

# Comparing statistical and process-based flow duration curve models in ungauged basins and changing rain regimes

Marc F. Müller<sup>1</sup> and Sally E. Thompson<sup>1</sup>

<sup>1</sup>Department of Civil and Environmental Engineering, Davis Hall, University of California, Berkeley CA, USA

*Correspondence to:* Marc F. Müller (marc.muller@berkeley.edu)

**Abstract.** The prediction of flow duration curves (FDCs) in ungauged basins remains an important task for hydrologists given the practical relevance of FDCs for water management and infrastructure design. Predicting FDCs in ungauged basins typically requires spatial interpolation of statistical or model parameters. This task is complicated if climate becomes non-stationary, as the prediction challenge now also requires extrapolation through time. In this context, process-based models for FDCs that mechanistically link the streamflow distribution to climate and landscape factors may have an advantage over purely statistical methods to predict FDCs.

This study compares a stochastic (process-based) and statistical method for FDC prediction in both stationary and non-stationary contexts, using Nepal as a case study. Under contemporary conditions, both models perform well in predicting FDCs, with Nash-Sutcliffe coefficients above 0.80 in 75% of the tested catchments. The main drivers of uncertainty differ between the models: parameter interpolation was the main source of error for the statistical model, while violations of the assumptions of the process-based model represented the main source of its error. The process-based approach performed better than the statistical approach in numerical simulations with non-stationary climate drivers. The predictions of the statistical method under non-stationary rainfall conditions were poor if (i) local runoff coefficients were not accurately determined from the gauge network, or (ii) streamflow variability was strongly affected by changes in rainfall. A Monte Carlo analysis shows that the streamflow regimes in catchments characterized by frequent wet-season runoff and a rapid, strongly non-linear hydrologic response are particularly sensitive to changes in rainfall statistics. In these cases, process-based prediction approaches are favored over statistical models.

## 1 Introduction

The flow duration curve (FDC) provides a compact summary of the variability of daily streamflow by indicating what proportion of the flow regime exceeds a given flow rate. FDCs have considerable practical relevance, particularly in supporting decisions that are affected by the availability and reliability of surface water. Common applications of FDCs include the design and management of hydropower infrastructure (e.g., Basso and Botter, 2012; Müller, 2015), the determination of environmental flow standards for ecosystem protection (e.g., Lazzaro et al., 2013), the allocation of water resources for consumptive uses (e.g., Alaouze, 1989) or the prediction of streamflow time series in ungauged or poorly gauged catchments (e.g., Hughes and Smakhtin, 1996; Westerberg et al., 2014).

Despite their utility, empirical FDCs are unavailable for many basins, primarily because they require extensive on-site observations of daily streamflow (Vogel and Fennessey, 1994). Globally, the majority of catchments remain ungauged (or the gauge data that exist are subject to significant quality assurance and data availability constraints). Furthermore, the global number of stream gauges continues to decline because of ongoing budgetary constraints faced by water monitoring agencies (Stokstad, 1999; United States Geological Survey, 2015). Therefore FDCs must typically be estimated in data-scarce areas. The most widely used techniques for FDC estimation are simple, graphical methods. Such empirical methods are easy to implement but often rely on overly simplistic assumptions that lead to substantial prediction errors. For instance in Nepal, the regionalization method prescribed in official design manuals (e.g., Chitrakar, 2004; Alternative Energy Promotion Center, 2014) relies on *one* in-situ observation of streamflow during the dry season to scale standardized regional indices for monthly flows. The procedure neglects the inter-annual variability of low-flows, which leads to important biases in the predicted flow distributions (see Section S1 of Supplementary Materials). Even in gauged catchments, FDCs constructed from historical observations may not represent current flow conditions well, because flow regimes are impacted by climate change and anthropogenic alterations of the catchments (e.g. Botter et al., 2013; Mu et al., 2007). Predicting streamflow in ungauged basins, particularly in the context of environmental change, remains both a fundamental necessity for water managers and a major research challenge (Blöschl et al., 2013; Montanari et al., 2013).

Recent efforts to predict FDCs in ungauged catchments focus on statistical approaches that predict the flow distribution based on the catchment's similarity to nearby, gauged watersheds (Castellarin et al., 2013). Index flow approaches, which regionalize specific index flows (typically the mean flow), and use those indices to rescale empirical FDCs from similar catchments, are particularly popular (e.g., Chalise et al., 2003; Castellarin et al., 2004b; Sauquet and Catalogne, 2011; Arora et al., 2005). While differing in methodological details, all index flow approaches assume that FDCs do not vary within homogeneous regions, except by a scaling factor. Because they do not assume any specific runoff-generating process, statistical methods are versatile. They have been successfully been applied globally to predict FDCs in a variety of climates and catchment types (Blöschl et al.,

2013). However, methods are also insensitive to the diversity of controls on the shape of the FDC exerted by climate processes and catchment characteristics. This may affect their reliability under non-stationary conditions (Milly et al., 2008). Finally, the calibration of statistical methods relies on extensive streamflow observations from a large number of representative and well characterized catchments (e.g., Cheng et al., 2012; Coopersmith et al., 2012). Their performance is therefore sensitive to the spatial density of available gauges Blöschl et al. (2013), and their reliability in regions where streamflow data is truly scarce is uncertain.

Stochastic, process-based models that mechanistically link the drivers, state and response of the system are a promising avenue to address these issues. In these models, basic assumptions about the stochastic structure of rainfall and the (deterministic) response of catchments allow the analytic derivation of streamflow probability density functions (PDFs). (Note that because the FDC can be obtained directly by transforming the PDF, a predictive technique that yields the streamflow PDF will also allow the FDC to be estimated). Botter et al. (2007b) show that runoff follows a gamma distribution if catchments behave as a linear reservoir, forced by stochastic rainfall that follows a marked Poisson process. The resulting gamma distribution depends on two parameters that are determined by the recession characteristics of the catchment, and by the frequency and intensity of effective rain. This process-based approach to the streamflow PDF has been extended to include the fast flow component of streamflow (Muneepeerakul et al., 2010), non-linearities in subsurface storage-runoff relationships (Botter et al., 2009), the effects of short-term snowmelt (Schaeffli et al., 2013) and the carryover of subsurface storage between seasons in seasonally dry climates (Müller et al., 2014). Although the stochastic framework allows the effects of changes in climate or landscape to be independently modeled, it relies on strong simplifying assumptions about the spatial homogeneity of catchments. These assumptions makes the existing process models less versatile than statistical methods. Nonetheless, the approach has low calibration requirements because it relies on a small number of parameters, which can be determined using rainfall, climate and geomorphological characteristics of the catchments (Doulatyari et al., 2015). This information is increasingly available in ungauged basins, thanks to remote-sensing technologies, even when ground-based measurements are sparse.

Process-based models successfully reproduce streamflow PDFs in numerous gauged catchments worldwide (Botter et al., 2007a; Ceola et al., 2010), including Nepal (Müller et al., 2014). Yet their predictive performance in ungauged basins remains largely unassessed, particularly in regions where the local gauge density is globally representative (as opposed to densely monitored catchments in developed countries such as, e.g., France and Austria in Castellarin et al. (2013). For lower gauge densities, it is unclear whether the advantages of the process-based approaches, which are derived from an explicit representation of flow-generating processes, are outweighed by the limitations imposed by the restrictive assumptions underlying these methods - and whether this trade-off is altered by non-stationarity in climate drivers.

95 Using Nepal as a test case, this study compares the process-based and statistical approaches on the basis of (i) their ability to predict FDCs in ungauged basins, (ii) their sensitivity to data-scarcity, represented both by the spatial density of the stream gauge network and by the temporal extent (length) of the available streamflow records, and (iii) their ability to accommodate changes in the rainfall regime.

100 Nepal provides an ideal setting to compare the two approaches, for four reasons. First, the country is representative of global availability of streamflow data, as measured by the density of its stream gauge network (Figure 1(a)). Second, methods drawn from both statistical and process-based approaches have been developed and validated in Nepal. Here we compare the stochastic-dynamic framework developed in Müller et al. (2014), with the index flow model described in Chalise et al. 105 (2003). Third, flow generation processes in Nepalese Himalayan catchments are complex, particularly with respect to the spatial and temporal properties of precipitation. Rainfall derives from the Indian Summer Monsoon and is strongly affected by topography. As a result, local rainfall is temporally autocorrelated, spatially heterogeneous and highly seasonal. There is also significant carryover of groundwater storage between the wet and dry seasons, so that dry season discharge reflects the 110 features of the antecedent wet season. These characteristics violate many of the assumptions that underlie the process-based method. The analysis in Nepal is therefore likely to provide a conservative estimate of the potential performance of the process-based method in ungauged basins. Finally, developing reliable methods for FDC prediction in Nepal represents an opportunity for ‘use-inspired science’ (Thompson et al., 2013b). Nepal has an enormous untapped hydropower potential and is 115 in dire need of electrical power, particularly in rural areas. A reliable method to estimate FDCs in ungauged catchments would be a valuable tool to support the development of micro hydropower, a sustainable technology for rural electrification (Müller, 2015).

Section 2 describes the two models and the procedures used to estimate their parameters from streamflow and rainfall observations. Section 3 presents the results of the comparative analysis in 120 Nepal. Section 4 examines the key sources of errors for both models and discusses implications for both Prediction in Ungauged Catchments (PUB) and Predictions Under Change (PUC) beyond Nepal.

