COST-EFFECTIVENESS OF THE STREAM-GAGING PROGRAM IN MISSOURI

By Loyd A. Waite

U.S. GEOLOGICAL SURVEY

Water Resources Investigations Report 87-4254

Rolla, Missouri
1987

DEPARTMENT OF INTERIOR

DONALD PAUL HODEL, Secretary

U.S. GEOLOGICAL SURVEY

Dallas L. Peck, Director

For additional information
write to:

District Chief
U.S. Geological Survey
1400 Independence Road
Mail Stop 200
Rolla, Missouri 65401

Copies of this report can
be purchased from:

U.S. Geological Survey
Books and Open-File Reports
Federal Center, Bldg. 810
Box 25425
Denver, Colorado 80225

CONTENTS

ILLUSTRATIONS

TABLES

CONVERSION FACTORS

For readers who prefer to use metric (International System) units, conversions for inch-pound units used in this report are listed below.

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 STREAM-GAGING PROGRAM IN MISSOURI

By

Loyd A. Waite

ABSTRACT

This report documents the results of an evaluation of the cost effectiveness of the 1986 stream-gaging program in Missouri. Alternative methods of developing streamflow information and cost-effective resource allocation were used to evaluate the Missouri program. Alternative methods were considered statewide, but the cost effective resource allocation study was restricted to the area covered by the Rolla field headquarters.

The average standard error of estimate for records of instantaneous discharge was 17 percent; assuming the 1986 budget and operating schedule, it was shown that this overall degree of accuracy could be improved to 16 percent by altering the 1986 schedule of station visitations. A minimum budget of $203,870, with a corresponding average standard error of estimate of 17 percent, is required to operate the 1986 program for the Rolla field headquarters; a budget of less than this would not permit proper service and maintenance of the stations or adequate definition of stage-discharge relations. The maximum budget analyzed was $418,870, which resulted in an average standard error of estimate of 14 percent. Improved instrumentation can have a positive effect on streamflow uncertainties by decreasing lost record.

An earlier study of data uses found that data uses were sufficient to justify continued operation of all stations. One of the stations investigated, Current River at Doniphan (07068000) was suitable for the application of alternative methods for simulating discharge records. However, the station was continued because of data use requirements.

INTRODUCTION

The U.S. Geological Survey is the principal Federal agency collecting surface-water data in the Nation. The collection of these data is a significant activity of the Geological Survey. The data are collected in cooperation with State and local governments and other Federal agencies. Currently (1986), the Geological Survey is operating approximately 7,500 continuous-record gaging stations throughout the Nation. Some of these records date back to the turn of the century.

Any activity of long standing, such as the collection of surface-water data, should be re-examined at intervals, if not continually, 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 Missouri contribution to that evaluation was done by Skelton and Homyk (1970). In 1983, the Geological Survey undertook another nationwide analysis of the streamflow-gaging program. The analysis is to be completed over a 5-year period; 20 percent of the program is to be analyzed each year. The objective of the nationwide analysis is to define and document the most cost-effective way to furnish streamflow information. Most of the sections of this report that describe techniques or methodology are taken directly from earlier reports (Fontaine and others, 1984, and Engel and others, 1984).

Phases of the Analysis

The nationwide analysis of the streamflow-gaging program is designed to comprise three major phases of analysis. The first phase is to analyze data use and availability, the second is to identify less costly alternative methods of furnishing streamflow information, and the third phase is to use statistical techniques to evaluate the operation of gaging station networks using associated uncertainty in streamflow records for various operating budgets.

The first phase of the analysis for Missouri -- to analyze data use and availability -- was completed in a report by Waite (1984). The report "Data Uses and Funding of the Stream-Gaging Program in Missouri", documents a survey that identified local, State, and Federal uses of data from 100 continuous-record, surface-water stations that were operated in 1983 by the Missouri District of the U.S. Geological Survey. The report also identified sources of funding pertaining to collection of streamflow data, and presented frequency of data availability. The uses of data from the stations were categorized into seven classes: Regional Hydrology, Hydrologic Systems, Legal Obligations, Planning and Design, Project Operation, Hydrologic Forecasts, and Water-Quality Monitoring. The report noted that there were sufficient uses of the surface-water data collected from the stations to justify continuous operation of all stations.

The purpose of this report is to present the second and third phases of the nationwide analysis as applied to Missouri. The second phase of the analysis -- to identify less costly alternate methods of furnishing streamflow information -- was applied to those stations in the Statewide network that were highly correlated with other stations. The third phase of the analysis -- to evaluate the uncertainty in streamflow records for various operating budgets -- was limited to the network of stations operated by the Rolla field headquarters of the Missouri District, U.S. Geological Survey. This network consists of stations in the Osage, Gasconade, Meramec, St. Frances, Missouri, Mississippi, White, and Arkansas River basins in southern Missouri and represents approximately half the total surface-water stations operated within the Missouri District. The evaluation of that network was considered an adequate sample to address the cost effectiveness of the overall streamgaging program in Missouri and to provide a basis for considering changes in operating procedures.

Missouri Streamflow-Gaging Program

The Missouri streamflow-gaging program has evolved through the years to meet Federal, State, and local needs for surface-water data. The streamflow-gaging network of stations (table 1) as described by Waite (1984) and as evaluated in this report is shown in figure 1.

Table 1.--Selected hydrologic data for 101 continuous-record streamflow-gaging stations in the 1983 surface-water program (from Waite, 1984

[mi², square miles; ft³/s, cubic feet per second]

Map no. (fig. 1)	USGS station no.	Streamflow-gaging station name	Drainage area (mi²)	Period of record	Mean annual flow (ft³/s)
1	05495000	Fox River at Wayland, Missouri	400	February 1922-	251
2	05496000	Wyaconda River above Canton, Missouri	393	October 1932-September 1972, October 1979-	238
3	05497000	North Fabius River at Monticello, Missouri	452	February 1922-	285
4	05498000	Middle Fabius River near Monticello, Missouri	393	July 1945-	265
5	05500000	South Fabius River near Taylor, Missouri	620	October 1934-	401
6	05501000	North River at Palmyra, Missouri	373	December 1934-	252
7	05502000	Bear Creek at Hannibal, Missouri	31.0	October 1938-September 1942, October 1947-	20.0
8	05502300	Salt River at Hagers Grove, Missouri	365	September 1974-	280
9	05503500	North Fork Salt River near Hunnewell, Missouri	626	April 1930-September 1931, October 1931-September 1940 [a], March 1966-October 1970, October 1979-	403
10	05503800	Crooked Creek near Paris, Missouri	80.0	March 1966-(Prior to 1979 published by U.S. Army, Corps of Engineers)	(b)
11	05505000	South Fork Salt River at Santa Fe, Missouri	298	October 1939- [c]	195
12	05506500	Middle Fork Salt River at Paris, Missouri	356	October 1939-	246
13	05506800	Elk Fork Salt River near Madison, Missouri	200	October 1968-	172
14	05507600	Lick Creek at Perry, Missouri	104	March 1966- [d]	(b)
15	05507800	Salt River near Center, Missouri	2,350	April 1963-	(b)
16	05508000	Salt River near New London, Missouri	2,480	February 1922-	1,691
17	05508805	Spencer Creek below Plum Creek near Frankford, Missouri	206	March 1930-September 1936, October 1961-(Prior to 1978, fragmentary record)	(b)
18	05514500	Cuivre River near Troy, Missouri	903	February 1922-July 1972, May 1979-	651
19	05587500	Mississippi River at Alton, Illinois	171,500	October 1927-	101,300
20	06813000	Tarkio River at Fairfax,	508	March 1922-	201

Table 1.--Selected hydrologic data for 101 continuous-record streamflow-gaging stations in the 1983 surface-water program (from Waite, 1984--Continued

Map no. (fig. 1)	USGS station no.	Streamflow-gaging station name	Drainage area (mi^2)	Period of record	Mean annual flow (ft^3/s)
21	06817500	Nodaway River near Burlington Junction, Missouri	1,240	March 1922-	558
22	06818000	Missouri River at St. Joseph, Missouri	420,300	August 1928-	40,200
23	06819500	One-Hundred and Two River at Maryville, Missouri	515	October 1932-	226
24	06820500	Platte River near Agency, Missouri	1,760	May 1924-August 1930, May 1932- [e]	921
25	06821150	Little Platte River at Smithville, Missouri	234	June 1965-	158
26	06821190	Platte River at Sharps Station, Missouri	2,380	December 1978-	(b)
27	06893000	Missouri River at Kansas City, Missouri	485,200	October 1897-	54,950
28	06893500	Blue River near Kansas City, Missouri	188	May 1939-	148
29	06893793	Little Blue River below Longview damsite, Missouri	50.7	August 1966-	38.3
30	06893890	East Fork Little Blue River near Blue Springs, Missouri	34.4	October 1974-	24.8
31	06894000	Little Blue River near Lake City, Missouri	184	March 1948-	147
32	06895500	Missouri River at Waverly, Missouri	487,200	October 1928-	49,650
33	06897500	Grand River near Gallatin, Missouri	2,250	June 1921-	1,170
34	06899500	Thompson River at Trenton, Missouri	1,670	June 1921-September 1923, August 1928- [f]	951
35	06900000	Medicine Creek near Galt, Missouri	225	July 1921-September 1975, October 1977-	145
36	06902000	Grand River near Sumner, Missouri	6,880	October 1923-	3,863
37	06904050	Chariton River at Livonia, Missouri	864	May 1974-	719
38	06904500	Chariton River at Novinger, Missouri	1,370	October 1930-September 1952, October 1954-	825
39	06905500	Chariton River near Prairie Hill, Missouri	1,870	October 1928- [g]	1,210
40	06906000	Mussel Fork near Musselfork, Missouri	267	October 1948-December 1951, October 1962-	232

