Probabilistic methods in weather forecasting

Sándor Baran

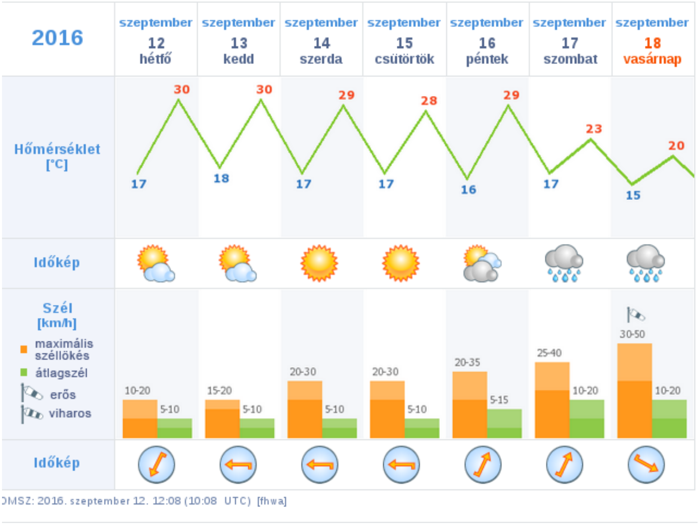
Faculty of Informatics, University of Debrecen

13 September 2016

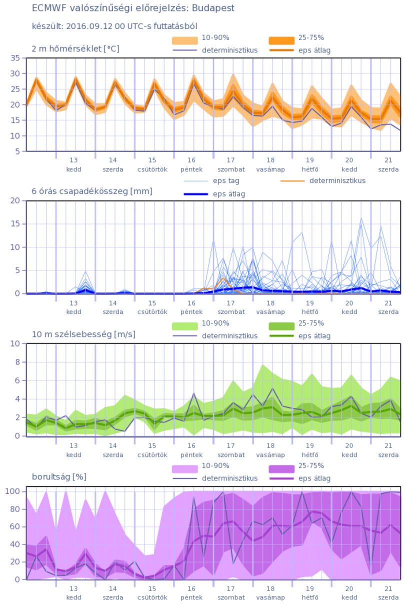
Outline

Point forecasts

Budapest



Ensemble forecasts



Ensemble of forecasts: forecasts obtained from different runs of a numerical weather prediction model. Initial conditions or model physics are changed.

Mesoscale models:

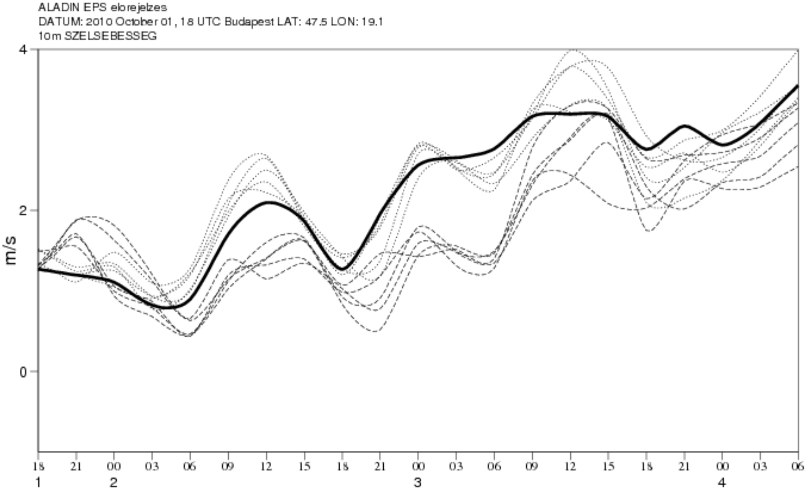
10 – 1000 km scales.

Short range prediction: 0 – 72 h.

Examples:

- ECMWF (European Centre for Medium-Range Weather Forecasting) ensemble: 51 members.
- Cosmo-DE ensemble (Deutscher Wetterdienst): 30 members.
- UWME (University of Washington mesoscale ensemble): 8 members.
- ALADIN-HUNEPS ensemble (Hungarian Meteorological Service): 11 members.

ALADIN-HUNEPS ensemble



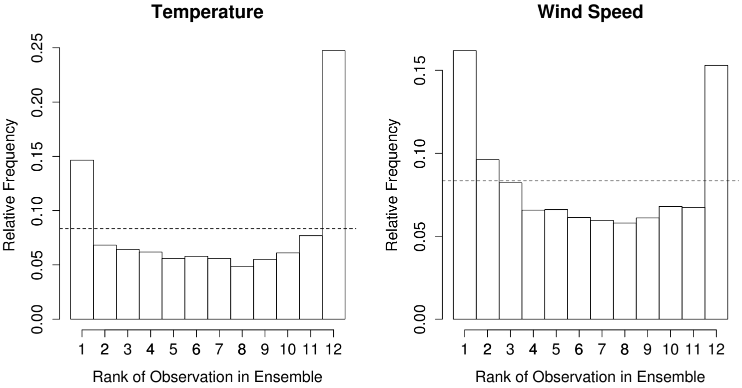
Plume diagram of ensemble forecast of 10m wind speed for Budapest initialized at 18 UTC, 01.10.2010.

Probabilistic weather forecasting

Quantities calculated from ensemble forecasts:

- Point forecasts, e.g., ensemble mean, ensemble median.
- Dispersion, e.g., standard deviation.
- Probabilities of events, e.g., 10m wind speed is greater than 12 m/s.

Problem: ensembles are often under-dispersive, uncalibrated.



Verification rank histograms of ALADIN-HUNEPS ensemble forecasts. Period: 01.04.2012 – 31.03.2013.

Statistical post-processing

Aim: find the distribution of the future weather quantity usually with the help of ensembles and verifying observations.

Predictive PDFs should be

- Calibrated: e.g., around 90% of the verifying observations should be contained between the lower and upper 5% quantiles of the predictive PDF.
- Sharp: e.g., the 90% central prediction intervals are narrower on average than the classical prediction intervals based on raw ensembles.

