Norsk Regnesentral
NORWEGIAN COMPUTING CENTER

www.nr.no/~thordis
www.climatefutures.no

Machine learning vs. statistical methods for climate data analysis

Thordis L. Thorarinsdottir
Norwegian Computing Center

Tekna

Data driven climate insight
January 19, 2021

Norsk Regnesentral
NORWEGIAN COMPUTING CENTER

www.nr.no/~thordis
www.climatefutures.no

Machine learning or statistical methods for climate data analysis?

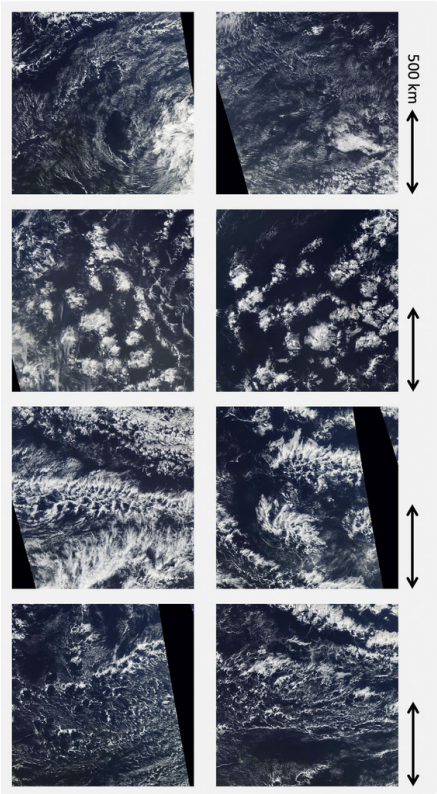
Thordis L. Thorarinsdottir
Norwegian Computing Center

