Doctoral Thesis

Ensemble precipitation nowcasting: limits to prediction, localization and seamless blending

Author(s):
Nerini, Daniele

Publication Date:
2019

Permanent Link:

https://doi.org/10.3929/ethz-b-000391932 →

DISS. ETH NO. 26091

Ensemble precipitation nowcasting: limits to prediction, localization and seamless blending

A thesis submitted for the degree of
DOCTOR OF SCIENCES OF ETH ZURICH
(DR. SC. ETH ZURICH)

presented by

DANIELE NERINI
MSc Hydrology and Water Resources Management, IC London
born 29 May 1988
citizen of Melide TI

accepted on the recommendation of

Prof. Dr. Heini Wernli, examiner
Dr. Loris Foresti, co-examiner
Dr. Urs Germann, co-examiner
Prof. Dr. Marc Berenguer, co-examiner

2019

Contents

Abstract

Nowcasting precipitation, that is, accurately predicting its location and intensity minutes to a few hours ahead, is a difficult task. Extreme rainfall events can be a threat to the population and challenge the limits of weather monitoring and prediction systems. For practical use, the forecasting techniques at our disposal range from physically-based numerical models to heuristic extrapolation procedures based on weather radars. Radar-based nowcasting models rely on the high spatial and temporal resolution of radar measurements and thus benefit from the best possible initial conditions, but the assumption of persistence leads to a rapid decay of predictive skill with increasing lead time and decreasing spatial scale. Numerical models provide physically consistent precipitation forecasts, but their practical relevance for nowcasting can be undermined by uncertain initial conditions, spinup issues, model approximations, or simply by their computational limits.

Analyses presented in this study show that, on average, radar-based nowcasting outperforms numerical simulations during the first three hours and that it is particularly useful for forecasting precipitation patterns with a horizontal dimension below 60 kilometers, as state-of-the-art numerical models cannot provide useful skill at those small spatial scales. An interesting finding is that the **predictive uncertainty** of numerical predictions relatively to radar nowcasting improves during warm convective days, which is explained by the combined effect of shorter precipitation lifetimes and more effective model assimilation of locally triggered air mass convection. After a lead time of 4.5 hours, we observed that precipitation on all scales below 150 kilometers is poorly predictable by all forecasting means.

Such serious limits to predictability determine the need to represent forecast uncertainty as accurately as possible. In radar nowcasting, ensemble methods use stochastic simulations to perturb a deterministic extrapolation and thus quantify forecast errors. We found that a more precise representation of the statistical properties of the precipitation field through **localization** can have a positive impact on the realism of the simulations as well as in terms of probabilistic forecast skill. We showed that localized nowcasts perform better in terms of ensemble reliability and resolution, as well

as conditional bias. However, we also found that a too strong a localization can lead to lower skill as it implicitly relies on a stricter assumption of persistence.
The quantitative estimation of forecast uncertainty provided by ensembles was finally used to design a **seamless blending** procedure that integrates all available sources of predictive skill. Implemented using a recursive formulation of the Bayesian update equations, the blending scheme involves a prediction step through a stochastic radar extrapolation, while a subsequent correction step updates the extrapolation using information from the most recent numerical model run. It is found that such an approach is able to capture the flow dependence of both the numerical forecast and the radar nowcast ensemble spreads resulting in an adaptive blending scheme that depends on the relative uncertainty of the individual forecasts. Despite the non-Gaussian nature of rainfall data, we were able to produce blended precipitation forecasts that are at least as skillful as the radar-only or the numerical model-only forecasts at any lead time.

Sommario

Il *nowcasting* delle precipitazioni, cioè prevederne accuratamente posizione e intensità a pochi minuti o ore, è un compito difficile. Piogge estreme sono una minaccia per la popolazione e rappresentano una sfida per i sistemi di monitoraggio e previsione meteorologici. I sistemi di previsione delle precipitazioni a nostra disposizione si compongono principalmente di modelli numerici basati sulla fisica e tecniche euristiche di estrapolazione delle osservazioni provenienti dai radar meteorologici. La previsione radar si affida alla elevata risoluzione spaziale e temporale delle misure telerilevate in modo da beneficiare delle migliori condizioni iniziali, per quanto l'ipotesi di persistenza durante l'evoluzione nel tempo ne determini un rapido deterioramento. I modelli numerici, invece, forniscono una previsione coerente con le leggi della fisica, ma il loro valore pratico per il nowcasting è limitato dalle maggiori incertezze nelle condizioni iniziali, imprecisioni nel modello stesso, o più semplicemente dai lunghi tempi di calcolo.

Le analisi presentate in questo studio mostrano che, in media, la previsione delle precipitazione tramite radar è superiore al modello numerico durante le prime tre ore di previsione e che il suo uso è soprattutto rilevante per prevedere dettagli al di sotto dei 60 chilometri, dato che una tale qualità della previsione non è ancora raggiungibile dai modelli numerici. È stato interessante notare come i **limiti di previdibilità** migliorino, in termini relativi, utilizzando il modello numerico durante eventi convettivi. Tale effetto è attribuito alla ridotta persistenza delle precipitazioni e maggior efficacia del modello nell'assimilare i dati relativi alla convezione da massa d'aria. Dopo 4.5 ore, abbiamo osservato che tutte le precipitazioni su scale spaziali inferiori ai 150 chilometri sono poco prevedibili per entrambi i sistemi di previsione.

Limiti di prevedibilità tanto severi giustificano il bisogno di rappresentare l'incertezza di previsione in modo accurato. Nel caso della previsione radar, i cosiddetti "metodi di insieme" usano un approccio stocastico per perturbare l'estrapolazione deterministica e così quantificarne l'errore di previsione. Abbiamo scoperto che una più precisa rappresentazione delle proprietà statistiche del campo di precipitazione tramite **localizzazione** determina un miglioramento del realismo e della qualità della previsione probabilistica. Abbiamo dimostrato che la previsione radar localizzata ottiene

un miglioramento in diverse misure di qualità probabilistiche, ma allo stesso tempo osservato che una localizzazione troppo azzardata può produrre il risultato opposto, siccome implicitamente si affida ad una ipotesi di persistenza più marcata.
La stima quantitativa dell'incertezza di previsione fornita dai modelli d'insieme è stata infine usata per elaborare un sistema di previsione capace d'integrare in modo coerente le varie informazioni disponibili. Basandosi su un'implementazione ricorsiva del teorema di Bayes, il sistema si compone di un primo passo di previsione tramite estrapolazione, seguito da un passo di correzione sulla base della più recente corsa del modello numerico. Abbiamo potuto dimostrare che un tale approccio permette di catturare la dipendenza di flusso dell'errore nei modelli di previsione. Il risultato è una **previsione combinata** capace di tenere conto delle singole incertezze. Nonostante la natura non gaussiana dei dati di precipitazione, le previsioni combinate hanno dimostrato una qualità di previsione tanto buona quanto la migliore tra le due stime iniziali prese singolarmente durante tutto l'orizzonte di previsione.

Chapter 1

Introduction

The introduction motivates this thesis and provides some background to its main topics. Section 1.2 presents the two forecasting approaches that are relevant for precipitation nowcasting, namely the extrapolation of weather radar observations and the use of numerical simulations. Section 1.3 is dedicated to a review of the forecast uncertainties in both nowcasting methods and thus introduces the need for probabilistic forecasting and blending procedures, which are the subjects of Section 1.4 and Section 1.5, respectively. Section 1.6 lists the main research objectives and outlines the subsequent chapters, while Section 1.7 concludes the introduction with a brief note on the scope of this research work.

1.1 The importance of nowcasting precipitation

Precipitation extremes, in terms of both positive and negative anomalies, represent the main driving force in the occurrence and amplitude of many water-related natural hazards. With growing concern for the adverse effects of changing climate and land use patterns, the risk associated with such natural phenomena is further set to rise.
As natural extremes cannot be prevented, the use of early warning systems has been recognized as an essential measure of risk reduction (Pappenberger et al., 2015). In this context, risk mitigation is achieved by providing sufficient warning time to reduce exposure and vulnerability to natural hazards (Sättele et al., 2012). It is thus not surprising that decision makers have become keen consumers of weather-related information, in particular concerning quantitative forecasts of precipitation events (Cuo et al., 2011; Alfieri et al., 2012).
To ensure effective warnings, the forecast must satisfy a set of requirements that ultimately depend on the natural hazard process and its scale of interest. Figure 1.1 illustrates the typical spatial and temporal scale of a number of hydrological extremes. It is easy to see how local phenomena such as debris flows occur on shorter time scales

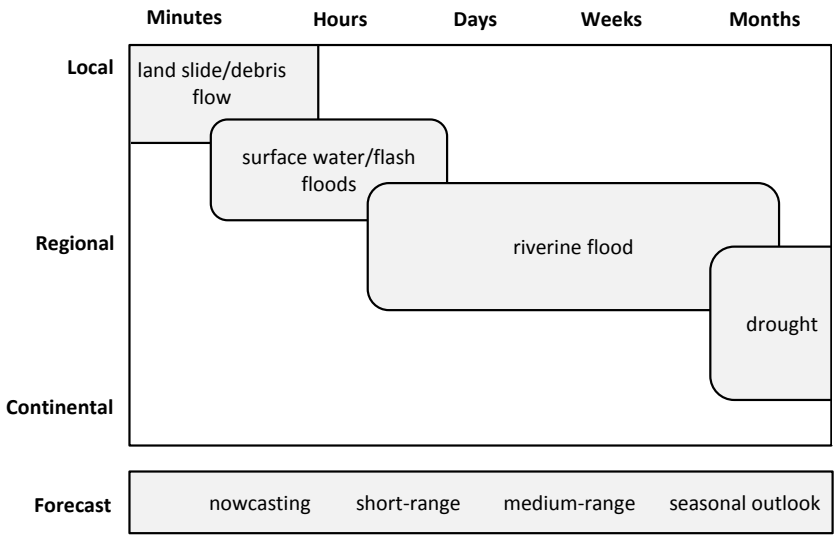


Figure 1.1: Space-time scale of interest for hydrological extremes and associated forecast range. Adapted from Alfieri et al. (2012).

than, for example, riverine flooding, and consequently define higher requirements on the level of detail that is to be provided to achieve risk mitigation.
The limited predictability of the atmosphere is a critical element to consider with respect to weather predictions. To put it simply, it is more difficult to provide accurate forecasts of local phenomena. This fact sets a hard limit on the forecast range that can be achieved for the accurate prediction of local natural phenomena such as flash floods or debris flows. The bottom panel in Fig. 1.1 indicates the typical forecast range for the space-time scale of interest associated with a given hydrological extreme.
We thus note that the nowcasting range is highly relevant for applications that are concerned with sub-regional spatial scales and time scales in the order of minutes to a few hours ahead. It is also important to note that the increasing availability of communication technologies and the development of automatic warning procedures is determining a reduction of warning time requirements. As a result, the demand for high-resolution forecasts targeting local, small-scale natural phenomena is on the rise (Mass, 2012).

1.2 Forecasting approaches for precipitation nowcasting

In the nowcasting range (typically 0 to 6 hours, Wang et al., 2017), precipitation is traditionally predicted using heuristic extrapolation techniques based on observations from weather radars. More recently, the advances in horizontal resolution, data assimilation and model physics in limited area NWP models have determined an interest for the application of numerical forecasts in the nowcasting range, too.

For the sake of completeness, we note that Wilson et al. (1998) listed a third approach for nowcasting precipitation, namely the use of conceptual or statistical models of the evolution of precipitation. This class of algorithms has so far found little use in precipitation nowcasting applications, perhaps because of the limited success in improving over simpler techniques based on persistence (e.g., Grecu and Krajewski, 2000) and it is thus not discussed further in this work. However, we note the existence of a new interest for such methods, which is mainly driven by the achievements of so-called *deep-learning* methods. At the end of this thesis (Section 5.1), we provide some considerations with respect to this topic.

1.2.1 Radar-based nowcasting

Radar-based precipitation nowcasting is an empirical forecasting procedure that is typically based on the assumption of persistence in the storm's frame of reference [*Lagrangian persistence*, as in Zawadzki et al. (1994), *status-quo* or *steady-state* assumption, as in Bellon and Austin (1978) and Wilson et al. (1998), respectively]. The conservation equation for an incompressible flow can be written as in Germann and Zawadzki (2002) by denoting a precipitation parcel by R and its displacement vector during time interval τ by $\boldsymbol{\alpha}(\tau)$

$$R(\boldsymbol{x}_0, t+\tau) = R(\boldsymbol{x}_0 - \boldsymbol{\alpha}(\tau), t), \tag{1.1}$$

or equivalently in differential form as

$$\frac{dR}{dt} = \frac{\partial R}{\partial t} + u\frac{\partial R}{\partial x} + v\frac{\partial R}{\partial y}, \quad u = \frac{dx}{dt}, \quad v = \frac{dy}{dt}, \tag{1.2}$$

where $dR/dt = 0$, and u and v are the x- and y-components of the apparent motion of the rainfall field (or echo motion).

In practice, the Lagrangian persistence approach consists of an extrapolation or advection forecast, in which u and v are estimated for a given location by solving equation (1.2) numerically based on a sequence of precipitation intensity fields. Typically, a constraint on the spatial continuity of nearby u and v is imposed to guarantee a unique solution. Once the motion field is known, the radar echoes are extrapolated by means of an advection scheme. Such a procedure is computationally efficient, allowing to fully exploit the high temporal sampling of weather radars.

Weather radars are ideally suited for providing the input data for precipitation nowcasting at high resolution (Berne et al., 2004). 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 surface. 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).
The extrapolation of radar echoes for QPF was first proposed by Ligda (1953) and later by Wilson and Kessler (1963), who considered that "the echo-area movement arrows on the [radar] charts, based on the observed movement during the preceding hour, are valuable for predicting the location of the areas [of precipitation] 3, 6 and 9 hr later". Early operational implementations can be traced back to the McGill weather radar in Montreal in 1976, when for the first time three-hourly precipitation forecasts were automatically produced by the geometrical translation of 3-km CAPPI[1] maps along the overall storm motion (Bellon and Austin, 1978, see also Fig. 1.2a). Later, Rinehart and Garvey (1978) presented a cross-correlation approach named TREC to estimate the three-dimensional motion of a storm. The study represented a first use of a pattern recognition technique for the estimation of the differential motion field within radar observations.
In computational sciences, the approximation of the apparent motion from a sequence of images is known as *optical flow* (e.g., Fleet and Weiss, 2006). This represents an important and broad class of algorithms which has found countless applications. Among these, many studies have used optical flow methods for precipitation nowcasting purposes. Global variational methods (e.g., Proesmans et al., 1994), with or without a multiscale approach, were proposed in Laroche and Zawadzki (VET, 1994) and Li et al. (COTREC, 2002), while local optical flow techniques (e.g., Lucas and Kanade, 1981; Farnebäck, 2003) were for example applied in Hohti et al. (2000) and Liu et al. (2015). Open-source implementations of optical flow methods for nowcasting applications have emerged in the past years, as Com- SWIRLS by the Regional Specialized Meteorological Centre (RSMC) for Nowcasting operated by the Hong Kong Observatory (HKO) (Wong et al., 2016), the rainymotion library at the University of Potsdam (Ayzel et al., 2019), and the community-based nowcasting library pysteps (Pulkkinen et al., 2019b, see also Fig. 1.2b).

1.2.2 Convective-scale NWP

By simulating the thermo-dynamical evolution of the atmosphere, numerical weather prediction (NWP) models provide physically-consistent forecasts for all relevant hydrometeorological variables and are thus ideal candidates for improving the effectiveness of early warning systems (Cuo et al., 2011; Alfieri et al., 2012).
NWP models are increasingly able to meet the nowcasting requirements in terms of skill at the convective scale, owing to constant improvements in terms of model

[1] constant-altitude-plan-position-indicator

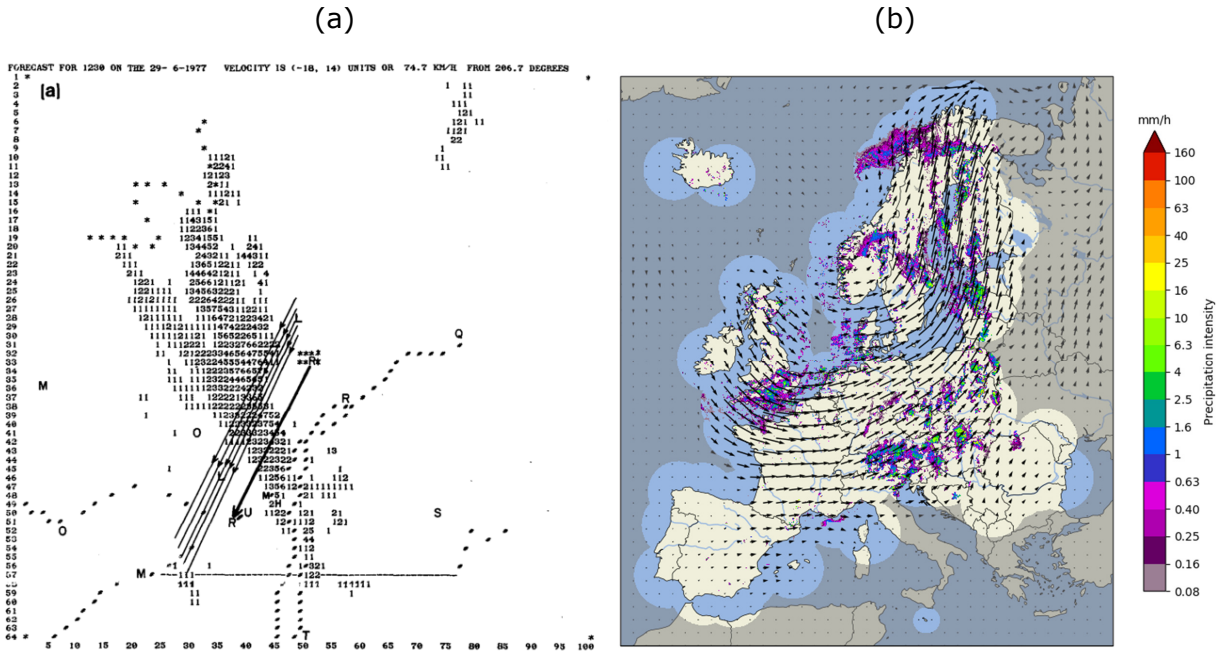


Figure 1.2: Radar-based nowcasting yesterday and today. (a) An early application at the McGill weather radar in Montreal; reproduced from Bellon and Austin (1978), their Fig. 3a. (b) An example using the open-source nowcasting library *pysteps* (Pulkkinen et al., 2019b) with the European composite OPERA valid at 1830 UTC 24 August 2018.

resolution, data assimilation, and physic parameterization. Sun et al. (2014) give an overview of the achievements and challenges in this field.

Nowadays, typical applications for (very) short-term precipitation forecasts use embedded, limited-area (mesoscale), high-resolution and non-hydrostatic models with grid point representation of the meteorological variables at grid sizes of 1-10 km. Compared to global spectral models, such configuration allows for a better representation of the small-scale processes and forcings that are crucial for the accurate forecasting of local weather as the vertical accelerations in deep convective cells (Clark et al., 2016), and are thus also known as *convection-permitting models* (CPMs). It is also generally expected that the greater detail of convective-scale simulations will feed back to the whole atmospheric system and eventually provide improvements for a wide range of spatial and temporal scales (Zeng et al., 2019).

For practical applications, CPMs are often run using rapid update cycles (RUC) in combination with data assimilation (DA) schemes, so that the forecast can be frequently updated using the latest available observations. A recent review article by

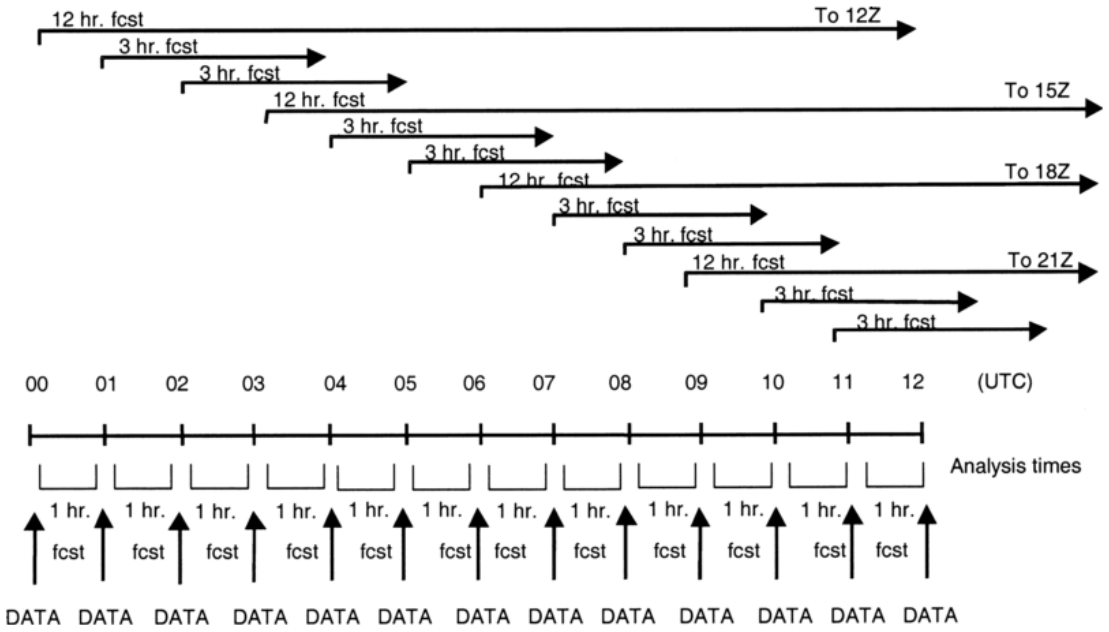


Figure 1.3: Overview of a 1 h RUC operational scheme in the early 2003 at NCEP. Reproduced from Benjamin et al. (2004), their Fig. 1. Rapid-update cycling is the key to reduce spinup time in NWP-based nowcasting applications.

Gustafsson et al. (2018) noted the benefits of DA for the initialization of mesoscale models. In this regard, radar DA is often seen as crucial for the successful prediction of precipitation events (e.g., Sun et al., 2012), which has been motivated by the fact that Doppler radar measurements constitute a significant data source about the spatial distribution of wind and precipitation at the convective scale (Sun, 2005; Craig et al., 2012). Convective-scale DA methods currently in operation are surveyed in Gustafsson et al. (2018) and include variational methods (3D- and 4D-Var), ensemble methods based on the ensemble Kalman filter (EnKF, Evensen, 1994), and diabatic initialization methods such as latent heat nudging (e.g., Stephan et al., 2008).

Non-hydrostatic models were first introduced in the mid-1970s as research tools to study individual convective systems, as in Miller and Pearce (1974), and other mesoscale phenomena. In 1989 the Center for Analysis and Prediction of Storms (CAPS) was established at the University of Oklahoma to demonstrate the practicability of storm-scale forecasting. The result of their studies is one of the first CPM-based forecasting systems: the Advanced Regional Prediction System (ARPS) (Xue et al., 2003). The first operational RUC was established in 1994 at the National Centers for Environmental Prediction (NCEP) with a 3-h assimilation cycle, which was updated to an hourly cycle in 2002 (Benjamin et al., 2004), see also Fig. 1.3. Operational CPM-based RUC forecasting system with radar DA are nowadays used in all major national weather services and are particularly important for issuing high-impact weather alerts.

1.3 Review of nowcasting uncertainties

There is a fundamental limit to what is possibly predictable in a real system such as the atmosphere. This is known as *intrinsic* predictability and is a consequence of the fact that a real system cannot be observed with absolute precision in one instantaneous state nor expressed without approximations in its governing laws (Lorenz, 1996). Because the atmosphere is a divergent system, small differences between two close initial states will grow exponentially. Such behavior is commonly referred to as *chaos*. Hence, the chaotic nature of the atmosphere determines the exponential growth of small errors, which eventually set a limit to its intrinsic predictability.
In practice, these errors are related to the specific assumptions and limitations of the forecasting procedure under study. The growth of errors which originate from a given forecasting method is referred to as the *practical* predictability. This section focuses on the main factors limiting the practical predictability of the radar-based and NWP-based approaches to precipitation nowcasting.

1.3.1 Errors in the initial conditions

The first important limit to predictability is given by errors in the initial conditions, often called observation errors. In this sense, initial condition errors are a consequence of the fact that our capacity to measure the current state of the atmosphere is limited by the sparseness, precision, and accuracy of observations.
The initialization problem is also closely related to the definition of the model state under study and its relationship to the observation space. Therefore, it makes perhaps more sense to refer to a more general *model initialization uncertainty*, as this concept includes errors in both measurements and their assimilation into the model.

Initial errors in radar-based nowcasting

The radar-based extrapolation approach to nowcasting has the advantage of considering surface precipitation as a prognostic variable, thus considerably reducing the degrees of freedom of the model initialization problem. The uncertainty of the initial conditions can be ascribed to errors in the radar QPE, which were estimated to account for approximately half of the total forecast uncertainty in the first hour into the nowcast (Browning et al., 1982; Bellon and Austin, 1984; Fabry and Seed, 2009).
Assuming a stable and well-calibrated radar system, Fabry (2015) divided radar QPE errors in three broad categories: (a) errors in radar measurements aloft – mainly driven by attenuation and beam broadening, (b) errors in extrapolating aloft radar quantities at ground level – a problem mainly related to the estimation of the vertical profile of reflectivity (VPR), and (c) errors in converting the extrapolated radar quantities at

the ground into rain rates, where the natural variability of the rain drop-size distribution (DSD) represents an important source of uncertainty in the empirical relationship between reflectivity and rain rates (Z-R relation). Additional errors can relate to the presence of non-meteorological echoes and the beam shielding, which represent particular challenges for real-time applications in mountainous regions (Germann et al., 2006a). Extensive reviews of the sources of uncertainty in single polarization radar-based QPE can be found in Wilson and Brandes (1979), Joss et al. (1990), and Villarini and Krajewski (2010), while Berne and Krajewski (2013) and Thorndahl et al. (2017) focus on the implications for hydrological applications.

Mandapaka and Germann (2010) provided a comprehensive review of radar-rainfall error models and ensemble generators. Several authors investigated so-called *source-specific* error models (e.g., Jordan et al., 2000, 2003; Berenguer and Zawadzki, 2008) by characterizing individual sources of errors and then superposing them to produce a model of the overall uncertainty. Because of the complex structure of all the individual error contributions and their interactions, another common approach is to characterize the total residual error in radar QPE against ground reference observations (e.g., Ciach et al., 2007; Habib et al., 2008; Germann et al., 2009; AghaKouchak et al., 2010). The empirically derived error structure can then be used to generate stochastic ensembles of precipitation fields (e.g., Germann et al., 2009; Cecinati et al., 2017). As highlighted by Berenguer and Zawadzki (2008), the total error approach is very convenient, but it bears its own limitations, as the mismatch in sampling volume between rain gauge and radar estimates and the need to interpolate the point residuals.

Finally, the initial condition uncertainty in radar-based nowcasting can also include errors in the estimation of the echo motion field. For nowcasting applications, the optical flow problem becomes particularly ill-posed and ambiguous, as it is difficult to separate the pure motion component of precipitation at the surface from changes due to growth and decay processes. Besides, the estimation can be affected by radar artifacts such as residual ground clutter or the limited visibility of the radar network. Error models for optical flow algorithms were proposed by Proesmans et al. (1994), based on the consistency between two motion fields computed forward and backward from the same sequence of images, and more recently by Kybic and Nieuwenhuis (2011), based on a bootstrap resampling approach. An interesting recent finding by Pulkkinen et al. (2019b) is that the application of three fundamentally different optical flow techniques provided very similar forecast skill. The authors concluded that for nowcasting applications, the optical flow uncertainty plays a less significant role compared to other sources of forecast error.

Initial errors in CPM-based nowcasting

The initialization of high-resolution NWP models requires an accurate description of the current state of the atmosphere at the convective scale. In a so-called *cold start*, initial conditions are provided by simply downscaling the analysis of a coarser model state, so that processes at the convective scale need to be "spun up" by the numerical simulations, which typically takes 3 to 6 h (Sun et al., 2014). The spinup time is obviously problematic for nowcasting applications, so most forecasting applications use a *warm start*, which consists in providing initial conditions for all prognostic variables and at the same scales resolved by the numerical model.
A warm start is typically produced with a data assimilation (DA) scheme which bears the difficult task of combining the prediction from a previous model run, the first guess, with an incomplete and inaccurate set of observations of the many atmospheric prognostic variables in order to produce a comprehensive model initial state, the analysis. In situ observations provide accurate measurements of prognostic variables such as temperature or pressure. On the other hand, such measurements are typically affected by the sparseness of the measuring network and its limited representativeness. Conversely, remote sensing observations can guarantee high measurement density at spatial scales matching those of the model, but their accuracy is generally lower. Moreover, remotely-sensed variables often relate to model prognostic variables in complex and strongly non-linear ways, which greatly affects the overall model initialization uncertainty.
In the case of precipitation, which is only an indirect output of numerical modeling (i.e., a diagnostic variable), radar DA has achieved mixed results and still undergoes active research (Sun, 2005). Diabatic initialization techniques offer an efficient way to assimilate surface radar reflectivities, reduce precipitation spinup times and increase the NWP skill in the nowcasting range (e.g., Jacques et al., 2018), but their success is limited by the capacity to simulate and support the observed rainfall intensities. Stephan et al. (2008) linked such problems to imperfections in the physics parameterization and errors in the large-scale environmental conditions of the background. Craig et al. (2012) investigated the impact of LHN and found that the assimilation of radar reflectivities had longer-lasting benefits during non-equilibrium convection events, that is, cases of weak synoptic forcing. 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 non-linear relationship between radar observations and model state (Tong and Xue, 2005; Dowell et al., 2011) and it is therefore often restricted to research activities (Gustafsson et al., 2018).

1.3.2 Model errors

The second relevant limit to practical predictability relates to sources of error within the forecasting model itself. Model errors thus enclose all uncertainties that can be associated with the assumptions and limitations in the model representation of the atmospheric dynamics and physics (Leutbecher and Palmer, 2008).

Model errors in radar-based nowcasting

Germann et al. (2006b) defined model errors in the Lagrangian persistence approach to nowcasting solely as the inaccuracies due to the practical implementation of the method, as the discretization in time, space and reflectivity, or the numerical diffusion of the advection scheme. However, here we also include errors that derive from the assumptions of the forecasting method itself, as already considered by Browning et al. (1982).

By following this definition, the main source of model errors in an extrapolation approach is intuitively related to the assumption of persistence of the rainfall field in Lagrangian coordinates. The extrapolation forecast can only represent temporal variations due to advection and consequently cannot capture the evolution of rainfall in terms of initiation, growth, decay, and dissipation processes. The limitations of advection forecasts were very well documented in the pioneering studies of Bellon and Austin (1978, 1984) and Browning et al. (1982), which showed that the predictive skill was clearly limited to the nowcasting range.

In addition, practitioners quickly recognized that removing noisy small-scale patches of precipitation would lead to improved forecasts. In their nowcasting procedure, Browning et al. (1982) applied an upscaling of the radar field to remove unpredictable scales of motions, while Mecklenburg et al. (2000b) found that the smoothing of the radar images produced better forecasts for convective precipitation events over Switzerland. Similarly, Grecu and Krajewski (2000) found that the predictability of extrapolation forecasts improved with coarser resolutions. Bellon and Zawadzki (1994) formalized and extended these intuitions by parameterizing spatial averaging as a function of lead time to filter out unpredictable small-scale precipitation features.

These practical considerations were preceded and supplemented with theoretical studies, some of which are of difficult access now (e.g., Wilson, 1966). In his seminal work, Zawadzki (1973) formulated a theoretical framework to study the statistical properties of precipitation patterns in the Lagrangian frame of reference which would later be adopted by several predictability studies based on radar observations (e.g., Zawadzki et al., 1994; Germann and Zawadzki, 2002, 2004; Germann et al., 2006b; Turner et al., 2004; Radhakrishna et al., 2012; Surcel et al., 2015). Within the field of multifractals (see Foufoula-Georgiou and Krajewski, 1995, for a review), Venugopal et al. (1999)

studied the variability of precipitation across spatial and temporal scales and found a power-law relationship consistent with the hypothesis of dynamic scaling, that is, the space (L) and time (t) organization of rainfall can be described by a simple scaling relationship in the form of $t \sim L^z$, with z being the dynamic scaling exponent.

It is thus clear that the predictive skill of radar-based nowcasting is mainly limited by the fact that the Lagrangian approach neglects the (scale-dependent) temporal evolution of precipitation patterns. Consequently, several authors aimed at characterizing the systematic biases associated with areas, times of the day, and flow conditions of preferential enhancement or dissipation of precipitation. For instance, Foresti and Seed (2015) investigated the effects of orography on the spatial distribution of the nowcasting errors and found a systematic underestimation on the windward side, followed by overestimation on the leeward side. Foresti et al. (2018) found similar patterns of systematic growth and decay over the Swiss Alps and highlighted the importance of environmental variables such as freezing level height, used as an indication of the stability of the air mass. Mandapaka et al. (2012) concluded that the lower predictability of radar extrapolation techniques over the Alps (around 3.5 h) compared to the US (5 h, Germann et al., 2006b) was also to link to the presence of predominantly complex terrain. Concerning systematic errors in time, Berenguer et al. (2012) noted a clear dependence on initialization time, which they attributed the effect of the diurnal cycle of precipitation, and Atencia et al. (2017) provided a systematic analysis of forecast biases as a function of time of the day, spatial scale and location.

In addition to the Lagrangian assumption of precipitation patterns, we also note that the extrapolation approach assumes temporal stationarity of the motion field itself, that is, constant displacement vectors during the forecast. Bowler et al. (2006) analyzed the errors due to changes in the motion field and found that these are negligible for the first 3 h, and only account for around 10% of the forecast error at +6 h. In a similar analysis, Germann et al. (2006b) reported a case-to-case variability of the relative importance of the changes in the storm motion field, but also noted that the contribution does not overcome the limit set by unpredictable growth and decay.

Finally, the advection routine used to extrapolate the radar image can represent a more or less important source of model errors, depending on the choice of the scheme. Germann and Zawadzki (2002) performed a detailed analysis of the problem, which is here summarized. In forward advection schemes, that is forward in time and downstream in space, precipitation at the origin needs to be redistributed to grid points at destination where often some sort of spreading is necessary to avoid artifacts in areas of diverging flow. In backward schemes, an interpolation is instead required at the origin to determine the precipitation value at destination, which means that the backward scheme is not, strictly speaking, mass conservative. Both spreading and in-

terpolation result in numerical diffusion, namely a loss of power at small scales. The authors concluded that a modified version of the semi-Lagrangian backward scheme by Robert (1981) provided the best performance. The modified routine reduces to one the number of interpolations needed, thus minimizing numerical diffusion, while flow divergence is limited by a smoothness constraint in the optical flow estimation so that the backward scheme can be considered approximately conservative in mass.

Model errors in CPM-based nowcasting

Model errors associated with the numerical approximations of the atmospheric dynamics and the parameterization of subgrid physical processes account for an important contribution to the overall forecast uncertainty in NWPs (Weisheimer et al., 2011; Shutts and Pallarés, 2014). Model errors are believed to become even more important for short-term and high-resolution applications (Baker et al., 2014; Zeng et al., 2019), especially in terms of biases (Clark et al., 2016). Besides, their importance was found to vary depending on the weather regime (e.g., Groenemeijer and Craig, 2012; Kober and Craig, 2016).

CPMs can generally resolve most of the convective-scale motion, which in principle should remove the need for deep convection schemes. However, other unresolved physical processes still rely on parameterization, in particular at the microscale. For instance, the microphysical processes driving the formation of clouds and precipitation are approximated by parameterization schemes that are notably associated with great uncertainty (e.g., Gilmore et al., 2004).

It should also be noted that as model resolution increases, it might enter a so-called *gray zone*, that is, a range of scales where a given physical process is only partially resolved. The presence of a gray zone poses a great challenge to the formulation of model physics and consequently affects model errors (Berner et al., 2017). Convective parametrizations, for example, aim at representing all the convective flow in the model, but the practical difficulty to separate it from the large-scale vertical motion might question the role that convective parameterization schemes play in CPMs (Clark et al., 2016). In this regard, Yu and Lee (2010) analyzed the interactions between convective parameterization and model resolution and found that only horizontal grid spacings of 3 km or below were sufficient to resolve convection bands fully. The authors also noted that for lower resolutions, false triggering of parameterized convection could in some cases suppress convection that was resolved by the model.

1.4 The ensemble approach to forecasting

As seen above, initial and model errors are amplified by the chaotic nature of the atmosphere, eventually determining a limit to predictability. By recognizing that a unique, deterministic solution to the weather prediction problem is impossible, the NWP community embraced the probabilistic approach to forecasting (Slingo and Palmer, 2011), which has been acknowledged as one of its greatest achievements (Bauer et al., 2015). Following the pioneering works of Epstein (1969) and Leith (1974), the representation of the forecast uncertainty was introduced as the stochastic perturbation of the set of initial conditions (e.g., Toth and Kalnay, 1993). The key idea behind ensemble forecasting is to provide random samples from the probability density function (PDF) of the *current* model state. These are then integrated forward in time to produce a set of equally-probable solutions, known as the members of the ensemble forecast. In this way, the model PDF is implicitly advanced and its *future* state can be estimated probabilistically (Zhang and Pu, 2010).
Similarly to the approach of perturbing the initial conditions, the representation of model uncertainties aims at providing multiple realizations of either the parameterization scheme itself (the multimodel method), the total tendency of the parameterization scheme (SPPT, Buizza et al., 1999) or the individual parameters themselves (e.g., Baker et al., 2014). Such perturbations can be applied either once at the beginning of the forecast or periodically throughout the model integration, so to explore the model parameter space better (Baker et al., 2014).
NWP-based ensemble prediction systems (EPS) are typically composed of 20-50 members at relatively lower resolution in order to compensate for the increased computational demand. Although successfully employed for years within coarser NWP global models, ensemble forecasting has only recently become a viable approach for CPMs thanks to increased supercomputer capability. For example, the Swiss national weather service MeteoSwiss introduced in early 2016 COSMO-E, a 21-member limited-area EPS with convection-permitting horizontal resolution (2.2 km), thanks to the breakthrough in performance that followed the introduction of GPUs technology for the numerical computations (Fuhrer et al., 2014; Lapillonne et al., 2016). In a comparative study with the global ECMWF EPS model (32 km resolution), Klasa et al. (2018) showed that COSMO-E was able to outperform its global counterpart in forecasting precipitation and that it provided a better spread-skill consistency.
Inspired perhaps by the NWP achievements in the field, the radar nowcasting community acknowledged the importance of quantifying predictive uncertainty, too. At first, heuristic techniques were developed to complement deterministic extrapolations with point exceedance probabilities. In the so-called *local Lagrangian* approach (Germann and Zawadzki, 2004), the rainfall intensities in the neighborhood of the pixel

to be extrapolated are used as an approximation of its future PDF. The method thus assumes that spatial variability around a given point of interest is a good indication of its forecast uncertainty.

Besides applications where predictions of *point* rainfall rates are needed, probability forecasts as presented above might be of limited use. This is because they cannot provide information about the *joint* PDF of precipitation, that is, they do not consider the existence of spatial and temporal dependencies in the rainfall field.

Furthermore, if one considers that the potential of radar nowcasting is first and foremost as forcing for hydrological applications (Berne and Krajewski, 2013; Thorndahl et al., 2017), then it makes sense to devise ways to quantify and *propagate* uncertainties throughout the whole modeling chain (Rossa et al., 2010; Zappa et al., 2010). As highlighted in Berenguer et al. (2011), the ensemble approach provides the best (and perhaps only) way to convey uncertainty and at the same time respect the spatiotemporal structure of the variable of interest.

