REGIONALIZATION OF HARMONIC-MEAN STREAMFLOWS IN KENTUCKY

By Gary R. Martin and Kevin J. Ruhl

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 92-4173

Prepared in cooperation with the

KENTUCKY NATURAL RESOURCES AND ENVIRONMENTAL PROTECTION CABINET

Louisville, Kentucky

1993

(U.S. Geological Survey, 1958a, 1958b, 1964a, 1964b, 1962-65, 1966-75, and 1976-90). Many of the stations in Kentucky are affected by regulation and (or) local diversions. Regulation by multipurpose or flood-control reservoirs reduces peak flows and generally augments low flows. Local diversions--localized transfers of water such as water supply withdrawals or wastewater discharges--artificially increase or decrease streamflows within a reach. Local diversions are common near municipalities and in urbanized areas. The extent of alterations in natural streamflows caused by regulation and local diversions is variable and difficult to quantify accurately. Periods of streamflow record affected by regulation and local diversions were considered separately from periods unaffected by regulation. Therefore, each continuous-record gaging station was screened to identify any periods of record affected by regulation (Melcher and Ruhl, 1984; Ruhl and Martin, 1991).

The retrievals included a total of 230 continuous-record streamflow-gaging stations (pl. 1), of which 54 also have regulated periods of record. Included with the stations located in Kentucky were several unregulated stations located nearby in adjacent States. The streamflow data in bordering States were retrieved to provide additional information for use in the regionalization of Q_h values. The period of record at the streamflow-gaging stations ranged from 1 to 78 years.

Unregulated Streamflow-Gaging Stations

Daily mean streamflow at unregulated stations and for unregulated periods at subsequently regulated stations were used to compute values of Q_h . To compare station values of Q_h, these values were standardized to basin drainage area (reported in units of $(ft^3/s)/mi^2$) at each unregulated station. Where nearby hydrologically similar index stations were available and suitable relations were obtained, the MOVE.1 record extension techniques were applied.

Record-extension adjustments to Q_h values were attempted at all unregulated stations with less than 10 years of record. The accuracy of results depend on the availability of a well-correlated index station in a hydrologically similar setting. Adjusted Q_h values that were not consistent with other nearby long-term stations were not used. The correlation coefficient (r) for the concurrent flows, though not used in the MOVE.1 calculation, is a measure of the strength of the linear relation; and r exceeded 0.80 for each station where the record extensions were used. Extensions were used at all the unregulated stations not affected by local diversion that had less than 6 years of record.

The extensions, while reducing time-sampling error at the station, generally improved consistency in Q_h and drainage-area-standardized Q_h values along stream reaches and among neighboring streams. A generalized depiction of the values of Q_h standardized to drainage area for unregulated stations not affected by local diversions is shown in figure 4. The largest values of drainage-area-standardized Q_h occur in karst areas (fig. 2) and also near Kentucky's eastern border with Virginia and West Virginia (an area roughly straddling the boundary between the Appalachian Plateaus and Valley and Ridge physiographic provinces). Sustained base flows are characteristic of streams in these areas (Ruhl and Martin, 1991; Hayes, 1991; and Friel and others, 1988) that have relatively high drainage-area-standardized Q_h .

Values of Q_h , drainage-area-standardized Q_h , total drainage area, and periods of record for stations not affected by regulation are listed in table 1 (back of report). The stations affected by local diversions and stations where the MOVE.1 record extensions were applied are identified by footnote in table 1.

