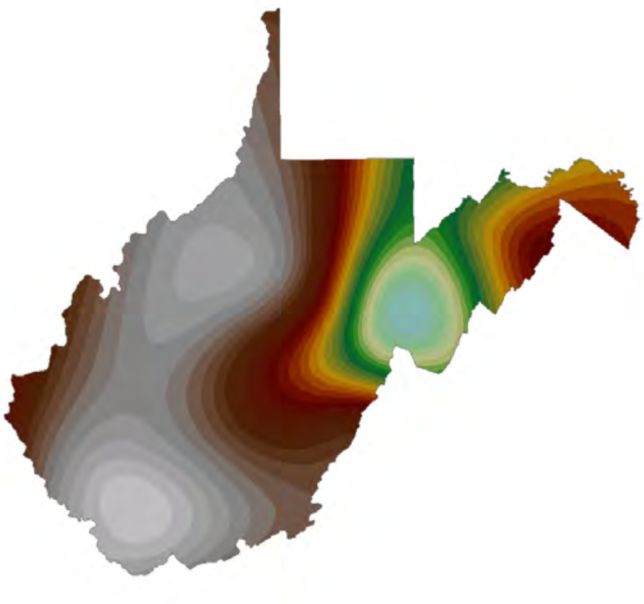
Prepared in cooperation with West Virginia Department of Transportation, Division of Highways

Generalized Skew Coefficients of Annual Peak Flows for Rural, Unregulated Streams in West Virginia



Open-File Report 2008–1304

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

This page intentionally left blank.

Generalized Skew Coefficients of Annual Peak Flows for Rural, Unregulated Streams in West Virginia

By John T. Atkins, Jr., Jeffrey B. Wiley, and Katherine S. Paybins

Prepared in cooperation with West Virginia Department of Transportation, Division of Highways

Open-File Report 2008–1304

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: 2009

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

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

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

Suggested citation:
Atkins, J.T., Jr., Wiley, J.B., and Paybins, K.S., 2009, Generalized skew coefficients of annual peak flows for rural, unregulated streams in West Virginia: U.S. Geological Survey Open-File Report 2008–1304, 13 p.

Contents

Figures

Tables

Conversion Factors and Datums

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)	2.590	square kilometer (km^2)
	Flow rate	
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)

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). Intermediate results in the North American Datum of 1927 (NAD 27) were re-projected to NAD 83 using the NADCON option of ArcView 3.3. For assistance with conversions, the reader is directed to either the National Geodetic Survey website for NADCON at *http://www.ngs.noaa.gov/TOOLS/Nadcon/Nadcon.html* or the U.S. Army Corps of Engineers website for Corpscon at *http://crunch.tec.army.mil/software/corpscon/corpscon.html*.

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

Generalized Skew Coefficients of Annual Peak Flows for Rural, Unregulated Streams in West Virginia

By John T. Atkins, Jr., Jeffrey B. Wiley, and Katherine S. Paybins

Abstract

Generalized skew was determined from analysis of records from 147 streamflow-gaging stations in or near West Virginia. The analysis followed guidelines established by the Interagency Advisory Committee on Water Data described in Bulletin 17B, except that stations having 50 or more years of record were used instead of stations with the less restrictive recommendation of 25 or more years of record. The generalized-skew analysis included contouring, averaging, and regression of station skews. The best method was considered the one with the smallest mean square error (MSE). MSE is defined as the following quantity summed and divided by the number of peaks: the square of the difference of an individual logarithm (base 10) of peak flow less the mean of all individual logarithms of peak flow. Contouring of station skews was the best method for determining generalized skew for West Virginia, with a MSE of about 0.2174. This MSE is an improvement over the MSE of about 0.3025 for the national map presented in Bulletin 17B.

Introduction

In the United States, the log-Pearson Type III distribution is widely used to calculate flood recurrences by fitting it to annual series of observed peak flows. This generally accepted method was published in Bulletin 17B by the Interagency Advisory Committee on Water Data (IACWD) (1982). Flood-frequency results are sensitive to the fit parameter, “skew,” so Bulletin 17B includes guidelines for studies of skew (Interagency Advisory Committee on Water Data, 1982). Bulletin 17B contains a nationwide map of the generalized skew that has been used to determine skews in West Virginia. However, Bulletin 17B recommends that users develop generalized skew coefficients for their area of interest using the procedures outlined in the Bulletin. Generalized skews are to be developed using a minimum of 40 streamflow-gaging stations with 25 or more years of record. This is now possible for West Virginia because sufficient data have been collected in or near the State.

In an effort to provide updated information for planning and design activities related to flood-frequency calculations, the U.S. Geological Survey (USGS), in cooperation with the West Virginia Department of Transportation, Division of Highways, conducted an investigation to develop generalized skew coefficients for West Virginia. The results of this study supercede the longstanding method in West Virginia of using the nationwide generalized skew-coefficient map in Bulletin 17B. This report (1) describes the development of a statewide generalized skew-coefficient map, (2) presents the methods used, and (3) summarizes error analyses. The new generalized skew-coefficient map is specific to magnitude and frequency analysis of peak flows in West Virginia only.

Station Skew Coefficients

Skewness is a measure of the degree of asymmetry of a distribution. As used in this report, the “coefficient of skewness,” “skew coefficient,” or simply “skew” is a parameter used to match a series of logarithms of annual peak flows recorded at a streamflow-gaging station to the log-Pearson Type III statistical distribution. Statistically, skew is defined as (1) a numerical measure or an index of the lack of symmetry in a frequency distribution, (2) a function of the third moment of magnitudes about their mean, and (3) the adjusted third moment divided by the cube of the standard deviation. The skew coefficient for a single streamflow-gaging station is termed the “station skew.”

The log-Pearson Type III distribution requires computation of the first moment (mean), the second moment (standard deviation), and the third moment (skew) of the base 10 logarithms of the station’s annual peak flows. In this case, the first, second, and third moments for a sample are computed from the following:

$$\bar{X} = (\Sigma X) / N$$ (first moment, mean, eq. 2 of Bulletin 17B),

$$S = \{[\Sigma X^2 - (\Sigma X)^2 / N] / [N - 1]\}^{1/2}$$ (second moment, standard deviation, eq. 3b of Bulletin 17B),

and

$$G = [N^2 (\Sigma X^3) - 3N (\Sigma X) (\Sigma X^2) + 2 (\Sigma X)^3] / [N (N - 1) (N - 2) S^3]$$

(third moment, skew, eq. 4b of Bulletin 17B),

where

X	is the logarithm of an annual peak flow, in ft^3/s, at a station for a given year;
$\Sigma X, \Sigma X^2, \Sigma X^3$	is the sum of X, X^2, X^3 for N years, respectively;
$\bar{X}$	is the mean of the logarithms of annual peak flows, in ft^3/s;
N	is the number of peak flows;
S	is the standard deviation of the logarithms of annual peak flows, in ft^3/s; and
G	is the skew coefficient (station skew, unitless).

