SPR-Part B
MD-20-SP910B4G-2

MARYLAND DEPARTMENT
OF TRANSPORTATION

STATE HIGHWAY
ADMINISTRATION

Larry Hogan
Governor
Boyd K. Rutherford
Lt. Governor
Pete K. Rahn
Secretary
Gregory Slater
Administrator

MARYLAND DEPARTMENT OF TRANPORTATION
STATE HIGHWAY ADMINISTRATION

RESEARCH REPORT

REGRESSION EQUATIONS FOR ESTIMATING FLOOD DISCHARGES FOR THE EASTERN COASTAL PLAIN REGION OF MARYLAND

Wilbert O. Thomas, Jr. and Windsor Sanchez-Claros

Michael Baker International

FINAL REPORT

November 2019

This material is based upon work supported by the Federal Highway Administration under the State Planning and Research program. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the Federal Highway Administration or the Maryland Department of Transportation. This report does not constitute a standard, specification, or regulation.

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. MD-20-SP910B4G-2	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle Regression Equations for Estimating Flood Discharges for the Eastern Coastal Plain Region of Maryland		**5. Report Date** November 2019
		6. Performing Organization Code
7. Author(s) Wilbert O. Thomas, Jr. and Windsor Sanchez-Claros		**8. Performing Organization Report No.**
9. Performing Organization Name and Address Michael Baker International 1306 Concourse Drive, Suite 500 Linthicum Heights MD 21090		**10. Work Unit No.**
		11. Contract or Grant No. SP910B4G
12. Sponsoring Agency Name and Address Maryland Department of Transportation (SPR) State Highway Administration Office of Policy & Research 707 North Calvert Street Baltimore MD 21202		**13. Type of Report and Period Covered** SPR-B Final Report (December 2018-November 2019)
		14. Sponsoring Agency Code (7120) STMD - MDOT/SHA
15. Supplementary Notes		

16. Abstract

Regression equations were updated for estimating the 1.25-. 1.50-, 2-, 5-, 10-, 25-, 50-, 100-, 200- and 500-year flood discharges for the Eastern Coastal Plain Region of Maryland. The regression equations were based on flood discharges for 36 stations and drainage area, in square miles; land slope, in percent; and hydrologic soils group A, in percent of the watershed area (five stations were identified as outliers). For the 36 stations used in the analysis, 16 stations are still active and 20 are discontinued stations as of 2017. Many of the small stream stations (less than 10 square miles) were only operated for the period 1965 to 1976 and experienced large floods in this short period of record particularly in 1967 and 1975. Historical information available in the U.S. Geological Survey Peak Flow File were used to adjust the frequency curves to be more representative of a longer record. The updated regression equations will be used by the Maryland Department of Transportation State Highway Administration in the design of bridges and culverts in Maryland. The updated regression equations will also be included in the fifth version of the Maryland Hydrology Panel report entitled "Application of Hydrologic Methods in Maryland" that will be published in 2020.

17. Key Words Flood discharges, regression equations, frequency analysis		**18. Distribution Statement** This document is available from the Research Division upon request.	
19. Security Classif. (of this report) None	**20. Security Classif. (of this page)** None	**21. No. of Pages** 14	**22. Price**

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Table of Contents

Appendices

Regression Equations for Estimating Flood Discharges for the Eastern Coastal Plain Region of Maryland

Executive Summary

Updated regression equations were developed for the Eastern Coastal Plain (ECP) Region for estimating the 1.25-, 1.5-, 2-, 5-, 10-, 25-, 50-, 100-, 200- and 500-year flood discharges. Flood frequency analyses were performed for 41 gaging stations using Bulletin 17C (England and others, 2018) and the U.S. Geological Survey (USGS) PeakFQ program (https://water.usgs.gov/software/PeakFQ/). Of the 41 gaging stations, 22 stations were in Maryland and 19 stations in Delaware with 16 active stations and 25 discontinued stations as of 2017.

A regional skew analysis was performed by analyzing the station skew using data for 15 rural stations in the Eastern Coastal Plain (ECP) Region and eight rural stations in the Western Coastal Plain (WCP) Region. A mean skew of 0.38 with standard error of 0.38 was determined in this analysis. The regional skew was weighted with the station skew for all watersheds with the following exceptions. Graphical frequency analyses were performed for three gaging stations because the frequency curves were S-shaped likely due to floodplain storage. In addition, record length was extended for two short record stations using graphical analysis with nearby long-term stations.

There were nine stations that had 62 to 74 years of record and seven of these stations had statistically significant upward trends due to large floods near the end of the record in 1989, 1999, 2011 and 2016. The impervious area of the watersheds was less than 10 percent of the drainage area so the upward trends were not related to urbanization. The upward trends were assumed to be climatic persistence or variability (not a permanent change in climate) and the entire period of record was used in the frequency analysis.

Watershed characteristics shown to be statistically significant in previous studies were evaluated in the regional regression analysis. Legacy SSURGO soils data in GISHydroNXT and SSURGO soils data dated May 2018 from the Natural Resources Conservation Service (NRCS) Soil Survey web site were evaluated and the May 2018 SSURGO data provided the more accurate regression equations. The final regression equations were based data on 36 gaging stations (five outlier stations were omitted) using drainage area, in square miles; land slope, in percent, and A soil, in percent, based on the May 2018 SSURGO data as explanatory variables. Regression estimates from the 100- and 10-year equations were compared to gaging station data and shown to be reasonably unbiased. The 100- and 10-year regression estimates were also compared to regression estimates from equations developed in 2010 and documented in the July 2016 version of the Maryland Hydrology Panel report. The 2010 regression equation provided 100-year discharges about 10 percent higher, on average, than the 2019 equations and 10-year discharges about seven percent higher, on average, than the 2019 equations.

The updated regression equations will be used by the Maryland Department of Transportation State Highway Administration (MDOT SHA) in the design of bridges and culverts in Maryland. The updated regression equations will be included in the fifth version of the Maryland Hydrology Panel report entitled “Application of Hydrologic Methods in Maryland” that will be published in 2020.

Regression Equations for Estimating Flood Discharges for the Eastern Coastal Plain Region of Maryland

Introduction

Fixed region regression equations are used to estimate flood discharges for bridge and culvert design and floodplain mapping in Maryland by several state and local agencies. These empirical equations are developed based on relations between flood discharges at gaging stations and watershed characteristics that can be estimated from available digital data layers. For ungaged locations, the watershed characteristics are used in the regression equations to predict the flood discharges. The Maryland Department of Transportation State Highway Administration (MDOT SHA) uses the regression equations to primarily evaluate the reasonableness of flood discharges estimated using the TR-20 watershed model (Maryland Hydrology Panel, 2016). The objective of the current analysis is to update the Fixed Region regression equations for the Eastern Coastal Plain Region for estimating the 1.25-, 1.5-, 2-, 5-, 10-, 25-, 50-, 100-, 200-, and 500-year flood discharges using the following data:

- Annual peak flow data through the 2017 water year if available,
- Flood frequency analyses using Bulletin 17C (England and others, 2018),
- Watershed characteristics computed using GISHydroNXT (land use data for various time periods, DEM data and SSURGO data currently in GISHydroNXT), and
- SSURGO data downloaded from the Natural Resources Conservation Service (NRCS) web site in May 2018.

Both sets of SSURGO data will be evaluated as explanatory variables to see which data set is most appropriate for estimating flood discharges using regression equations.

Previous Studies

Several studies have been completed since 1980 that developed regional regression equations for Maryland. Following is a brief description of the data used in previous regression equations for the Eastern Coastal Plain Region:

- U.S. Geological Survey (USGS) Open-File Report 80-1016 (Carpenter, 1980) – used drainage area, channel slope, percent storage, percent forest cover and percent A and D soils based on STATSGO soils and annual peak flow data through the 1977 water year,
- USGS Water-Resources Investigations Report 95-4154 (Dillow, 1996) – used drainage area, runoff curve number (STATSGO data), basin relief, percent forest, percent storage and annual peak flow data through the 1990 water year,
- Maryland Hydrology Panel report (2006) and Moglen and others (2006) – used drainage area, basin relief and percent A soils based on STATSGO data and annual peak flow data through the 1999 water year,
- USGS Scientific Investigations Report 2006-5146 (Ries and Dillow, 2006) (report for Delaware streams) – used drainage area, land slope and percent A soils based on STATSGO data and annual peak flow data through the 2004 water year, and
- Maryland Hydrology Panel report (2010) – used drainage area, land slope, percent A soils based on SSURGO data and annual peak flow data through the 2006 water year.

A water year is from October 1 to September 30 with the ending month determining the water year. For example, the 2017 water year is from October 1, 2016 to September 30, 2017. The 2016 Maryland

Hydrology Panel report has the same equations for the Eastern Coastal Plain (ECP) Region as the 2010 report because the regression equations for the coastal plain regions have not been updated since 2010.

Flood Frequency Analyses at Gaging Stations

Flood frequency estimates were updated through 2017 if data were available using the USGS PeakFQ program (https://water.usgs.gov/software/PeakFQ/) that implements Bulletin 17C (England and others, 2018). Flood data were compiled and analyzed for 41 gaging stations in the ECP: 22 stations in Maryland and 19 stations in Delaware. The location of the gaging stations is shown in Figure 1 that defines the four major hydrologic regions in Maryland: Appalachian Plateau and Allegheny Ridge, Blue Ridge-Piedmont, Western and Eastern Coastal Plains. The gaging stations are numbered in downstream order as listed in Appendix 1.

