Author/s:
Su, CH;Costelloe, JF;Peterson, TJ;Western, AW

Title:
On the structural limitations of recursive digital filters for base flow estimation

Date:
2016-06-01

Citation:
Su, C. H., Costelloe, J. F., Peterson, T. J. & Western, A. W. (2016). On the structural limitations of recursive digital filters for base flow estimation. Water Resources Research, 52 (6), pp.4745-4764. https://doi.org/10.1002/2015WR018067.

Persistent Link:
https://hdl.handle.net/11343/291375

On the structural limitations of recursive digital filters for baseflow estimation

Chun-Hsu Su,[1] Justin F. Costelloe,[1] Tim J. Peterson,[1] Andrew W. Western[1]

Corresponding author: C.-H. Su, Department of Infrastructure Engineering, University of Melbourne, Victoria 3010, Australia. (csu@unimelb.edu.au)

[1]Department of Infrastructure Engineering, University of Melbourne, Victoria 3010, Australia.

Abstract. Recursive digital filters (RDFs) are widely used for estimating baseflow from streamflow hydrographs, and various forms of RDFs have been developed based on different physical models. Numerical experiments have been used to objectively evaluate their performance, but they have not been sufficiently comprehensive to assess a wide range of RDFs. This paper extends these studies to understand the limitations of a generalized RDF method as a pathway for future field calibration. Two formalisms are presented to generalize most existing RDFs, allowing systematic tuning of their complexity. The RDFs with variable complexity are evaluated collectively in a synthetic setting, using modelled daily baseflow produced by *Li et al.* [2014] from a range of synthetic catchments simulated with HydroGeoSphere. Our evaluation reveals that there are optimal RDF complexities in reproducing baseflow simulations, but shows that there is an inherent physical inconsistency within the RDF construction. Even under the idealized setting where true baseflow data are available to calibrate the RDFs, there is persistent disagreement between true and estimated baseflow over catchments with small baseflow components, low saturated hydraulic conductivity of the soil and larger surface runoff. The simplest explanation is that low baseflow 'signal' in the streamflow data is hard to distinguish, although more complex RDFs can improve upon the simpler Eckhardt filter at these catchments.

1. Introduction

Partitioning of streamflow into different fluxes, such as overland flow, subsurface flow and groundwater discharge, is needed to achieve a deeper understanding of the underlying processes in river basins. While streamflow can be directly and continuously monitored, this is difficult for individual flow paths at a catchment scale [*Kalbus et al.*, 2013]. In the presence of a continuum of subsurface flow paths, it is common to consider only two flow components in the makeup of the total discharge – storm runoff versus groundwater discharge, or quickflow versus slowflow [*Chapman*, 1999]. Slow flow is generally treated as synonymous to baseflow, although drainage from other sources, such as wetlands, bank storage, surface water bodies, snowpacks, and the unsaturated zone, can contribute to slow flow [*Halford and Mayer*, 2000]. Here the term baseflow is used specifically to describe groundwater discharge from aquifers that reaches the stream.

Streamflow hydrographs are commonly analyzed to distinguish their baseflow components, and numerous automated separation methods have been developed. For instance, baseflow has been estimated from computing moving-window statistics of the streamflow [*Pettyjohn and Henning*, 1979; *Sloto and Crouse*, 1996], extrapolation of recession curves [*Wittenberg*, 1999; *Van Dijk*, 2010], low-pass filtering [*Lyne and Hollick*, 1979], and using recursive digital filters (RDFs) [*Boughton*, 1993; *Jakeman and Hornberger*, 1993; *Chapman and Maxwell*, 1996; *Furey and Gupta*, 2001; *Eckhardt*, 2005; *Huyck et al.*, 2005; *Croke*, 2010]. Correct baseflow estimates are critical to support a wide range of applications, e.g., water management, low-flow forecasting, hydrologic modelling [*Ferket*, 2010], infrastructure design purposes, catchment classification and identification of climatic and

physiographic factors [*Van Dijk*, 2010; *Carrillo et al.*, 2011; *Beck et al.*, 2013], dominant processes and source areas generating runoff [*Gonzales et al.*, 2009; *Cartwright et al.*, 2014; *Costelloe et al.*, 2015].

RDFs are commonly used in practice because they can be easily automated to yield reproducible results. Their algorithms involve relating baseflow at a given time step, to total discharge at the current and previous time steps and baseflow at previous time steps. The existing RDFs, which differ in degree of complexity, were developed based on a simplified model of a linear reservoir [*Eckhardt*, 2005], cascaded stores [*Croke*, 2010], hillslope mass balance [*Furey and Gupta*, 2001], or a groundwater aquifer [*Huyck et al.*, 2005]. To develop better understanding of hydrological processes, it is desirable to move towards a process-based interpretation of flow components by using these physically-based filters; however, the choice of RDFs used in practice is often ad-hoc, dependent on the applicability of underlying assumptions regarding the aquifer characteristics, and based on preferences (with default filter parameter values) and availability of observational data (e.g., discharge and precipitation) to partially calibrate the filter parameters.

There is now increasing effort to use independent baseflow estimates from field measurements of groundwater levels [*Peters and Van Lanen*, 2005], and natural and artificial tracers [*Stewart et al.*, 2007; *Zhang et al.*, 2013; *Gonzales et al.*, 2009; *Rimmer and Hartmann*, 2014] in RDF calibration. As more observational data are applied to improve the objectivity of filter calibration, a clear guidance for different choices of RDFs is needed. In particular, the objective choice of these filters should be based on their ability to recover all, or part of, the characteristics of baseflow.

At present, controlled numerical experiments are the only practical tools available to test different baseflow separation methods at the catchment scale. *Furey and Gupta* [2003] evaluated filter-estimated baseflow against simulations from a numerical model of a 2-dimensional hillslope. More recent studies take into account the influence of channel routing and losses on flow generation by using simulations generated by HydroGeoSphere (HGS) [*Therrien et al.*, 2009] – a fully integrated 3-dimensional physically-based surface-water/groundwater model, representing a range of physical catchment characteristics and hydrological forcings [*Partington et al.*, 2011, 2012; *Li et al.*, 2013, 2014].

Within this context, this paper extends these evaluation studies to analyse the structural adequacy of RDFs for reproducing HGS simulations generated by *Li et al.* [2014]. Our evaluation framework closely follows *Li et al.* [2013] by using simulated daily baseflow to calibrate and evaluate the RDFs. The novel contributions of this work are as follows. We evaluate a generalized form of RDFs that encompasses filters of *Chapman and Maxwell* [1996], *Boughton* [1993], *Jakeman and Hornberger* [1993], *Eckhardt* [2005], *Furey and Gupta* [2001], *Huyck et al.* [2005], and *Croke* [2010]. It closes the gap in *Partington et al.* [2012] and *Li et al.* [2014] by testing the more complex physically-based RDFs of *Furey and Gupta* [2001], *Huyck et al.* [2005] and *Croke* [2010].

We first extend the formalism of *Croke* [2010] to highlight a more generic construction of RDFs with arbitrary numbers of quickflow and slowflow pathways. Then we further extend the RDF construction using digital signal processing theory. Through generalization of these filters, the fundamental structures of recursive filtering can be assessed collectively without restricting to a specific filter design. That is, the results are applicable to any form of RDF with arbitrary complexity developed with different linear physical models.

Using calibration and split-sample evaluation under idealised scenarios where true baseflow is known, we demonstrate that there are optimal RDF complexities but also inherent limitations of the RDFs in reproducing HGS simulations, and by extension, baseflow from actual streamflow measurements. Finally, we associate the performance and limitations of the optimal RDFs and benefits of using higher-order RDFs with catchment characteristics.

2. Theory

Two theoretical formalisms are presented in Sec. 2.2 to generalize the existing RDFs. Before proceeding to their full mathematical exposition, we review their general concepts. The first formalism extends the model of *Croke* [2010], which treats a single 'pathway' (as a linear store drainage) to quickflow generation, and a separate pathway to recharge that cascades to generate baseflow. We generalize this model by assuming multiple independent pathways to quickflow generation and similarly, multiple independent pathways recharge and baseflow. The second formalism conceptualizes the problem of filter construction as a timeseries modelling problem, where the RDF is a particular timeseries model that describes the influence of past baseflow values, and past and present streamflow values on the current baseflow values. In this way, the resultant RDFs are even more flexible than those constructed from the first formalism. The key mathematical results are as follows: the model of *Croke* [2010] yields the RDF equation in Eq. 6, the first formalism leads to EX(M,N) filters prescribed by Eq. 28, and the second formalism leads to GEN(I,J) filters given in Eq. 30. Tab. 1 and Tab. 2 translate the generalized filters into the aforementioned existing RDFs.

2.1. Review of *Croke* [2010] Formalism

The model uses effective rainfall (u), which is rainfall available for streamflow generation, to produce quickflow q and recharge r by passing effective rainfall through individual exponentially decaying (linear) stores. Slowflow s, which is treated as synonymous with baseflow in this formulation, is generated by passing recharge through another exponentially decaying store. Consequently, the system of equations is,

$$y_n = q_n + s_n, \tag{1}$$

$$q_n = -\alpha_q q_{n-1} + \beta_q u_{n-\delta_q}, \tag{2}$$

$$r_n = -\alpha_r r_{n-1} + \beta_r u_{n-\delta_r}, \tag{3}$$

$$s_n = -\alpha_s s_{n-1} + \beta_s r_{n-\delta_s}, \tag{4}$$

where y_n, q_n and s_n are total streamflow, quickflow and baseflow at time step n, respectively. Each store is associated with a recession constant $-\alpha_* \in [0,1]$. $\beta_q = (1+\alpha_q)\gamma_q \in [0,1]$, $\beta_r = (1+\alpha_r)\gamma_r \in [0,1]$, and $\beta_s = (1+\alpha_s) \in [0,1]$. $\gamma_r \in [0,1]$ is the fraction of effective rainfall that becomes recharge and baseflow, whereas the remaining fraction γ_q of rainfall becomes quickflow. Hence, the mass conservation for effective rainfall requires that,

$$\gamma_q + \gamma_r = 1. \tag{5}$$

The integer-valued indices δ_* are used to specify the time step delay between rainfall and quickflow response (δ_q), between effective rainfall and recharge response (δ_r), and between recharge and the baseflow response (δ_s).

With some algebra, Eqs. 1–4 lead to the RDF algorithmic equation where the baseflow timeseries s_n is a function of streamflow y_n,

$$s_n = -(\alpha_s + \alpha_r)s_{n-1} - \alpha_s\alpha_r s_{n-2} -$$

$$\beta'(s_{n-\Delta} + \alpha_q s_{n-\Delta-1}) +$$

$$\beta'(y_{n-\Delta} + \alpha_q y_{n-\Delta-1}) \quad (6)$$

where $\beta' = \beta_s\beta_r/\beta_q$ and $\Delta = \delta_s + \delta_r - \delta_q$. This time-domain (difference) equation specifies a filtering process acting on the input (streamflow) data, and is a recursion as the first three terms (feedback) use past filter outputs. It is a linear, time-invariant (LTI) filter because the filter output is a linear function, and the filter parameters do not vary with time. It is also a causal filter because only past and present information on effective rainfall and streamflow are used in the estimation of present baseflow. In this way, it is distinctive from a commonly-used, non-causal filter of *Lyne and Hollick* [1979] where its backward filtering pass uses future streamflow information.

