Journal of Hydrology 539 (2016) 640–647

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Quantification of uncertainties in the 100-year flow at an ungaged site near a gaged station and its application in Georgia

Huidae Cho *, Emma Bones

Water Resources Engineer, Dewberry, 2835 Brandywine Road Suite 100, Atlanta, GA 30341, United States

ARTICLE INFO

Article history:
Received 30 January 2016
Received in revised form 30 April 2016
Accepted 31 May 2016
Available online 7 June 2016
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Felix Frances, Associate Editor

Keywords:
Uncertainty analysis
100-year flow
Regression equation
Flood
Risk assessment
Floodplain

SUMMARY

The Federal Emergency Management Agency has introduced the concept of the "1-percent plus" flow to incorporate various uncertainties in estimation of the 100-year or 1-percent flow. However, to the best of the authors' knowledge, no clear directions for calculating the 1-percent plus flow have been defined in the literature. Although information about standard errors of estimation and prediction is provided along with the regression equations that are often used to estimate the 1-percent flow at ungaged sites, uncertainty estimation becomes more complicated when there is a nearby gaged station because regression flows and the peak flow estimate from a gage analysis should be weighted to compute the weighted estimate of the 1-percent flow. In this study, an equation for calculating the 1-percent plus flow at an ungaged site near a gaged station is analytically derived. Also, a detailed process is introduced for calculating the 1-percent plus flow for an ungaged site near a gaged station in Georgia as an example and a case study is performed. This study provides engineers and practitioners with a method that helps them better assess flood risks and develop mitigation plans accordingly.

1. Introduction

This study introduces a new uncertainty concept in flood insurance studies by the Federal Emergency Management Agency (FEMA), investigates how uncertainties in flow prediction are propagated through the flow weighting method developed by U.S. Geological Survey (USGS), and derives an analytical solution that quantifies those uncertainties.

The National Flood Insurance Program (NFIP) was created by the United States Congress through the passage of the National Flood Insurance Act of 1968 (FEMA, 2002). The Act was created in response to Hurricane Betsy which caused over one billion dollars in damage to the Gulf States. It was the first hurricane to cause damages in excess of one billion dollars and received the nickname "Billion Dollar Betsy" (Holladay and Schwartz, 2010). The devastation caused by poorly communicated risk can also be noted in recent flooding in South Carolina in October 2015, where 17 people were killed by flood waters (CNBC Weather, 2015), and in Missouri in January 2016, where 7100 buildings were affected by flooding and 25 people in Illinois and Missouri were killed by flood waters (ABC News, 2016). Currently, the NFIP insures over 5.5 million properties, affecting over 10,000 communities (Holladay and Schwartz, 2010), and since insurance payouts became widespread in 1978, it has provided 51.7 billion dollars in funds for homeowners (FEMA, 2015).

As the NFIP requires the purchase of flood insurance by property owners, a standard had to be established "to be used as the basis for risk assessment, insurance rating, and floodplain management" (FEMA, 2002). Based on these criteria, the "1-percent annual chance flood" (i.e., 100-year or 1% flood) was recommended for use as the NFIP standard (FEMA, 2002). The Department of Housing and Urban Development (HUD) initially oversaw the mapping of the 1% floodplain in flood-prone communities until 1979 when the responsibilities of the NFIP were taken over by FEMA. Since the inception of the NFIP in 1968, the 1% floodplain has been used to communicate the extent of the risk associated with flooding. However, there are many uncertainties associated with predicting the 1% floodplain, and when those uncertainties are accounted for, flooding risk has the potential to expand linearly or non-linearly depending on the channel geometry (Jung and Merwade, 2015). Since uncertainty analysis for floodplain mapping provides more resilient and reliable information for flood risk management (Ntelekos et al., 2006; Xu and Booij, 2007; Jung and Merwade, 2015) and, as computer models and topography have greatly improved since 1968, FEMA has realized the importance of evaluating those uncertainties in order to communicate the potential risk to communities. To facilitate the communication of risk, FEMA

* Corresponding author.
E-mail address: hcho@isnew.info (H. Cho).

http://dx.doi.org/10.1016/j.jhydrol.2016.05.070
0022-1694/

has added the "1-percent plus" flood elevation to the Risk Mapping, Assessment, and Planning (Risk MAP) flood risk products for all riverine analyses (FEMA, 2011, 2013). FEMA (2013) defines the "1-percent plus" flood elevation as follows:

> The 1% plus flood elevation is defined as a flood elevation derived by using discharges that include the average predictive error for the regression equation discharge calculation for the Flood Risk Project. This error is then added to the 1% annual chance discharge to calculate the new 1% plus discharge. The upper 84-percent confidence limit is calculated for Gage and rainfall-runoff models for the 1% annual chance event.

