

Verif – Comparison of atmospheric simulations to observations

Mikołaj Okrzesa – mikolaj@belgingur.is

Ólafur Rögnvaldsson – or@belgingur.is



Importance of simulation verification

- **Crucial in assessing accuracy and reliability of weather predictions and/or atmospheric re-analysis data**
- **Can help quantify model performance, identify problems and guide improvements**
- **Regular verification help monitoring model performance, independent of dynamic conditions**
- **Helps to identify problems with observations stations**

What is VERIF

Forecast verification software

release v1.3.0 Python package passing DOI 10.1175/BAMS-D-22-0253.1

Verif is a command-line tool that lets you verify the quality of weather forecasts for point locations. It can also compare forecasts from different forecasting systems (that have different models, post-processing methods, etc).

The program reads files with observations and forecasts in a specific format (see "Input files" below). The input files contain information about dates, forecast lead times, and locations such that statistics can be aggregated across different dimensions. To ensure a fair comparison among files, Verif will discard data points where one or more forecast systems have missing forecasts. Since Verif is a command-line tool, it can be used in scripts to automatically create verification figures.

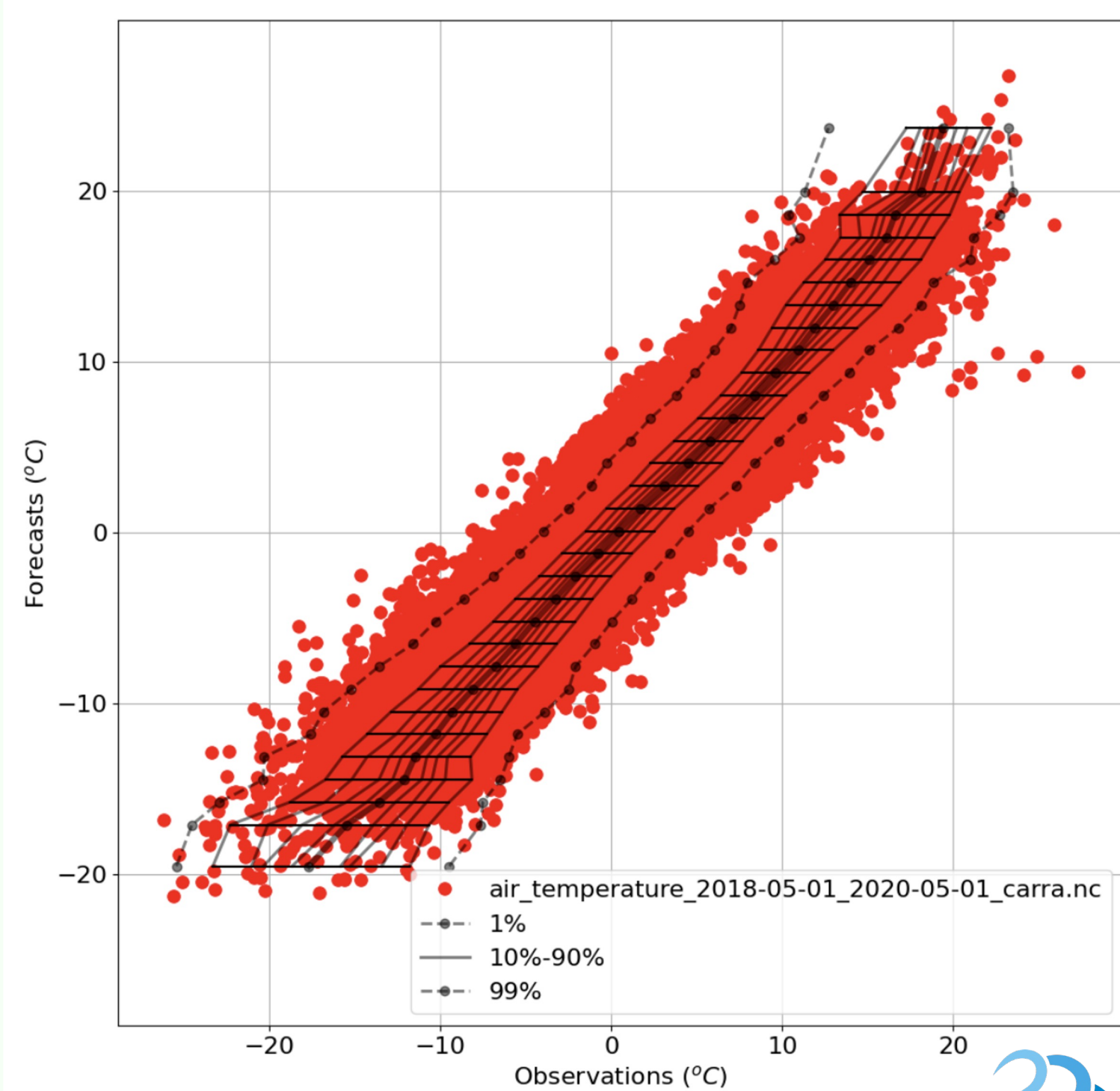
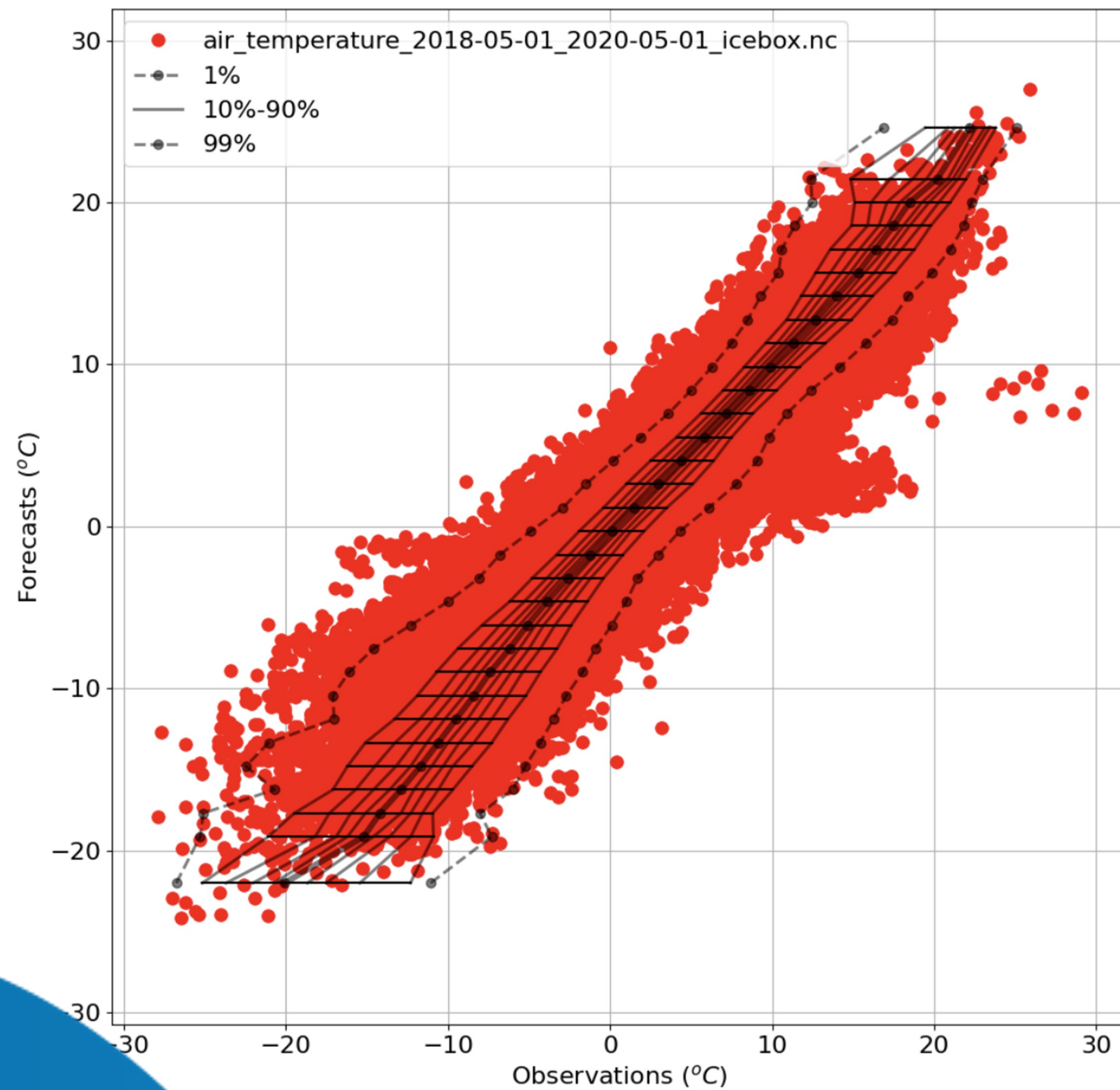
Verif version 1.3 has been released (see "Installation Instruction" below). We welcome suggestions for improvements. Verif is developed by Thomas Nipen (thomasn@met.no), with contributions from many.

