Bayesian spatial hierarchical modeling for temperature extremes

Indriati Bisono
Dr. Andrew Robinson
Dr. Aloke Phatak

Mathematics and Statistics Department
The University of Melbourne
Maths, Informatics and Statistics, CSIRO WA

October 6-7th, 2011

Outline

Problem Discussion

What we observe?

Downscaled annual maximum temperatures, z_{it}, for grid cells $i = 1, \ldots, 2856$ covering all Tasmania, during year $t = 1, \ldots, 49$ (1961 to 2009).

Problem Discussion

What we observe?

Downscaled annual maximum temperatures, z_{it}, for grid cells $i = 1, \ldots, 2856$ covering all Tasmania, during year $t = 1, \ldots, 49$ (1961 to 2009).

What we do?

Developing a three stage Bayesian spatial hierarchical model, following Schliep E. and Dan Cooley's (2010) model. The spatial structure is depicted by means of the random effect which is modelled using conditional autoregressive (CAR).

Problem Discussion

Aim of study

Constructing a return level map of weather extremes .

Problem Discussion

Aim of study

Constructing a return level map of weather extremes .

p-Year Return Level

- is an expected value to be exceeded once every $1/p$ years with probability p. In other words, a p-quantile that associated with return period 1/p.

Problem Discussion

Aim of study

Constructing a return level map of weather extremes .

p-Year Return Level

- is an expected value to be exceeded once every $1/p$ years with probability p. In other words, a p-quantile that associated with return period 1/p.
- can be find simply by inverted the following equation;

Problem Discussion

Aim of study

Constructing a return level map of weather extremes .

p-Year Return Level

- is an expected value to be exceeded once every $1/p$ years with probability p. In other words, a p-quantile that associated with return period 1/p.
- can be find simply by inverted the following equation;

$$P(Z \geq z) = \exp\left\{-\left(1 + \xi\frac{x - \mu}{\sigma}\right)^{-1/\xi}\right\} = p$$

Problem Discussion

Aim of study

Constructing a return level map of weather extremes .

p-Year Return Level

- is an expected value to be exceeded once every $1/p$ years with probability p. In other words, a p-quantile that associated with return period 1/p.
- can be find simply by inverted the following equation;

$$P(Z \geq z) = \exp\left\{-\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-1/\xi}\right\} = p$$

- For $p = 0.01$, return level has 1% chance of being exceeded during return period of $1/p = 100$ years.

Stage 1 of 3

Data level

Assume z_{it} follows GEV distribution

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

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

Stage 2 of 3

Process level

- Let $\theta = (\mu, log(\sigma), \xi)$
- Assume θ follows Normal distribution

$$\theta \sim N\left(X\beta + U, \frac{1}{\tau^2}\right)$$

where

- X is a 2856×3 matrix of covariates (latitude and longitude)
- β is a 3×3 matrix of regression coefficients
- U is a 2856×3 matrix of random effect
- τ^2 is a fixed precision matrix

Stage 2 of 3

Process level

- Let $\theta = (\mu, log(\sigma), \xi)$
- Assume θ follows Normal distribution

$$\theta \sim N\left(X\beta + U, \frac{1}{\tau^2}\right)$$

where

- X is a 2856×3 matrix of covariates (latitude and longitude)
- β is a 3×3 matrix of regression coefficients
- U is a 2856×3 matrix of random effect
- τ^2 is a fixed precision matrix

U is modelled spatially using conditional autoregressive model

Conditional autoregressive model for random effects U

Assume random effect U is a Gaussian Markov random field (GMRF) that satisfies conditional independence assumptions.

$$U_i | u_j, j \neq i \sim N\left(\sum_j b_{ij} u_j, t_i^2\right), i = 1, \ldots, n$$

where

- b_{ij} is a spatial dependence parameter, i.e. $b_{ij} = \frac{w_{ij}}{w_{i+}}$ and $w_{i+} = \sum_j w_{ij}$.

