

Francisco J. Tapiador, Ziad S. Haddad and Joe Turk

# PROBABILISTIC RDSD

# Why a probabilistic approach?

Empirical, statistical fits to the gamma distribution are not Probability Distribution Functions (PDFs)

$$n(D)_{MP} = N_0 e^{-\Lambda D}$$

$$n(D)_{Ulb} = N_0 D^\mu e^{-\Lambda D}$$

# Normalized RDSD

Empirical, statistical fits to the gamma function  
are not proper PDFs

$$n(D)_{Tes} = N_0^* F_\mu(D/D_m)$$

$$F_\mu(X) = \frac{\Gamma(4)}{4^4} \frac{(4 + \mu)^{4+\mu}}{\Gamma(4 + \mu)} X^\mu e^{-(4+\mu)X}$$

# What is a PDF, then?

- Empirical fits to gamma functions are not PDFs
- What is a PDF?

$$\left. \begin{array}{l} p(D) \in [0, 1] \\ \int p(D)dD = 1 \\ p(D_1 \cup D_2 \cup \dots \cup D_n) = \sum_{i=1}^n p(D_i) \end{array} \right\} p(D) \in \mathbb{R}, p(D) \geq 0$$

# What is a PDF, then?

- Empirical fits to gamma functions are not PDFs
- What is a PDF?
- So for instance, this expression qualifies as a (proper) PDF:

$$p(D) = D^\mu \Lambda^{\mu+1} \frac{e^{-\Lambda D}}{\Gamma(\mu + 1)}$$

This is the probability of finding a drop of diameter D into the population

## How do you move from PDFs to RDSDs?

- A Rain Drop Size Distribution  $n(D)$  is simply PDF times the number of drops:

$$p(D) = D^\mu \Lambda^{\mu+1} \frac{e^{-\Lambda D}}{\Gamma(\mu + 1)}$$

$$n(D) = N_T \cdot p(D) = N_T D^\mu \Lambda^{\mu+1} \frac{e^{-\Lambda D}}{\Gamma(\mu + 1)}$$

This is the number of drops of diameter  $D$  we expect to find in the population

# How this differs from other RDSDs?

- Only apparently similar to for instance Ulbrich's RDSD:

$$n(D) = N_T \cdot p(D) = N_T D^\mu \Lambda^{\mu+1} \frac{e^{-\Lambda D}}{\Gamma(\mu + 1)}$$

$$n(D)_{Ulbrich} = N_0 D^\mu e^{-\Lambda D}$$

# How this differs from other RDSDs?

- Only apparently similar to for instance Ulbrich's RDSD:

$$n(D) = N_T \cdot p(D) = N_T D^\mu \Lambda^{\mu+1} \frac{e^{-\Lambda D}}{\Gamma(\mu + 1)}$$

$$n(D)_{Ulb} = N_0 D^\mu e^{-\Lambda D}$$



# Why do we really need RDSDs to be probabilistic (and $N_T$ -linked with PDFs)?

1. Mathematical consistency: robust parameter estimation requires PDFs.
2. Physical modeling: RDSD comes from a random process.
3. We want a coherent set of units.
4. To build a Z/R relationship independent on  $N_T$
5. To analyze microphysics in terms of physically-meaningful parameters [not a and b].

# Integral parameters

from

$$n(D) = N_T \cdot p(D) = N_T D^\mu \Lambda^{\mu+1} \frac{e^{-\Lambda D}}{\Gamma(\mu + 1)}$$

to

Z, R, W, Z<sub>e</sub>, P<sub>s</sub>, P<sub>e</sub>, P<sub>a</sub> ...

# Z in terms of [gamma] PDFs

$$Z = k_1 \int n(D) D^6 dD = k_1 \int N_{Tp}(D) D^6 dD =$$

# Z in terms of [gamma] PDFs

$$\begin{aligned} Z &= k_1 \int n(D) D^6 dD = k_1 \int N_T p(D) D^6 dD = \\ &= k_1 \int N_T D^\mu \Lambda^{\mu+1} \frac{e^{-\Lambda D}}{\Gamma(\mu + 1)} D^6 dD = \end{aligned}$$

## Z in terms of [gamma] PDFs

$$\begin{aligned} Z &= k_1 \int n(D) D^6 dD = k_1 \int N_T p(D) D^6 dD = \\ &= k_1 \int N_T D^\mu \Lambda^{\mu+1} \frac{e^{-\Lambda D}}{\Gamma(\mu+1)} D^6 dD = \\ &= k_1 N_T \frac{\Lambda^{\mu+1}}{\Gamma(\mu+1)} \int D^{\mu+6} e^{-\Lambda D} dD \end{aligned}$$

It happens that the integral is analytically solvable

$$\int D^{\mu+6} e^{-\Lambda D} dD = \Lambda^{-(\mu+7)} \Gamma(\mu+7)$$

It happens that the integral is analytically solvable

$$\int D^{\mu+6} e^{-\Lambda D} dD = \Lambda^{-(\mu+7)} \Gamma(\mu+7)$$

$$\begin{aligned} Z &= k_1 N_T \frac{\Lambda^{\mu+1}}{\Gamma(\mu+1)} \Lambda^{-(\mu+7)} \Gamma(\mu+7) = \\ &= k_1 N_T \Lambda^{-6} \frac{\Gamma(\mu+7)}{\Gamma(\mu+1)} \end{aligned}$$

# The same for R

Assuming this fall speed modeling  $v(D) = v_1 D^{v_2}$

$$R = k_3 \cdot v_1 N_T \Lambda^{-(3+v_2)} \frac{\Gamma(\mu + v_2 + 4)}{\Gamma(\mu + 1)}$$



But instruments have limits for min and max diameters

$$Z = k_1 N_T \frac{\Lambda^{\mu+1}}{\Gamma(\mu+1)} \Lambda^{-(\mu+7)} \Gamma(\mu+7) =$$
$$= k_1 N_T \Lambda^{-6} \frac{\Gamma(\mu+7)}{\Gamma(\mu+1)}$$

$$Z = k_1 N_T \Lambda^{-6} \frac{\gamma(\mu+7, D_{max} \cdot \Lambda) - \gamma(\mu+7, D_{min} \cdot \Lambda)}{\Gamma(\mu+1)}$$

$$R = k_3 \cdot v_1 N_T \Lambda^{-(3+v_2)} \frac{\Gamma(\mu+v_2+4)}{\Gamma(\mu+1)}$$

$$R = k_3 v_1 N_T \Lambda^{-(3+v_2)} \frac{\gamma(\mu+4+v_2, D_{max} \Lambda) - \gamma(\mu+4+v_2, D_{min} \Lambda)}{\Gamma(\mu+1)}$$

# What about parameters?

$N_T$  is clear, but what is the physical meaning of  $\Lambda$  and  $\mu$ ?

