Global patterns in base flow index and recession based on streamflow observations from 3394 catchments

Hylke E. Beck,[1,2] Albert I. J. M. van Dijk,[3,4] Diego G. Miralles,[5] Richard A. M. de Jeu,[1] L. A. (Sampurno) Bruijnzeel,[1] Tim R. McVicar,[4] and Jaap Schellekens[6]

Received 2 April 2013; revised 29 October 2013; accepted 30 October 2013; published 3 December 2013.

[1] Numerous previous studies have constructed models to estimate base flow characteristics from climatic and physiographic characteristics of catchments and applied these to ungauged regions. However, these studies generally used streamflow observations from a relatively small number of catchments (<200) located in small, homogeneous study areas, which may have led to less reliable models with limited applicability elsewhere. Here, we use streamflow observations from a highly heterogeneous set of 3394 catchments (<10,000 km^2) worldwide to construct reliable, widely applicable models based on 18 climatic and physiographic characteristics to estimate two important base flow characteristics: (1) the base flow index (BFI), defined as the ratio of long-term mean base flow to total streamflow; and (2) the base flow recession constant (k), defined as the rate of base flow decay. Regression analysis results revealed that BFI and k were related to several climatic and physiographic characteristics, notably mean annual potential evaporation, mean snow water equivalent depth, and abundance of surface water bodies. Ensembles of artificial neural networks (ANNs; obtained by subsampling the original set of catchments) were trained to estimate the base flow characteristics from climatic and physiographic data. The catchment-scale estimation of the base flow characteristics demonstrated encouraging performance with R^2 values of 0.82 for BFI and 0.72 for k. The connection weights of the trained ANNs indicated that climatic characteristics were more important for estimating k than BFI. Global maps of estimated BFI and k were obtained using global climatic and physiographic data as input to the derived models. The resulting global maps are available for free download at http://www.hydrology-amsterdam.nl.

Citation: Beck, H. E., A. I. J. M. van Dijk, D. G. Miralles, R. A. M. de Jeu, L. A. Bruijnzeel, T. R. McVicar, and J. Schellekens (2013), Global patterns in base flow index and recession based on streamflow observations from 3394 catchments, *Water Resour. Res.*, *49*, 7843–7863, doi:10.1002/2013WR013918.

1. Introduction

[2] Base flow is defined in this study as the slowly varying portion of streamflow (Q), originating from groundwater storage and/or other delayed sources such as channel bank storage, lakes, wetlands, and melting snow and ice [*Hall*, 1968; *Griffiths and Clausen*, 1997; *Smakhtin*, 2001]. Knowledge of the base flow regime is important for a number of purposes, such as water resources management, aquatic ecosystem preservation, hydropower generation, contaminant transport, and low-flow forecasting [e.g., *Campolo et al*., 1999; *Brauman et al*., 2007; *Cyr et al*., 2011, and references therein]. Such knowledge is not directly available for ungauged catchments and hence for most of the terrestrial land surface [*Fekete and Vörösmarty*, 2007]. Regionalization procedures are required to transfer model parameters or Q characteristics from gauged to ungauged catchments [e.g., *Parajka et al*., 2005; *Yadav et al*., 2007; *Oudin et al*., 2008; *Zhang et al*., 2008; *Blöschl et al*., 2013]. Two important base flow characteristics are: (1) the base flow index (BFI (–)), defined as the ratio of long-term mean base flow to total Q [*Smakhtin*, 2001]; and (2) the recession constant (k (d^{-1})), defined as the rate of base flow decay [*Vogel and Kroll*, 1996].

[3] Several regression-based regionalization studies have established models to estimate BFI or k from climatic and physiographic characteristics of catchments [e.g., *Mazvimavi et al*., 2005; *Brandes et al*., 2005; *Longobardi and Villani*, 2008; *Van Dijk*, 2010; *Peña-Arancibia et al*., 2010; *Krakauer and Temimi*, 2011; *Ahiablame et al*., 2013]. Geology and soils were generally among the most important catchment characteristics identified in these

[1]Faculty of Earth and Life Sciences, VU University Amsterdam, Amsterdam, Netherlands.
[2]Now at Water Resources Unit, Joint Research Centre, Ispra (VA), Italy.
[3]Fenner School of Environment and Society, The Australian National University, Canberra, ACT, Australia.
[4]CSIRO Land and Water, Canberra, ACT, Australia.
[5]School of Geographical Sciences, University of Bristol, Bristol, UK.
[6]Inland Water Systems Unit, Deltares, Delft, Netherlands.

Corresponding author: H. E. Beck, Water Resources Unit, Joint Research Centre, Via Enrico Fermi 2749, 21027 Ispra (VA), Italy. (hylke.beck@jrc.ec.europa.eu)

0043-1397/13/10.1002/2013WR013918

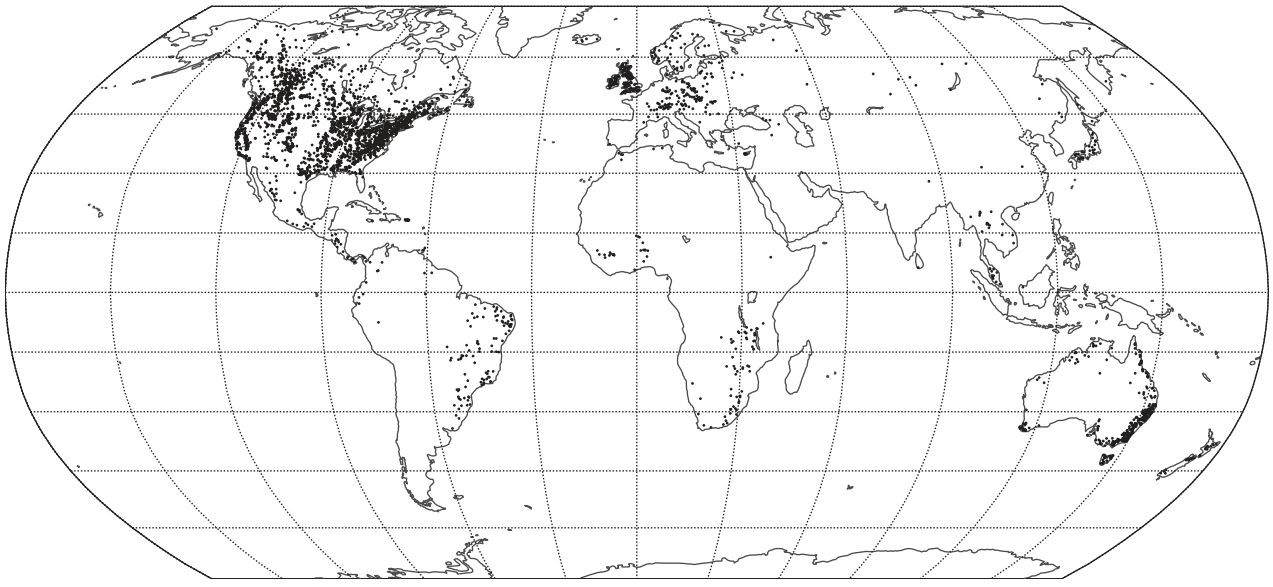

Figure 1. Locations of the catchments that passed the three quality control criteria and were used in the analysis here. Each data point represents a catchment centroid ($n = 3394$). The global maps in this paper are presented in the Robinson projection (80°S–80°N and 180°W–180°E) with grid lines at every 15° latitude and 30° longitude.

studies, although topography, climate, and/or land cover have also proven useful in some cases. Although previously developed models may, in theory, be used to estimate base flow characteristics for ungauged catchments, they suffer from one or more of the following shortcomings that limit their use in macro and global-scale applications. First, they were mainly based on Q observations from a relatively small number of catchments (<200), which can lead to less reliable and/or overfitted models. Second, they focused on a regional scale and used regional data sets to characterize geology or soils, thereby potentially restricting their larger-scale applicability. Third, most studies did not evaluate the generalization ability of the model using an independent set of catchments, and thus it is difficult to judge the true performance of the models. Finally, they reached conflicting conclusions regarding the importance of certain physiographic characteristics, notably mean surface slope and the fraction of forest cover.

[4] Groundwater, surface water, snow, and ice-related parameters of hydrological models generally control the generation of base flow. Current macro and global-scale baseflow-related parameterizations, such as those used in land surface schemes (a component of global circulation models) and global hydrological models, vary widely [cf. *Duan et al*., 2001]. Some use globally fixed parameters, such as the Community Land Model (CLM) [*Oleson et al*., 2010], Noah-MP [*Niu et al*., 2011], and Mac-PDM [*Gosling and Arnell*, 2011]. Others use nearest-neighbor interpolation of calibrated parameters, such as the Variable Infiltration Capacity (VIC) model [*Liang et al*., 1994; *Nijssen et al*., 2001] and WASMOD-M [*Widén-Nilsson et al*., 2007]. Yet others rely on expert opinion and hydrologic interpretation of global geological data sets, such as WaterGAP [*Döll and Fiedler*, 2008] and PCR-GLOBWB [*Van Beek et al*., 2011], or extrapolate a relationship between catchment climate and base-flow-related parameters found for Australia [*Peña-Arancibia et al*., 2010] to the global domain (W3RA) [*Van Dijk et al*., 2013]. Recently, the wider availability of Q, climatic, and physiographic data, thanks to (on-going) efforts by organizations such as the U.S. Geological Survey (USGS) and the Global Runoff Data Centre (GRDC; Koblenz, Germany; http://grdc.bafg.de) and advances in computing and remote-sensing technology, has created the possibility to explore the estimation of BFI and k at larger scales. Examples include *Schneider et al*. [2007] and *Santhi et al*. [2008], who investigated the regionalization of BFI for Europe and the conterminous USA, respectively, and *Peña-Arancibia et al*. [2010], who investigated the regionalization of k across the tropics.

[5] The current study uses a large set of 3394 catchments (<10,000 km^2) that covers all continents and a wide range of hydrologic, climatic, and physiographic conditions. This allows the construction of reliable, globally applicable models to estimate BFI and k. To our knowledge, this study is the first attempt to estimate these base flow characteristics from climatic and physiographic data globally using such a large Q data set. Specific objectives are to: (1) analyze the relationships between catchment characteristics and the selected base flow characteristics; (2) train empirical (as opposed to conceptual) models to estimate the base flow characteristics from climatic and physiographic characteristics of catchments and assess to what extent these models can be generalized; and (3) investigate the feasibility and accuracy of global-scale estimation of the base flow characteristics.

2. Data

2.1. Observed Streamflow (Q)

[6] Daily observed Q data were derived from three sources. First, Q data for 1862 USA stations that were part of the Model Parameter Estimation Experiment (MOPEX) [*Duan et al*., 2006] were downloaded from the USGS

National Water Information System (http://waterdata.usgs.gov) and the associated catchment boundaries from the MOPEX webserver (ftp://hydrology.nws.noaa.gov/pub/gcip/mopex/). Second, Q data for 4047 stations from the GRDC were considered. Corresponding catchment boundary data were provided by the GRDC. Third and finally, Q data and associated catchment boundaries for 321 Australian stations part of a database compiled by *Peel et al*. [2000] were used. Together, this resulted in an initial data set comprising 6230 Q gauging stations.

[7] For catchments to be included here, three requirements needed to be satisfied. First, due to the importance of channel routing in large catchments [e.g., *McGlynn et al*., 2004] the catchment area was limited to $<10{,}000$ km^2 [cf. *Peña-Arancibia et al*., 2010; *Van Dijk et al*., 2013]. Second, to reduce anthropogenic influences, catchments were only included if they had $<2\%$ urban [using the "artificial areas" class of the GlobCover v2 map; *Bontemps et al*., 2011] and/or subject to irrigation (using the Global Irrigated Area Map; http://www.iwmigiam.org). Third and finally, to ensure reliable estimates of the base flow characteristics the Q record length had to be >10 yr (but not necessarily consecutive). This resulted in a set of 3394 Q gauging stations, the locations of which are shown in Figure 1. Figure 2 shows frequency histograms of Q record length, area, mean humidity index (HI (–)), and major Köppen-Geiger climate type [*Peel et al*., 2007] for the catchments. HI was calculated as $\mathrm{HI} = P/\mathrm{PET}$, where P (mm yr^{-1}) is the mean annual precipitation derived from WorldClim [*Hijmans et al*., 2005] and PET (mm yr^{-1}) the mean annual potential evaporation derived from CGIAR-CSI [*Trabucco et al*., 2008]. Figure 3 shows the distribution of the major Köppen-Geiger climate types over the globe. All Q data were converted to mm d^{-1} using the catchment areas.

2.2. Climatic and Physiographic Characteristics

[8] Table 1 lists the climatic and physiographic characteristics used as predictors to model BFI and k. Predictor selection was guided by previous regionalization studies, expert knowledge, and data availability. Among the selected catchment characteristics, eight were related to climate, two to topography, three to land cover, one to geology, and four to soils—bringing the total number of predictors to 18. For the catchment-scale estimation of the base flow characteristics the full-resolution data were used. However, for the computation of global maps the data were resampled to 0.25° using averaging.

[9] A number of other predictors were considered, but not included in the analysis. The topographic wetness index (TWI) [*Beven and Kirkby*, 1979] was not used because a global high-resolution TWI data set is not (yet) available. Drainage density (total length of streams per unit catchment area) was not used either, due to the lack of globally consistent river-network data [cf. *Benstead and Leigh*, 2012]. Permafrost extent [*Brown et al*., 1997] was tested but not used as its inclusion did not result in better model performance and since its spatial patterns closely matched those of SNOW (cf. Table 1). Five global data sets providing information on soils or geology were considered [*Dürr et al*., 2005; *Batjes*, 2006; *Gleeson et al*., 2011; *Hartmann and Moosdorf*, 2012; *FAO/IIASA*, 2012]. For soils, we selected the Harmonized World Soil Database (HWSD)

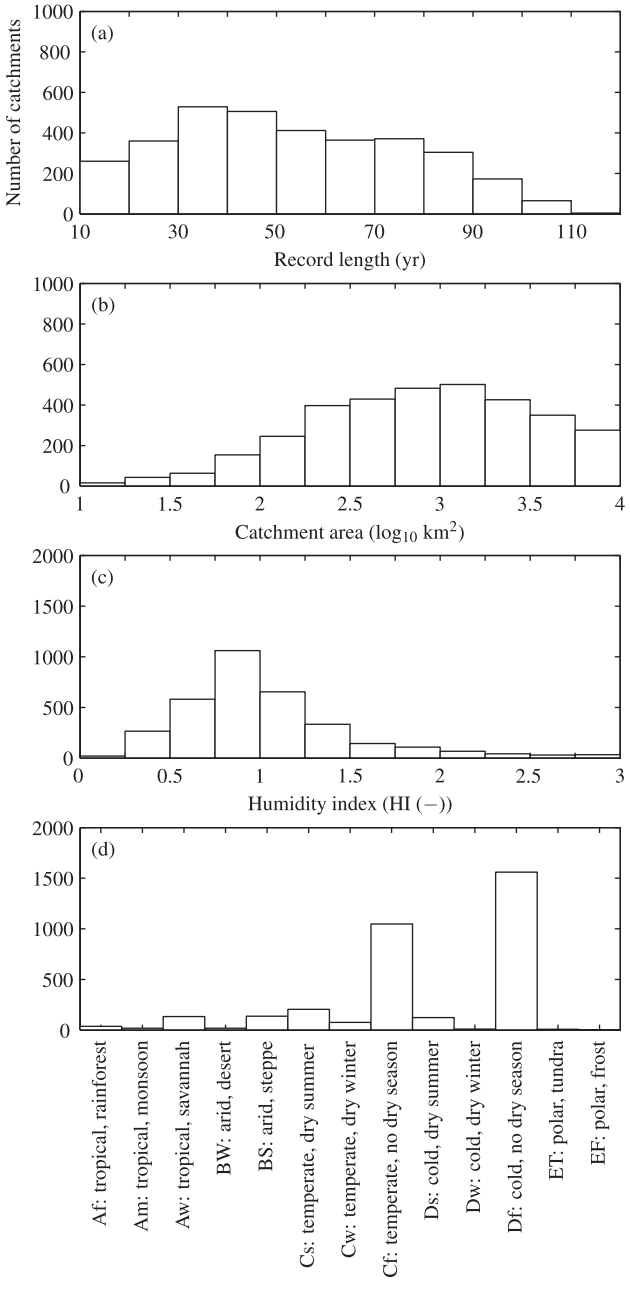

Figure 2. Frequency histograms of (a) Q record length, (b) catchment area, (c) mean catchment humidity index (HI), and (d) major Köppen-Geiger climate type for each catchment ($n = 3394$). To determine the major Köppen-Geiger climate type for each catchment the most dominant class was used. See Figure 3 for the Köppen-Geiger climate type map.