Most popular post-processing methods:

- Bayesian Model Averaging[1] (BMA); `R` package `ensembleBMA`.
- Non-homogeneous regression or Ensemble Model Output Statistics[2] (EMOS); `R` package `ensembleMOS`.

[1] Raftery, A. E., Gneiting, T., Balabdaoui, F. and Polakowski, M. (2005) Using Bayesian model averaging to calibrate forecast ensembles. *Mon. Wea. Rev.* **133**, 1155–1174.

[2] Gneiting, T., Raftery, A. E., Westveld, A. H. and Goldman, T. (2005) Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon. Wea. Rev.* **133**, 1098–1118.

Basic BMA model

X: weather quantity (vector) of interest (temperature, wind speed, precipitation accumulation, wind vector).

$f_1, f_2, \ldots, f_K$: ensemble of forecasts for X.

$g_k(x \,|\, f_k; \theta_k)$: conditional PDF of X given f_k is the best forecast.

θ_k: parameter to be estimated with the help of training data.

Training data: a sliding window containing ensemble members and verifying observations for the preceding n time points.

BMA predictive model

$$p\big(x \,\big|\, f_1, \ldots, f_K; \theta_1, \ldots, \theta_K\big) = \sum_{k=1}^{K} \omega_k g_k(x \,|\, f_k; \theta_k).$$

ω_k: posterior probability of forecast k being the best one based on the performance of the forecast in the training period.

$\omega_k \geq 0, \; k = 1, 2, \ldots, K, \; \omega_1 + \omega_2 + \ldots + \omega_K = 1.$

Exchangeable ensemble members

M ensemble members divided into K exchangeable groups.

Ensemble: $f_{k,\ell}$, $k = 1, 2, \ldots, K$, $\ell = 1, 2, \ldots M_k$. $M_1 + M_2 + \ldots + M_K = M$.

$$p\big(x \,\big|\, f_{k,\ell}, k = 1, \ldots, K,\ \ell = 1, \ldots M_k; \theta_1, \ldots, \theta_K\big) = \sum_{k=1}^{K} \sum_{\ell=1}^{M_k} \omega_k g_k(x \,|\, f_{k,\ell}; \theta_k).$$

Example. ALADIN-HUNEPS System of the HMS, Continental Europe, 8 km grid. Dynamical downscaling of the global ARPEGE based PEARP system of Météo-France. 10+1 ensemble members. 10 forecasts from perturbed initial conditions, one control from unperturbed analysis.

42h predictions for 10m wind speed (m/s), 10 stations, $K = 2$, $M_1 = 1$, $M_2 = 10$.

"control"	"lam01"	"lam02"	...	"lam10"	"observation"	"date"	"station"	"latitude"	"longitude"
2.96837	1.61277	3.25756	...	3.79590	3.10	2012040112	12772	48.10	20.80
5.00778	6.08366	6.99126	...	4.37388	6.10	2012040112	12812	47.20	16.60
7.34007	6.41086	8.73414	...	5.88894	4.10	2012040112	12822	47.70	17.60
7.14338	5.52808	8.07670	...	6.45395	6.90	2012040112	12843	47.50	19.10
4.61355	1.30060	5.87290	...	3.85628	3.40	2012040112	12882	47.50	21.60
5.89391	2.94215	6.57465	...	4.91185	3.50	2012040112	12892	47.90	21.60
5.35703	4.48092	6.36201	...	5.00581	4.20	2012040112	12925	46.50	17.00
5.63003	2.65106	6.92411	...	4.83982	6.50	2012040112	12942	46.10	18.20
7.01321	4.59867	9.12443	...	7.19802	8.00	2012040112	12970	46.90	19.70
5.83536	3.47728	7.63234	...	5.72420	8.10	2012040112	12982	46.30	20.20
5.03390	4.98639	4.56233	...	5.14226	7.00	2012040212	12772	48.10	20.80

Temperature and pressure

X: temperature (K) or pressure (mb,hPa).

$f_1, f_2, \ldots, f_K$: ensemble of forecasts for X.

BMA predictive model for X:

$$p\big(x \,\big|\, f_1, \ldots, f_K\big) = \sum_{k=1}^{K} \omega_k g_k(x \,|\, f_k).$$

$g_k(x \,|\, f_k)$: PDF of normal distribution. Mean: $b_{k,0} + b_{k,1} f_k$; variance: σ^2.

Parameters to be estimated: $b_{k,0},\ b_{k,1},\ \omega_k,\ k = 1, 2, \ldots, K,$ and σ^2.

Estimation of $b_{0k},\ b_{1k},\ k = 1, 2, \ldots, K$: linear regression using training data.

- Dependent variable: validating observation.
- Independent variable: f_k.

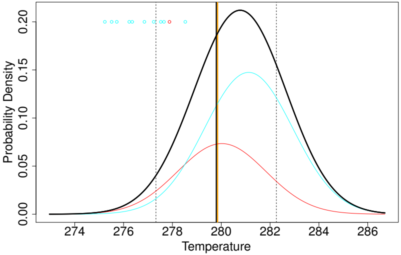
Estimation of $\omega_1, \ldots, \omega_K$ and σ^2: ML with EM algorithm.

Both expectation (E) and maximization (M) steps lead to closed formulae.

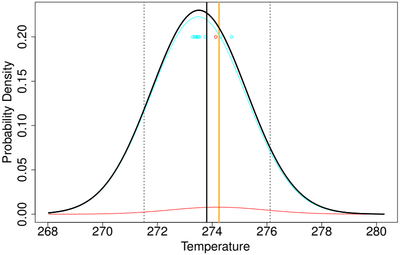
Raftery, A. E., Gneiting, T., Balabdaoui, F. and Polakowski, M. (2005) Using Bayesian model averaging to calibrate forecast ensembles. *Mon. Wea. Rev.* **133**, 1155–1174.

Example

ALADIN-HUNEPS forecasts for surface (2m) temperature.



Debrecen, 24.10.2010



Debrecen, 21.12.2010

PDFs: black: overall; red: control; light blue: exchangeable.

Vertical lines: black: median forecast; orange: verifying observation; dashed: deciles.

Circles: red: control; light blue: exchangeable ensemble members.

