AGU ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1029/2019WR026940

Key Points:
- A comprehensive statistical framework for regional hydrologic regression is presented with spatially correlated flow data and varying record lengths
- Framework correctly attributes variability to sampling errors in computed statistic, variability explained by the model, and model error
- New diagnostics includes Bayesian plausibility value, pseudo adjusted R^2, pseudo ANOVA, and Bayesian metrics for leverage and influence

Correspondence to:
D. S. Reis Jr,
dirceureis@unb.br

Citation:
Reis, D. S., Jr, Veilleux, A. G., Lamontagne, J. R., Stedinger, J. R., & Martins, E. S. (2020). Operational Bayesian GLS regression for regional hydrologic analyses. *Water Resources Research*, *56*, e2019WR026940. https://doi.org/10.1029/2019WR026940

Received 10 DEC 2019
Accepted 13 FEB 2020
Accepted article online 19 FEB 2020

Operational Bayesian GLS Regression for Regional Hydrologic Analyses

Dirceu S. Reis Jr[1], Andrea G. Veilleux[2], Jonathan R. Lamontagne[3], Jery R. Stedinger[4], and Eduardo S. Martins[5]

[1]Department of Civil and Environmental Engineering, University of Brasilia, Brasilia, Brazil, [2]U.S. Geological Survey, Water Mission Area, Analysis and Prediction Branch, Reston, VA, USA, [3]Department of Civil and Environmental Engineering, Tufts University, Medford, MA, USA, [4]School of Civil and Environmental Engineering, Cornell University, Ithaca, NY, USA, [5]Research Institute for Meteorology and Water Resources, FUNCEME, Fortaleza, Brazil

Abstract This paper develops the quasi-analytic Bayesian analysis of the generalized least squares (GLS) (B-GLS) model introduced by Reis et al. (2005, https://doi.org/10.1029/2004WR003445) into an operational and statistically comprehensive GLS regional hydrologic regression methodology to estimate flood quantiles, regional shape parameters, low flows, and other statistics with spatially correlated flow records. New GLS regression diagnostic statistics include a Bayesian plausibility value, pseudo adjusted R^2, pseudo analysis of variance table, and two diagnostic error variance ratios. Traditional leverage and influence are extended to identify rogue observations, address lack of fit, and support gauge network design and region-of-influence regression. Formulas are derived for the Bayesian computation of estimators, standard errors, and diagnostic statistics. The use of B-GLS and the new diagnostic statistics are illustrated with a regional log-space skew regression analysis for the Piedmont region in the Southeastern U.S. A comparison of ordinary, weighted, and GLS analyses documents the advantages of the Bayesian estimator over the method-of-moment estimator of model error variance introduced by Stedinger and Tasker (1985, https://doi.org/10.1029/WR021i009p01421). Of the three types of analyses, only GLS considers the covariance among concurrent flows. The example demonstrates that GLS regional skewness models can be highly accurate when correctly analyzed: The B-GLS average variance of prediction is 0.090 for South Carolina (92 stations), whereas a traditional ordinary least squares analysis published by the U.S. Geological Survey yielded 0.193 (Feaster & Tasker, 2002, https://doi.org/10.3133/wri024140). B-GLS provides a statistical valid framework for the rigorous analysis of spatially correlated hydrologic data, allowing for the estimation of parameters and their actual precision and computation of several diagnostic statistics, as well as correctly attributing variability to the three key sources: time sampling error, model error, and signal.

1. Introduction

Hydrologists often need to estimate hydrologic quantities for water resources planning and floodplain management, including rainfall intensity-duration-frequency (IDF) curves, annual mean flow, flood quantiles, and low-flow statistics. Because these estimates are commonly required at ungauged sites, they need to be based upon gauged data collected elsewhere in the region. Even if the estimates are required at a gauged site, the record length may not be long enough to provide estimates with (good) precision, in which case, the combination of both regional and at-site estimators, based on the precision of each, can provide a more precise estimate of the variable of interest (IACWD, 1982; Kuczera, 1982, 1983; Merz & Bloschl, 2008; Micevski & Kuczera, 2009; Vicens et al., 1975).

Various strategies have been developed over the years to obtain regional estimates of hydrologic variables at ungauged sites (Cunnane, 1988; GREHYS, 1996). The most popular approach, which is the one adopted here, consists of building an empirical relationship between the hydrologic variable of interest and catchment attributes at gauged sites using regional regression analysis (e.g., Thomas & Benson, 1970; Jennings et al., 1994; Griffis & Stedinger, 2007b). The rationale is that catchment attributes would be closely related to the patterns of many hydrological processes that occur in the basin (Merz & Bloschl, 2005).

An alternative to traditional regional regression methods is the region-of-influence (ROI) approach proposed by Burn (1990a, 1990b). The ROI approach is not conceptually different from regional regression, in

that it relies on the idea that gauges with some degree of similarity with the ungauged site of interest should be employed to obtain a regional estimator. Similarity among sites is often measured in two different ways: catchment attributes (e.g., slope, area, soil type, and land use) and basin proximity. The former assumes that basins with similar characteristics should have similar hydrologic responses. However, because the scientific understanding of how catchment characteristics influence water fluxes and storages in a basin is incomplete, it is possible that an important basin attribute is not considered in the set of possible explanatory variables. This justifies the use of spatial proximity, a simple and reasonable measure of similarity that assumes that sites close to each other share similar hydrologic responses. The rationale behind this idea is based on the assumption that hydrologic variables of interest vary smoothly over the region. Merz and Bloschl (2005) and Eng, Milly, et al., (2007), Eng, Stedinger, et al., (2007) provide evidence that suggests that these similarity measures are in fact complementary and show that a combination of both similarity measures results in a set of gauged sites that provides more precise regional estimators for ungauged sites.

Regional estimators of hydrologic variables can also be obtained by Bayesian hierarchical models (Lima et al., 2016; Renard, 2011; Viglione et al., 2013), which can be seen as a generalization of regression models. Hierarchical models are structured in layers. The idea is to have one layer to model the at-site data variability and another layer to model the regional dependence. The hydrologic variable of interest observed in gauged sites can be modeled as coming from a specific probability distribution with unknown parameters in one layer, and the spatial variability of these parameters can then be modeled in another layer as linearly dependent on some basin characteristics. This hierarchical approach differs from the traditional regression approach because the whole estimation process is carried out simultaneously, allowing for possible uncertainty interactions in the different layers; that may be an advantage over the traditional approach where two independent steps are usually taken: (1) at-site estimation of the variable of interest at all gauged sites and its corresponding covariance matrix and (2) derivation of the regional regression model based on basin characteristics. Then, perhaps (3), use the regional regression model as a prior in an at-site analysis at another time or a place not included in the original analysis (Martins & Stedinger, 2000).

Spatial interpolation techniques, such as canonical (Chokmani & Ouarda, 2004) and topological kriging (Archfield et al., 2013; Skøien et al., 2006; Skøien & Bloschl, 2007), can provide regional estimators of hydrologic variables. Canonical kriging is applied to a two-dimensional spatial representation of catchment and climate attributes, which is often obtained by some multivariate statistical method, such as canonical correlation analysis (Chokmani & Ouarda, 2004; Castiglioni et al., 2009) or principal component analysis. All gauged and ungauged sites in the region can be represented in this two-dimensional space (x, y). The observed or estimated values of the hydrologic variable for each of the gauged sites are represented in a third dimension (z), which are interpolated to obtain a regional estimator at ungauged sites. Topological kriging, on the other hand, is an interpolation method that does not work through basin characteristics space but across a river network taking into account the drainage areas and their proximity. Topological kriging can be very much like region-of-influence regression in that the estimator for any point is a unique weighting of values at nearby gauges. Topological and canonical krigings have been used to provide regional estimators for flood quantiles (Skøien & Bloschl, 2007) and low-flow statistics (Castiglioni et al., 2009, 2011).

Archfield et al. (2013) compare canonical and topological kriging against the traditional regression approach for flood quantile estimation advance here. They concluded that topological kriging outperformed the traditional regression approach for the analyzed case and suggested that future studies should evaluate how the different treatments of spatial correlation affect the performance of these quantile estimators. It is important to note that their data set was not an example of strictly "ungauged" watersheds as one might expect from the title and text. In a regional analysis for the same region, using many of the same flood records, Feaster et al. (2009) and Veilleux (2011) note that many stream gauges were located on the same stream and corresponded to very nearly the same watershed. Archfield et al.'s map shows how many watersheds were nested. Topological kriging is designed to take advantage of such spatially relationships, including "both the area and nested nature of catchments" (Castiglioni et al., 2011). In Archfield et al. (2013), and many other studies, "ungauged" functionally meant there was not a gauge exactly at the stream cross section of interest, though there could be gauges not far upstream or downstream or both. However, traditional regression has been formulated to estimate hydrologic statistics at ungauged sites, meaning the watersheds of nearby sites do not contain or are contained in the watershed of concern; thus, an accurate flow record for the site of concern cannot be obtained by a modest adjustment of the record of flows at one or more nearby gauges. The

Archfield et al. (2013) kriging results are what is to be expected when many gauges are located near each other in a stream network; that relatively accurate flood quantile estimates were obtained in such cases is an important result.

This paper represents advances in regional regression analyses for hydrologic variable estimation. For many years, regional regression analyses used ordinary least squares (OLS) that considers the residual errors to be homoscedastic and independently distributed (Riggs, 1973). However, the estimates of the variable of interest at different gauged sites have different precision due to differences in record length (Tasker, 1980; Kuczera, 1983) and possible differences in the precision of measurements and their variability (Tasker & Stedinger, 1989).

A further complication is that concurrent hydrologic measurements in nearby or similar basins are generally correlated so sampling errors may not be independent. To address this, Stedinger and Tasker (1985, 1986a, 1986b) developed a generalized least squares (GLS) framework, which considers both differences in precisions and cross correlation among station estimators. They showed that a GLS analysis provides better estimates of the model parameters and the model error variance (in terms of mean square errors) than does OLS. See also Kroll and Stedinger 1998). The GLS procedure has been widely used in many hydrologic studies, including the regionalization of flood quantiles (Tasker et al., 1986; Curtis, 1987; Landers & Wilson, 1991; Rosbjerg & Madsen, 1995; GREHYS, 1996; Madsen & Rosbjerg, 1997; Robson & Reed, 1999; Kjeldsen & Rosbjerg, 2002; Feaster & Tasker, 2002; Micevski & Kuczera, 2009), water quality parameters (Tasker & Driver, 1988), low-flow statistics (Ludwing & Tasker, 1993), extreme rainfall (Madsen et al., 2002), and the design of hydrologic network (Moss & Tasker, 1991).

A weighted least squares (WLS) procedure, which considers only differences in record lengths, has been used for the regionalization of the shape parameter (the skewness coefficient) for the State of Kansas (Rasmussen & Perry, 2000) and the State of North Carolina (Pope et al., 2001). GLS has been used as a regression method in various studies using ROI techniques to regionalize flood quantiles (Tasker et al., 1996; Law & Tasker, 2003, Eng, Milly, et al., 2007; Eng, Stedinger, et al., 2007; Haddad et al., 2011), L-moments (Laio et al., 2011), and parameters of a theoretical probability distribution, such as the log-Pearson type 3 (Micevski et al., 2015). GLS has also been used as the basis of hydrologic network design (Medina, 1987; Moss & Tasker, 1991; Soenksen et al., 1999; Tasker, 1986; Tasker & Stedinger, 1989). Kjeldsen and Jones (2007) explore the spatial correlation of regression model error variance. Building on that work, Kjeldsen and Jones (2009a, 2009b, 2010) developed a recursive GLS procedure, which takes into account the cross correlation of the regression errors in hydrological regression models. This work may be particularly useful in situations where only a few catchment descriptors are available and data from nearby stations are used to improve regression models (Kjeldsen et al., 2014).