Analogously to the use of model error schemes in NWP EPS, most radar ensemble nowcasting systems are based on stochastic simulations in which fields of correlated noise are used to perturb a deterministic nowcast and thus simulate the loss of predictability for increasing lead time (e.g., Bowler et al., 2006; Metta et al., 2009; Berenguer et al., 2011; Atencia and Zawadzki, 2014; Pulkkinen et al., 2019b). The relative simplicity of these models allows for the simulation of large ensembles. Using 8 CPU cores, Pulkkinen et al. (2019b) showed that they were able to produce a +3 h, 48-member ensemble nowcast at 1 km and 5 min resolutions over a 710x640 grid in around three minutes.

1.5 Radar-NWP blending

If radar-based nowcasting is designed to take advantage of the high spatiotemporal resolution of radar measurements and thus benefit from the best possible initial conditions, the assumption of Lagrangian persistence determines a rapid decay of its skill with increasing lead time and decreasing spatial scale. The NWP approach provides physically consistent precipitation fields, but its practical relevance in the nowcasting range can be undermined by outdated initial conditions, spinup issues, model approximations, or more simply by the computational limits of high-resolution numerical simulations.

We thus note that no single forecasting procedure can provide the best forecast in the whole nowcasting range, across all spatial scales, and for all weather situations. Instead, there is the potential to produce a better prediction by integrating the information from the individual forecasts. Blending is intended as a data integration

procedure that aims at maximizing the skill of a combined forecast by weighting the single components by their performance. From the point of view of NWP, blending can be seen as a post-processing step aimed at updating NWP forecasts with real-time observations. More generally, the term *seamless* is often used when blending multi-source forecasts to indicate a consistent prediction regardless of location, lead time, or forecasting procedure (Brunet et al., 2015).

Precipitation nowcasting systems have traditionally relied on deterministic blending schemes to increase their forecast range. In most cases, blending consisted of a simple linear-weighting function in time to produce a smooth transition from the radar extrapolation to a deterministic NWP run. A first example of an operational scheme using such an approach is the Nimrod system at the UK Met Office (Golding, 1998). Technically, the approach corresponds to an image cross-dissolve operation, namely a linear interpolation in the rain rate space between two QPF fields, where weights are typically pre-computed based on long-term forecast error analysis and then applied in real time. The lead time that marks the transition to a larger weight for the NWP compared to the weight assigned to the extrapolation nowcast is usually referred to as *crossover lead time*. The technique is simple and robust, which explains its appeal for operational use and the fact that it is still employed for blending in several nowcasting systems (e.g., Haiden et al., 2011; Ridal et al., 2011; Li and Lai, 2004). On the other hand, it cannot account for variability in both the NWP and the radar extrapolation skills, resulting in a sub-optimal blended forecast. Most notably, 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 over the past years, as the short-term ensemble prediction system (STEPS, Bowler et al., 2006; Seed et al., 2013). STEPS quantifies the scale-dependent 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. Liguori et al. (2012) found that STEPS blended nowcasts provided an extended lead time for the forecasting of flow in urban catchments.

Atencia et al. (2010) aimed at considering the spatial variability of the forecast skill. The authors recognized the limitation of linear cross dissolve techniques in terms of spatial uniformity of the weighting and introduced the concept of spatial blending in order to represent better the skill of NWP in generating new areas of precipitation. Spatial blending is a helpful concept, particularly for mountainous regions, where radar-based nowcasting systems are expected to perform poorly in areas of orographic forcing. In order to account for this effect, Yu et al. (2015) tried to distinguish between

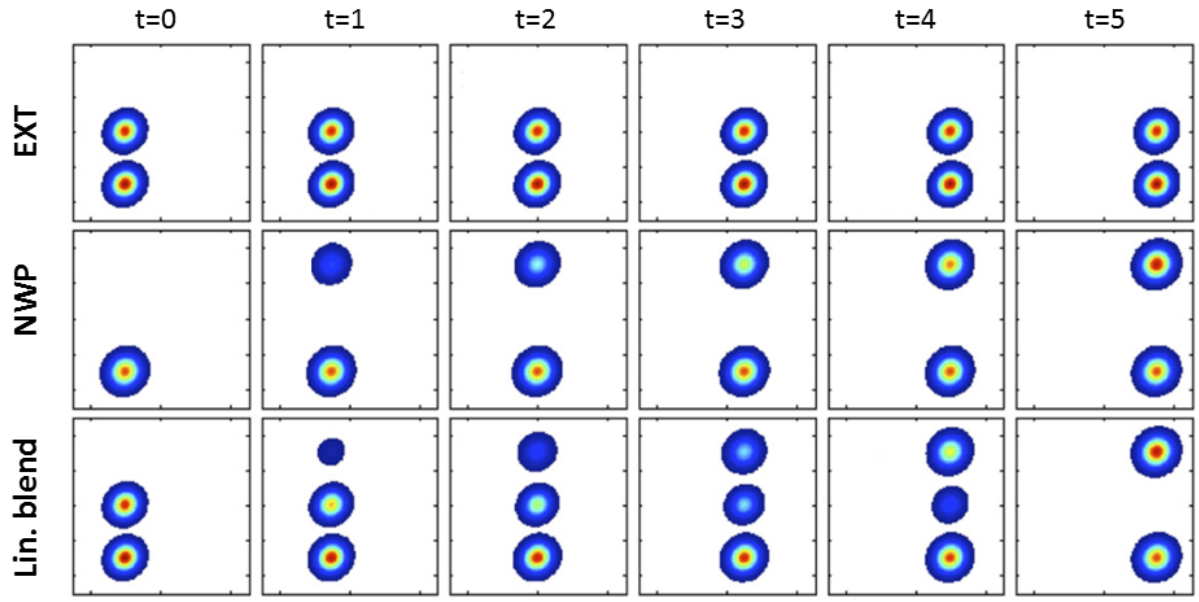


Figure 1.4: Example of double image effect and smoothing during a linear blending procedure of idealized storm cells in an extrapolation (EXT) and numerical model (NWP) forecasts. Adapted from Hwang et al. (2015), their Fig. 2.

orographic and non-orographic rainfall so that different blending weights could be applied.

A critical challenge in blending precipitation fields is the so-called *double image effect*, which appears when combining objects that are not aligned in both precipitation forecasts because of space-time displacement errors. This effect typically results in a blended forecast that displays a larger fraction of rain and lower precipitation intensities, i.e. a *smooth* field. The double image effect is illustrated in Fig. 1.4 with a sequence of idealized storm cells in extrapolation and NWP-based nowcasts where the tapering of peak intensities in case of non-aligned objects is clearly visible. Hwang et al. (2015) proposed a morphing approach to try aligning the objects in the two fields before the cross-dissolve blending. Similar techniques have previously been investgated in data assimilation (e.g., Beezley and Mandel, 2008). Thus, it can be seen that the blending of nowcasting sources remains a challenging and applied field of study.

1.6 Aim and outline

This thesis aims to contribute to a better understanding of precipitation nowcasting methods and their synergies within the applied context of the Swiss national weather

service MeteoSwiss. The work is underpinned by the following research questions which are addressed in the three main chapters.

(a) The **limits to predictability** of precipitation patterns depends on spatial scales. It is, therefore, of particular interest to adopt a verification framework based on scale decomposition to investigate to what extent high-resolution numerical simulations can provide useful skill compared to traditional radar nowcasting (Chapter 2).

(b) Besides scale dependence, the statistical properties of precipitation can vary in space and thus determine the heterogeneous distribution of the forecast uncertainty. This rises to the question of whether a **localization** approach can improve radar nowcasting and the ability to quantify the forecast error (Chapter 3).

(c) The differences in terms of practical predictability among nowcasting procedures suggests an integrated and **seamless blending** approach to precipitation forecasting. The final part of this thesis (Chapter 4) is thus dedicated to understanding the benefits of a Bayesian approach to the blending of radar and NWP-based precipitation nowcasts.

1.7 Research scope

This thesis focuses on hydrometeorological predictions within the nowcasting range. Consequently, the primary variable of interest is precipitation rates, typically denoted as R. The term precipitation is used to refer to the flux of liquid and solid phases of precipitating hydrometeors at the surface and expressed in $\mathrm{mm\,h^{-1}}$of (equivalent) liquid water. Despite strictly referring to liquid rain only, rainfall is loosely adopted here as a synonym for precipitation.
The term nowcasting is used to refer to the forecast range from 0 to around +6 h. In the context of this thesis, radar nowcasting implies an advection or extrapolation forecast of the radar QPE field. We do not discuss object-based nowcasting methods such as cell tracking, which are more relevant for hazards related to severe convection (e.g., hail, wind gusts, or lightning).
Most of our research work is conducted using data from the MeteoSwiss weather radar QPE composite. With an area of 454,400 $\mathrm{km^2}$, the Swiss radar domain covers the whole of Switzerland, Liechtenstein, and the neighboring regions of Germany, France, Italy and Austria (Fig. 1.5). Besides, the numerical forecasts from the COSMO-1 and COSMO-E models operated at MeteoSwiss since May 2016 were retrieved from an archive of operational runs and interpolated onto the radar composite grid. The

model domain for COSMO-1 and COSMO-E covers a broader area enclosing the whole Alpine region (Fig. 1.5).

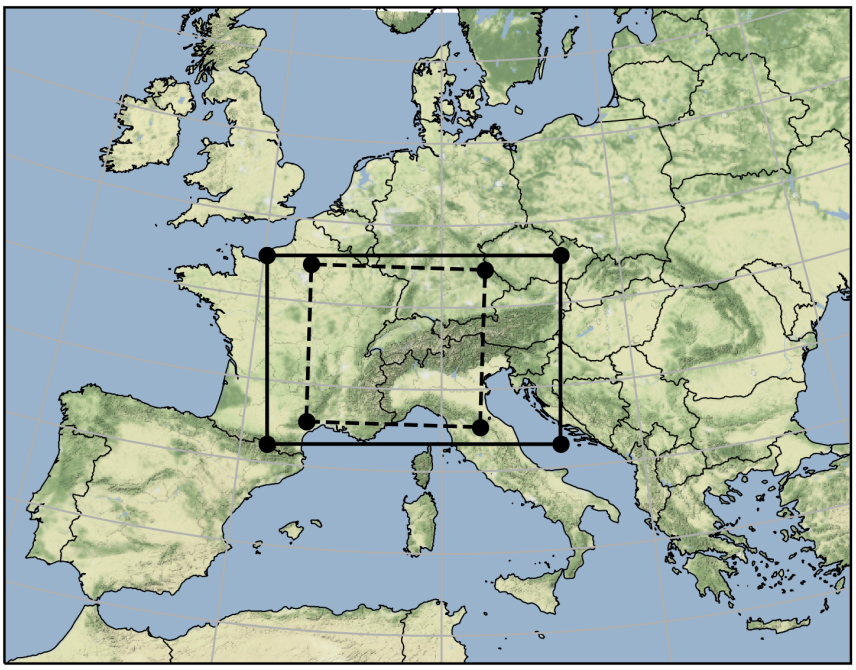


Figure 1.5: The Swiss radar composite domain (dashed line) and the COSMO model Alpine domain (solid line) with respect to the rotated lat/lon-grid with coordinates 43°N and 170°W used in COSMO.

This thesis was part of a wider research project funded by the Swiss National Science Foundation and hosted at the Radar, Satellite and Nowcasting Division within the Swiss Federal Office of Meteorology and Climatology MeteoSwiss. The project named *Precipitation attractor from radar and satellite data archives and implications for seamless very short-term forecasting* (grant PZ00P2_161316) aimed at constructing an attractor for precipitation directly from the large archives of radar data to gain insight into the intrinsic predictability of precipitation and thus design a seamless probabilistic quantitative precipitation estimation, nowcasting and short-term forecasting system.

Chapter 2

Scale-aware verification of precipitation nowcasting

This chapter has been adapted from the manuscript:

- Nerini, D., L., Foresti, D., Leuenberger, D., U. Germann, and H. Wernli: Scale-aware verification of high-resolution precipitation forecasts, **in preparation**.

It presents a verification of deterministic precipitation nowcasts produced with a radar-based extrapolation approach and the numerical model COSMO-1, whereby the forecast errors are analyzed through a scale separation technique based on the Fourier transform.

2.1 Summary

Owing to the increased realism of numerical simulations, NWP-based applications at the convective scale are rapidly growing and thus call for an in-depth characterization of the NWP predictive uncertainty and comparison with traditional heuristic radar-extrapolation techniques. The predictive skill of deterministic precipitation forecasts from the convective-scale model COSMO-1 is thus compared to the predictability of precipitation by the Lagrangian persistence approach within the nowcasting range (0-6 h) and at 10 min, 1 km resolution.

The analysis is performed for a full year (2018) and within a large domain (512 km x 512 km) centered over Switzerland. The forecast errors with respect to a high-resolution precipitation analysis are stratified by lead time and weather regime. The main contribution is the generalization of a traditional grid-scale verification method such as the RMSE to consider spatial scales through a scale separation technique based on the Fourier transform.

This study highlights the relevance of the scale-aware approach for the verification of high-resolution precipitation forecasts. The scale-aware predictability of precipitation by the Lagrangian persistence approach to nowcasting is found to be superior to COSMO-1 up to 3 h, while in relative terms the contribution of the numerical model is more important during warm convective events. It is shown that radar-based nowcasting is indispensable for applications requiring skill in predicting precipitation patterns with a horizontal dimension below 60 km, as currently this cannot be provided by state-of-the-art numerical models. After 4.5 h, precipitation on all scales below 150 km is poorly predictable by all forecasting means.

2.2 Introduction

High-resolution precipitation forecasts from numerical models have achieved a remarkable degree of realism. The convective rain of CPM-based forecasts is, in fact, no more to be seen just as general likelihood over an area – a "chance of showers" as in global atmospheric models, but it can be informative of the convective structure and details (Clark et al., 2016).
This rather new quantitative value has led to the rapid development of applications that can now benefit from the realism of convective-scale simulations. Among others, Panziera et al. (2016) combined radar observations with precipitation forecasts from the mesoscale prediction system COSMO-model to set-up an automatic warning system for heavy precipitation events over Switzerland, while Antonetti et al. (2019) used COSMO to force a flash-flood forecasting model for two medium-scale alpine catchments.
Just as the realism of simulations grows, so too the complexity of the forecast error structure increases. The link between improved model realism and better model skill is far from being trivial, as the higher level of detail determines the amplification of displacement errors and the need for high-quality, dense observations (Mass et al., 2002). As a result, some authors have advised against a blind increase in the model horizontal resolution and warned about the risks of the illusory improvements provided by highly realistic simulations (Wedi, 2014). With the growing role that NWP forecasts have to play in science-based decision making today, it is not surprising that it still exists a strong demand for adequate verification methods (Sun et al., 2014; Dorninger et al., 2018).

2.2.1 Brief review of spatial verification

The underlying assumption of standard verification methods at the grid scale is the existence of a perfect forecast from which any actual deviation is to be quantified (Brown et al., 2012). In the case of the CPM-based forecasts, the simulation of weather details beyond the range of predictable scales questions the existence of such a perfect forecast (e.g. Hohenegger and Schär, 2007).
With the above issue in mind, spatial verification methods have been designed to relax the point-to-point constraint of traditional skill scores, so that prescribed space-time errors can be forgiven. The spatial approach has rapidly found extensive use in the verification of high-resolution precipitation forecasts (as recent examples, Simonin et al., 2017; Chen et al., 2018; Gofa et al., 2018). Brown et al. (2012) provides an overview of spatial methods, while inter-comparison studies such as Gilleland et al.

(2009) or more recently Dorninger et al. (2018) aim at understanding the strengths and weaknesses of the broad spectrum of spatial verification methods at our disposal.
The scale decomposition approach is a subclass within the spatial verification family. Differently to neighborhood verification approaches, which rely on the *smoothing* of the error field (e.g. Roberts and Lean, 2008), scale decomposition methods aim at *isolating* spatial scales in order to quantify their skill. This aspect is particularly convenient in the sense that it can be used to generalize established verification methods in order to characterize the scale-dependence of the forecast error.
Briggs and Levine (1997) first introduced the use of wavelet analysis as a noise removal technique in the verification process. However, the authors already recognized the inherent value of the scale analysis in a diagnostic sense. Later, Casati et al. (2004) proposed the combination of wavelets and thresholding to provide an intensity-scale characterization of the forecast error. The scale analysis also motivated the work of Surcel et al. (2015) who introduced a methodology based on a spectral decomposition technique in order to study the scale dependence of precipitation predictability from NWP ensembles. The approach estimates a decorrelation scale λ_0, that is, the smallest spatial scale below which precipitation patterns are fully decorrelated and hence unpredictable.

2.2.2 Contribution of this study

This chapter focuses on depicting the predictability of high-resolution precipitation forecasts within the nowcasting range (0-6 h). The term *predictability* is used here in the sense introduced by Lorenz (1973). That is to say, by demonstrating that a given forecasting model can exhibit useful skill for a given range. The predictability of the numerical model is thus compared to a straightforward extrapolation-based technique which represents the ideal benchmark for highlighting the presence of predictive skill beyond the Lagrangian persistence (e.g. Berenguer et al., 2012; Foresti et al., 2019). The results are also stratified into contrasting weather regimes in order to gain some additional insights into the flow-dependent behavior of the model predictability. Building on similar scale decomposition verification methods, the approach used in this analysis introduces a generalization to spatial scales of traditional point-to-point error metrics through the Fourier transform. By doing so, the method also takes advantage of a high-quality precipitation analysis available at the same spatial and temporal resolution of the numerical model.
The chapter is organized as follows. Section 2.3 describes the precipitation analysis used as a reference and the two forecasting procedures under study. Section 2.4 presents the verification methods, including the formulation of the scale-aware error metrics, and the weather type classification that is used. In Sect. 2.5, the results of the

traditional grid-scale approach are presented first, and then the scale analysis is introduced. The section starts by discussing the Lagrangian and COSMO-1 approaches separately and then provides a comparative analysis. A brief sensitivity analysis on data transformation is provided at the end of the section, while an overview of rainfall statistics is provided in the beginning. Finally, Sect. 2.6 summarizes the main findings, discusses the limitations and foresees some future work.

2.3 Data and models

2.3.1 Radar precipitation analysis

The skill of the precipitation forecasts is evaluated against the CombiPrecip precipitation analysis, which is computed with 1 h accumulation and 1 km spatial resolution. The product is based on a geostatistical combination of weather radar observations and hourly rain gauge accumulations using kriging with external drift. Details can be found in Sideris et al. (2014). The hourly analysis is further disaggregated into 10 min rainfall accumulations based on the radar data (Barton et al., 2019). The precipitation analysis spans a domain of 640 km x 710 km that is centered over Switzerland (Fig. 2.1).

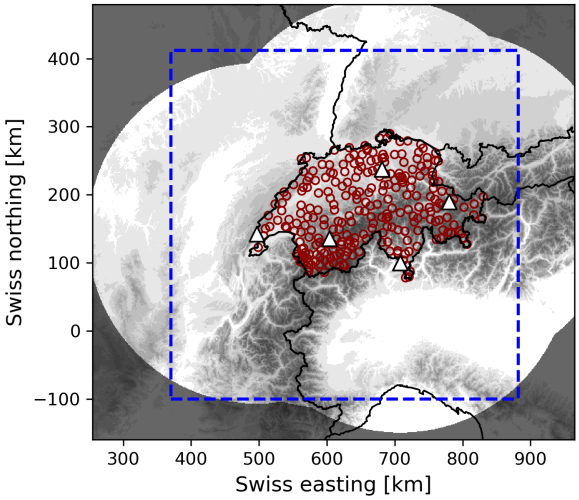


Figure 2.1: The full extent (640 km x 710 km) of the precipitation analysis used in this study and the 512 km x 512 km study area (blue dashed box) centered over Switzerland. The five radars are indicated with white triangles, while the red circles represent the location of the rain gauges used in the precipitation analysis. The composite radar mask, topography, and national borders are also included as a reference.

For 2018 the number of rain gauges available in CombiPrecip was 264, all located within the Swiss national borders (Fig. 2.1). Outside Switzerland, the precipitation analysis is designed to converge to the radar-only estimate. We also note that the

radar observation quality is range dependent. For these reasons, the analysis domain was further reduced to a 512 km x 512 km box centered over Switzerland (dashed box in Fig. 2.1), which focused the analysis on a region with higher observation density and quality.
The forecast verification was performed using quality-checked weather radar observations, that is, only when all five radars were in operation. Additionally, radar images displaying anomalous precipitation statistics were discarded. For more details on the radar quality check, see Foresti et al. (2018).

2.3.2 The Lagrangian persistence model

The radar-based precipitation nowcasts (Section 1.2.1) were generated using the open-source nowcasting library pysteps (Pulkkinen et al., 2019b), release V1.0.0. Pysteps includes several routines for the computation of deterministic advection forecasts, as well as modules for the reading and processing of radar QPE composites. For estimating the motion field, we selected the pysteps implementation of the local Lucas-Kanade optical flow method (Lucas and Kanade, 1981; Bouguet, 2001), as it provides the best trade-off between computational time and accuracy (Pulkkinen et al., 2019b). The extrapolation of the radar observation was performed with the backward-in-time semi-Lagrangian scheme described in Germann and Zawadzki (2002), which is available in the pysteps library. Default model parameters were used.
The Lagrangian persistence model was run with the same frequency of the COSMO-1 numerical forecasts, that is, every 3 h starting at 00 UTC (see below). It used a 30 min assimilation window (i.e., a sequence of three radar images) and produced +6 h nowcasts. The estimation of the optical flow at 1 km and 10 km resolution took around 2 seconds, while the advection routine was completed in less than 5 seconds.

2.3.3 The COSMO model

The Consortium for Small-scale Modelling (COSMO) develops and maintains the COSMO-model, a non-hydrostatic limited-area atmospheric model (Baldauf et al., 2011). Since early 2016, the Swiss national weather service MeteoSwiss operates the COSMO-model with a high-resolution setup for the Alpine region, known as COSMO-1. COSMO-1 provides short-range deterministic forecasts (+33 h), 8 times a day (00, 03, ..., 21 UTC) with 0.01°horizontal resolution ($\approx$ 1.1 km) and 80 vertical levels. Currently, a full COSMO-1 forecast is available around 1.5 h after analysis time.
Model initialization is based on nudging against conventional observations using a cut-off time of 45 min. The conventional observations types include surface observa-

tions from land and sea stations (SYNOP, SHIP and BUOY), upper air profiles (TEMP and PILOT), aircraft observations (AMDAR) and wind profilers (WPROF).
Additionally, surface radar reflectivities are assimilated using latent heat nudging (LHN, Stephan et al., 2008). LHN adjusts the model latent heating using a correction factor which depends on the ratio of radar QPE and model precipitation analysis. The change in buoyancy that follows affects the microphysical processes driving the formation of clouds and precipitation (Leuenberger and Rossa, 2007). LHN is applied only for the first 55 min into forecast time and assimilates the latest available radar-only QPE images.
Finally, COSMO-1 downscales its boundary conditions directly from the global IFS-HRES model of the European Centre for Medium-range Weather Forecasts (ECMWF), available at 0.1°horizontal resolution ($\approx$ 9 km), four times a day. The initialization of the soil state (TERRA, Schulz et al., 2016) is based on the background run only.
The numerical forecasts were extracted from an archive of operational runs and interpolated by nearest neighbor from the native latitude-longitude grid onto the projected grid and domain of the precipitation analysis.

2.4 Verification methods

2.4.1 Univariate predictability

The univariate approach to forecast verification considers point-to-point forecast-observation pairs as independent samples of the underlying error distribution.
In this study, the univariate verification procedure includes common continuous errors metrics such as the mean error (ME), root-mean-squared error (RMSE), mean absolute error (MAE) and correlation coefficient (r). We note that the latter is computed without subtraction of the mean as in the product-moment correlation coefficient introduced by Zawadzki (1973). In addition, ME, RMSE and MAE are normalized by the mean observed precipitation intensity and are thus denoted by the subscript $_n$. The formulas are as follows.

$$\mathrm{ME}_n(\tau) = \frac{\sum_{i=1}^{N} F(\tau,i) - O(i)}{\sum_{i=1}^{N} O(i)}, \tag{2.1}$$

$$\mathrm{RMSE}_n(\tau) = \sqrt{\frac{\sum_{i=1}^{N} \left[F(\tau,i) - O(i)\right]^2}{\sum_{i=1}^{N} O(i)^2}}, \tag{2.2}$$

$$\mathrm{MAE}_n(\tau) = \frac{\sum_{i=1}^{N} |F(\tau,i) - O(i)|}{\sum_{i=1}^{N} O(i)}, \tag{2.3}$$

and

$$r(\tau) = \frac{\sum_{i=1}^{N} F(\tau,i)O(i)}{\sqrt{\sum_{i=1}^{N} F(\tau,i)^2 \sum_{i=1}^{N} O(i)^2}}, \tag{2.4}$$

where τ and i are a given lead time and grid point. N corresponds to the total number of forecasts at grid points where either the value for the forecast F or the observation O exceeded 0.1 mm h^{-1}(i.e., single or weak conditioning). We note that the RMSE is normalized by the root-mean-squared observations.

The normalized scores can be used to define a limit to predictability by using a maximum error threshold after which predictability is considered lost, as the forecast does not exhibit useful skill. In the case of the correlation coefficient, Germann and Zawadzki (2002) used the value of $1/e \approx 0.37$ which, by assuming exponential decay of $r(\tau)$, corresponds to a complete decorrelation for $\tau \geq \tau_0$, the decorrelation time. For MAE$_n$ and RMSE$_n$, a more pragmatic approach denotes as *poorly predictable* a forecast associated with a normalized error larger than 1 (e.g. Keil et al., 2014). We applied and compared both approaches.

2.4.2 Scale-aware predictability

To investigate the scale-dependence of forecast errors, we compute the discrete energy spectrum of a given (L x L)-pixel field of errors $\mathbf{X} = \mathbf{F} - \mathbf{O}$, namely the square of the norm of its complex Fourier transform

$$G(k_x, k_y, \tau) = |\mathcal{F}\{\mathbf{X}(\tau)\}(k_x, k_y)|^2, \tag{2.5}$$

with k_x and k_y being the spatial wavenumbers in the x and y directions, respectively. By virtue of the Parseval's theorem, (2.5) conserves the total power of the original spatial field. That is to say, the sum over all wavenumbers of $\mathbf{G}(\tau)$ will equal the total variance of $\mathbf{X}(\tau)$, hence the sum of the squared errors over all grid points. It follows that $G(k_x, k_y, \tau)$ can be interpreted as the portion of error variance explained by wavenumbers k_x and k_y.

The 2D spectrum can be simplified into a 1D spectrum by computing its radial average

$$G(|\mathbf{k}|, \tau) = \frac{1}{|\mathbf{Z}|} \sum_{z=1}^{|\mathbf{Z}|} G(Z_z, \tau), \tag{2.6}$$

where $\mathbf{Z} = \{(k_x, k_y)_1, \ldots, (k_x, k_y)_{|\mathbf{Z}|}\}$ is the set of wavenumbers for which $|\mathbf{k}| \leq \sqrt{k_x^2 + k_y^2} < |\mathbf{k}| + 1$, namely all spatial wavenumbers lying within one radial wavenumber unit.

By combining (2.5) with (2.6) and then applying the square root, we obtain the relation for the RMSE as a function of radial wavenumber and lead time:

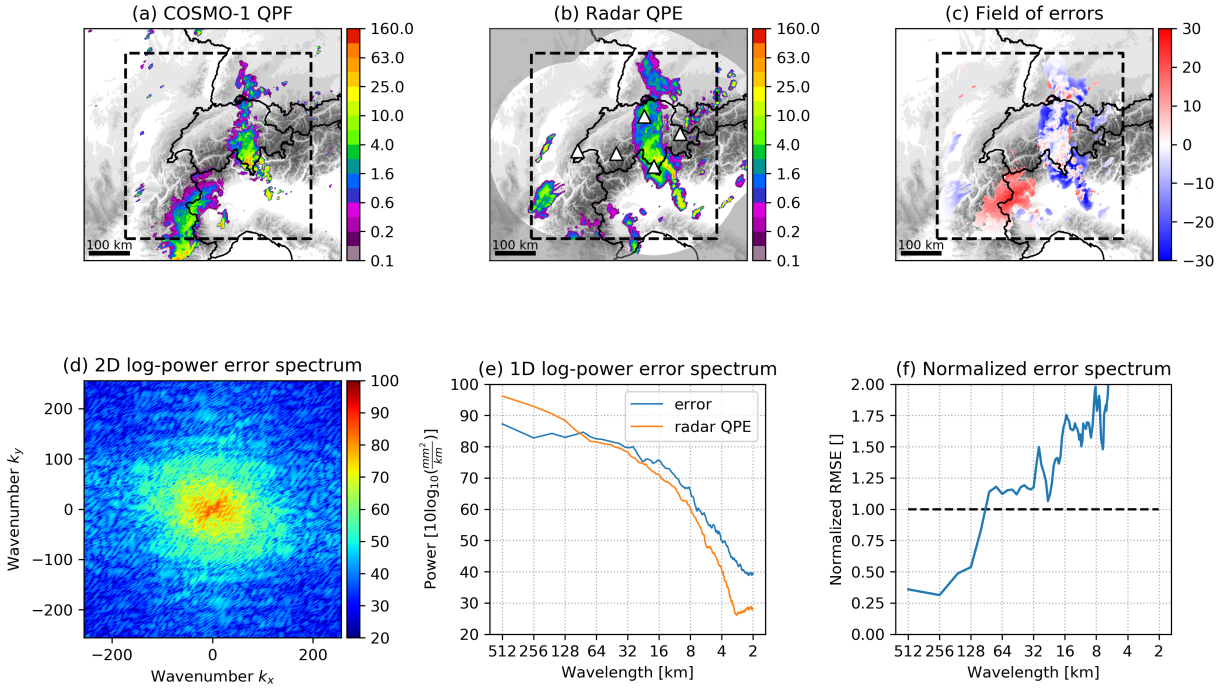


Figure 2.2: Illustration of the scale-aware verification procedure. In rain rates, the 1500 UTC COSMO-1 forecast for +2 h (a) and the corresponding radar-based precipitation analysis valid at 1700 UTC 9 August 2018 (b) define the linear additive error field (c, here defined as $10\log_{10}(\text{fct}/\text{obs})$, only for visualization purpose). The error field is Fourier transformed into the frequency domain, its power spectrum computed (d), and then radially-averaged (e). Finally, the normalized RMSE spectrum is obtained as the square root of the ratio between the error and observation power spectra (f).

$$\text{RMSE}(|\mathbf{k}|,\tau) = \sqrt{\frac{1}{|\mathbf{Z}|N}\sum_{z=1}^{|\mathbf{Z}|}\sum_{j=1}^{M}|\mathcal{F}\{\mathbf{X}_j(\tau)\}(\mathbf{Z}_z)|^2}, \tag{2.7}$$

where M is the number of forecast fields included in the analysis. To obtain the normalized scale-aware RMSE, one needs to consider the average spectrum of the observation field:

$$\text{RMSE}_n(|\mathbf{k}|,\tau) = \sqrt{\frac{\sum_{z=1}^{|\mathbf{Z}|}\sum_{j=1}^{M}|\mathcal{F}\{\mathbf{X}_j(\tau)\}(\mathbf{Z}_z)|^2}{\sum_{z=1}^{|\mathbf{Z}|}\sum_{j=1}^{M}|\mathcal{F}\{\mathbf{O}_j(\tau)\}(\mathbf{Z}_z)|^2}}. \tag{2.8}$$

Finally, the radial wavenumber $|\mathbf{k}|$ relates to spatial scale as $s = L\Delta x/2/|\mathbf{k}|$ where Δx is the grid resolution. A visual illustration of the above procedure is provided in Fig. 2.2.

We note that before applying the Fourier transform, the field of errors $\mathbf{X}(\tau)$ is tapered with a Tukey window (Fig. 2.2c) to remove the discontinuities at the edges.

2.4.3 Classification of weather regimes

Several automatic weather type classifications (WTCs) are available for the Alpine region. Weusthoff (2011) provides a comprehensive overview. For this study, the 11-class GWTWS weather type classification was used. The method employs fields of 500 hPa geopotential heights to differentiate between the eight main wind directions (N, NE, E, SE, SW, W, NW) and applies a $7\,\mathrm{m\,s^{-1}}$ threshold on the mean wind speed at the 500 hPa to differentiate between advective and convective situations. For convective cases (wind speed $< 7\,\mathrm{m\,s^{-1}}$), a distinction between low, high and flat pressure situations is made based on the mean sea level pressure.
Despite the name, convective situations as defined above can very well occur outside the warm season and thus include events that are not, strictly speaking, convective. To better capture this difference, the advective and convective situations are further classified in cold and warm types based on the month. The cold season is defined from January to March and from October to December (JFMOND), while the warm season extends from April to September (AMJJAS), thus covering the full convective season for Central Europe (Nisi et al., 2016).

2.5 Results

2.5.1 Rainfall statistics

Figure 2.3a shows the frequency of all 11 weather situations for 2018, along with their occurrence within the cold months (JFMOND) or warm months (AMJJAS). The fraction of wet analyses by weather type is also represented. A 10-minute precipitation analysis is considered wet if its fractional coverage over the study domain exceeds 5%.
During 2018, weather situations from West were common advective types and brought precipitation roughly 40% of the time. Conversely, advective classes from the southern sector were more often associated with precipitation (more than 80% of the time for class South), but had lower occurrence. In case of convective weather types, Low Pressure systems were associated to a higher occurrence of precipitation (40%) despite being a relatively rarer convective type. The prevalence of convective weather type situations during the warm season is well captured in Fig. 2.3b, where it is also possible to notice a spike in the fraction of wet analyses in May (60%). This aspect is

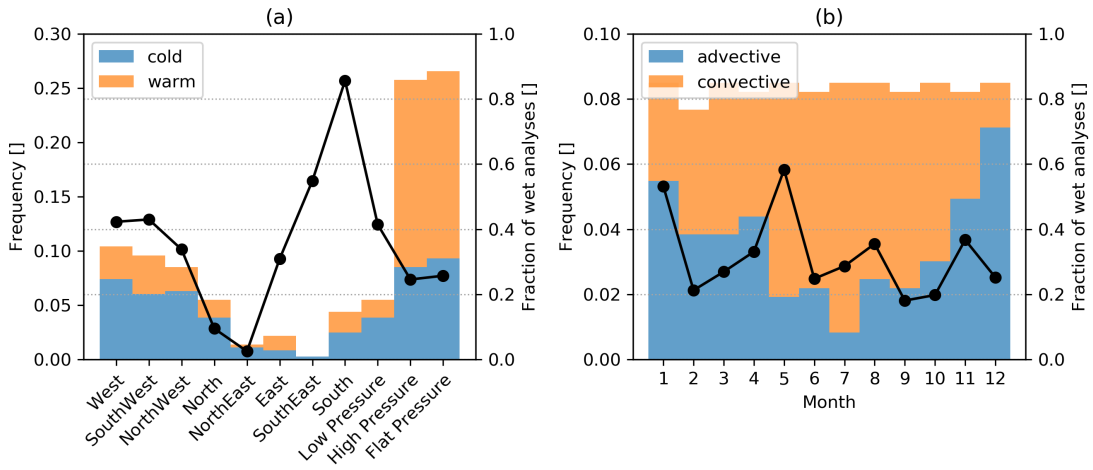


Figure 2.3: (a) The 11-class GWTWS weather type classification for 2018 and the fraction of wet analyses for each of the weather types situations. The months are grouped in one *cold* season (months JFM and OND) and one *warm* season (AMJJAS). (b) The monthly distribution of advective vs. convective situations and the monthly fraction of wet 10-min analyses. The *advective* type includes situations NE, E, SE, SW, W, NW, while LP, HP and FP are classified as *convective*.

well backed by the official climate reports which described May 2018 as anomalously warm and characterized by an important convective activity (e.g. Cattani et al., 2019). The diurnal cycle of precipitation statistics for the whole of 2018 is presented in Fig. 2.4, where the 25-50-75 quantiles in terms of conditional mean precipitation intensity, that is, excluding the zeros, and fractional coverage are given as a function of day time and for the two contrasting weather types, namely advective and convective situations. The results are further stratified into the warm and cold seasons.

Compared to advective cases, precipitation associated with warm convective situations is on average stronger, more localized, and displays a more pronounced diurnal cycle. The variability of mean precipitation intensities is also larger for warm convective cases, while precipitation in advective cases exhibits more variability in terms of fractional coverage.

In the cold months, precipitation intensity is significantly lower and does not follow any distinct diurnal pattern. However, the fractional coverage of cold advective cases is markedly higher than any other weather type considered in this analysis.

2.5.2 Predictability by Lagrangian persistence

The practical predictability of precipitation by the Lagrangian persistence approach to nowcasting is investigated first. The nowcast errors are analyzed using the univariate approach, which is then generalized to account for different spatial scales.

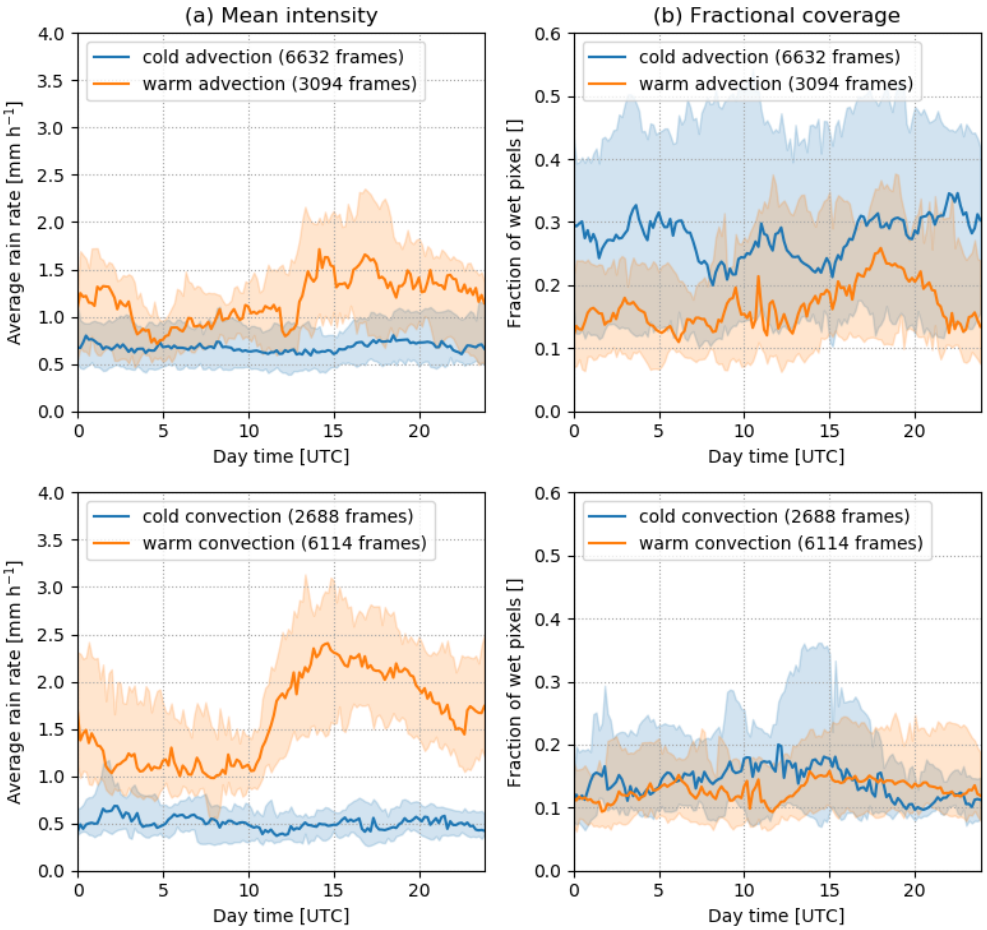


Figure 2.4: Diurnal cycle of two rainfall statistics in 2018. The analysis includes (a, left column) the conditional mean precipitation intensity and (b, right column) the fraction of wet pixels over the verification domain, and it is performed separately for the advective (top row) and convective (bottom row) weather types. The 25-50-75 quantiles are represented as a function of day time, and the statistics are further stratified into the cold (blue) and warm (orange) seasons.

