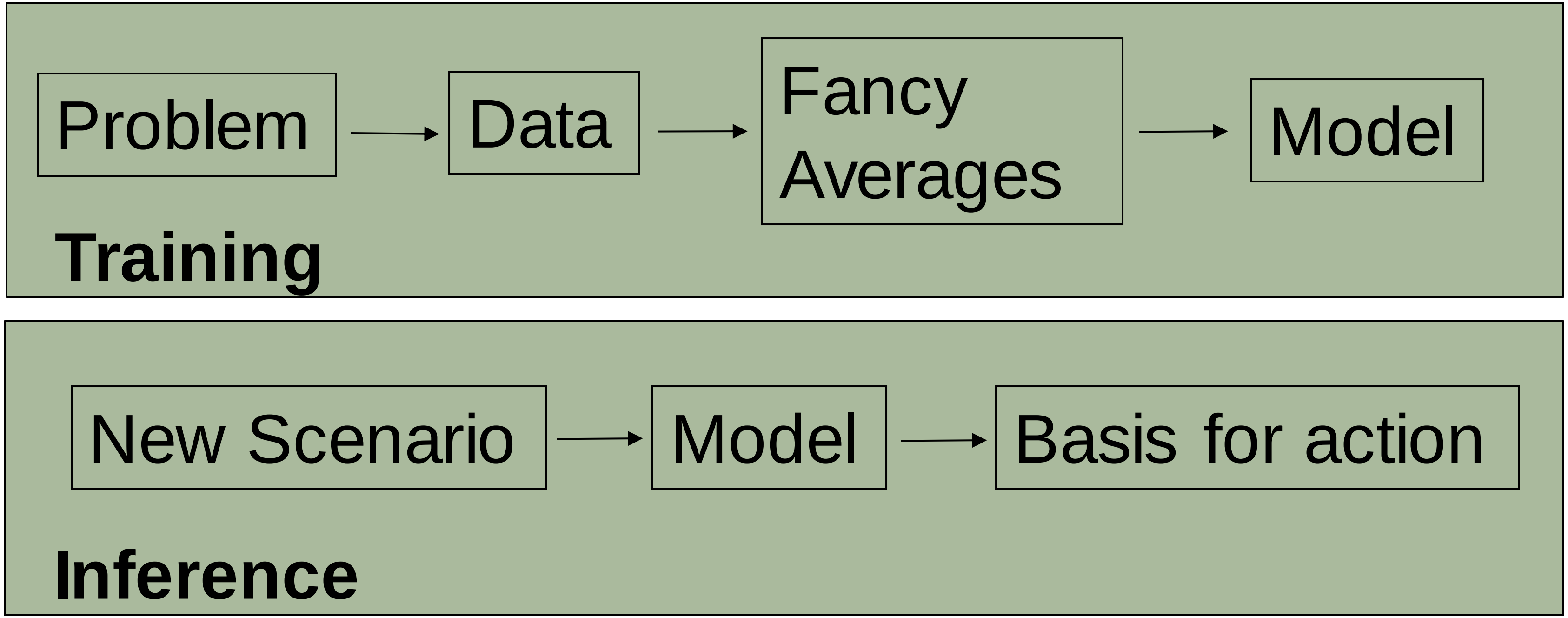
NORSK REGNESENTRAL
NORWEGIAN COMPUTING CENTER

Muligheter med maskin læring og 'stor data' i klimatilpasning

Alex Lenkoski, Forskningsleder + Colleagues

Modeling and Data – What is it honestly?



