Received: 8 June 2017 | Accepted: 9 December 2017
DOI: 10.1002/hyp.11426

RESEARCH ARTICLE

WILEY

A comparison of methods for low streamflow estimation from spot measurements

Timothy J. Stagnitta[1] | Charles N. Kroll[1] | Zhenxing Zhang[2]

[1] Environmental Resources Engineering, SUNY ESF, Syracuse 13210NY, USA
[2] Illinois State Water Survey, University of Illinois at Urbana-Champaign, Champaign 61820IL, USA

Correspondence
Charles N. Kroll, Environmental Resources Engineering, 1 Forestry Drive, Baker 402, SUNY ESF, Syracuse 13210, NY, USA.
Email: cnkroll@esf.edu

Funding information
Department of Environmental Resources Engineering at SUNY ESF

Abstract

Low streamflow statistic estimators at ungauged river sites generally have large errors and uncertainties. This can be due to many reasons, including lack of data, complex hydrologic processes, and the inadequate or improper characterization of watershed hydrogeology. One potential solution is to take a small number of streamflow measurements at an ungauged site to either estimate hydrogeologic indices or transfer information from a nearby site using concurrent streamflow measurements. An analysis of four low streamflow estimation techniques, regional regression, regional plus hydrogeologic indices, baseflow correlation, and scaling, was performed within the Apalachicola–Chattahoochee–Flint watershed, a U.S. Geological Survey WaterSMART region in the south-eastern United States. The latter three methods employ a nominal number of spot measurements at the ungauged site to improve low streamflow estimation. Results indicate that baseflow correlation and scaling methods, which transfer information from a donor site, can produce improved low streamflow estimators when spot measurements are available. Estimation of hydrogeologic indices from spot measurements improves regional regression models, with the baseflow recession constant having more explanatory power than the aquifer time constant, but these models are generally outperformed by baseflow correlation and scaling.

KEYWORDS

baseflow correlation, hydrogeologic indices, low streamflow estimation, scaling, spot measurements, ungauged river sites

1 | INTRODUCTION

Low streamflow estimates are required for a variety of water quality and water quantity management purposes. Low streamflow estimates are used to assist water managers in planning for low flow conditions for river ecology and flow management, wastewater treatment plants, water withdrawal limitations and discharge permitting, and hydropower operations (Smakhtin, 2001). Often, these estimates are represented by low streamflow statistics such as 7Q10, the annual average 7-day minimum flow that is not exceeded on average once every 10 years (Riggs, 1980), or a quantile from an annual or monthly flow duration curve, such as the Q95, the daily average streamflow that is exceeded 95% of the time (Susquehanna River Basin Commission, 2012). The U.S. Geological Survey (USGS) has over 20,000 gauged locations across the United States, and one can relatively easily generate low streamflow statistics from the data provided at these sites. The question addressed here is how best to estimate low streamflow statistics at streamflow sites where only a nominal number of streamflow measurements are available.

A number of techniques are used in practice to estimate flow statistics at ungauged river sites. The simplest techniques are scaling methods, such as the drainage area ratio method, which transfers flow statistics from a donor site using drainage area as a scaling factor (Hirsch, 1979). Hirsch (1979) explored this method in two small basins in Virginia and found that this method works relatively well if the donor site has similar hydrologic characteristics, similar low streamflow drivers and response, and if the low streamflow statistics are strongly related to drainage area. Because streamflow characteristics are generally unknown at the ungauged river site and low streamflow drivers and response are difficult to determine at gauged or ungauged sites, there can be high degree of uncertainty in drainage area ratio methods (Hirsch, 1979). A related method when no streamflow data are available at a site of interest is regional regression (Kroll, Luz, Allen, & Vogel, 2004; Thomas & Benson, 1970; Vogel & Kroll, 1992). Regional regression requires a database of geomorphic, geologic, climatic, and topographic basin characteristics to develop a model (Thomas & Benson, 1970). Regional regression models often perform poorly estimating low flow quantiles (Kroll et al., 2004). Vogel and Kroll (1992), Kroll et al. (2004), and Eng and Milly (2007) found that there is often a lack of hydrogeologic information within the watershed characteristic database that could potentially improve low streamflow regional regression models.

Another potential method to estimate low streamflow statistics at an ungauged river site is to develop a rainfall–runoff model (Broda, Larocque, & Paniconi, 2014; Matonse & Kroll, 2013). Rainfall–runoff models typically focus on capturing average flow or flood events and often provide a simplistic representation of groundwater discharge processes. Such groundwater discharges (baseflow) typically have a large impact on low streamflow conditions. However, baseflow processes are complex, and rainfall–runoff models often do not accurately capture the physical processes and heterogeneous subsurface characteristics that are important to low streamflow generation. In addition, rainfall–runoff models often require a large number of streamflow measurements to calibrate and verify the model.

Another method to estimate low streamflow statistics utilizes flow duration curves (FDCs) to transfer information from a gauged river site to an ungauged river site (Archfield et al., 2009; Archfield & Vogel, 2010; Vogel & Fennessey, 1994). Using an FDC at an ungauged site, a relationship between the exceedance probability and daily streamflows at a donor site is formed. This technique is called the quantile-probability-probability-quantile technique (Fennessey, 1994) and is used to recreate a sequence of daily streamflows at the ungauged site, from which low streamflow statistics can then be derived. Although the quantile-probability-probability-quantile method often works well to create a streamflow record at the ungauged site, it also requires a technique to create the FDC at the ungauged site (typically regional regression) and thus suffers from some of the same problems as regional regression models used to directly estimate low streamflow statistics. Previous studies have employed these techniques with the assumption that there are no data available at the ungauged site (Archfield et al., 2009; Vogel & Fennessey, 1994).

Laaha and Blöschl (2005) examined a number of methods to improve low flow estimation with short records (SR) ranging from a few measurements to multiple years of daily streamflow data. Laaha and Blöschl used a ratio method between low flow measurements at an ungauged site and a donor site to determine Q95. They found that their spot gauging technique was a slight improvement over a simple regional regression technique but was highly dependent on how the donor site was chosen. Spatial variability of low flow characteristics and timing of spot measurements caused large uncertainties with the method, and they suggested that spot measurements are often not representative of the Q95 low flow. They suggested that the method could improve if they expanded the number of spot measurements and analysed more than one low flow period.

Stedinger and Thomas (1985) developed a baseflow correlation technique with use of a donor site to predict low flow statistics at partially gauged sites with a nominal number of streamflow measurements. This technique estimates the log-space mean and variance of d-day annual minimum flows using regression between concurrent baseflow measurements and performed well in the small Virginia study area where it was first applied. Potter (2001) followed similar assumptions as Stedinger and Thomas (1985), with an additional assumption that the log-variances at the gauged and ungauged site are the same, to estimate the central moments of the daily baseflow record using four or less discharge measurements. Potter applied the method to two watershed pairs in Wisconsin and found that the use of a donor site performed well to estimate the long-term baseflow mean, median, and lower decile. Reilly and Kroll (2003) expanded on Stedinger and Thomas' (1985) method at 1,300 river sites throughout the United States. They found that the baseflow correlation technique generally is a good method for estimating low streamflow statistics across the United States and is an improvement on methods such as regional regression. Zhang and Kroll (2007a) examined the impact of assumptions employed in the baseflow correlation method, whereas Zhang and Kroll (2007b) expanded this technique by using multiple gauged river sites. These studies indicated that baseflow correlation should be further examined for estimating low streamflow statistics in other study areas.

Eng, Kiang, Chen, Carlisle, and Granato (2011) examined the bias for estimating the 7Q10 using three index-streamflow approaches (i.e., maintenance of variance, baseflow correlation, and a scaling method) and compared these to regional regression augmented by hydrogeologic indices. They explored the impact of the range of streamflow used with the index approaches, the areal density of gauges, and two donor site selection methods. Eng et al. (2011) found that baseflow correlation and a maintenance of variance method produced 7Q10 estimators with a lower bias than regional regression and that only a small portion of this bias is explained by the areal density of stream gauges and hydrologic similarity. Eng et al. (2011) used very large hydrologic regions, which may have adversely impacted the performance of regional regression. In addition, they only examined the use of 10 spot measurements at the partial record gauge.

In this experiment, regional regression, regional regression with added hydrogeological indices, baseflow correlation, and scaling methods are compared when a nominal number of baseflow measurements are available at the ungauged river site. This experiment expands on the research of Eng et al. (2011) by studying a smaller region with a larger density of gauges, examining estimation of a variety of low streamflow statistics, varying the number of spot measurements taken at the partial record site, using two different hydrologic indices with regional regression, and comparing a number of different donor site selection methods. The study area includes unregulated gauged streamflow sites within the Apalachicola–Chattahoochee–Flint (ACF) watershed in the south-eastern United States. An analysis estimating hydrogeological indices from spot measurements is first performed. This is followed by a comparison of different methods for estimating low streamflow statistics when spot measurements are available. An analysis of how best to choose spot measurements and donor sites is also included, with a focus on developing a methodology that practitioners could use to measure streamflows and estimate low streamflow statistics.