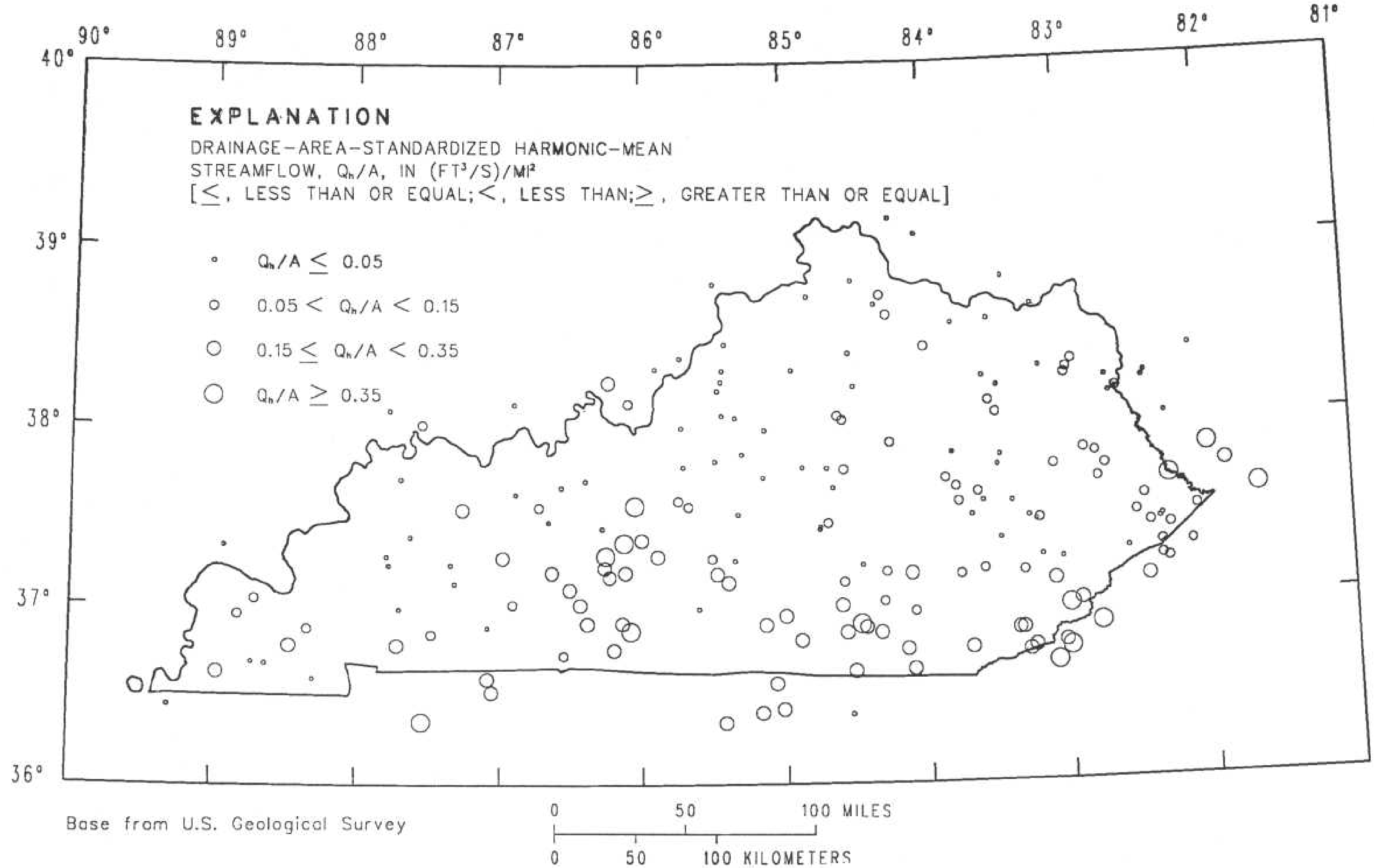

Figure 4.--Generalized drainage-area-standardized harmonic-mean streamflow at selected unregulated continuous-record streamflow-gaging stations in the study area.

At stations on the main stem of the Cumberland River, downstream of Harlan, Kentucky (station 03401000), the entire period of record was used for computation of Q_h. The effect of Martins Fork Lake, constructed in the basin headwaters in 1978, on the long-term Q_h at these downstream locations was considered to be insignificant. Streamflow at several of these stations is affected to varying degrees by local diversions.

Regulated Streamflow-Gaging Stations

Most major drainage basins in Kentucky are affected to some degree by regulation from reservoirs (lakes). All of the major reservoirs in Kentucky, except Herrington Lake on the Dix River and Kentucky Lake on the Tennessee River, are operated by the U.S. Army Corps of Engineers (COE). Herrington Lake is operated by Kentucky Utilities Company, and Kentucky Lake is operated by the Tennessee Valley Authority. The major reservoirs affecting streamflows in Kentucky are shown in figure 5. Continuous-record streamflow-gaging stations affected by these reservoirs are also shown in figure 5.

Values of Q_h were determined for both the regulated and, if applicable, unregulated periods at streamflow-gaging stations located downstream of these reservoirs. The Q_h values for periods of natural flow prior to regulation presented previously (table 1) are useful for regionalization of Q_h, but they do not reflect current streamflow conditions downstream of the reservoirs. Therefore, the values of Q_h have also been determined for the regulated record at each of the streamflow-gaging stations affected by a major reservoir. The Q_h values obtained for the regulated record and the period of regulation analyzed are given in table 2 at the back of the report.

The periods of regulated record analyzed were defined so as to reflect current operating conditions. A survey was made of each of the COE Districts operating reservoirs in Kentucky to determine the date when reservoir filling and normal release patterns were initiated. The end of the unregulated period coincides with the date that reservoir filling was initiated. The beginning of the regulated period, as used in this report, corresponds to that time when normal reservoir release patterns were initiated. Thus, the period of time when the reservoir was being filled has been excluded in computation of the Q_h.

Alterations of operating policy of the hydraulic structures at large reservoirs can significantly change streamflow characteristics downstream. Each of the COE Districts was surveyed regarding changes in operating policy and the consistency of low-flow releases from the reservoirs, because values of Q_h are likely to be sensitive to changes in low-flow release practices. Each reservoir has a target minimum release flow that may be different in summer and winter. Occasionally, the target release cannot be achieved because outlet gates must be closed either when repairs are needed or when reservoir inflow is not adequate to maintain the desired pool elevation. To analyze the streamflow-gaging stations where streamflow is affected by one or more major reservoirs, the annual 1-day, 3-day, and 7-day low-flow time series for the regulated periods of record were plotted. These annual low-flow values are the lowest mean flow for the specified number of consecutive days in each climatic year (April 1 through March 31). These lowest mean annual flows for station 03311000, Barren River near Finney, Kentucky, are shown in figure 6. This station is located just downstream from the outlet works at Barren River Lake. As indicated by the upper plot, the minimum 1-day releases were generally greater than or equal to 20 ft^3/s, except for climatic years 1977 and 1986. The minimum target release at this site is 20 ft^3/s (Paul Roberson, U.S. Army Corps of Engineers, oral commun., 1992). This was achieved during most years, and the low-flow release tends not to vary much over longer periods as indicated by the 3-day and 7-day low-flow plots (fig. 6). A review of the daily mean streamflow data (U.S. Geological Survey, 1978) indicated that the gates at Barren River Lake were closed from October through mid-December 1976 because of repairs on the outlet works. The only flow was leakage through the gates, and the flow ranged from 0 to 4.9 ft^3/s. During the period August 12-15, 1985, the flows ranged from

