A new Bayesian hierarchical geostatistical model based on two spatial fields with case studies with short records of annual runoff in Norway

Ingelin Steinsland[1], Thea Roksvåg[1] & Kolbjørn Engeland[2]

1) NTNU (Norwegian University of Science and Technology),

2) NVE (Norwegian water resources and energy directorate)

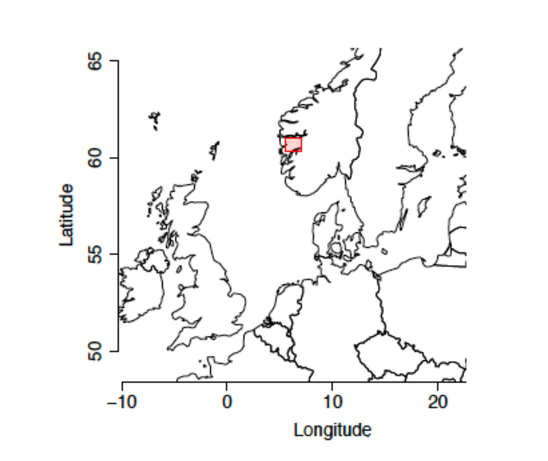
Norwegian University of Science and Technology

NVE

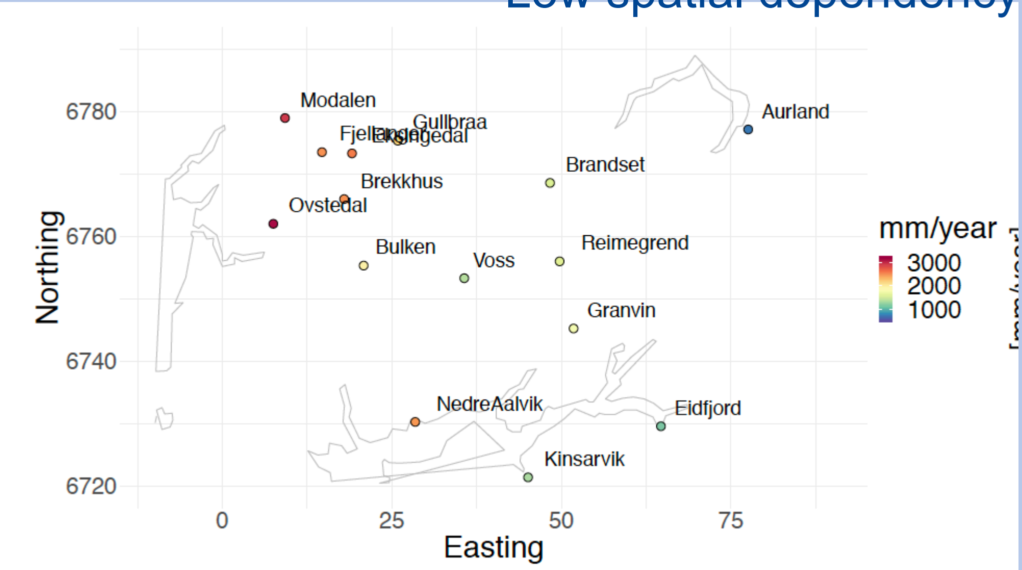
NTNU

Motivating example

Annual precipitation in Western Norway

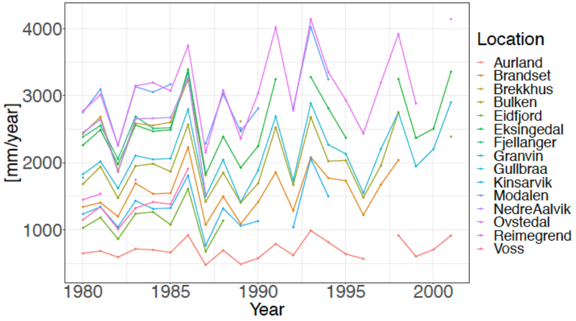


Low spatial dependency

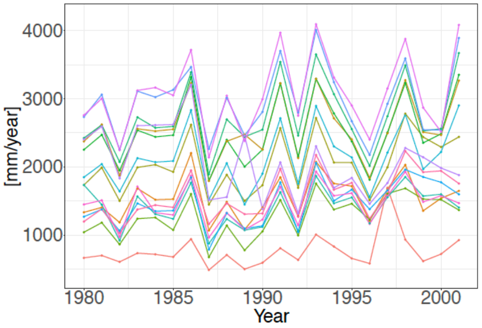


Low temporal dependency, but we find the same pattern each year

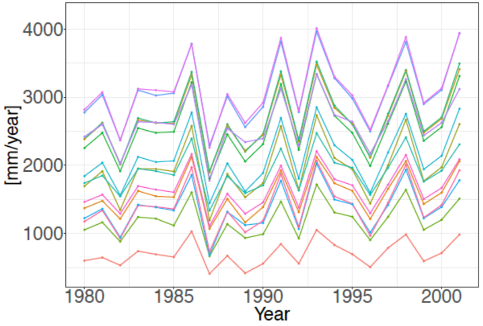
Data:



Traditional geostatistics:



Desired solution?



We have constructed a geostatistical method that is able to detect and capture such data patterns.

Model run-off location *u* year *j*