2 | METHODS

Four techniques to estimate low flow quantiles at ungauged sites with spot measurements are explored: regional regression, regional regression with estimated hydrogeologic indices, baseflow correlation, and scaling methods. In the following sections, each of these methods is briefly described.

2.1 | Regional regression

The first technique is least squares regression (Thomas & Benson, 1970). Regional regression uses the relationship between flow statistics and geomorphic, geologic, climatic, and topographic parameters to estimate flow statistics at ungauged sites (Thomas & Benson, 1970; Vogel & Kroll, 1992). These models typically have the form

$$Q = \beta_o X_1^{\beta_1} X_2^{\beta_2}, \tag{1}$$

where Q is the flow statistic of interest, X_i are basin characteristics, and β_i are parameters obtained from multivariate regression procedures. Vogel and Kroll (1992) showed that for low streamflow estimation, the form of this model is consistent with a Boussinesq-derived groundwater discharge model based on a linear reservoir hypothesis. The logarithm of both sides of Equation 1 is taken, resulting in a log-linear regression model. Here, ordinary least squares (OLS) regression procedures are used to develop low streamflow regional regression models. Although weighted least squares or generalized least squares (GLS) regression procedures could be employed to construct regional regression equations (Tasker, 1980; Tasker & Stedinger, 1989), when the model error variance is large, which is typical of low streamflow models, it overwhelms the time sampling error, and OLS, weighted least squares, and GLS regression procedures produce similar results (Kroll & Stedinger, 1998). Because a concurrent record is used at all sites, the record length and thus the time sampling error of the at-site estimators will be less variable across sites. Under these conditions, OLS should perform similarly to GLS.

2.2 | Regression with estimated hydrogeologic indices

The second technique again utilizes OLS regression to estimate low flow statistics at the site of interest, except that baseflow indices from small samples are included as potential explanatory variables in the model; these indices are often not included in the database of watershed characteristics. There are many potential hydrogeologic indices that can be derived from streamflow series, such as the baseflow recession constant (K_b; Vogel & Kroll, 1996), the aquifer time constant (τ; Eng & Milly, 2007), and the baseflow index (Insititute of Hydrology, 1980). Here, K_b and τ are considered, as both have been shown to improve low streamflow regional regression models (Eng & Milly, 2007; Vogel & Kroll, 1996). For a review of baseflow recession analysis, see Tallaksen (1995).

Equation 2 defines K_b and τ as an estimator of the daily percentage decline in streamflow during times with no surface or shallow subsurface run-off (Eng & Milly, 2007; Kroll et al., 2004):

$$Q_{t+\Delta t} = Q_t {K_b}^{\Delta t} = Q_t e^{-\Delta t/\tau}, \tag{2}$$

where Q_t is the daily streamflow on day t and $Q_{t+\Delta t}$ is the daily streamflow Δt days after t. K_b and τ have been shown to be related to basin hydraulic conductivity and drainable soil porosity (Brutsaert & Lopez, 1998; Eng & Milly, 2007; Vogel & Kroll, 1996). Vogel and Kroll (1996) recommended K_b be estimated as

$$K_b = \exp\left\{-\exp\left[\frac{1}{m}\sum_{t=1}^{m}\left\{\ln\left(\frac{Q_t - Q_{t+\Delta t}}{\Delta t}\right) - \ln\left(\frac{Q_t + Q_{t+\Delta t}}{2}\right)\right\}\right]\right\}, \tag{3}$$

where Q_t is the first chosen streamflow in the baseflow recession, Δt is the number of days from the first to the second chosen streamflow in the baseflow recession, and m is the total number of recession pairs. This equation is based on assuming that groundwater discharge to a stream is linearly related to the storage within the aquifer (Brutsaert & Nieber, 1977; Vogel & Kroll, 1992; Vogel & Kroll, 1996). Eng and Milly (2007) recommended τ be calculated as

$$\tau = \frac{\Delta t}{\ln(Q_t) - \ln(Q_{t+\Delta t})}, \tag{4}$$

where τ is estimated for each recession and then averaged across all recessions at a site. Estimators of K_b and τ require two measurements from each baseflow recession. Section 4.3 explores estimation of K_b and τ using a small number of spot measurements, and Section 4.4 examines whether these small-sample estimators can improve low streamflow regional regression models.

2.3 | Baseflow correlation

Baseflow correlation also uses spot measurement at the ungauged site as well as concurrent baseflow measurements at a donor site (Stedinger & Thomas, 1985). This method assumes that there is a linear relationship between the logarithm of the annual minimum d-day flows at the donor and ungauged sites, the relationship between annual d-day minimum flows is similar to the relationship between instantaneous baseflows measurements, annual minimum flows are well described by the log Pearson type 3 (LP3) distribution (Barnes, 1986; Rumenik and Grubbs, 1996; Wandle and Randall, 1994), and the log-skew of the d-day flows (and thus the frequency factors) at the donor and ungauged sites are the same. Zhang and Kroll (2007a) examined these assumptions in regions across the United States and generally found them to be valid. Using these assumptions, the log-space mean and variance of the d-day low flows are estimated, and then, the quantile of interest from the LP3 distribution is estimated (Reilly & Kroll, 2003; Stedinger & Thomas, 1985). Stedinger and Thomas (1985) also derived the variance of the logarithm of the quantile estimator. This method was used to estimate 7Q10, 7Q2, 30Q10, and 30Q2; because Q95 and Q99 are not based on a distributional assumption, this method could not be used to estimate these statistics.

2.4 | Scaling methods

Laaha and Blöschl (2005) proposed estimating Q95 at SR sites using a ratio method:

$$\frac{\sum Q_{i,\mathrm{UG}}}{\sum Q_{i,\mathrm{DS}}} = \frac{Q95_{\mathrm{UG}}}{Q95_{\mathrm{DS}}}, \tag{5}$$

where $Q_{i,\mathrm{UG}}$ and $Q_{i,\mathrm{DS}}$ are concurrent baseflows at the ungauged and donor sites, and Q95 at the donor site is calculated from the entire record. This method was used to estimate Q95, Q99, 7Q10, 7Q2, 30Q10, and 30Q2 using both real- and log-space flows. In addition to this method, the performance of both real- and log-space drainage area scaling (Hirsch, 1979) was explored. Because neither of these methods was better than regional regression without hydrogeologic indices, these methods were not included in our results.

3 | EXPERIMENTAL DESIGN

3.1 | Study region and streamflow data

This analysis uses streamflow data from a U.S. Department of Interior WaterSMART (Sustain and Manage America's Resources for Tomorrow) study area (USGS, 2014). The primary WaterSMART study areas are the ACF, Colorado, and Delaware River Basins (USGS, 2014). This study focuses on the ACF Basin, which includes 182 gauged streamflow sites in the south-eastern United States that have limited amounts of regulation.

Eight sites from the original 182 gauged sites have flow quantiles estimated as zero. Although intermittent streamflow sites have important information, streamflow and flow quantiles estimated as zero complicate analyses requiring logarithmic transformations, such as log-linear regional regression models employed here (Kroll & Stedinger, 1999). Sites with zero quantiles are removed from this analysis, thus leaving 174 sites. In addition, two streamflow sites with exceptionally large drainage areas have been removed because of their lack of hydrologic homogeneity with other sites, leaving a total of 172 sites. To limit the impact of climate from non-concurrent periods of record, a common period of record from 1980 to 2010 was employed at all sites. At least 10 years of continuous record over this period at each site was required, leaving 152 sites. Farmer et al. (2014) use the same study region using the same concurrent record, though they filled in missing streamflows at sites; this analysis does not fill in missing streamflows and uses only recorded streamflow measurements. In addition, three sites were removed because of a limited number of baseflow days within the concurrent record, leaving a total of 149 sites for our analysis (Figure 1).

At each streamflow site, low streamflow statistics are estimated using the low flow water years from April 1981 to March 2010. To estimate the 7Q10, 7Q2, 30Q10, and 30Q2, annual 7-day or 30-day low streamflow series are determined in each water year, and then, an LP3 distribution is fit using method of moment parameter estimators (Stedinger, Vogel, & Foufoula-Georgiou, 1993). The LP3 distribution is generally employed to describe annual low streamflow

