Published in Journal of Hydrology, Vol. 458–459, 2012, p. 28-39
which should be used for any reference to this work

Evaluation of outputs from automated baseflow separation methods against simulated baseflow from a physically based, surface water-groundwater flow model

D. Partington [a,*], P. Brunner [b], C.T. Simmons [c], A.D. Werner [c], R. Therrien [d], H.R. Maier [a], G.C. Dandy [a]

[a] School of Civil, Environmental and Mining Engineering, The University of Adelaide, Adelaide, South Australia 5005, Australia
[b] Centre of Hydrogeology and Geothermics (CHYN), Rue Emile-Argand 11, CP 158, CH-2009 Neuchâtel, Switzerland
[c] School of the Environment and National Centre for Groundwater Research and Training, Flinders University, GPO Box 2100, Adelaide, South Australia 5001, Australia
[d] Department of Geology and Geological Engineering, Université Laval, Quebec, Canada G1K 7P4

Keywords:
Baseflow separation
Surface water–groundwater interaction
Physically based surface–subsurface model
HydroGeoSphere

S U M M A R Y

Baseflow is often considered to be the groundwater discharge component of streamflow. It is commonly estimated using conceptual models, recursive filters or a combination of the two. However, it is difficult to validate these methods due to the current challenges of measuring baseflow in the field. In this study, simulation of a synthetic catchment's response to rainfall is carried out using a fully integrated surface water-groundwater flow model. A series of rainfall events with differing recovery periods and varied antecedent moisture conditions is considered to span a range of different streamflow generation dynamics. Baseflow is estimated for the outlet hydrograph of the synthetic catchment using a selection of commonly used automated baseflow separation methods. These estimates are compared to the baseflow signal obtained from the numerical model, which serves as the control experiment. Results from these comparisons show that depending on the method used, automated baseflow separation underestimates the simulated baseflow by as much as 28%, or overestimates it by up to 74%, during rainfall events. No separation method is found to be clearly superior to the others, as the performance of the various methods varies with different soil types, antecedent moisture conditions and rainfall events. The differences between the various approaches clearly demonstrate that the baseflow separation methods investigated are not universally applicable.

1. Introduction

Quantifying baseflow contributions to streamflow is of great interest in the understanding, identification and quantification of streamflow generation processes, in particular where baseflow supports important ecosystems and/or provides critical dry season water supply (e.g. Smakhtin, 2001; Werner et al., 2006). The term baseflow is often referred to as the groundwater contribution to streamflow (e.g. Freeze, 1972; Brutsaert and Nieber, 1977; Eckhardt, 2005), although it is also referred to as the release from both groundwater and other natural storages of water that sustain streamflow between rainfall events (e.g. Hall, 1968; Smakhtin, 2001; Piggott et al., 2005). In this paper, the term baseflow is used to describe groundwater discharge that reaches the stream, not including interflow through the vadose zone.

Baseflow can be inferred through field measurements of temperature, artificial and natural tracer concentrations, and flow in seepage meters installed in stream beds (Becker et al., 2004; Cook et al., 2003, 2008). However, for practical reasons, it is very difficult to apply these techniques over an entire catchment. Furthermore, the required end members in chemical mass balance approaches are difficult to characterise (McCallum et al., 2010), which complicates baseflow estimates using measurement of tracers. Consequently, the available field methods do not currently allow accurate determination of spatially and temporally distributed baseflow. In the absence of detailed field data, but where a streamflow hydrograph is available, baseflow is therefore often estimated using simple baseflow separation methods.

Since the early twentieth century, a variety of methods has been developed to estimate baseflow. The earliest methods and some of the more recent ones are based on a linear storage–discharge relationship between aquifer and stream (e.g. Maillet, 1905; Barnes, 1939; Hall, 1968; Boughton, 1993). More recently, non-linear

* Corresponding author. Tel.: +61 (0)8 8303 4323; fax: +61 (0)8 8303 4359.
E-mail addresses: daniel.partington@adelaide.edu.au (D. Partington), philip.brunner@unine.ch (P. Brunner), craig.simmons@flinders.edu.au (C.T. Simmons), adrian.werner@flinders.edu.au (A.D. Werner), rene.therrien@ggl.ulaval.ca (R. Therrien), holger.maier@adelaide.edu.au (H.R. Maier), graeme.dandy@adelaide.edu.au (G.C. Dandy).

storage–discharge relationships have also been applied to baseflow separation (e.g. Wittenberg, 1994, 2003; Wittenberg and Sivapalan, 1999) following theoretical studies suggesting that non-linear recessions are appropriate for some catchments. Also, other methods that use some form of hydrological reasoning have been developed without a physically based mathematical framework. Currently, the separation of baseflow from the streamflow hydrograph can be carried out utilising methods that can be grouped into the following four categories: (1) graphical separation (Sloto and Crouse, 1996), (2) recession analysis (Tallaksen, 1995), (3) conceptual models (Barnes, 1939; Singh and Stall, 1971; Furey and Gupta, 2001; Eckhardt, 2005; Huyck et al., 2005) and (4) recursive digital filters (Nathan and McMahon, 1990; Arnold and Allen, 1999).

The different categories of separation approaches as noted above have been compared and reviewed in several previous studies (Hall, 1968; Nathan and McMahon, 1990; Arnold et al., 1995; Chapman, 1999; Smakhtin, 2001; Schwartz, 2007; Eckhardt, 2008). The reviews of Hall (1968), Smakhtin (2001) and Schwartz (2007) provide a history of methods for baseflow separation, and discuss the problems related to the definition of baseflow, as well as the underlying assumptions of the different separation methods. In the context of identifying groundwater recharge from streamflow records, the underlying assumptions that underpin many methods were examined by Halford and Mayer (2000). They concluded that identifying the groundwater contribution from streamflow records can be ambiguous due to drainage exponentially decreasing from other sources, such as bank storage, wetlands and the unsaturated zone. Furthermore, they noted that simple automated methods are highly subjective with respect to their algorithmic structure, and affected by the same underlying assumptions as other more complex methods.

The analyses used in comparative studies to evaluate baseflow separation methods (e.g. Nathan and McMahon, 1990; Arnold et al., 1995; Mau and Winter, 1997; Chapman, 1999; Halford and Mayer, 2000; Schwartz, 2007; Eckhardt, 2008) are often based on subjective measures, such as the plausibility of hydrological behaviour, rather than a quantitative comparison to a known and well-quantified baseflow hydrograph. This point was emphasised by Mau and Winter (1997) who highlighted the need to validate baseflow estimates to avoid issues related to subjective measures and other shortcomings of simplified methods. Unfortunately, to date, no measured baseflow hydrograph at the catchment scale is available. Until comprehensive data and better observation techniques come into existence, numerical models, although theoretical, provide the best independent conceptualisation of baseflow dynamics in catchments under different forcing functions and hydrological conditions.

Some studies (e.g. Szilagyi, 2004; Fenicia et al., 2006; Ferket et al., 2010) have compared baseflow estimated by separation methods with simulated baseflow from lumped and semi-distributed catchment models. However, a critical analysis of separation methods is inhibited by the lack of a reliable estimate of baseflow, as well as some simplifications in the lumped and semi-distributed models, such as the aquifer storage–discharge relationship (linear reservoir in Fenicia et al. (2006), the sum of multiple linear reservoirs in Szilagyi (2004), and the Boussinesq-equation in Ferket et al. (2010)).

More recently, physically based separation methods have been developed based on process-based formulations of fluid mass balance equations of an aquifer (e.g. Furey and Gupta, 2001, 2003; Huyck et al., 2005). They constitute an important step in overcoming the subjective elements of earlier simpler methods, and attempt to alleviate some of their simplifying assumptions. As well as the development of two physically based baseflow separation methods, Furey and Gupta (2003) evaluated their methods against a complex physically based numerical model of a hill-slope. Their study appears to be the first to critically compare baseflow separation methods with a physically based numerical model. However, by considering only the discharge from a single 2D synthetic hill-slope (rather than a catchment), Furey and Gupta (2003) neglected important catchment-scale processes, such as channel routing (e.g. streamflow attenuation and translation) and channel losses (through losing sections, abstraction and evaporation).

Fully integrated surface and subsurface flow models, some examples of which are HydroGeoSphere (HGS) (Therrien et al., 2009), MODHMS (Hydrogeologic Inc., 2006) and Parflow (Kollet and Maxwell, 2006), are useful for evaluating simpler models because they do not need to assume a functional relationship between baseflow and streamflow, or simple empirical relations. These models typically represent 3D variably saturated subsurface flow with Richard's equation, and 1D and 2D surface flow with the diffusion wave approximation to the St. Venant equations. A unique feature of fully integrated models is that water that is derived from rainfall is allowed to partition into overland flow, streamflow, evaporation, infiltration and recharge, whilst subsurface discharge to surface water features, such as lakes and streams, occurs in a physically based fashion (Therrien et al., 2009). Therefore, physically based numerical models provide an excellent means for comparison of baseflow separation methods if the modelled baseflow component of streamflow can be extracted.

Using physically based numerical models of 3D systems to evaluate baseflow separation methods is difficult because the baseflow component of streamflow is not a standard output. For a 2D hill-slope model, the baseflow component of outflow is simply groundwater discharge. However, in the extension beyond 2D hill-slopes, the baseflow component of simulated streamflow must be calculated in some other way. As highlighted by Partington et al. (2011), the available integrated models do not explicitly report the groundwater contribution to streamflow. This problem is of particular importance for catchments where the flow regime between surface water and groundwater is changing (e.g. gaining to losing sections of a stream and vice versa). To overcome these difficulties, Partington et al. (2011) developed a hydraulic mixing-cell (HMC) approach that allows extraction of the groundwater contribution to streamflow within integrated surface and subsurface flow models. Combining the HMC approach with the HydroGeoSphere model, they demonstrated that spatiotemporal fluxes into and out of a stream can be translated to a point along the stream allowing for meaningful hydrograph separation. The HMC method allows for theoretical examination of baseflow dynamics within a 3D catchment model, thus providing a platform for comparison to automated baseflow separation methods.

