Modelling annual maximum 24 hour precipitation in Iceland using block sampling and SPDE models

Óli Páll Geirsson[1]
Birgir Hrafnkelsson[1]
Daniel Simpson[2]

School of Engineering and Natural Sciences,
University of Iceland.[1]
Department of Mathematical Sciences,
Norwegian University of Science and Technology[2]

June 13, 2013

Outline

1 Introduction

2 Model structure

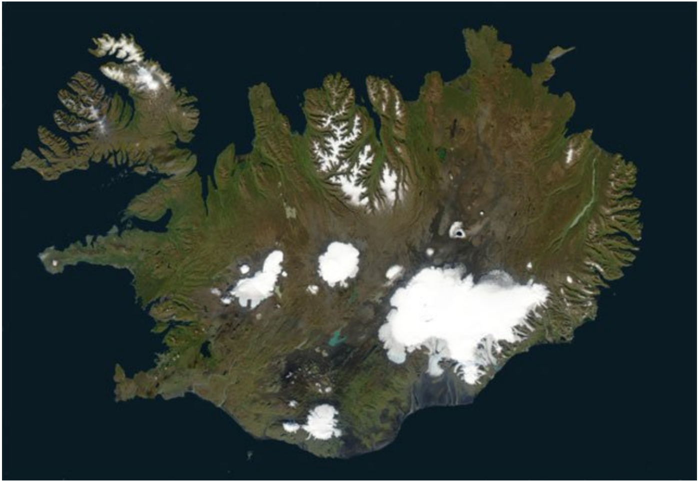
3 Results

4 Conclusions

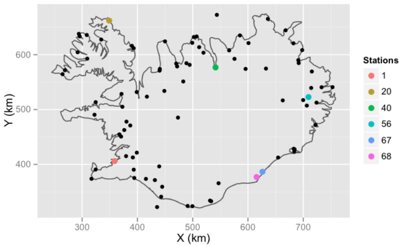
Introduction

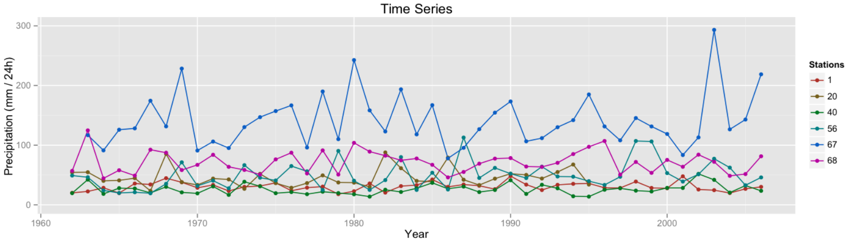
- The main goal is to obtain distributional properties of extreme precipitation on a high resolution grid over Iceland
- We rely on
  - Observations of 24 hour annual extreme precipitation from 86 sites over the years 1961 to 2006
  - Covariates based on outputs from a linear model of orographic precipitation on a 1 km^2 regular grid [Crochet et all, 2007]
- A method for spatial quantile predictions of extreme precipitation combining these two sources is presented

I HAVE IN MY POSSESSION A MAP (OF ICELAND)



Observation stations and time series





Data level [Stage 1]

- The data are modelled with a Bayesian hierarchical model assuming a generalized extreme value distribution for the observations.
- That is, let y_{it} denote the annual maximum 24 hour precipitation at station i at year t, with a cumulative distribution function of the form

$$F(y_{it}) = \exp\left\{ -\left(1 + \xi\left(\frac{y_{it} - \eta_i}{\sigma_i}\right)\right)^{-1/\xi}\right\},$$

where η_i, σ_i and ξ are location, scale and shape parameters, respectively.

