MULTIPLE-REGRESSION EQUATIONS FOR ESTIMATING LOW FLOWS AT UNGAGED STREAM SITES IN OHIO

By G. F. Koltun and Ronald R. Schwartz

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 86-4354

Prepared in cooperation with the

OHIO ENVIRONMENTAL PROTECTION AGENCY

U.S. DEPARTMENT OF THE INTERIOR

March 3, 1849

Columbus, Ohio

1987

DEPARTMENT OF THE INTERIOR

DONALD PAUL HODEL, Secretary

U.S. GEOLOGICAL SURVEY

Dallas L. Peck, Director

For additional information
write to:

District Chief
Water Resources Division
U.S. Geological Survey
975 W. Third Avenue
Columbus, Ohio 43212

Copies of this report can
be purchased from:

Books and Open-File Reports
U.S. Geological Survey
Building 41
Box 25425, Federal Center
Denver, Colorado 80225

CONTENTS

ILLUSTRATIONS

TABLES

4-8. Equations for annual low flows of unregulated streams in the five regions:

CONVERSION FACTORS

For the convenience of readers who may prefer to use metric (International System) units rather than the inch-pound units used in this report, values may be converted by using the following factors:

Multiply inch-pound unit	By	To obtain metric units
inch (in.)	25.40	millimeter (mm)
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
foot per mile (ft/mi)	0.1894	meter per kilometer (m/km)
square mile (mi^2)	2.590	square kilometer (km^2)
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)

Sea level: In this report "sea level" refers to the National Geodetic Vertical Datum of 1929 (NGVD of 1929)--a geodetic datum derived from a general adjustment of the first-order level nets of both the United States and Canada, formerly called "Mean Sea Level of 1929."

MULTIPLE-REGRESSION EQUATIONS FOR ESTIMATING LOW FLOWS AT UNGAGED STREAM SITES IN OHIO

By G. F. Koltun and Ronald R. Schwartz

ABSTRACT

This report presents multiple-regression equations for estimating selected low-flow characteristics for most unregulated Ohio streams at sites where little or no discharge data are available. The equations relate combinations of drainage area, main-channel length, main-channel slope, average basin elevation, forested area, average annual precipitation, and an index of infiltration to low flows with durations of 7 and 30 days and average recurrence intervals of 2 and 10 years. Data from 132 long-term continuous-record gaging stations and partial-record sites in Ohio were used in the analyses.

Multiple-regression analyses were first performed by using data from all 132 sites in an attempt to develop equations that would be applicable statewide. Standard errors for the statewide equations were too high (111 to 189 percent) for them to be of practical use in estimating low streamflows.

Data for the state were then subdivided into five regions, and multiple-regression equations were developed for each region. Standard errors for four of the five regions improved, and ranged from 43 to 106 percent. Standard errors for region 5 remained high (74 to 129 percent).

The multiple-regression equations presented in this report are not applicable to streams with significant low-flow regulation. The equations also are not applicable if (1) the site has been gaged and low-flow estimates have been developed from gaging-station records, (2) low flows can be estimated by the drainage-area transference method from data for a nearby gaged site, or (3) a sufficient number of partial-record measurements made at the site can be adequately correlated with concurrent base flows at a suitable index station.

INTRODUCTION

Information on the magnitude and frequency of low flows of streams is of critical importance to government planners and water managers in addressing such issues as the reliability of a surface-water supply or a stream's suitability as a destination for waste effluent. In Ohio, for example, the permissible rate of waste discharge into a stream generally is based on the 7-day, 10-year low flow (Ohio Environmental Protection Agency, 1978, p. 67).

Johnson and Metzker (1981) reported low-flow characteristics for a total of 237 gaged sites on regulated and unregulated streams throughout Ohio. Johnson and Metzker also provided multiple-regression equations for estimating certain low-flow characteristics at ungaged sites on unregulated streams in the Great Miami River basin upstream from Dayton, Ohio. Unfortunately, there are many ungaged sites in Ohio outside the Great Miami River basin for which no reliable means of estimating low flows has been available.

Researchers have reported various degrees of success in developing multiple-regression equations for estimation of low flows throughout a large area, such as a major river basin or a State (Thomas and Benson, 1970; Armbruster, 1976a; Flippo, 1982). Where successful, the resulting equations provide a useful tool for evaluating water-supply and water-quality problems related to low flows. In order to expand on the work of Johnson and Metzker, an investigation was conducted in cooperation with the Ohio Environmental Protection Agency to develop multiple-regression equations for estimating low-flow characteristics at sites on most unregulated streams in Ohio.

The purpose of this report is to document the development of these equations and to explain their use. The equations, which are applicable to drainage basins ranging in area from 1.35 to 1,250 square miles, are offered as a supplement to the more reliable site-specific low-flow data reported by Johnson and Metzker (1981).

SELECTION OF DATA FOR ANALYSIS

Low-Flow Data

Johnson and Metzker (1981) report low-flow discharges for several combinations of duration and recurrence interval. This report considers four of these "j-day, k-year" discharges ($Q_{j,k}$), where j equals 7 or 30 days, and k equals 2 or 10 years. These combinations of duration and recurrence interval were chosen because of their frequent use by water planners and regulatory agencies. In general terms, the minimum mean discharge for j consecutive days that occurs within a given year will be equal to or less than $Q_{j,k}$ an average of once in k years.

Low-flow data for 132 of the 237 sites reported on by Johnson and Metzker were used in developing the multiple-regression equations (fig. 1, in pocket at back of report; table 1). Sites were excluded from the analyses for the following reasons: 62 sites because of some form of significant stream regulation; 31 sites because of insufficient soils information (primarily because soils information for drainage areas that extend into Michigan, Indiana, and Pennsylvania was not readily available or compatible with data reported for the corresponding soils in Ohio); 8 sites because their drainage areas were considered too large (greater than 1,250 square miles); and 4 sites because of questionable low-flow data.

Table 1.--Descriptions and basin characteristics of low-flow sites used in multiple-regression analyses

[ft, feet; ft/mi, feet per mile; in., inches; mi^2, square miles; %, percent. Station types: C, continuous record; P, partial record. Dash, not applicable.]

Site number	USGS station number	Station name	Station type	Region	Drainage area (mi^2)	Length (mi)	Elevation (thousands of ft)	Main channel slope (ft/mi)	Forested area (%)	Average annual precipitation (in.)	Infiltration index
1	03086500	Mahoning River at Alliance	C	1	89.2	16.1	1.030	10.4	13.0	37.0	--
2	03089500	Mill Creek near Berlin Center	C	1	19.1	8.71	1.160	11.1	20.0	37.0	--
3	03092000	Kale Creek near Pricetown	C	1	21.9	10.6	.984	11.4	28.0	38.0	--
4	03092090	West Branch Mahoning River near Ravenna	C	1	21.8	10.3	1.110	19.0	27.0	38.5	--
5	03092099	Hinkley Creek at Charlestown	P	1	7.85	8.57	1.110	18.7	20.0	38.0	--
6	03093000	Eagle Creek at Phalanx Station	C	1	97.6	18.9	.958	10.7	16.9	38.0	--
7	03102950	Pymatuning Creek at Kinsman	C	1	96.7	18.8	.969	4.0	17.0	39.0	--
8	03109000	Lisbon Creek at Lisbon	C	5	6.19	5.23	1.190	55.6	15.0	39.0	--
9	03110000	Yellow Creek near Hammondsville	C	5	147	28.9	1.090	9.81	44.1	40.0	--
10	03110600	North Fork Yellow Creek at Hammondsville	P	5	59.4	16.3	1.070	22.86	14.8	39.5	--
11	03110850	Island Creek near Toronto	P	5	26.4	11.1	1.080	45.6	65.4	38.5	--
12	03111000	Cross Creek at Mingo Junction	P	5	127	28.6	1.180	13.0	26.0	38.5	--
13	03111500	Short Creek near Dillonvale	C	5	123	25.8	1.110	14.4	13.5	39.0	--
14	03111550	Wheeling Creek at Brookside	P	5	103	30.4	1.160	14.1	13.6	40.0	--
15	03114000	Captina Creek at Armstrongs Mills	C	5	134	26.2	1.140	16.0	8.3	41.0	--
16	03115400	Little Muskingum River at Bloomfield	C	5	210	43.3	.974	7.0	50.0	41.5	--
17	03115650	East Fork Duck Creek at Lower Salem	P	5	111	28.2	.993	4.7	24.7	40.0	--
18	03116200	Chippewa Creek at Easton	C	4	146	21.8	1.130	5.0	10.0	37.0	--
19	03117150	Sandy Creek at Minerva	P	4	61.9	14.0	1.190	7.63	13.2	39.5	--
20	03117160	Still Fork near Minerva	P	4	36.2	8.33	1.120	6.4	25.5	39.5	--
21	03117300	Sandy Creek at Malvern	P	4	163	21.4	1.170	6.3	7.9	37.5	--
22	03117310	Pipe Run at Malvern	P	4	27.7	11.0	1.150	7.28	27.8	37.5	--
23	03117500	Sandy Creek at Waynesburg	C	4	253	25.6	1.180	7.61	10.0	37.5	--
24	03118100	East Branch Nimishillen Creek near Canton	P	4	33.4	6.55	1.120	8.14	.6	37.0	--
25	03119700	Conotton Creek at Jewett	P	4	14.3	4.78	1.150	22.3	9.8	40.0	--
26	03122850	Sugar Creek near Orrville	P	4	47.2	13.5	1.120	11.0	3.6	36.5	--
27	03122900	Sugar Creek near West Lebanon	P	4	69.8	22.3	1.130	7.18	5.7	41.0	--
28	03125900	Boggs Fork at Piedmont	P	4	36.6	13.0	1.080	6.18	26.2	40.5	--
29	03127100	Crooked Creek near Stillwater	P	4	47.5	8.49	1.030	12.56	14.2	40.0	--
30	03129100	White Eyes Creek near Fresno	P	4	52.1	10.2	1.010	21.0	11.8	39.0	--
31	03134000	Jerome Fork at Jeromeville	C	4	120	18.5	1.090	9.56	.9	35.0	--
32	03134300	Muddy Fork near Rowsburg	P	4	66.2	21.8	1.140	10.4	16.1	35.5	--
33	03136500	Kokosing River at Mount Vernon	C	4	202	31.1	1.180	10.1	2.3	37.0	--
34	03137000	Kokosing River at Millwood	C	4	455	50.5	1.140	7.57	2.1	38.0	--
35	03138790	Killbuck Creek at Burbank	P	4	42.4	12.7	1.100	17.8	8.1	36.0	--

