Application of Bayesian latent Gaussian models to precipitation extremes in the context of extreme value theory

Kári Hlynsson

July 2024

Abstract

Effective modelling of extreme precipitation events have the potential to reduce casualty and infrastructural damage as a result of natural disasters such as those caused by floods. The application of Bayesian hierarchical models (BHMs), particularly Bayesian latent Gaussian models (BLGMs), has seen substantial success in mitigating the effects of these events and motivating preventative measures. In this article, BLGMs are applied to precipitation data where the response layer is modelled using the generalized extreme value (GEV) distribution. Three models are compared, namely (i) a baseline model, (ii) a time-trend model, and (iii) a piecewise-linear time trend model. Important concepts from extreme value theory are reviewed to give the reader perspective.

1 Introduction

1.1 Motivation and statement of purpose

Disastrous and often unlikely events such as floods, heatwaves or financial market crashes can occur suddenly and without warning, leading to loss of life and damage to infrastructure. Accurate modelling of such events has been the subject of great interest in the statistical community, and has recently seen substantial success when considered in the context of extreme value theory. In particular, Bayesian hierarchical models (BHMs) have proven to be especially desirable due to their ability to incorporate prior knowledge and to obtain a measure of uncertainty, for example by using the posterior predictive density when predicting likely future outcomes.

In this article, models of maximum cumulative precipitation in the period of 24 hours by year is modelled using the generalized extreme value (GEV) distribution, which is a natural distribution to block maxima data (such as yearly maxima), as follows from the Fisher-Tippett-Gnedenko theorem, briefly discussed in the next section. The distribution has three parameters, μ, σ, and ξ, which are often referred to as the *location*, *scale*, and *shape* parameters, respectively. Three models are considered, with increasing complexity. The first model, the null model, claims that the precipitation data are modelled with no trend, ignoring climate effects. The second model uses a time trend in the location parameter of the GEV distribution. Lastly, the third model uses piecewise linear trends to predict changes in location parameter with time. Comparison of the aforementioned models has the potential to eludicate effects of climate change on extreme precipitation events.

2 Extreme value theory

Fisher-Tippett theorem states that [...]. The GEV has likelihood

$$\mathcal{L}(y \mid \mu, \sigma, \xi) = \sigma^{-1}\eta(y)^{\xi+1} \exp\big(-\eta(y)\big),$$

where

$$\eta(y) = \begin{cases} \left[1 + \xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-1/\xi} & \text{if } \xi \neq 0, \\ \exp\left(\frac{y-\mu}{\sigma}\right) & \text{if } \xi = 0. \end{cases}$$

Thus, the likelihood of the observed data is

$$\begin{aligned} \mathcal{L}(\boldsymbol{y} \mid \mu, \sigma, \xi) &= \prod_{i=1}^{n} \mathcal{L}(y_i \mid \mu, \sigma, \xi) \\ &= \sigma^{-n} \prod_{i=1}^{n} \eta(y_i)^{\xi+1} \exp\big(-\eta(y_i)\big), \end{aligned}$$

and by taking the log of the above we arrive at the log likelihood:

$$\ell(\boldsymbol{y} \mid \mu, \sigma, \xi) = -n \log(\sigma) + \sum_{i=1}^{n} (1+\xi) \log \eta(y_i) - \sum_{i=1}^{n} \eta(y_i).$$