



The EUPPBench postprocessing benchmark

The EUPPBench Team *

* **Jonathan Demaeyer**, Jonas Bhend, Sebastian Lerch, Cristina Primo, Bert Van Schaeybroeck, Aitor Atencia, Zied Ben Bouallègue, Jieyu Chen, Markus Dabernig, Gavin Evans, Jana Faganeli Pucer, Ben Hooper, Nina Horat, David Jobst, Janko Merše, Peter Mlakar, Annette Möller, Olivier Mestre, Maxime Taillardat, and Stéphane Vannitsem



EUPPBench - a benchmark for NWP postprocessing

- Statistical postprocessing is an important component of modern forecasting systems
- Many new methods for postprocessing - not least from the rapidly growing field of deep learning
 - Relative performance of new methods often unclear.
- Proposition of benchmark from EUMETNET module postprocessing + scientific community
 - Fair quantitative comparison of methods on real-world data
 - Stimulate scientific progress and collaboration by lowering barriers of entry for different communities
 - Set of benchmarks to continuously and automatically evaluate new methods against pool of published ones
- Here we present the first analysis-ready dataset and an example intercomparison

The EUPPBench dataset v1.0

Forecasts:

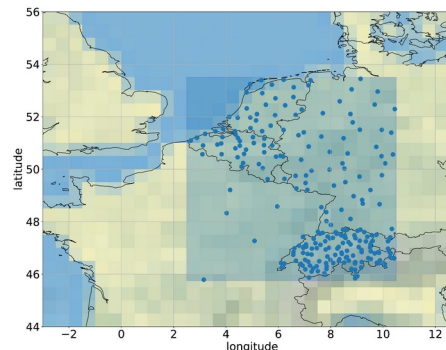
- ECMWF IFS: 0.25° (~25 km), 6-hourly
- Ensemble and high-res
- 730 forecast initializations for 2017-2018
- 209 x 20 corresponding re-forecasts
- Surface and pressure level data (> 20 par.)

Observations:

- Station data from BeNeLux, Germany, France, Austria, and Switzerland
- ERA-5 re-analysis data for gridded analyses

Data access:

- Climetlab: (github.com/EUPP-benchmark)
- Zenodo:
<https://zenodo.org/record/7429236> (gridded)
<https://zenodo.org/record/7708362> (station)



Structure of the dataset

Data stored on the ECMWF European Weather Cloud (EWC) :

- Accessible through a ECMWF climetlab plugin
- Return xarray objects
- Conversion possible to netCDF and other formats

```
In [5]: fcs
```

```
Out[5]: xarray.Dataset
```

```
- Dimensions:      (station_id: 51, number: 51, time: 730, step: 20, surface: 1)
```

```
▼ Coordinates:
```

model_latitude	(station_id)	float64	dask.array<chunksi...		
model_longitude	(station_id)	float64	dask.array<chunksi...		
number	(number)	int64	0 1 2 3 4 5 6 ... 45 4...		
station_altitude	(station_id)	float64	dask.array<chunksi...		
station_id	(station_id)	int32	460 662 691 704		
station_latitude	(station_id)	float64	dask.array<chunksi...		
station_longitude	(station_id)	float64	dask.array<chunksi...		
station_name	(station_id)	<U20	dask.array<chunksi...		
step	(step)	timedelta64[ns]	0 days 06:00:00		
surface	(surface)	float64	0.0		
time	(time)	datetime64[ns]	2017-01-01 ... 2018...		
valid_time	(time, step)	datetime64[ns]	dask.array<chunksi...		

```
▼ Data variables:
```

cp6	(station_id, number, time, step, surface)	float32	dask.array<chunksi...		
mn2t6	(station_id, number, time, step, surface)	float32	dask.array<chunksi...		
mx2t6	(station_id, number, time, step, surface)	float32	dask.array<chunksi...		
p10fg6	(station_id, number, time, step, surface)	float32	dask.array<chunksi...		
silh6	(station_id, number, time, step, surface)	float32	dask.array<chunksi...		
sshf6	(station_id, number, time, step, surface)	float32	dask.array<chunksi...		
ssr6	(station_id, number, time, step, surface)	float32	dask.array<chunksi...		
ssrd6	(station_id, number, time, step, surface)	float32	dask.array<chunksi...		
str6	(station_id, number, time, step, surface)	float32	dask.array<chunksi...		

