

Statistical and machine learning methods for postprocessing ensemble forecasts of wind gusts

Benedikt Schulz and Sebastian Lerch

Karlsruhe Institute of Technology (KIT)

vEGU – NP5.3, 28 April 2021

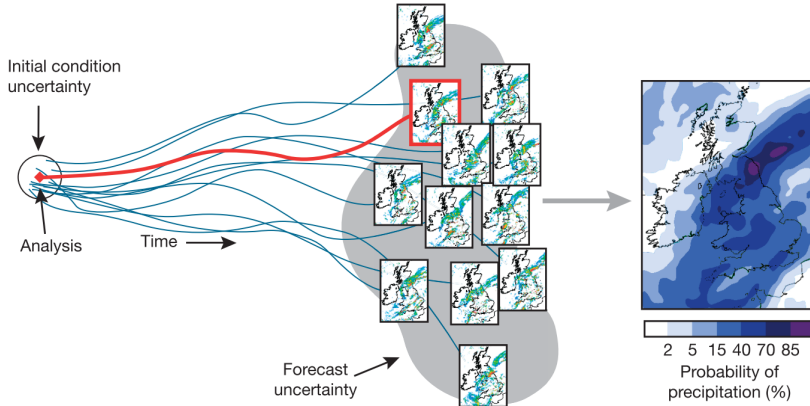
Advances in statistical post-processing, blending and verification
of deterministic and ensemble forecasts

Agenda

- 1 Motivation
- 2 Basic Postprocessing Methods
- 3 Machine Learning Techniques
- 4 Locally Adaptive Networks
- 5 Forecast Performance

Postprocessing of Ensemble Forecasts

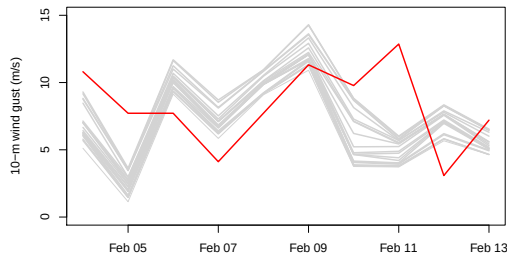
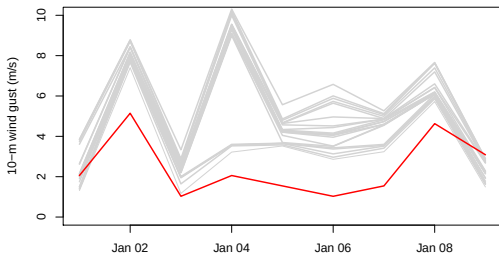
Ensemble forecasts try to quantify the **uncertainty** in numerical weather prediction...



Source: Bauer, P., Thorpe, A. and Brunet, G. (2015): **The quiet revolution of numerical weather prediction**, *Nature*, **525**, 47–55.

Postprocessing of Ensemble Forecasts

Ensemble forecasts try to quantify the **uncertainty** in numerical weather prediction...
...but are subject to **biases** and **dispersion errors**.



COSMO-DE ensemble forecast (sorted) and **wind gust observations**.

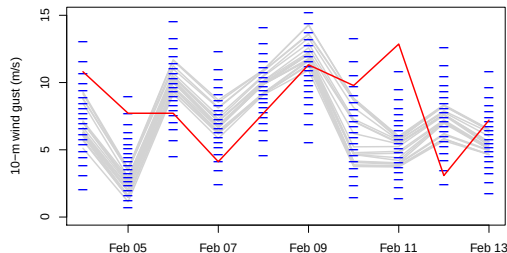
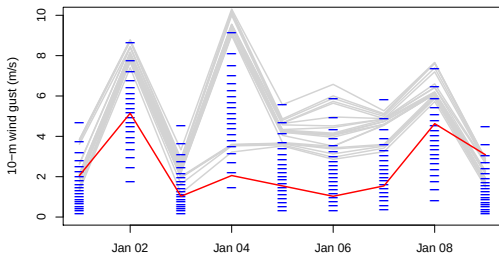
Left: 6h-forecasts in Muehlacker from 01/01/2016 to 09/01/2016.

Right: 5h-forecasts in Stoetten from 04/02/2016 to 13/02/2016.

