QUARTERLY JOURNAL OF THE ROYAL METEOROLOGICAL SOCIETY
Q. J. R. Meteorol. Soc. **135**: 445–456 (2009)
Published online 6 February 2009 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/qj.375

REAL – Ensemble radar precipitation estimation for hydrology in a mountainous region

Urs Germann[a]*, Marc Berenguer[b], Daniel Sempere-Torres[c] and Massimiliano Zappa[d]
[a]*MeteoSwiss, Locarno-Monti, Switzerland*
[b]*McGill University, Montréal, Québec, Canada*
[c]*GRAHI-UPC, Barcelona, Spain*
[d]*WSL, Swiss Federal Institute for Forest Snow and Landscape Research, Birmensdorf, Switzerland*

ABSTRACT: An elegant solution to characterise the residual errors in radar precipitation estimates is to generate an *ensemble* of precipitation fields. The paper proposes a radar ensemble generator designed for usage in the Alps using LU decomposition (REAL), and presents first results from a real-time implementation coupling the radar ensemble with a semi-distributed rainfall–runoff model for flash flood modelling in a steep Alpine catchment. Each member of the radar ensemble is a possible realisation of the unknown true precipitation field given the observed radar field and knowledge of the space–time error structure of radar precipitation estimates. Feeding the alternative realisations into a hydrological model yields a distribution of response values, the spread of which represents the sensitivity of runoff to uncertainties in the input radar precipitation field.

The presented ensemble generator is based on singular value decomposition of the error covariance matrix, stochastic simulation using the LU decomposition algorithm, and autoregressive filtering. It allows full representation of spatial dependence of the mean and covariances of radar errors. This is of particular importance in a mountainous region with large uncertainty in radar precipitation estimates and strong dependence of error structure on location. The real-time implementation of the radar ensemble generator coupled with a semi-distributed hydrological model in the framework of the forecast demonstration project MAP D-PHASE is one of the first experiments of this type worldwide, and is a fully novel contribution to this evolving area of applied research.

KEY WORDS QPE; Alpine radar; orographic precipitation; MAP

Received 21 January 2008; Revised 31 October 2008; Accepted 18 December 2008

1. Introduction

The idea of using reflectivity measurements to estimate precipitation at the ground is appealing given the high space–time resolution of weather radar observations. However, there are many sources of error which require both careful system design and sophisticated data processing (Austin, 1987; Joss and Waldvogel, 1990; Yuter, 2002). We have to deal with beam shielding and the vertical profile of reflectivity (Joss and Waldvogel, 1990; Kitchen, 1995; Joss and Lee, 1995; Pellarin *et al.*, 2002; German and Joss, 2002; Bellon *et al.*, 2007), beam smoothing and post-detection integration (Zawadzki, 1982), variability in raindrop size distributions and related uncertainty in the relation between reflectivity and rain rate (Joss and Gori, 1978; Lee and Zawadzki, 2005), attenuation by water on the radome (Germann, 1999), attenuation in heavy rain and hail (Delrou *et al.*, 2000), and overestimation in hail (Austin, 1987). Illingworth (2004) discusses to what extent polarimetric measurements can improve the accuracy of radar precipitation estimates, Friedrich *et al.* (2007) investigate the sensitivity of polarimetric rainfall estimation to partial beam shielding, Gabella *et al.* (2001) promote weighted multiple regression for radar–rain-gauge merging, Mittermaier *et al.* (2004) use mesoscale model winds to correct wind-drift errors, and Wesson and Pegram (2006) propose using geostatistics for proper interpolation and extrapolation. Of course the design of error correction depends on the site, climate and application. The best solution for a cold climate (Koistinen *et al.*, 2004) may differ from that found for a mountainous region (Germann *et al.*, 2006b).

This list of error sources and approaches for correction is far from complete, but illustrates both the range of problems and the large efforts that have been dedicated to improve quantitative precipitation estimates by radar. In spite of significant progress, the residual errors are still large and need to be taken into account, in particular in the context of operational hydrogeological applications such as issuing river runoff forecasts or flash flood and debris flow warnings.

A complete description of radar data quality includes various types of descriptor, such as system status reports, hardware monitoring messages, visibility and ground clutter maps, plausibility checks and verification statistics (Hollemann *et al.*, 2006). For hydrological applications,

*Correspondence to: U. Germann, MeteoSwiss, Via ai Monti 146, CH-6605 Locarno-Monti, Switzerland.
E-mail: urs.germann@meteoswiss.ch

we may think of a probability density function (pdf) that describes the range of possible values in space and time for each radar estimate. One single pdf per pixel, however, is not sufficient as errors are correlated in space and time, and one would need a pdf conditional on the values in the neighbourhood, or alternatively the full error variance–covariance matrix, the dimension of which corresponds to the total number of pixels in space and time. In practice, neither the conditional pdf nor the full error covariance matrix can be directly used in present hydrological models.

A promising solution to express the residual uncertainties in radar estimates is to generate an *ensemble* of precipitation fields, (e.g. Krajewski and Georgakakos, 1985). Each ensemble member is a possible realisation given the reflectivity measurements and knowledge on the radar error structure (Germann *et al.*, 2006a; Lee *et al.*, 2007). The original (deterministic) radar precipitation field is perturbed with a stochastic component, which has the correct space–time covariance structure as defined by the radar error covariance matrix. The advantage of the ensemble as opposed to more classical approaches is the simple interface with hydrology: each member can directly be fed into the hydrological model. Instead of running the hydrological model only once, we run it several times. We thus get an ensemble of possible hydrological outputs, the spread of which represents the sensitivity of the hydrological system to the uncertainty in the radar precipitation fields on input.

The idea to express uncertainty by adding a stochastic component has already been explored for radar precipitation *nowcasting* by Seed (2003) and Bowler *et al.* (2006), and evaluated in a hydrological context by Berenguer *et al.* (2005, 2006). The use of a pdf to express the uncertainty in radar nowcasts, on the other hand, was proposed in Germann and Zawadzki (2004). As a next step, from the pdf nowcasts one may calculate an ensemble which is in agreement with the individual pdfs of all pixels. A prerequisite of these stochastic nowcasting approaches is a rigorous study of the sources of uncertainty and their relative importance (e.g. Germann and Zawadzki, 2006). Clark and Slater (2006) and Bellerby and Sun (2007) provide examples of ensemble techniques for rain-gauge and satellite data.

There is an important difference between the *stochastic* ensemble proposed here, and the *dynamic* ensemble used in numerical weather prediction models (Palmer, 2002). Here, each ensemble member is the sum of the deterministic radar precipitation field and a stochastic perturbation. The stochastic term represents the measurement uncertainty and is generated such that it has the correct space–time covariance structure as defined by the radar error covariance matrix. In the case of an ensemble system of a numerical weather prediction model, on the other hand, we have a highly nonlinear dynamic system of equations, and the initial conditions are usually perturbed such that the resulting trajectories exhibit maximum perturbation growth in phase space. The size of the ensemble is limited because of limited computing time.

1.1. Outline

This paper discusses a powerful solution to generate ensembles of radar precipitation field time series, demonstrates its applicability for the mountainous context of Switzerland, and presents first results of a prototype implementation for real-time radar ensemble rainfall–runoff modelling in the southern Alps. The scheme is referred to as REAL (Radar Ensemble generator designed for usage in the Alps using LU decomposition). The concept has already been presented by Germann *et al.* (2006a). The stochastic simulation technique selected here, as opposed to other approaches, offers full flexibility regarding variances and covariances; i.e. we do not assume second-order stationarity of residual uncertainties in radar precipitation estimates. This allows us to take into account the strong dependence of radar errors on location. The prototype implementation (briefly presented in Zappa *et al.*, 2008) is, as far as the authors know, the first real-time experiment of this type in a mountainous region, and is a fully novel contribution to this evolving area of applied research. A hydrological verification will be presented in a later paper.

Section 2 presents the Alps as an ideal test site for developing and testing a radar ensemble generator. The core of the generator is desribed in section 3. The subsequent sections 4 and 5 explain how to estimate the mean and covariances of radar errors, and how to generate perturbation fields with the correct statistical properties, respectively. Section 6 presents the implementation and first results of the prototype generator coupled with a semi-distributed rainfall–runoff model. Discussion and outlook are given in section 7.

2. Test site: the Alps

The Swiss Alps are an ideal site for developing and testing a prototype radar ensemble generator for several reasons. First, MeteoSwiss has a long tradition in the operational usage of weather radar technology in a mountainous region (Joss and Waldvogel, 1990), and has implemented a systematic approach for the development, validation and deployment of algorithms (Germann *et al.*, 2006b). This includes detailed studies of the residual uncertainty of radar precipitation estimates and its space–time dependence and correlation, which is a prerequisite for generating ensembles. For the ensemble discussed in this paper, *all* algorithms presented in Germann *et al.* (2006b) including local bias correction were implemented for real-time radar precipitation estimation.

