | 01473169 | 30.0 | 44.60 | 67.0 | 18.71 | 61.42 |
| 21 | 01474000 | 12.5 | 45.00 | 16.0 | 21.46 | 61.73 |
| 22 | 01475300 | 37.1 | 44.99 | 0 | 19.84 | 60.82 |
| 23 | 01475850 | 31.2 | 45.00 | 0 | 18.58 | 60.73 |
| 24 | 01480300 | 25.1 | 45.00 | 3.0 | 22.89 | 59.80 |
| 25 | 01480675 | 24.7 | 45.01 | 0 | 22.00 | 59.97 |
| 26 | 01481000 | 13.8 | 45.00 | 8.0 | 21.22 | 60.54 |
| 27 | 01516350 | 37.4 | 34.09 | 0 | 25.39 | 53.45 |
| 28 | 01532000 | 12.2 | 36.27 | 0 | 24.82 | 54.75 |
| 29 | 01534000 | 17.7 | 40.32 | 0 | 25.85 | 55.06 |
| 30 | 01538000 | 63.8 | 43.68 | 0 | 26.07 | 55.98 |
| 31 | 01539000 | 17.9 | 42.06 | 0 | 26.39 | 56.22 |
| 32 | 01541000 | 15.1 | 44.29 | 0 | 25.34 | 56.24 |
| 33 | 01542000 | 32.5 | 39.24 | 0 | 37.70 | 55.77 |
| 34 | 01543000 | 23.6 | 43.40 | 0 | 29.29 | 55.32 |
| 35 | 01544500 | 35.7 | 39.91 | 0 | 31.85 | 54.75 |

**Table 4.** Basin characteristics for streamflow-gaging stations used to develop a regression equation for estimating ground-water recharge in Pennsylvania.—Continued

[Sources used to derive basin characteristics are referenced in table 1. ft/mi, feet per mile; in., inches; %, percent; °F, degrees Fahrenheit]

| Identifier on figure 2 | U.S. Geological Survey streamflow-gaging-station number | Channel slope (ft/mi) | Mean annual precipitation (in.) | Carbonate rock in basin (%) | Sand in the soil (%) | Average daily maximum temperature (°F) |
|---|---|---|---|---|---|---|
| 36 | 01545600 | 76.1 | 41.75 | 0 | 43.37 | 54.83 |
| 37 | 01546500 | 19.5 | 38.76 | 72.3 | 26.11 | 57.58 |
| 38 | 01547100 | 22.4 | 38.48 | 73.1 | 27.22 | 57.86 |
| 39 | 01547200 | 19.4 | 38.19 | 44.2 | 29.50 | 57.53 |
| 40 | 01547700 | 59.6 | 39.03 | 0 | 28.20 | 57.21 |
| 41 | 01547950 | 37.1 | 40.11 | 0 | 44.32 | 55.85 |
| 42 | 01549500 | 54.6 | 36.17 | 0 | 24.41 | 54.25 |
| 43 | 01550000 | 44.7 | 36.70 | 0 | 29.08 | 53.29 |
| 44 | 01552500 | 111.0 | 45.40 | 0 | 25.04 | 53.46 |
| 45 | 01555000 | 13.3 | 43.75 | 24.0 | 35.99 | 57.46 |
| 46 | 01555500 | 9.3 | 42.67 | 0 | 30.22 | 58.30 |
| 47 | 01556000 | 12.0 | 39.56 | 21.0 | 35.76 | 57.77 |
| 48 | 01557500 | 48.5 | 38.89 | 5.0 | 33.64 | 56.36 |
| 49 | 01558000 | 40.3 | 39.65 | 20.0 | 35.46 | 55.96 |
| 50 | 01560000 | 38.0 | 38.73 | 4.0 | 27.36 | 57.44 |
| 51 | 01564500 | 12.9 | 38.29 | 4.0 | 33.84 | 58.95 |
| 52 | 01567500 | 42.4 | 39.89 | 21.3 | 32.08 | 59.35 |
| 53 | 01568000 | 15.9 | 39.65 | 9.2 | 36.74 | 58.80 |
| 54 | 01573160 | 10.1 | 42.77 | 76.0 | 19.25 | 59.98 |
| 55 | 01573560 | 9.3 | 44.10 | 13.0 | 31.03 | 60.02 |
| 56 | 01574000 | 5.1 | 39.72 | 6.0 | 22.76 | 61.34 |
| 57 | 01576754 | 5.3 | 42.40 | 59.0 | 20.60 | 60.32 |
| 58 | 03007800 | 16.3 | 40.39 | 0 | 28.18 | 54.00 |
| 59 | 03009680 | 22.4 | 44.45 | 0 | 33.67 | 53.89 |
| 60 | 03010500 | 8.7 | 42.03 | 0 | 28.76 | 53.04 |
| 61 | 03010655 | 27.1 | 39.13 | 0 | 32.62 | 54.39 |
| 62 | 03011800 | 34.0 | 45.00 | 0 | 37.02 | 53.13 |
| 63 | 03015500 | 6.6 | 46.80 | 0 | 21.30 | 55.50 |
| 64 | 03020500 | 9.4 | 44.54 | 0 | 20.19 | 55.53 |
| 65 | 03021350 | 11.2 | 47.03 | 0 | 18.06 | 55.90 |
| 66 | 03021410 | 8.8 | 46.94 | 0 | 19.68 | 55.00 |
| 67 | 03022540 | 25.9 | 45.00 | 0 | 19.49 | 56.33 |
| 68 | 03026500 | 80.7 | 45.01 | 0 | 37.14 | 53.10 |
| 69 | 03028000 | 48.0 | 45.00 | 0 | 36.43 | 53.40 |
| 70 | 03032500 | 5.7 | 43.24 | 0 | 23.66 | 55.86 |

