Journal of Hydrology 522 (2015) 203–210

Contents lists available at ScienceDirect

Journal of Hydrology

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

A new approach for continuous estimation of baseflow using discrete water quality data: Method description and comparison with baseflow estimates from two existing approaches

Matthew P. Miller [a,*], Henry M. Johnson [b], David D. Susong [a], David M. Wolock [c]

[a] U.S. Geological Survey, Utah Water Science Center, 2329 Orton Circle, Salt Lake City, UT 84119, USA
[b] U.S. Geological Survey, Oregon Water Science Center, 2130 SW 5[th] Ave, Portland, OR 97201, USA
[c] U.S. Geological Survey, Kansas Water Science Center, 4821 Quail Crest Place, Lawrence, KS 66049, USA

ARTICLE INFO

Article history:
Received 6 August 2014
Received in revised form 12 December 2014
Accepted 15 December 2014
Available online 24 December 2014
This manuscript was handled by Corrado Corradini, Editor-in-Chief, with the assistance of Rao S. Govindaraju, Associate Editor

Keywords:
Baseflow
Conductivity mass balance
Specific conductance
Hydrograph separation
Groundwater

SUMMARY

Understanding how watershed characteristics and climate influence the baseflow component of stream discharge is a topic of interest to both the scientific and water management communities. Therefore, the development of baseflow estimation methods is a topic of active research. Previous studies have demonstrated that graphical hydrograph separation (GHS) and conductivity mass balance (CMB) methods can be applied to stream discharge data to estimate daily baseflow. While CMB is generally considered to be a more objective approach than GHS, its application across broad spatial scales is limited by a lack of high frequency specific conductance (SC) data. We propose a new method that uses discrete SC data, which are widely available, to estimate baseflow at a daily time step using the CMB method. The proposed approach involves the development of regression models that relate discrete SC concentrations to stream discharge and time. Regression-derived CMB baseflow estimates were more similar to baseflow estimates obtained using a CMB approach with measured high frequency SC data than were the GHS baseflow estimates at twelve snowmelt dominated streams and rivers. There was a near perfect fit between the regression-derived and measured CMB baseflow estimates at sites where the regression models were able to accurately predict daily SC concentrations. We propose that the regression-derived approach could be applied to estimate baseflow at large numbers of sites, thereby enabling future investigations of watershed and climatic characteristics that influence the baseflow component of stream discharge across large spatial scales.

Published by Elsevier B.V.

1. Introduction

Scientists and managers are often interested in identifying how watershed characteristics (e.g. geology, land use, soil type, etc.) and climatic conditions influence baseflow discharge to streams. Addressing such processes requires quantitative estimates of baseflow discharge across a gradient of watershed types. The development of quantitative methods for baseflow estimation is also necessary to understand water budgets (Stewart et al., 2007), estimate groundwater discharge (Arnold and Allen, 1999) and associated effects on stream temperature (Hill et al., 2013), and address questions of the vulnerability and response of the water cycle to natural and human-induced change in environmental conditions, such as stream vulnerability to legacy nutrients (Tesoriero et al., 2013). Given the importance of baseflow, many methods have been used to quantify the baseflow component of stream discharge beginning with Boussinesq (1877).

Approaches for baseflow estimation can be grouped into two general categories: graphical hydrograph separation (GHS) methods, which rely on stream discharge data alone, and tracer mass balance (MB) methods, which rely on chemical constituents in the stream, stream discharge, and the streamflow end-member constituent concentrations (runoff and baseflow). Many different approaches for GHS exist, including recession curve methods and digital filter methods. Recession curve methods are generally considered more objective than digital filter methods because they provide an assumed integrated signal of basin hydrologic and geologic characteristics through identification of a linear recession-constant based on the falling limb of the hydrograph (Barnes, 1939; Hall, 1968; Gardner et al., 2010). However, the ability of recession curve methods to quantify groundwater discharge to streams has been questioned because of the accuracy of the method assumptions (Halford and Mayer, 2000). Digital filter

* Corresponding author. Tel.: +1 801 908 5065.
E-mail address: mamiller@usgs.gov (M.P. Miller).

http://dx.doi.org/10.1016/j.jhydrol.2014.12.039
0022-1694/Published by Elsevier B.V.

methods either filter high frequency (assumed to be surface runoff) signals from low frequency (assumed to be baseflow) signals (Nathan and McMahon, 1990), or identify and connect successive minima on a stream hydrograph, and define baseflow as the line connecting the minima (Wahl and Wahl, 1988; Wolock, 2003). The definitions of basin-specific parameters used in these methods are generally subjective and not based on hydrologic processes (Stewart et al., 2007).

It has been suggested that MB methods for baseflow estimation are more objective than GHS because measured stream water concentrations, and either measured or estimated end-member concentrations, are related to physical and chemical processes and flow paths in the basin (Stewart et al., 2007; Zhang et al., 2013). One type of MB method that is commonly applied is the conductivity mass balance (CMB) method, which uses specific conductance (SC) as a chemical tracer for hydrograph separation. One advantage of CMB over other types of MB methods is that SC is relatively easy and inexpensive to measure. Additionally, high frequency SC measurements can be obtained using in-situ SC probes. High frequency SC data and CMB methods have been used to estimate baseflow across gradients of watershed size and land use settings (Covino and McGlynn, 2007; Miller et al., 2014; Pellerin et al., 2007; Stewart et al., 2007).

