

Estimating **mean annual runoff** by using a geostatistical spatially varying coefficient model that incorporates process-based simulations and short records

T.Roksvåg^{1,2}, I. Steinsland² and K.Engeland³

¹ Norwegian Computing Center (NR), Oslo, Norway

² Norwegian University of Science and Technology (NTNU), Trondheim, Norway

³ The Norwegian Water Resources and Energy Directorate (NVE), Oslo, Norway



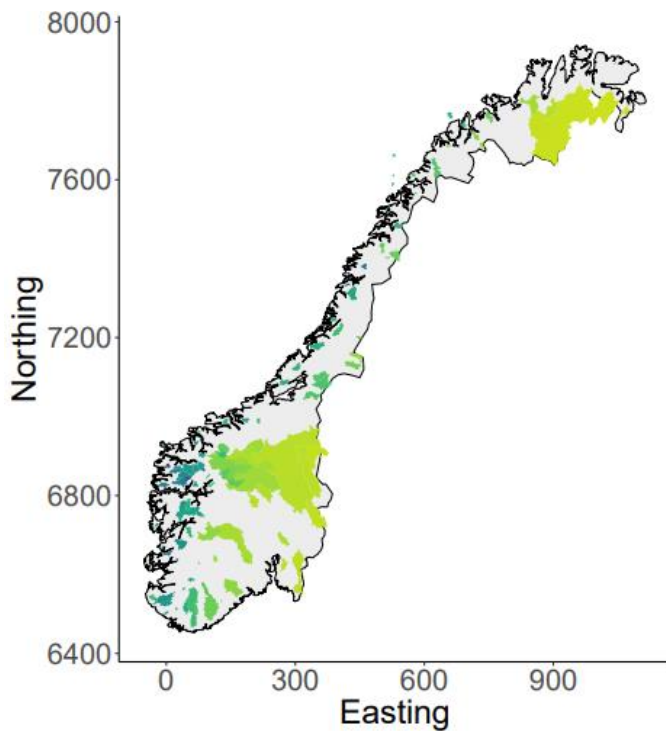
Motivation

Main goal: Estimate mean annual runoff in ungauged and partially gauged catchments.

How: We want to exploit several data sources and suggest to use a **process-based model** in combination with a **geostatistical approach**.

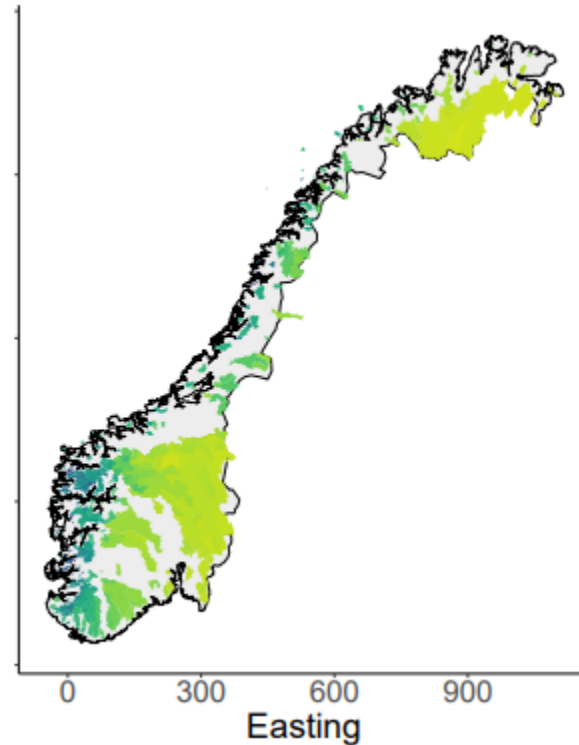


Data available for Norway (1981-2010)