24.10.2010: observation: 279.85 K; ens. median: 276.85 K; BMA median: 279.80 K

21.12.2010: observation: 274.25 K; ens. median: 273.50 K; BMA median: 273.80 K

Gamma BMA model for wind speed

X: maximal or instantaneous wind speed (m/s).

$g_k(x \,|\, f_k)$: PDF of a Gamma distribution.

$$g_k(x \,|\, f_k) = \frac{1}{\beta_k^{\alpha_k}\Gamma(\alpha_k)} x^{\alpha_k - 1} \exp\big(-x/\beta_k\big) \qquad \text{if} \ \ x \geq 0.$$

Mean: $\mu_k = \alpha_k \beta_k$; variance: $\sigma_k^2 = \alpha_k \beta_k^2$.

$$\mu_k = b_{k,0} + b_{k,1} f_k, \qquad \sigma_k = c_0 + c_1 f_k.$$

Parameters to be estimated: $b_{k,0},\ b_{k,1},\ \omega_k,\ k = 1, 2, \ldots, K,$ and $c_0,\ c_1$.

Estimation of $b_{k,0},\ b_{k,1},\ k = 1, 2, \ldots, K$: linear regression using training data.

Estimation of $\omega_1, \ldots, \omega_K$ and $c_0,\ c_1$: ML with EM algorithm.

Maximization (M) step requires numerical optimization.

Sloughter, J. M., Gneiting, T. and Raftery, A. E. (2010) Probabilistic wind speed forecasting using ensembles and Bayesian model averaging. *J. Amer. Stat. Assoc.* **105**, 25–35.

Truncated normal BMA model for wind speed

X: maximal or instantaneous wind speed (m/s).

$g_k(x \,|\, f_k)$: PDF of a truncated normal (TN) distribution $\mathcal{N}_0\big(\mu, \sigma^2\big)$ with cut-off at zero.

$$g_k(x \,|\, f_k) := \frac{\frac{1}{\sigma}\varphi\big((x-\mu_k)/\sigma\big)}{\Phi\big(\mu_k/\sigma\big)}, \quad \text{if } \ x \geq 0. \qquad \mu_k = \alpha_k + \beta_k f_k.$$

μ_k: location parameter; σ: scale parameter.

Mean and variance of $\mathcal{N}_0\big(\mu, \sigma^2\big)$:

$$\kappa_k = \mu_k + \frac{\sigma\varphi\big(\mu_k/\sigma\big)}{\Phi\big(\mu_k/\sigma\big)} \quad \text{and} \quad \varrho_k^2 = \sigma^2\left(1 - \frac{\mu_k}{\sigma}\frac{\varphi\big(\mu_k/\sigma\big)}{\Phi\big(\mu_k/\sigma\big)} - \left(\frac{\varphi\big(\mu_k/\sigma\big)}{\Phi\big(\mu_k)/\sigma\big)}\right)^2\right).$$

Parameters to be estimated: $\alpha_k, \ \beta_k, \ \omega_k, \ k = 1, 2, \ldots, K$, and σ.

All parameters are estimated using ML, the likelihood function is maximized with the help of EM algorithm.

Closed formulae. 2 - 3 times faster than Gamma BMA.

Baran, S. (2014) Probabilistic wind speed forecasting using Bayesian model averaging with truncated normal components. *Comput. Stat. Data. Anal.* **75**, 227–238.

Further BMA models

Precipitation accumulation:

- Discrete-continuous model. Point mass at zero, gamma distribution for modelling positive precipitation accumulation.

  Sloughter, J. M., Raftery, A. E., Gneiting, T. and Fraley, C. (2007) Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *Mon. Wea. Rev.* **135**, 3209–3220.

Wind direction:

- Von-Mises distribution.

  Bao, L., Gneiting, T., Raftery, A. E., Grimit, E. P. and Guttorp, P. (2010) Bias correction and Bayesian model averaging for ensemble forecasts of surface wind direction. *Mon. Wea. Rev.* **138**, 1811–1821.

Wind vector:

- Bivariate normal distribution.

  Sloughter, J. M., Gneiting, T. and Raftery, A. E. (2013) Probabilistic wind vector forecasting using ensembles and Bayesian model averaging. *Mon. Wea. Rev.* **141**, 2107–2119.

Wind speed and temperature:

- Bivariate normal distribution truncated from below at zero in the wind coordinate.

  Baran, S. and Möller, A. (2015) Joint probabilistic forecasting of wind speed and temperature using Bayesian model averaging. *Environmetrics* **26**, 120–132.

Scoring rules

P: predictive distribution.

x: observation of a random quantity X.

Scoring rule: a function $S(P, x)$ that assigns a numerical score to the forecast-observation pair (P, x).

We consider negatively oriented scoring rules: the smaller the better.

Proper scoring rule: for all predictive distributions P, Q

$$\int S(P,x)\,\mathrm{d}P(x) =: \mathrm{E}_P\left[S(P,X)\right] \leq \mathrm{E}_P\left[S(Q,X)\right] := \int S(Q,x)\,\mathrm{d}P(x). \quad (1)$$

Strictly proper scoring rule: equality in (1) holds iff P coincides with Q.

Example. Logarithmic score for a predictive PDF p:

$$\mathrm{LogS}(p,x) := -\log\left[p(x)\right].$$

Strictly proper scoring rule with

$$\mathrm{E}_p\left[\mathrm{LogS}(p,X)\right] := -\int \log\left[p(x)\right] p(x)\,\mathrm{d}x \qquad \text{(entropy of } p\text{)}.$$

Gneiting, T. and Raftery, A. E. (2007) Strictly proper scoring rules, prediction and estimation. *J. Amer. Statist. Assoc.* **102**, 359–378.