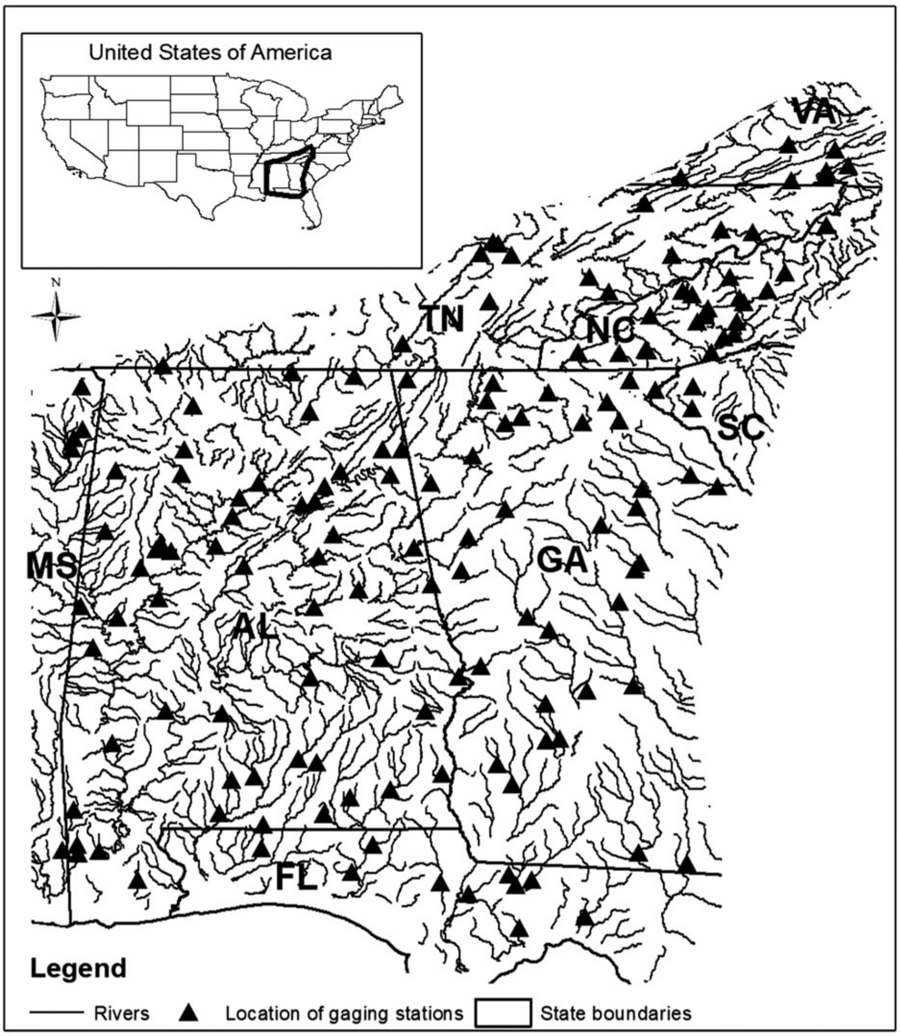

FIGURE 1 Map of gauge locations in the Apalachicola–Chattahoochee–Flint study region

series in the United States (Risley, Stonewall, & Haluska, 2008; Rossman, 1990). From the fitted LP3 distributions, either the 10th or 50th percentile is used to estimate 7Q10 or 7Q2 (and 30Q10 or 30Q2). To estimate Q95 and Q99, period-of-record FDCs are developed for each site, and the quantiles with either the 95th or 99th exceedance probability is used to estimate Q95 or Q99 by linearly interpolating between Weibull plotting positions (method 2 of Vogel & Fennessey, 1994).

These "at-site" low streamflow statistic estimates from historic records are used as the observations of these low flow statistics; all methods are assessed by comparing predictions from the other methods to these estimates. Forty-nine watershed characteristics that were selected by Farmer et al. (2014) as the most promising explanatory variables in this region were used in the low flow regional regression analyses. These explanatory variables were provided by the USGS database developed by Falcone, Carlisle, Wolock, and Meador (2010) and have been employed in other studies in this region (see Pugliese, Farmer, Castellarin, Archfield, & Vogel, 2016 and Croteau, Kroll, Over, & Archfield, 2016 for a description of variables).

3.2 | Spot measurements

This experiment requires spot measurements, instantaneous streamflow measurements taken at a specific time, from an ungauged site to calculate low flow statistics. Because such spot measurements are not readily available, in this experiment, average daily streamflows are used as a replacement for spot measurements. At a USGS gauge, daily streamflow data are an average of instantaneous streamflow data measured at intervals of 5 to 60 min (USGS, 2012) and then averaged across a day to estimate the daily average streamflow. In practice, a water manager would obtain spot measurements from an ungauged site at a specific time, and the instantaneous streamflow at the donor site at the same time. For this experiment, the assumption is that the daily average streamflow is similar to instantaneous spot measurements on the same day, which generally should be a valid assumption during baseflow conditions.

3.3 | Baseflow conditions

To estimate low flow statistics from spot measurements, one needs to designate baseflow conditions to separate days where streamflow is primarily from baseflow (baseflow days) and from non-baseflow days. Baseflow is defined as the portion of streamflow from groundwater (Arnold & Allen, 1999). Zhang and Kroll (2007a) found that for the baseflow correlation method, baseflow measurements should be taken in the late summer and early fall months (July to October), as far from run-off events as possible, and should be nearly independent from other baseflow measurements. Across all study sites, 92% of all 7-day annual minimum flows occur from July to October, so this period was also used here. Following Vogel and Kroll (1992, 1996) and Kroll and Stedinger (1999), for this experiment, baseflow conditions occurs after a 3-day drop from peak streamflow, and the baseflow recession continues until the streamflow increases (Figure 2). Note that other non-climatic methods have been employed for defining baseflow conditions, including a 5-day drop (Aksoy & Wittenberg, 2011; Wittenberg, 2003) and a drop calculated for each site as a function of watershed area (Bras, 1990; Reilly & Kroll, 2003). Although a 3-day drop is a relatively short duration, its use substantially increases the number of pairs of current baseflow conditions at the donor and ungauged sites. An analysis of precipitation events during baseflow conditions is explored in Section 4.1.

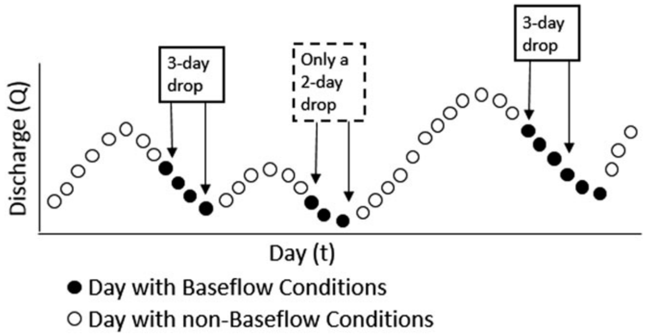

FIGURE 2 Illustration of baseflow definition in daily streamflow record with recession pairs

In addition, baseflow filters are used to reduce the effect of precipitation and large flows that may remain after a 3-day drop of streamflow. To reduce the effect of precipitation, the last day of each recession could be removed (L), given there is a chance that precipitation may occur on this day because the following day there is an increase in streamflow. To reduce the effect of recessions that occur with large streamflows, the FDC at the donor site is utilized. If the streamflow at the donor site the day before a recession starts is above a specific FDC threshold (25% exceedance probability = E_0.25 or 50% exceedance probability = E_0.5), then the entire recession is removed. This analysis explores these filtering methods as well as a combination of both methods (L_0.25 and L_0.5). The impact of these filtering techniques is discussed in Section 4.1.

3.4 | Cross validation of regional regression without hydrogeology

Here, a repeated sequential delete-1/3 cross validation procedure is used to assess performance of low flow regional regression estimators. In this technique, the data set is randomly divided into thirds. One of the thirds is removed, the other two-thirds is used to calibrate the model, and then, the fitted model is used to estimate flow statistics at the removed sites. The other thirds are then sequentially removed, and the process is repeated. Five hundred iterations of this random selection are performed to reduce the impact of randomly selecting sites; performance statistics stabilized after 100 iterations.

3.5 | Cross validation of regional regression with hydrogeologic indices

A similar strategy as presented by Vogel and Kroll (1996) and Eng and Milly (2007) to select baseflow pairs to estimate hydrogeologic indices was used in this experiment. This method chooses a random starting year and a random starting recession within that year. A baseflow recession requires at least a 6-day period where the streamflow does not increase. The 3rd day of the drop is considered the first day of the baseflow recession (Q_t), and the 6th day of the streamflow drop is considered as the second baseflow measurement ($Q_{t+\Delta t}$), where $\Delta t = 3$ (Figure 2). The number of total measured flows (two per recession) examined in this experiment was 4, 6, 8, 10, and 12 days. Eng et al. (2011) used longer recessions (at least an 8-day drop) and more baseflow measurements (20 measurements to estimate 10 values of τ) for regional regression with τ. A less stringent baseflow criterion and fewer measurements to reduce the total duration of sampling needed to obtain low streamflow estimates were used. It was surmised that sampling protocol could be performed over one or two low flow seasons. Estimators of K_b and τ from the entire record are used to develop regression models, and K_b and τ estimators from the measured flows are used in the leave one third out cross validation procedure (Section 3.3). SR estimators of K_b and τ are also compared to those calculated from the entire record to assess the performance of these estimators.

