Bayesian Modelling of Extreme Event with the Link to the Climatological Model Data

Darmesah Gabda and Jonathan Tawn

Universiti Malaysia Sabah,Malaysia, Lancaster University, UK

Scope of study

AIM: To improve inferences for a model for observational maxima data.

Idea: Exploit the extra information of the trend of extreme from the numerical climate model, i.e global climate model (GCM) and a linked regional climate model (RCM).

Data structure

- Annual maximum temperature 25km gridded observations, X_t from year (t), 1959 to 2008 at site $s = 1, \ldots, h_r$ in region $r = 1, \ldots, 5$.
- Simulated annual temperature maxima, $Y_t^{(j)}$ at year, (t) 1949 to 2099 from different members of an ensemble, $j = 1, \ldots, 11$ using a single RCM with the same fine 25 km grid simulations.
- Simulated annual temperature maxima, $Z_t^{(j)}$ at year, (t) 1949 to 2099 from different members of an ensemble, $j = 1, \ldots, 17$ using a GCM with a grid of 300km.
- Observed global mean temperature, $g_{X,t}^{(o)}$.
- The global annual mean temperature with respect to year 2000, $g_{M,t}^{(j)}$ from j^{th} GCM ensemble members.
- *Both RCM and GCM models were driven under the A1B Special Report on Emission Scenarios (SRES) from year 2000 to 2099.*

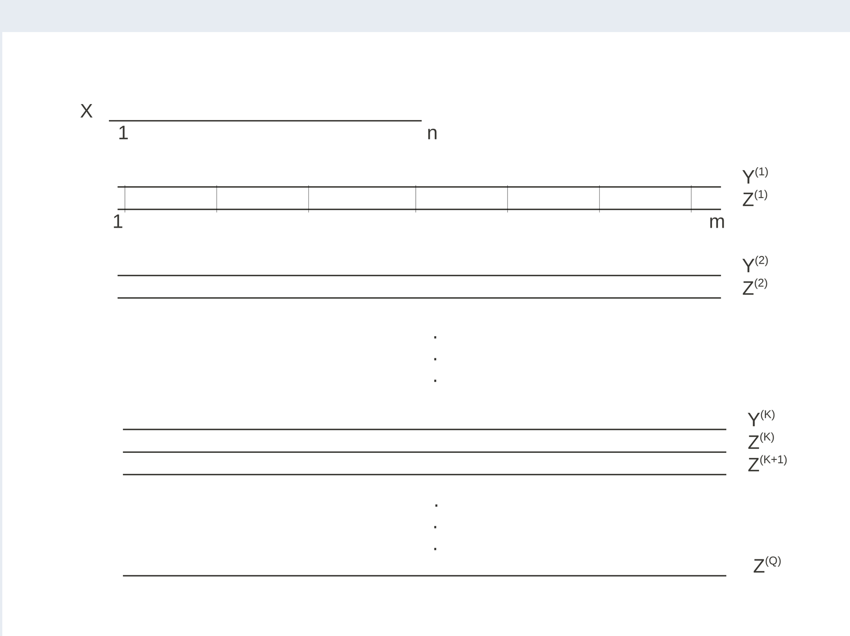


Figure 1: The data structure.

Model formulation

We fit a generalized extreme value distribution (GEV) at each site as follows:

$$X_t \sim GEV\left(\boldsymbol{\theta}_{X,t}\right),$$
$$Y_t^{(j)} \sim GEV\left(\boldsymbol{\theta}_{Y,t}^{(j)}\right)$$
$$Z_t^{(j)} \sim GEV\left(\boldsymbol{\theta}_{Z,t}^{(j)}\right)$$

where:

$$\boldsymbol{\theta}_{X,t} = \left(\mu_{X,t} = \alpha_X + \beta_{X,t} g_{X,t}^{(o)}, \sigma_{X,t}, \xi_X\right)$$
$$\boldsymbol{\theta}_{Y,t}^{(j)} = \left(\mu_{Y,t}^{(j)} = \alpha_Y^{(j)} + \beta_{Y,t}^{(j)} g_{M,t}^{(j)}, \sigma_{Y,t}^{(j)}, \xi_Y^{(j)}\right)$$
$$\boldsymbol{\theta}_{Z,t}^{(j)} = \left(\mu_{Z,t}^{(j)} = \alpha_Z^{(j)} + \beta_{Z,t}^{(j)} g_{M,t}^{(j)}, \sigma_{Z,t}^{(j)}, \xi_Z^{(j)}\right)$$

By pooling all these data $\left(X, Y^{(j)}, Z^{(j)}\right)$ in one model fit leads to a complex structure with 92 number of parameters.

Strategy

Explore possible link between the parameters, $\boldsymbol{\theta}_{X,t}, \boldsymbol{\theta}_{Y,t}^{(j)}, \boldsymbol{\theta}_{Z,t}^{(j)}$.