$$p(D) = D^\mu \Lambda^{\mu+1} \frac{e^{-\Lambda D}}{\Gamma(\mu + 1)}$$

Since  $p(D)$  is a PDF, we can apply the MOM

Method of Moments:

$$\Lambda = \frac{m}{\sigma^2}$$
$$\mu = \frac{m^2 - \sigma^2}{\sigma^2}$$

All integral parameters will now depend on  $m$ ,  $\sigma^2$ , and  $N_T$

$$\Lambda = \frac{m}{\sigma^2}$$
$$\mu = \frac{m^2 - \sigma^2}{\sigma^2}$$

$$Z = k_1 N_T \Lambda^{-6} \frac{\gamma(\mu + 7, D_{max} \cdot \Lambda) - \gamma(\mu + 7, D_{min} \cdot \Lambda)}{\Gamma(\mu + 1)}$$

$$R = k_3 v_1 N_T \Lambda^{-(3+v_2)} \frac{\gamma(\mu + 4 + v_2, D_{max} \Lambda) - \gamma(\mu + 4 + v_2, D_{min} \Lambda)}{\Gamma(\mu + 1)}$$

# Generalized moments

$$M_{x,y} \equiv \frac{\int_0^{\infty} N(D) D^x dD}{\int_0^{\infty} N(D) D^y dD} :$$

$D_m$  [ $N_T$  independent]

$$D_m = \Lambda^{-1} \frac{\gamma(\mu + 5, D_{max}\Lambda) - \gamma(\mu + 5, D_{min}\Lambda)}{\gamma(\mu + 4, D_{max}\Lambda) - \gamma(\mu + 4, D_{min}\Lambda)}$$

NB:  $D_m = (\mu+4)/\Lambda = m+3(\sigma^2/m)$

# W [and scaled intercept param $N_w$ ]

$$W = k_2 N_T \frac{\Lambda^{\mu+1}}{\Gamma(\mu+1)} \Lambda^{-(\mu+3)} \Gamma(\mu+4) =$$
$$= k_2 N_T \Lambda^{-3} \frac{\Gamma(\mu+4)}{\Gamma(\mu+1)}$$

$$N_w \equiv \frac{4^4}{\pi \rho_w} \frac{W}{D_m^4}$$

$$N_w = \frac{4^4 k_2}{\pi \rho_w} \cdot N_T \cdot \Lambda \frac{[\gamma(\mu+4+v_2, D_{max}\Lambda) - \gamma(\mu+4+v_2, D_{min}\Lambda)]^5}{\Gamma(\mu+1) [\gamma(\mu+5, D_{max}\Lambda) - \gamma(\mu+5, D_{min}\Lambda)]^4}$$

$N_w$  is now made explicitly dependent on  $m$ ,  $\sigma^2$  and  $N_T$

# Polarimetric radar

$$Z_e = \frac{\lambda^4}{\pi^5 |K_w|^2} k_4 \frac{\Lambda^{\mu+1}}{\Gamma(\mu+1)} N_T \int_{D_{min}}^{D_{max}} D^\mu e^{-\Lambda D} \sigma_b(D, \lambda) dD$$

$$k = 4.343 \cdot k_5 \frac{\Lambda^{\mu+1}}{\Gamma(\mu+1)} N_T \int_{D_{min}}^{D_{max}} D^\mu e^{-\Lambda D} \sigma_e(D, \lambda) dD$$

$$P_s = k_6 \frac{\Lambda^{\mu+1}}{\Gamma(\mu+1)} N_T \int_{D_{min}}^{D_{max}} D^\mu e^{-\Lambda D} \sigma_s(D, \lambda) dD$$

$$P_e = k_6 \frac{\Lambda^{\mu+1}}{\Gamma(\mu+1)} N_T \int_{D_{min}}^{D_{max}} D^\mu e^{-\Lambda D} \sigma_e(D, \lambda) dD$$

$$P_a = N_T \frac{\int_{D_{min}}^{D_{max}} D^\mu e^{-\Lambda D} \sigma_s(D, \lambda) a(D, \lambda) dD}{\int_{D_{min}}^{D_{max}} D^\mu e^{-\Lambda D} \sigma_s(D, \lambda) dD}$$

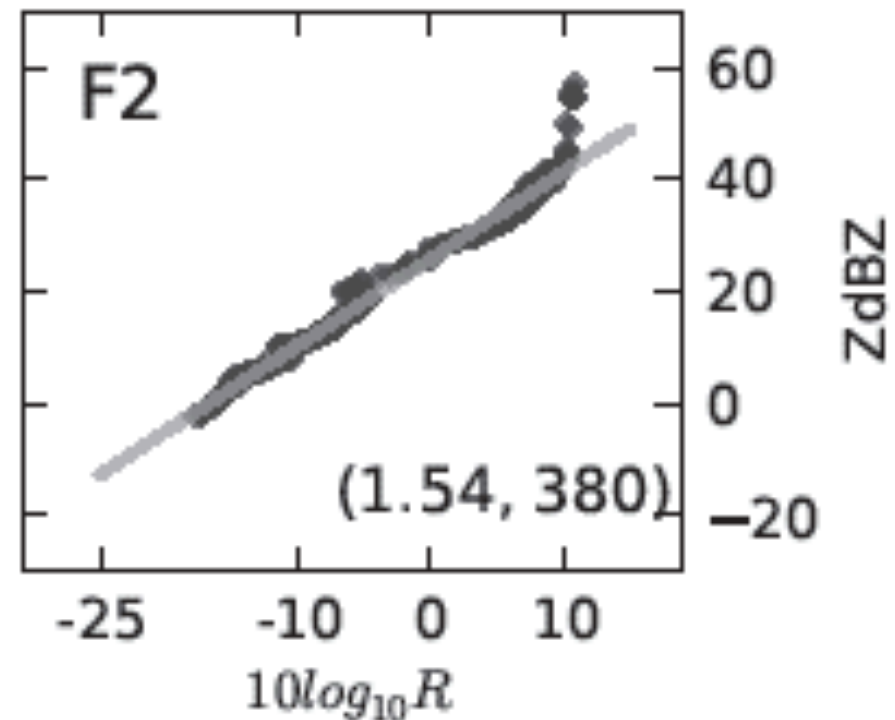
Using those equations, it is possible to retrieve  $N_T$  and build tables for  $m$  and  $\sigma^2$ .

