Statistical Characterization and Modeling of Raindrop Spectra Time Series for Different Climatological Regions

Mario Montopoli, Frank Silvio Marzano, *Senior Member, IEEE*, Gianfranco Vulpiani, Marios N. Anagnostou, *Member, IEEE*, and Emmanouil N. Anagnostou

***Abstract*—A large data set of raindrop size distribution (RSD) measurements collected with the Joss–Waldvogel disdrometer (JWD) and the 2-D video disdrometer (2DVD) in the U.K., Greece, Japan, and the U.S. are analyzed and modeled. This work extends a previous effort devoted to the exploitation of U.K. data and the design of a stochastic procedure to randomly generate synthetic RSD intermittent time series. This study seeks to: 1) explore the differences of RSD-derived moments for distinct hydroclimate regions, ranging from tropics to subtropics and mid and northern latitudes; 2) compare the governing parameters of the normalized gamma RSD for both stratiform and convective events and perform a sensitivity analysis by using different best fitting techniques; 3) exploit the time-correlation structure of the estimated RSD parameters as the input of a vector autoregressive stationary model used to simulate time series (or horizontal profiles) of RSDs and, consequently, its moments as the rain rate and concentration; and 4) characterize the distribution of the inter-rain duration and rain duration to design a semi-Markov chain to represent the intermittency feature of the rainfall process in a climatological framework. This climatological analysis and the related stochastic RSD generation model may find useful applications within both hydrometeorology and radio propagation.**

***Index Terms*—Autoregressive process, climatology, disdrometer data, Markov chain, modeling, radar applications, radar meteorological factors, radio propagation, rain, raindrop size distribution (RSD), signal synthesis, weather radar.**

Manuscript received October 15, 2007; revised March 1, 2008 and March 26, 2008. Current version published October 1, 2008. This work was supported in part by the Italian Ministry of University and Research, by Region Abruzzo, and by the Italian Department of Civil Protection. The work of E. N. Anagnostou and M. N. Anagnostou was supported by the European Union under (FP6) Marie Curie Excellence Grant MEXT-CT-2006-038331.

M. Montopoli is with the Department of Electrical and Information Engineering and Center of Excellence (CETEMPS), University of L'Aquila, 67100 L'Aquila, Italy (e-mail: mmontop@ing.univaq.it).

F. S. Marzano is with the Department of Electronic Engineering, University of Rome "La Sapienza," 00184 Rome, Italy, and also with the Center of Excellence (CETEMPS), University of L'Aquila, 67100 L'Aquila, Italy (e-mail: marzano@die.uniroma1.it).

G. Vulpiani was with the Department of Electrical and Information Engineering and Center of Excellence (CETEMPS), University of L'Aquila, 67100 L'Aquila, Italy. He is now with the Italian Department of Civil Protection, Rome, Italy (e-mail: gianfranco.vulpiani@gmail.com).

M. N. Anagnostou is with the Institute for Inland Waters, Hellenic Center for Marine Research, 19013 Anavissos, Greece (e-mail: managnostou@ath.hcmr.gr).

E. N. Anagnostou is with the Department of Civil and Environmental Engineering, University of Connecticut, Storrs, CT 06269 USA, and also with the Institute for Inland Waters, Hellenic Center for Marine Research, 19013 Anavissos, Greece (e-mail: manos@engr.uconn.edu; manos@ath.hcmr.gr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2008.2000652

I. Introuction

RAINDROP size distribution (RSD) plays a fundamental role in governing both the atmospheric precipitation microphysics and phenomenology and the response of microwave remote sensors and radio-propagation links [1]–[3]. In this respect, the variability of the RSD is significantly dependent on the meteorological, geographical, and climatological conditions [4]. This intrinsic variability may be even noted within each storm's space–time evolution and obviously affects the macroscopic relationship between the rain rate and microwave observables such as weather radar reflectivity and radio-link specific attenuation [5], [6].

The knowledge of RSD and its variability is also important to design model-based retrieval algorithms for remote sensing and prediction methods for radio propagation since an accurate RSD characterization, coupled with an electromagnetic scattering and absorption model, allows us to simulate polarimetric radar and radio observables [7]–[9]. To this aim, it may be also of relevant interest the capability to generate synthetic time series of RSDs, constrained by the measured statistics in a climatological context, in order to increase the available databases and test new algorithms and processing techniques [5], [6], [10].

This work starts from a previous effort devoted to the exploitation of three years of disdrometer data in the U.K. and to the design of a stochastic procedure to randomly generate synthetic RSD intermittent time series [14]. The latter methodology was named vector autoregressive Markov synthesizer (VARMS). In this further investigation, we have extended the sources of the available RSD time series, including the Joss–Waldvogel disdrometer (JWD) and the 2-D video disdrometer (2DVD) from Greece, Japan (JP), and the U.S. This choice tends to cover different geographical and climatological areas in order represent as much as possible the intrinsic RSD variability. A new technique to retrieve the parameters of the best fitting normalized gamma RSD for each data set is also tested, and the results are classified with respect to the stratiform and convective regimes. Finally, the VARMS approach is characterized for each distinct hydroclimate area in order to design a general tool for both hydrometeorology and radio-propagation purposes.

The paper is organized as follows. In Section II, we describe the disdrometer measurement characteristics and data processing procedures. In the same section, the retrieval methods for estimating the RSD parameters are introduced and applied to the available raindrop spectra to show how the moment techniques may give results substantially different from those of the maximum likelihood (ML) method. A discussion on the

TABLE I
STATISTICAL CHARACTERISTICS OF DISDROMETER-DERIVED RAIN-RATE DATA SETS

	Number samples	*Mean (mm/h)*	*Std (mm/h)*	*Max (mm/h)*	*Kurtosis*	*Skewness*
ATH	8022	3.35	7.73	188.34	129.18	9.04
UK	40065	1.29	2.49	115.51	295.15	12.40
JP_{TMH}	25390	4.73	9.06	153.4	37.58	4.90
JP_{KAS}	22705	3.53	6.86	154.8	74.07	6.59
JP_{SBS}	11876	3.42	6.54	114.00	82.05	7.20
JP_{HRJ}	4704	1.94	3.30	43.48	35.35	4.60
Iowa	1067	0.75	2.23	56.43	392.63	17.30
Virginia	6148	2.18	3.89	56.03	51.16	5.91
Florida	6148	7.33	13.96	120.58	14.14	3.08

various RSD features as a function of the available hydroclimate zones is carried out in Section III. Finally, in Section IV, the retrieved RSD parameters are exploited to specify the RSD stochastic simulator used to randomly generate time series of normalized gamma RSD parameters.

II. RSD DATA SETS

The disdrometer measurements used in this study were collected at Chilbolton (U.K.), JP, Iowa, Virginia, and Florida (U.S.) with JWD and in Athens (ATH; Greece) with 2DVD, spanning a period from 2001 to 2006, and are all available at 1-min temporal resolution. The U.K. data set, based on JWD measurements, is extensively described in [14].

A. Globally Distributed Disdrometer Measurements

The JP data set is composed of different disdrometric measurements gathered in four places, which are: 1) Tamaho (JP_{TMH}), 2) Kitaashigara (JP_{KAS}), 3) Hiratsuka, (JP_{HRJ}), and 3) Shibusawa (JP_{SBS}). In order to mitigate the effects of isolated spikes and noise in the measured RSD sequences (N_m), we applied a selection procedure on the RSD-derived rain rate (R_m), which has been computed through the moment of order 3.67 of N_m. This selection procedure individuates the wet periods (WPs) and, correspondingly, the dry periods (DPs), which have been defined as the instants where consecutive samples of R_m overcome the threshold of 0.1 mm/h for at least 10 min. Details on the obtained sample sizes and associated rain-rate statistics are listed in Table I.

