Simultaneous ensemble post-processing for multiple lead times with standardized anomalies

Markus Dabernig, Georg J. Mayr, Jakob W. Messner, Achim Zeileis

Working Papers in Economics and Statistics

2016-31

University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

- Department of Banking and Finance
- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact address of the editor:
Research platform “Empirical and Experimental Economics”
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: + 43 512 507 7171
Fax: + 43 512 507 2970
E-mail: eeecon@uibk.ac.at

The most recent version of all working papers can be downloaded at
http://eeecon.uibk.ac.at/wopec/

For a list of recent papers see the backpages of this paper.

Simultaneous Ensemble Post-Processing for Multiple Lead Times with Standardized Anomalies

Markus Dabernig Universität Innsbruck
Georg J. Mayr Universität Innsbruck
Jakob W. Messner Universität Innsbruck
Achim Zeileis Universität Innsbruck

Abstract

Statistical post-processing of ensemble predictions is usually adjusted to a particular lead time so that several models must be fitted to forecast multiple lead times. To increase the coherence between lead times, we propose to use standardized anomalies instead of direct observations and predictions. By subtracting a climatological mean and dividing by the climatological standard deviation, lead-time-specific characteristics are eliminated and several lead times can be forecasted simultaneously. The results show that forecasts between +12 and +120 h can be fitted together with a comparable forecast skill to a conventional method. Furthermore, forecasts can be produced with a temporal resolution as high as the observation interval e.g., up to ten minutes.

Keywords: standardized anomalies, non-homogeneous regression, ensemble post-processing, probabilistic temperature forecasts.

1. Introduction

Weather forecasts are important for many aspects of life whether professional or private. These weather forecasts rely mainly on numerical weather prediction (NWP) models, which solve partial differential equations. However, it is impossible to solve these equations exactly and errors occur, which grow with forecasting horizon. Additionally, due to limited computational power, the numerical grid of these forecast models has to be coarse and sub-grid processes have to be parameterized. These parameterizations lead to forecast errors especially over complex terrain, such as the Alps. Therefore, starting from Glahn and Lowry (1972) statistical methods are employed to correct these systematic errors. Statistical post-processing learns the relationship between numerical forecasts and observations from a historical dataset to correct future forecasts.

Additionally, weather centers are calculating several perturbations of the models with different parameterizations and initial conditions to capture uncertainties of the atmosphere (Lorenz 1982; Leutbecher and Palmer 2008). However, not all error sources can be considered so that ensembles are often still biased and underdispersive. To get unbiased and calibrated probabilistic forecasts similar methods as Glahn and Lowry (1972) can be applied. Examples of such models are nonhomogeneous Gaussian regression (Gneiting, Raftery, Westveld, and Golfman 2005) or Bayesian model averaging (Raftery and Gneiting 2005).

NWP models produce forecasts with typical lead time intervals between one and six hours for a certain grid location. They are usually post-processed for every lead time and station individually, which can be time consuming for several lead times or stations. Scheuerer and Büermann (2014) proposed a method to forecast several stations and a certain lead time with anomalies. Stauffer, Messner, Mayr, Umlauf, and Zeileis (2016) and Dabernig, Mayr, Messner, and Zeileis (2016) adjusted this method by standardizing the anomalies to produce spatial forecasts.

Their basic idea is to fit a statistical model on standardized anomalies (Wilks 2011) instead of the direct observations and ensemble forecasts. Because these standardized anomalies should not have any station specific characteristics anymore one statistical model can be used for all locations within a certain region.

We adapt the method of Dabernig *et al.* (2016) but instead of forecasting several stations simultaneously we forecast several lead times simultaneously. Additional to saving computation time, this approach allows to provide calibrated forecasts in a much higher temporal resolution as the NWP forecasts. Therefore, only a observation climatology at a high resolution has to be available which captures all lead-time specific effects.

Another advantage of these standardized anomalies is that they do not only remove lead-time-specific characteristics but also seasonal specifics. Conventional methods are usually trained on a sliding window (Gneiting *et al.* 2005) to account for seasonal differences in the relationship between forecasts and observations. Standardized anomalies do not contain these season-specific characteristics any more so that less limited training data sets can be used.

As a result this method has three distinct advantages: All data can be used as training data, not only a subset, all lead times can be forecasted simultaneously, and forecasts are available at a higher temporal resolution as provided by the numerical model.

The rest of this article is structured as follows: The data is described in the next chapter, followed by the methodology. Subsequently, results for a single station and for multiple stations are presented with a conclusion at the end.

2. Data

The proposed method is tested with measurements from 39 automatic weather stations located in the province of South Tyrol in northern Italy. Here we use 2 m temperatures that are available in ten-minute resolution.

The ensemble NWP forecasts are taken from the European Centre for Medium-Range Weather Forecasts (ECMWF) and are bilinearly interpolated to the measurement sites. Lead times between +12 and +120 hours are used with an interval of three hours.

The forecast variable (y) is the temperature measured 2 m above ground. The predictor variable used from the 51 members of the NWP model is also the 2 m temperature. Data from 2010-02-01 to 2016-01-31 are used with a constant horizontal resolution of 32 km (T639).

