Quantification of Variabilities of Baseflow of Watersheds.

REU Site: Interdisciplinary Program in High Performance Computing

Christian Dixon[1], Gabriel Martinez[2], Hwan Hee Park[3], Maddie Rainey[4],
Graduate assistant: Sai Popuri[1] and Nadeesri Wijekoon[1],
Faculty mentor: Kofi Adragni[1], and Client: Jeff Raffensperger[5].

[1]Department of Mathematics and Statistics, UMBC,
[2]Department of Mathematics, California State University, Fullerton,
[3]Department of Computational and Applied Mathematics, Rice University,
[4]Department of Mathematics, Grand Valley State University,
[5]U.S. Geological Survey (USGS)

Abstract

The U.S. Geological Survey National Water-Quality Assessment Project conducted a study of 225 sites in the Chesapeake Bay watershed to estimate base flow. Baseflow is the estimated volumetric discharge of water, primarily from groundwater sources, that is relayed to the measurement sites. The study is necessary in order to address the nation's water supply for changes in the environment. Baseflow is estimated using a recursive digital filter. Calculating the variability of baseflow water discharge is important to make informed decisions about water regulation. We explored the estimation of variability of baseflow using two methods: the bootstrap method and the Delta method. Each method has its own limitations and requirements. Ultimately, bootstrapping was shown to be a reasonable recommendation for estimating baseflow variability. The bootstrapping algorithm was parallelized in order to compute numerous iterations on multiple processors for big data analysis. The derivation of the variability of a non-constant streamflow was also considered for further study, but not implemented.

1 Introduction

The U.S. Geological Survey National Water-Quality Assessment Project collected daily-mean streamflow and daily estimated baseflow data from 1913 to 2016 from 225 sites in the Chesapeake Bay watershed. Baseflow is the estimated volumetric discharge of water that is relayed to the measurement sites. Estimation of water discharge is essential in addressing questions of aquatic ecosystems to environmental changes. Several methodologies have been used to estimate water discharge. Particularly, the following hydrography separation methods have been used to estimate baseflow: PART (Rutledge, 1998), HYSEP (Sloto and Crouse, 1996), and BFI along with the Recursive Digital Filter (Eckhardt, 2005). PART method obtains the mean rate of groundwater discharge though linear interpolation of daily values of streamflow. The implementation of the HYSEP computer program separates streamflow into baseflow and quickflow components. BFI method estimates baseflow by connecting a smooth minima technique on hydrograph separation to a stream hydrograph. A generalized Recursive Digital Filter (RDF) estimates baseflow on the assumption of a linear reservoir with two adjustable parameters.

In a recent study by a team of researchers at USGS (Raffensperger et al, 2017) these hydrography separation methods were evaluated. The study recommended RDF as the optimal method. The RDF method is a theoretical framework for filter algorithms introduced by Eckhardt (2005) that developed a two-parameter filter equation. The parameters are α, the recessive constant, and β, the maximum value of the baseflow index (BFI). The parameter α is determined through data analysis, and β is estimated based on characteristics of the site, and depends on the value of α. The recursive digital filter equation is

$$Q_{Bj} = \frac{(1-\beta)\alpha Q_{Bj-1} + (1-\alpha)\beta Q_j}{(1-\alpha\beta)} \tag{1.1}$$

where Q_j is the streamflow (L^3/t), Q_B is the baseflow (L^3/t), where $j = 1, \cdots, T$ are the days streamflow was recorded. With the recursive digital filter, the estimation of the baseflow is carried out using the α, recessive constant. Although it is assumed to be a constant, α is actually an estimate that carries an uncertainty. While the value of β depends on α, the empirical distribution of α is skewed. So far in the study of Raffensperger et al. (2017), the median value has been utilized in the estimation of Q_{B_j}. The primary purpose of this project is to quantify the variability around the estimates of the baseflow Q_{B_j} induced by the variability of the recessive constant α. This helps determine a margin of error around the estimate of baseflow.

Two statistical methods are considered to estimate the variability of Q_{Bj}: the bootstrap method and the Delta method. The bootstrap method is a re-sampling method that is computationally intensive, while the Delta method using Taylor series expansion is faster and simpler, but likely unreliable for the setting herein presented.

The remainder of this report is as follows. In section 2, we describe the statistical methods we considered for the estimation of the variability. In section 3, we present the computational implementation of these methods. The application of the methods to the streamflow data are presented in section 4 along with some evaluation of the performance of the distributed computing system for parallel computing.