# Z-R relationship

$Z=a \cdot R^b$  expressions are the result of correlating Z with R (statistical)

a and b do not have physical meaning

There is not such thing called 'instantaneous Z-R' [but a radar beam is pretty instantaneous]



# Natural Z-R relationship for a gamma distribution

$$\frac{Z}{R} = \frac{k_1}{k_3} \cdot v_1^{-1} \cdot \Lambda^{-(3+v_2)} \frac{\Gamma(\mu + 7)}{\Gamma(\mu + 4 + v_2)}$$

$$R = T \cdot Z$$

$$T \equiv \frac{k_3}{k_1} \cdot v_1 \cdot \Lambda^{(3-v_2)} \frac{\gamma(\mu+4+v_2, D_{max}\Lambda) - \gamma(\mu+4+v_2, D_{min}\Lambda)}{\gamma(\mu+7, D_{max}\Lambda) - \gamma(\mu+7, D_{min}\Lambda)}$$

T doesn't depend on  $N_T$ . It only depends on  $m$  and  $\sigma^2$



Gamma  
modeling  
means

You have three fundamental  
variables for the integral  
parameters:

$$N_T, \{m, \sigma^2\}$$

Z/R depends on

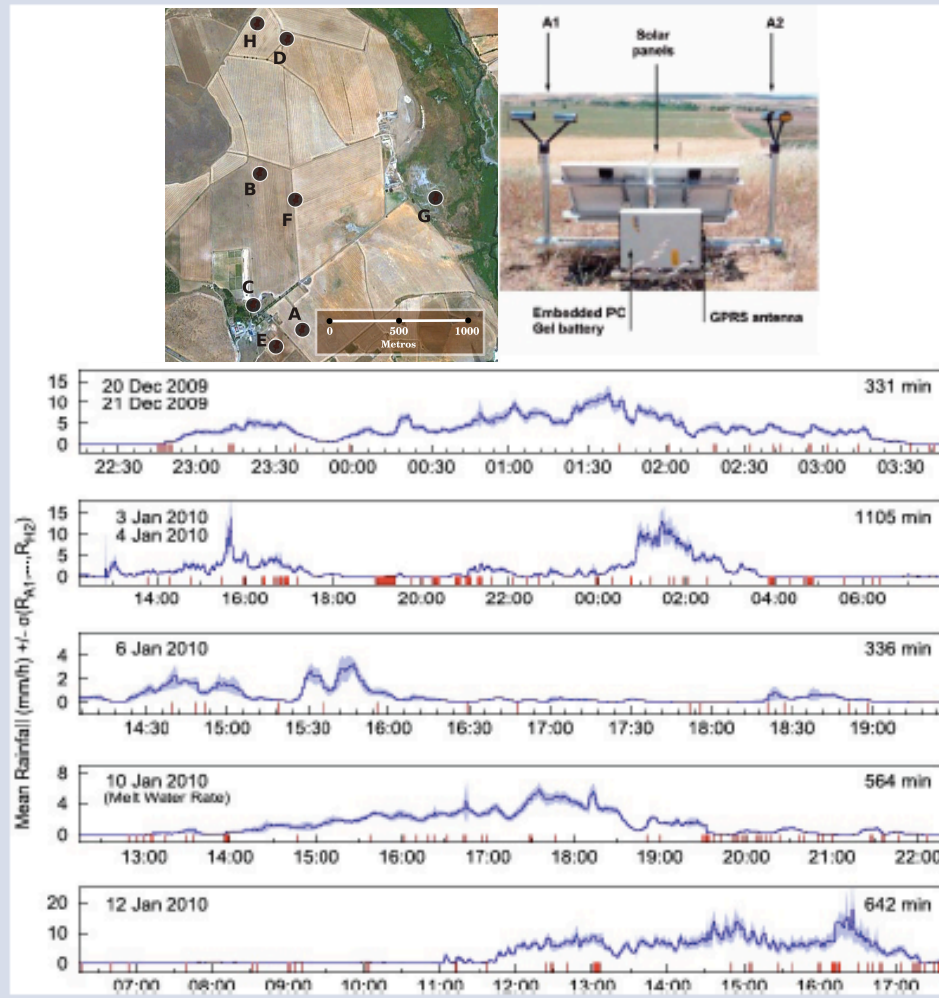
$$m, \sigma^2$$

The proof is in the eating



# Experimental setup 1

## Medium-scale variability



In 2010, we used 16 Parsivel disdrometers (in a dual setup to ensure consistency) to analyze the spatial variability of the RSD within a DPR-size pixel.

The experiments were made in central Spain, which has a semiarid climate with moderate rain rates, and thus within Parsivels' known limitations.

# Experimental setup 2

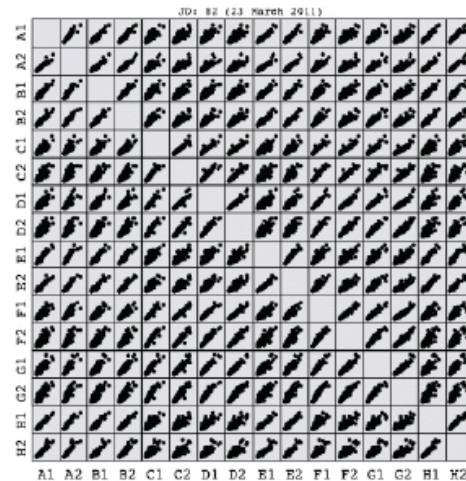
## Small-scale variability



In 2011, we located 16(+2) Parsivels to analyze the consistency of the instruments, the spatial variability of the RSD at decimeter scale, and to cross-compare the new Parsivel<sup>2</sup> instruments.

The experiments were made in Toledo, and included a sonic anemometer.

We found that the Parsivels provided consistent estimates of the RSD for moderate rainfall rates such as those found in Toledo.



Tapiador, F.J., Turk, J., Petersen, W., Hou, A.Y., García-Ortega, E., Machado, L.A.T, Angelis, C.F., Salio, P., Kidd, C., Huffman, G.J. and de Castro, M. 2011. Global Precipitation Measurement: Methods, Datasets and Applications. *Atmospheric Research*, accepted October 2011

