UNC-WRRI-82-180

A RAPID, LOW-COST TECHNIQUE FOR ESTIMATING PEAK FLOW FOR SELECTED FLOOD EVENTS

by

Nelson R. Nunnally
Department of Geography and Earth Sciences
University of North Carolina at Charlotte
Charlotte, North Carolina 28223

The work upon which this publication is based was supported in part by funds provided by the Office of Water Research and Technology, U.S. Department of the Interior, Washington, D.C., through the Water Resources Research Institute of The University of North Carolina as authorized by the Water Research and Development Act of 1978.

Project No. A-119-NC

Agreement No. 14-34-0001-1135

ACKNOWLEDGMENTS

This report would not have been possible without the effort of Patrick Hallett, who served as Research Associate for the project. Patrick spent hundreds of hours assembling data sources, measuring variables, and performing statistical analyses. Though much of this effort produced no direct, fruitful results, they were necessary tasks, and the data generated will be a useful source for further research and for teaching.

Thanks are due also to the staff of the Computing Center at UNCC who spent many hours reading tapes and writing special programs, and to Becky Guy who typed the manuscript.

DISCLAIMER STATEMENT

ABSTRACT

Water resources data are often unavailable for small watersheds in the Southeast, yet accurate estimates of the magnitude of discharge events are needed for engineering design and floodplain delineation. A regionalized streamflow estimation technique was developed for North Carolina and Virginia for estimating the two, ten and twenty-five year return interval peak flows. Separate equations were generated for the Mountain, Piedmont, Coastal Plain regions and for urbanized watersheds. Comparison with previous techniques used in North Carolina indicated that the new equations produce more accurate peak discharge estimates in most instances.

LIST OF TABLES

LIST OF FIGURES

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

The need for improved estimates of peak discharges for small watersheds is well established. The problem is most acute in basins affected by urban development where discharge data are needed for designing stormwater drainage, detention ponds, channel improvements and other engineering applications.

The potential users of peak discharge data seldom have the expertise, time, or money to employ sophisticated modes. Regression equations provide an acceptable compromise between the more accurate, but harder to use, process-response models and the simpler, but very inaccurate deterministic equations that are still widely used.

Several investigators have developed regression equations for predicting small basin stream discharges in North Carolina (Hinson, 1965; Jackson, 1976; and, Putnam, 1972). Jackson's and Hinson's equations are applicable, essentially, only to basins that have not been subjected to radical modifications such as channelization, regulation or urbanization. Putnam's study, though confined to urbanizing or urbanized basins, produced equations that appear to give unsatisfactory results when tested against urban basins with gaging records of 10 years or longer (the average error exceeded 50 percent).

In this study we sought to produce improved regression equations by confining the study to small watersheds, expanding the sample size by using stream basins outside North Carolina, using longer data records (due to the time elapsed since Jackson's study), incorporating more variables which influence peak discharge, and controling for some of the variance in the data set by creating more homogeneous regions through the use of multivariate techniques.

Regionalization schemes employing basin attributes produced no better equations than grouping by physiographic provinces. Grouping of basins using discharge characteristics yielded equations with very low standard errors of estimates, but the basins in each group possessed no distinct attributes which could be identified with the groups so that ungaged basins could be assigned to groups.

The final equations for two, ten and twenty-five year returns were based on four groups--one for each of the three physiographic provinces, mountain, piedmont, and coastal plain, and one for basins which were substantially urbanized. These equations employ only size and a lag coefficient as dependent variables. None of the other climatic, land use, or geomorphic variables proved to be useful in predicting discharge, probably because they were controlled to some extent by the regional groupings. Prediction errors were calculated for each basin for each return interval. These were compared to the prediction errors using U.S.G.S. equations (Jackson's equations for non-urban basins and Putnam's equation for urban basins).

Equations developed for urban basins were the most accurate, overall, and the results (Table III) were clearly superior to results using Putnam's equation. Our equations produced an average percentage error less than thirty percent for all recurrence intervals compared to an average error ranging from 56 percent to over 100 percent for the Putnam equation. In addition, there were fewer large errors of overprediction.

For rural basins the results were mixed. Estimates of peak discharges were improved, in most cases, over those using Jackson's equations. However, the U.S.G.S. equations gave slightly better results for the twenty-five year return interval flows for the mountain regions and for the ten and twenty-five year discharges in the piedmont. Even in these cases, however, there was a better balance between numbers of stations in each group that were over-predicted and under-predicted compared to the U.S.G.S. equations. This implies that the U.S.G.S. equations, which were developed from a much broader range of watershed sizes, may be somewhat biased when used on small watersheds.

The equations produced by this study seem to be logically consistent in that successively larger peak discharges are predicted for successively larger basins and successively higher return intervals. Furthermore, if return interval, basin size, and lag are held constant urban basins produce the largest discharges, followed by mountain, piedmont, and coastal plain. The lag effect is negative, except for the mountain basins. This is contradictory to the prevailing notion of lag effect. There is reason to believe, nevertheless, that the effect may be real and may be a product of the dramatic reduction in the time of concentration of overland flow in small, steep mountain basins. Further research using detailed process-response models is needed to clarify this issue.

It is recommended that individuals who need two, ten, or twenty-five year return interval discharge estimates for small, ungaged watersheds use the procedures presented in this report. The only measures required are drainage basin size, main stream length and average main stream slope between points 10 percent and eighty-five percent upslope from the point where the discharge estimate is required. Detailed procedures for calculating the lag variable and for using the equation are given on pages 15 and 16. Nomographs are provided (Figures 6, 7 and 8) for those who prefer a graphical solution.

TABLE OF CONTENTS

Introduction

Water resource data are often unavailable for small watersheds, yet accurate estimates of the magnitude of discharge events are needed for engineering design and floodplain delineation (Lichty and Liscum, 1978). The deficiency of streamflow data for small ungaged watersheds is particularly severe in the Southeast, where rapid development of rural areas and urban expansion demands adequate streamflow data for planning purposes. To compound the problem this development is occurring in a region characterized by small communities which have neither the computer hardware, qualified personnel nor funds to make use of sophisticated simulation methods. These potential users need a technique which accurately estimates specific discharge events for small, ungaged watersheds and which is simple and inexpensive to implement. The model should be flexible enough to work with equal accuracy in the range of physical conditions encountered from mountains to coastal plain and for urban and rural areas.

Streamflow simulation techniques vary considerably in data requirements, desired outputs, geographical focus, necessary expertise and cost. Models such as the U.S. Army Corps of Engineers Streamflow Synthesis (SSARR) and the Stanford Watershed Model Series reflect great technological sophistication in reproducing watershed processes (Fleming, 1975). However, the very sophistication which enables these models to simulate drainage basin responses to particular events is also their main weakness. The instrumentation, expertise, time and cost requirements inhibit the general usefulness to many local governments and private concerns who often have immediate need for specific, yet inexpensively obtained discharge data.

Other so-called "quick and dirty" techniques are often used but the large inherent errors make them unsuitable for many design purposes. The Rational

Formula, for example, is restricted to very small watersheds and despite its reputation as a conservative approach, it consistently underestimates peak discharges in some areas (Dunne and Leopold, 1978).

Another set of streamflow estimation techniques lies somewhere between the highly sophisticated deterministic models and the "quick and dirty" methods. This third group utilizes regression analysis to mathematically define the relationships between geomorphic, climatic and land use variables and discharge data to yield a statistically based estimate of a particular discharge event. Horton (1945) conceived of the feasibility of using geomorphic parameters to estimate discharge of ungaged watersheds and did much to identify important variables. Initially, basin area was found to be a relatively good surrogate for mean annual discharge (Hack, 1957), and it is relied upon as the sole parameter in some models (Jackson, 1976).

Regression has persisted as the dominant analytical tool since the first attempts to utilize the relation between streamflow and geomorphic and climatic parameters (Benson, 1959; Carlston, 1963; Wong, 1973; Thomas and Benson, 1970; de Coursey, 1972; Reynolds, 1972; Graf, 1975; Guevera, 1975; Ozga-Zielinski, 1975; Miller and Newson, 1975; Osborn, 1974; Osborn, 1975; Bevens and Kirby, 1979; Wong, 1979). The vast majority of these models have focused on refining and combining variables for efficiency, while few have contributed to the refinement of the regression technique itself. Wong (1963) and others (Haan, 1972; Shelton and Sewell, 1969; Wong, 1979) have used principal components analysis to reduce the numbers of variables while retaining most of the explanatory power of the set.

Error in regression analysis is due to several factors including sampling variance, failure to include all pertinent variables, measurement errors, and improper specification of the model form. As a general rule, the more variable the environments from which samples are drawn the larger the errors of prediction. But through the use of principal components analysis with grouping analysis (deCoursey, 1972; Gustafson, 1973) or mapping of regression residuals (Jackson, 1976; Miller and Newson, 1975; Newson, 1978) this variability can be significantly reduced (deCoursey, 1972).