Continuous ranked probability score (CRPS) of a predictive CDF P at x:

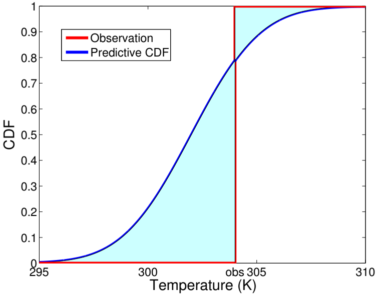
$$\mathrm{crps}\,(P,x) = \int_{-\infty}^{\infty} \big(P(y) - \mathbb{1}_{\{y\geq x\}}\big)^2 \mathrm{d}y = \int_{-\infty}^{x} P^2(y)\mathrm{d}y + \int_{x}^{\infty} \big(1 - P(y)\big)^2 \mathrm{d}y.$$

X, X': independent random variables with CDF P and finite first moment.

$$\mathrm{crps}\,(P,x) = \mathsf{E}_P|X - x| - \frac{1}{2}\mathsf{E}_P|X - X'|.$$

Properties of the CRPS

- Strictly proper scoring rule.
- Reported in the same unit as the observation.
- In the case of a point forecast it reduces to the absolute error.



Example:

$$\mathrm{crps}\left(\mathcal{N}(\mu,\sigma^2), x\right) = \sigma\left[\frac{x-\mu}{\sigma}\Big(2\Phi\big((x-\mu)/\sigma\big) - 1\Big) + 2\varphi\big((x-\mu)/\sigma\big) - \frac{1}{\sqrt{\pi}}\right].$$

Verification scores

$x_{s,t}$: verifying observation of the weather quantity at location s and time t.

$p_{s,t}(x)$, $P_{s,t}(x)$: estimated predictive PDF and CDF at location s and time t.

$\widehat{x}_{s,t}$: point forecast based on $p_{s,t}(x)$ or on the ensemble (mean, median).

n: total number of forecast cases.

Deterministic forecasts:

MAE: $\frac{1}{n}\sum_{s,t}\left|x_{s,t}-\widehat{x}_{s,t}\right|$;   RMSE: $\sqrt{\frac{1}{n}\sum_{s,t}\left(x_{s,t}-\widehat{x}_{s,t}\right)^2}$.

MAE is optimal for median, RMSE for mean.

Probabilistic forecasts:

Mean CRPS: $\frac{1}{n}\sum_{s,t}\mathrm{crps}\left(P_{s,t},x_{s,t}\right)$, $\mathrm{crps}\left(P,x\right)=\int_{-\infty}^{\infty}\left(P(y)-\mathbb{1}_{\{y\geq x\}}\right)^2\mathrm{d}y$.

Mean log-score: $-\frac{1}{n}\sum_{s,t}\log p_{s,t}\left(x_{s,t}\right)$.

Coverage: percentage of observations being in, e.g., 90 % prediction interval.

Average width of central prediction interval.

Ensemble forecasts:

Mean CRPS: empirical CDF of the ensemble should be considered.

Coverage and average width of central prediction interval.

Ensemble Model Output Statistics (EMOS)

Predictive PDF is a single parametric density where the parameters are functions of the ensemble. Parameters of these functions are estimated by optimizing the average value of some verification score over the training data.

Temperature and pressure

X: temperature (K) or sea level pressure (mb, hPa).

Predictive distribution for X:

$$\mathcal{N}\big(a_0 + a_1 f_1 + \cdots + a_K f_K, b_0 + b_1 S^2\big) \qquad \text{with} \qquad S^2 := \frac{1}{K-1}\sum_{k=1}^{K}\big(f_k - \overline{f}\big)^2.$$

Parameters to be estimated: $a_0, \ a_1, \ldots, a_K \in \mathbb{R}$ and $b_0, b_1 \geq 0$.

Minimum points of the mean value of an appropriate verification score over the training data. Usually mean CRPS or mean log-score. CRPS can be given in a closed form.

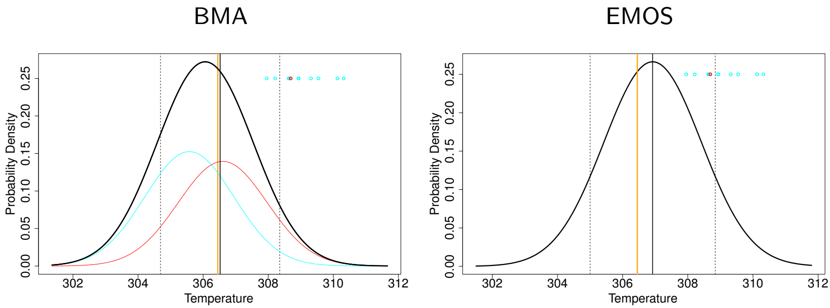
Exchangeable ensemble members:

$$\mathcal{N}\Bigg(a_0 + a_1 \sum_{\ell_1=1}^{M_1} f_{1,\ell_1} + \cdots + a_m \sum_{\ell_K=1}^{M_K} f_{K,\ell_K}, b_0 + b_1 S^2\Bigg).$$

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

Example

ALADIN-HUNEPS forecasts for surface temperature, Debrecen, 02.07.2012.



PDFs: black: overall; red: control; light blue: exchangeable.

Vertical lines: black: median forecast; orange: validating observation; dashed: deciles.

Circles: red: control; light blue: exchangeable ensemble members.

Validating observation: 306.45 K.

Ensemble median: 308.93 K; BMA median: 306.52 K; EMOS median: 306.92 K.

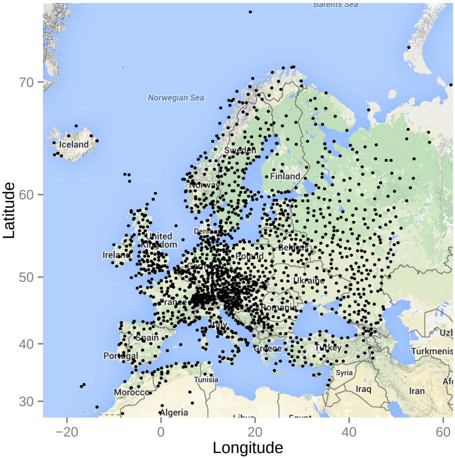
Training data

Regional EMOS: parameters are estimated using all available forecast cases from the training period. A single universal set of parameters across the entire ensemble domain. Short training periods.

Local EMOS: distinct parameter estimates for the different stations using only the training data of the given station.

Semilocal EMOS: parameters for a given station are estimated using training data of similar stations[3].

