Analysis and Synthesis of Raindrop Size Distribution Time Series From Disdrometer Data

Mario Montopoli, Frank Silvio Marzano, *Senior Member, IEEE*, and Gianfranco Vulpiani, *Member, IEEE*

***Abstract*—Hydrometeorological and radio propagation applications can benefit from the capability to model the time evolution of raindrop size distribution (RSD). A new stochastic vector autoregressive semi-Markov model is proposed to randomly synthesize (generate) the temporal series of the three driving parameters of a normalized Gamma RSD. Rainfall intermittence is reproduced through a discrete semi-Markov process, modeled from disdrometer measurements using two-state analytical statistics of rain and dry period duration. The overall model is set up by means of a large set of disdrometer measurements, collected from 2003 to 2005 at Chilbolton, U.K. The driving parameters of the retrieved RSD are estimated using three approaches: the Gamma moment method and the 1-D and 3-D maximum-likelihood methods. Interestingly, these methodologies lead to quite different results, particularly when one is interested in evaluating RSD higher order moments such as the rain rate. The accuracy of the proposed RSD time-series generation technique is evaluated against available disdrometer measurements, providing excellent statistical scores.**

***Index Terms*—Autoregressive process, disdrometer, estimation techniques, intermittent rain process, raindrop size distribution (RSD), semi-Markov chain.**

I. Introduction

MEASUREMENTS of raindrop size distribution (RSD) are largely exploited to investigate the microphysics of precipitation and to improve rainfall remote-sensing estimation techniques. One major source of uncertainty in the precipitation remote-sensing methodologies is the variability of the RSD [1]. This variability is due to both climatological context and storm cyclogenesis, and it strongly affects the relationship between the weather radar reflectivity and the rain rate [2]. Indeed, accurate RSD measurements, coupled with a backscattering model, allow to simulate not only radar reflectivity but also all radar polarimetric observables and quantities of interest such as water content and rain rate [3]–[6]. On the other hand, accurate RSD data are required not only for hydrometeorological applications. Since rain attenuation is strongly correlated with rain rate intensity, the detailed analysis of rainfall size–time distribution is crucial for satellite communication design at high microwave frequencies and spaceborne meteorological product validation [7], [8]. In this respect, several efforts are currently devoted to the development of techniques to randomly synthesize (or generate) rain rate time-space series [9], reproducing real rain rate measurement properties in terms of statistics and intermittence features [10]. This capability is essential to set up radio system simulators and numerically test them under a large set of synthetic, but realistic, rainfall scenarios.

Early ground-based RSD measurements used filter paper (or the so-called flour method). One of the current most widespread techniques is the disdrometer such as the Joss–Waldvogel instrument. In the last decade, there has been the introduction of new disdrometer techniques like the optical spectropluviometer or the recent 2-D video disdrometer. Above the ground, RSD observations may be obtained from particle-measuring systems optical array probes on aircraft. A common difficulty with all of these disdrometers is related to the small collection area of the instrument, which makes the drop counting in each diameter class very noisy, particularly for large drops characterized by the smallest concentration. Moreover, the estimation of best fitting analytical RSD parameters may be affected by these instrumental features, and, in this respect, various RSD estimation techniques may have a different behavior and a different level of accuracy with respect to a specific purpose, e.g., estimate of the RSD shape, the rain rate, rain water contents, or specific attenuation.

The objectives of this paper are manifold.

1) Characterize a large set of Joss–Waldvogel impact disdrometer measurements, collected from 2003 to 2005 at Chilbolton, U.K.
2) Retrieve from this disdrometer data set the driving parameters of the normalized Gamma RSD and perform a sensitivity analysis of these results by using different best fitting techniques.
3) Exploit the correlation structure of the estimated RSD parameters to extend and further investigate the properties of the vector autoregressive (VAR) stationary model proposed in [9] to simulate time series (or horizontal profiles) of RSDs and, consequently, of either the rain rate or the path attenuation.
4) Characterize the distribution of the inter-rain duration (or dry periods DP) and rain duration (or wet periods WP) to design a simple semi-Markov chain to represent the intermittence feature of the rainfall process.

Manuscript received February 19, 2007; revised June 26, 2007. This work was supported in part by the Italian Ministry of University and Research, Region Abruzzo, and the Italian Department of Civil Protection.

M. Montopoli is with the Department of Electrical Engineering and Information, and also with Centro di Eccellenza per l'intagrazione di tecniche di Telerilevamento e Modellistica per la Previsione di eventi meteorologici Severi (CETEMPS), University of L'Aquila, 67040 L'Aquila, Italy (e-mail: mmontop@ing.univaq.it).

F. S. Marzano is with the Department of Electronic Engineering, Sapienza University of Rome, 00184 Rome, Italy, and also with CETEMPS, University of L'Aquila, 67040 L'Aquila, Italy (e-mail: marzano@die.uniroma1.it).

G. Vulpiani was with CETEMPS, University of L'Aquila, 67040 L'Aquila, Italy. He is now with the Prime Ministry, Department of Civil Protection, 00189 Rome, Italy (e-mail: gianfranco.vulpiani@protezionecivile.it).

Digital Object Identifier 10.1109/TGRS.2007.909102

The overall stochastic procedure to randomly synthesize (or generate) RSD time series is named the VAR raindrop Markov synthesizer (VARMS) model.

This paper is organized as follows. In Section II, the disdrometer measurement characteristics and the data processing procedure are described. In Section III, the retrieval methods for estimating the RSD parameters are shown, whereas in Section IV, the retrieved parameters are exploited to define the RSD stochastic simulation VARMS model. Finally, in Section V, conclusions are given.

II. Disdrometer Data

The Joss–Waldvogel disdrometer (JWD) measurements, collected from 2003 to 2005 at Chilbolton, U.K., and provided by the British Atmospheric Data Centre (BADC), have been thoroughly analyzed. A description of the instrument together with its data processing is illustrated in Appendix A.

The measured RSD at the discrete instant t (in seconds) has been calculated from JWD counts using the following equation:

$$N_m(D_i,t)=\frac{n_i(t)}{A\cdot\Delta t\cdot\nu_i\cdot\Delta D_i} \tag{1}$$

where the subscript m indicates a measured quantity; D_i (in millimeters) is the central raindrop diameter of the "channel" c_i, which has been computed as the mean value between the drop diameters D_{ci} and D_{ci+1}; $N_m(D_i,t)$ (in units per millimeter per cubic meter) is the number of raindrops per unit of volume in the channel c_i at the discrete instant t; $n_i(t)$ is the number of drops reckoned in the ith channel at the instant t; A (in square meters) is the sensor area; Δt (in seconds) is the sample period; ν_i (in meters per second) is the fall speed of a raindrop, in which the diameter is given by D_i; and ΔD_i (in millimeters) is the ith channel width. The parameters setup to calculate $N_m(D_i,t)$ from (1) are as follows: $A=0.005$ m^2, $\Delta t=10$ s, $\Delta D_i=D_{ci+1}-D_{ci}$ for i ranging from 1 to $n_c=104$ with minimum and maximum diameter equal to 0.5 and 5 mm, respectively, uniformly spaced in a logarithmic scale (see Appendix A). The rainfall speed ν_i is assumed to be the same as that proposed in [18], i.e.,

$$\nu_i=3.78\cdot D_i^{0.67}. \tag{2}$$

At Chilbolton, JWD measurements of $N_m(D_i,t)$ were available every $\Delta t=10$ s. This time series may exhibit fluctuations in D_i and t. To reduce the intrinsic noise of N_m with respect to D_i at each time step, a moving average filter on N_m with a span of five raindrop diameters D_i has been applied over the entire range of $n_c=104$ samples (or channels). Moreover, to smoothen the temporal fluctuations of N_m, the latter has been subsequently averaged, as done in [19], over a time interval of 2 min. Indeed, it is not easy to find a satisfactory compromise for the integration time: If long, it may smoothen the actual physical variations; if short, the observed RSD may be dominated by counting fluctuations. The interval of 2 min seems, in this case, to be a reasonable choice.

From (1) and (2) and after time–size integration, the computation of the measured rain rate R_m is straightforward through the moments of $N_m(D_i,t)$ of order 3.67, as specified by

$$R_m(t)=3.78\cdot\frac{\pi}{6}\cdot m_{3.67}(t) \tag{3}$$

where $m_n(t)$ is the general expression of $N_m(D_i,t)$ moment of order n, defined as

$$m_n(t)=\int_0^\infty D^n\cdot N(D,t)\cdot dD=\sum_{i=1}^{n_c}D_i^n\cdot N_m(D_i,t)\cdot\Delta D_i. \tag{4}$$

In (4), the third term underlines the discrete nature of the measured RSD, whereas the extremes of the integral point out the untruncated range of diameters.

From (3), the beginning of a rainy event has been defined as the instant where $R_m(t)$ exceeds the threshold of 0.1 mm $\cdot$ h^{-1} at least for 12 min, and the corresponding end has been defined as the instant where $R_m(t)$ is lower than the aforementioned threshold. The duration of a rainy event implicitly defines a wet period, and, accordingly, the time interval separating two consecutive WP's defines a dry period DP. By selecting only the 24-h measurements in which at least one WP takes place, a set of 224 days (corresponding to 161 280 samples, spaced each other of 2 min), among the available 785, has been obtained. Within these 224 days, we have identified 656 rain events or WP's, corresponding to 15 710 samples (about 10% of the total number of samples).

III. Analysis of the RSD Data

A systematic comparison among the three distinct approaches for best fitting the observed RSD and estimating the parameters of an assumed normalized Gamma distribution is discussed in this section. Many authors adopt the following expression to describe the volumetric size distribution of raindrops [1], [20], [21]:

$$N(D,\mathbf{p})=N_w\cdot f(\mu)\cdot\left(\frac{D}{D_m}\right)^{\mu}\cdot\exp\left[-(4+\mu)\cdot\frac{D}{D_m}\right] \tag{5}$$

where $N(D,\mathbf{p})$ (in units per cubic meter per millimeter) is the number of drops per unit volume per unit size interval; D (in millimeters) is the sphere-equivalent drop diameter; N_w (in units per cubic meter per millimeter), μ, and D_m (in millimeters) are the intercept, the shape, and the mass weighted mean diameter parameters, respectively; and $f(\mu)$ takes the following form:

$$f(\mu)=\frac{6}{4^4}\cdot\left[\frac{(4+\mu)^{(4+\mu)}}{\Gamma(4+\mu)}\right] \tag{6}$$

where Γ is the complete Gamma function. The vector $\mathbf{p}=[N_w,D_m,\mu]^T$, with "$T$" referring to matrix transpose, stands for the *time-dependent* parameter set of the modeled RSD.

