

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a preprint version which may differ from the publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/168920>

Please be advised that this information was generated on 2020-01-24 and may be subject to change.

Developing and testing a global-scale regression model to quantify mean annual streamflow

Valerio Barbarossa^{a,*}; Mark A.J. Huijbregts^{a,b}; A. Jan Hendriks^a; Arthur H.W. Beusen^c; Julie Clavreul^d; Henry King^d; Aafke M. Schipper^b

^aRadboud University, Institute for Water and Wetland Research, Department of Environmental Science, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands.

^bPBL, Netherlands Environmental Assessment Agency, Department of Nature and Rural Areas, P.O. Box 303, 3720 AH Bilthoven, The Netherlands.

^cPBL, Netherlands Environmental Assessment Agency, Department of Information, Data and Methodology, P.O. Box 303, 3720 AH Bilthoven, The Netherlands.

^dUnilever R&D, Safety and Environmental Assurance Centre, Colworth Science Park, Sharnbrook, Bedfordshire MK44 1LQ, UK.

*Corresponding author: vbarbarossa@science.ru.nl

This document is the unedited Author's version of a Submitted Work that was subsequently accepted for publication in the Journal of Hydrology, Copyright © 2017 Elsevier B.V. after peer review. To access the final edited and published work see <http://dx.doi.org/10.1016/j.jhydrol.2016.11.053>.

Abstract

Quantifying mean annual flow of rivers (MAF) at ungauged sites is essential for assessments of global water supply, ecosystem integrity and water footprints. MAF can be quantified with spatially explicit process-based models, which might be overly time-consuming and data-intensive for this purpose, or with empirical regression models that predict MAF based on climate and catchment characteristics. Yet, regression models have mostly been developed at a regional scale and the extent to which they can be extrapolated to other regions is not known. In this study, we developed a global-scale regression model for MAF based on a dataset unprecedented in size, using observations of discharge and catchment characteristics from 1,885 catchments worldwide, measuring between 2 and 10^6 km². In addition, we compared the performance of the regression model with the predictive ability of the spatially explicit global hydrological model PCR-GLOBWB by comparing results from both models to independent measurements. We obtained a regression model explaining 89% of the variance in MAF based on catchment area, mean annual precipitation and air temperature, average slope and elevation. The regression model performed better than PCR-GLOBWB for the prediction of MAF, as root-mean-square error (RMSE) values were lower (0.29 - 0.38 compared to 0.49 - 0.57) and the modified index of agreement (*d*) was higher (0.80 - 0.83 compared to 0.72 - 0.75). Our regression model can be applied globally to estimate MAF at any point of the river network, thus providing a feasible alternative to spatially explicit process-based global hydrological models.

Keywords: mean annual discharge; river discharge; global hydrology; empirical modelling; predictions in ungauged basins; scaling relationships; model comparison; PCR-GLOBWB; spatial error model.

1. Introduction

Mean annual discharge or flow of rivers (hereafter abbreviated as MAF) is an important indicator of global water supply, with applications in irrigation supply assessment, climate change vulnerability assessment (Chang, 2003; Santini and di Paola, 2015), hydropower assessment (Hall et al., 2004), water footprinting (Hanafiah et al., 2011; Hoekstra et al., 2011; Jefferies et al., 2012; Pfister et al., 2009; Tendall et al., 2014), and for quantifying sediment fluxes (Syvitski et al., 2003). It also represents one of the most important

factors determining the ecosystem integrity of freshwater biodiversity (Oberdorff et al., 1995; Oberdorff et al., 2011; Poff and Zimmerman, 2010; Xenopoulos and Lodge, 2006; Xenopoulos et al., 2005). Despite its importance, streamflow data availability is limited, and monitoring is in rapid decline since the mid-1980s (Shiklomanov et al., 2002). Modelling approaches have long been used to estimate MAF at ungauged sites and are generally divided into two categories: spatially explicit process-based models and regression-based empirical models.

State-of-the-art spatially explicit numerical models for global-scale calculations of streamflow are Macroscale Hydrological Models (MHM) or Global Hydrology and Water Resources Models (GHWM) (Alcamo et al., 2003; Gosling and Arnell, 2011; Hanasaki et al., 2008; van Beek and Bierkens, 2008; Van Der Knijff et al., 2010; Widén-Nilsson et al., 2007; Wisser et al., 2010). As these models account for the spatial variability of the physical processes involved within catchment hydrology and are capable of predicting streamflow even at the daily time scale, they are computationally and data intensive.

Regression-based approaches to calculate MAF are less time-consuming and computationally less intensive. Moreover, regression equations relating streamflow to explanatory catchment characteristics like upstream drainage area, precipitation and temperature may help to better understand general hydrological patterns and processes across different scales (Burgers et al., 2013; Farmer et al., 2015). However, to date, regression-based approaches relating mean annual streamflow to catchment characteristics have been mainly applied at a regional scale (Hortness and Berenbrock, 2001; Stuckey, 2006; Tran et al., 2015; Verdin and Worstell, 2008; Vogel et al., 1999) or to specific climate zones (Syvitski et al., 2003), and the extent to which these models can be extrapolated to other regions is not known. Regression relationships at the global scale have hardly been established so far. An exception is Burgers et al. (2013), who derived MAF relationships at a global scale using precipitation and catchment area as predictors. However, their model explained only 56% of the variance in MAF, which is low compared to the range of 77-99% achieved by regional regression models (e.g. Verdin and Worstell (2008)). Yet, the regional studies typically included a larger number of predictors, which suggests that the explanatory power of a global-scale regression model may increase if relevant predictors are added. In addition, the applicability of global regression relationships

for the prediction of mean annual streamflow has not yet been tested. Therefore, the aim of this study was twofold: 1) to establish an empirical regression model relating MAF to easily retrievable catchment characteristics at the global scale; 2) to test the predictive ability of the regression model in a backcasting analysis and compare its performance with the predictive performance of PCR-GLOBWB, a spatially explicit MHM (van Beek et al., 2011). To our knowledge, our study is the first to make an explicit comparison of the predictive abilities of a process-based and a regression-based global-scale model.

We based our regression model on measured long-term average MAF from 1,885 catchments worldwide, ranging from 2 km² to 10⁶ km² in size. We used five predictor variables, including two climatic variables – mean annual precipitation and air temperature – and three geomorphologic variables – area, mean slope and mean elevation of the catchment. Drainage area, mean annual precipitation and mean annual temperature are often used as predictors of MAF in regional regression modelling studies (Verdin and Worstell, 2008; Vogel and Sankarasubramanian, 2000; Vogel et al., 1999). The dependence of MAF on drainage area is a well-accepted power relationship reflecting the self-similarity of river systems (Rodríguez-Iturbe and Rinaldo (2001). Mean annual precipitation represents the potential runoff of the catchment, as it equals the amount of water supplied to the catchment (Thomas and Benson, 1970). We selected the mean annual temperature as a proxy for the potential evapotranspiration (PET), because temperature is a major determinant of evapotranspiration (Hamon, 1963; Lu et al., 2005; Thornthwaite, 1948). Furthermore, previous regression analyses of MAF have shown an increased explained variance when additional geomorphologic parameters were considered (Hortness and Berenbrock, 2001; Stuckey, 2006; Vogel et al., 1999). Therefore, we included average slope and elevation of the catchment as additional predictors in our study. Although elevation and slope alone may not directly influence MAF, they may serve as proxies for other factors causing inter-basin streamflow variation which are difficult to measure, e.g. radiation, wind, vegetation and basin ruggedness (Thomas and Benson, 1970).

