Atmospheric Research 216 (2019) 1–10

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmosres

Nowcasting of the probability of accumulated precipitation based on the radar echo extrapolation

Lukáš Pop, Zbyněk Sokol*, Jana Minářová

Institute of Atmospheric Physics, Czech Academy of Sciences, Bocni II, 1401, 141 31 Prague, Czech Republic

ABSTRACT

The study presents a new method nowcasting precipitation called the Ensemble Tree Method (ETM), which gives probability forecast of accumulated precipitation based on the extrapolation of radar reflectivity. ETM combines a tree model with a Bootstrap technique. It forecasts the probability that the hourly accumulated precipitation exceeds a given threshold for cells of 3 by 3 km size.

ETM was tested using radar reflectivity data from July 2012 in a domain of 489 km by 291 km covering the Czech Republic (Central Europe). While forecasting, we considered a lead time of up to 180 min having a time step of 30 min and four precipitation thresholds (0.1, 1.0, 5.0, and 10.0 mm). ETM provided us forecasts of the probability of exceeding an hourly precipitation threshold from 0 to 60 min, 30 to 90 min, …, and 120 to 180 min. The performance of ETM was assessed using a skill score derived from the mean-square-error, and was compared with the performance of forecasts based on a logistic regression that was used as reference forecast. We demonstrated that the prediction of ETM is better than that of the reference forecast. The main advantage of ETM is that the ETM reflects the uncertainty of forecast better as compared to the overconfident reference forecasts, which is particularly true for the higher precipitation thresholds. Thus, despite low predicted probabilities, the forecasts given by ETM seem more realistic.

1. Introduction

Frequently used techniques for rainfall nowcasting consist in the extrapolation along Lagrangian trajectories of the measured radar reflectivity due to its low computational costs and easy implementation. Diverse extrapolation methods have been developed that differ in the calculation of motion fields and the application of extrapolation. The majority of the extrapolation methods is designed for a deterministic forecast (e.g. Germann and Zawadzki, 2002; Novák, 2007; Reyniers, 2008; Haiden et al., 2011; Sokol and Pesice, 2012; Sokol and Zacharov, 2012; Foresti et al., 2015; Bližňák et al., 2017). Nevertheless, several extrapolation methods have also been developed for a probabilistic forecast (e.g. Kitzmiller, 1996; Bowler et al., 2007; Berenguer et al., 2011; Atencia and Zawadzki, 2014; Atencia and Zawadzki, 2015; Sokol et al., 2017). The major advantage of a probabilistic forecast is that it enables to express the uncertainty of a forecast explicitly. The uncertainty of forecasts has two main reasons: (i) inaccurate calculation of Lagrangian trajectories and (ii) non-explicit modelling of the rainfall growth and decay. The uncertainties of forecasts are usually modelled using simple statistical models.

Often, the probabilistic extrapolation methods that forecast precipitation extrapolate the radar reflectivity data. Likely, the simplest probabilistic extrapolation method is the method that considers the forecasted values in the neighbourhood of a given point as possible forecasts; the possible forecasts then give an ensemble forecast (Schmid et al., 2000; Germann and Zawadzki, 2004; Theis et al., 2005). Kitzmiller (1996) developed a probabilistic precipitation forecast based on the Model Output Statistic approach (MOS; e.g. Sokol, 2003). MOS was used to statistically derive the regression equations that combine the rainfall probability with initial-time predictors (calculated from radar reflectivity data and selected NWP model variables), forecasted remote-sensor fields, and NWP model fields (Kitzmiller, 1996; Sokol and Pesice, 2012). Currently, the probabilistic forecasting methods almost exclusively proceed from ensemble predictions, tough they differ in a manner how the ensembles are formed. For instance, stochastic algorithms enable to simulate the stochastic perturbation and reproduce the spatial and temporal structure of precipitation fields (Bowler et al., 2007; Atencia and Zawadzki, 2014). Among other statistical methods, one can cite the String of Beads Model (SBM; Pegram and Clothier, 2001), which was used by Berenguer et al. (2011) in the ensemble nowcasting technique called SBMcast (String of Beads Model cast). The SBMcast enables one to generate the ensemble members of the rainfall forecast. Sokol et al. (2017) presented a different approach to form the ensembles. They calculated the probabilistic forecast in two steps: (i) they generated an ensemble of Lagrangian trajectories by using a covariance structure of the advection errors that were derived from historical data, and (ii) they estimated the error due to the neglect of both the growth and the decay of precipitation by dissociating the

