In cooperation with the West Virginia Department of Environmental Protection, Division of Water and Waste Management and the West Virginia Department of Transportation, Division of Highways

Basin Characteristics for Selected Streamflow-Gaging Stations In and Near West Virginia

Open-File Report 2008-1087
Version 1.1, July 2021

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

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

U.S. Geological Survey
Mark Myers, Director

U.S. Geological Survey, Charleston, West Virginia 2008
First release: 2008
Revised: July 2021 (ver 1.1)

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:
Paybins, K.S., 2008, Basin characteristics for selected streamflow-gaging stations in and near West Virginia (ver. 1.1, July 2021): U.S. Geological Survey Open-File Report 2008–1087, 9 p., https://doi.org/10.3133/ofr20081087.

Contents

Figures

Tables

Conversion Factors

Multiply	By	To obtain
inch (in.)	25.4	millimeter (mm)
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
square mile (mi^2)	2.590	square kilometer (km^2)
acre	0.004047	square kilometer (km^2)

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

°F=(1.8×°C)+32

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

°C=(°F-32)/1.8

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

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

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

Basin Characteristics for Selected Streamflow-Gaging Stations In and Near West Virginia

By Katherine S. Paybins

Abstract

Basin characteristics have long been used to develop equations describing streamflow. In the past, flow equations used in West Virginia were based on a few hand-calculated basin characteristics. More recently, the use of a Geographic Information System (GIS) to generate basin characteristics from existing datasets has refined the process for developing equations to describe flow values in the Mountain State. These basin characteristics are described in this document for streamflow-gaging stations in and near West Virginia. The GIS program developed in ArcGIS Workstation by Environmental Systems Research Institute (ESRI®) used data that included National Elevation Dataset (NED) at 1:24,000 scale, climate data from the National Oceanic and Atmospheric Agency (NOAA), streamlines from the National Hydrologic Dataset (NHD), and LandSat-based land-cover data (NLCD) for the period 1999-2003. Full automation of data generation was not achieved due to some inaccuracies in the elevation dataset, as well as inaccuracies in the streamflow-gage locations retrieved from the National Water Information System (NWIS). A Pearson's correlation examination of the data indicates that several of the basin characteristics are correlated with drainage area. However, the GIS-generated data provide a consistent and documented set of basin characteristics for resource managers and researchers to use.

Introduction

Basin characteristics are descriptors of watersheds bounded by topographical limits. The descriptors include topographic, hydrologic, land-use, and climatic variables. The topographic limits of a basin are maximum elevations bounding an area where the minimum elevation is located at the basin outlet. The topographical limits describe the contributing drainage area, which is the area that contributes to surface runoff. A basin can have ground water entering and exiting the topographical boundary, but for the purposes of this study, the contributing area is assumed to be equal to the area within the basin topographical boundary.

Basin characteristics are useful for statistical analysis of hydrologic quantities. They are typically used for determining correlations and equations for investigations of streamflow, geomorphology, and water quality. Basin characteristics historically have been determined from measurements taken from physical maps. The study was conducted by the U.S. Geological Survey (USGS) West Virginia Water Science Center, in cooperation with the West Virginia Department of Environmental Protection, Division of Water and Waste Management, and the West Virginia Department of Transportation, Division of Highways. The study employs a geographic information system (GIS) to measure and compile basin characteristics for West Virginia and surrounding areas. The data are compiled for basins upstream from USGS streamflow-gaging stations in and near West Virginia.

