A probabilistic forecast approach for daily precipitation totals

Petra Friederichs *

Andreas Hense

Meteorological Institute, Bonn, Germany

Submitted to Weather and Forecasting

Revised version November 27, 2007

*Corresponding author address: Petra Friederichs, Meteorological Institute, Auf dem Hügel 20, 53121 Bonn, Germany and Interdisciplinary Centre of Complex Systems (IZKS), Römerstr. 164, 53117 Bonn.

E-mail: pfried@uni-bonn.de

Abstract

Commonly, post-processing techniques are employed to calibrate a model forecast. Here, we present a probabilistic post-processor that provides calibrated probability and quantile forecasts of precipitation on the local scale. The forecasts are based on large-scale circulation patterns of the 12h forecast from the NCEP High Resolution Global Forecast System. The censored quantile regression is used to estimate selected quantiles of the precipitation amount and the probability of the occurrence of precipitation. The approach accounts for the mixed discrete-continuous character of daily precipitation totals. The forecasts are verified using a new verification score for quantile forecasts, namely the censored quantile verification (CQV) score.

The forecast approach is as follows. First, a canonical correlation is employed to correct systematic deviations in the GFS large-scale patterns compared to the NCEP or ERA40 reanalysis. Secondly, the statistical quantile model between the large-scale circulation and the local precipitation quantile is derived using NCEP and ERA40 reanalysis data. Then, the statistical quantile model is applied to 12h forecasts provided by the GFS forecast system. The probabilistic forecasts are reliable and the relative gain in performance of the quantile as well as the probability forecasts compared to the climatological forecasts range between 20% and 50%. The importance of the various parts of the post-processing are assessed, and the performance is compared to forecasts based on the direct precipitation output from the ECMWF forecast system.

1. Introduction

Quantitative forecasts of precipitation including its extremes are of high socio-economic interest. Although present-day global weather forecast models provide reliable forecasts of the atmospheric large-scale circulation, they cannot provide realistic descriptions of local weather variability. This is due to the horizontally restricted resolution (30-100km at best for global models), but also due to missing cloud dynamical and microphysical processes. For this reason, a description and forecast of local weather phenomena and particularly extreme events can only be achieved through a combination of dynamical and statistical analysis methods, where a stable and significant statistical model based on a-priori physical reasoning establishes a-posterior a calibrated model between the local condition and the large-scale circulation.

The perfect prog method (PP) (Klein 1971) and model output statistics (MOS) (Glahn and Lowry 1972) have been successfully applied in numerical weather prediction to re-calibrate the direct model output to local conditions. In contrast to MOS, PP ignores the model forecast error, but has the advantage of available large data sets. Kalnay (2003) and Marzban et al. (2005) discuss the pros and cons of both approaches. They propose a combination of PP and MOS, which uses reanalysis for the development of the regression equations and denote this approach RAN as an acronym for reanalysis. MOS is generally based on multiple linear regression and therefore applies to the post-processing and forecasting of expectation values. Other approaches use statistical correction based on the analog method (Zorita and von Storch 1998; Hamill and Whitaker 2006). Vislocky and Young (1989) used different PP models based on an analog model and logistic regression as predictors in an MOS approach. Bremnes (2004) was the first who applied quantile regression for precipitation forecasts in the context of numerical weather prediction.

The stochastic character of weather requires a probabilistic treatment. Furthermore, probabilistic forecasts provide a measure of uncertainty that might be important for decision makers. Statistical post-processing of deterministic model forecasts should thus also provide probabilistic forecast measures. Here, our focus is on daily precipitation totals. Of interest are e.g. the probability of the occurrence of precipitation, and the expected amount. Particular attention is also given to extreme precipitation events.

A complete probabilistic description of a variable is obtained by an estimate of the conditional distribution function - conditional on the large-scale model forecast. However, there is no general agreement that precipitation can be adequately modeled by a single parametric distribution. Vrac and Naveau (2007) proposed a mixture of Gamma and Generalized Pareto

distribution for modeling precipitation and combined this with a non-homogeneous stochastic weather typing approach for statistical downscaling. Instead of a parametric estimate of the conditional distribution function, the conditional distribution function can be estimated at given values or thresholds. One possibility to do so is to use the logistic multiple regression (Hamill et al. 2004), which gives an estimate of the conditional probability that an a-priori fixed threshold is exceeded. This is a valuable approach for a-priori defined thresholds, particularly for the zero precipitation threshold. Another approach would be to estimate the conditional quantile function at a given probability τ. The method estimates the precipitation amount that is exceeded with a probability of $1 - \tau$. In that case, there is no need to a-priori define thresholds which could be very site- or user-specific, but representative quantiles such as the 0.5 quantile (median), bounds for "normal" conditions such as the 0.25 and 0.75 quantiles, and extreme quantiles, e.g. the 0.05 and 0.95 quantiles can be used.

Here we demonstrate a probabilistic forecast approach that derives such probabilistic measures from the output of a single deterministic model forecast. The probabilistic measures are the probability of precipitation above zero, and the quantile function at given probabilities τ. The method employs censored quantile regression (QR) (Koenker 2005; Powell 1986) and logistic regression (Fahrmeir and Tutz 1994). Our statistical post-processing estimates the probability of precipitation and selected quantiles conditional on the forecasts of the large-scale circulation.

The forecast approach follows Friederichs and Hense (2007) (hereafter FH), who presented a downscaling approach for daily station rainfall data using NCEP reanalysis data. Downscaling, or more general statistical post-processing, seeks for a statistical model between the local variable and the large-scale model output. In FH it is shown how the statistical post-processing derives conditional quantiles of precipitation at one station given the large-scale circulation of the NCEP reanalysis.

In this paper we will extend and modify the post-processing in several ways. The conditional quantile model is trained on either the NCEP or the ERA40 reanalysis data, and then applied to GFS model 12h forecasts. A canonical correlation analysis (CCA) is used to correct for systematic model forecast errors relative to the reanalysis in the large-scale circulation patterns. As this approach uses reanalysis for the training of the probabilistic model, we denote this approach RAN approach as in Marzban et al. (2005), although they use multiple linear regression instead of CCA.

Forecast skill is assessed using the censored quantile verification (CQV) score (FH) and the

Brier score. A skill score is defined with respect to a climatological forecast. The performance of various approaches is compared. Those approaches are the RAN approach, a PP approach where no calibration of the large-scale patterns is applied, a probabilistic MOS approach (P-MOS), where the training of the statistical post-processing relies on 5 years of short-term GFS model forecasts, and a downscaling (D) approach which is similar to the PP approach but uses reanalysis instead of forecasts. The differences assess the role of the various steps in the RAN post-processing. The results are also compared to a somewhat simpler approach that uses calibrated precipitation forecasts from the ECMWF direct model output. The comparison should provide some reference skill in order to rank the skill obtained with the RAN approach. For comparability, the calibration of the ECMWF direct model output again uses censored quantile regression, but no information of the large scale circulation is included.

Section 2 shortly describes the censored quantile regression approach. For a detailed description the reader is referred to FH. Section 3 introduces the data used in this study, and the forecast approach is presented in section 4. Section 5 presents the results using NCEP and ERA40 reanalysis, and compares the various alternative approaches. Section 6 concludes the article. An appendix discusses in more detail the quantile regression and the QV score.

2. Censored quantile regression

Our forecast approach employs censored quantile regression. The concept of quantile regression (QR) has been developed by Koenker and Bassett (1978) as a comprehensive strategy in order to complete the regression picture. While standard regression estimates conditional mean surfaces, QR gives conditional quantile estimates, and thus a more complete measure of the conditional distribution particularly in the case of non-normality. The conditional quantile model coefficients are estimated such that they minimize a loss function derived from the absolute deviations (least absolute deviation (LAD) or L_1 method) (Appendix A). This is equivalent to the maximization of the likelihood function of the data, which can be formulated by independently distributed asymmetric Laplace densities (Yu and Moyeed 2001). In this section, censored QR is introduced in a nutshell. For more insight see FH, and references herein. A very comprehensive description of quantile regression and related subjects is given in the monograph by Koenker (2005). Calculations are performed using the R programming language (R Development Core Team 2003).

In order to account for the mixed discrete-continuous character of daily precipitation totals,

precipitation is represented by a so-called censored variable. Censoring is a term, that originates from survival analysis. In the case of life data, the variable of interest is the time until failure. If the failure occurs during the test period, the value is complete. However, if the test ends before the failure has occurred, the value is incomplete and is said to be censored. The censoring line is the end time of the test period and represents an upper bound of observation, although the underlying process has no upper bound.

In the case of precipitation we assume a hypothetic process Y^* that is observed through the amount of precipitation Y. The zero precipitation line is assumed to represent a censoring line, which acts as a lower bound of observation. The hypothetic process could well have values below zero, although it is not observed through the observable 'precipitation'. Note that this is a statistical construct and does not assume a real physical process.