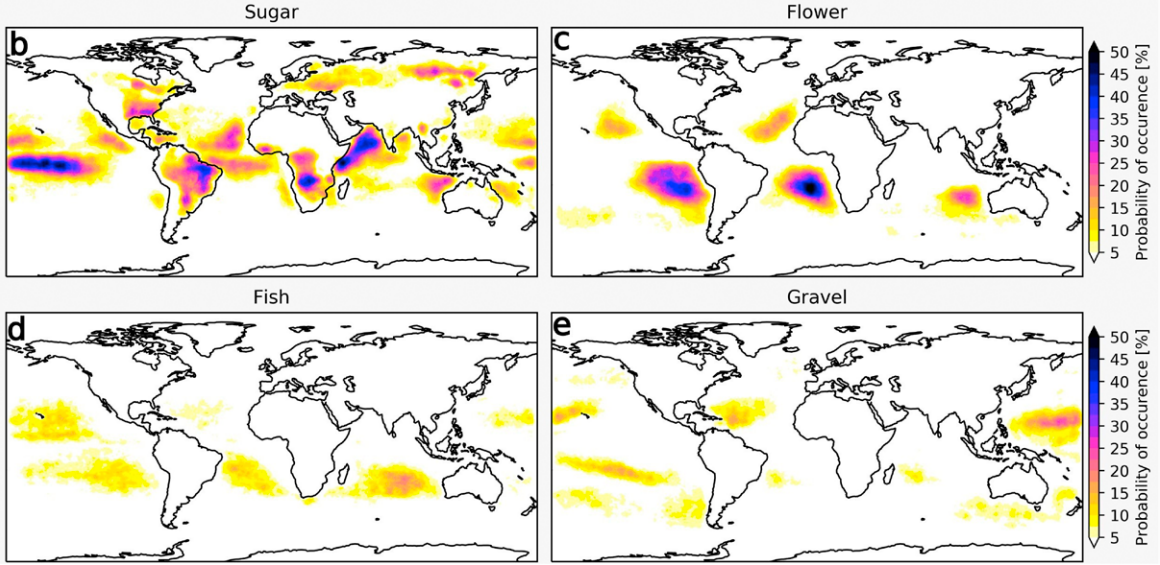
Tekna

Data driven climate insight
January 19, 2021

It depends…

Sugar, Flower, Fish and Gravel: Finding clouds with deep learning





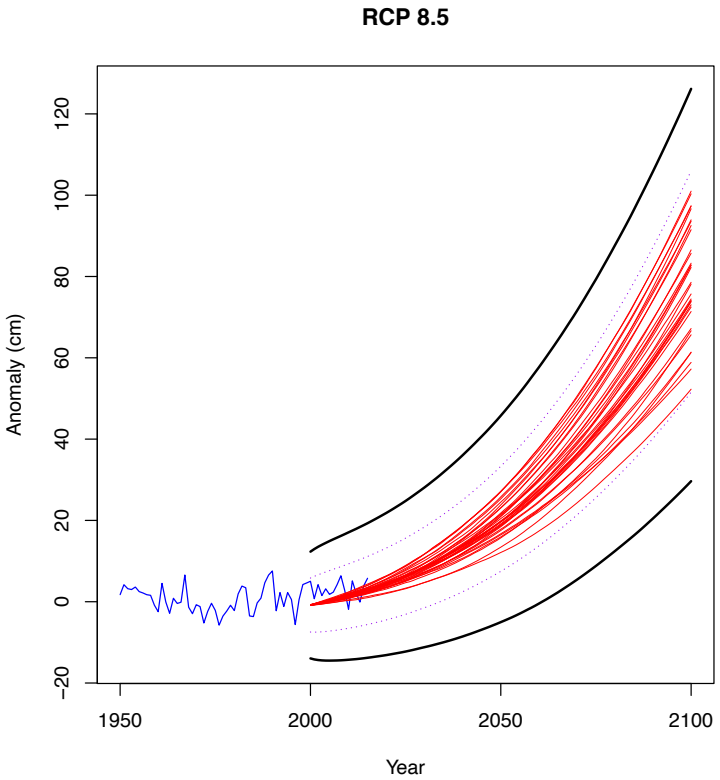
DNNs are great for **detecting features in images**.

Issue: Need large hand-labelled training data sets.

Proposed solution: Crowdsourcing.

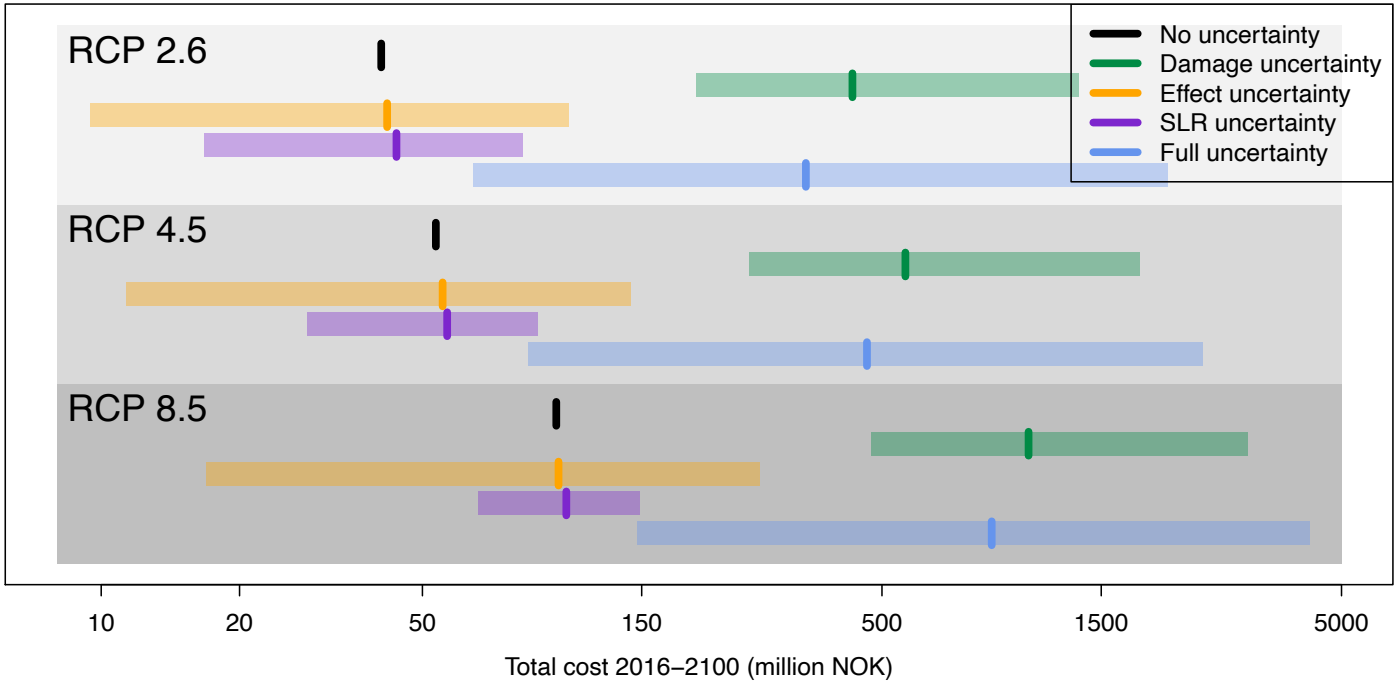
Rasp et al. (BAMS, 2020)

