

Report version 8

Tony Zheng

21th June 2020

1 Introduction

The study of environmental conditions is crucial in the development of large scale construction projects. In particular, the study of long term behaviour of environmental variables is necessary to understand the risks of hazardous meteorological events such as floods, storms and droughts. This can be done using the well established theory of extreme values, using asymptomatic models such as a Process point process model described in Coles [4]. However, the analysis of extreme events is often hindered by a scarcity of data, and an inadequate treatment of model and prediction uncertainties. To overcome these problems, a Bayesian methodology has been proposed. This approach allows us to exploit prior information, and provides a framework for taking account of the various uncertainties in the prediction process.

A review of the use of Bayesian methods in extreme value theory has been carried out in Coles and Powell [5], and analyses of rainfall data have been performed in both Coles et al. [3] and Coles et al. [6]. In the latter, Coles et al. use the Poisson point process model, and elicit expert-specified prior information in the form of independent Gamma distributions on the difference of three quantiles of the annual maxima. In this paper, we will attempt to replicate this analysis using alternative prior distributions.

In § 2, we will outline the Poisson point process model. In § 3, we will describe the original prior distribution used by Coles et al., and propose alternative prior distributions. In § 5, we will explain how Markov chain Monte Carlo (MCMC) methods can be used in the analysis to avoid intractable calculations. to sample from the Finally, in § 6, we will implement and compare our priors in two simulation studies.

2 Likelihood model

Assume that we have daily observations of some strictly positive random variable X over M years, and some (sufficiently large) threshold u . Denote by $\mathbf{x} = (x_1, \dots, x_{n_u})$ the exceedances of our data by the threshold. Using the Poisson point process characterisation of extremes detailed in Coles [4], the likelihood of the data is

$$L(\mu, \sigma, \xi \mid \mathbf{x}) = \exp(-M\Lambda[u, \infty)) \prod_{i=1}^{n_u} \lambda(x_i), \quad (1)$$

with intensity function

$$\lambda(x) := \begin{cases} \frac{1}{\sigma} \left\{ 1 + \xi \left(\frac{x-\mu}{\sigma} \right) \right\}_+^{-\frac{\xi+1}{\xi}} & \text{if } \xi \neq 0, \\ \frac{1}{\sigma} \exp\left(-\frac{x-\mu}{\sigma}\right) & \text{if } \xi = 0, \end{cases} \quad (2)$$

and

$$\Lambda[u, \infty) := \int_u^\infty \lambda(x) dx \quad (3)$$

$$= \begin{cases} \left\{ 1 + \xi \left(\frac{u-\mu}{\sigma} \right) \right\}_+^{-\frac{1}{\xi}} & \text{if } \xi \neq 0, \\ \exp\left(-\frac{u-\mu}{\sigma}\right) & \text{if } \xi = 0. \end{cases} \quad (4)$$

The parameters $\theta := (\mu, \sigma, \xi)$ are defined in the set

$$\Theta := \{(\mu, \sigma, \xi) \in \mathbb{R}^3 : \sigma > 0\}, \quad (5)$$

and are known as Generalised Extreme Value (GEV) parameters. This is because the annual maxima of the data follow a GEV distribution with parameters μ, σ, ξ :

$$\Pr(\max\{X_1, \dots, X_{365} \leq x\}) = \begin{cases} \exp\left(-\left\{1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right\}_+^{-\frac{1}{\xi}}\right) & \text{if } \xi \neq 0, \\ \exp(-\exp(-\frac{x-\mu}{\sigma})) & \text{if } \xi = 0. \end{cases} \quad (6)$$

Inverting (6), we get the quantile function

$$q(p \mid \mu, \sigma, \xi) = \begin{cases} \mu + \frac{\sigma}{\xi}(\exp(s\xi) - 1) & \text{if } \xi \neq 0, \\ \mu + \sigma s & \text{if } \xi = 0, \end{cases}$$

where

$$s := -\log(-\log(1 - p)).$$

Of interest to us are the quantiles corresponding to very small probabilities. The *return level* of a *return period* r in years is defined as the quantile $q(1 - 1/r \mid \mu, \sigma, \xi)$.

3 Eliciting prior information

In Bayesian methodology, we elicit prior information about the model in the form of a distribution π_θ on (μ, σ, ξ) . In our case, we assume that this information arises from the opinion of an expert. Unfortunately, it is not feasible to specify joint distribution of (μ, σ, ξ) . Instead, we will consider a prior for three quantiles of the annual maxima.

Let $0 < p_1 < p_2 < p_3 < 1$ be three probabilities, and consider the corresponding quantiles

$$(q_1, q_2, q_3) \in \mathcal{Q} := \{(q_1, q_2, q_3) \in \mathbb{R}^3 : q_1 < q_2 < q_3\},$$

defined by

$$q_i := q(1 - p_i \mid \mu, \sigma, \xi), \quad i = 1, 2, 3. \quad (7)$$

A prior for (q_1, q_2, q_3) is equivalent to a prior for (μ, σ, ξ) through the inverse of the transformation

$$\begin{aligned} g: \Theta &\rightarrow \mathcal{Q} \\ (\mu, \sigma, \xi) &\mapsto (q_1, q_2, q_3). \end{aligned}$$

Using the change of variables formula,

$$\pi_\theta(\mu, \sigma, \xi) = \pi_q(g(\mu, \sigma, \xi)) |\det J(g)(\mu, \sigma, \xi)| \mathbb{1}_{\sigma > 0}, \quad (8)$$

where $J(g)$ is the Jacobian of g . We will now derive an expression for the determinant. Recall that

$$s_i = -\log(-\log(1 - p_i)), \quad i = 1, 2, 3$$

and

$$q_i = \mu + \sigma \frac{\exp(s_i \xi) - 1}{\xi}, \quad i = 1, 2, 3, \quad (9)$$

where we are assuming that $\xi \neq 0$. Therefore

$$\begin{aligned}\frac{dq_i}{d\mu}(\mu, \sigma, \xi) &= 1, \\ \frac{dq_i}{d\sigma}(\mu, \sigma, \xi) &= \frac{\exp(s_i \xi) - 1}{\xi}, \\ \frac{dq_i}{d\xi}(\mu, \sigma, \xi) &= \frac{\sigma(s_i \xi \exp(s_i \xi) - \exp(s_i \xi) + 1)}{\xi^2},\end{aligned}$$

and therefore

$$\begin{aligned}\det J(\mu, \sigma, \xi) &= \frac{\sigma}{\xi^3} \begin{vmatrix} 1 & 1 \\ \exp(s_1 \xi) - 1 & \exp(s_2 \xi) - 1 \\ s_1 \xi \exp(s_1 \xi) - \exp(s_1 \xi) + 1 & s_2 \xi \exp(s_2 \xi) - \exp(s_2 \xi) + 1 \end{vmatrix} \\ &= \frac{\sigma}{\xi^3} \begin{vmatrix} 1 & 1 \\ \exp(s_3 \xi) - 1 & \\ s_3 \xi \exp(s_3 \xi) - \exp(s_3 \xi) + 1 & \end{vmatrix} \\ &= \frac{\sigma}{\xi^3} \begin{vmatrix} 1 & 1 & 1 \\ \exp(s_1 \xi) & \exp(s_2 \xi) & \exp(s_3 \xi) \\ s_1 \xi \exp(s_1 \xi) & s_2 \xi \exp(s_2 \xi) & s_3 \xi \exp(s_3 \xi) \end{vmatrix} \\ &= \frac{\sigma}{\xi^2} \sum_{i=1}^3 s_i \exp(s_i \xi) (\exp(s_{i+2} \xi) - \exp(s_{i+1} \xi)) > 0,\end{aligned}$$

where $i \equiv j \pmod 3$ implies $s_i = s_j$. Therefore the formula (8) becomes

$$\pi_\theta(\mu, \sigma, \xi) = \pi_q(g(\mu, \sigma, \xi)) \sigma \Psi(\xi) \mathbb{1}_{\sigma > 0}, \quad (10)$$

with

$$\Psi(\xi) := \frac{1}{\xi^2} \sum_{i=1}^3 s_i \exp(s_i \xi) (\exp(s_{i+2} \xi) - \exp(s_{i+1} \xi)).$$

As the data are assumed to be strictly positive, the support of the quantiles implies that a distribution for (q_1, q_2, q_3) is a valid prior if and only if

$$\Pr(0 < q_1 < q_2 < q_3) = 1.$$

In order to satisfy this constraint, Coles et al. [6] specify marginal priors for the quantile differences

$$\begin{aligned}\tilde{q}_1 &:= q_1, \\ \tilde{q}_2 &:= q_2 - q_1, \\ \tilde{q}_3 &:= q_3 - q_2,\end{aligned}$$

and construct a joint distribution using an independent copula. In particular, the expert provided 0.5-quantile and 0.9-quantile estimates of each distribution, and Gamma distributions were fitted. We will denote this prior π^{QD} for Quantile Differences.

The choice of a distribution for the quantile differences can be motivated using the principle of maximum entropy, which was introduced by Jaynes [8]. Entropy is a measure of how much information we have about a distribution. It is defined by Shannon [14], for a continuous probability density p , as

$$\mathcal{E}(p) = - \int_{\mathcal{D}(p)} p(x) \log_2(p(x)) \, dx.$$

According to the principle of maximum entropy, if we are to choose a prior from a class of distributions satisfying certain constraints, then we should choose a distribution which maximises the entropy in that class, as it will be the least informative. As quantile differences are positive, we are looking for distributions with the support constraint $(0, +\infty)$, and functional constraints defined by certain quantities which are fixed by an expert. The maximum entropy distributions for various fixed quantities are shown below:

Fixed quantities	Maximum entropy distribution
Any number of quantiles	Does not exist
1 st moment	Exponential
1 st , 2 nd moments	Truncated normal

Specifying 0.5 and 0.9-quantiles, as in π^{QD} , does not have a maximum entropy solution. However, specifying the mean leads to an exponential distribution, and specifying the mean and variance leads to a truncated normal distribution. We will denote these priors π^{QDE} and π^{QDTN} respectively. In order to implement and compare these priors, we will first specify 0.5 and 0.9-quantiles and fit Gamma distributions, then obtain the mean and variance of these Gamma distributions as specifications of the mean and variance.

Although we have thus far been considering an independent copula due to simplicity, the choice of copula is also one which can be made using the principle of maximum entropy. We will consider a prior π^{ME} which is the joint distribution with maximum entropy for fixed marginals q_1, q_2, q_3 satisfying $\Pr(q_1 < q_2 < q_3) = 1$. This order constraint allows us to specify priors on the quantiles instead of the quantile differences, which may be more convenient for the expert. However, in order to compare this prior to π^{QD} , we will choose these priors as approximations of the marginals of π_q^{QD} using numerical minimisation of the Kullback–Leibler divergence.

In the case that we are only able to elicit information on two quantile differences, we will consider a prior π^{ULS} (Uniform on Log Sigma) which includes an improper uniform prior on $\log(\sigma)$. As the maximum entropy distribution with a compact support $[a, b]$ is uniform, this prior may be thought of as the maximum entropy distribution when a and b tend to $-\infty$ and $+\infty$ respectively.

Finally, we will consider a variant of π^{QD} where ξ has a non-zero probability to be zero. This will be denoted π^{SAS} (Spike And Slab) due to the shape of the marginal of ξ .

