In Reply Refer To:

Mail Stop 415

June 3, 2010

OFFICE OF SURFACE WATER TECHNICAL MEMORANDUM 2010.05

Subject: Weighted estimates of peak flow frequency statistics

To obtain a better estimate of a peak flow statistic, such as the one-percent annual exceedance probability (AEP) flow at a streamgage, the at-site estimate can be combined with a regional estimate to obtain a weighted estimate of the flow statistic at the streamgage, as recommended in previous memorandums[1]. This memorandum updates previous recommendations by requiring that weighted estimates be calculated for unregulated rural peak flow analyses and by specifying the use of variances to assign weights to the at-site and regional estimates. For unregulated rural peak flows, weighting by variance must be used for regionalization reports, individual data requests, and within the National Streamflow Statistics Program (NSS) and StreamStats. This policy is consistent with the recommendations in Bulletin 17B for peak flows (Interagency Advisory Committee on Water Data, 1982). The purpose of this memorandum is to describe current policy for how to calculate the weighted estimate of the peak flow statistic, and to provide an update on software available to assist with this calculation.

The at-site estimate of the peak flow statistic is calculated using the available records of peak flows at the streamgage, typically following the methods described in Bulletin 17B. The regional estimate is calculated based on regional information from many streamgages in the area, usually employing multiple linear regression techniques. If proper weights can be determined, then a weighted average of these two estimates provides a better overall estimate of the flow statistic than either one individually. The optimal weights for this purpose are inversely proportional to the variance of each estimate.

Background

In the past, the weights for each peak flow statistic estimate have often been approximated based on the number of years of record used to determine the at-site frequency estimate, and the equivalent years of record estimated for the regression equation. However, the length of record and equivalent years of record often fail to account for the true variance of their respective peak

[1] Use of weighted estimates was initially described as an option in SW73.16, recommended in SW76.07, and reiterated in SW89.12.

flow statistic estimates. For example, the number of years of record fails to account for the improvement in information content provided by the regional skew. Furthermore, flood probability distributions computed from two different streamgage records of the same record length may not be of equal reliability owing to differences in the underlying variances of their flow records. For example, smaller drainages may have flashier, more highly varied records, or may be more difficult to accurately gage than a larger basin; hence estimates for these smaller drainages could be expected to have larger variances. The equivalent years of record of a regression model is used to evaluate the regression model in a similar framework as the number of years of record, but is itself an approximation.

True optimal weights are inversely proportional to the variance of each peak flow statistic estimate. The variance can be thought of as a measure of the uncertainty of each estimate. When the variance corresponding to one of the estimates is high, the uncertainty is also high, and so the optimal weight for that estimate is relatively small. Conversely, when the variance is low, the uncertainty is also low and so the optimal weight is correspondingly large.

For both frequency analysis and regression, the peak flow statistic is commonly transformed using base-10 logarithms:

$$X = \log_{10} Q$$

Where Q is the peak flow statistic in cubic feet per second (cfs) and X is the log-transformed variable. All subsequent operations are performed on the transformed variable, X. As described in Appendix 8 of Bulletin 17B (with slightly different notation), the weighted estimate can be calculated using variances as:

$$X_{weighted} = \frac{X_{Site} * V_{Reg} + X_{Reg} * V_{Site}}{V_{Site} + V_{Reg}} \qquad \text{(all } X \text{ and } V \text{ in log-10 units)}$$

Where $X_{weighted}$ is the weighted estimate; X_{Site} is the at-site estimate; X_{Reg} is the regression estimate; V_{Site} is the variance of the at-site estimate; and V_{Reg} is the variance of the regression estimate. For independent X_{Site} and X_{Reg}, the variance of the weighted estimate can be calculated as:

$$V_{weighted} = \frac{V_{Site} * V_{Reg}}{V_{Site} + V_{Reg}} \qquad \text{(all } V \text{ in log-10 units)}$$

According to Bulletin 17B, X_{Site} and X_{Reg} can generally be considered independent, so the above equation usually gives acceptable estimates of the variance of the weighted estimate. Exceptions may occur for regional regression equations based on clusters of streamgages in close proximity or with uniformly short periods of record, as may be the case for urban regression equations. In

these cases, the at-site and regression estimates of the flow statistic may not be independent and the variance of a weighted estimate will be larger than that calculated by the above equation.

Tools for computing weighted flow statistics

One difficulty in applying the equations above has been in obtaining an estimate of the variance of the at-site estimate, V_{Site}. PeakFQ is being updated so that it will output the variance of its estimates of flow statistics. NSS is being updated to perform the weighting procedure, using appropriate inputs from the user.

In the interim, the program Weighted Independent Estimates (WIE) was developed by the Office of Surface Water to implement the process described above. WIE performs the following tasks:

A. calculates the variance of the at-site estimate based on user inputs;
B. calculates the variance of the regression estimate as either;
    1. variance calculated for a specific streamgage, requiring user-input, or
    2. an average variance for the regression model, as input by the user; and
C. calculates the weighted estimate, $X_{weighted}$, using (A) and (B) and outputs the result, along with the variance of the weighted estimate, $V_{weighted}$.

The WIE program is available for download on-line at: http://water.usgs.gov/usgs/osw/FreqReg/ Complete instructions for the use of WIE are available at the above website also. Information on updated versions of PeakFQ and NSS will be distributed as the programs become available.

WIE and PeakFQ are designed specifically for use with annual peak flow estimates, and weighting by variance is required for unregulated rural peak flow analyses. This policy supersedes any in previous memorandums. The use of weighted estimates is not standard practice for low flow statistics, but the equations provided above can be used for low flow analyses as well, provided appropriate estimates of the variance of the at-site estimate are available. These will be provided in future versions of the program SWSTAT (http://water.usgs.gov/software/SWSTAT/). It is less clear whether or not weighting of estimates is useful for urban sites and urban regression equations.

For questions about calculating weighted estimates, please contact Julie Kiang (jkiang@usgs.gov, 703-648-5364). For questions about the use of the WIE program, please contact Charles Berenbrock (ceberenb@usgs.gov, 703-648-6876).

Reference

Interagency Advisory Committee on Water Data, 1982, Guidelines for determining flood flow frequency, Bulletin 17B of the Hydrology Subcommittee, Reston, Virginia, U.S. Geological Survey, Office of Water Data Coordination, 183p.