Motivations for modelling extremes (by A. Davison)

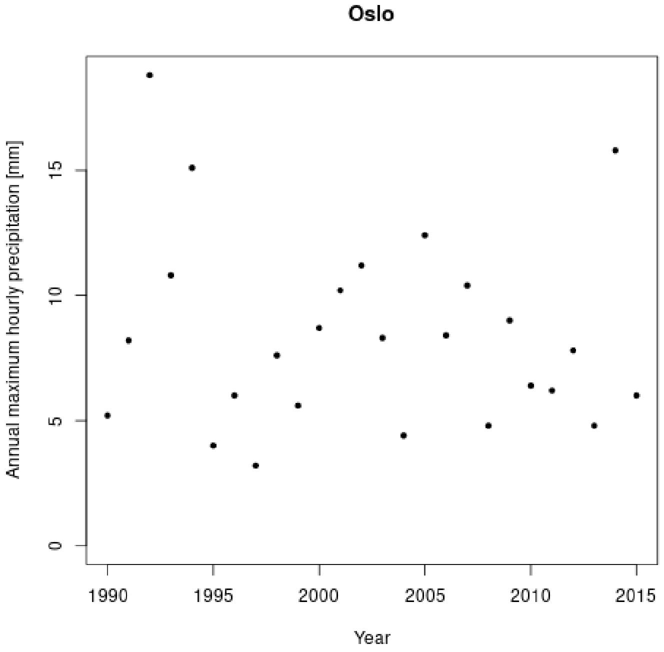
- pointwise maps of return levels (*”should I buy this house?”*)
- estimation of probability of rare complex events (*“what is the risk of insuring these houses?”*)
- detection/attribution (*“is the floodwater in my hall caused by climate change?”*)
- short range forecasting (*“should I leave the house?”*)

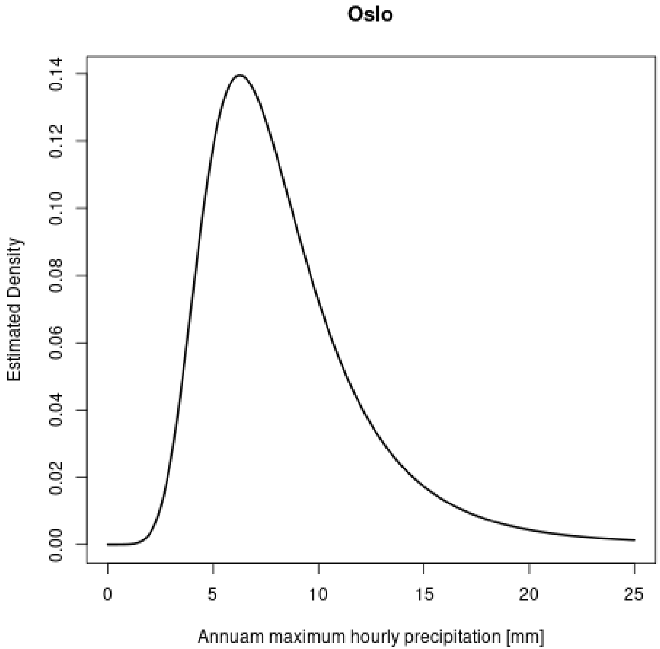
A down-in-the trenches rule of thumb for method choice

Sample Size	Method	"Discipline"
Less than 30	Average	Statistics
30 to 1.000	Regression	Statistics
1.000 to 10.000	• GAMs • Lasso	Statistics ML
10K to 1M	• Random Forests • Gradient Boosting	ML
More than 1M	Neural Nets	AI