Let Y be the univariate censored response variable (e.g., daily precipitation totals) and $\mathbf{X}$ the conditioning multivariate variable. Then the statistical censored linear model is

$$Y|\mathbf{X} \;=\; \max(0, \boldsymbol{\beta}^T\mathbf{X} + \boldsymbol{\gamma}^T\mathbf{X}\;\epsilon), \qquad \epsilon \sim \text{IID}. \tag{1}$$

The non-censored process Y^* is modeled by $Y^*|\mathbf{X} = \boldsymbol{\beta}^T\mathbf{X} + \boldsymbol{\gamma}^T\mathbf{X}\;\epsilon$, where $\vec{\beta}$ are the unknown regression coefficients of the model, and identical to $\vec{\beta}$ in (1). The term $\boldsymbol{\gamma}^T\mathbf{X}\;u$ is an error term that accounts for a linear dependency of the square root error variance on the covariate $\mathbf{X}$ (heteroscedasticity). Censoring is applied using the maximum function by taking only values of zero and above, so that censoring is expressed as $Y = \max(0, Y^*)$.

As far as the ordering of the data is not changed, or more general, for every non-decreasing function h, the equality

$$Q_{h(y^*)}(\tau|\mathbf{X}) = h(Q_{y^*}(\tau|\mathbf{X})) \tag{2}$$

holds. This property of the quantile function is used in censored quantile regression. The conditional quantile function $Q_y(\tau|\mathbf{X})$ at a probability τ is derived by

$$\hat{y}_\tau = \hat{Q}_y(\tau|\mathbf{X}) = \hat{Q}_{\max(0,y^*)}(\tau|\mathbf{X}) = \max(Q_{y^*}(\tau|\mathbf{X})) = \max(0, \hat{\boldsymbol{\beta}}_\tau^T\mathbf{X}). \tag{3}$$

The coefficients $\hat{\boldsymbol{\beta}}_\tau$ are estimated by minimizing a piecewise linear censored least absolute deviation function (LAD) (Powell 1986)

$$\hat{\boldsymbol{\beta}}_\tau = \arg\min \sum_n \rho_\tau[y_n - \max(0, \boldsymbol{\beta}_\tau^T\mathbf{x}_n)], \tag{4}$$

where $\rho_\tau(u)$ is the so-called check function, with $\rho_\tau(u) = \tau u$ if $u \geq 0$, and $\rho_\tau(u) = (\tau - 1)u$ if $u < 0$. The index n denotes the members of the training sample. More details on the quantile function and its relation to the optimization problem are given in Appendix A.

The minimization of (4) is much more complex than minimizing the LAD function of the non-censored QR model $\sum_n \rho_\tau[y_n - \boldsymbol{\beta}_\tau^T \mathbf{x}_n]$. We thus use a 3-step approach following Chernozhukov and Hong (2002) which is described in more detail in FH. This approach estimates the non-censored QR model, but on a sub-sample for which the estimated conditional probability of the occurrence of precipitation is larger than the respective probability τ. The procedure is as follows. First, the probability of the occurrence of precipitation $P(y > 0|\mathbf{X})$ is estimated using a logistic regression (with a probit function). A sub-sample $\{y_n\}$ is selected defined by $\{y|P(y > 0|\mathbf{X}) > \tau\}$. In a second step, the $\boldsymbol{\beta}_\tau$ coefficients are estimated minimizing the non-censored LAD function over the sub-sample $\{y_n\}$. A third step repeats the second step, but now on an updated sub-sample $\{y_n\}$ defined by $\{y|\hat{\boldsymbol{\beta}}_\tau^T \mathbf{X} > 0\}$. The coefficients $\boldsymbol{\beta}_\tau$ are updated minimizing the non-censored LAD function on the new sub-sample. Optionally, this step can be repeated several times. Thus, besides the conditional τ-quantile, the method also estimates the probability of non-censoring, hence the conditional probability of precipitation above zero, which is estimated in step one.

Additionally, the censored quantile regression naturally provides a scoring rule, the censored quantile verification (CQV) score

$$\text{CQV} = \sum_n \rho_\tau[y_n - \max(0, \hat{\boldsymbol{\beta}}_\tau^T \mathbf{x}_n)], \tag{5}$$

where y_n and $\mathbf{x}_n$, $n = 1, \ldots, N$, are taken from the forecast and verification sample of Y and $\mathbf{X}$. Further explication on the separation of training and forecast samples is given in Section 4.

The CQV score is a proper scoring (Gneiting and Raftery 2007; Bröcker and Smith 2007), which discourages hedging on the part of the forecaster (Murphy 1973). The CQV score is a positive definite function, and its expected minimum is obtained if the forecasts correspond to the conditional τ-quantile (see Appendix A). Its expectation is zero only if the forecast is perfect and the underlying process deterministic. In order to assess the relative gain in performance of a forecast with respect to a reference forecast, we construct a skill score analogously to the Brier skill score as

$$\text{CQVSS}(\tau) = 1 - \frac{\text{CQV}(\tau)}{\text{CQV}_{ref}(\tau)}. \tag{6}$$

The CQVSS can takes values on the half-open interval $(-\infty, 1]$. A zero CQVSS indicates no gain with respect to the reference forecast, while a value of one indicates a perfect and deterministic forecast. Likewise, the probability forecast of the occurrence of precipitation is verified using the Brier skill score (Brier 1950). Note that the CQVSS as well as the Brier skill score are only asymptotically proper for very large samples (Murphy 1973).

3. Data

A very common problem to model output statistic is the lack of a sufficiently long and homogeneous training data set. We have access to 5 years (2001-2005) of forecasts produced for the 12h period after the 0000UTC initialization time from the NCEP High Resolution Global Forecast System (GFS) (NOAA-EMC 2003). The data were archived at and kindly provided by WetterOnline GmbH. The analysis (0h forecast) was not stored, so no analysis with the GFS model are available. Three variables of the GFS forecasts are available over a region of 1°E-20°E and 45°N-59°N over western Europe: precipitable water (PWat), relative vorticity (ζ_{850}) on the 850hPa pressure level, and vertical velocity (ω) on the 850hPa pressure level.

Our study aims at providing reliable forecasts of daily totals of precipitation at German weather stations. The daily totals are measured from 7.30 GMT to 7.30 GMT of the next day. The observational period thus starts 7.30 hours after the initialization of the model forecasts. Observations at German weather stations are provided by the German Weather Service (DWD) within the priority project 'Quantitative precipitation forecasts' of the German Research Foundation. We have chosen 50 stations in the region of Rhineland-Palatinate, that are almost complete for the forecast period from 2002 to 2005, and that have a sufficient data coverage for the training period from 1958 to 2000. In order to account for seasonal non-stationarities, the data are divided into a cold season, November to March (winter), and warm season, May to September (summer).

The statistical post-processing is derived on the basis of reanalysis data, both from the NCEP re-analysis project (Kalnay and et al. 1996) and ERA40 re-analysis project (Uppala and et al. 2005). Unfortunately, no GFS precipitation forecasts are available. Instead, we used 5 years of the deterministic ECMWF precipitation forecasts (European Centre for Medium-Range Weather Forecasts 2006) over Germany for the period from 2001 to 2005. The ECMWF forecast system provides precipitation forecasts on a $0.4^{\circ} \times 0.4^{\circ}$ horizontal grid. As a predictor we use the 6h-30h accumulated precipitation of the 0000UTC initial time forecast, which approximately corresponds to the observational accumulation period. The ECMWF forecasts are intended to provide some reference skill in order to rank the skill obtained with the RAN approach.

4. The forecast approach

Except when using ECMWF forecasts, the forecast approaches are based on a combined phase state vector ($\mathbf{X}_{GFS}$, $\mathbf{X}_{Re}$) of ζ_{850}, ω_{850}, and PWat at grid points in the area of 1°E-20°E

and $45°$N-$59°$N. In order to give equal weight to the different meteorological variables, the time series have to be weighted. This is done by multiplying each variable with the inverse of its spatial mean temporal standard deviation, which is of order $10^{-4}s^{-1}$ for ζ_{850}, 10^{-1} Pa s^{-1} for ω_{850}, and $50kg\ m^{-2}$ for PWat. The ECMWF state vector $\mathbf{X}_{EC}$ contains the total precipitation amount at 4 neighboring grid points.