In the current study, the HMC method is used in conjunction with HGS in order to compare the outputs from a series of commonly used automated baseflow separation methods. A numerical control experiment is developed using the integrated model to simulate hydrological processes within a synthetic catchment. Multiple simulations are carried out using differing initial, hydrologic and forcing conditions in order to generate a series of outflow and baseflow hydrographs. Baseflow separation methods are then applied to the outflow hydrographs from the simulations. This allows comparison of the baseflow obtained from the separation methods to the simulated baseflow. The commonly used separation methods are based on graphical, conceptual and digital filter methods. The analysis is limited to automated methods that are readily available and that only rely on streamflow discharge data and catchment area. This analysis does not include an assessment of more complex physically based methods. However, it is noteworthy that this approach could also be used to test physically based methods, e.g. those developed by Furey and Gupta (2001).

2. Methodology

The HydroGeoSphere (HGS) model used here is a fully integrated, physically based model that simultaneously simulates 3D variably saturated subsurface flow and 2D surface flow (Therrien et al., 2009). Water is exchanged between the surface and subsurface domains through a first-order leakage relation based on the head difference between the domains. The model also accounts for evapotranspiration as a function of the leaf area index, soil moisture and root depth. For further details on the numerical formulation and a review of the code, the reader is referred to Park et al. (2009), Therrien et al. (2009) and Brunner and Simmons (2011).

2.1. The synthetic catchment

The geometry of the catchment is loosely based on the tilted V-catchment employed by Panday and Huyakorn (2004). As in Panday and Huyakorn (2004), the catchment is symmetrical and therefore only half of the catchment is modelled (shown in Fig. 1). This particular geometry is an ideal synthetic framework for generating hydrographs, because a range of hydrological processes control the catchment's behaviour. These processes include 3D saturated/unsaturated groundwater flow, infiltration/exfiltration, overland flow and streamflow. An analysis of the Panday and Huyakorn (2004) synthetic catchment highlighting some of the issues associated with their model setup was undertaken by Gaukroger and Werner (2011), and in response to these, several modifications to the original setup are adopted here. The steep slopes and initially horizontal water table in the Panday and Huyakorn (2004) case cause all groundwater discharge to be concentrated around the outlet. Reducing the slope of the catchment (particularly along the stream) creates a greater spatial distribution of the surface–subsurface exchanges throughout the catchment. Therefore, the slopes perpendicular and parallel to the stream are decreased from 0.05 m/m and 0.02 m/m to 0.002 m/m and 0.0005 m/m, respectively. Furthermore, the horizontal water table represents an unrealistic (overly dry) initial condition. To start the model from more realistic initial conditions, the catchment is saturated and allowed to drain for between 7 and 9 months without any precipitation events. The original roughness coefficients used for the hill-slope and stream domains cause overland flow to be dominant parallel and adjacent to the stream, rather than in the stream (Gaukroger and Werner, 2011). In order to allow overland flow to discharge into the stream as it reaches the banks (rather than flowing alongside the stream), the same roughness (0.015 m/s$^{1/3}$) is used in both the overland flow and stream domains. Also, the plane adjacent to the stream is raised by 0.6 m over a 5 m length to promote direct discharge of groundwater to the stream as opposed to upslope exfiltration or return flow. Finally, the areal extent of the catchment is increased from 810,000 m^2 to 3,220,000 m^2 by doubling the original length and width of the catchment (keeping the stream width at 10 m). This reduces boundary effects and increases aquifer storage capacity, which promotes sustained baseflow contributions to the stream. Given the modifications outlined above, a wide range of hydrographs can be generated by changing the forcing functions (e.g. rainfall, groundwater pumping and evapotranspiration).

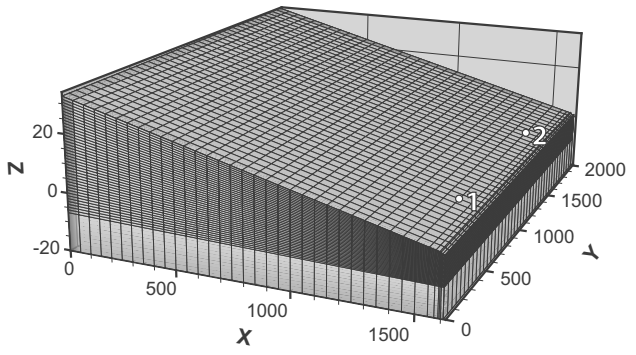

Fig. 1. Modified tilted V-catchment used for simulation of the synthetic catchment's rainfall response. Points 1 and 2 denote locations of groundwater pumps. Note that due to the symmetry of the catchment, only half of it is shown.

The bottom elevation of the model domain is set at −20 m relative to the 0 m elevation of the streambed at the outlet. The aquifer properties are homogeneous and isotropic. In separate model scenarios, two different sets of properties of the aquifer material are considered (Table 1). Properties for evaporation and transpiration are also included in Table 1.

The spatial discretisation in the catchment model is as follows: grid spacing along the x axis is 50 m from x = 0–1550 m, 25 m from x = 1550–1575 m, 15 m from x = 1575–1590 m, 5 m from x = 1590–1600 m and 10 m from x = 1600–1610 m. The grid spacing along the y axis is 50 m. The vertical grid discretisation increases in thickness according to the slopes perpendicular and parallel to the stream. Vertical discretisation along the z axis ranges from 0.25 m to 1 m for the first 10 m below the surface. The time steps used in the model vary in accordance with an adaptive time-stepping approach with a maximum step of 1000 s. A no flow boundary is applied at the bottom and sides of the model domain. A critical depth boundary is used at (x, y, z) = (1600 m, 0 m, 0 m) and (1610 m, 0 m, 0 m) to control the outflow at the stream outlet.

Table 1
Surface and subsurface parameters for the synthetic catchment model. For a detailed description of these model parameters see Therrien et al. (2009).

Parameter	Value
Surface	
Manning's roughness	0.015 s/m$^{1/3}$
Rill storage height	0.001 m
Obstruction storage height	0.0 m
Transpiration fitting parameter c1	0.3
Transpiration fitting parameter c2	0.2
Transpiration fitting parameter c3	10
Leaf area index	2.08
Wilting point	0.2
Field capacity	0.32
Oxic limit	0.76
Anoxic limit	0.9
Limiting saturation (minimum)	0.2
Limiting saturation (maximum)	0.32
Canopy storage parameter	0.0 m
Initial interception storage	0.0 m
Subsurface	
Sand – porosity	0.1
Sand – saturated hydraulic conductivity	8.25×10^{-5} m/s
Sand – Van Genuchten α	14.5 m^{-1}
Sand – Van Genuchten β	2.68
Sand – residual saturation θ_r	0.045
Loamy sand – porosity	0.1
Loamy sand – saturated hydraulic conductivity	4.05×10^{-5} m/s
Loamy sand – Van Genuchten α	12.4 m^{-1}
Loamy sand – Van Genuchten β	2.28
Loamy sand – residual saturation θ_r	0.057
Evaporation depth (quadratic decay function)	3 m
Root depth (quadratic decay function)	3 m
Surface/subsurface coupling	
Coupling length	0.5 m

2.2. Baseflow calculation using the HMC method

The tracking of the streamflow generation mechanisms within a model simulation requires the tracking of the spatiotemporal fluxes into, out of, and along the stream. Parcels of groundwater discharging directly to the stream are "tracked" using the HMC method (Partington et al., 2011) to allow determination of when (and if) groundwater contributes to streamflow, as measured at the outlet (or at any other location along the stream). The HMC method accounts for the travel time along the stream and the spatiotemporal variation in the surface–subsurface exchange fluxes, thereby separating the simulated streamflow hydrograph into baseflow, overland flow and direct rainfall to the stream. A version of HGS that has the HMC method incorporated into it is used to simulate the outlet hydrograph of the synthetic catchment in response to a series of rainfall events, and considering groundwater pumping and evapotranspiration. The calculated HMC baseflow is used as the control experiment with which baseflow separations of the simulated hydrograph are compared.

2.3. Baseflow separation using automated methods

The automated methods for baseflow separation used in this study are implemented using the programs HYSEP (Sloto and Crouse, 1996), PART (Rutledge, 1998) and BFLOW (Arnold and Allen, 1999). The Eckhardt filter (Eckhardt, 2005) is also used. All of these approaches are well established methods, are readily available, and were previously compared in the study of Eckhardt (2008). However, they were judged subjectively based on hydrological plausibility. Detailed descriptions of all approaches can be found in the above cited literature, hence only a very brief overview of these methods is provided below.

The HYSEP program allows the use of three curve fitting methods of hydrograph separation; sliding interval (HYSEP1), fixed interval (HYSEP2) and local minimum (HYSEP3), as detailed in Pettyjohn and Henning (1979). For these three methods, an empirical relationship is used, which relates the catchment area to the number of days until baseflow makes up all streamflow, after a streamflow peak.

PART uses a form of streamflow partitioning based on antecedent streamflow recession (similar to the local minimum method of HYSEP), details of which are given in Rutledge (1998). The determination of the antecedent recession requirement in PART is done in three ways (see Rutledge, 1998) and thereby provides three baseflow estimates: PART1, PART2 and PART3.

The BFLOW program uses the Lyne and Hollick (1979) filter, which is a low-pass filter. This separation method uses signal processing theory, and is based on the hydrological reasoning that baseflow is the low frequency component of streamflow. The filter equation for baseflow is expressed as (Eckhardt, 2005):

$$b_t = \alpha b_{t-1} + \frac{1-\alpha}{2}(Q_t + Q_{t-1}) \text{ subject to } b_t \leqslant Q_t \qquad (1)$$

where b_t [L^3/T] is the baseflow at time step t [T], α [dimensionless] is the filter parameter and Q_t [L^3/T] is the streamflow at time step t. It is worth noting the constraint $b_t \leqslant Q_t$, which is required in applying (1) to avoid predictions of baseflow greater than streamflow (Chapman, 1991; Eckhardt, 2005). This constraint is discussed further in Section 4.3. The BFLOW program carries out three passes of the filter: forwards (BFLOW1), backwards (BFLOW2) and forwards again (BFLOW3) and uses a filter parameter $\alpha = 0.925$ as suggested by Nathan and McMahon (1990). Each pass of the filter acts to attenuate the baseflow signal. Despite having no physical basis, the baseflow separation of BFLOW has been found to agree well with manual separation techniques (Arnold and Allen, 1999).