Table 1.--Descriptions and basin characteristics of low-flow sites used in multiple-regression analyses--Continued

Site number	USGS station number	Station name	Station type	Region	Drainage area (mi^2)	Length (mi)	Elevation (thousands of ft)	Main channel slope (ft/mi)	Forested area (%)	Average annual precipitation (in.)	Infiltration index
36	03138800	Killbuck Creek at Wooster	P	4	128	28.3	1.070	7.01	14.4	36.0	--
37	03140000	Mill Creek near Coshocton	C	4	27.2	10.4	1.000	21.1	27.0	39.5	--
38	03141900	Leatherwood Creek near Cambridge	P	4	88	22.7	1.050	7.04	25.0	39.0	--
39	03144000	Wakatomika Creek near Frazeysburg	C	4	140	31.6	.963	10.3	5.0	39.0	--
40	03146000	North Fork Licking River at Utica	C	4	116	25.4	1.150	14.0	7.0	38.0	--
41	03148450	Jonathan Creek at East Fultonham	P	4	125	27.5	1.000	7.76	9.4	38.0	--
42	03149500	Salt Creek near Chandlersville	C	4	75.7	23.8	.921	8.98	9.4	38.5	--
43	03150250	Meigs Creek near Beverly	P	4	136	22.7	.938	5.88	14.6	39.0	--
44	03150480	West Branch Wolf Creek near Waterford	P	4	144	48.5	.919	4.4	33.4	39.0	--
45	03150490	South Branch Wolf Creek near Waterford	P	4	79.3	25.9	.839	5.15	22.5	39.0	--
46	03156400	Hocking River at Lancaster	C	5	48.2	13.2	.942	19.8	1.0	39.0	--
47	03157000	Clear Creek near Rockbridge	C	5	89.0	22.5	.993	9.20	9.2	39.0	--
48	03157500	Hocking River at Enterprise	C	5	459	29.2	.962	10.6	5.6	39.0	--
49	03159540	Shade River near Chester	C	5	156	29.8	.749	4.0	42.0	40.5	--
50	03201800	Sandy Run near Lake Hope	C	5	4.99	4.24	.913	22.1	98.0	40.5	--
51	03201990	Little Raccoon Creek near Vinton	P	5	154	41.3	.855	3.23	30.6	40.5	--
52	03202000	Raccoon Creek at Adamsville	C	5	585	71.6	.829	2.81	21.9	40.0	--
53	03205210	Indian Guyan Creek near Bradrick	P	5	67.5	27.6	.785	6.8	73.5	41.0	--
54	03217400	Scioto River near Kenton	P	2	129	32.9	1.020	3.45	5.3	34.0	1.38
55	03217500	Scioto River at LaRue	C	3	255	45.2	.944	2.50	.6	34.5	--
56	03218000	Little Scioto River above Marion	C	3	72.4	17.7	.980	4.30	.3	35.5	--
57	03219500	Scioto River near Prospect	C	3	567	70.7	.997	1.53	.4	35.0	--
58	03220000	Mill Creek near Bellepoint	C	3	178	45.3	1.020	5.22	.1	37.0	--
59	03223000	Olentangy River at Claridon	C	3	157	36.6	1.040	6.75	.4	35.5	--
60	03224500	Whetstone Creek near Ashley	C	3	98.7	29.6	1.130	11.7	.9	35.5	--
61	03228200	Big Walnut Creek above Sunbury	P	3	77.8	24.6	1.250	7.04	13.1	37.0	--
62	03228700	Blacklick Creek near Groveport	P	3	57.4	20.1	.996	17.2	13.2	38.0	--
63	03230200	Big Darby Creek at Plain City	P	3	151	33.1	1.040	5.84	7.0	37.0	--
64	03230300	Little Darby Creek at Chuckery	P	3	71.4	16.7	1.070	10.27	6.6	37.5	--
65	03230500	Big Darby Creek at Darbyville	C	3	534	70.5	.986	3.87	.7	37.0	--
66	03230800	Deer Creek at Mount Sterling	C	3	228	38.1	.905	7.0	4.0	38.0	--
67	03231300	Kinnikinnick Creek near Kinnikinnick	P	3	36.2	14.5	.808	11	2.8	39.0	--
68	03235500	Tar Hollow Creek at Tar Hollow State Park	C	3	1.35	1.67	1.015	140	96	39.5	--
69	03236000	Salt Creek near Londonderry	C	3	286	38.0	.906	13.0	64.0	39.5	--
70	03237280	Upper Twin Creek at McGaw	C	5	12.2	6.91	1.050	67	96	43	--

Table 1.--Descriptions and basin characteristics of low-flow sites used in multiple-regression analyses--Continued

Site number	USGS station number	Station name	Station type	Region	Drainage area (mi^2)	Length (mi)	Elevation (thousands of ft)	Main channel slope (ft/mi)	Forested area (%)	Average annual precipitation (in.)	Infiltration index
71	03237500	Ohio Brush Creek near West Union	C	5	387	45.9	.877	8.3	9.0	42.0	--
72	03238200	Eagle Creek near Ripley	P	5	137	26.2	.865	15.78	16.6	42.0	--
73	03238500	Whiteoak Creek near Georgetown	C	5	218	44.4	.984	7.92	3.9	41.5	--
74	03240000	Little Miami River near Oldtown	C	2	129	26.4	1.060	13.2	1.3	38.0	3.01
75	03240500	North Fork Massies Creek at Cedarville	C	2	28.9	13.9	1.090	6.98	3.0	38.5	1.53
76	03241000	South Fork Massies Creek near Cedarville	C	2	17.1	10.8	1.090	7.17	2.0	38.5	1.26
77	03241500	Massies Creek at Wilberforce	C	2	63.2	20.5	1.070	14.2	1.0	38.5	1.33
78	03242050	Little Miami River near Spring Valley	C	2	366	43.5	.991	10.0	2.6	38.5	2.41
79	03242150	Caesar Creek near Xenia	C	2	71.4	16.7	1.030	13.0	5.0	39.0	1.27
80	03242200	Anderson Fork near New Burlington	C	2	77.8	25.6	1.040	7.0	5.0	42.0	1.47
81	03243400	Cowan Creek at Clinton County AFB	P	2	29.7	11.1	1.070	7.81	4.1	43.0	1.49
82	03246500	East Fork Little Miami River at Williamsburg	C	5	237	52.7	1.000	5.27	1.3	40.5	--
83	03255500	Mill Creek at Reading	C	5	73.0	15.9	.715	7.6	13.6	39.0	--
84	03260620	Muchinippi Creek near Russells Point	P	2	86.2	17.0	1.010	3.53	6.0	40.5	1.91
85	03260700	Bokengehalas Creek near DeGraff	C	2	36.3	13.7	1.180	28.6	1.0	40.5	2.24
86	03260800	Stoney Creek near DeGraff	C	2	59.1	19.0	1.130	22.8	1.2	40.5	1.88
87	03262000	Loramie Creek at Lockington	C	2	257	39.6	.938	2.53	8.0	41.5	1.26
88	03262500	Great Miami River at Piqua	P	2	866	58.1	.950	3.2	7.0	41.0	1.69
89	03262700	Great Miami River at Troy	C	2	926	67.1	.930	3.2	7.0	37.5	1.69
90	03262800	Lost Creek near Troy	P	2	55.3	20.1	1.040	14.9	3.0	42.5	1.00
91	03262900	Honey Creek near New Carlisle	P	2	72.8	21.5	.964	17.9	4.0	42.5	1.72
92	03264000	Greenville Creek near Bradford	C	2	193	37.3	1.040	5.79	4.3	42.5	1.48
93	03265000	Stillwater River at Pleasant Hill	C	2	503	43.2	.949	3.06	3.0	42.0	1.40
94	03266000	Stillwater River at Englewood	C	2	650	62.6	.894	4.11	3.0	42.0	1.42
95	03266500	Mad River near Zanesfield	C	2	7.31	4.90	1.300	49.8	32.0	41.5	2.76
96	03267000	Mad River near Urbana	C	2	162	26.3	1.110	11.0	1.5	42.0	3.76
97	03267400	Cedar Run near Tremont City	P	2	2.08	3.52	.962	9.10	20.0	42.5	7.50
98	03267600	Chapman Creek at Tremont City	P	2	24	13.3	1.080	20.1	8.0	42.5	1.37
99	03267900	Mad River at St. Paris Pike, at Eagle City	C	2	310	36.6	1.050	8.0	5.0	42.0	3.18
100	03268000	Buck Creek at New Moorefield	C	2	65.3	12.7	1.120	18.4	2.2	43.0	4.05
101	03269500	Mad River near Springfield	C	2	490	40.4	1.030	8.32	1.7	38.0	3.33
102	03270800	Wolf Creek at Trotwood	C	2	22.7	12.3	.830	19.0	6.0	38.0	1.80
103	03271300	Holes Creek near Kettering	P	2	18.7	9.32	.948	18.6	10.1	39.0	1.46
104	03271700	Clear Creek at Franklin	P	2	51.6	14.4	.880	19.46	6.5	40.0	1.46
105	03271800	Twin Creek near Ingomar	C	2	197	28.2	1.010	10.0	5.0	38.0	1.76

Table 1.--Descriptions and basin characteristics of low-flow sites used in multiple-regression analyses--Continued