Reis et al. (2003, 2005) introduced a Bayesian approach to estimation for the model error variance of the GLS regression analysis developed by Stedinger and Tasker (1985, 1986a, 1986b) for regional hydrologic analysis, which they used to get the posterior distribution of estimated $\boldsymbol{\beta}$ parameters. A Bayesian analysis (Gelman et al., 2004; Zellner, 1971) provides both an exact measure of precision of the model error variance that the method-of-moments (MM) and maximum likelihood (ML) estimators lack and a more reasonable description of the possible values of the model error variance in cases where the MM and ML model error variance estimators are 0 or nearly 0 (examples in Madsen & Rosbjerg, 1997). Reis et al. (2005) show that for cases in which the model error variance is small compared to the sampling error of the at-site estimates, which is often the case for regionalization of a shape parameter, the Bayesian posterior provides a more reasonable description of the model error variance than either the MM or ML point estimator. The MM estimator of the model error variance can be 0 if the observed variability in the data is explained by the sampling error in the at-site estimates, causing a distortion in the uncertainty of the regional estimate. Similarly, the ML estimator of the model error variance may not be a good representation of the possible values of the model error variance when its value is small or 0 because the likelihood function is often highly skewed; this results in the mode being a less appropriate summary statistic than the center of mass. Sometimes, the mode is at the origin, which results in an ML estimate of 0. Qian et al. (2005) employ a similar Bayesian analysis for a watershed loading model with three error terms representing independent observational errors, a structural correlated spatial dependency, and the impact of errors in one reach on the distribution of the estimated loads downstream. Jeong et al. (2007) use Bayesian GLS (B-GLS) for regionalization of the L-CV

and L-skew for the generalized extreme value (GEV) distribution in Korea, while Haddad, Johnson, et al. (2015) employ a B-GLS framework to obtain regional regression equations to predict rainfall L-moments in ungauged sites in Australia. Seidou et al. (2006) use a parametric Bayesian methodology to estimate GEV flood quantiles, Kim and Lee (2010) use a Bayesian OLS procedure to estimate low-flow statistics in Korea, and Haddad et al. (2013) and Egodawatta et al. (2014) employ OLS, WLS, and Bayesian WLS procedures to estimate uncertainty in urban storm water quality models. More recently, the Australian Government, in its updated Australian rainfall-runoff guideline on regional peak flow estimation (Rahman et al., 2016), recommended the use of Bayesian GLS for regional estimation of the parameters of the log-Pearson type 3 distribution. The decision was made based on comparative studies in Australia (Haddad et al., 2012; Haddad & Rahman, 2012; Haddad, Rahman, & Ling, 2015). In the United States, the U.S. Geological Survey (USGS) has recently conducted regional skew studies in several states employing the Bayesian GLS framework suggested in Reis et al. (2005) and further developed in Gruber et al. (2007) (Eash et al., 2013; Feaster et al., 2009; Gotvald et al., 2009; Gotvald et al., 2012; Lamontagne et al., 2012; Mastin et al., 2016; Olson, 2014; Over et al., 2016; Parrett et al., 2011; Paretti et al., 2014; Southard & Veilleux, 2014; Weaver et al., 2009; Wagner et al., 2016; Wood et al., 2016; Zarriello, 2017).

This paper starts with the Bayesian GLS model error variance estimator in Reis et al. (2005) and develops a complete analysis framework including a range of new regression diagnostic statistics for both B-WLS and B-GLS analyses. New diagnostic statistics include a description of significance, a pseudo R^2, a pseudo analysis of variance (ANOVA) table, and two error variance ratios (EVRs) that quantify the need for WLS and GLS analyses. Leverage and influence statistics for GLS identify rogue observations, address lack of fit, and support gauge network design and ROI regression. This paper also compares the results of OLS, WLS, and GLS analyses and the use of MM and Bayesian model error variance estimators to derive regional OLS/WLS/GLS models of the shape parameter (skewness coefficient) of the log-Pearson type III distribution for the Piedmont region in the Southeastern U.S. The paper by Reis et al. (2005) proposed the idea of a Bayesian GLS model error variance estimator and analyzed data for the Tibagi River and the Muskingum River basins. This paper builds on that idea and provides a full suite of diagnostic statistics, thus providing a statistically rigorous and comprehensive framework for the analysis of hydrologic information consisting of cross-correlated streamflow records of different length from stations across a region.

2. GLS Regression for Hydrologic Statistics

Streamflow data sets can be used to derive an empirical relationship between hydrologic characteristics at a site, such as the T-year flood or log-space skewness coefficient, and physiographic variables, such as drainage area and channel slope. Our GLS analysis assumes that the actual value of the quantity of interest y_i (or some transformation of y_i) for a given site i can be described by a linear function of physiographic basin characteristics with an additive error

$$y_i = \beta_0 + \sum_{s=1}^{k} \beta_s X_{is} + \delta_i \qquad i = 1, 2, ..., \ \ n \ \text{ sites}, \tag{1}$$

wherein X_{is} $(s = 1 \ldots, k)$ are the elements of a matrix of k explanatory variables based upon the physical characteristics at each site i, β_s are the model parameters, and δ_i are the independently distributed model errors with the following properties:

$$E[\delta_i] = 0 \qquad \mathrm{Cov}(\delta_i, \delta_j) = \begin{cases} \sigma_\delta^2 & i = j \\ 0 & i \neq j \end{cases}. \tag{2}$$

However, in most analyses, only an at-site estimate of y_i, $\hat{y}_i$, is available, and thus, a sampling error η_i should be introduced into the model, such that

$$\hat{y}_i = y_i + \eta_i \quad i = 1, 2, ..., n \ \text{ sites}, \tag{3}$$

with

$$E[\eta_i] = 0 \qquad \mathrm{Cov}\left(\eta_i, \eta_j\right) = \begin{cases} \sigma_{\eta_i}^2 & i = j \\ \sigma_{\eta_i}\sigma_{\eta_j}\rho_{ij} & i \neq j \end{cases}, \tag{4}$$

wherein $\sigma_{\eta_i}^2$ is the at-site sampling error variance for $\hat{y}_i$ and ρ_{ij} is the sampling error correlation coefficient due to correlation among the statistic of interest at stations i and j (cross-site correlation).

In matrix notation, the GLS model is

$$\hat{\mathbf{y}} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\eta} + \boldsymbol{\delta} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{5}$$

where the $(n \times [k+1])$ matrix $\mathbf{X}$ contains ones in the first column and values of the k explanatory variables in the remaining columns, the vector $\boldsymbol{\beta}$ has the $(k+1)$ parameters of the model that must be estimated, the vector $\boldsymbol{\eta}$ contains the sampling errors in the sample estimators, and the vector $\boldsymbol{\delta}$ contains the model errors for the n sites used in the analysis.

The errors ε_i are a combination of (i) sampling error η_i in the sample estimators of y_i and (ii) underlying model error δ_i. The total error vector $\boldsymbol{\varepsilon}$ has mean 0 and covariance matrix

$$E\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\right] = \boldsymbol{\Lambda}\left(\sigma_\delta^2\right) = \sigma_\delta^2\mathbf{I} + \boldsymbol{\Sigma}(\hat{\mathbf{y}}), \tag{6}$$

where $\boldsymbol{\Sigma}(\hat{\mathbf{y}})$ is the covariance matrix of the sampling errors in the sample estimators whose elements are given by Equation (4) and σ_δ^2 is the underlying model error variance, which must be determined. The value of σ_δ^2 can be viewed as a heterogeneity measure (Madsen et al., 2002; Madsen & Rosbjerg, 1997).

WLS and OLS analyses are special cases of a GLS analysis. When $\hat{\rho}\left(\hat{y}_i, \hat{y}_j\right) = 0$ for every pair of sites $(i \neq j)$, GLS reduces to WLS. WLS reduces to OLS when the diagonal covariance matrix has elements on the diagonal equal to a common value.

The GLS estimator of $\boldsymbol{\beta}$ and its respective covariance matrix for known σ_δ^2 are given by

$$\mathbf{b} = \left[\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right]^{-1}\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\hat{\mathbf{y}}, \tag{7a}$$

$$\boldsymbol{\Sigma}[\mathbf{b}] = \left[\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right]^{-1}. \tag{7b}$$

The model error variance σ_δ^2 can be estimated by either generalized MM or ML estimator, as described by Stedinger and Tasker (1986). The MM generalized estimator is determined by iteratively solving Equation (7a) along with the generalized residual mean square error equation:

$$(\hat{\mathbf{y}} - \mathbf{X}\mathbf{b})^T\left[\hat{\sigma}_\delta^2\mathbf{I} + \Lambda(\hat{\mathbf{y}})\right]^{-1}(\hat{\mathbf{y}} - \mathbf{X}\mathbf{b}) = n - (k+1), \tag{8}$$

for n sites and $k+1$ parameters. In some situations, the sampling covariance matrix explains all the variability observed in the data, which means that the left-hand side of Equation (8) will be less than $n-(k+1)$ even if $\hat{\sigma}_\delta^2$ is 0. In these circumstances, the MM estimator of the model error variance is generally taken to be 0 (Stedinger & Tasker, 1985, 1986a, 1986b).

The ML estimators of $\boldsymbol{\beta}$ and σ_δ^2 can be obtained by minimizing the negative of the log-likelihood function of the residuals, which are assumed to be normally distributed with zero mean and the covariance matrix in Equation (6):

$$\min\left\{\ln\left[\left|\Lambda\left(\sigma_\delta^2\right)\right|\right] + (\hat{\mathbf{y}} - \mathbf{X}\boldsymbol{\beta})^T\Lambda\left(\sigma_\delta^2\right)^{-1}(\hat{\mathbf{y}} - \mathbf{X}\boldsymbol{\beta})\right\}, \tag{9}$$

subject to $\hat{\sigma}_\delta^2 \geq 0$, wherein $|\mathbf{A}|$ denotes the determinant of a matrix $\mathbf{A}$. The ML estimate of $\boldsymbol{\beta}$ is the same as the one computed by using Equations (7a) and (7b), except that the value of $\hat{\sigma}_\delta^2$ would be different. The variance of $\boldsymbol{\beta}$ is the same as in Equations (7a) and (7b) because $\boldsymbol{\beta}$ and $\hat{\sigma}_\delta^2$ are asymptotically independent (Rencher, 2000). The inverse of the second derivative of the likelihood function could be used to estimate the variance of the model error variance estimator when the constraint $\hat{\sigma}_\delta^2 \geq 0$ is not binding (Bickel & Doksum, 1977).

3. Bayesian Approach

Reis et al. (2005) developed a Bayesian analysis of the GLS model. In particular, they computed the posterior moments of the $\boldsymbol{\beta}$ parameters and the full posterior distribution of the model error variance σ_δ^2.

The Bayesian approach requires the specification of prior distributions for both the $\boldsymbol{\beta}$ parameters and model error variance σ_δ^2. A multivariate normal distribution with a mean of 0 and a large variance was used for the prior for $\boldsymbol{\beta}$. This almost noninformative prior produces a pdf that is relatively flat in the region of interest. The prior information for the model error variance σ_δ^2 was represented by an exponential distribution with parameter λ, which represents the reciprocal of the prior mean of the model error variance. Following Reis et al. (2005) for the regionalization of skews, we employ a value for λ equal to 6, though as experience accumulates a smaller value or a different distribution may be justified.

The likelihood function for the data $\widehat{\mathbf{y}}$ is considered to be a multivariate normal distribution; thus, the marginal posterior distribution of the model error variance can be computed by integrating the joint posterior distribution over the possible values of $\boldsymbol{\beta}$ (Zellner, 1971, equation 8.14; Kitanidis, 1986) to obtain

$$f(\sigma_\delta^2|\widehat{\mathbf{y}}) = \int f(\boldsymbol{\beta}, \sigma_\delta^2|\widehat{\mathbf{y}})\, d\boldsymbol{\beta} \propto \int f(\widehat{\mathbf{y}}|\boldsymbol{\beta}, \sigma_\delta^2)\, \xi(\boldsymbol{\beta}, \sigma_\delta^2)\, d\boldsymbol{\beta}, \tag{10}$$

where $f(\boldsymbol{\beta}, \sigma_\delta^2|\widehat{\mathbf{y}})$ is the joint posterior of the parameters and the model error variance, $f(\widehat{\mathbf{y}}|\boldsymbol{\beta}, \sigma_\delta^2)$ is the likelihood function for the data $\widehat{\mathbf{y}}$, and $\xi(\boldsymbol{\beta}, \sigma_\delta^2)$ is the joint prior for $\boldsymbol{\beta}$ and σ_δ^2. If one uses a noninformative prior on $\boldsymbol{\beta}$, the marginal posterior distribution for the model error variance, except for the normalizing constant, is

$$f(\sigma_\delta^2|\widehat{\mathbf{y}}) \propto \left[|\boldsymbol{\Lambda}(\sigma_\delta^2)| \left| \mathbf{X}^T \boldsymbol{\Lambda}(\sigma_\delta^2)^{-1} \mathbf{X} \right| \right]^{-1/2} \exp\left[-0.5 (\widehat{\mathbf{y}} - \mathbf{X}\mathbf{b})^T \boldsymbol{\Lambda}(\sigma_\delta^2)^{-1} (\widehat{\mathbf{y}} - \mathbf{X}\mathbf{b}) \right] \xi(\sigma_\delta^2). \tag{11}$$

The importance of using the correct likelihood function for a data set is illustrated by Stedinger et al. (2008). With Equation (11), one can compute the marginal pdf, mean, and variance of σ_δ^2 by a numerical evaluation of one-dimensional integrals (Reis et al., 2005). Similarly, posterior moments of $\boldsymbol{\beta}$ can also be computed by a one-dimensional numerical integration using the pdf in Equation (11) where the conditional distribution of $(\boldsymbol{\beta}|\sigma_\delta^2, \widehat{\mathbf{y}})$ is normal with mean and variance given in (7a) and (7b); thus, we have

$$\boldsymbol{\mu}_\beta = E(\boldsymbol{\beta}|\widehat{\mathbf{y}}) = \int \boldsymbol{\beta} f(\boldsymbol{\beta}|\widehat{\mathbf{y}})\, d\boldsymbol{\beta} = \int E(\boldsymbol{\beta}|\sigma_\delta^2, \widehat{\mathbf{y}}) f(\sigma_\delta^2|\widehat{\mathbf{y}})\, d\sigma_\delta^2 = \int \mathbf{b}(\sigma_\delta^2) f(\sigma_\delta^2|\widehat{\mathbf{y}})\, d\sigma_\delta^2, \tag{12}$$

$$\begin{aligned} Var(\boldsymbol{\beta}|\widehat{\mathbf{y}}) &= \int\int \left(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta\right)\left(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta\right)^T f(\boldsymbol{\beta}|\sigma_\delta^2, \widehat{\mathbf{y}}) f(\sigma_\delta^2|\widehat{\mathbf{y}})\, d\boldsymbol{\beta}\, d\sigma_\delta^2 \\ &= \int \left\{ \left(\mathbf{b}(\sigma_\delta^2) - \boldsymbol{\mu}_\beta\right)\left(\mathbf{b}(\sigma_\delta^2) - \boldsymbol{\mu}_\beta\right)^T + \left(\mathbf{X}^T \boldsymbol{\Lambda}(\sigma_\delta^2)^{-1} \mathbf{X}\right)^{-1} \right\} f(\sigma_\delta^2|\widehat{\mathbf{y}})\, d\sigma_\delta^2. \end{aligned} \tag{13}$$

Here the posterior variance of $\boldsymbol{\beta}$ equals the variance of the conditional mean $\mathbf{b}(\sigma_\delta^2)$ plus the average of the conditional variance of $\boldsymbol{\beta}$ for a given σ_δ^2 (Reis et al., 2005).

