Deep Neural Networks for the reconstruction of near-surface hourly temperature

Manuel Carrer, Cristian Lussana, Line Båserud, Gianluca Turin and John Bjørnar Bremnes

The Norwegian Meteorological Institute, Oslo, Norway

Norwegian Meteorological Institute

SCREEN CAPTURE WELCOME

2025-09-10

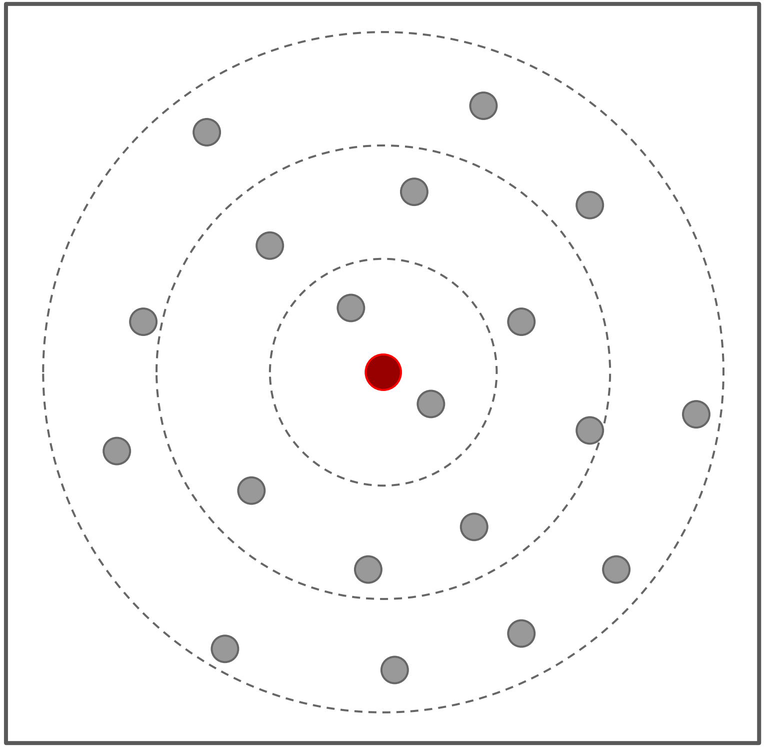
EMS2025-219 OSA3.2

Deep Neural Networks for the reconstruction of near-surface temperature

¿Can we apply well-established deep neural network (DNN) regression to reconstruct daily or hourly temperature at a specific location? given:

1. Hourly temperatures from the nearest observation stations
   a. Target application: Automatic DQC of a dense network of in-situ stations
2. Daily temperatures from geographical, weather, and climate variables
   a. Target application: Spatial analysis for climatological applications

Hourly temperatures from the nearest observation stations



Can we reconstruct hourly 2 m temperature at a target location (red dot) using data from the 15 nearest stations (gray dots) in a sparse in-situ observational network?

The dataset uses the network of MET Norway stations, plus private stations (Netatmo) in the Fennoscandia region.

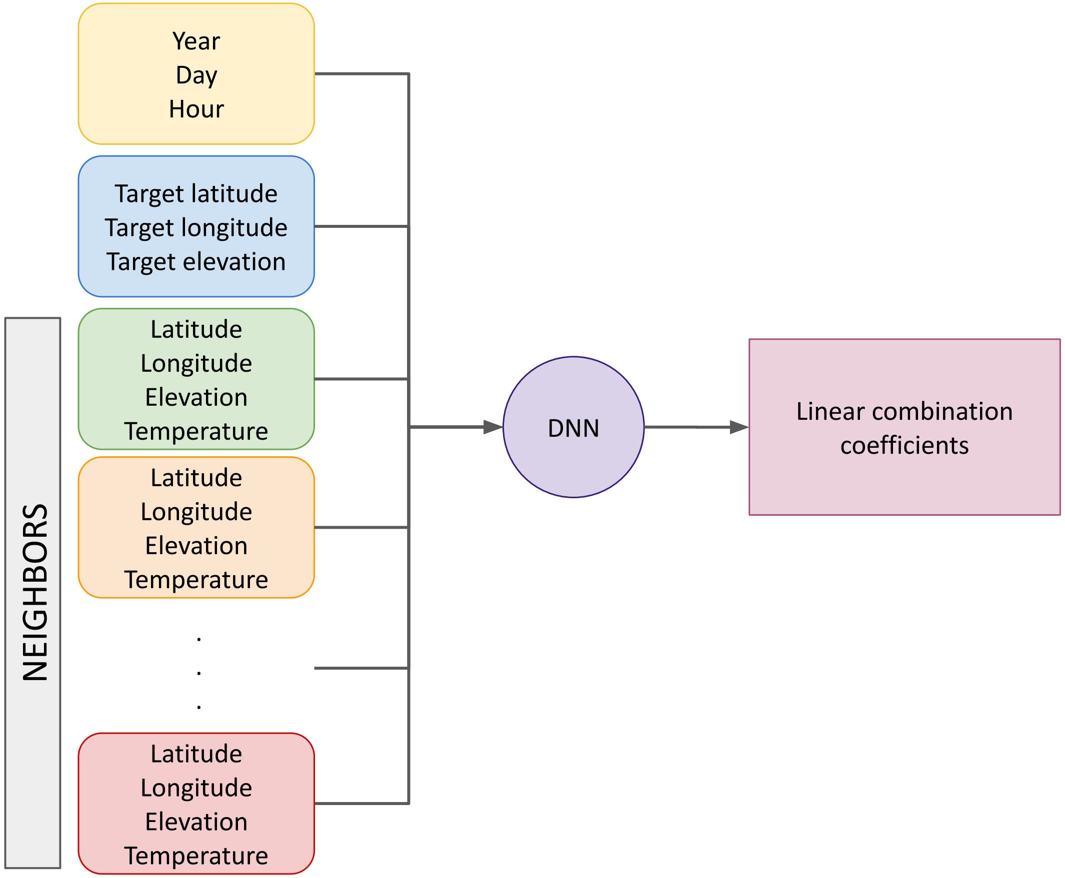
Hourly temperatures from the nearest observation stations

Quantile function as a linear combination:

$$Q(\tau, x) = \sum_{j=0}^{d} \alpha_j(x) B_{jd}(\tau)$$

Bernstein polynomials (binomial distribution):

$$B_{jd} = \binom{d}{j} \tau^j (1-\tau)^{d-j}$$



Bremnes, J. B. (2020).
Monthly Weather Review, *148*(1), 403-414.