Table 1.--Selected hydrologic data for 101 continuous-record streamflow-gaging stations in the 1983 surface-water program (from Waite, 1984--Continued

Map no. (fig. 1)	USGS station no.	Streamflow-gaging station name	Drainage area (mi^2)	Period of record	Mean annual flow (ft^3/s)
41	06906200	East Fork Little Chariton near Macon, Missouri	112	September 1971-	107
42	06906300	East Fork Little Chariton River near Huntsville, Missouri	220	October 1962-	243
43	06908000	Blackwater River at Blue Lick, Missouri	1,120	June 1922-September 1933, May 1938-	712
44	06909000	Missouri River at Boonville, Missouri	501,700	October 1925-	59,260
45	06918440	Sac River near Dadeville, Missouri	257	June 1966-	213
46	06918460	Turnback Creek above Greenfield, Missouri	252	September 1965-	235
47	06918740	Little Sac River near Morrisville, Missouri	237	October 1968-	210
48	06919000	Sac River near Stockton, Missouri	1,160	July 1921-	959
49	06919020	Sac River at Highway J below Stockton, Missouri	1,292	October 1973-	1,012
50	06919500	Cedar Creek near Pleasant View, Missouri	420	April 1923-September 1926, October 1948	294
51	06919900	Sac River near Caplinger Mills, Missouri	1,810	October 1974-	1,230
52	06921070	Pomme De Terre River near Polk, Missouri	276	October 1968-	241
53	06921200	Lindley Creek near Polk, Missouri	112	April 1957-	87.4
54	06921350	Pomme De Terre River near Hermitage, Missouri	615	August 1960-	457
55	06921590	South Grand River at Archie, Missouri	356	October 1969-	248
56	06922450	Osage River below Truman Dam at Warsaw, Missouri	7,856	May 1978-	(b)
57	06926000	Osage River near Bagnell, Missouri	14,000	October 1880-	9,574
58	06926500	Osage River near St. Thomas, Missouri	14,500	August 1931-	9,933
59	06932000	Little Piney Creek at Newburg, Missouri	200	October 1928-	151
60	06933500	Gasconade River at Jerome, Missouri	2,840	April 1903-July 1906,[h] January 1923-	2,485

Table 1.--Selected hydrologic data for 101 continuous-record streamflow-gaging stations in the 1983 surface-water program (from Waite, 1984--Continued

Map no. (fig. 1)	USGS station no.	Streamflow-gaging station name	Drainage area (mi^2)	Period of record	Mean annual flow (ft^3/s)
61	06934500	Missouri River at Hermann, Missouri	524,200	October 1897-	80,050
62	07010000	Mississippi River at St. Louis, Missouri	697,000	January 1861-	178,300
63	07013000	Meramec River near Steelville, Missouri	781	October 1922-	564
64	07014500	Meramec River near Sullivan, Missouri	1,475	September 1921-September 1933, October 1943-	1,166
65	07015720	Bourbeuse River near High Gate, Missouri	135	July 1965-	118
66	07016500	Bourbeuse River at Union, Missouri	808	June 1921-	634
67	07017200	Big River at Irondale, Missouri	175	July 1965-	173
68	07018000	Big River near DeSoto, Missouri	718	October 1948-	659
69	07018500	Big River at Byrnesville, Missouri	917	October 1921-	835
70	07019000	Meramec River near Eureka, Missouri	3,788	August 1903-July 1906, October 1921-	3,060
71	07020500	Mississippi River at Chester, Illinois	708,600	October 1927-	186,700
72	07021000	Castor River at Zalma, Missouri	423	January 1920-	511
73	07022000	Mississippi River at Thebes, Illinois	713,200	October 1932-[i]	192,500
74	07037500	St. Francis River near Patterson, Missouri	956	October 1920	1,103
75	07039500	St. Francis River at Wappapello, Missouri	1,311	October 1940-	1,536
76	07042500	Little River ditch 251 near Lilbourn, Missouri	235	October 1945-	324
77	07043500	Little River ditch 1 near Morehouse, Missouri	450	October 1945-	527
78	07050580	James River near Strafford, Missouri	165	October 1973-	153
79	07050700	James River near Springfield, Missouri	246	October 1955-	215
80	07052500	James River at Galena, Missouri	987	October 1921-	942

Table 1.--Selected hydrologic data for 101 continuous-record streamflow-gaging stations in the 1983 surface-water program (from Waite, 1984--Continued

Map no. (fig. 1)	USGS station no.	Streamflow-gaging station name	Drainage area (mi^2)	Period of record	Mean annual flow (ft^3/s)
81	07053500	White River near Branson, Missouri	4,022	October 1951-	3,513
82	07057500	Northfork River near Tecumseh, Missouri	561	October 1944-	706
83	07058000	Bryant Creek near Tecumseh, Missouri	570	October 1944-	511
84	07061300	East Fork Black River at Lesterville, Missouri	94.5	January 1960-	108
85	07061500	Black River at Annapolis, Missouri	484	April 1939-	568
86	07062500	Black River at Leeper, Missouri	987	June 1921-	954
87	07063000	Black River at Poplar Bluff, Missouri	1,245	October 1936-September 1937, October 1939-	1,286
88	07066000	Jacks Fork at Eminence, Missouri	398	October 1921-	441
89	07067000	Current River at Van Buren, Missouri	1,667	October 1912-	1,859
90	07068000	Current River at Doniphan, Missouri	2,038	October 1918-	2,712
91	07068250	Middle Fork Little Black River at Grandin, Missouri	6.85	October 1980-	(b)
92	07068300	North Prong Little Black River near Grandin, Missouri	39.4	April 1980-	(b)
93	07068380	Little Black River near Grandin, Missouri	79.5	May 1980-	(b)
94	07068510	Little Black River below Fairdealing, Missouri	194	May 1980-	(b)
95	07068540	Logan Creek at Oxly, Missouri	37.5	August 1980-	(b)
96	07068600	Little Black River at Success, Arkansas	386	October 1980-	(b)
97	07068863	Fourche Creek near Poynor, Missouri	87.2	January 1976-	104
98	07071500	Eleven Point River near Bardley, Missouri	793	October 1921-	752
99	07186000	Spring River near Waco, Missouri	1,164	April 1924-	844
100	07186400	Center Creek near Carterville, Missouri	232	June 1962-	188
101	07187000	Shoal Creek above Joplin, Missouri	427	October 1941-	390

a Published as "Salt River near Hunnewell."
b No mean annual flow published, less than 5 years of streamflow record.
c October 1969 to September 1975 published as "near Santa Fe."
d Prior to October 1979, gage heights only by St. Louis District Corps of Enginners.
e May 1924 to August 1930 published as "at Agency."
f June 1921 to September 1923 published as "near Hickory."
g Prior to October 1953 published as "near Keytesville."
h April 1903 to July 1906 published as "at Arlington."
i Prior to April 1941 published as "at Cape Girardeau."

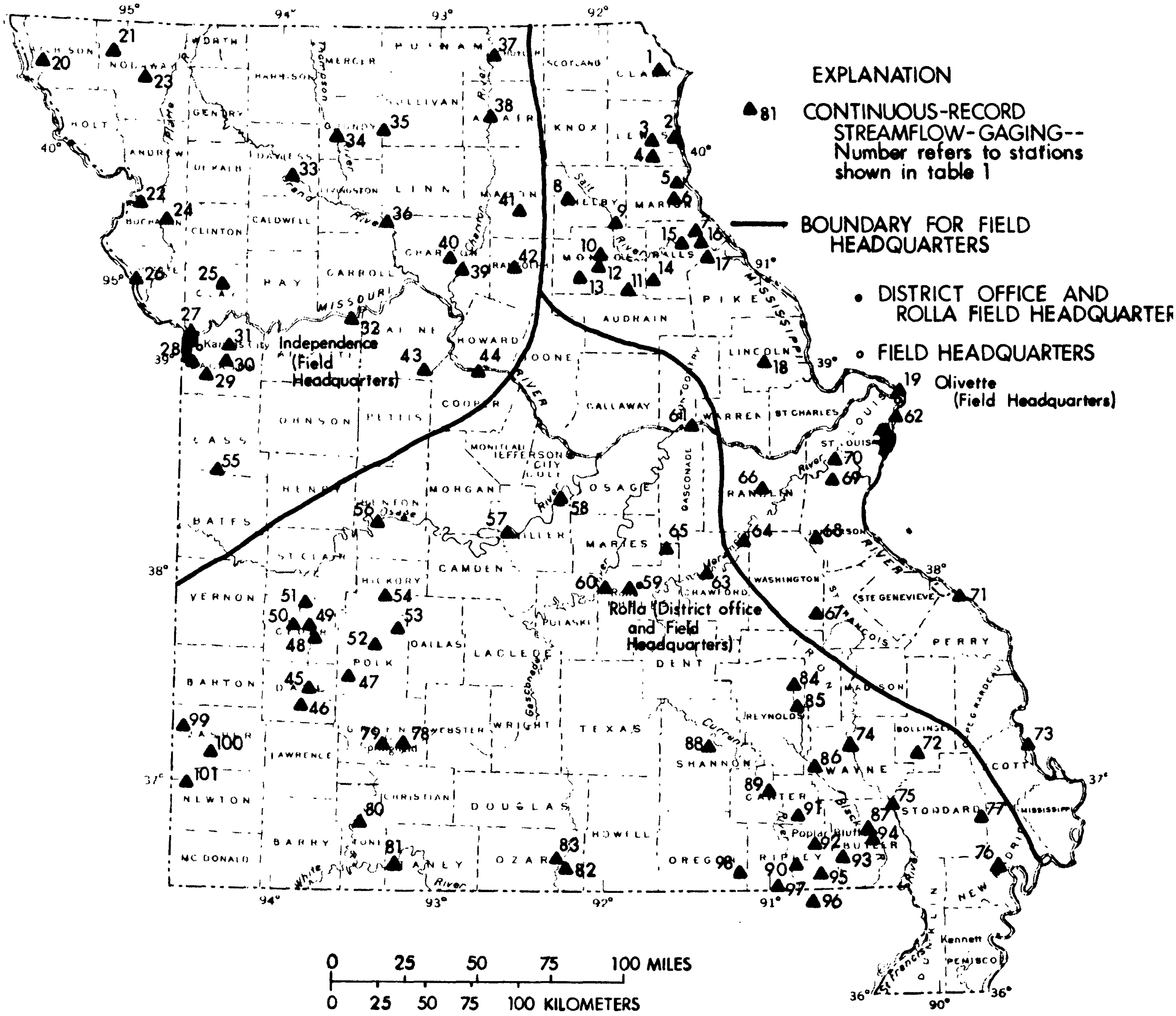

Figure 1.--Location of continuous-record streamflow-gaging stations, district office, field headquarters, and areas of responsibility.

