Proceedings of ERAD (2004): 223–229 © Copernicus GmbH 2004

ERAD 2004

# Spatial characterization and classification of rainfall fields derived from operational c-band weather radar data

**F. S. Marzano**[1], **G. Vulpiani**[1], **P. P. Alberoni**[2], **L. Ferraris**[3], **L. Provenzale**[3], **and N. Rebora**[3]

[1]Centro di Eccellenza CETEMPS, Dip. di Ingegneria Elettrica e Dipartimento di Fisica, Univ. dell'Aquila, L'Aquila, Italy
[2]Servizio Idro Meteororologico (SIM), ARPA Emilia Romagna, Bologna, Italy
[3]CIMA, Centro di ricerca Interuniversitario in Monitoraggio Ambientale, Univ. di Genova e della Basilicata, Savona, Italy

**Abstract.** Preliminary results of a spatial characterization of rain fields, performed by means of an analysis of spatial power spectra, spatial correlation functions and statistical moments, is here presented. To this aim, a large rainfall dataset derived from an operational C-band dual-polarized radar, placed in S.Pietro Capofiume near Bologna (Italy), is analysed. The data set consists of radar-derived rain intensities on a grid of 256 by 256 km with a spatial resolution of 1 km and a time sampling of 15 min. from 1996 to 1999. A data quality control and a rain event selection is systematically carried identifying more than 40 events of interest. Power spectrum slopes, spatial correlation lengths, mean rain intensity and variation coefficient, event duration are investigated as potential indicators of rain event features. A supervised classification scheme, based on a Gaussian metrics, is finally proposed in order to categorize in an automatic way stratiform, convective and hybrid rain events.

## 1 Introduction

Spatial and temporal characterization of rainfall patterns is not only of meteorological interest, but it is becoming more and more important in various fields (Austin and Houze, 1972; Bell, 1987; Crane 1990). For hydrological purposes, the interest to model rain fields is mostly due to the need of initializing run-off models and to devise downscaling methods in order to properly deal with sub-basin scales for flood forecast (Veneziano et al., 1996; Perica and Foufoula-Georgoiu, 1996; Willems, 2001). Recently, radiopropagation community has also started to explore the possibility to model and generate rain fields in a synthetic way in order to accomplish a more accurate link budget design at Ka frequency band and above (Gremont and Filip, 2004).

Apart from the specific purpose, the only choice to model rain fields in a realistic way is to resort to statistically significant, fairly accurate and highly spatially resolved rain measurements (Meija and Rodriguez-Iturbe, 1974; Wheater et al., 2000; Ferraris et al., 2003). In this respect, weather microwave radars can offer a unique tool due to their characteristics of wide areal coverage, high sensitivity to hydrometeor electromagnetic backscattering and fairly high resolution, especially at short to medium ranges (i.e. up to 125 km from the radar site). Indeed, the accuracy of surface rain fields estimation can be questionable and should be carefully considered under some circumstances, as in the case of beam blockage, enhanced ground-clutter, anomalous propagation, bright band contamination, strong path attenuation, second-trip echo ambiguity and long-range error bias (e.g. Alberoni et al., 2001; Marzano et al., 2004). Moreover, the relation between radar backscattered power and rain intensity is far from being constant and can depend on rainfall microphysics (Sauvageot, 1992). Nevertheless, the radar-derived estimates can be considered nowadays the major source for a robust spatial-temporal characterization of rainfall features.

In this work, a preliminary spatial characterization of the rain fields has been performed by means of an analysis and modelling of spatial power spectra, spatial correlation functions and histogram statistical moments. A large rainfall dataset derived from an operational C-band dual-polarized radar, placed in S. Pietro Capofiume near Bologna (Italy), has been collected and analysed for this purpose. The data set consists of radar-derived rain intensities on a grid of 256x256 km with a spatial resolution of 1 km from 1996 to 1999. A preliminary supervised classification scheme has been finally investigated in order to cluster stratiform, convective and hybrid rain events.

