

# The role of climatic and terrain attributes in estimating baseflow recession in tropical catchments

J. L. Peña-Arancibia<sup>1,2,3</sup>, A. I. J. M. van Dijk<sup>3</sup>, M. Mulligan<sup>1</sup>, and L. A. Bruijnzeel<sup>2</sup>

<sup>1</sup>CSIRO Land and Water, GPO 1666, Black Mountain ACT, Australia

<sup>2</sup>Environmental Monitoring and Modelling Research Group, Department of Geography, King's College London, Strand, London WC2R 2LS, UK

<sup>3</sup>Faculty of Earth and Life Sciences, VU University Amsterdam, De Boelelaan 1085-1087, 1081 HV, Amsterdam, The Netherlands

Received: 13 June 2010 – Published in Hydrol. Earth Syst. Sci. Discuss.: 1 July 2010

Revised: 15 October 2010 – Accepted: 27 October 2010 – Published: 4 November 2010

**Abstract.** The understanding of low flows in rivers is paramount more than ever as demand for water increases on a global scale. At the same time, limited streamflow data to investigate this phenomenon, particularly in the tropics, makes the provision of accurate estimations in ungauged areas an ongoing research need. This paper analysed the potential of climatic and terrain attributes of 167 tropical and sub-tropical unregulated catchments to predict baseflow recession rates. Daily streamflow data ( $\text{m}^3 \text{s}^{-1}$ ) from the Global River Discharge Center (GRDC) and a linear reservoir model were used to obtain baseflow recession coefficients ( $k_{\text{bf}}$ ) for these catchments. Climatic attributes included annual and seasonal indicators of rainfall and potential evapotranspiration. Terrain attributes included indicators of catchment shape, morphology, land cover, soils and geology. Stepwise regression was used to identify the best predictors for baseflow recession coefficients. Mean annual rainfall (MAR) and aridity index (AI) were found to explain 49% of the spatial variation of  $k_{\text{bf}}$ . The rest of climatic indices and the terrain indices average catchment slope (SLO) and tree cover were also good predictors, but co-correlated with MAR. Catchment elongation (CE), a measure of catchment shape, was also found to be statistically significant, although weakly correlated. An analysis of clusters of catchments of smaller size, showed that in these areas, presumably with some similarity of soils and geology due to proximity, residuals of the regression could be explained by SLO and CE. The approach used provides a potential alternative for  $k_{\text{bf}}$  parameterisation in ungauged catchments.

## 1 Introduction

The gradual depletion of water stored in a catchment during dry weather constitutes the drainage or baseflow recession (Tallaksen, 1995). The understanding of quantities and temporal patterns of baseflow are central to water resources management, particularly in catchments with marked streamflow seasonality (Vogel and Kroll, 1992; Bruijnzeel, 2004; Brandes et al., 2005).

In recent years, several assessments of global water resources have been conducted using hydrological models and land surface models (LSMs); mainly in response to increase in water demand and potential impacts of climatic and land use change (Vörösmarty et al., 2000; Oki and Kanae, 2006). Linear conceptual storage-discharge models have been used to simulate baseflow recession in many of these models. In many cases, the linear reservoir application in global hydrological models used fixed parameter values, e.g. the routing HD model (Hagemann and Dümenil, 1998), macro-PDM (Arnell, 1999, 2003) and WGHM (Döll et al., 2003). Values obtained from drainage theory have been used in PCR-GLOBWB (Van Beek and Bierkens, 2008) whereas calibrated values were used in the global application of WASMOD-M (Widen-Nilsson et al., 2007) and in an application of the Catchment Land Surface Model (CLSM) to the Somme River Basin (Gascoïn et al., 2009). The use of drainage theory is questionable at large scales and hindered by the uncertain quality of data needed to estimate various parameters. For example, the theoretical approach of Brutsaert and Nieber (1977), one of the few analytical ways to obtain aquifer parameters from hillslope to catchment scales, was used by Zecharias and Brutsaert (1988) to advance a proportionality relationship between the recession coefficient and aquifer characteristics:



Correspondence to:

J. L. Peña-Arancibia

(jorge.pena\_arancibia@kcl.ac.uk)

$$k_{\text{bf}} \propto \frac{K D \alpha}{Y L} \quad (1)$$

where  $K$  is hydraulic conductivity,  $D$  is aquifer thickness,  $\alpha$  is slope,  $Y$  is storativity and  $L$  a characteristic flow path length. Many of the aquifer parameters are not readily available in the tropics (and elsewhere); in particular data on aquifer hydraulic conductivity and thickness are sparse and scattered and cannot be considered representative of large areas. Also, the density of streamflow station data – used on a routine basis to calibrate conceptual models – is not spatially uniform, particularly in remote forested areas. Moreover, calibration approaches are not practical for global applications because of the large number of locations for which separate calibrations would be needed (Nijssen et al., 2001). Nijssen et al. (2001) modelled the seasonal discharge of 26 large basins in the world (including the Amazon, Congo and Mekong) using data from the Global River Discharge Center (GRDC) based in Koblenz-Germany. From these data, they estimated baseflow recession coefficients ( $k_{\text{bf}}$ ) to parameterise the conceptual quasi-linear baseflow reservoir component of the VIC model (Liang et al., 1994). Baseflow recession coefficients were determined for 347 stations that had good quality data using a linear regression on the log-transformed discharges and then interpolated to the nearest areas.

On the other hand, several studies have correlated terrain attributes – including catchment morphology and soil type – to estimate  $k_{\text{bf}}$  in different climatic and physiographic regions or for geological formations across the world (e.g. Post and Jakeman, 1996; Yu et al., 2002; Brandes et al., 2005). Most studies have focused on catchments with areas  $<200 \text{ km}^2$  and located in common physiographic regions of similar climate. Van Dijk (2010) included climatic attributes in addition to terrain attributes to analyse the relationship with  $k_{\text{bf}}$  for 183 mainly temperate Australian catchments with a large geographical spread and encompassing different climates. The results showed that baseflow recession from a linear reservoir was best explained by climatic attributes, with catchment aridity index (AI, the ratio of rainfall to potential evapotranspiration) explaining 27% of the variation in derived recession coefficients. No correlations were found with catchment morphology or geology; however, spatial coherence of the residual unexplained variation showed that another 53% of the variation was spatially correlated over distances of 100–150 km. This was probably associated with terrain factors not captured by the available data and the large geographical spread of individual catchments (Van Dijk, 2010).

Motivated by the latter results, the objective of the present study is to identify the dominant climatic and terrain attributes that control the variance of  $k_{\text{bf}}$  in tropical and subtropical catchments. There is a dearth of studies that have investigated these relationships in the tropics and most of them were limited to small geographic regions (e.g. Yu et al., 2002; Mwakalila et al., 2002). In this study, catch-

ment baseflow recession coefficients were determined from GRDC daily streamflow data ( $\text{m}^3 \text{ s}^{-1}$ ) using a linear reservoir model. Although research indicates that low flows during dry periods can be adequately approximated by linear reservoirs (Zecharias and Brutsaert, 1988; Vogel and Kroll, 1992; Chapman, 1999; Fenicia et al., 2006; Van Dijk, 2010), this may not be the case in every catchment, in which some hydrological processes may lead to a non-linear behaviour. The choice of a linear model was made by taking into account that the aforementioned global hydrological models use a linear reservoir to estimate baseflow. Building equations based on the climatic and terrain indices that best explain  $k_{\text{bf}}$  in gauged catchments to estimate this parameter in ungauged catchments is a subsidiary objective of this study. These equations can be potentially used to parameterise  $k_{\text{bf}}$  in global hydrological models.

