Autoregressive Extensions of EMOS with Application to Surface Temperature Ensemble Postprocessing

Annette Möller[1], David Jobst[2], Jürgen Groß[2]

[1] Bielefeld University, [2] University of Hildesheim

19. April 2024

UNIVERSITÄT BIELEFELD
Faculty of Business Administration and Economics

Contents

1 Introduction

The Need for Ensemble Postprocessing

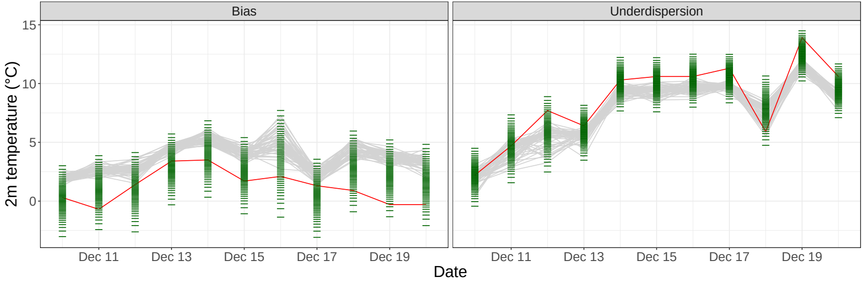


Figure: Station Fürstenzell (left) und Saarbrücken (right) in 2020.

- Ensemble forecasts allow to quantify **forecast uncertainty**,
- however, they typically exhibit **biases and dispersion errors**.
- **Ensemble postprocessing models** correct forecasts in coherence with past forecast errors, yield **better calibrated and sharp predictive distributions**.

2 Postprocessing Models

Ensemble Model Output Statistics (EMOS)

- Let $x_1(t), \ldots, x_m(t)$ denote **ensemble of m forecasts**, and $Y(t)$ **observed weather quantity**, at time points $t = 1, \ldots, T$.
- In subsequent case study: Y represents **surface temperature**, assumption of **Gaussian predictive distribution**.

$$Y(t)|x_1(t), \ldots, x_m(t) \sim \mathcal{N}\left(\mu(t), \sigma^2(t)\right)$$

- **Predictive mean** $\mu(t)$ and **predictive variance** $\sigma^2(t)$ linked to ensemble forecasts via

$$\mu(t) := a_0 + a_1 \overline{x}(t),$$
$$\log(\sigma(t)) := b_0 + b_1 \log(s(t)),$$

  with $\overline{x}(t)$ ensemble mean, $s(t)$ empirical ensemble standard deviation at day t, and coefficients $a_0, a_1, b_0, b_1 \in \mathbb{R}$.
- Estimation via **minimization of proper scoring rule** (typically CRPS) with **sliding training period** (e.g. 30 days length).

Smooth EMOS (SEMOS)

- Introduce **seasonally varying intercept and slope** to account for **seasonality/trend**, instead of employing sliding window for parameter estimation.

⇝ Link of predictive mean and variance to ensemble forecasts:

$$\mu_S(t) := a_0 + f_0(t) + (a_1 + f_1(t)) \cdot \overline{x}(t)$$
$$\log(\sigma_S(t)) := b_0 + g_0(t) + (b_1 + g_1(t)) \cdot s(t)$$

- f_0, f_1, g_0, g_1 – **cyclic regression splines** conditional on t,
- a_0, a_1, b_0, b_1 – **global/unconditional** intercepts and slopes.
- Modification based on truncated Fourier series:

$$f_i(t) := \alpha_{i1} \sin\left(\frac{2\pi t}{365.25}\right) + \alpha_{i2} \cos\left(\frac{2\pi t}{365.25}\right) + \alpha_{i3} \sin\left(\frac{4\pi t}{365.25}\right) + \alpha_{i4} \cos\left(\frac{4\pi t}{365.25}\right),$$
$$g_i(t) := \beta_{i1} \sin\left(\frac{2\pi t}{365.25}\right) + \beta_{i2} \cos\left(\frac{2\pi t}{365.25}\right) + \beta_{i3} \sin\left(\frac{4\pi t}{365.25}\right) + \beta_{i4} \cos\left(\frac{4\pi t}{365.25}\right),$$

  where $\alpha_{i,j}, \beta_{i,j} \in \mathbb{R}$ for $i = 0, 1,\ j = 1, 2, 3, 4$.
- Drawback: **Autoregressive behaviour** not considered.