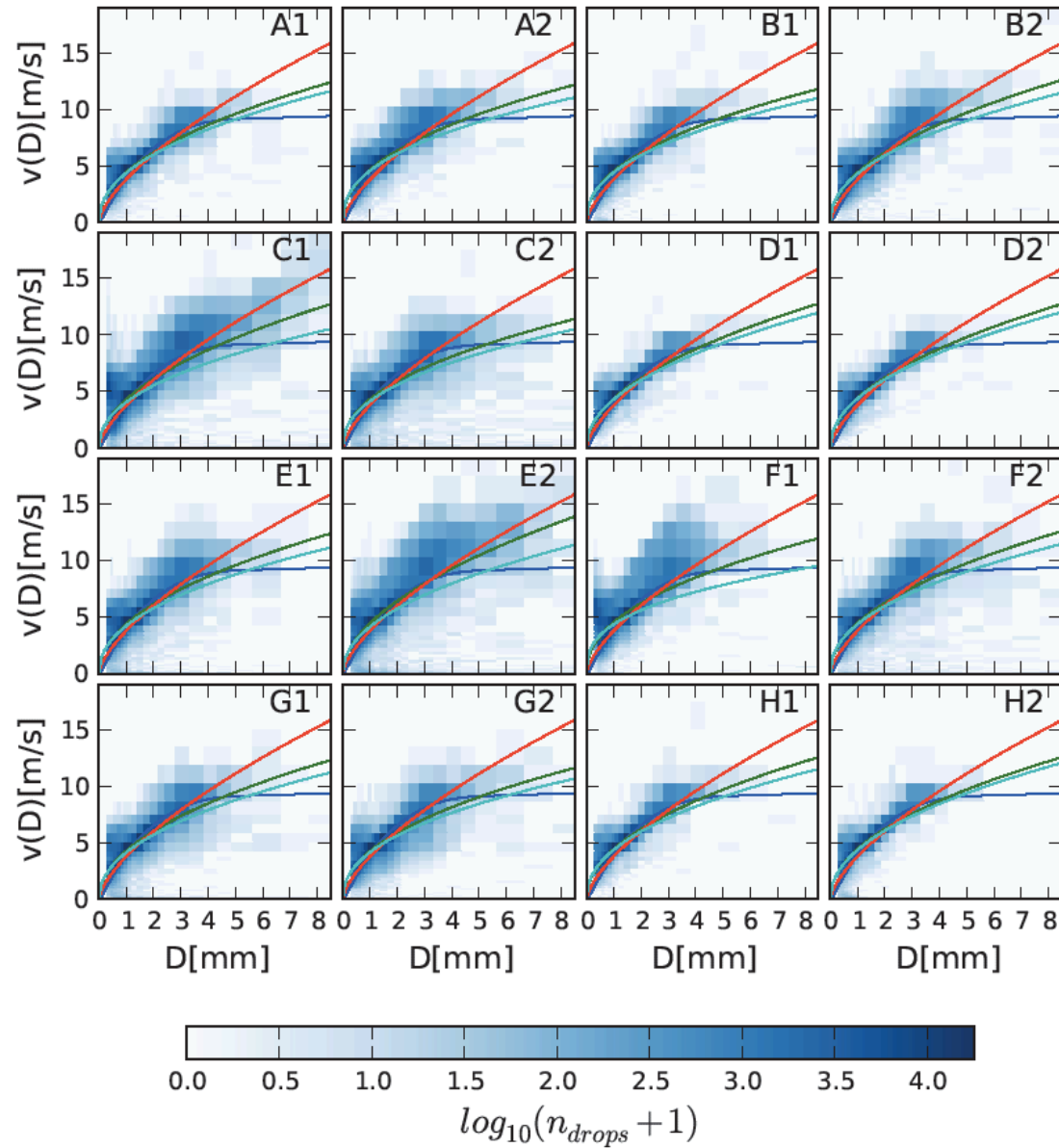
# Transient setup 3



$$v(D) = 9.65 - 10.3e^{-0.6D}$$

## Without the filter

- Red:  $v(D): 3.78D^{0.67}$
- Green: linear  $v(D) = \gamma D^\delta$
- Light Blue: non-linear  $v(D) = \gamma D^\delta$
- Dark Blue: non-sphericity correction

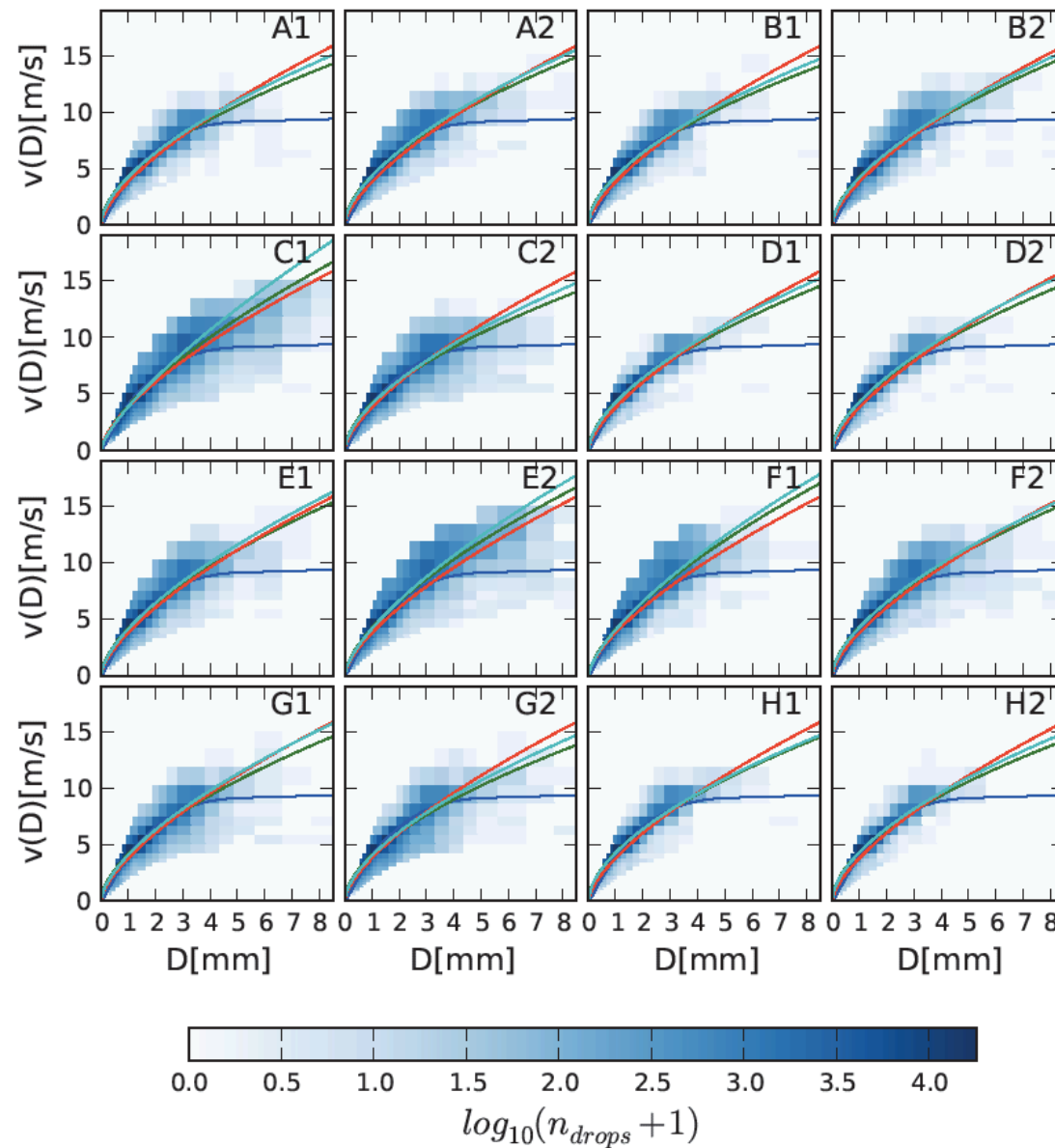


Filtering with

$$v(D) = 9.65 - 10.3e^{-0.6D}$$

With the filter

- Red:  $v(D): 3.78D^{0.67}$
- Green: linear  $v(D) = \gamma D^\delta$
- Light Blue: non-linear  $v(D) = \gamma D^\delta$
- Dark Blue: non-sphericity correction