Generalized Skew Coefficients for West Virginia

The 88 selected streamflow-gaging stations on rural, unregulated streams in West Virginia and 59 similar and nearby stations were combined to yield 147 stations for analysis (table 1, fig. 1). Records for a subset of 75 stations having at least 50 years of record also were analyzed (filled symbols in figure 1). To obtain a consistent set of skew values for all sites, new flood-frequency computations were performed using the PeakFQ computer program, version 4.1 (February 2002) as described by Flynn and others (2006). Annual peak flows from previous compilations for the 147 stations were updated from NWISWeb, the Web interface of the U.S. Geological Survey (USGS) National Water Information System (U.S. Geological Survey, 2008). Skew values were verified by recalculating them in PeakFQ and then reconciled to published values. Where there were no published skew values, the published 100-year flood flows were reconciled (see flood reconciliation, in percent, in table 1); nine stations are flagged as having 100-year flood values with more than a 1-percent difference, and values at all but two stations are within 2.3 percent of the 100-year flood flow. In this study, skew values were computed and used for Prater Creek at Vasant, Va., and Potomac River at Point of Rocks, Md., that failed to reproduce the previously published values of 100-year flood flows. This was done so that station skew values were in all cases calculable from the available peak-flow series. Overall, reconciliation problems were fewer than expected.

Confidence in longer streamflow records is usually well justified and has resulted in the development of methods that give greater weight to longer records, such as that described by Tasker and Stedinger (1986). However, if there is a general trend in skew as a function of record length (not chronological time), shorter records might reasonably be rejected, because more confidence can be placed in longer records. A simple step test for trend (same as a two-sample test) can be obtained by using a non-parametric statistical test comparing the means of shorter and longer records.

The mean of the skews for 72 stations with less than 50 years of record (0.0615) and the mean for 75 stations with 50 or more years of record (0.2449) differ by 0.1834 (see table 1, footnote 1). This difference in the skew would make an important change in any flood-frequency calculation; therefore, its statistical significance needs verification. Can dividing the data set this way be ignored? That is, is the data set homogeneous with respect to record length?

A Wilcoxon-Mann-Whitney rank-sum test was used to answer this question. The Wilcoxon-Mann-Whitney rank-sum test statistic is the sum of the ranks for observations from the smaller of the two samples (in "Wilcoxon two-sample test" at the end of table 1). The null hypothesis is that the means are equal or that the two samples are drawn from the same population. The alternate hypothesis is that the samples have different "locations" (in the rank of values). The inputs to the test are the smallest sum of squares of the ranks of skews, its expected value, the standard deviation, and a small adjustment (table 1, footnote 2). The approximate form of the rank-sum test is considered good if the number of values is more than 20. The SAS procedure NPAR1WAY (SAS Institute Inc., 1990) was used to compute the approximate Wilcoxon two-sample test, yielding a p-value of 0.0445 (details in "Wilcoxon two-sample test" at the end of table 1), which is significant at the alpha=0.05 level. The null hypothesis is therefore rejected. In other words, there is only about a 4.45-percent chance of observing a difference in mean station skew as large as 0.1834 if the two samples were from the same population. For enhanced reliability, (1) station skews for the group of 75 stations with 50 or more years of record were used to contour the generalized skew map and for regression, but were not the only basis for computing regional averages (as explained farther on); and (2) values for the 147 stations were used for all comparisons between methods.

Bulletin 17B recommends comparison of mean square errors (MSE) between station skews and generalized skews to select the best generalization method. In this study, MSE developed by contouring, averages, and regression analysis for the 147 streamflow-gaging stations were compared to select the best generalized method. MSE is equal to the "standard error of estimate" squared, as follows:

$$MSE = \Sigma(SK_{ST} - SK_{GEN})^2 / N,$$

where

N	is the number of stations (147),
SK_{ST}	is the station skew (unitless), and
SK_{GEN}	is the generalized skew (unitless).

The nationwide generalized skew map (Interagency Advisory Committee on Water Data, 1982), which can be viewed on the Humbolt State University Web site at URL *http://www.humboldt.edu/~geology/for_download/hydrology/*

Table 1. Station skews, differences between published and new (this study) flood-frequency computations of 100-year floods (reconciliation), and results of a Wilcoxon two-sample test of station skews for 147 streamflow-gaging stations in or near West Virginia with at least 25 years of record.

[A rank with a decimal place indicates a tie with a station with the same skew value. Shading indicates station with 50 or more years of record.[1] MD, Maryland; WV, West Virginia; VA, Virginia]