Second, uncertainties in radar precipitation fields in an Alpine region are large because of severe shielding of the radar beam by mountains, orographic precipitation mechanisms not fully seen by the radar, and strong mountain returns, i.e. clutter. As a result, radar errors exhibit complex structure with strong dependence on the location. This is an ideal set-up to explore the benefits of the ensemble compared with a simpler approach. In Switzerland, 98% of all cluttered pixels are eliminated with the

DOI: 10.1002/qj

standard procedure (Joss and Lee, 1995; Germann, 2000; Germann *et al.*, 2006b). The remaining 2% correspond to 0.07% of the whole radar volume. For hydrological applications these are further reduced by about a factor of ten by means of additional statistical filters. Clutter is thus not a serious problem any more. The dominant components of the total residual error in the operational Swiss radar precipitation fields are related to beam shielding not fully corrected with the profile correction scheme, orographic enhancement not seen by the radar (Germann and Joss, 2002), and, to some extent, beam smoothing, attenuation and uncertainty in the $Z-R$ relation caused by variability of the drop size distribution. It is the superposition of the space–time structure of these types of residual errors which needs to be modelled with the radar ensemble generator. The ensemble is the most promising solution to incorporate this complex error structure into present hydrological applications.

Third, strong and persistent orographic precipitation combined with rapid runoff in steep mountainous catchments often leads to flash flood events causing considerable damage. There is a strong need for more accurate runoff nowcasts which include quantitative assessment of the uncertainty.

Finally, in the Swiss Alps, there is a high-resolution network of quality-checked rain-gauges (Figure 1). Average spacing between two rain-gauges is 25 km for the automatic network with 10 mins resolution, and, 10 km for the climate network with daily reports. These surface measurements are used as a reference to determine the radar error covariance matrix, which is needed in the ensemble generator.

Uncertainties in radar precipitation estimates depend on location, especially in a mountainous region (e.g. Figures 8 and 10 in Germann *et al.*, 2006b). This reflects the strong relation between radar errors and factors such as distance from the antenna, radar horizon and orographic precipitation regimes. This space-dependence needs to be taken into account when generating an ensemble of radar precipitation fields. The next question is to what extent an error at a given location and time is correlated with the error at a *different* location and time. Hydrological applications typically combine many radar pixels such as a time series of all measurements over a given catchment. The sensitivity of the hydrological system to radar errors thus strongly depends on the autocorrelation of radar errors. If radar errors are completely uncorrelated in space and time, the impact on the output of the hydrological model will be relatively small, as errors average out when many samples of rainfall estimates are integrated in time and space to simulate runoff in a basin. However, if radar errors exhibit high autocorrelation at the characteristic scales of the catchment, we have to expect a large impact on the hydrological simulation. The dependence of radar errors on location is widely accepted. A preliminary study based on a large-sample radar–rain-gauge comparison shows that *spatial autocorrelation* of radar errors also depends on the location with respect to the radar and the orography (Germann *et al.*, 2006a). Autocorrelation for a station in the eastern Swiss Alps (Chur, Figure 1) exhibits high correlation with many locations in the Alpine bow and the Jura mountains in western Switzerland, whereas autocorrelation for a station in the relatively flat Swiss plateau (Reckenholz, near Zurich,

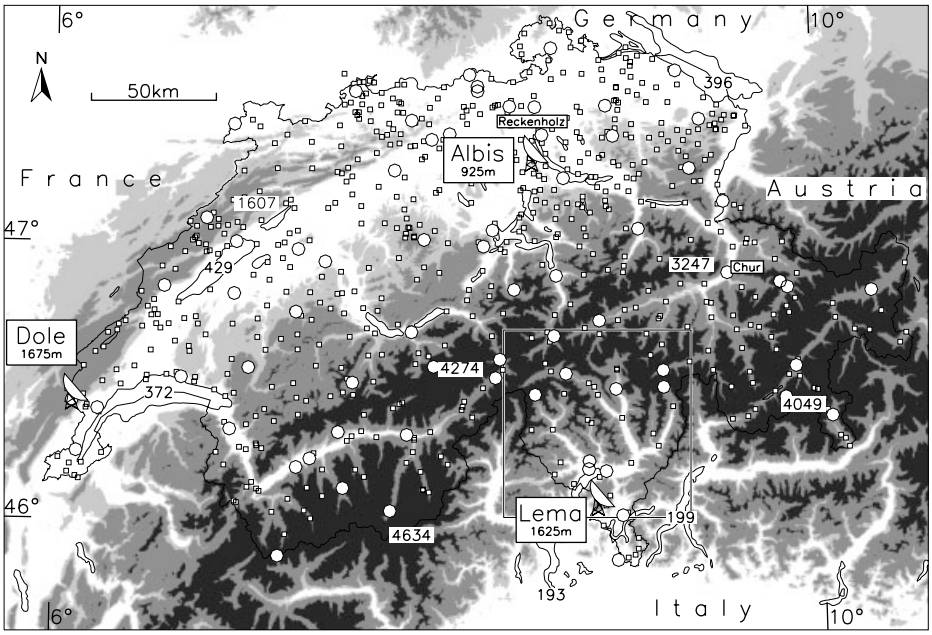

Figure 1. Situation plan of MeteoSwiss radar and rain-gauge networks. The three operational C-band radars are located on mountain tops near Zurich (Albis radar, 925 m), Geneva (La Dôle radar, 1675 m) and Lugano (Monte Lema radar, 1625 m). The white circles and small white squares indicate the locations of the automatic and climate rain-gauge networks, respectively. Levels of shading correspond to terrain height below 700 m (white), between 700 and 1000 m (light grey), between 1000 and 2000 m (dark grey), and above 2000 m (black). The three-digit and four-digit numbers indicate heights of some lakes and mountain peaks, respectively. All heights are in metres above sea level. The ensemble generator presented in this paper was implemented for a catchment in the southern Alps (in the square box north of the Lema radar), shown in Figure 3.

DOI: 10.1002/qj

Figure 1) rapidly decreases with distance from the station and shows no remote correlation at all. Autocorrelation for Reckenholz is approximately isotropic with a decorrelation length of about 30 km. Autocorrelation for Chur is clearly anisotropic with regions with high correlation well beyond 100 km. To model the spatial autocorrelation pattern of Chur with a simple distance-dependent isotropic or anisotropic variogram approach would not be satisfactory.

3. Ensemble generator

The task of the ensemble generator is to model a number of realistic error perturbation fields, $\boldsymbol{\delta}_{t,i}$, and to superimpose these onto the original unperturbed radar precipitation field

$$\underbrace{\boldsymbol{\Phi}_{t,i}}_{\text{probabilistic}} = \underbrace{\mathbf{R}_t}_{\text{deterministic}} + \underbrace{\boldsymbol{\delta}_{t,i}}_{\text{stochastic}}, \quad (1)$$

where $\mathbf{R}_t$ is the original unperturbed radar precipitation field at time t, $\boldsymbol{\delta}_{t,i}$ is the perturbation field for ensemble member i, and $\boldsymbol{\Phi}_{t,i}$ is the resulting precipitation field for ensemble member i. By precipitation we mean either precipitation rates in $\mathrm{mm\,h^{-1}}$, or precipitation amounts in mm accumulated over a pre-defined period. $\mathbf{R}_t$ is a vector that contains the original unperturbed radar precipitation estimates at time t and locations $\mathbf{x}^{\mathrm{T}} = (x_1, x_2, x_3, \ldots, x_M)$:

$$\mathbf{R}_t{}^{\mathrm{T}} = (R_{t,x_1}, R_{t,x_2}, R_{t,x_3}, \ldots, R_{t,x_M}). \quad (2)$$

M is the number of locations (typically radar pixels) in the study area Ω (typically a catchment). Analogously, $\boldsymbol{\Phi}_{t,i}$ and $\boldsymbol{\delta}_{t,i}$ are vectors the components of which are $\Phi_{t,x_k,i}$ and $\delta_{t,x_k,i}$ for all M locations $\mathbf{x}$ in Ω. $\mathbf{R}$ and $\boldsymbol{\delta}$ are henceforth referred to as the 'deterministic' and the 'stochastic' component, respectively. The result is an ensemble of precipitation fields $\boldsymbol{\Phi}_{t,i}$ with $i = 1, \ldots, N$, each member i of which is in agreement with the deterministic component $\mathbf{R}_t$ and our knowledge of the space–time structure of radar uncertainties, which goes into the stochastic component $\boldsymbol{\delta}_{t,i}$. The output $\boldsymbol{\Phi}_{t,i}$ is 'probabilistic' in the sense that for a given area Ω and time period T we have N realisations of precipitation field time series, on the basis of which one can calculate probability density of precipitation rates for that space–time window. Let Ω be a catchment and T the corresponding rainfall–runoff response time. Then, the alternative N realisations $\boldsymbol{\Phi}_{t,i}$ can be assimilated in a rainfall–runoff model, yielding a distribution of response values, the spread of which represents the sensitivity of runoff to uncertainties in the input radar precipitation field. Of course, N needs to be large enough so that $\boldsymbol{\Phi}$ adequately samples the range of uncertainties.

