

Stauffer, Reto; Mayr, Georg J.; Messner, JakobW.; Zeileis, Achim

Working Paper

Hourly probabilistic snow forecasts over complex terrain: A hybrid ensemble postprocessing approach

Working Papers in Economics and Statistics, No. 2018-05

Provided in Cooperation with:

Institute of Public Finance, University of Innsbruck

Suggested Citation: Stauffer, Reto; Mayr, Georg J.; Messner, JakobW.; Zeileis, Achim (2018) : Hourly probabilistic snow forecasts over complex terrain: A hybrid ensemble postprocessing approach, Working Papers in Economics and Statistics, No. 2018-05, University of Innsbruck, Research Platform Empirical and Experimental Economics (eeecon), Innsbruck

This Version is available at:

<https://hdl.handle.net/10419/184983>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Hourly probabilistic snow forecasts over complex terrain: A hybrid ensemble postprocessing approach

Reto Stauffer, Georg J. Mayr, Jakob W. Messner, Achim Zeileis

Working Papers in Economics and Statistics

2018-05



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 71022
Fax: + 43 512 507 2970
E-mail: eeecon@uibk.ac.at

The most recent version of all working papers can be downloaded at
<https://www.uibk.ac.at/eeecon/wopec/>

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

Hourly Probabilistic Snow Forecasts over Complex Terrain: A Hybrid Ensemble Postprocessing Approach

Reto Stauffer
Universität Innsbruck

Georg J. Mayr
Universität Innsbruck

Jakob W. Messner
Technical University of Denmark

Achim Zeileis
Universität Innsbruck

Abstract

Accurate and high-resolution snowfall and fresh snow forecasts are important for a range of economic sectors as well as for the safety of people and infrastructure, especially in mountainous regions. In this article a new hybrid statistical postprocessing method is proposed, which combines standardized anomaly model output statistics (SAMOS) with ensemble copula coupling (ECC) and a novel re-weighting scheme to produce spatially and temporally high-resolution probabilistic snow forecasts. Ensemble forecasts and hindcasts of the European Centre for Medium-Range Weather Forecasts (ECMWF) serve as input for the statistical postprocessing method, while measurements from two different networks provide the required observations.

This new approach is applied to a region with very complex topography in the Eastern European Alps. The results demonstrate that the new hybrid method not only allows to provide reliable high-resolution forecasts, it also allows to combine different data sources with different temporal resolutions to create hourly probabilistic and physically consistent predictions.

Keywords: meteorology, ensemble postprocessing, standardized anomalies, copula coupling, high-resolution, fresh snow, snowfall.

1. Introduction

Large parts of our daily social and economic life strongly rely on weather forecasts. In this article we focus on the governmental area of Tyrol, Austria, which is located in the Eastern Alps and consists of a large number of narrow valleys surrounded by high mountains. The economic backbone of Tyrol is tourism with more than 5.3 million visitors and more than 25 million overnight stays recorded during the winter season 2013/2014 ([Amt der Tiroler Landesregierung 2014](#)). In winter tourism strongly focuses on alpine outdoor sports such as skiing and back-country skiing, for which resorts and skiing areas need sufficient amounts of snow and good snow conditions. On the other hand, the “white gold” can also cause hazardous situations. During the winter seasons 2009–2016 145 people died in avalanche accidents in Austria ([Lawinenwarndienst Tirol 2009–2017](#)), of which more than half of all

events and deaths occurred in Tyrol. Furthermore, severe snow events can obstruct traffic on roads, train tracks and at airports. Accurate and reliable forecasts of fresh snow and snowfall for the region of Tyrol are therefore of high importance for the public and also for decision makers or warning services (see e.g., [Zhu *et al.* 2002](#); [Palmer 2002](#); [Neal *et al.* 2014](#); [Knox *et al.* 2015](#); [Raftery 2016](#)).

Weather forecasts are typically provided by numerical weather prediction (NWP) models predicting the future atmospheric state on a global or regional scale. Due to different influencing factors such as the model resolution, necessary approximations and parameterizations but also imperfect initial conditions and the chaotic behavior of the atmosphere these forecasts are never fully exact. Ensemble prediction systems (EPSs) address these issues by running several independent forecasts for the same day using different and slightly perturbed initial conditions and model formulations to provide valuable additional information about the uncertainty of a specific weather forecast. However, it has been shown ([Hagedorn *et al.* 2012](#); [Mullen and Buizza 2001](#)) that EPS forecasts typically show too little spread on a sub-grid scale and require additional correction to enhance the predictive skill for specific locations. One widely accepted procedure to reduce possible forecast errors and to adjust the uncertainty information is statistical ensemble postprocessing. Statistical postprocessing methods use historical weather forecasts and the corresponding observations to detect and correct possible systematic EPS errors.

A wide range of different ensemble postprocessing methods have been proposed including analog methods ([Hamill *et al.* 2006, 2015](#)), ensemble dressing methods ([Roulston and Smith 2003](#)), Bayesian model averaging (BMA; [Sloughter *et al.* 2007](#); [Fraley *et al.* 2010](#)), extended logistic regression ([Wilks 2009](#); [Bouallègue and Theis 2014](#); [Messner *et al.* 2014b](#)), or distributional regression methods. Distributional regression models optimize the parameters of a pre-specified response distribution to correct for both errors in the mean and errors in the uncertainty given a set of covariates. One of the earliest and most well known approach is the ensemble model output statistics (EMOS) approach first published by [Gneiting *et al.* \(2005\)](#) applied to near-surface temperature. This approach has further been extended by [Thorarinsdottir and Gneiting 2010](#), [Lerch and Thorarinsdottir 2013](#), [Scheuerer 2014](#), [Scheuerer and Hamill 2015](#), [Messner *et al.* 2014a](#), [Scheuerer 2014](#), [Scheuerer and Hamill 2015](#) and many others for different meteorological quantities using different response distributions and optimization approaches.

Originally distributional regression was only applied to specific locations but has also been extended for spatial and even spatio-temporal corrections of the ensemble forecasts. Many of these extensions are based on anomalies ([Scheuerer and Büermann 2014](#)) or standardized anomalies ([Dabernig *et al.* 2017](#); [Stauffer *et al.* 2017b](#)) to account for location-specific characteristics in mean and variance and create corrected and fully probabilistic spatial predictions of temperature and daily precipitation sums over potentially complex terrain.

In terms of snow prediction several difficulties have to be considered. The availability and quality of good and reliable snow observations is sparse, even in the region of Tyrol. Measuring snow can be tricky due to possible snow drift, melting processes, or liquid water input (rain) between two observation times which can yield large measurement errors ([Rasmussen *et al.* 2012](#)). Overall, the amount and quality of snow measurements make it very difficult to train reliable spatial postprocessing models.

An alternative approach to forecasting fresh snow amounts is to make use of precipitation and temperature forecasts. Rather than predicting fresh snow amounts directly, they can be derived from postprocessed precipitation and temperature forecasts. Temperature forecasts are on the one hand required to determine whether precipitation reaches the ground as rain or snow and on the other hand to estimate the snow density. Snow density and its alteration are affected by the prevalence of inversions, additional cooling effects due to melting and evaporation of hydrometeors, and other local effects and is thus an extremely complex issue itself. For simplicity we will only regard the problem of whether precipitation occurs as snow or rain and assume that precipitation will fall as snow as soon as the 2 meter dry air temperature falls below +1.2 degrees Celsius, which is a threshold used in literature for the European Alps (Rohregger 2008; Bellaire *et al.* 2011).

A major challenge of converting probabilistic precipitation and temperature forecasts into fresh snow predictions are the very different temporal resolutions of ensemble predictions, temperature observations, and precipitation observations. European Centre for Medium-Range Weather Forecast (ECMWF) hindcast and EPS forecasts, which we use in this study, have a temporal resolution of 6 and 1 hours, respectively, temperature observations are usually available hourly, while precipitation or snow heights are often only measured once or few times a day.

In this article we propose a new hybrid approach that combines standardized anomaly model output statistics (SAMOS; Dabernig *et al.* 2017; Stauffer *et al.* 2017b) with ensemble copula coupling (ECC; Schefzik *et al.* 2013) and a novel re-weighting scheme to combine these data to:

- i create full probabilistic spatial predictions,
- ii provide probabilistic temperature and precipitation forecasts on an hourly temporal scale,
- iii create a physically consistent copula (pairs of temperature and precipitation) which can be used to
- iv create spatially and temporally high-resolution snowfall and fresh snow amount forecasts.

The structure of this article is as follows. Section 2 introduces the different statistical methods required to achieve the desired goal. The methods section is followed by the description of the different data sets used in this study (Section 3) and the explicit specification of the statistical models (Section 4) used in the results section (Section 5). At the end the results and limitations of this approach will be discussed (Section 6).

2. Methods

This section contains the three methodological blocks required to create probabilistic snow forecasts. Distributional regression is explained in Section 2.1 followed by the required extensions for the standardized anomaly model output statistics (SAMOS) in Section 2.2. Section 2.3 shows the ensemble copula coupling (ECC) approach to generate a postprocessed ensemble followed by the re-weighting procedure required to transform daily precipitation

sums into hourly predictions. Finally, hourly temperature and precipitation sums will be converted into probabilities of snowfall and fresh snow amounts in Section 2.5.

2.1. Distributional Regression

Statistical methods considering all parameters of a specific response distribution can be summarized as “distributional regression models”. The EMOS for temperature using a normal response distribution as originally suggested by Gneiting *et al.* (2005) can be seen as a classical example of this family.

Imagine a time series of 2 m temperature observations $\mathbf{y} = \{y_i\}_{i=1,\dots,N}$ at a specific site and the corresponding ensemble forecasts of the 2 m temperature from the EPS $\mathbf{x} = \{x_{im}\}_{i=1,\dots,N}^{m=1,\dots,M}$ where N denotes the total sample size of the data set and M the number of ensemble members. x_{im} is the individual 2 m temperature prediction of the NWP for date/time i of member m . The classical EMOS is then specified as:

$$y_i \sim \mathcal{N}(\mu_i, \sigma_i) \quad (1)$$

$$\mu_i = \beta_0 + \beta_1 \cdot \bar{\mathbf{x}}_i \quad (2)$$

$$\log(\sigma_i) = \gamma_0 + \gamma_1 \cdot \langle \mathbf{x}_i \rangle \quad (3)$$

The response y_i is assumed to follow a normal distribution \mathcal{N} with the two distributional parameters μ_i (location or mean) and σ_i (scale or standard deviation). Both parameters are expressed by a linear predictor including an intercept (β_0/γ_0) and a slope coefficient (β_1/γ_1) for a covariate. While the location μ_i depends on the ensemble mean $\bar{\mathbf{x}}_i$ over all members $m = 1, \dots, M$ for each individual sample i the log-scale depends on the logarithm of the corresponding ensemble standard deviation denoted as $\langle \mathbf{x}_i \rangle$. The log-link on σ_i ensures positive variance in predictions.

The coefficients $\theta = (\beta_0, \beta_1, \gamma_0, \gamma_1)$ can be estimated by using an appropriate M-estimator such as the maximum-likelihood estimator maximising the likelihood:

$$\hat{\theta} = \operatorname{argmax}_{\theta} \left(\prod_{i=1}^N \phi \left(\frac{y_i - \mu_i}{\sigma_i} \right) \right), \quad (4)$$

where $\phi \left(\frac{y_i - \mu_i}{\sigma_i} \right)$ denotes the standard normal probability density function (PDF) evaluated at each individual $i = 1, \dots, N$ in the data set.

For the daily precipitation sums the model shown in Equations 1-3 can be improved by replacing the response distribution and adding an additional covariate z which allows to account for EPS forecasts where the majority of all EPS members predict no precipitation. Following the work of Gebetsberger *et al.* (2017) and Stauffer *et al.* (2017a) the model specification can be written as follows:

$$y_i^{1/p} = \mathcal{L}_0(\mu_i, \sigma_i) \quad (5)$$

$$\mu_i = \beta_0 + \beta_1 \cdot \overline{\mathbf{x}_i^{1/p}} \cdot (1 - z_i) + \beta_2 \cdot z_i \quad (6)$$

$$\log(\sigma_i) = \gamma_0 + \gamma_1 \cdot \langle \mathbf{x}_i^{1/p} \rangle \cdot (1 - z_i). \quad (7)$$

The power-transformed observations y_i are assumed to follow a left-censored logistic distribution \mathcal{L}_0 censored at 0 and a power parameter $p = 1.35$. The additional covariate z_i takes 1 if 80 % or more of all ensemble members predict less than 0.05 mm over 24 h and 0 otherwise and is used to handle unanimous predictions (cf. Gebetsberger *et al.* 2017). The corresponding M-estimator can be written as

$$\hat{\theta} = \operatorname{argmax}_{\theta} \left(\prod_{i=1}^N f \left(\frac{y_i^{1/p} - \mu_i}{\sigma_i} \right) \right)$$

$$\text{with } f = \begin{cases} \Lambda \left(\frac{-\mu_i}{\sigma_i} \right) & \text{if } y_i = 0 \\ \lambda \left(\frac{y_i^{1/p} - \mu_i}{\sigma_i} \right) & \text{else} \end{cases} \quad (8)$$

where λ is the probability density function and Λ the cumulative distribution function of the standard logistic distribution.

