Journal of Hydrology xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

A hierarchical Bayesian GEV model for improving local and regional flood quantile estimates

Carlos H.R. Lima [a,*], Upmanu Lall [b], Tara Troy [c], Naresh Devineni [d]

[a] Civil and Environmental Engineering, University of Brasilia, Brazil
[b] Earth & Environmental Engineering, Columbia University, New York, USA
[c] Civil and Environmental Engineering, Lehigh University, Bethlehem, USA
[d] Department of Civil Engineering, NOAA-Cooperative Remote Sensing Science and Technology Center, City University of New York, New York, USA

A R T I C L E I N F O

Article history:
Received 23 March 2016
Received in revised form 1 July 2016
Accepted 27 July 2016
Available online xxxx
This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Antonino Cancelliere, Associate Editor

Keywords:
Floods
GEV distribution
Hierarchical Bayesian
Regional flood frequency analysis
Scaling

A B S T R A C T

We estimate local and regional Generalized Extreme Value (GEV) distribution parameters for flood frequency analysis in a multilevel, hierarchical Bayesian framework, to explicitly model and reduce uncertainties. As prior information for the model, we assume that the GEV location and scale parameters for each site come from independent log-normal distributions, whose mean parameter scales with the drainage area. From empirical and theoretical arguments, the shape parameter for each site is shrunk towards a common mean. Non-informative prior distributions are assumed for the hyperparameters and the MCMC method is used to sample from the joint posterior distribution. The model is tested using annual maximum series from 20 streamflow gauges located in an 83,000 km^2 flood prone basin in Southeast Brazil. The results show a significant reduction of uncertainty estimates of flood quantile estimates over the traditional GEV model, particularly for sites with shorter records. For return periods within the range of the data (around 50 years), the Bayesian credible intervals for the flood quantiles tend to be narrower than the classical confidence limits based on the delta method. As the return period increases beyond the range of the data, the confidence limits from the delta method become unreliable and the Bayesian credible intervals provide a way to estimate satisfactory confidence bands for the flood quantiles considering parameter uncertainties and regional information. In order to evaluate the applicability of the proposed hierarchical Bayesian model for regional flood frequency analysis, we estimate flood quantiles for three randomly chosen out-of-sample sites and compare with classical estimates using the index flood method. The posterior distributions of the scaling law coefficients are used to define the predictive distributions of the GEV location and scale parameters for the out-of-sample sites given only their drainage areas and the posterior distribution of the average shape parameter is taken as the regional predictive distribution for this parameter. While the index flood method does not provide a straightforward way to consider the uncertainties in the index flood and in the regional parameters, the results obtained here show that the proposed Bayesian method is able to produce adequate credible intervals for flood quantiles that are in accordance with empirical estimates.

1. Introduction

Floods are one of the most interesting hydrologic phenomenon. The understanding of their spatio-temporal patterns is of practical importance for society and their scientific studies along the years have contributed to the advance of other fields in science such as theoretical statistics (e.g. the introduction of the extreme value theory by Gumbel (1958)) and climate (e.g. the study of floods in a hydroclimate perspective introduced by Hirschboeck (1988)). The empirical study of local and regional flood occurrence has been the main tool to help understand the spatial and temporal patterns of such extreme events and to provide information (e.g. flood quantiles) for flood risk management. A probability model is usually fit to block maximum or peak over threshold (POT) streamflow series in order to obtain flood quantiles for a specified frequency or return period (see for instance Stedinger et al. (1993)). For ungauged basins, with limited records (prediction in ungauged basins - PUB, see Gupta et al., 2007) regional flood frequency analysis(see, for instance Stedinger et al., 1993 and the references

* Corresponding author at: Civil and Environmental Engineering, University of California, Irvine, USA.

E-mail addresses: chrlima@unb.br (C.H.R. Lima), ula2@columbia.edu (U. Lall), tara.troy@lehigh.edu (T. Troy), ndevineni@ccny.cuny.edu (N. Devineni).

http://dx.doi.org/10.1016/j.jhydrol.2016.07.042
0022-1694/© 2016 Elsevier B.V. All rights reserved.

therein) with data from other basins are used to estimate the model parameters. The traditional assumption of stationarity in such analyses has been relaxed by explicitly modeling the time variation of the parameters (Kwon et al., 2008; Villarini et al., 2010; Villarini et al., 2009; Cunderlik and Burn, 2003; Cheng et al., 2014; Towler et al., 2010; Lima et al., 2015).

A number of challenges are presented when estimating flood frequency using Generalized Extreme Value (GEV) models. Small sample sizes often lead to an unstable likelihood function resulting in estimates with high bias and variance. For instance, Martins and Stedinger (2000) show in a simulation study how the shape parameter of a generalized extreme value (GEV) distribution is highly biased for small samples. Merz and Blöschl (Merz and Blöschl, 2008a; Merz and Blöschl, 2008b) proposed the term flood frequency hydrology in lieu of flood frequency statistics to highlight the importance of adding to the inference problem the best use of wide hydrological information and hydrological reasoning, which they divided into three types: temporal, spatial and causal information. Temporal information includes paleoflood (Benito and Thorndycraft, 2005) and historical (Reis and Stedinger, 2005) data. Spatial information refers to data from neighboring sites which have similar physical/hydrological/climate characteristics (Merz and Blöschl, 2005) and by making use of them in the inference process (Lima and Lall, 2010). Finally, causal information is the extra information that can be potentially added to the inference routine from the knowledge of the local flood dynamics (Viglione et al., 2013). The choice of parameter estimation method (e.g. L-moments, generalized maximum likelihood, Bayesian methods, etc) has been also subject of much research (Katz et al., 2002; Kwon et al., 2008; Madsen and Rosbjerg, 1997; Parent and Bernier, 2003; Martins and Stedinger, 2000; Rao and Hamed, 2000; Renard, 2011; Renard et al., 2013; Smith, 1989; Smith et al., 2011; Hosking and Wallis, 1997).

Uncertainty quantification and modeling continues to be a challenge (Sivapalan et al., 2003). In this paper, we use spatial information to *reduce predictive uncertainty* (Sivapalan et al., 2003) through a multilevel, hierarchical Bayesian (HB) flood frequency model. The scaling theory of flood statistics, distribution parameters and flood events with basin drainage area (Sivapalan et al., 2002; Gupta et al., 2007), which has found several applications in hydrology (Gupta and Dawdy, 1995; Gupta et al., 1994; Gupta and Waymire, 1990; Lima and Lall, 2010; Ishak et al., 2011; Villarini and Smith, 2010; Villarini et al., 2011) is used. The model is tested using annual maximum series from 20 streamflow gauges located in the Rio Doce basin in southeastern Brazil. A comparison with flood quantile estimates from the traditional GEV distribution and from the index flood method (Hosking and Wallis, 1997) is made.