Latent level [Stage 2]

- Working within the LGM framework, decompose $\boldsymbol{\eta}$ into

$$\boldsymbol{\eta} = \boldsymbol{X}_\eta \beta_\eta + \boldsymbol{A}\boldsymbol{u}_\eta + \boldsymbol{v}_\eta, \tag{1}$$

  where
  - $\boldsymbol{X}_\eta$ is a design matrix based on outputs from the local climate model
  - β_η is the corresponding weight
  - $\boldsymbol{u}_\eta$ denotes a Matérn type spatial field constructed with the SPDE approach of [Lindgren et al., 2011]
  - $\boldsymbol{A}$ is a projection matrix
  - $\boldsymbol{v}_\eta$ is an unstructured random effect.
- We assign the following priors

$$\beta_\eta \sim \mathcal{N}(0, \kappa_\beta^{-1})$$
$$\boldsymbol{u}_\eta \sim \mathcal{N}(\mathbf{0}, (\boldsymbol{Q}_u(\tau_\eta, \theta_\eta))^{-1})$$
$$\boldsymbol{v}_\eta \sim \mathcal{N}(\mathbf{0}, \kappa_v^{-1}\boldsymbol{I}).$$

- Similar structure is implemented for $\nu_i := \log \sigma_i$

The hunt for covariates - the $X_\eta \beta_\eta$ part in (1)

Reasonable covariates for extreme precipitation should contain information such as

- The topology of the domain
- Underlying physical processes
- Preferably on a high resolution grid

It has been suggested (Benestad et al, 2012)[7] that observed mean values are useful predictors for extreme precipitation.
The main idea: How about using simulated mean values from meteorological models instead?

Construct covariates

- A covariate based on the orographic model can be computed at each grid point which has some of the desired properties.
- Namely, by calculating the mean of the simulated precipitation in each grid point, which yields a 521 km x 361 km regular grid of covariates.
- Observations tend not to be on regular grid points. However, a simple smoother can be used to map the covariates from grid points to observation points.

SPDE SPATIAL MODEL - THE Au_η PART IN (1)

- Triangulation of the domain, Iceland, is constructed
- An approximate solutions of the SPDE

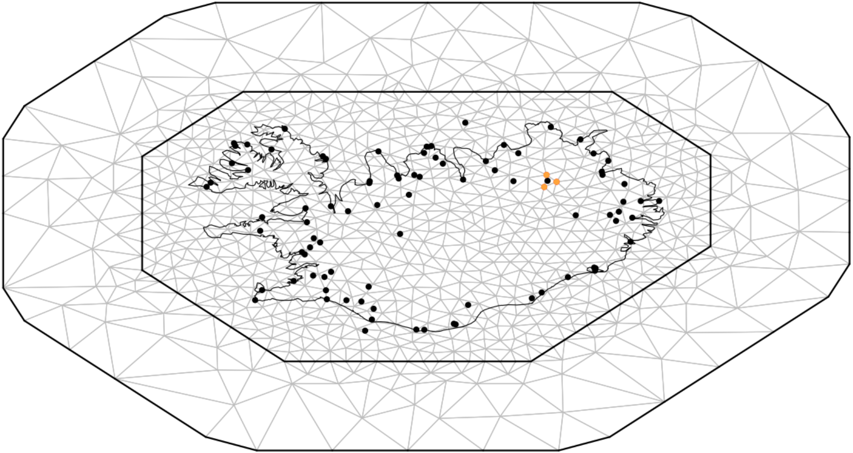
$$(\kappa^2 - \Delta)^{\alpha/2} (\tau x(\boldsymbol{s})) = \mathcal{W}(\boldsymbol{s}),$$
$$\boldsymbol{s} \in \mathbb{R}^d, \quad \alpha = \nu + d/2, \quad \kappa > 0, \quad \nu > 0$$

  can be found on the triangular grid, which is a Gaussian Markov random field (GMRF) representation of a Matérn field
- The approximate solution $x(\boldsymbol{s})$ is represented by

$$x(\boldsymbol{s}) = \sum \psi(\boldsymbol{s}) w_k$$

  where ψ_k are piecewise linear in each triangle, and w_k are Gaussian weights.