The Eckhardt filter is a two-parameter filter based on the assumption that aquifer outflow is linearly proportional to storage. This filter limits the maximum ratio of baseflow to streamflow. Eckhardt (2005) describes this as potentially beneficial following the demonstration of Spongberg (2000) that runoff has a non-negligible low-frequency component. The equation for this filter is given by (Eckhardt, 2005):

$$b_t = \frac{(1-BFI_{\max})ab_{t-1} + (1-a)BFI_{\max}Q_t}{1-aBFI_{\max}} \text{ subject to } b_t \leqslant Q_t \qquad (2)$$

where a [dimensionless] is the baseflow recession constant and $BFI_{\max}$ [dimensionless] is the maximum value of the baseflow index.

As the $BFI_{\max}$ cannot be identified prior to separation, Eckhardt (2005) suggests using a value of 0.80 for perennial streams with porous aquifers, 0.50 for ephemeral streams with porous aquifers, and 0.25 for perennial streams with hard rock aquifers. The use of $BFI_{\max} = 0.50$ yields an equivalent filter to that proposed by Chapman (1991). In the formulation of Chapman (1991), a filter is developed to overcome baseflow being constant in the absence of direct runoff (similar to the Lyne and Hollick (1979) filter; Nathan and McMahon (1990)). This gives the filter parameter physical meaning in the form of the baseflow recession constant a. The recession constant for the Eckhardt (2005) filter is determined using the method outlined in Eckhardt (2008). This method involves plotting the flow Q_t against Q_{t-1} for periods where streamflow is decreasing for five consecutive days. A linear regression that passes through the origin is then calculated for these data points. The slope of this regression gives the recession constant a.

All streamflow data generated from the numerical model are translated to daily time-steps before being processed by HYSEP, PART, BFLOW and the Eckhardt filter. This translation is done in order to be compatible with the automated methods. The translation is carried out by calculating the average flow for each day. The influence of this constraint is discussed in Section 5.

3. Model simulations

To provide varied catchment responses and streamflow regimes, the synthetic catchment is subjected to varied rainfall and antecedent moisture conditions. These different conditions are used to examine the extent to which the HYSEP, BFLOW, PART and the Eckhardt separation methods reproduce and capture the HMC calculated baseflow signal. Two of the simulations are also subject to near-stream groundwater pumping. The simulations with groundwater pumping allow investigation into the common scenario of a modified catchment. Although some separation methods are specified for use in undisturbed catchments, they are still applied in this study to simulations with pumping. However, for this reason, the simulations with pumping are considered separately from the simulations without pumping. The baseflow separation methods are evaluated quantitatively and qualitatively

Table 2
Scenarios for simulating catchment response. Scenarios with an asterisk denote where groundwater pumping is applied in the catchment.

Scenario	Soil	Initial water table	Pumping
1	Sand	WT3	–
2*	Sand	WT3	Pump 1
3*	Sand	WT3	Pump 1and 2
4	Sand	WT2	–
5	Sand	WT1	–
6	Loamy sand	WT3	–
7	Loamy sand	WT2	–
8	Loamy sand	WT1	–

against the simulated baseflow using measures that account for total baseflow volume, as well as baseflow dynamics.

Initially, eight model scenarios are simulated (Table 2). The first three scenarios (1–3) consider the influence of pumping for the sandy catchment. Scenarios 4 and 5 consider different initial conditions as the starting points for scenarios 4 and 5, respectively. Scenarios 6, 7 and 8 consider a change in soil properties of the sand to loamy sand. As well as providing the controlled baseflow signal for evaluation of baseflow separation methods, the variation of aquifer properties provides insight into their influence on streamflow generation mechanisms.

The initial hydraulic heads and water table elevations are obtained by draining the fully saturated catchment for a period of 7, 8 and 9 months without applying any rainfall forcing or subsurface boundary recharge. The initial conditions used are: water Table 1 (WT1) = 7 months drainage, water Table 2 (WT2) = 8 months drainage and water Table 3 (WT3) = 9 months drainage. For each of these initial conditions the stream is still flowing at the end of the drainage period, whilst providing significantly different initial conditions.

The catchment's response is controlled by modifying the forcing functions (e.g. rainfall, pumping and ET) to scenarios with a series of different initial conditions. The rainfall varies in intensity and duration over three rain events throughout each of the simulations. The same rainfall boundary is applied in each of the eight scenarios as follows:

(1) 10 days without rainfall, then rainfall at a rate of 2 mm/h for 24 h followed by 10 days without rainfall; and
(2) rainfall at a rate of 4 mm/h for 48 h followed by 5 days without rainfall; and
(3) rainfall at a rate of 40 mm/h for 3 h followed by 30 days without rainfall (58 days and 3 h total)

The idealised rainfall events are uniform and constant with sufficiently large recovery periods such that the streamflow generation processes resulting from individual events can be clearly identified. The rainfall rates and durations are chosen to represent a range of streamflow generation mechanisms.

Pumping is applied in scenarios 2 and 3 at two locations (shown in Fig. 1): Pump 1 located at $(x, y, z) = (1550\text{ m}, 500\text{ m}, 0\text{ m})$, and pump 2 located at $(x, y, z) = (1550\text{ m}, 1500\text{ m}, 0\text{ m})$. The pumping rate is increased linearly from 0.00 to 0.01 m^3/s for pump 1 over the first day of simulation in scenario 2, with pump 2 inactive. In scenario 3, the pumping rate is increased linearly from 0.00 to 0.015 m^3/s for both pumps 1 and 2 over the first day of simulation. For scenarios 2 and 3, pumping is applied over the entire simulation. This pumping rate induces losing conditions locally along the stream near the pumping location. Pumping therefore allows the effect of varied flow regimes (i.e. gaining and losing sections) on streamflow generation to be explored with respect to baseflow.

Based on the initial simulations, evapotranspiration (ET) is applied to 5 additional scenarios (denoted as 9, 10, 11, 12 and 13; not listed in Table 3). The setup and forcing functions of scenarios 9, 10, 11, 12 are the same as for scenario 1, except that constant specified evaporation rates of 2, 5, 10 and 26 mm/day are applied, respectively. Some high evaporation rates (10 and 26 mm/day) are chosen to explore the influence that high ET in the catchment has on the baseflow separation methods. ET is also applied to scenario 13 with the same setup and forcing functions as for scenario 6, but with a constant specified evaporation rate of 5 mm/day. The simulations with ET are performed in order to examine the influence of ET on baseflow dynamics, baseflow recession and performance of separation methods against the simulated baseflow.

4. Results

4.1. Fully integrated model simulations

The simulated streamflow hydrograph at the outlet and the corresponding streamflow generation components (calculated from the HMC method) are shown for scenarios 1 (Fig. 2) and 2 (Fig. 3). The streamflow generation mechanisms varied in response

Table 3
BFI, *NSE* and *PBIAS* for simulated baseflow and estimated baseflow during event 1 using HYSEP, PART, BFLOW and the Eckhardt separation methods. Lightly shaded cells highlight a *NSE* < 0.5 and darkly shaded cells highlight |*PBIAS*| > 25%. Scenarios with an asterisk denote where groundwater pumping is applied in the catchment.

Scenario		Simulated	HYSEP1	HYSEP2	HYSEP3	PART1	PART2	PART3	BFLOW1	BFLOW2	BFLOW3	ECKHARDT
1	BFI	0.789	0.659	0.691	0.643	0.772	0.694	0.673	0.692	0.675	0.640	0.630
	NSE		0.607	0.698	0.591	0.872	0.742	0.692	0.751	0.722	0.577	0.567
	PBIAS		16.47%	12.40%	18.48%	2.12%	11.96%	14.66%	12.20%	14.36%	18.90%	20.15%
2*	BFI	0.815	0.666	0.706	0.646	0.798	0.711	0.683	0.693	0.679	0.623	0.653
	NSE		0.547	0.677	0.519	0.878	0.724	0.655	0.693	0.660	0.423	0.584
	PBIAS		18.30%	13.35%	20.68%	2.00%	12.74%	16.15%	14.93%	16.71%	23.52%	19.89%
3*	BFI	0.832	0.651	0.705	0.627	0.811	0.713	0.674	0.676	0.659	0.551	0.667
	NSE		0.426	0.635	0.372	0.890	0.686	0.574	0.591	0.537	−0.014	0.604
	PBIAS		21.73%	15.25%	24.62%	2.54%	14.25%	19.01%	18.73%	20.73%	33.81%	19.84%
4	BFI	0.878	0.717	0.763	0.699	0.850	0.770	0.737	0.735	0.731	0.694	0.675
	NSE		0.561	0.701	0.538	0.899	0.748	0.669	0.668	0.660	0.513	0.461
	PBIAS		18.34%	13.07%	20.32%	3.17%	12.23%	16.06%	16.26%	16.75%	21.00%	23.14%
5	BFI	0.929	0.753	0.803	0.735	0.902	0.814	0.772	0.762	0.762	0.727	0.703
	NSE		0.546	0.735	0.526	0.947	0.761	0.660	0.629	0.629	0.491	0.414
	PBIAS		18.87%	13.54%	20.82%	2.86%	12.40%	16.83%	17.91%	17.91%	21.71%	24.32%
6	BFI	0.286	0.255	0.265	0.250	0.267	0.267	0.260	0.333	0.270	0.250	0.286
	NSE		0.817	0.871	0.804	0.903	0.903	0.873	0.335	0.963	0.802	0.984
	PBIAS		10.86%	7.13%	12.63%	6.76%	6.76%	9.21%	−16.42%	5.42%	12.70%	−0.09%
7	BFI	0.310	0.272	0.283	0.266	0.285	0.285	0.277	0.348	0.288	0.266	0.302
	NSE		0.765	0.829	0.753	0.864	0.864	0.829	0.561	0.933	0.752	0.977
	PBIAS		12.35%	8.60%	14.14%	8.23%	8.23%	10.70%	−12.25%	7.11%	14.16%	2.60%
8	BFI	0.341	0.295	0.308	0.289	0.310	0.310	0.301	0.369	0.312	0.289	0.324
	NSE		0.732	0.803	0.720	0.839	0.839	0.800	0.721	0.906	0.716	0.955
	PBIAS		13.32%	9.53%	15.08%	9.11%	9.11%	11.64%	−8.37%	8.47%	15.23%	4.90%