* Corresponding author.
E-mail address: sokol@ufa.cas.cz (Z. Sokol).

https://doi.org/10.1016/j.atmosres.2018.09.019
Received 26 July 2018; Received in revised form 19 September 2018; Accepted 19 September 2018
Available online 21 September 2018
0169-8095/

Brier Score using the historical data. Several authors also applied the analogue-based approach, which consists in seeking the analogues (ensemble members), i.e. similar weather states to the current state, in the historical dataset (Panziera et al., 2011; Foresti et al., 2015; Atencia and Zawadzki, 2015). Studies based on analogues differ in the way how the analogues are defined because the main difficulty of the analogue-based approach is in searching for the suitable analogues. On the other hand, Atencia and Zawadzki (2015) concluded that any analogue-based probabilistic forecast has a better forecasting skill than the stochastic Lagrangian ensemble approach.

In this study, we present a new probabilistic extrapolation method for precipitation nowcasting that we call Ensemble Tree Method (ETM). ETM is based on the extrapolation of measured radar reflectivity data and an ensemble approach. A decision tree method is used to generate the ensemble members. In contrast to the majority of above-mentioned methods focused on probabilistic forecasting of precipitation, the presented ETM aims at a probabilistic forecasting of accumulated precipitation. The main reason for the probabilistic forecasting of the accumulated precipitation (instead of "simple" precipitation) is that we assume that for a given lead time the accumulated precipitation forecast is generally more successful as compared to [simple] precipitation forecast due to large spatial and temporal variability of precipitation, especially in summer. Moreover, even the users usually require the accumulated precipitation forecast instead of [simple] precipitation forecast.

After this introductory Section 1, Section 2 describes the used radar reflectivity data and the verification area situated in Central Europe. Section 3 provides an insight into the ETM, i.e. presented method/model for the probabilistic forecasting of accumulated precipitation. Section 4 describes the verification methods of ETM, whereas Section 5 discusses the results of the verification and provides a comparison of ETM with a model based on logistic regression. Several examples of precipitation forecasts by ETM are also given in Section 5. Section 6 summarizes the major findings of the new method/model for the probabilistic nowcasting of accumulated precipitation.

2. Radar reflectivity data and the verification area

In the study, we used radar reflectivity measurements from two Czech C-band radars (Novák, 2007) from May to September during 2009—2012. The two C-band radars are operated by the Czech Hydrometeorological Institute (CHMI). The radar reflectivity is measured in the two radar domains with a diameter of 256 km each (Fig. 1).

Two standard operational products of the CHMI are used:

(i) Radar reflectivity interpolated to a level of 2 km above sea level (*CAPPI 2 km*; Constant Altitude Plan Position Indicator),
(ii) Maximum reflectivity measured in the vertical column for each radar pixel (*MAX3D*).

Radar composites of both the CAPPI 2 km and the MAX3D are calculated using the maximum values of the two radars in the area where the two radars overlap. The resulting radar composites cover a domain of 728 km by 528 km (Fig. 1) and contain the data with a horizontal resolution of 1 km and a temporal resolution of 10 min.

The operational procedures of the CHMI also include a data quality control, which consists of ground clutter removal with a Doppler filter and the reduction of anomalous propagation artefacts (Novák et al., 2009). It also comprises the removal of artificial echoes such as WIFI interference in the field structures of reflectivity (Žejdlík and Novák, 2010). The radars are located in the highest parts of highlands, thus the blockage of radar echoes by terrain is insignificant in the radar domain and does not contaminate the quality of data.