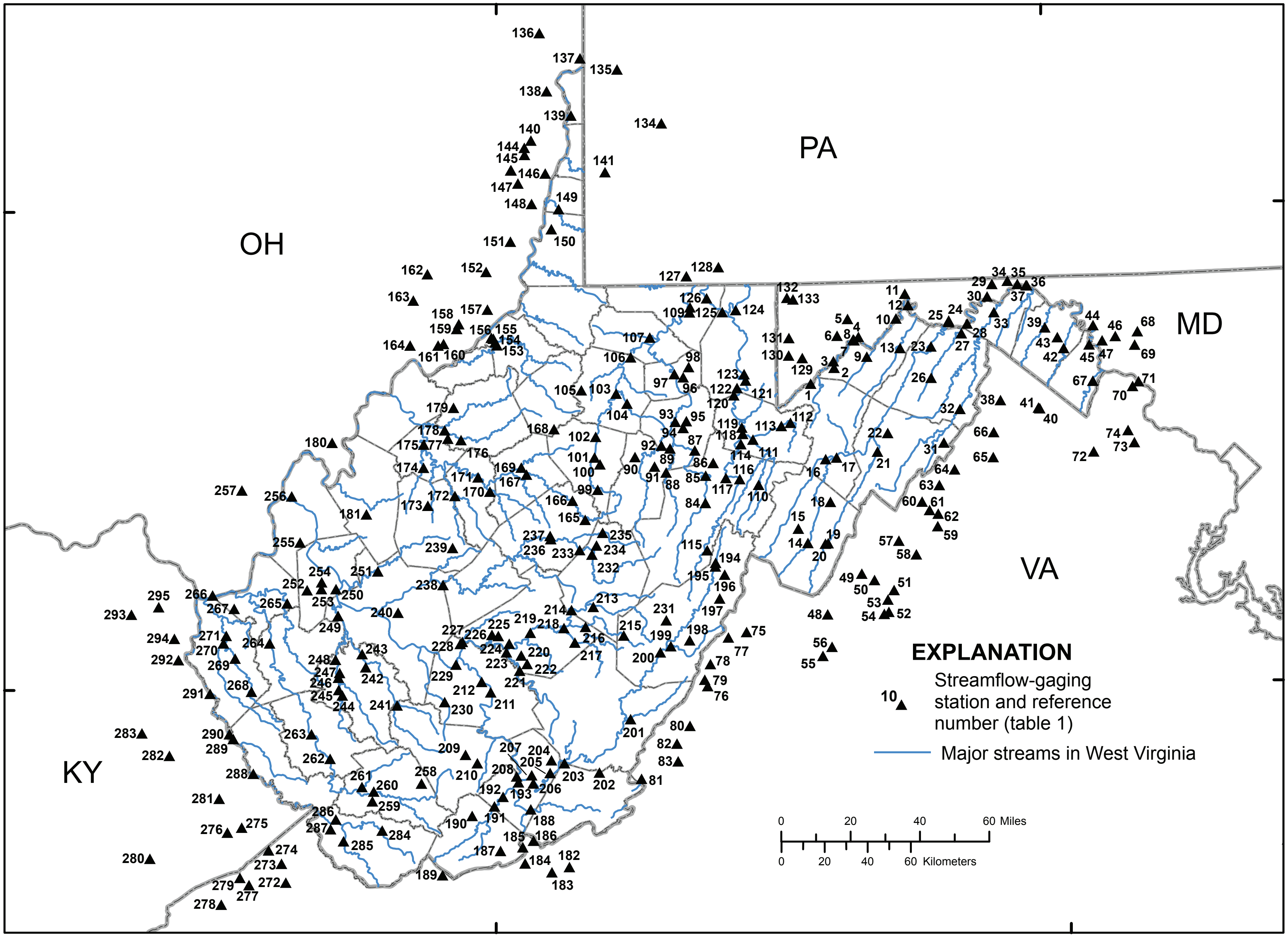
Acknowledgements

This publication would not be complete if not for the data and programming contributions of the following individuals: Jeffrey Wiley, USGS WV; Amy Downs, USGS KY; John Brakebill, USGS MD; Greg Koltun, USGS OH; Scott Hoffman, USGS PA; Jennifer Krstolic, USGS VA; Curtis Price, USGS SD; and Douglas Freehafter, USGS NY. My appreciation is offered to you all.

Basin Characteristics

Basin characteristics were determined for 295 streamflow-gaging stations in or near West Virginia (fig. 1 and table 1 [available at URL *http://pubs.water.usgs.gov/of/2008/1087/table1.csv*) by using a GIS to measure and compute values from available digital and/or geospatial datasets. Initial project design was based on using the Watershed Characterization Management System (WCMS). WCMS is a map-based web applications system developed by the National Resource Analysis Center (NRAC) at West Virginia University for the West Virginia Division of Environmental Protection (WVDEP) (Strager, 2005). WCMS is similar to the USGS "StreamStats" program (Ries and others, 2004), and is used by State agencies for management of the natural resources of West Virginia. However, WCMS does not currently calculate some of the basin characteristics needed for this study.

Equations for predicting surface-water flows for low and high flow conditions will be incorporated in the WCMS program in the future. In order to generate the basin data that will inform these statistical analyses for the 295 sites, Arc Macro



Streamlines and state boundaries from the National Atlas of the United States (2005*a*, 2005*b*); county lines from West Virginia Department of Environmental Protection (2005).

Figure 1. Streamflow-gaging stations where basin characteristics were determined in and near West Virginia.

Language (AML) programming was used in Arc Workstation 9.2 by Environmental Systems Research Institute (ESRI). The elevation and land cover data sources of WCMS and this project were of differing time periods, but the same horizontal scale. Additionally, the elevation and land-cover data used in this project extend beyond the borders of WV.

Basin boundaries were generated from USGS National Elevation Dataset (NED) 30-meter data, and these basin boundaries enclosed the focus areas for which all the other data were generated. Basin boundaries were obtained for all the sites using several methods. Multiple, nested drainage basins for several sites were generated using BASINS, an ArcView 3.0 program made for use with surface-water and water-quality models by the EPA (U.S. Environmental Protection Agency, 2006). WCMS was used to generate some drainage basin boundaries within West Virginia. USGS Water Science Centers in Kentucky, Maryland, Ohio, Pennsylvania, and Virginia contributed drainage basin boundaries for sites within their states (Amy Downs, John Brakebill, Greg Koltun, Scott Hoffman, U.S. Geological Survey, personal communications; and Krstolic and Hayes, 2006). Basin characteristics generated for sites in VA, OH, KY, PA, and MD for this study do not supersede any basin information originating from the listed sources.