To apply the filter in Eq. 6, 5 filter parameters $\{\alpha_q, \alpha_r, \alpha_s, \gamma_r, \Delta\}$ must be determined beforehand, e.g., either a priori or through calibration. *Croke* [2010] noted that the filters of *Chapman and Maxwell* [1996] (C-M), *Boughton* [1993], *Jakeman and Hornberger* [1993] (IHACRES), and *Furey and Gupta* [2001] (F-G) are special cases of Eq. 6. Tab. 1 provides the parameterization and values of the filter parameters to reduce Eq. 6 to these filters. The Boussinesq filter of *Huyck et al.* [2005] and Eckhardt filter [*Eckhardt*, 2005] are also included, although the Eckhardt and Boughton filters are mathematically equivalent. All these filters assume instantaneous recharge with $\alpha_r = 0$. With $\alpha_q = 0$, C-M, Boughton, Eckhardt and F-G filters assume that quickflow is drained at time delay δ_q after rainfall; in contrast, IHACRES and Boussinesq filters allow transient drainage of quickflow. With $\delta_s = \delta_q$, C-M, Boughton, Eckhardt and IHACRES filters also assume that the drainage of quickflow and baseflow begin simultaneously, while F-G and Boussinesq filters relaxes this restriction to allow a time lapse between them.

Before proceeding to further generalization of Eq. 6, it is instructive to cast the time-domain filtering equation as a transfer function $H(z)$ in the z-domain. This allows us to examine the overall properties of the filter in the frequency domain, and it is used to simplify the derivation of generalized filter designs in Sec. 2.2. This involves applying z-transform, denoted by the $\mathcal{Z}$ operator, to both sides of the Eq. 6 using the following properties:

$$\mathcal{Z}(a_n) = A(z), \tag{7}$$

$$\mathcal{Z}(a_{n-m}) = A(z)z^{-m}, \tag{8}$$

where $A(z)$ is the transform of an arbitrary timeseries a_n, $z \equiv \exp(i\omega) = \cos(\omega) + i\sin(\omega)$ and ω is the analysis angular frequency. Given the inverse transform, the timeseries a_n can be expressed as a sum of sinusoidal functions with a range of frequencies,

$$a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} A(z)z^n d\omega. \tag{9}$$

For instance, the transform $S(z) \equiv \mathcal{Z}(s_n)$ informs the relative magnitude of the frequency contents in the baseflow timeseries. Similarly for streamflow timeseries, its transform is $Y(z) \equiv \mathcal{Z}(y_n)$.

Applying the z-transform to both sides of Eq. 6 leads to,

$$\begin{aligned} S(z) = & -(\alpha_s + \alpha_r)S(z)z^{-1} - \alpha_s\alpha_r S(z)z^{-2} - \\ & \beta'(S(z)z^{-\Delta} + \alpha_q S(z)z^{-\Delta-1}) + \\ & \beta'(Y(z)z^{-\Delta} + \alpha_q Y(z)z^{-\Delta-1}) \end{aligned} \tag{10}$$

The transfer function $H(z)$ of the RDF relates the input $Y(z)$ to output $S(z)$, i.e. $S(z) = H(z)Y(z)$. Collecting terms in Eq. 10, we arrive at,

$$H(z) = \frac{\beta'[z^{-\Delta} + \alpha_q z^{-(\Delta+1)}]}{1 + (\alpha_s + \alpha_r)z^{-1} + \alpha_s\alpha_r z^{-2} + \beta'[z^{-\Delta} + \alpha_q z^{-(\Delta+1)}]}. \tag{11}$$

Notice that the coefficients of z and its exponents in the numerator of Eq. 11 correspond to the coefficients and the time lags of the input terms (y_n) in Eq. 6, respectively. On the other hand, the minus of the coefficients of z (ignoring the unity) and its exponents in the denominator are those of the feedback terms (s_n). In other words, the coefficients and exponents of $H(z)$ can be read off directly to construct the time-domain equation (see also Eq. 30 and Eq. 31).

The modulus $|H(z)|$, known as the gain function, describes the attenuation of different frequency components that make up y. At low frequencies, $\lim_{\omega\to 0} |H(z)| = \gamma_r$ so that the parameter γ_r informs the level of attenuation on the low-frequency components of y, and thus influences the overall volume of baseflow. More specifically, there is an identity relation

$$H(\omega = 0) = \gamma_r, \tag{12}$$

which stipulates that the ratio of the temporal mean of the filter output ($\bar{y}$), which includes the transient response of the filter after the input is switched off, to the mean of the filter input ($\bar{s}$) must be given by γ_r. This is simply an input-output mass conservation condition where the fraction of baseflow in streamflow is directly related to the recharge component of effective rainfall. Consequently, specifying the filter parameter γ_r is equivalent to specifying the baseflow index (BFI) in the filtering period. In the current LTI form, the filter in Eq. 6 permits analytical sensitivity analysis of its parameters, such as that

performed for the Eckhardt filter by *Eckhardt* [2012]. However, the filter also allows negative flows, i.e. estimated $s_n > y_n$.

To avoid negative flows, it is a common practice to apply thresholding to Eq. 6, leading to different baseflow timeseries $\tilde{s}$. In particular, the baseflow estimate $\tilde{s}_n$ at a time step n is given by taking the minimum value between streamflow y_n and the baseflow estimate from Eq. 6,

$$\tilde{s}_n = \min(y_n, s_n). \tag{13}$$

The thresholding can be conducted either when filtering of the entire timeseries is completed or at each step of the recursive filtering algorithm. For the latter approach, which this study adopts, the RDF algorithm equation becomes,

$$\begin{aligned}\tilde{s}_n \;=\; \min\Big[y_n, & -(\alpha_s+\alpha_r)\tilde{s}_{n-1} - \alpha_s\alpha_r\tilde{s}_{n-2} - \\ & \beta'(\tilde{s}_{n-\Delta} + \alpha_q\tilde{s}_{n-\Delta-1}) + \\ & \beta'(y_{n-\Delta} + \alpha_q y_{n-\Delta-1})\Big].\end{aligned} \tag{14}$$

In applying thresholds, there is an implicit assumption that there is zero measurement error in streamflow. With this constraint, the filtering becomes mathematically nonlinear and time-variant, even though the filter construction is based on linear assumptions. Analytical solutions for the transfer function and sensitivity analysis are no longer valid. The most problematic issue is that Eq. 13 leads to internal inconsistency of the filter construction framework because the input-output mass conservation conditions in Eq. 5 and Eq. 12 will be violated. In other words, the constraint is not compatible with the physical meanings associated with the model parameters in Eq. 1–4 and the parameters of the physically-based filters listed in Tab. 1. Indeed, *Furey and Gupta* [2001] suggested

that Eq. 13 should not be imposed on physically-based methods; steps should be taken to modify the method so that these errors are reduced or fully removed for physical consistency. Alternatively, the possible values of the filter parameters can be restricted to produce $s_n \leq y_n$ for all time steps without using Eq. 13, but this may severely limit the flexibility of the filter to reproduce certain baseflow dynamics. This issue will be explored further in this work through systematic modifications of RDFs and by using different filter calibration scenarios.

Finally, when accurate catchment average rainfall data is available, the explicit modelling of effective rainfall allows one to place a constraint on the effective rainfall u_n based on observed rainfall p_n. Assuming that there is no timing difference (e.g., due to routing) between u_n and p_n, the filtering process can require that $u_n \leq p_n$ for all time steps n [*Croke*, 2010]. This is implemented through another thresholding (of estimated u_n) in addition to the above Eq. 13, leading to a two-stage filtering algorithm (see details in Appendix A). In this paper, we focus on the more typical cases where observed rainfall is unavailable or not used and ignore such filter implementation.

2.2. Generalization of linear RDF construction

The first approach to extend the formalism in Eq. 1–4 is to include multiple pathways for quickflow, recharge and baseflow by considering a multiplicity of exponentially decaying stores. The system of equations then becomes,

$$y_n = s_n + q_n = \sum_{i=1}^{M} s_{i,n} + \sum_{j=1}^{N} q_{j,n}, \tag{15}$$

$$q_{i,n} = -\alpha_{q,i} q_{i,n-1} + \beta_{q,i} u_{n-\delta_{q,i}}, \tag{16}$$

$$r_{i,n} = -\alpha_{r,i} r_{i,n-1} + \beta_{r,i} u_{n-\delta_{r,i}}, \tag{17}$$

$$s_{i,n} = -\alpha_{s,i} s_{i,n-1} + \beta_{s,i} r_{n-\delta_{s,i}}, \tag{18}$$

where $i = 1, 2, ..., M$ each labels a pathway to recharge and baseflow and $j = 1, 2, ..., N$ labels a pathway to quickflow. $\beta_{q,i} = (1 + \alpha_{q,i})\gamma_{q,i}$, $\beta_{r,i} = (1 + \alpha_{r,i})\gamma_{r,i}$, and $\beta_{s,i} = 1 + \alpha_{s,i}$. For mass conservation, we require that

$$\sum_{i=1}^{M} \gamma_{r,i} + \sum_{j=1}^{N} \gamma_{q,j} = 1. \tag{19}$$

To solve for the filtering equation, it is easiest to do so in the conjugate z-space. Under z-transform, Eq. 15–18 becomes,

$$Y = S + Q = \sum_{i=1}^{M} S_i + \sum_{j=1}^{N} Q_j, \tag{20}$$

$$Q_i = -\alpha_{q,i} Q_i z^{-1} + \beta_{q,i} U z^{-\delta_{q,i}}, \tag{21}$$

$$R_i = -\alpha_{r,i} R_i z^{-1} + \beta_{r,i} U z^{-\delta_{r,i}}, \tag{22}$$

$$S_i = -\alpha_{s,i} S_i z^{-1} + \beta_{s,i} R_i z^{-\delta_{s,i}}. \tag{23}$$

where Y, Q, S, U, Q_i, R_i, and S_i are transforms of y_n, q_n, s_n, u_n, $q_{i,n}$, $r_{i,n}$, and $s_{i,n}$, respectively. Substituting Eq. 22 and 23 to $S = \sum_i S_i$ yields,

$$S = \mathcal{C} U, \tag{24}$$

with

$$\mathcal{C} = \sum_{i=1}^{M} \frac{\beta_{s,i}\beta_{r,i} z^{-\delta_{s,i}-\delta_{r,i}}}{(1 + \alpha_{s,i} z^{-1})(1 + \alpha_{r,i} z^{-1})}. \tag{25}$$

Then, substituting Eq. 21 to $S = Y - \sum_{j=1}^{N} Q_i$ yields,

$$U = \frac{Y - S}{\mathcal{D}}, \tag{26}$$

given

$$\mathcal{D} = \sum_{j=1}^{N} \frac{\beta_{q,j} z^{-\delta_{q,i}}}{1 + \alpha_{q,i} z^{-1}}. \tag{27}$$

Finally, we combine Eq. 26 with Eq. 24 to arrive at the transfer function of the system $H(z) \equiv S/Y$ with multiple quickflow, recharge and baseflow pathways,

$$H_{\mathrm{EX}(M,N)}(z) = \frac{\mathcal{C}}{\mathcal{C}+\mathcal{D}}. \tag{28}$$

where the subscript EX(M,N) is used to identify this manner of filter construction using exponentially decaying stores, and integers M, N are included to denote the number of pathways in the construction. As with Eq. 12, the transfer function of the resultant EX(M,N) filter satisfies the input-output mass conservation,

$$H_{\mathrm{EX}(M,N)}(\omega = 0) = \sum_{i=1}^{M} \gamma_{r,i}. \tag{29}$$

For $M, N = 1$, Eq. 28 reduces to Eq. 11, i.e., $H_{\mathrm{EX}(1,1)}(z)$. As before, the coefficients and exponents of z in the numerator and denominator of Eq. 28 can be extracted to prepare a time-domain equation. For $M, N \geq 2$, the equation is not sufficiently compact to be presented here. It should be noted that the number of effective filter parameters grows with $4M + 3N - 2$.