3. Method

3.1. Nonhomogeneous Gaussian regression

Gneiting *et al.* (2005) introduced nonhomogeneous Gaussian regression (NGR) for statistically post-processing of ensemble forecasts. The observations are assumed to follow a normal

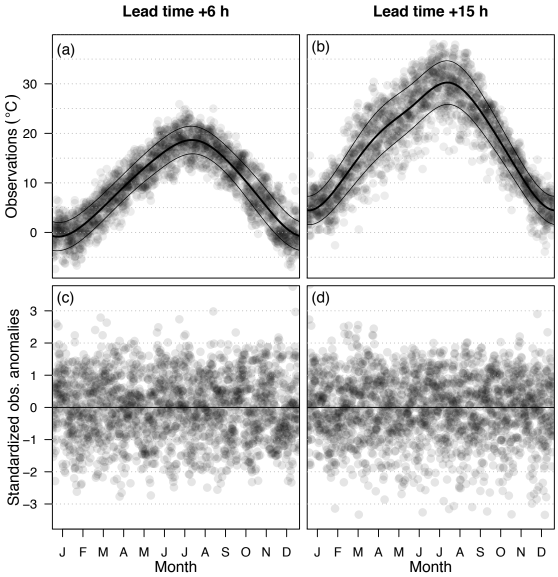

Figure 1: Annual cycle of the observations with the climatological mean (thick line) and the climatological standard deviation (thin lines) for (a) 6 UTC and (b) 15 UTC. (c) Annual cycle of the standardized observation anomalies for 6 UTC (d) and 15 UTC.

distribution with regression equations for both, the mean (μ) and the standard deviation (σ):

$$y \sim N(\mu, \sigma), \tag{1}$$

$$\mu = b_0 + \; b_1 m, \tag{2}$$

$$\log(\sigma) = c_0 + c_1 \log(s), \tag{3}$$

with regression coefficients b_0, b_1, c_0 and c_1. μ depends on the ensemble mean (m), and σ on the ensemble spread (s). We adapted the method of Gneiting *et al.* (2005) slightly by using the logarithmic link function in Equation 3 to ensure positive values for the standard deviation σ. Another adaption is that we fit Equation 1 with a maximum likelihood estimation provided by the R package *crch* (Messner, Mayr, and Zeileis 2016) instead of CRPS minimization.

3.2. Standardized anomalies

The basic idea of using standardized anomalies is visualized in Figure 1. Whereas the annual cycles clearly differ for two different lead times (top), the standardized anomalies for both lead times are centered around zero and have no pronounced annual cycle anymore (bottom panel). This may allow to use the same coefficients for both lead times and to fit only one statistical model.

To obtain standardized anomalies, the climatological mean (μ_y) is subtracted from the observations and divided by a climatological standard deviation (σ_y):

$$\widetilde{y} = \frac{y - \mu_y}{\sigma_y}. \tag{4}$$

The standardized anomalies $\widetilde{m}$ and the corresponding $\widetilde{s}$ are calculated as follows:

$$\widetilde{m} = mean\left(\frac{\text{ens} - \mu_{\text{ens}}}{\sigma_{\text{ens}}}\right) \tag{5}$$

and

$$\widetilde{s} = sd\left(\frac{\text{ens} - \mu_{\text{ens}}}{\sigma_{\text{ens}}}\right), \tag{6}$$

where ens is the 51 ensemble member forecast and μ_{ens} and σ_{ens} are the climatological mean and standard deviation, respectively. The members of the ensemble forecasts are exchangeable and therefore μ_{ens} is the same for all members. $\widetilde{m}$ is then the mean of all 51 standardized ensemble member anomalies and $\widetilde{s}$ their standard deviation. The details for computing μ_y, μ_{ens}, σ_y, and σ_{ens} will be shown in Subsection 33.4.

3.3. Standardized anomaly model output statistics (SAMOS)

Standardized anomalies remove seasonal and lead-time-specific characteristics and as a result several lead times of the same station can be fitted and forecasted together, if the climatology is able to capture these different characteristics. Therefore, Equation 1 is fitted with standardized anomalies ($\widetilde{y}$, $\widetilde{m}$ and $\widetilde{s}$) instead of direct observations and predictions. The fitted coefficients are then representative for all fitted lead times and since the coefficients (b_0[1], b_1, c_0 and c_1) are fitted on data without lead-time specifics they are also representative for every lead time. However, the ensemble spread has to provide the information to capture differences in the forecast uncertainty. Otherwise, the model can not adapt to e.g., higher forecast uncertainties at longer lead times.

To use this fitted model for forecasting, Equation 2 and 3 have to be restructured into:

$$\hat{\mu} = (b_0 + \ b_1\widetilde{m}) \cdot \sigma_y + \mu_y \tag{7}$$

and

$$\hat{\sigma} = \exp(c_0 + c_1 \log(\widetilde{s})) \cdot \sigma_y. \tag{8}$$

Since the coefficients do not depend on any lead time, the use of these equations are not restricted to the lead times of the NWP forecasts. By using a higher temporal resolution of the observation climatology (μ_y and σ_y) and bilinear interpolated ensemble forecasts ($\widetilde{m}$, $\widetilde{s}$) forecasts for any lead time can be computed.