In general, the parameters N_w, μ, and D_m can be retrieved using the moments of the RSD such as, for example, the second,

the third, the fourth, and the sixth. The mass weighted mean diameter D_m is calculated as the ratio between the fourth and third empirical moments of the RSD, i.e.,

$$D_m = \frac{m_4}{m_3}. \tag{7}$$

The generalized intercept parameter N_w is also computed from the fourth and third moments of the RSD and can be derived from

$$N_w = \frac{256}{6} \cdot \frac{m_3^5}{m_4^4} \tag{8}$$

where the moment m_n of order n is expressed by (4).

In the next two subsections, the expression of the shape parameter μ in (5) is first estimated by means of the Gamma moment method [24] and then by means of the maximum-likelihood (ML) method in two versions [25]. Indeed, in the second version of the ML method, the parameters N_w and D_m are differently estimated with respect to (7) and (8).

A. GM Estimation Method

For each 2-min averaged RSD sample, the μ parameter has been computed by following the procedure suggested by Ulbrich and Atlas [24] for an untruncated range of diameters. Following that procedure, the shape parameter μ can be resumed from

$$\mu_{\mathrm{GM}} = \frac{(7-11\eta) - \sqrt{(7-11\eta)^2 - 4\cdot(\eta-1)\cdot(30\eta-12)}}{2\cdot(\eta-1)} \tag{9}$$

where η depends from the moments m_2, m_4, and m_6, i.e.,

$$\eta = \frac{m_4^2}{m_2 \cdot m_6}. \tag{10}$$

Note that the temporal dependence is omitted in (7)–(10) for simplifying the notation.

Thus, for the GM method, the estimated RSD parameter vector is given by

$$\mathbf{p}_{\mathrm{GM}}(t) = [N_{w\,\mathrm{GM}}(t), D_{m\,\mathrm{GM}}(t), \mu_{\mathrm{GM}}(t)]^T \tag{11}$$

where $N_{w\mathrm{GM}}$ and $D_{m\mathrm{GM}}$ are retrieved from (7) and (8), respectively, and the time dependence of each RSD parameter (and, thus, of $\mathbf{p}$) has been explicitly indicated.

It is worth observing that the values of μ can reach infinity when η is close to unity. This circumstance occurs when the measured RSD is nearly monodisperse. In fact, if we model the monodisperse condition replacing the measured RSD $N_m(D_i, t)$ with a shifted delta of Dirac $\delta(D_i - D_k)$ and solving the integral in (4), we obtain, by means of (10), $\eta = 1$ and, correspondingly, $\mu = \infty$. The increase of μ for the decreasing spread of RSD is confirmed by the trend of the average standard deviation of the drop diameters D_i versus the interval of μ ($\Delta\mu$), as shown in Fig. 1. The intervals $\Delta\mu$ are spaced each other of 1 unit when μ range from -1.6 to 60. It is worth mentioning that high values of μ take place when the measured intensities of the rain rate R_m lie below 1 mm $\cdot$ h^{-1}. We remark that, according to (4), this paper deals with an untruncated range of diameters, whereas in [24], a truncated version of the GM method is also introduced. A further extension of the truncated method proposed in [24], called the truncated Gamma moment (TGM) method, has been also considered, as described in [29]. Numerical results indicate that the TGM estimates of the RSD parameters are highly correlated with the GM ones (for N_w and D_m, the correlation coefficient is 0.97, whereas for μ estimates, it is 0.89). Considering the simplicity and the efficiency of the GM formulation, we have limited our attention to GM only in this paper.

To have a comprehensive picture of RSD best fitting, the scaled data $N_m(D_i)/N_w$ have been plotted versus the normalized drop diameter D/D_m in Fig. 2. As in [22], superimposed

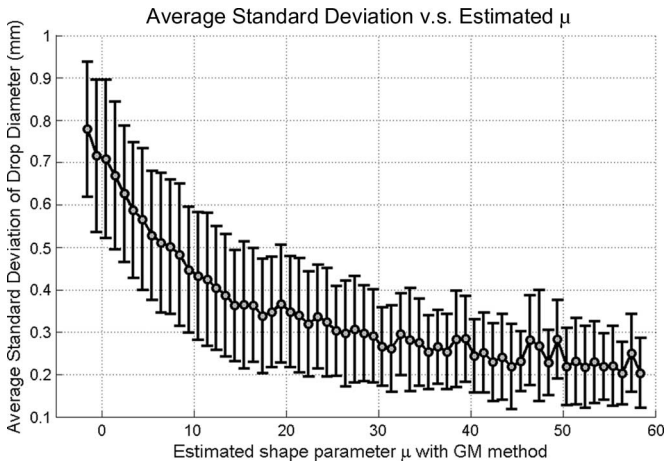

Fig. 1. (Filled circle) Estimated μ by means of moment method versus average standard deviation of drop diameters of measured RSDs. The vertical bars indicate the $\pm$ standard deviation with respect to the average trend.

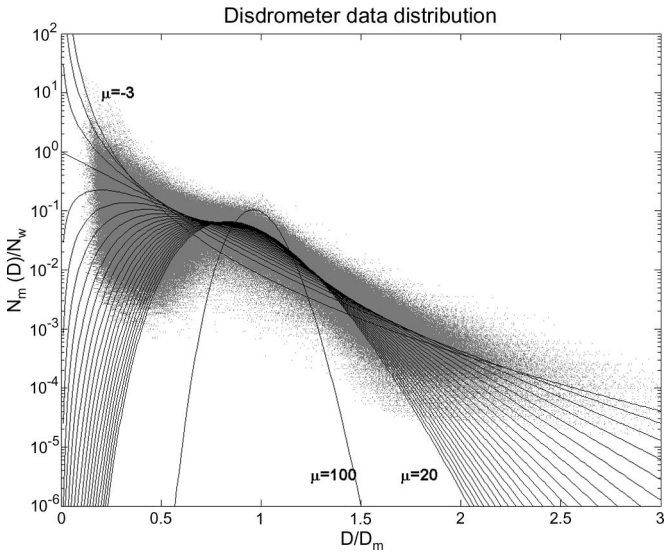

Fig. 2. Gray dots indicate the scaled RSD $(N_m(D)/N_w)$ versus normalized diameter (D/D_m) of the 15 710 measured samples. Black solid lines indicate the normalized gamma distributions for values of μ ranging between -3 and 20 at step of 1 plus the curve at $\mu = 100$.

on the scaled disdrometer data, the corresponding scaled Gamma distributions for some values of μ are shown. The measured RSDs, when scaled as previously shown, are well bounded by the family of scaled Gamma functions as μ varies over the range from -3 to 20; however, values of μ outside this range may be also represented.

B. ML Estimation Methods

In past works, various alternative methods for estimating the three parameters of Gamma distribution $N(D)$ in (5) were proposed. For example, in [22], the parameters N_w and D_m were estimated by means of the moment method, whereas the shape parameter μ was estimated by minimizing the absolute deviation between the scaled RSD data and the scaled Gamma form $f(\mu) \cdot x^{\mu} \cdot \exp[-(4+\mu) \cdot x]$.

In this section, we remove this restriction by enabling the three parameters in (5) to range within a prescribed bound to minimize, at each discrete instant t, an error measure between the measured RSD $N_m(D_i, t)$ and the Gamma distribution $N(D)$. We label this 3-D estimation technique as ML3 for distinguishing it from the 1-D approach described after.

We can estimate the RSD parameter vector $\mathbf{p}(t)$ at each instant t by minimizing, in a likelihood sense [23], the following expression:

$$\mathbf{p}_{\mathrm{ML3}}(t) = \min_p \left\{ \sum_{i=1}^{n_c} \left[N_m(D_i, t) - N(D_i, \mathbf{p}(t)) \right]^2 \right\} \quad (12)$$

where the time dependence of $\mathbf{p}$ within the expression (5) of N has been pointed out, D_i is the discrete drop diameter, and $\min_p$ is an operator returning the value of $\mathbf{p}$ corresponding to the minimum of its argument.

The ML approach, as already mentioned, can be applied to estimate only the parameter μ, exploiting the GM method to evaluate the other parameters N_w and D_m [22]. Using the formalism of (12), μ is derived as follows:

$$\mu_{\mathrm{ML1}}(t) = \min_{\mu} \left\{ \sum_{i=1}^{n_c} \left[N_m(D_i, t) - N(D_i, \mu(t)) \right]^2 \right\}. \quad (13)$$

Consequently, the parameter vector $\mathbf{p}$ using the hybrid 1-D ML technique (ML1) is

$$\mathbf{p}_{\mathrm{ML1}}(t) = \left[N_{w\mathrm{GM}}(t), D_{m\mathrm{GM}}(t), \mu_{\mathrm{ML1}}(t) \right]^T \quad (14)$$

where $D_{m\mathrm{GM}}$ and $N_{w\mathrm{GM}}$ are derived from (7) and (8), respectively.

C. Results

The goal of this section is to describe the differences and the peculiarities of the RSD estimation methods so far exposed. Past works have already discussed the errors associated to the GM methods in estimating the RSD parameters of a Gamma distribution on the basis of computer simulations (e.g., [26] and [27]). In this section, we focused our attention to compare the GM estimation method with the ML1 and ML3 methods.

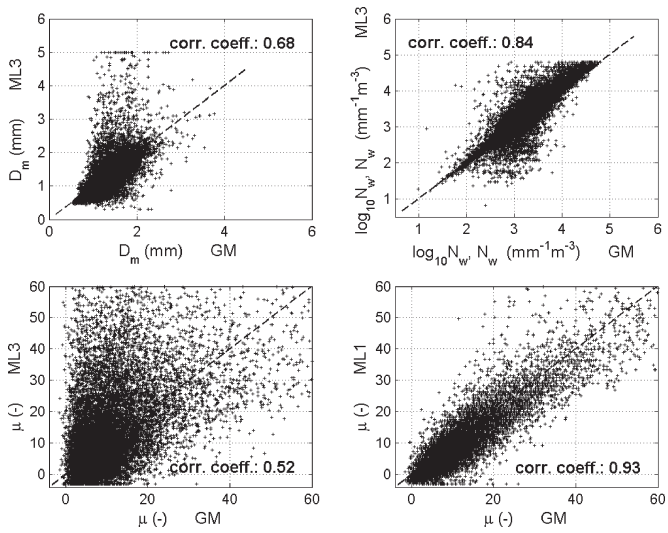

Fig. 3. Scatter plots of the 15 710 estimated RSD parameters by means of the moment method GM and the ML methods ML3 and ML1.