All basin boundaries were checked with respect to USGS 1:24,000 topographic maps and National Hydrologic Dataset (NHD) at the 1:24,000 scale to verify relative accuracy. Several digital drainage areas were edited in cases where polygons were excluded from a drainage area that was determined to be contributing flow. Additionally, stream-gage locations were relocated manually in the GIS to enable the automation of computing the basin characteristics. That is, latitude and longitude of the basin outlet (table 1) is reflective of the streamflow-gage location, and not the lowest elevation within the boundary of the drainage area, or the basin outlet.

Basin characteristics for each drainage basin were generated by the GIS using data layers developed from existing datasets downloaded from a variety of websites serving earth observation data, and some published values. These characteristics include elevation, land use, and stream density. The NED was used to identify or calculate several characteristics (U.S. Geological Survey, 2007). NED data are developed by merging the best available elevation datasets at the 1:24,000 scale, and at the 1:63,360 scale for Alaska, into a single projection and datum for the nation, and updating the dataset bimonthly. The NED process includes smoothing the artifacts found in DEMs, such as banding, which are generated by the process of DEM development, but streamlines are not burned into the final NED dataset.

The National Land Cover Dataset 2001 (NLCD) data were used for land-use characteristics (Multiple Resolution Land Characteristics Consortium, 2007). Several climate-related datasets were available from the National Oceanographic and Atmospheric Administration (NOAA), and from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) Group at Oregon State University, which generates gridded data from point locations representing NOAA climate stations (PRISM Group, 2007; National Oceanic and Atmospheric Association, 2007a and 2007b). The NHD at the 24,000 scale was used to calculate some basin characteristics, such as total stream length inside each basin boundary (U.S. Geological Survey, 2006).

Characteristics Generated from NWIS Surface-Water Station Data, NED Data, and Basin Boundaries

Drainage area (*DA*).—Drainage areas, in mi^2, are published values for surface drainage area for each basin in the study. Drainage areas range from 0.04 mi^2 to 9,651 mi^2. The median drainage area size is 93.6 mi^2.

Latitude of basin outlet (LAT_O).—Latitude of streamflow-gaging station, in decimal degrees, as published in the National Water Information System (NWIS) database.

Longitude of basin outlet ($LONG_O$).—Longitude of streamflow-gaging station, in decimal degrees, as published in the National Water Information System (NWIS) database.

DA, LAT_O, and $LONG_O$ were available from published values for stations in West Virginia (Mathes, 1977; Wilson, 1979; Mathes and others, 1982; Preston and Mathes, 1984; Stewart and Mathes, 1995; Wiley, 1997; Wiley and others, 2006) and were available from NWIS (Ward and Crosby, 2006) for stations near West Virginia.

There were two stations that were identified during this study with incorrect drainage areas or USGS station identification numbers in NWIS- Reeds Creek Tributary near Franklin, and Rich Creek near Peterstown. Reeds Creek Tributary near Franklin (station number 01605700) was previously published as being at 38°41’47” Latitude 79°23’55” Longitude and with a drainage area of 0.45 mi^2 (Wiley and others, 2000; U.S. Geological Survey, 1996). The correct location is 38°41’52” Latitude 79°24’18” Longitude and the drainage area is 0.23 mi^2. Historic basin characteristics for 01605700 published by Wiley and others (2000) were accurate, but are superseded by values published in this report.

Rich Creek near Peterstown (03176400) was previously published as station number 03177000 (Ward and Crosby, 2006). In 2007, the station number was updated to reflect a more accurate station number in comparison to other sites in the area. All other information for the streamflow-gaging station is correct, and remains the same as previously published.

Basin perimeter (BP)—Basin perimeter, in mi, is the distance measured around a basin boundary. The median basin perimeter was 62.0 mi., ranging from a low of 0.80 to 930 mi.

Basin slope (BS)—Average basin slope, in ft/mi, is measured by the “contour-band” method within the contributing drainage area, where: BS = (length of all elevation contours, in mi) (contour interval, in ft) / DA (in mi^2) (Harvey and Eash, 1996). Contour interval was set at 20 ft for all basins. Basin slope varies from 78.9 to 862 ft/mi, with a median value of 416 ft/mi.

Basin relief (BR)—Basin relief, in ft, is the difference between the highest and lowest elevation. Basin relief varies from 135 to 5,104 ft, with a median of 1,443 ft.

Basin Orientation (BOr)—Basin orientation, in degrees, is defined as an angle between the X axis and the major axis of a flow direction grid for each basin boundary (see below for more information on flow direction). *BOr* is thus an average of the orientation of flow direction in a basin, and not a true basin azimuth. The values of the orientation angle increase counter-clockwise, starting from 0 in the east (horizontal, to the right) and going through 90 when the major axis is vertical. Using this method, *BOr* can range from 0 to 180 degrees. This data-set produced basin orientation values of 2.08 to 178 degrees, with a median value of 75.3 degrees.

Channel length (CL)—Channel length, in mi, is the distance measured along a line from the basin outlet following the main channel of the stream to the farthest point on the basin divide if the stream channel were extended. Synthetic streamlines were generated for each basin from a flow-accumulation grid created from a flow-direction grid, and based on elevation data that was preprocessed to remove elevation sinks. Flow-direction grids are based on the DEM for a given area. Each cell in the resulting flow-direction grid has a value indicating the direction to which surface water would flow. The grid cells in a flow-accumulation grid are a reference to the flow-direction grid. Each value of flow accumulation is a count of how many upstream grid cells are pointing flow in the direction of each grid cell. The process of creating a synthetic stream network from the flow-accumulation grid assumed that a minimum area of at least 15 grid cells or more was contributing to a single grid cell for basins less than 5 mi^2 in area, and that a minimum area of at least 100 cells was contributing to a single cell for basins equal to or greater than 5 mi^2.