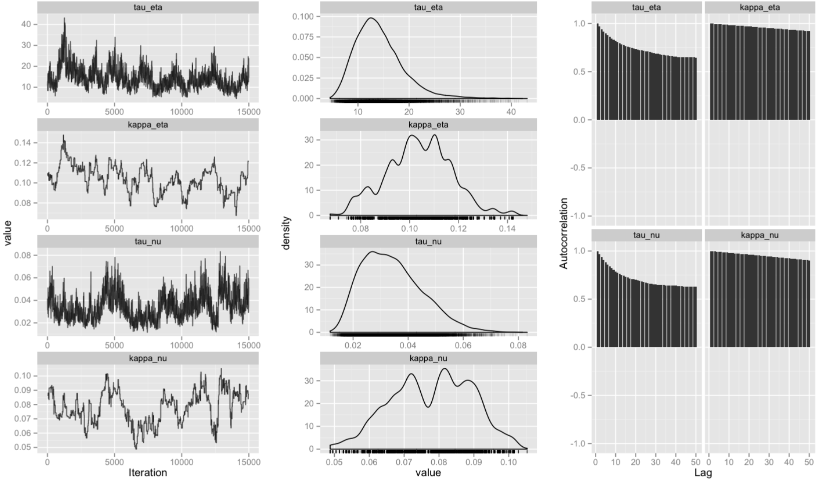
Triangulation



RESULTS

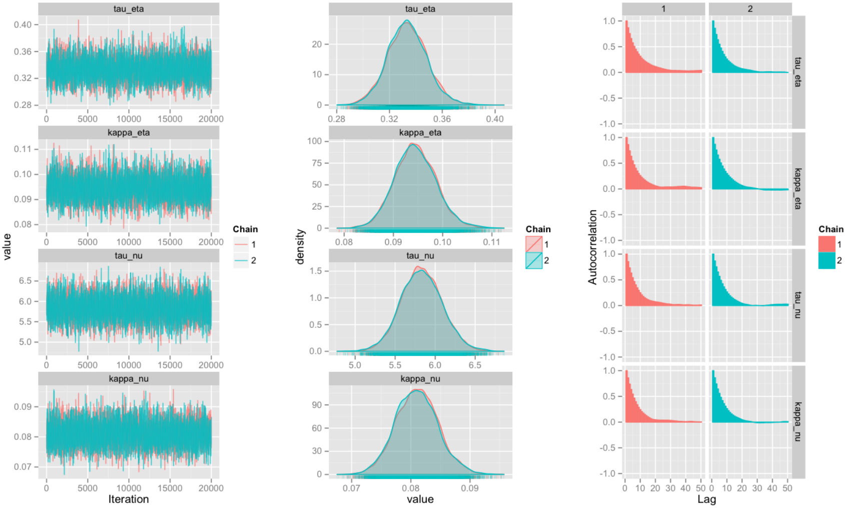
- The following results are based on two MCMC chains, 25.000 iterations each.
- The MCMC chains are inferred with a novel two block updating scheme
- Runtime is around 4 hours on a laptop (i7 core, 4 Gb ram)

- Trace, density and autocorrelation for $\boldsymbol{\theta}$



Two block updating scheme

- Trace, density and autocorrelation for $\boldsymbol{\theta}$



Spatial effects - Posterior estimates

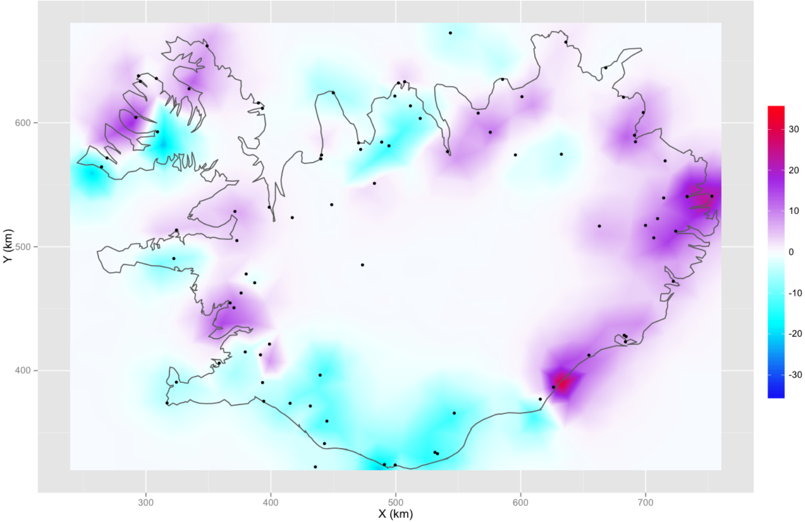
- Let $\hat{\boldsymbol{u}}$ denote a posterior estimate for the spatial effects on the *triangulated mesh*.
- Since every point in the domain of interest belongs to some triangle, a posterior estimate for the spatial effect can be obtained at every point with a convex linear combination of $\hat{\boldsymbol{u}}$.
- Let $\hat{\boldsymbol{u}}_{hd}$ denote a posterior estimate for the spatial effects on the *high resolution grid*. It is obtained by projecting $\hat{\boldsymbol{u}}$ with the linear transformation