3.4. Climatology

To obtain forecasts with a high temporal resolution, a temporally highly resolved climatology is necessary. It can be produced with generalized additive models for location, scale and shape (GAMLSS, Rigby and Stasinopoulos 2005; Stasinopoulos and Rigby 2007)) as in Dabernig *et al.* (2016). With GAMLSS one model for climatological mean and standard deviation for all lead times and all points in between can be calculated simultaneously. GAMLSS is similar to NGR but can also include nonlinear effects,

$$y \sim N(\mu_y, \sigma_y). \tag{9}$$

[1]Due to the centering of the standardized anomalies b_0 is not significant and could be left out.

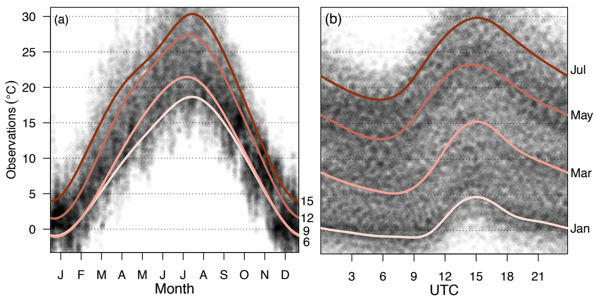

Figure 2: (a) All observations are plotted against the day of the year with climatological mean for different lead times from +6 (brightest, morning) to +15 h (darkest line, afternoon). (b) All observations against the minute of the day with climatological mean for mid January (brightest) to mid July (darkest line).

To calculate a climatology the daily, variation has to be considered as well as the seasonal change. The nonlinear models for μ_y and σ_y are therefore:

$$\mu_y = \beta_0 + f(\text{yd}, \text{dm}) \text{ and } \sigma_y = \gamma_0 + g(\text{yd}, \text{dm}), \tag{10}$$

with β_0 and γ_0 as regression coefficients and f and g as smooth functions which capture the interaction between the day of the year (yd) and the minute of the day (dm, 0–1439) with a cyclic cubic regression spline (fitted with the R package *gamlss*). The degree of freedom for the day of the year is eight and ten for the minute of the day.

The resulting effects of the fitted climatological mean based on ten minute *measurements* are illustrated in Figure 2. Figure 2a shows the seasonal effect for four different times of the day with colder temperatures in the winter than in the summer. The colors indicate that different lead times have different seasons. The season at the morning hours (bright colors) has a smaller amplitude than during the day (darker colors).

Figure 2b shows the diurnal variation for four different months. Whereas all lines show a minimum in the morning hours and a maximum during day, the minimum is in January (brightest line) at around 9 UTC whereas in July (darkest line) it is at around 6 UTC. The interaction is necessary to capture the changing diurnal cycles over the year and on the other hand the changing annual cycles over the day. With these effects a climatology for every day of the year and every minute of the day is available.

Additional to observation climatologies, climatologies of the ensemble forecasts are required. Forecasts are only available at a much lower temporal resolution (i.e., 3 hours) and there is no information in between these lead times. Therefore, it makes little sense to compute a full climatology for all lead times. Instead, simpler climatologies without interactions are calculated at each lead time individually. To get forecasts in between these lead times, the standardized anomalies are simply linearly interpolated. For these ensemble climatologies, Equation 10 reduces to

$$\mu_{\text{ens}} = \beta_0 + f(\text{yd}) \text{ and } \sigma_{\text{ens}} = \gamma_0 + g(\text{yd}), \tag{11}$$

where only one nonlinear effect for the day of the year is necessary using eight degrees of freedom.

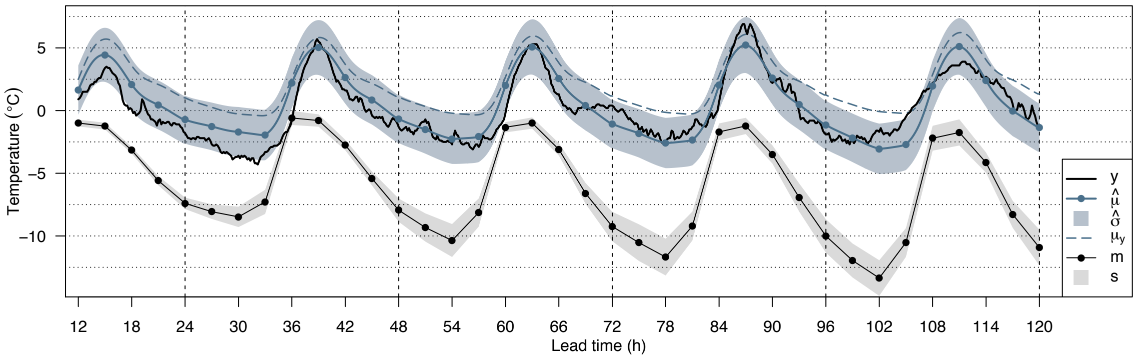

Figure 3: SAMOS full forecast for a certain initial day (2011-01-25) with lead times from +12 to +120 h compared to the observations (y) at the station in Auer. Additionally, the ensemble forecasts.