Site number	USGS station number	Station name	Station type	Region	Drainage area (mi^2)	Length (mi)	Elevation (thousands of ft)	Main channel slope (ft/mi)	Forested area (%)	Average annual precipitation (in.)	Infiltration index
106	03272000	Twin Creek near Germantown	C	2	275	44.0	.890	9.27	3.9	39.0	1.77
107	03272200	Elk Creek at Miltonville	P	2	46.2	14.9	.930	23.4	10.9	40.0	1.78
108	03272300	Dick's Creek near Excello	P	2	44.8	9.78	.767	32.1	2.3	41.0	1.50
109	03272800	Sevenmile Creek at Collinsville	C	2	120	31.4	.954	16.0	10.0	39.0	2.07
110	04187500	Ottawa River at Allentown	C	1	160	40.2	.840	3.90	.7	36.0	--
111	04188300	Blanchard River at Mount Blanchard	P	1	109	26.5	.920	4.5	7.5	34.5	--
112	04188500	Eagle Creek near Findlay	C	1	55	46.6	.763	3.69	2.5	35.0	--
113	04196000	Sandusky River near Bucyrus	C	1	88.8	29.4	1.020	7.37	1.1	36.0	--
114	04196200	Broken Sword Creek at Nevada	P	1	83.8	30.3	.980	4.4	7.7	35.5	--
115	04196500	Sandusky River near Upper Sandusky	C	1	298	56.3	.987	6.58	1.5	35.0	--
116	04197000	Sandusky River near Mexico	C	1	774	89.9	.831	4.34	.7	35.0	--
117	04198000	Sandusky River near Fremont	C	1	1,251	118	.825	4.03	1.7	35.0	--
118	04198500	East Branch Huron River near Norwalk	C	1	85.5	15.5	.878	14.8	3.2	34.0	--
119	04199000	Huron River at Milan	C	1	371	55.1	.852	9.31	3.7	35.5	--
120	04199300	Vermilion River at Clarksfield	P	1	130	34.0	1.060	6.27	10.2	35.0	--
121	04199500	Vermilion River near Vermilion	C	1	262	64.1	.782	6.99	8.8	34.5	--
122	04200000	East Branch Black River at Elyria	C	1	217	59.7	.859	6.78	9.8	36.0	--
123	04200500	Black River at Elyria	C	1	396	63.4	.907	6.69	3.4	35.0	--
124	04201400	West Branch Rocky River at West View	P	1	147	25.8	1.110	9.7	13.5	36.0	--
125	04201500	Rocky River near Berea	C	1	267	30.8	.971	9.45	6.4	36.5	--
126	04207200	Tinkers Creek at Bedford	C	1	83.9	24.3	1.070	5.0	24.0	38.5	--
127	04208900	Aurora Branch near Chagrin Falls	P	1	54.7	17.1	1.120	17.9	31.9	41.0	--
128	04209000	Chagrin River at Willoughby	C	1	246	32.6	1.040	12.40	22.4	40.0	--
129	04210000	Phelps Creek near Windsor	C	1	25.6	11.1	1.010	20.7	36.0	39.5	--
130	04212000	Grand River near Madison	C	1	581	52.2	.965	1.45	17.2	39.5	--
131[a]	03261500	Great Miami River at Sidney	C	2	541	45.2	.995	2.95	1.0	36.5	1.84
132[a]	03262650	Spring Creek near Troy	P	2	21.0	7.67	.958	18.3	1.0	37.0	1.00

[a] These stations were added after the original data set.

Streams larger than 1,250 square miles were eliminated because of the potential bias that they could introduce into the model. The eight sites that were eliminated had drainage areas that ranged from 2,129 to 6,330 square miles. These values are approximately two to five times larger than the largest drainage area used in the analysis. Elimination of these sites was not considered a problem because low-flow characteristics on most large streams are fairly well defined.

The 132 sites used in these analyses are composed of 85 long-term continuous-record gaging stations and 47 partial-record stations. A long-term continuous-record gaging station is a site where daily discharge has been systematically observed for 10 or more years. A partial-record station is defined as either (1) a site where sparse low-flow data have been collected and correlated with a hydrologically similar long-term continuous-record station or, (2) a site where systematic observations of daily discharge data have been obtained for at least 2 years, but less than 10 years.

Basin-Characteristic Data

The low-flow characteristics of a given stream are highly dependent on certain physical and climatological characteristics of the drainage basin. Not all of the characteristics that affect low flows are easily quantified. For multiple-regression analysis, a set of measurable basin characteristics must be chosen that explains the observed variability in the low streamflows within a specified error range.

Seven basin characteristics were used in various combinations in the multiple-regression models. These characteristics can be placed into four basic categories--basin morphology, land use, climate, and geology. The characteristics used in the multiple-regression equations are described below:

Drainage area (A)--in square miles; determined from U.S. Geological Survey topographic maps.

Main-channel length (L)--in miles; measured along the channel from the basin divide to the site where low flow is to be estimated.

Mean basin elevation (E)--in thousands of feet above sea level; measured from topographic maps by the transparent-grid sampling method (elevations were determined at 20 to 80 equally spaced unique points in the basin and averaged).

Main-channel slope (S)--in feet per mile; computed as the difference between the elevations, in feet, at 10 and 85 percent of the main-channel length (as defined above), divided by the distance in miles between these two points, as determined from U.S. Geological Survey topographic maps.

Forested area (F)--expressed as a percentage of the drainage area, as determined from U.S. Geological Survey topographic maps.

Average annual precipitation (P)--in inches; determined for each drainage basin by interpolation between isohyetal lines on the precipitation map (fig. 2, in pocket at back of report).

Index of relative infiltration (I)--dimensionless; the relative capacity of soils to accept and release water, determined from the soil-group map (fig. 3, in pocket at back of report) and the following formula originally presented by Armbruster (1976b) in a different but algebraically equivalent form:

$$I = (Af_1 \times Wf_1) + (Af_2 \times Wf_2) + \ldots + (Af_k \times Wf_k)$$

where I is the index of relative infiltration;

Af_i is the fraction of the drainage area covered by soil grouping i (i=1,2,...k);

Wf_i is the weighting factor for soil group i (i=1,2,...k); and

k is the total number of unique soil groups present in the basin.

Soil-group weighting-factor maps were compiled from county soil maps prepared by the Ohio Department of Natural Resources, Division of Lands and Soils. Weighting factors for each of the eight major soil groups are presented in table 2. Weighting factors were based on generalized relative infiltration capacities for the different soil groups. The weighting factors were optimized by trial and error in the regression analyses to minimize the standard error of regression. (For additional detail, see Armbuster, 1976b.) As Armbruster states, the numerical value of the index has no true physical significance; it is simply a means of accounting for varying soil characteristics.

Table 2.--Weighting factors for use with soil maps to determine infiltration index

Soil type	Description	Weighting factor
1	Ground moraine; end moraine; kame----------	1.0
2	Outwash------------------------------------	8.0
3	Alluvium-----------------------------------	7.5
4	Lake beds----------------------------------	2.3
5	Glacial stream bottom; soils from bedrock--	.9
6	Urban; spoil-------------------------------	.1
7[a]	Beach ridges-------------------------------	.8
8[a]	Colluvium and residuum---------------------	.3

[a] These soil types are not present in region 2, the only region for which infiltration index is a significant independent variable.

DEVELOPMENT OF MULTIPLE-REGRESSION EQUATIONS

In multiple linear regression, the dependent variable is assumed to be a linear function of two or more independent variables. The general form of the multiple-regression model used in this study is:

$$\log(Y + 0.1) = \log(C) + b_1\log(X_1) + b_2\log(X_2) + \ldots + b_n\log(X_n)$$

where Y is low streamflow of a specified duration and recurrence interval;

C is a constant;

b_i is the regression coefficient for the ith independent variable;

X_i is the ith independent variable; and

n is the total number of independent variables in the equation.

For computational convenience, the above equation is presented in this report in the algebraically equivalent form given below:

$$Y = C(X_1)^{b_1}(X_2)^{b_2} \; . \; . \; . \; (X_n)^{b_n} - 0.1$$

The constant (0.1) in the above equations was added to the dependent variable prior to taking logarithms so that sites having $Q_{j,k}$ flows equal to zero could be included in the analyses. The constant was subtracted from both sides in the exponential form of the equation to isolate the dependent variable.

Selection of Regression Variables

Multiple-regression analyses were performed using the SAS[1] statistical programs (SAS Institute, 1982). The RSQUARE and STEPWISE procedures were used as an aid in determining the optimum set of regressor variables to include in the multiple-regression equations. The final selection of regressor variables, however, was based on the following criteria:

1.--The choice of regressor variables, as well as the signs and magnitudes of their associated regression coefficients, must be hydrologically plausible in the context of low flows. This criterion takes precedence over all other criteria.

2.--All regressor variables should be statistically significant at the 5-percent level.

3.--The choice of regressor variables, within the constraints of criteria 1 and 2, should minimize the standard error (root mean-square error) and maximize the coefficent of determination (R^2, a measure of the proportion of the variation in the dependent variable accounted for by the regression equation).

After the potential set of regressor variables was chosen, the REG procedure was used so that tests for bias, constancy of residual variances, and multicollinearity of regressors could be performed.

---

[1] Use of trade and firm names in this report is for identification purposes only and does not constitute endorsement by the U.S. Geological Survey.

Attempts to Analyze Data on a Statewide Basis

Multiple-regression analyses were first performed using data from all 132 unregulated low-flow sites in Ohio (tables 1 and 3) in an attempt to develop equations that would be applicable statewide. Standard errors for the statewide equations were too high (111 to 189 percent) for them to be of practical use in estimating low streamflows.

Analysis of Data by Selected Regions

Data for the state were then subdivided into large drainage basins (regions), and multiple-regression analyses performed on the data sets for each region.

Regional boundaries were adjusted, if necessary, after an analysis of the residuals of the multiple-regression equations developed for each region. The residuals, which are the differences between the observed and predicted values of the dependent variable ($Q_{j,k}$), were plotted on a map at the location of the corresponding low-flow site. Geographic trends in residuals were then used as an aid in repartitioning the data. The multiple-regression analysis was repeated for the newly formed regions to insure that boundary adjustments did, in fact, result in an improved fit to the data, as determined from the standard error and coefficient of determination.