In order to assess the importance of the different parts in the post-processing, we apply several approaches additional to the RAN approach. All approaches are summarized in Table 1. For the reason of comparability, the verification period always extends from 2002 to 2005. The quantile forecasts are verified using the CQV and the Brier score. A skill score is derived with respect to a climatological forecast. The climatological probability of precipitation and the climatological quantiles are estimated from the forecast period from 2002 to 2005 for the respective season. The climatology is estimated for each station separately, to avoid artificial skill (Hamill and Juras 2007). The sampling error of the skill scores is estimated using the bootstrap method (Efron and Tibshirani 1993).

a. The RAN approach

Our forecast approach follows the RAN approach of Marzban et al. (2005), i.e. it uses reanalysis for the development of the censored QR model. The access to a large training data base is important for the estimation of a stable QR model as shown in FH. The RAN forecast approach is illustrated in Fig. 1. In a first step, a canonical correlation analysis (CCA) is performed to derive patterns (denoted as canonical patterns) in the GFS forecasts and the NCEP/ERA40 reanalysis with the maximum correlation. Prior to the CCA, a reduction of spatial degrees of freedom is needed for the gridded data sets (NCEP reanalysis, ERA40 reanalysis, GFS forecasts). This reduction is performed using a principal component analysis (Barnett and Preisendorfer 1987). The CCA patterns are derived for the daily fields of the respective season in the year 2001 (Fig. 1). The CCA model between the GFS forecasts and the reanalysis (Re) reads as follows

$$\mathbf{U}\,\mathbf{E}_{Re}\mathbf{X}_{Re} = \mathbf{V}\,\mathbf{E}_{GFS}\mathbf{X}_{GFS} + \boldsymbol{\epsilon}_{CCA}, \tag{7}$$

where $\mathbf{X}_{Re}$ and $\mathbf{X}_{GFS}$ are the gridded and scaled data of reanalysis and GFS forecasts, $\mathbf{E}_{Re}$ and $\mathbf{E}_{GFS}$ are the matrices composed by empirical orthogonal functions, and $\mathbf{E}_{Re}\mathbf{X}_{Re}$ and $\mathbf{E}_{GFS}\mathbf{X}_{GFS}$ constitute the corresponding principal components (PCs) or new variables. $\boldsymbol{\epsilon}_{CCA}$ denotes the error term of the CCA model. $\mathbf{U}$ and $\mathbf{V}$ are matrices containing the canonical pattern or transformations from the CCA. The transpose is denoted by ^{T} and matrices are bold. The

estimates $\hat{\mathbf{U}}$, $\hat{\mathbf{V}}$, and $\hat{\mathbf{E}}_{GFS}$ are derived from the daily data of the year 2001, whereas $\hat{\mathbf{E}}_{Re}$ is derived from the reanalysis period 1948/1958-2001.

In a second step, the NCEP/ERA40 reanalysis of the period from 1948/1958 to 2000 are projected onto the canonical patterns from step 1. This period serves as the training period for the censored QR model between the projections of the NCEP/ERA40 reanalysis onto their canonical patterns ($\mathbf{U}$) and the DWD station precipitation totals at a selected station, notably the estimation of the coefficients $\boldsymbol{\beta}_\tau$. The QR model reads

$$Y|\mathbf{X}_{Re} = \max(0, \boldsymbol{\beta}_\tau^T \, \hat{\mathbf{U}} \, \hat{\mathbf{E}}_{Re}\mathbf{X}_{Re} + \epsilon_\tau), \tag{8}$$

where ϵ_τ accounts for the error in the QR model. The model coefficients $\boldsymbol{\beta}_\tau$ are estimated using data from the training period 1948/1958 to 2000.

The third step is the forecast step for the remaining period from 2002 to 2005. The forecast first corrects for systematic errors in the large-scale variables in the GFS forecasts using the CCA model. Then the censored QR model is applied to derive estimates of the conditional τ-quantile of the station precipitation y. So the forecasts of the conditional τ- quantiles of precipitation are derived as

$$\hat{y}_\tau = \hat{Q}_y(\tau|\mathbf{X}_{GFS}) = \max(0, \hat{\boldsymbol{\beta}}_\tau^T \, \hat{\mathbf{V}} \, \hat{\mathbf{E}}_{GFS}\mathbf{X}_{GFS}), \tag{9}$$

for the remaining period from 2002 to 2005. The RAN approaches are denoted as **RAN ERA40/GFS** and **RAN NCEP/GFS**. Note, that the CCA model avoids the interpolation of the data onto a common grid.

b. The perfect prog approach

Here we want to assess the importance of the CCA correction. The reanalysis derived model is directly applied to the GFS forecasts without the CCA step. As both, the ERA40 and the NCEP grid, have lower resolution than the GFS forecasts, the GFS forecasts are interpolated onto the respective reanalysis grid. The QR model is now trained between the first leading EOFs of the reanalysis data and the DWD station precipitation. The forecasts are estimated through

$$\hat{y}_\tau = \max(0, \hat{\boldsymbol{\beta}}_\tau^T \, \hat{\mathbf{E}}_{Re}\mathbf{X}_{GFS}). \tag{10}$$

Here, the systematic model and forecast errors are ignored (prefect prog), hence the approaches are denoted as **PP ERA40/GFS** and **PP NCEP GFS**. Differences in skill are due to forecasts errors as well as to systematic differences between reanalysis and interpolated GFS forecasts.

c. The downscaling approach

Another approach is included that is obtained for downscaled precipitation using the NCEP reanalysis. The approach is denoted as **D NCEP**. The training of the QR model is based on the NCEP reanalysis from 1948-2001. The quantile estimates for the period from 2002 to 2005 are obtained from

$$\hat{y}_\tau = \max(0, \hat{\boldsymbol{\beta}}_\tau^T \, \hat{\mathbf{E}}_{Re} \mathbf{X}_{Re}). \tag{11}$$

The differences in skill between RAN NCEP/GFS and D NCEP are mainly due to forecast errors.

d. The P-MOS approach

In order to emphasize the benefit of the RAN approach we included a probabilistic MOS approach, where only the relatively short GFS forecasts are used for training the post-processing. The training and verification only uses the GFS forecasts, and is denoted as **P-MOS GFS**. The forecasts are derived using cross-validation, where one target season (NDJFM or MJJAS) is withheld from the training data, and the forecast

$$\hat{y}_\tau = \max(0, \hat{\boldsymbol{\beta}}_\tau^T \, \mathbf{E}_{GFS} \mathbf{X}_{GFS}). \tag{12}$$

is derived for each target season in the period from 2002 to 2005. The training data also include the data from the year 2001.

e. Using precipitation (ECMWF) as predictor

This censored QR approach, denoted as **DMO ECMWF**, uses only precipitation forecasts. As no GFS precipitation forecasts were available, we use the ECMWF forecasts instead. This approach assesses the skill obtained using direct precipitation forecasts. This approach is not directly comparable with the GFS forecasts based on the large-scale information only, as it introduces a new model with a higher resolution. Rather it should give some reference to rank the obtained skill scores.

The covariate to derive the forecast is again a multivariate covariate, however, it only consists of forecasts of total precipitation at the four nearest grid points around the respective station, therefore no reduction of degrees of freedom is needed. However, a simple calibration is needed to derive probabilities and quantiles from the deterministic DMO, which are comparable. As for the P-MOS GFS approach we use cross-validation for training and verification. The

conditional quantiles of the target season are estimated as

$$\hat{y}_\tau = \max(0, \hat{\boldsymbol{\beta}}_\tau^T \, \mathbf{X}_{EC}). \tag{13}$$

$\mathbf{X}_{EC}$ now is a 5-dimensional vector containing the forecasts of total precipitation at the four grid points nearest to the station location and the one.

5. Results

We first discuss the results obtained with the RAN post-processing. In order to present meaningful results, the sensitivity of the quantile forecasts and the CQV skill scores to different settings has to be assessed. First, the optimal number of EOF and canonical patterns that enter the QR model has to be defined. It turned out that an optimal number of EOF is of the order of 12 to 14, and the QR model then gives best results with about the same number of CCA modes (not shown). We have chosen 14 EOFs to represent the GFS forecasts and the NCEP/ERA40 reanalysis in the CCA model. All 14 CCA modes enter the QR model. Although the length of the training period for the CCA correction is only one year, about 150 days for each season, it turned out to suffice for our purpose. A longer period (2 seasons) for the CCA correction in turn reduces the period for the verification and did not significantly increase the skill.

Figure 2 shows as an example the CQV skill scores for 4 stations in Rhineland-Palatinate, Wachtberg-Berkum (50.62°N, 7.13°E, 220m), Schmelz-Hüttersdorf(49.42°N, 6.83°E, 223m), Helmbach-Düttling (50.60°N, 6.55°E, 380m), and Hermeskeil (49.65°N, 6.93°E, 480m). Wachtberg-Berkum (Helmbach-Düttling) represents the station with lowest, and Schmelz-Hüttersdorf (Hermeskeil) with largest skill in winter (summer). Skill in winter is generally larger than in summer. The Brier skill score ranges between 23% (19%) and 50% (40%) in winter (summer). The CQV skill score is generally largest for the 0.9 or 0.95 quantile and is of the order of the Brier skill score. The CQV skill score of the 0.99 quantile is smaller and has a large uncertainty. Least skill is obtained for the 0.25 quantile. However, this quantile lies only above the zero line, if the probability of precipitation is above 0.75. The error bars given in Fig. 2 represent the 95% confidence interval of the CQV skill score estimate derived by a bootstrap approach. No significant differences exist between the forecasts that are derived either using ERA40 or NCEP reanalysis.

Figure 2 also represents the CQV skill score obtained using the calibrated DMO ECMWF forecasts. For all stations, the DMO ECMWF approaches outperforms the RAN approaches. The CQV skill score reaches 70% (50%) for some stations in winter (summer). The sampling

error of the CQV skill score is comparable to that of the RAN approaches. Further studies are needed in order to assess whether the differences are due to the different model performance (e.g. due to the higher horizontal resolution of the ECMWF forecasts), or the choice of the covariates, or both.

