COST-EFFECTIVENESS OF THE FEDERAL STREAM-GAGING PROGRAM IN VIRGINIA

By D. H. Carpenter

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 85-4345

Towson, Maryland

1985

UNITED STATES DEPARTMENT OF THE INTERIOR

DONALD PAUL HODEL, Secretary

GEOLOGICAL SURVEY

Dallas L. Peck, Director

For additional information write to:

District Chief
U.S. Geological Survey
208 Carroll Building
8600 La Salle Road
Towson, Maryland 21204

Copies of this report can be purchased from:

Open-File Services Section
Western Distribution Branch
U.S. Geological Survey
Box 25425, Federal Center
Denver, Colorado 80225
Telephone: (303) 234-5888

PREFACE

The collection of surface-water data is a major activity of the U.S. Geological Survey's (USGS) Water Resources Division (WRD). Approximately $40 million was spent in 1982 by WRD in cooperation with State and local governments and other Federal agencies in the collection of these data. This major expenditure of funds for hydrologic data collection should be evaluated periodically with respect to the needs of the data users and the utility of the data. It is essential that a rigorous analysis be made of the stream-gaging program to assure maximum cost-effectiveness. The USGS is undertaking a nationwide analysis of its stream-gaging program over a 5-year period. The results from such an analysis should satisfy both local and national water-data needs within budget constraints while maintaining quality control.

This report for the State of Virginia is one in a series of statewide reports describing this analysis. The techniques and methods being utilized in the nationwide analysis are described and documented in this report as applied to the Virginia stream-gaging program.

Analysis of the stream-gaging program is designed to define and document the most cost-effective means of furnishing streamflow information. The stream-gaging activity is no longer considered a network of observation points, but, rather, an information system in which data are provided by both observation and synthesis. Alternative methods of providing streamflow information such as flow routing and statistical methods are investigated as to their cost-effectiveness, accuracy, and information content.

Recently, new techniques for evaluating the cost-effectiveness of data-collection programs have been developed. These techniques, Kalman filtering and mathematical programming, are utilized to define strategies for operating the stream-gaging program so that the uncertainty in the streamflow records is minimized. The USGS first applied these techniques to a stream-gaging program in the Lower Colorado River Basin. Subsequently, the techniques have been expanded and improved, and are being applied to the present nationwide study of the USGS stream-gaging program. No doubt these techniques will continue to be modified and improved over the duration of the study.

The analysis of the stream-gaging program is a part of the continuing effort of the USGS to evaluate the Nation's water resources. The national stream-gaging program that results from this analysis should be responsive to the needs of local, State, and Federal agencies and provide streamflow information in the most cost-effective manner.

CONTENTS

CONTENTS--Continued

ILLUSTRATIONS

TABLES

CONVERSION FACTORS AND ABBREVIATIONS

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 unit
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
square mile (mi^2)	2.590	square kilometer (km^2)
cubic foot (ft^3)	0.02832	cubic meter (m^3)
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)

COST-EFFECTIVENESS OF THE FEDERAL STREAM-GAGING PROGRAM IN VIRGINIA

By David H. Carpenter

ABSTRACT

This report documents the results of a cost-effectiveness study of the stream-gaging program in Virginia. Data uses and funding sources were identified for the 77 continuous stream gages currently being operated in Virginia by the U.S. Geological Survey with a budget of $446,000. Two stream gages were identified as producing data which may not be used sufficiently to warrant continuing their operation. Operation of these stations should be considered for discontinuation. Data collected at two other stations were identified as having uses primarily related to short-term studies; these stations also should be considered for discontinuation at the end of the data-collection phases of the studies. The remaining 73 stations should be kept in the program for the foreseeable future.

The current policy for operation of the 77-station program requires a budget of $446,000 per year. The average standard error of estimation of streamflow records is 10.1 percent. It was shown that this overall level of accuracy at the 77 sites could be maintained with a budget of $430,500 if resources were redistributed among the gages.

A minimum budget of $428,500 is required to operate the 77-gage program; a smaller budget would not permit proper service and maintenance of the gages and recorders. At the minimum budget, with optimized operation, the average standard error would be 10.4 percent. The maximum budget analyzed was $650,000, which resulted in an average standard error of 5.5 percent.

The study indicates that a major component of error is caused by lost or missing data. If perfect equipment were available, the standard error for the current program and budget could be reduced to 7.6 percent. This also can be interpreted to mean that the streamflow data have a standard error of this magnitude during times when the equipment is operating properly.

INTRODUCTION

The U.S. Geological Survey (USGS) is the principal Federal agency collecting surface-water data in the Nation. The collection of these data is a major activity of the Water Resources Division of the USGS. The data are collected in cooperation with State and local governments and other Federal agencies. The USGS presently (1984) operates approximately 8,000 continuous-record gaging stations throughout the Nation. Some of these records extend back to the turn of the century. Any activity of long standing, such as the collection of surface-water data, should be reexamined at intervals, if not continuously, because of changes in objectives, technology, or external constraints. The last systematic nationwide evaluation of the streamflow information program was completed in 1970 and is documented by Benson and Carter (1973). The USGS is presently undertaking another nationwide analysis of the stream-gaging program that will be completed over a 5-year period with 20 percent of the program being analyzed each year. The object of this analysis is to define and document the most cost-effective means of furnishing streamflow information.

As a first phase, for every continuous-record gaging station, the analysis identifies the principal uses of the data and relates these uses to funding sources. Gaged sites for which data are no longer needed are identified, as are deficient or unmet data demands. In addition, gaging stations are categorized as to whether the data are available to users in a real-time sense, on a provisional basis, or at the end of the water year.

The second phase of the analysis is to identify less costly methods of furnishing the needed information; among these are flow-routing models and statistical methods. The stream-gaging activity no longer is considered a network of observation points, but rather an integrated information system in which data are provided both by observation and synthesis.

The final part of the analysis involves the use of Kalman-filtering and mathematical-programming techniques to define strategies for operation of the necessary stations that minimize the uncertainty in the streamflow records for given operating budgets. Kalman-filtering techniques are used to compute uncertainty functions (relating the standard errors of computation or estimation of streamflow records to the frequencies of visits to the stream gages) for all stations in the analysis. A steepest descent optimization program uses these uncertainty functions, information on practical stream-gaging routes, the various costs associated with stream gaging, and the total operating budget to identify the visit frequency for each station that minimizes the overall uncertainty in the streamflow. The stream-gaging program that results from this analysis will meet the expressed water-data needs in the most cost-effective manner.

This report is organized into five sections; the first is an introduction to the stream-gaging activities in Virginia and to the study itself. The middle three sections each contain discussions of individual steps of the analysis. Because of the sequential nature of the steps and the dependence of subsequent steps on previous results, summaries of conclusions are given at the end of each middle section. The complete study is summarized in the final section.

Parts of this text have been excerpted from the prototype report for the nationwide network analyses; U.S. Geological Survey Water-Supply Paper 2244 (Fontaine and others, 1984).

The standard errors of estimate given in the report are those that would occur if daily discharges were computed through the use of methods described in this study. No attempt has been made to estimate standard errors for discharges that are computed by other means. Such errors could differ from the errors computed in the report. The magnitude and direction of the differences would be a function of methods used to account for shifting controls and for estimating discharges during periods of missing record.

History of the Stream-Gaging Program in Virginia

The U.S. Geological Survey program of surface-water investigations in Virginia began in 1895 when gaging stations were established on the Maury, North, and South Rivers. A few additional stations were established during the late 1890's. All work was financed with Federal funds.

In 1905, a short-term cooperative agreement with the Virginia Geological Survey resulted in the publication by the Virginia Geological Survey in 1906 of Bulletin 3, Hydrography of Virginia, which was a compilation of streamflow records prior to that date as well as water-surface profiles of some of the larger streams in the State.

By 1924, the surface-water network had gradually grown to 17 gaging stations. In February 1925, an office of the U.S. Geological Survey was established in Charlottesville. During 1925, an enlarged stream-gaging program developed in cooperation with the Virginia Geological Survey and the number of stations in operation nearly doubled, to 32 sites. The cooperative program continued with the Virginia Division of Water Resources and the Virginia Division of Mineral Resources with the number of streamflow stations increasing rather rapidly to 76 in 1930 and then more slowly to a peak of 166 in 1955.

In July 1957, the cooperative program was terminated by the State. The U.S. Geological Survey continued a surface-water data-collection program on a reduced scale with other local and Federal agencies assuming the role of cooperator. The network was reduced to about 100 gages while another network of about 60 gages was being operated by the Virginia Division of Water Resources.

A small-streams project was begun with the Virginia Department of Highways in 1964 and full-scale cooperative funding with the State was resumed in 1966, with independent (though coordinated) stream-gaging program operations. With the resumption of cooperation with the State, the District office was moved from Charlottesville to Richmond to be closer to the headquarters of the principal cooperator, the State Water Control Board.

The data-collection network expanded somewhat in recent years in response to the demand for water-resources information. The stream-gaging program reached another plateau in 1980 with 201 stations being operated, including 93 by the State. Excluding short-term special sites, the current surface-water program in Virginia is comprised of 77 gaging stations operated by the U.S. Geological Survey and 100 operated by the State Water Control Board.

The numbers of continuous stream gages operated through the history of the program are presented in figure 1.

Current Virginia Stream-Gaging Program

The areal distribution of the 77 stream gages in Virginia currently operated by the Mid-Atlantic District of the USGS is shown in figure 2. The cost of operating the stream-gaging program in fiscal year 1984 was $446,000.

Selected hydrologic data, including drainage area, period of record, and mean annual flow, for the 77 stations are given in table 1. Station identification numbers used throughout this report are the last seven digits of the USGS's eight-digit downstream-order station numbers; the first digit of the standard USGS station numbers for all stations used in this report is zero.

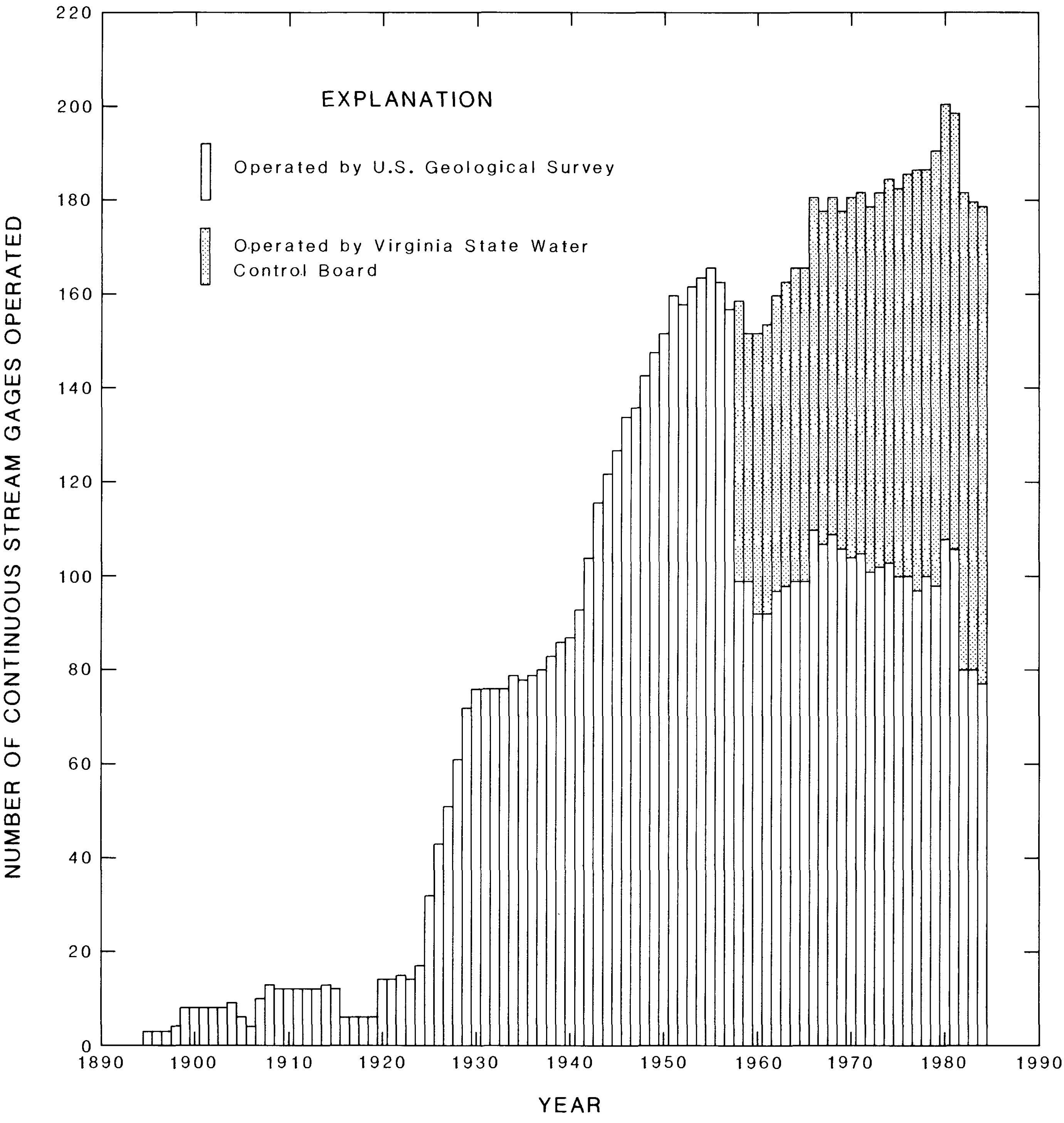


Figure 1.-- History of continuous stream gaging in Virginia.

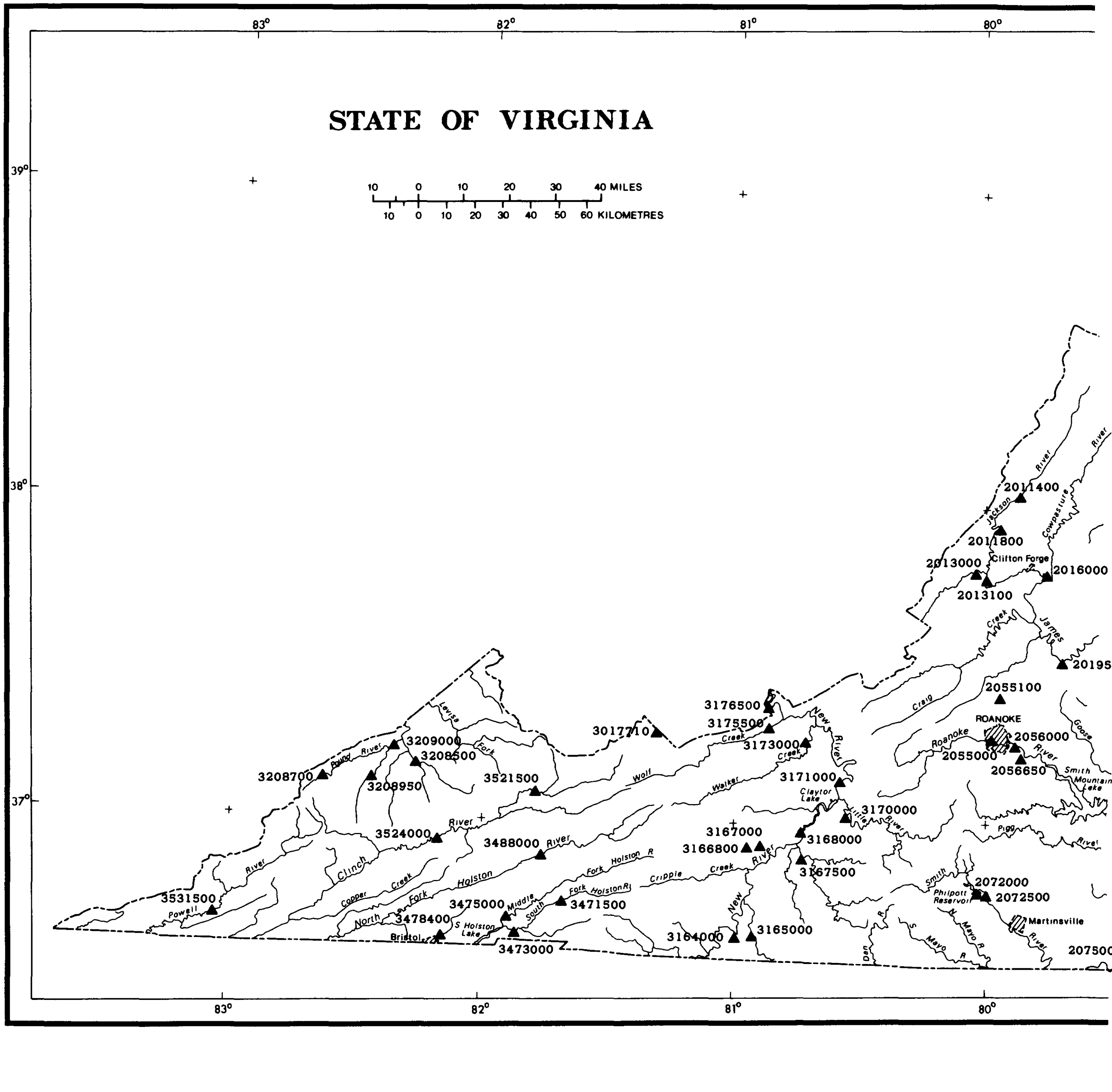
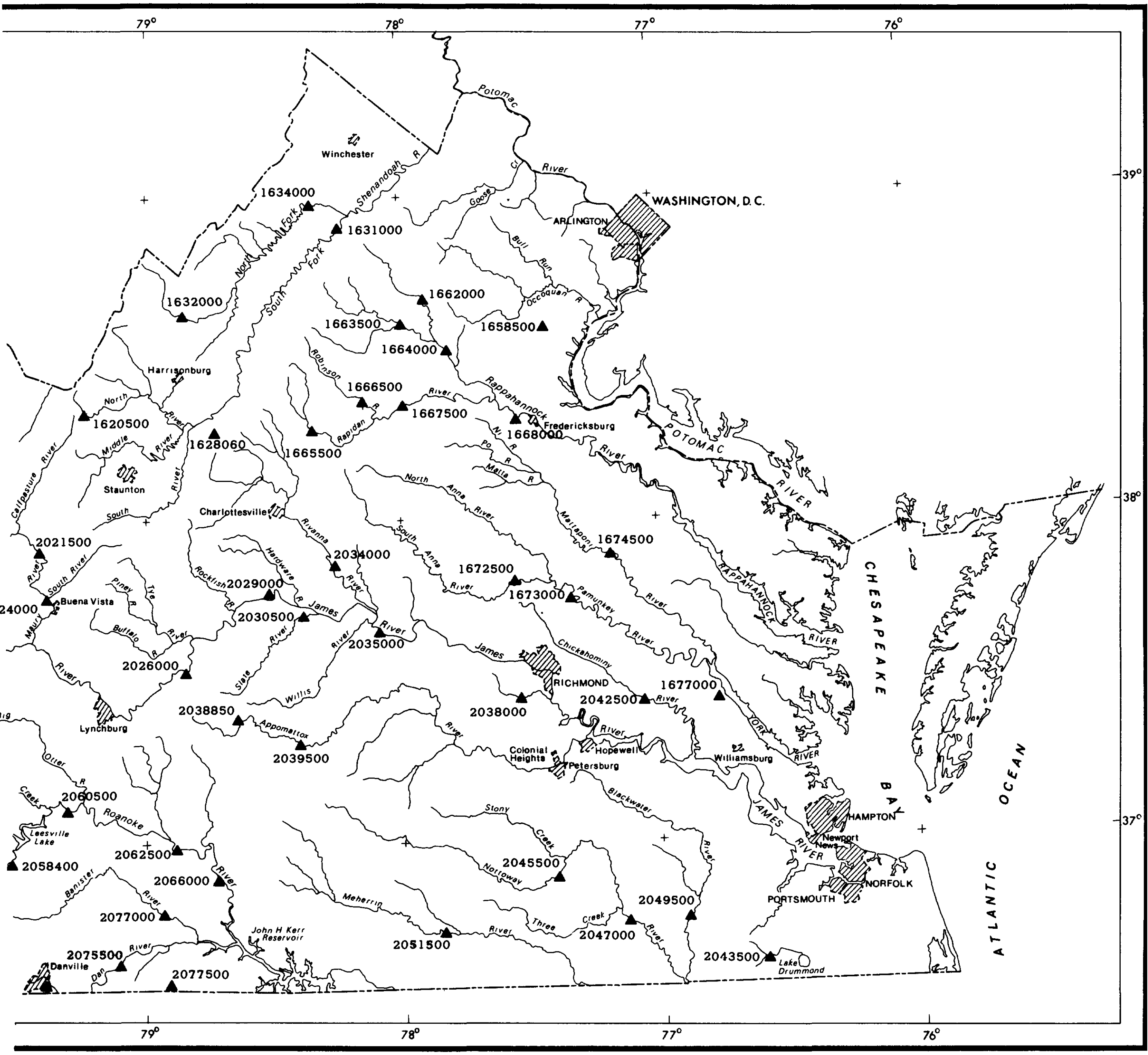


Figure 2.-- Location of stream gages.