The latter comparison has been also tackled in [28] on different disdrometric measurement.

Using the whole RSD time series available, Fig. 3 shows the scatter plots of the estimated parameters N_w, D_m, and μ for all the 15 710 samples using the GM, ML1, and ML3 methods. These fitting methods have been tested on the entire data set and not only on selected homogeneous RSD sequences since, as it will be explained in the next section, our final goal is to reproduce some overall characteristics (e.g., the probability density function (pdf) and the correlation properties of wet periods and the intermittence properties of the rainfall phenomena) of the measured disdrometer data.

Fig. 3 is completed by Fig. 4, which shows the distribution of the three RSD parameters, obtained for the GM, ML1, and ML3 methods, whereas the main statistical indicators of these distributions have been listed in Table I. The plots of Fig. 3 show a dispersion around the bisector line, indicating that each implemented method provides a different estimate of the Gamma RSD parameters. However, all methods tend to agree on the result that the most probable values of N_w, D_m, and μ in a typical southern British climate are about 10^3 to $10^{3.5}$ $\mathrm{m}^{-3} \cdot \mathrm{mm}^{-1}$, 1 mm, and 0–5, respectively. Because ML1 differs from GM only for the estimation of μ [see (11) and (14)], in Figs. 3 and 4, we limit the comparison to this parameter only. The GM–ML1 correlation diagram denotes a high correlation coefficient of μ of about 0.9 (see the lower right panel of Fig. 3). On the contrary, the correlation coefficient results to be significantly less (about 0.5–0.7) when we compare the estimates of μ and D_m from GM and ML3 (see left panels of Fig. 3). Indeed, the comparison between GM and ML3 tends to show relevant differences also for the other parameter N_w, although it shows a high correlation coefficient of about 0.8 (see the upper right panel of Fig. 3).

It is worth stating that a statistical χ test, based on the analysis of the sum of the square differences (SSD) between $N_m(D_i, t)$ and $N(D_i, \mathbf{p}_X)$, where X labels the adopted estimation method (i.e., GM, ML3, and ML1), has proved for every discrete instant t that the SSD of ML3 is always smaller than

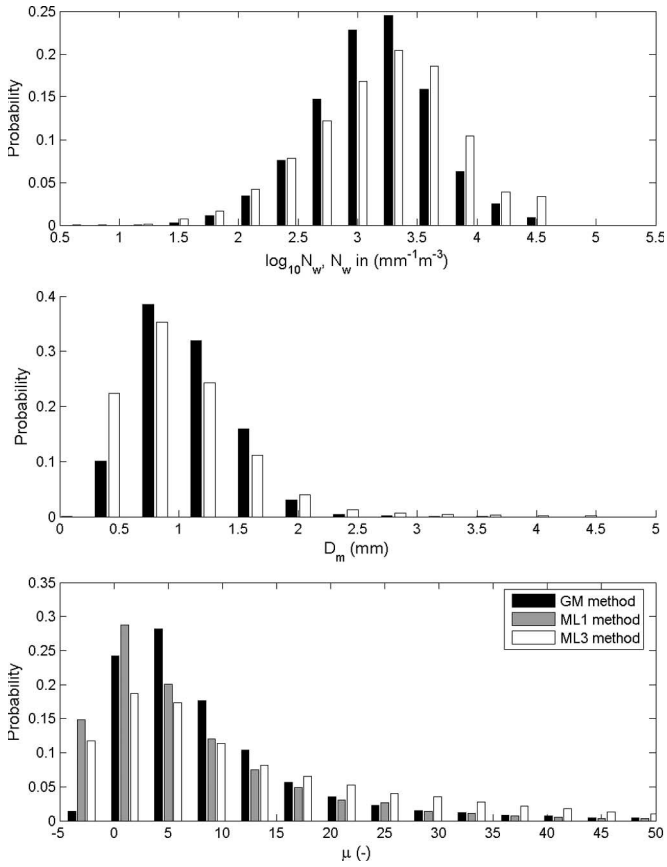

Fig. 4. Histograms of estimated parameters (top panel) $\log_{10}(Nw)$, (middle panel) D_m, and (bottom panel) μ by (gray bins) the ML1 method, (white bins) the ML3 method, and (black bins) the GM method.

TABLE I
MEAN AND STANDARD DEVIATION OF RSD PARAMETER DISTRIBUTIONS

	GM		ML1		ML3	
	Mean	std	mean	std	mean	std
$\log_{10}(N_w)$	3.27	0.50	3.27	0.50	3.36	0.60
D_m	1.26	0.38	1.26	0.38	1.22	0.55
μ	12.44	15.31	10.50	15.71	16.81	20.69

N_w is expressed in ($mm^{-1}m^{-3}$), D_m in (mm), and μ is adimensional.

that of ML1 and much smaller than that of GM. This result is, in a way, expected, as ML3 is based on the minimization of the SSD with respect to all parameters, as in (12). On the other hand, it means that the ML3 estimates of the RSD parameters always provide a better fit of the measured RSDs with respect to the other methods. To underline the discrepancies among the estimation methods, some examples of measured RSDs together with the best fitted Gamma distributions $N(Di, \mathbf{p}_{\rm GM})$, $N(D_i, \mathbf{p}_{\rm ML1})$, and $N(D_i, \mathbf{p}_{\rm ML3})$ are shown in Fig. 5. In each subplot of this figure, the values of the rain rate both from the measured RSDs, namely, R_m, using (3) and from the estimated RSDs, namely, R_X, where the $m_{3.67}$ moment has been derived from the best fitted RSD using the three methods, i.e., from $N(D_i, \mathbf{p}_X(t))$, have been emphasized. The plots of Fig. 5 highlight how the ML3-estimated RSD curves tend to follow the measured RSDs very closely, whereas the best fitted RSDs from ML1 and GM show the tendencies to underestimate the peaks

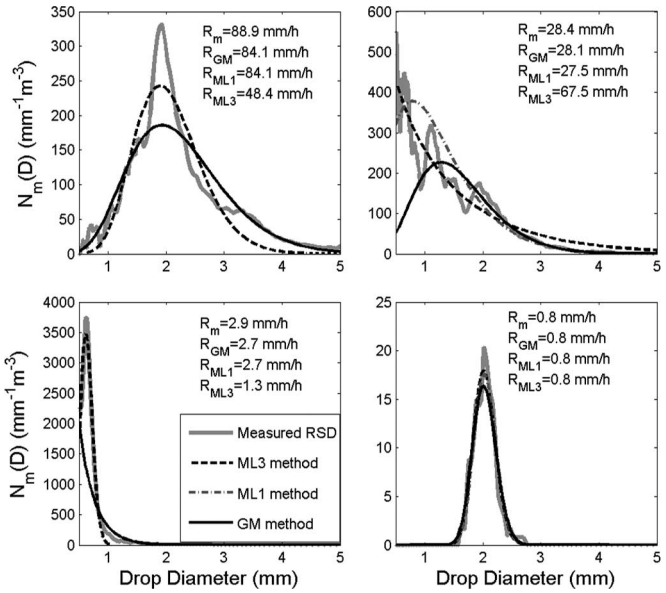

Fig. 5. Examples of Gamma RSD fits at four different instants. Gray line represents the measured RSD from the disdrometer, whereas the black solid line, gray shade, and dashed black stand for the Gamma fit with parameters estimated by means of GM, ML1, and ML3, respectively.

(this behavior is more accentuated for GM). On the other hand, it seems that the ML3-estimated curves, when compared to GM and ML1 and particularly for high values of R_m, are prone to the underestimation of the measured RSD right tails. This could be expected because the ML3 method takes into account the entire RSD, whereas the ML1 and GM methods focus on high-order moments that mainly depend on the tail of the RSD. The latter correspond to larger raindrop diameters. However, cases where the ML3 method overestimates the measured large raindrop diameters and, accordingly, the rain rate are possible (see, for example, the upper right panel of Fig. 5).

To quantify the discrepancies between the Gamma RSD estimates and the measured RSDs, the rmse between $N_m(D_i, t)$ and each retrieved RSD, expressed as $N(D_i, \mathbf{p}_X(t))$, for each diameter (channel) D_i, has been computed as follows:

$$\text{rmse}_X(D_i) = \left[\frac{1}{n_s}\sum_{j=1}^{n_s}\left[N_m(D_i, t_j) - N(D_i, \mathbf{p}_X(t_j))\right]^2\right]^{\frac{1}{2}} \quad (15)$$

where $n_s = 15\,710$ is the number of discrete instants t_j, obtained from the disdrometer measurements, and the index i ranges from 1 to the number of channels n_c (see Section II).

Fig. 6 shows the behavior of rmse_X, computed by (15), versus the discrete drop diameters D_i. Values of $\text{rmse}_{\rm ML3}$ smaller than $\text{rmse}_{\rm ML1}$ and $\text{rmse}_{\rm GM}$ are noted up to D_i about 0.9 mm; however, there are larger values of $\text{rmse}_{\rm ML3}$ for larger diameters beyond about 1 mm. For the latter large raindrop diameters, the curves of $\text{rmse}_{\rm ML1}$ and $\text{rmse}_{\rm GM}$ appear very close to each other, whereas for diameters lower than 0.9 mm, $\text{rmse}_{\rm ML1}$ is positioned between $\text{rmse}_{\rm ML3}$ and $\text{rmse}_{\rm GM}$. The analysis of Fig. 6 would imply that the ML3 estimation method is an accurate technique for the overall measured RSD best

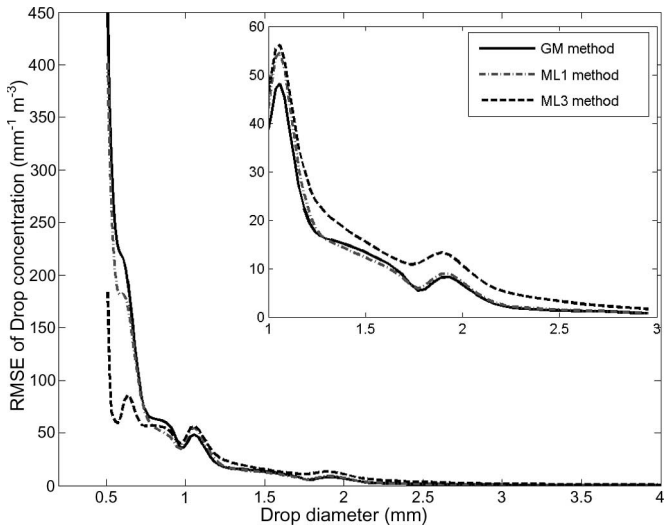

Fig. 6. RMSE between the whole of 15 710 measured RSDs N_m and the fitted ones N for each channel D_i using the GM, ML1, and ML3 methods.

fitting, but not a good candidate for the computation of higher moments of RSD where the larger diameters play a major role.

To confirm the previous consideration, the rain rates R_m, $R_{\rm GM}$, $R_{\rm ML1}$, and $R_{\rm ML3}$ have been compared in terms of correlation coefficients and rmse indicator. The latter has been obtained by replacing, in (15), N_m and N with R_m and R_X, respectively. We found values of 0.99, 0.99, and 0.78 for the correlation coefficient and 0.14, 0.15, and 1.79 for the rmse for GM, ML1, and ML3, respectively. As expected, nevertheless, the ML3 method provides the best overall fitting of measured RSDs (see Figs. 5 and 6); the respective estimated rain rate is not in good agreement with the measured one. The failure of the ML3 method to reproduce RSD high-order moments underlines that, in most of the considered cases, the conjectured Gamma distribution may not be the best choice for representing the tails of the measured RSDs. As shown by Figs. 3–6, a good compromise in terms of RSD shape and moment best fitting is the use of the ML1 approach expressed by (14) because N_w and D_m are estimated using relatively high-order moments (mainly controlled by the tail of the RSD). A tentative to use other RSD shapes such as the lognormal and Weibull RSDs has been carried out without obtaining better results. The optimization of the RSD analytical function form for the Chilbolton JWD data set is worth of further investigation, which is, however, beyond the scope of this paper.

IV. Synthesis of RSD Time Series

Given the stochastic nature of the RSD, rainfall may be thought of as a time signal of a random process [30]. It is usually characterized by a significant space-time variability; therefore, within a rainfall process, we can basically distinguish two different macroscopic phenomena: the rain or wet periods WP and the no-rain or dry periods DP.

To model this "intermittent" behavior, the rainfall random process is supposed to be a renewal process, which is a generalization of the Poisson process, the latter being a continuous-time Markov process [31]. Within each WP, the time series

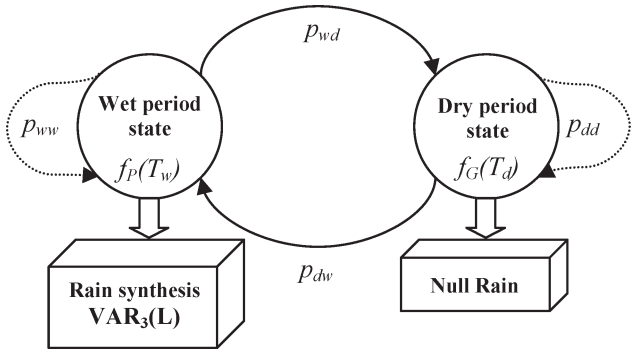

Fig. 7. General scheme of VARMS procedure, modeled as a semi-Markov chain between two states with a holding time distribution f_P and f_G and a transition probability p_{ij}.

of RSD parameters $\mathbf{p}(t)$ is modeled as a VAR process of order L (i.e., the future time behavior is conditioned by L past time), thus taking into account the correlation properties of the RSD parameters. Indeed, the problem of modeling the RSD parameters in the WP's with a VAR process has been already tackled and described in [9] for $L = 1$ and for a different parameterization of the RSD with respect to that used in this paper. Here, we propose an extension and a generalization of the approach adopted in that work.

On the other hand, within each DP, the void values of RSD parameters $\mathbf{p}(t)$ have been considered to describe the no-rain events. The overall procedure, named VARMS, is fully described in the next subsections.

A. Intermittent Rainfall Random Process

The intermittent rainfall process can be modeled as a discrete state machine switching between DP and WP states (see Fig. 7). This process may be modeled as a Markov chain, which is defined as a discrete-time stochastic process with the Markov property [32]. A Markov chain describes at successive times the states of a system. The changes of state are called transitions. The Markov property means that the conditional probability distribution of the state in the future, given the state of the process currently and in the past, depends only on its current state and not on its state in the past. In a word, a Markov chain is memoryless. It is worth remarking that a Markov model imposes the state duration to follow an exponential law [32]. In particular, in a Markov chain, the probability that the machine stays in the particular state k for a duration T is equal to $f(T) = p_{kk}^{T-1}(1 - p_{kk})$, with p_{kk} the probability to remain in state k during two consecutive samples. The function $f(T)$ represents then an exponential pdf of duration T in state k. To remove this limitation when describing the rainfall intermittence, in the VARMS approach, we have adopted a semi-Markov chain with the two states the DP and WP states.

A semi-Markov chain is used to enable any possible duration T_d for the DP state and T_w for the WP state, described by arbitrary pdfs $f(T_d)$ and $f(T_w)$, respectively (see Fig. 7). Table II lists the moments of measured WP and DP duration pdfs, derived from the rain events defined as in Section II. From this table, it is clear that the exponential pdf is only a crude

TABLE II
MOMENTS OF THE MEASURED WET AND DRY PERIOD DISTRIBUTIONS (EXPRESSED IN MINUTES)

	WP	DP
Mean	45.82	444.94
Deviation	68.99	609.18
Skewness	4.69	1.44
Kurtosis	32.03	4.12

TABLE III
RMSE VALUES BETWEEN THE BINS OF THE MEASURED AND SYNTHESIZED DRY AND WET PERIOD DISTRIBUTIONS

	Gamma	Pareto	Exponential	Weibull
WP	$107 \cdot 10^{-4}$	$28 \cdot 10^{-4}$	$108 \cdot 10^{-4}$	$108 \cdot 10^{-4}$
DP	$215 \cdot 10^{-4}$	$55 \cdot 10^{-4}$	$205 \cdot 10^{-4}$	$205 \cdot 10^{-4}$

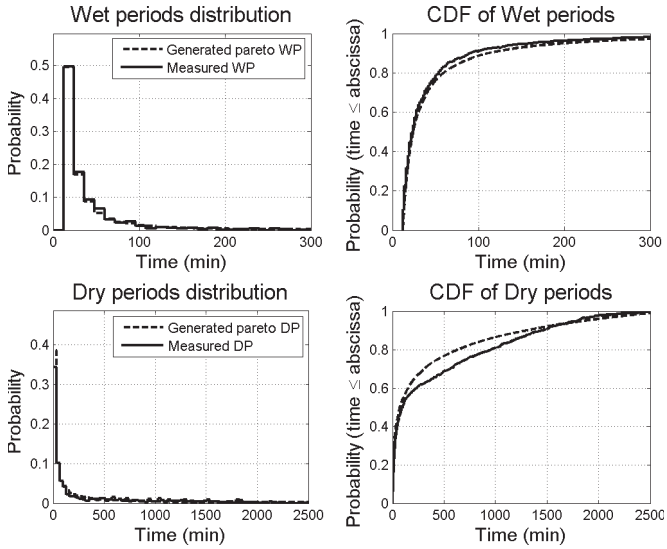

Fig. 8. (Left panels) PDF and (right panels) CDF of (solid line) measured and (dotted line) synthesized duration for (bottom panels) dry and (upper panels) wet periods.

approximation (e.g., the measured mean is not equal to the standard deviation).

The optimal pdf describing the measured WP and DP duration distributions is assessed first by using an exponential best fitting. Other probability distributions such as the Weibull, the Gamma, and the Pareto pdfs have been also tested. The latter pdf form for a generic duration T is

$$f_P(T) = a(b^a) \cdot T^{-(a+1)}, \qquad \text{for } T \geq b \tag{16}$$

where f_P is the Pareto pdf. The rmse's between the measured and conjectured distributions for the WP and DP duration are shown in Table III. The results in this table indicate that a Pareto distribution for the WP duration T_w and the DP duration T_d provides a good agreement with respect to the measured pdfs.

This consideration, based on Table III, is confirmed by Fig. 8, which compares the pdfs due to the generated and measured WP and DP duration. The same figure shows the corresponding cumulative distribution functions (cdfs) for a better comparison. For WP and DP, the Pareto distribution is, respectively, in an excellent and reasonable agreement with the measured WP and DP distributions. The optimal values for (16) are given in Table IV. Accordingly, with the definition of the Pareto distribution in (16) and the definition of WP and DP given in Section II, the parameter b represents the minimum duration of WP and DP (i.e., 12 and 2 min, respectively). The transition matrix $\mathbf{T}$ of the semi-Markov chain between DP and WP states is, by construction, a banded unit matrix [10], [32], where $p_{dd} = 0$ and $p_{ww} = 0$ are the probabilities to remain into DP and WP states, respectively, whereas $p_{dw} = 1$ and $p_{wd} = 1$ are the probabilities of transition for a DP to a WP state and vice versa, respectively (see Fig. 7).

TABLE IV
PARAMETERS OF THE PARETO DISTRIBUTION OF DRY AND WET PERIOD DURATION, RESPECTIVELY

	a	b
WP	0.96	12
DP	$3.9 \cdot 10^{-4}$	2

The parameter a is adimensional, b is expressed in (min).

B. Raindrop Distribution Autoregressive Generation

When the semi-Markov discrete system is in the *wet* period state, the three RSD parameters are generated through an appropriate VAR model. On the contrary, when the discrete system is in the *dry* period state, we can consider that it generates null rain rate samples for each of the RSD parameters (see Fig. 7). In this section, we will describe the methodology developed to synthesize (or generate, henceforth, used as synonymous) the correlated time series of the three RSD parameters $\mathbf{p}_X(t)$ of the Gamma distribution $N(D_i, \mathbf{p}_X(t))$, estimated by means of the X method on the whole of the 15 710 values of $\mathbf{p}_X$. To this aim, we assume, as in [9], that the stochastic process governing the RSD variability can be modeled as a discrete stationary VAR process. It is worth observing that, in our case, unlike [9], the order L of the autoregressive process is set to 7 rather than 1, and the analysis is carried out for all the three RSD parameters rather than only on N_w and D_m. The choice of the optimum order $L_{opt} = 7$ to generate realistic RSD parameters will be motivated in Section IV-C.