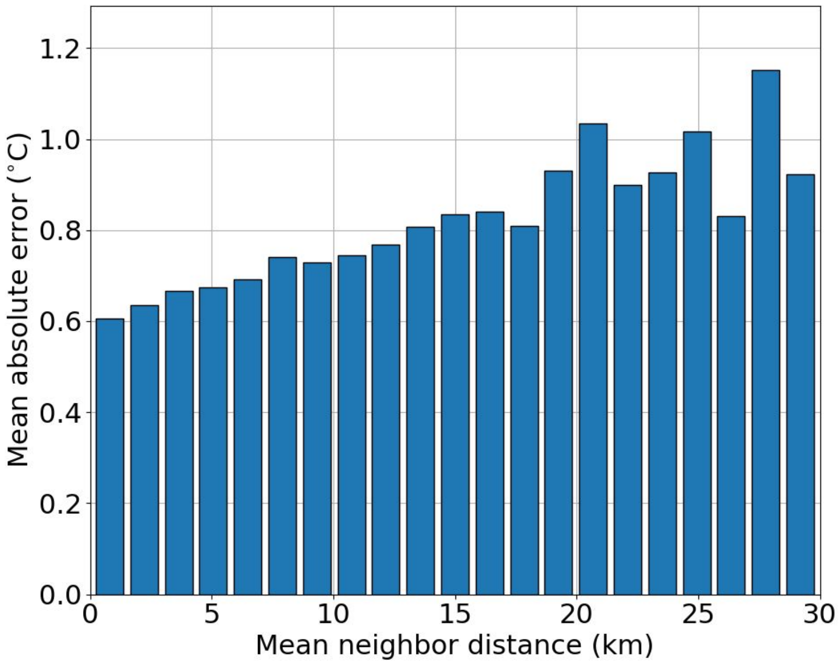
Training setup

- Inputs: 58 features (using 15 neighbors, distances, time encoding)
- Layers: 4 fully connected layers with size 128
- Outputs: 9, since we bound the linear combination to the 8th Bernstein polynomial
- Batch size: 128
- Optimizer: Adam
- Learning rate: 1e-4
- Epochs: 500
- Train one model per month

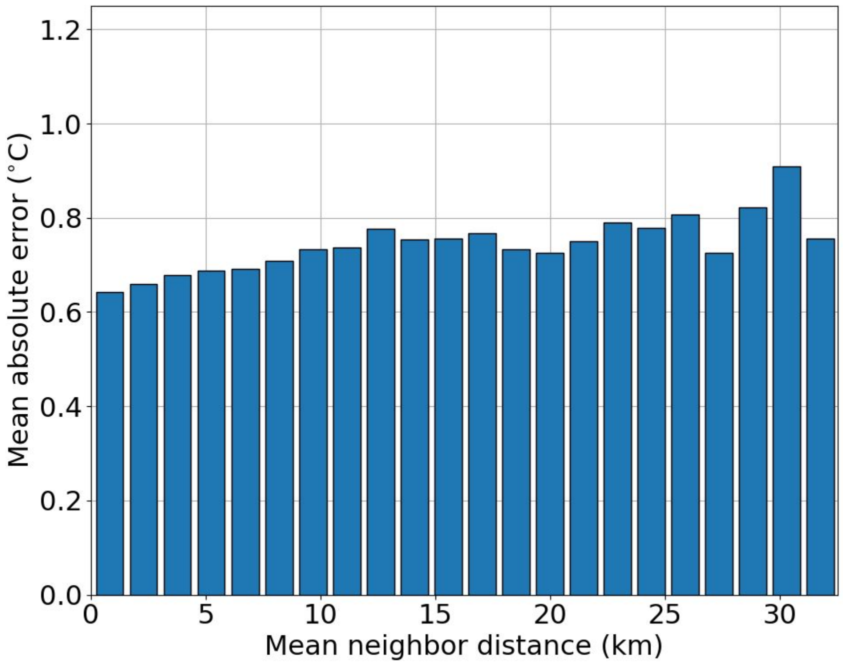
    We might expand later to include multiple/all months

