A Reduced-Space Ensemble Kalman Filter Approach for Flow-Dependent Integration of Radar Extrapolation Nowcasts and NWP Precipitation Ensembles

DANIELE NERINI

*Federal Office of Meteorology and Climatology MeteoSwiss, Locarno-Monti,
and Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland*

LORIS FORESTI AND DANIEL LEUENBERGER

Federal Office of Meteorology and Climatology MeteoSwiss, Locarno-Monti, Switzerland

SYLVAIN ROBERT

Swiss Re, Zurich, Switzerland

URS GERMANN

Federal Office of Meteorology and Climatology MeteoSwiss, Locarno-Monti, Switzerland

(Manuscript received 24 July 2018, in final form 11 December 2018)

ABSTRACT

A Bayesian precipitation nowcasting system based on the ensemble Kalman filter is formulated. Starting from the last available radar observations, the prediction step of the filter consists of a stochastic radar extrapolation technique, while the correction step updates the radar extrapolation nowcast using information from the most recent forecast by the numerical weather prediction model (NWP). The result is a flow-dependent and seamless blending scheme that is based on the spread of the nowcast and NWP ensembles, used as the definition of the forecast error. To simplify the matrix operations, the Bayesian update is performed in the subspace spanned by the principal components, hence the term *reduced space*. Synthetic data experiments demonstrated that the Bayesian nowcast correctly captures the flow dependency in both the NWP forecast and the radar extrapolation skills. Four experiments with real precipitation data and a relatively small ensemble size (21 members) represented a first test under realistic conditions, such as stratiform wintertime precipitation and localized summertime convection. The skill was quantified in terms of fractions skill score at 32-km scale and 2.0 mm h^{-1} intensity. The results indicate that the system is able to produce blended forecasts that are at least as skillful as the nowcast-only or the NWP-only forecasts at any lead time.

1. Introduction

Very short-term forecasts up to a lead time of about 6 h have become an important tool for issuing high-impact weather warnings. Commonly referred to as *nowcasting*, such forecasts cover a time range where the persistence of the most recent observations can produce

significant skill, while the performance of numerical weather prediction (NWP) can suffer from outdated initial conditions, spinup issues, and model approximations. In this sense, nowcasting can be seen as a post-processing step aimed at updating NWP forecasts with real-time observations. The term *seamless* is often used in this context to indicate a consistent prediction regardless of location, lead time or forecasting procedure (World Meteorological Organization 2015). To obtain a seamless forecast that does not degrade the quality of its individual components, the combination should be based on an objective procedure that explicitly accounts for the respective forecast uncertainties.

Denotes content that is immediately available upon publication as open access.

Corresponding author: Daniele Nerini, daniele.nerini@meteoswiss.ch

DOI: 10.1175/MWR-D-18-0258.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) (www.ametsoc.org/PUBSReuseLicenses).

In the context of quantitative precipitation forecasting, nowcasting systems are typically based on observations from a network of weather radars. The radar quantitative precipitation estimation (QPE) consists of a data processing chain that transforms raw polar-coordinate measurements of backscattered signal from volumes of hydrometeors at a certain height in the atmosphere into estimates of precipitation intensity at the ground. For hydrological applications, the radar QPE is usually presented as a sequence of fields on a regular Cartesian grid at 1–4-km and 1–6-min resolutions, down to 100 m and 1 min in case of urban catchments (Thorndahl et al. 2017). By assuming persistence of precipitation patterns in moving coordinates, known as the *Lagrangian persistence* assumption (Zawadzki et al. 1994), the radar observations can be used to produce extrapolation forecasts whose skill rapidly decreases with lead time as the persistence assumption weakens. Because the lifetime of precipitation relates to its spatial scale (Venugopal et al. 1999; Seed 2003), predictability in the Lagrangian persistence approach can range from several hours at the scale of hundreds of kilometers down to time scales less than 1 h when considering the extent of a localized thunderstorm (e.g., Germann et al. 2006; Foresti and Seed 2014).

Thanks to the improvements in horizontal resolution that have occurred in the past decades, NWP precipitation forecasts are increasingly able to meet nowcasting requirements in terms of skill at the convective scale. Sun et al. (2014) give an overview of the achievements and challenges in this field. While diabatic initialization techniques offer an efficient way to assimilate radar reflectivities, reduce precipitation spinup problems, and increase the NWP skill in the nowcasting range (e.g., Jacques et al. 2018), their success is limited by the capacity to simulate and support the observed rainfall intensities. Stephan et al. (2008) link such problems to imperfections in the physics parameterization and in the accuracy of the large-scale environment conditions provided by the NWP. The direct assimilation of radar volume scans is also possible (e.g., Johnson et al. 2015), but the approach is particularly challenging owing to the strong nonlinear relationship between radar observations and model state (Tong and Xue 2005; Dowell et al. 2011).

Beyond the challenges related to the assimilation of radar observations, the computational cost of numerical simulations limits the update frequency of NWP runs and hence their ability to exploit real-time radar observations for initializing a new forecast. In fact, only a few national weather services can currently afford to operationally run rapid update cycles, and this usually applies to deterministic runs only. Hence, today, the

potential of high-spatiotemporal-resolution measurements from weather radars is largely untapped in NWP schemes.

Precipitation nowcasting systems have traditionally relied on deterministic blending schemes, in most cases using a simple linear-weighting function in time to produce a smooth transition from the radar extrapolation to a deterministic NWP forecast. A first example of an operational scheme is the NIMROD system at the Met Office (Golding 1998). In this approach, weights are typically precomputed based on long-term forecast error statistics and then applied in real time. For example, the blending system currently in operations at MeteoSwiss (INCA; Haiden et al. 2011) performs the blending by starting to give weight to the NWP forecast at +1 h and then linearly increasing the weight given to the NWP model until +4 h. The technique is simple and robust, which explains its appeal for operational use. On the other hand, it cannot account for variability in both the NWP and the radar extrapolation skills, resulting in a suboptimal blended forecast. Moreover, linear blending generally produces weights that are only a function of lead time, ignoring other dependencies such as location or spatial scales.

More advanced nowcasting blending schemes have thus been developed in recent years: for example, the Short-Term Ensemble Prediction System (STEPS; Bowler et al. 2006; Seed et al. 2013). STEPS quantifies the skill of the NWP forecast in real time and uses the information to adjust the weights that combine the radar-based extrapolation and NWP forecast as a function of lead time and spatial scale. As the actual forecast skill is unknown, the blending weights at analysis time are gradually relaxed to climatological skill values.

Following the success encountered in the NWP community (Bauer et al. 2015), a number of nowcasting systems have introduced the ensemble approach by means of stochastic perturbations as a mean to quantify the forecast uncertainty (e.g., Bowler et al. 2006; Berenguer et al. 2011; Atencia and Zawadzki 2014). The definition of forecast error as the spread of the ensemble members represents an uncertainty estimate that can be readily propagated into several applications—for example, hydrological forecasting (Zappa et al. 2010)—and ultimately used for informed decision-making (Todini 2018). The same nowcast error models can be also used to find the optimal combination of radar extrapolation and NWP forecasts. In fact, Bayes's theorem represents an attractive mathematical framework for the combination of information in the presence of uncertainty. In essence, one needs to represent the two precipitation forecasts as probability density functions to construct the prior and likelihood distributions and

thus compute their posterior distribution. Bayesian filtering refers to the Bayesian formulation of optimal filtering, which seeks to estimate the state of a time-varying system based on observations. For this class of methods, “optimal” is intended in the sense of a minimum mean-squared error, thus corresponding to the least squares solution, while “filtering” indicates the estimation of the current state given past observations.

The widely used Kalman filter (KF; Kalman 1960) algorithm is a recursive solution to the linear Gaussian optimal filtering problem that can be formulated from a purely Bayesian point of view, that is, to produce a solution for the posterior Gaussian distribution (e.g., Särkkä 2013). In this sense, the KF becomes a powerful tool to implement the Bayesian update for a time-varying system under the assumption of a linear Gaussian model. To relax the linear assumption and in order to allow its application to high-dimensional state vectors, Evensen (1994) introduced the ensemble Kalman filter (EnKF), which is essentially a Monte Carlo implementation of the KF equations that allows us to sample the error distribution without the need to fully represent it.

It should be noted that the use of precipitation data in an EnKF system is not straightforward. This can be related to the presence of non-Gaussian error statistics and the limited ensemble size that is available in most applications. An important question of this study is therefore whether a Bayesian approach based on the EnKF can produce reasonable results despite such challenges.

The main contribution of this study is the formulation of the precipitation nowcasting and blending with NWP in a Bayesian framework through the use of an ensemble Kalman filter. The idea is to apply the EnKF from a nowcasting perspective: the prediction step is computed as an extrapolation nowcast, while the correction step uses NWP precipitation forecasts as pseudomeasurements. Thus, the usual application of EnKF in data assimilation is reversed, as observations represent the state to be updated with NWP pseudomeasurements.

Following a similar approach, Buil (2017) investigated the use of a local ensemble transform Kalman filter (LETKF) to improve ensemble precipitation nowcasts produced with the SBMcst model (Berenguer et al. 2011). Our work generalizes the same concept to the use of NWP ensembles and introduces two technical solutions that are necessary to deal with high-resolution precipitation fields. These include the use of principal component analysis (PCA) to perform a reduced-space EnKF and a resampling technique addressing the non-Gaussianity of precipitation fields. In formulating this approach, we have tried to present a clean mathematical

development that can be generalized to analogous problems and datasets.

