Technical note: Analytical sensitivity analysis and uncertainty estimation of baseflow index calculated by a two-component hydrograph separation method with conductivity as a tracer

Weifei Yang[1], Changlai Xiao[1], Xiujuan Liang[1]

[1]Key Laboratory of Groundwater Resources and Environment, Ministry of Education, and National-Local Joint Engineering Laboratory of In-situ Conversion, Drilling and Exploitation Technology for Oil Shale, and College of New Energy and Environment, Jilin University, No 2519, Jiefang Road, Changchun 130021, PR China

Correspondence to: Changlai Xiao (xcl2822@126.com)

Abstract. The two-component hydrograph separation method with conductivity as a tracer is favored by hydrologists owing to its low cost and easy application. This study analyzes the sensitivity of the baseflow index (BFI, long-term ratio of baseflow to streamflow) calculated using this method to errors or uncertainties of two parameters (BF_C, the conductivity of baseflow, and RO_C, the conductivity of surface runoff) and two variables (y_k, streamflow, and SC_k, specific conductance of streamflow, where k is the time step), and then estimates the uncertainty of BFI. The analysis shows that for time series longer than 365 days, random measurement errors of y_k or SC_k will cancel each other, and their influence on BFI can be neglected. An uncertainty estimation method of BFI is derived on the basis of the sensitivity analysis. Representative sensitivity indices (the ratio of the relative error of BFI to that of BF_C or RO_C) and BFI' uncertainties are determined by applying the resulting equations to 24 watersheds in the United States. These dimensionless sensitivity indices can well express the propagation of errors or uncertainties of BF_C or RO_C into BFI. The results indicate that BFI is more sensitive to BF_C, and the conductivity two-component hydrograph separation method may be more suitable for the long time series in a small watershed. When the mutual offset of the measurement errors of conductivity and streamflow is considered, the uncertainty of BFI is reduced by half.

1 Introduction

Hydrograph separation (also called baseflow separation), aims to identify the proportion of water in different runoff pathways in the export flow of a basin, which helps in identifying the conversion relationship between groundwater and surface water; in addition, it is a necessary condition for optimal allocation of water resources (Cartwright et al., 2014; Miller et al., 2014; Costelloe et al., 2015). Some researchers indicated that tracer-based hydrograph separation methods yield the most realistic results because they are the most physically based methods (Miller et al., 2014; Mei and Anagnostou, 2015; Zhang et al., 2017). Many hydrologists have suggested that electrical conductivity can be used as a tracer in hydrograph separation (Stewart et al., 2007; Munyaneza et al., 2012; Cartwright et al., 2014; Lott and Stewart, 2016; Okello et al., 2018). Conductivity is a suitable tracer because its measurement is simple and inexpensive, and it has distinct applicability in long-series hydrograph separation (Okello et al., 2018).

The two-component hydrograph separation method with conductivity as a tracer (also called conductivity mass balance (CMB) method (Stewart et al. 2007)) calculates baseflow through a two-component mass balance equation. The general equation is shown in Eq. (1), which is based on the following assumptions:

a) Contributions from end-members other than baseflow and surface runoff are negligible.

b) The specific conductance of runoff and baseflow are constant (or vary in a known manner) over the period of record.

c) Instream processes (such as evaporation) do not change specific conductance makedly.

d) Baseflow and surface runoff have significantly different specific conductance.

$$b_k = \frac{y_k(SC_k - \mathrm{RO_C})}{\mathrm{BF_C} - \mathrm{RO_C}} \quad (1)$$

where b is baseflow (L^3/T), y is streamflow (L^3/T), SC is the electrical conductivity of streamflow, and k is time step number. The two parameters BF_C and RO_C represent the electrical conductivity of baseflow and surface runoff, respectively.

Stewart et al. (2007) conducted a field test in a drainage basin of $12km^2$ area in southeast Hillsborough County, Florida and showed that the maximum conductivity of streamflow can be used to replace BF_C, and the minimum conductivity can be used to replace RO_C. However, Miller et al. (2014) pointed out that the maximum conductivity of streamflow may exceed the real BF_C; Therefore, they suggested that the 99th percentile of the conductivity of each year should be used as BF_C to avoid the impact of high BF_C estimates on the separation results and assumed that baseflow conductivity varies linearly between years. The determination of the parameters (BF_C, RO_C) of the conductivity two-component hydrograph separation method involves some uncertainties (Miller et al., 2014; Okello et al., 2018). Therefore, sensitivity analysis of parameters and quantitative analysis of the uncertainties will contribute towards further optimization of the CMB method and improving the accuracy of hydrograph separation.

Most existing parameter sensitivity analysis methods are empirical methods that usually substitute varying values of a certain parameter into the separation model and then compare the range of the separation results produced by these varying parameter values (Eckhardt, 2005; Miller et al., 2014; Okello et al., 2018). Eckhardt (2012) indicated that "An empirical sensitivity analysis is only a makeshift if an analytical sensitivity analysis, that is an analytical calculation of the error propagation through the model, is not feasible". Eckhardt (2012) derived sensitivity indices of equation parameters by the partial derivative of a two-parameter recursive digital baseflow separation filter equation. Until now, the parameters' sensitivity indices of the CMB equation have not been derived.

At present, the uncertainty of the separation results of the CMB method is mainly estimated using an uncertainty transfer equation based on the uncertainty of BF_C, RO_C, and SC_k (Genereux, 1998; Miller et al., 2014). See Sect. 3.1 for details. In this uncertainty estimation method, the uncertainty of the baseflow ratio (f_{bf}, the ratio of baseflow to streamflow in a single calculation process) is estimated, and the average uncertainty of multiple calculation processes is then used to estimate the uncertainty of the baseflow index (BFI, long-term ratio of baseflow to total streamflow). This method can neither directly estimate the uncertainty of BFI nor consider the randomness and mutual offset of conductivity measurement errors, and thus, it does not provide accurate estimates of BFI uncertainty.

The main objectives of this study are as follows: (i) analyze the sensitivity of long-term series of baseflow separation results (BFI) to parameters and variables of the CMB equation (Sect. 2); (ii) derive the uncertainty of BFI (Sect.3). The derived solutions were applied to 24 basins in the United States, and the parameter sensitivity indices and BFI uncertainty characteristics were analyzed (Sect. 4).