4. Different models

Three different models with standardized anomalies are compared against a reference to investigate the advantages of SAMOS. As reference, an individual NGR (EMOS, Gneiting *et al.* 2005) is fitted for each three-hourly lead time with a sliding window of the previous 30 days as training data.

The SAMOS variations differ in the calculation of the observation climatology and NGR. The observation climatology can either be calculated at every lead time individually as in Equation 11 or at all lead times simultaneously as in Equation 10. As second difference, Equation 7 and 8 can either be calculated at all lead times individually or simultaneously. The three SAMOS variations are:

- *SAMOS lead-time-wise* (*SAMOS ltw*), where the climatologies and the NGR equations are calculated at every three-hourly lead time individually. This model shows the difference between an NGR with standardized anomalies and without.

- For the second model *SAMOS lead-time-wise simultaneous* (*SAMOS ltw-S*) the climatology is also fitted on all lead times individually but the NGR equation is calculated for all lead times simultaneously. This model gives insight if fitting all lead times with on NGR gives a similar result to lead-time-wise forecasts.

- The third model is *SAMOS full* as described in the methods section with one climatology over all available observations and one NGR for all three-hourly lead times simultaneous.

A forecast example of SAMOS full for a particular initial day is shown in Figure 3. Whereas conventional methods such as EMOS only produce forecasts at the three-hourly lead times, SAMOS full also provides forecasts in between. The forecasts in between the three-hourly lead times follow the shape of the observation climatology (dashed line in Figure 3). Additionally, SAMOS full inflates the uncertainty of the raw NWP ensemble to correct for its underdispersion (Hopson 2014; Möller and Groß 2016; Dabernig *et al.* 2016).

The performance of these forecasts is evaluated in the next section.

5. Results

To use all available data, ten-fold cross-validation is performed on the time series for the SAMOS models resulting in training data sets of approximately 5.5 years. In contrast, EMOS is fitted daily with 30 days prior to the forecast day. Additionally, every station has been forecasted individually.

The mean absolute error (MAE) is used as deterministic measure while the continuous rank probability scores (CRPS) is used to test the probabilistic performance. Both are calculated for every lead time and model individually. For a better comparison, a skill score is used,

$$\text{skill score} = 1 - \frac{\text{score}}{\text{score}_{\text{ref}}}, \tag{12}$$

where the score of one model is compared with a reference model.

To investigate calibration, the root mean squared error (RMSE) is compared to the standard deviation.

In the following, we first show results for a single station (Auer) before aggregating results of all stations.

5.1. Single station

To evaluate how well the models are calibrated, the RSME and the standard deviation of the forecasts are compared in Figure 4. A well calibrated forecasted should have a similar predicted standard deviation as the RMSE of the predicted mean. Whereas EMOS is underdispersive at all lead times, SAMOS lead-time-wise (Figure 4b) is almost perfectly calibrated. SAMOS full seems also to be well calibrated but is slightly overdispersive in the afternoons of the first four days.This pattern of overdispersion is also apparent for SAMOS lead-time-wise simultaneous.

The MAE/CRPS results are illustrated in Figure 5. Forecasts in between the three-hourly lead times for EMOS, SAMOS lead-time-wise and SAMOS lead-time-wise simultaneous are interpolated with a natural spline function to the ten-minute resolution. A natural spline was used instead of a linear interpolation to better capture the diurnal cycle of the temperature. A distinct daily oscillating pattern is shown by all the models where during the day the forecasts are worse than during night. However, all three SAMOS models have similar MAE and CRPS whereas EMOS is slightly worse at almost all lead times.

The advantage of SAMOS full is better shown by the bottom panel in Figure 5, where the skill score of SAMOS full with SAMOS lead-time-wise as reference is visualized. Whereas the reference model is almost always better at the lead times provided by the ECMWF model (red dots), the skill of the SAMOS full model in between increases the further away the points are from a provided three-hourly lead time.

So far, we have only regarded the performance at a single forecast location (Auer). To get more general results, in the following we aggregate the results of the 39 stations in the region of South Tyrol.