[*FAO/IIASA*, 2012] and for geology the global permeability map of *Gleeson et al*. [2011] as these data sets have a high resolution of 1 km and are based on comprehensive collections of soil and geologic data, respectively. In addition, BFI has been found to be strongly related to the sand content of the soil (SAND) [*Boorman et al*., 1995; *Santhi et al*., 2008] as well as indices related to the permeability of the underlying geology [*Longobardi and Villani*, 2008; *Bloomfield et al*., 2009]. To improve the HWSD data set, the SAND data for the USA were replaced with SAND

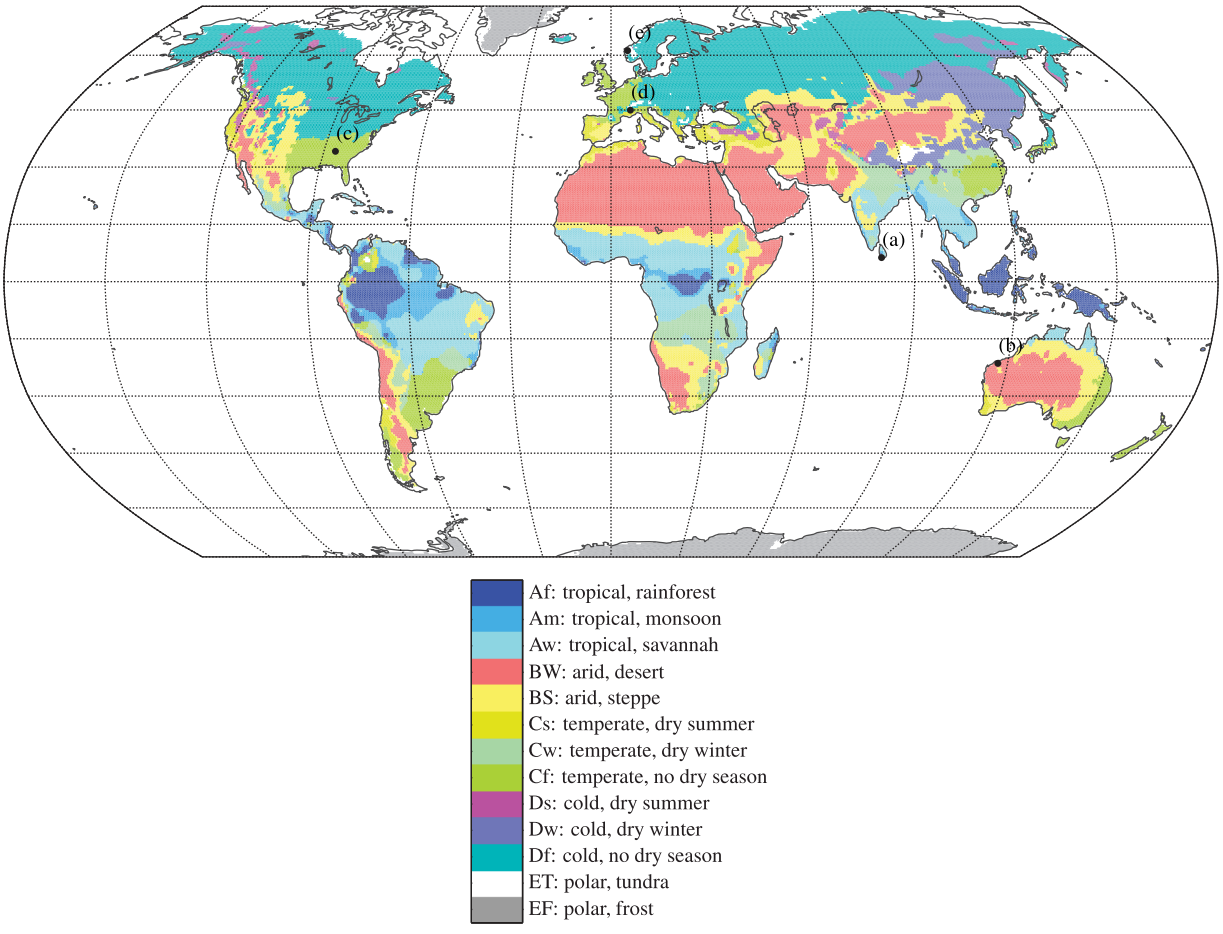

Figure 3. The Köppen-Geiger climate type map [*Peel et al.*, 2007]. The original classification consisted of 30 climate types, which were condensed here to 13 major climate types by omitting the third letter of the climate symbols. The five markers denote the locations of the Q gauging stations used to illustrate the base flow separation procedure (see Figure 4).

data as derived from the STATSGO data set [*Wolock et al.*, 2004] and the CLAY data for Australia with CLAY data as derived from the ASRIS data set [*Johnston et al.*, 2003], with the sand, silt, and clay contents changed proportionally to ensure that the sum equals 100% (i.e., if the STATSGO SAND value was 10% higher than the HWSD SAND value then the HWSD SILT and CLAY values would each be reduced by 5%).

3. Methodology

3.1. Computation of BFI and k

[10] Single values of BFI and k were computed from the Q record of each catchment following *Van Dijk* [2010]. A linear reservoir model was assumed as this is generally considered to be a good approximation [e.g., *Chapman*, 1999; *Fenicia et al.*, 2006; *Van Dijk*, 2010]. The linear reservoir model is given by:

$$Q(t)=kS(t), \tag{1}$$

where Q (mm d^{-1}) is the streamflow, k (d^{-1}) is the base flow recession constant, S (mm) is the reservoir storage, and t (d) is time. The continuity equation reads:

$$\frac{dS(t)}{dt}=-Q(t). \tag{2}$$

[11] After differentiation, equation (1) can be substituted in equation (2). By solving the resulting differential equation and introducing $t=1$, we can derive the following recurrence relation:

$$Q(t)=Q(t-1)\exp(-k). \tag{3}$$

[12] Equation (3) describes the falling limbs of hydrographs. The final k for a catchment was computed by reformulating equation (3) as follows:

$$k=-\ln\left(\frac{\mathrm{Q}}{\mathrm{Q}_*}\right), \tag{4}$$

where Q (mm d^{-1}) and Q_* (mm d^{-1}) are calculated from $Q(t)$ and $Q(t-1)$ data pairs of the Q record according to [*Van Dijk*, 2010]:

$$\mathrm{Q}=\exp\left(\overline{\ln(Q(t=1,2,3,\ldots))}\right) \text{ and} \tag{5}$$

$$\mathrm{Q}_*=\exp\left(\overline{\ln(Q(t=0,1,2,\ldots))}\right). \tag{6}$$

Table 1. The Climatic and Physiographic Characteristics Used as Predictor of the Base Flow Index (BFI) and Recession Constant (k)

Type	Predictor(s)	Description	Calculation and Data Source	Resolution
Climate	HI (–)	Humidity index	Calculated as: HI $=P/\mathrm{PET}$, where P is the mean annual precipitation derived from WorldClim [*Hijmans et al*., 2005] and PET is the mean annual potential evaporation derived from CGIAR-CSI [*Trabucco et al*., 2008]	~1 km
	P (mm yr^{-1})	Mean annual precipitation	WorldClim [*Hijmans et al*., 2005]	~1 km
	P_{si} (–)	Precipitation seasonality	Calculated following *Walsh and Lawler* [1981] as: $P_{si}=P_{yr}^{-1}\sum\lvert P_m-P_{yr}/12\rvert$, where P_{yr} and P_m are, respectively, the mean annual and monthly precipitation derived from WorldClim [*Hijmans et al*., 2005] and the summation is over all months	~1 km
	PET (mm yr^{-1})	Mean annual potential evaporation	CGIAR-CSI [*Trabucco et al*., 2008]	~1 km
	PET_{si} (–)	Potential evaporation seasonality	Calculated following *Walsh and Lawler* [1981] as: $\mathrm{PET}_{si}=\mathrm{PET}_{yr}^{-1}\sum\lvert\mathrm{PET}_m-\mathrm{PET}_{yr}/12\rvert$, where PET_{yr} and PET_m are, respectively, the mean annual and monthly potential evaporation derived from CGIAR-CSI [*Trabucco et al*., 2008] and the summation is over all months	~1 km
	CORR (–)	Seasonal correlation between water supply and demand	Correlation coefficient calculated between monthly climatic values of P and PET [*Petersen et al*., 2012] derived from WorldClim [*Hijmans et al*., 2005] and CGIAR-CSI [*Trabucco et al*., 2008], respectively	~1 km
	TA (K)	Mean annual air temperature	WorldClim [*Hijmans et al*., 2005]	~1 km
	SNOW (mm)	Mean snow water equivalent depth	GlobSnow L3A prototype with mountains included v1 [mean of 2008–2010; *Luojus et al*., 2010] for latitudes >35°N and AMSR-E/Aqua L3 v10 [mean of 2003–2011; *Chang* and *Rango*, 2000] for latitudes ≤35°N	~25 km
Topography	ELEV (m asl)	Mean elevation	CGIAR-CSI SRTM v4.1 data were used for the catchment-scale analysis, while IIASA-LUS data [*Fischer et al*., 2008] were used for the global-scale analysis	~90 m, 0.08°
	SLO (°)	Mean surface slope	Idem	~90 m, 0.08°
Land cover	fW (–)	Fraction of open water	GlobCover v2 [*Bontemps et al*., 2011]	~300 m
	fTC (–)	Fraction of forest	MODIS MOD44B collection 4 v3 [*Hansen et al*., 2003]	~250 m
	NDVI (–)	Mean Normalized Difference Vegetation Index (NDVI) [*Tucker*, 1979]	MODIS MOD13C2 collection 5 [*Huete et al*., 2002], mean of 2001–2012	0.05°
Geology	PERM ($\log_{10}$ m^2)	Mean permeability of consolidated and unconsolidated geologic units below the soil	Global permeability map [*Gleeson et al*., 2011]	~1 km
Soils	GRAV (%)	Mean gravel content	HWSD v1.21 [*FAO/IIASA*, 2012], mean of topsoil and subsoil values	~1 km
	SAND (%)	Mean sand content	HWSD v1.21 [*FAO/IIASA*, 2012], mean of topsoil and subsoil values, with the USA data replaced by STATSGO data [*Wolock et al*., 2004]	~1 km
	SILT (%)	Mean silt content	HWSD v1.21 [*FAO/IIASA*, 2012], mean of topsoil and subsoil values	~1 km
	CLAY (%)	Mean clay content	HWSD v1.21 [*FAO/IIASA*, 2012], mean of topsoil and subsoil values, with the Australian data replaced by ASRIS data [*Johnston et al*., 2003]	~1 km

[13] Data pairs with t during the first 5 days after a peak-flow were excluded in the calculation of Q and Q_* to reduce the influence of quick flows on the computed k [cf. *Peña-Arancibia et al*., 2010]. In addition, data pairs with zero flow were excluded. Typical values of k range from ~0.02 d^{-1} for slowly receding streams to ~0.3 d^{-1} for rapidly receding streams.

[14] The distribution of the derived k values for the different catchments had a strong positive skew that might confound the modeling exercise. Hence, to make the data

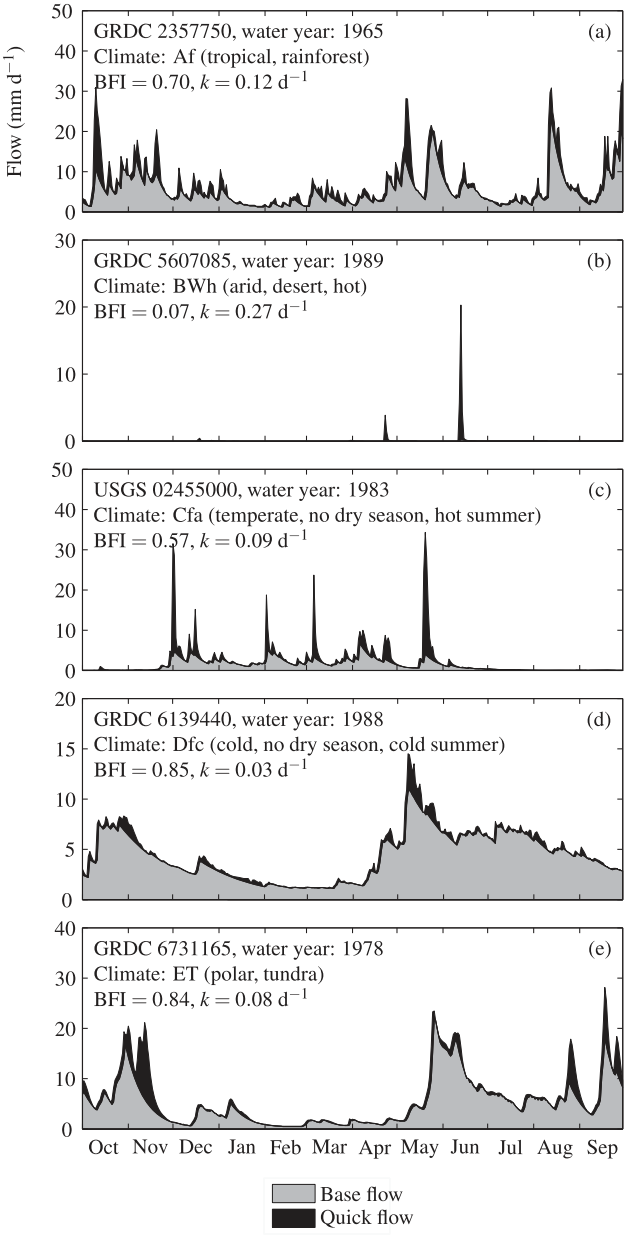

Figure 4. Examples of computed base flow and quick flow for the major Köppen-Geiger climate types. Figure 3 indicates the locations of the five selected Q gauging stations. The computed quick flow is the observed Q minus the computed base flow. The BFI and k values listed in each panel were based on the complete record of each station. A water year is defined from 1 October of one year to 30 September of the following year, and is designated by the year in which it ends.

better conform to a normal distribution the following transformation was applied:

$$k_{\text{trans}} = \ln(k), \tag{7}$$

where k_{trans} (–) is the transformed k. The inverse of this transformation is given by:

$$k = \exp(k_{\text{trans}}). \tag{8}$$

[15] The obtained k was subsequently used to separate the Q record into base flow and quick flow using a combination of forward- and backward-recursive digital filters following *Van Dijk* [2010]. Figure 4 gives examples of base flow computed in this way. The BFI was calculated as the ratio of long-term mean base flow to total Q, ranging from 0 to 1. Since the distribution of the derived BFI values for the different catchments showed a weak negative skew the following transformation was applied:

$$\text{BFI}_{\text{trans}} = \text{BFI}^2, \tag{9}$$

where $\text{BFI}_{\text{trans}}$ (–) is the transformed BFI. The inverse of this transformation is given by:

$$\text{BFI} = \sqrt{\text{BFI}_{\text{trans}}}. \tag{10}$$

[16] It is noted that the choice of techniques to compute BFI and k tends to affect the results, and since the "true" BFI and k are not known the choice of techniques remains somewhat subjective. Multicatchment studies comparing different techniques to compute BFI and k have generally concluded that there is high correlation, but sometimes relatively large systematic differences in BFI values among techniques [e.g., *Nathan and McMahon*, 1990; *Vogel and Kroll*, 1996; *Chapman*, 1999; *Sujono et al*., 2004; *Eckhardt*, 2008]. For instance, *Eckhardt* [2008] computed BFI values for 65 catchments using seven techniques and reported coefficient of determination (R^2) values ≥ 0.85 between BFI values computed using the different techniques (i.e., there was high correlation among techniques), while the mean BFI values computed using the different techniques ranged from 0.49 to 0.70 (i.e., there were relatively large systematic differences among techniques).