3.1 QD: Gamma priors for independent quantile differences

Coles et al. [6] specify Gamma priors for the quantile differences

$$\begin{aligned}\tilde{q}_1 &:= q_1 \sim \Gamma(\alpha_1, \beta_1), \\ \tilde{q}_2 &:= q_2 - q_1 \sim \Gamma(\alpha_2, \beta_2), \\ \tilde{q}_3 &:= q_3 - q_2 \sim \Gamma(\alpha_3, \beta_3),\end{aligned}$$

with shape and rate parametrisation, and multiply these together to form a joint distribution π_q^{QD} with independent marginals. This is then converted into a prior for the quantiles π_q^{QD} using the transformation

$$\begin{aligned}(\tilde{q}_1, \tilde{q}_2, \tilde{q}_3) &\mapsto (\tilde{q}_1, \tilde{q}_1 + \tilde{q}_2, \tilde{q}_1 + \tilde{q}_2 + \tilde{q}_3) \\ &= (q_1, q_2, q_3).\end{aligned}$$

Explicitly, the prior for (q_1, q_2, q_3) is given by

$$\pi_q^{\text{QD}}(q_1, q_2, q_3) = C q_1^{\alpha_1 - 1} \exp(-\beta_1 q_1)$$

$$\times \prod_{i=2}^3 (q_i - q_{i-1})^{\alpha_i - 1} \exp(-\beta_i(q_i - q_{i-1})) \mathbb{1}_{0 < q_1 \leq q_2 \leq q_3}, \quad (11)$$

where

$$C := \prod_{i=1}^3 \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)}.$$

Substituting (11) into (10) gives us

$$\begin{aligned} \pi_{\theta}^{\text{QD}}(\mu, \sigma, \xi) &= C q_1^{\alpha_1 - 1} \exp(-\beta_1 q_1) \prod_{i=2}^3 \left(\sigma \frac{\exp(s_i \xi) - \exp(s_{i-1} \xi)}{\xi} \right)^{\alpha_i - 1} \\ &\quad \times \exp \left(-\beta_i \sigma \frac{\exp(s_i \xi) - \exp(s_{i-1} \xi)}{\xi} \right) \sigma \Psi(\xi) \mathbb{1}_{q_1 > 0} \mathbb{1}_{\sigma > 0} \\ &= C q_1^{\alpha_1 - 1} \exp(-\beta_1 q_1) \sigma^{\alpha_2 + \alpha_3 - 1} \prod_{i=2}^3 \left(\frac{\exp(s_i \xi) - \exp(s_{i-1} \xi)}{\xi} \right)^{\alpha_i - 1} \\ &\quad \times \exp \left(-\frac{\sigma}{\xi} \sum_{i=2}^3 \beta_i (\exp(s_i \xi) - \exp(s_{i-1} \xi)) \right) \Psi(\xi) \mathbb{1}_{q_1 > 0} \mathbb{1}_{\sigma > 0} \\ &= C q_1^{\alpha_1 - 1} \exp(-\beta_1 q_1) \sigma^{\alpha_2 + \alpha_3 - 1} \Phi_1(\xi) \exp(-\sigma \Phi_2(\xi)) \mathbb{1}_{q_1 > 0} \mathbb{1}_{\sigma > 0}, \end{aligned}$$

where

$$\begin{aligned} \Psi(\xi) &= \frac{1}{\xi^2} \sum_{i=1}^3 s_i \exp(s_i \xi) (\exp(s_{i+2} \xi) - \exp(s_{i+1} \xi)), \\ \Phi_1(\xi) &:= \prod_{i=2}^3 \left(\frac{\exp(s_i \xi) - \exp(s_{i-1} \xi)}{\xi} \right)^{\alpha_i - 1} \Psi(\xi), \\ \Phi_2(\xi) &:= \frac{1}{\xi} \sum_{i=2}^3 \beta_i (\exp(s_i \xi) - \exp(s_{i-1} \xi)) > 0. \end{aligned}$$

The marginal of (σ, ξ) is

$$\begin{aligned} \pi_{\sigma, \xi}^{\text{QD}}(\sigma, \xi) &= \int_{-\infty}^{\infty} \pi_{\theta}^{\text{QD}}(\mu, \sigma, \xi) d\mu \\ &= C \int_0^{\infty} q_1^{\alpha_1 - 1} \exp(-\beta_1 q_1) \sigma^{\alpha_2 + \alpha_3 - 1} \Phi_1(\xi) \exp(-\sigma \Phi_2(\xi)) \mathbb{1}_{\sigma > 0} dq_1 \\ &= C \sigma^{\alpha_2 + \alpha_3 - 1} \Phi_1(\xi) \exp(-\sigma \Phi_2(\xi)) \beta_1^{-\alpha_1} \\ &\quad \times \int_0^{\infty} (\beta_1 q_1)^{\alpha_1 - 1} \exp(-\beta_1 q_1) \beta_1 dq_1 \mathbb{1}_{\sigma > 0} \\ &= C \sigma^{\alpha_2 + \alpha_3 - 1} \Phi_1(\xi) \exp(-\sigma \Phi_2(\xi)) \beta_1^{-\alpha_1} \Gamma(\alpha_1) \mathbb{1}_{\sigma > 0} \\ &= C_2 \sigma^{\alpha_2 + \alpha_3 - 1} \Phi_1(\xi) \exp(-\sigma \Phi_2(\xi)) \mathbb{1}_{\sigma > 0}, \end{aligned} \quad (12)$$

where

$$C_2 = \frac{\beta_2^{\alpha_2} \beta_3^{\alpha_3}}{\Gamma(\alpha_2) \Gamma(\alpha_3)}.$$

The marginal of ξ is

$$\pi_{\xi}^{\text{QD}}(\xi) = \int_{-\infty}^{\infty} \pi_{\sigma, \xi}^{\text{QD}}(\sigma, \xi) d\sigma$$

$$\begin{aligned}
&= C_2 \int_0^\infty \sigma^{\alpha_2 + \alpha_3 - 1} \Phi_1(\xi) \exp(-\sigma \Phi_2(\xi)) \, d\sigma \\
&= C_2 \Phi_1(\xi) \Phi_2(\xi)^{-\alpha_2 - \alpha_3} \\
&\quad \times \int_0^\infty (\sigma \Phi_2(\xi))^{\alpha_2 + \alpha_3 - 1} \exp(-\sigma \Phi_2(\xi)) \, d\sigma \\
&= C_2 \Phi_1(\xi) \Phi_2(\xi)^{-\alpha_2 - \alpha_3} \Gamma(\alpha_2 + \alpha_3) \\
&= C_3 \frac{\Phi_1(\xi)}{\Phi_2(\xi)^{\alpha_2 + \alpha_3}}.
\end{aligned} \tag{13}$$

where

$$C_3 = \frac{\Gamma(\alpha_2 + \alpha_3) \beta_2^{\alpha_2} \beta_3^{\alpha_3}}{\Gamma(\alpha_2) \Gamma(\alpha_3)}.$$

3.2 ME: Prior from maximum entropy distribution of quantiles

For this prior, we specify marginal priors on the quantiles q_1, q_2, q_3 directly. In particular, we choose Gamma distributions

$$q_i \sim \Gamma(\alpha_i, \beta_i), \quad i = 1, 2, 3$$

with shape and rate parametrisation. Let f_i and F_i be their PDFs and CDFs respectively. Let

$$\begin{aligned}
h_i(x) &= \frac{f_i(x)}{F_{i-1}(x) - F_i(x)} \\
&= \frac{\Gamma(\alpha_{i-1}) \beta_i^{\alpha_i} x^{\alpha_i - 1} \exp(-\beta_i x)}{\gamma(\alpha_{i-1}, \beta_{i-1} x) \Gamma(\alpha_i) - \gamma(\alpha_i, \beta_i x) \Gamma(\alpha_{i-1})}
\end{aligned}$$

if $x > 0$, and $h(x) = 0$ otherwise. According to Butucea et al. [2], under certain conditions on the F_i , the joint distribution of (q_1, q_2, q_3) with maximum entropy which satisfies $\Pr(q_1 \leq q_2 \leq q_3) = 1$ has the unique density

$$\begin{aligned}
\pi_q^{\text{ME}}(q_1, q_2, q_3) &= f_1(q_1) \prod_{i=2}^3 h_i(q_i) \exp\left(-\int_{q_{i-1}}^{q_i} h_i(s) \, ds\right) \mathbb{1}_{q_1 \leq q_2 \leq q_3} \\
&= \frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)} q_1^{\alpha_1 - 1} \exp(-\beta_1 q_1) \prod_{i=2}^3 h_i(q_i) \exp\left(-\int_{q_{i-1}}^{q_i} h_i(s) \, ds\right) \mathbb{1}_{0 < q_1 \leq q_2 \leq q_3}.
\end{aligned}$$

See § 8.1 for more detail on one of these conditions. This distribution can be obtained in Python using the class `MaximumEntropyOrderStatisticsDistribution` from the library `OpenTURNS` [1]. Therefore, from equation (10), the corresponding prior on the GEV parameters becomes

$$\pi_\theta^{\text{ME}}(\mu, \sigma, \xi) = \frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)} q_1^{\alpha_1 - 1} \exp(-\beta_1 q_1) \prod_{i=2}^3 h_i(q_i) \exp\left(-\int_{q_{i-1}}^{q_i} h_i(s) \, ds\right) \sigma \Psi(\xi) \mathbb{1}_{q_1 > 0} \mathbb{1}_{\sigma > 0}, \tag{14}$$

where

$$\Psi(\xi) = \frac{1}{\xi^2} \sum_{i=1}^3 s_i \exp(s_i \xi) (\exp(s_{i+2} \xi) - \exp(s_{i+1} \xi)).$$

3.3 ULS: Uniform prior on log sigma

This is a variant of π^{QD} , in which an improper uniform prior is specified for $\log \sigma$, and Gamma priors are specified for independent quantile differences \tilde{q}_1 and \tilde{q}_2 given σ . That is,

$$\begin{aligned}\pi_{\sigma}^{\text{ULS}}(\sigma) &\propto \frac{1}{\sigma} \mathbb{1}_{\sigma>0}, \\ \pi_{q_1, q_2 | \sigma}^{\text{ULS}}(q_1, q_2 | \sigma) &\propto q_1^{\alpha_1-1} \exp(-\beta_1 q_1) (q_2 - q_1)^{\alpha_2-1} \exp(-\beta_2 (q_2 - q_1)) \mathbb{1}_{0 < q_1 \leq q_2} \\ \implies \pi_{\sigma, q_1, q_2}^{\text{ULS}}(\sigma, q_1, q_2) &\propto q_1^{\alpha_1-1} \exp(-\beta_1 q_1) (q_2 - q_1)^{\alpha_2-1} \exp(-\beta_2 (q_2 - q_1)) \frac{1}{\sigma} \mathbb{1}_{\sigma>0} \mathbb{1}_{0 < q_1 \leq q_2}.\end{aligned}$$

Using the inverse transformation

$$g: (\mu, \sigma, \xi) \mapsto (\sigma, q_1, q_2)$$

whose determinant of the Jacobian is

$$\det J(g) = \frac{\sigma}{\xi^2} ((s_1 \xi - 1) \exp(s_1 \xi) - (s_2 \xi - 1) \exp(s_2 \xi)),$$

the prior on (μ, σ, ξ) can be expressed as

$$\begin{aligned}\pi_{\theta}^{\text{ULS}}(\mu, \sigma, \xi) &\propto q_1^{\alpha_1-1} \exp(-\beta_1 q_1) \sigma^{\alpha_2-1} \left(\frac{\exp(s_2 \xi) - \exp(s_1 \xi)}{\xi} \right)^{\alpha_2-1} \\ &\quad \times \exp \left(-\beta_2 \sigma \frac{\exp(s_2 \xi) - \exp(s_1 \xi)}{\xi} \right) \Psi(\xi) \mathbb{1}_{\sigma, q_1 > 0},\end{aligned}$$

with

$$\Psi(\xi) := \frac{1}{\xi^2} |(s_1 \xi - 1) \exp(s_1 \xi) - (s_2 \xi - 1) \exp(s_2 \xi)|.$$

The posterior, with a single observation x such that

$$1 + \xi \left(\frac{x - \mu}{\sigma} \right) > 0,$$

is written

$$\begin{aligned}\pi_{\theta|x}^{\text{ULS}}(\mu, \sigma, \xi | x) &\propto q_1^{\alpha_1-1} \exp(-\beta_1 q_1) \sigma^{\alpha_2-1} \left(\frac{\exp(s_2 \xi) - \exp(s_1 \xi)}{\xi} \right)^{\alpha_2-1} \\ &\quad \times \exp \left(-\beta_2 \sigma \frac{\exp(s_2 \xi) - \exp(s_1 \xi)}{\xi} \right) \Psi(\xi) \\ &\quad \times \exp \left(-M \left\{ 1 + \xi \left(\frac{u - \mu}{\sigma} \right) \right\}^{-\frac{1}{\xi}} \right) \frac{1}{\sigma} \left\{ 1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right\}^{-\frac{\xi+1}{\xi}} \mathbb{1}_{\sigma, q_1 > 0}.\end{aligned}$$

In order to show that this is proper, we need to show that the following is integrable:

$$\begin{aligned}\pi_{\sigma|x, \mu, \xi}^{\text{ULS}}(\sigma | x, \mu, \xi) &\propto q_1^{\alpha_1-1} \exp(-\beta_1 q_1) \sigma^{\alpha_2-2} \exp \left(-\beta_2 \sigma \frac{\exp(s_2 \xi) - \exp(s_1 \xi)}{\xi} \right) \\ &\quad \times \exp \left(-M \left\{ 1 + \xi \left(\frac{u - \mu}{\sigma} \right) \right\}^{-\frac{1}{\xi}} \right) \left\{ 1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right\}^{-\frac{\xi+1}{\xi}} \mathbb{1}_{\sigma, q_1 > 0}.\end{aligned}$$

Northrop and Attalides [11] show that for a GEV likelihood model with at least 2 observations, if uniform priors are specified for μ and $\log \sigma$, and a proper prior is placed on ξ , then the joint prior will yield a proper posterior.