79°
78°
77°
76°
39°
38°
37°
Winchester
WASHINGTON, D.C.
ARLINGTON
Potomac
River
Shenandoah
R.
North
Fork
South
Fork
Goose
Cr.
Bull
Run
Occoquan
R.
1634000
1631000
1632000
1662000
1663500
1658500
1664000
Harrisonburg
Robinson
1666500
River
R.
Rappahannock
1667500
1620500
North
River
1668000
Fredericksburg
Rapidan
1628060
1665500
Middle
River
POTOMAC
RIVER
Staunton
Calfpasture
River
Ni
R.
Po
R.
Matta
R.
North
Anna
River
Charlottesville
South
Rivanna
South
Anna
Mattaponi
1674500
2021500
2034000
River
South River
1672500
Piney
R.
Tye
Hardware
R.
Rockfish
2029000
1673000
Pamunkey
Buena Vista
2024000
Maury
James
2030500
River
Buffalo
R.
River
Slate
River
River
2035000
James
Chickahominy
River
RAPPAHANNOCK
RIVER
CHESAPEAKE
2026000
River
Willis
RICHMOND
1677000
2042500
River
Lynchburg
Big
2038850
Appomattox
2038000
River
YORK
RIVER
2039500
River
Colonial Heights
Hopewell
Petersburg
Williamsburg
Otter
R.
Creek
2060500
Roanoke
Blackwater
JAMES
RIVER
BAY
HAMPTON
OCEAN
Leesville
Lake
Stony
Creek
Newport
News
2058400
2062500
River
Nottoway
2045500
Banister
River
2066000
Meherrin
Three
Creek
2049500
NORFOLK
PORTSMOUTH
2077000
River
John H Kerr
Reservoir
River
2047000
River
ATLANTIC
2051500
2075500
Danville
Dan
River
2077500
2043500
Lake
Drummond

Table 1.--Selected hydrologic data for stations in the Virginia surface-water program

Station number	Station name	Drainage area (mi^2)	Period of record	Mean annual flow (ft^3/s)
1620500	North River near Stokesville, VA	17.2	October 1946-	25.7
1631000	South Fork Shenandoah River at Front Royal, VA	1,642	June 1899-September 1906, September 1930-	1,585
1632000	North Fork Shenandoah River at Cootes Store, VA	210	February 1925-	191
1634000	North Fork Shenandoah River near Strasburg, VA	768	March 1925-	583
1658500	South Fork Quantico Creek near Independent Hill, VA	7.64	May 1951-	6.85
1662000	Rappahannock River near Warrenton, VA	195	August 1942-	194
1663500	Hazel River at Rixeyville, VA	287	August 1942-	339
1664000	Rappahannock River at Remington, VA	620	October 1942-	674
1665500	Rapidan River near Ruckersville, VA	114	September 1942-	149
1666500	Robinson River near Locust Dale, VA	179	July 1943-	218
1667500	Rapidan River near Culpeper, VA	472	October 1930-	523
1668000	Rappahannock River near Fredericksburg, VA	1,596	September 1907-	1,655
1672500	South Anna River near Ashland, VA	394	October 1930-	364
1673000	Pamunkey River near Hanover, VA	1,081	October 1941-	992
1674500	Mattaponi River near Beulahville, VA	601	September 1941-	587
1677000	Ware Creek near Toano, VA	6.29	October 1979-October 1981 March 1982-	——1/
2011400	Jackson River near Bacova, VA	158	October 1974-	169
2011800	Jackson River below Gathright Dam, near Hot Springs, VA	345	October 1973-	455
2013000	Dunlap Creek near Covington, VA	164	October 1928-	166
2013100	Jackson River below Dunlap Creek, at Covington, VA	614	October 1974-	721
2016000	Cowpasture River near Clifton Forge, VA	461	March 1925-	524
2019500	James River at Buchanan, VA	2,075	February 1898-	2,466
2021500	Maury River at Rockbridge Baths, VA	329	October 1928-	371
2024000	Maury River near Buena Vista, VA	646	October 1938-	655
2026000	James River at Bent Creek, VA	3,683	October 1924-	4,174
2029000	James River at Scottsville, VA	4,584	October 1924-	5,128

1/ No mean annual flow published; less than 5 years of streamflow record.

Table 1.--Selected hydrologic data for stations in the Virginia surface-water program-- Continued

Station number	Station name	Drainage area (mi^2)	Period of record	Mean annual flow (ft^3/s)
2030500	Slate River near Arvonia, VA	226	April 1926-	230
2034000	Rivanna River at Palmyra, VA	664	October 1933-	716
2035000	James River at Cartersville, VA	6,257	October 1898-	7,053
2038000	Falling Creek near Chesterfield, VA	32.8	October 1955-	34.1
2038850	Holiday Creek near Andersonville, VA	8.53	April 1966-	9.28
2039500	Appomattox River at Farmville, VA	303	March 1926-	286
2042500	Chickahominy River near Providence Forge, VA	248	January 1942-	262
2043500	Cypress Swamp at Cypress Chapel, VA	23.8	October 1953-September 1971, March 1978-	27.9
2045500	Nottoway River near Stony Creek, VA	579	October 1929-	560
2047000	Nottoway River near Sebrell, VA	1,421	September 1941-	1,353
2049500	Blackwater River near Franklin, VA	617	August 1944-	641
2051500	Meherrin River near Lawrenceville, VA	552	October 1928-	496
2055000	Roanoke River at Roanoke, VA	395	February 1899-	371
2055100	Tinker Creek near Daleville, VA	11.7	April 1956-	11.4
2056000	Roanoke River at Niagara, VA	512	July 1926-	509
2056650	Back Creek near Dundee, VA	56.8	July 1974-	59.0
2058400	Pigg River near Sandy Level, VA	350	May 1963-	363
2060500	Roanoke River at Altavista, VA	1,789	August 1930-	1,786
2062500	Roanoke River at Brookneal, VA	2,415	April 1923-	2,379
2066000	Roanoke River at Randolph, VA	2,977	August 1900-September 1906, October 1927-September 1930, October 1950-	3,048
2072000	Smith River near Philpott, VA	216	August 1946-	276
2072500	Smith River at Bassett, VA	259	April 1939-	328
2075000	Dan River at Danville, VA	2,050	August 1934-	2,314
2075500	Dan River at Paces, VA	2,550	November 1950-	2,699
2077000	Banister River at Halifax, VA	547	September 1904-December 1905, October 1928-	504
2077500	Hyco River near Denniston, VA	289	October 1928-September 1934, October 1950-	250
3164000	New River near Galax, VA	1,131	October 1929-	1,895
3165000	Chestnut Creek at Galax, VA	39.4	October 1944-	67.1
3166800	Glade Creek at Grahams Forge, VA	7.15	May 1976-	0.94

Table 1.--Selected hydrologic data for stations in the Virginia surface-water program-- Continued

Station number	Station name	Drainage area (mi^2)	Period of record	Mean annual flow (ft^3/s)
3167000	Reed Creek at Grahams Forge, VA	247	July 1908-September 1916, January 1927-	267
3167500	Big Reed Island Creek near Allisonia, VA	278	August 1908-September 1916, April 1939-	400
3168000	New River at Allisonia, VA	2,202	September 1929-	3,198
3170000	Little River at Graysonton, VA	300	October 1928-	363
3171000	New River at Radford, VA	2,748	October 1907-September 1915, August 1939-	3,848
3173000	Walker Creek at Bane, VA	305	March 1938-	326
3175500	Wolf Creek near Narrows, VA	223	July 1908-September 1916, March 1938-	300
3176500	New River at Glen Lyn, VA	3,768	August 1927-	4,986
3177710	Bluestone River at Falls Mills, VA	44.2	October 1980-	—1/
3207800	Levisa Fork at Big Rock, VA	297	October 1967-	393
3208500	Russell Fork at Haysi, VA	286	July 1926-	333
3208700	NF Pound River at Pound, VA	18.5	October 1961-	29.3
3208950	Cranes Nest River near Clintwood, VA	66.5	October 1963-	82.9
3209000	Pound River below Flannagan Dam, near Haysi, VA	221	July 1926-	277
3471500	SF Holston River at Riverside, near Chilhowie, VA	76.1	October 1920-December 1931, July 1942-	112
3473000	SF Holston River near Damascus, VA	301	October 1931-	479
3475000	MF Holston River near Meadowview, VA	211	October 1931-September 1953, May 1976-	242
3478400	Beaver Creek at Bristol, VA	27.7	July 1957-	35.8
3488000	NF Holston River near Saltville, VA	222	June 1907-December 1908, October 1920-	302
3521500	Clinch River at Richlands, VA	137	October 1945-	194
3524000	Clinch River at Cleveland, VA	528	October 1920-	714
3531500	Powell River near Jonesville, VA	319	October 1931-	541

1/No mean annual flow published; less than 5 years of streamflow record.

USES, FUNDING, AND AVAILABILITY OF CONTINUOUS STREAMFLOW DATA

The relevance of a stream gage is determined by the uses that are made of data produced from the gage. The uses of the data from each gage in the Virginia program were identified by a survey of approximately 430 recipients of the annual water-data reports. The survey documented the relative importance of each gage and identified gaging stations that may be considered for discontinuation.

Data uses identified by the survey were categorized into nine classes, defined below, and were tabulated in table 2 by gaging station. The users and uses within most of the categories cover a broad spectrum of the community involved in water-resources activities. The individual users and uses are described in the footnotes to table 2. The sources of funding for each gage and the frequency at which data are provided to the users are included in table 2.

Data-Use Classes

The following definitions were used in categorizing the uses of streamflow for each continuous stream gage.

Regional Hydrology

For data to be useful in defining regional hydrology, a stream gage must be largely unaffected by manmade storage or diversion. In this case of uses, man's effects on streamflow are not necessarily small, but the effects considered are limited to those of land-use and climate changes. Large amounts of manmade storage may exist in the basin provided the outflow is uncontrolled. These stations are useful in developing regionally transferable information about the relationship between basin characteristics and streamflow.

Sixty-one stations in the Virginia network are classified in the regional hydrology data-use category. One of these stations is a special case in that it is a designated bench-mark station. Such stations serve as indicators of hydrologic conditions in watersheds relatively free of manmade alterations.

Table 2.--Data-use table

STATION NUMBER	DATA USE									FUNDING				FREQUENCY OF DATA AVAILA-BILITY
	REGIONAL HYDROLOGY	HYDROLOGIC SYSTEMS	LEGAL OBLIGATIONS	PLANNING AND DESIGN	PROJECT OPERATION	HYDROLOGIC FORECASTS	WATER-QUALITY MONITORING	RESEARCH	OTHER	FEDERAL PROGRAM	OFA PROGRAM	CO-OP PROGRAM	OTHER NON-FEDERAL	
1620500	*	1 2					1		3 4			5		A P
1631000	*	1 2 6 7			8 9	10	1 7 11 12 13 14 15 16		3 17		18			A T
1632000	*	1 2 19			9	10	1 11		3			5		A T
1634000	*	1 2 20 7			8 9	10	1 12 20 7 16		3 17			5		A T
1658500	*	21 22					21 23 24	25	24		26			A P
1662000	*	27 1 28					29		29		27			A
1663500	*	27 1 28							3		27			A
1664000	*	27 1 7 28 30 31				10	12 7 23 29 31		3 17 29 32 33		27			A P T
1665500	*	27 1 2 28							3 4 34		27			A
1666500	*	27 1 2 28					23		3 34		27			A
1667500	*	27 1 2 28 30				10	23		3 17		27			A T
1668000	*	27 1 2 7				10	12 7 29 35 36 37		3 29 35 33 38		27			A P
1672500	*	27 2				10	39 35		3 40 41	*				A P
1673000		27				10	5 12 16 35 37		35 41 42		27			A P
1674500	*	27					12 16 35 37		35		27			A P
1677000	*	27					5 35		35 43			44		A
2011400	*	1 11 27			45		46 47		3 47		27			A P T
2011800		27 46			45 48		16 45 46 47 49		3 50 48 49		27			A P T
2013000	*	1 27 46			45 48		46 48		3 48		27			A P T
2013100		5 27			45 48		12 46 48 51		48		27			A P T
2016000	*	1 2					46 51		3		27			A P
2019500	*	27 46				10	5 47 51		3 17 33 0		27			A P T
2021500		27					11		3		27			A P
2024000	*	2 11 27				10	11 12		3		27			A P T
2026000	*	7 27 46				10 52	7 46		3 17 33		27			A P T

See footnotes at end of table.

Table 2.--Data-use table--Continued

STATION NUMBER	DATA USE: REGIONAL HYDROLOGY	HYDROLOGIC SYSTEMS	LEGAL OBLIGATIONS	PLANNING AND DESIGN	PROJECT OPERATION	HYDROLOGIC FORECASTS	WATER-QUALITY MONITORING	RESEARCH	OTHER	FUNDING: FEDERAL PROGRAM	OFA PROGRAM	CO-OP PROGRAM	OTHER NON-FEDERAL	FREQUENCY OF DATA AVAILABILITY
2029000	*	27 31				10 52	11 31		3 17		27			A P T
2030500	*	1 27							3 53		27			A P
2034000	*	27				10 52			3 17 54		27			A P T
2035000	*	7 27				10 52 55	5 12 7 32 37 52 55		3 17 33	*	27			A P T
2038000	*	2 27					35 55 56		56			5		A P
2038850	57	27 57					12 35 57		58	*				A P
2039500	*	1 2 27				10	12 29		29 40 53 58 59		27			A P T
2042500		2 27					12 16 35 56 60		60			5	61	A P
2043500	*	27					56	62	56 63			5		A
2045500	*	2 27				10	12 56 64		3		27			A P
2047000	*	2 27 56			65	10	12 37 56 66 67		56 67		27	68		A P T
2049500	*	5 27					12 16 37 56 66 67		56 67		27	68		A P
2051500	*	27				10	56		40 56 59 69	*				A
2055000	*	46			70 71 72	10 73	12 46 49		3 17 38 49 69 74 75		76		77	A P T
2055100	*	46			78		46 51		74 79			80		A
2056000		5 46			70 71		46 49		38 49 81				77	A P
2056650	*	46		82			46		69			80		A
2058400	*	46			70 72		46		17 38 53 59 69		76		77	A P T
2060500		46 76			70 72	10 73	16 46 83		3 17 38 53 75 83		76		77	A P T
2062500		46 76			70 72 84	10 73	46 84		3 17 38 53 75 85		76		77	A P T
2066000		5 76			72	10 73	5 16 32 86 87		17 38 53 86 87		76			A P T
2072000		5 46 76			72	10 73	46		3 38 53 75 88 89 90		76			A P
2072500		5 46 76			72	10	46		3 17 38 53 88 91		76			A P T

See footnotes at end of table.

Table 2.--Data-use table--Continued

STATION NUMBER	DATA USE: REGIONAL HYDROLOGY	HYDROLOGIC SYSTEMS	LEGAL OBLIGATIONS	PLANNING AND DESIGN	PROJECT OPERATION	HYDROLOGIC FORECASTS	WATER-QUALITY MONITORING	RESEARCH	OTHER	FUNDING: FEDERAL PROGRAM	OFA PROGRAM	CO-OP PROGRAM	OTHER NON-FEDERAL	FREQUENCY OF DATA AVAILA-BILITY
2077000	*	46			93	10 73	46		53 59 69 75 88		76			A P
2077500		5				10 73			53 59		76			A P
3164000	*	1			70 94	10	16 95		3 17		96			A P T
3165000	*	1			94		95		3 69		96			A P
3166800	*						95					5		A
3167000	*	1					95		69	*				A
3167500	*	1 46			94		46		3 69		96			A P
3168000	*	46			70 94	10	46 97		17 50 97		96		77	A P T
3170000	*	1 46					46		3				98	A
3171000		46			70	10	46 97 99		3 17 38 50 69 97 99				77	A P T
3173000	*	1 46			94		46		69 100		96			A P
3175500	*	46			94		46				96			A P
3176500		46			70 94	10	12 37 46 97 99		3 17 38 97 99		96			A P T
3177710	*	46					5 12		38 69			5		A P
3207800	*				101		12 16 95		17 100		96			A P T
3208500	*				101	10	16 95 102		3 17 100 102		96			A P T
3208700		5			101		95 102		17 100 102		96			A P T
3208950	*				101		16 95 102		17 102		96			A P T
3209000		5			101		12 95		17 100		96			A P T
3471500	*	1			103	10	95			*				A P
3473000	*				103	10	12 95		3 17			5		A P T
3475000	*	5			103	10	95		3 69			5		A P
3478400	*					10	95		104			5		A P
3488000	*	1			103	10	12 95		3			5		A P
3521500	*	1			103		16 95		3 69 100		105			A P
3524000	*	1			103	10	16 95 102		3 69 100 102		105			A P T
3531500	*	1			103	10	95 106		3 69 100 106			5		A P T

1. Interstate Commission on the Potomac River Basin.
2. Virginia State Water Control Board--safe yield analysis.
3. Streamflow data requests for recreational planning.
4. University of Virginia--Shenandoah watershed acidification study.
5. Virginia State Water Control Board.
6. Virginia State Water Control Board, Valley Regional Office--ground-water quantification.
7. Virginia Commission of Game and Inland Fisheries.
8. U.S. Army Corps of Engineers--flood control and Washington, D.C., water-supply system.
9. Interstate Commission on the Potomac River Basin--low-flow coordination and planning.
10. U.S. National Weather Service--flood forecasting.
11. Virginia State Water Control Board, Valley Regional Office.
12. U.S. Environmental Protection Agency.
13. Virginia Polytechnic Institute, Biology Department.
14. Lawler, Matusky & Skelly (consulting engineers)--sediment composition analysis.
15. E. I. duPont Corporation--mercury pollution study.

