

Appendix A. DATA QUALIFICATION CODES FOR USGS PEAK FLOW FILE

The Peak Flow File allows data qualification codes to be stored along with the basic numerical data on dates, gage heights, and discharges. These codes presently are used by the flood-frequency program (PEAKFQ) to control the processing of records containing historic information and other data requiring special interpretation. Proper use of these codes can enhance the accuracy of statistical analysis output from PEAKFQ. Conversely, improper use or inconsistent application and interpretation of these codes can result in significant errors in the analysis and interpretation of the output from PEAKFQ. The codes therefore must have unambiguous definitions that promote consistent use and interpretation by data providers and data analysts.

Appropriate qualification codes should be assigned to any observations that otherwise would result in a misleading impression of flood risk. Because they complicate the course of analysis, however, *codes should not be assigned unless they are necessary to avoid misinterpretation of the data.*

In addition to the constraints described in the explanations that follow, there are a few simple rules on the concurrent use of peak discharge qualification codes. Peak discharge values may be qualified with from none to twelve codes. Code values of 1, 2, 3, 7, and 8 cannot be used with code 4. Codes 5 and 6 cannot be used at the same time.

The User's Manual for the National Water Information System of the U.S. Geological Survey (Version 3_1, Open-file Report 97-635) provides much-abbreviated definitions of the qualification codes in the instructions for input to the file (Chapter 3 (ADAPS), section 12.8). Documentation for this manual can be found at http://wwwnwis.er.usgs.gov/conversion/nwisdocs3_2/adaps/adaps.book.html. Although most of these definitions seem self-explanatory, experience has shown that some of them are subject to different interpretations by different readers. The instructions for preparation of water-resources data reports (Novak, 1985, p. 96-102) offer general advice on qualification of published peak-flow data. The following paragraphs supplement the above references with specific suggestions for the proper assignment and interpretation of qualification codes in the Peak Flow File of the U.S. Geological Survey.

While the expanded definitions are intended to promote more consistent National usage of the codes, some WSCS have evolved internally consistent, long-term practices that may differ from these definitions. Except for codes 3, 4, 7, 8, and C (which have a direct effect on the handling of individual peaks by PeakFQ) and codes A, B, and C (which are well defined), WSCs that have such established practices, may continue to assign codes based on those practices pending future guidance. However, these practices and definitions should be well documented in internal memorandum and published flood-frequency regionalization reports.

Explanation of Discharge Codes

Code 1. –Discharge is a maximum daily average – Use code 1 only if the instantaneous value cannot be estimated and is believed to be significantly larger than the daily mean. A code 1 implies that the peak estimate is likely biased low. Do not

use code 1 if the daily mean is approximately equal to the instantaneous peak, as may be the case for large rivers or basins with little topographic relief. If at all possible, enter an estimate of instantaneous peak discharge, qualified by code 2, if necessary, in lieu of a daily average. Peaks qualified with a code 1 are included in PEAKFQ frequency analyses.

Code 2. –Discharge is an estimate – Use this code to characterize peak flow values that are generally less accurate or less reliable than other values in the database. This might include indirect measurements rated less than fair, larger-than-recommended extrapolations of the rating curve, or a routing/modeling technique not involving measured stage-discharge relations at the site. The number should still be defensible and reasonable in light of available field evidence and current comprehension of flow hydraulics and processes. Also use this code for discharges determined when the stage-discharge relation was affected by ice and for discharges determined from high-water marks located some distance from the gage. PEAKFQ includes peaks qualified with a code 2 in the frequency analyses. Occasionally, the need to rely on an estimate arises from an unusual circumstance relating to the origin of the peak (failure of ice jams, debris dam, etc.) If so, the peak may also be accompanied by a code 9.