Map number (fig. 1)	Station number	Station name	State	Years of systematic record	Flood reconciliation, in percent[2]	Station skew, unitless	Rank of station skew
1	01595000	North Branch Potomac River at Steyer	MD	36	0	0.400	101.5
2	01595300	Abram Creek at Oakmont	WV	27	0	0.355	96.5
3	01595500	North Br. Potomac River at Kitzmiller	MD	41	0	1.633	147
4	01596000	North Br. Potomac River at Bloomington	MD	25	0	0.378	99
5	01596500	Savage River near Barton	MD	42	0	0.610	119
6	01597000	Crabtree Creek near Swanton	MD	33	0	0.925	134
7	01599000	Georges Creek at Franklin	MD	61	0	0.492	111
8	01601500	Wills Creek near Cumberland	MD	61	0	0.983	137
9	01603000	N. Br. Potomac River near Cumberland	MD	52	0	0.823	129
10	01604500	Patterson Creek near Headsville	WV	59	0	-0.264	33
11	01605500	South Branch Potomac River at Franklin	WV	51	0	0.604	117
12	01606000	N. Fk. S. Br. Potomac River at Cabins	WV	41	0	0.466	110
13	01606500	S. Br. Potomac River near Petersburg	WV	70	0	0.961	136
14	01607500	S. Fk. S. Br. Potomac R. at Brandywine	WV	54	0	0.777	127
15	01608000	S. Fk. S. Br. Potomac R. near Moorefield	WV	68	0	0.898	132
16	01608500	S. Br. Potomac River near Springfield	WV	74	0	0.667	122
17	01609500	Sawpit Run near Oldtown	MD	25	-0.5	-0.760	5
18	01610000	Potomac River at Paw Paw	MD	46	0.2	-0.473	18
19	01611500	Cacapon River near Great Cacapon	WV	74	0	-0.074	54.5
20	01613000	Potomac River at Hancock	MD	65	0	0.502	112
21	01613900	Hogue Creek near Hayfield	VA	31	-0.7	-1.401	1
22	01614000	Back Creek near Jones Springs	WV	46	0	-0.531	14
23	01615000	Opequon Creek near Berryville	VA	49	0.4	0.015	64
24	01616500	Opequon Creek near Martinsburg	WV	50	0	0.176	78.5
25	01617800	Marsh Run at Grimes	MD	27	0	0.584	116
26	01618000	Potomac River at Shepherdstown	MD	64	1.4	0.288	86
28	01620500	North River near Stokesville	VA	45	0	0.864	131
29	01621000	Dry River at Rawley Springs	VA	31	0	1.365	145
30	01621200	War Branch near Hinton	VA	28	-0.4	0.392	100
31	01622000	North River near Burketown	VA	63	-0.8	0.945	135
32	01632000	N. Fk. Shenandoah R. at Cootes Store	VA	69	-0.9	0.087	68
33	01632900	Smith Creek near New Market	VA	32	-0.7	-0.494	16
34	01633000	N. Fk. Shenandoah R. at Mount Jackson	VA	48	0	-0.108	50
35	01633500	Stony Creek at Columbia Furnace	VA	32	-2.1	0.073	67
36	01634000	N. Fk. Shenandoah River near Strasburg	VA	65	-0.2	0.461	108
37	01634500	Cedar Creek near Winchester	VA	55	-0.6	-0.074	54.5
38	01636500	Shenandoah River at Millville	WV	80	0	0.146	76
39	01637000	Little Catoctin Creek at Harmony	MD	30	0	0.355	96.5
40	01637500	Catoctin Creek near Middletown	MD	43	0.2	0.664	121
41	01638500	Potomac River at Point Of Rocks	MD	96	9.8	0.187	80

Table 1. Station skews, differences between published and new (this study) flood-frequency computations of 100-year floods (reconciliation), and results of a Wilcoxon two-sample test of station skews for 147 streamflow-gaging stations in or near West Virginia with at least 25 years of record.—Continued

[A rank with a decimal place indicates a tie with a station with the same skew value. Shading indicates station with 50 or more years of record.[1] MD, Maryland; WV, West Virginia; VA, Virginia]

Map number (fig. 1)	Station number	Station name	State	Years of systematic record	Flood reconciliation, in percent[2]	Station skew, unitless	Rank of station skew
42	01644000	Goose Creek near Leesburg	VA	65	-1.2	0.523	114
43	02011500	Back Creek near Mountain Grove	VA	34	0	-0.445	20
44	02012500	Jackson River at Falling Spring	VA	55	1.8	-1.184	2
45	02013000	Dunlap Creek near Covington	VA	58	-0.7	0.369	98
46	02014000	Potts Creek near Covington	VA	53	-0.8	-0.209	37
47	03050000	Tygart Valley River near Dailey	WV	70	0	0.334	92
48	03050500	Tygart Valley River near Elkins	WV	53	0	0.715	124
49	03051000	Tygart Valley River at Belington	WV	90	0	0.289	87
50	03051500	Middle Fork River at Midvale	WV	27	0	0.176	78.5
51	03052000	Middle Fork River at Audra	WV	46	0	0.227	83.5
52	03052500	Sand Run near Buckhannon	WV	51	0	0.402	103
53	03053500	Buckhannon River near Hall	WV	83	0	-0.349	27
54	03054500	Tygart Valley River at Phillipi	WV	57	0	0.352	95
55	03056500	Tygart Valley River at Fetterman	WV	32	0	0.251	85
56	03057500	Skin Creek near Brownsville	WV	40	0	-0.295	30
57	03058000	West Fork River at Brownsville	WV	45	0	0.337	93
58	03058500	West Fork River at Butcherville	WV	82	0	-0.085	52
59	03059000	West Fork River at Clarksburg	WV	60	0	-0.488	17
60	03059500	Elk Creek at Quiet Dell	WV	27	0	-0.062	56
61	03061000	West Fork River at Enterprise	WV	65	0	-0.183	42
62	03061500	Buffalo Creek at Barrackville	WV	73	0	-0.178	43
63	03062400	Cobun Creek at Morgantown	WV	32	0	0.854	130
64	03062500	Deckers Creek at Morgantown	WV	30	0	0.324	91
65	03065000	Dry Fork at Hendricks	WV	57	0	1.310	144
66	03066000	Blackwater River at Davis	WV	76	0	0.746	125
67	03069000	Shavers Fork at Parsons	WV	73	0	0.992	138
68	03069500	Cheat River near Parsons	WV	84	0	1.144	141
69	03070000	Cheat River at Rowlesburg	WV	74	0	0.796	128
70	03070500	Big Sandy Creek near Rockville	WV	83	0	0.692	123
71	03071000	Cheat River near Pisgah	WV	47	0	0.130	74
72	03072000	Dunkard Creek at Shannopin	PA	35	-0.8	-0.277	31
73	03075500	Youghiogheny River near Oakland	MD	49	0	0.095	69
74	03076600	Bear Creek at Friendsville	MD	26	0	0.214	81.5
75	03108000	Raccoon Creek at Moffatts Mill	PA	51	0	-0.216	36
76	03109000	Lisbon Creek at Lisbon	OH	35	0	0.441	106
77	03109500	Little Beaver Creek near East Liverpool	OH	72	0	0.214	81.5
78	03110000	Yellow Creek near Hammondsville	OH	47	0	0.296	88
79	03111500	Short Creek near Dillonvale	OH	46	-0.1	-0.735	7
80	03112000	Wheeling Creek at Elm Grove	WV	46	0	-0.031	59
81	03114000	Captina Creek at Armstrongs Mills	OH	38	0	0.061	66

Table 1. Station skews, differences between published and new (this study) flood-frequency computations of 100-year floods (reconciliation), and results of a Wilcoxon two-sample test of station skews for 147 streamflow-gaging stations in or near West Virginia with at least 25 years of record.—Continued

[A rank with a decimal place indicates a tie with a station with the same skew value. Shading indicates station with 50 or more years of record.[1] MD, Maryland; WV, West Virginia; VA, Virginia]