In this table, we note that all data sets have a significant number of samples (see first column of Table I), and maximum rain rates up to 188 mm/h registered in ATH. All these distributions show positive values of skewness (obtained as the third-order moment of the distribution normalized to the third power of the standard deviation). This latter feature implies that the right tails of the distribution of the rain rate are longer so that the distribution is concentrated on small values of rain intensities. The highest values for the skewness are registered in U.K. and Iowa: this indicates the prevalence of stratiform events in these data sets. On the other hand, the lowest value of skewness is registered in the Florida data set, and this may be interpreted as the prevalence of convective events in that region. These conclusions, based on the moment analysis of the rain-rate distribution, will be also confirmed in Section III-A, where an algorithm for identifying the stratiform and the convective part of the RSD sequences is applied on all the available data sets.

B. Analysis of RSD Data

In this section, a systematic comparison of the four distinct approaches for best fitting the observed RSD and estimating the parameters of an assumed normalized gamma distribution is discussed. Many authors adopt the following expression to describe the volumetric size distribution of raindrops [1], [6], [14]:

$$N(D,\mathbf{p}) = N_w \cdot f(\mu) \cdot \left(\frac{D}{D_m}\right)^{\mu} \cdot \exp\left[-(4+\mu)\cdot\frac{D}{D_m}\right] \quad (1)$$

where $N(D,\mathbf{p})$ ($\mathrm{m}^{-3}\cdot\mathrm{mm}^{-1}$) is the number of drops per unit volume per unit size interval, D (mm) is the sphere-equivalent drop diameter, N_{w} ($\mathrm{m}^{-3}\cdot\mathrm{mm}^{-1}$), μ, and D_m (mm) are the intercept, shape, and mass-weighted mean diameter parameters, respectively. The vector $\mathbf{p} = [N_w, D_m, \mu]^{\mathrm{T}}$, with "T" indicating a matrix transpose, stands for the *time-dependent* parameter set of the modeled RSD.

The three parameters of the vector $\mathbf{p}$ in (1) are estimated by means of the gamma moment (GM) method [1], the ML method in two versions [4], [14], and the truncated GM (TGM) method [8]. The latter method is a numerical version of GM, taking into account the finite integration over a minimum and a maximum diameter. ML methods can be implemented in 1-D (ML1) and 3-D (ML3) versions. In the first case, the parameters N_w and D_m are estimated by using the GM method, whereas μ is retrieved by minimizing, for each diameter interval, the square difference between the observed RSD N_m and N given in (1). Eventually, if the latter minimization is performed by varying all the components of $\mathbf{p}$, the ML3 estimation is accomplished [14].

C. Results

All RSD-parameter estimation methods introduced in the previous paragraph have been implemented on the whole data set of disdrometer measurements (see Table I). The performance of each estimation technique is evaluated by computing the root-mean-square error (RMSE) between $N_m(D)$ and $N(D,\mathbf{p})$, as shown in Fig. 1. In the following, RMSE_X will indicate the RMSE of the method X.

In most cases, values of $\mathrm{RMSE}_{\mathrm{ML3}}$ smaller than $\mathrm{RMSE}_{\mathrm{ML1}}$, $\mathrm{RMSE}_{\mathrm{GM}}$, and $\mathrm{RMSE}_{\mathrm{TGM}}$ are noted in the range [0.5, 1.5] mm, but there are larger values of $\mathrm{RMSE}_{\mathrm{ML3}}$ for diameters beyond about 1.2 mm. For large raindrop diameters, the curves of $\mathrm{RMSE}_{\mathrm{ML1}}$, $\mathrm{RMSE}_{\mathrm{GM}}$, and $\mathrm{RMSE}_{\mathrm{TGM}}$ appear very close to each other, whereas for smaller diameters, $\mathrm{RMSE}_{\mathrm{ML1}}$ is positioned between $\mathrm{RMSE}_{\mathrm{ML3}}$ and $\mathrm{RMSE}_{\mathrm{GM}}$. The analysis of Fig. 1 would imply that the ML3 estimation method may be an accurate technique for the overall measured RSD best fitting but not a good candidate for the computation of higher moments of RSD, where the larger diameters play a major role. In addition, the TGM method, in most cases, seems to have a behavior very similar to the $\mathrm{RMSE}_{\mathrm{GM}}$ curves, whereas for the Virginia and ATH data sets and smaller diameters, $\mathrm{RMSE}_{\mathrm{TGM}}$ is larger than the RMSE curves associated with the other methods.

The analysis of Fig. 1 suggests, as in [10], that the ML1 method is a relatively robust choice for estimating the RSD parameters when compared with both TGM and GM methods. Based on the considerations just exposed, hereafter, the ML1 method will be considered for estimating D_m, N_w, and μ in (1).

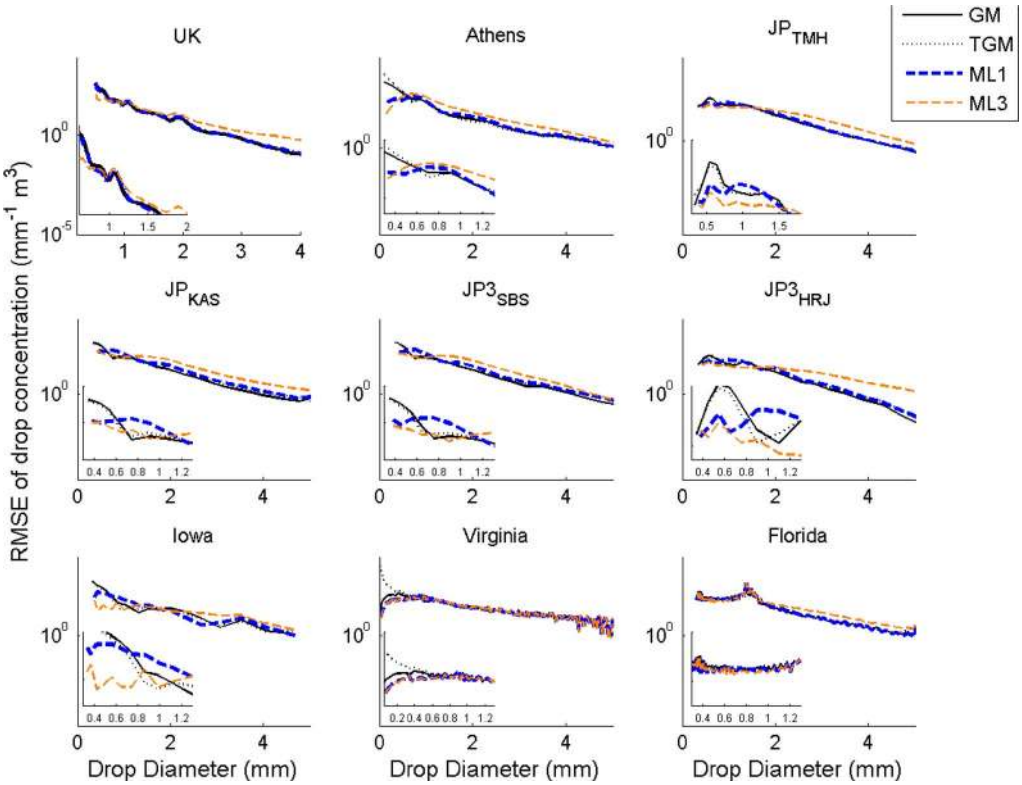

Fig. 1. RMSE between the measured RSD spectra and the estimated gamma ones as a function of raindrop diameter. The RMSE trend has been derived for the four implemented estimation methods: GM, TGM, ML1, and ML3. The inset subplots for each disdrometer measurement show an enlarged version of the whole view of the RMSE trend.