2.2. Standardized Anomaly Model Output Statistics (SAMOS)

While the model specifications in Equations 1-3 & 5-7 work well for single stations, an extension is required for spatial and/or spatio-temporal ensemble postprocessing. In the following, we will employ the SAMOS approach (Dabernig *et al.* 2017; Stauffer *et al.* 2017b) for this purpose. Its basic idea is to remove location and time specific characteristics from the observation and EPS data by transforming them into standardized anomalies. This transformation then allows to fit a single postprocessing model that is valid for the whole area and all-season and can thus be applied to any location and time.

Standardized anomalies of the observations (y^*) and EPS forecasts (\mathbf{x}^*) will be characterized by a superscript asterisk from here on and are defined as:

$$y_i^* = \frac{y_i - \tilde{\mu}_{y,i}}{\tilde{\sigma}_{y,i}} \quad \text{and} \quad \mathbf{x}_i^* = \frac{\mathbf{x}_i - \tilde{\boldsymbol{\mu}}_{x,i}}{\tilde{\boldsymbol{\sigma}}_{x,i}}. \quad (9)$$

$\tilde{\mu}_{\bullet,i}$ and $\tilde{\sigma}_{\bullet,i}$ are the estimates of the climatological location and scale for each required quantity and depend on the location and season respectively. A comprehensive description of how these climatologies are specified can be found in Appendix A. Once the climatologies and thus the standardized anomalies are known, the SAMOS regression coefficients can be estimated using Equation 1-8 by simply replacing y_i and \mathbf{x}_i by their corresponding standardized anomalies y_i^* and \mathbf{x}_i^* (except in the condition in Equation 8 where $y_i = 0$ is not replaced). Given a new EPS forecast, the postprocessed predictions can be obtained by applying the SAMOS correction and transforming the results back to the original scale (e.g., °C or mm). Algorithm 1 contains the pseudo-code for the SAMOS procedure as used for this article.

2.3. Ensemble Copula Coupling

The SAMOS procedure (Section 2.2) provides postprocessed probabilistic predictions for 2 m temperature as well as corrected probabilistic forecasts for 24 h precipitation sums. Due to the model specification, SAMOS allows to retrieve predictions for any arbitrary location within the area of interest (spatial prediction) and even for all forecast steps covered by the training data set (temporal predictions). This allows to create forecasts for +30/+54/+78 h for the 24 h precipitation sums, and hourly forecasts for 2 m temperature for the whole study area.

1. Observations:

i) Estimate spatio-temporal climatology

 $\tilde{\mu}_{y,i}, \tilde{\sigma}_{y,i}$ for the response

ii) Compute standardized anomalies

$$y_i^* = (y_i - \tilde{\mu}_{y,i}) / \tilde{\sigma}_{y,i}$$

2. Numerical weather forecasts:

i) Compute model climatology $\tilde{\mu}_{x,i}, \tilde{\sigma}_{x,i}$

based on hindcasts

ii) Interpolate hindcasts and model climatologies to station location

iii) Compute standardized anomalies for the covariates

$$x_i^* = (x_i - \tilde{\mu}_{x,i}) / \tilde{\sigma}_{x,i}$$

3. SAMOS estimation:

i) Combine y_i^* and \mathbf{x}_i^* , $i = 1, \dots, N$ to generate the training data setii) Estimate the statistical model $y_i^* \sim \mathcal{D}(\mu_i^*, \sigma_i^*)$, where \mathcal{D} is either \mathcal{N} or \mathcal{L}_0

4. Prediction given a new EPS forecast:

i) Compute standardized anomalies of new EPS forecast $\hat{\mathbf{x}}$ $\hat{x}_j^* = (\hat{x}_j - \tilde{\mu}_{x,j}) / \tilde{\sigma}_{x,j}$, where the index j describes the location and season of $\hat{\mathbf{x}}$

ii) Correct forecast anomalies using the estimated model parameters from Step 3

iii) Transform corrected $\hat{y}_j^* \rightarrow \hat{y}_j$ via

$$\hat{y}_j \sim \mathcal{D}(\mu_j^* \cdot \tilde{\sigma}_{y,j} + \tilde{\mu}_{y,j}, \sigma_j^* \cdot \tilde{\sigma}_{y,j})$$

Algorithm 1: Pseudo-code of the required steps for SAMOS using hindcasts for model training and latest EPS forecasts for prediction. A detailed description and a graphical representation of the algorithm can be found in Appendix A.

In order to retrieve probabilistic snowfall forecasts from the SAMOS predictions, the marginal predictive distributions of temperature and precipitation have to be combined such that correlations between them are considered. This can be achieved by using ensemble copula coupling (ECC) proposed by Schefzik *et al.* (2013). The basic idea is to restore the physical coupling between two or more quantities based on the raw EPS forecasts. As numerical predictions are based on physically consistent prognostic equations, each EPS member provides a distinct physically meaningful combination of temperature and precipitation. This property is lost during the SAMOS postprocessing since both quantities are corrected independently. However, the coupling can be restored by drawing a sample of the postprocessed predictive distributions and rearranging the sampled values in the rank order structure of the original EPS forecasts.

There are different ways to draw a new sample from the postprocessed distributions. It turned out (not shown) that the quantile approach (ECC-Q; Schefzik *et al.* 2013) yields best and most stable results for this application, which supports the findings of Schefzik *et al.* (2013). For ECC-Q, a set of $M = 50 + 1$ (ensemble size of ECMWF EPS) quantiles based on equally spaced probabilities $p = \frac{1}{M+1}, \dots, \frac{M}{M+1}$ are drawn from the marginal predictive distribution.

2.4. Precipitation Re-Weighting

Temperature and precipitation observation data are based on two different observational networks with different temporal resolutions. The 2 m temperature observations are available hourly while precipitation sums are only reported once a day (details in Section 3.3). This temporal resolution is maintained by the SAMOS postprocessing so that it also differs for the forecasts of the different quantities.

However, temperature has a clear diurnal cycle and thus the time of precipitation can highly affect the precipitation phase and thus the total fresh snow amount. Therefore, the precipitation forecasts have to be temporally downscaled before they can be combined with the temperature forecasts. For this purpose, we extend ECC (Section 2.3) with a novel re-weighting scheme where the daily precipitation sums are allocated to the hours of the day according to the time series of the raw EPS predictions. E.g., if an EPS member predicts 10% of its daily precipitation to fall between 10:00 am and 11:00 am, 10% of the corresponding precipitation forecast is allocated to this hour. Thus each of the $M = 50 + 1$ ECC-Q draws from the marginal precipitation distribution can be downscaled to an hourly temporal resolution and then combined with the respective draws from the marginal temperature distribution. Algorithm 2 shows the procedure how to draw a new copula and perform the re-weighting for the temporal downscaling.

1. Get 24 h precipitation sums from raw EPS: $(tp_1, \dots, tp_m, \dots, tp_M)$.
2. Draw a copula $(\hat{y}_1, \dots, \hat{y}_m, \dots, \hat{y}_M)$ of 24 h postprocessed precipitation sums using ECC-Q drawing from the predictive distribution $\mathcal{L}_0(\mu, \sigma)^{1.35}$ returned by SAMOS.
3. Compute correction weighting factors $\omega = (\hat{y}_1/tp_1, \dots, \hat{y}_m/tp_m, \dots, \hat{y}_M/tp_M)$.
4. Correct hourly EPS member forecasts using the weights ω such that the sum over 24 hourly precipitation sums of a specific copula member m is equal to \hat{y}_m (or $\omega_m \cdot tp_m$).

Algorithm 2: Re-weighting pseudo-code for temporal downscaling of probabilistic precipitation forecasts. Generate a new 50+1 member copula with an hourly temporal resolution from postprocessed probabilistic daily precipitation sum forecasts provided by SAMOS.

For stability reasons, the weights ω are computed using values for \hat{y}_m and tp_m rounded to two digits ($\frac{1}{100}$ mm d⁻¹) to avoid weights close to infinity. If \hat{y}_m or tp_m is 0, the corresponding weight is set to 0 as well. As a side note it has to be mentioned that due to the re-weighting, the ranks of the hourly copulas are no longer strictly preserved and might sometimes differ from the original rank structure of the EPS.

2.5. New Snow Amount and Probability of Snow

Once ECC-Q and re-weighting are applied to the marginal distributions, bi-variate time series of calibrated hourly precipitation sums and 2 m temperatures are available for each of the M ensemble members. For each individual pair of member m and forecast step s the “snow indicator function” SI_{ms} can be retrieved.

$$SI_{ms} = \begin{cases} \text{“dry” if:} \\ \quad \text{precipitation}_{ms} \leq 0.05 \frac{mm}{h} \\ \text{“rain” if:} \\ \quad \text{precipitation}_{ms} > 0.05 \frac{mm}{h} \wedge T_{2m} > 1.2^\circ C \\ \text{“snow” if:} \\ \quad \text{precipitation}_{ms} > 0.05 \frac{mm}{h} \wedge T_{2m} \leq 1.2^\circ C \end{cases} \quad (10)$$

The threshold of 0.05 mm h^{-1} has been chosen as the smallest recorded value of the rain gauges used for validation is 0.1 mm. To distinguish between rain and snow we use a fixed threshold of $1.2^\circ C$ as a rough approximation, following [Bellaire *et al.* \(2011, p. 1121\)](#). The empirical probabilities π_{cs} for each of the three classes (snow, rain, and dry, which are mutually exclusive for each individual member and forecast time step) or for combinations can be computed using:

$$\pi_{cs} = \frac{1}{M} \sum_{m=1}^M \mathbf{1}(SI_{ms} = c), \quad (11)$$

where s is a specific forecast step and c is the desired class (e.g., snow, rain, rain \vee snow). $\mathbf{1}(\cdot)$ is an indicator function which takes 1 if the argument in brackets is true or 0 otherwise. The conditional expectation can be derived similarly:

$$E[c] = \frac{\sum_{i=1}^M \text{precipitation}_{ms} \cdot \mathbf{1}(SI_{ms} = c)}{\sum_{i=1}^M \mathbf{1}(SI_{ms} = c)}. \quad (12)$$

If one is interested in the snow height of fresh snow ($E[\text{snow}]$ in centimeters) the snow density has to be taken into account. A rule of thumb is the “1 : 10 rule” where 1 millimeter of liquid water equivalent, the quantity forecasted by the postprocessing, corresponds to 1 centimeter of fresh snow. This is equivalent to a fresh snow density of 100 kg m^{-3} . In reality, fresh snow densities can vary strongly between 10 kg m^{-3} up to 526 kg m^{-3} given location and prevailing conditions (e.g., [Meister 1985](#); [Judson and Doesken 2000](#); [Roebber *et al.* 2003](#)). As reliable fresh snow height or density observations with the desired temporal resolution are not available for this study, a detailed verification cannot be performed. For visual representation we simply assume a mean density of 100 kg m^{-3} .

3. Data

This section presents the data sets used for this study. These consist of two different EPS forecast data sets (ECMWF hindcast and operational EPS) and three different sources of observation data for model training and verification.

3.1. Numerical Weather Prediction Data: Forecast Data

All predictions presented in this article are based on the ECMWF EPS. The ECMWF EPS consists of 50 perturbed ensemble members and 1 control run (50+1) and is initialized four times a day every six hours. For this study, the control run is treated the same way as the 50 perturbed members. We will solely focus on the 0000 UTC forecast run of EPS model version *43r1*. This version became operational on November 22, 2016 and the output is available at an hourly temporal resolution up to +90 h ahead on a 16 km·16 km grid.

The presented application will focus on the winter season 2016/2017 (December 1, 2016 through April 15, 2017) and on predictions from +6 h to +78 h in advance, spanning the first three days after EPS initialization (0600 UTC to 0600 UTC of three consecutive days).

3.2. Numerical Weather Prediction Data: Training Data

For training the SAMOS models we use ECMWF hindcasts, similar to the approach of [Stauffer *et al.* \(2017b\)](#). ECMWF hindcasts become available twice a week (Mondays and Thursdays) providing a 10+1 member ensemble for the same date over the previous 20 years, initialized at 0000 UTC. For example: on Monday January 2, 2017 hindcasts for January 2, 2016, 2015, ..., 1998, and 1997 become available. As for the EPS, the hindcast control run is treated as an additional member to increase the ensemble sample size. The hindcasts are available on the same spatial resolution as the EPS, but on a 6 hourly temporal resolution only.

For the statistical postprocessing methods of 2 m temperature, all 6 hourly intervals from +6 h to +78 h will be used. Beside the forecasted 2 m temperature the 2 m dew point temperature, 850 hectopascal temperature, and surface pressure forecasts are used as additional covariates (see Section 4). For precipitation, 24 h total precipitation sum hindcasts are used for the forecast time steps +30 h, +54 h, and +78 h.

3.3. Observational Data

Two major different observation networks will be used in the following. As in [Stauffer *et al.* \(2017b\)](#), daily liquid water equivalent observations from the network of the hydrographical service Tyrol (EHYD; [BMLFUW 2016](#)) are used for the postprocessing of daily precipitation sums. In comparison to other networks in the area, the hydrographical service maintains the highest density of stations (number of stations) with very long historical records (up to 47 years of data). The observation sites are well distributed up to an altitude of about 1800 m a.s.l.. However, observations are only made once a day (manually) at 0600 UTC. In the following, we will use these data for training the precipitation SAMOS models.

The second network consists of 89 automated weather stations operated by the national weather service (TAWES network; Zentralanstalt für Meteorologie und Geodynamik). 75 out of these 89 stations provide at least 6 years of data. Observations are recorded every ten minutes, of which all observations at every full hour are used for training and validation of

the 2 m air temperature SAMOS models.