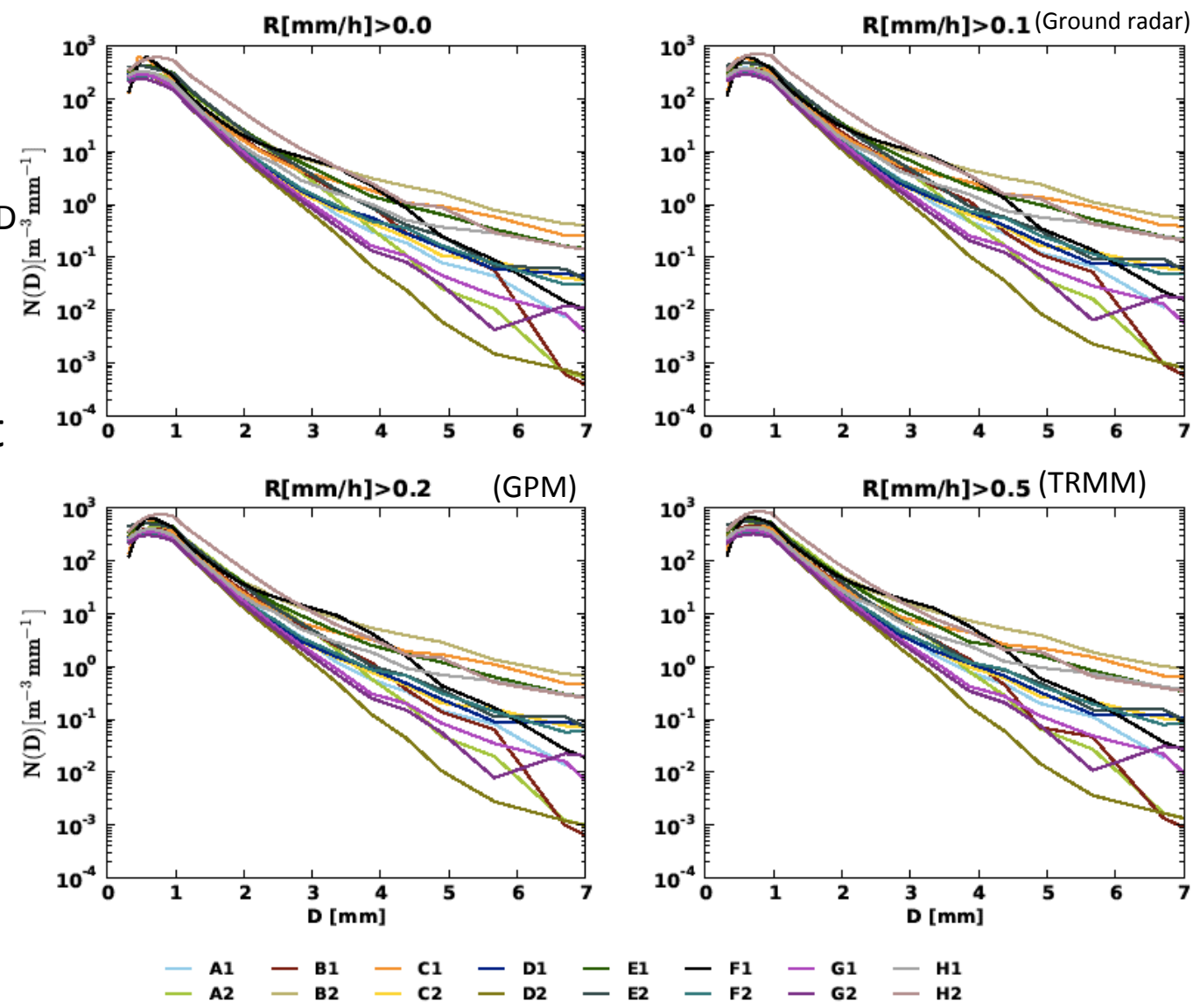


# Composite DSD

Filtering with  
 $v(D) = 9.65 - 10.3e^{-0.6D}$

Effects on different  
 R thresholds

No filter



Checa, R. 2012, PhD dissertation, Director: Francisco J. Tapiador

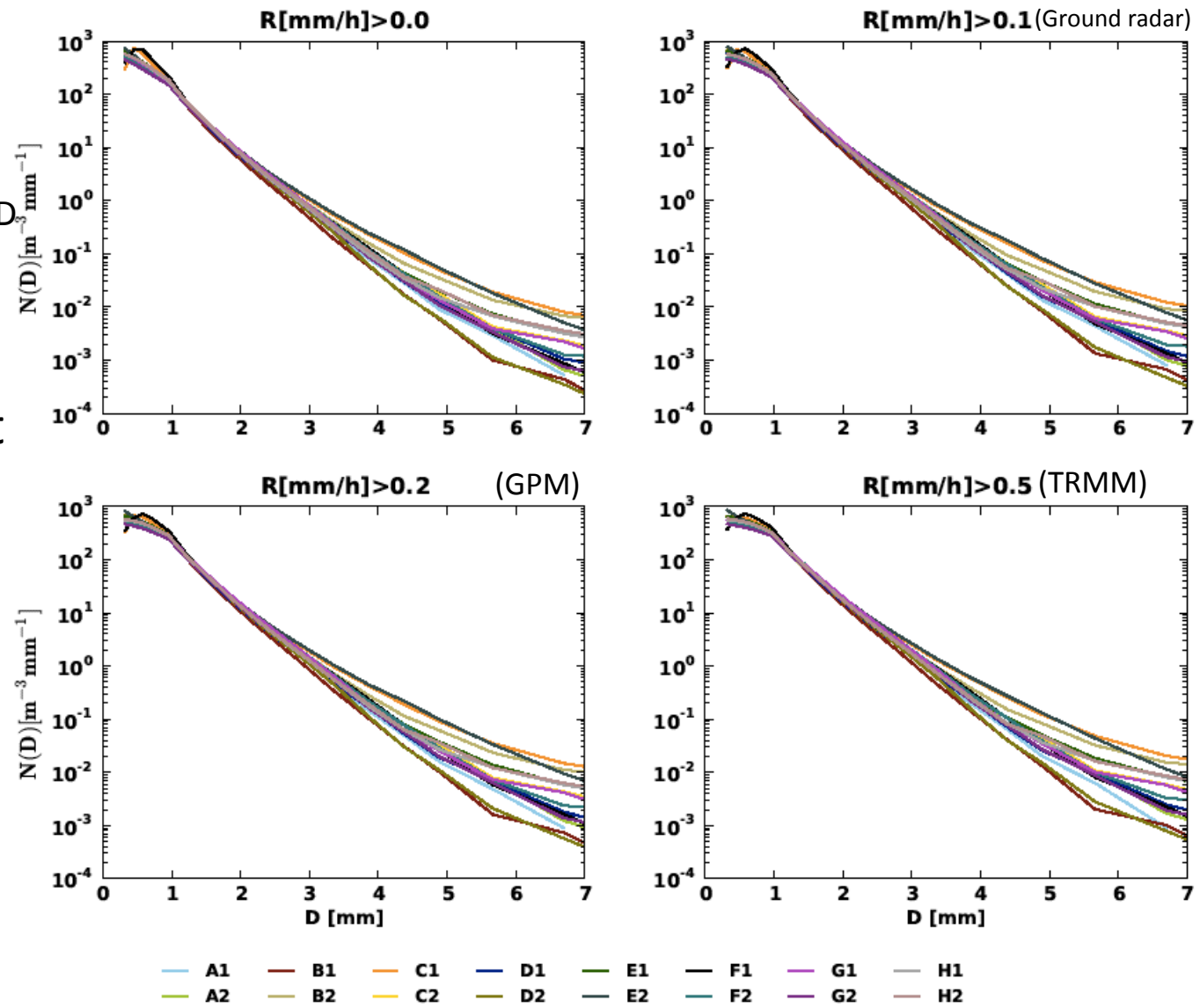


# Composite DSD

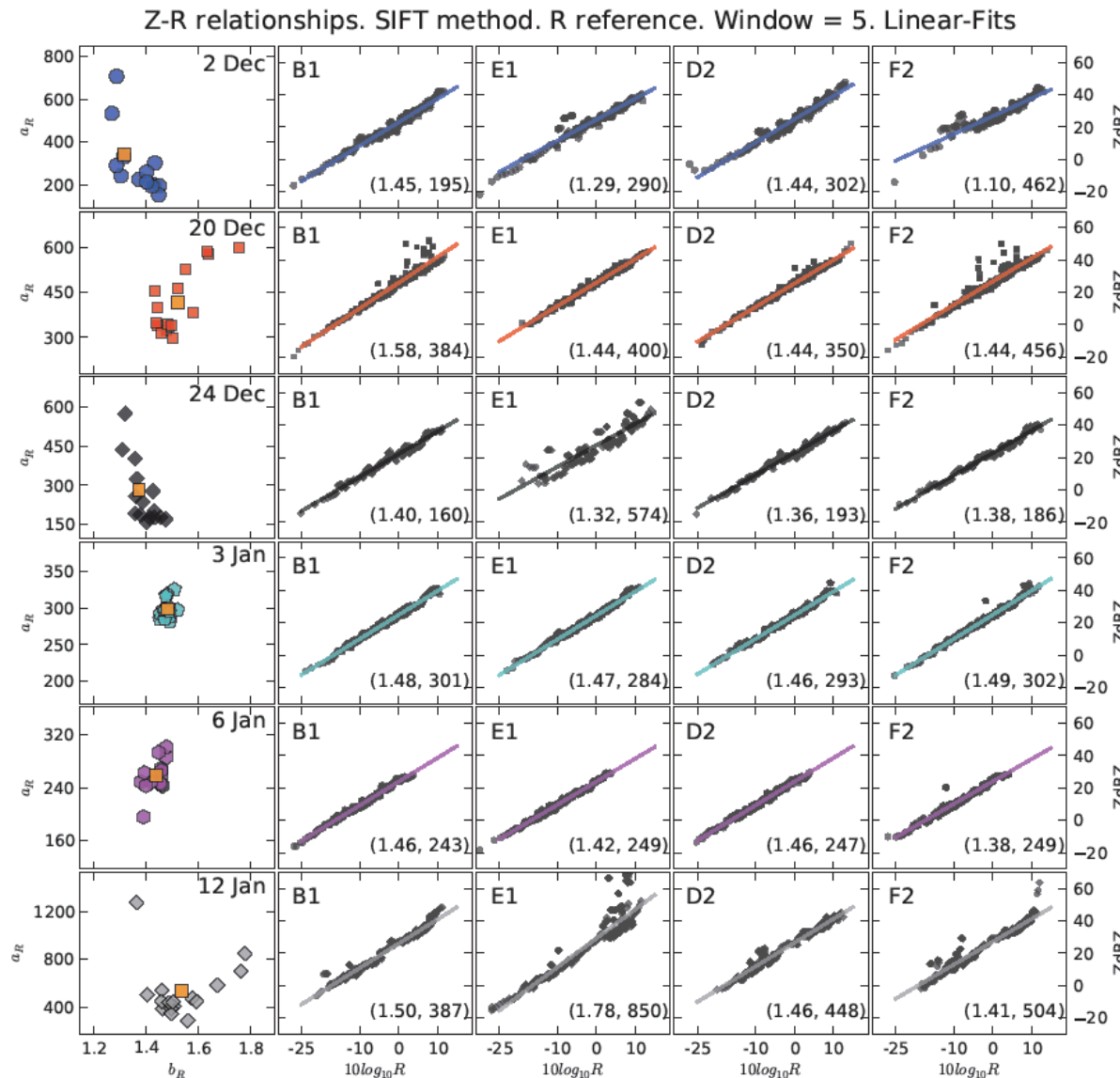