As most radar errors are *multiplicative*, it is sensible to define the residual errors $\boldsymbol{\epsilon}$ in dB

$$\boldsymbol{\epsilon}_t \equiv 10\log(\mathbf{S}_t/\mathbf{R}_t) \quad (3)$$

where $\mathbf{S}_t$ is the true (unknown) precipitation field, which has the same units as the radar precipitation field $\mathbf{R}_t$. Correspondingly, perturbations are added in the logarithmic domain and Equation (1) is replaced by

$$10\log[\boldsymbol{\Phi}_{t,i}] = 10\log[\mathbf{R}_t] + \boldsymbol{\delta}_{t,i}, \quad (4)$$

where $\boldsymbol{\delta}$ is now a perturbation in dB. An obvious example of a multiplicative type of error is signal attenuation caused by a water film on the radome, the protection sphere around the antenna. The water film attenuates all radar measurements in terms of received power or equivalent rain rates along a ray by the same reduction *factor*, hence the term *multiplicative*. Another example of a source of multiplicative errors are fluctuations of the $Z-R$ relation (Lee *et al.*, 2007, and references therein). Similarly, partial shielding of the radar beam by the horizon, attenuation by strong precipitation or hail, losses or instabilities in the receiver, and growth of precipitation below the lowest radar measurement to a first approximation all result in multiplicative errors.

The technique for generating the perturbation fields $\boldsymbol{\delta}_{t,i}$ can be summarised as follows. $\boldsymbol{\delta}_{t,i}$ is modelled using multivariate statistical simulation. The mathematical core is based on pre-multiplying a Gaussian random vector with zero mean and unit variance with the 'square root' of the error covariance matrix $\mathbf{C}$ and adding the mean error $\boldsymbol{\mu}$. The technique is known as 'the LU decomposition algorithm' (Goovaerts, 1997), or 'the Cholesky (decomposition) method' (Ripley, 1987; Cressie, 1993) or 'variate generation' (Johnson, 1987). The square root of the symmetric covariance matrix $\mathbf{C}$ is obtained by matrix decomposition using Cholesky factorisation or singular value decomposition (SVD; Press *et al.*, 1992). The resulting perturbation fields $\boldsymbol{\delta}_{t,i}$ are correlated Gaussian with pre-defined mean and covariances. Given the logarithmic definition of $\boldsymbol{\epsilon}$ in Equation (3), residual errors in radar precipitation fields $\mathbf{R}$ are assumed to be correlated Gaussian if expressed in dB.

If the distribution of residual errors expressed in dB can not be satisfactorily assumed Gaussian, we can add a normal score transform (e.g. Goovaerts, 1997). This is based on matching cumulative probabilities and converts the empirical distribution into Gaussian and vice versa. $\mathbf{C}$ is then calculated and imposed in normal space, i.e. after normal score transform. This way, the ensemble generator correctly reproduces the space–time error structure as defined in the normal space.

Residual errors in $\mathbf{R}$ are expected to be correlated both in space and time. Correlation in space is simulated by means of the LU decomposition algorithm, as described above. We propose two solutions to add error correlation in *time*:

1. LU decomposition algorithm. Extend the dimension of the vector $\boldsymbol{\mu}$ and the matrix $\mathbf{C}$ to include time, and model temporal correlation in the same way as spatial correlation, i.e. using the LU decomposition algorithm. The dimensions of $\boldsymbol{\mu}$ and $\mathbf{C}$ become MQ and $MQ \times MQ$, respectively, with M being

DOI: 10.1002/qj

the number of locations and Q being the number of time steps. This solution allows us to specify the mean and variance of errors for any arbitrary point in space and time (x_1, t_1), as well as the error covariance for any arbitrary pair of points (x_1, t_1) and (x_2, t_2). It is a powerful approach, but matrix decomposition puts an upper limit to the size of MQ.

2. Auto-regressive filtering. Let $\mathbf{C}_t$ be the $M \times M$ error covariance matrix in space at time t, and $\boldsymbol{\mu}_t$ be a vector of dimension M holding the mean error at locations $\mathbf{x}$ for the same time, t. Add correlation in time using temporal filtering techniques, i.e. we first generate $\boldsymbol{\delta}_{t,i}$ for given time t independently of $\boldsymbol{\delta}$ of other time steps by means of the LU decomposition algorithm, and then impose temporal correlation by filtering $\boldsymbol{\delta}$ using, for instance, a second-order autoregressive model AR(2) (Priestley, 1981). For using AR(2) we need to estimate two parameters, the lag-1 and lag-2 error correlations in time, r_1 and r_2. As solution 1, this approach allows full flexibility for specifying the covariances in space. To model *temporal* correlation, however, it uses for all locations the same two parameters, r_1 and r_2. With this solution the perturbation fields can be calculated step-by-step, which simplifies implementation in a real-time application. A similar application of an autoregressive AR(2) model was proposed by Seed *et al.* (1999) and Seed (2003).

We start by assuming temporal correlation of residual errors to be invariant in space, and adopt the second solution from above to add time, i.e. AR(2) auto-regressive filtering. For simplicity, we further assume $\boldsymbol{\mu}$, $\mathbf{C}$, r_1 and r_2 to be invariant in time. These parameters can all be calculated once offline and then saved in look-up tables. Given these assumptions, $\boldsymbol{\mu}$ and $\mathbf{C}$ are defined as

$$\begin{aligned} \mu_k &\equiv \mathrm{E}\left\{\epsilon_{x_k}\right\}, \\ \mathbf{C}_{kk} &\equiv \mathrm{Var}\left\{\epsilon_{x_k}\right\}, \\ \mathbf{C}_{kl} &\equiv \mathrm{Cov}\left\{\epsilon_{x_k}, \epsilon_{x_l}\right\}, \end{aligned} \tag{5}$$

where ϵ_{x_k} is the residual error at location x_k, and $\mathrm{E}\{\}$, $\mathrm{Var}\{\}$ and $\mathrm{Cov}\{\}$ are the expectation, variance and covariance operators, respectively. Note that the motivation for the above assumptions is of practical rather than theoretical nature: they simplify both coding of the prototype implementation and notation in the remaining part of the paper. The method can easily be extended such that the model parameters also depend on time t.

Table I summarises the two main steps of the ensemble generator. The steps are discussed in more detail in the following two sections.

4. Estimation of space–time structure of errors

There are two fundamentally different approaches to characterize the mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$ of residual errors $\boldsymbol{\epsilon}_t$ in radar precipitation estimates.

1. Comparison with ground reference. $\boldsymbol{\mu}$ and $\mathbf{C}$ are determined by comparing radar precipitation estimates $\mathbf{R}_t$ with independent ground reference, such as rain-gauge measurements $\mathbf{G}_t$. This approach is relatively simple and fast, and provides a direct estimate of the overall uncertainty in $\mathbf{R}_t$ of the operational radar product. The resulting $\boldsymbol{\mu}$ and $\mathbf{C}$ may, however, overestimate the real uncertainty $\boldsymbol{\epsilon}_t$, as the disagreement between $\mathbf{R}_t$ and $\mathbf{G}_t$ includes also rain-gauge and representativeness errors. A comprehensive analysis of radar uncertainties based on comparison against rain-gauge measurements can be found in Ciach *et al.* (2007) and references therein. To underline the fact that this approach directly gives an estimate of the total uncertainty in the operational radar product, Ciach *et al.* (2007) introduce the term 'product-error-driven uncertainty model'.
2. Systematic analysis of error sources. First, all relevant sources of residual error in $\mathbf{R}_t$ are identified and quantified individually by combining measurement theory, physical concepts, sample data and statistical simulation (e.g. Sempere-Torres *et al.*, 1994; Bellon *et al.*, 2005; Lee and Zawadzki, 2005; Lee *et al.*, 2007; Berenguer and Zawadzki, 2008). Second, the results of the individual error analyses are superimposed in order to determine the overall $\boldsymbol{\mu}$ and $\mathbf{C}$. The superposition is not trivial, because the different types of residual errors are expected to be correlated in a complex manner.

Both approaches give only approximate estimates of the real unknown $\boldsymbol{\mu}$ and $\mathbf{C}$. We opted for the first approach for two reasons:
(i) in the study area in southern Switzerland there is a dense and well-maintained rain-gauge network, and
(ii) in a mountainous region such as the European Alps, several sources of error in radar precipitation estimates need to be considered, and estimation and superposition of individual error covariance matrices is not obvious.
Thus we are more likely to get a sensible estimate of $\boldsymbol{\mu}$ and $\mathbf{C}$ by opting for the first solution and comparing $\mathbf{R}_t$ against $\mathbf{G}_t$.

Principally $\boldsymbol{\mu}$, $\mathbf{C}$, r_1 and r_2 can be updated in real time, if corresponding information is available. For instance, if the uncertainty in radar estimates exhibits significant dependence on the height of the melting layer and beam propagation, the model parameters can be conditioned to the vertical temperature and humidity profile.