Synthetic streamlines were then transformed into a line coverage, from which the main channel line for the synthetic stream was selected. These streamlines were also used to edit the basin outlet point data, to ensure smooth processing of the basin characteristics program. Synthetic streams were generally fairly good digital representations of streams in this study, when compared to the NHD. But commonly there were many more synthetic streams than digital streamlines in the NHD.

Two grids were then generated, from the flow-direction grid, which measured the length of the longest flow path from the outlet of a basin to the basin boundary; one looking from the basin outlet to the basin divide and the other from the basin divide to the basin outlet. These 2 grids were used to measure the main channel of the basin, from the basin outlet point to the point farthest upstream along the synthetic streams, and measure the distance from that farthest upstream point to the basin divide. Both values were used to calculate the total length for this distance (*CL*).

The main channel location was chosen from the two grids based on the basin outlet and the basin-divide point by analyzing the two flow-direction grids, isolating the maximum number of contributing cells for the upstream grid, and the minimum number of contributing grid cells in the downstream grid to identify a single point for the upstream limit of the synthetic streamline nodes. In many cases, the gage location was moved slightly to a site at the start of the synthetic streamline that contained links to the main channel stream sections to facilitate the automating the computation process. Some basins in this study contained errors in the synthetic streamlines that affected the choice of the channel length. These synthetic streamlines were edited, and the program rerun to generate *CL*. Channel length varied from 0.27 to 353 mi, with a median of 24.5 mi.

Valley length (*VL*)—Valley length, in mi, is the distance measured from the basin outlet along a line centered within the basin to the farthest point on the basin divide if the stream channel were extended. It was generated from the main-channel length data, using the variable of weed tolerance within ArcGIS Workstation. Weed tolerance is the minimum distance allowed between points in a line that represents the main channel. Weed tolerance for this characteristic was calculated as (*DA* / *CL*) x 500. Using this method, most valley length lines contained at least 5 points. Valley length varied from 0.21 mi to 178 mi, with a median of 17.4 mi.

Channel slope (CS) — Channel slope, in ft/mi, was calculated by determining the elevation at 10 and 85 percent of the distance along the main-channel length (*CL*), and then dividing the change in elevation between these two points by the length of that main-channel line segment connecting the two points. Channel slope varied from 0.50 to 1,338 ft/mi, with a median of 30.7 ft/mi.

Stream length (SL)—Stream length, in mi, is the sum of all stream segments in a watershed. Each basin boundary was used to clip the 1:24,000-scale NHD dataset. The NHD data are an enhanced digital representation of the blue lines for streams on the 1:24,000 scale topographic maps produced by the USGS. The NHD coverage used does not differentiate between intermittent and perennial streams. Total stream length varied from 0.27 to 11,275 mi, with a median of 146 mi. The NHD data contained no mapped streams for 22 basins.

Elevation (E)—Elevation, in ft above sea level, is the mean watershed altitude. It is based on the mean elevation of the NED grid before smoothing the data to remove spikes and sinks, and clipped to the basin extent. Elevation varied from 505 to 4,195 ft above sea level, with a median value of 2,008 ft above sea level.

Latitude of the basin centroid (LAT_C)—Latitude of the watershed centroid, measured in decimal degrees.

Longitude of the basin centroid ($LONG_C$)—Longitude of the watershed centroid, measured in decimal degrees.

Climatic characteristics

24-hour 2-year rainfall (I24-2)—Rainfall for the 24-hour 2-year recurrence interval, in in., was available from the National Oceanographic Atmospheric Administration (NOAA) (National Oceanographic Atmospheric Administration, 2007b). The rainfall frequency was determined by averaging all values from a data grid clipped to the boundary of each basin, and ranged from 2.36 in. to 3.42 in., with a median value of 2.63 in.

Precipitation (P)—Precipitation, in in., is the mean annual precipitation falling on the basin, averaged over the whole basin. Mean annual precipitation data were acquired from the PRISM Group at Oregon State University for the 30 year period 1971-2000 (PRISM Group, 2007).The data were acquired in Arc Grid format and are reported in hundredths of millimeters for each grid cell. Average annual precipitation data were calculated for each basin by clipping the precipitation data to the basin boundary, taking an average for all grid cells in the basin, and computing the basin average in in. Values for mean annual precipitation ranged from 35.8 to 61.7 in., with a median of 45.0 in.

January minimum temperature (JANMIN)—January minimum temperature, in degrees Fahrenheit (ºF), was computed as an average from a grid of minimum January temperature for each basin. The data were acquired from the PRISM Group for the 30 year period 1971-2000 (PRISM Group, 2007). *JANMIN* ranged from 14.0º F, to 24.0º F, with a median value of 20.0º F.