2. Materials and methods

2.1. Mean annual discharge data

We retrieved worldwide MAF data from the Global Runoff Data Centre (GRDC) database, which provides daily or monthly observations of 9,213 gauging stations monitored from 1806 to 2015, with variable record length (GRDC, 2015). The GRDC has spent more than 25 years gathering river discharge data from the National Hydrological Services of all the World Meteorological Organization (WMO) state members, which has resulted in a discharge dataset unprecedented in size. For example, the SAGE Global River Discharge Database (<http://nelson.wisc.edu/sage/data-and-models/riverdata/>) and the RivDis database (Vorosmarty et al., 1998) provide discharge data for 3,500 and 1,018 stations, respectively. The accuracy of the discharge measurements included in the GRDC database is estimated to be about 10 – 20% (Syvitski et al., 2005). For our model development we selected discharge data for the period 1981 – 2010. We used a 30 year period because this is in accordance with the recommendations for climate analyses (World Meteorological Organization, 1992). We excluded years after 2010 because of a decrease in data availability for the most recent years. We averaged the daily discharge data over each year, using only those years where 100% of the daily observations were available. Next, we averaged the yearly observations for the period 1981 - 2010 in order to obtain a long-term mean annual discharge for each catchment. We selected monitoring stations with at least 15 yearly average discharge values in order to obtain representative long-term mean values (Kennard et al., 2010). This resulted in a dataset of 1,885 observations out of the 2,759 available GRDC gauged catchments with observations in the time range 1981 - 2010 (**Figure 1**).

2.2. Catchment characteristics

We retrieved catchment-specific values for catchment area (A), average altitude (H), average slope (S), 30-years average temperature (T) and 30-years average precipitation (P) from a combination of data sources (**Table 1**). Catchment area was retrieved from the GRDC database, which includes a georeferenced map of the upstream catchment corresponding with each gauging station (GRDC, 2011).

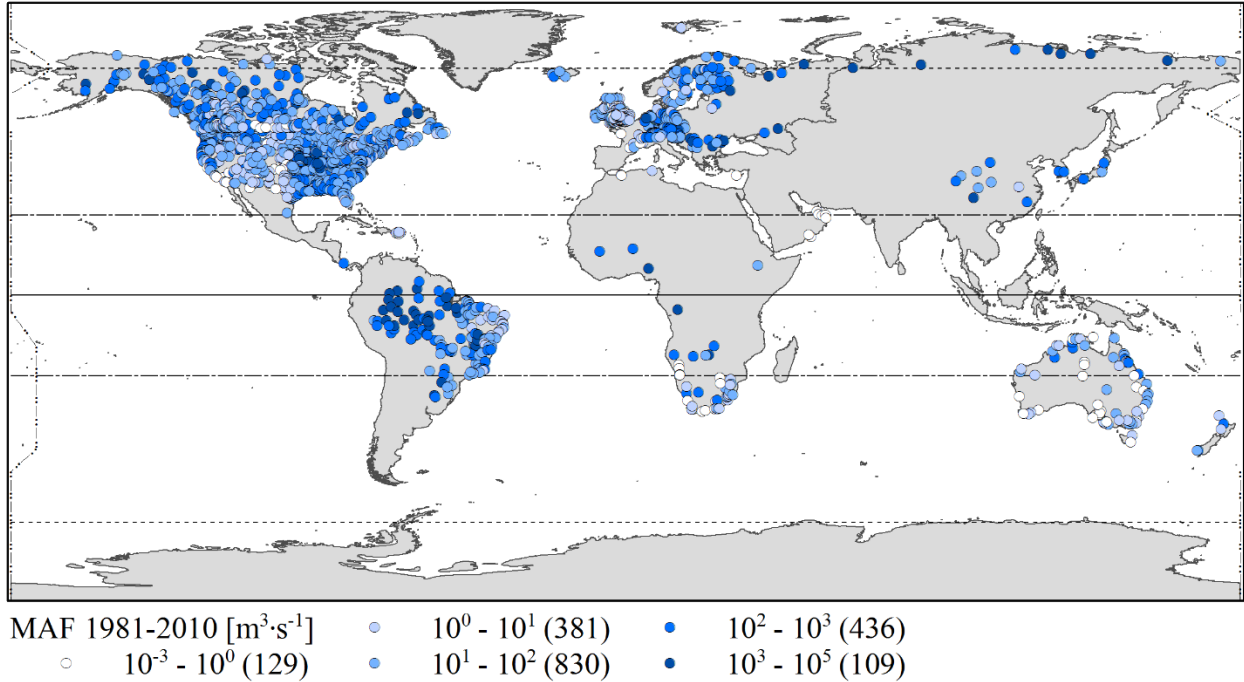


Figure 1. Distribution of the 1,885 GRDC gauging stations monitored for at least 15 years in the 1981-2010 period. The stations are grouped (graduated symbols) based on the mean annual flow (MAF) recorded at each station. Next to each MAF category, the number of observations are provided in brackets.

Catchment boundaries in this map have been established based on the HydroSHEDS drainage network, a 15 arc-seconds hydrological map derived from 3 arc-seconds elevation data of the National Aeronautics and Space Administration (NASA) Shuttle Radar Topography Mission (SRTM), extended with the hydro1k hydrological network for latitudes above 60N, which are not covered by the SRTM data (GRDC, 2011; Lehner et al., 2008).

We derived altitude from the WorldClim digital elevation model (DEM), which is a 30 arc-seconds DEM based upon the SRTM elevation data extended with GTOPO30 elevation data for latitudes above 60N (Hijmans et al., 2005). We derived a raster slope map from the WorldClim DEM using the ‘average maximum technique’ in ArcGIS, similarly to Hortness and Berenbrock (2001). For precipitation and temperature, we averaged 30 arc-minutes resolution monthly raster maps from the Climatic Research Unit time series (CRU TS) 3.23 to annual values, and consequently averaged over the period 1981-2010 (Harris et al., 2014; University of East Anglia Climatic Research Unit, 2013).

We resampled the raster maps obtained for precipitation, temperature, altitude and slope in order to match the 15 arc-seconds resolution of the HydroSHEDS drainage network, upon which the watershed boundaries were established. We then calculated a single mean value of each variable for each catchment corresponding with a GRDC gauging station, as described by Syvitski et al. (2003). Resampling and averaging were performed in ArcGIS 10.3. An overview of the summary statistics of the variables is available in **Table 1**.

Table 1. Summary statistics of the mean annual streamflow (MAF) and catchment characteristics of 1,885 gauging stations in the period 1981-2010

Variable	Symbol	Unit	Mean	Median	SD ^a	Min ^a	Max ^a	γ_1^a	Source database
MAF	Q	m ³ ·s ⁻¹	3.72·10 ²	3.38·10 ¹	1.99·10 ³	3.47·10 ⁻³	4.73·10 ⁴	14.2	GRDC ^b
Catchment area	A	m ²	4.58·10 ¹⁰	4.61·10 ⁹	2.05·10 ¹¹	2.00·10 ⁶	3.63·10 ¹²	9.7	GRDC ^c
Altitude	H	m	6.84·10 ²	4.52·10 ²	6.05·10 ²	1.35·10 ¹	4.76·10 ³	1.6	WorldClim ^d
Slope	S	(°)	2.31·10 ⁰	1.27·10 ⁰	2.65·10 ⁰	4.71·10 ⁻²	1.56·10 ¹	2.0	WorldClim ^d
Precipitation	P	m·s ⁻¹	2.78·10 ⁻⁸	2.46·10 ⁻⁸	1.46·10 ⁻⁸	3.53·10 ⁻⁹	1.09·10 ⁻⁷	1.3	CRU TS 3.23 ^e
Temperature	T	°C	8.96·10 ⁰	8.24·10 ⁰	7.51·10 ⁰	-1.67·10 ¹	2.76·10 ¹	0.4	CRU TS 3.23 ^e

^aSD = standard deviation; Min = minimum; Max = maximum; γ_1 = skewness

^b(GRDC, 2015)

^c(GRDC, 2011)

^d(Hijmans et al., 2005)

^e(Harris et al., 2014)

