Bayesian Spatial Hierarchical Modeling of Maximum Temperature

1Indriati Bisono
2Aloke Phatak
1Andrew Robinson

1University of Melbourne, 2CSIRO Maths, Informatics and Statistics WA
Outline

- Introduction
  - Extreme Values
  - Spatial Hierarchical Models
- The Data
- The Model
  - Data Level
  - Process Level
  - Prior Level
  - Implementation
- Results
- Conclusion
Introduction
Extreme Values
Hierarchical Spatial Models
Extreme Values

Data and distribution

- **Block maxima**, follows generalized extreme value (GEV) distribution;
  \[
  G(x) = \exp \left\{ -\left(1 + \frac{\xi(x - \mu)}{\sigma}\right)^{-1/\xi}\right\}
  \]

- The three parameters \((\mu, \sigma, \xi)\) are called **location**, **scale** and **shape** parameters.

- **Exceedances over threshold**, another type of extremes, follow generalized Pareto distribution (GPD).

Parameter Estimation

- **Maximum Likelihood.** A common and inexpensive method in estimation parameter

- **L-moments.** Expectations of certain **linear combinations of order statistics** that work analogously to method of moments, in some cases might **outperform** the excellent MLE (Hosking (1990).

- **Bayesian approach.** Suitable method for hierarchical model.
Hierarchical Gaussian Spatial Model

- Let $Y = (Y(\tilde{s}_1, \ldots, \tilde{s}_m))^T$ are $m$ observations of a spatial process, and define a latent spatial vector, $\eta = (\eta(s_1), \ldots, \eta(s_n))^T$ where $\eta(s_i)$ is from a Gaussian spatial process and $s_i$ is a spatial location. Observation locations $\{\tilde{s}_1, \ldots, \tilde{s}_m\}$ are not necessarily coincide with $\{s_1, \ldots, s_n\}$.

- A common hierarchical spatial model consist of three stages
  
  **Data model**: $Y|\beta, \eta, \sigma^2_\epsilon \sim Gau(X\beta + H\eta, \sigma^2_\epsilon I)$
  
  **Process model**: $\eta|\theta \sim Gau(0, \Sigma(\theta))$
  
  **Parameter model**: $[\beta, \sigma^2_\epsilon, \theta]$

- Bayes’ rule
  
  $P(\text{process, parameters}|\text{data})$
  
  $P(\text{data}|\text{process, parameters})P(\text{process}|\text{parameters})P(\text{parameters})$
The Data were retrieved from the Tasmanian Partnership of Advanced Computing portal:
The data

- **6 sets** monthly maximum temperature data, each are simulation RCM data driven by 6 different GCMs output.
- Choose **maximum** temperature for each year from 1961 to 2009. In that case, the data are simulated as if from 'control' run; no increase in CO2.
- The spatial domain are **56 by 51**, yielding 2856 grid cells covering Tasmania (Latitude: 44° – 39°S, Longitude: 143.5° – 154°E).

**Climate Future for Tasmania** used the CSIRO downscaling application CCAM to derive climate processes over Tasmania at scales of **10-15 km**.

Histogram of monthly and annual maximum temperature at some randomly chosen locations.
Annual time series data

Monthly time series data
The Model

follows Schliep, E. et.al. (2010) model
using three stages:
data – process – prior level
Data Level

- **Assume**, \( P(Z_{it} \leq z) = \exp \left\{ -\left( 1 + \xi_i \frac{z - \mu_i}{\sigma_i} \right)^{-1/\xi_i} \right\} \)

  provided \( \left( 1 + \xi_i \frac{z - \mu_i}{\sigma_i} \right) > 0 \) for each \( i \) and where \( \mu_i, \sigma_i \) and \( \xi_i \) are unknown location, scale and shape parameters in grid cell \( i \) respectively.