Note that this conclusion is coherent with what was found in [14], analyzing U.K. data only.

III. Raindrop Spectra of the Different Hydroclimate Regimes

To have a comprehensive picture of RSD best fitting, the scaled data $N_m(D)/N_w$ are plotted against the normalized drop diameter D/D_m and are shown in Fig. 2. Superimposed on the scaled disdrometer data, the corresponding scaled gamma distributions for values of μ equal to -3, 10, and 100 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 -3 to 10, but values of μ outside of this range may be also represented. It is interesting to note that the right tails of the normalized RSDs ($D/D_m \geq 3$ mm) are well characterized for values of μ that fluctuate around -3.

A. Convective–Stratiform Separation

To better understand the differences between the considered data sets, a classification procedure, similar to those adopted in [4] and [7], to separate the stratiform rain type from convective is applied. The classification procedure here adopted is based on the criterion that stratiform rain tends to be horizontally extended and weak in intensity as opposed to the convective regime, which generally shows high concentrated intensities. A sample of rain rate R_m at the instant t_i, denoted $R_m(t_i)$, is classified as stratiform if the values of R_m from $t_i - N$ to $t_i + N$ lie in the range [0.5, 10] mm/h, and their standard deviation is less than 1.5 [4], [6], [7]. On the other hand, a sample $R_m(t_i)$ is classified as convective if values of R_m from $t_i - N$ to $t_i + N$ are greater than 10 mm/h. Samples $R_m(t_i)$ that belong to neither the convective nor the stratiform cluster are classified as mixed type and will be excluded from the investigation. In this analysis, N has been set to five.

An example of the rain-regime classification procedure here adopted is shown in Fig. 3, where the sequences of RSD parameters and the rain rate extracted from the $\mathrm{JP_{KAS}}$ data set are considered. The analysis of this figure evidences that for convective samples (indicated with the red color), the time sequence of the mean diameter D_m evolves over the values of 1.5 mm, and the shape parameter μ is nearly zero (see the lower panel). Following this temporal evolution analysis, in the transition zone i.e., a mixture of convective, mixed-type (yellow color), and stratiform samples (blue color) D_m and μ slightly decrease and increase, respectively, with respect to the convective zone. On the other hand, in the stratiform time slots (horizontal shaded blue bars), D_m oscillates around 1 mm, whereas μ oscillates around five. It should be noted that when the rain rate intensity is very low (about 0.1–0.5 mm), the RSD parameters show a very high variability, as also found in [10].

B. Statistical Analysis of RSD Features

The conclusions, achieved from the example shown in Fig. 3, can be also extended to the whole disdrometer data set through the analysis of Fig. 4, where the RSD features are compared in mean term $\langle D_m \rangle$, $\langle \log_{10} N_w \rangle$, with the symbol "$\langle \cdot \rangle$" indicating the average operator. This figure suggests a clear separation between the convective and stratiform domains in the plane $\langle D_m \rangle$, $\langle \log_{10} N_w \rangle$. Indeed, a linear least square fit, applied

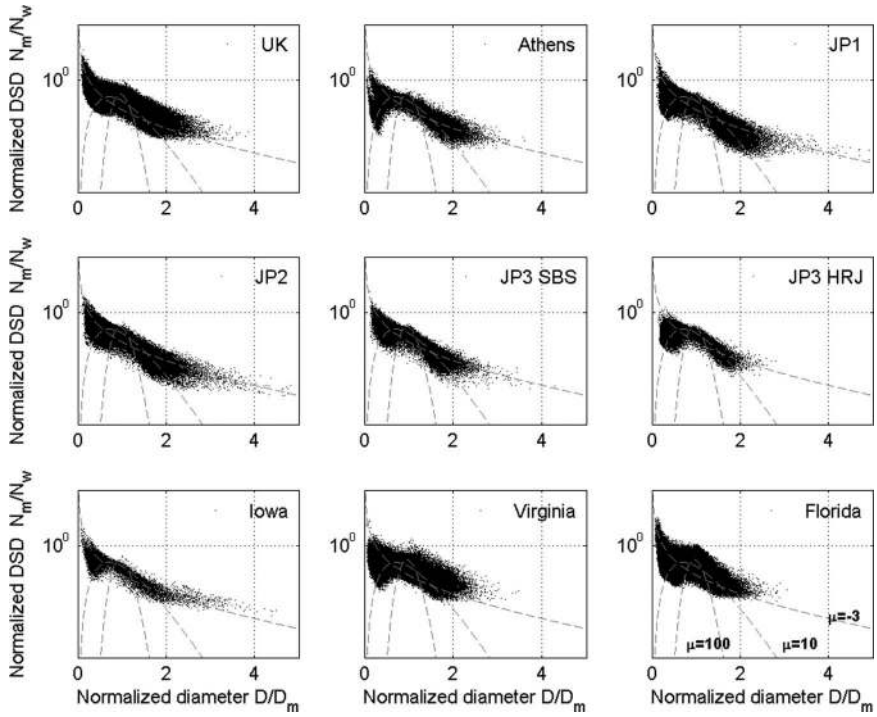

Fig. 2. Normalized RSD $N(D)/N_w$ versus D/D_m where N_w ($\mathrm{mm}^{-1}\mathrm{m}^{-3}$) and D_m (mm) are calculated by using the ML1 method. Dotted lines represent the gamma distribution, normalized as done for the measured DSD, when μ takes the values -3, 10, and 100 as indicated in the lower left panel.

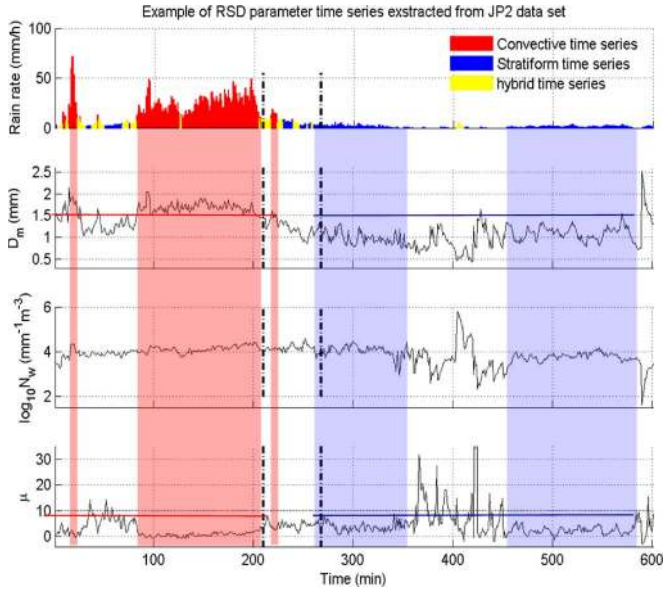

Fig. 3. Example of classification of RSD sequences.

