

Verification as a service to bring more transparency on forecast accuracy to weather services and users

Evelyn Müller

Jan Hoffmann, Dennis Schulze



SCREEN CAPTURE
WELCOME

Forecast verification as an independent service

Forecast User



Who does best for my location?

How many of your forecasts are correct?

What errors should I expect in my provider's wind forecast?

Weather Service .. Sales



How do we compare?

Scientist Developer

Do we predict the diurnal cycle well?



How is our error distributed geographically?

Forecaster

Was my forecast good yesterday?



Which model captured last week's development better?

Management

Has our quality improved after we updated the system?



Use cases

Compare forecasting success of different methods or providers – evaluate the cost-benefit ratio

Identify system weaknesses

Quality monitoring during operations

Case specific analysis

Service expectations

Well communicable measures

Comparison with other forecasting systems

Available the following day

Easily accessible (data already prepared)

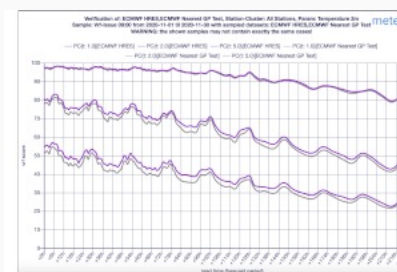
Continuously updated (no one-off project)



SCREEN CAPTURE
WELCOME

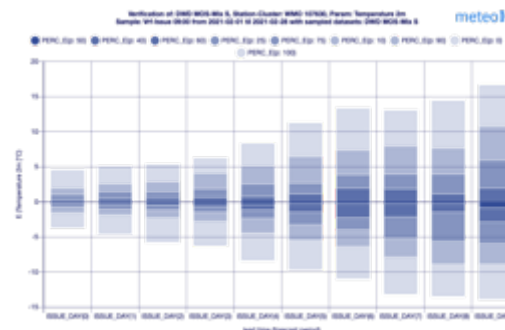
Present results in a web frontend

Sampling method



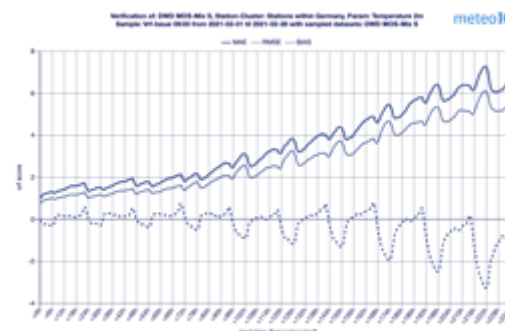
Examples: Answers for forecast users

What errors should I expect in the temperature forecast from supplier X for Nuremberg?



How many of your forecasts are correct?

„In March 2021, 91.3% of the temperature forecasts for Germany for the first day have been correct (within 2.0 °C).“



From which provider will I get the most accurate forecast for my location?

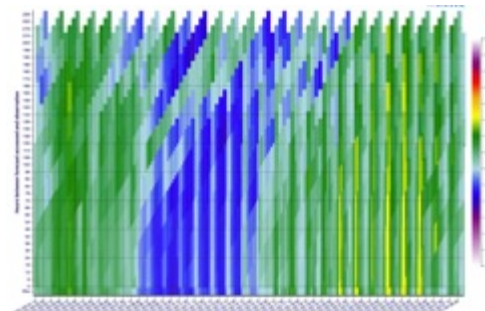
Provider	MAE	%	RMSE	%	BIAS	%	maxE	%	minE	%	Cases
1	0,93		1,17		-0,46		3,1		-2,0		31
2	0,81		1,06		-0,17		3,1		-1,5		30
3	1,00		1,33		-0,43		3,4		-2,8		29
4	0,81		1,13		-0,20		3,5		-2,7		30
5	0,97		1,32		0,10		4,1		-2,4		23
6	0,99		1,28		-0,06		3,5		-2,7		31

Check out more examples on verify.meteoiq.com!

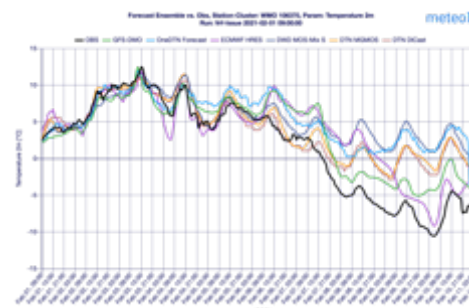


Examples: Answers for the weather room

How did the forecasts cover the cold wave in February?



Users have replied to our forecast from February 1. Have other providers captured the development better?