The paper is organized as follows. Section 2 formulates the Bayesian nowcasting system based on the reduced-space ensemble Kalman filter. The developed approach is tested using synthetic experiments in section 3 and real precipitation fields in section 4. Section 5 concludes the paper and discusses possible future developments.

2. Formulation of the Bayesian nowcasting technique

Sequential Bayesian filtering consists of three main steps:

- 1) *Initialization step*, where a vector characterizing the initial state of the model and a covariance matrix describing errors in the estimation of the model state are defined.
- 2) *Prediction step*, where a state transition function is used to move the system forward in time, while process noise can be included to account for the model uncertainty.
- 3) *Correction step*, where an observation is used to update the prediction according to the relative importance of the observation error and prediction error in a Bayesian fashion.

The algorithm is iterated between steps 2 and 3. Each of these main steps is now detailed: first in the terms of the original Kalman filter, then by looking at the approximations introduced by the EnKF, and finally by formulating the nowcasting application of this study. Figure 1 summarizes the algorithm and its application in the nowcasting context.

a. Initialization

The true but unknown model state vector $\mathbf{x}$ includes the m model variables that characterize the system of interest. In the Kalman filter, our knowledge of the unknown state vector $\mathbf{x}$ is represented as a multivariate Gaussian distribution $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{P})$, where $\boldsymbol{\mu} \in \mathbb{R}^{m \times 1}$ is a vector of mean state estimates representing the most likely model state, while the model state uncertainty is represented by the covariance matrix $\mathbf{P} \in \mathbb{R}^{m \times m}$.

The EnKF is a Monte Carlo approximation of the Kalman filter designed to address the computational limitations of working with the covariance matrix $\mathbf{P}$ in a high-dimensional space (Evensen 2003). The basic idea is to provide a random sample of the full model state distribution in what is called an *ensemble*. The model state estimate becomes matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]^T \in \mathbb{R}^{n \times m}$, namely, an ensemble of n equally probable realizations

FIG. 1. The flowchart summarizes the basic equations of the Bayesian nowcasting system. Detailed explanation and notation can be found in the text (section 2).

$\mathbf{x}_i \in \mathbb{R}^{m \times 1}$ of the unknown state vector $\mathbf{x}$. Note that in $\mathbf{X}$, the ensemble members are stacked as rows. Following the Gaussian assumption, the best estimate of the model state is approximated with the sample mean $\bar{\mathbf{x}}$ of the n ensemble members, while $\mathbf{P}$ is approximated with the sample covariance matrix:

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i,$$

$$\mathbf{P} = \frac{1}{n-1} (\mathbf{X} - \mathbf{1} \bar{\mathbf{x}}^T)^T (\mathbf{X} - \mathbf{1} \bar{\mathbf{x}}^T), \quad (1)$$

where $\mathbf{1} \in \mathbb{R}^{n \times 1}$ is a vector of all ones. The estimation of $\mathbf{P}$ in (1) stems from one central idea in the EnKF, namely, the dispersion of the ensemble members around their mean is representative of the uncertainty about the system state (Evensen 2003).

In precipitation nowcasting, the state variable of interest is the mean precipitation intensity in millimeters per hour, with time step Δt and over a regular grid with spacing Δx . The state variable $\mathbf{x}_0 \in \mathbb{R}^{m \times 1}$ includes values of precipitation intensity for all m grid points at the start of the forecast. The number of model variables is thus equal to the number of grid points in the two-dimensional precipitation field. The nowcast is initialized with the most recent radar QPE field available at observation time t_0 .

Following the ensemble approach introduced in the EnKF, the uncertainty in the radar estimates can be

included by means of multiple realizations of the QPE. In this way, observation errors are explicitly represented and propagated throughout the forecast. There exists a considerable research effort in the quantification of radar QPE uncertainty with the work of, among others, Ciach et al. (2007), Germann et al. (2009), or more recently Cecinati et al. (2017); a review can be found in Mandapaka and Germann (2010). A quantification of the QPE uncertainty is nevertheless outside the scope of this study, and it is thus assumed to be zero, meaning that all members of the nowcast ensemble are initialized with the same QPE field $\mathbf{X}_0 = [\mathbf{x}_0, \mathbf{x}_0, \dots, \mathbf{x}_0]^T \in \mathbb{R}^{n \times m}$.

b. Prediction

For the time update step of the filter, a prediction model is used to move the state forward in time to produce a first guess of the new state of the system, namely, the prior. We now include a time index and denote a prior estimate with the superscript f for forecast.

In the original KF, a linear prediction model $\mathbf{F} \in \mathbb{R}^{m \times m}$ is used to propagate both moments of the model state distribution $\mathbf{x}$ forward in time:

$$\begin{aligned} \boldsymbol{\mu}_{t+1}^f &= \mathbf{F} \boldsymbol{\mu}_t, \\ \mathbf{P}_{t+1}^f &= \mathbf{F} \mathbf{P}_t \mathbf{F}^T, \end{aligned} \quad (2)$$

where $\boldsymbol{\mu}_{t+1}^f \in \mathbb{R}^{m \times 1}$ and $\mathbf{P}_{t+1}^f \in \mathbb{R}^{m \times m}$ are the two moments of prior state distribution $\mathbf{x}_{t+1}^f$.

In EnKF, a prediction model $\mathcal{F}$, which can be nonlinear, is applied to each individual ensemble member in order to propagate the whole ensemble forward in time:

$$\mathbf{x}_{i,t+1}^f = \mathcal{F}(\mathbf{x}_{i,t}), \quad i = 1, \dots, n. \quad (3)$$

The term $\mathbf{x}_{i,t+1}^f$ is the i th member of the ensemble state $\mathbf{X}_{t+1}^f$ at time $t + 1$. By doing so, the uncertainty of the state estimate is implicitly propagated, since this can be approximated with the prior sample error covariance matrix $\mathbf{P}_{t+1}^f$ as in (1).

In radar-based nowcasting, the prediction is generally computed as a simple extrapolation forecast based on the motion field estimated from a sequence of radar observations by means of optical flow techniques. This nowcasting procedure is based on the assumption of persistence of precipitation patterns in coordinates moving with the storm. For this reason, the extrapolation approach is also referred to as *Lagrangian persistence* forecast, and it can be expressed as

$$\begin{aligned} \mathbf{x}_{i,j,t+1}^f &= x_{i,j+\delta_j,t}, \\ i &= 1, \dots, n \quad \text{and} \quad j = 1, \dots, m. \end{aligned} \quad (4)$$

The term on the left-hand side of (4) represents the i th member of the extrapolated ensemble for grid point j and time $t + 1$, while the right-hand-side term is the previous precipitation intensity at time t , where δ_j is the displacement scalar in index space accounting for the advection of the precipitation field occurred during the time update interval. The motion field δ is estimated from a sequence of radar observations using a Lucas–Kanade local tracking approach as implemented in the open-source nowcasting library pySTEPS (Pulkkinen et al. 2018).

Model errors in the extrapolation forecast are mainly related to the assumption of persistence (Germann et al. 2006). As a simple extrapolation cannot capture the evolution of the precipitation field in terms of storm initiation, growth, and decay, such uncertainty needs to be quantified in the nowcasting procedure. Following Seed (2003), the prediction is generated by using an autoregressive process of order 2, defined as AR(2):

$$\begin{aligned} \mathbf{x}_{i,t+1}^f &= \phi_1 \mathbf{x}_{i,t} + \phi_2 \mathbf{x}_{i,t-1} + \boldsymbol{\varepsilon}_{i,t+1}, \quad \boldsymbol{\varepsilon}_{i,t} \sim \mathcal{N}(0, \boldsymbol{\Sigma}) \\ i &= 1, \dots, n. \end{aligned} \quad (5)$$

In (5), $\mathbf{x}_{i,t}$ and $\mathbf{x}_{i,t-1}$ are the normalized precipitation fields of the i th member that have been advected forward to time $t + 1$ using (4). The autoregressive coefficients ϕ_1 and ϕ_2 and the variance of the noise term $\boldsymbol{\varepsilon}$ are related to the rate of evolution of the field in storm

coordinates and can be estimated in real time with the Yule–Walker equations based on the temporal autocorrelation function (e.g., Wilks 2011b). Particular care for the covariance structure of $\boldsymbol{\varepsilon}$ must be taken; for details on how the correct spatial structure of a rainfall field can be efficiently reproduced, the reader is referred to Nerini et al. (2017). Following the work of Seed (2003), the temporal evolution of precipitation fields is stochastically simulated by decomposing the precipitation field into a multiplicative cascade. The scale decomposition framework is very attractive in the context of precipitation nowcasting since it accounts for spatial and dynamic scaling properties of the precipitation field (large-scale features evolve more slowly and are more predictable than small-scale features). In practice, it means that the autoregressive model introduced in (5) is applied independently to each cascade level. As a consequence, the rate of temporal evolution imposed by the autoregressive process is consistent with the predictability of the spatial scale represented by the cascade level. More details can be found in Seed et al. (2013).

c. Correction

The correction step of the KF performs the update of the prior distribution $\mathbf{x}^f$ (we neglect the time index) using a related observation in order to produce a posterior distribution for the next prediction. Measurements from p sensors are included into the observation vector $\mathbf{z} \in \mathbb{R}^{p \times 1}$, while the observation uncertainty is represented by the covariance matrix $\mathbf{R} \in \mathbb{R}^{p \times p}$. As observations do not necessarily relate to all state variables directly, a linear observation operator $\mathbf{H} \in \mathbb{R}^{p \times m}$ is used to translate the state estimates into equivalent observations. It thus becomes possible to compute the residuals between model mean prediction and observation vectors into what is known as the innovation term, $\mathbf{z} - \mathbf{H}\boldsymbol{\mu}^f \in \mathbb{R}^{p \times 1}$.

The posterior mean vector $\boldsymbol{\mu}$ is constructed by updating the prior mean vector $\boldsymbol{\mu}^f$ with a weighted innovation term:

$$\boldsymbol{\mu} = \boldsymbol{\mu}^f + \mathbf{K}(\mathbf{z} - \mathbf{H}\boldsymbol{\mu}^f), \quad (6)$$

and the prior state covariance matrix $\mathbf{P}^f$ is also updated accordingly:

$$\mathbf{P} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^f, \quad (7)$$

where $\mathbf{I} \in \mathbb{R}^{m \times m}$ is the identity matrix, and the weighting matrix $\mathbf{K} \in \mathbb{R}^{m \times p}$ is the *Kalman gain*, which is defined as follows:

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1}. \quad (8)$$

In the EnKF, the correction step is performed individually on each ensemble member:

$$\begin{aligned}\mathbf{x}_i &= \mathbf{x}_i^f + \mathbf{K}(\mathbf{z}_i - \mathbf{H}\mathbf{x}_i^f), \\ \mathbf{z}_i &= \mathbf{z} + \boldsymbol{\xi}_i, \quad \boldsymbol{\xi}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), \\ i &= 1, \dots, n.\end{aligned}\quad (9)$$

The innovation term in (9) is computed with a perturbed version of the observation $\mathbf{z}_i = \mathbf{z} + \boldsymbol{\xi}_i$, where the added perturbation comes from the observation error distribution. This is necessary in order to produce the correct covariance of the analyzed ensemble (Evensen 2003).

In our Bayesian nowcasting system, the time-synchronous precipitation forecasts from the latest available NWP ensemble prediction system (EPS) run are assimilated as pseudo-observations $\mathbf{z}_i$. Assimilating a different NWP member $\mathbf{z}_i$ for each ensemble member $\mathbf{x}_i$ is equivalent to assimilating perturbed observations in (9). As observation error covariance matrix $\mathbf{R}$ in (8), we use the sample covariance matrix of the NWP ensemble fields:

$$\begin{aligned}\bar{\mathbf{z}} &= \frac{1}{n} \sum_{i=1}^n \mathbf{z}_i, \\ \mathbf{R} &= \frac{1}{n-1} (\mathbf{Z} - \mathbf{1}\bar{\mathbf{z}}^T)^T (\mathbf{Z} - \mathbf{1}\bar{\mathbf{z}}^T),\end{aligned}\quad (10)$$

where $\mathbf{Z}$ is the matrix of the n NWP members $\mathbf{z}_i$ stacked by rows.

In a preprocessing step, the original NWP EPS forecasts and the radar QPE product are interpolated onto the same grid consisting of m points; therefore, $\mathbf{x}_i$ and $\mathbf{z}_i$ include the same variable and unit over the same grid. As a result, $\mathbf{H}$ becomes the identity matrix and can be dropped from the equation for the Kalman gain, which becomes

$$\mathbf{K} = \mathbf{P}^f (\mathbf{P}^f + \mathbf{R})^{-1}. \quad (11)$$

It is also possible to include in (11) a covariance inflation factor, as is commonly done for practical applications of the EnKF in data assimilation (Hunt et al. 2007). Covariance inflation techniques can account for biases in the estimation of both the model state and the observation uncertainty by tuning the inflation factor against some measure of the analysis performance. For the sake of simplicity, this study does not include an inflation factor, and we thus assume that $\mathbf{P}^f$ and $\mathbf{R}$ are unbiased.

To summarize, the correction equation of our Bayesian nowcasting system is expressed as

$$\mathbf{x}_i = \mathbf{x}_i^f + \mathbf{K}(\mathbf{z}_i - \mathbf{x}_i^f), \quad i = 1, \dots, n, \quad (12)$$

where $\mathbf{K}$ is defined in (11), and $\mathbf{z}_i$ are obtained from the n members of the latest available NWP EPS run.

Finally, $\mathbf{X}$ represents the posterior ensemble forecast including information from both the extrapolation forecast and the numerical forecast. It is worth emphasizing again that this Bayesian nowcasting application of the EnKF reverses the standard interpretation of the NWP forecast by using it as pseudo-observation for the correction of the radar extrapolation forecast. It is also important to stress that the EnKF algorithm is here cycled over lead times, meaning that an analysis is produced at each integration step.

d. Reduced-space EnKF by means of principal component analysis

The high-resolution precipitation fields used in nowcasting can be seen as samples from high-dimensional probability distributions, where each dimension corresponds to one of the m grid points in the domain. Because we only have n members, the $m \times m$ covariance matrix has a maximum rank of $n - 1$, and it is therefore not invertible. Moreover, the cost of computing and storing a covariance matrix in such a highly dimensional space can become quickly prohibitive. To address such limitations, the operations in the EnKF correction step can be carried out in a subspace of much lower dimension.

The idea of computing a more efficient EnKF analysis in a space of reduced dimensionality has been already exploited in data assimilation methods [see, e.g., Ott et al. (2002) and references therein]. The definition of such subspace is, of course, the main question. The ensemble transform Kalman filter (Bishop et al. 2001; Hunt et al. 2007) makes direct use of the prior members in order to define the basis of an ensemble space wherein to perform the analysis. A limitation of this approach is that the basis is not linearly independent, so other approaches have used the singular vectors of $\mathbf{X}^f$, as in the ensemble adjustment Kalman filter of Anderson (2001). In this way, the analysis is performed in a space spanned by an orthogonal basis, although this involves the additional cost of computing the singular vectors.

Our approach is related to the above, as it attempts to simplify the EnKF update by performing the necessary matrix computations in a lower-dimensional space. PCA is a widely used tool that can fit our purpose. Because the orthogonal principal components that are extracted from the data are sorted by decreasing explained variance, truncation can be applied to reduce the number of dimensions while minimizing the reconstruction error (Tipping and Bishop 1999). The main idea is thus to compute the correction step of the EnKF

with $\mathbf{X}^f$ and $\mathbf{Z}$ projected onto the subspace spanned by the principal components. This projection step ensures a manageable problem size, regardless of the original data resolution or dimensionality. The analysis $\mathbf{X}$ can thus be computed in the reduced space and then transformed back to the spatial domain, ready for the next prediction step of the EnKF. The principal component analysis is applied to the concatenated matrix $\mathbf{Y} = [\mathbf{X}^f, \mathbf{Z}] \in \mathbb{R}^{2n \times m}$ as *S-mode* PCA (Wilks 2011a). The truncated matrix of eigenvectors $\mathbf{U} \in \mathbb{R}^{m \times d}$ with $d \leq 2n - 1 \ll m$ is defined in order to minimize the mean-squared error of the reconstructed data for both $\mathbf{X}^f$ and $\mathbf{Z}$ and to conserve at least 99% of the original variance. Notice that the size of the projection matrix $\mathbf{U}$ cannot be known in advance, as PCA is applied at each correction step. The columns of the projected matrices $\mathbf{X}^f \mathbf{U}$ and $\mathbf{Z} \mathbf{U} \in \mathbb{R}^{n \times d}$ are called PC scores. The state update equations are modified as follows:

$$\bar{\mathbf{x}}_d^f = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i^f \mathbf{U},$$

$$\mathbf{P}_d^f = \frac{1}{n-1} (\mathbf{X}^f \mathbf{U} - \mathbf{1} \bar{\mathbf{x}}_d^{fT})^T (\mathbf{X}^f \mathbf{U} - \mathbf{1} \bar{\mathbf{x}}_d^{fT}). \quad (13)$$

$$\bar{\mathbf{z}}_d = \frac{1}{n} \sum_{i=1}^n \mathbf{z}_i \mathbf{U},$$

$$\mathbf{R}_d = \frac{1}{n-1} (\mathbf{Z} \mathbf{U} - \mathbf{1} \bar{\mathbf{z}}_d^T)^T (\mathbf{Z} \mathbf{U} - \mathbf{1} \bar{\mathbf{z}}_d^T). \quad (14)$$

$$\mathbf{K}_d = \mathbf{P}_d^f (\mathbf{P}_d^f + \mathbf{R}_d)^{-1}. \quad (15)$$

$$\mathbf{x}_i = [\mathbf{x}_i^f \mathbf{U} + \mathbf{K}_d (\mathbf{z}_i \mathbf{U} - \mathbf{x}_i^f \mathbf{U})] \mathbf{U}^T, \quad i = 1, \dots, n. \quad (16)$$

In this way, now $\mathbf{P}_d^f \in \mathbb{R}^{d \times d}$, $\mathbf{R}_d \in \mathbb{R}^{d \times d}$, and $\mathbf{K}_d \in \mathbb{R}^{d \times d}$ represent the reduced-space sample covariance matrices and the reduced-space Kalman gain, denoted by the subscript d . By definition, the PC scores are mutually uncorrelated. Thus, the sample covariance matrices $\mathbf{P}_d^f$, $\mathbf{R}_d$, and hence $\mathbf{K}_d$ can safely be assumed to be diagonal.

The approach is illustrated in Fig. 2 for a radar-based nowcast at lead time +40 min trying to assimilate a NWP EPS forecast that was initialized almost 10 h before. The data in this example belong to case study ID01; see section 4b. In this figure, two of the precipitation fields in $\mathbf{X}^f$ and $\mathbf{Z}$ are projected by the projection matrix $\mathbf{U}$ into the PC space where the prior ensemble $\mathbf{X}^f$ is updated. After the update, the analysis $\mathbf{X}$ is projected back to the spatial domain. In this example, it is possible to qualitatively appreciate how $\mathbf{X}^f$ has lower spread; that is, the two prior members are closer to each other, compared to the two members from $\mathbf{Z}$. This is quantified

in terms of variances of the PC score distributions, where the $\mathbf{Z} \mathbf{U}$ distribution is clearly more spread than the $\mathbf{X}^f \mathbf{U}$ distribution. Consequently, the posterior $\mathbf{X} \mathbf{U}$ distribution results much closer to the prior distributions (notice the very low values for $\mathbf{K}$).

e. Dealing with non-Gaussian precipitation data

Precipitation exhibits a bimodal distribution, being either zero or lognormally distributed. Such non-Gaussianity is problematic to the reduced-space EnKF technique described above mainly in two ways. First, it introduces biases in the estimation of the sample covariance matrices during the Bayesian update. Second, it causes an overestimation of the wet area and underestimation of maximum intensity when precipitation fields are linearly combined to produce the analysis. We refer to the latter effect as *smoothing*.

To account for the skewed distribution of precipitation, positive rain rates R are log-transformed to produce units of dBR according to the following formula: $\text{dBR} = 10 \log_{10} R$. The influence of zeros on the estimation of the covariance matrices is limited by adapting the criterion described by Lien et al. (2013) so that the Kalman gain is computed by only considering grid points where at least 50% of the members in both ensembles have positive precipitation. Hence, (13), (14), and (15) are applied with a subset of $\mathbf{X}$, $\mathbf{Z}$, and $\mathbf{U}$, while (16) uses the whole domain.

To adjust the bias in the wet area and the underestimation of maximum precipitation intensity that is found in the analysis, we introduce a final resampling and probability matching step. Similar approaches can be found in many applications including recent developments in nonlinear data assimilation [e.g., the probability mapping of Poterjoy (2016)]. Our implementation is inspired by the procedure described by Ebert (2001) for the computation of probability matched ensemble medians. The basic idea is first to build a new empirical posterior distribution of precipitation intensities by resampling the prior and pseudo-observation empirical distributions. Finally, the original posterior distribution is matched to the resampled distribution. The algorithm can be described as follows.

- 1) For the current lead time t , compute the resampling weight p as

$$p = \frac{1}{m} \sum_{j=1}^m \frac{\sum_{i=1}^n x_{ij} - x_{ij}^f}{\sum_{i=1}^n z_{ij} - x_{ij}^f}, \quad (17)$$

which can be interpreted as the average probability of sampling the pseudo-observation $\mathbf{z}_{ij}$ based on the

FIG. 2. Example of reduced-space EnKF for one particular time step of a nowcast (+40 min; event ID01). The first two columns include the first and last members of the prior $\mathbf{X}^f$, pseudo-observations $\mathbf{Z}$, and posterior $\mathbf{X}$ ensembles and thus illustrate the effect of the update in the spatial domain. The last two columns illustrate the PCA space that it is used to compute the update. Only the eigenvectors and PC scores for the first three components are shown. The eigenvectors in $\mathbf{U}$ project the original fields into the space of PCs. The distributions of the ensembles for a given PC are represented by the histograms, while the dashed lines are the best fits for a normal distribution. The transformed posterior distribution ($\mathbf{X}\mathbf{U}$; in green) is a weighted average of the transformed prior ($\mathbf{X}^f\mathbf{U}$; red) and transformed pseudo-observation ($\mathbf{Z}\mathbf{U}$; blue) distributions according to the Kalman gain $\mathbf{K}$. The event is described in section 4b, and the domain is as in Fig. 5.

updated ensemble $\mathbf{X}$, where m is the number of grid points in the domain and n is the ensemble size.

- 2) Sort the m rain rates from $\mathbf{x}_i^f$, $\mathbf{x}_i$, and $\mathbf{z}_i$ to obtain their empirical cumulative probability density functions for member $i = 1$.
- 3) For rank $r = 1, \dots, m$, randomly sample the corresponding percentile from either $\mathbf{z}_{i(r)}$ with probability p or $\mathbf{x}_{i(r)}^f$ with probability $1 - p$. A new posterior empirical distribution is thus generated.
- 4) Assign the value of the highest rank $r = 1$ from the new distribution to the highest rank in the posterior, and so on, down to the lowest rank $r = m$.
- 5) Repeat steps 2–4 for the remaining ensemble members $i = 2, \dots, n$.

f. Summary of the algorithm

The basic steps of the Bayesian nowcasting algorithm are summarized by the flowchart in Fig. 1. As the diagram shows, the last available radar QPE ensemble, called $\mathbf{X}_0$, represents the starting point of the nowcast. A stochastic extrapolation technique is then used to predict the next state and at the same time quantify the uncertainty of the extrapolation with an ensemble of perturbed predictions. In the next step, the NWP ensemble is used to update the stochastic extrapolation in a Bayesian fashion. To address the high dimensionality of the system, the update is performed in a reduced space computed with PCA. A resampling and probability matching step is introduced at the end of each iteration to correct the higher-order statistics of the posterior

rain-rate distribution (not shown in the diagram). The resulting analysis ensemble is finally used as starting point for the next time step prediction.

3. Synthetic data experiments

The Bayesian nowcasting technique is first applied to synthetic data in a series of numerical experiments. Synthetic data use prescribed parameters and data distributions, thus simplifying the interpretation of the simulation results.

a. Synthetic data

The synthetic time series of true model states $\mathbf{x}$ is simulated as a multivariate Gaussian random process having zero mean, unit variance, and spatial autocorrelation function $\rho_x(u) = \exp(-u/r)$, where u is the lag distance, while the parameter r sets the decorrelation distance (Schiemann et al. 2011). The temporal correlation is produced with a first-order autoregressive model AR(1) controlled by the autocorrelation coefficient ρ_t :

$$\mathbf{x}_{t+1} = \rho_t \mathbf{x}_t + \boldsymbol{\varepsilon}_{t+1}, \quad \boldsymbol{\varepsilon}_{t+1} \sim \mathcal{N}(0, \mathbf{C}), \quad (18)$$

where $\mathbf{C}(u) = \rho_x(u)$ is the covariance matrix of $\mathbf{x}$. The synthetic initial prior ensemble $\mathbf{X}_0$ is thus constructed as a stack of $\mathbf{x}$ at $t = 0$. Data are generated over a 10×10 pixel grid and for 20 time steps.

The synthetic pseudo-observations $\mathbf{Z}$ are generated by adding Gaussian random noise $\mathcal{N}(\boldsymbol{\mu}, \mathbf{C}_n)$ to the synthetic sequence of true model states $\mathbf{x}$. Spatial and temporal correlation of the random noise are prescribed using the same approach described above. To simulate the error of the pseudo-observations $\mathbf{Z}$, a bias term is introduced by setting $\boldsymbol{\mu} > 0$. The bias of the random noise is set equal to its standard deviation so that spread-skill consistency is ensured (i.e., the ensemble forecast is reliable; Fortin et al. 2014).

b. Simulation results

Figure 3 shows the results of the Bayesian nowcast from a set of 100 synthetic simulations using 100 ensemble members. Please note that the posterior estimate at lead time t is computed based on the ensembles of prior and pseudo-observations at the same lead time. Since this is a recursive algorithm, the prior at lead time t is based on the posterior of time $t - 1$. The mean value over a target box of 5×5 pixels centered at the simulation grid is shown as function of time step. The ensemble medians for persistence (EXTR), pseudo-observations (NWP), and posterior (BAY) are represented as well as their ensemble spreads (interval

FIG. 3. Results of 100 synthetic simulations to test the Bayesian nowcasting based on reduced-space EnKF. The mean value over a 5×5 pixel target box is shown as a function of time step for a stochastic persistence forecast (red), synthetic NWP ensemble (blue), and Bayesian nowcasting (green). The graph represents at each lead time the median and the 10%–90% interval. All simulations used 100 member ensembles, $\rho(1) = 0.95$, and an NWP spread of 60%.

between 10th and 90th percentiles). It can be seen that the spread of the persistence forecast increases as a function of lead time, while the spread of the synthetic pseudo-observations remains more or less constant. The EnKF posterior estimate gradually converges to the ensemble of pseudo-observations, and after about 10 time steps, it closely matches it.

Figure 4 presents the results of two experiments that verify the ability of the Bayesian nowcast to account for flow dependence, here intended as the capacity to adapt to the uncertainty of both the pseudo-observations and radar-based nowcast. The root-mean-square errors (RMSE) are computed with respect to the simulated true model state $\mathbf{x}$. In the first experiment (Fig. 4a), the spread of the synthetic pseudo-observations is kept constant at 60% (NWP; horizontal dashed line), while the temporal autocorrelation coefficient ρ_t of the random field is set equal to 0.97, 0.95, and 0.90, which represents a typical range of values for actual radar precipitation fields. This reflects the fact that the temporal persistence of precipitation varies from case to case, thus affecting the skill of the persistence forecast (EXTR; dotted line). The Bayesian nowcast (BAY; solid line) starts with zero error and then correctly converges to the same error of pseudo-observations. In the second experiment (Fig. 4b), the persistence of the synthetic true model state is kept constant ($\rho_t = 0.95$), while the spread of the pseudo-observations is increased (40%, 60%, and 80%). This simulates the variability of the NWP skill. The Bayesian nowcast is

FIG. 4. Results from two numerical experiments with the Bayesian nowcast based on reduced-space EnKF. (a) The skill of the Bayesian nowcast ensemble mean (solid line) is verified for three levels of temporal persistence of the simulated true model state (dotted line; left to right: $\rho_t = 0.90, 0.95, 0.97$) and a fixed spread of the pseudo-observations set equal to 60%. (b) The skill of the Bayesian nowcast is verified for an increasing spread of the ensemble of pseudo-observations (bottom to top: 40%, 60%, and 80%) and for a fixed temporal persistence $\rho_t = 0.95$.

once again able to converge to the skill of the pseudo-observations at a rate that is consistent with its uncertainty.

4. Real data experiments

The Bayesian nowcasting technique is now applied and verified with a set of real precipitation events over the Swiss Alps.

a. Datasets and study domain

The weather radar observation dataset was produced by the fourth generation of the MeteoSwiss radar network. The radar QPE is a composite image derived from measurements of five dual-polarization, C-band

FIG. 5. The 512 km $\times$ 512 km study area (solid red line) and the 300 km $\times$ 300 km verification domain (dashed red line) are centered over Switzerland. The five radars are indicated with white triangles. The composite radar mask, the topography, and the national borders are also included as references.

weather radars distributed across the Swiss territory (white triangles in Fig. 5). The data processing chain is specifically conceived to address the many challenges that exist in operating weather radars in an Alpine region. The design of the fourth-generation Swiss radar network is presented by Germann et al. (2015, 2016, 2017). The precipitation product is based on a fixed $Z-R$ relationship between measured reflectivity and rainfall intensity $Z = 316R^{1.5}$. The original QPE product is available at 5-min and 1-km resolutions over a wide domain that covers Switzerland and the neighboring regions of Germany, France, Italy, Austria, and Liechtenstein. To match the resolution of the NWP model, the QPE product has been aggregated to 10-min and 2-km resolutions. The domain has been reduced to a 512 km $\times$ 512 km region centered over Switzerland (red solid box in Fig. 5).

A control experiment is produced through stochastic radar extrapolation. This ensemble forecast is based solely on the stochastic Lagrangian persistence approach described in section 2b. Therefore, it represents a benchmark against which the benefits of the Bayesian update can be assessed.

The mesoscale EPS forecasts that are included in this study are produced by the 21-member COSMO-E model that has been in operation at MeteoSwiss since 2016 (Klasa et al. 2018). In this study, the control member is treated as an additional member of the ensemble (i.e., 20 perturbed members and one unperturbed

FIG. 6. Maps of total accumulated radar precipitation corresponding to the events considered in section 4. The timestamp indicates the start of the event, while the event duration is given within brackets.

control). The COSMO-E model provides twice-daily (initialization times at 0000 and 1200 UTC) medium-range forecasts (up to 120 h) for the Alpine region at convection-permitting resolution (2.2-km mesh size). The ensemble of initial conditions is produced by the kilometer-scale ensemble data assimilation system (KENDA; Schraff et al. 2016), based on an LETKF (Hunt et al. 2007). The lateral boundary conditions are provided by the ECMWF ensemble model. The radar reflectivities are assimilated using a latent heat nudging scheme (Leuenberger and Rossa 2007; Stephan et al. 2008). Finally, the uncertainty of the

model parameterizations is included in COSMO-E with stochastically perturbed physics tendencies (SPPT; Buizza et al. 1999).

b. Case studies

Four precipitation events between 2016 and 2017 were selected. They are listed in Fig. 6, where their total accumulated precipitation is represented.

The first precipitation event ID01 (Fig. 6a) is linked to a frontal system causing convective activity to be organized in lines with a SW–NE orientation, while the general flow is predominately west-southwest.

The second event ID02 (Fig. 6b) concerns a winter precipitation case that is produced by a strong northwesterly flow bringing moist air toward the Alps. As a consequence, precipitation mainly occurs on the northern slopes while it is effectively blocked by the Alpine chain, leading to the formation of a rain shadow south of the Alps.

Event ID03 (Fig. 6c) is a summertime situation with widespread and localized convective activity under a westerly flow. Thunderstorms have a scattered distribution but only occurred in the southern half of the domain.

The last precipitation case ID04 (Fig. 6d) is again a summertime event, but this time the flow is from the southwest, resulting in an interesting case with pre-frontal convective activity on the southern side of the Alps and an active frontal passage with organized convection on the northern side of the Alps.

c. Verification method

Because of the large variability of precipitation fields, the verification of high-resolution ensemble precipitation forecasts represents a challenging task, particularly in the presence of convective precipitation. In this study, the fractions skill score (FSS; Roberts and Lean 2008) was chosen as reference error metric. Following Ebert et al. (2013), the FSS is defined as

$$\text{FSS} = 1 - \frac{\frac{1}{M} \sum_M (P_F - P_O)^2}{\frac{1}{M} \sum_M P_F^2 + \frac{1}{M} \sum_M P_O^2}, \quad (19)$$

where P_F and P_O are the forecast and observed fractions of grid points with intensity greater than a given threshold within a given square neighboring area, while M is the number of grid points in the verification domain. The FSS spans from 0 (worst possible score) to 1 (perfect match).

FSS conveniently depends on spatial scale and intensity, and it is thus particularly suitable for the verification of precipitation fields. To generalize the FSS to the ensemble skill, we applied the same approach described by Zacharov and Rezacova (2009), which consists of computing the FSS for each individual ensemble member and then averaging it to obtain the mean ensemble skill.

d. Bayesian precipitation nowcast examples

Figure 7 shows an example of a nowcast started at 2100 UTC 11 July 2016 (ID01). Each row represents one selected lead time (+00, +60, +120, or +180 min). Each column includes one of the forecasts, namely, the

COSMO-E control member (COE), one random member from the Bayesian nowcast (BAY), or stochastic radar extrapolation (EXTR) ensembles. The verifying observations from the radar QPE are also included in the first column (OBS).

An organized line of convection moving eastward with a southwest–northeast orientation is clearly visible in the radar images in the first column. In the second column, COSMO-E was initialized at 1200 UTC (i.e., 9 h before the start of this nowcast). In a real-time situation, this is currently the most recent EPS run that would be available for forecasting precipitation. The initialization time of the NWP EPS run is implicitly accounted for in the Bayesian nowcast, as this adapts to the spreads in both the NWP and extrapolation nowcasts, which also depend on their initialization time. We refer to this behavior as *flow dependence*. One can notice a delay in the initiation of convective activity in the COSMO-E control run, where the storms do not yet appear as developed and organized as in the observation. It is in such situations that the use of a nowcasting system is particularly valuable.

The third column presents the result of the Bayesian nowcasting introduced in this study. One can expect this forecast to be a combination of the forecasts in the second and fourth columns. In fact, the first and last frames are identical and very close to the extrapolation and COSMO-E fields, respectively, while in between the correction step of the EnKF uses the COSMO-E forecast to update the extrapolation forecast. The fourth column contains one member of the stochastic radar extrapolation ensemble forecast. The last observations are advected while being progressively replaced by correlated noise in order to simulate the loss of predictability of the Lagrangian persistence forecast. The radar mask is also advected along with the radar field in order to illustrate the influence of the domain boundaries on the extrapolation technique.

A more quantitative illustration of the Bayesian nowcasting technique is given in Fig. 8, where the mean precipitation intensity over the 300 km × 300 km verification box (Fig. 5) is shown as a function of forecast lead time. As in Fig. 3, the plot shows the ensemble median estimates with their 10%–90% prediction intervals, while the observed mean intensity over the box is also included as a black dotted line. It is interesting to note that not only does the mean precipitation intensity seamlessly merge from observations into the NWP forecast, but also the variability within the ensemble grows consistently with its transition from the observations to the numerical model.

A second example of a nowcast is presented in Fig. 9. This time, the nowcast is started at 1000 UTC 31 January

FIG. 7. An example for a forecast started at 2100 UTC 11 Jul 2016 (ID01) with (a) verifying radar QPE at 2100, 2200, 2300, and 0000 UTC; (b) the COSMO-E control run; (c) one realization (member) of the Bayesian nowcast; and (d) one member of the stochastic radar extrapolation ensemble at +00, +60, +120 and +180 min after 2100 UTC. The COSMO-E forecast was initialized at 1200 UTC. The shaded area indicates the radar QPE mask. All fields are 10-min equivalent precipitation intensities.

2017 (ID02), and it concerns the winter precipitation event with pronounced orographic precipitation, as can be seen in the radar images.

The COSMO-E run was initialized at 0000 UTC (i.e., 10 h before the start of the nowcast). While the general structure of the precipitation event seems well captured, the control run of the ensemble is missing most of the higher precipitation intensities, as seen by the radar observation.

By design, the stochastic extrapolation model conserves the rain-rate distribution over the whole domain (it is a persistence forecast). It is also interesting to observe how a simple extrapolation approach can be challenging in regions with complex orography, as it can happen that precipitation is unrealistically advected across a mountain ridge to areas that should remain cloud free. There are ongoing research efforts in trying to account for additional sources of predictability, such

FIG. 8. Mean areal precipitation intensity in dBR over the $300\text{ km} \times 300\text{ km}$ verification box as a function of forecast lead time for the nowcast started at 2100 UTC 11 Jul 2016 (as in Fig. 7). The median and the 10%–90% prediction interval are visualized.

as the diurnal cycle (Atencia et al. 2017) or orographic forcing (Panziera and Germann 2010; Foresti et al. 2018) for the correction of extrapolation nowcasts.

The Bayesian nowcasting at +60 min has already picked up the large-scale structure from COSMO-E, but at the same time, it still manages to produce higher intensity values, consistently with the distribution of rain rates at observation time. At +180 min, it has almost entirely converged to the numerical model.

The seamless transition from the radar observation to the COSMO-E forecast is perhaps easier to notice in Fig. 10, where the mean precipitation intensity over the $300\text{ km} \times 300\text{ km}$ verification box (Fig. 5) is shown as a function of forecast lead time. It is now clearer from this chart that the Bayesian nowcast starts exactly from the last radar observation and then moves toward the COSMO-E forecast. After 3 h, the Bayesian nowcast is very close to COSMO-E. Another interesting aspect to note is how the spread of the Bayesian nowcast ensemble is equal to zero by definition at the beginning of the nowcast (assuming no observation error) and increases to closely match the COSMO-E ensemble spread. In this sense, the Bayesian nowcasting can be considered seamless in terms of both mean and spread.

It is important to remember that the Bayesian update is based on the distribution of the PC scores (cf. Fig. 2). The spread in terms of PC scores is illustrated in Fig. 11, where the average ratio between the extrapolation and NWP prediction intervals is shown for the first five PC scores and as function of lead time. The ratio becomes larger than 1.0 (black dotted line) when the prediction interval of a given PC score in the extrapolation is larger than the corresponding prediction interval for the NWP.

For the examples included in this study, this happens on average between 1.5 and 2 h.

e. Verification results

For each event, a 4-h Bayesian nowcast was started every hour. Given that all events last 12 h, the number of runs per event is equal to 9. Therefore, there are in total $4\text{ events} \times 9\text{ runs} \times 4\text{ h} = 144\text{ h}$ that are included in this verification. If one further considers the time resolution of 10 min (hence, six precipitation fields for every hour), the size of the ensembles (21 members) and the number of forecasting procedures that are compared (i.e., radar-based nowcasting, NWP, and Bayesian estimate), there is a total of $144\text{ h} \times 6\text{ fields h}^{-1} \times 21\text{ members} \times 3\text{ procedures} = 54\,432$ precipitation fields that are verified.

The verification of the four precipitation events (ID01–04) is presented in Fig. 12, which illustrates the mean ensemble skill in terms of FSS at 32-km scale and 2.0 mm h^{-1} . These scale and intensity parameters are those typically used in comparison analysis of precipitation forecasts (e.g., Simonin et al. 2017). Scores are computed within the verification box as defined in Fig. 5. For each of the three ensemble forecasts, the average skill and its variability among all runs are represented.

We will refer to *crossover time* to indicate the lead time when the skill of the numerical model becomes better than the skill of the radar extrapolation, which is easily found in the verification plots as the point where the blue line crosses the red line. Hence, the challenge for the Bayesian nowcast is to correctly estimate the crossover point using the relative spread of the two ensemble forecasts.

The first verification concerns the convective event presented above (ID01; see Figs. 7 and 8). Interestingly, the skill of the NWP ensemble shows a considerable variability within the precipitation event, particularly at short forecast lead times, where it spans from 0.8 to less than 0.1. This is an indication that for this particular event, the NWP skill strongly varied among model runs. The crossover time is on average between +2 and +2.5 h, and it is well captured by the Bayesian nowcasting system. However, there is a tendency to converge a bit too early to COSMO-E.

The second verification case concerns the winter stratiform event (ID02; see Figs. 9 and 10). In Fig. 12b, the average ensemble skill for COSMO-E is higher and has a lower variability compared to event ID01, while the skill of the extrapolation drops more rapidly. Consequently, the average crossover time appears to occur already after about 1.5–2.0 h. In this case, the Bayesian approach is correctly weighting the contribution

FIG. 9. As in Fig. 7, but for the nowcast started at 1100 UTC 31 Jan 2017 (ID02). The COSMO-E forecast was initialized at 0000 UTC of the same day.

of the numerical model to produce an optimal combined forecast.

The intense and localized precipitation of event ID03 translates into the low skill of COSMO-E, as can be seen in Fig. 12c. This seems to reflect the fact that the numerical model cannot correctly forecast individual thunderstorms at a spatial scale of 32 km. As a consequence, the crossover point is not reached within the 4-h lead times included in the verification. The Bayesian nowcast manages to identify the poor skill of the

numerical model, and as a result, the convergence is also occurring after a lead time of 4 h. However, the performance of the Bayesian nowcast remains lower than the stochastic radar extrapolation for most of the forecast range, in particular between 1- and 3-h lead time, an indication that the skill of COSMO-E is somewhat overestimated.

Concerning the last precipitation case study (ID04) in Fig. 12d, it can be seen that the large-scale organization of this event leads to a slower decay of the skill of

FIG. 10. As in Fig. 8, but for the nowcast started at 1100 UTC 31 Jan 2017 (ID02).

the radar extrapolation compared with event ID03. COSMO-E also exhibits a reasonably good accuracy, and the average crossover is found at about 2.5–3 h, which is also well captured by the Bayesian nowcast.

The four examples illustrate how a combination of flow-dependent NWP and radar extrapolation skills requires a blending scheme that can recognize and then adapt to each situation accordingly. In fact, the observed crossover time for this limited sample of events and the specific definition of skill used here varied from 1.5 to more than 4 h. The Bayesian nowcasting has been shown to effectively use the ensemble spread as an estimate for the forecast uncertainty to derive optimal weights to integrate the NWP forecasts with the radar-based nowcasts. The blended product is a better forecast overall, since the total skill of the blended nowcast is equal to or larger than the total of both the radar-based and NWP-based nowcast skills.

Despite these promising results, here we did not observe the same improvement in terms of crossover time that was found in the numerical experiments based on synthetic Gaussian data (Fig. 4). Besides a larger variability due to the limited sample size, we probably also need to attribute such outcome to the challenging nature of real precipitation data.

5. Summary and conclusions

Nowcasting can be seen as a forecasting procedure that combines NWP forecasts with real-time observations in the very short-term range (from 0 to about 6 h). The result should be one seamless forecast that optimally integrates all available sources of predictability in a consistent way.

Such a nowcasting system should use the information from ensembles, as these can convey uncertainty and at

FIG. 11. The spread of the extrapolation nowcast (PI_{EXTR}) relative to the COSMO-E spread (PI_{COE}) in terms of the first five PC scores. The results are averaged over all the case studies. The spread PI is defined as the 10%–90% prediction interval.

the same time respect the spatiotemporal structure of the variable of interest (Berenguer et al. 2011). Already recognized as an established technique in data assimilation, the Kalman filter is an attractive tool for the nowcasting problem, too, thanks to its recursive implementation of the Bayesian update equations. However, its application to high-resolution precipitation fields is not trivial and requires a number of adjustments to relax some of its basic assumptions and efficiently work with large datasets. In this study, a reduced-space ensemble Kalman filter was introduced as a tentative answer to such requirements.

Synthetic data experiments were run in order to test the methodology. The results have demonstrated that the Bayesian nowcast correctly captures the flow dependence of both the NWP forecast and the Lagrangian persistence of the radar observations. The result is an adaptive blending scheme that depends on the forecast skill as estimated by the forecast ensemble spread.