## 2 Radar data

Data were provided by C-band operational Doppler of S. Pietro Capofiume (Bologna, Italy), placed along the Po river valley in northern Italy (Alberoni et al., 2001). This dual-polarization radar is placed on a tower with a Cassegrain

*Correspondence to:* F. S. Marzano
(marzano@ing.univaq.it)

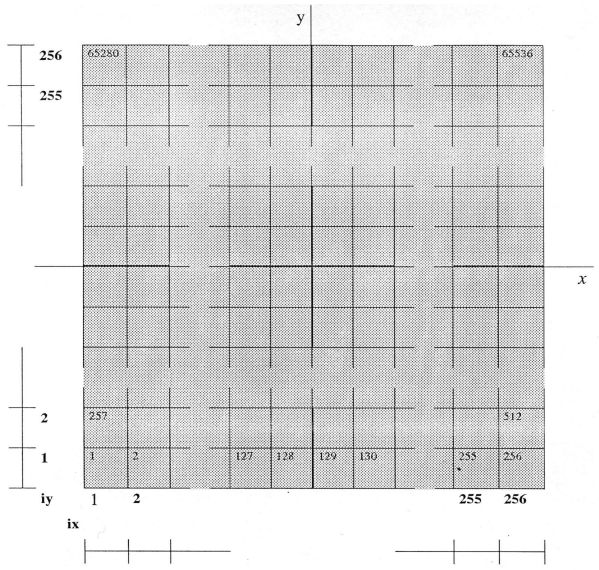


**Fig. 1.** Cartesian map to project polar radar maps.

parabolic antenna (without radome cover), providing a half-power beam-width of 1.0° and a directivity of about 45-dB. The klystron peak-power is 250 kW at 5.6 GHz with an alternating horizontal-vertical polarization transmission and dual pulse repetition frequency (PRF) system for unfolding capability. Pulse widths of 0.5 ms (i.e. short pulse with a resampled bin resolution of 250 m) and 1.5 ms (i.e. long pulse with a resampled bin resolution of 1500 m). The receiver sensitivity is equal to −113 dBm. The typically used maximum range is 250 km (with long pulse) and 125 km (with short pulse) for the intensity and velocity mode, respectively. A self-contained software is used to remotely operate and archive radar data.

Radar data are acquired with a prescribed scanning strategy during operational activity, consisting of 15 elevations with an angular spacing of 1°. Radial spatial resolution is set to 250 m for short ranges (i.e. 125 km) and to 1000 km for long range (i.e. 250 km) scans, the latter being carried out only twice per hour. Time sampling of radar volume data is such that there are 4 acquisitions per hour (i.e. every 15 min) being the dual-polarized one performed only twice per hour. Procedures to correct for gas absorption, to remove ground-clutter echoes and to identify anomalous propagation conditions are routinely applied. Side-lobe effects at very short ranges (less than 20 km) for low elevations are avoided by choosing higher elevations not affected by this effect.

## 2.1 Data processing

Three years of radar data, acquired from January 1996 to December 1999 in an operational mode, have considered here. For the present study we have used only radar reflectivity at horizontal resolution ($Z_{hh}$ or simply Z) at 250 m radial resolution and a maximum range of 125 km. Due to orographic blockage in the south-western sector of observation, at each range-azimuth bin location we have extracted the value of Z relative to the lowest available elevation using a radar visibility map.

The resulting polar map of measured lowest-bin Z has been then projected on a regular Cartesian grid having 256×256 pixel, each of them of 1 km resolution. The following transformation has been used from radar latitude-longitude ($\lambda_R$,$\varphi_R$) coordinates to ($x$, $y$) pixel centre:

$$\begin{aligned} y &= (\lambda - \lambda_R)\,\tfrac{\pi}{180} R_e \\ x &= (\varphi - \varphi_R)\,\tfrac{\pi}{180} \cos\left(\tfrac{\lambda+\lambda_R}{2}\right) R_e \end{aligned} \quad (1)$$

where $R_e$ is the effective Earth radius and $\lambda, \phi$ the pixel latitude-longitude. Figure 1 shows the regular grid with 65536 pixels. In case more than one Z value belonged to same Cartesian pixel, an average Z value has been attributed to the pixel itself.

For easy representation and interpretation, each reflectivity value of these gridded maps has been then converted to surface rainrate R, using a standard power law Z-R relation (Sauvageot, 1992):

$$Z = aR^b \quad (2)$$

where, for the coefficients $a$ and $b$, we adopted the Marshall-Palmer ones ($a$=200, $b$=1.6 with R in mm/h and Z in $mm^6\,m^{-3}$). As a result, the measured dataset has been converted into a temporal series of rainfall field maps R($x$, $y$).

In order to produce a hourly rainrate map we have also performed a time average of available maps doing a simple uniform weighting to compute a mean rain rate per hour $R_h$:

$$R_h = \frac{1}{4}\left(R^{00} + R^{15} + R^{30} + R^{45}\right) \quad (3)$$

where $R^{tt}$ are rainrate values available at *tt* starting minute of each volume scan within a given hour. Some anomalous features with values of R larger than 200 mm/h, probably due to receiver saturation, were also removed by a smoothing spatial filter.

