Reproducing different types of changes in hydrological indicators with rainfall-runoff models

Carolina Massmann

ABSTRACT

Hydrological indicators support analyses about the impact of climate and anthropogenic changes on riverine ecosystems. As these studies often rely on hydrological models for estimating the future value of the indicators, it is important to investigate how well, and under which conditions, we can replicate changes in the indicators. This study looks at these questions by investigating the performance that can be achieved depending on the objective function for calibrating the model, the direction of the change in the indicator, the magnitude of this change and the properties of the catchments. The results indicate that, in general, indicators describing the magnitude of discharge (monthly and annual) can be adequately estimated with hydrological models, but that there are difficulties when estimating the characteristics of flow pulses, flow reversals and timing variables. For some of these indicators, it is not even possible to correctly estimate the direction of large changes. The analysis showed further that these problems cannot be resolved by adjusting the calibrated parameters, but that the model structure is unsuitable for modelling these indicators.

Key words | ecohydrology, hydrologic change, hydrological indicators, indices of hydrological alteration, model evaluation, rainfall-runoff model

INTRODUCTION

The expected economic and population growth rates, coupled with the effect of climate change, raise fears that the number and severity of water-related conflicts might increase in the future. When dealing with competing uses of water, it is necessary that societies agree on the water quality and quantity allocated to fulfilling different needs and that they define, based on these agreements, clear targets for managing this resource (Chopra et al. 2005). The defined targets will depend, among other factors, on the current state and level of use of the aquatic resources as the closer the systems are to their natural state, the more likely that they will be managed for preserving ecosystems and their biodiversity (Acreman et al. 2014a). Accurate assessments of the current state of river ecosystems are thus important for supporting the definition of management objectives. Furthermore, such assessments constitute the basis for estimating how ecosystems could be affected by future interventions or changes (Acreman et al. 2014a), playing a prominent role in preparing for the future.

One important aspect these assessments need to consider is the health of river communities (i.e., plants, invertebrates and fish). However, due to the lack of available ecological data, such analyses are difficult to carry out (Acreman et al. 2014b). This is remedied by resorting to indicators describing the flow regime and using them as proxies for the ecological state by assuming that these indicators are
the main determinants of the ecosystem properties (Poff et al. 1997). This allows for comparing the values of the indicators, for instance, at two points in time, and concluding that there is no impact on the river ecosystem if these values do not vary significantly (Acreman et al. 2014b). While it is recognized that this approach leaves out important factors influencing riverine communities (e.g., water temperature, sediment load), it is not unreasonable to do so as these factors tend to be correlated with river discharge (Colby 1956; Sinokrot & Gulliver 2000; Acreman et al. 2014b), which is regarded as the most important factor.

While the hydrological indicators can be directly calculated from discharge data, it is necessary to rely on data-based statistical approaches or hydrological models when discharge data are unavailable. Data-based approaches have been widely used for estimating long-term hydrological indicators by establishing links between catchment properties (e.g., soil properties or average precipitation) and the hydrological indicators using, for instance, regional regressions (Sanborn & Bledsoe 2006; Carlisle et al. 2010; Knight et al. 2012). Hydrological models, on the other hand, use climate time series as model inputs and provide discharge time series as output. The hydrologic indicators can then be calculated from the simulated hydrographs (Shrestha et al. 2014; Caldwell et al. 2015). While both approaches have their merits and disadvantages (see detailed account in Murphy et al. 2013), hydrological models tend to be better suited for carrying out more detailed analyses (e.g., annual variability) at more local scales (see Olsen et al. 2013), while statistical models might be preferred for obtaining a general picture or a first approximation of the long-term averages at more regional scales (see Carlisle et al. 2010). The present study ultimately aims at improving our abilities for estimating changes in the hydrological properties of streams and using this information for adapting to climate change. As this requires detailed local analyses, the study focuses on model-based rather than on data-based indicator estimates.

Previous studies investigating the factors influencing the quality of estimated hydrological indicators obtained with modelled discharge time series have focused primarily on the impact of different model calibration strategies and objective functions. This is not surprising since this is a practical question practitioners and researchers are often confronted with when working with more than one indicator as it has been observed that models tend to replicate the characteristics to which they are calibrated well, while they might have a quite bad performance for estimating other metrics (Murphy et al. 2013). Many efforts have, therefore, gone into finding a small set of objective functions that allow a close reproduction of several indicators. While this question has not yet been settled, the results indicate that considering the indicator of interest in the objective function has a positive impact on the quality of the estimated indicators. It is possible to consider single indicators by calibrating the model directly to the indicator of interest (Kiesel et al. 2017; Pool et al. 2017) or to combine indicators in the context of a multi-objective Pareto calibration (Hernandez-Suarez et al. 2018) or with objective functions that combine various indicators (Vis et al. 2015; Kiesel et al. 2017; Pool et al. 2017).

An aspect that has not been thoroughly considered in this context is our ability for replicating changes in hydrologic indicators. This is somewhat surprising, given that the assessment of changes in river ecosystems is one of the main applications of hydrologic indicators and there is a vast body of literature looking at the performance of rainfall-runoff models in periods with different climate properties to those observed during the calibration period (Coron et al. 2012; Seiller et al. 2012; Brigode et al. 2013; Patil & Stieglitz 2015; Dakhlaoui et al. 2017; Vormoor et al. 2018). Many of these studies were inspired by the work of Klemeš (1986), who advocated for assessing the usefulness of models for estimating hydrographs under changed climate conditions with differential split-sample tests. In such tests, the model is calibrated for a subset of climate conditions and then tested on periods with contrasting climate conditions. Most of these studies compare ‘wetter’ periods with contrasting ‘drier’ periods as precipitation is the ultimate driver of discharge. If the analysis, however, does not focus on the annual or seasonal discharge, it might not be straightforward to identify contrasting periods a priori (Vormoor et al. 2018). This has implications when considering several indicators as it is expected that they will be sensitive to different climate characteristics, resulting in different sets of years characterizing contrasting climate properties depending on the indicator. In the only study addressing the ability of hydrological models for
reproducing changes for a set of hydrological indicators that was identified, Shrestha et al. (2016) considered different ‘cold’ and ‘warm’ ENSO (El Niño Southern Oscillation) and PDO (Pacific Decadal Oscillation) phases for four Canadian subbasins. While this might not necessarily result in highly contrasting periods for each indicator, it is an approach that can be easily justified and which avoids the need for a sensitivity analysis for each indicator. This study follows another route as it investigates our ability for modelling changes in the IHAs irrespective of the changes in climate. This is achieved by analysing changes in performance between different periods with contrasting IHA values and not periods with contrasting climate conditions.