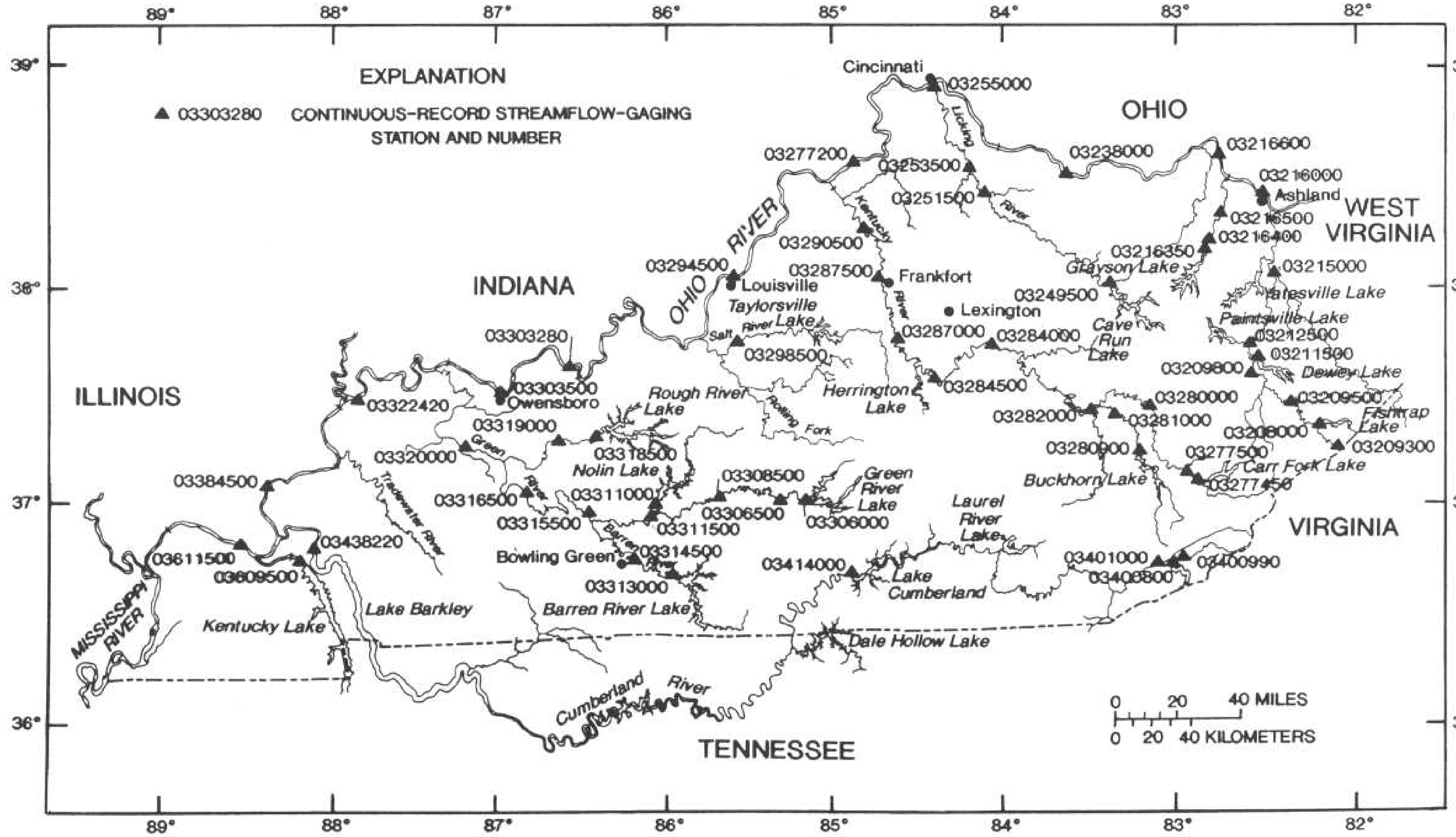

Figure 5.--Location of major reservoirs and selected regulated continuous-record streamflow-gaging stations in the study area [Adapted from U.S. Army Corps of Engineers, 1981].