The second approach of generalizing the RDFs is to use the formalism of infinite-impulse response (IIR) filters from the theory of digital signal processing. Without loss of generality, all the RDFs, including the filters in Eq. 6, Eq. 28 and Tab. 1 can be written as special cases of the following,

$$s_n = \sum_{i=1}^{I} a_i s_{n-i} + \sum_{j=0}^{J} b_j y_{n-j}. \tag{30}$$

The presence of the recursion (first) term means that the filter estimate at time step n is derived using streamflow information from all time steps before n. If there are any nonzero streamflow values before n, the value of s_n is strictly non-zero. Thus, such linear filters produce only perennial flows. Without the recursion term, the filter is regarded as

a finite-impulse response (FIR) filter, and can in theory produce zero baseflow estimates for ephemeral flows.

The filter order is given by the larger number of the feedforward order J and feedback order I. In contrast to Eq. 28, this expression relaxes the restriction on the number of flow pathways or the response time delays. Applying z-transform, its associated transfer function,

$$H_{\mathrm{GEN}(I,J)}(z) = \frac{\sum_{j=0}^{J} b_j z^{-j}}{1 - \sum_{i=1}^{I} a_i z^{-i}}, \tag{31}$$

where the correspondence between the coefficients and exponents of z with the coefficients and time lags in Eq. 30 is apparent. The subscript GEN(I,J) labels this filter as a full generalization of all forms of *linear* RDFs. To assess the stability of the GEN(I,J) filter, the polynomials in the numerator and denominator are factorized into their roots c_j and d_i,

$$H_{\mathrm{GEN}(I,J)}(z) = \frac{b_0 \prod_{j=1}^{J}(1 - c_j z^{-j})}{\prod_{i=1}^{I}(1 - d_i z^{-i})}. \tag{32}$$

For the linear filter to be stable, the magnitude of the roots $|d_i|$ must be less than unity. To ensure that $\bar{s}/\bar{y} \leq 1$, we also require that,

$$H_{\mathrm{GEN}(I,J)}(\omega = 0) = \frac{\sum_{j=0}^{J} b_j}{1 - \sum_{i=1}^{I} a_i} = \frac{b_0 \prod_{j=1}^{J}(1 - c_j)}{\prod_{i=1}^{I}(1 - d_i)} \leq 1. \tag{33}$$

In other words, specifying the filter parameters a_i and b_j is equivalent to specifying the BFI in the filtering period when the constraint in Eq. 13 is not applied. For completeness, Tab. 2 illustrates the direct mapping between the IIR GEN(I,J) filter (Eq. 30) and the existing RDFs.

In summary, the Venn diagram in Fig. 1 illustrates the relationships between baseflow estimated by different RDFs. Suppose the real world produces many possible baseflow

realizations in rivers ranging from perennial to ephemeral flow regimes. Within which, a subset of the baseflow realizations can be reproduced using numerical catchment or hydrologic models such as HydroGeoSphere (Sec. 3.1). Depending on the complexity of the models, a subset of the model's simulations can be reproduced using linear RDFs. Conversely, as the EX filter can in theory be used to model an arbitrary number of quickflow and recharge-baseflow pathways, their realizations of baseflow can possibly reside outside model realizations. Within the GEN(I,J) filters lies the nesting of EX(M,N) filters, and the existing simpler RDFs, as per Tab. 1 and 2.

When the EX and GEN filters are implemented with Eq. 13 and rendered nonlinear, their solutions are distinguishable from the solutions from linear filters. For convenience, these filters are referred to as EX$_{Th}$(M,N) and GEN$_{Th}$(I,J). Furthermore, zero baseflow values in ephemeral flows are retrieved via thresholding. We note that (not shown completely) the hierarchy of the GEN$_{Th}$ and EX$_{Th}$ filters are nested in the same way as the linear counterparts. It should be noted that in addition to the filters depicted in Fig. 1, there can be other nonlinear, time-variant filters that are structurally different from and more general than these filters, and use a non-linear storage-discharge relationship [e.g. *Wittenberg*, 1999] and use additional input timeseries data [e.g. *Furey and Gupta*, 2003].

3. Methodology

By following the evaluation framework of *Li et al.* [2013], this study aims to provide insights into the functionality and performance of the RDFs in the context where independent baseflow data are available for calibration. We recognize that independent calibration data are frequently not available in many catchments where RDFs are used in practice and we comment on this in our discussion in Sec. 5. Accordingly, we test the

ability of the generalised RDFs, EX(M,N) filter, and GEN(I,J) filter, to reproduce the model-simulated daily baseflow.

Three calibration approaches are considered to explore the aforementioned issue of internal inconsistency arising from applying thresholding to filter estimates (Eq. 13) (Sec. 3.2). Sec. 3.3 describes the metrics for evaluating the filter performance, and Sec. 3.4 presents the correlation analysis to associate performance to catchment properties.

3.1. Simulated hydrological data

Li et al. [2014] generated a test data set containing synthetic daily streamflow and baseflow timeseries by integrating the 3D surface water/groundwater model, HydroGeoSphere (HGS). A HGS model represents many physical processes, including 3D saturated/unsaturated groundwater flow, overland flow, streamflow, and infiltration/exfiltration, explicitly within a 3D catchment. The 2D St. Venant equations under diffusion wave approximation, 3D modified Richard's equation, and head continuity equation were solved simultaneously to simulate surface flow, subsurface flow, and surface-subsurface coupling at each time step, respectively. We therefore consider the HGS model to provide a superior numerical experimental structure than could be generated using a conceptual model structure.

In this case, the HGS model was for a tilted V-catchment with uniform soils above an impermeable basement at 20 m depth, and by symmetry, only one half of each catchment was simulated. It was forced by 10-year rainfall and potential evapotranspiration (PET) in-situ measurements from five Australian cities, including Adelaide, Brisbane, Darwin, Melbourne and Sydney, representing temperate, sub-tropical and tropical climatic conditions. In particular, 70 representative catchments were set up with randomly sampled

model parameter values, namely catchment area (A), hill slope ($S1$), channel slope ($S2$), aspect ratio (AR) and soil properties, using Latin Hypercube sampling strategy. However, four that generated zero streamflow are excluded. The values are listed in Tab. 3.

The catchment area is related to aquifier storage capacity that sustains baseflow generation. The hill and channel slopes influence the partitioning of rainfall to channel flow and groundwater recharge and the rate at which soil water moves downslope. The channel aspect ratio between catchment width to length affects the travel time of surface runoff and baseflow to the channel outlet. The soil type is controlled by varying saturated hydraulic conductivity K_s and van Genuchten parameters α_{vG} and β_{vG}, to represent sandy, sandy loam, and loamy sand.

The hydraulic mixing-cell (HMC) method of *Partington et al.* [2011, 2013] was used to determine the groundwater component of streamflow at each simulation time step in each cell of the stream. In particular, the streamflow generation at each stream cell, which is the solution to the HGS's fluid mass balance equation, is interpreted by HMC to distinguish groundwater, overland flow, unsaturated flow and preferential flow entering or leaving the cell, and rainfall contribution to and evaporative loss from the storage within the cell.

The 66 modelled catchments show different baseflow dynamics. Their baseflow components have variable BFI between 0.006–0.997 and the coefficient of quartile variation (CQV) of baseflow varies between 0.03–1 (see also Fig. S1 and S2 of *Supplementary Materials*). It should be noted that almost all (64) catchments are perennial with continuous flow almost through the entire simulation period. Readers are referred to *Therrien et al.* [2009] for details of the HGS and *Li et al.* [2014] for the simulation data set.

3.2. Filter implementations and calibration

Consider the idealized scenario where true (i.e., simulated in our case) baseflow data are available to calibrate the RDFs with a cost function between estimated baseflow (s) and the HGS-simulated baseflow ($\hat{s}$) defined as,

$$\mathcal{J}(s,\hat{s}) = \frac{\overline{(s_n^l - \hat{s}_n^l)^2}}{\overline{(\hat{s}_n^l - \bar{\hat{s}}^l)^2}}. \tag{34}$$

where $\overline{\circ}$ denotes time averaging, and s^l and $\hat{s}^l$ are the logarithmic transforms of s and $\hat{s}$, e.g., $s^l = \ln(s + 10^{-6})$, where 10^{-6} mm/day is the order of the lowest simulated non-zero flow across the catchments. Note that $\mathcal{J}$ is based on the Nash-Sutcliffe coefficient of efficiency (NSE) and penalizes mismatch in mean, variance and temporal shape of the baseflow timeseries. The model data are split equally into calibration (first half) and verification (second-half) data sets so that there are 5 years of data points in each period. The calibration with the above cost function is equivalent to using other least-squares metrics such as root-mean-square deviation.

The three schemes for filter implementation and calibration, denoted by *C1*, *C2*, and *C3*, are described as follows.

- *C1*: The filters are applied without thresholding (i.e., without Eq. 13) and the resultant linear filters are allowed to produce negative flows, i.e., $s_n > y_n$. The cost function is defined using all the data in the calibration period.

- *C2*: The filters are applied with thresholding as $\mathrm{GEN}_{Th}(I,J)$ and $\mathrm{EX}_{Th}(M,N)$ filters. Following *Li et al.* [2014], the cost function is defined using only the 'non-recession' data that satisfies the inequality $\hat{s} \neq y$ in the calibration period.

- *C3*: Same as *C1*, but we consider only $\mathrm{GEN}(I,J)$ filters with parameter values that strictly produce positive flows, i.e., $s_n \leq y_n, \forall_n$.

We follow *Furey and Gupta* [2001] to define the initial baseflow values as the minimum value of streamflow over the operating period so as to reduce the rise time of the filter response. For the same reason, the cost function and the following evaluation metrics (see Sec. 3.3) are computed with data starting from the second year of the individual filtering period.