3.4 SAS: Spike-and-slab prior

This is a prior with a non-zero probability mass (spike) at $\xi = 0$, and a flat slab elsewhere.

We use the method proposed by Kuo and Mallick [9]. We introduce an indicator random variable γ such that $\gamma = 0$ when ξ is “in” the spike and $\gamma = 1$ when ξ is “in” the slab. The probability that ξ is nonzero can then be calculated as the mean of γ . We also need a random variable β such that $\xi = \beta\gamma$, which determines the distribution of ξ in the slab. Our model now has parameters $\eta := (\mu, \sigma, \beta, \gamma)$, and we suppose that γ does not depend on the other parameters.

If we have a prior π_θ for (μ, σ, ξ) , then we can set

$$\begin{aligned}\pi_{\mu, \sigma, \beta}^{\text{SAS}} &:= \pi_\theta \\ \pi_\gamma^{\text{SAS}}(\gamma) &:= \alpha^\gamma (1 - \alpha)^{1-\gamma} \\ \implies \pi_\eta^{\text{SAS}}(\eta) &= \pi_\theta(\mu, \sigma, \beta) \alpha^\gamma (1 - \alpha)^{1-\gamma},\end{aligned}$$

with $\alpha \in [0, 1]$. The posterior is then proportional to

$$\pi_{\eta|\mathbf{x}}^{\text{SAS}}(\eta | \mathbf{x}) \propto L(\mu, \sigma, \beta\gamma | \mathbf{x}) \pi_\theta(\mu, \sigma, \beta) \alpha^\gamma (1 - \alpha)^{1-\gamma}.$$

Since, up to the same constant of proportionality, we have

$$\pi_{\gamma|\mu, \sigma, \beta, \mathbf{x}}^{\text{SAS}}(0 | \mu, \sigma, \beta, \mathbf{x}) \propto (1 - \alpha) L(\mu, \sigma, 0 | \mathbf{x})$$

and

$$\pi_{\gamma|\mu, \sigma, \beta, \mathbf{x}}^{\text{SAS}}(1 | \mu, \sigma, \beta, \mathbf{x}) \propto \alpha L(\mu, \sigma, \beta | \mathbf{x}),$$

the full conditional of γ is a Bernoulli distribution with probability of success

$$\frac{\alpha L(\mu, \sigma, \beta | \mathbf{x})}{\alpha L(\mu, \sigma, \beta | \mathbf{x}) + (1 - \alpha) L(\mu, \sigma, 0 | \mathbf{x})}.$$

We set $\pi_\theta = \pi_\theta^{\text{QD}}$, and for an uninformative prior on γ , we set $\alpha = 0.5$.

3.5 QDE: Exponential priors for independent quantile differences

...

3.6 QDTN: Truncated normal priors for independent quantile differences

...

4 Prior informativeness

The following recent papers investigate prior informativeness in general settings: [7, 10, 12]. Are they meaningful for our model, our priors?

5 Sampling using MCMC

For the priors π^{QD} , π^{ME} , and π^{ULS} , we can sample from prior and posterior distributions of (μ, σ, ξ) using a Metropolis-Within-Gibbs algorithm with independent symmetrical proposal distributions for each parameter. We can reparametrise to $(\mu, \log(\sigma), \xi)$, which allows us to

Algorithm 1: Metropolis-Within-Gibbs with Normal proposal distributions

Input: function $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ such that $\exp f$ is proportional to the target density
 proposal distribution standard deviations $\sigma_j \in \mathbb{R}^+$ for $j = 1, 2, 3$
 initial state $\theta_0 \in \mathbb{R}^3$
 number of iterations $n \in \mathbb{N}$
 length of burn-in $b \in \mathbb{N}$

Output: samples $s^{(1)}, \dots, s^{(n-b)} \in \mathbb{R}^3$

begin

$(\theta, y) \leftarrow (\theta_0, f(\theta_0))$
for $i \in 1, \dots, n$ **do**

for $j \in 1, 2, 3$ **do**

$\theta_j^* \sim \mathcal{N}(\theta_j, \sigma_j)$
 $y^* \leftarrow f(\theta^*)$
 $u \sim \mathcal{U}(0, 1)$
if $y^* - y > \log(u)$ **then**

$(\theta, y) \leftarrow (\theta^*, y^*)$

if $i > b$ **then**

$s^{(i-b)} \leftarrow \theta$

use Normal distributions for each parameter. The algorithm is described in Alg. 1. For π^{SAS} , as we have access to the full conditional for γ , we can replace the Normal proposal density for γ by its full conditional. These algorithms are special cases of the Metropolis–Hastings algorithm, whose convergence is studied in Roberts and Smith [13]. Details of the implementation of the algorithm are shown in tables at the end of the report.

Once we have samples of $(\mu, \log(\sigma), \xi)$, we can obtain samples of the three quantiles (q_1, q_2, q_3) . We can then use kernel density estimation with Gaussian kernels to visualise the marginals of $(\mu, \log(\sigma), \xi)$ and (q_1, q_2, q_3) using line and contour plots. The smoothing from the density estimation leads to the constraints on the quantiles not always being satisfied. The marginals for π_{ξ}^{QD} and $\pi_{\log(\sigma), \xi}^{\text{QD}}$ do not have to be estimated as we can calculate them analytically. We can use the samples of the posterior of $(\mu, \log(\sigma), \xi)$ to plot the mean return level against the return period, as well as 95% credibility intervals.

6 Results

6.1 PPP: Poisson point process data

According to the model described in § 2, the exceedances of the threshold u approximately follow a non-homogeneous Poisson point process with intensity function

$$\lambda(x) = \frac{1}{\sigma} \left\{ 1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right\}_+^{-\frac{\xi+1}{\xi}}.$$

We will simulate data for which this approximation is exact, with known parameters. We set $M = 54$ and $n_u = 86$, with parameters $(\mu, \sigma, \xi) = (25, 5, 0.2)$, and we calculated the threshold using the equation

$$\begin{aligned} M\Lambda[u, \infty) &= n_u \\ \iff u &= \mu + \sigma \frac{\left(\frac{n_u}{M}\right)^{-\xi} - 1}{\xi}, \end{aligned}$$

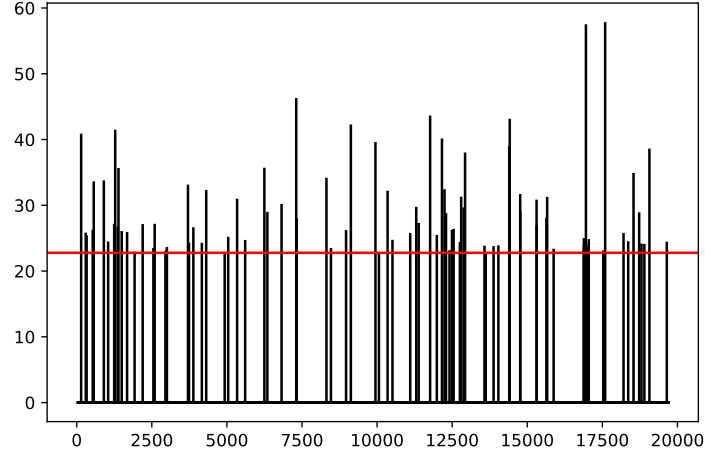


Figure 1: Data in Poisson point process simulation study with threshold $u = 22.778$ in red

with Λ defined in (4). This resulted in a threshold of $u = 22.778$. The intensity function is equal to the PDF of a Generalised Pareto distribution, except with support $\{x: x > u\}$. This means that we can simulate from the intensity by simulating from a Generalised Pareto distribution with parameters (u, σ, ξ) . We simulated 86 observations, and distributed them uniformly over the total time period of $365M$ observations, setting the remaining points to 0. This dataset, denoted \mathbf{x}^{PPP} , is illustrated in Fig. 1.

We constructed the priors with the knowledge of the true quantiles and quantile differences. Fixing $p = (0.1, 0.01, 0.001)$, then

$$\begin{aligned} q_1^* &= 39.211, & q_2^* &= 62.734, & q_3^* &= 99.517 \\ \implies \tilde{q}_1^* &= 39.211, & \tilde{q}_2^* &= 23.523, & \tilde{q}_3^* &= 36.783 \end{aligned}$$

are the true quantiles and quantile differences respectively. To construct π_q^{QD} , we chose

$$\begin{aligned} \tilde{q}_i &\sim \Gamma\left(\frac{(\tilde{q}_i^*)^2}{v}, \frac{\tilde{q}_i^*}{v}\right), \quad i = 1, 2, 3 \\ \implies \mathbb{E}(\tilde{q}_i) &= \tilde{q}_i^*, \quad \text{Var}(\tilde{q}_i) = v \end{aligned}$$

for $v = 100$.

The univariate and bivariate marginals of the priors are illustrated in Table ??, Table ??, and Table ??. For π^{SAS} , the posterior mean of γ was 0.467. We plotted the mean return level compared to the analytic true return level in Fig. 5. The dashed line was obtained by simulating 270,000 years of data and calculating the empirical quantiles.

6.2 PD: Pseudo-data

In order to compare our results to Coles et al. [6], we generated data, denoted \mathbf{x}^{PD} , from $\text{GEV}(17, 1.8, 0.28)$ to resemble their data, such that:

	\mathbf{x}^{PD}	Coles et al.
M	54	54
u	40.109	40
n_u	86	86

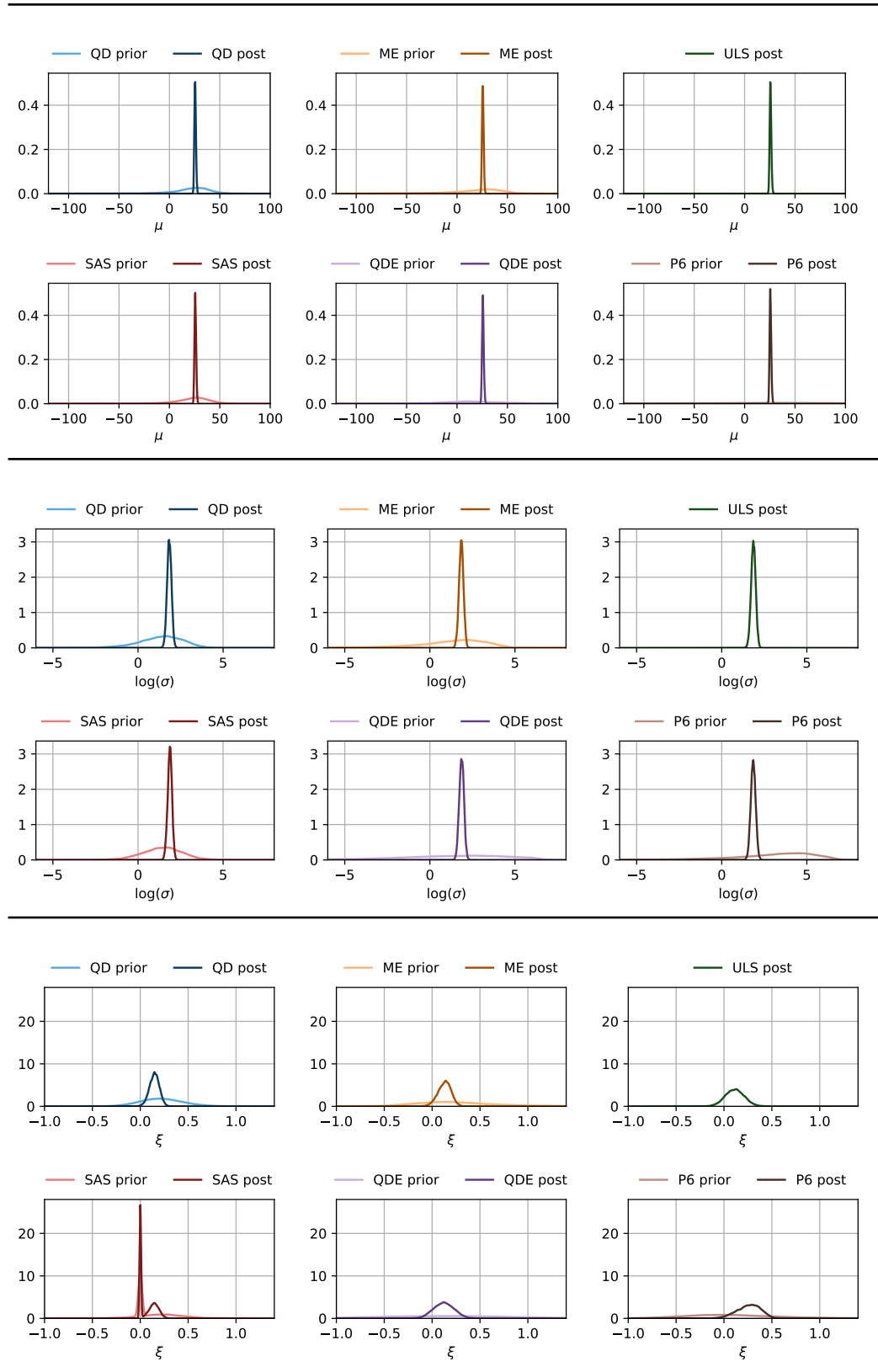


Table 1: Prior and posterior univariate marginals of (q_1, q_2, q_3) for Poisson process simulation study