The above procedure was repeated until no further improvement in fit was observed. Final delineation of region boundaries was made using geology and topographic divides as guides. Multiple-regression on the $Q_{7,2}$ discharge was used for purposes of subdivision.

Table 3.--Low-flow discharges for sites used in multiple regression analyses

Site number	USGS station number	Station name	Low-flow discharge, in cubic feet per second $Q_{7,2}$	$Q_{7,10}$	$Q_{30,2}$	$Q_{30,10}$
1	03086500	Mahoning River at Alliance, Ohio	5.4	2.3	7.7	3.9
2	03089500	Mill Creek near Berlin Center, Ohio	.2	0	.4	.1
3	03092000	Kale Creek near Pricetown, Ohio	.1	0	.2	.1
4	03092090	West Branch Mahoning River near Ravenna, Ohio	1.5	.8	2.1	1.1
5	03092099	Hinkley Creek at Charlestown, Ohio	.3	.2	.4	.2
6	03093000	Eagle Creek at Phalanx Station, Ohio	9.6	6.2	12.0	7.9
7	03102950	Pymatuning Creek at Kinsman, Ohio	1.2	.3	3.2	.9
8	03109000	Lisbon Creek at Lisbon, Ohio	.2	.1	.2	.1
9	03110000	Yellow Creek near Hammondsville, Ohio	5.3	1.9	8.6	3.5
10	03110600	North Fork Yellow Creek at Hammondsville, Ohio	1.5	.8	2.4	1.2
11	03110850	Island Creek near Toronto, Ohio	1.4	.6	2.1	1.0
12	03111000	Cross Creek at Mingo Junction, Ohio	3.7	.5	6.9	1.3
13	03111500	Short Creek near Dillonvale, Ohio	16	7.7	20	11
14	03111550	Wheeling Creek at Brookside, Ohio	11	4.5	15	7.0
15	03114000	Captina Creek at Armstrongs Mills, Ohio	1.0	0	4.7	.2
16	03115400	Little Muskingum River at Bloomfield, Ohio	1.3	.2	3.4	.6
17	03115650	East Fork Duck Creek at Lower Salem, Ohio	2.0	.4	4.3	1.1
18	03116200	Chippewa Creek at Easton, Ohio	9.3	5.6	12	6.8
19	03117150	Sandy Creek at Minerva, Ohio	4.4	2.7	5.6	3.2
20	03117160	Still Fork near Minerva, Ohio	.6	.3	1.0	.4
21	03117300	Sandy Creek at Malvern, Ohio	20	14	24	16
22	03117310	Pipe Run at Malvern, Ohio	.2	.1	.3	.1
23	03117500	Sandy Creek at Waynesburg, Ohio	27	17	34	20
24	03118100	East Branch Nimishillen Creek near Canton, Ohio	3.7	2.7	4.4	3.0
25	03119700	Conotton Creek at Jewett, Ohio	.3	.1	.5	.2
26	03122850	Sugar Creek near Orrville, Ohio	1.5	.7	2.2	1.0
27	03122900	Sugar Creek near West Lebanon, Ohio	2.3	.9	3.4	1.4
28	03125900	Boggs Fork at Piedmont, Ohio	.5	.1	1.1	.2
29	03127100	Crooked Creek near Stillwater, Ohio	.4	.1	1.0	.2
30	03129100	White Eyes Creek near Fresno, Ohio	.5	.1	1.2	.2
31	03134000	Jerome Fork at Jeromeville, Ohio	4.0	2.5	6.3	3.9
32	03134300	Muddy Fork near Rowsburg, Ohio	3.0	1.7	3.9	2.1
33	03136500	Kokosing River at Mount Vernon, Ohio	29	18	33	20
34	03137000	Kokosing River at Millwood, Ohio	62	42	71	48
35	03138790	Killbuck Creek at Burbank, Ohio	.4	.1	.6	.2

Table 3.--Low-flow discharges for sites used in multiple regression analyses--Continued

Site number	USGS station number	Station name	Low-flow discharge, in cubic feet per second			
			$Q_{7,2}$	$Q_{7,10}$	$Q_{30,2}$	$Q_{30,10}$
36	03138800	Killbuck Creek at Wooster, Ohio	3.8	2.0	4.9	2.5
37	03140000	Mill Creek near Coshocton, Ohio	.4	.1	.8	.2
38	03141900	Leatherwood Creek near Cambridge, Ohio	.8	.2	2.5	.8
39	03144000	Wakatomika Creek near Frazeysburg, Ohio	7.9	3.8	11.0	5.2
40	03146000	North Fork Licking River at Utica, Ohio	4.3	2.1	5.9	3.0
41	03148450	Jonathan Creek at East Fultonham, Ohio	5.4	3.3	6.6	3.8
42	03149500	Salt Creek near Chandlersville, Ohio	.7	.1	2.2	.7
43	03150250	Meigs Creek near Beverly, Ohio	3.3	.8	6.8	1.8
44	03150480	West Branch Wolf Creek near Waterford, Ohio	.6	.1	1.1	.2
45	03150490	South Branch Wolf Creek near Waterford, Ohio	.3	.1	.8	.2
46	03156400	Hocking River at Lancaster, Ohio	5.1	2.4	6.6	3.5
47	03157000	Clear Creek near Rockbridge, Ohio	13	8.5	15	11
48	03157500	Hocking River at Enterprise, Ohio	47	30	59	36
49	03159540	Shade River near Chester, Ohio	3.6	1.0	8.9	2.3
50	03201800	Sandy Run near Lake Hope, Ohio	0	0	.1	0
51	03201990	Little Raccoon Creek near Vinton, Ohio	4.6	1.3	6.9	1.9
52	03202000	Raccoon Creek at Adamsville, Ohio	14.0	3.5	22	5.3
53	03205210	Indian Guyan Creek near Beadrick, Ohio	.2	0	.4	0
54	03217400	Scioto River near Kenton, Ohio	3.2	2.0	4.2	2.4
55	03217500	Scioto River at LaRue, Ohio	5.3	3.1	7.3	3.9
56	03218000	Little Scioto River above Marion, Ohio	.1	0	.2	0
57	03219500	Scioto River near Prospect, Ohio	16	9.3	20	11
58	03220000	Mill Creek near Bellepoint, Ohio	1.8	.4	2.9	.9
59	03223000	Olentangy River at Claridon, Ohio	2.4	.3	4.5	.5
60	03224500	Whetstone Creek near Ashley, Ohio	2.5	.2	3.7	.6
61	03228200	Big Walnut Creek above Sunbury, Ohio	.1	0	.2	0
62	03228700	Blacklick Creek near Groveport, Ohio	2.4	1.0	3.5	1.3
63	03230200	Big Darby Creek at Plain City, Ohio	2.5	.6	3.5	1.0
64	03230300	Little Darby Creek at Chuckery, Ohio	2.4	.3	4.0	.6
65	03230500	Big Darby Creek at Darbyville, Ohio	18	5.3	24	8.4
66	03230800	Deer Creek at Mount Sterling, Ohio	12	7.4	15	9.3
67	03231300	Kinnikinnick Creek near Kinnikinnick, Ohio	7.9	5.8	8.8	7.0
68	03235500	Tar Hollow Creek at Tar Hollow State Park, Ohio	0	0	0	0
69	03236000	Salt Creek near Londonderry, Ohio	8.7	6.1	13	8.0
70	03237280	Upper Twin Creek at McGaw, Ohio	.1	0	.2	0

Table 3.--Low-flow discharges for sites used in multiple regression analyses--Continued

Site number	USGS station number	Station name	Low-flow discharge, in cubic feet per second $Q_{7,2}$	$Q_{7,10}$	$Q_{30,2}$	$Q_{30,10}$
71	03237500	Ohio Brush Creek near West Union, Ohio	2.4	0.3	6.8	0.9
72	03238200	Eagle Creek near Ripley, Ohio	.2	0	.4	.1
73	03238500	Whiteoak Creek near Georgetown, Ohio	.7	0	2.7	.2
74	03240000	Little Miami River near Oldtown, Ohio	15.0	8.1	16	9.0
75	03240500	North Fork Massies Creek at Cedarville, Ohio	1.0	0	1.1	.2
76	03241000	South Fork Massies Creek near Cedarville, Ohio	.2	0	.3	0
77	03241500	Massies Creek at Wilberforce, Ohio	3.8	1.1	4.8	1.7
78	03242050	Little Miami River near Spring Valley, Ohio	62	34	74	42
79	03242150	Caesar Creek near Xenia, Ohio	1.8	.7	2.9	1.3
80	03242200	Anderson Fork near New Burlington, Ohio	1.6	.4	3.1	.6
81	03243400	Cowan Creek at Clinton County AFB, Ohio	.1	0	.1	0
82	03246500	East Fork Little Miami River at Williamsburg, Ohio	.8	0	1.7	.1
83	03255500	Mill Creek at Reading, Ohio	7.1	4.0	11.0	5.7
84	03260620	Muchinippi Creek near Russells Point, Ohio	1.3	.7	1.7	.8
85	03260700	Bokengehalas Creek near DeGraff, Ohio	5.9	3.4	6.6	4.0
86	03260800	Stoney Creek near DeGraff, Ohio	11.0	6.4	12	7.1
87	03262000	Loramie Creek at Lockington, Ohio	5.6	3.3	6.9	3.8
88	03262500	Great Miami River at Piqua, Ohio	40	24	49	28
89	03262700	Great Miami River at Troy, Ohio	56	27	68	31
90	03262800	Lost Creek near Troy, Ohio	1.7	.7	2.5	.9
91	03262900	Honey Creek near New Carlisle, Ohio	7.0	3.8	8.1	4.7
92	03264000	Greenville Creek near Bradford, Ohio	18	9.9	21	12
93	03265000	Stillwater River at Pleasant Hill, Ohio	25	12	33	17
94	03266000	Stillwater River at Englewood, Ohio	34	15	42	20
95	03266500	Mad River near Zanesfield, Ohio	1.2	.7	1.4	.9
96	03267000	Mad River near Urbana, Ohio	51	32	55	35
97	03267400	Cedar Run near Tremont City, Ohio	3.3	2.1	3.5	2.3
98	03267600	Chapman Creek at Tremont City, Ohio	.9	.4	1.1	.5
99	03267900	Mad River at St. Paris Pike, at Eagle City, Ohio	114	77	121	88
100	03268000	Buck Creek at New Moorefield, Ohio	24	15	26	17
101	03269500	Mad River near Springfield, Ohio	166	118	182	130
102	03270800	Wolf Creek at Trotwood, Ohio	.4	0	.7	.2
103	03271300	Holes Creek near Kettering, Ohio	1.2	.3	1.5	.8
104	03271700	Clear Creek at Franklin, Ohio	.8	.2	1.2	.4
105	03271800	Twin Creek near Ingomar, Ohio	7.1	3.5	8.9	4.3