Considering (i) the importance of knowing how well we can reproduce changes in the IHAs and (ii) the lack of studies investigating this question systematically for a large number of catchments, the objectives of this study are:

- Investigating if models calibrated to achieve a good overall performance outperform models calibrated towards specific hydrological indicators when modelling different types of changes in the indicators.
- Investigating if the performance for estimating changes in the hydrologic indicators varies with the direction and magnitude of the considered changes.
- Identifying possible links between catchment properties and our ability for modelling changes in the indicators.

A strength of the analysis presented here is that it considers a large number of catchments exhibiting considerable variations in climate and catchment properties for obtaining more robust and representative results.

**METHODOLOGY**

**Datasets and catchments**

This study analyses 560 catchments located across the United States with no discharge gaps between 1981 and 2008. They are representative of a large spectrum in climate and catchment properties as shown by the ranges covered by the mean annual precipitation (between 304 mm and 3,250 mm), mean annual discharge (between 2 mm and 3,508 mm) and aridity index (between 0.21 and 4.65).

These catchments belong to the GAGES II (Geospatial Attributes of Gages for Evaluating Streamflow, version II) dataset (Falcone 2011), which consists of gauges maintained by the USGS (United States Geological Survey) with at least 20 years of data or active in 2009. The gauges used in the present study are reference basins, which means they are among the least disturbed basins.

The climate data for the daily precipitation and the minimum and maximum temperatures were obtained from the Daymet 3 dataset (Thornton et al. 1997; Thornton et al. 2014) produced by the NASA Oak Ridge National Laboratory (ORNL). This dataset covers North America from 1980 to the present at a 1 × 1 km² resolution. The daily discharge time series were downloaded from the USGS webpage.

The catchments were characterized with respect to seven variables describing the climate (Walsh seasonality, aridity index, maximum 1-day annual precipitation, percentage of precipitation falling as snow), subsurface properties (baseflow index (BFI)), topography (average slope) and the water balance (runoff coefficient). The climate properties were estimated with the Daymet 3 data. The approach described by Walsh & Lawler (1981) was used for calculating the seasonality index and the aridity index is defined as the quotient between precipitation and potential evapotranspiration. The BFI was taken from a gridded dataset provided by the USGS (Wolock 2003) and the average slope was estimated from a 10 m resolution digital elevation model obtained from the National Elevation Dataset (NED).

**Hydrological model**

The conceptual rainfall-runoff model used in this study is the most complex member of a set of eight models exhibiting a stepwise increase in complexity. The models were initially developed by Jothityangkoon et al. (2001), Atkinson et al. (2002) and Farmer et al. (2003) and later modified by Bai et al. (2009). The required model inputs consist of the daily precipitation and the maximum and minimum daily temperatures. Interception is modelled as a fraction of the precipitation input and described by the parameter cei. A day-degree-based snow accumulation and melt routine (described by the parameters tt and dd) implemented for 100 m elevation bands, computes then
the liquid input to the soil bucket of size $sb$. The mechanisms available for water removal from this bucket are saturation overflow, subsurface flow (described by parameter $ass$), groundwater recharge (described by the coefficient $kd$) and evapotranspiration. The water reaching the deep flow bucket is routed to the stream as defined by the baseflow recession coefficient $abf$. The evapotranspiration is modelled as a function of soil moisture at the current timestep and the potential evapotranspiration, estimated according to the Hargraeves–Samani approach (Hargraeves & Samani 1982). Unlike most conceptual models, this model uses a parameter representing the field capacity ($fc$) for distinguishing a saturated and an unsaturated zone in the soil bucket. These zones interact differently with the potential evapotranspiration and differ in their way of estimating the fraction of the catchment covered by vegetation ($m$) and the fraction of bare soil ($1 - m$). A sketch of the model structure can be found in Massmann (2019).

The hydrological model was run at a daily timestep between 1 October 1981 and 30 September 2008. A three-year warm-up period was considered for stabilizing the state variables, leaving the period between 1 October 1984 and 30 September 2008 (24 years) for the analysis. The model was calibrated under a Monte Carlo (MC) framework that relied on 40,000 model runs for each catchment. For each of these model runs, the parameter values were randomly sampled assuming a uniform distribution between the ranges shown in the Supplementary Material (Table S1).

### Indices of hydrological alteration

For keeping the study manageable, it was necessary to select a subset of hydrological indicators. One option would have been to build on a recent study identifying the indicators correlating with the integrity of invertebrates and fish populations in different regions in the United States (Carlisle et al. 2011). However, since the correlated indicators vary depending on the region and between the invertebrate and fish communities, it would have been possible to focus only on one region. Instead, it was decided to select a subset of Richter’s indices of hydrological alteration (IHAs) which characterize different flow conditions and components of the flow regime (Richter et al. 1996). Flow condition refers here to the magnitude of discharge which could be classified, for instance, into average, low and high discharges. The five components of the flow regime affecting the ecological processes in rivers are the magnitude of the discharge, the frequency, duration and timing of certain flow magnitudes and the rate of change of the discharge (Richter et al. 1996). This study considers 12 indicators which attempt to cover a wide range of flow conditions (average, high, very high, low, very low) and all five regime descriptors (Table 1).

### Table 1 | Description of the indices of hydrological alteration used in the study

| Name   | Description                                      | Unit          | Regime property              | Flow condition |
|--------|--------------------------------------------------|---------------|-----------------------------|----------------|
| $Q_{Jan}$ | Mean annual flow January (winter)                | mm            | Magnitude, timing           | Average        |
| $Q_{Jul}$ | Mean annual flow July (summer)                   | mm            | Magnitude, timing           | Average        |
| $Q_{95}$ | 95 exceedance percentile                         | mm            | Magnitude, duration         | Low extreme    |
| $Q_{5}$       | 5 exceedance percentile                          | mm            | Magnitude, duration         | High extreme   |
| $Q_{year}$  | Mean annual discharge                            | mm            | Magnitude                    | Average        |
| $dHFP$   | Duration of high flow pulses                      | days          | Magnitude, duration         | High           |
| $dLFP$   | Duration of low flow pulses                       | days          | Magnitude, duration         | Low            |
| $nHFP$   | Number of high flow pulses                        | –             | Magnitude, frequency        | High           |
| $nLFP$   | Number of low flow pulses                         | –             | Magnitude, frequency        | Low            |
| Reversals| Number of reversals                               | –             | Rate of change              | Average        |
| DOY$x$   | Day of year maximum discharge                     | –             | Timing                      | High extreme   |
| DOY$n$   | Day of year minimum discharge                     | –             | Timing                      | Low extreme    |
While the information provided in Table 1 suffices for calculating most indicators, additional information might be necessary for understanding how the number of high flow pulses (HFPs) and low flow pulses (LFPs) was estimated. The approach used in this study consisted of taking the discharge time series for the 24 years considered in the analysis and estimating the 25th and the 75th flow percentiles. A pulse is then defined as an uninterrupted period during which the discharge is above the 75th percentile (HFP) or below the 25th percentile (LFP).