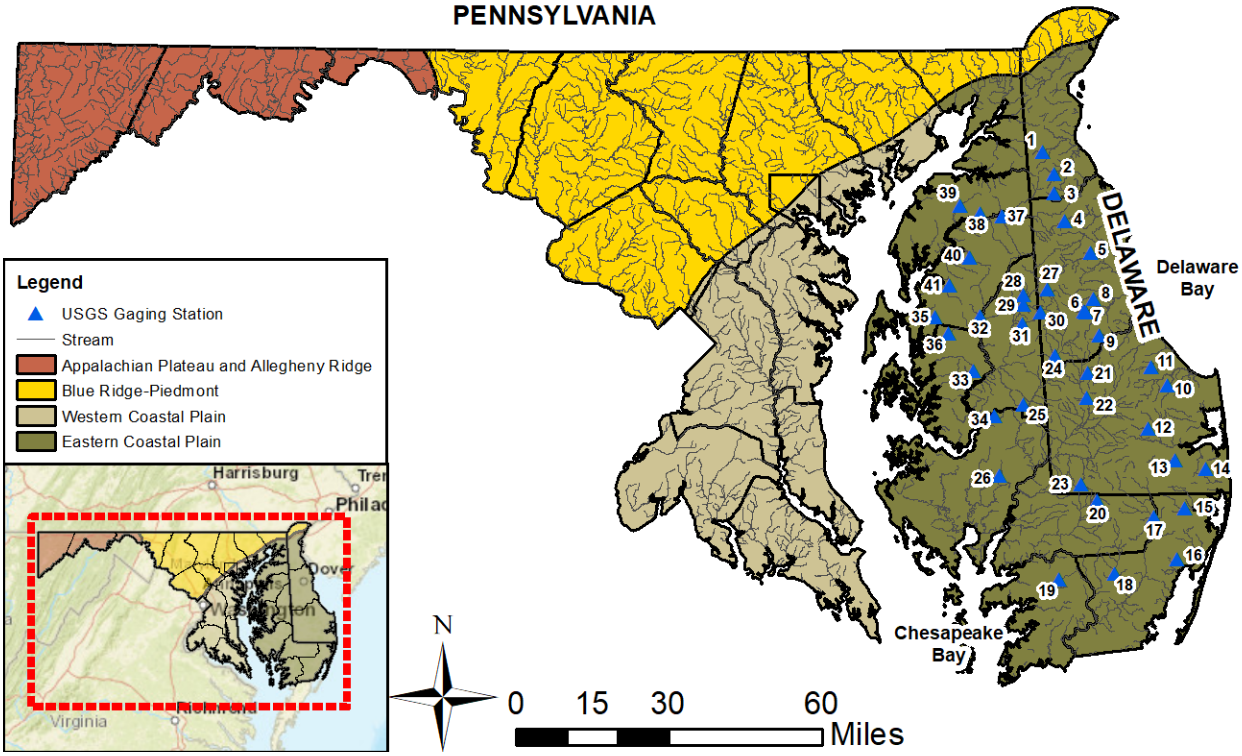


Figure 1. Location of gaging stations in the Eastern Coastal Plain Region of Maryland.

For the 41 stations shown in Figure 1, only 16 stations were still active in 2017. Record lengths ranged from 9 to 75 years with 19 stations having 20 or more years of record. As noted earlier, there are 22 stations in Maryland and 19 stations in Delaware shown in Figure 1. The 2010 analysis for the ECP Region evaluated 16 stations in Maryland and 15 stations in Delaware. Therefore, six new stations in Maryland and four new stations in Delaware were added to the current analysis.

Rural stations with 19 or more years of record were analyzed to obtain station skew to estimate a new regional skew. Using 15 stations from the ECP Region and eight stations from the Western Coastal Plain Region (WCP), a regional skew of 0.38 with a standard error of 0.38 was determined. The 2010 regional skew analysis for the ECP resulted in a regional skew of 0.45 with a standard error of 0.41 so the change in regional skew is minimal.

For three stations, the frequency curves were S-shaped (likely due to floodplain storage) and the plotting positions for the logarithms of the data did not fit a Pearson Type III distribution very well. A graphical frequency analysis was performed for: Blackbird Creek at Blackbird, DE (01483200), Manokin Branch near Princess Anne, MD (19486000) and Marshyhope Creek near Adamsville, DE (10488500). Records were extended for two short record stations using a graphical analysis with a nearby long-record station: Southeast Creek at Church Hill, MD (01494000) extended with records from Beaverdam Branch at Matthews, MD (01492000) and Three Bridges Branch at Centerville, MD (01494150) extended with records from Sallie Harris Creek near Carmichael, MD (01492500).

Thirteen of the 41 stations had flood data for the period 1965 to 1976 when the USGS small streams program was active (gaging stations less than 10 square miles). This was an active flood period with major floods in 1967 and 1975. Most of the small stream sites experienced their maximum flood in August 1967 that was generally known to be the highest flood since 1935. A historical period of 40 years (highest flood in the period 1936 to 1976) was used in the frequency analysis for those stations experiencing a major flood in 1967 to obtain more reasonable estimates of the flood discharges. Historical information was also available for some stations for the 1975 flood. There were 10 small stream stations where Carpenter (1980) extended the records using a rainfall-runoff model and nearby long-term climatic data. The frequency estimates from Carpenter (1980) were evaluated and it was determined that the frequency estimates based on observed and historical data as described above were more reasonable.

There were nine stations that had 62 to 74 years of record and seven of these stations had statistically significant upward trends due to large floods near the end of the record in 1989, 1999, 2011 and 2016. The impervious area of the watersheds was less than 10 percent of the drainage area so the upward trends were not related to urbanization. The upward trends were assumed to be climatic persistence or variability (not a permanent change in climate) and the entire period of record was used in the frequency analysis. The only exception was Marshyhope Creek near Adamsville, DE (01488500) where extensive channelization occurred during 1969-71. The period 1972 to 2017 was used for the frequency analysis that was based on a graphical analysis as noted above.

The final flood frequency estimates were based on a weighted skew (combining station and regional skew) and those flood discharges are provided in Appendix 1. The period of record and years of record at the gaging stations are given in Appendix 2.

Watershed Characteristics Evaluated for the Regression Analysis

The watershed characteristics evaluated for the regression analysis included those that were statistically significant in previous regression analyses and were estimated using the digital data in GISHydroNXT (http://www.gishydro.eng.umd.edu/document.htm). The watershed characteristics included:

- Drainage area (DA), in square miles,
- Channel slope (CSL), in feet per mile,
- Land slope (LANDSL), in feet per feet (this is basin or watershed slope perpendicular to the stream) but used in the regression analysis as a percent to reduce the regression constant,
- Basin Relief (BR), in feet,

- Forest cover (FOR), in percent of the drainage area,
- Percent A, B, C and D SSURGO soils based on the legacy data currently in GISHydroNXT, and
- Percent A, B, C and D SSURGO soils based on soils data downloaded from the NRCS Soil Survey web site in May 2018

The legacy SSURGO soils data in GISHydroNXT are shown in Figure 2 for the four Hydrologic Soil Groups A, B, C and D where A has the highest infiltration and D the lowest infiltration. These data were added to GIShydroNXT over time and were representative of different dates for each county in the two states.