- The **likelihood** of the data, assuming that \( Z_{it} \) given \( \mu_i, \sigma_i \) and \( \xi_i \) is conditionally independent to \( Z_{jt} \) given \( \mu_j, \sigma_j \) and \( \xi_j \) for \( i \neq j \), with Martins Stedinger penalized:

\[
\Pi \left( z_i | (\mu_i, \sigma_i, \xi_i) \right) = K \prod_{i=1}^{d} \prod_{t=1}^{49} \exp \left\{ - \left[ 1 + \xi_i \left( \frac{z_{it} - \mu_i}{\sigma_i} \right) \right]^{-1/\xi_i} \right\} \\
\times \frac{1}{\sigma_i} \left[ 1 + \xi_i \left( \frac{z_{it} - \mu_i}{\sigma_i} \right) \right]^{-1/\xi_i - 1} \\
\times \frac{\Gamma(15)}{\Gamma(9)\Gamma(6)} \left( 0.5 + \xi_i \right)^8 \left( 0.5 - \xi_i \right)^5
\]
Process Level

- Process level

\[ \mu_i \sim N \left( X_i^T \beta_\mu + U_{i\mu}, \frac{1}{\tau_\mu^2} \right) \]

\[ \log(\sigma_i) \sim N \left( X_i^T \beta_\sigma + U_{i\sigma}, \frac{1}{\tau_\sigma^2} \right) \]

\[ \xi_i \sim N \left( X_i^T \beta_\xi + U_{i\xi}, \frac{1}{\tau_\xi^2} \right) \]

where

- \( X_i \) is the covariate function for grid \( i \),
- \( \beta_\theta \) is a vector of regression coefficients,
- \( \tau_\theta^2 \) is a fixed precision values,
- \( U_\theta = \{ U_{1\theta}, \ldots, U_{d\theta} \} \) is a spatial random effects
- \( \theta \) is generically used to stand for \( \mu, \sigma, \) and \( \xi \)

- Random effect \( U_i \) were modelled spatially using multivariate intrinsic autoregression (IAR) model (Banerjee et al. 2004). IAR is a special case of conditional autoregressive (CAR) model.

- The relationship between \( U \), \( U \), and \( U \) is depicted through precision matrix.
Parameter Level

- For the two hyperparameters $\beta$ and $T$ we chose conjugate priors.
  - Beta priors $\beta_0 \sim N(\hat{\theta}_{mle}, 100)$
    $\beta_1 \sim N(0, 10)$
    $\beta_2 \sim N(0, 10)$
  - A Wishart prior with 3 degree of freedom is assigned to the precision matrix $T$. 
MCMC Implementation

- GEV parameters were drawn using a Metropolis-Hastings step with starting values are the corresponding cell-wise maximum likelihood estimates.

- Candidates for the three parameters are drawn in a block for each location $j$ using a uniform random walk.

- Random effects $U$ were updated following Rue & Held canonical parameterization.

- $\beta$ and $T$ were updated following its distribution.

- Codes for this work were provided by Daniel Cooley.

- The model was run for 15000 iterations for each data set, discarded the first 5000 to allow for burn-in, took only every $10^{th}$ iteration results to reduce dependence.
Results
draws from posterior
distribution of $\mu, \sigma, \xi, \beta, U$ and $T$
Location parameters,

Posterior mean of location parameter for each simulated data driven by each of GCMs.
Scale parameters,

Posterior mean of scale parameter for each simulated data driven by each of GCMs
Shape parameters,

Posterior mean of shape parameter for each simulated data driven by each of GCMs
100 years Return levels map

Return level basically is a quantile that associated with return period $1/p$. 
Hyperparameters Posterior

Histogram of beta posterior; the first column correspond to intercept term, second column and third correspond to latitude and latitude respectively. Each row is beta posterior for local, scale and shape parameters respectively.

Histogram of $U$ posterior for random selected cells.

Histogram of $T$ posterior.
Conclusion
Conclusion

- We modeled maximum temperature data; simulation output from RCM driven by 6 different GCMs, using three stages Gaussian hierarchical model. The spatial patterns are not directly modeled from the data but through parameters of the assume data distribution.

- Bayesian inference was carried out by Metropolis Hastings, and following Rue & Held algorithm.

- The 100years-return level map from 6 GCMs show slightly different spatial pattern in the posterior parameter distribution especially for shape parameters.

- Model improvement can be carried on by adding covariates or considering second order spatial. Modeling observed data and comparing the model to simulation data model might enable us to find correction factor for better simulation model.

- Possible extension of this work would be to combine temperature and precipitation data using multivariate spatial hierarchical model.
Bibliography


Bibliography