The operation of the streamflow-gaging network is shared by the field headquarters at Rolla, Independence and Maryland Heights (moved to Olivette, Missouri, November 1986). The Rolla field headquarters operates stations in the southern half of the State (fig. 1), Independence the northwest quadrant of the State, and Maryland Heights the eastern area along the Mississippi River.

The streamflow-gaging program has remained fairly stable since Waite (1984) reported on the 100 station network that was in place in 1983. The alternative methods section of this report will deal with selected stations from the 101 station network that was in operation in 1983. The cost-effective resource allocation phase of this report will analyze the 47 streamflow-gaging station network currently (1986) operated by the Rolla field headquarters.

ALTERNATIVE METHODS OF DEVELOPING STREAMFLOW INFORMATION

The second phase of the analysis of the stream-gaging program investigates alternative methods of providing daily streamflow information instead of operating continuous-flow gaging stations. The objective of this phase of the analysis was to identify gaging stations where alternative technology, such as flow-routing or statistical regression methods, could provide accurate estimates of daily mean streamflow efficiently. No guidelines exist concerning suitable accuracies for particular uses of the data; therefore, judgment was required in deciding whether the accuracy of the estimated daily flows would be adequate for the intended purpose.

The data uses at a station affect whether or not information can potentially be provided by alternative methods. For example, those stations for which flood hydrographs are required in a real-time sense, such as hydrologic forecasts and project operation, are not candidates for the alternative methods. Likewise, there might be a legal obligation to operate an actual gaging station that would preclude using alternative methods. Data uses for the U.S. Geological Survey stations in Missouri were previously defined by Waite (1984).

The primary candidates for alternative methods are stations that are operated upstream or downstream from other stations on the same stream. The accuracy of the estimated streamflow at these sites may be adequate if flows are correlated between sites. Gaging stations in similar watersheds, located in the same physiographic and climatic area, also may have potential for alternative methods.

Discussion of Methods

Desirable attributes of a proposed alternative method are: (1) computer oriented and easy to apply, (2) have an available interface with the U.S. Geological Survey's WATSTORE Daily Values File (Hutchison, 1975), (3) technically sound and generally acceptable to the hydrologic community, and (4) provide a measure of the accuracy of the simulated streamflow records. Because of the short duration of this analysis, only two methods were considered; hydrologic routing and regression.

Stations in the Missouri stream-gaging program were screened to determine their potential for use of alternative methods, and selected methods were applied at those stations where the potential was great. The applicability of alternative methods to specific stream-gaging stations is described in this section of this report.

Description of Flow-Routing Model

Hydrologic flow-routing methods use the law of conservation of mass and the relation between the storage in a reach and the outflow from the reach. The hydraulics of the system are not considered. The methods usually require only a few parameters, and the reach is not subdivided. A discharge hydrograph is required at the upstream end of the reach, and the computations produce a discharge hydrograph at the downstream end of the reach. Hydrologic routing methods include the Muskingum, Modified Puls, Kinematic Wave, and the unit-response flow-routing methods. The unit-response method uses one of two routing techniques--storage continuity (Sauer, 1973) and diffusion analogy (Keefer, 1974; Keefer and McQuivey, 1974).

The unit-response method has been widely used to route streamflow from one or more upstream locations to a downstream location is available as a documented computer program (Doyle and others, 1983). The model treats a stream reach as a linear one-dimensional system in which the downstream hydrograph is computed by multiplying (convoluting) the coordinates of the upstream hydrograph by a derived unit-response function and time lagging them appropriately for the channel routed distance. The model has the capability of combining hydrographs, multiplying a hydrograph by a ratio, and changing the timing of a hydrograph.

For most streams daily flows usually can be computed using a single unit-response function (linearization about a single discharge) to represent the system response. However, if the routing coefficients vary significantly with discharge, linearization about a low-range discharge results in overestimated high flows that arrive late at the downstream site, and linearization about a high-range discharge results in low-range flows that are underestimated and arrive too soon. Multiple linearization (Keefer and McQuivey, 1974), in which separate unit-response functions are defined for different ranges of discharge, minimizes this problem.

Determination of the system's response to an upstream pulse 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. Ungaged inflows usually are estimated by multiplying known flows at an index gaging station by an adjustment factor (for example, the ratio of drainage area at the point of interest to that at the index gage).

In both the storage-continuity and diffusion-analogy methods, the routing parameters are calibrated by trial and error. The analyst must decide if suitable parameters have been derived by comparing the simulated discharge to the observed discharge.

Description of Regression Analysis

Simple- and multiple-regression techniques also can be used to estimate daily flow records. Unlike hydrologic routing, regression methods are not limited to locations where an upstream station exists on the same stream. Regression equations can be computed that relate daily flows (or their logarithms) at a station (dependent variable) to daily flows at another station or at a combination of upstream, downstream, or tributary stations. The independent variables in the regression analysis can include stations from different watersheds.

The regression method is easy to apply, provides indices of accuracy, and is widely used and accepted in hydrology; the theory and assumptions are described in numerous textbooks such as Draper and Smith (1966) and Kleinbaum and Kupper (1978). The application of regression methods 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 commonly is used for estimating daily mean discharges:

$$Y_i = B_o + \sum_{j=1}^{n} B_j X_j + e_i \tag{1}$$

where

Y_i = daily mean discharge at station i (dependent variable),

X_j = daily mean discharge(s) at n station(s) j (independent variables); these values may be lagged to approximate travel time between stations i and j,

B_o and B_j = regression constant and coefficients, and

e_i = 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 (Hutchison, 1975). The values of discharge for the independent variables may be observed on the same day as discharges at the independent station or may be for previous or future days, depending on whether station j is upstream or downstream of station i. During calibration, the regression constant and coefficients (B_o and B_j) are tested to determine if they are significantly different from zero. A given independent variable is retained in the regression equation only if its regression coefficient is significantly different from zero.

The regression needs to be calibrated using one period of time and verified or tested using a different period of time to obtain a measure of the true predictive accuracy. Both the calibration and verification periods need to be representative of the expected range of flows. The equation can be verified by: (1) Plotting the residuals (difference between simulated and observed discharges) against both the dependent and the independent variables in the equation, and (2) plotting the simulated and observed discharges versus time. These tests are needed to confirm that the linear model is appropriate and that there are no time trends reflected in either the data or the equation. The presence of either nonlinearity or bias requires that the data be transformed (for example, by converting to logarithms) or that a different form of the model be used.

The use of a regression relation to produce a simulated record at a discontinued gaging station causes the variance of the simulated record to be less than the variance of an actual record of streamflow at the site. The reduction in variance is not a problem if the only concern is with deriving the best estimate of a given daily mean discharge record. If, however, the simulated discharges are to be used in additional analyses where the variance of the data are important, least-squares regression models are not appropriate. Hirsch (1982) discusses this problem and describes several models that preserve the variance of the original data.

Potential for Use of Alternative Methods

A two-level screening process was applied to gaging stations in Missouri to evaluate the potential for use of alternative methods. The first level was based only on hydrologic considerations; the only concern at this level was whether it was hydrologically possible to simulate flows at a given station from information at other gages. The first-level screening was subjective; there was no attempt at that level to apply any mathematical procedures. Those stations that passed the first level of screening (table 2) were then screened again to determine if simulated data would be acceptable in view of the data uses defined by Waite (1984). Even if simulated data were not acceptable for the given data uses, the analysis continued. Mathematical procedures were applied to determine if it were technically possible to simulate data. This was done under the assumption that the data uses may change in the future. Where data uses required continuation of gaging, however, the result was predetermined to be that although alternative methods were technically possible, they were unacceptable given the present uses of the data.

Combinations of stations identified in the first level of screening are listed in table 2. The location of these stations is shown in figure 1. Correlation coefficients were determined for the combinations of stations shown in table 2 to eliminate from consideration those stations that showed little correlation with corresponding stations. Combinations of stations that showed a correlation >0.90 were passed on to the regression analysis.

Regression Results

Linear-regression results were applied to two of the combinations shown in table 2. The two combinations considered were 06904050 (Chariton River at Livonia) and 06905500 (Chariton River near Prairie Hill); 07067000 (Current River at Van Buren) and 07068000 (Current River at Doniphon). The daily streamflow values for the primary station (the dependent variable) were related to concurrent daily streamflow values at the investigated station (explanatory variables) during a given period of record (the calibration period).

The results of regression for stations 06904050 and 06905500 were not presented here as 72 percent of the computed values departed more than 50 percent from the gaging station data. The results of regression for station 07068000 (Current River at Doniphan) are good and shown below. The regression equation for daily mean discharge, Q, in cubic feet per second was defined as:

$$(Q\ 07068000) = 1204 + (.89)\ (Q\ 07067000)$$

and standard error was 11 cubic feet per second. A summary of the regression analyses is shown in table 3.