[17] The remaining three methodological subheaders reflect the original three objectives of this study, and are used to structure the subsequent sections 4 and 5.

3.2. Climatic and Physiographic Controls of $\text{BFI}_{\text{trans}}$ and k_{trans}

[18] Using regression analysis the strength and shape of the relationships between climatic and physiographic characteristics of the catchments and the transformed base flow characteristics ($\text{BFI}_{\text{trans}}$ and k_{trans}) were evaluated. Linear, exponential, logarithmic, and power functions were fitted by least squares and the function with the highest R^2 was reported. Significance levels (or p values) were not calculated as these may be misleading [*Nicholls*, 2001], particularly when using such a large number of catchments [*Royall*, 1986; *Johnson*, 1999].

3.3. Catchment-Scale Estimation of $\text{BFI}_{\text{trans}}$ and k_{trans}

[19] Initial attempts to estimate $\text{BFI}_{\text{trans}}$ and k_{trans} from climatic and physiographic characteristics of the catchments using multivariate linear models gave poor results due to the often nonlinear nature of the relationships [cf. *Mazvimavi et al*., 2005; *Van Dijk*, 2010; *Peña-Arancibia et al*., 2010]. We therefore decided to use artificial neural networks (ANNs), models composed of interconnected neurons that are able to model complex nonlinear relationships between inputs and outputs [*Bishop*, 1995]. ANNs have been used successfully in many fields of science, including hydrology [*ASCE*, 2000a, 2000b; *Govindaraju and Rao*, 2000; *Maier and Dandy*, 2000; *Mazvimavi et al*.,

2005]. Here, feed-forward ANNs based on the multilayer perceptron (MLP) [*Bishop*, 1995] with one hidden layer were used to estimate BFI_{trans} and k_{trans} from the climatic and physiographic data. The inputs (i.e., climatic and physiographic data) and outputs (i.e., BFI_{trans} and k_{trans}) were standardized using the means and standard deviations of the catchment values. The MLP models were trained using the Levenberg-Marquardt algorithm [*Levenberg*, 1944; *Marquardt*, 1963], which is considered to be one of the most efficient learning algorithms [*Hagan*, 1994], in combination with the mean-squared error performance function.

[20] The tenfold cross-validation procedure [*Shao*, 1993] was used to estimate the generalization ability of the trained MLP models. This procedure randomly partitions the original set of 3394 catchments into 10 subsamples, of which each comprises 10% of the catchments ($n = 339$). For 10 iterations, each time a different subsample of catchments was used to independently test the model's performance, the other 90% were randomly partitioned further into a training subset consisting of 75% of the catchments ($n = 2546$) and a validation subset consisting of 15% of the catchments ($n = 509$). For each iteration, the MLP model was trained on the training subset, while to prevent overfitting, the training process was stopped once the error for the validation subset started to increase [*Sarle*, 1995; *Bishop*, 1995]. The generalization ability of the MLP models was derived by averaging R^2 and root mean square error (RMSE) statistics computed for each cross-validation iteration from the testing subset of catchments. The optimal number of neurons in the hidden layer was determined by trial and error, based on R^2 and RMSE values obtained for the testing subsets. The number of neurons in the hidden layer was set at 30, as no further performance improvement was gained beyond 30 neurons. Using such a high number of neurons in the hidden layer avoids the convergence to local minima [*Tetko et al*., 1995]. As a last step, the output was destandardized using the mean and standard deviation of the catchment values.

[21] ANNs are often labeled as being "black boxes" as it is difficult to gain insight into the nature of the relationships between the input and the output [*Sjöberg et al*., 1995]. To overcome this perception, several methods have been developed to quantify the relative importance of the different inputs [*Gevrey et al*., 2003; *Olden et al*., 2004]. We employed the connection weight approach (CWA) [*Olden and Jackson*, 2002], which was shown to perform best in a comparison of nine different methods [*Olden et al*., 2004]. The CWA computes the relative importance of inputs according to [*Olden and Jackson*, 2002]:

$$S_i = \sum_{h=1}^{30} a_{i,h} b_h, \tag{11}$$

where S_i (–) is the relative importance, $a_{i,h}$ (–) are the input-hidden connection weights, b_h (–) are the hidden-output connection weights, $i = 1, 2, \ldots, 18$ denote the inputs, and $h = 1, 2, \ldots, 30$ denote the hidden neurons. Figure 5 illustrates the computation of S for a specific input. A high absolute S value indicates that the input is important. In this study, the results of all 10 trained MLP models (one for each cross-validation iteration) were taken into account by calculating the mean relative input importance ($\overline{S}$ (–)). The inputs are subsequently ranked based on absolute values of the mean relative input importance ($|\overline{S}|$).

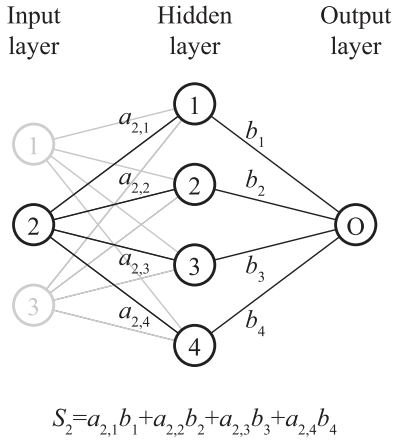

Figure 5. Computation of the relative importance of input 2 (S_2) illustrated using a feed-forward neural network consisting of three inputs and four hidden neurons.

3.4. Global Maps of BFI and k

[22] Global climatic and physiographic data were used as input to the trained MLP models, producing 10 maps of BFI_{trans} and k_{trans} (0.25° resolution). These maps were combined into single maps of BFI_{trans} and k_{trans} by calculating the per-pixel median of the 10 BFI_{trans} or k_{trans} values, respectively. The maps were subsequently transformed back to BFI and k using equations (8) and (10), respectively. The uncertainty of the estimates was quantified by calculating the per-pixel standard deviation of the 10 BFI_{trans} and k_{trans} values. The use of nontransformed values would artificially inflate the uncertainty in regions with low (high) BFI (k). Pixels attributed as ice (using the WWF terrestrial biomes map v2; *Olson et al*. [2001]) or open water (using GlobCover v2; *Bontemps et al*. [2011]) were excluded. To better understand the spatial patterns of the global estimates, for each 0.25° latitude band the 90th percentile, median, and 10th percentile of the global BFI and k estimates were computed. Additionally, the medians of the global BFI and k estimates were computed for the major Köppen-Geiger climate types (see Figure 3).

4. Results

4.1. Climatic and Physiographic Controls of BFI_{trans} and k_{trans}

[23] Figures 6 and 7 show scatterplots of catchment-mean values of the 18 selected climatic and physiographic characteristics versus observed values of the transformed base flow characteristics (BFI_{trans} and k_{trans}, respectively). The relationships were all rather weak ($R^2 \leq 0.22$) and often characterized by high degrees of nonlinearity and/or heteroscedasticity (i.e., uneven variability). Among the eight climate predictors, PET, PET_{si}, TA, SNOW, ELEV, and fW were moderately well related to BFI_{trans} (Figures 6d, 6e, 6g, 6h, 6i, and 6k, respectively), whereas HI, PET,

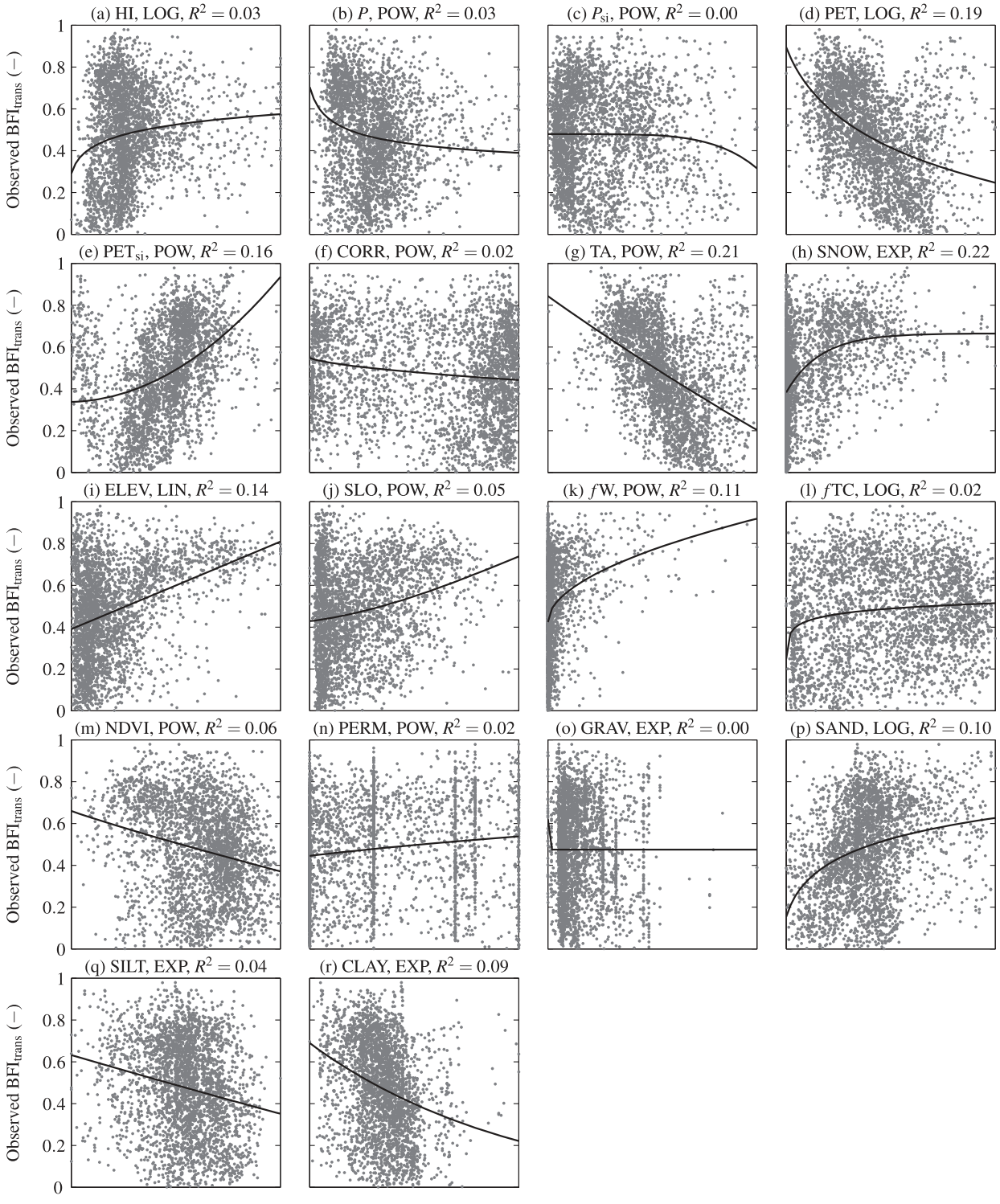

Figure 6. Scatterplots of catchment-mean values of the climatic and physiographic characteristics (along the x axis) versus the observed BFI_{trans} (along the y axis), including the best-fit regression line. Each data point represents a catchment ($n = 3394$). The x axis is linear and ranges from the minimum to maximum value of the data (not shown). Table 1 lists descriptions of the predictor variables. Abbreviations referring to the type of regression equation: EXP, exponential; LIN, linear; LOG, logarithmic; and POW, power.

and SNOW were moderately well related to k_{trans} (Figures 7a, 7d, and 7h, respectively). ELEV and SLO were positively related to BFI_{trans} and negatively related to k_{trans} (Figures 6i, 7i, 6j, and 7j, respectively). Among the three land-cover indices, fTC and NDVI were related to neither BFI_{trans} nor k_{trans} (Figures 6l, 7l, 6m, and 7m, respectively),

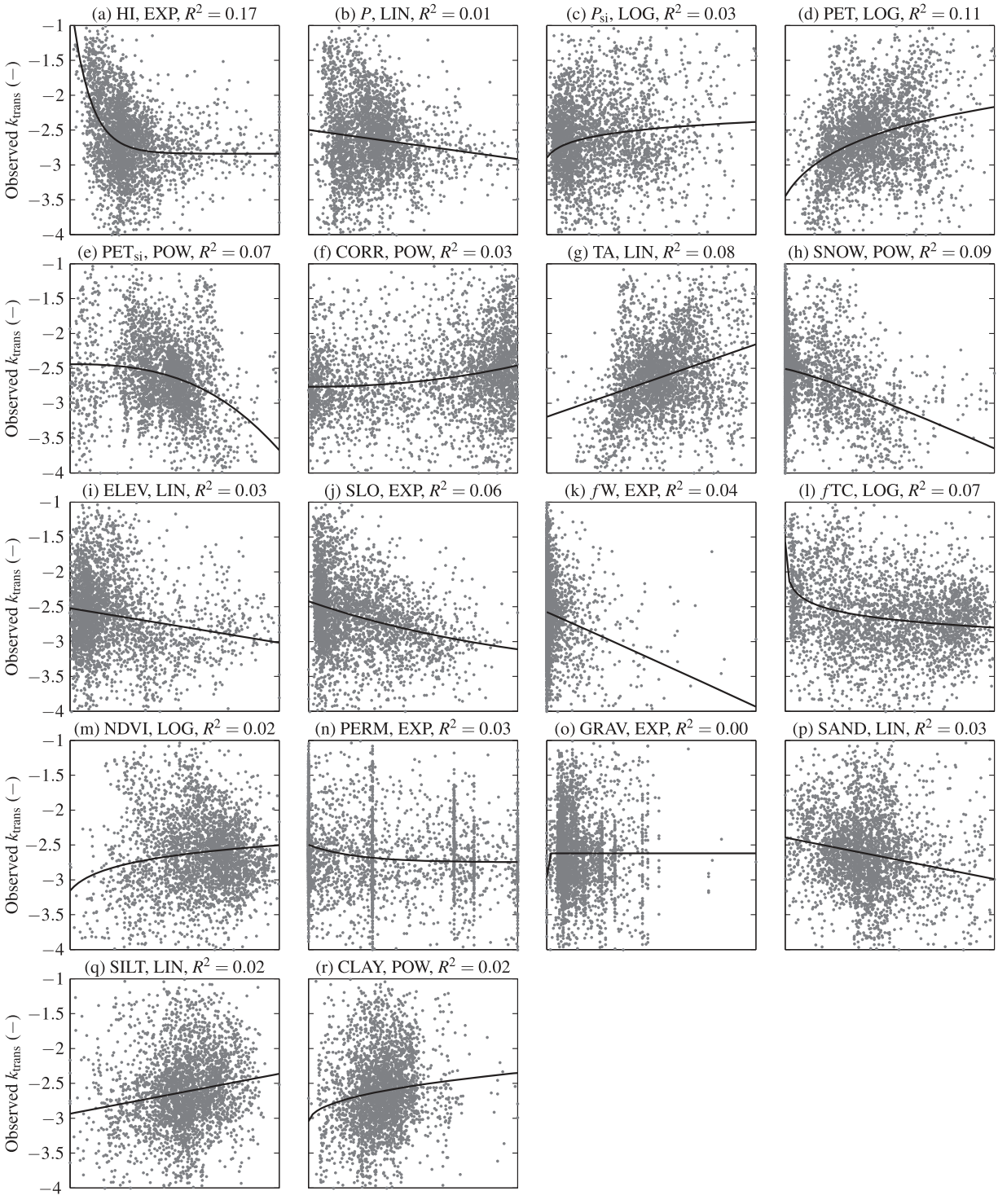

Figure 7. Scatterplots of catchment-mean values of the climatic and physiographic characteristics (along the x axis) versus the observed k_{trans} (along the y axis), including the best-fit regression line. Each data point represents a catchment ($n = 3394$). The x axis is linear and ranges from the minimum to maximum value of the data (not shown). Table 1 lists descriptions of the predictor variables. Abbreviations referring to the type of regression equation: EXP, exponential; LIN, linear; LOG, logarithmic; and POW, power.