## 2.2 Examples

A rain event has been defined as an episode where rain rates larger than 0.1 mm/h were at least 10% of the entire image and with a hourly latency not longer than 1 h. With this criterion, more than 40 events were detected within the 3-year dataset.

The following figures show some example of selected rain events. The total accumulated rainfall (AR) during the event is represented by a coded image and corresponding histogram. Figures 2 and 3 show the polar map of typical stratiform events where accumulated rainfall was not larger than 200 mm with a mean AR of about 39 mm in 16 h (see Fig. 2) and of 6 mm in 11 h (see Fig. 3).

Figures 4 and 5 show mostly convective events where accumulated rainfall was larger than 500 mm with a mean AR of about 124 mm in 17 h (see Fig. 4a) and of 46 mm/h in 3 h (see Fig. 4b).

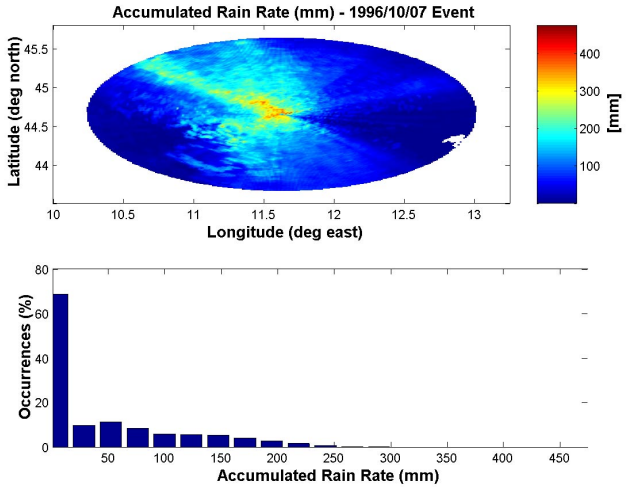


**Fig. 2.** Rain event of October 7, 1996.

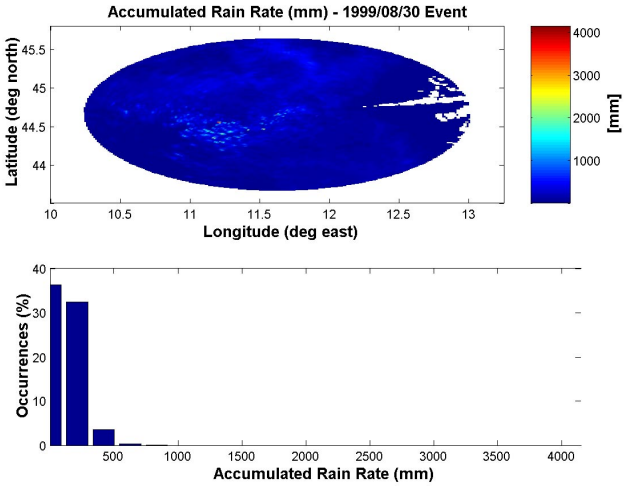


**Fig. 4a.** Rain event of August 30, 1999.

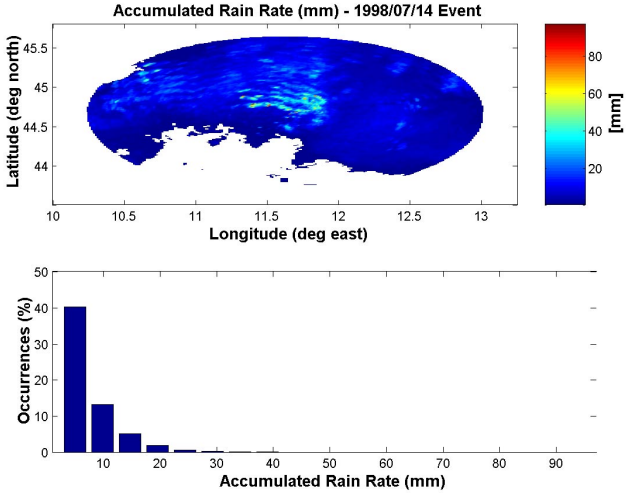


**Fig. 3.** Rain event of July 14, 1998.

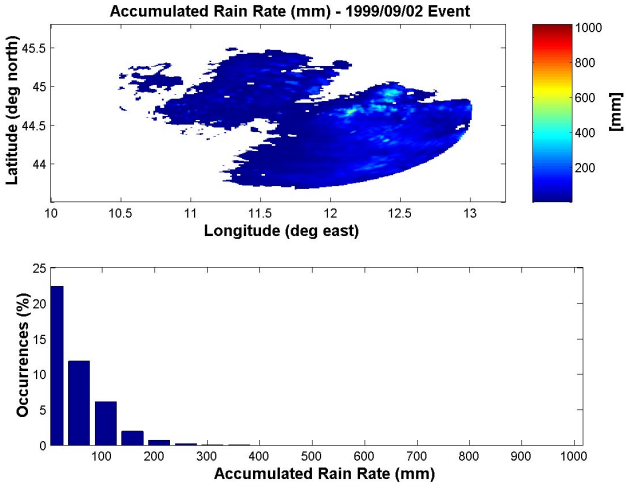


**Fig. 4b.** Rain event of September 2, 1999.

## 3 Spatial data analysis

The main goal of this radar data analysis is to achieve a spatial characterization of the observed events in terms of a statistical and spectral parameterisation. As already mentioned, this purpose may have manifold applications:

- to set up an automatic classification scheme where stratiform, convective and hybrid (convection embedded in stratiform rain) events can be detected;