Previous Work in North Carolina

Several flood flow frequency techniques have been developed for North Carolina in the past twenty years. Speer and Gamble (1964, 1965) developed procedures for estimating flood magnitudes for large streams (greater than 150 $mi.^2$). Their efforts were followed by Hinson (Hinson, 1965) who used the network of crest-gage stations established in 1952 to define flood frequencies for smaller basins (less than 150 $mi.^2$). Putnam (1972) examined the effect of urbanization on flood frequencies and magnitudes in the piedmont.

Though these studies encompassed a broad range of basin sizes and urbanized areas none of them produced very accurate results across the full range of physiographic and climatic variations representative of North Carolina streams. Jackson, however, found through mapping residuals from a general multiple regression equation that two hydrologic zones could be delineated in North Carolina. The piedmont and mountain regions were combined into one while the coastal plain formed the other.

The use of Jackson's regional model is restricted by several constraints. Streams must not be affected by regulation, tides, urbanization or channel improvement (channelization). But these last two factors disqualify many streams for which flood frequency data is most needed.

Considering these constraints coupled with the efforts of previous research (deCoursey, 1972; Gustafson, 1973) it was felt that an improved regional model could be developed. Virginia stations were included to increase sample size. Errors resulting from a large variance within the sample should be reduced by grouping basins based on geomorphic and hydrologic characteristics. Further reduction in error may be attained through increasing the number of predictor variables. Simultaneously, the need to provide data for small ungaged watersheds through the range of urban and rural areas could be fulfilled.

Data and Method

Historic records of streamflows at gaged streams provide the foundation on which statistical treatment of specific flows and predictor variables is based. The integrity of the magnitude-frequency relationship produced from historic records is partially dependent on the length of record at the gaging station. A minimum of ten years of record is recommended by the U.S. Water Resources Council to develop sufficiently accurate magnitude-frequency estimates using the log-Pearson Type III method (U.S. Water Resources Council, 1976). Extrapolating along frequency curves synthesized from short term records to predict extreme return interval events (greater than 50 years) can result in large errors. However, Moss (1979) has illustrated that extrapolating to medium return interval events (25 year) can be achieved if a sufficient number of stations is utilized. Although tasks such as floodplain mapping require one hundred year return interval data, there may be a greater need at the local level for more frequent return interval discharges which are necessary in designing storm drainage and stormwater detention facilities (Linsley and Franzini, 1979). Consequently, to supply useful recurrence interval discharges, but maximize the number of stations included in the study, while ensuring adequate estimates from the magnitude frequency curves, the two, ten and twenty-five year flood were chosen as the focus of this study.

Watershed Characteristics Related to Discharge

Numerous researchers have noted a relationship between streamflow and watershed characteristics (Horton, 1945; Benson, 1959; Carlston, 1963; Wong, 1963; Thomas and Benson, 1970; de Coursey, 1972; Reynolds, 1972; Graf, 1975; Guevera, 1975; Ozga-Zielinska, 1975; Miller and Newson, 1975; Osborn, 1974; Osborn, 1975; Bevens and Kirkby, 1979; Wong, 1979). Their efforts have served to isolate those variables which most influence the hydrology of a watershed. Initially, basin area was found to be a relatively good surrogate for mean annual discharge (Hack, 1957). Benson (1959), utilizing multiple regression, determined that area and valley slope combined produced significantly more effective estimates than area used alone. Trainer (1969) defined a close relationship between drainage density (the total stream's length divided by catchment area) and low flows along the Potomac River. Others have remarked on the importance of drainage density in characterizing streamflow (Horton, 1945; Schumm, 1956; Carlston, 1963; Patton and Baker, 1976; Gardiner, 1979). Many additional morphometric and land use variables have been inspected (Wong, 1963; Jackson, 1976; Thomas and Benson, 1970) but the final list of most effective variables remains small. These are enumerated in Table I.

The three land use variables (Table I) were measured from the Land use and Land Cover Series produced by the U.S. Geological Survey. This series, based on data from high altitude aerial photography and LANDSAT imagery, was found to present the best, most easily acquired source for current land use and land cover information.

Because many morphometric characteristics are somewhat affected by map scale (Gardiner and Parks, 1978) a consistent scale is required for data collection. Morisawa (1962) noted that the 1:24,000 scale U.S. Geological Survey topographic series served as well as aerial photographs to depict stream channels in Kentucky. Thus the 7.5 minute series was selected as the base from which all the morphometric variables were synthesized.

Table I

Watershed Characteristics

CHARACTERISTIC	MEASURE	SOURCE
1. Watershed size	Area	USGS 1:24,000 Topographic Map
2. Channel slope	10% - 85% slope along mainstream	USGS 1:24,000 Topographic Map
3. Lag time (k)	$K = \frac{\text{Mainstream Length}}{\sqrt{\text{channel slope}}}$	USGS 1:24,000 Topographic Map
4. Elongation ratio	$\frac{\text{Area of Basin}}{\text{Area of circle with same diameter as basin's long axis}}$	USGS 1:24,000 Topographic Map
5. Precipitation	2 yr., 24 hr. precipitation & 10 yr., 24 hour precipitation	U.S. Weather Service Technical Paper 40

Land Use and Land Cover[1]

1. Agricultural Land and Forest Cover
2. Urban Land
3. Water and Wetland Component (Storage)

---

[1] See Appendix A

When the U.S. Geological Survey water data system (NAWDEX) was accessed to determine the number of streams in Virginia and North Carolina with the designated size range[1] one hundred and forty-five watersheds were identified. However, the selection of the 1:24,000 series topographic map posed a restriction on the available streamflow stations since not all of North Carolina has been mapped at the 1:24,000 scale. Furthermore, when the data were inspected, some stations were found to be subject to flow regulation in twenty-five percent of the basin and were discarded. Other watersheds located in areas exhibiting karst topography were rejected. Inspection of the streamflow records revealed constant or near constant discharges for multiple years at some stations. These were removed from the sample on the assumption that they were subject to flow regulation or that serious errors existed. The final sample of basins thus held 95 cases.

Attempts to Improve Predictive Equations

The two main sources of error in regression models are model error and sampling error. Model errors include those due to incomplete or incorrect specification of the model, and measurement errors in the independent variables. Sampling errors are associated with spatial and temporal sampling problems (Moss, 1979). Small samples may not be representative of the population of watersheds for which discharge estimates are desired, even if the sample is considered random. The shorter the flow records the less accurate are the parameter estimates derived from them which are needed to fit the frequency distribution from which return interval discharges are estimated.

Unfortunately, little can be done about sampling errors since the number

---

[1] Heeding Ward's (Ward, 1971) suggestion that a twenty-five square mile upper limit set a range of sizes wherein data may be easily collected and analyzed with reasonable accuracy, a twenty-five square mile upper limit and one square mile lower limit was adopted for this study.

of gaged states and their periods of record are limited. There are some options, however, for reducing model error. The most common approaches involve the use of alternative mathematical forms, the inclusion of additional variables and more accurate measurement of independent variables. Other studies have shown that curvilinear models employing log variables can reduce error. However, the improvement is only on the order of three or four percent. This did not seem to be sufficient to warrant the use of these more complex models. But increasing the number of variables reduces the degrees of freedom for the equation, and if the ratio of variables to cases is low, may produce more inaccurate estimates. And the collection of additional variables increases the time and expense of data collection.

An alternative approach to increasing the number of variables in the equation is to control for some of the variables by grouping basins with similar topography, vegetation, rainfall or other basin attributes that influence peak discharges. Grouping reduces the error variance by restricting the range of the independent variables within each group.

Several methods exist which may accomplish this regionalization but the use of principal components analysis in conjunction with a clustering algorithm has been shown to be very effective in grouping similar basins based on their physical and climatic attributes (deCoursey, 1975; Gustafson, 1973). Homogenous regions are produced by inputting a set of physical and climatic parameters which adequately describe those facets of a watershed which influence its hydrology. These variables are standardized, correlated and grouped into components. One output is a standardized factor score which can be interpreted as a measure of the influence of each component in describing a watershed. Thus a large set of descriptors is combined and reduced to a few effective but uncorrelated variables. These factor scores are then used as input to a clustering algorithm which groups an observation's attributes. Factor scores are input rather than the raw measures because cluster analysis requires both standardized and uncorrelated variables to affectively group observations (Gustafson, 1973).

The basin characteristics in Table I were factored, and clustering was used to group the basins into relatively homogenous statistical regions. Their locations were mapped, and the groups were analyzed to determine whether a geographical pattern existed. However, no definite regions could be delineated, though a general trend based on slope and urbanization was evident. Nor could a reliable system be devised to allocate watersheds to a particular group on an a priori basis.

Consequently, it was felt that another approach utilizing characteristics of the peak discharges themselves might yield clearer results. And in this way certain pitfalls encountered in dealing with hydrologic-topographic relationships where topography may be a product of prior geologic, tectonic or climatic conditions could be avoided.