## 2 Theory

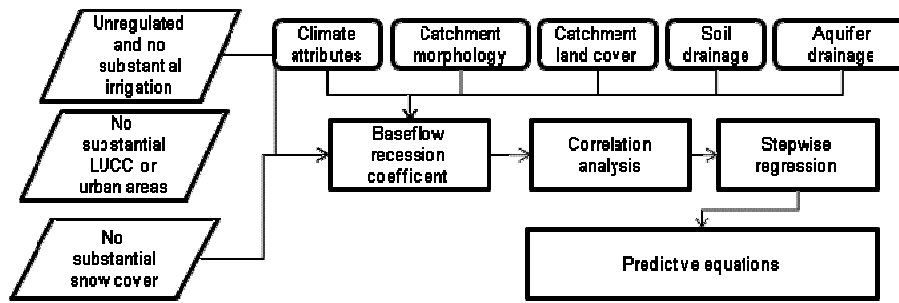
Several reviews on baseflows and recession analysis can be found in the existing literature (Tallaksen, 1995; Wittenberg, 1999; Smakhtin, 2001). In this paper, only a summary of the rationale and the main equations involved in baseflow recession analysis are presented. The theoretical framework of this study follows the one presented in Van Dijk (2010).

A linear reservoir model requires a recession coefficient ( $k_{\text{bf}}$ ) to separate daily streamflow data into baseflow and quickflow and is expressed as:

$$Q_{\text{bf}} = -k_{\text{bf}} S \quad (2)$$

where  $Q_{\text{bf}}$  (in  $\text{mm day}^{-1}$ ) is the flow rate during the baseflow recession period,  $S$  (mm) is reservoir storage. The constant  $k_{\text{bf}}$  is expressed in  $\text{day}^{-1}$ .

Streamflow data representative of baseflow needs not to be affected by stormflow. It is assumed that stormflow affects streamflow for a period of  $T_{\text{QF}}$  days after the event peak flow (Van Dijk, 2010). Van Dijk (2010) found that for catchments in Australia, the number of data pairs decreased exponentially with increasing  $T_{\text{QF}}$  period. Vogel and Kroll (1992) considered baseflow recession to start when the 3-day streamflow moving average begins to decrease, and the recession to end when the 3-day moving average start to increase. A period of 5 days ( $T_{\text{QF}5}$ ) was considered a useful compromise between representative low flow conditions and data availability. Increasing the window size to more days would have resulted in many catchments being dropped from the analysis. In addition, by using a large  $T_{\text{QF}}$ , results may be biased because only a small number of long baseflow recessions would be used to determine  $k_{\text{bf}}$  (Van Dijk, 2010). On the other hand, 5 days was assumed a sufficient time to avoid influence of storm flow on the hydrograph recession. This criterium was considered to construct  $Q$  and  $Q^*$  ( $Q$  of the previous day) data pairs representative of baseflow conditions for each gauging station. All days with zero streamflow



**Fig. 1.** Methodology flowchart. Unregulated catchments without substantial land use and land cover change (LUCC) and with snow, irrigation or urban extent <5% are selected for the analysis. Streamflow had to have at least 5 years of data, 30 runoff events and 30  $Q-Q^*$  data pairs.

and or missing data were also excluded. By using a representative number of  $Q-Q^*$  data pairs it was possible to estimate the recession coefficient  $k_{bf}$ . The procedure to obtain  $k_{bf}$  and the influence of using different windows sizes on  $k_{bf}$  will be described further on.

The relationship between initial storage ( $S_0$  in mm) and  $S$  after  $t$  days is defined by:

$$S = S_0 e^{-k_{bf}t} \tag{3}$$

By combining Eqs. (2) and (3) for time step  $t = 1$ , provided that both  $Q$  and  $Q^*$  represent baseflow and  $Q_0 = Q^*$ , baseflow recession can be represented by the exponential decay function:

$$Q = Q_0 e^{-k_{bf}t} \tag{4}$$

### 3 Methodology

Time-series of catchment streamflow hydrographs for 1175 tropical and sub-tropical stations with >5 years of data were obtained from the Global Runoff Database (GRDB) from the GRDC. Additional data for 272 stations in tropical and subtropical Queensland (Australia) were obtained from the Department of Environment and Resource Management Queensland (DERM). Stations with more than 30 runoff events (defined as the number of times that daily average streamflow is exceeded) and more than 30 data pairs  $Q-Q^*$  characteristic of low flow conditions were selected for the analysis. Data in  $m^3 s^{-1}$  were subsequently converted to  $mm d^{-1}$ . Furthermore, stations were geo-referenced using the Hydrosheds river network data (Lehner et al., 2008). Catchments not affected by regulation were identified using a pan-tropical dam dataset (Saenz and Mulligan, 2010) representing the only available dataset (to the authors' knowledge) providing the actual catchment areas of reservoirs on a pan-tropical scale. In addition, GLOBCOVER land use data (Arino et al., 2008) and the MODIS 500-m map of global urban extent (Schneider et al., 2009) were used to check for catchments that may have snowmelt influence and irrigation

areas, and large urban centres respectively. Only catchments with less than 5% snow cover, irrigation or urban extent were used in the analysis. Areas not affected by extensive deforestation during the period of analysis, which would likely have an impact on the recession coefficient trend, were determined from the map of areas of rapid land cover change provided by Lepers et al. (2005). Catchments complying with these criteria were considered unregulated for the purpose of this study.

Relevant catchment climatic, physiographic and geological attributes previously used in baseflow recession analysis were derived using terrain analysis and available climatic, geological or soils data. A preliminary analysis of frequency distributions for  $k_{bf}$  and various climatic indices and catchment attributes was conducted to assess applicable correlations methods. Furthermore, a non-parametric correlation matrix was used to determine the degree of correlation between recession constants and catchment attributes. Finally, predictive relationships were obtained using stepwise regression. Figure 1 presents a summary flowchart of the procedure described above.

#### 3.1 Climatic and terrain attributes of pan-tropical catchments

Several climatic and terrain attributes with a demonstrated correlation with baseflow parameters (e.g. Post and Jakeman, 1996; Brandes et al., 2005; Van Dijk, 2010) were derived for each catchment. A summary of parameters, their original resolution and source are summarised in Table 1. Climatic attributes included annual and seasonal descriptors of rainfall and potential evapotranspiration and were defined as follows:

- Mean annual rainfall (MAR) expressed in  $mm y^{-1}$  obtained from the WORLDCLIM dataset (Hijmans et al., 2005).
- Potential evapotranspiration (PET) in  $mm y^{-1}$  estimated using the Hargreaves et al. (1985) model formulation and parameterised as described in Trabucco et al. (2008).

**Table 1.** Summary of climatic and terrain attributes used in the present study.

Parameter	Resolution		Source
	Temporal	Spatial	
MAR (mm y <sup>-1</sup> )	Monthly average climatology 1950–2000	1 × 1 km grid	WORLDCLIM (Hijmans et al., 2005)
PET (mm y <sup>-1</sup> )	Monthly average climatology 1950–2000	1 × 1 km grid	Trabucco et al. (2008; available at: <a href="http://www.csi.cgiar.org">http://www.csi.cgiar.org</a> )
AI	NA	1 × 1 km grid	Calculated from MAE and PET
TMI	NA	1 × 1 km grid	Calculated from monthly rainfall and monthly PET
SI	NA	1 × 1 km grid	Calculated from monthly rainfall
CE (m <sup>2</sup> m <sup>-1</sup> )	NA	NA	Hydrosheds 1 km DEM (Lehner et al., 2010)
SLO (%)	NA	90 × 90 m grid	Hydrosheds 90 m DEM (Lehner et al., 2010)
DD (km km <sup>-2</sup> )	NA	90 × 90 m grid	Hydrosheds 90 m river network available at <a href="http://hydrosheds.cr.usgs.gov/hydro.php">http://hydrosheds.cr.usgs.gov/hydro.php</a>
TC (%)	NA	1 × 1 km grid	AVHRR Tree Cover Continuous fields (DeFries et al., 2000; available at: <a href="http://glcf.umd.edu/data/treecover/data.shtml">http://glcf.umd.edu/data/treecover/data.shtml</a> )
SDI	NA	9 × 9 km grid	ISRIC-WISE derived soil properties (Batjes, 2006)
DPI	NA	NA	WHYMAP (2010)

- Aridity index (AI = MAR/PET)
- Thornthwaite Moisture Index (TMI, Thornthwaite, 1948). An overall measure of precipitation effectiveness on a monthly basis. It is estimated using monthly rainfall and PET totals from the above mentioned datasets as follows:

$$TMI = \frac{\sum_{m=1}^{12} (100s_m - 60d_m)}{PET} \quad (5)$$

where  $s$  is the monthly water surplus and  $d$  is the monthly water deficit (mm mo<sup>-1</sup>).