In the absence of ground reference in the area under consideration, one needs a physical-statistical model (e.g. Gabella *et al.*, 2001) that allows us to estimate the radar error structure for any given location using predictors such as distance from the radar, radar visibility, terrain height, and height of melting layer. A systematic analysis of error sources mentioned above would help substantially to construct the physical-statistical error model. One difficulty lies in the fact that we would need a model of the structure of *residual* errors.

DOI: 10.1002/qj

Table I. Ensemble generator.

Step 1: Estimate space–time structure of errors
Estimate the error mean vector $\boldsymbol{\mu}$ and error covariance matrix $\mathbf{C}$ for all M locations $\mathbf{x}$ in the study area Ω.
The dimension of $\boldsymbol{\mu}$ is M, and the dimension of $\mathbf{C}$ is $M \times M$.
Decompose $\mathbf{C}$ into $\mathbf{C} = \mathbf{L}\mathbf{L}^{\mathrm{T}}$ using Cholesky factorisation or SVD.
Estimate the temporal lag-1 and lag-2 error correlation r_1 and r_2, averaged over the whole domain Ω.
Step 2: Generate perturbations and build ensemble
Generate the perturbations $\boldsymbol{\delta}_{t,i}$ using Gaussian random white noise, the LU decomposition algorithm with $\boldsymbol{\mu}$ and $\mathbf{L}$, and auto-regressive AR(2) filtering with the two parameters r_1 and r_2.
Add $\boldsymbol{\delta}_{t,i}$ to deterministic radar precipitation field $\mathbf{R}_t$, and build N-member ensemble $\boldsymbol{\Phi}_{t,i}$ by repeating step 2 for all ensemble members $i = 1, \ldots, N$.

Here, $\boldsymbol{\mu}$, $\mathbf{C}$, r_1 and r_2 are initialised once using radar–rain-gauge data from the past and are then kept constant. By doing this, they represent the error characteristics averaged over the selected calibration dataset, and decomposition of $\mathbf{C}$ needs to be done only once offline. We thus assume uncertainties to be independent of precipitation rates.

ϵ_{t,x_k}, μ_k and $\mathbf{C}_{kl}$ can be estimated at the rain-gauge locations using

$$\begin{aligned}
\widehat{\epsilon}_{t,x_k} &= 10\log(G_{t,x_k}/R_{t,x_k}),\\
\widehat{\mu}_k &= \frac{1}{\sum_{t=1}^{Q}\omega_{t,x_k}}\sum_{t=1}^{Q}\omega_{t,x_k}\widehat{\epsilon}_{t,x_k},\\
\widehat{\mathbf{C}}_{kk} &= \frac{1}{\sum_{t=1}^{Q}\omega_{t,x_k}^2}\sum_{t=1}^{Q}\omega_{t,x_k}^2(\widehat{\epsilon}_{t,x_k}-\widehat{\mu}_k)^2,\\
\widehat{\mathbf{C}}_{kl} &= \frac{\sum_{t=1}^{Q}\omega_{t,x_k}\omega_{t,x_l}(\widehat{\epsilon}_{t,x_k}-\widehat{\mu}_k)(\widehat{\epsilon}_{t,x_l}-\widehat{\mu}_l)}{\sum_{t=1}^{Q}\omega_{t,x_k}\omega_{t,x_l}},
\end{aligned} \tag{6}$$

where ω_{t,x_k} is a weight for time t and rain-gauge location x_k, and Q is the number of time steps. The hat over ϵ, μ and $\mathbf{C}$ denotes that these equations are estimators.

Samples with either $G = 0$ (false alarms) or $R = 0$ (misses) are excluded from the dataset. This is correct, as the task of the ensemble generator is to model uncertainty in radar observations of existing precipitation areas. Misses and false alarms have quite different error structures, and including them in the dataset would make little sense both from a conceptual and a statistical point of view. Fortunately misses and false alarms are rare during convective summer rainfall in Switzerland.

Weighting is necessary to avoid irrelevant samples erroneously having large influence. By setting the weight to the observed precipitation, $\omega_{t,x_k} = R_{t,x_k}$, we force the expected value of the ensemble for a given location to be equal to the original unperturbed component at that location. That is, the ensemble generator is bias-free in terms of water amounts. Strictly speaking, this presumes that ϵ_{t,x_k} is Gaussian, or was transformed to Gaussian using the normal score transform.

Figure 2 depicts cumulative distribution of residual errors $\widehat{\boldsymbol{\epsilon}}$ for hourly precipitation in southern Switzerland. Each $\widehat{\epsilon}_{t,x_k}$ (plotted on the abscissa) is weighted by the corresponding R_{t,x_k} (see ordinate). Gaussian distributions are also shown. It can be seen that $\widehat{\epsilon}_{t,x_k}$ weighted by

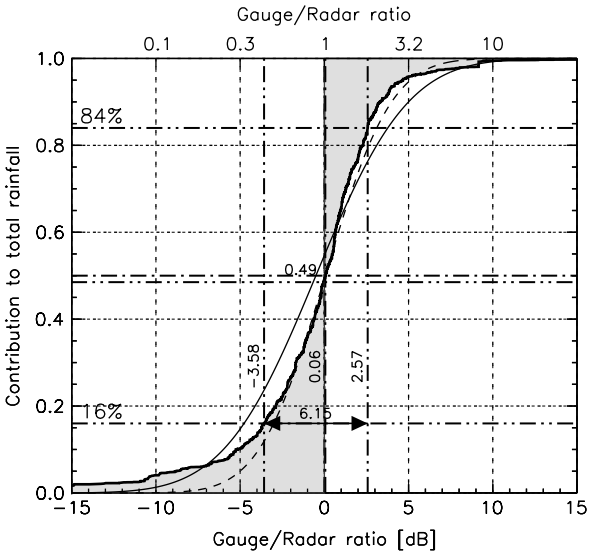

Figure 2. Residual errors in radar precipitation estimates expressed in dB and weighted by observed precipitation (thick solid s-shaped line) roughly follow a Gaussian distribution (thin s-shaped curves). Data shown here are radar–rain-gauge agreement for hourly accumulations from a 6-month calibration dataset. The two thin s-shaped lines are Gaussian distributions using different fitting techniques: the solid line is a Gauss fit using classical mean and standard deviation, whereas the dashed line is a Gauss fit using the median and the distance between the 16% and 84% percentiles. This type of error diagram was introduced by Joss and Waldvogel (1990) and further refined by Germann *et al.* (2006b).

R_{t,x_k} can be approximated by a Gaussian function. For precise reproduction of the observed error distribution, it is recommended to apply the normal score transform mentioned previously. Equation (6) needs to be extended if normal score transform is applied.

Equation (6) provides estimates of $\widehat{\mu}_k$ and $\widehat{\mathbf{C}}_{kl}$ at rain-gauge locations only. In order to calculate perturbations $\boldsymbol{\delta}$ for all M locations in Ω we need to interpolate between the rain-gauge locations. Interpolation can be done either at the level of $\widehat{\boldsymbol{\mu}}$ and $\widehat{\mathbf{C}}$, or alternatively at the level of $\boldsymbol{\delta}$ after invoking the LU decomposition algorithm to generate the perturbations at the rain-gauge locations. We tested both solutions. Interpolation is done linearly using the Delaunay triangulation.

If some of the rain-gauges do not provide the temporal resolution desired for the ensemble, the corresponding error variances and covariances need to be scaled down.

DOI: 10.1002/qj

This was the case when implementing the prototype generator in southern Switzerland. The temporal resolution of the climate rain-gauge network used to study the dependence of ϵ on location is 24 hours, whereas the desired temporal resolution of the ensemble is 1 hour. Measurements from eight automatic rain-gauges with 10-minute resolution were used to investigate how variances and covariances increase with increasing temporal resolution in that region. In the 6-month calibration dataset used later for the prototype implementation, variances and covariances scale approximately linearly if plotted against the logarithm of temporal resolution.

The high-resolution data from the automatic rain-gauges were also used to estimate temporal lag-1 and lag-2 error correlation averaged over domain Ω. Here we provide the estimators for r_1 and r_2 for one single rain-gauge location x_k. The equations can be easily extended to obtain domain-average estimates.

$$\begin{aligned}\widehat{r}_1 &= \frac{\sum_{t=1}^{Q-1} \omega_{t,x_k}\omega_{t+1,x_k}(\widehat{\epsilon}_{t,x_k}-\widehat{\mu}_k)(\widehat{\epsilon}_{t+1,x_k}-\widehat{\mu}_k)}{\widehat{\mathbf{C}}_{kk}\sum_{t=1}^{Q-1}\omega_{t,x_k}\omega_{t+1,x_k}} \\ \widehat{r}_2 &= \frac{\sum_{t=1}^{Q-2} \omega_{t,x_k}\omega_{t+2,x_k}(\widehat{\epsilon}_{t,x_k}-\widehat{\mu}_k)(\widehat{\epsilon}_{t+2,x_k}-\widehat{\mu}_k)}{\widehat{\mathbf{C}}_{kk}\sum_{t=1}^{Q-2}\omega_{t,x_k}\omega_{t+2,x_k}}.\end{aligned} \tag{7}$$

These two parameters are required for AR(2) filtering to impose the correct temporal correlation on $\boldsymbol{\delta}_{t,i}$.