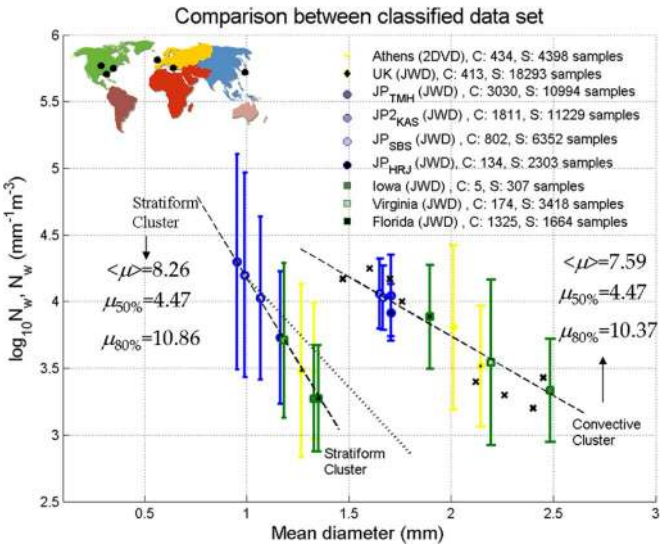

Fig. 4. Average of the mean diameter (D_m), against the mean intercept parameters $\log_{10}(N_w)$, both estimated by applying the ML1 method on the classified data set. The standard deviation bar of $\log_{10}(N_w)$ is shown as vertical bars as well. The least square fit of the average points is shown for both the stratiform (S) and the convective (C) cluster. In addition, the disdrometric measurements classified as convective in [4] have been indicated with the cross markers, whereas the dotted line is for their stratiform cases.

to the average values of D_m and $\log_{10} N_w$ for each climatic regime and for each cluster, points out an inverse relationship as $\langle \log_{10} N_w \rangle = p_1 \langle D_m \rangle + p_2$ for both the stratiform and convective rain types. The values of the p_1 and p_2 coefficients found in this analysis are listed in Table II. In addition, in order to compare these results with those shown in [4], the dotted-line curve for stratiform cases is shown in Fig. 4 together with the convective cluster (see the gray cross markers).

It is worth mentioning that for the analyzed data set, the two clusters are mainly discriminated by D_m [4]: for $\langle D_m \rangle > 1.5$ mm, the convective cluster is identified and vice versa. The average and standard deviation values of μ associated with convective and stratiform clusters (i.e., for the data sets altogether) are (7.59, 12.88) and (8.26, 12.57), respectively. The average μ of the stratiform cluster is slightly larger than that

TABLE II
COEFFICIENT FOR THE $\langle D_m \rangle$,$\langle \log_{10} N_w \rangle$ LINEAR RELATIONSHIP

	p_1	p_2
Stratiform cluster	-2.51	6.68
Convective cluster	-0.88	5.51

TABLE III
STATISTICAL INDICATORS OF THE WHOLE RSD PARAMETERS FOR THE STRATIFORM AND CONVECTIVE CLASSES

	mean		std		median		skewness		Kurtosis		80% percentile	
	C	S	C	S	C	S	C	S	C	S	C	S
D_m	1.76	1.13	0.42	0.39	1.68	1.14	1.83	0.17	9.28	3.31	2.00	1.44
$\log_{10}N_w$	3.96	3.84	0.37	0.76	4.02	3.72	-1.08	0.71	6.06	3.28	4.22	4.37
μ	7.59	8.26	12.88	12.57	4.47	4.47	5.73	4.28	40.88	27.01	10.37	10.86

(C): convective, (S): stratiform, D_m (mm), N_w (mm^{-1} m^{-3}).

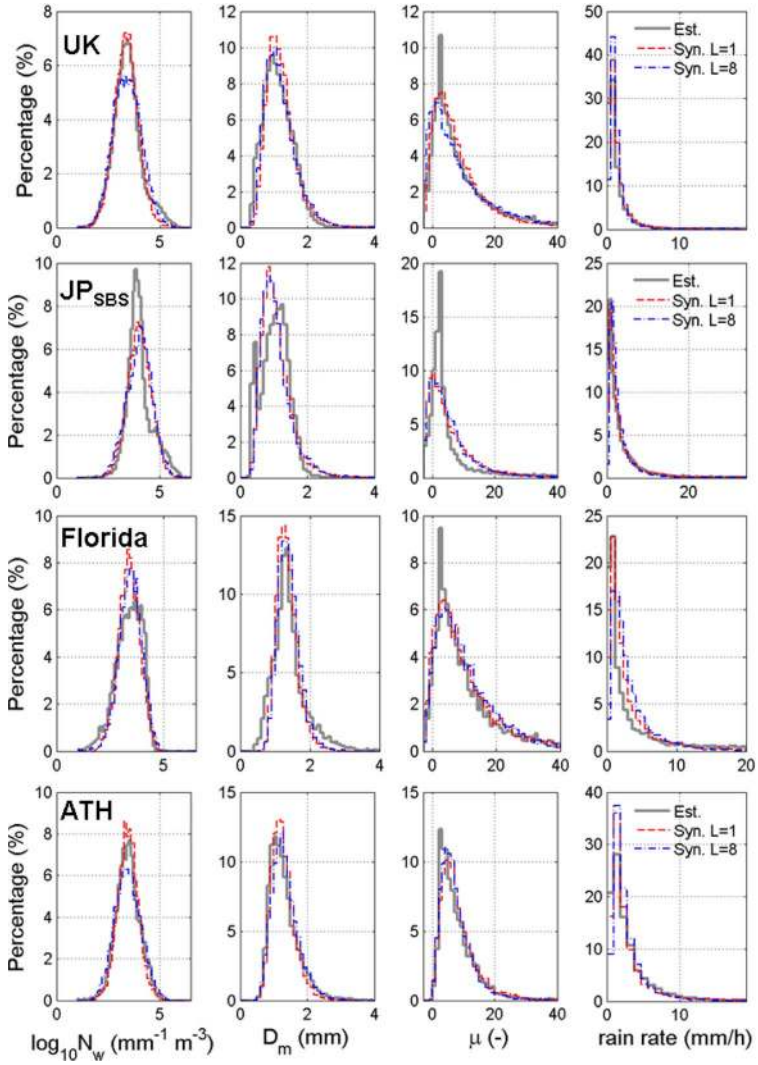

Fig. 5. Distributions of (gray lines) the estimated and synthesized RSD parameters and the synthesized ones for (dashed black lines) $L = 1$ and (dotted black lines) $L = 8$ for the UK, ATH, $\mathrm{JP_{SBS}}$, and Florida data sets.

associated with the convective one, and this behavior seems to be slightly different from that shown in [6] and [10]. A possible explanation is that low rain regimes may cause a large RSD variability [10] so that other definitions for the stratiform cluster may account for these different behaviors. In addition, as confirmed by the positive values of skewness and large values of kurtosis of μ, which are shown among other statistical indicators in Table III, the major variability of this parameter is due to extreme cases, which are aggregated on the right tail of the distribution. Indeed, if the median value of the distribution of μ is considered instead of the mean, no difference is noted between the convective and stratiform clusters (i.e., the median of μ is equal to 4.47 for both classes), and values of μ consistent with those found by other authors in past works [4], [6], [10] are obtained. On the other hand, it should be noted that the estimate of μ is the most critical, being strongly affected by disdrometer data quality.