Absolute error versus mean neighbor distance

January

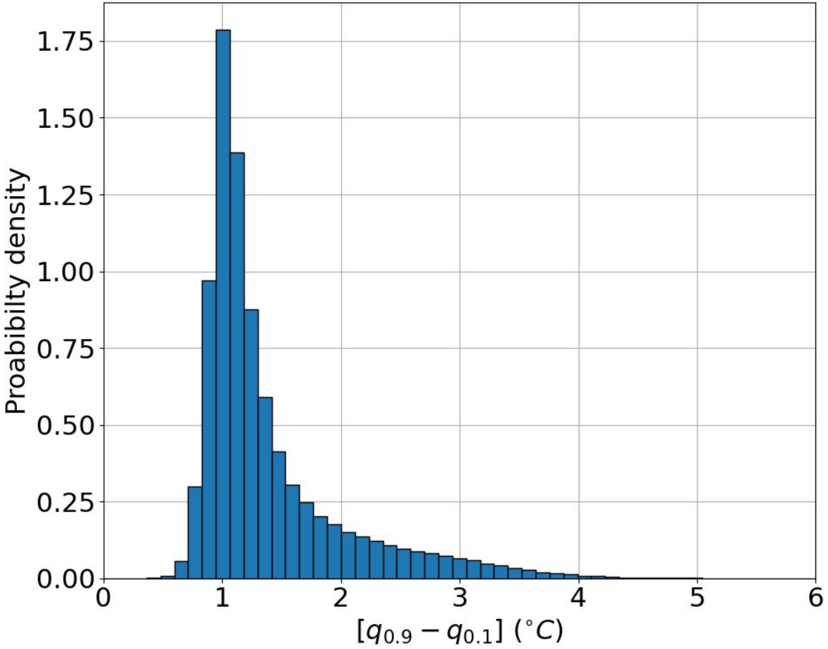


August

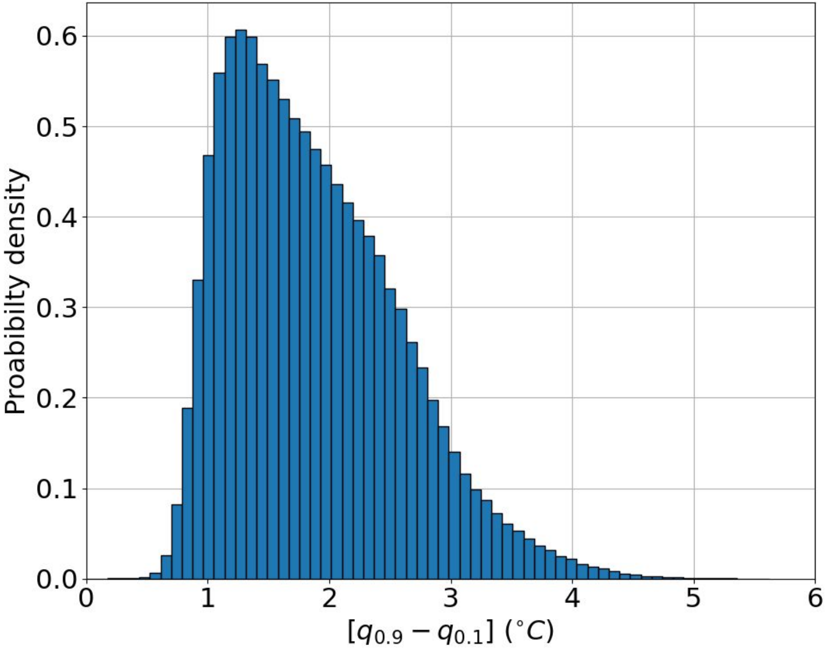


Quantile spread

January

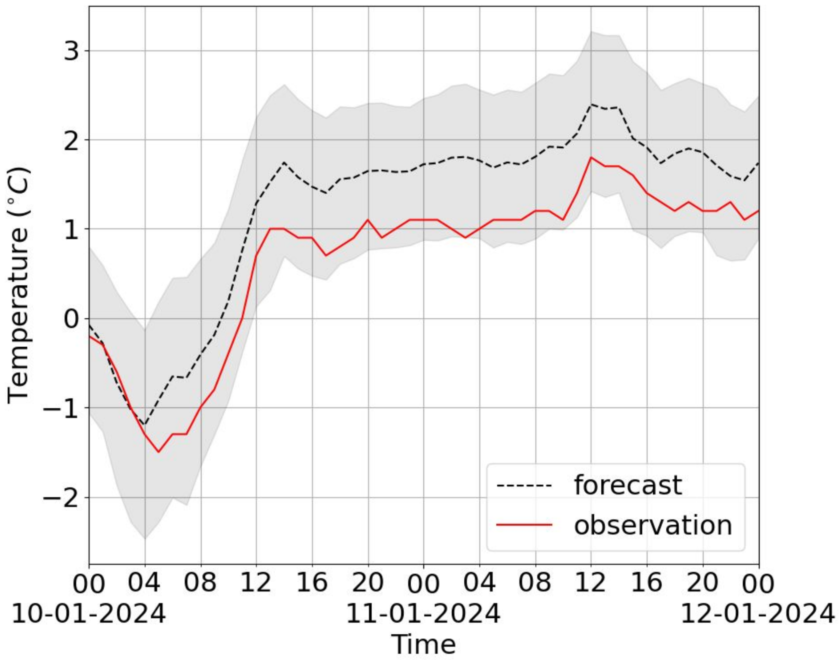


August

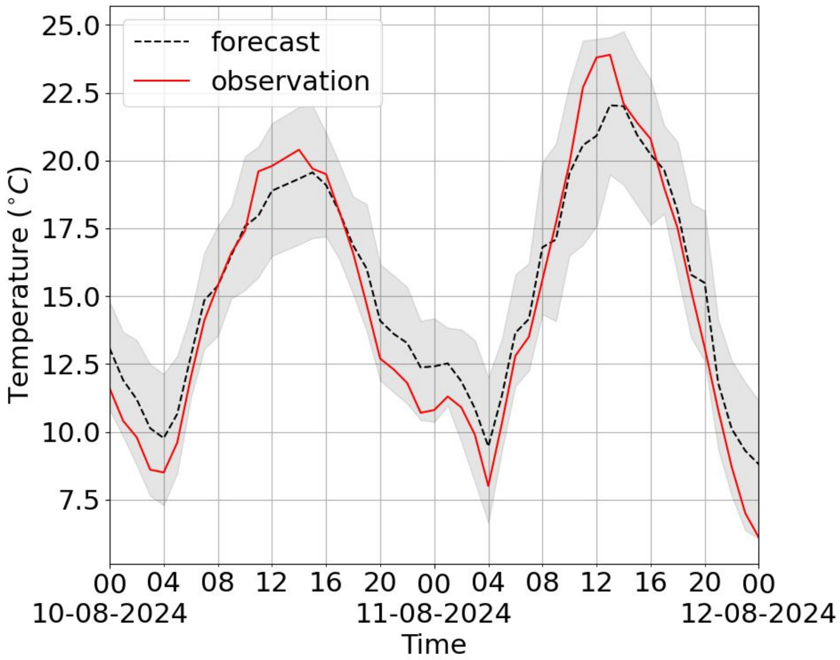


Timeseries forecast

January (Orsta, Norway West Coast)