$$q_j(\boldsymbol{u}) = \beta_c + c(\boldsymbol{u}) + \beta_j + x_j(\boldsymbol{u}); \qquad \boldsymbol{u} \in \mathcal{R}^2; \qquad j = 1, ..r.$$

Model r years simultaneously.

Climate (long-term averages):

β_c : Intercept.

$c(\boldsymbol{u})$: Spatial effect. GRF(ρ_c,σ_c).

Annual discrepancy from climate:

β_j : Intercept. year j

$x_j(\boldsymbol{u})$: Spatial effect. GRF(ρ_x,σ_x).

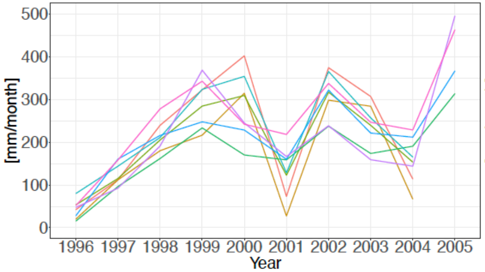
GRF: Gaussian random field. $\mathcal{N}(0, \boldsymbol{\Sigma}(\rho, \sigma^2))$.

$\boldsymbol{\Sigma}$: Covariance matrix.

ρ : Spatial range.

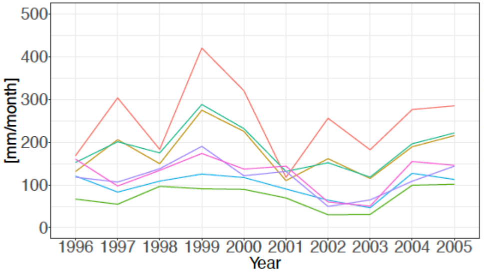
σ^2 : Spatial marginal variance.

Simulated examples from model



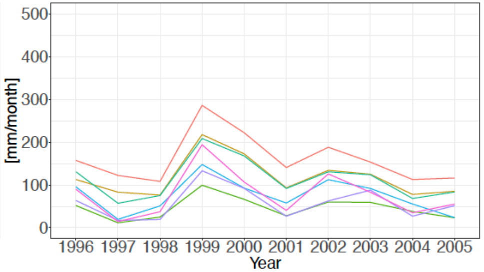
Climate < Annual

$\sigma_c < \sigma_x$



Climate ≈ Annual

$\sigma_x \approx \sigma_c$



Climate » Annual

$\sigma_c \gg \sigma_x.$

$$q_j(\boldsymbol{u}) = \beta_c + c(\boldsymbol{u}) + \beta_j + x_j(\boldsymbol{u}); \qquad \boldsymbol{u} \in \mathcal{R}^2; \qquad j = 1, ..r.$$

Model r years simultaneously.