IV. GENERATION OF RAINDROP SPECTRA TIME SERIES

A brief description of the autoregressive model, used to generate the time series of RSD parameters, is here discussed and then characterized for each data set. More details can be found in [14].

A. Autoregressive Generation of Raindrop Spectra Series

Rainfall is usually characterized by a significant space–time variability. Within a rainfall process, we can distinguish two

TABLE IV
PARAMETERS OF THE PARETO DISTRIBUTION FOR DP AND WP DURATIONS

	UK	ATH	JP_{TMH}	JP_{KAS}	JP_{SBS}	JP_{HRJ}	Iowa	Virginia	Florida
	a	a	a	a	a	a	a	a	A
WP	0.9	0.6	0.6	0.7	0.8	0.6	1.2	0.6	0.6
DP	$1.2\cdot10^{-4}$	$1.3\cdot10^{-4}$	$1.1\cdot10^{-4}$	$1.1\cdot10^{-4}$	$1.1\cdot10^{-4}$	$1.0\cdot10^{-4}$	$1.7\cdot10^{-4}$	$\cdot10^{-4}$	$1.8\cdot10^{-4}$

The parameter a is adimensional, b is equal to 10 min and 1 min for WP and DP respectively.

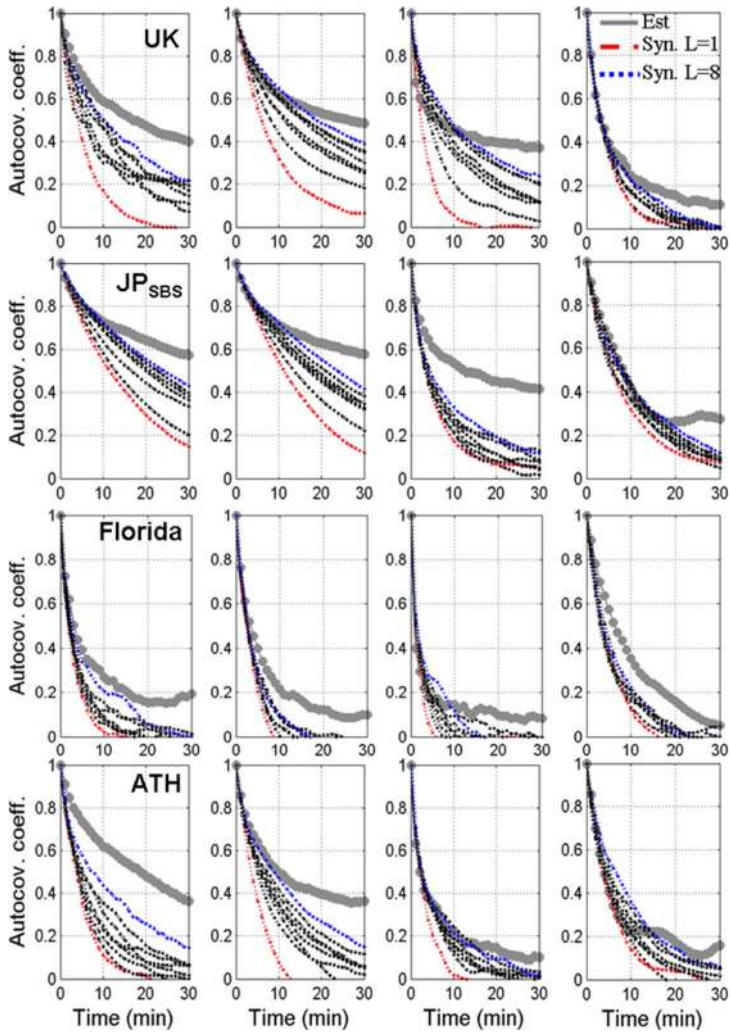

Fig. 6. Comparison between the autocovariance functions of (gray lines) the estimated RSD parameters and the synthesized ones for (dashed black lines) $L = 1$ and (dotted black lines) $L = 8$ for the U.K., ATH, JP_{SBS}, and Florida data sets.

different macroscopic phenomena: 1) the rain periods or WPs and 2) the no-rain periods or DPs. In order to model this "intermittent" behavior, the rainfall random process is considered as a renewal process, which is a generalization of the Poisson process, the latter being a continuous-time Markov process [12]. Within each WP, the time series of RSD parameters $\mathbf{p}(t)$ are modeled as a vector autoregressive (VAR) process of order L, named $\mathrm{VAR}_3(L)$ (i.e., future-time behavior is conditioned by L past time), thus taking into account the correlation properties of the RSD parameters. The general formulation of $\mathrm{VAR}_3(L)$ is given by

$$\mathbf{z}(t) = \sum_{i=1}^{L} \mathbf{D}(i) \cdot \mathbf{z}(t-i) + \boldsymbol{\varepsilon}(t) \tag{2}$$

where $\mathbf{z}(t-i)$ is the synthesized time sequence vector $\mathbf{z}$ at the ith time lag before, $\mathbf{D}(i)$ are 3×3 autoregressive coefficient matrices, and $\boldsymbol{\varepsilon}(t)$ is a 3×1 zero-mean Gaussian white noise vector, 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:

$$\mathbf{z}(t) = \ln\left[\mathbf{p}(t)\right] - \langle \ln\left[\mathbf{p}_X(t)\right] \rangle \tag{3}$$

where $\langle \ln[\mathbf{p}_X(t)] \rangle$ is the temporal ensemble mean of $\ln[\mathbf{p}_X(t)]$, and "X" indicates the chosen estimation method of RSD parameters.

Both the matrices $\mathbf{D}$ and the Gaussian noise ε can be retrieved from the estimated RSD parameter time series (see [14, Appendix B] for details). The Gaussian noise $\boldsymbol{\varepsilon}$ is described through its autocovariance matrix $\mathbf{S}_\varepsilon$ of the Gaussian white

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

	ATH				UK				JP_{TMH}				JP_{KAS}			
L	N_w	D_m	μ	*Rain*	N_w	D_m	μ	*Rain*	N_w	D_m	μ	*Rain*	N_w	D_m	μ	*Rain*
1	0.43	0.38	0.15	0.13	0.39	0.33	0.35	0.12	0.37	0.30	0.35	0.28	0.44	0.20	0.27	0.16
2	0.41	0.26	0.09	0.09	0.27	0.22	0.26	0.11	0.30	0.25	0.31	0.28	0.36	0.16	0.26	0.20
3	0.39	0.28	0.09	0.11	0.21	0.14	0.17	0.08	0.33	0.18	0.29	0.25	0.31	0.16	0.26	0.20
4	0.34	0.21	0.06	0.09	0.22	0.13	0.16	0.09	0.27	0.16	0.28	0.21	0.25	0.12	0.25	0.13
5	0.33	0.20	0.06	0.08	0.16	0.09	0.14	0.08	0.24	0.13	0.26	0.19	0.28	0.09	0.19	0.11
6	0.32	0.19	0.08	0.09	0.12	0.05	0.10	0.08	0.24	0.09	0.27	0.17	0.28	0.11	0.21	0.08
7	0.27	0.11	0.04	0.05	0.21	0.07	0.10	0.07	0.25	0.10	0.23	0.16	0.23	0.07	0.21	0.07
8	0.19	0.15	0.04	0.06	0.17	0.07	0.09	0.06	0.25	0.07	0.22	0.13	0.28	0.09	0.21	0.09