Decision support for decision-making on coastal adaptation



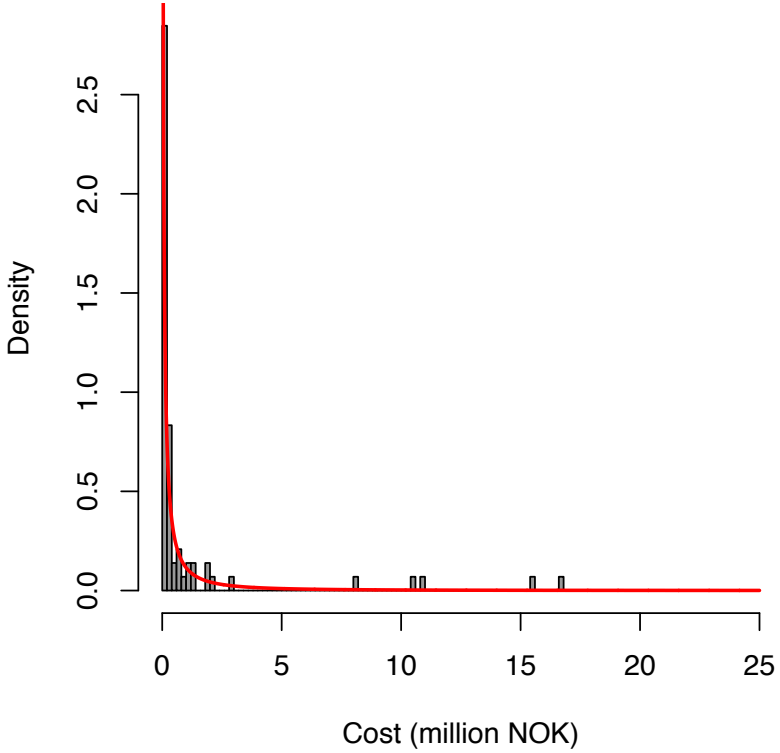
Thorarinsdottir et al. (WRR, 2017)

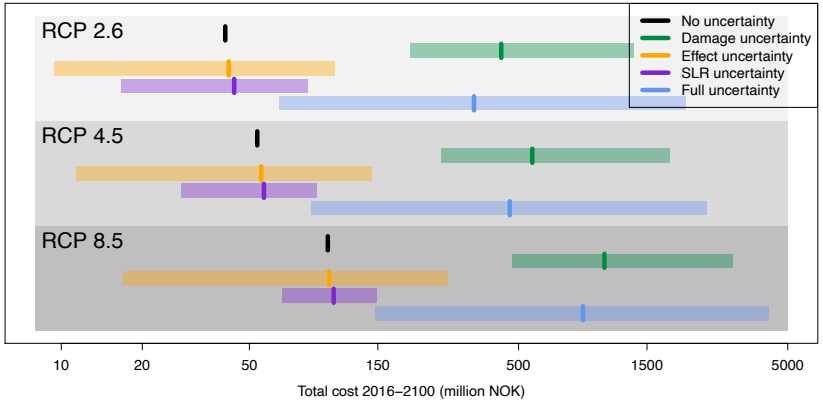
Total costs 2016-2100: damage costs + SLR + effect of SLR on damage



Thorarinsdottir et al. (WRR, 2017)