3.6 | Donor site selection for baseflow correlation and scaling methods

Baseflow correlation and scaling methods rely on a gauged donor site; selection of the donor site can have a large influence on the performance of low flow estimators (Eng et al., 2011; Laaha & Blöschl, 2005; Zhang & Kroll, 2007a). There are a wide range of techniques that can be used to select donor sites such as the nearest neighbour (NN), drainage area within a certain range, annual precipitation within a certain range, map correlation method, and various other basin characteristics and spatial methods (Archfield & Vogel, 2010; Clark & Evans, 1954; Laaha & Blöschl, 2005; Ries & Friesz, 2000; Zhang & Kroll, 2007a). For this analysis, the following donor site selection methods are explored: NN, most similar drainage area across the entire study area (area), site producing the minimum variance low flow estimator within 100 km from the ungauged site (100), minimum variance site within 200 km (200), gauges within 100 km, drainage area within ±50% and minimum estimated variance (100area), and gauges within 200 km, drainage area within ±50%, and minimum estimated variance (200area). The variance of 7Q10, 7Q2, 30Q10, and 30Q2 baseflow correlation estimators is estimated using methods presented in Stedinger and Thomas (1985). For Q95 and Q99, the minimum estimated variance above was replaced with the maximum correlation between the logarithm of concurrent baseflows.

A delete-1 jackknife simulation is performed to assess the performance of the baseflow correlation and scaling methods, where one site is designated as the ungauged basin and all other sites are considered possible donor sites. The "recession method" defined in Reilly and Kroll (2003) is used to pick independent baseflow measurements to form a baseflow segment. This method chooses a random starting year, a random starting recession, and a random baseflow from the recession to start the baseflow segment. Random baseflows from consecutive recessions (one flow per recession) are used until the baseflow segment reaches a specified length (again 4, 6, 8, 10, or 12 baseflow measurements). As suggested by Reilly and Kroll (2003), the total number of baseflow segments examined for an ungauged site is equal to the total number of baseflow days divided by the segment length. The filtering methods used previously to define baseflow days were also implemented for this technique, and the correlation between the log-space flows at the donor and ungauged sites needed to be at least 0.7.

3.7 | Performance metrics

Six performance metrics were used to assess each estimation method. Each performance metric was estimated at each site, and the average of each metric across all sites is reported. Four of these metrics are bias and mean square error (MSE) in real- and log-space (where all streamflows are in units of cubic feet per second). Real-space metrics generally are more influenced by an estimator's fit to larger streamflow values, whereas log-space metrics are more influenced by an estimator's fit to smaller streamflows. In addition, two relative metrics were calculated. The average relative absolute difference (ARAD) was calculated as

$$\mathrm{ARAD} = \frac{\sum_{i=1}^{N}\left|\frac{\widehat{Q_i} - Q_{\mathrm{obs}}}{Q_{\mathrm{obs}}}\right|}{N} \tag{6}$$

where $\widehat{Q_i}$ is the ith low streamflow estimate (in units of cubic feet per second), Q_{obs} is the at-site low flow statistic calculated using the entire record, and N is the number of estimates of $\widehat{Q_i}$ at the site. ARAD is a measure of the average per cent deviation of predicted values and is less influenced by the magnitude of the observations, though when estimating very small Q_{obs}, this metric can increase greatly. The unit area absolute difference (UAAD) is calculated as

$$\mathrm{UAAD} = \frac{\sum_{i=1}^{N}\left|\frac{\widehat{Q_i} - Q_{\mathrm{obs}}}{\mathrm{Area}_i}\right|}{N}, \tag{7}$$

and is a measure of the absolute difference scaled by drainage area (km^2) and is also less influenced by the magnitude of streamflow values. UAAD is useful when there is a large range of observed values within the data set, as it is not strongly influence by exceptionally small observations such as ARAD.

4 | RESULTS AND DISCUSSION

This section first provides an analysis of how well the proposed baseflow filters screen precipitation events during baseflow conditions (Section 4.1). Section 4.2 provides a comparison of possible donor site selection and baseflow filter methods using a robust rank-based evaluation (RRBE) and the identification of preferred donor site selection and baseflow filtering methods. Section 4.3 examines small-sample estimators of K_b and τ, including a donor site estimator not previously explored in the literature. Finally, Section 4.4 compares low streamflow estimators from regional regression, regression with estimated hydrogeologic indices, baseflow correlation, and scaling.

4.1 | Precipitation on baseflow days

The techniques examined in this analysis require identification of streamflow under baseflow conditions. Baseflow is determined using hydrologic separation. To assess the impact of small precipitation events on streamflow recessions, an analysis of precipitation and streamflow was performed for three watersheds: the French Broad River near Newport, TN (USGS#03455000), Talladega Creek at Alpine, AL (USGS#02406500), and the Flint River near Carsonville, GA (USGS#02347500). Precipitation data were obtained from a National Weather Service maintained monitoring site, where data were downloaded from the National Climatic Data Centre (Menne et al., 2012). All three watersheds exhibited similar results, so here, results are discussed only for the French Broad River watershed.

Figure 3 contains box plots for the French Broad River watershed for precipitation events during identified baseflow days for different baseflow filters. Of the 589 days originally identified as baseflow, 80 had measured precipitation events (14%), and five of these events were above 0.5 in. (1.27 cm). By dropping the last day identified as baseflow (L in Figure 3), only 36 of 404 baseflow days (9%) had precipitation events, with only two events above 0.5 in. By eliminating recessions where streamflow the day before the recession was larger than the 25th or 50th percentile (E_0.25 or E_0.50 in Figure 3), 64 of 492 baseflow days (13%) or 49 of 408 baseflow days (12%) had recorded precipitation events. Both of these methods failed to screen two precipitation events above 0.5 in. Finally, by both removing the last baseflow day and eliminating recessions preceded by large streamflows (L_0.25 or L_0.50 in Figure 3), 30 of 330 baseflow days (9%) or 22 of 273 baseflow days (8%) had recorded precipitation events. Both of these filters failed to screen one recorded precipitation event above 0.5 in. on a baseflow day.

The single precipitation event above 0.5 in. for L_0.25 and L_0.50 was a 1.2-in. (3.05 cm) event that occurred in the middle of a baseflow recession without an increase in measured streamflow at the stream gauge. This could be due to a measurement error at the precipitation gauge or an isolated storm, where the soil may be extremely dry, and all precipitation is absorbed with little or no response at the stream gauge (Faures, Goodrich, Woolhiser, & Sorooshian, 1995). To confirm this assumption, precipitation data were obtained on the same day (July 19, 2006) from all precipitation gauges within 100 km from the precipitation gauge where the 1.2 in. precipitation event was observed. Of the eight gauges observed, only one had a precipitation event, also recorded as 1.2 in. This suggests

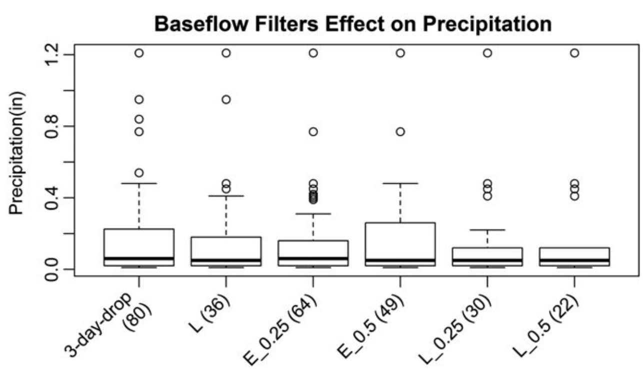

FIGURE 3 Precipitation events during baseflow for different baseflow filters at the French Broad River (USGS#03455000 and rain gauge USC00315356). Number of non-zero precipitation events in parentheses

that there was an isolated band of storms in this region at both of these stations. To confirm this, historic radar data were also observed (NOAA, 2006), which indicated isolated precipitation events on July 19, 2006 within this region.

Although the proposed filters do not remove all precipitation events during baseflow conditions, they do reduce the number and magnitude of these events. Of interest is whether such filters also help improve the performance of low streamflow estimation techniques. In Section 4.2, baseflow filters are paired with donor site selection techniques, and resulting low streamflow estimators are examined.

4.2 | Donor site selection

Donor sites are required to transfer information to calculate low streamflow statistics from spot measurements using baseflow correlation and scaling methods. Previously stated donor site selection methods (Section 3.5) are paired with a baseflow filtering technique (Section 4.1) and then compared using an RRBE, similarly used in Farmer et al. (2014). With this technique, each performance metric (Section 3.7) is calculated for each streamflow gauge for baseflow correlation (7Q10, 7Q2, 30Q10, and 30Q2) or scaling (Q95 and Q99), and the mean of each performance metric was calculated across all sites. Next, each donor site/filter selection method is ranked across each of the six performance metrics, where 1 is the rank of the best method. The mean and standard deviation of these six ranks are then calculated and plotted, creating a cross-metric RRBE point cloud of all donor site/filter selection methods (Farmer et al., 2014). This graphical tool helps identify more optimal donor site/filter selection methods and the trade-offs between methods.

Figure 4 shows the RRBE as a point cloud for each low streamflow statistic with the standard deviation of the ranks on the *y* axis, and the mean of the ranks across all performance metrics on the *x* axis. The optimal position is the lower left corner of each figure, where the specific donor site/filter selection method will have the lowest mean and standard deviation of the ranks.