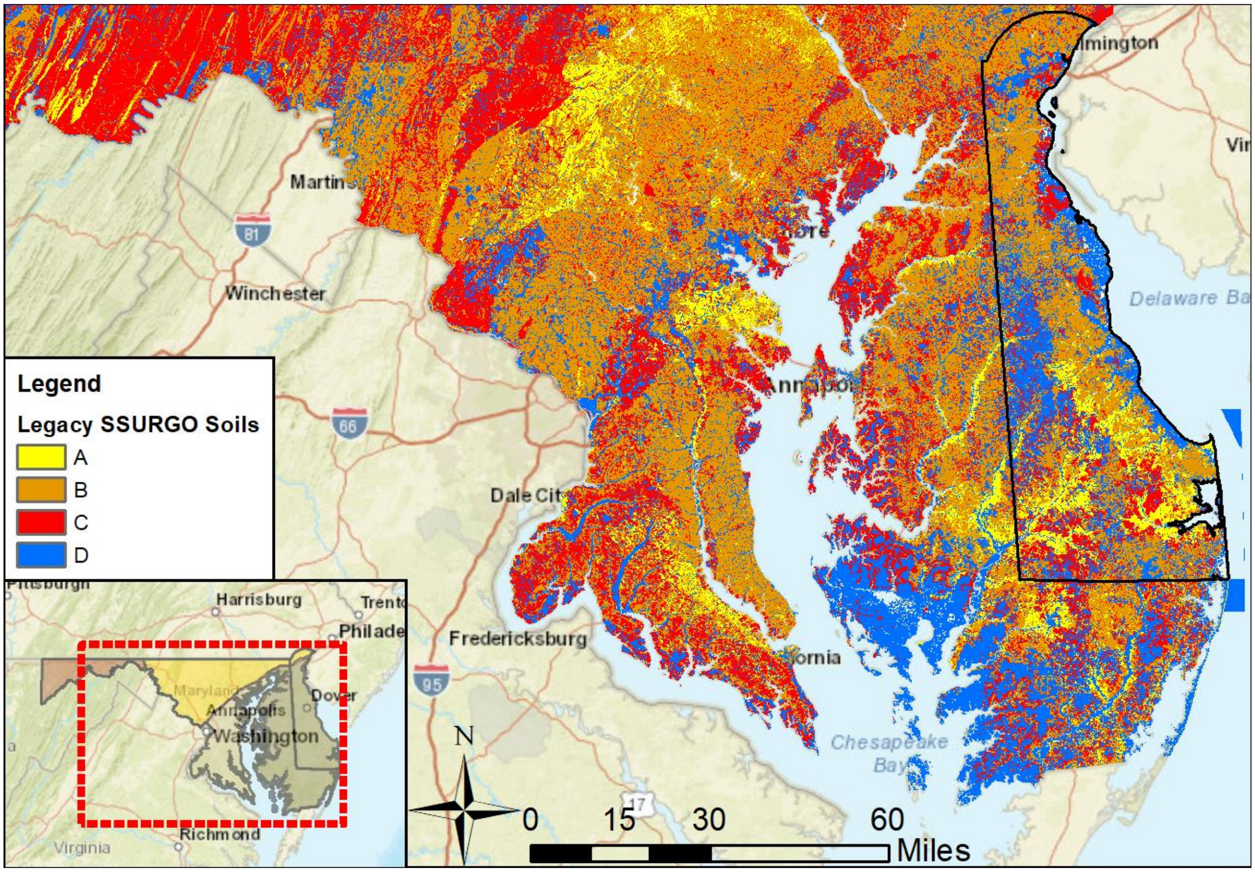


Figure 2. Legacy SSURGO soils data in GISHydroNXT.

The SSURGO soils data downloaded from the NRCS Soil Survey web site in May 2018 are shown in Figure 3. The NRCS procedures for estimating the Hydrologic Soil Groups (HSGs) were updated prior to 2009 and documented in the NRCS Part 630 Hydrology, National Engineering Handbook, Chapter 7, Hydrologic Soils Group (HSG) dated January 2009. The calculations for the new HSGs were completed for Maryland in 2014 and the updated HSGs were posted to the NRCS Soil Survey database in 2016. The new criteria for assigning HSGs use soil properties that influence runoff potential such as:

- Depth to a seasonal high-water table,
- Saturated hydraulic conductivity (Ksat) after prolonged wetting, and
- Depth to a layer with a very slow water transmission rate.

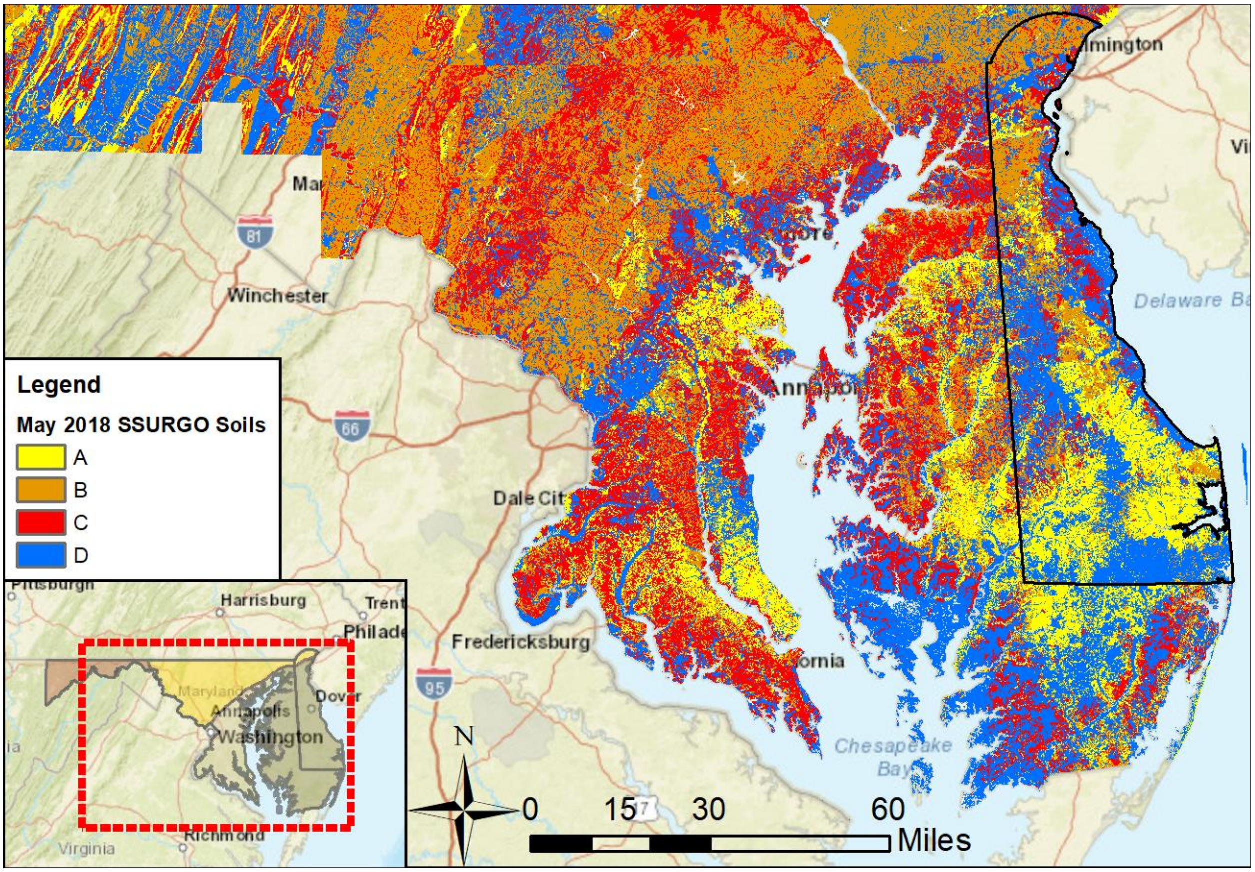


Figure 3. The May 2018 SSURGO soils data in GISHydroNXT.

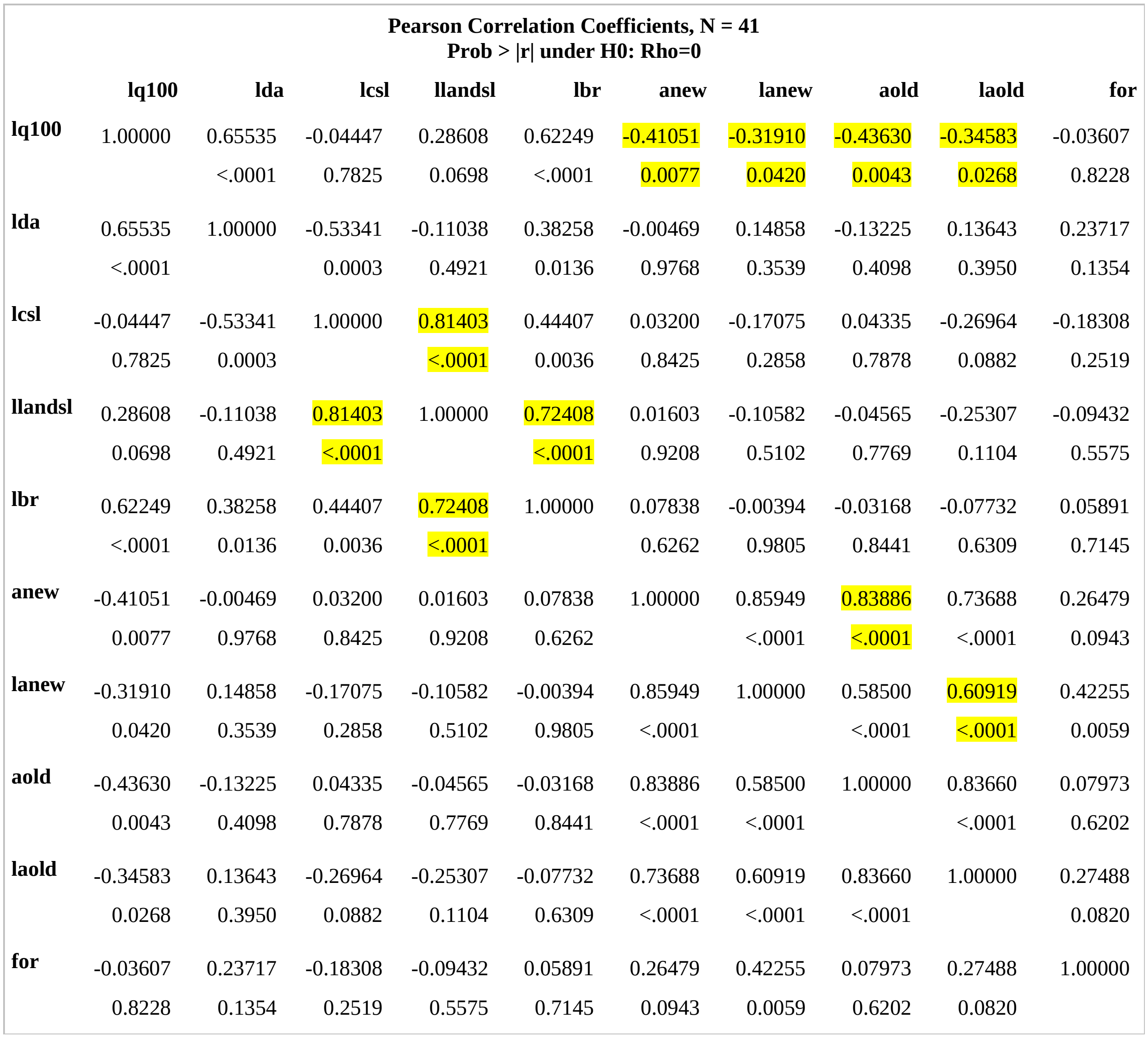
A correlation analysis was performed to determine which explanatory variables were highly correlated. The objective of the regression analysis is to have the explanatory variables as independent as possible. For the soils data, only the A and D soils (extreme values) were evaluated because these hydrologic soil groups were statistically significant in previous analyses.