	JP_{HRJ}				Iowa				Virginia				Florida			
L	N_w	D_m	μ	*Rain*	N_w	D_m	μ	*Rain*	N_w	D_m	μ	*Rain*	N_w	D_m	μ	*Rain*
1	0.30	0.26	0.25	0.32	0.27	0.13	0.08	0.10	0.38	0.41	0.16	0.35	0.18	0.18	0.18	0.18
2	0.15	0.26	0.19	0.37	0.25	0.09	0.08	0.11	0.37	0.34	0.15	0.31	0.20	0.20	0.20	0.20
3	0.23	0.22	0.23	0.32	0.19	0.08	0.07	0.10	0.25	0.24	0.10	0.32	0.16	0.16	0.16	0.16
4	0.20	0.21	0.21	0.41	0.19	0.07	0.06	0.12	0.32	0.26	0.13	0.30	0.18	0.18	0.18	0.18
5	0.08	0.21	0.21	0.27	0.17	0.07	0.09	0.10	0.22	0.17	0.12	0.33	0.15	0.15	0.15	0.15
6	0.09	0.18	0.10	0.28	0.22	0.06	0.06	0.06	0.20	0.13	0.10	0.29	0.09	0.09	0.09	0.09
7	0.21	0.12	0.06	0.28	0.18	0.06	0.05	0.10	0.22	0.11	0.08	0.29	0.12	0.12	0.12	0.12
8	0.17	0.11	0.13	0.26	0.20	0.06	0.05	0.09	0.26	0.13	0.07	0.23	0.14	0.14	0.14	0.14

	JP_{SBS}			
L	N_w	D_m	μ	*Rain*
1	0.22	0.28	0.31	0.14
2	0.21	0.19	0.30	0.14
3	0.21	0.18	0.30	0.06
4	0.20	0.16	0.22	0.14
5	0.14	0.15	0.24	0.10
6	0.13	0.11	0.21	0.09
7	0.06	0.04	0.18	0.06
8	0.10	0.05	0.21	0.06

noise error. In order to allow a practical implementation of the proposed autoregressive model through (2) and, at the same time, to limit the length of this paper, the matrices $\mathbf{D}(1)$ and $\mathbf{S}_\varepsilon$ for the order $L = 1$ are given below for some selected climatic zones, which are the U.K., ATH, JP_{SBS}, and Florida data sets, respectively:

$$D(1) = \begin{bmatrix} 0.8595 & -0.2858 & -0.1008 \\ 0.0071 & 0.9475 & 0.0286 \\ -0.0356 & -0.0556 & 0.8171 \end{bmatrix}$$
$$S_\varepsilon = \begin{bmatrix} 0.3202 & -0.0576 & 0.0772 \\ -0.0576 & 0.0208 & -0.0227 \\ 0.0772 & -0.0227 & 0.1792 \end{bmatrix} \quad (4)$$
$$D(1) = \begin{bmatrix} 0.8716 & -0.5401 & -0.2398 \\ 0.0149 & 0.9688 & 0.0271 \\ -0.0354 & -0.5219 & 0.6176 \end{bmatrix}$$
$$S_\varepsilon = \begin{bmatrix} 0.2992 & -0.0508 & 0.0854 \\ -0.0508 & 0.0190 & -0.0295 \\ 0.0854 & -0.0295 & 0.1286 \end{bmatrix} \quad (5)$$
$$D(1) = \begin{bmatrix} 0.9070 & -0.2587 & -0.0640 \\ 0.0119 & 1.0024 & 0.0145 \\ -0.0212 & -0.3025 & 0.7808 \end{bmatrix}$$
$$S_\varepsilon = \begin{bmatrix} 0.1825 & -0.0364 & 0.0736 \\ -0.0364 & 0.0153 & -0.0214 \\ 0.0736 & -0.0214 & 0.1660 \end{bmatrix} \quad (6)$$
$$D(1) = \begin{bmatrix} 0.8711 & -0.0672 & -0.1293 \\ 0.0279 & 0.8775 & 0.0187 \\ 0.0447 & -0.0572 & 0.6045 \end{bmatrix}$$
$$S_\varepsilon = \begin{bmatrix} 0.4699 & -0.0668 & 0.2033 \\ -0.0668 & 0.0280 & -0.0268 \\ 0.2033 & -0.0268 & 0.3527 \end{bmatrix} \quad (7)$$

The mean subtraction in (3) holds under the assumption, valid in this work, of considering the WPs and then the RSD parameters $\mathbf{p}$ as a quasi-stationary process. In addition, to avoid the log-transform failure for negative values of μ, in (3), we have added to μ the absolute value of its minimum so that the translated μ is positive defined (when inverting (3), this translation must be corrected).

The choice of developing the VAR algorithm in (2) by using the logarithm of $\mathbf{p}(t)$, instead of directly using $\mathbf{p}(t)$, has been prompted by the consideration that: 1) under the hypothesis that the joint probability density function (pdf) of the estimated RSD parameters $\mathbf{p}$ follows a log-normal distribution, the estimated $\mathbf{z}_X(t)$ time sequence follows a Gaussian distribution; this is a characteristic required to apply an autoregressive model under the Gaussian hypothesis on $\boldsymbol{\varepsilon}(t)$, and 2) as shown in [13], 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 Fig. 5, where the marginal pdf's of the RSD parameters are shown with the gray color. To keep the figure comprehensible, we limited our attention on the U.K., ATH, JP_{SBS}, and Florida data sets.

In summary, to generate a time series of correlated RSD parameters, (2) has to be iterated for a given number of time steps [14]. This 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 f_p. The Pareto pdf for a given duration T is described as

$$f_p(T) = a(b^a) \cdot T^{-(a+1)} \quad \text{for} \quad T \geq b \quad (8)$$

where a and b are the parameters of the distribution. The choice of modeling the duration of DPs and WPs with a Pareto distribution is due to the fact that it fits better, in terms of the RMSE between the measured and synthesized WP and DP distributions, than other tested distributions such as exponential, log normal, gamma, and Weibull. From the definition given in Section II-A for WPs and DPs, the latter must have durations of, at least, 10 min and 1 min, respectively. As a result, the parameter b in (8) is equal to 10 for WPs and 1 for DPs, whereas the values of the parameter a are listed in Table IV for all considered climatic zones. On the other hand, within each DP, void values of RSD parameters $\mathbf{p}(t)$ have been considered to describe the no-rain events. The overall procedure is named VARMS.

B. Results

In order to asses its features, VARMS is tested, taking as input the time series of RSD parameters estimated for the various climatological regimes considered in this work. The output of VARMS (i.e., the synthesized time series of RSD parameters) is discussed in terms of the autocovariance properties and distributions of RSD parameters. Figs. 5 and 6 show examples of the properties of the synthesized RSD parameters and rain rates from the U.K., ATH, $\mathrm{JP_{SBS}}$, and Florida data sets. A point to note is that VARMS can reproduce, with a good agreement, the distributions of the estimated N_w, D_m, μ, and rain rate R (see Fig. 5), the latter being retrieved by computing the moment of order 3.67 of (1).

