Bayesian modeling of monthly precipitation via integration of measurements and meteorological model output

Atli Norðmann Sigurðarson
Birgir Hrafnkelsson

Faculty of Physical Sciences
University of Iceland

May 31, 2013

Objectives

- Propose a novel Bayesian hierarchical model for observed monthly precipitation.
- Incorporate covariates based on an output from a meteorological model.
- Give spatial predictions of the parameters of the proposed data distribution.
- Compare these parameters to summaries of the meteorological output.

Rain gauge data

- Corrected observed monthly precipitation
- Comes from 40 sites across Iceland.
- Time period 1958-2006. Some data missing.
- Corrected for:
  - wind
  - wetting
  - evaporation

- A priori simplifications:
- Analyze each month separately.
  - Removes the necessity of temporal correlation.
  - Alleviates the problem of dimensionality.

Rain gauge data - Location of sites

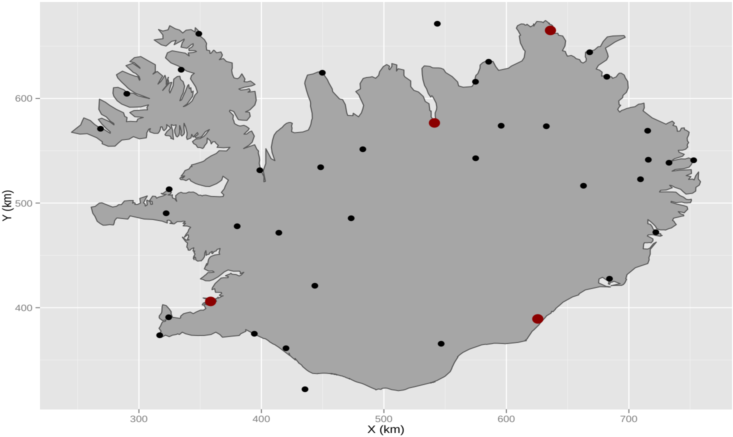


Figure: Location of rain gauge sites across Iceland.

Rain gauge data - Timeseries

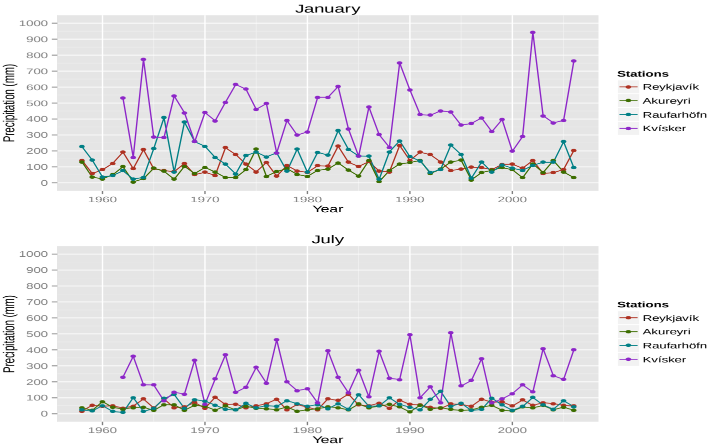


Figure: Time series for selected sites in January and July.

Rain gauge data - Boxplot

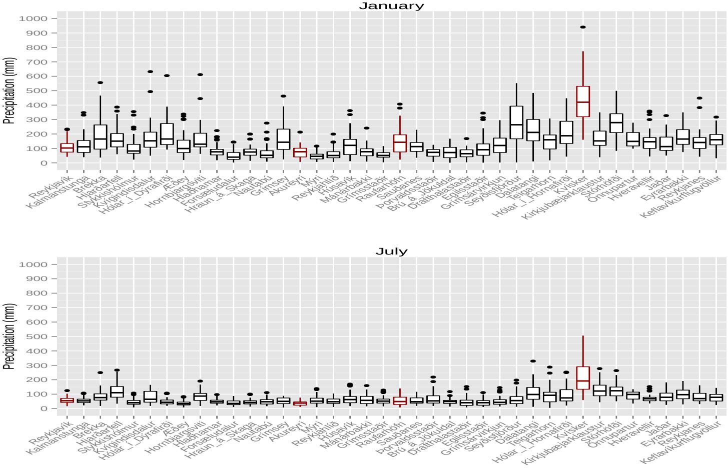


Figure: Boxplot of the data in January and July.