How good was our forecast yesterday?



Check out more examples
on verify.meteoiq.com!



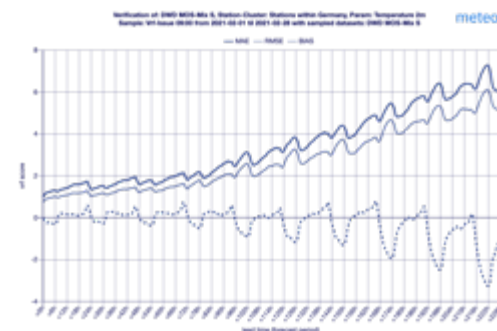
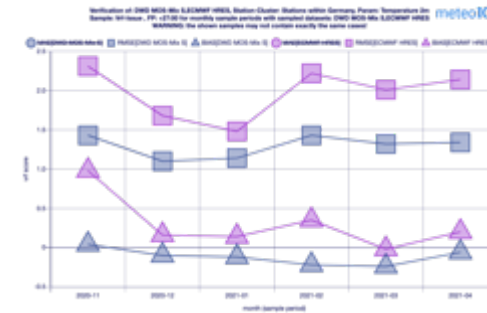
Examples: Answers for the management

How has our forecast quality developed since November?

And compared to other providers?

How good were our forecasts in February?

How were the errors distributed geographically?



Parameter: temperature / Score: all / Provider: all / Sampletype: single / Period: 2021-02 / Leadtime: <27h / Issue: all

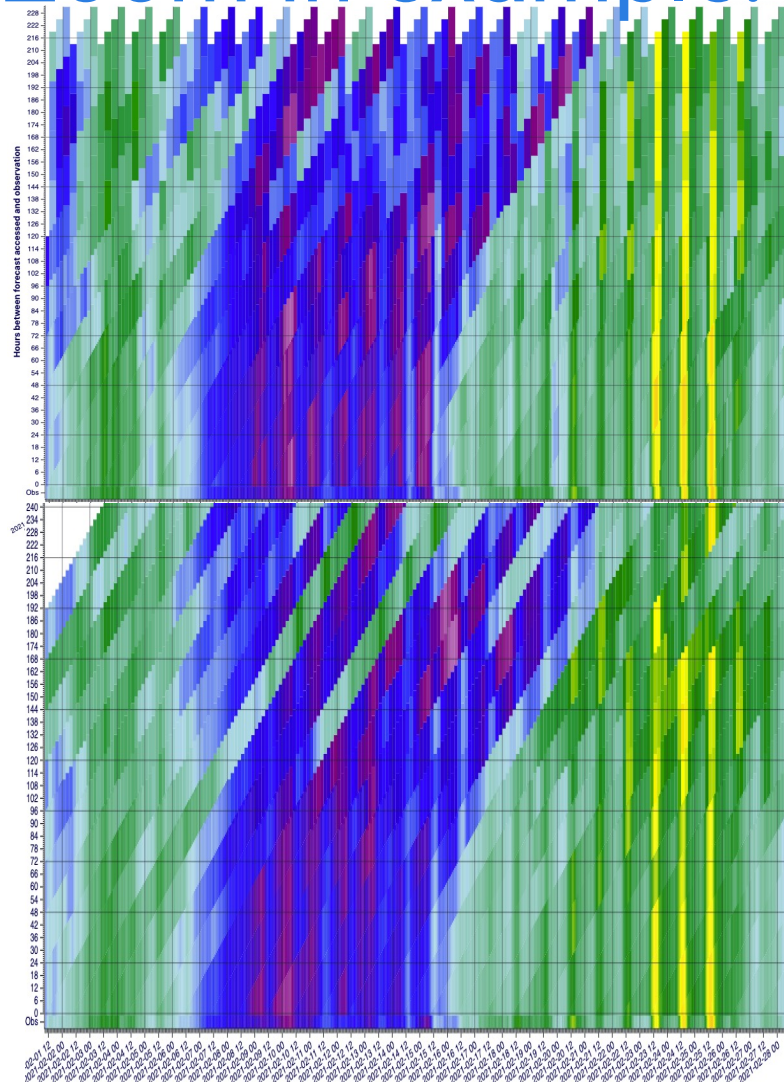
Provider: Score: Provider: Sampletype: Period: Leadtime: Issue: Presentation

GFS-DMO

WMO/ICAO ID	Name	%	MAE	%	RMSE	%	BIAS	%	maxE	%	minE	%	Cases
	Netherlands		1,35		1,79		-0,63		6,7		-5,6		447
06330	Hoek Van Holland		1,85		2,37		-1,65		0,7		-4,8		26
06235	De Kooy		1,68		2,16		-0,83		3,4		-4,9		26
06242	Vlissingen		1,68		2,08		-0,08		3,0		-5,6		27
06350	Gilze Rijen		1,53		1,97		-0,97		5,3		-3,9		26
06290	Tweetha		1,47		1,77		-0,99		2,4		-3,8		26
06260	De Bilt		1,42		1,79		-0,71		4,4		-3,4		26
06280	Eelde		1,42		2,03		-0,52		6,7		-3,4		26

Check out more examples on verify.meteoiq.com!