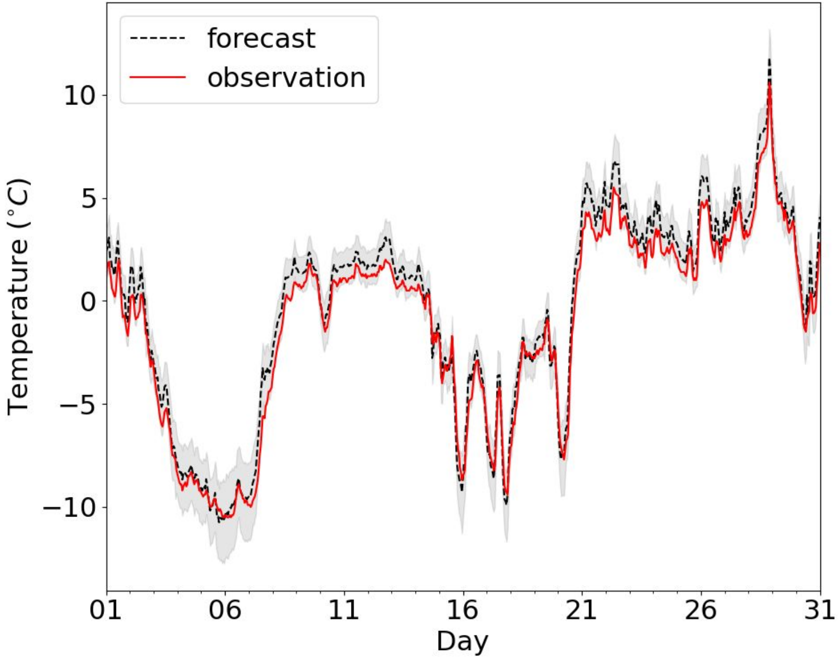


August (Roverud, North East of Oslo)

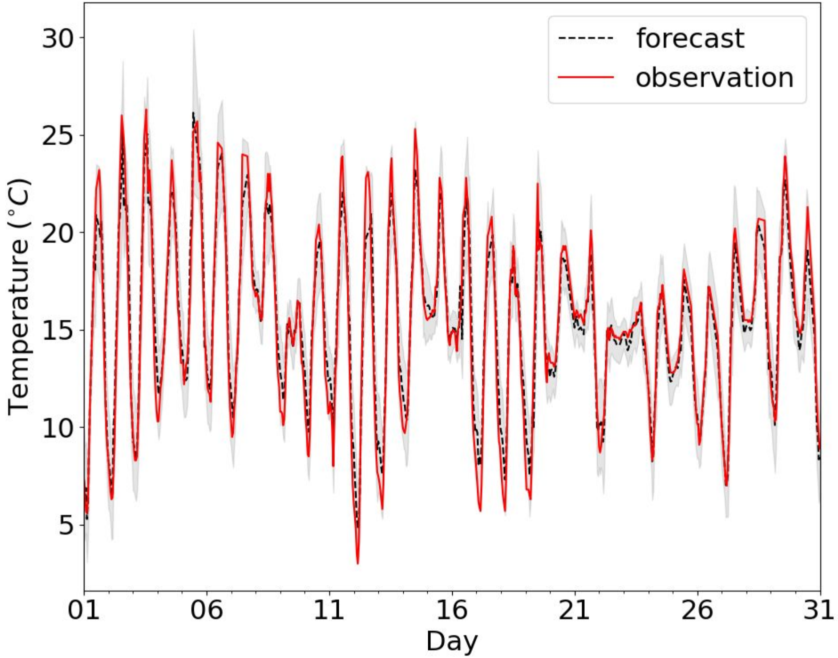


Timeseries forecast (whole month)

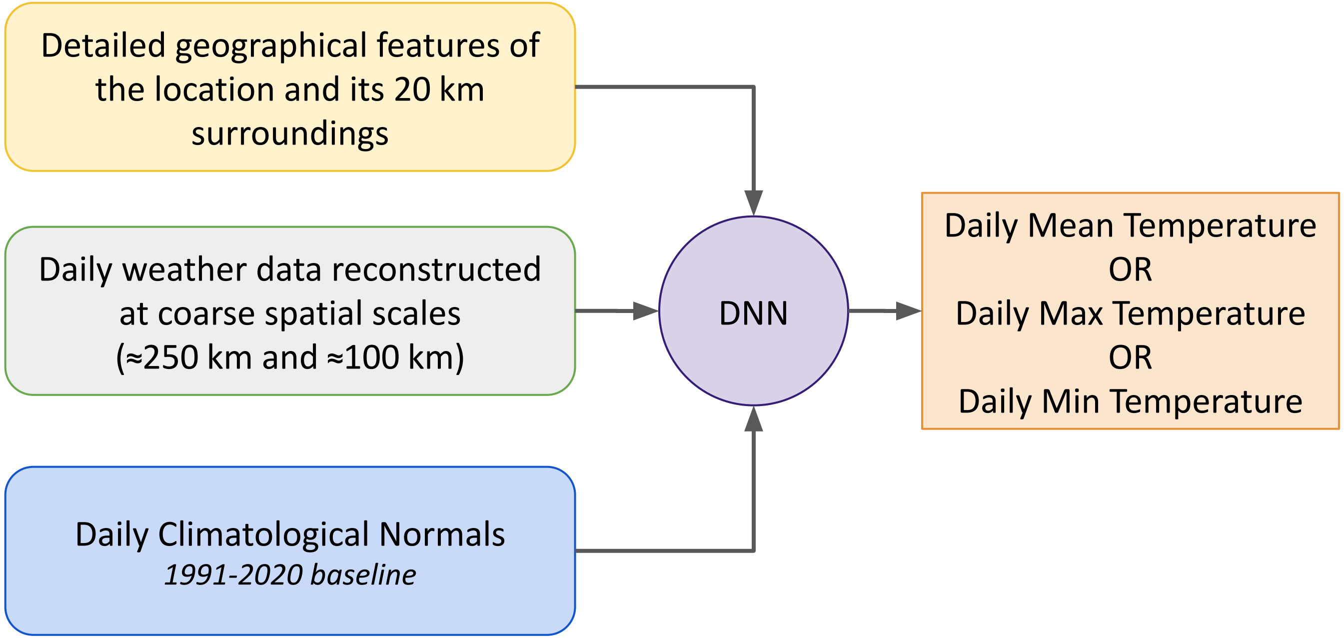
January (Orsta, Norway West Coast)



August (Roverud, North East of Oslo)



Daily temperatures from geographical, weather, and climate variables



Can a neural network learn the relationship between geography, large-scale weather conditions, and climatology to infer small-scale temperature patterns?