Autoregressive EMOS (AR-EMOS)

- Fit **AR process of order p_k to error series** $r_k(t) := Y(t) - x_k(t)$ of each ensemble forecast $x_k(t)$.
- **AR-adjusted forecast ensemble** obtained via

$$\widetilde{x}_k(t) := x_k(t) + \widehat{r}_k(t) = x_i(t) + \eta_k + \sum_{j=1}^{p_k} \tau_{k,j}\big(\widehat{r}_k(t-j) - \eta_k\big),$$

with $\eta_k, \tau_{k,j} \in \mathbb{R}$, $j = 1, \ldots, p_i$ AR(p_k) coefficients, $\widehat{r}_i(t)$, $i = 1, \ldots, m$ residuals obtained from $y(t-j)$.

- **Predictive mean** (location) estimated via $\mu(t) := \frac{1}{m}\sum_{k=1}^{m} \widetilde{x}_k(t)$,
- **Predictive variance** (scale) via $\sigma(t) := \omega\sigma_1(t) + (1-\omega)\sigma_2(t)$.
- $\sigma_1(t) := \sqrt{\frac{1}{m}\sum_{k=1}^{m} \gamma_k^2(t)}$, $\gamma_k^2(t)$ empirical variance of AR(p_k) process,
- $\sigma_2(t)$ empirical SD of AR-adjusted forecasts $\widetilde{x}_k(t)$,
- $\omega \in [0, 1]$ weight obtained by minimizing CRPS of predictive Gaussian distribution.

AR-EMOS – Disadvantages

- Requires 2 **rolling training periods** $\rightsquigarrow$ parameters of AR process, weights of linear combination
- Static period often preferred in operational settings
- Location and scale parameter not estimated simultaneously $\rightsquigarrow$ can reduce forecast performance
- Possible general autoregressive conditional heteroscedasticity (GARCH) behavior not taken into account

AR-EMOS and SEMOS – Modifications

- Different extensions of SEMOS within AR-EMOS philosophy:
  1. **Apply AR process to de-seasonalized forecast error of mean** $\rightsquigarrow$ de-seasonalized autoregressive SEMOS (DAR-SEMOS).
  2. Squared forecast errors of DAR-SEMOS show (G)ARCH effects $\rightsquigarrow$ generalized autoregressive conditional variance component (DAR-GARCH-SEMOS).
  3. **Apply AR process to standardized residuals** (anomalies) of SEMOS model $\rightsquigarrow$ standardized AR-SEMOS (SAR-SEMOS).
- Model **parameters estimated simultaneously** by minimizing a proper scoring rule in comparison to 2-step approach in original AR-EMOS model.
- Models able to postprocess **forecasts of arbitrary lead time**.

Deseasonalized Autoregressive SEMOS (DAR-SEMOS)

- **Motivation**: Include **autoregressive behaviour**, not only trend and seasonal patterns as in SEMOS.

$\rightsquigarrow$ Apply AR(p) model to residuals **after** removing mean/trend.

- Assume AR(p) process for error series $r(t) := Y(t) - \mu_S(t)$

$$r(t) := \eta + \sum_{j=1}^{p} \tau_j \big(r(t-j) - \eta\big) + \varepsilon(t), \quad \varepsilon(t) := \sigma_S(t) \cdot z(t)$$

$$\mu_S(t) := a_0 + f_0(t) + (a_1 + f_1(t)) \cdot \overline{x}(t),$$

$$\log(\sigma_S(t)) := b_0 + g_0(t) + (b_1 + g_1(t)) \cdot s(t),$$

- where $\eta, \tau_j \in \mathbb{R}$, $j = 1, \ldots, p$ are AR(p) coefficients,
- $z(t) \sim \mathcal{N}(0, 1)$ white noise,
- f_0, f_1, g_0, g_1 truncated Fourier series, $s(t)$ emp. ensemble SD.
- Parameters of predictive distribution $\mathcal{N}(\mu, \sigma^2)$ obtained by

$$\mu(t) := \mu_S(t) + \eta + \sum_{j=1}^{p} \tau_j \big(\widehat{r}(t-j) - \eta\big).$$