Code 3. –Discharge affected by dam failure – Use code 3 only for dam-burst floods in which the peak was influenced by the unplanned and uncontrolled release of stored water, with concurrent destruction of the dam or other storage barrier. Qualify emergency reservoir releases and similar recurrent phenomena with one of the regulation/diversion codes (code 5 or 6). Peaks qualified with this code are interpreted as unique non-recurrent events, unrepresentative of future flood risk at the site. Peaks qualified with a code 3 are always excluded in PEAKFQ frequency analysis

Code 4. –Discharge is less than the indicated value, which is the minimum recordable discharge at this site – Use this code only if there probably was an unmeasured, non-zero flow at some time during the year. Otherwise, report a discharge of zero without any qualification. For example, a crest stage gage may be set too high to record the largest peak during a dry year. Lacking other evidence collected in the field, the discharge corresponding to the pin elevation could be entered with a code 4. Peaks qualified with this code are interpreted as defining the lower limit of definition of the flood-frequency curve at this site. In PEAKFQ all peaks less than or equal to the highest peak flagged with code 4 are treated as if they were flagged with a code 4. This has the effect of setting them equal to zero. By default these peaks are not included in PEAKFQ frequency analysis. The user may include these peaks by specifying a value lower than the lowest peak qualified with a code 4 value for the gage base discharge.

Code 5. –Discharge affected to unknown degree by regulation or diversion – Use code 5 only when the exact magnitude of a “planned” [manmade] regulation or diversion is unknown and probably insignificant (less than 10 percent of the peak discharge). The effect may be either to decrease or to increase the regulated peak. Quantify minor overflows and emergency releases that sometimes are thought of as dam failures with code 5 if the degree of effect is unknown, but probably not

significant and if such operations have occurred repeatedly in the past or are likely to recur in the future.

At some sites, natural storage such as beaver dams, lakes, or swamps can affect some peaks. Code 5 is not intended to be applied in such cases.

PEAKFQ frequency analysis includes peaks qualified with a code 5. See also remarks for code 6 and code 9.

Code 6. –Discharge affected by regulation or diversion – Use code 6 for storage effects caused by “planned” [manmade] structures or operations in situations where at least 10 percent of the basin is controlled by regulation or when usable storage in the basin exceeds about 103 acre-feet per square mile (Benson, 1962, Water-Supply Paper 1580-B, pp. 7-8). The effect may be either to decrease or to increase the regulated peak. It should be noted that regulation of one peak discharge does not necessarily imply regulation of succeeding peaks; discontinuing code 6 or using it periodically may be appropriate. For example, the peaks of medium floods may be greatly affected by some smaller reservoirs, but a very large flood may involve a sufficient volume of water that reservoir storage may be filled without affecting the magnitude of the peak. Thus the occurrence of significant regulation or diversion should be verified independently for each peak. Such verification might be accomplished by comparing upstream and downstream hydrographs, the stability of reservoir contents during a flood (inflow equal outflow), etc. Qualify minor overflows and emergency releases that sometimes are thought of as dam failures with code 6 if the effect is significant and if such operations have occurred repeatedly in the past or are likely to recur in the future.

At some sites, natural storage such as beaver dams, lakes, or swamps can affect some peaks. Code 6 is not intended to be applied in such cases.

Code 6 serves as a criterion for excluding regulated peaks from a homogeneous statistical sample of natural-flow peaks. By default, PEAKFQ excludes peaks qualified with code 6 from the frequency analysis; the user may include these peaks by indicating “Yes” for urbanized/regulated peaks.

Code 7. –Discharge is a historic peak – Use this code only to identify ISOLATED peaks that occurred either before or after systematic record or during extended breaks within the systematic gaging program at the site, and that would not have been observed or recorded except for preliminary evidence indicating that they were of relatively large magnitude. Do not assign this code to any peak observed as part of a systematic gaging program at the site, regardless of the magnitude of the peak or information about its return period. In particular, DO NOT use this code to identify the peak of record at a site if that peak occurred during the period of systematic data collection. Peaks with code 7 form a biased sub-sample and have to be treated separately from systematic records in statistical analyses. PEAKFQ excludes peaks with a qualification code of 7 unless the user specifies the bulletin 17B historical period.