The balance of the article is organized as follows. In the next section we describe the structure of the model. In Section 3, we present the results from the application of the model to a flood prone region in Brazil. A summary is presented in Section 4.

2. A hierarchical Bayesian Generalized Extreme Value (GEV) model

2.1. Regional analysis

In ungauged basins, flood quantiles are estimated using regional flood frequency analysis, where data from gauged sites are used to estimate the flood attributes in ungauged sites. The index flood method which dates back to the 1960's (Dalrymple, 1960; Benson, 1962) is one of the most used. L-moments estimates (see for instance Hosking and Wallis, 1997) are typically used to guide parameter estimation and model choice. For a pre-identified homogeneous region, the key idea is to estimate regional parameters based on the local distribution parameters estimated from standardized series of annual maximum for the gauged sites. The index flood, usually taken as the average or median annual maximum for each site, is used as the index to standardize the annual maximum series and correlations of it with basin attributes to obtain a flood index for each ungauged site, which is then used to estimate the flood quantiles in association with the regional growth curve (we refer the reader to Hosking and Wallis (1997) for more details about the index flood procedure). However, a chain of uncertainties in model, parameters and input data are not directly or routinely transferred to the flood quantile estimates and this is one of the main limitations of the index flood method. The hierarchical Bayesian model proposed here addresses part of this issue.

Alternatives to overcome the drawbacks of the index flood method, for instance the uncertainties in parameters and violations of the homogeneity and independence assumptions (Katz et al., 2002), will generally rely on estimating distribution parameters or flood quantiles for gauged sites and model their spatial variability using as covariates specific physical and/or hydrometeorological attributes (e.g. drainage area, elevation, slope, precipitation, etc) of the corresponding basins, which in turn will provide a framework to estimate the desired flood statistics at the ungauged site (e.g. Kuczera, 1982; Lettenmaier and Potter, 1985; Hosking and Wallis, 1997; Leclerc and Ouarda, 2007; Adamowski, 2000; Merz and Blöschl, 2005; Gaume et al., 2010; Morrison and Smith, 2002). The GEV, Log-Pearson Type III, log-normal and generalized Pareto have been the most common distributions used in such studies. Standard errors for the estimates provided by the maximum likelihood method are often based on asymptotic assumptions and well known to produce unrealistic results (Katz et al., 2002), particularly when estimating flood quantiles for return periods beyond the length of observed data. Bayesian inference is an appropriate tool to deal with parameter uncertainty and to aggregate extra (e.g. historical) information. The application of Bayesian inference in hydrology has experienced an extraordinary growth in the last decade due to the advance of computational tools and methods. We refer the reader to the works of Gaume et al. (2010), Haddad et al. (2012), Kwon et al. (2008), Liang et al. (2011), Lima and Lall (2010), Martins and Stedinger (2000), Parent and Bernier (2003), Reis and Stedinger (2005), Renard (2011), Renard et al. (2006), Renard and Lall (2014), Viglione et al. (2013), Steinschneider and Lall (2015), Sun et al. (2015), Sun and Lall (2015), Halbert et al. (2016) and the references therein for some recent examples of Bayesian methods in flood frequency analysis. Here we undertake the physical/empirical well known scaling property of flow moments with drainage area (Gupta et al., 2007) to derive regional distribution parameters in a hierarchical Bayesian inference scheme, which will allow a reduction in uncertainty for local estimates as well as will propagate such uncertainties to flood quantile estimates in both local and regional analysis - a desired information in most flood frequency studies.

2.2. Local flood frequency analysis

The model presented in this paper considers the GEV distribution, which is theoretically appealing for use in block maximum data and is popular for flood frequency analysis. Let us assume then that $Q_{i,j}$ represents the annual maximum peak flow for site j in year i. Under some regularity and stationary conditions, extreme value theory (see for instance Coles, 2001) postulates that $Q_{i,j}$ follows a GEV distribution g with parameters μ_j (location), σ_j (scale) and ξ_j (shape):

$$Q_{i,j} \sim g(\mu_j, \sigma_j, \xi_j) \quad (1)$$

where the cumulative distribution function of g is given by:

$$G(x) = \exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{2}$$

Given a time series record of $Q_{i,j}$ for different locations, the parameters μ_j, σ_j and ξ_j can be estimated independently for each site using, for instance, standard procedures of maximum likelihood estimation, which will provide under asymptotic assumptions confidence intervals for such parameters (e.g. the delta method in chapter 3 of Coles (2001)). The limitations of point and interval maximum likelihood estimates for μ_j, σ_j and ξ_j are well known in the literature and alternative methods have been proposed along the years (e.g. the L-moments of Hosking and Wallis (1997) and the generalized maximum-likelihood of Martins and Stedinger (2000)).

2.3. The hierarchical Bayesian approach

In order to include more information when estimating the GEV model parameters, we consider the scaling theory and assume that all sites have identical location and scale parameters except for a scale factor. In this case, the scale factor is the drainage area and it is assumed that a log-log linear relationship of location and scale parameters and the drainage area exists, which has theoretical and empirical support (Gupta and Dawdy, 1995; Gupta et al., 1994; Gupta et al., 2007; Gupta and Waymire, 1990; Lima and Lall, 2010; Ishak et al., 2011; Villarini et al., 2011; Villarini and Smith, 2010; Morrison and Smith, 2002; Northrop, 2004). We develop the idea through independent normal prior distributions for μ_j and σ_j that reflects our knowledge about them before observing the data:

$$\log(\mu_j) \sim N(\alpha_1 + \alpha_2 \cdot \log(A_j), \tau_\mu^2) \tag{3}$$

$$\log(\sigma_j) \sim N(\beta_1 + \beta_2 \cdot \log(A_j), \tau_\sigma^2) \tag{4}$$

where A_j is the basin drainage area for site j and $\alpha_1, \alpha_2, \beta_1, \beta_2, \tau_\mu^2$ and τ_σ^2 are the so-called hyperparameters.

The scaling of the shape parameter ξ_j with the drainage area is not supported by empirical evidence (e.g. Morrison and Smith, 2002; Northrop, 2004; Villarini and Smith, 2010). Regional shape parameter estimates obtained from simple averages from the at-site estimates have been employed in the literature (e.g. Stedinger and Lu, 1995) in agreement with scaling theory (Burlando and Rosso, 1996; Gupta and Waymire, 1990; Morrison and Smith, 2002). Hence, we will assume as prior distribution that at-site shape parameters are drawn from a normal distribution with a common mean across all sites:

$$\xi_j \sim N(\overline{\xi}, \tau_\xi^2) \tag{5}$$

with $\overline{\xi}$ and τ_ξ^2 as hyperparameters. Note that such a prior is an intermediate model between separate or non-pooled estimates (i.e. individual at-site estimates) and a fully-pooled estimate (the shape parameter is the same across all sites). This is similar to the approach adopted in Gelman et al. (2013) to estimate the mean treatment effect in the eight schools problem.

For the hyperparameters presented in Eqs. (3)–(5), we will assume no prior knowledge and hence will choose independent uniform hyperprior distributions:

$$p(\alpha_1, \alpha_2, \beta_1, \beta_2, \overline{\xi}) \propto 1 \tag{6}$$

and

$$p(\log(\tau_\mu, \tau_\sigma, \tau_\xi)) \propto 1 \tag{7}$$

The joint posterior distribution of all model parameters $\theta = [\mu\, \sigma\, \xi\, \alpha_1\, \alpha_2\, \beta_1\, \beta_2\, \overline{\xi}\, \tau_\mu\, \tau_\sigma\, \tau_\xi]$ is determined by combining the likelihood function (or sampling distribution) $p(\mathbf{q}|\theta)$ with the prior distributions $p(\theta)$ presented above under the Bayesian rule:

$$p(\theta|\mathbf{q}) = \frac{p(\mathbf{q}|\theta)\cdot p(\theta)}{p(\mathbf{q})} \propto p(\mathbf{q}|\theta)\cdot p(\theta) \tag{8}$$

with $\mathbf{q}$ as the observed streamflow data, $p(\mathbf{q})$ the marginal likelihood of observations acting as a normalization constant to ensure that $p(\theta|\mathbf{q})$ has an integral equals to 1 and

$$p(\mathbf{q}|\theta) = \prod_{j=1}^{J}\prod_{i=1}^{n_j} g(q_{ij}|\mu_j, \sigma_j, \xi_j) \tag{9}$$

where q_{ij} is the observed annual maximum flow for site j at year i, n_j the recording length for site j and J the total number of sites in the region of interest with available data.

Substituting then Eq. (9) and (3)–(7) into Eq. (8) yields the joint posterior distribution of the parameter set θ:

$$p(\theta|\mathbf{q}) \propto \prod_{j=1}^{J}\prod_{i=1}^{n_j} g(q_{ij}|\mu_j, \sigma_j, \xi_j) \cdot N(\log(\mu_j)|\alpha_1 + \alpha_2 \cdot \log(A_j), \tau_\mu^2) \cdot N(\log(\sigma_j)|\beta_1 + \beta_2 \cdot \log(A_j), \tau_\sigma^2) \cdot N(\xi_j|\overline{\xi}, \tau_\xi^2) \tag{10}$$

Since the analytical integration of this joint posterior distribution (Eq. (10)) over all parameters is not tractable, we sample from it using the well known Markov Chain Monte Carlo (MCMC) method as implemented in the software OpenBUGS (**B**ayesian inference **U**sing **G**ibbs **S**ampling, Lunn et al. (2000)), which automatically selects an appropriate sampling algorithm (e.g. Metropolis Hasting) based on the combination of the likelihood function and the prior distributions. OpenBUGS is called from the R software using the R2WinBUGS package (Sturtz et al., 2005). The convergence of the posterior distribution was evaluated visually based on the mixture of 3 chains and based on the convergence diagnostic $\widehat{R}$ as suggested in Gelman et al. (2013). The R codes developed for the proposed Bayesian model are freely available as supplementary material.

Local flood quantiles and their associated uncertainties can be estimated by drawing samples of μ_j, σ_j and ξ_j from the joint posterior distribution using (1). Estimation of flood quantiles and their associated uncertainties for ungauged sites (regional flood frequency analysis) can be done by obtaining the predictive posterior distribution of the location, scale and shape parameters from Eqs. (3)–(5) using the drainage area A of the ungauged site and the joint posterior distribution of $\alpha_1, \alpha_2, \beta_1, \beta_2, \overline{\xi}, \tau_\mu, \tau_\sigma$ and τ_ξ. The sampling values of the GEV parameters can then be used in Eq. (1) to obtain any desired flood quantile for the ungauged site.

3. Application

3.1. Streamflow data and region of study

The model proposed here is applied to a series of mean daily flow for 20 streamflow gauges (details in Table 1) located in the flood-prone Rio Doce basin (southeaster Brazil, see Fig. 1) provided by the Brazilian National Water Agency (ANA). Data availability for the entire streamflow gauges is shown in Fig. 2 and the record length for each gauge is displayed in Table 1. The streamflow data spans the period from 1938 to 2012, but not all sites have a complete record, with some sites with very few data (e.g. sites # 1 and 4), which makes the estimation of model parameters in the traditional approach challenging, highlighting the need for the kind of model presented in this paper. For each series, the annual maximum of the daily flows was identified and this derived data set

Table 1
Location, drainage area and series length of streamflow gauges used in this work.

Site	Latitude (S)	Longitude (W)	Drainage area (km^2)	Series length (years)
1	−19.77	−42.48	16,100	3
2	−19.78	−42.07	162	28
3	−19.33	−42.37	24,200	25
4	−19.53	−40.63	76,400	7
5	−18.88	−41.95	40,500	43
6	−19.41	−40.06	82,100	27
7	−19.52	−42.64	5270	26
8	−19.19	−42.42	10,200	38
9	−19.77	−43.03	3060	35
10	−20.38	−42.90	6230	37
11	−19.34	−41.25	61,200	21
12	−19.49	−41.16	8720	72
13	−18.97	−41.64	55,100	38
14	−18.57	−41.92	9770	38
15	−19.99	−42.67	10,100	31
16	−19.90	−42.58	12,800	10
17	−19.33	−42.36	24,200	44
18	−19.32	−41.25	61,200	45
19	−20.65	−43.83	77	23
20	−19.93	−43.17	1160	75

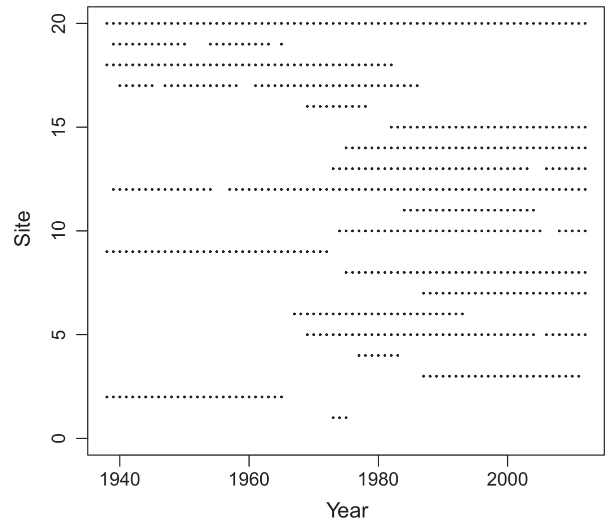

Fig. 2. Data availability for the entire set of streamflow gauges (Table 1) used in this work.