Table 3.--Low-flow discharges for sites used in multiple regression analyses--Continued

Site number	USGS station number	Station name	Low-flow discharge, in cubic feet per second $Q_{7,2}$	$Q_{7,10}$	$Q_{30,2}$	$Q_{30,10}$
106	03272000	Twin Creek near Germantown, Ohio	10	4.4	14	5.9
107	03272200	Elk Creek at Miltonville, Ohio	.8	.3	1.2	.4
108	03272300	Dick's Creek near Excello, Ohio	2.0	1.0	2.5	1.3
109	03272800	Sevenmile Creek at Collinsville, Ohio	4.0	2.0	5.6	2.3
110	04187500	Ottawa River at Allentown, Ohio	17	13	21	15
111	04188300	Blanchard River at Mount Blanchard, Ohio	.2	0	.3	.1
112	04188500	Eagle Creek near Findlay, Ohio	0	0	0	0
113	04196000	Sandusky River near Bucyrus, Ohio	2.4	1.1	3.8	1.7
114	04196200	Broken Sword Creek at Nevada, Ohio	.2	0	.6	.1
115	04196500	Sandusky River near Upper Sandusky, Ohio	5.1	1.7	8.4	2.8
116	04197000	Sandusky River near Mexico, Ohio	16	7.6	23	11
117	04198000	Sandusky River near Fremont, Ohio	26	12	35	17
118	04198500	East Branch Huron River near Norwalk, Ohio	2.4	.6	3.6	1.2
119	04199000	Huron River at Milan, Ohio	10	4.6	14	6.8
120	04199300	Vermilion River at Clarksfield, Ohio	.2	0	.5	0
121	04199500	Vermilion River near Vermilion, Ohio	1.5	.1	3.3	.3
122	04200000	East Branch Black River at Elyria, Ohio	.4	0	1.5	.1
123	04200500	Black River at Elyria, Ohio	7.2	3.6	9.7	4.9
124	04201400	West Branch Rocky River at West View, Ohio	1.7	.3	3.4	.6
125	04201500	Rocky River near Berea, Ohio	5.5	1.4	9.8	2.4
126	04207200	Tinkers Creek at Bedford, Ohio	13	7.5	20	9.9
127	04208900	Aurora Branch near Chagrin Falls, Ohio	7.6	3.9	11	5.6
128	04209000	Chagrin River at Willoughby, Ohio	26	13	36	19
129	04210000	Phelps Creek near Windsor, Ohio	.5	.2	.7	.4
130	04212000	Grand River near Madison, Ohio	5.4	.9	9.5	2.8
131[a]	03261500	Great Miami River at Sidney, Ohio	37.0	21.0	44.0	26.0
132[a]	03262650	Spring Creek near Troy, Ohio	.1	0	.2	0

[a] These stations were added after the original data set.

The State ultimately was divided into five regions using the techniques described above. These regions are shown on figure 1. Multiple-regression equations for estimating $Q_{7,2}$, $Q_{7,10}$, $Q_{30,2}$, and $Q_{30,10}$ were developed for each region individually. These equations are presented in tables 4 through 8.

The equations for region 2 in table 5 are seen to differ from the general form of the regression equation presented earlier. The variable I (infiltration index) was found to explain a larger proportion of the variation in the dependent variable when it was not log transformed; consequently, this variable appears as a power of ten in the exponential form of the equation.

The variables that best explained the variation in the $Q_{j,k}$ flows differ from region to region. Only the drainage-area variable (A) is in all equations. The mean basin elevation (E) was significant in equations for regions 3 and 4. It is likely that this variable acts as a surrogate for some other characteristic that is not explicitly represented in the equations. The infiltration index (I) was only significant for region 2; consequently, a hydrologic soil-grouping map (fig. 3) is included for region 2 only.

Site numbers 56 and 126 were not used in determining the multiple-regression equations for regions 3 and 1, respectively. These sites were found to exhibit considerably different low-flow characteristics than the other sites within their respective regions; however, there is no evidence to suggest that they were improperly partitioned. Site number 56 (Little Scioto River above Marion, Ohio) exhibited low-flow yields (in cubic feet per second per square mile) that were 14 to 18 times smaller than the median yield for region 3. Site number 126 (Tinkers Creek at Bedford, Ohio) exhibited low-flow yields that were seven to eight times larger than the median yield for region 1. Because these sites are not representative of their regions, the authors believe that their omission was justified.

Table 4.--Equations for annual low flows of unregulated streams in Region 1

$$Y = C(A)^{b_1}(S)^{b_2}(F+1)^{b_3}(P-30)^{b_4} - 0.1$$

Low-flow characteristic (Y)	Regression constant (C)	Regression coefficients				Coefficient of determination (R^2)	Standard error of estimate (percent)	Residual degrees of freedom
		b_1	b_2	b_3	b_4			
$Q_{7,2}$	8.75×10^{-7}	1.33	1.54	-0.91	3.93	0.65	79.4	22
$Q_{7,10}$	2.68×10^{-7}	1.13	1.69	-1.37	4.98	.51	106	22
$Q_{30,2}$	1.66×10^{-6}	1.37	1.42	-0.74	3.64	.67	73.2	22
$Q_{30,10}$	4.22×10^{-7}	1.18	1.50	-1.18	4.84	.60	90.0	22

where: A = drainage area, in square miles

S = main-channel slope, in feet per mile

F = percentage of drainage area occupied by forest

P = mean annual precipitation, in inches.

Twenty-two percent of the sites analyzed were partial-record sites.

Table 5.--Equations for annual low flows of unregulated streams in Region 2

$$Y = C(A)^{b_1}(S)^{b_2} \times 10^{(b_3 \times I)} - 0.1$$

Low-flow characteristic (Y)	Regression constant (C)	Regression coefficients			Coefficient of determination (R^2)	Standard error or estimate (percent)	Residual degrees of freedom
		b_1	b_2	b_3			
$Q_{7,2}$	5.76×10^{-4}	1.37	0.52	0.38	0.93	48.5	33
$Q_{7,10}$	7.76×10^{-5}	1.52	.63	.45	.93	53.8	33
$Q_{30,2}$	9.02×10^{-4}	1.35	.51	.35	.94	42.6	33
$Q_{30,10}$	1.47×10^{-4}	1.47	.62	.42	.92	55.3	33

where: A = drainage area, in square miles

S = main-channel slope, in feet per mile

I = index of relative infiltration.

Thirty-five percent of the sites analyzed were partial-record sites.

Table 6.--Equations for annual low flows of unregulated streams in Region 3

$$Y = C(A)^{b_1}(E)^{b_2} - 0.1$$

Low-flow characteristic (Y)	Regression constant (C)	Regression coefficients		Coefficient of determination (R^2)	Standard error of estimate (percent)	Residual degrees of freedom
		b_1	b_2			
$Q_{7,2}$	9.53×10^{-2}	0.74	-7.80	0.73	56.7	11
$Q_{7,10}$	7.22×10^{-2}	.60	-11.18	.67	71.3	11
$Q_{30,2}$	1.02×10^{-1}	.80	-7.06	.80	45.9	11
$Q_{30,10}$	7.74×10^{-2}	.66	-10.64	.76	57.9	11

where: A = drainage area, in square miles

E = mean basin elevation, in thousands of feet above sea level.

Thirty-six percent of sites analyzed were partial-record sites.

Table 7.--Equations for annual low flows of unregulated streams in Region 4

$$Y = C(A)^{b_1}(E)^{b_2}(F+1)^{b_3} - 0.1$$

Low-flow characteristic (Y)	Regression constant (C)	Regression coefficients				Coefficient of determination (R^2)	Standard error of estimate (percent)	Residual degrees of freedom
		b_1	b_2	b_3	b_4			
$Q_{7,2}$	8.54×10^{-3}	1.43	5.81	-0.44	--	0.93	51.1	24
$Q_{7,10}$	3.68×10^{-3}	1.48	7.72	-0.53	--	.94	54.4	24
$Q_{30,2}$	1.80×10^{-2}	1.36	4.02	-0.38	--	.94	45.3	24
$Q_{30,10}$	6.92×10^{-3}	1.44	5.99	-0.53	--	.92	58.3	24

where: A = drainage area, in square miles

F = percentage of drainage area occupied by forest

E = mean basin elevation, in thousands of feet above sea level.

Sixty-eight percent of the sites analyzed were partial-record sites.

Table 8.--Equations for annual low flows of unregulated streams in Region 5

$$Y = C(A)^{b_1}(L)^{b_2} - 0.1$$

Low-flow characteristic (Y)	Regression constant (C)	Regression coefficients		Coefficient of determination (R^2)	Standard error of estimate (percent)	Residual degrees of freedom
		b_1	b_2			
$Q_{7,2}$	2.67×10^{-1}	2.19	-2.49	0.75	88.3	21
$Q_{7,10}$	7.44×10^{-1}	2.57	-3.72	.69	129	21
$Q_{30,2}$	2.54×10^{-1}	2.12	-2.19	.75	73.9	21
$Q_{30,10}$	4.96×10^{-1}	2.60	-3.48	.79	93.6	21

where: A = drainage area, in square miles

L = main-channel length, in miles, from site to basin divide.