Climate (long-term averages):
β_c : Intercept.
$c(\boldsymbol{u})$: Spatial effect. GRF(ρ_c,σ_c).

Annual discrepancy from climate:
β_j : Intercept.
$x_j(\boldsymbol{u})$: Spatial effect. GRF(ρ_x,σ_x).

For catchment $\mathcal{A}$

Areal model:
$Q_j(\mathcal{A}) = \frac{1}{|\mathcal{A}|} \int_{\boldsymbol{u} \in \mathcal{A}} q_j(\boldsymbol{u}) d\boldsymbol{u}.$

Mass-conservation.

Centroid model:
$Q_j(\mathcal{A}) = q_j(\boldsymbol{u}_{\mathcal{A}})$
$\boldsymbol{u}_{\mathcal{A}}$ is the catchment centroid.

No mass conservation, but quicker.

With the model framework we can:

- Model overlapping catchments consistently
- Combining areal (runoff) and point (precipitation) observations.
- Utilize short records
- Use informative priors in a Bayesian setting
- Fast inference using INLA (Integrated nested Laplace Approximations)

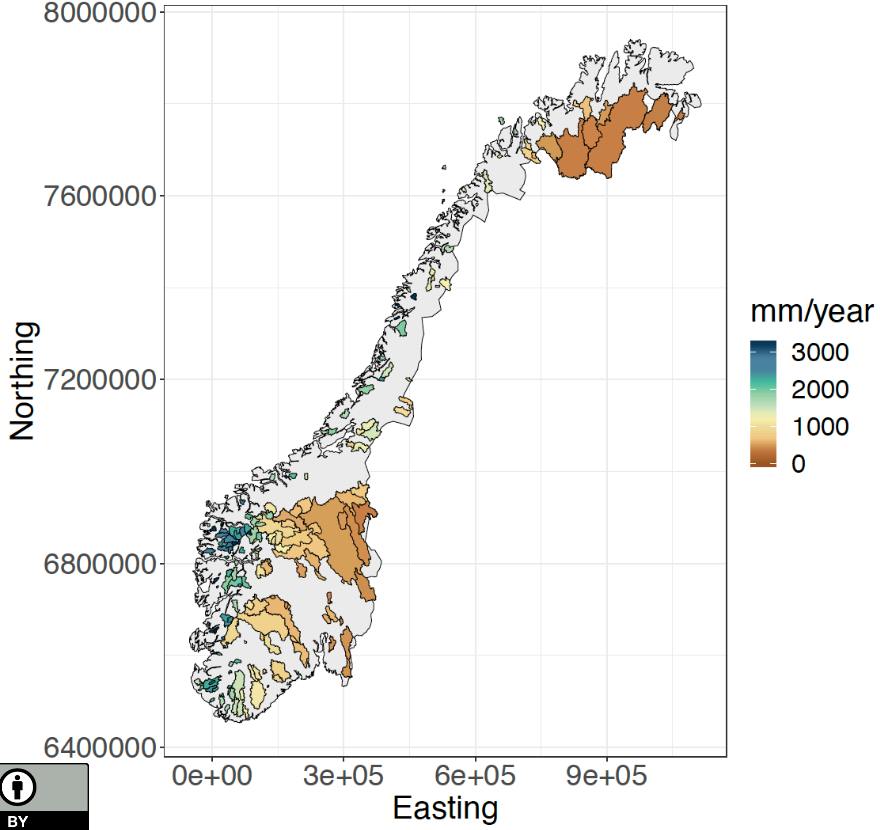
Papers available :