As a result of this preliminary evaluation by regression analysis, the application of streamflow routing was pursued to use in lieu of operating a complete record gaging station.

Table 2.--Seasonally adjusted correlation coefficients for stations considered in the alternative-methods analysis. Based on season April 1 through September 30

Map No. (fig. 1)[1]	Primary Station	Map No. (fig. 1)[1]	Station Investigated	Lag days	Data pairs	Correlation coefficient
4	05498000	3	05497000	0	10,732	0.8706
5	05500000	4	05498000	0	10,786	0.8461
9	05503500	8	05502300	1	1,460	0.7688
16	05508000	15	05507800	0	1,460	0.6223
30	06893890	29	06893793	0	2,920	0.7021
33	06897500	36	06902000	1	4,915	0.6507
39	06905500	37	06904050	0	5,336	0.9183
68	07018000	67	07017200	1	661	0.7918
87	07063000	86	07062500	1	10,949	0.8924
90	07068000	89	07067000	1	10,949	0.9469

[1] See table 1 for station names.

Table 3.--Summary of regression analyses for mean-daily streamflow for the period from April 1 to December 31

Gaging-station number and regression equation	Percent of days within indicated percentage deviation for calibration period 1981-84 water years ± 10	20	30	50	for verification period 1979-81 water years ± 10	20	30	50
Q07068000 = 1204 + (.89) (Q07067000)	50	70	93	100	50	70	90	100

Flow-Routing Model Results

The CONROUT model (Doyle and others, 1983) requires two parameters:

C_o = flood wave celerity (controls travel time), and

K_o = dispersion or damping coefficient (controls spreading of the wave).

C_o and K_o are approximated from the following expressions:

$$K_o = Q_o / (2 S_o W_o)$$

$$C_o = (1/W_o) \quad (dQ_o/dy)$$

where

W_o = average channel width (ft) in the reach

S_o = average bed slope (ft/ft) in the reach

Q_o = the stream discharge of interest (ft^3), and

dQ_o/dY = the slope of the stage-discharge curve.

These parameters were estimated for the reach of the Current River between Van Buren (07067000) and Doniphan (07068000) gages and were refined based on the application of the model to the calibration period, 1930-31 and 1980-81. The calibrated model was then used to simulate mean-daily discharges for the verification period, 1982-83.

The net contributing drainage areas are 1,667 sq mi for Van Buren and 2,038 sq mi for Doniphan. The model was used to simply route the flow at Van Buren to Doniphan as there is no single significant drainage contribution. Results of the calibration and verification are shown in table 2.

The flow routing model was applied to Current River at Van Buren (07067000) and Current River at Doniphan (07068000). The results are shown on table 4. It was determined that Current River at Doniphan could be computed using flow-routing techniques with acceptable results.

Summary of Second Phase of Analysis

None of the stations investigated presently are suitable for the application of alternative methods. Only at Current River at Doniphan (07068000) is the accuracy of the flow-routing model sufficient to consider discontinuing the gage; however, the data uses require the gage to be continued.

COST-EFFECTIVE RESOURCE ALLOCATION

Discussion of the Model

A set of techniques called K-CERA (Kalman filtering for Cost-Effective Resource Allocation) was developed by Moss and Gilroy (1980) to study the cost-effectiveness of networks of stream gages. The original application of the technique was to analyze a network of stream gages operated to determine water consumption in the Lower Colorado River Basin (Moss and Gilroy, 1980). Because of the water-balance nature of that study, the minimization of the total

Table 4.--Summary showing selected characteristics of results used for the routing model as applied to the reach of the Current River between the Van Buren (07067000) and Doniphan (07068000) gages

Daily discharge errors	Percent: Calibration 1930-31	Percent: Calibration 1980-81	Percent: Verification 1982-83
Less than or equal 5 percent	54	36	69
Less than or equal 10 percent	76	89	85
Less than or equal 15 percent	93	95	94
Less than or equal 20 percent	98	97	98
Less than or equal 25 percent	98	98	98
Greater than 25 percent	2	2	2
Mean error in percent for 365 days	6.3	6.7	5.0

Q_o[1]	W_o[2]	S_o[3]	C_o[4]	K_o[5]	X_o[6]
1860	240	.000602	5.42	12,870	202,800

[1] Q_o stream discharge in cubic feet per second.
[2] W_o average channel width for the study reach, in feet.
[3] S_o average bed slope in feet per feet.
[4] C_o flood wave celerity in feet per second.
[5] K_o wave dispersion or damping coefficient in feet squared per second.
[6] X_o length of the study channel in feet.

variance of errors of estimation of annual mean discharges was chosen as the measure of effectiveness of the network. This total variance is defined as the sum of the variances of errors of mean annual discharge at each site in the network. This measure of effectiveness tends to concentrate stream-gaging resources on the large rivers and streams where discharge and, consequently, potential errors (in cubic feet per second) are greatest. Although this may be acceptable for a water-balance network, considering the many uses of data collected by the U.S. Geological Survey, concentration of effort on larger rivers and streams is undesirable and inappropriate.

The original version of K-CERA was therefore altered to include as optional measures of effectiveness the sums of the variances of errors of estimation of the following streamflow variables: annual mean discharge, in cubic feet per second; annual mean discharge, in percent; average instantaneous discharge, in cubic feet per second; or average instantaneous discharge, in percent (Fontaine and others, 1984). The use of percentage errors effectively gives equal weight to large and small streams. In addition, 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 continuously gaged sites as the measure of the effectiveness of the data-collection activity.

The original version of K-CERA also did not 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 (Fontaine and others, 1984) and was incorporated into this study.

Brief descriptions of the mathematical program used to minimize the total error variance of the data-collection activity for given budgets and of the application of Kalman filtering (Gelb, 1974) to the determination of the accuracy of a stream-gaging record are presented by Fontaine and others (1984); that description is reproduced in the Supplemental Information section at the end of this report. For more detail on either the theory or the applications of the K-CERA model, see Moss and Gilroy (1980) and Gilroy and Moss (1981).

Application of the Model in Missouri

The first two phases of this analysis showed that operation of the current network of stream gages in Missouri needs to be continued. The Rolla field headquarters network was selected and analyzed by the K-CERA technique to evaluate the current operation and to consider alternative operating schedules. The results of this third and final phase of the analysis are described in the remainder of this section.

The model assumes the uncertainty of discharge records at a given gage to be derived from three sources: (1) errors that result because the stage-discharge relationship is not perfect (applies when the gage is operating); (2) errors in reconstructing records based on records from another gage when the primary gage is not operating; and (3) errors inherent in estimated discharge when the gage is not operating and no correlative data are available to aid in record reconstruction. These uncertainties are measured as the variance of the percentage errors in instantaneous discharge. The proportion of time that each source of error applies is dependent on the frequency at which the equipment is serviced.

Definition of Variance When the Station is Operating

The model used in this analysis assumes the difference (residual) between instantaneous discharge (measurement discharge) and rating curve discharge is a continuous first-order Markov process. The underlying probability distribution is assumed to be Gaussian (normal) with a zero mean; the variance of this distribution is referred to as process variance. Because the total variance of the residuals includes error in the measurements, the process variance is defined as the total variance of the residuals minus the measurement error variance.

Computation of the error variance about the stage-discharge relation was done in three steps. A long-term rating was defined, generally based on measurements made during 3 or more water years, and deviations (residuals) of the measured discharges from the rating discharge were determined. A time-series analysis of these residuals defined the 1-day lag (lag-one) autocorrelation coefficient and the process variance required by the K-CERA model. Finally, the error variance is defined within the model as a function of the lag-one autocorrelation coefficient, the process and measurement variances, and the frequency of discharge measurements.

In the Rolla field headquarters program analysis, definition of long-term rating functions was complicated by the fact that many stream gages in Missouri are affected by backwater from ice for about 3 months during the year. Rating curves based on open-water measurements are not applicable during the ice-affected periods.

In the pilot study for Maine, winter rating curves were replaced with regression relations relating the discharge at the ice-affected station to the discharge at an ice-free station. The model used this relationship in place of a standard stage-discharge relation, and uncertainties of the ice-affected and ice-free periods were evaluated separately (Fontaine and others, 1984). This approach does not work well in Missouri because of the distances between gages and the variability of flow resulting from the temporary storage and subsequent release of ice. Reliable discharge records during the winter can presently be produced only by making periodic visits and measurements to document the degree of ice effect.

Review of past discharge records indicates that significant ice effects generally occur intermittently from about mid-December to mid-March. The decision was made that, regardless of ice-free period visit requirements, three visits will continue to be made during the winter season. The model then was applied only to the approximately 9 months (275 days) that are virtually free from ice effect.

Long-term rating curves applicable to ice-free periods were defined for each station used in the evaluation. In some cases, existing ratings adequately defined the long-term condition and were used in the analysis. The rating function used was of the following form:

$$LQM = B1 + B3\,[\,\ln(GHT - B2)] \qquad (2)$$

where

LQM = the logarithmic (base e) value of the measured discharge, and
GHT = the recorded gage height corresponding to the measured discharge.

The constants B1, B2, and B3 were determined by a non-linear regression procedure (Helwig and Council, 1979) and have the following physical interpretation: B1 is the logarithm of discharge for a flow depth of one foot, B2 is the gage height of zero flow, and B3 is the slope of the logarithmic rating curve.

The residuals about the long-term rating curve for individual gages defined the total variance. A review of discharge measurements made in Missouri indicated that the average standard error of open-water measurements was about five percent. The measurement variance for all gages, therefore, was defined as equal to the square of the five-percent standard error so the process variance required in the model is the variance of the residuals about the long-term rating minus the constant measurement variance.