The Eckhardt filter is used as a benchmark to assess the benefits of using higher-order RDFs, and its two parameters (see Tab. 1) are calibrated simultaneously via optimization. The higher-order RDFs are assessed systematically by increasing the filter order of GEN(I,J) filter or increasing the number of pathways in EX(M,N) filter, for $I = J \in [1, 5]$ and $M = N \in [1, 4]$. For the implementation of the higher-order EX(M,N) filters with $M, N \geq 2$, we restrict ourselves to $\delta_{s,i} + \delta_{r,i} - \delta_{q,j} = 0$ and $\delta_{q,i} - \delta_{q,j} = 0, \forall_{i,j}$ in Eq. 28 to reduce the computational burden of mixed-integer optimization. That is, the filter parameters include δ's, which are integer-valued, and continuous variables α's and γ's. Consequently, these higher-order EX(M,N) filters operate as GEN($2M + N - 1$,$2M + N - 2$) filters. The filter calibration was performed with a range of optimization algorithms, including interior-point, genetic algorithm [*Goldberg*, 1989] and shuffled complex evolution [*Duan et al.*, 1992], and executed multiple times to determine the parameter sets that yield lowest costs.

3.3. Filter evaluation

The performance of the filters is measured by how well they can reproduce the HGS-simulated daily baseflow during calibration and verification periods. The filters are applied

to the streamflow data from the two periods independently. The performance metrics are,

$$\mathrm{NSE}_l = 1 - \frac{\overline{(s_n^l - \hat{s}_n^l)^2}}{\overline{(\hat{s}_n^l - \bar{\hat{s}}^l)^2}}, \tag{35}$$

$$\mathrm{NSE} = 1 - \frac{\overline{(s_n - \hat{s}_n)^2}}{\overline{(\hat{s}_n - \bar{\hat{s}})^2}}, \tag{36}$$

$$\mathrm{PBIAS} = \frac{\overline{(s - \hat{s})}}{\bar{\hat{s}}} \times 100, \tag{37}$$

and Pearson's correlation coefficient $R(s, \hat{s})$. Given the choice of the cost function in calibration, filters with $\mathrm{NSE}_l = 1$ are perfect, $\mathrm{NSE}_l \in [0.5, 1)$ are considered to have a 'good' performance, and $\mathrm{NSE}_l \in [0, 0.5)$ are considered 'average', otherwise 'unacceptable'. This is similar for NSE, which allows more weighting towards periods with higher baseflow. PBIAS is a measure of a bias in the estimated baseflow, or equivalently, in the estimated BFI since $\mathrm{PBIAS}/100 = (\bar{s}/\bar{y} - \bar{\hat{s}}/\bar{y})/(\bar{\hat{s}}/\bar{y})$. $R = 1$ means a perfect positive linear association, and zero indicates no linear association. For filters calibrated with the *C2* scheme, the metrics are computed with the 'non-recession' data as per the calibrating cost function, unless stated otherwise.

Increasing the filter orders can render the RDFs excessively complex with too many filter parameters and marginal gains in performance. The Bayesian information criterion (BIC), widely used to determine an optimal complexity during calibration, is applied to the results of the EX(N,M) filters and GEN(I,J) filters, in terms of the number of filter parameters θ,

$$\mathrm{BIC} = L \ln \left[\overline{(\hat{s}_n^l - s_n^l)^2} \right] + \theta \ln L, \tag{38}$$

where L is the sample size. The BIC penalizes the disagreement between true baseflow and estimated values (first term), but also penalizes the multiplicity of filter parameters. Thus, a suitable filter complexity with the lowest BIC is preferred. It is however of note

that the BIC is based on the assumption about prior model probabilities, and that the residual $\hat{s}_n^l - s_n^l$ is an independent and identically distributed random variable (iid). By contrast, the errors in the filter estimates are likely to be serially correlated and time-variant, which over-penalize the mismatch in the estimates. To reduce this, we follow the approach of *Westra* [2014] and compute BIC on the thinned data, i.e., every kth day of the record where k is the minimal value for which the lag-1 autocorrelation is insignificant (p-value > 0.05).

3.4. Correlation analysis on filter performance

The Spearman rank correlation analysis is used to identify possible associations of filter performance and enhancement via generalization with catchment characteristics. It is fit for our purpose as it does not rely on a linear association between variables or require pre-transformations of variables. For the exploratory variables, we consider the controlled catchment model parameters $\{K_s, \alpha_{vG}, \beta_{vG}, A, S1, S2, AR\}$ (Tab. 3), and forcing parameters, namely annual rainfall, PET and their ratios. The catchment hydrologic response properties given by runoff coefficient (RC), BFI, and CQV values of baseflow and streamflow are also included.

For the purpose of identifying the possible factors influencing the strengths and limitations of RDFs, the first set of response variables under consideration comprises the performance metrics $\{\mathrm{NSE}_l, \mathrm{NSE}, \mathrm{PBIAS}, R\}$ (based on calibration data) of the BIC-identified optimal filters. The absolute values or modulus of PBIAS are used. This is analogous to *Li et al.* [2014] who developed log-linear regression models for the performance of Lyne-Hollick and Eckhardt filters with some of the above explanatory variables. The second set of response variables measures the improvements to the Eckhardt

filter by using a higher-order RDF, and they are denoted by changes in the metrics $\{\Delta\text{NSE}_l, \Delta\text{NSE}, \Delta\text{PBIAS}, \Delta R\}$ and the filter order. Improvements are indicated by positive ΔNSE_l, ΔNSE and ΔR, and negative ΔPBIAS. This is to explore the circumstances under which one might benefit from using higher-order RDFs.

4. Results

The results are similar between the GEN and EX filters. We first present the evaluation results of the GEN filters in Sec. 4.1 because of its generality over the EX filters and later review their differences in Sec. 4.2. Subsequently in Sec. 4.3, the results from Spearman correlation analyses are shown.

4.1. GEN(I,J) filters

Fig. 2 shows the NSE_l, NSE, PBIAS and R scores for the GEN filters and Eckhardt filter calibrated with three distinct calibration schemes. The calibration metrics for the 66 catchments are summarized in the boxplots, showing their median, inter-quartile range and outliers. Readers are also referred to Fig. S3-5 of *Supplementary Materials* for the values of the metrics for individual models from calibration, Fig. S6-8 for examples of estimated baseflow timeseries, and Fig. S9 for verification results.

Several observations can be made for *C1* (Fig. 2a). Firstly, there is a significant spread in filter performance with calibration NSE_l values ranging from negative values to near unity. The range of the NSE_l values is between -59 and 1 for the Eckhardt filter and between 0.08 and 1 for GEN(2,2) filter. Negative NSE suggests that the filter performs worse than a time-constant baseflow timeseries, which does not necessarily exist in the

solution space during calibration because it requires filtering to be initialized at a specific, a priori unknown state.

Secondly, the Eckhardt filter and GEN(1,1) filter shows unacceptable $\text{NSE}_l < 0$ at 21 and 11 catchments, respectively. The biases present in their baseflow estimates are primarily responsible for their negative NSEs. The use of the GEN(2,2) filter leads to largest increases in calibration NSE, and the improvements are found in both the PBIAS and R. This is because the higher-order RDFs can reproduce some of the high-frequency fluctuations in the baseflow timeseries (see Fig. S6 of *Supplementary Materials*). They also reduce the rise time needed by the Eckhardt filter and produce baseflow levels closer to the truth.

Thirdly, the application of even higher-order GEN($> 2, > 2$) filters leads to relatively limited improvements in agreement with the synthetic truth. By using the GEN filters, the number of catchments with average-to-good performance (NSE > 0) increases from 45 (Eckhardt) to 66, but the number with good performance (NSE ≥ 0.5) only increases from 32 (Eckhardt) to 39.

Separate comparative BIC analyses on the calibration and verification data found that GEN(2,2) filters are 'optimal' with the minimum BIC values at $> 50\%$ (35–38) of catchments. The NSE_l, NSE, PBIAS and R measures computed with the verification data (not shown) are found to show similar trends as a function of filter complexity to those computed with the calibration data. There is some degradation in NSE_l and PBIAS in the verification data. This can be attributed to the filter parameter errors because the filters have been calibrated to a data period with different BFI and there is a direct relation between filter parameter values and specific BFI response as per Eq. 33. In particular,

the BFI values over the verification period differ by up to $50-70\%$ of the BFI values over the calibration period for some catchments.

For *C2* (Fig. 2b), we observe that the NSE_l and R over the rising hydrographs are generally high and the biases are mostly within $\pm 5\%$ range. There is limited capacity of the generalized filters of higher orders to improve upon the Eckhardt filter. There are roughly the same number of catchments with above zero NSE_l (63 for Eckhardt and 64 for EX(2,2)) and with good performance (46 and 48, respectively). Nevertheless, the BIC analyses show preference for the GEN_{Th}(1,1) filters, and higher-order filters are optimal at over 20 catchments.

Next, Fig. 2c summarizes the results of using the linear GEN filters but with *C3*, where the filters must strictly estimate positive flow. Even though R is moderately high with its $\text{IQR} = 0.3 - 0.9$, the calibration scheme leads to mass-conservative but biased baseflow estimates. The biases are the largest amongst the three calibration scenarios, and consequently, the IQR of the NSE_l values appears much worse. The Eckhardt and GEN filters show negative NSE_l at 31-34 catchments. The use of the GEN filter, especially GEN(1,1) and GEN(2,2), leads to some improvements in PBIAS and R. Accordingly, these two filters are generally preferred at 51 catchments based on the BIC.

4.2. EX(M,N) filters

The EX filters are special cases of the GEN filters as they require the filter parameters of the GEN filters to have fixed relations. The *C1* and *C2* calibration scenarios are used specifically and are shown in Fig. 3. Fig. S10–S12 of *Supplementary Materials* provides additional illustrations of the metric values. There are notable similarities between the results of the EX and GEN filters for both cases. Specifically for *C1*, increasing the

filter complexity improves upon the Eckhardt filter in terms of biases, R and number of catchments with acceptable and good performances. For *C2*, the advantages of using the generalized EX filters over the Eckhardt filter are limited.

BIC shows a clear preference for using the EX(1,1) filter over the Eckhardt filter for *C1*. And for *C2*, Eckhardt filter is preferred but EX(1,1) and EX(2,2) are preferred at 29–36 catchments. When using the EX(1,1) filter, we found that the values of the calibrated parameter Δ (Eq. 6) are zero at most (63) catchments. In other words, the EX(1,1) filter effectively functions as a (special case of) GEN(2,1) filter at these catchments, and with the restrictions on the filters' delay parameters (see Sec. 3.2), an EX(2,2) filter operates as a GEN(5,4) filter and an EX(3,3) filter as a GEN(8,7).