Snow (S)—Snow, in in., is the measure of the mean annual snow falling on the basin (averaged over the basin), based on a grid of data generated from climate record stations in the conterminous U.S. (National Oceanographic and Atmospheric Association, 2000a). The data are from the 30 year period 1961-1990. The data are presented in shapefile format for download, and each polygon is classified as a range of values. To create a grid, the median of each class of data was used as the single value for calculation, except in the case of the areas that receive more than 72 in. of snowfall a year. In those cases, the single value for computation was assumed to be 75 in.. Each basin boundary was used to clip the snow data, and then area-weighted mean annual snowfall for all values in the basin snow grid was computed. The mean annual snow for the data ranged from 17.9 to 72.7 in.; median *S* for all basins is 32.9 in..

Land cover characteristics

Land cover data were generated from the National Land Cover Dataset 2001 (NLCD), which is based on LandSat 7 data from the period 1997-2003 (Multiple Resolution Land Characteristics Consortium, 2007). The NLCD data has 30-meter grid cells. Percent of a given land use is calculated based on the number of grid cells in a given basin, multiplied by the cell size, and that product is divided by the drainage area of the basin to compute the percent of land cover. Because of this, the statistics presented below may sometimes, in a given basin, add up to slightly more or slightly less than 100 percent of total basin area. The land-cover data were also generalized from 21 to 7 groups as follows:

Forest (F)—Forest, in percent, is the measure of the basin covered by all types of forest, including deciduous, coniferous, and mixed forest types. Most West Virginia basins have a high percent of land area covered in forest. But for basins outside of West Virginia, more of the drainage area was covered by other land uses. Thus, the values range from 3.13 to 100 percent, with a median of 82.4 percent.

Grassland (G)—Grassland, in percent, is the measure of the basin covered by grassland or pasture. The values range from 0.00 to 67.0 percent, with a median of 9.20 percent.

Barren land (B)—Barren land, in percent, is the measure of the basin having barren land. Barren land ranges from 0.00 to 23.7 percent, with a median of 0.13 percent.

Urban land (U)—Urban land, in percent, is the measure of the basin covered by urban development. Urban land varies from 0.00 to 78.6 percent, with a median of 5.67 percent.

Wetland (W)—Wetland, in percent, is the measure of that portion of the basin covered by emergent and herbaceous wetlands. Wetland varies from 0.00 to 16.7 percent, with a median of 0.00 percent.

Open water (Wa)—Open water, in percent, is the measure of the basin covered by lakes, reservoirs, large rivers, and swamps. Open water varies from 0.00 to 4.72 percent, with a median of 0.16 percent.

Agricultural (A)—Agricultural land, in percent, is the measure of the basin in agricultural production. Agricultural land varies from 0.00 to 39.2, with a median of 0.57 percent.

Impervious (I)—Impervious area, in percent, includes areas such as parking lots and rock outcrops. Impervious area is a calculated value dependant upon weighted percents of each basin covered by each of the land uses, with one variant in classification. Urban land cover was split into low-resolution residential, high-resolution residential, and high-intensity commercial. Open water and wetlands are assumed to be entirely pervious, and are therefore always equal to 0.00.

I = (percent low-intensity residential x 25) + (percent high-intensity residential x 45) + (percent high-intensity commercial x 60) + (percent agricultural land x 1.5) + (percent grassland x 5) + (percent forest x 0.8) + (percent barren land x 1). Impervious area of basins ranged from 0.80 to 32.0 percent, with a median of 2.80 percent.

Computed Characteristics

The following computed basin characteristics are calculated from the GIS-generated characteristics that are documented above.

Basin width (BW)—Basin width, in mi, is computed by dividing the drainage area by the valley length (DA/*VL*). Basin width varies from 0.16 to 57.82 miles, with a median of 4.84 miles.

Shape factor (SF)—Shape factor, dimensionless, is computed by dividing valley length by basin width (*VL*/*BW*). Shape factor in this dataset varies from 1.04 to 18.9, with a median of 3.40.

Elongation ratio (ER)—Elongation Ratio, dimensionless, is computed by dividing the diameter of a circle with an area equal to that of the watershed by the valley length, or 1.13

x $(1/SF)^{0.5}$. The elongation ratio for the basins in this study ranges from 0.26 to 1.11, with a median of 0.61.

Rotundity of basin (RB)—Rotundity of basin, dimensionless, is computed as $\pi\,(VL)^2 / 4DA$, or 0.785*SF*. Rotundity of basin ranged from 0.82 to 14.8, with a median of 2.67.

Compactness ratio (CR)—Compactness ratio, dimensionless, is computed by dividing the basin perimeter by the circumference of a circle of equal area ($BP / 2(\pi\, DA)^{0.5}$). It represents the ratio of the perimeter of the basin to the circumference of a circle of equal area. Compactness ratio ranges from 0.14 – 75,367, with a median of 544.

Relative relief (RR)—Relative relief, in ft/mi, is computed by dividing the basin relief by the basin perimeter (*BR*/*BP*). Relative relief varies from 3.48 to 545 ft/mi, with a median of 28.1 ft/mi.

Sinuosity ratio (SR)—Sinuosity ratio, dimensionless, is computed by dividing the channel length by valley length (*CL*/*VL*). It is an approximate measurement of the degree to which a stream channel meanders within its valley. Sinuosity varies from 0.82 to 2.03, with a median of 1.34.