## 2 Methods

### 2.1 Compared Approaches

#### 125 2.1.1 Process-based model

The process-based approach models daily streamflow as a random variable. Subject to strong simplifying assumptions about rainfall stochasticity and runoff generation, the streamflow PDF can be analytically derived. During the wet season, daily rainfall is represented as a stationary marked Pois-

son process with exponentially distributed depths. Assuming linear evapotranspiration losses, Botter et al. (2007b) showed that effective rain, that is the portion of the total rainfall that contributes to streamflow generation, also follows a stationary marked Poisson process. For a spatially homogeneous catchments with an exponentially distributed response time (i.e. a catchment that behaves as a linear reservoir), this effective rainfall will produce gamma-distributed streamflow. The parameters of the gamma distribution are derived from the frequency ( $\lambda_P$ ) and mean depth ( $\alpha_P$ ) of rainfall, and from the recession constant ( $k$ ) of the catchment. If rainfall in the dry season is sufficiently minimal that effective rainfall does not contribute to runoff generation, then dry season streamflow represents only the discharge of groundwater stored during the previous wet season. This discharge is modeled as a single seasonal recession with stochastic initial conditions that depend on the wet season properties. Because groundwater is not replenished during the dry season, the water table is subject to a large transient drawdown, resulting in a nonlinear discharge behavior and a power law relation between recession rate and discharge (Brutsaert and Nieber, 1977). We showed in Müller et al. (2014) that the distribution of streamflow, and therefore the FDC, in seasonally dry climates that meet the assumptions above can be expressed analytically as a function of seven independent parameters: the frequency ( $\lambda_P$ ) and mean intensity ( $\alpha_P$ ) of wet season rainfall, maximum daily evapotranspiration during the wet season ( $ET$ ), the water storage capacity of the soil in the root zone ( $SSC$ ), the (linear) wet-season recession constant ( $k$ ), the duration of the dry season ( $T_d$ ) and the exponent of the power law recession during the dry season ( $b$ ). The model admits an additional input parameters, the scale  $a$  of the of the power-law seasonal recession, which we showed in Müller et al. (2014) can be expressed as a function of  $k$ ,  $b$ ,  $\lambda_P$  and  $\alpha_P$ . The formal derivation of the model is summarized in Appendix A.

The model was successfully validated in a variety of regions with seasonally dry climates worldwide, including Nepal, where observed FDCs were predicted in 24 gauged catchments with a median Nash Sutcliffe Coefficient of 0.90 on log-transformed flow quantiles (Müller et al., 2014). The approach successfully reproduced both the rain-driven distribution of flows during the wet season and the release of stored monsoon water during the dry season recession. In this study, we assess the operational performance of the process-based approach as a tool to predict streamflow in *ungauged* catchments. Therefore, we do not further attempt to attribute model errors to parameters versus the model structure in the results presented in Section 3, since in practice these errors are confounded in any real application. The relative significance of these two error sources is nonetheless discussed in Section 4.1.1.

In ungauged catchments, the process-based model is implemented as follows. Three of the seven parameters of the model ( $T_d$ ,  $\lambda_P$ ,  $\alpha_P$ ) are rainfall characteristics that can be estimated in ungauged basins using meteorological observations. Recession parameters ( $k$  and  $b$ ) describe aquifer properties that are challenging to observe at the catchment scale. They can be estimated using observed streamflow time series in nearby gauged basins, and subsequently interpolated from nearby gauges,

using the geostatistical approach described in Müller and Thompson (2015), which accounts for the topology of the stream network. The last two parameters ( $ET$ ) and ( $SSC$ ) describe catchment-scale soil moisture dynamics that are arduous to determine empirically. Previous applications of the model relied on reasonable values of  $ET$  and  $SSC$ , based on land use, soil and climate characteristics of the catchment (e.g., Botter et al., 2007a; Ceola et al., 2010). Alternatively, runoff coefficients can be used to directly relate rainfall statistics to streamflow increments (Doulatyari et al., 2015). Runoff coefficients describe the ratio of mean discharge to mean precipitation, and can be predicted in ungauged basins using water balance models and meteorological observations. This approach circumvents the need to estimate  $ET$  and  $SSC$ , but the accuracy of predicted runoff coefficients in ungauged catchment is critically dependent on the type of water balance model used and on the availability of appropriate calibration data (Doulatyari et al., 2015). Instead, this study follows the former procedure and uses reasonable estimates of  $ET$  and  $SSC$  for Nepal.

### 2.1.2 Statistical model

The statistical approach is entirely driven by observation data and does not assume any specific runoff generation process. Instead, it identifies and exploits statistical correlations that may occur between streamflow observed at existing gauges and the geology, topography and climate of the corresponding catchments. The index flow model used in this study was developed by Chalise et al. (2003) to regionalize FDCs in Nepal to assess the potential for small hydropower development. The model is based on local flow indices for mean ( $Q_m = E[Q]$ ) and low flows ( $q_{95} = Q_{95}/Q_m$ , where  $Q_{95}$  is the 95<sup>th</sup> streamflow percentile) and uses a non-parametric approach to represent the shape of the FDC. Empirical FDCs from available gauges are normalized by  $Q_m$  and pooled into equally-sized groups based on the  $q_{95}$  index of the gauge. A standardized curve is determined for each group by taking the average of the normalized flows corresponding to each duration, in order to represent the average catchment response in the group. The chosen statistical approach is considerably less complex than many alternative state-of-the-art methods using multiple (often non-linear) equations to relate multiple flow quantiles to a variety of observed covariates (see Castellarin et al. (2013) for a review). However, Chalise et al. (2003) is, to our knowledge, the most recent statistical method specifically developed and validated in the study region. The approach is parsimonious and adapted to situations, where in-situ observations of catchment characteristics are scarce. The method is therefore representative of the level of complexity of statistical approaches likely to be implemented in developing countries practical hydrological engineering purposes.

Predictions in ungauged catchments are obtained by first using linear regressions to predict  $Q_m$  and  $q_{95}$ . Although the original method calls for a stepwise multiple regressions approach to determine regression covariates inductively, we used the regression models obtained in Chalise et al. (2003):  $Q_m$  is regressed against annual rainfall ( $R_y$ ) and gauge elevation ( $z_{min}$ ) as a proxy for evapotranspiration; and  $q_{95}$  is regressed against the ratios of catchment area occupied by each of the

considered geological units. The two regressions loosely represent the long term water balance and short-term response of the catchment. The predicted low-flow index is then used to determine the standardized FDC shape, which is finally multiplied by the predicted mean flow to obtain the FDC.

205 An important assumption, inherent to the linear regression models, is that the dependent variable (here  $Q_m$  and  $q_{95}$ ) is not spatially correlated, when controlling for the considered covariates. This assumption is reasonable in Nepal, where the typical distance between stream gauges is much larger than the correlation scale of runoff (Müller and Thompson, 2015). In more densely gauged areas (or if runoff is correlated over larger distances), streamflow observations at neighboring or flow-  
210 connected gauges are likely to be correlated. In these regions, accounting for the effect of distance and stream network topology when interpolating flow indices (e.g., using TopREML (Müller and Thompson, 2015)) will improve predictions.

## 2.2 Study region and data

The two methods were evaluated using observed streamflow data from 25 Nepalese catchments mapped in Figure 1 (b). The gauges in this dataset (HKH-FRIEND, 2004; Department of Hydrology and Meteorology, 2011) have at least 10 years of daily streamflow records. They were checked for consistency, using double mass plots (Searcy and Hardison, 1960), and bias: we discarded non-glaciated catchments that had a precipitation deficit on their long term water balance. Watersheds were delineated using the ASTER GDEM v2 digital elevation model (NASA Land Processes Distributed Active Archive Center (LP DAAC), 2011). The study watersheds are located in central Nepal but cover a wide variety of catchment sizes, elevation ranges, precipitation characteristics and geological units (Table 1).  
220

We focused on the Chepe Kohla catchment in central Nepal (Figure 1 (b, insert)) as a case study for analyses requiring resampling (Section 2.3.1) or simulation (Section 2.3.2) of streamflow time series. The Chepe Kohla watershed has a long (by Nepalese standards) record of daily streamflow observations (31 years) and is representative of the full sample of gauges in terms of topography and recession behavior (Table 1). The catchment is also small (i.e. close to spatially homogenous) and local rainfall is well approximated by a marked Poisson process (First order autocorrelation coefficient of rainfall occurrence (AR): 0.09, Coefficient of variation of rainfall depths (CV): 1.09),  
225  
230 echoing the underlying assumptions of the process-based model.

Rainfall characteristics over the sampled catchments were obtained from 178 precipitation gauges (HKH-FRIEND, 2004; Department of Hydrology and Meteorology, 2011), also mapped on Figure 1 (b). The average duration of the dry season ( $T_d$ ) was estimated at each precipitation gauge by fitting a step function to the corresponding rainfall time series (Müller and Thompson, 2013), and wet-season precipitation records were used to compute the frequency and mean intensity of rainfall ( $\lambda_P$  and  $\alpha_P$ ). Rainfall characteristics were then aggregated at the catchment level by assuming that the rain process aggregates linearly within the basins. For rainfall occurrence, we assumed that the

duration between rain events caused by two consecutive storms can be estimated as the average of the inter-arrival times measured at the rain gauges within the catchment. This allows us to compute catchment level rainfall frequency as:

$$\lambda_P = \left( \frac{1}{N_g} \sum_i^{N_g} \frac{1}{\lambda_P^{(i)}} \right)^{-1}$$

where  $\lambda_P^{(i)}$  designates rainfall frequency observed at gauge  $i$  and  $N_g$  the number of rain gauges within the catchment. Similarly, the catchment-level duration between rainy seasons is assumed to be the average of the durations observed within the catchment:

$$T_d = \frac{1}{N_g} \sum_i^{N_g} T_d^{(i)}$$

Finally, the precipitation depth received on any given day by a catchment is assumed to be the average of the precipitation depths observed by individual rain gauges. It follows that the aggregated mean rainfall intensity can be expressed as:

$$\alpha_P = \lambda_P^{-1} \frac{1}{N_g} \sum_i^{N_g} \lambda_P^{(i)} \alpha_P^{(i)}$$

If no precipitation station is located within the catchment, rainfall characteristics observed at the rain station closest to the catchment centroid were considered. Although aggregating rainfall *time series* before computing their statistics would better account for spatial correlation in rainfall, aggregating rainfall *statistics* instead allows for non-overlapping observation periods (assuming rainfall is stationary). This is important in the context of Nepal, where rain gauges are scarce with sporadic observations. Unfortunately, the low density of rain gauges within the considered basins prevents a formal treatment of spatial correlation when aggregating frequencies. However, in a previous study (Müller and Thompson, 2013) we observed large spatial correlation ranges on rainfall occurrence in Nepal (125km during the monsoon). Under these conditions the selected method stands out as the most parsimonious approach to utilize multiple, yet sparse, rainfall observations.