Statistically, the logarithmic 1% plus flow is one standard error of prediction away from the mean logarithmic estimate of the 1% flow, which is equivalent to the upper 84% prediction limit in a one-tailed test. Modelers should derive the 1% plus flood elevation by using the 1% plus flow, which indicates how uncertain the estimate of the 1% flow is. Although the benefits of including uncertainty analyses in flood risk assessments have been discussed previously in the literature (Ntelekos et al., 2006; Xu and Booij, 2007; Jung and Merwade, 2012, 2015), it is not well established how to obtain the 1% plus flow for estimating the corresponding flood elevation. Ames (2006) proposed a bootstrap approach for obtaining confidence limits of low flow from data and the same approach could be applied to flood studies if there were enough peak flow data from which a large number of new "realizations" of peak flow data could be generated. However this resampling technique is not applicable when the area of interest is not gaged and no records of peak flows are available. Other than this similar work of Ames (2006) that addresses uncertainties in low flow estimates intead of peak flow estimates, the literature review has revealed no instructions on how to compute the 1% plus flow to the best of the authors' knowledge. Therefore, with no guidance for modelers, they are unable to properly communicate the flood risk to communities, endangering the persons and properties in flood-prone areas.

In order to accurately calculate the 1% plus flows and properly communicate risk to those affected by floods, this study derives an equation for the 1% plus flow for regression analyses weighted with historical gage records using the error propagation method (Birge, 1939; Ku, 1966; Tellinghuisen, 2001). Additionally, a case study is presented that develops the 1% plus flow based on the method discussed in this study. The floodplains for the 1% and 1% plus flows are delineated and compared to demonstrate how the 1% plus floodplain can be an effective tool to communicate risk to communities and their residents.

2. Background

2.1. 1-percent flow

For FEMA flood studies, modelers often use USGS Scientific Investigations Reports (SIR) for estimating the magnitude and frequency of floods for ungaged watersheds. USGS publishes regional flood-frequency equations for different exceedance probabilities including the 1% chance flow for different states. For example, Gotvald et al. (2009) derived the regional flood-frequency equations for rural ungaged streams in Georgia. USGS takes the logarithm of peak flows and performs a regression analysis using various watershed parameters including the drainage area, slope, percents of the watershed falling in different hydrologic regions, etc. Using the regression equations derived in this way, hydrologic modelers can estimate flows at ungaged sites for a 1% chance of exceedance. A typical equation for such log-linear regression analyses is as follows:

$$\log Q_P = \log K + \sum_{i=1}^{n} A_i \log X_i \tag{1}$$

where $\log(\cdot)$ is the logarithm function of base 10, Q_P is the flow of a $P\%$ probability, K and A_i's are regression coefficients, and X_i's are independent variables describing the watershed. One can obtain the final equation for Q_P by taking the exponential of both sides of Eq. (1) after a log-linear regression analysis as follows:

$$Q_P = K\prod_{i=1}^{n} X_i^{A_i}. \tag{2}$$

Eq. (2) can be used to estimate the $P\%$ flow at ungaged sites. USGS computes a $100(1-\alpha)\%$ prediction interval for the true peak flow at an ungaged site using the following inequation:

$$Q_P/C < Q_P < CQ_P \tag{3}$$

where C is defined as

$$C = 10^{t_{(\alpha/2,n-p)} S_{p,i}} \tag{4}$$

where $t_{(\alpha/2,n-p)}$ is the critical value of the Student's t-distribution at an α level and degrees of freedom $n-p$, where n is the number of observations used for the regression analysis and p is the number of regression variables; and $S_{p,i}$ is the standard error of prediction for site i. Gotvald et al. (2009) defines $S_{p,i}$ as follows:

$$S_{p,i} = \gamma^2 + \mathbf{x}_i \mathbf{U} \mathbf{x}_i^T \tag{5}$$

where γ^2 is the model error variance, $\mathbf{x}_i$ is a row vector of a 1 as the first element followed by regression equation parameter values for site i, $\mathbf{U}$ is the covariance matrix for the regression coefficients, and $\mathbf{x}_i^T$ is the transpose of $\mathbf{x}_i$.

When there is a streamflow gage at the outlet of the study watershed, one can analyze historical records of annual peak flows to estimate the $P\%$ flow without having to use the regression equation. Assuming that annual peak flows follow the log-Pearson Type III distribution as recommended by the Interagency Advisory Committee on Water Data (1982), one can fit a distribution curve to historical annual peak flows by adjusting its statistical parameters. The PeakFQ program (Flynn et al., 2006) implements this parameter estimation procedure, and calculates flows and confidence intervals for different probabilities. The user can obtain the upper one-tailed 84% prediction limit of the 1% flow by changing the confidence intervals parameter to 0.84 (i.e., 84%). This upper 84% prediction limit is located one standard prediction error away from the mean estimate of the 1% flow and corresponds to the definition of the 1% plus flow.

2.2. Weighted flow estimate for an ungaged site near a gaged station

Since the area of interest is most likely ungaged, the regression analysis is needed to estimate peak flows. However, if there is a gaged station near the ungaged site of interest, it is recommended by the Interagency Advisory Committee on Water Data (1982) to incorporate the gage data into the regression analysis to produce results that more closely resemble the real-world data. According to Gotvald et al. (2009), a gaged station is considered near an ungaged site if the drainage area of the ungaged site is within the range of 50% to 150% of the drainage area of the gaged station. To properly incorporate gaged data into flow estimates at nearby ungaged sites, Gotvald et al. (2009) combined two estimates of the peak flow from the gaged station and ungaged site by weighting both flows with the drainage area difference between the two locations, and defined the weighted estimate of the peak flow for a $P\%$ chance exceedance at the ungaged site as follows: