Nationwide Radar-Based Precipitation Nowcasting—A Localization Filtering Approach and its Application for Germany

Ricardo Reinoso-Rondinel [ORCID], Martin Rempel, Markus Schultze, and Silke Trömel

Abstract—One approach to improve advection methods is the short-term ensemble prediction system (STEPS). STEPS decomposes precipitation fields into different spatial scales and filters those having a short lifetime. The latter is achieved by using an auto-regressive (AR) model that considers a sequence of observations. However, such a model tends to smooth nowcasting fields especially in small but convective precipitation areas and at longer lead-times. With focus on the deterministic configuration of STEPS, i.e., the spectral prognosis model (SPROG), this article 1) extends the STEPS approach by estimating spatially localized parameters of the AR process, 2) conducts a sensitivity analysis of the SPROG model to the order of the AR process, the spatial decomposition levels, and post-processing, and 3) analyzes the forecast skill of the extended STEPS. For such purpose, the performance of the localized AR model was demonstrated and evaluated at several precipitation thresholds and window sizes using a varied set of precipitation events collected by the radar network of the German Weather Service. The statistical results exhibited an improved performance of the localized AR model over SPROG when both are evaluated at precipitation thresholds and window sizes larger than $0.1 \, \mathrm{mm \, h^{-1}}$ and 1 km, respectively, and for lead-times up to 2 h. The analysis suggested a first-order AR process, six cascade levels, and a mean adjustment post-processing procedure. Our results show a key role of the localization aspect when generating nationwide forecasts in scenarios that include large precipitation areas which are non-uniformly distributed having isolated convective features.

Index Terms—Precipitation modeling, prediction, quantitative precipitation nowcasting, weather radar.

I. INTRODUCTION

PROLONGED heavy precipitation induced by low pressure systems or short intense rainfall due to deep convection

Ricardo Reinoso-Rondinel is with the Institute for Geosciences, Department of Meteorology, University of Bonn, 53113 Bonn, Germany (e-mail: ricardoreinoso@uni-bonn.de).

Martin Rempel and Markus Schultze are with the Deutscher Wetterdienst, 63067 Offenbach, Germany (e-mail: martin.rempel@dwd.de; markus.schultze@dwd.de).

Silke Trömel is with the Institute for Geosciences, Department of Meteorology, University of Bonn, 53113 Bonn, Germany, and also with the Laboratory for Clouds and Precipitation Exploration, Geoverbund ABC/J, 53113 Bonn, Germany (e-mail: silke.troemel@uni-bonn.de).

are typical severe weather phenomena in Europe [1]. They represent one of the main triggers for water-related hazards such as flash floods or debris flows [2]. For example, a two-week period in May–June 2016 with widespread airmass convection in Central Europe led to an economic loss of EUR 2.6 billion in Germany alone [3]. These examples of precipitation illustrate the variability over a wide range of temporal and spatial scales, making it challenging to forecast precipitation accurately [4].

An important tool to forecast such phenomena is an observation-based precipitation nowcasting technique. The World Meteorological Organization (WMO) [5] defines this technique as a detailed description of the current state of the atmosphere in combination with forecasts based on extrapolation 6 h ahead. Despite the immense progress made in improving the quality of forecasts made by numerical weather prediction (NWP) over the last decades [6], forecast errors of the above described small-scale phenomena are quite high compared to extrapolation-based nowcasting methods [7]. This may be caused by inaccurate initial and outdated boundary conditions or effects that are not well captured by numerical models in particular with respect to parameterized processes (e.g., cloud microphysics; [8]). Consequently, observation-based precipitation nowcasting techniques typically outperform the NWP up to lead-times of the order of 4–6 h (e.g., [9]), especially when considering the low update frequency compared to extrapolation-based nowcasting. Deterministic radar-based precipitation nowcasting techniques are well-established and one essential tool for short-term prediction of precipitation. Precipitation nowcasts, for instance [10], are conducted by extrapolating the most recent observed precipitation field along the motion field estimated from optical-flow methods (e.g., [10]–[12]) up to a lead-time of approximately 3 h. One benefit of using radar-based precipitation nowcasts is the direct applicability as high-quality input for hydrological models [13] at least for short lead-times. The aim of such flood forecasting and derived warning services is to provide timely and accurate warnings in order that adequate response can be taken to mitigate the impact of flooding [13]. Therefore, the meteorological and hydrological communities require input from accurate short-term prediction of precipitation for applications on rainfall-runoff simulation [14]–[18]. One common feature of traditional precipitation nowcasting is the forecast in accordance to the concept of Lagrangian Persistence [10]. This approach describes a pure advective forecast that does not

take into account the growth and decay of precipitation. During the last two decades, a variety of ideas have been developed to overcome this limitation. Germann and Zawadzki [10] proposed the incorporation of a source/sink term to consider growth and decay of precipitation during the forecast based on the determined upstream intensity change of the precipitation field. The authors emphasize that the forecast is improved, if a spatial and temporal persistence of intensity changes is prevalent. This persistence mainly occurs in situations with synoptic scale precipitation. Atencia *et al.* [19] statistically analyze the forecast errors of the McGill Algorithm for Precipitation Nowcasting based on Lagrangian Extrapolation (MAPLE; [10]) in relation to the time of day. These effects of the diurnal cycle of precipitation were used for an adaptive correction of the nowcasts, which significantly reduced the forecast error.

In recent years, multidata approaches making use of machine learning techniques came up in the field of precipitation nowcasting. The systematic growth and decay of precipitation by orographic forcing in the Alpine region has been studied by [20] using a 10-year archive of radar data. In combination with time of day and airmass information from NWP, these data were used to train an artificial neural network (ANN). Hamann *et al.* [21] have utilized multisensor data (radar, lightning, satellite, NWP) to train a XGBoost model (gradient boosted trees) eventually containing several hundred of input parameters to forecast the evolution of cell severity. Probabilistic precipitation nowcasts are carried out for the continental United States using radar data and satellite data of GOES-16 by [22] applying deep neural networks. This approach is able to outperform traditional precipitation nowcasting and NWP up to 8 h lead time. Schaumann *et al.* [23] proposed an ANN that is able to produce consistent and calibrated probabilistic forecasts for multiple thresholds of rain rates by combining NWP data with precipitation nowcasts.

One typical drawback of statistical methods such as machine learning systems is the necessity of sufficient and homogeneous datasets for training purposes being delicate in case of radar data, for instance (e.g., [24], [25]). Another limitation of such systems is that accurate predictions are not expected if similar conditions were not given during the training process. Accordingly, the trained algorithms frequently tend toward the statistical mean value, also in case of rare events (i.e., high intensities) [26]. Thus, the added value compared to traditional methods commonly vanishes for higher thresholds [27]. Due to its "black box" character, the results of machine learning algorithms may be difficult to understand for a developer, but even more for end users (e.g., operational forecasters). In addition, generating realistic patterns in precipitation remains a major challenge, but the demand for using the data as input for hydrological models continues.

Another well-established precipitation nowcasting approach is the Short-Term Ensemble Prediction System (STEPS; [28]). It comprises the scale dependency of the predictability of precipitation, which has been shown in numerous studies [29]–[32]. STEPS is a stochastic extension of the nowcasting method Spectral Prognosis referred to as SPROG [30]. The latter aims to filter an observed precipitation field on small spatial scales,

expecting that precipitation on these scales has a short life time and thus a shorter predictability compared to larger spatial scales. The spatial decomposition of the precipitation field into a number of cascade levels is done through a filtering decomposition process, whereas the spatial filtering, at each cascade level, is performed during extrapolation using a scale-dependent autoregressive (AR) model. The parameters for this AR model are obtained from the latest radar observations.

In STEPS, the basis of the stochastic part is a Fourier-filtered white noise field that reproduces the spatial properties of the latest observation and represents the forecast uncertainty. Next, STEPS generates an ensemble nowcast, where each realization provides a realistic precipitation pattern. Common postprocessing tools, such as ensemble calibration or derivation of exceedance probabilities, can be applied to the ensemble. One reason why the supposedly outdated STEPS idea prevails against more advanced precipitation nowcasting schemes is the easy comprehensibility from a meteorologist's point of view. Over the years, the STEPS approach has been adapted for several regions with individual configurations and extensions (e.g., [33]–[35]), showing potential and suitability for operational applications. Recently, the pySTEPS[1] [36] initiative was established to develop and maintain an open-source python framework for short-term deterministic and probabilistic precipitation nowcasting techniques. For example, a recent work [37] developed a deterministic precipitation nowcasting model with focus on convective rainfall making use of volumetric radar observations and an autoregressive integrated (ARI) process to mitigate smoothing effects in the nowcasting caused by the abovementioned filtering of small spatial scales. Inspired by previous studies, the German Weather Service (Deutscher Wetterdienst; DWD) intends to apply the STEPS approach as a basis to seamlessly combine short-range NWP forecasts with radar-based extrapolation nowcasts in an operational environment; the Seamless Integrated Forecasting System.

Several preassumptions are needed to run the pySTEPS system. Assumptions that can be difficult to meet and model in complex space-time structures, for example, a homogeneity of the rain field. Additionally, the default model parameters (such as the number of decomposition levels and the order of the AR model) and postprocessing methods such as probability matching given by pySTEPS need to be investigated for its specific application. This work aims to 1) study the sensitivity of the STEPS approach in its deterministic configuration, i.e., the SPROG model, to the order of the AR process and the number of decomposition levels, 2) adapt the estimation of the AR parameters by considering a localization method to control the temporal evolution of precipitation while addressing the inherent homogeneity assumption, and 3) evaluate the precipitation nowcasting fields at various intensity and spatial precipitation thresholds as well as to show the dependency of the nowcasting models to the motion field vectors and postprocessing. For that, a set of configurations are tested and the results are analyzed using 10 events collected by the DWD radar network in order to suggest a suitable configuration for the generation of deterministic and

---

[1][Online]. Available: https://pysteps.github.io/

ensemble precipitation nowcasting in Germany under real-time operation constraints.

The rest of this article is organized as follows. Section II gives a background on advection-based precipitation nowcasting and describes the components of the SPROG nowcasting model [30]. Additionally, it presents our methodology to extend SPROG by means of a localization approach of the AR parameters, considering the spatial variability of observed precipitation. Section III introduces the dataset, configuration settings, and the evaluation metrics. Section IV performs a sensitivity analysis of SPROG to initiation parameters and postprocessing methods. Section V quantifies the performance and conducts a statistical evaluation of the extended SPROG-localized approach. Finally, Section VI concludes this article.

II. METHODOLOGY

A. Advection Nowcasting

We follow the advection approach given by [10], which extrapolates the last observed precipitation field R (mm h^{-1}) via estimated advection vectors, providing a sequence of nowcasting fields. The extrapolation is conducted in a semi-Lagrangian backward scheme. This scheme implies that for each grid box within the extrapolated nowcasting field the value of the origin grid box in the latest observation is advected along the trajectories of the vector field. The latter, however, remains stationary during the lead time. One approach for the estimation of advection vectors is given by the dynamic and adaptive radar tracking of storms DARTS introduced by [38]. DARTS is an area-wide approach in which the optical-flow equations are formulated and solved in the spectral domain, using the 3-D discrete Fourier transformation (3D-DFT). Three user-defined numbers, N_x, N_y, and N_t, are needed to truncate the number of DFT coefficients of the consecutive R fields in the horizontal, vertical, and temporal sequence axes, respectively. Two more parameters M_x and M_y determine the truncation of the DFT coefficients, which are estimated from the linear least square method, to solve the optical-flow equations and obtain the advection vector fields. Thus, the spatial detail of the horizontal and vertical components of the 2-D advection fields is controlled by M_x and M_y.

A second and distinct approach is the duality-based total variation L^1-norm (Dual TV-L^1) optical flow algorithm [39], [40], which is used for operational purposes at DWD (e.g., a deterministic radar extrapolation nowcasting). The algorithm solves the optical flow equations in spatial space by minimization of an energy functional containing a regularization and a L^1 attachment term. The regularization term describes herein the sum of the total variation of an initial displacement vector field. Due to a pyramidal scheme, the minimization takes place on different spatial scales so that the initial vector field is updated from coarse to fine scales. The smoothness of the results as well as the link between both terms can be controlled by an attachment and tightness parameter, respectively. The pyramidal scheme can be configured by a down-sampling factor and the number of scales. The basic idea of the Dual TV-L^1 algorithm is closely linked to the variational echo tracking method applied in MAPLE [10].

Although only one optical-flow technique will be sufficient to perform the analysis of distinct nowcasting methodologies, we will use the DARTS and Dual TV-L^1 techniques, mainly because the latter has been used operationally at DWD for nowcasting purposes but without a formal documentation of its performance, while the performance of the former one had been presented in related pySTEPS articles (e.g., [36]), allowing its use for reference purposes in this study.

B. Spatial-Decomposition Nowcasting Models

For the SPROG nowcasting, first, a sequence of observed precipitation fields are brought to Lagrangian-coordinates using, for instance, either of the advection schemes described in Section II-A. Further, these precipitation fields are decomposed in a number of cascade levels ($K > 1$). For the spatial decomposition and to handle nonprecipitation areas, Seed [30] suggested to express R in logarithm scale R_T (dB) such that

$$R_T = \begin{cases} 10 \log_{10} R & \text{for } R > R_{th} \\ R_{min} & \text{otherwise} \end{cases} \quad (1)$$

where R_{th} and R_{min}, such that $R_{min} < R_{th}$, are user-defined parameters.

The decomposition of a square-shaped R_T field is done by multiplying its Fourier spectrum by K Gaussian-shaped band-pass filters each centered in the frequency interval $[L^{-1}q^{k-2}, L^{-1}q^k]\Delta x^{-1}$, where L (grid boxes) is the side length of R_T, Δx (km) is the resolution of R_T, q is the ratio of scales between decomposition levels ($q > 1$; e.g., 1.62 for $K = 12$), and $k = 1, 2, \ldots K$ [28], [32]. In this way, R_T is expressed as

$$R_T = \sum_{k=1}^{K} \mu_k + \sigma_k R_{T,k} \quad (2)$$

where $R_{T,k}$ represents the kth level normalized by the mean μ_k and standard deviation σ_k. Both scaling parameters remain constant during extrapolation. An AR model of order $p \geq 1$, referred to as AR(p), is used to model the evolution of the decomposed precipitation field at each cascade level $R_{T,k}$ during extrapolation as

$$R_{T,k}(t + t_l) = \sum_{n=1}^{p} \phi_{k,n}(t) R_{T,k}(t + t_l - n\Delta t) + \phi_{k,0}(t)\epsilon_{c,k}(t + t_l). \quad (3)$$

In this equation, t indicates the time of the last observed rainfall field, t_l represents the lead time, and Δt the time between two consecutive observations of rainfall fields such that $t_l = m\Delta t$ and $m \geq 1$ is an integer number. The AR parameter $\phi_{k,n}(t)$ controls the rate of evolution of each cascade level and it is constant during extrapolation. Each AR parameter is obtained from the time lagged autocorrelation coefficients $\rho_{k,n}$ resulting from the last $p + 1$ advected frames of $R_{T,k}$ to time t by solving the Yule–Walker equations [28]. The term $\epsilon_{c,k}(t + t_l)$ represents the stochastic noise term of the AR process and is neglected in the SPROG model. However, in the STEPS model [41], $\epsilon_{c,k}(t + t_l)$ is included to simulate the uncertainties related to the growth and

decay process of precipitation at time $t + t_l$. It can be generated by multiplying the spectrum of the rainfall field and a white noise field, then transforming the product back to the spatial domain, and decomposing it into K levels. Afterward, the decomposed noise term is scaled at each level k by the parameter $\phi_{k,0}$, also known as the innovation variance. The latter describes the uncertainty not accounted for the serial correlation [42] and it is given by

$$\phi_{k,0} = \sqrt{1 - \sum_{n=1}^{p} \rho_{k,n} \phi_{k,n}}. \tag{4}$$

Last, the nowcasting field $R(t + t_l)$ is obtained by applying the following operations to $R_{\mathrm{T},k,e}(t + t_l)$: recomposition of the K cascade levels through μ_k and σ_k, postprocessing of the integrated field as indicated in Section II-D, advection of the postprocessed field to time t_l, and unit transformation to mm h^{-1}. In the STEPS model, the generation of the noise term and the described operations are iterated multiple times to generate multiple members resulting in an ensemble nowcast.