The TAWES network also provides automated precipitation measurements at a 10 minute resolution. However, the length of historical records is much shorter compared to the time series provided by EHYD data set. Furthermore, the measurement errors of the automated rain gauges are expected to be larger than from the daily manual records provided by the hydrographical service, especially during winter. Thus, we decided to not use the TAWES precipitation observations for model training and for the estimates of the spatio-temporal climatologies. Nevertheless, since observations from the hydrographical service are only available up to 2012 at this time (2018), we do use TAWES precipitation observations for validation. Therefore, daily precipitation sums are generated by taking the sum over all 10 minute intervals between 0610 UTC and 0600 UTC of the following day (yields 144 10-minute values). Periods for which more than four 10-minute values are missing are eliminated.

In addition to the temperature and precipitation observations from the hydrographical service and the TAWES network, meteorological aerodrome reports (METAR) from Innsbruck Airport are used in the verification section as it is the only longer-term source of temporally high-resolution *precipitation phase* observations available. The weather conditions from the METARs are classified into “snow” (if report contains SN, SG, IC, PL, SNRA, or RASN), “rain” (if message contains DZ, RA, SNRA, or RASN), and “dry” (else). Conditions with sleet (mixed rain/snow; SNRA/RASN) are attributed to both, “snow” and “rain”. METARs are available every 30 minutes either created by a human observer or an automated procedure if the airport is closed over night. These observations have been aggregated to an hourly temporal resolution and will be used to validate the forecasted probabilities of snowfall. Overall, 3318 observations are available for the time period of interest with 333 cases reporting rain or sleet (10%), 246 cases reporting snow or sleet (7.5%), and 2786 dry cases (84%).

Figure 1 shows an overview of the area of interest. The markers show the locations of the observational sites from the two networks (TAWES, EHYD) and the location of the airport (581 m a.m.s.l.). To the right the height-distribution of the stations from the two networks is shown.

4. Statistical Models

This section presents the specifications of the models that will be compared and tested in Section 5. During the preparation of this manuscript, a variety of slightly different models formulations have been tested and the presented models are only a subset that was selected because they performed well or showed interesting results.

All models follow the approaches presented in Section 2 but differ in their input variables and whether the data are transformed to standardized anomalies. Four models will be used for 2 m temperature and three for daily precipitation sums. The training data set to estimate the regression coefficients is composed of all forecast steps provided by the ECMWF hindcasts from +6 h up to +78 h on a regular 6 h interval. For precipitation, these forecasts are aggregated to 24 h sums resulting in forecast steps +30 h, +54 h, and +78 h. The power parameter was set to $p = 1.35$, found to have the best predictive cross-validated performance in [Stauffer et al. \(2017b\)](#).

Table 1 shows the different model assumptions and naming. The first two models named *EMOS* correspond to Equation 1-8 operating on the physical scale (not on standardized

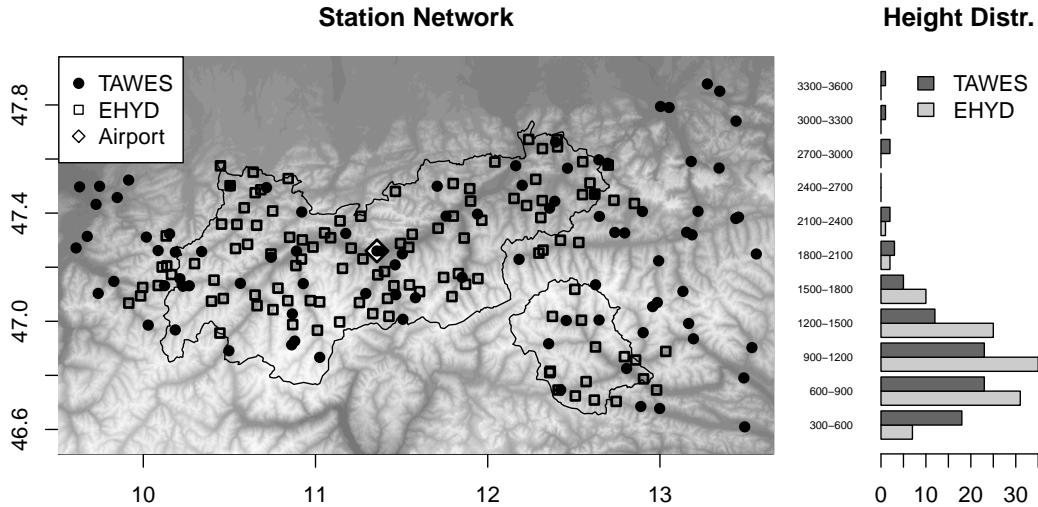


Figure 1: The left panel shows the topography of the area of interest. Overlays: governmental area of Tyrol (black outline), location of the TAWES stations (circles) and EHYD stations (squares). The airport is indicated by a diamond in the center of the map. The right panel shows the height distribution of the stations grouped into 300 m intervals: number of stations (abscissa) and altitude intervals (ordinate; meters *a.m.s.l.*).

anomalies). One crucial modification has to be made for the 2 m temperature: interactions with factors for the time of the day (*hour*; 0000/0600/1200/1800 UTC) and the station (*station*) are included to capture spatial and diurnal differences, yielding separate (and independent) coefficients for each station and each time of the day. For daily precipitation sums, this extension has not been made as only 0600 UTC observations are included (no diurnal effect required) and station-wise regression models partially returned highly unstable estimates due to the low number of observations for each single site. Please note that the *EMOS* models are *not designed for spatial or spatio-temporal predictions* even if spatial predictions would be possible in case of precipitation. These two models serve as reference for the performance of the *SAMOS* models.

All other models are spatio-temporal (in case of 2 m temperature) and spatial (in case of daily precipitation sums) *SAMOS* models operating on the standardized anomaly scale. Thus, the spatial and temporal characteristics among all stations and for all lead times are already removed from the data and do not have to be considered in the linear predictors for location μ^* and scale σ^* .

The second and third pair of models named *SAMOS_hom* and *SAMOS_het* are two *SAMOS* variations, both solely using the corresponding quantity from the EPS as covariate (i.e., 2 m temperature and total precipitation, respectively). While *SAMOS_het* is a full heteroscedastic model including the ensemble standard deviation in the linear predictor for the scale σ^* , *SAMOS_hom* is a homoscedastic model where the scale does not depend on any covariates. These two models allow to quantify the improvement in the predictive performance by including the ensemble spread information in the postprocessing methods. For 2 m temperature, a fourth model called *xSAMOS_het* (*x* for *extended*) is used, which includes additional covariates for both location μ^* and scale σ^* .

Models for 2 m temperature using a Gaussian response distribution.	Models for 24 h precipitation sums using a power-transformed left-censored logistic response distribution.
Heteroscedastic EMOS models (EMOS; cf. Equation 1-3 & 5-7)	
These models <i>are not designed to provide spatial or spatio-temporal predictions.</i>	
μ = hour / station + $\overline{T_{2m}}$ / hour / station	μ = $tp^{1/p} \cdot (1-z) + z$
$\log(\sigma)$ = hour / station + $\langle T_{2m} \rangle$ / hour / station	$\log(\sigma)$ = $\langle tp^{1/p} \rangle \cdot (1-z)$
Homoscedastic SAMOS models (SAMOS_hom)	
μ^* = $\overline{T_{2m}^*}$	μ^* = $tp^{1/p^*} \cdot (1-z) + z$
$\log(\sigma^*)$ = constant	$\log(\sigma^*)$ = constant
Heteroscedastic SAMOS models (SAMOS_het)	
μ^* = $\overline{T_{2m}^*}$	μ^* = $tp^{1/p^*} \cdot (1-z) + z$
$\log(\sigma^*)$ = $\langle T_{2m}^* \rangle$	$\log(\sigma^*)$ = $\langle tp^{1/p^*} \rangle \cdot (1-z)$
Extended Heteroscedastic SAMOS models (xSAMOS_het)	
μ^* = $\overline{T_{2m}^*} + \overline{Td_{2m}^*} + \overline{T_{850}^*} + \overline{P^*}$	—
$\log(\sigma^*)$ = $\langle T_{2m}^* \rangle + \langle Td_{2m}^* \rangle + \langle T_{850}^* \rangle + \langle P^* \rangle$	—

Table 1: Statistical model specification for 2 m temperature (left) and 24 h precipitation sums (right). For each model the linear predictors for μ and $\log(\sigma)$ are shown. Superscript asterisk indicate variables on the standardized anomaly scale (SAMOS). T_{2m} , Td_{2m} , T_{850} , P , and tp are the 2m temperature, 2m dew point temperature, temperature in 850hPa, surface pressure, and total precipitation ensemble forecasts respectively. \overline{X} are ensemble means, $\langle X \rangle$ denote ensemble log-standard deviations. X /hour and X /station are interactions between X and the ‘time of the day’ and/or the ‘station’.

5. Results

The first two subsections show the performance of the full predictive distributions of the 2m temperature (Section 5.1) and daily precipitation forecasts (Section 5.2). Section 5.3 shows the validation for hourly predictions and precipitation-type classification based on the 50+1 members generated via re-weighting and ECC. Last but not least, spatial forecasts for a specific forecast are shown in Section 5.5 to demonstrate the feasibility of high-resolution areal predictions.

5.1. Temperature (6h Intervals)

Figure 2 shows BIAS, continuous rank probability score (CRPS; Gneiting and Raftery 2007), mean width of the prediction interval between the 10% and 90% percentile, and CRPS skill scores, all based on the full predictive distribution returned by the statistical models. All results are temporally out-of-sample and validated on the TAWES network for all forecast steps +6h/+12h/.../+72h/+78h as used to train the statistical models on hindcasts. The box-whiskers show station-wise mean scores for the spatio-temporal climatology (CLIM; Equation 13), the raw EPS, and the four statistical postprocessing models (cf. Table 1).

The raw EPS performs poorly for the area of interest as the NWP model with its current spatial resolution is not able to represent the local topography. It performs even worse than the underlying climatology in terms of BIAS and CRPS. All statistical postprocessing models perform significantly better and are essentially bias-free. As expected, the station-wise statistical EMOS model performs best since it has separate model coefficients for each station location and is thus more flexible than the spatial models. In terms of CRPS, the

spatial models loose about 7–12% of skill (Figure 2d; *SAMOS_hom*: –12.3%; *SAMOS_het*: –12.3%; *xSAMOS_het*: –6.9%) but allow to predict at any arbitrary location within the area of interest and any desired time between +6 h and +78 h. The two models *SAMOS_hom* and *SAMOS_het* perform very similarly indicating that the uncertainty information from the EPS 2 m temperature forecast provides barely any additional information. Small improvements can be achieved by including additional covariates (model *xSAMOS_het*).

Overall, all statistical models show promising values in terms of CRPS (median 1.45–1.65° C) and mean absolute error (median 2.0–2.3° C; not shown) across all four methods. The median of the mean prediction interval width for the 10–90% interval is around 6.0° C for the station-wise *EMOS* model and around 6.9–7.2° C for the SAMOS models.

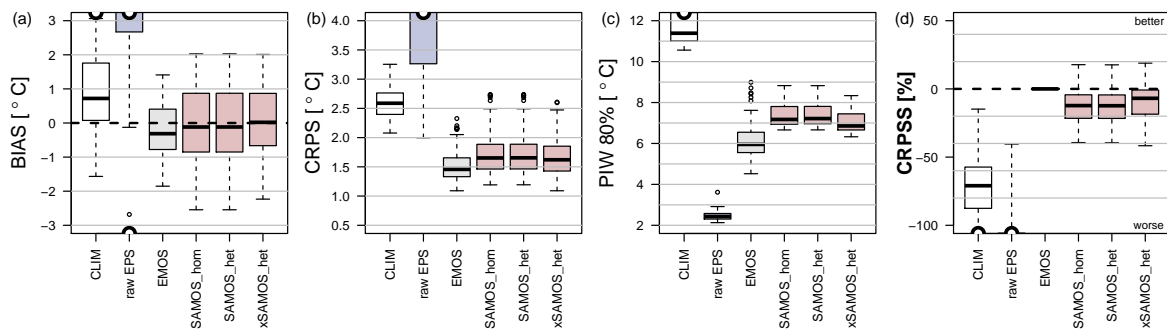


Figure 2: Scores for 2 m temperature forecasts based on the full predictive distribution based on 6 h/+12 h/.../+72 h/+78 h forecasts as used for model training. The box whisker show station-wise means for (a) BIAS, (b) CRPS, (c) width of the 80% prediction interval, and (d) CRPS skill scores with *EMOS* as reference. Scores are shown for the climatology (CLIM), raw EPS, and the four postprocessing models (cf. Table 1). Abscissa are set to manually specified ranges; the ‘semi-sphere’ marker (top/bottom) indicate data outside the plotted range.

5.2. Daily Precipitation Sums

Figure 3 shows the verification of the daily precipitation sum predictions for the forecast steps +30/+54/+78 h. Again, this analysis is based on the full predictive distribution returned by the statistical models. Here, the validation is done on different stations (TAWES) than used for model fitting (EHYD; Section 3) so that these results are spatially and temporally out of sample. The box-whiskers show station-wise mean scores for the spatio-temporal climatology (CLIM; Equation 14), the raw daily-accumulated total precipitation from the ECMWF EPS (raw EPS), and the three postprocessing methods shown in Table 1.