2.3 Model fitting

Methods available for correlative modelling range from parametric and non-parametric regression-based approaches to machine-learning techniques (Chen et al., 2015; Danandeh Mehr et al., 2013; Fan et al., 2015; Okkan and Serbes, 2012; Wang et al., 2015; Wu et al., 2009). For the present study we selected ordinary least squares (OLS) regression because it results in an explicit equation, which facilitates interpretation and comparison with other studies. All the variables except for temperature were log-transformed to avoid heteroscedasticity as they revealed a right-skewed distribution (**Table 1**), in agreement with the choices made in previous studies for similar variables (Burgers et al., 2013; Hendriks et al., 2012; Syvitski et al., 2003; Verdin and Worstell, 2008; Vogel et al., 1999). This resulted in the following linear regression equation:

$$\log_{10}Q = \beta_0 + \beta_A \cdot \log_{10}A + \beta_P \cdot \log_{10}P + \beta_T \cdot T + \beta_S \cdot \log_{10}S + \beta_H \cdot \log_{10}H + \varepsilon \quad (1)$$

where β_0 is the intercept, $\langle \beta_A, \beta_P, \beta_T, \beta_S, \beta_H \rangle$ is the vector of the regression coefficients associated with the predictor variables (symbols are described in **Table 1**) and ε is the error term. The back-transformed form to real scale of **Equation 1** yields the nonlinear formulation:

$$Q = 10^{\beta_0} \cdot A^{\beta_A} \cdot P^{\beta_P} \cdot 10^{\beta_T \cdot T} \cdot S^{\beta_S} \cdot H^{\beta_H} \cdot 10^\varepsilon \quad (2)$$

Prior to performing the regression analysis, we assessed multi-collinearity among the predictors using Variance Inflation Factors (VIFs), employing the function “vif” of the package “HH” (Heiberger, 2015) in the R environment (Development Core Team, 2005). We preferred VIFs over bivariate correlation analysis because pairwise correlation coefficients do not reveal more subtle forms of multicollinearity (Field, 2009). The maximum VIF was 2.8, well below the standard threshold of 5 (Zuur et al., 2009). We then fitted OLS regression models for an increasing number of predictors. To identify the best regression model for each of the five sets of predictors (first set with one predictor variable, second set with two predictor variables, etc.) as well as the best overall model, we employed the function “dredge” of the package “MuMIn” in the R environment (Barton, 2015). Within each set, the algorithm analyzes all possible combinations of predictor variables and ranks the regression models based on a user-defined criterion. To identify the most parsimonious model for each set of predictors, we used the Akaike Information Criterion (AIC) as well as the Bayesian Information Criterion (BIC), which employs a larger penalty term for additional predictor variables. Further, we used the Cooks D influence statistic in order to identify observations that may have biased the coefficients of the regression (Cook and Weisberg, 1982).

In order to assess potential bias in the regression coefficients induced by spatial autocorrelation resulting from the nested structure of the catchments, we compared the regression coefficients with the coefficients of a spatial error (SE) model. Since spatial autocorrelation can exist within either the residuals (spatial error) or the response variable (spatial lag), we performed a preliminary test for spatial autocorrelation based on the Lagrange Multiplier test (LM test) using the R package ‘spdep’ (Anselin, 1988; Bivand et al., 2013; Bivand and Piras, 2015). We preferred the LM test to the more commonly employed Moran’s I test, for the LM test has a higher power to discriminate among either spatial error autocorrelation or spatial lag (Anselin

and Rey, 1991). The LM test showed significantly higher autocorrelation in the error term (LM test value of 465, robust LM test value of 375) than in the response variable (with values of 115 and 25, respectively). Therefore we fitted an SE model that accounts for spatial autocorrelation in the residuals, expressing the error term of **Equation 1** as: $\varepsilon = \lambda W\varepsilon + \mu$, where λ is the coefficient in the spatial autoregressive structure, W is a weight matrix defined by the inverse distance between observations, and μ is the vector of identically distributed random errors (Ord, 1975). We calculated the distances between the GRDC stations across the stream network thereby considering as neighbors only those stations belonging to the same encompassing hydrologic basin. As we employed HydroSHEDS as rivers network for the calculation of W , we fitted and compared the OLS and SE regression coefficients based on a subset of observations within 58S - 60N latitude ($n = 1,748$).

2.4. Comparison with PCR-GLOBWB: backcasting analysis

We compared the predictive performance of our best regression model with the global hydrological model PCR-GLOBWB (van Beek and Bierkens, 2008; van Beek et al., 2011). Defined as a “leaky bucket” type of model, PCR-GLOBWB calculates changes in water storage between two different soil layers, groundwater reservoir and atmosphere, forced by CRU TS 2.1 data, on a cell-by-cell basis at 30 arc-minutes resolution, for daily time steps. PCR-GLOBWB has been widely employed for assessments of global surface water and groundwater availability, nutrient transport modelling and biodiversity impact calculations (Beusen et al., 2016; Gleeson et al., 2012; Janse et al., 2015; Wada et al., 2011; Wanders and Wada, 2015). Compared to other GHMs, PCR-GLOBWB is a purely process-based model, as opposed to for example WaterGAP which is partially calibrated (Alcamo et al., 2003; Döll et al., 2003). Therefore, we considered PCR-GLOBWB a more suitable benchmark for the comparison.

We considered monitoring data of GRDC stations continuously monitored from 1971 to 1980 as an independent and common basis for the comparison between the regression model and PCR-GLOBWB. From the 2,219 GRDC stations used for the testing of PCR-GLOBWB (van Beek et al., 2011), we selected the 543 stations that were continuously monitored from 1971 through 1980 (**Figure S2**). We derived mean

annual values of temperature and precipitation from the CRU TS 3.23 for each catchment, according to the approach described in section 2.2, and calculated mean annual streamflow for each year in the time span 1971-1980, and as a 10 years average.

We evaluated and compared the performances of the regression model and PCR-GLOBWB employing root mean square error (*RMSE*) and modified index of agreement (*d*). Thus, we employed an absolute as well as a relative error measure, following the recommendations for hydrological model evaluation as provided by Legates and McCabe (1999). We adopted the index of agreement d^2 in the modified form *d* to avoid inflation of errors by squared values (Legates and McCabe, 1999). In addition, the *d* represents an improvement over the coefficient of determination (R^2) (Legates and McCabe, 1999).

Given the six orders of magnitude covered by the data, we log transformed the mean annual streamflow values. The *RMSE* is calculated as:

$$RMSE = \sqrt{\frac{1}{n \cdot m} \cdot \sum_{x=1}^m \sum_{t=1}^n (\log_{10} O_{x,t} - \log_{10} P_{x,t})^2} \quad (3)$$

where $n \cdot m$ are the dimensions of the matrix of observations of the m GRDC stations over the n years of the backcasting period, $O_{x,t}$ is the observed value for the station x at time t , and $P_{x,t}$ is the predicted value for the station x at time t .

The *d* is formulated as:

$$d = \frac{\sum_{x=1}^m \sum_{t=1}^n |\log_{10} O_{x,t} - \log_{10} P_{x,t}|}{\sum_{x=1}^m \sum_{t=1}^n (|\log_{10} P_{x,t} - \log_{10} \bar{O}| + |\log_{10} O_{x,t} - \log_{10} \bar{O}|)} \quad (4)$$

where \bar{O} stands for the overall average of the observed streamflow of all the stations across the 10 years period. The index of agreement varies between 0 and 1, with higher values indicating a better fit. When the mean annual streamflow averaged over the 10 years is considered, n becomes 1 in **Equations 3** and **4**.

3. Results

3.1 Model fitting

The OLS regression analysis revealed that the model with the full set of predictors was the most parsimonious (i.e., lowest AIC; **Table S1**). Nearly 90% of the variation in MAF could be explained by the five catchment characteristics (**Table 2**), indicating that the most relevant predictors for MAF were covered by the regression model. According to the standardized regression coefficients, which can be compared across explanatory variables to assess their relative importance (Bring, 1994), catchment area was the most important predictor of MAF, followed by precipitation, temperature, slope and elevation.

Table 2. Coefficients (raw and standardized), goodness of fit (R^2) and number of underlying observations (m) of the most parsimonious regression model $Q = 10^{\beta_0} \cdot A^{\beta_A} \cdot P^{\beta_P} \cdot 10^{\beta_T \cdot T} \cdot H^{\beta_H} \cdot S^{\beta_S}$. CI = confidence interval.