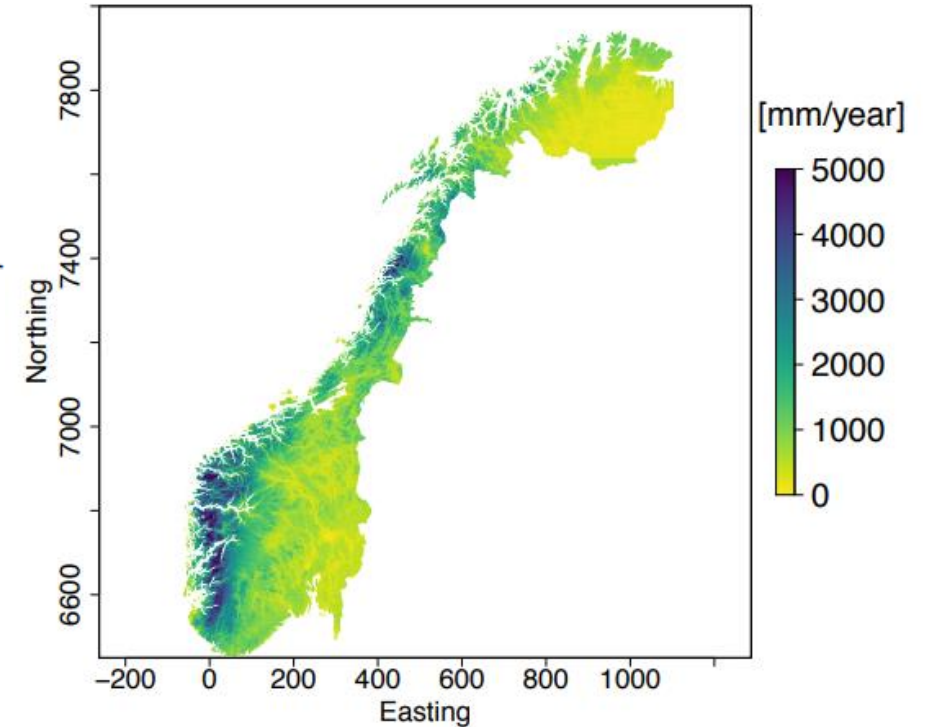
1) Streamflow obs. from fully gauged catchments:
127 catchments with annual streamflow obs. from the whole 30 year target period .

Areal referenced data



2) Streamflow obs. from partially gauged catchments (short records):
284 catchments have annual streamflow obs. for 1-29 years in the 30 year target period.

Areal referenced data



3) Simulations of mean annual runoff from a process-based hydrological model
The **HBV model** [1] was used to make simulations on a 1 km x 1 km grid. Land use, temperature and precipitation information was used as input.

Gridded data

Geostatistical model

The true mean annual runoff $q(\mathbf{u})$ at a location $\mathbf{u}$ is modeled as

$$q(\mathbf{u}) = \beta_0 + (\beta_1 + \alpha(\mathbf{u})) \cdot h(\mathbf{u}) + x(\mathbf{u}); \quad \mathbf{u} \in R^2$$

$$x(\mathbf{u}) \sim \text{GRF}(\sigma_x, \rho_x)$$

$$\alpha(\mathbf{u}) \sim \text{GRF}(\sigma_\alpha, \rho_\alpha)$$

$h(\mathbf{u})$: The simulated mean annual runoff for grid cell $\mathbf{u}$, generated by a process-based model.

$\text{GRF}(\sigma, \rho)$: Gaussian random field with range ρ and standard deviation σ that captures spatial variability.

β_0 is an intercept and β_1 is a regression coefficient.

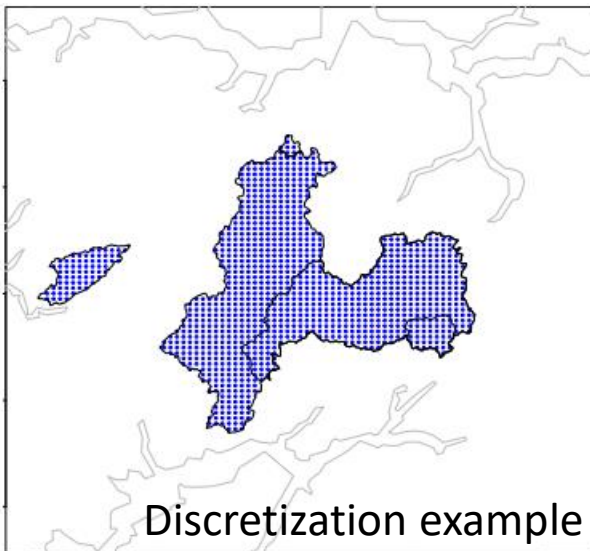
Spatially varying coefficient (SVC): The regression coefficient that models the relationship between the simulations $h(\mathbf{u})$ and the true runoff $q(\mathbf{u})$ varies in the study area according to $\alpha(\mathbf{u})$.

