

meteoblue continuously monitors forecast quality through rigorous verification, Comparing our forecast to data from over 100,000 global weather stations to ensure highest forecast accuracy. Our routine verification is the most comprehensive global study available. We choose proven data over expensive marketing campaigns.

Ground Truth Verification

Actual weather stations serve as our ground truth. Our independent baseline spans all national weather services. We also include over 100,000 partner network stations. Strict quality control filters out unreliable meteorological measurements. Only 60% of stations meet our standards. This rigorous screening ensures a reliable ground truth dataset.

Our benchmarking study evaluates every single continent. We verify meteoblue forecasts against major global models:

- **IFS** (ECMWF)
- **AIFS** (ECMWF)
- **GFS** (NOAA)
- **NEMSGLOBAL** (NOAA/meteoblue)
- **ICON** (DWD)
- **ARPEGE** (Météo-France)
- **UKMO** (UK Met Office)
- **GEM** (Environment Canada)

Precision Hourly Methodology

We compute all statistical scores at an hourly resolution. Predicting a precise temperature at 15:00 UTC is difficult. You must get the exact timing right. Daily maximums are much easier to predict. As a Swiss company, precision timing is our standard.

We calculate these daily scores using hourly data:

Mean absolute Error (MAE) tracks the average error magnitude.

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|$$

Root mean square Error (RMSE) similar to MAE but emphasizes larger forecasting errors.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}$$

- **Squaring mechanism:** Magnifies large discrepancies.
- **Same units:** Matches the original data units.
- **Variance indicator:** Highly sensitive to occasional outliers

Bias or Mean Error (ME) reveals systematic directional errors

$$ME = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)$$

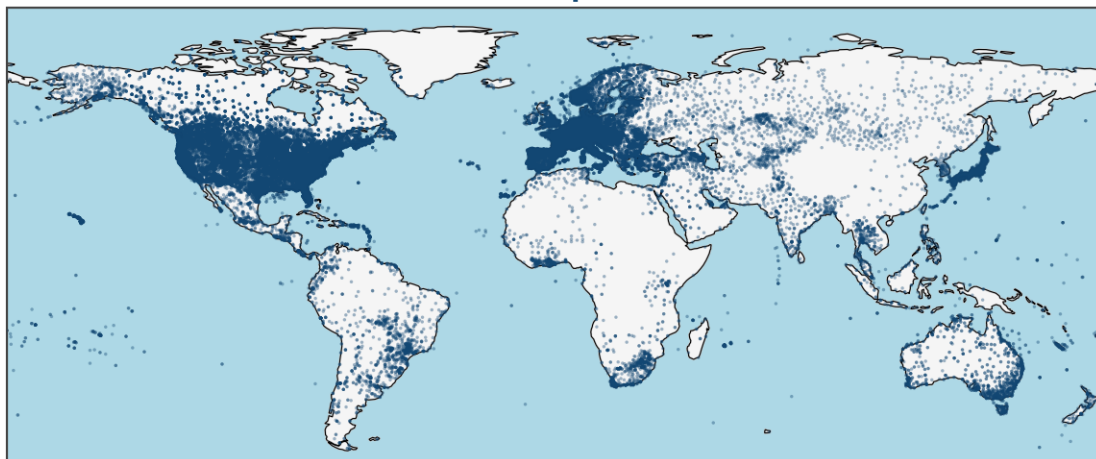
- $\hat{y}_i$: Predicted forecast value.
- y_i : Actual observed value.
- n : Total data points.
- **Positive value:** Systematic overestimation.
- **Negative value:** Systematic underestimation.
- **Cancellation effect:** Positive and negative errors offset.

Global Station Distribution

Global weather station density is not uniform.

- **Europe & USA:** Over 20,000 verified stations each.
 - **Africa:** Approximately 1,000 verified stations total.
- Consequently, verification data is much more robust in high-density regions. We share these infrastructure differences openly to provide the most transparent accuracy assessments possible.

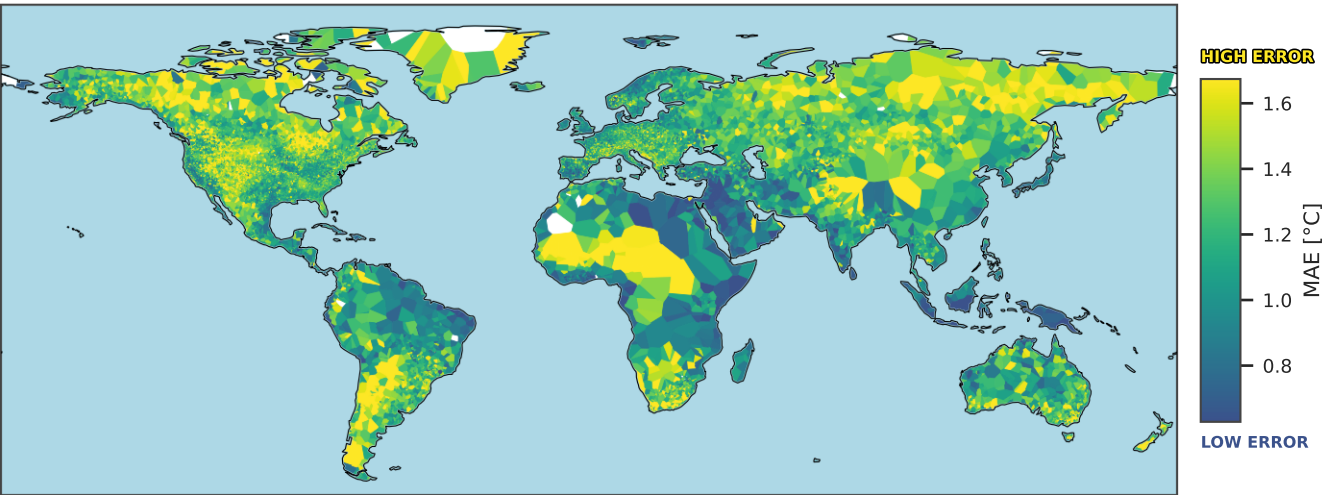
78,561 stations and 46,199,969 data points used for verification



Continent	Count
North America	44,047
South America	1,338
Europe	22,286
Africa	1,350
Asia	3,695
Australia	1,529

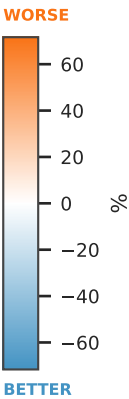
Forecast accuracy varies by location. The map provides a localized performance overview. This report aggregates data at the continental level. It evaluates every individual weather station. We cannot list every country or smaller administrative zone here. Please contact us for specific regional insights. We can access verification data for any individual station worldwide in real time.

Global MAE of meteoblue mLM - Temperature



MAE relative to best global raw model

	Europe	Asia	Africa	North America	South America	Australia
meteoblue mLM	-32.0%	-32.9%	-33.7%	-32.1%	-28.3%	-23.8%
model1	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
model2	+27.7%	+54.0%	+36.4%	+24.8%	+32.2%	+26.4%
model3	+43.6%	+71.6%	+41.5%	+42.5%	+57.0%	+46.2%
model4	+3.4%	+10.3%	+5.6%	+1.1%	+3.1%	-0.5%
model5	+25.3%	+29.2%	+21.3%	+41.2%	+22.9%	+20.3%
model6	+12.6%	+17.3%	+23.5%	+13.2%	+12.1%	+12.8%
model7	+17.5%	+34.2%	+25.4%	+13.5%	+19.5%	+16.2%
model8	+8.3%	+9.9%	+2.1%	+5.2%	+4.4%	+16.0%



Forecast Day Analysis - Temperature

Global MAE by forecast lead time (day 1-7)