- to reduce the variance of the parameters from the GEV fit of the observed data
- to reduce model complexity

The proposed model: link between observations, RCM and GCM data

The link between parameters of $\boldsymbol{\theta}_{X,t}, \boldsymbol{\theta}_{Y,t}^{(j)}, \boldsymbol{\theta}_{Z,t}^{(j)}$:

$$\beta_{X,(r,s)} = \frac{1}{11}\sum_{j=1}^{11} \beta_{Y,(r,s)}^{(j)}, \quad \beta_{Y,(r,s)}^{(j)} = \kappa_{\beta_0} + \kappa_{\beta_1}\beta_{Z,(r,s)}^{(j)}$$
$$\sigma_{X,(r,s)} = \frac{1}{11}\sum_{j=1}^{11} \sigma_{Y,(r,s)}^{(j)}, \quad \sigma_{Y,(r,s)}^{(j)} = \kappa_{\sigma_0} + \kappa_{\sigma_1}\sigma_{Z,(r,s)}^{(j)}$$
$$\xi_{Y,(r,s)}^{(j)} = \kappa_{\xi_0} + \kappa_{\xi_1}\xi_{Z,(r,s)}^{(j)}$$

We found both the trend parameter and the scale parameter of X are common to Y parameters so we can link these parameters and simplify our proposed model.

Bayesian inference

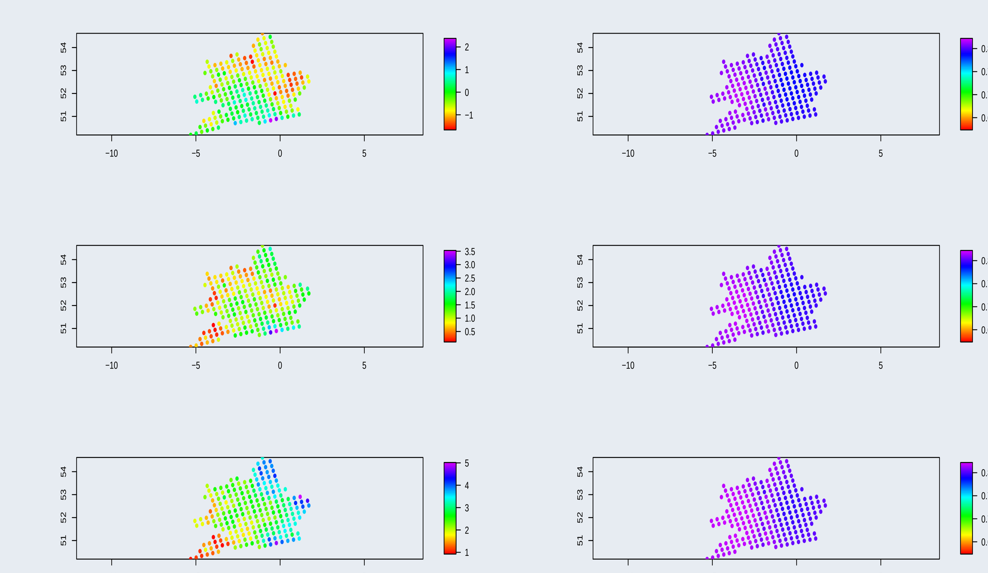
The remaining parameters to estimate are as follows:

$$\boldsymbol{\theta}_X^{**} = \left\{\alpha_{X,(r,s)}, \xi_{X,(r,s)}, \quad s = 1, \ldots, h_r\right\}$$
$$\boldsymbol{\theta}_{YZ} = \left\{\kappa_{\beta_0}, \kappa_{\beta_1}, \kappa_{\sigma_0}, \kappa_{\sigma_1}, \kappa_{\xi_0}, \kappa_{\xi_1}\right\}.$$

MCMC with Metropolis Hastings algorithm implemented as an alternative to the maximum likelihood and to produce uncertainty in the climatological data.

Results

The summary of the trend parameter estimates of the observed maxima temperature and the corresponding 95% confidence intervals/credible intervals using the maximum likelihood and the Bayesian inferences.



(a) (b)

Figure 2: The comparisons of the trend parameter estimates by (a) the maximum likelihood using X data only and (b) the Bayesian inference using $(X, Y^{(j)}, Z^{(j)})$. From top to bottom (a); the lower bound of the 95% confidence intervals, the trend parameter estimates by MLE and the upper bound of the 95% confidence intervals. From top to bottom of (b): the lower bound of the 95% credible intervals, the mean sampling distributions of trend parameter estimates by MCMC and the upper bound of the 95% credible intervals.

Region 3	Lower bound 95% CI	Estimate	Upper bound 95% CI
MLE	-0.506	1.311	3.129
MCMC	0.797	0.802	0.806
Region 5	Lower bound 95% CI	Estimate	Upper bound 95% CI
MLE	-0.237	0.868	1.973
MCMC	0.811	0.816	0.819

Table 1: The average of trend parameter estimates of the observed temperature maxima over the region and the corresponding of 95% confidence intervals/credible intervals using two different methods; maximum likelihood and the Bayesian inferences.

Conclusions

- The inference of the trend parameter of the observed temperature maxima are more efficient by pooling all the information of $(X, Y^{(j)}, Z^{(j)})$ instead of using observed temperature maxima data only.
- We have obtained an efficient trend estimate of the observed extreme temperature by also considering the uncertainty of the climatological model changes.

References

Brown, S.J., Murphy, J., Sexton, D. and Harris, G. (2013). Climate projections of future extreme events accounting for modelling uncertainties and historical simulation biases. *CliMathNet Conference 2013.*

Carlin, B.P. and Louis, T.A. (2000). *Bayes and empirical Bayes methods for data analysis, Second edition.* Chapman and Hall/CRC.

Acknowledgements

UK Met Office for provision of the data used in this research, specifically Simon Brown for advice on the problem setup. Ministry of Higher Education Malaysia and the Universiti Malaysia Sabah for the Gabda,D. scholarship.