Four flow parameters were chosen to characterize the watersheds' hydrology. These are the coefficient of variation and the Z-score of the ten year flood, both selected to represent the dispersion of the data, the mean discharge per unit area, and the ten year unit discharge. The mean unit discharge and ten year unit discharge were computed to describe interstation variations in the magnitude of the flows.

When these data were subjected to principal components analysis and the clustering algorithm four groups emerged. However, even though they were studied in combination with the topographic characteristics no dependable technique was found to determine the group membership of an independent sample. As with the watershed characteristic groups, the grouping scheme had to be abandoned.

Nonetheless, the three physiographic provinces, the mountains, piedmont and coastal plain were identifiable as separate entities when lag was plotted against basin area (Figures 1-4). Others (Linsley, Kohler and Paulhus, 1958) have noted similar relationships between basin lag and location. Similarly, the urbanized basins appeared consistently as a distinct group. Consequently, the sample stations were regrouped by physiographic province, and those watersheds exhibiting a large percentage (greater than 70%) of urbanized land were allocated to a separate group.

Results

Regression equations (Table II) were calculated for each group using the two, ten and twenty-five year return interval peak discharges as dependent variables and basin area and lag coefficients as independent or predictor variables. Examination of the equations produced several interesting observations. Predictably, the influence of area is positive on the equations. But, unlike in other studies (Jackson, 1976) the regression coefficient associated with area increases with increasing return interval when the lag coefficient is negative. This phenomenon is a product of the increasing negative influence of lag on the higher return interval events. In the mountain region, where the lag coefficients are positive, the coefficients for area present a more conventional decreasing trend.

Where the effects of area on the equations can be easily explained, lag presents a more complex situation. The lag coefficient is a measure of residence time. And though one of the primary effects of urbanization on small return interval events is to modify the residence time of storm runoff, this effect of urbanization is controlled for by allocating the urban watersheds to a single group. Therefore, lag, rather than being an index of land use change, solely reflects the basin's topography. Since the lag coefficient is a func-

Figure 1. Scatter plot of drainage area vs. lag coefficient for streams in the Coastal Plain

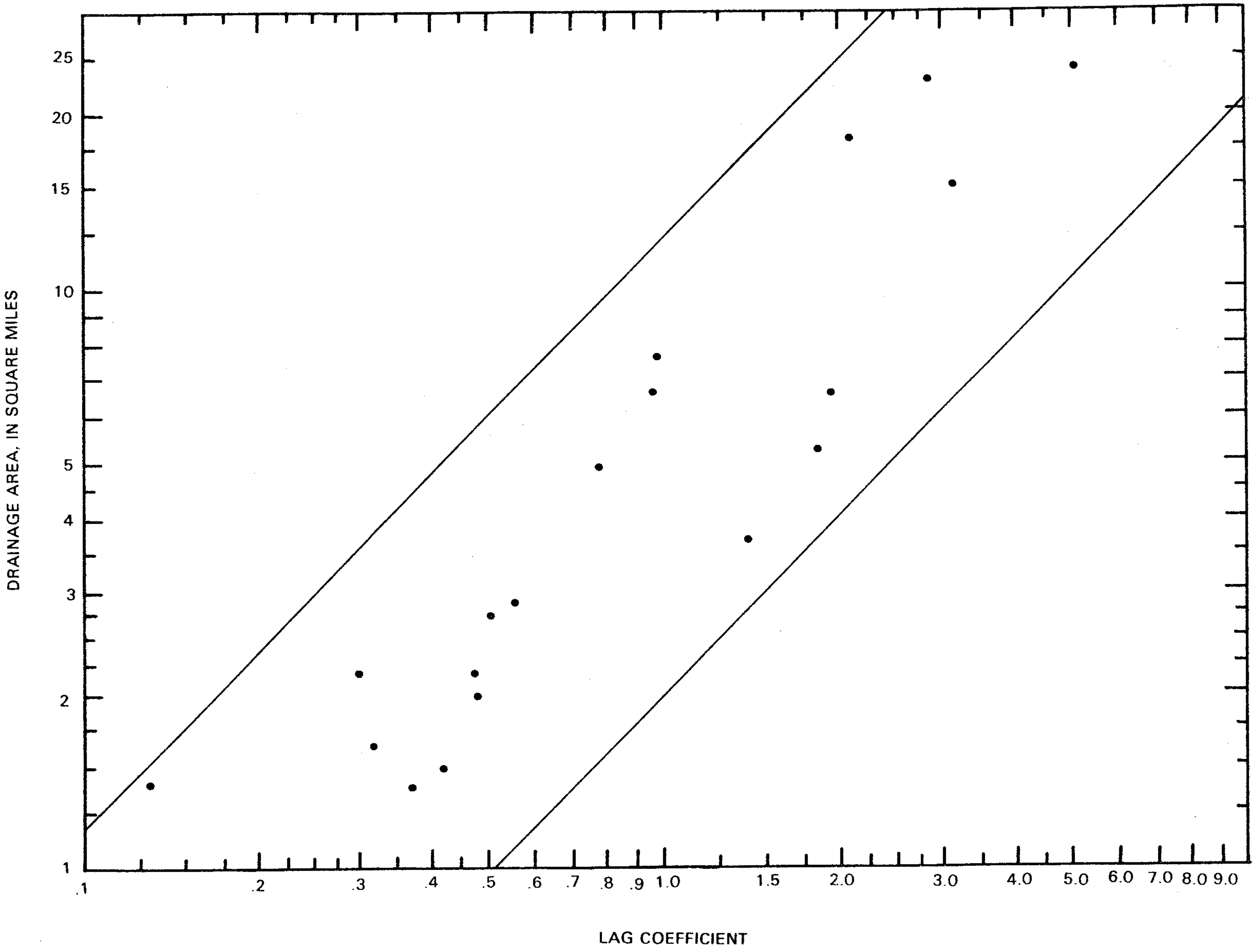


Figure 2. Scatter plot of drainage area vs. lag coefficient for streams in the Piedmont

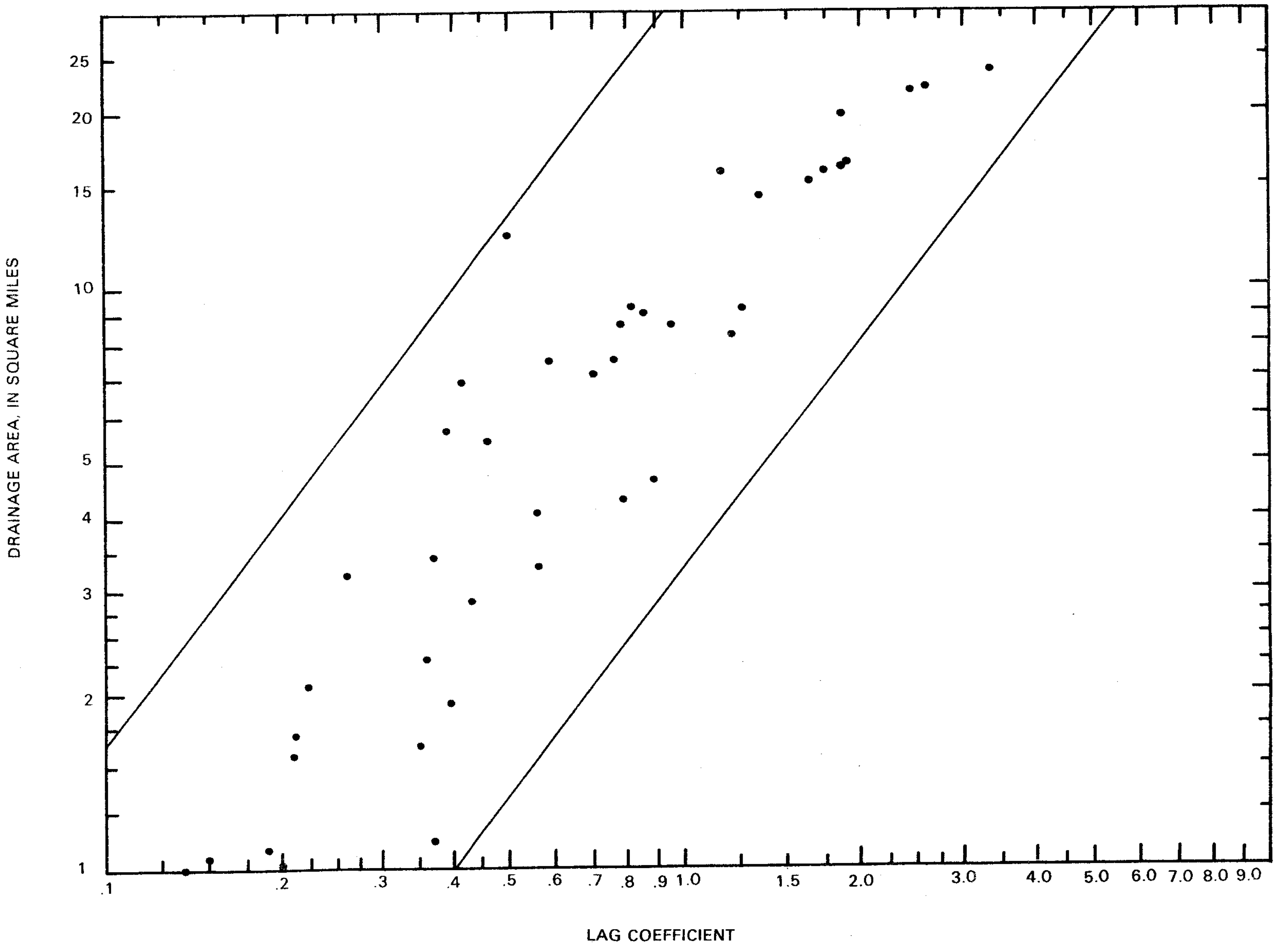


Figure 3. Scatter plot of drainage area vs. lag coefficient for streams in the Mountain region