5.2. All stations

For these results one MAE/CRPS at every lead time averaged over all stations is calculated and compared with a skill score. One set of skill scores are only aggregated over the three-

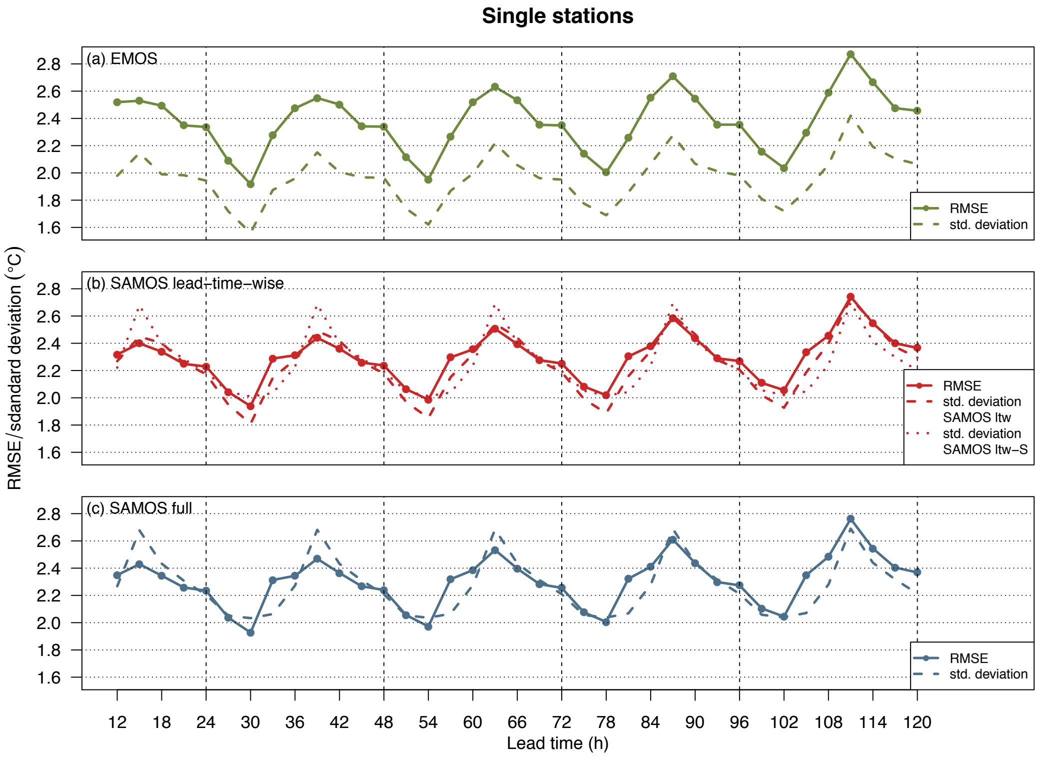

Figure 4: RMSE/standard deviation at the three-hourly lead times for (a) EMOS, (b) SAMOS lead-time-wise and (c) SAMOS full.

hourly ECMWF lead times and the other skill scores are computed for all forecasts at a ten minutes resolution. The results are shown in Figure 6.

Regarding the three-hourly MAE, SAMOS lead-time-wise, and SAMOS lead-time-wise simultaneous are slightly, but not significantly, better than the reference and SAMOS full is slightly worse than EMOS. In contrast, the CRPS of all SAMOS models on the three-hourly lead times is clearly better than that of EMOS. This difference between MAE and CRPS shows that the SAMOS models capture the forecast uncertainty better than EMOS which can also be seen in Figure 4. SAMOS full is still slightly worse than the other SAMOS variations.

In contrast, SAMOS full performs as good as the other SAMOS variants when regarding not only the three-hourly lead times but all forecasts at ten-minute resolution. In other words, the better performance of SAMOS full in between the three-hourly ECMWF lead times can compensate its worse performance at the lead times.

6. Conclusion

Statistical post-processing with standardized anomalies allows to use one single statistical model for several lead times. Furthermore, this approach allows to provide calibrated probabilistic forecasts at a much higher temporal resolution than the raw ensemble forecasts.

Additionally, these standardized anomalies are without season-specific characteristics so that the training data is not limited to similar seasons. Instead a of 30 days moving training

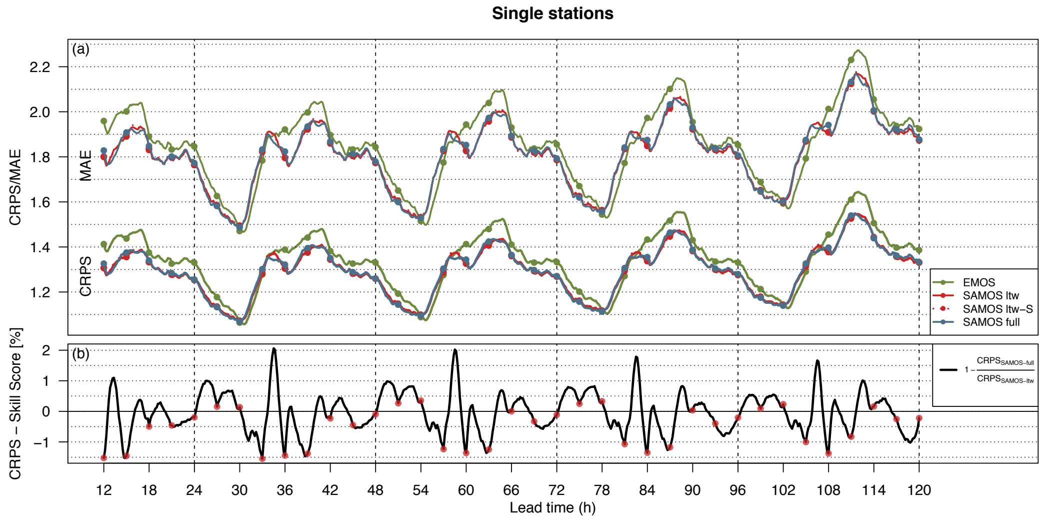

Figure 5: (a) MAE and CRPS for different models and lead times +12 to +120 h at ten minutes resolution. (b) CRPS skill score for SAMOS full with SAMOS lead-time-wise as reference at ten minutes resolution. The red dots mark the three-hourly lead times of the ensemble forecasts.

window that has been used frequently in the literature (Gneiting *et al.* 2005) we could use 6 years of available data for training. The training data could be expanded even further with reforecasts as shown by Stauffer *et al.* (2016).

