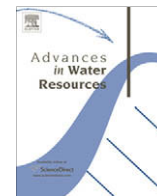




Contents lists available at ScienceDirect

Advances in Water Resources

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

A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data

Carlos A. Velasco-Forero ^{a,*}, Daniel Sempere-Torres ^a, Eduardo F. Cassiraga ^b, J. Jaime Gómez-Hernández ^b

^a Grup de Recerca Aplicada en Hidrometeorologia, ETSECCPB, Universitat Politècnica de Catalunya, C/Gran Capità, 2-4, Edifici NEXUS 102-106, 08034 Barcelona, Spain

^b Departamento de Ingeniería Hidráulica y Medio Ambiente, Universidad Politécnica de Valencia, C/Camino de Vera, s/n. 46071 Valencia, Spain

ARTICLE INFO

Article history:

Received 10 February 2008

Received in revised form 8 October 2008

Accepted 12 October 2008

Available online xxxx

Keywords:

Radar rainfall

Rainfall estimation

Rain gauges

Merging

Blending

Geostatistics

ABSTRACT

Quantitative estimation of rainfall fields has been a crucial objective from early studies of the hydrological applications of weather radar. Previous studies have suggested that flow estimations are improved when radar and rain gauge data are combined to estimate input rainfall fields. This paper reports new research carried out in this field. Classical approaches for the selection and fitting of a theoretical correlogram (or semivariogram) model (needed to apply geostatistical estimators) are avoided in this study. Instead, a non-parametric technique based on FFT is used to obtain two-dimensional positive-definite correlograms directly from radar observations, dealing with both the natural anisotropy and the temporal variation of the spatial structure of the rainfall in the estimated fields. Because these correlation maps can be automatically obtained at each time step of a given rainfall event, this technique might easily be used in operational (real-time) applications. This paper describes the development of the non-parametric estimator exploiting the advantages of FFT for the automatic computation of correlograms and provides examples of its application on a case study using six rainfall events. This methodology is applied to three different alternatives to incorporate the radar information (as a secondary variable), and a comparison of performances is provided. In particular, their ability to reproduce in estimated rainfall fields (i) the rain gauge observations (in a cross-validation analysis) and (ii) the spatial patterns of radar fields are analyzed. Results seem to indicate that the methodology of kriging with external drift [KED], in combination with the technique of automatically computing 2-D spatial correlograms, provides merged rainfall fields with good agreement with rain gauges and with the most accurate approach to the spatial tendencies observed in the radar rainfall fields, when compared with other alternatives analyzed.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Rainfall is one of the most important inputs in hydrological models. Rain gauges and weather radars are probably the two sensors that are most widely used in rainfall measurement. Unfortunately, typical densities of operational rain gauge networks are usually unable to fulfill the requirements for real-time hydrological modelling. Weather radar helps to improve our knowledge of rainfall fields, but radar rainfall fields are affected by systematic and local errors that should be corrected, in order to use these fields directly for hydrological purposes. Previous works have suggested that radar data are essential information in providing accurate flow estimates using a rainfall runoff model, even when a dense rain gauge network exists [1–3]. These results have led to the development of diverse methodologies for estimating rainfall fields by merging radar and rain gauge data.

These merging techniques range from the simplest formulation, i.e. finding a constant multiplicative calibration factor [4–6], to statistical approaches based on multivariate analysis [7,8], radar–rain gauge probability distribution analysis [9–12], geostatistical estimators [2,13–18], and Bayesian methods (see e.g. [19]).

The aim of this study is to define a methodology to estimate rainfall fields that fit locally with direct (but scarce) rain gauge measurements and spatially with the shapes of highly detailed (but indirect) radar fields in a real-time combination of radar and rain gauge data using geostatistical approaches.

The methodology proposed here introduces two main advances over classical approaches.

- (i) The use of a non-parametric automatic procedure for estimating valid spatial variability models, firstly proposed in [20], and applied for the first time to rainfall estimation by the authors [21,22] and fully described and developed in this paper. This automatic procedure uses FFT for estimating valid spatial variability models for rainfall fields that can

* Corresponding author. Tel.: +34 934 017 371; fax: +34 934 054 194.

E-mail address: velasco@grahi.upc.edu (C.A. Velasco-Forero).

be used to estimate rainfall fields merging radar and rain gauge data. The main advantages of this automatic technique are that:

- o it does not require the prior selection of an analytical model (i.e. covariance values do not need to be related by any analytical formulae and there are no parameters that must be adjusted),
 - o it includes the observed anisotropy of rainfall fields in the estimation because it is possible to compute two- or three-dimensional correlation maps,
 - o the spatial variability model can be changed for each time step of an event because this technique provides fast, automatic definitions and
 - o it guarantees a single, valid solution for the kriging equations system by definition.
- (ii) The second advance is the use of the radar data field to infer the spatial structure of the rainfall field. Thus, the high spatial and temporal resolutions of radar data are used to compute the correlogram (or variogram if desired), allowing a better definition of the high variability in space and time of the rainfall.

The incorporation of these two new advances into geostatistical techniques for blending radar and gauge data is the main originality of the methodology proposed here.

The goal of the paper is to show how this methodology can be used to improve selected geostatistical estimators of rainfall fields and to compare the performances of three alternative approaches that use the radar field data as supporting information. All three alternatives use rain gauge data to estimate rainfall fields, but supplementing these estimates with radar data in different ways: (1) using radar fields only to define the spatial variability models of rainfall (i.e., an ordinary kriging estimator), (2) using radar fields as in (1) but also driving rainfall estimation using the observed trends in radar fields (i.e., via kriging with external drift), and (3) using radar fields as in (1) but also including the radar observations at the target point in the estimation equations (i.e., via collocated cokriging). It is noted that in all these alternatives radar fields are used to infer the spatial variability model (as correlation maps) of true rainfall through the non-parametric automatic methodology based on FFT introduced in this paper. Finally, a fourth (reference) technique is used in the comparisons: classical ordinary kriging using only information from rain gauges (i.e., with correlation maps obtained directly from the rain gauge data), which serve to assess the improvements introduced by the use of radar data. The performance of each one of these estimators was analyzed, compared and discussed using several events observed in Catalonia (Spain).

The description of the automatic modelling technique used to automatically define spatial variability models for rainfall fields is presented in Section 2. A brief description of each blending alternative, as well as some examples showing how they are implemented, is included in Section 3. The case study is described in Section 4. Finally, Section 5 describes the analysis and the comparisons and discusses the results and performance of the proposed methodology and its different alternatives.

2. Automatic modelling of the spatial variability of rainfall

The traditional (parametric) modelling approach for determining valid spatial models (to be used in kriging) considers only positive, linear combinations of basic covariance or semivariogram models known to be positive definite under very restrictive conditions. All previous radar–rain gauge (co-)kriging applications have used this parametric modelling of the spatial covariance of exper-

imental radar and rain gauge data. Each author defined different spatial models appropriate to the specific data used. Krajewski [13] worked with simulated fields and defined an exponential covariance model for a radar–rain gauge cokriging system. Creutin et al. [14] defined three different variograms from radar and rain gauge data for each analyzed day. Seo [2] employed isotropic exponential models with a nugget effect to model both indicator and conditional correlation functions in his estimations of real-time rainfall fields. Sun et al. [3] worked on an Australian case study and defined a Gaussian model for both auto- and cross-covariance from radar and gauge data. Nicolau et al. [23] mapped the spatial distribution of rainfall in Portugal, interpolating rain gauge data using a spherical semivariogram model. Seo et al. [16] merged radar and rain gauge data by cokriging using a least squares criterion to find the best-fit empirical combination for defining the semivariogram model, choosing between spherical, exponential and Gaussian models. Germann and Joss [24] reported the effects of precipitation type on the spatial patterns of rainfall in various Swiss Alpine regions using semivariograms. Berne et al. [25] used spherical semivariograms to analyze the temporal and spatial correlation of Mediterranean rainfall events. Due to this diversity of results, and the difficulty of adjusting the covariance (correlation) models, it has not been possible to establish, until now, a reference methodology for the operational combination of radar and rain gauge data.

As an alternative, some authors have proposed non-parametric methodologies for finding valid spatial covariance (correlation) models [26,27]. Generally, non-parametric techniques look for covariance values that do not need to be related by any analytical formulae via numerical approaches. Nevertheless, all values must be consistent with each other and provide a good fit for the corresponding experimental covariance. The principal advantages of non-parametric techniques are that they: do not require the prior selection of an analytical model; free the modelling of coregionalization models when several variables are involved in the estimation process; and guarantee a single, valid solution for the kriging equations system. To our knowledge, there are no previous studies in which these non-parametric techniques were applied to the estimation of rainfall fields with a combination of radar and rain gauge data. Only preliminary results have previously been reported by the authors [21,22,28]. In this paper, we propose using a non-parametric technique based on FFT firstly proposed by Yao and Journel [20] (from now as YJ) for the fast and automatic definition of valid, two-dimensional correlograms from radar or rain gauge data. With this methodology it is possible to progressively compute valid, two-dimensional spatial correlograms for each time step in operational applications of the radar–rain gauge merging process.

The technique proposed by YJ consists of transforming the experimental (cross-)covariance values into spectral density maps in the frequency domain using FFT. Following Bochner's theorem [29], the spectral density maps are smoothed in the frequency domain under positivity constraints. Then, the inverse FFT transform of the smoothed spectral density provides (jointly) positive-definite (cross-)covariance matrices. This technique can be applied to define intrinsic cross-covariance models and to find valid variability models in one, two, or three dimensions.

YJ's algorithm proceeds as follows:

1. Calculate the experimental covariance (correlation) maps from the sample data. This original sample covariance matrix is generally not positive definite.
2. Because not all covariance values are experimentally available, the sample covariance values must be interpolated to fill it up. This preliminary smoothing of the experimental covariance maps fills in all missing entries of the covariance matrix and filters