2 Ssensitivity analysis

2.1 Parameters BF_C and RO_C

In order to calculate the sensitivity indices of the parameters, the partial derivatives of b_k in Eq. (1) with respect to BF_C and RO_C are required (the derivation process is expressed as Eq. (A1) and (A2)):

$$\frac{\partial b_k}{\partial \mathrm{BF_C}} = -y_k \frac{SC_k - \mathrm{RO_C}}{(\mathrm{BF_C} - \mathrm{RO_C})^2} \quad (2)$$

$$\frac{\partial b_k}{\partial \mathrm{RO_C}} = y_k \frac{SC_k - \mathrm{BF_C}}{(\mathrm{BF_C} - \mathrm{RO_C})^2} \quad (3)$$

For the convenience of comparison, the baseflow index (BFI) is selected as the baseflow separation result for long time series to analyze the influence of parameter uncertainty on BFI,

$$\mathrm{BFI} = \frac{\sum_{k=1}^{n} b_k}{\sum_{k=1}^{n} y_k} = \frac{b}{y} \tag{4}$$

where b and y denote the total baseflow and total streamflow, respectively, over the whole available streamflow sequences, and n is the number of available streamflow data.

Then, the partial derivatives of BFI to $\mathrm{BF_C}$ and $\mathrm{RO_C}$ should be calculated (the derivation process is presented in Eq. (A3) and (A4)):

$$\frac{\partial \mathrm{BFI}}{\partial \mathrm{BF_c}} = \frac{y\mathrm{RO_c} - \sum_{k=1}^{n} y_k SC_k}{y(\mathrm{BF_c} - \mathrm{RO_c})^2} \tag{5}$$

$$\frac{\partial \mathrm{BFI}}{\partial \mathrm{RO_c}} = \frac{\sum_{k=1}^{n} y_k SC_k - y\mathrm{BF_c}}{y(\mathrm{BF_c} - \mathrm{RO_c})^2} \tag{6}$$

The definition of the partial derivative suggests that the influence of the errors of the parameters ($\Delta\mathrm{BF_C}$ and $\Delta\mathrm{RO_C}$) in Eq. (1) on the BFI can be expressed by the product of the errors and its partial derivatives. Then the errors of BFI caused by small errors of $\mathrm{BF_C}$ and $\mathrm{RO_C}$ can be approximated by:

$$\Delta_{\mathrm{BF_c}}\mathrm{BFI} = \frac{\partial \mathrm{BFI}}{\partial \mathrm{BF_c}} \Delta \mathrm{BF_c} = \frac{y\mathrm{RO_c} - \sum_{k=1}^{n} y_k SC_k}{y(\mathrm{BF_c} - \mathrm{RO_c})^2} \Delta \mathrm{BF_c} \tag{7}$$

$$\Delta_{\mathrm{RO_c}}\mathrm{BFI} = \frac{\partial \mathrm{BFI}}{\partial \mathrm{RO_c}} \Delta \mathrm{RO_c} = \frac{\sum_{k=1}^{n} y_k SC_k - y\mathrm{BF_c}}{y(\mathrm{BF_c} - \mathrm{RO_c})^2} \Delta \mathrm{RO_c} \tag{8}$$

The dimensionless sensitivity indices (S) can be obtained by comparing the relative error of BFI caused by the small errors of $\mathrm{BF_C}$ and $\mathrm{RO_C}$ with that of $\mathrm{BF_C}$ and $\mathrm{RO_C}$, (see Eq. (B1), (B2)):

$$\mathrm{S(BFI|BF_c)} = \frac{\Delta_{\mathrm{BF_c}}\mathrm{BFI}}{\mathrm{BFI}} \Big/ \frac{\Delta \mathrm{BF_c}}{\mathrm{BF_c}} = \frac{\mathrm{BF_c}(y\mathrm{RO_c} - \sum_{k=1}^{n} y_k SC_k)}{y\mathrm{BFI}(\mathrm{BF_c} - \mathrm{RO_c})^2} \tag{9}$$

$$\mathrm{S(BFI|RO_c)} = \frac{\Delta_{\mathrm{RO_c}}\mathrm{BFI}}{\mathrm{BFI}} \Big/ \frac{\Delta \mathrm{RO_c}}{\mathrm{RO_c}} = \frac{\mathrm{RO_c}(\sum_{k=1}^{n} y_k SC_k - y\mathrm{BF_c})}{y\mathrm{BFI}(\mathrm{BF_c} - \mathrm{RO_c})^2} \tag{10}$$

where $\mathrm{S(BFI|BF_c)}$ represent the dimensionless sensitivity index of BFI (output) with $\mathrm{BF_c}$ (uncertain input), and $\mathrm{S(BFI|RO_c)}$ with $\mathrm{RO_c}$.

The dimensionless sensitivity index is also called the "elasticity index", and it reflects the proportional relationship between the relative error of BFI and the relative error of parameters (e.g. if $\mathrm{S(BFI|BF_c)} = 1.5$, and the relative error of $\mathrm{BF_c}$ is 5%, then the relative error of BFI is 1.5 times 5% = 7.5%). After determining the values of $\mathrm{BF_C}$, $\mathrm{RO_C}$, BFI, y, y_k and SC_k, the sensitivity indices $\mathrm{S(BFI|BF_c)}$ and $\mathrm{S(BFI|RO_c)}$ can be calculated and compared.

2.2 Variables y_k and SC_k

In addition to the two parameters, there are two variables (SC_k and y_k) in Eq. (1). This section describes the sensitivity analysis of BFI to these two variables. Similar to Sect. 2.1, the partial derivatives of b_k in Eq. (1) to SC_k and y_k are obtained (see Eq. (A5), (A6)), and the partial derivatives of BFI to SC_k and y_k are further obtained (see Eq. (A7), (A8)),

$$\frac{\partial \mathrm{BFI}}{\partial SC_k} = \frac{1}{\mathrm{BF_c} - \mathrm{RO_c}} \tag{11}$$

$$\frac{\partial \mathrm{BFI}}{\partial y_k} = \frac{\sum_{k=1}^{n}(SC_k - \mathrm{RO_c}) - \mathrm{nBFI}(\mathrm{BF_c} - \mathrm{RO_c})}{\mathrm{y}(\mathrm{BF_c} - \mathrm{RO_c})} \tag{12}$$

According to previous studies (Munyaneza et al., 2012; Cartwright et al., 2014; Miller et al., 2014; Okello et al., 2018) and this study (Table 1), the difference between $\mathrm{BF_C}$ and $\mathrm{RO_C}$ is often greater than 100 μs/cm. Therefore, $\partial \mathrm{BFI}/\partial SC_k$ is usually less than 0.01 cm/ μs. Appendix C shows that the value of $\partial \mathrm{BFI}/\partial y_k$ is usually far less than 1 d/m^3.