The increases in the number of quickflow and recharge pathways modelled in the EX($>2, >2$) filters allow access to more feedback and feedforward terms than the highest-order GEN filters considered in this study, but there is greater flexibility in the choices of filter parameter values for the GEN filters. Keeping these differences in mind, we compared the performance of the 'optimal' GEN and EX filters selected based on their lowest BIC values. Their calibration and verification NSE_l values are similar within 10% difference for most (49–61) catchments. Amongst the remaining catchments where two generalized filters show $>10\%$ differences in their calibration NSE_l values, the GEN filters outperform the EX filters in both *C1* and *C2*.

4.3. Relationships with catchment characteristics

Fig. 4 shows the Spearman's correlation matrices between exploratory and response variables identified in Section 3.4 for *C1* and *C2* scenarios. We use the evaluation results of the GEN filters, given their similarities with the EX filters. It is apparent that the per-

formance metrics of the RDFs generally show significant associations (p-value < 0.005) with K_s (saturated hydraulic conductivity), climate forcing and flow characteristics. The other catchment parameters, namely van Genuchten parameters and geometrical parameters, are poorer indicators of filter performance and filter enhancement metrics. This is consistent with the regression results of *Li et al.* [2014], who found that K_s and the ratio rainfall/PET are good indicators for the performance of the Lyne-Hollick and the Eckhardt filters.

The strong influence of K_s on the size of the simulated baseflow components (BFI) is expected because larger K_s values facilitate infiltration, leading to increased soil saturation and groundwater discharge [*Li et al.*, 2013]. By contrast, the soil permeability tuned with the van Genuchten parameters do not appear to have similarly strong influence on BFI. Higher K_s values are also associated with lower CQV of streamflow, where the streamflow has sustained flow owing to groundwater discharge and/or small temporal variations due to smaller surface runoff or quick flow component. These are compatible with observing higher BFIs. In other words, the filter performance of the optimal RDFs show significant positive associations with K_s (and BFI), and negative association with the CQV of streamflow.

The filter performance also shows strong negative association with rainfall, rainfall/PET, and RC. However given the particular choice of the forcing data used in the HGS modelling, there is already an inherent strong positive correlation between rainfall and PET, and between rainfall and rainfall/PET. Given that there is significant positive correlation between rainfall amount and RC and negative correlation between RC and

BFI, the increased rainfall tends to produce increased surface runoff at the simulated catchments.

The benefits of using higher-order RDFs over the Eckhardt filter are associated with the same exploratory variables that affect the above RDF performances. For *C1*, the larger magnitudes of performance improvements tend to be associated with catchments with larger areas, lower BFI, lower baseflow CQV and/or higher runoff coefficient. This is sensible given that the Eckhardt filter has under-performed at these catchments and is therefore more likely to benefit from using higher-order RDFs.

The results from the correlation analysis are supported by Fig. 5. The catchments where the GEN filters perform satisfactorily with $\text{NSE}_l \geq 0.5$ under *C1* have noticeably higher K_s, β_{vG}, S_2, PET and BFI, and lower rainfall/PET, RC and CQV of baseflow and streamflow, in comparison to remaining catchments. Similar observations can be made for *C2*, but with lower α_{vG}. Amongst these variables, BFI, rainfall/PET and RC provide the clearest distinction between these two sets of catchments for both calibration cases. Separating based on the lower or upper quartiles, the GEN filters under *C1* show good performance at the catchments with BFI ≥ 0.60, RC ≤ 0.15 and rainfall/PET ≤ 0.81. For *C2*, satisfactory performance is found at the catchments with BFI ≥ 0.40 and RC ≤ 0.25. Some of these threshold values are more restrictive for the Eckhardt filter; it yields satisfactory NSE_l at the catchments with BFI ≥ 0.78 and RC ≤ 0.11 for *C1*, and at the catchments with BFI ≥ 0.46 for *C2*.

5. Discussion

5.1. Ambiguities in RDF construction and implementation

We have considered three calibration scenarios because there is no clear best practice of implementing the RDFs. In some ways, it is related to the inherent dissociation between the construction and implementation of a RDF. On one hand, the consistency of the physically-based (linear) RDF formulation and associated filter parameters needs to be preserved (in *C1* and *C3*). On the other hand, the baseflow estimates must be conservative at the filter outputs (in *C2* and *C3*).

The consequences of *C1* scheme in producing negative flow components reflect structural errors of the RDFs. The dynamical thresholding in *C2* scheme avoids negative flow but the physical meanings of the calibrated parameters become highly ambiguous. The consideration that effective rainfall can be greater than observed rainfall in the EX formalism (Appendix A) further exacerbates this ambiguity. It could be argued that the thresholding is necessary to mitigate the structural errors of the RDFs. However, at some catchments and during recession periods, there is limited gain in information about baseflow from filtering than what already exists in the streamflow hydrographs. Fig. 6a shows that the amount of baseflow estimates derived from invoking Eq. 13 can be as high as 60–70% of the calibration period. The number of time steps at which thresholding was applied can exceed the number of time steps at which there is no quick flow component, because it was also applied at some of the time steps with quick flow. Fig. 6b also shows that baseflow values during significant fractions of the periods without quick flow were estimated via thresholding.

When the *C3* calibration scheme attempts to satisfy both conditions individually observed by *C1* and *C2*, the estimated baseflow values are biased to ensure that the estimated

baseflow values do not exceed streamflow at all times, and higher-order RDFs offer limited remedy. In other words, the combination of *C3* and more complex RDFs is not a viable approach. This is further illustrated by Fig. 7 where the three schemes are directly compared based on NSE_l of the coincident data; *C2* yields the highest NSE_l, followed by *C1*.

5.2. Limitations of the RDFs

There remains persistent disagreement between the baseflow simulated by HGS and estimated by generalised RDFs, even with high-order complexity, at many catchments. Their inability to reproduce modelled baseflow behaviours for some catchments in all three calibration scenarios reflects a structural limitation of the RDFs to perform satisfactorily for all cases.

From the correlation analyses, the catchments where the RDFs perform better have higher saturated hydraulic conductivity, higher BFI, lower surface runoff, and lower CQV of daily streamflow, which are inter-related. One interpretation for the observed dependency is based on BFI, which can be thought of as a measure of the amount of baseflow information or 'signal' in streamflow hydrographs. When the baseflow signal is low, it is generally more difficult to estimate, regardless of the estimation method. Another interpretation is based on our conceptual picture given by the Venn diagram of Fig. 1. On one hand, the catchments, where the RDFs show very high agreement with model simulations, constitute the overlapping subspace (i.e. intersection) between GEN(I,J) (or EX(M,N) and their nonlinear counterparts) and the model, and these catchments have high BFIs and its associated characteristics. On the other hand, the other catchments with lower BFIs are the model realizations that reside outside the subspace occupied by the linear

and nonlinear RDFs in the Venn diagram. Accordingly, this evaluation is inherently limited to identifying and characterizing the overlapping subspaces of baseflow realizations between the HGS model and the RDFs.

With this in mind, this synthetic evaluation with HGS catchments provides some guidance on identifying catchments where the RDFs can possibly perform with $\mathrm{NSE}_l \geq 0.5$, based on its BFI, runoff coefficient and rainfall/PET ratio beforehand. The latter two are most useful because they can be derived from accurate ancillary meteorological data. Given that the HGS models represent approximations to reality, further evaluation is warranted. The modelled catchments possess uniformity in geology and vegetation cover, vertically uniform soils, and simple and symmetric topography, which should lead to more linear hydrological responses to meteorological forcing compared with real catchments. These simplified and uniform characteristics will also limit the occurrence of other slow flow components, such as interflow and bank storage, that occur more strongly in heterogeneous catchments. For instance, interflow will commonly occur where there are permeability contrasts at the base of the soil profile, and bank storage where there are contrasts between near river sediments and the regional unconfined aquifer.

The influence of filter parameter errors is obvious in the linear implementations (i.e., *C1* and *C3*) of the RDFs. For these implementations, the chosen filter parameter values are directly associated with the estimated BFI of the filtering period through the identities in Eqs. 12, 29, and 33. In other words, a given BFI estimate is dependent only on the calibrating data. At catchments where the BFI values in calibration and filtering periods are different, the estimated BFI values for the filtering periods are expected to be biased. In practices where Eckhardt filter parameters are specified based on catchment

characteristics without calibration [*Eckhardt*, 2008, e.g.], the BFI value during the filtering period is already determined prior to filtering.

This work is restricted to the particular forms and implementations of the RDFs that use streamflow as their only filter input and filter parameters are time-invariant. Nonlinear, time-variant filters based on using additional input data, non-linear storages or time-varying storage-discharge relations are not considered and would occupy different solution subspaces in the Venn diagram (Fig. 1). In particular, contrary to the use of time-invariant filter parameters, it may be beneficial to consider time-varying filter parameters within an even more generalized RDF framework to overcome temporal changes in hydraulic properties of a catchment. This would come at the cost of requiring guidance of how to apply highly parameterized filters in practice.

5.3. Choosing between different RDFs

Our evaluation demonstrates the advantages of higher-order RDFs over the commonly-used Eckhardt filter in all three calibration scenarios, albeit to varying extents. Based on BIC, GEN(1,1), GEN(2,2), EX(1,1) and EX(2,2) filters have more flexible filter structures that better match the HGS simulations, especially for the catchments with small baseflow components. For example, we envisage that higher-order RDFs would better simulate catchments where baseflow is driven by multiple stores or where thresholding behaviour occurs within stores. Inspection of baseflow timeseries shows that the higher-order RDFs can reproduce strong high-frequency components existing in the daily HGS baseflow and, for the linear filtering implementation, this reduces the dependency of the baseflow estimates at later times on the choice of the initial baseflow value. These differences are responsible for reducing the biases in their baseflow estimates and improving their corre-

lation, relative to the Eckhardt filter. Although we have ignored the first-year of filtering data in computing the calibration cost function and evaluation metrics, the linear Eckhart filter took more than one year of filtering to recover from the mis-specification of the initial baseflow values.

The observed preference for higher-order RDFs over the Eckhardt filter is, however, conditioned on having perfect knowledge of baseflow for calibration and depends on the catchment and forcing characteristics. Limited and uncertain independent baseflow estimates (e.g., from tracers or salinity measurements) and errors in the streamflow data would challenge the calibration of higher-order RDFs because of increased numbers of unknown filter parameters. Moreover our results pertain to recovering daily baseflow estimates, with a view to understanding hydrologic processes across daily to longer time scales. Different applications of the RDFs, such as baseflow estimation at monthly or seasonal time scales for water resource management, would involve a different cost function. Compared to the calibration with daily data, this can be expected to lead to greater emphasis on correcting the biases in baseflow estimates, and may prefer the use of lower-order RDFs – these open questions are reserved for future investigation. Further, the application of BIC on thinned data for model selection is not ideal because the full data are not fully utilized and the issues such as time variation in autocorrelation in the residual and persistence in the autocorrelation are not directly addressed. Alternative methods based on autocorrelated and heteroscedastic error models for use in specifying likelihood functions [*Smith et al.*, 2010; *Evin et al.*, 2014] can be trialled. Compared to our simple least-squares based calibration, such models can also be expected to improve the calibration of the RDFs.