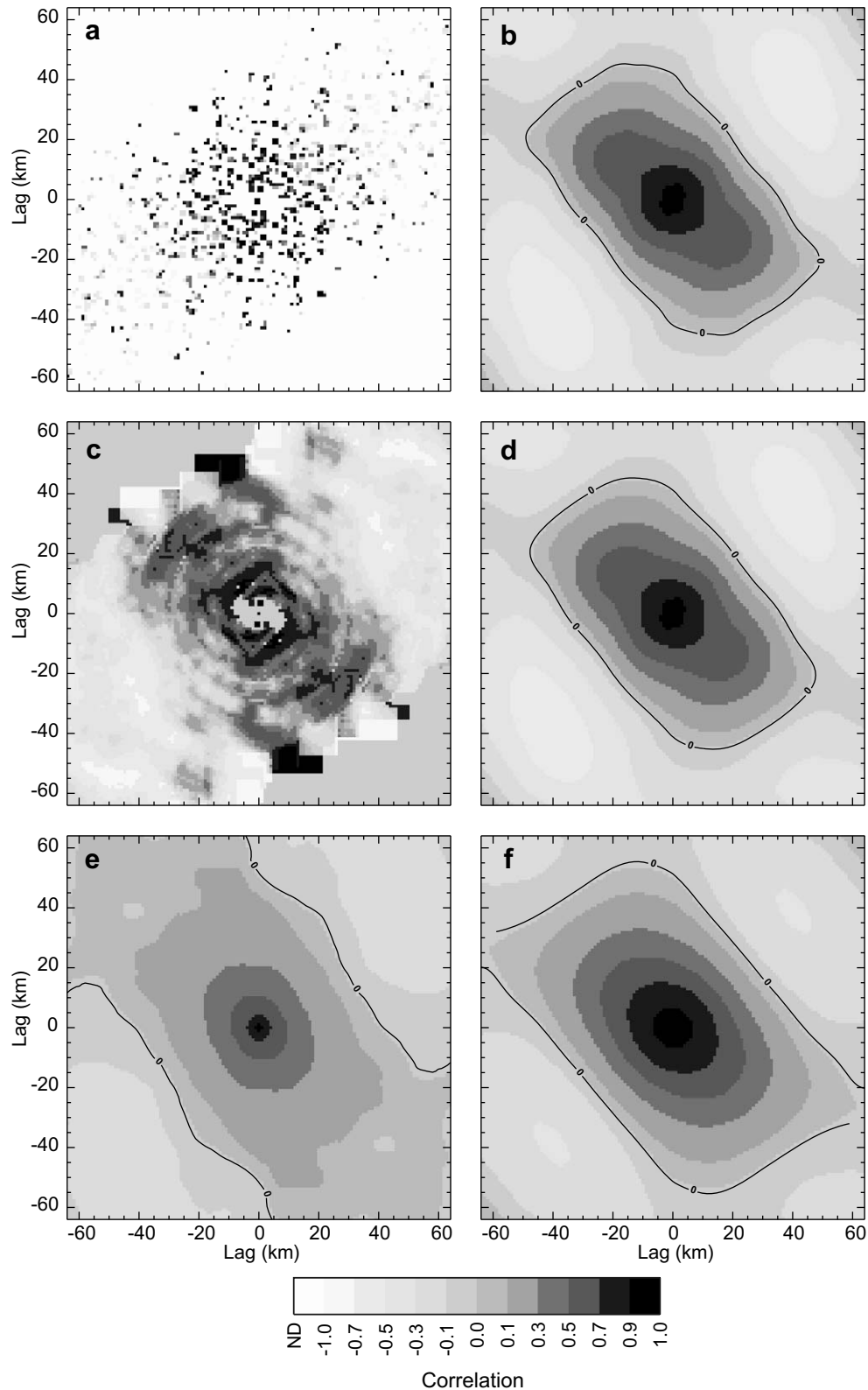


Fig. 1. Steps of the automatic methodology for computing 2-D spatial correlation maps. First row shows experimental correlation maps from gauge data (a) and from radar rainfall field (b). Second row shows correlation maps after a radial smoothing schema fulfilled all missing values in the experimental maps of first row (Step 2 of YJ's algorithm). Third row shows the 2-D valid spatial correlation maps obtained after the smoothing process in the Fourier domain (last steps of YJ's algorithm). Left column corresponds with correlation maps computed using gauge data ((a), (c), and (e)), and right column using a radar rainfall field ((b), (d) and (f)). Only central (127×127) sectors of correlograms are showed in this figure. Bold lines indicate the null correlation level in each case. Rain gauge and radar data used to compute these correlation maps were extracted from event number 5 at 1300 UTC 9 of October 2002, and can be seen in Fig. 8a and b.

the main data fluctuations. YJ proposed to weight the original values via a radial moving window scheme, although others

could be used. The weight associated with each original value is proportional to the number of data pairs used in its derivation

and decreases with the distance of the value from centre of the window. This interpolation method may produce a covariance matrix that does not meet the positive definiteness condition, so checks and corrections are applied as follows.

3. The correction for the positive definiteness condition of the experimental covariance matrix is done in the spectral domain. The transformation between the spatial and frequency domains is done using FFT. With FFT, an experimental covariance matrix $C(\mathbf{h})$ can be transformed into a gridded spectrum matrix $s(\mathbf{w})$, where $\mathbf{h}$ is the distance vector in the d -dimension space R^d and $\mathbf{w}$ is the frequency vector in the corresponding frequency domain.
4. Now, in the frequency domain, the only conditions are that: the experimental density spectrum table must be greater than zero and sum to a unitary value. Bochner [29] states that a function $C(\mathbf{h})$ is positive definite, if and only if it can be expressed as the Fourier transform of a bounded nondecreasing positive measure $S(\mathbf{w})$

$$C(\mathbf{h}) = \int_{\mathbb{R}^d} e^{2\pi i \mathbf{h} \cdot \mathbf{w}} dS(\mathbf{w}). \quad (1)$$

If the stationary random function model $Z(\mathbf{u})$ has unit variance (where $\mathbf{u}$ is the position vector in the function space), i.e. $C(\mathbf{0}) = \text{Var}\{Z(\mathbf{u})\} = 1$, then $S(\mathbf{w})$ can be seen as a cumulative distribution function (cdf), with $dS(\mathbf{w}) = s(\mathbf{w})d\mathbf{w}$. Bochner's theorem then requires that the density $s(\mathbf{w}) = dS(\mathbf{w})/d\mathbf{w}$ be a probability density function (pdf), i.e. such that

$$s(\mathbf{w}) \geq 0 \quad \forall \mathbf{w}, \quad \int_{\mathbb{R}^d} s(\mathbf{w})d\mathbf{w} = 1. \quad (2)$$

Therefore, the original spectrum values are smoothed under these last two constraints to obtain a valid (YJ call 'licit') density spectrum map, following Bochner's theorem.

5. That licit and smooth spectral density is then inverse transformed by FFT into a licit positive definite and smooth correlation map. This correlation map can be used for kriging estimation or simulation in the spatial domain instead of selecting an analytical model.

This technique does not require the prior selection of an analytical model; the transformation between the spatial and frequency domains is performed very quickly using FFT and the constraints for positive definiteness in the frequency domain are easy to implement. In summary, the final matrix of covariance values is thus built from a FFT "roundtrip" with multiple intermediate smoothing processes. Also, it is important to note that in the case of using radar rainfall fields as is proposed here, the experimental correlation maps of exhaustive sampled variables can be computed quickly using FFT by applying the Wiener–Khinchin theorem, so the first two steps of YJ's technique are unnecessary in these cases. A full description of this algorithm can be found in [20,26] and details on the computational implementation in [30].

In order to illustrate how YJ's algorithm is applied to estimate valid correlograms, Fig. 1 shows an example of the application of some steps of this algorithm to two experimental 2-D spatial correlation fields, one computed from gauge data (left column) and the other from the contemporaneous radar rainfall field (right column). The rain gauge and radar rainfall data used to create the correlation images shown in Fig. 1 were extracted from event number 5 at 1300 UTC of 9 October 2002, and can be seen in Fig. 8a and b, respectively.

Fields in the top row of Fig. 1 correspond with sample correlation maps (Step 1 of YJ's algorithm). The experimental spatial correlation map of rain gauge data (Fig. 1a) has a sparse set of available correlation values because gauge pairs are not located

at all separation distances (lags). Therefore the main tendencies of the spatial correlation of rainfall cannot be easily identified from this field. On the other hand, the sample spatial correlation map calculated from the radar rainfall field (Fig. 1b) shows correlation values for almost all lags due to the availability in the radar field of more points with observed rainfall values than the rain gauge network. An anisotropic distribution of the field of correlation values of rainfall can be observed here, with the main correlation trend in the NW–SE direction.

In order to infill the missing values (especially for the gauges), the second step of YJ's algorithm (radial smoothing) was applied to each one of these sample correlation fields obtaining the images shown in the second row of Fig. 1. The "smoothed" radar correlation map (Fig. 1d) shows no important changes because just a few values were initially missing in this case (Fig. 1b). However, the changes in the correlation map of rain gauge data (Fig. 1c) are quite remarkable. At this stage, the main trend of this new smoothed field can be inferred (it seems similar to the radar's one), but strong discontinuities remain after this step.

Finally, last three steps of YJ's algorithm were applied to each one of the smoothed fields, and valid spatial correlation fields were obtained in each case (last row of Fig. 1). A valid correlation field obtained from rain gauge data (Fig. 1e) now shows clear tendencies: (1) a high correlation region at shorter distances and (2) a decay of the correlation values with matching distances and orientations to those observed in the correlation map of the radar field (Fig. 1f). The main shape of the valid spatial correlation field of the radar data (Fig. 1f) is the same as the previous ones (Fig. 1a and c), but an increment in the correlation length in all directions can be observed after the smoothing process in the Fourier domain. However, the null correlation zones are located almost at the same distances and directions in both final correlation fields.

These results demonstrate the ability of the proposed technique to estimate valid spatial correlation fields from gauge data or radar rainfall fields, and illustrate how the high spatial definition of radar data can assist the estimation of more detailed correlation fields of rainfall than any rain gauge network. However, it is important to note that the estimation of the spatial correlation of rainfall from radar data will be accurate only if an adequate chain of corrections is applied to the raw radar data. Any uncorrected error (e.g., anomalous propagation, attenuation or sample regions with very low reflectivity values associated with noise) in the radar rainfall field will affect the quality of the estimation of the spatial correlation field.

3. Blending alternatives

One of the objectives of the paper is to show how the new methodology to automatically obtain correlation maps can be (i) incorporated into some geostatistical estimators of rainfall fields, and then (ii) to identify the best approach to introduce the radar information as supporting data. Thus, three different geostatistical alternatives were selected to estimate rainfall fields using rain gauge data but dealing with radar data in different ways. Details of the different kriging techniques used here can be found in the literature (see, e.g. [31–34]). However, in order to make clear which kriging estimators were used and how they were implemented in this study, descriptions of each one of them are outlined in the next subsections.

3.1. Ordinary kriging

Ordinary kriging (OK) is widely used to estimate rainfall fields using only one variable (in our case the rain gauge data) and is a well-established reference to assess the improvements that more

advanced techniques might provide. The OK estimator is a weighted linear estimator widely described in the literature, which guarantees that the weights defined for a target location, in which the value of rainfall is unknown, produce the best linear unbiased estimation (BLUE) in a least squares sense.

We take $\{x_1, x_2, \dots, x_n\}$ as n locations (in two dimensions) that have known rainfall values $\{Z_G(x_1), Z_G(x_2), \dots, Z_G(x_n)\}$ and x_0 as the location at which the rainfall value is unknown. The value of rainfall $Z_G(x_0)$, at the target point x_0 , is then estimated by OK as a weighted linear combination of the n observations through

$$Z_{OK}^*(x_0) = \sum_{i=1}^n \lambda_i^{OK} Z_G(x_i), \quad (3)$$

where the symbol $*$ indicates that the value is only an estimation of the real value and λ_i^{OK} are the weights associated with the estimation of the rainfall value at point x_0 . If the target location changes, so will the weights. Optimal weights are obtained by solving the following system of $(n+1)$ linear equations:

$$\begin{aligned} \sum_{j=1}^n \lambda_j^{OK} C(x_i, x_j) + \mu &= C(x_i, x_0), \\ \sum_{j=1}^n \lambda_j^{OK} &= 1, \end{aligned} \quad (4)$$

where the last equation corresponding to the null bias constraint of the estimation, $C(x_i, x_j)$ is the covariance value between points x_i and x_j and μ is an auxiliary Lagrange multiplier.

If $\tilde{\sigma}^2$ is the (field) variance of the observed values, the estimation (kriging) error variance, σ_{OK}^2 , can be computed by

$$\sigma_{OK}^2 = \tilde{\sigma}^2 - \left(\sum_{i=1}^n \lambda_i^{OK} C(x_i, x_0) + \mu \right), \quad (5)$$

which is zero when the target point is located at one of the observations.

In this study we will use the OK estimator in two ways, with different estimates of the variogram:

- To determine the reference field, with the rain gauges observations used (without assistance from radar) to obtain the spatial correlation model of rainfall (from now we call this OK_{RG}). This reference field will serve to assess the improvements introduced by the use of radar data.
- The second way of using ordinary kriging allows us to include the full spatial variability of the rainfall field measured by the radar in the OK estimates (from now OK_{RD}). In this OK_{RD} method, rain gauge data are only used as conditional values on the resulting field; they are interpolated using the spatial correlogram of the radar fields. This is the simplest one of the three blending alternatives proposed for the application of the new methodology to define a spatial variability model of rainfall, since the information of the radar is only introduced through the inference of the spatial correlogram.

3.2. Kriging with external drift

If extensive auxiliary information (e.g., radar, satellite or topographical data) is available and correlated with a target variable (e.g., rain gauge data), kriging with external drift (from now KED) is a simple method for including an external variable in the estimation process. In this paper, KED allows us to drive the rainfall estimations with the observed trends in radar fields. KED assumes that estimations are modelled as a drift term plus a residual term. The drift term is an unknown linear function defined externally

through an auxiliary variable. Previous works have estimated rainfall fields using KED merging radar and rain gauges (see e.g. [28,35]) or rain gauge and satellite data [36]. A full description of KED equations can be found in the literature [27,33,37].