Rain gauge data - Correlation

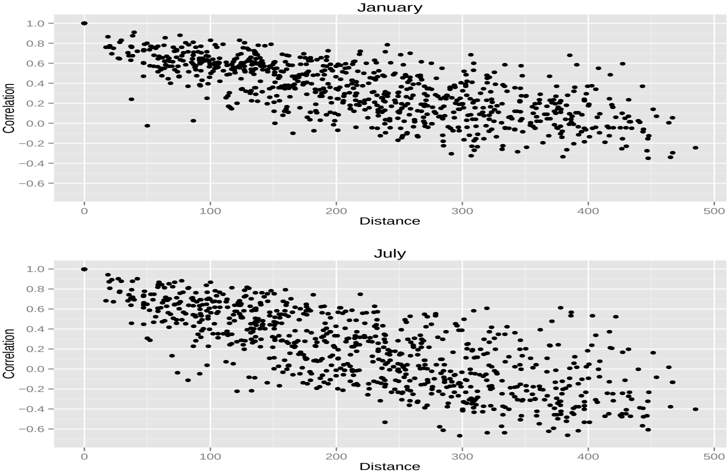


Figure: Correlation of the data in January and July.

- Corrected monthly precipitation, y_{obs}, not adequately described by normal or log-normal distribution.
- Decided to use Box-Cox transformation, with parameter $\lambda = 0.4$, to achieve normality.

$$y = \frac{y_{\text{obs}}^{\lambda} - 1}{\lambda}$$

- Gives a pretty good fit for most sites in each month.
- Much better fit than with no transformation or natural logarithm.

Distributional properties

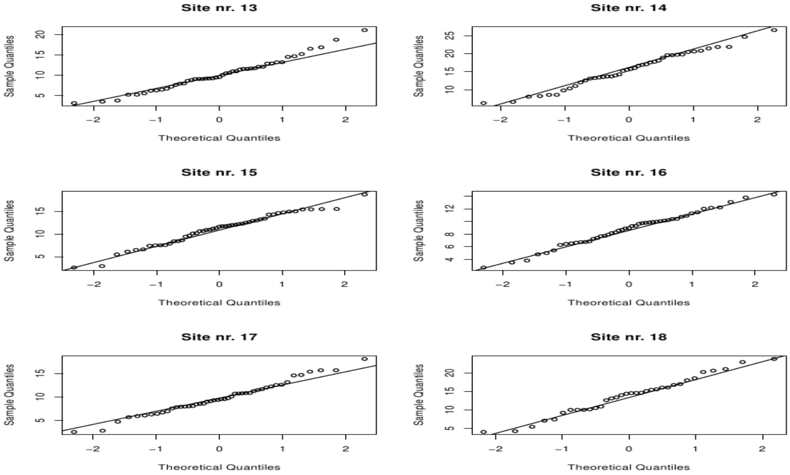


Figure: Q-Q plot of the Box-Cox transformed data, with parameter $\lambda = 0.4$, from selected sites in January.

Meteorological model output

- Based on a linear model of orographic precipitation.
- Driven using coarse-resolution 40-yr reanalysis data (ERA-40) over the period 1958-2002.
- Daily precipitation on a 1 km^2 grid over Iceland.
- Takes into account:
  - Topology.
  - Airflow dynamics.
  - Condensed water advection.
  - Downslope evaporation.

Meteorological model output

- Daily precipitation in each grid point added together within each month to get monthly precipitation.
- Monthly precipitation in each grid point transformed with the Box-Cox transformation with $\lambda = 0.4$.
- Mean and $\log(\sqrt{\text{variance}})$ of monthly precipitation calculated in each grid point and for each month over the whole time period.
- Mean and $\log(\sqrt{\text{variance}})$, on the grid, projected onto the rain gauge sites to be used as covariates.