Recession characteristics were estimated using streamflow observations as described in Müller et al. (2014). We computed wet season recession constants ( $k$ ) by regressing the logarithm of streamflow against time for each period of consecutively decreasing streamflow during the wet season. The recession constant was then obtained by taking the median value of the regression coefficients of recessions lasting more than four days. The power law exponent of dry season recessions ( $b$ ) was obtained by fitting a non-linear recession curve

$$Q(t) = (Q_0^{1-b} - a(1-b)t)^{\frac{1}{1-b}} \quad (1)$$

to base flow, which was computed from observed streamflow time series using the Lyne Hollick algorithm (Nathan and McMahon, 1990). The last streamflow peak of the wet season was taken as

250 initial flow condition  $Q_0$ , and we used a stochastic optimization algorithm (Simulated Annealing, B  lisle (1992)) to minimize least square fitting errors. In ungauged catchments, the scale exponent of the seasonal recession was approximated as (M  ller et al., 2014):

$$a \approx \frac{\lambda}{-r} \left( e^{\frac{-r}{m}} - 1 \right) (\alpha_Q \cdot (m + 1)), \quad (2)$$

where  $r = 1 - b$ ;  $m$  is the ratio between the frequency  $\lambda$  of effective rain events and the linear 255 recession constant  $k$ , and  $\alpha_Q$  is the average depth of effective rain events (see Appendix A).

Potential evapotranspiration was approximated by applying the empirical relation estimated by Lambert and Chitrakar (1989) for Nepal during the rainy season (July-September):

$$ET \approx 4.0 - 0.0008 \cdot z_{mean}$$

where  $ET$  is given in  $[mm/d]$  and  $z_{mean}$  is the average elevation of the catchment in meters. The formula provides daily average evapotranspiration estimates for each month. It accounts for elevation but assumes a spatially homogenous elevation gradient. A uniform soil moisture capacity of 50 mm was assumed throughout the country, based on empirical observations reported in Shrestha 260 (1997). By neglecting local variation in soil characteristics, this produces conservative estimates of the performance of the process-based model in ungauged basins.

## 2.3 Comparative Analyses

### 2.3.1 Predictions in Ungauged Basins

We used three cross-validation techniques to evaluate the predictive ability of both methods in un- 265 gauged basins. Firstly, a leave-one-out analysis was carried out to assess predictive performances in a realistic situation, where FDCs are predicted in Nepal using all streamflow gauges available in the region. Secondly, we examined the sensitivity of the methods to decreasing data-availability by reducing the number of gauges available to calibrate the models. Finally, we performed a similar data-degradation procedure, but in this case we reduced the number of daily streamflow observa- 270 tions, while holding the number of gauges constant. This final analysis accounts for the challenges posed by recent or temporary installation of stream gauges, which introduce uncertainties into the estimation of model parameters due to the short streamflow records used. These errors can propagate through the model and affect the prediction of FDCs.

In a leave-one-out analysis, one gauge is ‘left out’ of the dataset, and streamflow is predicted at 275 the ‘missing’ location using observations from the remaining gauges. The predicted FDC is then compared to observations from the omitted gauge. The resulting error between observation and prediction yields the prediction performance of the method at that catchment if it were not gauged. Repeating the procedure for all gauges offers an approximation to the overall prediction error of the method. To measure this error we constructed error duration curves (M  ller et al., 2014), where the 280 relative prediction error at each flow quantile is plotted against the corresponding duration. Error

duration curves allow the partitioning of prediction errors across flow quantiles to be visualized. General prediction performances (across all durations) at individual gauges were also determined using the Nash Sutcliffe coefficient (NSC) on log streamflow quantiles (Müller et al., 2014):

$$NSC = 1 - \frac{\sum_{t=1}^{364} \left( \ln Q_t^{(emp)} - \ln Q_t^{(mod)} \right)^2}{\sum_{t=1}^{364} \left( \ln Q_t^{(emp)} - E \left[ \ln Q_t^{(emp)} \right] \right)^2} \quad (3)$$

285 where  $Q_t^{(emp)}$  and  $Q_t^{(mod)}$  are the empirical and modeled streamflow quantile of duration  $t$ .

The effect of the number of calibration gauges was assessed using a Jack-Knife cross-validation analysis (Shao and Tu, 2012; Müller and Thompson, 2013). At each of 10,000 iterations, a selected fraction of the available gauges was randomly sampled (without replacement) and used to predict the FDC at one (randomly selected) remaining gauge. Prediction accuracies for flow duration curves  
290 (given by the  $NSC$ ) and uncertainties on the spatial interpolation of model parameters were reported for each iteration. The procedure was repeated for decreasing numbers of selected ‘training’ gauges.

The available streamflow data did not allow a direct evaluation of the effects of timeseries length through cross-validation, because such an analysis requires substantial overlaps in the monitoring periods of all gauges. Therefore we focused the final analysis on the Chepe Kohla catchment which  
295 has the longest observation record in our dataset. We evaluated the effect of the length of the available observation records on parameter estimation, and propagated the ensuing uncertainty in the parameters to the FDCs predicted by each model. To do this, we selected a fixed number of full years of streamflow observations, estimated the parameters, predicted the FDC using these parameters, and compared the results to the empirical FDC obtained from the full observation record. The procedure  
300 was repeated 10,000 times. The estimation errors in the model parameters and the resulting FDC prediction performances ( $NSC$ ) were recorded as a function of the number of sampled years. This analysis is not intended to describe the models’ ability to predict FDCs at catchments with short observation records: in this case, constructing an empirical FDC using the available (however short) observation record is likely to be the best course of action (Castellari et al., 2004a). Instead, the  
305 analysis is intended to simulate the effect of short observation records on FDC prediction at nearby, *ungauged* catchments. The underlying assumptions behind this analysis are that (i) the error associated with interpolation is independent of the flow record length, and (ii) the Chepe Kohla catchment is representative of Nepalese basins .

### 2.3.2 Predictions Under Change

310 We used numerical simulations to assess the ability of both models to predict streamflow when subject to changing rainfall regimes, as described in Figure 2.

Synthetic streamflow time series were generated by coupling the stochastic rainfall generator described in Müller and Thompson (2013) to a rainfall-runoff model. The generated wet-season rainfall is a first order Markov process (i.e. rainfall occurrence on a given day is correlated to rainfall oc-

315 currence on the previous day) with gamma-distributed rainfall intensities, and as such produces a  
 rainfall record that explicitly violates the assumptions under-pinning the process-based model. The  
 duration of the rainy season was assumed constant, and no rainfall was generated during the dry sea-  
 son. Wet-season streamflow was simulated by feeding synthetic rainfall into a linear reservoir (with a  
 recession constant  $k$ ) with linear evapotranspiration losses, as in Müller et al. (2014). Dry season dis-  
 320 charge was obtained by simulating non-linear seasonal recessions of duration  $T_d$  starting at randomly  
 selected runoff peaks in the (previously generated) wet-season streamflow. These assumptions are  
 close to the observed reality in Nepal, as seen in Figure 5 (a), where the FDC constructed from the  
 simulated streamflow is a close approximation to the empirical FDC in the Chepe Kohla watershed.  
 We translated the effect of shifts in precipitation regimes into changed streamflow for the Chepe  
 325 Kohla catchment by considering a range of future combinations for rainfall frequencies and intensi-  
 ties. In line with what is expected in Nepal (Turner and Slingo, 2009; Turner and Annamalai, 2012),  
 we considered negative changes in the frequency and positive changes in the mean daily rainfall  
 depth. We neglected changes in soil moisture capacity, evapotranspiration, rainfall autocorrelation  
 and the duration of the rainy season. These parameters are explicit in the process-based model, so  
 330 we expect differences in the sensitivity of the process-based and statistical models to climate change  
 to be underestimated by this procedure. For each rainfall scenario, we evaluated the performance of  
 the models in a changing climate by generating 1000 years of daily streamflow using future rainfall  
 frequencies and intensities.