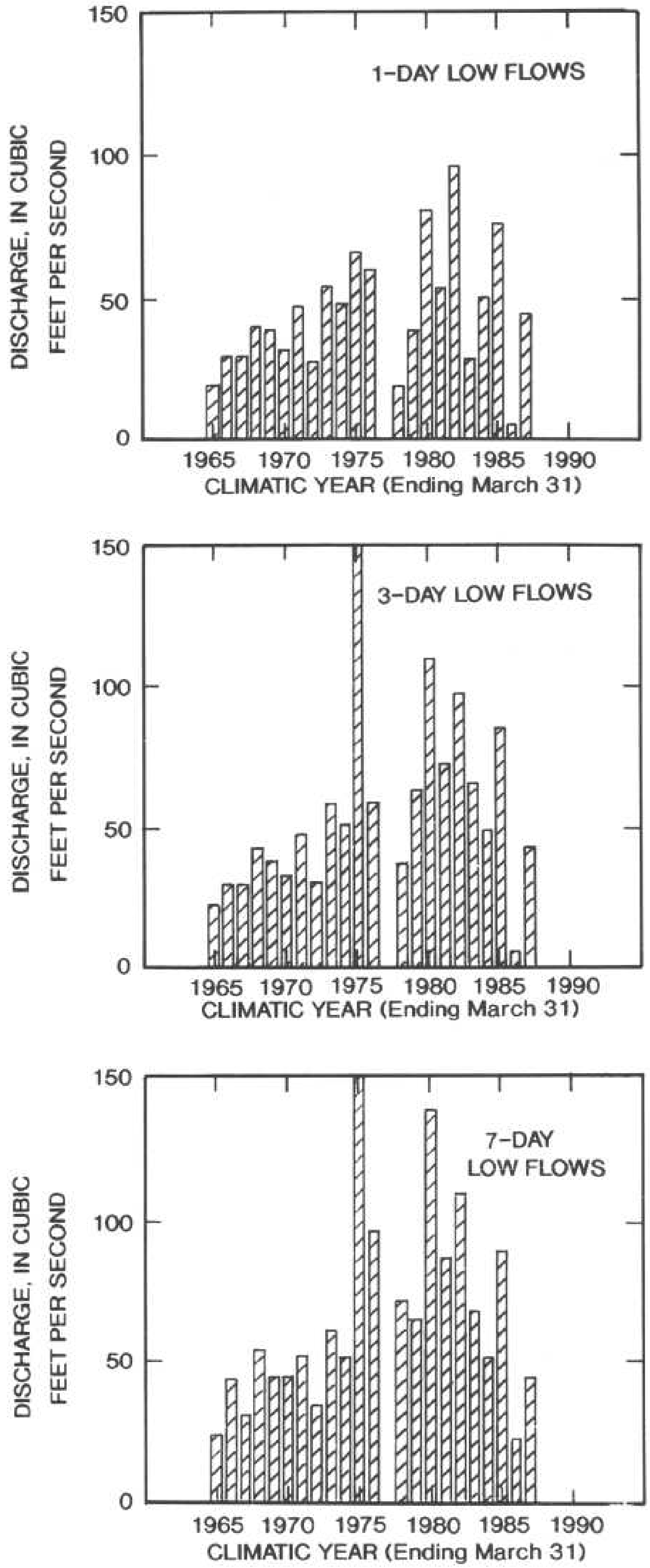

Figure 6.--The annual 1-day, 3-day, and 7-day lowest mean flows for the period of record affected by regulation at Barren River near Finney, Kentucky.

4.9 to 6.6 ft^3/s during minor maintenance activities. An upward trend in the releases occurs in the 1970's and early 1980's. This upward trend is climatic in nature, and it occurs in the flows at unregulated sites as well (Ruhl and Martin, 1991).

Zero- or near-zero-flow periods have occurred at almost all stations located directly below major reserviors, indicating gate closure. These gate closures were generally because of repair and (or) maintenance activities at the reservoirs. Some other low-release periods were the result of maintaining target pool elevations. For the stations displaying these random restrictions in reservoir releases, which are likely to continue to occur in the future, the entire period of regulated record was used in computing Q_h.

Only one station indicated a change in the release pattern over time. The plots of annual low flows for Cumberland River near Rowena (station 03414000), located just downstream of the outlet works at Lake Cumberland, indicated increased minimum releases starting around 1984. COE officials confirmed that the current power-operating policy was implemented in July, 1984. Because only 5 years of record (1984-89) are available for the current operating policy, a determination of Q_h was made using the entire regulated period 1951-89 (table 2).

The effects of regulation are most pronounced immediately below reservoir control structures and are gradually dampened with increasing distance downstream. At streamflow-gaging stations located downstream of more than one reservoir that were used in the study, it was unnecessary to delineate values of Q_h for different periods of regulation because the reservoirs most influencing streamflows were constructed first. Only a few of the stations are downstream of multiple reservoirs. Flows at the main-stem Kentucky River stations at Lock 14 (03282000) and Lock 10 (03284000) are regulated by both Buckhorn Lake, completed in 1960, and are affected to a lesser degree by Carr Fork Lake, completed in 1976 (fig. 5). Therefore, the entire period of record after 1960 was used to compute Q_h. Farther downstream, flows at the main-stem Kentucky River stations at Lock 6 (03287000), Lock 4 (03287500), and Lock 2 (03290500) are regulated by Herrington Lake, completed in 1925 (fig. 5). Buckhorn Lake and Carr Fork Lake have little effect on flows at these stations. Therefore, the entire period of record starting in 1925 was used for these three stations. For the main-stem Green River stations at Lock 6 (03311500), Lock 4 (03315500), and Lock 2 (03320000), the earliest date of regulation was used as the starting date in the determination of the Q_h. Streamflow at Lock 6 is regulated by Nolin Lake starting in 1963 and by Green River Lake starting in 1968. The addition of the Green River Lake outflows probably have had a minimal effect on low flows at Lock 6 (fig. 5). The same is true farther downstream at the Lock 4 station, which receives flows from Barren River Lake, Nolin Lake, and Green River Lake. Barren River Lake came into operation in 1964, just after Nolin Lake was completed. The next downstream station, at Lock 2, is also regulated by Rough River Lake, which was completed in 1959. For each station, the reservoir that would probably have the most effect on flows, especially low flows, was also the first to be completed. For this reason, the reservoir with the earliest completion date was used to determine the starting date of the period of regulated record at each station downstream of multiple reservoirs.

The Ohio River main-stem stations (fig. 5) were treated somewhat differently. Because reservoirs having different start dates and degrees of regulation are located throughout the Ohio River basin and the locks and dams in the main stem affect the flow in the Ohio River, the entire period of record was used for determining the Q_h. Two stations on the Ohio River have long-term, continuous streamflow record: Ohio River at Louisville, Kentucky (03294500), and Ohio River at Metropolis, Illinois (03611500). The stations at Louisville and Metropolis have daily flow record available from April 1928 through September 1989. Other stations on the Ohio River have been operated periodically, but for shorter periods of time. The Q_h was determined for the period of record at each of the Ohio River main-stem stations. Because of climatic variations in the different periods of record and different degrees of regulation, Q_h values varied considerably between stations. In an effort to make these values more consistent with the long-term record available at the Louisville and Metropolis gages (index stations), the MOVE.1 record-extension procedure was used to estimate long-term Q_h. The estimated long-term Q_h values are given in table 2. Streamflow record at the Ohio River at Louisville index station was used to provide an adjusted Q_h at

Ohio River at Ashland, Ky. (03216000),
Ohio River at Greenup Dam, Ky. (03216600),
Ohio River at Maysville, Ky. (03238000),
Ohio River at Cincinnati, Ohio (03255000),
Ohio River at Markland Dam, Ky. (03277200),
Ohio River at Cannelton Dam, Ky. (03303280),
Ohio River at Owensboro, Ky. (03303500), and
Ohio River at Uniontown, Ky. (03322420).

Streamflow record at the Ohio River at Metropolis, Illinois index station was used to provide an adjusted Q_h at Ohio River at Lock and Dam 51 at Golconda, Illinois (03384500).