2 Statistical Methods

To calculate the baseflow, Q_B, we investigated two statistical methodologies: bootstrapping (Efron, 1994) and the Delta method. The bootstrap method is an intensive process based on a resampling of the data points, while the Delta method is a less intensive process and built upon some mathematical assumptions. We will implement both methods and compare their performance and accuracy. Ideally, if both methods produce the same accuracy, then we will use the more efficient method.

We start by obtaining an explicit expression of equation (1.1) which can be used in the estimation of the variability of streamflow. Let τ_1 and τ_2 be functions of α and defined as

$$\tau_1 = \left[\frac{(1-\beta)\alpha}{1-\alpha\beta}\right] \text{ and } \tau_2 = \left[\frac{(1-\alpha)\beta}{1-\alpha\beta}\right],$$

then expression (1.1) can be written as $Q_{Bj} = \tau_1 Q_{Bj-1} + \tau_2 Q_j$. As a recursive expression, the first term can be obtained $Q_{B1} = \tau_1 Q_{B0} + \tau_2 Q_1$. The subsequent two terms can be expanded as

$$\begin{aligned} Q_{B2} &= \tau_1 Q_{B1} + \tau_2 Q_2 \\ &= \tau_1 \left[\tau_1 Q_{B0} + \tau_2 Q_1\right] + \tau_2 Q_2 \\ &= \tau_1^2 Q_{B0} + \tau_1 \tau_2 Q_1 + \tau_2 Q_2 \end{aligned}$$

$$\begin{aligned} Q_{B3} &= \tau_1 Q_{B2} + \tau_2 Q_3 \\ &= \tau_1 \left[\tau_1^2 Q_{B0} + \tau_1 \tau_2 Q_1 + \tau_2 Q_2\right] + \tau_2 Q_3 \\ &= \tau_1^3 Q_{B0} + \tau_1^2 \tau_2 Q_1 + \tau_1 \tau_2 Q_2 + \tau_2 Q_3. \end{aligned}$$

Generalizing the expansion, and letting $Q_{B0} = Q_1$, we obtain the explicit expression of the baseflow as

$$Q_{Bj} = \tau_1^j Q_1 + \tau_1^{j-1} \tau_2 Q_1 + \tau_1^{j-2} \tau_2 Q_2 + \cdots + \tau_1 \tau_2 Q_{j-1} + \tau_2 Q_j. \tag{2.1}$$

That is, the baseflow is estimated using the streamflow on the previous days including the current day of interest j, with $j = 1, \cdots, T$, where T is known.

2.1 The Bootstrap Method

Bootstrapping is used for making statistical inferences. It helps in estimating the variance of an estimator for example by estimating these variances when sampling from an approximating distribution. This is a newer method because of its use of modern computer power to help simplify and quicken the estimations of variances and other statistics in complex settings. It is a data-based method that will allow us to calculate the variance of the baseflow, Q_B.

The goal is to evaluate the variance of Q_{B_j} at time j. To do so, we assume that β is a known function of α. To distinguish the difference between a statistically defined 'parameter' and its estimate, we will denote by $\widetilde{\alpha}$, the median value of the observed α values. Furthermore, we assume the conditioning on streamflow $Q_1, \cdots, Q_j$. We thus rewrite the above expression (1.1) as

$$Q_{B_j}|(Q_1, \cdots, Q_j) = g(\widetilde{\alpha})$$

Again, the notation $\widetilde{\alpha}$ is to assert that the value is an actual estimate, and thus it is subject to variabilities. However, we assume that $Q_1, \cdots, Q_j$ are deterministic. We aim at estimating the variance of $Q_{B_j}|(Q_1, \cdots, Q_j)$, which means that $(Q_1, \cdots, Q_j)$ are treated as known, fixed and not random. We assume for now that the estimate $\widetilde{\alpha}$ is the sole source of variability in $Q_{B_j}|(Q_1, \cdots, Q_j)$.