We compared the synthetic FDCs to model predictions that were made with *future* rainfall statis-  
 335 tics, but *contemporary* recession and low flow parameters (Figure 2). The statistical method in  
 Chalise et al. (2003) uses a linear regression over a cross-sectional sample of observations to pre-  
 dict mean flow based on mean rainfall and altitude. The regression may fail to capture a variety of  
 unobserved characteristics affecting both rainfall and streamflow (e.g., local topographic features),  
 and hence may not capture the causal relation between the two variables. The extent of this bias  
 340 cannot be quantified a priori, so we considered two extreme cases: infinite and zero bias. The infinite  
 bias case (Case 1 on Figure 2) represents the case where no effective relationship can be determined  
 between rainfall and mean flow. The best estimator of future mean flow is then the *current* flow  
 condition. Conversely, if regression coefficients perfectly describe the effect of annual rainfall on  
 average flow (Case 2 on Figure 2), then the future flow conditions can be perfectly estimated using  
 345 the (known) future annual rainfall. We modeled this situation by estimating  $Q_m$  directly from the  
 (simulated) *future* flow conditions. While the two cases differed in the determination of mean flow  
 ( $Q_m$ ), the low-flow parameter ( $q_{95}$ ) was determined from current flow conditions in both cases. In  
 Chalise et al. (2003),  $q_{95}$  is normalized by  $Q_m$  and represents recession behavior, which is assumed  
 independant from rainfall. The process-based predictions were obtained by inserting future rainfall  
 350 statistics and contemporary recession constants into the analytical FDC equation described in Ap-

pendix A. The two models were compared by plotting prediction performances ( $NSC$ ) against the relative change in the frequency and intensity of synthetic rainfall.

Although the recession assumptions of the process-based model are taken to generate the synthetic streamflow used as control, we believe that the analysis is not biased against the statistical approach for three reasons. Firstly, the only parameter of the statistical approach that is influenced by rainfall ( $Q_m$ ) is also computed from synthetic streamflow (Case 2 on Figure 2). Secondly, although based on identical recession assumptions, the process-based model and the synthetic streamflow generator are driven by different stochastic rainfall processes (i.e. Poisson and Markov respectively). Lastly and most importantly, empirical observations reveal that synthetic streamflow distributions generated under contemporaneous rainfall conditions reproduce closely FDCs constructed from gauge records (Figure 5 (a)), showing that the underlying recession assumptions are, in fact, representative with runoff processes actually occurring in Nepal.

### 3 Results

#### 3.1 Prediction in Ungauged Basins

Results from the leave-one-out cross-validation analysis are presented in Figure 3 and show that both methods perform similarly in the prediction of FDCs in ungauged basins. Error duration curves (Figure 3 (a) and (b)) show comparable streamflow prediction uncertainties: 75% of the predicted flow quantiles are between half and double the observed streamflow for both models, although the low flows in the process-based model display an increasing upwards bias (Figure 3 (b)). Considering the Nash-Sutcliffe coefficients computed at the individual basin level, the mean and median performances are again comparable for both models, but the accuracy of the statistical model predictions are more variable across sites than the process model predictions, as indicated by the larger spread of the Nash Sutcliffe coefficients (Figure 3(c)).

Figure 4 (a, top) shows prediction performances of both models as the number of streamflow gauges available for predictions decreases, and indicates that the performance of both models is relatively insensitive to the gauge density, until it declines to less than approximately 0.6 gauges per 10,000 km<sup>2</sup>. For such situations, which represent discarding more than half the available gauges in Nepal, the statistical model performance declines rapidly compared to the process-based model. Prediction performances are strongly affected by uncertainties on the interpolation of model parameters, as seen in Figure 4 (a, bottom). Interpolation uncertainties are generally larger for the flow indices of the statistical model ( $Q_m$  and  $q_{95}$ ) than for the recession parameters of the process-based model ( $k$  and  $b$ ). This explains the larger spread in prediction performances of the former (Figure 3 (c) and error bars in Figure 4 (a, top)). The parameter uncertainties are also relatively insensitive to the total gauge density until about 60% of the originally available gauges are discarded. At this

point, the uncertainties associated with estimation of the flow indices increase significantly, while the process-based model parameters remain more reasonably estimated.

When considering short observation windows, parameter uncertainties also drive the performance of the models. Figure 4 (b, top) shows the prediction performance of both models at the Chepe Khola watershed, as the number of observation years used to estimate the model parameters is reduced. In this case, the statistical model outperforms the process-based model when less than 10 years of streamflow observations are available. The parameter uncertainties associated with the short time-series estimates (Figure 4 (b, bottom)) suggest that a longer time series of streamflow observations is needed to accurately estimate the wet-season recession parameter ( $k$ ), resulting in the lower performance of the process-based model for short streamflow records.

### 3.2 Prediction Under Change

Simulation results presented in Figure 5 (b) show both models' ability to predict a simulated future flow duration curve of the Chepe Kohla River, under a range of different possible changes in rainfall regimes. In all simulations, parameters describing the hydrological response of the basin ( $k$ ,  $b$ , and  $q_{95}$ ) are determined using current flow conditions, and evapotranspiration is assumed constant.

The results show that explicitly modeling rainfall-runoff processes allows the process-based model to accommodate the effects of the changing precipitation regime. In contrast, the performance of the statistical model is affected at various degrees by shifts in rainfall regimes, depending on how the model translates changes in annual precipitation to changes in average flows. If these shifts are perfectly represented by the model, then prediction errors arise solely from changes in the shape of the FDC, and the process and statistical models perform similarly in the Chepe Kohla watershed across the full range of considered rainfall scenarios (Figure 5 (b), dashed curve). If, however, average (future) streamflows cannot be reliably predicted from the predicted changes in annual rainfall, the statistical model does not accommodate flow regime changes at all. In this case, future FDCs are modeled using current streamflow observations, and the ensuing prediction errors can be substantial (Figure 5 (b), dotted curve). The simulated cases provide upper and lower bounds for the actual performance of the statistical model in future rainfall regimes. We evaluated the model's ability to predict  $Q_m$  by using cross sectional data (i.e. average streamflow and annual rainfall from the 25 catchments) to estimate the linear relation between  $Q_m$  and annual rainfall  $R_y$ . Applied to the Chepe Kohla watershed, the estimated regression coefficients allowed the annual streamflow to be estimated from annual precipitation with a bias of  $-13\%$  and a coefficient of determination of  $R^2 = 0.57$  (Figure 5 (c)). Regardless, prediction errors remained negligible for both bounds ( $NSC > 0.95$ ) for the range of changes actually anticipated in Nepal (e.g.,  $\Delta\lambda_P/\lambda_P \approx 0.98$  and  $\Delta\alpha_P/\alpha_P \approx 1.20$  for the  $2 \cdot CO_2$  scenario (Turner and Slingo, 2009)).

## 4 Discussion

### 4.1 Predictions in Ungauged Basins

The analysis suggests that both statistical and process-based methods to estimate FDCs in ungauged basins perform comparably in Nepal, over a wide range of gauge densities and observation durations. Yet prediction performances varied significantly between the models as data became increasingly sparse. The statistical method is more sensitive to spatially sparse data, which degrades the interpolation accuracy of  $Q_m$ . In contrast, the estimation method for recession parameters makes the process-based approach more sensitive to temporally restricted observations, which reduce the accuracy with which recession parameters can be estimated. This suggests that the performance of the two models in ungauged basins is affected by different sources of uncertainty. In this section, we investigate the source of prediction error in each methods and discuss the implications for their application in ungauged basins beyond Nepal.

#### 4.1.1 Sources of Uncertainty

The statistical model relies on two assumptions about the correlations of observed data. The first assumption is that catchments with similar low-flow indices ( $q_{95}$ ) have identical hydrological responses, and therefore identical FDC shapes. Second, the model assumes that the flow indices ( $Q_m$  and  $q_{95}$ ) at ungauged catchments can be best predicted using linear regressions against observable covariates (annual rainfall, elevation and geology). The latter assumption does not hold if the flow indices are spatially auto-correlated, or if the posited linear relations are spatially heterogeneous or, in fact, non-linear. Further, ‘omitted variable’ biases (Greene, 2003) will arise if an unobserved variable is correlated to both a covariate and a flow index. For instance, local topographic features may affect both the annual rainfall and the average streamflow in mountainous regions. Violation of the second assumption leads to substantial uncertainty in the interpolation of the flow indices in Nepal and drives the prediction errors of the statistical approach, as shown in Section S2 of the supplementary materials.

While the performance of the process-based model is also driven by parameter estimation uncertainties, these errors arise from simplifying assumptions about local hydrological processes (rather than uncertainties from their statistical interpolation from neighboring gauges). Additional cross-validation analyses (shown in Section S2 of the supplementary material) suggest that uncertainties caused by the aggregation of observed point-rainfall statistics at the catchment level drive prediction errors of high flow quantiles. While increasingly accurate remote sensing rainfall data will progressively allow such spatial heterogeneities to be resolved, current precipitation products (e.g., TRMM 3B42) remain substantially biased in mountainous regions like Nepal, where they do not outperform available rain gauges in predicting the frequency and intensity of areal rainfall (Müller and Thompson, 2013). A second source of error arises from the simplifying assumptions made about

streamflow recession that do not hold perfectly in the observed catchments. Because they describe  
455 the same watershed, the wet and dry recession parameters are assumed to be physically related. In  
Müller et al. (2014), the scale parameter of the non-linear seasonal recession ( $a$ ) is expressed as  
an explicit function of the two recession parameters ( $k$  and  $b$ ) for sufficiently short recession times,  
where power-law recessions can be approximated by exponential functions. We show in Supplemen-  
460 tary Materials (Section S2) that, although this approach provides more accurate estimates of  $a$  than  
would be obtained through spatial interpolation, estimation uncertainties remain, propagate through  
the model and result in prediction errors during the dry season.