5. Generation of perturbation fields

There are different techniques to produce correlated Gaussian fields, see e.g. the chapter on assessment of spatial uncertainty in Goovaerts (1997), or the discussion of multivariate distributions in Johnson (1987) and Ripley (1987). The most versatile technique for the given application is the LU decomposition algorithm briefly explained in section 3. It offers full flexibility regarding space–time dependence of errors, and allows for conditioning (Goovaerts, 1997) if we wish to incorporate additional external information. The possibility for conditioning is attractive but is not part of this first-generation ensemble generator. It will be explored in future work.

As the name says, the LU decomposition algorithm starts by decomposing the symmetric positive definite covariance matrix into a lower-triangular and an upper-triangular matrix

$$\mathbf{C} = \mathbf{L}\mathbf{L}^{\mathrm{T}} \tag{8}$$

This can be done with the Cholesky algorithm or SVD (Press *et al.* 1992). Cholesky is faster but may be numerically unstable if the matrix is near to singular. In order to increase numerical stability, we recommend use of the modified Cholesky algorithm proposed by Gill and Murray in 1974 (Gill *et al.*, 1981). The advantages of SVD are high stability and the possibility for matrix reduction by eliminating negligibly small singular values (e.g. Wesson and Pegram, 2004). Both modified Cholesky and SVD including matrix reduction have been tested.

The LU decomposition algorithm is based on pre-multiplying a Gaussian random vector $\mathbf{y}_{t,i}$ by $\mathbf{L}$ and adding $\boldsymbol{\mu}$:

$$\boldsymbol{\delta}_{t,i} = \boldsymbol{\mu} + \mathbf{L}\mathbf{y}_{t,i} \tag{9}$$

$\mathbf{y}_{t,i}$ is a zero-mean and unit-variance Gaussian random vector of dimension M

$$\mathbf{y}_{t,i} = N_M(\mathbf{0}, \mathbf{I}) \tag{10}$$

and can be generated by M successive calls to a univariate normal generator. The resulting $\boldsymbol{\delta}_{t,i}$ is a multivariate normal vector distributed $N_M(\boldsymbol{\mu}, \mathbf{C})$. Matrix reduction in SVD reduces $\mathbf{L}$ to that part of the covariance structure that is significant. This improves speed and prevents introduction of spurious noise in Equation (9).

We then use AR(2) filtering (Priestley, 1981) to impose the desired temporal structure. The AR(2) model calculates the perturbation field for time step t by combining $\mathbf{L}\mathbf{y}_{t,i}$ with the AR(2) filtered perturbation fields of the previous two time steps $t-1$ and $t-2$. The procedure is iterative and $\boldsymbol{\mu}$ is added only at the end. Combining AR(2) filtering with Equation (9) yields

$$\boldsymbol{\delta}'_{t,i} = \mathbf{L}\mathbf{y}_{t,i} - a_1\boldsymbol{\delta}'_{t-1,i} - a_2\boldsymbol{\delta}'_{t-2,i}\,, \tag{11}$$

$$\boldsymbol{\delta}_{t,i} = \boldsymbol{\mu} + v\boldsymbol{\delta}'_{t,i}\,, \tag{12}$$

with the AR(2) model parameters a_1 and a_2 estimated by means of the Yule-Walker equations (Priestley, 1981), and v, the square root of the AR(2) variance rescaling factor:

$$\begin{aligned} a_1 &= r_1\frac{r_2-1}{1-r_1^2} \\ a_2 &= \frac{r_1^2-r_2}{1-r_1^2} \\ v &= \left[\frac{1+a_2}{(1-a_2)(1-a_1+a_2)(1+a_1+a_2)}\right]^{-0.5}. \end{aligned} \tag{13}$$

Note that AR(2) filtering changes the variance by a linear factor that depends on the model parameters a_1 and a_2. That is why we need to rescale the perturbations $\boldsymbol{\delta}'_{t,i}$ with the scaling factor v in Equation (12). The ensemble $\boldsymbol{\Phi}_{t,i}$ is built by evaluating Equations (10), (11) and (12) for all ensemble members i, and adding the resulting $\boldsymbol{\delta}_{t,i}$ to $\mathbf{R}_t$ using Equation (4).

Equations (4), (6), (7), (8), (10), (11), (12) and (13) constitute the mathematical core of the ensemble generator. If space–time correlations of residual radar uncertainties are assumed to be constant, Equations (6), (7), (8) and (13) can be evaluated once offline and saved as look-up tables. This helps to save computing time in a real-time implementation.

6. Implementation

The ensemble generator was implemented, and coupled with a semi-distributed rainfall–runoff model as part of

DOI: 10.1002/qj

the Mesoscale Alpine Programme (MAP) hydrometeorological forecast demonstration project MAP D-PHASE (Zappa *et al.*, 2008) and the European concerted research action on the propagation of uncertainty in advanced hydrometeorological forecast systems (COST-731; http://cost731.bafg.de). This section briefly describes the implementation and first results, and demonstrates the applicability of the present approach to radar–hydrology coupling in a mountainous region.

The experiment is running in a steep Alpine 2800 km^2 area enclosing the rivers Maggia, Verzasca, Ticino and Moesa (Figures 3 and 4). The ensemble of precipitation field time series is fed in real time into the semi-distributed precipitation–runoff–evapotranspiration–hydrological response unit model PREVAH. For more information on PREVAH, the reader is referred to Zappa *et al.* (2003), Wöhling *et al.* (2006), Ranzi *et al.* (2007) and Viviroli *et al.* (2007).

The operational MeteoSwiss radar precipitation product with a resolution of 1 km and 5 min is taken as the deterministic component $\mathbf{R}_t$. Germann *et al.* (2006b) provides a detailed description. Perturbations $\boldsymbol{\delta}_{t,i}$ are generated at a resolution of 2 km and 1 h. Before evaluating Equation (4), $\mathbf{R}_t$ is aggregated in time and space such that it matches the resolution of $\boldsymbol{\delta}_{t,i}$. The number of ensemble members N is set to 25, which is sufficiently large for a proof of concept. The sensitivity of runoff modelling to ensemble size will be studied later. In order to simplify description, coding and verification of this experiment, we start by assuming $\boldsymbol{\mu}$, $\mathbf{C}$, a_1, a_2 and v to be invariant in time, as discussed at the end of section 3. Thus, Equations (6), (7), (8) and (13) need to be evaluated only once offline. The equations that remain to be executed in real time are Equations (10), (11), (12) and (4).

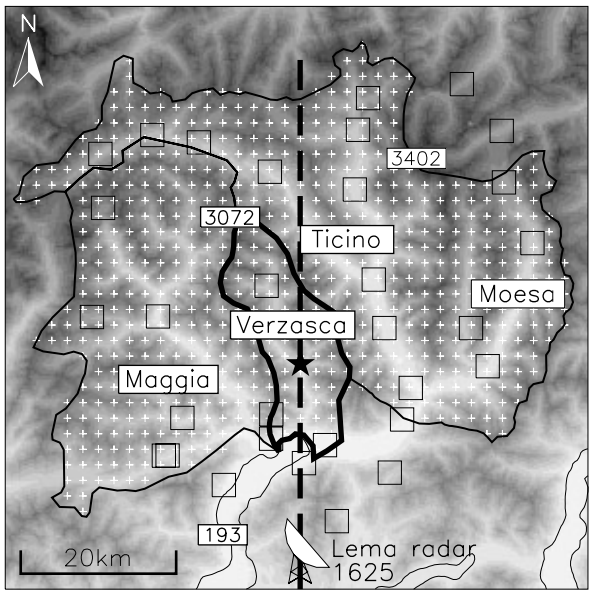

Figure 3. Test catchment 'Maggia–Verzasca–Ticino–Moesa' used for prototype implementation of radar ensemble runoff modelling. Three- and four-digit numbers are heights (m amsl) of lakes and mountain peaks. Square boxes indicate location of rain-gauges used to estimate mean and covariances of radar errors. The black star shows the location of the river-gauge 'Lavertezzo'. Figure 4 shows a vertical cross-section along the thick dashed line.

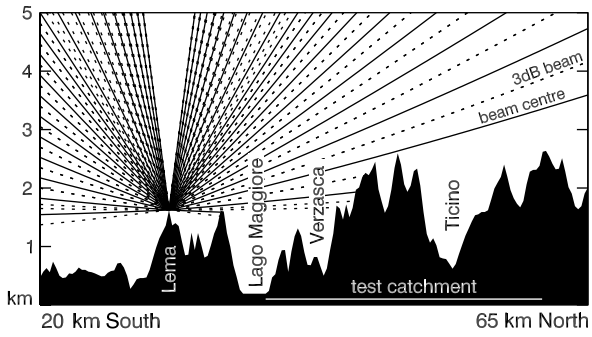

Figure 4. Vertical cross-section of radar scan geometry over the test catchment 'Maggia–Verzasca–Ticino–Moesa'. The full radar volume consisting of 20 elevation sweeps is updated every 5 minutes (Joss and Lee, 1995). The horizontal white line indicates the location of the test catchment. The cross-section clearly reveals severe shielding of the radar beam by mountains and increasing pulse volume with increasing distance from the radar antenna. As a consequence of beam blocking and increasing pulse volume, we get a strong dependence of radar errors on geographic location. The resulting spatial structure of radar errors needs to be taken into account in the ensemble generator.