Conditional autoregressive model for random effects U

Assume random effect U is a Gaussian Markov random field (GMRF) that satisfies conditional independence assumptions.

$$U_i | u_j, j \neq i \sim N\left(\sum_j b_{ij} u_j, t_i^2\right), i = 1, \ldots, n$$

where

- b_{ij} is a spatial dependence parameter, i.e. $b_{ij} = \frac{w_{ij}}{w_{i+}}$ and $w_{i+} = \sum_j w_{ij}$.
- $w_{ij} = 1$ if node i and j share the same boundary, and $w_{ij} = 0$ otherwise.

Conditional autoregressive model for random effects U

Assume random effect U is a Gaussian Markov random field (GMRF) that satisfies conditional independence assumptions.

$$U_i | u_j, j \neq i \sim N\left(\sum_j b_{ij} u_j, t_i^2\right), i = 1, \ldots, n$$

where

- b_{ij} is a spatial dependence parameter, i.e. $b_{ij} = \frac{w_{ij}}{w_{i+}}$ and $w_{i+} = \sum_j w_{ij}$.
- $w_{ij} = 1$ if node i and j share the same boundary, and $w_{ij} = 0$ otherwise.
- t_i^2 is conditional precision; set $t_i^2 = \frac{T^2}{w_{i+}}$.

Conditional autoregressive model for random effects U

- The setup suggests a joint multivariate normal distribution for $U = (U_1, \ldots, U_n)$ with mean 0 and precision $Q = (D_w - W)$ where D_w is diagonal with $(D_w)_{ii} = w_{i+}$

Conditional autoregressive model for random effects U

- The setup suggests a joint multivariate normal distribution for $U = (U_1, \ldots, U_n)$ with mean 0 and precision $Q = (D_w - W)$ where D_w is diagonal with $(D_w)_{ii} = w_{i+}$
- But, $(D_w - W)\mathbf{1} = \mathbf{0}$, i.e. Σ_U^{-1} is singular so that Σ_U does not exist

Conditional autoregressive model for random effects U

- The setup suggests a joint multivariate normal distribution for $U = (U_1, \ldots, U_n)$ with mean 0 and precision $Q = (D_w - W)$ where D_w is diagonal with $(D_w)_{ii} = w_{i+}$
- But, $(D_w - W)\mathbf{1} = \mathbf{0}$, i.e. Σ_U^{-1} is singular so that Σ_U does not exist
- Replacing each element in W with $\frac{w_{ij}}{w_{i+}}$ would restrict each row sum to one, $(D_w - W)$ would not be singular and Σ_U does exist. This is often referred to as an intrinsically autoregressive (IAR) model [Banerjee et al., 2004]

Stage 3 of 3

Parameter level

- The hyperparameters are β and T
- Choose conjugate priors for the two hyperparameters, i.e.
  - a normal distribution for β priors, $\beta \sim N(\beta_0, \kappa^{-1})$
  - a Wishart prior with 3 d.f. for precision matrix T

MCMC implementation

- Using hybrid Monte Carlo; combining Metropolis and Gibbs sampler algorithms
- Metropolis algorithm used for generating GEV parameters posterior distribution

$$\pi(\theta|z) = \pi(z|\theta)\pi(\theta)$$

  where θ represents μ, σ and ξ.
- Gibbs sampler employed for generating posterior distributions for U, β and T

$$\begin{aligned}
U|\theta, \beta &\sim N_C\left(\tau^2(\theta - X\beta), T + \tau^2\right) \\
\beta|\theta, U &\sim N_C\left(\tau^2(\theta - U) + \kappa\beta_0, \tau^2 + \kappa\right) \\
T|\beta, U &\sim W^{-1}(\Psi, 3 + k), where \Psi = U^T WU + T_0, k = 2856
\end{aligned}$$