Map number (fig. 1)	Station number	Station name	State	Years of systematic record	Flood reconciliation, in percent[2]	Station skew, unitless	Rank of station skew
82	03114500	Middle Island Creek at Little	WV	78	0	0.157	77
83	03115600	Barns Run near Summerfield	OH	33	0	0.121	71
84	03151500	Little Kanawha River near Burnsville	WV	35	0	-0.427	22
85	03152000	Little Kanawha River at Glenville	WV	69	0	-0.008	62
86	03153000	Steer Creek near Grantsville	WV	39	0	-0.878	4
87	03153500	Little Kanawha River at Grantsville	WV	69	0	-0.354	25
88	03154000	W. Fk. Little Kanawha River at Rocksdale	WV	63	0	-0.200	38
89	03154500	Reedy Creek near Reedy	WV	27	0	-0.596	12
90	03155000	Little Kanawha River at Palestine	WV	40	0	-0.738	6
91	03155500	Hughes River at Cisco	WV	59	0	-0.092	51
92	03171500	New River at Eggleston	VA	25	0	1.225	143
93	03173000	Walker Creek at Bane	VA	54	0	-0.302	29
94	03175500	Wolf Creek near Narrows	VA	62	2.3	-0.185	40.5
95	03176500	New River at Glen Lyn	VA	82	0	1.006	139
96	03178500	Camp Creek near Camp Creek	WV	29	0	-0.122	49
97	03179000	Bluestone River near Pipestem	WV	47	0	-1.000	3
98	03179500	Bluestone River at Lilly	WV	27	0	-0.191	39
99	03180000	New River at Bluestone Dam	WV	25	0	1.218	142
100	03180500	Greenbrier River at Durbin	WV	54	0	1.624	146
101	03182500	Greenbrier River at Buckeye	WV	68	0	0.609	118
102	03183000	Second Creek near Second Creek	WV	29	0	0.107	70
103	03183500	Greenbrier River at Alderson	WV	102	0	-0.223	34.5
104	03184000	Greenbrier River at Hilldale	WV	62	0	0.139	75
105	03185000	Piney Creek at Raleigh	WV	31	0	0.340	94
106	03186000	New River at Fayette	WV	35	0	-0.434	21
107	03186500	Williams River at Dyer	WV	68	0	0.522	113
108	03187000	Gauley River at Camden on Gauley	WV	66	0	0.227	83.5
109	03187500	Cranberry River near Richwood	WV	39	0	0.436	105
110	03189000	Cherry River at Fenwick	WV	43	0	1.077	140
111	03189100	Gauley River near Craigsville	WV	31	0	-0.035	58
112	03189500	Gauley River near Summersville	WV	37	0	0.912	133
113	03190000	Meadow River at Nallen	WV	51	0	0.029	65
114	03190400	Meadow River near Mount Lookout	WV	29	0	-0.028	60.5
115	03191500	Peters Creek near Lockwood	WV	31	0	0.400	101.5
116	03192000	Gauley River above Belva	WV	69	0	0.444	107
117	03193000	Kanawha River at Kanawha Falls	WV	120	0	-0.075	53
118	03194700	Elk River below Webster Springs	WV	68	0	0.769	126
119	03195000	Elk River at Centralia	WV	29	0	-0.043	57
120	03197000	Elk River at Queen Shoals	WV	70	0	0.654	120
121	03198500	Big Coal River at Ashford	WV	75	0	0.124	73

Table 1. Station skews, differences between published and new (this study) flood-frequency computations of 100-year floods (reconciliation), and results of a Wilcoxon two-sample test of station skews for 147 streamflow-gaging stations in or near West Virginia with at least 25 years of record.—Continued

[A rank with a decimal place indicates a tie with a station with the same skew value. Shading indicates station with 50 or more years of record.[1] MD, Maryland; WV, West Virginia; VA, Virginia]

Map number (fig. 1)	Station number	Station name	State	Years of systematic record	Flood reconciliation, in percent[2]	Station skew, unitless	Rank of station skew
122	03199000	Little Coal River at Danville	WV	54	0	-0.028	60.5
123	03200500	Coal River at Tornado	WV	42	0	-0.471	19
124	03201000	Pocatalico River at Sissonville	WV	52	0	-0.223	34.5
125	03202000	Raccoon Creek at Adamsville	OH	68	0	0.314	89
126	03202400	Guyandotte River near Baileysville	WV	29	0	-0.156	44.5
127	03203000	Guyandotte River at Man	WV	69	0	-0.413	23
128	03204000	Guyandotte River at Branchland	WV	76	0	-0.637	10
129	03204500	Mud River near Milton	WV	43	0	0.462	109
130	03206600	East Fork Twelvepole Creek near Dunlow	WV	33	0	-0.156	44.5
131	03207000	Twelvepole Creek at Wayne	WV	31	0	-0.138	47
132	03207020	Twelvepole Creek below Wayne	WV	52	0	-0.185	40.5
133	03207400	Prater Creek at Vasant	VA	27	15.1	-0.695	9
134	03207500	Levisa Fork near Grundy	VA	39	-0.3	-0.704	8
135	03208000	Levisa Fork below Fishtrap Dam	KY	55	0	-0.134	48
136	03208500	Russel Fork at Haysi	VA	65	2.2	-0.375	24
137	03208950	Cranes Nest River near Clintwood	VA	28	-0.5	0.429	104
138	03209000	Pond R. blw. Flannagan Dam near Haysi	VA	80	1.1	0.122	72
139	03210000	Johns Crek near Meta	KY	46	0	-0.541	13
140	03212000	Paint Creek at Staffordsville	KY	32	0	-0.615	11
141	03213000	Tug Fork at Litwar	WV	56	0	-0.276	32
142	03213500	Panther Creek near Panther	WV	39	0	-0.353	26
143	03213700	Tug Fork at Williamson	WV	30	0	-0.328	28
144	03214000	Tug Fork near Kermit	WV	51	0	0.006	63
145	03214500	Tug Fork at Kermit	WV	73	-0.1	-0.145	46
146	03215500	Blaine Creek at Yatesville	KY	52	0.5	0.582	115
147	03216500	Little Sandy River at Grayson	KY	30	0	-0.523	15