$\boldsymbol{\mu}$ and $\mathbf{C}$ are estimated for all M 2 km pixels (white crosses in Figure 3) by interpolating the mean, variances and covariances of ϵ_{t,x_k} (Equation (6)) of a 6-month calibration dataset at 31 rain-gauge locations in the test area (square boxes in Figure 3). $\mathbf{C}$ is decomposed into $\mathbf{C} = \mathbf{L}\mathbf{L}^{\mathrm{T}}$ using SVD. Plotting the resulting singular values shows that only about 29 are significant, whereas the remaining $M - 29$ are negligibly small. This suggests

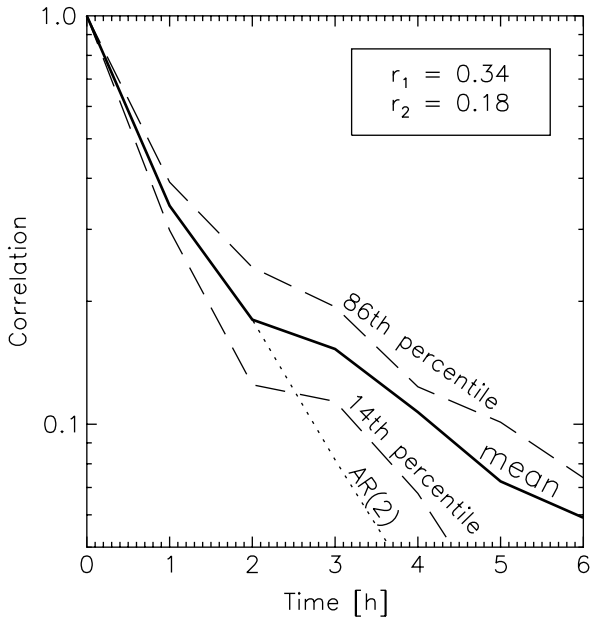

Figure 5. Temporal decorrelation of residual errors in radar precipitaton estimates averaged over the test catchment 'Maggia–Verzasca–Ticino–Moesa'. Mean correlations are 0.34 and 0.18 for 1-hour and 2-hour time lags, respectively. These two numbers are used to estimate the parameters for AR(2) filtering in the ensemble generator (Equation (13)). The dotted line shows the temporal decorrelation of the corresponding AR(2) model. The 14th and 86th percentiles (dashed) give some idea of the spatial variability of temporal decorrelation within the catchment.

DOI: 10.1002/qj

that a large part of the matrix **L** can be set to 0 by matrix reduction (Wesson and Pegram, 2004), which speeds up computation time significantly. The number 29 is not surprising, as this is a little less than the number of rain-gauges that were used to estimate **C** in the first place. In fact, as pointed out in section 4, one can alternatively evaluate $\boldsymbol{\mu}$, **C L**, and $\boldsymbol{\delta}$ for the small number of rain-gauge locations only, and determine the perturbations for all M pixels in a second step by interpolation using $\boldsymbol{\delta}$ at the rain-gauge locations. The results are virtually the same; the choice depends on practical aspects regarding coding. The solution with the dimension of **C** being the number of rain-gauge locations will be the preferred way when we start to vary **C** in time.

The AR(2) parameters a_1, a_2 and v are estimated from the temporal lag-1 and lag-2 correlation of radar–rain-gauge agreement in the catchment (Equations (7) and (13); Figure 5). In the area under consideration, spatial and temporal decorrelation length, defined as the lag at which correlation drops below 1/e, is roughly 1 hour (Figure 5), and a few tens of km (between 10 and 40) depending on location and direction. Figure 6 gives some idea of how δ changes in space, time and between different ensemble members: it shows $\boldsymbol{\delta}_{t,i}$ for three consecutive time steps and three different randomly chosen ensemble members.

A cross-validation experiment leaving out one rain-gauge at a time was performed to assess the uncertainty in our estimates of **C** given by the limited spatial resolution of the rain-gauge network. For most rain-gauges, the cross-validation difference is three to ten times smaller than the variance at that location. Largest uncertainty was found in regions with strongest radar beam shielding, that is towards the northnortheast. This is not surprising, as beam shielding results in strong spatial (azimuthal) gradients in the radar error structure. Average spacing between rain-gauges is 10 km, which is at the lower end of values found for the spatial decorrelation length of ϵ. We can conclude that the network with 31 rain-gauges is sufficiently dense to capture the main spatial features of the radar error structure.

The expected value of $\Phi_{t,x_k,i}$ for a given location x_k is equal to the determinstic component at that location R_{t,x_k}; i.e. the ensemble generator is bias-free in terms of precipitation amounts (discussion of Equation (6) in section 4). This can be verified numerically by evaluating $\mathrm{E}\left\{\Phi_{t,x_k,i}\right\}$ for large N. Similarly, comparing the covariance matrix **C** on input with the covariances calculated experimentally from several thousands of simulated perturbation fields confirms that the present ensemble generator correctly reproduces the desired covariance structure of uncertainties in radar precipitation estimates.

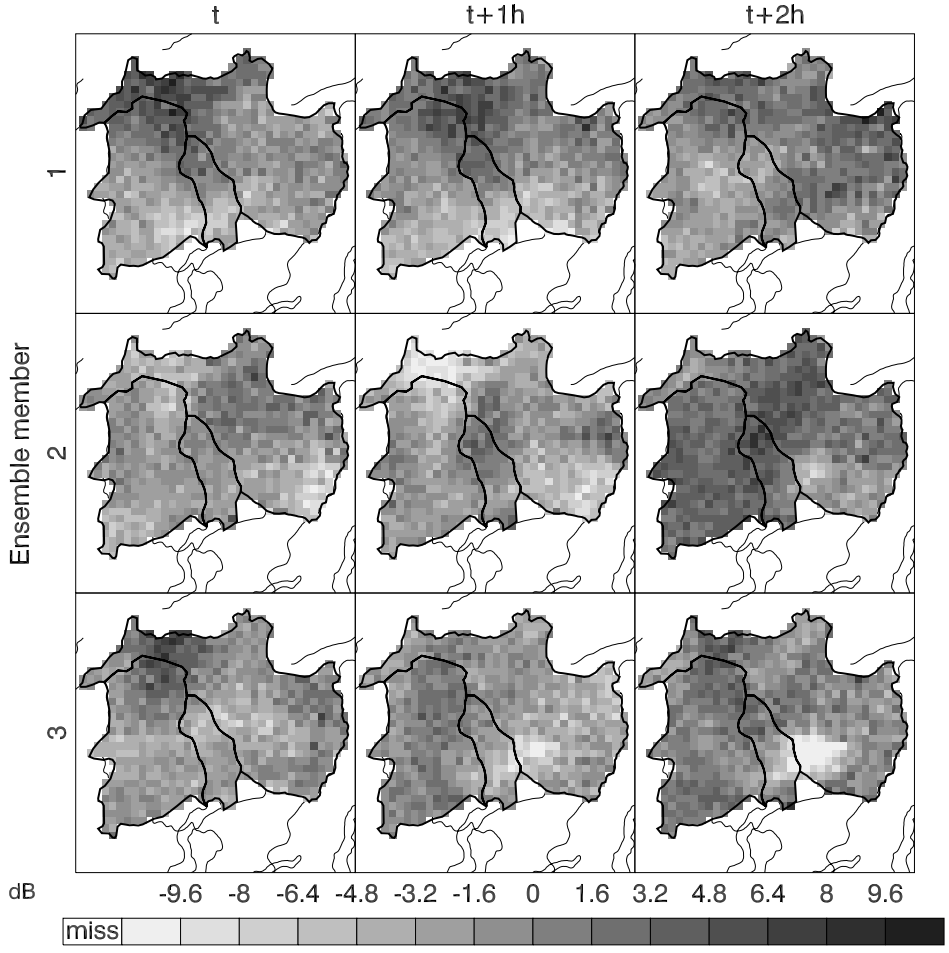

Figure 6. Example of perturbation field $\boldsymbol{\delta}_{t,i}$ for three time steps (columns) and three randomly chosen ensemble members (rows).

DOI: 10.1002/qj

The coupled radar ensemble runoff system is running automatically and being updated in real time every hour since May 2007. Figure 7 depicts an example of observed versus modelled runoff at the river-gauge at Lavertezzo (Verzasca river) for a major flood event in September 2008.

7. Discussion

The paper presents a novel probabilistic approach to represent the uncertainty in radar precipitation fields and to examine uncertainty propagation in a coupled radar–hydrological forecast system. The methodology was designed for a mountainous region with large uncertainties in radar precipitation estimates and strong dependence of uncertainty on geographic location. It assumes radar errors to be correlated Gaussian if expressed as factors, a reasonable assumption given the fact that most radar errors are actually multiplicative. The model allows full representation of spatial dependence of error mean and covariances and considers error correlation in space *and* time.