Features

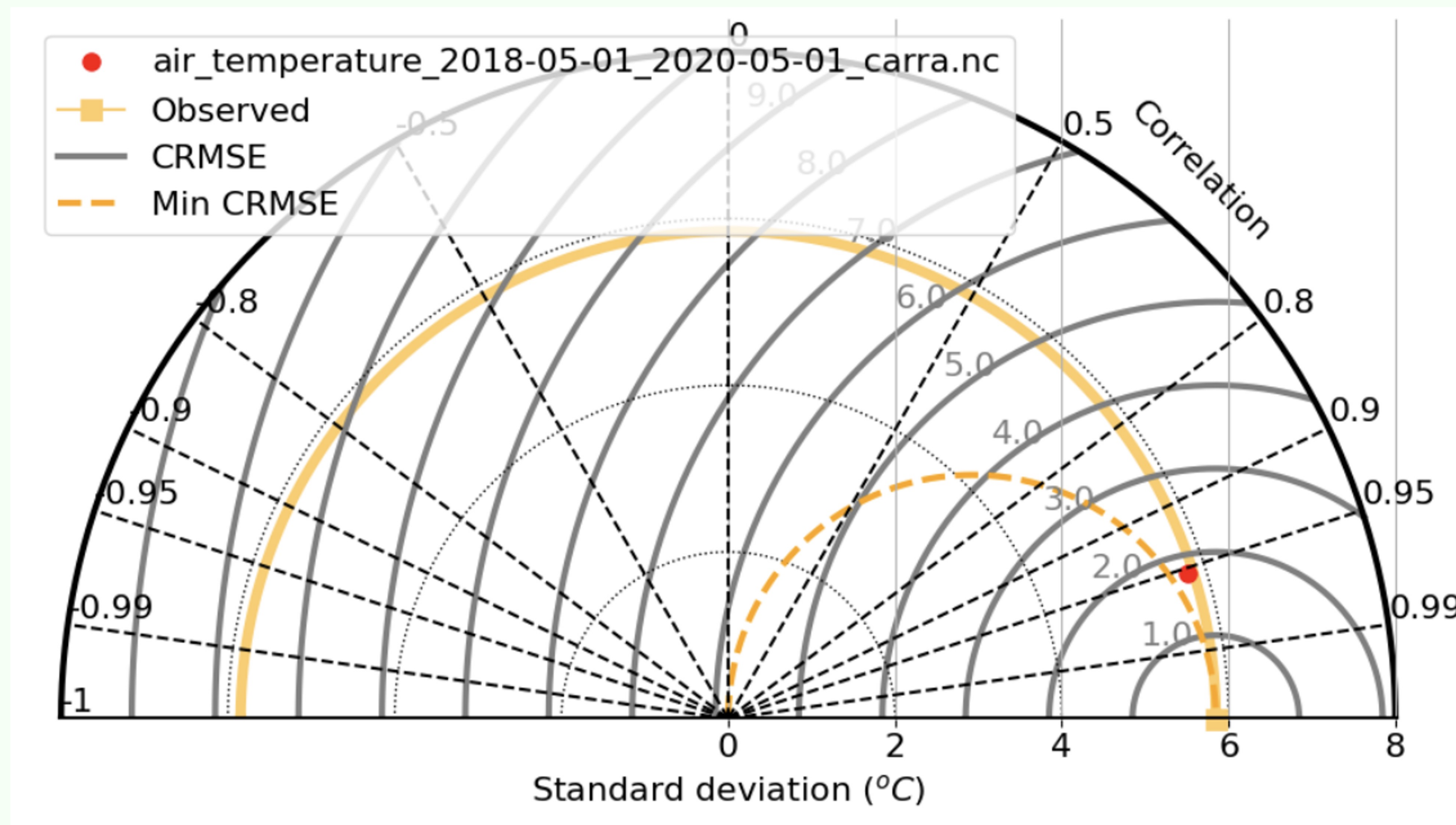
- Deterministic metrics such as MAE, bias, correlation, RMSE (e.g. `-m mae`)
- Threshold-based metrics such as the false alarm rate, ETS, EDI, Yule's Q (e.g. `-m ets`)
- Probabilistic metrics such as brier score, PIT-histogram, reliability diagrams (e.g. `-m bs`)
- Special plots like Taylor diagrams (`-m taylor`), quantile-quantile plots (`-m qq`).
- Plot scores as a function of date, lead time, station altitude/lat/longitude (e.g. `-x date`)
- Show scores on maps (`-type map`)
- Subset the data by specifying a date range and lat/lon range (`-latrange 58,60`)
- Export to text (`-type text`)
- Options to adjust font sizes, label positions, tick marks, legends, etc (`-labfs 14`)
- Anomaly statistics relative to a baseline like climatology (`-c climfile.txt`)
- Output to png, jpeg, eps, etc and specify dimensions and resolution (`-f image.png -dpi 300`)