Postprocessing of Ensemble Forecasts

Ensemble forecasts try to quantify the **uncertainty** in numerical weather prediction...

...but are subject to **biases** and **dispersion errors**. $\Rightarrow$ **Statistical postprocessing**



COSMO-DE ensemble forecast (sorted) , EMOS quantile forecast and wind gust observations.

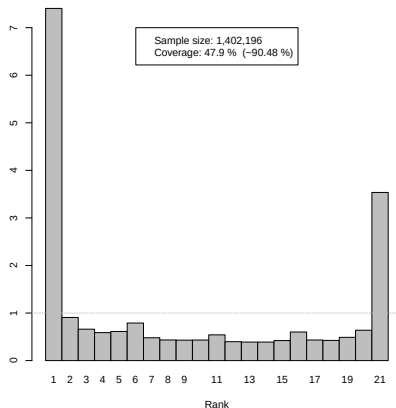
Left: 6h-forecasts in Muehlacker from 01/01/2016 to 09/01/2016.

Right: 5h-forecasts in Stoetten from 04/02/2016 to 13/02/2016.

Data

- 20-member COSMO-DE-EPS.
- 2.8km horizontal resolution.
- 61 variables incl. wind gusts.
- Initialization time: 00 UTC.
- Lead times: 00 – 21h.
- Time range: 08/12/2010 – 31/12/2016.
- Training in 2010–2015, evaluation in 2016.
- Observations at 175 SYNOP stations.
- Forecasts taken from closest grid point.

For operational postprocessing at DWD:
See Hess (2020).



Rank histogram of the ensemble forecasts of wind gusts for all lead times and locations in 2016.

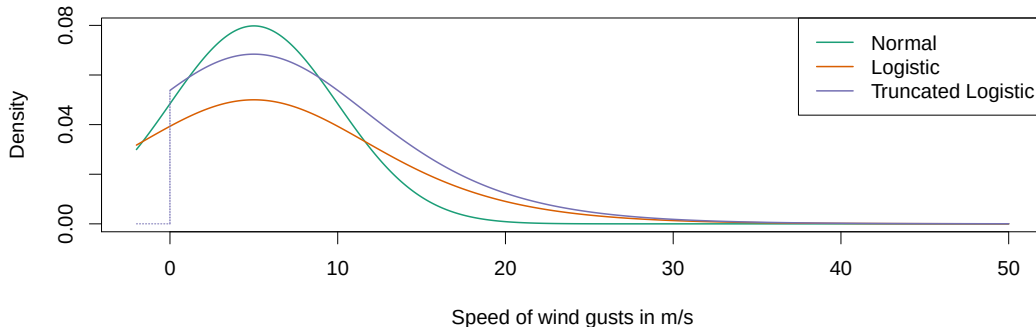
Hess, R. (2020): **Statistical postprocessing of ensemble forecasts for severe weather at Deutscher Wetterdienst**, *Nonlinear Processes in Geophysics*, **27**, 473–487.

Agenda

- 1 Motivation
- 2 Basic Postprocessing Methods**
- 3 Machine Learning Techniques
- 4 Locally Adaptive Networks
- 5 Forecast Performance

Ensemble Model Output Statistics (EMOS)

Forecast form: Zero-truncated logistic distribution with location μ and scale $\sigma > 0$.



Ensemble Model Output Statistics (EMOS)

Forecast form: Zero-truncated logistic distribution with location μ and scale $\sigma > 0$.

Method: Distribution parameters are connected to the wind gust ensemble via

$$\mu = a + b\bar{x}, \quad \sigma = \exp(c + d \log s(x)), \quad b > 0, \quad a, c, d \in \mathbb{R},$$

where $\bar{x}$ is the ensemble mean and $s(x)$ the ensemble standard deviation.

Estimation: Locally via CRPS minimization on seasonal training sets.

Gneiting, T., Raftery, A. E., Westveld, A. H. and Goldman, T. (2005): **Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation**, *Monthly Weather Review*, **133**, 1098–1118.

Member-by-Member (MBM)

Idea: Adjust each member individually to generate a calibrated ensemble.