4. Model Selection

A goal of model selection is to resolve which set of possible explanatory variables affords the most accurate prediction, while also searching for the simplest model possible. Several traditional statistics are available for model selection: R^2, likelihood ratios, Mallows C_p, Akaike information criterion (AIC), and the Bayesian information criterion (BIC) (Gelman et al., 2004; Linhart & Zucchini, 1986). Qian et al. (2005) employ Bayes factors and a Bayesian deviance information criterion (DIC). Many of these statistics penalize model complexity. Thus, a sufficient improvement in the model's prediction ability must result so as to support the inclusion of an additional independent variable. Below, the B-GLS regression statistics developed by the authors to guide model selection are presented.

4.1. Average Variance of Prediction

Because our interest is to make predictions at gauged and ungauged sites, a natural metric to evaluate a model is a variance of prediction, which penalizes the inclusion of extra independent variables because it accounts for the sampling variance of the parameters (Carlin & Louis, 2000). Because the variance of prediction generally depends upon the values of the independent variables at a given site, Tasker and Stedinger (1986) suggested the use of an average variance of prediction (AVP), computed across the x values of sites

used in the regression. This implicitly assumes that these sites are representative of the sites at which predictions will be made.

For a new and perhaps ungauged site with a row vector of explanatory characteristics $\mathbf{x}_o$ and a y value of y_o, the posterior expected value of y_o is $\mathbf{x}_o\boldsymbol{\mu}_\beta$, where $\boldsymbol{\mu}_\beta$ is the posterior expected value of $\boldsymbol{\beta}$. With a Bayesian analysis, the posterior sampling error variance for $\mathbf{x}_o\boldsymbol{\mu}_\beta$ is

$$Var\left(\mathbf{x}_o\boldsymbol{\mu}_\beta\right) = \underset{\sigma_\delta^2}{E}\left\{\mathbf{x_o}\left[\left(\mathbf{b}-\boldsymbol{\mu}_\beta\right)\left(\mathbf{b}-\boldsymbol{\mu}_\beta\right)^T + \left(\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\right]\mathbf{x_o}^T\right\} = \mathbf{x}_o Var[\boldsymbol{\beta}|\hat{\mathbf{y}}]\mathbf{x}_o^T. \tag{14}$$

The posterior variance of prediction for the unknown true value y_o associated with x_o is given by

$$VP = E\left[\left(y_0-\mathbf{x}_0\boldsymbol{\mu}_\beta\right)^2\right] = E\left[\sigma_\delta^2\right] + Var\left(\mathbf{x}_o\boldsymbol{\mu}_\beta\right) = E\left[\sigma_\delta^2\right] + \mathbf{x}_o Var[\boldsymbol{\beta}|\hat{\mathbf{y}}]\mathbf{x}_o^T. \tag{15}$$

A measure of how well OLS, WLS, and GLS regression analysis would predict a hydrologic statistic on average over a new region, whose $\mathbf{x}_o$ was like those in the $\mathbf{X}$ matrix, is the AVP for a new site AVP_{new}, introduced by Tasker and Stedinger (1986). For a Bayesian analysis, as shown in Gruber et al. (2007),

$$AVP_{\text{new}} = E\left[\sigma_\delta^2\right] + \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i Var[\boldsymbol{\beta}|\hat{\mathbf{y}}]\mathbf{x}_i{}^T. \tag{16}$$

However, if the prediction is for a site i that was used for estimation of the parameters of the model, the variance of prediction for such an old site, VP_{old} in Equation (A.13), requires an additional term. Appendix A provides derivations of the expressions used to calculate variance of predictions.

Generally, we anticipate that models will be used at new sites, so AVP_{new} is the critical statistic for model selection. However, if a regional estimator is nearly as precise as the at-site estimator, then one may use the regional estimator instead or a weighted average, in which case, VP_{old} becomes relevant, if there is any practical difference between the two.

4.2. Bayesian Plausibility Value

It is common in classical statistics to perform a hypothesis test to check if a given parameter is statistically different from 0. If one cannot reject the null hypothesis that the value of the parameter is equal to 0, the variable whose coefficient is being tested is generally dropped from the model. It is also common to report a P value, which reflects the probability, under the null hypothesis, of computing a parameter value as large or larger than the value obtained from the sample.

The Bayesian plausibility value, ψ, developed by Reis (2005) and expanded on by Gruber et al. (2007) describes whether 0 is a plausible value for each $\boldsymbol{\beta}$ parameter in a regression analysis given the prior and data. As discussed by Lindley (1965) and Zellner (1971), given the Bayesian posterior pdf of $\boldsymbol{\beta}$ and the available data, one can construct a credible region for the regression parameters. A credible region summarizes the posterior belief about a parameter and can be the basis of a hypothesis test that concludes that a parameter is 0 if 0 is included in a 90% or a 95% credible region. This allows one to perform the equivalent of a classical hypothesis test within a Bayesian framework using the posterior distribution of each parameter, which also reflects the prior.

Here we define the plausibility level for 0 to be the smallest probability ψ such that 0 is in a $100(1-\psi)\%$ credible region for a parameter. The plausibility value is computed as

$$\psi = 2\underset{\sigma_\delta^2}{E}\left\{\Phi\left[-\upsilon\frac{b\left(\sigma_\delta^2\right)}{\sigma_b\left(\sigma_\delta^2\right)}\right]\right\}, \tag{18}$$

wherein Φ is the standard normal cdf and the conditional mean $b\,(\sigma_{\delta 2})$ and standard error $\sigma_b(\sigma_{\delta 2})$ for $\boldsymbol{\beta_i}$ are both dependent on σ_δ^2 as indicated in Equations (7a) and (7b); $\upsilon = \text{sign}[\mu_\beta] = 1$ for $\mu_\beta \geq 0$ and -1 for $\mu_\beta < 0$.

The Bayesian P value discussed by Bayarri and Berger (2000) and Robins et al. (2000) corresponds to the probability that another random sample X would generate a more extreme value of a test statistic than

that which was observed and thus is a statistic more consistent with the classical *P* value. Those authors and others have tried to develop a Bayesian *P* value that strictly reflects the data and not the prior. The Bayesian plausibility value reflects the Bayesian point of view that the prior is also information about the parameters, and thus, it is appropriate to use such information when deciding when to include a parameter in the model.

5. Regression Diagnostics

This section develops goodness of fit and diagnostic statistics for a Bayesian GLS regression analysis. First, a pseudo adjusted R^2 is proposed, and analysis of variance (ANOVA) for a GLS analysis is discussed, along with two other diagnostic EVRs, leverage, influence, and σ-influence statistics.

5.1. R^2 and Analysis of Variance

The traditional R^2 statistic measures the extent to which a model explains the variability in the data. It uses the partitioning of the sum of squared deviations and associated degrees of freedom to describe the variance of the signal versus the model error. Traditionally, for OLS regression, the total sum of squared deviations about the mean (SST) is divided into two separate terms, the sum of squared errors explained by the regression model (SSR) and the residual sum of squared errors (SSE), where SST = SSR + SSE. The coefficient of determination R^2 and the adjusted R^2 (denoted $\overline{R}^2$) both describe the fraction of the total variability the model explains, computed as

$$R^2 = \frac{SST - SSE}{SST} = 1 - \frac{SSE}{SST}, \tag{19}$$

$$\overline{R}^2 = 1 - \frac{\mathrm{SSE(n-k-1)}}{\mathrm{SST(n-1)}} = 1 - \frac{S_\varepsilon^2}{S_y^2}.$$

Here n is the total number of observations, and k is the number of covariates used in the regression model in addition to a constant.

For WLS and GLS analyses, these formulas do not provide the intended insight. The error of most concern is the model error variance because the sampling error is unexplainable and represents noise that complicates the analysis. A new measure is needed in which the sampling error variance is separated from the total error variance, leaving behind the fraction of the variance accounted for by the model and by the model error.

Liu et al. (2005) and Han et al. (2009) use negative binomial regression to explain variations in the expected number of power distribution system failures. Their observed discrete count data include sampling variability, as does the $\hat{\mathbf{y}}$ in our example. Their pseudo R^2 value describes the fraction of the true variability in the expected number of failures that their regression model explains.

The natural estimator for WLS and GLS models of the expected value of the sum of squared model errors would be $n\sigma_\delta^2(k)$ for a model with k covariates plus a constant. A corresponding estimator of the expected total variation in the true y_i values, corresponding to SST neglecting sampling error, is $n\sigma_\delta^2(0)$ where $\sigma_\delta^2(0)$ is the model error variance estimate when no explanatory variables are included but sampling errors are correctly deducted. Then, a pseudo coefficient of determination describing the fraction of the expected variability in the true y_i values that is explained by the model equals

$$R_\delta^2 = \frac{n\left[\sigma_\delta^2(0) - \sigma_\delta^2(k)\right]}{n\sigma_\delta^2(0)} = 1 - \frac{\sigma_\delta^2(k)}{\sigma_\delta^2(0)}. \tag{20}$$

We published this idea earlier in Reis (2005) and further refined the proposal in Gruber et al. (2007) and Griffis and Stedinger (2007b). In actual practice, the Bayesian mean values of $\sigma_\delta^2(k)$ and $\sigma_\delta^2(0)$ are employed. R_δ^2 is a direct extension of $\overline{R}^2$ in that it uses the ratio of unbiased estimators of the variance of the error δ and the variance of y. However, a critical difference is that $\overline{R}^2$ is based upon the sample variance of the observed y_i and the computed residual errors $\hat{\varepsilon}_i$. Here, R_δ^2 is based upon the estimated variance of the unobserved y_i and of the unobserved δ_i values, so it is called a pseudo-R^2. If $\sigma_\delta^2(k) = 0$, then $R_\delta^2 = 1$ as it should.

Table 1
Pseudo ANOVA Table for WLS and GLS Regression Analyses

Source	Degrees of freedom	Sum of squares
Model	k	$n\left[\sigma_\delta^2(0)-\sigma_\delta^2(k)\right]$
Model error, δ_i	$n-k-1$	$n\left[\sigma_\delta^2(k)\right]$
Sampling Error, η_i	n	$\sum_{i=1}^{n} Var(\hat{y}_i)$
Total	$2n-1$	$n\left[\sigma_\delta^2[0]\right]+\sum_{i=1}^{n} Var(\hat{y}_i)$

$EVR = \dfrac{\sum_{i=1}^{n} Var(\hat{y}_i)}{n\left[\sigma_\delta^2[k]\right]}$

$MBV^* = \dfrac{Var\left[b_0^{WLS}|GLS\ analysis\right]}{Var\left[b_0^{WLS}|WLS\ analysis\right]} = \dfrac{w^T\Lambda w}{w^T v}$ where $w_i = 1/\Lambda_{ii}$, $\upsilon = (n \times 1)$ vector of 1's.

If no explanatory variables are employed, the R_δ^2 will be 0, as is expected. R_δ^2 fairly compares different WLS or different GLS models with varying numbers of parameters using the regional mean model ($k = 0$) as the base case.

Table 1 presents a pseudo ANOVA table for WLS or GLS. This table describes how much of the variation in the observations can be attributed to the model and how much to model error and sampling error, respectively. We can describe the total sampling error sum of squares by its mean value, which is $tr[\mathbf{\Sigma}(\widehat{\mathbf{y}})]$, where $tr[\mathbf{A}]$ is the trace of the matrix $\mathbf{A}$. As noted above, because there are n observations, the total variation due to the model error $\boldsymbol{\delta}$ for a model with k parameters has a mean equal to $n\sigma_\delta^2(k)$. These provide descriptions of two of the three sources of variation. This is called a pseudo ANOVA because the contributions of the three sources of error are estimated or constructed, rather than being determined from the computed residual errors and the observed model predictions. The impact of correlation among the sampling errors is ignored.