The estimate $\widetilde{\alpha}$ is the median of the empirical distribution, not the mean. We assume that $\widetilde{\alpha}$ follows a distribution D with true median α and variance σ^2. The empirical distribution of α is skewed. Following the method described by Dr. Raffensperger, the empirical distribution of α is generated. Let $A = \{\alpha_i^*, i = 1, \cdots, n\}$ be the values generated and let K be a large number, say $K = 10000$ for example. The bootstrap procedure is as follows:

For $k = 1, \cdots, K$

1. Sample n values out of A with replacement.
2. Obtain the median value $\widetilde{\alpha}_k$
3. Compute the corresponding estimate of β.
4. Compute $Q_{B_j}^{(k)}$ using $\widetilde{\alpha}_k$ and the estimate of β.

Once the K values of Q_{B_j} are obtained, compute $\overline{Q}_{B_j} = (1/K)\sum_{k=1}^{K} Q_{B_j}^{(k)}$. The estimated variance is then estimated as

$$\hat{\sigma}^2 = \hat{\sigma}^2(Q_{B_j}|(Q_1, \cdots, Q_j)) = \frac{1}{K}\sum_{k=1}^{K}[Q_{B_j}^{(k)} - \overline{Q}_{B_j}]^2$$

This last expression will be considered as an estimate of the variance of Q_{B_j} given $(Q_1, \cdots, Q_j)$.

2.2 Delta Method

Let X be a random variable with known variance, and let $Y = g(X)$ for some function g, with some continuity and derivability assumptions. The Delta method allow an estimation of the variance of Y by expanding the function $g(X)$ about its mean, with a one-step Taylor approximation, and then takes the variance. In our case, the Delta method uses Taylor series expansion to linearly approximate the variance of baseflow, Q_B, around the mean of the α values. Assume that the variance of Q_B is induced by the variance of α, and that β is a function of α, let $Q_{Bj} = g(\alpha)$ and $\bar{\alpha} = E(\alpha)$, the expectation or mean value of α. By Taylor series expansion on $g(\alpha)$, we get

$$g(\alpha) = g(\bar{\alpha}) + g'(\bar{\alpha})(\alpha - \bar{\alpha}) + \frac{g''(\bar{\alpha})(\alpha - \bar{\alpha})^2}{2!} + \cdots$$

By first order approximation, we can write $g(\alpha) \approx g(\bar{\alpha}) + g'(\bar{\alpha})(\alpha - \bar{\alpha})$. Assuming that $\bar{\alpha}$ is non-random, the variance of $g(\alpha)$ is given by

$$Var(Q_{B_j}) = Var(g(\alpha)) = [g'(\bar{\alpha})]^2 \, Var(\alpha) = [g'(\bar{\alpha})]^2 \, \sigma^2. \tag{2.2}$$

Once the variance of α is computed, the function $g'(.)$ is to be evaluated at $\bar{\alpha}$. The estimated variance is directly calculated using all observed values of α.

2.3 Estimation with Random Steamflow

Expression (2.1) gives the explicit form of Q_{B_j} as a function of streamflow values Q_1 through Q_j. If the streamflows carry some uncertainty, their variabilities can be used in computing the total variance of Q_{B_j}. However, if they are deterministic, then only the randomness in α is used for the computation of the total variance.

Considering Q_j as a fixed non-random constant and since β is dependent upon α, then Q_{Bj} becomes a function of α. Thus, the only variables in the explicit function (2.1) that produce variance are τ_1 and τ_2. Consider Q_{B2} for example. The variance of Q_{B2} is,

$$\begin{aligned}\mathrm{Var}(Q_{B2}) &= \mathrm{Var}(\tau_1^2 Q_1 + \tau_1\tau_2 Q_1 + \tau_2 Q_2) \\ &= Q_1^2 \mathrm{Var}(\tau_1^2) + Q_1^2 \mathrm{Var}(\tau_1\tau_2) + Q_2^2 \mathrm{Var}(\tau_2) + 2Q_1^2 \mathrm{Cov}(\tau_1^2, \tau_1\tau_2) \\ &+ 2Q_1 Q_2 \mathrm{Cov}(\tau_1^2, \tau_2) + 2Q_1 Q_2 \mathrm{Cov}(\tau_1\tau_2, \tau_2).\end{aligned}$$

The variances are based upon each individual term, while the covariances between the individual terms are based on the correlation between the two terms. And this process can be carried out to obtain an expression of the variance of Q_{B_j} for any given j.