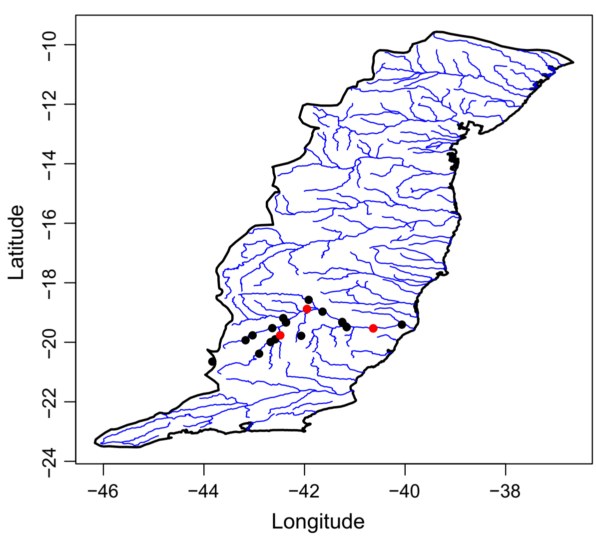

Fig. 1. Location of Rio Doce basin in southeaster Brazil. The streamflow gauges used to estimate the model parameters are indicated by the black dots, while out-of-sample gauges (sites # 1, 4 and 5 in Table 1) used in the cross-validation analysis are shown in red. Note that some sites are very close to each other (sites 3 and 17 and sites 11 and 18, see location in Table 1) and overlap in this figure.

was used for the application of the model. Drainage areas range from 77 to 82100 km^2.

3.2. Results

In order to verify whether a scaling law with drainage area is reasonable for the GEV parameters, we examine the scatter plot of the logarithm of the ordinary maximum likelihood estimates (MLE) of the GEV parameters versus the logarithm of the drainage area in Fig. 3 for a sub-set (i.e. training set) of 17 sites (black dots in Fig. 1). Sites # 1, 4 and 5 in Table 1 are chosen as cross-validation sites for the Bayesian regional flood frequency analysis and their data are not included in the Bayesian inference process. Both location μ and scale σ parameters present a well defined log-log linear relationship with the drainage area, with different intercept (α_1 and β_1) and slope (α_2 and β_2) parameters. As conjectured earlier, the shape parameter ξ does not present any strong evidence of scaling with the drainage area, with a surprising negative slope (−0.058) from the ordinary least square regression that is not statistically significant at the 5% significance level (p-value = 0.20). It is also interesting to note the presence of an outlier data in the bottom of the rightmost panel ($\hat{\xi} = -1.078$) that refers to training site # 16, which has only 10 years of data available for parameter estimation (see site # 16 in Fig. 2 and Table 1) and thus leads to a poor MLE of the shape parameter.

A comparison of the MLE and of the HB posterior distributions of model parameters, including the 95% credible interval and the Bayesian point estimates based on the median of the posterior distributions, is shown in Fig. 4. Notice that almost all HB credible intervals include the MLE, except for the shape parameter of site # 16, which, as discussed above, has a limited amount of data (10 years), and may not give a reliable estimate when the ML estimation process is used. This shows the advantage of the HB model over the ordinary MLE, which plausibly is underestimating ξ for site # 16, while the HB estimate lies close to the average estimate from all testing sites, exhibiting a strong degree of pooling. The range of the HB credible interval for the shape parameter suggests that most likely the entire set of streamflow sites will have a heavy tail distribution ($\xi > 0$). It is also worth mentioning that, besides the scale factor, as the number of data points increases, the uncertainty in the estimates as evidenced by the size of the credible interval tends to reduce (see for instance the estimates for sites # 12 and 20).

Estimation of flood quantiles for three randomly chosen sites (# 14, 16 and 18) based on the MLE (blue curve) and HB estimates (black curve) and their associated uncertainties are displayed in Fig. 5. For comparison purposes, empirical estimates of the flood quantiles based on the Weibull formula are also included in Fig. 5 (black dots). For sites 14 and 18 (left and right panels in Fig. 5) the lower bound of the uncertainty associated with the ML flood quantiles tends to reduce as the return period increases, which drastically augments the confidence interval, showing the difficulty in using this methodology to estimate confidence intervals for large return periods. On the other hand, the proposed method produces credible intervals that are narrower than the correspondent confidence intervals and still contain most empirical estimates. For site 16 (middle panel), due to the limited amount of data (10 years), the confidence interval produced by MLE is degraded and very thin, and cannot be seen in the figure. For site

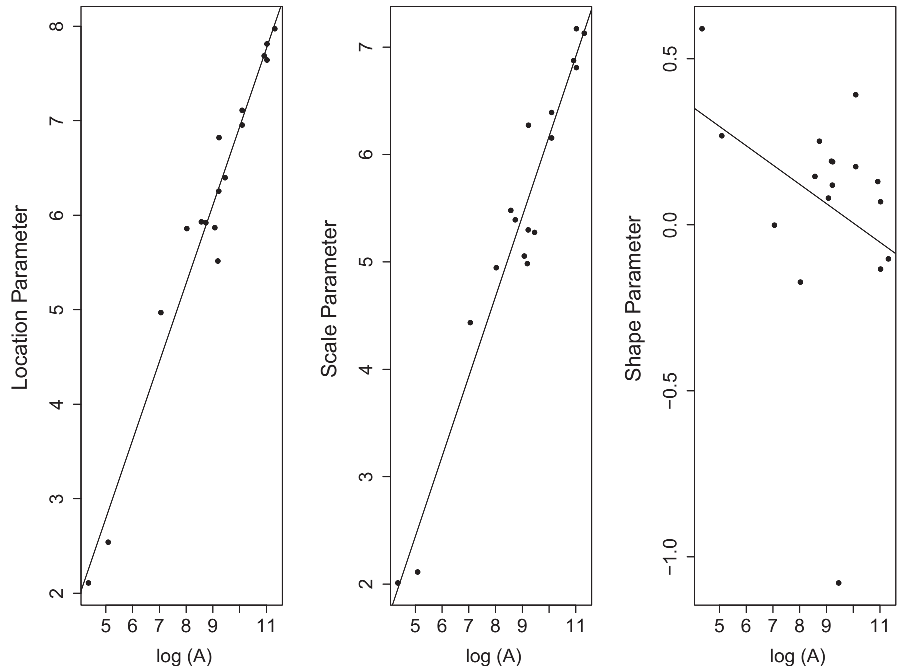

Fig. 3. Logarithm of GEV location μ (left panel), scale σ (middle panel) and shape ξ (right panel) parameter estimates versus logarithm of drainage area for the training sites. The black line shows an ordinary least square regression.