If only one secondary variable $Z_2(x)$ is available, KED assumes that the mean of the principal variable is linearly related to the tendencies of the secondary variable, such that

$$E(Z_1(x)) = m_1(x) = a_0 + a_1 \cdot Z_2(x), \quad (6)$$

where $m_1(x)$ is the mean value of the principal variable at the point x , $Z_2(x)$ is the value of the secondary variable at the same location, and a_0 and a_1 are unknown coefficients (which do not need to be estimated). Rainfall estimations using KED (Eq. (7)) are computed using the same expression as in the OK estimator (Eq. (3))

$$Z_{KED}^*(x_0) = \sum_{i=1}^n \lambda_i^{KED} Z_G(x_i). \quad (7)$$

However, KED weights, λ_i^{KED} , are unlike those used in the OK technique because the equation system of the KED estimator (Eq. (8)) has an additional constraint to satisfy the new drift hypothesis and there is a second auxiliary Lagrange multiplier, μ_1

$$\begin{aligned} \sum_{j=1}^n \lambda_j^{KED} C_R(x_i, x_j) + \mu_0 + \mu_1 Z_R(x_i) &= C_R(x_i, x_0), \quad i = 1, 2, \dots, n \\ \sum_{j=1}^n \lambda_j^{KED} &= 1, \\ \sum_{j=1}^n \lambda_j^{KED} Z_R(x_j) &= Z_R(x_0) \end{aligned} \quad (8)$$

where $Z_G(\cdot)$ and $Z_R(\cdot)$ are respectively gauge and radar observations at a given location and C_R is the covariance of the residuals $Z_G(x) - m_G(x)$, where $m_G(x)$ is the gauge drift field.

The final constraint in the equation system (Eq. (8)) introduces the variability of the radar data into the interpolation process of the rain gauges values. This constraint means that rain gauge data are weighted using the set of weights that add in the effect of estimating the radar value at the target point as the radar values at rain gauge locations were interpolated. Theoretically, KED requires that C_R , the covariance of the residuals $Z_G(x) - m_G(x)$, to be inferred. This is not straightforward, because neither the trend (drift) nor the residuals are known a priori. In this paper the following approach is used: first, a trend field $m_G(x)$ was estimated, at all target points from the radar rainfall field by OK_{RD} and then residuals were computed by subtracting this estimated trend field from the radar rainfall field, to estimate $Z_G(x) - m_G(x)$.

To obtain the kriging variance in KED, if $\tilde{\sigma}_R^2$ is the variance of the residuals calculated as above, the estimation (kriging) error variance, σ_{KED}^2 , can be computed as an extension of Eq. (5)

$$\sigma_{KED}^2 = \tilde{\sigma}_R^2 - \left(\sum_{i=1}^n \lambda_i^{KED} C_R(x_i, x_0) + \mu_0 + \mu_1 \cdot Z_R(x_0) \right). \quad (9)$$

To summarize the methodology for obtaining estimated rainfall fields using KED, we assume that the correlogram of the radar field is used to interpolate the rainfall field and the relationship between rain gauge and radar rainfall data is linear and high. Then

- Calculate a drift map from the radar rainfall field by OK for the whole estimation area, using only the radar rainfall data at the rain gauge locations. The valid correlogram of the rainfall field for this initial estimation is computed from the complete radar rainfall field using the non-parametric modelling technique described in the previous section.
- Obtain the radar residual map by subtracting the drift map calculated in step (a) from the radar field.

- (c) Apply YJ's technique to the radar *residual* map to calculate a valid radar residual correlogram map.
- (d) The KED algorithm can be used then to estimate rainfall fields using rain gauge data as the primary variable, radar data as the secondary variable and the radar residual correlogram as the spatial variability model.

This alternative might be more computationally expensive than the OK approach because the residual map must be computed before fitting the valid spatial model and constructing the equation system. However, the results might be more precise because the full map of radar values is included in the estimation process but in as unbiased a manner as possible. The proposed approach for defining residual fields is fairly similar to the conditional merging approach [18], but in this KED application, radar residuals are only used to define the spatial correlation and only indirectly affect the rainfall field estimated from the rain gauge values via the external drift.

3.3. Collocated cokriging

Collocated cokriging (from now ColCOK) is a simplified version of cokriging (COK) in which the secondary data used for the estimation are limited to the secondary datum value collocated at each target point. This simplification is used in this study since the full cokriging estimator is computationally more expensive than the other techniques, and is not likely to be appropriate for real-time applications where there are serious temporal constraints. However, several authors (see, e.g. [13,15,17]) have shown the advantages of the COK estimator among the other geostatistical techniques in the estimation of rainfall by merging radar and gauge data, and have proposed some schemes to allow its use in operational frameworks. Full details of the ColCOK technique can be found in the literature (see e.g. [33,34]).

The ColCOK estimator equation can be defined as

$$Z_{\text{ColCOK}}^*(x_0) = \sum_{i=1}^n \lambda_i^{\text{ColCOK}} Z_G(x_i) + \lambda_2' Z_R(x_0), \quad (10)$$

where $\lambda_i^{\text{ColCOK}}$ are the weights associated with the n known rainfall values of the gauges, $Z_G(x_i)$, and λ_2' is the weight related to the known rainfall value of the radar field $Z_R(x_0)$ at the target point. In this case, there is a single constraint

$$\sum_{i=1}^n \lambda_i^{\text{ColCOK}} + \lambda_2' = 1 \quad (11)$$

and the weights of this estimator can therefore be computed by solving the following linear equation system:

$$\begin{aligned} \sum_{k=1}^n \lambda_k^{\text{ColCOK}} C_{GG}(x_i, x_k) + \lambda_2' C_{GR}(x_i, x_0) + \mu &= C_{GG}(x_i, x_0), \quad i = 1, 2, \dots, n, \\ \sum_{k=1}^n \lambda_k^{\text{ColCOK}} C_{RG}(x_0, x_k) + \lambda_2' C_{RR}(0) + \mu &= C_{RG}(0), \\ \sum_{k=1}^n \lambda_k^{\text{ColCOK}} + \lambda_2' &= 1. \end{aligned} \quad (12)$$

Thus, in ColCOK it is only necessary to know the covariance function of the gauges, $C_{GG}(h)$, and the cross-covariance function between radar and gauges, $C_{RG}(h)$. The full radar covariance $C_{RR}(h)$ is not required, only the variance $C_{RR}(0)$ at $h = 0$. Often in practice, the cross-covariance $C_{RG}(h)$ is not computed directly but instead defined as a factor of the correlogram of the gauges, for example by applying a Markov-type model as suggested in [38]

$$\rho_{GR}(h) = \rho_{GR}(0) \cdot \rho_{GG}(h), \quad (13)$$

where $\rho_{GG}(h)$ is the standardized primary covariance or correlogram of gauges $Z_G(x)$, $\rho_{GR}(h)$ is the cross-correlogram and $\rho_{GR}(0)$ is the linear correlation coefficient inferred from co-located data $Z_G(x)$, $Z_R(x)$.

As an alternative, [39,40] suggested that when the secondary variable is defined on a large support volume (as in the case of the radar–rain gauge combination) Eq. (13) should be replaced by the following model:

$$\rho_{GR}(h) = \rho_{GR}(0) \cdot \rho_{RR}(h), \quad (14)$$

where the cross-correlogram is now a factor of the correlation of the radar.

In this study the ColCOK algorithm was applied to estimate rainfall fields using rain gauge data as the primary variable, the radar data as the collocated secondary variable and the radar spatial correlation map as the spatial correlation model of the rainfall. The cross-correlogram model was also approximated from the radar spatial correlation map, using the alternative Markov approach (Eq. (14)). Using ColCOK, both rain gauge and radar data are considered in the estimation of rainfall, as in the full cokriging approach, but without the need to calculate multiple (cross-)spatial correlation models or build a large equations system. We remark that ColCOK is not a full COK approach, but a simplified approximation that could have some advantages when it is operationally applied.

4. Case study

Practical applications of the trio of merging methodologies introduced in Section 3 are analyzed, compared and discussed based on a case study in Catalonia (Spain). Catalonia is a Mediterranean region in which rain events can cause catastrophic floods in a wide range of river basin sizes. The region has been the object of different pilot studies and has been used as a Hydrometeorological Observatory in the EC projects FLOODSite and HYDRATE. Four operational weather radars cover the region and one more is planned to enter into operation during 2008. There are about 200 telemetering rain gauges belonging to the Catalan Water Agency (ACA) and The Meteorological Service of Catalonia (SMC) distributed across the whole region.

The radar data used in this study were obtained from the single C-band radar in Puig de les Agulles (near Barcelona city), operated by the Spanish Meteorology Service (AEMET). Raw radar data that correspond to the first PPI were preprocessed using the full chain of processing algorithms of the EHIMI System (Integrated Hydrometeorological Forecasting Tool; for more details, see, e.g. [41]). Thus, ground clutter, partial screening and secondary lobes were corrected, and the ASCMORE algorithm [42] was applied to avoid attenuation by rain over the radome and to provide temporal stability. The widely used Z–R relationship proposed by Marshall and Palmer [43] was applied to compute the rainfall values associated with different levels of reflectivity of radar data. Accumulated radar images were then computed assuming that precipitation fields move at a constant velocity between images and vary linearly in intensity with time between each time interval [44]. A region of 140×140 km centred on the radar was defined as a study area in an attempt to reduce the effect of errors in radar data due to distance. Spatial resolution was defined as 1×1 km² and time resolution as 60 min. The radar data intensities were thus linearly morphed and accumulated and converted to hourly rainfall depth for compatibility with rain gauge data.

Seventy-seven rain gauges of the Spanish national programme SAIH (Automatic System of Hydrological Data) are located inside of the study area (a subset of the 200 gauges in the whole Catalan

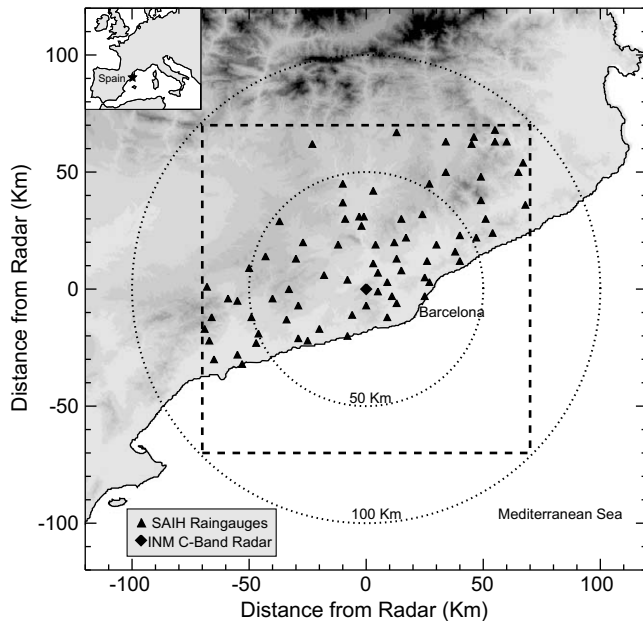


Fig. 2. Map of the Catalonia central area, showing the C-band radar location near Barcelona (diamond) operated by the Spanish Institute of Meteorology, rain gauges from the telemetered network of the Catalan Water Agency (triangles) and range rings at 50 and 100 km from radar. Dashed square delimits the 140×140 km region that was used as the study area.

region) and were used to provide precipitation data. Each observed precipitation value was centred on the pixel in which the corresponding rain gauge was located to define a precipitation grid with the same data format as the radar. If two or more rain gauges were located in the same pixel, the maximum precipitation value was accepted. Rain gauge data hourly accumulations were calculated to temporally match the radar hourly intervals. Fig. 2 shows the locations of radar, rain gauges and the square test area used for this study.

Data from six different rainfall events were used in this study and are described in chronological order. Their main characteristics are also summarized in Table 1.