In order to further verify the quantile forecasts, we investigated, whether the quantile forecasts are reliable, i.e. show systematic errors (Fig. 3). If a quantile forecast is reliable then it is expected that a fraction τ of the observed precipitation values lies below the conditional τ-quantiles. This fraction divided by the total number of observations is denoted as the empirical probability. Note that the empirical probabilities can only be estimated for those forecasts, where the conditional quantile forecast is above zero. If the quantile forecast is censored, which means that it is set to zero, and the observation is also zero, then it is not possible to decide whether it is below or above the quantile forecasts. Counting the zero quantile forecasts induces an artificial bias to the reliability, which is particularly strong for the lower quantiles, as they are more often censored. The reliability of the censoring, i.e. the reliability of the categorical forecast of precipitation above zero, should be assessed separately.

Fig. 3 compares the theoretical (horizontal lines) and the empirical probabilities (dots) of the conditional τ-quantile forecasts. The 95% sampling error intervals of the empirical probability estimates are indicated by vertical lines. They are estimated from a 1000 member bootstrap sample. Although some significant deviations from the theoretical values occur, particularly for the 0.25-quantile forecasts, a χ^2-tests for count data (Wilks 1995) does not reject the null-hypothesis of reliability for any of the forecast approaches. Similarly, the mean occurrence of precipitation lies within the uncertainty of the mean forecast occurrence (not shown).

As the conditional quantiles are estimated independently for the different levels of τ, it may occur that the quantile forecasts cross (i.e. the conditional 0.9 quantile is larger than 0.95 quantile). This would constitute an invalid distribution. FH have shown, that oversampling and a short training data set can lead to an increased number of crossing quantiles. For the four stations shown in Fig. 2, and for the six quantile forecasts for each day of the winter and summer seasons from 2002 to 2005 (1217 days), crossing occurred five times for the GFS/ERA40, never for the GFS/NCEP forecast approach, and 333 times for the forecasts using the QR model based on the ECMWF precipitation forecasts. Even though, the ECMWF forecast approach provides more skillful quantile forecasts, it predicts significantly more often an invalid distribution than the GFS/ERA40 (GFS/NCEP) reanalysis approach. These results are completely consistent with FH.

We now want to investigate the importance of the different parts of the RAN approach. Therefore, the CQV skill score of the remaining forecast approaches are displayed in Fig. 4. The PP ERA40/GFS and PP NCEP/GFS approaches highlight the benefit from the CCA correction, which should correct for systematic errors in the forecasts of the large-scale circulation, that are due to the forecast errors as well as to systematic differences between the reanalysis and the GFS model. For all stations and seasons, the skill of PP ERA40/GFS is comparable to RAN ERA40/GFS. The spatial resolution of both models is comparable ($1.25°$ for ERA40 and $1.0°$ for GFS), hence the systematic differences between the ERA40 and GFS large-scale patterns as well as systematic forecast errors seem small. In contrast, the performance of PP NCEP/GFS is reduced compared to RAN NCEP/GFS for all stations and seasons. This suggests that the systematic model differences are larger between NCEP ($2.5°$ resolution) and GFS.

Similar conclusions are suggested by the performance of the downscaling approach (D NCEP). The skill for the downscaling, which should not contain forecast errors, is comparable to those of the RAN approaches. Thus, the systematic model errors have a larger effect on the skill than the forecast errors. Note, that the forecast lead time is 12h, and that the forecast error will certainly be more pronounced with increasing lead time.

Finally, we assess the benefit of having large training samples. As the training of the P-MOS GFS approach only relies on 4 seasons of daily data, the data reduction of the large-scale fields uses the first leading 10 EOFs. The skill is reduced for almost all stations compared to the RAN ERA40/GFS or RAN NCEP/GFS approaches. More importantly, the uncertainty of the quantile forecasts is largely increased due to the short training period, which results in highly variable CQV and Brier skill scores. The amount of crossing quantiles which is of the order of 1/1000 is only slightly increased compared to the RAN approaches.

Fig. 5 shows an example of forecasts derived using the GFS/ERA40 approach for January (Fig. 5 a)) and August 2002 (Fig. 5 b)). The winter and summer climatologies of the daily precipitation totals as estimated from the period 2002 to 2005 are indicated in a box on the left-hand side of the panels. The climatological probability of precipitation amounts to 0.56 in winter and 0.44 in summer. In summer, the climatological 0.25 and 0.50 quantiles lie on the zero precipitation line. The 0.9 quantile amounts to $8.7mm$ ($8.2mm$) and the 0.99 quantile to $26.5mm$ ($27.9mm$) in winter (summer). The black dots indicate the observed precipitation amount. The first 10 days of January 2002 were very dry. Accordingly, the probability of precipitation estimated for those days lies below 0.2. In the second half of January 2002, the probability of precipitation was much higher, with a probability reaching almost one on January

27. On that day, the value that could be exceeded with a probability of 0.1 amounts to 24mm, and the 0.99 quantile lies at 40mm. The observed value amounts to 17mm. The forecasts follow the observations and capture nicely the range of variance of the daily precipitation totals. Similar good correspondence is obtained for August 2002 at the station Schmeltz-Hüttersdorf. The forecast captures the dry period from August 13th to August 18th as well as the extreme rainfall on August 20th of about 40mm.

We now look more closely at January 27, 2002 and August 20, 2002. During both days, one winter and one summer day, the total precipitation exceeded 40mm at one station. Figs. 6 a,c show the observed precipitation totals at 50 stations in Rhineland-Palatinate on January 27, 2002 and August 20, 2002. On January, 27th small cyclones passed over Germany within a distinct west-south-west drift. The situation on August, 20th was marked by two small low pressure systems coming from France that transported warm and humid air from the south-west leading to shower and thunderstorms. On both days, the observed precipitation totals strongly vary between the station with values ranging from 1.5mm (0.7mm) to 44.7mm (41mm) on January 27 (August 20).

The 0.95 quantile forecasts for the respective winter and summer days are shown in the right panels of Fig. 6. On January 27, the 0.95 quantile forecasts range between 14.7mm and 37.7mm. In contrast to the observed precipitation totals, the quantile forecasts are much smoother in space. But this can be expected since quantiles similar to expectation values characterize the probability distributions, while the observations constitute a sample of realizations of those distributions. The 0.95 quantile forecasts on August, 20th show even smaller spatial variance than in January and range around 20mm.

6. Conclusions

We presented a probabilistic post-processing that derives probabilistic measures of daily precipitation on the basis of the large-scale circulation patterns of a single deterministic model forecast. Here, 12h forecasts are taken from the NCEP High Resolution Global Forecast System (GFS). Observations are daily precipitation totals at 50 weather stations in the region of Rhineland-Palatinate, Germany, provided by the German Weather Service. The post-processing employs canonical correlation analysis, logistic regression, and censored quantile regression. It constitutes a probabilistic extention to standard model output statistics.

As Friederichs and Hense (2007) pointed out, a sufficiently long training period is needed in

order to reduce the probability of an invalid distribution forecast. Following the RAN approach of Marzban et al. (2005), the training of the probabilistic model, namely the censored quantile model, is derived on reanalysis data, either from ERA40 or NCEP. On the basis of more than 40 years of daily data a stable and meaningful QR model is estimated. A CCA corrects for systematic differences due to different grids and systematic model and forecast errors between the large-scale circulation patterns of the reanalysis and the GFS forecasts.

Forecast skill is assessed using the CQV and Brier skill score, where the reference forecast constitutes an estimate of the climatological distribution of precipitation. The reliability of the quantile forecasts is good, and the occurrence of crossing quantile forecasts is negligible. Skill is generally larger in winter than in summer, and ranges between 20% to 50% depending on the respective quantile, season and station. Although NCEP reanalysis have a lower resolution than the ERA40 data, no significant differences are obtained when using NCEP or ERA40 for the training of the quantile model.

The performance of the probabilistic RAN approach, which is based on the large-scale atmospheric patterns is compared to a DMO approach, that uses precipitation forecasts from the ECMWF forecast system. The ECMWF DMO approach provides significantly more skill than the RAN approach. Further studies are needed to assess those differences. Other forecasts variables should be investigated and particularly the GFS precipitation forecasts should be included. The occurrence of crossing quantiles of the DMO ECMWF forecasts is small but not negligible. It indicates that although the skill is large, the data basis is not large enough to estimate a stable QR model.

Several other forecast approaches are tested in order to assess the effect of the different components of the RAN approach. The advantage of the CCA correction is that an interpolation on a common grid is not necessary. It turns out, that the CCA correction of the large-scale patters is important when using NCEP reanalysis to train the censored QR model, whereas no skill is lost when using ERA40 reanalysis. So a correction of the systematic differences between the NCEP and GFS large-scale patterns is important. The forecast error seems small for the 12h forecasts, as the performance of the RAN approach is comparable to the performance of downscaling based on NCEP reanalysis. A P-MOS approach that relies solely on the large-scale patterns of the GFS forecasts provides a less stable and less meaningful statistical post-processing, as the training period is too short. Here, the RAN approach is clearly outperforming the P-MOS approach.