The results showed using a single model for all lead times provides similar good forecasts as separate models for each lead time. Compared to EMOS, with the sliding window, the SAMOS variations with all available training data are better calibrated and perform better in terms of MAE and CRPS.

Compared to simpler approaches for interpolating forecasts between lead times, the full SAMOS model shows clear advantages in between the lead times provided by the ECMWF. Overall, this better performance between the ensemble lead times compensates the slightly worse performance at the lead times themselves.

7. Acknowledgments

The authors thank the Austrian weather service ZAMG for providing access to ECMWF data and the South Tyrolean weather service for the observations. This study is supported by the Autonome Provinz Bozen (Abteilung Bildungsförderung, Universität und Forschung – ORBZ110725).

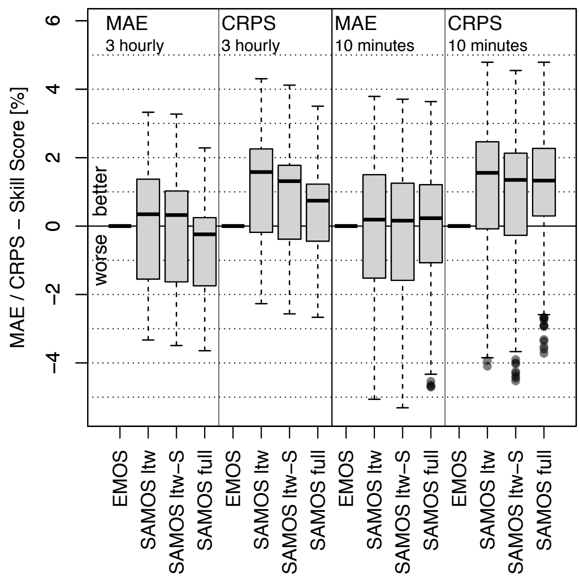

Figure 6: Comparison of different models with a skill score based on EMOS as reference. Boxplots on the left contain one MAE/CRPS point averaged over all stations for every three-hourly lead time (37 points). Boxplots on the right contain one mean MAE/CRPS point averaged over all stations for each ten minutes lead time (649 points).

References

Dabernig M, Mayr GJ, Messner JW, Zeileis A (2016). "Spatial Ensemble Post-Processing with Standardized Anomalies." *Working Papers*, Faculty of Economics and Statistics, University of Innsbruck. URL http://EconPapers.repec.org/RePEc:inn:wpaper:2016-08.

Glahn HR, Lowry DA (1972). "The Use of Model Output Statistics (MOS) in Objective Weather Forecasting." *Journal of Applied Meteorology*, **11**(8), 1203–1211. doi:10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2.

Gneiting T, Raftery AE, Westveld AH, Golfman T (2005). "Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation." *Monthly Weather Review*, **133**(5), 1098–1118. ISSN 1520-0493. doi:10.1175/MWR2904.1.

Hopson TM (2014). "Assessing the Ensemble Spread-Error Relationship." *Monthly Weather Review*, **142**(3), 1125–1142. ISSN 0027-0644, 1520-0493. doi:10.1175/MWR-D-12-00111.1.

Leutbecher M, Palmer TN (2008). "Ensemble Forecasting." *Journal of Computational Physics*, **227**(7), 3515–3539. doi:10.1016/j.jcp.2007.02.014.

Lorenz EN (1982). "Atmospheric Predictability Experiments with a Large Numerical Model." *Tellus A*, **34**(6), 505–513. ISSN 0280-6495. doi:10.3402/tellusa.v34i6.10836.

Messner JW, Mayr GJ, Zeileis A (2016). "Heteroscedastic Censored and Truncated Regression with crch." *R Journal*, **8**(1), 173–181.

Möller A, Groß J (2016). "Probabilistic Temperature Forecasting Based on an Ensemble Autoregressive Modification." *Quarterly Journal of the Royal Meteorological Society*, **142**(696), 1385–1394. ISSN 00359009. doi:10.1002/qj.2741. URL http://doi.wiley.com/10.1002/qj.2741.

Raftery A, Gneiting T (2005). "Using Bayesian Model Averaging to Calibrate Forecast Ensembles." *Monthly Weather Review*, **133**(5), 1155–1174. doi:10.1175/MWR2906.1.