Four experiments were run with real data produced by the Swiss weather radar network and the COSMO-E model. This represented an application of the EnKF blending approach with actual precipitation fields and a relatively small ensemble size (21 members). The four meteorological situations spanned from a stratiform wintertime precipitation case to very localized summertime convection. The first two case studies were detailed in a sequence of images for a qualitative assessment of the system. The skill for all four cases was quantified in terms of fractions skill score at 32-km scale and 2.0 mm h^{-1} . Scores were computed for different scales and intensities (not shown), yielding consistent results. The results indicate that the system can cope

FIG. 12. Mean ensemble FSS as a function of lead time for a 32-km scale, threshold 2.0 mm h^{-1} , and computed within the verification box (see Fig. 5) for all four events (ID01–04). The mean ensemble FSS is the average of the individual FSSs of each ensemble member. The bars represent the variability among the individual runs within a single event. The three ensembles are the COSMO-E in blue (COE), the Bayesian nowcasting in green (BAY), and the stochastic radar extrapolation in red (EXTR).

with realistic conditions to produce precipitation forecasts that perform as well as extrapolation nowcasts for short lead times and then converge to the skill of NWP at a plausible rate.

Compared to traditional blending schemes based on fixed linear weights, the Bayesian approach does not need a training period, and it can adapt to various meteorological situations and changes in the forecast skill. Furthermore, it provides a blending that is a function not only of lead time, but also of location and spatial scale, thanks to its formulation into the space of principal components. Its computational cost remains within the needs of a nowcasting system: with the data used in this study, the correction step takes approximately 1 s on a 2.6-GHZ core, while one prediction step demands approximately 6 s. In total, a 4-h Bayesian nowcast takes about 2 min of computations. The only parallelization achieved in this study is the one currently integrated in the Python library numpy via OpenMP.

The framework that was presented in this study has the potential to serve for other variables than precipitation. In principle, any observed meteorological parameter can be used as a state variable of a Bayesian nowcast, provided that an error model for its persistence forecast is available. Moreover, the nowcast of multiple variables, including their interdependencies, can be foreseen. For example, future research could consider the joint integration of multiple fields (e.g., precipitation, temperature, wind) by means of a combined PCA approach (CPCA; Wilks 2011a). In this sense, the Bayesian nowcasting approach can open the door to a truly integrated, multivariable, seamless, probabilistic nowcasting system.

One additional aspect calling for future research is the Gaussian assumption of the EnKF in presence of non-Gaussian quantities such as precipitation. The non-Gaussian nature of precipitation questions the use of an EnKF approach. In principle, a nonparametric approach

such as the particle filter (PF) would be more suitable, as it employs a nonlinear and non-Gaussian update. On the other hand, PF requires a very large ensemble size, which limits its application to high-resolution data (Snyder et al. 2008). For this reason, the present study included a resampling and probability matching step to correct for the biases in wet area and maximum intensities resulting from the EnKF analysis. However, recent developments in data assimilation using local particle filters (LPF; see, e.g., Penny and Miyoshi 2016; Poterjoy 2016) and hybrid local ensemble Kalman particle filters (Robert et al. 2018) are showing promising results and hence could also represent valuable developments for seamless forecasting systems.

Acknowledgments. This study was supported by the Swiss National Science Foundation Ambizione project *Precipitation attractor from radar and satellite data archives and implications for seamless very short-term forecasting* (PZ00P2_161316). Discussions with Isztar Zawadzki on the use of PCA on precipitation fields have been of great inspiration. We are grateful to Heini Wernli, Nikola Besic, Martin Rempel, and the three anonymous reviewers for their valuable comments and corrections to the manuscript.

REFERENCES

- Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903, [https://doi.org/10.1175/1520-0493\(2001\)129<2884:AEAKFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2).
- Atencia, A., and I. Zawadzki, 2014: A comparison of two techniques for generating nowcasting ensembles. Part I: Lagrangian ensemble technique. *Mon. Wea. Rev.*, **142**, 4036–4052, <https://doi.org/10.1175/MWR-D-13-00117.1>.
- , —, and M. Berenguer, 2017: Scale characterization and correction of diurnal cycle errors in MAPLE. *J. Appl. Meteor. Climatol.*, **56**, 2561–2575, <https://doi.org/10.1175/JAMC-D-16-0344.1>.
- Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, <https://doi.org/10.1038/nature14956>.
- Berenguer, M., D. Sempere-Torres, and G. G. Pegram, 2011: SBMcst—An ensemble nowcasting technique to assess the uncertainty in rainfall forecasts by Lagrangian extrapolation. *J. Hydrol.*, **404**, 226–240, <https://doi.org/10.1016/j.jhydrol.2011.04.033>.
- Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420–436, [https://doi.org/10.1175/1520-0493\(2001\)129<0420:ASWTET>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2).
- Bowler, N. E., C. E. Pierce, and A. W. Seed, 2006: STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Quart. J. Roy. Meteor. Soc.*, **132**, 2127–2155, <https://doi.org/10.1256/qj.04.100>.
- Buil, A., 2017: Nowcasting probabilístico basado en observaciones de lluvia con radar meteorológico. Ph.D. thesis, Universitat Politècnica de Catalunya, 175 pp., <https://upcommons.upc.edu/handle/2117/111231>.
- Buizza, R., M. Milleer, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908, <https://doi.org/10.1002/qj.49712556006>.
- Cecinati, F., M. A. Rico-Ramírez, G. B. Heuvelink, and D. Han, 2017: Representing radar rainfall uncertainty with ensembles based on a time-variant geostatistical error modelling approach. *J. Hydrol.*, **548**, 391–405, <https://doi.org/10.1016/j.jhydrol.2017.02.053>.
- Ciach, G. J., W. F. Krajewski, and G. Villarini, 2007: Product-error-driven uncertainty model for probabilistic quantitative precipitation estimation with NEXRAD data. *J. Hydrometeorol.*, **8**, 1325–1347, <https://doi.org/10.1175/2007JHM814.1>.
- Dowell, D. C., L. J. Wicker, and C. Snyder, 2011: Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Mon. Wea. Rev.*, **139**, 272–294, <https://doi.org/10.1175/2010MWR3438.1>.
- Ebert, E. E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.*, **129**, 2461–2480, [https://doi.org/10.1175/1520-0493\(2001\)129<2461:AOAPMS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2461:AOAPMS>2.0.CO;2).
- , and Coauthors, 2013: Progress and challenges in forecast verification. *Meteor. Appl.*, **20**, 130–139, <https://doi.org/10.1002/met.1392>.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, <https://doi.org/10.1029/94JC00572>.
- , 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367, <https://doi.org/10.1007/s10236-003-0036-9>.
- Foresti, L., and A. Seed, 2014: 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.*, **18**, 4671–4686, <https://doi.org/10.5194/hess-18-4671-2014>.
- , I. V. Sideris, L. Panziera, D. Nerini, and U. Germann, 2018: A 10-year radar-based analysis of orographic precipitation growth and decay patterns over the Swiss Alpine region. *Quart. J. Roy. Meteor. Soc.*, **144**, 2277–2301, <https://doi.org/10.1002/qj.3364>.
- Fortin, V., M. Abaza, F. Anctil, and R. Turcotte, 2014: Why should ensemble spread match the RMSE of the ensemble mean? *J. Hydrometeorol.*, **15**, 1708–1713, <https://doi.org/10.1175/JHM-D-14-0008.1>.
- Germann, U., I. Zawadzki, and B. Turner, 2006: Predictability of precipitation from continental radar images. Part IV: Limits to prediction. *J. Atmos. Sci.*, **63**, 2092–2108, <https://doi.org/10.1175/JAS3735.1>.
- , M. Berenguer, D. Sempere-Torres, and M. Zappa, 2009: REAL—Ensemble radar precipitation estimation for hydrology in a mountainous region. *Quart. J. Roy. Meteor. Soc.*, **135**, 445–456, <https://doi.org/10.1002/qj.375>.
- , M. Boscacci, M. Gabella, and S. Maurizio, 2015: Peak performance: Radar design for prediction in the Swiss Alps. *Meteor. Technol. Int.*, 42–45, <https://www.meteoswiss.admin.ch/content/dam/meteoswiss/en/Mess-Prognosesysteme/doc/peak-performance-radar-design-for-prediction.pdf>.