The correlation analysis was done for the logarithms of the topographic characteristics and flood discharges and the untransformed and the logarithmic transformation for the soils data. The D soil variable was not as statistically significant as the A soil so only the A soil variable is shown in the correlation matrix in Figure 4. The variable Aold is the A soil value from the legacy SSURGO data currently in GISHydroNXT (Figure 2) and Anew is the value based on the May 2018 SSURGO data (Figure 3). A “l” before the variable name implies it is a logarithmic value.

Some pertinent correlations are highlighted in yellow in Figure 4 and include:

- The 100-year discharge lq100 is more correlated with the untransformed A soil than the logarithmic transformed values,
- Drainage area (lda) and channel slope (lcsl) are negatively correlated (-0.533),
- Land slope (llandsl) and basin relief (lbr) are highly correlated (0.724),

- Land slope (llandsl) and channel slope (lcsl) are highly correlated (0.814),
- The 100-year discharge lq100 is more highly correlated with basin relief (lbr) (0.622) than land slope (llandsl) (0.286), and
- Anew and Aold soils are more correlated for the untransformed data (0.839) than the log transformed data (0.609).



Pearson Correlation Coefficients, N = 41
Prob > |r| under H0: Rho=0

	lq100	lda	lcsl	llandsl	lbr	anew	lanew	aold	laold	for
lq100	1.00000	0.65535	-0.04447	0.28608	0.62249	-0.41051	-0.31910	-0.43630	-0.34583	-0.03607
		<.0001	0.7825	0.0698	<.0001	0.0077	0.0420	0.0043	0.0268	0.8228
lda	0.65535	1.00000	-0.53341	-0.11038	0.38258	-0.00469	0.14858	-0.13225	0.13643	0.23717
	<.0001		0.0003	0.4921	0.0136	0.9768	0.3539	0.4098	0.3950	0.1354
lcsl	-0.04447	-0.53341	1.00000	0.81403	0.44407	0.03200	-0.17075	0.04335	-0.26964	-0.18308
	0.7825	0.0003		<.0001	0.0036	0.8425	0.2858	0.7878	0.0882	0.2519
llandsl	0.28608	-0.11038	0.81403	1.00000	0.72408	0.01603	-0.10582	-0.04565	-0.25307	-0.09432
	0.0698	0.4921	<.0001		<.0001	0.9208	0.5102	0.7769	0.1104	0.5575
lbr	0.62249	0.38258	0.44407	0.72408	1.00000	0.07838	-0.00394	-0.03168	-0.07732	0.05891
	<.0001	0.0136	0.0036	<.0001		0.6262	0.9805	0.8441	0.6309	0.7145
anew	-0.41051	-0.00469	0.03200	0.01603	0.07838	1.00000	0.85949	0.83886	0.73688	0.26479
	0.0077	0.9768	0.8425	0.9208	0.6262		<.0001	<.0001	<.0001	0.0943
lanew	-0.31910	0.14858	-0.17075	-0.10582	-0.00394	0.85949	1.00000	0.58500	0.60919	0.42255
	0.0420	0.3539	0.2858	0.5102	0.9805	<.0001		<.0001	<.0001	0.0059
aold	-0.43630	-0.13225	0.04335	-0.04565	-0.03168	0.83886	0.58500	1.00000	0.83660	0.07973
	0.0043	0.4098	0.7878	0.7769	0.8441	<.0001	<.0001		<.0001	0.6202
laold	-0.34583	0.13643	-0.26964	-0.25307	-0.07732	0.73688	0.60919	0.83660	1.00000	0.27488
	0.0268	0.3950	0.0882	0.1104	0.6309	<.0001	<.0001	<.0001		0.0820
for	-0.03607	0.23717	-0.18308	-0.09432	0.05891	0.26479	0.42255	0.07973	0.27488	1.00000
	0.8228	0.1354	0.2519	0.5575	0.7145	0.0943	0.0059	0.6202	0.0820	

Figure 4. Correlation matrix for the 100-year discharge (lq100) and selected watershed characteristics for 41 stations in the Eastern Coastal Plain Region of Maryland.

The correlation matrix is helpful in explaining why variables are statistically significant in the regression analysis. Of two highly correlated explanatory variables, only one will likely be statistically significant in the regression equation since the two variables are explaining the same variability in the discharge variable. For example, one would not expect channel slope (lcsl) and land slope (llandsl) to be statistically significant in the same equation since these variables are highly correlated (0.814).

Development of Regression Equations

Multiple regression analyses were run using the watershed characteristics described earlier and the Statistical Analysis System (SAS) computer software developed by the SAS Institute, Inc., Cary, NC (https://www.sas.com/en_us/company-information.html). A and D soils based on the legacy and May 2018 SSURGO data were used in the regression analysis. For the soil parameters, the percent A soil was more statistically significant than the D soil similar to previous analyses. The Anew soil data (May 2018) provided a lower standard error than Aold soil (legacy) data and was used in the regression equations. The most statistically significant variables used in the final regression equations include: drainage area, in square miles, ranging from 0.91 to 113.8 square miles; A soil, in percent, ranging from 0.2 to 82.3 percent; and land slope, in percent, ranging from 5.44 to 22.0 percent.

Land slope was estimated in feet per feet and then converted to percent to reduce the regression constant to a more reasonable value. **Therefore, the user must input land slope in percent in the regression equations.** The standard error is the same whether ft/ft or percent is used in the regression analysis. The watershed characteristics used in the regression equations are given in Appendix 2 for all 41 stations.

The explanatory variables used in this analysis for the ECP Region are the same as Ries and Dillow (2006) and the 2010 Maryland Hydrology Panel report. Basin relief provides equations with about the same standard error as land slope but land slope was chosen for the equations since it is more independent of drainage area than basin relief (see Figure 4) and has a more uniform range of values.

The following regression equations were based on 36 stations minus five outlier stations: Silver Lake Tributary at Middletown, DE (01483155), Murderkill River Tributary near Felton, DE (01484002), Birch Branch at Sowell, MD (0148471320), Andrews Branch near Delmar, MD (01486100) and Toms Dam Branch near Greensboro, MD (01486980). Murderkill River Tributary had a large flood in a short record and the gaging station estimates were conservatively high. Birch Branch had an indeterminate drainage area and the other three stations had very low annual peaks for the size of the watershed.

The regression equations, standard error of estimate (SE) and Equivalent Years of Record (EY) are given below. As discussed earlier, the Anew variable was not transformed to logarithms so this variable is represented in the equations as an exponent to the base 10. The Equivalent years of record is defined as the number of years of actual streamflow record required to achieve an accuracy equivalent to the standard error of the regression equation. Equivalent years of record are used to weight the gaging station estimates with the regression estimates following the approach described by Dillow (1996) and described in the Maryland Hydrology Panel report (2016). The computation of EY in years is described in Appendix 3.

$Q_{1.25} = 35.6\ DA^{0.757}\ LANDSL^{0.127}\ 10^{-0.00815*Anew}$ SE = 45.6 percent EY = 2.8 (1)

$Q_{1.5} = 48.0\ DA^{0.757}\ LANDSL^{0.202}\ 10^{-0.00871*Anew}$ SE = 43.6 percent EY = 3.0 (2)

$Q_{2} = 67.3\ DA^{0.751}\ LANDSL^{0.281}\ 10^{-0.00919*Anew}$ SE = 41.8 percent EY = 3.3 (3)

$Q_{5} = 134.8\ DA^{0.737}\ LANDSL^{0.473}\ 10^{-0.01027*Anew}$ SE = 39.5 percent EY = 6.9 (4)

$Q_{10} = 200.0\ DA^{0.725}\ LANDSL^{0.605}\ 10^{-0.01091*Anew}$ SE = 38.9 percent EY = 11 (5)

$Q_{25} = 314.5\ DA^{0.707}\ LANDSL^{0.793}\ 10^{-0.01151*Anew}$ SE = 39.0 percent EY = 19 (6)

$Q_{50} = 420.6\ DA^{0.700}\ LANDSL^{0.895}\ 10^{-0.01202*Anew}$ SE = 39.8 percent EY = 19 (7)

$Q_{100} = 551.2\ DA^{0.692}\ LANDSL^{0.991}\ 10^{-0.01249*Anew}$ SE = 41.5 percent EY = 22 (8)

$Q_{200} = 709.7\ DA^{0.684}\ LANDSL^{1.076}\ 10^{-0.01296*Anew}$ SE = 43.8 percent EY = 24 (9)

$Q_{500} = 989.4\ DA^{0.670}\ LANDSL^{1.177}\ 10^{-0.01347*Anew}$ SE = 47.4 percent EY = 25 (10)

For Equations 1-10, the drainage area exponent decreases with an increasing recurrence interval, consistent with earlier results. A possible explanation is that the storm rainfall for the larger storms varies considerably across a watershed and does not have a uniform impact across the entire watershed (that is, the effective drainage area is less). The exponent on land slope (LANDSL) increases as the recurrence interval increases implying this variable becomes more significant as rainfall depth increases over the watershed. The exponent on Anew increases from the 1.25-year flood up to the 500-year flood implying the soils become more significant as rainfall depth increases over the watershed.