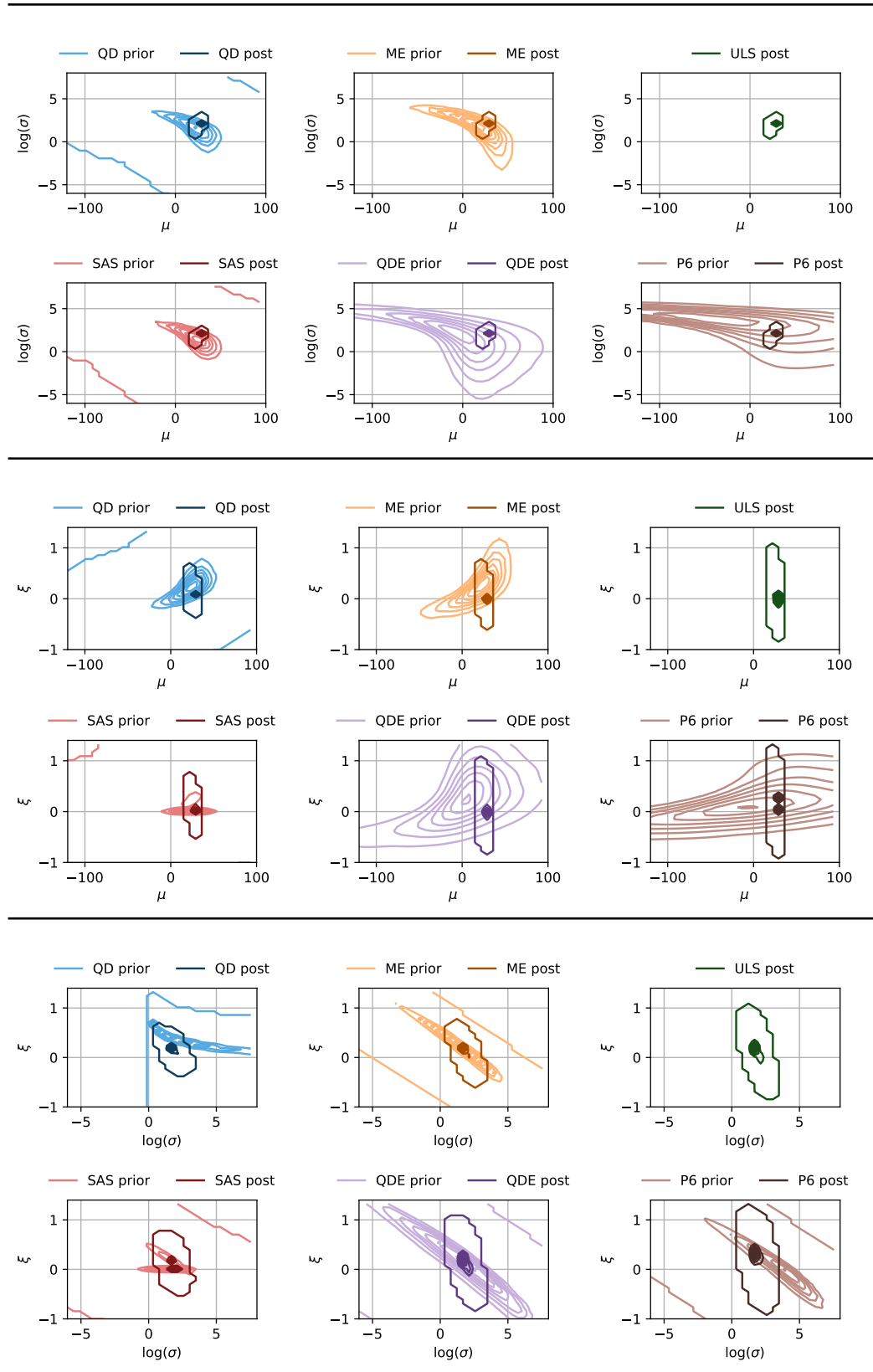


Table 2: Prior and posterior bivariate marginals of (q_1, q_2, q_3) for Poisson process simulation study

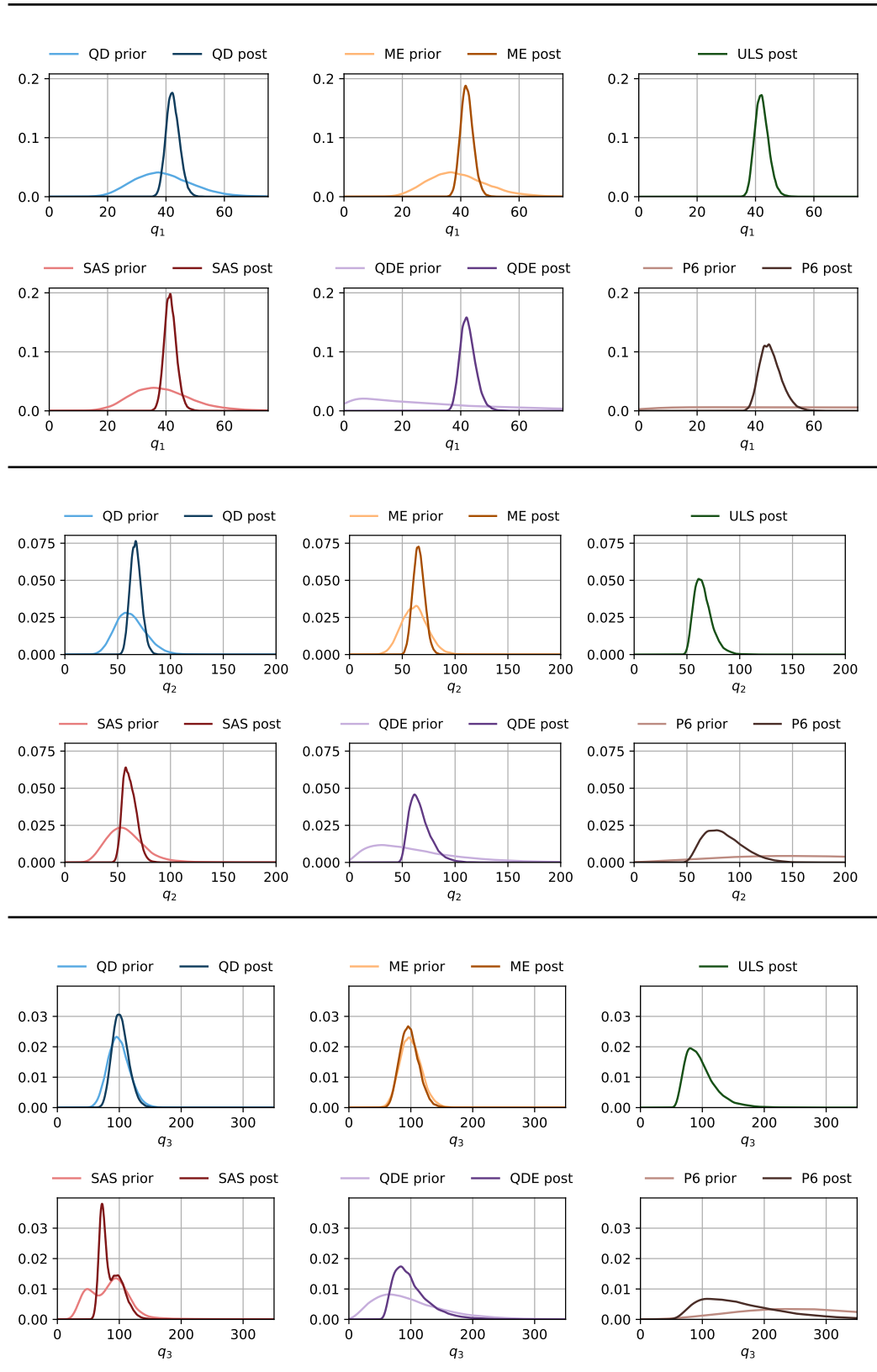


Table 3: Prior and posterior univariate marginals of $(\mu, \log \sigma, \xi)$ for Poisson process simulation study

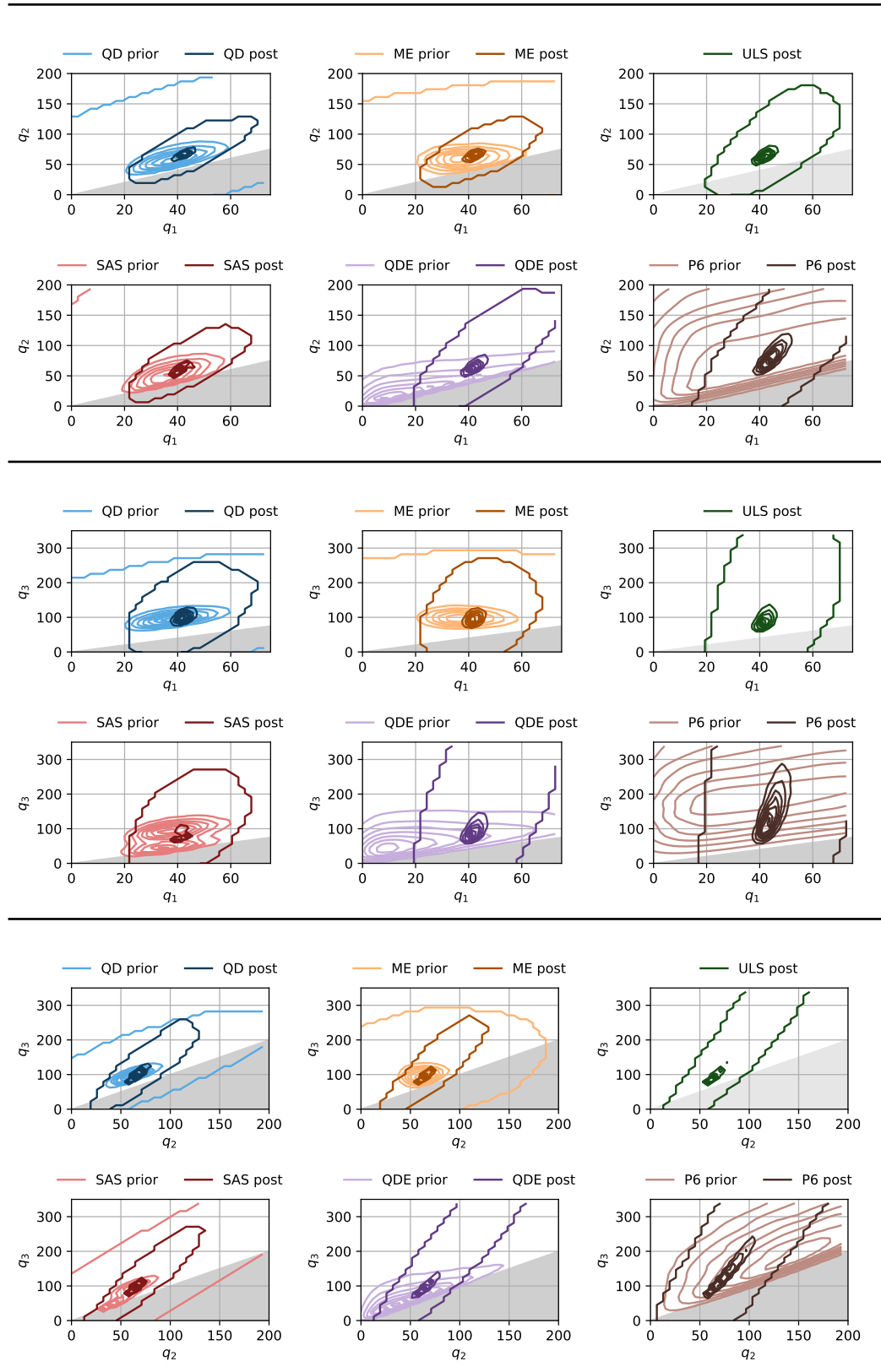


Table 4: Prior and posterior bivariate marginals of $(\mu, \log \sigma, \xi)$ for Poisson process simulation study