### Study set-up

#### Estimation of changes in the hydrological indices

The steps that need to be followed for calculating the changes in the IHAs are (Figure 1):

- Calculate the annual value of the indicator using the observed discharge (IHA_o).
- Calculate the annual value of the indicator (IHA_sim) and the annual value of the Nash–Sutcliffe Efficiency (NSE) for each of the 40,000 simulated discharge timeseries.
- Sort the annual IHA_o values and separate them into four groups with increasing IHA_o value. This is done for analysing different types of changes (e.g., from very low to high indicator values and vice versa). Randomly select five years in each group and calculate the average IHA_o, IHA_sim and NSE value for these 5 years. For IHA_o, these averages are identified as T_o,g with g standing for the group. For the IHA_sim and NSE, the 5-year averages are identified as T_h.g.c and T_n.g.c, respectively with c denoting the number of the MC run (i.e., ranging from 1 to 40,000). The usage of the average of five annual values, instead of one value accounting for the whole period, is inline of the use of the split KGE by Fowler et al. (2018), who found this results in consistently better performances during validation.
- Identify, for each group, the MC run within the highest NSE value and the lowest absolute IHA error (abs(IHA_sim - IHA_o)). Store these values in the vectors MCn and MCh, respectively, with MCn[g] and MCh[g] specifying the MC run with the highest NSE and lowest IHA error in group g, respectively.

- For all combination of the four groups (g1, g2, g3, g4) calculate the observed and simulated changes in the indicators. This results in 12 cases (i.e., g1 to g2, g1 to g3, g1 to g4, g2 to g1 and so on). The observed changes in IHAs are calculated as the difference in the observed IHA between both considered periods. For the modelled timeseries, two sets of changes are calculated. The first is based on the optimal Monte Carlo run with respect to the NSE, whereas the other set considers the optimal parameter set with respect to the IHA of interest.

#### Evaluation metrics

An essential aspect when assessing the performance of a model for reproducing changes is its ability to correctly reproduce the direction of change. This is assessed with the ‘agreement in direction’ (AiD) metric (Equation (1)):

$$AiD = \begin{cases} 1, & \text{if } (IHA_{sim,g2} - IHA_{sim,g1}) / (IHA_{o,g2} - IHA_{o,g1}) > 0 \\ 0, & \text{if } (IHA_{sim,g2} - IHA_{sim,g1}) / (IHA_{o,g2} - IHA_{o,g1}) \leq 0 \end{cases}$$

where IHA_{sim,g1} and IHA_{o,g2} stand for the simulated and observed IHA value in the groups specified by the subscripts g1 and g2. If the simulated and observed changes agree in their direction, the product will be positive and a will equal one. If the direction of change differs between the observed and simulated IHA, their product will be negative and AiD will equal zero.

It is further important to assess the magnitude of the error incurred when estimating changes in hydrological indicators. It can be distinguished between the nominal and the relative magnitude of the error. This differentiation is especially relevant when comparing catchments exhibiting large differences in discharge as errors in nominal units tend to be larger for catchments with higher discharge. On the other hand, the relative error (e.g., with respect to the mean annual discharge), tends to be larger for catchments with lower discharge (Murphy et al. 2013). This study considers both the nominal (NE) and the relative errors (RE),...
Figure 1 | Procedure for calculating the changes in the indicators for all 12 combinations of groups. \( T_{o,i} \) stands for the observed IHA value in group \( g \), while \( T_{n,g} \) and \( T_{h,g} \) stand for the 40,000 NSE and IHA values in group \( g \). \( MC_{n}[g] \) and \( MC_{h}[g] \) identify the optimal MC run with respect to the NSE and IHA in period \( g \).
described in Equations (2) and (3), respectively:

\[
NE = \frac{|(IHA_{sim,g2} - IHA_{sim,g1}) - (IHA_{o,g2} - IHA_{o,g1})|}{IHA_{o,md}}
\]

(2)

\[
RE = \frac{|(IHA_{sim,g2} - IHA_{sim,g1}) - (IHA_{o,g2} - IHA_{o,g1})|}{IHA_{o,md}}
\]

(3)

where \(IHA_{o,md}\) stands for the mean value of the observed IHA values in both considered groups.

**Data analysis**

It is expected that both the ability for estimating the direction and the magnitude of changes might vary depending on the magnitude and direction of the considered change. For instance, it might be more difficult to get the direction of change right, the smaller the change is. The ability for correctly reproducing the magnitude of change might also vary depending on the type of change (i.e., if it corresponds to a wet-to-dry or a dry-to-wet situation and to a large or small change). The analysis was therefore carried out for different percentiles of observed IHA changes.

For shedding some light on the factors affecting the performance of hydrological models for estimating IHA changes, the relationship between some catchment descriptors (see section ‘Datasets and catchments’) and the performance metrics (AiD, NE and RE) was analysed using Pearson’s correlation coefficient. This analysis was not carried out for the two variables describing the timing of events (DOYx and DOYn) as they are circular variables and thus require a different approach.

**RESULTS AND DISCUSSION**

**How do models calibrated with different objective functions differ in their ability to reproduce IHAs?**

The impact of the objective function on the calculated IHAs is shown in Figure 2, which depicts scatterplots relating the observed (x-axis) with the simulated IHAs (y-axis). The plots on the left were obtained with the Monte Carlo run achieving the highest NSE in each group (i.e., each plot has 560 (catchments) × 4 (groups) points), whereas the plots on the right side show the results when using the model run with the best value for the corresponding IHA.