The top row of Figure 3 shows BIAS, CRPS, and the Brier score for probability of precipitation (BS_{0mm}). The row below shows skill scores with the raw EPS as reference. The two SAMOS models (*SAMOS_hom* and *SAMOS_het*) show the best BIAS among all methods but less predictive skill in terms of MAE, CRPS, and BS_{0mm} than the *EMOS* model not using standardized anomalies. The distinct improvements in the BS_{0mm} are expected due to the well-known wet BIAS of the EPS when comparing interpolated data (spatial scale) to a specific site (point scale). As for 2 m temperature, the use of the forecasted EPS uncertainty in the heteroscedastic model (*SAMOS_het*) produces barely any improvement.

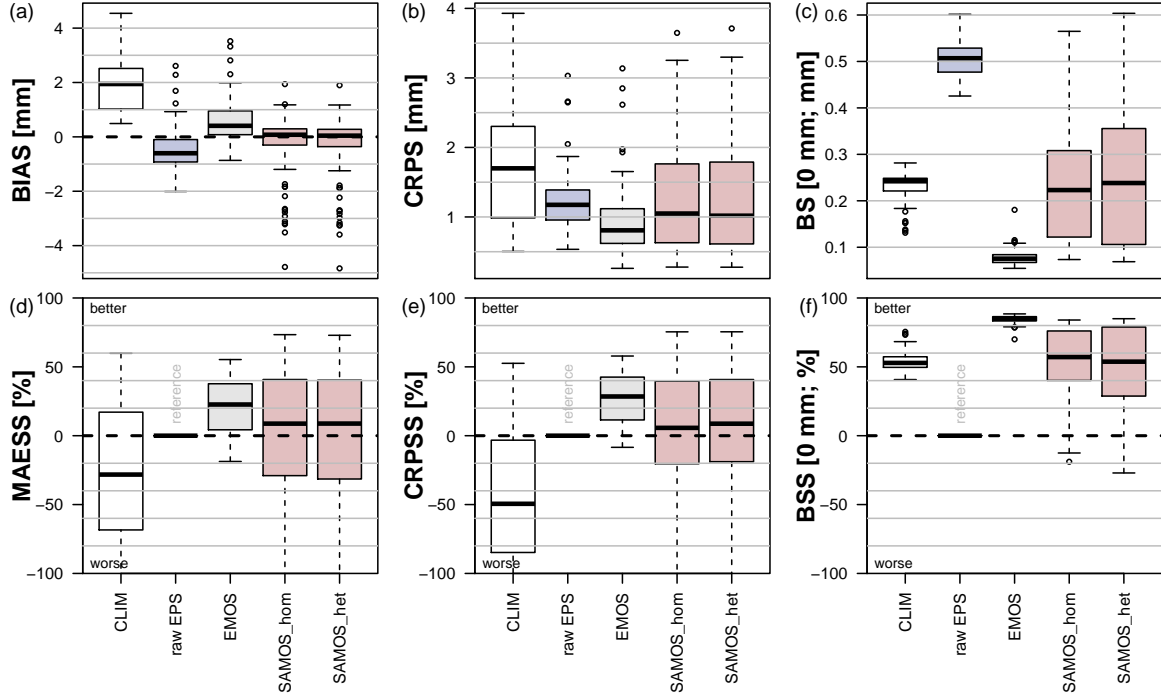


Figure 3: Scores for 24h precipitation sums based on the full predictive distribution for +30 h, +54 h, and +78 h forecasts as used for model training. Box-whiskers of station-wise mean scores for (a) BIAS, (b) CRPS, and (c) Brier scores for probability of precipitation. The scores are shown for the climatology (CLIM), raw EPS, and the three postprocessing models (cf. Table 1). The lower row shows skill scores for (d) mean absolute error, (e) CRPS, and (f) Brier score for probability of precipitation with *EMOS* as reference. Positive skill scores indicate an improvement compared to the *EMOS* model.

5.3. Hourly Temperature and Precipitation Sums

Sections 5.1 and 5.2 show that the postprocessing models are able to improve the predictive performance of the raw EPS for temperature and daily precipitation sums. The main goal of this study is to provide accurate and reliable snow predictions by combining hourly 2 m temperature and precipitation forecasts. Thus, an hourly temporal resolution for both, temperature and precipitation forecasts is required. This section therefore shows the verification of hourly forecasts. For temperature, the hourly forecasts are based on the spatio-temporal SAMOS model *xSAMOS_het* as it shows the overall best performance among all tested spatial models. The hourly precipitation sums are based on the predictions from the *SAMOS_het* model downscaled to the desired temporal resolution using the re-weighting approach presented in Section 2.4. Since the re-weighted precipitation forecasts are only available as ensembles but not as full predictive distributions, ensemble verification methods are employed in the following.

Figure 4 shows ensemble rank histograms (Hamill 2001) for hourly temperature predictions and hourly precipitation sums for the raw EPS and the postprocessed forecasts. Each observation is assigned to a rank where observations falling below the lowest member get rank 1

and observations higher than the highest member get rank 52 (50+1 members, 52 possible ranks). A perfectly uniform distribution would indicate perfect calibration. For temperature (Figure 4a+b), the postprocessing strongly improves calibration compared to the raw EPS. However, the pronounced U-shape indicates that the predicted uncertainty is lower than in reality (underdispersion). A similar picture can be seen for the hourly precipitation sums plotted as ‘stacked ensemble rank histograms’ (Figure 4c+d). The total height of the bars given the rank shows the rank histogram of the full verification data set. The faded colors show the calibration for all forecasts where at least 50 % of all members forecasted 0 mm h⁻¹ (dry cases). It can be seen that the dry cases are relatively well-calibrated and that the majority of the underdispersion results from the wet cases. Nevertheless, the asymmetry (decreasing density with increasing rank) indicates a small wet-bias also for the dry cases.

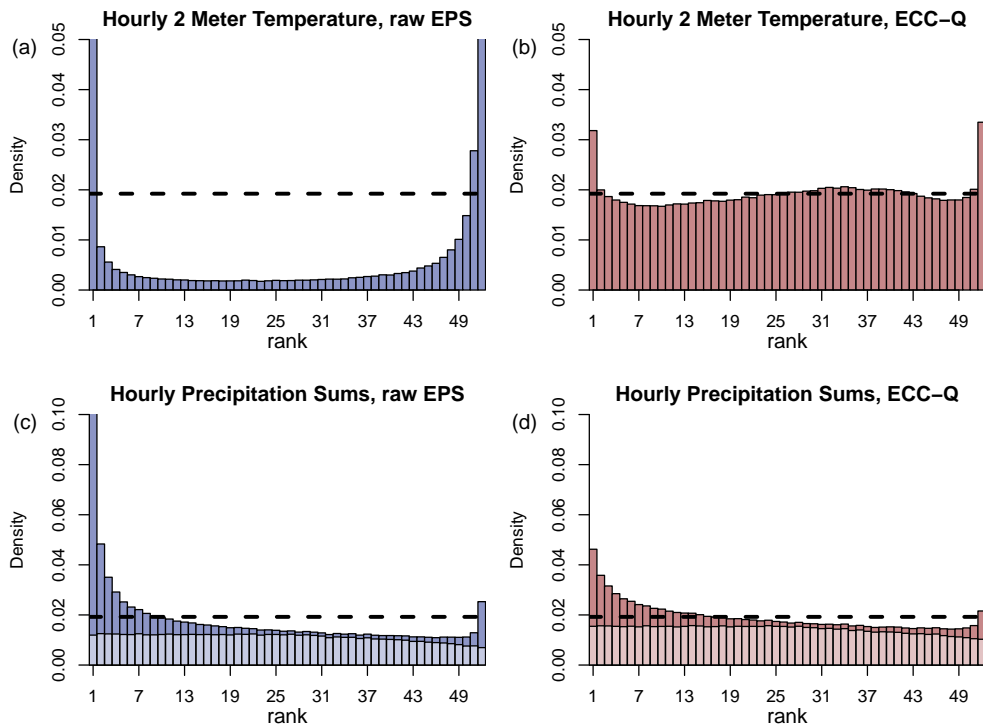


Figure 4: (Stacked) ensemble rank histograms for hourly 2 meter temperature (top row) and hourly precipitation sum forecasts (bottom row) of the raw EPS (left) and postprocessed copula (right) with 50 + 1 members each. The rank histograms contain all available forecasts for all stations and forecast steps +7h, +8h, . . . , +78h in advance. For precipitation, the faded colors show the rank histogram for all forecasts where 50% or more of all members predicted 0 mm h⁻¹. Please note that the y-axis is cut at 0.05 for temperature and 0.10 for precipitation respectively.

To investigate the predictive performance of hourly predictions for different forecast horizons, Figure 5 shows CRPS skill scores for all individual lead times. Each box-whisker contains station-wise mean skill scores over the verification period. While always on a high level, the 2m temperature forecasts for morning hours (+7–12,+31–36,+55–60; corresponds to 0700–1200 UTC) show slightly less skill. For precipitation, the skill scores are overall positive but

clearly decreasing with increasing forecast horizon. Lowest skill scores are found for early morning hours (+26–30,+50–54,+74–78; 0200–0600 UTC).

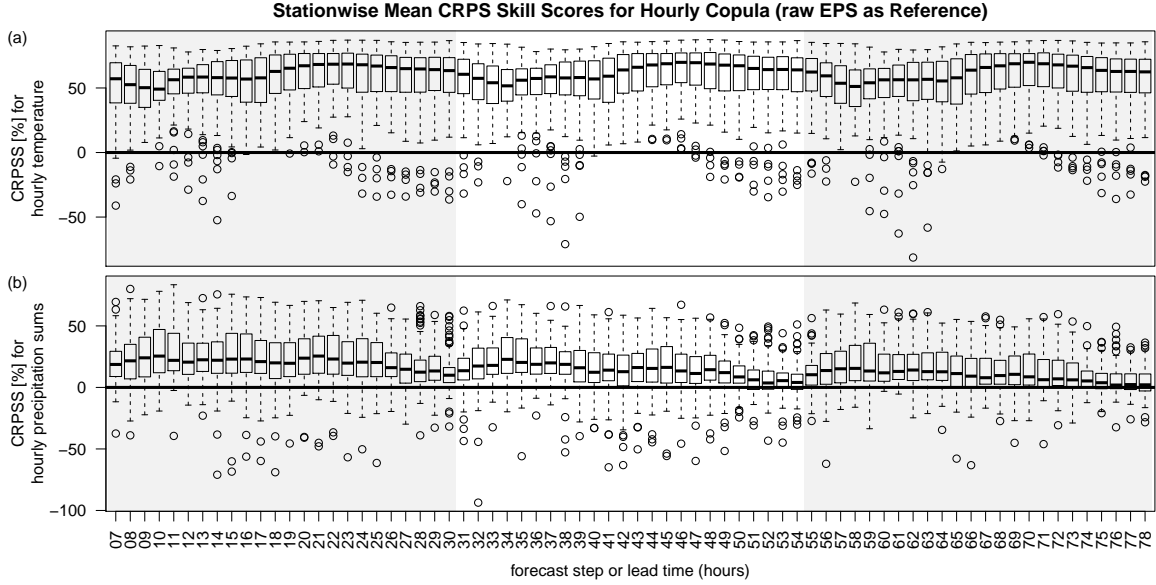


Figure 5: Continuous ranked probability skill scores (CRPS) for 2 m temperature (top) and hourly precipitation sums (bottom) based on station-wise mean empirical CRPS values (50+1 member). The raw EPS is used as reference. CRPSs are shown for each individual forecast step from +7 to +78 after model initialization. CRPSs above 0 (bold black line) show that the postprocessed hourly forecasts outperform the raw EPS.

5.4. Fresh Snow Amounts and Probability of Snowfall

This section shows the verification for the main target variable. Due to the limited availability of temporally high-resolution and reliable observations this can only be done for one site, the regional airport in Innsbruck (Figure 1). Figure 6 shows reliability diagrams (Bröcker and Smith 2007) for the probability of precipitation (rain/snow), rain, and snow using the classification from Section 2.5 and the aggregated METAR observations as described in Section 3.3. For all three precipitation classes the postprocessed forecasts are able to outperform the raw EPS (less off-diagonal, lower Brier score). Again, the wet BIAS of the raw EPS can be seen as the observed frequencies are typically lower than the forecasted probabilities for all classes. The probabilities for snowfall from the postprocessing seem to be slightly smaller than observed in reality, however, this result should not be over-interpreted as snowfall is relatively rare (7.5% of all cases).

Next, Figure 7 shows a forecast time series example for a random station and a day when the temperature is just around 1.2°C , the threshold used to decide whether the forecasted precipitation will fall as snow or rain. As no fresh snow measurements are available, a validation of the forecasted fresh snow amounts cannot be performed for this case.