If the sampling variance is not small in comparison to the model error variance, a WLS or GLS analysis is more appropriate than an OLS analysis. The EVR provides a measure of the relative importance of the sampling error compared to the model error (Griffis & Stedinger, 2007b; Gruber et al., 2007). Thus, it provides an indication of the need for a WLS or GLS analysis. Based on Table 1, the *EVR* is defined as

$$EVR = \frac{SS(\text{sampling error})}{SS(\text{model error})} = \frac{tr[\Sigma(\widehat{y})]}{n\sigma_\delta^2}. \tag{21}$$

As a rule of thumb, when the *EVR* is greater than 20%, one should employ a WLS or GLS analysis as opposed to OLS. If the EVR is less than 10%, the OLS results should be close to the WLS or GLS results depending upon the heterogeneity of the errors.

Although EVR distinguishes between the need for an OLS versus a WLS/GLS analysis, it does not determine whether a GLS regression is needed to address cross correlation. Thus, the misrepresentation of the beta variance (MBV) statistic was developed to determine whether a WLS regression is sufficient or if a GLS regression is needed (Griffis & Stedinger, 2007b). The MBV describes the error made by a WLS regression analysis in its evaluations of the precision of b_0^{WLS}, which is the estimator of the constant β_0. Covariance among the estimated y_i's generally has its greatest impact on the precision of the constant term (Stedinger & Tasker, 1985) and zero-one regional indicator variables. However, as stated in Veilleux (2011), the weights used by Griffis and Stedinger (2007b) to calculate the MBV are not the correct weights for determining the error made by a WLS regression error analysis in its evaluation of the precision of b_0^{WLS}. MBV was computed using weights that are the inverse of the standard deviation. Instead, a correct WLS analysis would weight each observation by the inverse of the variance. Thus, a corrected MBV is defined below as *MBV** using the correct weights (Veilleux, 2011).

$$MBV^* = \frac{Var\left[b_0^{WLS}|GLS\ analysis\right]}{Var\left[b_0^{WLS}|WLS\ analysis\right]} = \frac{w^T\Lambda w}{w^T v} \quad \text{where} \quad w_i = \frac{1}{\Lambda_{ii}}, \tag{22}$$

where ν is an (n × 1) vector of ones. If *MBV** is substantially larger than 1, then the GLS estimate of the variance of the constant term will be that much larger than the value provided by WLS. If all sites in the region have the same record length n, all records are concurrent, and all cross correlations among the $\widehat{y}_i$ are equal to ρ_η, *MBV** would be

$$MBV^* = 1 + (n-1)\rho_\eta \frac{EVR}{EVR+1}. \tag{23}$$

This special case shows the critical importance of the number of sites n in an analysis and the cross-correlation ρ_η of the at-site estimators $\widehat{y}_i$. This indicates that in the case of our skew regression example, wherein $EVR = Var(\widehat{y})/\sigma_\delta^2$ is generally greater than 2 or 3, the precision of our constant estimator will be particularly sensitive to cross-correlation ρ_η and GLS is likely to be required to obtain statistically valid results.

5.2. Leverage and Influence

Leverage, as adopted by Tasker and Stedinger (1989, Equation (23)), considers whether an observation or x value is unusual and thus likely to have a large effect on the estimated regression coefficients (Cook & Weisberg, 1982). How to measure leverage can be problematic. It is not clear how to describe how large a change in different residuals should be considered when model errors are heteroscedastic. The first leverage measure, suggested by Tasker and Stedinger (1989, Equation (23)), considers the effect of a unit change in each residual. If all the residuals have the same units and precision, then this is an appropriate measure of the effect of equivalent errors in the different observations. Thus, this leverage measures the marginal/unit impact of the residuals ε_i on the estimated y_i values. In a Bayesian context, as shown in Reis (2005), the leverage measure described by Tasker and Stedinger (1989) becomes

$$leverage\,(i) = \frac{\partial \widehat{y}_i}{\partial \varepsilon_i} = \underset{\sigma_\delta^2}{E}\left[\mathbf{x}_i\left(\mathbf{X}^T\Lambda\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\Lambda\left(\sigma_\delta^2\right)^{-1}e_i\right], \tag{24}$$

where $\widehat{y}_i$ is the estimator of y_i associated with $\boldsymbol{x}_i$ and e_i is a unit column vector with 1 at the ith row and 0 otherwise. Tasker and Stedinger (1989) show that the average value of this leverage statistic is (k + 1)/n. Generally, $2(k+1)/n$ is considered to be a large value.

This leverage statistic seems appropriate when lack of fit can be described by errors of the same magnitude. For example, this would be appropriate when modeling skew coefficients or the log 100-year flood estimates at different sites, as opposed to when some equations describe head and others flow characteristics in a groundwater model (Yager, 1998).

The second measure of leverage introduced here considers not a unit change in each residual but a change proportional to the standard deviation of that residual. Thus, this measure considers the likely statistical variation in each ε_i and the effect of such variation. Statistical leverage is computed as

$$S\text{–}leverage\,(i) = \omega(k+1)\frac{\partial \widehat{y}_i}{\partial \varepsilon_i}\sigma_{\varepsilon_i} = \omega(k+1)\left(leverage\,(i)\cdot\sqrt{\Lambda\left(\sigma_\delta^2\right)_{ii}}\right), \tag{25}$$

where $\omega = 1/\sum_{j=1}^{n} leverage\,(i)\cdot\sqrt{\Lambda\left(\sigma_\delta^2\right)_{jj}}$.

As it is defined, the average value of *S-leverage* is also equal to $(k+1)/n$. Twice the average value, $2(k+1)/n$, is considered to be a large value.

Statistical leverage is an appropriate statistic to consider when the concern is with the likely effect on the regression of probabilistic variation in each residual. Because the GLS weights depend upon the statistical precision of each ε_i, the leverage in (24) for a point often increases as the at-site record length increases because of the greater weight assigned to the observation, whereas statistical leverage in (25) is less dependent on record length. If an observation has small leverage, then, given its anticipated statistical precision and the leverage associated with the corresponding $\mathbf{x}$, the observation is unlikely to have any effect on estimated model parameters. The leverage in (24) may be more appropriate when one is concerned with the impact of gross errors in a model's structure, but it does not correct for differences in units among the ε_i. Examples in the application section illustrate the use of these statistics.

A third measure of leverage was developed for use with ROI regression. ROI creates a unique region, or set of gauged basins, for each ungauged basin of interest. The ROI regression can be used to predict hydrologic quantities such as flood quantiles for the target basin (Eng, Stedinger, et al., 2007). ROI leverage is computed as (Eng, Stedinger, et al., 2007)

$$ROI\text{–}leverage(i) = \frac{\partial \widehat{y}_0}{\partial \varepsilon_i} = \left[\mathbf{x}_0\left(\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\right]e_i, \tag{26}$$

where $\mathbf{x}_0$ is a vector of basin attributes for the ungauged sites of interest. ROI leverage measures the impact on the estimate of y_0 at site 0 with $\boldsymbol{x} = \boldsymbol{x}_0$ of a unit error e_i at other sites i, $i = 1 \ldots, n$. In an ROI regression, one would like all the ROI leverage statistics to be positive with about the same value.

Unlike leverage that highlights points, which are likely to affect the fit of the regression, influence describes those points that did have an unusual impact on the regression analysis. An influential observation is one with an unusually large residual that had a disproportionate effect on the fitted regression relationships. Influential observations often have high leverage. The following influence measure proposed by Tasker and Stedinger (1989) is based on Cook's D (Clarke, 1994; Cook & Weisberg, 1982),

$$D_i = \frac{K_{ii}\widehat{\varepsilon}_i^2}{(k+1)(\lambda_{ii}-K_{ii})}, \tag{27}$$

where $(k+1)$ is the dimension of β, λ_{ii} are the diagonal elements of Λ, and K_{ii} are the diagonal elements of

$$\boldsymbol{K} = \boldsymbol{X}\left(\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{\Lambda}^{-\boldsymbol{1}}\boldsymbol{X}\right)^{-\boldsymbol{1}}\boldsymbol{X}^{\boldsymbol{T}}, \tag{28a}$$

so that $\lambda_{ii} - K_{ii}$ is the variance of $\widehat{\varepsilon}_i$, as demonstrated in Appendix A. Tasker and Stedinger (1989) suggested that influence is large when D_i is greater than $4/n$ where n is the number of sites. In a Bayesian analysis, to be correct, one should employ $\mathbf{K_B}$, the average value of (28a),

$$\mathbf{K_B} = \mathbf{X}\underset{\sigma_\delta^2}{E}\left[\left(\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\right]\mathbf{X}^T. \tag{28b}$$

A second measure of influence, σ-influence (Gruber et al., 2007), determines if any observation had an unusual impact on the estimated model error variance. In developing regional skew models, for example, the model error variance is very important because it determines the weight placed on the regional skew relative to the at-site estimator. The σ-influence statistic describes the relative influence of each observation on the estimated model error variance. The influence statistic D_i described above identifies those observations with significant influence on the model predictions. D_i does not necessarily describe whether the point has a significant influence on the estimated model error variance. The σ-influence is calculated as

$$\sigma\text{-}Influence_i = \frac{2\sum_{j=1}^{n}\widehat{\varepsilon}_i\left(\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\right)_{ij}\widehat{\varepsilon}_j}{\sum_{i=1}^{n}\sum_{j=1}^{n}\widehat{\varepsilon}_i\left(\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\right)_{ij}\widehat{\varepsilon}_j} = \frac{2\widehat{\varepsilon}_i\left(\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\widehat{\varepsilon}\right)_i}{\widehat{\varepsilon}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\widehat{\varepsilon}}. \tag{29}$$

Here the standardized sum-of-squares $\widehat{\boldsymbol{\varepsilon}}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\widehat{\boldsymbol{\varepsilon}}$ used to compute the likelihood function for the data (and the generalized MM model error variance in Stedinger & Tasker, 1985) is divided among the n different sites. By construction, the average value of σ-influence is $2/n$, where n is the number of sites in the regression; thus, σ-influence values greater than $4/n$ are considered to be large, as is the case with D_i.

6. Application: Regional Skew Estimation

Bulletin 17C (England et al., 2018) recommends the use of the log-Pearson type III distribution for flood frequency analysis. The data available at a given site are often short, thus providing an inaccurate estimate of the skewness coefficient. In order to improve the precision of the skewness estimator, Bulletin 17C recommends combining a regional skew with the at-site skew (Hardison, 1975; Tasker, 1978; McCuen, 1979, 2001; IACWD, 1982; Stedinger et al., 1993; Griffis et al., 2004). Griffis and Stedinger (2009, Appendix A) show that the Bulletin 17B (IACWD, 1982) mean squared error (MSE) weighted skewness estimator results in the estimator with the smallest MSE provided that the regional skew is unbiased and independent of the at-site skew estimator. Griffis and Stedinger (2007a, 2009) illustrate the value of a good regional skewness estimator in terms of the precision of flood quantile estimates. McCuen and Smith (2008) developed a regression model for the skewness coefficient based on rainfall skew and basin storage variables. However, it is unclear whether they are addressing log-space or real-space skewness. Reis et al. (2005) evaluated regional skew for 17 sites in the Tibagi River basin and for 44 sites in the Muskingum River basin using different statistical models. For this study, regional skew was evaluated for the Piedmont region in the Southeastern U.S. with 92 stations with record lengths varying from 19 to 108 years. We compare the results of OLS, WLS, and GLS

Table 2
Skew Regression for the Piedmont Region in the Southeastern U.S. (Number of Sites = 92)

	Regression parameters		Regression diagnostics			
Model	Const	HSI	σ^2_δ	*ASV*	AVP_{new}	R^2_δ
MM-OLS	−0.109	0.869	0.279	0.006	0.285	0.108
	(0.055)	(0.263)				
	(5.3%)	(0.2%)				
MM-WLS	−0.095	0.814	0.086	0.006	0.092	0.253
	(0.052)	(0.252)				
	(7.3%)	(0.2%)				
B-WLS	−0.095	0.821	0.085	0.005	0.090	0.201
	(0.050)	(0.242)				
	(4.7%)	(0.1%)				
MM-GLS	−0.095	-	0.096	0.010	0.106	0.0
	(0.098)	-				
	(33.8%)	-				
B-GLS	−0.094	-	0.080	0.009	0.090	0.0
	(0.097)	-				
	(33.2%)	-				
MM-GLS2	−0.096	0.588	0.090	0.014	0.103	0.137
	(0.098)	(0.304)				
	(33.3%)	(6.7%)				
B-GLS2	−0.095	0.557	0.080	0.013	0.094	0.0004
	(0.097)	(0.299)				
	(32.6%)	(6.1%)				

Note. Table reports best models in terms of minimum average variance of prediction (AVP) for a new site. Standard errors, classical (OLS and MM) and Bayesian estimates, and Bayesian plausibility values (%) are presented in parentheses. Average sampling variance (ASV) is reported for each model.

regional regression analyses using MM and Bayesian model error variance estimators. The USGS conducted many regional skew studies, and citations are provided in section 1.

