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

Carlos H.R. Lima, Upmanu Lall, Tara Troy, Naresh Devineni

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

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 km² 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.

© 2016 Elsevier B.V. All rights reserved.

Research papers

A hierarchical Bayesian GEV model for improving local and regional flood quantile 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...
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 process 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 Thornycraft, 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 (Vigilone 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), Vigilone 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 \( Q_{ij} \) represents the annual maximum peak flow for site \( i \) in year \( j \). Under some regularity and stationary conditions, extreme value theory (see for instance Coles, 2001) postulates that \( Q_{ij} \) follows a GEV distribution \( g \) with parameters \( \mu_i \) (location), \( \sigma_i \) (scale) and \( \xi_i \) (shape):

\[
Q_{ij} \sim g(\mu_i, \sigma_i, \xi_i)
\]
where the cumulative distribution function of g is given by:

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

(2)

Given a time series record of \( Q_{ij} \) for different locations, the parameters \( \mu_j, \sigma_j \) and \( \xi_j \) 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 \( \mu_j, \sigma_j \) and \( \xi_j \) 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 \( \mu_j \) and \( \sigma_j \) that reflects our knowledge about them before observing the data:

\[ \log(\mu_j) \sim N(\alpha_1 + \alpha_2 \cdot \log(A_j), \tau_{\mu}^2) \]  

(3)

\[ \log(\sigma_j) \sim N(\beta_1 + \beta_2 \cdot \log(A_j), \tau_{\sigma}^2) \]  

(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 \( \tau_{\sigma}^2 \) are the so-called hyperparameters.

The scaling of the shape parameter \( \xi_j \) 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 (Burland 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:

\[ \xi_j \sim N(\xi, \tau_{\xi}^2) \]  

(5)

with \( \xi \) and \( \tau_{\xi}^2 \) 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:

\[ p(\alpha_1, \alpha_2, \beta_1, \beta_2, \xi) \propto 1 \]  

(6)

and

\[ p(\log(\tau_{\mu}, \tau_{\sigma}, \tau_{\xi})) \propto 1 \]  

(7)

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

\[ p(\theta|q) = \frac{p(q|\theta) \cdot p(\theta)}{p(q)} \propto p(q|\theta) \cdot p(\theta) \]  

(8)

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

\[ p(q|\theta) = \prod_{j=1}^{n} p(q_{ij}|\mu_j, \sigma_j, \xi_j) \]  

(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 \( n \) 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 \( \theta \):

\[ p(\theta|q) \propto \prod_{j=1}^{n} \prod_{i=1}^{n_j} p(q_{ij}|\mu_j, \sigma_j, \xi_j) \cdot N(\log(\mu_j)|\alpha_1 + \alpha_2 \cdot \log(A_j), \tau_{\mu}^2) \cdot N(\log(\sigma_j)|\beta_1 + \beta_2 \cdot \log(A_j), \tau_{\sigma}^2) \cdot N(\xi_j|\xi, \tau_{\xi}^2) \]  

(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 (Bayesian inference using Gibbs Sampling, Lunn et al. (2000)), which automatically selects an appropriate sampling algorithm (e.g. Metropolis Hastings) 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 \( R \) 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 \( \mu, \sigma, \xi \) 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 \( \alpha_1, \alpha_2, \beta_1, \beta_2, \tau_{\mu}, \tau_{\sigma}, \) and \( \tau_{\xi} \). 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 (southeastern 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.

<table>
<thead>
<tr>
<th>Site</th>
<th>Latitude (S)</th>
<th>Longitude (W)</th>
<th>Drainage area (km²)</th>
<th>Series length (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-19.77</td>
<td>-42.48</td>
<td>16,100</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>-19.78</td>
<td>-42.07</td>
<td>162</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td>-19.33</td>
<td>-42.37</td>
<td>24,200</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>-19.53</td>
<td>-40.63</td>
<td>76,400</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>-18.88</td>
<td>-41.95</td>
<td>40,500</td>
<td>43</td>
</tr>
<tr>
<td>6</td>
<td>-19.41</td>
<td>-40.06</td>
<td>82,100</td>
<td>27</td>
</tr>
<tr>
<td>7</td>
<td>-19.52</td>
<td>-42.64</td>
<td>5270</td>
<td>26</td>
</tr>
<tr>
<td>8</td>
<td>-19.19</td>
<td>-42.42</td>
<td>10,200</td>
<td>38</td>
</tr>
<tr>
<td>9</td>
<td>-19.77</td>
<td>-43.03</td>
<td>3060</td>
<td>35</td>
</tr>
<tr>
<td>10</td>
<td>-20.38</td>
<td>-42.99</td>
<td>6230</td>
<td>37</td>
</tr>
<tr>
<td>11</td>
<td>-19.34</td>
<td>-41.25</td>
<td>61,200</td>
<td>21</td>
</tr>
<tr>
<td>12</td>
<td>-19.49</td>
<td>-41.16</td>
<td>8720</td>
<td>72</td>
</tr>
<tr>
<td>13</td>
<td>-18.97</td>
<td>-41.64</td>
<td>55,100</td>
<td>38</td>
</tr>
<tr>
<td>14</td>
<td>-18.57</td>
<td>-41.92</td>
<td>9770</td>
<td>38</td>
</tr>
<tr>
<td>15</td>
<td>-19.99</td>
<td>-42.67</td>
<td>10,100</td>
<td>31</td>
</tr>
<tr>
<td>16</td>
<td>-19.90</td>
<td>-42.58</td>
<td>12,800</td>
<td>10</td>
</tr>
<tr>
<td>17</td>
<td>-19.33</td>
<td>-42.36</td>
<td>24,200</td>
<td>44</td>
</tr>
<tr>
<td>18</td>
<td>-19.32</td>
<td>-41.25</td>
<td>61,200</td>
<td>45</td>
</tr>
<tr>
<td>19</td>
<td>-20.65</td>
<td>-43.83</td>
<td>77</td>
<td>23</td>
</tr>
<tr>
<td>20</td>
<td>-19.53</td>
<td>-43.17</td>
<td>1160</td>
<td>75</td>
</tr>
</tbody>
</table>

Fig. 1. Location of Rio Doce basin in southeastern 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.

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 \( \mu \) and scale \( \sigma \) parameters present a well defined log-log linear relationship with the drainage area, with different intercept (\( \alpha_1 \) and \( \beta_1 \)) and slope (\( \alpha_2 \) and \( \beta_2 \)) parameters. As conjectured earlier, the shape parameter \( \xi \) 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 (\( \xi = -1.078 \)) 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 \( \xi \) 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 corresponding 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
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 \( \mu \) and \( \sigma \). The cross-validation results are displayed in Fig. 6. They show estimates of flood quantiles and associated uncertainties for return periods beyond the length of the streamflow data record.

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.

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.

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.
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 $\xi$ which reflects a shrinkage towards a common mean was able to constrain the shape parameter of sites in the GEV distribution. The homogeneity assumption used in the index flood quantile is not needed in the HB model as introduced here. The variability of $\xi$ across sites is controlled by $\tau_i$ in Eq. (7), which was set a uniform prior distribution and therefore is essentially estimated from data. We expect that $\tau_i$ will increase as the region under study becomes less homogeneous or has less data available. For both homogeneous and heterogeneous regions, assign $\xi$ for an ungauged site by sampling it from Eq. (5) and the posterior distribution of $\xi$ and $\tau_i$ 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 funded by AIG Insurance. The first author acknowledges a Postdoctoral Fellowship from the Brazilian Government Agency CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) 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