Estimation of DAR-SEMOS: 2-Stage Procedure

I. Initialization:

1. Estimate **initial coefficients of seasonal model** $\mu(t)$ by linear regression (OLS) $y(t) = \mu_S(t) + r^*(t)$, $r^*(t) \sim \mathcal{N}(0, \sigma^2_{r^*})$.
2. Use model residuals $\hat{r}^*(t)$ to estimate **initial order** p **and coefficients of corresponding AR**(p) **process** $r(t) = \eta + \sum_{j=1}^{p} \tau_j \big(r(t-j) - \eta \big) + \varepsilon^*(t)$ with white noise $\varepsilon^*(t) \sim \mathcal{N}(0, \sigma^2_{\varepsilon^*})$
3. Set initial coefficients for scale parameter $\sigma_S(t)$ all to 0, except for $b_1 = 1$.

Estimation of DAR-SEMOS: 2-Stage Procedure

II. Optimization: Fix order p of AR(p) process and optimize all parameters simultaneously with respect to CRPS.

1. Calculate $\widehat{\mu}_S(t) = \widehat{a}_0 + \widehat{f}_0(t) + (\widehat{a}_1 + \widehat{f}_1(t)) \cdot \overline{x}(t)$,
   $\widehat{r}(t) = y(t) - \widehat{\mu}_S(t)$.
2. Predict model residuals $\widehat{r}(t) = \widehat{\eta} + \sum_{j=1}^{p} \widehat{\tau}_j\big(\widehat{r}(t-j) - \widehat{\eta}\big)$.
3. Update $\widehat{\mu}(t) = \widehat{\mu}_S(t) + \widehat{r}(t)$,
   $\widehat{\sigma}(t) = \exp\left(\widehat{b}_0 + \widehat{g}_0(t) + (\widehat{b}_1 + \widehat{g}_1(t)) \cdot s(t)\right)$.
4. Calculate CRPS $(\mathcal{N}(\widehat{\mu}(t), \widehat{\sigma}^2(t)))$.

⇝ Procedure can be applied **to all lead times**.

⇝ To forecast lead times > 24h proceed as in AR-EMOS and predict necessary residuals by AR(p) model.

Standardized Autoregressive SEMOS (SAR-SEMOS)

- So far: autoregressive behavior of **residuals** $r(t)$ **and variance** $\sigma^2(t)$ **considered separately**.
- Modeling **autoregressive behavior of standardized residuals**

$$z(t) := \frac{r(t)}{\sigma_S(t)} = \frac{Y(t) - \mu_S(t)}{\sigma_S(t)}$$

generalizes previous approaches.

$$z(t) := \eta + \sum_{j=1}^{p} \tau_j\big(z(t-j) - \eta\big) + \xi(t),$$

- where $z \sim \mathrm{AR}(p)$, with coefficients $\eta, \tau_j \in \mathbb{R}$, $j = 1, \ldots, p$,
- and $\xi(t) \sim \mathcal{N}(0, \sigma_\xi^2)$ white noise.
- Parameters of predictive distribution $\mathcal{N}(\mu, \sigma^2)$ obtained via

$$\mu(t) := \mu_S(t) + \sigma_S(t) \cdot \left(\eta + \sum_{j=1}^{p} \tau_j\big(\widehat{z}(t-j) - \eta\big)\right).$$

Estimation of SAR-SEMOS: 2-Stage Procedure

I. Initialization:

1. Initialize coefficients for μ_S and σ_S as in DAR-SEMOS to obtain $\widehat{\mu}_S(t)$ and $\widehat{\sigma}_S(t)$.
2. Use standardized residuals $\widehat{z}(t) = \frac{y(t) - \widehat{\mu}_S(t)}{\widehat{\sigma}_S(t)}$ to estimate initial order p and coefficients of residual AR(p) process.