In our comparisons between the EX and GEN filters, both filters perform similarly at most catchments. In theory, the baseflow estimates by the EX filters can be replicated by the GEN filters, but not vice versa. However, by having fewer filter parameters requiring calibration, the EX filters have advantages over the GEN filters. Ignoring the integer-valued time delay parameters of the EX filters, a EX(M,N) filter has $3M + 2N - 1$ parameters, compared to $4M + 2N - 3$ parameters in an equivalent GEN filter. Inclusion of the delay parameters in the EX filters will increase the number of (equivalent) GEN's parameters more quickly than EX's. In our study, we have ignored the calibration of these delay parameters in the higher-order EX($\geq 2, \geq 2$) filters (Sec. 3.2) because of the computational costs associated with mixed-integer optimization. While there are many off-the-shelf optimization algorithms for calibrating the GEN filters' parameters, algorithms for mixed-integer optimization are far more limited at present; the optimization algorithm must search through integer values of the delay parameters and continuous values of the α and γ parameters while subject to the mass conservation constraint in Eq. 5. We have tested genetic algorithm for this purpose but found that the standard setup of the algorithm is not sufficiently robust to reproduce results obtained via brute force. Further work is needed to explore the benefits of the additional degree of freedom in the EX filters.

6. Conclusions

This paper extends the existing synthetic evaluations of RDFs for baseflow estimation. We unify the existing linear RDFs (Tabs. 1 and 2) through two theoretical frameworks. The first involves extending the formalism of [*Croke*, 2010] for multiple linear cascaded storages, and the second invokes the signal processing theoretic for IIR filters. The existing

RDFs can be seen as special cases of the EX filters, and by extension, the GEN filters. By nesting these existing RDFs within the two generalized forms, the complexity of a RDF can be systematically increased by increasing the number of flow pathways or by increasing the number of feedforward and feedback terms. This forms the basis for our collective assessment of RDFs against the daily HGS-simulated baseflow, leading to the following key conclusions.

First, the RDFs are capable of satisfactorily reproducing the simulated baseflow from virtual catchments with mid-range to high BFI $\geq 0.4 - 0.6$, lower runoff coefficients $\leq 0.15 - 0.25$ and average rainfall/PET ratios ≤ 0.81, depending on the specific calibration/implementation schemes. Despite increasing the orders of the RDFs, there remains persistent, and at times substantial, disagreement between the synthetic and estimated baseflow for many catchments with lower BFI, higher runoff coefficients and higher rainfall/PET ratios.

Second, the mass-balance condition within the RDF construction and the mass-balance condition at filter output cannot be satisfied simultaneously for any linear RDF acting on streamflow data only. This can be seen as an inherent structural flaw of the RDFs, which is mitigated by thresholding.

Third, the RDFs with higher filter order than the commonly-used Eckhardt filter are BIC-optimal at over 50 (out of 66) catchments. However, the magnitudes of performance improvements vary between the calibration/implementation schemes and are marginal at catchments where the Eckhardt filter already shows good performance. The higher-order RDFs have the advantages of producing short-time scale fluctuation and reducing the rise time of the filter response (i.e., the dependency on the initial baseflow state of the filter).

However, the observed preference for higher-order RDFs over the simpler filters is based on having perfect knowledge of baseflow for calibration and the specific application of estimating daily baseflow.

Further research is recommended on: (1) improving the sophistication of filter calibration and filter selection beyond least-squares metrics (e.g., defining suitable error models for likelihood functions); (2) extending the filter evaluation with other synthetic catchment simulations to validate our guidance for types of catchments where RDFs is a satisfactory method; (3) considering other practical filtering scenarios where baseflow is estimated at different spatio-temporal timescales; (4) investigating the trade-off between (calibration) data requirements and filter complexity; and finally (5) defining standards for applying RDFs, with the aim of finding appropriate interpretations of filter estimates in light of inherent ambiguities within filter construction and implementation.

Appendix A: Two-stage filtering for EX(M,N) filters

To explain the two-stage filtering process for EX filters, we first derive the generalized filtering equation for estimating effective rainfall timeseries. We combine Eq. 24 with Eq. 26 to find,

$$U = \frac{1}{\mathcal{C} + \mathcal{D}} Y. \tag{A1}$$

Upon substituting Eq. A1 into Eq. 24, it is straightforward to see that $H_{\mathrm{EX}(M,N)}$ in Eq. 28 is the product of two transfer functions, namely

$$H_{Y \to U} = \frac{1}{\mathcal{C} + \mathcal{D}}, \tag{A2}$$

$$H_{U \to S} = \mathcal{C}. \tag{A3}$$

In words, $H_{Y\to U}$ acts on Y to convert the streamflow to effective rainfall U, and $H_{U\to S}$ acts on the resultant U to convert the effective rainfall to baseflow S. Thus this can be treated as a two-stage filtering. The time-domain filtering equations for deriving u_n and, subsequently, s_n in the two stages can be derived by substituting Eqs. 25 and 27 to Eqs. A1 and 24, respectively, followed by inverse z-transform as before.

At the first stage, the constraint by observed rainfall can be imposed on u_n via thresholding, to derive a different effective rainfall timeseries $\tilde{u}_n$. Analogous to Eq. 13, we have

$$\tilde{u}_n = \min[p_n, u_n]. \tag{A4}$$

This assumes that there is no timing difference (e.g., due to routing) between u_n and p_n. Using $\tilde{u}_n$ in place of u_n as the input for the second stage modifies the resultant baseflow estimates. Each stage is therefore accompanied by thresholding to ensure effective rainfall and baseflow estimates are less than the observed values. *Croke* [2010] noted that this additional constraint gives the filter a stronger physical basis. However, as with the first thresholding defined in Eq. 13, it leads to nonlinear filtering and violates the internal consistency of and the input-output mass balance within the filter construction.

Acknowledgments. The HGS simulation data is available online as the Supplementary data to *Li et al.* [2014]; last accessed on March 1, 2016 via http://www.sciencedirect.com/science/article/pii/S1364815213003149. This research was conducted with financial support from the Australian Research Council (ARC Discovery Project No. DP120100253).

References

Beck, H. E., Van Dijk, A. I. J. M., Miralles, D. G., De Jeu, R. A. M., Bruijnzeel, L. A., McVicar, T. R., and Schellekens, J. (2013), Global patterns in baseflow index and recession based on streamflow observations from 3394 catchments, *Water Resour. Res.*, 49, 1-21, doi:10.1002/2013WR013918.

Boughton, W. C. (1993), A hydrograph-based model for estimating the water yield of ungauged catchments *In Hydrology and Water Resources Symposium*, Institution of Engineers Australia, Newcastle, NSW; 317-324.

Carrillo, G., Troch, P. A., Sivapalan, M., Wagener, T., Harman, C., and Sawicz, K. (2011), Catchment classification: hydrological analysis of catchment behavior through process-based modeling along a climate gradient, *Hydrol. Earth Syst. Sci.*, *15*, , 3411–3430, doi::10.5194/hess-15-3411-2011.

Cartwright, I., Gilfedder, B., and Hofmann, H. (2014), 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.

Chapman, T. (1999), A comparison of algorithms for stream flow recession and baseflow separation, *Hydrol. Processes*, *13*, 701–714, doi:10.1002/(SICI)1099-1085(19990415)13:5¡701::AID-HYP774¿3.0.CO;2-2.

Chapman, T. G., and Maxwell, A. I. (1996), Baseflow separation comparison of numerical methods with tracer experiments. Institute Engineers Australia National Conference. Publ. 96/05, 539–545.

Costelloe, J. F., Peterson, T. J., Halbert, K., Western, A. W., and McDonnell, J. J. (2015), Groundwater surface mapping informs sources of catchment baseflow, *Hydrol.*

Earth Syst. Sci., *19*, 1599–1613, doi:10.5194/hess-19-1599-2015.

Croke, B.F.W. (2010), Exploring changes in catchment response characteristics: application of a generic filter for estimating the effective rainfall and unit hydrograph from an observed streamflow timeseries. In: BHS Third International Symposium: Role of Hydrology in Managing Consequences of a Changing Global Environment. Newcastle University, UK.

Duan, Q., Sorooshian, S., and Gupta, V. (1992), Effective and effi- cient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, *28*(4), 1015–1031.

Eckhardt, K. (2005), How to construct recursive digital filters for baseflow separation, *Hydrol. Processes*, *19*, 507–515, doi:10.1002/hyp.5675.

Eckhardt, K. (2008), A comparison of baseflow indices, which were calculated using several different baseflow separation methods, *J. Hydrol.*, *352*(1-2), 168-173, doi: 10.1016/j.jhydrol.2008.01.005.

Eckhardt, K. (2012), 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.

Evin, G., M. Thyer, D. Kavetski, D. McInerney, and G. Kuczera (2014), Comparison of joint versus postprocessor approaches for hydrological uncertainty estimation accounting for error autocorrelation and heteroscedasticity, *Water Resour. Res.*, *50*, 2350–2375, doi:10.1002/2013WR014185.

Ferket, B. V. A., Samain, S., and Pauwels, V. R. N. (2010), Internal validation of conceptual rainfallrunoff models using baseflow separation, *J. Hydrol.*, *381*(1–2), 158–173, doi:10.1016/j.jhydrol.2009.11.038.

Furey, P. R., and Gupta, V. K. (2001), A physically based filter for separating base flow from streamflow time series, *Water Resour. Res.*, *37*(11), 2709–2722, doi: 10.1029/2001WR000243.

Furey, P. R., and Gupta, V. K. (2003), Test of two physically based filters for base flow separation, *Water Resour. Res.*, *39*(10), 1297–2722, doi:10.1029/2002WR001621.

Goldberg, D. E. (1989), Genetic Algorithms in Search, Optimiza- tion & Machine Learning, Addison-Wesley, Reading, MA.

Gonzales, A. I., Nonner, J., Heijkers, J., and Uhlenbrrok, S. (2009), Comparison of different base flow separation methods in a lowland catchment, *Hydrol. Earth Sys. Sci.*, *13*, 2055–2068, doi:10.5194/hess-13-2055-2009.

Halford, K. J. and Mayer, G. C. (2000), Problems associated with estimating ground water discharge and recharge from stream-discharge records. *Groundwater*, *38*, 331-342. doi:10.1111/j.1745-6584.2000.tb00218.x.

Huyck, A. A. O., Pauwels, V. R. N., and Verhoest, N. E. C. (2005), A base flow separation algorithm based on the linearized Boussinesq equation for complex hillslopes, *Water Resour. Res.*, *41*, W08415, doi:10.1029/2004WR003789.

Jakeman, A. J., and G. M. Hornberger (1993), How much complexity is warranted in a rainfall-runoff model? *Water Resour. Res.*, *29*(8), 2637-2649, doi:10.1029/93WR00877.

Kalbus, E., Reinstorf, F., and Schirmer, M. (2006), Measuring methods for groundwater-surface water interactions: A review, *Hydrol. Earth Syst. Sci.*, *10*, 873–887, doi: 10.5194/hess-10-873-2006.