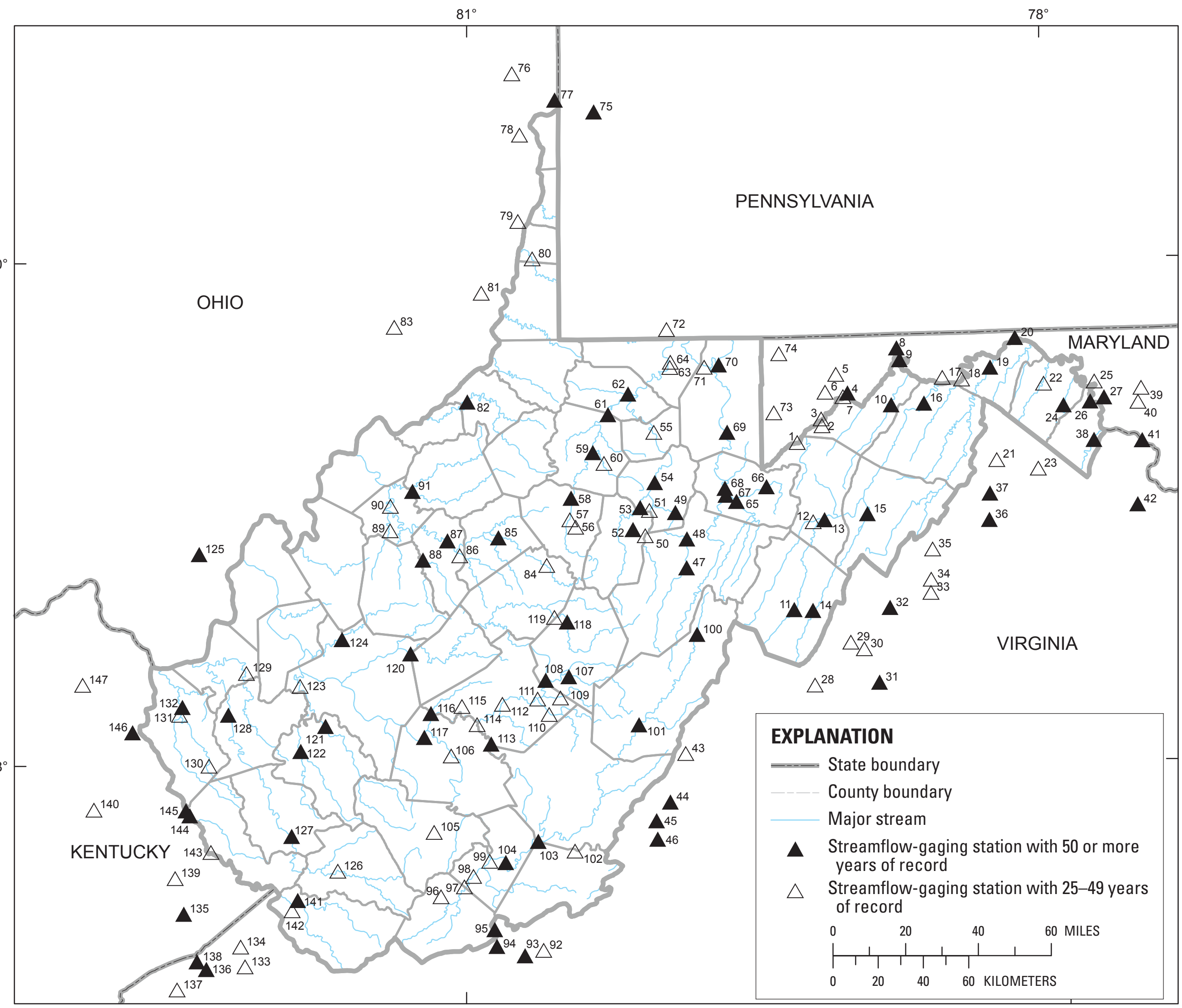
Wilcoxon two-sample test

Number of stations		Sum of ranks		Expected statistics under the null hypothesis					Conclusion
				Standard deviation	Sum of ranks		Normal approximation of Z-value[3]	Probability greater than absolute value of Z	
With 50 or more years of record	With less than 50 years of record	With 50 or more years of record	With less than 50 years of record		With 50 or more years of record	With less than 50 years of record			
75	72	6,069	4,809	258.067	5,550	5,328	-2.0092	0.045	Reject null hypothesis at alpha = 0.05

[1]Mean of station skews for shaded rows is 0.2449 and for unshaded rows is 0.0615, for a difference of 0.1834.

[2]Flood reconciliation, in percent = 100 * (published 100-year flood – recomputed 100-year flood) / published 100-year flood.

[3]Z = (test statistic – expected + continuity correction) / expected standard deviation = (4,809 – 5,328 + 0.05) / 258.067 = -2.0092.



Base from U.S. Geological Survey 1:2,000,000-scale state and county boundaries of the conterminous United States, 1994;
1:2,000,000 coverage of streams for the conterminous United States, 1991. Universal Transverse Mercator projection, zone 17, NAD 83.

Figure 1. Location of U.S. Geological Survey streamflow-gaging stations with 25 or more years of record in or near West Virginia. (Identification numbers refer to stations listed in table 1.)

maps_diagrams/regional_skewness_map.gif , has been used since its development as the means of selecting skew coefficients for use in predicting peak flows in West Virginia (U.S. Interagency Advisory Committee on Water Data, 1982). The nationwide map was developed by contouring station skew values plotted at the latitude and longitude of the streamflow-gaging station. Basin-centroid locations were not available for use in developing the Bulletin 17B skew map. The West Virginia generalized skew map, presented in this report, was developed using basin centroids rather than station locations.

In general, bias is removed from MSE by subtracting the degrees of freedom (DF) from N (in other words, the variance is typically used instead of MSE). However, this is not a factor in the Bulletin 17B methodology, thus allowing comparison of MSEs for contouring, averages, and regression analysis (whether or not the DF is known). For instance, in the case of the West Virginia generalized skew map, the DF is unknown because if smoothing were eliminated, the grid would touch each of the 147 data points, yielding zero total error—an absurd result. In lieu of applying a rule, a cautious minimum smoothing limit was used.

Contouring Station Skew Coefficients

Using the irregular grid of the 75 station skews as Z-values and basin centroids as X- and Y-points, a generalized regular grid was calculated; it consisted of 11,011 points, made up of 91 latitudes (36.5 to 41 degrees, by 0.05-degree increments) by 121 longitudes (-83.5 to -77.5 degrees, by 0.05-degree increments). The grid extent was chosen to be large enough so that the centroid of all basins for streamflow-gaging stations in or near West Virginia that are potentially representative of peak streamflows applicable to the State would fall within the grid. The SAS procedure G3GRID was used with the SPLINE option and with the parameter SMOOTH=0.1 (SAS Institute Inc., 1999). G3GRID creates a data set whose horizontal (X and Y) variable values form a complete and regular grid, and interpolates and smooths the values of the vertical (Z) variable for each point specified in the X-Y plane. SAS Institute Inc. (1999) includes the major equations of the G3GRID algorithms and explains: "The surface that is generated can be thought of as one that would be formed if a stiff, thin metal plate were forced through or near the given data points***. A smoothed spline trades closeness to the original data points for smoothness. To find a value that produces the best balance between smoothness and fit to the original data, you can try several values for the SMOOTH= option."