The first rain event occurred on 10 of June 2000. This event was a sharp band of precipitation that crossed Catalonia, showing the importance of the topographic enhancement caused by the Montserrat Mountain. It was mainly convective during the first 15 hours but showed stratiform characteristics at the end. The data selected consist of a 22-h series in which the maximum accumulated depth measured by rain gauges was close to 225 mm.

The second case study is a rain event that occurred on 28–29 of September 2000. This 17-h event showed mainly stratiform rainfall fields but with some nuclei of embedded convection in the first half of the period. The maximum accumulated precipitation reported by rain gauges was 40 mm.

The third event was a 25-h data set recorded on 18–19 of July 2001. This case was a pure stratiform rainfall event with a maximum accumulated precipitation of less than 40 mm.

The fourth case analyzed was a 96-h rainfall event that took place between 14 and 17 of November 2001. Successive stratiform rainfall fields moving in from the Mediterranean Sea crossed the study area during the selected period, dropping a maximum accumulated precipitation of 125 mm on rain gauge locations. Coastal mountains facilitated the formation of convective nuclei over the sea in the southern part of the study area, which caused a local increase in the total accumulated precipitation.

Between 8 and 10 of October 2002, successive rain fronts, again moving in from the sea with strong nuclei of embedded convection crossed the study area, which constitutes the fifth case. This event lasted 37 hours and the maximum accumulated precipitation reported by rain gauges was 193 mm.

Finally, between 29 and 31 of March 2004, first a broad front and then several stratiform rain formations covered the study area, leaving precipitation values close to 100 mm in 96 hours.

The available rain gauge observations during each event were analyzed to eliminate possible erroneous values before combining them with radar data. Rain gauges with accumulated precipitation values equal to zero for the whole event were discarded from the analysis when the total accumulated rainfall measured by radar on the collocated pixel was greater than 10 mm. Remaining rain gauges with long periods of inactivity or with total accumulated values that were noticeably lower than radar measurements were also discarded. This simple quality control stage for rain gauge data was made manually and independently for each event analyzed; we note that quality control of gauges was made in this study after the rainfall events ended. In real-time applications, therefore, other schemes proposed in literature (e.g. the methods proposed by Amitai [45], Habib et al. [46] or Upton and Rahimi [47]) should be used to automatically identify wrong or low quality gauges values. Finally, only time steps with a total number of non-null radar–rain gauge pairs greater than or equal to 10 were included in all analyses and estimations made herein. A total of 193 hourly time steps were finally analyzed in this paper. The eventual number of rain gauges used, the number of total and selected time steps are also shown in Table 1.

5. Results

In this section, the results of the comparison of the different geostatistic estimators are presented and analyzed. First, the characteristics of the correlograms computed in this paper are commented upon. The cross-validation technique used to evaluate the quantitative performance of each kriging method for estimating rainfall values, with reference to gauged data, is then described. Finally, the spatial difference/similarity between each of the estimated rainfall fields and the original radar fields is analyzed.

Table 1

Main characteristics of the events selected for the validation of the tested blending techniques (“RG” stands for rain gauges, “Radar at RG” stands for rainfall values of radar pixels located at rain gauge positions).

Case study				Maximum accumulated rainfall (mm)			Main rainfall type
Code	Date	Time-steps available (selected)	Number RG	RG	Radar at RG	Radar Field	
1	10 June 2000	22 (22)	70	223.8	65.7	106.1	Convective
2	28 September 2000	18 (14)	75	39.7	26.9	40.8	Stratiform
3	19 July 2001	26 (10)	72	38.6	25.5	32.9	Stratiform
4	15 November 2001	96 (62)	76	125.4	115.7	226.2	Stratiform
5	08 October 2002	38 (34)	76	193.8	119.5	134.8	Convective
6	30 March 2004	73 (51)	75	96.9	42.1	67.4	Stratiform

5.1. Automatically computed correlograms

This paper introduces an automatic non-parametric technique to define 2-D positive-definite correlograms for geostatistical esti-

mators from radar fields. These correlation maps condense the main tendencies of the data employed in the estimations and guarantee the solution of the systems of linear equations for all estimators tested in this study. Thus three different experimental 2-D

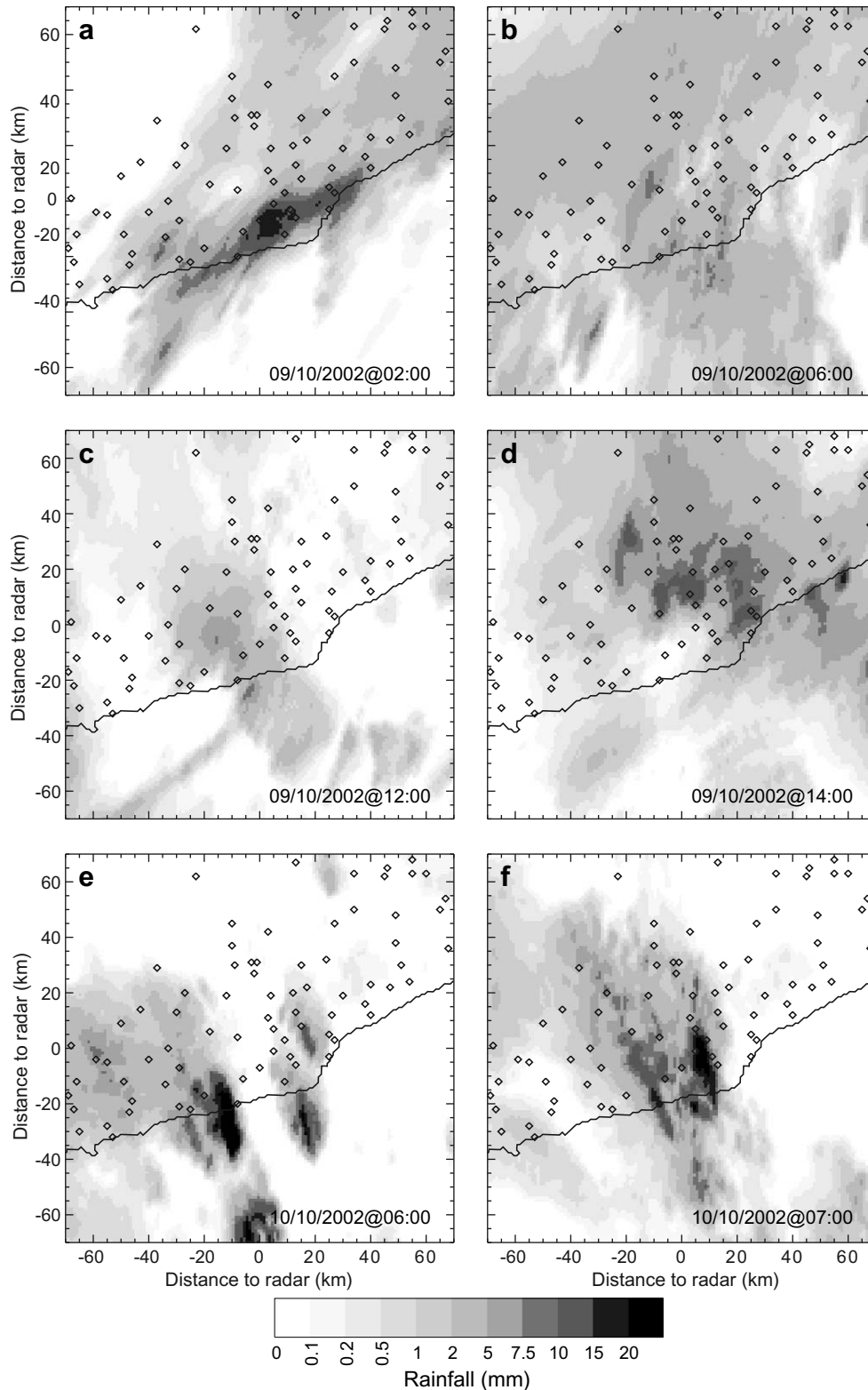


Fig. 3. Some examples of radar rainfall fields extracted from event number 5. Rainfall fields correspond to 1-h accumulation of rainfall at (a) 0200, (b) 0600, (c) 1200, (d) 1400 UTC 09 of October 2002, and (e) 0600, (f) 0700 UTC 10 of October 2002. Diamonds in rainfall field indicate the rain gauge locations. Bold lines show the Mediterranean coast near Barcelona.

correlation maps were computed for each time step of all the events: one from rain gauge data, one from the radar rainfall field and a third from the radar residuals using the drift field (see Section 3.2 on KED). Experimental correlation maps from rain gauge data were computed following the first two steps of YJ's technique

(because of the gaps in the raw correlation maps) but from radar and residual fields using FFT. Each of these sets of experimental correlation maps was then adjusted to obtain valid correlograms using the last three steps of YJ's technique in a completely automatic process. All computed correlograms have dimensions

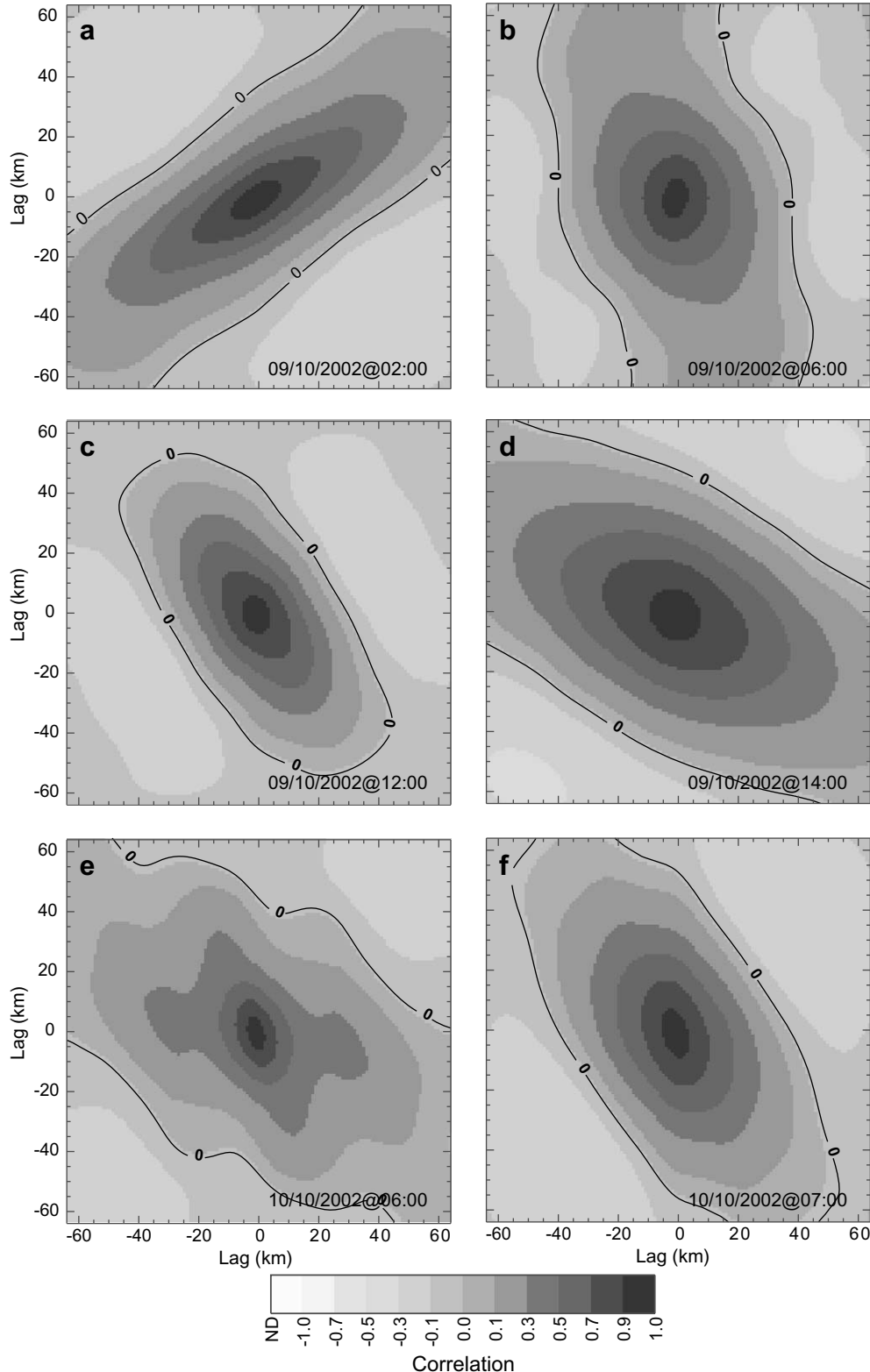


Fig. 4. Correlation maps computed automatically from radar rainfall fields shown in Fig. 3. Only central (127×127) sectors of correlograms are shown. Bold lines indicate the null correlation level in each case. Anisotropy and temporal changes in the spatial correlation of radar rainfall fields can be seen in the examples presented.

