Physically coherent probabilistic weather forecasts via ensemble copula coupling (ECC)

Roman Schefzik

Institute for Applied Mathematics, Heidelberg University

September 19, 2014

Joint work with Thordis Thorarinsdottir and Tilmann Gneiting

Outline

Outline

Weather prediction: Deterministic forecasting

- Until the early 1990s: Numerical weather prediction models based on a system of partial differential equations yield deterministic forecasts of future atmospheric states.
- However, two major sources of uncertainty:
  - Initial conditions: incomplete network of observations, shortcomings in data assimilation, measurement errors,...
  - Model formulation: incomplete or inaccurate numerical schemes, incomplete knowledge of physical processes,...

Weather prediction: Ensemble forecasting

- Change in weather prediction practice during the last twenty years: Employment of ensemble prediction systems
- Ensemble weather forecasts consist of multiple runs of numerical weather prediction models differing in the initial conditions and/or the parameterized representation of the atmosphere, thereby addressing the two major sources of uncertainty.
- Example: The 50-member global European Centre for Medium-Range Weather Forecasts (ECMWF) ensemble, operational since 1992

Example: ECMWF ensemble forecast

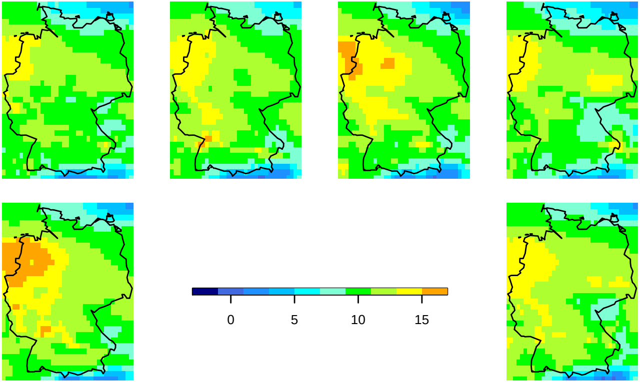


Figure: Six exemplary members of the 48 h ahead ECMWF ensemble forecast for temperature (in °C) over Germany valid on 24 April 2011, 2:00 am

Ensemble postprocessing

- Problem: Ensembles frequently reveal model biases and a lack of calibration, in that the verifying observations fall far too often outside the ensemble range.
- Statistical postprocessing of the raw output data is required, leading to full predictive distributions over future weather quantities instead of point forecasts.
- Goal: Maximize sharpness subject to calibration

Univariate ensemble postprocessing: BMA and EMOS

Two prominent **univariate** ensemble postprocessing methods for a weather quantity y based on an ensemble forecast $x_1, \ldots, x_M$:

- Bayesian model averaging (BMA; e.g. Raftery et al., 2005): Each ensemble member m is linked to a kernel function $f(y|x_m)$, using a weight w_m that reflects the member's relative skill.

$$y|x_1, \ldots, x_M \sim \sum_{m=1}^{M} w_m f(y|x_m)$$

- Ensemble model output statistics (EMOS; e.g. Gneiting et al., 2005): Fits a single parametric distribution g using summary statistics from the ensemble.

$$y|x_1, \ldots, x_M \sim g(y|x_1, \ldots, x_M)$$

Univariate ensemble postprocessing: BMA and EMOS

- The choice of the kernel and distribution, respectively, depends on the weather quantity one is interested in.

  Example: Normal distribution (temperature, pressure,...)

  - BMA:

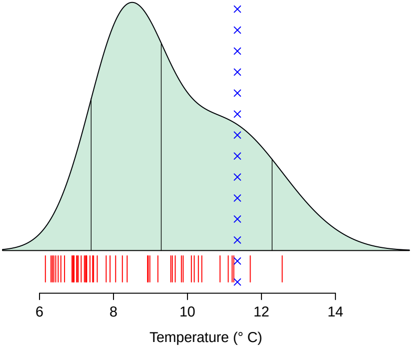
$$y | x_1, \ldots, x_M \sim \sum_{m=1}^{M} w_m \, \mathcal{N}(a_m + b_m x_m, \sigma_m^2)$$