Li, L., Maier, H. R., Lambert, M. F., Simmons, C. T., and Partington, D. (2013), Framework for assessing and improving the performance of recursive digital filters for baseflow

estimatin with application to the Lyne and Hollick filter, *Environ. Model. and Soft.*, *41*, 163–175, doi:10.1016/j.envsoft.2012.11.009.

Li, L., Maier, H. R., Partington, D., Lambert, M. F., and Simmons, C. T. (2014), Performance assessment and improvement of recursive digital baseflow filters for catchments with different physical characteristics and hydrological inputs, *Environ. Model. and Soft.*, *54*, 39–52, doi:10.1016/j.envsoft.2013.12.011.

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

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(7)*, 886–898.

Partington, D., P. Brunner, S. Frei, C. T. Simmons, A. D. Werner, R. Therrien, H. R. Maier, G. C. Dandy, and J. H. Fleckenstein (2013), Interpreting streamflow generation mechanisms from integrated surface-subsurface flow models of a riparian wetland and catchment, *Water Resour. Res.*, *49(9)*, 5501–5519.

Partington, D., Brunner, P., Simmons, C.T., Werner, A.D., Therrien, R., Maier, H.R., Dandy, G.C. (2012), Evaluation of outputs from automated baseflow separation methods against simulated baseflow from a physically based, surface water-groundwater flow model, *J. Hydrol.*, *458–459*, 28–39.

Peters, E., and Van Lanen, H. A. J. (2005), Separation of base flow from streamflow using groundwater levels illustrated for the Pang catchment (UK), *Hydrol. Processes*, *19*,

921–936, doi:10.1002/hyp.5548.

Pettyjohn, W. A., and Henning, R. (1979), Preliminary estimate of groundwater recharge rates, related streamflow and water quality in Ohio. Ohio State University Water Resources Centre Project Completion Report No 552, 323.

Rimmer, A., and Hartmann, A. (2014), Optimal hydrograph separation filter to evaluate transport routines of hydrological methods, *J. Hydrol.*, 514, 249–257, doi: 10.1016/j.jhydrol.2014.04.033.

Sloto, R. A., and Crouse, M. Y. (1996), HYSEP: A computer program for streamflow hydrograph separation and analysis, US Geological Survey, Water Resources Investigations Report 96–40.

Smith, T., A. Sharma, L. Marshall, R. Mehrotra, and S. Sisson (2010), Development of a formal likelihood function for improved Bayesian inference of ephemeral catchments, *Water Resour. Res.*, *46*, W12551, doi:10.1029/2010WR009514.

Stewart, M., Cimino, J., and Ross, M. (2007), Calibration of base flow separation methods with streamflow conductivity. *Ground Water*, *45*(1), 17–27, doi:10.1111/j.1745-6584.2006.00263.x.

Therrien, R., McLaren, R. G., Sudicky, E. A., Panday, S. M. (2009), HydroGeoSphere: A three-dimensional numerical model describing fully-integrated subsurface and surface flow and solute transport: Waterloo.

Van Dijk, A. I. J. M. (2010), Climate and terrain factors explaining streamflow response and recession in Australian catchments, *Hydrol. Earth Syst. Sci*, *15*, 159–169, doi: 10.5194/hess-14-159-2010.

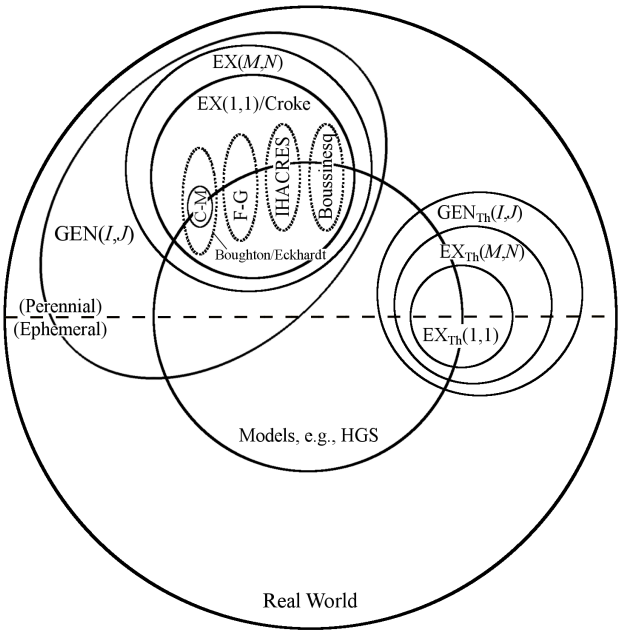

Figure 1. Venn diagram showing relationships between baseflow timeseries in the real world, estimated baseflow produced by different RDFs and models such as HGS. GEN(I,J) and EX(M,N) are linear filters, whereas GEN$_{Th}$(I,J) and EX$_{Th}$(M,N) denote their nonlinear counterparts when they were implemented with the constraint in Eq. 13.

Westra, S., M. Thyer, M. Leonard, D. Kavetski, and M. Lambert (2014), A strategy for diagnosing and interpreting hydrological model nonstationarity, *Water Resour. Res.*, *50(6)*, 5090-5113, doi:10.1002/2013WR014719.

Wittenberg, H (1999), Baseflow recession and recharge as nonlinear storage processes, *Hydrol. Process.*, *13*, 715–726, doi:10.1002/(SICI)1099-1085(19990415)13:5¡715::AID-HYP775¿3.0.CO;2-N.

Zhang, R., Li, Qiang, Chow, T. L., Li, S., and Danielescu, S. (2013), Baseflow separation in a small watershed in New Brunswick, Canada, using a recursive digital filter calibrated with the conductivity mass balance method, *Hydrol. Processes*, *27*, 2659–2665, doi: 10.1002/hyp.9417.

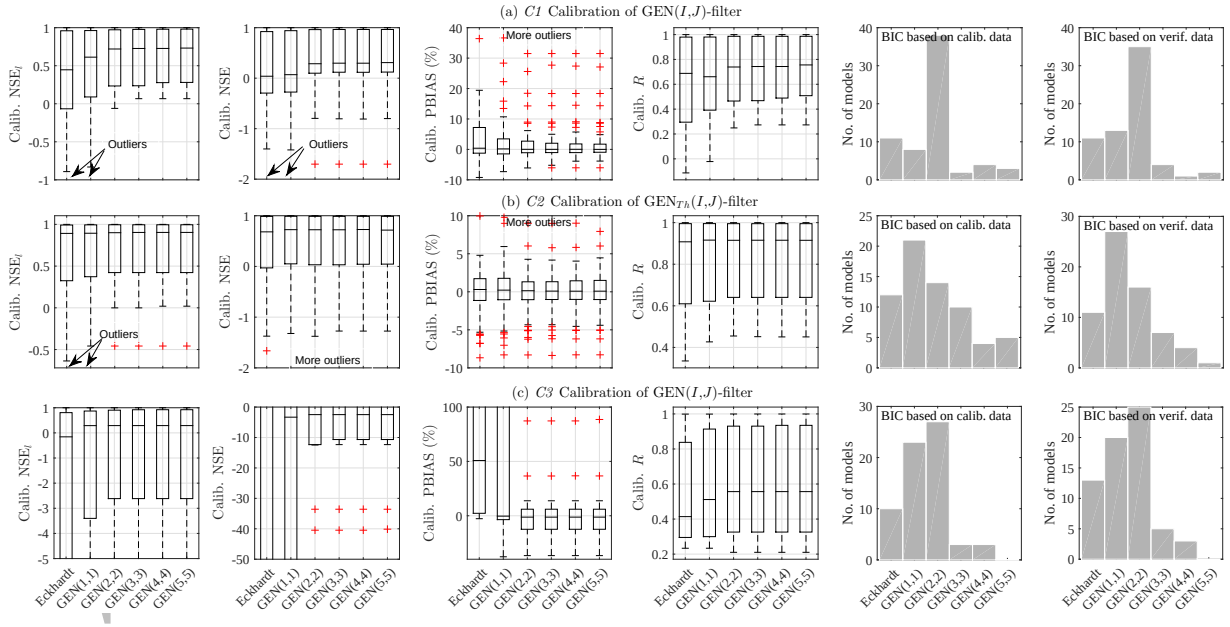

Figure 2. Comparison of the calibration results of the linear GEN(M,N) filters and nonlinear GEN$_{Th}$(M,N) filters under three different calibration scenarios. For each case, the NSE$_l$, NSE, PBIAS and correlation R for the 66 catchments are summarized with boxplots. Note that the calibration schemes are not directly comparable because the metrics were computed differently for *C2*. Filter selection based on the BIC analyses of the calibration and verification data are also shown in the last two columns.

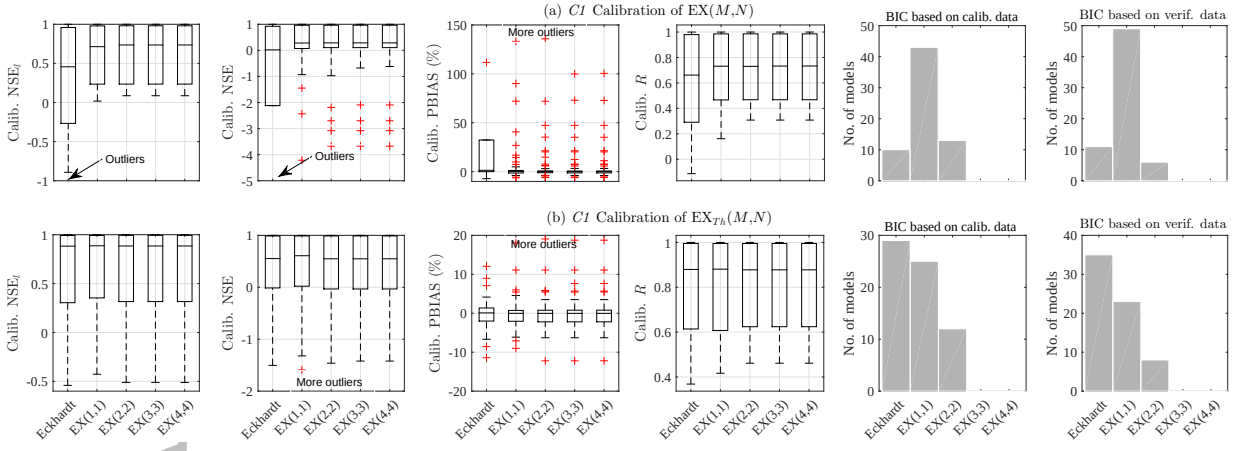

Figure 3. Same as Fig. 2, but for the linear EX(M,N) filters and nonlinear EX$_{Th}$(M,N) filters, and only *C1* and *C2* are considered. Note that the subscript 'Th' is dropped for *C2* plots for brevity.