The explanatory variables drainage area and Anew are statistically significant for all recurrence intervals at the 5-percent level of significance but land slope is only statistically significant at this level for the 5-year flood and larger. Land slope was maintained in the equations for the 1.25-, 1.5- and 2-year discharges to achieve consistency. The 5-percent level of significance is traditionally used in regression analysis for determining statistical significance because this implies there is less than a 5-percent chance of erroneously including a variable in the equation when it is not really significant.

The higher standard errors for the shorter recurrence interval (1.25- to 5-year) floods imply that explanatory variables other than drainage area, land slope, and percentage of A soils influence these floods. The time-sampling error (error in T-year flood discharge) is less for these smaller floods, so one would expect a lower standard error in the regression analysis. Instead, the standard errors of the regression equations for the smaller events are influenced by the model error, indicating that other important explanatory variables may be missing from the equations.

The regression equations should be unbiased and provide estimates consistent with the gaging station estimates. The 100-year regression estimates from Equation 8 are plotted versus the 100-year gaging station estimates in Figure 5 for the 36 gaging stations used to develop the regression equation. The line shown in Figure 5 is the equal discharge line and the data points should scatter uniformly about this line.

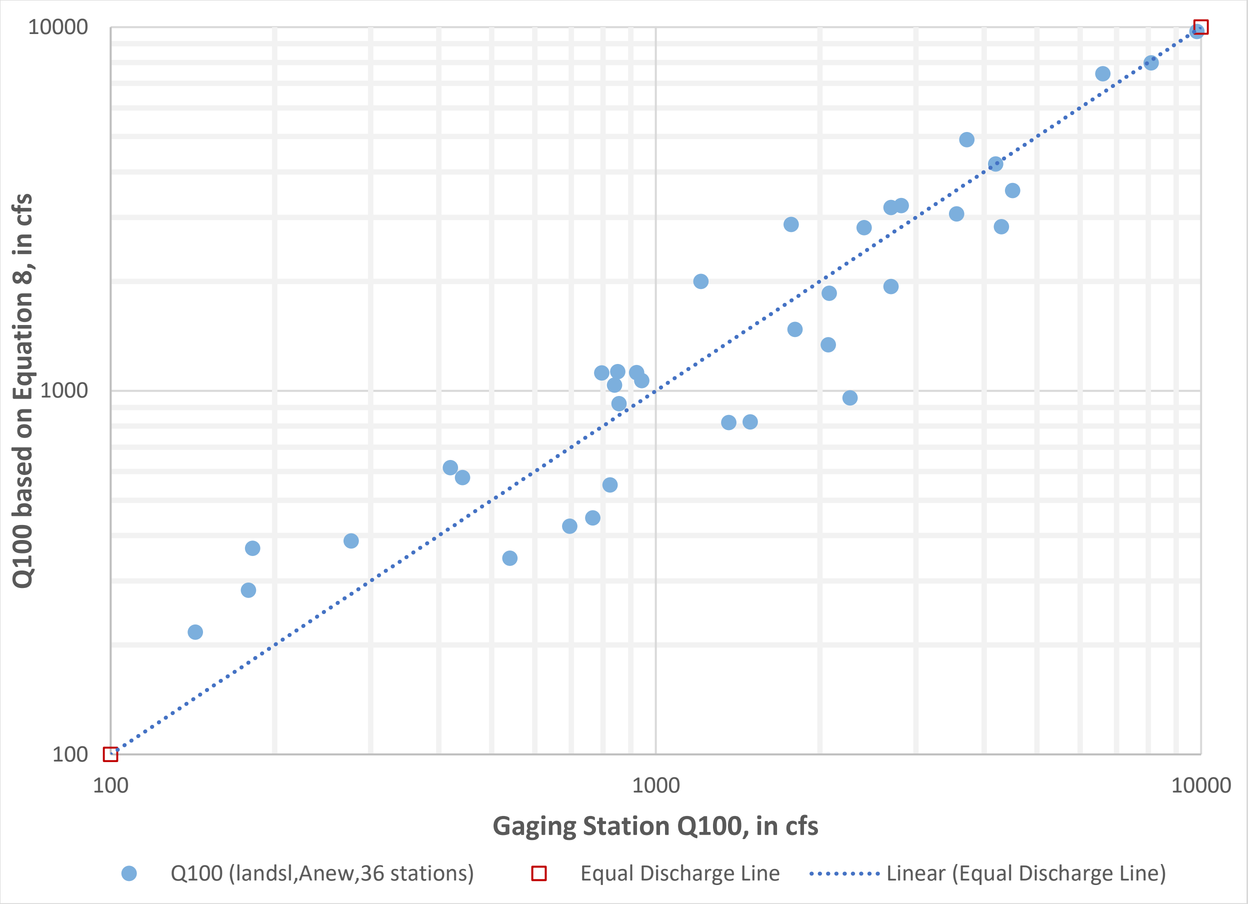


Figure 5. The 100-year regression estimates from Equation 8 plotted versus the 100-year estimates based on the gaging station data for 36 stations in the Eastern Coastal Plain Region.

In Figure 5, the data points to the right of the equal discharge line are stations where the regression equation is underestimating the 100-year discharge and points to the left are indicative of the regression equation overestimating the 100-year discharge based on gaging station data. Although there is considerable scatter in Figure 5, there is no indication of bias in the regression equation except for the four smallest 100-year discharges where all four stations plot to the left of the equal discharge line. Three of these stations are located in the same area near the coast of Delaware in a high A soil area and are:

- Beaverdam Branch at Houston, DE (01484100), site 9 in Figure 1, Anew = 30.3 percent,
- Beaverdam Creek near Milton, DE (01484270), site 10 in Figure 1, Anew = 73.3 percent, and
- Sowbridge Branch near Milton, DE (01484300), site 11 in Figure 1, A new = 82.3 percent.

The magnitude of the deviation from the equal discharge line is consistent with other stations so these three stations are not extreme outliers. Even though Anew is large for these stations, the regression equation still overestimates the 100-year discharge. As shown in Figure 1, the three stations are in close proximity so there must be some other explanatory variable not in the regression equation that is impacting these watersheds. The variable Anew soil is still statistically significant in the 100-year equation even if the three stations above are omitted from the analysis. The decision was to keep these stations in the regression analysis since they reflect the impact of A soil in a given watershed.

The same analysis was performed for the 10-year flood with the 10-year regression estimates from Equation 5 plotted versus the 10-year gaging station estimates as shown in Figure 6. The three largest 10-year discharges plot to the right of the equal discharge line indicating the regression equation is underestimating the 10-year flood in comparison to the gaging station data. However, the departures from equal discharge line are small and Equation 5 is considered unbiased.

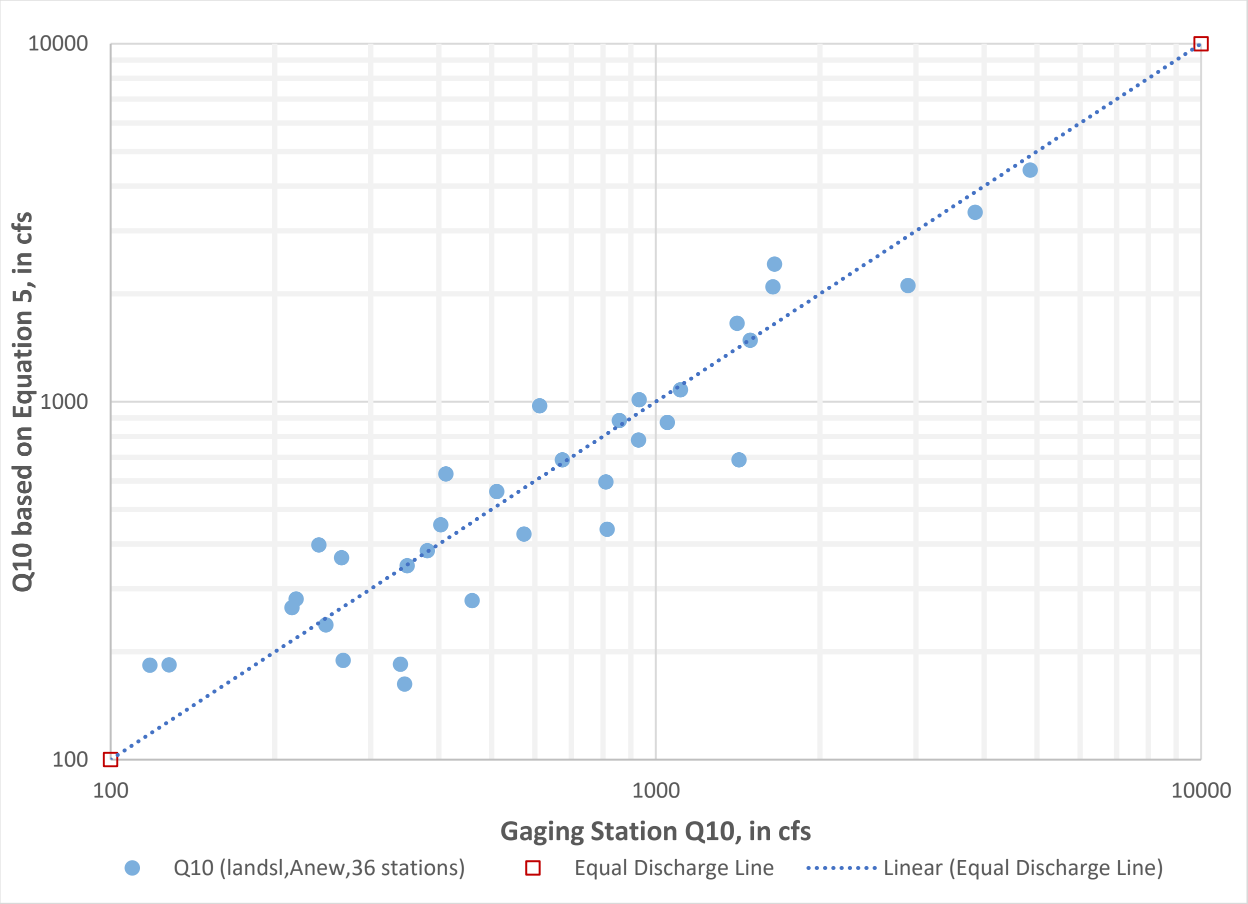


Figure 6. The 10-year regression estimates from Equation 5 plotted versus the 10-year estimates based on the gaging station data for 36 stations in the Eastern Coastal Plain Region.