A recent approach [37] aims to improve the SPROG model in convective precipitation using, besides gridded rainfall field inputs, vertically integrated liquid estimated from volumetric radar observations and an ARI process. The ARI model allows to simulate the growth and decay process using the differenced time series of rainfall fields representing a source/sink term as explained in [10]. Afterward, the AR equation [cf. (3)] is utilized but for the differenced time series. The resulting extrapolated tendencies are added to the original field. Similar to SPROG, the parameters of the ARI model are estimated at each spatial scale. The estimation of these parameters are done also localized at the grid point level.

C. SPROG(STEPS)-Localized Nowcasting

The proposed nowcasting model, hereafter SPROG-LOC, considers the spatio-temporal heterogeneity of precipitation systems that is expected in nationwide radar composites of observed precipitation. Therefore, this model is expected to improve the representation of higher rain rates throughout the extrapolation forecast while preserving long-living precipitation structures such as mesoscale convective systems. Thus, the essence of SPROG-LOC is a 2-D localization filtering process that is applied during the estimation of the autocorrelation coefficients and, consequently, 2-D AR parameters $\mathbf{\Phi}_{k,n}$ and $\mathbf{\Phi}_{k,0}$ are obtained as well as the scaling parameters μ_k and σ_k. Subsequently, only the estimation of the correlation coefficients is described as the scaling parameters are obtained in a similar manner.

The basis for the localization approach is a scale-dependent convolution square kernel $\mathbf{Q}$. For each spatial scale, the length extent l_k of the square kernel is defined as

$$l_k = \max(l_{\min}, \mathbf{k}_{\min} q^{K-k} h). \tag{5}$$

Here, $l_{\min}$ and h are user-defined values representing a minimum kernel length and a multiplicative factor, respectively. These values ensure on one hand that at least one half of an oscillation at the respective scale is covered by the kernel and, on the other hand, that still a sufficient large surrounding around each

point is used to reduce noise in the resulting fields of correlation coefficients. $\mathbf{k}_{\min} q^{K-k}$ corresponds to the central wavelength of the respective cascade level, whereas herein the user-defined $\mathbf{k}_{\min}$ depicts the scale of the second cascade levels and controls the multiplier between the cascade levels. In the case of a low number of spatial scales, $\mathbf{k}_{\min}$ maintains a reasonable symmetric aspect of the K band-pass filters and separation between the first two Gaussian filters. Further details on $\mathbf{k}_{\min}$ are given in [33]. If the h-fold of the respective central wavelength exceeds the side length of the domain, correlation coefficients are obtained globally as in SPROG using the entire field.

In a further step, the square kernel determined in this manner is normalized such that the sum of its elements equals 1 and is used to estimate the spatially and localized time-lagged correlation coefficients. For a cascade level k, a 2-D field of localized mean values of an arbitrary decomposed precipitation field R_T is obtained by applying a convolution with the square kernel $\mathbf{Q}_k$ so that

$$\langle R_{\mathrm{T},k} \rangle_{\mathrm{loc}} = R_{\mathrm{T},k} * \mathbf{Q}_k. \tag{6}$$

For a time-efficient computation, the convolution takes place in Fourier space. Next, standard deviations and covariances are computed for each cascade level of the last $p + 1$ frames beginning at the time of the latest observation t by (see e.g., [43], [44])

$$\sigma_{\mathrm{loc}}(R_{\mathrm{T},k}(t)) = \sqrt{\langle R_{\mathrm{T},k}(t)^2 \rangle_{\mathrm{loc}} - \langle R_{\mathrm{T},k}(t) \rangle_{\mathrm{loc}}^2} \tag{7}$$

and

$$\sigma_{\mathrm{loc}}(R_{\mathrm{T},k}(t), R_{\mathrm{T},k}(t - n\Delta t)) = \langle R_{\mathrm{T},k}(t) R_{\mathrm{T},k}(t - n\Delta t) \rangle$$
$$- \langle R_{\mathrm{T},k}(t) \rangle \langle R_{\mathrm{T},k}(t - n\Delta t) \rangle, \text{ with } n = 1, 2, \ldots p. \tag{8}$$

Finally, the 2-D correlation coefficients are obtained by

$$\mathbf{P}_{k,n} = \frac{\sigma_{\mathrm{loc}}(R_{\mathrm{T},k}(t), R_{\mathrm{T},k}(t - n\Delta t)}{\sigma_{\mathrm{loc}}(R_{\mathrm{T},k}(t))\sigma_{\mathrm{loc}}(R_{\mathrm{T},k}(t - n\Delta t))}, \tag{9}$$

with $n = 1, 2, \ldots p$ and $k = 1, 2, \ldots K$. The estimation of correlation coefficients may be affected in several ways by the motion vector field that is used. First, when using an optical-flow technique like the Dual TV-L^1 algorithm, the temporal evolution of precipitation in a sequence of observations is equal to an intensity change leading to a violation of the optical flow constraint. Second, a smoothing of the resulting motion vector field to obtain an advection velocity also in areas without precipitation leads to an overall reduction in amplitude. Both aspects result in a lower advection velocity. A spatial shift of oscillations in the spatial scales of the decomposed precipitation field caused by the reduction in advection velocity may lead to lower correlation coefficients, especially when they are estimated by (9). To diminish this effect, a scale-dependent logarithmically increasing optimization factor f_k is introduced such that $\tilde{\mathbf{P}}_{k,n,x,y} = f_k \mathbf{P}_{k,n,x,y}$. The factor is given by

$$f_k = (1 + I)^{\frac{k-1}{K-1}}. \tag{10}$$

Here, I is the desired maximum fractional increment given by a user.

Afterward, the 2-D parameter fields $\mathbf{\Phi}_{k,n}$ and $\mathbf{\Phi}_{k,0}$ are estimated from the optimized correlation coefficients in the same way as indicated in Section II-B. Although STEPS uses the standard FFT for the estimation of the spatial correlated noise field, we will use the short-space fourier transform (SSFT) on the latest observation as described in [34] for the generation of the stochastic term in both STEPS and STEPS-LOC.

Finally, the modified AR(p) process is applied to the normalized $R_{\mathrm{T},k}$ fields. Due to this localization approach, the 2-D coefficients of the AR(p) process is expected to overcome the spatial uniform filtering rate within each extrapolated cascade level. This has the advantage that different synoptic situations of the same spatial scale within the observed domain (e.g., isolated as well as embedded convection) are considered more precisely in terms of their individual decorrelation behavior. Moreover, the localized AR(p) process leads straightforward to a localized 2-D parameter field given by (4) that scales in STEPS-LOC the stochastic term and thereby the noise field.

D. Postprocessing Nowcasting Fields

A postprocessing step is often used at each time step to ensure that the predicted field $R_T(t + t_l)$, after recomposition of the K levels, maintains statistics similar to those from $R_T(t)$ such as the wet-area ratio and intensity distribution. The motivation here is to reduce systematic errors associated with unit transformation, spatial decomposition, and the AR modeling that could propagate through the nowcasting fields. Seed [30] suggested a mean-based adjustment to $R_T(t + t_l)$ in areas conditioned to keep the estimated wet-area ratio, hereafter μ-adjustment. This adjustment is done as $R_T(t + t_l) - \mu(t + t_l) + \mu(t)$, where $\mu(t + t_l)$ and $\mu(t)$ represent the mean of the predicted and the observed field, respectively. A further postprocessing method is to match the cumulative distribution function of $R_T(t + t_l)$ to that of $R_T(t)$, hereafter CDF-matching, proposed by [45].

III. DATA DESCRIPTION AND MODEL CONFIGURATION

A. Data Events

We utilize data from the DWD C-band radar network that currently consists of 17 dual-polarimetric radars that collect data, in azimuth and elevation, every 5 min with a range resolution of 250 m. The maximum range of each radar is 150 km for the terrain-following precipitation scan and 180 km for the volume scan mode. The obtained data provide the basis for the operational system for quantitative precipitation estimation (QPE) at DWD, called RADOLAN [46]. This system provides a set of several rain gauge-adjusted and real time products.

For this study, we use the RADOLAN real time product RY [47]. The product covers a gridded area of 900×900 km with a horizontal resolution of $\Delta x = 1$ km and a temporal resolution of 5 min. For the RY product, reflectivity values Z from each radar site are corrected for clutter contamination, path attenuation, and beam blockage [48]. Then, these corrected Z values are transformed into rain rates using a refined Z–R-relation [46]. The parameters of this relation are different for three reflectivity categories (low, medium, and high) and additionally, depend on a reflectivity structure metric. Afterward, the obtained rain rates

are mapped from the original polar grid onto the abovementioned 1 km polar stereo-graphic grid. A more detailed description of the composite steps and the performance of the RY product is given by [49]. As the focus of this study is on the sensitivity of individual SPROG components with regard to the heterogeneity of precipitation pattern sizes, we have chosen 10 precipitation events for the years 2012–2017 as described in Table I. For the sensitivity analysis, we compute within a uniform time period of 06:00–20:30 UTC half-hourly forecasts with a lead time range of 3 h for each configuration of SPROG.

To outline the development of precipitation within the considered domain and time period, Fig. 1 illustrates the diurnal cycle of normalized precipitation coverage. The latter represents the precipitation coverage normalized by its daily maximum. The resulting curves are additionally smoothed by a Gaussian filter to reduce short-term peaks and to emphasize the shape of the diurnal cycle. The investigated events were grouped by the time at which the diurnal cycle reached its maximum peak: 1) events E1-E4 and E9 show a peak in the evening around 19:00 UTC, 2) E5–E8 peak in the afternoon around 15:00 UTC, and 3) E10, which is a stratiform event, does not display a distinct peak. Four events are illustrated in Fig. 2. A prolonged linearly organized convective system that crossed Germany from southwest to northeast, E1 on June 21, 2012, is depicted in Fig. 2(a). Fig. 2(b) shows event E4 on May 29, 2016 by a large-scale convective-enhanced precipitation induced by a lee cyclone, i.e., orographic related, over South Germany with a further developing thunderstorm line eastward. Additionally, a few single convective cells are situated over the northern part of Germany. The event E7 on June 18, 2016 in Fig. 2(c) exhibits typical postfrontal scattered showers. The event E10 on 25 July 2017 is shown in Fig. 2(d) and it reveals a large-scale stratiform precipitation over Germany induced by a cold air pool. The interpretation of the advection vectors, also shown in Fig. 2, can be found at.[2]

B. Configuration Parameters

A summary of the main configuration parameters needed to execute SPROG(STEPS), SPROG(STEPS)-LOC, and advection+ARI are shown in Table II.

The optical-flow DARTS approach was applied to nine consecutive frames, $N_s = 9$, of precipitation fields R_{T} using the following parameters that truncate the Fourier coefficients: $N_x = 50$, $N_y = 50$, $N_t = 4$, and those related to the spatial detail $M_x = 2$ and $M_y = 2$. We tested a combination of different settings, for instance, M_x and M_y equal to eight while varying N_x and N_y within 10 and 100 and N_t within two and eight, however without a significant trend on the evaluation skills of predicted fields. The setting indicated above is equal to the default configuration given by the open-source pySTEPS motion module.

Motion vectors estimated by the Dual TV-L^1 technique are obtained in a time period of 30 min between all available consecutive frames of precipitation fields R_T. The attachment parameter λ determining the smoothness of the

---

[2][Online]. Available: https://weather.gov/hfo/windbarbinfo

TABLE I
LIST OF 10 PRECIPITATION EVENTS IN GERMANY OBSERVED BY THE DWD RADAR NETWORK DURING 2012–2017

Event	Date	Description
E1	21 Jun 2012	Linearly organized convective system crossing Germany from southwest towards northeast
E2	20 Jun 2013	Non-organized system moving north but associated with a flooding of the Mehlemer Bach MB catchment in Bonn
E3	06 Aug 2013	Small cells moving eastward followed by a convective front that extends along the north-south axis
E4	29 May 2016	Large scale convective-enhanced precipitation over south Germany induced by a lee cyclo-genesis with a further developing thunderstorm line eastward
E5	01 Jun 2016	Isolated and multi-cell regions moving southeast
E6	04 Jun 2016	Scattered precipitation in the south area followed by several isolated cells moving west which led to flooding of the MB catchment
E7	18 Jun 2016	Typical post-frontal widespread scattered showers
E8	25 Jun 2016	Convective line along the north-south axis moving north with a further scenario of scatter cells in the north-east side and then of an organized line in the south side moving northeast
E9	19 Jul 2017	Convective front with a long north-south extend moving eastward
E10	25 Jul 2017	Large scale stratiform precipitation induced by a cold air pool

Each event consists of observations between 06:00 and 24:00 UTC.
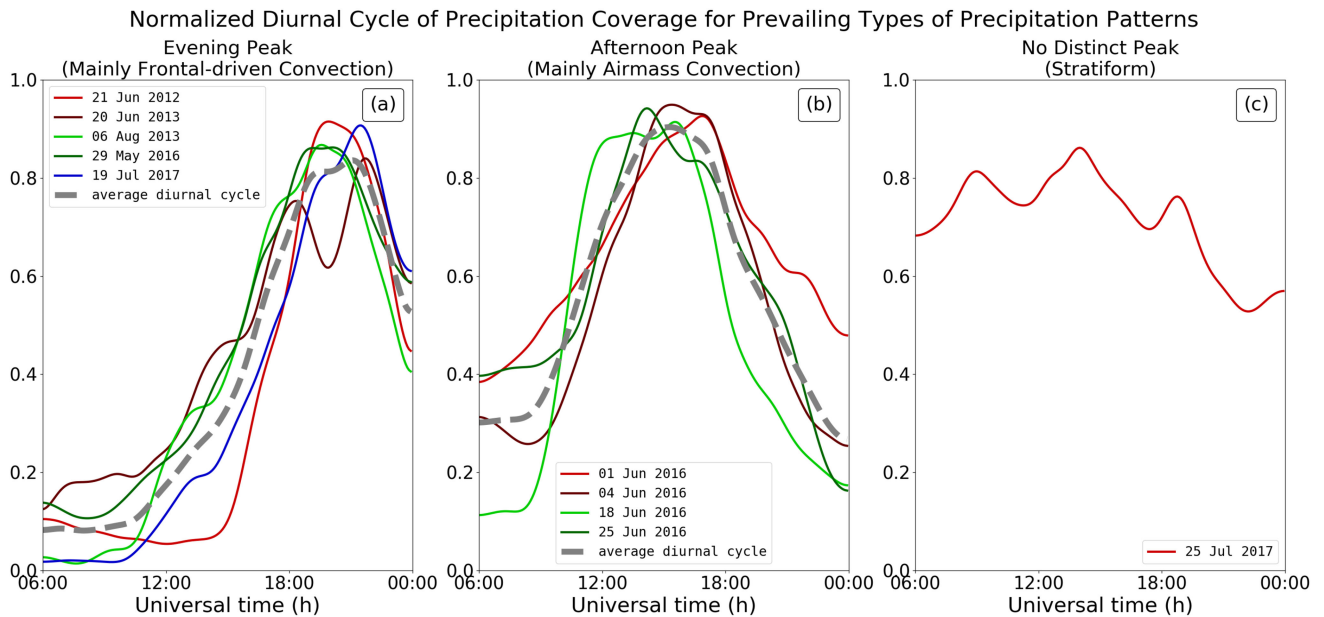

Fig. 1. Normalized diurnal cycle of the 10 events listed in Table I. (a) The maximum diurnal cycles are seen in the evening hours while in (b) they are seen during the afternoon hours. (c) A distinctive maximum is not observed. The dashed grey line shows the average of the diurnal cycle among the events if there are more than one.

output is set for our purposes to 0.02. The other parameters controlling this technique are set to values suggested by [40].