- Seasonality index. The seasonality index (SI, Walsh and Lawler, 1981) is defined as the sum of the absolute deviation of mean monthly rainfall ( $\bar{X}_m$ ) from the overall monthly mean divided by the mean annual rainfall (MAR):

$$SI = \frac{1}{MAR} \sum_{m=1}^{12} \left| \bar{X}_m - \frac{MAR}{12} \right| \quad (6)$$

The SI varies from zero (when all months have the same rainfall) to 1.83 (when all rainfall occurs in a single month): values <0.19 indicate a very equable rainfall regime, whereas values between 0.20 and 0.99 indicate a seasonal rainfall regime and values >1 a short wet season.

Terrain attributes included indicators of catchment shape, morphology, land cover, soils and geology.

- Catchment shape, defined by catchment elongation (CE) in km<sup>2</sup> surface area per km of catchment length, or by the ratio of a circle with the same area as the catchment to the catchment's length (Post and Jakeman, 1996).
- Mean catchment rainfall weighted slope (SLO) (%). To account for spatial variability in rainfall, each catchment slope pixel is scaled using normalised mean catchment rainfall data. By scaling SLO in this way, areas that may produce more runoff due to spatial differences in rainfall have more weight in the final mean catchment slope computation.

- Catchment drainage density (DD) in km per km<sup>2</sup>, defined by the total length of streams per square kilometre of catchment area.
- Catchment tree cover percentage (TC), from AVHRR Tree Cover Continuous fields (DeFries et al., 2000).
- Soil unit weighted infiltrability class (SDI), obtained from the ISRIC-WISE Soil Derived Properties database (Batjes, 2006). The final values were obtained using an area weighted average of dominant soils comprising each soil unit.

Categorical information on drainage potential index (DPI) according to geology and climate was obtained from WHYMAP (2010). WHYMAP included eleven classes: Class I to V correspond to the presence of a major groundwater basin with very high, high, medium, low and very low drainage rates respectively. Classes VI to IX correspond to complex hydrogeological structure and very high, high, medium and low drainage rates respectively. Class X and XI correspond to local and shallow aquifers with high and low drainage rates respectively.

### 3.2 Estimation of recession coefficient $k_{bf}$

Methods to obtain  $k_{bf}$  from baseflow data pairs include linear regression through the origin, linear regression on log-transformed baseflow data pairs and optimisation techniques (e.g. Wittenberg, 1999; Tularam and Ilahee, 2008). For this study, recession coefficients were estimated by fitting Eq. (4) to the baseflow data pairs using the mean relative error ( $\epsilon_{MRE}$ ) as the objective function and a multi-start downhill simplex search method (Van Dijk, 2010):

$$\epsilon_{MRE} = \frac{1}{n} \sum \left| \frac{Q}{Q_{est}} - 1 \right| \quad (7)$$

where  $Q_{est}$  is  $Q$  predicted from Eq. (4). By using the relative agreement between estimated and observed streamflows, this formulation does not use absolute values which could bias the results and gives equal weighting to all data pairs. However weighting influence by very low or very large values when using different objective functions cannot be entirely avoided. In addition, by estimating a mean baseflow recession constant from many observed recession segments, the problem of time variability (per event or seasonal) in baseflow recession is partially overcome (Tallaksen, 1995).

### 3.3 Statistical analysis

A correlation matrix was used to determine the correlation between various catchment attributes and the recession coefficients. The attributes with the best individual explanatory values were combined into a stepwise multiple regression equation. Exponential, logarithmic and power functions were computed to link potential predictors to  $k_{bf}$ , and the

best regression was selected to subsequently predict  $k_{bf}$ . Attributes that co-correlated were not considered in the subsequent stepwise regression. After selecting the best equation, the same types of regression were computed for both the absolute and relative residual variance and the remaining potential predictors, until no further variation was explained by adding these.

## 4 Results

### 4.1 Assembling a pan-tropical dataset for baseflow modelling

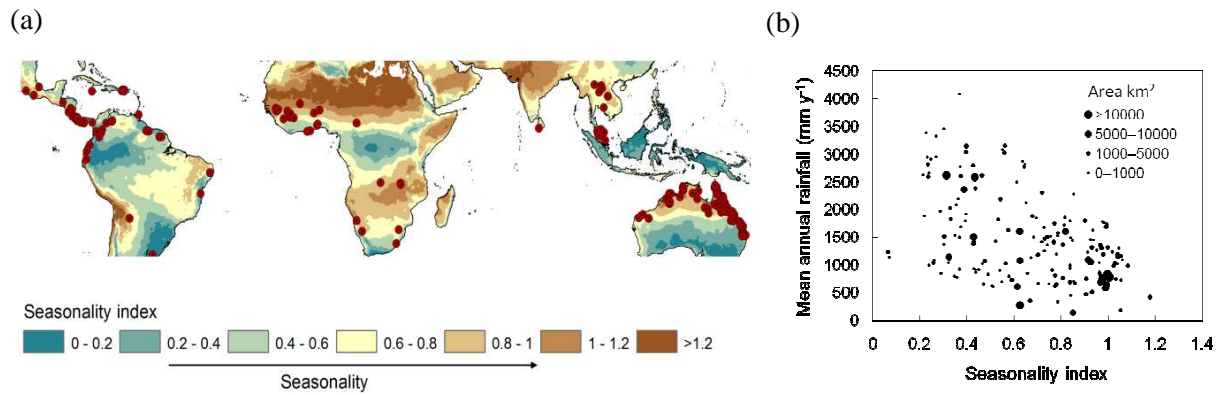
Catchment boundaries were obtained from the Hydrosheds 1 km river network (Lehner et al., 2010). Only catchments with a relative error of less than 10% between the GRDC reported surface areas and the river network derived areas were considered in the analysis. After controlling for regulation, snow and lake influence, urban and irrigation areas, and land use change; the analysis resulted in a database comprising 167 catchments worldwide (Fig. 2a). Of the 167 catchments, 50% had a catchment area <1000 km<sup>2</sup> and 90% <6000 km<sup>2</sup>. The median was 850 km<sup>2</sup>.

The catchment assemblage encompassed many tropical climates (Fig. 2b). A large number of stream gauging stations were located in Australia. No stations complying with the aforementioned requirements were found in the Amazon or Congo Basins. Most stations in these basins had monthly records or short daily records, which excluded them from the present analysis of daily flows.