#### 4.1.2 Applicability beyond Nepal

This study compares two specific methods on their ability to predict FDCs in the particular context  
of ungauged Nepalese basins. Results are thus not necessarily representative of the relative perfor-  
465 mance of process-based and statistical methods in general, particularly in regions where abundant  
field data allow more advanced statistical approaches to be implemented. Yet fundamentally, the  
statistical model relies on observed correlations, rather than assumptions about hydrologic mecha-  
nisms. Because FDC shapes are modeled non-parametrically, the approach is applicable to regions  
with highly variable catchment responses. However, prediction performance in ungauged basins is  
470 constrained by interpolation errors in the mean flow. This makes the method unsuitable for regions  
where the local determinants of mean flow (i.e. rainfall, evapotranspiration, glacial melt) cannot be  
accurately monitored at the catchment level. In contrast, a key advantage of the process-based model  
is its ability to exploit characteristics of the stochastic structure of rainfall that can be estimated from  
daily rainfall observations. The model is appropriate for regions where the spatial heterogeneity of  
475 runoff is driven by rainfall, and where the frequency and intensity of rainfall depths at the catchment  
level can be readily estimated (i.e. small catchments with numerous rain gauges, or places where  
satellite observations provide a good representation of rainfall statistics). Unlike rainfall, recession  
behavior arises from lumped and complex interactions between climate, vegetation and groundwa-  
ter processes that typically cannot be monitored in a spatially explicit manner. The process-based  
480 model is therefore inappropriate for regions where the hydrologic response of the catchment is the  
main source of runoff heterogeneity, or where the assumed recession behavior (in particular the  
relation between  $a$ ,  $k$  and  $b$ ) does not occur.

Conveniently, the appropriate implementation contexts for both methods appear to be comple-  
mentary, and the optimal method in a given region is determined by the driving source of runoff  
485 heterogeneity in the catchments. Ultimately, the performance of both methods is constrained by their  
ability to estimate their parameters in ungauged basins. This relation is apparent in Figure 4, where  
drops in prediction performances correspond to increases in the estimation uncertainty of model  
parameters. Under these conditions, the performance of each method is driven by the ability of the  
available observations to capture the variability of the model parameters. When interpolated from

neighboring gauges, uncertainties are governed by the interplay between the layout of the gauges and the spatial correlation range of the considered model parameter. When estimated from short observation records, accuracy is determined by the extent to which the available record is representative of the temporal variability of the parameter. These interactions between data availability and runoff variability are inherently local and will affect the determination of the most appropriate method for any given region.

## 4.2 Prediction under Change

Expected shifts in the frequency and intensity of Monsoon rainfall over Nepal only have a marginal impact on the streamflow distributions in the Chepe Kohla catchment, as shown by the numerical simulation presented in (Figure 5 (a, dashed curve)). Consequently, changes in rainfall regime do not appear to affect the performance of either model (Figure 5 (b)), unless they are significantly larger than expected. Climate change may nonetheless affect flow predictions elsewhere. It is therefore helpful to consider the conditions under which FDCs can be reliably predicted in a changing climate.

Although rainfall stationarity is an inherent assumption of the process-based approach, climate change can be incorporated by updating the relevant parameters to their future value to predict the (pseudo-) stationary future state of the system. The method accounts for otherwise confounding changes in the frequency and intensity of rainfall, which are expected in Nepal. By explicitly accounting for soil moisture dynamics and recession behavior, the model emulates the (causal) effect of rainfall on streamflow. As a result, the method reliably predicts the distribution of future streamflow, provided that governing flow generation processes are in line with the basic assumptions listed in Section 2.1.1.

In contrast, the statistical model is solely based on observed correlations, leading to two important sources of errors for predictions under change. First, the model only accommodates rainfall changes to the extent that the estimated statistical relation between rainfall and runoff is representative of local runoff coefficients. The model will not reliably predict future streamflows if runoff coefficients are strongly spatially heterogeneous, or if the cross sectional sample of gauges fails to capture important processes governing mean flow. This source of uncertainty appears to be significant in Nepal, as illustrated by the substantial bias in annual flow predictions on Figure 5(c). Secondly, the statistical model only considers the effect of *average* rainfall on *average* flow: the effect of rainfall *distribution* in streamflow *distribution* is ignored. As a result, the model cannot predict changes in the shape of FDC that are brought about by changing rainfall. The prediction performance of the statistical approach is therefore determined by the resilience of the flow regime, that is the extent to which streamflow distribution is affected by shifting rain signals (Botter et al., 2013): the method will perform poorly in catchments with non-resilient flow regimes. The Monte Carlo analysis presented in Supplementary Materials (Sections S3 and S4) shows that streamflow resilience in seasonally dry catchments depends on two distinct seasonal effects: a 'direct' effect driven by the ratio between

$\lambda_P$  and  $k$  during the wet season, and an 'indirect' effect during the dry season, when resilience is determined by the interplay between  $Q_0$  (i.e. wet-season rainfall) and  $b$ . In seasonally dry climates, we expect the statistical method to be most reliable in regions where wet seasons are short with limited total rainfall but persistent flow regimes, and where the recession behavior during the dry-season is close to linear.

Lastly, a key assumption in this study is that catchment response (in terms of low-flow or recession characteristics) is independent of climate. It is possible that shifts in climate have an effect on catchment response by affecting the partitioning of effective rainfall between storage and runoff. Although not quantitatively assessed in this study, we expect that this effect would negatively affect the performance of both approaches.

## 5 Conclusions

Stochastic, process-based models predicted the FDCs for ungauged catchments in Nepal well, with a performance that was comparable to that of statistical models. It suggests that in regions with globally representative gauge-densities, and under seasonally dry climates, the advantages of the statistical approaches relative to stochastic models noted in previous analyses (Blöschl et al., 2013) may not apply. Fundamentally, the performances of both approaches are strongly affected by the method chosen to estimate model parameters in ungauged basins, so this conclusion comes with the caveat that this study cannot be interpreted as a general benchmark to compare these approaches at a global level. Although we believe that the selected models are appropriate to compare process-based and statistical approaches for practical PUB application in Nepal, their relative performance may be different in other regions, where more abundant information on catchment characteristics allow more complex (and presumably more accurate) regionalization approaches to be applied. Thus, substantial research remains to be done to compare these approaches in other parts of the world, where locally appropriate methods should be carefully considered.

Nonetheless, this study finds a complementarity between the different sources of uncertainty in the stochastic and statistical methods. This suggests that model selection should be driven by a consideration of the main drivers of heterogeneity in any study catchment: Process-based models are advisable if climate is likely to be the main source of runoff heterogeneity. Conversely, statistical methods are more appropriate for regions with substantially different recession behaviors across catchments. These distinctions provide a potentially robust basis for model selection in any given application.

The results also suggest that the sensitivity of statistical approaches to changes in rainfall statistics is dependent on the 'resilience' of the flow regime as defined by Botter et al. (2013). Overall, the process-based models are more reliable in projecting FDCs into new rainfall regimes. This is particularly true for catchments characterized by a strong wet-season runoff and a rapid, strongly non-linear

hydrologic response, because their flow regime is particularly vulnerable to rainfall changes, making the assumptions of the statistical model inappropriate.

The excellent performance of both process-based and statistical models for the FDC and PDF in ungauged basins suggests that extending probabilistic analyses in such basins to also include flow-derived variables such as hydropower capacity (Basso and Botter, 2012) or ecological responses (Thompson et al., 2013a) may be feasible. While these prospects are enticing, we note that a model's ability to predict an FDC with high fidelity is not necessarily indicative of prediction performances on all derived stochastic properties. For instance, Dralle et al. (2015) demonstrate that the crossing properties of streamflow can be very poorly estimated by stochastic process-based models, even in applications where the same models predict the PDF of flow well. Further exploration of the potential opportunities and limitations afforded by use of probabilistic models in ungauged basins offers a promising avenue for future study.

## Appendix A: Process-based streamflow distribution model for seasonally dry climates

This appendix presents the analytical expression of FDC in seasonal climates derived in Müller et al. (2014). The approach assumes that rainfall can be represented as a marked Poisson process with exponentially distributed depths. Catchments are modeled as spatially homogenous linear reservoirs with linear evapotranspiration losses. Under these conditions, wet season streamflow can be represented as a gamma-distributed random variable (Botter et al., 2007b):

$$Q_w \sim \text{Gamma}(m, \alpha_Q^{-1})$$

with  $m = \lambda/k$  and  $\alpha_Q = \alpha_P k A$ , and where  $k$  is the linear recession constant,  $A$  the area of the contributing catchment and  $\alpha_P$  the mean intensity of wet season rainfall. The frequency  $\lambda$  of runoff events can be expressed as a function of the frequency ( $\lambda_P$ ) and intensity of rainfall (Botter et al., 2007b):

$$\lambda = \eta \frac{\exp(-\gamma) \gamma^{\frac{\lambda_P}{\eta}}}{\Gamma_L(\lambda_P/\eta, \gamma)} \quad (\text{A1})$$

where  $\Gamma_L(\cdot, \cdot)$  is the lower incomplete gamma function, and where  $\eta = ET/SSC$  and  $\gamma = SSC/\alpha_P$  are respectively the ratio between maximum evapotranspiration and soil storage capacity, and the ratio between soil storage capacity and mean rainfall intensity.