- Can be used to compute and visualize verification scores
- Comparison of a single or multiple model output data
- Open-source, Python-based, command-line tool
- Offers 80+ metrics including deterministic, threshold-based and probabilistic metrics
- Examples: Mean Absolute Error, Equitable Threat Score, Brier Score, Taylor Diagram

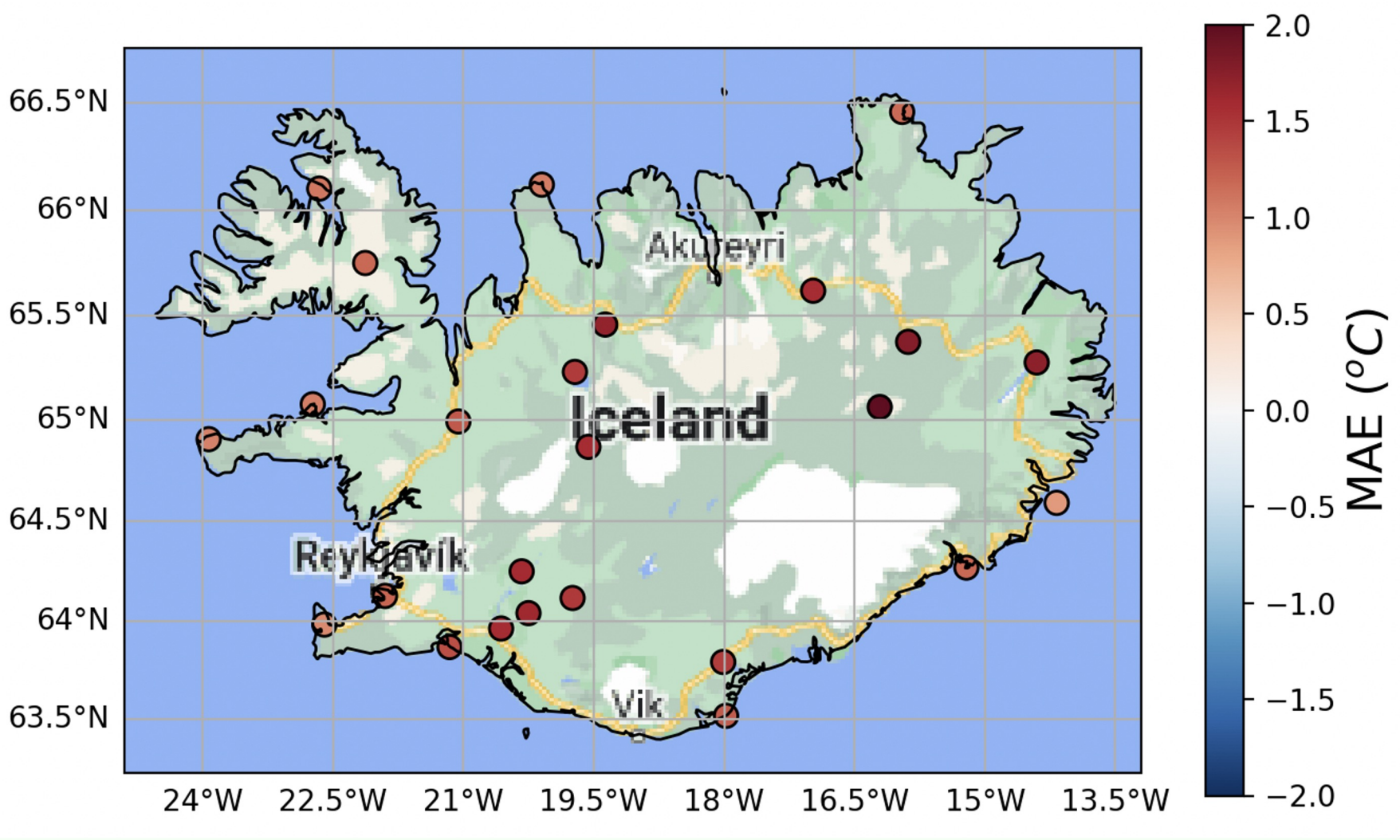
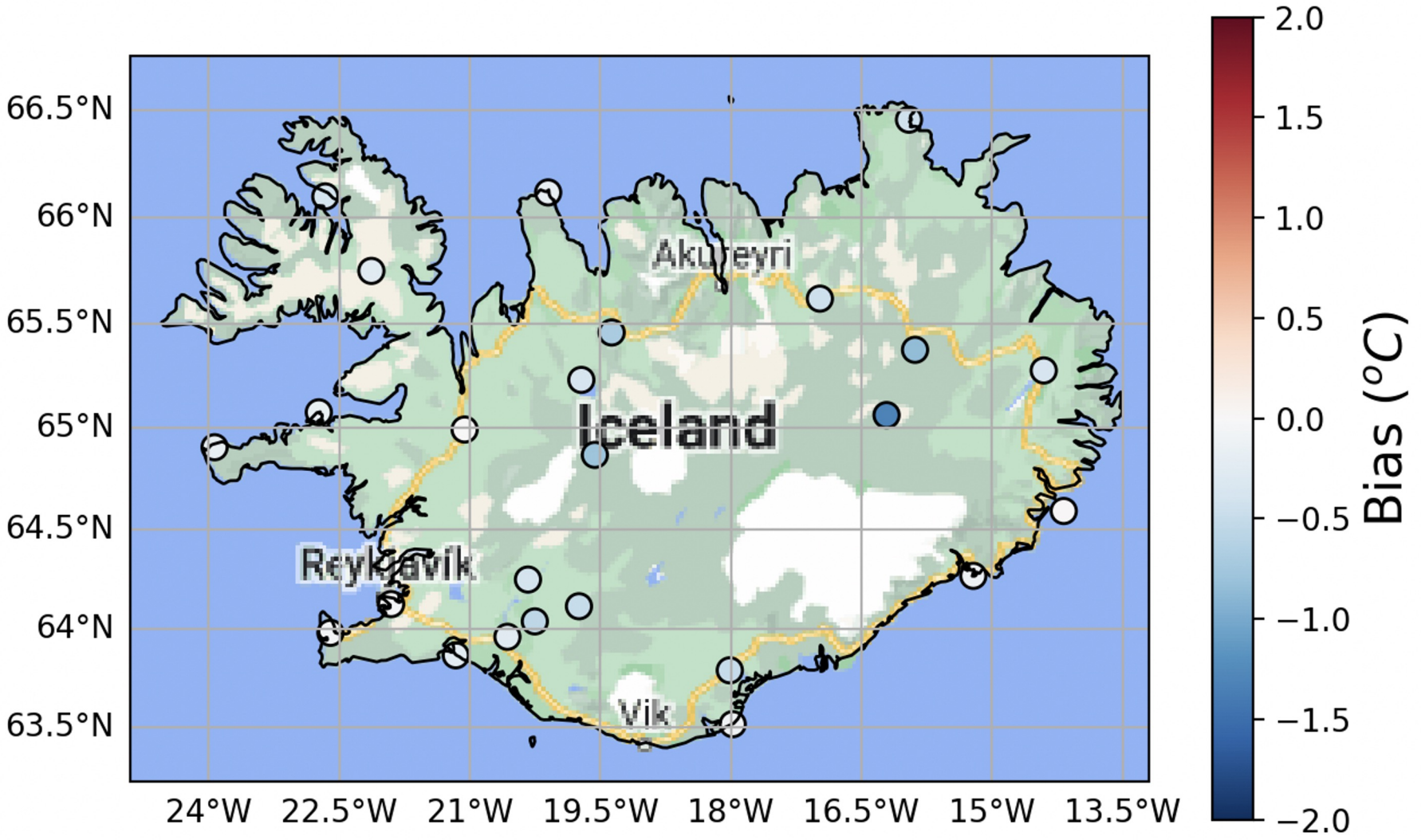
Plotting options – scatter plots