The 100-year regression estimates for the 2019 analysis (Equation 8) were also compared to the 100-year regression estimates for the equations developed in 2010 and currently in use and documented in the July 2016 version of the Maryland Hydrology Panel report. The data are compared in Figure 7 for all 41 stations where the trend line is the equal discharge line.

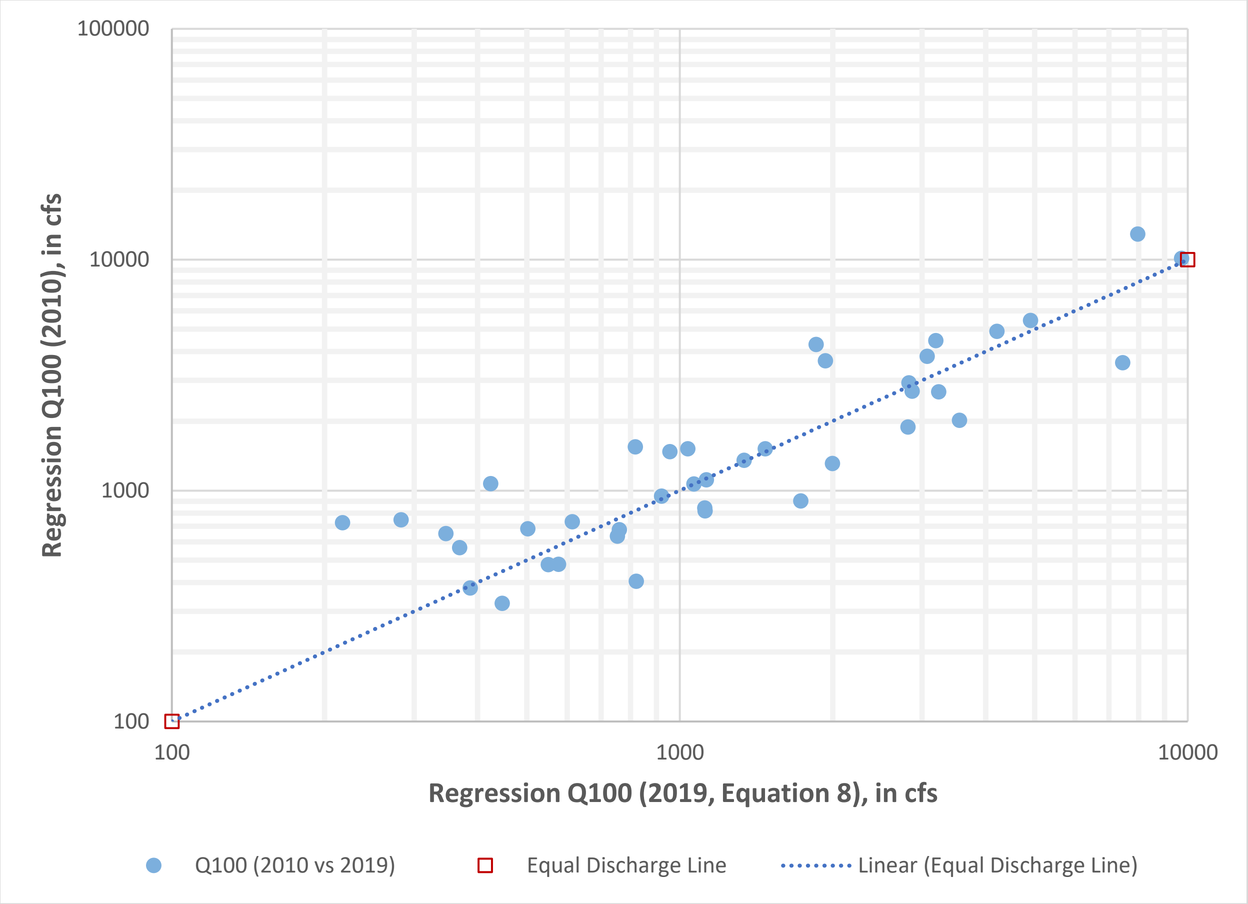


Figure 7. The 100-year regression estimates from the 2010 analysis versus the 100-year regression estimates based on the 2019 analysis (Equation 8) for 41 stations in the Eastern Coastal Plain Region.

As shown in Figure 7, there is a tendency for the 2010 equations for the 100-year discharge to predict slightly higher than the 2019 equation (Equation 8). On average, the 2010 equation is predicting a 100-year discharge about 10 percent higher than the 2019 equation (Equation 8). The largest differences between the 2010 and 2019 equations were evaluated and most of the time the 2019 equation gave estimates closest to the updated gaging station estimate. The 2010 and 2019 equations are based on the same variables (drainage area, land slope and A soil) but the 2019 estimates of drainage area and land slope are based on an updated DEM and the A soil data are from the May 2018 SSURGO data. Based on a comparison to updated gaging station data, the 2019 equation is considered more accurate for estimating the 100-year discharge.

The 10-year regression estimates for the 2019 analysis (Equation 5) were also compared to the 10-year regression estimates for the equations developed in 2010 and currently in use and documented in the July 2016 version of the Maryland Hydrology Panel report. The data are compared in Figure 8 for all 41 stations where the trend line is the equal discharge line.

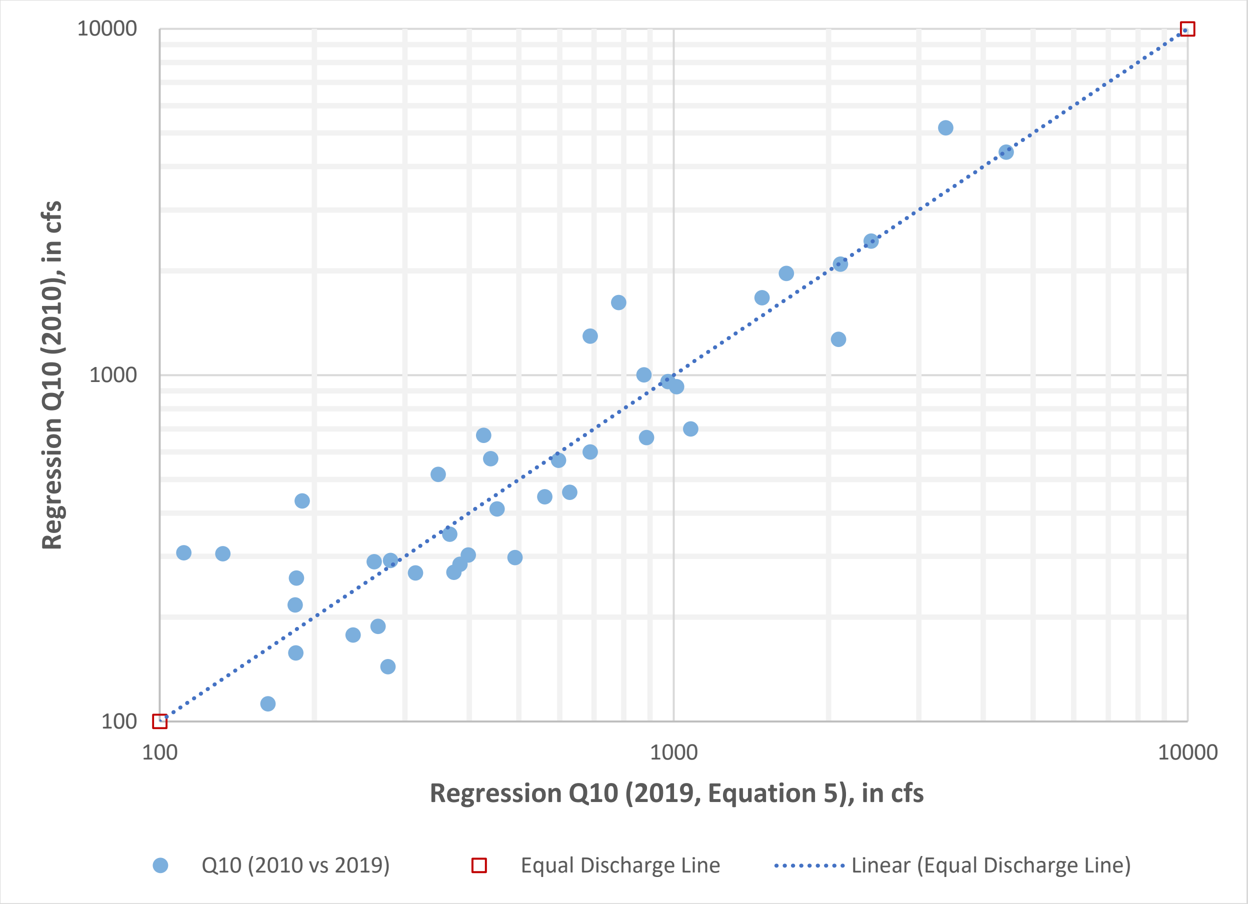


Figure 8. The 10-year regression estimates from the 2010 analysis versus the 10-year regression estimates based on the 2019 analysis (Equation 5) for 41 stations in the Eastern Coastal Plain Region.