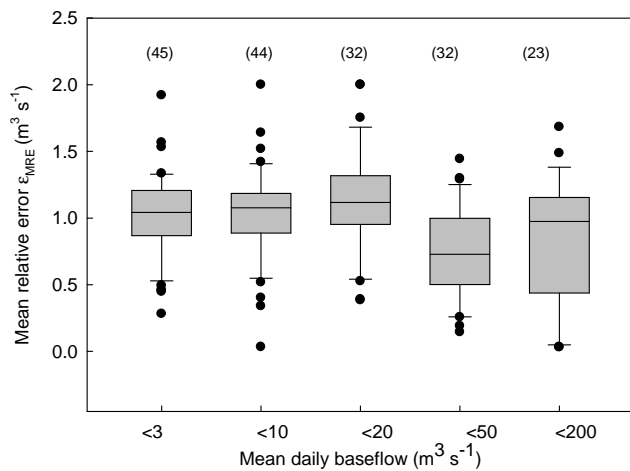
### 4.2 Estimation of recession coefficient $k_{bf}$

The overall mean  $k_{bf}$  for the 167 analysed catchments was of  $0.08 \pm 0.053$  (std. dev.) day<sup>-1</sup>. The distribution was positively skewed. Higher values were found in arid catchments and lower values in wetter catchments. In addition, lower values were generally found in catchment closer to the coastline. The mean relative error ( $\epsilon_{MRE}$ ) was  $0.97 \pm 0.38$  m<sup>3</sup> s<sup>-1</sup>. The distribution of  $\epsilon_{MRE}$  for different mean daily baseflow ranges is shown in Fig. 3.

In Australia, the lowest values of  $k_{bf}$  (0.02–0.08 day<sup>-1</sup>) were generally found in catchments that lie closer to the east and north coastlines. Catchments located in the more arid interior had values of 0.11–0.18 day<sup>-1</sup>. In Southeast Asia, the fully humid Malayan Peninsula had values of 0.02–0.06 day<sup>-1</sup>. Continental Southeast Asia showed values of 0.04–0.07 day<sup>-1</sup>. The highest values in Africa were found in Namibia (0.20 day<sup>-1</sup>) and in the catchments located in the northernmost of the Sahel (0.17 day<sup>-1</sup>). Catchments located closer to the coastline in West Africa and Central Africa (Congo and Zambia) generally showed values of  $\sim 0.035$  day<sup>-1</sup>, as did temperate catchments in South Africa. Catchments located in the Andes had values of 0.03–0.08 day<sup>-1</sup>. Catchments in Panama, Costa Rica Nicaragua



**Fig. 2.** Distribution of catchments in the dataset: (a) Geographic distribution (b) In terms of climate using the seasonality index (SI; Walsh and Lowler, 1981). Symbol sizes in (b) indicate catchment areas in km<sup>2</sup>.



**Fig. 3.** Distribution of the mean relative error for different mean daily baseflow ranges. The number in brackets is the sample size per range.

and Honduras had mostly values around 0.03–0.09 day<sup>-1</sup> whereas catchments in Puerto Rico had values of 0.05 day<sup>-1</sup>. In tropical Mexico, catchments close to the coastline had values of 0.03–0.10 days.

The stability of recession coefficients was assessed by varying the window size  $T_{QF}$  from 0–20 days. Results for six catchments with different climate regimes and geographical areas are illustrated in Fig. 4a–f. Increasing  $T_{QF}$  results in a reduction of  $k_{bf}$ , with the fastest decrease occurring in the first days 0–3 days. The rate of reduction diminishes after 5–10 days in most cases. Complex patterns occur when window size increased beyond 10 days. Similar variations of  $k_{bf}$  and  $Q-Q^*$  pairs were also observed for temperate Australian catchments in the study of Van Dijk (2010).

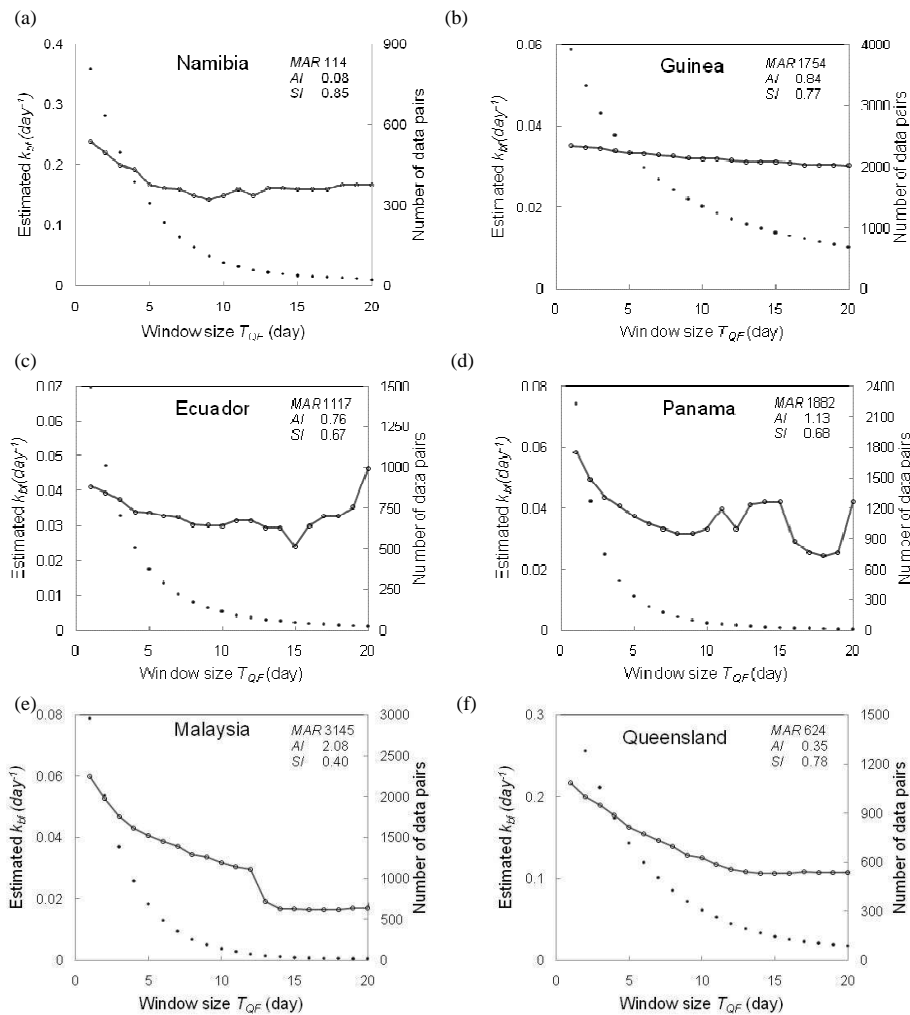
### 4.3 Statistical analyses

Visual inspection of scatter plots (Fig. 5) already suggested catchment recession coefficients to be correlated to various climatic attributes. Of the respective terrain attributes, only slope and tree cover appeared to show some correlation (Fig. 5e and f). The rest of the catchment attributes did not reveal a clear pattern (not shown). In addition, different aquifer drainage potential classes did not seem to have any influence on  $k_{bf}$  either (Fig. 5).

Recession coefficient data showed a positively skewed distribution and thus a non-parametric Spearman rho test was used in the correlation analysis. The correlation matrix is presented in Table 2. Significant strong correlations were found between  $k_{bf}$  and most climatic attributes; slope and tree cover. As expected, cross-correlations occurred between all climatic attributes. In addition, cross-correlations between slope, tree cover and climatic attributes were also observed. The best correlations for  $k_{bf}$  were with MAR and the Thornthwaite Moisture Index TMI (non-parametric  $r^* = -0.65$ ). AI also showed good correlation with  $k_{bf}$  ( $r^* = -0.64$ ). Regression equations were computed for  $k_{bf}$  vs. MAR and AI, results are shown in Fig. 6 (no power or exponential regression were possible for negative values of TMI).

A two-parameter exponential relationship of MAR and AI explained 49% of the variance in  $k_{bf}$ . Only marginal improvement was achieved with the stepwise regression when including the weakly correlated catchment elongation (CE,  $r^* = 0.138$ ). The other catchment terrain attributes with explanatory value were cross-correlated to climatic attributes and therefore not used in the multivariate analysis. The equations and summary statistics of all regressions are shown in Table 3.