Distribution of damage cost is highly skewed



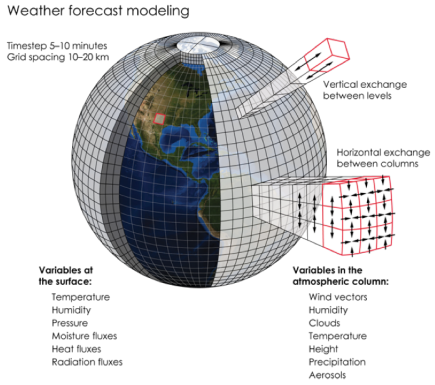


Thorarinsdottir et al. (WRR, 2017)

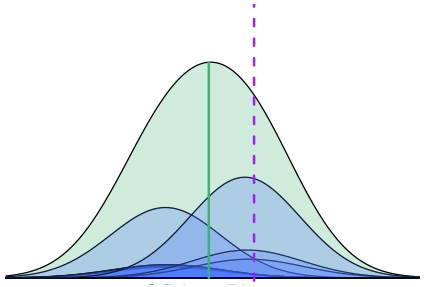
Decision support for decision-making on adaptation and mitigation

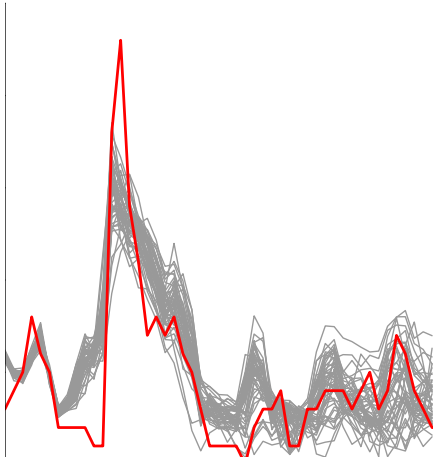
- We need **transparency** between scientists from different disciplines, decision makers, other practitioners, stakeholders and the general public.
- Including **uncertainty** is vital
- **Extrapolation** beyond observed data requires a model

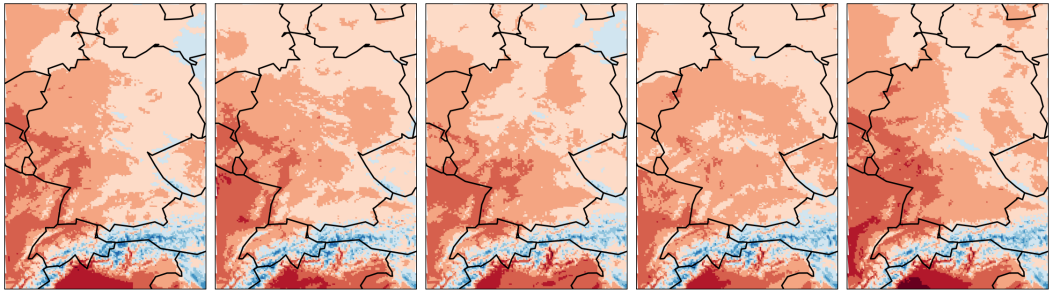
Post-processing of weather and climate predictions