Figure 2.5 presents the average error growth of Lagrangian nowcasts in terms of normalized ME, RMSE and MAE, and in terms of the correlation coefficient, r. The results are stratified in two contrasting weather regimes, namely advective and convective situations, and by warm and cold months. We first note that the Lagrangian persistence approach is often affected by a dry bias (ME_n on Fig. 2.5a) which can reach -50% by the end of the 6 h nowcast in case of cold advective cases. Such bias relates to the effect of extrapolation on the lateral boundaries, as precipitation is advected outside the domain. We also note that precipitation events in the cold season are generally associated with smaller forecast errors and are hence more predictable than precipitation occurring during the warm season. The same can be said of the advective versus convective situations, the former being characterized by higher predictability than the

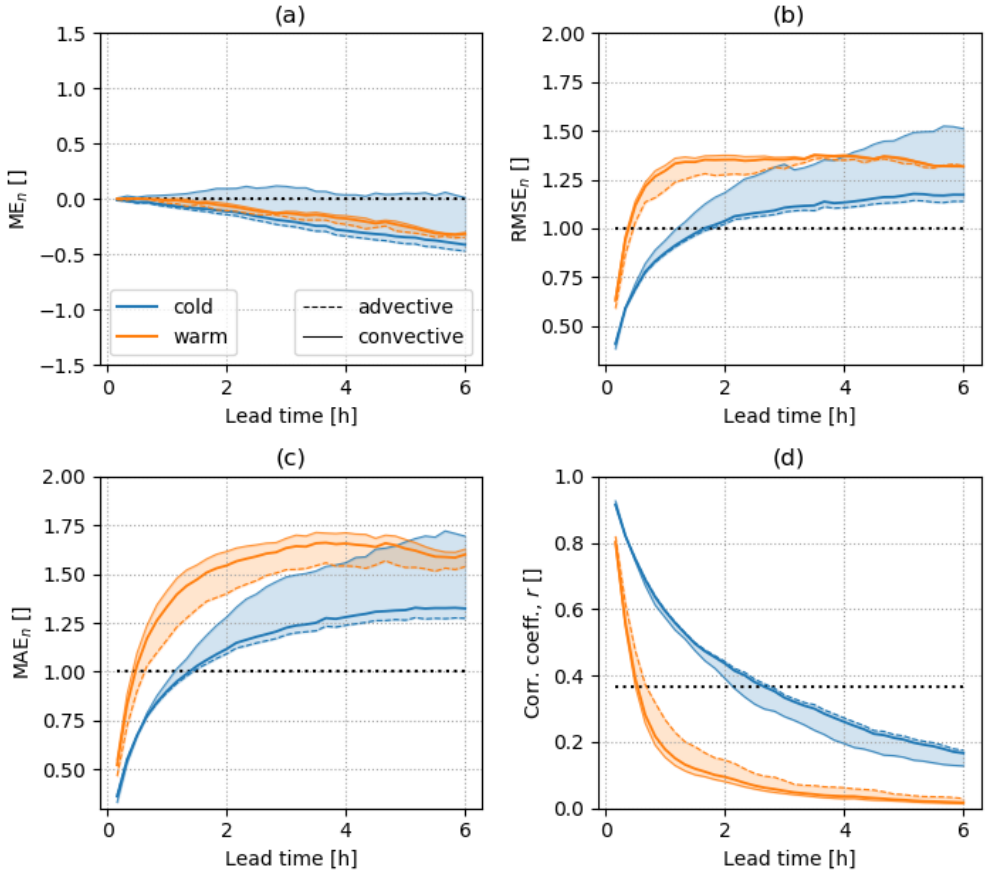


Figure 2.5: Univariate predictability of precipitation by Lagrangian persistence during 2018. The results are stratified according to two contrasting weather regimes (advective vs. convective types) and two main seasons (warm vs. cold). The thick lines represent the average results for the cold and convective seasons, while the intervals represent the variability between advective and convective cases. The limits to predictability in terms of normalized RMSE and MAE, and r are represented as horizontal dotted lines, while the horizontal dotted line in panel (a) is for an unbiased forecast.

latter. The growth of the normalized MAE and RMSE is similar, with both scores indicating poor predictability already after around 30 min during the warm season and slower error growth in the cold season (between 1.5 and 2 h, Figs. 2.5b-c). The correlation coefficient confirms the significant difference in skill between cold and warm seasons (Fig. 2.5d), as the threshold $1/e$ is reached on average within 30 min for warm months and within 2.5 h in cold months.

If such results may seem at first look rather on the pessimistic side, one has to consider the harsh requirements of the univariate verification approach. The errors are computed at 1 km and 10 min resolution, they exclude correct negatives (because of the single conditioning), and are defined in terms of rain rate units. It is thus not surprising to find limits to predictability below 1 h when considering warm convective

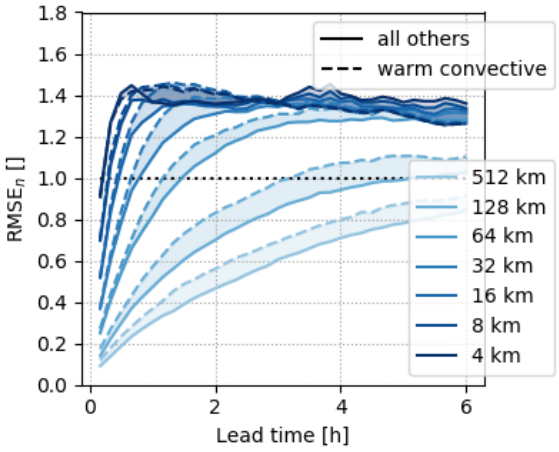


Figure 2.6: Scale-aware predictability of the Lagrangian nowcasts in 2018. The results are stratified according to two contrasting weather regimes (warm convective vs. all others) and for a set of spatial scales. The fill color is used to pair results for the same scales.

events. Foresti et al. (2016) reported similar results for convective features in their analysis over Belgium.

As highlighted above, the challenges of the univariate approach in the presence of high-resolution precipitation forecasts motivate the adoption of a verification framework that can naturally cope with the full spectrum of spatial scales. Figure 2.6 presents the scale-aware normalized RMSE. The prediction error has been decomposed by Fourier transform and plotted for a set of spatial scales. The results are now stratified into warm convective events and remaining events.

Because the approach is a relatively straightforward extension of well-established error metrics, the results offer an intuitive representation of the scale-dependent practical predictability of precipitation. This is found to range between 30 min and more than 6 h for horizontal spatial scales in the order of 10 km or 100 km, respectively. The scale separation approach thus provides insights into the actual predictive skill of the forecast by allowing the distinction between predictable and unpredictable scales. For example, we note that the relative lower predictability of warm convective events seems to be associated especially with larger scales, where the normalized errors are clearly larger with respect to the remaining events.

These results are reasonably aligned with other studies on the scale-dependent skill of the Lagrangian persistence approach. Germann et al. (2006b) used wavelets as a band-pass filter to perform a scale analysis of continental radar images in the US and found average predictability of around 4 h at 100 km, which dropped to less than 1 h at 10 km. More recently, Foresti and Seed (2014) employed a cascade decomposition framework to study the precipitation lifetime as a function of scale from the radar composites of

eastern Victoria, Australia and found average predictability of around 8 h for 80 km spatial wavelengths and around 50 min for 20 km. Using a scale analysis approach based on the discrete cosine transform, Surcel et al. (2015) estimated the predictive limit of radar extrapolation to be equal to 2 h for scales smaller than 200 km.

2.5.3 Predictability by COSMO-1

Figure 2.7 presents the univariate error growth of COSMO-1 precipitation forecasts in the 0-6 h range and in terms of normalized ME, RMSE and MAE, and correlation coefficient r.

We first note in the (a) panel of Fig. 2.7 that during cold months, COSMO-1 is on average affected by a rather systematic wet bias. This seems to be most important in the case of low-intensity and small-extent cold convective events when the overestimation can be as high as 100%.

The benefits of radar DA in the first hour into forecast time are easily recognized, as the RMSE, MAE and r scores consistently display better skill. However, we also note that this does not mean a perfect skill at time zero, the main reason being that in practice LHN assimilates only a proxy of precipitation, that is, latent heating. Moreover, the benefits of radar DA appear to quickly decay after LHN is switched off, particularly in case of warm events where the errors saturate within 30 min.

In general, the normalized RMSE and MAE scores (Fig. 2.7b-c) indicate poor grid-scale predictability for all considered lead times and error saturation within 1.5 h. We also note that the growth of COSMO-1 normalized MAE and RMSE exhibits a relatively weak dependence with lead time within the nowcasting range.

The Zawadzki's correlation coefficient is not insensitive to forecast biases in the same way as the Pearson's correlation, but it seems nonetheless to be less influenced than other error metrics considered in this study. As a result, in Fig. 2.7d we note the presence of some predictive skill during the cold season, which also happens to be the season being most affected by the positive bias (Fig. 2.7a). This result seems to suggest that the presence of systematic bias might have an impact on the analysis of predictability.

More generally, the above results once again emphasize the limits of a univariate verification approach as this tends to focus on the quick error saturation occurring at the grid scale. Instead, the scale-aware RMSE_n of COSMO-1 presented in Fig. 2.8 highlights more interesting contrasts in the model skill between predictable and unpredictable scales. Concerning large scales (> 100 km), we note that COSMO-1 exhibits skill during the whole nowcast range, particularly during non-convective cases. The result may be a consequence of longer-lasting effects of radar DA at large scales, which has also been suggested by some authors in idealized (e.g. Bachmann et al., 2019) as

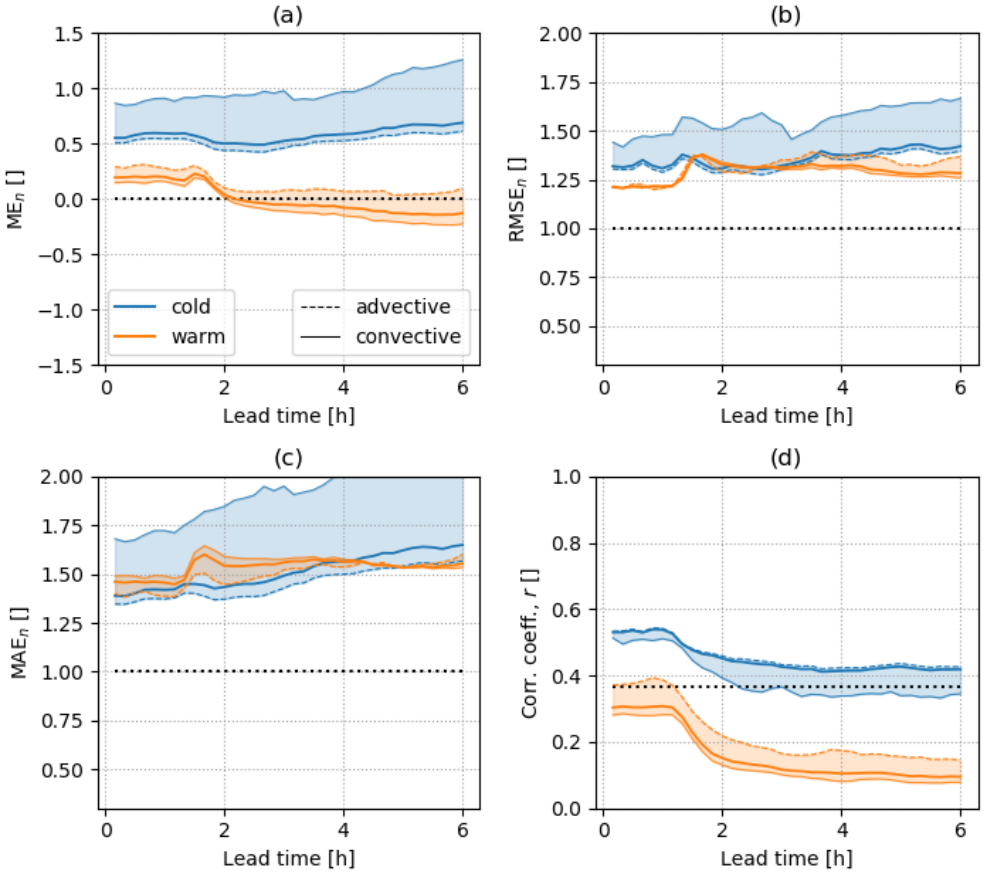


Figure 2.7: As in Fig. 2.5, but for COSMO-1 forecasts. Note that radar surface reflectivities are assimilated with LHN during the first hour into forecast time.

well as real modeling experiments (e.g. Stratman et al., 2012; Surcel et al., 2015). For smaller scales, the skill of the model is either lost within the first two hours or is poor from the very beginning. In particular, we observe for scales smaller than 32 km a clear error saturation peak at around 1.5 h. The predominant small-scale error in non-convective events might be associated with the same wet bias that was observed in a previous analysis (Fig. 2.7a).

Finally, we note that in the first hour into forecast time, warm convective cases display a consistently lower error than the remaining cases, which might be an indication of the improved performance of the LHN scheme in weather situations characterized by the presence of a marked atmospheric instability and weak synoptic forcing, since the assimilation of radar reflectivities is able to trigger convection in the model more effectively (Craig et al., 2012).

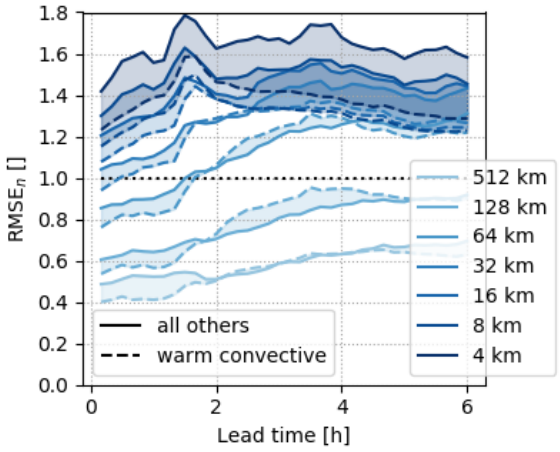


Figure 2.8: As in Fig. 2.6, but for COSMO-1 forecasts. Note that radar surface reflectivities are assimilated with LHN during the first hour into forecast time.

2.5.4 A comparative analysis

This section aims at providing a comparative analysis of the skill of the Lagrangian and COSMO-1 forecasts in the nowcasting range. It will, therefore, focus on the lead time, weather type, and spatial scale at which the numerical model manages to outperform the Lagrangian approach. Because of its practical implications (e.g., for radar-NWP blending), such a comparison must consider the requirements of a real-time setting. Hence, the COSMO-1 results are adjusted to account for the computational time of a numerical run, which currently lasts around 1.5 h. The COSMO-1 error curves are shifted in time accordingly (and the nowcasting range is consequently reduced to 4.5 h).

In a comparative study for the Alpine region, Mandapaka et al. (2012) computed the correlation between the observed and forecasted reflectivity fields for both COSMO and Lagrangian-based nowcasts and concluded that the crossover lead time was on average around 2.5 h. Here we will provide a similar analysis with the added value of a scale-aware verification procedure.

Figure 2.9 provides a summary of the scale-aware normalized RMSE in warm convective events versus all other situations.

In warm convective cases, the Lagrangian approach provides on average better forecasts at scales larger than 64 km up to around 2.5 h, at the point when COSMO-1 starts to outperforms. For smaller scales, the numerical model appears to outperform the Lagrangian approach already at shorter lead times, that is, it takes over as the best forecast within the first hour at scales smaller than approximately 30 km. The finding can be related to the combined effect of the improved radar DA performance and the lower Lagrangian predictability occurring during warm convective events. In a recent

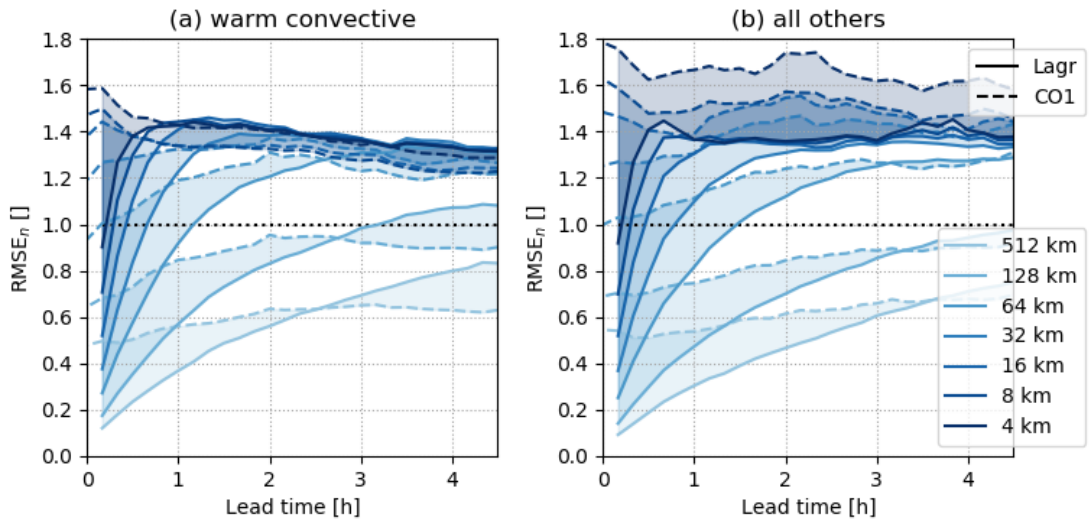


Figure 2.9: Comparison of Lagrangian and COSMO-1 scale-aware forecast errors during (a) warm convective events and (b) all other events. The COSMO-1 curves are shifted by 1.5 hours to account for the computational requirements of numerical simulations.

study for the UK, Simonin et al. (2017) found similar results by noting that the numerical model performed relatively better by providing 20 or 30 min additional benefit during convective days. We should nonetheless note how at those scales and lead times, our results indicate that both models provide poor predictive skill.

In the remaining cases, the Lagrangian approach manages to produce better forecasts up to around 3.5 h for scales down to 64 km. For smaller scales, the errors of the numerical model are larger than the Lagrangians errors for the whole nowcast range, presumably because of a high-frequency wet bias in the model, as already discussed above.

Figure 2.10 compares the limit to predictability for both forecasting procedures. This is illustrated in terms of the smallest predictable scale for a given lead time. The concept of the smallest predictable scale is similar to the decorrelation scale of Surcel et al. (2015) in the sense that it defines a spatial scale beyond which the forecast does not exhibit any predictive skill. In our case, this is produced by plotting for a given lead time the smallest scale at which the condition $\mathrm{RMSE}_n < 1$ is satisfied, while the decorrelation scale was defined as the scale at which the covariance between forecast and observation approaches zero.

The smallest predictable scale in COSMO-1 precipitation forecasts starts at around 60 km and doubles within 1 h. In the remaining forecast range, the predictive skill of COSMO-1 stabilizes around 150 km. The Lagrangian nowcasts can provide useful skill below 50 km during the first hour into the forecast, but then the predictable scale

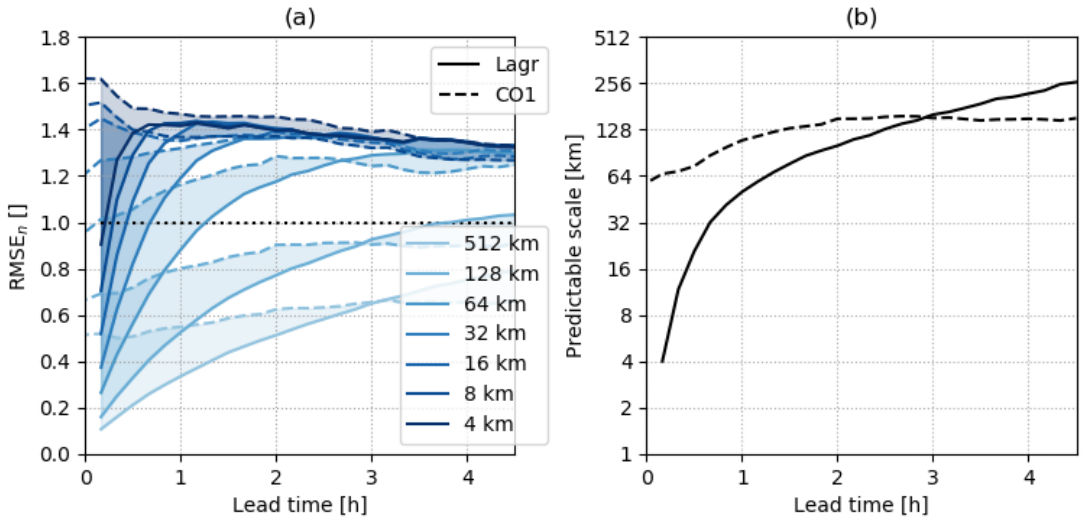


Figure 2.10: The scale-dependent limit to predictability in both COSMO-1 and Lagrangian nowcasts. The (a) scale-aware normalized RMSE for the whole 2018 is used to compute the spatial scale at which the RMSE$_n$ drops below 1, which is then used as indication of predictive skill. This minimum predictable scale is illustrated in (b) (note the logarithmic scale).

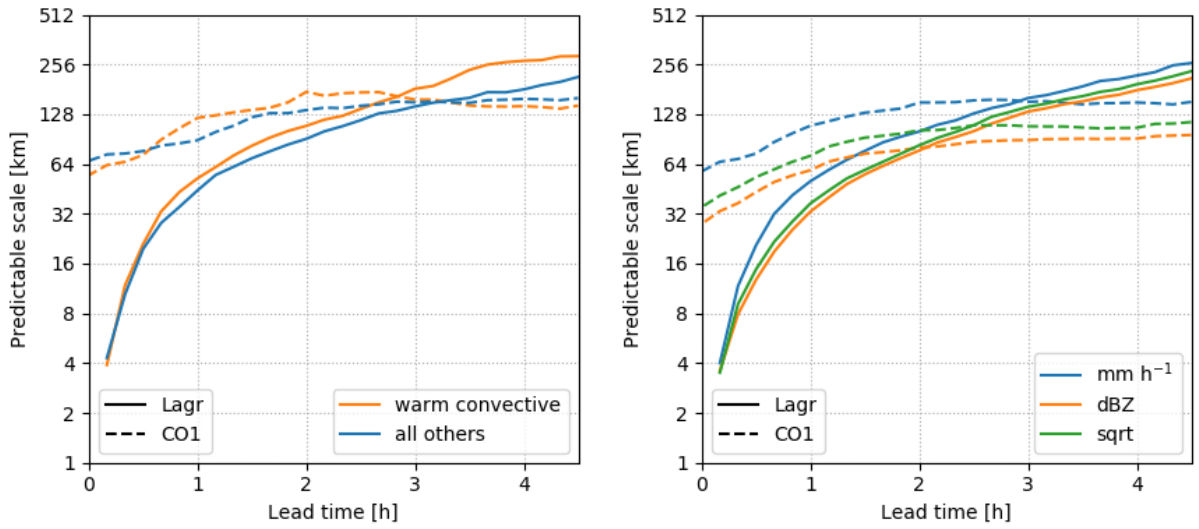
rapidly degrades. After 3 h, the scale-aware predictability provided by COSMO-1 is smaller than the one provided by the Lagrangian forecast. We thus note a good agreement between the scale-aware crossover lead time of this study and the findings of Mandapaka et al. (2012). After 4.5 h, all scales below 150 km are poorly predictable by all means.

A closer look at the flow dependence of the scale-aware practical predictability is taken in Fig. 2.9a. This confirms the shortened crossover leadtime (around 30 min) during warm convective days, which is mainly driven by the decrease of skill in radar nowcasting, while the predictable scales by COSMO-1 are not particularly affected.

2.5.5 Remarks on data transformations

The above verification results look at precipitation as a real continuous scalar quantity and provide a summary of the forecast errors using the sample mean. The underlying assumption of such an approach is that the error distribution is *symmetric*. However, it easy to show that the assumption of symmetry does not hold in the presence of the highly skewed distribution of precipitation intensities. One common approach consists of using data transformation techniques to enforce more symmetry into the data distribution before the analysis (e.g., Erdin et al., 2012).

However, one could question the need for data transformation in the context of scale analysis. The main argument is the intuition that larger scales will naturally exhibit



(a) **(b)**

Figure 2.11: Sensitivity analysis of the scale-aware predictability by COSMO-1 and Lagrangian nowcasts to (a) weather regimes and (b) data transformations.

more symmetry in their rain rate distribution. Figure 2.11b looks at the sensitivity of the scale-aware predictability analysis to the use two common data transformation approaches, namely the log (as dBZ units) and square-root transformations.
We note that data transformation generally improves the predictability of the forecast, arguably due to the lower weight given to errors associated with high intensities. Moreover, both transformations appear to reward the COSMO-1 forecasts by reducing the crossover lead time. The predictable scale of COSMO-1 is improved by almost a factor of two, while the benefits for the Lagrangian nowcasts are less important. This difference might be another indication of the bias contribution to the RMSE of COSMO-1, which is then more effectively reduced following the data transformation. With a larger sensitivity compared to weather regimes (Fig. 2.9a), we conclude that the impact of data transformation is non-negligible. The asymmetric distribution of the errors might still play a role and, therefore, represents a limitation of the present method. However, in the somewhat awkward position of having to choose on data transformation, we finally opted for not applying any. The main arguments are the advantages of working directly with precipitation rates, the plausibility of our results in the linear space, and, most of all, the rather arbitrary choice of which data transformation approach to use.

2.6 Conclusions

This work investigated the scale-dependent predictability of precipitation patterns within the nowcasting range by two different forecasting procedures, namely a radar extrapolation technique based on the Lagrangian assumption and the numerical forecasts from the limited-area COSMO-1 high-resolution model. The analysis covered the whole of 2018 and was restricted to a 512x512 km domain over the Swiss Alps. The results were stratified according to lead times and weather types. The main findings can be summarized as follows.

- If applied at the grid scale, traditional verification methods offer limited insights. The predictability of precipitation was found to be either poor from the very beginning or rapidly lost, particularly in case of warm convective events. However, the simple formulation of these error metrics is still advantageous in the sense that they provide an intuitive definition of predictability.

- Scale decomposition can be used as a generalization of such verification methods to account for the presence of unpredictable scales. Fourier transform provides a computationally-efficient way to decompose the error field, while thanks to the Parseval's theorem, we can interpret its power spectrum in terms of contribution to the total squared error. This scale-aware formulation provides an easily interpretable measure of skill which can deal with the complex structure of high-resolution forecast fields. The error normalization naturally fits within the method and introduces a suitable framework for the analysis of the scale-aware predictability.

- It was thus found that the predictability of precipitation patterns by both the Lagrangian approach and COSMO-1 can span beyond 6 h for $O(100)$ km horizontal scales, while it is generally lower than 2 h for $O(10)$ km. During warm convective events, the practical predictability of precipitation is reduced mainly at large scales and in particular for the Lagrangian approach, as a consequence of the lower persistence of convective storms. We also noted lower initial errors in the numerical model during warm convective days, which is explained as being the result of the more effective assimilation of radar data in the presence of marked atmospheric instability and weak synoptic forcing. In the remaining cases, the normalized errors of COSMO-1 at small scales can be 20% higher, arguably because of a systematic bias in the high-frequency components of the numerical simulations.

- For practical applications, radar-based nowcasting appears to be indispensable when predictive skill below the 60 km horizontal scale is required, as state-of-

the-art numerical simulations are currently unable to deliver skill at such small spatial scales. In terms of scale-aware predictive skill, the Lagrangian approach can provide added value to up to a 3 h lead time, after which the skill of COSMO-1 is generally higher. In warm convective days, COSMO-1 outperforms radar nowcasting around 30 min earlier than in the remaining days. After 4.5 h, precipitation patterns on all scales below 150 km are poorly predictable by all forecasting means.

We finally note that this study bears several limitations.
First, the sensitivity of the scale-aware analyses to data transformation is found to be non-negligible and represents, at the present state, a limitation of the approach in the presence of non-symmetric error distributions.
Second, the single conditioning applied during the verification procedure implies that the errors statistics are representative of the *wet* skill only, that is, the error conditional to the occurrence of either a wet forecast or analysis. Consequently, the results of this study do not describe the ability to discriminate between dry and wet events, which might be crucial in some applications.
Third, the scale-aware RMSE_n is influenced by the presence of systematic biases which, particularly considering COSMO-1, might be interesting to quantify separately. A scale-aware analysis of the model biases should be included in future work to better characterize the actual forecast uncertainty.
Fourth, by not explicitly accounting for observation uncertainty, the analysis is, in fact, rewarding the extrapolation nowcast. We tried to mitigate the issue by making use of the best available precipitation analysis.
Finally, the scale analysis was applied in terms of spatial analysis and hence neglected time as a dimension. A time scale analysis does not present technical problems, but it does complicate the study of the growth of errors.
Apart from addressing the limitations listed above, future work should consider the extension of the presented scale-aware framework to QPF ensembles. This would produce a scale-aware quantification of both the ensemble skill and spread and thus provide a scale-dependent analysis of the spread-error relationship, as already attempted by some authors (e.g. Zacharov and Rezacova, 2009; Chen et al., 2018).

Chapter 3

A local approach to ensemble radar nowcasting

This chapter has been adapted from the manuscripts:

- 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 (6)**, 2777 – 2797.
- Pulkkinen, S., D. Nerini, A. Perez Hortal, C. Velasco-Forero, U. Germann, A. Seed, and L. Foresti, 2019: Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1.0). *Geosci. Model Dev.*, **12 (10)**, 4185–4219.

It explores a moving window approach for model localization in the context of ensemble precipitation nowcasting.

3.1 Summary

The statistical properties of rainfall fields often exhibit significant spatial heterogeneity due to the natural variability of physical processes and the heterogeneous distribution of external forcings as topography. However, most ensemble techniques in radar-based precipitation nowcasting rely on the assumption of stationarity. For instance, the standard approach to simulate stochastic rainfall fields based on the Fourier filtering of white noise is only able to reproduce the global power spectrum and spatial autocorrelation of a precipitation field.

In this study, we explore the use of a short-space approach for the localization of the nowcasting procedure. First, a non-stationary stochastic generator based on the Short-Space Fourier Transform (SSFT) is presented. Conceptually similar to wavelet analysis, the SSFT is a simple and effective extension of the Fourier transform developed for space-frequency localization. Then, the same short-space approach is applied to the whole nowcasting procedure by using moving windows to better capture the local statistics of rainfall.

The potential of the localization procedure is demonstrated using synthetic and real data. The stochastic generator is verified in its ability to reproduce both the global and the local Fourier power spectra of the target field, while the overall forecast quality is verified using a set of precipitation events from the MeteoSwiss radar composite archive.

Localization generally produces a positive impact in terms of realism of the simulations and skill of the probabilistic forecasts. The local modulation of the stochastic precipitation fields allows for the seamless evolution from analysis into forecast, also in the presence of marked spatial inhomogeneities. We show that the localized nowcasts perform better in terms of ensemble reliability and resolution, as well as conditional bias. However, we also found that a too strong a localization can lead to lower skill as it implicitly relies on a stricter assumption of Lagrangian persistence.

3.2 Introduction

It has been long observed that rainfall fields exhibit spatial and temporal organization which is consistent with the scaling and multifractal framework (Schertzer and Lovejoy, 1987; Menabde et al., 1997). This observation is convenient since it allows developing parsimonious stochastic simulation methods which can generate realistic precipitation fields that depend only on a few parameters. Stochastic rainfall generators are effective in characterizing the uncertainty related to the measurement and forecasting of precipitation. They constitute the core of most probabilistic nowcasting methods, where they are used to perturb a deterministic nowcast and thus simulate the loss of predictability for increasing lead time (e.g. Bowler et al., 2006; Metta et al., 2009; Berenguer et al., 2011; Atencia and Zawadzki, 2014; Pulkkinen et al., 2019b).
One important class of such algorithms relies on the Fast Fourier Transform (FFT). The Wiener-Khintchine theorem (see Appendix B.3) states that the power spectral density (generally referred to as *power spectrum*) of a wide sense stationary random process can be obtained as the Fourier transform of the autocorrelation (autocovariance) function of that process. Thus, it becomes clear that stochastic simulation methods based on covariance or variogram (e.g., Lantuéjoul, 2002) can be easily reformulated by exploiting the FFT to achieve increased computational efficiency (Marcotte, 1996). Besides, the representation of data in Fourier space can enhance our understanding of the underlying generating mechanism and scaling behavior of rainfall.
The stochastic generators based on FFT assume spatial stationarity, that is, uniformity of the generator across space. As a consequence, the generator is only able to produce spatially homogeneous noise fields. More precisely, they cannot represent and simulate non-stationary fields that display a heterogeneous distribution of autocorrelation range or anisotropy.
Weather radar observations, however, reveal that stationarity rarely occurs in nature (Schleiss et al., 2014; Jameson et al., 2018). The non-stationary distribution of the statistical properties of precipitation becomes easily visible when stratiform and convective rain coexist in different regions of the domain or when there are clear local anisotropies in different directions, as in Figure 3.1. The need for a non-stationary stochastic noise generator is even more relevant in regions of complex orography, such as Switzerland, where the statistical properties and predictability of precipitation are influenced by orographic forcing (e.g., Panziera and Germann, 2010; Foresti et al., 2018).
A few approaches have been proposed to adapt the stochastic generators to capture the local variability of precipitation. Schleiss et al. (2014) introduced the concept of "dry drift" to capture the non-stationary nature of intermittent rainfall in stochastic simulations. Pathirana and Herath (2002) and Badas et al. (2006) proposed to filter or remove the heterogeneities of the precipitation field before applying a homogeneous

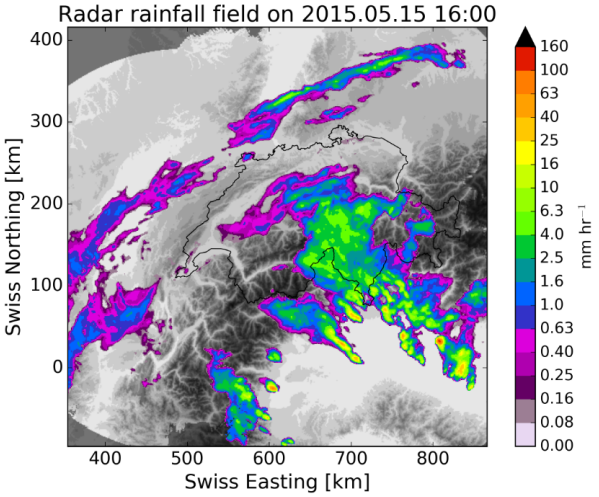


Figure 3.1: Example of non-stationarities in a radar rainfall field valid at 1600 UTC 15 May 2015. The domain spans 512x512 km and it is centered over Switzerland. Two distinct anisotropies coexist in the same image, while convection develops only to the south of the Alps.

multifractal downscaling technique and re-adding the trend after running the stochastic simulation. Wang et al. (2012) tried to detect and remove singularities in violation of the stationarity assumption. A similar strategy was employed in a recent STEPS implementation, which computes the local mean and variance of the rainfall field on a regular grid to remove the local non-stationarity before adding the homogeneous stochastic perturbations (Seed et al., 2013). However, these strategies are not able to generate noise fields that reflect the local 2D power spectrum and anisotropy of precipitation fields. When applied for nowcasting of weather situations presenting strong local anisotropies, the stochastic perturbations gradually destroy the original spatial structure, and important local information is thus rapidly lost.
The need for a more flexible approach to directly consider the non-stationarity of real-world geophysical fields such as rainfall was already pointed out by Lovejoy and Schertzer (2013). The same authors argue that the estimation of the generator over small sections of the field in a moving window fashion suffers from the small sample size. Therefore, they propose to model the spatial variation of the variables defining the generator using a nonlinear extension of Generalized Scale Invariance (GSI), whose estimation, however, is also not trivial (Lovejoy and Schertzer, 2013).

3.2.1 Contribution of this study

This chapter explores the concept of model localization in the context of stochastic precipitation nowcasting. The main contribution is the design of an ensemble precip-

itation nowcasting model which can deal with the non-stationary spatial statistics of precipitation fields.
To this end, we first present a non-stationary stochastic rainfall generator based on the Short-Space Fourier Transform (SSFT, Appendix B.4). The idea is to extend the Short-Time Fourier Transform (STFT) for two-dimensional spatial fields both for local Fourier analysis and stochastic simulation. Conceptually similar to the wavelet transform (Kumar and Foufoula-Georgiou, 1997), the STFT and SSFT allow time- and space-frequency localization respectively, that is, they enable performing a local moving window Fourier analysis to account for non-stationarity in the statistical properties of the signal (in our case a 2D precipitation field). The traditional global FFT-noise technique can be easily extended by performing local Fourier filtering of white noise in a moving window fashion using as filter the local Fourier transform of the precipitation field. The obtained locally correlated noise field (denoted as SSFT-noise) reproduces the local correlation structure of the target precipitation field. With this respect, we also want to investigate to what extent it is reasonable to estimate a local stochastic generator from the data considering the uncertainty due to the reduced sample size.
In the second part of this chapter, the localization approach is implemented within a state-of-the-art ensemble prediction system. To better show the impact of the localization procedure, the analysis includes a visual assessment for a precipitation event displaying marked non-stationarity. Finally, we performed a verification experiment using a set of 10 precipitation events from the MeteoSwiss radar composite archive.
The chapter is organized as follows. Section 3.3 describes the global Fourier-based method for generating spatially correlated stochastic noise fields (FFT-noise) and explains the generation of locally correlated stochastic noise based on the Short-Space Fourier Transform (SSFT-noise). Section 3.4 makes use of synthetic data produced with parametric methods in order to illustrate and test the local approach. Section 3.5 presents the nowcasting application and the results of the verification experiment on a set of real data. Section 3.6 discusses the findings and provides some general conclusions.

3.3 Stochastic generators

3.3.1 Global filtering of white noise with FFT

The straightforward way of generating a stationary correlated noise field, that is, a stochastic field with the overall spatial correlation of a natural rainfall field, is the non-parametric filtering of a white noise field ($\mathbf{n} \in \mathbb{R}^{M \times N}$). Non-parametric filtering implies the convolution of a generated white Gaussian noise field, characterized by

zero mean and unity variance, with a given rainfall field $\mathbf{r} \in \mathbb{R}^{M \times N}$. If we employ the FFT algorithm to calculate the DFT of a rainfall field as $\mathbf{R} = \text{FFT}\{\mathbf{r}\}$, and the DFT of a noise field as $\mathbf{N} = \text{FFT}\{\mathbf{n}\}$, a stationary correlated noise field ($\mathbf{n}_{scr}$) can be derived as:

$$\mathbf{n}_{scr} = \text{FFT}^{-1}\{\mathbf{R} \odot \mathbf{N}\}, \tag{3.1}$$

with $[\cdot] \odot [\cdot]$ being the pointwise (Hadamard) product. Given the quasi-constant power spectral density of white noise, this way we mostly alter the spectral phase of a rainfall field, without significantly modifying the amplitude. It makes the resultant spectrum of one realization of correlated noise field ($\mathbf{N}_{scr}$) almost identical to the spectrum of the considered rainfall field ($\mathbf{R}$), whose inverse FFT corresponds to the spatial autocorrelation function (ACF). The latter could also be achieved by generating a noise field which can be used to perturb the phase of $\mathbf{R}$ directly (Metta et al., 2009). In this case, we can avoid calculating the Fourier transform of the noise, $\text{FFT}\{\mathbf{n}\}$, which can slightly contribute to the computational efficiency.
The most important drawback of the obtained stationary correlated noise field is its spatial stationarity. Namely, as already pointed out, the obtained spectrum ($\mathbf{N}_{scr}$) corresponds to the one of the selected rainfall field ($\mathbf{R}$), which implies that the two fields have the same global field anisotropy properties and scaling behavior. This, however, does not imply that a correlated noise field exhibits the same local behavior as its rainfall field counterpart.