Some details about the general case of VAR processes of order L, named VAR$_3(L)$ in our case, are summarized in Appendix B. The model of VAR$_3(L_{opt})$ is given by

$$\mathbf{z}(t) = \sum_{i=1}^{L_{opt}} \mathbf{D}(i) \cdot \mathbf{z}(t-i) + \varepsilon(t) \tag{17}$$

where $\mathbf{z}(t-i)$ is the synthesized time sequence vector $\mathbf{z}$ at the ith time lag before, and the vector $\mathbf{z}(t)$ is related to the mean-centered value of the synthesized RSD parameters $\mathbf{p}(t)$ through a logarithmic relation, i.e.,

$$\mathbf{z}(t) = \ln[\mathbf{p}(t)] - \langle \ln[\mathbf{p}_X(t)] \rangle \tag{18}$$

where $\langle \ln[\mathbf{p}(t)] \rangle$ is the temporal ensemble mean of $\ln[\mathbf{p}(t)]$. The mean subtraction in (18) holds under the assumption, valid in this paper, of considering the rainy events and then

the RSD parameters $\mathbf{p}$ as a quasi-stationarity process. The compact expression $\ln[\mathbf{p}(t)]$ is equivalent to a column vector $[\ln(N_w(t))\ \ln(D_m(t)\ \ln(\mu(t))]^T$. It should be noted in (18) that μ may assume negative values (see the lower panel of Fig. 4) so that the log transformation may fail. To avoid this limitation, we have added to μ the absolute value of its minimum so that the translated μ is positive defined (when inverting (18), this translation must be corrected, of course).

In (17), the matrices $\mathbf{D}(i)$ are 3×3 autoregressive coefficient matrices, whereas $\boldsymbol{\varepsilon}(t)$ is a (3×1) zero-mean Gaussian white noise vector. The autoregressive coefficient matrices $\mathbf{D}(i)$ are arranged in the coefficient matrix $\mathbf{D}$, which has the following expression:

$$\mathbf{D} = \widetilde{\mathbf{S}}_{\mathbf{z}X}^{-1}(L_{opt}-1) \cdot \widetilde{\mathbf{S}}_{\mathbf{z}X}(L_{opt}) \tag{19}$$

where $\widetilde{\mathbf{S}}_{\mathbf{z}X}(L_{opt})$ and $\widetilde{\mathbf{S}}_{\mathbf{z}X}(L_{opt}-1)$ are matrices formed when the autocovariance matrix $\mathbf{S}_{\mathbf{z}X}(L)$ (3×3) is arranged up to the time lag L_{opt} and $L_{opt}-1$ respectively, as detailed in (B7). The autocovariance matrix $\mathbf{S}_{\mathbf{z}X}(L)$ has been computed as

$$\mathbf{S}_{\mathbf{z}X}(L) = \left\langle \mathbf{z}_X(t) \cdot \mathbf{z}_X^T(t-L) \right\rangle \tag{20}$$

where $\mathbf{z}_X = \ln[\mathbf{p}_X(t)] - \langle \ln[\mathbf{p}_X(t)] \rangle$ is specified by the chosen estimation method X. Once the coefficient matrices $\mathbf{D}(i)$ are estimated through (19), the synthesized RSD parameter vector $\mathbf{p}(t)$ can be recovered from the time series of $\mathbf{z}(t)$ by simply inverting (18), i.e.,

$$\mathbf{p}(t) = \exp\left[\mathbf{z}(t) + \langle \ln\left[\mathbf{p}_X(t)\right] \rangle\right]. \tag{21}$$

The choice of developing the VAR algorithm in (17) by using the logarithm of $\mathbf{p}(t)$, instead of directly $\mathbf{p}(t)$, has been prompted by the consideration that: 1) under the hypothesis that the joint pdf of the estimated RSD parameters $\mathbf{p}$ follows a lognormal distribution, the estimated $\mathbf{z}_X(t)$ time sequence follows a Gaussian distribution, and this is a characteristic that is required to apply an autoregressive model under the Gaussian hypothesis on $\boldsymbol{\varepsilon}(t)$; and 2) as shown in [30], the rainfall process in logarithmic coordinates can be regarded as a stationary Gaussian process. Indeed, the Gaussian assumption on the RSD parameters in the logarithmic plane seems to be confirmed by a visual inspection of the lower panel of Fig. 9, where the marginal pdfs of the RSD parameters are shown.

To generate a time series of correlated RSD parameters through the VARMS scheme of Fig. 7, (17) has to be iterated for a given number of time steps. The number n_t of steps is provided, every time the RSD generator is in a wet state, by a value extracted from the Pareto duration distribution, as discussed in Section IV-A.

C. Results

In this section, the overall performances of the proposed model VARMS are shown and discussed.

To generate a set of synthesized RSD samples by means of (17), the coefficient matrices $\mathbf{D}(i)$ and the autocovariance matrix $\mathbf{S}_{\varepsilon}$ of the white noise vector in (17) have been evaluated using the Chilbolton JWD data set. The choice to drive the

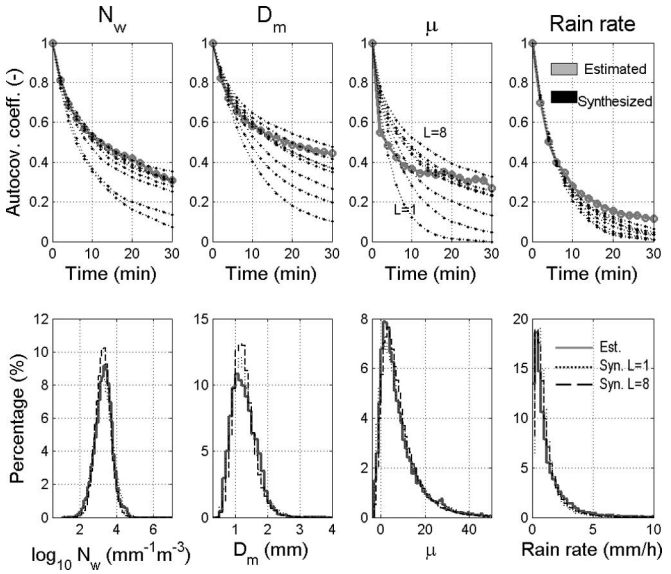

Fig. 9. (Upper panels) Autocovariance functions of RSD parameters estimated by the ML1 method and (gray line) rain rate superimposed to (black lines) the synthesized ones obtained by VARMS for different orders L of $\mathrm{VAR}_3(L)$ ranging from 1 to 8. (Lower panels) Distributions of synthesized and estimated RSD parameters and rain rate.

$\mathrm{VAR}_3(L)$ process using the whole WP data set is motivated by three main considerations.

1) About 98% of U.K. RSD parameter samples are classified as stratiform (i.e., they belong to the same type of rainfall process) when the classification procedure detailed in [22] is applied.
2) We tested the $\mathrm{VAR}_3(L)$ process by using only a single WP sequence of RSD parameters; however, the obtained results were worse than those presented later in this section.
3) To account the different types of rainfall processes, maintaining the same approach here adopted, the complexity of the semi-Markov chain shown in Fig. 7 should be increased.

For example, we could split up the WP state in three or more states (e.g., the stratiform, the convective, and the hybrid rain state [33]), calculate the probabilities of transition from one state to another one, and, for each state, drive the $\mathrm{VAR}_3(L)$ process with homogeneous preclassified RSD sequences. However, this possible improvement is beyond the scope of this paper.

The ML1 estimation method has been used to drive the VARMS tool and then generate 720 000 samples to synthesize WP and DP of the rainfall process. In addition, for each wet state of the VARMS, the synthesized wet period of $\mathbf{p}(t)$ has been computed for different orders L, ranging from 1 to 8, of the $\mathrm{VAR}_3(L)$ process. In this numerical experiment, which lasted a few minutes with a commercial computer, 90 236 WP samples of $\mathbf{p}(t)$ for each of the considered values of L, and of $\mathbf{z}(t)$ through (18), have been extracted from the whole sequence of 720 000 generated samples. The percentage of wet periods (about 12%) is fairly well reproduced with respect to the estimated ones (about 10%: see Section II).

TABLE V
RMSE VALUES BETWEEN ESTIMATED AND SYNTHESIZED AUTOCOVARIANCE FUNCTIONS OF RSD PARAMETERS AND RAIN RATE

L	N_w	D_m	μ	Rain Rate
1	0.19	0.26	0.23	0.12
2	0.15	0.19	0.20	0.13
3	0.05	0.15	0.14	0.10
4	0.03	0.12	0.09	0.10
5	0.02	0.08	0.09	0.08
6	0.02	0.08	0.08	0.07
7	0.02	0.06	0.08	0.07
8	0.08	0.10	0.09	0.08

To choose the optimal order L_{opt} of the VAR$_3(L)$, the autocovariance functions of the estimated RSD parameters $\mathbf{p}_{\mathrm{ML1}}$ and the synthesized ones $\mathbf{p}(t)$, for different orders L, have been compared, as shown in the upper panels of Fig. 9. From these plots, it seems that, as the order L increases, the autocovariance functions of the synthesized RSD parameters and the rain rate tend to approach the measured ones, particularly for their RSD tails. This behavior is partially confirmed by looking at the rmse's, given in Table V, between the estimated and synthesized curves, depicted in the upper panels of Fig. 9. The minimum discrepancy, in terms of rmse, between the measured and synthesized autocovariance functions of all the RSD parameters and the rain rate has been found for the order of the VAR process $L = 7$. The error analysis performed on the corresponding rain rate has confirmed the same trend. As a result of this comparison, L_{opt} has been set to 7. This behavior confirms the qualitative good agreement between the measured and synthesized autocovariance functions, plotted in the upper panels of Fig. 9. The autocovariance function of μ sharply decreases between lag 0 and lag 1 that corresponds to a time lag of 2 min. This behavior indicates that, for the U.K. disdrometer data, there is a significant variability of μ within a 2-min lag with respect to the other RSD parameters. If the RSD data are processed so that the time sampling of the RSD spectra is degraded to 1 min, the high variability of μ is confirmed as well. The lower panels of the same figure show the comparison between the estimated distributions of the RSD parameters and the synthesized ones for the orders of the VAR process $L = 1$ and $L = 8$. This figure suggests an acceptable agreement between the estimated and synthesized distributions of all quantities independently from the values taken from L. On the whole, Fig. 9 seems to indicate that orders L of the VAR process higher than 1 can only improve the correlation characteristics of the synthesized time series with respect to the estimated ones, but with a scarce influence on their probability distributions.

As a further comparison between estimated and synthesized characteristics, Fig. 10 shows the scatter plots between the 90 236 generated samples (associated to $L = 7$) of the components of $\mathbf{p}(t)$ superimposed to the $\mathbf{p}_{\mathrm{ML1}}(t)$ ones for the whole of the 15 710 selected instants (see Section II). From this figure, an overall good agreement between the estimated and synthesized D_m and $\log_{10}(N_w)$ is noted. The other correlation diagrams, representing D_m against μ and $\log_{10}(N_w)$ against μ, seem to confirm the capability of RSD generation to reproduce the measurement domain.