Now, let's consider the scenario where the streamflow Q_js are random, and carry some variability. This can be the case because the streamflows were recorded from measuring gauges at measuring sites. Gauge measurements of streamflows carry measurement errors that can be intrinsic the the measurement devices, and also from other natural phenomenons, creating some variability for Q_B. The recursive term, Q_{Bj-1}, has some variance that will accumulate over time. This contributes to the variance in estimating baseflow. In equation (2.1), Q_{B0} and $Q_1, Q_2, \cdots, Q_j$ all contribute to the variance in Q_{Bj}. Let $Q_{B0} \neq Q_1$, and suppose

$$
\begin{aligned}
Q_j &= \mu_j + \epsilon_j \\
Q_{B0} &= \mu_0 + e_0,
\end{aligned}
$$

where μ_j and μ_0 are the true streamflow and initial baseflow respectively. Let $\epsilon_j \sim D(0, \sigma_j^2)$ and $e_0 \sim D(0, \sigma_0^2)$ represent the uncertainty around streamflow and initial baseflow. By substitution, and after collecting like terms, we get

$$
\begin{aligned}
Q_{Bj} =&(\tau_1^j \mu_0 + \tau_1^{j-1}\tau_2\mu_1 + \cdots + \tau_1\tau_2\mu_{j-1} + \tau_2\mu_j) + \\
&(\tau_1^j e_0 + \tau_1^{j-1}\tau_2\epsilon_1 + \cdots + \tau_1\tau_2\epsilon_{j-1} + \tau_2\epsilon_j).
\end{aligned} \tag{2.3}
$$

Using vector notation to rewrite equation (2.3), we obtain:

$$
Q_{Bj} = \gamma^T X + \varepsilon^T X \tag{2.4}
$$

where $X = \left[\tau_1^j, \tau_1^{j-1}\tau_2, \cdots, \tau_1\tau_2, \tau_2\right]^T$, $\gamma = [\mu_0, \mu_1, \cdots, \mu_{j-1}, \mu_j]^T$ and $\varepsilon = [e_0, \epsilon_1, \cdots, \epsilon_{j-1}, \epsilon_j]^T$. Treating X as non-random and γ as a constant, the variance of the modified equation (2.4) is the variance of the error term $\varepsilon^T X$, $Var(Q_{Bj}|X) = Var(\varepsilon^T X)$, and consequently,

$$
\text{Var}(Q_{Bj}|X) = X^T \text{Cov}(\varepsilon^T) X \tag{2.5}
$$

where $\text{Cov}(\varepsilon^T)$ is the covariance matrix of ε^T. This is a matrix with size $(j+1) \times (j+1)$ composed of the variance of each term on the main diagonal and the covariance between each individual term on the off diagonal. To find the total estimated variance of Q_Bj, we will use the following formula:

$$
\text{Var}(Q_{Bj}) = \text{E}\left[\text{Var}(Q_{Bj}|X)\right] + \text{Var}\left[\text{E}(Q_{Bj}|X)\right]. \tag{2.6}
$$

Substituting equation (2.5) into equation (2.6) gives,

$$
\begin{aligned}
\text{Var}(Q_{Bj}) &= \text{E}\left[X^T \text{Cov}(\varepsilon^T) X\right] + \text{Var}\left[\text{E}(Q_{Bj}|X)\right] \\
&= X^T \text{Cov}(\varepsilon^T) X + \text{Var}\left[\text{E}(Q_{Bj}|X)\right].
\end{aligned}
$$

Now we need to find the expected value of $Q_{Bj}|X$. Since ε is distributed with a mean of 0, the expected value of that term is just 0, which leaves the expected value of $(Q_{Bj}|X)$ to be $\gamma^T X$, where γ is a vector of constants and X is a vector of τ_1 and τ_2. Substituting this value gives

$$\begin{aligned}\text{Var}(Q_{Bj}) &= X^T\text{Cov}(\varepsilon^T)X + \text{Var}(\gamma^T X)\\ &= X^T\text{Cov}(\varepsilon^T)X + \gamma^T\text{Cov}(X)\gamma\end{aligned}$$

where $\text{Cov}(X)$ is the covariance matrix of X with dimension $(j+1)\times(j+1)$ composed of the variance of each term on the main diagonal and the covariance between each individual term on the off diagonal.

We note that this procedure can be carried out provided that σ^2 and σ_0^2 are available or estimable. We do not have enough information to assess these parameters and thus we provide these details as a part of our recommendations.

3 Numerical Methods

We have implemented the two methods in the statistical software R. These codes are available and accessible upon request. The following following sections explain the functions we wrote. We use this `font` to represent these R functions and objects.