The following basin characteristics were available for use in the regression models: (1) drainage area, A, in squared miles; (2) main channel slope in feet per mile; (3) average basin slope in percentage; (4) length of the main channel in miles; (5) basin perimeter length in miles; (6) basin shape factor (BSF), defined as the ratio between the perimeter squared and the drainage basin; (7) the mean basin elevation in feet above sea level; (8) the maximum basin elevation in feet above NAVD88; (9) the minimum basin elevation in feet above NAVD88; (10) mean annual precipitation; (11) fraction of impervious surface area expressed in percentage of drainage area; (12) forest cover expressed as percentage of drainage area; (13) mean soil drainage index, which varies from 1 to 7, with 1 denoting excessively drained soil; and (14) mean hydrologic soil index (HSI), which varies from 1 to 4, with 1 denoting high infiltration rate.

6.1. The Sampling Covariance Matrix (Σ)

Estimates of $\sigma^2_{\eta i}$ and $\widehat{\rho}\left(\widehat{y}_i, \widehat{y}_j\right)$ are required. Griffis and Stedinger (2009) provide an accurate approximation of $\sigma^2_{\eta i}$ that, with the skewness estimator unbiasing factor in Tasker and Stedinger (1986), equals:

$$Var(G) = \left[1 + \frac{6}{m}\right]^2 \left(\frac{6}{m} + a(m)\right)\left(1 + \left\{\frac{9}{6} + b(m)\right\}\gamma^2 + \left\{\frac{15}{6*8} + c(m)\right\}\gamma^4\right), \tag{30}$$

wherein

$$a(m) = -\frac{17.75}{m^2} + \frac{50.06}{m^3},\ b(m) = \frac{3.92}{m^{0.3}} - \frac{31.1}{m^{0.6}} + \frac{34.86}{m^{0.9}},\ \text{and}\ c(m) = -\frac{7.31}{m^{0.59}} + \frac{45.9}{m^{1.18}} - \frac{86.5}{m^{1.77}}.$$

Table 3
Sensitivity Analysis for the B-GLS Model

	Regression parameters		Regression diagnostics			
Model	Constant	BSF	σ_δ^2	*ASV*	AVP_{new}	R_δ^2
Without Site 40						
B-GLS	−0.111	-	0.066	0.009	0.075	0.000
	(0.096)	-				
	(24.6%)	-				
Without Site 87						
B-GLS	−0.126	0.031	0.056	0.012	0.068	0.171
	(0.145)	(0.015)				
	(18.9%)	(3.8%)				
Without Site 70						
B-GLS	−0.078	-	0.070	0.009	0.079	0.000
	(0.097)	-				
	(41.8%)	-				

Note. Table reports best models in terms of minimum average variance of prediction (AVP) for a new site. Standard errors and Bayesian plausibility values (%) are presented in parentheses. Average sampling variance (ASV) is reported for each model.

Here m is the sample size, and γ is the true value of skew. The factor $[1+6/m]^2$ in Equation (30) should be employed when the bias correction factor proposed by Tasker and Stedinger (1986) is used in the estimation of the at-site skew, which is computed as

$$G = \left[1 + \frac{6}{m}\right] \frac{m\sum_{t=1}^{m}(u_t - \overline{u})^3}{(m-1)(m-2)s^3}, \tag{31}$$

where u_t is the logarithm of the annual peak flows in year t and s is the sample standard deviation of u_t. Because the true values of skews at each site are unknown, the regional mean of the skews is used in Equation (30).

Martins and Stedinger (2002) express the intersite correlation coefficient between two G_i in terms of the intersite correlation coefficient ρ_{ij} between concurrent flows as

$$\widehat{\rho}(G_i, G_j) = \left(cf_{ij}\right)\rho_{ij}^{\kappa}, \qquad cf_{ij} = m_{ij}/\sqrt{(m_{ij}+m_i)(m_{ij}+m_j)}, \tag{32}$$

where the exponent κ depends upon the regional value for γ and is equal to 2.8 for $\gamma = 0$, m_{ij} is the common record period, and m_i and m_j are the extra observation period for stations i and j, respectively. The factor (cf_{ij}) accounts for the sample size differences between stations and the concurrent record length: It worked reasonably well for the skew coefficient in the range considered $(-1 \le \gamma \le 1)$.

The use of sample estimates of ρ_{ij} may result in a covariance matrix $\mathbf{\Lambda}(\sigma_\delta^2)$ that is not positive definite due to sampling uncertainties and variations in concurrent record lengths (Tasker & Stedinger, 1989). Therefore, one can use a smoothed estimate of ρ_{ij} that depends on the distances between any two stations $\rho(d_{ij})$. For this area in the Piedmont region, a good model was

$$\rho(d_{ij}) = \theta^{\left(\frac{\tau d_{ij}}{\alpha d_{ij}+1}\right)}, \tag{33}$$

wherein $\theta = 0.993$, $\alpha = 0.00989$, and $\tau = 2.78$; d_{ij} is the distance between sites in miles.

Tables 2 and 3 present the results of the analysis for the Piedmont region in the Southeastern U.S. The following sections describe the results in detail.

Table 4
Pseudo ANOVA Table for the Piedmont Region in the Southeastern U.S. (B-GLS)

Source	Degrees of freedom			Sum of squares			
	Case 1	Cases 2 and 4	Case 3	Case 1 (all sites)	Case 2 (w/o 40)	Case 3 (w/o 87)	Case 4 (w/o 70)
Model	$k = 0$	$k = 0$	$k = 1$	0.0	0.0	1.1	0.0
Model error	$n - k - 1 = 91$	$n - k - 1 = 90$	$n - k - 1 = 89$	7.4	6.0	5.1	6.4
Sampling error	$n = 92$	$n = 91$	$n = 91$	16.9	16.8	16.8	16.8
Total	$2n - 1 = 183$	$2n - 1 = 181$	$2n - 1 = 181$	24.3	22.8	23.0	23.2
EVR				2.29	2.79	3.29	2.64
*MBV**				1.90	1.91	1.94	1.91
R_δ^2				0.000	0.000	0.171	0.000

6.2. OLS Regression Models

As expected, the MM-OLS model error variance estimate is much larger than those obtained with WLS and GLS. The OLS model with minimum AVP for a new site is composed of a constant and the HSI (MM-OLS in Table 2). As shown in Table 2, the model error variance, σ_δ^2, for MM-OLS is 0.279 with an AVP_{new} of 0.285. These values are much larger than any of the results obtained using WLS or GLS analysis.

The exaggerated variance of predictions occurs because the OLS regression analysis does not make any distinction between the variance due to the model error and the variance due to time sampling error. This concept is reinforced when viewing the EVR results from the pseudo ANOVA table (Table 4). The EVR for Case 1 is very large, 2.29, implying that the variation due to sampling error is over twice as large the variation due to model error. Clearly, a WLS or GLS model should be employed rather than an OLS model.

6.3. WLS Regression Models

In the case of the weighted least squares regression, both the MM and Bayesian estimators were used. Like the best OLS regression model, the best WLS model in terms of AVP_{new} also employs a constant and HSI as explanatory variable regardless of the estimator used (MM-WLS and B-WLS in Table 2). MM-WLS and Bayesian WLS estimates were very similar. The model MM-WLS (B-WLS), as presented in Table 2, has an AVP_{new} of 0.092 (0.090) and a model error variance equal to 0.086 (0.085). Both values are much smaller than those obtained with MM-OLS.

In order to determine if a WLS analysis is sufficient, the misrepresentation of beta variance (*MBV**) can be consulted. As shown in Table 4, Case 1 has an *MBV** of 1.9. These results clearly suggest that the correlation among estimators of skew in these regions should not be neglected by using WLS; otherwise, the model error variance as well as the AVP of the regional skew will be underestimated.

6.4. GLS Regression Models

In the case of GLS regression, both the MM and Bayesian estimators were tested. Unlike the best models obtained when OLS and WLS models are used, which consisted of a constant and the HSI as explanatory variables, the best GLS model in terms of AVP_{new} is a simple regional mean regardless of the estimator used in the analysis (MM-GLS and B-GLS in Table 2). In the GLS case, the HSI should not be used as explanatory variable (see MM-GLS2 and B-GLS2 models in Table 2) because its coefficient is either no longer significant at 5% level, in the case of MM-GLS2 model, or 0 is contained in its 95% credible interval, in the case of B-GLS2 model, as suggested by its Bayesian plausibility values of 6.1%. But that is not the only reason for choosing a simple regional mean model. The AVP_{new} of MM-GLS2 is almost the same as that of the regional mean, meaning the increase in model complexity yields no increase in precision and thus does not appear to be justifiable. In the Bayesian case, the situation is even more clear because the AVP_{new} of the B-GLS2 is larger than that of the regional mean model.

Although the MM-GLS and B-GLS regional mean models have almost identical parameter estimates, their model error variance and AVP_{new} are different. The B-GLS's model error variance and AVP for a new site are, respectively, 17% and 13% smaller than those associated with the MM-GLS.

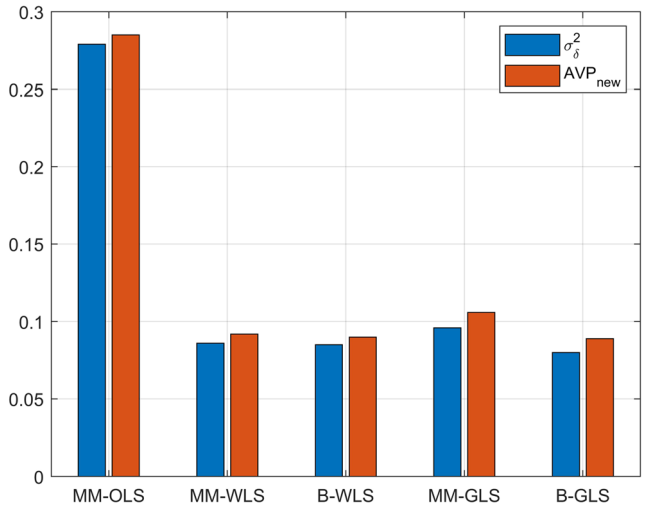

Figure 1. Model error variance (σ_δ^2) and average variance of prediction for a new site (AVP_{new}) for the Piedmont region in the Southeastern U.S.

Figure 1 provides an overall view of the differences among models (OLS, WLS, and GLS) and estimators (MM and Bayesian) with respect to model error variance and AVP for a new site, AVP_{new}. These two metrics are key for model selection and weights on regional and at-site estimates of log-space skews.

OLS, WLS, and GLS models are based on different assumptions about the data, so it is not a surprise that they provide different interpretations on how the skew varies over the region. The largest difference is the precision of regional estimates. One can clearly see that OLS estimates of σ_δ^2 and AVP_{new} are much larger than those obtained by WLS and GLS analyses. WLS models neglect the presence of cross correlation, resulting in estimates of σ_δ^2 and AVP_{new} that are different than those obtained by GLS models. The Bayesian estimator provided models with smaller σ_δ^2 and AVP_{new} than those obtained with MM estimator. A Bayesian estimator is more appropriate because it is based on the full posterior distribution of both the parameters and the error variance of the regional model, describing in a more reasonable fashion the possible values of the model error variance.

6.5. Sensitivity Analysis and Diagnostic Statistics for B-GLS Models

This section offers a diagnostic analysis of the best B-GLS model for the Piedmont region data in the Southeastern U.S. Figure 2 presents the leverage and influence results for the B-GLS model for the 20 sites with largest influence. The sites are ordered by decreasing influence, as it identifies those sites that had a large impact on the analysis.

Site 40 has the highest influence and σ-influence values, implying that it has a large impact on the model error variance. Site 40 was removed from the B-GLS analysis as a test (Case 2 in Table 4). As expected, σ_δ^2 of B-GLS (without site 40) decreased from 0.080 to 0.066, a 25% reduction, as shown in Table 3. Site 40 has such a large impact on the regression analysis because it has a large positive at-site skew value of 1.50 and the second largest residual of 1.60, although it has 47 years of record length, which is close to the average value of record lengths in the region. Other sites have also large at-site skews and large residuals, but they are not as influential as Site 40 because they have smaller record lengths. For instance, Site 18 has the most extreme at-site skew and the largest residual, but it is only the fifth most influential because it has a short record length of just 26 years, the fifth shortest among the 92 sites used in the study. Site 2 is also an example of how the record length affects the degree of influence of a site in the analysis. It has the second largest at-site skew and the third largest residual, but it is only the ninth most influential site because of its short record length of 25 years.

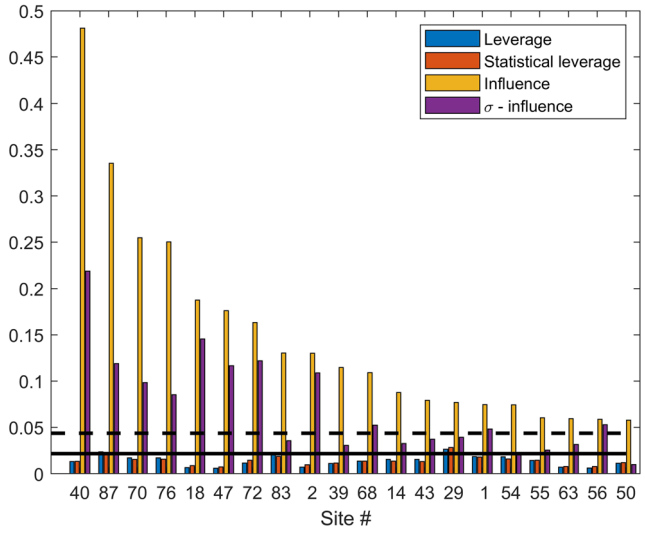

Figure 2. Regression diagnostics: leverage and influence for the Piedmont region in the Southeastern U.S. B-GLS model. The solid, horizontal line represents the threshold for high leverage and high statistical leverage, while the dashed, horizontal line represents the threshold for high influence and high σ-influence.