Observation model

The true mean annual runoff $q(u)$ is observed through **areal referenced** streamflow observations, with **uncertainty**:

$$y_i = \frac{1}{n_i} \sum_{u \in L_{A_i}} q(u) + \epsilon_i \quad \epsilon_i \sim N(0, \sigma_y^2)$$

L_{A_i} is a discretization of catchment A_i with n_i grid nodes.



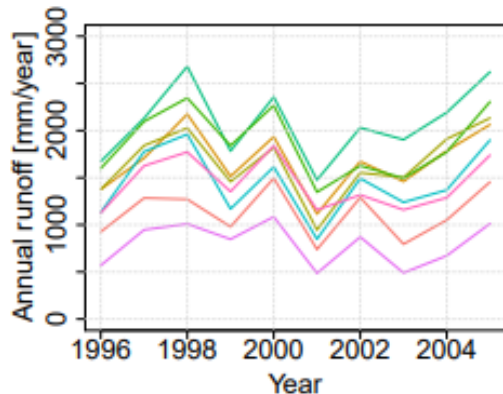
The nested structure of catchments is taken into account by modeling the observations as **areal referenced**.

How to include short records?

We perform **record augmentation** as a **preprocessing step** for partially gauged catchments that have less than 30 years of data. This gives us y_i for the partially gauged catchments.

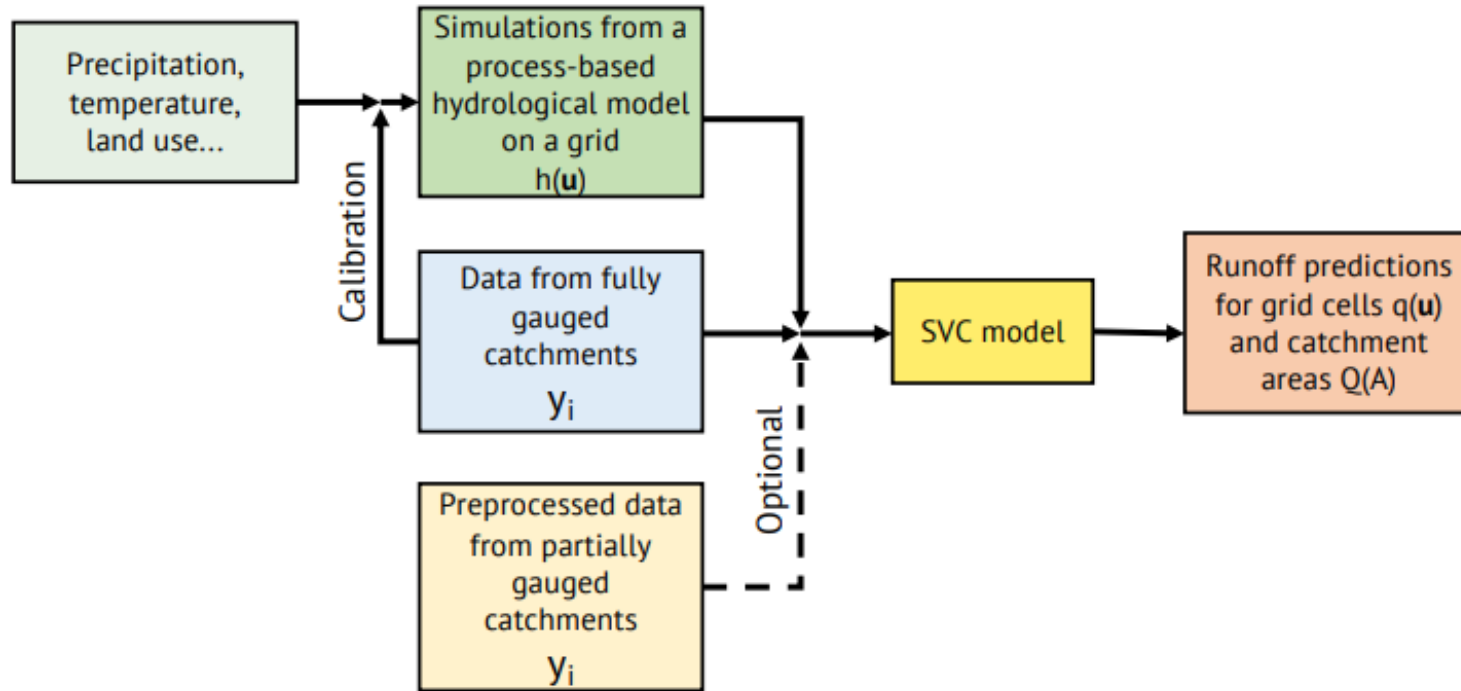
The approach from Roksvåg et. al (2020) is used [2]. Here, we exploit that the Norwegian annual runoff follows **repeated spatial patterns** due to e.g. orographic precipitation.