Coefficient	Value (95% CI)	Std. value	R^2	m
β_0	9.066 (8.503 - 9.630)	0	0.89	1,885
β_A	1.018 (1.001 - 1.035)	0.961		
β_P	2.070 (1.991 - 2.148)	0.486		
β_T	-0.038 (-0.040 - -0.035)	-0.290		
β_H	-0.509 (-0.565 - -0.454)	-0.212		
β_S	0.464 (0.421 - 0.507)	0.237		

The model performed better for higher MAF values (**Figure 2**). Furthermore, residual errors were slightly larger for catchments with lower precipitation values and at higher altitudes (**Figure S1**). Residuals tended to be randomly distributed in relation to catchment area, precipitation or slope (**Figure S1**). Only about 1.3% of the predicted values showed errors greater than one order of magnitude (**Figure 2**). For the Cooks D statistic, a maximum value of 0.03 was found, well below the threshold of 1 (Cook and Weisberg, 1982), meaning that none of the observations biased the regression coefficients.

The comparison of the regression coefficients between the OLS and SE regression models revealed a large overlap of the confidence intervals of the coefficients (49-86% CI overlap; **Table S2**). This indicated that the OLS regression coefficients were not significantly influenced by spatial autocorrelation (type I error).

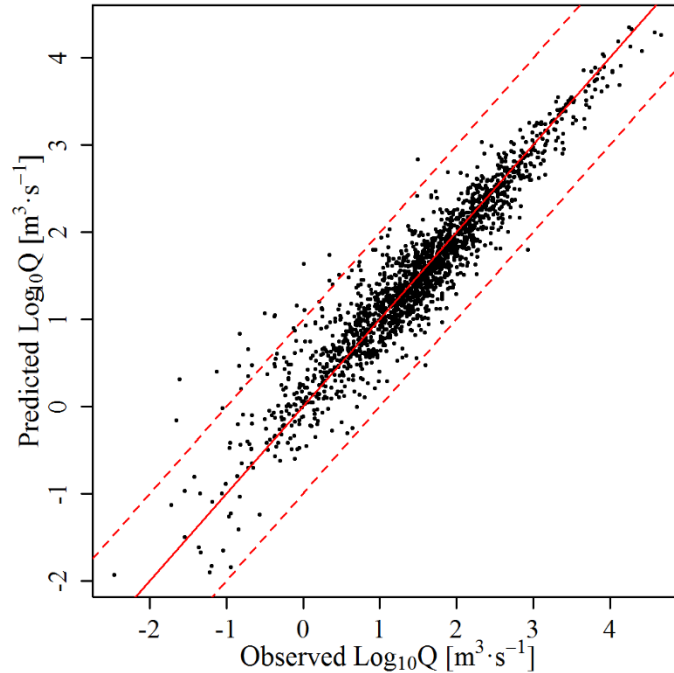


Figure 2. Predicted versus observed MAF values of the 1,885 GRDC stations employed in the regression analysis. The solid line represents perfect model fit (1:1 line) and the dashed lines represent a difference of one order of magnitude.

3.2 Performance testing on independent data and comparison with PCR-GLOBWB

The testing of the regression model on independent data in the time period 1971 - 1980 (backcasting analysis), for both single year and 10-years average MAF, revealed that the predictions of the regression model were characterized by lower RMSE values and higher d values than the predictions of the global hydrological model PCR-GLOBWB (**Figure 3**).

The PCR-GLOBWB simulation resulted in a greater number of outliers, in agreement with the higher RMSE values (**Figure 3**). For the OLS model, about 3% of the observations had residuals greater than one order of magnitude, whereas about 7% of the PCR-GLOBWB results deviated more than one order of magnitude from the measurements. PCR-GLOBWB performed slightly better than the OLS model for the highest MAF values ($> 10,000 \text{ m}^3/\text{s}$; **Figure 3**). Both models performed poorly for MAF values lower than $10 \text{ m}^3/\text{s}$. In general, residuals of PCR-GLOBWB revealed a cone-shaped distribution, with larger errors at lower discharge values. In contrast, the OLS model showed a tendency to overestimate low discharge values.

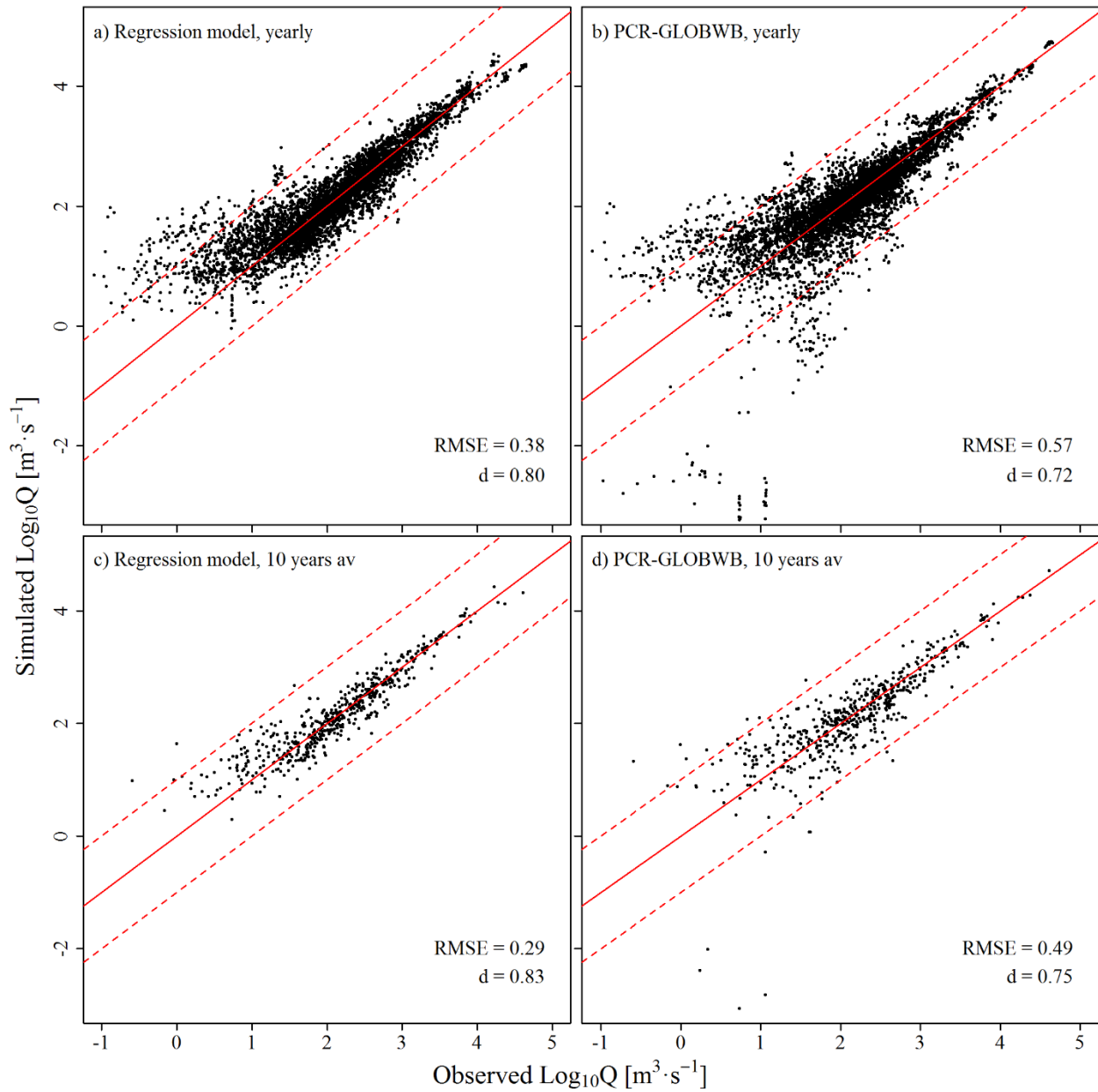


Figure 3. Results of the backcasting analysis for the period 1971-1980, showing predicted versus observed MAF for the OLS regression model (left) and the GHM PCR-GLOBWB (right), based on yearly values (top) as well as 10-year average values (bottom). Within each chart, the solid line represents perfect model fit (1:1 line) and the dashed lines define a range of accuracy of plus/minus one order of magnitude. RMSE = root mean square error; d = modified index of agreement.