3.1 Bootstrapping the recession constant

The Bootstrap function takes in four variables: a list of streamflow observations from one site (`Q`), a list of baseflow estimates from one site (`Q_0`), a sample of α values calculated from the streamflow (`dat`), and the number of `j` days to calculate the variance. Figure 3.1 displays the algorithm for estimating the variability of streamflow using the bootstrap method.

Bootstrapping the recession constant, α, involves re-sampling, with replacement, from the supplied alpha data to obtain a new sample of 572 α values. The median is obtained from the new sample, and then stored in a vector. This process is repeated K times to produce K different α values stored in a vector `alpha`. Then, a vector of K β values, `beta`, is calculated by calling the `beta` function and passing through `alpha, Q,` and `Q_0`.

Calling the `calc_Qb` function and passing through `Q, j, alpha`, and `beta` will calculate the respective baseflow. The `calc_Qb` function calculates baseflow using formula 2.1 and returns a vector of Q_B values for a single `j` at each `alpha` and `beta` pair. Then the variance is calculated using the `Q_B` vector and stored in a vector of variances. This process is then repeated for each `j`. The function returns the vector of variances.

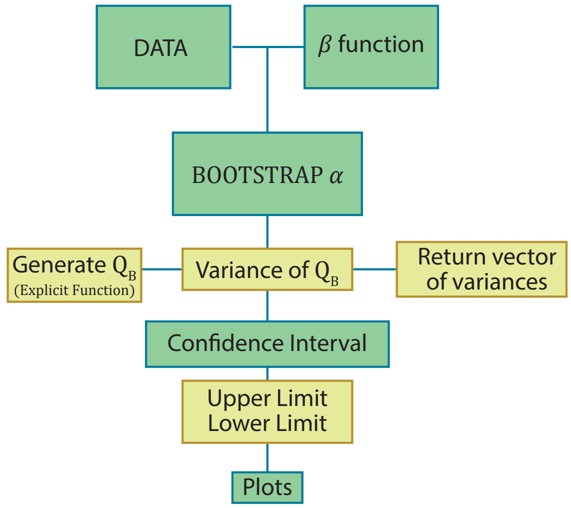

Figure 3.1: The bootstrap algorithm generates K samples of the recession constant to estimate the variability of the baseflow.

3.2 Delta Method

The function `final_var` for Delta method takes in the same four variables as the Bootstrap function: `Q_0, Q, dat`, and `j`. We first take the mean and standard deviation of `dat` and store it in the variables `alpha` and `std` respectively. We use the `numericDeriv` function in R to calculate the derivative of the `calc_Qb` function with respect to α, and then substitute in the singular alpha value along with the vector of `Q_0` and `Q`. This will return a vector of `Q_B (Qb_deriv)`, one `Q_B` for each `j`. Each value in `Qb_deriv` is then substituted into (2.2) for $g'(\alpha_0)$, and `std` is substituted in for σ, see Figure 3.2. This produces a vector of variances `(final_var)`, one variance for each `j`. Finally, the function returns the Delta method variances.

4 Results

4.1 Bootstrap Method vs. Delta Method

We computed the variance and determined a 95% confidence interval of the baseflow for each day. The plots for the first $K = 1000$ days shown in Figures 4.1 and 4.1 illustrate the variability obtained from each of the methods for the first 1000 days. From these plots, it can be seen that the margin of error for the Delta method is greater than the margin of error for the bootstrap method. The averages of the variances confirm this relationship

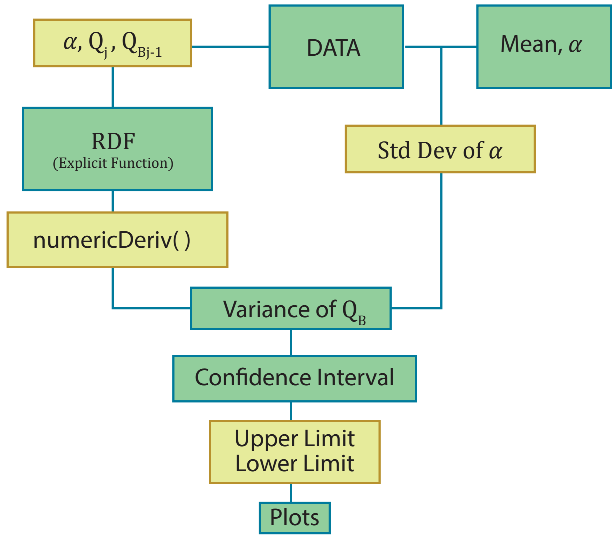

Figure 3.2: The Delta method relies on the derivative of the RDF and the standard deviation of the recessive constant to estimate the variability of the baseflow.

with an average of $447326.7(L^3/t)^2$ for the Delta method, which surpasses the average of the bootstrap method, which is $538.8(L^3/t)^2$.