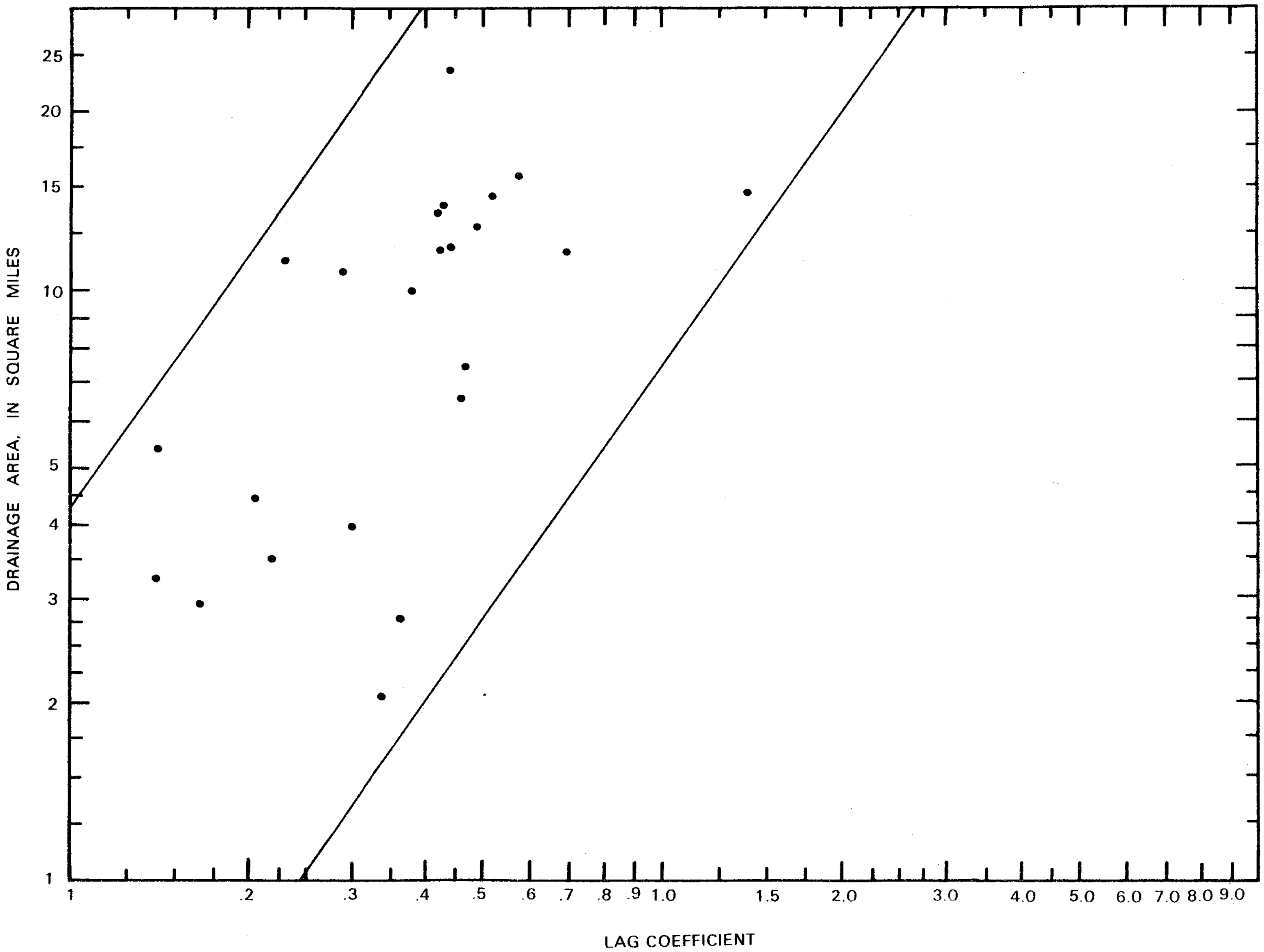


Figure 4. Scatter plot of drainage area vs. lag coefficient for urban streams.