Estimation of SAR-SEMOS: 2-Stage Procedure

II. Optimization: Fix order p of AR(p) process and optimize all parameters simultaneously with respect to CRPS.

1. Update

$$\widehat{\mu}_S(t) = \widehat{a}_0 + \widehat{f}_0(t) + (\widehat{a}_1 + \widehat{f}_1(t)) \cdot \overline{x}(t),$$
$$\widehat{\sigma}_S(t) = \exp\left(\widehat{b}_0 + \widehat{g}_0(t) + (\widehat{b}_1 + \widehat{g}_1(t)) \cdot s(t)\right).$$

2. Update $\widehat{z}(t) = \frac{y(t) - \widehat{\mu}_S(t)}{\widehat{\sigma}_S(t)}$ and predict model residuals

$$\widehat{z}(t) = \widehat{\eta} + \sum_{j=1}^{p} \widehat{\tau}_j \big(\widehat{z}(t-j) - \widehat{\eta}\big).$$

3. Update $\widehat{\mu}(t) = \widehat{\mu}_S(t) + \widehat{\sigma}_S(t) \cdot \widehat{z}(t)$, $\widehat{\sigma}(t) := \widehat{\sigma}_S(t)$.
4. Calculate CRPS $(\mathcal{N}(\widehat{\mu}(t), \widehat{\sigma}^2(t)))$.

$\rightsquigarrow$ Estimation procedure **can be applied to all lead times.**

$\rightsquigarrow$ To forecast lead times > 24h proceed as in DAR-SEMOS.

Overview of Models

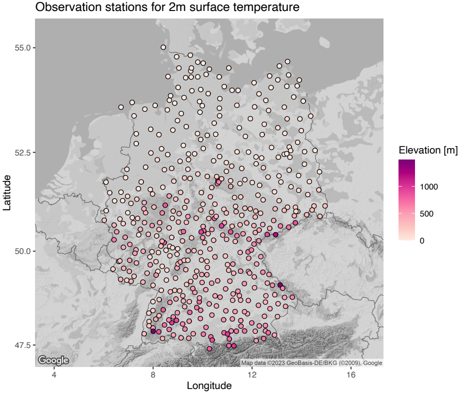
Gaussian distribution assumed for all methods for the weather quantity **2m surface temperature**, i.e. $Y|x_1, \ldots, x_k \sim \mathcal{N}(\mu, \sigma^2)$.

Parameter	μ			σ^2 or σ		
Method	season	trend	ar	season	trend	ar
EMOS						
AR-EMOS			✓			✓
SEMOS	✓	✓		✓	✓	
DAR-SEMOS	✓	✓	✓	✓	✓	
DAR-GARCH-SEMOS	✓	✓	✓	✓	✓	✓
SAR-SEMOS	✓	✓	✓	✓	✓	✓

3 Case Study

Data

- **Weather Variable:** 2m surface temperature at 462 stations in Germany (DWD)
- **Forecasts:** 50 member forecast ensemble of ECMWF
- **Forecast horizons:** 24 h, 48 h, 72 96 h, 120 h
- **Initialization time:** 12 UTC
- **Resolution:** $0.25^\circ \times 0.25^\circ$
- Bilinear **interpolation to station locations** from grid points
- **Training data:** 2015 - 2019
- **Test data:** 2020.



Seasonal and Autoregressive Effects Temperature

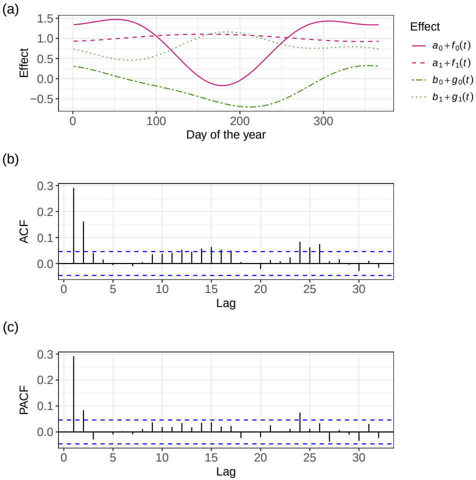


Figure: (a) Estimated seasonal intercept and slope effects for location (purple) and log-scale (green), (b) ACF and (c) PACF for forecast residuals at station Metzingen.