For 7Q10, 30Q2, and 30Q10, removing the last day of baseflow, excluding recessions with streamflow the day before the recession greater than the 50th percentile and using the NN as the donor site (L_0.5_NN) have the lowest mean rank and standard deviation. For 7Q2, using a donor site with the lowest estimator variance within 100 km was best, and removing the last day of the recession had less of an impact (E_0.5_100 had the lowest mean rank; E_0.5_100 and L_0.5_100 had the lowest standard deviation). Q95 and Q99 display different patterns from the other low streamflow statistics. For Q95, using NN as the donor site and the 50th percentile, baseflow filter (E_0.5_NN) had the lowest mean rank and standard deviation. For Q99, using gauges within 100 km with a drainage area within ±50% and the minimum estimator variance with the 50th percentile baseflow filter and removing the last baseflow (L_0.5_100area) had the lowest average mean rank, whereas E_0.5_NN had the lowest average standard deviation. Using these results, further analysis of methods to estimate hydrogeologic indices and low streamflow statistics is performed in the following sections using L_0.5_NN, L_0.5_100area, and E_0.5_NN donor site/filtering selection methods.

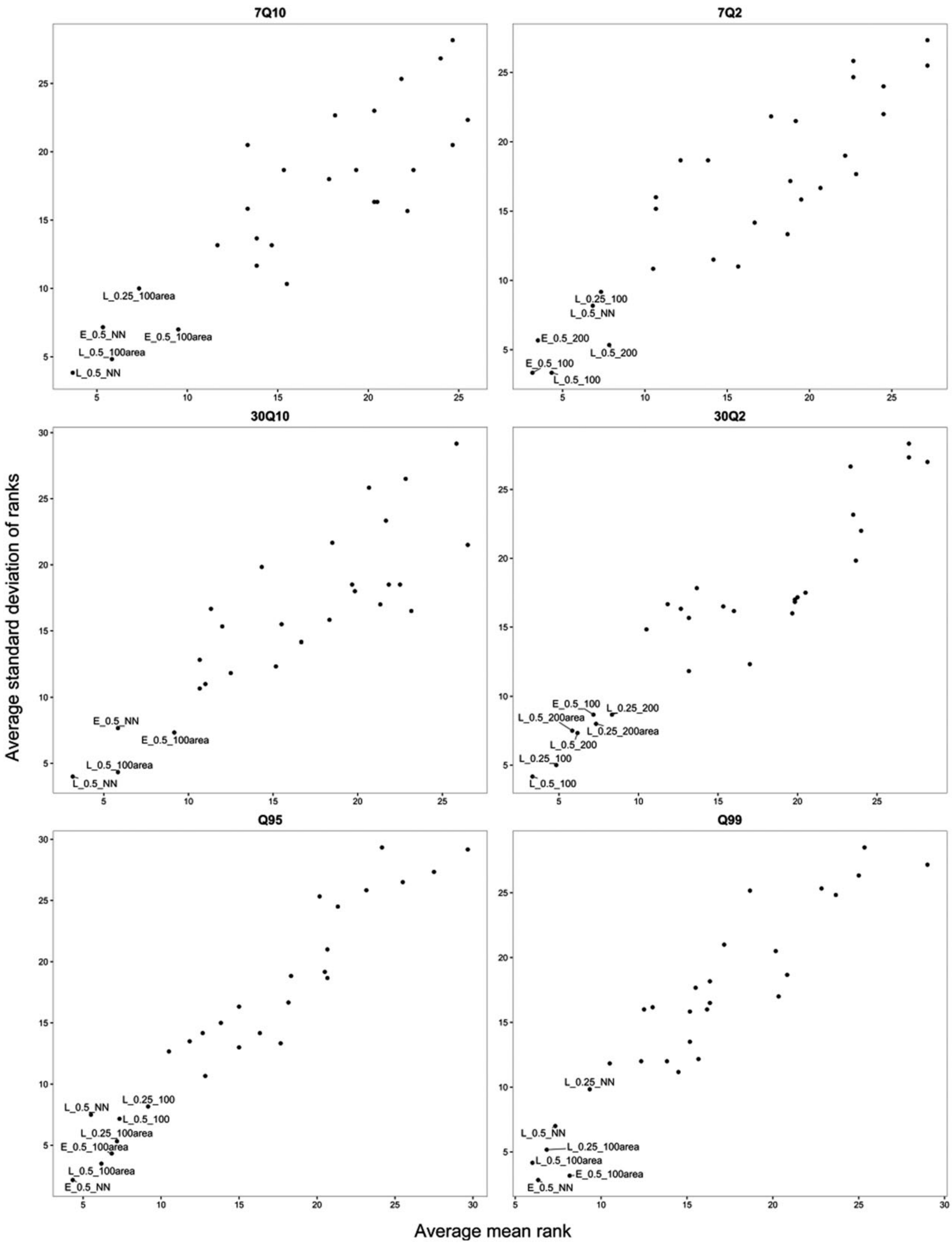

FIGURE 4 Robust rank-based evaluation of all donor site selection methods and baseflow filters for each low streamflow statistic

4.3 | Small-sample estimation of hydrogeologic indices

A comparison of K_b and τ estimators from SR with a donor site (DS) estimators is displayed in Figure 5. The donor site estimator for K_b is

$$K_{b,DS,UG} = K_{b,SR,UG}\left(K_{b,HR,DS}/K_{b,SR,DS}\right), \tag{8}$$

where $K_{b,SR,UG}$ and $K_{b,SR,DS}$ are the estimators of K_b from the SR at the ungauged site and donor site, respectively, and $K_{b,HR,DS}$ is the estimator of K_b from the historic record at the donor site. A

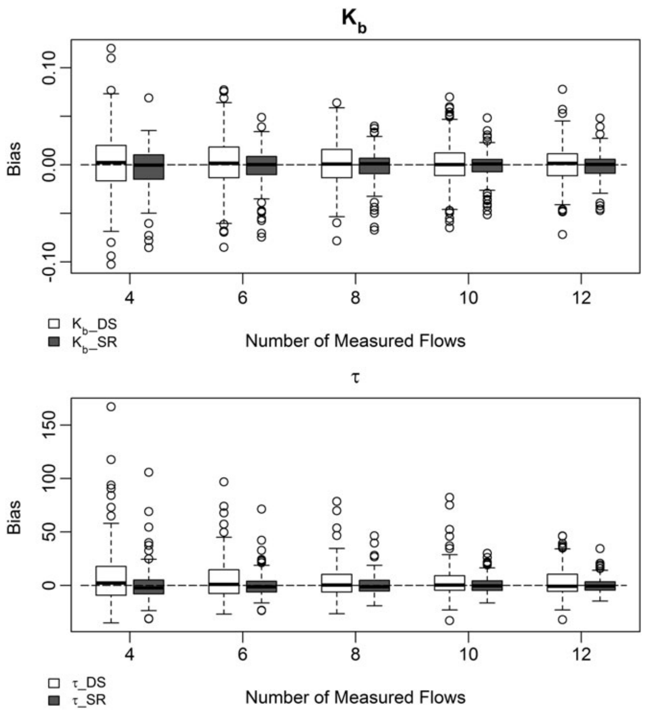

FIGURE 5 K_b and τ comparison between methods estimated from short record (SR) and using the L_0.5_NN donor site (DS)/filtering selection method

donor site estimator of τ had a similar form as Equation 8. The L_0.5_NN donor site/filtering selection method was used to select baseflows for this analysis on the basis of its good performance for estimating both 7Q10 and Q95. At each site, the real-space biases of K_b and τ are estimated by comparing SR and DS estimates to those from the historic record at the site. Figure 5 contains box plots of these biases across all sites. Results show that using the NN donor site to estimate K_b and τ always increased the range of bias in K_b and τ, regardless of the number of baseflows used in the analysis. Differences in hydrogeology at the donor site compared to the ungauged site may cause this increased bias. On the basis of this result, subsequent analyses of regression plus hydrogeology indices only use SR estimators of K_b and τ.

4.4 | Comparison of low streamflow estimation methods

For all low flow regional regression models, the first four entering explanatory variables were the same. These variables are summarized in Table 1. These are the same explanatory variables chosen by Pugliese et al. (2016) and Croteau et al. (2016) in their regional regression models for Q95 within this region. The adjusted coefficient of determination (Adj-R^2) of regression models without hydrogeology ranged from 0.57 for 7Q10 to 0.71 for 30Q2; when K_b from the historic record was added, Adj-R^2 ranged from 0.88 to 0.91 and with τ from 0.85 to 0.88.

Figures 6 (for 7Q10) and 7 (for Q95) present the average ARAD, UAAD, bias, log-bias, MSE, and log-MSE across all sites for regression, regression with hydrogeologic indices, and baseflow correlation and scaling for 4, 6, 8, 10, and 12 baseflow measurements. For 7Q10 and Q95, the two best combinations of baseflow filters and donor selection methods are presented: L_0.5_NN and L_0.5_100area for 7Q10 and E_0.5_NN and L_0.5_NN for Q95. Results for 7Q2, 30Q10, and 30Q2 were similar to those for 7Q10, and results for Q99 were similar to those for Q95, and thus, results are only shown for 7Q10 and Q95, both of which are commonly employed in practice.