whereas fW was positively related to BFI_{trans} (Figure 6k). The relationship between fW and BFI_{trans} was, however, highly heteroscedastic, demonstrating high variability at low fW and low variability at high fW (Figure 6k). The geology-related index PERM was related to neither BFI nor k (Figures 6n and 7n, respectively). Among the four soil indices, SAND was moderately well related to BFI_{trans} (Figure 6p).

Table 2. Mean R^2 and RMSE Statistics Obtained for the Training and Testing Subsets of Catchments

Base Flow Characteristic	Mean R^2		Mean RMSE (–)	
	Training	Testing	Training	Testing
BFI_{trans}	0.74	0.65	0.11	0.13
k_{trans}	0.65	0.53	0.36	0.41

4.2. Catchment-Scale Estimation of BFI_{trans} and k_{trans}

[24] Table 2 shows mean R^2 and RMSE values (mean of 10 cross-validation iterations) obtained by the MLP models for the training and testing subsets of catchments. The mean training and testing R^2 values for BFI_{trans} are 0.74 and 0.65, respectively, whereas the corresponding values for k_{trans} are somewhat poorer at 0.65 and 0.53 (Table 2). Values of the Nash-Sutcliffe efficiency [*Nash and Sutcliffe*, 1970] are not listed because they were nearly identical to the reported R^2 values, suggesting that there is no bias in the estimates. Figure 8 shows scatterplots of estimated versus observed values of BFI_{trans} and k_{trans}, including the linear regression line. The estimated values are the median estimates of the 10 MLP models (one for each cross-validation iteration). The associated R^2 values are 0.82 and 0.72 for the respective base flow characteristics (Figure 8), and thus are substantially higher than the mean training R^2 values (Table 2). The higher than unity slope of the regression equations (Figure 8) is due to the statistical phenomenon of regression toward the mean [*Galton*, 1886; *Bland and Altman*, 1994].

[25] Table 3 presents for both base flow characteristics the mean relative input importance ($\overline{S}$) calculated from the connection weights of the trained ANN models and the corresponding ranking based on absolute $\overline{S}$ values. For BFI_{trans} the three most important inputs are *f*W, SNOW, and ELEV, and for k_{trans} TA, P, and CLAY (Table 3). *f*W is important for BFI_{trans} (rank 1) but not k_{trans} (rank 9), whereas *P* is important for k_{trans} (rank 2) but not BFI_{trans} (rank 16; Table 3). The inputs related to geology and soils are relatively unimportant for both base flow characteristics, with the exception of CLAY for k_{trans} (rank 3; Table 3). Overall, the results indicate that climate-related inputs are more important for k_{trans} than for BFI_{trans} (Table 3).

4.3. Global Maps of BFI and k

[26] Figures 9a and 10a present global maps of BFI and *k*, respectively. These were produced in turn, using global climatic and physiographic data as input to the 10 trained MLP models (one for each cross-validation iteration), calculating the per-pixel median, and back-transforming the result. A large base flow fraction and slow recessions are found in the tropics, in the tundra-taiga zone, and for mountain ranges (notably the Chilean Andes, the Rocky Mts., and the Himalayas). To facilitate the comparison of spatial patterns, Figures 9b and 10b show the observed values. The BFI and *k* estimates appear to be relatively unbiased and the maps generally agree well with the observations (Figures 9 and 10).

[27] Figure 11 presents for each 0.25° latitude band the median and degree of dispersion of the global BFI and *k* estimates, illustrating how values vary according to latitude. High values of BFI are generally found between ~15°S and ~10°N and between ~45°N and ~65°N, whereas low *k* values are found between ~10°S and ~5°N and north of ~50°N (Figure 11). The spread in BFI values is relatively even across the whole latitudinal range, whereas there is a lower spread in *k* values around the equator and north of ~60°N (Figure 11). Table 4 lists for each base flow characteristic and major Köppen-Geiger climate type the median of the global maps. The variability in median BFI between climate types is relatively small, with values ranging from 0.49 for the "BW" climate type (arid, desert) to 0.77 for the "Ds" climate type (cold, dry summer). Values for *k* range from 0.04 for the "EF" climate type (polar, frost) to 0.27 for the "BW" climate type (arid, desert).

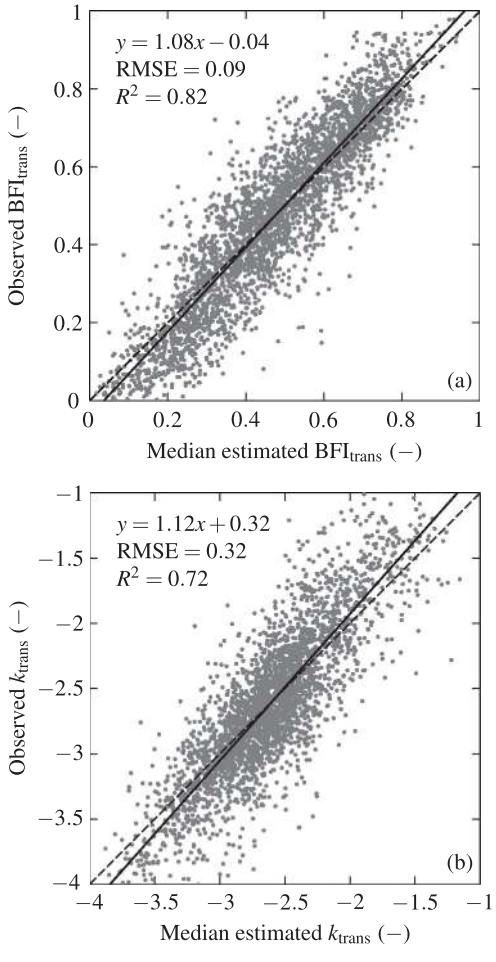

Figure 8. Scatterplots of estimated versus observed values of (a) BFI_{trans} and (b) k_{trans}. The linear regression is indicated by a solid line and the 1:1 relation by a dashed line. Each data point represents a catchment ($n = 3394$). The estimated values are the median estimates of the 10 artificial neural networks used (one for each cross-validation iteration). For each iteration, the training, validation, and testing subsets were included. Scatterplots using the nontransformed base flow characteristics (BFI and *k*) were not made as being based on non-normally distributed data these would result in nonrobust regressions.

Table 3. For BFI_{trans} and k_{trans} the Mean Relative Input Importance of the 18 Climatic and Physiographic Inputs as Expressed by $\overline{S}$ and the Corresponding Ranking Based on Absolute $\overline{S}$ Values

	BFI_{trans}		k_{trans}	
Input[a]	$\overline{S}$	Rank	$\overline{S}$	Rank
HI	−0.25	9	0.18	12
P	−0.07	16	−0.56	2
P_{si}	−0.10	15	−0.06	16
PET	0.01	18	0.27	6
PET_{si}	−0.34	7	0.29	5
CORR	−0.51	5	0.14	15
TA	0.20	13	1.41	1
SNOW	0.84	2	−0.25	7
ELEV	0.63	3	−0.15	13
SLO	0.24	10	−0.31	4
*f*W	1.02	1	0.21	9
*f*TC	−0.44	6	−0.05	17
NDVI	0.60	4	−0.14	14
PERM	−0.20	11	−0.22	8
GRAV	−0.14	14	−0.18	11
SAND	0.03	17	0.05	18
SILT	0.27	8	0.18	10
CLAY	−0.20	12	0.39	3

[a]Inputs are defined in Table 1.

[28] Figure 12 shows for the transformed base flow characteristics (BFI_{trans} and k_{trans}) the uncertainty of the global estimates, as computed from the estimation spread of the 10 trained MLP models. For both base flow characteristics, a lower uncertainty is found in North America, Europe, and southeastern Australia due to higher input data densities in these regions (Figure 1), whereas greater uncertainty is associated with (semi-)arid and (sub-)arctic regions (Figure 12).

5. Discussion

5.1. Climatic and Physiographic Controls of BFI_{trans} and k_{trans}

[29] Table 5 provides an overview of previous studies of (nontransformed) BFI or *k* regionalization. Although most of these studies were regional in nature, three had a continental scope [*Gustard and Irving*, 1994; *Schneider et al*., 2007; *Santhi et al*., 2008], and one covered the entire tropics [*Peña-Arancibia et al*., 2010]. The most commonly used climate-related indices were mean annual precipitation (*P*) [*Lacey and Grayson*, 1998; *Haberlandt et al*., 2001; *Mazvimavi et al*., 2005; *Longobardi and Villani*, 2008; *Peña-Arancibia et al*., 2010; *Krakauer and Temimi*, 2011], humidity index (HI) [*Lacey and Grayson*, 1998; *Mwakalila et al*., 2002; *Van Dijk*, 2010; *Peña-Arancibia et al*., 2010], and mean annual potential evaporation (PET) [*Lacey and Grayson*, 1998; *Van Dijk*, 2010]. The studies generally reported a positive (negative) relationship of BFI (*k*) with HI, an inconsistent relationship with *P*, and a negative (positive) relationship with PET. This is all in agreement with the present results (Figures 6a, 7a, 6b, 7b, 6d, and 7d, respectively). However, HI was only weakly related to BFI_{trans}, while the relationship between HI and k_{trans} was characterized by a high degree of heteroscedasticity (Figures 6a and 7a, respectively). Our results are consistent with the expectation that catchments with a high evaporative demand dry out faster after rainfall events, resulting in flows dominated by short-duration events (i.e., low BFI and high *k*). The positive relationship found between BFI_{trans} and mean snow water equivalent depth (SNOW) and the negative relationship found between k_{trans} and SNOW (Figures 6h and 7h, respectively) can be attributed to the gradual release of meltwater to *Q*.

[30] To date, no other studies have examined the relationship between BFI and precipitation seasonality (P_{si}). *Peña-Arancibia et al*. [2010] found no relationship between P_{si} and *k*, in agreement with this study (Figure 7c). The relationships obtained between potential evaporation seasonality (PET_{si}) and the base flow characteristics (Figures 6e and 7e) are probably not because PET_{si} is a controlling variable, but rather because of the strong similarity between PET and PET_{si} patterns. Several previous studies reported that the seasonal correlation between water supply and demand (CORR) is negatively related to the runoff coefficient (defined as the ratio of long-term *Q* to *P*) due to a strongly in-phase seasonality in supply and demand of water (i.e., high CORR) resulting in increased evaporation and hence drier soils [e.g., *Milly*, 1994; *Wolock and McCabe*, 1999; *Potter et al*., 2005]. These drier soils would favor lower BFI and higher *k* and we would therefore expect a negative (positive) relationship between CORR and BFI_{trans} (k_{trans}) to be present. The expected relationships are indeed present for both base flow characteristics, although the relationships are very weak (Figures 6f and 7f).

[31] The relationship obtained here between catchment-mean surface slope (SLO) and BFI_{trans} was positive and between SLO and k_{trans} negative (Figures 6j and 7j, respectively), which is in agreement with some other regionalization studies [*Lacey and Grayson*, 1998; *Mazvimavi et al*., 2005; *Longobardi and Villani*, 2008; *Peña-Arancibia et al*., 2010; *Van Dijk*, 2010; *Krakauer and Temimi*, 2011] and a sensitivity experiment using TOPMODEL [*Wolock et al*., 1989]. However, the relationships were rather weak and somewhat heteroscedastic (Figures 6j and 7j), indicating that topography is not a major control of base flow, particularly in catchments with gentle slopes. Consequently, the use of TOPMODEL-based runoff parameterizations in such catchments may be inappropriate [cf. *Beven*, 1997; *Li et al*., 2011]. Conversely, other regionalization studies found a negative relationship between SLO and BFI [*Haberlandt et al*., 2001] or a positive relationship between SLO and *k* [*Zecharias and Brutsaert*, 1988; *Post and Jakeman*, 1996; *Brandes et al*., 2005], although these studies used only a limited number of catchments (≤ 25). Additionally, the relationships found here (and by others) seems to contradict general drainage theory, which would predict a negative (positive) relationship between SLO and BFI (*k*) based on the premise that more steeply sloping aquifers drain faster [*Brutsaert and Nieber*, 1977; *Zecharias and Brutsaert*, 1988; *Vogel and Kroll*, 1992]. There are three potential explanations for this discrepancy. The first is that SLO may be a poor proxy for the aquifer flow gradient. However, this explanation fails to clarify why a positive (negative) relationship was obtained for BFI (*k*). The second is that SLO accounts for the effects of orography on climate, although the climate data sets used here (Table 1) did apply orographic corrections. The third is that SLO (which is based on relatively high-resolution SRTM data,