A real-time experiment coupling the radar ensemble with a rainfall runoff model for a 2800 km^2 area in the southern Alps demonstrates the applicability of the method for a steep mountainous context. It is one of the very first experiments of this type. Figure 7 shows an example of the output at the end of the chain for the Verzasca catchment. For this steep 186 km^2 catchment the time between a precipitation impulse and the main runoff response at the outlet is of the order of 1 to 2 hours. Hence, even if no precipitation *forecast* is available, coupling radar observations with a runoff model results in 1–2 hours lead time for predicting a flash flood peak.

A preliminary quantitative evaluation of modelled versus observed runoff was performed for the Pincascia river, a 44 km^2 sub-catchment of the Verzasca river, (Table II). The evaluation uses runoff data from the first year of the real-time experiment and is based on bias and scatter statistics as introduced in Germann *et al.* (2006b). The results allow three important conclusions:

1. Accuracy of radar-driven runoff is comparable to that of rain-gauge-driven runoff. As this result represents average conditions for all seasons, we

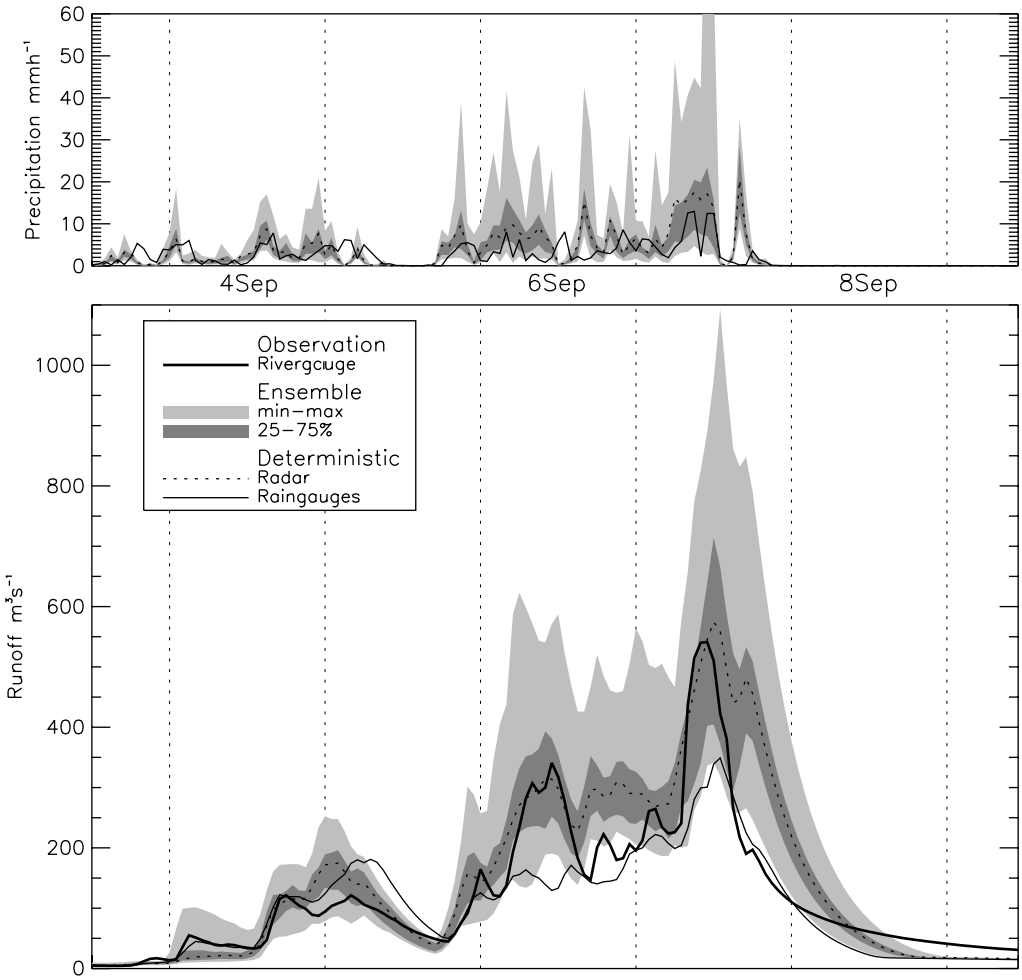

Figure 7. Lower panel: ensemble hourly runoff nowcasting of a major flood event in the Verzasca catchment at the Lavertezzo river-gauge (186 km^2) in September 2008. The thick solid line indicates runoff observed at Lavertezzo. Thin lines are modelled runoff forced by interpolated rain-gauge measurements (solid) and deterministic radar estimates (dashed). Grey shading shows modelled runoff forced by probabilistic radar estimates (radar ensemble). Upper panel: hourly precipitation amounts from both radar and rain-gauge datasets for the same period. Plots of this type have been routinely generated in real time since May 2007.

DOI: 10.1002/qj

Table II. Results from the first year of the real-time radar ensemble runoff modelling experiment.

Pincascia river	Radar deterministic R_t	Radar ensemble $\Phi_{t,i}$	Rain-gauges deterministic G_t
Bias (%)	–16	–16	–29
Scatter (dB)	2.1	2.2	2.0

Bias and scatter are evaluated for modelled versus observed runoff.

expect radar to outperform rain-gauges for convective situations with strong spatial rainfall variability. This is a remarkable result considering, on the one hand, the difficulties in radar precipitation estimation in a mountainous region and, on the other hand, the relatively dense automatic rain-gauge network in the area under consideration.

2. The scatter between observed runoff and runoff driven by the radar ensemble ($\mathbf{\Phi}_{t,i}$) is only slightly larger than the scatter between observed runoff and runoff driven by the deterministic radar field ($\mathbf{R}_t$). That means, in terms of runoff, the ensemble generator does not overstate the uncertainty that is already present in the deterministic component.
3. $\mathbf{\Phi}_{t,i}$ has same bias as $\mathbf{R}_t$. This confirms that the presented generator is bias-free in terms of water amounts (discussion of Equation 6 in section 4).

The probabilistic ensemble approach opens up new possibilities to examine the propagation of uncertainty in advanced hydrometeorological forecast systems. It allows us to evaluate the sensitivity of runoff to uncertainties in the radar precipitation field on input, and to represent the final uncertainty at the end of the chain in an easy-to-understand format.

Future work will include an extensive hydrological verification of the results of the real-time implementation, and refinements of the ensemble generator. For instance, we would like to stratify the error structure and make $\boldsymbol{\mu}$ and $\mathbf{C}$ conditional to additional external information, such as the height of the melting layer or orographic enhancement. Ciach *et al.* (2007) show dependence of errors on precipitation expressed in mm. By stratifying errors using physical factors mentioned above, the dependence on precipitation rates may already be captured. Another future refinement of the generator deals with conservation of variance at different scales. In the present implementation, the variance of the ensemble $\mathbf{\Phi}_{t,i}$ is greater than that of the observed field $\mathbf{R}_t$ due to the variance that is added by $\boldsymbol{\delta}_{t,i}$. From a practical point of view, however, this seems to be of little importance, as shown by the preliminary quantitative evaluation discussed above. Finally, we plan to relax the restriction of temporal correlation being invariant in space by applying also for time the LU decomposition algorithm successfully implemented for space.

Acknowledgements

The authors wish to thank M. Rotach, M. Arpagaus, M. Boscacci and A. Rossa for support in the MAP D-PHASE and COST-731, and I. Zawadzki, G. Pegram, E. Cassiraga, X. Llort, A. Seed and G. Lee for very stimulating discussions. Thanks to Ufficio dei corsi d'acqua (Canton Ticino) and Istituto Scienze della Terra (SUPSI) for additional rain-gauge data. This study is part of COST-731 and MAP D-PHASE, and was funded by MeteoSwiss and Universitad Politècnica de Catalunya.

References

Austin PM. 1987. Relation between measured radar reflectivity and surface rainfall. *Mon. Weather Rev.* **115**: 1053–1070.

Bellerby TJ, Sun J. 2007. Probabilistic and ensemble representations of the uncertainty in an IR/microwave satellite precipitation product. *J. Hydrometeorol.* in press.

Bellon A, Lee G, Zawadzki I. 2005. Error statistics of VPR corrections in stratiform precipitation. *J. Appl. Meteorol.* **44**: 998–1015.

Bellon A, Lee G, Kilambi A, Zawadzki I. 2007. Real-time comparisons of VPR-corrected daily rainfall estimates with a gauge mesonet. *J. Appl. Meteorol.* **46**: 726–741.

Berenguer M, Zawadzki I. 2008. A study on the error covariance matrix of radar rainfall estimates in stratiform rain. *Weather Forecasting* in press.

Berenguer M, Corral C, Sánchez-Diezma R, Sempere-Torres D. 2005. Hydrological validation of a radar-based nowcasting technique. *J. Hydrometeorol.* **6**: 532–549.