Both models performed better when backcasting the 10-years average of MAF than MAF for single years. For the 10-years average MAF, the residuals greater than 1 order of magnitude reduced to about 1% (4 observations) and 4% (22 observations) for the regression model and PCR-GLOBWB, respectively.

4. DISCUSSION

4.1 Regression coefficients interpretation

We developed a global-scale multiple regression model for predicting mean annual flow of rivers based on easily retrievable input parameters. We calibrated the model on 1,885 catchments worldwide based on long-term average discharge data (1981-2010), resulting in a model explaining 89% of the variance in MAF based on catchment area and catchment-averaged precipitation, temperature, slope and altitude. The analysis revealed the catchment area to be the most important predictor of MAF (**Table 2**). Indeed, a single regression analysis based on catchment area alone already explains 61% of the variation in MAF (**Table S1**), in agreement with scaling relationships reflecting self-similarity across catchments. Nevertheless, additional predictors considerably increased the explained variance (**Table S1**). This suggests that multiple regression improves the interpretation of the spatial scaling of MAF at a global level, in agreement with recent findings that point toward multiple regression (called “multiscaling”) to improve the interpretation of scaling behavior for daily streamflow of the Southeast United States (Farmer et al., 2015).

The MAF scaled to catchment area with an exponent of about 1, implying a linear relationship between MAF and drainage area. This is in agreement with the coefficients reported by a number of regional studies using multiple regression analysis (**Table 3**), and close to the value of 0.86 reported by the global study of Burgers et al. (2013). Further, MAF scaled to precipitation with an exponent of about two. This is highly similar to the findings reported by Tran et al. (2015) (**Table 3**), who conducted a regional study covering 533 catchments. On the other hand, the coefficient for precipitation found in this study is in contrast with the linear relationship observed in Burgers et al. (2013). However, in the current study we employed about three times more observations and our regression model explained about 33% more of the variance. In general, coefficients reported for precipitation tend to be greater than 1 and converging towards 2 (**Table 3**), reflecting nonlinearity in the physical process responsible for the runoff generation (see e.g. Yu et al. (2015)). Therefore, given the global coverage, the heterogeneity of the input variables values employed to calibrate the coefficients and the large explained variance, the value of about 2 is considered to be a reasonable estimate of the exponent for precipitation.

Table 3. Regression coefficients found in this study compared with coefficients reported in regional and global studies available from the literature. n represents the number of catchments employed to calibrate the regression coefficients; A range is the range of the catchment areas employed in the respective study.

	<i>This study</i>	<i>Burgers et al. (2013)</i>	<i>Hortness and Berenbrock (2001)</i>	<i>Stuckey (2006)</i>	<i>Tran et al. (2015)</i>	<i>Vogel et al. (1999)</i>	Gyawali et al. (2015)
β_A	1.02	0.86	0.83 - 1.10	1.01	1.01	0.58 - 1.14	0.87
β_P	2.07	1.01	1.64 - 2.70	1.80	2.04	1.21 - 6.42	3.68
β_T	-0.04	-	-	-	-0.49 ^a	-7.66 - -0.51 ^a	-
β_S	0.46	-	-3.44 - 7.52 ^b	-	-	0.33 - 0.51	0.30 ^b
β_H	-0.51	-	-2.36 - 2.30	0.13	-	1.66 ^c	-
n	1,885	663	200	195	533	1,553	93
A range [km²]	2.0·10 ⁰ - 3.6·10 ⁶	7.3·10 ³ - 4.6·10 ⁶	8.0·10 ⁰ - 3.5·10 ⁴	5.6·10 ⁰ - 4.5·10 ³	1.3·10 ¹ - 1.7·10 ²	-	6.2·10 ⁰ - 4.4·10 ³
Extent	Global	Global	Idaho, USA	Pennsylvania, USA	Upper Mississippi, USA	USA	western Great Lakes, USA

^avalues refer to the log-transformed form and are therefore not directly comparable with the coefficient obtained in this study

^bvalues refer to slope in percent instead of degrees and are therefore not directly comparable with the coefficient obtained in this study

^cused only for region 16, "Great Basin" in Vogel et al. (1999)

For temperature a negative coefficient was obtained, reflecting decreasing MAF with increasing temperature. Indeed, increases in temperature would lead to an increased evapotranspiration, which eventually implies that less water is routed through the drainage network and poured at the discharge point. The negative relationship is in agreement with the results of regional studies, but as these studies were based on log-transformed temperature data, the values of the coefficients are not directly comparable with ours (Tran et al., 2015; Verdin and Worstell, 2008; Vogel et al., 1999). The relationship between MAF and altitude is rather complex. The exponent resulting from the regression analysis was negative, which can be explained by the fact that at higher altitudes the solar radiation and wind are more intense, therefore enhancing the evapotranspiration process (Tran et al., 2015). In contrast, some studies reported positive exponents for altitude (**Table 3**). A possible explanation could be that these studies did not include temperature as a predictor, which implies that the positive exponent for the altitude term may reflect an effect of temperature. At high altitudes, the lower temperature likely results in less evapotranspiration hence an increase in MAF. However, if the model does not include temperature as predictor, the larger MAF at higher altitudes may result in a positive regression coefficient for altitude instead. The positive exponent of

about ½ obtained for slope (S) compares well with the values reported in regional studies and reflects that in catchments with steeper slopes, the runoff and consequently the MAF is enhanced.

4.2 Performance comparison with PCR-GLOBWB

Overall, our regression model performed better than PCR-GLOBWB when applied to an independent test dataset. Differences were apparent in particular for smaller catchments (**Figure 3**). The residuals (absolute values) of PCR-GLOBWB for the 10-years average MAF revealed a significant negative trend in relation to area (p-value < 0.01, $R^2 = 0.03$), while for the regression model no significant correlation was found. However, new global hydrological models with greater spatial resolution than the 30 arc-minutes version of PCR-GLOBWB employed in this study may achieve better results, especially for smaller catchments (e.g. see the list of models provided in Bierkens (2015)). Yet, such refined models are more demanding in terms of computational costs (Bierkens, 2015), and might therefore be more suitable when monthly or daily discharge values are needed.

Both the regression model and PCR-GLOBWB performed worse for water-scarce regions, as revealed by larger errors at higher dryness ratio values (**Figure 4**). The dryness ratio reflects water losses due to evapotranspiration relative to the amount of precipitation. It is defined as the actual annual evapotranspiration divided by the total annual precipitation, where the actual evapotranspiration is calculated as the annual precipitation minus the unit discharge, in turn obtained by dividing the discharge by the area of the catchment (Vogel et al., 1999). The larger errors with higher dryness ratios are likely due to a combination of higher uncertainty in the precipitation values for water-scarce regions and hydrological processes that are particularly relevant in dry regions yet not described by the models (Döll et al., 2003; van Beek et al., 2011; Vogel et al., 1999). Examples of such processes include the almost instantaneous evaporation from many ephemeral post-rainfall ponds and relatively large losses from the river channel to groundwater (Döll et al., 2003). In addition, water abstraction by human activities is likely to affect the natural flow in water scarce regions more than in wet regions, thus providing an additional possible explanation for the overestimation of MAF in dry regions.