- Distance-based similarity.
- Clustering-based similarity.



Domain of the GLAMEPS ensemble

[3] Lerch, S., Baran, S. (2016) Similarity-based semilocal estimation of EMOS models. *J. R. Stat. Soc. Ser. C Appl. Stat.*, doi:10.1111/rssc.12153.

Distance based similarity

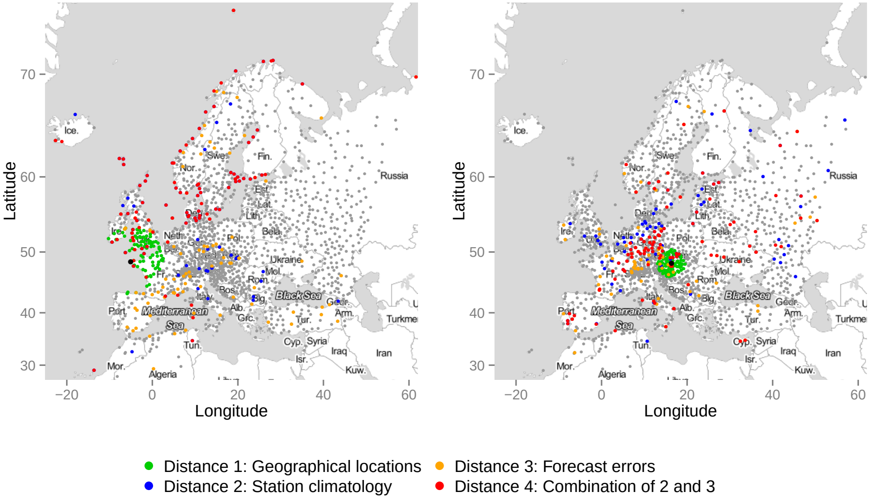


Illustration of the 100 most similar stations measured by the four distance functions for two reference stations at Ouessant, France (*left*) and Vienna, Austria (*right*).

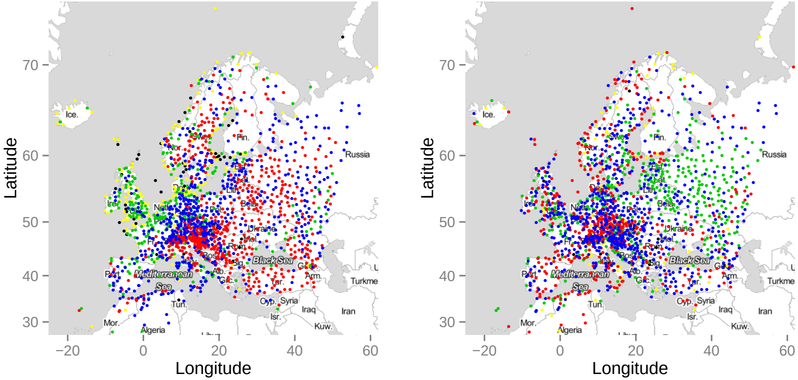
Clustering

k-means clustering, each station is characterized by an N-dimensional vector.

1. Station climatology. Equidistant $\frac{1}{N+1}, \frac{2}{N+1}, \ldots, \frac{N}{N+1}$ quantiles of the empirical CDF of observations over the training period.

2. Forecast errors. Equidistant quantiles of the empirical CDF of the forecast errors of the ensemble mean.

3. Combination of feature sets 1 and 2. 50 - 50 %.



Clustering based on feature sets 1. (*left*) and 2. (*right*). 5 clusters, 24 features.

Truncated normal EMOS model for wind speed

Predictive distribution for wind speed:

$$\mathcal{N}_0\big(a_0 + a_1 f_1 + \cdots + a_K f_K, b_0 + b_1 S^2\big) \qquad \text{with} \qquad S^2 := \frac{1}{K-1}\sum_{k=1}^{K}\big(f_k - \overline{f}\big)^2.$$

$\mathcal{N}_0\big(\mu, \sigma^2\big)$: TN distribution with cut-off at zero; location: μ; scale: σ.

Parameters to be estimated: $a_0 \in \mathbb{R}$, $a_1, \ldots, a_K \geq 0$ (loc.) and $b_0, b_1 \geq 0$ (scale).

Minimum points of the mean value of an appropriate verification score over the training data. Usually mean CRPS or mean log-score. CRPS can be given in a closed form.

Exchangeable ensemble members:

$$\mathcal{N}_0\Bigg(a_0 + a_1 \sum_{\ell_1=1}^{M_1} f_{1,\ell_1} + \cdots + a_m \sum_{\ell_K=1}^{M_K} f_{K,\ell_K}, b_0 + b_1 S^2\Bigg).$$

Thorarinsdottir, T. L. and Gneiting, T. (2010) Probabilistic forecasts of wind speed: ensemble model output statistics by using heteroscedastic censored regression. *J. Roy. Statist. Soc. Ser. A* **173**, 371–388.

Log-normal EMOS model

$\mathcal{LN}(\mu,\sigma)$: log-normal distribution (LN) with location μ and scale σ.

Probability density function:

$$h(x|\,\mu,\sigma) := \frac{1}{x\sigma}\varphi\big((\log x-\mu)/\sigma\big), \qquad x \geq 0.$$

Mean: $m = \mathrm{e}^{\mu+\sigma^2/2}$; variance: $v^2 = \mathrm{e}^{2\mu+\sigma^2}\big(\mathrm{e}^{\sigma^2}-1\big)$.

Reparametrization: $\mu = \log\Big(m^2/\sqrt{v^2+m^2}\Big)$, $\sigma = \sqrt{\log\big(1+v^2/m^2\big)}$.

Log-normal EMOS model:

$$m = \alpha_0 + \alpha_1 f_1 + \cdots + \alpha_K f_K, \qquad v^2 = \beta_0 + \beta_1 S^2.$$

Parameters to be estimated: $\alpha_0 \in \mathbb{R}$, $\alpha_1,\ldots,\alpha_K \geq 0$ and $\beta_0,\beta_1 \geq 0$.