**Table 4.** Basin characteristics for streamflow-gaging stations used to develop a regression equation for estimating ground-water recharge in Pennsylvania.—Continued

[Sources used to derive basin characteristics are referenced in table 1. ft/mi, feet per mile; in., inches; %, percent; °F, degrees Fahrenheit]

| Identifier on figure 2 | U.S. Geological Survey streamflow-gaging-station number | Channel slope (ft/mi) | Mean annual precipitation (in.) | Carbonate rock in basin (%) | Sand in the soil (%) | Average daily maximum temperature (°F) |
|---|---|---|---|---|---|---|
| 71 | 03034000 | 9.7 | 44.32 | 0 | 22.95 | 56.49 |
| 72 | 03034500 | 15.9 | 47.14 | 0 | 26.90 | 56.84 |
| 73 | 03042000 | 19.1 | 46.45 | 0 | 36.14 | 56.57 |
| 74 | 03049000 | 11.6 | 41.00 | 0 | 21.37 | 59.41 |
| 75 | 03072000 | 2.7 | 42.03 | 0 | 17.30 | 61.08 |
| 76 | 03074500 | 39.5 | 43.67 | 0 | 20.63 | 59.84 |
| 77 | 03079000 | 16.0 | 41.73 | 0 | 23.72 | 55.66 |
| 78 | 03080000 | 23.4 | 46.19 | 0 | 28.21 | 55.52 |
| 79 | 03102500 | 23.5 | 40.51 | 0 | 20.95 | 58.03 |
| 80 | 03106000 | 4.8 | 39.13 | 0 | 28.04 | 59.31 |
| | Minimum | 2.7 | 34.09 | 0 | 17.30 | 52.95 |
| | Mean | 27.2 | 43.03 | 10.2 | 27.39 | 57.34 |
| | Maximum | 111.0 | 48.72 | 83.0 | 44.32 | 61.99 |

**Table 5.** Observed and predicted mean-annual base-flow yield values, with residuals for 80 basins with streamflow-gaging stations used to develop the regression equation for estimating recharge rates in Pennsylvania, 1971-2000.