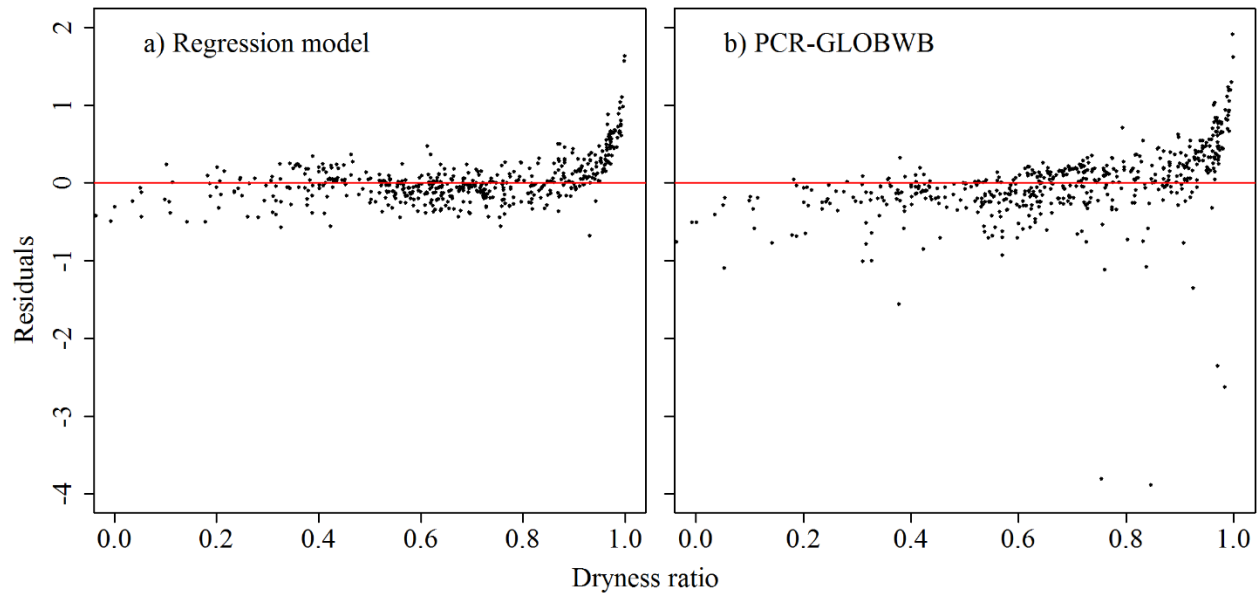


Figure 4. Plot of the residuals of the backcasting on the 10-year average MAF vs the dryness ratio for a) the regression model and b) PCR-GLOBWB. Note that 18 values (about 3% of total data) with dryness ratios smaller than 0 were excluded for clarity of representation.

4.3 Applicability of the regression model

The residuals of our regression model were not related to area (**Figure S1-b**), suggesting that the model is area-independent and maintains similar performance across catchments ranging across at least six orders of magnitude in size (2 to 10^6 km²). Although the distribution of the monitoring stations employed for the calibration of the regression coefficients was skewed towards America and Europe (**Figure 1**), the residuals of the model-application (yearly and 10 years average MAF) were consistent across different continents (**Figure S3**). This is supported by the wide range of latitudes covered by the monitoring stations, reflected by a large range in precipitation and temperature values (**Table 1**). This indicates that the model developed by this study can be applied to predict MAF at any point of any river network globally, taking into account the weaker prediction power for water scarce regions (see section 4.2). As such, the model is most suitable for assessments of water availability and ecological integrity in relation to changing future climatic conditions.

The model requires only a small number of input parameters, namely catchment area, catchment-averaged precipitation, catchment-averaged temperature, catchment-averaged altitude and catchment-averaged slope. Catchment area as well as slope and altitude can be easily derived from a flow direction raster map

and a digital elevation model with standard Geographic Information System (GIS) tools. Annual mean precipitation and temperature can be obtained from observations covering a given period of interest. If predictor variable values are within the range of values employed in the calibration phase (**Table 1**), the uncertainty of the predictions is known. Application of the model to predictor values outside the calibration domain results in MAF values with unknown uncertainty. In addition to this, given the fact that the model was calibrated on climatic data different from the CRU TS 3.23 (Harris et al., 2014), with a river network different from HydroSHEDS 15-sec (Lehner et al., 2008) or with another DEM else than the one provided by the WorldClim database (Hijmans et al., 2005), we acknowledge that the model is not valid with other input sources.

The regression model performed worse for extreme MAF values when applied at finer temporal scales, as exemplified by the decreased performance of the model on a year-by-year basis (**Figure 3**). This is due to the fact that the regression coefficients have been calibrated on 30 years averaged data and therefore when the model is applied at a finer temporal scale it would underestimate high MAF values and overestimate low MAF values. Yet, the bias due to temporal downscaling was relatively low at high MAF values (**Figure 3**), suggesting that the overestimation of MAF at low values is related to the decreased model performance in water-scarce regions (**Figure 4**) rather than the temporal downscaling. Implementing catchment aridity as predictor in a regression model is, however, a non-trivial problem. Attempts to include an interaction term of precipitation and temperature to describe catchment aridity within this study were inconclusive due to the scarce increase in the explained variance and multicollinearity issues. Furthermore, including complex multi-variable predictors would go beyond the scope of developing a conceptually simple and easily applicable model. In addition, process-based models like PCR-GLOBWB which implement the evapotranspiration process in a more mechanistic way reveal the same limitations when predicting MAF in arid catchments. Therefore, the scarcity and uncertainty of monitored precipitation and discharge values in such regions represent a major source of bias. Hence, further research should focus on improving models for predicting mean annual flow in water scarce regions, which is particularly relevant in assessments of

water availability and water footprints. In addition, future work would benefit from an increase in the quality and quantity of monitoring data, for precipitation and discharge in particular.

5. CONCLUSIONS

We presented a conceptually simple model for predicting mean annual flow of rivers globally at any point of the river network. The model explained 89% of the variance in MAF based on observations retrieved from 1,885 catchments worldwide. The regression coefficients obtained were within the ranges reported by previous regional-scale studies and indicated that MAF scales linearly to catchment area, while it scales nonlinearly to precipitation with an exponent of about 2. Temperature, slope and altitude, which have not been used before in global regression models for MAF, further improve the explained variance. Our model can be applied to estimate MAF at any ungauged site in the river network globally. However, it should be noted that the model is valid only for input parameters within the range of the calibration variables and therefore, outside the applicability domain the uncertainty in the estimation of MAF is unknown. Application of the model to predict 1-year and 10-year average MAF based on independent test data revealed that only 3% and 1% of the simulated MAF values deviated more than one order of magnitude from the measurements, respectively. In addition to this, our model performed slightly better than the widely employed macro hydrological model PCR-GLOBWB, particularly for smaller catchments. Both the model developed in this study and PCR-GLOBWB performed worse for water-scarce regions by overestimating the MAF, due to the increased uncertainty in rainfall and discharge observations and to the difficulty in describing the catchment hydrology in such regions. This implies that in dry regions, our model should be applied very cautiously.

It is recommended that forthcoming studies on global models for the prediction of MAF concentrate on including a better description of water-scarce regions within the model. Nonetheless, streamflow prediction models would in general benefit from the improvement of quality and quantity of monitored precipitation and discharge data especially in arid regions. In addition, future research should focus on whether such a

simplified framework for describing catchment hydrology worldwide can be derived for finer temporal resolutions (daily or monthly).

Acknowledgments

This project has received funding from the Europeans Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 641459. The authors would like to thank the Global Runoff Data Center for providing the data of daily rivers discharge from worldwide gauging stations, and Rens van Beek from Utrecht University for providing global monthly gridded discharge values from the model PCR-GLOBWB. The data used in this work are listed in the references and provided in the supporting information.