The advection method is configured to generate nowcasting fields with a temporal resolution of $\Delta t = 5$ min for a lead-time period t_l up to 3 h while maintaining the spatial grid Δx to 1 km. To represent a given precipitation field R in logarithmic scale R_T, R_{th} in (1) is set to 0.1 mm h^{-1}. To reduce the computational time of the Fourier transformations, the resulting R_T field is extended to the next power of 2, i.e., $L = 1024$ grid boxes, and it is filled by values of $R_{min} = -15$ dB, equivalent to 0.03 mm h^{-1}. The resulting difference, less than 0.1 mm h^{-1}, between the set values for R_{th} and R_{min} seems reasonable to avoid an unnecessary increase of power at high spatial frequencies, which can impact the AR process [34]. For the

next steps related to the AR process and extrapolation, however, R_T will be processed at its original size. To study the sensitivity to model parameters of SPROG, the AR(p) order will be set by AR(1) and AR(2) while the number of scale levels K are set at 3, 6, and 12. Such configurations to execute SPROG will be indicated as SPROG(p, K) in Sections IV and V. The additional parameters related to SPROG-LOC, considering our spatial domain size and grid resolution, are given by $k_{min} = 2$, $h = 3$, and $l_{min} = 100$ km for the determination of the kernel size while for the adjustment of the AR parameters, two values of I are tested: $I = 10\%$ and 20%. For implementation purposes of the spatial decomposition, the normalized coefficients of the kth filter, centered at the corresponding Fourier wavenumber $|\mathbf{k_w}(k)| = |\mathbf{k_w}(2)|q^{k-2}; k \geq 2$, were obtained from the pySTEPS decomposition module by giving $|\mathbf{k_w}(2)|$, i.e., the
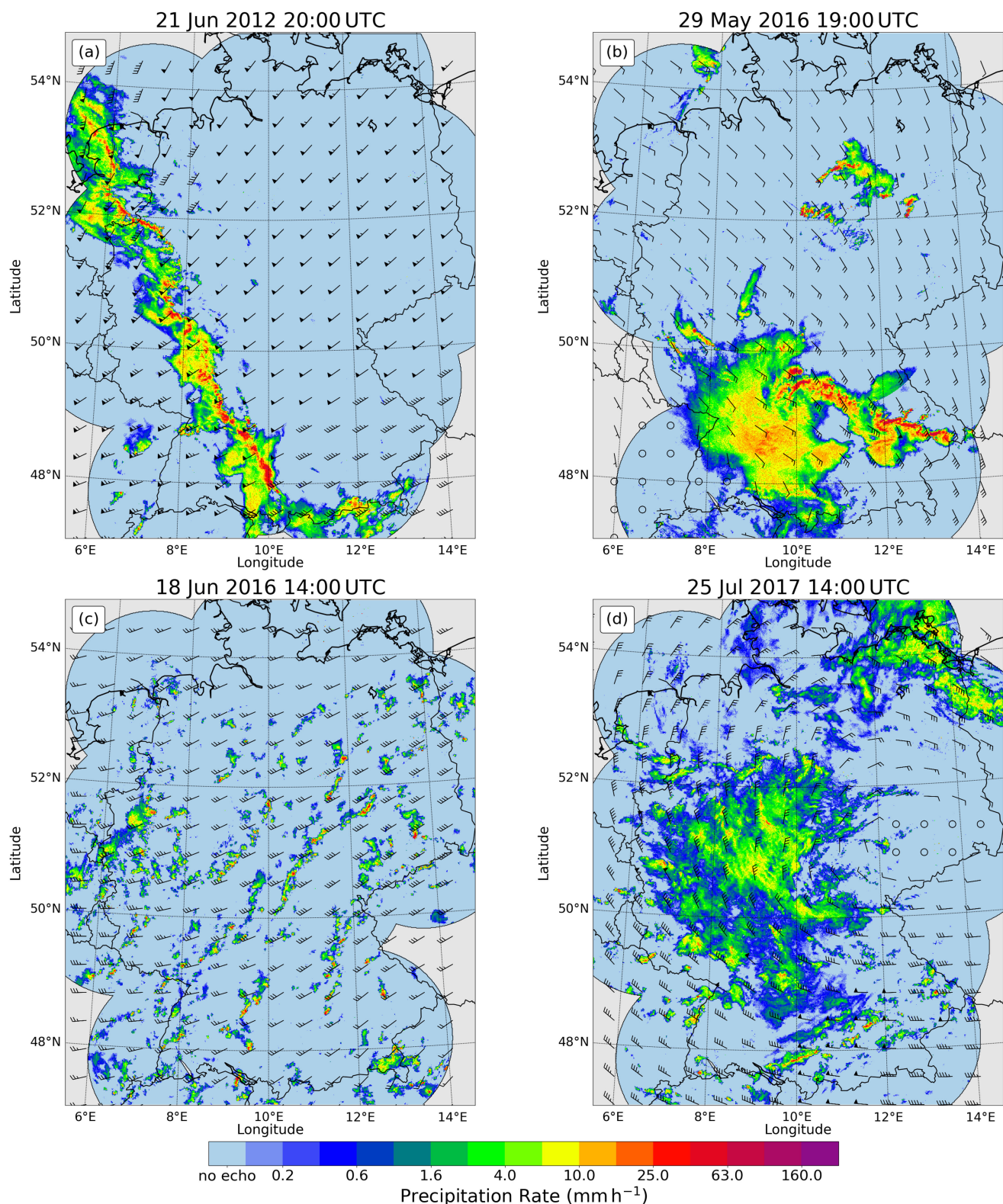
Fig. 2. Four events showing the rainfall rate RY product given by the DWD's operational radar network. (a)–(c) Displays three convective events E1, E4, and E7 at 20:00, 19:00, and 14:00 UTC while (d) shows the stratiform event E10 at 14:00 UTC. In addition, the motion vectors obtained by the Dual TV-L^1 technique are indicated by the conventional wind speed and direction symbol. The blue area represents the network coverage.

Model	Parameters	Values
Advection (DARTS)	Coefficients that control the truncation of the Fourier transformation N_s, N_x, N_y, and N_t. M_x and M_y control the spatial resolution of the advection vectors	$N_s = 9$, $N_x = 50$, $N_y = 50$, $N_t = 4$, $M_x = 2$ and $M_y = 2$
Advection (Dual TV-L^1)	Observation period for the calculation of and smoothness degrees (λ) of the advection vectors	30 min, equivalent to 6 observation frames, and $\lambda = 0.02$
Advection + ARI	AR(p) order and first order differentiation (d)	$p = 2$ and $d = 1$ and $K = 6$
SPROG	AR(p) order, number of decomposition levels (K), side-extent in pixels of the observation field (L), filling value over dry areas ($R_{\min}$), threshold value for wet areas (R_{th}), and the scaling parameter for the second Gaussian-shaped filters (β)	$p = 1, 2$ and $K = 3, 6, 12$ and $L = 1024$ and $R_{\min} = -15$ dB and $R_{th} = 0.1$ mm h^{-1} and $\beta = 0.5$ (at $K = 3$), 0.25 (at $K = 6$), 0.6 (at $K = 12$). In addition, the q ratios are given by 8 (at $K = 3$ and $|\mathbf{k_w}(2)| = 64$), 2.83 (at $K = 6$ and $|\mathbf{k_w}(2)| = 8$), and 1.62 (at $K = 12$ and $|\mathbf{k_w}(2)| = 4$)
STEPS	As these for SPROG and configuration for the stochastic generation using SSFT	$p = 1$ and $K = 6$ and 20 ensemble members, SSFT with 9 sub-domains with an edge length of $0.5L$ and an overlapped percent of 50%. The remaining parameters are the same as for SPROG but only using CDF-matching
SPROG-LOC	Configuration for the length of the convolution kernel and the 2-D correlation coefficients	In addition to the SPROG and STEPS parameters, $\mathbf{k}_{\min} = 2$ and $h = 3$, and $l_{\min} = 100$ km and $I = 10, 20\%$
STEPS-LOC	As these for SPROG-LOC and STEPS	Same values as for STEPS and SPROG-LOC but only using CDF-matching

central frequency of the second filter. In our setup q is given as

$$q = \left(\frac{L}{2|\mathbf{k_w}(2)|} \right)^{\frac{1}{K-2}} \qquad (11)$$

so that if K is equal to 3, 6, and 12 while $|\mathbf{k_w}(2)|$ equals 64, 8, and 4, then q is equal to 8.0, 2.83, and 1.62, respectively. Additionally, a user can also adjust the peak of the filter centered at $|\mathbf{k_w}(2)|$ by setting a scaling parameter $\beta \in [0, 1]$ to maintain a uniform peak of the K Gaussian-shaped filters. The values of β are set at 0.5, 0.25, and 0.6 for values of K equal to 3, 6, and 12, respectively. The same values for $|\mathbf{k_w}(2)|$, q, and β are also used in SPROG.

We have realized that the μ-adjustment applied to $R_{\mathrm{T}}(t + t_l)$ becomes necessary when the decomposition step includes areas with no precipitation, i.e., the areas filled by $R_{\min}$. Without this adjustment, the SPROG method will lead to nowcasting skills lower than those of the advection model since SPROG typically underestimates precipitation intensity as it is biased towards $R_{\min}$. This effect is more noticeable at a low wet-area ratio. Therefore, to run SPROG one could either use a mask that excludes these areas and avoid the μ-adjustment or keep these areas and use the μ-adjustment. This is because a mask will exclude these areas during the estimation of the scaling parameters and the autocorrelation coefficients. However, for SPROG-LOC, these areas will not be excluded since the estimation of the standard deviation as given by (7) and the μ-adjustment will be

applied. Last, the CDF-matching will be tested in both SPROG and SPROG-LOC.

For experiments with the ARI model, we use a second-order AR and a first-order differentiation at each of the six spatial scales. Since the evolution of intensity changes is directly modeled by the ARI(p, d) process, we go without a postprocessing step.

For the generation of the stochastic term needed to conduct STEPS and STEPS-LOC nowcasting, which will be discussed in Section V, we set the SSFT method to nine subdomains with an edge length of $\frac{1}{2}L$ and an overlap between each subdomain of 50 %. Finally, we set the number of ensemble members to 20, while the postprocessing is set by the CDF-matching for both STEPS and STEPS-LOC.

C. Nowcasting Skill Scores

In this study, we use a set of neighborhood verification metrics to evaluate the extrapolation forecasts. Neighborhood metrics are chosen for three reasons. First, to mitigate the double penalty problem occurring when high-resolution data are compared with traditional scores. Second, to evaluate at which spatial scales and up to which lead-time the forecasts provide skill and third, to consider the spatial variability of precipitation among the selected spatial scales. The fraction of grid boxes exceeding a given threshold within a surrounding size of a pixel can be

expressed as a probability [50], [51]. It is defined as

$$\mathbf{P}_m(i,j) = \frac{1}{m^2} \sum_{u,v=1}^{m} \mathbf{B} \left(i + u - 1 - \frac{m-s}{2}, \right.$$
$$\left. j + v - 1 - \frac{m-s}{2} \right), \tag{12}$$

where $\mathbf{B}$ describes the binary field resulting from exceeding a predefined precipitation threshold P_{th}, whereas m represents the side extent in km, hereafter the window size, of the considered squared neighborhood while i and j indicate the location of a pixel since the horizontal resolution is $\Delta x = 1$ km. In addition, if m is even then s equals 0, otherwise s equals 1. Based on these neighborhood probabilities, we construct classical categorical metrics, namely, the probability of detection (POD), the false alarm ratio and the critical success index (CSI). An overview of these metrics can be found in [52] or [42]. The required contingency table is constructed using the weakest constraint, i.e., if $\mathbf{P}_m$ is larger than zero, then it is counted as an occurred event. The table contains the following parameters: hits (H), false alarms (FA), misses (MI), and correct negatives (CN). The POD is given by $H(H+MI)^{-1}$ and quantifies the rate of observed events that were successfully predicted while the FAR, equal to $FA(H+FA)^{-1}$, indicates the fraction of incorrect prediction. The CSI is estimated as $H(H+FA+MI)^{-1}$ and represents the detection skill of the prediction model. When increasing the window size, the values of H will increase correspondingly while those of MI and FA will decrease. In addition to these categorical metrics, the well-known root mean squared error (RMSE) can also be used when evaluating nowcasting skill.

When using the values of $\mathbf{P}_m$ itself, the fractions skill score (FSS; [53]) can be constructed. The FSS evaluates the frequency that the predictions are similar to the observations at a given precipitation threshold and window size. This is done by comparing the fraction of predicted pixels exceeding P_{th} with that from observations within a neighborhood as

$$\text{MSE}_m = \frac{1}{L_x L_y} \sum_{i=1}^{L_x} \sum_{j=1}^{L_y} \left[\mathbf{P}_m^O(i,j) - \mathbf{P}_m^F(i,j) \right]^2$$
$$\text{FSS}_m = 1 - \frac{\text{MSE}_m}{\text{MSE}_{\text{ref},m}}. \tag{13}$$

L_x and L_y describe the domain dimensions in x- and y-direction, respectively, whereas $\mathbf{P}_m^O$ and $\mathbf{P}_m^F$ represent the fields of observed and forecasted neighborhood probabilities for the threshold P_{th} and the window size m. The mean squared error used as a reference (MSE_{ref}) indicates the largest possible MSE resulting from the forecasted and observed neighborhood probabilities.

IV. SPROG SENSITIVITY ANALYSIS

The performance of the SPROG(p, K) model and its sensitivity to the pth order of the AR model and the spatial decomposition number K are evaluated pixel-by-pixel (i.e., without the neighborhood probability method) and at P_{th} equal to 0.1 and 5.0 mm h^{-1}, as shown in Figs. 3 and 4, using the precipitation

events listed in Table I. This analysis will allow us to better initialize and evaluate the SPROG and SPROG-LOC models as described later in Section V. For the purpose of this analysis, the predicted fields were not postprocessed and the advection vectors were estimated using the DARTS optical-flow technique.

From Fig. 3, at $P_{\text{th}} = 0.1$ mm h^{-1}, it is seen that the configured parameters for the SPROG model led to a better performance than that of the advection model for the 3 h lead-time, i.e., higher values of CSI and POD and lower values of FAR and RMSE. The results from SPROG(2,K) show in general a slight improvement in comparison to SPROG(1,K). The number of decomposition levels does not seem to play a role at SPROG(1,K), whereas SPROG(2,6) leads to slightly higher skills than SPROG(2,3). However, a further increasing of K from 6 to 12 does not necessarily improve the prediction skill. The different setups regarding AR order and number of decomposition levels show that an AR process with extended memory is able to better represent the temporal evolution of the observed precipitation fields. However, this improved representation is limited to large-scale features as shown by the results of using different number of cascade levels. For the prediction of small-scale features, a first-order AR model is sufficient since their lifetime is very short. One may also see that a low number of cascade levels is not enough to accurately represent the scale-dependence of predictability. In terms of RMSE, the SPROG model exhibits a decreasing behavior compared to the advection model. This is because the AR process tends to smooth unpredictable components of precipitation. This smoothing may also explain the reduced penalty in terms of FAR. However, only when the smoothing is associated with an increase of observed precipitation within forecast lead-time (cf. Fig. 1). If there is a decrease of observed precipitation, it is likely that the smoothing spreads precipitation also in nonprecipitation areas, increasing FAR values associated with SPROG. Fig. 4 shows skill scores lower than in Fig. 3 because the evaluated areas are associated with values of precipitation larger than 5.0 mm h^{-1}, i.e., with less predictable regions of precipitation since phenomena with such intensities are typically very small and thus have a short lifetime. Note that SPROG(p,3) performs less well than the advection model as depicted by the CSI and POD curves. In terms of the CSI, the combinations of SPROG at $K = 6$ and 12 give similar skill, outperforming the advection model. In terms of POD, SPROG(2,6) leads to a lower performance than that of SPROG(2,12), however, both SPROG(1,6) and SPROG(1,12) are similar and also similar to that of SPROG(2,12).

The subplot in Fig. 4(b) indicates the accumulated number per lead-time N in which a precipitation nowcasting field is verified by an observed precipitation field. The maximum number of N is 298 and it is the total number of observed precipitation fields from the 10 events. One can see a decreasing tendency of N with lead-time. This is caused by the filtering nature of the SPROG model that leads to smoothed nowcasting fields and whose values are not high enough compared to those from observations. above P_{th}. More precisely, only these instances where at least one nowcasting pixel exceeds the precipitation threshold were considered. The lowest curve of N is obtained by SPROG at $p = 2$ and $K = 3$ while the three highest curves are
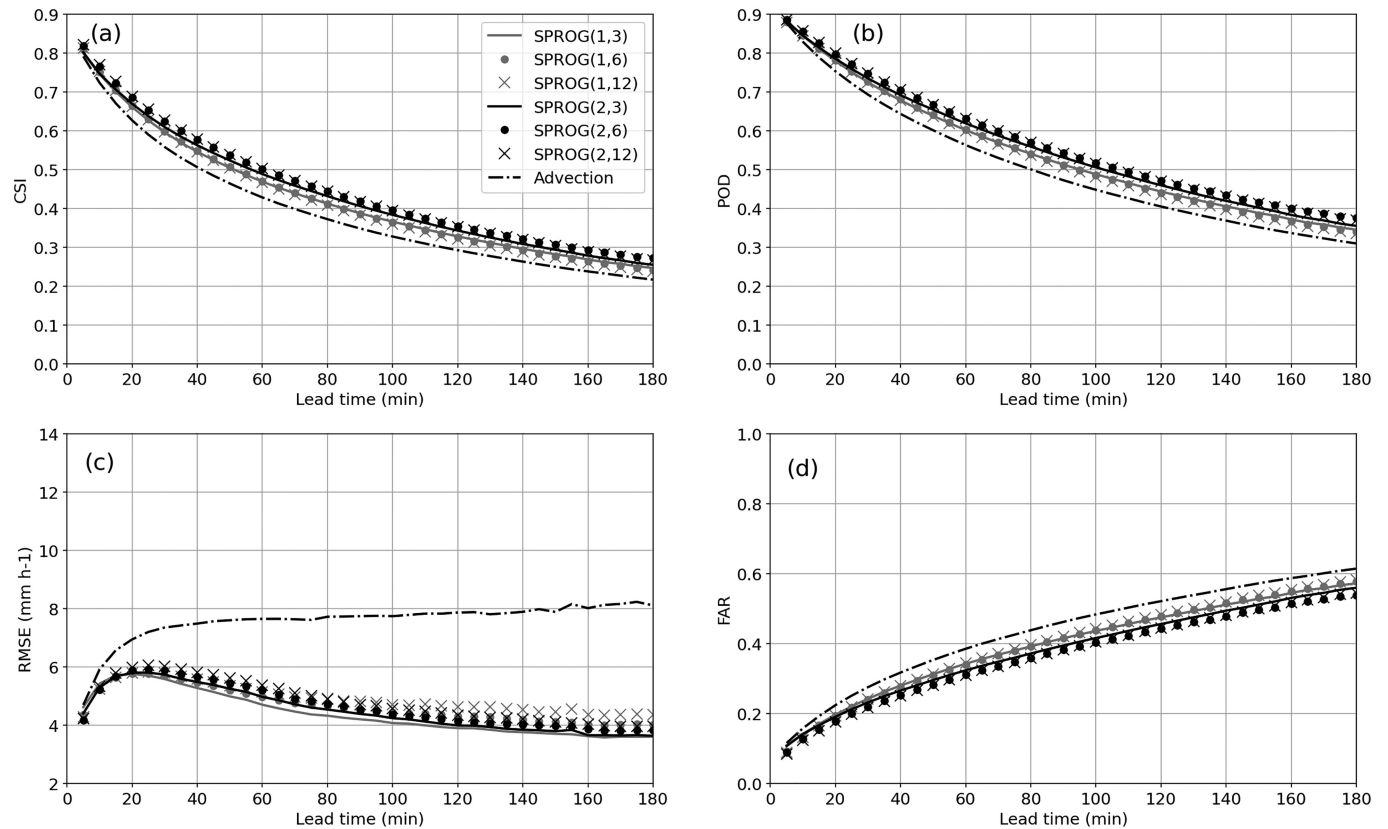
Fig. 3. Skill scores as a function of 3 h lead-time for different configuration of SPROG (AR-order and scale levels) or SPROG(p, K) at $P_{th} = 0.1$ mm h^{-1}, resulting from the 10 events listed in Table I. (a) CSI scores resulting from SPROG(1,3) in grey line, SPROG(1,6) in grey dots, and SPROG(1,12) in grey crosses. Those from SPROG at $p = 2$ are shown in black color. The scores from advection nowcasting (dotted line) are also shown. (b)–(d) As in (a) but for POD, RMSE, and FAR, respectively. In this experiment, probability matching was not applied but the SPROG fields were masked following the adapted approach of [30] as indicated in Section II.

obtained by SPROG at $p = 1$ and $K = 6$, SPROG at $p = 1$ and $K = 12$, and by the advection model. These curves allow us to see a connection among the decomposition levels, the AR order, and the precipitation values of the smoothed nowcasting fields. The N values from the advection model slightly deviate from 298 because there were episodes in which observation values were below the threshold of 5.0 mm h^{-1}. This was not the case when P_{th} was set to 0.1 mm h^{-1}. As the N values from SPROG at $p = 1$ are larger than those from SPROG at $p = 2$, for both $K = 6$ and 12, and the skill scores from SPROG at $p = 1$ and $K = 6$ are similar to those at at $p = 1$ and $K = 12$, for both P_{th} values, it is suggested to use an AR(1) process with six decomposition levels to configure SPROG. Therefore, the remainder of the analysis in this work will be conducted at SPROG(1,6).

The radially averaged power spectral density RAPSD allows to compute the 1-D spectrum of precipitation fields to study the impact of spatial filtering methods, such as SPROG, on the loss of small-scale variability. As an example, Fig. 5 shows that the power distributions resulting from SPROG at different lead-times are smaller than those from the advection model and observations because of the spatial filtering characteristic of SPROG driven by the AR coefficients and the iterated AR process. For instance, the RAPSD from SPROG is reduced

by at least 10 dB at 45 min lead-time for wavelengths up to 8 km and at 120 min lead-time for wavelengths up to 32 km. This is consistent with the decreasing tendency of N with lead-time shown in Fig. 4. The impact of CDF-matching on the RAPSD is clearly seen by a power increase of 10 dB for wavelengths down to 2 km at a lead-time of at least 45 min. This is, however, a successful adjustment assuming that the CDF of the observed precipitation remains constant during extrapolation.

To further illustrate the impact of CDF-matching, Fig. 6 shows the CSI, the normalized RMSE, and the FSS at $P_{th} = 0.1$ and 5.0 mm h^{-1} resulting from configuring SPROG with and without CDF-matching. The normalization of RMSE (nRMSE) is motivated to mitigate the penalty caused by the computation of more available samples from the predicted fields with CDF-matching. RMSE is normalized by the median of the observed field, masked by the condition that both predicted and observed values are larger than P_{th}. At $P_{th} = 0.1$ mm h^{-1}, the CSI indicates a similar performance from both configurations of SPROG but associated with increased nRMSE values, which are close to those resulting from the advection model. The reason for that is because the predicted values from SPROG with CDF-matching are redistributed to match the CDF from observations, experiencing a penalty from the redistributed
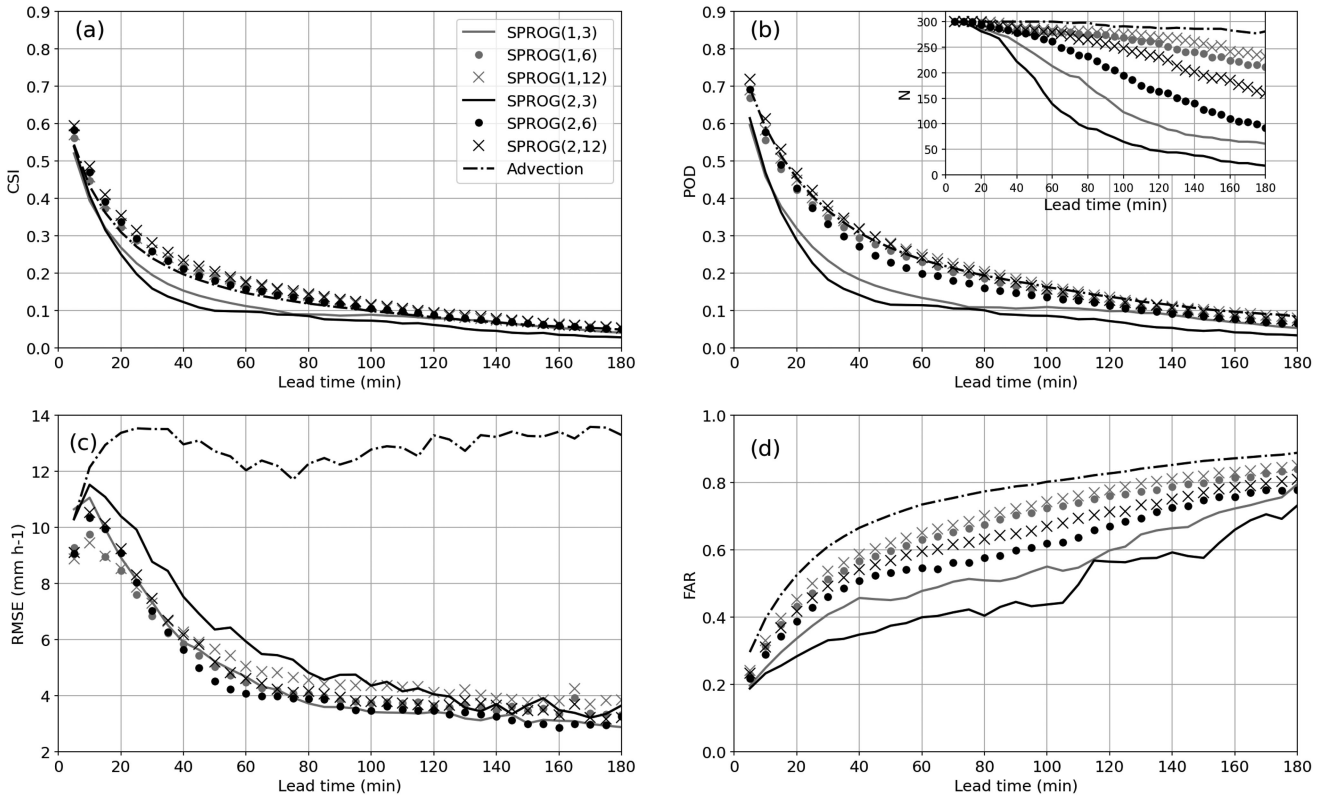
Fig. 4. (a) CSI, (b) POD, (c) RMSE, and (d) FAR as in Fig. 3, but evaluated at $P_{th} = 5.0$ mm h^{-1}. The number of available samples for evaluation as a function of lead-time is also shown in panel (b). Note that for lead-times larger than approximately 60 min, the scores do not represent a useful forecast skill.
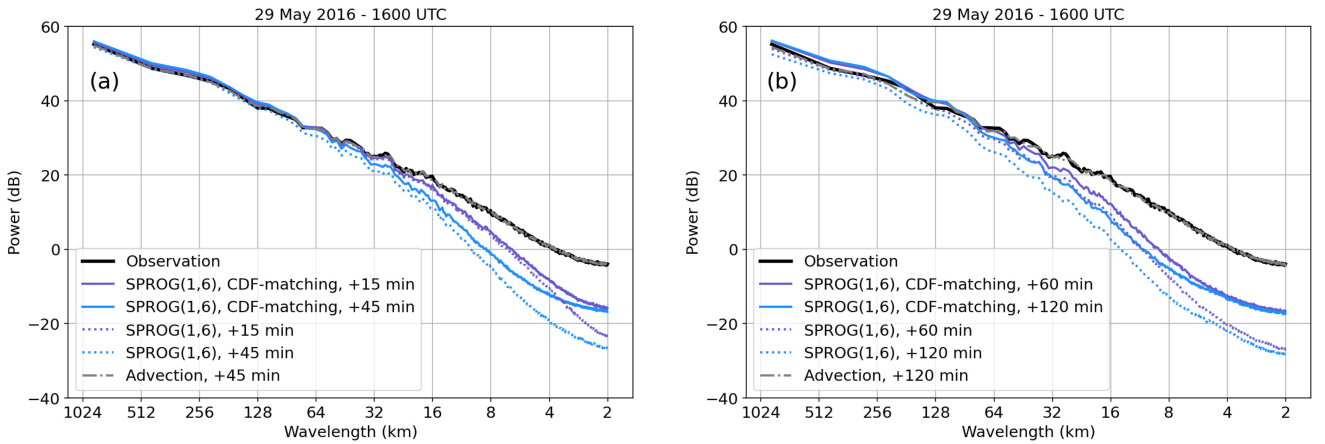

Fig. 5. Radial averaged power spectral density (RAPSD) of observed and predicted fields as a function of spatial wavelength, event E4 at 1600 UTC. (a) RAPSDs from SPROG with CDF-matching (continuous blue) and without matching (dashed blue) at 15 and 45 min lead-time. The RAPSD from advection is given at 45 min lead-time as gray dotted-line. (b) Similar as (a) but at 60 and 120 min lead-time.

areas. At 5.0 mm h^{-1}, a similar penalty is also observed, however, the CSI shows enhanced values compared to those from SPROG without CDF-matching. Thus, the CDF-matching configuration could be a reasonable attempt to partially address the smoothing behavior of SPROG tolerating nRMSE values similar to those from the advection model. Results related to the μ-adjustment, not shown here, led to similar skills to those from SPROG without CDF-matching since the computation

during the scale decomposition and correlation coefficients steps considered only values of R_T above R_{min}.

The FSS curves, also shown in Fig. 6, at both exceeding thresholds and at 1 km window size are consistent with the results depicted by the CSI curves. However, at $P_{th} = 0.1$ mm h^{-1}, the performance of SPROG and SPROG with CDF-matching equals that of the advection model for window sizes larger than 1 km. In contrast, at $P_{th} = 5.0$ mm h^{-1}, the performance of both
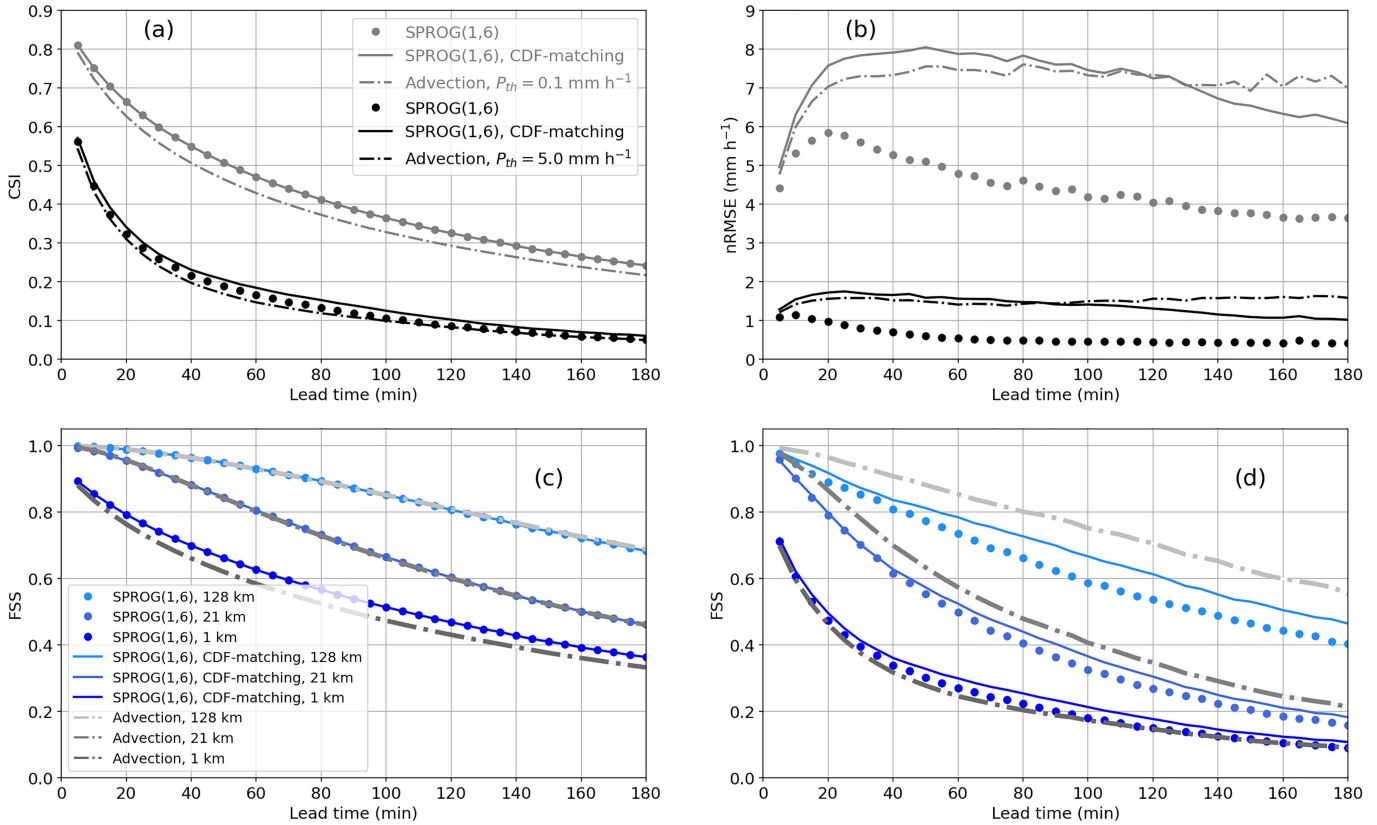
Fig. 6. (a) CSI scores resulting from SPROG(1,6) evaluated at $P_{th} = 0.1$ mm h^{-1} are shown in grey while these at $P_{th} = 5.0$ mm h^{-1} in black color. The scores related to SPROG with CDF-matching are indicated by the solid lines whereas these without matching are marked by dots. (b) As in (a) but for the normalized RMSE. FSS curves evaluated at (c) 0.1 and (d) 5.0 mm h^{-1} for window sizes of 1 (dark blue), 8 (medium blue), and 128 (light blue) km.

SPROG configurations at larger window sizes seems to be lower than that from the advection model. This means that at higher thresholds and larger window sizes, SPROG is penalized due to its spatial filtering essence. Given this and the fact the SPROG with CDF-matching only leads to a slightly improvement at $P_{th} > 0.1$ mm h^{-1} but with increased values of nRMSE, it is suggested to apply the μ-adjustment, when no initial mask is used, and avoid the CDF-matching as the goal of SPROG is to keep low nRMSE values. For visualization purposes, the resulting SPROG precipitation fields with μ-adjustment at 60 and 120 min are shown in Fig. 7 panels (b) and (e), respectively.

V. SPROG-LOC EVALUATION RESULTS

In this section, we validate the skill of the SPROG-LOC approach and evaluate its possible configurations. The skill scores will be calculated at window sizes equal to or larger than 1 km, i.e., $m \geq 1$, in which the neighborhood probability method allows a spatial oriented evaluation, and at multiple precipitation thresholds P_{th}. First, we verify the SPROG-LOC approach in correspondence with the SPROG and advection+ARI methods followed by a brief analysis of the nowcasting skill of the STEPS-LOC model in terms of FSS values. Second, we provide a statistical evaluation and discussion of the performance of the SPROG-LOC and advection+ARI methods.

A. Verification of SPROG-LOC

To compare the nowcasting fields from SPROG and SPROG-LOC, Fig. 7 displays an example of nowcasting fields from both approaches at 60 and 120 min lead-times from precipitation observed of event E4 at 16:00 UTC. As expected, the nowcasting fields from SPROG are characterized by a smoothing appearance, mainly over small convective precipitation regions. However, it can be seen that SPROG-LOC is able to maintain the intensity of small precipitation regions, i.e., reducing the excessive smoothing, while conducting the filtering AR process. For instance, the regions highlighted by the white circles seem to qualitatively match those regions in the corresponding observations fields better than those from SPROG.

Similar to Fig. 5, the RAPSDs from SPROG and SPROG-LOC are shown in Fig. 8. One can see that the nowcasting fields resulting from SPROG-LOC maintain a relative higher power of precipitation than those from SPROG for wavelengths up to 32 km and with a power difference of at least 10 dB. Note that the CDF-matching applied to SPROG-LOC led only to an increase of 5 dB at 2 km wavelength. As expected, the RAPSD of the observed precipitation field remains higher than the corresponding curves of SPROG-LOC (up to 8 km wavelength) and SPROG (up to 32 km wavelength).

Fig. 9 depicts the results given by CSI, POD, and FAR to compare different nowcasting models and configurations evaluated at 5.0 mm h^{-1} and at two window sizes, $m = 1$ and 21 km. At
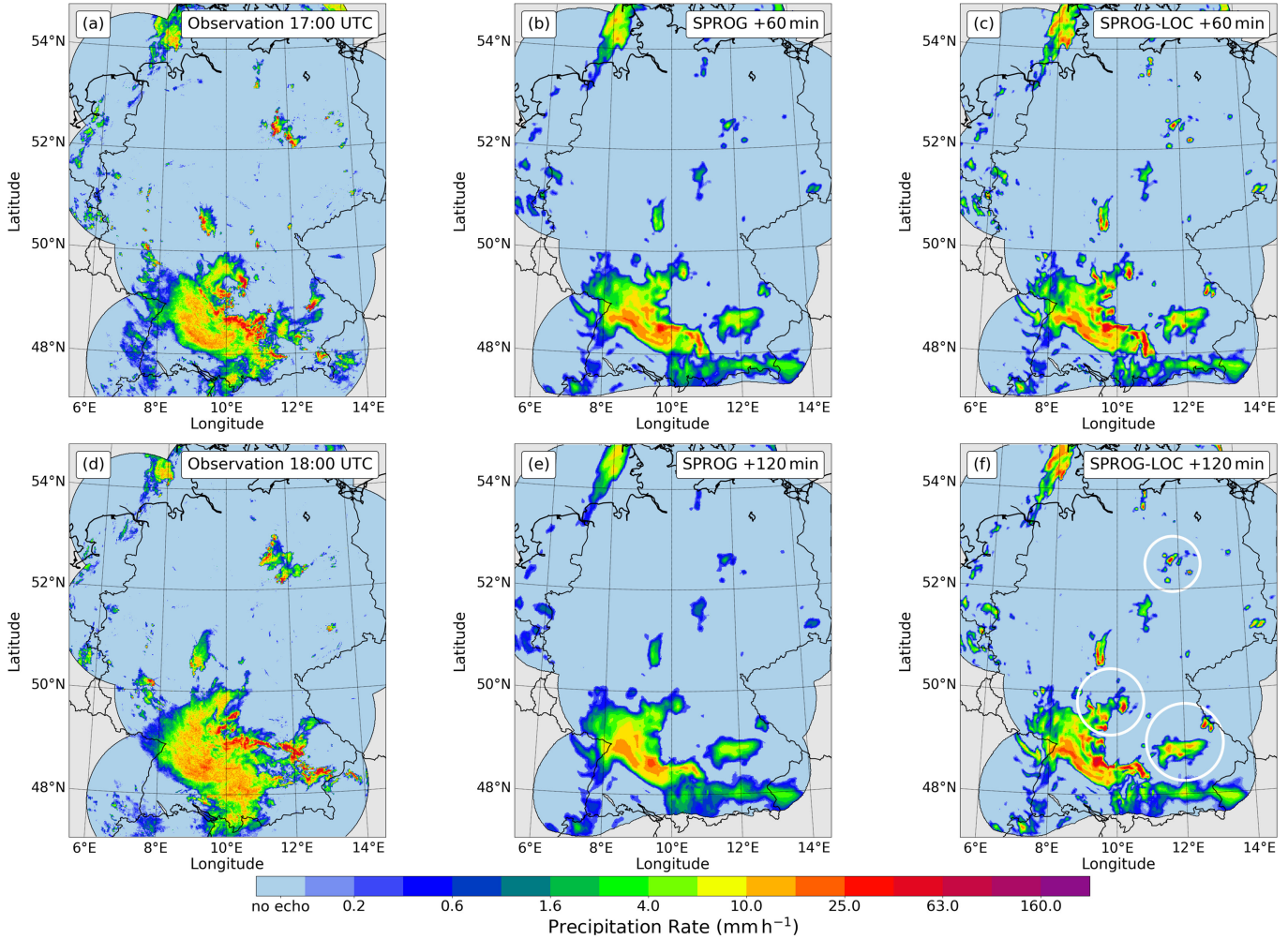
Fig. 7. RY product from event E4 observed at (a) 17:00 and (d) 18:00 UTC. SPROG(1,6) with μ-adjustment at (b) 60 and (e) 120 min lead-time from observations at 16:00 UTC. (c) and (f) As (b) and (e) but using SPROG-LOC(1,6) at $I = 20\%$, with μ-adjustment, and with a minimum kernel size, for the AR estimation parameters, of 100 km. The white circles localize areas of high intensities, matching these from observations.

1 km, the CSI curves drop quickly within the first 60 min in which the three lowest curves are given by the advection model and SPROG-LOC at $I = 10$ and 20% without any postprocessing step, except SPROG. This becomes clear when observing the POD results. Additionally, the highest curve is given by the SPROG-LOC at $I = 20\%$ with μ-adjustment. In terms of the FAR, SPROG-LOC with $I = 20\%$ and either post-processing options leads to similar results as that from the advection model whose FAR values are slightly larger than those of SPROG for lead-times larger than 30 min.

At 21 km window size, a much clearer distinction of the forecast skills among the SPROG, SPROG-LOC, advection, and advection+ARI models can be observed. The SPROG and SPROG-LOC at $I = 10\%$ perform similar, meaning that both are more conservative models, one because of its excessive smoothing and the other due to its mean bias during the spatial decomposition. A slight improvement is seen when I is set to 20% due to the slower decorrelation of the adjusted AR coefficients. Note that the SPROG-LOC at $I = 20\%$ and with either postprocessing

step, a balance is obtained as this configuration of SPROG-LOC conducts a spatial filtering process while considering the spatial variability of precipitation. Although the skills, at 1 km, given by the advection+ARI and SPROG-LOC models are comparable, at 21 km window size the advection+ARI shows a better skill than that of SPROG-LOC. This is due to the utilization of a source/sink term by the former method that allows capturing the growth and decay process of precipitation while avoiding smoothing nowcasting fields.

The nowcasting fields were extrapolated using the advection vectors derived from the Dual TV-L^1 technique. From a qualitative assessment, it is observed that the CSI from the advection model, Fig. 9(a), is nearly equal to the one shown in Fig. 6(a) at 5.0 mm h^{-1}, where the advection vectors were estimated using the DARTS approach. This is consistent with the results given by [54], who provided a detailed analysis on the impact of optical-flow methods to nowcasting.

Fig. 10 shows a similar verification but in terms of the FSS at 1) 1 km and 0.1 mm h^{-1} and 2) 21 km and 5.0 mm h^{-1}. Fig. 10(a)
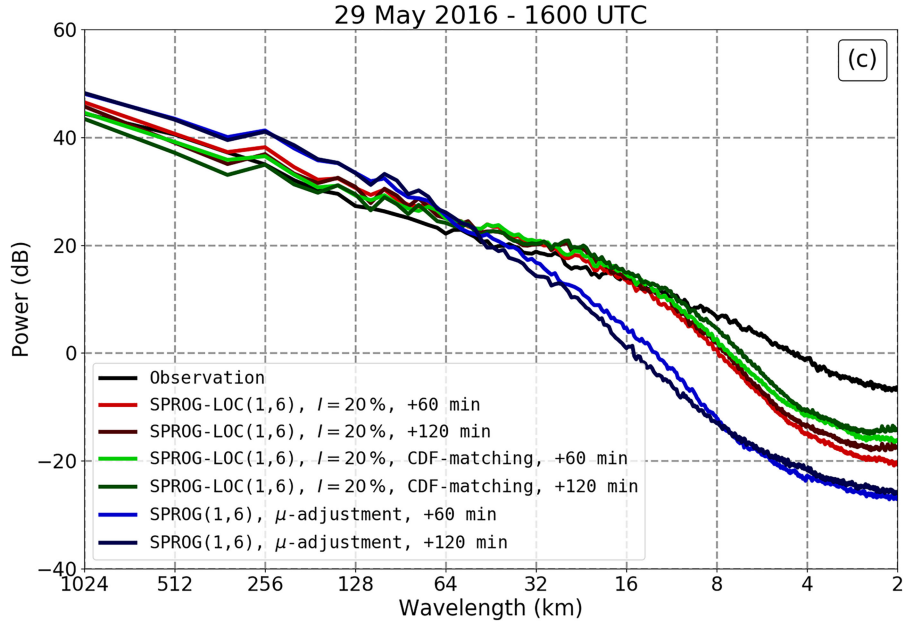
Fig. 8.   RAPSD as in Fig. 5 but for SPROG-LOC. The light green (light red) curve depicts SPROG-LOC with (without) CDF-matching at a lead-time of 60 min. For a lead-time of 120 min, RAPSD of both SPROG-LOC settings are illustrated by a dark green and dark red curve, respectively. Further, the RAPSD of the corresponding observation is shown in black and the one of SPROG with μ-adjustment is represented in light blue and dark blue at a lead-time of 60 min and 120 min, respectively.

highlights that the best performance is given by SPROG due to its better skill on predicting precipitation areas that consist of low intensities. As in the case of the CSI, the FSS curve from the advection model seems to match well with the one from the advection model but using the DARTS approach as seen in Fig. 6(c) at 1 km window size. Fig. 10(b) indicates a similar balance obtained by SPROG-LOC with either postprocessing option but with a narrow distinction from the SPROG and advection models. Note that, at 1 km and 0.1 mm h^{-1}, all nowcasting models give useful nowcasting skill for lead-times up to 3 h approximately. However, at 21 km and 5.0 mm h^{-1}, useful lead-times are reduced to 120 min (advection+ARI), 110 min (advection), and 100 min (SPROG-LOC). In general, the scores resulting from the SPROG-LOC and possibly advection+ARI models can further increase when most of the observations are related with large wet-area ratios. However, the presented evaluation also included precipitation periods in which the wet areas were reduced to less than 25% of the spatial domain with the goal of evaluating under operational conditions.