Time-series analysis of the residuals was used to compute sample estimates of the lag-one autocorrelation coefficient; this coefficient is required to compute the error variance during the time when the recorders are functioning.

The values of lag-one autocorrelation coefficient, measurement and process variance, length of season (275 days), and data from the definition of missing record probabilities are used jointly to define uncertainty functions for each gaging station. The uncertainty functions give the relation of error variance to the number of visits, assuming a measurement is made at each visit. Examples of typical uncertainty functions are given in figure 2. The uncertainty curve for station 07063000 is representative of stations with a large process variance and that for station 06919000 represents stations with relatively small process variance. Lag-one autocorrelation coefficients are approximately 0.95 for all three stations shown.

The residuals about rating curves for many stations serviced by the District do not approximate a continuous first-order Markov process. These stations have significant changes in ratings resulting from channel changes, usually caused by floods. These may shift with each flood, but will not necessarily return to the original rating after a change. In addition, several stations apparently have discontinuous ratings that change as the flow regime changes. These regime changes can occur as a result of changes in stage, water temperature, or suspended-sediment load. In either case (channel change or regime change), the process may be Markovian, but is not continuous as there is no meaningful long-term rating. In addition, records at nine stations were too short to define the process variance. A total of 24 of the 47 stations analyzed were excluded from the analysis because the records were either too short or did not meet the assumptions of the model. Those stations are listed in table 5.

Definition of Variance When Record is Lost

When stage record is lost at a gaging station, the model assumes that the discharge record is either reconstructed using correlation with another gage or estimated from historical discharge for that period. Fontaine and others (1984, p. 24) indicate that the fraction of time a record must be either reconstructed or estimated can be defined by a single parameter in a probability distribution of times to failure of the equipment. The reciprocal of the parameter defines the average time, since the last servicing visit, to failure. The value of average time to failure varies from site to site depending on the type of equipment at the site and on exposure to natural elements and vandalism. In addition, the average time to failure can be changed by advances in the technology of data collection and recording equipment.

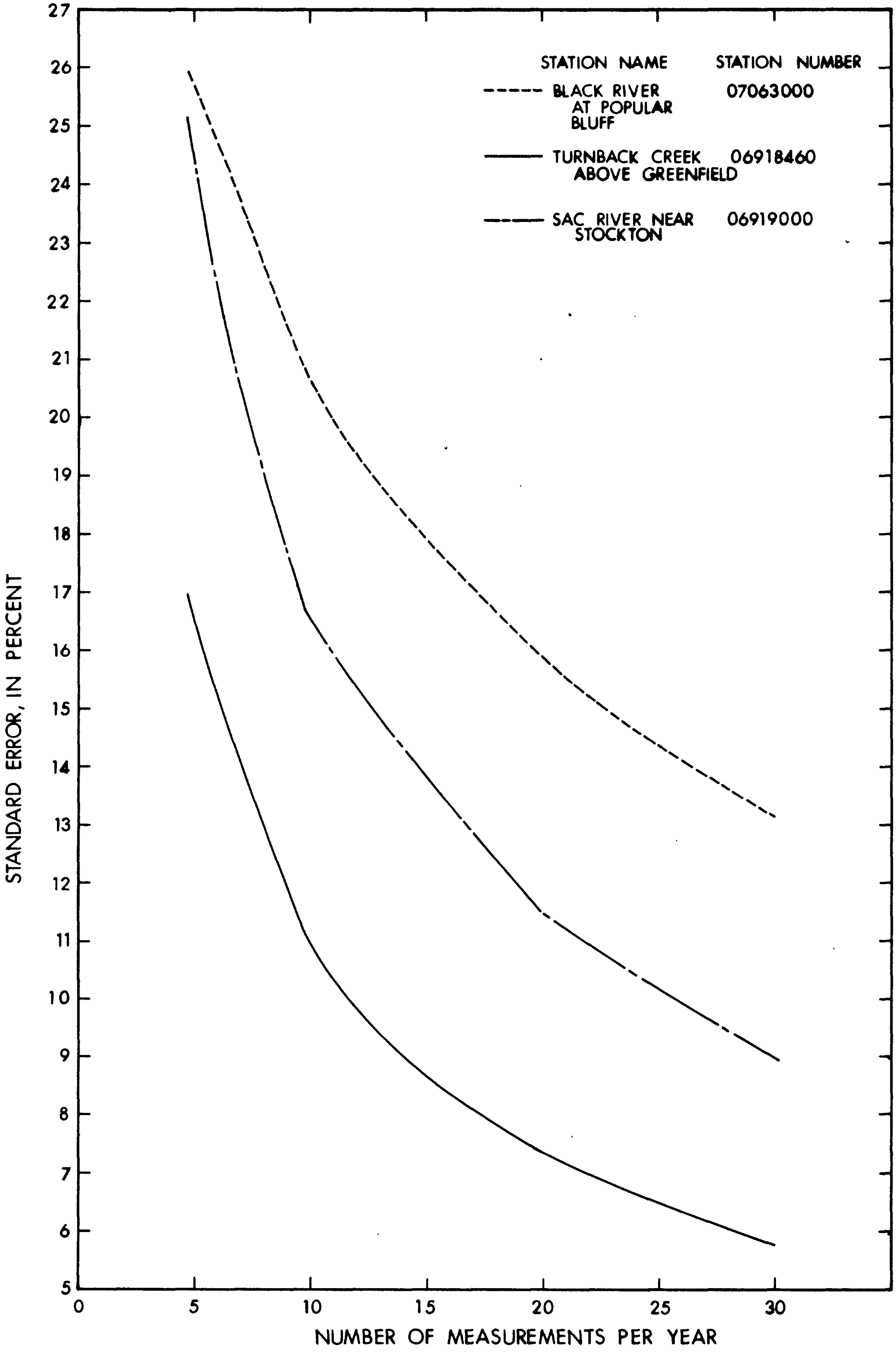

Figure 2.--Uncertainty functions for three gaging stations in Missouri.

Table 5.--Stations with no defined uncertainty function

Station number	Station name
06919900	Sac River near Caplinger Mills
06921200	Lindley Creek near Polk
06922450	Osage River below Harry S. Truman Dam at Warsaw
06926000	Osage River near Bagnell
06932000	Little Piney Creek at Newburg
06934500	Missouri River at Hermann
07013000	Meramec River near Steelville
07015720	Bourbeuse River near Highgate
07034000	St. Francis River near Roselle
07035000	Little St. Francis River at Fredericktown
07036100	St. Francis River near Saco
07037000	Big Creek at Des Arc
07039500	St. Francis River at Wappapello
07042500	Little River Ditch 251 near Lilbourn
07043500	Little River Ditch 1 near Morehouse
07053500	White River near Branson
07058000	Bryant Creek near Tecumseh
07061300	East Fork Black River at Lesterville
07061500	Black River near Annapolis
07066000	Jacks Fork at Eminence
07068000	Current River at Doniphan
07068510	Little Black River near Fairdealing
07068600	Little Black River at Success
07071500	Greer Spring at Greer

Data collected in Missouri in recent years were reviewed to define the average time to failure for recording equipment and stage-sensing devices. Few changes in technology occurred during the period examined, and stream gages were visited on a consistent pattern of about 12 visits per year. During this period, gages were found to be malfunctioning an average of about five percent of the time. Because the K-CERA analysis in Missouri was confined to a 9-month non-winter period, there was no reason to distinguish between gages on the basis of their exposure or equipment. The five percent lost record and a visit frequency of nine times in 9 months (275 days) were used to determine an average time to failure of 221 days after the last visit. This average time to failure was used to determine the fractions of time, as a function of the frequency of visits, that each of the three sources of uncertainty were applicable for individual stream gages.

The model defines the uncertainty as the sum of the multiples of the fraction of time each error source (rating, reconstruction, or estimation) is applicable and the variance of the error source (equation 4 in supplemental information). The variance associated with reconstruction and estimation of a discharge record is a function of the coefficient of cross correlation with the station(s) used in reconstruction and the coefficient of variation of daily discharges at the station. Daily streamflows for the last 30 water years were used to define seasonally-averaged coefficients of variation for each station. In addition, cross-correlation coefficients (with seasonal trends removed) were defined for various combinations with other stations.

In current practice, many different sources of information are used to reconstruct periods of missing record. These sources include, but are not limited to, recorded ranges in stage (for graphic recorders with clock stoppage), known discharges on adjacent days, recession analysis, observer's staff-gage readings, weather records, highwater-mark elevations, and comparison with nearby stations. However, most of these techniques are unique to a given station or to a specific period of lost record. Using all the information available, several days of lost record usually can be reconstructed quite accurately. Longer periods (more than a month) of missing record can be reconstructed with reasonable accuracy if observer's readings are available. If, however, none of these data are available, long reconstructions can be subject to large errors. The uncertainty associated with all the possible methods of reconstructing missing record at the individual sites could not be quantified reasonably for the present study.

Historically, operating procedures have caused most periods of missing record to be measured in days rather than months. Given the low cross-correlations and the relatively high variability of flow that usually occurs in Missouri, the model undoubtedly overstates the uncertainty associated with short periods of missing record. Therefore, in Missouri a lower limit of 0.75 was placed on the cross-correlation coefficient. This affected results at only four stations. In reconstructing records, the cross-correlation coefficient was, therefore, used as a surrogate for the knowledge of basin response that remains unquantified in the present model. This assumption is believed to be reasonable for short periods of missing record; it probably causes the uncertainty to be understated for long periods of lost record.

Uncertainty functions were defined for 23 of the 47 stations operated in the Rolla field headquarters streamflow information program. The statistics used to define those uncertainty functions are shown in table 6.