We noted that on Figure 4.1 the confidence interval for the bootstrap method is indistinguishable from the observed values of baseflow. Since the variability is minute, we can draw a conclusion that *using the median of the α values instead of the mean has very little to no effect on the observed variability.* This method is more reliable for determining variability around baseflow.

For the Delta method, the plot exhibits larger variances, as observed in Figure 4.1. We concluded that this is due to the distribution of α not being asymptotically normal, and a discontinuity in the derivative of the function when $\alpha\beta \to 1$. Based on the calculated α and β values, we would expect the denominator to be closer to 0 than 1. The resulting variances are large, and ergo the method is not recommended.

4.2 Parallelized Bootstrap vs. Serial Bootstrap

The plot on Figure fig:Perf is representative of a performance study done when parallelizing bootstrap for one, four, and 16 processes. Also compared are the running times for the Delta method. The Delta method consistently run in less time than the Bootstrap method. When run for the first 1000 days (as shown in Table 4.2) the serial bootstrap code runs for 0.87 minutes and the parallelized form for 4 processes runs for 1.35 minutes. However, when the code is run for 10000 days, the serial bootstrap runs much slower than the parallel using

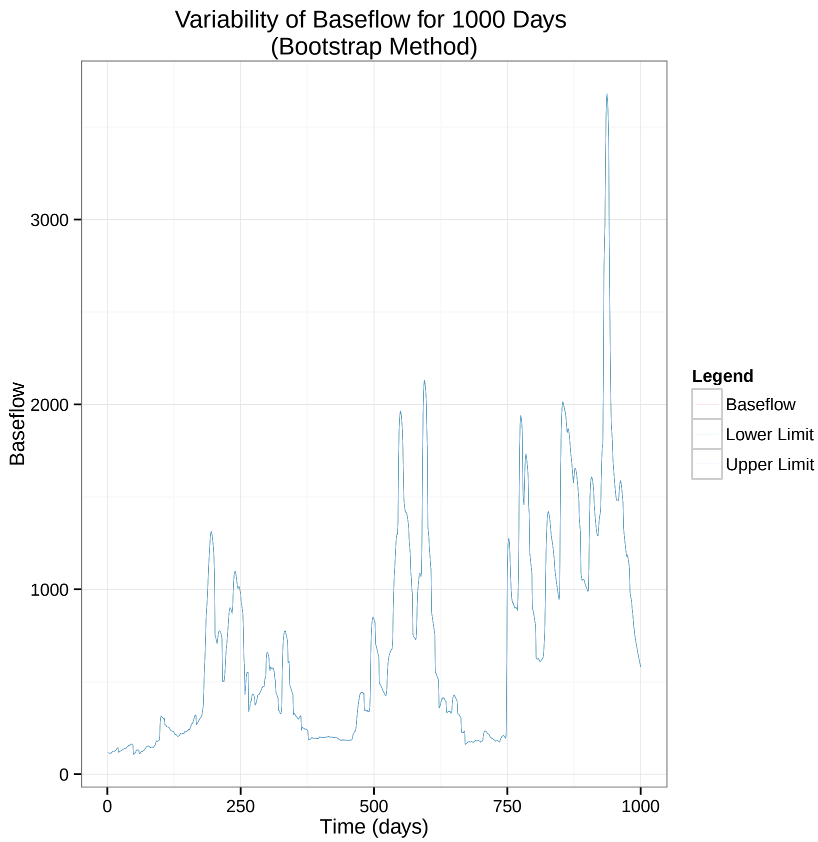

Figure 4.1: Plot of variance using bootstrap method for 1000 days. The three lines are not discernible. Using the median of alpha adds negligible variability for the baseflow using this method.

four processes by 54.44 minutes. The latter was the outcome that we had hoped to see from parallelization.