In Figures 6 and 7, the horizontal lines represent results for OLS regression without hydrogeology (Reg) and OLS regression with K_b and τ estimated from the entire historic (Reg + K_b and Reg + τ); these methods do not use spot measurements. Regression equations with SR hydrogeologic indices are represented by squares and triangles (Reg + Est K_b and Reg + Est τ), and baseflow correlation and scaling methods are represented by circles (BFC and Scaling). MSE is presented in terms of relative performance, where MSE relative performance$_i$ = $MSE_{Reg\ +\ Kb}/MSE_i$; this was done to avoid plotting MSE, which had some large values. When a figure is in terms of relative performance, greater values (i.e., larger values on the y axis) indicate methods that are performing better. Biases should be close to zero, whereas ARAD and UAAD should be as small as possible.

Figure 6 displays the results for 7Q10 for each performance metric where the donor site/filtering selection method is L_0.5_NN (solid symbols). For ARAD and log-MSE, baseflow correlation outperforms all other methods; for UAAD, baseflow correlation and scaling produce similar results. For ARAD and UAAD at 6 measurements, baseflow correlation outperforms regression with hydrogeologic indices estimated from the entire record. Baseflow correlation had a higher bias than the other methods, though its bias is similar to that of regression equations using the entire record. For log-bias, baseflow correlation performs well at 4 and 6 measurements, but when the number of measurements increases, the bias of baseflow correlation increases, which could be due to the decrease in the number of baseflow segments analysed at larger baseflow segments (8, 10, and 12 measurements). This may also be due to this method performing worse at sites with smaller at-site low streamflow statistics. For MSE and log-MSE after eight measurements, baseflow correlation outperforms all regression models.

Figure 6 also displays results for 7Q10 and baseflow correlation and scaling methods when the donor site/filtering selection method is L_0.5_100area (open symbols). Because L_0.5_NN and L_0.5_100area filter baseflows the same way (they only differ in donor site selection), the results for Reg + Est K_b and Reg + Est τ are the same for either filter. The results for this filter are identified as BCF_2 and Scaling_2. The MSE of 7Q10 for BFC and scaling with L_0.5_100area is less than with L_0.5_NN (higher relative performance), whereas for log-bias (and bias for 4 and 6 measurements), the opposite is true. This appears to indicate that choosing the donor site within 100 km, a

TABLE 1 Summary of watershed characteristics values that were used in the models

Characteristics	Median	Range across all sites
Drainage area (km^2)	423	10 to 4,799
Average basin precipitation (mm)	1,460	1,140 to 2,070
Rock depth (m)	1.40	0.48 to 1.52
Slope (%)	8.66	0.419 to 47.8

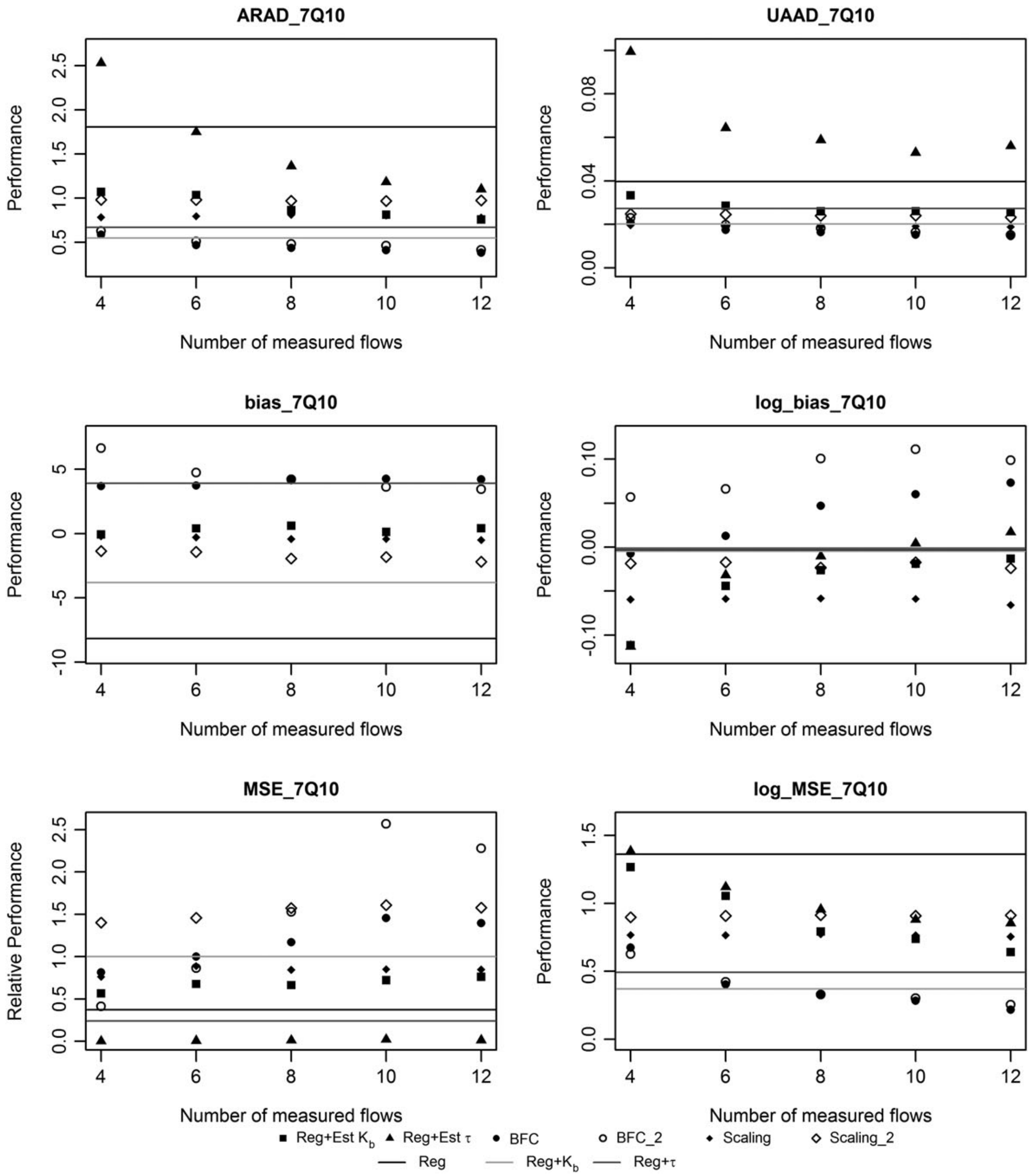

FIGURE 6 Comparison of all methods used to predict 7Q10 for all performance metrics, where baseflow correlation (BFC) and scaling use L_0.5_NN and BFC_2 and Scaling_2 use L_0.5_100area donor site/filtering selection methods; regression methods that do not employ a donor site use L_0.5. Reg + Est τ is removed for bias because of large values. Mean square error (MSE) relative performance = $MSE_{Reg\,+\,Kb}/MSE_i$, where higher performance is better. ARAD = rage relative absolute difference; UAAD = unit area absolute difference

drainage area ±50% from the site of interest, and the minimum variance 7Q10 estimator (L_0.5_100area) improves the fit over NN (L_0.5_NN) at sites with large 7Q10 values, but it performs worse at sites with smaller 7Q10 values. Baseflow correlation with either filter performs similarly.

In Figure 6, regression equations with hydrogeologic indices calculated using the entire record outperform regression equations without K_b and τ, except for the MSE of τ, which is much larger than for the other methods (thus decreasing the MSE relative performance). For 7Q10, across all performance metrics, Reg + K_b and Reg + Est K_b always outperforms Reg + τ and Reg + Est τ. In addition, even with only four measurements, Reg + Est K_b is always better than performing regression with no hydrogeologic indices, an important result for practitioners.

Figure 7 displays the results for Q95 for each performance metric where the donor site/filtering selection method is E_0.5_NN. Note that baseflow correlation is not analysed for Q95. For ARAD, UAAD, bias, and MSE, the scaling method outperforms or performs as well as regression equations plus hydrogeologic indices calculated from the entire record, even when only four measurements are used. Scaling has a slightly larger negative log-bias than the regression models regardless of the number of measurements, though it has a slightly smaller log-MSE than regression plus hydrogeologic indices calculated from the entire record when eight or more measurements are used.