+

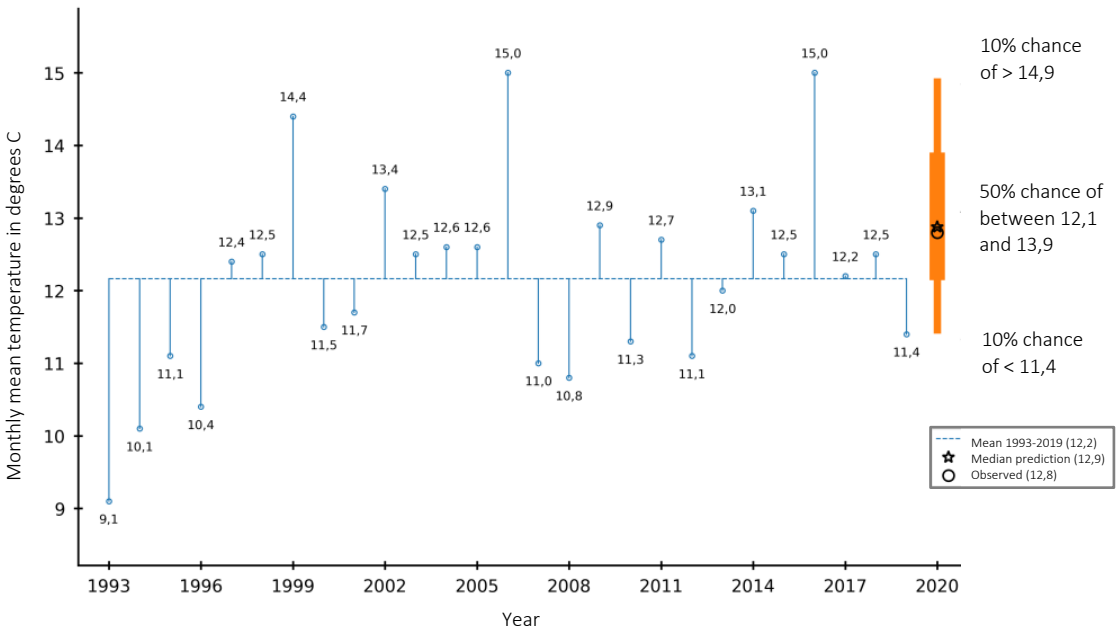






Mean monthly temperature in Oslo next month

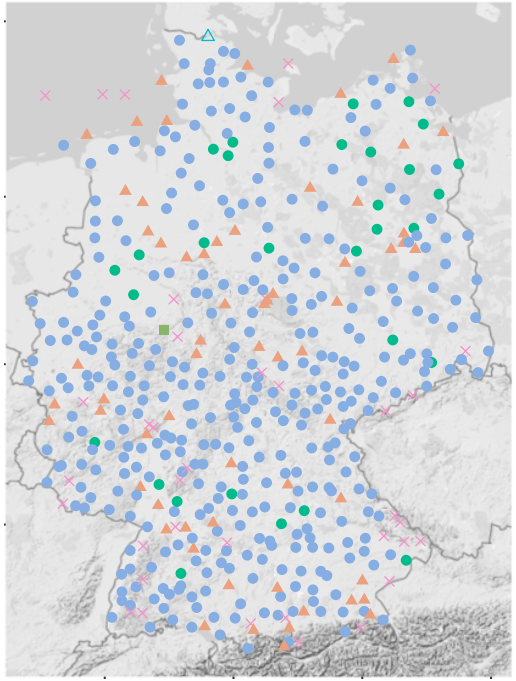
September 2020 forecast issued in August 2020



~200 ensemble members

Results by E. Kolstad and A. Lenkoski.

Daily temperature forecast for Germany



Best model

- EMOS−gl
- EMOS−loc
- EMOS−loc−bst
- QRF
- FCN−aux
- FCN−emb
- FCN−aux−emb
- NN−aux
- NN−aux−emb

Nonlinear relationships
Auxiliary predictors
Latent station features

Feature	Description
Ensemble predictions (mean and std dev)	
t2m	2-m temperature
cape	Convective available potential energy
sp	Surface pressure
tcc	Total cloud cover
sshf	Sensible heat flux
slhf	Latent heat flux
u10	10-m *U* wind
v10	10-m *V* wind
d2m	2-m dewpoint temperature
ssr	Shortwave radiation flux
str	Longwave radiation flux
sm	Soil moisture
u_pl500	*U* wind at 500 hPa
v_pl500	*V* wind at 500 hPa
u_pl850	*U* wind at 850 hPa
v_pl850	*V* wind at 850 hPa
gh_pl500	Geopotential at 500 hPa
q_pl850	Specific humidity at 850 hPa
Station-specific information	
station_alt	Altitude of station
orog	Altitude of model grid point
station_lat	Lat of station
station_lon	Lon of station

Rasp and Lerch (MWR, 2018)

Post-processing of weather and climate predictions

- The choice of methodology should depend on the **sample size**
- Post-processing can provide valuable input to weather/climate model developers if the results are **explainable**

There are many more applications for machine learning and statistical modelling in climate data analysis

- Downscaling numerical model output to higher resolutions
- Representation of subgrid-scale processes
- Combining multi-type information into a single prediction/projection

- For a list of weather and climate datasets preprocessed for AI research, see **mldata.pangeo.io**

Norsk Regnesentral
NORWEGIAN COMPUTING CENTER

www.nr.no/~thordis
www.climatefutures.no

Thank you for your attention!