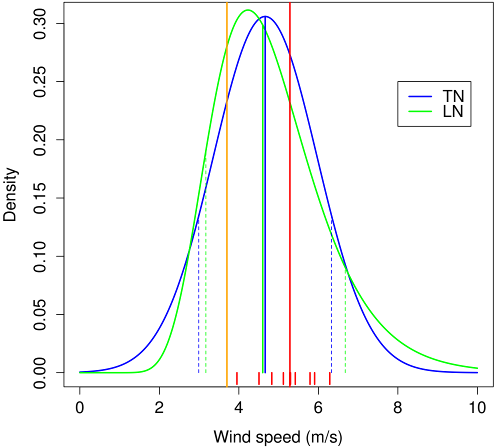
Minimum points of the mean CRPS or mean log-score. CRPS can be given in a closed form.

Exchangeable ensemble members:

$$m = \alpha_0 + \alpha_1 \sum_{\ell_1=1}^{M_1} f_{1,\ell_1} + \cdots + \alpha_K \sum_{\ell_K=1}^{M_K} f_{K,\ell_K}, \qquad v^2 = \beta_0 + \beta_1 S^2.$$

Example

Log-normal distribution has heavier right tail than truncated normal, better fit for high wind speed values.



Debrecen, 12.12.2012.

Observation (orange): 3.7 m/s

Ensemble median (red): 5.28 m/s

TN EMOS median (blue): 4.66 m/s

LN EMOS median (green): 4.60 m/s

Dashed lines: 80 % central prediction intervals.

Regime-switching model

Model choice depends on the value of ensemble median f_{med} compared to a threshold θ.

Predictive distribution:

$$\begin{cases} \mathcal{N}_0\big(\mu_{TN}, \sigma_{TN}^2\big), & \text{if } f_{med} < \theta; \\ \mathcal{LN}\big(\mu_{LN}, \sigma_{LN}\big), & \text{if } f_{med} \geq \theta. \end{cases}$$

$$\mu_{TN} = a_0 + a_1 f_1 + \cdots + a_K f_K, \qquad \sigma_{TN}^2 = b_0 + b_1 S^2;$$

$$\mu_{LN} = \log\left(m^2 / \sqrt{v^2 + m^2}\right), \qquad \sigma_{LN} = \sqrt{\log\big(1 + v^2/m^2\big)},$$

$$m = \alpha_0 + \alpha_1 f_1 + \cdots + \alpha_K f_K, \qquad v^2 = \beta_0 + \beta_1 S^2.$$

Large data set (e.g., UWME, ECMWF): parameters μ_{LN} and σ_{LN} $\big(\mu_{LN}$ and $\sigma_{LN}\big)$ are estimated from training data with $f_{med} \geq \theta$ ($f_{med} < \theta$).

Small data set (e.g., ALADIN-HUNEPS): parameters of both models are estimated from the same training data. Only the model choice is based on f_{med}.

Problem: static threshold value, reduced flexibility.

Baran, S. and Lerch, S. (2015) Log-normal distribution based EMOS models for probabilistic wind speed forecasting. *Q. J. R. Meteorol. Soc.* **141**, 2289–2299.

Mixture model

$g(x|\,\mu_{TN},\sigma_{TN})$: PDF of $\mathcal{N}_0\big(\mu_{TN},\sigma_{TN}^2\big)$.

$h(x|\,\mu_{LN},\sigma_{LN})$: PDF of $\mathcal{LN}\big(\mu_{LN},\sigma_{LN}\big)$.

$$\mu_{TN} = a_0 + a_1 f_1 + \cdots + a_K f_K, \qquad \sigma_{TN}^2 = b_0 + b_1 S^2;$$
$$\mu_{LN} = \log\Big(m^2/\sqrt{v^2+m^2}\Big), \qquad \sigma_{LN} = \sqrt{\log\big(1+v^2/m^2\big)},$$
$$m = \alpha_0 + \alpha_1 f_1 + \cdots + \alpha_K f_K, \qquad v^2 = \beta_0 + \beta_1 S^2.$$

TN-LN mixture EMOS predictive PDF:

$$\psi(x|\,\mu_{TN},\sigma_{TN};\mu_{LN},\sigma_{LN};\omega) := \omega g(x|\,\mu_{TN},\sigma_{TN}) + (1-\omega)h(x|\,\mu_{LN},\sigma_{LN}), \quad \omega\in[0,1].$$

Parameters of the TN and LN components and weight ω are estimated

- by minimizing the mean CRPS (no closed form, numerical integrals);
- by minimizing the mean log-score (maximum likelihood).

Baran, S. and Lerch, S. (2016) Mixture EMOS model for calibrating ensemble forecasts of wind speed. *Environmetrics* **27**, 116–130.

Generalized extreme value EMOS model

$\mathcal{GEV}(\mu, \sigma, \xi)$: generalized extreme value (GEV) distribution with location μ, scale σ and shape ξ. CDF:

$$G(x) := \begin{cases} \exp\Big(-\big[1+\xi(\frac{x-\mu}{\sigma})\big]^{-1/\xi}\Big), & \xi \neq 0, \\ \exp\Big(-\exp\big(-\frac{x-\mu}{\sigma}\big)\Big), & \xi = 0, \end{cases} \qquad 1+\xi(x-\mu)/\sigma > 0.$$

GEV EMOS model: $\mu = \gamma_0 + \gamma_1 f_1 + \cdots + \gamma_K f_K, \qquad \sigma = \sigma_0 + \sigma_1 \overline{f}.$

Parameters to be estimated: $\gamma_0,\ \gamma_1, \ldots, \gamma_K,\ \sigma_0, \sigma_1\ (\sigma > 0)$ and $\xi \geq 0$.

Closed form for the CRPS. ML method is more stable.

Exchangeable ensemble members:

$$\mu = \gamma_0 + \gamma_1 \sum_{\ell_1=1}^{M_1} f_{1,\ell_1} + \cdots + \gamma_K \sum_{\ell_K=1}^{M_K} f_{K,\ell_K}, \qquad \sigma = \sigma_0 + \sigma_1 \overline{f}.$$

TN-GEV mixture model based on f_{med} .

Problem: GEV distribution can assign mass to negative values!