Stream density (SD)—Stream density, in mi/mi^2, is computed by dividing the total length of all streams in a basin by the drainage area (*SL*/*DA*). There were no streamlines contained in the NHD dataset for 22 smaller basins. For all other basins, the stream density varies from 0.08 to 2.37 mi/mi^2, with a median of 1.30 mi/mi^2.

Constant of channel maintenance (CM)—The constant of channel maintenance, in mi, is computed by dividing the drainage area by stream length (*DA*/*SL*). For 22 smaller basins, this value cannot be calculated due to the lack of NHD streams in the basins. For all other basins, the values range from 0.42 to 12.7 mi, with a median value of 0.77 mi.

Slope proportion (SP)—Slope proportion, dimensionless, is computed as $CL / (CS)^{0.5}$. Values range from 0.02 to 384, with a median of 4.91.

Ruggedness number (RN)—Ruggedness number, in ft/mi, is computed by dividing the product of stream length and basin relief by drainage area ((*SLBR*)/*DA*), or computed as the product of the stream density and the basin relief (*SDBR*). Values range from 114 to 6618 ft./mi., with a median of 1,961 ft./mi..

Slope ratio (SR)—Slope ratio, dimensionless, is computed by dividing the channel slope by the basin slope (*CS*/*BS*). Slope ratios varied from 0.002 to 2.16, with a median value of 0.08.

Uses and limitations to the basin characteristics data

The basin characteristics dataset presented in this study has a nominal scale of 1:24,000, using the best available data for West Virginia and surrounding states. However, in this study, there were problems with accuracy related to scale and/or vertical accuracy for station locations, elevation data, and streamlines. Basin characteristics are used in statistical analyses for water resources applications that assume that the variables are independent, but in fact, there are some significant relations among variables.

Streamflow-gaging station locations are often mapped to the side of a stream, because that is where the gage house is located. Additionally, older NWIS records of streamflow-gaging station locations were hand-calculated from 1:24,000 or smaller scale maps. When the station location is an equal distance between 2 or more digital streams, the AML program may snap the station to the wrong streamline. Additionally, if the station is farther away from the proper stream than another, the station may be snapped to the wrong streamline. Finally, some stations were mapped just outside the basin boundary, and the AML program used analyzes data only within the basin boundary. Careful attention must be paid to these issues to produce accurate results.

The elevation and streamline datasets used in this project needed to be common to West Virginia and surrounding states. However, other datasets exist that are of a larger scale, higher resolution, and higher accuracy level; but these are limited to West Virginia, including 1:4,800 NED data (U.S. Geological Survey and State Addressing and Mapping Board, 2003). Light Detection and Ranging (LiDAR) elevation data also exists for some counties and communities located within the study area. The use of these datasets would likely improve the automation process and the accuracy of the output.

Basin characteristics are unique to each basin area, but some characteristics within a basin are highly correlated. This may be problematic when using these data in multiple regression analyses. Specifically, two or more highly correlated independent basin characteristics essentially explain the same variation in a dependant variable, such as streamflow (Blalock, 1972). The degree to which the independent basin characteristics are correlated will cause partial correlations and slope estimates to be increasingly sensitive to sampling and measurement errors.

In order to assess the degree of correlation, a Pearson correlation coefficient analysis, or Pearson's r, was used to compare the basin characteristic data (table 2). Correlations were considered high where the absolute value of Pearson's r was greater than or equal to 0.90. In this dataset, the correlation analysis shows that drainage area (*DA*) is significantly positively correlated with *BP*, *CL*, *SL*, *BW*, *CR*, and *SP*. Basin perimeter (*BP*) is also correlated with *CL*, *VL*, *SL*, *BW*, and *CR*. Channel length (*CL*) is correlated to *VL*, *WL*, and *BW*, in addition to *DA* and *BP*. Valley length (*VL*) is correlated with *BW*, in addition to *DA*, *BP*, and *BW*. Stream length (*SL*) is also correlated with *BW* and *CR*. Channel slope is correlated with relative relief (*RR*). Basin relief (*BR*) is correlated highly with basin ruggedness (*RN*). Percent impervious land cover *(I)* is correlated with percent urban land cover (*U*). Finally, shape factor (*SF*) is a factor used to calculate rotundity of the basin (*RB*), and thus there is perfect correlation (r = 1) between these two basin characteristics, and the r for both are the same in

Table 2. Pearson's coefficients (r) for relations among basin characteristics for streamflow-gaging stations in West Virginia and adjacent areas of Virginia, Maryland, Ohio, Pennsylvania, and Kentucky (significant levels < 0.0001 for all r values greater than 0.89 are shown in bold).