3.3.2 Non-parametric and parametric filters

In Section 3.3.1, the white noise field was filtered by the actual DFT of the rainfall field without any parametrization of the distribution of power as a function of frequency and direction. This is referred to as non-parametric filtering of white noise and allows automatic simulation of the observed anisotropy and variability of the spectral slope as a function of scale. This approach is particularly useful for cases with sufficient precipitation coverage over the radar domain and in the absence of measurement noise. On the other hand, for cases with a small fraction of precipitation, the Fourier transform suffers from the small sample size (Seed et al., 2013). As a consequence, the power spectrum becomes noisy and it may be hard to observe the anisotropy or the power law rainfall scaling behavior. In order to increase the sample size and reduce the noise, a more stable Fourier power spectrum can be computed by averaging the spectra of a sequence of precipitation fields as done by Niemi et al. (2014).
Alternatively, the rainfall field spectrum ($\mathbf{R}$) can be as well approximated by a suitable power-law scaling model, resulting in the parametric approach to white noise filtering. The parametrization of 2D power spectra can be achieved by constructing an anisotropic filter that considers isolines of constant power as a function of frequency

based on the principle of GSI (Niemi et al., 2014). A simpler isotropic scaling model can be constructed from the radially averaged 1D power spectrum by neglecting the field anisotropy. A power law fit of the rainfall field power vs. spatial frequency in the log-log plot is generally sufficient for the task. More refined models acknowledge the presence of a scaling break in the power law around 10-20 km by fitting two spectral exponents (β_1 and β_2), which represent the different scaling regimes of the large scale precipitation features and the small scale convective cells respectively (Gires et al., 2011; Seed et al., 2013).

3.3.3 Local filtering of white noise with SSFT

The most obvious drawback of a conventionally derived correlated noise field is its spatial stationarity. Tackling that issue, which is the main motivation and idea of this work, is conceptually quite simple: it comes down to the replacement of the Fourier transform, in both non-parametric and parametric filters, with the Short-Space Fourier transform (see Appendix B.4).
Namely, instead of deriving either a parametrized or non-parametrized Fourier spectrum of the entire rainfall field in order to filter white noise, here we derive a number of local spectra by focusing at different parts of the field. This is done by sliding and centering the window at positions[1]:

$$\mathbf{R}(p_1, p_2, m, n) = \frac{1}{\sqrt{MN}} \sum_{l=0}^{N-1} \sum_{k=0}^{M-1} \mathbf{r}(p_1 + k, p_2 + l)\mathbf{w}(k, l)e^{-j2\pi(km/M + ln/N)}, \tag{3.2}$$

where p_1 varies in the range $(1, N/\Delta)$ (equivalent for p_2), with $2T \times 2T$ being the size of the non-zero portion of the window function. The maximum value corresponds to the size of the window ($2T$) to avoid gaps, yet the overlapping ($\Delta < 2T$) can ensure a smoother representation of the non-stationarity. Although the integration could be performed only in the non-zero part of the matrix $\mathbf{w}$, the window is zero padded to the size of the field $\mathbf{r}_{M \times N}$, resulting in a better frequency sampling resolution.
In the following step, the local spectra are convolved with a white noise field $\mathbf{n}$, producing the local spatially correlated noise:

$$\mathbf{n}_{scr}(p_1, p_2) = \mathrm{FFT}^{-1}\{|\mathbf{R}(p_1, p_2)| \odot \mathbf{N}\}. \tag{3.3}$$

By using only the amplitude spectrum of the rainfall field (its absolute value), rather than perturbing the phase of the rainfall field, we are imposing the random phase of the white noise field, and slightly perturbing the amplitude. This way, we preserve the local anisotropy, while simultaneously benefiting from the phase coherence in the final

[1]See also Appendix B.5 for a discussion on the choice of the window function.

recomposition, that is, it decreases the need for the overlapping. The recomposition of the final non-stationary field is obtained by summing over all the local spatially correlated noise fields:

$$\mathbf{n}_{scr} = \sum_{p_1=1}^{N/\Delta} \sum_{p_2=1}^{M/\Delta} \mathbf{n}_{scr}(p_1, p_2). \tag{3.4}$$

where N/Δ is the number of windows in the k direction and M/Δ is the number of windows in the l dimension. In other words, if the rainfall field $\mathbf{r}_{M\times N}$ is represented by a set of 10x10 local FFTs (without overlapping), one must sum over the 100 correlated noise fields, which are weighted by the window for spatial localization of the noise.

3.4 Synthetic data experiments

The novel concept can be intuitively demonstrated, or rather illustrated, using a synthetic dataset. Namely, instead of using (3.1) (i.e., the non-parametric approach), here we rely on parametric filtering in order to prescribe a power spectrum to a field of Gaussian white noise. The most straightforward approach is to use the isotropic power law filter

$$H(k) = k^{-\beta/2}, \tag{3.5}$$

where $H(k)$ is the parametric filter as a function of k, the spatial frequency or wavenumber. It follows a stationary, isotropic field of correlated noise whose spectral slope is nearly equal to β. It can be seen in Fig. 3.2 that β controls the smoothness of field, or autocorrelation range, as a more negative β results in a more pronounced spectral slope and consequently in a higher power at larger wavelengths.

3.4.1 Non-stationary autocorrelation range

The above parametric filtering can be applied locally so that a set of β is used to control the autocorrelation range within successive windows over the same field of Gaussian white noise. Figure 3.3a presents a set of non-stationary ACFs in terms of autocorrelation range overlaid on the resulting correlated noise. In other words, these parametric 2D ACFs are used to filter white noise, locally, using a set of windows without overlapping. This image represents our target field. It is easy to notice that the autocorrelation range of the background noise field decreases when going from the top left to the bottom right corner of the image. In the following experiment, we compute the global and local Fourier transforms of the target noise field in order to filter the same white noise field but with non-parametric filters. In other words, we want to test

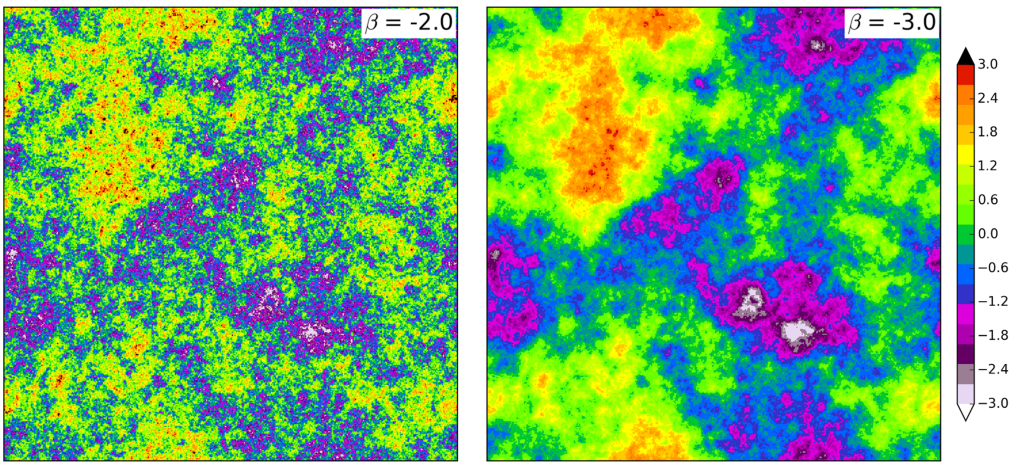


Figure 3.2: Example of 512x512 synthetic fields of correlated noise produced with an isotropic power law filtering of Gaussian with noise using two different values of β. Both noise fields have been drawn from the standard normal distribution and share the same random seed.

whether the local SSFT can learn the prescribed parametric spectra directly from the target noise field. Given such heterogeneous target field, the application of the global non-parametric generator will fail to capture the trend in the spectral slope. This can be seen in Fig. 3.3b, where the noise field has homogeneous characteristics throughout the image, and it is confirmed by the black contours representing the resulting local non-parametric 2D ACFs.
Conversely, the local non-parametric generator applied in Fig. 3.3c closely reproduces the autocorrelation structure that was prescribed in the target image. The spatial coherence between different windows, that is, the continuity in absolute values at the edges of consecutive windows, is the result of convolving the white noise field with the amplitude spectrum of the rainfall field only. A reasonable amount of overlapping between windows (e.g., 10%) would be sufficient to obtain a smoother representation of the non-stationary anisotropy or autocorrelation range in the presence of stronger gradients.

3.4.2 Non-stationary anisotropy

As the next step, the method was tested in its ability to reproduce the locally varying anisotropy of a synthetic target image. This target image was produced using the GSI model as presented by Niemi et al. (2014). The GSI method has the advantage of being able to model the scale dependence of anisotropy using only a few parameters. It should be noted that the parameterization of anisotropy can also be done with 2D anisotropic ACFs as can be found in the field of geostatistics. In such cases, the corre-

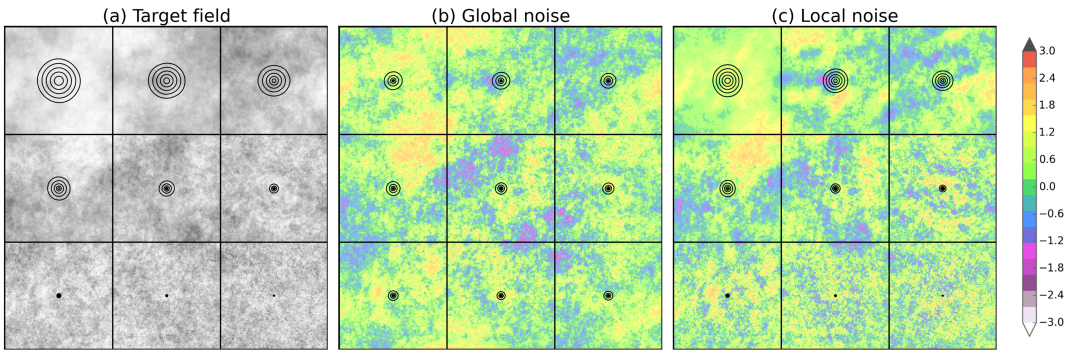


Figure 3.3: (a) An example of 512x512 pixel non-stationary correlated Gaussian noise generated using parametric filtering. The spectral slope β increases from -3.0 to -1.5 from the top towards the bottom of the image as represented by black contours at 0.5, 0.6, 0.7, 0.8 and 0.9 correlation coefficients of the corresponding ACF. Non-parametric stochastic realizations of the target field using a global generator (b) and a local SSFT generator (c) with Hanning windows centered over the grid boxes in the image. All simulations have been drawn from the standard normal distribution and share the same random seed.

sponding filter in the Fourier space is obtained with the Wiener-Khintchine theorem (Appendix B.3).

In a similar way as in Fig. 3.3, the set of arbitrary parameters of the GSI model was spatially varied in order to produce a target image with changing anisotropy. The result of the simulation is presented in Fig. 3.4a. We tested again the ability of the global and local non-parametric generators to reproduce the spatial heterogeneities introduced in the target field, and the results are presented in Fig. 3.4b and Fig. 3.4c. While generally the global generator could only learn and reproduce the same average anisotropy, the local approach managed to localize the distinct patterns of the field. In particular, the method is effective in producing the correct angle of anisotropy.

In the non-parametric local realizations presented in Fig. 3.3c and Fig. 3.4c, one can notice a certain contraction of the contours of the ACFs compared to the references in the target field. This means that the non-parametric realizations exhibit slightly lower correlation ranges. We believe that this is caused by a combination of minor sources of error associated with the use of the Hanning window and imperfections in the white noise. Hardly preventable, these minor deviations are an integral part of the method. It should also be noted that some degree of localization can still be found in the result of the global non-parametric simulations. This is believed to be the effect of the residual phase information contained in the global Fourier spectrum used for the white noise filtering. However, this will not be the case with a parametric filter, as no phase information will be available.

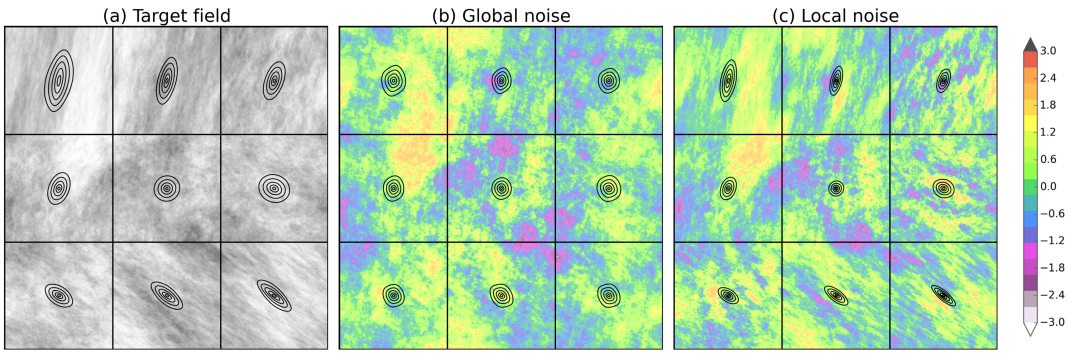


Figure 3.4: (a) An example of 512x512 pixel non-stationary correlated Gaussian noise generated using parametric filtering. The GSI parameters (see Niemi et al., 2014, for details) linearly evolve from $G(-0.12, -0.12, -0.12)$ and $I_s = 5$ at the upper left corner to $G(0.12, 0.12, 0.12)$ and $I_s = 5$ at the lower right corner of the image. The black contours show the corresponding ACFs at 0.5, 0.6, 0.7, 0.8 and 0.9 correlation coefficients. Non-parametric stochastic realizations of the target field using a global generator (b) and a local SSFT generator (c) with Hanning windows centered over the grid boxes in the image. All simulations have been drawn from the standard normal distribution and share the same random seed.

3.4.3 Effect of the window size

The side-length of the window function used to estimate the local Fourier spectrum is expected to affect the accuracy of the estimation. Intuitively, if a too large window is used, then the localization gets less and less informative and the assumption of stationarity within the window weakens. The global approach represents a particular case where the window size equals the image size. On the other hand, if a too small window is used, we lose information at wavelengths larger than the window itself and the filter may become ill-defined due to limited sample size.
In Fig. 3.5, nonzero synthetic realizations of 512x512 pixels were produced using (3.5) and β =-3.0, -2.5, -2.0 and -1.5. We then applied the SSFT method with decreasing window sizes of 512, 256, 128 or 64-pixel width to produce 20 stochastic realizations for each of the windows. Finally, the average 1D power spectra of the resulting fields of correlated noise were plotted in Fig. 3.5a, alongside the benchmark representing the 1D power spectra of the original image. In order to facilitate the comparison, all spectra were centered by removing the mean value. The idea was to investigate the impact of decreasing window sizes on the ability to reproduce the global power spectra of the original image. More generally, the experiment verified the presence of biases in the localization routine. For instance, any issue related to numerical diffusion

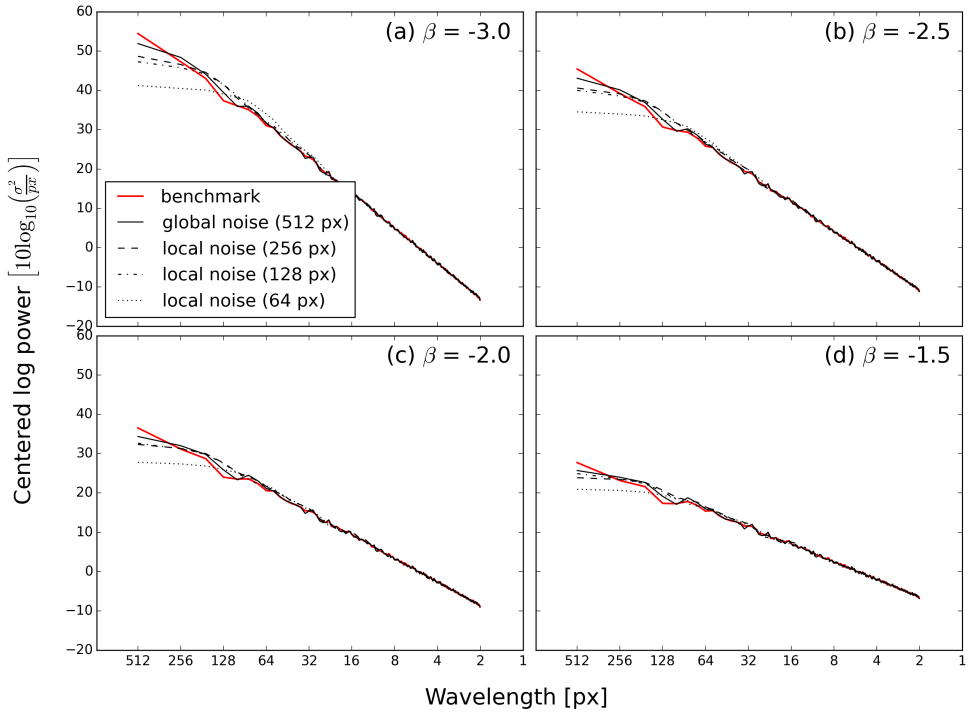


Figure 3.5: The radially averaged 1D centered power spectra of synthetic realizations with spectral slope (a) $\beta = -3.0$, (b) $\beta = -2.5$, (c) $\beta = -2.0$ and (d) $\beta = -1.5$ (in red). In black are the corresponding average spectra of 20 non-parametric realizations for decreasing window sizes.

(e.g., because of the overlapping between successive windows) would be detected as a loss of power at small wavelengths (e.g. Germann and Zawadzki, 2002).
The effect of reducing the window size is visible in Fig. 3.5, where we note a loss of power for wavelengths larger than the window itself. The magnitude of the deviation is more and more important for smaller window sizes and can, therefore, represent a limitation in the reproduction of scales larger than the chosen window size. However, Fig. 3.5 also shows that the effect gets less important as the spectral slope of the original field decreases. This outcome suggests that the choice of the window size, or rather its lower limit, should consider the spatial characteristics of the target field, namely the presence of large scale features. Finally, we do not note any loss of power for small wavelengths, which indicates the absence of numerical diffusion issues due to the windowing approach.

3.4.4 A nested Fourier transform

As seen above, a practical problem of the local Fourier analysis is a loss of power at wavelengths larger than the window size, which becomes more accentuated for decreasing window sizes and increasing spectral slopes.
We explored a few possible solutions to the problem. A promising approach consists in the use of a nested approach to Fourier Transform (NFT). In essence, the rainfall field is subsequently divided into smaller and smaller boxes for local Fourier analysis. In this way, the global Fourier spectrum can be successively updated with increasingly local Fourier spectra, while the information at the larger scales is conserved.
Figure 3.6 shows some preliminary results with the NFT approach. The problem of using a too small localization window in the SSFT method is easily noticeable both in the simulated field (lower left panel) and in the corresponding power spectrum (upper right) as a loss of power at large scales. In contrast, the nested approach is able to reproduce the large-scale structures of the original field better and consequently its power spectrum.
NFT is still at an early stage and it will not be included in the subsequent analyses of this chapter.

3.5 The short-space ensemble prediction system

The SSFT-noise fields discussed above can reproduce the local power spectrum and spatial ACF of an observed rainfall field, but they do not replicate the local marginal statistics (e.g., mean, variance and wet area ratio) which are essential for practical applications.
This experiment applies the same short-space approach on the whole precipitation nowcasting procedure and investigates the sensitivity of the forecast skill on localization. For this purpose, a state-of-art ensemble nowcasting model is run on a set of precipitation events over Switzerland with an increasing degree of localization. The forecast skill is then evaluated using a set of standard probabilistic verification methods.

3.5.1 Description of the nowcasting model

The Short-Term Ensemble Prediction System (STEPS) is an ensemble-based probabilistic precipitation nowcasting model that was developed at the Australian Bureau of Meteorology (BOM) in collaboration with the UK MetOffice. The detailed description of STEPS can be found in Bowler et al. (2006) and Seed et al. (2013), while the description of its recent open-source implementation is available in Pulkkinen et al.

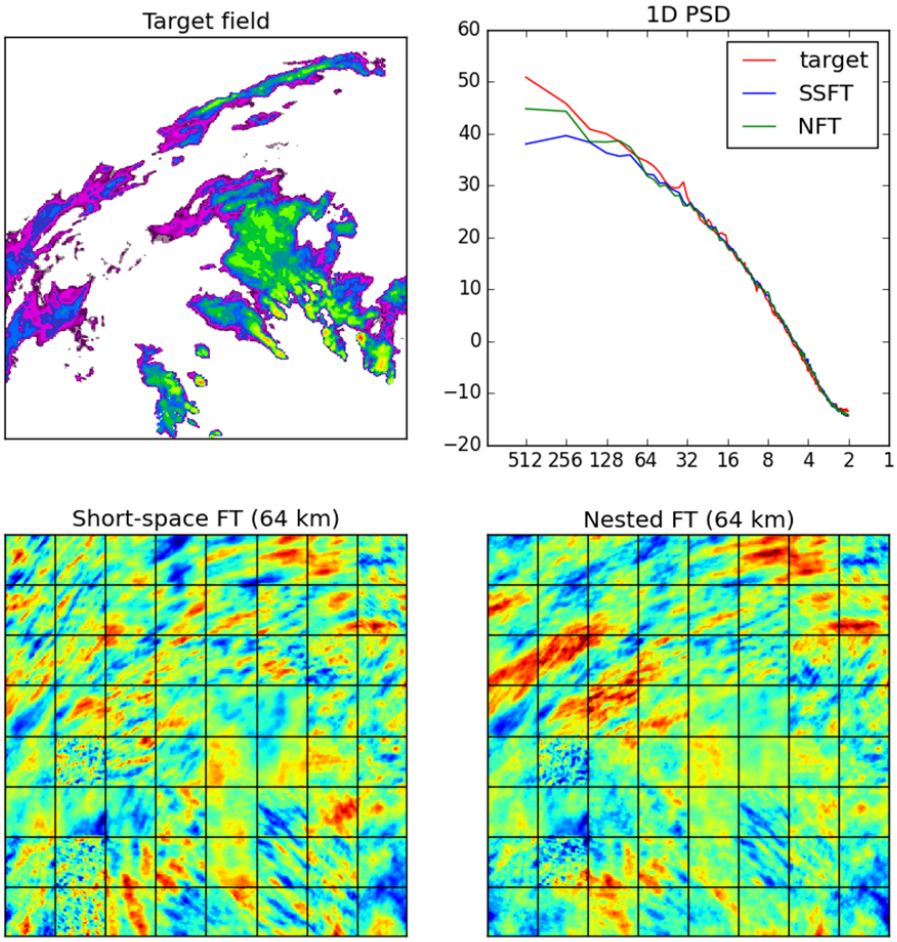


Figure 3.6: Preliminary results of the nested Fourier transform (NFT) method. (Upper left) The target radar field valid at 1630UTC 15 May 2015. (Lower left) One stochastic realization from the SSFT approach using a 64 km window. (Lower right) One stochastic realization from the NFT approach using a 64 km window. (Upper right) The corresponding 1D radially-averaged power spectra.

(2019b). Here we provide a summary and a description of the localization procedure. A summary of the model parameters is provided in Table 3.1.

The key idea of STEPS is to decompose the precipitation field into a multiplicative cascade in order to account for the wide range of temporal variability that is found across spatial scales (e.g., Seed, 2003, for technical details see Appendix B.6). This time-space relationship is better known as *dynamic scaling* (Venugopal et al., 1999).

The temporal evolution of precipitation within the nowcast is thus simulated by an auto-regressive (AR) process that combines the deterministic persistence component with a stochastic innovation term. The combination of an auto-regressive model in time and cascade decomposition in space allows controlling the rate of the temporal and spatial evolution of precipitation. In practice, for each cascade level j, the recursion formula for a general AR(p) process is given by

$$R_j(x,y,t) = \sum_{k=1}^{p} \phi_{j,k} R_j(x,y,t-k\Delta t) + \phi_{j,0}\varepsilon_j(x,y,t). \tag{3.6}$$

The first term corresponds to the predictable component at cascade level j (i.e., the Lagrangian persistence). The second term is a stochastic term that represents the unpredictable component at the same cascade level j, that is, mainly initiation, growth and decay of precipitation at a specific spatial scale. The symbol Δt denotes the time difference between consecutive precipitation fields R_j that are normalized to zero mean and unit variance.

It should be noted that equation (3.6) assumes a normal distribution of R which is enforced by applying the logarithmic transformation

$$R \to \begin{cases} 10\log_{10} R, & \text{if } R \geq 0.1 \text{ mm h}^{-1} \\ -15, & \text{otherwise.} \end{cases} \tag{3.7}$$

This results into logarithmic radar rain rates (units of dBR). The value of -15 dBR is equivalent to assigning a rain rate of approximately 0.03 mm h^{-1}to the zeros. Using the logarithmic transformation is motivated by the fact that rain rates are approximately log-normally distributed (Crane, 1990).

The parameters $\phi_{j,k}$ in equation (3.6) are estimated from time-lagged auto-correlation coefficients $\rho_{j,k}$ for $k = 1, 2, \ldots, p$ using the Yule-Walker equations (Hamilton, 1994). For $p = 2$, the correlation coefficients can be adjusted to ensure that the resulting AR(p) process is stationary in time and non-periodic (Box et al., 2013). Finally, the parameters $\phi_{j,0}$ are chosen as

$$\phi_{j,0} = \sqrt{1 - \sum_{k=1}^{p} \rho_{j,k}\phi_{j,k}}. \tag{3.8}$$

Given that the variance of the noise fields ε_j is one, this choice guarantees that the AR(p) process is normalized to unit variance (Hamilton, 1994).

The perturbation field ε in equation (3.6) can be produced with the approaches described in Sect. 3.3, namely the parametric, non-parametric and SSFT noise generators. Following the assumption of temporal stationarity, a final post-processing step is included to ensure that the forecast fields have the same statistical properties of the last observed precipitation field. In practice, the cumulative distribution function (CDF) of the forecast field is mapped to the observed one. This is defined as

$$R'(x,y) = F_{\text{obs}}^{-1}(F(R(x,y))), \tag{3.9}$$

where F_{obs} and F denote the CDFs of the observed and the input forecast field R, respectively.

The experiment of this section investigates the impact of localization on the nowcast quality. For this purpose, we seek at using a spatial subset of the observations in order to estimate the model parameters so that they can vary in space. The short-space approach is generalized to the whole nowcasting system. Conceptually similar to the box approach in Sideris et al. (2018), the method essentially implements a moving-window localization of the nowcasting procedure, whereby all parameters are estimated over a subdomain of prescribed size. The localization is applied to the cascade decomposition, the autoregressive process, the non-parametric Fourier filter, and the probability matching. This localized version of the STEPS model is named SSEPS, as in Short-Space Ensemble Prediction System. The size of the moving window controls the degree of localization. Some degree of overlapping is introduced to ensure a smooth transition between successive windows.

Table 3.1: The pysteps model configuration used in the experiments, see Pulkkinen et al. (2019b) for a detailed description of the parameters.

Parameter	**Value**
optical flow	Lucas-Kanade
extrapolation	semi-Lagrangian
cascade levels	6
order of the AR(p) model	2
precip. perturbations	SSFT
transformation	R to dBR
minimum precipitation	0.1 mm h^{-1}
value for dry pixels	-15 dBR
mask method	None
ensemble size	24
overlap	0.1
probability matching	yes
velocity perturbation	None

3.5.2 Description of the data

The latest 4th generation MeteoSwiss network consists of five polarimetric C-band Doppler radars (Germann et al., 2015). The quantitative precipitation estimation (QPE) product used in this study includes automatic hardware calibration, clutter filtering, correction for beam shielding, correction for VPR effects, Z-R relation $Z = 316R^{1.5}$, and bias adjustment (Germann et al., 2006a). The radar composite is cal-

culated on a 1 km grid every 5 minutes. The spatial resolution of the final Cartesian radar grid is 1 km, and the temporal sampling is 5 min. The size of the radar domain is 710x640 km, and it is centered over Switzerland.
Ten events consisting of predominantly convective precipitation were chosen from the Swiss radar archive and are described in Table 3.2.

3.5.3 Verification metrics

In verifying probabilistic forecasts, we need to assess the *reliability* of the ensemble, that is, its ability to provide samples from the same underlying distribution of the observation (Weigel, 2012). For a reliable ensemble, therefore, members and observations are statistically identical.
First, the rank histogram (Hamill, 2001) measures how well the ensemble spread corresponds to the observed uncertainty. For each nowcast grid pixel, the ensemble members are ranked in increasing order. A pooled histogram is computed by assigning each verifying observation a bin which it falls into among the ensemble members. The first and last bins are assigned for observations below or above all members, respectively. For a forecast ensemble whose distribution is consistent with the observations, the histogram is flat, and no observations fall into the first or last bin. To handle ties (e.g., when both the observed precipitation and several ensemble members are equal to 0), we implemented the method of Hamill and Colucci (1997). The method randomly chooses a bin between $(M+1)$ and $(M+M_{tied})+1$, where M is the number of members smaller than the observation and M_{tied} is the number of ties (ensemble members equal to the observation).
An additional metric that can be derived from rank histograms is the outlier percentage (OP). The OP measures the proportion of observations falling outside the ensemble, defined by

$$OP = \frac{h_1 + h_{n+1}}{\sum_{i=1}^{n+1} h_i}, \tag{3.10}$$

where h_i denotes the number of samples in the i-th bin of the rank histogram and n is the ensemble size. It follows that the reference OP for a perfectly calibrated ensemble (i.e., a flat histogram) is equal to the sum of the expected frequencies in the first and last bins, or $2/(n+1)$. The deviation of the OP value from the reference can be interpreted as a measure of over- or underdispersion.
Third, the reliability diagram (Bröcker and Smith, 2007) measures the bias (reliability) and the resolution of a probabilistic forecast. For a given intensity threshold, the diagram shows the forecast probability against the observed frequencies, where the probability range $[0,1]$ is divided into n bins. For a reliable forecast, the curve lies on the diagonal. The reliability diagram is often accompanied by a histogram showing

Table 3.2: Precipitation events in Switzerland (MeteoSwiss). The duration of each event is 12 hours.

No.	Date	Start time (UTC)	Description
1	16 Apr 2016	18:00	Prefrontal precipitation induced by a low pressure system over the North Sea. Lines of convection develop over western Switzerland.
2	11 Jul 2016	13:00	An approaching cold front causes widespread convective activity in a south-westerly flow.
3	31 Jan 2017	10:00	A strong north-westerly flow causes orographic blocking on the northern slopes of the Alps resulting in widespread precipitation.
4	14 Jun 2017	13:00	Fairly uniform pressure distribution across Central Europe, air mass convection develops in the afternoon.
5	24 Jun 2017	22:00	Prefrontal activity with intense thunderstorms south of the Alps. Measured peak intensity reached 33.5 mm in 10 min and presence of large size hail stones (3-5 cm) was observed.
6	27 Jun 2017	20:00	A frontal passage during the night induces organized convection over the domain and important prefrontal convective activity on the southern side of the Alps.
7	19 Jul 2017	13:00	In a south-westerly flow, development of large convective cells over central Switzerland.
8	21 Jul 2017	13:00	Flat pressure distribution across central Europe, South-westerly flow associated to a low over the British Islands. Clusters of intense thunderstorms develop over western Switzerland.
9	29 Jul 2017	13:00	South-westerly flow connected to large depression over British Islands. Large clusters of convection develop south of the Alps.
10	31 Aug 2017	14:00	Strong south-westerly flow over the Alps associated to a cold front. Important lines of stationary convection affect the southern Alps, while more stratiform precipitation occurs in the West and North of Switzerland.

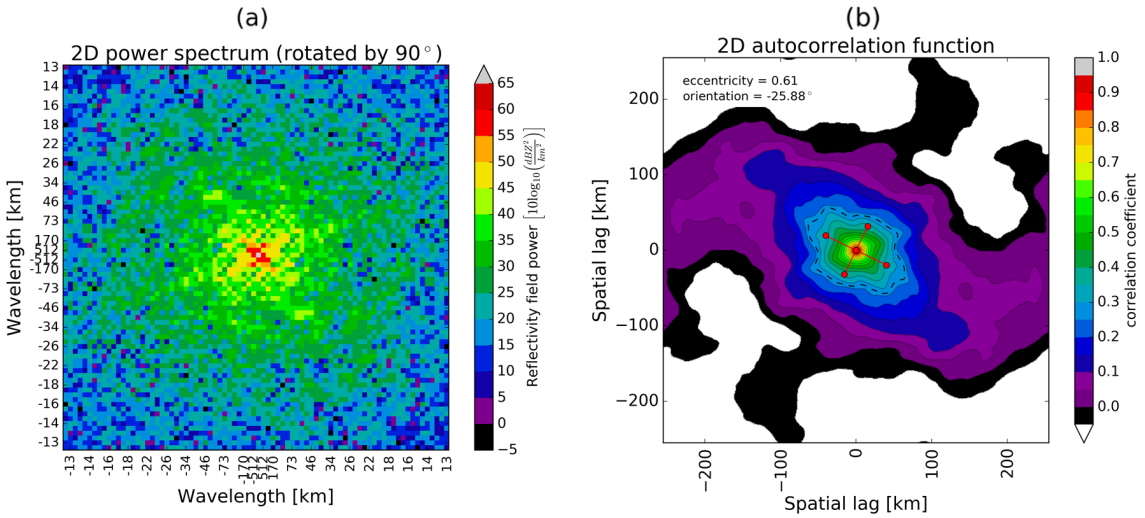


Figure 3.7: The spatial characteristics of the precipitation radar composite valid at 1600 UTC 15 May 2015 (Fig. 3.1). (a) 2D Fourier spectrum zoomed on wavelengths $>$ 13 km and rotated by 90° and (b) corresponding 2D ACF. The ACF is obtained as inverse FFT of the 2D power spectrum. The 90° rotation is performed in order to align the anisotropies of the 2D spectra and spatial ACFs.

the sample size in each bin (sharpness diagram). A sharp forecast has few samples in the middle of the histogram and many on the sides (probability of either 1 or 0).
Finally, the relative operating characteristic (ROC) curve (Jolliffe and Stephenson, 2003) measures the ability of a probabilistic forecast to discriminate the exceedance of a given intensity threshold (i.e., its resolution). For a set of probability thresholds, the ROC curve is constructed by plotting the probability of detection (POD) against the false alarm rate (POFD), which is not to be confused with the false alarm ratio (FAR). For a perfect forecast, the curve passes through the upper left corner (i.e., 100% hit rate and 0% false alarm rate). The area under the ROC curve can be used as a measure of potential skill.

3.5.4 An example of localized nowcasting

To demonstrate the advantage of the localized nowcasting approach, we selected an example of non-stationary precipitation case, which was already presented in Fig. 3.1. On 15 May 2015, a depression system moved from France towards the Ligurian Sea, establishing a south-easterly flow over the Alps and the presence of an active line of convergence in the western part of the Ticino region. As a result, frequent showers and convective storms were triggered over the Po Valley and the southern Prealps. At the same time, narrow bands of low-intensity precipitation were distributed on the northern and western portions of the domain while moving Southwest. Overall, the

apparent motion field had an anticlockwise rotation about the center of the image and was responsible for the appearance of the local anisotropies. The spatial characteristics of the precipitation radar composite valid at 1600 UTC 15 May 2015 are presented in Fig. 3.7. The two distinct precipitation structures have pronounced anisotropies at 90° from each other (Fig. 3.7a), which meant that they almost cancel out once averaged together over the field, finally resulting in a fairly isotropic 2D ACF (Fig. 3.7b).
Figure 3.8 presents a sequence of stochastic precipitation nowcasts started at 1600UTC 15 May 2015 and using the STEPS and SSEPS models, that is, with and without localization. The visual assessment for increasing lead times (+10, +30 and +60 min) highlights the advantages of the local SSEPS approach.

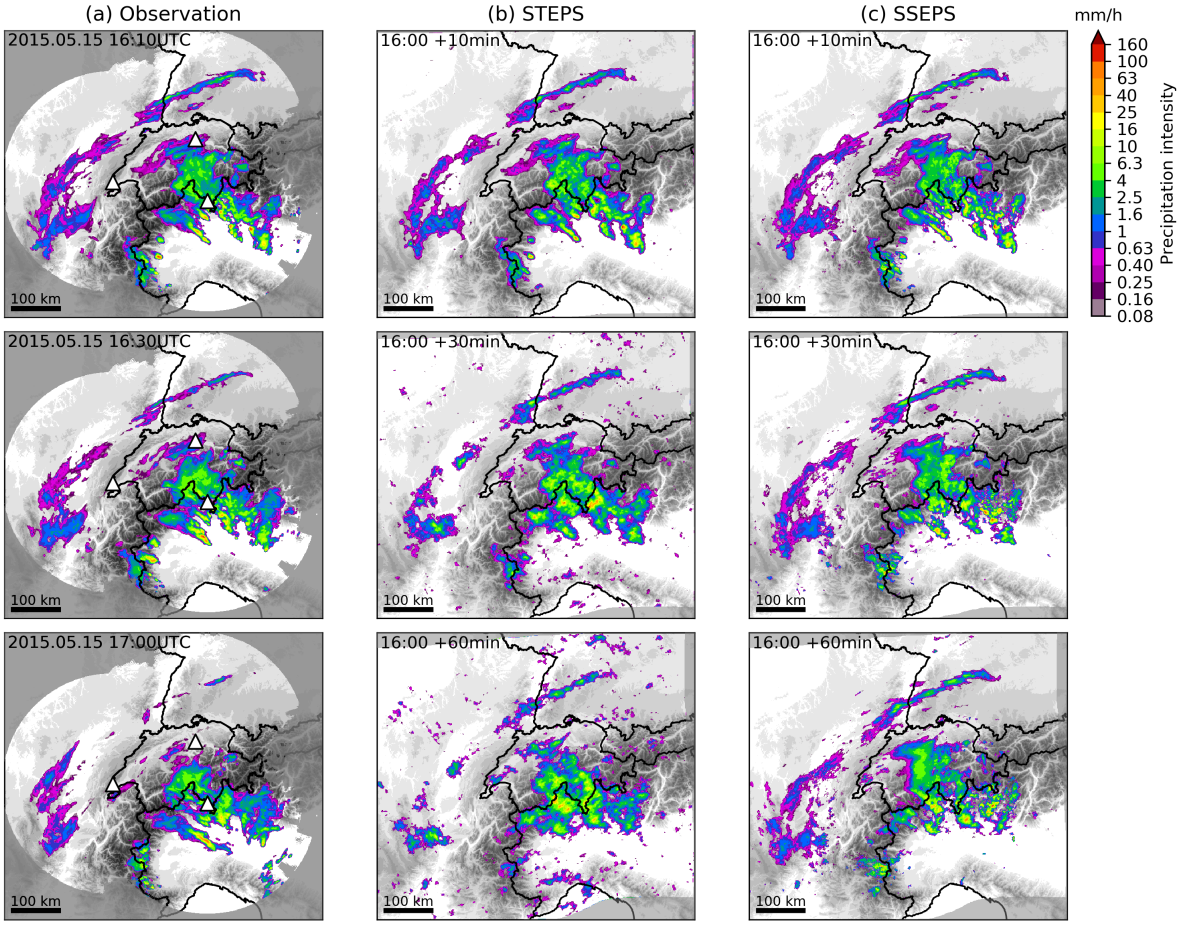


Figure 3.8: An example of the localized nowcasting procedure with the same precipitation event as in Fig. 3.1. (a) The sequence of radar observations valid at 1610, 1630 and 1700 UTC 15 May 2015. (b) The first member of a STEPS nowcast (i.e., global approach) initialized at 1600 UTC on the same day. (c) The first member of the SSEPS nowcast (i.e., local 100 km window) initialized at 1600 UTC on the same day.