Meteorological model output

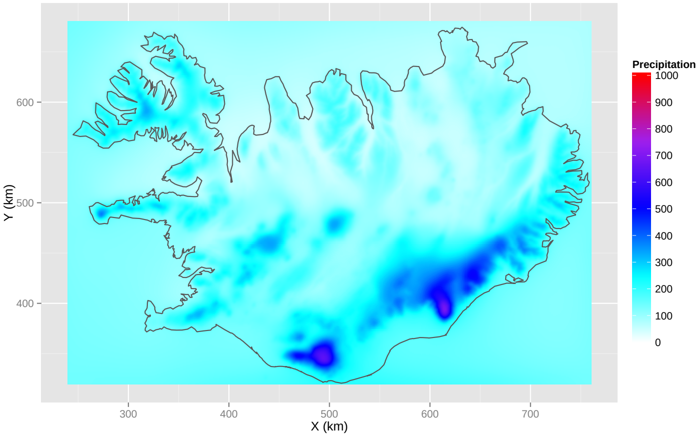


Figure: Meteorological covariate for the spatial mean on the 1-km grid over Iceland for January.

Meteorological model output

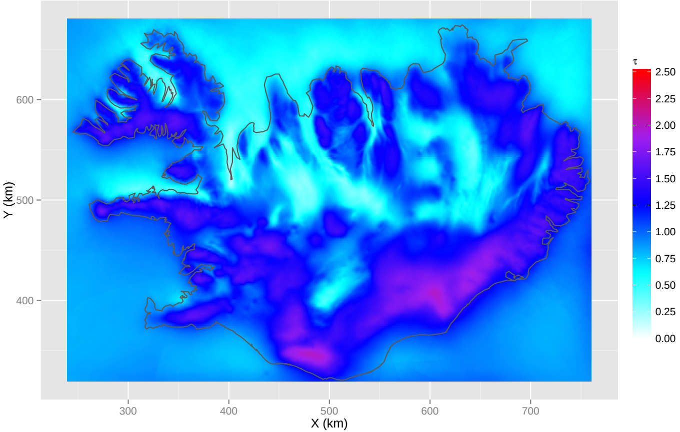


Figure: Meteorological covariate for the spatial variance on the 1-km grid over Iceland for July.

Statistical modeling

- Data: Box-Cox transformed corrected monthly precipitation, $\vec{\mathbf{y}}$.
- The model should include:
  - Correlation at the data level.
  - Measurement error.
  - Spatially varying mean and variance parameters.
  - Time main effect.
- Such a model has not been presented before.

The statistical model

- Data level

$$\vec{\mathbf{y}} \mid \vec{\boldsymbol{\alpha}}, \vec{\boldsymbol{\gamma}}, \vec{\boldsymbol{\tau}}, \psi, \kappa \quad \sim \quad N(Z_1\vec{\boldsymbol{\alpha}}+Z_2\vec{\boldsymbol{\gamma}},\, Q(\vec{\boldsymbol{\tau}})(R_\nu(\psi)+\kappa E)Q(\vec{\boldsymbol{\tau}})).$$

- Latent level

$$\begin{aligned}
\vec{\boldsymbol{\alpha}} \mid \beta, \sigma_\alpha^2, \phi &\sim N(X\beta,\, \sigma_\alpha^2 R_\nu(\phi)) \\
\vec{\boldsymbol{\tau}} \mid \eta, \sigma_\tau^2, \rho &\sim N(V\eta,\, \sigma_\tau^2 R_\nu(\rho)) \\
\vec{\boldsymbol{\gamma}} \mid \sigma_\gamma^2 &\sim N(0,\, \sigma_\gamma^2 I).
\end{aligned}$$

- $\vec{\boldsymbol{\alpha}}$ is a spatially varying mean.
- $\vec{\boldsymbol{\tau}}$ is $\log\left(\sqrt{\vec{\boldsymbol{\sigma}}^2}\right)$, where $\vec{\boldsymbol{\sigma}}^2$ is a spatially varying variance.
- $\vec{\boldsymbol{\gamma}}$ is the time main effect.

Results - β

Table: Posterior estimates of β

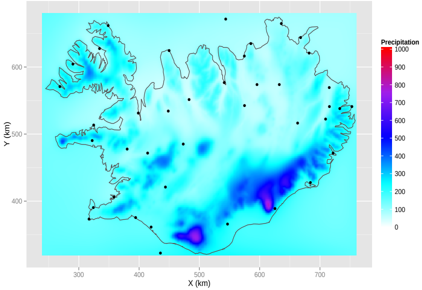
Month	median	2.5%	97.5%
January	1.06	1.02	1.11*
February	1.08	1.02	1.13*
March	1.05	1.00	1.09
April	1.04	1.00	1.08
May	0.99	0.94	1.05
June	1.01	0.96	1.06
July	1.01	0.97	1.04
August	0.98	0.93	1.03
September	1.04	1.00	1.08
October	1.05	1.00	1.09
November	1.08	1.04	1.12*
December	1.08	1.03	1.12*