What can be seen is that the ECC-Q temperature predictions (Figure 7a) show a much larger spread than the raw EPS. The postprocessed temperature uncertainty dominates the variation of the observed temperature over the whole forecast period (day 1–3). The observations,

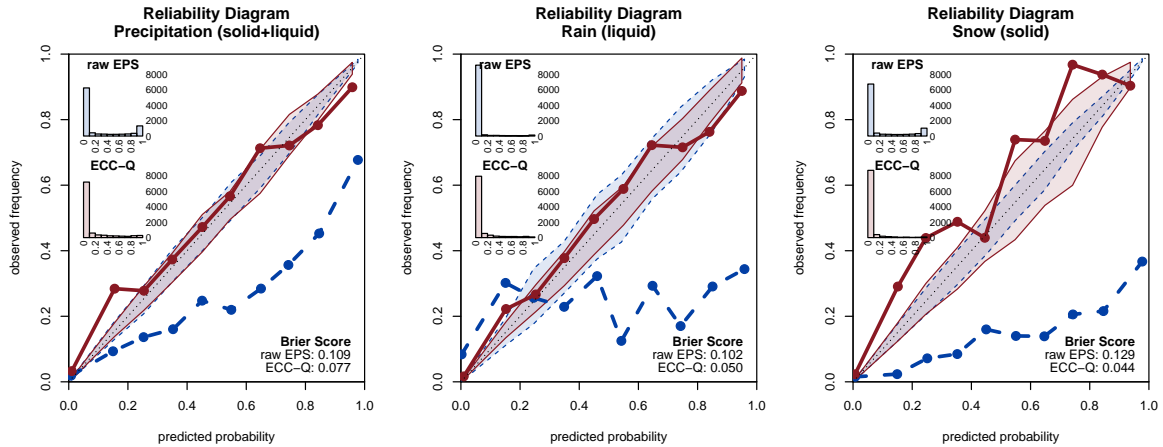


Figure 6: Reliability diagrams for hourly predictions of precipitation (snow+rain; left), snowfall (middle) and rain (right) at the Innsbruck Airport based on meteorological aerodrome reports (METAR) for the raw EPS (dashed) and the postprocessed forecasts (solid). The shaded area shows the 90% confidence interval. Histograms: counts of the number of observations in each bin in the reliability diagram. The analysis is based on ≈ 9700 observation/forecast pairs for each precipitation type. Mean Brier-scores printed in lower right corner.

however, are nicely falling into this interval which yields the overall well calibrated forecasts (see Figure 4). For precipitation (Figure 7c), the differences between the raw EPS and the postprocessed copula are less pronounced. Figure 7b shows the probability of snow+rain (precipitation), rain, and snow as defined by Equation 11. The expected amounts of snow+rain (precipitation) and snow from the postprocessed forecasts are plotted in Figure 7d. Rather than plotting each individual ECC member, the median and two confidence intervals are shown. For this specific date and location, the median shows 30.5 mm of precipitation (rain and/or snow liquid water equivalent) accumulated over the three consecutive days of which 8.4 mm are expected to fall as snow. When assuming the 1:10 rule (Section 2.5) and not taking the alteration of the aging snow into account, this corresponds to a median of 8.4 cm of fresh snow within 3 days.

5.5. Spatial Forecast Example

As a last result, Figure 8 and 9 show a spatial forecast example to demonstrate the ability to create high-resolution spatial predictions. These results show the +48 h forecast initialized March 8, 2017 on an approximately 500 m·500 m grid (corresponds to the +48 h forecast shown in Figure 7).

While Figure 8 shows the probability of precipitation (snow+rain), rain, and snow, Figure 9 shows the expected amount of precipitation for the period >47 h to +48 h. The color coding represents the dominant precipitation type based on $\pi_{snow,+48h}$ and $\pi_{rain,+48h}$ (cf. Equation 11). In addition, the snow line ($\pi_{snow,+48h} > \pi_{rain,+48h}$) is shown. The individual EPS and ECC-Q members used to derive probabilities and the expectation can be found in Appendix B.

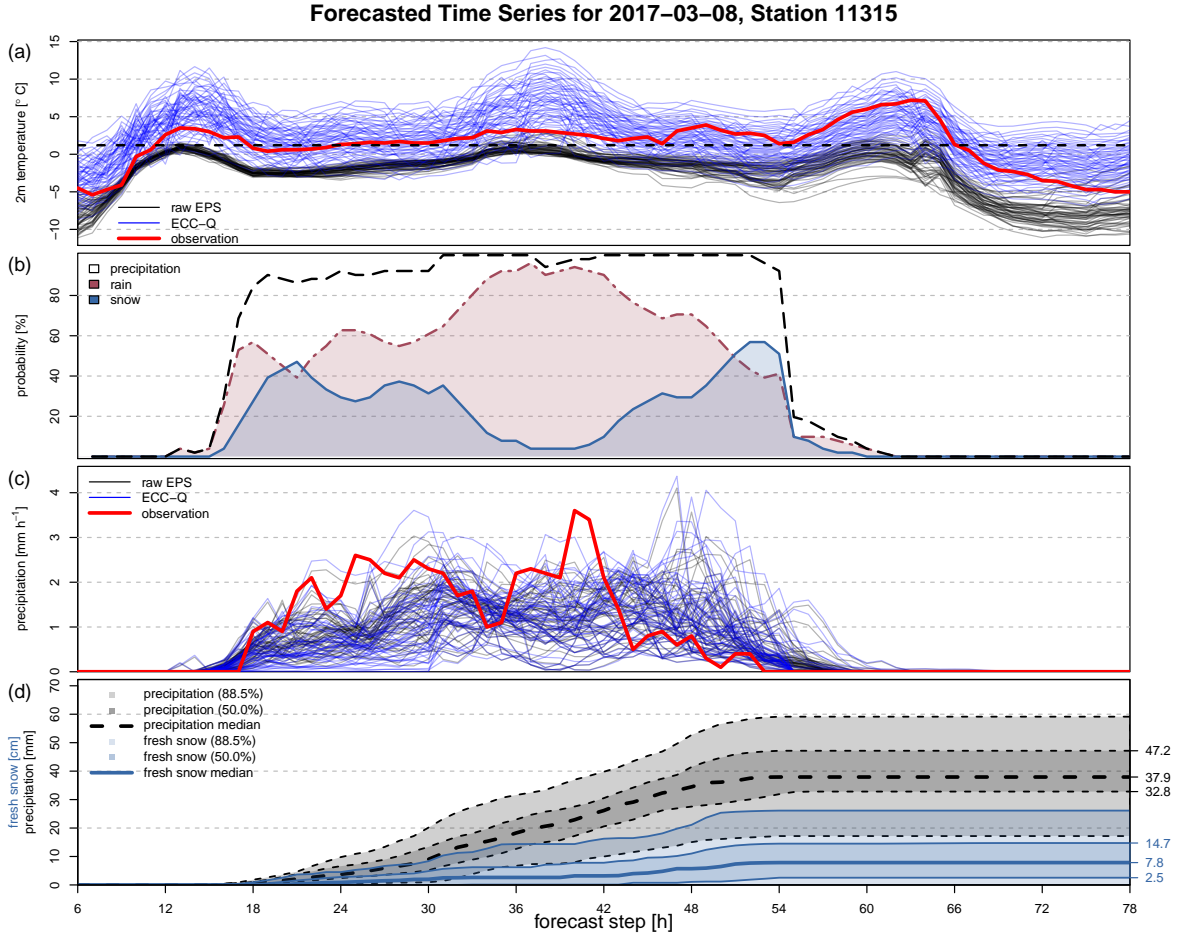


Figure 7: Example prediction for March 8, 2017 (station 11315, Holzgau) for the whole forecast horizon +6h up to +78h ahead. (a) Raw EPS forecast (black), postprocessed copula (blue), and observation (red; bold) for 2 m temperature. The black dashed line is the 1.2°C line used for precipitation type classification. (b) Probability of snow (blue solid), rain (red dotdash), and precipitation (snow \vee rain; black dashed). (c) Hourly precipitation forecasts and observations as in Panel (a). (d) Postprocessed forecasts for precipitation sum (dashed; mm), and fresh snow height (solid; cm) using the 1:10 rule (snow density of 100 kg m^{-3}). Predicted medians, predicted 50% intervals, and predicted 88.5% intervals are shown.

6. Discussion

This article presents a new hybrid approach to combine standardized anomaly output statistics (SAMOS) with ensemble copula coupling (ECC) and a novel re-weighting scheme for probabilistic snow forecasts. The results demonstrate that the new approach provides a framework for accurate high-resolution spatio-temporal probabilistic forecasts for 2 m temperature, precipitation, and snowfall over complex terrain.

The use of ECMWF hindcasts for model training and ECMWF EPS for prediction offers a computationally efficient way to get the required inputs for the SAMOS method (see Appendix A). Rather than estimating a complex spatio-temporal climatology for each covariate

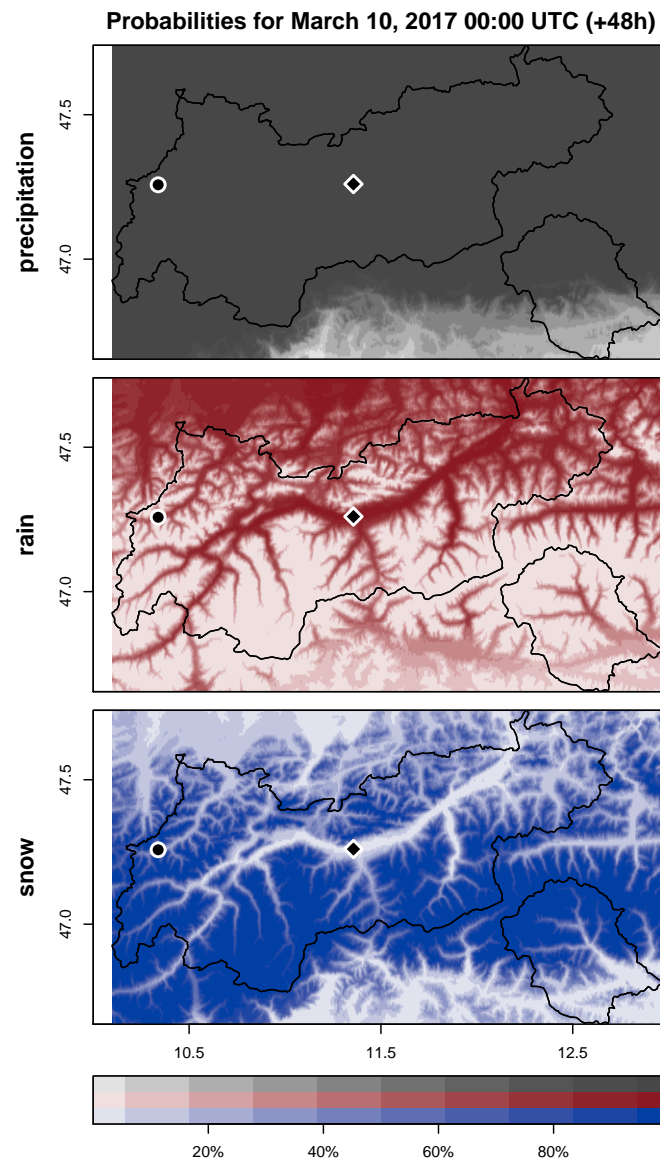


Figure 8: Top down: 1 h probability of precipitation (rain vs. snow), rain, and snow for March 10, 2017 0000 UTC (+48 h forecast initialized March 8, 2017). Overlays: the governmental area of Tyrol (solid line), Innsbruck Airport (diamond), and the location of the example station used in Figure 7 (circle).

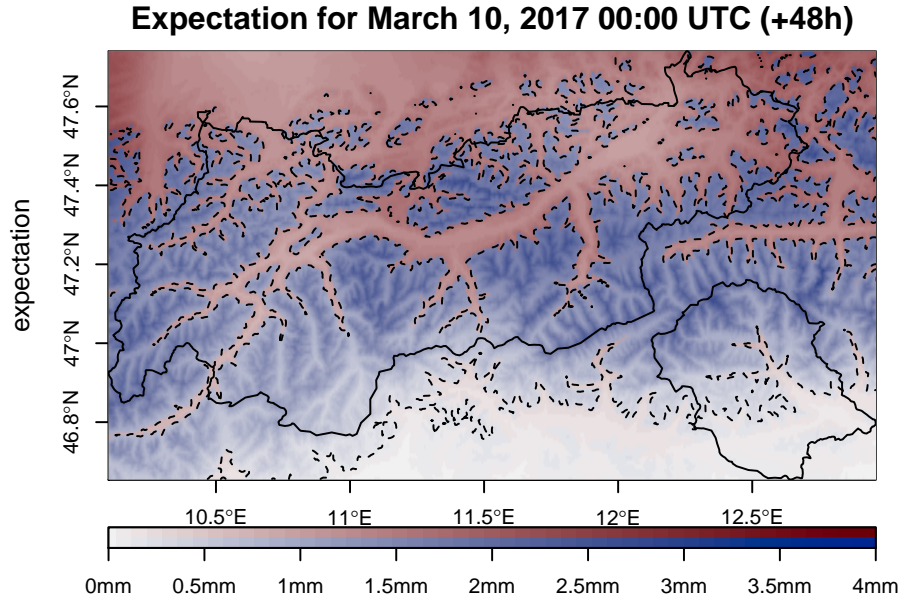


Figure 9: Expected 1 h amount of liquid water content for the March 10, 2017 0000 UTC (+48 h forecast initialized March 8, 2017). Areas with a higher chance to observe snow are shown in blue, those with a higher chance to observe rain in red. The dashed line shows the forecasted snow line with an equal chance to observe snow or rain (Equation 11). Overlay: governmental area of Tyrol (solid line).