  - EMOS:

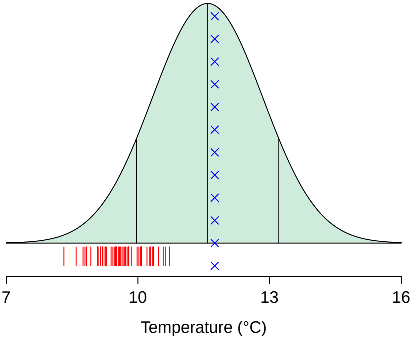
$$y | x_1, \ldots, x_M \sim \mathcal{N}(a + b_1 x_1 + \ldots + b_M x_m, c + dS^2)$$

- BMA or EMOS parameters are estimated from training data comprising forecasts and observations, using a sliding training window of typically the past 20 to 40 days.

Example: Univariate ensemble postprocessing BMA/EMOS



(a) **BMA**, 21 April 2011, 2:00 am



(b) **EMOS**, 25 April 2011, 2:00 am

Figure: 24 hour ahead predictive PDFs for temperature at Berlin. Red: 50-member ECMWF ensemble forecast, blue: verifying observation, black: 10th, 50th and 90th percentiles of the predictive distribution.

From univariate to multivariate ensemble postprocessing

- Disadvantage: These postprocessing methods frequently apply to a **single** weather variable at a **single** location for a **single** look-ahead time only.
- However, physical consistency of multivariate dependence structures across space, time and variables is required in numerous applications:
  - Air traffic management
  - Winter road maintenance
  - Ship routeing
  - ...

Parametric multivariate ensemble postprocessing

Several parametric multivariate ensemble postprocessing methods have emerged:

- Bivariate postprocessing of (u, v)- wind vectors (Pinson, 2012; Schuhen et al., 2012)
- Gaussian copula approach (Möller et al., 2013)
- Spatial EMOS for temperature (Feldmann et al., 2014)
- ...

Non-parametric multivariate ensemble postprocessing

Non-parametric alternatives:

- Schaake shuffle (Clark et al., 2004)

  $\longrightarrow$ Adoption of the rank dependence structure from specific historical observation data

- Our proposal: The **ensemble copula coupling (ECC)** approach (Schefzik et al., 2013)

Outline

Basic idea of ECC

- Multivariate postprocessing method for ensemble weather forecasts, which takes spatial, temporal and inter-variable dependencies into account

  $\longrightarrow$ Inheritance of the rank dependence structure from the unprocessed raw ensemble
- Implicit assumptions:
  - The ensemble members are exchangeable.
  - The ensemble is able to represent observed dependence structures appropriately.

The ECC method (Schefzik et al., 2013) I

1. **Raw ensemble:**
   Weather quantity $i \in \{1, \ldots, I\}$
   Location $j \in \{1, \ldots, J\}$
   Prediction horizon $k \in \{1, \ldots, K\}$
   Multi-index $\ell := (i, j, k)$ and $L := I \times J \times K$

   For each fixed ℓ, the M-member raw ensemble forecast

   $$x_1^\ell, \ldots, x_M^\ell$$

   with the order statistics $x_{(1)}^\ell \leq \ldots \leq x_{(M)}^\ell$ induces the permutation σ_ℓ of $1, \ldots, M$ given by

   $$\sigma_\ell(m) := \mathrm{rank}(x_m^\ell)$$

   for $m \in \{1, \ldots, M\}$, with ties resolved at random.

The ECC method (Schefzik et al., 2013) II

2. **Univariate postprocessing:**
   Use **univariate** state-of-the-art postprocessing methods, such as BMA or EMOS, to obtain calibrated and sharp predictive CDFs F_ℓ for each variable, location and look-ahead time **individually**.

The ECC method (Schefzik et al., 2013) III

3. **Quantization:**
   For each ℓ, draw M samples $\tilde{x}_1^\ell, \ldots, \tilde{x}_M^\ell$ from F_ℓ.