References

- Alcamo, J. et al., 2003. Development and testing of the WaterGAP 2 global model of water use and availability. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 48(3): 317-337. DOI:10.1623/hysj.48.3.317.45290
- Anselin, L., 1988. Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity. *Geographical Analysis*, 20(1): 1-17. DOI:10.1111/j.1538-4632.1988.tb00159.x
- Anselin, L., Rey, S., 1991. Properties of Tests for Spatial Dependence in Linear Regression Models. *Geographical Analysis*, 23(2): 112-131. DOI:10.1111/j.1538-4632.1991.tb00228.x
- Barton, K., 2015. MuMIn: Multi-Model Inference. R package version 1.15.1.
- Beusen, A.H.W., Bouwman, A.F., Van Beek, L.P.H., Mogollón, J.M., Middelburg, J.J., 2016. Global riverine N and P transport to ocean increased during the 20th century despite increased retention along the aquatic continuum. *Biogeosciences*, 13(8): 2441-2451. DOI:10.5194/bg-13-2441-2016
- Bierkens, M.F.P., 2015. Global hydrology 2015: State, trends, and directions. *Water Resources Research*, 51(7): 4923-4947. DOI:10.1002/2015WR017173
- Bivand, R., Hauke, J., Kossowski, T., 2013. Computing the Jacobian in Gaussian Spatial Autoregressive Models: An Illustrated Comparison of Available Methods. *Geographical Analysis*, 45(2): 150-179. DOI:10.1111/gean.12008
- Bivand, R., Piras, G., 2015. Comparing Implementations of Estimation Methods for Spatial Econometrics. 2015, 63(18): 36. DOI:10.18637/jss.v063.i18
- Bring, J., 1994. How to Standardize Regression Coefficients. *The American Statistician*, 48(3): 209-213. DOI:10.2307/2684719
- Burgers, H.E., Schipper, A.M., Hendriks, A.J., 2013. Size relationships of water discharge in rivers: scaling of discharge with catchment area, main-stem length and precipitation. *Hydrological Processes*, 28(23): 5769-5775. DOI:10.1002/hyp.10087
- Chang, H., 2003. Basin Hydrologic Response to Changes in Climate and Land Use: the Conestoga River Basin, Pennsylvania. *Physical Geography*, 24(3): 222-247. DOI:10.2747/0272-3646.24.3.222
- Chen, X.Y., Chau, K.W., Busari, A.O., 2015. A comparative study of population-based optimization algorithms for downstream river flow forecasting by a hybrid neural network model. *Engineering Applications of Artificial Intelligence*, 46, Part A: 258-268. DOI:http://dx.doi.org/10.1016/j.engappai.2015.09.010
- Cook, R.D., Weisberg, S., 1982. *Residuals and Influence in Regression*. Chapman & Hall.
- Danandeh Mehr, A., Kahya, E., Olyaie, E., 2013. Streamflow prediction using linear genetic programming in comparison with a neuro-wavelet technique. *Journal of Hydrology*, 505: 240-249. DOI:http://dx.doi.org/10.1016/j.jhydrol.2013.10.003
- Development Core Team, 2005. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

- Döll, P., Kaspar, F., Lehner, B., 2003. A global hydrological model for deriving water availability indicators: model tuning and validation. *Journal of Hydrology*, 270(1–2): 105-134. DOI:[http://dx.doi.org/10.1016/S0022-1694\(02\)00283-4](http://dx.doi.org/10.1016/S0022-1694(02)00283-4)
- Fan, Y.R. et al., 2015. A stepwise-cluster forecasting approach for monthly streamflows based on climate teleconnections. *Stochastic Environmental Research and Risk Assessment*, 29(6): 1557-1569. DOI:10.1007/s00477-015-1048-y
- Farmer, W.H., Over, T.M., Vogel, R.M., 2015. Multiple regression and inverse moments improve the characterization of the spatial scaling behavior of daily streamflows in the Southeast United States. *Water Resources Research*, 51(3): 1775-1796. DOI:10.1002/2014WR015924
- Field, A., 2009. *Discovering statistics using SPSS*. Sage, London, UK.
- Gleeson, T., Wada, Y., Bierkens, M.F.P., van Beek, L.P.H., 2012. Water balance of global aquifers revealed by groundwater footprint. *Nature*, 488(7410): 197-200. DOI:<http://www.nature.com/nature/journal/v488/n7410/abs/nature11295.html#supplementary-information>
- Gosling, S.N., Arnell, N.W., 2011. Simulating current global river runoff with a global hydrological model: model revisions, validation, and sensitivity analysis. *Hydrological Processes*, 25(7): 1129-1145. DOI:10.1002/hyp.7727
- GRDC, 2011. *Watershed Boundaries of GRDC Stations / Global Runoff Data Centre*. Federal Institute of Hydrology (BfG), Koblenz, Germany.
- GRDC, 2015. *Long-Term Mean Monthly Discharges and Annual Characteristics of GRDC Stations / Global Runoff Data Centre*. Koblenz, Germany: Federal Institute of Hydrology (BfG).
- Gyawali, R., Griffis, V.W., Watkins, D.W., Fennessey, N.M., 2015. Regional regression models for hydro-climate change impact assessment. *Hydrological Processes*, 29(8): 1972-1985. DOI:10.1002/hyp.10312
- Hall, D., Cherry, S., Reeves, K., Lee, R., Carroll, G., Sommers, G., Verdin, K., 2004. *Water Energy Resources of the United States With Emphasis on Low Head/Low Power Resources*, U.S. Department of Energy Efficiency and Renewable Energy, Wind and Hydropower Technologies, Idaho Falls, Idaho. Retrieved from <http://www1.eere.energy.gov/windandhydro/pdfs/doewater-11111.pdf>.
- Hamon, W.R., 1963. Computation of Direct Runoff Amounts from Storm Rainfall. *Int Assoc Sci Hydrol Publ*, 63: 52–62
- Hanafiah, M.M., Xenopoulos, M.A., Pfister, S., Leuven, R.S.E.W., Huijbregts, M.A.J., 2011. Characterization factors for water consumption and greenhouse gas emissions based on freshwater fish species extinction. *Environ. Sci. Technol.*, 45(12): 5272-5278. DOI:10.1021/es1039634
- Hanasaki, N. et al., 2008. An integrated model for the assessment of global water resources – Part 1: Model description and input meteorological forcing. *Hydrol. Earth Syst. Sci.*, 12(4): 1007-1025. DOI:10.5194/hess-12-1007-2008

- Harris, I., Jones, P.D., Osborn, T.J., Lister, D.H., 2014. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *Int. J. Climatol.*, 34(3): 623-642. DOI:10.1002/joc.3711
- Heiberger, R., M., 2015. HH: Statistical Analysis and Data Display: Heiberger and Holland. R package version 3.1-21.
- Hendriks, A.J., Schipper, A.M., Caduff, M., Huijbregts, M.A.J., 2012. Size relationships of water inflow into lakes: Empirical regressions suggest geometric scaling. *Journal of Hydrology*, 414: 482-490. DOI:10.1016/j.jhydrol.2011.11.025
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.*, 25(15): 1965-1978. DOI:10.1002/joc.1276
- Hoekstra, A.Y., Chapagain, A.K., Aldaya, M.M., Mekonnen, M.M., 2011. *The water footprint assessment manual: Setting the global standard*, London, UK.
- Hortness, J.E., Berenbrock, C., 2001. Estimating monthly and annual streamflow statistics at ungaged sites in Idaho. *Water-Resources Investigations Report 01-4093*, pp.36. US Department of the Interior, US Geological Survey 2001-4093, Reston, VA.
- Janse, J.H. et al., 2015. GLOBIO-Aquatic, a global model of human impact on the biodiversity of inland aquatic ecosystems. *Environmental Science & Policy*, 48: 99-114. DOI:http://dx.doi.org/10.1016/j.envsci.2014.12.007
- Jefferies, D. et al., 2012. Water Footprint and Life Cycle Assessment as approaches to assess potential impacts of products on water consumption. Key learning points from pilot studies on tea and margarine. *Journal of Cleaner Production*, 33: 155-166. DOI:http://dx.doi.org/10.1016/j.jclepro.2012.04.015
- Kennard, M.J., Mackay, S.J., Pusey, B.J., Olden, J.D., Marsh, N., 2010. Quantifying uncertainty in estimation of hydrologic metrics for ecohydrological studies. *River Research and Applications*, 26(2): 137-156. DOI:10.1002/rra.1249
- Legates, D.R., McCabe, G.J., 1999. Evaluating the use of “goodness-of-fit” Measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, 35(1): 233-241. DOI:10.1029/1998WR900018
- Lehner, B., Verdin, K., Jarvis, A., 2008. New Global Hydrography Derived From Spaceborne Elevation Data. *Eos, Transactions American Geophysical Union*, 89(10): 93-94. DOI:10.1029/2008EO100001
- Lu, J., Sun, G., McNulty, S.G., Amatya, D.M., 2005. A COMPARISON OF SIX POTENTIAL EVAPOTRANSPIRATION METHODS FOR REGIONAL USE IN THE SOUTHEASTERN UNITED STATES1. *JAWRA Journal of the American Water Resources Association*, 41(3): 621-633. DOI:10.1111/j.1752-1688.2005.tb03759.x
- Oberdorff, T. et al., 1995. Global Scale Patterns of Fish Species Richness in Rivers. *Ecography*, 18(4): 345-352.
- Oberdorff, T. et al., 2011. Global and Regional Patterns in Riverine Fish Species Richness: A Review. *International Journal of Ecology*, 2011: 12. DOI:10.1155/2011/967631