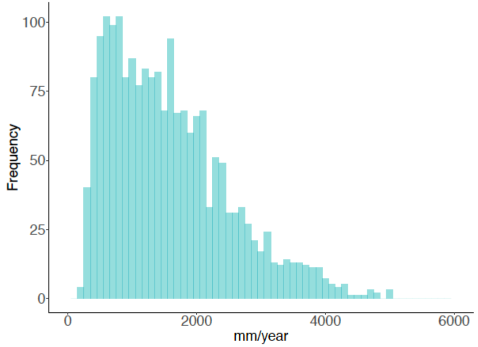
- Thea Roksvåg, Ingelin Steinsland, Kolbjørn Engeland , *A geostatistical two field model that combines point observations and nested areal observations, and quantifies long-term spatial variability -- A case study of annual runoff predictions in the Voss area* Under revision, available at arXiv:1904.02519

- Roksvåg, T., Steinsland, I., and Engeland, K.: *A geostatistical framework for estimating flow indices by exploiting short records and long-term spatial averages – Application to annual and monthly runoff*, Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-415, in review, 2019.

Example using short records

Test the areal and centroid model on Norwegian annual runoff data.





200 catchments.
Data from 1996-2005.

Divide the catchments into 20 groups/folds.

For each cross-validation fold we:

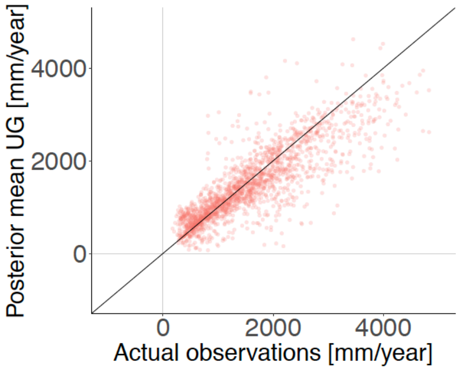
- Remove data.
- Predict the missing values for individual years $\in \{1996, 2005\}$.
- Evaluate the predictive performance of the areal and centroid model, compare to Top-Kriging.

Ungauged catchments (UG): Treat the target catchments as ungauged.

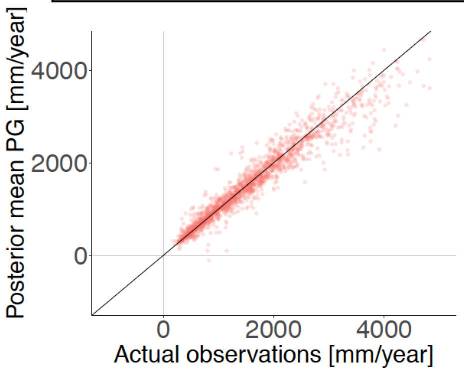
Partially gauged catchments (PG): Include one random observation from the target catchment from one of the years $\in \{1996, 2005\}$. Assess the value of short records.

Results: Partially gauged catchments (PG)

$\sigma_c = 902$ mm/year. $\sigma_c^2/(\sigma_c^2 + \sigma_x^2) = 90\%$.
$\sigma_x = 270$ mm/year. The climatic effects are dominating.