Dry season streamflow is modeled as a seasonal recession starting at the last discharge peak of the wet season. Because wet season streamflow is a gamma-distributed variable, streamflow at discharge peaks, and therefore the initial condition of the seasonal recession, is itself a gamma distributed variable (Müller et al., 2014):

$$Q_{peak} \sim \text{Gamma}(m + 1, \alpha_Q^{-1}).$$

Assuming a power law-relation between discharge and recession rate, the cumulative distribution function of dry season streamflow can be expressed as (Müller et al., 2014):

$$P_{Q_d}(Q) = \begin{cases} 1 + \frac{q_d^r \Gamma_1 - \alpha_Q^r \Gamma_2}{ar T_d \Gamma(m+1)}, & \text{if} \\ & Q > -(ar T_d)^{\frac{1}{r}} \\ & \text{and } r < 0 \\ 1 + \frac{q_d^r \Gamma_1 - \alpha_Q^r \Gamma_2}{ar T_d \Gamma(m+1)} + \frac{\alpha_Q^r \Gamma_4 + (Q^r - ar T_d) \Gamma_3}{ar T_d \Gamma(m+1)}, & \text{otherwise} \end{cases}$$

585 with

$$\begin{aligned} \Gamma_1 &= \Gamma_U(m+1, \alpha_Q^{-1} Q) \\ \Gamma_2 &= \Gamma_U(r+m+1, \alpha_Q^{-1} Q) \\ \Gamma_3 &= \Gamma_U(m+1, \alpha_Q^{-1} (Q^r + ar T_d)^{\frac{1}{r}}) \\ \Gamma_4 &= \Gamma_U(r+m+1, \alpha_Q^{-1} (Q^r + ar T_d)^{\frac{1}{r}}) \end{aligned}$$

590  $\Gamma(\cdot)$  and  $\Gamma_U(\cdot, \cdot)$  denote the complete and upper incomplete gamma functions;  $T_d$  is the duration of the dry season;  $r = 1 - b$  and  $a$  are the parameters of the non-linear recession, which are assumed stationary. Because they describe the same watershed, recession parameters for the wet and dry seasons are related. If power-law recessions can be approximated by an exponential function for sufficiently short recession times, we can express  $a$  as a function of  $k$  and  $b$  (Müller et al., 2014):

$$595 \quad a \approx \frac{\lambda}{-r} \left( e^{\frac{-r}{m}} - 1 \right) (\alpha_Q \cdot (m+1)) \quad (\text{A2})$$

The law of total probability can finally be used to combine seasonal streamflow distributions and derive the cumulative distribution function of streamflow for the whole year:

$$P_Q(Q) = \left( 1 - \frac{T_d}{365} \right) \cdot P_{Q_w}(Q) + \frac{T_d}{365} \cdot P_{Q_d}(Q) \quad (\text{A3})$$

The FDC for seasonally dry climates is finally obtained by plotting the streamflow quantiles  $Q$  600 against  $1 - P_Q(Q)$ , the complement of the the cumulative distribution function of streamflow.

*Acknowledgements.* The Swiss National Science Foundation are gratefully acknowledged for funding (M.F.M).

## References

- Alaouze, C. M.: Reservoir Releases to Uses with Different Reliability Requirements, *AWRA Water Resources Bulletin*, 25, 1163–1168, 1989.
- 605 Alternative Energy Promotion Center: Construction and Installation Manual for Micro hydropower Project Installers (In Nepalese), Government of Nepal, 2014.
- Arora, M., Goel, N., Singh, P., and Singh, R.: Regional flow duration curve for a Himalayan river Chenab., *Nordic hydrology*, 36, 193–206, 2005.
- Basso, S. and Botter, G.: Streamflow variability and optimal capacity of run-of-river hydropower plants, *Water*
- 610 *Resources Research*, 48, doi:10.1029/2012WR012 017, 2012.
- Bélisle, C. J.: Convergence theorems for a class of simulated annealing algorithms on  $\mathbb{R}^d$ , *Journal of Applied Probability*, 29, 885–895, 1992.
- Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., and Savenije, H.: *Runoff Prediction in Ungauged Basins: Synthesis across Processes, Places and Scales*, Cambridge University Press, 2013.
- 615 Botter, G., Peratoner, F., Porporato, A., Rodriguez-Iturbe, I., and Rinaldo, A.: Signatures of large-scale soil moisture dynamics on streamflow statistics across US climate regimes, *Water Resources Research*, 43, doi:10.1029/2007WR006 162, 2007a.
- Botter, G., Porporato, A., Rodriguez-Iturbe, I., and Rinaldo, A.: Basin-scale soil moisture dynamics and the probabilistic characterization of carrier hydrologic flows: Slow, leachingprone components of the hydrologic
- 620 response, *Water Resources Research*, 43, doi:10.1029/2006WR005 043, 2007b.
- Botter, G., Porporato, A., Rodriguez-Iturbe, I., and Rinaldo, A.: Nonlinear storage-discharge relations and catchment streamflow regimes, *Water resources research*, 45, doi: 10.1029/2008WR007 658, 2009.
- Botter, G., Basso, S., Rodriguez-Iturbe, I., and Rinaldo, A.: Resilience of river flow regimes, *Proceedings of the National Academy of Sciences*, 110, doi: 10.1073/pnas.1311920 110, 2013.
- 625 Brutsaert, W. and Nieber, J. L.: Regionalized drought flow hydrographs from a mature glaciated plateau, *Water Resources Research*, 13, doi: 10.1029/WR013i003p00 637, 1977.
- Castellarin, A., Galeati, G., Brandimarte, L., Montanari, A., and Brath, A.: Regional flow-duration curves: reliability for ungauged basins, *Advances in Water Resources*, 27, doi:10.1016/j.advwatres.2004.08.005, 2004a.
- Castellarin, A., Vogel, R., and Brath, A.: A stochastic index flow model of flow duration curves, *Water Re-*
- 630 *sources Research*, 40, doi:10.1029/2003WR002 524, 2004b.
- Castellarin, A., Botter, G., Hughes, D., Liu, S., Ouarda, T., Parajka, J., Post, D., Sivapalan, M., Spence, C., Viglione, A., and Vogel, R.: *Runoff prediction in ungauged basins: Synthesis across processes, places and scales. Chapter 7: Prediction of Flow Duration Curves in Ungauged Basins*, Cambridge University Press, 2013.
- 635 Ceola, S., Botter, G., Bertuzzo, E., Porporato, A., Rodriguez-Iturbe, I., and Rinaldo, A.: Comparative study of ecohydrological streamflow probability distributions, *Water Resources Research*, 46, doi: 10.1029/2010WR009 102, 2010.
- Chalise, S., Kansakar, S., Rees, G., Croker, K., and Zaidman, M.: Management of water resources and low flow estimation for the Himalayan basins of Nepal, *Journal of Hydrology*, 282, doi:10.1016/S0022–
- 640 1694(03)00 250–6, 2003.

Cheng, L., Yaeger, M., Viglione, A., Coopersmith, E., Ye, S., and Sivapalan, M.: Exploring the physical controls of regional patterns of flow duration curves—Part 1: Insights from statistical analyses, *Hydrology and Earth System Sciences*, 16, doi:10.5194/hess-16-4435-2012, 2012.

Chitrakar, P.: Micro-hydropower design aids manual, Small Hydropower Promotion Project (GTZ) and Mini-grid support program (Alternate Energy Promotion Center, Government of Nepal), 2004.

Coopersmith, E., Yaeger, M., Ye, S., Cheng, L., and Sivapalan, M.: Exploring the physical controls of regional patterns of flow duration curves—Part 3: A catchment classification system based on regime curve indicators, *Hydrology and Earth System Sciences*, 16, doi:10.5194/hess-16-4467-2012, 2012.

Department of Hydrology and Meteorology: Daily Streamflow and Precipitation Data,, Kathmandu, 2011.

Doulatyari, B., Betterle, A., Basso, S., Biswal, B., Schirmer, M., and Botter, G.: Predicting streamflow distributions and flow duration curves from landscape and climate, *Advances in Water Resources*, 83, doi:10.1016/j.advwatres.2015.06.013, 2015.

Dralle, D. N., Karst, N., and Thompson, S.: Dry season streamflow persistence in seasonal climates, *Water Resources Research*, In Press, doi:10.1002/2015WR017752, 2015.

Global Runoff Data Center: Global Runoff Data Base, Global Runoff Data Centre. Koblenz, Federal Institute of Hydrology (BfG), 2014.

Greene, W. H.: *Econometric analysis*, Pearson Education India, 2003.

HKH-FRIEND: Hindu Kush Himalayan - Flow Regimes from International Experimental and Network Data, UNESCO International Hydrological Programme, 2004.

Hughes, D. and Smakhtin, V.: Daily flow time series patching or extension: a spatial interpolation approach based on flow duration curves, *Hydrological Sciences Journal*, 41, 851–871, 1996.

Lambert, L. and Chitrakar, B.: Variation of potential evapotranspiration with elevation in Nepal, *Mountain Research and Development*, pp. 145–152, 1989.

Lazzaro, G., Basso, S., Schirmer, M., and Botter, G.: Water management strategies for run-of-river power plants: Profitability and hydrologic impact between the intake and the outflow, *Water Resources Research*, 49, doi:10.1002/2013WR014210, 2013.

Milly, P., Julio, B., Malin, F., Robert, M., Zbigniew, W., Dennis, P., and Ronald, J.: Stationarity is dead, *Science*, 319, doi:10.1126/science.1151915, 2008.

Montanari, A., Young, G., Savenije, H., Hughes, D., Wagener, T., Ren, L., Koutsoyiannis, D., Cudennec, C., Toth, E., and Grimaldi, S.: Panta Rhei—Everything Flows: Change in hydrology and society—The IAHS Scientific Decade 2013–2022, *Hydrological Sciences Journal*, 58, doi:10.1080/02626667.2013.809088, 2013.