256 × 256 (i.e., horizontal and vertical lags vary from –127 km to +127 km). However, these dimensions should be appropriately defined in other cases, according with the particular size of the estimation region.

To illustrate the results obtained with this technique, several individual radar rainfall fields (see Fig. 3) have been selected from event 5 and then used to compute their respective correlation maps (see Fig. 4). The examples are selected (time separated) hourly depth accumulations during the rainfall event that took place on 8 of October 2002 (event 5), and are shown in chronological order, from left to right and from top to bottom. At the early stages this rainfall event exhibited rainfall fields (Fig. 3a) with a strong structure and correlation (Fig. 4a) in the SW–NE direction. This was followed by a period of stratiform rainfall fields (Fig. 3b–d) that slowly changed the direction of the main correlation to NW–SE and produced broad, quasi-elliptic correlograms as indicated in Fig. 4b–d. Next, some small and detached nuclei of intense precipitation appeared from the south (Fig. 3e) and the associated correlograms therefore reflect a reduced size of the main correlation area, and secondary areas of significant correlations appeared, in keeping with new spatial tendencies observed in the radar fields (Fig. 4e). Finally, rainfall fields with a small number of intense precipitation nuclei (Fig. 3f) and a NW–SE orientation crossed the study area and the correlogram maps changed again to reflect this new direction (Fig. 4f). The size of the highest correlation area increases in response to the wider spatial patterns observed during this interlude in this event.

Changes in the main shapes of correlation maps were observed for each time step of all the events studied and their evolution closely follows that of the spatial patterns observed in the radar fields. These results are not surprising since the high spatial and temporal variability of precipitation is widely known [48,49]. However these observations contrast with previous works that have generally used classical approaches based on isotropic and temporally homogeneous correlograms (variograms). Thus the capacity to infer, from the measured data, the best correlogram at any time, without imposing any model or limitation, is indeed an improvement provided by the non-parametric, automatic methodology introduced here, which is able to fully deal with anisotropy and temporal evolution of the spatial correlation of the precipitation fields.

It is worth noting that in this case study, the reduced size of the estimation region and the duration of the time step used, allow us to consider that the one correlogram can be used in the estimation of the all points inside the study area at a given time step, assuming translation invariance. By contrast, in cases with larger estimation areas and shorter time steps this assumption might not be valid, because different types of rainfall or dissimilar spatial patterns in the same time step might be observed. In these cases a more complex methodology for dealing with these problems will be required. In that case, a suggestion is that “local” correlograms associated with each different region or rain type detected in the estimation area would be identified and then the correlogram that most closely represents the particular rainfall characteristics of the point for which the estimation is required would be used, allowing that the correlogram would vary from point to point in the studied area. Even in this case, the non-parametric, automatic technique for computing correlograms introduced here would be a suitable solution for this type of problem.

5.2. Agreement between the blending estimates and the rain gauge observations

The quantitative performance of each blending methodology for estimating rainfall values measured by the rain gauges is evaluated in this section using the cross-validation technique. In this iterative

technique the rain gauge data at each site in turn is omitted from the calculations and the estimated value of rainfall at the “hidden data” location is interpolated using the remaining data (see details in, e.g. [32]). The cross-validation error, defined as the difference between the estimated and known (but not used) value, can be then computed. The process is repeated N times for each time step where N is the number of gauge sites. It is worth noting that this technique only assesses the accuracy of estimations at the location of the observed points and not in other areas; nevertheless, these statistics allow us to compare the behaviour of the different estimation procedures systematically. Cross-validation estimates are independently calculated for each case study, for each time step and for each kriging estimator considered (OK_{RG}, OK_{RD}, KED and ColCOK).

As the cross-validation errors of a good estimator have to be unbiased, i.e. the mean and median of errors should be close to zero with minimum dispersion, we use these cross-validation statistics to analyze the quality of the estimates. Thus, the mean, median, maximum, minimum and range of the individual cross-validation errors were computed. Data pairs with both values zero were removed so that they did not influence the mean or median results. The results reported in Table 2 show that all estimators are relatively unbiased (mean and median errors were close to zero in all cases) but the range of errors varies between the different techniques. For the data analyzed, the ColCOK estimator has lowest mean cross-validation errors, although the differences between methodologies are not significant and much lower than the mean errors computed just using radar values. No significant differences are observed between the standard deviations of the cross-validation errors of the different estimators, but all values are lower than the corresponding radar value. Perhaps the main difference between these results can be seen in the range of the errors, where blending techniques that used the radar data in the interpolation (KED and ColCOK, in order of better performance) have a smaller range of error than the radar and both OK approaches.

Comparison of cross-validation estimates of the three proposed blending techniques as well as OK_{RG} (as reference technique) and the radar data were made for all events using the following four quality criteria: the Nash–Sutcliffe’s efficiency (see Eq. (15)), the correlation with rain gauge data, the correlation with collocated radar data at rain gauge locations and the slope of the best-fit line of estimates with rain gauge data. Scatter plots of estimates of blending and the reference technique as well as of the radar data for the event on 19 of July 2001 (event 3) are shown in Fig. 5. However, the statistics obtained for each specific estimator and for all events are summarized in Table 3.

The Nash–Sutcliffe’s efficiency statistics [50] (from now NS) is defined in Eq. (15)

$$\text{eff} = 1 - \frac{\sum_{i=1}^n [\text{RG}_i - \text{RE}_i]^2}{\sum_{i=1}^n [\text{RG}_i - \overline{\text{RG}}]^2} \quad (15)$$

where RG_i is the rain rate measured using the reference instrument (e.g. rain gauges) at a known point i , RE_i is the rain rate

Table 2

Statistics of cross-validation errors (against rain gauge data) for radar and each blending methodology tested using the results of all time-steps of all events. The time step used was 1 h.

Estimator	Cross-validation error (mm)					
	Mean	Median	St. Dev.	Minimum	Maximum	Range
Radar	–0.38	0.00	2.18	–58.0	24.1	82.1
OK _{RG}	0.05	0.03	1.89	–49.6	27.3	76.9
OK _{RD}	0.07	0.02	2.05	–34.4	49.7	84.1
KED	0.05	0.02	1.86	–34.3	31.8	66.1
ColCOK	–0.03	0.00	1.77	–35.5	37.0	72.5

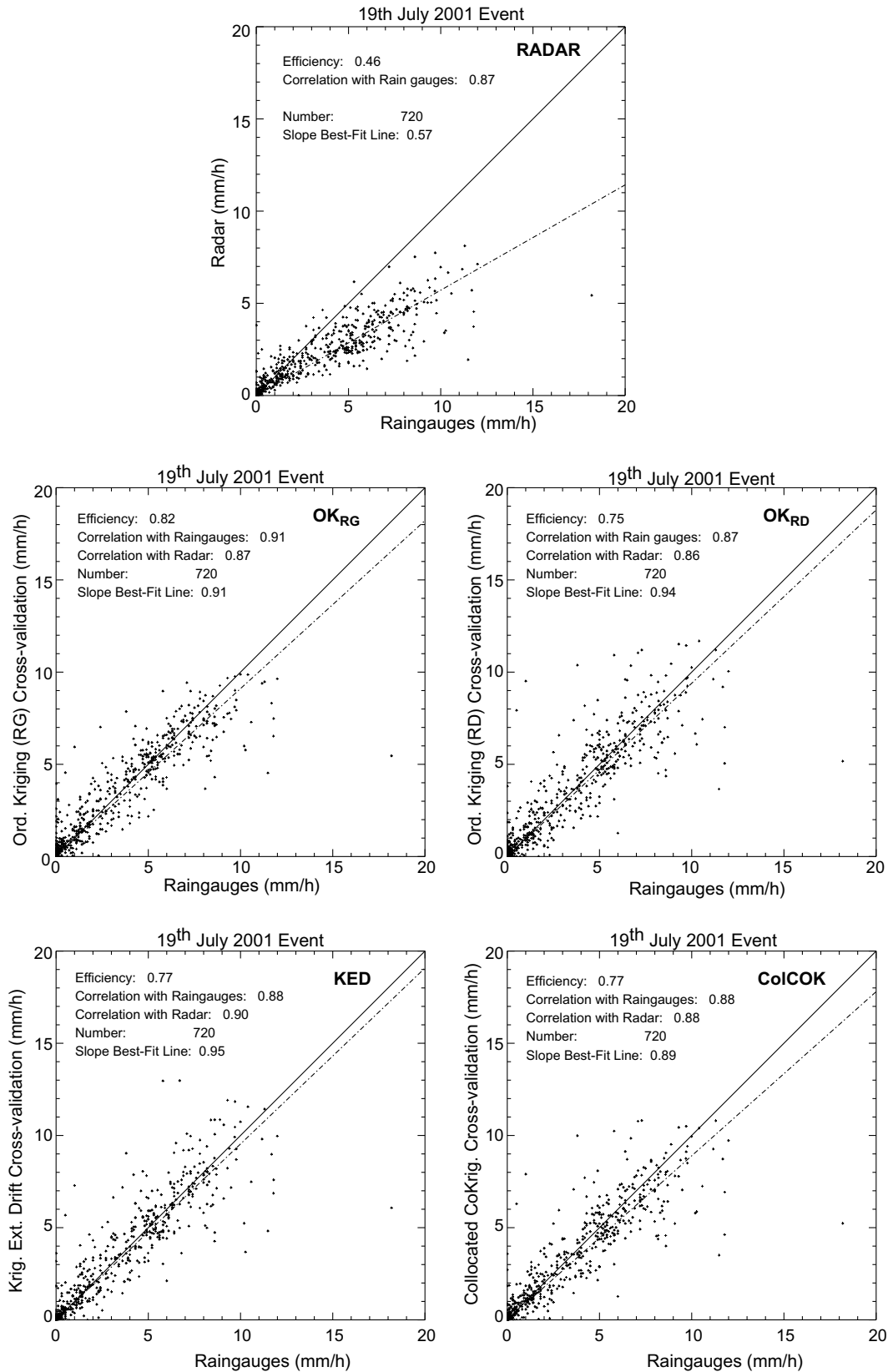


Fig. 5. Scatter plots showing the relationship between cross-validation estimations made with OK_{RG}, OK_{RD}, KED, and CoICOK, and the rain gauge data for event number 3 (19 July 2001). The results of each blending technique can be compared with original radar performance shown in the top of this figure.

estimated (measured) using the comparison technique (instrument) (e.g. radar rainfall or merged rainfall values), RG is the

mean rain rate measured by gauges and n is the number of known points.

Table 3

Statistics of the comparisons between cross-validation estimations of each estimate methodology and rain gauge observations. Radar estimations correspond to radar values of the collocated pixel at each rain gauge location. Results of each case study are presented separately. Last column shows the average value of each statistic for each estimator using all single event results. Cross-validation estimations and radar values correspond to 1-h rainfall. Values in bold indicate the best value for each quality criterion for each studied event, and for the average performance.