As with the 100-year comparison, the 2010 equation for the 10-year discharge is providing slightly higher estimates of the 10-year discharge than the 2019 analysis (Equation 5). On average, the 2010 estimates of the 10-year discharge are seven percent higher than the 2019 estimates. The largest differences between the 2010 and 2019 equations were evaluated and most of the time the 2019 equation gave estimates closest to the gaging station estimate. The 2010 and 2019 equations are based on the same variables (drainage area, land slope and A soil) but the 2019 estimates of drainage area and land slope are based on an updated DEM and the A soil data are from the May 2018 SSURGO data. Based on a comparison to gaging station data, the 2019 equation is considered more accurate for estimating the 10-year discharge.

Summary and Conclusions

The Fixed Region regression equations for the Eastern Coastal Plain Region were updated using annual peak flow data through the 2017 water year. The updated flood discharges were based on Bulletin 17C (England and others, 2018). The regression equations were based on 36 gaging stations in Maryland and Delaware: 19 gaging stations in Maryland and 17 gaging stations in Delaware; 16 active stations and 19 discontinued stations as of 2017. Five gaging stations were considered outliers primarily because the annual peak flows were very low for the drainage area size and were not included in the regression analysis. The most statistically significant explanatory variables were drainage area, in square miles; land slope, in percent; and A soil data, in percent, based on the SSURGO data dated May 2018 from the NRCS Soil Survey web site. The legacy SSURGO data in GISHydroNXT and the May 2018 SSURGO data were both evaluated in the regression analysis to determine which set of soils data provided the most accurate regression equations. The May 2018 SSURGO data provided the most accurate regression equations and is now the default soils data in GISHydroNXT.

There were nine stations that had 62 to 74 years of record and seven of these stations had statistically significant upward trends due to large floods near the end of the record in 1989, 1999, 2011 and 2016. The impervious area of the watersheds was less than 10 percent of the drainage area so the upward trends were not related to urbanization. The upward trends were assumed to be climatic persistence or variability (not a permanent change in climate) and the entire period of record was used in the frequency analysis.

The 2019 regression estimates for the 100- and 10-year flood discharges were compared to the respective gaging station estimates and shown to be reasonably unbiased. The 2019 regression estimates were also compared to the respective regression estimates from the 2010 analysis, regression equations that are currently in use and documented in the July 2016 version of the Maryland Hydrology Panel report. The 2010 equations provided 100-year discharges that are about 10 percent higher, on average, than the 2019 equations. For the 10-year discharges, the 2010 equations provided estimates about seven percent higher, on average, than the 2019 estimates.

The updated regression equations will be used by MDOT SHA in the design of bridges and culverts in Maryland. The updated regression equations will be included in the fifth version of the Maryland Hydrology Panel report entitled “Application of Hydrologic Methods in Maryland” that will be published in 2020.

References

Carpenter, D.H., 1980, ***Technique for estimating magnitude and frequency of floods in Maryland***: U.S. Geological Survey Water-Resources Investigations Open-File Report 80-1016, 79 p.

Dillow, J.J.A., 1996, ***Technique for estimating magnitude and frequency of peak flows in Maryland***: U.S. Geological Survey Water-Resources Investigations Report 95-4154, 55 p.

England, J.F., Jr., Cohn, T.A., Faber, B.A., Stedinger, J.R., Thomas, W.O. Jr., Veilleux, A.G., Kiang, J.E., and Mason, R.R., Jr., 2018. ***Guidelines for Determining Flood Flow Frequency – Bulletin 17C***: U.S. Geological Survey Techniques and Methods, Book 4, Chapter B5, 148 p.

Hardison, C.H., 1971, ***Prediction error of regression estimates of streamflow characteristics at ungaged sites***: U.S. Geological Survey Professional Paper 750-C, pp. C228-C236.

Helsel, D.R., and Hirsch, R.M., 2002. ***Statistical Methods in Water Resources***: Techniques of Water Resources Investigations of the U.S. Geological Survey, Book 4, Hydrologic Analysis and Interpretation, Chapter A3, 503 p.

Kilgore, R., Herrmann, G.R., Thomas, W.O., Jr., and Thompson, D.B., 2016. ***Highways in the River Environment – Floodplains, Extreme Events, Risk and Resilience***. Hydraulic Engineering Circular No. 17 (HEC-17), Federal Highway Administration, FHWS-HIF-16-018.

Maryland Hydrology Panel, 2006. ***Application of Hydrologic Methods in Maryland – Second Edition***: A report prepared by the Hydrology Panel for the Maryland State Highway Administration and Maryland Department of the Environment, August 2006.

Maryland Hydrology Panel, 2010. ***Application of Hydrologic Methods in Maryland -Third Edition***: A report prepared by the Hydrology Panel for the Maryland State Highway Administration and Maryland Department of the Environment, September 2010.

Maryland Hydrology Panel, 2016. ***Application of Hydrologic Methods in Maryland -Fourth Edition***: A report prepared by the Hydrology Panel for the Maryland State Highway Administration and Maryland Department of the Environment, July 2016.

Moglen, G.E., Thomas, W.O., Jr., and Cuneo, Carlos, 2006. ***Evaluation of alternative statistical methods for estimating frequency of peak flows in Maryland***: Final report (SP907C4B), Maryland State Highway Administration, Baltimore, Maryland.

Natural Resources Conservation Service, 2009. ***Chapter 7 Hydrologic Soil Groups***. Part 630 Hydrology National Engineering Handbook, U.S. Department of Agriculture, January 2009, 10 p.

Ries, K. G., III, and Dillow, J.J.A., 2006. ***Magnitude and Frequency of Floods on Nontidal Streams in Delaware:*** U.S. Geological Survey Scientific Investigations Report 2006-5146, 59 p.

Stedinger, J.R., Vogel, R.M., and Foufoula-Georgiou, E., 1993, Chapter 18 ***Frequency Analysis of Extreme Events***: *in* Handbook of Hydrology, David Maidment, Editor in Chief, McGraw-Hill, Inc.

Appendix 1. Flood discharges for the 1.25-, 1.5-, 2-, 5-, 10-, 25- 50-, 100-, 200- and 500-year events (in cubic feet per second) for 41 gaging stations on the Eastern Coastal Plain of Maryland.