Method: Calibrated ensemble $\tilde{x}$ is calculated via

$$\begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_{20} \end{pmatrix} := (a + b \cdot \bar{x}) + \gamma \cdot \left(\begin{pmatrix} x_1 \\ \vdots \\ x_{20} \end{pmatrix} - \bar{x} \right), \quad \gamma := c + \frac{d}{\delta}, \quad a, b, c, d \in \mathbb{R},$$

where x is the ensemble, $\bar{x}$ is the ensemble mean and δ the ensemble spread.

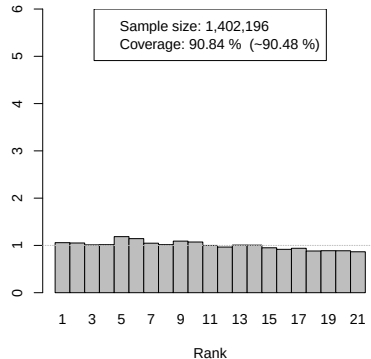
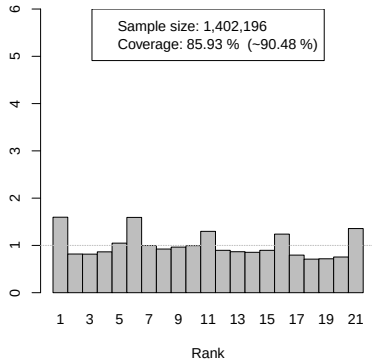
Estimation: Locally via CRPS minimization on seasonal training sets.

Van Schaeybroeck, B. and Vannitsem, S. (2015): **Ensemble post-processing using member-by-member approaches: theoretical aspects**, *Quarterly Journal of the Royal Meteorological Society*, **141**, 807–818.

MBM – Incorporating Sub-Models

COSMO consists of **four sub-models** with each five members.

⇒ Use separate b_k and d_k for each sub-ensemble mean and spread, $k = 1, \dots, 4$.



Rank histogram of general MBM forecasts (left) and MBM forecasts incorporating sub-models (right) for all lead times and locations in 2016 (here: non-seasonal training sets).

Isotonic Distributional Regression (IDR)

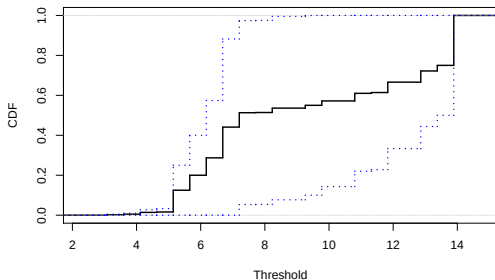
Non-parametric, **generic benchmark** for probabilistic forecasting.

Forecast: Data-driven, flexible CDF in form of a step-function.

Method: Far-reaching generalization of isotonic regression techniques based on the PAV-algorithm. Assumption of an isotonic relationship between predictors and response.

Predictors: Ensemble members.

Estimation: Local.



Exemplary predictive CDF (black) with boundaries indicating uncertainty (blue).

Henzi, A., Ziegel, J. F. and Gneiting, T. (2019): **Isotonic Distributional Regression**, working paper, <https://arxiv.org/abs/1909.03725>.

Additional Predictors

Feature	Unit	Description
VMAX	m/s	Maximum wind, i.e. wind gusts (10m)
U	m/s	U-Component of wind (10m, 500–1,000 hPa)
V	m/s	V-Component of wind (10m, 500–1,000 hPa)
WIND	m/s	Wind speed (10m, 500–1,000 hPa), derived from U and V via $\sqrt{U^2 + V^2}$
T	K	Temperature (Ground-level, 2m, 500–1,000 hPa)
TMIN	K	Minimum temperature (2m)
TMAX	K	Maximum temperature (2m)
TD	K	Dew point temperature (2m)
TOT_PREC	kg/m ²	Total precipitation (Accumulation)
RAIN_GSP	kg/m ²	Large scale rain (Accumulation)
SNOW_GSP	kg/m ²	Large scale snowfall - water equivalent (Accumulation)
W_SNOW	kg/m ²	Snow depth water equivalent
CLCT	%	Total cloud cover
CLCL	%	Cloud cover (800 hPa - soil)
CLCM	%	Cloud cover (400 hPa - 800 hPa)
CLCH	%	Cloud cover (000 hPa - 400 hPa)
HBAS_SC	m	Cloud base above mean sea level, shallow connection
HTOP_SC	m	Cloud top above mean sea level, shallow connection
ASOB_S	W/m ²	Net short wave radiation flux (at the surface)
ATHB_S	W/m ²	Net long wave radiation flux (m) (at the surface)
ALB_RAD	%	Albedo (in short-wave)
PMSL	Pa	Pressure reduced to mean sea level
RELHUM	%	Relative humidity (500–1,000 hPa)
FI	m ² /s ²	Geopotential (500–1,000 hPa)
OMEGA	Pa/s	Vertical velocity (Pressure) (500–1,000 hPa)
W_S01	kg/m ²	Column integrated soil moisture (multilayers) (also for 2, 6, 18, 54 instead of 1)