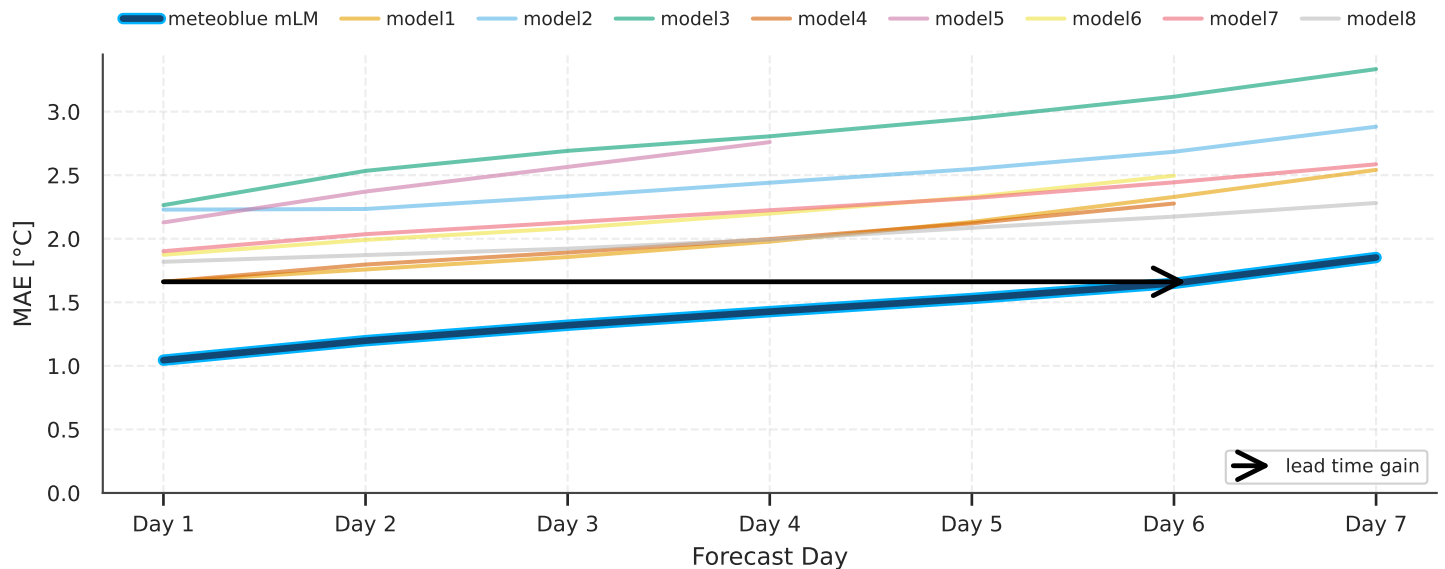
Forecasting error generally increases with lead time. The further into the future, the larger the expected error. Despite this, meteoblue mLM remains highly competitive: on day 6 it is still as accurate as the best raw model on day 1.

The general ranking of models over time stays quite stable, so for the remainder of this report we use forecast day 1 as the basis for all comparisons.

Global MAE by Forecast Day

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
meteoblue mLM	1.04	1.20	1.32	1.43	1.53	1.65	1.85
model1	1.66	1.76	1.86	1.98	2.13	2.33	2.54
model2	2.23	2.23	2.33	2.44	2.55	2.68	2.88
model3	2.26	2.53	2.69	2.81	2.95	3.12	3.33
model4	1.66	1.80	1.89	2.00	2.12	2.28	-
model5	2.13	2.37	2.57	2.76	-	-	-
model6	1.87	1.99	2.08	2.20	2.33	2.50	-
model7	1.90	2.04	2.13	2.22	2.32	2.44	2.59
model8	1.82	1.87	1.92	1.99	2.09	2.17	2.28

Global MAE by Forecast Day - Temperature



The black arrow on the chart above illustrates the lead-time gain: 6 days ahead, meteoblue mLM is still as accurate as the best raw model forecast is on day 1. The table below reports this gain for every model.

meteoblue mLM Lead-Time Advantage

Reference model	MAE on day 1	Lead time gained by meteoblue mLM
meteoblue mLM	1.04	-
model4	1.66	+5 days
model1	1.66	+5 days
model8	1.82	+5 days
model6	1.87	+6 days
model7	1.90	+6 days
model5	2.13	+6 days
model2	2.23	+6 days
model3	2.26	+6 days

Global Summary - Temperature

Monthly scores by continent, daily MAE timeseries, and model comparison

MAE

	Europe	Asia	Africa	North America	South America	Australia
meteoblue mLM	1.17	0.98	1.17	1.23	1.21	1.16
model1	1.72	1.46	1.77	1.81	1.69	1.52
model2	2.20	2.25	2.41	2.26	2.23	1.92
model3	2.47	2.51	2.50	2.58	2.65	2.22
model4	1.78	1.61	1.87	1.83	1.74	1.51
model5	2.16	1.89	2.15	2.55	2.08	1.83
model6	1.94	1.71	2.18	2.05	1.89	1.71
model7	2.02	1.96	2.22	2.05	2.02	1.76
model8	1.86	1.60	1.81	1.90	1.76	1.76

RMSE

	Europe	Asia	Africa	North America	South America	Australia
meteoblue mLM	1.59	1.35	1.61	1.68	1.66	1.53
model1	2.31	1.98	2.38	2.40	2.28	1.99
model2	2.88	2.94	3.14	2.99	2.99	2.50
model3	3.21	3.31	3.25	3.39	3.50	2.90
model4	2.37	2.14	2.42	2.42	2.33	1.98
model5	2.82	2.56	2.80	3.34	2.75	2.36
model6	2.56	2.33	2.86	2.69	2.58	2.25
model7	2.68	2.74	2.95	2.71	2.72	2.30
model8	2.44	2.22	2.41	2.50	2.36	2.25

ME

	Europe	Asia	Africa	North America	South America	Australia
meteoblue mLM	0.03	0.03	-0.02	0.11	0.08	-0.10
model1	-0.19	-0.38	-0.33	0.13	0.10	-0.05
model2	-0.51	-0.89	0.28	-0.03	0.04	0.06
model3	-0.16	-0.97	-0.11	0.61	0.52	0.42
model4	-0.29	-0.61	-0.27	0.27	-0.06	0.03
model5	0.70	0.16	0.08	1.42	0.31	-0.00
model6	-0.33	-0.44	-0.96	0.18	-0.65	-0.01
model7	-0.47	-1.10	-1.34	-0.24	-0.21	-0.47
model8	-0.21	-0.23	-0.36	0.12	-0.07	-0.12

Understanding Forecast Variability

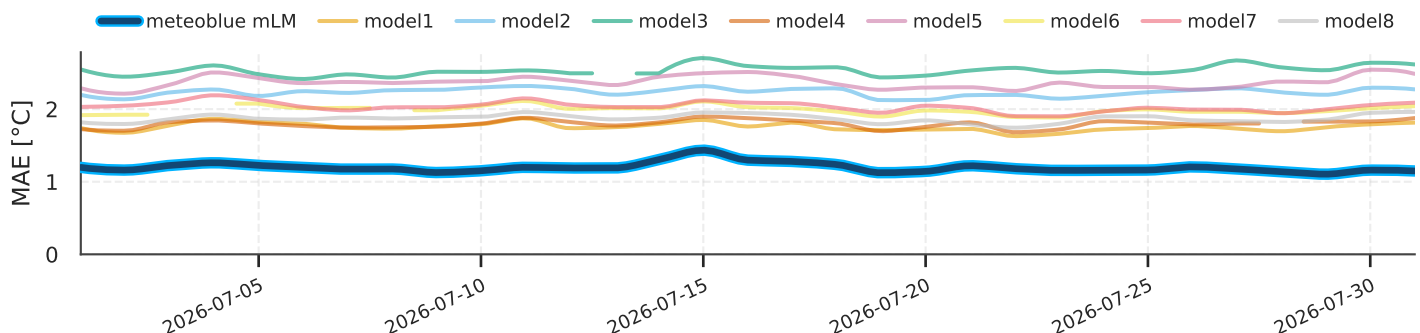
Monthly summaries can create the impression that one weather model is always superior. In reality, a model that performs perfectly today may fluctuate tomorrow as weather situations change. These graphs illustrate how forecast errors shift over time. Because these figures represent continental averages, variability at individual station levels is much higher.