Table 3 shows the Q_h for various periods, including the periods of record, at the Ohio River at Louisville and Metropolis gages. The periods chosen approximate the periods of record for the other main-stem stations (table 2). The records for stations at Ashland, Maysville, Cincinnati, and Owensboro were collected mainly in the period 1939-64. As shown in table 3, the Q_h for the period 1939-64 at Louisville is approximately equal to the Q_h for the entire period of record. Similarly, the Q_h at Metropolis for the period 1940-52 (which corresponds to the period of record at Golconda) is approximately equal to the Q_h for the entire period of record at Metropolis. However, later periods at the two index stations show a marked increase in Q_h values. The period 1970-89 corresponds closely to periods of record at Greenup, Markland, and Cannelton Dams. The period 1975-89 corresponds to the period of record at the Uniontown station. The increased values of Q_h observed at the index stations during these later periods is probably climatic in nature.

Table 3.--Harmonic-mean streamflows for selected periods for Ohio River at Louisville, Kentucky and Ohio River at Metropolis, Illinois

[ft^3/s, cubic feet per second]

Station	Start date (month/year)	End date (month/year)	Harmonic-mean flow (ft^3/s)
Ohio River at Louisville, Ky.	04/1928	09/1989	39,700
(03294500)	10/1939	05/1964	38,100
	05/1970	09/1989	54,800
Ohio River at Metropolis, Ill.	04/1928	09/1989	135,000
(03611500)	10/1940	09/1952	134,100
	10/1975	09/1989	164,500

Annual precipitation records for this period show a similar increase (Conner, 1982; Conner, written commun., 1991). Record extensions at the short-term stations helped to minimize this time-sampling error caused by such climatic variation and resulted in a consistent downstream increase in the values of Q_h (table 2).

Flows at Cumberland River at Grand Rivers (03438220) and Tennessee River at Paducah (03609500) are affected by releases from Lake Barkley and Kentucky Lake, respectively (fig. 5). Since June 1966, the flows at these stations have also been affected by interbasin transfer. A canal connecting the two lakes was constructed to facilitate barge traffic. Values of Q_h (table 2) for the regulated period at these two stations were computed using streamflow record after June 1966 only.

DEVELOPMENT OF PROCEDURE FOR ESTIMATING HARMONIC-MEAN STREAMFLOW AT UNGAGED STREAMFLOW SITES

Drainage-basin characteristics, including climate, influence streamflow patterns. Relations among selected basin characteristics and computed Q_h were investigated by methods of linear correlation and multiple-linear regression. A regional relation to estimate Q_h at ungaged sites was defined by regression analysis.

Selected Drainage-Basin Characteristics

Several drainage-basin characteristics were tested for applicability in the regionalization of Q_h. Selection of basin characteristics for inclusion in exploratory scatter plots, linear correlation analysis, and subsequent multiple-linear-regression analysis was based on (1) the possible hydrologic significance of the characteristic in relation to the Q_h statistic, (2) the availability of previously determined basin characteristics for the study basins, and (3) results of previous regionalization studies of other streamflow statistics (Beaber, 1970; Wetzel and Bettandorff, 1986; Choquette, 1988; Ruhl and Martin, 1991).

Basin characteristics obtained from the basin and streamflow characteristics file of the National Water Data Storage and Retrieval System (Dempster, U.S. Geological Survey, written commun., 1983) and tested for significance in the regression analysis included the following:

1. Total drainage area, in square miles, is the area measured in a horizontal plane that is enclosed by a drainage divide.

2. Contributing drainage area, in square miles, is the total drainage area excluding any parts characterized by internal drainage.

3. Main-channel length, in miles, is the length measured along the main stream channel from the gage to the basin divide, following the longest tributary.

4. Main-channel slope, in feet per mile, is computed as the difference in elevation between points located at 10 and 85 percent of the main-channel length from the gage, divided by the stream length between these two points.

5. Basin length, in miles, is the straight-line distance from the gage to the basin divide (defined by the main-channel length).

6. Mean basin width, in miles, is calculated by dividing the total drainage area by basin length.

7. Main-channel sinuosity is the ratio of main-channel length, in miles, to basin length, in miles.

8. Mean basin elevation, in feet, is measured as the average elevation of 20 to 80 points per basin using the transparent grid sampling method.

9. Mean annual precipitation, in inches, is estimated from Conner (1982).

10. Soils index, in inches ("S"; U.S. Department of Agriculture, 1969), is a measure of potential infiltration based on basin vegetative cover, soil infiltration rate, and soil water storage.

11. Soil infiltration index, in inches per hour, is based on minimum infiltration rates for the U.S. Soil Conservation Service hydrologic soil groups (Musgrave, 1955) for soil series in Kentucky (U.S. Department of Agriculture, 1975 and 1984).

12. Forested area, as a percentage of contributing drainage area, is measured from topographic maps using the transparent grid sampling method.

13. Streamflow-recession index at a station is defined as the number of days it takes base streamflow to decrease one log cycle, or one order of magnitude, as determined graphically from hydrograph plots of daily mean streamflow during representative periods of streamflow recession (Riggs, 1964; Bingham, 1982, Ruhl and Martin, 1991).

14. Streamflow-variability index (Lane and Lei, 1950) at a station ("station" value) is computed as the standard deviation of the logarithms of the 19 discharges at 5-percent class intervals from 5 to 95 percent on the flow-duration (cumulative-frequency) curve (Searcy, 1959; Dempster, 1990) of daily mean streamflow for the entire period of record. Like the streamflow-recession index, this streamflow index is a measure of basin capacity to sustain base flow in a stream. In Kentucky, streamflow-variability indexes have been mapped by delineating areas of similar station streamflow-variability index and similar geologic features (Ruhl and Martin, 1991). The "map" values of streamflow-variability index for stations in Kentucky were computed as an area-weighted mean of the basin streamflow-variability indexes for use in the regression analysis. For the stations located outside Kentucky, station values of streamflow-variability index were used. The values of streamflow-variability index for the stations used in the regression are listed in table 1.