Thirty-five percent of sites analyzed were partial-record sites.

Assessment of the Equations

Tests for Constant Residual Variance and Collinearity

Tests for constancy of residual variance were conducted by plotting the regression residuals against the corresponding estimate of the dependent variable. Constant residual variance is characterized by a uniform band of points around the line corresponding to the zero residual. All equations presented in this report exhibited constant residual variance.

Tests for collinearity, the condition where regressor variables exhibit nearly linear dependencies, were conducted by requesting the collinearity diagnostics in the SAS procedure REG. Collinearities, if severe, can cause appreciable round-off errors in the regression calculations. More moderate collinearities have a destabilizing effect on estimates of the regression coefficients. This, however, does not necessarily have a negative effect on the ability of the regression model to satisfactorily predict the dependent variable.

A moderate level of collinearity was found to exist between the variables Area (A) and Length (L), which are present together in equations for region 5. The extent to which this harms the accuracy with which these equations can be used to estimate the $Q_{j,k}$ discharges is unknown.

No other equations exhibited multicollinearities of a sufficient magnitude to degrade estimates.

Accuracy of the Equations

The standard error of estimate is a measure of the average deviation of the observed values of the dependent variable from the regression line. The standard error of estimate reported for each of the multiple-regression equations in tables 4 trough 8 can be used as a crude indicator of the level of accuracy with which future estimates can be made; however, the actual accuracy with which future estimates can be made will likely be somewhat lower.

It is apparent from the standard errors in table 8 that low-flow discharge estimates for region 5 are likely to be very inaccurate. Consequently, the region 5 equations should be used only when this level of inaccuracy is tolerable for the desired application.

The coefficient of determination (R^2), a measure of the effectiveness of the chosen basin characteristics in explaining the observed variations in the $Q_{j,k}$ discharge, is reported for each of the multiple-regression equations in tables 4 through 8 as an aid in assessing their adequacy.

Sensitivity Analysis

The basin-characteristics data used in the multiple-regression equations must generally be computed or estimated from maps. These data are therefore subject to errors in measurement or judgment. Sensitivity analyses were performed to illustrate the effects of such errors on the estimated $Q_{7,2}$ discharge.

Sensitivity analyses were conducted using the $Q_{7,2}$ equations for each of the five regions. For a given regional equation, one variable at a time was allowed to vary about its regional mean while the rest of the variables in the equation were held constant at their regional mean values. In this way, it was possible to measure the effects on the $Q_{7,2}$ discharge that are attributable to a change in each of the independent variables. Figures 4 through 8 show the percentage of change in the estimated $Q_{7,2}$ value as a result of varying the independent variables from +50 percent to -50 percent of their regional means. (These data also are presented in abbreviated form in table 9.) The dependent variable ($Q_{7,2}$) will be least affected by changes in independent variables that plot closest to the horizontal dashed line. Conversely, the dependent variable is most sensitive to changes in variables that plot farthest from the horizontal dashed line. For example, the sensitivity plot for region 1 (fig. 4) shows that the 7-day, 2-year discharge is very sensitive to changes in the precipitation variable (P). Consequently, it would be wise to spend extra time and effort to accurately measure this variable.

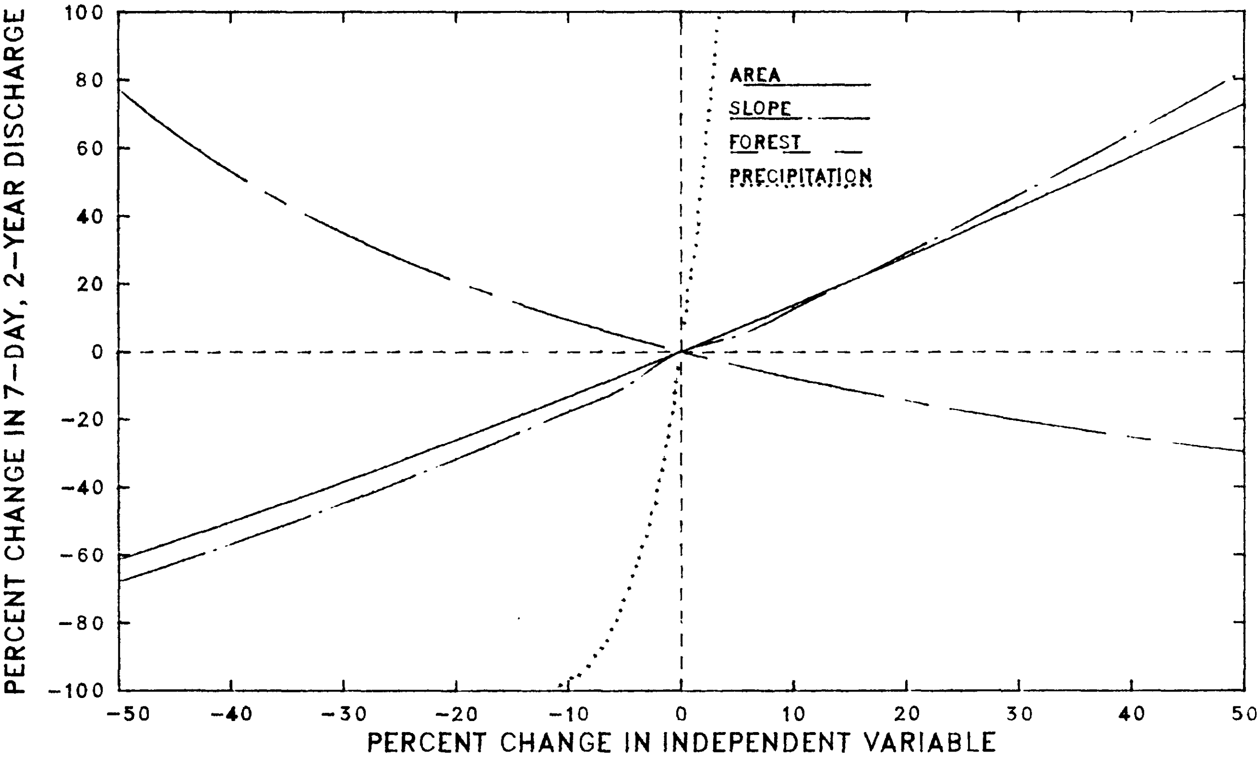

Figure 4.--Sensitivity of computed 7-day, 2-year discharge to change in independent variables in the multiple-regression equation for region 1.

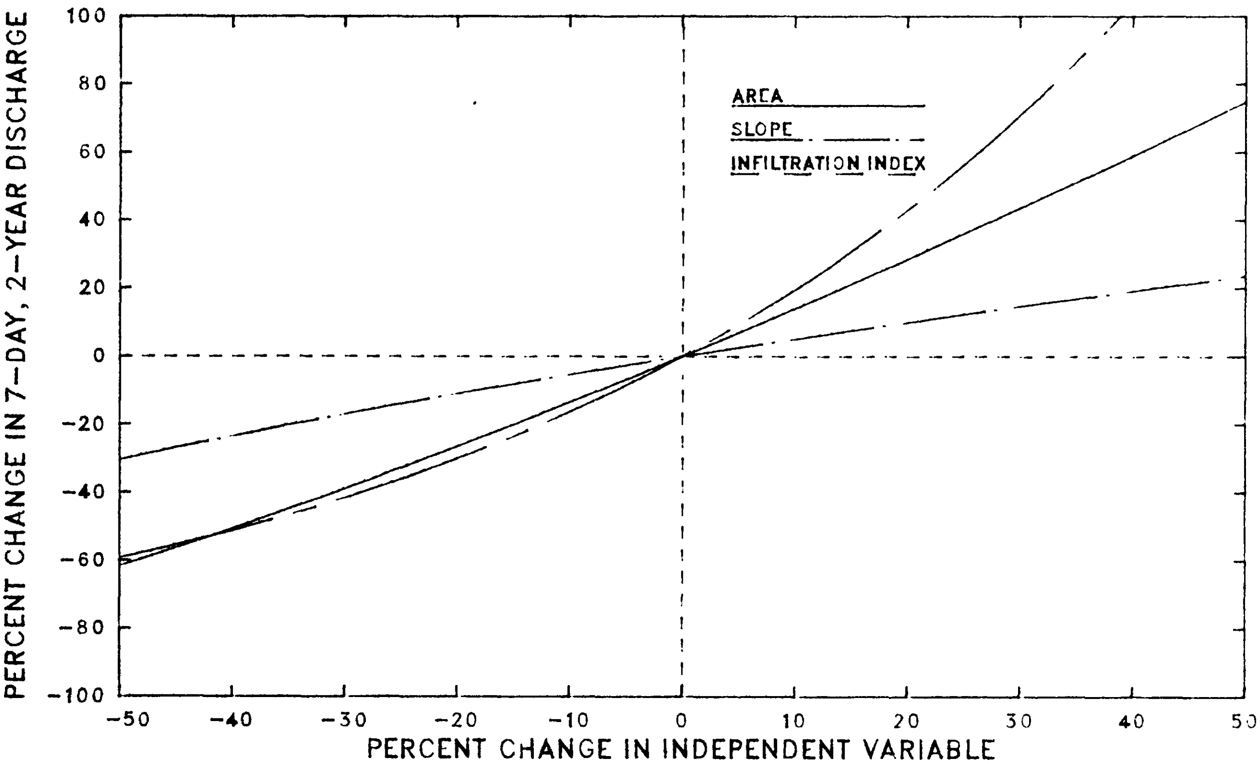

Figure 5.--Sensitivity of computed 7-day, 2-year discharge to change in independent variables in the multiple-regression equation for region 2.