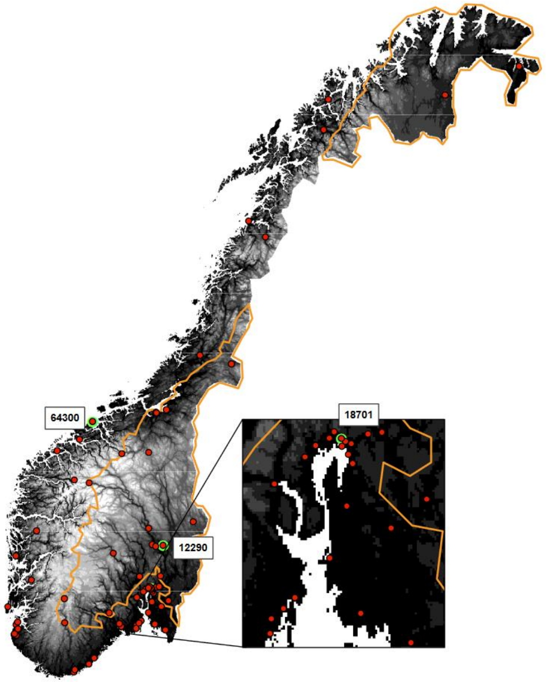
- When we discuss climate extremes, we never have 1M observations
- This means ML/AI methods are not immediately, directly relevant.