The plots for the Q5 exceedance are similar to the plots for Q Jan, Q Jul, Q95 and Q year (Figure S1 in the Supplementary material), all of which describe the magnitude of the discharge (i.e., their units are mm). It can be seen that the IHAs calculated with the NSE calibrated model agree relatively well with the observed IHAs (left side plots). As the plots on the right side show the best IHA value that was achieved with the 40,000 Monte Carlo runs, it is not surprising that the observed and simulated values lie almost on the 1:1 line.

The patterns observed for the number of low pulses and reversals are similar to the patterns observed for the number of high pulses (Figure S1 in the Supplementary Material). While the plots on the right side agree better with the observed IHAs than the plots calibrated with the NSE, there is a considerable underestimation of the observed values even when the IHAs are used for calibrating the model. This underestimation is more pronounced for higher IHA values as the model is able to reproduce the values of the indicators at the lower end of the distribution to a certain extent, but there are large deficiencies when the IHAs exhibit higher values. For example, the model was unable to reach a value above 50 for the LFP indicator, while the observed values rise above 50. The fact that the hydrological model is unable to reproduce these features with any of the 400,000 MC runs indicates that the performance cannot be improved with more sophisticated calibration approaches, but that the problem lies in the model structure. The recent surge of interest in model structure-related issues has led to the development of modular hydrological models which enable researchers to compare and test different model structures (Clark et al. 2008; Fenicia et al. 2011) and to investigate the impact of different structural components on the hydrograph (Lane et al. 2019; Massmann 2020).

The identification of an adequate model structure for a given catchment is, however, still an unresolved problem. The reason for this is that besides model structural uncertainty there are also measurement errors (e.g., in the discharge time-series) and uncertainties in the climate data which is used for forcing the models. Since model parameters can compensate for errors in the forcing, discharge measurements and model structure, it is difficult to isolate the effect of model structures on model performance (Gupta & Govindaraju 2019).
Due to the similarity of the plots for the indicators describing the duration of the HFPs and LFPs, only the results for the HFPs are presented in the manuscript. The model is able to correctly simulate this indicator as there is good agreement between the observed and simulated values on the right-hand side plot. However, it is observed...
that the optimal parameter set with respect to the NSE is not a good choice for modelling these indicators as it results in large overestimations of the duration of the pulses.

As the plots for the DOY of maximum and minimum discharge have similar patterns, the results are only shown for the day of maximum discharge. It is seen that the NSE is not able to reproduce this feature well, but there is no consistent over- or underestimation. The results reached with the IHA calibrated models agree in general with the observed IHA, but there is a large scatter. It is further interesting to note that the model calibrated with the IHA can model the day with minimum discharge better than the day with maximum discharge, but that this is the other way around when the NSE is used as objective function during calibration.

In summary, the hydrological model is able to simulate the indicators describing the magnitude of the flow and the timing of extreme events, but it is inadequate for modelling the number of pulses and reversals, which tend to be underestimated. These results agree well with previous studies. Shrestha et al. (2014) found, for instance, that the correlation between observed and modelled IHAs describing the magnitude tended to be above 0.5, while most IHAs related to flow pulses and reversals had values below 0.5.

The next sections investigate if our ability for correctly modelling the IHAs describing the magnitude and timing properties of hydrographs is reflected in good estimates of the changes in these IHAs. Analogously, we want to find out if the difficulties of hydrological models for reproducing the number of pulses and reversals result in low performances for modelling changes in these indicators as it might be possible that estimates of the IHAs obtained with the models are biased, but that their changes can be nevertheless modelled with some confidence.

**Which objective function is most suitable for modelling IHA changes?**

Table 2 allows for assessing the performance of the NSE and the IHAs as objective functions for calibrating the hydrological model used for estimating the IHAs. The table shows, for each IHA, the percentage of cases in which the model calibrated with the corresponding IHA has a better performance than the model calibrated with the NSE. The performance is evaluated with respect to two criteria: the

| Type of change | Direction of change | Magnitude of change |
|----------------|---------------------|---------------------|
|                | − − − + + ++       | − − − + + ++       |
| Q Jan          | 13 19 39 28 33      | 36 29 30 41 45      |
| Q Jul          | 15 22 35 37 32      | 27 37 51 61 51      |
| Q95            | 24 37 30 46 38      | 45 44 69 60 60      |
| Q5             | 55 27 46 30 38      | 42 34 41 36 33      |
| Q year         | 71 15 51 29 100     | 45 44 40 44 33      |
| nHFP           | 9 17 44 54 46       | 13 20 41 63 53      |
| nLFP           | 50 60 45 28 34      | 68 65 57 84 86      |
| dHFP           | 29 34 59 68 61      | 32 35 48 68 65      |
| dLFP           | 52 55 43 39 48      | 81 85 91 93 84      |
| Reversals      | 65 58 60 76 79      | 64 56 46 74 80      |
| DOYx           | 60 56 51 63 57      | 56 57 54 59 55      |
| DOYn           | 40 46 50 67 62      | 42 47 61 63 58      |

Cells with white and grey background indicate where the NSE and the IHA achieved the best performance, respectively. Cells in light grey colour show no significant difference (p-value <0.05) according to the z-score of proportions test. The results are shown for five change percentiles. The first column shows the results for the largest negative changes, the central column for the smallest (negative and positive) changes and the last column for large positive changes.
agreement in the direction and magnitude of the modelled and observed IHA changes. The results are shown for five different types of changes, each comprising 20% of the total cases. The first (second) column shows the results for very large (large) negative changes, the central column shows the result for small changes (both negative and positive) and the last two columns show the results for large and very large positive changes. It is seen, for example, that all values for the January discharge (Q Jan) are below 50. This means that the model calibrated with the IHA tends to have a better performance than the model calibrated with the NSE in less than 50% of the cases.

There is a tendency for the objective function that is best for modelling the direction of change to have also the best performance for modelling the magnitude of change. It is further seen that for the direction of change there is a tendency of having values closer to 50 in the central columns and more extreme values (either closer to zero or closer to 100) towards the edges. This indicates that the differences between the two objective functions are less pronounced for smaller than for larger changes. This relationship is, however, less evident for the magnitude of change.