Overall PIT Histograms

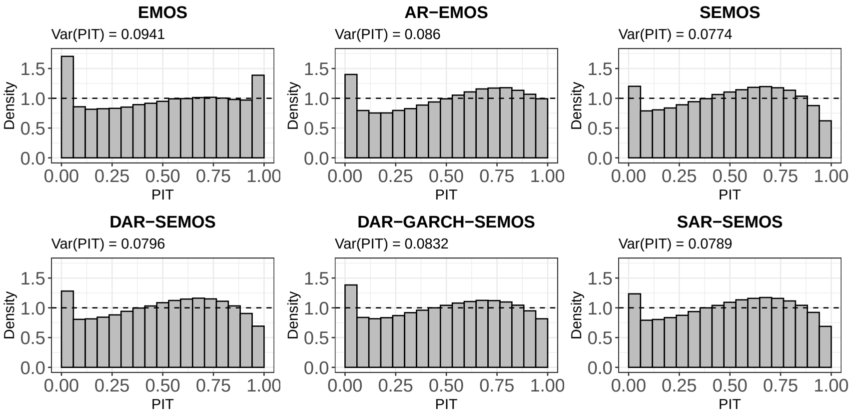


Figure: PIT histograms of the methods aggregated over all time points, stations and lead times in the validation period.

Overall Verification Scores

Method	CRPS	LogS	MAE	Width	Coverage
Raw ensemble	1.165	–	1.513	4.935	74.69
EMOS	1.007	2.183 (0%)	1.376	6.407	90.08
AR-EMOS	0.943	1.952 (0%)	1.309	6.442	93.41
SEMOS	0.908	1.880 (0%)	1.259	6.688	**95.25**
DAR-SEMOS	0.902	1.885 (0.23%)	1.249	6.458	94.67
DAR-GARCH-SEMOS	0.902	1.906 (0%)	1.249	**6.282**	93.93
SAR-SEMOS	**0.890**	**1.865** (0%)	**1.233**	6.401	94.94

Figure: Verification scores aggregated over all time points, stations and lead times in the validation period. Bold value represents the best value for each score. The number in brackets for the LogS denote the percentage of omitted infinite values.

Lead-time specific Results

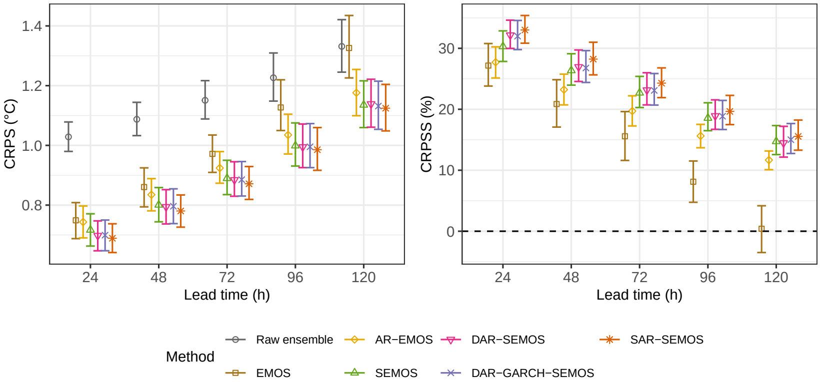


Figure: Mean CRPS (left) and CRPSS over raw ensemble (right), including the 95%-confidence intervals for the validation period.

Lead-time specific Results

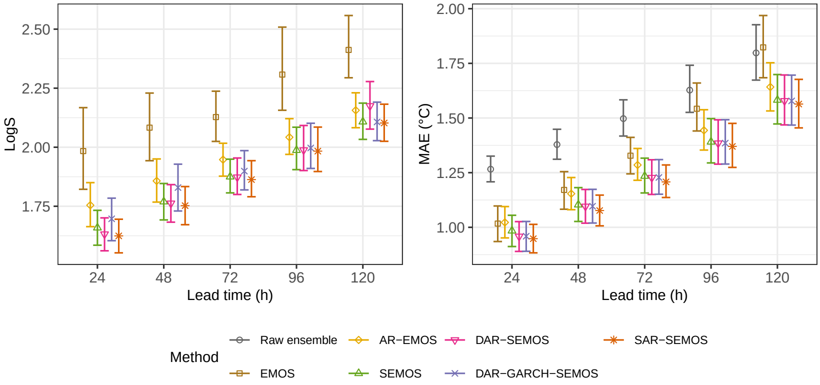


Figure: Mean LogS (left) and MAE (right), including the 95%-confidence intervals for the validation period.