The 11,011 X-Y points were then projected to UTM (Universal Transverse Mercator, zone 17, NAD 27) and, along with Z-values, were used to construct a Triangulated Irregular Network (TIN) using the following software: ArcInfo 9.2 and ArcMap 9.2, and ArcView 3.3 (Environmental Systems Research Institute, Inc., 2006 and 2002, respectively). Digitally and conceptually, the skew map exists as a TIN (with interpolation provided to get a Z-value for any X- and Y-value). A shapefile of contour lines was generated using the TIN Contour Tool of the 3D Analyst from the ArcMap software then re-projected to Universal Transverse Mercator, zone 17, NAD 83. The contour lines can be made at any selected contour interval from this conceptual skew map; a contour interval of 0.1 is displayed in figure 2. The MSE of the West Virginia generalized skew map, developed from the 147 streamflow-gaging-station skews, is 0.2174 (fig. 2).

Regional Averages

Average station skews for test regions and the resulting regional MSE for the full 147-station data set are presented in table 2. Data in the column labeled "average of systematic skews" are optimum means for the full 147-station data set, meaning that these optimums cannot be excelled by the means of any subset. The best result (lowest MSE) was obtained using the means of four regions (described in the first row of table 2), yielding a MSE of 0.2546.

Regional Regression Analysis

A screening was done to evaluate the relations of 38 basin variables (Paybins, 2008) to the station skew. Results are presented in table 3.

One-Variable Results

The single best descriptive variable is annual snowfall (SNOW), with a coefficient of determination, r^2, of 0.303 (0.294 adjusted for degrees of freedom). However, this may not mean that snow itself is the best describer of skew. Mean elevation (E) is the next best descriptive variable, with r^2 of 0.168 (0.156 adjusted for degrees of freedom). SNOW may be proportional to an important effect of the topography that E fails to describe. The 147-site MSE for the SNOW model is 0.2334 for the equation:

$$\text{SKEW} = -0.4730142 + 0.0164339\ \text{SNOW},$$

where

SKEW is the station skew (unitless), and

SNOW is the annual average snowfall in inches.

SNOW is an annual average of map-grid snow values that fall within each drainage basin. The reference used by Paybins (2008) is no longer available. A readily accessible mean total snowfall for the "Lower 48 States" (SNOW14) was used. SNOW14 signifies the specific methodology and the specific data set explained in U.S. Department of Commerce (1999). A national map of SNOW14 is available at URL *http://gis.ncdc.noaa.gov/website/ims-climatls/ussnow14/viewer.htm* (U.S. Department of Commerce, 2008). This method differs from that used in previous flood studies in West Virginia, including Wiley and others (2000), in which SNOW was determined by

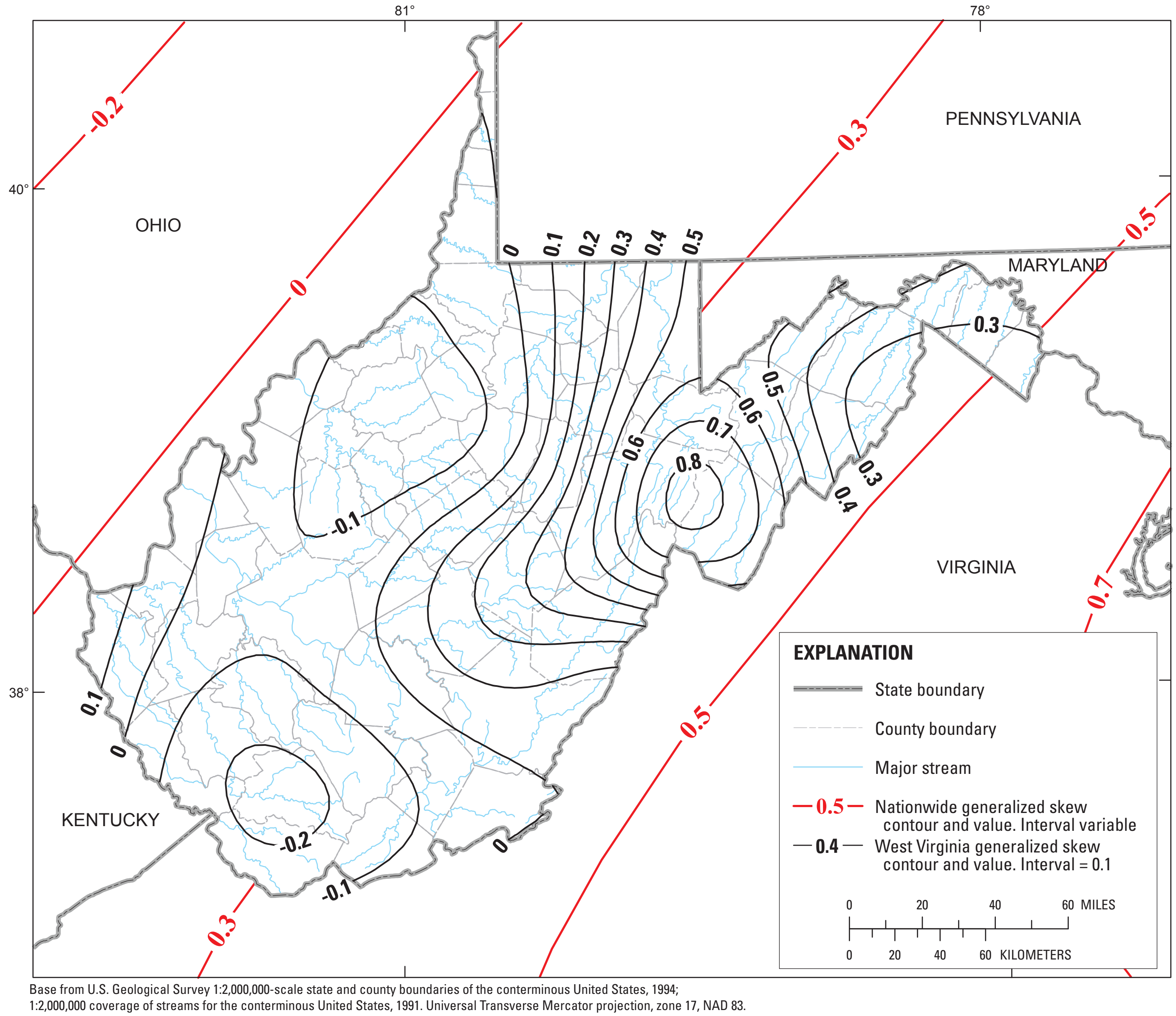


Figure 2. Generalized skew map of West Virginia. (Nationwide generalized skew refers to Generalized Skew Coefficients of Logarithms of Annual Streamflow, Bulletin 17B of the Hydrology Subcommittee, Interagency Advisory Committee on Water Data, 1982.)