First, we note that the precipitation patterns in the STEPS forecast evolve towards an average spatial structure that, in the presence of a marked spatial non-stationarity as in this case, appears unreal. Instead, the evolution gains realism when the local SSEPS

approach is employed, as non-stationarity is better reproduced in terms of anisotropy and correlation length. In other words, thanks to the local modulation of the shock-term, growth and decay processes take place in a spatially consistent way so that new cells appear with the same anisotropy and spatial autocorrelation as existing nearby cells.

Second, the localization can also capture the non-stationarity in the rain rate distribution, as the low-intensity precipitation system in the West and the presence of convective cells in the southwestern quarter of the domain. In such cases, the localized probability matching is able to conserve the same rain rate distribution of the last radar observation, including the absence of rain in areas far from the precipitation system.

Figure 3.9 provides a view on the ensemble variability at the +60 min lead time. The first 15 members of both the STEPS and SSEPS ensembles are visualized along with the probability fields for the 0.1, 1.0, and 5.0 $\mathrm{mm\,h^{-1}}$thresholds. We generally note that as a result of the localization, the SSEPS ensemble appears more confident than its non-localized counterpart, particularly in the discrimination of wet and dry areas. For instance, the narrow band of precipitation in the northern portion of the domain is reproduced by all members.

Despite the promising result of the visual assessment, a marked improvement of the forecast quality is yet to be demonstrated, as by localizing the nowcasting procedure we also implicitly rely on a more local and hence more strict assumption of Lagrangian persistence. That is to say, by relaxing the assumption of stationarity in space, we enforce stronger stationarity in time, which might as well result in lower skill, especially for longer lead times.

3.5.5 Verification results

The verification experiment consisted of three-hour and 24-member ensemble nowcasts initialized every 3 h within the ten selected precipitation events of 12 h duration each (Table 3.2). The results from all the runs were pooled together and stratified by lead time. To test the impact of localization, the verification experiment was repeated for several localization window sizes (100, 200, 400 and 700 km) and the results compared. Smaller windows imply a stronger localization of the nowcasting procedure. With a 700 km window, the set up corresponds to the STEPS reference run, that is, no localization.

Figure 3.10 shows the percentage of observations falling outside the predicted precipitation range (outlier percentage, OP) as a function of lead time. We first note that localization is effective in correcting the spread of the ensemble in the first hours into the nowcast, where it can reduce the percentage of outliers in the reference run by around 30% [i.e., $100(0.22 - 0.15)/0.22$]. On the other hand, we also notice that a too

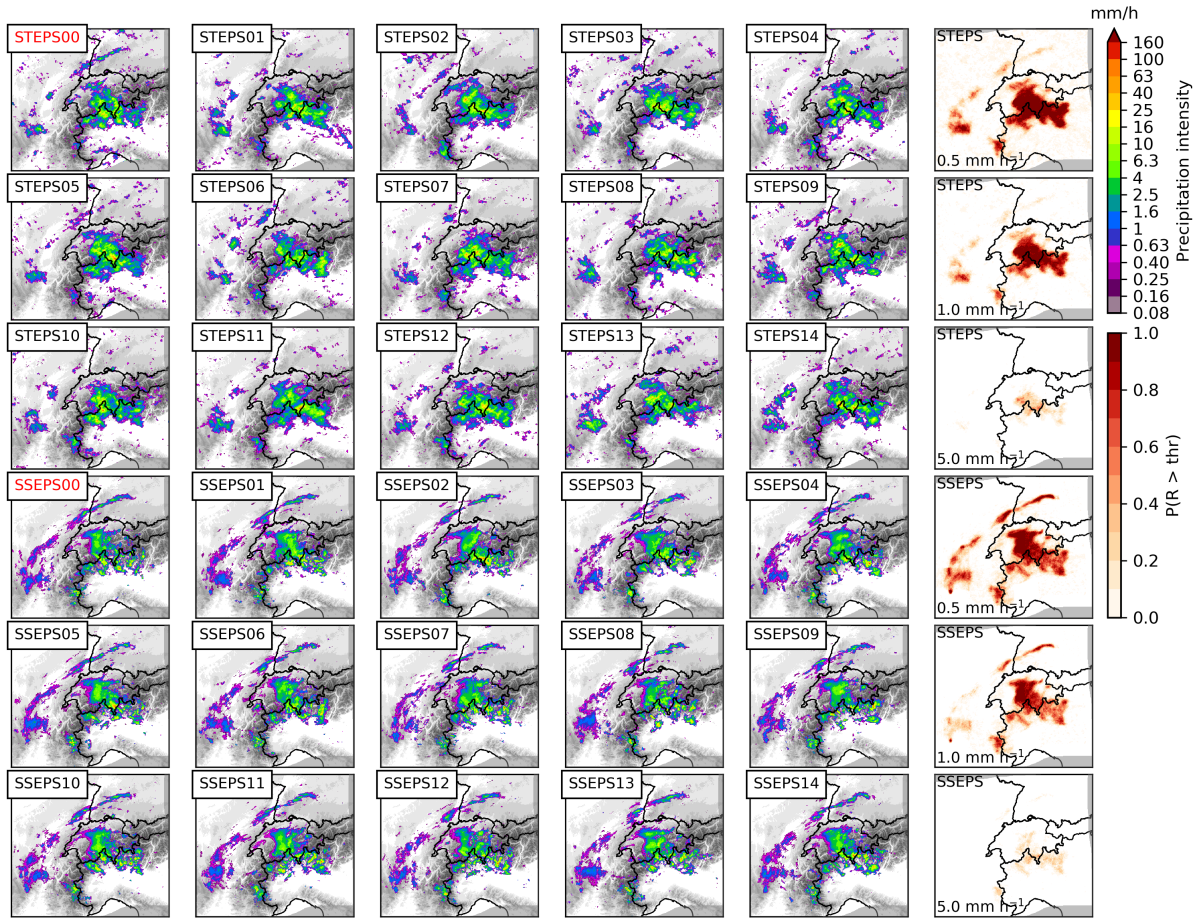


Figure 3.9: Stamp maps of precipitation intensities valid at 1600 UTC +60 min on 15 May 2015. For both STEPS (i.e., global approach) and SSEPS (i.e., local 100 km window), 15 members are illustrated. The last column on the right includes a representation of the corresponding probability fields for the 0.1, 1.0, and 5.0 mm h^{-1}thresholds. The STEPS00 and SSEPS00 maps correspond to the +60 min forecasts in columns (b) and (c) in Fig. 3.8, respectively.

strong localization can produce the opposite effect for longer lead times. The benefits of a 100 km localization window are evident at +30 min, but these vanish within the +110 min lead time. At +180 min, the localized run displays a percentage of outliers that is approximately 30% larger than the reference run [i.e., $100(0.19 - 0.15)/0.15$]. Larger localization windows produce longer-lasting benefits in terms of OP.

A closer look is provided by Fig. 3.11, where the rank histograms at +30 min and +180 min are shown. The rank histograms at +30 min tend to get flatter as the localization window gets smaller (Fig. 3.11a), confirming the positive effect of localization on ensemble reliability. By using a 100 km window, the ensemble underdispersion reaches a minimum with only around 6% of the observations lying below the ensemble and 9% above. By reducing the number of outliers in both ends of the rank histograms, localization can thus be seen as an effective way to correct for the presence

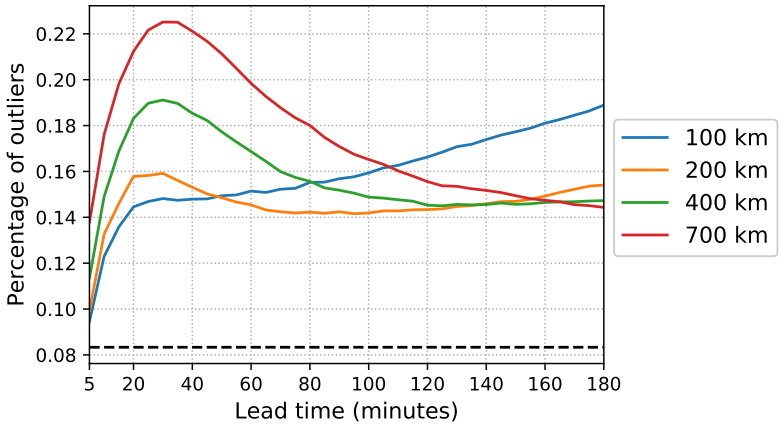


Figure 3.10: The impact of localization on the outlier percentages (OP) computed for all lead times and 1.0 mm h^{-1}. The dashed line represents the reference for a perfectly reliable ensemble.

of conditional biases in the ensemble nowcasts. Conversely, at the end of the nowcast range (+180 min, Fig. 3.11b) the 100 km localization window produces more underdispersion, indicated by the higher frequency of observations falling in the first and last bins of the histogram. The asymmetric shape of the rank histograms is due to the high number of ties occurring when the observation and two or more members are equal to zero.

Figure 3.12 shows the reliability diagrams for the prediction at +30 min of the exceedance probabilities for two intensity thresholds, namely 0.1 and 1.0 mm h^{-1}. In both cases, the localization appears to improve the reliability of the probability forecasts both in terms of reduction of conditional biases and, especially for the higher threshold, overforecasting (i.e., the reduction of a positive bias). On the other hand, we notice a slight decrease of sharpness as localization increases, which can be linked to the higher dispersion of the ensemble (Fig. 3.11a).

We also looked at higher rainfall intensities by means of the ROC curves. Figure 3.13 shows for a 5.0 mm h^{-1}intensity threshold, ROC curves (Fig. 3.13a, +30 min) and ROC areas (Fig. 3.13b, whole nowcasting range). In the ROC curves, the probability threshold that maximizes POD-POFD is also showed. It appears that localization provides some better skill in predicting the occurrence of higher precipitation intensities, which indicates an improvement in terms of probabilistic forecast resolution. Remarkably, in this case the positive impact of localization is present over the whole nowcasting range. After 3 h, the ROC areas are all above 0.5, meaning that the forecast can still provide some potential skill. However, the low probability threshold maximizing POD-POFD is an indication of poor calibration. The optimal threshold is found

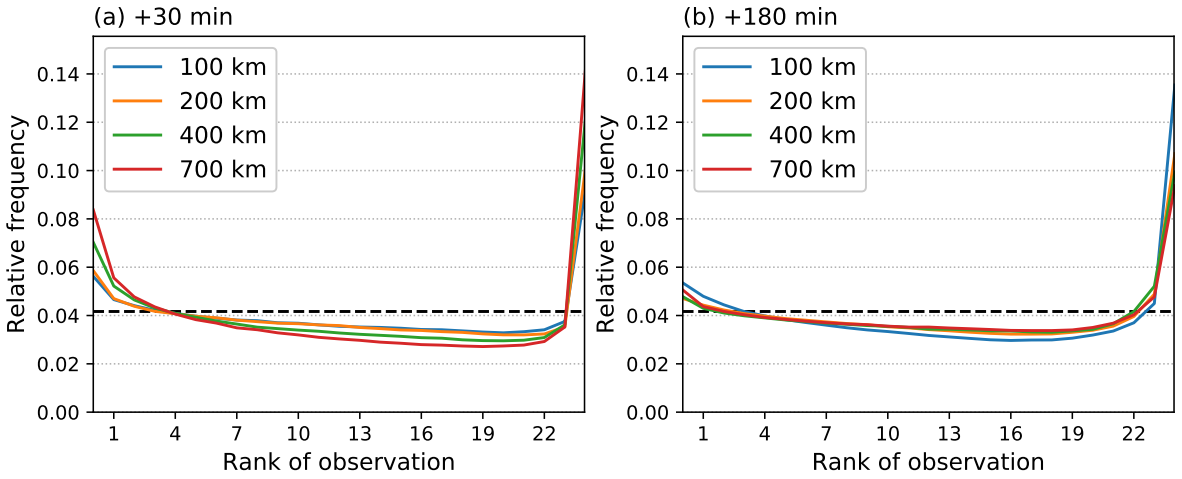


Figure 3.11: Effect of localization on the rank histograms of observation among the ensemble members computed for (a) +30 min and (b) +180 min lead times and 1.0 mm h^{-1}. The dashed line represents the reference for a perfectly reliable ensemble.

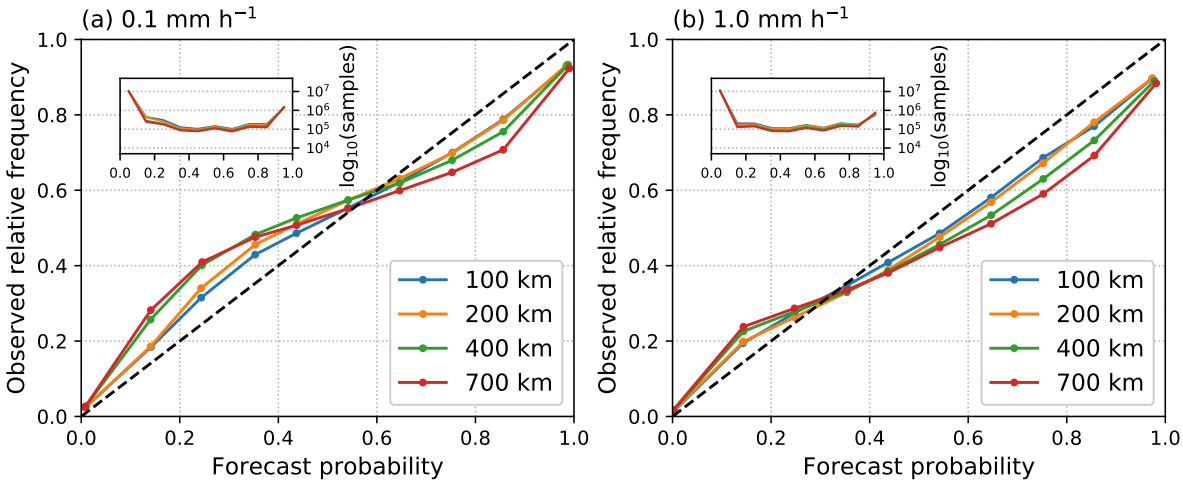


Figure 3.12: Effect of localization in terms of reliability diagrams at +30 min for exceedance probabilities of (a) 0.1 mm $h^{-1}$and (b) 1.0 mm h^{-1}. The dashed lines indicate perfect reliability.

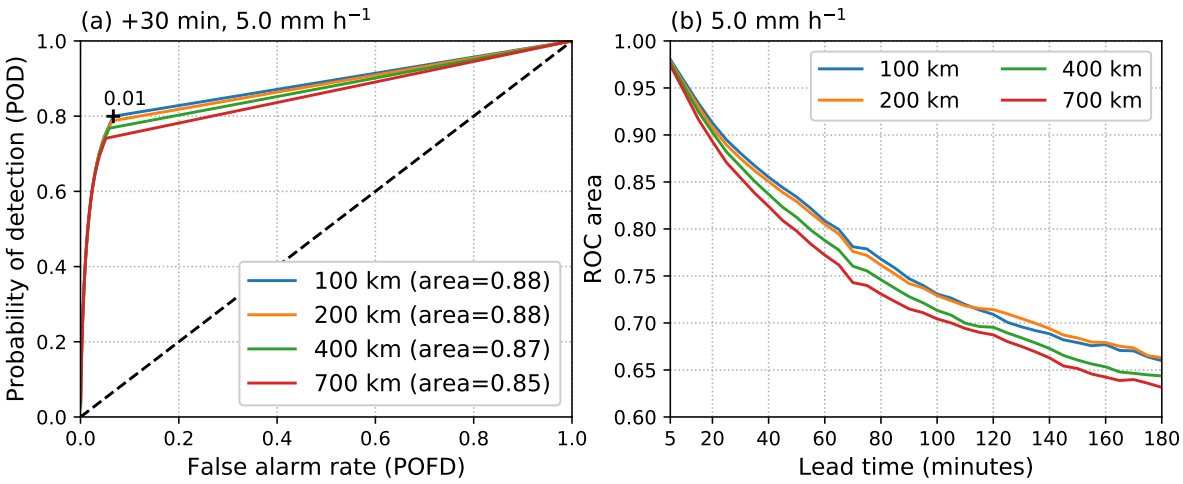


Figure 3.13: Effect of localization in terms of (a) ROC curves at +30 min and 5.0 mm h^{-1}and (b) ROC areas for 5.0 mm h^{-1}. The optimal probability threshold that maximizes POD-POFD is marked in the ROC curves with a black cross. The dashed line indicates a no-skill forecast and corresponds to a ROC area of 0.5.

to be 0.01 at +30 min, while it should be 0.5 for perfectly calibrated probabilities. The probability threshold does not depend on localization (not shown).
We finally analyzed the computation time needed to generate localized nowcast ensembles. In a real-time setting, it is essential to know whether the full nowcast ensemble can be produced before the arrival of the next input radar rainfall image (typically every 5 min). To this end, one-hour ensemble nowcasts were computed with different localization windows and number of parallel threads. The results are shown in Fig. 3.14, which demonstrates that for the input grid of 710x640 pixels used in the MeteoSwiss radar composite, it is possible to generate one-hour SSEPS nowcast with 24 members and a 100 km localization window in less than two minutes using 12 processor cores.

3.6 Conclusions

We proposed a localization approach for the generation of non-stationary stochastic rainfall fields and presented a practical application in the context of probabilistic radar-based precipitation nowcasting. The main idea behind this work consisted in the simulation of a stochastic rainfall field in order to preserve not only the global spatial correlation properties of the target precipitation field but also the spatial distribution of its local correlation structure. This was achieved by employing a moving window in order to estimate the model parameters over contiguous subdomains of prescribed size. This way, we could obtain simulations which locally reproduced the

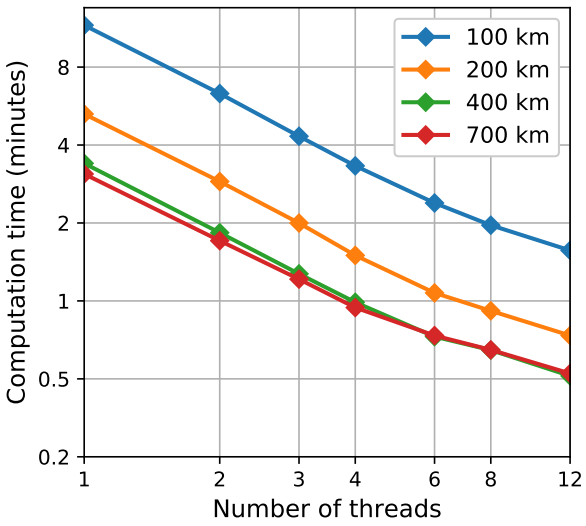


Figure 3.14: Computation time for a 24-member SSEPS nowcast with different window sizes and number of parallel threads. The grid size of the domain is 710x640. One-hour nowcasts with 12 time steps of 5 minutes were computed. The computation times include only the ensemble computation, excluding the optical flow, the initialization of the model and writing the results to disk.

spectral properties of the observed rainfall field, that is, they conserved local correlation lengths and anisotropies. Besides, the local marginal statistics of the precipitation field can be easily accounted for.

First, we illustrated the benefits of the non-stationary generator with synthetic experiments, but also found that the negative effect of localizing the Fourier transform is a loss of power at scales larger than the window size, whose magnitude depends on the global spectral slope. In other words, precipitation fields displaying large scale patterns suffer more from such loss of power. We have shown that ways to account for such effect are currently under development using a nested approach. Other sources of errors associated with the method itself, such as the use of a Hanning window, were found to affect the reproduction of local autocorrelation structure slightly. However, we believe that these limitations are compensated by the benefits introduced by the new approach.

Second, we assessed the impact of the localized ensemble nowcasting approach using a set of 10 precipitation events over Switzerland. We showed that the local shock term helps to preserve the spatial heterogeneities of the original field during the simulation of the temporal evolution of precipitation. More importantly, we demonstrated that localization is beneficial in terms of probabilistic forecast skill, especially during the first few hours into the forecast. We found improvements in terms of forecast reliability and resolution, which were attributed to lower conditional bias in the simulations.

Localization was able to achieve up to a 30% reduction of the ensemble underdispersion. For longer lead times, high degrees of localization were found to reduce skill, arguably because of the stronger assumption of the Lagrangian persistence of the statistical properties of rainfall. For instance, at +180 min, the localization at the scale of 100 km produced an ensemble forecast that was 30% less dispersive than in the reference run.

The need for localization is intuitively important on large domains, where stratiform and convective rain can coexist in different geographical regions, but also on smaller domains characterized by complex terrain, as demonstrated in this study. Besides, these results highlight the importance of defining an appropriate model domain for probabilistic radar nowcasting. That is to say, one that compromises between the need for homogeneous statistical properties (i.e. a small domain) and the need for a robust estimation of model parameters (i.e. large domain). Despite the higher computational requirements, we showed that the localization approach can still meet the needs of a real-time nowcasting application.

Chapter 4

An ensemble approach to seamless blending

This chapter has been adapted from the manuscript:

- Nerini, D., L. Foresti, D. Leuenberger, S. Robert, and U. Germann, 2019: A Reduced-Space Ensemble Kalman Filter Approach for Flow-Dependent Integration of Radar Extrapolation Nowcasts and NWP Precipitation Ensembles. *Mont. Weather Rev.*, **147 (3)**, 987–1006.

It presents a nowcasting blending scheme based on the ensemble Kalman filter, whereby the prediction step is implemented as an ensemble radar extrapolation, while the update step uses the spread of the NWP ensemble (COSMO-E, in this study) to integrate the radar nowcast with the NWP forecast in a Bayesian fashion.

4.1 Summary

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. In order 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 can produce blended forecasts that are at least as skillful as the nowcast-only or the NWP-only forecasts at any lead time.

4.2 Introduction

Following the success encountered in the NWP community (Bauer et al., 2015), several nowcasting systems have introduced the ensemble approach employing stochastic perturbations as a means to quantify the forecast uncertainty (e.g., Bowler et al., 2006; Metta et al., 2009; Berenguer et al., 2011; Atencia and Zawadzki, 2014; Pulkkinen et al., 2019a). 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 also be used to find the optimal combination of radar extrapolation and NWP forecasts.

The 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. In order 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 represent it fully.

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 essential question of this study is, therefore, whether a Bayesian approach based on the EnKF can produce consistent 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 pseudo-measurements. Thus, the typical application

of EnKF in data assimilation is reversed, as observations represent the state being updated with NWP pseudo-measurements.
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 SBMcast 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 data sets.
This chapter is organized as follows. Section 4.3 formulates the Bayesian nowcasting system based on reduced-space ensemble Kalman filter. The developed approach is tested using synthetic experiments in Sect. 4.4 and real precipitation fields in Sect. 4.5. Section 4.6 concludes the paper and discusses possible future developments.

4.3 Formulation of the Bayesian nowcasting technique

Sequential Bayesian filtering consists of three main steps:

a) *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.

b) *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.

c) *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 b) and c). 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 4.2 summarizes the algorithm and its application in the nowcasting context.

4.3.1 Initialization

The true but unknown model state vector $\mathbf{x}$ includes the m model variables that characterize the system of interest. In 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 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]^\mathsf{T} \in \mathbb{R}^{n\times m}$, namely an ensemble of n equally probable realizations $\mathbf{x}_i \in \mathbb{R}^{m\times 1}$ 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 $\overline{\mathbf{x}}$ of the n ensemble members, while $\mathbf{P}$ is approximated with the sample covariance matrix:

$$\begin{aligned} \overline{\mathbf{x}} &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i, \\ \mathbf{P} &= \frac{1}{n-1}(\mathbf{X} - \mathbf{1}\overline{\mathbf{x}}^\mathsf{T})^\mathsf{T}(\mathbf{X} - \mathbf{1}\overline{\mathbf{x}}^\mathsf{T}), \end{aligned} \tag{4.1}$$

where $\mathbf{1} \in \mathbb{R}^{n\times 1}$ is a vector of all ones. The estimation of $\mathbf{P}$ in (4.1) stems from one central idea in the EnKF, namely that 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 mm h^{-1}, 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 using 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, among others with the work of 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, $\mathbf{X}_0 = [\mathbf{x}_0, \mathbf{x}_0, \ldots, \mathbf{x}_0]^\mathsf{T} \in \mathbb{R}^{n\times m}$.

4.3.2 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:

$$\begin{aligned} \boldsymbol{\mu}_{t+1}^f &= \mathbf{F}\boldsymbol{\mu}_t, \\ \mathbf{P}_{t+1}^f &= \mathbf{F}\mathbf{P}_t\mathbf{F}^\mathsf{T}. \end{aligned} \tag{4.2}$$

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 non-linear, is applied to each individual ensemble member in order to propagate the whole ensemble forward in time:

$$\mathbf{x}_{i,t+1}^f = \mathcal{F}(\mathbf{x}_{i,t}), \qquad i = 1, \ldots, n. \tag{4.3}$$

$\mathbf{x}_{i,t+1}^f$ 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 (4.1).
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 using 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:

$$\begin{aligned} &x_{i,j,t+1}^f = x_{i,j+\delta_j,t}, \\ &i = 1, \ldots, n \text{ and } j = 1, \ldots, m. \end{aligned} \tag{4.4}$$

The term on the left-hand side of (4.4) 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 $\boldsymbol{\delta}$ 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., 2019b).

Model errors in the extrapolation forecast are mainly related to the assumption of persistence (Germann et al., 2006b). 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):

$$\begin{aligned} &\mathbf{x}^f_{i,t+1} = \phi_1 \mathbf{x}_{i,t} + \phi_2 \mathbf{x}_{i,t-1} + \boldsymbol{\epsilon}_{i,t+1}, \quad \boldsymbol{\epsilon}_{i,t} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}) \\ &i = 1, \ldots, n. \end{aligned} \tag{4.5}$$

In (4.5), $\mathbf{x}_{i,t}$ 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 (4.4). The global autoregressive coefficients ϕ_1 and ϕ_2 and the variance of the noise term $\boldsymbol{\epsilon}$ 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). Special care for the covariance structure of $\boldsymbol{\epsilon}$ must be taken, and 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 (see Appendix B.6). 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 (4.5) 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 (2003).

4.3.3 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:

$$\boldsymbol{\mu} = \boldsymbol{\mu}^f + \mathbf{K}(\mathbf{z} - \mathbf{H}\boldsymbol{\mu}^f), \tag{4.6}$$

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

$$\mathbf{P} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^f, \tag{4.7}$$

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

$$\mathbf{K} = \mathbf{P}^f\mathbf{H}^\mathsf{T}(\mathbf{H}\mathbf{P}^f\mathbf{H}^\mathsf{T} + \mathbf{R})^{-1}. \tag{4.8}$$

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

$$\begin{aligned} &\mathbf{x}_i = \mathbf{x}_i^f + \mathbf{K}(\mathbf{z}_i - \mathbf{H}\mathbf{x}_i^f), \\ &\mathbf{z}_i = \mathbf{z} + \boldsymbol{\xi}_i, \quad \boldsymbol{\xi}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), \\ &i = 1, \ldots, n. \end{aligned} \tag{4.9}$$

The innovation term in (4.9) 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 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 (4.9). As observation error covariance matrix $\mathbf{R}$ in (4.8) we use the sample covariance matrix of the NWP ensemble fields:

$$\begin{aligned} \overline{\mathbf{z}} &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i, \\ \mathbf{R} &= \frac{1}{n-1}(\mathbf{Z} - \mathbf{1}\overline{\mathbf{z}}^\mathsf{T})^\mathsf{T}(\mathbf{Z} - \mathbf{1}\overline{\mathbf{z}}^\mathsf{T}), \end{aligned} \tag{4.10}$$

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

In a pre-processing 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:

$$\mathbf{K} = \mathbf{P}^f (\mathbf{P}^f + \mathbf{R})^{-1}. \tag{4.11}$$

It is also possible to include in (4.11) a covariance inflation factor, as it 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:

$$\mathbf{x}_i = \mathbf{x}_i^f + \mathbf{K}(\mathbf{z}_i - \mathbf{x}_i^f), \qquad i = 1, \ldots, n. \tag{4.12}$$

where $\mathbf{K}$ is defined in (4.11) 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.

4.3.4 Reduced-space EnKF by means of PCA

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 high dimensional space can become quickly prohibitive. In order 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 (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. Principal component analysis (PCA, see Appendix B.7) 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 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 *principal component (PC) scores*. The state update equations are modified as follows:

$$\begin{aligned} \bar{\mathbf{x}}_d^f &= \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i^f \mathbf{U}, \\ \mathbf{P}_d^f &= \frac{1}{n-1} (\mathbf{X}^f \mathbf{U} - \mathbf{1}\bar{\mathbf{x}}_d^{f\mathsf{T}})^\mathsf{T} (\mathbf{X}^f \mathbf{U} - \mathbf{1}\bar{\mathbf{x}}_d^{f\mathsf{T}}). \end{aligned} \tag{4.13}$$

$$\begin{aligned} \bar{\mathbf{z}}_d &= \frac{1}{n} \sum_{i=1}^{n} \mathbf{z}_i \mathbf{U}, \\ \mathbf{R}_d &= \frac{1}{n-1} (\mathbf{Z}\mathbf{U} - \mathbf{1}\bar{\mathbf{z}}_d^\mathsf{T})^\mathsf{T} (\mathbf{Z}\mathbf{U} - \mathbf{1}\bar{\mathbf{z}}_d^\mathsf{T}). \end{aligned} \tag{4.14}$$

$$\mathbf{K}_d = \mathbf{P}_d^f (\mathbf{P}_d^f + \mathbf{R}_d)^{-1}. \tag{4.15}$$

$$\begin{aligned} \mathbf{x}_i = \left[\mathbf{x}_i^f \mathbf{U} + \mathbf{K}_d (\mathbf{z}_i \mathbf{U} - \mathbf{x}_i^f \mathbf{U}) \right] \mathbf{U}^\mathsf{T}, \\ i = 1, \ldots, n. \end{aligned} \tag{4.16}$$

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. 4.1 for a radar-based nowcast at lead time +40 min trying to assimilate an NWP EPS forecast that was initialized almost 10 hours before. The data in this example belong to case study ID01, see Sect. 4.5.2. 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 $\mathbf{ZU}$ distribution is more spread than the $\mathbf{X}^f\mathbf{U}$ distribution. Consequently, the posterior $\mathbf{XU}$ distribution results much closer to the prior distributions (notice the low values of K).

4.3.5 Dealing with non-Gaussian precipitation data

Precipitation exhibits a bimodal distribution, being either zero or log-normally distributed. Such non-Gaussianity is problematic to the reduced-space EnKF technique described above in mainly 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: $\text{dBR} = 10\log_{10}\text{R}$. The influence of zeros on the estimation of the covariance matrices is limited by adapting the criterion described in 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, (4.13), (4.14) and (4.15) are applied with a subset of $\mathbf{X}$, $\mathbf{Z}$ and $\mathbf{U}$, while (4.16) uses the whole domain.
In order to adjust the positive 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 in Ebert (2001) for the computation of probability matched ensemble medians. The basic idea is first to build a new empirical posterior distribution

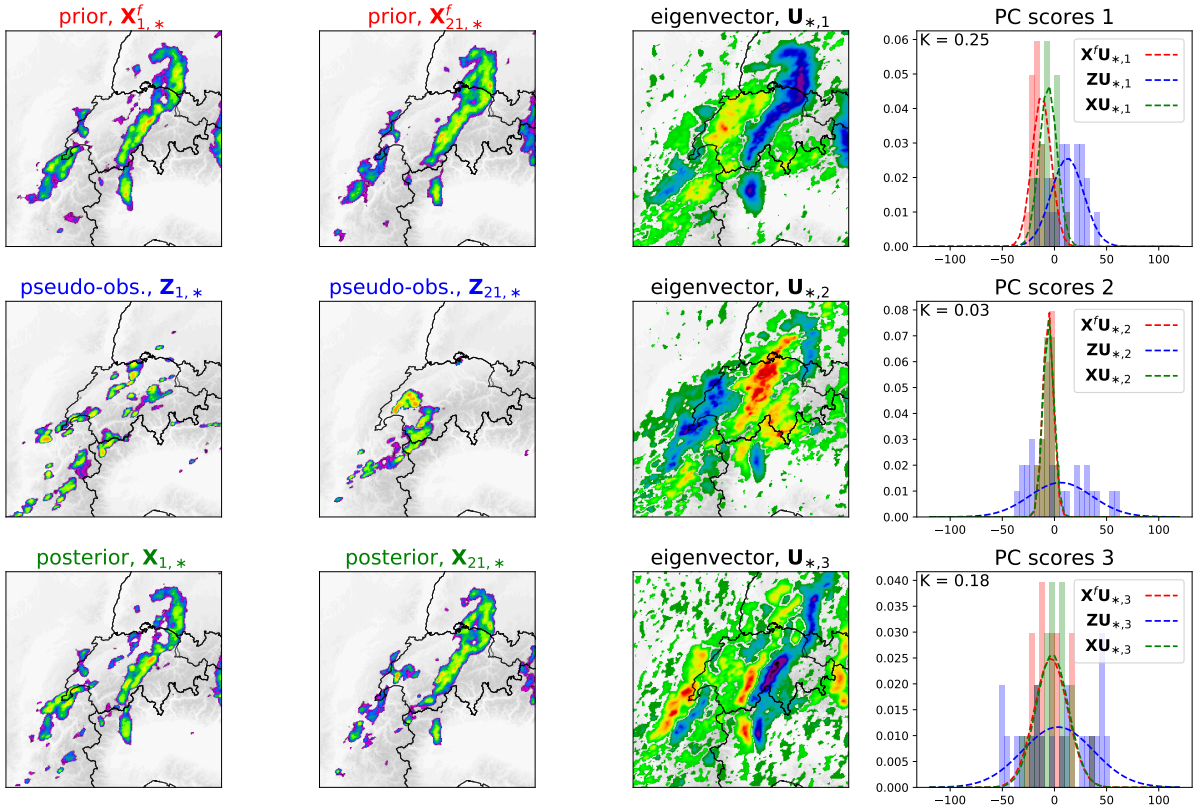


Figure 4.1: Example of reduced-spaced EnKF for one particular time step of a nowcast (+40 min, event ID01). The first two columns include the first and last members of the prior $\mathbf{X}^f$, pseudo-observations $\mathbf{Z}$ and posterior $\mathbf{X}$ ensembles and thus illustrate the effect of the update in the spatial domain. The last two columns illustrate the PCA space that 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 principal components. The distributions of the ensembles for a given PC component are represented by the histograms, while the dashed lines are the best fits for a normal distribution. The transformed posterior distribution ($\mathbf{XU}$, in green) is a weighted average of the transformed prior ($\mathbf{X}^f\mathbf{U}$, red) and transformed pseudo-observations ($\mathbf{ZU}$, blue) distributions according to the Kalman gain K. The event is described in Sect. 4.5.2 and the domain is as in Fig. 4.5.

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 probability p as

$$p = \frac{1}{m} \sum_{j=1}^{m} \frac{\sum_{i=1}^{n} x_{i,j} - x_{i,j}^{f}}{\sum_{i=1}^{n} z_{i,j} - x_{i,j}^{f}}, \tag{4.17}$$

   which represents the average probability of sampling the pseudo-observation $\mathbf{z}_{i,j}$ based on the updated ensemble $\mathbf{X}$. m is the number of grid points in the domain, 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 $r = 1, \ldots, m$, 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 to 4 for the remaining ensemble members $i = 2, \ldots, n$.

4.3.6 Summary of the algorithm

The flow-chart in Fig. 4.2 summarizes the basic steps of the Bayesian nowcasting algorithm. 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. In order 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 the starting point for the next prediction step.

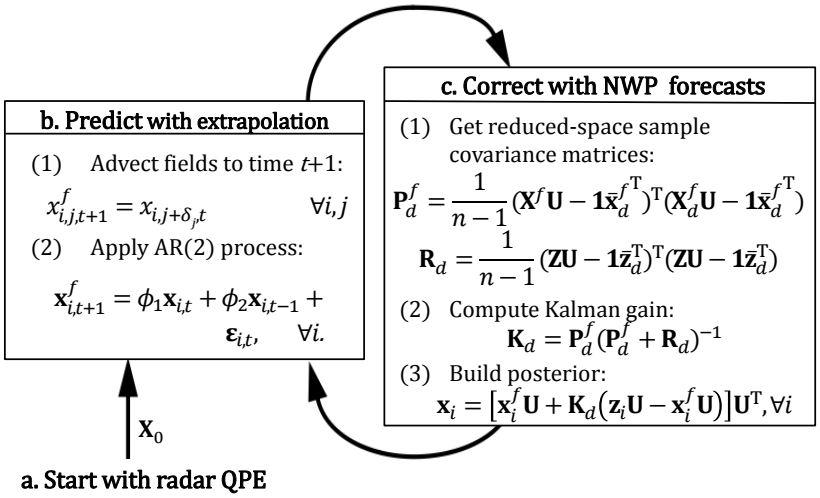


Figure 4.2: The flow chart summarizes the basic equations of the Bayesian nowcasting system. Detailed explanation and notation can be found in the text (Sect 4.3).

4.4 Synthetic data experiments

The Bayesian nowcasting technique is first applied to synthetic data in a series of numerical experiments.

4.4.1 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:

$$\mathbf{x}_{t+1} = \rho_t \mathbf{x}_t + \boldsymbol{\epsilon}_{t+1}, \quad \boldsymbol{\epsilon}_{t+1} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}) \tag{4.18}$$

where $C(u) = \rho_x(u)$ 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 10x10 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. In order to simulate the error of the pseudo-observations $\mathbf{Z}$, a bias term is introduced by setting $\mu > 0$. 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).

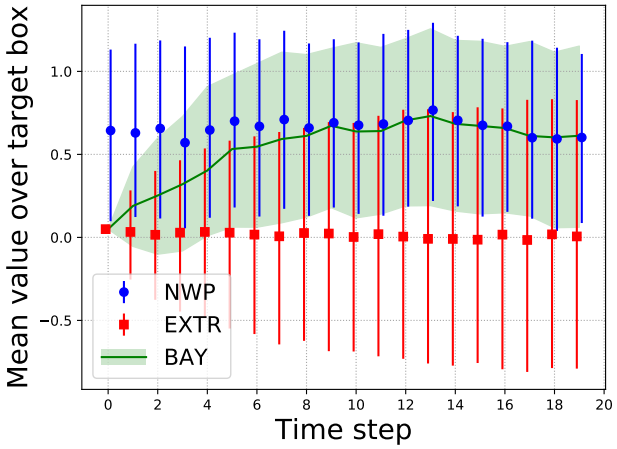


Figure 4.3: Results of 100 synthetic simulations to test the Bayesian nowcasting based on reduced-space EnKF. The mean value over a 5x5 pixel target box is shown as a function of time step for a stochastic persistence forecast (red), synthetic NWP ensemble (blue) and Bayesian nowcasting (green). The graph represents at each lead time the median and the 10-90% interval. All simulations used 100 member ensembles, $\rho(1) = 0.95$ and a NWP spread of 60%.

4.4.2 Simulation results

Figure 4.3 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 in based on the posterior of time $t-1$. The mean value over a target box of 5x5 pixels centered at the simulation grid is shown as a function of the time step. The ensemble medians for persistence (EXTR), pseudo-observations (NWP) and posterior (BAY) are represented as well as their ensemble spreads (interval between 10^{th} and 90^{th} 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 ten time steps, it closely matches it.