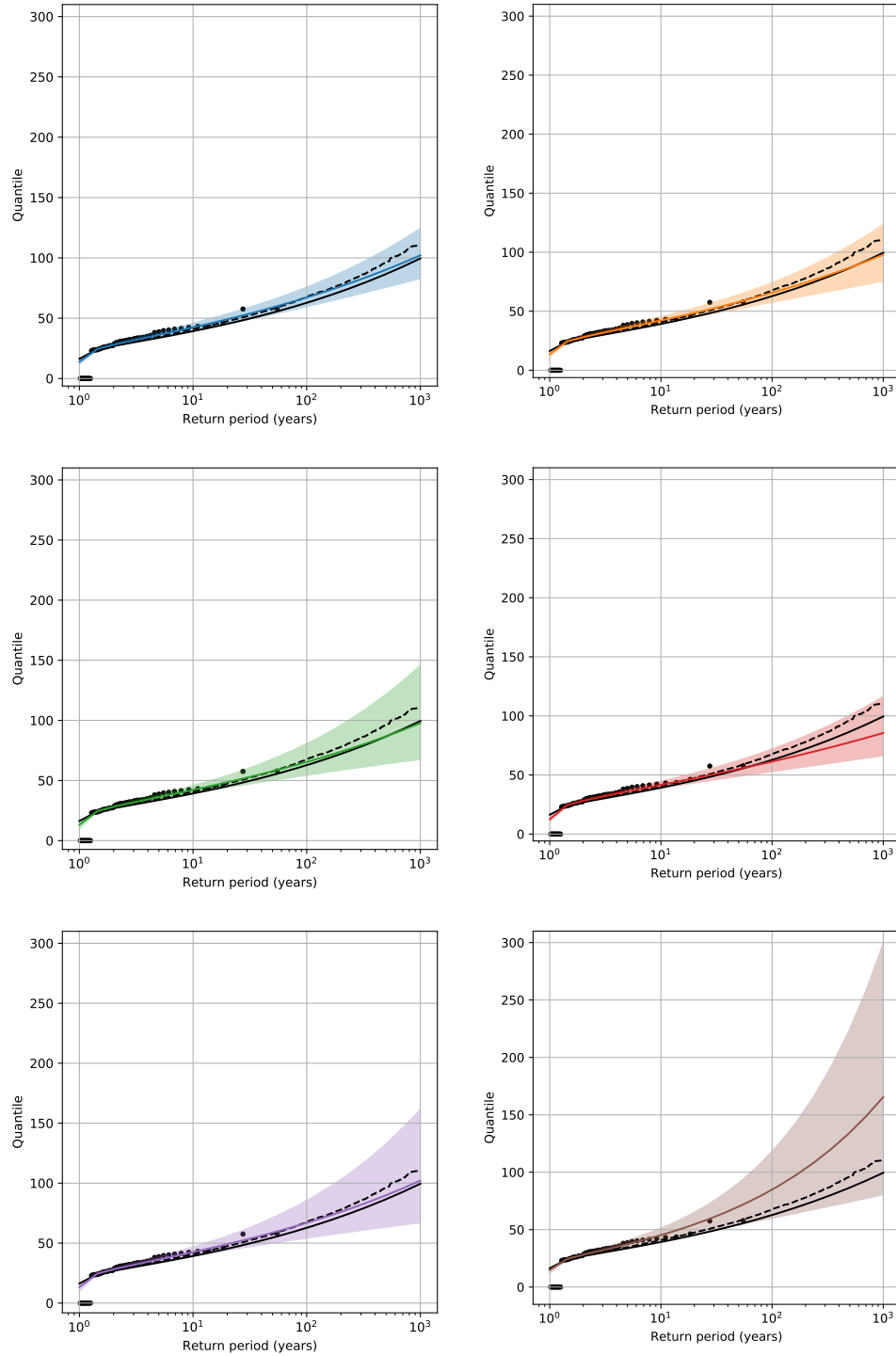


Table 5: Mean return level with 95% credibility intervals estimated using posterior distributions of π^{QD} (blue), π^{ME} (orange), π^{ULS} (green), and π^{SAS} (red), with analytic return level (black solid), simulated return level (black dashed) and empirical quantiles (black dots) for Poisson process simulation study

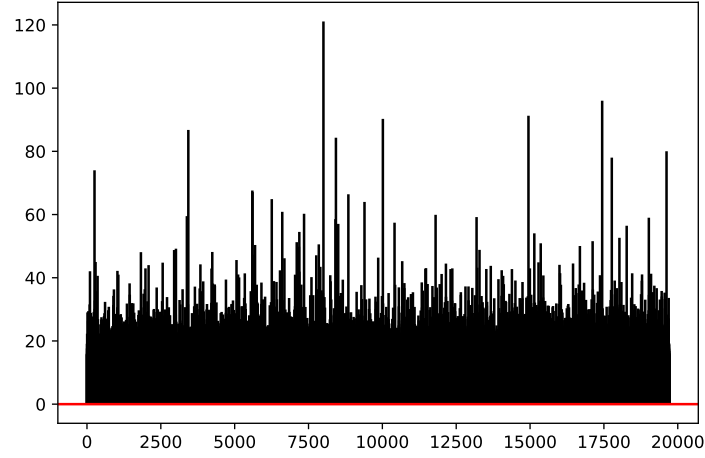


Figure 2: Data in pseudo-data simulation study with threshold $u = 40.109$ in red

	\tilde{q}_1	\tilde{q}_2	\tilde{q}_3
Expert	(57.563, 70.263)	(42.310, 66.805)	(119.09, 142.836)
π_{θ}^{QD}	(57.592, 70.239)	(41.652, 65.907)	(119.868, 143.807)
π_{θ}^{ME}	(57.774, 70.694)	(42.601, 71.7)	(121.04, 162.726)
$\pi_{\theta}^{\text{ULS}}$	(57.611, 70.266)	(42.191, 66.498)	—
$\pi_{\theta}^{\text{SAS}}$	(54.696, 67.987)	(26.169, 58.588)	(79.505, 135.024)

Table 6: Pairs of (median, 0.9-quantile) of quantile differences for pseudo-data simulation study

This dataset is illustrated in Fig. 2.

We used the same prior $\pi_{\tilde{q}}^{\text{QD}}$ for independent quantile differences as specified by Coles et al., given by

$$\begin{aligned}
 p &= (0.1, 0.01, 0.001), \\
 \tilde{q}_1 &\sim \Gamma(38.9, 0.67), \\
 \tilde{q}_2 &\sim \Gamma(7.1, 0.16), \\
 \tilde{q}_3 &\sim \Gamma(47, 0.39).
 \end{aligned}$$

The univariate and bivariate marginals of the priors are illustrated in Table ??, Table ??, and Table ?. For π^{SAS} , the posterior mean of γ was 0.638. In Table 6, we compare the statistics of the quantile differences specified by the expert, to the statistics estimated using the MCMC samples of (μ, σ, ξ) for each of the priors. In Fig. ??, the mean return level is plotted with 95% credibility intervals and empirical quantiles. The dashed black line was obtained by simulating 270,000 years of data and calculating the empirical quantiles. We know that the annual maximum has CDF F^{365} , where F is the CDF of $\text{GEV}(17, 1.8, 0.28)$, and the solid black line shows the quantiles calculated analytically from this CDF.

In Fig. 3, we varied the threshold, and plotted the number of exceedances against the return level for return periods 10^2 , 10^3 , and 10^4 , all estimated using the prior $\pi_{\theta|\mathbf{x}^{\text{PD}}}^{\text{QD}}$.

6.3 WS: Daily average wind speed at Tours

The data in Fig. 4 represents daily average wind speed at Tours over a period of X years, from 1981 to 2011. We chose a threshold of $u = 0$.

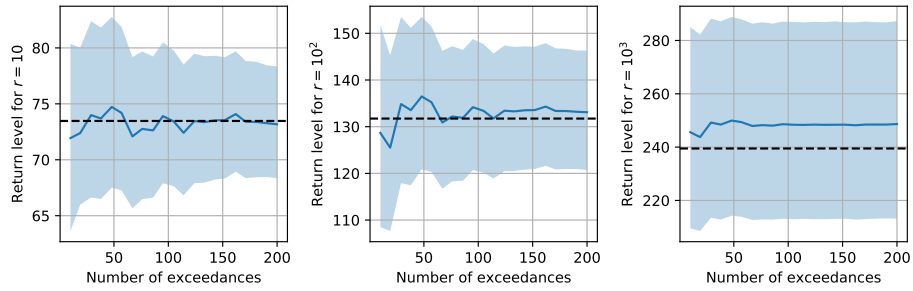


Figure 3: Mean return levels (blue) with 95% credibility intervals estimated using $\pi_{\theta|\mathbf{x}^{PD}}^{QD}$ for various thresholds, with simulated return levels (black dashed)

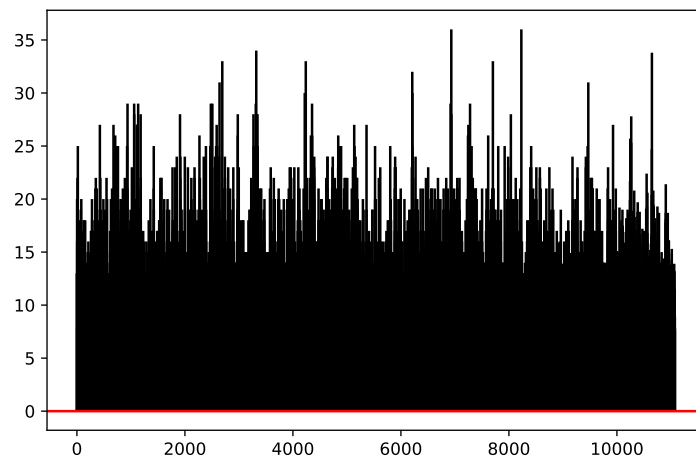


Figure 4: Daily average wind speed at Tours with threshold $u = 0$ in red

7 Conclusion

...

8 Appendix

8.1 Condition (C2) of the maximum entropy distribution for Weibull and Gamma distributed marginals

The construction of the maximum entropy distribution in § 3.2 requires the following condition on the marginals: for a sequence of distributions $(F_i)_{1 \leq i \leq d}$,

$$(C2) := \forall 1 \leq i \leq d-1 \forall x \in \{x: 1 > F_i(x), F_{i+1}(x) > 0\} (F_i(x) > F_{i+1}(x)).$$

We will investigate this condition for Weibull and Gamma distributed marginals.