Structure of the dataset

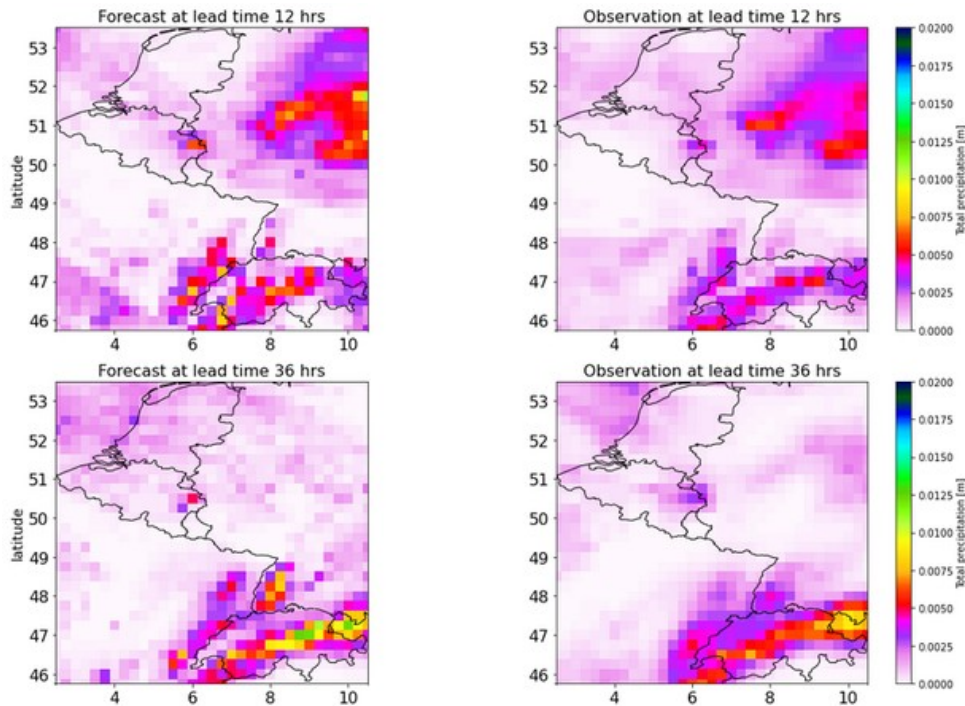
Data stored on the ECMWF European Weather Cloud (EWC) :

- Accessible through a ECMWF climetlab plugin
- Return xarray objects
- Conversion possible to netCDF and other formats
- Time-aligned observation and (re)forecasts

```
In [6]: obs
Out[6]: xarray.Dataset
```

```
- Dimensions:      (station_id: 51, time: 730, step: 20)
- Coordinates:
  altitude          (station_id)      float64  dask.array<chunksize=(1,), meta=np...
  latitude          (station_id)      float64  dask.array<chunksize=(1,), meta=np...
  longitude         (station_id)      float64  dask.array<chunksize=(1,), meta=np...
  station_id       (station_id)      int32    460 662 691 704 ... 5839 5871 5906
  station_name     (station_id)      <U20    dask.array<chunksize=(1,), meta=np...
  step             (step)             timedelta64[ns]  0 days 06:00:00 ... 5 days 00:00:00
  time            (time)             datetime64[ns]  2017-01-01 ... 2018-12-31
- Data variables:
  p10fg6          (time, step, station_id)  float64  dask.array<chunksize=(730, 20, 1), m...
  tp6             (time, step, station_id)  float64  dask.array<chunksize=(730, 20, 1), m...
- Attributes:
  full_dataset_me...
  history :      Retrieved from https://opendata.dwd.de/climate_environment/CDC/, March 2022
  source :      DWD, Deutscher Wetterdienst, https://www.dwd.de/
```

Example of a gridded forecast - total precip. (high-res)