- to initialise hydrological rainfall disaggregation models in order to downscale rainfall features to lower spatial scales for flood forecast purpose;
- to define rainfall spatial patterns, typical of stratiform and convective events, useful for the development of radio-communication link budgets at Ka band and above.

### 3.1 Spectral analysis

Spatial rain fields R$(x, y)$ in a logarithmic scale can be characterised by a power spectrum S$(k_x, k_y)$ by performing a Fourier analysis. Indeed, thanks to the "frozen storm" assumption, some of the conclusions derived with respect to spatial properties can be also attributed to temporal spectra (Veneziano et al., 1996). A typical result of this spectral analysis, due to Crane (1990), is shown in Fig. 5 where $k$ (and $w$) stands for the spatial (temporal) wave number for the 1-dimensional (1-D) case.

The major result from Fig. 5 is the identification of a stepwise power law behaviour, namely for each interval in a logarithmic scale (Crane, 1990; Veneziano et al., 1996):

$$Log[S(k)] = \alpha + \beta Log(k) \tag{4}$$

where the coefficients $\alpha$ and $\beta$ depend on the $k$ or $w$ range. For wave numbers lower than below $k_1$ (typically 1/15 to 1/7 km$^{-1}$ and due to energy injection associated to

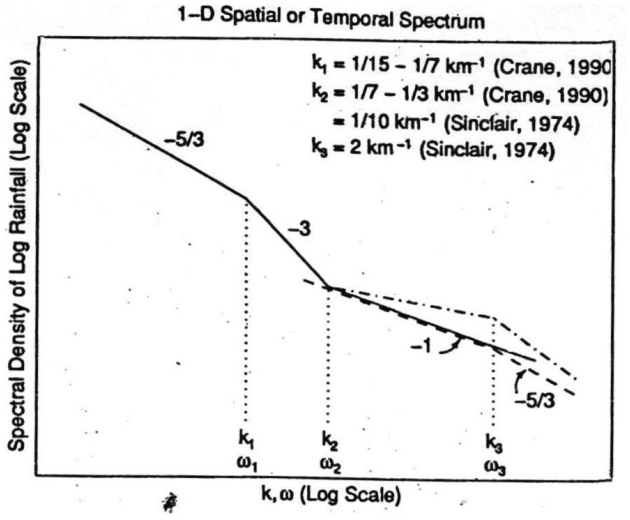


**Fig. 5.** Typical power spectrum of a log-rain field (Veneziano et al., 1996).

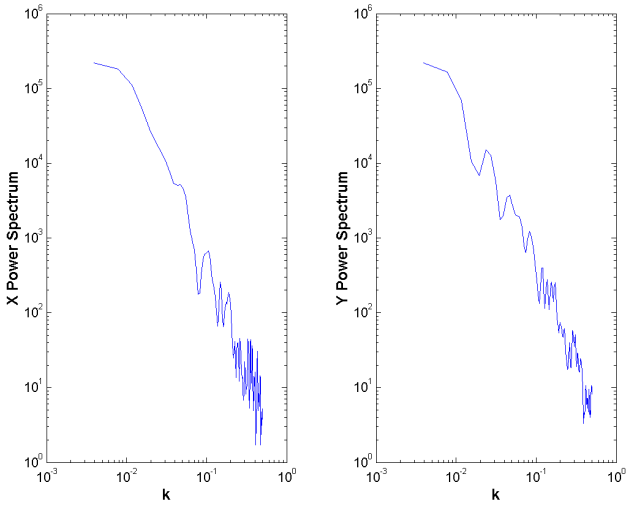


Fig. 6. 1-D power spectrum along x and y for rain event on Sept. 9, 1999.