Mu, X., Zhang, L., McVicar, T. R., Chille, B., and Gau, P.: Analysis of the impact of conservation measures on stream flow regime in catchments of the Loess Plateau, China, *Hydrological Processes*, 21, doi:10.1002/hyp.6391, 2007.

Müller, M. F. and Thompson, S. E.: Bias adjustment of satellite rainfall data through stochastic modeling: Methods development and application to Nepal, *Advances in Water Resources*, 60, doi:10.1016/j.advwatres.2013.08.004, 2013.

Müller, M. F. and Thompson, S. E.: TopREML: a topological restricted maximum likelihood approach to regionalize trended runoff signatures in stream networks, *Hydrology and Earth System Sciences*, 19, doi:10.5194/hess-19-2925-2015, 2015.

Müller, M. F., Dralle, D. N., and Thompson, S. E.: Analytical model for flow duration curves in seasonally dry climates, *Water Resources Research*, 50, doi:10.1002/2014WR015301, 2014.

Müller, M. F.: Bridging the Information Gap: Remote Sensing and Micro Hydropower Feasibility in Data-Scarce Regions, Doctoral Dissertation, University of California, Berkeley, CA, 2015.

Muneepeerakul, R., Azale, S., Botter, G., Rinaldo, A., and Rodriguez-Iturbe, I.: Daily streamflow analysis based on a two-scaled gamma pulse model, *Water Resources Research*, 46, doi: 10.1029/2010WR009286, 2010.

NASA Land Processes Distributed Active Archive Center (LP DAAC): ASTER GDEM v2, NASA Land Processes Distributed Active Archive Center (LP DAAC). ASTER L1B. USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls., 2011.

Nathan, R. and McMahon, T.: Evaluation of automated techniques for base flow and recession analyses, *Water Resources Research*, 26, 1465–1473, 1990.

Sauquet, E. and Catalogne, C.: Comparison of catchment grouping methods for flow duration curve estimation at ungauged sites in France, *Hydrology and Earth System Sciences*, 15, doi:10.5194/hess-15-2421-2011, 2011.

Schaefli, B., Rinaldo, A., and Botter, G.: Analytic probability distributions for snow-dominated streamflow, *Water Resources Research*, 49, doi:10.1002/wrcr.20234, 2013.

Searcy, J. and Hardison, C.: Double-mass curves. *Manual of hydrology: Part I, General surface water techniques*, US Geol Surv Water Supply Paper, 1960.

Shao, J. and Tu, D.: *The Jackknife and Bootstrap*, Springer-Verlag, New York, 2012.

Shrestha, D. P.: Assessment of soil erosion in the Nepalese Himalaya: a case study in Likhu Khola Valley, Middle Mountain Region, *Land Husbandry*, 2, 59–80, 1997.

Stokstad, E.: Scarcity of rain, stream gages threatens forecasts, *Science*, 285, doi: 10.1126/science.285.5431.1199, 1999.

Thompson, S., Levin, S., and Rodriguez-Iturbe, I.: Linking plant disease risk and precipitation drivers: a dynamical systems framework, *The American Naturalist*, 181, doi: 10.1086/668572, 2013a.

Thompson, S., Sivapalan, M., Harman, C., Srinivasan, V., Hipsey, M., Reed, P., Montanari, A., and Blöschl, G.: Developing predictive insight into changing water systems: use-inspired hydrologic science for the Anthropocene, *Hydrology and Earth System Sciences*, 17, doi: 10.5194/hess-17-5013-2013, 2013b.

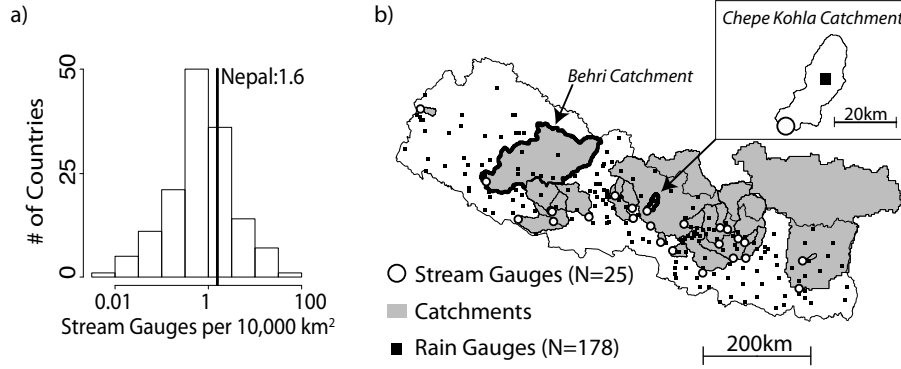
Turner, A. G. and Annamalai, H.: Climate change and the South Asian summer monsoon, *Nature Climate Change*, 2, doi:10.1038/nclimate1495, 2012.

Turner, A. G. and Slingo, J. M.: Subseasonal extremes of precipitation and active-break cycles of the Indian summer monsoon in a climate-change scenario, *Quarterly Journal of the Royal Meteorological Society*, 135, 549–567, 2009.

United States Geological Survey: USGS Threatened and Endangered Streamgages, [http://streamstats09.cr.usgs.gov/ThreatenedGages/ThreatenedGages\\_str.html](http://streamstats09.cr.usgs.gov/ThreatenedGages/ThreatenedGages_str.html), 2015.

Vogel, R. and Fennessey, N.: Flow-duration curves. I: New interpretation and confidence intervals, *Journal of Water Resources Planning and Management*, 120, 485–504, 1994.

- 720 Westerberg, I. K., Gong, L., Beven, K. J., Seibert, J., Semedo, A., Xu, C.-Y., and Halldin, S.: Regional water balance modelling using flow-duration curves with observational uncertainties, *Hydrology and Earth System Sciences*, 18, 2993–3013, 2014.

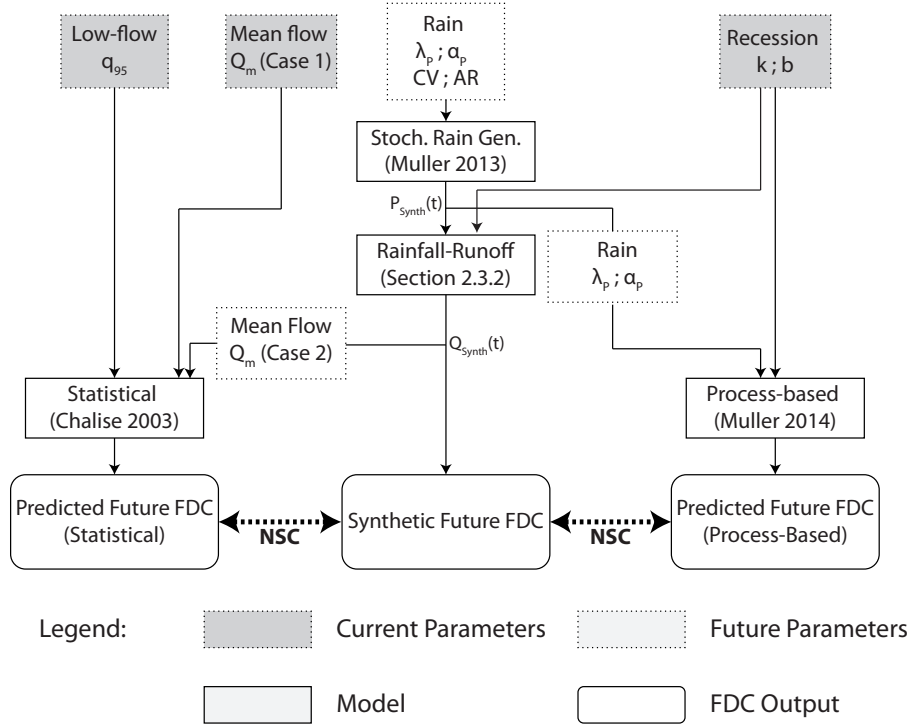


**Figure 1.** (a) Global histogram of the approximate spatial density of streamflow gauges by nation, represented by the sample of 8540 gauges indexed by the Global Runoff Data Center for 146 countries (Global Runoff Data Center, 2014). With a density of 1.6 gauges per 10,000  $km^2$ , Nepal falls close to the mode of the global distribution. (b) Location of the rain gauges, streamflow gauges and corresponding Nepalese catchments used in the analysis.

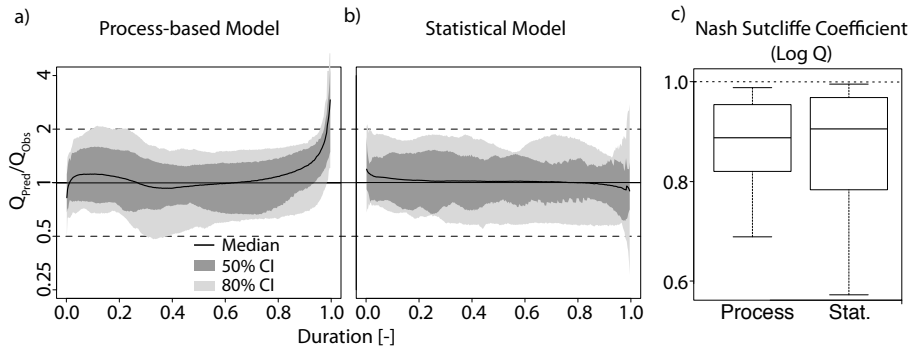
**Table 1.** Catchment characteristics. Median values and interquartile distances (IQD) are given for the whole sample of 25 gauges. The table also presents characteristics of the Chepe Kohla watershed considered in the analysis as a case study.