The objective of this study was to draw attention to the method of censored quantile re-

gression that provides a tool to formulate a probabilistic post-processing of numerical weather forecasts. The utilization of reanalysis for the training of a probabilistic MOS system together with the RAN-type approach is appropriate to avoid problems due to the data shortage problem. The RAN approach represents a valuable and efficient post-processing to derive local probabilistic precipitation forecasts from a single deterministic model forecast.

Acknowledgments.

We greatfully acknowledge fruitful discussions with Martin Goeber, Armin Mathes and Jan Keller. Special thanks are also due to three anonymous reviewers for their detailed comments, and to WetterOnline GmbH for providing us with the GFS forecasts. Parts of the research has been carried out within the DFG funded project PP1167-PQP and at the Interdisciplinary Centre of Complex Systems (IZKS) of the University of Bonn.

APPENDIX A

Quantiles and the quantile verification (QV) score

This appendix gives some preliminaries about quantile estimation, notably, that the quantiles can be expressed as the solution of an optimization problem (Koenker 2005). Standard linear regression minimizes the least square error, and the solution of an optimization problem that aims at minimizing the expected square error is the expectation value. And so the mean square error is an adequate measure of the performance of such a forecast. Analogously, quantile regression minimizes a function of the least absolute differences. This optimization problem is the basic principle for quantile regression, and the solution is the conditional quantile function at a probability τ. Likewise, the optimization problem defines a verification score, the quantile verification (QV) (or censored quantile verification (CQV)) score, which is discussed in some detail here.

Let Y be a univariate random variate with distribution $F(y) = Prob(Y \leq y)$. The quantile function is defined as

$$Q(\tau) = F^{-1}(\tau) = \inf\{y|F(y) \geq \tau\}, \tag{A1}$$

and $F^{-1}(\tau) = y_\tau$ is called the τ-quantile. Quantiles arise from a simple optimization problem, that is defined through the so-called check function

$$\rho_\tau(u) = u(\tau - I(u < 0)) = (\tau - 1)uI(u < 0) + \tau uI(u \geq 0) \quad \text{for some} \quad \tau \in (0, 1). \tag{A2}$$

The check function $\rho_\tau(u)$ is a function of the least absolute deviations with $u = y - \hat{y}_\tau$ and is displayed in Fig. 7. $I(.)$ is an indicator function, which takes the value 1 if the condition in the bracket is valid, and 0 if the condition is not valid. The estimate $\hat{y}_\tau$ should be determined such that it minimizes the expected loss ($E[.]$ indicates the expectation under $F(y)$)

$$E[\rho_\tau(Y - \hat{y}_\tau)] = \int_{-\infty}^{\hat{y}_\tau} (\tau - 1)(y - \hat{y}_\tau)dF(y) + \int_{\hat{y}_\tau}^{\infty} \tau(y - \hat{y}_\tau)dF(y). \tag{A3}$$

Differentiating with respect to $\hat{y}_\tau$ results in

$$\frac{\partial}{\partial \hat{y}_\tau} E[\rho_\tau(y - \hat{y}_\tau)] = (1 - \tau) \int_{-\infty}^{\hat{y}_\tau} dF(y) - \tau \int_{\hat{y}_\tau}^{\infty} dF(y) = F(\hat{y}_\tau) - \tau \stackrel{!}{=} 0, \tag{A4}$$

so a necessary condition of $\hat{y}_\tau$ to minimize the loss function is that $F(\hat{y}_\tau) = \tau$, which is indeed the τ-quantile. Analogously, it can be shown that the value that minimizes $E[(Y - \hat{y})^2]$ is the expectation value $E[Y]$ of Y.

The loss function $\rho_\tau(y - \hat{y}_\tau)$ is proposed as a scoring rule for quantile forecast verification, the QV score (for censored data CQV score). It will be shown in the following, that the expected loss attains a minimum if the forecast is perfect. This proves that the QV score is proper, but also that $\hat{y}_\tau = F^{-1}(\tau)$ is a necessary and sufficient condition for a minimum of the expected loss. We first recall the definition of a proper score (for details see Gneiting and Raftery (2007)). Let $S(y, \hat{y}_\tau) = \rho_\tau(y - \hat{y}_\tau)$ denote the QV score of a forecast $\hat{y}_\tau$ of the τ-quantile of the random variate $Y \sim F(y)$. Let y_τ be the τ-quantile, with $F(y_\tau) = \tau$ or $y_\tau = F^{-1}(\tau)$, then $S(y, .)$ is proper if for each $\hat{y}_\tau$

$$E[S(Y, \hat{y}_\tau)] \geq E[S(Y, y_\tau)], \tag{A5}$$

and strictly proper if equality is only obtained if $\hat{y}_\tau = y_\tau$. The expectation of each score can be trivially decomposed into

$$E[S(Y, \hat{y}_\tau)] = E[S(Y, y_\tau)] + \{E[S(Y, \hat{y}_\tau)] - E[S(Y, y_\tau)]\}\,. \tag{A6}$$

If $S(Y, .)$ is a proper score, then the term in the braces is positive definite. Such a decomposition for the QV score can be derived as follows.

$$\begin{aligned}
E[\rho_\tau(y - \hat{y}_\tau)] &= \int_{-\infty}^{\infty} \rho_\tau(y - \hat{y}_\tau) dF(y) \\
&= \int_{-\infty}^{\hat{y}_\tau} (\tau - 1)(y - \hat{y}_\tau) dF(y) + \int_{\hat{y}_\tau}^{\infty} \tau(y - \hat{y}_\tau) dF(y) \\
&= \int_{-\infty}^{\infty} \tau(y - y_\tau) dF(y) - \int_{-\infty}^{\hat{y}_\tau} (y - y_\tau) dF(y) \\
&\quad - \int_{-\infty}^{\infty} \tau(\hat{y}_\tau - y_\tau) dF(y) + \int_{-\infty}^{\hat{y}_\tau} (\hat{y}_\tau - y_\tau) dF(y) \\
&= \int_{-\infty}^{\infty} \tau(y - y_\tau) dF(y) - \int_{-\infty}^{y_\tau} (y - y_\tau) dF(y) - \int_{y_\tau}^{\hat{y}_\tau} (y - y_\tau) dF(y) \\
&\quad - \tau(\hat{y}_\tau - y_\tau) \int_{-\infty}^{\infty} dF(y) + (\hat{y}_\tau - y_\tau) \int_{-\infty}^{\hat{y}_\tau} dF(y) \\
&= E\,\rho_\tau(Y - y_\tau) + \left\{ (F(\hat{y}_\tau) - \tau)(\hat{y}_\tau - y_\tau) - \int_{y_\tau}^{\hat{y}_\tau} (y - y_\tau) dF(y) \right\}
\end{aligned} \tag{A7}$$

Due to the mean value theorem for integration, the integral in the braces is bounded by

$$0 \leq \int_{y_\tau}^{\hat{y}_\tau} (y - y_\tau) dF(y) \leq (\hat{y}_\tau - y_\tau)(F(\hat{y}_\tau) - \tau). \tag{A8}$$

As $F(y)$ is a positive definite and monotonically increasing function of y, the term in the brackets is positive definite. Thus the QV score is a proper score.

References

Barnett, T. P. and R. Preisendorfer, 1987: Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. *Mon. Wea. Rev.*, **115**, 1825–1850.

Bremnes, J. B., 2004: Probabilistic forecasts of precipitation in terms of quantiles using NWP model output. *Mon. Wea. Rev.*, **132**, 338–347.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1–3.

Bröcker, J. and L. A. Smith, 2007: Scoring probabilistic forecasts: The importance of being proper. *Weather and Forecasting*, **22**, 382–388.

Chernozhukov, V. and H. Hong, 2002: Three-step censored quantile regression, with an application to extramarital affairs. *J. Am. Stat. Assoc.*, **97**, 872–882.

Efron, B. and R. J. Tibshirani, 1993: *An Introduction to the Bootstrap*. Chapman & Hall, 436 pp.

European Centre for Medium-Range Weather Forecasts, 2006: ECMWF operational analysis data. British Atmospheric Data Centre. 2006-2007. Available from http://badc.nerc.ac.uk/data/ecmwf-op/.

Fahrmeir, L. and G. Tutz, 1994: *Multivariate Statistical Modelling Based on Generalized Linear Models*. Springer series in statistics, Springer, New-York, 425 pp.

Friederichs, P. and A. Hense, 2007: Statistical downscaling of extreme precipitation using censored quantile regression. *Mon. Wea. Rev.*, **135**, 2365–2378.

Glahn, H. R. and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203–1211.

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

Hamill, T. M. and J. Juras, 2007: Measuring forecast skill: is it real skill or is it the varying climatology? *Q. J. R. Meteorol. Soc.*, **132**, 2905–2923.

Hamill, T. M. and J. S. Whitaker, 2006: Probabilistic quantitative precipitation forecasts based on reforecast analogs: theory and application. *Mon. Wea. Rev.*, **134**, 3209–3229.