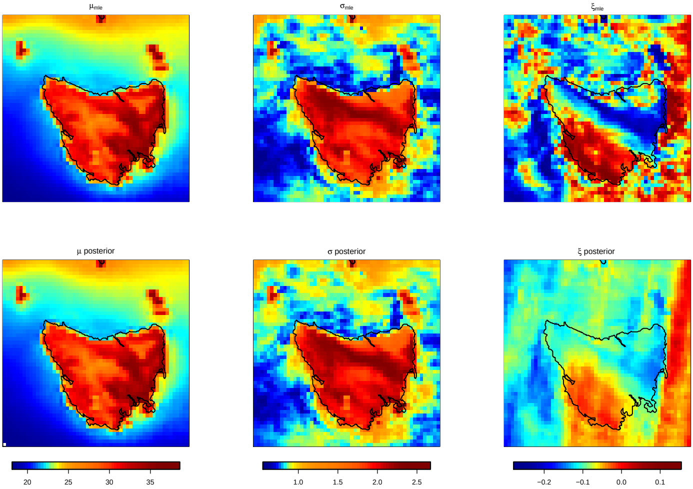
Metropolis Algorithm

1. Start with MLE estimates of corresponding parameters $\theta^{(0)}$. Set $k = 1$
2. Generate a proposal θ^* from proposal distribution as follow
   - $\mu^* = \mu^{(k-1)} + \mathrm{scale} \cdot \mathrm{rt}(1, 2)$
   - $\sigma^* = \sigma^{(k-1)} + \mathrm{scale} \cdot (\mathrm{runif} - 0.5)$
   - $\xi^* = \xi^{(k-1)} + \mathrm{scale} \cdot \mathrm{rt}(1, 5)$
3. Set $\theta^{(k)} = \theta^*$ with probability

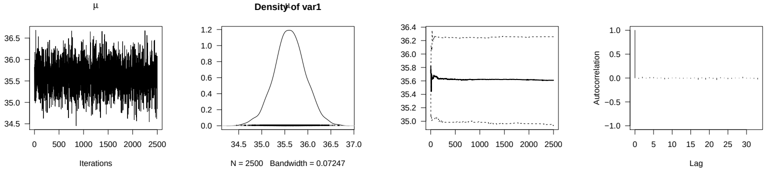
   $$\alpha = \min\left\{1, \frac{\pi(\theta^*)}{\pi(\theta^{(k-1)})}\right\}$$

   Otherwise set $\theta^{(k)} = \theta^{(k-1)}$
4. Set $k = k + 1$ and return to 2.

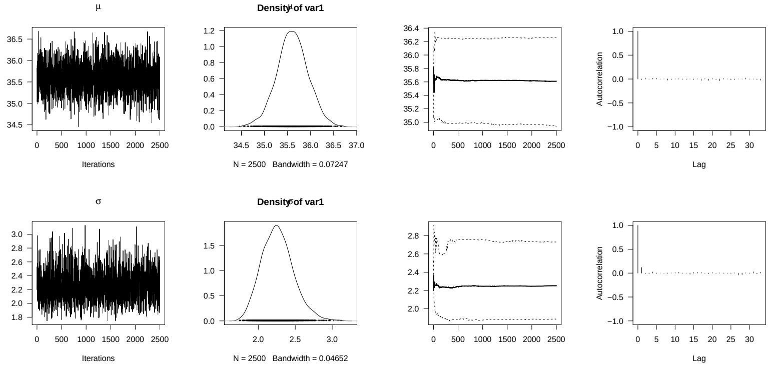
Prior and posterior GEV parameters



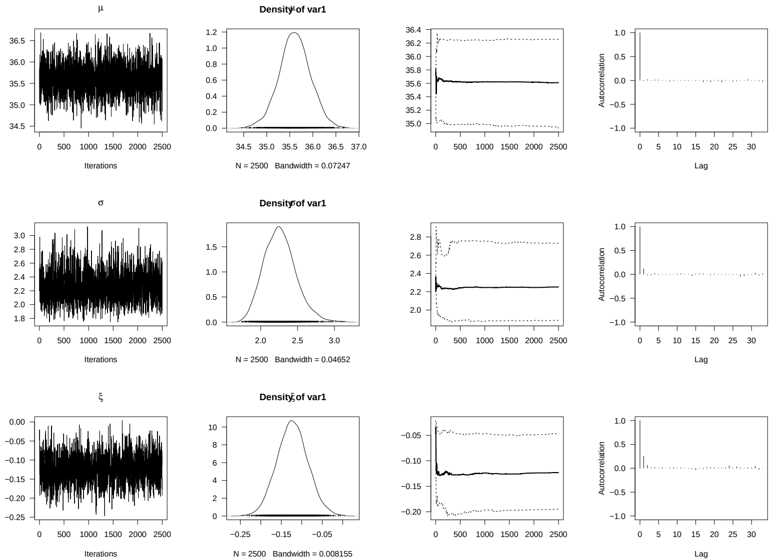
Diagnostic plots for GEV parameters at one grid cell