The previous methods only use **VMAX_10M**.

Use **Machine Learning** to incorporate additional features.

Agenda

- 1 Motivation
- 2 Basic Postprocessing Methods
- 3 Machine Learning Techniques**
- 4 Locally Adaptive Networks
- 5 Forecast Performance

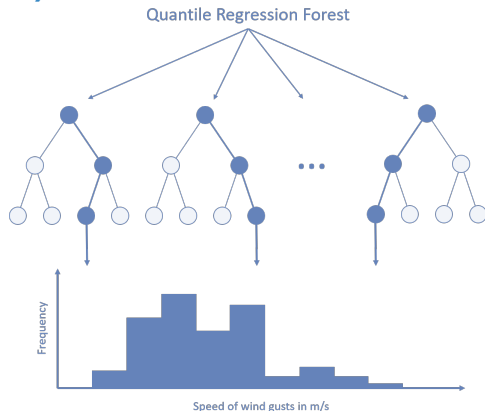
Quantile Regression Forests (QRF)

Forecast form: Set of quantiles given by past observations.

Estimation: Local.

Predictors:

- Mean and standard deviation of ensemble forecast of wind gusts.
- Cosine-transformed day of the year.
- Ensemble mean of additional variables.



Taillardat, M., Mestre, O., Zamo, M. and Naveau, P. (2016): **Calibrated ensemble forecasts using quantile regression forests and ensemble model output statistics**, *Monthly Weather Review*, **144**, 2375–2393.

EMOS – Gradient Boosting Extension (EMOS-GB)

Recall EMOS: Distribution parameters are connected to the predictors via

$$\mu = a + b\bar{x}, \quad \sigma = \exp(c + d \log s(x)), \quad b > 0, \quad a, c, d \in \mathbb{R}.$$

Question: How can we include additional features like temperature or precipitation?

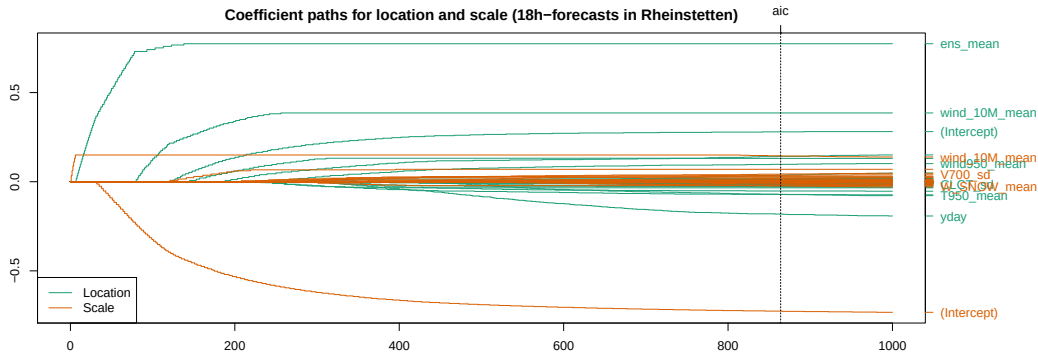
Boosting: Update coefficients of predictors iteratively depending on improvement of current fit until **stopping criterion** is reached.

Messner, J. W., Mayr, G. J. and Zeileis, A. (2017): **Nonhomogeneous boosting for predictor selection in ensemble postprocessing**, *Monthly Weather Review*, **145**, 137–147.