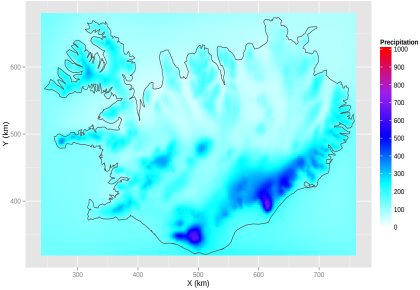
Results - η

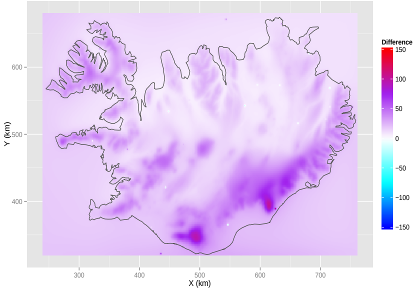
Table: Posterior estimates of η

Month	median	2.5%	97.5%
January	1.23	1.07	1.37*
February	1.08	0.87	1.28
March	1.02	0.84	1.18
April	0.90	0.70	1.08
May	0.82	0.65	0.98*
June	0.78	0.61	0.95*
July	0.70	0.57	0.83*
August	0.82	0.67	0.97*
September	1.00	0.84	1.17
October	1.09	0.92	1.27
November	0.96	0.79	1.11
December	0.95	0.71	1.17

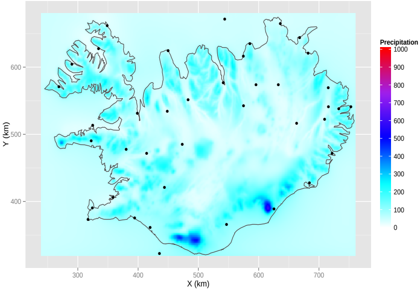
Results - Median precipitation in January