Daily temperatures from geographical, weather, and climate variables

Detailed geographical features of the location and its surroundings
[13 parameters]

Digital elevation model (DEM)

Elevation-based statistics (minimum, 10th, 25th, 50th, 75th, 90th percentiles, and maximum) within a 20 km radius

Land area fraction

Mean and standard deviation of land area fraction within a 20 km circle

Percentage of land occurrence within a 20 km radius

Distance to the coastline

Daily weather data reconstructed at coarse spatial scales (≈250 km and ≈100 km)
[8 parameters]

Optimal interpolation (OI) values of daily mean temperature, daily maximum temperature, daily minimum temperature, and precipitation at spatial scales of 250 km and 100 km.

Model inputs are OI(250 km) and the difference: OI (250 km) – OI (100 km)

Daily Climatological Normals
[4 parameters]

Long-term climatological means from the seNorge_2018 dataset for:

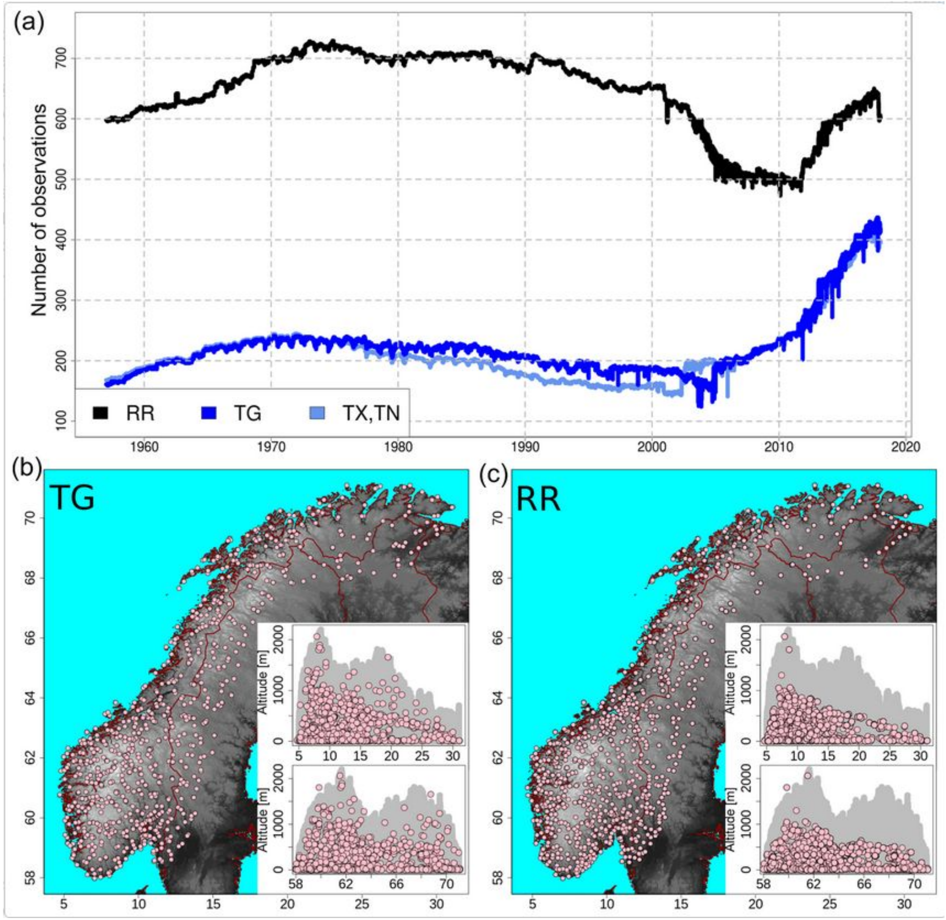
Daily mean temperature

Daily maximum temperature

Daily minimum temperature

Precipitation

Daily temperatures: Observational Network of in-situ data



we use only traditional weather stations either directly managed by MET or from other public institutions

Model is evaluated in Winter and Summer periods:

Summer 2023-06 / 2023-08 Hans extreme weather
number of samples: tx=4587 tg=7411 tn=4471

Winter 2020-12 / 2021-02 Gyda extreme weather
number of samples: tx=6833 tg=7806 tn=6825

Experiment 1: Training strategy

30 days before “D”	Day “D” Year “Y”	30 days after “D”

Cross-validation holdout: 10% of stations on day D reserved; their data excluded for the full 61 days.

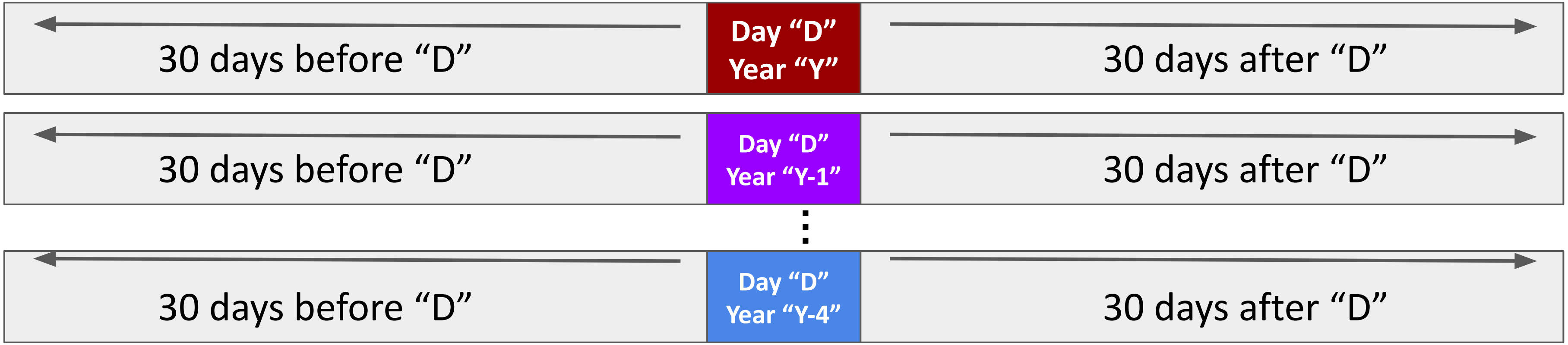
Remaining data split:

80% training

20% validation (to reduce overfitting).