**Fig. 6.** 1-D power spectrum along x and y for rain event on Sept. 9, 1999.

convective rain clusters), the spectrum behaves like $k^{-5/3}$, while after $k_2$ (typically 1/7 to 1/3 km$^{-1}$), like $k^{-3}$. Between $k_2$ and $k_3$ (i.e. for scales between 0.5 and 10 km) the behaviour is $k^{-1}$, while for smaller scales (i.e. beyond $k_3$ in Fig. 5) it changes to $k^{-5/3}$. The latter transition may be associated to a transition in the vertical velocity regime.

Keeping in mind Fig. 5 and that 5/3≅1.67, we have performed a spectral power analysis using the Fourier periodogram method. The latter has been applied to the average rain field along x and along y, separately, in order to obtain a 1-D power spectra. Different averaging procedures may produce slightly different results.

An example is shown in Fig.6 for the rain event on Sept. 9, 1999. It is interesting to note that the spectrum is not necessarily symmetrical with respect to x and y wave numbers. A power law behaviour seems to be confirmed by this analysis. The spectrum slopes have been computed around k$\prime$=1/10 km$^{-1}$ and k$\prime\prime$=1/20 km$^{-1}$ for both spectra, obtaining: $\beta = -1.8$ at k$\prime$ and $\beta = -3.3$ at k$\prime\prime$ along x, while $\beta = -1.9$ at k$\prime$ and $-2.5$ at k$\prime\prime$ along y.

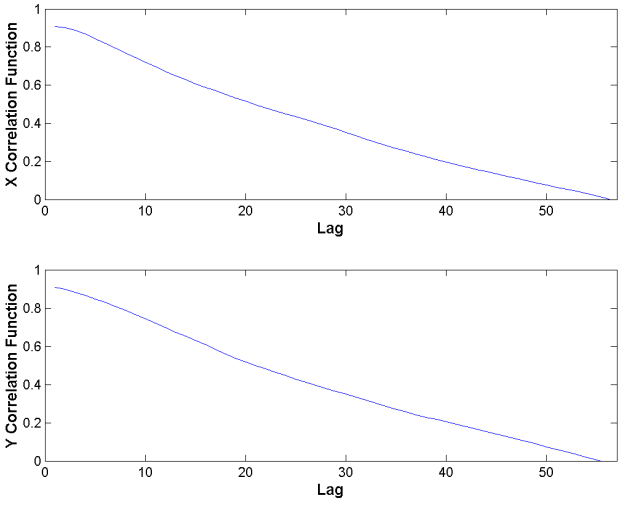


**Fig. 7.** 1-D correlation function along x and y for event on Sept. 9, 1999.

### 3.2 Spatial correlation analysis

Similarly to spectral analysis, we have considered here 1-D spatial auto-correlation function $\rho_x(\Delta x)$ along x of each rain field $R(x, y)$ defined as follows:

$$\rho_x(\Delta x) = \frac{< [R(x,y) - m_{Rx}][R(x+\Delta x, y) - m_{Rx}] >}{\sigma^2_{Rx}} \qquad (5)$$

where the angle brackets stand for ensemble averaging of rain field along y, $m_{Rx}$ is the mean of $R$ along x and $\sigma_{Rx}$ is its standard deviation. Note that the numerator of (5) is the auto-covariance function of $R$. Previous equation is similarly defined for $\rho_y(\Delta y)$ along y. An example is shown in Fig. 7 for the same event of Fig. 6 where lag distance is in km.

The correlation scales $L_x$ and $L_y$ can be derived from $\rho_x(\Delta x)$ and $\rho_y(\Delta y)$ by adopting the following definition:

$$L_x = \int_0^\infty \rho_x(\Delta x) d\Delta x \qquad (6)$$

and similarly for $L_y$. Other definitions can be adopted as a percentage of the $\rho_x(\Delta x{=}0){=}1$, but (6) can offer some computational advantages if an exponential form of the correlation function is assumed, i.e. $\rho_x(\Delta x){=}\exp(-\Delta x)/L_x)$.