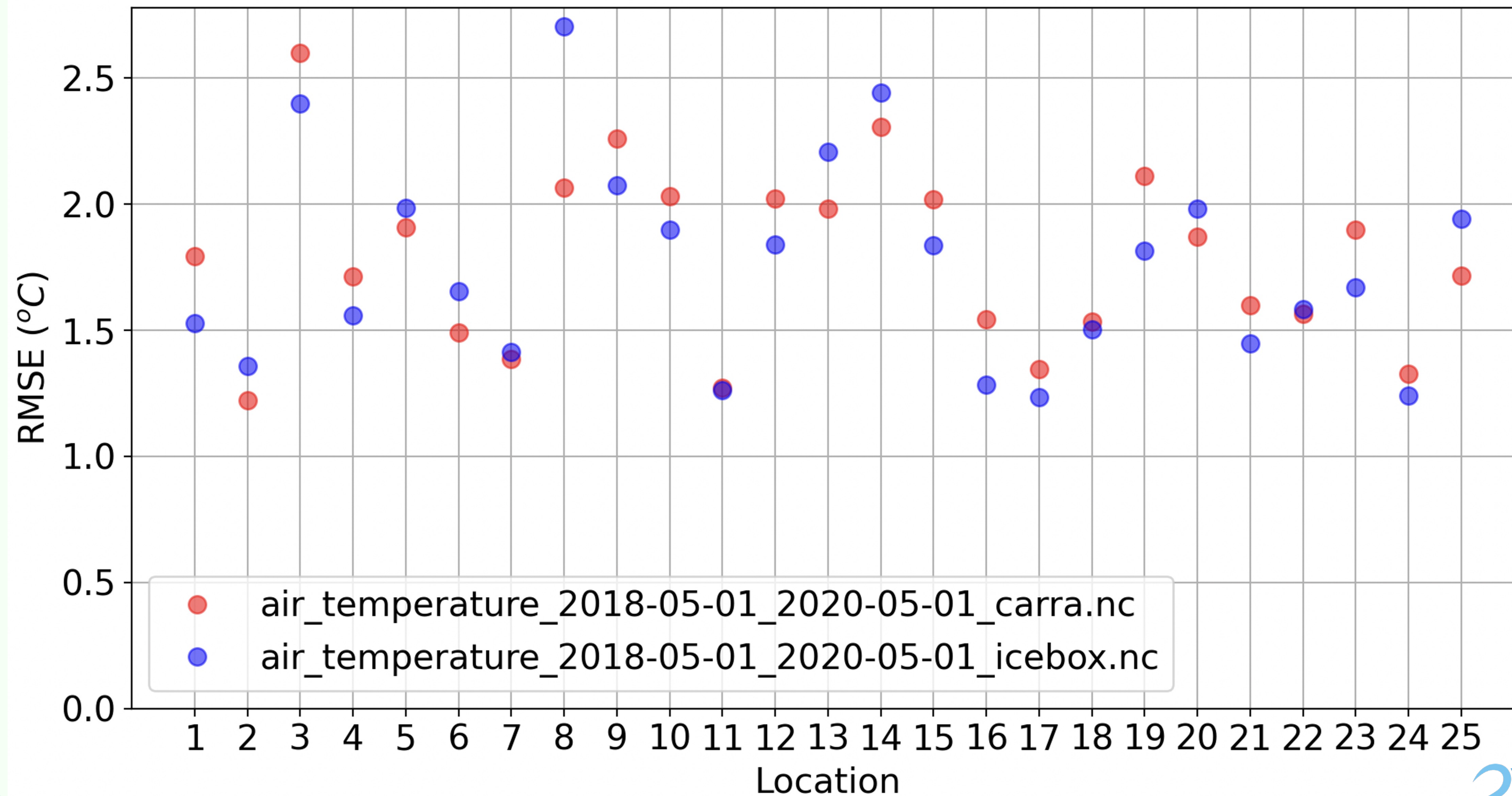
Plotting options - Taylor diagrams