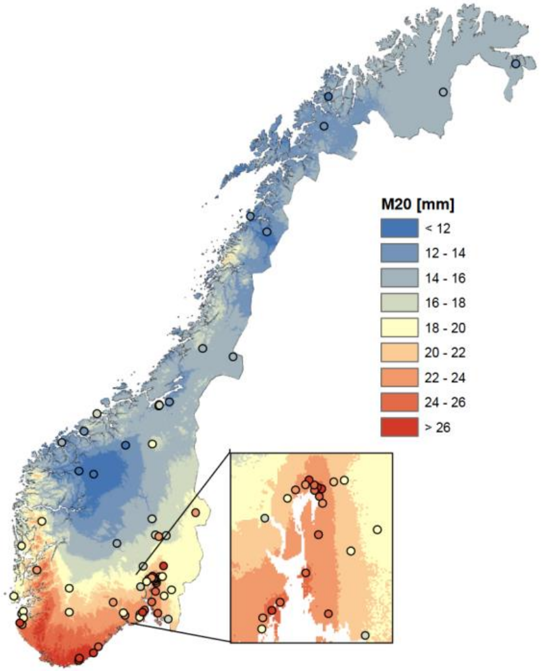
One dimensional extreme





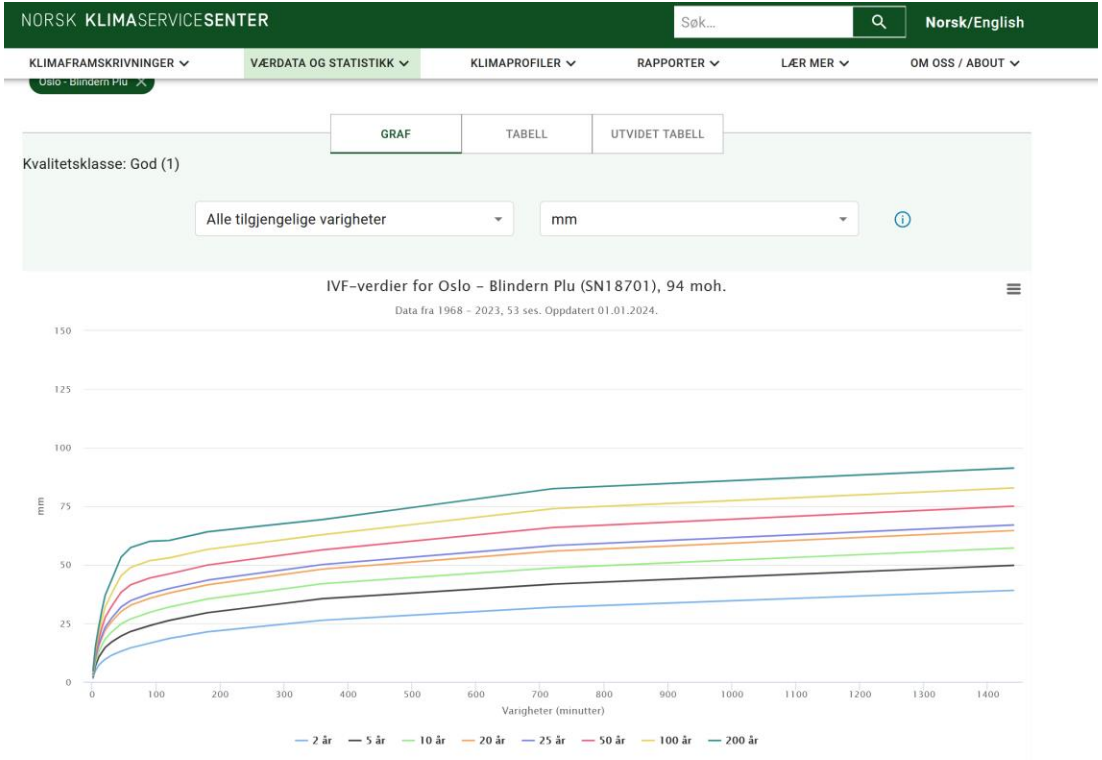
Spatial Extremes – Combining data from 70 stations