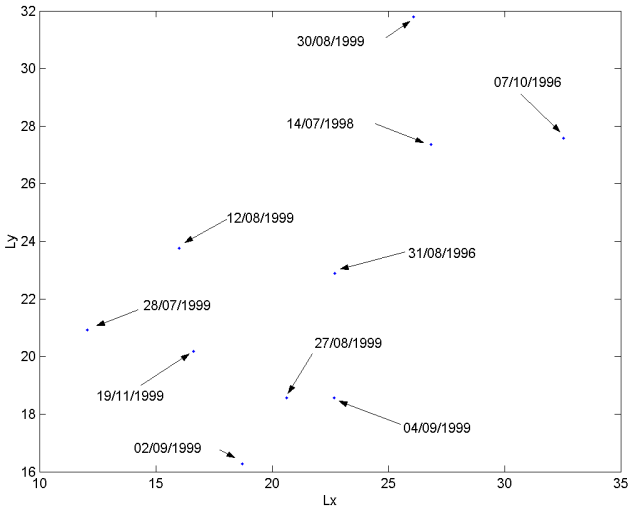


**Fig. 8a.** Mean correlation lengths (km) along x and y for some rainfall events (data 1996–1999 from S. Pietro Capofiume radar).

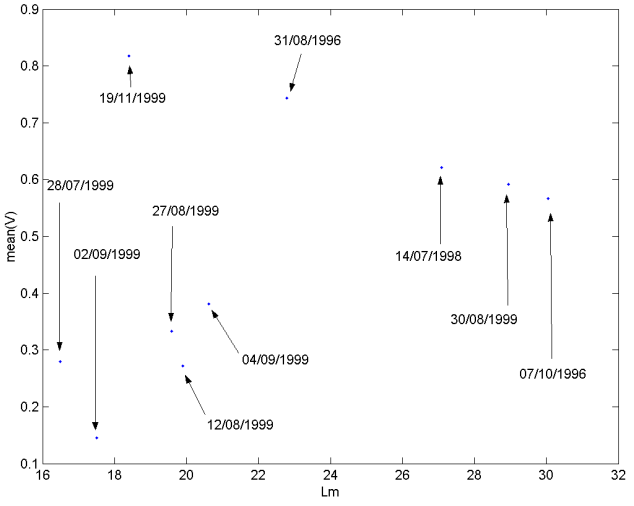


**Fig. 8b.** Variation coefficient vs. mean correlation length (h) for some rainfall events (data 1996–1999 from S. Pietro Capofiume radar).

The latter assumption is corroborated by the results in Fig. 7. Thus $\rho_x(\Delta x)$ can be estimated from the solution of the following transcendental equation (Ferraris et al., 2003):

$$L_x(1 - e^{-\Delta_{xM}/L_x}) = \int_0^{\Delta_{xM}} \rho_x(\Delta x) d\Delta x \qquad (7)$$

where $\Delta x_{xM}$ is the maximum lag where $\rho_x$ can be computed from data (in our case $\Delta x_{xM}$=128 km). From (7), average values of $L_x$ and $L_y$ equal to 18.7 km and 16.3 km have been found for the auto-correlation function of the event in Fig. 7.

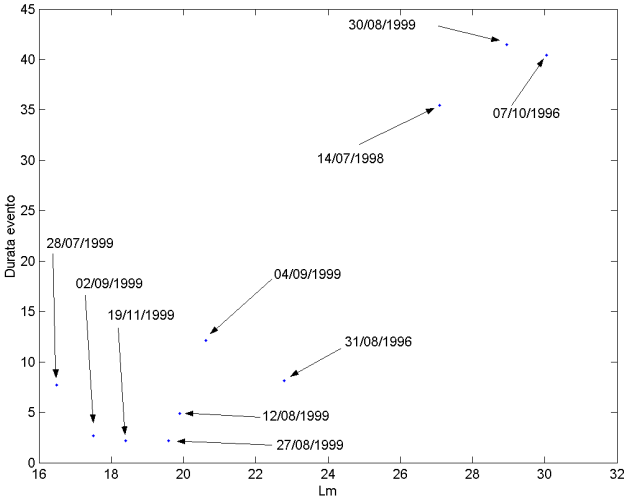


**Fig. 8c.** Event duration (h) vs. mean correlation length (km) for some rainfall events (data 1996–1999 from S. Pietro Capofiume radar).

### 3.3 Statistical analysis

As statistical indicators of $R$(x,y) of each rain event, we have considered the first and the second moment of its histogram, i.e. the mean value <R> and standard deviation $\sigma_R$. A threshold of 0.1 mm/h has been applied to select rain pixels, while the ensemble averaging is performed both in space and in time.