Uncertainty estimate: Ensemble approach — train 10 times with random splits.

Experiment 2: Training strategy



Cross-validation holdout: 10% of stations on day D reserved; their data excluded for the full 61 days.

Remaining data split:

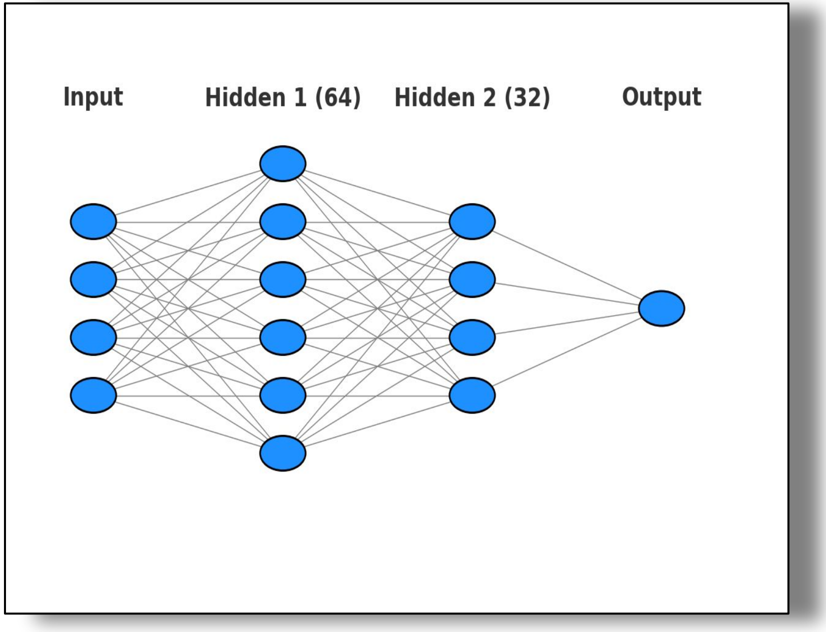
80% training

20% validation (to reduce overfitting).

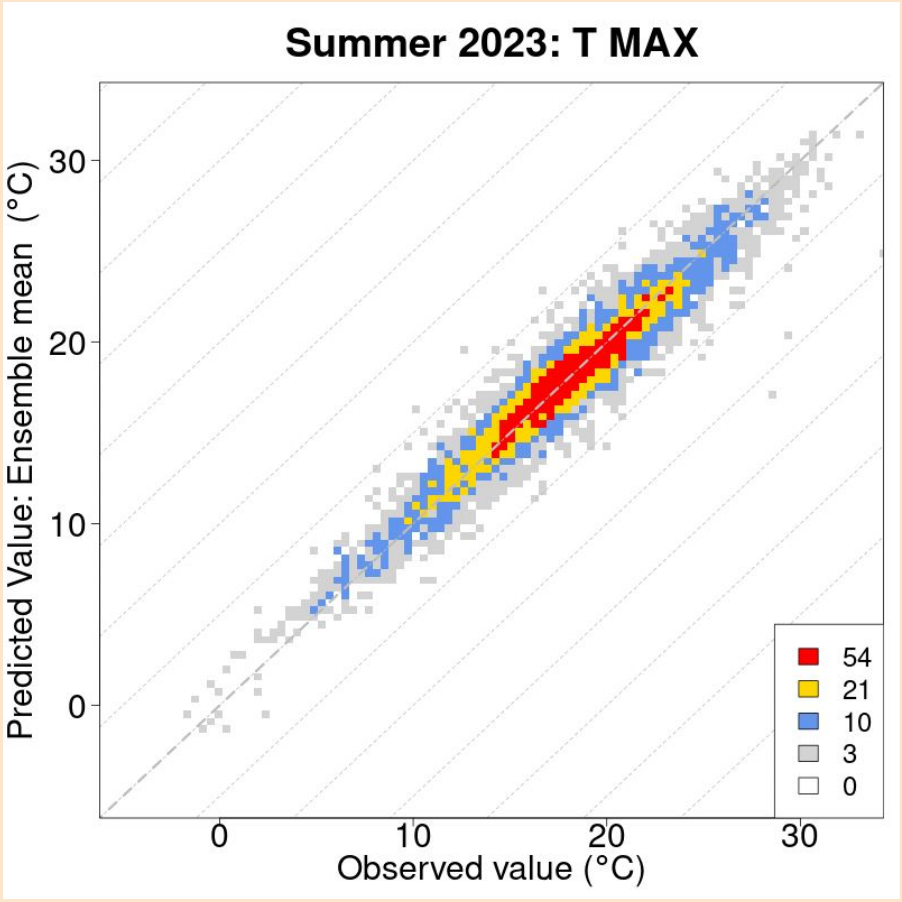
Uncertainty estimate: Ensemble approach — train 10 times with random splits.

Neural Network Configuration

- Architecture: 2 hidden layers (64 & 32 neurons)
- Model: Basic feed-forward neural network with BatchNorm
- Batch size: 2048
- Learning rate: 0.001
- Epochs: 200