| Identifier on figure 2 | U.S. Geological Survey streamflow-gaging-station number | "Observed" mean annual base-flow yield 1971-2000 from streamflow data (expressed as inches over basin area) | Predicted mean annual base-flow yield, 1971-2000 from regression equation (expressed as inches over basin area) | Residuals, observed minus predicted (expressed as inches over basin area) |
|---|---|---|---|---|
| 1 | 01440400 | 22.16 | 19.85 | 2.31 |
| 2 | 01442500 | 22.41 | 19.22 | 3.19 |
| 3 | 01447500 | 22.28 | 20.12 | 2.16 |
| 4 | 01447720 | 24.33 | 22.53 | 1.80 |
| 5 | 01449360 | 23.79 | 19.61 | 4.18 |
| 6 | 01451500 | 15.80 | 16.97 | -1.17 |
| 7 | 01451800 | 15.51 | 14.31 | 1.20 |
| 8 | 01452000 | 13.79 | 14.29 | -.50 |
| 9 | 01452500 | 15.99 | 17.48 | -1.49 |
| 10 | 01467048 | 13.55 | 12.22 | 1.33 |
| 11 | 01467086 | 12.39 | 12.45 | -.06 |
| 12 | 01468500 | 22.88 | 22.60 | .28 |
| 13 | 01470500 | 21.18 | 19.16 | 2.02 |
| 14 | 01470756 | 14.69 | 15.05 | -.36 |
| 15 | 01470779 | 18.41 | 15.96 | 2.45 |
| 16 | 01471980 | 13.94 | 15.57 | -1.63 |
| 17 | 01472157 | 14.04 | 13.56 | .48 |
| 18 | 01472198 | 14.37 | 14.41 | -.04 |
| 19 | 01472199 | 14.46 | 14.74 | -.28 |
| 20 | 01473169 | 15.82 | 16.40 | -.58 |
| 21 | 01474000 | 13.09 | 12.79 | .30 |
| 22 | 01475300 | 15.53 | 13.13 | 2.40 |
| 23 | 01475850 | 12.86 | 12.82 | .04 |
| 24 | 01480300 | 12.13 | 13.90 | -1.77 |
| 25 | 01480675 | 12.31 | 13.44 | -1.13 |
| 26 | 01481000 | 14.97 | 12.97 | 2.00 |
| 27 | 01516350 | 11.38 | 11.84 | -.46 |
| 28 | 01532000 | 10.38 | 11.25 | -.87 |
| 29 | 01534000 | 12.27 | 13.87 | -1.60 |
| 30 | 01538000 | 15.87 | 17.03 | -1.16 |
| 31 | 01539000 | 16.61 | 14.26 | 2.35 |
| 32 | 01541000 | 15.60 | 15.20 | .40 |
| 33 | 01542000 | 17.21 | 15.07 | 2.14 |
| 34 | 01543000 | 14.01 | 16.47 | -2.46 |
| 35 | 01544500 | 15.57 | 15.51 | .06 |

**Table 5.** Observed and predicted mean-annual base-flow yield values, with residuals for 80 basins with streamflow-gaging stations used to develop the regression equation for estimating recharge rates in Pennsylvania, 1971-2000.—Continued