Moreover, for evaluating the relative variability of $R$(x,y) we have defined the variation coefficient $v_R$:

$$v_R = \frac{< R(x, y) >}{\sigma_R} \qquad (8)$$

Finally, the duration $T_d$ of the rain event has been included into the statistical description vector.

### 3.4 Results

The spectral, correlation and statistical analysis has been performed for all the events considered in the 3-year datasets. A subset of these is shown in Figs. 8a, b and c. In the latter two figures, for simplicity, we have plotted the average correlation length $L_m$=($L_x$+$L_y$)/2 in km. The label of each point indicates the date (day/month/year) of each rain event.

The analysis of Fig.8 tends to suggest a clustering of the events into two main groups. The first one, including the rain events 27/08/1999, 02/09/1999 and 04/09/1999, is characterized by mean correlation lengths below 20 km with mean rain rate (not shown) larger than 15 mm/h and a mean variation coefficient less than 0.6. The second group, including the rain events 07/10/1996, 14/07/1998 and 30/08/1999, denotes a mean correlation larger than 25 km, a mean rain rate (not shown) less than 5 and a variation coefficient less than 0.5. A third class may be also identified, represented by the event 31/08/1999, can be considered in the middle of the two previous groups.

The analysis of the power spectra for the same events reveals a mean values (along x and y) of $\beta \cong -2.4$ at k' and $\beta \cong -4.3$ at k" for the second group, and $\beta \cong -1.9$ at k' and $\beta \cong -2.4$ at k" for the first group. The average for all considered events gives $\beta \cong -2.1$ at k' and $\beta \cong -3.3$ at k".

From these preliminary results, the first class, as also confirmed by looking at mean duration equal to about 6.4 h (see Fig. 8c), could be classified as convective precipitation, while the second class with mean duration o about 39.1 h could be labelled as stratiform rainfall. The third class, having characteristics which are an average between the two, may be classified as hybrid and associated to scenarios where convective rain is embedded in stratiform precipitation. The visual inspection of each rain event evolution tend to confirm this preliminary supervised clustering.

## 4 Rainfall classification technique

Previous analysis suggests a supervised classification scheme where 3 classes can be envisaged, stratiform (S), convective (C) and hybrid (H).

In order to set up an automatic classifier, we need to defined the metrics. If a maximum likelihood method is adopted and a Gaussian statistics is chosen, then it can be shown that the objective function $d(c)$ to be minimized with respect to the class $c$ has the following form (e.g. Marzano et al., 2002):

$$d(c = S, C, H) = \sqrt{(\boldsymbol{x} - \boldsymbol{x}_c)^T \, \boldsymbol{S}_{xc}^{-1} \, (\boldsymbol{x} - \boldsymbol{x}_c)} \tag{9}$$

where $\mathbf{x}$ is the classification vector, $\mathbf{x}_c$ the centroid of the class $c$ and $\mathbf{S}_{xc}$ its auto-covariance. For simplicity, the latter can be also assumed diagonal by neglecting possible cross-correlations among the elements of $\mathbf{x}$ or even as an identity matrix by disregarding the variances of the elements of $\mathbf{x}$. The classification vector is defined as the set of previously analysed spectral, spatial and statistical parameters, that is:

$$\boldsymbol{x} = \begin{bmatrix} L_x \\ L_y \\ \beta' \\ \beta'' \\ < R > \\ v_R \\ T_d \end{bmatrix} \tag{10}$$

where $\beta\prime$ and $\beta\prime\prime$ stand for the power spectrum exponent $\beta$ computed at k$\prime$ and k$\prime\prime$.

By applying (9), preliminary results have shown that the convective class C has, for instance, as centroid elements the following values: $L_{xC} \cong 19.6$ km, $L_{xC} \cong 18.4$ km, $\beta\prime_C \cong -2.4$, $\beta\prime\prime_C \cong -4.3$, $< R >_C \cong 4.3$ mm/h, $v_{RC} \cong 0.6$ and $T_{dC} \cong 6.4$ h. To the elements of the centroid of the stratiform class S can be attributed the following values: $L_{xS} \cong 28.5$ km, $L_{xS} \cong 28.9$ km, $\beta\prime_S \cong -1.9$, $\beta\prime\prime_S \cong -2.4$, $< R >_S \cong 14.5$ mm/h, $v_{RS} \cong 0.3$ and $T_{dS} \cong 39.1$ h. A tentative definition of the hybrid class may be associated to a centroid with: $L_{xH} \cong 22.7$ km, $L_{xH} \cong 22.5$ km, $\beta\prime_H \cong -2.1$, $\beta\prime\prime_H \cong -3.3$, $< R >_H \cong 10.5$ mm/h, $v_{RH} \cong 0.7$ and $T_{dH} \cong 9.1$ h.