Precipitation is derived from the radar measurements by a standard formula, which is operationally used in the CHMI:

$$Z = 200R^{1.6} \tag{1}$$

where R is the rain rate in mm/h and Z is the reflectivity in mm^6/mm^3. Precipitation is calculated using CAPPI 2 km.

Although we are aware that the radar derived precipitation is less accurate than the precipitation obtained by merging of the radar data with rain gauge measurements, we use radar data only since the CHMI does not use the merged data for precipitation nowcasting and we aim at precipitation nowcasting and its verification. Moreover, there are two main reasons for using the radar data only. First reason is that the rain gauge measurements are available later than the radar data in the CHMI, therefore the use of rain gauge measurements would slow down the operational forecast, which is inconvenient for nowcasting and thus for our study. Second reason is that including rain gauge measurements in only the verification of the extrapolation forecasting method might significantly affect the results. Thus in our study, both the forecast and the verification are based on radar reflectivity data only. Resulting forecasts are evaluated in a verification area (489 km by 291 km), which covers the Czech Republic (Fig. 1).

The radar reflectivity data are averaged into an area covering 3 × 3 pixels (i.e. 3 × 3 km) because our experience showed that a horizontal resolution of 1 km is too high to obtain satisfactory results for convective precipitation forecasts. Moreover, as desired in this study, the averaging to a lower horizontal resolution (3 × 3 km) leads to a suppression of individual local extremes that are usually incorrect, and it smooths the precipitation field. Sokol et al. (2017) provided a detailed description of the reasons for averaging of the radar reflectivity data.

3. Ensemble Tree Method (ETM) for forecasting the accumulated precipitation

Ensemble Tree Method (ETM), which we propose, forecasts the probability that the precipitation accumulated from a time $T_0 + T$ to the time $T_0 + T + 1$ h exceeds a given precipitation threshold T_r in cells covering 3 km by 3 km (i.e. 3 by 3 pixels) in the verification area (Fig. 1). T_0 is the time of the last known radar reflectivity measurement (i.e. beginning of a forecast) and T represents the lead time. We considered T varying from 0 to 2 h with a time step of 30 min, and following precipitation thresholds T_r = 0.1, 1, 5, 10, 15, and 20 mm. We calculated the accumulated precipitation using the trapezoidal rule on radar derived precipitation with a time step of 10 min.

The ETM for forecasting the accumulated precipitation comprises of two steps: (i) computation of two predictors (predictor R_{as} and SV; Section 3.1), and (ii) run of a probabilistic forecasting model using the calculated predictors (Section 3.2).

3.1. Computation of two predictors

First, we calculate motion fields from a time T_0 and the time T_0–10 min using MAX3D data (Section 2) and an algorithm similar to COTREC, i.e. the Continuity of TREC (Tracking Radar Echo by Correlation) vectors (e.g., Novák et al., 2009), that has been detailed by Sokol et al. (2017). Instead of CAPPI 2 km, which we use to derive immediate rain rates (Eq. 1), we use MAX3D data at full horizontal resolution (i.e. 1 km) to derive the motion fields, because the MAX3D data have a more pronounced structure as compared to that of CAPPI 2 km, which is suitable for the algorithm of Sokol et al. (2017).

Subsequently, we accumulate precipitation from T_0–60 min to T_0 and we extrapolate it over a lead time T. The precipitation is accumulated and extrapolated based on the backward-in-time Lagrangian trajectories (e.g., Germann and Zawadzki, 2002). This way, we computed the extrapolated hourly precipitation R_a for the whole dataset covering the warm period (May—September) during 2009—2012.

3.1.1. Predictor R_{as}

In the next step, we smooth the R_a to obtain the value of a predictor