5 Conclusion

The U.S. Geological Survey's study of Raffensperger et al. (2017) provided the Recursive Digital Filter (RDF), an estimation method for baseflow. The RDF is an equation that depends on the previous day's baseflow estimate, streamflow measurement of the current day, a recession constant (α), and the maximum baseflow index (β). Assuming streamflow values to be constant, the variability of baseflow relies on the parameters α and β as well as the estimated baseflow of the previous day.

In order to quantify variability, we used the bootstrap method and the Delta method.

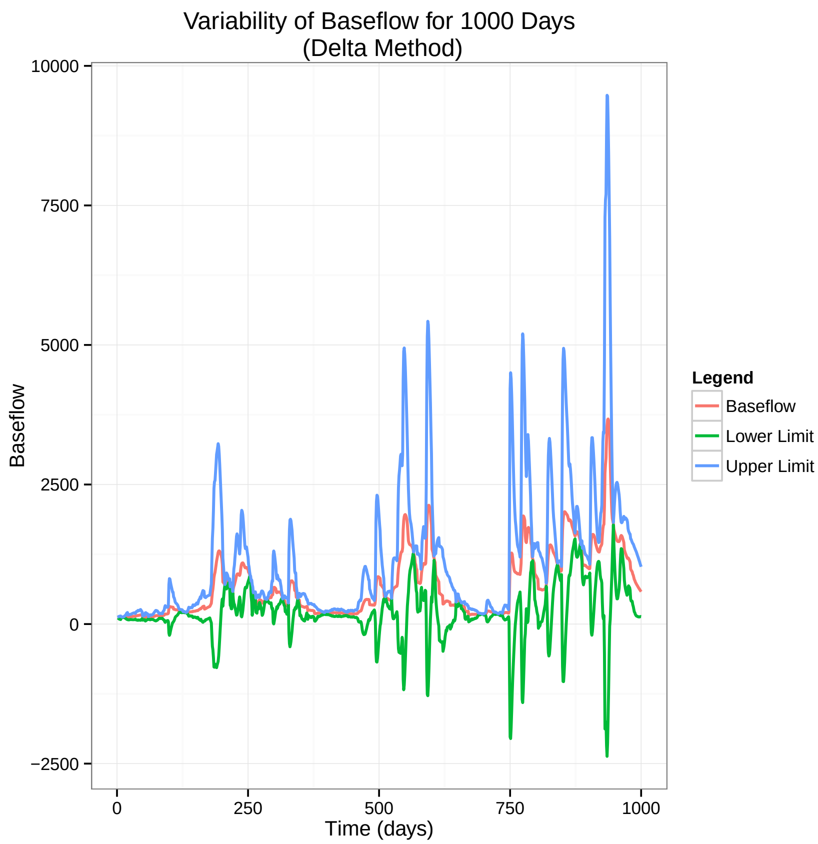

Figure 4.2: Plot of variance using Delta method for 1000 days.

The bootstrap method is a resampling method, which we used to iterate a large number of samples of α values for a given day. From the resulting estimates of baseflow, we calculated the average variance to be $538.75(L^3/t)^2$. Compared to the BFI measurements, this variance was concluded to be negligible. We have also considered the Delta method which relies on the derivative of the baseflow function along with the mean of the measured α values. This method produced an average variance of $447{,}326.7(L^3/t)^2$. With the current α and β values, the denominator of the baseflow estimator is closer to 0 than 1, which results in larger variances, thus the Delta method is not recommended.

A performance study was done for the serial and parallelized bootstrap method and the Delta method. Even when compared with the parallelized bootstrap method, the Delta method code runs significantly faster. While the study shows that the Delta method runs faster for the same set of data, we do not recommend this approach because of the possible discontinuity that occurs when the product of α and β approach 1. While the parallelized bootstrap method shows a negligible margin of error for the variability of baseflow, it would

Table 4.1: Perfomance Study of Statistical Algorithms (data times are in minutes)

j (Days)	Bootstrap (1 process)	Bootstrap (4 processes)	Bootstrap (16 processes)	Delta (1 process)
100	0.19	1.08	1.09	0.003
1000	0.87	1.35	1.15	0.007
10000	76.8	22.36	7.17	0.5
20000	343.2	92.4	26.85	2.31
30000	781.2	207.6	58.95	5.6
31046	838.2	223.2	63	6.37