Code 8. –Discharge actually greater than indicated value – Use code 8 only if the actual discharge cannot be estimated but is known to be much larger than the indicated value such as would occur if a gage was overtopped and high water marks were not available to support more exact estimates. Rather than use code 8, users are encouraged to supply an estimated value (accompanied by code 2), if at all possible. Values qualified with this code are always excluded from PEAKFQ statistical analyses because they may grossly understate the actual peak flow.

Code 9. –Discharge due to snowmelt, hurricane, ice-jam or debris dam breakup – Use this code to identify flood peaks generated by hydrometeorological mechanisms or natural storage phenomenon different from the predominant flood-generating mechanism at the site but which are still common enough to be representative of overall future flood risks at this site. The definition of different or unusual mechanisms will vary from region to region as the prevalence of snowmelt, hurricanes, and other rainfall-runoff generation mechanisms differs. Where code 9 is used to a significant degree, particularly when it is used to create subsamples of flood peak data, the WSC should explain in flood-frequency regionalization reports the process that it employed to assign code 9 to individual peaks.

At some sites, beaver dams lakes, or swamps may affect all peaks. Do not use Code 9 in such cases.

PEAKFQ always includes a peak qualified with a code 9 in the frequency analysis.

Code A. –Year of occurrence is unknown or not exact - If a month is not reported, the year reported should be the water year the peak occurred in. If a month is reported, the year should be the calendar year of that month. Generally, such peaks will have occurred sometime in December or January, but might also occur in September or October. When the year is unknown or not exact, the month and day should also be unknown or not exact (code B.) Peaks qualified with a code A are included in PEAKFQ frequency analysis.

Code B. –Month or day of occurrence is unknown or not exact –If a month is reported, the year should be the calendar year of that month. If no month is reported, the year should be the water year the peak occurred in. When the month or day is unknown or not exact, the year of occurrence may be known or unknown (code A.) Generally the year will be unknown for peaks occurring sometime in December or January. Peaks qualified with a code B are included in PEAKFQ frequency analysis.

Code C. –All or part of the record affected by urbanization, mining, agricultural changes, wildfire, deforestation (land cover changes), or channelization – Peak discharge is influenced significantly by basin development or drainage system modifications. Once applied, continue to apply code C until such time as land use, land cover, or drainage characteristics of the basin revert to natural conditions. Regulation/diversion code C is used to develop a homogeneous record for statistical analysis by segregating the peaks influenced by channel improvements, changes in land use, or other modifications that alter natural flow peaks. By default, PEAKFQ analysis excludes peaks qualified with a code C from the frequency

analysis; the user may include these peaks by indicating YES for Urban/Reg Peaks.

Code D. –Gage base discharge changed during this year – For most sites, the goal is to enter 3-4 events each year (on average, over a 10- to 20-year period) that exceed a selected partial-duration base discharge. This information may be important in the analyses of partial-duration time series. Such computations are usually made to better define flood magnitudes that have a probability of occurrence of 10 percent or more in any given year. If the base is lowered, the previous record may be rendered incomplete, making it necessary to add any peaks that were below the previous base discharge but above the new base discharge. Similarly, if the base is increased, some partial duration peaks that are less than the new base will need to be removed from the PFF database. Peaks qualified with a code D are included in a PEAKFQ frequency analysis.

Code E. –Only the annual maximum peak exceeded the flood base this year – Use code E to identify those years in a partial-duration series in which no secondary peaks exceeded the base. The code is needed in this case to resolve uncertainty as to whether the absence of secondary peaks from the record means that, except for the annual peak, the flow remained below the base OR that no search was made for secondary peaks. The code is used in the former case; do not use it unless a search for the secondary peaks was made and none were found to be greater than the partial-duration base value. A peak qualified with a code E has no effect on PEAKFQ and will be included in frequency analyses.