A subset of catchments that were smaller and geographically close to each other, contiguous in some cases, were analysed to see whether the correlation of catchment terrain properties with  $k_{bf}$  was confounded by the large geographical



**Fig. 4.** Variation of estimated recession coefficient  $k_{bf}$  (closed lines) and number of  $Q-Q^*$  data pairs (dots) with an increase in  $T_{QF}$  from 0 to 20 days. Values for mean annual rainfall (MAR in  $\text{mm y}^{-1}$ ), aridity index (AI) and seasonality index (SI) is show for each catchment gauge data.

area and the different climates covered by the overall dataset (cf. Fig. 2a). Relatively smaller groups of catchments ( $<300 \text{ km}^2$ ) were selected in north, central and south Queensland and in Puerto Rico. Only clusters of larger catchments ( $500\text{--}3000 \text{ km}^2$ ) were left for analysis and so they were selected in the absence of data more suited to the purpose, in any case only two catchments were larger than  $1500 \text{ km}^2$ . These were located in Panama, Senegal and Malaysia. Relative residuals of the original regression of  $k_{bf}$  and MAR were analysed using scatter plots and non-parametric correlation. Only slope and catchment elongation showed significant correlations ( $r^* = 350$  and  $-250$  respectively). Although correlations were weak, scatter plots of such properties versus relative residuals showed some degree of spatial organisation by location (Fig. 7).

## 5 Discussion

### 5.1 Pan tropical catchment dataset

In the present study, great care was taken in producing a good quality daily streamflow dataset of unregulated flows (using the georeferenced dam dataset of Saenz and Mulligan) for tropical landscapes. A good range of climatic landscapes and rainfall regimes has been covered, but data from hydrologically important areas such as the Amazon and Congo basins are not yet represented in the analysis. Needless to say, their inclusion is highly desirable.

### 5.2 Characteristics of recession coefficients

In general, higher (faster) recession coefficients were observed for drier catchments. In the most arid catchments (e.g. Namibia, arid parts of Australia) streamflow is typically

**Table 2.** Spearman rank correlation matrix of recession coefficients and catchment attributes.  $k_{bf}$  correlations with climatic and catchment attributes are shown in bold.

	$k_{bf}$	MAR	PET	AI	TMI	SI	CE	SLO	DD	TC	SD	DPI
$k_{bf}$												
MAR	<b>-0.650**</b>											
PET	<b>0.291**</b>	-0.453**										
AI	<b>-0.639**</b>	0.979**	-0.608**									
TMI	<b>-0.649**</b>	0.987**	-0.564**	0.996**								
SI	<b>0.170*</b>	-0.436**	0.649**	-0.534**	-0.469**							
CE	<b>0.138*</b>	0.119	-0.068	0.105	0.108	0.008						
SLO	<b>-0.380**</b>	0.528**	-0.693**	0.613**	0.587**	-0.499**	0.100					
DD	<b>0.016</b>	0.064	-0.210**	0.088	0.086	-0.005	0.027	0.171*				
TC	<b>-0.425**</b>	0.578**	-0.577**	0.636**	0.618**	-0.436**	0.161*	0.592**	0.353**			
SD	<b>0.007</b>	0.003	0.294**	-0.056	-0.019	0.390**	0.064	-0.304**	0.191**	-0.02		
DPI	<b>0.003</b>	0.025	-0.005	0.012	-0.084	0.019	0.048	-0.017	0.287**	-0.003	-0.01	

\* Correlation is significant at the 0.05 level.

\*\* Correlation is significant at the 0.01 level.

**Table 3.** Summary of results, including  $r^2$  and standard error of estimate (SEE) for the exponential, logarithmic and power regressions linking  $k_{bf}$  with MAR and AI ( $n = 167$ ).

Equation	$r^2$	SEE
$k_{bf} = 0.0356 + 0.2273 \times e^{-0.0014MAR}$	0.4850	0.0382
$k_{bf} = 0.5247 - 0.0619 \times \ln(MAR)$	0.4447	0.0396
$k_{bf} = 10.4370 \times MAR^{-0.7050}$	0.3850	0.0510
$k_{bf} = 0.0394 + 0.2087 \times e^{-2.2282AI}$	0.4865	0.0382
$k_{bf} = 0.0692 - 0.0552 \times \ln(AI)$	0.4414	0.0397
$k_{bf} = 0.0580 \times AI^{-0.6210}$	0.3720	0.0515

ephemeral and consequently mainly event driven. The presence of fast-draining perched aquifers may also explain higher  $k_{bf}$ . By contrast, lower recession coefficients (slower drainage) were found for most of the humid tropics. Although there were no good quality data to account for the effects of soil depth and aquifer porosity, deep soils and permeable regoliths are widely present in tropical landscapes; and are likely to represent an important source of baseflow (Chappell et al., 2007). A recent three-year study in a small catchment underlain by very deep soils in the central Amazon Basin by Tomasella et al. (2008) showed an important contribution to the groundwater system by the extended unsaturated zone. Both unsaturated and groundwater flow showed a delayed response to rainfall and most of the seasonal variability in streamflow tended to be dampened by either one or the other.

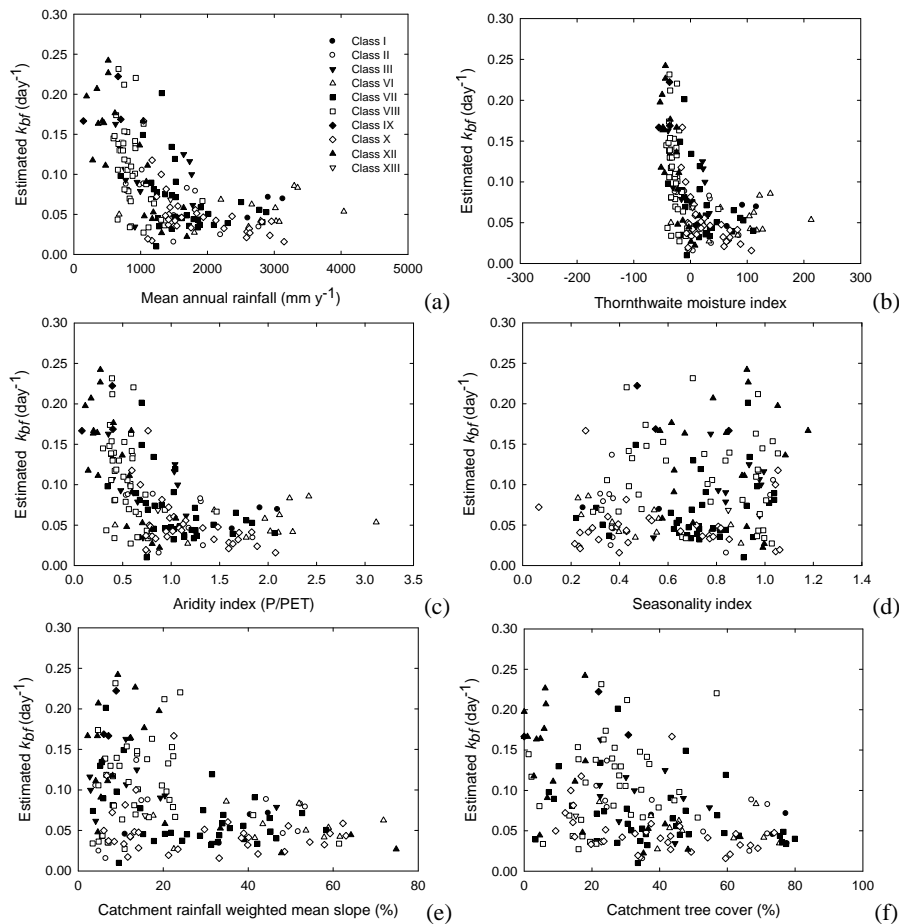
### 5.3 Predictors of recession coefficients