   Various possible procedures:

   - **ECC-R**: Take **R**andom samples from F_ℓ.
   - **ECC-T**: **T**ransformation approach:

     $$\tilde{x}_1^\ell := F_\ell^{-1}(S_\ell(x_1^\ell)), \ldots, \tilde{x}_M^\ell := F_\ell^{-1}(S_\ell(x_M^\ell)),$$

     where S_ℓ is a continuous CDF fitted to the raw ensemble $x_1^\ell, \ldots, x_M^\ell$
   - **ECC-Q**: Take equidistant **Q**uantiles from F_ℓ:

     $$\tilde{x}_1^\ell := F_\ell^{-1}\left(\frac{1}{M+1}\right), \ldots, \tilde{x}_M^\ell := F_\ell^{-1}\left(\frac{M}{M+1}\right).$$

The ECC method (Schefzik et al., 2013) IV

4. **Reordering:**
   For each ℓ, the ECC ensemble $\hat{x}_1^\ell, \ldots, \hat{x}_M^\ell$ is given by

$$\hat{x}_1^\ell := \tilde{x}_{(\sigma_\ell(1))}^\ell, \ldots, \hat{x}_M^\ell := \tilde{x}_{(\sigma_\ell(M))}^\ell.$$

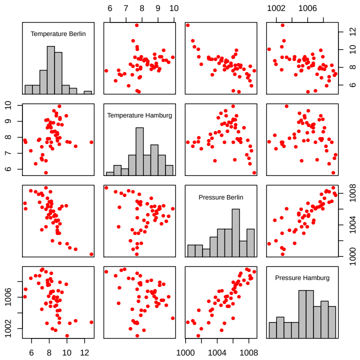
   It retains the spatial, temporal and inter-variable rank dependence structure from the raw ensemble.

ECC – a four-stage procedure (Schefzik et al., 2013)

1. Generate an M-member raw ensemble with a forecast output in $\mathbb{R}^L$.
2. Apply univariate postprocessing to get predictive distributions for each location, prediction horizon and weather quantity individually.
3. Draw a sample of size M from each predictive distribution.
4. Rearrange the sampled values in the rank order structure of the raw ensemble.

Illustration: Basic idea of ECC

24 h ahead **ECMWF ensemble** forecast of temperature and pressure at Berlin and Hamburg valid on 27 May 2010, 2:00 am, **before** and **after** individual BMA **postprocessing**



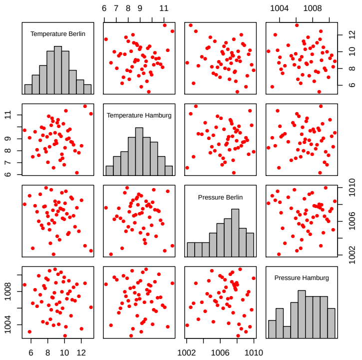
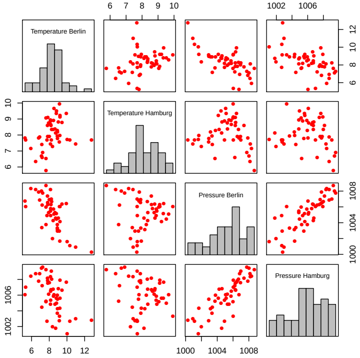
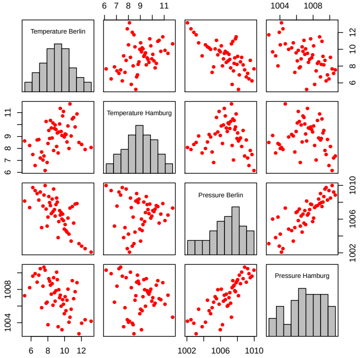


Illustration: Basic idea of ECC

24 h ahead **ECMWF ensemble** forecast of temperature and pressure at Berlin and Hamburg valid on 27 May 2010, 2:00 am, **before** and **after** individual BMA **postprocessing** and **ECC**





Benefits of ECC

- Works with any univariate postprocessing technique
- Can be employed in any dimension
- Easy to implement with negligible computational effort, once the univariate postprocessing is done
- Overarching concept for existing techniques scattered in the literature
- Theoretical backing via discrete/empirical copulas
- May be used analogously in many other settings apart from weather prediction