8.1.1 Weibull distribution

Let $1 \leq i \leq d-1$. We have that

$$1 > F_i(x), F_{i+1}(x) > 0 \iff x > 0.$$

The CDF when $x > 0$ is

$$F_i(x) = 1 - \exp\left(-\left(\frac{x}{\lambda_i}\right)^{k_i}\right),$$

with $k_i, \lambda_i > 0$. Let $x > 0$. Then

$$\begin{aligned} (C2) &\iff 1 - \exp\left(-\left(\frac{x}{\lambda_i}\right)^{k_i}\right) > 1 - \exp\left(-\left(\frac{x}{\lambda_{i+1}}\right)^{k_{i+1}}\right) \\ &\iff \left(\frac{x}{\lambda_i}\right)^{k_i} > \left(\frac{x}{\lambda_{i+1}}\right)^{k_{i+1}}. \end{aligned}$$

If $k_i = k_{i+1} = k$,

$$\begin{aligned} \left(\frac{x}{\lambda_i}\right)^k > \left(\frac{x}{\lambda_{i+1}}\right)^k &\iff \frac{x}{\lambda_i} > \frac{x}{\lambda_{i+1}} \\ &\iff \frac{1}{\lambda_i} > \frac{1}{\lambda_{i+1}} \\ &\iff \lambda_i < \lambda_{i+1}. \end{aligned}$$

If $k_i \neq k_{i+1}$,

$$\begin{aligned} F_i(x) = F_{i+1}(x) &\iff \left(\frac{x}{\lambda_i}\right)^{k_i} = \left(\frac{x}{\lambda_{i+1}}\right)^{k_{i+1}} \\ &\iff \exp(k_i(\log x - \log \lambda_i)) = \exp(k_{i+1}(\log x - \log \lambda_{i+1})) \\ &\iff k_i(\log x - \log \lambda_i) = k_{i+1}(\log x - \log \lambda_{i+1}) \\ &\iff k_i \log x - k_i \log \lambda_i = k_{i+1} \log x - k_{i+1} \log \lambda_{i+1} \\ &\iff k_i \log x - k_{i+1} \log x = k_i \log \lambda_i - k_{i+1} \log \lambda_{i+1} \\ &\iff (k_i - k_{i+1}) \log x = k_i \log \lambda_i - k_{i+1} \log \lambda_{i+1} \end{aligned}$$

$$\begin{aligned}
&\Longleftrightarrow \log x = \frac{k_i \log \lambda_i - k_{i+1} \log \lambda_{i+1}}{k_i - k_{i+1}} \\
&\Longleftrightarrow x = \exp \left(\underbrace{\frac{k_i \log \lambda_i - k_{i+1} \log \lambda_{i+1}}{k_i - k_{i+1}}}_{=: h(k_i, \lambda_i, k_{i+1}, \lambda_{i+1})} \right).
\end{aligned}$$

Therefore the two CDFs intersect, and so (C2) cannot be satisfied. In conclusion,

$$(C2) \Longleftrightarrow (k_i = k_{i+1}) \wedge (\lambda_i < \lambda_{i+1}) \quad \forall 1 \leq i \leq d-1.$$

In practice however, as $y := h(k_i, \lambda_i, k_{i+1}, \lambda_{i+1}) \rightarrow \pm\infty$, the intersection will tend to the bounds of the set $\{x: 1 > F_i(x), F_{i+1}(x) > 0\}$ and so (C2) will be asymptotically satisfied. When does this happen? Consider a reparametrisation

$$(k_i, \lambda_i, k_{i+1}, \lambda_{i+1}) \mapsto (k_i, \lambda_i, k_i - k_{i+1}, \lambda_{i+1}),$$

which implies that

$$h(k_i, \lambda_i, z_i, \lambda_{i+1}) = \frac{k_i \log \lambda_i + (z_i - k_i) \log \lambda_{i+1}}{z_i}, \quad z_i < k_i, z_i \neq 0.$$

Then

- As $z_i \rightarrow 0, y \rightarrow +\infty$,
- As $k_i \rightarrow +\infty, y \rightarrow \pm\infty$, with sign depending on the sign of $\log(\frac{\lambda_i}{\lambda_{i+1}})$,
- As $\lambda_i \rightarrow +\infty$ or $\lambda_{i+1} \rightarrow 0, y \rightarrow +\infty$,
- As $\lambda_i \rightarrow 0$ or $\lambda_{i+1} \rightarrow +\infty, y \rightarrow -\infty$.

8.1.2 Gamma distribution

Let $1 \leq i \leq d-1$. We have that

$$1 > F_i(x), F_{i+1}(x) > 0 \Longleftrightarrow x > 0.$$

Suppose that we have two such distributions,

$$F_1(x) = \frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)} x^{\alpha_1} \exp(\beta_1 x) \quad \text{and} \quad F_2(x) = \frac{\beta_2^{\alpha_2}}{\Gamma(\alpha_2)} x^{\alpha_2} \exp(\beta_2 x),$$

with $\alpha_i, \beta_i > 0$ the shape and rate parameters. Denote their respective PDFs by f_1 and f_2 . Let $x > 0$, and define

$$\phi(x) = F_2(x) - F_1(x),$$

so that if $(i, i+1) = (1, 2)$,

$$(C2) \Longleftrightarrow \phi(x) < 0,$$

and if $(i, i+1) = (2, 1)$,

$$(C2) \Longleftrightarrow \phi(x) > 0.$$

We have that

$$\phi'(x) = f_2(x) - f_1(x)$$

$$\begin{aligned}
&= f_1(x) \left(\frac{f_2(x)}{f_1(x)} - 1 \right) \\
&= f_1(x) \left(\frac{\Gamma(\alpha_1)\beta_2^{\alpha_2}}{\Gamma(\alpha_2)\beta_1^{\alpha_1}} x^{\alpha_2-\alpha_1} \exp((\beta_1 - \beta_2)x) - 1 \right).
\end{aligned}$$

Let

$$C := \frac{\Gamma(\alpha_1)\beta_2^{\alpha_2}}{\Gamma(\alpha_2)\beta_1^{\alpha_1}} > 0$$

and

$$f(x) := Cx^{\alpha_2-\alpha_1} \exp((\beta_1 - \beta_2)x) - 1.$$

Then

$$\begin{aligned}
f'(x) &= C [(\alpha_2 - \alpha_1)x^{\alpha_2-\alpha_1-1} \exp((\beta_1 - \beta_2)x) + x^{\alpha_2-\alpha_1} \exp((\beta_1 - \beta_2)x)(\beta_1 - \beta_2)] \\
&= \underbrace{Cx^{\alpha_2-\alpha_1-1} \exp((\beta_1 - \beta_2)x)}_{>0} [(\alpha_2 - \alpha_1) + x(\beta_1 - \beta_2)]
\end{aligned}$$

Case 1: $(\alpha_1 \leq \alpha_2 \wedge \beta_1 \geq \beta_2) \wedge \neg(\alpha_1 = \alpha_2 \wedge \beta_1 = \beta_2)$

We have that $\alpha_2 - \alpha_1 \geq 0$ and $\beta_1 - \beta_2 \geq 0$, and therefore

$$\begin{aligned}
(\alpha_2 - \alpha_1) + x(\beta_1 - \beta_2) &> 0 \\
\implies f'(x) &> 0.
\end{aligned}$$

This implies that f is strictly increasing on \mathbb{R}^+ .

1. If $\alpha_1 < \alpha_2$, we have that

$$\lim_{x \rightarrow 0^+} f(x) = -1 \quad \text{and} \quad \lim_{x \rightarrow +\infty} f(x) = +\infty.$$

2. If $\alpha_1 = \alpha_2 = \alpha$ and $\beta_1 > \beta_2$, we have that

$$\lim_{x \rightarrow 0^+} f(x) = \left(\frac{\beta_2}{\beta_1} \right)^\alpha - 1 < 0 \quad \text{and} \quad \lim_{x \rightarrow +\infty} f(x) = +\infty.$$

Therefore in both cases, there exists a unique $z_0 > 0$ such that $f(z_0) = 0$. This implies that

$$\begin{cases} \phi'(x) < 0 & \text{if } x < z_0, \\ \phi'(x) = 0 & \text{if } x = z_0, \\ \phi'(x) > 0 & \text{if } x > z_0. \end{cases}$$

In order to proceed we will need the following Lemma.

Lemma 8.1. 1. Let $f: (a, b] \rightarrow \mathbb{R}$ be continuous on $(a, b]$ and differentiable on (a, b) , with $a \in \mathbb{R} \cup \{+\infty\}$, $b \in \mathbb{R}$. If for all $x \in (a, b)$, $f'(x) > 0$ (resp. $<$) and $\lim_{x \rightarrow a^+} f(x) = L \in \mathbb{R}$, then for all $z \in (a, b]$, $f(z) > L$ (resp. $<$).

2. Let $f: [a, b) \rightarrow \mathbb{R}$ be continuous on $[a, b)$ and differentiable on (a, b) , with $a \in \mathbb{R}$, $b \in \mathbb{R} \cup \{+\infty\}$. If for all $x \in (a, b)$, $f'(x) > 0$ (resp. $<$) and $\lim_{x \rightarrow b^-} f(x) = L \in \mathbb{R}$, then for all $z \in [a, b)$, $f(z) < L$ (resp. $>$).

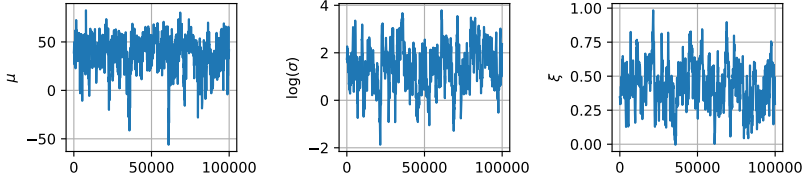
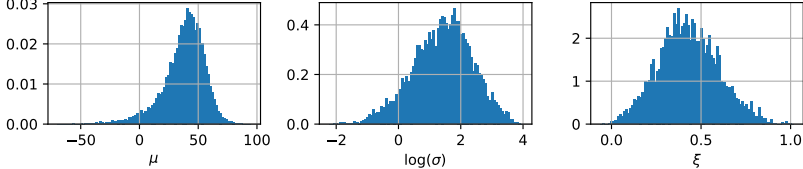
Proposal distr.	$\mu^{\text{new}} \sim \mathcal{N}(\mu^{\text{old}}, 40^2), \quad \log \sigma^{\text{new}} \sim \mathcal{N}(\log \sigma^{\text{old}}, 1.2^2), \quad \xi^{\text{new}} \sim \mathcal{N}(\xi^{\text{old}}, 0.3^2)$		
Initialisation	50, 1.5, 0.5		
Sample size	100,000 (after burn-in of 1,000)		
Traceplots			
Histograms			
Accept. rates	0.274, 0.135, 0.101		

Table 7: Metropolis algorithm for π_{θ}^{QD} for pseudo-data simulation study

Proof. We will prove the first case, for $f'(x) > 0$, as both cases are symmetrical. Suppose that $z \in (a, b)$. We need to show that $f(z), f(b) > L$. Let $y \in (a, z)$. As f' is strictly positive on (a, b) , f is strictly increasing (by the mean value theorem), and so $f(z) - f(y) = c > 0$. Therefore $f(z) > f(y) + \frac{c}{2}$. Then $f(z) = \lim_{y \rightarrow a^+} f(z) \geq \lim_{y \rightarrow a^+} f(y) + \frac{c}{2} = L + \frac{c}{2} > L$. Furthermore, $f(b) = \lim_{z \rightarrow b^-} f(z) \geq \lim_{x \rightarrow b^-} (L + \frac{c}{2}) = L + \frac{c}{2} > L$. \square

By Lemma 8.1, for all $z \in (0, z_0]$, $\phi(z) < \lim_{x \rightarrow 0^+} \phi(x) = 0$, and for all $z \in [z_0, +\infty)$, $\phi(z) < \lim_{x \rightarrow +\infty} \phi(x) = 0$. Therefore for all $x > 0$, $\phi(x) < 0$, and so if $(i, i+1) = (1, 2)$, (C2) is satisfied, and if $(i, i+1) = (2, 1)$, (C2) is violated.

Case 2: $\alpha_1 < \alpha_2, \beta_1 < \beta_2$

Since

$$\lim_{x \rightarrow 0^+} f(x) = -1,$$

there exists a $\delta > 0$ such that for all $z \in (0, \delta]$, $f(z) = \phi'(z) < 0$, and so by Lemma 8.1, $\phi(z) < \lim_{x \rightarrow 0^+} \phi(x) = 0$. Since

$$\lim_{x \rightarrow +\infty} f(x) = -1,$$

there exists a $\delta' > 0$ such that for all $z \in [\delta', +\infty)$, $f(z) = \phi'(z) < 0$, and so by Lemma 8.1, $\phi(z) > \lim_{x \rightarrow 0^+} \phi(x) = 0$. Therefore ϕ is neither strictly positive nor strictly negative, and so if either $(i, i+1) = (1, 2)$ or $(i, i+1) = (2, 1)$, (C2) is violated.