Fig. 10(c) and (d) shows the FSS results when the AR process, both globally and locally, is perturbed by the stochastic term, i.e., the STEPS and STEPS-LOC model. The CDF-matching was selected for both models because the spatial resolution of the ensemble nowcasting fields is similar to that of observations. It can be seen that at both evaluating set-ups related to the window size and precipitation threshold, the average FSS resulting from STEPS-LOC, represented by a randomly selected member, show values higher than those obtained by STEPS for the 3 h lead-time. As the STEPS-LOC leads to FSS values near to those of advection, it can be said that the ensemble members generated

by STEPS-LOC are able to reduce the penalty from the spatial filtering aspect by imprinting unpredictable precipitation structures at high spatial resolution during extrapolation to model the resulting forecast uncertainty. In contrast, the perturbation term in STEPS is driven globally by a single AR coefficient and when it is evaluated at P_{th} values higher than 0.1 mm h^{-1}, the average skill reduces even lower than that of the advection model. Namely, because the stochastic term in STEPS is correlated to the spatial distribution of the observed precipitation with values equal to or larger than 0.1 mm h^{-1}. Thus, the benefit of the localization approach is also imprinted on the distributed spatially localized field of AR coefficients $\Phi_{k,0}$ that model the evolution of the stochastic term. As these coefficient fields are straightforward derived from $\Phi_{k,n}$, it is expected that the computational burden required to generate localized ensemble members will be comparable to that of STEPS, which is suitable for operational purposes. A detailed analysis on the performance of ensemble members resulting from STEPS and STEPS-LOC and their forecast skill based on probabilistic scores (e.g., CRPS, reliability diagram, ROC curve) is encouraged but it goes beyond the scope of this work.

B. Statistical Evaluation of SPROG-LOC and Advection+ARI

To extend the analysis of SPROG-LOC at several P_{th} and window size m values, the heatmaps shown in Fig. 11 display in a) the median CSI of SPROG as basis for the average skill difference versus SPROG-LOC and advection+ARI, respectively, at 60 min lead-time. The areas in red indicate a better performance of SPROG, whereas the areas in green represent improvements from SPROG-LOC at different configurations
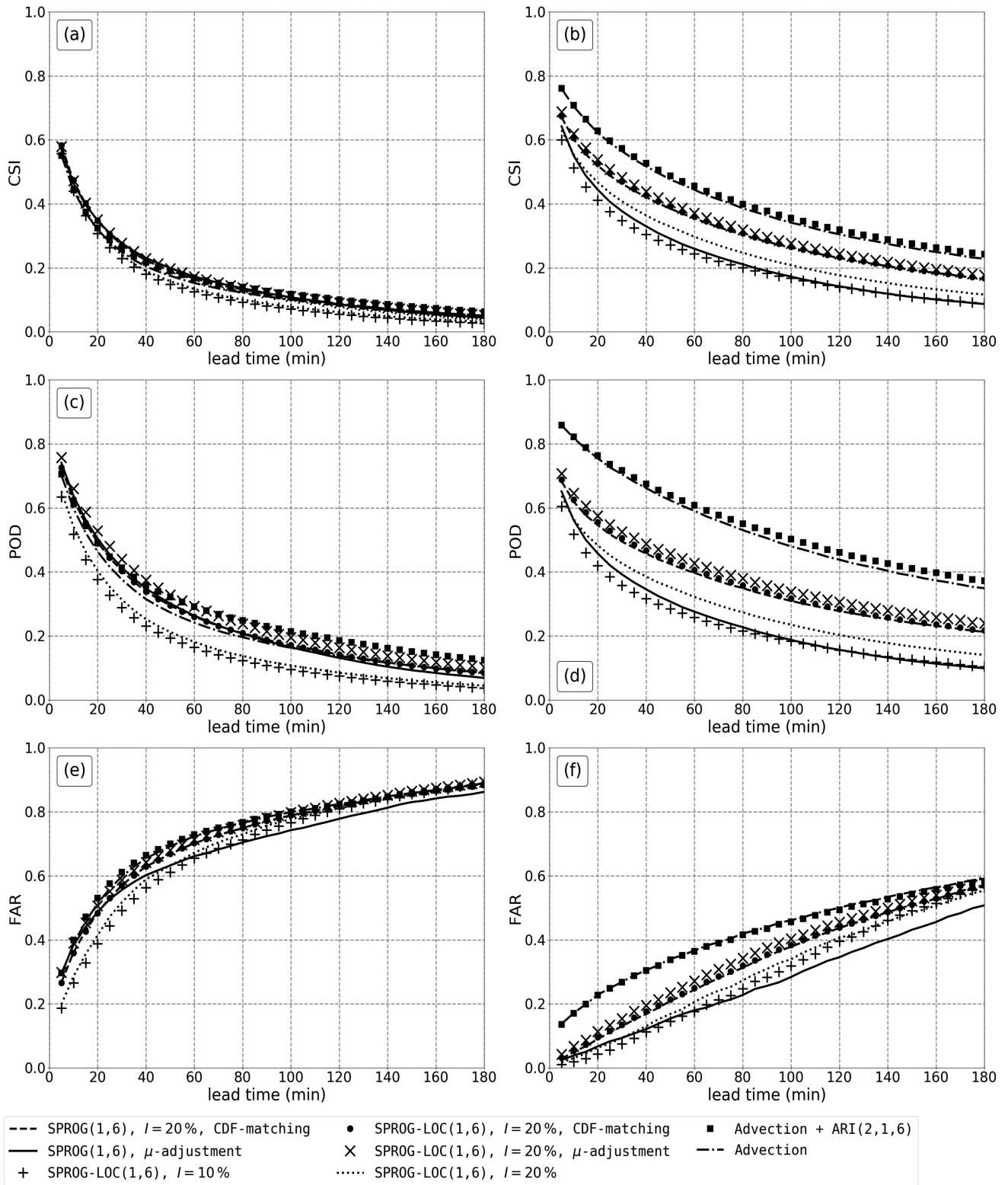
Fig. 9.   Skill scores for different configurations of SPROG-LOC(1,6) as well as SPROG(1,6), advection with ARI(2,1,6) and advection only. Left column: (a) CSI, (c) POD, and (e) FAR curves evaluated at $P_{th} = 5.0$ mm h^{-1} and at 1 km window size. Right column: same es left column but evaluated at 21 km window size using a neighborhood approach. Scores resulting from SPROG(1,6) and advection are indicated by the solid and dash-dotted lines, respectively. Results of the ARI model are depicted with filled squares. One should note that values of CSI and POD scores smaller than approx. 0.2 and 0.3 do not represent an useful forecast skill.
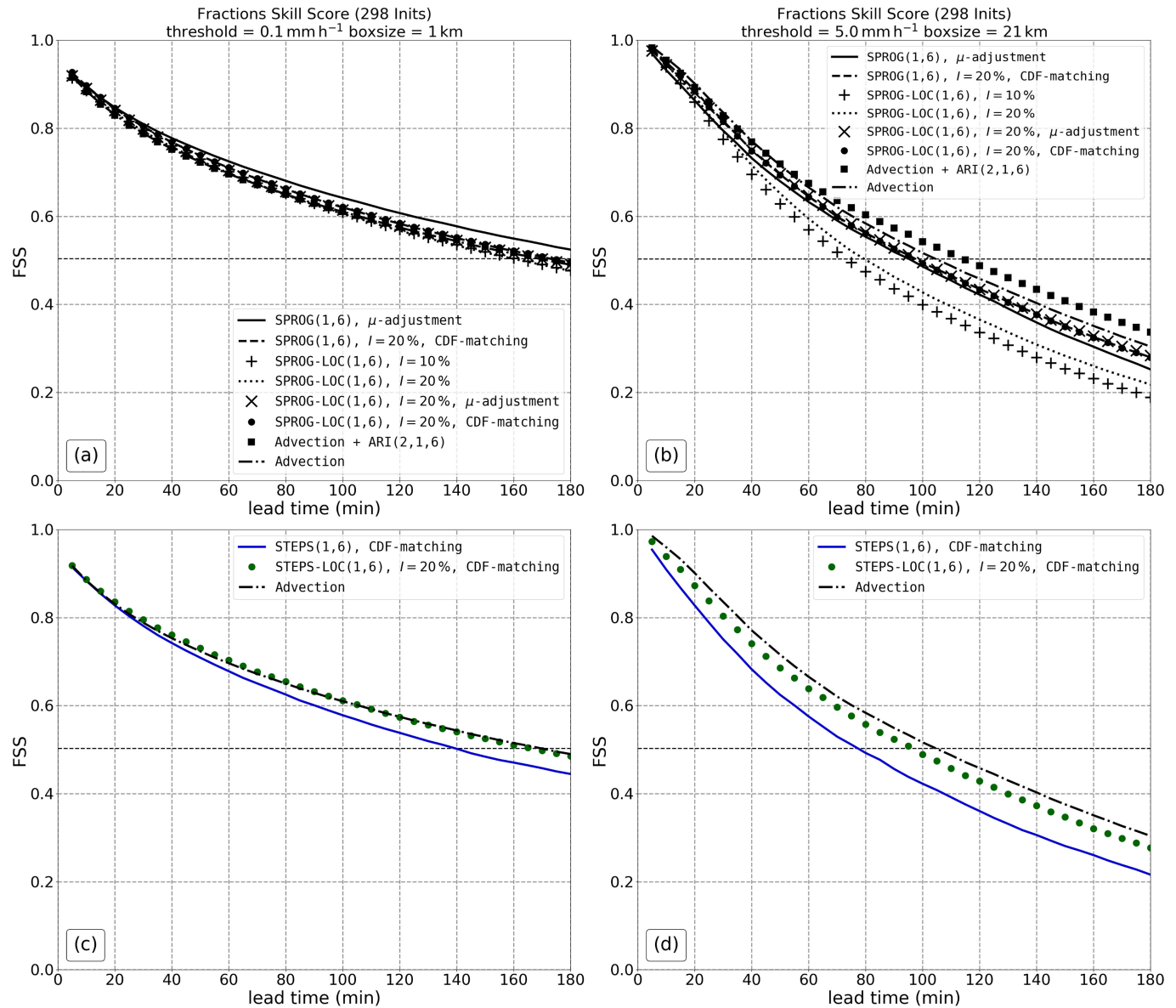
Fig. 10. Top: As in Fig. 9 but for FSS at (a) $P_{th} = 0.1$ mm h^{-1} and 1 km window size and (b) $P_{th} = 5.0$ mm h^{-1} and 21 km window size. Bottom: Mean FSS curves resulting from one randomly chosen member of each CDF-matched STEPS ensemble forecast. The blue line depicts STEPS whereas the green dots represent STEPS-LOC with $I = 20\%$ in (c) for $P_{th} = 0.1$ mm h^{-1} and 1 km window size and in (d) for $P_{th} = 5.0$ mm h^{-1} and 21 km window size. As in the top row, the dash-dotted line indicates the advection.

and advection+ARI, respectively. For example, for an increased value of I from to 10% to 20%, it is seen that SPROG-LOC outperforms SPROG at P_{th} and window size m values of at least 1.5 mm h^{-1} and 11 km. This is because an increase on the adjustment of the AR coefficients mitigates the tendency of faster decorrelation rate, mainly at lower spatial scales. The green areas are further expanded and enhanced when a postprocessing step is applied to SPROG-LOC. For instance, the μ-adjustment allows a better performance of SPROG-LOC when it is evaluated at values of P_{th} and window size m as low as 0.1 mm h^{-1} and 11 km or at a smaller window size of 5 km but at larger P_{th} of 1.5 mm h^{-1}. A similar result is obtained when SPROG-LOC is followed by CDF-matching. Since an initial mask cannot be used for SPROG-LOC, either the μ-adjustment or CDF-matching is

needed. The advection+ARI extrapolations model solely precipitation tendencies and not the whole field itself. Therefore, the SPROG typical smoothing is prevented. This behavior could be seen in Fig. 11(b). Advection+ARI extrapolations exhibit at grid-scale and up to a threshold of 5.0 mm h^{-1} a slightly lower score compared to SPROG. However, at larger window sizes a strong improvement is apparent ranging over all considered thresholds.