$$\boldsymbol{A}\hat{\boldsymbol{u}} = \hat{\boldsymbol{u}}_{hd}$$

  where $\boldsymbol{A}$ is a matrix which describes the piecewise linear functions within each triangle.

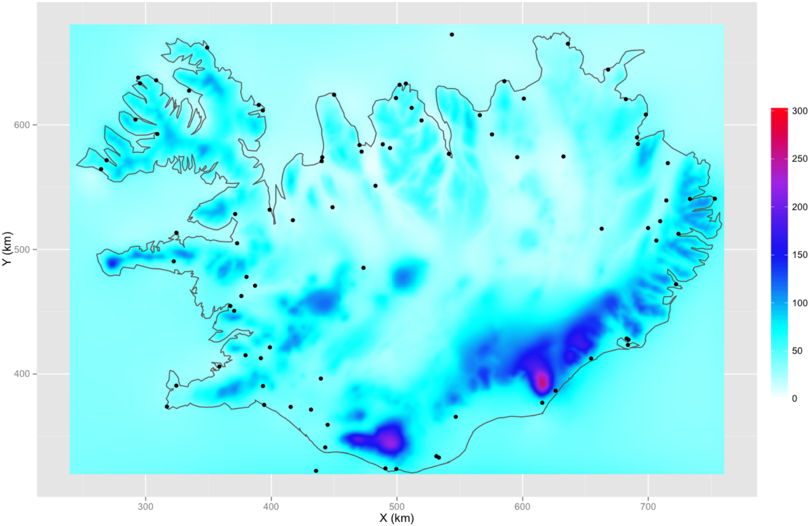
Location parameter

- Posterior mean of the spatial field $\boldsymbol{u}_\eta$ on the high resolution net.



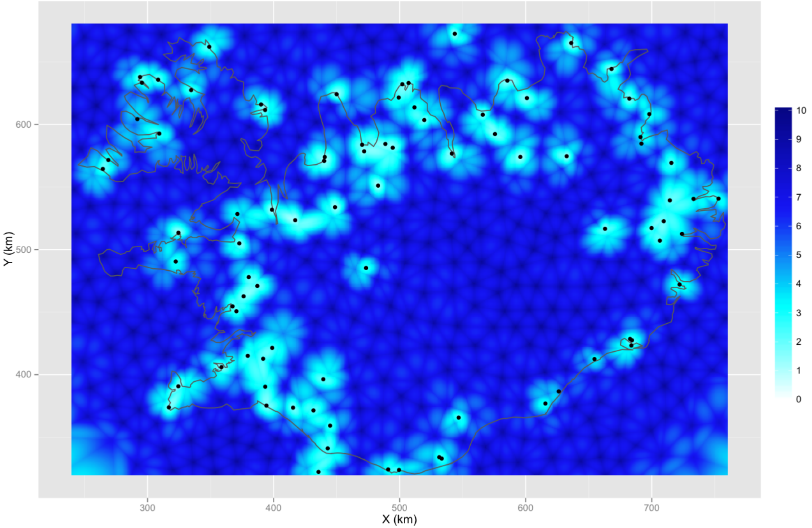
Location parameter

- Posterior mean of the spatial field $\boldsymbol{X}_\eta \beta_\eta + \boldsymbol{u}_\eta$