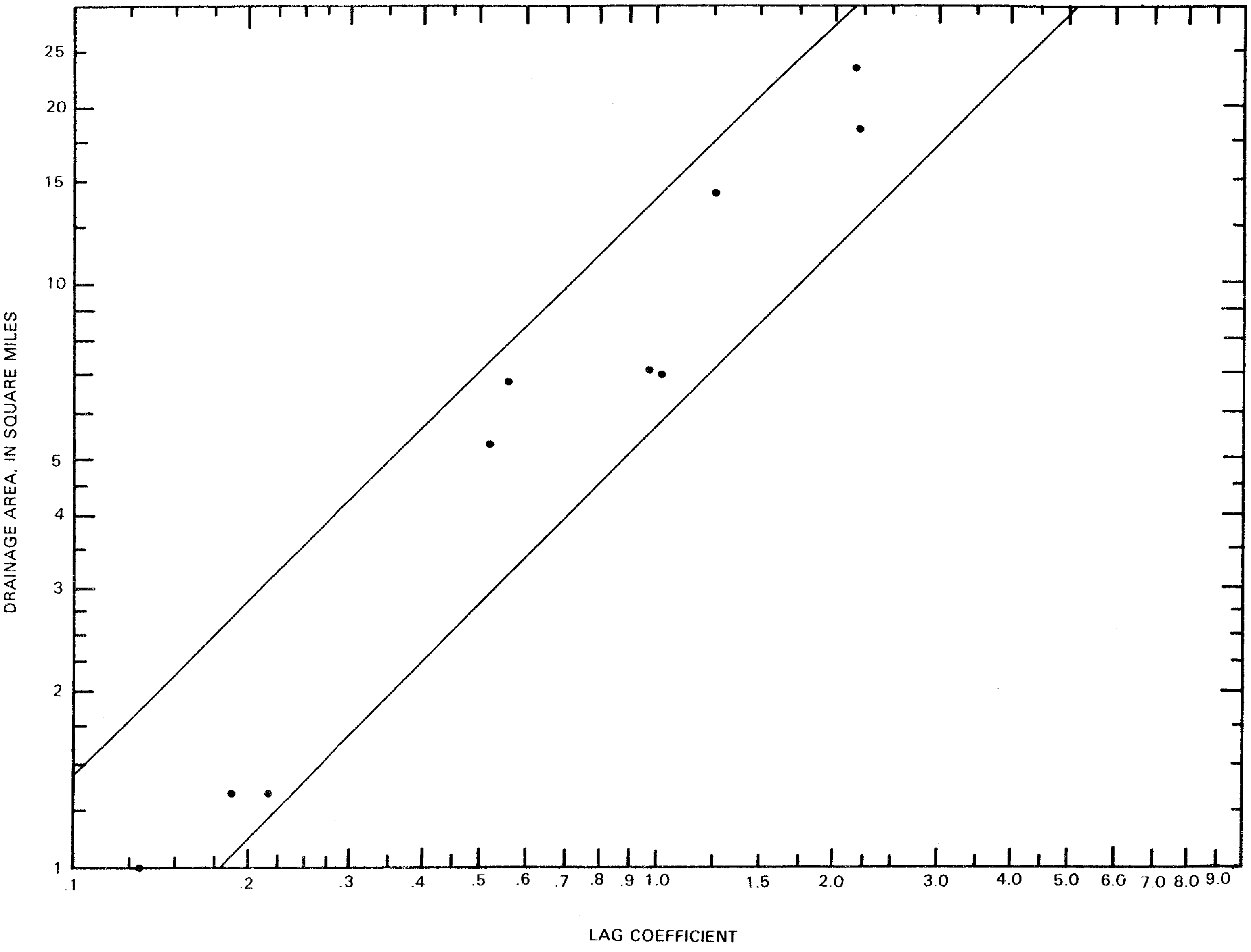


Table II

Regression Equations and Standard Errors of Estimate

COASTAL PLAIN REGION

$\log Q_2 = 1.781 + .643 \log AREA - .084 \log K$

standard error of estimate = .23738

overestimation = 72%   underestimation = 42%   average = 57%

$\log Q_{10} = 2.127 + .687 \log AREA - .168 \log K$

standard error of estimate = .23896

overestimation = 42%   underestimation = 73%   average = 58%

$\log Q_{25} = 2.263 + .712 \log AREA - .206 \log K$

standard error of estimate = .26935

overestimation = 86%   underestimation = 46%   average = 66%

PIEDMONT REGION

$\log Q_2 = 2.005 + .809 \log AREA - .205 \log K$

standard error of estimate = .16486

overestimation = 46%   underestimation = 32%   average = 39%

$\log Q_{10} = 2.446 + .809 \log AREA - .264 \log K$

standard error of estimate = .18797

overestimation = 54%   underestimation = 35%   average = 45%

$\log Q_{25} = 2.623 + .818 \log AREA - .304 \log K$

standard error of estimate = .23205

overestimation = 70.6%   underestimation = 35%   average = 56%

MOUNTAIN REGION

$\log Q_2 = 2.253 + .606 \log AREA + .255 \log K$

standard error of estimate = .20147

overestimation = 59%   underestimation = 37%   average = 48%

$\log Q_{10} = 2.712 + .599 \log AREA + .293 \log K$

standard error of estimate = .23694

overestimation = 42%   underestimation = 73%   average = 58%

$\log Q_{25} = 2.912 + .585 \log AREA + .308 \log K$

standard error of estimate = .26619

overestimation = 85%   underestimation = 46%   average = 66%

URBAN

$\log Q_2 = 1.682 + 1.482 \log \text{AREA} - 1.001 \log K$

standard error of estimate = .17429
overestimation = 49%   underestimation = 33%   average = 41%

$\log Q_{10} = 1.337 + 2.222 \log \text{AREA} - 1.831 \log K$

standard error of estimate = .16308
overestimation = 46%   underestimation = 31%   average = 39%

$\log Q_{25} = 1.154 + 2.576 \log \text{AREA} - 2.228 \log K$

standard error of estimate = .2046
overestimation = 60%   underestimation = 38%   average = 49%

tion of slope and shape, as measured by mainstream length, the greater the slope and less elongate the basin the smaller the lag coefficient tends to be, and the shorter the residence time within the basin. The influence of a lag coefficient should be negative on the equations. In the mountains, however, the opposite effect is noted. No statistical quirk was found and it is suggested that this anomolous behavior may be characteristic of watersheds with steep slopes and long main tributaries. Nonetheless, though this relationship is intuitively acceptable, further analysis is beyond the scope of this study and merits additional research. In the other equations lag behaved as expected, influencing the equation negatively.

Since the purpose of developing a regionalization scheme for estimating peak flows was to improve on existing techniques, the test of success was to compare the results of the new equations with previous work. Therefore, the residuals were calculated for each station and averaged to ascertain the general predictive errors of the equations. It is important to note that these are not standard errors, but are mean absolute errors. Standard error of estimates are given in Table II. The two, ten and twenty-five year peak flows were calculated for the sample using the U.S. Geological Survey equations developed by Jackson (1976) for the primarily rural basins and the equations recommended for estimating peak flows in urbanized watersheds developed by Putnam (1972). The results were then compared (Table III). Estimates of peak discharges were improved upon in most cases. However, the U.S. Geological Survey estimations were slightly better for the twenty-five year peak flow in the mountain region and for the ten and twenty-five year peak discharge in the piedmont. The estimates for the coastal plain were improved for all other recurrence intervals, with substantial improvement for the twenty-five year discharge.

The direction of error was also calculated to verify that the equations were

Table III

Comparative Errors for Peak Flow Estimation

		2 Year				10 year				25 Year			
		% Model Error	No. Cases	% USGS Error	No. Cases	% Model Error	No. Cases	% USGS Error	No. Cases	% Model Error	No. Cases	% USGS Error	No. Cases
COASTAL PLAIN	Overprediction	56.0	10	65.8	9	53.4	10	76.2	11	53.4	11	84.8	15
	Underprediction	31.8	10	29.5	11	31.3	10	29.8	9	35.1	9	42.6	5
	Total	43.9	20	45.9	20	42.3	20	55.3	20	45.2	20	74.2	20
PIEDMONT	Overprediction	30.0	22	37.3	22	36.2	24	19.8	17	48.0	23	19.3	14
	Underprediction	25.8	19	26.7	19	31.7	17	37.6	24	33.3	18	41.3	27
	Total	28.0	41	32.6	41	34.3	41	30.2	41	41.5	41	33.8	41
MOUNTAINS	Overprediction	53.0	11	61.2	17	49.6	14	73.7	9	64.6	13	84.9	8
	Underprediction	21.4	13	22.1	7	34.7	10	30.9	15	36.2	11	32.2	16
	Total	34.6	24	48.9	24	43.4	24	50.0	24	51.9	24	49.8	24
		% Model Error	No. Cases	% Putnam Error	No. Cases	% Model Error	No. Cases	% Putnam Error	No. Cases	% Model Error	No. Cases	% Putnam Error	No. Cases
URBAN	Overprediction	35.2	6	179.0	4	25.2	6	51.1	6	28.0	7	61.3	6
	Underprediction	23.2	4	88.9	6	25.7	4	63.5	4	34.5	3	47.8	4
	Total	29.2	10	106.9	10	25.4	10	56.0	10	30.0	10	55.9	10

not biased to underpredict or overpredict flows. No such bias was noted in the new equations, (Table III) but in several instances the U.S. Geological Survey equations appear to exhibit some bias. The Geological Survey equations overpredicted flows in two-thirds of the sample for the twenty-five year peak flows in the coastal plain group. This tendence to overpredict also contributes to the large error term. Similarly, a tendency to underpredict flows for the twenty-five year peak flows in the mountain and piedmont groups contributed to the relatively lower absolute mean error terms for the Geological Survey technique.

Procedure for Using the Technique

The equations developed in this study can be used to estimate peak flows for small ungaged streams in North Carolina and Virginia. In order to estimate the two, ten and twenty-five year return interval peak discharges follow the procedure listed below.

1. Determine that the stream is not regulated, located in a region of karst topography (limestone) or affected strongly by tides.
2. Calculate the drainage area in square miles from the best available map and check that the watershed area is not greater than 25 $mi.^2$ nor smaller than one $mi.^2$.
3. Calculate the lag coefficient by measuring the slope of the mainstream from the point 10 percent and 85 percent upstream from the site where the discharge estimate is desired. Determine its square root and divide the total mainstream length by this figure.
4. Determine whether 70 percent of the watershed is urbanized. If it is, use the equations or graphs developed for urban basins.
5. Ascertain whether the site is located in the coastal plain, piedmont, or mountain areas, then select the appropriate equation or graph.

Figure 5 Graphical solution of equations for streams in the Piedmont region.