Dyrrdal, A. V., Lenkoski, A., Thorarinsdottir, T. L., & Stordal, F. (2015). Bayesian hierarchical modeling of extreme hourly precipitation in Norway. Environmetrics, 26(2), 89-106

Return levels across durations



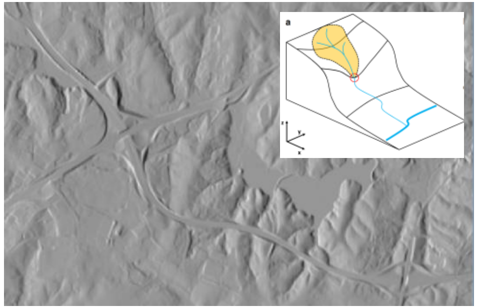
Roksvåg, T., Lutz, J., Grinde, L., Dyrrdal, A. V., & Thorarinsdottir, T. L. (2021). Consistent intensity-duration-frequency curves by post-processing of estimated Bayesian posterior quantiles. *Journal of Hydrology*, *603*

Physical Climate Risk Example – Collaboration with Gjensidige

- **Insurance data** (Gjensidige):
- 12 years of data, 1.7 million contracts, 32 500 claims
- Individual building attributes and coordinates

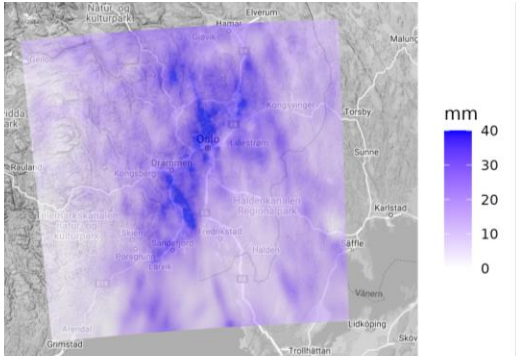
Insurance data

Topography



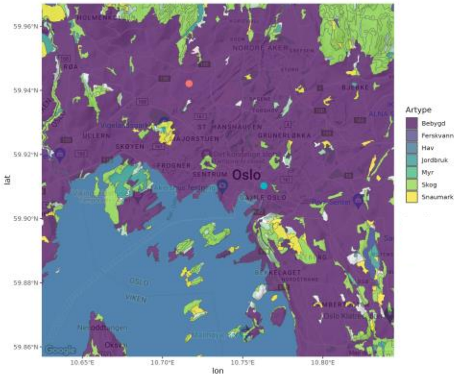
Risk model

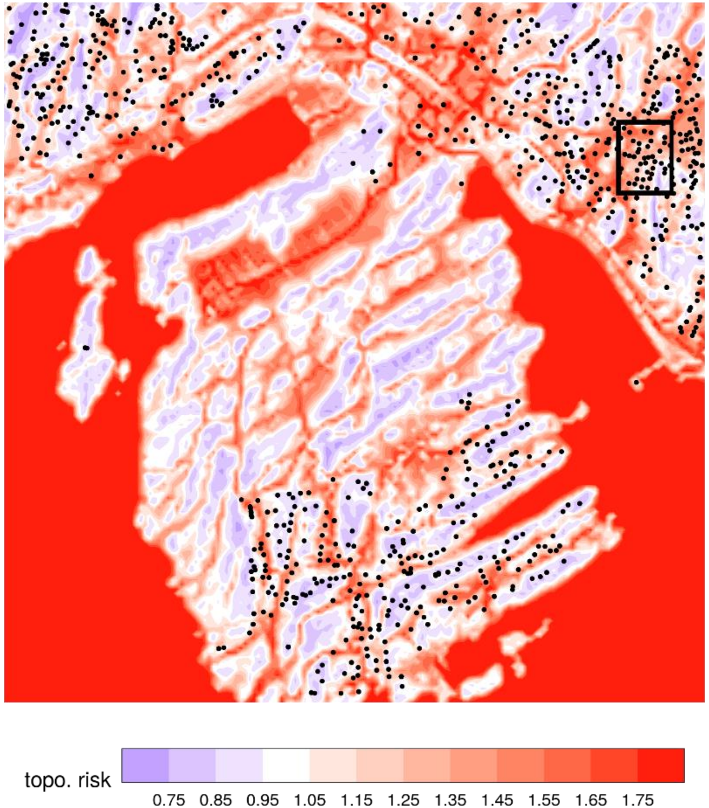
Rainfall

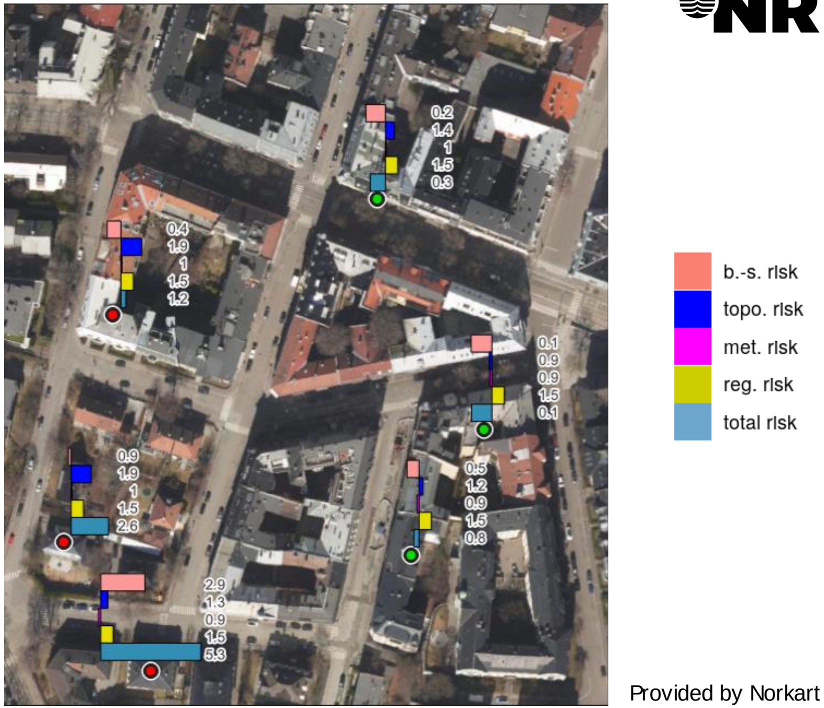


Regional unit

Area type





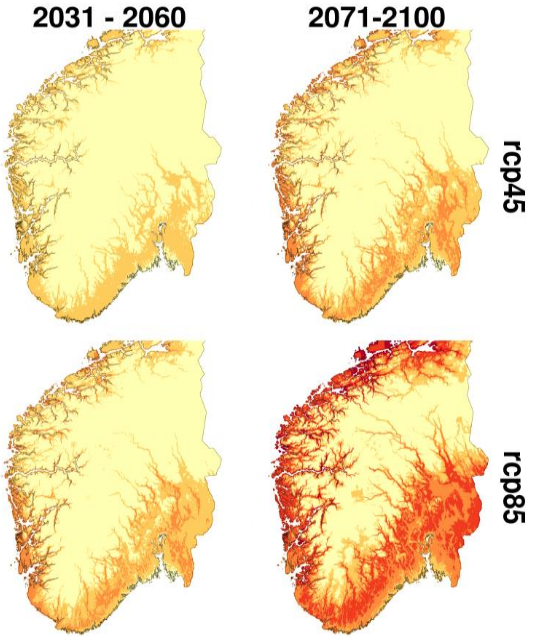


Provided by Norkart

Full resolution contract risk values are fictitious and meant for illustration!

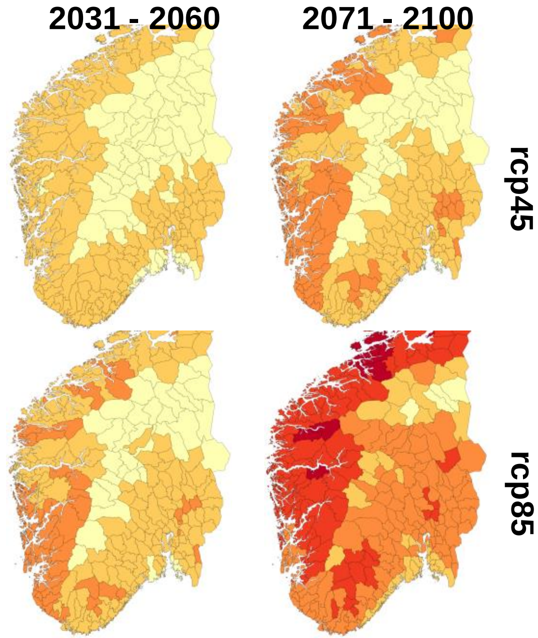
Projected change of climatological risk

$$r_i = r_0 r_i^{\text{clim}} r_i^{\text{topo}} r_i^{\text{b-s}} r_i^{\text{reg}}$$



Projected change of total risk

$$r_i = r_0 r_i^{\text{clim}} r_i^{\text{topo}} r_i^{\text{b-s}} r_i^{\text{reg}}$$