Changes in the IHAs describing flow magnitudes (i.e., the first five indicators in Table 1) are generally better modelled with parameter sets obtained with the NSE (rather than with the IHA) as objective function. This indicates that the focus of the NSE on achieving a good overall performance in reproducing the hydrograph makes this a good objective function for modelling changes in the annual flows, monthly flows and flow percentiles. On the other hand, it suggests that single hydrological indicators might not constrain well enough all parameters to which the indicators are sensitive.

For the indicators describing the pulses and reversals, no clear trend could be observed. Finally, the indicators describing the day of the year in which the maximum and minimum values are observed, tend to be better modelled when the calibration is done with the IHA. As the DOY is calculated using the discharge of only one day in each year, it is important that the model reproduces adequately the relative magnitudes of the peaks and low period flows. This seems to be best achieved by focusing just on the days with extreme values rather than with a model calibrated to reproduce adequately overall model performance.

How well can we estimate the direction of IHA changes?

Figure 3 informs about the agreement in direction between the observed and simulated changes for the IHAs calculated using the NSE as objective function. It can be seen that the direction of change can be best reproduced for the Q year indicator. The ability for reproducing the direction of change increases with the magnitude of the change until changes of about 1,500 mm. From there on, the ability of the model for identifying the direction of changes starts to decline and it is observed, for the largest changes, that the model is only capable of getting the direction of changes right about 50% of the time.

The patterns for the HFPs and LFPs are similar, although the direction of change of the LFPs is less likely to be correctly reproduced than the direction of change of the HFPs. It is further interesting to note that for the number of pulses there is a clear increase in the ability for modelling the direction of changes as the magnitude of the changes increases, while this is not the case for the duration of the pulses, where the ability for reproducing the direction seems to be independent of the magnitude of the change.

The size of the considered change has only a small impact on the ability for reproducing the indicators describing the timing of extreme flows (DOYx and DOYn). In agreement with the results shown for the pulses, which showed that difficulties for reproducing the direction of change were larger for the indicators describing low flows, it is seen here that the day of minimum discharge has a lower agreement in direction than the day of maximum discharge. Moreover, as the agreement in direction of the DOYn varies around 0.5, it can be concluded that the model has no skill for modelling changes of this indicator.

The impact of the objective function used for defining the best parameter set is observed in Figure 4, which shows the results when using the IHA as objective function. The most striking difference between these plots at the ones presented in Figure 3 is their asymmetry, which indicates that the results vary depending on the type of change (i.e., from high to low indicator values or vice versa), something which is not observed when the models are calibrated with the NSE.
How well can we estimate the magnitude of IHA changes?

The first four rows of plots in Figure 5 show the absolute nominal error (NE) incurred when modelling changes in the IHAs with the model run optimizing the NSE. The plot for Q year is representative for all indicators describing the discharge magnitude. It shows an increase in the error as the modelled changes become larger. This increase is small for smaller changes but rises more sharply for more extreme changes. The error of the number of pulses and reversals shows an almost linear increase with little scatter as the magnitude of the change increases. The reason for this can be understood when looking at Figure 1, which shows only small differences in the simulated values of these IHAs regardless of the observed IHA values.

Besides an analysis of the NE, it is important to look at the relative error as it might be easier to grasp the relevance of an error when comparing its magnitude, for example, to the mean value of the indicator. Another advantage of using the relative error is that it facilitates comparisons of model performance between indicators. The last three rows of plots in Figure 5 show the relative error for some indicators. The relative error for the indicators describing
the discharge magnitude is high for small changes. For large changes, the errors are low and almost independent of the size of the modelled change. The REs for the number of pulses tend to increase as the modelled changes increase, which can again be explained by the little variability of the modelled IHA, even as the observed values increase (see Figure 1). The relative error of the timing IHAs is not considered as all catchments have the same 365 days and there is thus no need for normalizing the DOYx and DOYn.

**Analysing the impact of the direction of change on the reproducibility of IHA changes**

Sections ‘How well can we estimate the direction of IHA changes?’ and ‘How well can we estimate the magnitude of IHA changes?’ show how our ability for modelling IHA changes varies with the magnitude and direction of the considered IHA change. In most cases, it was possible to observe differences as a function of the magnitude of the change; the impact of the direction of change is, however, more subtle and requires additional analyses.

Table 3 shows the quotient between the mean absolute negative error and the mean absolute positive error for five change percentiles. For example, the percentiles considered for the largest changes were 0.1–1 (negative changes) and 99–99.9 (positive changes). The Q Jan indicator has an average error of 125.2 mm in the 0.1–1 percentile and an error of 63.5 mm in the 99–99.9 percentile (Table S4 in the Supplementary Material). The quotient of the negative and positive errors is thus 1.97 (Table 3), indicating that the error for negative changes is almost twice as large as the error for positive changes. Analogously, all other values larger than one in Table 3 indicate that the errors for negative changes are larger than the errors for positive changes, whereas this is the other way around when the values in Table 3 are smaller than one. Since negative changes are encountered when the indicator goes from high to lower values, values above one in Table 3 indicate that changes can be better modelled when they are from low to high indicator values.

For almost all considered indicators and change percentiles, the model is able to better model changes from low to high indicator values. These differences on performance as a function of the direction of the modelled change are more pronounced when the model is calibrated with respect to the IHA than to the IHAs. For the indicators describing flow magnitudes (i.e., Q year), this means that it is easier to model changes when models are calibrated in dryer periods and then run in wetter periods than the other way around. This is in agreement with previous studies (Vaze et al. 2010; Coron et al. 2012).

Besides analysing the impact of the direction of change on the mean error (Table 3), it was investigated how the
Figure 5 | Absolute nominal error (NE) in the first four rows and relative error (RE) in the last three rows for all 12 cases in the 560 catchments and using the optimal model run with respect to the NSE. The results are shown for the running mean of 30 (dark colour) and 7 (light colour) consecutive simulated errors. The x-axis shows the observed value of the corresponding change in IHA in the units of the IHA (specified in the title of each plot).
The direction of the change impacts our ability for reproducing the direction of change (Tables S2 and S3 in the Supplementary Material). There is almost no influence of the direction of change on our ability for identifying the direction of change if the NSE is used as objective function. The differences are more pronounced when the model is calibrated with the IHAs and when modelling the indicators describing the flow dynamics (i.e., pulses and reversals). It is interesting

![Table 3](image)

![Table 4](image)

The results were estimated for five different change percentiles and for the NSE and IHA calibrated models.