Managing Risk and Uncertainty

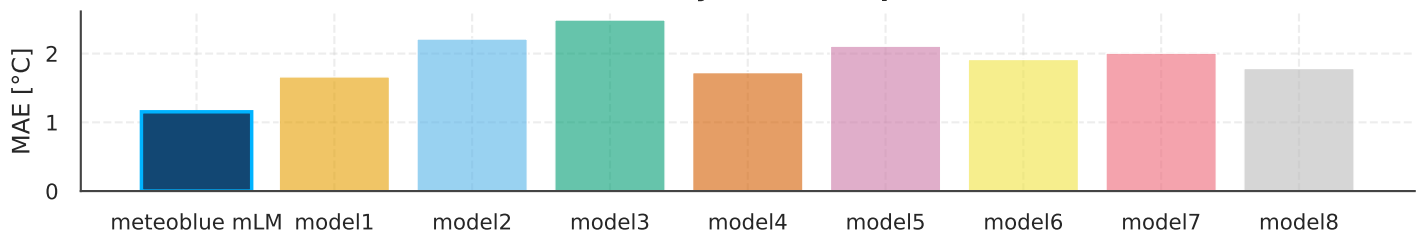
Weather forecasting is highly complex. The meteoblue Learning MultiModel (mLM) achieves the highest overall accuracy in these summaries by leading in most locations, yet no single model wins every time.

This variability is exactly why we provide dozens to hundreds of alternative forecasts from other institutions alongside our own best estimate. Having access to all available forecasts ensures you are covered even when a specific model underperforms. Depending on your business needs and technical expertise, evaluating this uncertainty is often critical for managing operational risk. Our experts are available to help you build a specially tailored solution.

Global daily MAE - Temperature



Global monthly MAE - Temperature

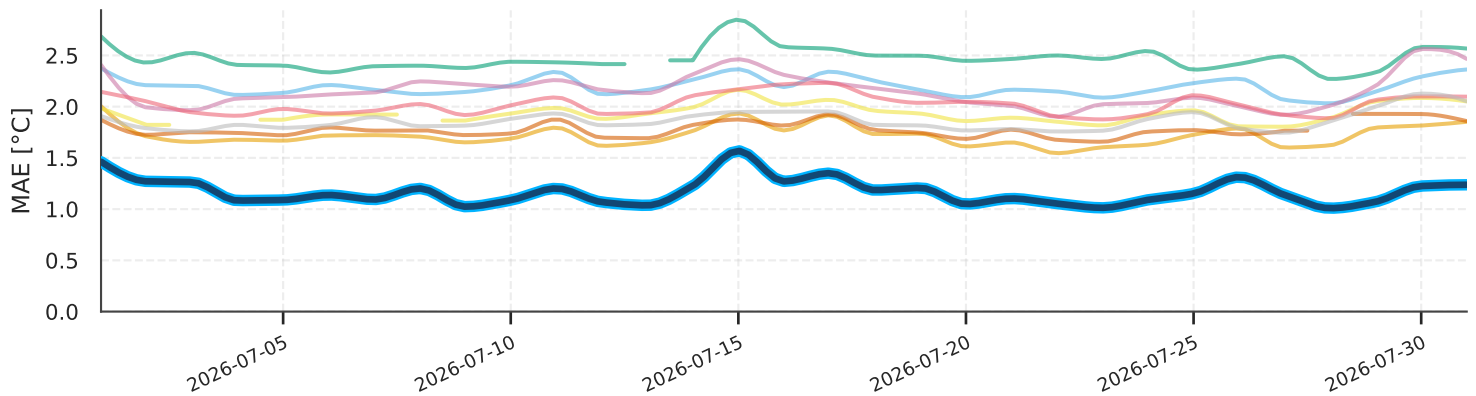


Europe - Temperature

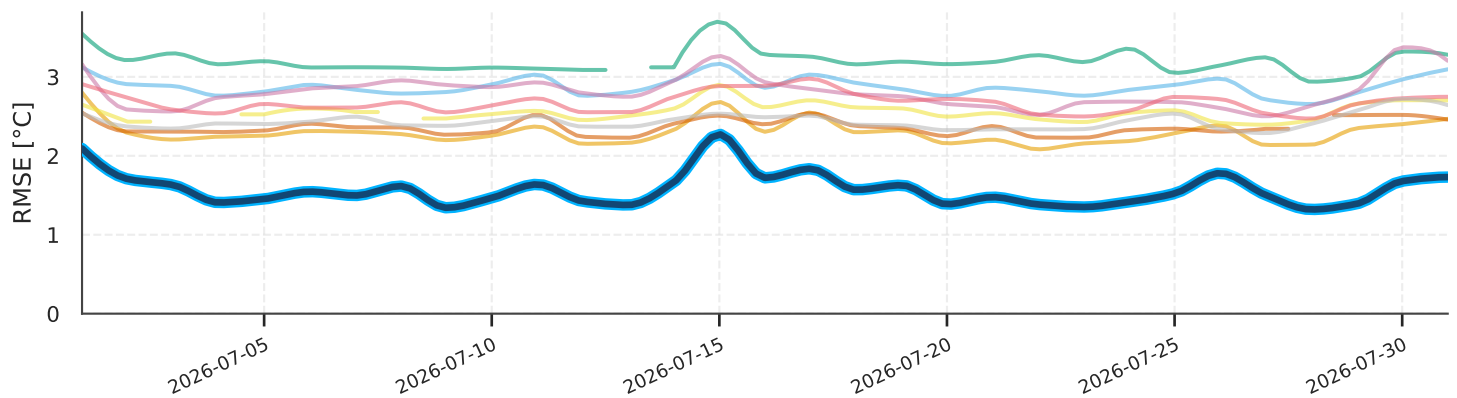
Daily verification scores and overall model ranking