[number of sites used in this Pearson's r test is 295, except for stream length and ruggedness number, which have a sample size of 273]

NAME	Drainage area	Basin perimeter	Basin relief	Channel length	Valley length	Channel slope	Stream length	Urban land
Drainage area	1.00	**0.92**	0.60	0.90	0.87	-0.18	1.00	-0.03
Basin perimeter	**0.92**	1.00	0.74	**0.98**	**0.97**	-0.30	**0.93**	-0.08
Basin relief	0.60	0.74	1.00	0.77	0.81	-0.26	0.60	-0.22
Channel length	**0.90**	**0.98**	0.77	1.00	**0.99**	-0.32	**0.90**	-0.08
Valley length	0.87	**0.97**	0.81	**0.99**	1.00	-0.35	0.88	-0.10
Channel slope	-0.18	-0.30	-0.26	-0.32	-0.35	1.00	-0.18	-0.09
Stream length	**1.00**	**0.93**	0.60	**0.90**	0.88	-0.18	1.00	-0.03
Urban land	-0.03	-0.08	-0.22	-0.08	-0.10	-0.09	-0.03	1.00
Impervious land	-0.01	-0.06	-0.21	-0.06	-0.08	-0.12	-0.01	**0.98**
Basin width	**0.90**	**0.93**	0.72	0.92	**0.92**	-0.35	**0.91**	-0.07
Shape factor	0.08	0.23	0.41	0.29	0.35	-0.19	0.07	-0.12
Rotundity of basins	0.08	0.23	0.41	0.29	0.35	-0.19	0.07	-0.12
Compactness ratio	**0.98**	**0.93**	0.58	**0.89**	0.86	-0.17	**0.99**	-0.03
Relative relief	-0.22	-0.36	-0.29	-0.37	-0.40	0.92	-0.24	-0.08
Slope proportion	**0.90**	0.81	0.54	0.83	0.82	-0.22	0.87	-0.03
Ruggedness number	0.62	0.74	0.94	0.77	0.79	-0.21	0.62	-0.22

NAME	Impervious land	Basin width	Shape factor	Rotundity of basin	Compactness ratio	Relative relief	Slope proportion	Ruggedness number
Drainage area	-0.01	**0.90**	0.08	0.08	0.98	-0.22	**0.90**	0.62
Basin perimeter	-0.06	**0.93**	0.23	0.23	**0.93**	-0.36	0.81	0.74
Basin relief	-0.21	0.72	0.41	0.41	0.58	-0.29	0.54	**0.94**
Channel length	-0.06	**0.92**	0.29	0.29	**0.89**	-0.37	0.83	0.77
Valley length	-0.08	**0.92**	0.35	0.35	0.86	-0.40	0.82	0.79
Channel slope	-0.12	-0.35	-0.19	-0.19	-0.17	**0.92**	-0.22	-0.21
Stream length	-0.01	**0.91**	0.07	0.07	0.99	-0.24	0.87	0.62
Urban land	**0.98**	-0.07	-0.12	-0.12	-0.03	-0.08	-0.03	-0.22
Impervious land	1.00	-0.05	-0.11	-0.11	-0.01	-0.11	0.00	-0.22
Basin width	-0.05	1.00	0.06	0.06	0.87	-0.41	0.84	0.72
Shape factor	-0.11	0.06	1.00	**1.00**	0.10	-0.18	0.14	0.35
Rotundity of basins	-0.11	0.06	**1.00**	1.00	0.10	-0.18	0.14	0.35
Compactness ratio	-0.01	0.87	0.10	0.10	1.00	-0.20	0.83	0.60
Relative relief	-0.11	-0.41	-0.18	-0.18	-0.20	1.00	-0.25	-0.24
Slope proportion	0.00	0.84	0.14	0.14	0.83	-0.25	1.00	0.52
Ruggedness number	-0.22	0.72	0.35	0.35	0.60	-0.24	0.52	1.00

relation to all other basin characteristics. Stream density (*SD*) is significantly positively correlated with *R* and *SF*.

Summary

Basin characteristics were automatically generated for sites in West Virginia and adjacent surrounding areas using digital data and GIS software. Basin characteristics were generated from digital data available for elevation, land cover, stream centerlines, and climate. The nominal scale of the available data for West Virginia and the surrounding states is 1:24,000 for NED, NHD, and NLCD data. The climatic data is based on the NED 30-meter grid data for extrapolating from point data (climate stations) to gridded data (PRISM Group, 2007). Some secondary basin characteristics were calculated from the data genereated directly from the GIS datasets, such as rotundity of basin, or stream density in a drainage basin.

All basin characteristics generated for this report are based on GIS data at the nominal scale of 1:24,000. Larger scale and/or higher resolution and accuracy datasets will increase the accuracy of basin characteristics and of the automated GIS programming, but these data were not available during this study for the entire region of concern. There were a few problems with the data automation that were related to gage locations, the elevation data, and streamline data. Several basin characteristics are highly correlated, suggesting some limitations for use in multiple regression analyses. Drainage area (*DA*) is highly correlated with basin perimeter (*BP*), main channel length *(CL)*, total stream length (*SL*), basin width (*BW*), the compactness ratio (*CR*), and the main channel length slope ratio (*SR*). But these data remain valuable to researchers and resource managers for developing statistical equations describing a variety of streamflow conditions.

References Cited

Blalock, Jr., H.M., 1972, Social Statistics, 2nd Edition: McGraw-Hill Book Company, Series in Sociology, 583 p.