Regression Analysis

A multiple-linear-regression model was developed to relate Q_h (dependent variable) to selected basin characteristics ("independent" or explanatory variables). Only stations with streamflow data collected during unregulated periods with streamflows not significantly modified by local diversions were included in the analysis (table 1). The station located at Cumberland River at Williamsburg, Kentucky (03404000), only minimally affected by regulation at Martins Fork Lake in the basin headwaters, was also included in the regression.

Inspection of scatter plots showing relations among dependent and explanatory variables and plots of residuals from initial linear regressions indicated that logarithmic (base 10) transformation of the dependent and most of the explanatory variables would be appropriate. This transformation generally helped make the relations more linear and the residuals more uniform in variance about the regression line than before transformation. The relations between dependent and explanatory variables after transformation were consistent with the assumed linear form of the model.

Several factors were considered in evaluating alternative regression models including (1) the coefficient of determination, the proportion of the variation in the dependent variable explained by the regression equation, (2) the standard error of the estimate, a measure of model-fitting error, (3) the PRESS statistic, a measure of model-prediction error, (4) the statistical significance of each alternative explanatory variable, (5) potential multicollinearity as indicated by correlation of explanatory variables and the value of the variance inflation factor (Montgomery and Peck, 1982), and (6) the effort and modeling benefit of determining the values of each additional explanatory variable.

Ordinary-least-squares regression techniques were used to fit the linear model. The alternative models were generated by all-possible-regression and stepwise-regression procedures (Statistical Analysis System Institute, 1985) using the prospective explanatory variables listed previously.

Weighted-least-squares (WLS) regression procedures could be used to compensate for differences in the reliability of estimates of Q_h (because of time-sampling error) based on station record length. However, WLS was not applied because of (1) the relatively large number of daily mean streamflow observations used to compute Q_h , and (2) record-extension procedures were applied to short-term stations to provide improved estimates of long-term values of Q_h . Of the stations included in the regression analysis, 6 years was the shortest station record length that was not extended using the MOVE.1 procedure.

Review of residual plots from regressions that included mapped values of streamflow-variability index (Ruhl and Martin, 1991) as an explanatory variable revealed a group of outliers located in the south central portion of the Mississippi Embayment in southern Calloway and Graves counties. The models significantly overpredicted harmonic-mean streamflow at the stations on Clarks River at Murray, Kentucky (0361000), Perry Creek near Mayfield, Kentucky (07022500), and Obion Creek at Pryorsburg, Kentucky (07023500). Geologic maps and information relating to ground-water/surface-water interactions in the area were reviewed for indications that the mapped values of variability index assigned for this area might be too low. Ground-water levels (Davis and others, 1983) and modeling simulations (Grubb and Arthur, 1991) indicate that significant net aquifer recharge is occurring in drainage basins in this area, particularly in Graves County. Also, a thin, shallow aquifer underlain by an aquitard (Porters Creek Clay) is present in the upper Clarks River basin (Murray, Kentucky vicinity), providing minimal ground-water discharge to streams. Perennial flow of streams in the Mississippi Embayment occurs only downstream of the intersection of stream channels with the water table. This intersection and resulting perennial streamflow occurs downstream of these three stations (Davis and others, 1983). Based on this information and the station values of streamflow-variability index computed for these three stations (1.34, 1.08, and 1.50), map variability-index boundaries from Ruhl and Martin (1991) were redefined, and a region of mapped streamflow-variability index of 1.25 was delineated (pl. 1).

The best group of two-variable models (in terms of predictive accuracy) included (1) total drainage area, or alternatively, drainage-area-related variables such as main-channel length, contributing drainage area, mean basin width, and basin length and (2) streamflow-variability index as the explanatory variables. Following this top group of models was a set of two-variable models that included streamflow-recession index coupled with various drainage-area-related basin characteristics. Thus, streamflow-variability index and streamflow-recession index, characteristics relating to basin hydrogeology and low-flow regime, are apparently closely associated with the Q_h, as they are associated with other low-flow statistics such as $7Q_{10}$ (Ruhl and Martin, 1991).

A two-variable model containing total drainage area and streamflow-variability index as explanatory variables was judged superior to alternative models including more than two variables based on comparisons of PRESS values, multicollinearity diagnostics, tests of significance of additional explanatory variables, and the extent of improvement in model predictive ability. A sensitivity analysis of the model indicated that sensitivity to the streamflow-variability index, which can vary significantly over short distances (pl. 1), would generally be reduced by including variability index without log transformation. Though model error is increased somewhat without the log transformation of the streamflow-variability index, potential model-application errors may be reduced.

The regression model selected is

$$Q_h = 1.65A^{1.02}10^{-1.85V}, \tag{4}$$

where

Q_h is the estimated harmonic-mean streamflow, in ft^3/s;
A is the total drainage area, in mi^2; and
V is the mapped streamflow-variability index on plate 1.

Noting the sign of the exponents, the estimated Q_h increases with increasing drainage area and decreasing streamflow-variability index. The estimate is a near-linear relation to drainage area because the exponent is approximately one. The estimate is more sensitive to a percentage change in the value of the streamflow-variability index than to a like change in total drainage area. The drainage area and the value of streamflow-variability index for each unregulated streamflow-gaging station not affected by local diversion that was used in the regression are listed in table 1.

Equation 4 can be solved graphically using the nomograph shown in figure 7. Example calculations are presented in the section "Estimating Harmonic-Mean Streamflow at Stream Sites in Kentucky."