Figure 4.4 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. 4.4a), 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,

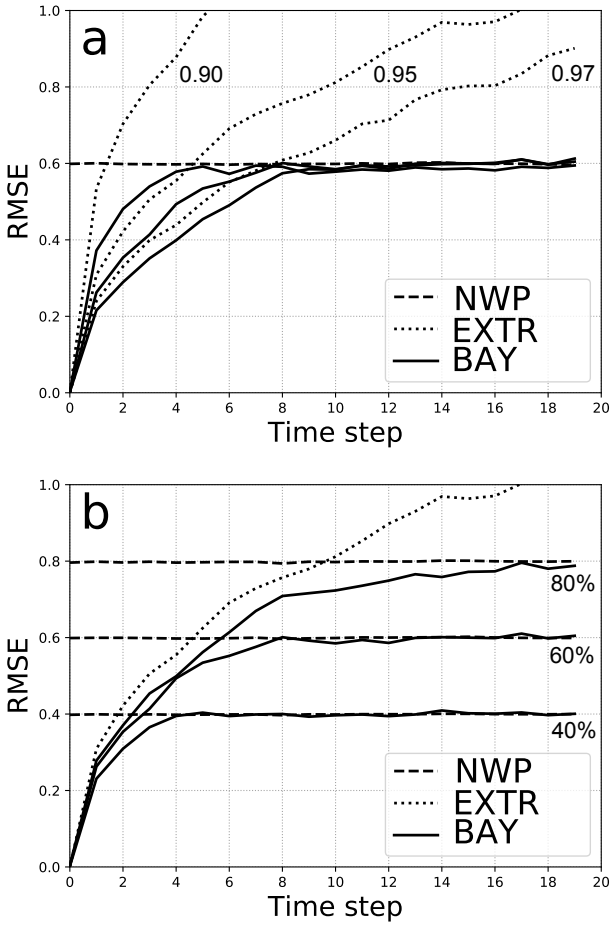


Figure 4.4: Results from two numerical experiments with the Bayesian nowcast based on reduced-space EnKF. (a) The skill of the Bayesian nowcast ensemble mean (solid line) is verified for three levels of temporal persistence of the simulated true model state (dotted line, left to right: $\rho_t = 0.90, 0.95, 0.97$) 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, $\rho_t = 0.95$.

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. 4.4b), 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 once again able to converge to the skill of the pseudo-observations at a rate that is consistent with its uncertainty.

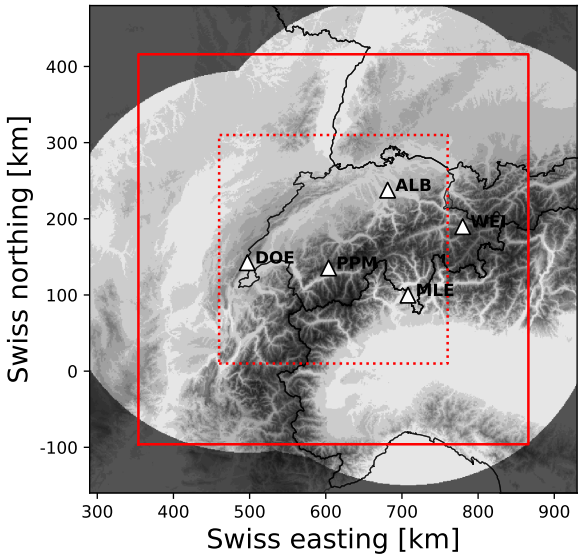


Figure 4.5: The 512x512 km study area (solid red line) and the 300x300 km verification domain (dashed red line) are centered over Switzerland. The five radars are indicated with white triangles. The composite radar mask, the topography and national borders are also included as reference.

4.5 Real data experiments

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

4.5.1 Data sets and study domain

The weather radar observation data set was produced by the 4^{th} generation of the MeteoSwiss radar network. The radar quantitative precipitation estimate (QPE) is a composite image derived from measurements of five dual-polarization, C-band weather radars distributed across the Swiss territory (white triangles on Fig. 4.5). 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 4^{th} generation Swiss radar network is presented in 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 Lichtenstein, and the neighboring regions of Germany, France, Italy, and Austria. In order 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 512x512 km region centered over Switzerland (red solid box in Fig. 4.5).

A control experiment is produced through stochastic radar extrapolation. This ensemble forecast is based solely on the stochastic Lagrangian persistence approach described in Sect. 24.3.2. 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 is 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 control). The COSMO-E model provides twice a day (initialization times at 00 and 12 UTC), medium-range forecasts (up to 120h) for the Alpine region at convection-permitting resolution (2.2 km mesh-size). The KENDA assimilation system (Schraff et al., 2016) produces the ensemble of initial conditions, based on a local ensemble transform Kalman filter (LETKF, Hunt et al., 2007), while the lateral boundary conditions are provided by the global ECMWF EPS 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 parametrizations is included in COSMO-E with stochastically perturbed physics tendency (SPPT, Buizza et al., 1999).

4.5.2 Case studies

Four precipitation events between 2016 and 2017 were selected. They are listed in Fig. 4.6, where their total accumulated precipitation is represented.
The first precipitation event ID01 (Fig. 4.6a) is linked to a frontal system causing the convective activity to be organized in lines with a SW-NE orientation, while the general flow is predominately from West, South-West.
The second event ID02 (Fig. 4.6b) concerns a winter precipitation case that is produced by a strong northwesterly flow bringing moist air towards 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. 4.6c) 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. 4.6d) is again a summertime event, but this time the flow is from 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.

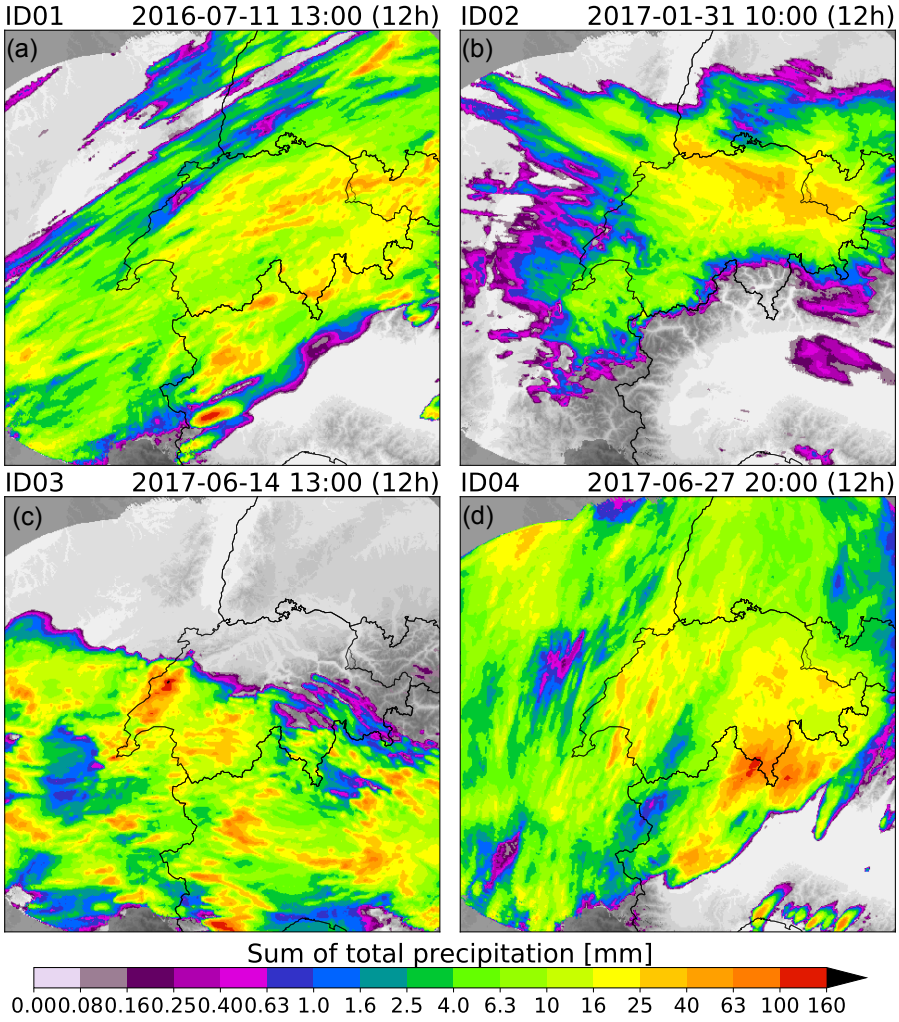


Figure 4.6: Maps of total accumulated precipitation corresponding to the events considered in Section 4.5. The timestamp indicates the start of the event, while the event duration is given within brackets.

4.5.3 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:

$$\mathrm{FSS} = 1 - \frac{\frac{1}{M}\sum_{M}(P_F - P_O)^2}{\frac{1}{M}\sum_{M}P_F^2 + \frac{1}{M}\sum_{M}P_O^2}, \tag{4.19}$$

where P_F and P_O are the forecasted and observed fractions of grid points with intensity higher 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. In order to generalize the FSS to the ensemble skill, we applied the same approach described in Zacharov and Rezacova (2009), which consists in computing the FSS for each ensemble member and then average it to obtain the mean ensemble skill.

4.5.4 Bayesian precipitation nowcast examples

Figure 4.7 shows an example of nowcast started at 2100 UTC 11 July 2016 (ID01). Each row represents one selected lead time (+00, +60, +120 or +180 minutes). 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 SW-NE orientation is visible in the radar images in the first column. In the second column, COSMO-E was initialized at 1200 UTC, that is, 9 hours 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.

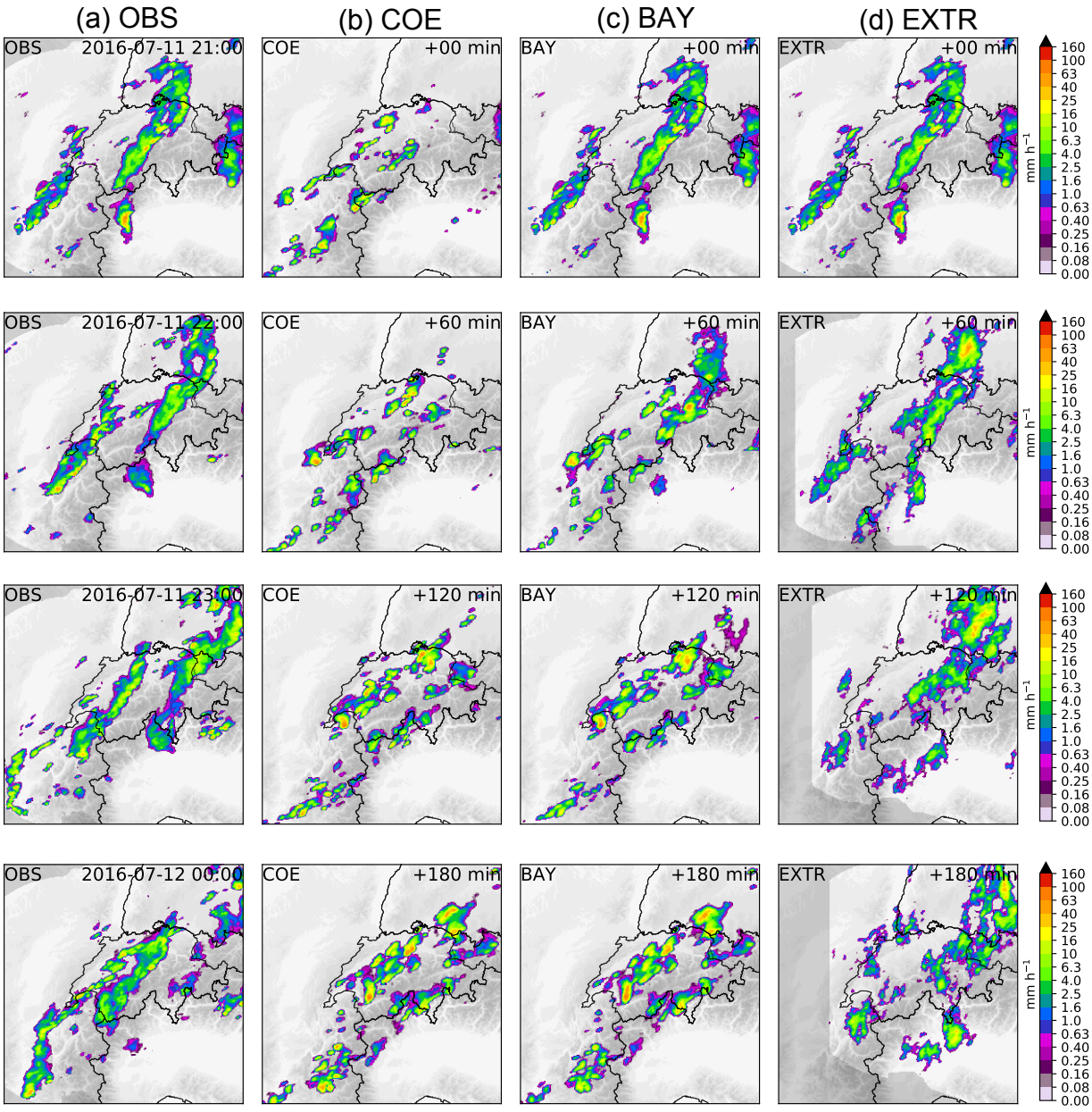


Figure 4.7: An example for a forecast started at 2100 UTC 11 July 2016 (ID01) with (a) verifying radar QPE at 2100 UTC, 2200 UTC, 2300 UTC and 0000 UTC; (b) the COSMO-E control run, (c) one realization (member) of the Bayesian nowcast and (d) one member of the stochastic radar extrapolation ensemble at +00, +60, +120 and +180 minutes 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 precipitation accumulations.

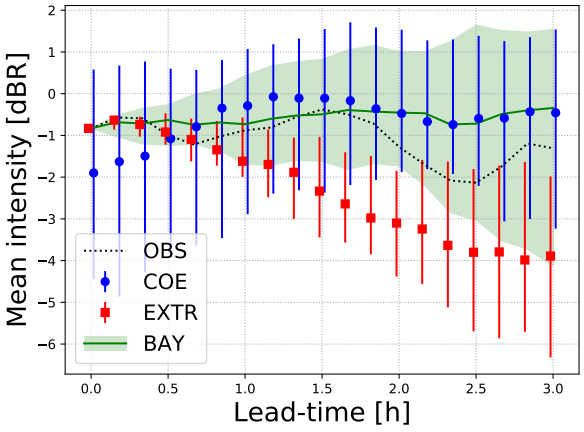


Figure 4.8: Mean areal precipitation intensity in dBR over the 300x300 km verification box as a function of forecast lead time for the nowcast started at 2100 UTC 11 July 2016 (as in Fig. 4.7, ID01). The median and the 10-90% prediction interval are visualized.

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. 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 to illustrate the influence of the domain boundaries.

A more quantitative illustration of the Bayesian nowcasting technique is given in Fig. 4.8 where the mean precipitation intensity over the 300x300 km verification box (Fig. 4.5) is shown as a function of forecast lead time. As in Fig. 4.3, 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 notice that not only the mean precipitation intensity seamlessly merges from observations into the NWP forecast, but also that the variability within the ensemble grows consistently with its transition from the observations to the numerical model.

The second example of nowcast is presented in Fig. 4.9. This time the nowcast is started at 1000 UTC 31 January 2017 (ID02), and it concerns the winter precipitation event with pronounced orographic precipitation, as it can be seen in the radar images. The COSMO-E run was initialized at 0000 UTC, that is, 10 hours before the start of the nowcast. While the general structure of the precipitation event seems well captured,

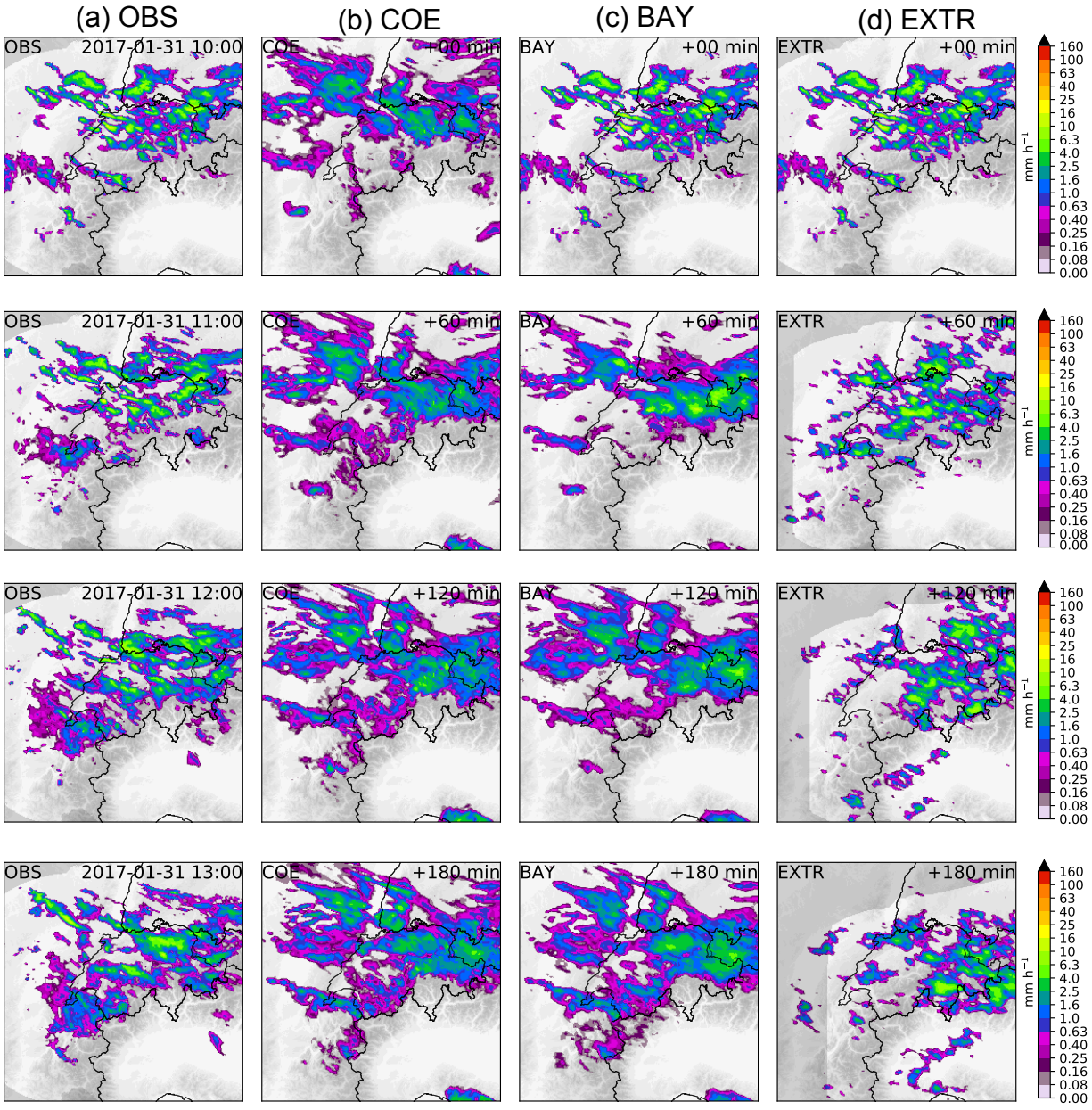


Figure 4.9: Same as in Fig. 4.7 but for the nowcast started at 1100 UTC 31 January 2017 (ID02). The COSMO-E forecast was initialized at 0000 UTC of the same day.

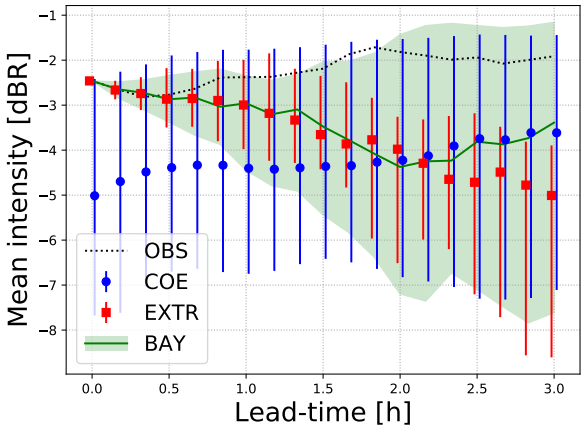


Figure 4.10: Same as in Fig. 4.8, but for the nowcast started at 1100 UTC 31 January 2017 (ID02).

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 precipitation can be 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 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 minutes 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 minutes 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. 4.10, where the mean precipitation intensity over the 300x300 km verification box (Fig. 4.5) is shown as a function of forecast lead time. It is now clearer from this chart that the Bayesian nowcast starts precisely from the last radar observation and then moves towards the COSMO-E forecast. After three hours, the Bayesian nowcast is very close to COSMO-E. Another interesting aspect to notice 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 match the COSMO-E ensemble spread closely. In this sense, the Bayesian nowcasting can be considered seamless in terms of both mean and spread.

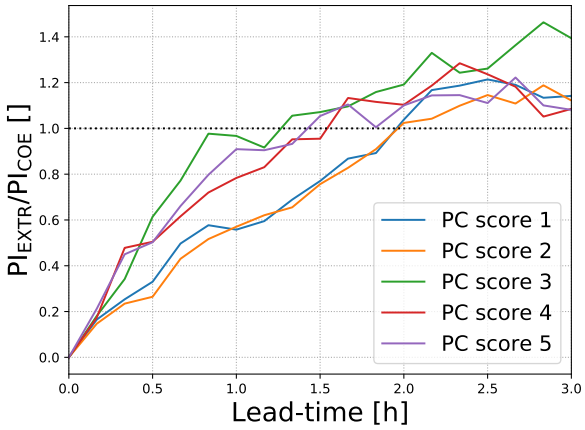


Figure 4.11: The spread of the extrapolation nowcast (PI_{EXTR}) relative to the COSMO-E spread (PI_{COE}) in terms of the first five PC scores. The results are averaged over all the case studies. The spread PI is defined as the 10-90% prediction interval.

It is important to remember that the Bayesian update is based on the distribution of the PC scores (cfr. Fig 4.1). The spread in terms of PC scores is illustrated in Fig 4.11, where the average ratio between the extrapolation and NWP prediction intervals is shown for the first five PC scores and as a 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 hours.

4.5.5 Verification results

For each event, a 4-hour Bayesian nowcast was started every hour. Given that all events last 12 hours, the number of runs per event is equal to 9. Therefore, there are in total 4 events x 9 runs x 4 hours = 144 hours that are included in this verification. If one further considers the time resolution of 10 minutes (hence 6 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 hours x 6 fields hour^{-1} x 21 members x 3 procedures = 54432 precipitation fields that are verified.
The verification of the four precipitation events (ID01-04) is presented in Fig. 4.12, 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 comparative analysis of precipitation forecasts (e.g. Simonin et al., 2017). Scores are computed within the verification box as defined in Fig. 4.5. For each of the three ensemble forecasts, its 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. 4.7 and 4.8). 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 with time. The crossover time is on average between +2 and +2.5 h, and the Bayesian nowcasting system well captures it. 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. 4.9 and 4.10). In Fig. 4.12b 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 hours. In this case, the Bayesian approach is correctly weighting the contribution of the numerical model to produce an optimally combined forecast.

The intense and localized precipitation of event ID03 translates into the low skill of COSMO-E as it can be seen in Fig. 4.12c. 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 four hour 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 four hours. 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 hour lead time, an indication that the skill of COSMO-E is somewhat overestimated.

Concerning the last precipitation case study (ID04) in Fig. 4.12d, it can be seen that the large-scale organization of this event leads to a slower decay of the skill of 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. 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 hours. The Bayesian nowcasting has 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 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. 4.4). 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.

4.6 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 hours). 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 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 several adjustments to relax some of its underlying assumptions and efficiently work with large data sets. 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 two first 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)

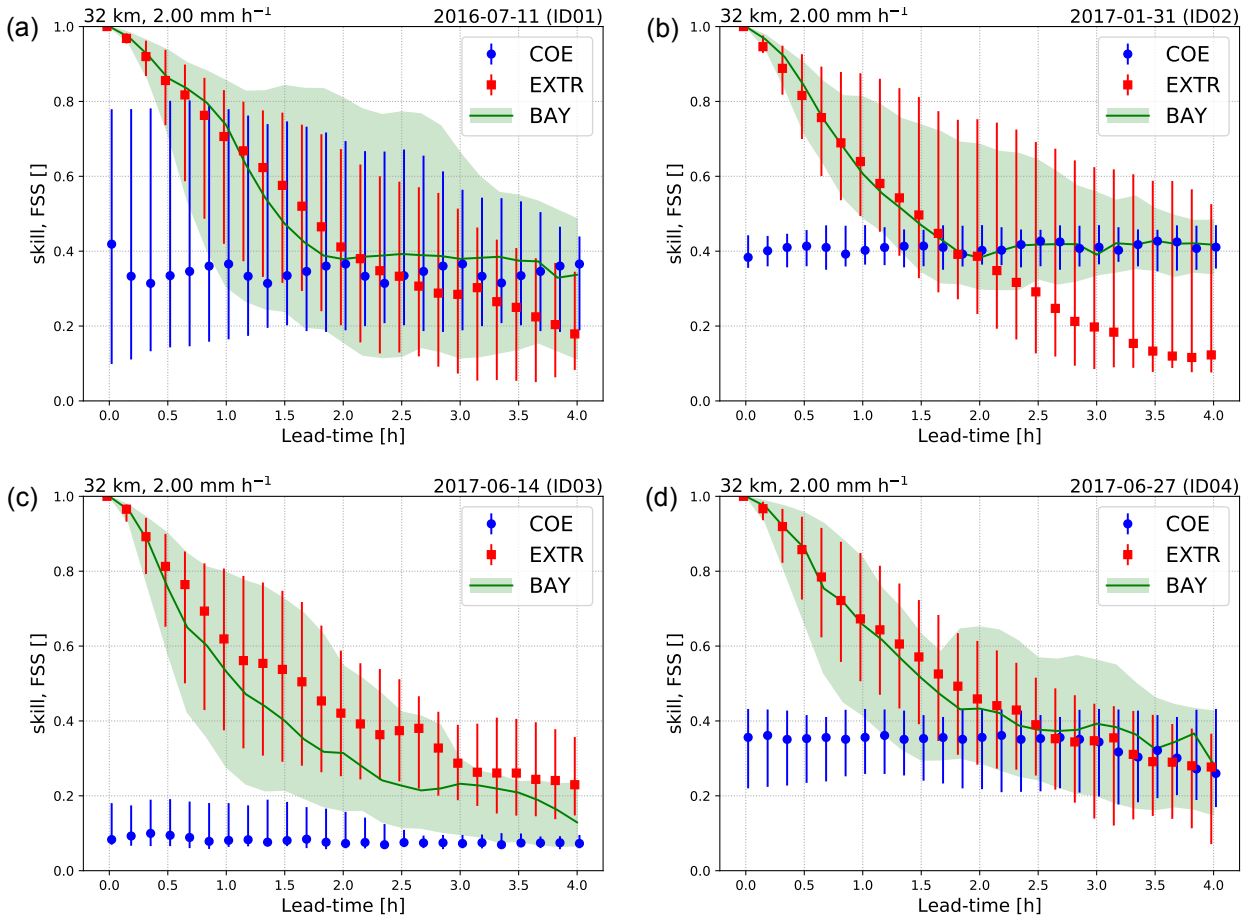


Figure 4.12: Mean ensemble Fractions skill scores (FSS) as a function of lead time for a 32 km scale, threshold 2.0 mm h^{-1} and computed within the verification box (see Fig. 4.5) for all four events (ID01-04). The mean ensemble FSS is the average of the individual FSS scores of each ensemble member. The bars represent the variability among the individual runs within a single event. The three ensembles are COSMO-E in blue (COE), the Bayesian nowcasting in green (BAY) and the stochastic radar extrapolation in red (EXTR).

yielding consistent results. The results indicate that the system can cope with realistic conditions to produce precipitation forecasts that perform as good as extrapolation nowcasts for short lead times and then converge to the skill of NWP at a plausible rate. The crossover lead time varies from 1.5 to more than 4 hours.

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 not only a function of lead time but also the 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 second on a 2.6 GHz Core, while one prediction step demands approximately 6 seconds. In total, a 4-hour Bayesian nowcast takes about 2 minutes of computations. The only parallelization achieved in this study is the one currently integrated into 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 inter-dependencies, can be foreseen. For example, future research could consider the joint integration of multiple fields (e.g., precipitation, temperature and wind) using a combined PCA approach (CPCA, Wilks, 2011a). In this sense, the Bayesian nowcasting approach can open the door to a truly integrated, multi-variable, seamless, probabilistic nowcasting system.

One additional aspect calling for future research is the Gaussian assumption of the EnKF in the presence of non-Gaussian quantities such as precipitation. The non-Gaussian nature of precipitation questions the use of an EnKF approach. In principle, a non-parametric 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 large biases in the wet area and maximum intensities resulting from the EnKF analysis. However, recent developments in data assimilation using local particle filters (LPF, e.g., Penny and Miyoshi, 2016; Poterjoy, 2016) and hybrid local ensemble Kalman particle filters (Robert et al., 2017) are showing promising results and hence could also represent valuable developments for seamless forecasting systems.

Chapter 5

Conclusions and outlook

The work presented in this thesis provided several contributions to the precipitation nowcasting science. The main findings can be summarized as follows.

In Chapter 2, we highlight the relevance of a scale decomposition framework for the spatial verification of high-resolution precipitation forecasts. The Fourier transform provides a computationally-efficient way to decompose the forecast error field, while, thanks to the Parseval's theorem, we can interpret its power spectrum in terms of fractional contribution to the total error. The scale-aware formulation provides an easily interpretable measure of skill which can deal with the complex structure of high-resolution forecast fields. For the Alpine region, our analyses show that the scale-aware predictability of precipitation patterns is at maximum 2 h for horizontal scales in the order of tens of kilometers, while it can span beyond 6 h for scales of hundreds of kilometers. Radar-based nowcasting is essential for applications requiring predictive skill for precipitation scales below 60 km, as currently this cannot be provided by state-of-the-art numerical models. In terms of scale-aware predictability, the extrapolation of radar images is superior to the numerical forecasts from the convection-permitting model COSMO-1 up to 3 h, although this crossover lead time is shortened during warm convective events, arguably because of the lower persistence of precipitation patterns and the more effective assimilation of locally triggered air mass convection into the numerical model. After 4.5 h, precipitation on scales below 150 km is poorly predictable by all forecasting means.

In Chapter 3, we demonstrate the importance of considering the heterogeneous distribution of the statistical properties of precipitation. The short-space approach provides an effective way to account for spatially-varying model parameters in a moving window fashion. A localized version of the radar-based Short-Term Ensemble Prediction System (STEPS) model is verified against a set of ten predominantly convective pre-

cipitation events within the Swiss radar domain (710x640 km). Localized simulations generally provide more reliable and more realistic probabilistic nowcasts, especially during the first couple of hours into the forecast. Increasing the degree of localization reduces conditional biases, and improves the reliability and resolution in the ensemble of simulations. However, it also implies a stronger assumption in the persistence of the statistical properties of precipitation, which can lead to less dispersive ensemble forecasts. For instance, using localization at the scale of 100 km reduces underdispersion by around 30% at +30 min, but it increases it by 30% at +180 min. The computational cost of the localization procedure is still affordable for nowcasting applications, provided that CPU parallel computing is enabled.

In Chapter 4, we show the feasibility of adopting a Bayesian view for the seamless blending of radar and NWP forecasts. The spread of the ensembles provides a measure of the forecast uncertainty that responds to changing meteorological situations and includes information on the scale and spatial dependence of the errors. The recursive implementation of the Bayesian update equations using the ensemble Kalman filter represents an attractive framework for the integration of uncertain data. The combination with principal component analysis provides a means to efficiently extract the information contained in the high-resolution ensemble forecasts by projecting the data into a space of lower dimensionality. Without the need for a training period, the blending weights are computed in real time based on the relative uncertainty between radar and model forecasts. However, a post-processing step is necessary in order to correct the higher-order statistics of the posterior rainfall distribution, especially in the presence of displacement errors in the numerical model. During the four precipitation events considered in the analysis, the probabilistic radar nowcasts generated with the STEPS model blended into the ensemble forecasts of the convection-permitting model COSMO-E in between 1.5 and 4 hours or more, providing blended forecasts that were at least as skillful as the nowcast-only or the NWP-only forecasts at any lead time.

5.1 Outlook

A detailed discussion of the limitations and possible improvements of the analyses in this thesis is available at the end of the relevant chapters. In a more general sense, here we want to discuss the broad research directions that could be interesting to explore once considered the work presented here in the wider context of the recent scientific and technical developments.

First, it exists a general trend towards data-driven models based on large archives of data. Statistical methods are getting much attention thanks to the successes and technical improvements of machine learning methods (e.g., Schmidhuber, 2015, for a review on deep learning). Related research work has been conducted in the broader context of this project (Appendix A.1), which has also contributed to the development of operational nowcasting tools (Sideris et al., 2018). However, there is still considerable work to be done. One critical aspect calling for further research relates to the difference between the estimation and simulation approaches. In the case of *estimation*, which represents the standard framework in machine learning, a deterministic model is trained to minimize the total prediction error, often defined as the sum of the squared errors. The error minimization framework inevitably leads to smoothing and conditional biases (e.g., Frei and Isotta, 2019; Foresti et al., 2019). Conversely, the *simulation* approach takes a stochastic stance and aims at providing realizations of the future precipitation state PDF. Simulation methods in machine learning are usually known as **generative models**. Over the past few years, the topic has become very popular with the appearance of so-called *GAN*s (Generative Adversarial Networks, Goodfellow et al., 2014), which have already shown to bear great potential for weather-related applications (e.g., Gagne et al., 2018).

Second, the use as forcing data for **hydrological forecasting** represents a privileged application of precipitation nowcasting methods (e.g., Mecklenburg et al., 2000a; Berenguer et al., 2005; Werner and Cranston, 2009; Liguori et al., 2012; Liechti et al., 2013; Silvestro et al., 2017; Antonetti et al., 2019). In the context of this thesis, some preliminary research work has been possible thanks to the work of two Master students. A summary is provided in Appendix A.2. The coupling of the meteorological and hydrological forecasting chains is, however, complex and more work is necessary in order to improve our understanding of the individual contributions to the total forecast uncertainty. Such effort is relevant and needed given that over the past few years a great interest in the use of impact-based warning and forecasting systems has emerged.

Finally, the work of this thesis aligned very well with the emergence over the past few years of **open-source software** in the atmospheric sciences (e.g., Heistermann et al., 2015; Raspaud et al., 2018). The nowcasting library pysteps (Pulkkinen et al., 2019b, see Appendix A.3) represents the computational tool that, in a very pragmatic sense, would have been needed at the very start of this project. It is also by thinking to the future generations of Ph.D. students that pysteps was designed and developed. Pysteps has facilitated research and enabled stimulating collaborations. The use of

open source has shown to be worth the investment and has rapidly become an integral part of our research activity. It is therefore foreseen that more work will be dedicated to the maintenance and further development of pysteps and other open-source tools, while the whole community is warmly welcomed to join the effort towards a more open nowcasting science.

Appendix A

Related research

A.1 The precipitation attractor

The quest for a precipitation attractor within the MeteoSwiss archive of weather radar data has been a central task of the wider research project to which this thesis belonged (Section 1.7).
First, the concept of precipitation attractor was applied in the analysis of predictable patterns of precipitation growth and decay in the historic radar data. The findings are summarized in a climatological study that was recently published.

- Foresti, L., I. 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, *Q.J.R. Meteorol. Soc.*, **144 (716)**, 2277–2301.

Second, we investigated the practical implications of the precipitation attractor using machine learning techniques to predict growth and decay depending on flow direction, freezing level height, and geographical location. We found that machine learning techniques provide a convenient solution to the problem of finding good analogs in a highly dimensional space by performing a non-linear interpolation of analog states within the finite-size dataset at our disposal. This work is presented in the following publication.

- Foresti, L., I. Sideris, D. Nerini, L. Beusch, and U. Germann, 2019: Using a 10-year radar archive for nowcasting precipitation growth and decay – a probabilistic machine learning approach. *Weather Forecast.*, **34 (5)**, 1547–1569.

In parallel, different ways to construct and study a low-dimensional phase space were explored using radar data from Switzerland and the US. We investigated the use of both summary statistics and principal component analysis, which resulted in the definition of two types of empirical attractors for precipitation. An example of precipi-

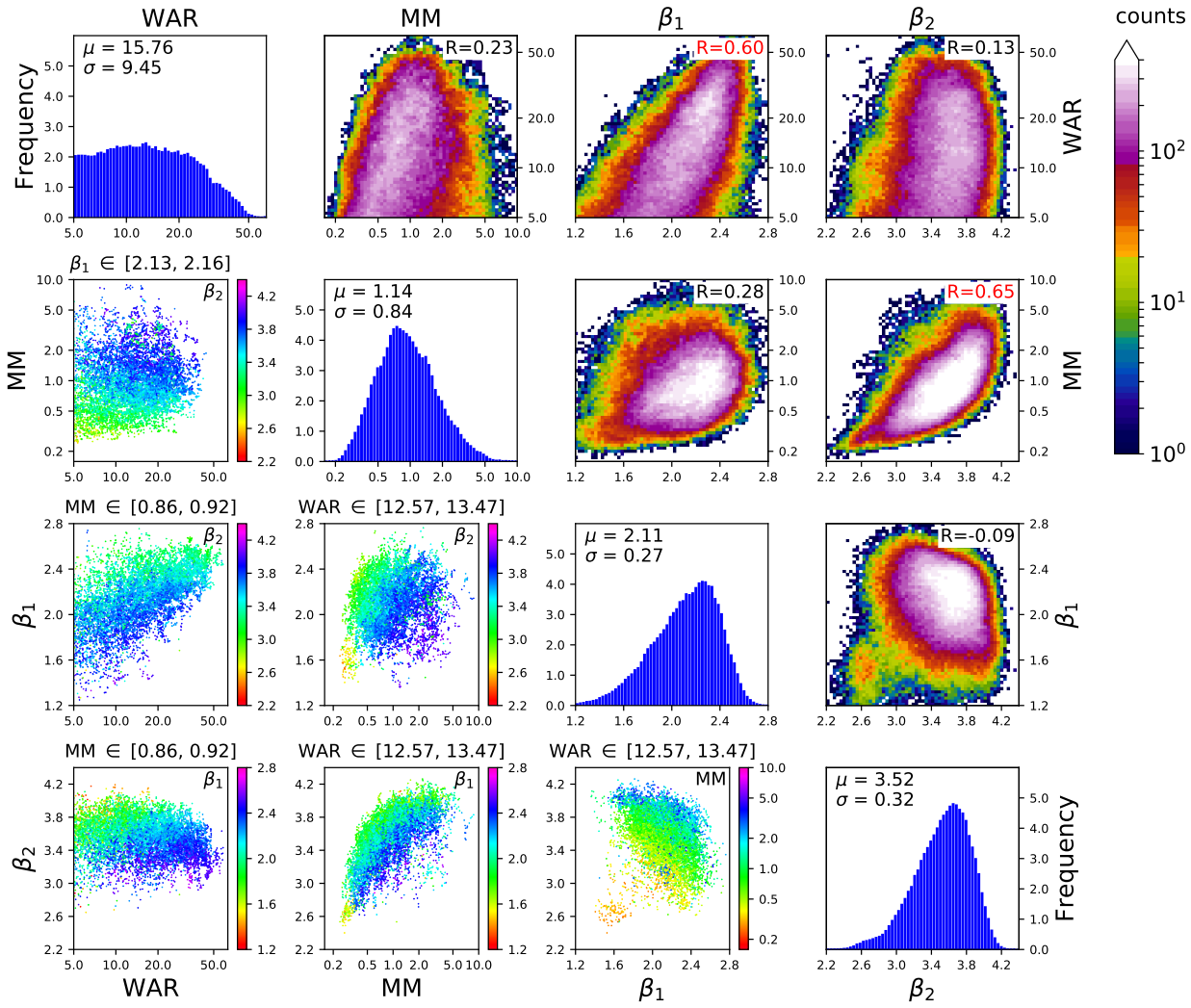


Figure A.1: Swiss 4D precipitation attractor using as phase space dimensions the marginal mean precipitation (MM), the wet area ratio (WAR), and the spectral slopes of the radially averaged Fourier power spectra (β_1 and β_2). MM and WAR are shown in log scale to account for the asymmetry of the distribution.

tation attractor for the Swiss data using summary statistics is illustrated in Fig. A.1, while a publication on the subject is currently in preparation.