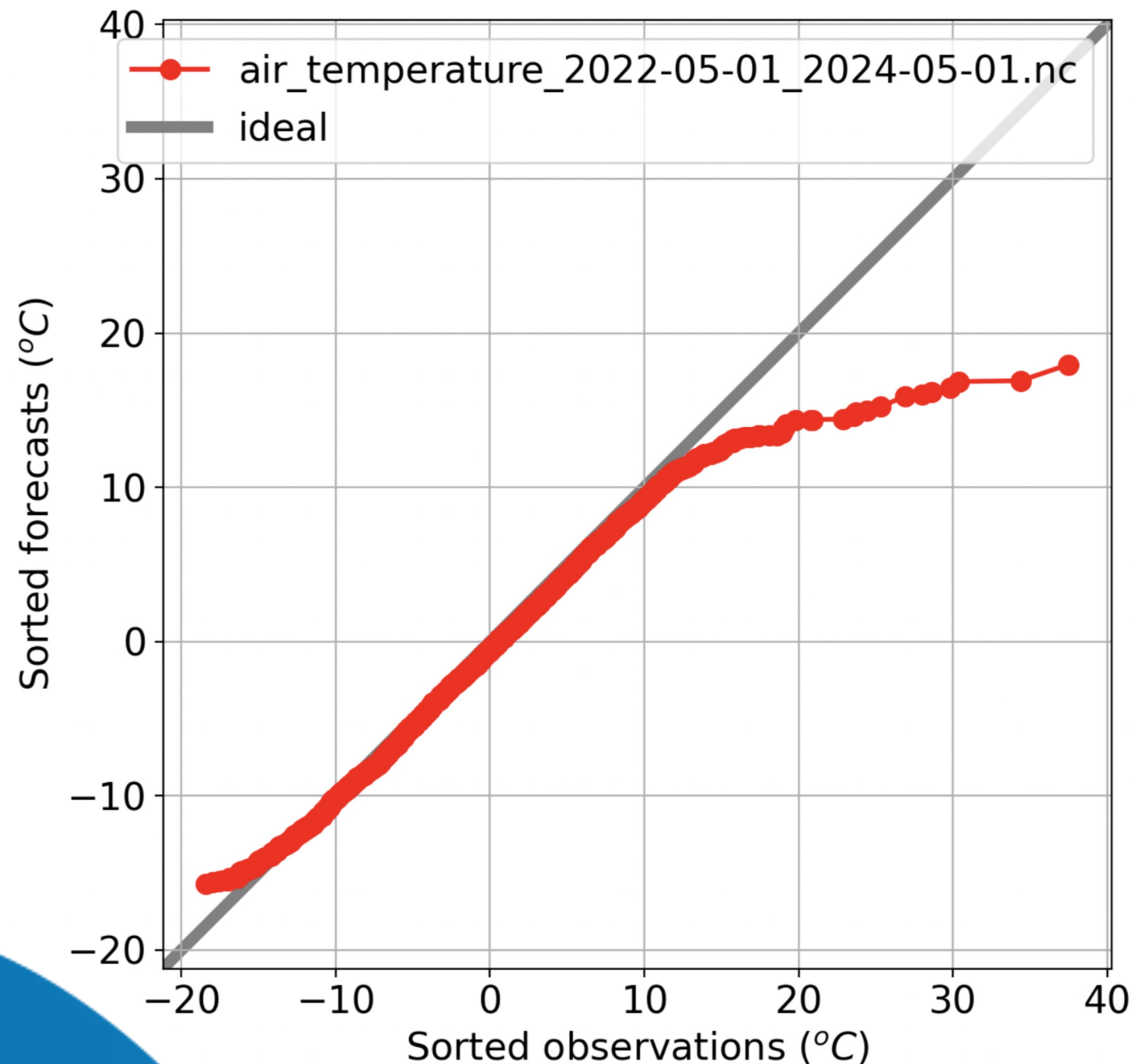
Plotting options - maps (MAE, BIAS)