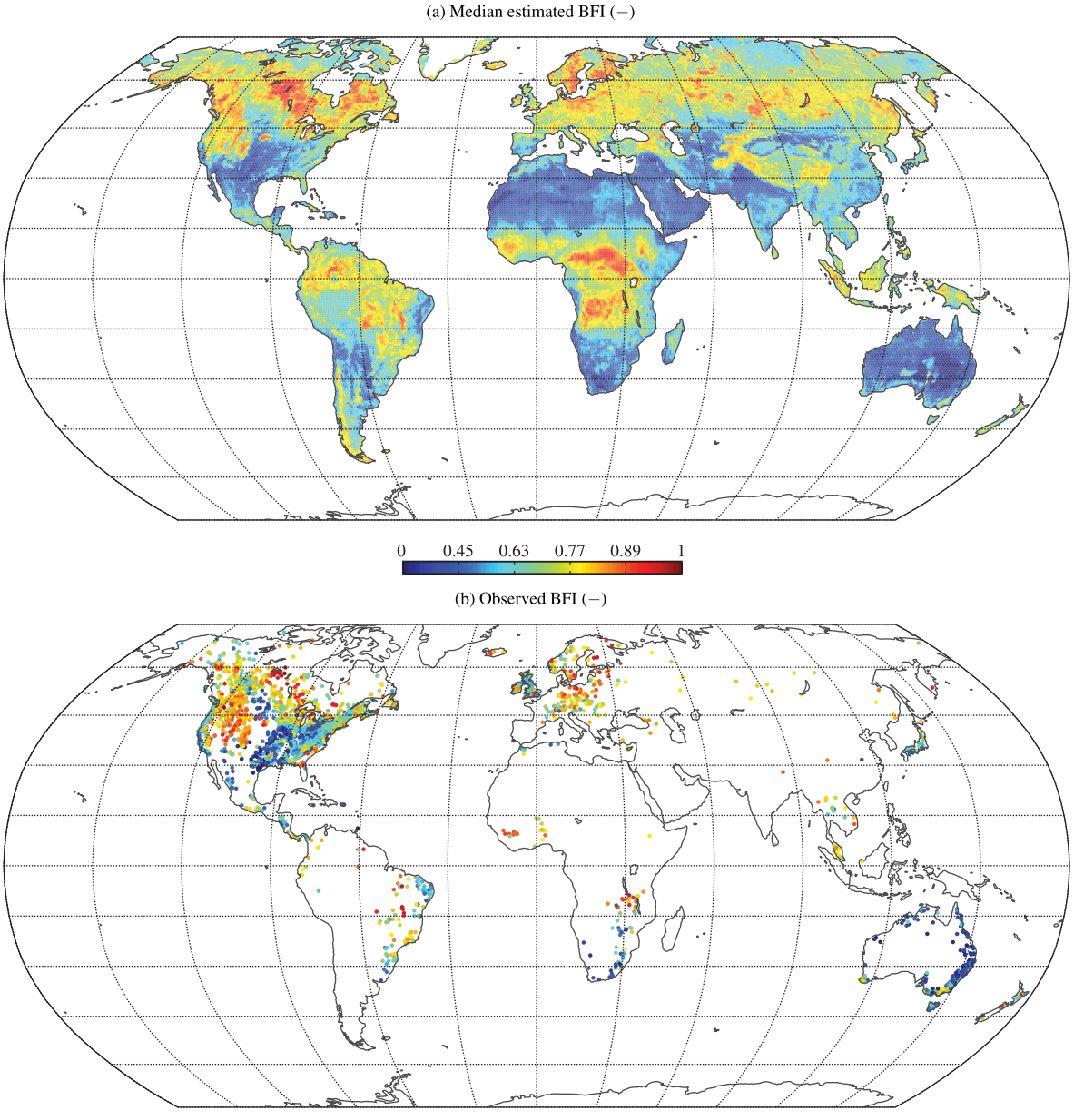

Figure 9. Global maps of (a) median estimated and (b) observed BFI. The estimated values in Figure 9a are the back-transformed medians of the 10 cross-validation iterations. Each data point in Figure 9b represents a catchment centroid ($n = 3394$).

cf. Table 1) acts as a surrogate for hydrologic characteristics of soils and geology in the absence of more detailed data on substrates. Topography is one of the primary influences on pedogenesis [*Price*, 2011], with soils forming on steep slopes often being more permeable than their counterparts on gentle slopes [e.g., *Ciolkosz et al*., 1989; *Janeau et al*., 2003; *Soulsby and Tetzlaff*, 2008]. Since permeable soils promote base flow contributions relative to less permeable soils [*Boorman et al*., 1995], this could explain the positive relationship obtained between SLO and BFI and the negative relationship obtained between SLO and *k*. A counter argument could be that bedrock usually occurs at more shallow depths on steep slopes than on gentler slopes, although bedrock is not necessarily impermeable due to fissures and fractures [*Davis*, 1969; *Tromp-van Meerveld et al*., 2007] and can contribute a considerable portion of total *Q* [*Uchida et al*., 2003].

[32] No clear evidence of relationships between *f*TC and the transformed base flow characteristics was obtained here (Figures 6l and 7l), nor in several other BFI regionalization studies [*Demuth and Hagemann*, 1994; *Mazvimavi et al*., 2005; *Longobardi and Villani*, 2008], suggesting that the presence of forests has little effect on the base flow fraction and recession rate. Although other regionalization studies did find relationships between *f*TC and BFI [*Lacey and Grayson*, 1998] or *k* [*Brandes et al*., 2005; *Peña-Arancibia*

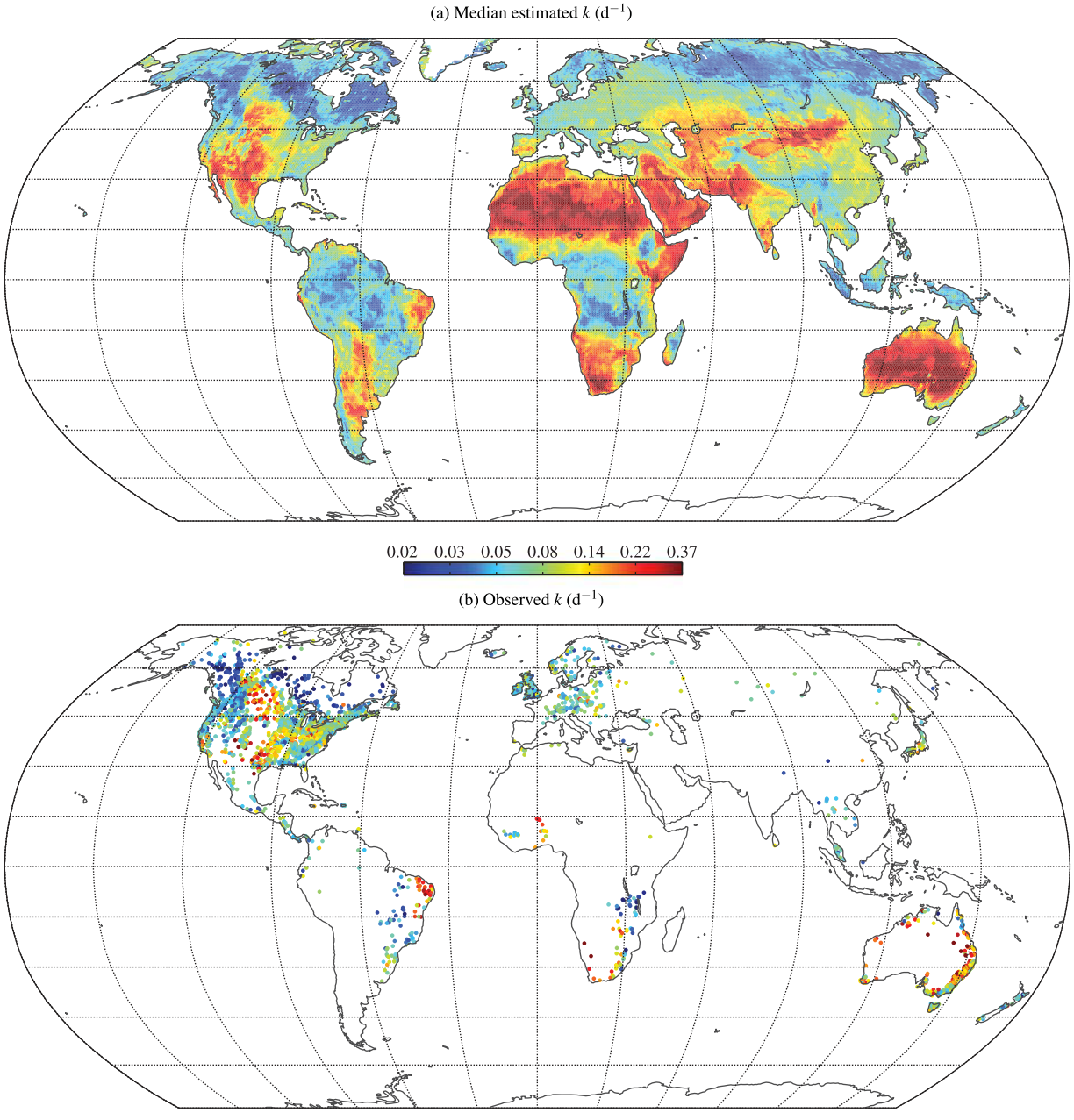

Figure 10. Global maps of (a) median estimated and (b) observed k. The estimated values in Figure 10a are the back-transformed medians of the 10 cross-validation iterations. Each data point in Figure 10b represents a catchment centroid ($n = 3394$).

et al., 2010; *Krakauer and Temimi*, 2011], the direction and strength of the relationships varied [cf. *Price*, 2011]. By contrast, the relationship between the fraction of open water (fW) and BFI_{trans} was positive (Figure 6k), suggesting that surface waters provide a steady supply of base flow. Similar relationships were obtained in BFI regionalization studies for the Great Lakes region of North America [*Neff et al*., 2005; *Ahiablame et al*., 2013].

[33] Soils and geology are undoubtedly two of the dominant controls of base flow [*Farvolden*, 1963; *Davis*, 1969; *Tague and Grant*, 2004] as confirmed by the fact that 16 of the 20 considered regionalization studies incorporated one, or more, indices related to soils or geology into their models (Table 5). Here, an index related to the permeability of the geology (PERM) and four indices related to soil texture (GRAV, SAND, SILT, and CLAY) were used. No relationships with PERM were found (Figures 6n and 7n), which does not mean that geology does not influence BFI or k, but rather that the data are not sufficiently accurate or that other hydrogeological properties are more important. A moderate relationship between the mean sand content of the soil (SAND) and BFI was found (Figure 6p), in agreement with BFI regionalization studies for the conterminous USA [*Santhi et al*., 2008] and the UK [*Boorman et al*., 1995]. The weaker relationship found here (Figure 6p) is most likely due to the quality of the global HWSD data set compared to the regional soil data sets used in the two previously mentioned studies.

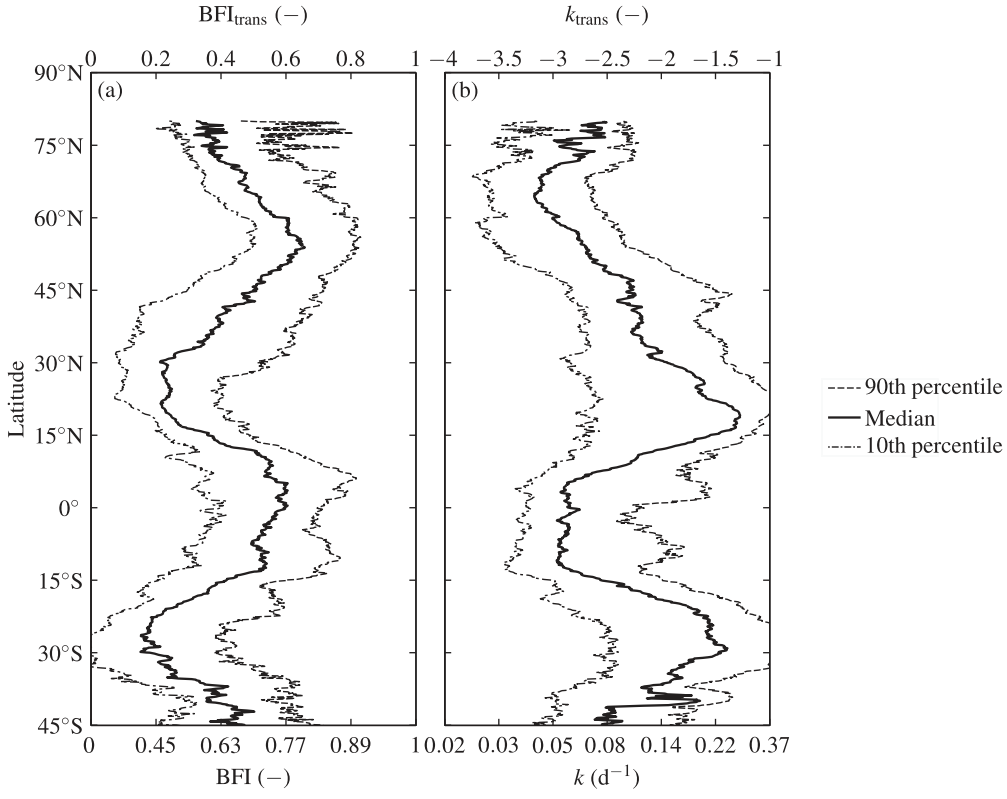

Figure 11. The 90th percentile, median, and 10th percentile values for (a) BFI and BFI_{trans} and (b) k and k_{trans} for each 0.25° latitude band derived from the global maps. Values are plotted only if there are >60 pixels of 0.25° resolution with a value within the latitudinal band.

5.2. Catchment-Scale Estimation of BFI_{trans} and k_{trans}

[34] Most regionalization studies used multivariate linear regression (Table 5) and nontransformed values of the base flow characteristics (BFI and k). However, the pronounced skewness of the BFI and k distributions found here suggests the need for transformation of the data to achieve greater normality. Additionally, the nonlinear relationships between the predictors and the (transformed) base flow characteristics obtained here [cf. Figures 6 and 7; *Mazvimavi et al*., 2005; *Van Dijk*, 2010; *Peña-Arancibia et al*., 2010] suggest the usefulness of ANNs. Many regionalization studies have used a relatively small number of catchments (<100; Table 5), which may have led to less reliable or overfitted models. In general, studies with <100 catchments obtained higher training R^2 values than studies based on a larger number of catchments (≥ 100; Table 5). For BFI_{trans}, the training R^2 value obtained here (0.74; Table 2) falls in the upper range of values reported for other studies with ≥ 100 catchments (0.34–0.79; Table 5). For k_{trans}, the training R^2 value obtained here (0.65; Table 2) is much higher than the corresponding values reported for other studies with ≥ 100 catchments (0.25–0.49; Table 5). Two of these studies [*Van Dijk*, 2010; *Peña-Arancibia et al*., 2010, Table 5] incorporated only one (climate related) predictor in their model, whereas better results might have been achieved using multiple predictors.

[35] Assessing a model's performance on an independent data set (i.e., to provide an estimate of generalization) is a crucial aspect of model development. Here, mean testing R^2 values of 0.65 and 0.53 were obtained for BFI_{trans} and k_{trans}, respectively (Table 2). We consider these results to be acceptable given the large geographic spread and the wide range of geology, soils, topography, climate, and land use covered by the catchments. Only three other (BFI) regionalization studies conducted a generalization assessment of their established models [*Haberlandt et al*., 2001; *Schneider et al*., 2007; *Bloomfield et al*., 2009]. Of these three studies, only one reported the associated statistical measures [*Haberlandt et al*., 2001]. We would recommend

Table 4. Median Estimated BFI and k for the Major Köppen-Geiger Climate Types

Climate Type[a]	BFI (–)	k (d^{-1})
Af: tropical, rainforest	0.77	0.05
Am: tropical, monsoon	0.74	0.06
Aw: tropical, savannah	0.71	0.08
BW: arid, desert	0.49	0.27
BS: arid, steppe	0.57	0.17
Cs: temperate, dry summer	0.65	0.09
Cw: temperate, dry winter	0.67	0.07
Cf: temperate, no dry season	0.63	0.08
Ds: cold, dry summer	0.77	0.05
Dw: cold, dry winter	0.74	0.07
Df: cold, no dry season	0.74	0.05
ET: polar, tundra	0.64	0.05
EF: polar, frost	0.70	0.04

[a]The locations of the major Köppen-Geiger climate types are shown in Figure 3.