Table 2. Average station skews and mean square errors for test regions using the 147 streamflow-gaging stations in or near West Virginia.

[mi^2, square miles]

Test-region description	Test regions	Average of systematic skews, unitless	Number of stations	Mean square error, unitless
Regions defined by Wiley and others (2000) with the South region divided into Central and South regions, where South region is all stations downstream from station number 03202400	East	0.409	51	0.256
	North	-0.018	36	
	Central	0.193	38	
	South	-0.217	22	
Two quadrants defined by the average longitude of basin centroids for 147 stations	East quadrant	0.347	75	0.267
	West quadrant	-0.045	72	
Regions defined by Wiley and others (2000)	East	0.409	51	0.271
	North	-0.018	36	
	South	0.043	60	
A combination of regions defined by Wiley and others (2000)	East	0.409	51	0.271
	Combined North and South	0.020	96	
Within and outside the Potomac River Basin	Within the basin	0.351	42	0.290
	Outside the basin	0.077	105	
Stations with 50 or more years of record, or stations with less than 50 years of record	50 or more years	0.245	75	0.297
	Less than 50 years	0.062	72	
Within and outside the Ohio River Basin	Within the basin	0.094	101	0.297
	Outside the basin	0.289	46	
Use of the nationwide generalized skew-coefficient map in Bulletin 17B ranks here.[1]	The full data set	0.155	147	0.301
Stations with 50 or more years of record and drainage areas of 500 mi^2 or greater, or stations with less than 50 years of record or less than 500 mi^2	50 or more years and 500 mi^2 or greater	0.264	27	0.3037
	Less than 50 years or 500 mi^2	0.131	120	
Any part of the basin boundary within Virginia, or basin entirely outside Virginia	Within Virginia	0.187	97	0.3034
	Outside Virginia	0.093	50	
Two quadrants defined by the average latitude of basin centroids for 147 stations	North	0.192	77	0.3039
	South	0.114	70	
Stations with drainage areas of 500 mi^2 or greater, or stations with drainage areas less than 500 mi^2	500 mi^2 or greater	0.212	39	0.304
	Less than 500 mi^2	0.134	108	
East or west of the Climatic Divide (Wiley and others, 2000)	East of Climatic Divide	0.171	50	0.3052
	West of Climatic Divide	0.147	97	
Station located within or outside West Virginia	Within West Virginia	0.158	88	0.3054
	Outside West Virginia	0.151	59	

[1]The pre-existing procedure is to use generalized skew values from the national skew map (Interagency Advisory Committee on Water Data, 1982). Average mean square error for the 147 stations using the national skew map is 0.3014 (the national mean square error is 0.3025).

Table 3. Results of screening relations of 38 basin variables to the generalized skew, determined from the West Virginia generalized skew map, for 147 streamflow-gaging stations using SAS procedure REG (with the "by r^2" option).

[SAS procedure REG is described in SAS Institute Inc., 1999; r^2, correlation coefficient; - -, less than 0.15]

Variable number	r^2	[1]Adjusted r^2	Variable name	Description[2]
One-variable screening				
1	0.3030	0.2935	SNOW	Mean annual snowfall, in inches
2	0.1675	0.1561	E	Mean basin elevation, in feet
3	- -	- -	LONGC	Longitude of basin centroid, in decimal degrees
4	- -	- -	LONDEG	Longitude of basin outlet, in decimal degrees
5	- -	- -	SLR	Slope ratio, dimensionless
6	- -	- -	CS	Channel slope, in feet per mile
7	- -	- -	BS	Basin slope, in feet per mile
8	- -	- -	LATDEG	Latitude, in decimal degrees
9	- -	- -	ORENTATION	Basin azimuth, in degrees
10	- -	- -	BR	Basin relief, in feet
11	- -	- -	SD	Stream density, in miles per square mile
12	- -	- -	RR	Relative relief, in feet per mile
13	- -	- -	LATC	Latitude of basin centroid, in decimal degrees
14	- -	- -	P	Mean annual precipitation, in inches
15	- -	- -	U	Urban land, in percent
16	- -	- -	I24-2	24-hour 2-year rainfall, in inches
17	- -	- -	W	Wetland, in percent
18	- -	- -	RN	Ruggedness number, in feet per mile
19	- -	- -	CM	Channel maintenance, in square miles per mile
20	- -	- -	SIR	Sinuosity ratio, dimensionless
21	- -	- -	I	Impervious, in percent
22	- -	- -	JANMIN	January minimum temperature, in degrees Fahrenheit
23	- -	- -	B	Barren land, in percent
24	- -	- -	WA	Open water, in percent
25	- -	- -	SF	Shape factor, dimensionless
26	- -	- -	RB	Rotundity of basin, dimensionless
27	- -	- -	SP	Slope proportion, dimensionless
28	- -	- -	ER	Elongation ratio, dimensionless
29	- -	- -	F	Forest, in percent
30	- -	- -	GRASS	Grassland, in percent
31	- -	- -	CL	Channel length, in miles
32	- -	- -	BP	Basin perimeter, in miles
33	- -	- -	CR	Compactness ratio, dimensionless
34	- -	- -	BW	Basin width, in miles
35	- -	- -	SL	Stream length, in miles
36	- -	- -	VL	Valley length, in miles
37	- -	- -	DA	Drainage area, in square miles
38	- -	- -	A	Agriculture, in percent
Two-variable screening				
- -	0.3612	0.3434	SNOW- ORENTATION	Basin orientation, in degrees, and mean annual snowfall, in inches
- -	0.3538	0.3358	SNOW- GRASS	Grassland, in percent, and mean annual snowfall, in inches

[1] Adjusted r^2 is an r^2 that is adjusted for the number of degrees of freedom in the screening equation and the number of degrees of freedom of the data set.

[2] Additional information is available in Paybins (2008).

visually integrating an isohyetal map published by the U.S. Department of Commerce (1968) within the area of a delineated basin.

Multi-Variable Results

Only two of the two-variable regression models yielded a 0.05 improvement in r^2 beyond the best single-variable model—the SNOW-ORIENTATION and SNOW-GRASS models (table 3). However, no two-variable model gave a 0.05 improvement in *adjusted* r^2. No three-variable model was better than the best two-variable model by more than 0.05 for r^2; no four-variable model was better than the best three-variable model by more than 0.05 for r^2; and no five-variable model was better than the best four-variable model by more than 0.05 for r^2.