Rigby RA, Stasinopoulos DM (2005). "Generalized Additive Models for Location, Scale and Shape." *Journal of the Royal Statistical Society. Series C: Applied Statistics*, **54**(3), 507–554. ISSN 00359254. doi:10.1111/j.1467-9876.2005.00510.x.

Scheuerer M, Büermann L (2014). "Spatially Adaptive Post-Processing of Ensemble Forecasts for Temperature." *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **63**(3), 405–422. doi:10.1111/rssc.12040.

Stasinopoulos DM, Rigby RA (2007). "Generalized Additive Models for Location Scale and Shape (GAMLSS) in R." *Journal of Statistical Software*, **23**(7), 1–46. doi:10.18637/jss.v023.i07.

Stauffer R, Messner JW, Mayr GJ, Umlauf N, Zeileis A (2016). "Ensemble Post-Processing of Daily Precipitation Sums over Complex Terrain Using Censored High-Resolution Standardized Anomalies." *Working Papers*, Faculty of Economics and Statistics, University of Innsbruck. URL http://EconPapers.repec.org/RePEc:inn:wpaper:2016-21.

Wilks DS (2011). *Statistical Methods in the Atmospheric Sciences*, volume 100. Academic Press.

Affiliation:

Markus Dabernig, Georg J. Mayr, Jakob W. Messner
Institute of Atmospheric and Cryospheric Sciences
Faculty of Geo- and Atmospheric Sciences
Universität Innsbruck
Innrain 52
6020 Innsbruck, Austria
E-mail: Markus.Dabernig@uibk.ac.at, Georg.Mayr@uibk.ac.at, Jakob.Messner@uibk.ac.at

Jakob W. Messner, Achim Zeileis
Department of Statistics
Faculty of Economics and Statistics
Universität Innsbruck
Universitätsstraße 15
6020 Innsbruck, Austria
E-mail: Jakob.Messner@uibk.ac.at, Achim.Zeileis@uibk.ac.at

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

http://eeecon.uibk.ac.at/wopec/

2016-31 **Markus Dabernig, Georg J. Mayr, Jakob W. Messner, Achim Zeileis:** Simultaneous ensemble post-processing for multiple lead times with standardized anomalies

2016-30 **Alexander Razen, Stefan Lang:** Random scaling factors in Bayesian distributional regression models with an application to real estate data

2016-29 **Glenn Dutcher, Daniela Glätzle-Rützler, Dmitry Ryvkin:** Don't hate the player, hate the game: Uncovering the foundations of cheating in contests

2016-28 **Manuel Gebetsberger, Jakob W. Messner, Georg J. Mayr, Achim Zeileis:** Tricks for improving non-homogeneous regression for probabilistic precipitation forecasts: Perfect predictions, heavy tails, and link functions

2016-27 **Michael Razen, Matthias Stefan:** Greed: Taking a deadly sin to the lab

2016-26 **Florian Wickelmaier, Achim Zeileis:** Using recursive partitioning to account for parameter heterogeneity in multinomial processing tree models

2016-25 **Michel Philipp, Carolin Strobl, Jimmy de la Torre, Achim Zeileis:** On the estimation of standard errors in cognitive diagnosis models

2016-24 **Florian Lindner, Julia Rose:** No need for more time: Intertemporal allocation decisions under time pressure

2016-23 **Christoph Eder, Martin Halla:** The long-lasting shadow of the allied occupation of Austria on its spatial equilibrium

2016-22 **Christoph Eder:** Missing men: World War II casualties and structural change

2016-21 **Reto Stauffer, Jakob Messner, Georg J. Mayr, Nikolaus Umlauf, Achim Zeileis:** Ensemble post-processing of daily precipitation sums over complex terrain using censored high-resolution standardized anomalies

2016-20 **Christina Bannier, Eberhard Feess, Natalie Packham, Markus Walzl:** Incentive schemes, private information and the double-edged role of competition for agents

2016-19 **Martin Geiger, Richard Hule:** Correlation and coordination risk

2016-18 **Yola Engler, Rudolf Kerschbamer, Lionel Page:** Why did he do that? Using counterfactuals to study the effect of intentions in extensive form games

2016-17 **Yola Engler, Rudolf Kerschbamer, Lionel Page:** Guilt-averse or reciprocal? Looking at behavioural motivations in the trust game

2016-16 **Esther Blanco, Tobias Haller, James M. Walker:** Provision of public goods: Unconditional and conditional donations from outsiders

2016-15 **Achim Zeileis, Christoph Leitner, Kurt Hornik:** Predictive bookmaker consensus model for the UEFA Euro 2016

2016-14 **Martin Halla, Harald Mayr, Gerald J. Pruckner, Pilar García-Gómez:** Cutting fertility? The effect of Cesarean deliveries on subsequent fertility and maternal labor supply

2016-13 **Wolfgang Frimmel, Martin Halla, Rudolf Winter-Ebmer:** How does parental divorce affect children's long-term outcomes?