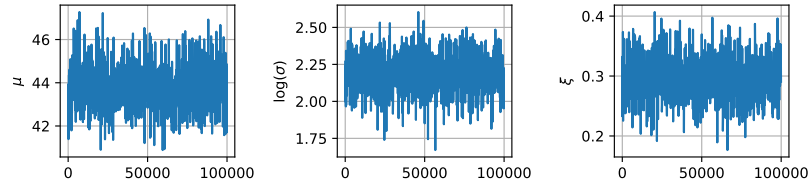
In conclusion, we have shown that

$$(C2) \iff (\alpha_i \leq \alpha_{i+1} \wedge \beta_i \geq \beta_{i+1}) \wedge \neg(\alpha_i = \alpha_{i+1} \wedge \beta_i = \beta_{i+1}) \quad \forall 1 \leq i \leq d-1.$$

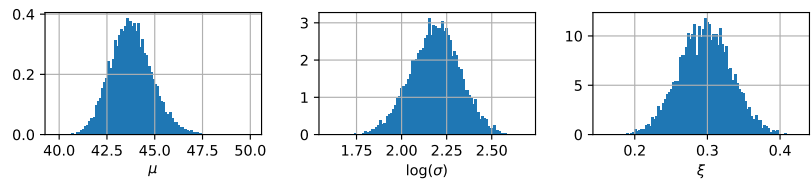
Like in the Weibull case, the condition can be asymptotically satisfied. For instance, if the means $\frac{\alpha}{\beta}$ are sufficiently separated, and the variances $\frac{\alpha}{\beta^2}$ are small, there will be less significant overlap between the CDFs.

Proposal distr.	$\mu^{\text{new}} \sim \mathcal{N}(\mu^{\text{old}}, 4^2), \quad \log \sigma^{\text{new}} \sim \mathcal{N}(\log \sigma^{\text{old}}, 0.5^2), \quad \xi^{\text{new}} \sim \mathcal{N}(\xi^{\text{old}}, 0.3^2)$
Initialisation	43, 2.2, 0.25
Sample size	100,000 (after burn-in of 1,000)

Traceplots



Histograms

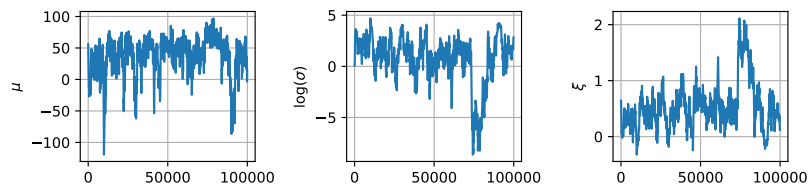


Accept. rates 0.25, 0.2, 0.101

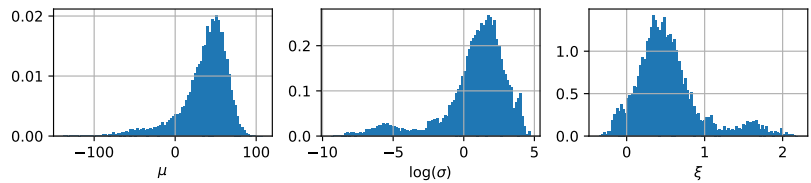
Table 8: Metropolis algorithm for $\pi_{\theta|\mathbf{x}^{\text{PD}}}^{\text{OD}}$ for pseudo-data simulation study

Proposal distr.	$\mu^{\text{new}} \sim \mathcal{N}(\mu^{\text{old}}, 40^2), \quad \log \sigma^{\text{new}} \sim \mathcal{N}(\log \sigma^{\text{old}}, 0.6^2), \quad \xi^{\text{new}} \sim \mathcal{N}(\xi^{\text{old}}, 0.2^2)$
Initialisation	50, 1.5, 0.5
Sample size	100,000 (after burn-in of 1,000)

Traceplots



Histograms

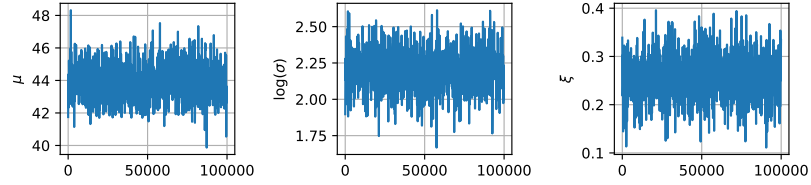


Accept. rates 0.242, 0.251, 0.169

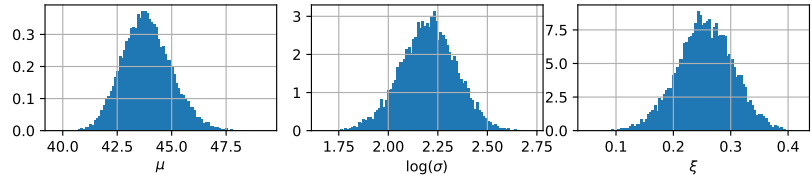
Table 9: Metropolis algorithm for π_{θ}^{ME} for pseudo-data simulation study

Proposal distr.	$\mu^{\text{new}} \sim \mathcal{N}(\mu^{\text{old}}, 4^2)$, $\log \sigma^{\text{new}} \sim \mathcal{N}(\log \sigma^{\text{old}}, 0.5^2)$, $\xi^{\text{new}} \sim \mathcal{N}(\xi^{\text{old}}, 0.3^2)$
Initialisation	43, 2.2, 0.25
Sample size	100,000 (after burn-in of 1,000)

Traceplots



Histograms

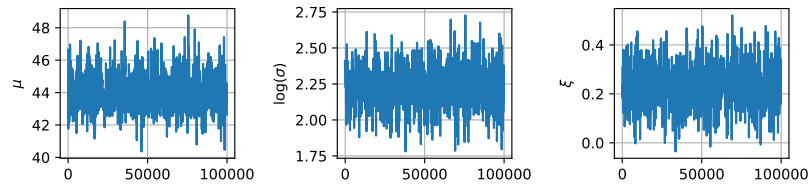


Accept. rates 0.261, 0.195, 0.126

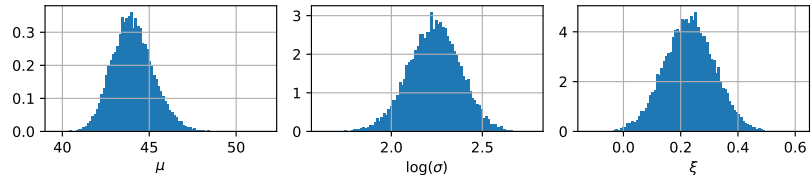
Table 10: Metropolis algorithm for $\pi_{\theta|\mathbf{x}_{\text{PD}}}^{\text{ME}}$ for pseudo-data simulation study

Proposal distr.	$\mu^{\text{new}} \sim \mathcal{N}(\mu^{\text{old}}, 4^2)$, $\log \sigma^{\text{new}} \sim \mathcal{N}(\log \sigma^{\text{old}}, 0.5^2)$, $\xi^{\text{new}} \sim \mathcal{N}(\xi^{\text{old}}, 0.3^2)$
Initialisation	43, 2.2, 0.25
Sample size	100,000 (after burn-in of 1,000)

Traceplots



Histograms

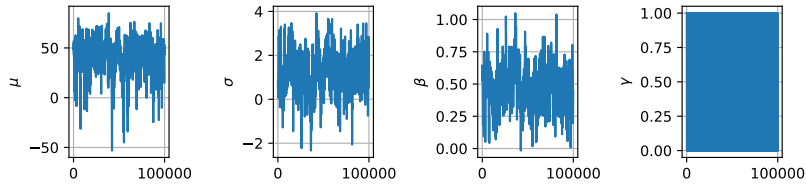


Accept. rates 0.271, 0.235, 0.275

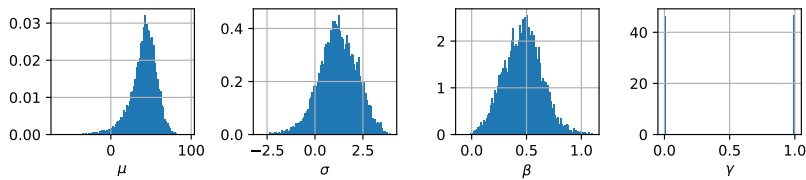
Table 11: Metropolis algorithm for $\pi_{\theta|\mathbf{x}_{\text{PD}}}^{\text{ULS}}$ for pseudo-data simulation study

Proposal distr.	$\mu^{\text{new}} \sim \mathcal{N}(\mu^{\text{old}}, 40^2), \quad \log \sigma^{\text{new}} \sim \mathcal{N}(\log \sigma^{\text{old}}, 0.6^2), \quad \xi^{\text{new}} \sim \mathcal{N}(\xi^{\text{old}}, 0.2^2)$
Initialisation	50, 1.5, 0.5, 1
Sample size	100,000 (after burn-in of 1,000)

Traceplots



Histograms

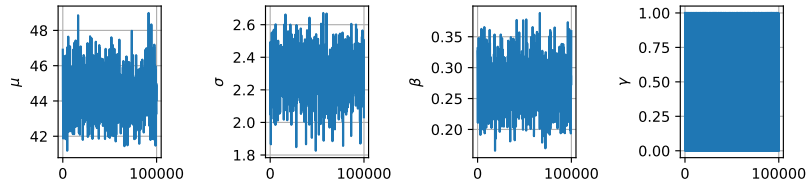


Accept. rates 0.274, 0.26, 0.151, 1

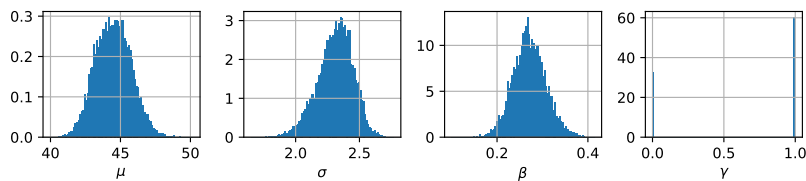
Table 12: Metropolis algorithm for $\pi_{\theta}^{\text{SAS}}$ for pseudo-data simulation study

Proposal distr.	$\mu^{\text{new}} \sim \mathcal{N}(\mu^{\text{old}}, 5^2), \quad \log \sigma^{\text{new}} \sim \mathcal{N}(\log \sigma^{\text{old}}, 0.5^2), \quad \xi^{\text{new}} \sim \mathcal{N}(\xi^{\text{old}}, 0.3^2)$
Initialisation	43, 2.2, 0.25, 1
Sample size	100,000 (after burn-in of 1,000)

Traceplots



Histograms

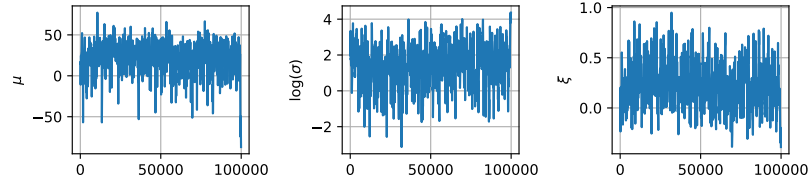


Accept. rates 0.25, 0.204, 0.107, 1

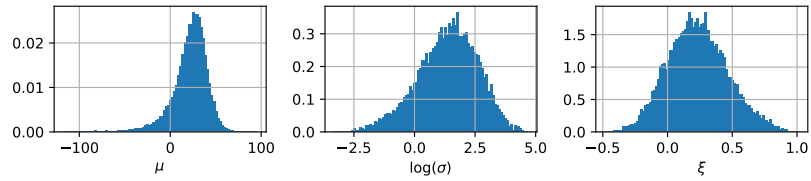
Table 13: Metropolis algorithm for $\pi_{\theta|\mathbf{x}^{\text{PD}}}^{\text{SAS}}$ for pseudo-data simulation study

Proposal distr.	$\mu^{\text{new}} \sim \mathcal{N}(\mu^{\text{old}}, 40^2), \quad \log \sigma^{\text{new}} \sim \mathcal{N}(\log \sigma^{\text{old}}, 1^2), \quad \xi^{\text{new}} \sim \mathcal{N}(\xi^{\text{old}}, 0.2^2)$
Initialisation	25, 0.2, 0.25
Sample size	100,000 (after burn-in of 1,000)