meteoblue mLM model1 model2 model3 model4 model5 model6 model7 model8

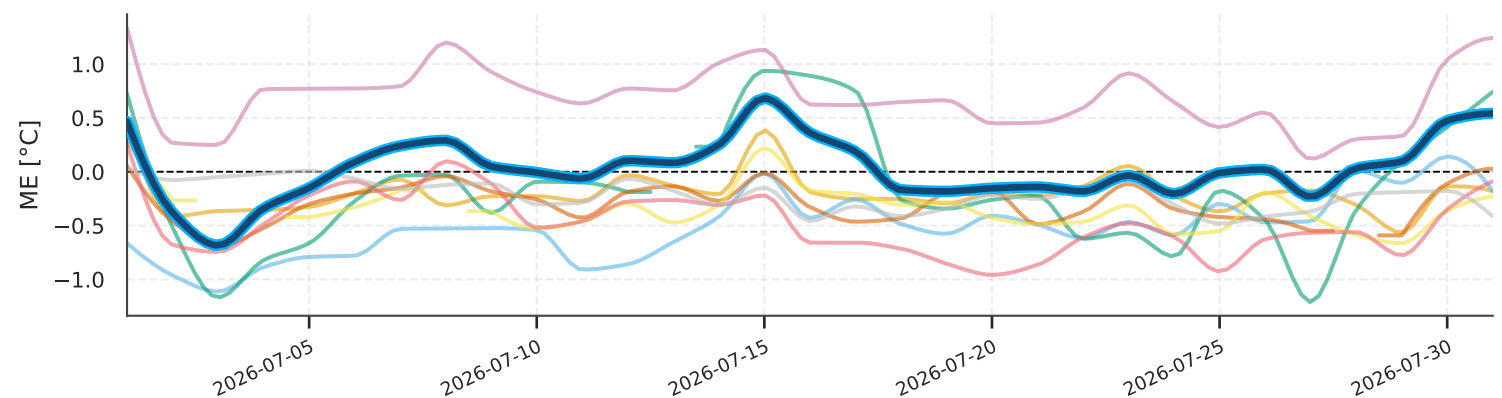
Daily MAE



Daily RMSE

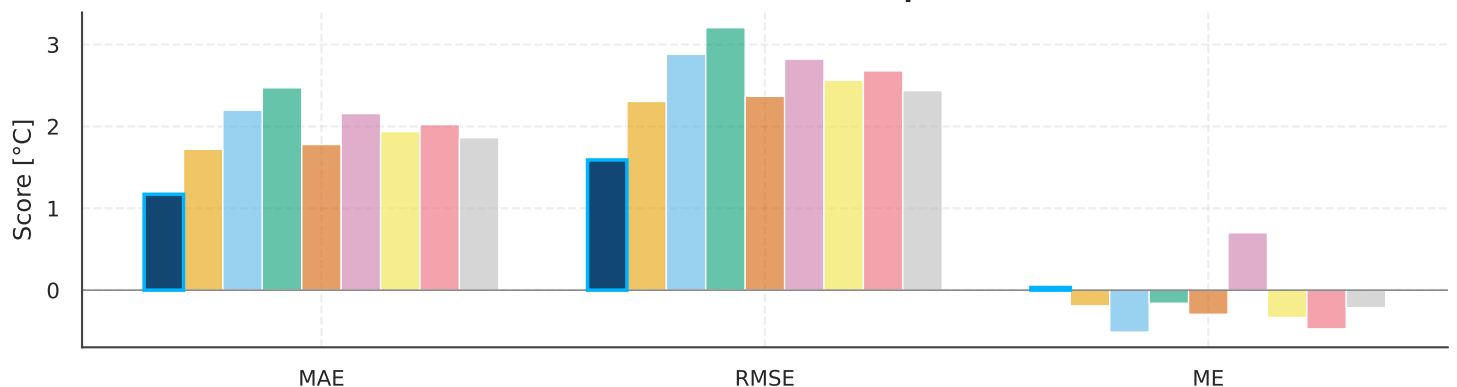


Daily ME



meteoblue mLM model1 model2 model3 model4 model5 model6 model7 model8

Overall scores - Europe

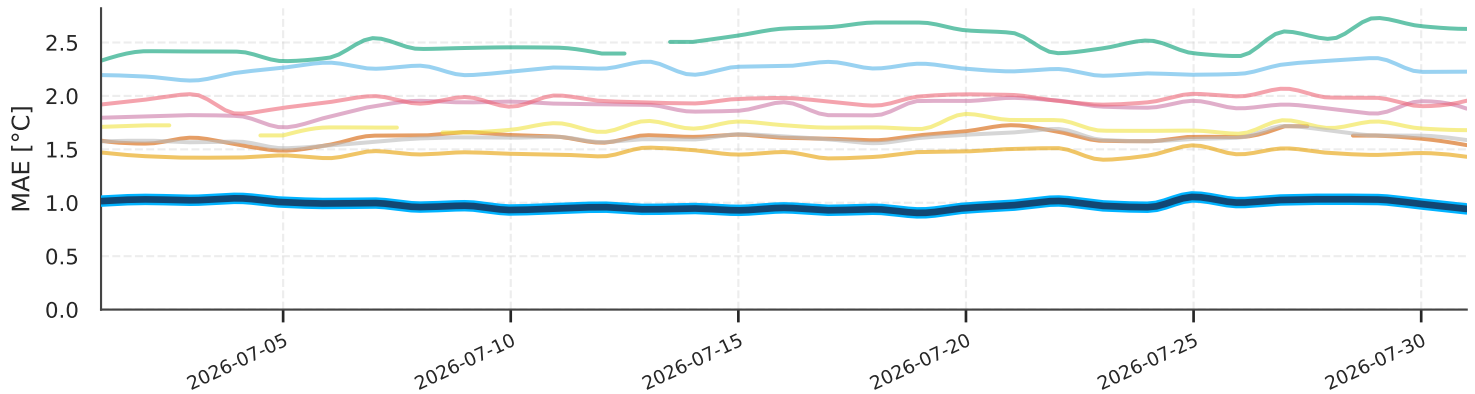


Asia - Temperature

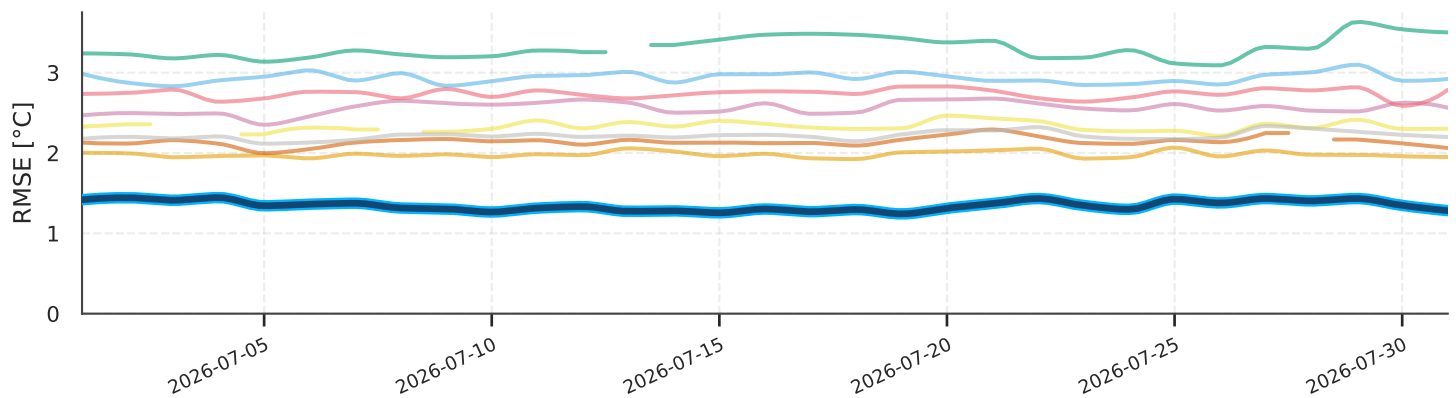
Daily verification scores and overall model ranking