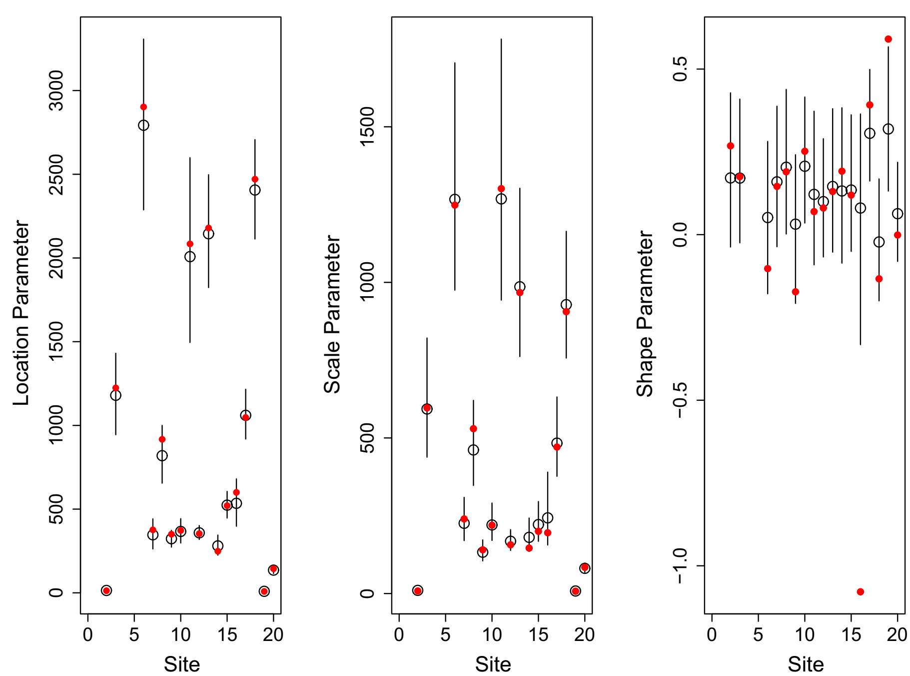

Fig. 4. Maximum likelihood estimates (red dots) of location (left panel), scale (middle panel) and shape (right panel) GEV parameters versus 95% hierarchical Bayesian credible intervals (vertical lines) for the same parameters, including the point estimates (black circles) based on the median of the posterior distributions. Note that data from the out-of-sample sites # 1, 4 and 5 (testing sites) are not used during the Bayesian inference and thus their parameter estimates are not included here.

18 (right panel of Fig. 5), the Bayesian credible interval visually appears wider than the frequentist confidence interval, but for the majority of sites (the quantile estimates for all training sites are available as supplementary material) the credible intervals are visually narrower than the confidence intervals. In fact, the most substantial gain of the proposed Bayesian methodology over traditional MLE in local flood frequency analysis is seen in estimating flood quantiles and associated uncertainties for return periods beyond the length of the streamflow data record.

In order to evaluate the performance of the proposed HB model for regional flood frequency analysis, we estimate flood quantiles for three out-of-sample sites (# 1, 4 and 5) whose streamflow data were not included in the Bayesian inference. We would call this procedure as cross-validation in space and it was done as explained in Section 2.3, given the drainage areas of the three sites and the Bayesian posterior distributions of the shape parameter and of the scaling parameters for μ and σ. The cross-validation results are displayed in Fig. 6. They show estimates of flood

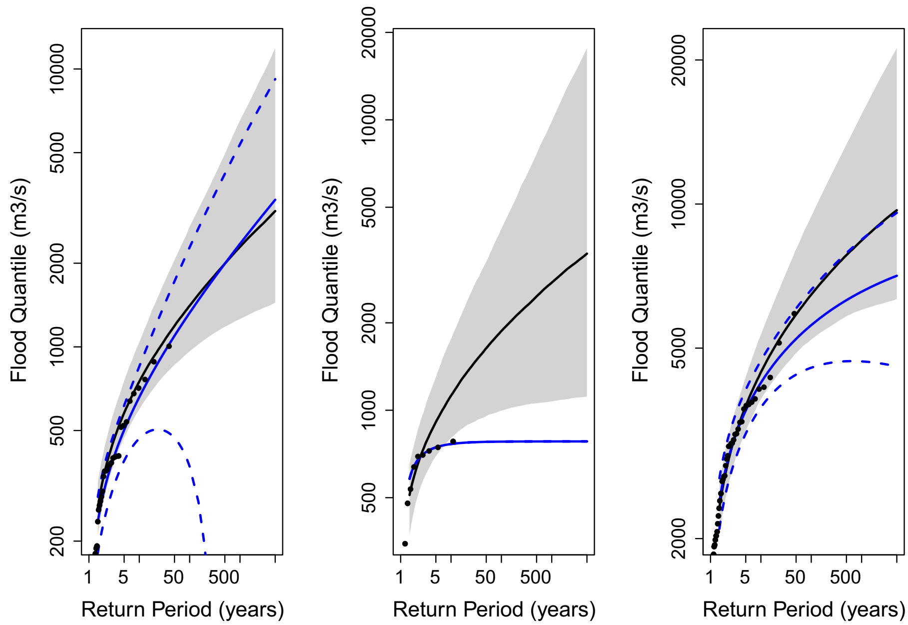

Fig. 5. Flood quantile estimates for three randomly chosen sites from the training set (sites # 14, 16 and 18) based on the proposed hierarchical Bayesian model (black line) and ordinary GEV (blue line). The gray region shows the 95% Bayesian credible interval while the dashed blue lines show the 95% confidence interval for the ordinary GEV. Black dots are empirical estimates. Results for the remaining training sites are in supplementary material.