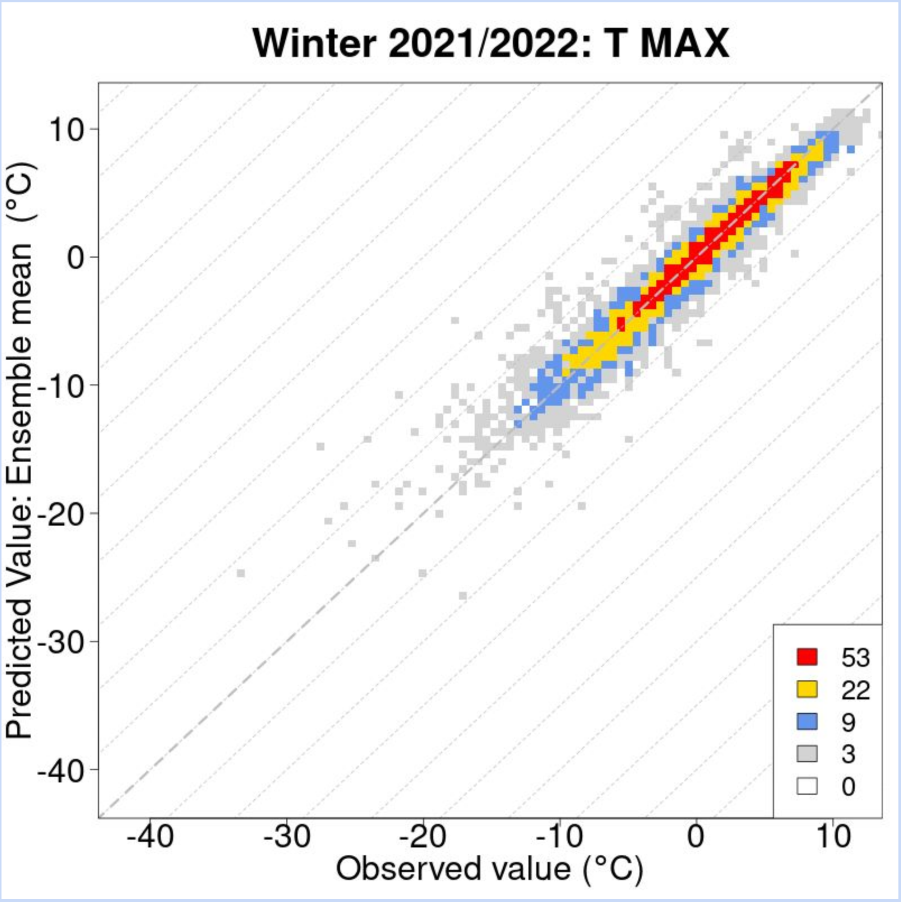


Daily Temperatures: Cross-Validation (EXP 1: 1 year, ±30 days)



Var	Bias (°C)	MAE (°C)	RMSE (°C)	% DEV > 3°C
T Max	0.0	1.0	1.3	3 %
T Mean	0.0	0.6	0.9	1 %
T Min	0.0	1.0	1.4	4 %

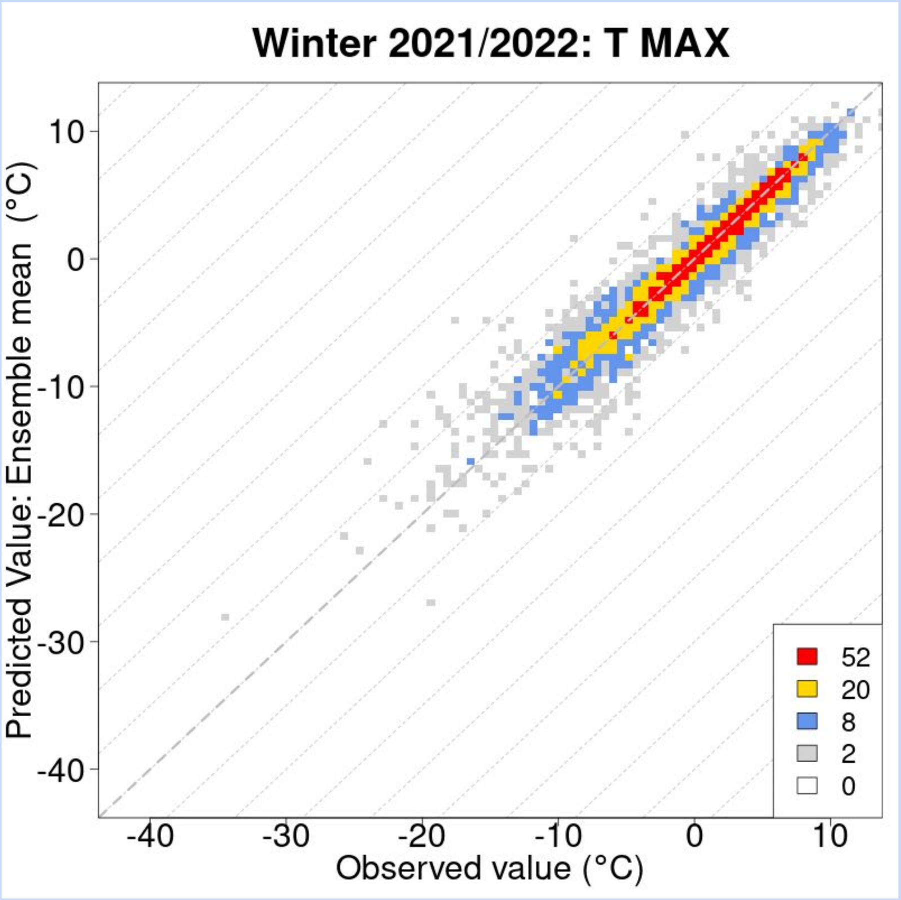
Daily Temperatures: Cross-Validation (EXP 1: 1 year, ±30 days)



Var	Bias (°C)	MAE (°C)	RMSE (°C)	% DEV > 3°C
T Max	0.0	1.0	1.3	3 %
T Mean	0.0	0.6	0.9	1 %
T Min	0.0	1.0	1.4	4 %

Var	Bias (°C)	MAE (°C)	RMSE (°C)	% DEV > 3°C
T Max	0.0	1.1	1.6	6 %
T Mean	-0.1	1.2	1.7	8 %
T Min	-0.1	1.8	2.5	17 %