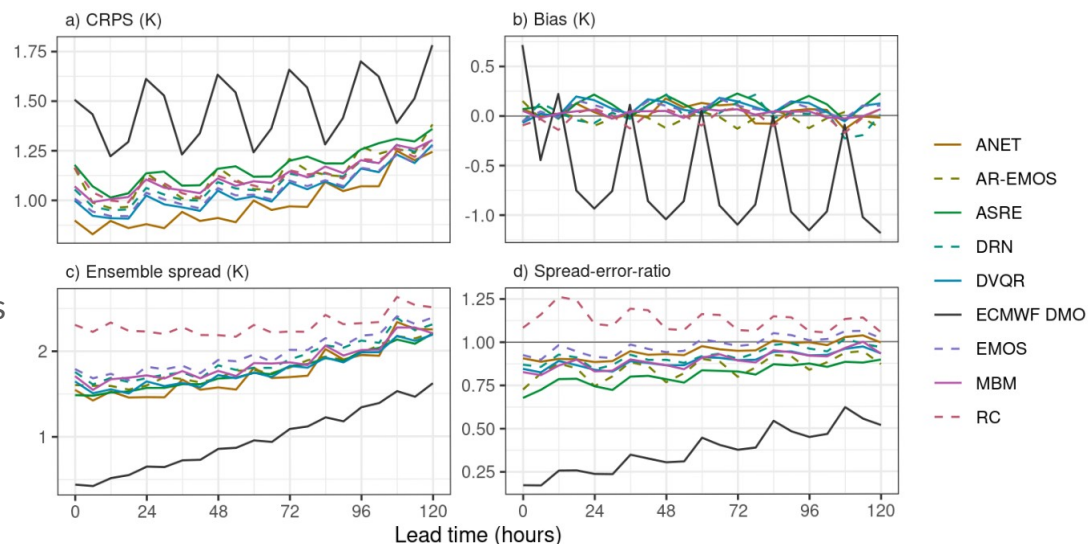
A first postprocessing experiment - 2m temperature

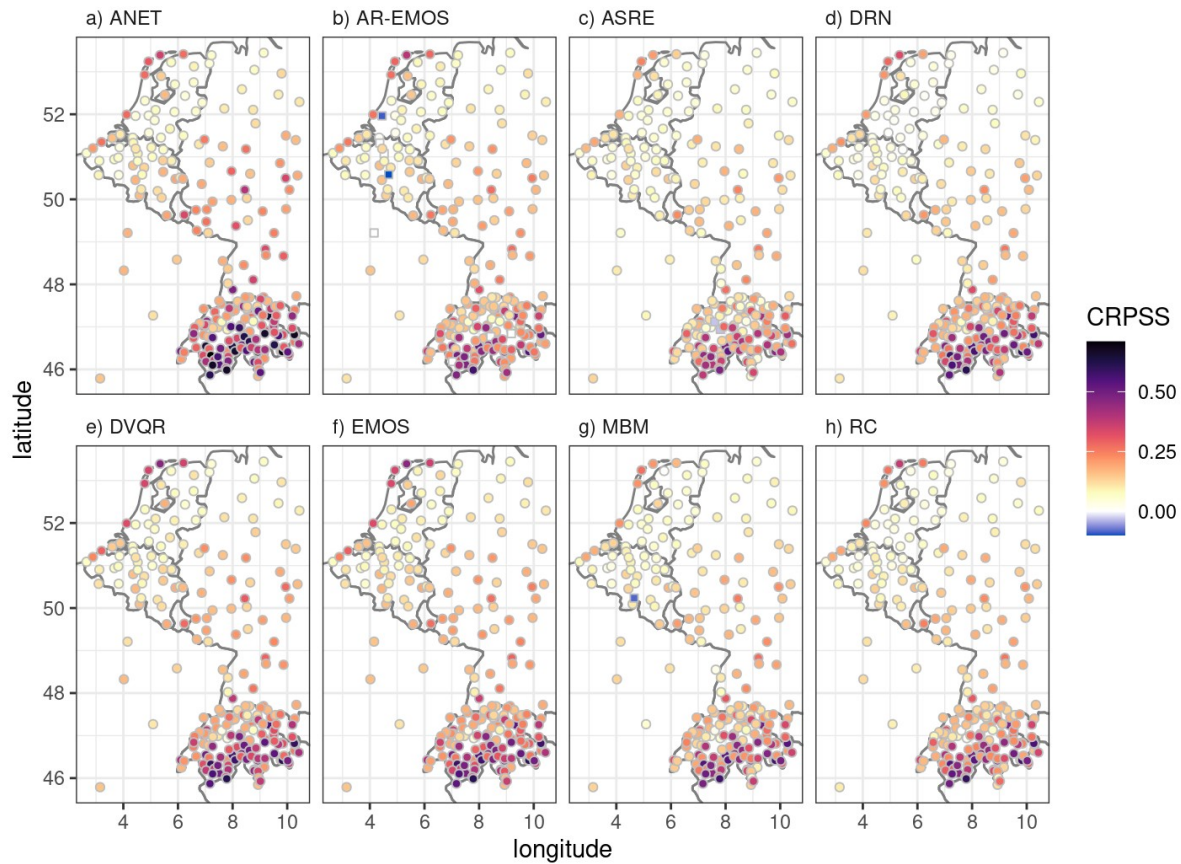
- Postprocessing of instantaneous temperature at observed locations
 - Forecasts from nearest ECMWF IFS grid point
- Ensemble temperature as the sole predictor
 - Most methods use ensemble mean and standard deviation
 - Some methods use information from each member
 - Static metadata (lon, lat, model orography, lead time, day of year) is implicitly or explicitly used
 - Many methods are local, some are global
- Only 11 members in training set, but 51 members in test set