Table 2.--Data-use table--Continued

16. Virginia Department of Highways and Transportation.
17. Virginia State Water Control Board--real time flooding analysis for flood-insurance program.
18. U.S. Army Corps of Engineers, Baltimore District.
19. Town of Broadway, Virginia.
20. Virginia State Water Control Board, Valley Regional Office--waste-load allocation.
21. Metropolitan Washington Council of Governments.
22. Prince William County.
23. Virginia State Water Control Board, Northern Regional Office, Surveillance Division.
24. Northern Virginia Planning District Commission--planning projects.
25. Investigation of sediment transfer and deposition.
26. U.S. National Park Service.
27. U.S. Army Corps of Engineers, Norfolk District.
28. Rappahannock-Rapidan Planning District Commission.
29. Greely and Hansen, Engineers--project planning and design.
30. Virginia State Water Control Board, Northern Regional Office--index station for regional flow-monitoring report.
31. Virginia Institute of Marine Science.
32. U.S. Geological Survey, Water Resources Division, Central Region Research--study of regional sediment loads.
33. U.S. Fish and Wildlife Service--planning and design of fish-passage facilities.
34. U.S. National Weather Service--rainfall-runoff research.
35. Virginia Institute of Marine Science--investigation of water quality and salt intrusion in estuaries.
36. Virginia State Water Control Board Northern Region Office--base station for Fredericksburg area.
37. NASQAN station.
38. U.S. Department of Energy, Southeastern Power Administration--analysis and planning, water availability for powerplants.
39. Virginia Highway and Transportation Research Council--short-term monitoring.
40. Wiley and Wilson, Engineers--project planning and design.
41. Hanover County--long-range water-supply withdrawal planning.
42. Virginia State Water Control Board--analysis for water-supply potential.
43. Gloucester County--project planning and design, and future flow analysis.
44. James City County.
45. U.S. Army Corps of Engineers, Norfolk District--operation of Gathright Dam.
46. Virginia State Water Control Board, West Central Regional Office.
47. Virginia Game Commission, Fish Division--fisheries management.
48. Westvaco Corporation (paper manufacturer)--monitoring for flood warning and to manage effluent discharge for compliance.
49. Fifth Planning District Commission--planning projects.
50. Virginia Game Commission, Fish Division--predict flows for fisheries management.
51. Fifth Planning District Commission.
52. Albemarle County, Office of County Engineer.
53. U.S. Soil Conservation Service, Chase City--predicting variability in streamflow.
54. Albemarle County, Office of County Engineer--project planning and design.
55. Chesterfield County, Utilities Department.
56. Southeastern Public Service Authority of Virginia--project planning and design.
57. Hydrologic benchmark station.
58. U.S. Soil Conservation Service, Farmville--real-time flooding and low-flow analyses for local warning system.
59. Virginia Commission of Game and Inland Fisheries--assess potential environmental impacts from dams, hydropower plants, etc.
60. Newport News, Virginia, Department of Public Works--evaluate inflow to Newport News water supply.
61. Newport News, Virginia.
62. Regional low-flow hydrology assessment.
63. U.S. Fish and Wildlife Service--investigations of hydrology of Great Dismal Swamp.
64. Georgia-Pacific Corporation.
65. Norfolk, Virginia, Utilities Department--water-supply diversion management.
66. Union Camp Corporation--management of effluent discharges for compliance.
67. Virginia State Water Control Board, Tidewater Regional Office--nutrient export loading study.
68. Southeastern Public Service Authority of Virginia.
69. Virginia State Water Control Board--water-supply forecasting.
70. Appalachian Power Company--powerplant operation management.
71. Roanoke, Virginia--manage operation of regional wastewater treatment plant.
72. U.S. Army Corps of Engineers, Wilmington District--manage operation of Roanoke River reservoir system.
73. Halifax County.
74. U.S. Army Corps of Engineers, Wilmington District--Roanoke-Salem study.
75. Halifax County--monitor reservoir operation.
76. U.S. Army Corps of Engineers, Wilmington District.
77. Appalachian Power Company.

Table 2.--Data-use table--Continued

78. Roanoke, Virginia--water-supply diversion management.
79. Virginia State Water Control Board--water-supply forecasting and monitoring flow-by for compliance.
80. Roanoke, Virginia.
81. Virginia Game Commission, Fish Division--monitor reservoir releases for downstream fishery management.
82. Roanoke, Virginia--water-supply source planning study.
83. Virginia Commission of Game and Inland Fisheries--assess potential environmental impacts from dams, hydropower plants, etc.; also striped bass research and management, and flood warnings to protect fish hatchery.
84. Virginia Game Commission, Fish Division--manage flow for striped bass spawning run and operation of fish hatchery.
85. Virginia Commission of Game and Inland Fisheries--assess potential environmental impacts from dams, hydropower plants, etc.; also striped bass research and management.
86. Virginia Polytechnic Institute, Civil Engineering Department.
87. Virginia Commission of Game and Inland Fisheries--striped bass management and research.
88. South Boston, Virginia--future flow analysis.
89. Virginia Commission of Game and Inland Fisheries--assess potential environmental impacts from dams, hydropower plants, etc.; trout research and management of trout fishery.
90. Virginia State Water Control Board--monitor releases from Lake Philpott.
91. Virginia Commission of Game and Inland Fisheries--assess potential environmental impacts from dams, hydropower plants, etc.; also trout management and research.
92. Virginia Commission of Game and Inland Fisheries--assess potential environmental impacts from dams, hydropower plants, etc.; also research and management of striped bass population in Kern reservoir.
93. South Boston, Virginia--manage water-supply plant operation.
94. U.S. Army Corps of Engineers, Huntington District--manage operation of Kanawha River reservoir system.
95. Virginia State Water Control Board--Southwestern Regional Office.
96. U.S. Army Corps of Engineers, Huntington District.
97. Virginia Polytechnic Institute, Biology Department--New River organic transport study.
98. Radford, Virginia.
99. Biological Monitoring, Inc. (consultants)--low-flow related study.
100. Virginia Polytechnic Institute, Civil Engineering Department--water-supply study for coal slurry pipeline.
101. U.S. Army Corps of Engineers, Huntington District--manage operation of Big Sandy River reservoir system.
102. Pittston Clinchfield Coal Company--coal hydrology study.
103. Tennessee Valley Authority--manage operation of Tennessee River reservoir system.
104. Tennessee Valley Authority, Flood Hazard Analysis Branch--evaluate effectiveness of upstream flood control projects.
105. Tennessee Valley Authority.
106. Westmoreland Coal Company--coal hydrology study.

Hydrologic Systems

Stations that can be used for accounting—that is, to define current hydrologic conditions and the sources, sinks, and fluxes of water through hydrologic systems including regulated systems--are designated hydrologic systems stations. They include diversions and return flows and stations that are useful for defining the interaction of water systems.

Seventy-one of the 77 gaging stations in the Virginia program have one or more uses in this category. The bench-mark station is included in the hydrologic systems category because it accounts for current and long-term conditions of the hydrologic system that it gages.

Legal Obligations

Some stations provide records of flows for the verification or enforcement of treaties, compacts, and decrees. The legal obligation category contains only those stations the USGS must operate to satisfy legal responsibilities of the Federal Government.

No stations in the Virginia program are operated to fulfill legal responsibilities of the USGS.

Planning and Design

Gaging stations in this category of data use are used for the planning and design of a specific project (for example, a dam, levee, floodwall, navigation system, water-supply diversion, hydropower plant, or waste-treatment facility) or group of structures. The planning and design category is limited to those stations that were instituted for such purposes and for which this purpose is still valid.

Currently, only one station in the Virginia program is operated for planning and design purposes.

Project Operation

Gaging stations in this category are used, on an ongoing basis, to assist water managers in making operational decisions such as reservoir releases, hydropower operations, or diversions. The project operation use generally implies that data are routinely available to operators on a rapid-reporting basis. For projects on large streams, data may only be needed every few days.

There are 40 stations in the Virginia program that are used in this manner. Thirty-three of these are used to aid operators in the management of reservoir systems.

Hydrologic Forecasts

Gaging stations in this category are regularly used to provide information for hydrologic forecasting. Such information might include flood forecasts for a specific river reach, or periodic (daily, weekly, monthly, or seasonal) flow-volume forecasts for a specific site or region. The hydrologic forecast use generally implies that the data are routinely available to the forecasters on a rapid-reporting basis. On large streams, data may only be needed every few days.

There are 40 stations in the Virginia program used for hydrologic forecasts. Stations in the program that are included in the hydrologic forecast category are used primarily for flood forecasting. Data are used mainly by the U.S. National Weather Service, but also by some local agencies for flood forecasting.

Water-Quality Monitoring

Gaging stations where regular water-quality or sediment-transport monitoring is conducted, and where the availability of streamflow data contributes to the utility or is essential to the interpretation of the water-quality or sediment data, are designated as water-quality-monitoring sites.

One station in the program is a designated bench-mark station and eight are National Stream Quality Accounting Network (NASQAN) stations. Water-quality samples from bench-mark stations are used to indicate water-quality characteristics of streams that have been and probably will continue to be relatively free of manmade influence. NASQAN stations are part of a nationwide network designed to assess water-quality trends of significant streams. In addition to the bench-mark and NASQAN utilization, a wide variety of other uses within the water-quality monitoring category are made of stations in the Virginia program. Water-quality monitoring activities of one type or another are conducted at 72 of the stations.

Research

Gaging stations in this category are operated for specific research or water-investigations studies. Typically, these are operated for only a few years.

Two stations in the Virginia program are being operated primarily in support of research activities. One station is being operated in conjunction with an investigation of sediment transfer and deposition for the U.S. National Park Service. The other station was established for a regional low-flow hydrology assessment in cooperation with the Virginia State Water Control Board. Data from this site also are being used in an investigation of the hydrology of the Great Dismal Swamp.

Other

In addition to the eight data-use classes described above, the gaging stations in the Virginia program are used to provide streamflow information for a wide variety of other purposes. One widespread use is for recreational planning, primarily for canoeists, rafters, and fishermen.

Funding

The four sources of funding for the streamflow-data program are as follows:

1. Federal program.--Funds that have been directly allocated to USGS.
2. Other Federal Agency (OFA) program.--Funds that have been transferred to the USGS by OFA's.
3. Coop program.--Funds that come jointly from USGS cooperative-designated funding and from a non-Federal cooperating agency. Cooperating agency funds may be in the form of direct services or cash.
4. Other non-Federal.--Funds that are provided entirely by a non-Federal agency or a private concern under the auspices of a Federal agency. Funds in this category are not matched by USGS cooperative funds.

In all four categories, identified sources of funding pertain only to the collection of streamflow data; sources of funding for other activities, particularly collection of water-quality samples, that might be carried out at the site are not necessarily the same as those identified herein.

Ten organizations, other than the USGS, currently contribute funds to the Virginia stream-gaging program.

Frequency of Data Availability

Frequency of data availability refers to the times at which streamflow data may be furnished to the users. In this category, three distinct possibilities exist. Data can be furnished by direct-access telemetry equipment for immediate use, by periodic release of provisional data, or in publication format through annual data reports published by the USGS for Virginia (U.S. Geological Survey, 1983). These three categories are designated T, P, and A, respectively, in table 2. In the current Virginia program, data for all 77 stations are made available through annual reports, data from 37 stations are available on a real-time basis, and data are released on a provisional basis at 61 stations.

Data-Use Presentation

Data-use and ancillary information for each continuous gaging station are presented in table 2. The entry of an asterisk in the table indicates that the station is used for regional hydrology purposes, and(or) the station is operated from Federal funds appropriated directly to the Geological Survey.

Conclusions Pertaining to Data Uses

A review of the data-use and funding information presented in table 2 indicates that, in general, the data from the gaging stations in this program are needed by significant numbers of organizations involved in water-resources related activities (such as consulting engineering firms). Nearly all the stations' data have multiple significant uses.

There are, however, based on known data uses, four stations that could be considered for discontinuation. Data from two of the stations, 3166800 and 3167000, have only minimal known applications, 3166800 in particular with only one acknowledged user and only one use other than for regional hydrology. Two other stations, 1658500 and 2043500, are operated primarily in conjunction with research activities that should be completed in 1985. However, other significant uses may preclude discontinuing these stations after their present primary research usage has been completed.

Based on the discussion above and the information presented in table 2, up to four stations could be discontinued. Regardless, in the interests of completeness and continuity, all 77 gaging stations will be included in the analysis in the following sections of this report.

ALTERNATIVE METHODS OF DEVELOPING STREAMFLOW INFORMATION

The second step of the analysis of the stream-gaging program is to investigate alternative methods of providing daily streamflow information in lieu of operating continuous-flow gaging stations. The objective of the analysis is to identify gaging stations where alternative technology, such as flow-routing or statistical methods, will provide information about daily mean streamflow in a more cost-effective manner than operating a continous stream gage. No guidelines concerning suitable accuracies exist for particular uses of the data; therefore, judgment is required in deciding whether the accuracy of the estimated daily flows is suitable for the intended purpose. The data uses at a station will influence whether a site has potential for alternative methods. For example, stations for which real-time flood hydrographs are required, for uses such as hydrologic forecasts and project operation, are not candidates for the alternative methods. Likewise, legal obligations to operate a gaging station would preclude utilizing alternative methods. The primary candidates for alternative methods are stations operated upstream or downstream of other stations on the same stream. The accuracy of the estimated streamflow at these sites may be suitable because of the high redundancy of flow information between sites. Similar watersheds, in the same physiographic and climatic area, also may have potential for alternative methods.

All stations in the Virginia stream-gaging program were evaluated regarding their potential for utilization of alternative methods. Selected methods were applied at six of the stations. These applications are described later in this section of the report. This section also briefly describes the two alternative methods used in the Virginia analysis and documents why these methods were chosen.

Desirable attributes of a proposed alternative method are (1) the proposed method should be computer oriented and easy to apply, (2) the proposed method should have an available interface with the USGS WATSTORE Daily Values File (Hutchinson, 1975), (3) the proposed method should be technically sound and generally acceptable to the hydrologic community, and (4) the proposed method should permit easy evaluation of the accuracy of the simulated streamflow records. The desirability of the first attribute above is obvious. Second, the interface with the WATSTORE Daily Values File is needed to easily calibrate the proposed alternative method. Third, the alternative method selected for analysis must be technically sound or it will not be able to provide data of suitable accuracy. Fourth, the alternative method should provide an estimate of the accuracy of the streamflow to judge the adequacy of the simulated data. The above selection criteria were used to select two methods--a flow-routing model and multiple-regression analysis.

Description of Flow-Routing Model

Hydrologic flow-routing methods use the law of conservation of mass and the relationship between storage in a reach and outflow from the reach. The hydraulics of the system are not considered. The method usually requires only a few parameters and treats the reach in a lumped sense without subdivision. The input is usually a discharge hydrograph at the upstream end of the reach and the output a discharge hydrograph at the downstream end. Several different types of hydrologic routing, such as Muskingum, Modified Puls, Kinematic Wave, and the unit-response flow-routing method, are available. The latter method was selected for this analysis. This method uses two techniques--storage continuity (Sauer, 1973) and diffusion analogy (Keefer, 1974; Keefer and McQuivey, 1974). These concepts are discussed below.

The unit-response method was selected because it fulfilled the criteria noted above. Computer programs for the unit-response method can be used to route streamflow from one or more upstream locations to a downstream location. Downstream hydrographs are produced by the convolution of upstream hydrographs with their appropriate unit-response functions. This method can be applied only at a downstream station where an upstream station exists on the same stream. An advantage of this model is that it can be used for regulated stream systems. Reservoir routing techniques are included in the model so flows can be routed through reservoirs if the operating rules are known. Calibration and verification of the flow-routing model is achieved using observed upstream and downstream hydrographs and estimates of tributary inflows. The convolution model treats a stream reach as a linear one-dimensional system in which the system output (downstream hydrograph) is computed by multiplying (convoluting) the ordinates of the upstream hydrograph by the unit-response function and lagging them appropriately. The model has the capability of combining hydrographs, multiplying a hydrograph by a ratio, and changing the timing of a hydrograph. In this analysis, the model is used only to route an upstream hydrograph to a downstream point. Routing can be accomplished using hourly data, but only daily data are used in this analysis.

Three options are available for determining the unit (system) response function. Selection of the appropriate option depends primarily upon the variability of wave celerity (traveltime) and dispersion (channel storage) throughout the range of discharges to be routed. Adequate routing of daily flows usually can be accomplished using a single unit-response function (linearization about a single discharge) to represent the system response. However, if the routing coefficients vary drastically with discharge, linearization about a low-range discharge results in overestimated high flows that arrive late at the downstream site; whereas linearization about a high-range discharge results in low-range flows that are underestimated and arrive too soon. A single unit-response function may not provide acceptable results in such cases. Therefore, the option of multiple linearization (Keefer and McQuivey, 1974), which uses a family of unit-response functions to represent the system response, is available.

Determination of the system's response to the input at the upstream end of the reach is not the total solution for most flow-routing problems. The convolution process makes no accounting of flow from the intervening area between the upstream and downstream locations. Such flows may be unknown or estimated by some combination of gaged and ungaged flows. An estimating technique that should prove satisfactory in many instances is the multiplication of known flows at an index gaging station by a factor (for example, a drainage-area ratio).

The objective in both the storage-continuity and the diffusion analogy flow-routing methods is to calibrate two parameters that describe the storage-discharge relationship in a given reach and the traveltime of flow passing through the reach. In the storage-continuity method, a response function is derived by modifying a translation hydrograph technique developed by Mitchell (1962) to apply to open channels. A triangular pulse (Sauer, 1973) is routed through reservoir-type storage and then transformed by a summation curve technique to a unit response of desired duration. The two parameters that describe the routing reach are K_s, a storage coefficient that is the slope of the storage-discharge relation, and W_s, the translation hydrograph time base. These two parameters determine the shape of the resulting unit-response function.

In the diffusion-analogy method, the two parameters requiring calibration are K_o, a wave dispersion or damping coefficient, and C_o, the floodwave celerity. K_o controls the spreading of the wave (analogous to K_s in the storage-continuity method) and C_o controls the traveltime (analogous to W_s in the storage-continuity method). In the single-linearization method, only one K_o and one C_o value are used. In the multiple-linearization method, C_o and K_o are varied with discharge so a table of wave celerity (C_o) versus discharge (Q) and a table of dispersion coefficient (K_o) versus discharge (Q) are used.

In both the storage-continuity and diffusion-analogy methods, the two parameters are calibrated by trial and error. The analyst must decide whether suitable parameters have been derived by comparing the simulated discharge with the observed discharge. The application of flow-routing techniques to four watersheds in Virginia is described in a subsequent section of this report.

Description of Regression Analysis

Simple- and multiple-regression techniques can also be used to estimate daily flow records. Regression equations can be computed that relate daily flows (or their logarithms) at a single station to daily flows at a combination of upstream, downstream, and(or) tributary stations. This statistical method is not limited, like the flow-routing method, to stations where an upstream station exists on the same stream. The explanatory variables in the regression analysis can be stations from different watersheds, or downstream and tributary watersheds. The regression method has many of the same attributes as the flow-routing method in that it is easy to apply, provides indices of accuracy, and is generally accepted as a good tool for estimation. The theory and assumptions of regression analysis are described in several textbooks such as those by Draper and Smith (1966) and Kleinbaum and Kupper (1978). The application of regression analysis to hydrologic problems is described and illustrated by Riggs (1973) and Thomas and Benson (1970). Only a brief description of regression analysis is provided in this report.