EMOS – Gradient Boosting Extension (EMOS-GB)

Estimation: Locally via maximum likelihood (ML).

Predictors: QRF predictors + ensemble standard deviations of additional variables.



Agenda

- 1 Motivation
- 2 Basic Postprocessing Methods
- 3 Machine Learning Techniques
- 4 Locally Adaptive Networks**
- 5 Forecast Performance

Locally Adaptive Networks

Use **Neural Networks** to generate a locally adaptive joint model.

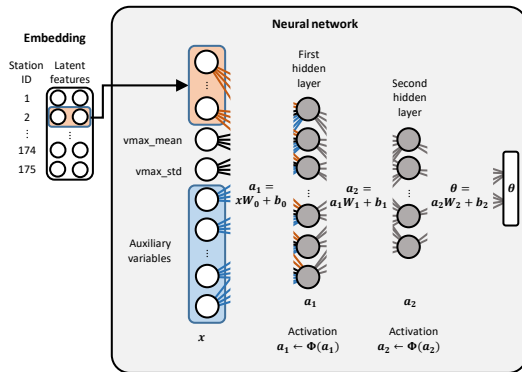
⇒ **Station Embedding**

Three different network variants:

- Distributional Regression Network (DRN)
- Bernstein Quantile Network (BQN)
- Histogram Estimation Network (HEN)

Each variant is based on an ensemble of ten networks, whose forecasts are aggregated.

Predictors: QRF predictors + additional spatial predictors (e.g. station altitude).

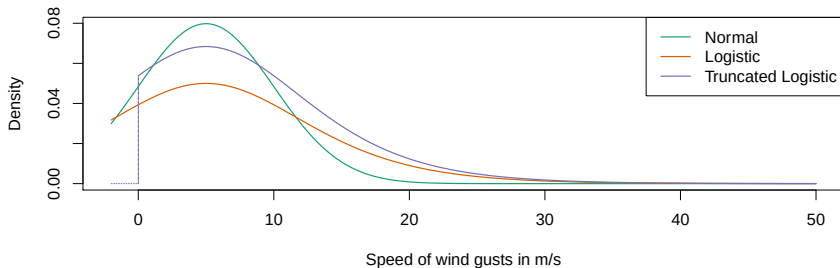


Adapted from:

Rasp, S. and Lerch, S. (2018): **Neural networks for post-processing ensemble weather forecasts**, *Monthly Weather Review*, **146**, 3885-3900.

Distributional Regression Network (DRN)

Forecast form: Zero-truncated logistic distribution.



Loss function: CRPS (only for $\mu \geq 0$ due to numerical reasons).

NN output: Location parameter $\mu \geq 0$ and scale parameter $\sigma > 0$.

Aggregation: Parameter averaging ($\sim$ quantile aggregation).

Rasp, S. and Lerch, S. (2018): **Neural networks for post-processing ensemble weather forecasts**, *Monthly Weather Review*, **146**, 3885–3900.

Bernstein Quantile Network (BQN)

Forecast form: Quantile function.

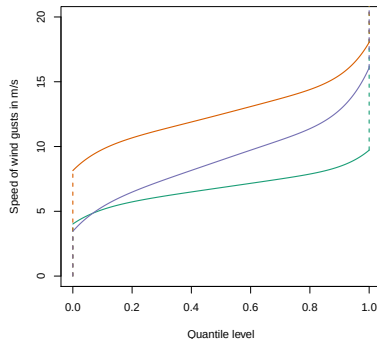
Forecast is defined by the coefficients α_j via

$$Q(\tau) := \sum_{j=0}^d \alpha_j \binom{d}{j} \tau^j (1 - \tau)^{d-j}, \quad \tau \in [0, 1].$$

Loss function: Quantile loss at levels 1%, 2%, ..., 99%.

NN output: Coefficients $0 \leq \alpha_0 \leq \dots \leq \alpha_d$.

Aggregation: Coefficient averaging (quantile aggr.).