The cells without values had non-significant correlation coefficients (p-value $\approx 0.05$).
to see that changes in the HFPs are better modelled when they are from high to low flows, while changes in LFPs can be better modelled when they are from low to high values.

Identification of catchment descriptors influencing the quality of the IHA estimates

Table 4 shows the results of the correlation analysis between the catchment descriptors and two evaluation metrics (the agreement in direction and the NE). In general, there is a high agreement in the sign of the correlation coefficient observed for different IHAs. A higher agreement in the modelled direction of change is observed for wetter catchments (i.e., higher maximum 1-day precipitation and lower aridity), as well as for catchments with lower BFI. As previous studies have found that it is more difficult getting a good performance in more arid (van Esse et al. 2015; Weingartner et al. 2013) catchments and in catchments with higher BFI (Massmann 2020), these results indicate that the ability for getting the direction of the changes right is related to our ability of getting a good overall model performance.

The NE of the IHAs describing the magnitude of discharge has a negative correlation with aridity and a positive correlation with the maximum 1-day precipitation, indicating that dryer catchments, with lower discharges, tend to also have lower nominal errors. There is further a positive correlation with the BFI. The correlation for the other catchment descriptors is difficult to interpret and it might be more helpful to analyse the patterns visually. Figure 6 shows the spatial distribution of the average NE for each catchment and some representative IHAs. The IHAs describing the flow magnitude (Q year) tend to have low errors in the dryer central Great Plains and higher errors in the wet western coast. The number of reversals, on the other hand, has higher errors in dryer catchments and lower errors in the wetter ones. The error in the number of HFPs is low in the snowy Rocky Mountains area, while reaching high values in the eastern part characterized by a more homogeneous precipitation pattern across the year. The error in the number of LFPs is lowest in the western coast, which has a high seasonality in the precipitation and higher in the remaining United States. The errors in the duration of the pulses are highest in the central part and decrease towards the coasts. Finally, it is seen that changes in the day of the maximum discharge (DOYx) can be best modelled in areas with a large influence of snow and in the western coast, with clear seasonal precipitation patterns. The errors when modelling changes in the day on minimum discharge (DOYn) do not have a distinct spatial pattern.

Limitations of the study

Investigating the ability of hydrological models to model changes in IHAs is important in the context of climate change studies. This study sheds some light on the ability of hydrological models for modelling changes in IHAs, but the results cannot be used directly for inferring about the ability for modelling the impact of climate change as the link between IHAs and the climate dynamics (i.e., information about climate sensitivity) was not considered.

SUMMARY AND CONCLUSIONS

This study investigated the performance of a conceptual rainfall-runoff model for estimating changes in indicators of hydrological alteration.

- The results agreed with previous studies showing that indicators describing the magnitude of discharge are better modelled than indicators describing the number of pulses and reversals. An interesting finding in this context was that the difficulties for modelling the number of flow pulses and reversals cannot be attributed to calibration deficiencies, but that the model is unable to reach the observed values. This suggests that, instead of focusing on different calibration alternatives, we should investigate and address the causes of model failure. It would be further interesting to find out if this is a limitation of the specific rainfall-runoff model used in this study or a more general deficiency of the type of model (e.g., are distributed models able to achieve better results than lumped models?). Another question that could be investigated is if the quality of the IHAs estimates can be improved by redefining the indicators, for example, by constraining the number of considered pulses and reversals.
It was found that our ability for modelling changes in hydrological indicators is correlated with our ability for modelling the corresponding indicators.

The quality of the estimates of IHA changes depends strongly on the considered indicator. The hydrological model was able to provide reliable estimates of the direction of the changes for the IHAs describing the discharge. While there are some difficulties in estimating the direction of the change when changes are small, the model achieves good performances as the magnitude of the changes increases. On the contrary, the model was not able to correctly estimate the direction of the changes for the duration of the pulses, the number of reversals and the day on minimum discharge, independently of the size of the considered change. For the number of pulses, the model had some skill for estimating the direction of

Figure 6 | Spatial distribution of the average nominal error (of all 12 cases) for each catchment and 8 representative IHAs.
large changes, but the estimates for smaller changes were unreliable.

- The magnitude of changes might be adequately estimated across the entire range of variation of the changes for indicators describing the discharge magnitude. For pulses and reversals, on the other hand, there is a clear reduction in performance with increasing magnitude of the considered changes, decreasing the potential applicability of this approach for estimating the impacts of climate change on the discharge dynamics. These results do not point to a shortcoming of the models for general hydrology, but to limitations when used for modelling hydrologic indicators under change. The results highlight, furthermore, the need for developing alternative methodologies for understanding how climate change could affect riverine ecosystems.

- Regarding the influence of the objective function used for calibrating the hydrological model on the ability of reproducing changes in the IHAs, it was observed that the results vary depending on the considered IHA. For indicators describing discharge magnitude, the results tended to be better when the NSE was used as objective function. For flow pulses, reversals and timing variables, on the other hand, the magnitude of the error tended to be lower when the IHA was used for calibrating the model, while there seemed to be no clear tendency when estimating the direction of the changes.

- An analysis of differences in the quality of the IHA estimates depending on the direction of the changes showed that changes from low to high indicator values could be better modelled than changes from high to low values. This effect was more pronounced when the models were calibrated with respect to the IHAs. This has some implications when modelling the impacts of climate change as we will calibrate our models for the current climate and the use for estimating the discharge for dryer conditions. Since it was found that these changes (wet to dry) are more difficult to get right, it is important to avoid overestimating our ability for modelling dryer periods.

- An exception is found for the agreement in direction for the indicators describing HFPs and calibrated with the IHA, where the direction of changes could be best predicted for changes from high to low indicator values.

- An important strength of this study is that it relies on many catchments and that it considers different types of changes (with respect to the magnitude and direction of the changes). Such a stratification of the changes allows for additional insight. For example, we found that models are able to model adequately low IHA values of the indicators describing the discharge dynamics, but have problems in reproducing high values.

**DATA AVAILABILITY STATEMENT**

The primary data used in the study are publicly available from U.S. organizations. More detailed information about how it can be accessed is found in the text. Processed data are available from the author upon request.

**ACKNOWLEDGEMENTS**

Three anonymous reviewers provided numerous comments and suggestions that helped to greatly improve the manuscript. This research was funded by the Austrian Science Fund (FWF): project J-3689-N29.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this paper is available online at https://dx.doi.org/10.2166/nh.2020.073.