The data y_i from the fully gauged catchments are weighted more than the data from the partially gauged catchments in our observation likelihood.



Example: Time series of annual runoff from 8 Norwegian catchments. The ranking between the catchments, from wet to dry, is approximately constant over time. According to [2], this is a setting for which the value of short records is high.

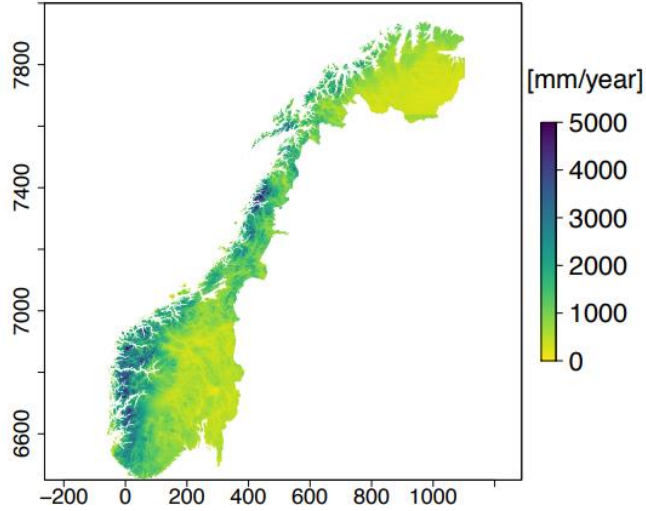
The workflow of the SVC model



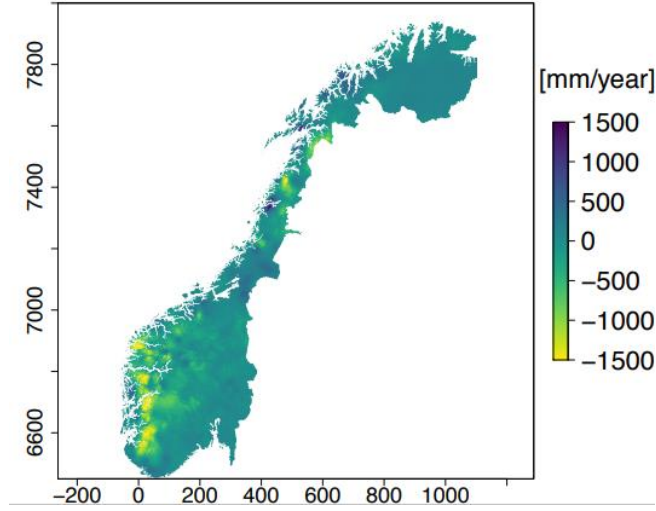
Statistical inference

Fast and approximative **Bayesian inference** by using **INLA** [3], and the **SPDE** [4] approach to spatial modeling

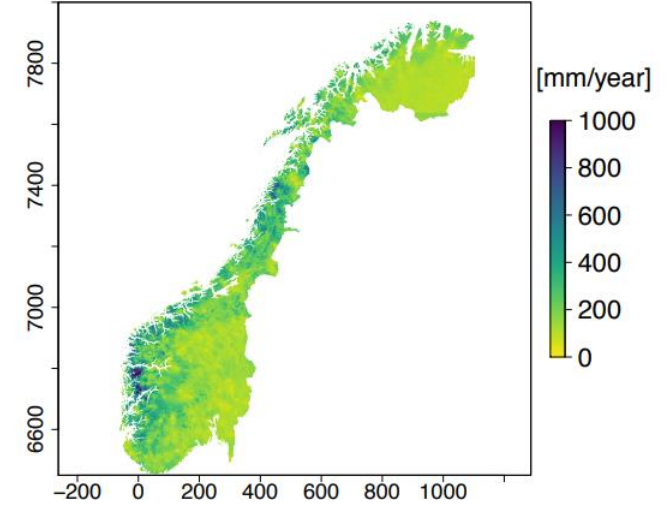
Results: A gridded mean annual runoff map