Lerch, S. and Thorarinsdottir, T. L. (2013) Comparison of non-homogeneous regression models for probabilistic wind speed forecasting. *Tellus A* **65**, 21206.

Further EMOS models

Precipitation accumulation:

- Censored generalized extreme value distribution.

  Scheuerer, M. (2014) Probabilistic quantitative precipitation forecasting using ensemble model output statistics. *Q. J. R. Meteorol. Soc.* **149**, 1086–1096.

- Censored and shifted gamma distribution.

  Baran, S. and Nemoda, D. (2016) Censored and shifted gamma distribution based EMOS model for probabilistic quantitative precipitation forecasting. *Environmetrics* **27**, 280–292.

Wind vector:

- Bivariate normal distribution. In operational use at DWD.

  Schuhen, N., Thorarinsdottir, T. L. and Gneiting, T. (2012) Ensemble model output statistics for wind vectors. *Mon. Wea. Rev.* **140**, 3204–3219.

Wind speed and temperature:

- Bivariate normal distribution truncated from below at zero in the wind coordinate.

  Baran, S. and Möller, A. (2016) Bivariate ensemble model output statistics approach for joint forecasting of wind speed and temperature. *Meteorol. Atmos. Phys.*, doi:10.1007/ s00703-016-0467-8.

ALADIN-HUNEPS ensemble

Wind speed data of 10 major cities (Miskolc, Szombathely, Győr, Budapest, Debrecen, Nyíregyháza, Nagykanizsa, Pécs, Kecskemét, Szeged) of Hungary. Source: HMS.

Data: 11 member ensembles of 42h forecasts of surface (10m) instantaneous wind speed produced by the ALADIN-HUNEPS system of the HMS and the corresponding validating observations.

Prediction interval: 42 hours.

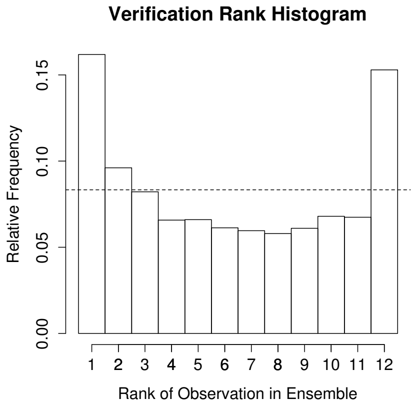
Ensemble: 10 exchangeable and one control forecast initialized at 18 UTC (give forecasts for 12 UTC two days later).

Period: 01.04.2012 – 31.03.2013.

Missing data: 6 days when no forecasts are available (excluded).

Horányi, A., Mile, M. and Szűcs, M. (2011) Latest developments around the ALADIN operational short-range ensemble prediction system in Hungary. *Tellus A* **63**, 642–651.

Raw ensemble



Ensemble range contains the observed wind speed only in 61.21 % of the cases.

Under-dispersive and uncalibrated. Nominal coverage: 83.33 %.

BMA models for the ALADIN-HUNEPS ensemble

Predictive PDF of 10m wind speed X:

$$p(x|\, f_c, f_{\ell,1}, \ldots, f_{\ell,10}) = \omega g_c(x|f_c) + \sum_{j=1}^{10} \frac{1-\omega}{10} g_\ell(x|f_{\ell,j}).$$

- f_c: control member of the ensemble.
- $f_{\ell,1}, \ldots, f_{\ell,10}$: exchangeable ensemble members.
- $\omega \in [0,1]$: weight of the control member.
- $g_c(x|f),\ g_\ell(x|f)$: gamma or truncated normal PDFs.

Gamma BMA

Mean: $b_0 + b_1 f$; standard deviation: $c_0 + c_1 f$.

Parameters to be estimated: $b_0, b_1,\ \omega$ and c_0, c_1.

Truncated normal BMA

Locations: $\alpha_c + \beta_c f$ and $\alpha_\ell + \beta_\ell f$; scale: σ.

Parameters to be estimated: $\alpha_c, \beta_c,\ \alpha_\ell, \beta_\ell$ and $\omega,\ \sigma$.

EMOS models for the ALADIN-HUNEPS ensemble

Truncated normal EMOS

$\mathcal{N}_0\big(\mu_{TN}, \sigma^2_{TN}\big)$, where $\mu_{TN} = a + a_c f_c + a_\ell \sum_{j=1}^{10} f_{\ell,j}$, $\sigma^2_{TN} = b_0 + b_1 S^2$.

Parameters to be estimated: a, a_c, a_ℓ (location) and b_0, b_1 (scale).

Log-normal EMOS

$\mathcal{LN}\big(\mu_{LN}, \sigma_{LN}\big)$, where $\mu_{LN} = \log\big(m^2/\sqrt{v^2+m^2}\big)$, $\sigma_{LN} = \sqrt{\log\big(1+v^2/m^2\big)}$,

$$m = \alpha + \alpha_c f_c + \alpha_\ell \sum_{j=1}^{10} f_{\ell,j}, \qquad v^2 = \beta_0 + \beta_1 S^2.$$

Parameters to be estimated: $\alpha, \alpha_c, \alpha_\ell$ (mean) and β_0, β_1 (variance).

Generalized extreme value EMOS

$\mathcal{GEV}\big(\mu_{GEV}, \sigma_{GEV}, \xi\big)$, where $\mu_{GEV} = \gamma + \gamma_c f_c + \gamma_\ell \sum_{j=1}^{10} f_{\ell,j}$, $\sigma^2_{GEV} = \sigma_0 + \sigma_1 \overline{f}$.

Parameters to be estimated: $\gamma, \gamma_c, \gamma_\ell$ (loc.), σ_0, σ_1 (scale) and ξ (shape).

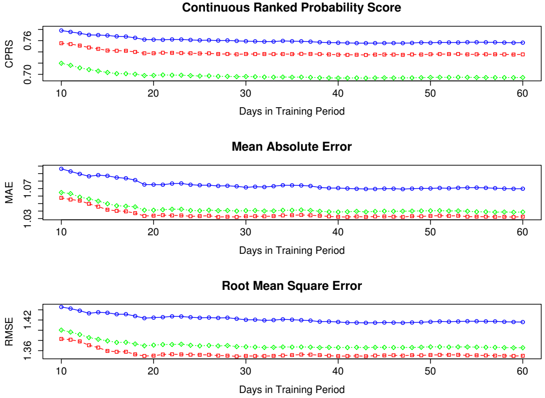
TN-LN and TN-GEV regime-switching EMOS

Parameters of TN, LN and GEV distributions are estimated using the same training data, only the model choice is based on the ensemble median.