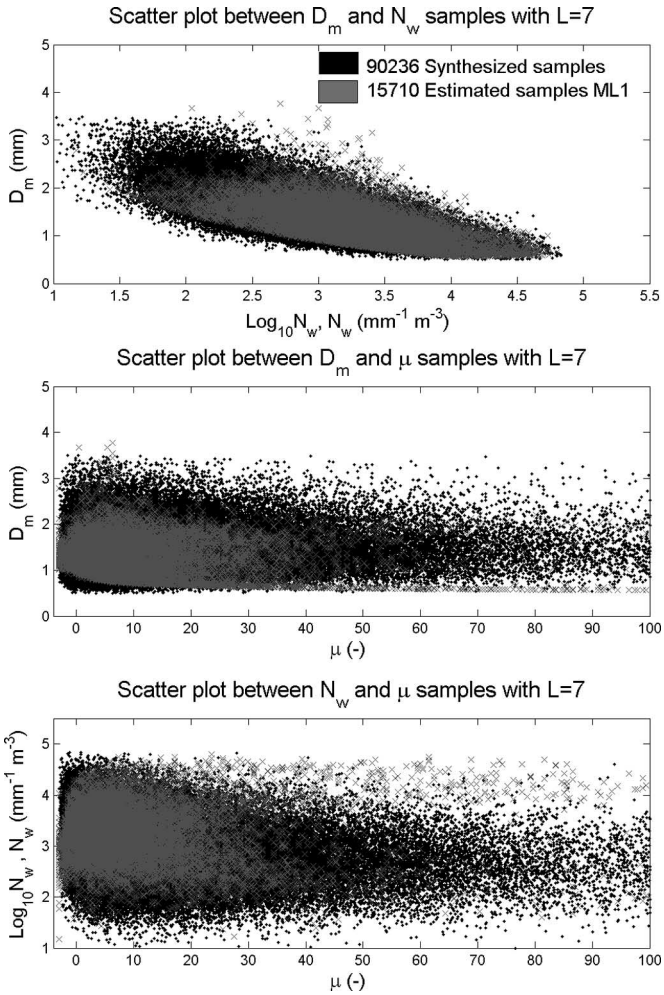

Fig. 10. Scatter plots between (top panel) mean diameter D_m and log concentration $\log_{10}(N_w)$, (middle panel) mean diameter D_m and shape parameter, and (lower panel) log concentration $\log_{10}(N_w)$ and shape parameter μ. Gray crosses are measured samples, whereas black dots indicate synthesized samples.

Summarizing the scheme in Fig. 7, the VARMS approach of the Gamma RSD parameter time series consists of the following steps.

1) Define a total duration T_s of the RSD parameter synthetic generation.
2) Start with a dry period DP, extracting a realization of T_d from (16) and setting $\mathbf{p}(t)$ equal to void values at each instant within this dry period.
3) After the period T_d, extract a realization of T_w from (16) with its proper WP coefficient values to determine the wet period WP duration.
4) Within WP, use (20) to set up all unknown matrices to apply (17) and determine $\mathbf{z}(t)$ by means of the VAR$_3(7)$ algorithm.
5) Convert $\mathbf{z}(t)$ into $\mathbf{p}(t)$ through (21) for each instant belonging to WP.
6) At the end of WP, go to step 2 until the sum of previous T_d and T_w is larger than T_s.

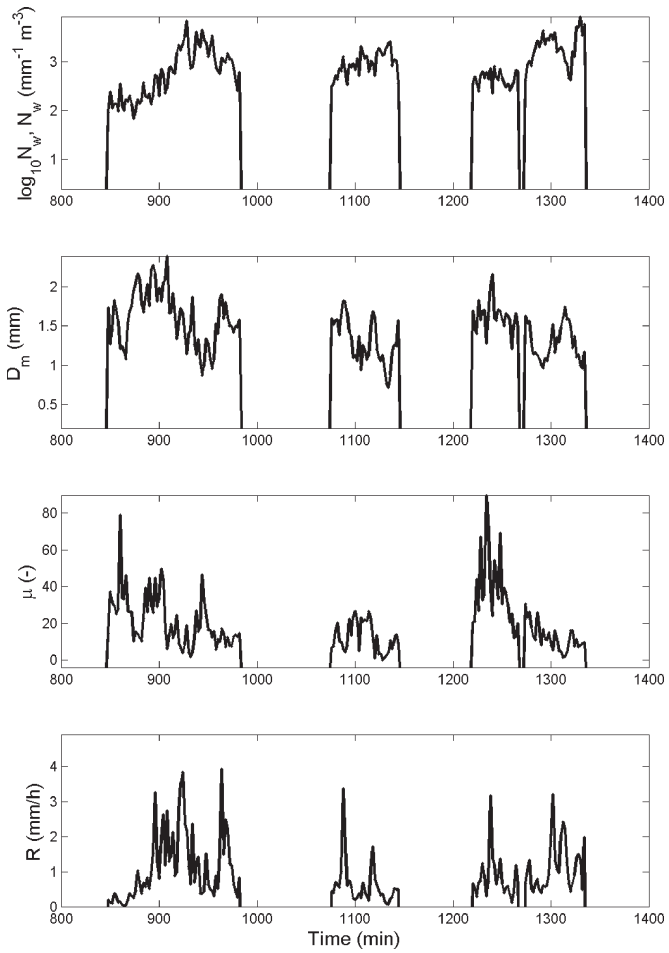

Fig. 11. Example of synthesized RSD parameters and rain rate by using VARMS and the ML1 approach for RSD parameter retrieval. From the top to the bottom panels, a realization for each of the synthesized RSD parameters $\log_{10} N_w$, D_m, μ, and the synthesized rain rate is shown.

An example of a synthesized time series of the Gamma RSD parameters $N_w(t)$, $D_m(t)$, and μ is shown in Fig. 11 together with the derived rain rate time series. The synthesized rain rate has been estimated by using (3) and by considering, in (4), $N(D, \mathbf{p}_{\mathrm{ML1}})$ in place of N_m. Fig. 11 is calibrated on available Chilbolton time-series JWD data and is derived from the ML1 approach for RSD parameter retrieval. Of course, the same VARMS methodology can be applied to any available disdrometer data set and with any method to estimate RSD parameters from disdrometer data.

V. Conclusion

A large set of Joss–Waldvogel impact disdrometer measurements, collected from 2003 to 2005 at Chilbolton, U.K., have been analyzed in this paper. The RSD derived from previous measurements has been approximated by means of the well-known normalized Gamma distribution. The driving parameters of the latter have been estimated using three methods. The first one is based on the consolidated GM approach, whereas the second and third ones are founded on the ML procedure in 1-D or 3-D configuration, respectively. The results of the comparison between these three RSD estimation methods have shown that they may lead to quite different results, particularly when one is interested in evaluating RSD higher order moments such as the rain rate. Based on this RSD estimation, a stochastic VAR model has been developed to randomly synthesize (generate) temporal series of normalized Gamma RSD parameters. To account for rainfall intermittence, a two-state semi-Markov chain has been modeled from measurements using two-state analytical statistics of rain- and dry-period duration. The accuracy of the overall methodology, named VARMS, has been evaluated on Chilbolton data, providing excellent statistical scores.

This stochastic RSD generation tool may find useful applications in hydrometeorology and radio propagation (e.g., [6] and [9]). The ambition of VARMS is to provide an easy tool, fully documented, to generate the time series of realistic RSDs to, for example, simulate polarimetric radar observables (e.g., reflectivity, differential reflectivity, linear depolarization ratio, specific differential phase) and/or radio propagation parameters (e.g., specific attenuation and differential attenuation, cross-polarization phase shift) by using a rain scattering and absorption model. In addition, the use of VARMS together with further development, based on spectral approaches, will make it possible to extend the RSD generation to a spatial case or a 2-D case.

Appendix A
Disdrometer Measurements

A brief description of the JWD instrumentation, the measurement process, and the performed data processing is here provided.

A disdrometer is an instrument designed to measure RSDs. There are various classes of disdrometers, including optical beam devices [11], video devices [12], and impact disdrometers [13], [14]. The latter is known as JWD too. The JWD RD-69 type [15], which is exploited in this paper, consists of two units the sensor, which is exposed to the rain, and the processor for analog processing and digitizing of the sensor-measured information.

The sensor transforms the mechanical momentum of an impacting drop into an electric pulse, whose amplitude is roughly proportional to the mechanical momentum. The sensor consists of an electromechanical unit and an amplifier housed in a common case. A conical Styrofoam body is used to transmit the mechanical impulse of an impacting drop to a set of two moving coil systems in magnetic fields. At the impact of a drop, the Styrofoam body together with the two coils moves downward, and a voltage is induced in the first sensing coil. This voltage, which is also the output information, is amplified and applied to the second driving coil, producing a force that counteracts the movement. As a consequence, the excursion is very small, and it takes very little time for the system to return to its original resting position and, therefore, to get ready for the next impact of a drop. The voltage amplitude of the pulse at the amplifier output is a measure for the size of the drop that caused it, and it is related to the drop diameter with the following relation [16], [17]:

$$V_L = k \cdot D^n \tag{A1}$$

where D is the drop diameter in the range 0.5–5 mm, V_L is the output voltage, and k and n are the calibration constants. A usual value for k is 0.02586, and a usual value for n ranges between 3.1 and 4.3 [17]. Afterward, a pulse height analyzer, embedded in the processor unit, classifies the peaks amplitude V_L and, accordingly, the impacting drop diameters into n_c classes (sometimes named channels or bins or size category) whose boundaries are specified by the following law:

$$D_{ci} = \left[\frac{10^{\{1-\alpha\cdot(n_c-i)\}}}{\gamma}\right]^{\beta} \tag{A2}$$

where D_{ci} is the lower boundary of the channel c_i, the index i assumes values from 1 to n_c, $\alpha = 0.014253$, $\beta = 0.6803$, and $\gamma = 0.94$. In this way, small drop diameters are classified by means of a fine channel with respect to the large ones. When raindrops fall on the sensor surface, the JWD records the number of counts in each of the n_c channels for a given averaging period.

APPENDIX B
RSD AUTOREGRESSIVE PROCESS

In this appendix, we provide a brief overview of the autoregressive process theory to explain the expressions of the coefficient matrix (19) and error covariance matrices.