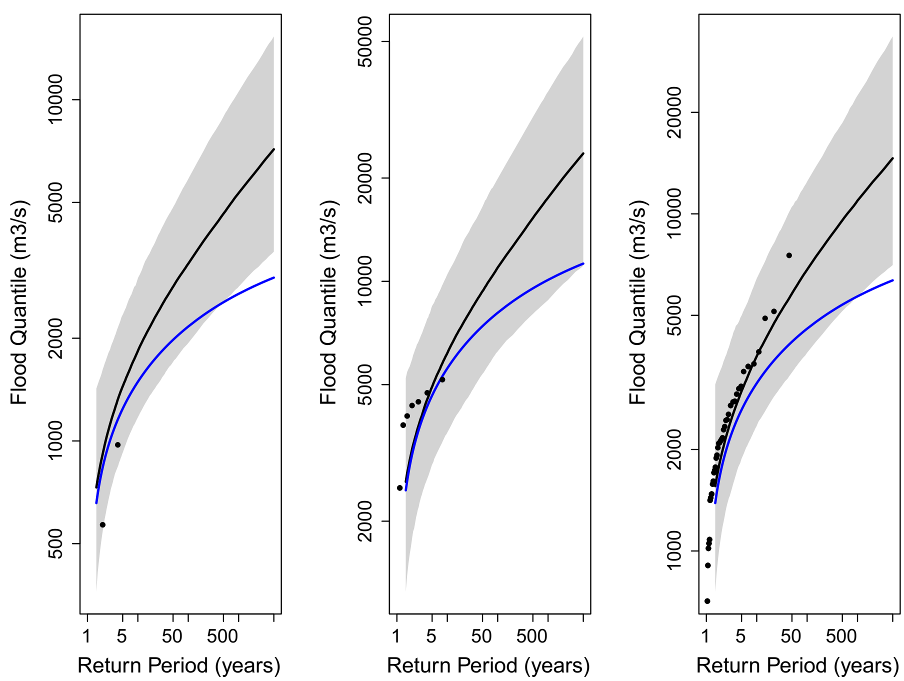

Fig. 6. Flood quantile estimates for out-of-sample sites (# 1, 4 and 5) based on the proposed hierarchical Bayesian model (black line) and on the L-moments index flood method (blue line). The gray region shows the 95% Bayesian credible interval. Black dots are empirical estimates.

quantiles (blue line) based on the index flood method with L-moments (Hosking and Wallis, 1997; Viglione, 2013) and based on the Hierarchical Bayesian method (black line) along with the associated 95% credible interval (gray region in Fig. 6). The L Moments method does not lend itself to a rigorous uncertainty analysis in a direct way. For all three sites, the HB credible interval includes the empirical estimates (black dots in Fig. 6). The index flood estimates deviate significantly from the empirical ones, particularly for the site showed in the rightmost panel.

4. Summary

In this paper we developed and applied a method for local and regional flood frequency analysis where the distribution parameters are estimated in a hierarchical Bayesian framework, to quantify and reduce parameter uncertainties for flood quantile estimates. The Brazil application shows the efficacy of the method in using data from multiple stations with missing values and variable periods of record, while appropriately pooling the information across the locations.

Please cite this article in press as: Lima, C.H.R., et al. A hierarchical Bayesian GEV model for improving local and regional flood quantile estimates. J. Hydrol. (2016), http://dx.doi.org/10.1016/j.jhydrol.2016.07.042

The scaling of the location and scale parameters of the GEV appears to be an effective way to model the dependence of the flood process on this parameter, while the pooling of the shape parameter across the sites used in the example is seen to be effective in constraining the estimates of the parameter while allowing some variation across the sites. This pooling stabilized the estimates and reduced the uncertainty in the flood quantile estimates, that arises largely from the poor estimation of this parameter of the GEV distribution.

For sites with short records, the HB estimates for the shape parameter tend to be closer to the average estimates. In this sense, the prior distribution of ξ which reflects a shrinkage towards a common mean was able to constrain the shape parameter of sites with short records to plausible values ($-0.5 < \xi < 0.5$, e.g. Martins and Stedinger, 2000) which would otherwise not be the case using traditional ML estimates. The homogeneity assumption used in the index flood quantile is not needed in the HB model as introduced here. The variability of ξ across sites is controlled by τ_ξ in Eq. (7), which was set a uniform prior distribution and therefore is essentially estimated from data. We expect that τ_ξ will increase as the region under study becomes less homogeneous or has less data available. For both homogeneous and heterogeneous regions, assign ξ for an ungauged site by sampling it from Eq. (5) and the posterior distribution of $\overline{\xi}$ and τ_ξ seems to be an appropriate approach based on the cross-validated results obtained here.

An extension of the approach to consider various covariates is possible. For instance, spatial factors may relate to drainage area as expected from scaling considerations, while the scale and location parameters across the region could also be allowed to vary in time as functions of appropriate climate variables, and one could model this variation as homogeneous or heterogeneous through partial pooling across stations, as has been done in prior work. The Hierarchical Bayesian approach is attractive since it provides a way to consider a variety of model structures, and exposes the uncertainty associated with each of them. Consequently, one can compare not just the estimates from nested models with different structures, but also the uncertainty estimated with each level and formulation. Despite the use of "objective" criteria, model building is invariably an interactive and exploratory activity, where the analyst seeks to learn from the data. Applications of the multilevel model in this context are revealing as to when and where different formulations and predictors may be effective.

Acknowledgments

We thank ANA for making the streamflow data available. This work was partially funding by AIG Insurance. The first author acknowledges a Postdoctoral Fellowship from the Brazilian Government Agency CNPq (Conselho Nacional de Desenvolvimento Cientifico e Tecnologico) during part of this work. We also thank the two anonymous reviewers for the insightful comments that greatly improved the original manuscript.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jhydrol.2016.07.042.

References

Adamowski, K., 2000. Regional analysis of annual maximum and partial duration flood data by nonparametric and L-moment methods. J. Hydrol. 229 (34), 219–231. http://www.sciencedirect.com/science/article/pii/S0022169400001566.

Benito, G., Thorndycraft, V., 2005. Palaeoflood hydrology and its role in applied hydrological sciences. J. Hydrol. 313, 3–15. http://www.sciencedirect.com/science/article/pii/S0022169405000259.

Benson, M.A., 1962. Evolution of methods for evaluating the occurrence of floods. Tech. rep., US Geological Survey water-supply paper 1580-A.