Small errors in SC_k and y_k cause errors in BFI

$$\Delta_{SC_k}\mathrm{BFI} = \frac{\partial \mathrm{BFI}}{\partial SC_k}\Delta SC_k = \frac{\Delta SC_k}{\mathrm{BF_c - RO_c}} \quad (13)$$

$$\Delta_{y_k}\mathrm{BFI} = \frac{\partial \mathrm{BFI}}{\partial y_k}\Delta y_k = \frac{\sum_{k=1}^{n}(SC_k - \mathrm{RO_c}) - \mathrm{nBFI}(\mathrm{BF_c - RO_c})}{\mathrm{y}(\mathrm{BF_c - RO_c})}\Delta y_k \quad (14)$$

The errors of BFI caused by SC_k and y_k are summed up to obtain the error of BFI caused by $\sum_{k=1}^{n} SC_k$ and $\sum_{k=1}^{n} y_k$ in the whole time series:

$$\Delta_{\sum_{k=1}^{n} SC_k}\mathrm{BFI} = \sum_{k=1}^{n}\Delta_{SC_k}\mathrm{BFI} = \sum_{k=1}^{n}\frac{\Delta SC_k}{\mathrm{BF_c - RO_c}} = \frac{1}{\mathrm{BF_c - RO_c}}\sum_{k=1}^{n}\Delta SC_k \quad (15)$$

$$\Delta_{\sum_{k=1}^{n} y_k}\mathrm{BFI} = \sum_{k=1}^{n}\Delta_{y_k}\mathrm{BFI} = \sum_{k=1}^{n}\left(\frac{\sum_{k=1}^{n}(SC_k - \mathrm{RO_c}) - \mathrm{nBFI}(\mathrm{BF_c - RO_c})}{\mathrm{y}(\mathrm{BF_c - RO_c})}\Delta y_k\right) = \frac{\sum_{k=1}^{n}(SC_k - \mathrm{RO_c}) - \mathrm{nBFI}(\mathrm{BF_c - RO_c})}{\mathrm{y}(\mathrm{BF_c - RO_c})}\sum_{k=1}^{n}\Delta y_k \quad (16)$$

Wagner et al. (2006) reported that the uncertainty of instruments is usually less than 5% for SC_k (<100 μs/cm) and less than 3% for SC_k (>100 μs/cm). According to Hamilton et al. (2012) streamflow data from USGS are often assumed by analysts to be accurate and precise within ±5% at the 95% confidence interval. In this study, the error ranges of SC_k and y_k are all considered to be ±5%. The errors in $\mathrm{SC_k}$ and $\mathrm{y_k}$ mainly comprise random measurement errors which mostly follow a normal distribution or a uniform distribution (Huang and Chen, 2011). Considering the mutual offset of random errors, when the time series (n) is sufficiently long, $\sum_{k=1}^{n}\Delta SC_{ck}$ in Eq. (15) and $\sum_{k=1}^{n}\Delta y_k$ in Eq. (16) will approach zero.

The analysis of $\sum_{k=1}^{n}\Delta SC_k$ and $\sum_{k=1}^{n}\Delta y_k$ under different time series (n) and different error distributions (normal distribution or uniform distribution) of a surface water station (USGS site number 0297100) showed that the random errors of daily average conductivity and streamflow have a negligible effect on BFI when the time series is greater than 365 days (See Supplement S1 for detail).

3 Uncertainty estimation

3.1 Previous attempts

According to previous studies, in the case where a variable g is calculated as a function of several factors x_1, x_2, x_3, …, x_n (e.g. g= G(x_1, x_2, x_3, …, x_n)) and based on the assumptions that the factors are uncorrelated and have a Gaussian distribution, the transfer equation (also known as Gaussian error propagation) between the uncertainty of the independent factors and the uncertainty of g is:

$$W_g = \sqrt{\left(\frac{\partial g}{\partial x_1}W_{x_1}\right)^2 + \left(\frac{\partial g}{\partial x_2}W_{x_2}\right)^2 + \cdots + \left(\frac{\partial g}{\partial x_n}W_{x_n}\right)^2} \quad (17)$$

where W_g, W_{x1}, W_{x2}, and W_{xn} are the same type of uncertainty values (e.g. all average errors or all standard deviations) for g, x_1, x_2, and x_n, respectively. A more detailed description of this equation can be found in Taylor (1982), Kline (1985), and Ernest (2005).

According to Genereux (1998), "While any set of consistent uncertainty (W) values may be propagated using Gaussian error propagation, using standard deviations multiplied by t values from the Student's t distribution (each t for the same confidence level, such as 95%) has the advantage of providing a clear meaning (tied to a confidence interval) for the computed uncertainty would correspond to, for example, 95% confidence limits on BFI".

Based on the above principle, Genereux (1998) substituted Eq. (18) into Eq. (17) to derive the uncertainty estimation equation (Eq. (19)) of the CMB method:

$$f_{bf} = \frac{SC_k - \mathrm{RO_c}}{\mathrm{BF_c - RO_c}} \quad (18)$$

$$W_{f_{bf}} = \sqrt{\left(\frac{f_{bf}}{\mathrm{BF_c - RO_c}}W_{\mathrm{BF_C}}\right)^2 + \left(\frac{1 - f_{bf}}{\mathrm{BF_c - RO_c}}W_{\mathrm{RO_c}}\right)^2 + \left(\frac{1}{\mathrm{BF_c - RO_c}}W_{SC}\right)^2} \quad (19)$$

where f_{bf} is the ratio of baseflow to streamflow in a single calculation process, W_{fbf} is the uncertainty in f_{bf} at the 95% confidence interval, W_{BFC} and W_{ROC} are the standard deviation of the $\mathrm{BF_C}$ and $\mathrm{RO_C}$ multiplied by the t-value (α=0.05; two-tail) from the Student's distribution, and W_{SC} is the analytical error in conductivity multiplied by the t-value (α=0.05; two-tail) (Miller et al., 2014).