Limitations and Accuracy

The regional regression model for estimating Q_h at ungaged sites is applicable to unregulated streams in Kentucky that are not significantly affected by local diversions. Caution is warranted when applying the regression model in areas where streamflows are affected by hydrologic discontinuities such as large springs and sinks common to karstic terrain in areas underlain by limestone. Streamflows in these areas may vary unpredictably over short reaches.

The regression model was developed using basin characteristics within a certain range of values. Drainage areas of stations used in the regression analysis ranged from 3.89 to 1,607 mi^2. Values of streamflow-variability index ranged from 0.40 to 1.35. Application of the regression model for a basin outside these ranges is an extrapolation and is, therefore, not recommended. Estimates of Q_h for stream sites with basin characteristics not in these ranges should be based on streamflow data collected at the site. A scatter plot showing the sampling space for drainage areas and variability indexes for the regression model is shown in figure 8. Note the absence of observations having a combination of low streamflow-variability index and low drainage area (V less than 0.50 and A less than about 65 mi^2). Application of the model in this region falls outside the sampling space, and this extrapolation is, likewise, not recommended.

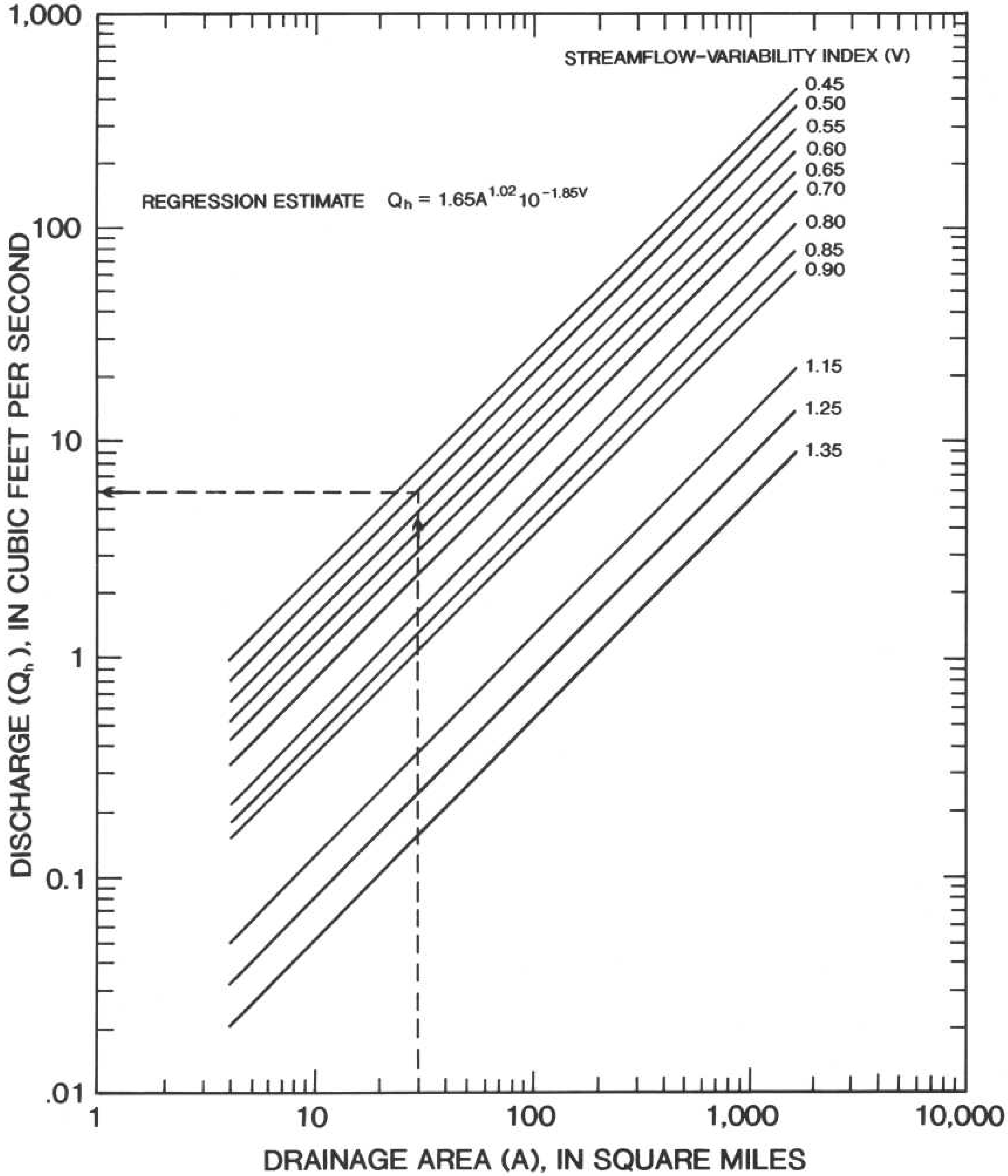

Figure 7.--Graphical solution of the regression equation for estimating harmonic-mean streamflow in Kentucky.