A linear regression model of the following form was used for estimating daily mean discharges in Virginia:

$$y_i = B_o + \sum_{j=1}^{\rho} B_j x_j + e_i$$

where

y_i is the logarithm of daily mean discharge at station i (dependent variable),

x_j is the logarithms of daily mean discharges at nearby stations (explanatory variables),

B_o and B_j are the regression constant and coefficients, and

e_i is the random error term.

The above equation is calibrated (B_o and B_j are estimated) using observed values of y_i and x_j. These observed daily mean discharges can be retrieved from the WATSTORE Daily Values File. The values of x_j may be discharges observed on the same day as discharges at station i or may be for previous or future days, depending on whether station j is upstream or downstream of station i. Once the equation is calibrated and verified, future values of y_i are estimated using observed values of x_j. The regression constant and coefficients (B_o and B_j) are tested to determine whether they are significantly different from zero. A given station j should be retained in the regression equation only if its regression coefficient (B_j) is significantly different from zero. The regression equation should be calibrated using one period of time and then verified or tested for a different period of time to obtain a measure of the true predictive accuracy. Both the calibration and verification period should be representative of the range of flows that could occur at station j. The equation should be verified by plotting the residuals e_i (difference between simulated and observed discharges) against the dependent and all explanatory varibles in the equation, and by plotting the simulated and observed discharges versus time. These tests are intended to determine whether the linear model is appropriate or some transformation of the variables is needed and whether there is any bias in the equation such as overestimating low flows. These tests might indicate, for example, that a logarithmic transformation is desirable, that a nonlinear regression equation is appropriate, or that the regression equation is biased in some way. In this report, these tests indicated that a linear model was appropriate, with y_i and x_j, in cubic feet per second, logarithmically transformed. The application of linear-regression techniques to three watersheds in Virginia is described in a subsequent section of this report.

It should be noted that the use of a regression relation to synthesize data at a discontinued gaging station entails a reduction in the variance of the streamflow record relative to that which would be computed from an actual record of streamflow at the site. The reduction in variance expressed as a fraction is approximately equal to one minus the square of the correlation coefficient that results from the regression analysis.

Categorization of Stream Gages by Their Potential for Alternative Methods

Analyses of the areal distribution of the gaging stations in the program and of the data uses presented in table 2 identified the six most promising stations at which to test alternative methods for obtaining the needed streamflow information. The six stations were 2029000, 2051500, 2056000, 2062500, 2075000, and 3173000. Based on the capabilities and limitations of the methods and data availability, flow-routing techniques were applied at four stations and regression methods were used at three sites. One station (2062500) was considered amenable to modeling with both techniques.

Dan River Flow-Routing Analysis

The purpose of this flow-routing analysis was to investigate the potential for use of the unit-response model for streamflow routing to simulate daily-mean discharges at station 2075000, Dan River at Danville. This station was chosen for demonstration purposes, to illustrate the most accuracy that reasonably can be expected from simulation of streamflow using currently available practical techniques. This station (and river system) was chosen primarily because of the minimal intervening ungaged drainage area and the lack of apparent regulation in that intervening area. In reality, this site would not be a suitable target for discontinuing because of the heavy and high-priority usage summarized in table 2.

This flow-routing analysis, of station 2075000, will be referred to hereafter as the Dan River flow-routing study. In this application, a best-fit model for the entire flow range is the desired product. A schematic diagram of the Dan River study area is presented in figure 3. Streamflow data available for this analysis are summarized in table 3.

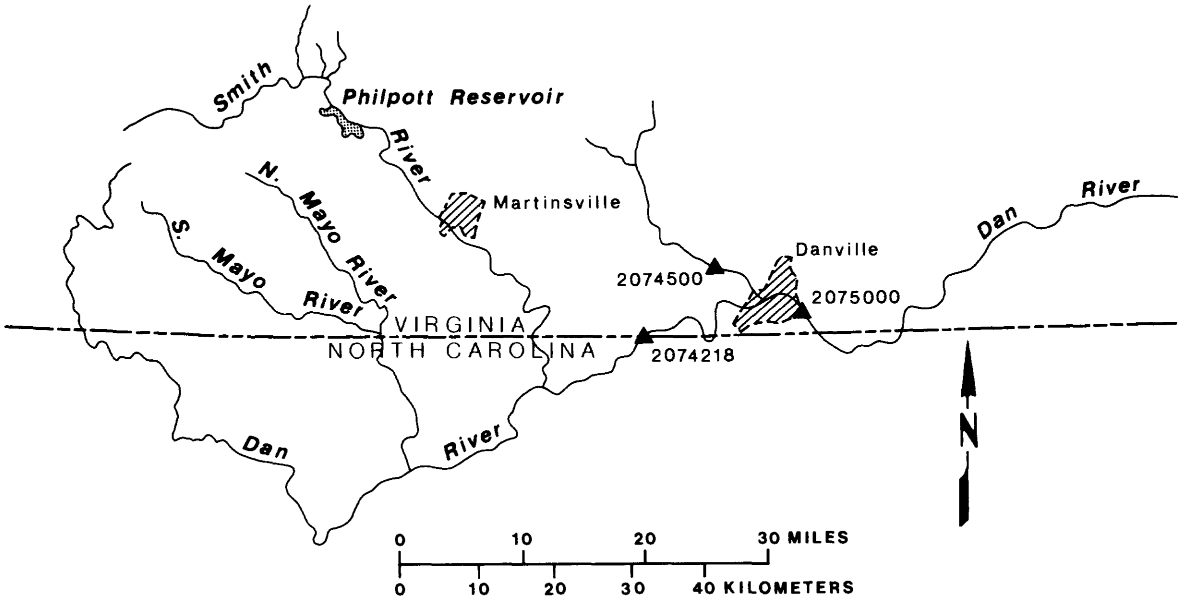


Figure 3.-- The Dan River flow-routing study area.

The Dan River gage being modeled (2075000) is located 18.3 mi downstream from the next upstream gage (2074218), Dan River near Mayfield, N.C. In this reach there is some diurnal fluctuation caused by cotton mills, which does not appear to have any significant effect on the daily discharge at 2075000. The intervening drainage area between stations 2074218 and 2075000 is 272 mi^2, or 13 percent of the total drainage area contributing to the downstream station. Station 2074500 is the only gage within the 272-mi^2 intervening area. The gaged area of station 2074500 reduces the ungaged intervening area between the two primary stations to 160 mi^2, or only 7.8 percent of the total downstream station drainage area.

To simulate the daily-mean discharges, flows were routed from station 2074218 to station 2075000 using the diffusion analogy method with a single linearization. The intervening drainage area was accounted for by using data from station 2074500 adjusted by drainage area ratio. The total discharge at 2075000 was the summation of the routed discharge from 2074218 and adjusted discharge from 2074500. Daily streamflow data for the 1980 water year were used to calibrate the model.

Table 3.--Gaging stations used in the Dan River flow-routing study

Station No.	Station name	Drainage area (mi^2)	Period of record
2074218	Dan River near Mayfield, N.C.	1,778	October 1976 - present
2074500	Sandy River near Danville	112	October 1929 - present
2075000	Dan River at Danville	2,050	August 1934 - present

To route flow from station 2074218 to station 2075000, it was necessary to determine the model parameters C_o (floodwave celerity) and K_o (wave-dispersion coefficient). The coefficients C_o and K_o are functions of channel width (W_o) in feet, channel slope (S_o) in feet per foot (ft/ft), the slope of the stage discharge relation (dQ_o/dY_o) in square feet per second (ft^2/s), and discharge (Q_o) in cubic feet per second (ft^3/s) representative of the reach in question and are determined as follows:

$$C_o = \frac{1}{W_o} \cdot \frac{dQ_o}{dY_o} \qquad (1)$$

$$K_o = \frac{Q_o}{2S_oW_o} \qquad (2)$$

The discharge, Q_o, for which initial values of C_o and K_o were linearized, was the mean daily discharge for the period of record for stations 2074218 and 2075000 (U.S. Geological Survey, 1982). The channel width, W_o, was the width at each station corresponding to the Q_o discharge and determined from field discharge measurement data. Channel slope, S_o, was determined by converting the corresponding gage heights of the initial discharges, Q_o, taken from the stage-discharge relationships at each gage to a common datum. The difference between these values was then divided by channel length to obtain a slope. The slope of the stage discharge relations, dQ_o/dY_o, was determined from the rating curves at each gage by using a 0.1-ft increment that bracketed the mean discharge, Q_o. The difference in the discharge divided by the 0.1-ft increment therefore represents the slope of the function at that point. The model parameters as determined above are listed in table 4.

Table 4.--Selected reach characteristics used in the Dan River flow-routing study

Station No.	Q_o (ft^3/s)	W_o (ft)	S_o (ft/ft)	$\frac{dQ_o}{dY_o}$ (ft^2/s)	C_o (ft/s)	K_o (ft^2/s)
2074218	2,077	184	8.377×10^{-4}	700	3.81	6,740
2075000	2,314	365		1,600	4.38	3,780

For the first routing trial, average values for the model parameters, $C_o = 4.10$ and $K_o = 5{,}260$ were used. To simulate the contribution of the intervening drainage area of 272 mi^2, the drainage basin above station 2074500 was assumed to be representative of the total intervening drainage area. Therefore, a drainage area ratio of 272 mi^2 divided by 112 mi^2 (2.43) was applied to the flow at station 2074500 and added in at the downstream end to simulate the input from the intervening drainage area. Unfortunately, there were no other stations in the vicinity to test as alternatives.

Using the 1980 water year as a calibration data set, several trials were made, adjusting the values of both C_o and K_o over a wide range (C_o over one order and K_o over two orders of magnitude). No adjustments were attempted with the intervening drainage area ratio factor because the volume error of the initial trial was nearly negligible at 0.47 percent. The best-fit single linearization model was determined to be that with the initially set values of $C_o = 4.10$ and $K_o = 5{,}260$ as no significant improvement could be made by varying these parameters. A summary of the simulation of mean-daily discharge at station 2075000 on the Dan River for water year 1980 is given in table 5.

Figure 4 is a comparison of the observed and simulated discharge for the Dan River station 2075000 during October, when base flow is generally low, and April when base flow is generally high. The fit for October is reasonably good except for October 1 and 2. On October 1, for which the flow was not plotted, the model was being initialized. On October 2, a storm over the drainage basin

Table 5.--Results of Dan River flow-routing model

Mean absolute error for 366 days	=	5.54 percent
Mean negative error (188 days)	=	-4.20 percent
Mean positive error (178 days)	=	6.95 percent
Total volume error	=	0.47 percent

62 percent of the total observations had errors	≤ 5	percent
86 percent of the total observations had errors	≤ 10	percent
92 percent of the total observations had errors	≤ 15	percent
96 percent of the total observations had errors	≤ 20	percent
97 percent of the total observations had errors	≤ 25	percent
3 percent of the total observations had errors	> 25	percent

upstream of station 2074500 was probably localized and consequently was over-represented by the model at station 2075000 (flow overestimated by 32 percent). The fit for April was better with the maximum difference occurring on April 11 when the flow was overestimated by 18 percent.

Flow-Routing Analysis Summary

The Dan River flow-routing analysis resulted in a fairly effective model for simulating daily streamflow at station 2075000 (see table 5). The other three flow-routing models (which also were calibrated over the 1980 water year, except 2062500; 1981) met with more limited success as summarized in table 6. Much of the difficulty stems from known and other apparent regulation which cannot be effectively quantified. Table 6 provides flow-routing modeling data for the four stations where this technique was applied. Location, drainage area, and other information about the stations involved are available from table 1 and figure 2. None of the four models is considered accurate enough to justify discontinuing any of the gaging stations. In addition, the uses, summarized in table 2, in some cases would have precluded discontinuing their operation.

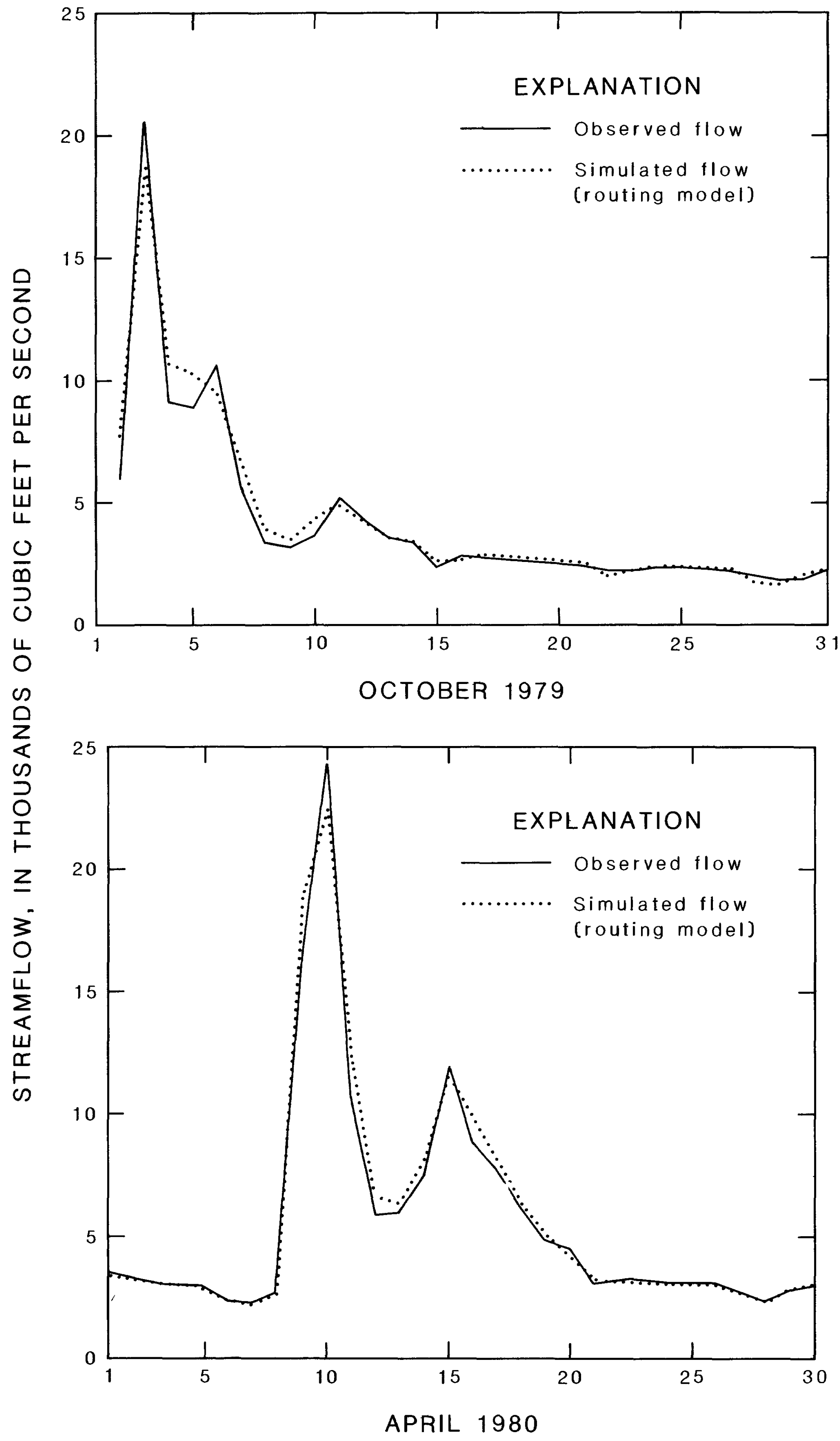


Figure 4.-- Daily hydrograph, Dan River, October 1979 and April 1980.

Table 6.--Summary of daily streamflow flow-routing models

Station No.	Model	C_o (ft/s)	K_o (ft^2/s)	Percentage of simulated flow within stated percent of actual		
				5 %	10 %	20 %
2029000	$Q_{2029000} = (Q_{2026000} + Q_{2027000} + Q_{2027800})_{Routed} + Q_{2028500} + 1.25(Q_{2027000} + Q_{2027800} + Q_{2028500})$	8.80	6,760	57	78	94
2056000	$Q_{2056000} = (Q_{2055000})_{Routed} + 2.06(Q_{2056650})$	4.15	866	40	57	69
2062500	$Q_{2062500} = (Q_{2060500} + Q_{2061500})_{Routed} + 1.00(Q_{2064000})$	3.54	5,140	40	62	80
2075000	$Q_{2075000} = (Q_{2074218})_{Routed} + Q_{2074500} + 1.43(Q_{2074500})$	4.10	5,260	62	86	96

Regression Analysis Results

Linear regression techniques were applied to three of the six selected modeling sites: stations 2051500, 2062500, and 3173000. The streamflow record for each station considered for simulation (the dependent variable) was regressed against streamflow records at other stations (explanatory variables) during a given period of record (the calibration period), which was water years 1979-81 for the gaging stations (systems) analyzed. Best-fit linear regression models were developed and used to provide a daily streamflow record that was compared to the observed streamflow record. The percent difference between the simulated and actual record for each day was calculated.

Station 2062500, Roanoke River at Brookneal, provides a fairly representative example of the regression modeling technique. This site also provides the only example where both the flow-routing and the regression analyses were made. A schematic diagram of the study area related to station 2062500 is presented in figure 5, which shows the stations included in the final model. All the streamflow data considered (some of which did not prove helpful) for this analysis are summarized in table 7. A daily mean-flow hydrograph is presented in figure 6

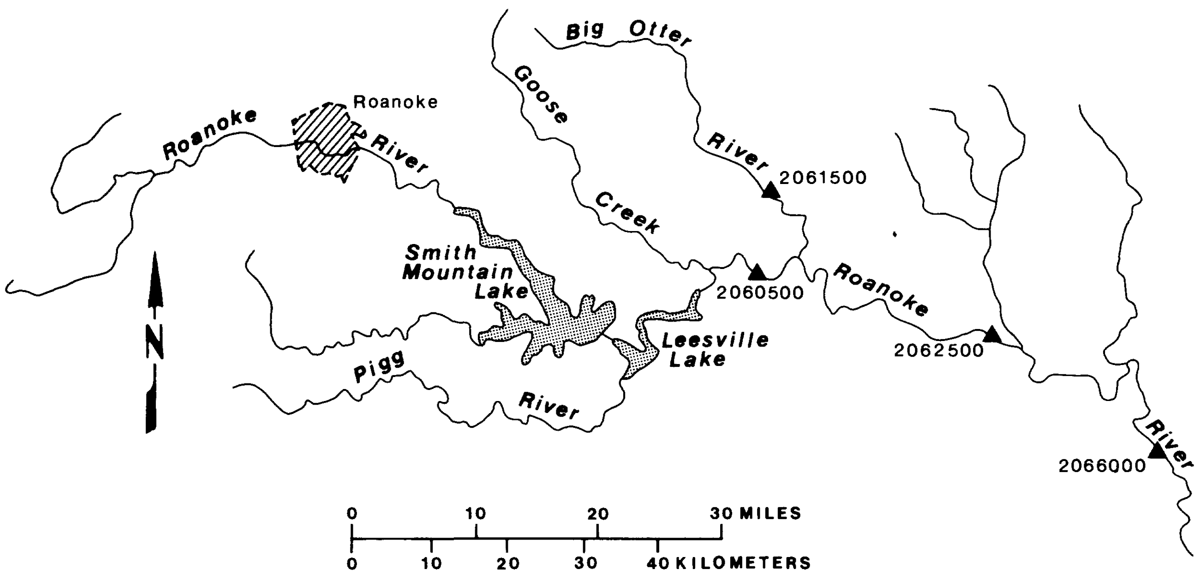


Figure 5.-- The Roanoke River, station 2062500, study area.