A similar pattern is observed in Fig. 12 but in terms of the FSS. For example, the SPROG-LOC outperforms SPROG when the nowcasting fields are evaluated using at least a P_{th} value of 0.5 mm h^{-1} and a window size of 21 km or at 5.0 mm h^{-1} and smaller window size of 5 km. This is observed when either the μ-adjustment or CDF-matching are applied to SPROG-LOC's
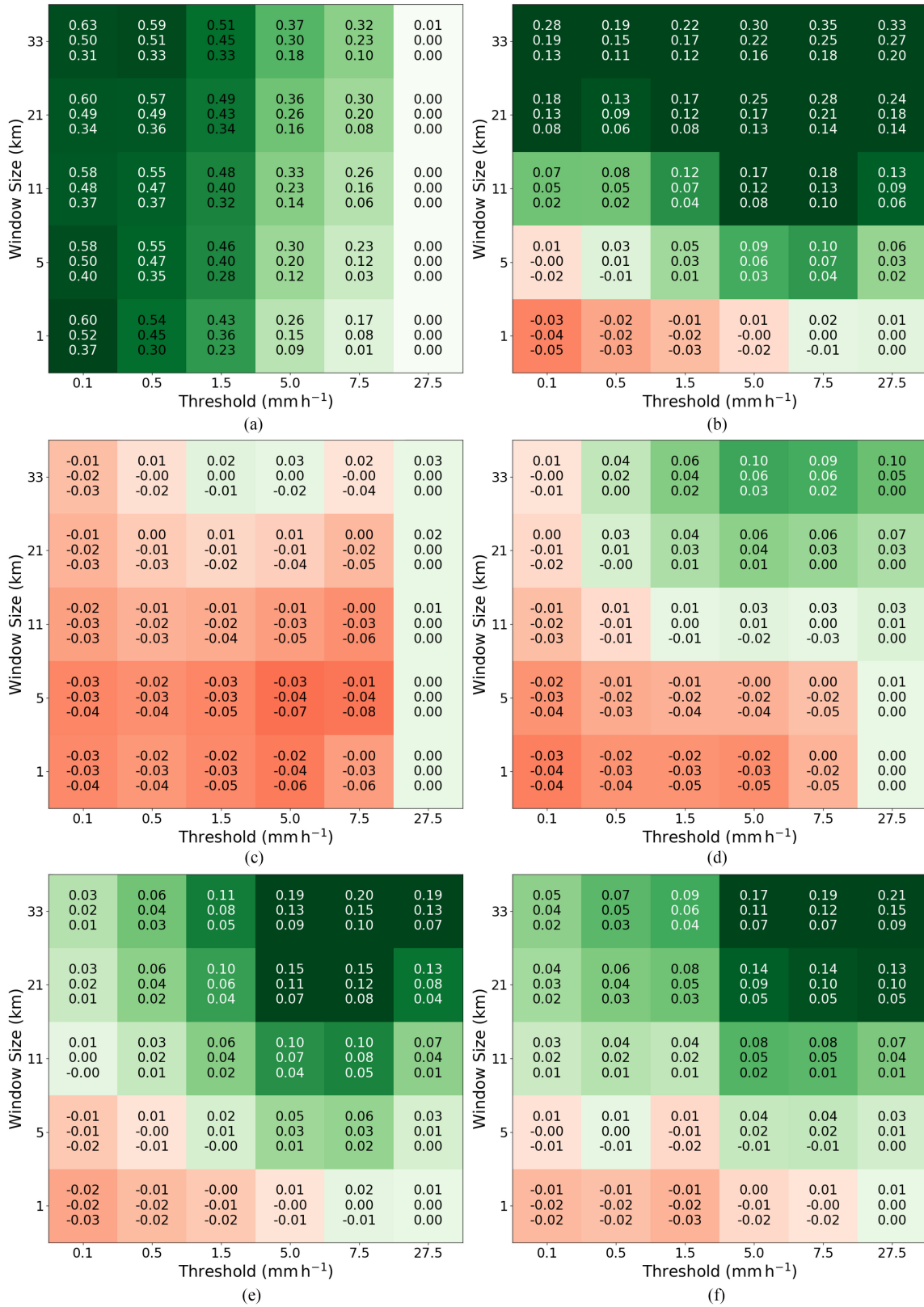
Fig. 11. Heatmaps representing (a) the median CSI of SPROG with μ-adjustment as basis for (b)–(f) the median containing differences in CSI of all forecasts (Q2(ΔCSI)) versus CSI(SPROG-LOC) and CSI (advection + ARI), respectively, at 60 min lead-time as a function of P_{th} threshold (x-axis) and window size (y-axis). Greenish colors indicate higher scores compared to SPROG. (b) Advection + ARI(2,1,6), (c) SPROG-LOC(1,6) with $I = 10\%$, (d) $I = 20\%$, (e) and (f) as in (d) but with μ-adjustment and CDF-matching, respectively. The three values in each tile indicate the Q1, Q2, and Q3 quartile of the aforementioned differences in CSI in ascending order. (a) SPROG(1,6), μ-adjustment. (b) Advection + ARI(2,1,6). (c) SPROG-LOC(1,6), $I = 10\%$. (d) SPROG-LOC(1,6), $I = 20\%$. (e) SPROG-LOC(1,6), $I = 20\%$, μ-adjustment. (f) SPROG-LOC(1,6), $I = 20\%$, CDF-matching.
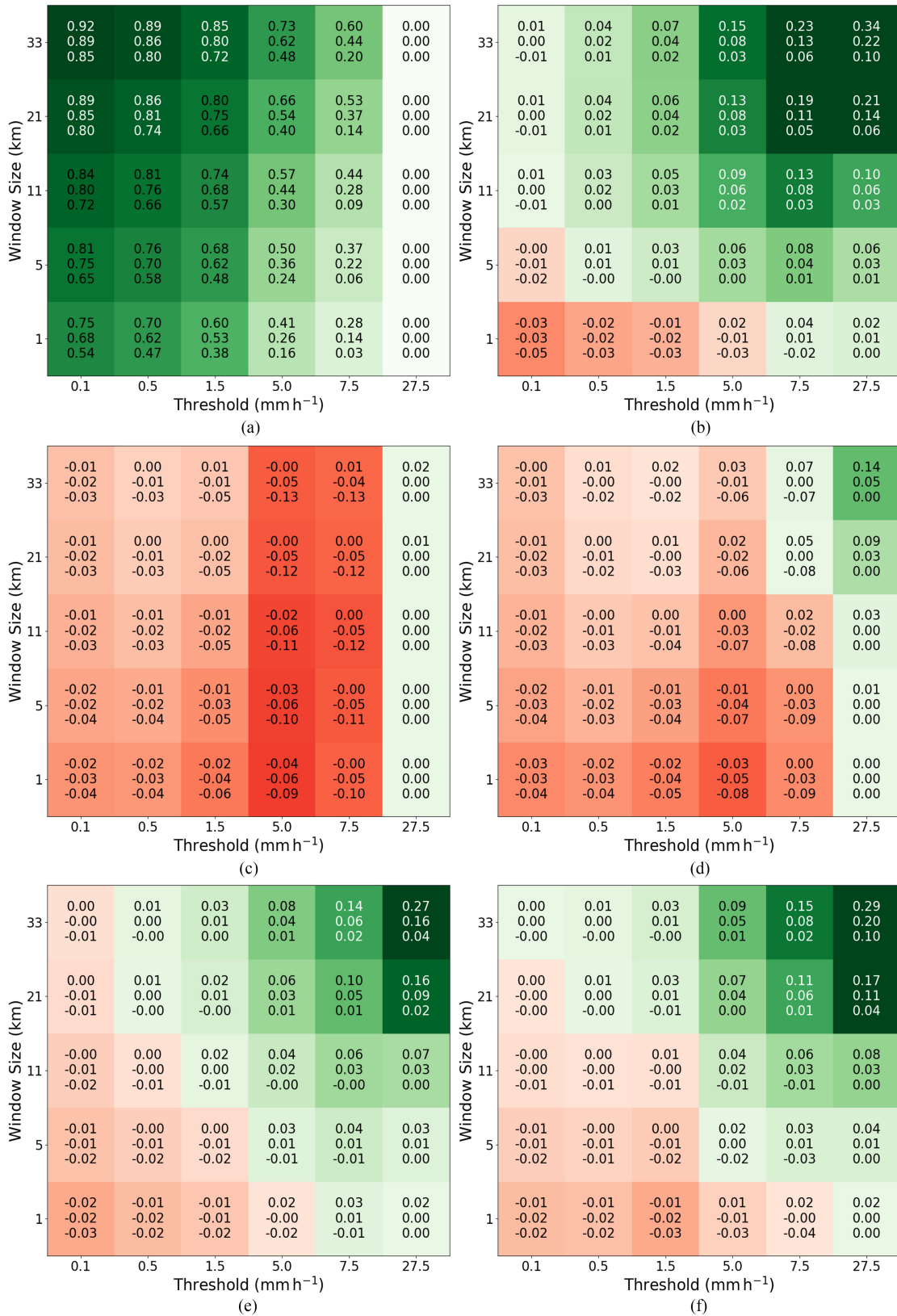
Fig. 12.   As in Fig. 11 but for Q2(ΔFSS). (a) SPROG(1,6), μ-adjustment. (b) Advection + ARI(2,1,6). (c) SPROG-LOC(1,6), $I = 10\%$. (d) SPROG-LOC(1,6), $I = 20\%$. (e) SPROG-LOC(1,6), $I = 20\%$, μ-adjustment. (f) SPROG-LOC(1,6), $I = 20\%$, CDF-matching.

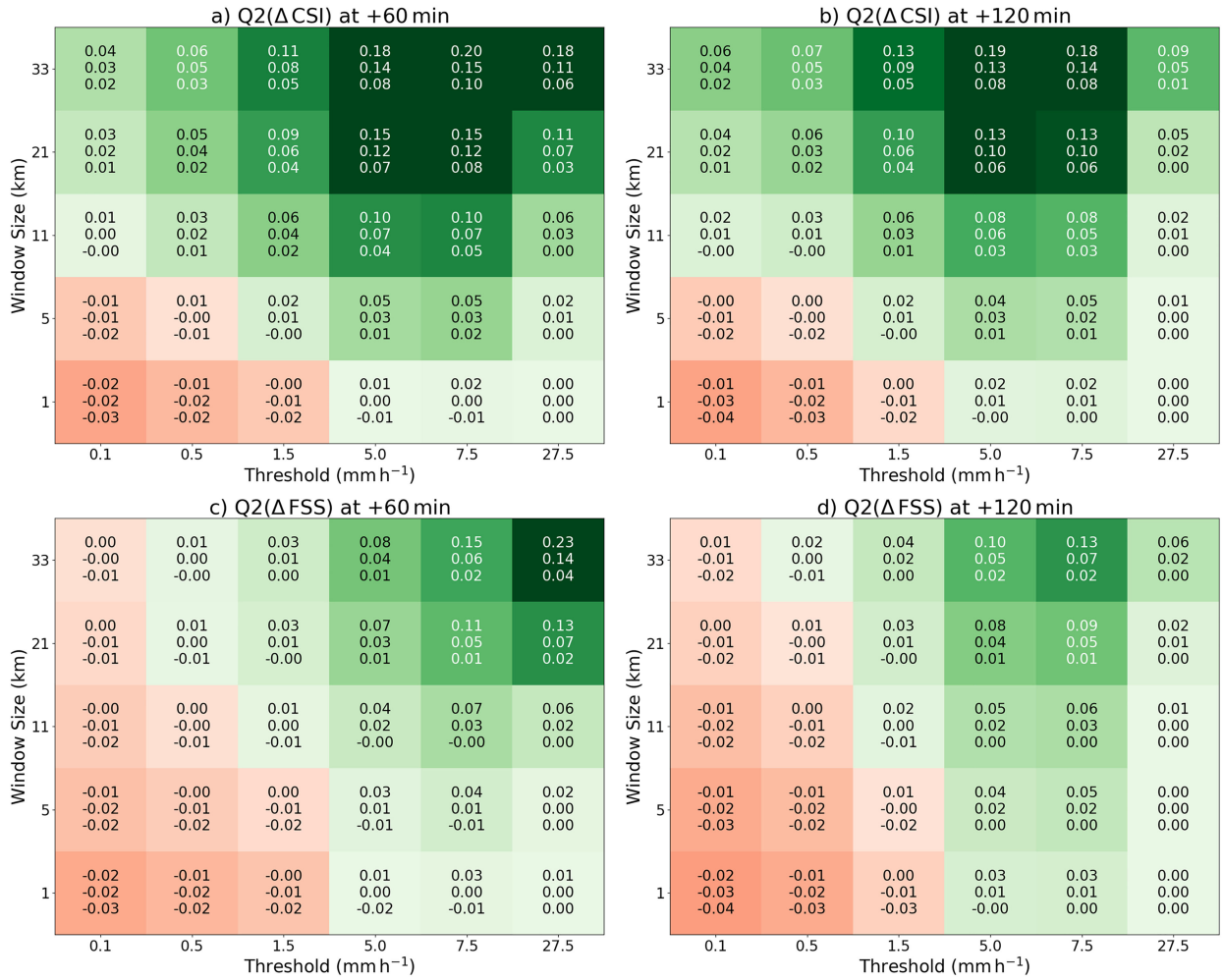Q2(ΔCSI) and Q2(ΔFSS) at SPROG-LOC(1,6), $I = 20\%$, μ-adjustment

Fig. 13. Top: Q2(ΔCSI) heatmaps at (a) 60 min and (b) 120 min lead-time. Bottom: Q2(ΔFSS) heatmaps at (c) 60 and (d) 120 min lead-time. SPROG-LOC(1,6) with $I = 20\%$ and μ-adjustment. (a) Q2(ΔCSI) at +60 min. (b) Q2(ΔCSI) at +120 min. (c) Q2(ΔFSS) at +60 min. (d) Q2(ΔFSS) at +120 min.

nowcasting fields. To summarize the statistical evaluation and to include these at 120 min lead-time, Fig. 13 illustrates analogous heatmaps at a suggested configuration of SPROG-LOC with $I = 20\%$ and μ-adjustment. One can see that at 120 min lead-time the skill from SPROG-LOC outperforms that of SPROG for values of P_{th} and window size as low as 5.0 mm h^{-1} and 1 km, respectively. These results highlight the scenario in which the localization extension of the SPROG model is most beneficial. Moreover, such extension was seen without any significant burden on the computational processes of the distributed AR coefficients because of the convolutional scheme explained in Section II.

VI. CONCLUSION

Radar-based precipitation nowcasting methods have offered the capability of prediction at high temporal and spatial resolutions with increasing lead-time, providing support to, among others, decision makers in the hydrological and meteorological communities. One of these predicting methods is the SPROG

model which uses spatio-temporal properties of rain, and thereby improves the skill of the nowcast compared to that from classic advection techniques. However, the SPROG model tends to smooth the nowcasting fields especially in small but convective precipitation areas and at longer lead-times. In this work, we explored and presented 1) a model referred to as the SPROG-LOC to improve the skill of the SPROG model by estimating localized 2-D parameters of the AR process, 2) a sensitivity analysis of the SPROG model to its parameters, such as the order of the AR process and the spatial decomposition levels, and 3) the degree of dependency to postprocessing. For these purposes, the performance of the SPROG-LOC model was demonstrated and evaluated at several precipitation thresholds using 10 precipitation events collected by the DWD radar network.

The analysis on the AR order and the spatial decomposition levels of SPROG suggested a first order AR process, AR(1), and six cascade levels. This configuration allows SPROG to perform equal to or better than the advection model at precipitation thresholds up to 5.0 mm h^{-1} while maintaining a reasonable smoothing degree and computational burden compared with

other possible combinations. More specifically, it was observed that the nowcasting fields related to an AR(2) led to smoother fields than those from an AR(1) process because values of at least $5.0 \, \text{mm} \, \text{h}^{-1}$ were more difficult to maintain. We also showed that one can reduce the loss of these values by increasing the number of cascade levels although at a cost of increasing computational time, or by including CDF-matching, however, with an inherent penalty in terms of normalized RMSE. For the CDF-matching, the extrapolated nowcasting fields are redistributed to match the CDF of the latest observed precipitation field. This redistribution could be observed by the radially averaged power spectral density in which precipitation components at short wavelengths were corrected much stronger than those at longer wavelengths. However, the SPROG fields with CDF-matching revealed that such enhancement was not necessarily occurring in isolated areas.

The SPROG-LOC model is based on the localization of the AR parameters, through a running kernel, to consider the nonhomogeneity of precipitation while controlling its temporal evolution (i.e., the spatial variability of predictability). That is slowing down the decorrelation in areas, where spatial gradients of observed precipitation are significant. Overall, the statistical results exhibited an improved performance of SPROG-LOC over SPROG when both are evaluated at precipitation thresholds and window sizes larger than $0.1 \, \text{mm} \, \text{h}^{-1}$ and 1 km and for lead-times up to 2 h. Besides the recommended AR(1) process and six decomposition levels, this study found a suitable configuration of SPROG-LOC that mainly consists of a minimum kernel size of 100 km, an adjusting factor of 20%, and the μ-adjustment. Furthermore, the SPROG-LOC model also included a stochastic term (STEPS-LOC) that simulates the uncertainty of precipitation during extrapolation by multiple realizations. The results indicated that the STEPS-LOC model can increase the lead-time by 20 min, with respect to STEPS, to keep an FSS value equal to 0.5 at a precipitation threshold of $5.0 \, \text{mm} \, \text{h}^{-1}$ and window size of 21 km. Thus, the ensemble members of STEPS-LOC are capable to lessen the penalty from the AR spatial filtering process by adding localized stochastic perturbations at high spatial resolution to consider the uncertainty induced by growth and decay processes of precipitation during the extrapolation.

Although SPROG-LOC showed improvements in better modeling the evolution of precipitation while addressing, to some extent, the spatial heterogeneity of precipitation, it is still a filtering-based approach (i.e., a conservative approach) and it is driven by user-defined values. Further improvements can be done when estimating the AR parameters, for example, by formulating the scale-dependent optimization factor as a spatial-distributed factor through an adaptive weighting obtained from the observed precipitation field. Another approach to improve the nowcasting skill is by including the well known moving-average model to consider previous errors resulting from the observations and the nowcasting fields. External processes such as the diurnal cycle of precipitation can also enhance the balance between the filtering attribute of SPROG-LOC and the possible growing tendency of precipitation areas. Regarding STEPS-LOC, further research and depth evaluation on the generation of the perturbation term is essential, for instance, by considering a localized-variance model for the stochastic field. Nevertheless,

we can say that the added value of the localization filtering approach is key in scenarios, where observed events encompass large precipitation areas that are nonuniformly distributed having either an embedded or isolated convective feature.

A comparison of SPROG-LOC with advection+ARI showed that the latter technique is able to outperform the advection nowcasting and further on SPROG-LOC. Despite this, the SPROG-LOC model offers a rather simple 2-D AR model with the capability of maintaining precipitation areas at small scale for lead-times longer than those resulting from the SPROG model. Moreover, it is expected that the generation of nowcasting ensemble members, obtained by the STEPS-LOC model at national scales, will be operational feasible because the 2-D AR parameter that scales the stochastic noise field could be obtained directly from the optimized 2-D correlation coefficients, as shown in Section II-C. At DWD, this will become highly advantageous for an ongoing seamless combination approach among ensemble members from the STEPS-LOC model and those resulting from the short-range NWP model known as the Icosahedral Nonhydrostatic Rapid Updated Cycle (ICON-D2-RUC) model. Yet, it is also of interest to the forecasting community to extend the advection+ARI model to include the stochastic perturbation term during extrapolation and investigate its operational performance.