As a result, by applying the automatic classification procedure to rain events not considered in the supervised training, we can, for instance, categorize the rain events on 28/07/1999, 12/08/1999 and 19/11/1999, shown in Fig. 8, as convective ones.

## 5 Conclusions

A three-year rainfall dataset derived from a C-band dual-polarized radar, placed in S.Pietro Capofiume near Bologna (Italy), has been processed and systematically analysed. The data set consists of radar-derived rain intensities on a grid of 256 by 256 km with a spatial resolution of 1 km and a time sampling of 15 min from 1996 to 1999. A rain event selection has been systematically carried identifying more than 40 events of interests.

A preliminary spatial characterization of the rain fields has been performed by means of computing power spectrum slopes, spatial correlation lengths, mean rain intensity and variation coefficient, event duration. This quantitative investigation has given some insights into rainfall characterization, putting in evidence the impact of rain regime on the parameters values. From these results, a supervised classification scheme, based on a Gaussian metrics , has been finally proposed in order to categorize in an automatic way stratiform, convective and hybrid rain events.

These results should be considered only preliminary as a conclusive data quality control and sensitivity analysis to different analysis procedures has been not concluded yet. Some fundamental issues, such as the hypotheses of a stationary field implicit in the definition of power spectra and autocorrelation functions, should be more deeply investigated. Future work will be also devoted to the analysis of the logarithm of rain-rate fields and to the overall evaluation of a 6-year dataset, where more than 100 cases are expected. Implications for rainfall space-time synthetic generators will be also investigated.

*Acknowledgement.* This work has been funded by CNR-GNDCI project under RAM research line.

## References

Alberoni, P. P, Andersson, T., Mezzasalma, P., Michelson, D. B, and Nanni, S.: Use of the vertical reflectivity profile for identification of anomalous propagation, Meteorol. Appl., 8, 257–266, 2001.

Austin, P. M. and Houze, R. A.: Analysis of structure of precipitation patterns in New England, J. Appl. Meteor., 11, 926–934, 1972.

Bell T. L.: A space-time stochastic model of rainfall for satellite remote sensing studies, J. Geophys. Res., 92, n. D8, 9631–9643, 1987.

Crane, R. K.: Space-time structure of rain rate fields, J. Geophys. Res., 95, 2001–2020, 1990.

Ferraris, L., Gabellani, S., Rebora, N., and Provenzale, A.: A comparison of stochastic models for spatial rainfall downscaling, Water Resources Res., 39, 12, 1368, doi:10.1029/2003WR002504, 2003.

Gremont, B. and Filip, M.: Spatio-temporal attenuation model for application to fade mitigation techniques, IEEE Trans. Ant. Propagat., 52, 5, 1245–1256, 2004.

Marzano, F. S., Fionda, E., Ciotti, P., and Martellucci, A.: Ground-based multi-frequency microwave radiometry for rainfall remote sensing, IEEE Trans. Geosci. Rem. Sens., 40, 742–759, 2002.

Marzano, F. S., Picciotti, E., and Vulpiani, G.: Rain field and reflectivity vertical profile reconstruction from C-band radar volumetric data, IEEE Trans. Geosci. Remote Sens., 42, 4, 1033–1046, 2004.

Meija, J. and Rodriguez-Iturbe, I.: On the synthesis of random fields sampling from the spectrum: an application to the generation of hydrologic spatial processes, Water Resources Res., 10, 705–711, 1974.

Sauvageot, H.: Radar Meteorology, Artech House, Norwood (MA), Boston-London, 1992.

Perica, S. and Foufoula-Georgiou, E.: Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions, J. Geophys. Res., 101, D21, 26 347–26 361, 1996.

Veneziano, D., Bras, R. L., and Niemann, J. D.: Nonlinearity and self-similarity of rainfall in time and a stochastic model, J. Geophys. Res., 101, D21, 26 371–26 392, 1996.

Wheater, H. S. et al.: Spatial-temporal rainfall fields: modelling and statistical aspects, Hydrology and Earth Syst. Science, 4, 581–601, 2000.

Willems, P.: A spatial rainfall generator for small spatial scales, J. Hydrology, 52, 126–144, 2001.