| Identifier on figure 2 | U.S. Geological Survey streamflow-gaging-station number | "Observed" mean annual base-flow yield 1971-2000 from streamflow data (expressed as inches over basin area) | Predicted mean annual base-flow yield, 1971-2000 from regression equation (expressed as inches over basin area) | Residuals, observed minus predicted (expressed as inches over basin area) |
|---|---|---|---|---|
| 36 | 01545600 | 16.87 | 19.71 | -2.84 |
| 37 | 01546500 | 14.79 | 15.97 | -1.18 |
| 38 | 01547100 | 16.88 | 15.98 | .90 |
| 39 | 01547200 | 16.85 | 14.56 | 2.29 |
| 40 | 01547700 | 12.42 | 13.50 | -1.08 |
| 41 | 01547950 | 18.55 | 16.70 | 1.85 |
| 42 | 01549500 | 14.52 | 12.91 | 1.61 |
| 43 | 01550000 | 15.80 | 14.24 | 1.56 |
| 44 | 01552500 | 18.96 | 21.13 | -2.17 |
| 45 | 01555000 | 16.58 | 17.45 | -.87 |
| 46 | 01555500 | 13.30 | 13.14 | .16 |
| 47 | 01556000 | 13.26 | 14.12 | -.86 |
| 48 | 01557500 | 16.77 | 14.81 | 1.96 |
| 49 | 01558000 | 18.20 | 16.83 | 1.37 |
| 50 | 01560000 | 11.44 | 12.89 | -1.45 |
| 51 | 01564500 | 9.72 | 11.53 | -1.81 |
| 52 | 01567500 | 13.44 | 14.41 | -.97 |
| 53 | 01568000 | 13.09 | 13.23 | -.14 |
| 54 | 01573160 | 16.25 | 15.31 | .94 |
| 55 | 01573560 | 13.54 | 13.98 | -.44 |
| 56 | 01574000 | 8.75 | 9.37 | -.62 |
| 57 | 01576754 | 13.31 | 13.51 | -.20 |
| 58 | 03007800 | 16.58 | 14.82 | 1.76 |
| 59 | 03009680 | 17.55 | 19.00 | -1.45 |
| 60 | 03010500 | 15.66 | 15.84 | -.18 |
| 61 | 03010655 | 15.52 | 15.00 | .52 |
| 62 | 03011800 | 20.05 | 21.44 | -1.39 |
| 63 | 03015500 | 14.73 | 15.47 | -.74 |
| 64 | 03020500 | 14.54 | 14.45 | .09 |
| 65 | 03021350 | 15.32 | 15.50 | -.18 |
| 66 | 03021410 | 15.81 | 16.02 | -.21 |
| 67 | 03022540 | 13.55 | 15.34 | -1.79 |
| 68 | 03026500 | 18.71 | 23.10 | -4.39 |
| 69 | 03028000 | 20.09 | 21.67 | -1.58 |
| 70 | 03032500 | 15.09 | 13.45 | 1.64 |

**Table 5.** Observed and predicted mean-annual base-flow yield values, with residuals for 80 basins with streamflow-gaging stations used to develop the regression equation for estimating recharge rates in Pennsylvania, 1971-2000.—Continued

| Identifier on figure 2 | U.S. Geological Survey streamflow-gaging-station number | "Observed" mean annual base-flow yield 1971-2000 from streamflow data (expressed as inches over basin area) | Predicted mean annual base-flow yield, 1971-2000 from regression equation (expressed as inches over basin area) | Residuals, observed minus predicted (expressed as inches over basin area) |
|---|---|---|---|---|
| 71 | 03034000 | 16.84 | 14.17 | 2.67 |
| 72 | 03034500 | 13.97 | 16.81 | -2.84 |
| 73 | 03042000 | 17.45 | 18.35 | -.90 |
| 74 | 03049000 | 11.56 | 10.93 | .63 |
| 75 | 03072000 | 8.03 | 9.05 | -1.02 |
| 76 | 03074500 | 13.80 | 13.15 | .65 |
| 77 | 03079000 | 14.45 | 13.92 | .53 |
| 78 | 03080000 | 18.78 | 17.98 | .80 |
| 79 | 03102500 | 10.92 | 11.98 | -1.06 |
| 80 | 03106000 | 10.22 | 10.05 | .17 |
| | Minimum | 8.03 | 9.05 | -4.39 |
| | Mean | 15.43 | 15.36 | .07 |
| | Maximum | 24.33 | 23.10 | 4.18 |

Risser and others— **Regression Method for Estimating Long-Term Mean Annual Recharge Rates from Base Flow**—Scientific Investigations Report 2008-5185