Discussion of Routes and Costs

Although there are only 47 continuous-record surface-water stations in the Rolla field headquarters network, crest-stage gages (operated to record peak stages) and low-flow partial-record stations are serviced on the same field trips. The operating budgets for these other types of stations are not included in the surface-water budget being analyzed; however, the additional mileage required to include these stations on field trips could not be ignored. These stations were, therefore, added to the 47 continuous surface-water stations to define the mileages associated with practical operating routes. These added stations acted as null stations in the analysis in that there were no uncertainty functions or annual operating costs defined. There were 10 null stations included in the analysis, and routes were defined for a total of 57 stations, including the null stations.

As indicated in a preceding section, uncertainty functions could not be defined for 24 of the 47 continuous surface-water stations. These 24 stations were treated as null stations except that all operating costs were included in the analysis.

Minimum visit constraints were defined for each of the 57 stations before defining the practical service routes. Minimum visits are dependent on the types of equipment and uses of the data. For example, crest stage gages generally are serviced on a monthly basis, so those stations must be visited at least once a month (or nine times in the 275-day open-water season). Missouri personnel estimated that visits to each gage were required about every other month to maintain the equipment. Therefore, unless a more stringent requirement existed, a minimum of four visits during the 275-day season were specified for all gages.

Practical routes to service the 57 stations were determined after consultation with personnel responsible for maintaining the stations and with consideration of the uncertainty functions and minimum visit requirements. A total of seven routes were identified to service all the stream gages in the Rolla field headquarters area. 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 stations, and combinations that grouped approximate gages where the levels of uncertainty indicated more frequent visits might be useful.

The costs associated with the practical routes are divided into three categories. Those categories are fixed costs, visit costs, and route costs, and are defined in the following paragraphs. Overhead costs are, of course, added to the total.

Fixed costs typically include charges for equipment rental, batteries, electricity, data processing and storage, maintenance, and miscellaneous supplies, in addition to supervisory charges and the costs of computing the record. Average values for Missouri generally were applied to individual stations. However, costs of record computation and supervision form a large percentage of the cost at each gaging station and can vary widely. These costs and unusual equipment costs were determined on a station-by-station basis from past experience.

Table 6.--Summary of the Kalman-Filtering analysis

[B_1, B_2, B_3, coefficients as shown for the rating function LQM = $B_1 + B_3$ Ln (GHT-B_2); RHO, 1-day auto correlation coefficient; VPROC, process variance (log base e); uncertainty, uncertainty function in percent error, with respect to number of measurements per period (275 day); C_v, coefficient of variation of daily values; P_c, cross correlation of daily streamflow values with nearby station as listed (log base 10)]

Station number	Station name	Rating coefficients B_1	B_2	B_3	RHO	VPROC	Uncertainty 2	5	10	20	C_v	P_c	Correlated station number
06918440	Sac River near Dadeville	1.83	4.0	0.62	0.980	0.0051	31.1	16.9	11.0	7.3	1.141	0.93	06918460
06918460	Turnback Creek above Greenfield	1.36	4.7	.54	.982	.0009	31.5	16.8	10.8	7.2	1.170	.93	06918440
06918740	Little Sac River near Morrisville	1.77	1.9	.59	.978	.0107	43.9	26.4	18.2	12.7	1.363	.82	06918460
06919000	Sac River near Stockton	1.78	4.8	.71	.995	.0046	45.0	25.2	16.7	11.3	1.571	.90	06919020
06919020	Sac River (Hwy J) below Stockton	2.18	5.4	.61	.993	.0035	33.7	18.9	12.5	8.5	1.172	.90	06919000
06919500	Cedar Creek near Pleasant View	1.42	2.6	.64	.974	.0321	92.8	60.4	43.0	30.5	2.277	.51	06919900
06921070	Pomme de Terre River near Polk	1.85	1.4	.57	.954	.0617	56.4	37.5	27.4	19.7	1.116	.84	06919020
06921350	Pomme de Terre River near Hermitage	1.78	1.0	.53	.915	.0206	65.5	43.6	31.5	22.5	1.511	.86	06921070
06926500	Osage River near St. Thomas	2.30	1.0	.57	.579	.0096	34.2	18.7	13.4	10.9	1.340	.97	06926000
06932000	Little Piney Creek at Newburg	1.00	1.6	.47	.897	.0024	50.6	30.6	21.2	14.9	1.319	.80	06933500
06933500	Gasconade River at Jerome	2.94	1.0	.73	.978	.0003	37.0	22.3	15.3	10.6	0.968	.80	06932000
07021000	Castor River at Zalma	1.18	1.2	.45	.986	.0090	58.5	36.4	25.4	17.9	1.647	.72	07043500
07037500	St. Francis River near Patterson	1.76	2.6	.48	.988	.0935	76.6	49.2	34.9	24.7	1.944	.62	07039500
07050580	James River near Strafford	1.82	2.0	.48	.652	.1665	57.4	47.2	43.3	39.9	1.379	.83	07050700
07050700	James River near Springfield	1.62	2.3	.44	.993	.0065	51.0	29.7	20.7	15.2	1.720	.88	07052500
07052500	James River at Galena	2.15	1.9	.41	.405	.0002	41.7	23.8	15.9	10.9	1.422	.88	07050700
07057500	North Fork River near Tecumseh	2.67	1.6	.59	.979	.0174	30.8	19.3	13.6	9.5	.890	.79	07067000
07062500	Black River at Leeper	2.00	0.0	.50	.400	.0037	28.2	16.6	11.8	9.1	.963	.89	07063000
0706300	Black River at Poplar Bluff	2.65	-1.0	.93	.941	.0644	33.6	25.8	20.7	15.6	.789	.80	07062500
07067000	Current River at Van Buren	1.93	0.4	.62	.992	.0035	25.7	13.4	8.5	5.4	.979	.95	07068000

Table 6.--Summary of the Kalman-Filtering analysis--Continued

Station number	Station name	Rating coefficients B_1	B_2	B_3	RHO	VPROC	Uncertainty 2	5	10	20	C_v	P_c	Correlated station number
07186000	Spring River near Waco	1.30	1.3	.61	.999	.0120	59.5	36.3	25.1	18.0	1.763	.77	07186400
07186400	Center Creek near Carterville	2.01	0.2	.57	.513	.0381	39.6	27.3	22.9	20.5	1.259	.87	07187000
07187000	Shoal Creek above Joplin	2.28	1.8	.57	.990	.0039	35.5	20.4	13.7	9.4	1.882	.87	07186400

Visit costs are those associated with paying the hydrographer for the time making a discharge measurement. These costs vary from station to station depending on the difficulty of the measurement, size of the channel, and quantity of and complexity of equipment serviced. Average visit times were estimated for each station based on past operations. This time was multiplied by the average hourly salary of the hydrographers in Missouri to determine total visit costs.

Route costs include the vehicle cost associated with driving the number of miles required to cover the route, the cost of the hydrographer's time while in transit, the time actually spent at a station servicing the equipment, and any per diem associated with the time needed to complete the trip.

The model was run on a 275-day period with the added requirement that three visits would continue to be made during the remaining 90 days of the year. The fixed costs were computed on an annual basis, but the visit and route costs were only applied when a trip was made. So that all costs could be applied on an annual basis, the visit and route costs for the three winter visits were added to the fixed costs for each station.

Results

The "Traveling Hydrographer Program" uses the uncertainty functions along with the appropriate cost data, route definitions, and minimum visit constraints to optimize the operation of the stream-gaging program. The objective function in the optimization process is the sum of the variances of the errors of instantaneous discharge (in percent squared) for the entire gaging-station network.

The current practices were simulated to define the total uncertainty associated with present practice and to calibrate the model. This was done by restricting the specific routes and number of visits to each stream gage to those now (1986) being used. This was done only to compute the standard errors of present practice; no optimization was done. The restrictions were then released and the model was allowed to define optimal visit schedules for the current budget. The optimization procedure was repeated for other possible budgets. The results for both the present operation and the optimal solutions are shown in figure 3 and in table 7.

The Equivalent Gaussian Spread (EGS) is shown in table 7 (Fontaine and others, 1984, p. 26) and is defined in the Supplemental Information section of this report. The approximate interpretation of EGS is, "Two-thirds of the errors in instantaneous streamflow data will be within plus or minus EGS percent of the reported value."

The analysis was repeated for each budget under the assumption that no stage record is lost. Those results, labeled "No missing record" in figure 3, show the average standard errors of estimate for instantaneous discharge attainable if perfectly reliable systems were available to measure and record stage.

The results in figure 3 and table 7 are based on the assumption that a discharge measurement is made each time that a station is visited. The percentage values also represent only the nine months that are virtually free from ice effect. No estimate is made of the probable errors during ice-affected periods. The upper curve in figure 3 represents the minimum level of uncertainty that can be obtained for a given budget and existing technology.