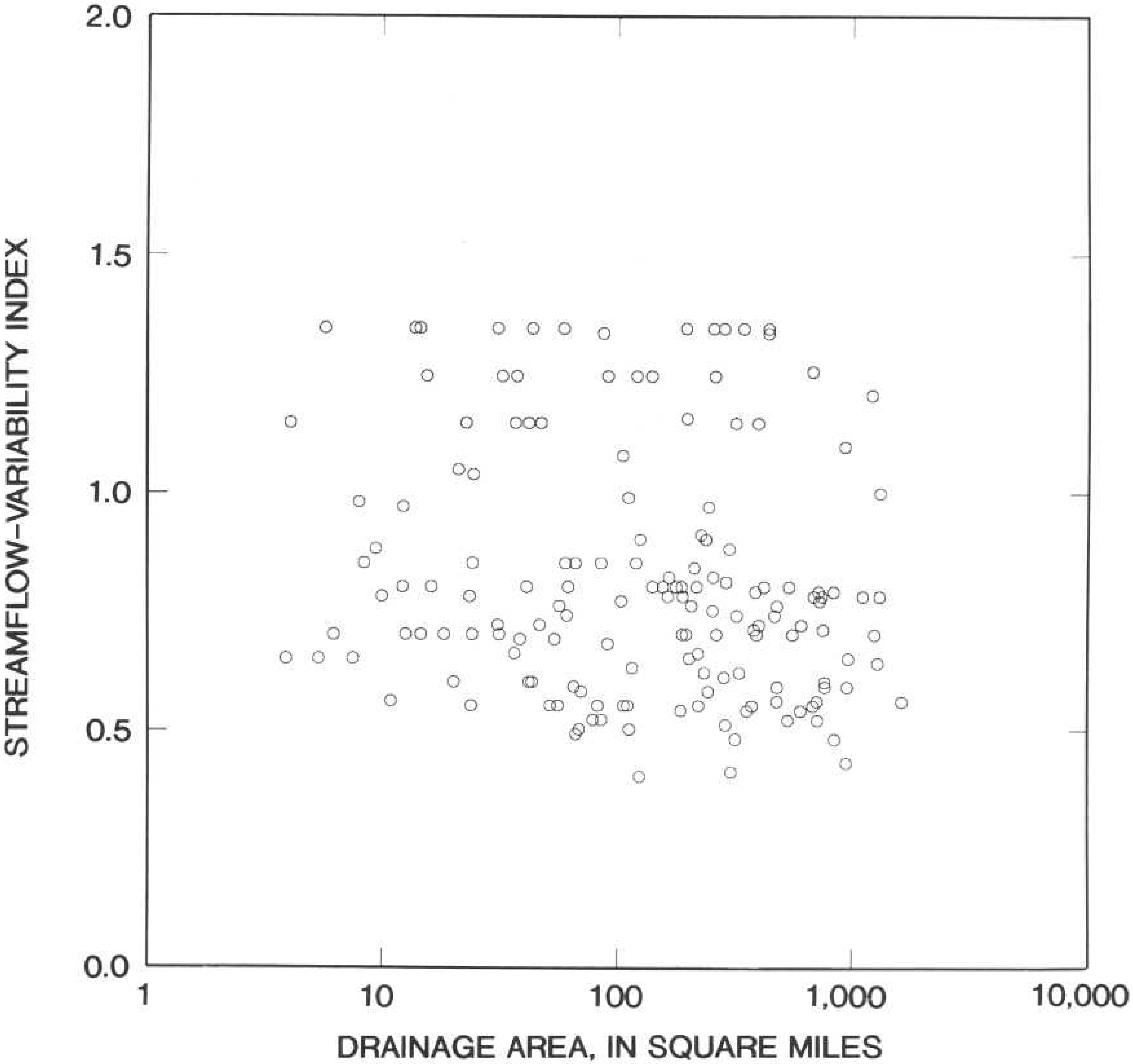

Figure 8.--Drainage area and streamflow-variability-index sampling space for the regression model.

The coefficient of multiple determination (R^2) for the regression model is 0.90. The standard error of estimate (of log Q_h)--a measure of the accuracy of the regression-model estimates compared to the observed values used in the regression--is 76 percent. The standard error of estimate was computed using the model root-mean-square error (Statistical Analysis System Institute, 1982) and information from Hardison (1971). The standard error of prediction (of log Q_h)--a measure of the accuracy of the regression estimates compared to observed data for stations excluded from the regression--is 78 percent, which is slightly higher than the standard error of estimate. Standard error of prediction was estimated as the square root of the PRESS divided by the error degrees of freedom (Statistical Analysis System Institute, 1982; Montgomery and Peck, 1982; Choquette, 1988). The procedure used for computing PRESS is considered a form of data splitting and can be applied as a model-validation tool. The accuracy of the model predictions for ungaged sites similar to those used in the regression could be expected to compare favorably to the standard error of prediction. If all the assumptions for applying regression are met, two-thirds of the observations lie within one standard error of a regression line. For this regression, a 0.293 log units standard error, when untransformed, would place two-thirds of the observations within plus 96 percent and minus 49 percent of the regression line.

A scatter plot of the values of Q_h computed from the streamflow-gaging station data and values computed using the regression model (fig. 9) shows reduced residuals and a slight tendency of the model to underpredict the values of Q_h above about 50 ft^3/s. The underprediction tendency may be associated with increased error and bias in the values of mapped streamflow-variability index for large basins. The reduced residuals are probably related to generally reduced time-sampling error (long periods of record) for the stations having large values of Q_h. Also, less variability in the streamflow response would be expected for large basins as compared to small basins.

ESTIMATING HARMONIC-MEAN STREAMFLOW AT STREAM SITES IN KENTUCKY

Procedures for obtaining Q_h estimates differ depending on the location of the stream site in relation to stream gage locations where Q_h has been determined. The appropriate procedures and examples are presented in the following sections.

Stream Sites With Gage Information

When streamflow-gaging information is available on the reach where an estimate of Q_h is desired, the gage information is used where appropriate in making the estimate, as discussed below.

Sites at Gage Locations

Estimates of Q_h values for 230 continuous-record streamflow-gaging stations are presented in tables 1 and 2. When an estimate of Q_h is required at a stream site, refer to table 1 (if the site is unregulated), or to table 2 (if the site is regulated), to determine whether values have previously been estimated for the site.

Sites Near Gage Locations

If information is available for an unregulated stream where an estimate is desired, but not at the specific location, a weighting procedure can be employed (Carpenter, 1983). The first constraint to the use of this method is that the drainage area of the ungaged site differ by no more than 50 percent from that of the gaged