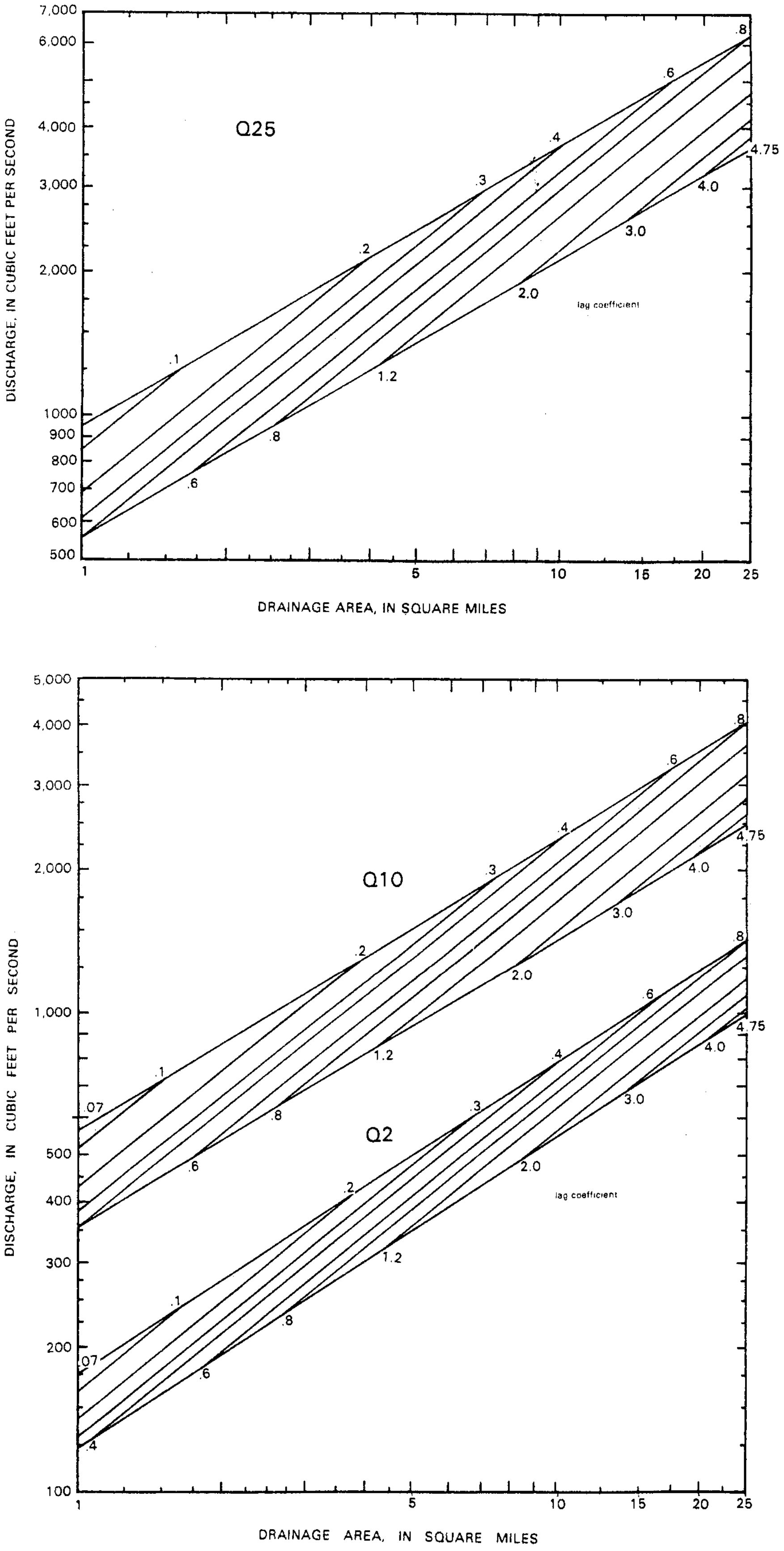


Figure 6. Graphical solution of equations for streams in the Mountain region.

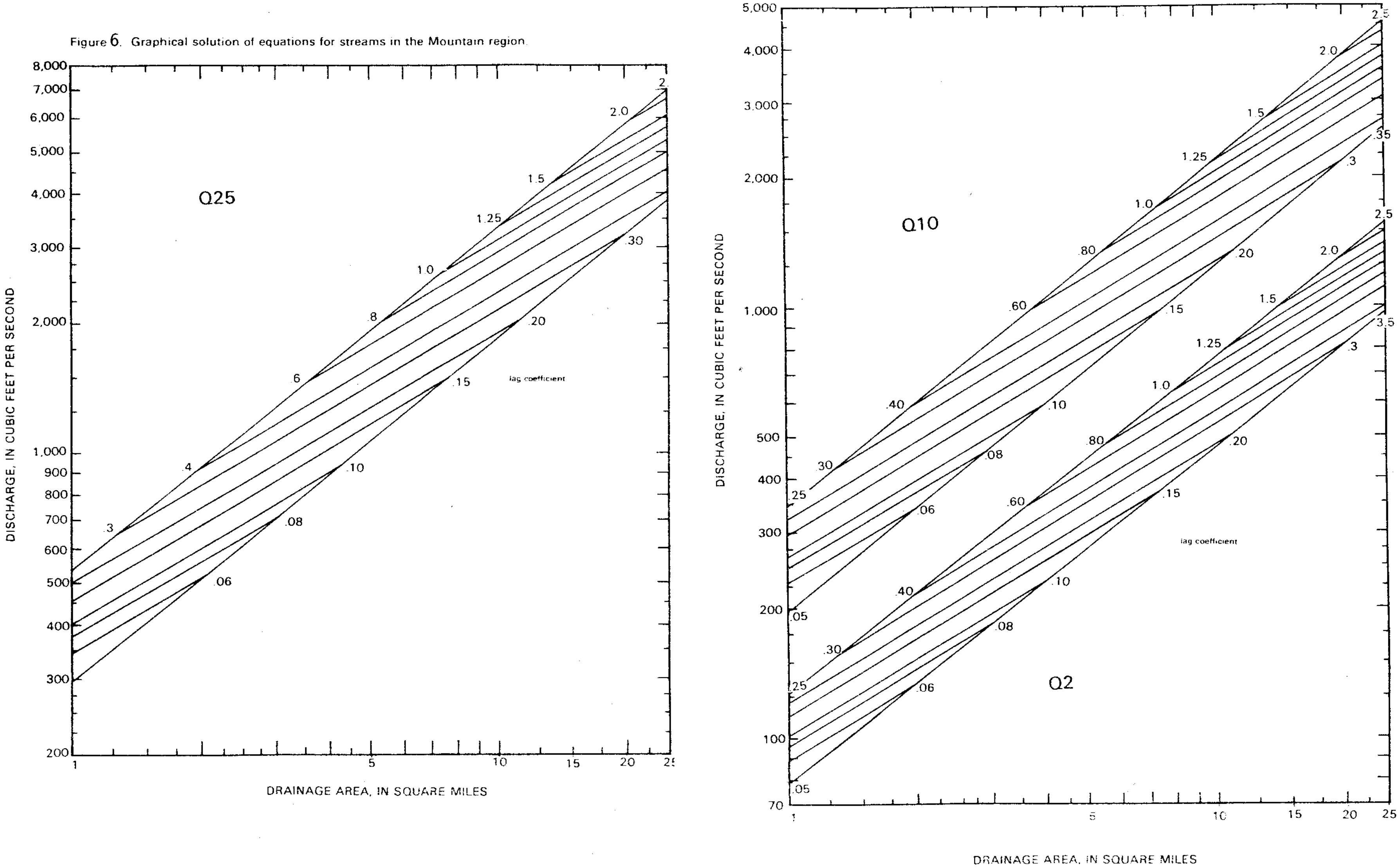


Figure 7. Graphical solution of the equations for urban streams. (Q25)

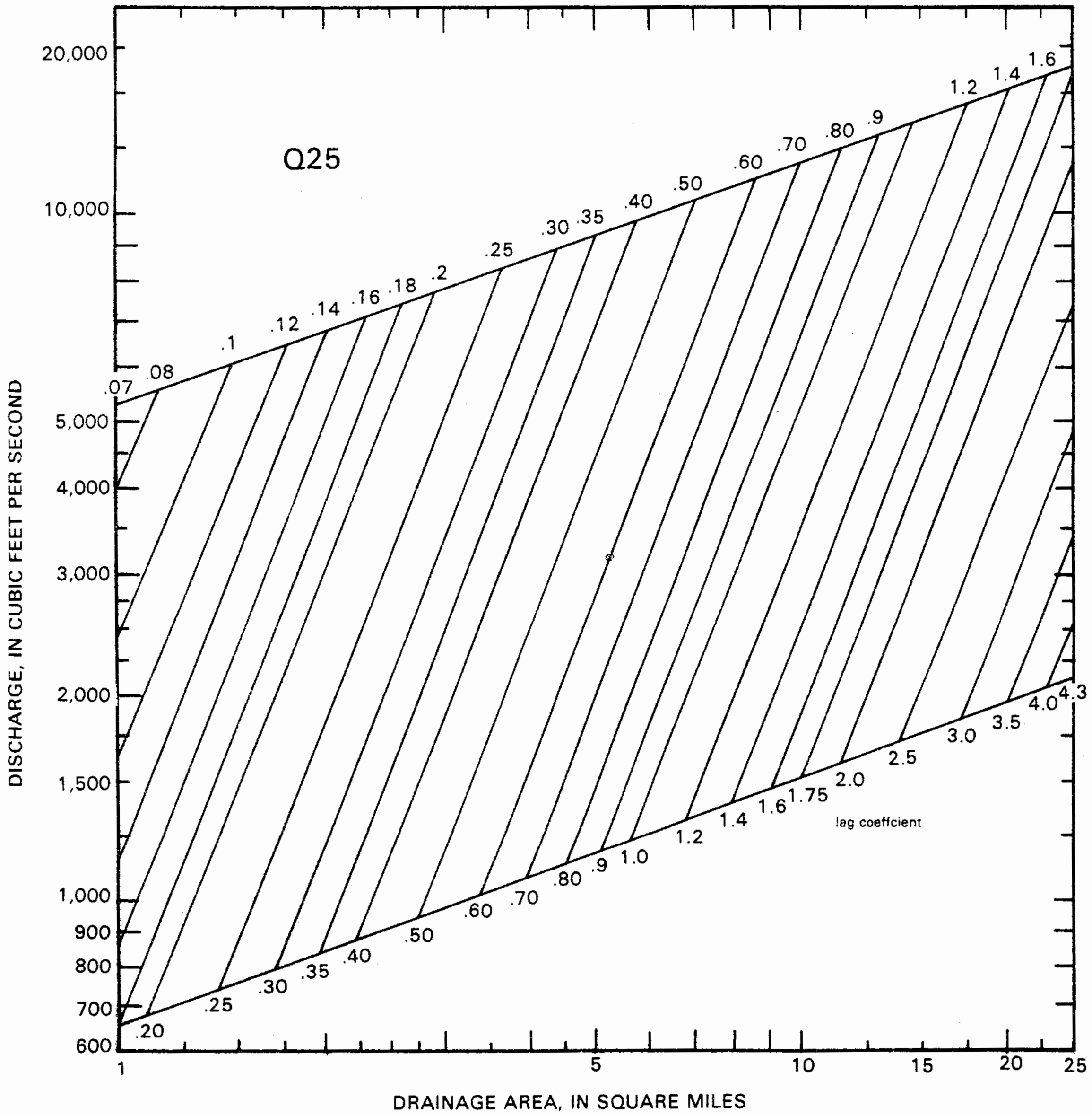


Figure 8. Graphical solution of the equations for urban streams. (Q2 and Q10)