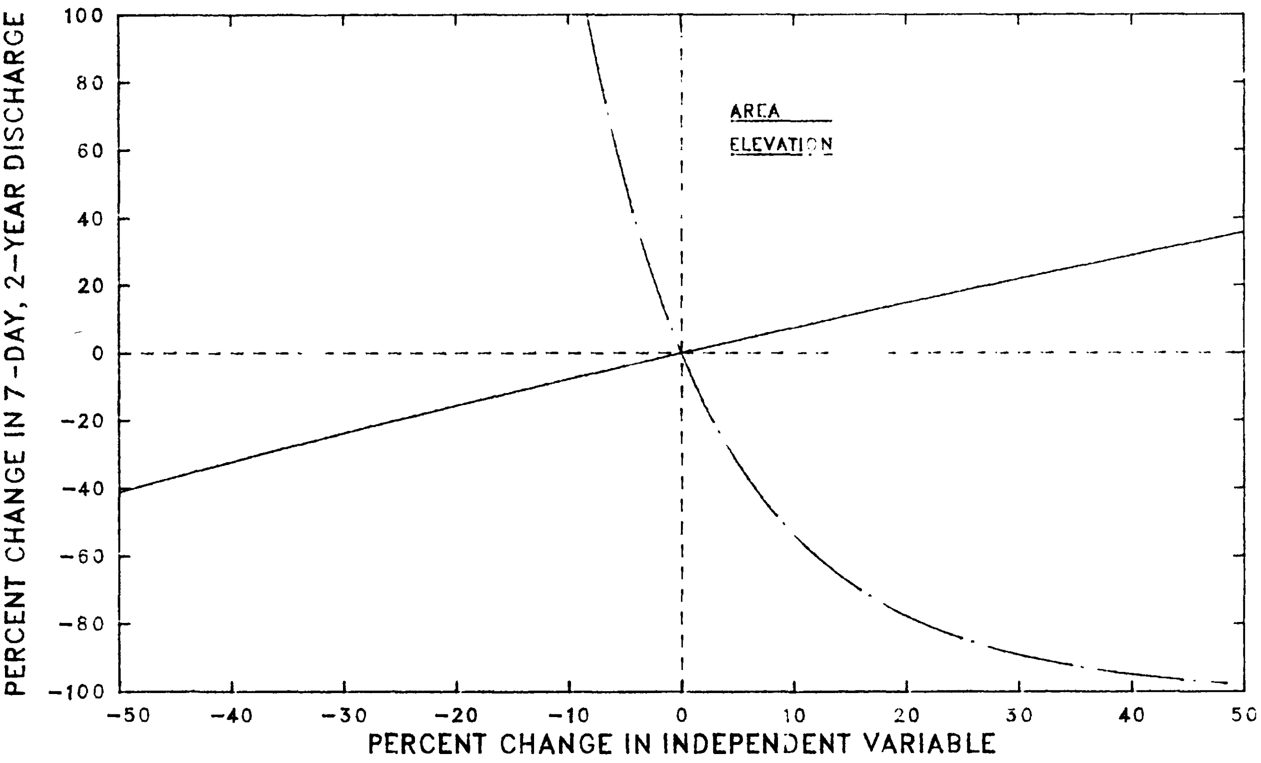

Figure 6.--Sensitivity of computed 7-day, 2-year discharge to change in independent variables in the multiple-regression equation for region 3.

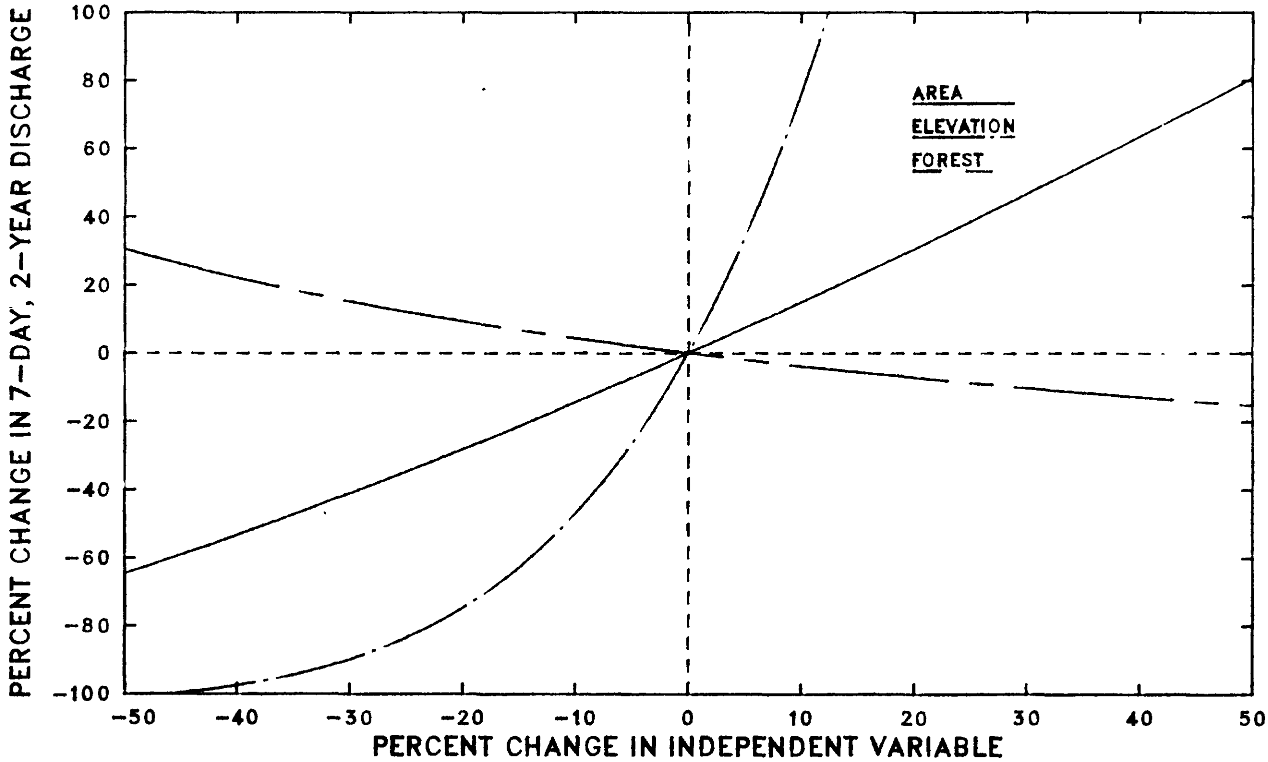

Figure 7.—Sensitivity of computed 7-day, 2-year discharge to change in independent variables in the multiple-regression equation for region 4.

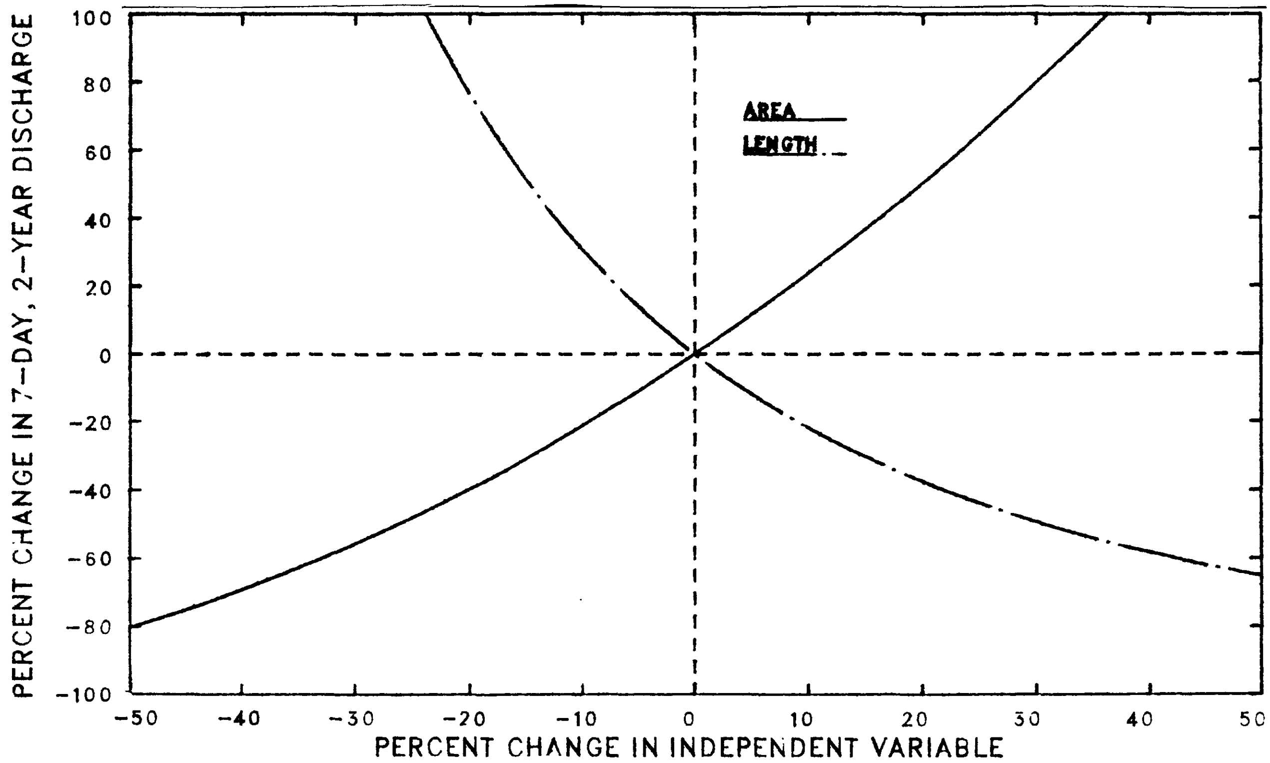

Figure 8.—Sensitivity of computed 7-day, 2-year discharge to change in independent variables in the multiple-regression equation for region 5.