Quality criteria	Estimator	Case study						Average
		1	2	3	4	5	6	
Nash–Sutcliffe efficiency	Radar	0.30	0.70	0.46	0.56	0.54	0.18	0.46
	OK _{RG}	0.72	0.56	0.82	0.60	0.62	0.71	0.67
	OK _{RD}	0.76	0.55	0.75	0.53	0.49	0.66	0.62
	KED	0.80	0.74	0.77	0.56	0.68	0.71	0.71
	ColCOK	0.79	0.69	0.77	0.63	0.61	0.68	0.70
Correlation with rain gauges	Radar	0.71	0.85	0.87	0.75	0.77	0.71	0.78
	OK _{RG}	0.85	0.75	0.91	0.78	0.79	0.84	0.82
	OK _{RD}	0.88	0.76	0.87	0.75	0.73	0.82	0.80
	KED	0.90	0.86	0.88	0.78	0.84	0.85	0.85
	ColCOK	0.89	0.83	0.88	0.79	0.79	0.83	0.84
Correlation with radar	Radar	–	–	–	–	–	–	–
	OK _{RG}	0.67	0.75	0.87	0.69	0.71	0.72	0.74
	OK _{RD}	0.68	0.78	0.86	0.70	0.68	0.70	0.73
	KED	0.79	0.89	0.90	0.79	0.79	0.76	0.82
	ColCOK	0.72	0.86	0.88	0.78	0.78	0.71	0.79
Slope best-fit line	Radar	0.28	0.68	0.57	0.64	0.59	0.34	0.52
	OK _{RG}	0.71	0.67	0.91	0.68	0.74	0.83	0.76
	OK _{RD}	0.88	0.78	0.94	0.79	0.80	0.86	0.84
	KED	0.86	0.89	0.95	0.84	0.87	0.87	0.88
	ColCOK	0.82	0.77	0.89	0.78	0.77	0.83	0.81

The analysis of Table 3 shows that KED is the best estimator for all quality criteria except on events 3 and 4 (19 of July 2001 and 15 of November 2001) in which it is not the best for two out of the four quality criteria. In these cases (NS and correlation with rain gauges) the OK_{RG} estimations (i.e. using only rain gauges) provides higher NS and correlation with rain gauges in event number 3, and the ColCOK provides slightly better performances in event 4, although KED is still the best in correlation with radar and bias (slope of best-fit line closer to one). It is worth noting that in event 3 the radar is showing strong underestimation and that the registered intensities are low. This could explain why, in this case, the reference technique just using rain gauge data can perform slightly better in cross-validation than the blending techniques. However, note also that even if radar underestimation is quite severe in this case, all three blending techniques are able to compensate for this underestimation, and looking at the scatter plots (see Fig. 5) the differences are not too large. Finally, even in this case, the KED blending technique still shows the least bias. This result, which is general for all the analyzed events, is one of the main results at this stage, showing that even if there are slight differences between the three proposed blending methodologies, KED is in all the cases the most effective technique in correcting the underestimation tendency of the radar data in this study.

A comparison of the OK_{RD} and OK_{RG} results on bias criterion in Table 3 (given by the slope of the best-fit line closer to one) is quite noticeable. Since the only difference between both techniques is the source of the correlogram, the fact that OK_{RG} always shows worse results than OK_{RD} in bias criterion can be interpreted as evidence that the use of correlograms computed from radar are more useful in reducing systematic bias even when the radar is severely underestimating.

Moreover, the last column of Table 3 shows the average value of the statistics for each methodology. These average results show clear tendencies in all statistics. For all the quality criteria each of the estimators show greater values than the corresponding values for radar data, which illustrates the effectiveness of the adjustment process of rain gauges used in all of them. Also, for all quality criteria, the KED and ColCOK techniques provide better values than

those of the OK_{RG} and OK_{RD} estimates, showing that the inclusion of radar data in the interpolation process is valuable. In addition, it can be noted that the optimum values for all the statistics are obtained by the KED estimates. Finally, the values of the slope of best-fit line statistic (the systematic bias reduction) indicate that the underestimation tendency of radar data is considerably reduced after merging the two types of data with all tested methodologies. Again, the KED estimates using cross-validation provide the best fit to rain gauges observations.

Cross-validation errors should be uncorrelated in space. Therefore, if the (semi-)variogram of a given dataset of errors is computed and follows a pure nugget effect model (i.e., a constant level of variance for all distances, the same as that at the null range), then these errors can be considered spatially uncorrelated [32]. However, the methodologies used in this study change the spatial variability model for each time step. For this reason, semi-variograms of the cross-validation errors were computed for each time step and were then divided by the corresponding field variance for the purpose of comparison. Fig. 6 summarizes the cross-validation error semivariogram results for each methodology tested here with data from event 6 (29–30 of March 2004). The bold line shows the median value for each distance calculated from the individual results of all time steps. The bars show the inter-quartile range for each distance. In all events, the semivariogram error results are similar to Fig. 6, in the sense that all blending techniques have a good fit (median of each distance) to a pure nugget effect model (all values are close to one and do not exhibit much trend), which shows that the blended fields' cross-validation errors are uncorrelated in space, unlike that of radar that shows a definite trend.

Finally, we note that the results analyzed above correspond to cross-validation estimates at rain gauge locations. All statistics and plots presented in this section are intended to measure the performance of the different blending techniques proposed in this study associated with the correct and accurate estimation of rain gauge observations. However, one of the goals of this study is to identify a technique that estimates rainfall fields conditioned by both known rain gauge values and also the spatial variability re-

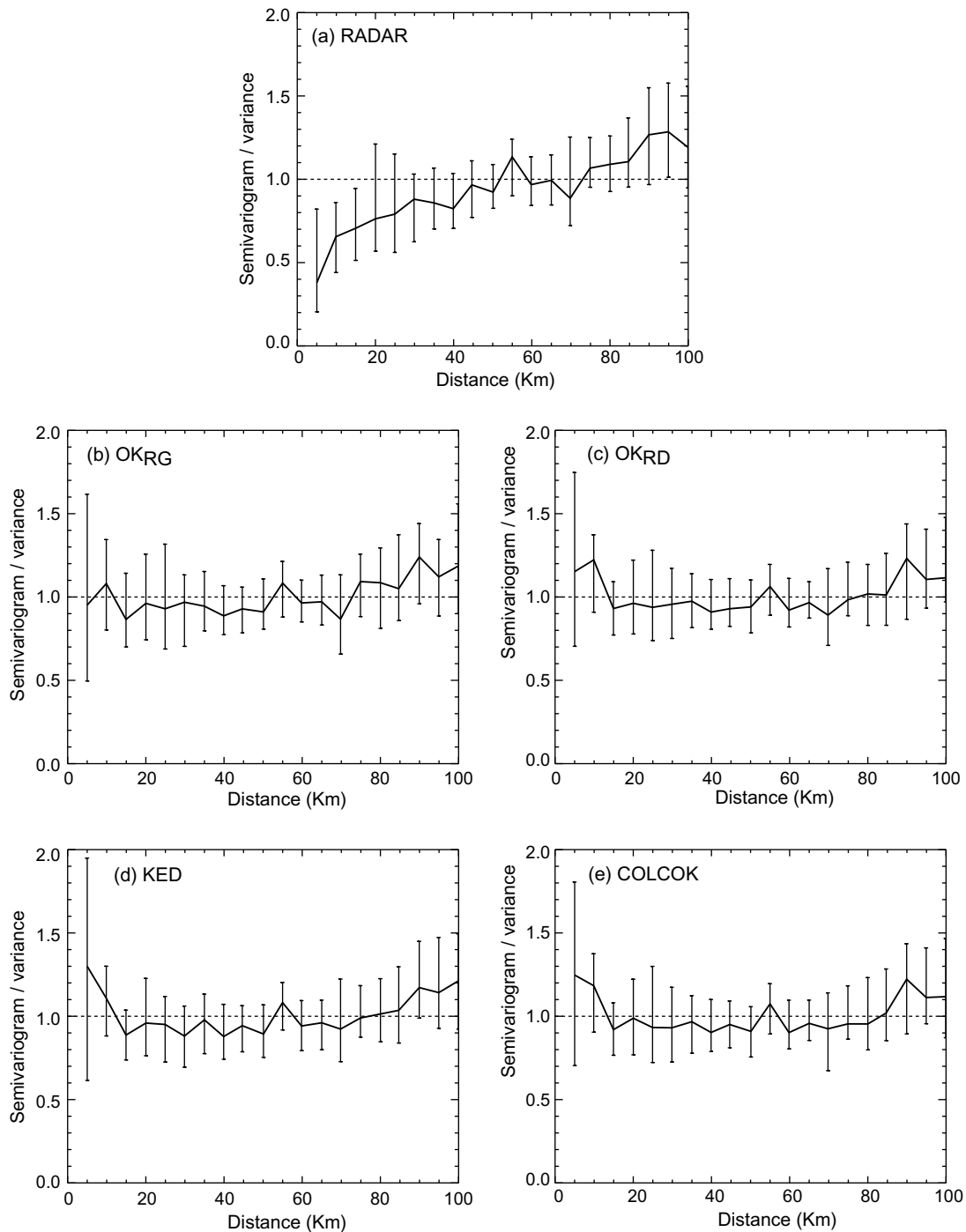


Fig. 6. Semivariogram of cross-validation errors for the event 30 of March 2004. In all estimation methodologies, semivariogram error could be modelled statistically with a pure nugget effect model indicating that estimation errors are uncorrelated in space.

ported by meteorological radar. Therefore, the agreement between the spatial properties of the estimated rainfall fields and the original radar fields is analyzed in the next section.

5.3. Spatial agreement between the blending estimates and radar fields

The most desirable characteristic for estimated rainfall fields is that they agree with the “known” values of precipitation, i.e. rain gauge observations, as evaluated in the previous section. However, different studies have shown that the spatial variability of rainfall fields is also important when seeking to improve hydrological esti-

mates. We understand that radar data provide an adequate representation (or at least a much better one than any rain gauge network) of these spatial patterns, so estimated rainfall fields should reproduce these patterns. Therefore, to evaluate the spatial characteristics of the estimates, rainfall fields were computed for all time-steps of all studied events using each blending methodology and were then compared with the original radar rainfall fields. As estimated rainfall fields would preserve only the general scales and not the point values of radar fields, comparisons were made using spectral analysis. Standard spectral methods and analysis can be used to test for scale-invariance of rainfall fields. The most

familiar result of scaling is the power-law behaviour that is expected in the power spectra of scaling processes. Therefore, the power spectrum derived from the rainfall fields should follow the power law in frequency (see e.g. [51–53])

$$E(\omega) \propto \omega^{-\beta} \quad (16)$$

where ω is the wave-number, $E(\omega)$ is the power spectrum of the rainfall field and β is the spectral power-law model exponent. The comparison variable used in this paper is thus the slope of the best linear fit to the logarithm of the power spectrum of the rainfall field, i.e. β . If the Fourier power spectra of two different rainfall fields have similar slopes (β), their spatial tendencies will be similar. It is worth noting that the energy spectrum is only a second-order statistic (i.e. the spectrum is related to the Fourier transform of the autocorrelation function) and is therefore not particularly robust. However, if only the fractal performance of rainfall fields is analyzed, the scaling of the second-order moments (characterized by the spectral exponents) is expected to provide almost complete information about their scaling properties, although for multifractal behaviour, this information alone may not be sufficient.