Climatic attributes proved to be the best predictors of  $k_{bf}$ , with MAR and AI together explaining 49% of the vari-

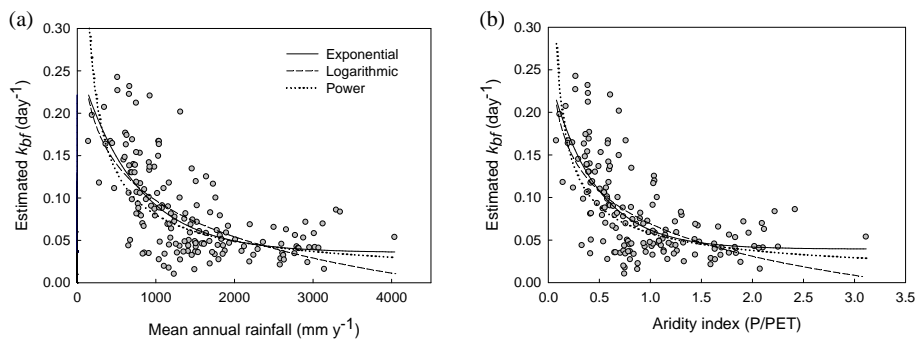
ance. The exponential and logarithmic regression equations for AI and MAR had very similar goodness-of-fit statistics but due to the nature of the fitted equations estimation errors appeared higher for drier catchments in all equations. For wetter catchments, both logarithmic and power relations approached the asymptotic value of 0.05 too gradually. The robustness of the equations for MAR and AI intervals was checked using box and whisker plots of relative residuals for all equations. The exponential equations for MAR and AI were slightly more robust than the other equations for all intervals, and MAR was only marginally better than AI (Fig. 8).

Of the terrain attributes, rainfall weighted slope (SLO), tree cover percentage (TC) and catchment elongation (CE) showed significant although weaker correlations with  $k_{bf}$ . The first two attributes were not included in stepwise regressions due to the cross-correlation with MAR. A relationship between tree cover, catchment slope and MAR is intuitively possible, for instance in the case of steep mountainous terrain where difficulty of access increases the chances of forest conservation and topography and altitude lead to enhanced orographic rainfall. The negative weaker correlation between catchment elongation (CE) and  $k_{bf}$  was also intuitively possible, implying the rounded catchments will drain faster than wider catchments of similar area (e.g. Post and Jakeman, 1996). Rainfall weighted slope also showed a negative correlation, opposite to the one expected from theory as in Zecharias and Brutsaert (1988, Eq. 1) and to results from similar correlation studies (Mwakalila et al., 2002; Brandes et al., 2005). The negative correlation between recession coefficients and rainfall weighted slope is counterintuitive; what common sense tells is that rugged catchments drain quicker than flatter ones (e.g. Post and Jakeman, 1996). However this effect may have less relevance when catchment areas are larger and other effects such as climate or complex topography/geology override the effects of slope, this was





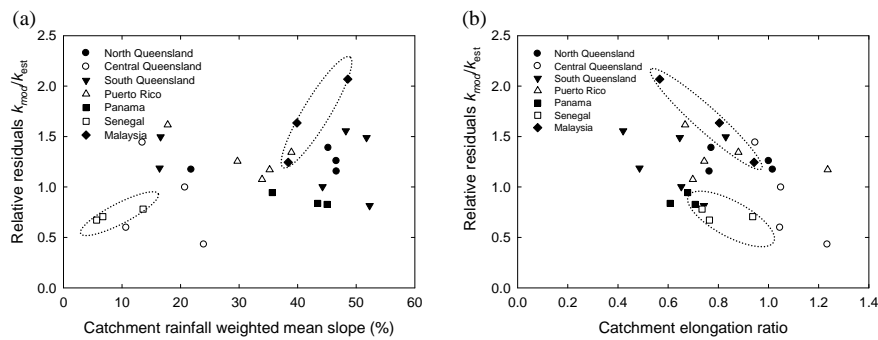
**Fig. 5.** Scatter plots of recession coefficient  $k_{bf}$  versus (a) MAR, (b) TMI, (c) AI, (d) SI, (e) SLO and (f) TC. Symbols denote a proxy for aquifer drainage potential from WHYMAP (2010).



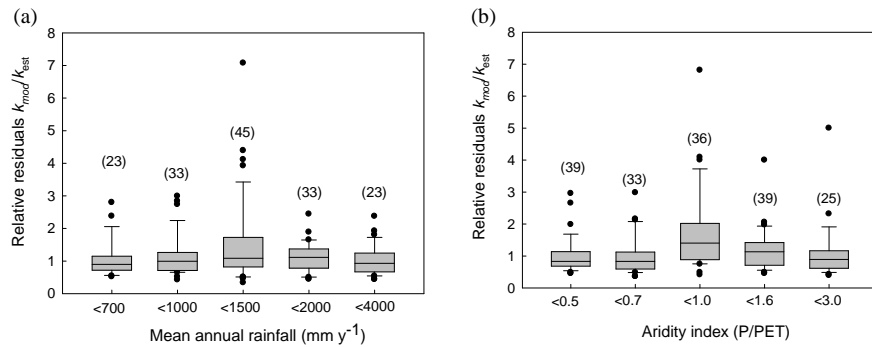
**Fig. 6.** Regression equations for (a) MAR versus  $k_{bf}$  and (b) AI versus  $k_{bf}$ .

also observed in Post and Jakeman (1996) but their results were not conclusive. A similar negative correlation was reported in Van Dijk (2010). The geology proxy used in the analysis, aquifer drainage potential, did not reveal any pattern with  $k_{bf}$  (Fig. 3) and its low correlation value indicated no influence on recession coefficients (Table 2). One would expect that geology and associated derivatives play a central

role in groundwater recession rates. This may well be ascribed to the lack of detailed geology at a global scale and the variety and geographic extent and distribution of catchments use in the study. Besides more detailed data, more robust methodologies could be used to include these attributes and the other catchment terrain attributes as covariates in regression equations. For example Detenbeck et al. (2005) used



**Fig. 7.** Scatter plot of relative residuals (ratio of modelled to hydrograph-based estimated  $k_{bf}$ ) versus (a) rainfall weighted slope of catchment (SLO) and (b) catchment elongation (CE). Elongated ellipses around Malaysian and Senegal data points are shown to illustrate possible correlations of residuals at smaller scales for geographically close catchments.



**Fig. 8.** Box and whiskers plot of relative residuals for exponential equations linking  $k_{bf}$  to (a) a range of mean annual rainfall classes (MAR). (b) Idem for aridity index (AI) range. The number in brackets is the sample size per range.

principal component analysis (PCA) to reduce dimensionality in an analysis to determine correlations between flow and velocity metrics in the North and South Shores of western Lake Superior (USA).