**REFERENCES**

Acreman, M., Arthington, A. H., Colloff, M. J., Couch, C., Crossman, N. D., Dyer, F., Overton, I., Pollino, C. A., Stewardson, M. J. & Young, W. 2014a Environmental flows for natural, hybrid, and novel riverine ecosystems in a changing world. *Frontiers in Ecology and the Environment* 12 (8), 466–473. doi:10.1890/130134.

Acreman, M. C., Overton, I. C., King, J., Wood, P. J., Cowx, I. G., Dunbar, M. J., Kendy, E. & Young, W. J. 2014b The changing role of ecohydrological science in guiding environmental flows. *Hydrological Sciences Journal* 59, 3–4, 433–450. doi: 10.1080/02626667.2014.886019.

Atkinson, S. E., Woods, R. A. & Sivapalan, M. 2002 Climate and landscape controls on water balance model complexity over
changing timescales. *Water Resources Research* **38** (12). doi:10.1029/2002WR001487.

Bai, Y., Wagener, T. & Reed, P. 2009 A top-down framework for watershed model evaluation and selection under uncertainty. *Environmental Modelling and Software* **24**, 901–916. doi:10.1016/j.envsoft.2008.12.012.

Brigode, P., Oudin, L. & Perrin, C. 2015 Hydrological model parameter instability: a source of additional uncertainty in estimating the hydrological impacts of climate change? *Journal of Hydrology* **476**, 410–425. doi:10.1016/j.jhydrol.2012.11.012.

Caldwell, P. V., Kennen, J. G., Sun, G., Kiang, J. E., Butcher, J. B., Eddy, M. C., Hay, L. E., LaFontaine, J. H., Hain, E. F., Nelson, S. A. C. & McNulty, S. G. 2015 A comparison of hydrologic models for ecological flows and water availability. *Ecohydrology* **8**, 1525–1546. doi:10.1002/eco.1602.

Caldwell, D. M., Falcone, J., Wolock, D. M., Meador, M. R. & Norris, R. H. 2010 Predicting the natural flow regime: models for assessing hydrological alteration in streams. *River Research and Applications* **26**, 118–156. doi:10.1002/rra.1247.

Carlisle, D. M., Grantham, T. E., Eng. K. & Wolock, D. M. 2017 Biological relevance of streamflow metrics: regional and national perspectives. *Freshwater Science* **36** (4), 927–940. doi:10.1086/669413.

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., Wagener, T. & Hay, L. E. 2008 Framework for understanding structural errors (FUSE): a modular framework to diagnose differences between hydrological models. *Water Resources Research* **44** (12). doi:10.1029/2007WR006735.

Colby, B. R. 1956 Relationship of Sediment Discharge to Streamflow. USGS Open-File Report 56-27, U.S. Dept. of the Interior, Geological Survey, Water Resources Division.

Coron, L., Andréassian, V., Perrin, C., Lerat, J., Vaze, J., Bourqui, M. & Hendrickx, F. 2012 Crash testing hydrological models in contrasted climate conditions: an experiment on 216 Australian catchments. *Water Resources Research* **48**, W05552. doi:10.1029/2011WR011721.

Chopra, K., Leemans, R., Kumar, P. & Simons, H. (Eds.) 2005 *Ecosystems and Human Well-Being: Policy Responses*. Findings of the Condition and Trends Working Group of the Millennium Assessment. Island Press, p. 654.

Dakhlaoui, H., Ruelland, D., Tramblay, Y. & Bargaoui, Z. 2017 Evaluating the robustness of conceptual rainfall-runoff models under climate variability in northern Tunisia. *Journal of Hydrology* **550**, 201–217. doi:10.1016/j.jhydrol.2017.04.032.

Falcone, J. A. 2011 *GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow*. USGS Report. doi:10.3133/70046617.

Farmer, D., Sivapalan, M. & Jothityangkoon, C. 2009 Climate, soil, and vegetation controls upon the variability of water balance in temperate and semiarid landscapes: downward approach to water balance analysis. *Water Resources Research* **39** (2), 1035. doi:10.1029/2001WR000328.

Fenicia, F., Kavetski, D. & Savenije, H. H. G. 2011 Elements of a flexible approach for conceptual hydrological modeling: I. Motivation and theoretical development. *Water Resources Research* **47**, 11. doi:10.1029/2010WR010174.

Fowler, K., Peel, M., Western, A. & Zhang, L. 2018 Improved rainfall-runoff calibration for drying climate: choice of objective function. *Water Resources Research* **54** (5), 3392–3408. doi:10.1029/2017WR022466.

Gupta, A. & Govindaraju, R. S. 2019 Propagation of structural uncertainty in watershed hydrologic models. *Journal of Hydrology* **575**, 66–81. doi:10.1016/j.jhydrol.2019.05.026.

Hargraves, G. H. & Samani, Z. A. 1982 Estimating potential evapotranspiration. *Journal of the Irrigation and Drainage Division* **108** (3), 225–230.

Hernandez-Suarez, J. S., Nejadhashemi, A. P., Kropp, I. M., Abouali, M., Zhang, Z. & Deb, K. 2018 Evaluation of the impacts of hydrologic model calibration methods on predictability of ecologically-relevant hydrologic indices. *Journal of Hydrology* **564**, 758–772. doi:10.1016/j.jhydrol.2018.07.056.

Jothityangkoon, C., Sivapalan, M. & Farmer, D. L. 2001 Process controls on water balance variability in a large semi-arid catchment: downward approach to hydrologic model development. *Journal of Hydrology* **254** (1–4), 174–198. doi:10.1016/S0022-1694(01)00496-6.

Kiesel, J., Guse, B., Pfannenstiel, M., Kakouei, K., Jähnig, S. C. & Fohrer, N. 2017 Improving hydrological model optimization for riverine species. *Ecological Indicators* **80**, 376–385. doi:10.1016/j.ecolind.2017.04.032.

Klemelä, V. 1986 Operational testing of hydrological simulation models. *Hydrological Sciences Journal* **31** (1), 13–24. doi:10.1080/02626668609491024.

Knight, R. N., Gain, W. S. & Wolfe, W. J. 2012 Modelling ecological flow regime: an example from the Tennessee and Cumberland River basins. *Ecological Flow 5*, 613–627. doi:10.1002/eco.246.