Burlando, P., Rosso, R., 1996. Scaling and multiscaling models of depth-duration-frequency curves for storm precipitation. J. Hydrol. 187, 45–64.

Cheng, L., AghaKouchak, A., Gilleland, E., Katz, R.W., 2014. Non-stationary extreme value analysis in a changing climate. Climatic Change 127 (2), 353–369 http://dx.doi.org/10.1007/s10584-014-1254-5.

Coles, S., 2001. An Introducton to Statistical Modeling of Extreme Values. Springer.

Cunderlik, J.M., Burn, D.H., 2003. Non-stationary pooled flood frequency analysis. J. Hydrol. 276 (1–4), 210–223. http://linkinghub.elsevier.com/retrieve/pii/S0022169403000623.

Dalrymple, T., 1960. Flood-frequence analyses. Tech. rep., US Geological Survey water-supply paper 1543-A.

Gaume, E., Gaál, L., Viglione, A., Szolgay, J., Kohnová, S., Blöschl, G., 2010. Bayesian MCMC approach to regional flood frequency analyses involving extraordinary flood events at ungauged sites. J. Hydrol. 394 (1–2), 101–117. http://linkinghub.elsevier.com/retrieve/pii/S0022169410000193.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013. Bayesian Data Analysis, third ed. CRC Press.

Gumbel, E.J., 1958. Statistics of Extremes. Columbia University Press.

Gupta, V.K., Dawdy, D.R., 1995. Physical interpretations of regional variations in the scaling exponents of flood quantiles. Hydrol. Process. 9, 347–361.

Gupta, V.K., Mesa, O.J., Dawdy, D.R., 1994. Multiscaling theory of flood peaks: regional quantile analysis. Water Resour. Res. 30 (12), 3405–3421.

Gupta, V.K., Troutman, B.M., Dawdy, D.R., 2007. Towards a nonlinear geophysical theory of floods in river networks: an overview of 20 years of progress. In: Tsonis, A., Elsner, J. (Eds.), Nonlinear Dynamics in Geosciences. Springer, pp. 121–151.

Gupta, V.K., Waymire, E., 1990. Multiscaling properties of spatial rainfall and river flow distributions. J. Geophys. Res. 95 (D3), 1999–2009.

Haddad, K., Rahman, A., Stedinger, J.R., 2012. Regional flood frequency analysis using Bayesian generalized least squares: a comparison between quantile and parameter regression techniques. Hydrol. Process. 26 (7), 1008–1021. http://doi.wiley.com/10.1002/hyp.8189.

Halbert, K., Nguyen, C., Payrastre, O., Gaume, E., 2016. Reducing uncertainty in flood frequency analyses: a comparison of local and regional approaches involving information on extreme historical floods. J. Hydrol. http://wwwsciencedirect.com/science/article/pii/S0022169416000342.

Hirschboeck, K.K., 1988. Flood hydroclimatology. In: Baker, V., Kochel, R.C., Patton, P.C. (Eds.), Flood Geomorphology. John Wiley & Sons, Inc.

Hosking, J.R.M., Wallis, J.R., 1997. Regional Frequency Analysis. Cambridge University Press.

Ishak, E., Haddad, K., Zaman, M., Rahman, A., 2011. Scaling property of regional floods in New South Wales Australia. Nat. Hazards 58, 1155–1167.

Katz, R.W., Parlange, M.B., Naveau, P., 2002. Statistics of extremes in hydrology. Adv. Water Resour. 25, 1287–1304.

Kuczera, G., 1982. Combining site-specific and regional information: an empirical Bayes approach. Water Resourc. Res. 18 (2), 306–314. http://dx.doi.org/10.1029/WR018i002p00306.

Kwon, H.-H., Brown, C., Lall, U., 2008. Climate informed flood frequency analysis and prediction in Montana using hierarhical Bayesian modeling. Geophys. Res. Lett. 35, 1–6.

Leclerc, M., Ouarda, T.B., 2007. Non-stationary regional flood frequency analysis at ungauged sites. J. Hydrol. 343 (34), 254–265. http://www.sciencedirect.com/science/article/pii/S0022169407003824.

Lettenmaier, D.P., Potter, K.W., 1985. Testing flood frequency estimation methods using a regional flood generation model. Water Resour. Res. 21 (12), 1903–1914. http://dx.doi.org/10.1029/WR021i012p01903.

Liang, Z., Chang, W., Li, B., 2011. Bayesian flood frequency analysis in the light of model and parameter uncertainties. Stochast. Environ. Res. Risk Assess. 26 (5), 721–730. http://link.springer.com/10.1007/s00477-011-0552-y.

Lima, C.H., Lall, U., Troy, T.J., Devineni, N., 2015. A climate informed model for nonstationary flood risk prediction: application to Negro River at Manaus, Amazonia. J. Hydrol. 522, 594–602. http://www.sciencedirect.com/science/article/pii/S0022169415000153.

Lima, C.H.R., Lall, U., 2010. Spatial scaling in a changing climate: a hierarchical bayesian model for non-stationary multi-site annual maximum and monthly streamflow. J. Hydrol. 383, 307–318.

Lunn, D., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. Stat. Comput. 10, 325–337.

Madsen, H., Rosbjerg, D., 1997. Generalized least squares and empirical Bayes estimation in regional partial duration series index-flood modeling. Water Resour. Res. 33 (4), 771–781. http://www.agu.org/pubs/crossref/1997/96WR03850.shtml.

Martins, E.S., Stedinger, J.R., 2000. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. Water Resour. Res. 3 (3), 737–744.

Merz, R., Blöschl, G., 2005. Flood frequency regionalisation - spatial proximity vs. catchment attributes. J. Hydrol. 302, 283–306.

Merz, R., Blöschl, G., 2008a. Flood frequency hydrology: 1. Temporal, spatial, and causal expansion of information. Water Resour. Res. 44 (8), w08432, n/a–n/a. http://dx.doi.org/10.1029/2007WR006744.

Merz, R., Blöschl, G., 2008b. Flood frequency hydrology: 2. Combining data evidence. Water Resour. Res. 44 (8), w08433, n/a–n/a. http://dx.doi.org/10.1029/2007WR006745.

Morrison, J.E., Smith, J.A., 2002. Stochastic modeling of flood peaks using the generalized extreme value distribution. Water Resour. Res. 38 (12), 41-1–41-12, 1305. http://dx.doi.org/10.1029/2001WR000502.

Northrop, P.J., 2004. Likelihood-based approaches to flood frequency estimation. J. Hydrol. 292, 96–113. http://www.sciencedirect.com/science/article/pii/S0022169404000253.