ECC as an empirical copula approach

- A copula $C : [0,1]^L \to [0,1]$ is a multivariate cumulative distribution function (CDF) with standard uniform margins.
- We seek a multivariate predictive CDF H with margins $F_1, \ldots, F_L$.
- Sklar's theorem (1959):

$$H(y_1, \ldots, y_L) = C(F_1(y_1), \ldots, F_L(y_L))$$

  for $y_1, \ldots, y_L \in \mathbb{R}$
- Take C to be the empirical copula

$$C : \left\{0, \frac{1}{M}, \ldots, \frac{M-1}{M}, 1\right\}^L \to \left\{0, \frac{1}{M}, \ldots, \frac{M-1}{M}, 1\right\}$$

  defined by the raw ensemble forecast, and $F_1, \ldots, F_L$ to be the empirical marginal CDFs given by the discrete samples from ECC step 3.

Empirical raw ensemble copula

The raw ensemble forecast $(x_m^\ell)_{m=1,\ldots,M}^{\ell=1,\ldots,L}$ defines an empirical copula

$$C\left(\frac{i_1}{M},\ldots,\frac{i_L}{M}\right) := \frac{1}{M}\sum_{m=1}^{M}\prod_{\ell=1}^{L}\mathbb{1}\{\text{rank}(x_m^\ell) \leq i_\ell\},$$

where $0 \leq i_\ell \leq M$ for $\ell \in \{1,\ldots,L\}$, with ties resolved at random.

By applying C to the samples from step 3, ECC inherits the multivariate rank dependence structure from the raw ensemble.

Thus, raw and ECC ensemble are associated with the same empirical copula modeling the dependence.

Case study: The ECMWF ensemble

- European Centre for Medium-Range Weather Forecasts (ECMWF) ensemble consisting of $M = 50$ members
- Forecasts available on a grid over Germany
- For the individual stations Berlin, Hamburg and Frankfurt, use bilinearly interpolated forecasts and real observations.
- On the grid, use the corresponding 0 h nowcast of the ECMWF control run (best estimate of the atmospheric state) as the ground truth instead of real observations.

Implementation and reference techniques

- Derive a univariate predictive CDF for each weather variable, location and prediction horizon individually via BMA, using a sliding training period of the past 30 days.
  $\longrightarrow$ R package `ensembleBMA`
- From each CDF, draw the $M = 50$ equidistant $\frac{1}{51}-, \ldots, \frac{50}{51}-$ **Q**uantiles as samples.
- Compare
  - Unprocessed raw ECMWF ensemble
  - Individual BMA-**Q** ensemble: quantiles randomly ordered
  - ECC-**Q** ensemble: quantiles ordered according to rank structure of the raw ensemble members

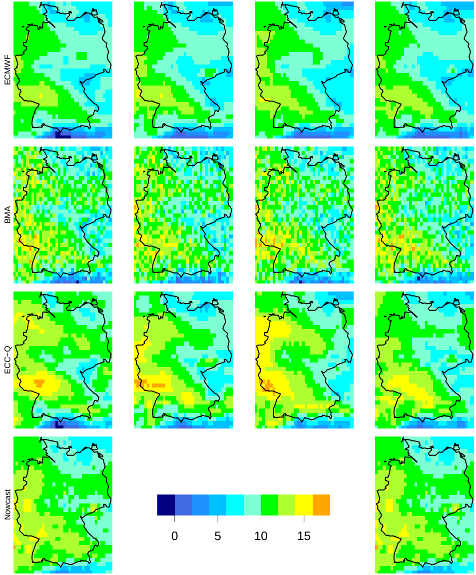


Figure: 24h temperature forecasts valid on 25 April 2011, 2:00 am

Multivariate verification methods

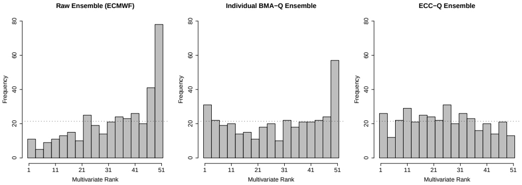
Goal: Maximize sharpness subject to calibration