The 147-station MSE for the SNOW-ORIENTATION model was 0.2215, where ORIENTATION (basin average streamflow orientation) was calculated as the orientation of a line averaged from the flow grid, colinear with the direction of flow, as determined by the direction of steepest descent from each cell. ORIENTATION, therefore, is an angle, not an azimuth. ORIENTATION is expressed in degrees, with a possible range of 0º to 180º, defined as its angle with the x-axis (an east-west or horizontal line). ORIENTATION increases counterclockwise starting at 0º in the east and going through 90º when the major axis is vertical, so there is a discontinuity at 0º = 180º (at 180º the function drops suddenly to 0º).

The 147-station MSE for the SNOW-GRASS model was 0.2426 where GRASS is a percentage of the area (30-meter grid cells) designated as grassland within a drainage basin. The National Land Cover Database (NLCD 2001) data were used to compute GRASS (Multiple Resolution Land Characteristics Consortium, 2007).

Comparison of Generalization Results

The 147-station MSE results show (1) the lowest MSE was 0.2174 for the West Virginia generalized contour skew map, (2) the MSE for the SNOW-ORIENTATION model was 0.2215, (3) the MSE for the SNOW model was 0.2334, (4) the MSE for the SNOW-GRASS model was 0.2426, and (5) merely using regional means for four regions (described in the first row of table 2) yielded a MSE of 0.2546. The MSE of 0.2174 for the West Virginia generalized contour skew map is an improvement over the MSE of 0.3025 for the national map presented in Bulletin 17B. The West Virginia generalized skew map is the best of these estimators and is a reasonable, smoothed contouring of the station skews that meets Bulletin 17B guidelines.

Summary

In an effort to provide updated information related to flood-frequency calculations, the U.S. Geological Survey (USGS), in cooperation with the West Virginia Department of Transportation, Division of Highways, conducted an investigation to develop generalized skew coefficients for rural, unregulated streams in West Virginia.

A skew coefficient is used to match peak-streamflow data to the log-Pearson Type III statistical distribution. An analysis of records from 147 streamflow-gaging stations in or near West Virginia having a minimum of 25 years of record was used to determine the best method for calculating generalized skew for the State. The procedures for the analysis generally followed guidelines established by the Interagency Advisory Committee on Water Data described in Bulletin 17B.

The difference between the mean skew for a subset of 75 stations having 50 years of record and the mean skew for the 147 stations was statistically significant. Therefore, the analysis of skews was based on the 75 stations rather than the initial 147 stations. The use of stations having 50 or more years of record was a more restrictive exception to the established guidelines.

The generalized-skew analysis included contouring, averaging, and regression of station skews. The best method was determined as the smallest mean square error (MSE), determined by comparing the station value to the regionalized value. The contouring of station skews was the best method for determining generalized skew, with a MSE of 0.2174. The MSE of 0.2174 is an improvement over the MSE of 0.3025 for the national map presented in Bulletin 17B.

Acknowledgments

The authors thank colleagues Thomas P. Suro and William H. Asquith for their technical reviews of this report. The authors are also grateful to William P. Bartlett, USGS Surface-Water Specialist for the Northeast and Midwest Areas, for helpful review and comments. The technical assistance of USGS West Virginia Water Science Center colleagues Ronald D. Evaldi and Hugh E. Bevans also is appreciated.

References Cited

Environmental Systems Research Institute (ESRI), Inc., 2006, ArcInfo 9.2 and ArcMap 9.2: Redlands, Calif., ESRI, Inc.

Environmental Systems Research Institute (ESRI), Inc., 2002, ArcView 3.3: Redlands, Calif., ESRI, Inc.

Flynn, K.M., Kirby, W.H., and Hummel, P.R., 2006, User's manual for program PeakFQ annual flood-frequency analysis using Bulletin 17B guidelines: U.S. Geological Survey Techniques and Methods, book 4, chap. B4, 42 p., accessed January 17, 2008, at URL *http://pubs.usgs.gov/tm/2006/tm4b4/tm4b4.pdf* and *http://water.usgs.gov/software/PeakFQ/*

Interagency Advisory Committee on Water Data, 1982, Guidelines for determining flood flow frequency, Bulletin 17B of the Hydrology Subcommittee: Reston, Va., U.S. Geological Survey, Office of Water Data Coordination, 183 p. (available from National Technical Information Service, Springfield, Va., as report no. PB 86 157 278, and from Federal Emergency Management Agency at URL *http://www.fema.gov/mit/tsd/dl_flow.htm*)

Multi-Resolution Land Characteristics Consortium, 2007, National land cover dataset, accessed September 20, 2007, at URL *http://www.mrlc.gov/mrlc2k_nlcd.asp*

Paybins, K.S., 2008, Basin characteristics for selected streamflow-gaging stations in and near West Virginia: U.S. Geological Survey Open-File Report 2008–1087, 9 p. (available at URL *http://pubs.usgs.gov/of/2008/1087/*).

SAS Institute Inc., 1990, SAS/STAT user's guide, v. 2, version 6 (4th ed.): Cary, N.C., SAS Institute Inc.

SAS Institute Inc. 1999, SAS/GRAPH® software: Reference, version 8: Cary, N.C., SAS Institute Inc., accessed November 2, 2007, at URL *http://www2.stat.unibo.it/ManualiSas/gref/c30.pdf*

Tasker, G.D., and Stedinger, J.R., 1986, Regional skew with weighted LS regression: Journal of Water Resources Planning and Management, v. 112, no. 2, p. 225–237.

U.S. Department of Commerce, 1968, Climatic atlas of the United States: Environmental Data Service, 80 p.

U.S. Department of Commerce, 1999, U.S. snow climatology (Data set documentation TD-9641M): Asheville, N.C., U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Climatic Data Center, accessed January 30, 2008, at URL *http://www1.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td9641m.pdf*

U.S. Department of Commerce, 2005, Climate maps of the United States: U.S. Department of Commerce, National Oceanic and Atmospheric Association, National Climatic Data Center, accessed January 30, 2008, at URL *http://cdo.ncdc.noaa.gov/cgi-bin/climaps/climaps.pl?directive=order_details&subrnum=&region=Lower%2048%20States&filename=snow1413*

U.S. Geological Survey, 2008, National Water Information System: Web interface—Peak streamflow for the Nation, accessed January 17, 2008, at URL *http://nwis/waterdata.usgs.gov/usa/nwis/peak*

Wiley, J.B., Atkins, J.T., Jr., 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.

For additional information call or write to:

Director, U.S. Geological Survey
West Virginia Water Science Center
11 Dunbar Street, Charleston, WV 25301
(304) 347-5130
http://wv.usgs.gov

Document prepared by the West Trenton Publishing Service Center

Atkins and others—**Generalized Skew Coefficients of Annual Peak Flows for Rural, Unregulated Streams in West Virginia**—Open-File Report 2008–1304

Printed on recycled paper