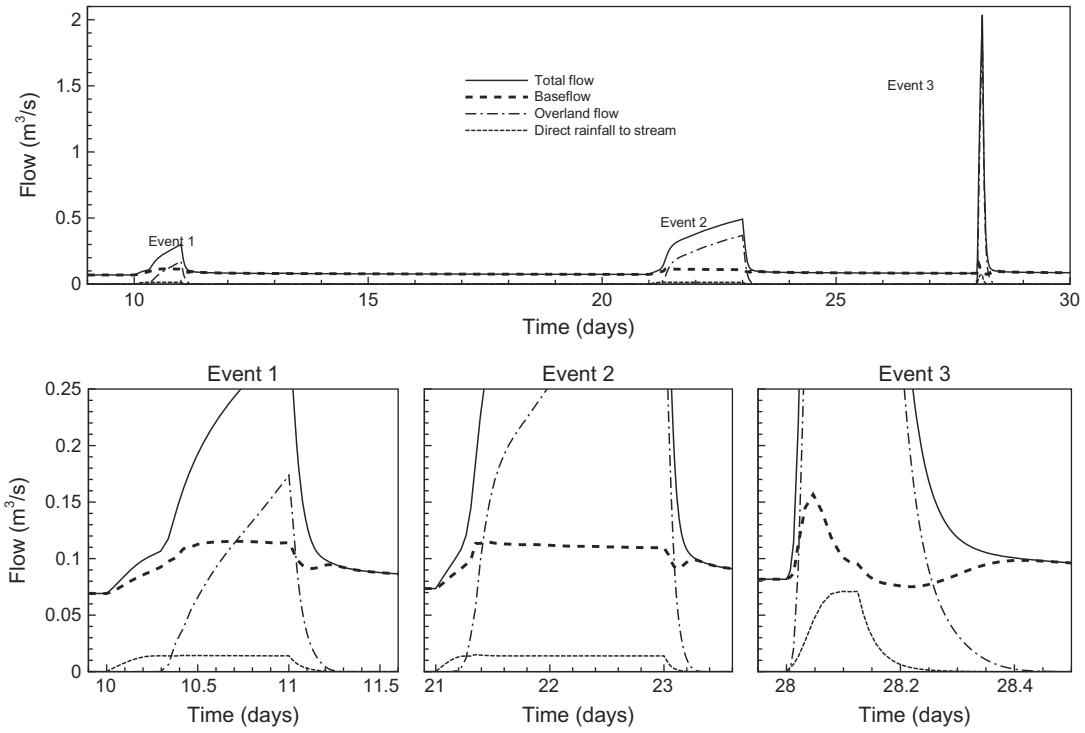

Fig. 2. Streamflow hydrograph at the outlet and HMC flow components for scenario 1 (without pumping), with highlighted events. An apparent steady-state baseflow rate was observed in the first event (10.5–11 days) and second event (21.4–23 days).

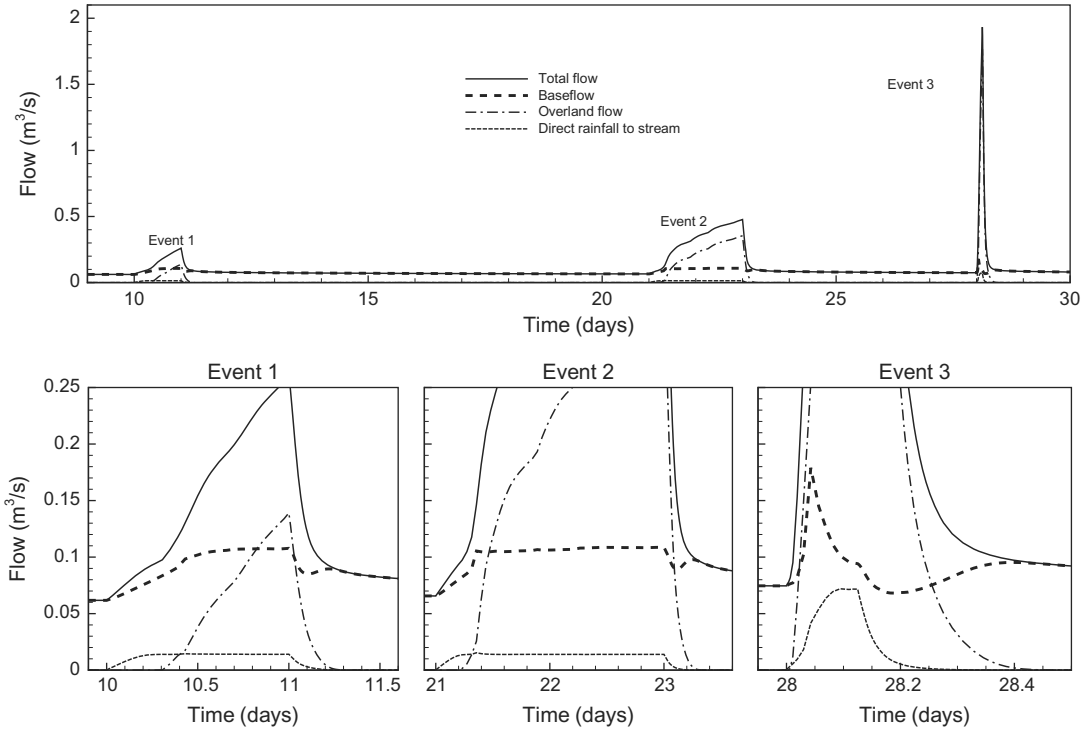

Fig. 3. Streamflow hydrograph at the outlet and HMC flow components for scenario 2 (with pumping), with highlighted events. An apparent steady-state baseflow rate was observed in the first event (10.5–11 days) and second event (21.4–23 days).

to different rainfall events. In all scenarios that were based on sandy material properties, streamflow was dominated by baseflow because the high infiltration capacity of the sand allowed for quick recharge. Consequently, there was only a small overland flow component due to saturation excess runoff. The almost horizontal slope of the catchment limited the vertical extent of the unsaturated zone to less than 1 m, resulting in a short delay between infiltration and recharge. As the timing between infiltration and recharge was short, there was a rapid response in the baseflow component of streamflow. As illustrated in Fig. 2, after a short and rapid initial increase, baseflow did not change significantly during the first two rainfall events and reached an apparent steady-state. As opposed

to baseflow, streamflow changed during this apparent steady-state. Therefore, the ratio of streamflow to baseflow changes as a function of time, but not consistently across events.

The apparent steady-state baseflow was only observed during the first two events. However, in the third event no apparent steady-state was reached. In this event, an initial rapid increase in head in the subsurface quickly increased the hydraulic gradient from the aquifer to the stream. During this rainfall event, the time delay from rainfall starting to the onset of overland flow is much slower than the time delay to the increase in groundwater discharge caused by the rapid aquifer response. After 1 h, the overland flow and accumulating direct rainfall to the stream increased the stream stage, thus reducing the hydraulic gradient between the aquifer and stream. This is a clear demonstration of the forcing functions controlling the baseflow dynamics.

The baseflow response from all three rainfall events did not follow the typical pattern of baseflow response as presented in standard textbooks, e.g. McCuen (2005) and Linsley et al. (1958). This is an important observation because these patterns are the basis for graphical approaches of baseflow separation. The pattern of baseflow during rainfall events obtained from the HGS model demonstrated a fast and transient response in stream-aquifer interaction. This was apparent at the beginning and cessation of the rainfall events, where an abrupt change in baseflow occurs, rather than a smooth and delayed response. The high transience of the stream aquifer interaction was also apparent in scenarios 2 and 3. The drawdown around the pump induced a loss in the adjacent stream, creating a variable flow regime with dynamic losing and gaining sections. The drawdown also increased the time between infiltration and recharge, further affecting the system dynamics.

The effect of ET (at a rate of 5 mm/day) is shown in Fig. 4 for the example of scenario 10. In comparison to scenario 1, the inclusion of ET in scenarios 9, 10, 11 and 12 slightly reduced event peaks and the baseflow component. These changes are due to the reduction in storage through losses from ET. However, it can be seen by comparison of Figs. 2 and 4, that the baseflow dynamics were very similar; the reduced overland flow component lead to a slightly higher proportion of baseflow with respect to streamflow.

4.2. Recession analysis

Following the approach of Eckhardt (2008), the recession periods were identified as periods in which streamflow was decreasing for 5 consecutive days. These periods were used to calculate the recession constant a (as defined in section 2.3). The slope for each linear regression of Q_{t+1} vs. Q_t passing through the origin was used as the recession constant a, which was then applied using the Eckhardt separation method. Fig. 5 shows the resulting recession constant a, and R^2 value obtained from the linear regression for each

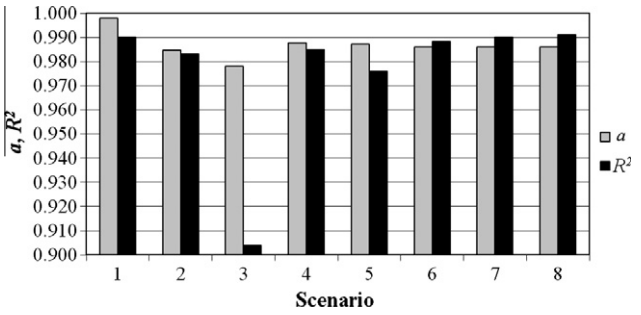

Fig. 5. Values of recession constant a and R^2 value for the linear regression of Q_t vs. Q_{t+1}, for sand and initial conditions WT1, WT2, WT3 and without/with pumps 1 and 2 active. The high value of R^2 suggests a linear storage–discharge relationship at the outlet during recession periods.