While CMB methods are generally considered to be more objective than GHS methods, their application is limited by the fact that they require high frequency SC records that are not always widely available over long time periods or spanning large numbers of watersheds. Multiple studies have developed methods to calibrate GHS estimates of baseflow to CMB estimates of baseflow (Lott and Stewart, 2013; Stewart et al., 2007; Zhang et al., 2013). Once calibrated at a specific stream location, and assuming that the end-member SC concentrations are constant over time, the GHS methods can be applied to long term stream discharge records at that location to estimate baseflow for time periods that span date ranges greater than those for which high frequency SC data are available. Li et al. (2014) showed that as little as six months of high frequency SC data can be used to calibrate a recursive digital filter model, which can then be applied to long term stream discharge records to estimate baseflow. This approach overcomes the CMB limitations associated with the lack of long-term SC records, but is only applicable at sites that have high frequency SC data available for GHS calibration. Unfortunately, high frequency SC data are not generally available at large numbers of sites within a given region. Therefore, the use of CMB, or calibration of GHS to CMB, to estimate baseflow and quantify environmental drivers of baseflow discharge across broad spatial scales is limited.

We propose that discrete SC concentration data and daily mean discharge data, which are frequently available at large numbers of sites, can be used with a CMB method to estimate baseflow at a daily time-step for the period of record of discharge data, thereby increasing the number of sites at which CMB can be used to estimate baseflow. The proposed approach involves the calibration of site-specific regression models that relate discrete SC concentrations to stream discharge and time to predict daily SC concentrations, and subsequently regression-derived CMB baseflow estimates, for the period of stream discharge record. A similar regression approach has been used to estimate water quality data, and subsequently groundwater discharge to a tropical stream for time periods when no water quality data exist (Genereux et al., 2005), but has not been applied to a number of sites and compared with other baseflow estimates from the same sites. The objective of this study is to test the proposed approach by comparing the regression-derived baseflow estimates with CMB baseflow estimates calculated using measured high frequency SC data (assumed to be the most objective estimates of baseflow) at twelve snowmelt dominated streams and rivers in the Upper Colorado River Basin (UCRB). As previously reported by Miller et al. (2014), CMB methods are well suited for estimating baseflow in snowmelt dominated watersheds. Baseflow estimates calculated using a commonly applied GHS model were also compared with measured CMB baseflow estimates.

2. Materials and methods

2.1. Site description

The UCRB is a heavily regulated watershed located in the western United States and drains an area of 294,000 km^2. The headwaters are high elevation catchments in the Rocky Mountains and the downstream end of the UCRB is located at Page, AZ, downstream of Lake Powell on the Colorado River (Fig. 1). Miller et al. (2014) estimated baseflow discharge at a daily time step for the period of record at fourteen sites draining large watersheds in the UCRB characteristic of snowmelt dominated hydrology using measured high frequency SC data with a CMB approach. As part of this process sites were screened for impacts due to anthropogenic activities. Twelve of these fourteen sites are included in the present methods comparison (Fig. 1, Table 1). Two of the fourteen sites – The Gunnison River at Delta, CO and The Uncompahgre River at Colona, CO – are not included in the present study because the short periods of record for which high frequency SC data are available at these sites resulted in a limited discrete SC data set that was not adequate for development of regression models to estimate daily SC concentrations. Drainage areas range from 1500 km^2 at PLAT to 62,000 km^2 at CO_3. Average baseflow estimates range from 1.0 ± 1.2 m^3/s to 103 ± 9.6 m^3/s, and the fraction of total streamflow estimated to be baseflow ranges from 11% to 59% (Table 1). Detailed site descriptions for these twelve locations are available in Miller et al. (2014).

2.2. Data sources

Daily mean discharge, daily mean SC, and discrete SC data were obtained from the U.S. Geological Survey (USGS) National Water Information System (NWIS) database. The date ranges for which data were acquired were limited to date ranges for which both daily mean discharge and daily mean SC data were available. Periods of records ranged from 3 to 37 years and the number of discrete samples used in regression model calibration (for estimation of daily regression-derived SC concentrations) ranged from 17 to 623 (Table 1). Detailed information regarding the periods of record, average discharge, and average SC at each site are available in Miller et al. (2014).

2.3. Regression-derived daily SC

Discrete SC values were related to daily discharge, time, and up to 7 additional variables that describe annual seasonality and variability in stream discharge of varying length. Nine different models were fitted at sites having more than 10 years of discrete SC data and 7 models were fitted at sites having less than 10 years of discrete SC data. The general form of the regression equations is described by Eq. (1). Table 2 shows the nine permutations of Eq. (1) that were used to simulate SC. Regressions were conducted in R (R Development Team, 2014).

$$\ln SC = \ln Q + \ln Q^2 + T + \sin 2\pi T + \cos 2\pi T + \sin 4\pi T + \cos 4\pi T + FA \quad (1)$$

where SC is the estimated discrete daily specific conductance (μS/cm), Q is daily discharge (m^3/s), T is time expressed as decimal years (*e.g.* 2005.25 = April 1, 2005), and FA is an additive combination of one of the groups of flow anomalies generated by