TN-LN mixture EMOS

Parameter estimates optimize the mean CRPS or mean log-score.

Optimal training period length



Verification period: 01.06.2012 – 31.03.2013 (298 days)

A **43** days training period seems reasonable.

Gamma BMA

Minima of CRPS and RMSE at day 43. Minimum of MAE at day 47. Value at day 43 is practically the same.

Truncated normal BMA

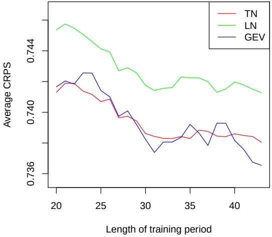
Minimum of CRPS at day 43. Minima of MAE and RMSE at day 59, values at day 43 are practically the same.

Truncated normal EMOS

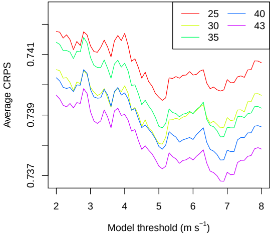
Minimum of CRPS at day 43. Minima of MAE and RMSE at days 59 and 29, resp. Values at day 43 are the same.

Baran, S., Horányi, A. and Nemoda, D. (2014) Comparison of the BMA and EMOS statistical methods in calibrating temperature and wind speed forecast ensembles. *Időjárás* **118**, 217–241.

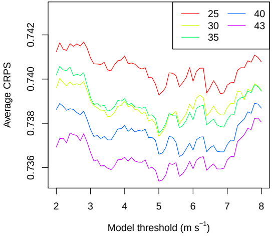
Studies based on CRPS, MAE and RMSE scores of BMA and TN EMOS models: optimal training period length is 43 days.



CRPS vs. training period length



CRPS vs. threshold, TN-LN



CRPS vs. threshold, TN-GEV

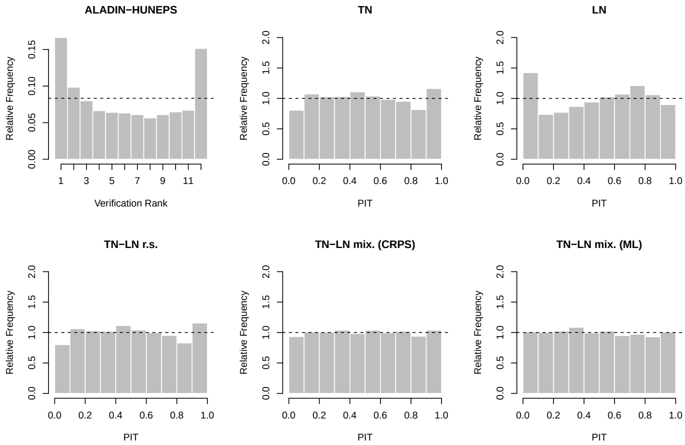
43 days training period length can be kept. BMA and EMOS predictive PDFs and validating observations for 313 calendar days (3130 forecast cases).

Optimal threshold values: TN-LN $\theta = 6.9\,\mathrm{m/s}$; TN-GEV $\theta = 5.0\,\mathrm{m/s}$.

LN proportion in TN-LN: 4 %; GEV proportion in TN-GEV: 15 %.

Baran, S. and Lerch, S. (2015) Log-normal distribution based EMOS models for probabilistic wind speed forecasting. *Q. J. R. Meteorol. Soc.* **141**, 2289–2299.

PIT histograms



Bootstrap estimates of rejection rates of the α^0_{1234} test of uniformity based on 10 000 random samples of size 2 500 each at the 0.05 level.

TN	LN	TN-LN r.-s.	TN-LN mix. (CRPS)	TN-LN mix. (ML)
100 %	100 %	100 %	0 %	1 %

Verification scores

Mean CRPS of probabilistic forecasts, MAE (median) and RMSE (mean) of point forecasts (m/s), coverage (%) and average width (m/s) of 83.33 % central predictions intervals.

Predictive model	CRPS	MAE	RMSE	Coverage	Av. w.
TN	0.738	1.037	1.357	83.59	3.53
LN	0.741	1.038	1.362	80.44	3.57
TN-LN mix. (CRPS)	0.736	1.037	1.358	83.02	3.62
TN-LN mix. (ML)	0.737	1.040	1.360	83.14	3.58
TN-LN r.s., $\theta = 6.9$	0.737	1.035	1.356	83.59	3.54
GEV	0.737	1.041	1.355	81.21	3.54
TN-GEV r.s., $\theta = 5.0$	0.735	1.039	1.355	85.59	3.72
Gamma BMA	0.760	**1.075**	**1.427**	81.87	3.72
TN BMA	0.698	1.045	1.377	85.46	3.76
Ensemble	0.803	1.069	1.373	68.22	2.88
Climatology	1.046	1.481	1.922	82.54	3.43

- Differences between the CRPS values of mixture and regime-switching models are not significant (Diebold-Mariano test).
- Mean (maximal) probability of forecasting a negative wind speed: GEV model 0.33 % (9.46 %); TN-GEV model 2.74×10^{-3} % (0.15 %).

Conclusions, further directions

Conclusions

- Statistical post-processing significantly improves the calibration of probabilistic and accuracy of point forecasts.
- Mixture and regime-switching EMOS models outperform the traditional single distribution based methods.

Further directions

- Development of efficient algorithms for estimating the parameters of mixture models (ongoing research).
- Extension of the `ensembleMOS` package with wind speed and precipitation models (ongoing work).
- Investigation of semi-local versions of the existing EMOS models for ECMWF global forecasts for temperature, wind speed and precipitation.
- Development of predictive models incorporating spatial dependence.
- Development of parallel algorithms for ensemble post-processing.