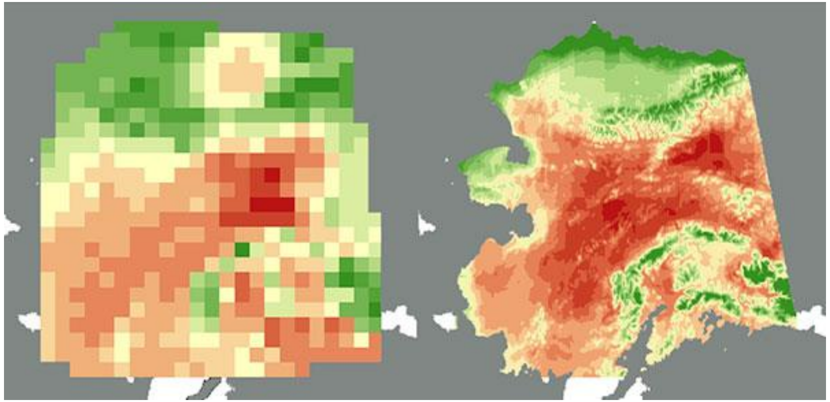


Ratio: 0.95-1.05 1.05-1.15 1.15-1.25 1.25-1.35 >1.35

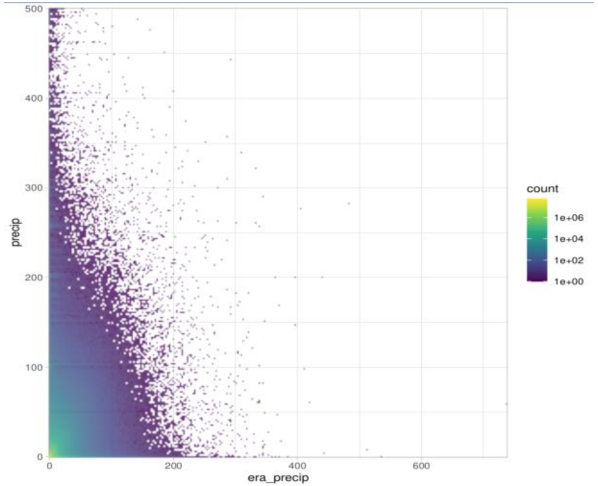
Heinrich-Mertsching, C., Wahl, J. C., Ordoñez, A., Stien, M., Elvsborg, J., Haug, O., & Thorarinsdottir, T. L. (2023). Assessing present and future risk of water damage using building attributes, meteorology, and topography. *Journal of the Royal Statistical Society Series C: Applied Statistics*, *72*(4), 809-828.

Stochastic downscaling

https://uaf-snap.org/how-do-we-do-it/downscaling/

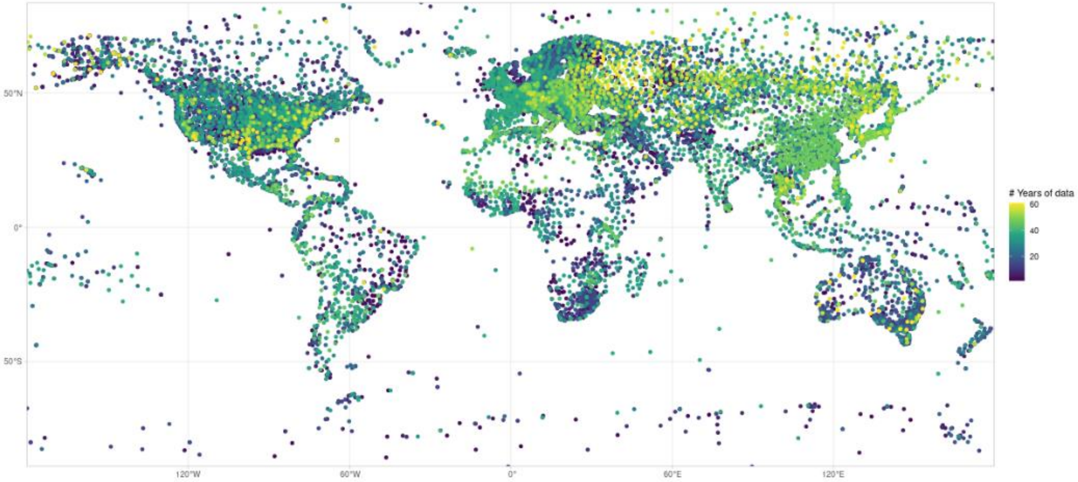


- **What it is:** Guessing what is happening inside of each grid cell
- **Why we need it:** Grid cell averages are too smooth. this leads to
  - Underestimation of variance
  - Underestimation of extremes
  - Mean bias in temperature
  - Etc.