Better estimates of the uncertainty of f_{bf} within a single calculation step can be obtained using Eq. (19). Hydrologists usually approximate the uncertainty of BFI by averaging the uncertainty of all steps (Genereux, 1998; Miller et al., 2014). However, this method does not consider the mutual offset of the conductivity measurement errors and cannot accurately reflect the uncertainty of BFI. In this study, an uncertainty estimation equation of BFI is derived on the basis of the parameter sensitivity analysis.

3.2 Uncertainty estimation of BFI

BFI is a function of $\mathrm{BF_c}$, $\mathrm{RO_c}$, SC_k and y_k. In addition, the uncertainties of $\mathrm{BF_c}$, $\mathrm{RO_c}$, SC_k and y_k are independent of each other. As explained earlier (Sect. 2.2), the random errors of daily average conductivity and streamflow have a negligible effect on BFI when the time series (n) is greater than 365 days (1 year), Therefore, the uncertainty of BFI can be expressed as:

$$W_{BFI} = \sqrt{(\frac{\partial \mathrm{BFI}}{\partial \mathrm{BF_c}} W_{\mathrm{BF_C}})^2 + (\frac{\partial \mathrm{BFI}}{\partial \mathrm{RO_c}} W_{\mathrm{RO_C}})^2} \quad (20)$$

$$\frac{\partial \mathrm{BFI}}{\partial \mathrm{BF_c}} = \mathrm{S(BFI|BF_c)} \frac{\mathrm{BFI}}{\mathrm{BF_c}} \quad (21)$$

$$\frac{\partial \mathrm{BFI}}{\partial \mathrm{RO_c}} = \mathrm{S(BFI|RO_c)} \frac{\mathrm{BFI}}{\mathrm{RO_c}} \quad (22)$$

Then, Eq. (20) can be rewritten as:

$$W_{BFI} = \sqrt{(\mathrm{S(BFI|BF_c)} \frac{\mathrm{BFI}}{\mathrm{BF_c}} W_{\mathrm{BF_C}})^2 + (\mathrm{S(BFI|RO_c)} \frac{\mathrm{BFI}}{\mathrm{RO_c}} W_{\mathrm{RO_C}})^2} \quad (23)$$

where W_{BFI}, $W_{\mathrm{BFC,}}$ and W_{ROC} are the same type of uncertainty values for BFI, $\mathrm{BF_{C,}}$ and $\mathrm{RO_C}$, respectively, as described above.

4 Application

4.1 Data and processing

The above sensitivity analysis and uncertainty estimation methods were applied to 24 catchments in the United States (Table 1). All basins used in this study are perennial streams, with drainage areas ranging from 10 $\mathrm{km^2}$ to 1258481 $\mathrm{km^2}$. Each gage has about at least 1 year of continuous streamflow and conductivity for the same period of records. All streamflow and conductivity data are daily average values retrieved from the United States Geological Survey's (USGS) National Water Information System (NWIS) website, http://waterdata.usgs.gov/nwis.

The daily baseflow of each basin was calculated using Eq. (1). The 99th percentile of the conductivity of each year was used as $\mathrm{BF_C}$, and linear variation of baseflow conductivity between years was assumed. The 1st percentile of the conductivity of the whole series of streamflow in each basin was used as the $\mathrm{RO_C}$. The total baseflow *b*, total streamflow *y* and baseflow index BFI of each watershed were then calculated. According to the results of the hydrograph separation, the parameter sensitivity indices of BFI for mean $\mathrm{BF_C}$ ($\mathrm{S(BFI|BF_c)}$) and $\mathrm{RO_C}$ ($\mathrm{S(BFI|RO_c)}$) were calculated using Eq. (9) and Eq. (10), respectively.

Finally, the uncertainty of f_{bf} in each step was calculated using Eq. (19) and averaged to obtain the mean W_{fbf} for each basin. The uncertainty (W_{BFI}) of BFI was directly calculated using Eq. (23), and then the values of mean W_{fbf} and W_{BFI} were compared. For each basin, W_{BFC} is the standard deviation of the $\mathrm{BF_C}$ of the whole series multiplied by the t-value (α=0.05; two-tail) from the Student's distribution, W_{ROC} is the standard deviation of the lowest 1% of measured conductivity multiplied by the t-value

(α=0.05; two-tail) from the Student's distribution, and W_{SC} is the analytical error in the conductivity (5%) multiplied by the t-value (α=0.05; two-tail).

4.2 Results and discussion

The calculation results are shown in Table 1. The average baseflow index of the 24 watersheds is 0.34, the average sensitivity index of BFI for mean BF_C ($S(BFI|BF_c)$) is -1.40, and the average sensitivity index of BFI for RO_C ($S(BFI|RO_c)$) is -0.89. The negative sensitivity indices indicate a negative correlation between BFI and BF_C, RO_C. The absolute value of the sensitivity index for BF_C is generally greater than that for RO_C, indicating that BFI is more affected by BF_C (for example, if there are 10% uncertainties in both BF_C and RO_C, then BF_C leads to -1.40 times 10% of uncertainty in BFI (-14.0%), while RO_C leads to -0.89 times 10% (-8.9%)). Therefore, the determination of BF_C requires more caution, and any small error may lead to greater uncertainty in BFI. Miller et al. (2014) reported that anthropogenic activities over long periods of time, or year to year changes in the elevation of the water table may result in temporal changes in BF_C. They recommended taking different BF_C values per year based on the conductivity values during low flow periods to avoid the effects of temporal fluctuations in BF_C.

Table 1. Basic information, parameter sensitivity analysis, and uncertainty estimation results for 24 basins in the United States. Footnote "a" in the "Area" column indicates that the values are estimated based on data from adjacent sites.

The sensitivity index of BFI for BF_C shows a decreasing trend with the increase of time series (n) (Fig. 1(a)) and an increasing trend with increasing watershed area (Fig. 1(b)), with correlation coefficients of 0.1492 and 0.3577, respectively. Although the correlations are not obvious, they have important guiding significance. Large basin, comprise many different subsurface flow paths contributing to streams (Okello et al., 2018), each of which has a unique conductivity value (Miller et al., 2014). Furthermore, it is difficult to represent the conductivity characteristics of subsurface flow with a special value. Therefore, the CMB method has higher applicability to long time series of small watershed.

The sensitivity index of BFI for RO_C did not change significantly with the increase of time series and watershed area (Fig. 1(c), Fig. 1(d)). During rainstorms, the conductivity of streams became similar to that of the rainfall (Stewart et al., 2007). The electrical conductivity of regional rainfalls varies slightly, usually at a fixed value, and it has no significant relationship with the basin area and year (Munyaneza et al., 2012). Therefore, the temporal and spatial variation characteristics of BFI for RO_C are not obvious.