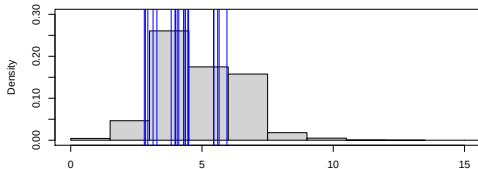
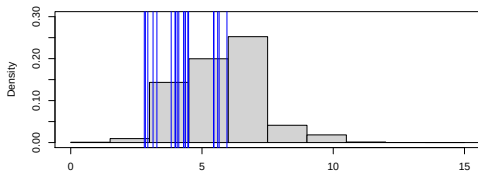
Bremnes, J. B. (2020): **Ensemble postprocessing using quantile function regression based on neural networks and Bernstein polynomials**, *Monthly Weather Review*, **148**, 403–414.

Histogram Estimation Network (HEN)

Idea: Approximate predictive density with a histogram consisting of N bins.

NN output: N bin probabilities, i.e. probability of observing wind gusts within the bin range.

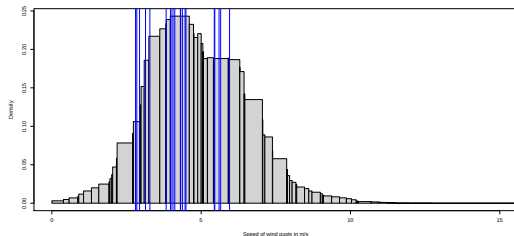
Loss function: Categorical cross-entropy (equivalent to ML).



Speed of wind gusts in m/s

Ensemble of ten networks...

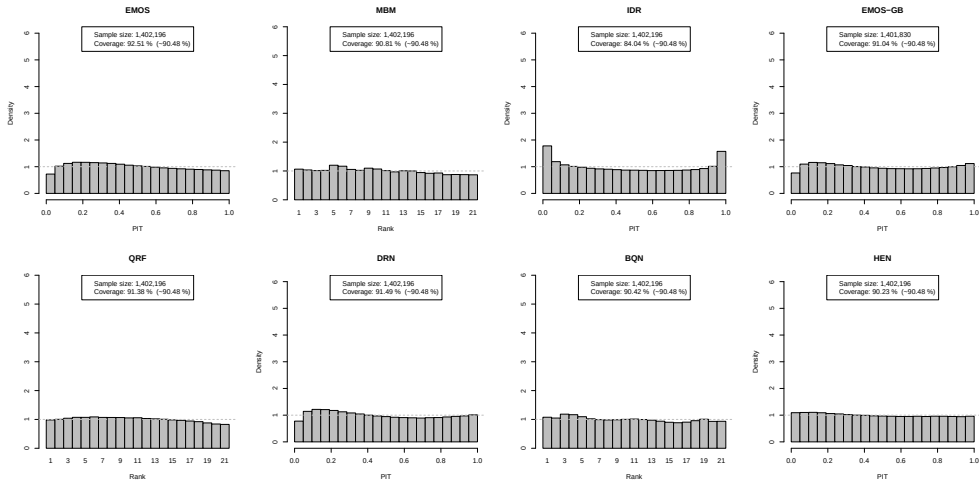
...combination via **quantile aggregation**.



Agenda

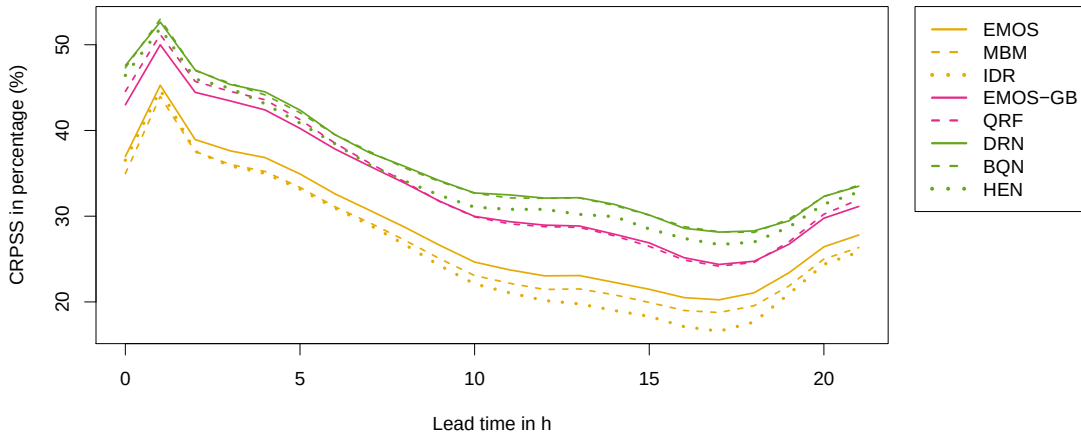
- 1 Motivation
- 2 Basic Postprocessing Methods
- 3 Machine Learning Techniques
- 4 Locally Adaptive Networks
- 5 Forecast Performance