Fig. 7 shows scatter plots of Fourier spectrum slopes for radar rainfall fields against Fourier spectrum slopes for estimated rainfall fields for each technique used in this paper. The data from all case studies are included in these scatter plots (one point for each time

step). The values of β in the OK_{RG} and OK_{RD} rainfall fields are very different to the computed β from radar fields, indicating that these rainfall fields generally show different spatial patterns than the radar fields. In addition, a higher dispersion of values is observed with the OK_{RD} technique. The effect of including the radar data as a secondary variable in the interpolation process can be seen by comparing the values of β from OK_{RD} and ColCOK rainfall fields. The only difference in the definition of these estimators is the inclusion of radar data in the interpolation process of the ColCOK technique. In general, it can be seen that the inclusion of radar values produces values of β that are closer to those of radar fields as seen in Fig. 7c and d. The poor fit of the OK_{RD} field slopes (Fig. 7b) to the radar field slopes is contrasted by a closer correlation between ColCOK and radar rainfall fields (Fig. 7d), but a significant reduction in the dispersion of the results is not observed. However in agreement with previous results, the values of β for KED rainfall fields show the least amount of bias and scatter when compared to the values of β for radar rainfall fields. This indicates that the KED estimator should provide a more faithful reproduction of the spatial patterns of the rainfall fields than the other techniques. The good performance of KED was expected because this technique uses both the spatial trends and spatial correlogram of the radar field to define the estimated values. However, it is interesting to see the high consistency and low dispersion of KED re-

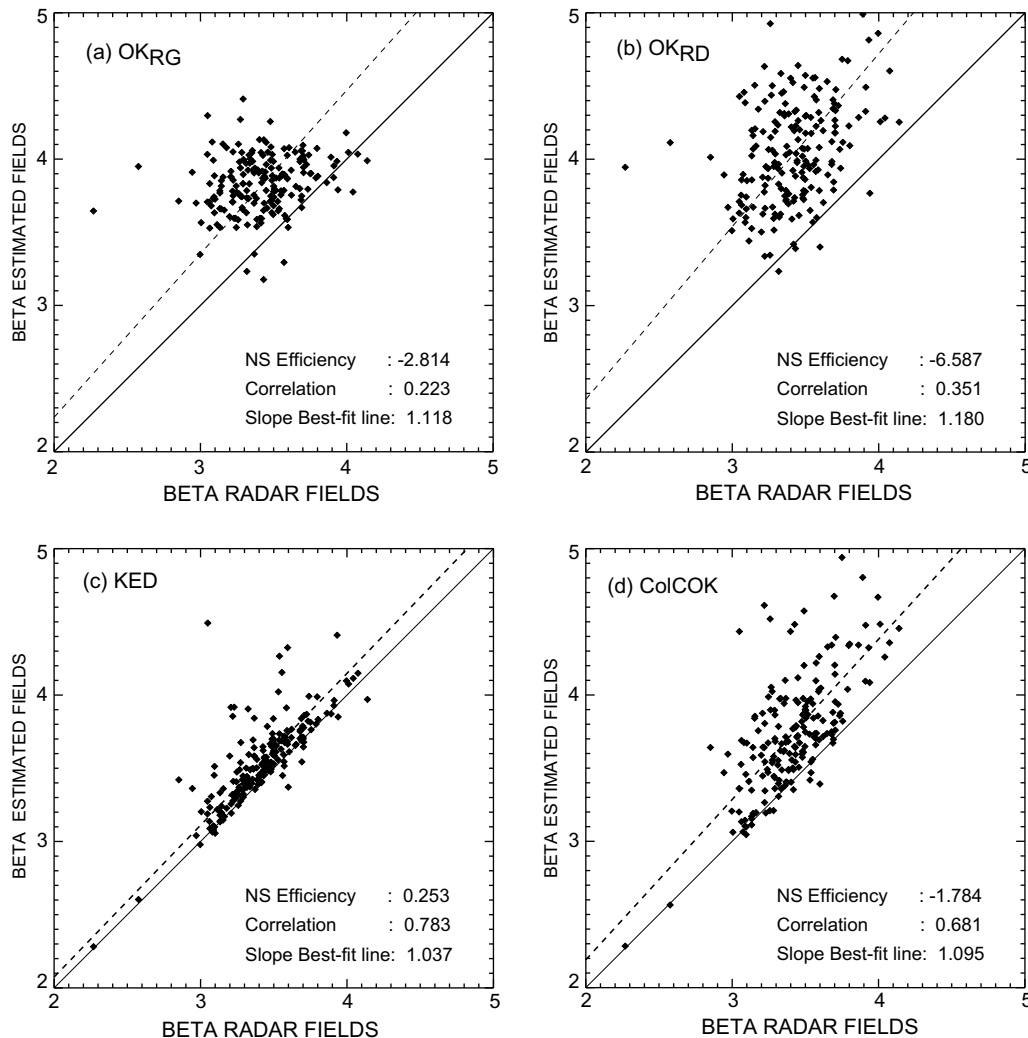


Fig. 7. Comparison between slopes (β) of Fourier Spectrum of radar rainfall fields and rainfall fields estimated by (a) OK_{RG}, (b) OK_{RD}, (c) KED and (d) ColCOK techniques. Rainfall fields for all time-steps of all case studies are shown.

sults when compared with the other estimators that also use the radar correlogram, especially with ColCOK.

Finally examples of estimated rainfall fields using each approach tested here are shown in Fig. 8. The estimated rainfall fields

were computed using the rain gauge data and the radar rainfall field shown in Fig. 8a and b, respectively. They correspond with 1-h accumulated rainfall values collected on 9 of October 2002 at 1300 UTC.

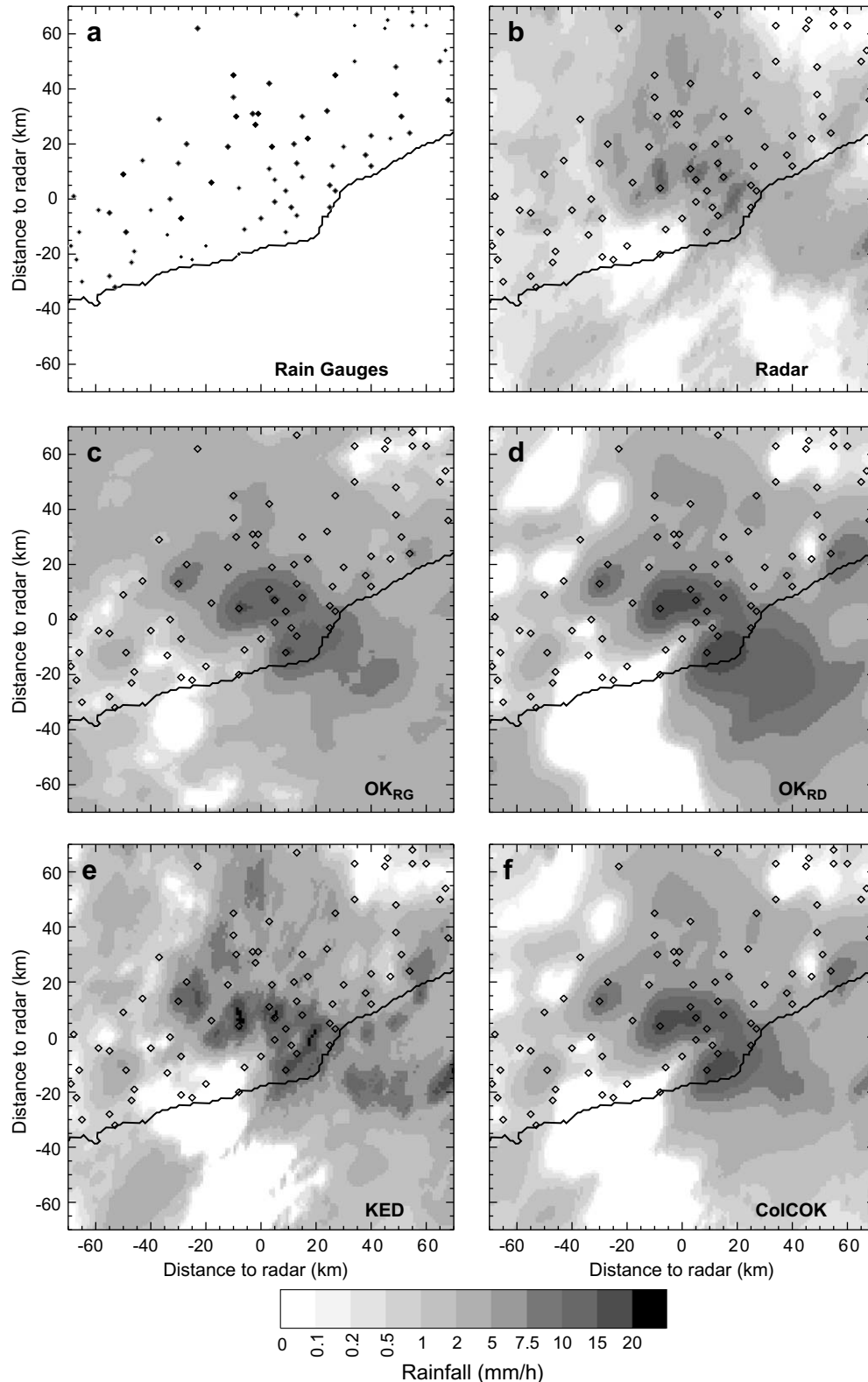


Fig. 8. Examples of estimated rainfall fields at 1300 UTC 9 of October 2002 (event 5). Rain gauge data and radar rainfall field are shown in (a) and (b), respectively. Then, estimated rainfall fields computed by (c) ordinary Kriging using the correlation map of rain gauge data (OK_{RG}), by (d) ordinary Kriging using the correlation map of radar data (OK_{RD}), by (e) Kriging with external drift (KED), or by (f) collocated cokriging (ColCOK), are shown as well. Values correspond to a 1-h accumulation period. Diamonds in rainfall fields indicate the rain gauge locations. Bold lines show the Mediterranean coast near Barcelona (Spain).

Rainfall fields estimated by OK_{RG} and OK_{RD} are shown in Fig. 8c and d, respectively. Both estimated rainfall fields have similar main tendencies and values in the regions situated close to rain gauge locations. Differences between these rainfall fields are due to the discrepancy between the correlation distances of each correlation map. Since the correlation distances of the rain gauge correlation map are shorter in all directions (see Fig. 1e and f), the OK_{RG} rainfall field tends to fill most estimation points located far from known values with the mean rainfall value. The OK_{RD} rainfall field interpolates rain gauge values in similar way to OK_{RG} , but their distance of influence is greater when OK_{RD} is used. Therefore, each known point produces broad rainfall areas of influence around it, e.g. the high rainfall value measured near the coast extends its high value over the sea further than in the OK_{RG} rainfall field. On the other hand, the rainfall field estimated using the ColCOK technique, shown in Fig. 8f, has similar spatial tendencies to those observed in the OK_{RD} rainfall field (Fig. 8d). However the inadequate coverage of rainfall values in zones without rain gauge observations shown in the OK_{RD} rainfall field is reduced with ColCOK since radar values provide useful additional information that controls the estimation process when fewer rain gauge data are available. The rainfall field estimated using the KED technique is shown in Fig. 8e. The spatial tendencies of the KED rainfall field are closer to those observed in the radar rainfall field than in the other estimated rainfall fields, but the rain gauge observations are maintained. Once more, this example shows the great potential of the KED technique in the radar–rain gauge blending context.

6. Summary and conclusions

This paper has reported work on the use of new non-parametric geostatistical methodologies for estimating rainfall fields using radar and rain gauge data. The goal was to identify an advanced geostatistical technique that makes possible the estimation of rainfall fields blending radar and rain gauge data in real time, under the hypothesis that estimated rainfall fields should be consistent with rain gauge observations but showing spatial patterns similar to the rainfall fields observed by radar.

Spatial variability models (correlograms) for the estimates were not defined according to the traditional method but rather estimated automatically using the new proposed technique based on a non-parametric FFT-smoothing process. This automatic technique for computing correlograms has considerable potential advantages when compared with the conventional method of defining spatial variability models. Specifically, with this method, correlation fields can be obtained without previous hypotheses about either the theoretical correlation model or average temporal behaviour, allowing a full appraisal of the anisotropy and the temporal variability of the fields. Thus, in this study, individual two-dimensional spatial correlograms were defined automatically for each time step and then used to estimate rainfall fields, avoiding the classical assumptions of spatial isotropy and without making previous steady-state “climatological” analyses. In all cases, the correlograms computed with this technique guaranteed the solution of the geostatistical equation systems. This automatic technique for estimating correlograms is extremely fast and therefore suitable for real-time applications of radar and rain gauge data for operational purposes. To our knowledge, this is the first implementation of this automatic technique in this field.

Rainfall fields were estimated by blending radar and rain gauge data taken from six rainfall events by using three geostatistical alternatives that differ in the way in which the radar data are included in the estimation. The automatic technique for defining correlograms indicated above was used to incorporate the spatial correlograms computed from radar data into the estimates. The re-

sults have been compared against the widely used reference technique: ordinary kriging with only rain gauge data.