Figure 1. Scatter plots of sensitivity indices vs. time series (n) and drainage area of the 24 US basins. The watershed area uses a logarithmic axis, while the others are linear axes.

Genereux's method (Eq.19) estimates the average uncertainty of BFI in the 24 basins (average of mean W_{fbf}) to be 0.20, whereas the average uncertainty of BFI (average of W_{BFI}) calculated directly using the proposed method (Eq. 23) is 0.11 (Table 1). Mean W_{fbf} in each basin is generally larger than W_{BFI} (W_{BFI} is about 0.51 times of mean W_{fbf}), and there is a significant linear correlation (Fig. 2). This shows that the two methods have the same volatility characteristics for BFI uncertainty estimation, but Genereux's method (Eq. 19) often overestimates the uncertainty of BFI. This also means that when the time series is longer than 365 days, the measurement errors of conductivity and streamflow will cancel each other and thus reduce the uncertainty of BFI (about half of the original).

Figure 2. Scatter plot of uncertainty in BFI (W_{BFI}) and mean uncertainty in f_{bf} (mean W_{fbf}).

The conductivity of shallow subsurface and soil flow in real watersheds is sensitive to climatic conditions and usually shows obvious fluctuations (Miller et al., 2014). The CMB method classifies high conductivity flow (e.g., deep subsurface flow) as baseflow and low conductivity flow (e.g., local shallow soil flow) as surface runoff (Cartwright et al., 2014). Therefore, in the watershed containing a large number of low-conductivity soil flows, the BFI calculated by the CMB method comprised only the baseflow index of the deep subsurface flow. The parameter sensitivity indices and uncertainty of the deep subsurface flow were also calculated by the methods of this paper. Cartwright et al. (2014) showed that the ratio of low-conductivity soil flow to high-conductivity subsurface flow in the Barwon basin in southeastern Australia is close to 1. If only the BFI doubles and other parameters remain unchanged, then the sensitivity indices calculated by Eq. (9) and Eq. (10) are halved, whereas the uncertainty calculated by Eq. (23) remains unchanged. Therefore, non-constant soil flow conductivity may lead to an overestimation of sensitivity, but it has less impact on uncertainty estimates.

5 Conclusions

This study analyzed the sensitivity of BFI, calculated using the CMB method, to errors or uncertainties of parameters BF_C and RO_C and variables y_k and SC_k. In addition, the uncertainty of BFI was calculated. The equations derived in this study (Eq. (9) and Eq. (10)) could calculate the sensitivity indices of BFI for BF_C and RO_C. For time series longer than 365 days, the measurement errors of conductivity and streamflow exhibited a mutual offset effect, and their influence on BFI could be neglected. Considering the mutual offset, the uncertainty of BFI would be halved. From this perspective, Eq. (23) could estimate the uncertainty of BFI for time series longer than 365 days. The application of the method to 24 basins in the United States showed that BFI is more sensitive to BF_C. Future studies should dedicate more effort into determining the value of BF_C. In addition, the CMB method may be more suitable for long time series of small watersheds.

Systematic errors in specific conductance and streamflow as well as temporal and spatial variations in baseflow conductivity may be the main sources of BFI uncertainty. Better rating curves are probably more important than better loggers, and understanding the specific conductance of baseflow is likely more important than understanding that of surface runoff.

The above conclusions were drawn only from the average of the studied 24 basins, and further research in other countries or in more watersheds is thus required. This study focused on the two-component hydrograph separation method with conductivity as a tracer, but parameter sensitivity analysis and uncertainty analysis methods involving other tracers are similar. Therefore, similar equations can easily be derived by referring to the findings of this study.

Appendix A

Calculation of the partial derivatives

$$\frac{\partial b_k}{\partial \mathrm{BF_c}} = \frac{\partial}{\partial \mathrm{BF_c}} \frac{y_k(SC_k - \mathrm{RO_c})}{\mathrm{BF_c} - \mathrm{RO_c}} = y_k(SC_k - \mathrm{RO_c}) \frac{\partial}{\partial \mathrm{BF_c}} \frac{1}{\mathrm{BF_c} - \mathrm{RO_c}} = -y_k \frac{SC_k - \mathrm{RO_c}}{(\mathrm{BF_c} - \mathrm{RO_c})^2} \quad \text{(A1)}$$

$$\frac{\partial b_k}{\partial \mathrm{RO_c}} = \frac{\partial}{\partial \mathrm{RO_c}} \frac{y_k(SC_k - \mathrm{RO_c})}{\mathrm{BF_c} - \mathrm{RO_c}} = y_k \frac{\partial}{\partial \mathrm{RO_c}} \frac{SC_k - \mathrm{RO_c}}{\mathrm{BF_c} - \mathrm{RO_c}} = y_k \frac{-(\mathrm{BF_c} - \mathrm{RO_c}) + (SC_k - \mathrm{RO_c})}{(\mathrm{BF_c} - \mathrm{RO_c})^2} = y_k \frac{SC_k - \mathrm{BF_c}}{(\mathrm{BF_c} - \mathrm{RO_c})^2} \quad \text{(A2)}$$

$$\frac{\partial BFI}{\partial BF_c} = \frac{\partial}{\partial BF_c} \frac{b}{y} = \frac{1}{y} \sum_{k=1}^{n} \frac{\partial b_k}{\partial BF_c} = \frac{1}{y} \sum_{k=1}^{n} \left(-y_k \frac{SC_k - \mathrm{RO_c}}{(\mathrm{BF_c} - \mathrm{RO_c})^2}\right) (\text{see Eq. A1}) = \frac{1}{y(\mathrm{BF_c} - \mathrm{RO_c})^2} \sum_{k=1}^{n} (y_k \mathrm{RO_c} - y_k SC_k) = \frac{yRO_c - \sum_{k=1}^{n} y_k SC_k}{y(BF_c - RO_c)^2} \quad \text{(A3)}$$

$$\frac{\partial BFI}{\partial RO_c} = \frac{\partial}{\partial RO_c} \frac{b}{y} = \frac{1}{y} \sum_{k=1}^{n} \frac{\partial b_k}{\partial RO_c} = \frac{1}{y} \sum_{k=1}^{n} \left(y_k \frac{SC_k - \mathrm{BF_c}}{(\mathrm{BF_c} - \mathrm{RO_c})^2}\right) (\text{see Eq. A2}) = \frac{1}{y(\mathrm{BF_c} - \mathrm{RO_c})^2} \sum_{k=1}^{n} (y_k SC_k - y_k \mathrm{BF_c}) = \frac{\sum_{k=1}^{n} y_k SC_k - yBF_c}{y(BF_c - RO_c)^2} \quad \text{(A4)}$$