We also removed Site 87 from the B-GLS model as a test (Case 3 in Table 4). Although Site 87 has only the eighth largest at-site skew and the sixth largest residual, it is the second most influential because it has a record length of 70 years. As shown in Table 3, when Site 87 is removed, the best Bayesian GLS model includes the basin shape factor as explanatory variable, resulting in a new σ_δ^2 of 0.056, a 30% reduction when compared to the regional mean model based on the 92 sites.

The last action taken based on the sensitivity analysis run in the Piedmont region data was the removal of Site 70 (Case 4 in Table 4), which had the third largest influence in the study, the seventh at-site skew, and the eight largest residual. As shown in Table 3, when Site 70 is removed from the B-GLS model, there is a decrease in σ_δ^2 from 0.080 to 0.070.

Table 4 contains the pseudo ANOVA results for both the B-GLS model and the three sensitivity analyses. The pseudo ANOVA table clearly demonstrates that for all four cases in the Piedmont region, the sampling error is at least twice as large as the model error, $EVR \geq 2.29$.

7. Conclusions

Reis et al. (2005) developed quasi-analytic Bayesian model error variance and $\boldsymbol{\beta}$ estimators for a GLS regression model. That analysis is the foundation of a new operational and comprehensive GLS regional hydrologic statistical methodology. New regression diagnostic statistics for B-WLS and B-GLS models include pseudo analysis of variance tables, a pseudo adjusted R^2, and the EVR and MBV quantifying the need for a WLS or GLS model. Bayesian leverage, influence, and σ-influence help identify rogue and influential data points and thus help complete the framework.

Regionalization of the log-space skew, the shape parameter of the log-Pearson type III distribution in the Southeastern U.S., illustrates the methodology. Results obtained from OLS, WLS, and GLS analyses were compared, as well as the results using the Bayesian and MM model error variance estimators. The OLS analysis provides misleading results because it does not make a distinction between the variance due to the model error and the variance due to time sampling error in estimated at-site skewness coefficients. GLS was the best framework because the cross correlation of the skewness estimators, which is neglected by WLS, proved to be important in the estimation of the standard errors of parameters, which guides model selection. This example demonstrates that the true model error variance for regional skew models is on the order of 0.10 or less.

An analysis of the best B-GLS models illustrates that the roles of record length, skew, and residual values play in the estimation of the model error variance and AVP, which affects model selection and the regression coefficients. One sees how leverage, influence, and σ-influence statistics can be useful in identifying stations that actually did have a significant impact on the analysis. The new B-GLS framework provides an operational, statistically rigorous, and comprehensive methodology for analysis of hydrologic regional hydrologic information consisting of cross-correlated streamflow records of different length from stations in a region; regression diagnostics correctly attribute variability to different sources (variation explained by time sampling error, the model, and model error and total sampling error) and identify potentially influential and rogue observations.

Appendix A.

This appendix provides a clean and consistent derivation of key expressions employed to compute the AVP for both new and old sites, as well as leverage and influence.

A.1. Variance of the Residuals

Let $\mathbf{y}$ be the vector of the true value of the statistic of interest, $\widehat{\mathbf{y}}$ the at-site estimate, and $\widehat{\mathbf{y}}_p$ the prediction of $\mathbf{y}$ produced by a fitted regression model. Then, the residual vector is given by

$$\widehat{\boldsymbol{\varepsilon}} = \widehat{\mathbf{y}} - \mathbf{y}_p = \widehat{\mathbf{y}} - \mathbf{X}\mathbf{b} = \widehat{\mathbf{y}} - \mathbf{X}\left(\mathbf{X}^{\mathrm{T}}\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\widehat{\mathbf{y}} = (\mathbf{I} - \mathbf{H})\widehat{\mathbf{y}}, \tag{A.1}$$

wherein the GLS hat matrix $\mathbf{H}$ is defined as

$$\mathbf{H} = \mathbf{X}\left(\mathbf{X}^{T}\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^{T}\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}. \tag{A.2}$$

That the leverages in Equation (24) are the average values of the diagonal elements h_{ii} of $\mathbf{H}$ follows from $\widehat{\mathbf{y}}_p = \mathbf{H}\widehat{\mathbf{y}}$. The average of the n leverage values $\mathrm{tr}[\mathbf{H}]/n$ equals $(k+1)/n$ because

$$\begin{aligned}\mathrm{tr}[\mathbf{H}] &= \mathrm{tr}\left[\mathbf{X}\left(\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\right] \\ &= \mathrm{tr}\left[\left(\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right] = \mathrm{tr}[\mathbf{I}_{k+1}] = (k+1).\end{aligned} \tag{A.3}$$

For the OLS case, the hat matrix $\mathbf{H}$ has simply the traditional value $\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$.

Substituting the equality $\widehat{\mathbf{y}} = \mathbf{X}\mathbf{b} + \varepsilon$ into (A.1), the computed residuals can be written as a function of the total error $\boldsymbol{\varepsilon}$:

$$\widehat{\varepsilon} = (\mathbf{I}-\mathbf{H})\varepsilon. \tag{A.4}$$

Given $E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T] = \boldsymbol{\Lambda}\left(\sigma_\delta^2\right)$, one finds that $\mathbf{H}\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)\mathbf{H}^T = \boldsymbol{\Lambda}\left(\sigma_\delta^2\right)\mathbf{H}^T$ as can be demonstrated by substitution for $\mathbf{H}$ the expression in (A.2). Thus, the covariance matrix of the estimated residuals is

$$\begin{aligned}E\left[\widehat{\boldsymbol{\varepsilon}}\widehat{\boldsymbol{\varepsilon}}^T\right] &= E\left[(\mathbf{I}-\mathbf{H})\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T(\mathbf{I}-\mathbf{H}^T)\right] = E\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T-\mathbf{H}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T-\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\mathbf{H}^T+\mathbf{H}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\mathbf{H}^T\right] \\ &= \boldsymbol{\Lambda}\left(\sigma_\delta^2\right)-\mathbf{H}\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)-\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)\mathbf{H}^T+\mathbf{H}\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)\mathbf{H}^T = (\mathbf{I}-\mathbf{H})\boldsymbol{\Lambda}\left(\sigma_\delta^2\right) = \boldsymbol{\Lambda}\left(\sigma_\delta^2\right)-\mathbf{K},\end{aligned} \tag{A.5}$$

which is clearly symmetric and where $\mathbf{K}$ is defined in Equations (28a) and (28b). For every σ_δ^2, the mean of $\widehat{\boldsymbol{\varepsilon}}$ is 0, so in a Bayesian framework, the covariance of $\widehat{\boldsymbol{\varepsilon}}$ is $E\left[\widehat{\boldsymbol{\varepsilon}}\widehat{\boldsymbol{\varepsilon}}^T\right] = E\left[\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)-\mathbf{K}\right]$.

Similarly, the variance of β for a given $\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)$ is obtained by noting that

$$\begin{aligned}\mathbf{b} &= \left(\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\widehat{\mathbf{y}} \;= \left(\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}(\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}) \\ &= \boldsymbol{\beta} + \left(\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\boldsymbol{\varepsilon}.\end{aligned} \tag{A.6}$$

Thus, for given $\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)$,

$$\begin{aligned}E\left[(\boldsymbol{\beta}-\mathbf{b})(\boldsymbol{\beta}-\mathbf{b})^T\right] &= E\left[\left(\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\left(\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\right] \\ &= \left(\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}.\end{aligned} \tag{A.7}$$

A.2. Variance of Prediction

The vector of differences between the true and predicted values is

$$\left(\mathbf{y}-\mathbf{y}_p\right) = \mathbf{X}\boldsymbol{\beta}+\boldsymbol{\delta}-\mathbf{X}\mathbf{b} = \mathbf{X}\boldsymbol{\beta}+\boldsymbol{\delta}-\mathbf{X}\left(\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}(\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}) = \boldsymbol{\delta}-\mathbf{H}\boldsymbol{\varepsilon}. \tag{A.8}$$

Thus, the covariance matrix for the prediction errors for given Λ is just

$$E\left[\left(\mathbf{y}-\mathbf{y}_p\right)\left(\mathbf{y}-\mathbf{y}_p\right)^T\right] = E\left[(\delta-\mathbf{H}\varepsilon)(\delta-\mathbf{H}\varepsilon)^T\right] = E\left[\delta\delta^T-\delta\varepsilon^T\mathbf{H}^T-\mathbf{H}\varepsilon\delta^T+\mathbf{H}\varepsilon\varepsilon^T\mathbf{H}^T\right]. \tag{A.9}$$

Noting that $\boldsymbol{\delta}$ for new sites and $\boldsymbol{\varepsilon}$ for old sites are uncorrelated, the covariance matrix for the predictions at new sites is simply

$$E\left[\left(\mathbf{y}-\mathbf{y}_p\right)\left(\mathbf{y}-\mathbf{y}_p\right)^T\right] = E\left[\delta\delta^T+\mathbf{H}\varepsilon\varepsilon^T\mathbf{H}^T\right]. \tag{A.10}$$

Therefore, substituting $E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T] = \Lambda$ with $\mathbf{H}$ from ((A.2)), the variance of prediction at a new site is given by

$$VP_{new}(i) = \sigma_\delta^2 + \mathbf{x}_i\left(\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\mathbf{x}_i{}^T. \quad (A.11)$$

However, if the predictions are made for those n old sites used in the regression, the covariance matrix of the predictions becomes

$$E\left[\left(\mathbf{y}-\mathbf{y}_p\right)\left(\mathbf{y}-\mathbf{y}_p\right)^T\right] = E\left[\boldsymbol{\delta}\boldsymbol{\delta}^T-2\boldsymbol{\delta}\boldsymbol{\delta}^T\mathbf{H}^T+\mathbf{H}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\mathbf{H}^T\right] = \sigma_\delta^2\mathbf{I}+\mathbf{H}\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)\mathbf{H}^T-2\sigma_\delta^2\mathbf{H}^T, \quad (A.12)$$

because the model error δ_t for the site is also part of the sampling error of the estimator.

Thus, the variance of prediction for an old site is

$$VP_{old}(i) = \sigma_\delta^2 + \mathbf{x}_i\left(\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\mathbf{x}_i^T-2\sigma_\delta^2\mathbf{x}_i\left(\mathbf{X}^T\boldsymbol{\Lambda}\left(\sigma_\delta^2\right)^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^Te_i, \quad (A.13)$$

wherein e_i is a column vector with one at the ith row and 0 otherwise. In a Bayesian analysis wherein σ_δ^2 is a random variable, one should employ the appropriate expected values in Equations (A.2), (A.7), (A.11), and (A.13), as in Equations (12)–17.

Acknowledgments
We thank the Brazilian government agency CNPq (Conselho Nacional de Desenvolvimento Cientifico e Tecnologico) who provided support for this research and Fundação de Apoio à Pesquisa do Distrito Federal (FAP-DF) for partially funding the sabbatical leave of the first author. We express our appreciation to Chuck Parrett (USGS), Ken Eng (USGS), Tim Cohn (USGS), and Veronica Webster (Michigan Technological University) for useful comments and encouragement. We appreciate support provided by a Water Resources Institute Internship Award #07HQAG0161 from the U.S. Geological Survey, U.S. Department of the Interior. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. government. Sample skewness coefficients and record lengths for different sites, concurrent record lengths for each pair of sites, the sampling covariance matrix of the skew estimates, and values of explanatory variables for the example in the paper are available in Cornell eCommons (https://doi.org/10.7298/X4348HHJ.2).

References

Archfield, S. A., Pugliese, A., Castellarin, A., Skoien, J. O., & Kiang, J. E. (2013). Topological and canonical kriging for design flood prediction in ungauged catchments: An improvement over a traditional regional regression approach? *Hydrology and Earth System Sciences*, *17*, 1575–1588. https://doi.org/10.5194/hess-17-1575-2013

Bayarri, M. J., & Berger, J. O. (2000). P-values for composite null models. *Journal of the American Statistical Association*, *95*, 1127–1142.

Bickel, P., & Doksum, K. (1977). *Mathematical statistics: Basic ideas and selected topics*. San Francisco, CA: Holden-Day.

Burn, H. B. (1990a). Evaluation of regional flood frequency analysis with a region of influence approach. *Water Resources Research*, *26*(10), 2257–2265. https://doi.org/10.1029/WR026i010p02257

Burn, H. B. (1990b). An appraisal of the "region of influence" approach to flood frequency analysis. *Hydrological Sciences Journal*, *35*(2), 149–165. https://doi.org/10.1080/02626669009492415

Carlin, B. P., & Louis, T. A. (2000). *Bayes and empirical Bayes methods for data analysis*. Boca Raton, FL: Chapman & Hall/CRC.