Table 7.--Gaging stations used in the Roanoke River, station 2062500, regression model study

Station No.	Station name	Drainage area (mi^2)	Period of record
2060500	Roanoke River at Altavista	1,789	August 1930-present
2061500	Big Otter River near Evington	320	October 1936-present
2062500	Roanoke River at Brookneal	2,415	April 1923-present
2064000	Falling River near Naruna	173	July 1929-January 1935, September 1941-present
2065500	Cub Creek at Phenix	980	August 1946-present
2066000	Roanoke River at Randolph	2,977	August 1900-September 1906, October 1927-September 1930, October 1950-present
2077000	Banister River at Halifax	547	September 1904-December 1905, October 1928-present

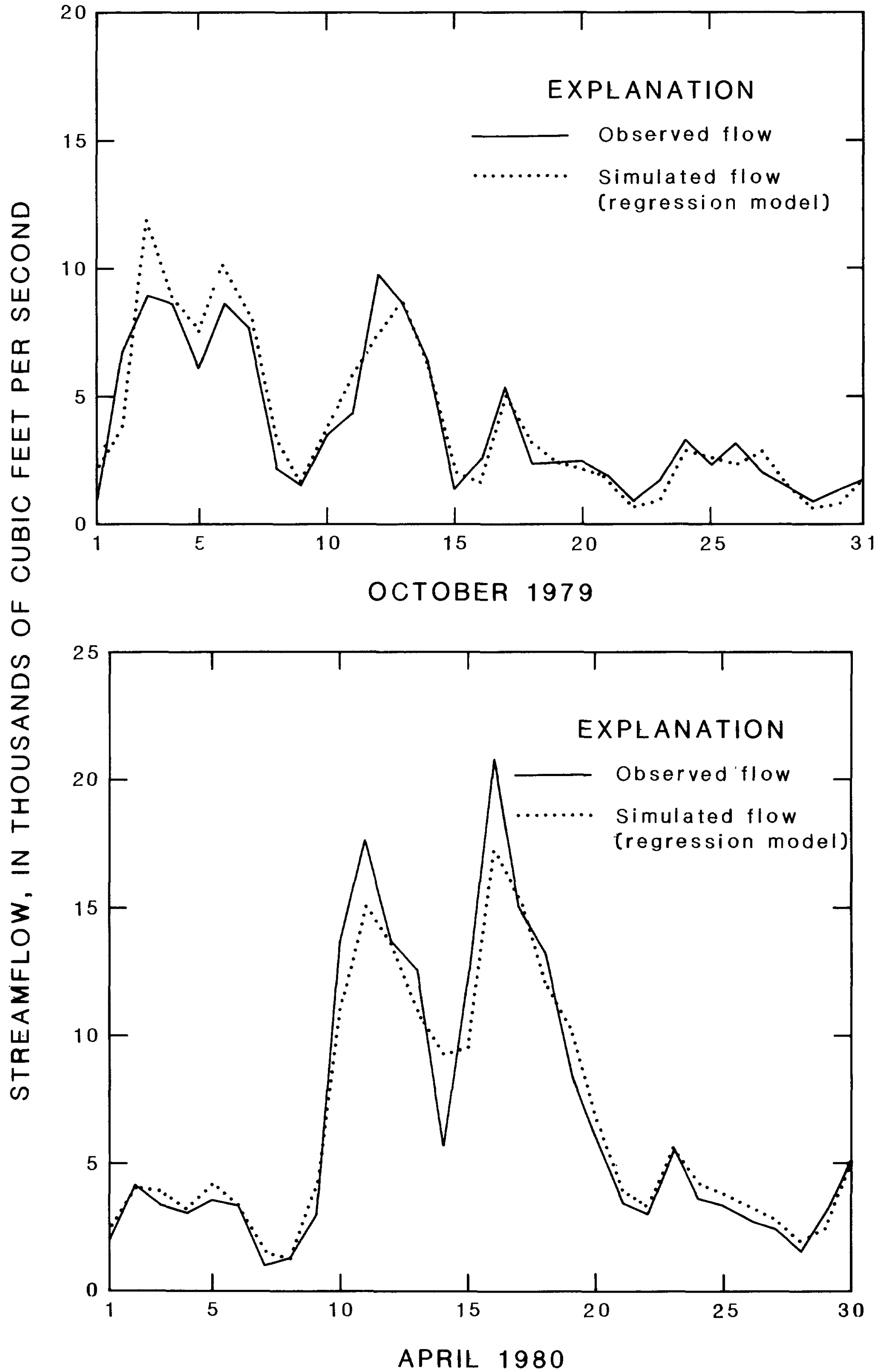


Figure 6.-- Daily hydrograph, Roanoke River, station 2062500, October 1979 and April 1980.

for October and April of water year 1980 for comparison of the observed discharge at the station with the discharge simulated by the regression model developed for this streamflow system. The model, $Q_{2062500} = 0.953(Q_{2060500})^{0.0358}$ $(LAG1\ Q_{2060500})^{0.460}$ $(Q_{2061500})^{0.151}$ $(Q_{2066000})^{0.155}$ $(LAG\text{-}1\ Q_{2066000})^{0.254}$, is included in table 8 which presents a summary of the regression models for the three sites which were analyzed thereby.

The somewhat more successful simulations of streamflow records at stations 2051500 and 2062500 were both produced from regressions with explanatory variables (stations) mostly located in the same basin as the dependent variable (station). The streamflows at these stations experience varying degrees of regulation. The dependent streamflow records were regressed against upstream and downstream records on the main stems of the rivers as well as tributaries to the main stems. Special explanatory variables specified as LAG1 Q and LAG-1 Q were created by lagging the discharges by plus 1 day or minus 1 day. The interaction in a regression of the lagged and unlagged values for a given streamflow record acts to statistically route the flow from an upstream to a downstream site. The lagged discharge values account for the traveltime between the two sites.

Conclusions Pertaining to Alternative Methods of Data Generation

The simulated data from both the flow-routing and regression methods for the six modeled stream systems were not considered sufficiently accurate to substitute for the operation of any continuous stream gages. In general, models based only on streamflow in adjacent drainage basins (which are necessarily regression models) are not very effective in simulating streamflow to reasonable levels of accuracy. The most successful simulation of streamflow was obtained from the flow-routing model of the Dan River (station 2075000) system, and even that model was able to produce flow records within 5 percent of the actual values only 62 percent of the time.

In summary, all six stations considered in this section should remain in operation as part of the Virginia stream-gaging program and will be included in the next step of this analysis.

Table 8.--Summary of daily streamflow regression models

Station No.	Model	Percentage of simulated flow within stated percent of actual		
		5%	10%	20%
2051500	$Q_{2051500} = 1.59(Q_{2044500})^{0.111}(\text{LAG1 } Q_{2051000})^{0.155}(Q_{2052000})^{0.437}(\text{LAG-1 } Q_{2052000})^{0.255}$	28	53	85
2062500	$Q_{2062500} = 0.953(Q_{2060500})^{0.0358}(\text{LAG1 } Q_{2060500})^{0.460}(Q_{2066000})^{0.155}(\text{LAG-1 } Q_{2066000})^{0.254}(Q_{2061500})^{0.151}$	36	58	82
3173000	$Q_{3173000} = 0.546(Q_{3167000})^{0.636}(Q_{3175500})^{0.487}$	28	49	74

COST-EFFECTIVE RESOURCE ALLOCATION

Introduction to Kalman-Filtering for Cost-Effective Resource Allocation (K-CERA)

In a study of the cost-effectiveness of a network of stream gages operated to determine water consumption in the Lower Colorado River Basin, a set of techniques called K-CERA was developed (Moss and Gilroy, 1980). Because that study concerned water balance, the network's effectiveness was measured in terms of the extent to which it minimized the sum of error variances in estimating annual mean discharges at each site in the network. This measure of effectiveness tends to concentrate stream-gaging resources on the larger streams where potential errors in absolute volume of flow are greatest. While such a tendency is appropriate for a water-balance network, in the broader context of the multitude of uses of the streamflow data collected in the U.S. Geological Survey's Streamflow Information program, this tendency causes undue concentration on large streams. Therefore, the original version of K-CERA was extended to include, as optional measures of effectiveness, the sums of the variances of errors in estimating the following streamflow variables; annual mean discharge in cubic feet per second, annual mean discharge in percentage, average instantaneous discharge in cubic feet per second, and average instantaneous discharge in percentage. Using percentage errors does not unduly weight activities at large streams to the detriment of records on small streams. In addition, the instantaneous discharge is the basic variable from which all other streamflow data are derived. For these reasons, this study used the K-CERA techniques with the sums of the variances of the percentage errors of the instantaneous discharges at all continuously gaged sites to measure the effectiveness of the data-collection activity.

The original version of K-CERA also failed to account for error contributed by missing stage or other correlative data that are used to compute streamflow data. The probabilities of missing correlative data increase as the period between service visits to a stream gage increases. A procedure for dealing with the missing record has been developed and was incorporated into this study.

Brief descriptions of the mathematical program used to optimize cost effectiveness of collecting data and of the techniques of applying Kalman filtering (Gelb, 1974) to determine stream-gage record accuracy are presented below.

For more detail on the theory or the applications of K-CERA, see Moss and Gilroy (1980), Gilroy and Moss (1981), and Fontaine and others (1984).

Description of Mathematical Program

The program, called "The Traveling Hydrographer," attempts to allocate among stream gages a predefined budget for the collection of streamflow data in such a manner that the field operation is the most cost-effective possible. The measure of effectiveness is discussed above. The set of decisions available to the manager is the frequency of use (number of times per year) of each of a number of routes that may be used to service the stream gages and to make discharge measurements. The range of options within the program is from zero usage to daily usage for each route. However, for this analysis, an upper limit was set at 72 trips per year. This constraint was considered a reasonable maximum and is discussed later under K-CERA results.

A route is defined as a set of one or more stream gages and the least-cost travel that takes the hydrographer from his base of operations to each of the gages and back to base. A route will have associated with it an average cost of travel and average cost of servicing each stream gage visited along the way. The first step in this part of the analysis is to define the set of practical routes. This set of routes frequently will contain the path to an individual stream gage with that gage as the sole stop and return to the home base so that the individual needs of stream gages can be considered in isolation from the other gages.

Another step in this part of the analysis is the determination of any special requirements for visits to each of the gages for such purposes as necessary periodic maintenance, rejuvenation of recording equipment, or required periodic sampling of water-quality data. Such special requirements are considered to be inviolable constraints in terms of the minimum number of visits to each gage.

The final step is to use all of the above to determine the number of times, N_i, that the i^{th} route for i=1, 2, ..., NR, where NR is the number of practical routes, is used during a year such that the budget for the network is not exceeded, the minimum number of visits to each station is made, and the total uncertainty in the network is minimized. Figure 7 represents this step in the form of a mathematical program. Figure 8 presents a tabular layout of the problem. Each of

$$\underset{\underline{N}}{\text{Minimize}}\ V = \sum_{j=1}^{MG} \phi_j\ (M_j)$$

$V \equiv$ total uncertainty in the network

$\underline{N} \equiv$ vector of annual number times each route was used

$MG \equiv$ number of gages in the network

$M_j \equiv$ annual number of visits to station j

$\phi_j \equiv$ function relating number of visits to uncertainty at station j

Such that

Budget $\geq T_c \equiv$ total cost of operating the network

$$T_c = F_c + \sum_{j=1}^{MG} \alpha_j M_j + \sum_{i=1}^{NR} \beta_i N_i$$

$F_c \equiv$ fixed cost

$\alpha_j \equiv$ unit cost of visit to station j

$NR \equiv$ number of practical routes chosen

$\beta_i \equiv$ travel cost for route i

$N_i \equiv$ annual number times route i is used (an element of $\underline{N}$)

and such that

$$M_j \geq \lambda_j$$

$\lambda_j \equiv$ minimum number of annual visits to station j

Figure 7.-- Mathematical-programing form of the optimization of the routing of hydrographers.

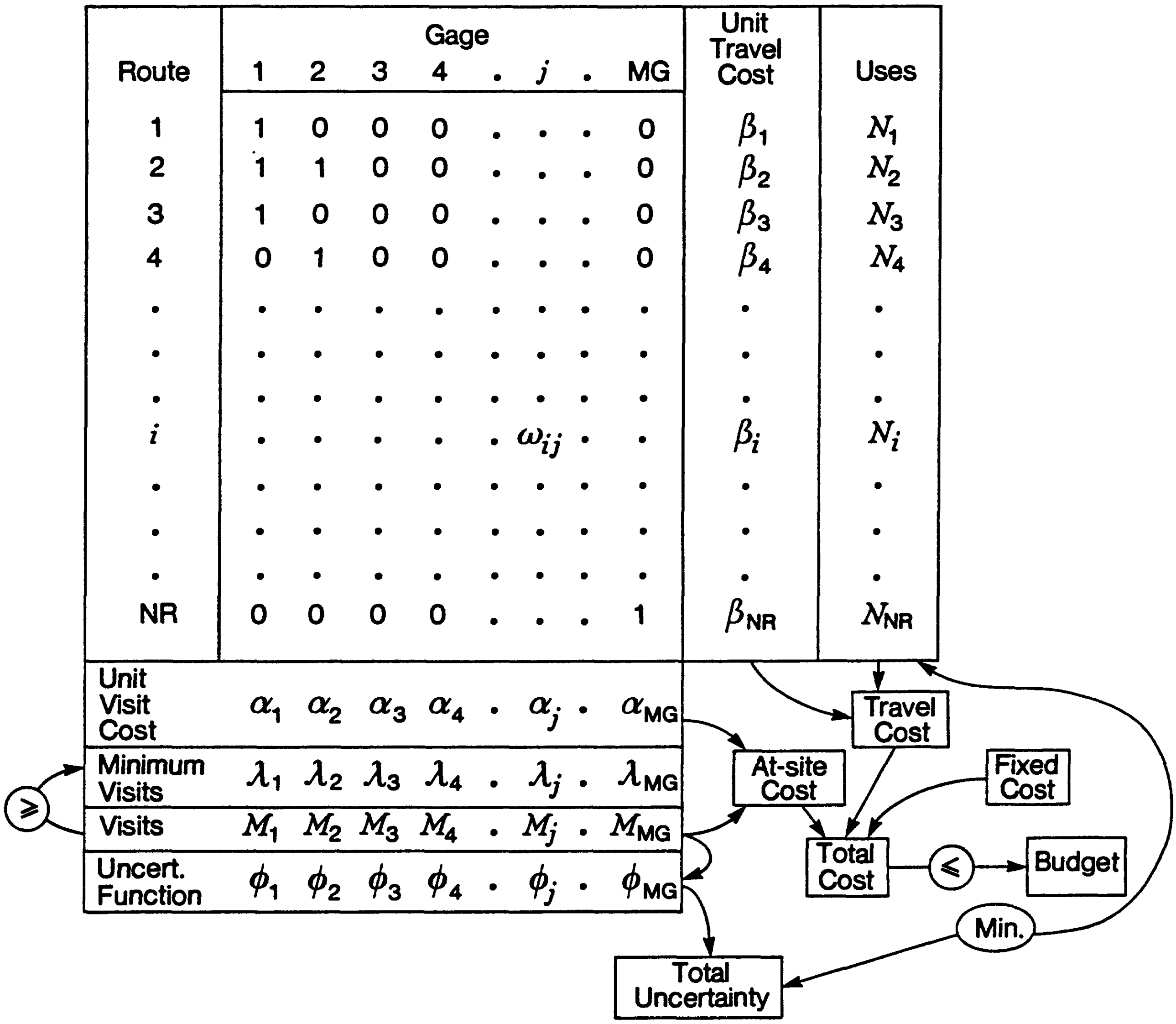


Figure 8.-- Tabular form of the optimization of the routing of hydrographers.

the NR routes is represented by a row of the table and each of the stations is represented by a column. The zero-one matrix, (ω_{ij}), defines the routes in terms of the stations that compose it. A value of one in row i and column j indicates that gaging station j will be visited on route i; a value of zero indicates that it will not. The unit-travel costs, β_i, are the per-trip costs of the hydrographer's travel time, average servicing and maintenance costs at the gaging stations, and any related per diem. The sum of the products of β_i and N_i for i = 1, 2, ..., NR is the total travel and servicing costs associated with the set of decisions $\underline{N} = (N_1, N_2, ..., N_{NR})$.

The unit-visit cost, α_j, is the average cost of making a discharge measurement on a visit to the station. The set of minimum visit constraints is denoted by the row λ_j, j = 1, 2, ..., MG, where MG is the number of stream gages. The row of integers M_j, j = 1, 2, ..., MG specifies the number of visits to each station. M_j is the sum of the products of ω_{ij} and N_i for all i and must equal or exceed λ_j for all j if $\underline{N}$ is to be a feasible solution to the problem.

The total cost expended for discharge measurements at the stations is equal to the sum of the products of α_j and M_j for all j. The cost of record computation, documentation, and publication is assumed to be influenced negligibly by the number of visits to the station and is included along with overhead in the fixed cost of operating the network. The total cost of operating the network equals the sum of the travel costs, the at-site costs, and the fixed cost, and must be less than or equal to the available budget.

The total uncertainty in the estimates of discharges at the MG stations is determined by summing the uncertainty functions, ϕ_j, evaluated at the value of M_j from the row above it, for j = 1, 2, ..., MG.

As pointed out in Moss and Gilroy (1980), the steepest descent search used to solve this mathematical program does not guarantee a true optimum solution. However, the locally optimum set of values for N obtained with this technique specify an efficient strategy for operating the network, which may be the true optimum strategy. The true optimum cannot be guaranteed without testing all undominated, feasible strategies.

Description of Uncertainty Functions

As noted earlier, uncertainty in streamflow records is measured in this study as the average relative variance of estimation of instantaneous discharges. The accuracy of a streamflow estimate depends on how that estimate was obtained. Three situations are considered in this study: (1) streamflow is estimated from measured discharge and correlative data using a stage-discharge relation (rating curve), (2) the streamflow records is reconstructed using secondary data at nearby stations because primary correlative data are missing, and (3) primary and secondary data are unavailable for estimating streamflow. The variances of the errors of the estimates of flow that would be employed in each situation were weighted by the fraction of time each situation is expected to occur. Thus, the average relative variance would be

$$\bar{V} = \varepsilon_f V_f + \varepsilon_r V_r + \varepsilon_e V_e \qquad (3)$$

with

$$1 = \varepsilon_f + \varepsilon_r + \varepsilon_e$$

where

$\bar{V}$ is the average relative variance of the errors of streamflow estimates,

ε_f is the fraction of time that the primary recorders are functioning,

V_f is the relative variance of the errors of flow estimates from primary recorders,

ε_r is the fraction of time that secondary data are available to reconstruct streamflow records given that the primary data are missing,

V_r is the relative variance of the errors of estimation of flows reconstructed from secondary data,

ε_e is the fraction of time that primary and secondary data are not available to compute streamflow records, and

V_e is the relative error variance of the third situation.

The fractions of time that each source of error is relevant are functions of the frequencies at which the recording equipment is serviced.

The time τ since the last service visit until failure of the recorder or recorders at the primary site is assumed to have a negative-exponential probability distribution truncated at the next service time; the distribution's probability density function is

$$f(\tau) = ke^{-k\tau}/(1-e^{-ks}) \quad (4)$$

where

k is the failure rate in units of $(\text{day})^{-1}$,

e is the base of natural logarithms, and

s is the interval between visits to the site in days.

It is assumed that, if a recorder fails, it continues to malfunction until the next service visit. As a result,

$$\varepsilon_f = (1-e^{-ks})/(ks) \quad (5)$$

(Fontaine and others, 1984, eq. 21).

The fraction of time ε_e that no records exist at either the primary or secondary sites can also be derived assuming that the time between failures at both sites are independent and have negative exponential distributions with the same rate constant. It then follows that

$$\varepsilon_e = 1 - [2(1-e^{-ks}) - 0.5(1-e^{-2ks})]/(ks)$$

(Fontaine and others, 1984, eqs. 23 and 25).

Finally, the fraction of time ε_r that records are reconstructed based on data from a secondary site is determined by the equation

$$\varepsilon_r = 1 - \varepsilon_f - \varepsilon_e$$
$$= [(1-e^{-ks}) - 0.5(1-e^{-2ks})]/(ks). \quad (6)$$

The relative variance, V_f, of the error derived from primary record computation is determined by analyzing a time series of residuals that are the differences between the logarithms of measured discharge and the rating curve discharge. The rating curve discharge is determined from a relationship between discharge and some correlative data, such as water-surface elevation at the gaging station. The measured discharge is the discharge determined by field observations of depths, widths, and velocities. Let $q_T(t)$ be the true instantaneous discharge at time t and let $q_R(t)$ be the value that would be estimated using the rating curve. Then

$$x(t) = \ln q_T(t) - \ln q_R(t) = \ln [q_T(t)/q_R(t)] \quad (7)$$

is the instantaneous difference between the logarithms of the true discharge and the rating curve discharge.

In computing estimates of streamflow, the rating curve may be continually adjusted on the basis of periodic measurements of discharge. This adjustment process results in an estimate, $q_c(t)$, that is a better estimate of the stream's discharge at time t. The difference between the variable $\hat{x}(t)$, which is defined

$$\hat{x}(t) = \ln q_c(t) - \ln q_R(t) \qquad (8)$$

and x(t) is the error in the streamflow record at time t. The variance of this difference over time is the desired estimate of V_f.

Unfortunately, the true instantaneous discharge, $q_T(t)$, cannot be determined and, thus, x(t) and the difference, $x(t) - \hat{x}(t)$, cannot be determined as well. However, the statistical properties of $x(t) - \hat{x}(t)$, particularly its variance, can be inferred from the available discharge measurements. Let the observed residuals of measured discharge from the rating curve be z(t) so that

$$z(t) = x(t) + v(t) = \ln q_m(t) - \ln q_R(t) \qquad (9)$$

where

v(t) is the measurement error, and

$\ln q_m(t)$ is the logarithm of the measured discharge equal to $\ln q_T(t)$ plus v(t).

In the Kalman-filter analysis, the z(t) time series was analyzed to determine three site-specific parameters. The Kalman filter used in this study assumes that the time residuals x(t) arise from a continuous first-order Markovian process that has a Gaussian (normal) probability distribution with zero mean and variance (subsequently referred to as process variance) equal to p. A second important parameter is β, the reciprocal of the correlation time of the Markovian process giving rise to x(t); the correlation between $x(t_1)$ and $x(t_2)$ is $\exp[-\beta|t_1-t_2|]$. Fontaine and others (1984) also define q, the constant value of the spectral density function of the white noise which drives the Gauss-Markov x-process. The parameters, p, q, and β are related by

$$\mathrm{Var}[x(t)] = p = q/(2\beta) \qquad (10)$$

The variance of the observed residuals z(t) is

$$\mathrm{Var}[z(t)] = p + r \qquad (11)$$

where r is the variance of the measurement error v(t). The three parameters, p, β, and r, are computed by analyzing the statistical properties of the z(t) time series. These three site-specific parameters are needed to define this component of the uncertainty relationship. The Kalman filter utilizes these three parameters to determine the average relative variance of the errors of estimation of discharges as a function of the number of discharge measurements per year (Moss and Gilroy, 1980).

If the recorder at the primary site fails and there are no concurrent data at other sites that can be used to reconstruct the missing record at the primary site, there are at least two ways of estimating discharges at the primary site. A recession curve could be applied from the time of recorder stoppage until the gage was once again functioning or the expected value of discharge for the period of missing data could be used as an estimate. The expected-value approach is used in this study to estimate V_e, the relative error variance during periods of no concurrent data at nearby stations. If the expected value is used to estimate discharge, the value that is used should be the expected value of discharge at the time of year of the missing record because of the seasonality of the streamflow processes. The variance of streamflow, which also is a seasonally varying parameter, is an estimate of the error variance that results from using the expected value as an estimate. Thus, the coefficient of variation squared $(C_v)^2$ is an estimate of the required relative error variance V_e. Because C_v varies seasonally and the times of failures cannot be anticipated, a seasonally averaged value of C_v is used:

$$\bar{C}_v = \left(\frac{1}{365} \sum_{i=1}^{365} \left(\frac{\sigma_i}{\mu_i} \right)^2 \right)^{1/2} \tag{12}$$

where

σ_i is the standard deviation of daily discharges for the i^{th} day of the year,

μ_i is the expected value of discharge on the i^{th} day of the year, and

$(\bar{C}_v)^2$ is used as an estimate of V_e.

The variance V_r of the relative error during periods of reconstructed streamflow records is estimated on the basis of correlation between records at the primary site and records from other gaged nearby sites. The correlation coefficient

ρ_c between the streamflows with seasonal trends removed at the site of interest and detrended streamflows at the other sites is a measure of the goodness of their linear relationship. The fraction of the variance of streamflow at the primary site that is explained by data from the other sites is equal to ρ_c^2. Thus, the relative error variance of flow estimates at the primary site obtained from secondary information will be

$$V_r = (1-\rho_c^2)\,\overline{C}_v^{\,2} \tag{13}$$

Because errors in streamflow estimates arise from three different sources with widely varying precisions, the resultant distribution of those errors may differ significantly from a normal or log-normal distribution. This lack of normality causes difficulty in interpretation of the resulting average estimation variance. When primary and secondary data are unavailable, the relative error variance V_e may be very large. This could yield correspondingly large values of $\overline{V}$ in equation 3 even if the probability that primary and secondary information are not available, ε_e, is quite small.

A new parameter, the equivalent Gaussian spread (EGS), is introduced here to assist in interpreting the results of the analyses. If it is assumed that the various errors arising from the three situations represented in equation 3 are log-normally distributed, the value of EGS is determined by the probability statement that

$$\text{Probability}\ [e^{-EGS} \leq (q_c(t)\,/\,q_T(t)) \leq e^{+EGS}] = 0.683 \tag{14}$$

Thus, if the residuals $\ln q_c(t) - \ln q_T(t)$ were normally distributed, $(EGS)^2$ would be their variance. Here EGS is reported in units of percent because EGS is defined so that nearly two-thirds of the errors in instantaneous streamflow data will be within plus or minus EGS percent of the reported values.

The Application of K-CERA in Virginia

As discussed in the first two parts of this analysis, data are currently being used from all 77 stream gages being operated in the Virginia Federal program. Also, there is no effective way to generate from other sources of hydrologic data reasonably accurate records at these sites. Therefore, anticipating these gages will continue to be operated, at least for the time being, they were included in the K-CERA analysis. Acceptable uncertainty functions were determined for 75 of the 77 stations and results are described below. At the other two sites, there were too few discharge measurements available to develop stable uncertainty functions. These two stations were not included in the accuracy optimization process described in a subsequent section.

Determination of Missing Record Probabilities

As described earlier, the statistical characteristics of missing stage or other correlative data for computation of streamflow records can be defined by a single parameter, the value of k in the truncated negative exponential probability distribution of times to failure of the equipment. In the representation of $f(\tau)$, as given in equation 4, the average time to failure is $1/k$. The value of $1/k$ will vary from site to site depending upon the type of equipment at the site and upon its exposure to natural elements and vandalism. The value of $1/k$ can be changed by advances in the technology of data collection and recording. To estimate $1/k$ in Virginia, the most recent records available, 1983, were used to represent the most current technology and the most current general pattern in which the stream gages are serviced. During this period, the gages malfunctioned on the average 1.95 percent of the time.

Gaging stations equipped with both digital and analog recorders experienced significantly less record loss than those with only digital recorders. Therefore, such dual stations were grouped and their average record loss of 0.24 percent and average visit frequency were used to determine a $1/k$ of 7,590 days.

The remaining gages, with single digital recorders, were grouped, in general, according to their physiographic provinces because their record losses tended to be patterned according to these regions. Streambed material composition and other physical and water-quality characteristics of the different regions were assumed to be related to loss of record in such ways as channel aggradation covering intakes, sediment filling and plugging intake pipes, and heat and humidity affecting recorder and battery operation. For the stations with single recorders, the average monthly visit frequencies and the average record losses of 4.7 percent (Coastal Plain), 1.9 percent (Piedmont), and 1.4 percent (Valley and Ridge, Blue Ridge, and Appalachian Plateau) were used to determine $1/k$ values of 470 days, 948 days, and 1,430 days, respectively. The 1/k values given above were used to determine ε_f, ε_e, and ε_r for each of the 77 stream gages as a function of the individual frequencies of visit.

Determination of Cross-Correlation Coefficients and Coefficients of Variation

To compute the values of V_e and V_r of the uncertainty functions, daily streamflow records for each of the 77 stations for the last 10 years, or the part thereof for which daily streamflow values are stored in WATSTORE (Hutchinson, 1975), were retrieved. For each of the stream gages that had 3 or more complete water years of data, the value of C_v was computed and various options, based on combinations of other stream gages, were explored to determine the maximum ρ_c. For one station, that had less than 3 water years of data, values of C_v and ρ_c were estimated subjectively.

Linear and multiple regression analyses were performed to determine the cross-correlation relationships between daily discharges at the gaging stations and various nearby stations. The nearby stations were considered to be the independent variables (auxiliary stations) in the relationships and were analyzed both singly and in combination, both coincident and, in some instances, lagged in time. The analyses were performed on the stations' records subsequent to 1974 to help insure the relationships reflected current field conditions.

The coefficient of variation and the highest cross-correlation coefficient for each station, and the auxiliary station(s) whose records gave the highest cross-correlation coefficient, are listed in table 9.

Kalman-Filter Definition of Variance

The determination of the variance V_f for each of the 77 stream gages required the execution of three distinct steps: (1) long-term rating analysis and computation of residuals of measured discharges from the long-term rating, (2) time-series analysis of the residuals to determine the input parameters of the Kalman-filter streamflow records, and (3) computation of the error variance, V_f, as a function of the time-series parameters, the discharge-measurement-error variance, and the frequency of discharge measurement.

Table 9.--Statistics of record reconstruction

Station number	C_v (percent)	ρ_c	Source of reconstructed records	
1620500	134	0.811	1632000	
1631000	93.8	.944	1636500	
1632000	144	.876	1634000	
1634000	103	.897	1631000	
1658500	133	.745	1664000	
1662000	111	.865	1664000	
1663500	111	.934	1662000	
1664000	112	.883	1662000	1663500
1665500	111	.940	1667500	
1666500	103	.964	1665500	1667500
1667500	104	.974	1665500	1666500
1668000	113	.946	1664000	1667500
1672500	112	.909	1673000	
1673000	113	.909	1672500	
1674500	103	.833	1673000	
1677000	81.7	.524	2038000	
2011400	97.0	.808	2013000	
2011800	101	.889	2013000	2013100
2013000	108	.808	2011400	
2013100	98.1	.952	2011800	2013000
2016000	107	.934	2021500	
2019500	101	.931	2016000	
2021500	121	.967	2024000	
2024000	102	.967	2021500	
2026000	91.7	.929	2029000	
2029000	91.7	.978	2026000	2035000
2030500	103	.894	2034000	2039500
2034000	111	.843	2030500	
2035000	92.7	.927	2029000	
2038000	134	.687	2042500	
2038850	91.5	.862	2030500	2039500
2039500	99.5	.803	2038850	
2042500	105	.820	1673000	1674500
2045500	116	.767	2047000	
2047000	118	.767	2045500	
2049500	123	.845	2047000	
2051500	113	.878	2045500	
2055000	100	.970	2056000	
2055100	99.7	.830	2055000	
2056000	89.3	.970	2055000	
2056650	105	.919	2055000	2055100
2058400	83.4	.810	2077000	
2060500	98.0	.901	2062500	
2062500	92.9	.949	2060500	2066000
2066000	89.4	.823	2062500	

Table 9.--Statistics of record reconstruction--Continued

Station number	C_v (percent)	ρ_c	Source of reconstructed records	
2072000	86.2	.963	2072500	
2072500	80.9	.963	2072000	
2075000	83.2	.952	2075500	
2075500	84.3	.952	2075000	
2077000	96.4	.715	2077500	
2077500	135	.888	2077303	
3164000	70.5	.957	3168000	
3165000	80.6	.825	3164000	3167500
3167000	78.9	.801	3166800	3167500
3167500	69.8	.899	3167000	3170000
3168000	71.1	.957	3164000	
3170000	74.2	.917	2058400	3167500
3171000	71.4	.945	3176500	
3173000	96.1	.881	3175500	
3175500	94.5	.881	3173000	
3176500	74.0	.945	3171000	
3177710*	97	.90		
3207800	108	.916	3208500	
3208500	116	.916	3207800	
3208700	113	.625	3208950	
3208950	108	.900	3208500	
3209000	106	.566	3208500	
3471500	79.8	.899	3473000	
3473000	78.5	.899	3471500	
3475000	75.4	.879	3473000	
3478400	57.6	.664	3475000	
3488000	100	.897	3473000	3475000
3521500	98.9	.928	3524000	
3524000	95.9	.928	3521500	
3531500	102	.750	3524000	

*Less than 3 water years of data are available. Estimates of C_v and ρ_c are subjective.

In the Virginia program analysis, definition of long-term rating functions was accomplished by a least squares curve-fitting process using a three-parameter non-linear model. The rating model is of the form

$$LQM = B_1 + B_3 \, [\ln(GHT - B_2)], \qquad (15)$$

in which

- LQM is the logarithmic (base e) value of the measured discharge,
- GHT is the recorded gage height corresponding to the measured discharge,
- B_1 is the logarithm of discharge for a flow depth of 1 foot,
- B_2 is the gage height of zero flow, and
- B_3 is the slope of the rating curve.

For example, the parameters for the rating function at station 1663500 are

$$B_1 = 4.42$$
$$B_2 = 1.69$$
$$B_3 = 1.72$$

and the residuals of the discharge measurements about that rating function are presented in table 10.

The time series of residuals (in logarithmic units) of the discharge measurements is used to compute sample estimates of q and β, two of the three parameters required to compute V_f, by means of an autocovariance analysis. Measurement variance, the third parameter, is determined from an assumed constant-percentage standard error. For the Virginia program, all open-water measurements were assumed to have a measurement error of 2 percent. Ice measurements do not constitute a significant part of the rating effort in Virginia and therefore were not included in the analysis.

The parameters q and β can be expressed as functions of p (process variance) and Rho (1-day autocorrelation coefficient). The process variance (p) is the total variance, minus the measurement variance, of the shifts about the rating curve. As discussed earlier, $p = q/2\beta$ (eq. 10). Rho is the 1-day autocorrelation coefficient of the shifts about the rating curve:

$$Rho = e^{-\beta |t_1 - t_2|}. \qquad (16)$$

Table 10.--Residual data for station 1663500

Observation number	Date	Measured discharge (ft^3/s)	Measured discharge (log base e)	Residuals (log base e)
1	Jan. 23, 1974	779	6.6580	0.0480
2	Mar. 6, 1974	254	5.5373	.0885
3	Apr. 23, 1974	284	5.6490	.0642
4	May 28, 1974	200	5.2983	.1587
5	July 9, 1974	101	4.6151	.2827
6	Aug. 28, 1974	141	4.9488	.1042
7	Oct. 17, 1974	132	4.8828	-.0275
8	Feb. 10, 1975	338	5.8231	-.0051
9	Apr. 1, 1975	580	6.3630	.1086
10	May 22, 1975	209	5.3423	.0309
11	July 23, 1975	258	5.5530	-.0492
12	Sep. 3, 1975	215	5.3706	-.0592
13	Nov. 3, 1975	267	5.5872	.1011
14	Dec. 8, 1975	202	5.3083	.1240
15	Feb. 11, 1976	381	5.9428	.0479
16	Apr. 1, 1976	3770	8.2348	-.0368
17	May 26, 1976	164	5.0999	.0892
18	July 21, 1976	64.1	4.1604	.1020
19	Sep. 9, 1976	21.0	3.0445	.3812
20	Oct. 21, 1976	1290	7.1624	-.0174
21	Dec. 14, 1976	340	5.8289	-.0804
22	Mar. 15, 1977	692	6.5396	-.0752
23	May 17, 1977	86.0	4.4543	-.1608
24	July 5, 1977	11.8	2.4681	-.3767
25	Aug. 19, 1977	41.6	3.7281	.0217
26	Oct. 14, 1977	24.7	3.2068	.0844
27	Nov. 7, 1977	2820	7.9445	-.0874
28	Dec. 12, 1977	189	5.2417	-.3167
29	Mar. 8, 1978	279	5.6312	.0816
30	Apr. 20, 1978	420	6.0402	.0812
31	May 23, 1978	643	6.4661	.0867
32	July 5, 1978	332	5.8051	-.0897
33	Aug. 16, 1978	198	5.2883	.0499
34	Oct. 11, 1978	46.8	3.8459	.1924
35	Nov. 30, 1978	655	6.4846	-.0469
36	Jan. 17, 1979	319	5.7652	-.0855
37	Mar. 12, 1979	822	6.7117	.0406
38	Apr. 23, 1979	435	6.0753	.0954
39	June 12, 1979	460	6.1312	-.0323
40	July 18, 1979	112	4.7185	.0729
41	Oct. 15, 1979	1190	7.0817	.0603
42	Dec. 20, 1979	310	5.7366	-.0142
43	Jan. 28, 1980	532	6.2766	.0580
44	Mar. 10, 1980	353	5.8665	-.0644
45	Apr. 21, 1980	743	6.6107	.0297

Table 10.--Residual data for station 1663500--Continued

Observation number	Date	Measured discharge (ft³/s)	Measured discharge (log base e)	Residuals (log base e)
46	June 30, 1980	126	4.8363	0.0463
47	Aug. 25, 1980	96.7	4.5716	.1003
48	Nov. 25, 1980	427	6.0568	-.1192
49	Feb. 19, 1981	183	5.2095	-.0814
50	May 27, 1981	134	4.8978	-.0634
51	July 22, 1981	27.6	3.3178	-.2811
52	Nov. 19, 1981	67.9	4.2180	-.1678
53	Apr. 13, 1982	260	5.5607	.0653
54	May 26, 1982	157	5.0562	.0825
55	July 13, 1982	137	4.9200	.0357
56	Sep. 9, 1982	53.8	3.9853	-.0943
57	Oct. 20, 1982	55.4	4.0146	-.0858

Table 11 presents a summary of the autocovariance analysis expressed in terms of process variance and 1-day autocorrelation coefficients. Autocorrelation coefficients were not determined at 2 of the 77 stations included in the analysis. At these sites (2043500 and 3166800), there were too few measurements available relative to the residual scatter encountered to enable definition of stable coefficients. Therefore, only the remaining 75 stations were included in table 11 and used in the optimizing process of the Traveling Hydrographer program. The other two stations (2043500 and 3166800) were used (as dummy stations) in the determination of costs, but not in the accuracy/optimization procedure for the overall network operation.

The autocovariance parameters, summarized in table 11, and data from the determination of missing record probabilities, summarized in table 9, are used jointly to define uncertainty functions for each gaging station. The uncertainty functions give the relationship of total error variance to the number of visits and discharge measurements.

Station 1663500 provides a typical example of an uncertainty function, shown in figure 9. These uncertainty functions are based on the assumption that measurements are made during each visit to the stations.

Table 11.--Summary of the autocovariance analysis

Station number	Rho*	Process variance (log base e)[2]
1620500	0.992	0.2490
1631000	.998	.0029
1632000	.879	.0074
1634000	.972	.0048
1658500	.981	.0588
1662000	.982	.0074
1663500	.976	.0156
1664000	.977	.0098
1665500	.994	.2500
1666500	.997	.0880
1667500	.990	.0046
1668000	.621	.0118
1672500	.988	.0050
1673000	.969	.0081
1674500	.975	.0095
1677000	.927	.0425
2011400	.988	.0101
2011800	.987	.0072
2013000	.975	.0031
2013100	.986	.0046
2016000	.972	.0043
2019500	.981	.0045
2021500	.622	.0048
2024000	.983	.0034
2026000	.980	.0084
2029000	.960	.0013
2030500	.853	.0021
2034000	.953	.0078
2035000	.965	.0085
2038000	.964	.0892
2038850	.972	.0069
2039500	.987	.0148
2042500	.678	.0065
2045500	.983	.0098
2047000	.629	.0094
2049500	.912	.0189
2051500	.972	.0035
2055000	.932	.0066
2055100	.969	.0592
2056000	.933	.0052

*One-day autocorrelation coefficient.

Table 11.--Summary of the autocovariance analysis--Continued

Station number	Rho*	Process variance (log base e)^2
2056650	.968	.0113
2058400	.979	.0026
2060500	.584	.0030
2062500	.981	.0086
2066000	.989	.0085
2072000	.969	.0020
2072500	.643	.0040
2075000	.722	.0370
2075500	.971	.0016
2077000	.980	.0023
2077500	.984	.0435
3164000	.981	.0051
3165000	.970	.0098
3167000	.985	.0078
3167500	.963	.0015
3168000	.636	.0010
3170000	.942	.0040
3171000	.977	.0027
3173000	.651	.0037
3175500	.610	.0047
3176500	.996	.0075
3177710	.976	.0171
3207800	.992	.0852
3208500	.695	.0078
3208700	.959	.0201
3208950	.969	.0062
3209000	.966	.0048
3471500	.960	.0017
3473000	.973	.0040
3475000	.704	.0027
3478400	.583	.0229
3488000	.952	.0052
3521500	.963	.0038
3524000	.979	.0046
3531500	.982	.0181

*One-day autocorrelation coefficient.

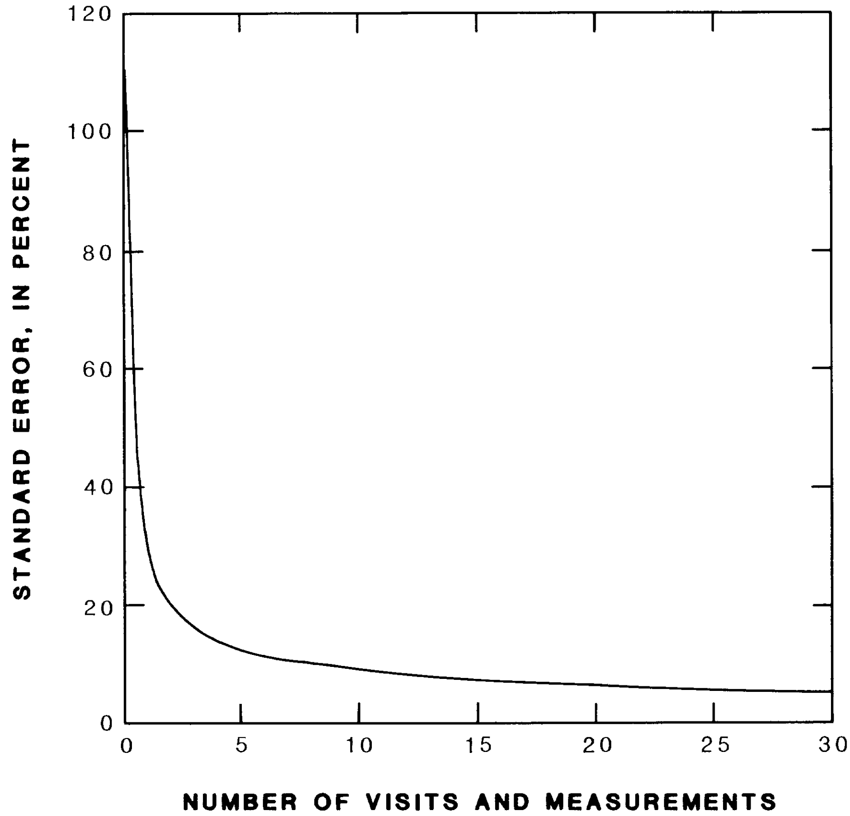


Figure 9.-- Uncertainty function for instantaneous discharge at station 1663500.

Network Operation - Routes and Cost Determination

In Virginia, feasible routes to service the 77 stream gages were determined in consultation with the chief of the Hydrologic Data Section of the Virginia office after review of the uncertainty functions. In summary, 134 routes were selected to service the stream gages in Virginia. These routes included all possible combinations that describe the current operating practice, alternatives that were under consideration as future possibilities, routes that visited certain key individual stations, and combinations that grouped proximate gages where the levels of uncertainty indicated more frequent visits might be useful. These routes and the stations visited on each are summarized in table 12. Visits to sites (such as ground-water wells) other than the primary network stations were included in the route determinations, but are not listed in table 12 to save space and avoid confusion.

Table 12.--Summary of the routes that may be used to visit stations in Virginia

Route number	Stations serviced on the route			
1	1620500			
2	1620500	1632000	1634000	1631000
3	1631000			
4	1631000	1634000		
5	1632000	1620500		
6	1632000			
7	1634000	1631000	1664000	
8	1634000			
9	1658500			
10	1658500	1668000		
11	1662000			
12	1662000	1663500		
13	1663500			
14	1663500	1662000	1664000	
15	1664000			
16	1665500			
17	1665500	1666500	2034000	
18	1665500	1666500		
19	1666500			
20	1667500			
21	1668000			
22	1668000	1627500		
23	1672500			
24	1672500	1673000	1674500	
25	1673000			
26	1673000	1674500		
27	1674500			
28	1674500	2042500		
29	1677000			
30	2011400			
31	2011800			
32	2013000			
33	2013000	2013100	2011800	2011400
	2016000	2019500		
34	2013100			
35	2016000			
36	2019500			
37	2021500			
38	2024000	2021500		
39	2024000			
40	2026000			
41	2026000	2029000		
42	2029000			
43	2030500			
44	2034000			
45	2035000			

Table 12.--Summary of the routes that may be used to visit stations in Virginia--Continued

Route number	Stations serviced on the rqute			
46	2035000	2045500		
47	2035000	2038000		
48	2038000			
49	2038850			
50	2039500			
51	2039500	2035000		
52	2039500	2038000		
53	2039500	2077500	2075000	2075500
	2077000	2051500		
54	2039500	2077500	2075000	2075500
55	2042500			
56	2042500	1677000		
57	2043500			
58	2045500			
59	2047000			
60	2049500			
61	2049500	2047000		
62	2051500			
63	2051500	2077000	2075000	2075500
64	2055000			
65	2055100			
66	2056000			
67	2056650			
68	2058400			
69	2060500			
70	2062500			
71	2062500	2066000	2058400	2072000
	2072500	2060500		
72	2066000			
73	2072000			
74	2072500			
75	2075000			
76	2075500			
77	2077000			
78	2077500			
79	3164000			
80	3164000	3165000		
81	3165000			
82	3166800			
83	3167000			
84	3167000	3166800		
85	3167500			
86	3168000			
87	3168000	3167500	3170000	3171000
	2055000	2056000	2056650	2055100
	3173000	3175500	3176500	3177710

Table 12.--Summary of the routes that may be used to visit stations in Virginia--Continued

Route number	Stations serviced on the route			
88	3170000			
89	3171000			
90	3173000			
91	3175500			
92	3176500			
93	3177710			
94	3207800			
95	3208500			
96	3208700			
97	3208950			
98	3209000			
99	3471500			
100	3471500	3473000	3475000	
101	3473000			
102	3475000			
103	3478400			
104	3488000			
105	3521500			
106	3524000			
107	3524000	3521500	3207800	3208500
	3209000	3208950	3208700	3531500
	3478400			
108	3531500			

The costs associated with the practical routes must be determined. Fixed costs to operate a gage typically include equipment rental, batteries, electricity, data processing and storage, computer charges, maintenance, miscellaneous supplies, and analysis and supervisory charges. For Virginia, average values were applied to each station in the program for all the above categories except analysis and supervisory costs. Costs of analysis and supervision form a large percentage of the cost at each gaging station and can vary widely. These costs were determined on a station-by-station basis from past experience.

Visit costs are those associated with paying the hydrographer for the time actually spent at a station making discharge measurements. These costs vary from station to station as a function of the difficulty encountered in servicing the equipment and the time required to make the discharge measurement. Average visit times were calculated for each station based on an analysis of discharge measurement data available, equipment at the site, and accessibility of the site. This time was then multiplied by the average hourly salary of hydrographers in the Virginia office to determine total visit costs.

Route costs include the cost of the hydrographer's time while in transit and while servicing equipment at the gaging stations. Route costs also include any per diem associated with the time it takes to complete the trip. Because of the accounting system used in the District, vehicle costs are included in overhead.

K-CERA Results

The Traveling Hydrographer program utilizes the uncertainty functions along with the appropriate cost data and route definitions to compute the most cost-effective way of operating the stream-gaging program. In this application, the first step was to simulate the current practice and determine the associated total uncertainty. To accomplish this, the number of visits made to each stream gage and the specific routes used to make these visits were fixed. The resulting average error of estimation for the current practice in Virginia is plotted as a point in figure 10 and is 10.1 percent.

The curve in figure 10 represents the minimum average standard error that can be obtained for a given budget with the existing instrumentation and technology. The line was defined by several runs of the Traveling Hydrographer program

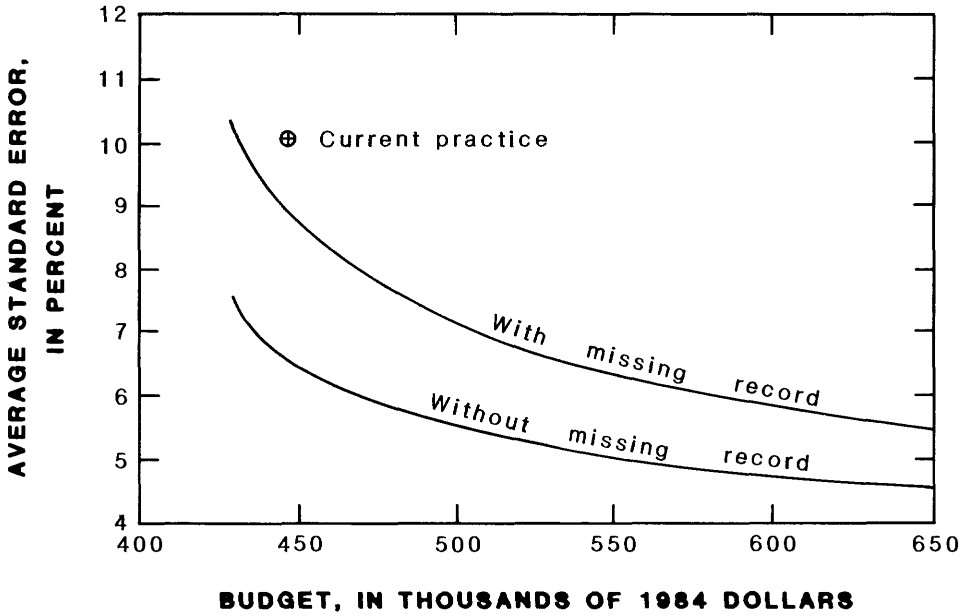


Figure 10.-- Temporal average standard error per stream gage.

with different budgets. Constraints on the operations other than budget were defined as described below.

To determine the minimum number of times each station must be visited, consideration was given primarily to the physical limitations of the method used to record data. The effect of visitation frequency on the accuracy of the data and amount of lost record is taken into account in the uncertainty analysis. In Virginia, a minimum of six visits per year was applied to all gaging stations. This value was based on limitations of the batteries used to drive recording equipment, capabilities of the uptake spools on the digital recorders, problems related to humidity, and the need to protect gages from freezing winter conditions and their intakes from sediment plugging and debris obstructions. At some stations, the six-visit minimum was increased to reflect additional requirements such as water-quality sampling.

A constraint also was placed on the maximum number of visits at each site. A limit of 72 visits per year was put in effect (as being a reasonable maximum) at all stations by flattening the uncertainty function beyond 72 visits.

At eight of the gaging stations which have minimum measurement constraints set above the basement level (6), measurements are not made with every visit. A probability of making measurements was assigned to those stations and was reflected in the visit costs and subsequently in the budget costs of the Traveling Hydrographer program.

The results of the K-CERA analysis are summarized in figure 10 and table 13. It can be seen that the current policy results in an average standard error of estimate of streamflow of 10.1 percent. This policy requires a budget of $446,000 to operate the 77-station stream-gaging program. The range in standard errors is from a low of 3.0 percent for station 3176500 to a high of 24.8 percent at station 2038000. It is possible to obtain the same average standard error with a reduced budget of $430,500 with a change of policy in the field activities of the stream-gaging program. This policy and budget change would result in an increase in standard error from 3.0 to 4.8 percent at station 3176500, while the standard error at station 2038000 would decrease from 24.8 percent to 17.0 percent. However, these two stations would no longer have the greatest extremes of standard error. Station 2013100 would have the lowest standard error at 4.3 percent, while station 2075000 would have the highest at 19.2 percent.

It also would be possible to reduce the average standard error by a policy change while maintaining the $446,000 budget. In this case, the average standard error would decrease from 10.1 to 9.0 percent. Extremes of standard errors for individual sites would be 4.3 and 19.1 percent for stations 2013100 and 2075000, respectively.

A minimum budget of $428,500 is required to operate the 77-station program; a smaller budget would not permit effective service and maintenance of the gages and recorders. Stations would have to be eliminated from the program if the budget fell below this minimum. At the minimum budget, the average standard error is 10.4 percent. The minimum standard error of 4.3 percent would occur at station 2013100, while the maximum of 19.7 percent would occur at station 1677000.

The maximum budget analyzed was $650,000, which resulted in an average standard error of estimate of 5.5 percent. Thus, increasing the budget about 50 percent in conjunction with a policy change would almost halve the average standard error that results from the current policy and current budget. For the $650,000 budget, the extremes of standard error are 2.7 percent for station 2029000 and 13.7 percent at station 2075000. Thus, it is apparent that significant improvements in accuracy of streamflow records can be obtained if larger budgets become available.

Table 13.--Selected results of K-CERA analysis

Station number	Standard error of instantaneous discharge (SE), in percent [Equivalent Gaussian spread (EGS)] (Number of visits per year to site)					
	Current operation	Budget, in thousands of 1984 dollars				
		428.5	430.5	446	500	650
Average per station	10.1* [5.6]	10.4* [6.0]	10.1* [6.0]	9.0* [5.5]	7.2* [4.6]	5.5* [3.4]
1620500	22.2 [19.9] (7)	18.5 [16.5] (10)	16.8 [15.0] (12)	13.3 [11.8] (19)	9.8 [8.6] (35)	6.9 [6.0] (71)
1631000	5.8 [2.9] (6)	4.4 [2.3] (10)	4.4 [2.3] (10)	4.4 [2.3] (10)	4.4 [2.3] (10)	3.4 [1.8] (17)
1632000	12.4 [8.1] (7)	11.8 [8.0] (8)	10.6 [7.7] (11)	10.3 [7.6] (12)	7.6 [6.1] (29)	5.5 [4.5] (62)
1634000	8.5 [5.1] (6)	6.7 [4.2] (10)	6.7 [4.2] (10)	6.7 [4.2] (10)	6.1 [3.9] (12)	4.1 [2.7] (27)
1658500	11.8 [11.2] (12)	11.8 [11.2] (12)	11.8 [11.2] (12)	10.6 [10.0] (15)	8.0 [7.6] (26)	5.8 [5.5] (50)
1662000	11.8 [5.5] (6)	10.9 [5.1] (7)	10.9 [5.1] (7)	8.3 [4.0] (12)	6.3 [3.0] (21)	4.2 [2.1] (46)
1663500	11.7 [8.8] (6)	10.9 [8.3] (7)	10.9 [8.3] (7)	8.4 [6.6] (12)	6.4 [5.0] (21)	4.3 [3.4] (46)
1664000	7.1 [6.5] (8)	6.8 [6.2] (9)	7.1 [6.5] (8)	6.3 [5.8] (11)	4.8 [4.4] (20)	3.2 [2.9] (46)
1665500	16.9 [16.1] (8)	14.3 [13.6] (11)	13.1 [12.4] (13)	10.5 [9.9] (20)	7.6 [7.2] (38)	5.6 [5.3] (71)
1666500	7.8 [6.4] (9)	7.8 [6.4] (9)	9.0 [7.2] (7)	6.7 [5.5] (12)	4.8 [4.0] (23)	3.9 [3.2] (36)
1667500	4.0 [2.2] (14)	4.4 [2.4] (12)	4.4 [2.4] (12)	4.4 [2.4] (12)	4.4 [2.4] (12)	3.2 [1.8] (21)
1668000	11.6 [10.8] (12)	11.6 [10.8] (12)	11.6 [10.8] (12)	11.6 [10.8] (12)	11.6 [10.8] (12)	10.9 [10.5] (22)
1672500	7.4 [3.0] (10)	9.7 [3.8] (6)	8.9 [3.6] (7)	6.7 [2.7] (12)	5.0 [2.1] (21)	3.3 [1.4] (47)
1673000	10.9 [6.3] (12)	10.9 [6.3] (12)	10.9 [6.3] (12)	8.7 [5.2] (19)	6.5 [4.0] (34)	4.5 [2.8] (69)
1674500	12.2 [6.3] (12)	12.2 [6.3] (12)	12.2 [6.3] (12)	9.7 [5.1] (19)	7.2 [3.8] (35)	5.1 [2.7] (70)

* Square root of seasonally averaged station variance.