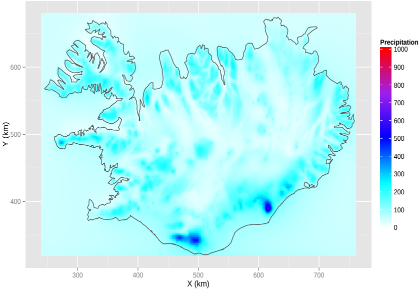


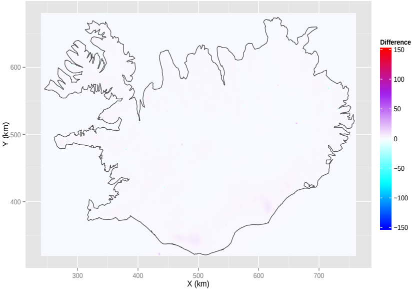




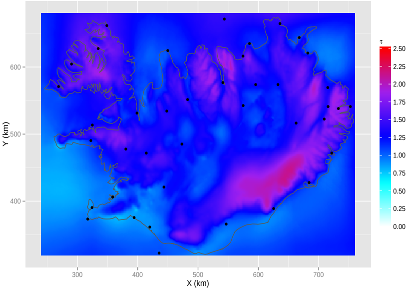
Results - Median precipitation in July

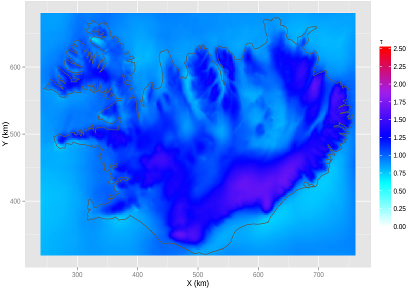


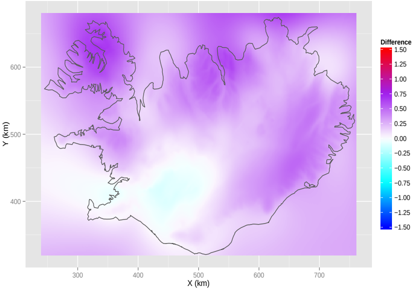




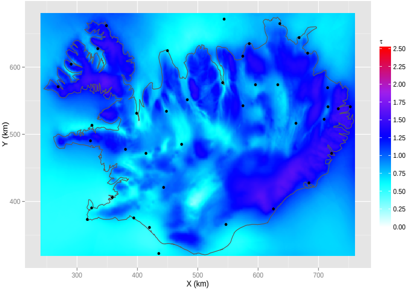
Results - The τ-field in January

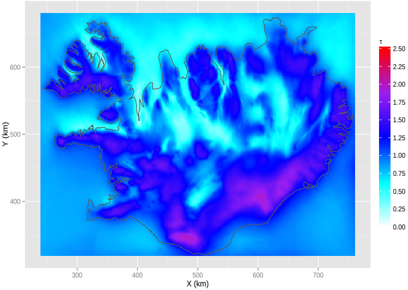


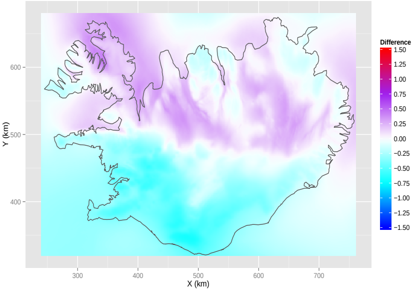




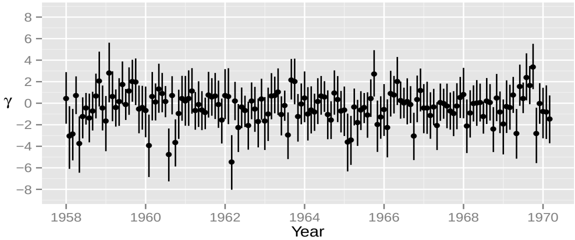
Results - The τ-field in July







Results - $\vec{\gamma}$



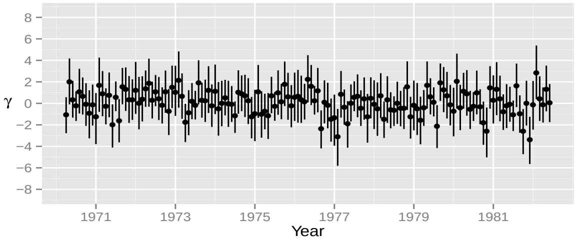


Figure: The parameter $\vec{\gamma}$ over 1958 to 1982.

RESULTS - $\vec{\gamma}$

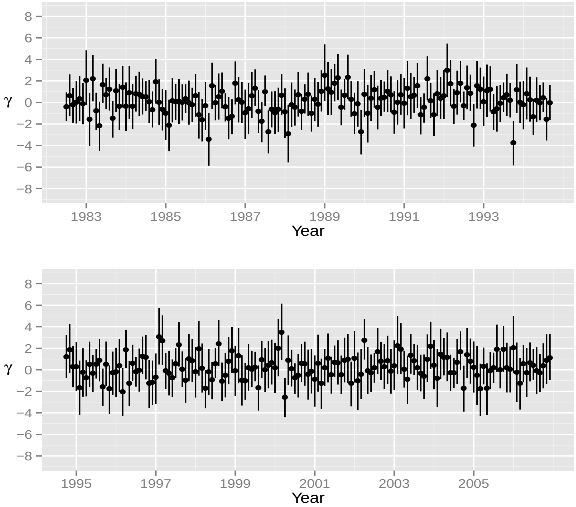


Figure: The parameter $\vec{\gamma}$ over 1983 to 2006.

Modeling conclusions

- The model is detailed and accurate in describing the data at hand.
- Including a spatial correlation at the data level is worth the effort.
- Spatially varying variance is just as natural as a spatially varying mean.
- Using a meteorological covariate is quite beneficial.

Comparison with meteorological covariates

- The meteorological covariate constructed for the mean monthly precipitation describes the mean behavior well.
  - It describes it very well for the summer months, May to August.
  - It underestimates the precipitation in the winter months, November to February.
  - It also gives a slight underestimation for the months in spring and autumn.
- The meteorological covariate constructed for the variance of monthly precipitation does not describe the variation as well.
  - A discrepancy is generally displayed for large areas.
  - For most of the months is underestimates the the variability in the North-East part of Iceland and overestimates it in the South-West corner.
- The meteorological model provides good covariates for the observed data.

Thank you for your attention.