meteoblue mLM model1 model2 model3 model4 model5 model6 model7 model8

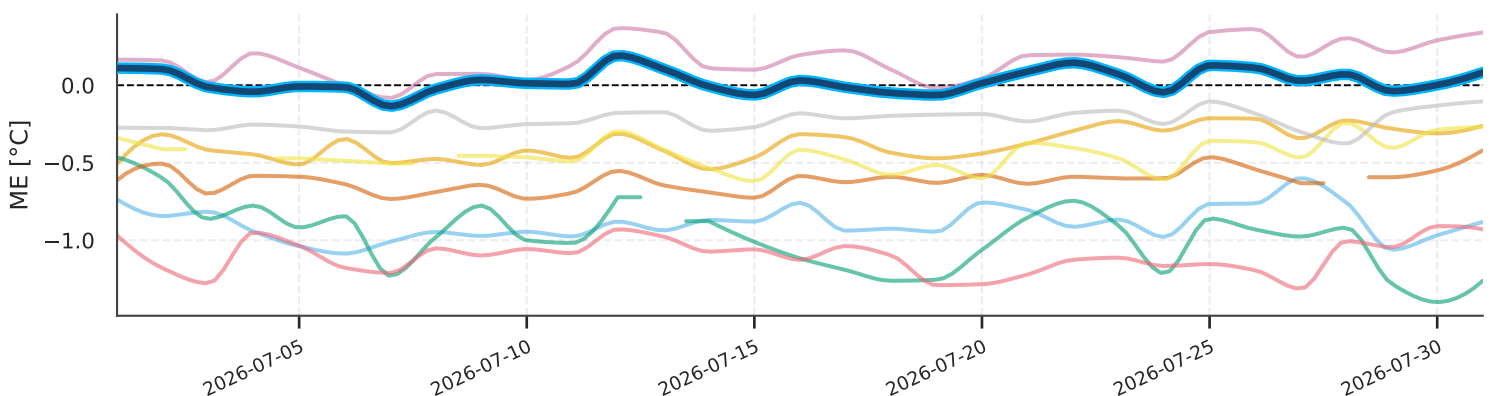
Daily MAE



Daily RMSE

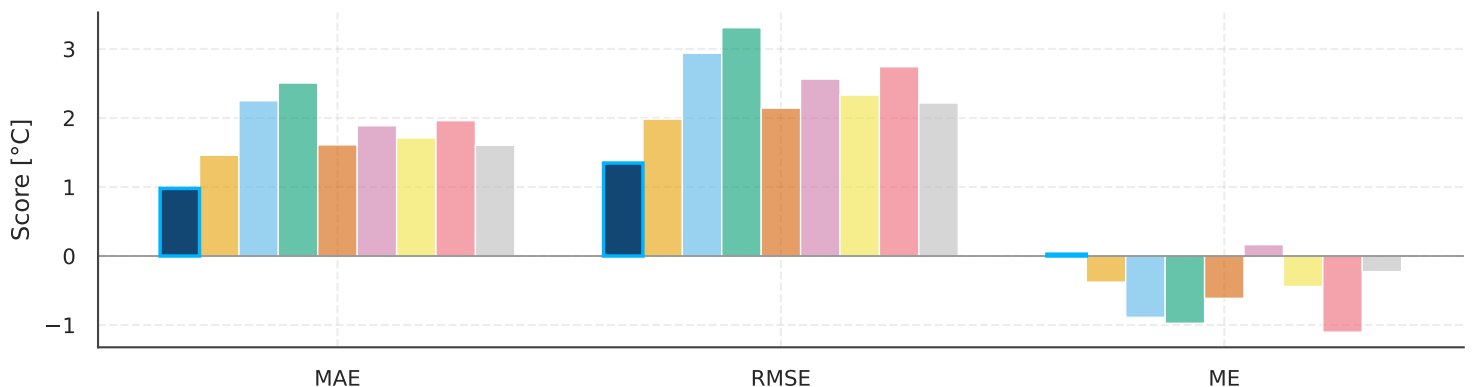


Daily ME



meteoblue mLM model1 model2 model3 model4 model5 model6 model7 model8

Overall scores - Asia

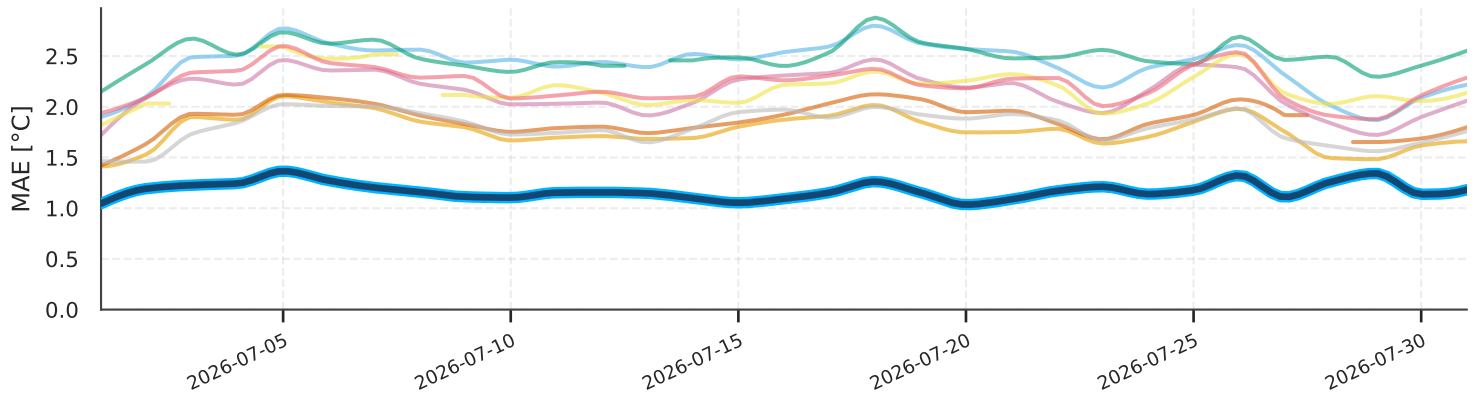


Africa - Temperature

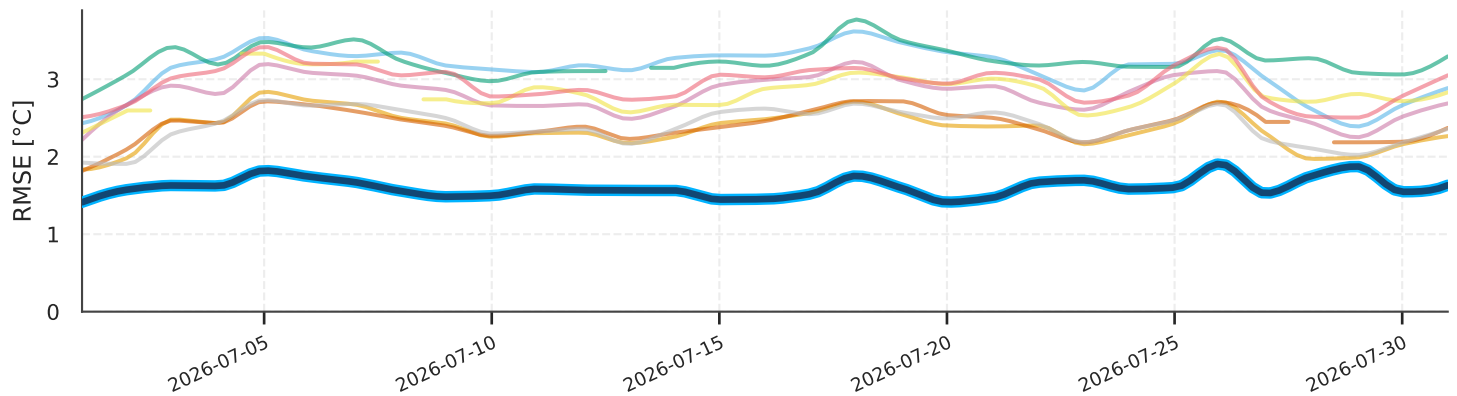
Daily verification scores and overall model ranking