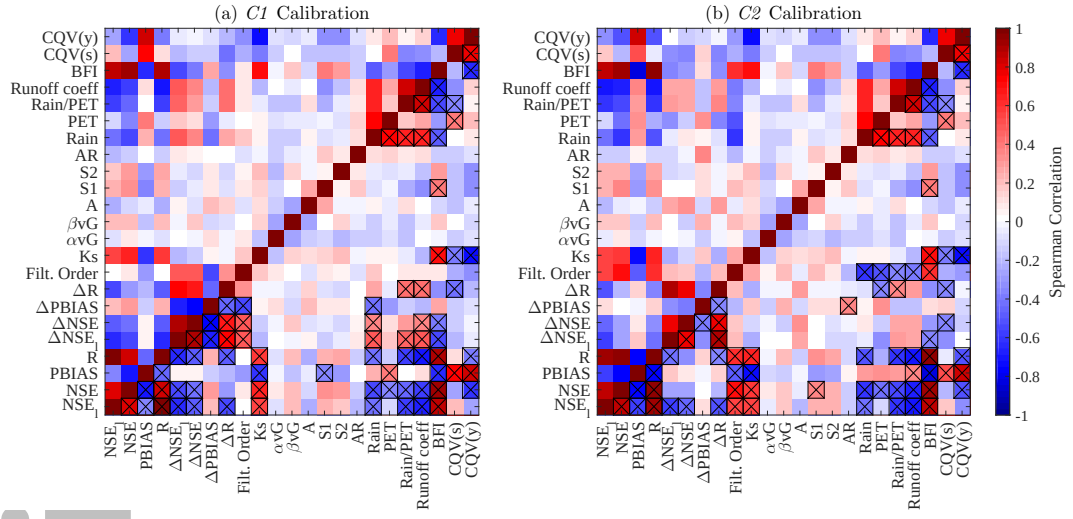

Figure 4. Spearman correlation matrices between optimal filter absolute performance, relative performance with respect to the Eckhardt filter, and catchment and flow characteristics. The cross-correlation values marked with a cross have a significance p-value < 0.005.

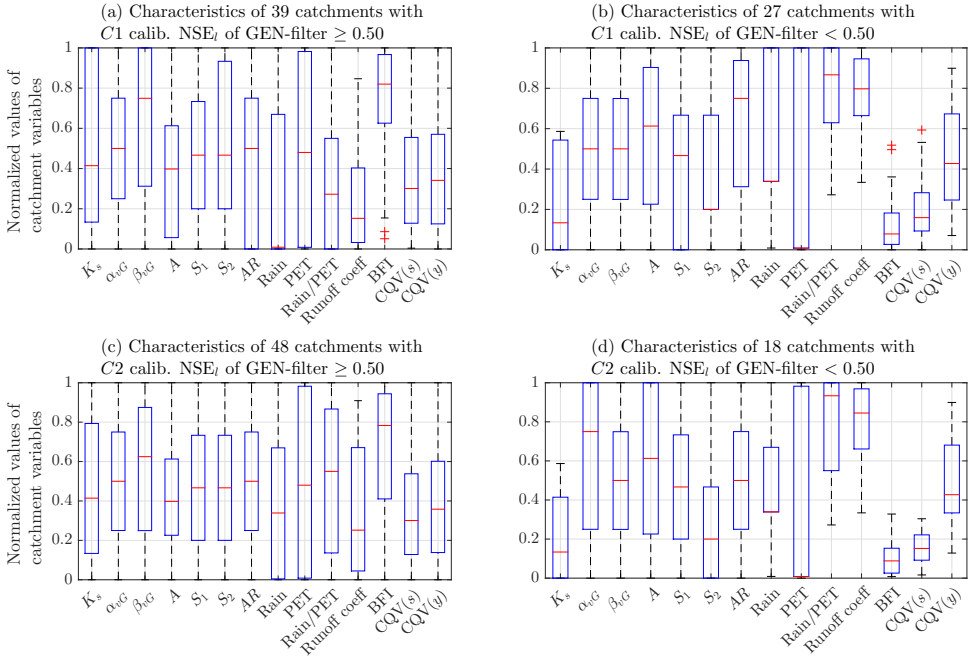

Figure 5. Characteristics of the HGS catchments for which the GEN filters does not perform satisfactorily with $\text{NSE}_l < 0.5$ (a,c) and those where $\text{NSE}_l \geq 0.5$ (b,d). The values of the catchment variables are linearly normalized between 0 and 1, except for K_s that is log-normalized.

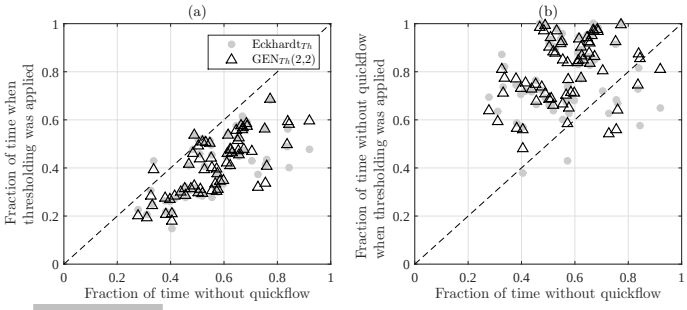

Figure 6. The extent of thresholding (Eq. 13) that was applied in the nonlinear filter implementations. (a) compares the fraction of baseflow values estimated via thresholding compared to the fraction of time when there is zero quickflow in discharge. (b) compares the fraction of baseflow values estimated in this way during the zero-quickflow periods.

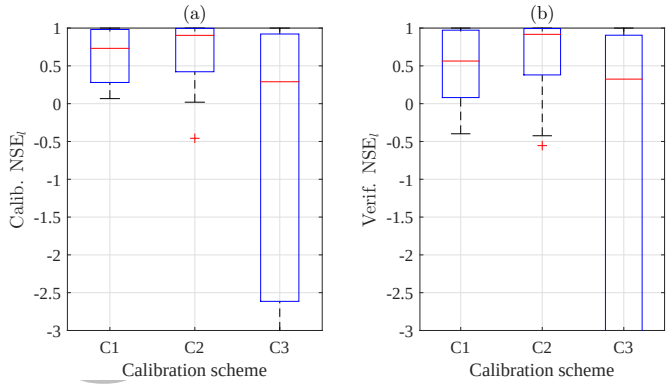

Figure 7. Summary statistics of NSE$_l$ of the baseflow estimates at the 66 catchments, derived from the three calibration schemes. Contrary to Fig. 2, the performance of the three schemes is directly comparable here as the metric is computed on coincident calibration and verification data.

Table 1. Mapping the EX filter in Eq. 6 and Eq. 28 to other recursive filters in the literature. The coefficients $\alpha_q, \alpha_s, \gamma_r$ and γ_s are expressed in terms of the variables found in the associated references.

Filters	Ref.	α_q	α_r	α_s	γ_r	δ_r	δ_s	EX(M,N)	Filt. Order
C-M	Eq. 6 in *Chapman and Maxwell* [1996]	0	0	$-k$	0.5	0	δ_q	EX(1,1)	1
Boughton	Eq. 7 in *Chapman and Maxwell* [1996]	0	0	$-k$	$\frac{C}{C-k+1}$	0	δ_q	EX(1,1)	1
Eckhardt	Eq. 19 in *Eckhardt* [2005]	0	0	$-a$	$\mathrm{BFI}_{\max}$	0	δ_q	EX(1,1)	1
F-G	Eq. 22 in *Furey and Gupta* [2001]	0	0	$\gamma - 1$	$\frac{c_3}{c_1+c_3}$	0	$\delta_q + d + 1$	EX(1,1)	$d+1$
IHACRES	*Jakeman and Hornberger* [1993]	const.	0	const.	const.	0	δ_q	EX(1,1)	1
Boussinesq	Eq. 19 in *Huyck et al.* [2005]	$-\frac{\psi}{\eta}$	0	$-\exp(B\Delta t)$	$\frac{c_3(\eta+\psi)}{c_3(\eta+\psi)+A\beta c_1(1-\exp(B\Delta t))}$	0	$\delta_q + d$	EX(1,1)	$d+1$

Table 2. Mapping the GEN filter in Eq. 30 to other recursive filters in the literature. The coefficients a_i and b_j are expressed in terms of the variables found in the associated references listed in Tab. 1.

Filters	a_i	b_j	GEN(I,J)	Filt. Order
C-M	$a_1 = \frac{a}{2-a}$	$b_0 = \frac{1-a}{2-a}$	GEN(1,0)	1
Boughton	$a_1 = \frac{C}{1+C}$	$b_0 = \frac{C}{1+C}$	GEN(1,0)	1
Eckhardt	$a_1 = \frac{a(1-\mathrm{BFI_{max}})}{1-a\mathrm{BFI_{max}}}$	$b_0 = \frac{\mathrm{BFI_{max}}(1-a)}{1-a\mathrm{BFI_{max}}}$	GEN(1,0)	1
F-G	$a_1 = (1-\gamma), a_{d+1} = -\gamma\frac{c_3}{c_1}$	$b_{d+1} = \gamma\frac{c_3}{c_1}$	GEN($d+1$,$d+1$)	$d+1$
IHACRES	$a_1 = \text{const.}$	$b_0 = \text{const.}, b_1 = \text{const.}$	GEN(1,1)	1
Boussinesq	$a_1 = \exp(B\Delta t), a_d = -\frac{c_3\eta}{c_1 A\beta}, a_{d+1} = \frac{c_3\psi}{c_1 A\beta}$	$b_d = \frac{c_3\eta}{c_1 A\beta}, b_{d+1} = -\frac{c_3\psi}{c_1 A\beta}$	GEN($d+1$,$d+1$)	$d+1$

Table 3. Parameters for the synthetic HGS catchment models from *Li et al.* [2014]. The parameters A, $S1$, $S2$, AR, K_s, α_{vG}, and β_{vG} are varied to produce 66 test catchments for this evaluation study.

Parameter	Value(s)
Geometry	
Catchment area A (km^2)	6, 48, 80, 120, 192
Hill slope $S1$	0.005, 0.008, 0.012, 0.016, 0.02
Channel slope $S2$	0.0025, 0.004, 0.006, 0.008, 0.01
Aspect ratio of width to length AR	0.5, 0.75, 1.0, 1.25, 1.5
Surface	
Manning's roughness (s/m$^{1/3}$)	0.015
Rill storage height (m)	0.001
Obstruction storage height (m)	0.0
Transpiration fitting parameter c1, c2, c3	0.3, 0.2, 10
Leaf area index	2.08
Wilting point	0.1
Field capacity	0.15
Oxic limit	0.9
Anoxic limit	1.0
Limiting saturation min, max	0.2, 0.32
Canopy storage parameter (m)	0.0
Initial interception storage (m)	0.0
Subsurface	
Evapotranspiration depth[a](m)	3
Root depth[a](m)	6
Saturated hydraulic conductivity (K_s) [m/s]	2.44E-5, 3.99E-5, 1.12E-4, 2.11E-4, 9.66E-4
van Genuchten α_{vG} (1/m)	0.572, 3.366, 6.161, 8.955, 11.75
van Genuchten β_{vG}	1.32, 1.56, 1.79, 2.03, 2.27
Surface-subsurface coupling	
Coupling length (m)	0.001

[a] Quadratic decay function