- , J. Figueras i Ventura, M. Gabella, A. Hering, I. Sideris, and B. Calpini, 2016: Triggering innovation. *Meteor. Technol. Int.*, 62–65, <https://www.meteoswiss.admin.ch/content/dam/meteoswiss/de/Mess-und-Prognosesysteme/Atmosferaere/Radarnetz/doc/MTI-April2016-Rad4Alp.pdf>.
- , D. Nerini, I. Sideris, L. Foresti, A. Hering, and B. Calpini, 2017: Real-time radar: A new Alpine radar network. *Meteor. Technol. Int.*, 88–92, <https://www.meteoswiss.admin.ch/content/dam/meteoswiss/en/Mess-Prognosesysteme/Atmosferaere/doc/MTI-April2017-Rad4Alp.pdf>.
- Golding, B. W., 1998: Nimrod: A system for generating automated very short range forecasts. *Meteor. Appl.*, **5**, 1–16, <https://doi.org/10.1017/S1350482798000577>.
- Haiden, T., A. Kann, C. Wittmann, G. Pistotnik, B. Bica, and C. Gruber, 2011: The Integrated Nowcasting through Comprehensive Analysis (INCA) system and its validation over the eastern Alpine region. *Wea. Forecasting*, **26**, 166–183, <https://doi.org/10.1175/2010WAF2222451.1>.
- Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126, <https://doi.org/10.1016/j.physd.2006.11.008>.
- Jacques, D., D. Michelson, J.-F. Caron, and L. Fillion, 2018: Latent heat nudging in the Canadian regional deterministic prediction system. *Mon. Wea. Rev.*, **146**, 3995–4014, <https://doi.org/10.1175/MWR-D-18-0118.1>.
- Johnson, A., X. Wang, J. R. Carley, L. J. Wicker, and C. Karstens, 2015: A comparison of multiscale GSI-based EnKF and 3DVar data assimilation using radar and conventional observations for midlatitude convective-scale precipitation forecasts. *Mon. Wea. Rev.*, **143**, 3087–3108, <https://doi.org/10.1175/MWR-D-14-00345.1>.
- Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. *J. Fluids Eng.*, **82D**, 35–45, <https://doi.org/10.1115/1.3662552>.
- Klasa, C., M. Arpagaus, A. Walser, and H. Wernli, 2018: An evaluation of the convection-permitting ensemble COSMO-E for three contrasting precipitation events in Switzerland. *Quart. J. Roy. Meteor. Soc.*, **144**, 744–764, <https://doi.org/10.1002/qj.3245>.
- Leuenberger, D., and A. Rossa, 2007: Revisiting the latent heat nudging scheme for the rainfall assimilation of a simulated convective storm. *Meteor. Atmos. Phys.*, **98**, 195–215, <https://doi.org/10.1007/s00703-007-0260-9>.
- Lien, G. Y., E. Kalnay, and T. Miyoshi, 2013: Effective assimilation of global precipitation: Simulation experiments. *Tellus*, **65A**, 19915, <https://doi.org/10.3402/tellusa.v65i0.19915>.
- Mandapaka, P. V., and U. Germann, 2010: Radar-rainfall error models and ensemble generators. *Rainfall: State of the Science, Geophys. Monogr.*, Vol. 191, Amer. Geophys. Union, 247–264, <https://doi.org/10.1029/2010GM001003>.
- Nerini, D., N. Besic, I. Sideris, U. Germann, and L. Foresti, 2017: A non-stationary stochastic ensemble generator for radar rainfall fields based on the short-space Fourier transform. *Hydrol. Earth Syst. Sci.*, **21**, 2777–2797, <https://doi.org/10.5194/hess-21-2777-2017>.
- Ott, E., B. R. Hunt, I. Szunyogh, M. Corazza, E. Kalnay, D. Patil, and J. A. Yorke, 2002: Exploiting local low dimensionality of the atmospheric dynamics for efficient ensemble Kalman filtering. *2002 Fall Meeting*, San Francisco, CA, Amer. Geophys. Union, Abstract NG71A-06, <http://abstractsearch.agu.org/meetings/2002/FM/NG71A-06.html>.
- Panziera, L., and U. Germann, 2010: The relation between air-flow and orographic precipitation on the southern side of the Alps as revealed by weather radar. *Quart. J. Roy. Meteor. Soc.*, **136**, 222–238, <https://doi.org/10.1002/qj.544>.
- Penny, S. G., and T. Miyoshi, 2016: A local particle filter for high-dimensional geophysical systems. *Nonlinear Processes Geophys.*, **23**, 391–405, <https://doi.org/10.5194/npg-23-391-2016>.
- Poterjoy, J., 2016: A localized particle filter for high-dimensional nonlinear systems. *Mon. Wea. Rev.*, **144**, 59–76, <https://doi.org/10.1175/MWR-D-15-0163.1>.
- Pulkkinen, S., D. Nerini, and L. Foresti, 2018: pySTEPS: The nowcasting initiative. pySTEPS, <https://pysteps.github.io/>.
- Robert, S., D. Leuenberger, and H. R. Küsch, 2018: A local ensemble transform Kalman particle filter for convective-scale data assimilation. *Quart. J. Roy. Meteor. Soc.*, **144**, 1279–1296, <https://doi.org/10.1002/qj.3116>.
- Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, <https://doi.org/10.1175/2007MWR2123.1>.
- Särkkä, S., 2013: *Bayesian Filtering and Smoothing*. Cambridge University Press, 232 pp., <https://doi.org/10.1017/CBO9781139344203>.
- Schiemann, R., R. Erdin, M. Willi, C. Frei, M. Berenguer, and D. Sempere-Torres, 2011: Geostatistical radar-raingauge combination with nonparametric correlograms: Methodological considerations and application in Switzerland. *Hydrol. Earth Syst. Sci.*, **15**, 1515–1536, <https://doi.org/10.5194/hess-15-1515-2011>.
- Schraff, C., H. Reich, A. Rhodin, A. Schomburg, K. Stephan, A. Periañez, and R. Potthast, 2016: Kilometre-scale ensemble data assimilation for the COSMO model (KENDA). *Quart. J. Roy. Meteor. Soc.*, **142**, 1453–1472, <https://doi.org/10.1002/qj.2748>.
- Seed, A. W., 2003: A dynamic and spatial scaling approach to advection forecasting. *J. Appl. Meteor.*, **42**, 381–388, [https://doi.org/10.1175/1520-0450\(2003\)042<0381:ADASSA>2.0.CO;2](https://doi.org/10.1175/1520-0450(2003)042<0381:ADASSA>2.0.CO;2).
- , C. E. Pierce, and K. Norman, 2013: Formulation and evaluation of a scale decomposition-based stochastic precipitation nowcast scheme. *Water Resour. Res.*, **49**, 6624–6641, <https://doi.org/10.1002/wrcr.20536>.
- Simonin, D., C. Pierce, N. Roberts, S. P. Ballard, and Z. Li, 2017: Performance of Met Office hourly cycling NWP-based nowcasting for precipitation forecasts. *Quart. J. Roy. Meteor. Soc.*, **143**, 2862–2873, <https://doi.org/10.1002/qj.3136>.
- Snyder, C., T. Bengtsson, P. Bickel, and J. Anderson, 2008: Obstacles to high-dimensional particle filtering. *Mon. Wea. Rev.*, **136**, 4629–4640, <https://doi.org/10.1175/2008MWR2529.1>.
- Stephan, K., S. Klink, and C. Schraff, 2008: Assimilation of radar-derived rain rates into the convective-scale model COSMO-DE at DWD. *Quart. J. Roy. Meteor. Soc.*, **134**, 1315–1326, <https://doi.org/10.1002/qj.269>.
- Sun, J., and Coauthors, 2014: Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bull. Amer. Meteor. Soc.*, **95**, 409–426, <https://doi.org/10.1175/BAMS-D-11-00263.1>.
- Thorndahl, S., T. Einfalt, P. Willems, J. E. Nielsen, M.-C. ten Veldhuis, K. Arnbjerg-Nielsen, M. R. Rasmussen, and P. Molnar, 2017: Weather radar rainfall data in urban hydrology. *Hydrol. Earth Syst. Sci.*, **21**, 1359–1380, <https://doi.org/10.5194/hess-21-1359-2017>.
- Tipping, M. E., and C. M. Bishop, 1999: Mixtures of probabilistic principal component analyzers. *Neural Comput.*, **11**, 443–482, <https://doi.org/10.1162/089976699300016728>.
- Todini, E., 2018: Paradigmatic changes required in water resources management to benefit from probabilistic forecasts. *Water Secur.*, **3**, 9–17, <https://doi.org/10.1016/j.wasec.2018.08.001>.

- Tong, M., and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Mon. Wea. Rev.*, **133**, 1789–1807, <https://doi.org/10.1175/MWR2898.1>.
- Venugopal, V., E. Foufoula-Georgiou, and V. Sapozhnikov, 1999: Evidence of dynamic scaling in space-time rainfall. *J. Geophys. Res.*, **104**, 31 599–31 610, <https://doi.org/10.1029/1999JD900437>.
- Wilks, D., 2011a: Principal component (EOF) analysis. *Statistical Methods in the Atmospheric Sciences*, D. S. Wilks, Ed., International Geophysics Series, Vol. 100, Academic Press, 519–562, <https://doi.org/10.1016/B978-0-12-385022-5.00012-9>.
- , 2011b: Time series. *Statistical Methods in the Atmospheric Sciences*, D. S. Wilks, Ed., International Geophysics Series, Vol. 100, Academic Press, 395–456, <https://doi.org/10.1016/B978-0-12-385022-5.00009-9>.
- World Meteorological Organization, 2015: *Seamless Prediction of the Earth System: From Minutes to Months*. WMO, 483 pp.
- Zacharov, P., and D. Rezacova, 2009: Using the fractions skill score to assess the relationship between an ensemble QPF spread and skill. *Atmos. Res.*, **94**, 684–693, <https://doi.org/10.1016/j.atmosres.2009.03.004>.
- Zappa, M., and Coauthors, 2010: Propagation of uncertainty from observing systems and NWP into hydrological models: COST-731 Working Group 2. *Atmos. Sci. Lett.*, **11**, 83–91, <https://doi.org/10.1002/asl.248>.
- Zawadzki, I., J. Morneau, and R. Laprise, 1994: Predictability of precipitation patterns: An operational approach. *J. Appl. Meteor.*, **33**, 1562–1571, [https://doi.org/10.1175/1520-0450\(1994\)033<1562:POPPAO>2.0.CO;2](https://doi.org/10.1175/1520-0450(1994)033<1562:POPPAO>2.0.CO;2).