Traceplots



Histograms

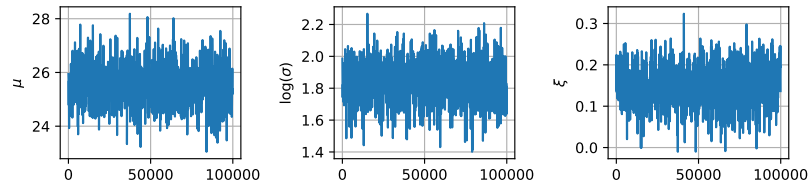


Accept. rates 0.288, 0.251, 0.264

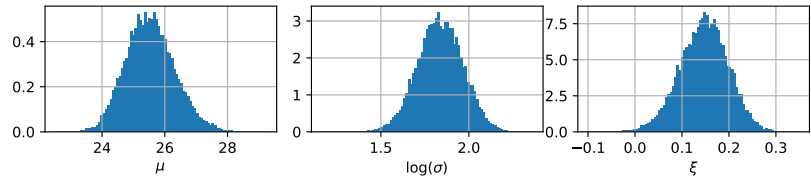
Table 14: Metropolis algorithm for π_{θ}^{OD} for Poisson process simulation study

Proposal distr.	$\mu^{\text{new}} \sim \mathcal{N}(\mu^{\text{old}}, 4^2), \quad \log \sigma^{\text{new}} \sim \mathcal{N}(\log \sigma^{\text{old}}, 0.5^2), \quad \xi^{\text{new}} \sim \mathcal{N}(\xi^{\text{old}}, 0.2^2)$
Initialisation	26, 1.8, 0.15
Sample size	100,000 (after burn-in of 1,000)

Traceplots



Histograms

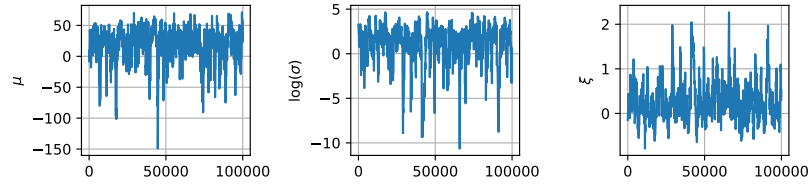


Accept. rates 0.194, 0.224, 0.257

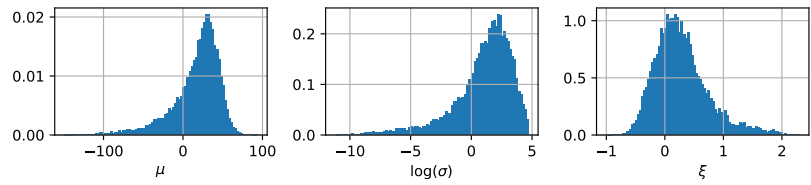
Table 15: Metropolis algorithm for $\pi_{\theta|\mathbf{x}^{\text{pp}}}^{\text{OD}}$ for Poisson process simulation study

Proposal distr.	$\mu^{\text{new}} \sim \mathcal{N}(\mu^{\text{old}}, 40^2)$, $\log \sigma^{\text{new}} \sim \mathcal{N}(\log \sigma^{\text{old}}, 1^2)$, $\xi^{\text{new}} \sim \mathcal{N}(\xi^{\text{old}}, 0.3^2)$
Initialisation	25, 0.2, 0.25
Sample size	100,000 (after burn-in of 1,000)

Traceplots



Histograms

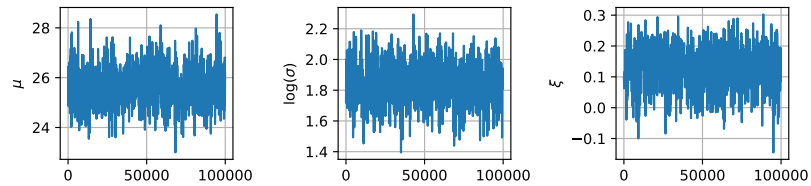


Accept. rates 0.244, 0.244, 0.218

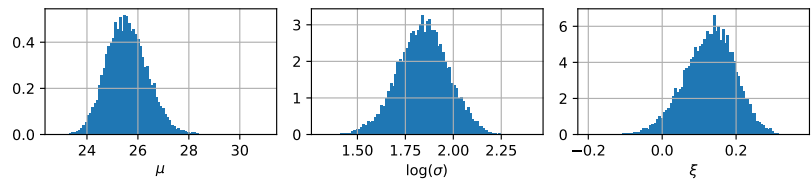
Table 16: Metropolis algorithm for π_{θ}^{ME} for Poisson process simulation study

Proposal distr.	$\mu^{\text{new}} \sim \mathcal{N}(\mu^{\text{old}}, 4^2)$, $\log \sigma^{\text{new}} \sim \mathcal{N}(\log \sigma^{\text{old}}, 0.5^2)$, $\xi^{\text{new}} \sim \mathcal{N}(\xi^{\text{old}}, 0.3^2)$
Initialisation	26, 1.8, 0.15
Sample size	100,000 (after burn-in of 1,000)

Traceplots



Histograms

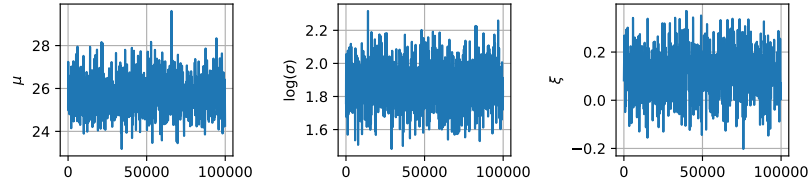


Accept. rates 0.199, 0.223, 0.212

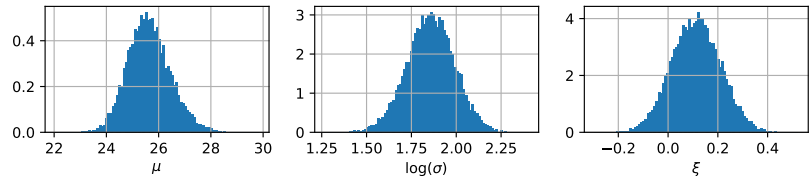
Table 17: Metropolis algorithm for $\pi_{\theta|\mathbf{x}^{\text{pp}}}^{\text{ME}}$ for Poisson process simulation study

Proposal distr.	$\mu^{\text{new}} \sim \mathcal{N}(\mu^{\text{old}}, 4^2), \quad \log \sigma^{\text{new}} \sim \mathcal{N}(\log \sigma^{\text{old}}, 0.5^2), \quad \xi^{\text{new}} \sim \mathcal{N}(\xi^{\text{old}}, 0.4^2)$
Initialisation	26, 1.8, 0.15
Sample size	100,000 (after burn-in of 1,000)

Traceplots



Histograms

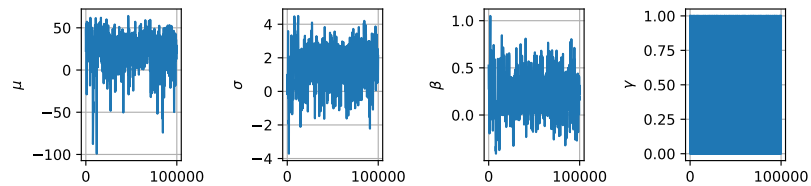


Accept. rates 0.204, 0.234, 0.252

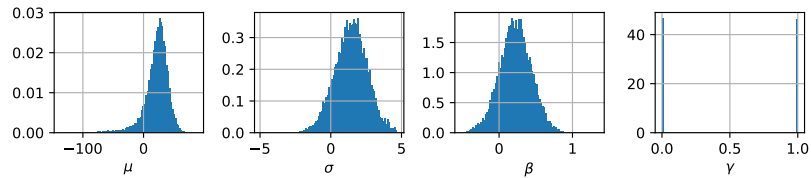
Table 18: Metropolis algorithm for $\pi_{\theta|\mathbf{x}^{\text{PPP}}}^{\text{ULS}}$ for Poisson process simulation study

Proposal distr.	$\mu^{\text{new}} \sim \mathcal{N}(\mu^{\text{old}}, 40^2), \quad \log \sigma^{\text{new}} \sim \mathcal{N}(\log \sigma^{\text{old}}, 1^2), \quad \xi^{\text{new}} \sim \mathcal{N}(\xi^{\text{old}}, 0.2^2)$
Initialisation	25, 0.2, 0.25, 1
Sample size	100,000 (after burn-in of 1,000)

Traceplots



Histograms



Accept. rates 0.29, 0.25, 0.261, 1

Table 19: Metropolis algorithm for $\pi_{\theta}^{\text{SAS}}$ for Poisson process simulation study

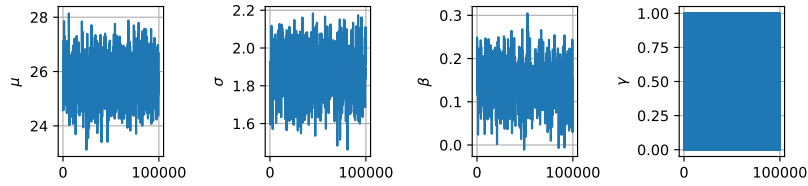
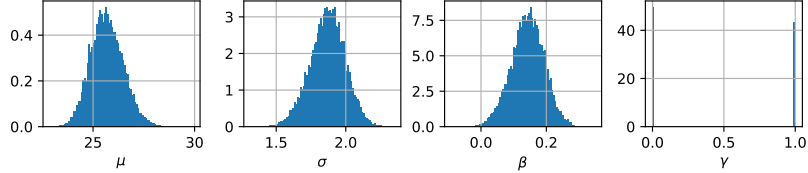
Proposal distr.	$\mu^{\text{new}} \sim \mathcal{N}(\mu^{\text{old}}, 4^2)$, $\log \sigma^{\text{new}} \sim \mathcal{N}(\log \sigma^{\text{old}}, 0.5^2)$, $\xi^{\text{new}} \sim \mathcal{N}(\xi^{\text{old}}, 0.2^2)$		
Initialisation	26, 1.8, 0.15, 1		
Sample size	100,000 (after burn-in of 1,000)		
Traceplots			
Histograms			
Accept. rates	0.205, 0.225, 0.259, 1		

Table 20: Metropolis algorithm for $\pi_{\theta|\mathbf{x}^{\text{PPP}}}^{\text{SAS}}$ for Poisson process simulation study

References

- [1] Michaël Baudin, Anne Dutfoy, Bertrand Iooss, and Anne-Laure Popelin. Openturns: An industrial software for uncertainty quantification in simulation. 2015.
- [2] Cristina Butucea, Jean-François Delmas, Anne Dutfoy, and Richard Fischer. Maximum entropy distribution of order statistics with given marginals. *Bernoulli*, 24(1):115–155, 2018.
- [3] Stuart Coles, Luis Pericchi, and Scott Sisson. A fully probabilistic approach to extreme rainfall modeling. *Journal of Hydrology*, 273:35–50, 03 2003.
- [4] Stuart G. Coles. *A Point Process Characterization of Extremes*. In: *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics. Springer, London, 2001.
- [5] Stuart G. Coles and Elwyn A. Powell. Bayesian methods in extreme value modelling: A review and new developments. *International Statistical Review / Revue Internationale de Statistique*, 64(1):119–136, 1996.
- [6] Stuart G. Coles and Jonathan A. Tawn. A Bayesian analysis of extreme rainfall data. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 45(4):463–478, 1996.
- [7] Fatemeh Ghaderinezhad and Christophe Ley. A general measure of the impact of priors in bayesian statistics via stein’s method. *arXiv preprint arXiv:1803.00098*, 2018.
- [8] E. T. Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 106:620–630, May 1957.
- [9] Lynn Kuo and Bani Mallick. Variable selection for regression models. *Sankhyā: The Indian Journal of Statistics, Series B (1960-2002)*, 60(1):65–81, 1998.
- [10] Ulrich K Müller. Measuring prior sensitivity and prior informativeness in large bayesian models. *Journal of Monetary Economics*, 59(6):581–597, 2012.

- [11] Paul J. Northrop and Nicolas Attalides. Posterior propriety in bayesian extreme value analyses using reference priors. *Statistica Sinica*, 26(2):721–743, 2016.
- [12] Matthew Reimherr, Xiao-Li Meng, and Dan L. Nicolae. Prior sample size extensions for assessing prior informativeness and prior–likelihood discordance, 2014.
- [13] G.O. Roberts and A.F.M. Smith. Simple conditions for the convergence of the gibbs sampler and metropolis-hastings algorithms. *Stochastic Processes and their Applications*, 49(2):207 – 216, 1994.
- [14] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.