Zoom-in example: Cold spell Feb 2021 Germany



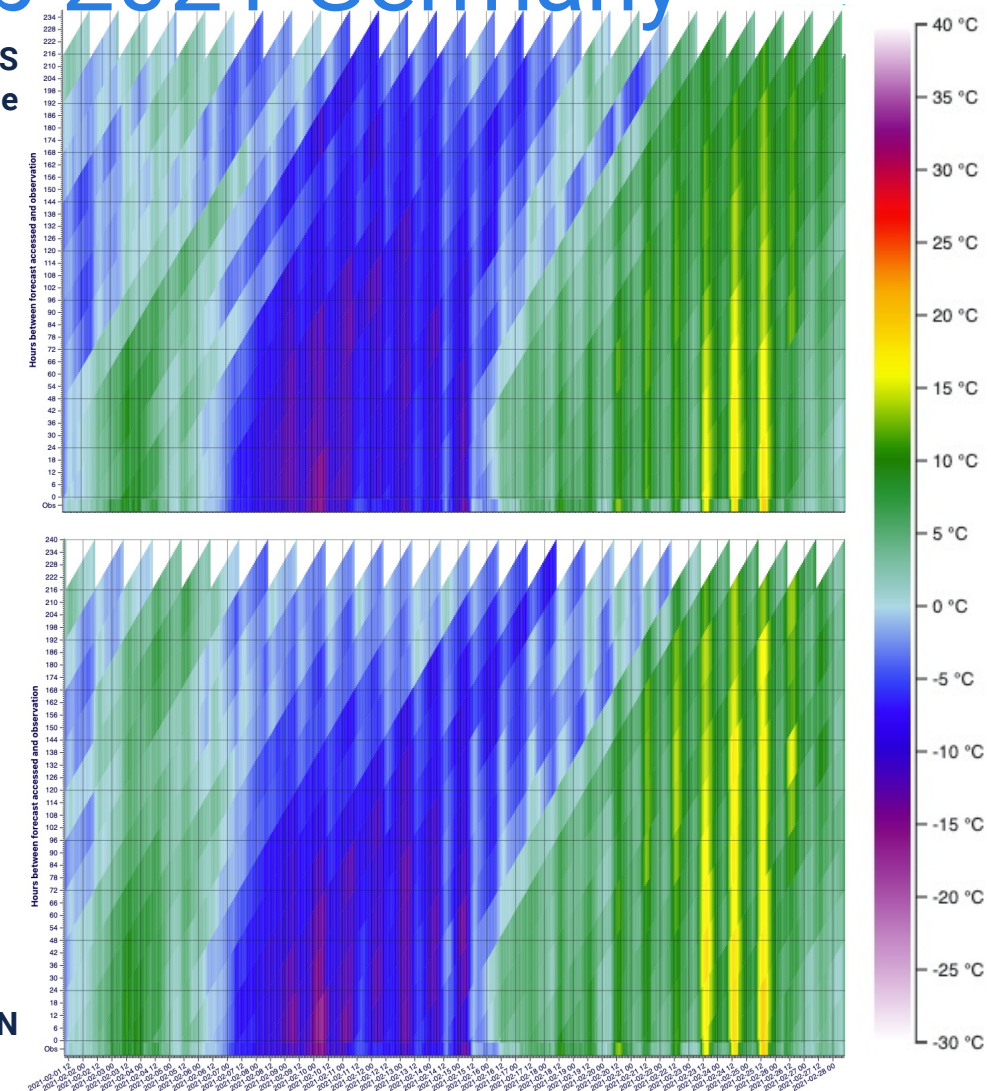
ECMWF HRES

DWD MOSMIX S
opendata.dwd.de

10488 Dresden-Klotzsche
February 2021 2m air temperature
Forecast vs observation (lowest
line).
X-axis: (calendar) time
Y-axis: forecast as available Y
hours before the observation was
made.

Left: Direct model output

Right: Post-processing systems



DTN