Rank and PIT Histograms



Rank and PIT histograms of the postprocessing methods for all lead times.
 Training in 2010–2015, evaluation in 2016. Each method is applied for each lead time separately.

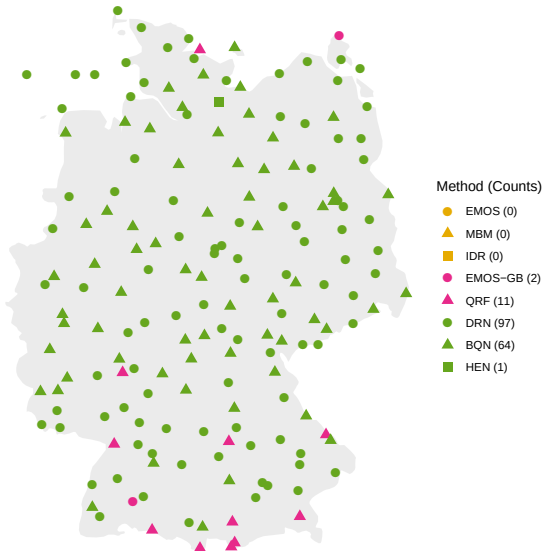
CRPSS over Lead Time



CRPS skill of the postprocessing methods w.r.t. the ensemble dependent on the lead time. Training in 2010–2015, evaluation in 2016. Each method is applied for each lead time separately.

Higher means better.

Stationwise Performance



Best method per location
in terms of the CRPS.

Training in 2010–2015,
evaluation in 2016.

Each method is applied
for each lead time separately.

Locally adaptive networks
perform best at
162 of 175 stations (ca. 92.6%).

Conclusions and Outlook

Conclusions:

- Basic techniques improve forecasts by up to **45%**.
- Using **additional predictors** increases skill by 5–10%.
- Locally adaptive networks **outperform** locally estimated state-of-the-art methods at more than 90% of the stations.

Outlook:

- Feature importance: What variables are relevant for corrections?
- Postprocessing performance in **winter storms**.
- Regime-dependent postprocessing.

Backup: Comparison of Postprocessing Methods

Overall performance of the EPS and the postprocessing methods. PI refers to the corresponding prediction interval with a coverage corresponding to a calibrated 20-member ensemble ($\sim 90.48\%$). The best methods are given in bold, where the best coverage is the closest to the nominal ensemble coverage.

Methods	CRPS	MAE	RMSE	Bias	PI length	Coverage	Runtime
EPS	1.33 m/s	1.63 m/s	2.16 m/s	0.47 m/s	2.85 m/s	47.90 %	-
EMOS	0.95 m/s	1.32 m/s	1.79 m/s	0.05 m/s	5.94 m/s	92.51 %	19 min
MBM	0.97 m/s	1.34 m/s	1.80 m/s	0.04 m/s	6.10 m/s	90.81 %	51255 min
IDR	0.98 m/s	1.36 m/s	1.84 m/s	0.01 m/s	4.72 m/s	84.04 %	8102 min
EMOS-GB	0.88 m/s	1.23 m/s	1.68 m/s	-0.06 m/s	5.24 m/s	91.04 %	510 min
QRF	0.87 m/s	1.22 m/s	1.65 m/s	-0.03 m/s	5.41 m/s	91.38 %	282 min
DRN	0.84 m/s	1.18 m/s	1.61 m/s	0.03 m/s	5.05 m/s	91.49 %	399 min
BQN	0.84 m/s	1.18 m/s	1.61 m/s	0.00 m/s	4.94 m/s	90.42 %	387 min
HEN	0.86 m/s	1.21 m/s	1.63 m/s	-0.04 m/s	5.07 m/s	90.23 %	321 min