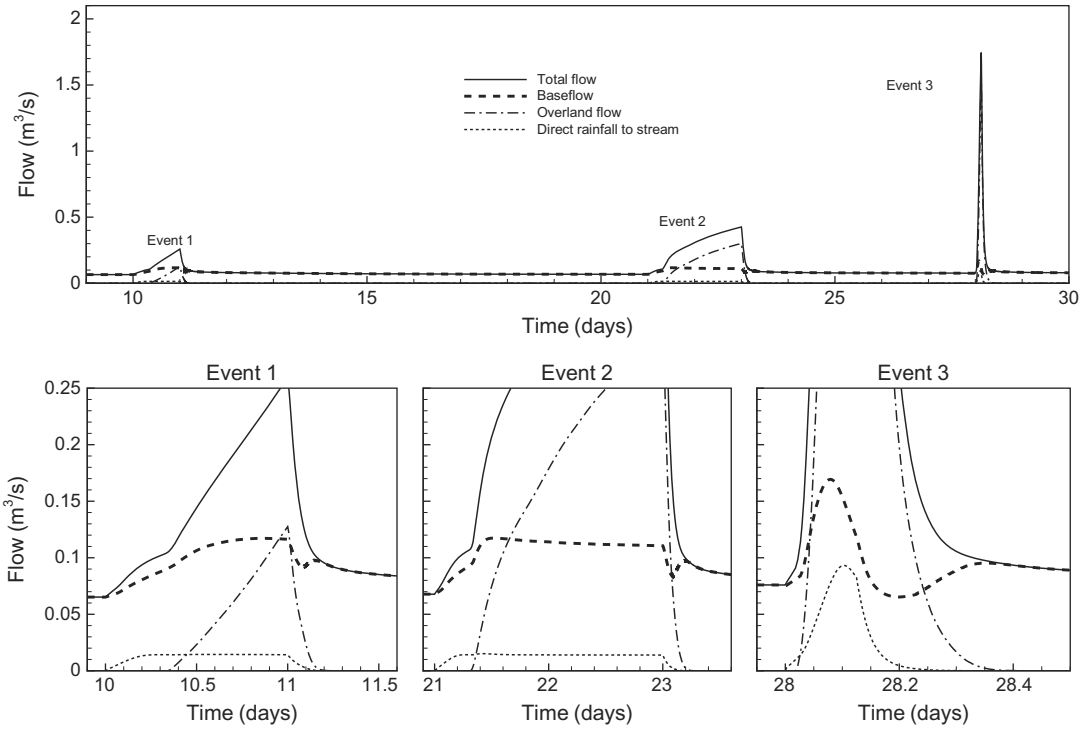

Fig. 4. Streamflow hydrograph at the outlet and HMC flow components for modified scenario 1 (with 5 mm/day ET), with highlighted events. An apparent steady-state baseflow rate was still observed in the first two rainfall events.

scenario. It can be seen in Fig. 5 that the R^2 for each regression was very close to 1. For all scenarios, this high R^2 value supports the assumption of a linear reservoir during recession periods, which is inherent in the Eckhardt method.

Adding ET in scenarios 9–13 reduced slightly the recession constant *a*, by less than 2% and it also slightly reduced the linearity of the storage–discharge relationship ($R^2 > 0.96$) of the catchment during recession periods. The linearity assumption for the storage–discharge relationship for this synthetic catchment was therefore still deemed reasonable.

4.3. Comparison of baseflow separation methods

All of the automated separation methods used are subject to the condition that baseflow cannot exceed streamflow. This is imposed because without this constraint, all of these methods can yield baseflow estimates above streamflow. By contrast, Furey and Gupta (2001) suggested that such a condition should not be imposed on physically-based methods. This way it is possible to identify time periods where estimated baseflow exceeds streamflow and diagnose these estimation errors. Steps can then be taken to modify the method while honouring physical processes so that these errors are reduced or fully removed. This constraint has repercussions for our analysis. In recession periods, the baseflow calculated from the automated separation methods is perfectly matched to the HMC calculated baseflow, because streamflow is entirely composed of baseflow. Therefore, an assessment of the differences between the simulated and approximated baseflow hydrographs is only meaningful during rainfall events. Consequently, this comparison is carried out during the rainfall events (i.e. 10–13 days, 21–25 days and 28–30 days) with the commonly used ratio of baseflow to streamflow, and two statistical measures, as follows:

(a) the baseflow index (*BFI*), which is the ratio of the total baseflow volume to the total streamflow volume over a given period:

$$BFI = \frac{V_{baseflow}}{V_{streamflow}} \tag{3}$$

where $V_{baseflow}$ [L^3] and $V_{streamflow}$ [L^3] are the total volume of baseflow and streamflow, respectively, over the simulation period.

(a) the Nash–Sutcliffe model efficiency (*NSE*) coefficient:

$$NSE = 1 - \frac{\sum_{t=1}^{T}(b_o^t - b_m^t)^2}{\sum_{t=1}^{T}(b_o^t - \bar{b}_o)^2} \tag{4}$$

where b_o^t [L^3/T] is the HMC calculated baseflow at time step t [T], b_m^t [L^3/T] is the baseflow from automated separation at time step t [T], and $\bar{b}_o$ [L^3/T] is the mean HMC calculated baseflow. The *NSE* provides a measure with values ranging from $-\infty$ to 1 (where 1 indicates a perfect match), of how well the separation methods compare to the HMC calculated baseflow. Moriasi et al. (2007) suggest that a $NSE > 0.5$ is satisfactory.

(a) the percent bias (*PBIAS*) (Gupta et al., 1999):

$$PBIAS = \frac{\sum_{t=1}^{T}(b_o^t - b_m^t) \times 100}{\sum_{t=1}^{T}(b_o^t)} \tag{5}$$

The *PBIAS* provides a measure of over or underestimation for each event, with an optimal value of 0%. Positive values of *PBIAS* indicate an underestimation, whereas negative values indicate overestimation of baseflow. Moriasi et al. (2007) suggest that a *PBIAS* of up to ±25% for streamflow is satisfactory, and this is used to guide acceptable *PBIAS* results in this study.

The baseflow index (*BFI*), Nash–Sutcliffe efficiency (*NSE*) and percent bias (*PBIAS*) are given for each of the three rainfall events in Tables 3–5, respectively, for scenarios 1–8. The values for $NSE < 0.5$ and $|PBIAS| > 25\%$ are highlighted. The *BFI*s calculated across the entire simulation for each scenario are ranked for each separation method in Fig. 6. Ranking is in order of best to worst *BFI* relative to the *BFI* observed from the HGS model (based on the HMC calculated baseflow). The results for the testing of the inclusion of ET for scenario 1 are shown in Table 6.

Table 4
BFI, *NSE* and *PBIAS* for simulated baseflow and estimated baseflow during event 2 using HYSEP, PART, BFLOW and the Eckhardt separation methods. Lightly shaded cells highlight a $NSE < 0.5$ and darkly shaded cells highlight $|PBIAS| > 25\%$. An asterisk denotes scenarios where groundwater pumping was applied in the catchment.

Scenario		Simulated	HYSEP1	HYSEP2	HYSEP3	PART1	PART2	PART3	BFLOW1	BFLOW2	BFLOW3	ECKHARDT
1	BFI	0.486	0.531	0.453	0.394	0.415	0.415	0.404	0.470	0.418	0.379	0.409
	NSE		0.741	0.707	0.337	0.501	0.501	0.422	0.693	0.616	0.177	0.504
	PBIAS		−9.45%	6.64%	18.83%	14.51%	14.51%	16.82%	3.09%	13.93%	21.93%	15.85%
2*	BFI	0.499	0.549	0.459	0.386	0.415	0.415	0.398	0.463	0.410	0.365	0.434
	NSE		0.702	0.696	0.140	0.399	0.399	0.268	0.677	0.461	−0.106	0.635
	PBIAS		−10.07%	8.03%	22.68%	16.75%	16.75%	20.12%	7.13%	17.71%	26.89%	12.94%
3*	BFI	0.527	0.588	0.474	0.373	0.419	0.419	0.393	0.452	0.400	0.336	0.474
	NSE		0.621	0.683	−0.186	0.244	0.244	0.035	0.532	0.202	−0.645	0.764
	PBIAS		−11.62%	9.97%	29.22%	20.52%	20.52%	25.36%	14.20%	23.97%	36.16%	9.98%
4	BFI	0.559	0.586	0.506	0.442	0.467	0.467	0.454	0.510	0.467	0.419	0.437
	NSE		0.877	0.692	0.222	0.402	0.402	0.314	0.622	0.488	−0.002	0.228
	PBIAS		−4.86%	9.38%	20.88%	16.33%	16.33%	18.80%	8.76%	16.37%	25.03%	21.77%
5	BFI	0.635	0.650	0.566	0.492	0.524	0.524	0.507	0.552	0.519	0.458	0.461
	NSE		0.907	0.687	0.137	0.342	0.342	0.246	0.508	0.397	−0.162	−0.077
	PBIAS		−2.40%	10.94%	22.54%	17.54%	17.54%	20.15%	13.14%	18.23%	27.84%	27.45%
6	BFI	0.163	0.282	0.197	0.148	0.155	0.155	0.152	0.248	0.170	0.144	0.200
	NSE		−12.697	−3.793	0.835	0.922	0.922	0.890	−7.843	0.916	0.767	−0.718
	PBIAS		−73.36%	−21.35%	8.99%	4.59%	4.59%	6.81%	−52.42%	−4.42%	11.23%	−23.21%
7	BFI	0.174	0.285	0.203	0.155	0.163	0.163	0.159	0.253	0.177	0.151	0.208
	NSE		−9.248	−2.599	0.763	0.868	0.868	0.826	−5.774	0.973	0.678	−0.315
	PBIAS		−64.03%	−17.11%	10.77%	6.27%	6.27%	8.56%	−45.59%	−2.02%	13.16%	−19.75%
8	BFI	0.186	0.289	0.211	0.163	0.171	0.171	0.167	0.259	0.186	0.158	0.217
	NSE		−6.686	−1.690	0.700	0.820	0.820	0.770	−4.173	0.994	0.605	−0.012
	PBIAS		−55.91%	−13.52%	12.14%	7.57%	7.57%	9.92%	−39.58%	−0.04%	14.60%	−16.77%

Table 5
BFI, *NSE* and *PBIAS* for simulated baseflow and estimated baseflow during event 3 using HYSEP, PART, BFLOW and the Eckhardt separation methods. Lightly shaded cells highlight a *NSE* < 0.5 and darkly shaded cells highlight |*PBIAS*| > 25%. An asterisk denotes scenarios where groundwater pumping was applied in the catchment.