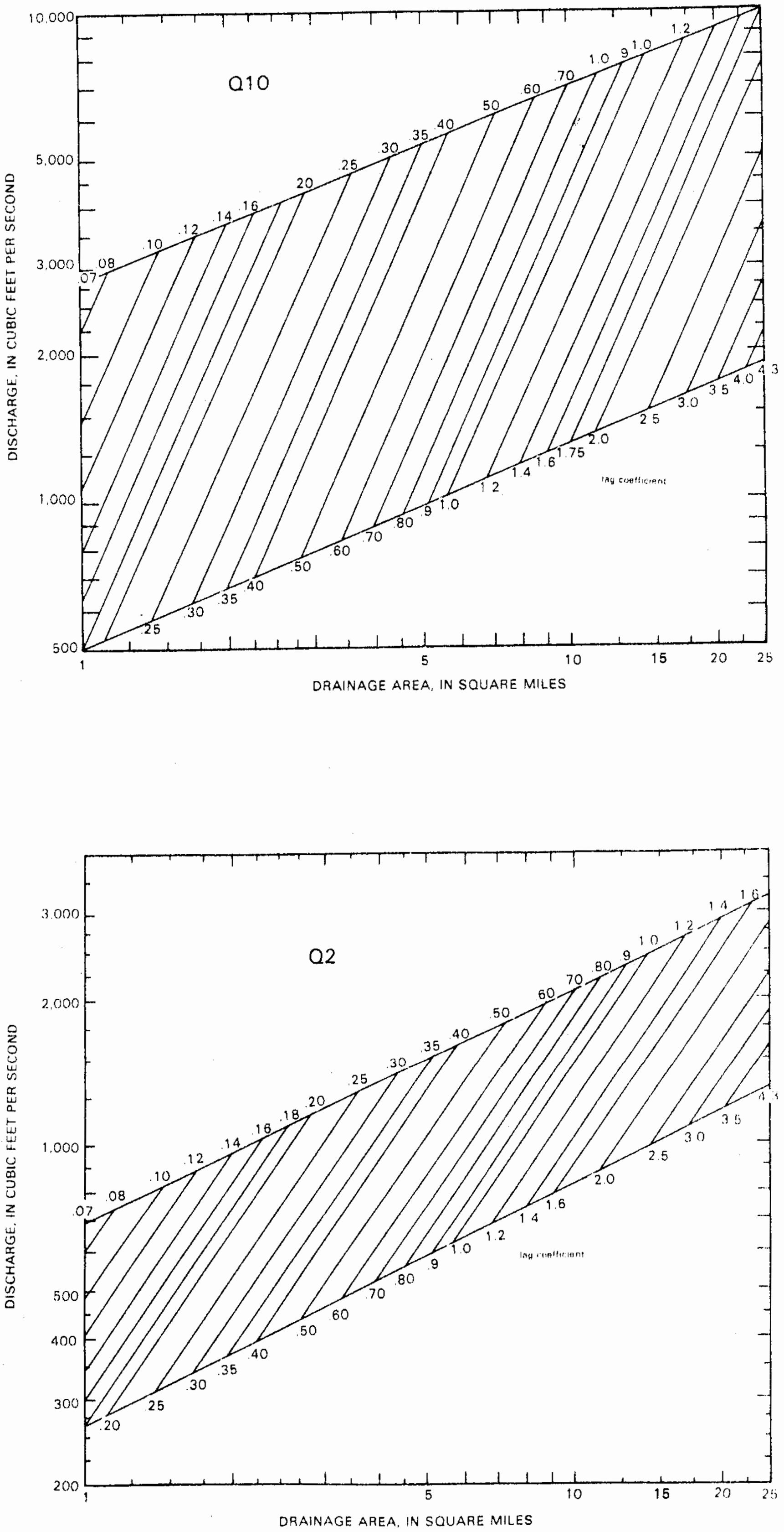
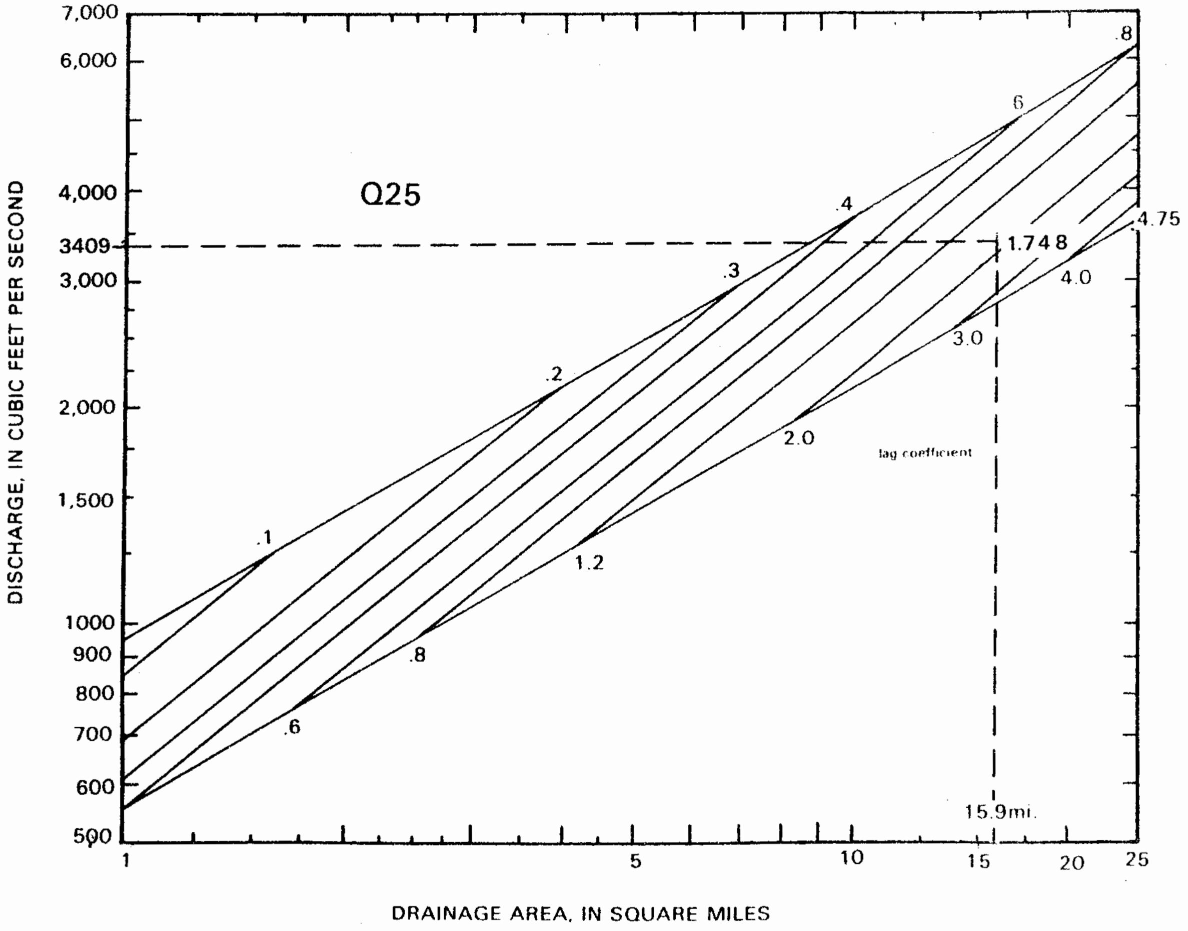


Figure 9. Example of the use the nomograph to graphically estimate the peak flow for a stream.



6. In certain cases streams requiring peak flow estimates may be influenced by inadvertent detention due to undersized culverts. As well, the discharges in a highly channelized system may be underpredicted by these equations. If some characteristic is known to affect the peak flows, the predicted peak discharges should be adjusted using judgment based on the influencing conditions.

Example of the Use of the Estimating Procedure

The twenty-five year peak flow is desired for a point on Horsepen Creek at Battleground, North Carolina. No regulation, karst topography or tidal influence was present. The drainage area has been determined to be 15.9 mi.^{2}. The lag coefficient was calculated as 1.748. Less than 70 percent of the watershed was urbanized and the site was located in the Piedmont region. Therefore the equation developed for the Piedmont region was used and the twenty-five year return interval peak flow was estimated to be 3,409 cubic feet per second. Instead of using the equation the discharge value can be calculated graphically (Figures 5-8). Locating the drainage area of the basin on the abscissa of the graph (Figure 9) move upward until that line intersects with the calculated lag coefficient. Then read the discharge from the left.