(as in [Dabernig *et al.* 2017](#)), only empirical moments (mean and standard deviation) of an appropriate hindcast subset have to be derived. The latest 8 hindcast runs (4 weeks) centered around the date of interest are used to capture the seasonality. As this processing step is very cheap in terms of computational costs, one can easily derive hindcast climatologies for a range of possible covariates which allows for a simple and low-cost multi-linear extension of the SAMOS approach. Furthermore, due to the use of a rolling 4 week training period, the postprocessing procedure automatically adapts itself to possible changes in the underlying NWP model within a few weeks. However, the rank histograms (Figure 4) for both, the 2 m temperature and daily precipitation sums, show a pronounced U-shape. The same characteristics can be seen for all tested postprocessing models (not shown) whether or not standardized anomalies are used. The rank histograms for in-sample predictions based on the training data set itself (not shown) do not show this distinct pattern. This indicates that the forecasted uncertainty of the hindcasts and the uncertainty information from the current EPS seem to differ distinctly. The EPS seems to overall provide sharper forecasts so that applying the regression coefficients estimated on the hindcasts and applied to the EPS yields slightly underdispersive predictions.

Nevertheless, the method is still able to strongly improve calibration and reliability of the forecasts, especially for 2 m temperature, even though the sharpness is rather low. The mean 80 % prediction interval width for temperature is between 6.9–7.2° C for the SAMOS methods. On a rainy/snowy day this interval is quite likely wider than the overall diurnal temperature variation. The relatively wide predictive intervals are a result of the input data. Due to the current spatial resolution, the EPS is not able to represent the area of interest in all its details.

Consequently, a wide range of local features are not yet included. To mention one specific feature: the EPS shows a far-too-strong near-surface cooling over night, especially over snow. Errors of 15°C between the forecasted 2 m temperature and the corresponding observation are relatively frequent for alpine grid points. Furthermore, the forecasted EPS uncertainty does not seem to be very informative as almost no improvements can be seen when including it in the statistical models.

To improve the temperature forecasts, we include the temperature from the 850 hPa level as an additional covariate, which can be seen as a ‘free atmosphere’ prediction over the area of interest. Furthermore, the 850 hPa temperature is a prognostic quantity which should be less strongly affected by possible unrealistic surface processes (cooling/heating effects). As shown, this could slightly improve the forecast skill, even with only simple linear effects. A more flexible SAMOS model specification might bring further improvements, e.g., by including other covariates, including interactions between the different covariates, or including more flexible effects such as elevation-dependent effects (or dependent on difference between true and model topography). This would allow the SAMOS model to rely more strongly on different covariates for sites at different altitude levels and could be investigated in future research.

One of the biggest advantages of the proposed hybrid approach is that forecasts can be produced on the same temporal scale as the current EPS even if the underlying data set used for model training (hindcasts and observations) are available on coarser temporal scales or even different time scales for different variables. This allows to combine the best information from (location-)independent sources to get the most reliable probabilistic predictions possible. For the present study, two observation networks have been combined, one providing long-term daily precipitation records, while the second one provides temporally highly resolved temperature measurements.

Overall the 2 m temperature and precipitation forecasts serve as a good proxy for probabilistic snowfall forecasts which is the main target variable of this study. The results show very promising results in terms of calibration and reliability of both, the expected amount of precipitation and fresh snow, but also for the probability to observe snowfall on an hourly temporal resolution.

Computational details

The main parts of this study are based on the R package **bamlss** (Umlauf *et al.* 2017) to compute the spatio-temporal observation climatologies and the R package **crch** (Messner *et al.* 2016) to estimate the (censored) non-homogeneous regression models. The continuous ranked probability scores are based on the R package **scoringRules** (Jordan *et al.* 2017).

Observations from the hydrographical service (BMLFUW 2016) can be downloaded from the website of the Bundesministerium für Land und Forstwirtschaft und Wasserwirtschaft (<http://ehyd.gv.at>).

Acknowledgments

The work of this article was partially funded by the Austrian Science Fund (FWF), grant TRP 290, and the Austrian Research Promotion Agency (FFG), grant No. 858537. The data sets are provided by the Zentralanstalt für Meteorologie und Geodynamik Vienna (ZAMG; <https://zamg.ac.at>), and the Federal Ministry of Agriculture, Forestry, Environment and Water Management (BMLFUW), Abteilung IV/4 – Wasserhaushalt (<http://ehyd.gv.at>)

References

- Amt der Tiroler Landesregierung (2014). “Statistisches Jahrbuch Bundesland Tirol.” Accessed: July 15 2016, URL https://www.tirol.gv.at/fileadmin/themen/statistik-budget/statistik/downloads/Statistisches_Handbuch_2014.pdf.
- Bellaire S, Jamieson JB, Fierz C (2011). “Forcing the Snow-Cover Model SNOWPACK with Forecasted Weather Data.” *The Cryosphere*, **5**(4), 1115–1125. doi:10.5194/tc-5-1115-2011.
- BMLFUW (2016). “Bundesministerium für Land und Forstwirtschaft, Umwelt und Wasserwirtschaft (BMLFUW), Abteilung IV/4 - Wasserhaushalt.” Available at <http://ehyd.gv.at>. Accessed: February 29 2016.
- Bouallègue ZB, Theis SE (2014). “Spatial Techniques Applied to Precipitation Ensemble Forecasts: from Verification Results to Probabilistic Products.” *Meteorological Applications*, **21**(4), 922–929. doi:10.1002/met.1435.
- Bröcker J, Smith LA (2007). “Increasing the Reliability of Reliability Diagrams.” *Weather and Forecasting*, **22**(3), 651–661. doi:10.1175/WAF993.1.
- Dabernig M, Mayr GJ, Messner JW, Zeileis A (2017). “Spatial Ensemble Post-Processing with Standardized Anomalies.” *Quarterly Journal of the Royal Meteorological Society*, **143**(703), 909–916. doi:10.1002/qj.2975.
- Fraley C, Raftery AE, Gneiting T (2010). “Calibrating Multimodel Forecast Ensembles with Exchangeable and Missing Members Using Bayesian Model Averaging.” *Monthly Weather Review*, **138**(1), 190–202. doi:10.1175/2009MWR3046.1.
- Gebetsberger M, Messner JW, Mayr GJ, Zeileis A (2017). “Fine-Tuning Non-Homogeneous Regression for Probabilistic Precipitation Forecasts: Unanimous Predictions, Heavy Tails, and Link Functions.” *Monthly Weather Review*. doi:10.1175/MWR-D-16-0388.1.
- Gneiting T, Raftery AE (2007). “Strictly Proper Scoring Rules, Prediction, and Estimation.” *Journal of the American Statistical Association*, **102**(477), 359–378. doi:10.1198/016214506000001437.
- Gneiting T, Raftery AE, Westveld III AH, Goldman T (2005). “Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation.” *Monthly Weather Review*, **133**(5), 1098–1118. doi:10.1175/MWR2904.1.
- Hagedorn R, Buizza R, Hamill TM, Leutbecher M, Palmer TN (2012). “Comparing TIGGE Multimodel Forecasts with Reforecast-Calibrated ECMWF Ensemble Forecasts.” *Quarterly Journal of the Royal Meteorological Society*, **138**(668), 1814–1827. doi:10.1002/qj.1895.
- Hagedorn R, Hamill TM, Whitaker JS (2008). “Probabilistic Forecast Calibration Using ECMWF and GFS Ensemble Reforecasts. Part I: Two-Meter Temperatures.” *Monthly Weather Review*, **136**(7), 2608–2619. doi:10.1175/2007MWR2410.1.

- Hamill TM (2001). “Interpretation of Rank Histograms for Verifying Ensemble Forecasts.” *Monthly Weather Review*, **129**(3), 550–560. doi:[10.1175/1520-0493\(2001\)129<0550:IORHFV>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2).
- Hamill TM, Scheuerer M, Bates GT (2015). “Analog Probabilistic Precipitation Forecasts Using GEFS Reforecasts and Climatology-Calibrated Precipitation Analyses.” *Monthly Weather Review*, **143**(8), 3300–3309. doi:[10.1175/MWR-D-15-0004.1](https://doi.org/10.1175/MWR-D-15-0004.1).
- Hamill TM, Whitaker JS, Mullen SL (2006). “Reforecasts: An Important Dataset for Improving Weather Predictions.” *Bulletin of the American Meteorological Society*, **87**(1), 33–46. doi:[10.1175/BAMS-87-1-33](https://doi.org/10.1175/BAMS-87-1-33).
- Jordan A, Krueger F, Lerch S (2017). *scoringRules: Scoring Rules for Parametric and Simulated Distribution Forecasts*. R package version 0.9.4, URL <https://CRAN.R-project.org/package=scoringRules>.
- Judson A, Doesken N (2000). “Density of Freshly Fallen Snow in the Central Rocky Mountains.” *Bulletin of the American Meteorological Society*, **81**(7), 1577–1587. doi:[10.1175/1520-0477\(2000\)081<1577:DOFFSI>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<1577:DOFFSI>2.3.CO;2).
- Knox T, Gerhold L, Ulbrich U (2015). “Perception and Use of Uncertainty in Severe Weather Warnings by Emergency Services in Germany.” *Atmospheric Research*, **158–159**, 292–301. doi:[10.1016/j.atmosres.2014.02.024](https://doi.org/10.1016/j.atmosres.2014.02.024).
- Lawinenwarndienst Tirol (2009–2017). “Winterberichte 2009/2010 bis 2015/2016.” Available at <https://lawine.tirol.gv.at/archiv/winterberichte/>. Accessed: November 14 2017.
- Lerch S, Thorarinsdottir T (2013). “Comparison of Non-Homogeneous Regression Models for Probabilistic Wind Speed Forecasting.” *Tellus A*, **65**. doi:[10.3402/tellusa.v65i0.21206](https://doi.org/10.3402/tellusa.v65i0.21206).
- Meister R (1985). “Density of New Snow and its Dependence on Air Temperature and Wind.” *Zürcher Geographische Schriften*, (23), 73–79.
- Messner JW, Mayr GJ, Wilks DS, Zeileis A (2014a). “Extending Extended Logistic Regression: Extended versus Separate versus Ordered versus Censored.” *Monthly Weather Review*, **142**(8), 3003–3014. doi:[10.1175/MWR-D-13-00355.1](https://doi.org/10.1175/MWR-D-13-00355.1).
- Messner JW, Mayr GJ, Zeileis A (2016). “Heteroscedastic Censored and Truncated Regression with crch.” *The R Journal*, **8**(1), 173–181. URL <https://journal.r-project.org/archive/2016-1/messner-mayr-zeileis.pdf>.
- Messner JW, Mayr GJ, Zeileis A, Wilks DS (2014b). “Heteroscedastic Extended Logistic Regression for Postprocessing of Ensemble Guidance.” *Monthly Weather Review*, **142**(1), 448–456. doi:[10.1175/MWR-D-13-00271.1](https://doi.org/10.1175/MWR-D-13-00271.1).
- Mullen SL, Buizza R (2001). “Quantitative Precipitation Forecasts over the United States by the ECMWF Ensemble Prediction System.” *Monthly Weather Review*, **129**(4), 638–663. doi:[10.1175/1520-0493\(2001\)129<0638:QPFOTU>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0638:QPFOTU>2.0.CO;2).

- Neal RA, Boyle P, Grahame N, Mylne K, Sharpe M (2014). “Ensemble Based First Guess Support Towards a Risk-based Severe Weather Warning Service.” *Meteorological Applications*, **21**(3), 563–577. doi:10.1002/met.1377.
- Palmer TN (2002). “The Economic Value of Ensemble Forecasts as a Tool for Risk Assessment: From Days to Decades.” *Quarterly Journal of the Royal Meteorological Society*, **128**(581), 747–774. doi:10.1256/0035900021643593.
- Raftery AE (2016). “Use and Communication of Probabilistic Forecasts.” *Statistical Analysis and Data Mining: The ASA Data Science Journal*, **9**(6), 397–410. doi:10.1002/sam.11302.
- Rasmussen R, Baker B, Kochendorfer J, Meyers T, Landolt S, Fischer AP, Black J, Thériault JM, Kucera P, Gochis D, Smith C, Nitu R, Hall M, Ikeda K, Gutmann E (2012). “How Well Are We Measuring Snow: The NOAA/FAA/NCAR Winter Precipitation Test Bed.” *Bulletin of the American Meteorological Society*, **93**(6), 811–829. doi:10.1175/BAMS-D-11-00052.1.
- Roebber PJ, Bruening SL, Schultz DM, Cortinas Jr JV (2003). “Improving Snowfall Forecasting by Diagnosing Snow Density.” *Weather and Forecasting*, **18**(2), 264–287. doi:10.1175/1520-0434(2003)018<0264:ISFBDS>2.0.CO;2.
- Rohregger JB (2008). *Methoden zur Bestimmung der Schneefallgrenze*. Master’s thesis, Universität Wien.
- Roulston MS, Smith LA (2003). “Combining Dynamical and Statistical Ensembles.” *Tellus A*, **55**(1), 16–30. doi:10.1034/j.1600-0870.2003.201378.x.
- Schefzik R, Thorarinsdottir TL, Gneiting T (2013). “Uncertainty Quantification in Complex Simulation Models Using Ensemble Copula Coupling.” *Statistical Science*, **28**(4), 616–640. doi:10.1214/13-STS443.
- Scheuerer M (2014). “Probabilistic Quantitative Precipitation Forecasting Using Ensemble Model Output Statistics.” *Quarterly Journal of the Royal Meteorological Society*, **140**(680), 1086–1096. doi:10.1002/qj.2183.
- 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.
- Scheuerer M, Hamill TM (2015). “Statistical Postprocessing of Ensemble Precipitation Forecasts by Fitting Censored, Shifted Gamma Distributions.” *Monthly Weather Review*, **143**(11), 4578–4596. doi:10.1175/MWR-D-15-0061.1.
- Sloughter JML, Raftery AE, Gneiting T, Fraley C (2007). “Probabilistic Quantitative Precipitation Forecasting Using Bayesian Model Averaging.” *Monthly Weather Review*, **135**(9), 3209–3220. doi:10.1175/MWR3441.1.
- Stauffer R, Mayr GJ, Messner JW, Umlauf N, Zeileis A (2017a). “Spatio-Temporal Precipitation Climatology over Complex Terrain Using a Censored Additive Regression Model.” *International Journal of Climatology*, **37**(7), 3264–3275. doi:10.1002/joc.4913.