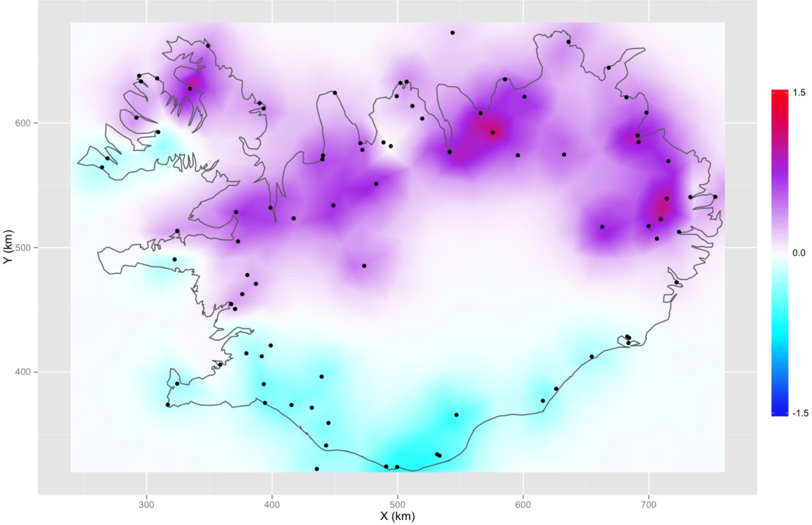


Location parameter

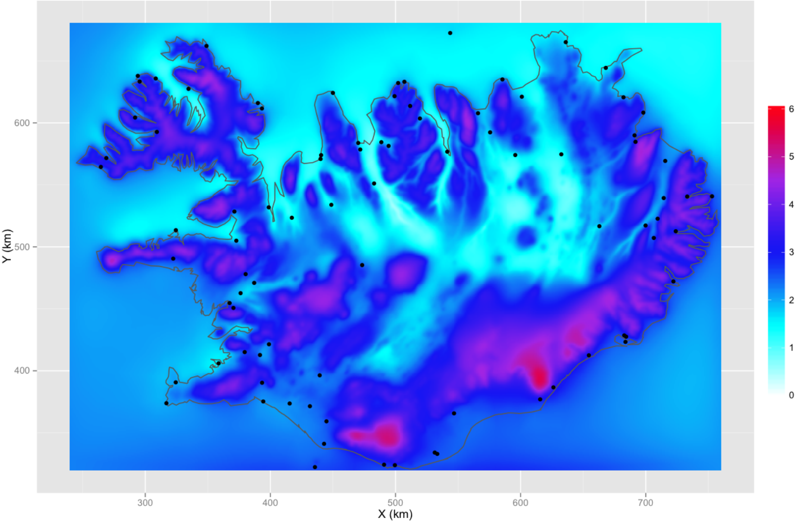
- Posterior standard deviation of the spatial field $\boldsymbol{u}_\eta$



- Posterior mean of the spatial field $\boldsymbol{u}_\nu$

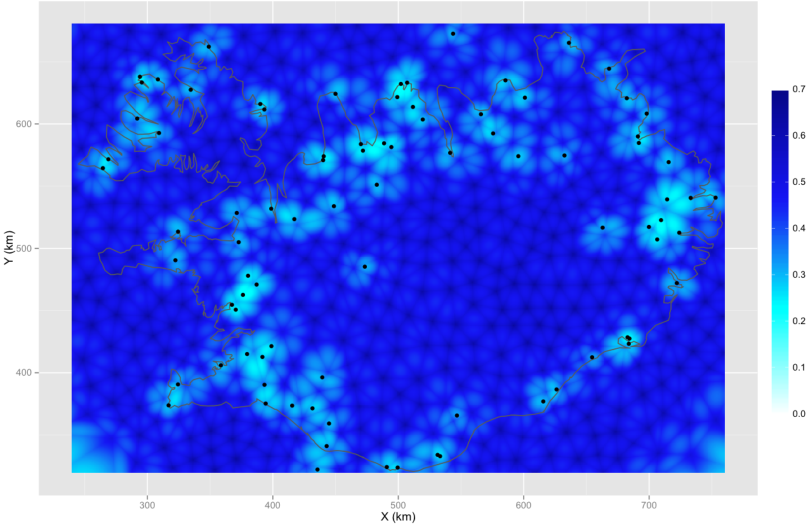


- Posterior mean of the spatial field $\boldsymbol{X}_\nu \beta_\nu + \boldsymbol{u}_\nu$