- Foresti, L., B. Puigdomènech Treserras, D. Nerini, M. Gabella, I.V. Sideris, A. Atencia, U. Germann, I. Zawadzki, 2019: A precipitation attractor based on weather radar observations: predictability, analogues and principal component analysis, **in preparation**.

A.2 Hydrological applications

With respect to hydrological applications, some collaborative research has been possible thanks to the hard work of two Master's students, who dedicated their final MSc projects to explore the use of precipitation nowcasting for hydrological simulations.

- Herrmann, Vincenz, 2019: Exploring hydro-meteorological uncertainties in hydrological flow forecasting. Master thesis. ETH Zurich, Zurich, Switzerland.

The first Master thesis by Vincenz Herrmann was co-supervised in collaboration with Prof. Dr. Peter Molnar and Dr. Nadav Peleg from ETH Zurich. The objective was to quantify the uncertainties of radar-based probabilistic nowcasting and explore how these uncertainties are propagated through a semi-distributed rainfall-runoff model.
To this end, a radar-based nowcasting model (pysteps, Pulkkinen et al., 2019b) was employed to generate ensembles of precipitation nowcasts for 12 manually selected convective precipitation events in a flash-flood-prone catchment in the Swiss Prealps. The precipitation forecasts were used to force the rainfall-runoff model KINEROS2, an event-oriented, semi-distributed, physically based hydrological model.
The student demonstrated the feasibility and potential of precipitation nowcasting for hydrological forecasts. Of particular interest was the finding that the hydrological model reacts sensitively to the uncertainties in the precipitation forcing. For instance, the overestimation by the nowcasts in case of low-intensity precipitation events was found to be amplified by the hydrological model.

- Koller, Severin, 2019: Performance analysis of different flash flood nowcasting chains based on ensemble precipitation nowcasting in the Emme and Verzasca catchments. Master thesis. University of Zurich, Zurich, Switzerland.

The second Master thesis by Severin Koller was co-supervised in collaboration with Prof. Dr. Jan Seibert from the University of Zurich and Dr. Massimiliano Zappa from the Swiss Federal Institute for Forest, Snow and Landscape Research WSL. The objective was to quantify the impact of using precipitation nowcasting techniques to trigger and force a fully-distributed, event-based flash flood forecasting system.
In hindcast mode, the rainfall accumulations over a given target basin were monitored by an automatic alert system operated by MeteoSwiss (NowPAL, Panziera et al., 2016). The exceedance of a predetermined threshold triggered the hydrological forecast by a novel process-oriented and fully-distributed rainfall-runoff (RGM-PRO, Antonetti et al., 2019). Hydrological simulations were forced with ensemble precipitation nowcasts from a seamless forecasting system that optimally combines the information and respective uncertainty from radar and NWP-based precipitation nowcasts (Nerini et al., 2019).

By using different combinations of the precipitation forcing data and a broad set of skill metrics, the student provided a robust analysis of the impact of precipitation nowcasting methods on the hydrological forecasts. It was found that including radar-nowcasting with blending to the numerical model improves the hydrological forecast skill compared to the use of numerical forecasts only and that the frequent update of the initial conditions is key to the success of the flood nowcasting chain.

A.3 The pysteps open-source initiative

Similarly to other research fields, the nowcasting community has invested a significant amount of time to re-implement from scratch routines and algorithms that have been around for decades, for example, optical flow and advection schemes.
Initially, this thesis did not represent an exception.
Part of the problem can be associated with the unavailability of software, which is often proprietary or too poorly documented to be understood, trusted, and used. As a consequence, the lack of common standards, transparency, code availability, and well-documented workflows often delivers non-reproducible results in many computational studies, which ultimately questions their scientific value (Hutton et al., 2016).
Recognizing that nowcasting methods and related applications can be further developed and distributed by promoting universal access to existing knowledge, a Python-based software package, called *pysteps*, is being developed as a community-driven effort. This effort fits well into the weather radar community with the emergence of open data and an increasing number of open-source software projects (Heistermann et al., 2015), for instance, in radar data processing (Heistermann et al., 2013; Helmus and Collis, 2016, see also `https://openradarscience.org/`).
The objective of pysteps is twofold. First, it aims at providing a well-documented and modular framework for the development of new nowcasting methods, thus promoting the adoption of open-science practices. Second, pysteps aims at providing an easily accessible software package for practitioners ranging from weather forecasters to hydrologists.
The pysteps library is freely available at `https://github.com/pySTEPS/pysteps/`, while a thorough model description is provided in a recent publication.

- Pulkkinen, S., D. Nerini, A. Perez Hortal, C. Velasco-Forero, U. Germann, A. Seed, and L. Foresti, 2019: Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1.0). *Geosci. Model Dev.*, **12 (10)**, 4185–4219.

Appendix B

Some mathematical tools

The content of this Appendix has been adapted from the manuscripts:

- 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 (6)**, 2777 – 2797.
- Pulkkinen, S., D. Nerini, A. Perez Hortal, C. Velasco-Forero, U. Germann, A. Seed, and L. Foresti, 2019: Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1.0). *Geosci. Model Dev.*, **12 (10)**, 4185–4219.
- Foresti, L., B. Puigdomènech Treserras, D. Nerini, M. Gabella, I.V. Sideris, A. Atencia, U. Germann, I. Zawadzki, 2019: A precipitation attractor based on weather radar observations: predictability, analogues and principal component analysis, **in preparation**.

B.1 Fourier Transform

The corner stone of signal processing theory (Stankovic et al., 2013) is the idea that a signal $x(t)$ can be represented as an infinite weighted sum of harmonic sine waves of infinitesimally close frequencies f:

$$x(t) = \int_{-\infty}^{+\infty} X(f) e^{j2\pi f t} df, \tag{B.1}$$

where $X(f)$ represents the complex Fourier frequency representation. It contains both the amplitude ($|X(f)|$) and the phase (ϕ) information ($X(f) = |X(f)|e^{j\phi}$). Obtained by means of the Fourier transform (FT), it is actually the same signal, only represented in the frequency domain:

$$X(f) = \mathcal{F}\{x(t)\} = \int_{-\infty}^{+\infty} x(t) e^{-j2\pi f t} dt, \tag{B.2}$$

where $\mathcal{F}\{x(t)\}$ denotes the Fourier transform of the signal $x(t)$. Therefore, the Fourier transform could be essentially interpreted as the transformation of a signal from a given space to the space of frequencies, where the energy/average power of a signal is conserved (Parseval's theorem). Signal $x(t)$ and its Fourier counterpart $X(f)$ are often called *FT pair*.

If the signal $x(t)$ is deterministic (or transient stochastic), its energy spectral density can be defined as $S(f) = |X(f)|^2$, while in the case of a non-transient stochastic signal we ought to define its power spectral density (commonly referred to as *power spectrum*), being $S(f) = \lim_{T\to\infty} \frac{|X(f)|^2}{2T}$.

That leads us to the next step, which is the expansion of the Fourier transform to the analysis of stochastic spatial signals stochastic fields. The expanded, two-dimensional Fourier transform (2D FT) would be, as a matter of fact, the corner stone of image processing theory (Stankovic et al., 2012). Namely, using 2D FT a field $f(x, y)$ can be transformed from a two-dimensional space to the two-dimensional frequency domain:

$$F(f_1, f_2) = \mathcal{F}\{f(x, y)\} = \iint_{-\infty}^{+\infty} f(x, y) e^{-j2\pi(f_1 x + f_2 y)} dx dy, \tag{B.3}$$

with $f_1 = 1/x$ and $f_2 = 1/y$ being the spatial frequencies in x and y directions. Analogously to the one-dimensional signal, which can be represented as a weighted sum of one-dimensional sinusoids, a field is represented by a weighted sum of sinusoids that vary in the x-y plane:

$$f(x, y) = \mathcal{F}^{-1}\{F(f_+, f_2)\} = \iint_{-\infty}^{+\infty} F(f_1, f_2) e^{j2\pi(f_1 x + f_2 y)} df_1 df_2, \tag{B.4}$$

with the complex Fourier representation $F(f_1, f_2) = |F(f_1, f_2)|e^{j\phi}$. The power spectral density of a stochastic field is estimated on $\Delta x \times \Delta y$ portion of a plane as $S(f_1, f_2) = \lim_{\Delta x \Delta y \to \infty} \frac{|F(f_1, f_2)|^2}{\Delta x \Delta y}$.

Aside from the complete two-dimensional information about field power spectral density ($S(f_1, f_2)$), it is common to use the one-dimensional, simplified representation ($S_1D(f)$), obtained by radially averaging $S(f_1, f_2)$ around the $S(0, 0)$ point. This allows us to analyze the scaling behavior of rainfall, by characterizing the power spectral curve with the spectral exponent , obtained by its power-law approximation in the log-log plot $S^a_{1D}(f) = 1/f^\beta$.

One of the most noteworthy properties of the Fourier transform is a significantly facilitated convolution in the spatial (or temporal) domain. Namely, the convolution of two fields $f_1(x, y)$ and $f_2(x, y)$ forms the FT pair with the product of their respective Fourier transforms $F_1(x, y)$ and $F_2(x, y)$:

$$\mathcal{F}\{f_1(x, y) * f_2(x, y)\} = F_1(f_1, f_2) F_2(f_1, f_2). \tag{B.5}$$

B.2 Discrete Fourier Transform

In analyzing real, observed fields, we actually never deal with the continuous functions (e.g. $r(x, y)$) but rather with the discrete ones. That would mean that a two dimensional field can be represented by a matrix of finite size ($\mathbf{r}_{MN}$), with the same being true for its Fourier pair $\mathbf{R}_{MN}$. In this case, the transformation from Eq. B.3 takes the form of a Discrete Fourier Transform (DFT):

$$\mathbf{R}(m, n) = \frac{1}{\sqrt{MN}} \sum_{l=0}^{N-1} \sum_{k=0}^{M-1} \mathbf{r}(k, l) e^{-j2\pi(km/M + ln/N)}, \tag{B.6}$$

which, particularly if M and N are powers of 2 or have only small prime factors, can be efficiently implemented using Fast Fourier Transform (FFT) algorithms (Cooley and Tukey, 1965). The transformation preserves all the properties presented in the preceding section, including the facilitated convolution.

B.3 Wiener-Khintchine theorem

Both the energy spectral density and the power spectral density form the Fourier transformation pair with the autocorrelation function of a corresponding signal:

$$R(\tau) = \int_{-\infty}^{+\infty} x(t)x(t + \tau)dt = \mathcal{F}^{-1}\{S(f)\}. \tag{B.7}$$

In the latter case, of non-transient stochastic signals, this relation is called *Wiener-Khintchine theorem* (Wiener, 1930; Khintchine, 1934), and it is very relevant in the analysis of stochastic fields. In other words, the autocorrelation function of a signal can be obtained as the inverse FT of its power spectral density.
Analogously to the 1D case, the power spectral density of a stochastic field ($S(f_1, f_2)$) forms the Fourier transformation pair with the two-dimensional autocorrelation function:

$$R(\xi, \zeta) = \iint_{-\infty}^{+\infty} f(x, y)f(x + \xi, y + \zeta)dxdy = \mathcal{F}^{-1}\{S(f_1, f_2)\}. \tag{B.8}$$

To obtain the autocorrelation function, the signal needs to be standardized by removal of the mean and division by the standard deviation, otherwise only the non-centered autocovariance is obtained. Using the Wiener-Khintchine theorem and exploiting the speed of FFT, one can compute the autocorrelation functions of regularly sampled signals very efficiently (Marcotte, 1996; Velasco-Forero et al., 2009; Schiemann et al., 2011). On the other side, the classical approach to derive the spatial autocorrelation function would require computing all the Euclidean distances between pairs of points located at different ranges and along with different directions.

B.4 Short-Space Fourier Transform

The concept of Short-Space Fourier Transform is introduced through its more intuitive 1D temporal equivalent and then extended to the 2D spatial case. Namely, the phase of the Fourier representation of a time series contains information about the temporal distribution of spectral components, but its interpretation proves to be complicated. Therefore, due to the increasing need to face the non-stationarity of signal, numerous advanced methods have been proposed (Stankovic et al., 2013), all of which can be gathered under the umbrella of time-frequency signal analysis (TFSA).
The most intuitive and representative TFSA method is the Short-Time Fourier Transform (STFT), which provides Fourier spectra of localized parts of a signal:

$$X^{ST}(t,f) = \int_{-\infty}^{+\infty} x(t+\tau)w(\tau)e^{-j2\pi f\tau}d\tau, \tag{B.9}$$

where $w(\tau)$ is the window localizing the particular part of a signal around the moment $t(x(t+\tau)w(\tau))$. The introduced properties of FT (power spectral density, convolution, autocorrelation), are valid for the obtained $X^{ST}(t,f)$, in the same way, they are for the representation of the entire signal.
The original signal can be obtained by integrating all individual inverse FT:

$$x(\tau) = \int_{-\infty}^{+\infty} x(t+\tau)w(\tau)dt = \iint_{-\infty}^{+\infty} X^{ST}(t,f)e^{j2\pi f\tau}dfdt, \tag{B.10}$$

which, due to the overlapping of segments, provides a better reconstruction, particularly important in case of highly non-stationary signals. If $\int_{-\infty}^{+\infty} w(\tau) = 1$, despite the overlapping, there is no need to normalize the obtained result. Theoretically, it would be enough to perform the inverse FT at the time step $\Delta t = 2T$, which corresponds to the width of the window.
Analogously to the Fourier Transform, the Short-Time Fourier Transform can be expanded to the spatial domain, becoming the Short-Space Fourier Transform (SSFT):

$$F^{SS}(x,y,f_1,f_2) = \iint_{-\infty}^{+\infty} f(x+\epsilon,y+\mu)w(\epsilon,\mu)e^{-j2\pi(f_1\epsilon+f_2\mu)}d\epsilon d\mu, \tag{B.11}$$

with the window $w(\epsilon,\mu)$ now localizing the particular segment of a field around the position $x,y(f(x+\epsilon,y+\mu))$. The original field is obtained as:

$$\begin{aligned} f(\epsilon,\mu) &= \iint_{-\infty}^{+\infty} f(x+\epsilon,y+\mu)w(\epsilon,\mu)dxdy \\ &= \iiiint_{-\infty}^{+\infty} F^{SS}(x,y,f_1,f_2)e^{(}j2\pi(f_1\epsilon+f_2\mu))df_1df_2dxdy. \end{aligned} \tag{B.12}$$

The critical part concerns choosing the size of the window, as well as its shape. Namely, due to the uncertainty principle, a smaller window in the space domain,

that is, a better spatial resolution, results in a worse frequency resolution, and vice versa. Choosing an appropriate function $w(\epsilon, \mu)$ can decrease the otherwise inevitable uncertainty.

B.5 Window function

Equivalently to the convolution in the space domain, which corresponds to the product in its frequency counterpart (Eq. B.5), the product in space domain forms a Fourier pair with the convolution of the spectra. Therefore, the localization of the field in space, with the window function $w(\epsilon, \mu)$, implies the localization in frequency through the convolution with the FT of the window function $W(f_1, f_2)$.
The most intuitive choice of a window would be a simple rectangular function of width 2T (Fig. B.1), which in case of a field analysis, through the product of two separable functions, becomes a square:

$$w_R(\epsilon, \mu) = w_R(\epsilon) w_R(\mu) = \begin{cases} 1 & \text{for } |\epsilon| < T \text{ and } |\mu| < T \\ 0 & \text{elsewhere,} \end{cases} \tag{B.13}$$

and whose FT is as well a product of the FTs of two separable functions:

$$W_R(f_1, f_2) = \mathcal{F}\{w_R(\epsilon)\}\mathcal{F}\{w_R(\mu)\} = \frac{\sin(2\pi f_1 T)\sin(2\pi f_2 T)}{\pi^2 f_1 f_2}. \tag{B.14}$$

However, as it can be seen in the first two panels in Fig. B.1, the side lobes of $W_R(f_1, f_2)$ are somewhat strong, causing the poor localization in the frequency domain. Hence the need to come up with a more suitable form of a window, which can reduce the overall uncertainty, through the good localization in both space and frequency domain. Although the best choice in terms of uncertainty would be the Gaussian window, due to the performance in the context of our application, we have adopted the Hanning window:

$$w_H(\epsilon, \mu) = \begin{cases} \frac{1}{4}(1 + \cos(\frac{\epsilon\pi}{T}))(1 + \cos(\frac{\mu\pi}{T})) & \text{for } |\epsilon| < T \text{ and } |\mu| < T \\ 0 & \text{elsewhere.} \end{cases} \tag{B.15}$$

The corresponding FT decreases proportionally to the f^3,

$$W_H(f_1, f_2) = \frac{\pi^2 \sin(2\pi f_1 T)\sin(2\pi f_2 T)}{4\pi^2 f_1 f_2 (\pi^2 - 4\pi^2 f_1^2 T^2)(\pi^2 - 4\pi^2 f_2^2 T^2)}, \tag{B.16}$$

leading to the better localization in the frequency domain with respect to the rectangular window (see Fig. B.1).

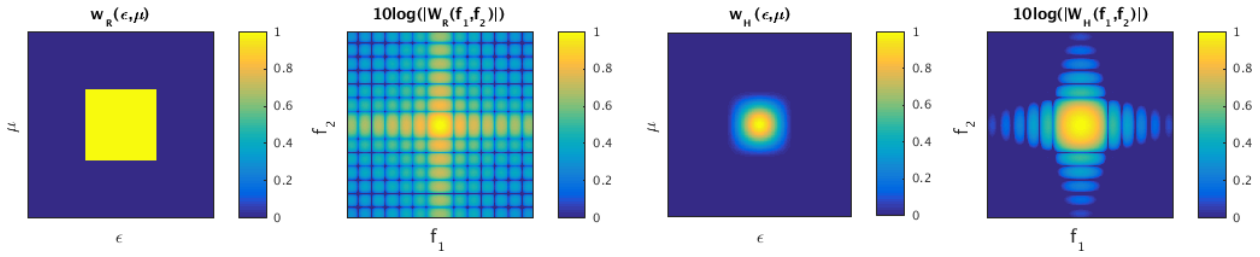


Figure B.1: FT pairs for rectangular and Hanning windows (values normalized to unity).

Finally, one can consider the use of more sophisticated window types. For instance, the Tukey window, also known as the *tapered cosine window* might help to improve the reproduction of the true local correlation structure of rainfall (Velasco-Forero, 2017).

B.6 Cascade decomposition

The scale decomposition of a two-dimensional rainfall field R can be obtained by means of its FFT (Seed, 2003). This is motivated by the fact that for a grid of size $L \times L$ pixels, the radial Fourier wavenumbers $|\boldsymbol{k}| = \sqrt{k_x^2 + k_y^2}$ are related to spatial scales via

$$\underbrace{\begin{matrix}\text{radial}\\ \text{wavenumber}\\ \text{(pixels)}\end{matrix}}_{|\boldsymbol{k}|} \rightarrow \underbrace{\begin{matrix}\text{wavelength}\\ \text{(pixels)}\end{matrix}}_{\frac{L}{|\boldsymbol{k}|}} \rightarrow \underbrace{\begin{matrix}\text{wavelength}\\ \text{(km)}\end{matrix}}_{\frac{L\Delta x}{|\boldsymbol{k}|}} \rightarrow \underbrace{\begin{matrix}\text{scale}\\ \text{(km)}\end{matrix}}_{\frac{L\Delta x}{2|\boldsymbol{k}|}}, \tag{B.17}$$

where Δx denotes the grid resolution (km), thus, the spatial scale is half the wavelength.

As done in Pulkkinen et al. (2018), Gaussian weight functions can be used for separating the Fourier spectrum into a set of radial bands. An example of the weight functions for the domain covered by the MeteoSwiss radars is shown in Fig. B.2. After the FFT and Gaussian filtering, each frequency band is transformed back to the spatial domain, which results in a cascade with n levels each representing a different scale (see an example in Fig. B.3).

The weight functions are given by

$$g_{|\boldsymbol{k}|,j} = \exp\left(-\frac{(\log_a |\boldsymbol{k}| - b_j)^2}{2c_j^2}\right), \quad 0 < |\boldsymbol{k}| \leq \frac{L}{2}, \tag{B.18}$$

where a detailed description of the parameters is given in Pulkkinen et al. (2018).

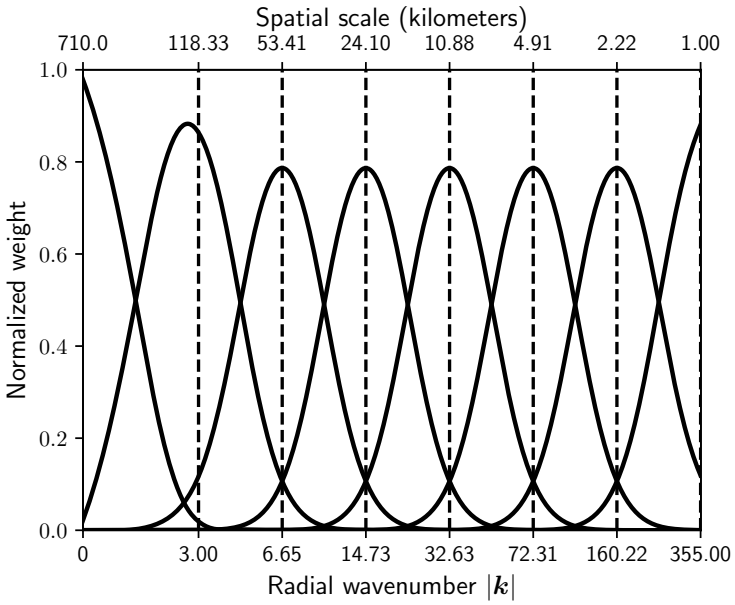


Figure B.2: Normalized weight functions with corresponding Fourier wavenumbers and spatial scales for the MeteoSwiss domain. The domain is a 710x640 grid at 1 km resolution.

This results into the cascade decomposition (Bowler et al., 2006)

$$R(x,y,t) = \sum_{j=1}^{n} \left[\sigma_j Y_j(x,y,t) + \mu_j\right], \tag{B.19}$$

of the input rainfall field R, where each level j is normalized with respect to the mean μ_j and standard deviation σ_j according to

$$Y_j(x,y,t) = \frac{R_j(x,y,t) - \mu_j}{\sigma_j}. \tag{B.20}$$

B.7 Principal component analysis

Principal components analysis (PCA) is a method to compress and summarize the information contained in a dataset of correlated variables. The procedure consists in finding an orthogonal transformation that linearly combines the variables to form a set of uncorrelated variables, which are called *principal components*. The principal components are sorted by decreasing explained variance while remaining orthogonal to each other.

The data array of size $N \times M$:

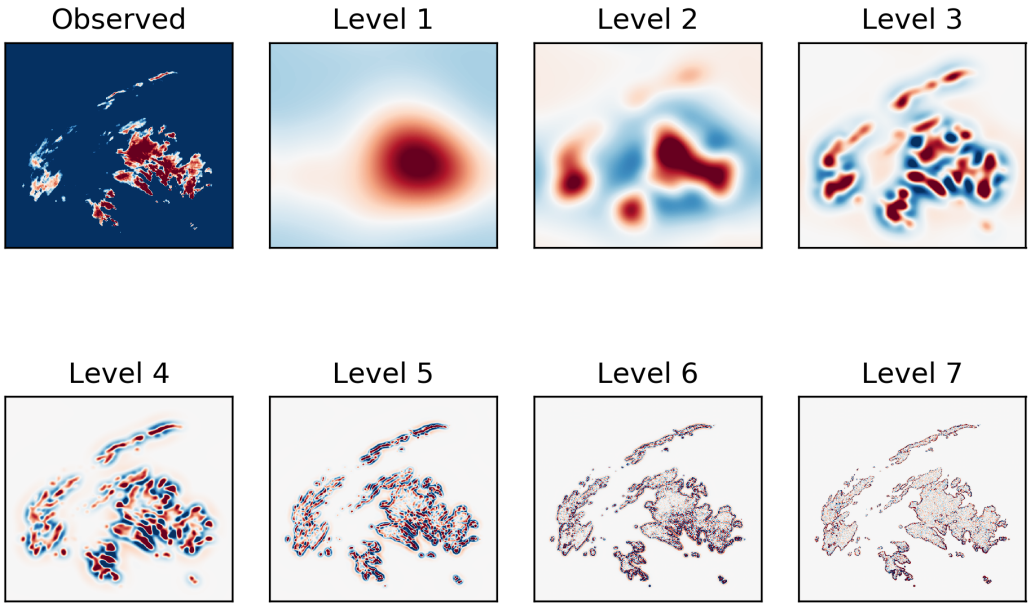


Figure B.3: The radar observations in dBR and 7 first levels of the cascade decomposition of a MeteoSwiss radar composite valid at 1630 UTC 15 May 2015. Values below -10 dBR were set to -10.1 dBR before applying the decomposition in order to reduce the discontinuity at the boundaries of precipitation areas. The observed field and the cascade levels have been normalized to zero mean and unit variance.

$$\mathbf{X}_{N,M} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,M} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N,1} & x_{N,2} & \cdots & x_{N,M} \end{pmatrix} \tag{B.21}$$

with N being the number samples and M the number variables (dimensions). These variables are often correlated, for example, because of spatial dependence. In the case of high-resolution precipitation data, rows include all values of a single precipitation field and columns contain time series at contiguous grid points. PCA exploits any existing spatial dependence to compress the information contained in the data array into a smaller set of principal components, which are used as phase space dimensions. The lower dimensional phase space is obtained by projecting the data matrix as follows:

$$\mathbf{Y}_{N,D} = \mathbf{X}_{N,M}\mathbf{U}_{M,D} \tag{B.22}$$

where $\mathbf{U}_{M,D}$ is the truncated matrix of eigenvectors and $\mathbf{Y}_{N,D}$ the matrix of principal component scores.
In more detail, PCA consists of the following steps:

1. Centering the data matrix $\mathbf{X}_{N,M}$ by the column means, i.e.
   $\hat{\mathbf{X}}_{N,M} = \mathbf{X}_{N,M} - \mathbf{1}\bar{\mathbf{x}}^{\mathbf{T}}$.

2. Computing the covariance matrix to estimate the linear dependence among variables, i.e.
   $\mathbf{C}_{M,M} = \hat{\mathbf{X}}^T_{N,M}\hat{\mathbf{X}}_{N,M}$.

3. Diagonalizing the covariance matrix by eigenvalue decomposition (EVD), i.e. $\mathbf{C} = \mathbf{U}\mathbf{V}\mathbf{U}^T$,
   where $\mathbf{U}_{M,M}$ is the orthogonal matrix of eigenvectors (each column being one vector) and $\mathbf{V}_{M,M}$ is the diagonal matrix of eigenvalues v_i.

4. Projecting the original data matrix into the space spanned by eigenvectors, i.e.
   $\mathbf{Y}_{N,M} = \hat{\mathbf{X}}_{N,M}\mathbf{U}_{M,M}$.

5. Truncating matrix and transforming back to original units, i.e.
   $\mathbf{X}'_{N,M} = \mathbf{Y}_{N,P}\mathbf{U}^T_{P,M} + \mathbf{1}\bar{\mathbf{x}}^{\mathbf{T}}$.

The eigenvectors are sorted by decreasing explained variance, which can be used as criterion for truncation such that $D \ll M$.
An alternative way to perform PCA is by singular value decomposition of the data matrix (e.g. Jolliffe, 2002). The SVD factorization of the (centered) data matrix is defined as:

$$\hat{\mathbf{X}}_{N,M} = \mathbf{L}_{N,N}\mathbf{S}_{N,M}\mathbf{R}^T_{M,M} \tag{B.23}$$

where $\mathbf{L}$ is the matrix of left singular vectors, $\mathbf{R}$ the matrix of right singular vectors, and $\mathbf{S}$ the diagonal matrix of singular values s_i. The eigenvalues can be obtained from the singular values as $v_i = s_i^2$.
Since SVD does not require the computation of the covariance matrix, it has larger numerical stability than EVD. However, SVD is slower than EVD if $N \gg M$. In such a case, one can perform a reduced SVD to avoid storing the large matrix $\mathbf{L}_{N,N}$. Finally, the projected data matrix can be obtained as

$$\mathbf{Y}_{N,M} = \mathbf{L}_{N,N}\mathbf{S}_{N,M} = \mathbf{X}_{N,M}\mathbf{R}_{M,M}. \tag{B.24}$$

References

AghaKouchak, A., E. Habib, and A. Bárdossy, 2010: Modeling Radar Rainfall Estimation Uncertainties: Random Error Model. *J. Hydrol. Eng.*, **15 (4)**, 265–274.

Alfieri, L., P. Salamon, F. Pappenberger, F. Wetterhall, and J. Thielen, 2012: Operational early warning systems for water-related hazards in europe. *Environ. Sci. Policy*, **21**, 35–49.

Anderson, J. L., 2001: An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon. Weather Rev.*, **129 (12)**, 2884–2903.

Antonetti, M., C. Horat, I. V. Sideris, and M. Zappa, 2019: Ensemble flood forecasting considering dominant runoff processes – Part 1: Set-up and application to nested basins (Emme, Switzerland). *Nat. Hazard Earth Sys.*, **19 (1)**, 19–40.

Atencia, A., and I. Zawadzki, 2014: A comparison of two techniques for generating nowcasting ensembles. Part I: Lagrangian ensemble technique. *Mon. Weather Rev.*, **142 (11)**, 4036–4052.

Atencia, A., I. Zawadzki, and M. Berenguer, 2017: Scale characterization and correction of diurnal cycle errors in MAPLE. *J. Appl. Meteor. Climatol.*, **56 (9)**, 2561–2575.

Atencia, A., and Coauthors, 2010: Improving QPF by blending techniques at the meteorological service of Catalonia. *Nat. Hazards Earth Syst. Sci.*, **10 (7)**, 1443–1455.

Ayzel, G., M. Heistermann, and T. Winterrath, 2019: Optical flow models as an open benchmark for radar-based precipitation nowcasting (rainymotion v0.1). *Geosci. Model Dev.*, **12 (4)**, 1387–1402.

Bachmann, K., C. Keil, and M. Weissmann, 2019: Impact of radar data assimilation and orography on predictability of deep convection. *Q. J. R. Meteorol. Soc.*, **145 (718)**, 117–130.

Badas, M. G., R. Deidda, and E. Piga, 2006: Modulation of homogeneous space-time rainfall cascades to account for orographic influences. *Nat. Hazard Earth Sys.*, **6 (3)**, 427–437.

Baker, L. H., A. C. Rudd, S. Migliorini, and R. N. Bannister, 2014: Representation of model error in a convective-scale ensemble prediction system. *Nonlinear Process. Geophys.*, **21 (1)**, 19–39.

Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities. *Mon. Weather Rev.*, **139 (12)**, 3887–3905.

Barton, Y., I. Sideris, U. Germann, and O. Martius, 2019: A method for real-time temporal disaggregation of blended radar-raingauge precipitation fields. *Meteorol. Appl.*, in review.

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525 (7567)**, 47–55.

Beezley, J. D., and J. Mandel, 2008: Morphing ensemble Kalman filters. *Tellus, Ser. A Dyn. Meteorol. Oceanogr.*, **60 A (1)**, 131–140.

Bellon, A., and G. Austin, 1978: The evaluation of two years of real-time operation of a short-term precipitation forecasting procedure (SHARP). *J. Appl. Meteorol.*, **17 (12)**, 1778–1787.

Bellon, A., and G. L. Austin, 1984: The accuracy of short-term radar rainfall forecasts. *J. Hydrol.*, **70**, 35–49.

Bellon, A., and I. Zawadzki, 1994: Forecasting of hourly accumulations of precipitation by optimal extrapolation of radar maps. *J. Hydrol.*, **157 (1-4)**, 211–233.

Benjamin, S. G., and Coauthors, 2004: An Hourly Assimilation–Forecast Cycle: The RUC. *Mon. Weather Rev.*, **132 (2)**, 495–518.

Berenguer, M., C. Corral, R. Sánchez-Diezma, and D. Sempere-Torres, 2005: Hydrological Validation of a Radar-Based Nowcasting Technique. *J. Hydrometeorol.*, **6 (4)**, 532–549.

Berenguer, M., D. Sempere-Torres, and G. G. Pegram, 2011: SBMcast – An ensemble nowcasting technique to assess the uncertainty in rainfall forecasts by Lagrangian extrapolation. *J. Hydrol.*, **404 (3)**, 226–240.

Berenguer, M., M. Surcel, I. Zawadzki, M. Xue, and F. Kong, 2012: The diurnal cycle of precipitation from continental radar mosaics and numerical weather prediction models. part ii: Intercomparison among numerical models and with nowcasting. *Mon. Weather Rev.*, **140 (8)**, 2689–2705.

Berenguer, M., and I. Zawadzki, 2008: A Study of the Error Covariance Matrix of Radar Rainfall Estimates in Stratiform Rain. *Weather Forecast.*, **23 (6)**, 1085–1101.

Berne, A., G. Delrieu, J.-D. Creutin, and C. Obled, 2004: Temporal and spatial resolution of rainfall measurements required for urban hydrology. *J. Hydrol.*, **299 (3)**, 166–179.

Berne, A., and W. F. Krajewski, 2013: Radar for hydrology: Unfulfilled promise or unrecognized potential? *Adv. Water Resour.*, **51**, 357–366.

Berner, J., and Coauthors, 2017: Stochastic parameterization toward a new view of weather and climate models. *Bull. Am. Meteorol. Soc.*, **98 (3)**, 565–587.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Mon. Weather Rev.*, **129 (3)**, 420–436.

Bouguet, J.-Y., 2001: Pyramidal implementation of the affine Lucas Kanade feature tracker description of the algorithm. *Intel Corp.*, **5 (1-10)**, 4.

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. *Q. J. R. Meteorol. Soc.*, **132 (620)**, 2127–2155.

Box, G. E. P., G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, 2013: *Time series analysis: forecasting and control*. 5th ed., John Wiley & Sons, New Jersey, USA.

Briggs, W. M., and R. A. Levine, 1997: Wavelets and Field Forecast Verification. *Mon. Weather Rev.*, **125**, 1329–1341.

Bröcker, J., and L. A. Smith, 2007: Increasing the reliability of reliability diagrams. *Weather Forecast.*, **22 (3)**, 651–661.

Brown, B., E. Gilleland, and E. Ebert, 2012: Chapter 6 - Forecasts of spatial fields. *Forecast verification: a practitioner's guide in atmospheric sciences*, I. T. Jolliffe, and D. B. Stephenson, Eds., Wiley-Blackwell, 95–118.

Browning, K. A., C. G. Collier, P. R. Larke, P. Menmuir, G. A. Monk, and R. G. Owens, 1982: On the Forecasting of Frontal Rain Using a Weather Radar Network. *Mon. Weather Rev.*, **110 (6)**, 534–552.

Brunet, G., S. Jones, and P. M. Ruti, 2015: Seamless Prediction of the Earth System: from minutes to months. Tech. Rep. WMO - No.1156.

Buil, A., 2017: Nowcasting probabilístico basado en observaciones de lluvia con radar meteorológico. Doctoral thesis, Universitat Politècnica de Catalunya, 175 pp.

Buizza, R., M. Miller, T. N. Palmer, and M. Milleer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Q. J. R. Meteorol. Soc.*, **125 (560)**, 2887–2908.

Casati, B., G. Ross, and D. B. Stephenson, 2004: A new intensity-scale approach for the verification of spatial precipitation forecasts. *Meteorol. Appl.*, **11 (2)**, 141–154.

Cattani, D., P. Kaufmann, S. Bader, A. Hering, and L. Nisi, 2019: Analysis of Variations in Forecast Quality during 2018. Tech. rep., Federal Office of Meteorology and Climatology MeteoSwiss, Zurich, 31 pp.

Cecinati, F., M. A. Rico-Ramirez, 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.

Chen, X., H. Yuan, and M. Xue, 2018: Spatial spread-skill relationship in terms of agreement scales for precipitation forecasts in a convection-allowing ensemble. *Q. J. R. Meteorol. Soc.*, **144 (710)**, 85–98.

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 (6)**, 1325–1347.

Clark, P., N. Roberts, H. Lean, S. P. Ballard, and C. Charlton-Perez, 2016: Convection-permitting models: A step-change in rainfall forecasting. *Meteorol. Appl.*, **23 (2)**, 165–181.

Cooley, J. W., and J. W. Tukey, 1965: An algorithm for the machine calculation of complex fourier series. *Math. Comput.*, **19 (90)**, 297–301.

Craig, G. C., C. Keil, and D. Leuenberger, 2012: Constraints on the impact of radar rainfall data assimilation on forecasts of cumulus convection. *Q. J. R. Meteorol. Soc.*, **138 (663)**, 340–352.

Crane, R., 1990: Space-time structure of rain rate fields. *J. Geophys. Res.*, **95 (D3)**, 2011–2020.

Cuo, L., T. C. Pagano, and Q. J. Wang, 2011: A Review of Quantitative Precipitation Forecasts and Their Use in Short- to Medium-Range Streamflow Forecasting. *J. Hydrometeorol.*, **12 (5)**, 713–728.

Dorninger, M., E. Gilleland, B. Casati, M. P. Mittermaier, E. E. Ebert, B. G. Brown, and L. J. Wilson, 2018: The Setup of the MesoVICT Project. *Bull. Am. Meteorol. Soc.*, **99 (9)**, 1887–1906.

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. Weather Rev.*, **139 (1)**, 272–294.

Ebert, E., and Coauthors, 2013: Progress and challenges in forecast verification. *Meteor. Appl.*, **20 (2)**, 130–139.

Ebert, E. E., 2001: Ability of a Poor Man's Ensemble to Predict the Probability and Distribution of Precipitation. *Mon. Weather Rev.*, **129 (10)**, 2461–2480.

Epstein, E. S., 1969: Stochastic dynamic prediction. *Tellus*, **21 (6)**, 739–759.

Erdin, R., C. Frei, and H. R. Künsch, 2012: Data Transformation and Uncertainty in Geostatistical Combination of Radar and Rain Gauges. *J. Hydrometeorol.*, **13 (1987)**, 1332–1346.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.-Oceans*, **99 (C5)**, 10 143–10 162.

Evensen, G., 2003: The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53 (4)**, 343–367.

Fabry, F., 2015: *Radar Meteorology: Principles and Practice*. Cambridge University Press, 256 pp.

Fabry, F., and A. W. Seed, 2009: Quantifying and predicting the accuracy of radar-based quantitative precipitation forecasts. *Adv. Water Resour.*, **32 (7)**, 1043–1049.

Farnebäck, G., 2003: Two-Frame Motion Estimation Based on Polynomial Expansion. *Image Analysis*, J. Bigun, and T. Gustavsson, Eds., Springer Berlin Heidelberg, Berlin, Heidelberg, 363–370.