The general expression of an autoregressive model of order L allows predicting the actual observation as a linear combination of the L past observations plus a random error as follows:

$$\mathbf{z}(t) = \sum_{i=1}^{L} \mathbf{D}(i)\cdot\mathbf{z}(t-i) + \varepsilon(t) \tag{B1}$$

where the vector $\mathbf{z}(t)$ is the $(k\times 1)$ predicted zero-mean vector at the instant t, L is the order of the autoregressive process, $\mathbf{D}(i)$ are $(k\times k)$ autoregressive coefficient matrices, $\boldsymbol{\varepsilon}(t)$ is a $(k\times 1)$ zero-mean Gaussian white noise vector, which represents the part of $\mathbf{z}(t)$ that is not linearly dependent on L past observations $\mathbf{z}(t-i)$, and k is the number of variables that have to be generated. Equation (B1) expresses the general form of the k VAR process of order L [VAR$_k(L)$]. Knowing the matrices $\mathbf{D}(i)$ and the Gaussian white noise $\boldsymbol{\varepsilon}(t)$, we can generate the correlated sequence $\mathbf{z}(t)$ at each desired instant. To calculate $\mathbf{D}(i)$ and $\boldsymbol{\varepsilon}(t)$, the estimated sequence $\mathbf{z}_X$, from which we want to extract the correlation structure and then generate correlated RSD parameter realizations, has to be exploited. As specified in Section IV-B, $\mathbf{z}_X(t)$ is the logarithm of the instantaneous vector of the RSD parameters, estimated by the X method, where $X =$ ML1, ML3, or GM. Therefore, the calculus of $\mathbf{D}(i)$ and $\boldsymbol{\varepsilon}(t)$ can be done replacing in (B1) $\mathbf{z}(t)$ with $\mathbf{z}_X(t)$ and rewriting (B1) as follows:

$$\boldsymbol{\varepsilon}(t) = \sum_{i=0}^{L} (-\mathbf{D}(i))\cdot\mathbf{z}_X(t-i) \tag{B2}$$

where $\mathbf{D}(i)$ is equal to the integer -1 when $i = 0$. Multiplying both terms of (B2) by $\mathbf{z}_X(t-h)$ (with the shift $h > 0$) and applying on them the temporal average operator $\langle\cdot\rangle$, we obtain the following equation:

$$\left\langle \boldsymbol{\varepsilon}(t)\cdot\mathbf{z}_X^T(t-h)\right\rangle = \sum_{i=0}^{L} (-\mathbf{D}(i))\cdot\left\langle \mathbf{z}_X(t-i)\cdot\mathbf{z}_X^T(t-h)\right\rangle \tag{B3}$$

where the symbol T indicates matrix transposition. Assuming that the zero-mean Gaussian white noise ε is uncorrelated at different time lags, or, equivalently, assuming that its autocovariance matrix at the generic lag h, i.e., $\mathbf{S}_{\varepsilon}(h)$, can be expressed by

$$\mathbf{S}_{\boldsymbol{\varepsilon}}(h) = \left\langle \boldsymbol{\varepsilon}(t)\cdot\boldsymbol{\varepsilon}^T(t-h)\right\rangle = \begin{cases} \mathbf{S}_{\boldsymbol{\varepsilon}}(0) & \text{for } h = 0 \\ 0 & \text{for } h \neq 0 \end{cases} \tag{B4}$$

then the first term in (B3) can be set to zero. Therefore, (B3) becomes

$$\sum_{i=0}^{L} (-\mathbf{D}(i))\cdot\mathbf{S}_{\mathbf{z}X}^T(h-i) = 0 \tag{B5}$$

where $\mathbf{S}_{\mathbf{z}X}(h-i)$ is the autocovariance matrix of $\mathbf{z}_X(t)$ at lag $(h-i)$ computed as in (B4), but replacing $\boldsymbol{\varepsilon}(t)$ with $\mathbf{z}_X(t)$ and where the property $\mathbf{S}_{\mathbf{z}X}(i) = \mathbf{S}_{\mathbf{z}X}^T(-i)$ has been exploited. Varying the indexes i and h from 0 to L in (B5), we obtain an equation system that can be written in a matrix form as follows:

$$\widetilde{\mathbf{S}}_{\mathbf{z}X}(L) = \widetilde{\mathbf{S}}_{\mathbf{z}X}(L-1)\cdot\mathbf{D} \tag{B6}$$

where $\mathbf{D}$, $\widetilde{\mathbf{S}}_{\mathbf{z}X}(L)$, and $\widetilde{\mathbf{S}}_{\mathbf{z}X}(L-1)$ are given in (B7), shown at the bottom of the page.

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}(1) \\ \mathbf{D}(2) \\ \vdots \\ \mathbf{D}(L) \end{bmatrix}_{kL\times k} \qquad \widetilde{\mathbf{S}}_{\mathbf{z}X}(L) = \begin{bmatrix} \mathbf{S}_{\mathbf{z}X}^T(1) \\ \mathbf{S}_{\mathbf{z}X}^T(2) \\ \vdots \\ \mathbf{S}_{\mathbf{z}X}^T(L) \end{bmatrix}_{kL\times k}$$

$$\widetilde{\mathbf{S}}_{\mathbf{z}X}(L-1) = \begin{bmatrix} \mathbf{S}_{\mathbf{z}X}(0) & \mathbf{S}_{\mathbf{z}X}(1) & \cdots & \mathbf{S}_{\mathbf{z}X}(L-1) \\ \mathbf{S}_{\mathbf{z}X}^T(1) & \mathbf{S}_{\mathbf{z}X}(0) & \cdots & \mathbf{S}_{\mathbf{z}X}(L-2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{S}_{\mathbf{z}X}^T(L-1) & \mathbf{S}_{\mathbf{z}X}^T(L-2) & \cdots & \mathbf{S}_{\mathbf{z}X}(0) \end{bmatrix}_{kL\times kL} \tag{B7}$$

At this point, from (B6), taking the inverse matrix of $\widetilde{\mathbf{S}}_{\mathbf{z}X}(L-1)$, the autoregressive coefficient matrix is derived from

$$\mathbf{D} = \widetilde{\mathbf{S}}_{\mathbf{z}X}^{-1}(L-1) \cdot \widetilde{\mathbf{S}}_{\mathbf{z}X}(L). \tag{B8}$$

We observe that when $L = 1$, $\mathbf{D}$ reduces to $\mathbf{D}(1)$, $\widetilde{\mathbf{S}}_{\mathbf{z}X}(L-1)$ reduces to $\mathbf{S}_{\mathbf{z}X}(0)$, and $\widetilde{\mathbf{S}}_{\mathbf{z}X}(L)$ reduces to $\mathbf{S}_{\mathbf{z}X}^{T}(1)$, and only the autocovariance matrix of $\mathbf{z}_X(t)$ at lags 1 and 0 has to be computed as

$$\mathbf{D}(1) = \mathbf{S}_{\mathbf{z}X}^{-1}(0) \cdot \mathbf{S}_{\mathbf{z}X}(1). \tag{B9}$$

The final step to describe the autoregressive generation process concerns the modeling of the white noise $\boldsymbol{\varepsilon}(t)$ that must have a zero mean to generate a vector time sequence with an autocorrelation structure similar to the measured one. If we consider the case $h = 0$ in (B3), we can rewrite it as follows:

$$\left\langle \boldsymbol{\varepsilon}(t) \cdot \mathbf{z}_X^T(t) \right\rangle = \sum_{i=0}^{L} \left(-\mathbf{D}(i)\right) \cdot \left\langle \mathbf{z}_X(t-i) \cdot \mathbf{z}_X^T(t) \right\rangle. \tag{B10}$$

Substituting (B1) with $\mathbf{z}_X(t)$, exploiting the properties of $\boldsymbol{\varepsilon}(t)$ and the definition of autocovariance matrix as in (B4) for the vector $\mathbf{z}_X(t)$, (B10) transforms as follows:

$$\mathbf{S}_{\boldsymbol{\varepsilon}}(0) = \sum_{i=0}^{L} \left(-\mathbf{D}(i)\right) \cdot \mathbf{S}_{\mathbf{z}X}(i). \tag{B11}$$

In case of VAR$_k(1)$, $\mathbf{S}_{\boldsymbol{\varepsilon}}(0)$ takes the following simplified form:

$$\mathbf{S}_{\boldsymbol{\varepsilon}}(0) = \mathbf{S}_{\mathbf{z}}(0) - \mathbf{D}(1) \cdot \mathbf{S}_{\mathbf{z}X}(1). \tag{B12}$$

In summary, to synthesize a vector sequence $\mathbf{z}(t)$, which shows the same time correlation properties as a given set of estimated data $\mathbf{z}_X(t)$, we have to compute the autocovariance matrix of $\mathbf{z}_X(t)$ at different lags L, i.e., $\mathbf{S}_{\mathbf{z}X}(L)$. Then, we have to apply (B8) and (B11), and, finally, compute (B1) for a desired number of iterations. In this paper, all 15 710 estimated RSD samples of $\mathbf{z}_X(t)$, described in Section III, have been exploited as measured data to compute the autocovariance matrix at lags from 0 to $L = 8$. Eventually, we can provide VAR matrix details for the special case of $L = 1$ order, assumed as a reference framework in [9]. The autoregressive coefficient matrix $\mathbf{D}$ and the error covariance matrix $\mathbf{S}_{\varepsilon}(0)$ for $L = 1$ are

$$\mathbf{D}(1) = \begin{bmatrix} 0.7794 & -0.4069 & -0.1410 \\ 0.0129 & 0.9093 & 0.0345 \\ -0.0674 & -0.0637 & 0.7335 \end{bmatrix} \tag{B13a}$$

$$\mathbf{S}_{\boldsymbol{\varepsilon}}(0) = \begin{bmatrix} 0.3461 & -0.0510 & 0.0972 \\ -0.0510 & 0.0229 & -0.0326 \\ 0.0972 & -0.0326 & 0.2460 \end{bmatrix} \tag{B13b}$$

$$\langle \ln[\mathbf{p}(t)] \rangle = \begin{bmatrix} 7.8686 \\ 0.1063 \\ 2.2720 \end{bmatrix} \tag{B13c}$$

where the mean vector is needed to reconstruct (17).

ACKNOWLEDGMENT

The authors would like to thank the BADC for providing the disdrometer data used in this paper. They would also like to thank Prof. J. Goddard (Rutherford Appleton Laboratory, U.K.) for his helpful suggestions to properly process the JWD data and Prof. F. Santucci for his useful suggestions.

REFERENCES

[1] A. Tokay, A. Kruger, W. F. Krajewski, P. A. Kucera, A. Jose, and P. Filho, "Measurements of drop size distribution in the southwestern Amazon basin," *J. Geophys. Res.*, vol. 107, no. D20, 8052, 2002. DOI:10.1029/2001JD000355.