meteoblue mLM model1 model2 model3 model4 model5 model6 model7 model8

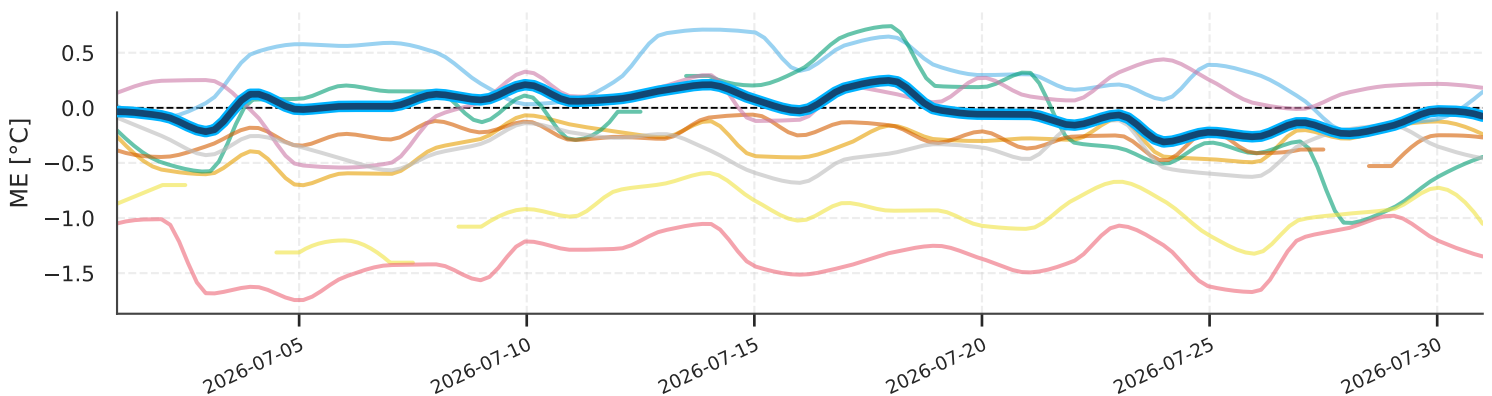
Daily MAE



Daily RMSE

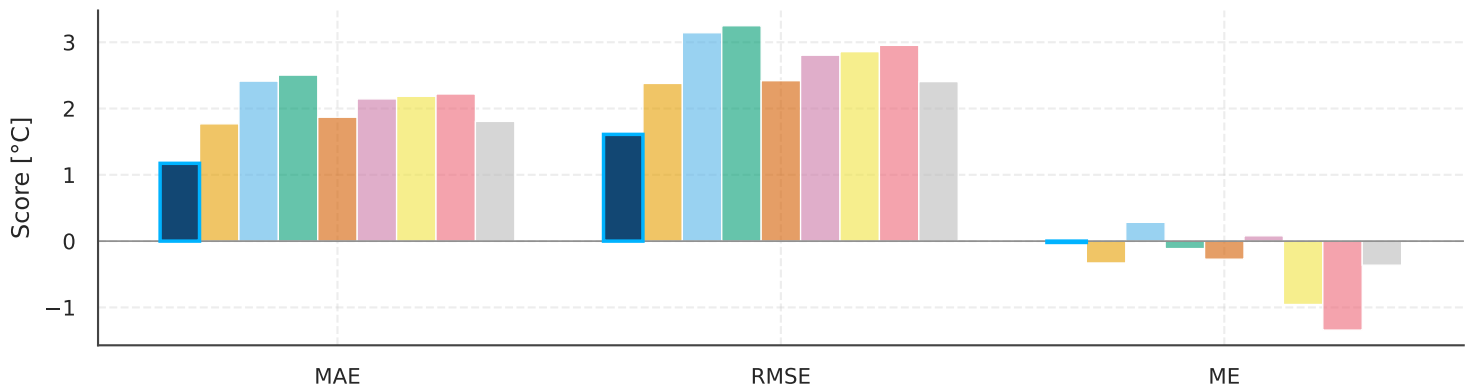


Daily ME



meteoblue mLM model1 model2 model3 model4 model5 model6 model7 model8

Overall scores - Africa

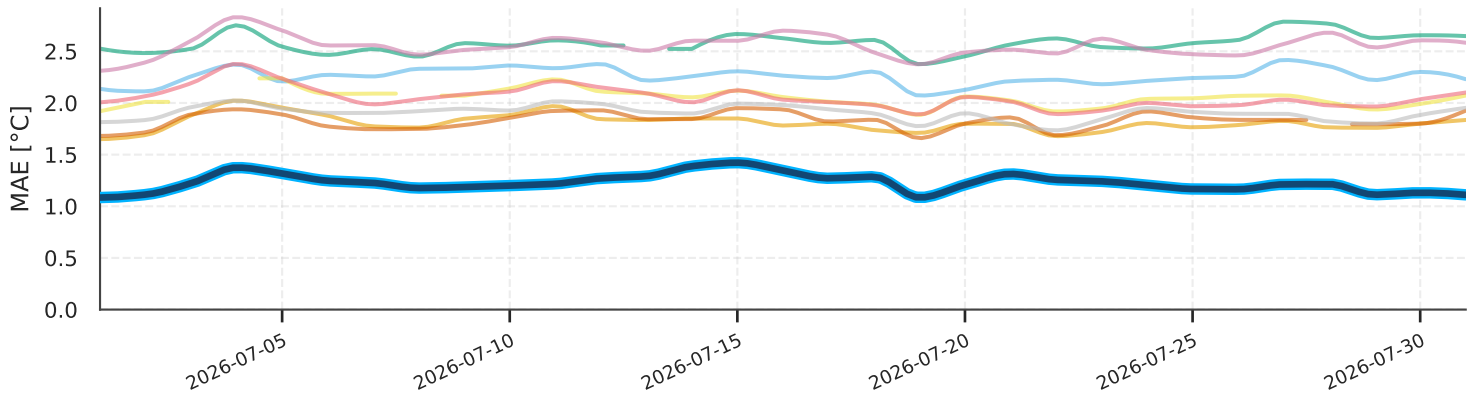


North America - Temperature

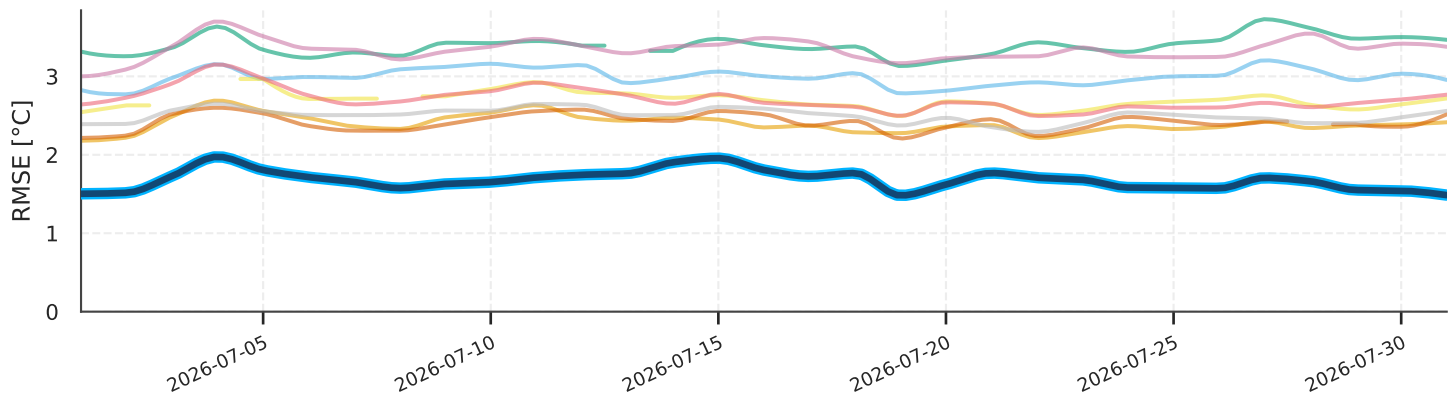
Daily verification scores and overall model ranking