On the other hand, in Fig. 6, the normalized autocovariance function of the synthesized WP time series is plotted in the case of the orders from $L = 1$ (red dotted lines) and $L = 8$ (blue dotted lines). VARMS reconstruction follows the measured data (gray lines) with a discrete accuracy up to 10 min for N_w, D_m, or R. Indeed, for the Florida data set, we note that the autocovariance curves of synthesized R are all positioned below the measured ones, and this behavior indicates that in this case, VARMS techniques cannot accurately reproduce the correlation of the data. This conclusion can be supported by the fact that the Florida data show a high temporal variation of N_w, D_m, and μ (see how the autocorrelation functions of the Florida RSD parameters rapidly decrease with respect to the other data sets), which results in being difficult to follow.

In addition, we observe that in the case of $\mathrm{JP_{SBS}}$, shown in the same figure, the correlation of μ is not well captured, and this may be probably due to the fact that it is difficult, in some cases, to reproduce the high variability shown by the shape parameter μ with respect to the other parameters. Figs. 5 and 6 seem to indicate that as the order L of the VAR process increases, the autocovariance function of RSD parameters and the rain rate are better described (see Fig. 6) but without any appreciable improvement on their distributions (see Fig. 5). An error analysis between the measured and synthesized autocovariance functions for orders of the VAR process ranging from one to eight and for all available disdrometric data sets has revealed that the optimum order L_{opt} oscillates between five and eight. This can be observed in more detail in Table V, where the values of the RMSE between the measured and synthesized autocovariance functions for the four selected climatic ones are listed.

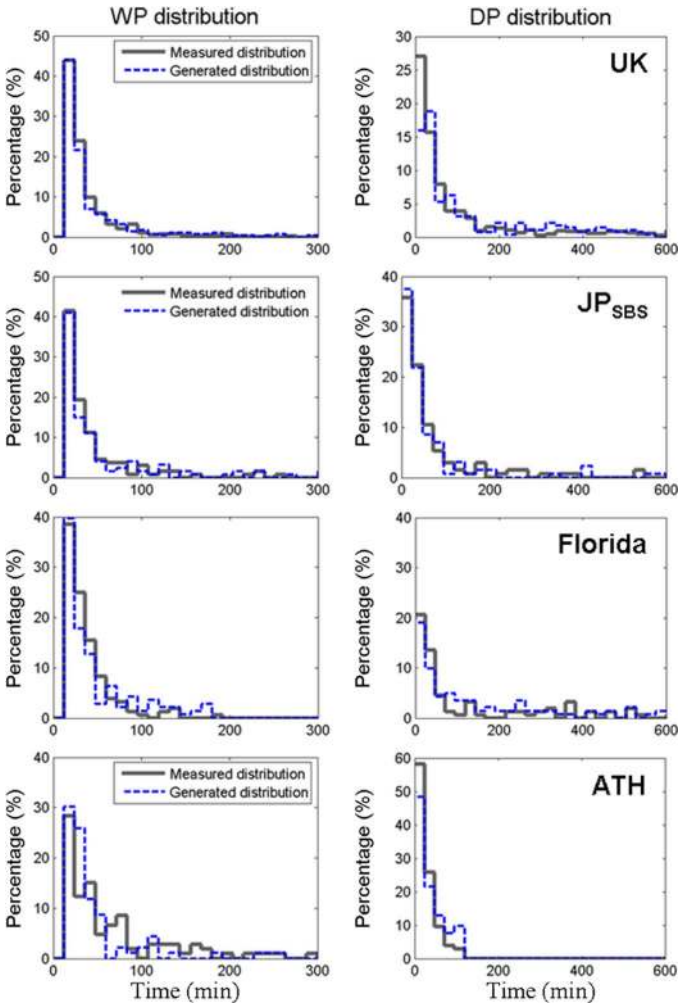

Fig. 7. Distributions of (gray lines) measured and (blue dotted line) synthesized (left panel) WPs and (right panel) DPs extracted from the U.K., $\mathrm{JP_{SBS}}$, Florida, and ATH data sets.

To qualitatively show the ability of VARMS to reproduce the intermittent behavior of the rainfall process, the synthesized and measured WP and DP distributions are shown in Fig. 7 for some examples extracted from the U.K., ATH, $\mathrm{JP_{SBS}}$, and Florida data sets. In this figure, it seems that the Pareto distribution describes well the WP and DP distributions, respectively. Behaviors similar to those shown in Fig. 7 have been found for the other data sets.

Finally, to further investigate the performances of VARMS, the estimated and synthesized RSD parameters with $L = 7$ for the U.K., ATH, $\mathrm{JP_{SBS}}$, and Florida data sets have been compared to each other and shown in Fig. 8 as scatterplots. By visual inspection of the plots in this figure, we can appreciate how VARMS is able to generate time sequences where the inverse proportional relationship between the intercept parameter N_w and the mean diameter D_m, evidenced in Fig. 3 for all measured RSD spectra, seems to be properly reproduced. The relationships between the other parameters (i.e., D_m versus μ and N_w versus μ) are not always reproduced with excellent accuracy but some exception are possible (e.g., D_m versus μ

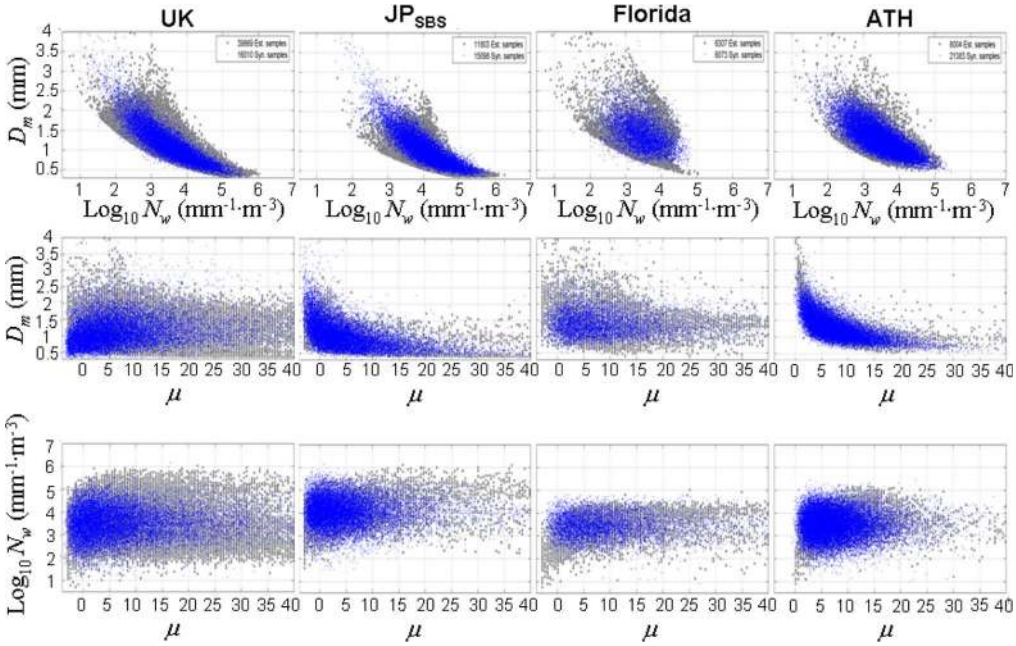

Fig. 8. Scatterplots of (blue markers) synthesized and (gray markers) estimated RSD parameters N_w, D_m, and μ using VARMS, which has been driven with (from the left to the right) the U.K., $\mathrm{JP_{SBS}}$, Florida, and ATH data.

for $\mathrm{JP_{SBS}}$ and ATH). The capability of VARMS to reproduce the mutual correlation of the temporal evolution of the three RSD parameters, particularly D_m and N_w, is, in some way, expected given the fairly good agreement in reproducing, in most cases and at least up to time lags of about 10 min, the measured autocovariance functions shown in Fig. 6. The aforementioned conclusions can be also extended to all other disdrometric data sets.