Filtering with  
 $v(D) = 9.65 - 10.3e^{-0.6D}$

Effects on different  
 R thresholds

Filter applied



Checa, R. 2012, PhD dissertation, Director: Francisco J. Tapiador



Variability of a and b

SIFT method  
 Window=5

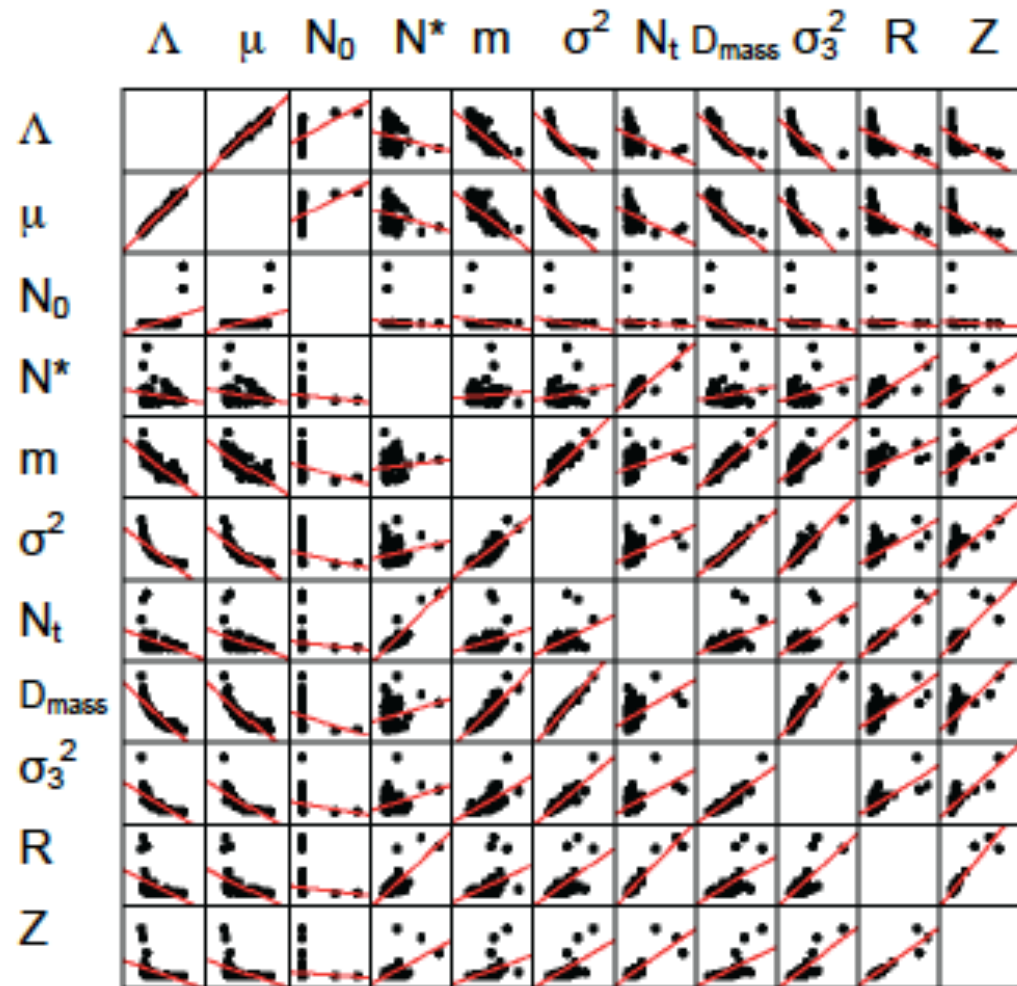
[Sequential Intensity  
 Filtering Technique, Lee  
 and Zawadzki 2005]