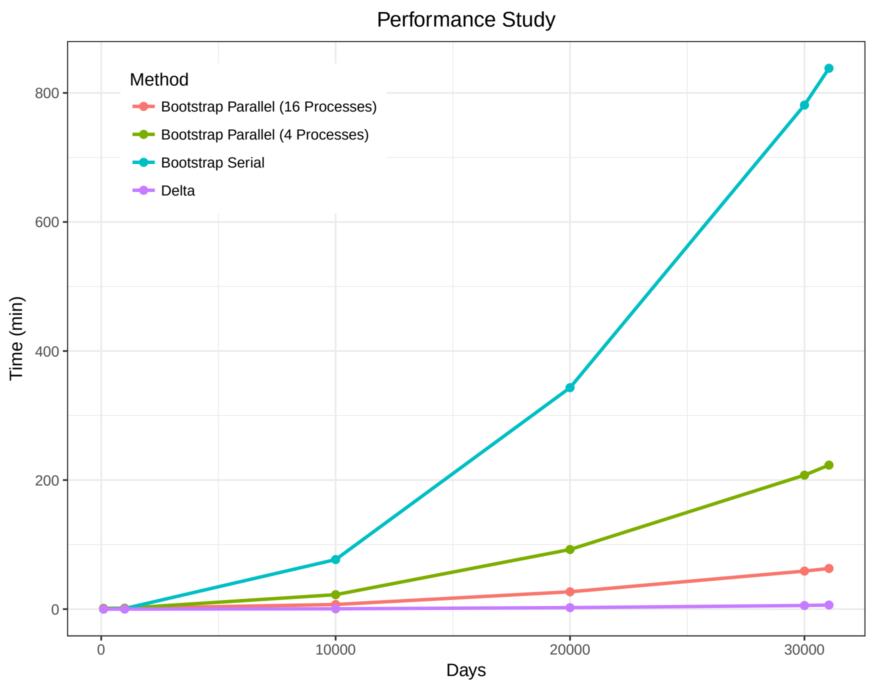

Figure 4.3: Performance study for Bootstrap method using 1, 4, and 16 processes compared to Delta method (unparallelized).

be worthwhile to implement the algorithm for multiple sites for a non-constant streamflow.

Acknowledgments

These results were obtained as part of the REU Site: Interdisciplinary Program in High Performance Computing (hpcreu.umbc.edu) in the Department of Mathematics and Statistics at the University of Maryland, Baltimore County (UMBC) in Summer 2017. This program is funded by the National Science Foundation (NSF), the National Security Agency (NSA), and the Department of Defense (DOD), with additional support from UMBC, the Department of Mathematics and Statistics, the Center for Interdisciplinary Research and Consulting (CIRC), and the UMBC High Performance Computing Facility (HPCF). HPCF is supported by the U.S. National Science Foundation through the MRI program (grant nos. CNS-0821258 and CNS-1228778) and the SCREMS program (grant no. DMS-0821311), with additional substantial support from UMBC. Author Christian Dixon was supported, in part, by the UMBC National Security Agency (NSA) Scholars Program through a contract with the NSA. Graduate assistants Nadeesri Wijekoon and Sai Popuri were supported by UMBC.

References

[1] Eckhardt, K., 2005, How to construct recursive digital filters for baseflow separation Hydrol. Process. 19, 507515 (2005) Published online 8 December 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/hyp.5675

[2] Raffensperger, J.P., Baker, A.C., Blomquist, J.D., and Hopple, J.A., 2017, Optimal hydrograph separation using a recursive digital filter constrained by chemical mass balance, with application to selected Chesapeake Bay watersheds: U.S. Geological Survey Scientific Investigations Report 2017 – 5034, 51 p., https://doi.org/10.3133/sir20175034.

[3] Rutledge, A.T., 1998, Computer programs for describing the recession of ground-water discharge and for estimating mean ground-water recharge and discharge from streamflow recordsupdate: U.S. Geological Survey Water-Resources Investigations Report 984148, 43 p., accessed March 16, 2017, at https://pubs.usgs.gov/wri/wri984148/.

[4] Sloto, R.A., and Crouse, M.Y., 1996, HYSEP: a computer program for streamflow hydrograph separation and analysis: U.S. Geological Survey Water-Resources Investigations Report 964040, 46 p., accessed March 16, 2017, at https://pubs.usgs.gov/wri/1996/4040/report.pdf.

[5] Efron, B. and Tibshirani, R.J., 1994. An introduction to the bootstrap. CRC press.