Table 9.--Sensitivity of computed 7-day, 2-year discharge to changes in the independent variables

[A, drainage area, in square miles; E, mean basin elevation, in thousands of feet above sea level; F, percentage of area occupied by forest; I, index of relative infiltration L, main-channel length, in miles, from site to basin divide; P, average annual precipitation, in inches; S, main-channel slope, in feet per mile]

Region	Independent variable	Percent change in the 7-day, 2-year low-flow corresponding to percent change in independent variable						Observed range (percent of mean)
		-50	-30	-10	10	30	50	
1	A	-61	-38	-13	14	42	73	-97 / +466
	S	-68	-45	-18	13	46	83	-84 / +127
	F	77	35	9.4	-7.9	-20	-38	-94 / +193
	P	--	--	-97	460	4,500	18,000	-8 / + 11
2	A	-62	-39	-14	14	44	75	-99 / +387
	S	-30	-17	-5.4	5.1	15	24	-81 / +280
	I	-59	-42	-16	20	71	144	-49 / +369
3	A	-41	-24	-7.6	7.5	22	36	-99 / +194
	E	23,000	1,600	130	-53	-89	-98	-20 / + 24
4	A	-65	-41	-14	15	47	81	-86 / +333
	E	-100	-90	-47	76	370	980	-22 / + 11
	F	31	15	4.3	-3.8	-10	-15	-95 / +157
5	A	-80	-56	-21	24	80	147	-97 / +280
	L	470	150	31	-22	-49	-65	-85 / +150

ESTIMATING LOW FLOWS FOR OHIO STREAMS

Determining the Proper Method of Estimation

Low-flow estimates for a particular point on a stream differ in reliability depending on the methods used to obtain the estimates. The most reliable means of estimating low-flows is to use frequency curves developed for a long-term continuous-record gaging station located at or near the point of interest. If the point of interest is not located at the same position on the stream as the gaging station, then the flow estimate should be made by multiplying the $Q_{j,k}$ discharge at the gaged site by the drainage-area ratio of the two sites. Transference by the drainage-area ratio method should be attempted only if (1) the drainage areas at the point of interest and at the gaging station differ by no more than 10 to 15 percent, and (2) no significant differences in the geology of the intervening area exist.

Under the proper conditions, the next most reliable method of estimating low flows requires that low-flow partial-record measurements be made at the point of interest and correlated to concurrent base flows at a suitable index station (Riggs, 1972). Accuracy of partial-record estimates is a function of many factors such as the number and magnitudes of the low-flow measurements, the correlation coefficient between low-flow measurements at the partial-record site and baseflow at the index station, and the number of years of record at the index station (for more information see Stedinger and Thomas, 1985).

If lack of suitable data precludes both of the above methods, then and only then, should the multiple-regression equations be used to estimate the low-flow discharge.

Limitations on Use of the Multiple-Regression Equations

General limitations on use of the multiple-regression equations include:

1. The equations are not applicable to streams with significant low-flow regulation.

2. The equations are only applicable in the region for which they were developed. Within a region, the equations should only be applied to sites having basin characteristics within the range of values employed in their development. The level of accuracy implied by the standard errors of estimate may not be realized if values outside this range are used. Maximum, minimum, mean, and median values of the basin characteristic variables are presented for each of the regions in table 10 so that the above limitation may be considered.

3. The multiple-regression equations presented in this report are not intended to act as a substitute for more reliable methods of estimating low-flows described in the previous section of this report.

Procedures for Using the Multiple-Regression Equations

The procedure for using the multiple-regression equations to estimate low-flow discharges on unregulated streams is as follows:

1. Identify the $Q_{j,k}$ flow characteristic of interest.

2. Determine the region in which the point of interest is located (fig. 1).

Table 10.--Statistics of regional basin characteristics

[A, drainage area, in square miles; E, mean basin elevation, in thousands of feet above sea level; F, percentage of area occupied by forest; I, index of relative infiltration; L, main-channel length, in miles, from site to basin divide; P, average annual precipitation, in inches; S, main-channel slope, in feet per mile. Dash, not applicable.]

		Variable						
Region	Statistic	A	L	E	S	F	P	I
1	Maximum---	1250	118	1160	20.7	36.0	41.0	--
	Mean------	221	36.7	964	9.13	12.3	36.7	--
	Median----	109	30.3	971	7.37	9.80	36.0	--
	Minimum---	7.85	8.57	763	1.45	0.70	34.0	--
2	Maximum---	926	67.1	1300	49.8	32.0	43.0	7.50
	Mean------	190	25.8	1010	13.1	5.51	40.1	2.03
	Median----	72.8	20.5	1010	10.1	4.10	40.5	1.69
	Minimum---	2.08	3.52	767	2.53	1.00	34.0	1.00
3	Maximum---	567	70.7	1250	140	96.0	39.5	--
	Mean------	193	34.6	1010	17.4	15.0	37.1	--
	Median----	154	34.9	1010	7.02	3.4	37.0	--
	Minimum---	1.35	1.67	808	1.53	0.10	34.5	--
4	Maximum---	455	50.5	1190	22.3	33.4	41.0	--
	Mean------	105	20.8	1070	9.84	13.0	38.2	--
	Median----	77.5	21.8	1110	7.70	10.0	38.3	--
	Minimum---	14.3	4.78	839	4.40	0.60	35.0	--
5	Maximum---	585	71.6	1190	67.0	98.0	43.0	--
	Mean------	153	28.1	979	16.4	29.1	40.2	--
	Median----	125	27.9	989	10.2	15.8	40.0	--
	Minimum---	4.99	4.24	715	2.81	1.00	34.0	--

3. Measure the basin characteristics required as input to the proper regional multiple-regression equation. The basin characteristics are defined in this report on pages 10 and 11.

4. Substitute the necessary basin characteristics into the appropriate multiple-regression equation and make the indicated calculation.

Example

Determine the $Q_{7,10}$ discharge for a site on Buck Creek near New Moorefield, Ohio, using the procedure outlined above:

1. The low-flow characteristic of interest has been identified as the $Q_{7,10}$ discharge.

2. Buck Creek is tributary to the Mad River, which is tributary to the Great Miami River. Buck Creek is in region 2 (fig. 1).

3. Table 5 shows that drainage area (A), main-channel slope (S), and the infiltration index (I) are the basin characteristics required as input to the multiple-regression equations for region 2. The basin characteristics are as follows:

   a. Drainage area (A) = 65.3 square miles

b. Main-channel slope (S):

Main-channel length (L) = 12.7 miles

Elevation of channel at 0.85(L) = 1,205 feet

Elevation of channel at 0.10(L) = 1,030 feet

$$S = \frac{E_{0.85(L)}-E_{0.10(L)}}{0.75(L)} = \frac{1,205-1,030}{0.75(12.7)}$$

S = 18.4 feet per mile

c. Index of relative infiltration (I):

Soil group	Fraction of area (Af)	Weighting factor (Wf)	Af x Wf
1	0.52	1	0.52
2	.10	8	.80
3	.36	7.5	2.7
4	.02	2.3	.046
Total	1	--	4.066

I = 4.07

4. Substitute the values obtained above into the appropriate multiple-regression equation to determine the $Q_{7,10}$ discharge:

$$Q_{7,10} = 7.76x10^{-5}\ (A)^{1.52}\ (S)^{0.63}\ 10^{(0.45(I))} - 0.1$$

$$= 7.76x10^{-5}\ (65.3)^{1.52}\ (18.4)^{0.63}\ 10^{(0.45(4.07))} - 0.1$$

$Q_{7,10}$ = 18.8 cubic feet per second

The $Q_{7,10}$ discharge determined for Buck Creek at New Moorefield, Ohio, from continuous-record gaging station data is 15.0 cubic feet per second (ft^3/s). In actual practice, the 15.0 ft^3/s value would be the accepted value because it was determined using the most reliable method. The sole purpose for using the multiple-regression equation in this instance was to illustrate the computational process.[1]

SUMMARY

Multiple-regression equations were developed to estimate selected low-flow characteristics for streams in Ohio where little or no discharge data have been collected. Attempts at developing equations that would be applicable statewide were not successful because of excessively large standard errors; consequently, regional equations were developed that relate low flows to selected characteristics of the basins. Five regional equations were developed for estimating low flows with average recurrence intervals of 2 and 10 years and durations of 7 and 30 days.

Standard errors of equations for regions 1 through 4 ranged from 43 to 106 percent. These equations should be sufficiently accurate for many uses. Large standard errors for region 5 equations (74 to 129 percent) severely limit their useful applications.

Use of the regression equations is not recommended if a more reliable means exists for estimating low-flow characteristics for the site of interest. The equations are intended to be used as a

---

[1] A companion program is available (for IBM-PC and compatible computers) to aid in performing these calculations. For information on obtaining copies, contact District Chief, U.S. Geological Survey, 975 W. Third Avenue, Columbus, OH 43212-3192.

supplement to the more reliable low-flow estimation techniques, not as a substitute.

The equations are not applicable to streams with significant low-flow regulation, nor are they applicable outside the region for which they were developed. Within their respective regions, the equations should only be applied to sites having basin characteristics within the approximate range of values employed in the development of the equations.

REFERENCES CITED

Armbruster, J. T., 1976a, Technical manual for estimating low-flow frequency characteristics of streams in the Susquehanna River basin: U.S. Geological Survey Water-Resources Investigations Report 76-51, 51 p.

----- 1976b, An infiltration index useful in estimating low-flow characteristics of drainage basins: U.S. Geological Survey Journal of Research, v. 4, no. 5, p. 533-38.

Flippo, H. N., Jr., 1982, Technical manual for estimating low-flow characteristics of Pennsylvania Streams: Pennsylvania Department of Environmental Resources Bulletin 15, 86 p.

Johnson, D. P., and Metzker, K. D., 1981, Low-flow characteristics of Ohio streams: U.S. Geological Survey Open-File Report 81-1195, 285 p.

Riggs, H. C., 1972, Low-flow Investigations: U.S. Geological Survey Techniques of Water-Resources Investigations, bk. 4, chap. B1, 18 p.

SAS Institute Inc., 1982, SAS User's Guide--Statistics, 1982 edition: Cary, N.C., SAS Institute, 584 p.

Stedinger, J. R., and Thomas, W. O., Jr., Low-flow Frequency Estimation Using Base-Flow Measurements, U. S. Geological Survey Open-File Report 85-95, 22 p.

☆U. S. GOVERNMENT PRINTING OFFICE 1987 - 750-3