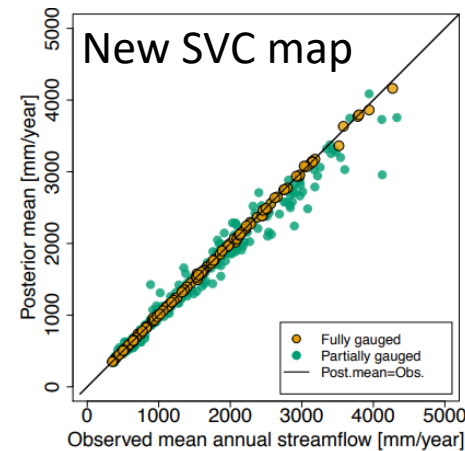
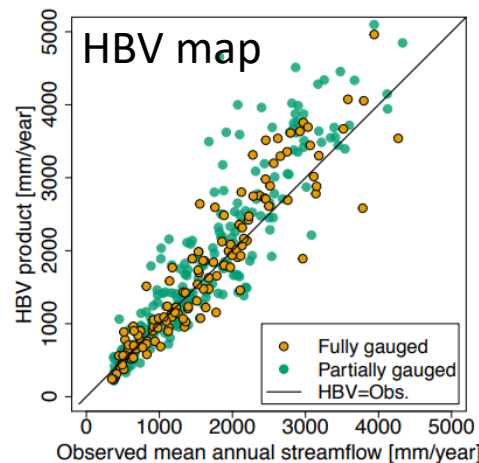
Posterior mean for SVC model



Difference between new SVC map and original HBV map (SVC – HBV)



Posterior standard deviation



The fit between the observations (x-axis) and the estimated runoff (y-axis) is considerably better for the new SVC map

Cross-validation for 127 catchments

We compare three methods:

SVC: The proposed approach.

GS: The geostatistical method from Roksvåg et. al. (2020) [\[2\]](#).

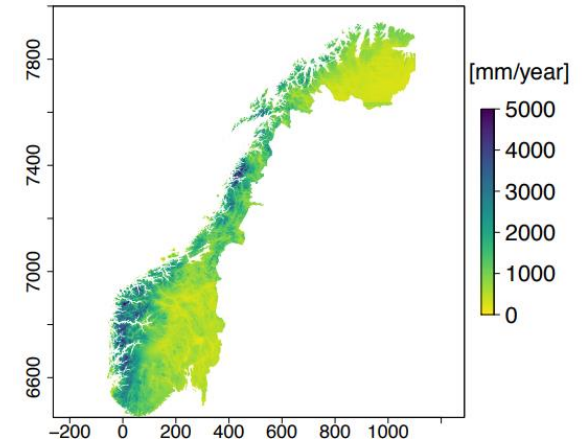
HBV: A process-based model.

Two evaluation settings:

Ungauged: The target catchments are treated as ungauged with zero annual observations between 1981 and 2010.

Partially gauged: The target catchments are treated as partially gauged and have 3 annual observations between 1981 and 2010 (a short record).

<u>Results</u> Criterion	HBV	Ungauged catch.		Partially gauged catch.	
		SVC	GS	SVC	GS
RMSE (mm/yr)	394	315	389	166	134
ANE (1)	0.18	0.11	0.192	0.054	0.047
CRPS (mm/yr)	235	145	209	73	71



Main conclusions: (i) The SVC model performs better than the process-based model (HBV) and the geostatistical model (GS) for ungauged catchments. (ii) For partially gauged catchments a purely geostatistical method performs better, but the SVC model is indeed able to exploit short records.

References

- [1] S. Bergström. Development and Application of a Conceptual Runoff Model for Scandinavian Catchments, volume 134, 1976.
- [2] T. Roksvåg, I. Steinsland, and K. Engeland: Estimation of annual runoff by exploiting long-term spatial patterns and short records within a geostatistical framework, Hydrol. Earth Syst. Sci., 24, 4109–4133, <https://doi.org/10.5194/hess-24-4109-2020>, 2020.
- [3] H. Rue, S. Martino, and N. Chopin. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations. Journal of the Royal Statistical Society: Series B, 2009.
- [4] F. Lindgren, H. Rue, and J. Lindström. An explicit link between Gaussian fields and Gaussian markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society: Series B, 2011