Castiglioni, S., Castellarin, A., & Montanari, A. (2009). Prediction of low-flow indices in ungauged basins through physiographical space-based interpolation. *Journal of Hydrology*, *378*(3-4), 272–280. https://doi.org/10.1016/j.jhydrol.2009.09.032

Castiglioni, S., Castellarin, A., Montanari, A., Skøien, J. O., Laaha, G., & Bloschl, G. (2011). Smooth regional estimation of low-flow indices: Physiographical space based interpolation and top-kriging. *Hydrology and Earth System Sciences*, *15*(3), 715–727. http://doi.org/10.5194/hess-15-715-2011

Chokmani, K., & Ouarda, T. B. M. J. (2004). Physiographical space-based kriging for regional flood frequency estimation at ungauged sites. *Water Resources Research*, *40*, W12514. https://doi.org/10.1029/2003WR002983

Clarke, R. T. (1994). *Statistical modeling in hydrology*. Chichester, UK: John Wiley & Sons Inc.

Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression*. New York, NY: Chapman and Hall.

Cunnane, C. (1988). Methods and merits of regional flood frequency analysis. *Journal of Hydrology*, *100*, 269–290. https://doi.org/10.1016/0022-1694(88)90188-6

Curtis, G.W. (1987). Technique for estimating flood-peak discharges and frequencies on rural streams in Illinois (Water-Resources Investigations Report 87-4207). U.S. Geological Survey.

Eash, D.A., Barnes, K.K., & Veilleux, A.G. (2013). Methods for estimating annual exceedance-probability discharges for streams in Iowa, based on data through water year 2010 (Water-Resources Investigations Report 2013-5086). U.S. Geological Survey.

Egodawatta, P., Haddad, K., Rahman, A., & Goonetilleke, A. (2014). A Bayesian regression approach to assess uncertainty in pollutant wash-off modelling. *Science of the Total Environment*, *479-480*(C), 233–240. https://doi.org/10.1016/j.scitotenv.2014.02.012

Eng, K., Milly, P. C. D., & Tasker, G. D. (2007). Flood regionalization: A hybrid geographic and predictor-variable region-of-influence regression method. *Journal of Hydrologic Engineering*, *12*, 585–591.

Eng, K., Stedinger, J.R., & Gruber, A.M. (2007). Regionalization of streamflow characteristics for the Gulf-Atlantic Rolling Plains using leverage guided region-of-influence regression. Paper presented at the World Environmental and Water Resources Congress, American Society of Civil Engineers, Tampa, Florida.

England, J.F., Jr., Cohn, T.A., Faber, B.A., Stedinger, J.R., Thomas, W.O., Jr., Veilleux, A.G., et al. (2018). *Guidelines for determining flood flow frequency—Bulletin 17C*, U.S. Geological Survey techniques and methods (book 4, ch. B5, pp. 1–148). Reston, VA: U.S. Geological Survey. https://doi.org/10.3133/tm4B5

Feaster, T.D., Gotvald, A.J., Weaver, J.C. (2009). Magnitude and frequency of rural floods in the southeastern United States, 2006, Volume 3—South Carolina (Water-Resources Investigations Report 2009–5156). U.S. Geological Survey.

Feaster, T.D. & Tasker, G.D. (2002). Techniques for estimating the magnitude and frequency of floods in rural basins of South Carolina, 1999 (Water-Resources Investigations Report 02-4140). Columbia, South Carolina: U.S. Geological Survey.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis*. Boca Raton, FL: Chapman & Hall/CRC.

Gotvald, A.J., Barth, N.A., Veilleux, A.G., & Parrett, C. (2012). Methods for determining magnitude and frequency of floods in California, based on data through water year 2006 (Water-Resources Investigations Report 2012–5113). U.S. Geological Survey.

Gotvald, A.J., Feaster, T.D., & Weaver, J.C. (2009). Magnitude and frequency of rural floods in the southeastern United States, 2006—Volume 1, Georgia (Water-Resources Investigations Report 2009–5043). U.S. Geological Survey.

Griffis, V. W., & Stedinger, J. R. (2007a). Log-Pearson type 3 distribution and its application in flood frequency analysis. II: Parameter estimation methods. *Journal of Hydrologic Engineering*, *12*(5), 492–500. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:5(492

Griffis, V. W., & Stedinger, J. R. (2007b). The use of GLS regression in regional hydrologic analyses. *Journal of Hydrology*, *344*(1-2), 82–95. https://doi.org/10.1016/j.jhydrol.2007.06.023

Griffis, V. W., & Stedinger, J. R. (2009). Log-Pearson type 3 distribution and its application in flood frequency analysis. III: Sample skew and weighted skew estimators. *Journal of Hydrologic Engineering*, *14*(2), 121–120. https://doi.org/10.1061/(ASCE)1084-0699(2009)14:2(121)

Griffis, V. W., Stedinger, J. R., & Cohn, T. A. (2004). LP3 quantile estimators with regional skew information and low outlier adjustments. *Water Resources Research*, *40*, W07503. https://doi.org/10.1029/2003WR002697

Groupe de recherche en hydrologie statistique (GREHYS) (1996). Presentation and review of some methods for regional flood frequency analysis. *Journal of Hydrology*, *186*(1–4), 63–84.

Gruber, A. M., Reis, D.D. Jr., & Stedinger, J.R. (2007). Models of regional skew based on Bayesian GLS regression. Paper presented at the World Environmental & Water Resources Conference—Restoring our Natural Habitat, American Society of Civil Engineers, Tampa, Florida.

Haddad, K., Egodawatta, P., Rahman, A., & Goonetilleke, A. (2013). Uncertainty analysis of pollutant build-up modelling based on a Bayesian weighted least squares approach. *Science of the Total Environment*, *449*(C), 410–417. https://doi.org/10.1016/j.scitotenv.2013.01.086

Haddad, K., Johnson, F., Rahman, A., Green, J., & Kuczera, G. (2015). Comparing three methods to form regions for design rainfall statistics: Two case studies in Australia. *Journal of Hydrology*, *527*(C), 62–76. https://doi.org/10.1016/j.jhydrol.2015.04.043

Haddad, K., & Rahman, A. (2012). Regional flood frequency analysis in Eastern Australia: Bayesian GLS regression-based methods within fixed region and ROI framework—Quantile regression vs. parameter regression technique. *Journal of Hydrology*, *430-431*, 142–161. https://doi.org/10.1016/j.jhydrol.2012.02.012

Haddad, K., Rahman, A., & Kuczera, G. (2011). Comparison of ordinary and generalised least squares regression models in regional flood frequency analysis: A case study for New South Wales. *Australian Journal of Water Resources*, *15*(1), 59–70.

Haddad, K., Rahman, A., & Ling, F. (2015). Regional flood frequency analysis method for Tasmania, Australia: A case study on the comparison of fixed region and region-of-influence approaches. *Hydrological Sciences Journal*, *60*(12), 2086–2101. https://doi.org/10.1080/02626667.2014.950583

Haddad, K., Rahman, A., & Stedinger, J. R. (2012). Regional flood frequency analysis using Bayesian generalized least squares: A comparison between quantile and parameter regression techniques. *Hydrological Processes*, *26*(7), 1008–1021. https://doi.org/10.1002/hyp.8189

Han, S., Guikema, S., Quiring, S., Lee, K., Rosowsky, D., & Davidson, R. (2009). Estimating the spatial distribution of power outages during hurricanes in the gulf coast region. *Reliability Engineering & System Safety*, *94*(2), 199–210.

Hardison, C. H. (1975). Generalized skew coefficients of annual floods in the United States and their application. *Water Resources Research*, *11*(6), 851–854. https://doi.org/10.1029/WR010i004p00745

Interagency Advisory Committee on Water Data (IACWD) (1982). Guidelines for determining flood flow frequency (Bull. 17B), Washington, D. C.

Jennings, M. E., Thomas, W. O. Jr., & Riggs, H. C. (1994). *Nationwide summary of U.S. Geological Survey regional regression estimates for estimating magnitude and frequency of floods for ungaged sites* (Water Resources Investigation Report 94-4002). U.S. Geological Survey.

Jeong, Dae Il, Stedinger, JR., Kim, Y-O, & Sung, J.H. (2007). Bayesian GLS for regionalization of flood characteristics in Korea. Paper presented at the World Environmental & Water Resources Conference—Restoring our Natural Habitat, American Society of Civil Engineers, Tampa, Florida.

Kim, S. U., & Lee, K. S. (2010). Regional low flow frequency analysis using Bayesian regression and prediction at ungauged catchment in Korea. *KSCE Journal of Civil Engineering*, *14*(1), 87–98. https://doi.org/10.1007/s12205-010-0087-7

Kitanidis, P. K. (1986). Parameter uncertainty in estimation of spatial functions: Bayesian analysis. *Water Resources Research*, *22*(4), 499–507. https://doi.org/10.1029/WR022i004p00499

Kjeldsen, T. R., & Jones, D. (2007). Estimation of an index flood using data transfer in the UK. *Hydrological Sciences Journal*, *52*(1), 86–98. https://doi.org/10.1623/hysj.52.1.86

Kjeldsen, T. R., & Jones, D. A. (2009a). An exploratory analysis of error components in hydrological regression modeling. *Water Resources Research*, *45*, W02407. https://doi.org/10.1029/2007WR006283

Kjeldsen, T. R., & Jones, D. A. (2009b). A formal statistical model for pooled analysis of extreme floods. *Hydrologic Research*, *40*(5), 465–480. https://doi.org/10.2166/nh.2009.055

Kjeldsen, T. R., & Jones, D. A. (2010). Predicting the index flood in ungauged UK catchments: On the link between data-transfer and spatial model error structure. *Journal of Hydrology*, *387*(1–2), 1–9. https://doi.org/10.1016/j.jhydrol.2010.03.024

Kjeldsen, T. R., Jones, D. A., & Morris, D. G. (2014). Using multiple donor sites for enhanced flood estimation in ungauged catchments. *Water Resources Research*, *50*(8), 6646–6657. https://doi.org/10.1002/2013WR015203

Kjeldsen, T. R., & Rosbjerg, D. (2002). Comparison of regional index flood estimation procedures based on the extreme value type I distribution. *Stochastic Environmental Research and Risk Assessment*, *16*(5), 358–373.

Kroll, C. N., & Stedinger, J. R. (1998). Regional hydrologic analysis: Ordinary and generalized least squares revisited. *Water Resources Research*, *34*(1), 121–128. https://doi.org/10.1029/97WR02685

Kuczera, G. (1982). Combining site-specific and regional information: An empirical Bayes approach. *Water Resources Research*, *18*(2), 306–314. https://doi.org/10.1029/WR018i002p00306

Kuczera, G. (1983). Effect of sampling uncertainty and spatial correlation on an empirical Bayes procedure for combining site and regional information. *Journal of Hydrology*, *65*(4), 373–398. https://doi.org/10.1016/0022-1694(83)90088-4

Laio, F., Ganora, D., Claps, P., & Galeati, G. (2011). Spatially smooth regional estimation of the flood frequency curve (with uncertainty). *Journal of Hydrology*, *408*(1-2), 67–77. https://doi.org/10.1016/j.jhydrol.2011.07.022

Lamontagne, J.R., Stedinger, J.R., Berenbrock, C., Veilleux, A.G., Ferris, J.C., Knifong, D.L. (2012). Development of regional skews for selected flood durations for the Central Valley Region, California, based on data through water year 2008 (Water-Resources Investigations Report 2012–5130). U.S. Geological Survey.

Landers, M.N., & Wilson, K.V. Jr. (1991). Flood characteristics of Mississippi streams. (Water Resources Investigations Report 91-4037). U. S. Geological Survey.

Law, G.S., & Tasker, G.D. (2003). Flood-frequency prediction methods for unregulated streams of Tennessee, 2000. (Water Resources Investigations Report 03-4176). U.S. Geological Survey.

Lima, C. H. R., Lall, U., Troy, T., & Devineni, N. (2016). A hierarchical Bayesian GEV model for improving local and regional flood quantile estimates. *Journal of Hydrology*. https://doi.org/10.1016/j.jhydrol.2016.07.042

Lindley, D. V. (1965). *Introduction to probability and statistics from a Bayesian viewpoint, part2. Inference*. Cambridge: Cambridge University Press.

Linhart, H., & Zucchini, W. (1986). *Model selection*. New York, NY: John Wiley and Sons, Inc.

Liu, H., Davidson, R. A., Rosowsky, D. V., & Stedinger, J. R. (2005). Negative binomial regression of electric power outages in hurricanes. *ASCE Journal of Infrastructure Systems*, *11*(4), 258–267.

Ludwing, A.H. & Tasker, G.D. (1993). Regionalization of low-flow characteristics of Arkansas streams. (Water-Resources Investigations Report 93-4013). U.S. Geological Survey.