$$\frac{\partial b_k}{\partial SC_k} = \frac{\partial}{\partial \mathrm{Q_{ck}}} \frac{y_k(SC_k - \mathrm{RO_c})}{\mathrm{BF_c} - \mathrm{RO_c}} = \frac{1}{\mathrm{BF_c} - \mathrm{RO_c}} \frac{\partial}{\partial \mathrm{SC_k}} y_k(SC_k - \mathrm{RO_c}) = \frac{y_k}{\mathrm{BF_c} - \mathrm{RO_c}} \quad \text{(A5)}$$

$$\frac{\partial b_k}{\partial y_k} = \frac{\partial}{\partial y_k} \frac{y_k(SC_k - \mathrm{RO_c})}{\mathrm{BF_c} - \mathrm{RO_c}} = \frac{(SC_k - \mathrm{RO_c})}{\mathrm{BF_c} - \mathrm{RO_c}} \frac{\partial}{\partial y_k} y_k = \frac{SC_k - \mathrm{RO_c}}{\mathrm{BF_c} - \mathrm{RO_c}} \quad \text{(A6)}$$

$$\frac{\partial \mathrm{BFI}}{\partial SC_k} = \frac{\partial}{\partial SC_k} \frac{b}{y} = \frac{1}{y} \sum_{k=1}^{n} \frac{\partial b_k}{\partial SC_k} = \frac{1}{y} \sum_{k=1}^{n} \frac{y_k}{\mathrm{BF_c} - \mathrm{RO_c}} (\text{see Eq. A5}) = \frac{1}{y(\mathrm{BF_c} - \mathrm{RO_c})} \sum_{k=1}^{n} \mathrm{y}_k = \frac{1}{\mathrm{BF_c} - \mathrm{RO_c}} \quad \text{(A7)}$$

$$\frac{\partial \mathrm{BFI}}{\partial y_k} = \frac{\partial}{\partial y_k} \frac{b}{y} = \frac{\partial}{\partial y_k} \frac{\sum_{k=1}^{n} b_k}{\sum_{k=1}^{n} y_k} = \frac{(\sum_{k=1}^{n} b_k)'(\sum_{k=1}^{n} y_k) - (\sum_{k=1}^{n} b_k)(\sum_{k=1}^{n} y_k)'}{(\sum_{k=1}^{n} y_k)^2} = \frac{y(\sum_{k=1}^{n} b_k)' - b(\sum_{k=1}^{n} y_k)'}{\mathrm{y}^2} = \frac{y \sum_{k=1}^{n} (\frac{SC_k - \mathrm{RO_c}}{\mathrm{BF_c} - \mathrm{RO_c}}) - \mathrm{n}b}{\mathrm{y}^2} (\text{see Eq. A6}) =$$

$$\frac{y \sum_{k=1}^{n} (SC_k - \mathrm{RO_c}) - \mathrm{n}b(\mathrm{BF_c} - \mathrm{RO_c})}{\mathrm{y}^2(\mathrm{BF_c} - \mathrm{RO_c})} = \frac{\sum_{k=1}^{n} (SC_k - \mathrm{RO_c}) - \mathrm{nBFI}(\mathrm{BF_c} - \mathrm{RO_c})}{\mathrm{y}(\mathrm{BF_c} - \mathrm{RO_c})} \quad \text{(A8)}$$

Appendix B

Calculation of the sensitivity indices

$$\mathrm{S(BFI|BF_c)} = \frac{\Delta_{\mathrm{BF_c}} \mathrm{BFI}}{\mathrm{BFI}} \Big/ \frac{\Delta \mathrm{BF_c}}{\mathrm{BF_c}} = \frac{y\mathrm{RO_c} - \sum_{k=1}^{n} y_k SC_k}{y(\mathrm{BF_c} - \mathrm{RO_c})^2} \Delta \mathrm{BF_c} \frac{\mathrm{BF_c}}{\mathrm{BFI} \Delta \mathrm{BF_c}} (\text{see Eq. 7}) = \frac{\mathrm{BF_c}(y\mathrm{RO_c} - \sum_{k=1}^{n} y_k SC_k)}{y\mathrm{BFI}(\mathrm{BF_c} - \mathrm{RO_c})^2} \quad \text{(B1)}$$

$$\mathrm{S(BFI|RO_c)} = \frac{\Delta_{\mathrm{RO_c}} \mathrm{BFI}}{\mathrm{BFI}} \Big/ \frac{\Delta \mathrm{RO_c}}{\mathrm{RO_c}} = \frac{\sum_{k=1}^{n} y_k SC_k - y\mathrm{BF_c}}{y(\mathrm{BF_c} - \mathrm{RO_c})^2} \Delta \mathrm{RO_c} \frac{\mathrm{RO_c}}{\mathrm{BFI} \Delta \mathrm{RO_c}} (\text{see Eq. 8}) = \frac{\mathrm{RO_c}(\sum_{k=1}^{n} y_k SC_k - y\mathrm{BF_c})}{y\mathrm{BFI}(\mathrm{BF_c} - \mathrm{RO_c})^2} \quad \text{(B2)}$$

Appendix C

Prove that $\partial \mathrm{BFI}/\partial y_k$ is far less than 1 d/m^3.

$$\frac{\partial \mathrm{BFI}}{\partial y_k} = \frac{\sum_{k=1}^{n} (SC_k - \mathrm{RO_c}) - \mathrm{nBFI}(\mathrm{BF_c} - \mathrm{RO_c})}{\mathrm{y}(\mathrm{BF_c} - \mathrm{RO_c})} (\text{see Eq. A8}) \quad \text{(C1)}$$

Because of n>0, BFI>0, ($\mathrm{BF_C}$-$\mathrm{RO_C}$)>0, the above formula can be simplified:

$$\frac{\partial \mathrm{BFI}}{\partial y_k} < \frac{\sum_{k=1}^{n} (SC_k - \mathrm{RO_c})}{\mathrm{y}(\mathrm{BF_c} - \mathrm{RO_c})} \quad \text{(C2)}$$

Since $\mathrm{BF_C}$ is usually much larger than SC_k, the above formula can be rewritten as:

$$\frac{\partial \mathrm{BFI}}{\partial y_k} < \frac{\sum_{k=1}^{n} (\mathrm{BF_c} - \mathrm{RO_c})}{\mathrm{y}(\mathrm{BF_c} - \mathrm{RO_c})} = \frac{\mathrm{n}(\mathrm{BF_c} - \mathrm{RO_c})}{\mathrm{y}(\mathrm{BF_c} - \mathrm{RO_c})} = \frac{\mathrm{n}}{\mathrm{y}} = \frac{1}{\bar{\mathrm{y}}} \quad \text{(C3)}$$

The daily average streamflow ($\bar{y}$) is usually much larger than 1 m^3/d, so $\partial BFI/\partial y_k$ is far less than 1 d/m^3.