Scenario		Simulated	HYSEP1	HYSEP2	HYSEP3	PART1	PART2	PART3	BFLOW1	BFLOW2	BFLOW3	ECKHARDT
1	BFI	0.567	0.530	0.537	0.530	0.548	0.537	0.533	0.559	0.519	0.492	0.557
	NSE		0.972	0.976	0.972	0.987	0.979	0.975	0.998	0.959	0.920	0.996
	PBIAS		6.46%	5.24%	6.46%	3.35%	5.25%	5.96%	1.37%	8.53%	13.24%	1.80%
2*	BFI	0.581	0.530	0.539	0.530	0.556	0.539	0.534	0.565	0.513	0.481	0.572
	NSE		0.949	0.957	0.949	0.977	0.962	0.954	0.990	0.925	0.863	0.997
	PBIAS		8.81%	7.27%	8.81%	4.42%	7.28%	8.24%	2.89%	11.70%	17.31%	1.69%
3*	BFI	0.596	0.527	0.539	0.527	0.561	0.539	0.531	0.567	0.492	0.446	0.592
	NSE		0.912	0.926	0.912	0.960	0.934	0.920	0.971	0.844	0.721	0.999
	PBIAS		11.68%	9.64%	11.68%	5.90%	9.68%	11.02%	5.02%	17.53%	25.24%	0.70%
4	BFI	0.620	0.577	0.584	0.577	0.596	0.584	0.580	0.605	0.562	0.524	0.571
	NSE		0.967	0.971	0.967	0.982	0.974	0.970	0.993	0.950	0.892	0.970
	PBIAS		6.87%	5.78%	6.87%	3.85%	5.79%	6.44%	2.40%	9.42%	15.51%	7.93%
5	BFI	0.684	0.636	0.644	0.636	0.657	0.644	0.639	0.663	0.612	0.562	0.583
	NSE		0.966	0.970	0.966	0.981	0.973	0.969	0.989	0.942	0.863	0.907
	PBIAS		6.98%	5.85%	6.98%	3.97%	5.86%	6.50%	2.99%	10.42%	17.72%	14.73%
6	BFI	0.257	0.251	0.254	0.251	0.258	0.254	0.252	0.283	0.247	0.236	0.275
	NSE		0.996	0.999	0.996	1.000	0.999	0.998	0.877	0.989	0.974	0.941
	PBIAS		2.32%	1.05%	2.32%	−0.51%	1.07%	1.71%	−10.46%	3.78%	7.82%	−7.39%
7	BFI	0.267	0.259	0.262	0.259	0.267	0.262	0.260	0.292	0.255	0.243	0.284
	NSE		0.994	0.997	0.994	1.000	0.998	0.996	0.904	0.987	0.967	0.954
	PBIAS		3.08%	1.76%	3.08%	0.12%	1.77%	2.43%	−9.27%	4.44%	8.82%	−6.49%
8	BFI	0.278	0.267	0.271	0.267	0.276	0.271	0.269	0.300	0.264	0.251	0.294
	NSE		0.991	0.995	0.991	0.999	0.996	0.994	0.925	0.984	0.960	0.964
	PBIAS		3.77%	2.41%	3.77%	0.65%	2.42%	3.10%	−8.16%	5.02%	9.65%	−5.70%

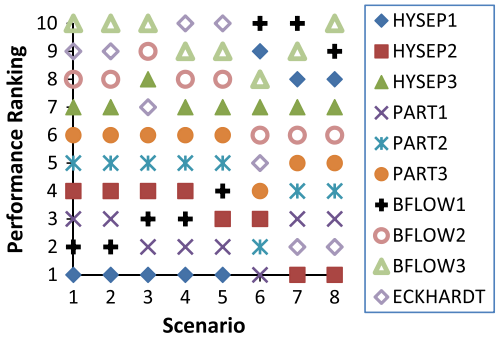

Fig. 6. Performance based ranking using BFI over the whole simulation for HYSEP, PART, BFLOW and the Eckhardt separation methods. 1 indicates best performance, 10 indicates worst performance.

For event 1, the *NSE*s were satisfactory over all scenarios for HYSEP2, PART1, PART2, PART3, BFLOW1, BFLOW2 and the Eckhardt separation methods. However, in scenario 3, HYSEP1, HYSEP3 and BFLOW3 had a *NSE* less than 0.5, with BFLOW3 showing a very unsatisfactory performance indicated by a negative *NSE*. The BFLOW1 separation had a single instance of *NSE* less than 0.5 and the Eckhardt separation had two instances of *NSE* less than 0.5. However, these were only slightly below this value, indicating that the performance of these methods was almost satisfactory. The *PBIAS* was at a maximum of 33.8% for event 1 of scenario 3 for the BFLOW3 separation method, showing a large underestimation of the HMC calculated baseflow. All separation methods for event 1 tended to underestimate baseflow. Only BFLOW1 overestimated baseflow, which occurred for the sandy loam in scenarios 6–8.

Table 6
Comparison of *BFI*, *NSE* and *PBIAS* for (scenario 1 with and without ET) simulated baseflow and estimated baseflow during event 2 using HYSEP, PART, BFLOW and the Eckhardt separation methods. Lightly shaded cells highlight a *NSE* < 0.5 and darkly shaded cells highlight |*PBIAS*| > 25%.

ET (mm/day)		Simulated	HYSEP1	HYSEP2	HYSEP3	PART1	PART2	PART3	BFLOW1	BFLOW2	BFLOW3	ECKHARDT
0	BFI	0.486	0.531	0.453	0.394	0.415	0.415	0.404	0.470	0.418	0.379	0.373
	N-S		0.741	0.707	0.337	0.501	0.501	0.422	0.693	0.616	0.177	0.130
	PBIAS		−9.45%	6.64%	18.83%	14.51%	14.51%	16.82%	3.09%	13.93%	21.93%	23.19%
2	BFI	0.513	0.539	0.465	0.408	0.428	0.428	0.417	0.482	0.429	0.393	0.436
	N-S		0.868	0.673	0.235	0.390	0.390	0.313	0.629	0.494	0.094	0.509
	PBIAS		−5.13%	9.40%	20.49%	16.48%	16.48%	18.67%	6.04%	16.28%	23.28%	14.94%
5	BFI	0.553	0.568	0.490	0.430	0.453	0.453	0.440	0.500	0.448	0.414	0.458
	N-S		0.896	0.642	0.144	0.305	0.305	0.224	0.560	0.366	−0.007	0.404
	PBIAS		−2.79%	11.24%	22.18%	18.07%	18.07%	20.31%	9.53%	18.85%	25.11%	17.19%
10	BFI	0.703	0.699	0.606	0.508	0.562	0.562	0.533	0.575	0.523	0.491	0.529
	N-S		0.817	0.722	−0.117	0.225	0.225	0.061	0.315	0.050	−0.253	0.133
	PBIAS		0.52%	13.74%	27.68%	20.02%	20.02%	24.19%	18.08%	25.54%	30.08%	24.67%
26	BFI	0.852	0.790	0.696	0.591	0.654	0.654	0.618	0.633	0.597	0.563	0.597
	N-S		0.705	0.602	−0.202	0.113	0.113	−0.054	0.035	−0.117	−0.387	−0.088
	PBIAS		7.34%	18.27%	30.60%	23.28%	23.28%	27.54%	25.75%	29.91%	33.97%	29.96%

For event 2, the *NSE*s were below 0.5 for each separation method in at least one of the eight scenarios. The BFLOW3 separation showed very poor performance with negative *NSE*s in scenarios 2–5. In each of these scenarios, BFLOW3 had a *PBIAS* showing underestimation of baseflow by more than 25%. The HYSEP1, HYSEP2, BFLOW1 and Eckhardt methods showed poor performance for sandy loam (scenarios 6–8) with negative *NSE*s in each scenario. The *PBIAS* for HYSEP1 and BFLOW1 separation showed overestimation of baseflow ranging from 40% to 73% in the sandy loam scenarios. It is interesting to note that the scenarios in which HYSEP1, HYSEP2 and BFLOW1 performed well, HYSEP3, PART1, PART2, PART3, BFLOW2 and BFLOW3 performed poorly and vice versa. More interestingly, where the HYSEP1, HYSEP2 and BFLOW1 methods performed poorly, these methods largely overestimated baseflow whereas where the HYSEP3, PART1, PART2, PART3, BFLOW2 and BFLOW3 methods performed poorly, these particular methods largely underestimated baseflow.

For event 3, the *NSE* was greater than 0.5 for every method in each scenario, with values close to 1. The largest *PBIAS* was for BFLOW3 in scenario 3, showing underestimation of baseflow by just over 25%.