Lane, R. A., Coxon, G., Freer, J. E., Wagener, T., Johnes, P. J., Bloomfield, J. P., Greene, S., Maceod, C. J. A. & Reaney, S. M. 2019 Benchmarking the predictive capability of hydrological models for river flow and flood peak predictions across over 1000 catchments in Great Britain. *Hydrology and Earth System Sciences* **23**, 4011–4032. doi:10.5194/hess-23-4011-2019.

Massmann, C. 2019 Modelling snowmelt in ungauged catchments. *Water* **11** (2), 301. doi:10.5390/w11020301.

Massmann, C. 2020 Identification of factors influencing hydrologic model performance using a top-down approach in a large number of U.S. catchments. *Hydrological Processes* **34**, 4–20. doi:10.1002/hyp.15366.

Murphy, J. C., Knight, R. R., Wolfe, W. J. & Gain, W. S. 2013 Predicting ecological flow regime at ungauged sites: a
comparison of methods. *River Research and Applications* **29** (5), 660–669. doi:10.1002/rra.2570.

Olsen, M., Troldborg, L., Henriksen, H. J., Conallin, J., Refsgaard, J. C. & Boegh, E. 2013 Evaluation of a typical hydrological model in relation to environmental flows. *Journal of Hydrology* **507**, 52–62. doi: 10.1016/j.jhydrol.2013.10.022.

Patil, S. D. & Stieglitz, M. 2015 Comparing spatial and temporal transferability of hydrological model parameters. *Journal of Hydrology* **525**, 409–417. doi:10.1016/j.jhydrol.2015.04.003.

Poff, N. L., Allan, J. D., Bain, M. B., Karr, J. R., Prestegaard, K. L., Richter, B. D., Sparks, R. E. & Stromberg, J. C. 1997 The natural flow regime. *BioScience* **47** (11), 769–784. doi:10.2307/1313099.

Pool, S., Vis, M. J. P., Knight, R. R. & Seibert, J. 2017 Streamflow characteristics from modeled runoff time series – importance of calibration criteria selection. *Hydrology and Earth System Sciences* **21**, 5443–5457. doi:10.5194/hess-21-5443-2017.

Richter, B. D., Baumgartner, J. V., Powell, J. & Braun, D. P. 1996 A method for assessing hydrologic alteration within ecosystems. *Conservation Biology* **10**, 1163–1174. doi:10.1046/j.1523-1739.1996.10041163.x.

Sanborn, S. C. & Bledsoe, B. P. 2006 Predicting streamflow regime metrics for ungauged streams in Colorado, Washington, and Oregon. *Journal of Hydrology* **325** (1–4), 241–261. doi:10.1016/j.jhydrol.2005.10.018.

Seiller, G., Ancill, F. & Perrin, C. 2012 Multimodel evaluation of twenty lumped hydrological models under contrasting climate conditions. *Hydrology and Earth System Sciences* **16** (4), 1171–1189. doi:10.5194/hess-16-1171-2012.

Shrestha, R. R., Peters, D. L. & Schnorbus, M. A. 2014 Evaluating the ability of a hydrologic model to replicate hydro-ecologically relevant indicators. *Hydrological Processes* **28** (14), 4294–4310. doi:10.1002/hyp.9997.

Shrestha, R. R., Schnorbus, M. A. & Peters, D. L. 2016 Assessment of a hydrologic model’s reliability in simulating flood regime alterations in a changing climate. *Hydrological Processes* **30** (15), 2628–2643. doi:10.1002/hyp.10812.

Sinokrot, B. A. & Gulliver, J. S. 2000 In-stream flow impact on river water temperatures. *Journal of Hydraulic Research* **38** (5), 339–349. doi:10.1080/0022168009498315.

Thornton, P. E., Running, S. W. & White, M. A. 1997 Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology* **190**, 214–251. doi:10.1016/S0022-1694(96)03128-9.

Thornton, P. E., Thornton, M. M., Mayer, B. W., Wilhelmi, N., Wei, Y., Devarakonda, R. & Cook, R. B. 2014 Daymet: Daily Surface Weather Data on A 1-km Grid for North America, Version 2. ORNL DAAC, Oak Ridge, Tennessee, USA (accessed June 2015). doi:10.3334/ORNLDAAC/1219.

van Esse, W. R., Perrin, C., Booj, M. J., Augustijn, D. C. M., Fenicia, F., Kovetski, D. & Lobigegas, F. 2013 The influence of conceptual model structure on model performance: a comparative study for 237 study catchments. *Hydrology and Earth System Sciences* **17**, 4227–4239. doi:10.5194/hess-17-4227-2013.

Vaze, J., Post, D. A., Chiew, F. H. S., Perraud, J.-M., Viney, N. R. & Teng, J. 2010 Climate non-stationarity – validity of calibrated rainfall–runoff models for use in climate change studies. *Journal of Hydrology* **394** (3–4), 447–457. doi:10.1016/j.jhydrol.2010.09.018.

Vis, M., Knight, R., Pool, S., Wolfe, W. & Seibert, J. 2015 Model calibration criteria for estimating ecological flow characteristics. *Water* **7**, 2358–2381. doi:10.3390/w7052358.

Vormoor, K., Heistermann, M., Bronstert, A. & Lawrence, D. 2018 Hydrological model parameter (in)stability – ‘crash testing’ the HBV model under contrasting flood seasonality conditions. *Hydrological Sciences Journal* **63** (7), 991–1007. doi:10.1080/02626667.2018.1466056.

Walsh, R. P. D. & Lawler, D. M. 1981 Rainfall seasonality: description, spatial patterns and change through time. *Weather* **36**, 201–208. doi:10.1002/j.1477-8696.1981.tb05400.x.

Weingartner, R., Blöschl, G., Hannah, D. M., Marks, D. G., Parajka, J., Pearson, C. S., Rogger, M., Salinas, J. L., Sauquet, E., Srikantan, R., Thompson, S. E. & Viglione, A. 2013 Prediction of seasonal runoff in ungauged basins. In: *Runoff Prediction in Ungauged Basins. Synthesis Across Processes, Places and Scales* (G. Blöschl, M. Sivapalan, T. Wagener, A. Viglione & H. Savenije, eds). Cambridge University Press, pp. 102–134.

Wolock, D. M. 2003 *Base-Flow Index Grid for the Conterminous United States*. USGS Report 2003-263. Available from: http://pubs.er.usgs.gov/publication/ofr03263

First received 16 May 2019; accepted in revised form 26 January 2020. Available online 30 March 2020