meteoblue mLM model1 model2 model3 model4 model5 model6 model7 model8

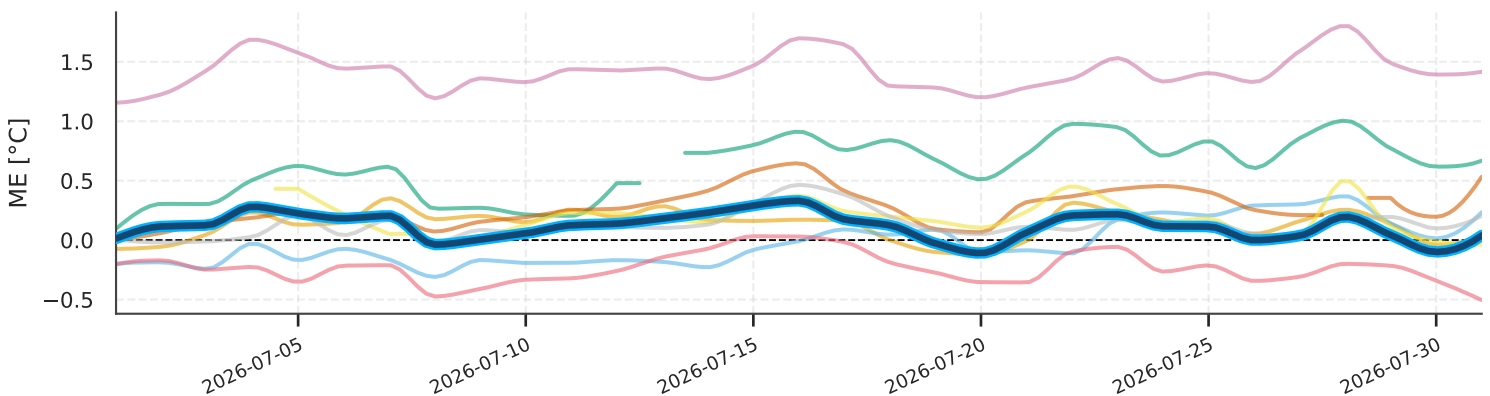
Daily MAE



Daily RMSE

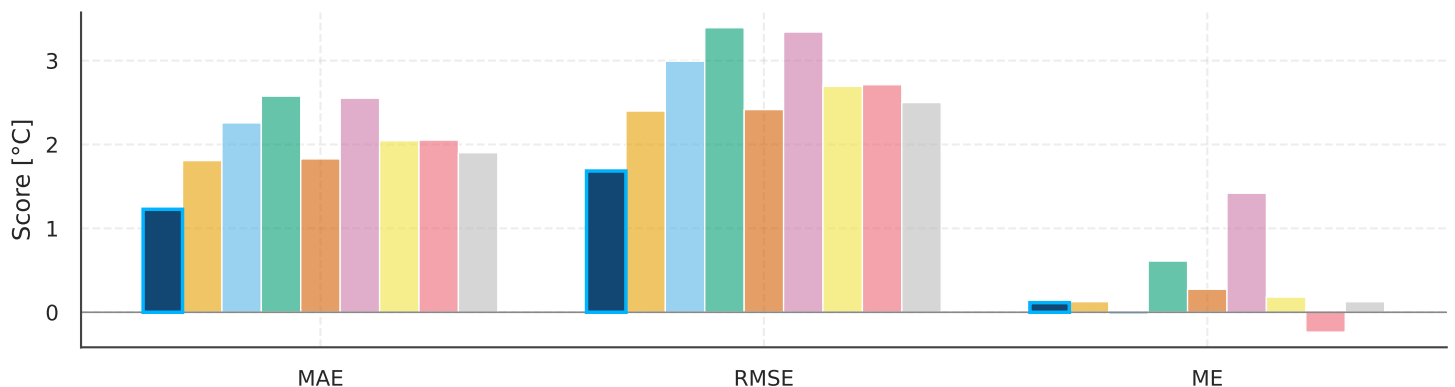


Daily ME



meteoblue mLM model1 model2 model3 model4 model5 model6 model7 model8

Overall scores - North America

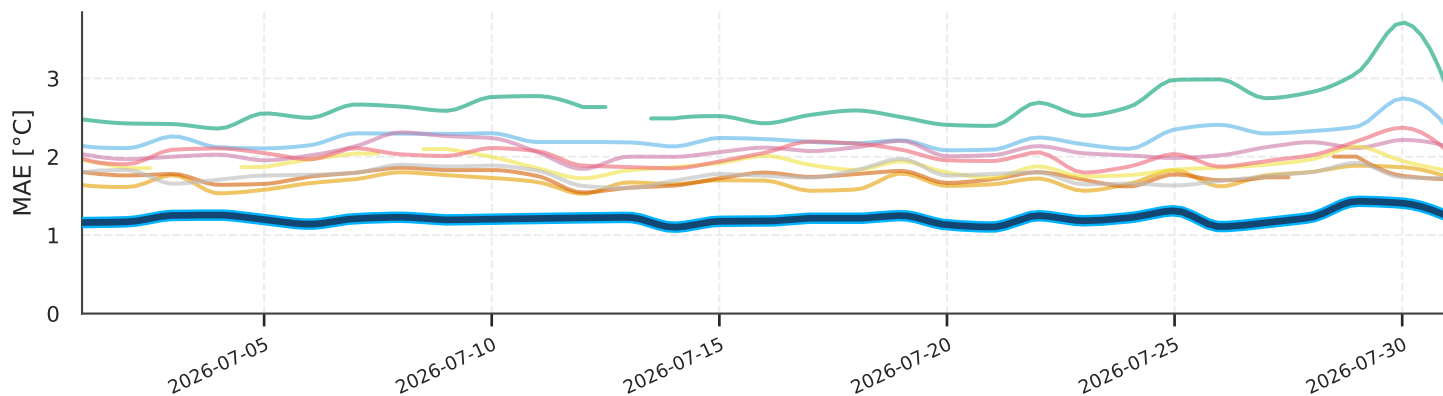


South America - Temperature

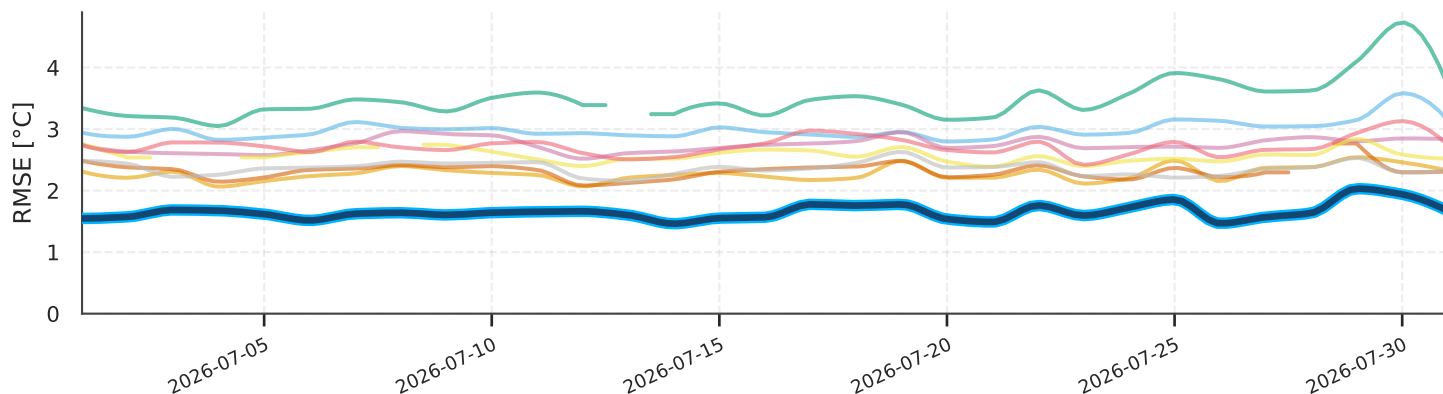
Daily verification scores and overall model ranking