The inclusion of ET in scenarios 9–13 showed that as the ET rate was increased, the *BFI* increased. ET also lead to a reduction in performance (both *NSE* and *PBIAS*) for every separation method, except HYSEP1.

The rankings of separation methods (shown in Fig. 6) based on *BFI* over the entire simulation provide a summary of the performance of each of the separation methods. The best replication of *BFI* resulted from the HYSEP1 method in scenarios 1–5, from PART1 in scenario 6, and from HYSEP2 in scenarios 7–8. The BFLOW3 method was worst in scenarios 1–3 and 8. The Eckhardt method performed worst in scenarios 4 and 5.

The baseflow separations from the streamflow hydrograph in scenario 1 obtained using HYSEP, PART, BFLOW and the Eckhardt separation methods are shown in Fig. 7. Visual inspection of baseflow curves in Fig. 7 shows that the ability of these separation methods to match the simulated baseflow was poor in almost all cases, despite the fact that they had satisfactory *NSE* and *PBIAS* values. It is clear that despite reasonable estimates of the *BFI* for each scenario, the dynamics of baseflow during rainfall were missed.

5. Discussion

The baseflow hydrographs obtained using HGS (with the HMC method) were used as a control experiment to test the performance of a series of automated baseflow separation methods. The initial conditions (antecedent moisture), forcing functions (rainfall patterns, pumping and ET) as well as the physical properties (soil properties) of the catchment were varied across simulation scenarios. The varied conditions across the different model scenarios allowed the generation of unique baseflow behaviour, controlled by a range of hydrological processes. The application of the HMC method allowed quantification of the relative importance of hydrological processes to the streamflow hydrograph. While the structure and geology of the synthetic model used in this study were simple, the hydrological processes considered were simulated in a physically based way. Despite the simplified nature of the catchment, the baseflow separation methods consistently failed to perform satisfactorily. This is easily attributed the variability of the baseflow dynamics as observed across all simulation scenarios. Increasing the complexity of the catchment (e.g. heterogeneous geology, more realistic topography and rainfall patterns), is likely to lead to an even more complex baseflow response. With increased complexity, it is expected that the variability seen in baseflow dynamics will remain and hence that the simple automated baseflow separation methods examined will not perform any better in estimating baseflow.

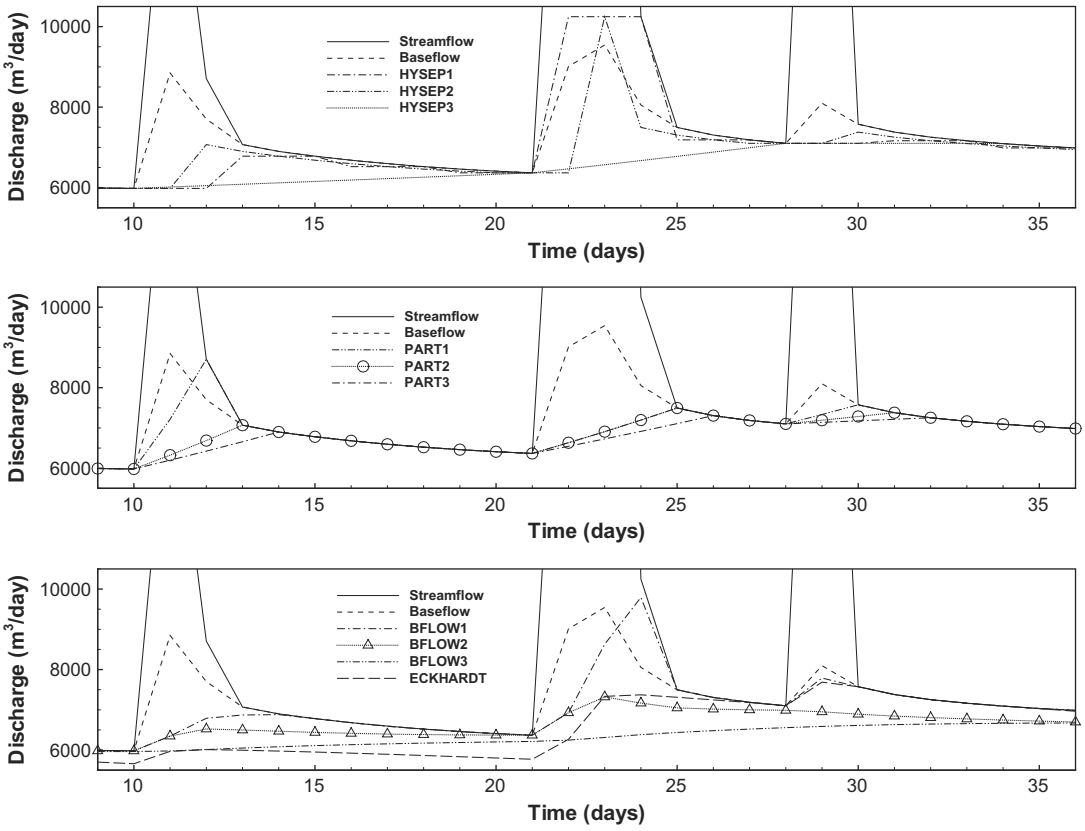

Fig. 7. Comparison of simulated daily baseflow and baseflow estimated using HYSEP, PART, BFLOW and the Eckhardt separation methods for scenario 1.

An initial analysis of the hydrographs revealed that the behaviour of baseflow was fundamentally different between rainfall events. For the first two events, baseflow remained constant and reached an apparent steady-state despite the changing forcing functions. In contrast, baseflow dynamics during the third event were highly transient. This illustrates that the controlling processes of baseflow are not always static, but instead change in response to different forcing functions. It also challenges the common assumption (based on hydrological reasoning) of the simple automated baseflow separation methods, that a simple fixed relation between baseflow and streamflow exists, for all rainfall events and antecedent moisture conditions. It is observed that this is not necessarily the case.

In the synthetic catchment used in this study, an apparent steady-state of baseflow discharge was reached for certain rain events. Once this apparent steady-state was reached, streamflow only increased with increasing overland flow. This led to a baseflow response that is dependent upon the rate and duration of rainfall. For all three events, the relationship between baseflow and streamflow was not consistent, as assumed by the BFLOW1, BFLOW2, BFLOW3 and Eckhardt separation methods. The variation of the ratio of baseflow to streamflow was observed in response to different rainfall events, as well as for the different initial antecedent moisture conditions. The variation observed in these simulations highlights an inability to accurately capture the average baseflow with the various baseflow separation methods examined. Moreover, it demonstrates the often acknowledged, but seldom addressed ambiguity of the separation methods used. The results from this numerical experiment suggest that quantifying baseflow in catchments with non-stationary processes, such as varied climatic conditions that are outside of seasonality, will alter the streamflow generation mechanisms and hence the *BFI*.

The *NSE*s for rainfall events 1 and 3 for each of the scenarios indicated satisfactory results for the baseflow separations. The agreement between the simulated baseflow method and separation methods for event 3 was significantly better than for the other two events. For rainfall event 2, the *NSE*s showed a poor match between the automated separation methods and HMC calculated baseflow for different methods in each scenario, with both large overestimation and underestimation of the HMC calculated baseflow in some cases. This variability in each of the methods' ability to match the HMC calculated baseflow across scenarios and corresponding rainfall events highlights that no single separation method performed consistently well in the control experiment. The *BFI* based rankings of the separation methods show that, on average, PART1 and HYSEP2 performed best overall in capturing the baseflow volume across the eight scenarios.

One of the limitations found in the use of the automated separation methods was the constraint of using daily streamflow data. It can be seen from the results in Figs. 2 and 3 that the behaviour of the baseflow varied on at least an hourly time scale, much smaller than could be captured in a daily time step. However, this is only a limitation when it is essential to accurately capture baseflow behaviour at finer timescales (e.g. flood modelling). In the context of low flow hydrology, where estimates of annual baseflow contributions are required, the nuances seen in the baseflow behaviour during rainfall events is not important as long as the average baseflow is captured. However, it is possible that the nuances seen in the hourly time step of this catchment present themselves in a larger catchment at the daily time step, in which case use of these filters would be problematic and would fail to capture even the average behaviour.

The automated separation methods result in the largest difference during rainfall events in which recharge is also occurring. This means that any perennial streams that are subject to significant and extended rainfall periods will be the most difficult to accurately determine the *BFI* for. This is because the proportion of time that streamflow is not driven purely by baseflow affects the relative magnitude of the potential baseflow error.

6. Conclusions

Whilst commonly used automated baseflow separation methods are known to be somewhat ambiguous and arbitrary, the potential errors have not been quantified previously using a 3D fully integrated physically based flow model. The numerical experiments in this study strongly suggested that baseflow dynamics are complex, even in a simple catchment. The complexity of baseflow dynamics was seen to affect the performance of the simple automated separation methods. The frequently used automated baseflow separation methods could not perform satisfactorily across all events and scenarios considered. This suggests that caution should be used when applying these methods, depending on the flow dynamics of the catchment being studied. Unfortunately, there are no clear indicators as to which separation methods are most and least appropriate under particular conditions, which is not surprising given the absence of a true physical basis in the simple methods examined. This is cause for concern because baseflow separation is an important tool influencing decisions and outcomes of the various applications it is used for, such as the analysis of event runoff; recharge estimation; low flow forecasting; hydrogeologic parameter estimation; hydrologic model calibration; and the identification of source areas; and dominant processes producing runoff (Schwartz (2007)). Large errors will undermine the many applications baseflow separation is used for.

Further work is required to understand the appropriate use of baseflow separation methods. More complex baseflow separation methods than those considered in this study should be tested in future studies. Physically based filters (e.g. Furey and Gupta, 2001, 2003; Huyck et al., 2005) could prove to be more robust. This is because they provide a physically based relation of rainfall and ET (and other physical parameters) to baseflow. However, such methods clearly require more data (e.g. rainfall time series) which may not be readily available. It is perhaps the case that the uncertainty associated with simple automated methods precludes their use for providing anything more than very rough estimates of baseflow.