A first postprocessing experiment - 2m temperature

- Range of methods, varying complexity
- All methods
 - reduce bias
 - improve calibration
 - improve accuracy
- Some methods perform better than others
- Methods published on github:

<https://github.com/EUPP-benchmark>







Lessons learned from first experiment

- Quality control matters
 - Are the requirements for data format specific enough and do submissions adhere to it?
 - Verification of submissions uncovers obvious issues (e.g. indexing errors, partial downloads)
 - Additional diagnostics (not used for verification) provide further insight
 - An issue with the verifying observations (and also for the training data) was only discovered late in the process (fix was included with the revised manuscript)
- Establish a quick feedback loop to resolve remaining issues with submitters



The future of EUPPBench

In the course of this year:

- Additional time-boxed experiments grouped around challenges
- Gather experience for future automation

In the coming years:

- Automation of (parts of) the pipeline to allow continued submissions and evaluation
- Addition of new datasets to address specific challenges (e.g. postprocessing of high-res NWP)



How can I participate?

- Most data of the benchmark are public and available → Reproducibility
- Closed data → Access through participation to the EUMETNET PP activities
- Participation open to anyone for research purposes
- New forthcoming challenges for the next phase:

- Postprocessing at unseen locations
- Freestyle postprocessing at observed locations
- Multivariate postprocessing (time)
- Multivariate postprocessing (space)
- Multivariate postprocessing (parameters)

(still under discussion)

- Contact Jonathan Demaeyer (jodemaey@meteo.be) if you are interested

Article in ESSD

<https://doi.org/10.5194/essd-15-2635-2023>



The EUPPBench postprocessing benchmark dataset v1.0

Jonathan Demaeyer^{1,2}, Jonas Bhend³, Sebastian Lerch⁴, Cristina Primo⁵, Bert Van Schaeybroeck¹, Aitor Atencia⁶, Zied Ben Bouallègue⁷, Jieyu Chen⁴, Markus Dabernig⁶, Gavin Evans⁸, Jana Faganeli Pucer⁹, Ben Hooper⁸, Nina Horat⁴, David Jobst¹⁰, Janko Merše¹¹, Peter Mlakar^{9,11}, Annette Möller¹², Olivier Mestre¹³, Maxime Taillardat¹³, and Stéphane Vannitsem^{1,2}

¹Royal Meteorological Institute of Belgium, Brussels, Belgium

²European Meteorological Network (EUMETNET), Brussels, Belgium

³Federal Office of Meteorology and Climatology MeteoSwiss, Zurich, Switzerland

⁴Karlsruhe Institute of Technology, Karlsruhe, Germany

⁵Deutscher Wetterdienst, Offenbach, Germany

⁶GeoSphere Austria, Vienna, Austria

⁷European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom

⁸Met Office, Exeter, United Kingdom

⁹University of Ljubljana, Faculty of Computer and Information Science, Slovenia

¹⁰University of Hildesheim, Hildesheim, Germany

¹¹Slovenian Environment Agency, Ljubljana, Slovenia

¹²Bielefeld University, Bielefeld, Germany

¹³Meteo France, Ecole Nationale de la Meteorologie, Toulouse, France

Correspondence: Jonathan Demaeyer (Jonathan.Demaeyer@meteo.be)