- Stauffer R, Umlauf N, Messner JW, Mayr GJ, Zeileis A (2017b). “Ensemble Postprocessing of Daily Precipitation Sums over Complex Terrain Using Censored High-Resolution Standardized Anomalies.” *Monthly Weather Review*, **145**(3), 955–969. doi:10.1175/MWR-D-16-0260.1.
- Thorarinsdottir TL, Gneiting T (2010). “Probabilistic Forecasts of Wind Speed: Ensemble Model Output Statistics by Using Heteroscedastic Censored Regression.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **173**(2), 371–388. doi:10.1111/j.1467-985X.2009.00616.x.
- Umlauf N, Klein N, Zeileis A (2017). “BAMLSS: Bayesian Additive Models for Location, Scale and Shape (and Beyond).” *Journal of Computational and Graphical Statistics*, **NA**(NA), NA. doi:10.1080/10618600.2017.1407325.
- Wilks DS (2009). “Extending Logistic Regression to Provide Full-Probability-Distribution MOS Forecasts.” *Meteorological Applications*, **16**(3), 361–368. doi:10.1002/met.134.
- Zhu Y, Toth Z, Wobus R, Richardson D, Myln eK (2002). “The Economic Value Of Ensemble-Based Weather Forecasts.” *Bulletin of the American Meteorological Society*, **83**(1), 73–83. doi:10.1175/1520-0477(2002)083<0073:TEVOEB>2.3.CO;2.

A. Standardized Anomaly Model Output Statistics (SAMOS)

For spatio-temporal ensemble postprocessing we followed the approach of Dabernig *et al.* (2017) and Stauffer *et al.* (2017b), which we summarize in the following. In contrast to other statistical postprocessing methods SAMOS uses standardized anomalies for both, the response and the covariates. This allows to remove location-specific and time-specific characteristics from the data and to estimate one single regression model for all stations and forecast lead times at once. For this study we closely follow the original articles (Dabernig *et al.* 2017; Stauffer *et al.* 2017b) but slightly modify the specification, especially for the temperature SAMOS to adapt to the different study area.

Observation climatologies: Two separate spatio-temporal models have been estimated for 2 m air temperature observations and daily precipitation sums. Both models have effects to capture seasonal, altitudinal, and spatial climatological features represented by (multi-dimensional) non-linear functions. The 2 m temperature observations are available at an hourly temporal resolution. Therefore, additional non-linear cyclic effects have to be included to capture the diurnal effects in the climatological estimates.

The spatio-temporal model for the 2 m temperature uses the geographical location (longitude *lon*, latitude *lat*, and altitude *alt*), the ‘hour of the day’ (*hour*), and the ‘day of the year’ (*doy*) as covariates and is specified as follows:

$$\begin{aligned}
 \text{temperature} &\sim \mathcal{N}(\tilde{\mu}_y, \tilde{\sigma}_y), \\
 \tilde{\mu}_y &= f_1(\text{hour}, \text{doy}, \text{alt}) + \\
 &f_2(\text{hour}, \text{doy}) + f_3(\text{doy}, \text{lon}, \text{lat}) + \\
 &f_4(\text{hour}) + f_5(\text{doy}) + f_6(\text{doy}, \text{alt}) + f_7(\text{alt}), \\
 \log(\tilde{\sigma}_y) &= g_1(\text{hour}, \text{doy}, \text{alt}) + g_2(\text{hour}, \text{doy}) + \\
 &g_3(\text{doy}, \text{lon}, \text{lat}) + \\
 &g_4(\text{hour}) + g_5(\text{doy}) + g_6(\text{doy}, \text{alt}) + g_7(\text{alt}),
 \end{aligned} \tag{13}$$

where f_\bullet and g_\bullet are up to three-dimensional smooth spline effects. Cyclic P-splines are used for all effects depending on the ‘day of the year’ or the ‘hour of the day’; all other effects use penalized thin plate splines with a varying number of possible degrees of freedom. Following the same concept, the spatio-temporal model for daily precipitation sums is defined as:

$$\begin{aligned}
 \text{precipitation}^{1/p} &\sim \mathcal{L}_0(\tilde{\mu}_y, \tilde{\sigma}_y), \\
 \tilde{\mu}_y &= f_1(\text{alt}) + f_2(\text{doy}) + f_3(\text{lon}, \text{lat}) + \\
 &f_4(\text{doy}, \text{lon}, \text{lat}), \\
 \log(\tilde{\sigma}_y) &= g_1(\text{alt}) + g_2(\text{doy}) + g_3(\text{lon}, \text{lat}) + \\
 &g_4(\text{doy}, \text{lon}, \text{lat}).
 \end{aligned} \tag{14}$$

As for Equation 13, cyclic P-splines are used for effects which depend on the ‘day of the year’ while all others use penalized thin plate splines. The major difference to the temperature

climatology (Equation 13) is that a left-censored logistic response distribution \mathcal{L}_0 is used on power-transformed observations of precipitation^{1/p} ($p = 1.35$; cf. Stauffer *et al.* 2017b). The complexity of the linear predictors in Equation 14 is lower than in Equation 13 as no effects for diurnal variation have to be considered.

Model climatology: Similar spatio-temporal climatologies as for the observations could be estimated for all quantities from the EPS which are used as covariates in the SAMOS models. This would have to be done for each quantity separately using a reasonably large data set of historical EPS forecasts. However, instead of this approach we extract the model climatologies directly from ECMWF hindcasts. These hindcasts are produced operationally twice a week and consist of 10 + 1 members using the same model version and model specification as the current EPS. For each hindcast run the forecasts for the same date over the most recent 20 years are computed. These hindcasts are designed to represent the climatology of the current EPS model and are used to calibrate EPS forecasts and for postprocessing applications (e.g., Hagedorn *et al.* 2012, 2008). For our SAMOS approach we can thus simply derive the empirical mean and empirical standard deviation over a set of hindcasts to get the climatological estimates $\tilde{\mu}_x$ and $\tilde{\sigma}_x$ required to compute the standardized anomalies for variable \mathbf{x} (Equation 9). Climatologies for lead times when no hindcast output is available (between the regular 6 h interval) are created using simple grid-point-wise linear interpolation.

Hindcasts are produced on Mondays and Thursdays (available Tuesday/Friday) computed two weeks in advance. Taking hindcasts for ± 2 weeks around the date of interest yields 8 independent hindcast runs with 11 members and 20 years of (re-)forecasts each, which yields $8 \cdot 11 \cdot 20 = 1760$ forecasts. With this large number of independent predictions these climatological estimates are fairly robust. They follow seasonal changes due to the centered 4 week moving window. It should be noted that separate climatologies for each forecast step are required to capture diurnal cycles (for temperature) and to account for changes in the model climate with increasing forecast horizon such as drifting means or increasing ensemble standard deviation. For this study 13 separate climatologies for the temperature models ($[+6h, +12h, \dots, +72h, +78h]$) and three climatologies for the precipitation forecasts ($[+30h, +54h, +78h]$) are needed.

Estimation of the SAMOS models (see Table 1): Equations 1-3 & 5-7 show the basic heteroscedastic models used for *SAMOS_hom* and *SAMOS_het*. The only modification is to replace the response \mathbf{y} and the covariate \mathbf{x} with the corresponding standardized anomaly \mathbf{y}^* and \mathbf{x}^* (Equation 9). For the *xSAMOS_het* model the linear predictors in Equations 1-3 are extended by simply adding additional covariates, which results in a multilinear SAMOS.

Once the regression coefficients of the SAMOS model have been estimated, future ensemble forecasts can be corrected by first computing standardized anomalies using the same model climatology as for model training and applying the SAMOS correction. As the outcomes μ_i^* and σ_i^* will be then on the standardized anomaly scale they have to be rescaled with respect to the observation climatology to obtain physical values. The final predictive distribution is then

$$y_i \sim \mathcal{D}(\mu_i^* \cdot \tilde{\sigma}_{y,i} + \tilde{\mu}_{y,i}, \sigma_i^* \cdot \tilde{\sigma}_{y,i}), \quad (15)$$

where \mathcal{D} represents the normal distribution \mathcal{N} in case of the 2 m temperature postprocessing SAMOS and \mathcal{L}_0 in case of the power-transformed daily precipitation sums.

Algorithm 1 presents pseudo-code for all steps required for the SAMOS method. The same is shown in Figure 10 as a graphical representation of this procedure, visualizing the required

data sets, the required steps, and their dependencies.

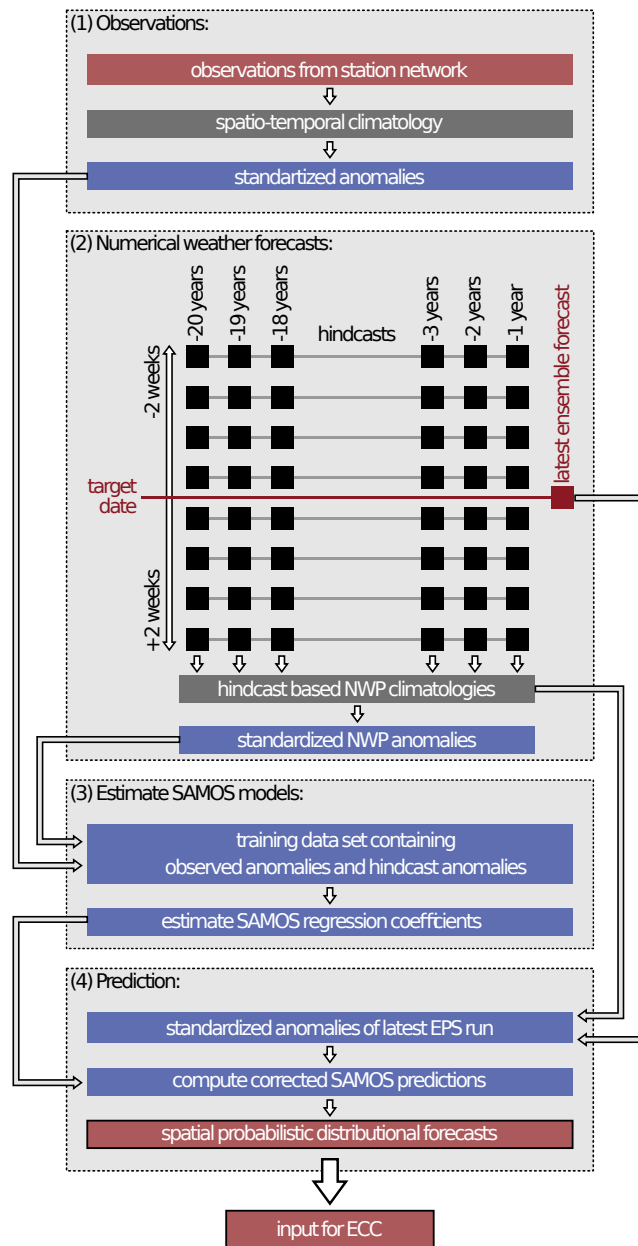


Figure 10: Schematic concept of the SAMOS postprocessing based on ECMWF hindcasts (black), ECMWF EPS forecasts (red), and observations (orange). Background climatologies (gray) are used to convert the data from the physical scale into standardized anomalies (blue) used to estimate the regression coefficients of the SAMOS postprocessing method. The SAMOS correction can be applied to the standardized anomalies of a new EPS forecast to obtain spatial or spatio-temporal probabilistic forecasts (full distribution). These results are used as input for the ECC approach.

B. Individual Copula Members

Figures 11 and 12 show the individual EPS members (Figure 11) and the corresponding re-weighted ensemble copula coupling members (Figure 12) for the +48 h forecast for March 10, 2017 1000 UTC as used to derive the probabilities and expectation plotted in Figure 8 and Figure 9. For easier comparison the NWP forecasts are bilinearly interpolated to $\approx 500 \cdot 500 \text{ m}^2$ to match the resolution of the postprocessed predictions.