V. Conclusion

A large set of disdrometer measurements, collected in the period from 2001 to 2006 in the U.K., Japan, the U.S., and Greece have been analyzed in this work. The governing parameters of the normalized gamma distribution have been estimated using four methods: 1) the consolidated GM; 2) TGM; 3) ML1; and 4) ML3. The results of the comparison between these four RSD estimation methods have shown that they may lead to quite different results, and ML1 may represent a good compromise for estimating RSD higher order moments.

The characteristics of the various climatologic regions have been shown, and after a convective–stratiform classification procedure, a marked separation between these two classes has been found in the plane $\langle D_m \rangle$, $\langle \log_{10} N_w \rangle$. Based on RSD ML1-based estimation, a stochastic VAR model (VARMS) has been developed to randomly synthesize (generate) time series of normalized gamma RSD parameters. The performance of this model has been discussed in terms of autocovariance functions, marginal probability distributions, and intermittence properties of the generated RSD parameters. Further development of this work should be focused on the improvement of rain classification and to the extension of RSD generation to a spatial domain. A tentative work about the latter point can be found in [15].

References

[1] 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.

[2] D. Atlas, C. W. Ulbrich, F. D. Marks, Jr., E. Amitai, and C. R. Williams, "Systematic variation of drop size and radar-rainfall relation," *J. Geophys. Res.*, vol. 104, no. D6, pp. 6155–6169, 1999.

[3] 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.

[4] 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. Appl. Meteorol.*, vol. 60, no. 2, pp. 354–365, Jan. 2003.

[5] 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. 2004.

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

[7] J. Testud, S. Oury, R. A. Black, P. Amayenc, and X. Dou, "The concept of 'normalized' distribution 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.

[8] V. N. Bringi and V. Chandrasekar, *Polarimetric Doppler Weather Radar: Principles and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2001.

[9] 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.

[10] 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, pp. L10 803.1–L10 803.4, May 2005.

[11] M. Maki, T. D. Keenan, Y. Sasaki, and K. Nakamura, "Characteristics of the raindrop size distribution in tropical continental squall lines observed in Darwin," *J. Appl. Meteorol.*, vol. 40, no. 8, pp. 1393–1412, Aug. 2001.

[12] R. Lawrence Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.

[13] 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.

[14] M. Montopoli, F. S. Marzano, and G. Vulpiani, "Analysis and synthesis of raindrop size distribution time series from disdrometer data," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 2, pp. 466–478, Feb. 2008.
[15] F. S. Marzano and M. Montopoli, "Microwave modelling of rain attenuation fields using disdrometer measurements and stochastic methods," in *Proc. EuCAP*, Nov. 2007, pp. 1–5.

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 University of Rome "La Sapienza," Rome, Italy.

In 2005, he joined the Center of Excellence (CETEMPS), University of L'Aquila, as a Research Scientist on ground-based radar meteorology, with a special focus on C-band applications and processing techniques. Since 2006, he has also been a Research Assistant with the Department of Electrical Engineering and Information, University of L'Aquila.

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 University of Rome "La Sapienza," Rome, Italy, in 1988 and 1993, respectively.

After being a Lecturer with the University of Perugia, Perugia, Italy, in 1997, he joined the Department of Electrical Engineering and cofounded the Center of Excellence (CETEMPS), University of L'Aquila, L'Aquila, Italy, where he is currently the Vice Director. In 2005, he joined the Department of Electronic Engineering, University of Rome "La Sapienza," where he teaches courses on antennas and remote sensing. He is the author of more than 70 papers in international refereed journals and more than 120 extended abstracts in conference proceedings. 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 a member of the Italian Society of Electromagnetics (SIEM). He received the Young Scientist Award of the Twenty-Fourth General Assembly of the International Union of Radio Science (URSI) in 1993. In 1998, he received the Alan Berman Publication Award (ARPAD) from the Naval Research Laboratory, Washington, DC. Since 2001, he has been the Italian National Delegate for the European Cooperation in the Field of Scientific and Technical Research (COST) Action 720 on atmospheric profiling by ground-based remote sensing and Action 280 on satellite fade mitigation techniques. Since January 2004, he has been acting as an Associated Editor of the IEEE GEOSCIENCE REMOTE SENSING LETTERS. In 2004 and 2006, he was a Co-Guest Editor of the Special Issues on MicroRad for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.

Gianfranco Vulpiani 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 May 2005, respectively.

In 2001, he joined the Department of Physics and the Center of Excellence (CETEMPS) at the University of L'Aquila, as a Research Scientist on ground-based radar meteorology, with a special focus on C-band applications and polarimetric applications. He was a Visiting Scientist with Colorado State University, Fort Collins, during 2004. In February 2006, he was with the Météo-France radar group as a Postdoctoral Research Scientist within the framework of the European project FLYSAFE. Since March 2007, he has been with the Italian Department of Civil Protection, Rome, Italy, working on the national radar network management and development within the national early-warning system for multirisks management. He is a Reviewer for several international journals in remote sensing topics.

Marios N. Anagnostou (S'01–M'04) received the B.S. and M.Eng. degrees in electrical engineering from York University, Toronto, ON, Canada, and the Ph.D. degree in environmental engineering from the University of Connecticut, Storrs.

He is currently a Contractor Research Associate with the Institute for Inland Waters, Hellenic Center for Marine Research, Anavissos, Greece. His main research experience and interests are in rainfall microphysics and precipitation remote sensing on the basis of new radar remote-sensing systems such as X-band dual polarization (polarimetric). He has participated in numerous international field experiments, deploying mobile dual-polarization radar. Lately, he has been involved in underwater acoustics for quantitative estimation of rain and wind. He has authored or coauthored more than 13 journal papers in the areas of dual polarization, radar rainfall estimation, and hydrometeorological applications.

Emmanouil N. Anagnostou received the B.S. degree in civil and environmental engineering from the National Technical University, Athens, Greece, and the M.S. and Ph.D. degrees from the University of Iowa, Iowa.

He is currently the Team Leader of a Marie Curie Excellence Team with the Institute of Inland Waters, Hellenic Center for Marine Research, Anavissos, Greece, and an Associate Professor with the Department of Civil and Environmental Engineering, University of Connecticut, Storrs. His research is on precipitation estimation from space and ground-based sensors and the optimum assimilation of remote sensing data in atmospheric and hydrological models for the prediction of floods and other hydrological variables.

Dr. Anagnostou is a member of the Precipitation Constellation Subgroup, International Committee on Earth Observation Satellites, and of NASA's International Precipitation Science Team. He is the recipient of several prestigious awards, including the 1999 NASA Earth Sciences Directorate New Investigator Award, the 2002 NSF Geosciences Program CAREER Award, and the 2003 Outstanding Junior Faculty Award by the School of Engineering of the University of Connecticut. He is also the recipient of two top-level international awards: the 2002 EGU Plinius Medal for outstanding achievements in the field of natural hazards and important interdisciplinary activity involving remote sensing and hydrometeorological research areas and the 2005 Marie Curie Excellence Award in recognition to pioneering work on precipitation remote sensing and his innovations in multisensor global rainfall measurement.