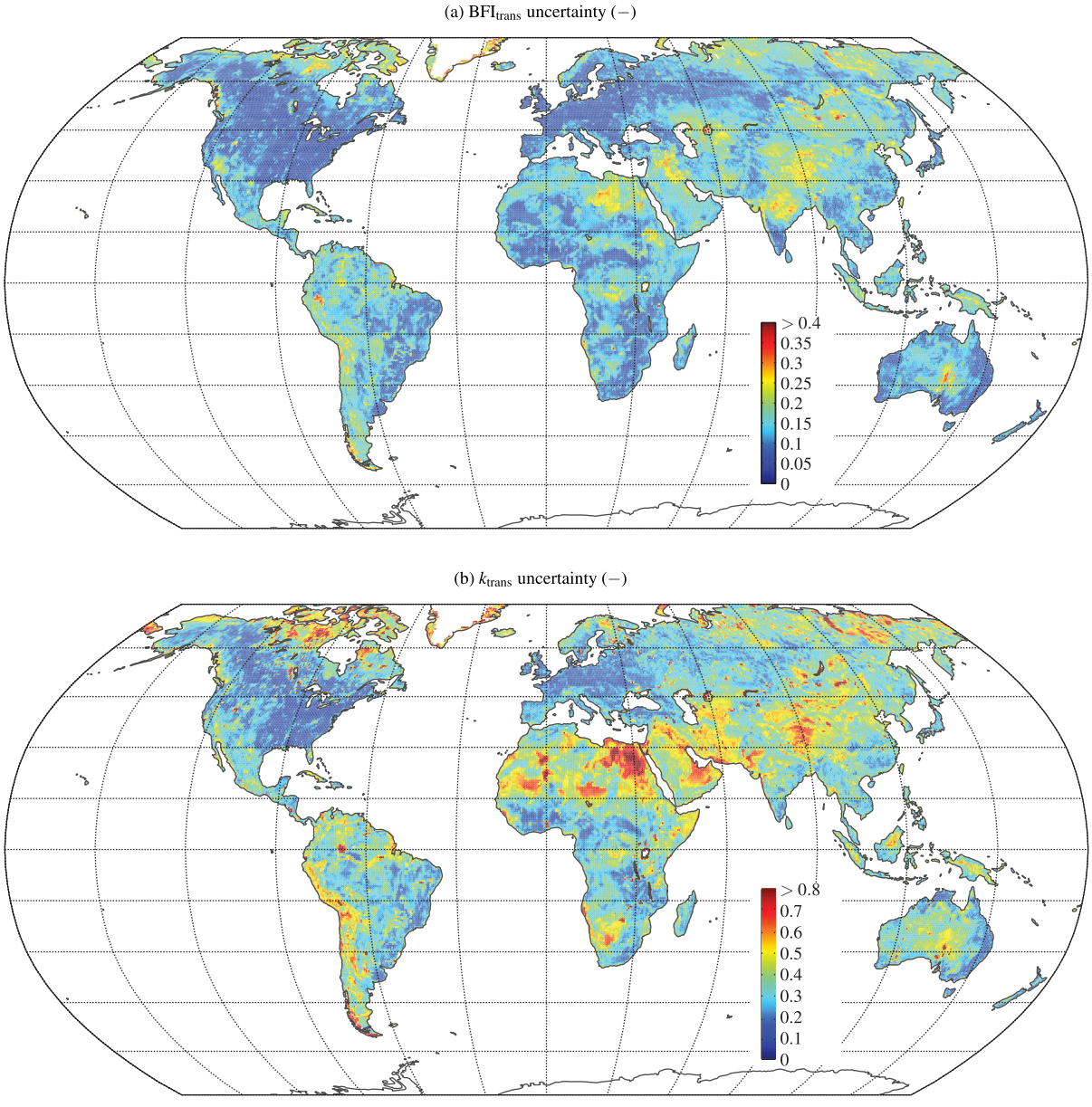

Figure 12. Global maps of the uncertainty of (a) BFI_{trans} and (b) k_{trans}. The data represent the per-pixel standard deviation of the 10 transformed estimates (one for each cross-validation iteration). Note that the use of untransformed values would artificially inflate the uncertainty in regions with low (high) BFI (k).

all future researchers to explicitly report generalization assessment statistics in their papers. *Haberlandt et al.* [2001] used multivariate linear regression, ordinary kriging, and external drift kriging to estimate BFI from 10 climatic and physiographic characteristics for 25 catchments located in the German part of the Elbe River Basin. Based on a leave-one-out cross-validation, they obtained RMSE values of 0.09, 0.11, and 0.08 for the respective approaches. Their RMSE values are comparable to the mean RMSE value of 0.11 computed here from back-transformed observed and estimated BFI for the testing subsets (noting that the value of 0.13 in Table 2 was derived from transformed BFI values).

[36] The performance statistics obtained for BFI_{trans} were better than those for k_{trans} (cf. Table 2 and Figure 8), in agreement with previous studies using nontransformed values (Table 5). The better performance of BFI may be attributable to several factors. First, k has been found to vary somewhat seasonally in response to changes in actual evaporation [*Czikowsky and Fitzjarrald*, 2004]. Moreover, if the base flow recession of a catchment is nonlinear, the assumption of a linear reservoir to derive k as used here will lead to different estimates of k depending on the flow rate [*Wittenberg*, 1999; *Krakauer and Temimi*, 2011]. Additionally, k is calculated from low-flow periods in the Q record that are subject to relatively lower instrument precision than intermediate flows [*Carter*, 1963] and are affected by rating-curve uncertainty [*Tomkins*, 2013]. Finally, k suffers from a greater sampling error than BFI as it is computed from only parts of the overall Q record.

Table 5. Overview of BFI or *k* Regionalization Studies, Listed in Order of the Number of *Q* Gauges Used, by BFI Then *k*

Base Flow Characteristic	Reference	Number of Gauges	Region	Climate	Model[a]	R^{2b}	RMSE[b]	Predictor(s)[c]
BFI	*Mwakalila et al*. [2002]	15	Tanzania	Semiarid	LIN	0.89	–	HI, catchment fraction underlain by permeable geology
	Ahiablame et al. [2013]	22	Indiana, USA	Temperate	LOG	0.91	–	*f*W, catchment fractions of two soil infiltration capacity classes
	Haberlandt et al. [2001]	25	NE Germany	Temperate	LIN	0.87	0.07	SLO, TWI, soil hydraulic conductivity, *P*
	Longobardi and Villani [2008]	28	Southern Italy	Mediterranean	LIN	0.68	0.17	Catchment fraction underlain by permeable geology
					LIN	0.80	–	Catchment fraction underlain by permeable geology, *P*, ELEV, SLO, *f*TC
	Bloomfield et al. [2009]	44	Thames Basin, UK	Temperate	LIN	0.89	0.09	Catchment fractions of six hydrogeological classes
	Mazvimavi et al. [2005]	52	Zimbabwe	Subtropical	LIN	0.75	0.09	*P*, drainage density, 75th percentile of slopes in the catchment
					ANN	0.77	0.02	*P*, drainage density, 10th percentile of slopes in the catchment
	Schneider et al. [2007]	103	Europe	Temperate	LIN	–	–	Catchment fractions of 29 soil classes
	Lacey and Grayson [1998]	114	VIC, Australia	Temperate	LIN	0.72	–	Catchment area, ELEV, PET, *f*TC, catchment fraction underlain by sedimentary rock, *P*, channel length
	Van Dijk [2010]	183	SE Australia	Temperate, subtropical	EXP	0.34	0.16	PET
	Boorman et al. [1995]	575	UK	Temperate	LIN	0.79	0.09	Catchment fractions of 29 soil classes
	Neff et al. [2005]	959	Great Lakes Basin, North America	Temperate	EXP	0.53	0.11	Surficial geology, *f*W
	Gustard and Irving [1994]	1530	Europe	Temperate	LIN	0.46	–	Catchment fractions of nine soil classes
	Current study[d]	3394	Global	Various	ANN	0.74	0.10	HI, P, P_{si}, PET, PET_{si}, CORR, TA, SNOW, ELEV, SLO, *f*W, *f*TC, NDVI, PERM, GRAV, SAND, SILT, CLAY
	Santhi et al. [2008]	~8600	Conterminous USA	(Semi)arid, temperate, subtropical	LIN	0.79	–	SAND, elevation range of catchment
k	*Post and Jakeman* [1996]	16	VIC, Australia	Temperate	LIN	0.53	–	SLO
					LIN	0.41	–	Catchment shape
	Brandes et al. [2005]	24	Pennsylvania, USA	Temperate	POW	0.80	0.02	Drainage density, soil hydraulic conductivity, SLO
	Demuth and Hagemann [1994]	57	SW Germany	Temperate	POW	0.70	–	Catchment fractions of 14 hydrogeological classes
	Krakauer and Temimi [2011]	61	USA	Temperate	LIN	0.3–0.5	–	Longitude, soil infiltration capacity, latitude, channel length, *f*TC, *P*
	Hughes [1997]	134	South Africa	Subtropical, Mediterranean	LIN	0.25	–	Rainfall seasonality classes, modeled estimates of recharge, geological index
	Peña-Arancibia et al. [2010]	167	Entire tropics	(Sub)tropical	LOG	0.49	–	HI
					EXP	0.49	–	*P*
	Van Dijk [2010]	183	SE Australia	Temperate, subtropical	POW	0.27	–	HI
	Current study[d]	3394	Global	Various	ANN	0.65	0.03	HI, P, P_{si}, PET, PET_{si}, CORR, TA, SNOW, ELEV, SLO, *f*W, *f*TC, NDVI, PERM, GRAV, SAND, SILT, CLAY

[a]Lists the type of (multivariate) regression equation used. Abbreviations used: LIN, linear; POW, power; LOG, logarithmic; EXP, exponential; ANN, artificial neural network.

[b]Dash indicates the information was not reported in the study or is not directly comparable to other studies.

[c]Lists the predictors included in the model only and not all predictors considered. See Table 1 for descriptions of the predictor variables.

[d]The current study is added for the sake of completeness. The R^2 of the current study represents the mean training R^2 (cf. Table 2), whereas the RMSE represents the mean RMSE computed from back-transformed observed and estimated values of the base flow characteristics for the training subsets. Note that the RMSE in Table 2 was computed from transformed observed and estimated values of the base flow characteristics.

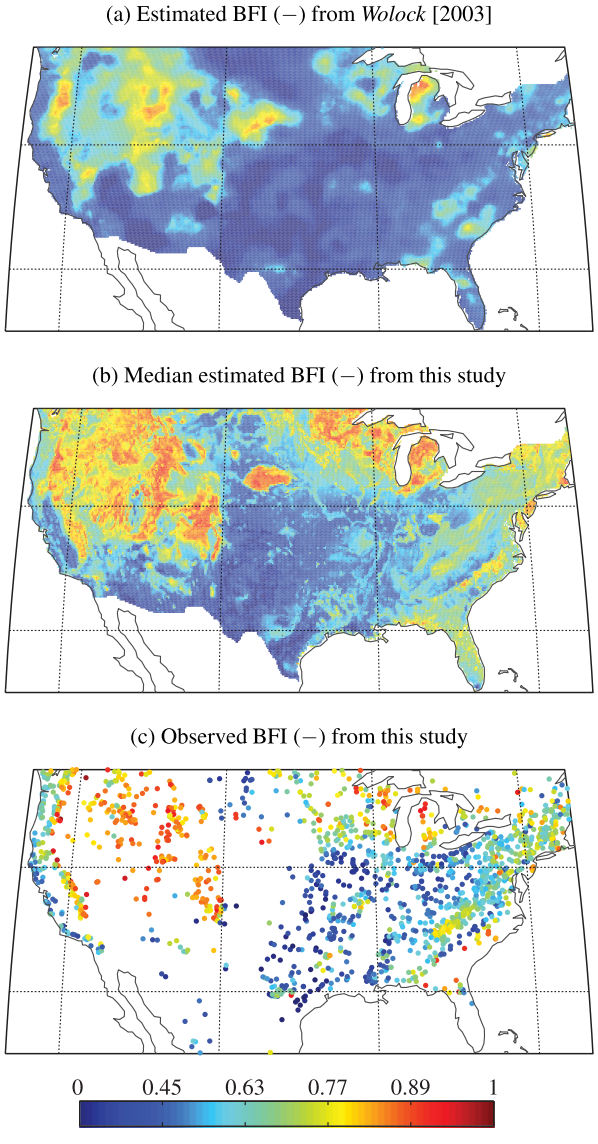

Figure 13. Map of the conterminous USA showing (a) estimated BFI from *Wolock* [2003], (b) median estimated BFI from this study, and (c) observed BFI. In (b) only values for the conterminous USA are shown. The BFI values in (a) were computed using a different method than the BFI values in (b) and (c). The maps have limits 25–48°N and 125–70°W and grid lines at every 10° latitude and 15° longitude.

[37] The R^2 values of the scatterplots exceed the mean training R^2 values (cf. Figure 8 and Table 2) because the estimated values in the scatterplots represent the mean of the 10 MLP models (one for each cross-validation iteration). Ensemble modeling (i.e., using the outputs from multiple models or from different realizations of the same model) is widely used in atmospheric, climate, and hydrologic sciences and is known to typically improve predictive accuracy [e.g., *Wandishin et al*., 2001; *Tebaldi and Knutti*, 2007; *Breuer et al*., 2009; *Viney et al*., 2009]. Additionally, several studies have reported that using ensembles of neural networks also improves the accuracy compared to using single neural networks [*Hansen and Salamon*, 1990; *Tetko et al*., 1995; *Shu and Burn*, 2004; *Ouarda and Shu*, 2009].

[38] The fraction of open water (*f*W) was identified as an important input for estimating BFI_{trans} but not k_{trans} (Table 3), probably because the BFI computation also considers peak flows which are delayed in catchments with high *f*W. Mean elevation (ELEV) was identified as important for the estimation of BFI_{trans} (Table 3), but since ELEV merely indicates the relative position it cannot physically influence the base flow characteristics. However, it is possible that ELEV, which was based on relatively high-resolution data (cf. Table 1), accounts for the effects of orography on climate and/or acts as a surrogate for soils and geology in the absence of more detailed data on substrates. The inputs related to geology and soils were mostly identified as unimportant (Table 3), probably due to the quality of the data. Note that the CWA results should be interpreted with caution as the approach does not account for [*Sarle*, 2000]: (1) the "squashing" effect of the hidden-layer activation function; (2) biases in the hidden and output layers; and (3) dependency among the input variables.

5.3. Global Maps of BFI and k

[39] To the best of our knowledge, our study is the first attempt to estimate BFI and *k* globally. However, several authors have produced regional BFI maps using various methods to compute and spatially estimate BFI [*Bullock et al*., 1997; *Haberlandt et al*., 2001; *Wolock*, 2003; *Neff et al*., 2005; *Lee et al*., 2006; *Santhi et al*., 2008]. *Wolock* [2003] and *Santhi et al*. [2008] produced BFI maps for the conterminous USA by interpolating the BFI values of 8249 and ~8600 catchments, respectively, using the inverse-distance weighting (IDW) technique. The BFI map of *Wolock* [2003] was available to us and is compared to the newly derived map in Figure 13. Our newly derived map and the map of *Wolock* [2003] exhibit similar spatial patterns, although the latter has a smoother appearance owing to the use of IDW interpolation (Figure 13). A per-pixel comparison of the maps was possible by resampling the map of *Wolock* [2003] using averaging to the resolution of the newly derived map (0.25°). The least squares linear regression equation associated with this comparison yields an R^2 of 0.50 and a RMSE of 0.14, indicating moderate agreement. For the newly derived map, the mean and standard deviation of BFI values for the conterminous USA are 0.61 and 0.18, respectively; for the map of *Wolock* [2003], the same figures are 0.43 and 0.19, respectively. The systematic difference in BFI values between the two maps is attributable to the use of different base flow separation procedures.