This study evaluated the capabilities of a combined localization-filtering approach, which can serve for the generation of real-time nowcasting fields in Germany or in countries having a similar or larger radar network coverage. Moreover, it is foreseen that this work can be extended to include the generation of radar-based nowcasting ensemble members and, at a later stage, can be combined with NWP models, thereby enhancing the skill of a nationwide forecasting system.

REFERENCES

[1] N. Dotzek, P. Groenemeijer, B. Feuerstein, and A. M. Holzer, "Overview of ESSL's severe convective storms research using the European severe weather database ESWD," *Atmospheric Res.*, vol. 93, no. 1–3, pp. 575–586, 2009.

[2] L. Alfieri, P. Salamon, F. Pappenberger, F. Wetterhall, and J. Thielen, "Operational early warning systems for water-related hazards in Europe," *Environ. Sci. Policy*, vol. 21, pp. 35–49, 2012.

[3] D. Piper et al., "Exceptional sequence of severe thunderstorms and related flash floods in May and June. 2016 in Germany-Part 1: Meteorological background," *Natural Hazards Earth Syst. Sci.*, vol. 16, no. 12, 2016, Art. no. 2835.

[4] M. Berenguer, C. Corral, R. Sánchez-Diezma, and D. Sempere-Torres, "Hydrological validation of a radar-based nowcasting technique," *J. Hydrometeorol*, vol. 6, no. 4, pp. 532–549, 2005.

[5] F. Schmid, Y. Wang, and A. Harou, "Nowcasting guidelines—A summary," *WMO—No. 1198*, vol. 68, no. 2, pp. 63–68, 2019.

[6] P. Bauer, A. Thorpe, and G. Brunet, "The quiet revolution of numerical weather prediction," *Nature*, vol. 525, no. 7567, pp. 47–55, 2015.

[7] J. Sun *et al.*, "Use of NWP for nowcasting convective precipitation: Recent progress and challenges," *Bull. Amer. Meteorol. Soc.*, vol. 95, no. 3, pp. 409–426, 2014.

[8] K. Stephan, S. Klink, and C. Schraff, "Assimilation of radar-derived rain rates into the convective-scale model COSMO-DE at DWD," *Quart. J. Roy. Meteorol. Soc., J. Atmos. Sci., Appl. Meteorol. Phys. Oceanogr.*, vol. 134, no. 634, pp. 1315–1326, 2008.

[9] K. Kober, G. C. Craig, C. Keil, and A. Dörnbrack, "Blending a probabilistic nowcasting method with a high-resolution numerical weather prediction ensemble for convective precipitation forecasts," *Quart. J. Roy. Meteorol. Soc.*, vol. 138, no. 664, pp. 755–768, 2012.

[10] U. Germann and I. Zawadzki, "Scale-dependence of the predictability of precipitation from continental radar images. Part I: Description of the methodology," *Monthly Weather Rev.*, vol. 130, pp. 2859–2873, 2002.

[11] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. DARPA Imag. Understanding Workshop*, 1981, pp. 121–130.

[12] M. Proesmans, L. van Gool, E. Pauwels, and A. Oosterlinck, "Determination of optical flow and its discontinuities using non-linear diffusion," in *Proc. J.-O Eklundh*, 1994, vol. 801, pp. 294–304.

[13] M. Werner and M. Cranston, "Understanding the value of radar rainfall nowcasts in flood forecasting and warning in flashy catchments," *Meteorol. Appl., J. Forecasting, Practical Appl., Training Techn. Modelling*, vol. 16, no. 1, pp. 41–55, 2009.

[14] S. Ochoa-Rodriguez *et al.*, "Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation," *J. Hydrol.*, vol. 531, pp. 389–407, 2015.

[15] T. Poméon, N. Wagner, C. Furusho, S. Kollet, and R. Reinoso-Rondinel, "Performance of a PDE-based hydrologic model in a flash flood modeling framework in sparsely-gauged catchments," *Water*, vol. 12, no. 8, pp. 2157–2178, 2020. [Online]. Available: https://www.mdpi.com/2073-4441/12/8/2157

[16] D. Heuvelink, M. Berenguer, C. C. Brauer, and R. Uijlenhoet, "Hydrological application of radar rainfall nowcasting in the Netherlands," *Environ. Int.*, vol. 136, 2020, Art. no. 105431.

[17] R. Reinoso-Rondinel, R. Evaristo, M. Schmidt, F. Crijnen, S. Trömel, and C. Simmer, "Storm cell observation and prediction using polarimetric weather radars," in *Proc. 21st IEEE Int. Radar Symp.*, Germany, Jun. 2021, pp. 1–7. [Online]. Available: https://ieeexplore.ieee.org/document/9466168

[18] S. Trömel *et al.*, "Near-realtime quantitative precipitation estimation and prediction," *Bull. Amer. Meteorol. Soc.*, vol. 35, pp. 1–13, 2021.

[19] A. Atencia, I. Zawadzki, and M. Berenguer, "Scale characterization and correction of diurnal cycle errors in MAPLE," *J. Appl. Meteorol. Climatol.*, vol. 56, no. 9, pp. 2561–2575, 2017.

[20] L. Foresti, I. V. Sideris, D. Nerini, L. Beusch, and U. Germann, "Using a 10-Year radar archive for nowcasting precipitation growth and decay: A probabilistic machine learning approach," *Weather Forecasting*, vol. 34, no. 5, pp. 1547–1569, 2019.

[21] U. Hamann *et al.*, "Nowcasting of thunderstorm severity with machine learning in the alpine region," Agencia Estatal de Meteorología, 2019.

[22] C. K. Sønderby *et al.*, "Metnet: A neural weather model for precipitation forecasting," 2020, *arXiv:2003.12140*. [Online]. Available: https://arxiv.org/abs/2003.12140

[23] P. Schaumann, R. Hess, M. Rempel, U. Blahak, and V. Schmidt, "A calibrated and consistent combination of probabilistic forecasts for the exceedance of several precipitation thresholds using neural networks," *Weather Forecasting*, vol. 36, no. 3, pp. 1079–1096, 2021.

[24] T. M. Hamill *et al.*, "Noaa's second-generation global medium-range ensemble reforecast dataset," *Bull. Amer. Meteorol. Soc.*, vol. 94, no. 10, pp. 1553–1565, 2013.

[25] M. Scheuerer and T. M. Hamill, "Statistical postprocessing of ensemble precipitation forecasts by fitting censored, shifted gamma distributions," *Monthly Weather Rev.*, vol. 143, no. 11, pp. 4578–4596, 2015.

[26] M. Mathieu, C. Couprie, and Y. LeCun, "Deep multi-scale video prediction beyond mean square error," 2016, *arXiv:1511.05440*.

[27] R. Prudden *et al.*, "A review of radar-based nowcasting of precipitation and applicable machine learning techniques," 2020, *arXiv:2005.04988*.

[28] N. E. Bowler, C. E. Pierce, and A. W. Seed, "STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP," *Quart. J. Roy. Meteorol. Soc.*, vol. 132, pp. 2127–2155, 2006.

[29] V. Venugopal, E. Foufoula-Georgiou, and V. Sapozhnikov, "Evidence of dynamic scaling in space-time rainfall," *J. Geophys. Res., Atmos.*, vol. 104, no. D24, pp. 31599–31610, 1999.

[30] A. Seed, "A dynamic and spatial scaling approach to advection forecasting," *J. Appl. Meteorol.*, vol. 42, pp. 381–388, 2003.

[31] S. Lovejoy and D. Schertzer, "Multifractals, cloud radiances and rain," *J. Hydrol.*, vol. 322, nos. 1–4, pp. 59–88, 2006.

[32] L. Foresti and A. Seed, "The effect of flow and orography on the spatial distribution of the very short-term predictability of rainfall from composite radar images," *Hydrol. Earth Syst. Sci.*, vol. 18, no. 11, 2014, Art. no. 4671.

[33] S. Pulkkinen, V. Chandrasekar, and A.-M. Harri, "Nowcasting of precipitation in the high-resolution dallas-fort worth (DFW) urban radar remote sensing network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 8, pp. 2773–2787, Aug. 2018.

[34] D. Nerini, N. Besic, I. Sideris, U. Germann, and L. Foresti, "A non-stationary stochastic ensemble generator for radar rainfall fields based on the short-space Fourier transform," *Hydrol. Earth Syst. Sci.*, vol. 21, pp. 2777–2797, 2017.

[35] L. Foresti, M. Reyniers, A. Seed, and L. Delobbe, "Development and verification of a real-time stochastic precipitation nowcasting system for urban hydrology in Belgium," *Hydrol. Earth Syst. Sci.*, vol. 20, pp. 505–527, 2016.

[36] S. Pulkkinen *et al.*, "Pysteps: An open-source Python library for probabilistic precipitation nowcasting (v1. 0)," *Geoscientific Model Develop.*, vol. 12, no. 10, pp. 4185–4219, 2019.

[37] S. Pulkkinen, V. Chandrasekar, A. von Lerber, and A.-M. Harri, "Nowcasting of convective rainfall using volumetric radar observations," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 58, no. 11, pp. 7845–7859, Nov. 2020.

[38] E. Ruzanski, V. Chandrasekar, and Y. Wang, "The CASA nowcasting system," *J. Atmospheric Ocean. Technol.*, vol. 28, pp. 640–655, 2011.

[39] C. Zach, T. Pock, and H. Bischof, "A duality based approach for real-time TV-L^1 optical flow," in *Proc. Joint Pattern Recognit. Symp.*, 2007, pp. 214–223.

[40] J. Sánchez, E. Meinhardt-Llopis, and G. Facciolo, "TV-L^1 optical flow estimation," *Image Process. Line*, vol. 2013, pp. 137–150, 2013.

[41] A. Seed, C. Pierce, and K. Norman, "Formulation and evaluation of a scale decomposition-based stochastic precipitation nowcast scheme," *Water Resour. Res.*, vol. 49, pp. 6624–6641, 2013.

[42] D. S. Wilks, *Statistical Methods in the Atmospheric Sciences*. New York, NY, USA: Academic, 2011.

[43] R. F. Ling, "Comparison of several algorithms for computing sample means and variances," *J. Amer. Stat. Assoc.*, vol. 69, no. 348, pp. 859–866, 1974.

[44] T. F. Chan, G. H. Golub, and R. J. LeVeque, "Algorithms for computing the sample variance: Analysis and recommendations," *Amer. Statistician*, vol. 37, no. 3, pp. 242–247, 1983.

[45] E. E. Ebert, "Ability of a poor man's ensemble to predict the probability and distribution of precipitation," *Monthly Weather Rev.*, vol. 129, no. 10, pp. 2461–2480, 2001.

[46] H. Bartels *et al.*, "RADOLAN routineverfahren zur online-aneichung der radarniederschlagsdaten mit hilfe von automatischen bodenniederschalagsstationen," 2004. [Online]. Available: https://www.dwd.de/DE/leistungen/radolan/radolan_info/abschlussbericht_pdf.pdf?__blob=publicationFile&v=2

[47] E. Weigl and T. Winterrath, "Radargestützte niederschlagsanalyse und-vorhersage (RADOLAN, RADVOR-OP)," *Promet*, vol. 35, pp. 78–86, 2009.

[48] M. Werner, "A new radar data post-processing quality control workflow for the DWD weather radar network," in *Proc. 8th Eur. Conf. Radar Meteorol. Hydrol. ERAD*, Garmisch-Partenkirchen, Germany, Sep. 2014, pp. 1–9.

[49] V. Pejcic, P. S. Garfias, K. Mühlbauer, S. Trömel, and C. Simmer, "Comparison between precipitation estimates of ground-based weather radar composites and GPM's DPR rainfall product over Germany," *Meteorologische Zeitschrift*, vol. 29, no. 6, pp. 451–466, 2020.

[50] S. Theis, A. Hense, and U. Damrath, "Probabilistic precipitation forecasts from a deterministic model: A pragmatic approach," *Meteorological Appl.: J. Forecasting, Practical Appl., Training Techn. Modelling*, vol. 12, no. 3, pp. 257–268, 2005.

[51] C. S. Schwartz *et al.*, "Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership," *Weather Forecasting*, vol. 25, no. 1, pp. 263–280, 2010.

[52] I. T. Jolliffe and D. B. Stephenson, *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. New York, NY, USA: Wiley, 2012.

[53] N. M. Roberts and H. W. Lean, "Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events," *Monthly Weather Rev.*, vol. 136, no. 1, pp. 78–97, 2008.

[54] G. Ayzel, M. Heistermann, and T. Winterrath, "Optical flow models as an open benchmark for radar-based precipitation nowcasting (rainymotion v0.1) Geosci," *Geoscientific Model Develop.*, vol. 12, pp. 1387–1402, 2019.

Markus Schultze received the B.Sc. and M.Sc. degrees in meteorology from Hamburg University, Hamburg, Germany, in 2011 and 2013, respectively.

From 2014 to 2017, he worked in the regional atmospheric modelling group with the Institute of Coastal Research of the Helmholtz Centre, Geesthacht, Germany. Since 2017, he has joint the research and development department of Deutscher Wetterdienst, Offenbach, to conduct research on the topic of probabilistic precipitation nowcasting. His main research interests include the application of dual-polarimetric radar data and the radar-based precipitation nowcasting.

Ricardo Reinoso-Rondinel received the B.Sc. degree in electrical engineering from the Pontifical Catholic University of Peru, Lima, in 2006, the M.Sc. degree in electrical and computer engineering, with the Atmospheric Radar Research Center (ARRC) on phased-array radar studies from University of Oklahoma, Norman, OU, USA, in 2011, and the Ph.D. degree on the research topic of differential phase processing for polarimetric radars from the Delft University of Technology of Mekelweg, CD Delft, Netherlands, in 2018.

In 2018, he started his first one-year post-doctoral research on Doppler spectra processing of vertically pointing radars at the Department of Geosciences and Remote Sensing of the Delft University of Technology (TU Delft), The Netherlands. Since 2019, he has joined the University of Bonn, Germany, to conduct research on the topic of radar-based quantitative precipitation nowcasting. His research interests include weather radar signal processing, phased-array radar, radar polarimetry, and radar- and satellite-based observation of severe weather.

Silke Trömel received the diploma and the PhD degree in meteorology from the Goethe University in Frankfurt, Germany, in 1999 and 2005, respectively.

After that she joined the Meteorological Institute, University of Bonn, Bonn, Germany, first as Coordinator and Research Scientists of the DFG project bundle AQUARadar (Advances in Quantitative Areal Precipitation Estimation by Radar), later as chair of the research group OASE (Object-based Analysis and SEamless prediction) in the Hans Ertel Centre for Weather Research programme (HErZ). In 2016, she received her Venia Legendi in meteorology from the Faculty of Mathematics and Natural Sciences, University of Bonn. She is a Lecturer and Senior Scientist with the Department of Meteorology, Institute for Geosciences, University of Bonn, and is the Spokesperson of the DFG priority programme "Polarimetric Radar Observations meet Atmospheric Modelling" (PROM) and the research unit "Near-Realtime Quantitative Precipitation Estimation and Prediction" (RealPEP). She is a Director of the competence centre CPEX-LAP (Clouds and Precipitation Exploration Laboratory) of the Geoverbund ABC/J. Her research interests are radar polarimetry and its applications including quantitative precipitation estimation, nowcasting, model evaluation, and data assimilation.

Martin Rempel received the B.S. and M.S. degrees in meteorology from Leipzig University, Leipzig, Germany, in 2013 and 2015, respectively.

From 2015 to 2016, he worked in the satellite remote sensing group of the Leibniz Institute for Tropospheric Research, Leipzig, Germany. Since 2017, he has been working as a Scientific Assistant in the research and development department of Deutscher Wetterdienst, Offenbach, Germany. His main research interests include the seamless combination of multiple forecasts, the improvement of nowcast accuracy and the forecast verification with remote sensing data.