Plotting options - dividing by location

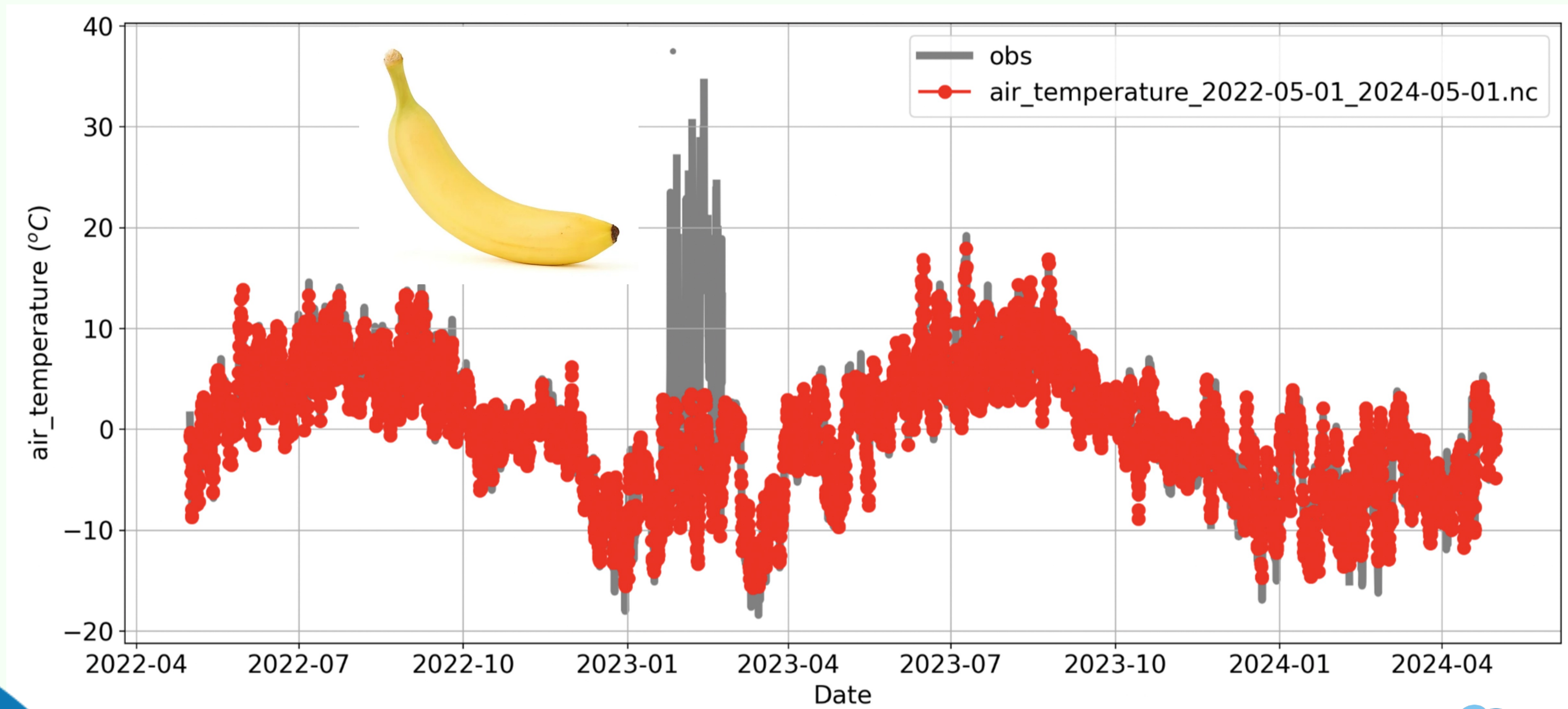


Finding faulty station



- **Quantile-Quantile Plot (QQ) compares the distribution of observations and simulations independently**
- **It can help identify with just a quick look if both fit ideal line**
- **In this case we can see a significant deviation from the ideal line for one of the stations**
- **On the next slide - time series plot showing when the observations go bananas**

Finding faulty data (and bananas)



Summary

- **Introduction to forecast verification**
- **Demonstration of a number of use-case: QQ plots, maps, scatter plots for observations-simulation/forecast data**
- **Example of how we managed to find a faulty station using Verif**
- **Using forecast verification can enhance the forecast quality and help us pin-point problems with observations and/or models alike**