[40] *Neff et al*. [2005] presented a BFI map for the Great Lakes region of North America based on exponential relationships with the fraction of open water (*f*W) and a geological index. Comparing the BFI map of *Neff et al*. [2005] with our map (Figure 9a) reveals that the mean BFI for this region is similar at ~0.75. *Bullock et al*. [1997] produced a BFI map for Southern Africa by assigning the BFI values computed from the *Q* records of ~650 gauging stations to the associated catchments. Some of the catchments used by *Bullock et al*. [1997] were rather large (8% of the catchments were >10,000 km^2 and 2% >100,000 km^2), which

may have led to inflated BFI values due to channel routing effects. Nevertheless, the BFI map of *Bullock et al*. [1997] and the corresponding part of our map (Figure 9a) agree well in terms of spatial patterns—both maps showed markedly higher BFI values north of ~15°S. *Lee et al*. [2006] produced a BFI map for Taiwan by interpolation of BFI values computed from the Q records of 174 gauging stations. Again, our map and that of *Lee et al*. [2006] showed very similar spatial patterns. Similarly, the BFI maps of *Haberlandt et al*. [2001] for the German part of the Elbe River Basin and our map agreed well; both maps give a mean BFI of ~0.75 and place the highest BFI values in the central part of the Elbe.

[41] *Peña-Arancibia et al*. [2010] produced a k map for the entire tropics extending between 35°S and 30°N based on an exponential relationship with mean annual precipitation (P). They used the same method to compute k, the same data to derive P, and many catchments that are also used in this study. Consequently, the map of *Peña-Arancibia et al*. [2010] and our newly produced map (Figure 10a) show good consistency in terms of both spatial patterns and absolute magnitude.

[42] The global maps of the estimation uncertainty for BFI and k exhibit similar patterns (Figure 12). The generally greater uncertainty in (semi-)arid and (sub-)arctic regions (Figure 12) is due to application of the MLP models outside the climatic and physiographic domain of the catchment data. Additional efforts are recommended to validate and/or improve the present results for these regions. In light of this, the declining number of Q gauging stations in operation around the globe is cause for concern [*Stokstad*, 1999; *Shiklomanov et al*., 2002; *Fekete and Vörösmarty*, 2007].

[43] The global maps of BFI and k produced here may be useful for a range of large-scale hydrological applications, including the diagnosis and parameterization of land surface schemes and global hydrological models, water resource assessments, and catchment classification. However, some important characteristics and limitations should be noted. First, the maps reflect flows under natural, unregulated conditions. Second, in cold regions the main source of base flow is a combination of snow and ice melt. Third, the outflow from surface waters, such as lakes, wetlands, and reservoirs, can comprise a large part of the base flow. Fourth, the effects of river channel routing should be accounted for when using the maps for catchments $>10{,}000$ km^2. Finally, the BFI values in arid regions (notably the Sahara) are perhaps higher than expected, but this can be attributed to the approach used to compute BFI and k. The maps are available for free download from http://www.hydrology-amsterdam.nl.

6. Conclusion

[44] This study is the first attempt to estimate two important base flow characteristics (BFI and k) globally. A highly heterogeneous set of 3394 catchments was used to construct widely applicable models relating climatic and physiographic characteristics to BFI and k. The main findings are:

[45] 1. Since the BFI and k distributions showed negative and positive skewness, respectively, a data transformation was needed to better approximate a normal distribution, required to avoid bias in the estimates. The relationships between climatic and physiographic characteristics of the catchments and the transformed base flow characteristics (BFI_{trans} and k_{trans}) were often highly nonlinear and heteroscedastic. Among the catchment characteristics pertaining to climate, the mean annual potential evaporation (PET), PET seasonality (PET_{si}), mean annual air temperature (TA), and mean snow water equivalent depth (SNOW) were best related to BFI_{trans}, whereas the humidity index (HI), PET, and SNOW were best related to k_{trans}. The mean surface slope (SLO) was positively related to BFI_{trans} and negatively related to k_{trans}, which seems to contradict classical drainage theory and may represent a spurious relationship due to underlying patterns in soil hydrology and/or hydrogeology. Among the predictors pertaining to land cover, the fraction of forest (fTC) and the mean normalized difference vegetation index (NDVI) were related to neither BFI_{trans} nor k_{trans}, whereas the fraction of open water (fW) showed a moderate (positive) relationship with BFI_{trans}. A moderate (positive) relationship was also found between the mean sand content of the soil (SAND) and BFI_{trans}.

[46] 2. The nonlinear relationships obtained between climatic and physiographic characteristics of the catchments and BFI_{trans} or k_{trans} justified the use of artificial neural networks to estimate BFI_{trans} and k_{trans}. It proved possible to satisfactorily estimate BFI_{trans} and k_{trans} from climatic and physiographic data of catchments, yielding training R^2 values of 0.74 and 0.65, respectively. It was found that averaging the estimates of the ten artificial neural-network models (one for each cross-validation iteration) resulted in more accurate estimates. The results further show that artificial neural networks can be considered a viable and perhaps better alternative to the commonly used multivariate linear regression. The connection weights of the trained artificial neural networks indicated that climate-related inputs are more important for estimating k_{trans} than BFI_{trans}.

[47] 3. Global maps of BFI and k were obtained by using global climatic and physiographic data as input to the established models and back-transforming the result. The BFI and k values showed higher uncertainty in comparatively data-poor (semi-)arid and (sub-)arctic regions, and lower uncertainty in the more data-rich North America, Europe, and southeastern Australia. The global maps will prove useful for a variety of large-scale hydrological applications, although further validation of the maps is recommended, particularly in poorly gauged and ungauged regions.

[48] **Acknowledgments.** The first author was supported by Deltares (Delft, Netherlands). The Global Runoff Data Centre (GRDC; Koblenz, Germany) is thanked for providing most of the flow and catchment boundary data used here. We are also most grateful to Tom Gleeson (University of British Columbia, Vancouver, Canada) for providing the global permeability map and Jens Hartmann (Universität Hamburg, Hamburg, Germany) for providing the GLiM data set. Finally, sincere thanks are due to the Associate Editor, two anonymous reviewers, and Anne van Loon for their valuable comments that improved an earlier version of the manuscript.

References

Ahiablame, L., I. Chaubey, B. Engel, K. Cherkauer, and V. Merwade (2013), Estimation of annual baseflow at ungauged sites in Indiana USA, *J. Hydrol.*, *476*, 13–27.

ASCE (2000a), Artificial neural networks in hydrology. I: Preliminary concepts, *J. Hydrol. Eng.*, *5*, 115–123.

ASCE, T. (2000b), Artificial neural networks in hydrology. II: Hydrologic applications, *J. Hydrol. Eng.*, *5*, 124–137.

Batjes, N. H. (2006), *ISRIC-WISE Derived Soil Properties on a 5 by 5 Arc-Minutes Global Grid (Version 1.0)*, ISRIC—World Soil Inform., Wageningen, Neth.

Benstead, J. P., and D. S. Leigh (2012), An expanded role for river networks, *Nat. Geosci.*, *5*, 678–679.

Beven, K. J. (1997), TOPMODEL: A critique, *Hydrol. Processes*, *11*, 1069–1085.

Beven, K. J., and M. J. Kirkby (1979), A physically based, variable contributing area model of basin hydrology, *Hydrol. Sci. Bull.*, *24*, 43–69.

Bishop, C. M. (1995), *Neural Networks for Pattern Recognition*, Clarendon, Oxford, U. K.

Bland, J. M., and D. G. Altman (1994), Statistic notes: Regression towards the mean, *Br. Med. J.*, *308*, 1499.

Bloomfield, J. P., D. J. Allen, and K. J. Griffiths (2009), Examining geological controls on baseflow index (BFI) using regression analysis: An illustration from the Thames Basin, UK, *J. Hydrol.*, *373*, 164–176.

Blöschl, G., M. Sivapalan, T. Wagener, A. Viglione, and H. Savenije (Eds.) (2013), *Runoff Prediction in Ungauged Basins: Synthesis Across Processes, Places and Scales*, Cambridge Univ. Press, New York.

Bontemps, S., P. Defourny, and E. van Bogaert (2011), GlobCover 2009, products description and validation report, technical report, ESA GlobCover Project. [Available at http://ionia1.esrin.esa.int.]

Boorman, D. B., J. M. Hollist, and A. Lilly (1995), *Hydrology of soil types: A hydrologically based classification of the soils of the United Kingdom, Tech. Rep. 126*, Inst. of Hydrol., Wallingford, U. K.

Brandes, D., J. G. Hoffmann, and J. T. Mangarillo (2005), Base flow recession rates, low flows, and hydrologic features of small watersheds in Pennsylvania, USA, *J. Am. Water Resour. Assoc.*, *41*, 1177–1186.

Brauman, K. A., G. C. Daily, T. K. Duarte, and H. A. Mooney (2007), The nature and value of ecosystem services: An overview highlighting hydrologic services, *Ann. Rev. Environ. Resour.*, *32*, 67–98.

Breuer, L., et al. (2009), Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). I: Model intercomparison with current land use, *Adv. Water Resour.*, *32*, 129–146.

Brown, J., O. J. Ferrians, J. A. Heginbottom, and E. S. Melnikov (1997), *Circum-arctic map of permafrost and ground-ice conditions, Version 2, technical report*, Natl. Snow and Ice Data Cent., Boulder, Colo.

Brutsaert, W., and J. L. Nieber (1977), Regionalized drought flow hydrographs from a mature glaciated plateau, *Water Resour. Res.*, *13*, 637–643.

Bullock, A., A. J. Andrews, and R. Mngodo (1997), Regional surface water resources and drought assessment, in Southern African FRIEND: Flow Regimes from International Experimental and Network Data, pp. 40–93, UNESCO, Paris.

Campolo, M., A. Soldati, and P. Andreussi (1999), Forecasting river flow rate during low-flow periods using neural networks, *Water Resour. Res.*, *35*, 3547–3552.

Carter, R. W. (1963), Accuracy of current meter measurements, *J. Hydraul. Div. Am. Soc. Civ. Eng.*, *4*, 105–115.

Chang, A. T. C., and A. Rango (2000), Algorithm theoretical basis document for the AMSR-E snow water equivalent algorithm, technical report, NASA Goddard Space Flight Cent., Greenbelt, Md.

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

Ciolkosz, E. J., W. J. Waltman, T. W. Simpson, and R. R. Dobos (1989), Distribution and genesis of soils of the northeastern United States, *Geomorphology*, *2*, 285–302.

Cyr, J., M. Landry, and Y. Gagnon (2011), Methodology for the large-scale assessment of small hydroelectric potential: Application to the province of New Brunswick (Canada), *Renew. Energy*, *36*, 2940–2950.

Czikowsky, M. J., and D. R. Fitzjarrald (2004), Evidence of seasonal changes in evapotranspiration in eastern U.S. hydrological records, *J. Hydrometeorol.*, *5*, 974–988.

Davis, S. N. (1969), *Porosity and Permeability of Natural Materials*, pp. 53–89, Academic, New York.

Demuth, S., and I. Hagemann (1994), Estimation of flow parameters applying hydrogeological area information, in FRIEND: Flow Regimes from International Experimental and Network Data (Proceedings of the Braunschweig Conference, October 1993), IAHS Publ. 221, edited by P. Seuna et al., IAHS, Wallingford, UK, pp. 151–157.

Döll, P., and K. Fiedler (2008), Global-scale modeling of groundwater recharge, *Hydrol. Earth Syst. Sci.*, *12*, 863–885.

Duan, Q., J. Schaake, and V. Koren (2001), A Priori estimation of land surface model parameters, in *Land Surface Hydrology, Meteorology, and Climate: Observations and Modeling, vol. 3, Water Sci. and Appl.*, edited by V. Lakshmi, J. Albertson, and J. Schaake, pp. 77–94, AGU, Washington, D. C.

Duan, Q., et al. (2006), Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops, *J. Hydrol.*, *320*, 3–17.

Dürr, H. H., M. Meybeck, and S. H. Dürr (2005), Lithologic composition of the Earth's continental surfaces derived from a new digital map emphasizing riverine material transfer, *Global Biogeochem. Cycles*, *19*, GB4S10, doi:10.1029/2005GB002515.

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

FAO/IIASA, F. (2012), *Harmonized world soil database (version 1.2), technical report*, FAO, Rome and IIASA, Laxenburg, Austria.

Farvolden, R. N. (1963), Geologic controls on ground-water storage and base flow, *J. Hydrol.*, *1*, 219–249.

Fekete, B. M., and C. J. Vörösmarty (2007), The current status of global river discharge monitoring and potential new technologies complementing traditional discharge measurements, in Predictions in Ungauged Basins: PUB Kick-off (Proceedings of the PUB Kick-off Meeting Held in Brasilia, 20–22 November 2002), IAHS Publ. 309, pp. 129–136. IAHS, Wallingford, UK.

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

Fischer, G., F. Nachtergaele, S. Prieler, H. T. van Velthuizen, L. Verelst, and D. Wiberg (2008), *Global agro-ecological zones assessment for agriculture (GAEZ 2008), technical report*, IIASA, Laxenburg, Austria and FAO, Rome.

Galton, F. (1886), Regression towards mediocrity in hereditary stature, *J. Anthropol. Inst. Great Britain Ireland*, *15*, 246–263.

Gevrey, M., I. Dimopoulos, and S. Leka (2003), Review and comparison of methods to study the contribution of variables in artificial neural network models, *Ecol. Modell.*, *160*, 249–264.

Gleeson, T., L. Smith, N. Moosdorf, J. Hartmann, H. H. Dürr, A. H. Manning, L. P. H. van Beek, and A. M. Jellinek (2011), Mapping permeability over the surface of the earth, *Geophys. Res. Lett.*, *38*, L02401, doi:10.1029/2010GL045565.

Gosling, S. N., and N. W. Arnell (2011), Simulating current global river runoff with a global hydrological model: Model revisions, validation, and sensitivity analysis, *Hydrol. Processes*, *25*, 1129–1145.

Govindaraju, R. S., and A. R. Rao (2000), *Artificial Neural Networks in Hydrology*, Kluwer Acad., Dordrecht, Netherlands.

Griffiths, G. A., and B. Clausen (1997), Streamflow recession in basins with multiple water storages, *J. Hydrol.*, *190*, 60–74.

Gustard, A., and K. M. Irving (1994), Classification of the low flow response of European soils, in FRIEND: Flow Regimes from International Experimental and Network Data (Proceedings of the Braunschweig Conference, October 1993), IAHS Publ. 221, edited by P. Seuna, A. Gustard, N. W. Arnell, and G. A. Cole, IAHS, Wallingford, UK, pp. 113–117.

Haberlandt, U., B. Klöcking, V. Krysanova, and A. Becker (2001), Regionalisation of the base flow index from dynamically simulated flow components—A case study in the Elbe River Basin, *J. Hydrol.*, *248*, 35–53.

Hagan, M. T. (1994), Training feedforward networks with the Marquardt algorithm, *IEEE Trans. Neural Networks*, *5*, 989–993.

Hall, F. R. (1968), Base-flow recessions—A review, *Water Resour. Res.*, *4*, 973–983.