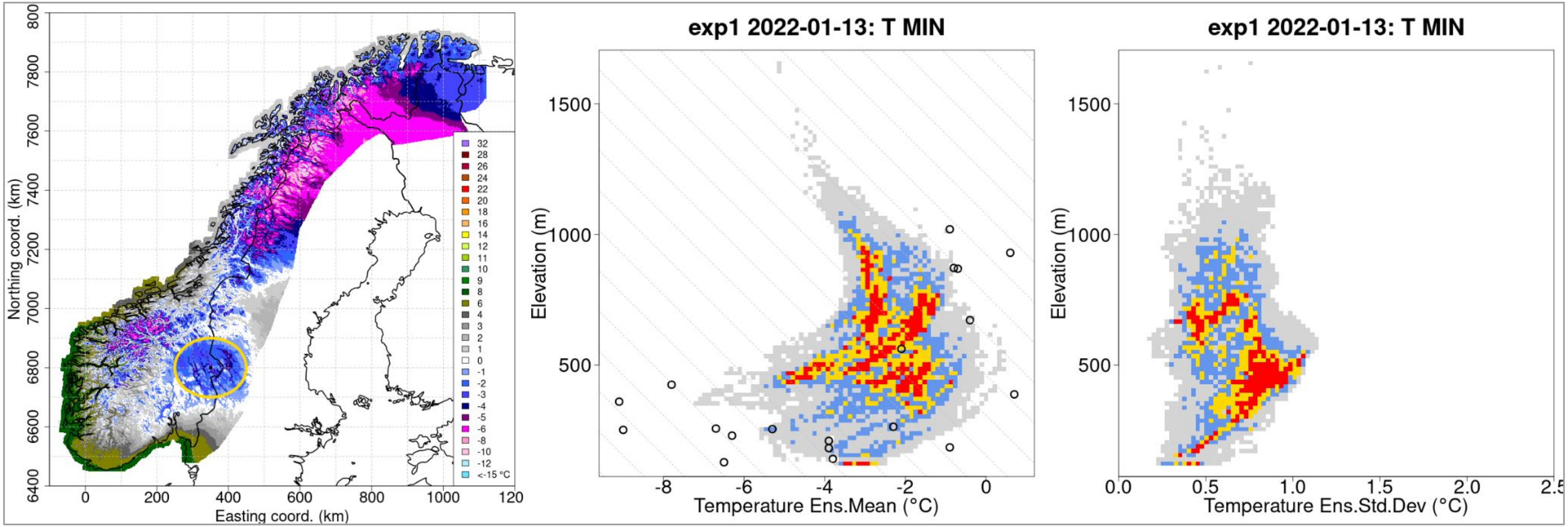
Daily Temperatures: Cross-Validation (EXP 2: 5 years, ±30 days)



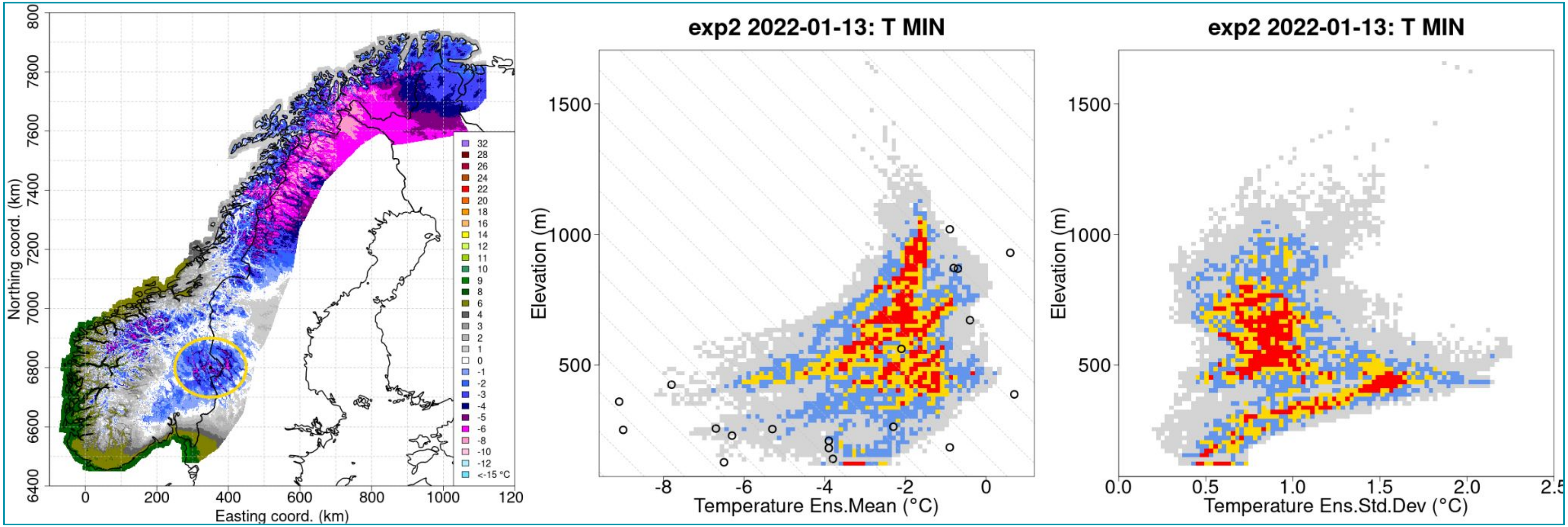
Var	Bias (°C)	MAE (°C)	RMSE (°C)	% DEV > 3°C
T Max	0.0	1.0	1.3	3 %
T Mean	0.0	0.6	0.9	1 %
T Min	0.0	1.0	1.4	5 %

Var	Bias (°C)	MAE (°C)	RMSE (°C)	% DEV > 3°C
T Max	0.0	1.2	1.7	7 %
T Mean	0.0	1.4	1.9	10 %
T Min	0.0	1.8	2.5	19 %

Daily Minimum Temperature: Gridded fields



Daily Minimum Temperature: Gridded fields



Conclusions

1. Hourly temperatures from the nearest observation stations
    a. The reconstruction quality of point observations is comparable to that obtained with traditional spatial analysis.
    b. Automatic DQC shows promising results, but further comparison with a baseline method is needed.
2. Daily temperatures predicted from geography, weather, and climate data
    a. Accuracy and precision match current operational methods
    b. One year of training gives decent performance; multi-year data better captures uncertainty
    c. Simple ensembling falls short in representing prediction uncertainty

Funded by the European Union

Coordinated by ECMWF

Weather Generator

Weathergenerator.eu

The WeatherGenerator project (grant agreement No101187947) is funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them.

WG Applications