meteoblue mLM model1 model2 model3 model4 model5 model6 model7 model8

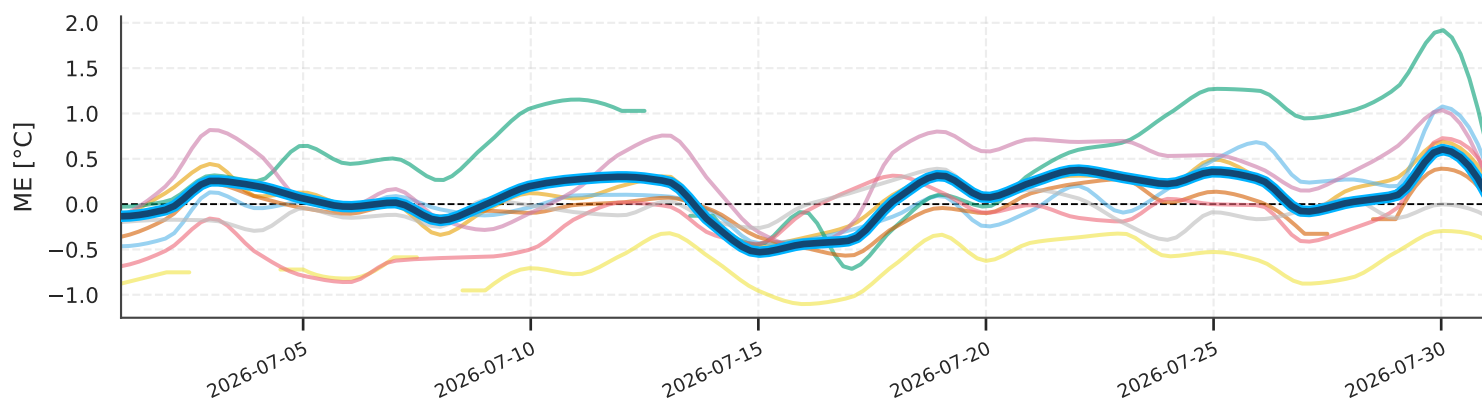
Daily MAE



Daily RMSE

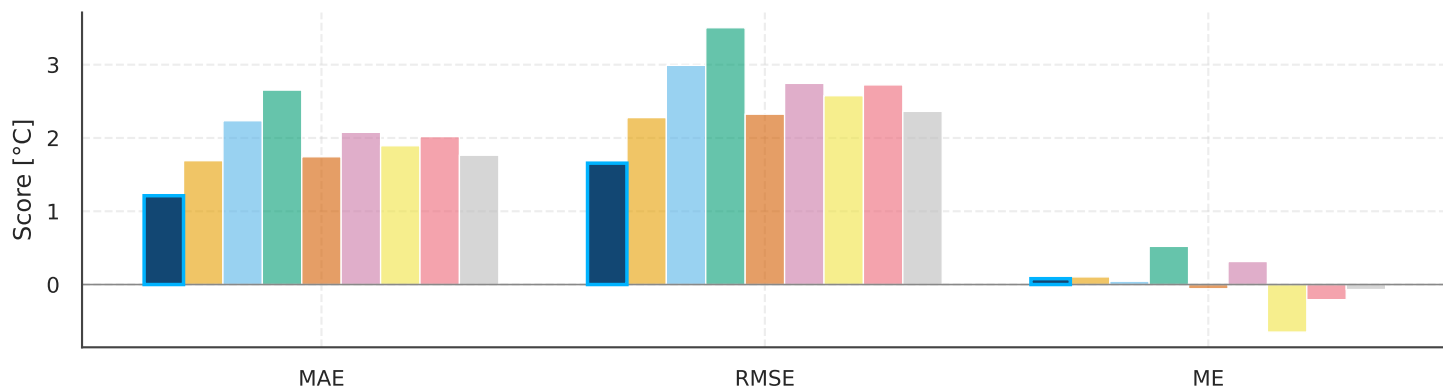


Daily ME



meteoblue mLM model1 model2 model3 model4 model5 model6 model7 model8

Overall scores - South America

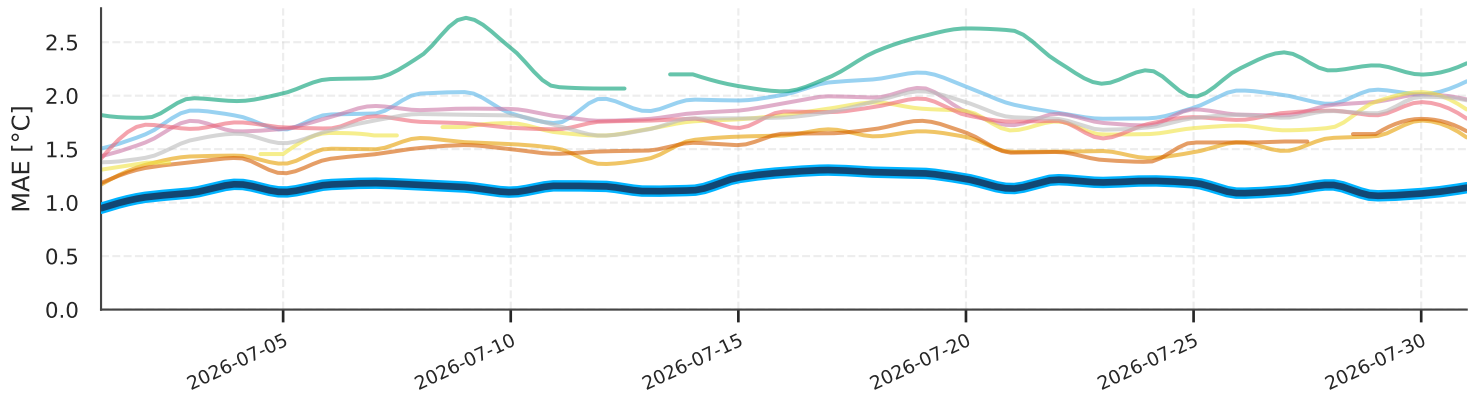


Australia - Temperature

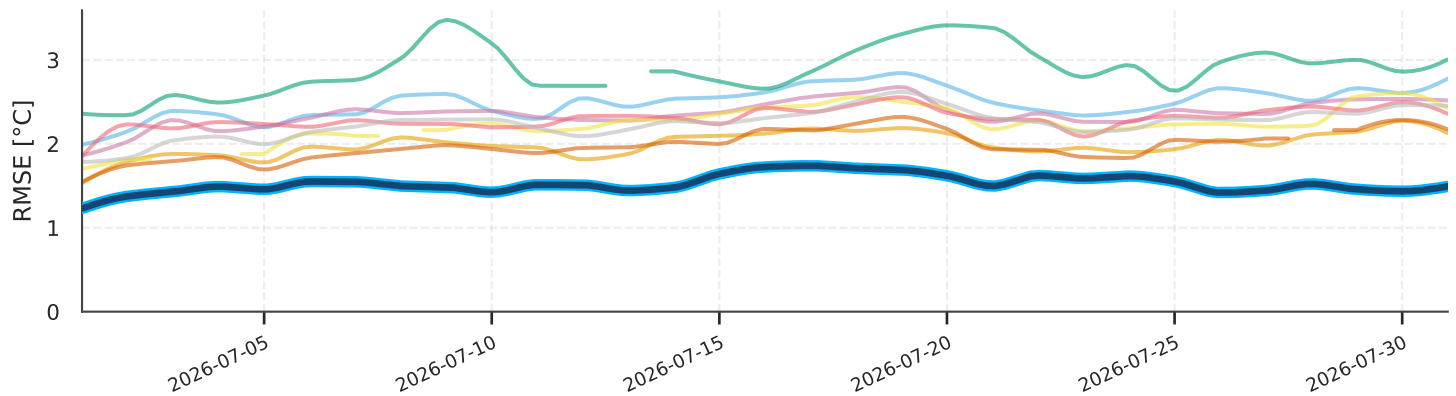
Daily verification scores and overall model ranking

meteoblue mLM model1 model2 model3 model4 model5 model6 model7 model8

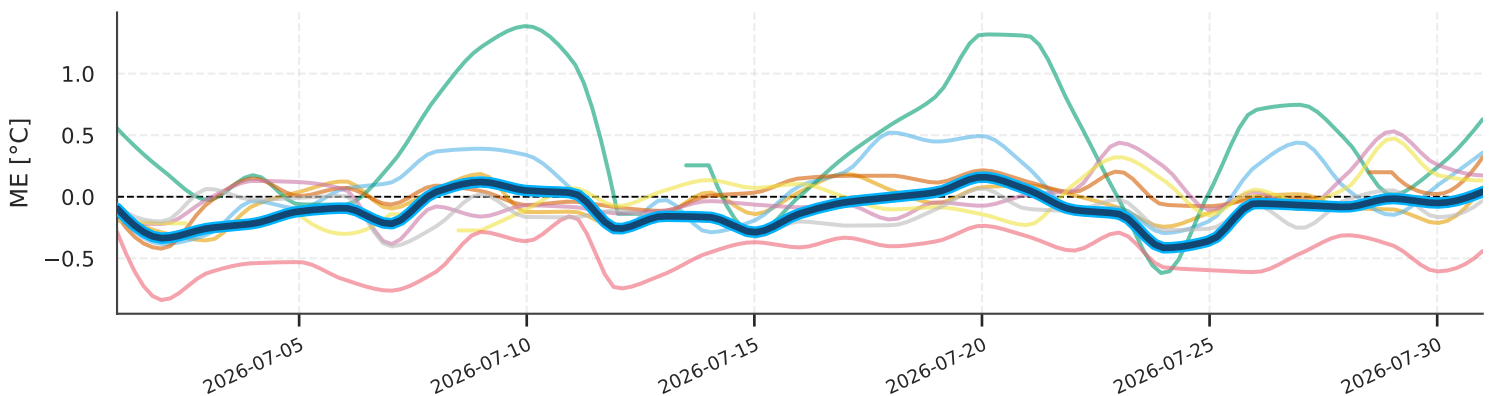
Daily MAE



Daily RMSE



Daily ME



meteoblue mLM model1 model2 model3 model4 model5 model6 model7 model8

Overall scores - Australia