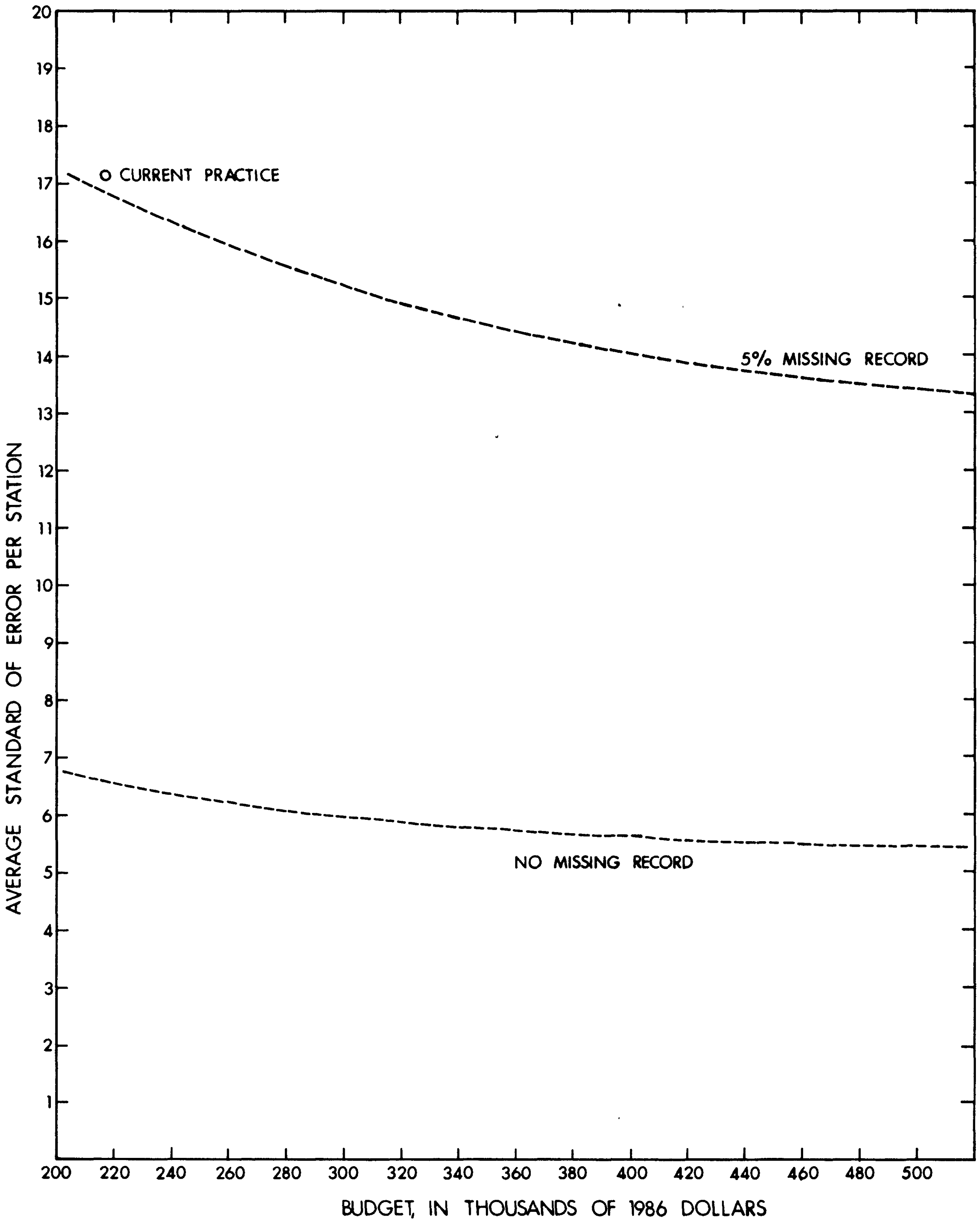

Figure 3.--Relationship between average standard of error per station and budget.

Table 7.--Selected results of the K-CERA analysis

[A, standard error of instantaneous discharge, in percent; B, Equivalent Gaussian Spread (EGS), in percent; C, number of visits per 275 day open water period to site; --, not averaged]

Station		Budget, in 1986 dollars														
		Current			Optimized values											
		$218,870			$203,870			$218,870			$318,870			$418,870		
number	name	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
06918440	Sac River near Dadeville	8.2	2.9	22	8.0	2.9	20	8.0	2.9	22	6.4	2.2	36	6.0	1.8	54
06918460	Turnback Creek above Greenfield	8.0	1.2	22	8.0	1.2	20	8.0	1.2	22	6.0	0.9	36	4.8	0.7	54
06918740	Little Sac River near Morrisville	14.0	4.2	22	14.3	4.3	20	14.0	4.2	22	10.7	3.2	36	8.9	2.7	54
06919000	Sac River Stockton	15.0	1.6	16	17.0	1.9	11	16.5	1.8	11	15.0	1.7	14	13.2	1.5	18
06919020	Sac River (Hwy J) below Stockton	11.3	1.7	16	12.5	1.9	11	12.5	1.9	11	11.0	1.7	14	10.0	1.5	18
06919500	Cedar Creek near Pleasant View	34.4	8.0	16	39.4	9.2	11	39.4	9.2	11	35.0	8.2	14	31.0	7.3	18
06921070	Pomme de Terre River near Polk	17.1	14.2	22	17.3	14.3	20	17.1	14.2	22	13.1	10.9	36	10.8	8.9	54
06921350	Pomme de Terre River near Hermitage	32.5	12.4	13	37.0	12.5	9	35.0	12.3	10	30.5	11.5	9	28.0	10.8	18
06926500	Osage River near St. Thomas	12.6	9.8	16	12.3	9.6	11	12.3	9.6	11	12.3	9.5	14	12.0	9.4	18
06932000	Little Piney Creek at Newburg	13.3	3.6	25	13.4	3.6	22	13.5	3.6	23	10.7	2.9	41	9.8	2.7	54
06933500	Gasconade River at Jerome	9.4	0.7	25	10.6	0.8	22	9.5	0.7	23	8.0	0.6	41	6.3	0.5	54
07021000	Castor River at Zalma	20.7	3.3	15	16.9	2.7	20	16.9	2.7	19	16.4	2.6	22	19.6	3.2	41
07037500	St. Francis River near Patterson	28.6	9.6	15	23.5	7.9	20	24.5	8.2	19	22.2	7.6	22	20.0	7.5	41
07050580	James River near Strafford	45.4	43.0	9	43.5	41.2	10	43.0	41.0	11	41.0	38.9	27	39.3	37.6	36
07050700	James River near Springfield	16.5	8.3	22	16.9	8.5	20	16.5	8.3	22	15.3	7.8	36	15.0	7.8	54
07052500	James River at Galena	12.1	1.5	22	12.1	1.5	20	12.1	1.5	22	11.9	1.4	36	11.7	1.4	54
07057500	North Fork River near Tecumseh	11.4	5.6	19	11.0	5.4	19	10.4	5.1	21	7.7	3.9	36	6.9	3.4	50
07062500	Black River at Leeper	10.0	6.2	15	9.5	5.9	20	9.5	5.9	19	9.4	5.8	22	9.4	5.9	41
07063000	Black River at Poplar Bluff	14.9	13.7	24	13.0	12.0	30	13.0	12.0	30	10.0	9.5	49	9.9	9.4	50
07067000	Current River at Van Buren	6.8	1.7	19	6.5	1.6	19	6.4	1.5	21	5.0	1.2	36	4.7	1.0	50

Table 7.--Selected results of the K-CERA analysis--Continued

Station		Budget, in 1986 dollars														
		Current			Optimized values											
		$218,870			$203,870			$218,870			$318,870			$418,870		
number	name	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
07186000	Spring River near Waco	19.4	1.2	22	19.4	1.2	20	19.4	1.2	22	14.0	0.9	36	13.5	0.9	54
07186400	Center Creek at Carterville	21.6	19.4	22	21.6	19.0	20	21.6	19.4	22	20.4	18.3	36	19.8	17.6	54
07187000	Shoal Creek at Joplin	10.4	1.8	22	10.5	1.8	20	10.4	1.8	22	8.4	1.4	36	6.4	1.1	54
Average per station		17.1	7.6	--	17.1	7.4	--	16.9	7.4	--	14.8	6.6	--	13.8	6.3	--

Assumptions made in the model need to be kept in mind when interpreting these results. In the author's opinion, residuals about the ratings for 20 of the 47 stations in the surface-water network did not follow the first-order Markov process assumed in the model, and records at four stations were too short to analyze. At about one-third of the remaining 27 stations, the assumption of a Markov process was questionable, but the stations were retained in the analysis. This was done believing that while the absolute values of standard error may be incorrect, the values have relative significance. Perhaps of more importance, these 24 stations without uncertainty functions had little impact on the optimization procedure. Because uncertainty functions were undefined, the 24 stations were treated as null stations and were visited monthly, the specified minimum number of times. If the budget changed, the number of visits for these 24 stations stayed at the minimum because increasing or decreasing the visits had no impact on the objective function. In practice, significant parts of any budget increase or decrease would be directed toward those stations.

The current operating policy results in an average standard error of estimate of non-winter streamflow of about 17 percent. This policy is based on a budget of $218,870 to operate the 47-station stream-gaging program. For periods without missing record, the present standard error is slightly less than 10 percent. These figures are within about one percent of the optimum values of standard error for the present budget. Average standard errors apparently could be improved about one percentage point by altering the route schedules to more frequent visits to the sites where uncertainty is large and less frequent visits to sites where uncertainty is small.

A minimum budget of about $203,870 is required to operate the program; a budget of less than this does not permit proper service and maintenance of the gages and recorders, and optimal solutions could not be reached. Stations would have to be eliminated from the program if the budget was less than this minimum. At the minimum budget, the average standard error is about 17 percent, an increase of about 1 percent compared to the accuracy possible under the present budget.

The maximum budget analyzed was $418,870, an increase of about 91 percent compared to the present budget. This resulted in an average standard error of estimate of about 14 percent. Thus, a 91 percent increase in the budget would give a standard error of estimate about 5 percent less than the optimum average standard error obtainable under the current budget.

For the minimal operational budget of $203,870, the impacts of lost record add about 10 percent to the average standard error. At present budget levels, missing record adds about 11 percentage points to the average standard error. With a budget of $418,870, stations would be visited more frequently, and missing record would add about 7 percentage points to the average standard errors. Thus, improvements in equipment can have a positive impact on uncertainties of instantaneous discharges.