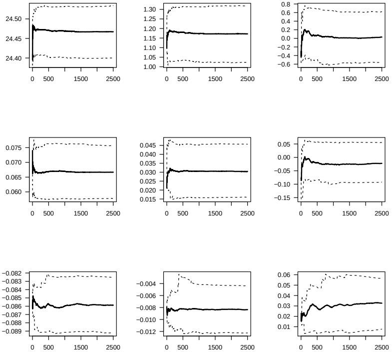
Diagnostic plots for GEV parameters at one grid cell



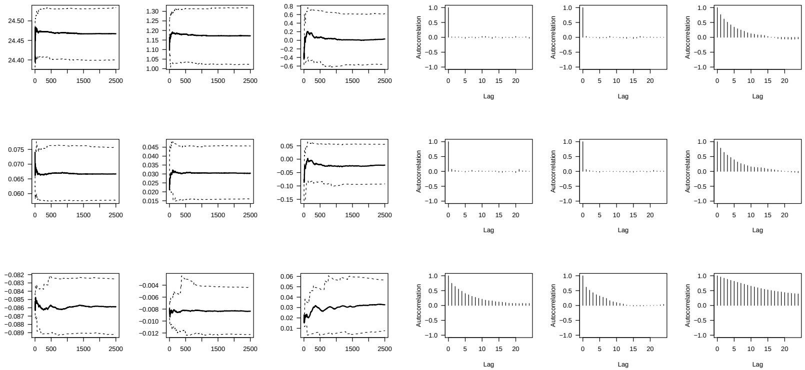
Diagnostic plots for GEV parameters at one grid cell