Lead-time specific Results

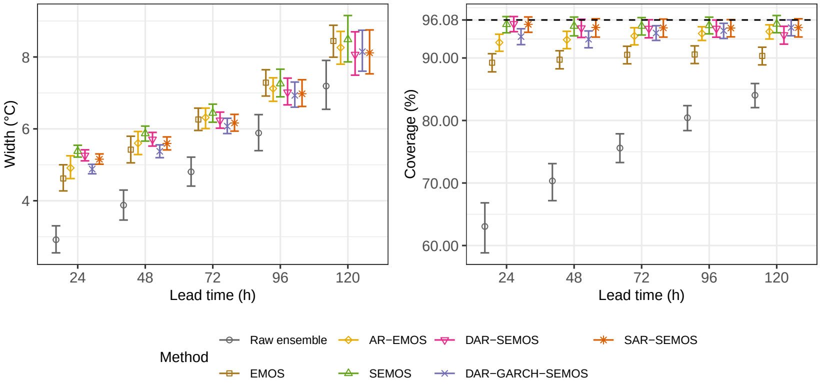


Figure: Mean width (left) and coverage (right), including the 95%-confidence intervals for the validation period.

Station-wise Results

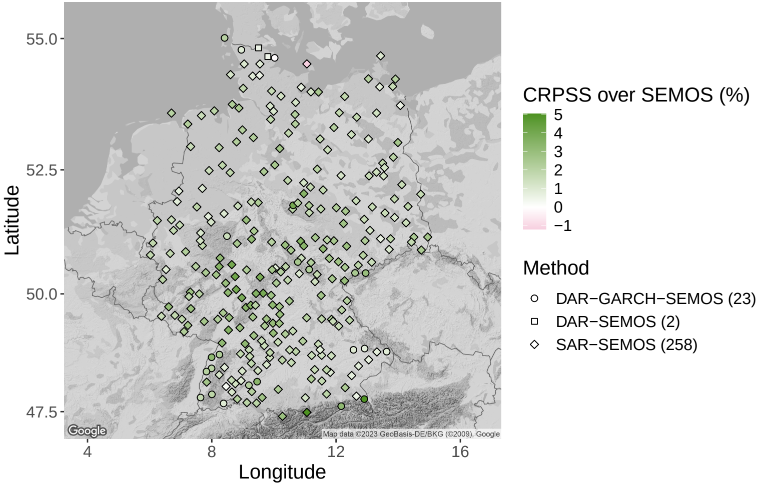


Figure: Station-wise highest CRPSS of the considered methods across all lead times over SEMOS in % in the validation period.

4 Outlook

Outlook

Further possible extensions:

- Allowing **more adjustable distribution functions**, as e.g. skew Gaussian distribution or skew Student-t
- Taking account of autoregressive behavior in (standardized) residuals in postprocessing techniques and at the same time **allowing a large set of predictor variables** such as e.g. gradient-boosted EMOS.
- Modification for **other weather variables**, such as wind speed and precipitation.
- **Combination with multivariate techniques such as ECC**, Gaussian copulas, vine copulas

Research supported by DFG within Projects

- Statistical postprocessing of ensemble forecasts for various weather quantities (Grant No 395388010)
- Mitigating climate risks by improving weather forecasts using copula based approaches for postprocessing of forecast ensembles (Grant No 520017589)

DFG Deutsche Forschungsgemeinschaft

Thank you for your attention!

Dabernig. M., Mayr. G.J., Messer, J.W. and Zeileis, A. (2017).
Spatial ensemble post-processing with standardized anomalies.
Quarterly Journal of the Royal Meteorological Society, Volume 143, Issue 703, 909 – 916.

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

Gneiting, T. and A. E. Raftery (2007).
Strictly proper scoring rules, prediction, and estimation.
Journal of the American Statistical Association 102, 359–378.

Jobst, D., Möller, A., and Groß, J. (2024).
Time Series based Ensemble Model Output Statistics for Temperature Forecasts Postprocessing.
arXiv:2402.00555

Jobst, D., Möller, A., and Groß, J. (2023).
tsEMOS. Version 0.1.0.
`https://github.com/jobstdavid/tsEMOS`.

Möller, A. and Groß, J. (2016).
Probabilistic temperature forecasting based on an ensemble AR modification.
Quarterly Journal of the Royal Meteorological Society, Volume 142, Issue 696, 1385 – 1394.

Möller, A. and Groß, J. (2020).
Probabilistic Temperature Forecasting with a Heteroscedastic Autoregressive Ensemble Postprocessing model.
Quarterly Journal of the Royal Meteorological Society, Volume 146, Issue 726, 211 – 224.