2016-12 **Michael Kirchler, Stefan Palan:** Immaterial and monetary gifts in economic transactions. Evidence from the field

2016-11 **Michel Philipp, Achim Zeileis, Carolin Strobl:** A toolkit for stability assessment of tree-based learners

2016-10 **Loukas Balafoutas, Brent J. Davis, Matthias Sutter:** Affirmative action or just discrimination? A study on the endogenous emergence of quotas *forthcoming in Journal of Economic Behavior and Organization*

2016-09 **Loukas Balafoutas, Helena Fornwagner:** The limits of guilt

2016-08 **Markus Dabernig, Georg J. Mayr, Jakob W. Messner, Achim Zeileis:** Spatial ensemble post-processing with standardized anomalies

2016-07 **Reto Stauffer, Jakob W. Messner, Georg J. Mayr, Nikolaus Umlauf, Achim Zeileis:** Spatio-temporal precipitation climatology over complex terrain using a censored additive regression model

2016-06 **Michael Razen, Jürgen Huber, Michael Kirchler:** Cash inflow and trading horizon in asset markets

2016-05 **Ting Wang, Carolin Strobl, Achim Zeileis, Edgar C. Merkle:** Score-based tests of differential item functioning in the two-parameter model

2016-04 **Jakob W. Messner, Georg J. Mayr, Achim Zeileis:** Non-homogeneous boosting for predictor selection in ensemble post-processing

2016-03 **Dietmar Fehr, Matthias Sutter:** Gossip and the efficiency of interactions

2016-02 **Michael Kirchler, Florian Lindner, Utz Weitzel:** Rankings and risk-taking in the finance industry

2016-01 **Sibylle Puntscher, Janette Walde, Gottfried Tappeiner:** Do methodical traps lead to wrong development strategies for welfare? A multilevel approach considering heterogeneity across industrialized and developing countries

2015-16 **Niall Flynn, Christopher Kah, Rudolf Kerschbamer:** Vickrey Auction vs BDM: Difference in bidding behaviour and the impact of other-regarding motives

2015-15 **Christopher Kah, Markus Walzl:** Stochastic stability in a learning dynamic with best response to noisy play

2015-14 **Matthias Siller, Christoph Hauser, Janette Walde, Gottfried Tappeiner:** Measuring regional innovation in one dimension: More lost than gained?

2015-13 **Christoph Hauser, Gottfried Tappeiner, Janette Walde:** The roots of regional trust

2015-12 **Christoph Hauser:** Effects of employee social capital on wage satisfaction, job satisfaction and organizational commitment

2015-11 **Thomas Stöckl:** Dishonest or professional behavior? Can we tell? A comment on: Cohn et al. 2014, Nature 516, 86-89, "Business culture and dishonesty in the banking industry"

2015-10 **Marjolein Fokkema, Niels Smits, Achim Zeileis, Torsten Hothorn, Henk Kelderman:** Detecting treatment-subgroup interactions in clustered data with generalized linear mixed-effects model trees

2015-09 **Martin Halla, Gerald Pruckner, Thomas Schober:** The cost-effectiveness of developmental screenings: Evidence from a nationwide programme *forthcoming in Journal of Health Economics*

2015-08 **Lorenz B. Fischer, Michael Pfaffermayr:** The more the merrier? Migration and convergence among European regions

2015-07 **Silvia Angerer, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** Cooperation and discrimination within and across language borders: Evidence from children in a bilingual city *forthcoming in European Economic Review*

2015-06 **Martin Geiger, Wolfgang Luhan, Johann Scharler:** When do Fiscal Consolidations Lead to Consumption Booms? Lessons from a Laboratory Experiment *forthcoming in Journal of Economic Dynamics and Control*

2015-05 **Alice Sanwald, Engelbert Theurl:** Out-of-pocket payments in the Austrian healthcare system - a distributional analysis

2015-04 **Rudolf Kerschbamer, Matthias Sutter, Uwe Dulleck:** How social preferences shape incentives in (experimental) markets for credence goods *forthcoming in Economic Journal*

2015-03 **Kenneth Harttgen, Stefan Lang, Judith Santer:** Multilevel modelling of child mortality in Africa

2015-02 **Helene Roth, Stefan Lang, Helga Wagner:** Random intercept selection in structured additive regression models

2015-01 **Alice Sanwald, Engelbert Theurl:** Out-of-pocket expenditures for pharmaceuticals: Lessons from the Austrian household budget survey

University of Innsbruck

Working Papers in Economics and Statistics

2016-31

Markus Dabernig, Georg J. Mayr, Jakob W. Messner, Achim Zeileis

Simultaneous ensemble post-processing for multiple lead times with standardized anomalies

Abstract
Statistical post-processing of ensemble predictions is usually adjusted to a particular lead time so that several models must be fitted to forecast multiple lead times. To increase the coherence between lead times, we propose to use standardized anomalies instead of direct observations and predictions. By subtracting a climatological mean and dividing by the climatological standard deviation, lead-time-specific characteristics are eliminated and several lead times can be forecasted simultaneously. The results show that forecasts between +12 and +120 h can be fitted together with a comparable forecast skill to a conventional method. Furthermore, forecasts can be produced with a temporal resolution as high as the observation interval e.g., up to ten minutes.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)