Map No.	Station No	Stream name	DA (mi^2)	Q1.25	Q1.5	Q2	Q5	Q10	Q25	Q50	Q100	Q200	Q500
1	01483155	Silver Lake Tributary at Middletown, DE	2.03	58	74	96	169	232	333	424	532	658	857
2	01483200	Blackbird Creek at Blackbird, DE	4.06	90	120	160	240	350	680	760	840	900	980
3	01483290	Paw Paw Branch Tributary near Clayton, DE	0.91	93	116	149	255	346	489	617	766	940	1210
4	01483500	Leipsic River near Cheswold, DE	9.21	120	156	211	412	612	964	1320	1770	2340	3340
5	01483720	Puncheon Branch at Dover, DE	2.41	77	103	141	268	381	562	727	922	1150	1510
6	01484000	Murderkill River near Felton, DE	12.64	137	191	271	552	810	1230	1620	2070	2610	3470
7	01484002	Murderkill River Trib near Felton, DE	0.96	11	15	22	51	81	137	196	273	372	550
8	01484050	Pratt Branch near Felton, DE	3.1	35	49	70	153	241	404	573	796	1090	1600
9	01484100	Beaverdam Branch at Houston, DE	3.31	34	43	56	95	128	179	224	276	335	428
10	01484270	Beaverdam Creek near Milton, DE	6.21	25	31	39	65	86	118	146	179	216	274
11	01484300	Sowbridge Branch near Milton, DE	7.45	25	29	36	56	72	96	118	143	171	215
12	01484500	Stockley Branch at Stockley, DE	4.8	76	88	111	171	219	290	352	420	497	859
13	01484550	Pepper Creek at Dagsboro, DE	8.31	166	204	254	400	511	669	800	942	1100	1320
14	01484695	Beaverdam Ditch near Millville, DE	2.71	57	72	94	160	215	295	365	442	529	660
15	0148471320	Birch Branch at Sowell, MD	6.38	418	495	598	887	1110	1420	1680	1960	2270	2720
16	01484719	Bassett Creek near Ironshire, MD	1.39	66	92	133	293	460	765	1080	1490	2020	2960
17	01485000	Pocomoke River near Willards, MD	51.61	494	589	717	1100	1410	1860	2260	2700	3200	3960
18	01485500	Nassawango Creek near Snow Hill, MD	45.47	362	459	600	1070	1490	2170	2800	3560	4470	5940
19	01486000	Manokin Branch near Princess Anne, MD	3.98	75	100	145	270	340	425	480	540	590	660
20	01486100	Andrews Branch near Delmar, MD	4.54	64	78	95	143	179	230	271	316	363	432
21	01486980	Toms Dam Branch near Greensboro, MD	5.97	25	31	38	59	75	98	116	136	157	188
22	01487000	Nanticoke River near Bridgeville, DE	71.99	391	506	670	1200	1650	2360	2990	3720	4570	5880
23	01487900	Meadow Branch near Delmar, DE	2.73	52	61	72	99	118	143	162	182	202	229
24	01488500	Marshyhope Creek near Adamsville, DE	46.47	970	1300	1650	2400	2900	3400	3800	4200	4500	5000
25	01489000	Faulkner Branch at Federalsburg, MD	8.06	108	159	238	532	814	1290	1730	2270	2910	3940
26	01490000	Chicamacomico River near Salem, MD	16.96	123	162	219	407	573	834	1080	1360	1700	2220
27	01490600	Meredith Branch near Sandtown, DE	8.76	136	178	241	464	674	1030	1380	1800	2330	3210
28	01490800	Oldtown Branch at Goldsboro, MD	4.45	111	139	178	301	403	558	695	851	1030	1300
29	01491000	Choptank River near Greensboro, MD	113.8	1090	1460	1990	3590	4860	6680	8190	9820	11600	14100
30	01491010	Sangston Prong near Whiteleysburg, DE	1.94	34	48	70	157	248	418	594	824	1120	1650
31	01491050	Spring Branch near Greensboro, MD	3.76	38	53	77	169	265	441	622	856	1160	1690
32	01491500	Tuckahoe Creek near Ruthaburg, MD	87.67	1100	1370	1740	2890	3850	5320	6620	8100	9800	12400
33	01492000	Beaverdam Branch at Matthews, MD	6.05	156	209	289	577	857	1340	1810	2410	3140	4390
34	01492050	Gravel Run at Beulah, MD	8.53	56	73	98	186	267	404	535	695	889	1210
35	01492500	Sallie Harris Creek near Carmichael, MD	8	133	188	274	602	932	1510	2090	2820	3730	5280
36	01492550	Mill Creek near Skipton, MD	4.24	70	94	132	273	412	657	901	1210	1600	2260
37	01493000	Unicorn Branch near Millington, MD	20.67	198	264	360	668	929	1330	1680	2080	2530	3220
38	01493112	Chesterville Branch near Crumpton, MD	6.14	99	151	241	640	1110	2040	3080	4510	6450	10100
39	01493500	Morgan Creek near Kennedyville, MD	12.73	192	272	405	976	1640	2980	4490	6600	9540	15200
40	01494000	Southeast Creek at Church Hill, MD	12.6	400	500	640	1110	1420	1850	2250	2700	3100	3800
41	01494150	Three Bridges Branch at Centerville, MD	8.24	100	155	250	640	1050	2000	3000	4300	5800	8800

Appendix 2. Watershed characteristics for 41 gaging stations in the Eastern Coastal Plain Region of Maryland. Asoil used in the regression equations is based on the May 2018 SSURGO data for the NRCS Soil Survey web site. Land slope was estimated in ft/ft and converted to percent for use in the regression analysis.

Map No.	Station No.	Period of record	Years of record	DA (mi^2)	LANDSL (ft/ft)	LANDSL (%)	Asoil (%)
1	01483155	2001-2016	16	2.03	0.02045	2.045	1.9
2	01483200	1952-2017	65	4.06	0.01898	1.898	33.8
3	01483290	1966-1975	10	0.91	0.01053	1.053	6.8
4	01483500	1943-1975, 2017	34	9.21	0.0161	1.61	12.5
5	01483720	1966-1975	10	2.41	0.01334	1.334	6.4
6	01484000	1932-33, 1960-99, 2007-09, 2017	35	12.64	0.00949	0.949	28.4
7	01484002	1966-1975	10	0.96	0.01201	1.201	86.2
8	01484050	1966-1975	10	3.1	0.01292	1.292	11.4
9	01484100	1958-2017	60	3.31	0.0073	0.73	30.3
10	01484270	1966-1980, 2002-2005	19	6.21	0.01195	1.195	73.3
11	01484300	1957-1980	22	7.45	0.01045	1.045	82.3
12	01484500	1943-2004	62	4.8	0.00805	0.805	26.5
13	01484550	1960-1975	16	8.31	0.00463	0.463	1.5
14	01484695	1999-2017	19	2.71	0.0062	0.62	5.9
15	0148471320	2000-2017	18	6.38	0.00619	0.619	31.3
16	01484719	2003-2009, 2011-2013	10	1.39	0.01248	1.248	1.7
17	01485000	1950-2004, 2007-2017	66	51.61	0.00667	0.667	19.9
18	01485500	1950-2017	68	45.47	0.00841	0.841	26.2
19	01486000	1951-1971, 1975-2017	64	3.98	0.00544	0.544	28.4
20	01486100	1967-1976	10	4.54	0.01044	1.044	26.7
21	01486980	1966-1975	10	5.97	0.00593	0.593	14.1
22	01487000	1935, 1943-2017	75	71.99	0.00768	0.768	17.8
23	01487900	1967-1975	9	2.73	0.00575	0.575	19.1
24	01488500	1972-2017	45	46.47	0.00636	0.636	6.1
25	01489000	1950-1991, 2011	42	8.06	0.00805	0.805	23.6
26	01490000	1951-1980, 2001-2017	46	16.96	0.00757	0.757	44.8
27	01490600	1966-1975	10	8.76	0.00643	0.643	2.8
28	01490800	1967-1976	10	4.45	0.00951	0.951	9.3
29	01491000	1948-2017	71	113.8	0.00922	0.922	11.3
30	01491010	1966-1975	10	1.94	0.00699	0.699	3.6
31	01491050	1967-1976	10	3.76	0.01008	1.008	14.3
32	01491500	1952-1956, 2001-2017	22	87.67	0.01189	1.189	20.7
33	01492000	1950-1981, 2010-2011	34	6.05	0.01794	1.794	6.8
34	01492050	1966-1976	11	8.53	0.01385	1.385	71.9
35	01492500	1952-1981, 2001-2017	47	8	0.01948	1.948	11.5
36	01492550	1966-1976	11	4.24	0.01814	1.814	10.5
37	01493000	1948-2005, 2007-2017	69	20.67	0.0127	1.27	38.9
38	01493112	1997-2017	13	6.14	0.01857	1.857	0.2
39	01493500	1951-2005, 2007-2017	66	12.73	0.02445	2.445	1.5
40	01494000	1952-1965	14	12.6	0.01893	1.893	39.3
41	01494150	2007-2017	11	8.24	0.022	2.2	21.1

Appendix 3. Computation of the Equivalent Years of Record for Regression Equations for the Eastern Coastal Plain Region.

Computational Procedure

The variance (standard error squared (SE^2)) of the x-year flood at a gaging station is estimated as

$$SE_x^2 = (S^2/N) * R_x^2 \qquad \text{(A3-1)}$$

where S is the standard deviation of the logarithms (log units) of the annual peak discharges at the gaging station, N is the actual record length in years and R_x is a function of recurrence interval x and skew (G) at the gaging station. The standard error increases as the recurrence interval increases, given the same record length.

In Equation A3-1, the standard error of the x-year flood at a gaging station is inversely related to record length N and directly related to the variability of annual peak flows represented by S (standard deviation) and G (skew). If the standard error of the x-year flood is interchanged with the standard error of estimate (SE) of the regression equation, then Equation A3-1 can be used to estimate the years of record needed to obtain that standard error of estimate. Rearranging Equation A3-1 and solving for N gives Equation A3-2 below.

The equivalent years of record of the regression estimate is defined as the number of years of actual streamflow record required at a site to achieve an accuracy equivalent to the standard error of the regional regression equation. Equivalent years of record is used to weight the gaging station and regression estimates. The equivalent years of record (N_r) of a regression equation is computed as follows (Hardison, 1971):

$$N_r = (S/SE)^2 * R^2 \qquad \text{(A3-2)}$$

where S is an estimate of the standard deviation of the logarithms of the annual peak discharges at the ungaged site, SE is the standard error of estimate of the regional regression estimates in logarithmic units, and R^2 is a function of recurrence interval and skew and is computed as (Stedinger and others, 1993):

$$R^2 = 1 + G*K_x + 0.5 *(1+0.75*G^2)*K_x^2 \qquad \text{(A3-3)}$$

where G is an estimate of the average skew for a given hydrologic region, and K_x is the Pearson Type III frequency factor for the x-year flood and skew G.

Computational Details

The equivalent years of record are estimated for the regional regression equations and using Equations A3-2 and A3-3 and an estimate of the average standard deviation and average skew for all gaging stations in a given region. For the Western Coastal Plain Region, the average standard deviation (S) is 0.3104 log units and the average skew (G) is 0.330.

Recurrence Interval (years)	K_x value	SE^2 (log units squared)	Equivalent years of record
1.25	-0.853519	0.03564	2.8
1.50			(3.0) Estimated
2	-0.054904	0.03042	3.3
5	0.821553	0.02732	6.9
10	1.311565	0.02618	11
25	2.18039	0.02670	19
50	2.225966	0.02774	19
100	2.565564	0.02995	22
200	2.881452	0.03311	24
500	3.280295	0.03821	25