Aim: filtering out stochastic  
 variability due to poor  
 sampling of large drops

# Testing the probabilistic PDF

Cross-correlations for one disdrometer, four episodes, 5 min accumulation

March 4, 2011



- a) We measure  $m$ ,  $\sigma^2$ ,  $R$  and  $Z$  with a disdrometer  
 b<sub>1</sub>) We calculate  $T=R/Z$  from measured  $R$  and  $Z$   
 [call it **T empirical**]  
 b<sub>2</sub>) We also use the analytical formulae to derive  $T$   
 [call it **T analytical**] using two values:  $m$  and  $\sigma^2$

$$R = T \cdot Z$$

$$T \equiv \frac{k_3}{k_1} \cdot v_1 \cdot \Lambda^{(3-v_2)} \frac{\gamma(\mu+4+v_2, D_{max}\Lambda) - \gamma(\mu+4+v_2, D_{min}\Lambda)}{\gamma(\mu+7, D_{max}\Lambda) - \gamma(\mu+7, D_{min}\Lambda)}$$

$$\Lambda = \frac{m}{\sigma^2}$$

$$\mu = \frac{m^2 - \sigma^2}{\sigma^2}$$

---

Parameters

---

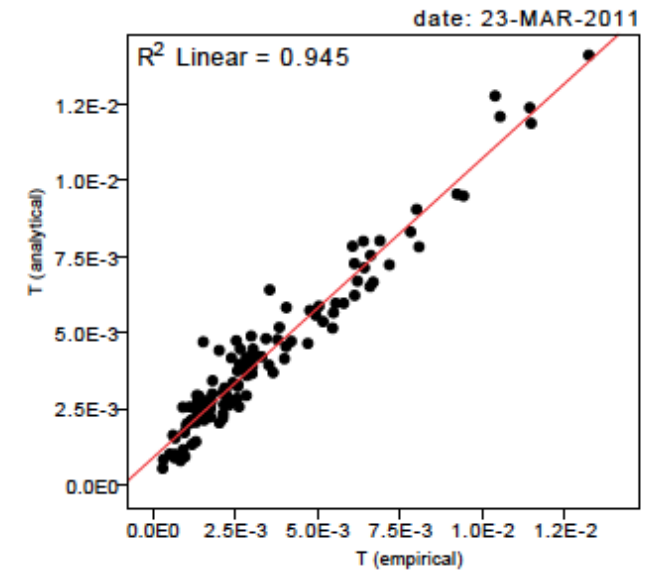
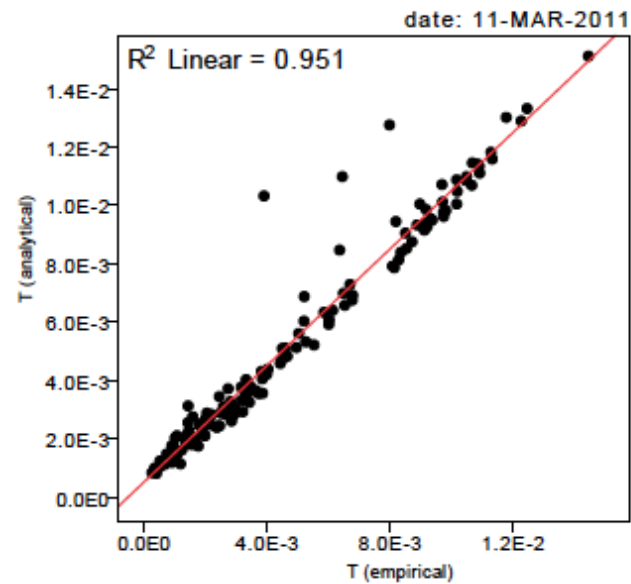
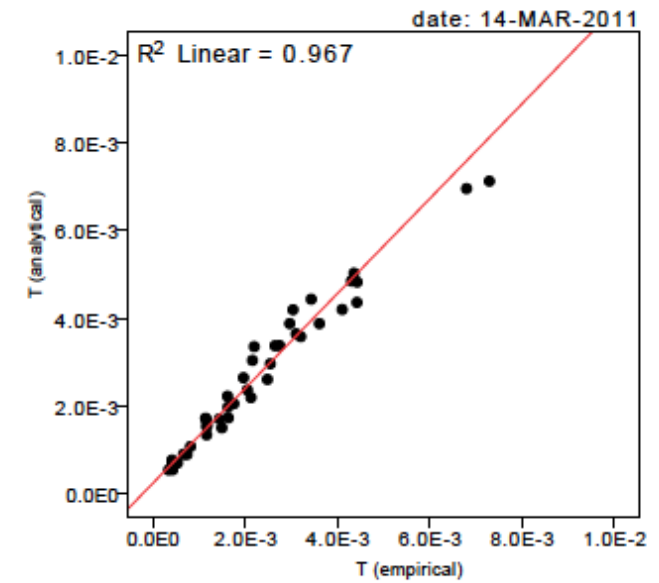
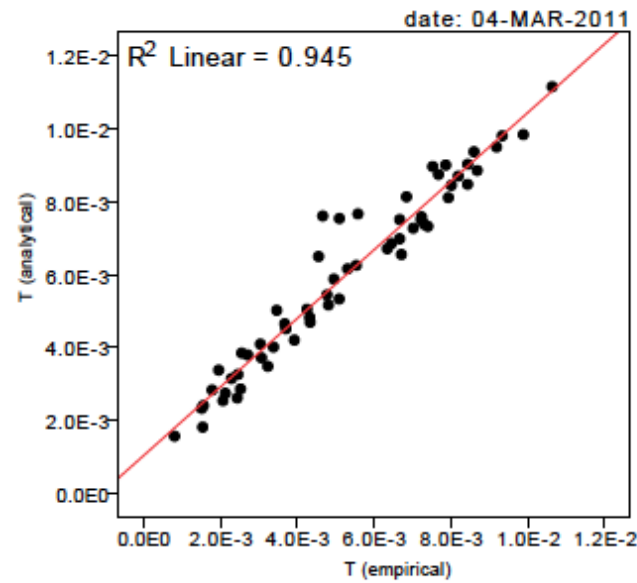
$$k_1 = 1.0 \text{ for } [Z] = mm^6 \cdot m^{-3}$$