Data availability

All streamflow and conductivity data can be retrieved from the United States Geological Survey's (USGS) National Water Information System (NWIS) website use the special gage number, http://waterdata.usgs.gov/nwis.

Author contributions

Weifei Yang, Changlai Xiao and Xiujuan Liang designed the research train of thought. Weifei Yang and Changlai Xiao completed the parameters' sensitivity analysis. Xiujuan Liang completed the uncertainty estimate of BFI. Weifei Yang carried out most of the data analysis and prepared the manuscript with contributions from all co-authors.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (41572216), the Provincial School Co-construction Project Special -- Leading Technology Guide (SXGJQY2017-6), the China Geological Survey Shenyang Geological Survey Center "Changji Economic Circle Geological Environment Survey" project (121201007000150012), and the Jilin Province Key Geological Foundation Project (2014-13). We thank the anonymous reviewers for useful comments to improve the manuscript.

References

Cartwright, I., Gilfedder, B., and Hofmann, H.: Contrasts between estimates of baseflow help discern multiple sources of water contributing to rivers, Hydrol. Earth Syst. Sci., 18, 15-30, doi:10.5194/hess-18-15-2014, 2014.

Costelloe, J. F., Peterson, T. J., Halbert, K., Western, A. W., and McDonnell, J. J.: Groundwater surface mapping informs sources of catchment baseflow, Hydrol. Earth Syst. Sci., 19, 1599-1613, doi:10.5194/hess-19-1599-2015, 2015.

Eckhardt, K.: How to construct recursive digital filters for baseflow separation, Hydrol. Process., 19, 507-515, doi:10.1002/hyp.5675, 2005.

Eckhardt, K.: Technical Note: Analytical sensitivity analysis of a two parameter recursive digital baseflow separation filter, Hydrol. Earth Syst. Sci., 16, 451-455, doi:10.5194/hess-16-451-2012, 2012.

Ernest, L.: Gaussian error propagation applied to ecological data: Post-ice-storm-downed woody biomass, Ecol. Monogr., 75, 451-466, doi.org/10.1890/05-0030, 2005.

Genereux, D.: Quantifying uncertainty in tracer-based hydrograph separations, Water Resour. Res, 34, 915-919, doi:10.1029/98wr00010, 1998.

Hamilton, A.S., and Moore R.D.: Quantifying Uncertainty in Streamflow Records, Can. Water Resour. J., 37, 3-21, doi:10.4296/cwrj3701865, 2012

Huang, Z. P., Chen, Y. F.: Hydrological statistics, China Water&Power Press, Beijing, China, 2011.

Kline, S. J.: The purposes of uncertainty analysis, J. Fluids Eng., 107, 153-160, 1985.

Lott, D. A., Stewart, M. T.: Base flow separation: A comparison of analytical and mass balance methods, J. Hydrol., 535, 525-533, doi:10.1016/j.jhydrol.2016.01.063, 2016.

Mei, Y., Anagnostou, E. N.: A hydrograph separation method based on information from rainfall and runoff records, J. Hydrol., 523, 636-649, doi:10.1016/j.jhydrol.2015.01.083, 2015.

Miller, M. P., Susong, D. D., Shope, C. L., Heilweil, V. M., and Stolp, B. J.: Continuous estimation of baseflow in snowmelt-dominated streams and rivers in the Upper Colorado River Basin: A chemical hydrograph separation approach, Water Resour. Res., 50, 6986–6999, doi:10.1002/2013WR014939, 2014.

Munyaneza, O., Wenninger, J., and Uhlenbrook, S.: Identification of runoff generation processes using hydrometric and tracer methods in a meso-scale catchment in Rwanda, Hydrol. Earth Syst. Sci., 16, 1991-2004, doi:10.5194/hess-16-1991-2012, 2012.

Okello, A. M. L. S., Uhlenbrook, S., Jewitt, G. P. W., Masih, L., Riddell, E. S., and Zaag, P.V.: Hydrograph separation using tracers and digital filters to quantify runoff components in a semi-arid mesoscale catchment, Hydrol. Process., 32, 1334-1350, doi:10.1002/hyp.11491, 2018.

Stewart, M., Cimino, J., and Rorr, M.: Calibration of base flow separation methods with streamflow conductivity, Ground Water, 45, 17-27, doi:10.1111/j.1745-6584.2006.00263.x, 2007.

Taylor, J. R.: An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements, Univ. Sci. Books, Mill Valley, Calif., 1982.

Wagner, R. J., Boulger R. W. Jr., Oblinger, C. J., and Smith, B. A.: Guidelines and standard procedures for continuous water-quality monitors-Station operation, record computation, and data reporting, U.S. Geol. Surv. Tech. Meth., 1-D3, 51 pp, 2006.

Zhang, J., Zhang, Y., Song, J., and Cheng, L.: Evaluating relative merits of four baseflow separation methods in Eastern Australia, J. Hydrol., 549, 252-263, doi:10.1016/j.jhydrol.2017.04.004, 2017.

Tables

Table 1. Basic information, parameter sensitivity analysis, and uncertainty estimation results for 24 basins in the United States. Footnote "a" in the "Area" column indicates that the values are estimated based on data from adjacent sites.