An improved understanding of baseflow dynamics is required for a broader range of catchments. With respect to baseflow dynamics, future studies should aim to elucidate: (a) scale dependence of baseflow generation to test if the baseflow response seen in the hourly time step of this small synthetic catchment occurs at the daily time step for larger catchments (e.g. >20 km^2); (b) testing how the *BFI* varies as a result of non-stationary processes; (c) testing of the impact of variations in geology, topography and vegetation, by incrementally adding layers of complexity to similar models in order to try and understand baseflow dynamics. Further investigation within numerical models should play a role in establishing physically based recommendations as to the appropriateness of commonly used automated baseflow separation methods in different catchment types and settings. Given the reality of the physical interpretations and subsequent calculations that such simple automated separation methods are used to support, there is a need to establish either stricter guidelines for such methods, develop improved methods (e.g. physically based methods) or at least provide error bounds on such estimates.

Acknowledgements

The authors gratefully acknowledge the reviewers' comments, which greatly improved the final manuscript. This work is supported by the Australian Research Council through its Linkage

grant scheme and the South Australian Department for Water as industry partners under grant number LP0668808. Part of this research was funded by the Swiss National Foundation, Ambizione grant PZ00P2_126415.

References

Arnold, J.G., Allen, P.M., Muttiah, R., Bernhardt, G., 1995. Automated base-flow separation and recession analysis techniques. Ground Water 33, 1010–1018.

Arnold, J.G., Allen, P.M., 1999. Automated methods for estimating baseflow and ground water recharge from streamflow records. J. Am. Water Resour. Assoc. 35, 411–424.

Barnes, B., 1939. The structure of discharge-recession curves. Trans. Am. Geophys. Union 20, 721–725.

Becker, M.W., Gerogian, T., Ambrose, H., Siniscalchi, J., Frederick, K., 2004. Estimating flow and flux of ground water discharge using water temperature and velocity. J. Hydrol. 296, 221–233.

Brunner, P., Simmons, C.T., 2011. HydeoGeoSphere: a fully integrated physically based hydrological model. Ground Water. http://dx.doi.org/10.1111/j.1745-6584.2011.00882.x.

Boughton, W.C. 1993. A hydrograph-based model for estimating the water yield of ungauged catchments. In: Paper presented at Hydrology and Water Resources Symposium, Inst. of Eng. Aust., Newcastle, NSW.

Brutsaert, W., Nieber, J.L., 1977. Regionalized drought flow hydrographs from a mature glaciated plateau. Water Resour. Res. 13, 637–644.

Chapman, T., 1991. Evaluation of automated techniques for base-flow and recession analyses – comment. Water Resour. Res. 27, 1783–1784.

Chapman, T., 1999. A comparison of algorithms for stream flow recession and baseflow separation. Hydrol. Process. 13, 701–714.

Cook, P.G., Favreau, G., Dighton, J.C., Tickell, S., 2003. Determining natural groundwater influx to a tropical river using radon, chlorofluorocarbons and ionic environmental tracers. J. Hydrol. 277, 74–88.

Cook, P.G., Wood, C., White, T., Simmons, C.T., Fass, T., Brunner, P., 2008. Groundwater inflow to a shallow, poorly-mixed wetland estimated from a mass balance of radon. J. Hydrol. 354, 213–226.

Eckhardt, K., 2005. How to construct recursive digital filters for baseflow separation. Hydrol. Process. 19, 507–515.

Eckhardt, K., 2008. A comparison of baseflow indices, which were calculated with seven different baseflow separation methods. J. Hydrol. 352, 168–173.

Fenicia, F., Savenije, H.H.G., Matgen, P., Pfister, L., 2006. Is the groundwater reservoir linear? Learning from data in hydrological modelling. Hydrol. Earth Syst. Sci. 10, 139–150.

Ferket, B.V.A., Samain, B., Pauwels, V.R.N., 2010. Internal validation of conceptual rainfall–runoff models using baseflow separation. J. Hydrol. 381, 158–173.

Freeze, R.A., 1972. Role of subsurface flow in generating surface runoff 1. Base flow contributions to channel flow. Water Resour. Res. 8, 609–623.

Furey, P.R., Gupta, V.K., 2001. A physically based filter for separating base flow from streamflow time series. Water Resour. Res. 37 (11), 2709–2722.

Furey, P.R., Gupta, V.K., 2003. Tests of two physically based filters for base flow separation. Water Resour. Res. 39 (10), 1297.

Gaukroger, A.M., Werner, A.D., 2011. On the Panday and Huyakorn surface–subsurface hydrology test case: analysis of internal flow dynamics. Hydrol. Process. 25, 2085–2093.

Gupta, H.V., Sorooshian, S., Yapo, P.O., 1999. Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration. J. Hydrol. Eng. 4, 135–143.

Halford, K.J., Mayer, G.C., 2000. Problems associated with estimating ground water discharge and recharge from stream-discharge records. Ground Water. 38, 331–342.

Hall, F.R., 1968. Base-flow recessions – a review. Water Resour. Res. 4, 973–983.

Huyck, A.A.O., Pauwels, V.R.N., Verhoest, N.E.C., 2005. A base flow separation algorithm based on the linearized Boussinesq equation for complex hillslopes. Water Resour. Res., 41.

HydroGeoLogic, 2006. MODHMS: a comprehensive MODFLOW based hydrologic modeling system, Version 3.0, Code documentation and user's guide. HydroGeoLogic Inc., Herndon, VA, 2006.

free-surface overland flow boundary condition in a parallel groundwater flow model. Adv. Water Resour. 29, 945–958.

Linsley, R.K., Kohler, M.A., Paulhus, J.L.H., Wallace, J.S., 1958. Hydrology for Engineers. McGraw Hill, New York.

Lyne, V., Hollick, M. 1979. Stochastic time-variable rainfall–runoff modelling. In: Hydrology and Water Resources Symposium. Institution of Engineers, Australia, Perth, pp. 89–92.

Maillet, E., 1905. Essai d'Hydraulique Souterraine et Fluviale. Hermann, Paris, 218.

Mau, D.P., Winter, T.C., 1997. Estimating ground-water recharge from streamflow hydrographs for a small mountain watershed in a temperate humid climate, New Hampshire, USA.. Ground Water. 35, 291–304.

McCallum, J.L., Cook, P.G., Brunner, P., Berhane, D., 2010. Solute dynamics during bank storage flows and implications for chemical base flow separation. Water Resour. Res., 46.

McCuen, R.H., 2005. Hydrologic Analysis and Design. Prentice Hall.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. Am. Soc. Agric. Biol. Eng. 50, 885–900.

Nathan, R.J., McMahon, T.A., 1990. Evaluation of automated techniques for base-flow and recession analyses. Water Resour. Res. 26, 1465–1473.

Panday, S., Huyakorn, P.S., 2004. A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow. Adv. Water Resour. 27, 361–382.

Park, Y.J., Sudicky, E.A., Panday, S., Matanga, G., 2009. Implicit subtime stepping for solving nonlinear flow equations in an integrated surface–subsurface system. Vadose Zone J. 8, 825–836.

Partington, D., Brunner, P., Simmons, C.T., Therrien, R., Werner, A.D., Dandy, G.C., Maier, H.R., 2011. A hydraulic mixing-cell method to quantify the groundwater component of streamflow within spatially distributed fully integrated surface water – groundwater flow models. Environ. Model. Softw. 26, 886–898.

Pettyjohn, W.A., Henning, R. 1979. Preliminary estimate of ground-water recharge rates, related streamflow and water quality in Ohio. Ohio State University Water Resources Center Project Completion Report Number 552.

Piggott, A.R., Moin, S., Southam, C., 2005. A revised approach to the UKIH method for the calculation of baseflow. Hydrol. Sci. J.–J. Sci. Hydrol. 50, 911–920.

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 data. US Geological Survey Water-Resources Investigations, Report 98–4148.

Schwartz, S.S., 2007. Automated algorithms for heuristic base-flow separation. J. Am. Water Resour. Assoc. 43, 1583–1594.

Singh, K.P., Stall, J.B., 1971. Derivation of base flow recession curves and parameters. Water Resour. Res. 7, 292–303.

Sloto, R., Crouse, M., 1996. HYSEP: A computer program for streamflow hydrograph separation and analysis. US Geological Survey, Water Resources Investigations, Report 96-4040.

Smakhtin, V.U., 2001. Low flow hydrology: a review. J. Hydrol. 240, 147–186.

Spongberg, M.E., 2000. Spectral analysis of base flow separation with digital filters. Water Resour. Res. 36 (3), 745–752.

Szilagyi, J., 2004. Heuristic continuous base flow separation. J. Hydrol. Eng. 9, 311–318.

Tallaksen, L.M., 1995. A review of baseflow recession analysis. J. Hydrol. 165, 349–370.

Therrien, R., McLaren, R.G., Sudicky, E.A. Panday, S., 2009. HydroGeoSphere. A Three-dimensional Numerical Model Describing Fully-integrated Subsurface and Surface Flow and Solute Transport. <hydrogeosphere.org>.

Werner, A.D., Gallagher, M.R., Weeks, S.W., 2006. Regional-scale, fully coupled modelling of stream-aquifer interaction in a tropical catchment. J. Hydrol. 328, 497–510.

Wittenberg, H. 1994. Nonlinear Analysis of Flow Recession Curves. Pages 61–67 in Friend: Flow Regimes from International Experimental and Network Data. International Association Hydrological Sciences, Wallingford.

Wittenberg, H., Sivapalan, M., 1999. Watershed groundwater balance estimation using streamflow recession analysis and baseflow separation. J. Hydrol. 219, 20–33.

Wittenberg, H., 2003. Effects of season and man-made changes on baseflow and flow recession: case studies. Hydrol. Process. 17, 2113–2123.