- Multivariate and band depth rank histogram to check calibration (Gneiting et al., 2008; Thorarinsdottir et al., 2014): Flatness/uniformity indicates calibration
- Proper scoring rules as overall performance measures

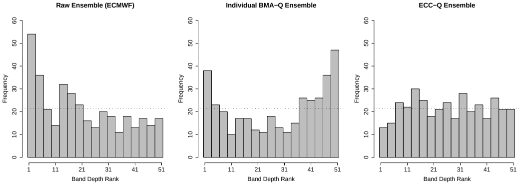
  Energy score for a discrete distribution P_{ens} based on the ensemble forecast $\mathbf{x}_1, \ldots, \mathbf{x}_M \in \mathbb{R}^L$ (the lower the better):

$$\mathrm{ES}(P_{ens}, \mathbf{y}) := \frac{1}{M} \sum_{m=1}^{M} ||\mathbf{x}_m - \mathbf{y}|| - \frac{1}{2M^2} \sum_{m=1}^{M} \sum_{\mu=1}^{M} ||\mathbf{x}_m - \mathbf{x}_\mu||,$$

  with observation vector $\mathbf{y} \in \mathbb{R}^L$ and Euclidean norm $|| \cdot ||$ (Gneiting et al., 2008)



(a) Multivariate rank histograms



(b) Band depth rank histograms

Figure: (a) Multivariate and (b) band depth rank histograms for 24 h ahead pressure forecasts jointly at Berlin, Hamburg and Frankfurt over the test period from 1 May 2010 to 30 April 2011.

Case study: Energy scores

	Temperature (°C)	Pressure (hPa)
Raw ECMWF Ensemble	2.28	1.01
Individual BMA-Q Ensemble	1.73	0.83
ECC-Q Ensemble	**1.72**	**0.81**

Table: Average energy scores for 24 h forecasts of temperature and pressure, respectively, jointly at Berlin, Hamburg and Frankfurt ($L = 3$) over the test period from 1 May 2010 to 30 April 2011.

Summary

- Uncertainty quantification in weather prediction is typically performed via ensemble forecasts.
- Ensembles call for statistical postprocessing, in that biases and dispersion errors need to be addressed.
- Postprocessing via ECC retains the spatial, temporal and inter-variable rank dependence pattern of the raw ensemble.
- ECC is an empirical copula-based method, has a couple of benefits and performs well in our case studies.

Future work

- ECC variant applicable to ensembles with non-exchangeable members
- Comparison of ECC and the Schaake shuffle (Clark et al., 2004), similar to the recent work of Wilks (2014)

Selected references

- Clark, M. P., Gangopadhyay, S., Hay, L. E., Rajagopalan, B. and Wilby, R. L. (2004). The Schaake shuffle: A method for reconstructing space-time variability in forecasted precipitation and temperature fields. Journal of Hydrometeorology, 5, 243-262.
- Feldmann, K., Scheuerer, M. and Thorarinsdottir, T. L. (2014). Spatial postprocessing of ensemble forecasts for temperature using nonhomogeneous Gaussian regression. `arXiv:1407.0058`
- 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.
- Möller, A., Lenkoski, A. and Thorarinsdottir, T. L. (2013). Multivariate probabilistic forecasting using ensemble Bayesian model averaging and copulas. Quarterly Journal of the Royal Met. Soc., 139, 982-991.
- Raftery, A. E., Gneiting, T., Balabdaoui, F. and Polakowski, M. (2005). Using Bayesian model averaging to calibrate forecast ensembles. Monthly Weather Review, 133, 1155-1174.
- Schefzik, R., Thorarinsdottir, T. L. and Gneiting, T. (2013). Uncertainty quantification in complex simulation models using ensemble copula coupling. Statistical Science, 28, 616-640.
- Wilks, D. S. (2014). Multivariate ensemble Model Output Statistics using empirical copulas. To appear in Quarterly Journal of the Royal Met. Soc.