Table 13.--Selected results of K-CERA analysis--Continued

Station number	Standard error of instantaneous discharge (SE), in percent [Equivalent Gaussian spread (EGS)] (Number of visits per year to site)					
	Current operation	Budget, in thousands of 1984 dollars				
		428.5	430.5	446	500	650
1677000	17.8 [14.9] (17)	19.7 [16.3] (13)	18.3 [15.2] (16)	14.8 [12.4] (27)	10.6 [8.9] (55)	9.3 [7.9] (71)
2011400	5.0 [4.1] (10)	5.5 [4.6] (8)	5.5 [4.6] (8)	5.5 [4.6] (8)	5.5 [4.6] (8)	4.1 [3.4] (15)
2011800	4.3 [3.9] (17)	4.4 [4.0] (16)	4.4 [4.0] (16)	4.4 [4.0] (16)	4.4 [4.0] (16)	4.4 [4.0] (16)
2013000	7.9 [3.3] (10)	8.8 [3.6] (8)	8.8 [3.6] (8)	8.8 [3.6] (8)	6.7 [2.8] (14)	4.6 [2.0] (30)
2013100	3.9 [2.5] (15)	4.3 [2.8] (12)	4.3 [2.8] (12)	4.3 [2.8] (12)	4.3 [2.8] (12)	3.5 [2.3] (18)
2016000	6.4 [4.2] (9)	6.7 [4.4] (8)	6.7 [4.4] (8)	6.7 [4.4] (8)	6.1 [4.0] (10)	4.0 [2.8] (23)
2019500	4.0 [3.6] (11)	4.6 [4.1] (8)	4.6 [4.1] (8)	4.6 [4.1] (8)	4.6 [4.1] (8)	3.4 [3.1] (16)
2021500	8.0 [6.9] (8)	8.5 [7.0] (6)	8.5 [7.0] (6)	8.5 [7.0] (6)	7.8 [6.9] (9)	6.9 [6.5] (22)
2024000	4.7 [3.2] (8)	5.5 [3.6] (6)	5.5 [3.6] (6)	5.5 [3.6] (6)	4.5 [3.0] (9)	2.8 [2.0] (22)
2026000	7.7 [5.4] (8)	8.9 [6.1] (6)	8.9 [6.1] (6)	8.2 [5.7] (7)	6.0 [4.3] (13)	4.0 [2.9] (29)
2029000	4.6 [2.7] (8)	5.4 [2.9] (6)	5.4 [2.9] (6)	5.4 [2.9] (6)	3.8 [2.4] (11)	2.7 [1.8] (22)
2030500	7.4 [4.2] (12)	7.4 [4.2] (12)	7.4 [4.2] (12)	7.4 [4.2] (12)	7.0 [4.1] (14)	4.9 [3.4] (33)
2034000	13.3 [7.6] (6)	12.4 [7.3] (7)	11.7 [7.0] (8)	9.5 [6.0] (13)	7.0 [4.5] (25)	4.7 [3.1] (56)
2035000	5.8 [4.8] (22)	6.9 [5.7] (15)	6.9 [5.7] (15)	6.9 [5.7] (15)	6.4 [5.3] (18)	4.3 [3.6] (41)
2038000	24.8 [21.0] (9)	18.3 [15.6] (18)	17.0 [14.5] (21)	13.2 [11.2] (35)	9.8 [8.3] (63)	9.3 [7.8] (71)
2038850	9.2 [5.6] (8)	10.6 [6.2] (6)	10.6 [6.2] (6)	9.8 [5.9] (7)	7.3 [4.5] (13)	4.8 [3.0] (31)
2039500	9.3 [5.0] (11)	11.7 [6.3] (7)	10.9 [5.9] (8)	8.9 [4.8] (12)	6.6 [3.5] (22)	4.5 [2.4] (47)

Station number	Standard error of instantaneous discharge (SE), in percent [Equivalent Gaussian spread (EGS)] (Number of visits per year to site) Current operation	Budget, in thousands of 1984 dollars 428.5	430.5	446	500	650
2042500	14.5 [8.2] (10)	13.5 [8.0] (12)	12.4 [7.9] (15)	10.3 [7.4] (27)	8.2 [6.5] (55)	7.5 [6.1] (71)
2045500	10.4 [4.5] (12)	10.4 [4.5] (12)	10.4 [4.5] (12)	9.6 [4.1] (14)	7.1 [3.0] (26)	5.0 [2.2] (53)
2047000	18.5 [10.0] (9)	17.1 [9.8] (11)	16.6 [9.7] (12)	13.9 [9.3] (20)	11.2 [8.6] (41)	9.4 [7.7] (71)
2049500	18.6 [12.4] (9)	17.2 [11.9] (11)	16.6 [11.6] (12)	13.6 [10.0] (20)	9.9 [7.5] (41)	7.6 [5.8] (71)
2051500	10.2 [4.2] (7)	11.0 [4.4] (6)	11.0 [4.4] (6)	8.9 [3.8] (9)	6.7 [3.0] (16)	4.6 [2.1] (33)
2055000	7.4 [6.8] (9)	8.2 [7.3] (6)	8.2 [7.3] (6)	8.2 [7.3] (6)	6.8 [6.3] (12)	4.7 [4.5] (31)
2055100	17.0 [16.0] (9)	18.7 [17.5] (7)	16.3 [15.3] (10)	12.9 [12.1] (17)	9.3 [8.7] (33)	6.6 [6.1] (66)
2056000	6.2 [5.7] (11)	7.3 [6.5] (6)	7.3 [6.5] (6)	7.3 [6.5] (6)	6.8 [6.2] (8)	4.7 [4.4] (24)
2056650	8.6 [7.1] (9)	10.1 [8.2] (6)	10.1 [8.1] (6)	9.0 [7.4] (8)	6.9 [5.7] (15)	4.8 [4.0] (32)
2058400	8.3 [3.1] (8)	9.5 [3.5] (6)	9.5 [3.5] (6)	9.5 [3.5] (6)	7.0 [2.7] (11)	4.8 [1.9] (24)
2060500	8.7 [5.5] (8)	9.7 [5.6] (6)	9.7 [5.6] (6)	9.7 [5.6] (6)	7.9 [5.4] (11)	6.4 [5.2] (24)
2062500	6.4 [4.8] (10)	8.3 [6.0] (6)	8.3 [6.0] (6)	8.3 [6.0] (6)	6.4 [4.8] (10)	4.2 [3.2] (23)
2066000	8.0 [3.7] (10)	10.4 [4.8] (6)	10.4 [4.8] (6)	9.6 [4.4] (7)	7.3 [3.4] (12)	4.9 [2.3] (27)
2072000	5.0 [3.0] (9)	6.2 [3.5] (6)	6.2 [3.5] (6)	6.2 [3.5] (6)	4.7 [2.9] (10)	3.1 [2.0] (23)
2072500	6.9 [6.2] (11)	7.8 [6.4] (6)	7.8 [6.4] (6)	7.8 [6.4] (6)	7.0 [6.2] (10)	6.2 [5.9] (23)
2075000	18.7 [18.6] (11)	19.2 [19.0] (7)	19.2 [19.0] (7)	19.1 [18.9] (8)	18.7 [18.6] (11)	13.7 [13.7] (71)
2075500	4.1 [2.3] (13)	5.7 [2.9] (7)	5.7 [2.9] (7)	5.3 [2.8] (8)	4.5 [2.4] (11)	3.2 [1.8] (21)

Table 13.--Selected results of K-CERA analysis--Continued

Station number	Standard error of instantaneous discharge (SE), in percent [Equivalent Gaussian spread (EGS)] (Number of visits per year to site)					
	Current operation	Budget, in thousands of 1984 dollars				
		428.5	430.5	446	500	650
2077000	11.7 [3.0] (7)	12.6 [3.2] (6)	12.6 [3.2] (6)	9.8 [2.6] (10)	7.1 [1.9] (19)	5.0 [1.4] (38)
2077500	12.4 [9.4] (11)	15.6 [11.7] (7)	14.6 [11.0] (8)	11.9 [9.0] (12)	9.0 [6.8] (21)	6.3 [4.7] (42)
3164000	4.2 [3.5] (11)	5.5 [4.6] (6)	5.5 [4.6] (6)	5.5 [4.6] (6)	4.8 [4.1] (8)	3.3 [2.8] (18)
3165000	8.0 [6.2] (10)	9.9 [7.4] (6)	9.9 [7.4] (6)	8.4 [6.4] (9)	6.3 [4.9] (17)	4.3 [3.4] (37)
3167000	7.1 [4.3] (9)	8.6 [5.2] (6)	8.6 [5.2] (6)	7.1 [4.3] (9)	5.3 [3.3] (16)	3.6 [2.2] (35)
3167500	4.6 [2.8] (9)	5.6 [3.1] (6)	5.6 [3.1] (6)	5.6 [3.1] (6)	4.6 [2.8] (9)	3.3 [2.1] (19)
3168000	4.0 [3.1] (9)	4.5 [3.1] (6)	4.5 [3.1] (6)	4.5 [3.1] (6)	4.5 [3.1] (6)	3.5 [3.0] (16)
3170000	6.2 [5.1] (9)	7.0 [5.6] (6)	7.0 [5.6] (6)	7.0 [5.6] (6)	5.7 [4.8] (11)	3.9 [3.4] (29)
3171000	3.4 [2.5] (15)	5.2 [3.6] (6)	5.2 [3.6] (6)	5.2 [3.6] (6)	4.3 [3.1] (9)	3.0 [2.2] (19)
3173000	8.1 [6.0] (9)	9.1 [6.1] (6)	9.1 [6.1] (6)	9.1 [6.1] (6)	7.6 [5.9] (12)	6.3 [5.5] (29)
3175500	8.6 [6.8] (9)	9.5 [6.9] (6)	9.5 [6.9] (6)	9.5 [6.9] (6)	8.3 [6.7] (11)	6.8 [6.3] (32)
3176500	3.0 [1.8] (15)	4.8 [2.7] (6)	4.8 [2.7] (6)	4.8 [2.7] (6)	4.1 [2.4] (8)	2.9 [1.7] (16)
3177710	8.4 [7.1] (11)	9.7 [8.2] (8)	9.7 [8.2] (8)	8.8 [7.4] (10)	6.1 [5.2] (21)	4.3 [3.6] (43)
3207800	10.3 [10.1] (9)	12.6 [12.4] (6)	12.6 [12.4] (6)	9.7 [9.6] (10)	7.0 [6.9] (19)	5.1 [4.9] (37)
3208500	10.2 [8.7] (9)	11.1 [8.8] (6)	11.1 [8.8] (6)	11.1 [8.8] (6)	9.6 [8.5] (13)	7.4 [7.1] (52)
3208700	11.1 [10.2] (9)	12.6 [11.4] (6)	12.6 [11.4] (6)	10.7 [9.8] (10)	7.4 [6.8] (24)	5.0 [4.6] (53)
3208950	5.7 [5.1] (9)	6.6 [5.9] (6)	6.6 [5.9] (6)	6.6 [5.9] (6)	5.2 [4.7] (11)	3.7 [3.4] (24)

Table 13.--Selected results of K-CERA analysis--Continued

Station number	Standard error of instantaneous discharge (SE), in percent [Equivalent Gaussian spread (EGS)] (Number of visits per year to site)					
	Current operation	Budget, in thousands of 1984 dollars				
		428.5	430.5	446	500	650
3209000	6.5 [4.7] (9)	7.6 [5.3] (6)	7.6 [5.3] (6)	7.6 [5.3] (6)	5.9 [4.3] (11)	3.9 [2.9] (27)
3471500	5.0 [2.9] (10)	6.3 [3.4] (6)	6.3 [3.4] (6)	5.2 [3.0] (9)	4.1 [2.5] (15)	2.8 [1.8] (34)
3473000	5.4 [4.0] (11)	6.2 [4.5] (8)	6.2 [4.5] (8)	6.2 [4.5] (8)	4.9 [3.6] (14)	3.4 [2.6] (30)
3475000	6.4 [5.0] (10)	7.4 [5.2] (6)	7.4 [5.2] (6)	7.0 [5.1] (7)	5.9 [4.9] (15)	4.8 [4.4] (39)
3478400	15.6 [15.0] (9)	16.1 [15.3] (6)	16.1 [15.3] (6)	15.8 [15.1] (8)	14.3 [14.1] (29)	12.5 [12.4] (71)
3488000	6.5 [4.9] (13)	6.8 [5.0] (12)	6.8 [5.0] (12)	6.8 [5.0] (12)	5.5 [4.2] (19)	3.8 [3.0] (41)
3521500	6.2 [4.3] (9)	7.4 [5.0] (6)	7.4 [5.0] (6)	7.0 [4.7] (7)	5.3 [3.8] (13)	3.7 [2.7] (28)
3524000	5.3 [3.5] (11)	7.1 [4.6] (6)	7.1 [4.6] (6)	6.6 [4.3] (7)	4.8 [3.3] (13)	3.4 [2.3] (27)
3531500	10.6 [7.1] (9)	12.8 [8.4] (6)	12.8 [8.4] (6)	10.1 [6.7] (10)	7.6 [5.0] (18)	5.1 [3.4] (39)

In order to estimate the amount of uncertainty in the stream-gaging records as a result of less than perfect instrumentation, the analysis also was performed under the assumption that no correlative data at a stream gage were lost. The curve, labeled "Without missing record" on figure 10, shows the average standard errors of estimation of streamflow that could be obtained if perfectly reliable systems were available to measure and record the correlative data.

For the minimal operating budget of $428,500, the effect of having perfect equipment would be the greatest, reducing the average standard error from 10.4 to 7.6 percent. At the other budgetary extreme of $650,000, under which stations are visited more frequently and less record should be lost, the standard error would be reduced from 5.5 percent, with the current system for sensing and recording hydrologic data, to 4.5 percent with ideal equipment. For the current operation and budget ($446,000), the use of completely reliable equipment would reduce the standard error from 10.1 to 7.6 percent. Thus, it is apparent that improved equipment can have a very positive impact on uncertainties in streamflow data throughout the range of operating budgets that might be anticipated for the stream-gaging program in Virginia.

Conclusions from the K-CERA Analysis

As a result of the K-CERA analysis, the following conclusions were drawn:

1. The policy for conducting field operations in the stream-gaging program could be altered to maintain the current average standard error of estimate of streamflow records of 10.1 percent with a budget of $430,500. This shift would result in some increases and some decreases in accuracy of records at individual sites.
2. The funding for stations with unacceptable accuracies for the data uses should be renegotiated with the data users.
3. The K-CERA analysis should be repeated with new stations included whenever sufficient information about the characteristics of new stations has been obtained.
4. Schemes for reducing the probabilities of missing record, such as increased use of dual recorders, local gage observers, and satellite relay of data, should be studied with respect to their cost-effectiveness in providing streamflow information.

SUMMARY

Currently, 77 continuous stream gages are operated in Virginia by the U.S. Geological Survey at a cost of $446,000. Eleven separate sources of funding contribute to this program and 18 separate uses were identified for data from a single gage.

In an analysis of the uses made of the data, two stations were identified as producing data with somewhat limited uses, possibly insufficient to warrant continuing their operation. Two other stations were identified as having uses primarily related to short-term studies. All four of these stations should be considered for discontinuing in the near future. Data from the remaining 73 stations in the program probably will continue to have multiple significant uses for the foreseeable future.

The current (1985) policy for operation of the 77-station program requires a budget of $446,000 per year. It was shown that the overall level of accuracy of the records at these 77 sites could be maintained with a budget of $430,500, if the allocation of gaging resources among the gages were altered.

The study indicates that a major component of error in the streamflow records results from lost or missing data. If perfect equipment were available, the standard error for the current program and budget could be reduced from 10.1 to 7.6 percent. This also can be interpreted to mean that the current streamflow data have a standard error of 7.6 percent during times when the equipment is operating properly. Upgrading equipment and developing strategies to minimize lost record appear to be key actions available to improve the reliability and accuracy of the streamflow data generated in the State.

Studies of the cost-effectiveness of the stream-gaging program should be continued in relation to increases in data available from gaging-station operations and streamflow measurements, and especially in relation to ongoing changes in stage-recording and discharge-measuring equipment. Future studies also should be made to reflect subsequent addition and deletion of stream gages in relation to changing demand for streamflow information. Such changes will affect the operation of other stations in the program both because of the interdependence (among stations) of the information that is generated (data redundancy) and because of the interdependence of the costs of collecting the data from which the information is derived.

SELECTED REFERENCES

Benson, M. A., and Carter, R. W., 1973, A national study of the streamflow data-collection program: U.S. Geological Survey Water-Supply Paper 2028, 44 p.

Carter, R. W., and Benson, M. A., 1970, Concepts for the design of streamflow data programs: U.S. Geological Survey Open-File Report, 33 p.

Draper, N. R., and Smith, H., 1966, Applied regression analysis (2d ed.): New York, John Wiley and Sons, 709 p.

Fontaine, R. A., Moss, M. E., Smath, J. A., and Thomas, W. O., Jr., 1984, Cost effectiveness of the stream-gaging program in Maine - A prototype for nation-wide implementation: U.S. Geological Survey Water-Supply Paper 2244, 39 p.

Gelb, A., ed., 1974, Applied optimal estimation: The Massachusetts Institute of Technology Press, Cambridge, Mass., 374 p.

Gilroy, E. J., and Moss, M. E., 1981, Cost-effective stream-gaging strategies for the Lower Colorado River Basin: U.S. Geological Survey Open-File Report 81-1019, 38 p.

Hutchinson, N. E., 1975, WATSTORE User's guide, volume 1: U.S. Geological Survey Open-File Report 75-426.

Keefer, T. N., 1974, Desktop computer flow routing: American Society of Civil Engineers Proceedings, Journal of the Hydraulics Division, v. 100, no. HY7, p. 1047-1058.

Keefer, T. N., and McQuivey, R. S., 1974, Multiple linearization flow routing model: American Society of Civil Engineers Proceedings, Journal of the Hydraulics Division, v. 100, no. HY7, p. 1031-1046.

Kleinbaum, D. G., and Kupper, L. L., 1978, Applied regression analysis and other multivariable methods: North Scituate, Mass., Duxbury Press, 556 p.

Mitchell, W. D., 1962, Effect of reservoir storage on peak flow: U.S. Geological Survey Water-Supply Paper 1580, p. C1-C25.

Moss, M. E., and Gilroy, E. J., 1980, Cost-effective stream-gaging strategies for the Lower Colorado River Basin; the Blythe field office operations: U.S. Geological Survey Open-File Report 80-1048, 111 p.

Moss, M. E., Gilroy, E. J., Tasker, G. D., and Karlinger, M. R., 1982, Design of surface-water data networks for regional information: U.S. Geological Survey Water-Supply Paper 2178, 33 p.

Riggs, H. C., 1973, Regional analysis of streamflow characteristics: U.S. Geological Survey Techniques of Water-Resources Investigations, Book 4, Chapter B3, 15 p.

Sauer, V. B., 1973, Unit response method of open-channel flow routing: American Society of Civil Engineers Proceedings: Journal of the Hydraulics Division, v. 99, no. HY1, p. 179-193.

Thomas, D. M., and Benson, M. A., 1970, Generalization of streamflow characteristics from drainage-basin characteristics: U.S. Geological Survey Water-Supply Paper 1975, 55 p.

U.S. Geological Survey, 1983, Water resources data - Virginia, water year 1983: U.S. Geological Survey Water Data Report VA-83-1, 442 p.