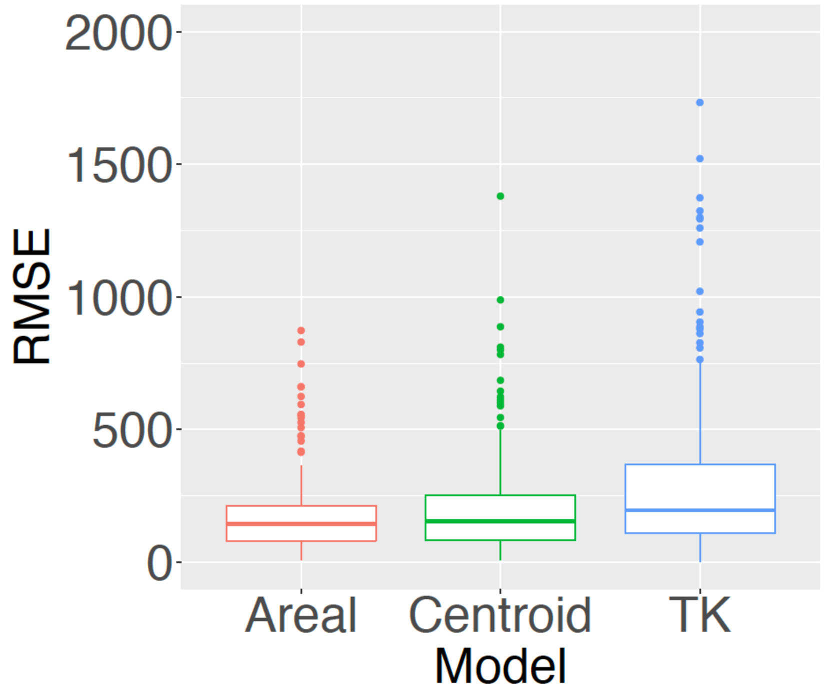


UG for the areal model.
RMSE=363 mm/year.



PG for the areal model.
RMSE= 184 mm/year.

The average reduction in RMSE is around 50 % when adding a short-record of length 1.



The areal and centroid models outperform Top-Kriging for partially gauged catchments.

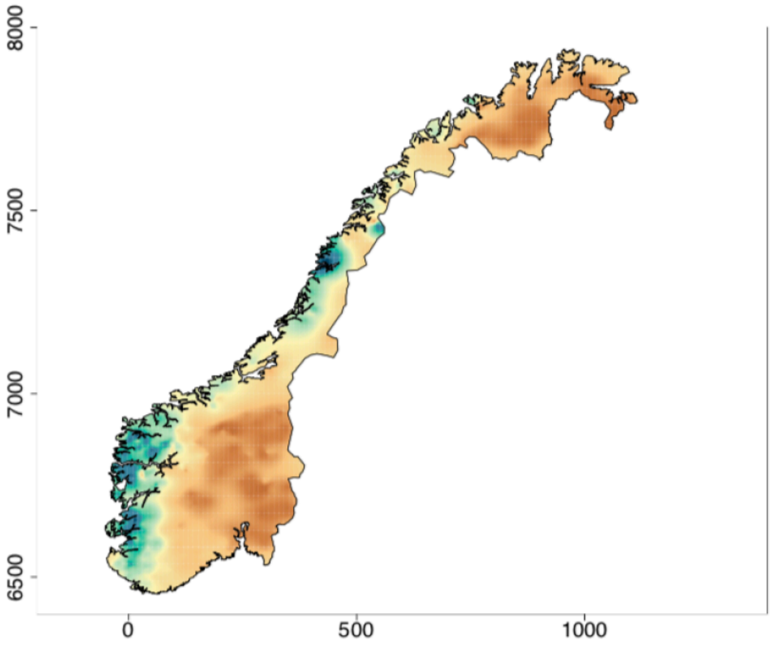
Top-Kriging treats each year individually, while we model 10 years simultaneously.

Conclusions short records

The spatial variability is stable in Norway over years ($\sigma_c >> \sigma_x$).

→ There is a lot of information stored in the short runoff records in Norway.

→ Comparing UG and PG, the reduction in RMSE was on average 50% when adding a short-record of length 1.



Thank you!

Want to read more?

- Thea Roksvåg, Ingelin Steinsland, Kolbjørn Engeland , *A geostatistical two field model that combines point observations and nested areal observations, and quantifies long-term spatial variability -- A case study of annual runoff predictions in the Voss area* Under revision, available at arXiv:1904.02519

- Roksvåg, T., Steinsland, I., and Engeland, K.: *A geostatistical framework for estimating flow indices by exploiting short records and long-term spatial averages – Application to annual and monthly runoff*, Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-415, in review, 2019.

NVE NTNU