

# C++ implementation for the paper Bayesian hierarchical clustering of extreme precipitation

Christian Rohrbeck and Jonathan A. Tawn

May 8, 2019

## 1 Introduction

The executable `Cluster` can be generated using the provided Makefile. Then, the software starts by alling `./Cluster` directory. We describe which files have to be specified in order to run the software to perform Bayesian clustering of precipitation extremes via a reversible jump MCMC algorithm. Together with the code, we provide the files used in the simulation study. Section 2 details the required files while Section 3 describes the result files generated by the implementation.

## 2 Input

All input files have to be in the same directory which has to be specified in the terminal when executing the `./Cluster` command.

### 2.1 Setup.txt

This file specifies the settings for the reversible jump MCMC algorithm. The following have to be defined

- Number of iterations, the burn-in period and the rate with which samples are stored.
- Proposal probabilities for the different moves. These don't have to add up to 1 but have to be positive.
- The initial number of clusters with which the algorithm starts.
- Block sizes for the likelihood adjustment. For instance, a value of 10 means that the covariance matrix of the gradient of the log-likelihood function is estimated based on splitting the observations into blocks of time length 10.

## 2.2 PeaksOverThreshold.txt

This file provides the observed threshold exceedances for all sites. The  $k$ th row contains all observations for site  $k$ , separated by a space.

## 2.3 P.txt and Q.txt

This file specifies the empirical pairwise measures for extremal dependence. Each is a  $K \times K$  matrix, where  $K$  is the total number of sites.

## 2.4 Neighbours.txt and Distances.txt

These provide the neighbourhodd structure and the distances between sites. Neighbours.txt is in the form of a list while distances takes the form of a  $K \times K$  matrix. Distances.txt does have to exist but we provided it in all simulations and case studies.

# 3 Output

The ouput files are stored in the same directory in which the input files are provided.

## 3.1 Sigma.txt, Xi.txt and Epsilon.txt

This files provides the sampled parameters of the cluster parameters  $\epsilon$ ,  $\sigma$  and  $\xi$  for each site. Each line corresponds to one site.

## 3.2 Number of clusters and hyperparameters.txt

This files gives the sampled number of clusters and cluster-independent parameters  $\gamma_0$ ,  $\beta$  and  $\kappa$ . Each line corresponds to one sample.

## 3.3 Similarity.txt

The pairwise frequency with which sites are allocated to the same cluster is stored herein.