Hamill, T. M., J. S. Whitaker, and X. Wei, 2004: Ensemble re-forecasting: improving medium-range forecast skill using retrospective forecasts. *Mon. Wea. Rev.*, **132**, 1434–1447.

Kalnay, E., 2003: *Atmospheric modeling, data assimilation and predictability*. Cambridge University Press, 341 pp.

Kalnay, E. and et al., 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.

Klein, W. H., 1971: Computer prediction of precipitation probability in the United States. *J. Appl. Meteorol.*, **10**, 903–915.

Koenker, R., 2005: *Quantile regression*, Econometric Society Monographs, Vol. 38. Cambridge University Press, 349 pp.

Koenker, R. and B. Bassett, 1978: Regression quantiles. *Econometrica*, **46**, 33–49.

Marzban, C., S. Sandgathe, and E. Kalnay, 2005: MOS, perfect prog, and reanalysis data. *Mon. Wea. Rev.*, **134**, 657–663.

Murphy, A. H., 1973: Hedging and skill scores for probability forecasts. *J. Appl. Meteor.*, **12**, 215–223.

NOAA-EMC, 2003: The GFS atmospheric model. *NCEP Office Note No. 442*, [http://www.emc.ncep.noaa.gov/officenotes/newernotes/on442.pdf].

Powell, J. L., 1986: Censored regression quantiles. *Journal of Econometrics*, **32**, 143–155.

R Development Core Team, 2003: R: A language and environment for statistical computing. available at http://www.R–project.org.

Uppala, S. M. and et al., 2005: The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012.

Vislocky, R. L. and G. S. Young, 1989: The use of perfect prog forecasts to improve model output statistics forecasts of precipitation probability. *Weather and Forecasting*, **4**, 202–209.

Vrac, M. and P. Naveau, 2007: Stochastic downscaling of precipitation: From dry events to heavy rainfalls. *Water Resour. Res.*, in press.

Wilks, D., 1995: *Statistical Methods in the Atmospheric Sciences: an Introduction*, International Geophysics Series, Vol. 59. Academic Press, 464 pp.

Yu, K. and R. Moyeed, 2001: Bayesian quantile regression. *Statistics and Probability Letters*, **54**, 437–447.

Zorita, E. and H. von Storch, 1998: The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. *J. Climate*, **12**, 2474–2489.

List of Figures

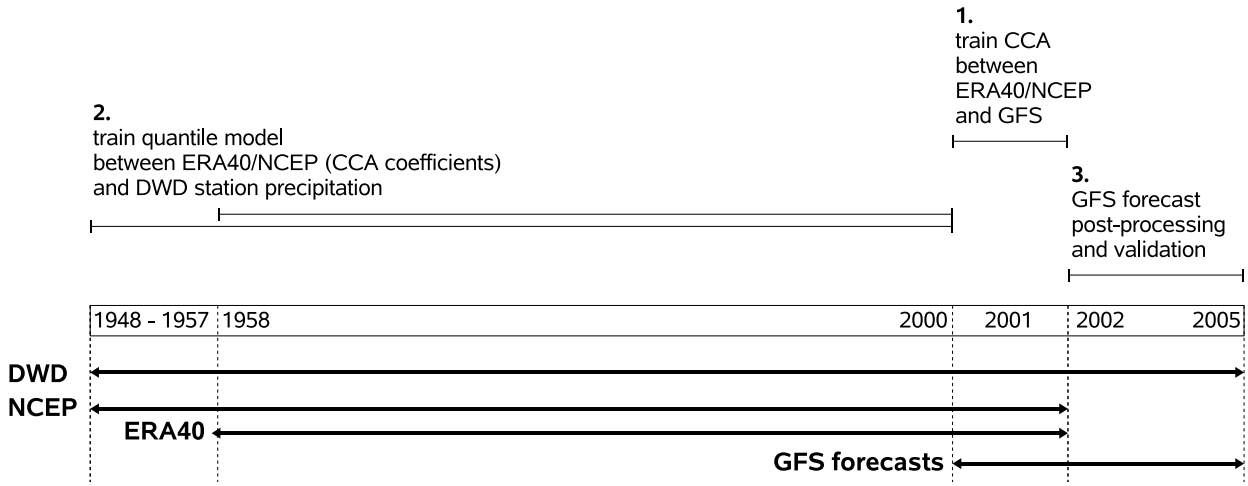

FIG. 1. Illustration of the three step RAN forecast approaches (GFS/ERA40 and GFS/NCEP).

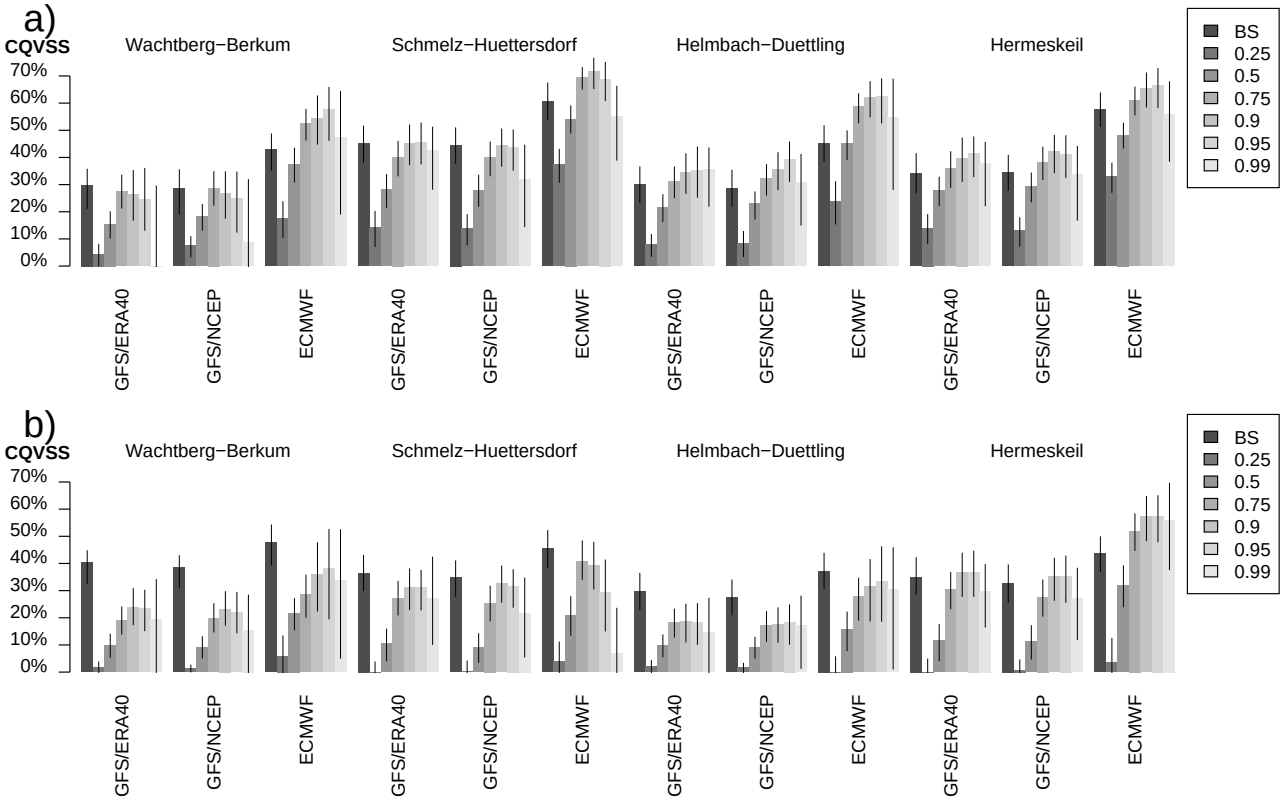

FIG. 2. CQV and Brier skill score for censored QR forecasts for 4 DWD stations for a) winter months and b) summer months. Training of QR model is performed on ERA40 and NCEP reanalysis ζ_{850}, ω_{850}, and PWat, and on ECMWF precipitation forecasts. Forecast is performed on the basis of GFS (GFS/ERA40, GFS/NCEP) or ECMWF precipitation forecasts (see Table 1). The shading of the Brier skill score for the probability forecast of the occurrence of precipitation is dark, the CQV skill score is shown for the 0.25, 0.5, 0.75, 0.90, 0.95, and 0.99 quantile forecast with grey shading. The error bars indicate the sampling errors of the CQV and Brier skill score.