Madsen, H., Mikkelsen, P. S., Rosbjerg, D., & Harremoes, P. (2002). Regional estimation of rainfall intensity-duration-frequency curves using generalized least squares regression of partial duration series statistics. *Water Resources Research*, *38*(11), 1239. https://doi.org/10.1029/2001WR001125

Madsen, H., & Rosbjerg, D. (1997). Generalized least squares and empirical Bayes estimation in regional partial duration series index-flood modeling. *Water Resources Research*, *33*(4), 771–782. https://doi.org/10.1029/96WR03850

Martins, E. S., & Stedinger, J. R. (2000). Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research*, *36*(3), 737–744. https://doi.org/10.1029/1999WR900330

Martins, E. S., & Stedinger, J. R. (2002). Cross correlations among estimators of shape. *Water Resources Research*, *38*(11), 1252. https://doi.org/10.1029/2002WR001589

Mastin, M.C., Konrad, C.P., Veilleux, A.G., & Tecca, A.E. (2016). Magnitude, frequency, and trends of floods at gaged and ungaged sites in Washington, based on data through water year 2014. (*Scientific Investigations Report 2016–5118*). Reston, VA: U.S. Geological Survey. https://doi.org/10.3133/sir20165118

McCuen, R. H. (1979). Map skew. *Journal of the Water Resources Planning and Management Division*, *105*(2), 269–277. [with Closure 107(WR2), 582, 1981]

McCuen, R. H. (2001). Generalized flood skew: Map versus watershed skew. *Journal of Hydrologic Engineering, ASCE*, *6*(4), 293–299.

McCuen, R. H., & Smith, E. (2008). Origin of flood skew. *Journal of Hydrologic Engineering*, *12*(9). https://doi.org/10.1061/(ASCE)1084-0699(2008)13:9(771)

Medina, K.D. (1987). Analysis of surface-water data network in Kansas for effectiveness in providing regional streamflow information (with a section on "Theory and application of generalized least squares" by G.D. Tasker). (USGS Water-Supply Paper 2303). Reston, VA: US Geological Survey.

Merz, R., & Bloschl, G. (2005). Flood frequency regionalisation—Spatial proximity vs. catchment attributes. *Journal of Hydrology*, *302*(1-4), 283–306. https://doi.org/10.1016/j.jhydrol.2004.07.018

Merz, R., & Bloschl, G. (2008). Flood frequency hydrology: 2. Combining data evidence. *Water Resources Research*, *44*, W08433. https://doi.org/10.1029/2007WR006745

Micevski, T., Hackelbusch, A., Haddad, K., Kuczera, G., & Rahman, A. (2015). Regionalisation of the parameters of the log-Pearson 3 distribution: A case study for New South Wales, Australia. *Hydrological Processes*, *29*(2), 250–260. https://doi.org/10.1002/hyp.10147

Micevski, T., & Kuczera, G. (2009). Combining site and regional flood information using a Bayesian Monte Carlo approach. *Water Resources Research*, *45*, W04405. https://doi.org/10.1029/2008WR007173

Moss, M. E., & Tasker, G. D. (1991). An intercomparison of hydrological network-design technologies. *Hydrological Sciences Journal*, *36*(3), 209.

Olson, S.A. (2014). Estimation of flood discharges at selected annual exceedance probabilities for unregulated, rural streams in Vermont (with a section on Vermont regional skew regression, by Veilleux, A.G.). (Scientific Investigations Report 2014–5078). U.S. Geological Survey.

Over, T. M., Saito, R. J., Veilleux, A. G., Sharpe, J. B., Soong, D. T., & Ishii, A. L. (2016). Estimation of peak discharge quantiles for selected annual exceedance probabilities in northeastern Illinois (Scientific Investigations Report 2016-5050). U.S. Geological Survey. https://doi.org/10.3133/sir20165050

Paretti, N. V., Kennedy, J. R., Turney, L. A., & Veilleux, A. G. (2014). Methods for estimating magnitude and frequency of floods in Arizona, developed with unregulated and rural peak-flow data through water year 2010 (Scientific Investigations Report 2014-5211). U.S. Geological Survey. https://doi.org/10.3133/sir20145211

Parrett, C., Veilleux, A., Stedinger, J.R., Barth, N.A., Knifong, D.L., & Ferris, J.C. (2011). Regional skew for California, and flood frequency for selected sites in the Sacramento–San Joaquin River Basin, based on data through water year 2006 (Scientific Investigations Report 2010–5260). U.S. Geological Survey.

Pope, F., Tasker, G.D., & Robbins, J.C. (2001). Estimating the magnitude and frequency of floods in rural basins of North Carolina—Revised. (Water Resources Investigations Report 01-4207). Raleigh, North Carolina: U.S. Geological Survey.

Qian, S. S., Reckhow, K. H., Zhai, J., & McMahon, G. (2005). Nonlinear regression modeling of nutrient loads in streams: A Bayesian approach. *Water Resources Research*, *41*, W07012. https://doi.org/10.1029/2005WR003986

Rahman, A., Haddad, K., Kuczera, G., & Weinmann, E. (2016). Regional flood methods. In *Commmonwealth of Australia, Australian rainfall and runoff: A guide to flood estimation* (book 3, ch. 3, pp. 105–146). Canberra: Geoscience Australia. Retrived from http://book.arr.org.au.s3-website-ap-southeast-2.amazonaws.com

Rasmussen, P.P., & Perry, C.A. (2000). Estimation of peak streamflows for unregulated rural streams in Kansas (Water Resources Investigations Report 00-4079). Lawrence, Kansas: U.S. Geological Survey.

Reis, D. S., Jr. (2005). Flood frequency analysis employing Bayesian regional regression and imperfect historical information, (Doctoral dissertation). Location: Cornell University.

Reis, D. S., Jr., Stedinger, J.R., & Martins, E.S. (2003). Bayesian GLS regression with application to LP3 regional skew estimation. Paper presented at the World Water & Environmental Resources Congress, American Society of Civil Engineers, Philadelphia, PA.

Reis, D. S. Jr., Stedinger, J. R., & Martins, E. S. (2005). Bayesian generalized least squares regression with application to log Pearson type 3 regional skew estimation. *Water Resources Research*, *41*, W10419. https://doi.org/10.1029/2004WR003445

Renard, B. (2011). A Bayesian hierarchical approach to regional frequency analysis. *Water Resources Research*, *47*, W11513. https://doi.org/10.1029/2010WR010089

Rencher, A. C. (2000). Linear models in statistics. In *Wiley series in probability and statistics* (pp. 1–578). New York, NY: John Wiley & Sons, Inc.

Riggs, H. C. (1973). Regional analyses of streamflow characteristics. (Techniques of water-resources investigations: Book 4, chapter B3). United States Geological Survey.

Robins, J. M., van der Vaart, A., & Ventura, V. (2000). The asymptotic distribution of p-values in composite null models. *Journal of the American Statistical Association*, *95*, 1143–1156.

Robson, A., & Reed, D. (1999). Statistical procedure for flood frequency estimation. In *Flood estimation handbook* (Vol. 3, pp. 1–338). Oxfordshire, UK: Institute of Hydrology.

Rosbjerg, D., & Madsen, H. (1995). Uncertainty measures of regional flood frequency estimators. *Journal of Hydrology*, *167*, 209–224.

Seidou, O., Ouarda, T. B. M. J., Barbet, M., Bruneau, P., & Bobeé, B. (2006). A parametric Bayesian combination of local and regional information in flood frequency analysis. *Water Resources Research*, *42*, W11408. https://doi.org/10.1029/2005WR004397

Skøien, J. O., & Bloschl, G. (2007). Spatiotemporal topological kriging of runoff time series. *Water Resources Research*, *43*, W09419. https://doi.org/10.1029/2006WR005760

Skøien, J. O., Merz, R., & Bloschl, G. (2006). Top-kriging—Geostatistics on stream networks. *Hydrology and Earth System Sciences*, *2*(6), 2253–2286. https://doi.org/10.5194/hess-10-277-2006

Soenksen, P. J., Miller, L. D., Sharpe, J. B., & Watton, J. R. (1999). Peak-flow frequency relations and evaluation of the peak-flow gaging network in Nebraska (Water-Resources Investigation Report 99-4032). Lincoln, Nebraska: U.S. Geological Survey.

Southard, R.E., & Veilleux, A.G. (2014). Methods for estimating annual exceedance-probability discharges and largest recorded floods for unregulated streams in rural Missouri: (Scientific Investigations Report 2014-5165). U.S. Geological Survey. https://doi.org/10.3133/sir20145165

Stedinger, J. R., & Tasker, G. (1986a). Correction to "Regional hydrologic analysis, 1, Ordinary, weighted and generalized least squares compared". *Water Resources Research*, *22*(5), 844. https://doi.org/10.1029/WR022i005p00844

Stedinger, J. R., & Tasker, G. (1986b). Regional hydrologic analysis, 2: Model-error estimators, estimation of sigma and log-Pearson type 3 distributions. *Water Resources Research*, *22*(10), 1487–1499. https://doi.org/10.1029/WR022i010p01487

Stedinger, J. R., & Tasker, G. D. (1985). Regional hydrologic analysis, 1. Ordinary, weighted and generalized least squares compared. *Water Resources Research*, *21*(9), 1421–1432. https://doi.org/10.1029/WR021i009p01421

Stedinger, J.R., Vogel, R.M., & Foufoula-Georgiou, E. (1993). Frequency analysis of extreme events. In D.R. Maidment (Ed.), *Handbook of hydrology* (chap. 18, pp. 18.2-18.66). New York, NY: McGraw-Hill.

Stedinger, J. R., Vogel, R. M., Lee, S. U., & Batchelder, R. (2008). Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. *Water Resources Research*, *44*, W00B06. https://doi.org/10.1029/2008WR006822

Tasker, G. D. (1978). Flood frequency analysis with a generalized skew coefficient. *Water Resources Research*, *14*(2), 373–376. https://doi.org/10.1029/WR014i002p00373

Tasker, G. D. (1980). Hydrologic regression with weighted least squares. *Water Resources Research*, *16*(6), 11107–11113. https://doi.org/10.1029/WR016i006p01107

Tasker, G.D. (1986). Generating efficient gaging plans for regional information. Paper presented at the Integrated Design of Hydrologic Networks Budapest Symposium, International Association of Hydrological Sciences, Budapeste, Hungary.

Tasker, G. D., & Driver, N. E. (1988). Nationwide regression model for predicting urban runoff water quality at unmonitored sites. *Water Resources Bulletin*, *24*(5), 1091–1101.

Tasker, G.D., Eychaner, J.H., & Stedinger, J.R. (1986). Application of generalized least squares in regional hydrologic regression analysis. (Water Supply Paper 2310 – Selected Papers in the Hydrologic Sciences). Reston, VA: U.S. Geological Survey.

Tasker, G. D., Hodge, S. A., & Barks, C. S. (1996). Region of influence regression for estimating the 50-year flood at ungauges sites. *Water Resources Bulletin*, *32*(1), 163–170. https://doi.org/10.1111/j.1752-1688.1996.tb03444.x

Tasker, G. D., & Stedinger, J. R. (1986). Estimating generalized skew with weighted least squares regression. *Journal of Water Resources Planning and Management*, *112*(2), 225–237.

Tasker, G. D., & Stedinger, J. R. (1989). An operational GLS model for hydrologic regression. *Journal of Hydrology*, *111*(1-4), 361–375.

Thomas, D. M., & Benson, M. A. (1970). Generalization of streamflow characteristics from drainage-basin characteristics. *U.S. Geological Survey Water-Supply Paper 1975*.

Veilleux, A.G. (2011). Bayesian GLS Regression, Leverage and influence for regionalization of hydrologic statistics (Doctoral Dissertation). Ithaca, NY: Cornell University.

Vicens, G. J., Rodríguez-Iturbe, I., & Schaake, J. C. Jr. (1975). A Bayesian framework for the use of regional information in hydrology. *Water Resources Research*, *11*(3), 405–414.

Viglione, A., Merz, R., Salinas, J. L., & Blöschl, G. (2013). Flood hydrology: 3. A Bayesiain analysis. *Water Resources Research*, *49*, 1–18. https://doi.org/10.1029/2011WR010782

Wagner, D. M., Krieger, J. D., & Veilleux, A. G. (2016). Methods for estimating annual exceedance probability discharges for streams in Arkansas, based on data through water year 2013 (Scientific Investigations Report 2016–5081). U.S. Geological Survey. https://doi.org/10.3133/sir20165081

Weaver, J.C., Feaster, T.D., & Gotvald, A.J. (2009). Magnitude and frequency of rural floods in the southeastern United States, through 2006—Volume 2, North Carolina. (Scientific Investigations Report 2009–5158). U.S. Geological Survey.

Wood, M. S., Fosness, R. L., Skinner, K. D., & Veilleux, A. G. (2016). Estimating peak-flow frequency statistics for selected gaged and ungaged sites in naturally flowing streams and rivers in Idaho (Scientific Investigations Report 2016–5083). U.S. Geological Survey. https://doi.org/10.3133/sir20165083

Yager, R. M. (1998). Detecting influential observations in nonlinear regression modeling of groundwater flow. *Water Resources Research*, *34*(7), 1623–1633. https://doi.org/10.1029/98WR01010

Zarriello, P. J. (2017). Magnitude of flood flows at selected annual exceedance probabilities for streams in Massachusetts (Scientific Investigations Report 2016–5156). U.S. Geological Survey. https://doi.org/10.3133/sir20165156

Zellner, A. (1971). *An introduction to Bayesian inference in econometrics*. New York: John Wiley and Sons, Inc.