log SCALE PARAMETER

- Posterior standard deviation of the spatial field $\boldsymbol{u}_\nu$

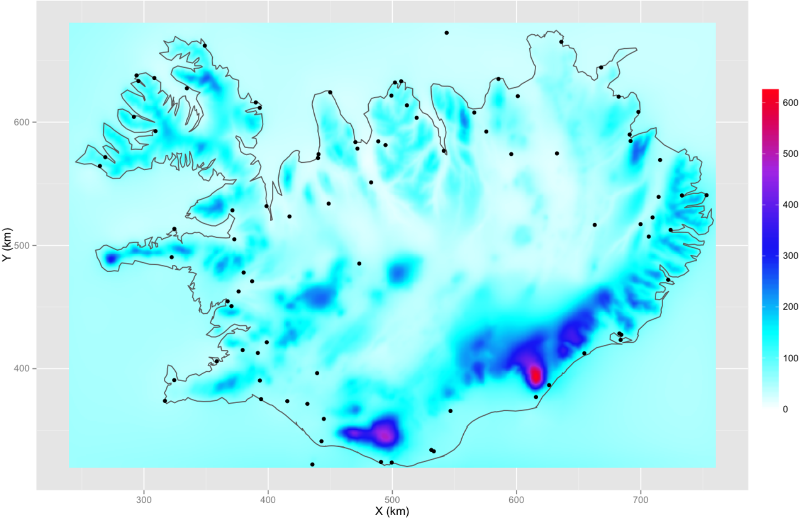


Quantile estimates

- Once the spatial estimates $\boldsymbol{\eta}_{hd}$ and $\boldsymbol{\tau}_{hd}$ at every grid point on the high resolution grid are obtained along with an estimate for ξ , an estimate can be given for distribution of maximum annual 24 hour precipitation at every grid point.
- In particaular, a simple estimate of the p-th quantile can be obtained by

$$y_{p,un,j} = \hat{\eta}_{hd,j} + \frac{\exp(\hat{\tau}_{hd,j})}{\hat{\xi}} \left(- \log(p)^{-\hat{\xi}} - 1 \right)$$

Estimation of the 95-th quantile



CONCLUSIONS

- The SPDE spatial models yield a computationally efficient prediction scheme
- The two block sampling scheme is beneficial in terms of
  - Efficient sampling of the non-likelihood latent parameters regardless of the choice of likelihood
  - Efficient sampling of the hyper parameters which dramatically reduces autocorrelation
- Different likelihood functions call for appropriate sampling schemes for the likelihood parameters.

References

1. Crochet, P., T. Jóhannesson, T. Jónsson, O. Sigurðsson, H. Björnsson, F. Pálsson and I. Barstad (2007): *Estimating the spatial distribution of precipitation in Iceland using a linear model of orographic precipitation*. J. of Hydrometeorol., Vol. 8 (6), 1285-1306.
2. Hrafnkelsson, B., Morris J. S. and Baladandayuthapani V. (2011) Spatial modeling of annual minimum and maximum temperatures in Iceland. *Meteorology and Atmospheric Physics*, to appear.
3. Banerjee, Carlin, Gelfand. (2004). *Hierarchical Modeling and Analysis for Spatial Data*. Chapman & Hall/CRC.
4. Sang, H. Y. and Gelfand A. E. (2009) Hierarchical modeling for extreme values observed over space and time. *Environmental and ecological statistics*, 16, 407-426.
5. Rue, H., and Held, L. (2005) Gaussian Markov Random Fields: Theory and Applications, New York: Chapman Hall/CRC
6. Lindgren, F., Lindström. J. and Håvard R. (2011) An explicit link between Gaussian fields and Gaussian Markov random fields; The SPDE approach. *Journal of the Royal Statistical Society*, 71 ,319-392.

Special thanks to you

- Thank you for your attention