- Okkan, U., Serbes, Z.A., 2012. Rainfall–runoff modeling using least squares support vector machines. *Environmetrics*, 23(6): 549-564. DOI:10.1002/env.2154
- Ord, K., 1975. Estimation Methods for Models of Spatial Interaction. *Journal of the American Statistical Association*, 70(349): 120-126. DOI:10.1080/01621459.1975.10480272
- Pfister, S., Koehler, A., Hellweg, S., 2009. Assessing the Environmental Impacts of Freshwater Consumption in LCA. *Environ. Sci. Technol.*, 43(11): 4098-4104. DOI:10.1021/es802423e
- Poff, N.L., Zimmerman, J.K.H., 2010. Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows. *Freshwater Biology*, 55(1): 194-205. DOI:10.1111/j.1365-2427.2009.02272.x
- Rodríguez-Iturbe, I., Rinaldo, A., 2001. *Fractal River Basins: Chance and Self-Organization*. Cambridge University Press, UK.
- Santini, M., di Paola, A., 2015. Changes in the world rivers' discharge projected from an updated high resolution dataset of current and future climate zones. *Journal of Hydrology*, 531, Part 3: 768-780. DOI:http://dx.doi.org/10.1016/j.jhydrol.2015.10.050
- Shiklomanov, A.I., Lammers, R.B., Vörösmarty, C.J., 2002. Widespread decline in hydrological monitoring threatens Pan-Arctic Research. *Eos, Transactions American Geophysical Union*, 83(2): 13-17. DOI:10.1029/2002EO000007
- Stuckey, M.H., 2006. Low-flow, base-flow, and mean-flow regression equations for Pennsylvania streams. U.S. Geological Survey Scientific Investigations Report 2006-5130, p. 84.
- Syvitski, J.P., Peckham, S.D., Hilberman, R., Mulder, T., 2003. Predicting the terrestrial flux of sediment to the global ocean: a planetary perspective. *Sedimentary Geology*, 162(1-2): 5-24. DOI:http://dx.doi.org/10.1016/S0037-0738(03)00232-X
- Syvitski, J.P., Vörösmarty, C.J., Kettner, A.J., Green, P., 2005. Impact of Humans on the Flux of Terrestrial Sediment to the Global Coastal Ocean. *Science*, 308(5720): 376-380. DOI:10.1126/science.1109454
- Tendall, D.M., Hellweg, S., Pfister, S., Huijbregts, M.A.J., Gaillard, G., 2014. Impacts of River Water Consumption on Aquatic Biodiversity in Life Cycle Assessment—A Proposed Method, and a Case Study for Europe. *Environ. Sci. Technol.*, 48(6): 3236-3244. DOI:10.1021/es4048686
- Thomas, D.M., Benson, M.A., 1970. Generalization of streamflow characteristics from drainage-basin characteristics, pp.62. US Geological Survey WaterSupply Paper 1975. Washington, DC.
- Thornthwaite, C.W., 1948. An approach toward a rational classification of climate. *Geographical Review*, 38: 55-94.
- Tran, L.T., O'Neill, R.V., Bruins, R.J.F., Smith, E.R., Harden, C., 2015. Linking land use/land cover with climatic and geomorphologic factors in regional mean annual streamflow models with geospatial regression approach. *Progress in Physical Geography*, 39(2): 258-274. DOI:10.1177/0309133314562441

- University of East Anglia Climatic Research Unit, Jones, P. D., Harris, I., 2013. CRU TS3.00: Climatic Research Unit (CRU) Time-Series (TS) Version 3.00 of High Resolution Gridded Data of Month-by-month Variation in Climate (Jan. 1901 - Dec.2006). NCAS British Atmospheric Data Centre.
- van Beek, L.P.H., Bierkens, M.F.P., 2008. The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification, Utrecht, The Netherlands. Retrieved from <http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf>.
- van Beek, L.P.H., Wada, Y., Bierkens, M.F.P., 2011. Global monthly water stress: 1. Water balance and water availability. *Water Resources Research*, 47(7). DOI:10.1029/2010WR009791
- Van Der Knijff, J.M., Younis, J., De Roo, A.P.J., 2010. LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science*, 24(2): 189-212. DOI:10.1080/13658810802549154
- Verdin, K.L., Worstell, B., 2008. A Fully Distributed Implementation of Mean Annual StreamFlow Regional Regression Equations1. *JAWRA Journal of the American Water Resources Association*, 44(6): 1537-1547. DOI:10.1111/j.1752-1688.2008.00258.x
- Vogel, R.M., Sankarasubramanian, A., 2000. Spatial scaling properties of annual streamflow in the United States. *Hydrological Sciences Journal*, 45(3): 465-476. DOI:10.1080/02626660009492342
- Vogel, R.M., Wilson, I., Daly, C., 1999. Regional Regression Models of Annual Streamflow for the United States. *Journal of Irrigation and Drainage Engineering*, 125(3): 148-157. DOI:doi:10.1061/(ASCE)0733-9437(1999)125:3(148)
- Vorosmarty, C.J., Fekete, B.M., Tucker, B.A., 1998. Global River Discharge, 1807-1991, V[ersion]. 1.1 (RivDIS). Data set. Available on-line [<http://www.daac.ornl.gov>] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A. DOI:10.3334/ORNLDAAAC/199
- Wada, Y. et al., 2011. Global monthly water stress: 2. Water demand and severity of water stress. *Water Resources Research*, 47(7): n/a-n/a. DOI:10.1029/2010WR009792
- Wanders, N., Wada, Y., 2015. Human and climate impacts on the 21st century hydrological drought. *Journal of Hydrology*, 526: 208-220. DOI:<http://dx.doi.org/10.1016/j.jhydrol.2014.10.047>
- Wang, W.-c., Chau, K.-w., Xu, D.-m., Chen, X.-Y., 2015. Improving Forecasting Accuracy of Annual Runoff Time Series Using ARIMA Based on EEMD Decomposition. *Water Resources Management*, 29(8): 2655-2675. DOI:10.1007/s11269-015-0962-6
- Widén-Nilsson, E., Halldin, S., Xu, C.-y., 2007. Global water-balance modelling with WASMOD-M: Parameter estimation and regionalisation. *Journal of Hydrology*, 340(1-2): 105-118. DOI:<http://dx.doi.org/10.1016/j.jhydrol.2007.04.002>
- Wisser, D., Fekete, B.M., Vörösmarty, C.J., Schumann, A.H., 2010. Reconstructing 20th century global hydrography: a contribution to the Global Terrestrial Network- Hydrology (GTN-H). *Hydrol. Earth Syst. Sci.*, 14(1): 1-24. DOI:10.5194/hess-14-1-2010
- World Meteorological Organization, 1992. *International Meteorological Vocabulary*, Secretariat of the World Meteorological Organization, Geneva.

- Wu, C.L., Chau, K.W., Li, Y.S., 2009. Methods to improve neural network performance in daily flows prediction. *Journal of Hydrology*, 372(1–4): 80-93.
DOI:<http://dx.doi.org/10.1016/j.jhydrol.2009.03.038>
- Xenopoulos, M.A., Lodge, D.M., 2006. Going with the Flow: Using Species-Discharge Relationships to Forecast Losses in Fish Biodiversity. *Ecology*, 87(8): 1907-1914. DOI:10.1890/0012-9658(2006)87[1907:GWTFUS]2.0.CO;2
- Xenopoulos, M.A. et al., 2005. Scenarios of freshwater fish extinctions from climate change and water withdrawal. *Global Change Biology*, 11(10): 1557-1564. DOI:10.1111/j.1365-2486.2005.001008.x
- Yu, K.-x., Gottschalk, L., Xiong, L., Li, Z., Li, P., 2015. Estimation of the annual runoff distribution from moments of climatic variables. *Journal of Hydrology*, 531, Part 3: 1081-1094.
DOI:<http://dx.doi.org/10.1016/j.jhydrol.2015.11.012>
- Zuur, A., Ieno, E.N., Walker, N., 2009. *Mixed Effects Models and Extensions in Ecology with R*. Springer, New York.