[2] D. Atlas, C. W. Ulbrich, F. D. Marks, Jr., E. Amitai, and C. R. Williams, "Systematic variation of drop size distribution and radar–rainfall relation," *J. Geophys. Res.*, vol. 104, pp. 6155–6169, 1999.

[3] L. D. Carey, S. A. Rutledge, D. A. Ahijevych, and T. D. Keenan, "Correcting propagation effects in C-band polarimetric radar observations of tropical convection using differential propagation phase," *J. Appl. Meteorol.*, vol. 39, no. 9, pp. 1405–1433, Sep. 2000.

[4] E. A. Brandes, G. Zhang, and J. Vivekanandan, "Comparison of polarimetric radar drop size distribution retrieval algorithms," *J. Atmos. Ocean. Technol.*, vol. 21, no. 4, pp. 584–598, Apr. 2003.

[5] L. D. Carey and S. A. Rutledge, "The relationship between precipitation and lightning in tropical island convection: A C-band polarimetric radar study," *Mon. Weather Rev.*, vol. 128, no. 8, pp. 2687–2710, Aug. 2000.

[6] G. Vulpiani, F. S. Marzano, V. Chandrasekar, A. Berne, and R. Uijlenhoet, "Polarimetric weather radar retrieval of raindrop size distribution by means of a regularized artificial neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3262–3275, Nov. 2006.

[7] T. Maseng and P. M. Bakken, "A stochastic-dynamic model of rain attenuation," *Int. J. Satell. Commun.*, vol. 29, no. 5, pp. 660–669, May 1981.

[8] C. Kummerow, W. Barnes, T. Kozu, J. Shiue, and J. Simpson, "The Tropical Rainfall Measuring Mission (TRMM) sensor package," *J. Atmos. Ocean. Technol.*, vol. 15, no. 3, pp. 809–817, Jun. 1998.

[9] A. Berne and R. Uijlenhoet, "A stochastic model of range profiles of raindrop size distributions: Application to radar attenuation correction," *Geophys. Res. Lett.*, vol. 32, no. 10, L10 803, May 2005. DOI:10.1029/2004GL021899.

[10] C. Allasseur, L. Husson, and F. Pèrez-Fontàn, "Simulation of rain events time series with Markov model," in *Proc. 15th IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.*, Sep. 2004, vol. 4, pp. 2801–2805.

[11] M. Grossklaus, K. Uhlig, and L. Hasse, "An optical disdrometer for use in high wind speeds," *J. Atmos. Ocean. Technol.*, vol. 15, no. 4, pp. 1051–1059, Aug. 1998.

[12] M. Schönhuber, H. Urban, J. P. V. Poiares Baptista, W. L. Randeu, and W. Riedler, "Measurements of precipitation characteristics by a new disdrometer," in *Proc. Atmos. Phys. Dyn. Anal. Prognosis Precipitation Fields*, Nov. 1994, pp. 15–18.

[13] J. Joss and A. Waldvogel, "A raindrop spectrograph with automatic analysis," *Pure Appl. Geophys.*, vol. 68, pp. 240–246, 1967.

[14] B. E. Sheppard and P. I. Joe, "Comparison of raindrop size distribution measurements by a Joss–Waldvogel disdrometer, a PMS 2DG spectrometer, and POSS Doppler radar," *J. Atmos. Oceanic.*, vol. 11, no. 4, pp. 874–887, Aug. 1994.

[15] Distromet LTD., "Disdrometer RD-69," *Instruction Manual*, Basel, Switzerland, 1997.

[16] J. Joss and A. Waldvogel, "Comments on some observations on the Joss–Waldvogel rainfall disdrometer," *J. Appl. Meteorol.*, vol. 16, no. 1, pp. 112–113, Jan. 1977.

[17] B. E. Sheppard, "Effect of irregularities in the diameter classification of raindrops by the Joss–Waldvogel disdrometer," *J. Atmos. Ocean. Technol.*, vol. 7, no. 1, pp. 180–183, Feb. 1990.

[18] D. Atlas and C. W. Ulbrich, "Path- and area-integrated rainfall measurement by microwave attenuation in 1–3 cm band," *J. Appl. Meteorol.*, vol. 16, no. 12, pp. 1322–1331, Dec. 1977.

[19] E. Gorgucci, V. Chandrasekar, V. N. Bringi, and G. Scarchilli, "Estimation of raindrop size distribution parameters from polarimetric radar measurements," *J. Atmos. Sci.*, vol. 59, no. 15, pp. 2373–2384, Aug. 2002.

[20] J. Testud, S. Oury, R. A. Black, P. Amayenc, and X. Dou, "The concept of 'normalized' distributions to describe raindrop spectra: A tool for cloud physics and cloud remote sensing," *J. Appl. Meteorol.*, vol. 40, no. 6, pp. 1118–1140, Jun. 2001.

[21] A. J. Illingworth and T. M. Blackman, "The need to represent raindrop size spectra as normalized gamma distributions for the interpretation of polarization radar observations," *J. Appl. Meteorol.*, vol. 41, no. 3, pp. 286–297, Mar. 2002.
[22] V. N. Bringi, V. Chandrasekar, J. Hubbert, E. Gorgucci, W. L. Randeu, and M. Schoenhuber, "Raindrop size distribution in different climatic regimes from disdrometer and dual-polarized radar analysis," *J. Atmos. Sci.*, vol. 60, no. 2, pp. 354–365, Jan. 2003.
[23] J. M. Bernardo and A. F. M. Smith, *Bayesian Theory*. New York: Wiley.
[24] C. W. Ulbrich and D. Atlas, "Rainfall microphysics and radar properties: Analysis methods for drop size spectra," *J. Appl. Meteorol.*, vol. 37, no. 9, pp. 912–923, Sep. 1998.
[25] V. N. Bringi and V. Chandrasekar, *Polarimetric Doppler Weather Radar: Principles and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2001.
[26] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics*, 5th ed, vol. 2. London, U.K.: Griffin, 1991.
[27] P. L. Smith and D. V. Kliche, "The bias in moment estimators for parameters of drop size distribution functions: Sampling from the exponential distributions," *J. Appl. Meteorol.*, vol. 44, no. 8, pp. 1195–1205, Aug. 2005.
[28] M. Montopoli, G. Vulpiani, M. N. Anagnostou, E. N. Anagnostou, and F. S. Marzano, "Processing disdrometer raindrop spectra time series from various climatological regions using estimation and autoregressive methods," in *Proc. IGARSS*, Barcelona, Spain, Jul. 23–27, 2007.
[29] J. Vivekanandan, G. Zhang, and E. Brandes, "Polarimetric radar estimators based on a constrained gamma drop size distribution model," *J. Appl. Meteorol.*, vol. 43, no. 2, pp. 217–230, Feb. 2004.
[30] D. Veneziano, R. L. Bras, and J. D. Niemann, "Nonlinearity and self-similarity of rainfall in time and a stochastic model," *J. Geophys. Res.*, vol. 101, no. D21, pp. 26 371–26 392, 1996.
[31] L. Kleinrock, *Queueing Systems*, vol. 1. New York: Wiley, 1975.
[32] R. L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.
[33] F. Fontan, U.-C. Fiebig, and C. Enjamino, "Improving rain rate time-series generation for system simulation applications," in *Proc. Int. Workshop COST Actions 272, 280*, Noordwijk, The Netherlands, May 26–28, 2003.

Mario Montopoli received the Laurea degree in electronic engineering from the University of L'Aquila, L'Aquila, Italy, in 2004. He is currently working toward the Ph.D. degree in radar meteorology in a joint program between the University of Basilicata, Potenza, Italy, and the Sapienza University of Rome, Rome, Italy.

In 2005, he joined the Centro di Eccellenza per l'intagrazione di tecniche di Telerilevamento e Modellistica per la Previsione di eventi meteorologici Severi as a Research Scientist on ground-based radar meteorology with a special focus on C-band applications and processing techniques. Since 2006, he has been with the Department of Electrical Engineering and Information, University of L'Aquila, as a Research Assistant.

Frank Silvio Marzano (S'89–M'99–SM'03) received the Laurea degree (*cum laude*) in electrical engineering and the Ph.D. degree in applied electromagnetics from the Sapienza University of Rome, Rome, Italy, in 1988 and 1993, respectively.

In 1993, he was with the Institute of Atmospheric Physics, National Research Council, Rome. From 1994 to 1996, he was with the Italian Space Agency, Rome, as a Postdoctorate Researcher. In 1997, after being a Lecturer with the University of Perugia, Perugia, Italy, he joined the Department of Electrical Engineering and cofounded the Centro di Eccellenza per l'intagrazione di tecniche di Telerilevamento e Modellistica per la Previsione di eventi meteorologici Severi, University of L'Aquila, L'Aquila, Italy, coordinating the Satellite and Radar Remote Sensing Laboratory. Since 2005, he has been with the Department of Electronic Engineering, Sapienza University of Rome, where he currently teaches courses on antennas, propagation, and remote sensing. His current research concerns passive and active remote sensing of the atmosphere from ground-based, airborne, and spaceborne platforms, with a particular focus on precipitation using microwave and infrared data, development of inversion methods, radiative transfer modeling of scattering media, and radar meteorology issues. He is also involved in radio propagation topics in relation to incoherent wave modeling, scintillation prediction, and rain fading analysis along satellite microwave links.

Dr. Marzano is the recipient of the Young Scientist Award of the XXIV Union Radio Scientifique Internationale General Assembly in 1993 and the Alan Berman Research Publications Award from the Naval Research Laboratory, Washington, DC, in 1998. Since 2001, he has been the Italian national delegate for the European Cooperation in the Field of Scientific and Technical Research Actions 720 and 280. Since January 2004, he has been acting as an Associated Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.

Gianfranco Vulpiani (M'06) received the Laurea degree in physics and the Ph.D. degree in radar meteorology from the University of L'Aquila, L'Aquila, Italy, in 2001 and 2005, respectively.

In 2001, he joined the Department of Physics and the Centro di Eccellenza per l'intagrazione di tecniche di Telerilevamento e Modellistica per la Previsione di eventi meteorologici Severi, University of L'Aquila, as a Research Scientist on ground-based radar meteorology with a special focus on C-band and polarimetric applications. In 2004, he was a Visiting Scientist with the Colorado State University, Fort Collins. In 2006, he joined the radar research group at Météo France for work on the development of attenuation correction and hail detection algorithms within the European project FLYSAFE. Since March 2007, he has worked for the Italian Department of Civil Protection, Rome, Italy, on the definition, development, and testing of the data processing chain employed by the national radar network.