Hansen, L. K., and P. Salamon (1990), Neural network ensembles, *IEEE Trans. Pattern Anal. Mach. Intel.*, *12*, 993–1001.

Hansen, M., R. S. DeFries, J. R. G. Townshend, M. Carroll, C. Dimiceli, and R. A. Sohlberg (2003), Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS Vegetation Continuous Fields algorithm, *Earth Interact.*, *7*, 1–15.

Hartmann, J., and N. Moosdorf (2012), The new global lithological map database GLiM: A representation of rock properties at the Earth surface, *Geochem. Geophys. Geosyst.*, *13*, Q12004, doi:10.1029/2012GC004370.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis (2005), Very high resolution interpolated climate surfaces for global land areas, *Int. J. Climatol.*, *25*, 1965–1978.

Huete, A. R., K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira (2002), Overview of the radiometric and biophysical

performance of the MODIS vegetation indices, *Remote Sens. Environ.*, *83*, 195–213.

Hughes, G. O. (1997), An analysis of baseflow recession in the Republic of South Africa, master's thesis, Dep. Agric. Eng., Univ. of Natal, Pietermaritzburg, South Africa.

Janeau, J. L., J. P. Bricquet, O. Planchon, and C. Valentin (2003), Soil crusting and infiltration on steep slopes in northern Thailand, *Eur. J. Soil Sci.*, *54*, 543–554.

Johnson, D. H. (1999), The insignificance of statistical significance testing, *J. Wildlife Manage.*, *63*, 763–772.

Johnston, R. M., et al. (2003), ASRIS: The database, *Aust. J. Soil Res.*, *41*, 1021–1036.

Krakauer, N. Y., and M. Temimi (2011), Stream recession curves and storage variability in small watersheds, *Hydrol. Earth Syst. Sci.*, *15*, 2377–2389.

Lacey, G., and R. Grayson (1998), Relating baseflow to catchment properties in south-eastern Australia, *J. Hydrol.*, *204*, 231–250.

Lee, C., W. Chen, and P. Lee (2006), Estimation of groundwater recharge using water balance coupled with base-flow-record estimation and stable-base-flow analysis, *Environ. Geol.*, *51*, 73–82.

Levenberg, K. (1944), A method for the solution of certain problems in least squares, *Q. Appl. Math.*, *5*, 164–168.

Li, H., M. Huang, M. S. Wigmosta, Y. Ke, A. M. Coleman, L. R. Leung, A. Wang, and D. M. Ricciuto (2011), Evaluating runoff simulations from the Community Land Model 4.0 using observations from flux towers and a mountainous watershed, *J. Geophys. Res.*, *116*, doi:10.1029/2011JD016276.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land surface water and energy fluxes for general circulation models, *J. Geophys. Res.*, *99*, 14,415–14,428.

Longobardi, A., and P. Villani (2008), Baseflow index regionalization analysis in a Mediterranean area and data scarcity context: Role of the catchment permeability index, *J. Hydrol.*, *355*, 63–75.

Luojus, K., J. Pulliainen, M. Takala, J. Lemmetyinen, C. Derksen, and L. Wang (2010), Snow water equivalent (SWE) product guide, technical report, GlobSnow Consortium.

Maier, H. R., and G. C. Dandy (2000), Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications, *Environ. Modell. Software*, *15*, 101–124.

Marquardt, D. (1963), An algorithm for least-squares estimation of nonlinear parameters, *SIAM J. Appl. Math.*, *11*, 431–441.

Mazvimavi, D., A. M. J. Meijerink, H. H. G. Savenije, and A. Stein (2005), Prediction of flow characteristics using multiple regression and neural networks: A case study in Zimbabwe, *Phys. Chem. Earth*, *30*, 639–647.

McGlynn, B. L., J. J. McDonnell, J. Seibert, and C. Kendall (2004), Scale effects on headwater catchment runoff timing, flow sources, and groundwater-streamflow relations, *Water Resour. Res.*, *40*, W07504, doi:10.1029/2003WR002494.

Milly, P. C. D. (1994), Climate, soil water storage, and the average annual water balance, *Water Resour. Res.*, *30*, 2143–2156.

Mwakalila, S., J. Feyen, and G. Wyseure (2002), The influence of physical catchment properties on baseflow in semi-arid environments, *J. Arid Environ.*, *52*, 245–258.

Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models. Part I: A discussion of principles, *J. Hydrol.*, *10*, 282–290.

Nathan, R. J., and T. A. McMahon (1990), Evaluation of automated techniques for baseflow and recession analysis, *Water Resour. Res.*, *26*, 1465–1473.

Neff, B. P., S. M. Day, A. R. Piggott, and L. M. Fuller (2005), *Base flow in the Great Lakes Basin*, U.S. Geol. Surv. Sci. Invest. Rep. 2005–5217, Reston, Va.

Nicholls, N. (2001), Commentary and analysis: The insignificance of significance testing, *Bull. Am. Meteorol. Soc.*, *82*, 981–986.

Nijssen, B., G. M. O'Donnell, D. P. Lettenmaier, D. Lohmann, and E. F. Wood (2001), Predicting the discharge of global rivers, *J. Clim.*, *14*, 3307–3323.

Niu, G.-Y., et al. (2011), The community Noah land surface model with multiparameterization options (Noah-MP). 1: Model description and evaluation with local-scale measurements, *J. Geophys. Res.*, *116*, D12109, doi:10.1029/2010JD015139.

Olden, J. D., and D. A. Jackson (2002), Illuminating the ''black box'': Understanding variable contributions in artificial neural networks, *Ecol. Modell.*, *154*, 135–150.

Olden, J. D., M. K. Joy, and R. G. Death (2004), An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data, *Ecol. Modell.*, *178*, 389–397.

Oleson, K. W., D. M. Lawrence, G. B. Bonan, M. G. Flanner, E. Kluzek, P. J. Lawrence, S. Levis, S. C. Swenson, and P. E. Thornton (2010), Technical description of version 4.0 of the Community Land Model (CLM), technical report, Clim. and Global Dyn. Div., Natl. Center for Atmos. Res., Boulder, Colo.

Olson, D. M., et al. (2001), Terrestrial ecoregions of the world: A new map of life on Earth, *Bioscience*, *51*, 933–938.

Ouarda, T. B. M. J., and C. Shu (2009), Regional low-flow frequency analysis using single and ensemble artificial neural networks, *Water Resour. Res.*, *45*, W11428, doi:10.1029/2008WR007196.

Oudin, L., V. Andréassian, C. Perrin, C. Michel, and N. Le Moine (2008), Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments, *Water Resour. Res.*, *44*, W03413, doi:10.1029/2007WR006240.

Parajka, J., R. Merz, and G. Blöschl (2005), A comparison of regionalisation methods for catchment model parameters, *Hydrol. Earth Syst. Sci.*, *9*, 157–171.

Peel, M. C., F. H. S. Chiew, A. W. Western, and T. A. McMahon (2000), Extension of unimpaired monthly streamflow data and regionalisation of parameter values to estimate streamflow in ungauged catchments, Rep. prepared for the Australian National Land and Water Resources Audit, Cent. for Environ. Appl. Hydrol., Univ. of Melbourne, Parkville, Victoria.

Peel, M. C., B. L. Finlayson, and T. A. McMahon (2007), Updated world map of the Köppen-Geiger climate classification, *Hydrol. Earth Syst. Sci.*, *11*, 1633–1644.

Peña-Arancibia, J. L., A. I. J. M. Van Dijk, M. Mulligan, and L. A. Bruijnzeel (2010), The role of climatic and terrain attributes in estimating baseflow recession in tropical catchments, *Hydrol. Earth Syst. Sci.*, *14*, 2193–2205.

Petersen, T., N. Devineni, and A. Sankarasubramanian (2012), Seasonality of monthly runoff over the continental United States: Causality and relations to mean annual and mean monthly distributions of moisture and energy, *J. Hydrol.*, *468–469*, 139–150.

Post, D. A., and A. J. Jakeman (1996), Relationships between catchment attributes and hydrological response characteristics in small Australian mountain ash catchments, *Hydrol. Processes*, *10*, 877–892.

Potter, N. J., L. Zhang, P. C. D. Milly, T. A. McMahon, and A. J. Jakeman (2005), Effects of rainfall seasonality and soil moisture capacity on mean annual water balance for Australian catchments, *Water Resour. Res.*, *41*, W06007, doi:10.1029/2004WR003697.

Price, K. (2011), Effects of watershed topography, soils, land use, and climate on baseflow hydrology in humid regions: A review, *Prog. Phys. Geogr.*, *35*, 465–492.

Royall, R. M. (1986), The effect of sample size on the meaning of significance tests, *Am. Stat.*, *40*, 313–315.

Santhi, C., P. M. Allen, R. S. Muttiah, J. G. Arnold, and P. Tuppad (2008), Regional estimation of base flow for the conterminous United States by hydrologic landscape regions, *J. Hydrol.*, *351*, 139–153.

Sarle, W. S. (1995), Stopped training and other remedies for overfitting, in paper presented at 27th Symposium on the Interface of Computing Science and Statistics.

Sarle, W. S. (2000), How to measure importance of inputs?, technical report, SAS Inst., Cary, N. C.

Schneider, M. K., F. Brunner, J. M. Hollis, and C. Stamm (2007), Towards a hydrological classification of European soils: Preliminary test of its predictive power for the base flow index using river discharge data, *Hydrol. Earth Syst. Sci.*, *11*, 1501–1513.

Shao, J. (1993), Linear model selection by cross-validation, *J. Am. Stat. Assoc.*, *88*, 486–494.

Shiklomanov, A. I., R. B. Lammers, and C. J. Vörösmarty (2002), Widespread decline in hydrological monitoring threatens pan-Arctic research, *EOS Trans. AGU*, *83*, 13–17.

Shu, C., and D. H. Burn (2004), Artificial neural network ensembles and their application in pooled flood frequency analysis, *Water Resour. Res.*, *40*, W09301, doi:10.1029/2003WR002816.

Sjöberg, J., Q. Zhang, L. Ljung, A. Benveniste, B. Deylon, P. Glorennec, H. Hjalmarsson, and A. Juditsky (1995), Nonlinear black-box modeling in system identification: A unified overview, *Automatica*, *31*, 1691–1724.

Smakhtin, V. U. (2001), Low flow hydrology: A review, *J. Hydrol.*, *240*, 147–186.

Soulsby, C., and D. Tetzlaff (2008), Towards simple approaches for mean residence time estimation in ungauged basins using tracers and soil distributions, *J. Hydrol.*, *363*, 60–74.

Stokstad, E. (1999), Scarcity of rain, stream gages threatens forecasts, *Science*, *285*, 1199–1200.

Sujono, J., S. Shikasho, and K. Hiramatsu (2004), A comparison of techniques for hydrograph recession analysis, *Hydrol. Processes*, *18*, 403–413.

Tague, C., and G. E. Grant (2004), A geological framework for interpreting the low-flow regimes of Cascade streams, Willamette River Basin, Oregon, *Water Resour. Res.*, *40*, W04303, doi:10.1029/2003WR002629.

Tebaldi, C., and R. Knutti (2007), The use of the multi-model ensemble in probabilistic climate projections, *Philos. Trans. R. Soc. A*, *365*, 2053–2075.

Tetko, I. V., D. J. Livingstone, and A. I. Luik (1995), Neural network studies. 1: Comparison of overfitting and overtraining, *J. Chem. Inform. Model.*, *35*, 826–833.

Tomkins, K. M. (2013), Uncertainty in streamflow rating curves: Methods, controls and consequences, *Hydrol. Processes*, in press.

Trabucco, A., R. J. Zomer, D. A. Bossio, O. van Straaten, and L. V. Verchot (2008), Climate change mitigation through afforestation/reforestation: A global analysis of hydrologic impacts with four case studies, *Agric. Ecosyst. Environ.*, *126*, 81–97.

Tromp-van Meerveld, H. J., N. E. Peters, and J. J. McDonnell (2007), Effect of bedrock permeability on subsurface stormflow and the water balance of a trenched hillslope at the Panola Mountain Research Watershed, Georgia, USA, *Hydrol. Processes*, *21*, 750–769.

Tucker, C. J. (1979), Red and photographic infrared linear combinations for monitoring vegetation, *Remote Sens. Environ.*, *8*, 127–150.

Uchida, T., Y. Asano, N. Ohte, and T. Mizuyama (2003), Seepage area and rate of bedrock groundwater discharge at a granitic unchanneled hillslope, *Water Resour. Res.*, *39*, 1018, doi:10.1029/2002WR001298.

Van Beek, L. P. H., Y. Wada, and M. F. P. Bierkens (2011), Global monthly water stress. 1: Water balance and water availability, *Water Res. Res.*, *47*, W07517, doi:10.1029/2010WR009791.

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

Van Dijk, A. I. J. M., J. L. Peña-Arancibia, E. F. Wood, J. Sheffield, and H. E. Beck (2013), Global analysis of seasonal streamflow predictability using an ensemble prediction system and observations from 6192 small catchments worldwide, *Water Resour. Res.*, *49*, 2729–2746, doi:10.1002/wrcr.20251.

Viney, N. R., et al. (2009), Assessing the impact of land use change on hydrology by ensemble modelling (LUCHEM). II: Ensemble combinations and predictions, *Adv. Water Resour.*, *32*, 147–158.

Vogel, R. M., and C. N. Kroll (1992), Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics, *Water Resour. Res.*, *28*, 2451–2458.

Vogel, R. M., and C. N. Kroll (1996), Estimation of baseflow recession constants, *Water Resour. Manage.*, *10*, 303–320.

Walsh, R. P. D., and D. M. Lawler (1981), Rainfall seasonality: Description, spatial patterns and change through time, *Weather*, *36*, 201–208.

Wandishin, M. S., S. L. Mullen, D. J. Stensrud, and H. E. Brooks (2001), Evaluation of a short-range multimodel ensemble system, *Mon. Weather Rev.*, *129*, 729–747.

Widén-Nilsson, E., S. Halldin, and C. Xua (2007), Global water-balance modelling with WASMOD-M: Parameter estimation and regionalisation, *J. Hydrol.*, *340*, 105–118.

Wittenberg, H. (1999), Baseflow recession and recharge as nonlinear storage processes, *Hydrol. Processes*, *13*, 715–726.

Wolock, D. M. (2003), Base-flow index grid for the conterminous United States, *U.S. Geol. Surv., Open File Rep.* 03–263.

Wolock, D. M., and G. J. McCabe (1999), Explaining spatial variability in mean annual runoff in the conterminous United States, *Clim. Res.*, *11*, 149–159.

Wolock, D. M., G. M. Hornberger, K. J. Beven, and W. G. Campbell (1989), The relationship of catchment topography and soil hydraulic characteristics to lake alkalinity in the northeastern United States, *Water Resour. Res.*, *25*, 829–837.

Wolock, D. M., T. C. Winter, and G. McMohan (2004), Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses, *Environ. Manage.*, *34*, 71–88.

Yadav, M., T. Wagener, and H. Gupta (2007), Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins, *Adv. Water Resour.*, *30*, 1756–1774.

Zecharias, Y. B., and W. Brutsaert (1988), The influence of basin morphology on groundwater outflow, *Water Resour. Res.*, *24*, 1645–1650.

Zhang, Z., T. Wagener, P. Reed, and R. Bhushan (2008), Reducing uncertainty in predictions in ungauged basins by combining hydrologic indices regionalization and multiobjective optimization, *Water Resour. Res.*, *44*, W00B04, doi:10.1029/2008WR006833.