Recommendations for Further Research

1. The regional equations improved estimates of the selected peak flows in most cases but errors for some sites were quite large (greater than 100 percent). Similar errors were produced by the U.S. Geological Survey equations for these same watersheds. And though these basins were examined to ascertain the cause for the anomolous flow characteristics no definite cause was found. Research should be conducted to determine the factors influencing the flows and some means should be developed to better identify regionally anomolous basins and modify existing streamflow estimation

procedures.

2. There is need for further research to determine whether the positive effect of the lag variable in the equations for mountain streams is real or is an artifact of the statistical technique. In most of the literature on drainage basin hydrology and in most discussions of hydrograph analysis, increased lag times are assumed to decrease peak discharges. Yet, examination of the data revealed no obvious reason why the result should be viewed as an artifact of the statistical procedures used.

   Further research is clearly warranted to determine the real cause of this apparently anomolous result. In a recent article two Chinese investigators (Lui and Wang, 1980) presented a very detailed model for predicting peak discharges for small watersheds in China (.006 to 163 Km^2). They indicate that "the characteristics of the formation of small watershed peak flows are different than that of large or medium river basins...." and that, in mountainous areas, the rate of overland flow is the most critical factor. Their set of simultaneous equations for calculating peak discharge appear to lend logical support to the equation for mountainous basins. However, no final conclusion can be reached until the model is actually applied, using data from the three geographic regions of North Carolina.

Appendix A

Since the Land Use and Land Cover series is published at the 1:250,000 scale for Virginia and North Carolina the watersheds (once identified and traced from the 1:24,000 scale topographic maps) were reduced approximately ten times before the boundaries were located on the land use and land cover sheets. The actual area of each category was determined using an overlay grid. Then this square mile figure was converted to a percentage by dividing it by the basin's area. As these three categories encompass the range of land uses and covers, the percentages total to one hundred.

References Cited

1. Benson, M.A., 1959, "Channel Slope Factor in Flood Frequency Analysis:" Am. Soc. Civil Eng. Proc. Jour. Hydraulics Div. V. 85 pt. 1 p. 1-9.

2. Bevens, K.J. and M.J. Kirkby, 1979, "A Physically Based, Variable Contributing Area Model of Basin Hydrology:" Hydrological Sciences Bull. V. 24 no. 1 p. 43-69.

3. Carlston, C.W., 1963, "Drainage Density and Streamflow:" U.S. Geol. Surv. Prof. Pap. 422-C, p. 1-8.

4. De Coursey, D.G., 1973, "Objective Regionalization of Peak Flow Rates:" in Schulz, E.F., V.A. Koelzer and Khalid Mahmood, eds. Floods and Droughts, Proceedings of the Second International Symposium in Hydrology, September, 1972. Fort Collins, Colo.: Water Resource Publications p. 395-405.

5. Dunne, T . and L.B. Leopold, 1978, Water in Environmental Planning: San Francisco, W.H. Freeman and Co., 817 p.

6. Fleming, G., 1975, Computer Simulation Techniques in Hydrology. New York: American Elsevier Publishing Co. 333 p.

7. Gardiner, V., 1979, "Estimation of Drainage Density from Topological Variables:" Water Resour. Res. V. 15 No. 4 p. 909-918.

8. Gardiner, V., and C.C. Parks, 1978, "Drainage Basin Morphometry: Review and Assessment:" Progress in Physical Geography, V. 2, p. 1-35.

9. Graf, W., 1975, "A Cumulative Stream Ordering System:" Geographical Analysis, V. 7, p. 335-340.

10. Guevera, E., 1975,"Indirect Determination of Maximum Flood Discharges by Means of Geomorphological Parameters in Twenty Catchments of the Peruvian Coast:"Hydrologic Characteristics of River Basins Symposium, International Association of Scientific Hydrology - UNESCO. 882 p.

11. Gustafson, G.C., 1973, Quantitative Investigation of the Morphology of Drainage Basins Using Orthophotography: Munich, Geographisches Institut der Universitat Munchen, 119 p.

12. Haan, C.T. and D.M. Allen, 1973, "Comparison of Multiple Regression and Principal Components Regression for Predicting Water Yields in Kentucky:" Water Resour. Res. V. 8, No. 6, pp. 1593-1596.

13. Hack, J.T., 1957, "Studies of Longitudinal Stream Profiles in Virginia and Maryland:" U.S. Geol. Surv. Prof. Pap. 294-B, p. 45-97.

14. Hinson, H.G., 1965, "Floods on Small Streams in North Carolina, Probable Magnitude and Frequency:" U.S. Geological Survey Circ. 517, 7 p.

15. Horton, R.E., 1945,"Erosional Development of Streams and Their Drainage Basins; Hydrophysical Approach to Quantitative Morphology:" Geol. Soc. America Bull., V. 56, p. 275-370.

16. Jackson, N.M. Jr., 1976, Magnitude and Frequency of Floods in North Carolina: Raleigh, N.C., U.S. Geological Survey Water Resources Investigations, 26 p.

17. Lichty, R.W. and F. Liscum, 1978, "A Rainfall-Runoff Modeling, Procedure for Improving Estimates of T-Year (Annual) Floods for Small Drainage Basins:" U.S. Geol. Surv. Water-Resour. Invest. 78-7., 44 p.

18. Linsley, R.K. and J.B. Franzini, 1979, Water Resources Engineering: New York, McGraw-Hill Book Co., 716 p.

19. Linsley, R.K. Jr., Kohler, M.A., and J.L.H. Paulhus, 1958, Hydrology for Engineers: New York, McGraw-Hill Book Co., 340 p.

20. Lui Changming and Wang Guante, 1980, "The Estimation of Small-Watershed Peak Flows in China:" Water Resources Research, v. 16, pp. 881-886.

21. Miller, J.B. and M.D. Newson, 1975, "Flood Estimation from Catchment Characteristics:" in T.J. Darwent, ed. Flood Studies Conference Proceedings: London, Institution of Civil Engineers. p. 57-60.

22. Morisawa, M., 1962, "Quantitative Geomorphology of Some Watersheds in the Appalachian Plateau:" Geological Soc. Amer. Bull., v. 73, pp. 1025-1046.

23. Moss, M.E., 1979, "Space, Time and the Third Dimension (Model Error):" Water Resources Res., V. 15, No. 6, pp. 1797-1800.

24. Osborn, J.F., 1974, "Determining Streamflows from Geomorphic Parameters:" Am. Soc. Civil Eng. Proc. V. 100, p. 455-475.

25. Osborn, J., 1975, "Predicting Ungaged Low Flows in Diverse Hydrologic Provinces Using River Basin Geomorphic Characteristics:" Hydrologic Characteristics of River Basins Symposium, International Association of Scientific Hydrology - UNESCO. 882 p.

26. Ozga-Zielinski, M., 1975, "Statistical Analysis of Relationships Among Characteristics of Streamflow and the Physiographic Characteristics of the River Basin:" Hydrologic Characteristics of River Basins Symposium, International Association of Scientific Hydrology - UNESCO. 822 p.

27. Patton, P.C. and V.R. Baker, 1976, "Morphometry and Floods in Small Drainage Basins Subject to Diverse Hydrogeomorphic Controls:" Water Resour. Res. V. 12, No. 5, p. 941-952.

28. Putnam, A.L., 1972,"Effects of Urban Development on Floods in the Piedmont Province of North Carolina:" U.S. Geological Survey Open File Report, 87 p.

29. Reynolds, E.R., 1972, The Discharge of South Carolina Streams as it Relates to Link Magnitude: Columbia, SC., M.A. Thesis, Department of Geology, University of South Carolina.

30. Schumm, S.A., 1956, "Evolution of Drainage Systems and Slopes in Badlands at Perth Amboy, New Jersey:" Geol. Soc. America Bull., V.67, p. 597-646.

31. Shelton, C.H., and J.I. Sewell, 1969, "Parameter Screening for Watershed Analysis:" Trans. Amer. Soc. Agr. Eng. V. 12, No. 4., p. 533-539.

32. Thomas, D.M., and M.A. Benson, 1970, "Generalization of Streamflow Characteristics from Drainage Basin Characteristics:" *U.S. Geol. Surv. Water Supply Paper 1975*, 55 p.

33. Trainer, F.W., 1969, "Drainage Density as an Indicator of Baseflow in Part of the Potomac River:" *U.S. Geol. Surv. Prof. Pap. 650-C.*, p. 177-183.

34. Ward, R.C., 1971, *Small Watershed Experiments: An Appraisal of Concepts and Research Developments*: Hull, England, Occasional Papers in Geography No. 18, University of Hull Publications, 254 p.

35. Wong, S.T., 1963, "A Multivariate Statistical Model for Predicting Mean Annual Flood in New England:" *Annals of the Association of American Geographers*, V. 53, p. 298-311.

36. Wong, S.T., 1979, "A Dimensionally Homogeneous and Statistically Optimal Model for Predicting Mean Annual Flood:" *Jour. of Hydrology*, V. 42, p. 269-280.