Fleet, D. J., and Y. Weiss, 2006: Optical Flow Estimation. *Handbook of Mathematical Models in Computer Vision*, N. Paragios, Y. Chen, and O. D. Faugeras, Eds., Springer, 237–257.

Foresti, L., M. Reyniers, A. Seed, and L. Delobbe, 2016: Development and verification of a real-time stochastic precipitation nowcasting system for urban hydrology in Belgium. *Hydrol. Earth Syst. Sci.*, **20 (1)**, 505–527.

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 (11)**, 4671–4686.

Foresti, L., and A. Seed, 2015: On the spatial distribution of rainfall nowcasting errors due to orographic forcing. *Meteorol. Appl.*, **22 (1)**, 60–74.

Foresti, L., I. Sideris, D. Nerini, L. Beusch, and U. Germann, 2019: Using a 10-year radar archive for nowcasting precipitation growth and decay – a probabilistic machine learning approach. *Weather Forecast.*, **34 (5)**, 1547–1569.

Foresti, L., I. 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. *Q. J. R. Meteorol. Soc.*

Fortin, V., M. Abaza, F. Anctil, and R. Turcotte, 2014: Why Should Ensemble Spread Match the RMSE of the Ensemble Mean? *J. Hydrometeorol.*, **15 (4)**, 1708–1713.

Foufoula-Georgiou, E., and W. Krajewski, 1995: Recent advances in rainfall modeling, estimation, and forecasting. *Rev. Geophys.*, **33 (S2)**, 1125–1137.

Frei, C., and F. A. Isotta, 2019: Ensemble Spatial Precipitation Analysis from Rain-Gauge Data – Methodology and Application in the European Alps. *J. Geophys. Res. Atmos.*, 2018JD030004.

Fuhrer, O., C. Osuna, X. Lapillonne, T. Gysi, M. Bianco, A. Arteaga, and T. C. Schulthess, 2014: Towards a performance portable, architecture agnostic implementation strategy for weather and climate models. *Supercomput. Front. Innov.*, **1 (1)**, 44–61.

Gagne, D. J., S. E. Haupt, D. Nychka, and H. Christensen, 2018: Generation of spatial weather fields with generative adversarial networks. *4th Conference on Stochastic Weather Generators (SWGEN 2018)*, University Corporation for Atmospheric Research (UCAR), Boulder, CO, US.

Germann, U., M. Berenguer, D. Sempere-Torres, and M. Zappa, 2009: REAL – Ensemble radar precipitation estimation for hydrology in a mountainous region. *Q. J. R. Meteorol. Soc.*, **135 (639)**, 445–456.

Germann, U., M. Boscacci, M. Gabella, and S. Maurizio, 2015: Peak performance – Radar design for prediction in the Swiss Alps. *Meteorological Technology International*, 42–45.

Germann, U., J. Figueras i Ventura, M. Gabella, A. Hering, I. Sideris, and B. Calpini, 2016: Triggering innovation. *Meteorological Technology International*, 62–65.

Germann, U., G. Galli, M. Boscacci, and M. Bolliger, 2006a: Radar precipitation measurement in a mountainous region. *Q. J. R. Meteorol. Soc.*, **132 (618 A)**, 1669–1692.

Germann, U., D. Nerini, I. Sideris, L. Foresti, A. Hering, and B. Calpini, 2017: Real-time radar – A new Alpine radar network. *Meteorological Technology International*, 88–92.

Germann, U., and I. Zawadzki, 2002: Scale-dependence of the predictability of precipitation from continental radar images. Part I: Description of the methodology. *Mon. Weather Rev.*, **130 (12)**, 2859–2873.

Germann, U., and I. Zawadzki, 2004: Scale Dependence of the Predictability of Precipitation from Continental Radar Images. Part II: Probability Forecasts. *J. Appl. Meteorol.*, **43 (1)**, 74–89.

Germann, U., I. Zawadzki, and B. Turner, 2006b: Predictability of precipitation from continental radar images. Part IV: Limits to prediction. *J. Atmos. Sci.*, **63 (8)**, 2092–2108.

Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert, 2009: Intercomparison of spatial forecast verification methods. *Weather Forecast.*, **24 (5)**, 1416–1430.

Gilmore, M. S., J. M. Straka, and E. N. Rasmussen, 2004: Precipitation uncertainty due to variations in precipitation particle parameters within a simple microphysics scheme. *Mon. Weather Rev.*, **132 (11)**, 2610–2627.

Gires, A., I. Tchiguirinskaia, D. Schertzer, and S. Lovejoy, 2011: Multifractal and spatiotemporal analysis of the rainfall output of the Meso-NH model and radar data. *Hydrolog. Sci. J.*, **56 (3)**, 380–396.

Gofa, F., D. Boucouvala, P. Louka, and H. A. Flocas, 2018: Spatial verification approaches as a tool to evaluate the performance of high resolution precipitation forecasts. *Atmos. Res.*, **208 (September 2017)**, 78–87.

Golding, B. W., 1998: Nimrod: A system for generating automated very short range forecasts. *Meteor. Appl.*, **5 (1)**, 1–16.

Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, 2014: Generative adversarial nets. *Advances in Neural Information Processing Systems 27*, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, Eds., Curran Associates, Inc., 2672–2680.

Grecu, M., and W. F. Krajewski, 2000: A large-sample investigation of statistical procedures for radar-based short-term quantitative presipitation forecasting. *J. Hydrol.*, **239**, 69–84.

Groenemeijer, P., and G. C. Craig, 2012: Ensemble forecasting with a stochastic convective parametrization based on equilibrium statistics. *Atmos. Chem. Phys.*, **12 (10)**, 4555–4565.

Gustafsson, N., and Coauthors, 2018: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Q. J. R. Meteorol. Soc.*, **144 (713)**, 1218–1256.

Habib, E., A. V. Aduvala, and E. A. Meselhe, 2008: Analysis of radar rainfall error characteristics and implications for streamflow simulations uncertainty. *Hydrol. Sci. J.*, **53 (3)**, 568–587.

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. *Weather Forecast.*, **26 (2)**, 166–183.

Hamill, T. M., 2001: Interpretation of Rank Histograms for Verifying Ensemble Forecasts. *Mon. Weather Rev.*, **129 (3)**, 550–560.

Hamill, T. M., and S. J. Colucci, 1997: Verification of Eta-RSM Short-Range Ensemble Forecasts. *Mon. Weather Rev.*, **125 (6)**, 1312–1327.

Hamilton, J. D., 1994: *Time Series Analysis*. Princeton University Press, Princeton, New Jersey, USA.

Heistermann, M., S. Jacobi, and T. Pfaff, 2013: Technical Note: An open source library for processing weather radar data (wradlib). *Hydrol. Earth Syst. Sci.*, **17**, 863–871.

Heistermann, M., and Coauthors, 2015: The emergence of open-source software for the weather radar community. *Bull. Am. Meteorol. Soc.*, **96**, 117–128.

Helmus, J. J., and S. M. Collis, 2016: a library for working with weather radar data in the Python programming language. *J. Open Source Softw.*, **4 (1)**.

Hohenegger, C., and C. Schär, 2007: Atmospheric predictability at synoptic versus cloud-resolving scales. *Bull. Am. Meteorol. Soc.*, **88 (11)**, 1783–1793.

Hohti, H., J. Koistinen, P. Nurmi, E. Saltikoff, and K. Holmlund, 2000: Precipitation nowcasting using radar-derived atmospheric motion vectors. *Phys. Chem. Earth Part B P*, **25 (10–12)**, 1323–1327.

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 (1)**, 112–126.

Hutton, C., T. Wagener, J. Freer, D. Han, C. Duffy, and B. Arheimer, 2016: Most computational hydrology is not reproducible, so is it really science? *Water Resour. Res.*, **52 (10)**, 7548–7555.

Hwang, Y., A. J. Clark, V. Lakshmanan, and S. E. Koch, 2015: Improved Nowcasts by Blending Extrapolation and Model Forecasts. *Weather Forecast.*, **30 (5)**, 1201–1217.

Jacques, D., D. Michelson, J.-F. Caron, L. Fillion, D. Jacques, D. Michelson, J.-F. Caron, and L. Fillion, 2018: Latent Heat Nudging in the Canadian Regional Deterministic Prediction System. *Mon. Weather Rev.*, 3995–4014.

Jameson, A. R., M. L. Larsen, and A. B. Kostinski, 2018: On the detection of statistical heterogeneity in rain measurements. *J. Atmos. Ocean. Technol.*, **35 (7)**, 1399–1413.

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. Weather Rev.*, **143 (8)**, 3087–3108.

Jolliffe, I. T., 2002: *Principal Component Analysis*. Springer Series in Statistics, Springer-Verlag, New York, 488 pp.

Jolliffe, I. T., and D. B. Stephenson, 2003: *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. John Wiley & Sons, Chichester, UK.

Jordan, P., A. Seed, and G. Austin, 2000: Sampling errors in radar estimates of rainfall. *J. Geophys. Res.*, **105 (D2)**, 2247–2257.

Jordan, P., A. Seed, and P. Weinmann, 2003: A stochastic model of radar measurement errors in rainfall accumulations at catchment scale. *J. Hydrometeor.*, **4**, 841–855.

Joss, J., A. Waldvogel, and C. G. Collier, 1990: Precipitation Measurement and Hydrology. *Radar in Meteorology*, American Meteorological Society, Boston, MA, 577–606.

Kalman, R. E., 1960: A New Approach to Linear Filtering and Prediction Problems. *J. Fluids Eng.*, **82 (Series D)**, 35–45.

Keil, C., F. Heinlein, and G. C. Craig, 2014: The convective adjustment time-scale as indicator of predictability of convective precipitation. *Q. J. R. Meteorol. Soc.*, **140 (679)**, 480–490.

Khintchine, A., 1934: Korrelationstheorie der stationären stochastischen Prozesse. *Math. Ann.*, **109 (1)**, 604–615.

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. *Q. J. R. Meteorol. Soc.*, **144 (712)**, 744–764.

Kober, K., and G. C. Craig, 2016: Physically Based Stochastic Perturbations (PSP) in the Boundary Layer to Represent Uncertainty in Convective Initiation. *J. Atmos. Sci.*, **73 (7)**, 2893–2911.

Kumar, P., and E. Foufoula-Georgiou, 1997: Wavelet Analysis for Geophysical Applications. *Rev. Geophys.*, **(97)**, 385–412.

Kybic, J., and C. Nieuwenhuis, 2011: Bootstrap optical flow confidence and uncertainty measure. *Comput. Vis. Image Underst.*, **115 (10)**, 1449–1462.

Lantuéjoul, C., 2002: *Geostatistical simulation: Models and Algorithms*. Springer, Berlin, Heidelberg.

Lapillonne, X., and Coauthors, 2016: Operational numerical weather prediction on a GPU-accelerated cluster supercomputer. *EGU General Assembly Conference Abstracts*, EGU General Assembly Conference Abstracts, Vol. 18, 13554.

Laroche, S., and I. Zawadzki, 1994: A Variational Analysis Method for Retrieval of Three-Dimensional Wind Field from Single-Doppler Radar Data. *J. Atmos. Sci.*, **51 (18)**, 2664–2682.

Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Weather Rev.*, **102 (6)**, 409–418.

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 (3-4)**, 195–215.

Leutbecher, M., and T. N. Palmer, 2008: Ensemble forecasting. *J. Comput. Phys.*, **227 (7)**, 3515–3539.

Li, L., W. Schmid, and J. Joss, 2002: Nowcasting of Motion and Growth of Precipitation with Radar over a Complex Orography. *J. Appl. Meteorol.*, **34 (6)**, 1286–1300.

Li, P. W., and E. S. T. Lai, 2004: Short-range quantitative precipitation forecasting in Hong Kong. *J. Hydrol.*, **288 (1-2)**, 189–209.

Liechti, K., L. Panziera, U. Germann, and M. Zappa, 2013: The potential of radar-based ensemble forecasts for flash-flood early warning in the southern Swiss Alps. *Hydrol. Earth Syst. Sci.*, **17 (10)**, 3853–3869.

Lien, G. Y., E. Kalnay, and T. Miyoshi, 2013: Effective assimilation of global precipitation: Simulation experiments. *Tellus, Ser. A Dyn. Meteor. Oceanogr.*, **65 (1)**.

Ligda, M. G., 1953: Horizontal motion of small precipitation areas as observed by radar. Tech. rep., Department of Meteorology. M.I.T.

Liguori, S., M. Rico-Ramirez, A. Schellart, and A. Saul, 2012: Using probabilistic radar rainfall nowcasts and NWP forecasts for flow prediction in urban catchments. *Atmos. Res.*, **103**, 80–95.

Liu, Y., D.-G. Xi, Z.-L. Li, and Y. Hong, 2015: A new methodology for pixel-quantitative precipitation nowcasting using a pyramid lucas kanade optical flow approach. *J. Hydrol.*, **529**, 354 – 364.

Lorenz, E. N., 1973: On the Existence of Extended Range Predictability. *J. Appl. Meteor.*, **12 (3)**, 543–546.

Lorenz, E. N., 1996: Predictability: a problem partly solved. *Predict. Weather Clim.*, T. Palmer, and R. Hagedora, Eds., Cambridge University Press, 40–58.

Lovejoy, S., and D. Schertzer, 2013: *The Weather and Climate: Emergent Laws and Multifractal Cascades*. Cambridge University Press.

Lucas, B. D., and T. Kanade, 1981: An iterative image registration technique with an application to stereo vision. *Proceedings of the 1981 DARPA Imaging Understanding Workshop*, 121–130.

Mandapaka, P. V., and U. Germann, 2010: Radar-Rainfall Error Models and Ensemble Generators. *Rainfall: State of the Science (Geophysical Monograph 191)*, F. Y. Testik, and M. Gebremichael, Eds., American Geophysical Union, 247–264.

Mandapaka, P. V., U. Germann, L. Panziera, and A. Hering, 2012: Can lagrangian extrapolation of radar fields be used for precipitation nowcasting over complex alpine orography? *Weather Forecast.*, **27 (1)**, 28–49.

Marcotte, D., 1996: Fast variogram computation with FFT. *Comput. & Geosci.*, **22 (10)**, 1175–1186.

Mass, C., 2012: Nowcasting: The Promise of New Technologies of Communication, Modeling, and Observation. *Bull. Am. Meteorol. Soc.*, **93 (6)**, 797–809.

Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Am. Meteorol. Soc.*, **83 (3)**, 407–430.

Mecklenburg, S., V. A. Bell, R. J. Moore, and J. Joss, 2000a: Interfacing an enhanced radar echo tracking algorithm with a rainfall-runoff model for real-time flood forecasting. *Phys. Chem. Earth, Part B Hydrol. Ocean. Atmos.*, **25 (10-12)**, 1329–1333.

Mecklenburg, S., J. Joss, and W. Schmid, 2000b: Improving the nowcasting of precipitation in an Alpine region with an enhanced radar echo tracking algorithm. *J. Hydrol.*, **239 (1-4)**, 46–68.

Menabde, M., D. Harris, A. Seed, G. Austin, and D. Stow, 1997: Multiscaling properties of rainfall and bounded random cascades. *Water Resour. Res.*, **33 (12)**, 2823–2830.

Metta, S., J. von Hardenberg, L. Ferraris, N. Rebora, and A. Provenzale, 2009: Precipitation nowcasting by a spectral-based nonlinear stochastic model. *J. Hydrometeorol.*, **10 (5)**, 1285–1297.

Miller, M. J., and R. P. Pearce, 1974: A three-dimensional primitive equation model of cumulonimbus convection. *Q. J. R. Meteorol. Soc.*, **100**, 133–154.

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 (6)**, 2777–2797.

Nerini, D., L. Foresti, D. Leuenberger, S. Robert, and U. Germann, 2019: A reduced-space ensemble kalman filter approach for flow-dependent integration of radar extrapolation nowcasts and nwp precipitation ensembles. *Mon. Weather Rev.*, **147**, 987–1006.

Niemi, T. J., T. Kokkonen, and A. W. Seed, 2014: A simple and effective method for quantifying spatial anisotropy of time series of precipitation fields. *Water Resour. Res.*, **50 (7)**, 5906–5925.

Nisi, L., O. Martius, A. Hering, M. Kunz, and U. Germann, 2016: Spatial and temporal distribution of hailstorms in the Alpine region: A long-term, high resolution, radar-based analysis. *Q. J. R. Meteorol. Soc.*, **142 (697)**, 1590–1604.

Ott, E., B. R. Hunt, I. Szunyogh, M. Corazza, E. Kalnay, D. Patil, and J. A. Yorke, 2002: Expoliting Local Low Dimensionality of the Atmospheric Dynamics for Efficient Ensemble Kalman Filtering. *Am. Geophys. Union, Fall Meet. 2002*.

Panziera, L., M. Gabella, S. Zanini, A. Hering, U. Germann, and A. Berne, 2016: A radar-based regional extreme rainfall analysis to derive the thresholds for a novel automatic alert system in Switzerland. *Hydrol. Earth Syst. Sci.*, **20 (6)**, 2317–2332.

Panziera, L., and U. Germann, 2010: The relation between airflow and Orographic precipitation on the southern side of the Alps as revealed by weather radar. *Q. J. R. Meteorol. Soc.*, **136 (646)**, 222–238.

Pappenberger, F., H. L. Cloke, D. J. Parker, F. Wetterhall, D. S. Richardson, and J. Thielen, 2015: The monetary benefit of early flood warnings in europe. *Environ. Sci. Policy*, **51**, 278–291.

Pathirana, A., and S. Herath, 2002: Multifractal modelling and simulation of rain fields exhibiting spatial heterogeneity. *Hydrol. Earth Syst. Sci.*, **6 (4)**, 695–708.

Penny, S. G., and T. Miyoshi, 2016: A local particle filter for high-dimensional geophysical systems. *Nonlinear Process. Geophys.*, **23 (6)**, 391–405.

Poterjoy, J., 2016: A Localized Particle Filter for High-Dimensional Nonlinear Systems. *Mon. Weather Rev.*, **144 (1)**, 59–76.

Proesmans, M., L. Van Gool, E. Pauwels, and A. Oosterlinck, 1994: Determination of optical flow and its discontinuities using non-linear diffusion. *Computer Vision — ECCV '94*, J.-O. Eklundh, Ed., Springer Berlin Heidelberg, Berlin, Heidelberg, 294–304.

Pulkkinen, S., V. Chandrasekar, and A.-m. Harri, 2018: Nowcasting of Precipitation in the High-Resolution Dallas–Fort Worth (DFW) Urban Radar Remote Sensing Network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, **11 (8)**, 2773–2787.

Pulkkinen, S., V. Chandrasekar, and A.-M. Harri, 2019a: Stochastic Spectral Method for Radar-Based Probabilistic Precipitation Nowcasting. *J. Atmos. Ocean. Technol.*

Pulkkinen, S., D. Nerini, A. Pérez Hortal, C. Velasco-Forero, U. Germann, A. Seed, and L. Foresti, 2019b: Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1.0). *Geosci. Model Dev.*, **12 (10)**, 4185–4219.

Radhakrishna, B., I. Zawadzki, and F. Fabry, 2012: Predictability of precipitation from continental radar images. Part V: growth and decay. *J. Atmos. Sci.*, **69**, 3336–3349.

Raspaud, M., and Coauthors, 2018: An open-source, community-driven python framework to process earth observation satellite data. *Bull. Am. Meteorol. Soc.*, **99 (7)**, 1329–1336.

Ridal, M., M. Lindskog, N. Gustafsson, and G. Haase, 2011: Optimized advection of radar reflectivities. *Atmos. Res.*, **100**, 213–225.

Rinehart, R. E., and E. T. Garvey, 1978: Three-dimensional storm motion detection by conventional weather radar. *Nature*, **273**, 287–289.

Robert, A., 1981: A stable numerical integration scheme for the primitive meteorological equations. *Atmosphere-Ocean*, **19 (1)**, 35–46.

Robert, S., L. Daniel, and H. R. Künsch, 2017: A local ensemble transform Kalman particle filter for convective-scale data assimilation. *Q. J. R. Meteorol. Soc.*

Roberts, N. M., and H. W. Lean, 2008: Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Mon. Weather Rev.*, **136 (1)**, 78–97.

Rossa, A., F. Laudanna Del Guerra, M. Borga, F. Zanon, T. Settin, and D. Leuenberger, 2010: Radar-driven high-resolution hydro-meteorological forecasts of the 26 September 2007 Venice flash flood. *J. Hydrol.*, **394 (1-2)**, 230–244.

Särkkä, S., 2013: *Bayesian filtering and smoothing*. Cambridge University Press, 232 pp.

Sättele, M., M. Bründl, and D. Straub, 2012: Classification of warning systems for natural hazards. *Proc. 10th Int. Probabilistic Work.*, C. Moormann, M. Huber, and D. Proske, Eds., Institut für Geotechnik der Universitt Stuttgart, Stuttgart, 257–270.

Schertzer, D., and S. Lovejoy, 1987: Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. *J. Geophys. Res.*, **92(D8)**, 9693–9714.

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 (5)**, 1515–1536.

Schleiss, M., S. Chamoun, and A. Berne, 2014: Nonstationarity in Intermittent Rainfall: The "Dry Drift". *J. Hydrometeorol.*, **15 (3)**, 1189–1204.

Schmidhuber, J., 2015: Deep learning in neural networks: an overview. *Neural Networks*, **61**, 85–117.

Schraff, C., H. Reich, A. Rhodin, A. Schomburg, K. Stephan, A. Periáñez, and R. Potthast, 2016: Kilometre-scale ensemble data assimilation for the COSMO model (KENDA). *Q. J. R. Meteorol. Soc.*, **142 (696)**, 1453–1472.

Schulz, J.-P., G. Vogel, C. Becker, S. Kothe, U. Rummel, and B. Ahrens, 2016: Evaluation of the ground heat flux simulated by a multi-layer land surface scheme using high-quality observations at grass land and bare soil. *Meteorol. Z.*, **25 (5)**, 607–620.

Seed, A. W., 2003: A Dynamic and Spatial Scaling Approach to Advection Forecasting. *J. Appl. Meteor.*, **42 (3)**, 381–388.

Seed, A. W., C. E. Pierce, and K. Norman, 2013: Formulation and evaluation of a scale decomposition-based stochastic precipitation nowcast scheme. *Water Resour. Res.*, **49 (10)**, 6624–6641.

Shutts, G., and A. C. Pallarés, 2014: Assessing parametrization uncertainty associated with horizontal resolution in numerical weather prediction models. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, **372 (2018)**.

Sideris, I., L. Foresti, D. Nerini, and U. Germann, 2018: Nowprecip: an algorithm for localized probabilistic precipitation nowcasting in the complex terrain of Switzerland. *Proceedings of the 10th European Conference on Radar in Meteorology and Hydrology (ERAD), Ede-Wageningen, The Netherlands*.

Sideris, I. V., M. Gabella, R. Erdin, and U. Germann, 2014: Real-time radar–rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland. *Q. J. R. Meteorol. Soc.*, **140 (680)**, 1097–1111.

Silvestro, F., N. Rebora, G. Cummings, and L. Ferraris, 2017: Experiences of dealing with flash floods using an ensemble hydrological nowcasting chain: implications of communication, accessibility and distribution of the results. *J. Flood Risk Manag.*, **10 (4)**, 446–462.

Simonin, D., C. Pierce, N. Roberts, S. Ballard, and Z. Li, 2017: Performance of Met Office hourly cycling NWP-based nowcasting for precipitation forecasts. *Q. J. R. Meteorol. Soc.*, **143**, 2862–2873.

Slingo, J., and T. Palmer, 2011: Uncertainty in weather and climate prediction. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, **369 (1956)**, 4751–4767.

Snyder, C., T. Bengtsson, P. Bickel, and J. Anderson, 2008: Obstacles to High-Dimensional Particle Filtering. *Mon. Weather Rev.*, **136 (12)**, 4629–4640.

Stankovic, L., M. Dakovic, and T. Thayaparan, 2013: *Time-frequency signal analysis with applications*. Artech house.

Stankovic, L., I. Orovic, and E. Sejdic, 2012: *Multimedia Signals and Systems*. Springer-Verlag.

Stephan, K., S. Klink, and C. Schraff, 2008: Assimilation of radar-derived rain rates into the convectivescale model COSMO-DE at DWD. *Q. J. R. Meteorol. Soc.*, **134**, 1315–1326.

Stratman, D. R., M. C. Coniglio, S. E. Koch, and M. Xue, 2012: Use of Multiple Verification Methods to Evaluate Forecasts of Convection from Hot- and Cold-Start Convection-Allowing Models. *Weather Forecast.*, **28 (1)**, 119–138.

Sun, J., 2005: Convective-scale assimilation of radar data: Progress and challenges. *Q. J. R. Meteorol. Soc.*, **131 (613)**, 3439–3463.

Sun, J., S. B. Trier, Q. Xiao, M. L. Weisman, H. Wang, Z. Ying, M. Xu, and Y. Zhang, 2012: Sensitivity of 0–12-h Warm-Season Precipitation Forecasts over the Central United States to Model Initialization. *Weather Forecast.*, **27 (4)**, 832–855.

Sun, J., and Coauthors, 2014: Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bull. Am. Meteorol. Soc.*, **95 (3)**, 409–426.

Surcel, M., I. Zawadzki, and M. K. Yau, 2015: A study on the scale dependence of the predictability of precipitation patterns. *J. Atmos. Sci.*, **72 (1)**, 216–235.

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 (3)**, 1359–1380.

Tipping, M. E., and C. M. Bishop, 1999: Mixtures of Probabilistic Principal Component Analyzers. *Neural Comput.*, **11 (2)**, 443–482.

Todini, E., 2018: Paradigmatic changes required in water resources management to benefit from probabilistic forecasts. *Water Secur.*, **3 (August)**, 9–17.

Tong, M., and M. Xue, 2005: Ensemble Kalman Filter Assimilation of Doppler Radar Data with a Compressible Nonhydrostatic Model: OSS Experiments. *Mon. Weather Rev.*, **133 (7)**, 1789–1807.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Am. Meteorol. Soc.*, **74 (12)**, 2317–2330.

Turner, B. J., I. Zawadzki, and U. Germann, 2004: Predictability of precipitation from continental radar images. Part III: Operational nowcasting implementation (MAPLE). *J. Appl. Meteorol.*, **43 (2)**, 231–248.

Velasco-Forero, C., 2017: Interactive comment on "A non-stationary stochastic ensemble generator for radar rainfall fields based on the Short-Space Fourier Transform" by Nerini et al. *Hydrol. Earth Syst. Sci. Discuss.*

Velasco-Forero, C., D. Sempere-Torres, E. F. Cassiraga, and J. Jaime Gómez-Hernández, 2009: A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data. *Adv. Water Resour.*, **32 (7)**, 986–1002.

Venugopal, V., E. Foufoula-Georgiou, and V. Sapozhnikov, 1999: Evidence of dynamic scaling in space-time rainfall. *J. Geophys. Res. Atmos.*, **104 (D24)**, 31 599–31 610.

Villarini, G., and W. F. Krajewski, 2010: Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall. *Surv. Geophys.*, **31 (1)**, 107–129.

Wang, L. P., C. Onof, S. Ochoa, and N. Simões, 2012: Analysis of kriged rainfields using multifractals. *9th Int. Work. Precip. Urban Areas*, 138–142, Urban Challenges in Rainfall Analysis.

Wang, Y., and Coauthors, 2017: Guidelines for nowcasting techniques. Tech. Rep. WMO - No. 1198, World Meteorological Organization.

Wedi, N. P., 2014: Increasing horizontal resolution in numerical weather prediction and climate simulations: illusion or panacea? *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, **372 (2018)**, 20130 289.

Weigel, A., 2012: Chapter 8 - Ensemble forecasts. *Forecast verification: a practitioner's guide in atmospheric sciences*, I. T. Jolliffe, and D. B. Stephenson, Eds., Wiley-Blackwell, 141–166.

Weisheimer, A., T. N. Palmer, and F. J. Doblas-Reyes, 2011: Assessment of representations of model uncertainty in monthly and seasonal forecast ensembles. *Geophys. Res. Lett.*, **38 (16)**, 1–5.

Werner, M., and M. Cranston, 2009: Understanding the Value of Radar Rainfall Nowcasts in Flood Forecasting and Warning in Flashy Catchments. *Meteorol. Appl.*, **16 (1)**, 41–55.

Weusthoff, T., 2011: Weather Type Classification at MeteoSwiss - Introduction of new automatic classification schemes. Tech. Rep. 235, Federal Office of Meteorology and Climatology MeteoSwiss, 46 pp.

Wiener, N., 1930: Generalized harmonic analysis. *Acta mathematica*, **55 (1)**, 117–258.

Wilks, D., 2011a: Chapter 12 - Principal Component (EOF) Analysis. *Statistical Methods in the Atmospheric Sciences*, D. S. Wilks, Ed., International Geophysics, Vol. 100, Academic Press, 519–562.

Wilks, D., 2011b: Chapter 9 - Time Series. *Statistical Methods in the Atmospheric Sciences*, D. S. Wilks, Ed., International Geophysics, Vol. 100, Academic Press, 395–456.

Wilson, J. W., 1966: Movement and predictability of radar echoes. Tech. rep., 40 pp. U.S. Weather Bureau, Contract CWB-11093, The Center for Environment and Man Inc., 40 pp.

Wilson, J. W., and E. A. Brandes, 1979: Radar Measurement of Rainfall – A Summary. *Bull. Am. Meteorol. Soc.*, **60 (9)**, 1048–1058.

Wilson, J. W., N. A. Crook, C. K. Mueller, J. Sun, and M. Dixon, 1998: Nowcasting thunderstorms: A status report. *Bull. Am. Meteorol. Soc.*, **79 (10)**, 2079–2099.

Wilson, J. W., and E. Kessler, 1963: Use of radar summary maps for weather analysis and Forecasting. *J. Appl. Meteorol.*, **2 (1)**, 1–11.

Wong, W.-k., V. T.-l. Cheng, and W.-c. Woo, 2016: Community SWIRLS Nowcasting System (Com-SWIRLS). *WMO WWRP 4th International Symposium on Nowcasting and Very-short-range Forecast (WSN16)*, 113.

Xue, M., D. Wang, J. Gao, K. Brewster, and K. K. Droegemeier, 2003: The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteorol. Atmos. Phys.*, **82 (1-4)**, 139–170.

Yu, W., E. Nakakita, S. Kim, and K. Yamaguchi, 2015: Improvement of rainfall and flood forecasts by blending ensemble NWP rainfall with radar prediction considering orographic rainfall. *J. Hydrol.*, **531**, 494–507.

Yu, X., and T. Y. Lee, 2010: Role of convective parameterization in simulations of a convection band at grey-zone resolutions. *Tellus, Ser. A Dyn. Meteorol. Oceanogr.*, **62 (5)**, 617–632.

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 (4)**, 684–693.

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 (2)**, 83–91.

Zawadzki, I., 1973: Statistical properties of precipitation patterns. *J. Appl. Meteorol.*, **12**, 459–472.

Zawadzki, I., J. Morneau, and R. Laprise, 1994: Predictability of Precipitation Patterns: An Operational Approach. *J. Appl. Meteor.*, **33 (12)**, 1562–1571.

Zeng, Y., T. Janjić, M. Sommer, A. de Lozar, U. Blahak, and A. Seifert, 2019: Representation of Model Error in Convective-Scale Data Assimilation: Additive Noise Based on Model Truncation Error. *J. Adv. Model. Earth Syst.*

Zhang, H., and Z. Pu, 2010: Beating the Uncertainties: Ensemble Forecasting and Ensemble-Based Data Assimilation in Modern Numerical Weather Prediction. *Adv. Meteorol.*, **2010**, 1–10.

Acknowledgements

The whole story concerning the "Precip Attractor" project started in late 2014 with Loris Foresti approaching Urs Germann, head of the Radar, Satellite and Nowcasting Division at MeteoSwiss in Locarno-Monti, to discuss a possible scientific collaboration in the field of precipitation nowcasting. The collaboration materialized in early 2015 with a successful proposal to the Swiss National Science Foundation for an Ambizione grant, which also included the request for funding one Ph.D. thesis in co-supervision with Heini Wernli. I am first and foremost grateful to Loris, Urs and Heini for giving me the opportunity of being part of this very special project.

Loris, the past three years have represented a privileged time of exploration and growth, thank you. I have highly appreciated your enthusiasm for the project and your excellent guidance through the mysteries and marvels of the nowcasting science. Thank you for your patience, the stimulating discussions, and your positive energy.

Urs, I am very grateful for the opportunity of being part of your group, where you have managed to create a unique balance between the operational and research worlds. It truly is a magical place where to work. Thank you for all the unwavering support and trust throughout my Ph.D. years.

Heini, thank you for always being so welcoming and supportive. I appreciated a lot your curiosity for the project, and your precious feedback and help, also when it came to the more administrative matters at ETH.

To the past and present colleagues and friends in MeteoSwiss Locarno-Monti, thank you for all the exchanges, help, and companion throughout these three wonderful years. Thanks for all the mokas, the cookies and cakes, the swims in the Maggia – when the heatwave struck, thanks for keeping the bees and helping out with the wing, thanks for the hikes (& flies), the inspired talks, the office chats, the skitours, the relaxed lunches under the pergola. Thank you.

A special mention to Isztar Zawadzki and Alan Seed for all the precious wisdom and inspiring discussions. To Massimiliano Zappa for welcoming me for a month-long visit to his hydrology group in Birmensdorf. To the pysteps core development team, Seppo Pulkkinen, Andrés Pérez Hortal and Carlos Velasco-Forero — I have learned a lot from your incredible skills with Python!

Thanks to Marco Boscacci, Lorenzo Clementi, Marco Gabella, and Ioannis Sideris for their excellent support with the radar data, as well as to Petra Baumann, Daniel Leuenberger, and André Walser (osm team) concerning the COSMO data, with special thanks to Daniel for all the inputs, help, and consideration throughout this project. My gratitude and appreciation also go to the administrative staff of MeteoSwiss and ETH Zurich, as well as to the IT support within MeteoSwiss and the CSCS Swiss National Supercomputing Centre in Lugano.
I would also like to thank Bertrand Calpini, Deputy Director of MeteoSwiss, for his support to Ph.D. and research projects at Meteoswiss, and his general enthusiasm for innovation.
Last but not least, my deepest appreciation goes to Giada – thank you for standing by me through thick and thin, and to my family for the unconditional support and constant encouragements.

Curriculum Vitae

Daniele Nerini
born in Lugano TI, Switzerland, on 29 May 1988

Education

Jun. 2016 – Jun. 2019	PhD ETH Zurich, Atmospheric and Climate Science *Ensemble precipitation nowcasting: limits to prediction, localization and seamless blending.* Supervisors: Prof. H. Wernli, Dr. L. Foresti, Dr. U. Germann.
Oct. 2012 – Oct. 2013	M.Sc. Imperial College London, Hydrology and Water Resources Management *Using rainfall merging techniques to improve hydrological modelling in tropical upland basins.* Supervisors: Dr. C. Onof, Dr. Z. Zulkafi.
Sept. 2008 – Jul. 2011	B.Sc. University of Lausanne, Geosciences and Environment Subject area: Geology.

Research experience

Sept. 2019 – Dec. 2019	Visiting Scientist, Alpiq, Lausanne
Oct. 2018 – Nov. 2018	Visiting PhD student, WSL, Birmensdorf
Jun. 2016 – Dec. 2019	SNSF researcher, MeteoSwiss, Locarno SNSF project no. 161316: *Precipitation attractor from radar and satellite data archives and implications for seamless very short-term forecasting.*
Jan. 2015 – May. 2016	Flood Specialist, Swiss Re, Zurich
Jan. 2014 – Dec. 2014	Intern, MeteoSwiss, Locarno

Supervision of students

2019	Anand, Mohit: ConvLSTM vs Spectral Prognosis – A comparative statistical analysis of nowcasting predictions. Master project. ETH Zurich.
2019	Koller, Severin: Performance analysis of different flash flood nowcasting chains based on ensemble precipitation nowcasting in the Emme and Verzasca catchments. Master thesis. University of Zurich.
2019	Herrmann, Vincenz: Exploring hydro-meteorological uncertainties in hydrological flow forecasting. Master thesis. ETH Zurich.

Publications

Sideris, I., L. Foresti, D. Nerini, and U. Germann, 2019: NowPrecip: Localized nowcasting in the complex terrain of Switzerland. *Q.J.R. Meteorol. Soc.*, **in review**.

Foresti, L., I. Sideris, D. Nerini, L. Beusch, and U. Germann, 2019: Using a 10-year radar archive for nowcasting precipitation growth and decay – a probabilistic machine learning approach. *Wea. Forecasting*, **34**, 1547–1569.

Pulkkinen, S., D. Nerini, A. Perez Hortal, C. Velasco-Forero, U. Germann, A. Seed, and L. Foresti, 2019: Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1.0). *Geosci. Model Dev.*, **12 (10)**, 4185–4219

Nerini, D., L. Foresti, D. Leuenberger, S. Robert, and U. Germann, 2019: A Reduced-Space Ensemble Kalman Filter Approach for Flow-Dependent Integration of Radar Extrapolation Nowcasts and NWP Precipitation Ensembles. *Mont. Weather Rev.*, **147 (3)**, 987–1006.

Foresti, L., I. 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. *Q.J.R. Meteorol. Soc.*, **144 (716)**, 2277–2301.

Germann, U., D. Nerini, I. Sideris, L. Foresti, A. Hering, and B. Calpini, 2017: Real-timeradar – A new Alpine radar network. Meteorological Technology International, 88–92.

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 (6)**, 2777 – 2797.

Nerini, D., Z. Zulkafi, L.-P. Wang, C. Onof, W. Buytaert, W. Lavado-Casimiro, and J.-L. Guyot, 2015: A Comparative Analysis of TRMM–Rain Gauge Data Merging Techniques at the Daily Time Scale for Distributed Rainfall–Runoff Modeling Applications. *J. Hydrometeor.*, 16, 2153–-2168.

Talks

D. Nerini, L. Foresti, I. Sideris, L. Beusch, and U. Germann, November 2019: Using the radar archive to predict precipitation growth and decay over Switzerland - a machine learning approach. *International Conference on Weather Forecast and Hydrological Applications of Radar*, Seoul, South Korea.

Nerini, D.,L. Foresti, D. Leuenberger, H. Wernli, and U. Germann, March 2019: Scale-aware verification of COSMO-E precipitation forecasts. *MeteoSwiss NWP seminar*, Kloten, Switzerland.

Nerini, D., S. Koller, L. Foresti, I. Sideris, U. Germann, and M. Zappa, December 2018: A flash-flood forecasting system based on high-resolution ensemble precipitation nowcasting. *International Workshop on Precipitation in Urban Areas (UrbanRain)*, Pontresina, Switzerland.

Nerini, D., L. Foresti, and U. Germann, July 2018: A Bayesian combination of radar-based nowcasts with COSMO-E precipitation forecasts. *10th European conference on radar in meteorology and hydrology (ERAD2018)*, Ede-Wageningen, Netherlands.

Nerini, D., L. Foresti, and U. Germann, May 2018: A reduced-space ensemble Kalman filter approach for the Bayesian integration of radar-based nowcasts and COSMO-E precipitation forecasts. *Joint EUMETNET Workshop Connecting Nowcasting and mesoscale EPS*, Bologna, Italy.

Nerini, D., L. Foresti, and U. Germann, April 2018: A Bayesian framework based on the ensemble Kalman filter for flow-dependent integration of weather radar extrapolation nowcasts and NWP precipitation fields. *EGU General Assembly 2018*, Vienna, Austria.

Nerini, D., N. Besic, I. Sideris, U. Germann, and L. Foresti, April 2017: Introducing spatial non-stationarity in stochastic simulation of radar precipitation fields and implications for probabilistic nowcasting. *Weather Radar and Hydrology International Symposium*, Seoul, South Korea.