State	Gage Number	N	Area	Mean BF_C	RO_C	Mean Baseflow	BFI	S(BFI|BF_C)	S(BFI|RO_C)	W_{BFI}	Mean W_{fbf}
		days	km^2	µs/cm	µs/cm	m^3/s					
FL	2298202	1808	966	1149.1	292.5	2.12	0.31	-1.32	-0.76	0.05	0.12
FL	2310545	1218	119[a]	6404.7	531.5	0.65	0.17	-1.11	-0.44	0.05	0.06
FL	2310650	779	77[a]	6558.7	3210.0	0.90	0.57	-1.84	-0.79	0.18	0.27
FL	2303000	728	570	432.7	120.5	2.32	0.34	-1.30	-0.77	0.06	0.14
FL	2298488	1303	76	737.3	194.0	0.14	0.38	-1.32	-0.58	0.14	0.18
FL	2298554	899	207[a]	969.2	320.5	0.50	0.30	-1.25	-1.22	0.13	0.27
FL	2298492	1478	16	1238.2	304.0	0.05	0.30	-1.11	-0.82	0.13	0.31
FL	2298495	330	10	1870.0	662.0	0.29	0.25	-1.52	-1.65	0.03	0.08
FL	2298527	807	23	1410.7	201.5	0.10	0.19	-1.03	-0.74	0.06	0.18
FL	2298530	1510	17	1460.8	348.0	0.14	0.29	-1.27	-0.77	0.08	0.13
FL	2297100	2979	342	1260.6	221.5	0.92	0.25	-1.18	-0.64	0.08	0.20
FL	2313000	787	4727	407.2	173.0	5.89	0.51	-1.71	-0.71	0.19	0.28
FL	2300500	821	386	447.9	83.0	0.30	0.20	-1.21	-0.89	0.05	0.11
ND	5057000	1401	16757	1420.6	610.0	2.08	0.51	-1.75	-0.74	0.14	0.21
ND	5056000	1277	5361	1681.4	546.0	3.61	0.44	-1.50	-0.60	0.07	0.14
TX	8068275	2801	482	361.7	65.0	0.57	0.15	-1.18	-1.23	0.06	0.11
GA	2336300	1235	225	230.4	63.0	0.79	0.31	-1.28	-0.88	0.16	0.33
GA	2207120	1383	417	312.5	59.0	1.48	0.24	-1.14	-0.76	0.09	0.20
SC	2160105	1363	1966	124.7	51.0	6.36	0.36	-1.56	-1.30	0.14	0.27
SC	2160700	1392	1150	148.7	51.0	4.45	0.37	-1.40	-0.94	0.15	0.28
MO	6894000	1375	477	1031.9	334.0	0.79	0.25	-1.40	-1.50	0.13	0.22
MO	6895500	802	1258481	786.7	428.0	939.98	0.57	-2.17	-0.90	0.06	0.20
ND	5082500	1274	77959	1390.6	427.0	77.19	0.38	-1.30	-0.77	0.15	0.26
KS	7144780	575	1847	1389.1	678.0	1.73	0.54	-1.73	-0.91	0.14	0.26
				Mean			0.34	-1.40	-0.89	0.11	0.20
				Standard deviation (STDEV)			0.13	0.28	0.29	0.05	0.08

Figures

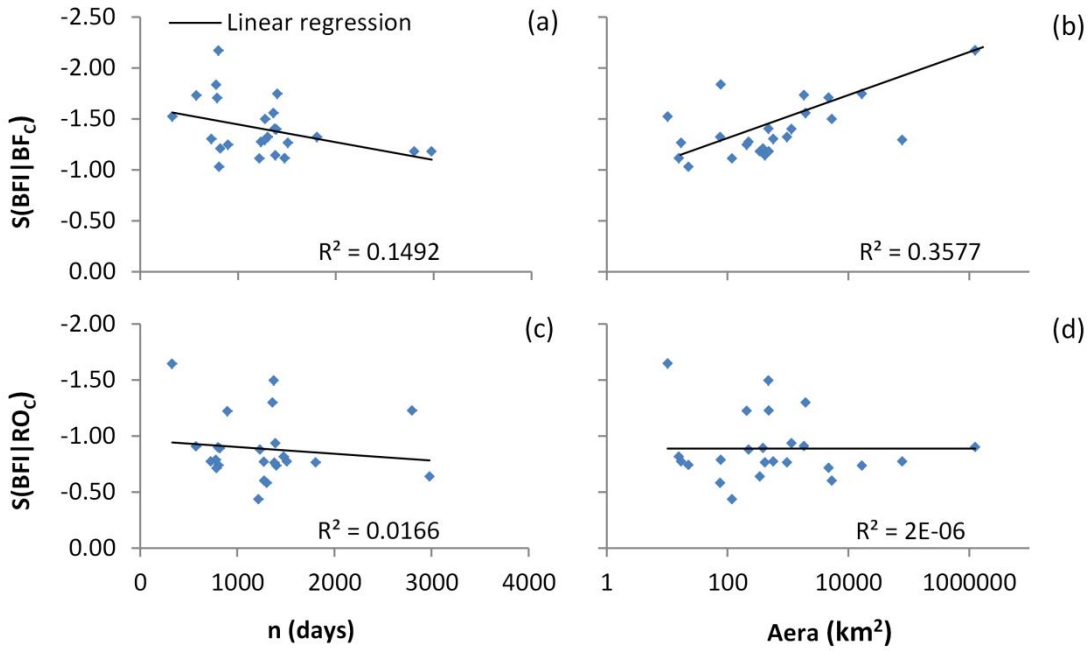

Figure 1. Scatter plots of sensitivity indices vs. time series (n) and drainage area of the 24 US basins. The watershed area uses a logarithmic axis, while the others are linear axes.

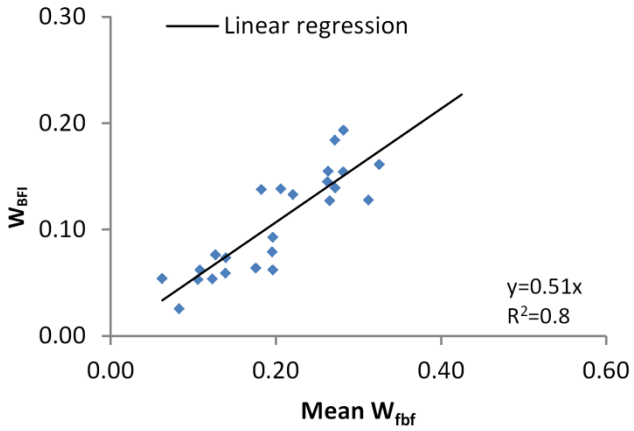

Figure 2. Scatter plot of uncertainty in BFI (W_{BFI}) and mean uncertainty in f_{bf} (Mean W_{fbf}).