	Streamflow:		Topography:				Climate:						Recession:			
	$Q_m$	$q_{95}$	$N_y$	$A$	$z_m$	$z_M$	$P_y$	$T_{\text{mons}}$	$\lambda_P$	$\alpha_P$	$AR$	$CV$	$ET$	$k$	$b$	
All gauges																
<i>Median</i>	76.1	0.14	22	1355	481	5209	1952	99	0.71	18.8	0.29	0.92	2.5	0.17	2.38	
<i>Min</i>	7.3	0.06	10	130	116	1913	1260	88	0.54	12.1	0.09	0.61	0.40	0.07	1.99	
<i>Max</i>	1462.4	0.25	41	32817	1641	8369	4030	152	0.91	33.0	0.51	1.53	3.27	0.32	2.99	
Chepe Kohla	23.0	0.14	31	277	475	4711	3050	100	0.84	26.5	0.09	1.03	2.1	0.20	2.41	

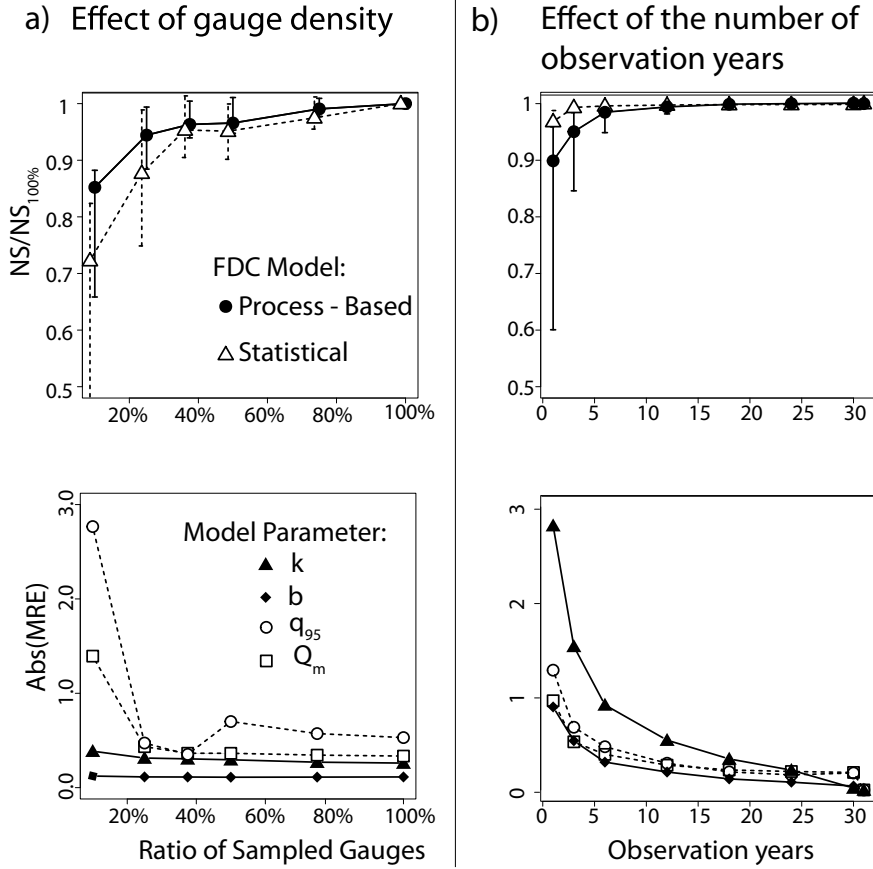
$Q_m$  is mean annual flow in  $m^3 s^{-1}$ ;  $q_{95}$  is the 95<sup>th</sup> flow percentile normalized by  $Q_m$ ;  $N_y$  indicates the number of observation years;  $A$  is the catchment area in  $km^2$ ;  $z_m$  and  $z_M$  are respectively the minimum and maximum elevation of the basins meters;  $P_y$  is mean precipitation in  $mm/y$ ;  $T_{mons}$  is the estimated duration of the monsoon in days;  $\lambda_P$  is rainfall frequency during the monsoon (in  $d^{-1}$ );  $\alpha_P$  is mean rainfall intensity in  $mm/d$ ;  $AR$  is the first-order autocorrelation coefficient of rainfall occurrence ( $AR = 0$  if rainfall follows a Poissonian process),  $CV$  is the coefficient of variation of rainfall intensity on rainy days ( $CV=1$  if depths are exponentially distributed);  $ET$  [mm/d] is the reference evapotranspiration during the rainy season Lambert and Chitrakar (1989);  $k$  is the linear recession constant estimated during the monsoon (in  $d^{-1}$ ) and  $b$  is the non linear exponent of the seasonal recession. A soil moisture capacity of 16mm are assumed throughout the country (Müller et al., 2014).



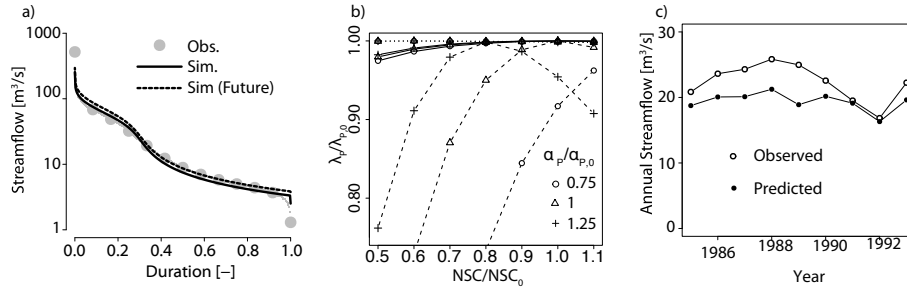
**Figure 2.** Numerical simulation analysis to assess predictions under change. Future rainfall characteristics (frequency  $\lambda_P$ , mean intensity  $\alpha_P$ , auto-correlation coefficient  $AR$  and coefficient of variation  $CV$ ) are determined according to expected changes in rain regimes in Nepal (see Section 2.3.2), and fed into a stochastic rainfall generator. The resulting 1000 years of synthetic daily rainfall values ( $P_{Synth}(t)$ ) are fed into a rainfall-runoff model that simulates the processes described in Section 2.1.1. The rainfall-runoff model uses current recession, soil and evapotranspiration conditions observed at the Chepe Kohla catchment. The resulting 1000 years of synthetic daily flow values ( $Q_{Synth}(t)$ ) are then reordered to construct an empirical synthetic (future) FDC, which was compared (in terms of the Nash Sutcliffe Coefficient) to modeled FDCs predicted by the statistical and process-based models. The process-based model admits *current* recession conditions, but *future* estimates for rainfall frequency ( $\lambda_P$ ) and mean intensity ( $\alpha_P$ ). Note that unlike the numerically generated empirical FDC, the process-based model assumes Poissonian rainfall with exponentially distributed depths, that is  $CV = 1$  and  $AR = 0$ . Current low flow characteristics ( $q_{95}$ ) are fed into the statistical model, as well as the current or future (i.e. computed from synthetic streamflow time series) mean flow, depending on the extent to which mean rainfall is an unbiased predictor of mean flow (Cases 1 and 2 described in Section 2.3.2).



**Figure 3.** Flow duration curve prediction performance in ungauged basins. The error duration curves of the leave-one-out cross-validation analysis using the process-based and statistical models are presented in panels (a) and (b) respectively. Relative errors are plotted on a log scale in order to allow the graphs to be balanced on the y-axis: a relative prediction error of 2 (the model predicts double the observed value) is at the same distance from  $y=1$  (perfect prediction) than a relative error of  $1/2$  (the model predicts half the observed value). Durations are plotted on the x-axis, with  $x=0$  and  $x=1$  for the highest and lowest flow quantiles respectively. Panel (c) shows box plots of Nash Sutcliffe coefficients computed from log-transformed flow quantiles.



**Figure 4.** Sensitivity of models to data scarcity. (a) Cross-validation analysis showing the sensitivity of both models to a decreasing number of calibration gauges. (b) Resampling analysis of streamflow observations in the Chepe Kohla (N=10,000) catchment showing the effect of the number of observation years. In panels (a) and (b), the effects on FDC prediction performances (top) are shown by plotting the ratio of calibration gauges sampled (or the number of observation years) against the relative Nash Sutcliffe coefficient (with the NSC for the full set of available data as reference). The plot shows the median value for all iterations, and the error bars indicate the interquartile (25 - 75%) range. The prediction uncertainties of model parameters (bottom) are given in absolute values of relative prediction errors.



**Figure 5.** Sensitivity of models to changes in the precipitation regime. (a) Empirical and simulated flow duration curves at Chepe Kohla. The simulated FDC obtained from the stochastic rainfall generator and the bucket watershed model (solid) reproduces the empirical FDC constructed from the observed streamflow well (grey dots). Rainfall changes expected in Nepal ( $\alpha_P/\alpha_{P,0} = 1.2$ ,  $\lambda_P/\lambda_{P,0} = 0.98$ ) do not have a substantial influence on the simulated flow distribution (dashed).  $\alpha_P$  and  $\lambda_P$  designate the mean depth and frequency of wet season rainfall, respectively. (b) Sensitivities to relative changes in rainfall frequency and intensity over the Chepe Kohla catchment. The performance of the process-based model is not affected by rainfall changes (dotted). The sensitivity of the statistical model depends on its ability to predict changes in mean flow from annual rainfall. The model is highly sensitive to rain changes if average streamflow cannot be predicted (dashed), and is robust to moderate changes if average flow is perfectly predicted (solid). (c) The linear regression of the statistical model underestimates annual flows at the Chepe Kohla when using a cross-sectional sample (25 gauges) to estimate the local relation between average rainfall and average runoff.