Moreover, cross-validation was used to quantify the accuracy of the estimation alternatives, taking into account the known precipitation values. Analysis of the results for each event reveals that the estimates using the three alternatives of the proposed technique were in general more in agreement with observed rain gauge data than the radar data. Also these results showed that estimates made with the kriging with external drift (KED) alternative are more accurate than those obtained with the other estimators. Estimates made with KED also provided the highest correlation with radar data and the best bias reduction in comparison with rain gauge data. These results indicate that radar data need to be included in the interpolation process of rain gauge data because they increase the accuracy of estimates significantly.

Finally, comparisons using spectral analysis were made to analyze the spatial behaviour of the merged, improved rainfall fields compared to the original radar fields. The comparison indicator used was the slope of the best linear fit to the logarithm of the power spectrum of the rainfall fields. In this case, the results indicate again that the KED alternative provides better reproductions of the spatial patterns of the radar rainfall fields than the other techniques. Although these results were expected, because the KED methodology makes intensive use of the radar data, it was surprising to see the high consistency and low dispersion of KED results when compared with the other alternatives that also include radar data in their interpolations.

For this case study and in the region of application, it is possible to conclude that results seem to indicate that KED methodology in combination with the technique to compute 2-D spatial correlograms provides the best agreement with rain gauges and the most accurate approach to the spatial tendencies observed in the radar rainfall fields than the other alternatives analyzed in this paper.

Although further analysis of more cases is needed to fully document the sensitivity of these techniques to systematic errors in calibration (over- or underestimation) and other errors of radar fields, to the density of the rain gauge network used for the merging process, and also to shorter time-steps, the results presented here show that the proposed automatic, non-parametric geostatistical technique is suited to the problem of blending radar and rain gauge data to obtain field estimates in real time.

Acknowledgments

The authors thank the Spanish Meteorological Institute (Agencia Estatal de Meteorología) and the Catalan Water Agency (Agencia Catalana de l'Aigua) for providing radar and rain gauge data, respectively. Thank are due to Prof. Geoff Pegram, Dr. Alan Seed, and the three anonymous reviewers for their extraordinary help giving suggestions and reviewing the manuscript. The studies in which this paper is based on have been partially funded by the Catalan Water Agency as well as the EC projects FLOODSite (GOCE-CT-2004-505420) and HYDRATE (GOCE-CT-2006-037024), and the Spanish MEC ESP2007-62417 project. The first author is also grateful to the Technical University of Catalonia (UPC) for supporting his doctoral studies through the scholarship “UPC per a la Formació de Personal Investigador”.

References

- [1] Sempere-Torres D, Corral C, Raso J, Malgrat P. Use of weather radar for combined sewer overflows monitoring and control. *J Environ Eng* 1999;125(4):372–80.
- [2] Seo D-J. Real-time estimation of rainfall fields using radar rainfall and rain gauge data. *J Hydrol* 1998;208(1–2):37–52.
- [3] Sun X, Mein RG, Keenan TD, Elliot JF. Flood estimation using radar and raingauge data. *J Hydrol* 2000;239:4–18.

- [4] Harrold TW, Austin PM. The structure of precipitation systems. *J Res Atmos* 1974;8:41–57.
- [5] Wilson JW, Brandes EA. Radar measurement of rainfall – a summary. *Bull Am Meteor Soc* 1979;60:1048–58.
- [6] Chumchean S, Seed A, Sharma A. Correcting of real-time radar rainfall bias using a Kalman filtering approach. *J Hydrol* 2006;137(1–2):123–37.
- [7] Hevesi JA, Flint AL, Istok JD. Precipitation estimation in Mountainous Terrain using multivariate geostatistics. Part II: Isohyetal Maps. *J Appl Meteorol* 1992;31(7):677–88.
- [8] Hevesi JA, Istok JD, Flint AL. Precipitation Estimation in Mountainous Terrain using multivariate geostatistics. Part I: Structural analysis. *J Appl Meteorol* 1992;31(7):661–76.
- [9] Calheiros RV, Zawadzki I. Reflectivity-rain rate relationship for radar hydrology in Brazil. *J Climate Appl Meteorol* 1987;26:118–32.
- [10] Rosenfeld D, Amitai E, Wolff D. Improved accuracy of radar WPMM-estimated rainfall upon application of objective classification criteria. *J Appl Meteorol* 1995;33(11):212–23.
- [11] Rosenfeld D, Atlas D, Wolff DB, Amitai E. The window probability matching method for rainfall measurements with radar. *J Appl Meteorol* 1994;34:682–93.
- [12] Rosenfeld D, Wolff D, Atlas D. General probability-matched relations between radar reflectivity and rain rate. *J Appl Meteorol* 1993;32:50–72.
- [13] Krajewski WF. Cokriging radar-rainfall and rain gage data. *J Geophys Res* 1987;92(D8):9571–80.
- [14] Creutin JD, Delrieu G, Lebel T. Rain measurement by rain gauge–radar combination: a geostatistical approach. *J Atmos Oceanic Technol* 1988;5:102–15.
- [15] Azimi-Zonooz A, Krajewski WF, Bowles DS, Seo D-J. Spatial rainfall estimation by linear and non-linear co-kriging of radar-rainfall and rain gauge data. *Stoch Environ Res Risk Assess* 1989;3:51–67.
- [16] Seo D-J, Krajewski WF, Bowles DS. Stochastic interpolation of rainfall data from rain gages and radar using cokriging. I: Design of experiments. *Water Resour Res* 1990;26(3):469–77.
- [17] Seo D-J. Real-time estimation of rainfall fields using rain gage data under fractional coverage conditions. *J Hydrol* 1998;208(1–2):25–36.
- [18] Sinclair S, Pegram GGS. Combining radar and rain gauge rainfall estimates using conditional merging. *Atmos Sci Lett* 2005;6(1):19–22.
- [19] Todini E. A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements. *Hydrol Earth Syst Sci* 2001;5(2):187–99.
- [20] Yao T, Journel AG. Automatic modeling of (cross) covariance tables using fast Fourier transform. *Math Geol* 1998;30(6):589–615.
- [21] Velasco C, Cassiraga EF, Sempere-Torres D, Sánchez-Diezma R, Gómez-Hernández JJ. Merging radar and raingauges data to estimate rainfall fields: an improved geostatistical approach using non-parametrical spatial models. In: Sixth international symposium on hydrological applications of weather radar. Melbourne: 2004.
- [22] Velasco-Forero C, Cassiraga E, Sempere-Torres D, Sánchez-Diezma R, Gómez-Hernández JJ. Merging radar–raingauge rainfall estimates: an improved geostatistical approach based in non-parametric Kriging. In: 31st conference on radar meteorology. Seattle: AMS; 2003.
- [23] Nicolau R, Ribeiro L, Rodrigues RR, Pereira HG, Câmara AS. Mapping the Spatial Distribution of Rainfall in Portugal. In: Geostats 2000. South Africa: 2000.
- [24] Germann U, Joss J. Variograms of radar reflectivity to describe the spatial continuity of Alpine precipitation. *J Appl Meteorol* 2001;40(6):1042–59.
- [25] Berne A, Delrieu G, Creutin JD, Obled C. Temporal and spatial resolution of rainfall measurements required for urban hydrology. *J Hydrol* 2004;299(3–4):166–79.
- [26] Yao T. Automatic covariance modeling and conditional spectral simulation with fast Fourier transform. PhD thesis, 1998. 173 Pages. Department of Geological and Environmental Sciences, Stanford University. Stanford, California.
- [27] Wackernagel H. Multivariate Geostatistics an introduction with applications. Berlin: Springer-Verlag; 1995. p. 256.
- [28] Cassiraga EF, Guardiola-Albert C, Gómez-Hernández JJ. Automatic modeling of cross-covariances for rainfall estimation using raingage and radar data in geoENV II. Barcelona, Spain: Kluwer Academic Publishers; 2002.
- [29] Bochner S. Fourier transforms. London: Princeton University Press; 1949. p. 219.
- [30] Ma X, Yao T. A program for 2D modeling (cross) correlogram tables using fast Fourier transform. *Comput Geosci* 2001;27(7):763–74.
- [31] Journel AG, Huijbregts CJ. Mining geostatistics. London (UK): Academic Press; 1978. p. 600.
- [32] Isaaks EH, Srivastava RM. Applied geostatistics. Oxford (UK): Oxford University Press; 1989.
- [33] Goovaerts P. Geostatistics for natural resources evaluation. Oxford University Press; 1997. p. 483.
- [34] Deutsch CV, Journel AG. GSLIB: geostatistical software library and user's guide. 2nd ed. Oxford: Oxford University Press; 1997. p. 368.
- [35] Raspa G, Tucci M, Bruno R. Reconstruction of rainfall fields by combining ground raingauges data with radar maps using external drift method. In: Baafi EY, Schofield NA, editors. Geostatistics wollongong'96. Kluwer Academic Publishers; 1997. p. 1306–15.
- [36] Grimes DIF, Pardo-Iguzquiza E, Bonifacio R. Optimal areal rainfall estimation using raingauges and satellite data. *J Hydrol* 1999;222(1–4):93–108.
- [37] Hengl T, Heuvelink GBM, Stein A. Comparison of Kriging with external drift and regression-Kriging. Technical Note, ITC 2003.
- [38] Almeida AS, Journel AG. Joint simulation of multiple-variables with a Markov-type coregionalization model. *Math Geol* 1994;26(5):565–88.
- [39] Journel AG. Markov models for cross-covariances. *Math Geol* 1999;31(8):955–64.
- [40] Shmaryan LE, Journel AG. Two Markov models and their application. *Math Geol* 1999;31(8):965–88.
- [41] Bech J, Rigo T, Pineda N, Segalà S, Vilaclara E, Sánchez-Diezma R, et al. Implementation of the EHIMI software package in the weather radar operational chain of the catalan meteorological service. In: 32nd conference on radar meteorology. Albuquerque-EUA: AMS; 2005.
- [42] Sempere-Torres D, Sánchez-Diezma R, Berenguer M, Pascual R, Zawadzki I. ASCMORE: an algorithm to control radar rainfall measurement stability using mountain returns in real time. *Geophys Res Abstr* 2002;4:6319.
- [43] Marshall JS, Palmer WK. The distribution of raindrops with size. *J Meteorol* 1948;5:165–6.
- [44] Bellon A, Fabry F, Austin GL. Errors due to space/time sampling strategies in high resolution radar data used in hydrology. In: 25th Radar Meteor. Conference. Paris: AMS; 1991.
- [45] Amitai E. Systematic variation of observed radar reflectivity-rainfall rate relations in the tropics. *J Appl Meteorol* 2000;39(12):2198–208.
- [46] Habib E, Ciach GJ, Krajewski WF. A method for filtering out rain gauge representativeness errors from the verification distributions of radar and raingauge rainfall. *Adv Water Resour* 2004;27(10):967–80.
- [47] Upton GJG, Rahimi AR. On-line detection of errors in tipping-bucket raingauges. *J Hydrol* 2003;278(1–4):197–212.
- [48] Zawadzki I. On radar–rain gauge comparison. *J Appl Meteorol* 1975;14(8):1430–6.
- [49] Zawadzki I. Statistical properties of precipitation patterns. *J Appl Meteorol* 1973;12:459–72.
- [50] Nash JE, Sutcliffe JV. River flow forecasting through conceptual models, 1. A discussion of principles. *J Hydrol* 1970;10:282–90.
- [51] Lovejoy S, Mandelbrot B. Fractal properties of rain and a fractal model. *Tellus* 1985;37A:209–32.
- [52] Lovejoy S, Schertzer D. Multifractals and rain. In: Kundzewicz AW, editor. New uncertainty concepts in hydrology and hydrological modelling. Cambridge Press; 1995. p. 62–103.
- [53] Pegram GGS, Clothier AN. High resolution space–time modelling of rainfall: the “String of Beads” model. *J Hydrol* 2001;241(1–2):26–41.