Harvey, C.A., and Eash, D.A., 1996, Description, Instructions, and Verification for Basinsoft, a Computer Program to Quantify Drainage-Basin Characteristics: U. S. Geological Survey Water Resources Investigations Report 95-4187, 25 p.

Krstolic, J, and Hayes, D.C., 2006, Drainage Basin Delineations for selected USGS streamflow-gaging stations in Virginia: U.S. Geological Survey Open File Report 2006-1308, online publication found at *http://water.usgs.gov/GIS/metadata/usgswrd/XML/OFR2006-1308_Drainage_Basin.xml* .

Mathes, M.V. Jr., 1977, Drainage areas of the Guyandotte River Basin, West Virginia: U.S. Geological Survey Open-File Report 77-801, 56 p.

Mathes, M.V., Kirby, J.R., Payne, D.D., and Schultz, R.A., 1982, Drainage areas of the Kanawha River Basin, West Virginia: U.S. Geological Survey Open-File Report 82-351 222 p.

Multiple Resolution Land Characteristics Consortium, 2007, National Land Cover Dataset, accessed online at *http://www.mrlc.gov/mrlc2k_nlcd.asp* on September 20, 2007.

National Atlas of the United States, 2005a, Streams and Waterbodies of the United States: National Atlas of the United States, Reston, VA; accessed online at *http://nationalatlas.gov/atlasftp.html* .

National Atlas of the United States, 2005b, State Boundaries of the United States: National Atlas of the United States, Reston, VA; accessed online at *http://nationalatlas.gov/atlasftp.html* .

National Oceanographic and Atmospheric Administration, 2007a, National Climatic Data Center Annual snow average for the lower-48 U.S., 1961 to 1990; accessed January 19, 2008 online at *http://hurricane.ncdc.noaa.gov/cgi-bin/climaps/climaps.pl* .

National Oceanographic Atmospheric Administration, 2007b, National Weather Service Hydrometerological Design Studies Center, Precipitation frequency data server, *http://hdsc.nws.noaa.gov/hdsc/pfds/orb/pfds_gis.html* accessed January 1, 2007.

Preston, J.S., and Mathes, M.V., 1984, Stream drainage areas for the Little Kanawha River Basin, West Virginia: U.S. Geological Survey Open-File Report 84-861, 171 p.

PRISM Group, 2007, Oregon State University, Mean annual snowfall, Average minimum temperature in January, and Mean annual precipitation from 1971 to 2000, accessed online at *http://prism.oregonstate.edu/products/matrix.phtml* on June 14, 2007.

Ries, K.G., III, Steeves, P.A., Coles, J.D., Rea, A.H., and Stewart, D.W., 2004, StreamStats: A U.S. Geological Survey web application for stream information: U.S. Geological Survey Fact Sheet 2004-3115, 4 p.

Stewart, D.K., and Mathes, M.V., 1995, Drainage areas of the Monongahela River Basin, West Virginia: U.S. Geological Survey Open-File Report 95-170, 79 p.

Strager, Michael P., 2005. Watershed Characterization and Modeling System Version 9.0 Technical Documentation: accessed online at *http://www.nrac.wvu.edu/projects/nrac244/NRAC244Technical_Documentation_07-05.pdf*

U.S. Environmental Protection Agency, 2006, Better Assessment Science Integrating Point and Non-Point Sources, accessed online at *http://www.epa.gov/waterscience/BASINS*, on June 6, 2006.

U.S. Geological Survey, 1996, National Water Information System, *http://waterdata.usgs.gov/nwis*, accessed October 27, 1996.

U.S. Geological Survey, 2006, National hydrologic dataset, accessed online at *http://nhdgeo.usgs.gov/viewer.htm* , on September 12, 2006.

U.S. Geological Survey, 2007, National Elevation Dataset, accessed online at *http://seamless.usgs.gov* , on March 20, 2007.

U.S. Geological Survey and State Addressing and Mapping Board, 2003, West Virginia Statewide Digital Elevation Models, available at *http://seamless.usgs.gov* .

Ward, S.M., and Crosby, G.R., 2006, Water Resources Data for West Virginia Water Year 2006: U. S. Geological Survey Water Data Report WV-06, 278 p.

West Virginia Department of Environmental Protection, 2005, County boundaries from 1:24,000 digital raster graphics of the USGS; accessed online at *http://wvgis.wvu.edu/data/dataset.php?action=search&ID=136* .

Wiley, J.B., 1997, Drainage areas of West Virginia streams tributary to the Ohio River: U.S. Geological Survey Open-File Report 97-231, 70 p.

Wiley, J.B., Atkins, J.T., and Tasker, G.D., 2000, Estimating magnitude and frequency of peak discharges for rural, unregulated, streams in West Virginia: U.S. Geological Survey Water-Resources Investigations Report 00-4080, 93 p.

Wiley, J.B., Hunt, M.L., and Stewart, D.K., 2006, Drainage areas of the Potomac River Basin, West Virginia: U.S. Geological Survey Open-File Report 95-292, 63 p.

Wilson, M.W., 1979, Drainage areas of the Twelvepole Creek Basin, West Virginia; Big Sandy River Basin, West Virginia; Tug Fork Basin, Virginia, Kentucky, West Virginia: U.S. Geological Survey Open-File Report 79-746, 49 p.