Berenguer M, Sempere-Torres D, Sánchez-Diezma R, Pegram G, Zawadzki I, Seed A. 2006.. 'Modelization of the uncertainty associated with radar-based nowcasting techniques. Impact on flow simulation'. Pp 575–578 in Proc. 4th European Conf. on Radar in Meteorology and Hydrology, 18–22 September 2006, Barcelona, Spain.

Bowler NE, Pierce CE, Seed AW. 2006. STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Q. J. R. Meteorol. Soc.* **132**: 2127–2155.

Ciach GJ, Krajewski WF, Villarini G. 2007. Product-error-driven uncertainty model for probabilistic quantitative precipitation estimation with NEXRAD data. *J. Hydrometeorol.* **8**: 1325–1347.

Clark MP, Slater AG. 2006. Probabilistic quantitative precipitation estimation in complex terrain. *J. Hydrometeorol.* **7**: 3–22.

Cressie NAC. 1993. *Statistics for spatial data*. *Wiley Series in Probability and Mathematical Statistics*. John Wiley & Sons.

Delrieu G, Andrieu H, Creutin JD. 2000. Quantification of path-integrated attenuation for X- and C-band weather radar systems operating in Mediterranean heavy rainfall. *J. Appl. Meteorol.* **39**: 840–850.

Friedrich K, Germann U, Gourley JJ, Tabary P, 2007. Effects of radar beam shielding on rainfall estimation for polarimetric C-band radar. *J. Atmos. Oceanic Technol.* **24**: 1839–1859.

Gabella M, Joss J, Perona G, Galli G. 2001. Accuracy of rainfall estimates by two radars in the same Alpine environment using gage adjustment. *J. Geophys. Res.* **106**: 5139–5150.

Germann U. 1999. Radome attenuation – A serious limiting factor for quantitative radar measurements? *Meteorol. Z.* **8**: 85–90.

Germann U. 2000. 'Spatial continuity of precipitation, profiles of radar reflectivity and precipitation measurements in the Alps'. PhD thesis No. 13 932, Swiss Federal Institute of Technology (ETH): Zurich. Available at http://e-collection.ethbib.ethz.ch.

Germann U, Joss J. 2002. Mesobeta profiles to extrapolate radar precipitation measurements above the Alps to the ground level. *J. Appl. Meteorol.* **41**: 542–557.

Germann U, Zawadzki I. 2004. Scale-dependence of the predictability of precipitation from continental radar images. Part II: Probability forecasts. *J. Appl. Meteorol.* **43**: 74–89.

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

Germann U, Berenguer M, Sempere-Torres D, Salvadè G. 2006a. 'Ensemble radar precipitation estimation – A new topic on the radar

horizon'. Pp 559–562 in Proc. 4th European Conf. on Radar in Meteorology and Hydrology, 18–22 September 2006, Barcelona, Spain.
Germann U, Galli G, Boscacci M, Bolliger M. 2006b. Radar precipitation measurement in a mountainous region. *Q. J. R. Meteorol. Soc.* **132**: 1669–1692.
Gill PE, Murray W, Wright MH. 1981. *Practical optimization*. Academic Press.
Goovaerts P. 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press.
Holleman I, Michelson D, Galli G, Germann U, Peura M, Hohti H. 2006. 'Quality information for radars and radar data'. OPERA 2 deliverable OPERA_2005_19, EUMETNET: Brussels. http://www.knmi.nl/opera/opera2/OPERA_2005_19_DataQuality.pdf.
Illingworth A. 2004. Improved precipitation rates and data quality by using polarimetric measurements. Pp 130–166 in *Weather Radar: Principles and Advanced Applications*, Meischner P. (ed.) Vol. XVII of *Physics of Earth and Space Environment*, Springer-Verlag: Berlin.
Johnson ME. 1987. *Multivariate Statistical Simulation*. John Wiley & Sons.
Joss J, Gori EG. 1978. Shapes of raindrop size distributions. *J. Appl. Meteorol.* **17**: 1054–1061.
Joss J, Lee R. 1995. The application of radar-gauge comparisons to operational precipitation profile corrections. *J. Appl. Meteorol.* **34**: 2612–2630.
Joss J, Waldvogel A. 1990. Precipitation measurement and hydrology. Pp 577–597 in *Radar in Meteorology: Battan Memorial and 40th Anniversary Radar Meteorology Conference*. Amer. Meteorol. Soc: Boston.
Kitchen M. 1995. 'Estimation of surface precipitation rate from radar using a variational method'. Pp 228–238 in *COST-75 Weather Radar Systems – International Seminar, Brussels 1994* vol. EUR 16013 EN.
Koistinen J, Michelson DB, Hohti H, Peura M. 2004. Operational measurement of precipitation in cold climates. Pp 78–114 in *Weather Radar: Principles and Advanced Applications* Meischner P. (ed.) Vol. XVII of *Physics of Earth and Space Environment*. Springer-Verlag: Berlin.
Krajewski WF, Georgakakos KP. 1985. Synthesis of radar rainfall data. *Water Resour. Res.* **21**: 764–768.
Lee G-W, Zawadzki I. 2005. Variability of drop size distributions: Time-scale dependence of the variability and its effects on rain estimation. *J. Appl. Meteorol.* **44**: 241–255.
Lee G, Seed AW, Zawadzki I. 2007. Modeling the variability of drop size distributions in space and time. *J. Appl. Meteorol.* **46**: 742–756.
Mittermaier MP, Hogan RJ, Illingworth AJ. 2004. Using mesoscale winds for correcting wind-drift errors in radar estimates of surface rainfall. *Q. J. R. Meteorol. Soc.* **130**: 2105–2123.
Palmer TN. 2002. Predicting uncertainty in numerical weather forecasts. Pp 3–13 in *Meteorology at the Millenium*, Pearce RP. (ed.) Vol. 83 of *International Geophysics Series*. Academic Press.
Pellarin T, Delrieu G, Saulnier G-M, Andrieu H, Vignal B, Creutin J-D. 2002. Hydrologic visibility of weather radar systems operating in mountainous regions: Case study for the Ardèche catchment (France). *J. Hydrometeorol.* **3**: 539–555.
Press WH, Flannery BP, Teukolsky SA, Vetterling WT. 1992. *Numerical Recipes in C – The Art of Scientific Computing*. Cambridge University Press.
Priestley MB. 1981. *Spectral Analysis and Time Series*. Academic Press.
Ranzi R, Zappa M, Bacchi B. 2007. Hydrological aspects of the Mesoscale Alpine Programme: Findings from field experiments and simulations. *Q. J. R. Meteorol. Soc.* **133**: 867–880.
Ripley BD. 1987. *Stochastic simulation. Wiley series in probability and mathematical statistics*. John Wiley & Sons.
Seed AW. 2003. A dynamic and spatial scaling approach to advection forecasting. *J. Appl. Meteorol.* **42**: 381–388.
Seed AW, Srikanthan R, Menabde M. 1999. A space and time model for design storm rainfall. *J. Geophys. Res.* **104**: 31623–31630.
Sempere-Torres D, Porra JM, Creutin JD. 1994. A general formulation for rain drop size distribution. *J. Appl. Meteorol.* **33**: 1494–1502.
Viviroli D, Gurtz J, Zappa M. 2007. 'The Hydrological Modelling System PREVAH. Model documentation and user manual'. Geographica Bernensia P40, Department of Geography, University of Berne, Switzerland.
Wesson SM, Pegram GGS. 2004. Radar rainfall image repair techniques. *Hydrol. Earth Syst. Sci.* **8**: 220–234.
Wesson SM, Pegram GGS. 2006. Improved radar rainfall estimation at ground level. *Nat. Hazards Earth Syst. Sci.* **6**: 1–20.
Wöhling T, Lennartz F, Zappa M. 2006. Technical Note: Updating procedure for flood forecasting with conceptual HBV-Type Models. *Hydrol. Earth Syst. Sci.* **10**: 783–788. http://www.hydrol-earth-syst-sci.net/10/783/2006/hess-10-783-2006.html.
Yuter SE. 2002. Precipitation radar. Pp 1833–1852 in *Encycl. Atmos. Sci.* Holton JR, Curry JA, Pyle JA. (eds.) Academic Press: London.
Zappa M, Pos F, Strasser U, Warmerdam P, Gurtz J. 2003. Seasonal water balance of an Alpine catchment as evaluated by different methods for spatially distributed snowmelt modelling. *Nordic Hydrol.* **34**: 179–202.
Zappa M, Rotach M, Arpagaus M, Dorninger M, Hegg C, Montani A, Ranzi R, Ament F, Germann U, Grossi G, Jaun S, Rossa A, Vogt S, Walser A, Wunram C. 2008. MAP D-PHASE: Real-time demonstration of hydrological ensemble prediction systems. *Atmos. Sci. Lett.* **9**: 80–87.
Zawadzki I. 1982. The quantitative interpretation of weather radar measurements. *Atmos.–Ocean* **20**: 158–180.

DOI: 10.1002/qj