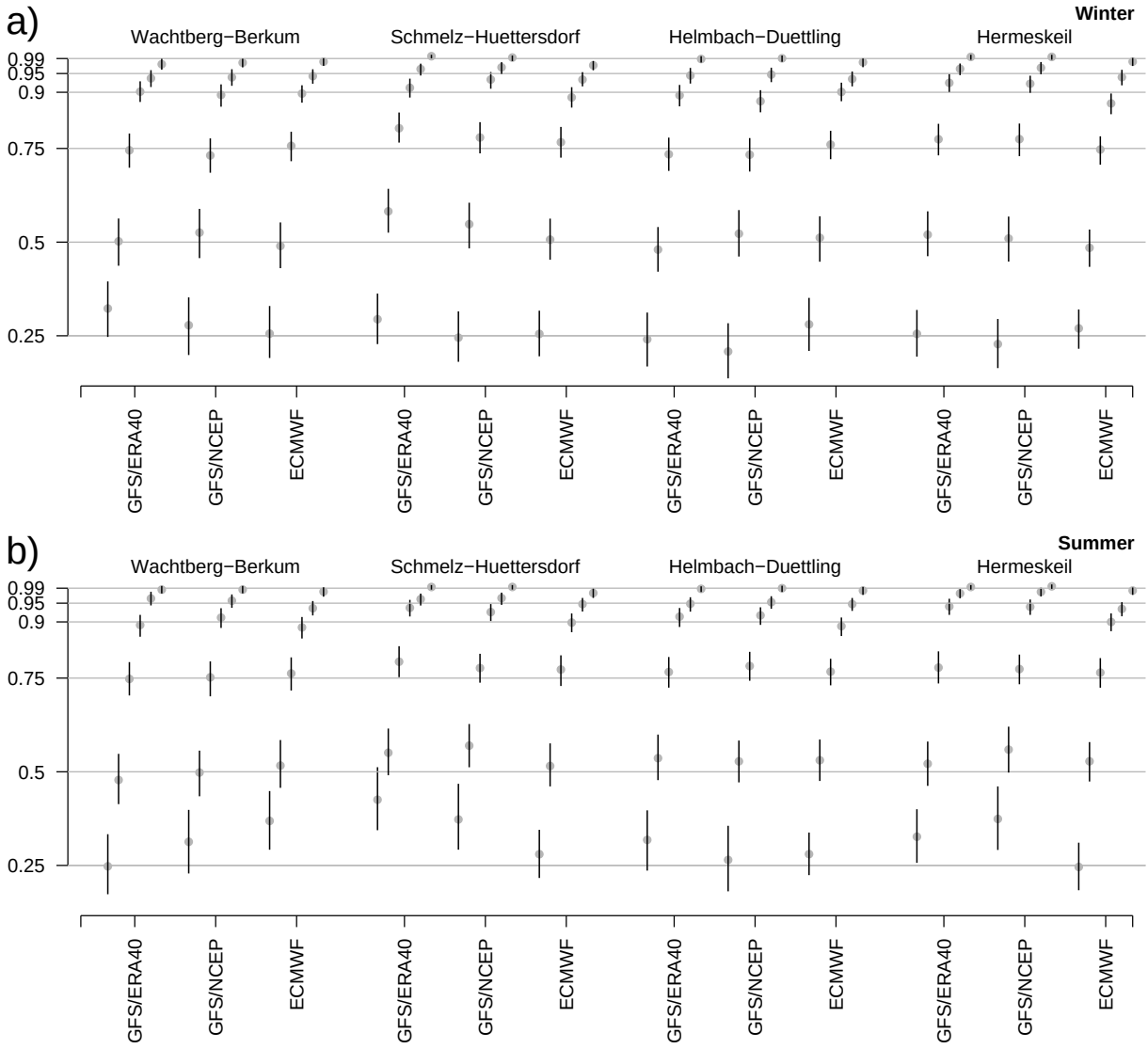

FIG. 3. Reliability for censored QR forecasts for 4 DWD stations for a) winter months and b) summer months. Training and forecast approaches as in Fig. 2. Horizontal lines indicate the theoretical probability, and the gray dots the empirical probability of conditional quantile forecasts (τ-values as in Fig. 2). The vertical bars represent the 95% error interval of the empirical probabilities estimated from a 1000 member bootstrap sample.

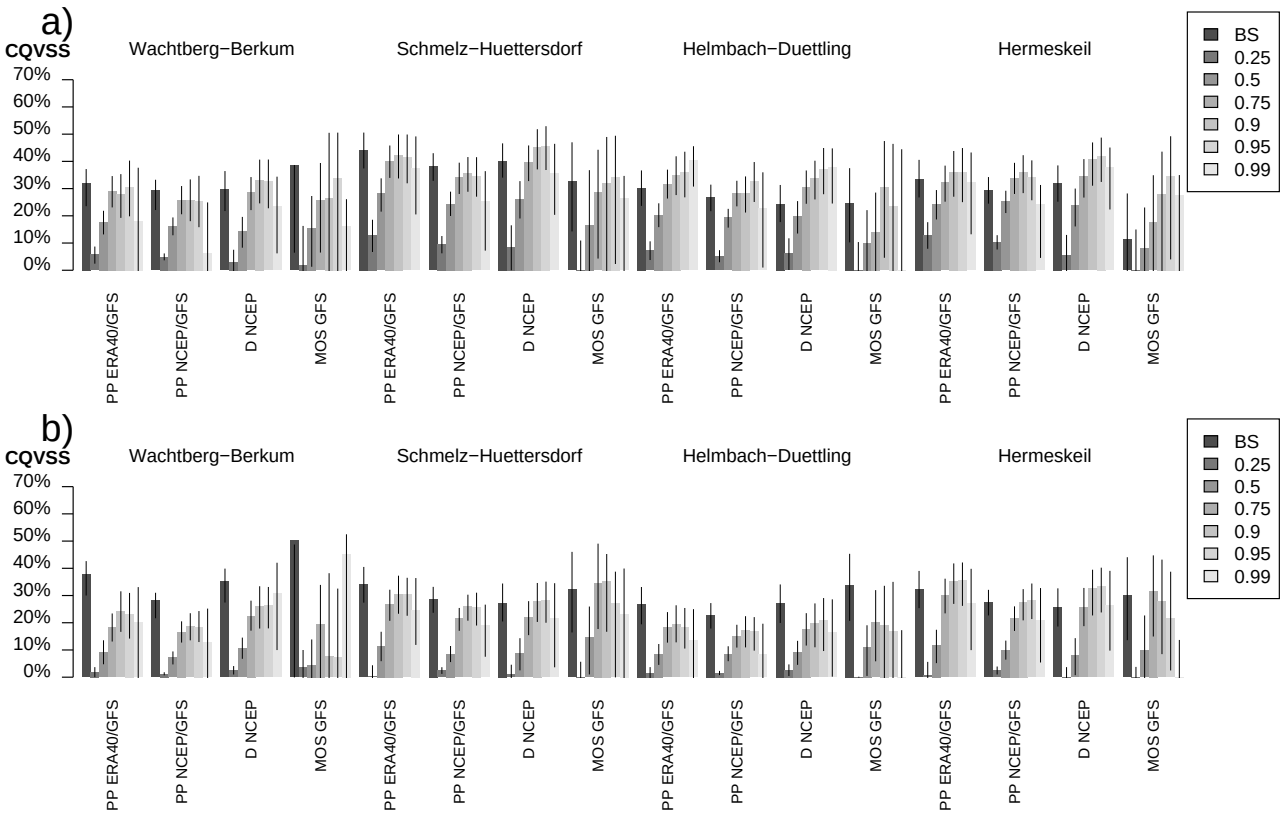

FIG. 4. CQV and Brier skill score for censored QR forecasts for 4 DWD stations for a) winter months and b) summer months as in Fig. 2), but for different forecast approaches (see Table 1).