Parent, E., Bernier, J., 2003. Bayesian POT modeling for historical data. J. Hydrol. 274, 95–108 http://www.sciencedirect.com/science/article/pii/S0022169402003967.

Rao, A.R., Hamed, K.H., 2000. Flood Frequency Analysis. CRC.

Reis Jr., D.S., Stedinger, J.R., 2005. Bayesian MCMC flood frequency analysis with historical information. J. Hydrol. 313, 97–116. http://www.sciencedirect.com/science/article/pii/S0022169405000314.

Renard, B., 2011. A Bayesian hierarchical approach to regional frequency analysis. Water Resour. Res. 47 (11). http://doi.wiley.com/10.1029/2010WR010089.

Renard, B., Kochanek, K., Lang, M., Garavaglia, F., Paquet, E., Neppel, L., Najib, K., Carreau, J., Arnaud, P., Aubert, Y., Borchi, F., Soubeyroux, J.-M., Jourdain, S., Veysseire, J.-M., Sauquet, E., Cipriani, T., Auffray, a., 2013. Data-based comparison of frequency analysis methods: a general framework. Water Resour. Res. 49 (2), 825–843. http://doi.wiley.com/10.1002/wrcr.20087.

Renard, B., Lall, U., 2014. Regional frequency analysis conditioned on large-scale atmospheric or oceanic fields. Water Resour. Res. 50 (12), 9536–9554. http://dx.doi.org/10.1002/2014WR016277.

Renard, B., Lang, M., Bois, P., 2006. Statistical analysis of extreme events in a non-stationary context via a Bayesian framework: case study with peak-over-threshold data. Stochast. Environ. Res. Risk Assess. 21 (2), 97–112. http://link.springer.com/10.1007/s00477-006-0047-4.

Sivapalan, M., Jothityangkoon, C., Menabde, M., 2002. Linearity and nonlinearity of basin response as a function of scale: discussion of alternative definitions. Water Resour. Res. 38 (2).

Sivapalan, M., Takeuchi, K., Franks, S.W., Gupta, V.K., Karambiri, H., Lakshmi, V., Liang, X., Mcdonnell, J.J., Mendiondo, E.M., O'Connell, P.E., Oki, T., Pomeroy, J.W., Schertzer, D., Uhlenbrook, S., Zehe, E., 2003. IAHS decade on Predictions in Ungauged Basins (PUB), 2003-2012: shaping an exciting future for the hydrological sciences. Hydrol. Sci. J. 48 (6), 857–880. http://dx.doi.org/10.1623/hysj.48.6.857.51421.

Smith, J., 1989. Regional flood frequency analysis using extreme order statistics of the annual peak record. Water Resour. Res. 25 (2), 311–317. http://onlinelibrary.wiley.com/doi/10.1029/WR025i002p00311/full.

Smith, J.A., Villarini, G., Baeck, M.L., 2011. Mixture distributions and the hydroclimatology of extreme rainfall and flooding in the Eastern United States. J. Hydrometeorol. 12 (2), 294–309. http://journals.ametsoc.org/doi/abs/10.1175/2010JHM1242.1.

Stedinger, J., Lu, L., 1995. Appraisal of regional and index flood quantile estimators. Stochast. Hydrol. Hydraul. 9, 49–75. http://link.springer.com/article/10.1007/BF01581758.

Stedinger, J.R., Vogel, R.M., Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events. In: Maidment, D.R. (Ed.), Handbook of Hydrology. McGRAW-HILL, INC., pp. 18.1–18.6 (Chapter 18).

Steinschneider, S., Lall, U., 2015. A hierarchical bayesian regional model for nonstationary precipitation extremes in northern california conditioned on tropical moisture exports. Water Resour. Res. 51 (3), 1472–1492. http://dx.doi.org/10.1002/2014WR016664.

Sturtz, S., Ligges, U., Gelman, A., 2005. R2WinBUGS: a package for running WinBUGS from R. J. Stat. Softw. 12 (3), 1–16.

Sun, X., Lall, U., 2015. Spatially coherent trends of annual maximum daily precipitation in the United States. Geophys. Res. Lett. 42 (22), 9781–9789. http://dx.doi.org/10.1002/2015GL066483 2015GL066483.

Sun, X., Lall, U., Merz, B., Dung, N.V., 2015. Hierarchical Bayesian clustering for nonstationary flood frequency analysis: application to trends of annual maximum flow in Germany. Water Resour. Res. 51 (8), 6586–6601 http://dx.doi.org/10.1002/2015WR017117.

Towler, E., Rajagopalan, B., Gilleland, E., Summers, R.S., Yates, D., Katz, R.W., 2010. Modeling hydrologic and water quality extremes in a changing climate: a statistical approach based on extreme value theory. Water Resour. Res. 46 (11), n/a–n/a. http://doi.wiley.com/10.1029/2009WR008876.

Viglione, A., 2013. nsRFA: non-supervised Regional Frequency Analysis. R package version 0.7-10. <http://CRAN.R-project.org/package=nsRFA>.

Viglione, A., Merz, R., Salinas, J.L., Blöschl, G., 2013. Flood frequency hydrology: 3. A Bayesian analysis. Water Resour. Res. 49 (2), 675–692. http://doi.wiley.com/10.1029/2011WR010782.

Villarini, G., Smith, J.A., 2010. Flood peak distributions for the eastern United States. Water Resour. Res. 46 (6), w06504. http://dx.doi.org/10.1029/2009WR008395.

Villarini, G., Smith, J.a., Napolitano, F., 2010. Nonstationary modeling of a long record of rainfall and temperature over Rome. Adv. Water Resour. 33 (10), 1256–1267. http://linkinghub.elsevier.com/retrieve/pii/S030917081000062X.

Villarini, G., Smith, J.A., Serinaldi, F., Bales, J., Bates, P.D., Krajewski, W.F., 2009. Flood frequency analysis for nonstationary annual peak records in an urban drainage basin. Adv. Water Resour. 32 (8), 1255–1266. http://www.sciencedirect.com/science/article/pii/S0309170809000761.

Villarini, G., Smith, J.A., Serinaldi, F., Ntelekos, A.A., 2011. Analyses of seasonal and annual maximum daily discharge records for central Europe. J. Hydrol. 399, 299–312.

Please cite this article in press as: Lima, C.H.R., et al. A hierarchical Bayesian GEV model for improving local and regional flood quantile estimates. J. Hydrol. (2016), http://dx.doi.org/10.1016/j.jhydrol.2016.07.042