Affiliation:

Reto Stauffer, Achim Zeileis
Department of Statistics
Faculty of Economics and Statistics
Universität Innsbruck
Universitätsstr. 15
6020 Innsbruck, Austria
E-mail: Reto.Stauffer@uibk.ac.at, Achim.Zeileis@uibk.ac.at
URL: <https://retostauffer.org>, <https://eeecon.uibk.ac.at/~zeileis/>

Georg J. Mayr
Department of Atmospheric and Cryospheric Sciences
Faculty of Geo- and Atmospheric Sciences
Universität Innsbruck
Innrain 52f
6020 Innsbruck, Austria
E-mail: Georg.Mayr@uibk.ac.at

Jakob W. Messner
Department of Electrical Engineering
Technical University of Denmark
Elektrovej
Building 325
2800 Kgs. Lyngby
E-mail: jwmm@elektro.dtu.dk

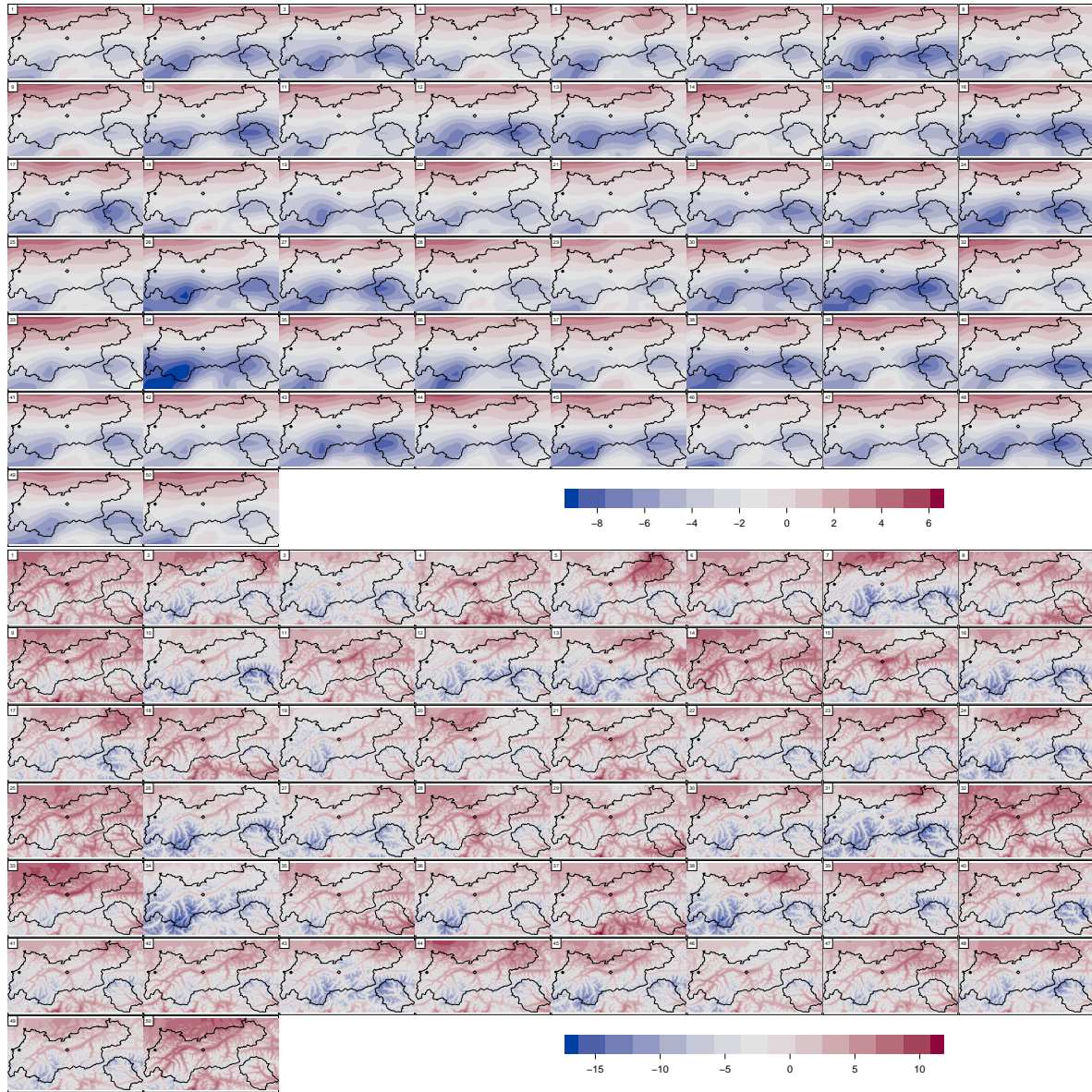


Figure 11: Stamps for +48 h forecasts initialized March 8, 2017 0000 UTC (valid for March 10, 2017 0000 UTC). Individual EPS members for 2 m temperature (top) and the corresponding copula members (bottom). Please note that the color scale for all members of one type (EPS/copula) is identical, but the scales between the raw EPS and the results from the postprocessing differ.

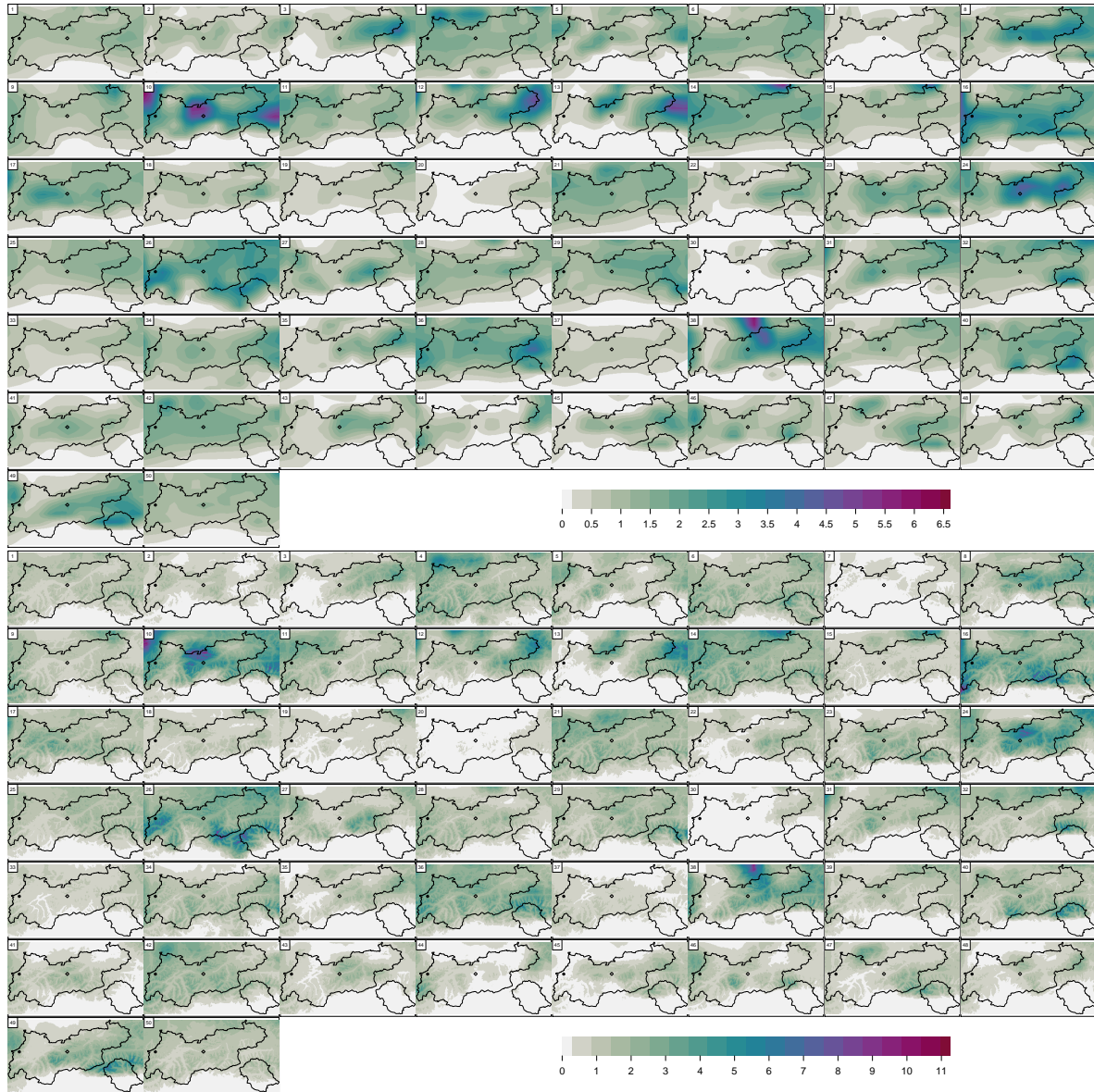


Figure 12: Stamps for +48 h forecasts initialized March 8, 2017 0000 UTC (valid for March 10, 2017 0000 UTC). Individual EPS members for 1 h precipitation sums (top) and the corresponding copula members (bottom). Please note that the color scale for all members of one type (EPS/copula) is identical, but the scales between the raw EPS and the results from the postprocessing differ.

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

<https://www.uibk.ac.at/eeecon/wopec/>

- 2018-05 **Reto Stauffer, Georg J. Mayr, Jakob W. Messner, Achim Zeileis:** Hourly probabilistic snow forecasts over complex terrain: A hybrid ensemble postprocessing approach
- 2018-04 **Utz Weitzel, Christoph Huber, Florian Lindner, Jürgen Huber, Julia Rose, Michael Kirchler:** Bubbles and financial professionals
- 2018-03 **Carolin Strobl, Julia Kopf, Raphael Hartmann, Achim Zeileis:** Anchor point selection: An approach for anchoring without anchor items
- 2018-02 **Michael Greinecker, Christopher Kah:** Pairwise stable matching in large economies
- 2018-01 **Max Breitenlechner, Johann Scharler:** How does monetary policy influence bank lending? Evidence from the market for banks' wholesale funding
- 2017-27 **Kenneth Harttgen, Stefan Lang, Johannes Seiler:** Selective mortality and undernutrition in low- and middle-income countries
- 2017-26 **Jun Honda, Roman Inderst:** Nonlinear incentives and advisor bias
- 2017-25 **Thorsten Simon, Peter Fabsic, Georg J. Mayr, Nikolaus Umlauf, Achim Zeileis:** Probabilistic forecasting of thunderstorms in the Eastern Alps
- 2017-24 **Florian Lindner:** Choking under pressure of top performers: Evidence from biathlon competitions
- 2017-23 **Manuel Gebetsberger, Jakob W. Messner, Georg J. Mayr, Achim Zeileis:** Estimation methods for non-homogeneous regression models: Minimum continuous ranked probability score vs. maximum likelihood
- 2017-22 **Sebastian J. Dietz, Philipp Kneringer, Georg J. Mayr, Achim Zeileis:** Forecasting low-visibility procedure states with tree-based statistical methods
- 2017-21 **Philipp Kneringer, Sebastian J. Dietz, Georg J. Mayr, Achim Zeileis:** Probabilistic nowcasting of low-visibility procedure states at Vienna International Airport during cold season
- 2017-20 **Loukas Balafoutas, Brent J. Davis, Matthias Sutter:** How uncertainty and ambiguity in tournaments affect gender differences in competitive behavior
- 2017-19 **Martin Geiger, Richard Hule:** The role of correlation in two-asset games: Some experimental evidence

- 2017-18 **Rudolf Kerschbamer, Daniel Neururer, Alexander Gruber:** Do the altruists lie less?
- 2017-17 **Meike Köhler, Nikolaus Umlauf, Sonja Greven:** Nonlinear association structures in flexible Bayesian additive joint models
- 2017-16 **Rudolf Kerschbamer, Daniel Muller:** Social preferences and political attitudes: An online experiment on a large heterogeneous sample
- 2017-15 **Kenneth Harttgen, Stefan Lang, Judith Santer, Johannes Seiler:** Modeling under-5 mortality through multilevel structured additive regression with varying coefficients for Asia and Sub-Saharan Africa
- 2017-14 **Christoph Eder, Martin Halla:** Economic origins of cultural norms: The case of animal husbandry and bastardy
- 2017-13 **Thomas Kneib, Nikolaus Umlauf:** A primer on bayesian distributional regression
- 2017-12 **Susanne Berger, Nathaniel Graham, Achim Zeileis:** Various versatile variances: An object-oriented implementation of clustered covariances in R
- 2017-11 **Natalia Danzer, Martin Halla, Nicole Schneeweis, Martina Zweimüller:** Parental leave, (in)formal childcare and long-term child outcomes
- 2017-10 **Daniel Muller, Sander Renes:** Fairness views and political preferences - Evidence from a large online experiment
- 2017-09 **Andreas Exenberger:** The logic of inequality extraction: An application to Gini and top incomes data
- 2017-08 **Sibylle Puntcher, Duc Tran Huy, Janette Walde, Ulrike Tappeiner, Gottfried Tappeiner:** The acceptance of a protected area and the benefits of sustainable tourism: In search of the weak link in their relationship
- 2017-07 **Helena Fornwagner:** Incentives to lose revisited: The NHL and its tournament incentives
- 2017-06 **Loukas Balafoutas, Simon Czermak, Marc Eulerich, Helena Fornwagner:** Incentives for dishonesty: An experimental study with internal auditors
- 2017-05 **Nikolaus Umlauf, Nadja Klein, Achim Zeileis:** BAMLSS: Bayesian additive models for location, scale and shape (and beyond)
- 2017-04 **Martin Halla, Susanne Pech, Martina Zweimüller:** The effect of statutory sick-pay on workers' labor supply and subsequent health
- 2017-03 **Franz Buscha, Daniel Müller, Lionel Page:** Can a common currency foster a shared social identity across different nations? The case of the Euro.
- 2017-02 **Daniel Müller:** The anatomy of distributional preferences with group identity

2017-01 **Wolfgang Frimmel, Martin Halla, Jörg Paetzold:** The intergenerational causal effect of tax evasion: Evidence from the commuter tax allowance in Austria