Stochastic downscaling

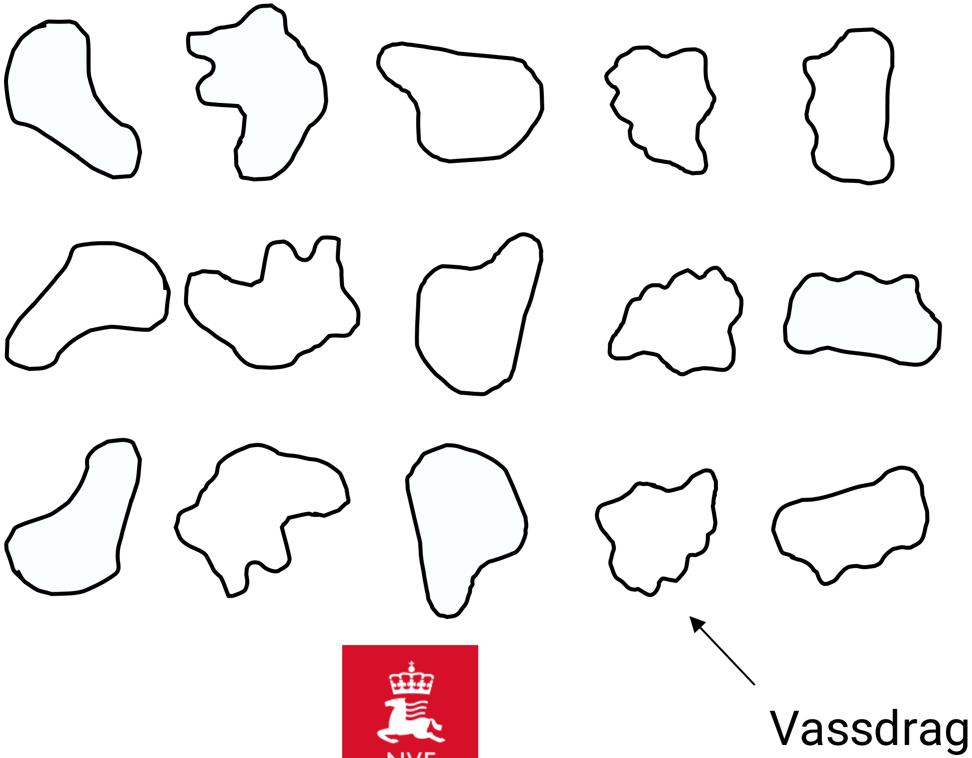
- Model the statistical distribution of point-observations from weather stations, using information from gridded reanalysis data as covariates
- Reanalysis data: ERA5
- Weather station data: GSOD[1]
- DEM data: ETOPO[2]
- Method: Work in progress



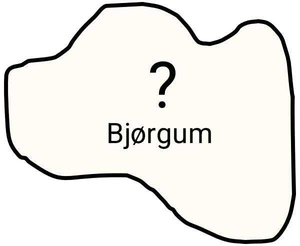
1) Global Summary of the Day, https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00516
2) Earth TOPOgraphy, https://www.ncei.noaa.gov/products/etopo-global-relief-model

Fra NVE-vassdrag til Småkraft-vassdrag

NR



NVE



småkraft©

Posisjon, areal, klimaforhold, gradient, elvelengde, % skog, fjell, innsjø

Senter for forskningsdrevet innovasjon

climatefutures

Replacing Physical Hydrological with Neural Nets

The Journal of Open Source Software

NeuralHydrology — A Python library for Deep Learning research in hydrology

Frederik Kratzert[1], **Martin Gauch**[2], **Grey Nearing**[1], **and Daniel Klotz**[2]

1 Google Research **2** Institute for Machine Learning, Johannes Kepler University Linz, Linz, Austria

DOI: 10.21105/joss.04050

Software

- Review
- Repository
- Archive

Summary and statement of need

Since ancient times humans have strived to describe environmental processes related to water (Angelakis et al., 2012; Biswas, 1970). Throughout this history, hydrologists built various process-based prediction models that simulate processes from soil moisture to streamflow

- Experimenting with replacing physical hydrological models with LSTMs
- Have been positively surprised with the quality of results, especially for "left out" basins

Conclusions

- There can be no debating that data will play a substantial role in helping address the problems of climate adaptation
- The Stats/ML/AI distinction is largely meaningless – which method to use is driven by the size of the data at hand
- Yet, extremes by their very nature will be small in size
- Therefore, the most advanced techniques will be used to process the data sources that feed into our extreme analyses, not our extreme analysis itself.