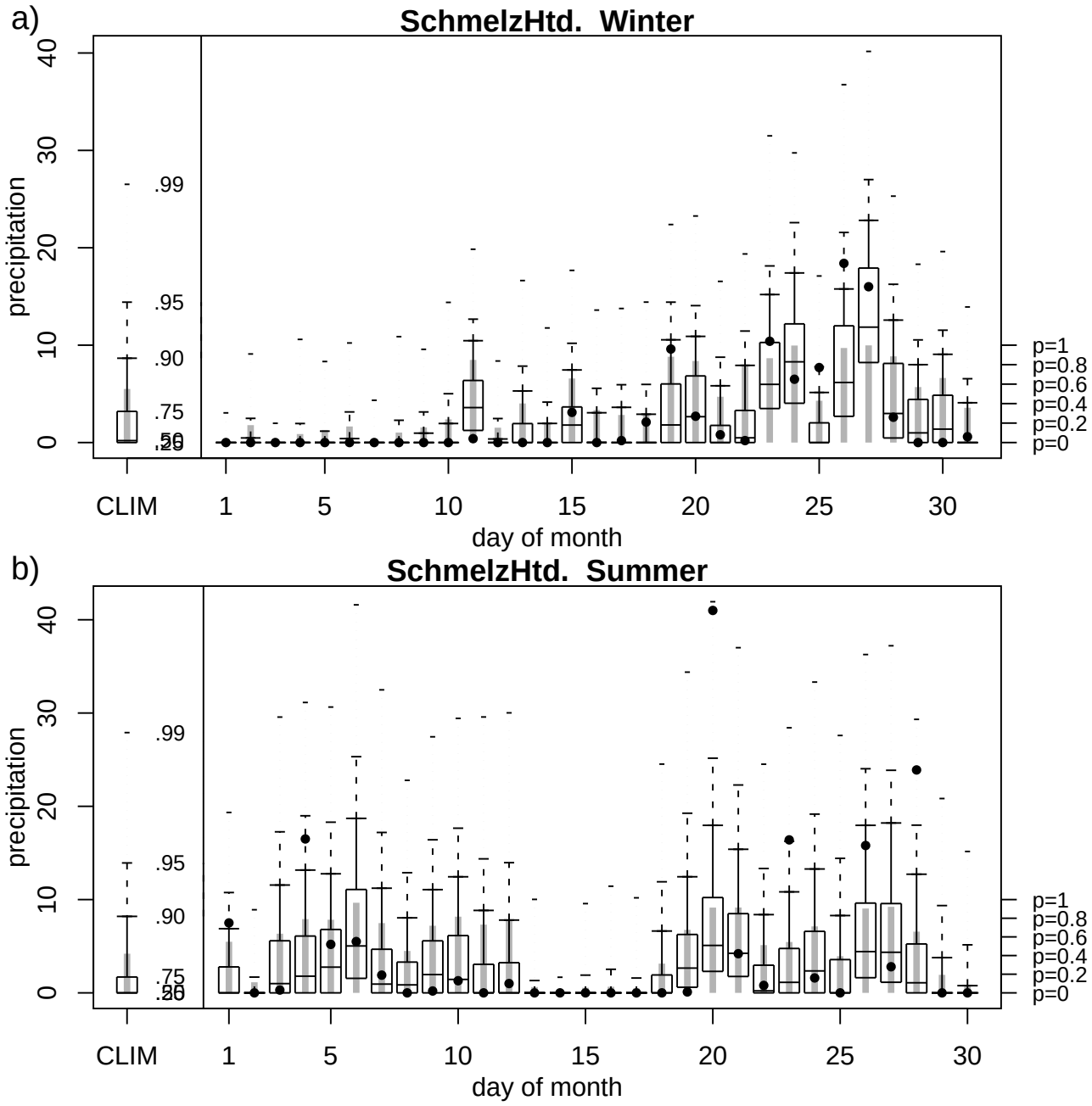

FIG. 5. Example of censored QR forecasts at station Schmelz-Huettersdorf for the months a) January 2002 and b) August 2002 using ERA40 reanalysis and GFS forecast. The outer left part of the panels shows the climatological forecast. Grey bars give the forecast probability of precipitation (right axis), black circles are the observed precipitation sums, and the Box-Whisker graphs indicate the 0.25, 0.50, and 0.75 quantiles as well as the extreme quantiles 0.90, 0.95, 0.99 as indicated for the climatological forecast. The left axis is in mm.

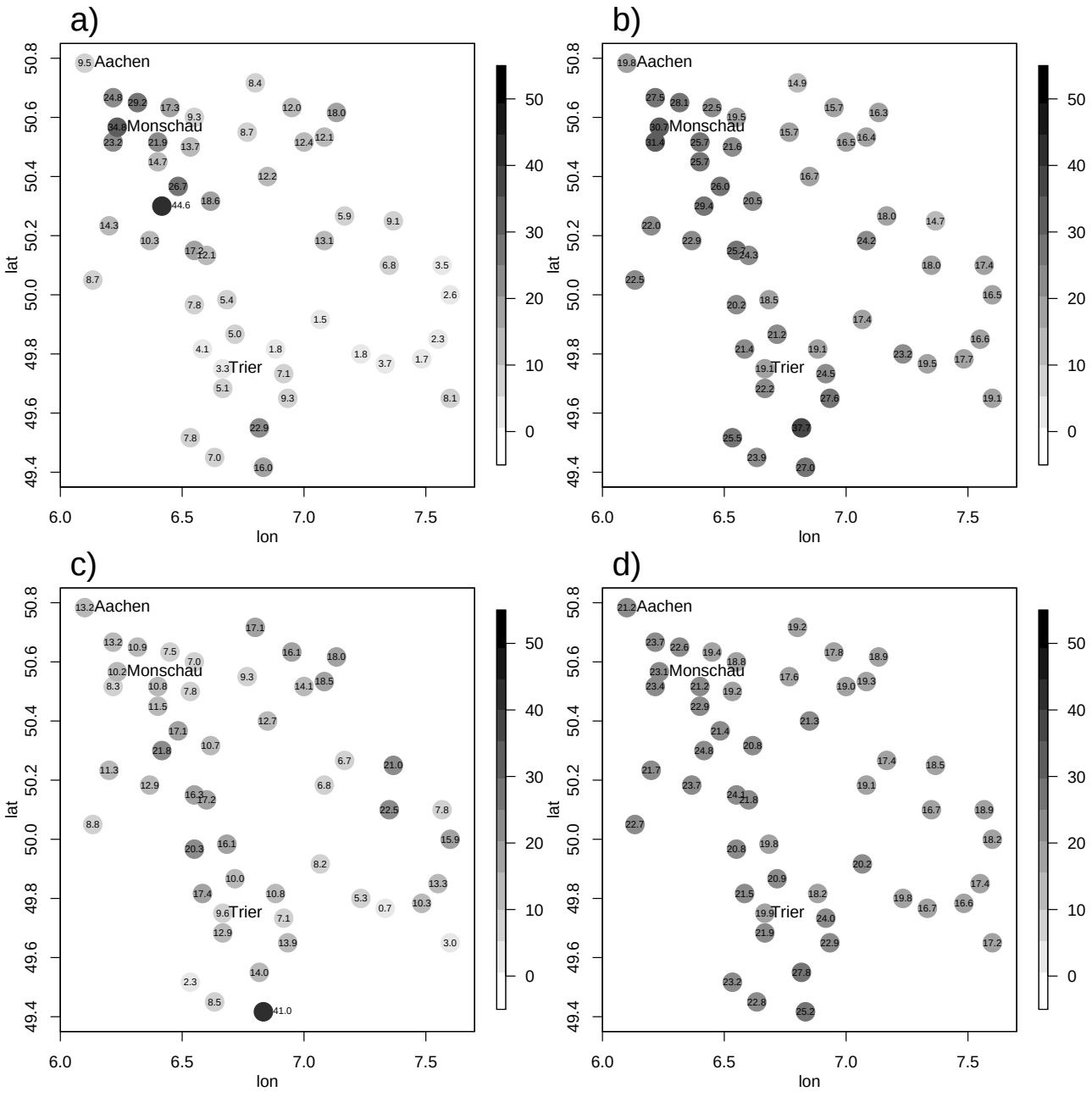

FIG. 6. Observed precipitation for all 50 stations on a) January 27, 2002 and c) August 20, 2002, and 0.95 quantile forecasts at 50 stations, b) January 27, 2002 and d) August 20, 2002.

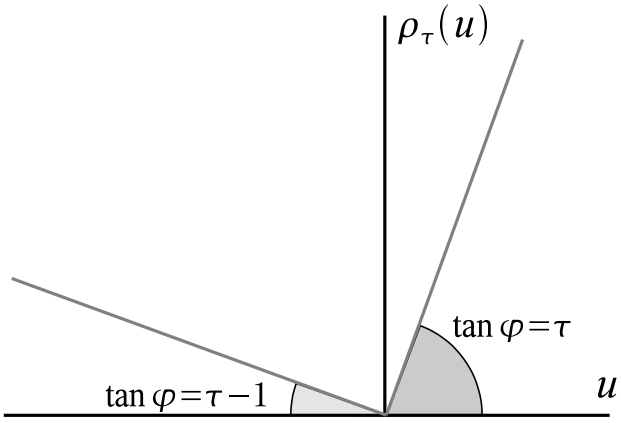

FIG. 7. Check function $\rho_\tau(u)$.

List of Tables

TABLE 1. Forecast approaches employed in this study.

Acronym	Training QR model	Forecast of quantiles	Variables
RAN ERA40/GFS	ERA40 1958-2000	GFS 2002-2005	14 CCA modes (ζ_{850}, ω_{850}, PWat) between ERA40 and GFS (2001)
RAN NCEP/GFS	NCEP 1948-2000	GFS 2002-2005	14 CCA modes (ζ_{850}, ω_{850}, PWat) between NCEP and GFS (2001)
PP ERA40/GFS	ERA40 1958-2001	GFS 2002-2005 (interpol. on ERA40 grid)	14 EOFs (ζ_{850}, ω_{850}, PWat) of ERA40 (1958-2001)
PP NCEP/GFS	NCEP 1948-2001	GFS 2002-2005 (interpol. on NCEP grid)	14 EOFs (ζ_{850}, ω_{850}, PWat) of NCEP (1948-2001)
D NCEP	NCEP 1948-2001	NCEP 2002-2005	14 EOFs (ζ_{850}, ω_{850}, PWat) of NCEP
P-MOS GFS	GFS 2001-2005	GFS 2002-2005	10 EOFs (ζ_{850}, ω_{850}, PWat) of GFS (cross-validation)
DMO ECMWF	ECMWF 2001-2005	ECMWF 2001-2005	4 grid points of ECMWF precipitation (cross-validation)