Summary of Third Phase of Analysis

As a result of this phase of the analysis, the following conclusions can be made:

1. The schedule of visits in the Rolla field headquarters stream-gaging program could be altered to decrease the average standard error of estimate of stream-flow records from 17 percent to 16 percent at a budget of approximately $218,870 by changes of frequency in visitation. This shift could result in some increases in accuracy of records at individual sites.

2. Stations with accuracies that are not acceptable would require much higher funding levels to significantly improve those accuracies.
3. An exploration of methods or means of including all of the stations in the K-CERA analysis could provide sufficient information about the characteristics of each station so that it can be weighted for its possible decrease of the total standard error of estimate of streamflow records.
4. Methods for decreasing the probablilities of missing record, such as increased use of local gage observers, satellite relay of data, and improved instrumentation, need to be explored and evaluated for their cost effectiveness in providing streamflow information.

SUMMARY

Currently (1986) there are 47 continuous stream gages being operated by the Rolla field headquarters at a cost of $218,870. Data from most stations have multiple uses. Present uses of the data require that operation of all gages be continued. Only 23 of the 47 complete-record stations could be evaluated as to their contribution to decreasing the errors and increasing the cost effectiveness of the program. This is one area that may deserve consideration for further study as funds become available.

It was shown that the overall level of accuracy of the records at 23 of the 47 stations could be improved at the current budget if the frequency of visits was altered in a cost-effective manner. A major component of the error in streamflow records is caused by loss of primary record (stage or other correlative data) at the stream gages because of malfunctions of sensing and recording equipment. Upgrading of equipment and development of strategies to minimize lost record seem to be key actions required to improve the reliability and accuracy of the streamflow data.

Any decrease in the current budget would be accompanied by discontinuing gaging stations because increasing the standard error of estimate is unacceptable. The minimum budget for which a solution could be obtained was $203,870, but that budget results in about a one percent increase in the presently attainable average standard error of estimate.

Future studies of the stream-gaging program need to include investigation of the optimum ratio of discharge measurements to total site visits as well as investigation of cost-effective ways of decreasing the probabilities of lost record.

One station was identified for which streamflow records probably could be simulated on the basis of an upstream station. However, that station, Current River at Doniphan, currently is used in forecasting and needs to be continued. If data uses for this station change so that simulated data are acceptable, alternative methods could be explored.

REFERENCES CITED

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.

Doyle, W.H., Jr., Shearman, J.O., Stiltner, G.J., and Krug, W.R., 1983, Digital model for streamflow routing by convolution methods: U.S. Geological Survey Water-Resources Investigations Report 83-4160, 130 p.

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

Engberg, R.A., 1980, A statistical analysis of the quality of surface water in Nebraska: U.S. Geological Survey Water-Resources Investigations 80-43, 277 p.

Engel, G.B., Wahl, K.L., and Boohar, J.A., 1984, Cost-effectiveness of the stream-gaging program in Nebraska: U.S. Geological Survey Water-Resources Investigations Report 84-4098.

Fontaine, R.A., Moss, M.E., Smath, J.A., and Thomas, W.O., Jr., 1984, Cost-effectiveness of the stream-gaging program in Maine: U.S. Geological Survey Water-Supply Paper 2244, 39 p. Gelb, A., ed., 1974, Applied optimal estimation: Cambridge, Mass., The Massachusetts Institute of Technology Press, 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.

Helwig, J.T., and Council, K.A., eds., 1979, SAS user's guide, 1979 edition: Raleigh, N.C., SAS Institute, Inc., 494 p.

Hirsch, R.M., 1982, A comparison of four streamflow record extension techniques: Water Resources Research, v. 18, no. 4, p. 1081-1088.

Hutchison, 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.

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

Skelton, J., and Homyk, A.J., 1970, A proposed streamflow data program for Missouri: U.S. Geological Survey Open-File Report, 44 p.

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.

Waite, L.A., 1984, Data users and funding of the stream-gaging program in Missouri: U.S. Geological Survey Open-File Report, 32 p.

SUPPLEMENTAL INFORMATION

Description of Uncertainty Functions and the Mathematical Program

In a study of the cost effectivenesss of a network of stream gages operated in the lower Colorado River basin, a methodology called K-CERA was developed (Moss and Gilroy, 1980). The K-CERA methodology considers the cost effectiveness of a network of stream gages to be determined by the total variance, uncertainty, in either the annual mean discharge or the instantaneous discharge at all sites involved in the stream-gaging program and the cost of achieving that uncertainty. For the present (1986) study, the measure of uncertainty at each site was taken to be the variance of the percent error in the instantaneous discharge. (See Fontain and others, 1984, for the argument for this measure of uncertainty).

The first step in estimating a site-specific uncertainty function, a relation between variance and number of visits to the site, is to determine a logarithmic discharge rating curve relating instantaneous discharge to some correlative data, such as gage height, for each station in the stream-gaging program. The sequence of discharge residuals (in logarithmic units) from this rating, the discharge measurement minus the rating value, is analyzed as a time series.

The second step is to fit a lag-one-day autoregressive model to this temporal sequence of discharge residuals. The three parameters obtained from this analysis are (1) the measurement variance, actually estimated a priority, (2) the process variance, a measure of the variability about the rating in the absence of measurement error, and (3) RHO, the lag-one-day autocorrelation coefficient, a measure of the memory in the sequence of discharge residuals. These three parameters determine the variance, V_f, of the percentage error in the estimation of instantaneous discharge whenever the primary correlative data at the site is available for use in the rating equation. Kalman-Filter theory, along with the assumption of a first-order Markovian process, is used to determine this variance V_f as a function of the number of discharge measurements per year (Moss and Gilroy, 1980).

If the primary correlative data at the site is not available, the discharge may be estimated by correlation with nearby sites. The correlation coefficient, r_c , between the streamflows with seasonal trends removed (detrended) at the site of interest and detrended streamflows at the other sites is a measure of the soundness of their linear relation. The fraction of the variance of the streamflow at the primary site that is explained by data from other sites is $r_c{}^2$. The variance of the percent error in streamflows at the primary site in the absence of primary data at both the principal site and nearby sites is taken to be

$$C_v = [\frac{1}{365} \sum_{i=1}^{365} (\frac{\sigma_i}{n_i})^2]^{1/2}, \qquad (1)$$

where s_i is the square root of the variance of daily discharges for the ith day of the year and u_i is the expected value of discharge on the ith day of the year. Thus the variance, V_r, of the percentage error during periods of reconstructed streamflow records is

$$V_r = (1 - P_c{}^2)\ C_r{}^{-2}, \qquad (2)$$

and the variance, V_e, of the percentage error during periods when neither primary correlative data nor reconstructed streamflow from nearby sites is

$$V_e = (Z_v)^2. \qquad (3)$$

If the fraction of time when primary correlative data are available is denoted by e_f and the fraction of time when secondary streamflow data is available for reconstruction is e_r and $e_e = 1 - e_f - e_r$, the total percentage error variance, VT is given by

$$VT = e_f V_f + e_r V_r + e_e V_e. \qquad (4)$$

The fraction uptime, e_f, of the primary recorders at the site of interest is modeled by a truncated negative exponential probability distribution which depends on t, the average time between service visits, and K, which is the reciprocal of the average time to failure when no visits are made to the site. The fraction concurrent downtime of the primary and secondary site is found by assuming independence of downtimes between sites (Fontaine and others, 1984).

The variance VT given by equation 4, and which is a function of the number of visits to the site, is determined for each site in the streamgaging network. For a given site visitation strategy, the sum of the variance, VT, over all sites is taken as the measure of the uncertainty of the network. The variance VT given by equation 4 is one measure of the spread of a probability density function, gT. The function gT is a mixture of three probability density functions, gf, gr, and ge, each of which is assumed to be a normal, or Gaussian, probability density with mean zero and variance V_f, V_r, and v_e, respectively. Such a mixture is denoted by

$$gT = e_f\, gf + e_r\, gr + e_e\, ge. \qquad (5)$$

In general, the density gT will not be a Gaussian probability density and the interval from the negative square root of VT to the positive square root of VT may include much more than 68.3 percent of the errors. This will occur because, while e_e may be small, V_e may be extremely large. Actually, this standard error interval may include up to 99 percent of the errors.

To assist in interpreting the results of the analyses, a new parameter, equivalent Gaussian spread (EGS), is introduced. The parameter EGS specifies the range in terms of equal positive and negative logarithmic units from the mean that would encompass errors with the same a priority probability as would a Gaussian distribution with a standard deviation equal to EGS; in other words, the range from -1 EGS to +1 EGS contains about two-thirds of the errors. For Gaussian distributions of logarithmic errors, EGS and standard error are equivalent. EGS is reported herein in units of percentage and an approximate interpretation of EGS is "two-thirds of the errors in instantaneous streamflow data will be within plus or minus EGS percent of the reported value." Note that the value of EGS always is less than or equal to the square root of VT and ordinarily is closer to V_f, the measure of uncertainty applicable during periods of no lost record, the greatest part of the time.

The cost part of the input to the K-CERA methodology consists of determining practical routes to visit the stations in the network, the costs of each route, the cost of a visit to each station, the fixed cost of each station, and the overhead cost associated with the stream-gaging program.

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

All these costs, routes, constraints, and uncertainty functions are then used in an iterative search program, called the traveling hydrographer program (figs. 4 and 5), to determine the number of times that each route is used during a year such that (1) the budget for the network is not exceeded, (2) at least the minimum number of visits to each station are made, and (3) the total uncertainty in the network is minimized. This allocation of the predefined budget among the stream gages is taken to be the optimal solution to the problem of cost-effective resource allocation. Because of the large dimensionality and non-linearity of the problem, the optimal solution may really be "near optimal." (See Moss and Gilroy, 1980, or Fontaine and others, 1984.)

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

$V \equiv$ 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; and

$N_i \equiv$ 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 4.--Mathematical programing form of the optimization of the routing of hydrographers.