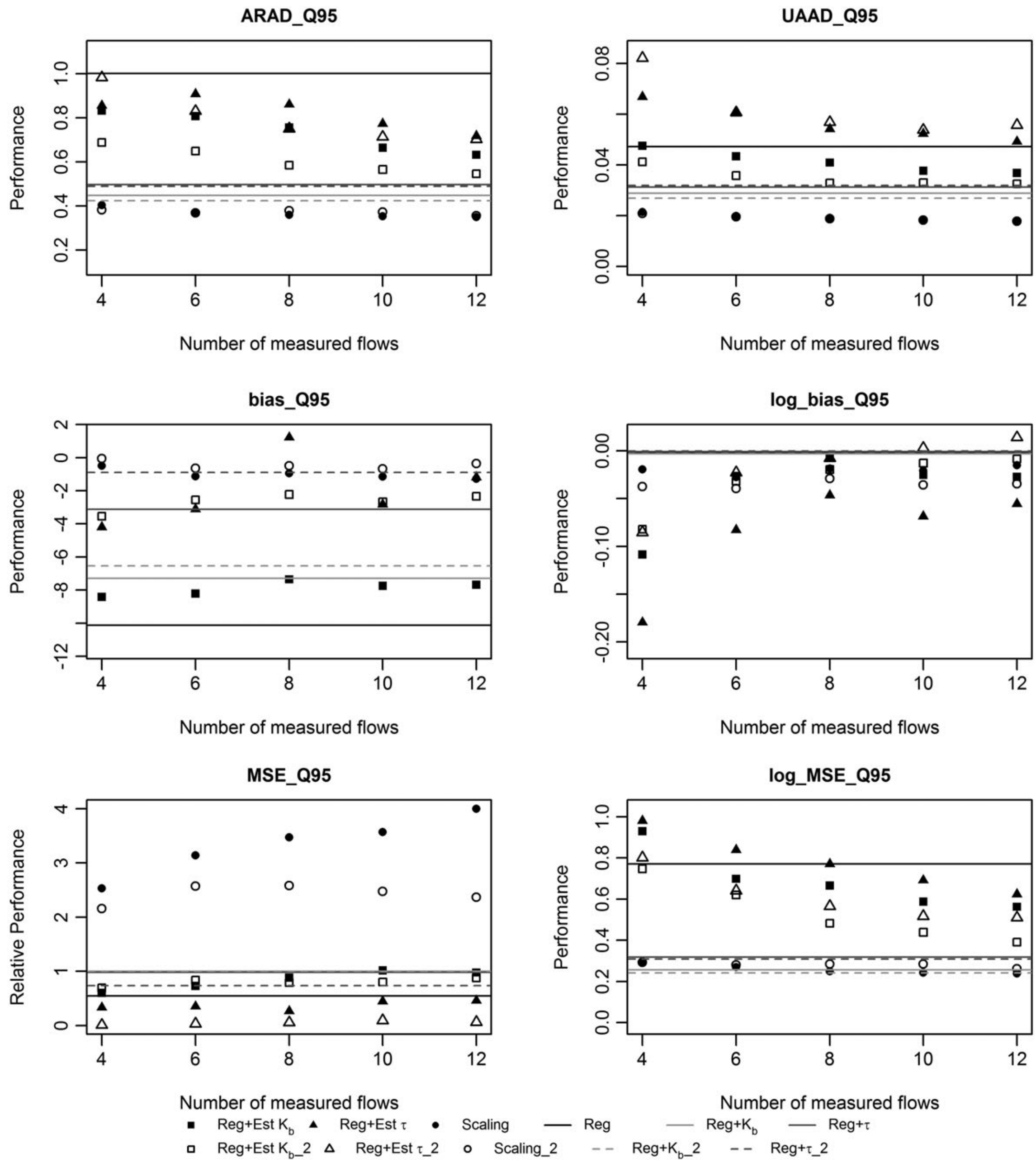

FIGURE 7 Comparison of all methods used to predict Q95 for all performance metrics, where Reg + Est K_b, Reg + Est τ, and scaling use E_0.5_NN; and Reg + Est K_b_2, Reg + Est τ _2, and Scaling_2 use L_0.5_NN. Reg + Est τ _2 is removed for bias because of large values. Mean square error (MSE) relative performance = $MSE_{Reg\ +\ Kb}/MSE_i$, where higher performance is better. ARAD = rage relative absolute difference; BFC = baseflow correlation; UAAD = unit area absolute difference

This result indicates that scaling may perform better as sites with larger flows. Across all performance metrics except for bias, for Q95, Reg + K_b and Reg + Est K_b always outperforms Reg + τ and Reg + Est τ. Reg + Est K_b performs better than regression without hydrogeologic indices for all performance metrics except log-bias when six or more measurements are available.

Also included in Figure 7 are the results for the L_0.5_NN donor site/filtering selection method (all results with this filter are indicated as "_2" and have open symbols). Because this method filters baseflows differently than E_0.5_NN, results are presented for all methods. Across all performance metrics, all methods that use the L_0.5_NN donor site/filtering method are outperformed by methods that use E_0.5_NN, except for bias.

The results presented in Figures 6 and 7 are averages across all study sites. Figure 8 contains box plots of UAAD for 4, 8, and 12 measurements at individual sites for each estimation method with the L_0.5_NN donor site/filtering method; note that the scale on the *y* axis is in log-space. These plots are used to assess if any sites are performing exceptionally poorly and thus have a large influence on the average results in Figures 6 and 7. Again, horizontal lines represent regional regression; ideally, box plots would be below these lines. Across all measured flows and all methods, there are some sites that perform poorly (worse than regression), but baseflow correlation and scaling methods have fewer of these sites. For regression methods, this could be due to the variability of small-sample estimators of hydrogeologic indices, whereas for baseflow correlation and scaling

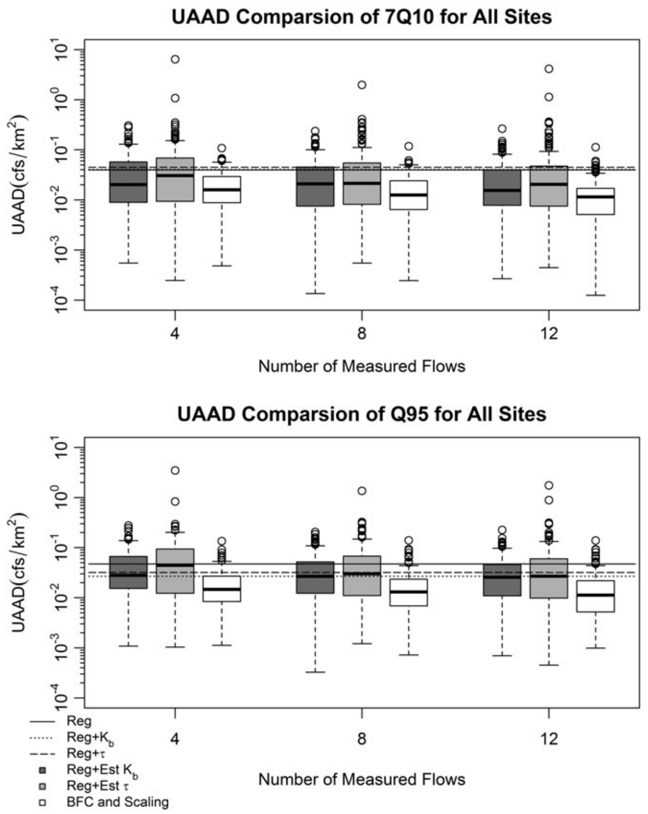

FIGURE 8 Comparison of box plots of unit area absolute difference (UAAD) for all estimation methods using spot measurements with L_0.5_NN donor site/filtering method. *Y* axis is log-scale

methods, this could be due to poor donor site selection or unusual baseflow observations (such as when small precipitation events occur). All methods that use spot measurements improve estimators of low streamflow statistics as the number of baseflow measurements is increased.

5 | CONCLUSIONS

This experiment examined estimators of low streamflow statistics at ungauged river sites when a nominal number of streamflow measurements are available at the ungauged site. Four estimation techniques were examined: regional regression, regional regression plus estimated hydrogeologic indices, baseflow correlation, and scaling (that latter two use a donor site to improve estimation). These techniques were used to predict the six different streamflow statistics, and results were presented for 7Q10 and Q95. The following conclusions can be made from this analysis:

- With eight measurements (and sometimes as few as four), baseflow correlation and scaling methods always outperform regional regression even when at-site estimators of hydrogeologic indices are available. The good performance of scaling compared to regional regression contradicts the findings of other researchers.
- The best donor site/baseflow filtering selection method is for the streamflow the day before the first measured streamflow to be smaller than the median daily streamflow at the donor site, for the nearest gauged streamflow site to be used as the donor site.
- Adding hydrogeologic indices improves low streamflow regional regression models, with baseflow recession constants, K_b, always produced better regression models than the aquifer time constant, τ.
- Use of a small-sample estimator of K_b improved low streamflow regional regression models even when only four measurements are taken; this was generally not true for τ.

These results reaffirm the importance of hydrogeology in low streamflow prediction. Results from this experiment overwhelmingly indicate that using a donor site to transfer information with concurrent streamflow measurements during baseflow conditions, such as baseflow correlation and scaling, is preferred to regional regression for estimation of low streamflow statistics.

ACKNOWLEDGMENTS

The authors would like to thank the Department of Environmental Resources Engineering at SUNY ESF for providing financial support for this experiment. In addition, we would like to recognize Charity Nyelele who helped developed Figure 1.

ORCID

Timothy J. Stagnitta http://orcid.org/0000-0001-8903-428X
Charles N. Kroll http://orcid.org/0000-0003-1157-8353

REFERENCES

Aksoy, H., & Wittenberg, W. (2011). Nonlinear baseflow recession analysis in watersheds with intermittent streamflow. *Hydrological Sciences Journal*, *56*(2), 226–237.

Archfield, S. A., & Vogel, R. M. (2010). Map correlation method: Selection of a reference streamgage to estimate daily streamflow at ungaged catchments. *Water Resources Research*, *46*, 1–15.