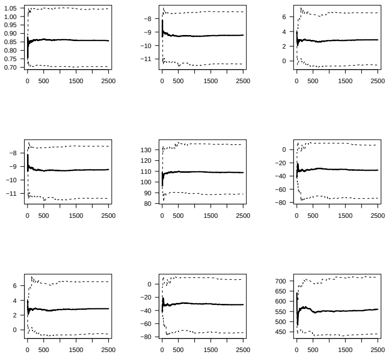
Cumulative and AR plots for β



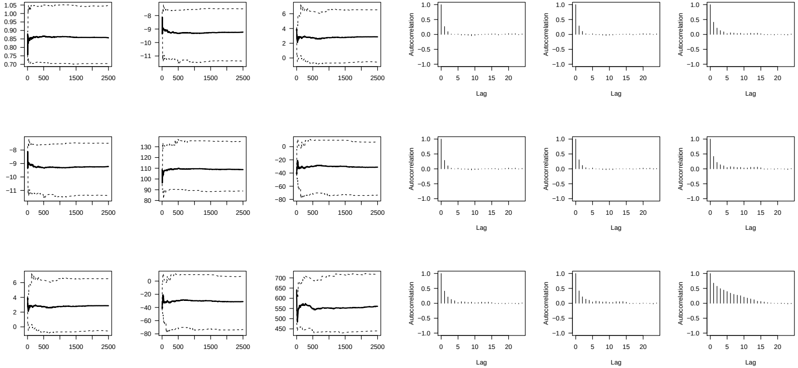
Cumulative and AR plots for β



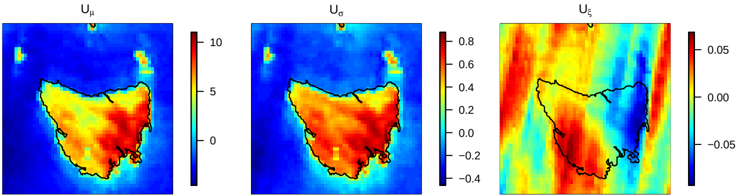
Cumulative and AR plots for T



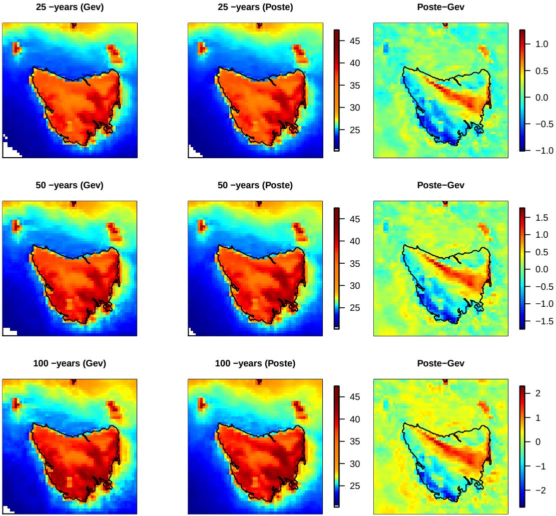
Cumulative and AR plots for T



Maps of U-posterior



Return level maps



Conclusion

- We used a hierarchical model with three stages.

Conclusion

- We used a hierarchical model with three stages.
- The spatial patterns are not directly modeled from the data but through parameters of the assume data distribution.

Conclusion

- We used a hierarchical model with three stages.
- The spatial patterns are not directly modeled from the data but through parameters of the assume data distribution.
- Conjugate priors were chosen for hyperparameters to ease the computation.

Conclusion

- We used a hierarchical model with three stages.
- The spatial patterns are not directly modeled from the data but through parameters of the assume data distribution.
- Conjugate priors were chosen for hyperparameters to ease the computation.
- Bayesian inference was carried out by hybrid MCMC, and following canonical parameterization of Rue and Held [Rue and Held, 2005].

Conclusion

- We used a hierarchical model with three stages.
- The spatial patterns are not directly modeled from the data but through parameters of the assume data distribution.
- Conjugate priors were chosen for hyperparameters to ease the computation.
- Bayesian inference was carried out by hybrid MCMC, and following canonical parameterization of Rue and Held [Rue and Held, 2005].
- Determination of τ^2; a fixed precision matrix for GEV parameters, considerably affects the rate of convergence.

Conclusion

- We used a hierarchical model with three stages.
- The spatial patterns are not directly modeled from the data but through parameters of the assume data distribution.
- Conjugate priors were chosen for hyperparameters to ease the computation.
- Bayesian inference was carried out by hybrid MCMC, and following canonical parameterization of Rue and Held [Rue and Held, 2005].
- Determination of τ^2; a fixed precision matrix for GEV parameters, considerably affects the rate of convergence.
- Other variables that greatly improved the produced chains are the choice of proposal distribution for Metropolis algorithm and the jump of for proposed parameters; too small or too big jump results in slower convergence and higher autocorrelation.

Future Works

- Estimate the best proposal distribution for Metropolis algorithm using an MCMC algorithm; MCMC within MCMC.

Future Works

- Estimate the best proposal distribution for Metropolis algorithm using an MCMC algorithm; MCMC within MCMC.
- Model the weather extremes based on observed data, and compare it to that of downscaled data.

Future Works

- Estimate the best proposal distribution for Metropolis algorithm using an MCMC algorithm; MCMC within MCMC.
- Model the weather extremes based on observed data, and compare it to that of downscaled data.
- Develop a joint distribution of temperature and wind extremes as multivariate spatial hierarchical model, possibly using copulas.

References

- Schliep, Erin M., Dan Cooley, Stephan R. Sain and Jennifer A. Hoeting (2010). A comparison study of extreme precipitation from six different regional climate models via spatial hierarchical modeling. *Extremes*, 13:219-239.
- Havard Rue and Leonhard Held (2005). *Gaussian Markov Random Fields: Theory and Applications*, Chapman & Hall/CRC.
- Sudipto Banerjee, Bradley P. Carlin, and Alan E. Gelfand, 2004. *Hierarchical Modeling and Analysis for Spatial Data*, Chapman & Hall/CRC.

Acknowledgements

Special thanks for Dan Cooley who kindly provides his R-script for this research.

Thanks to my supervisor and Department of Mathematics and Statistics, The University of Melbourne for providing financial support for attending this conference.

I would also like to convey thanks to the Indonesia Higher Education Ministry for providing a scholarship.

Thank you!