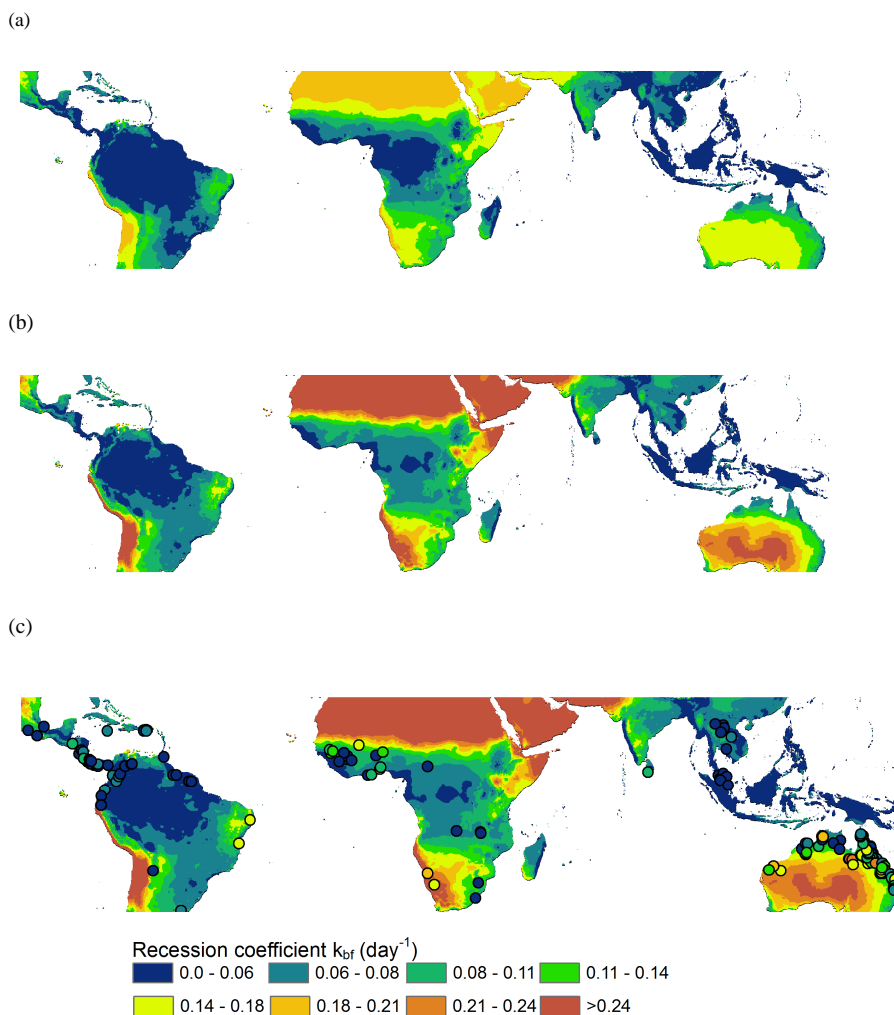
Pan-tropical maps of  $k_{bf}$  extending to  $30^{\circ}$  N and  $35^{\circ}$  S were derived using the MAR regression equation and the lower and upper bounds of the 95% confidence interval. The resulting catchment  $k_{bf}$  values are plotted in the map showing the original value (Fig. 9). A reasonable agreement is observed between original values and the ones using the regression equation.

The analysis of relative residuals for smaller catchments showed that catchment attributes such as slope (SLO) and elongation ratio (CE) had weak correlations with  $k_{bf}$ . Studies in catchments  $<100$  km<sup>2</sup> (e.g. Post and Jakeman, 1996; Brandes et al., 2005) also showed the explanatory power of terrain attributes and soils with respect to  $k_{bf}$  or other baseflow associated parameters. The present study and Van Dijk (2010) have demonstrated a more important role of climatic characteristics in relation to baseflow recessions across the tropics and Australia at catchment scales  $>100$  km<sup>2</sup>. Van Dijk (2010, Fig. 7 for the AI vs.  $k_{bf}$  plot) obtained similar power relationships between MAR, AI and baseflow recessions respectively for temperate Australian catchments. Esti-

mates of  $k_{bf}$  using the equations derived in the present study produced slightly higher estimates in these catchments, but the form of the relationships were similar. Differences between the rainfall data, and the Priestley-Taylor PET formulation in Van Dijk (2010) with the Hargreaves formulation in the present study may explain these differences.

The current empirical equations for the estimation of  $k_{bf}$  are necessarily subject to the limitations and uncertainties of the data used to derive them. There is no dense network of streamflow gauges in much of the tropics; the same can be said of the spatial density of gauges used to interpolate rainfall surfaces in the WORLDCLIM dataset (Hijmans et al., 2005). Furthermore, the estimation of  $k_{bf}$  was performed using long-term monthly climatologies, which would also yield long-term  $k_{bf}$  estimates. Coefficients will vary seasonally and interannually, with different  $k_{bf}$  estimates for dry-wet years and seasons.

It is expected that better rainfall data will result in more accurate relationships in the future. In addition, better soil and geological data may also improve the predictions.



**Fig. 9.** Pan-tropical map of baseflow recession coefficient using the exponential regression equation and mean annual rainfall (MAR): (a) equation representing the lower and (b) upper bounds of the 95% confidence interval and (c) original regression equation. Symbol colours represent estimated  $k_{bf}$  values of the 167 catchments used in this study.

## 6 Summary and conclusions

This study analysed the potential of various climatic and terrain attributes to estimate baseflow recession coefficients ( $k_{bf}$ ) for 167 unregulated tropical and subtropical catchments with areas  $>200 \text{ km}^2$ . Linear reservoir theory was used to estimate  $k_{bf}$  from daily streamflow data ( $\text{m}^3 \text{ s}^{-1}$ ) obtained from the Global River Discharge Center (GRDC). Stepwise regression showed the overriding importance of climatic attributes over terrain ones at this scale. The best predictors for baseflow recession coefficient were mean annual rainfall (MAR) and aridity index (AI) together explaining 49% of the variance.

The interaction between climate and surface and subsurface attributes also plays an important role at smaller scales. For example, catchment elongation (CE), a measure of catchment shape, was also found to be statistically significant, al-

though weakly correlated. An analysis of clusters of catchments of smaller size, showed that in these areas with presumably similar soils and geology, residuals of the regression could be explained by average catchment slope (SLO) and CE.

Although climatic characteristics explained a great deal of the variation in  $k_{bf}$ , baseflow is catchment-specific and dependent on the rainfall spatial and temporal patterns, land cover and land use, catchment morphology, infiltration opportunities and soil water holding capacity, configuration of the groundwater system and timing of groundwater discharge to the stream. The differences in  $k_{bf}$  found in humid and arid catchments show the interconnection of climate and surface and subsurface characteristics of catchments: ephemeral and consequently mainly event-driven streamflow as well as the occurrence of fast-draining perched aquifers may explain the higher recession coefficients observed in drier catchments.

The lowest recession coefficients in the humid tropics may be attributed to excess rainfall recharging deep soils and porous aquifers present in these areas (e.g. volcanic belts Central Amazonia, sandstone basin forms in Northeast Thailand). These sources may be an important source of baseflow during dry weather.

If better data are obtained for these surface and subsurface attributes, the prediction of baseflow in ungauged areas can be improved accordingly.

*Acknowledgements.* The authors gratefully acknowledge the funding from Microsoft Research and the Basin Focal Project Andes, a project of the CGIAR Challenge Program on Water and Food (CPWF). The GRDC Runoff Data Center and DERM Queensland are also thanked for providing the streamflow data. The authors are also grateful to Hannah Cloke from King's College London for providing valuable comments on the research. This study was conducted during an internship carried on in CSIRO Land and Water, Canberra-Australia and is part of the PhD dissertation of the first author.