Archfield, S. A., Vogel, R. M., Steeves, P. A., Brandt, S. L., Weiskel, P. K., & Garabedian, S. P. (2009). The Massachusetts sustainable-yield estimator: A decision-support tool to assess water availability at ungaged stream locations in Massachusetts. U.S. Geological Scientific Investigations Report 2009–5227.

Arnold, J. G., & Allen, P. M. (1999). Automated methods for estimating baseflow and ground water recharge from streamflow records. *American Water Resources Association*, *35*, 411–424.

Barnes, C. R. (1986). Method for estimating low-flow statistics for ungaged streams in the lower Hudson River Basin. New York. U.S. Geological Survey Water Resorces Investigations Report 85–4070.

Bras, R. L. (1990). Hydrology: An introduction to hydrologic science, Addison-Wesley-Longman, Reading, Massachusetts.

Broda, S., Larocque, M., & Paniconi, C. (2014). Simulation of distributed base flow contributions to streamflow using a hillslope-based catchment model coupled to a regional-scale groundwater model. *Journal of Hydrologic Engineering*, https://doi.org/10.1061/(ASCE)HE.1943-5584.0000877, 907–917.

Brutsaert, W., & Lopez, J. P. (1998). Basin-scale geohydrologic drought flow features of riparian aquifers in the southern Great Plains. *Water Resources Research*, *34*, 233–240.

Brutsaert, W., & Nieber, J. L. (1977). Regionalized drought flow hydrographs from a mature glaciated plateau. *Water Resources Research*, *13*, 637.

Clark, P. J., & Evans, F. C. (1954). Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecological Society of America*, *35*, 445–453.

Croteau, K. E., Kroll, C. N., Over, T. M., & Archfield, S. A. (2016). Estimation of flow duration curve quantiles in ungauged river sites. *Journal of Hydrology*, submitted March 2016.

Eng, K., Kiang, J. E., Chen, Y. Y., Carlisle, D. M., & Granato, G. E. (2011). Causes of systematic over- or underestimation of low streamflows by use of index-streamgage approaches in the United States. *Hydrological Processes*, *25*, 2211–2220.

Eng, K., & Milly, P. C. D. (2007). Relating low-flow characteristics to the base flow recession time constant at partial record stream gauges. *Water Resources Research*, *43*, 1–8.

Falcone, J. A., Carlisle, D. M., Wolock, D. M., & Meador, M. R. (2010). GAGES: A stream gage database for evaluating natural and altered flow conditions in the conterminous United States: Ecological archives E091-045. *Ecology*, *91*, 621.

Farmer, W., Archfield, S., Over, T., Hay, L., Jacob, L., & Kiang, J. (2014). A comparison of methods to predict historical daily streamflow time series in the southeastern United States. U.S. Geological Survey Scientfic Investigations Report 2014–5231.

Faures, J.-M., Goodrich, D. C., Woolhiser, D. A., & Sorooshian, S. (1995). Impact of small-scale spatial rainfall variability on runoff modeling. *Journal of Hydrology*, *173*, 309–326.

Fennessey, N. M. (1994). A hydro-climatological model of daily streamflow for the northeast United States. Ph.D Dissertation Tufts Univeristy.

Hirsch, R. M. (1979). An evaluation of some record reconstruction techniques. *Water Resources Research*, *15*, 1781–1790.

Institute of Hydrology (1980). *Low flow studies*. Wallingford, U.K: Institute of Hydrology.

Kroll, C., Luz, J., Allen, B., & Vogel, R. M. (2004). Developing a watershed characteristics database to improve low streamflow prediction. *Journal of Hydrologic Engineering*, *9*, 116–125.

Kroll, C. N., & Stedinger, J. R. (1998). Regional hydrologic analysis: Ordinary and generalized least squares revisited. *Water Resources Research*, *34*, 121–128.

Kroll, C. N., & Stedinger, J. R. (1999). Development of regional regression relationships with censored data. *Water Resources Research*, *35*, 775–784.

Laaha, G., & Blöschl, G. (2005). Low flow estimates from short stream flow records – A comparison of methods. *Journal of Hydrology*, *306*, 264–286.

Matonse, A. H., & Kroll, C. N. (2013). Applying hillslope-storage models to improve low flow estimates with limited streamflow data at a watershed scale. *Journal of Hydrology*, *494*, 20–31.

Menne, M. J., Durre, I., Korzeniewski, B., McNeal, S., Thomas, K., Yin, X., Anthony, S., … Houston, T. G. (2012). Global Historical Climatology Network – Daily (GHCN-Daily), Version 3. Subset used: October 1980 – September 2010. NOAA National Climatic Data Center.

NOAA National Weather Service (NWS) (2006). Radar operations center (1991): NOAA Next Generation Radar (NEXRAD) level II base data. (Knoxville,). NOAA National Centers for Environmental Information.

Potter, K. (2001). A simple method for estimating baseflow at ungages locations. *American Water Resources Association*, *37*, 177–184.

Pugliese, A., Farmer, W. H., Castellarin, A., Archfield, S. A., & Vogel, R. M. (2016). Regional flow duration curves: Geostatistical techniques versus multivariate regression. *Advances in Water Resources*, *96*, 11–22 https://doi.org/10.1016/j.advwatres.2016.06.008.

Reilly, C. F., & Kroll, C. N. (2003). Estimation of 7-day, 10-year low streamflow statistics using baseflow correlation. *Water Resources Research*, *39*. https://doi.org/10.1029/2002WR001740.

Ries, K., & Friesz, P. (2000). Methods for estimating low-flow statistics for Massachusetts streams. U.S. Geological Survey Water-Resources Investigations Report 00–4135.

Riggs, H. C. (1980). Characteristics of low flows. *Journal of the Hydraulics Division*, *106*, 717–731.

Risley, J., Stonewall, A., & Haluska, T. (2008). Estimating flow-duration and low-flow frequency statistics for unregulated streams in Oregon. U.S. Geological Survey Scientific Investigations Report 2008–5126.

Rossman, L. A. (1990). DFLOW user's manual. US EPA risk reduction Engineering Laboratory, Office of Research and Development, Cincinnati, OH, 31 pp.

Rumenik, R. P., & Grubbs, J. W. (1996). Methods for estimating low-flow characteristics of ungaged streams in selected areas, Northern Florida. U.S. Geological Survey Water Resorces Investigations Report 96–4124.

Smakhtin, V. U. (2001). Low flow hydrology: A review. *Journal of Hydrology*, *240*, 147–186.

Susquehanna River Basin Commission (2012). Technical guidance for low flow protection related to withdrawal approvals, Harrisburg, Pennsylvania.

Stedinger, J. R., & Thomas, W. (1985). Low-flow frequency estimation using base-flow measurements. U.S. Geological Survey Open-File Report 85–95.

Stedinger, J. R., Vogel, R. M., & Foufoula-Georgiou, E. (1993). Frequency analysis of extreme events. Chapter 18 in handbook of hydrology, edited by DR Maidment.

Tallaksen, L. M. (1995). A review of baseflow recession analysis. *Journal of Hydrology*, *165*, 349–370. https://doi.org/10.1016/0022-1694(94)02540-R.

Tasker, G. D. (1980). Hydrologic regression with weighted least squares. *Water Resources Research*, *16*, 1107–1113.

Tasker, G. D., & Stedinger, J. R. (1989). An operational GLS model for hydrologic regression. *Journal of Hydrology*, *111*, 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 55.

USGS (2012). Instantaneous Data Archive – IDA. Available from: http://ida.water.usgs.gov/ida/index_usgs.cfm

USGS (2014). WaterSMART. Available from: http://water.usgs.gov/watercensus/WaterSMART.html

Vogel, R. M., & Fennessey, N. M. (1994). Flow-duration curves. I: New interpretation and confidence intervals. *Journal of Water Resources Planning and Management*, *120*(4), 485–504.

Vogel, R. M., & Kroll, C. N. (1992). Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics. *Water Resources Research*, *28*, 2451–2458.

Vogel, R. M., & Kroll, C. N. (1996). Estimation of baseflow recession constants. *Water Resources Management*, *10*, 303–320.

Wandle, Jr, S. W., & Randall, A. D. (1994). Effects of surficial geology, lakes and swamps, and annual water availability on low flows of streams in Central New England, and their use of low-flow estimation. U.S. Geological Survey Water-Resources Investigations Report 93–4092.

Wittenberg, H. (2003). Effects of season and man-made changes on baseflow and flow recession: Case studies. *Hydrological Processes*, *17*, 2113–2123.

Zhang, Z., & Kroll, C. (2007a). Closer look at the baseflow correlation method. *Journal of Hydrologic Engineering*, *12*, 190–196.

Zhang, Z., & Kroll, C. (2007b). The baseflow correlation method with multiple gauged sites. *Journal of Hydrology*, *347*, 371–380.

How to cite this article: Stagnitta TJ, Kroll CN, Zhang Z. A comparison of methods for low streamflow estimation from spot measurements. *Hydrological Processes*. 2018;32:480–492. https://doi.org/10.1002/hyp.11426