Edited by: N. Verhoest

## References

- Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R., and Weber, J. L.: GLOBCOVER The most detailed portrait of Earth, *Esa Bulletin-European Space Agency*, 24–31, 2008.
- Arnell, N. W.: A simple water balance model for the simulation of streamflow over a large geographic domain, *J. Hydrol.*, 217, 314–335, 1999.
- Arnell, N. W.: Effects of IPCC SRES\* emissions scenarios on river runoff: a global perspective, *Hydrol. Earth Syst. Sci.*, 7, 619–641, doi:10.5194/hess-7-619-2003, 2003.
- Batjes, N. H.: ISRIC-WISE derived soil properties on a 5 by 5 arc-minutes global grid (version 1.0). Report 2006/02 ISRIC World Soil Information, Wageningen 2006. (Available at: [http://www.isric.org/isric/webdocs/Docs/ISRIC\\_Report\\_2006\\_02.pdf](http://www.isric.org/isric/webdocs/Docs/ISRIC_Report_2006_02.pdf), data available at: <http://www.isric.org> last access: 17 May 2010), 2010.
- Brandes, D., Hoffmann, J. G., and Mangarillo, J. T.: Base flow recession rates, low flows, and hydrologic features of small watersheds in Pennsylvania, USA, *J. Am. Water Resour. Assoc.*, 41, 1177–1186, 2005.
- Bruijnzeel, L. A.: Hydrological functions of tropical forests: not seeing the soil for the trees?, *Agr. Ecosyst. Environ.*, 104, 185–228, doi:10.1016/j.agee.2004.01.015, 2004.
- Brutsaert, W. and Nieber, J. L.: Regionalized Drought Flow Hydrographs From A Mature Glaciated Plateau, *Water Resour. Res.*, 13, 637–644, 1977.
- Chapman, T.: A comparison of algorithms for stream flow recession and baseflow separation. *Hydrological Processes*, 13, 601–604, 1999.
- Chappell, N. A., Sherlock, M., Bidin, K., Macdonald, R., Najman, Y., and Davies, G.: Runoff processes in Southeast Asia: Role of soil, regolith, and rock type, *Forest Environments in the Mekong River Basin, Part 1*, 3–23, 2007.
- Defries, R. S., Hansen, M. C., Townshend, J. R. G., Janetos, A. C., and Loveland, T. R.: A new global 1-km dataset of percentage tree cover derived from remote sensing, *Glob. Change Biol.*, 6, 247–254, 2000.
- Detenbeck, N. E., Brady, V. J., Taylor, D. L., Snarski, V. M., and Batterman, S. L.: Relationship of stream flow regime in the western Lake Superior basin to watershed type characteristics, *J. Hydrol.*, 309, 258–276, 2005.
- Doll, P., Kaspar, F., and Lehner, B.: A global hydrological model for deriving water availability indicators: model tuning and validation, *J. Hydrol.*, 270, 105–134, 2003.
- Fenicia, F., Savenije, H. H. G., Matgen, P., and Pfister, L.: Is the groundwater reservoir linear? Learning from data in hydrological modelling, *Hydrol. Earth Syst. Sci.*, 10, 139–150, doi:10.5194/hess-10-139-2006, 2006.
- Gascoïn, S., Ducharme, A., Ribstein, P., Carli, M., and Habets, F.: Adaptation of a catchment-based land surface model to the hydrogeological setting of the Somme River Basin (France), *J. Hydrol.*, 308, 105–116, 2009.
- Hagemann, S. and Dumenil, L.: A parametrization of the lateral waterflow for the global scale, *Clim. Dynam.*, 14, 17–31, 1998.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A.: Very high resolution interpolated climate surfaces for global land areas, *Int. J. Climatol.*, 25, 1965–1978, doi:10.1002/joc.1276, 2005.
- Lehner, B., Verdin, K., and Jarvis, A.: Hydrosheds technical documentation. Version 1.1. (Data available from: <http://hydrosheds.cr.usgs.gov>, last access 17 May 2010), 2008.
- Lepers, E., Lambin, E. F., Janetos, A. C., DeFries, R., Achard, F., Ramankutty, N., and Scholes, R. J.: A synthesis of information on rapid land-cover change for the period 1981–2000, *Bioscience*, 55, 115–124, 2005.
- Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J.: A Simple Hydrologically Based Model Of Land-Surface Water And Energy Fluxes For General-Circulation Models, *J. Geophys. Res.-Atmos.*, 99, 14415–14428, 1994.
- Mwakalila, S., Feyen, J., and Wyseure, G.: The influence of physical catchment properties on baseflow in semi-arid environments, *J. Arid Environ.*, 52, 245–258, doi:10.1006/jare.2001.0947, 2002.
- Nijssen, B., O'Donnell, G. M., Lettenmaier, D. P., Lohmann, D., and Wood, E. F.: Predicting the discharge of global rivers, *J. Clim.*, 14, 3307–3323, 2001.
- Oki, T. and Kanae, S.: Global hydrological cycles and world water resources, *Science*, 313, 1068–1072, doi:10.1126/science.1128845, 2006.
- Post, D. A. and Jakeman, A. J.: Relationships between catchment attributes and hydrological response characteristics in small Australian mountain ash catchments, *Hydrol. Process*, 10, 877–892, 1996.
- Saenz, L. and Mulligan, M.: Development and validation of a georeferenced tropics-wide database of dams, *Water Resour. Res.*, to be submitted, 2010.
- Schneider, A., Friedl, M. A., and Potere, D.: A new map of global urban extent from MODIS satellite data, *Environ. Res. Lett.*, 4, 044003, doi:10.1088/1748-9326/4/4/044003, 2009.
- Smakhtin, V. U.: Low flow hydrology: a review, *J. Hydrol.*, 240, 147–186, 2001.
- Tallaksen, L. M.: A Review Of Baseflow Recession Analysis, *J.*

- Hydrol., 165, 349–370, 1995.
- Thornthwaite, C. W.: An Approach toward a Rational Classification of Climate, *Geogr. Rev.*, 38, 55–94, 1948.
- Tomasella, J., Hodnett, M. G., Cuartas, L. A., Nobre, A. D., Waterloo, M. J., and Oliveira, S. M.: The water balance of an Amazonian micro-catchment: the effect of interannual variability of rainfall on hydrological behaviour, *Hydrol. Process.*, 22, 2133–2147, doi:10.1002/hyp.6813, 2008.
- Tularam, G. U. and Ilahee, M.: Exponential smoothing method of base flow separation and its impact on continuous loss estimates, *American Journal of Environmental Sciences*, 4, 136–144, 2008.
- van Beek, L. P. H., and Bierkens, M. F. P.: The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification, 2008. Report Department of Physical Geography, Utrecht University, Utrecht, The Netherlands, available at: <http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf>, last access: 17 May 2010.
- van Dijk, A. I. J. M.: Climate and terrain factors explaining streamflow response and recession in Australian catchments, *Hydrol. Earth Syst. Sci.*, 14, 159–169, doi:10.5194/hess-14-159-2010, 2010.
- Vogel, R. M. and Kroll, C. N.: Regional Geohydrologic-Geomorphic Relationships for the Estimation of Low-Flow Statistics, *Water Resour. Res.*, 28, 2451–2458, 1992.
- Vorosmarty, C. J., Green, P., Salisbury, J., and Lammers, R. B.: Global water resources: Vulnerability from climate change acid population growth, *Science*, 289, 284–288, 2000.
- Walsh, R. P. D. and Lawler, D. M.: Rainfall seasonality: description, spatial patterns and change through time, *Weather*, 36, 201–209, 1981.
- WHYMAP – World-wide Hydrogeological Mapping and Assessment Programme. Groundwater resources of the world 2010, Data available from <http://www.bgr.de/app/fishy/GoogleEarth/whymap.kml>, last access: 17 May 2010.
- Widen-Nilsson, E., Halldin, S., and Xu, C. Y.: Global water-balance modelling with WASMOD-M: Parameter estimation and regionalisation, *J. Hydrol.*, 340, 105–118, doi:10.1016/j.jhydrol.2007.04.002, 2007.
- Wittenberg, H.: Baseflow recession and recharge as nonlinear storage processes, *Hydrol. Process.*, 13, 715–726, 1999.
- Yu, P. S., Yang, T. C., and Liu, C. W.: A regional model of low flow for Southern Taiwan, *Hydrol. Process.*, 16, 2017–2034, 2002.
- Zecharias, Y. B. and Brutsaert, W.: Recession Characteristics Of Groundwater Outflow And Base-Flow From Mountainous Watersheds, *Water Resour. Res.*, 24, 1651–1658, 1988.