$$k_2 = \frac{\pi}{6000} \text{ for } [W] = mm^3 \cdot m^{-3}$$

$$k_3 = \frac{6\pi}{10^4} \text{ for } [R] = mm \cdot h^{-1}$$

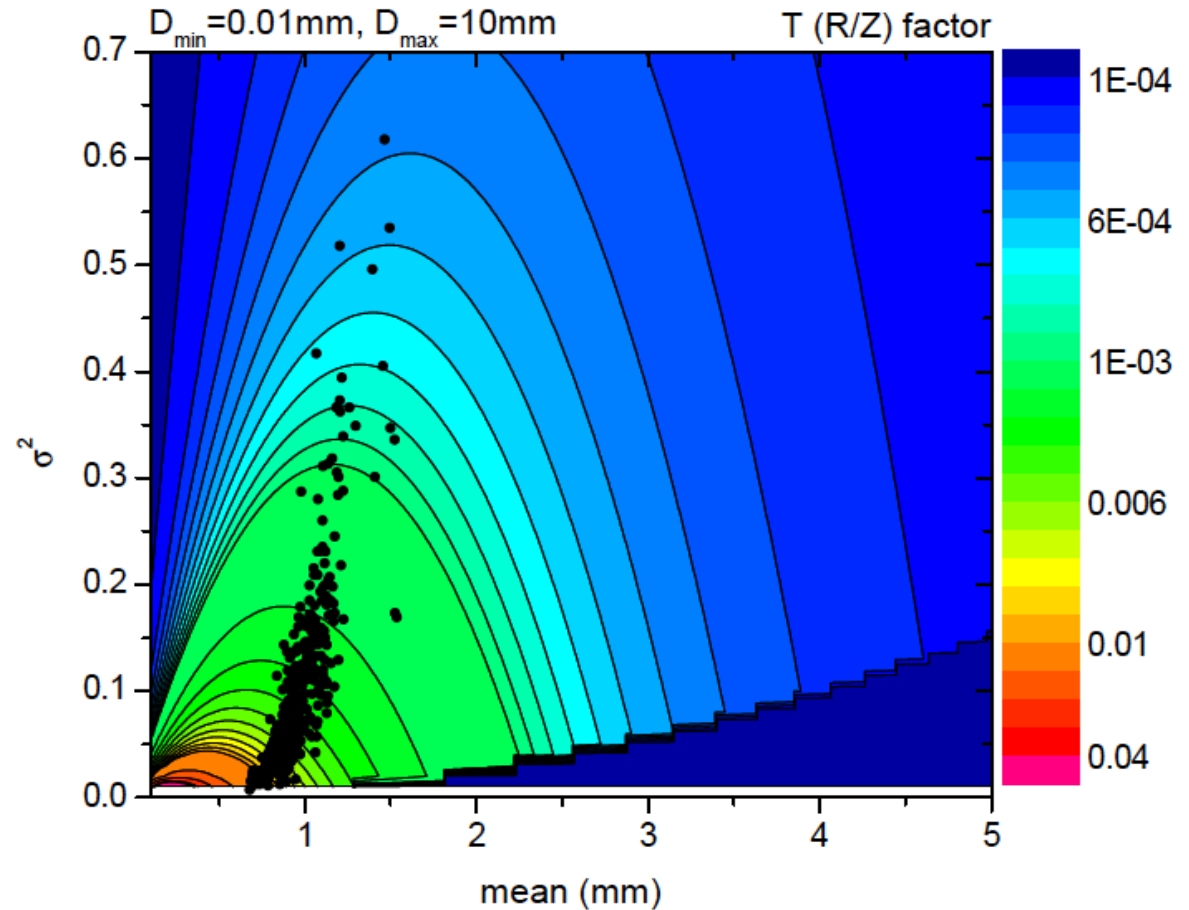
$$v_1 = 3.78, v_2 = 0.67$$

R/Z (empirical)  
vs.  
R/Z (analytical)

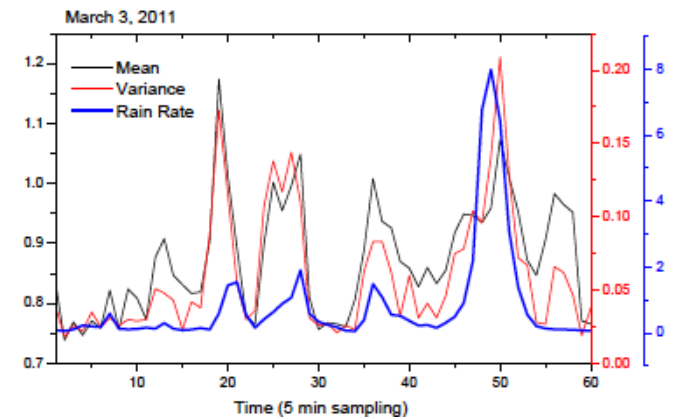
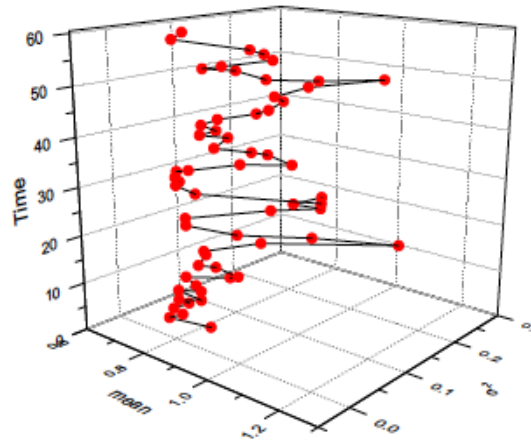
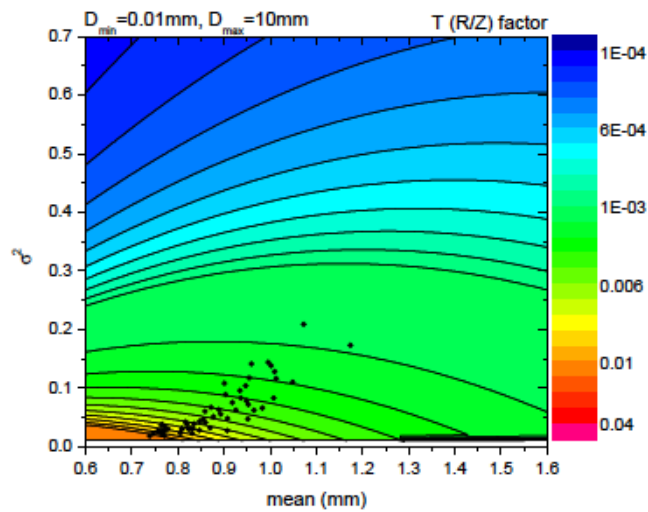


## Real rainfall in $m$ and $\sigma^2$ space

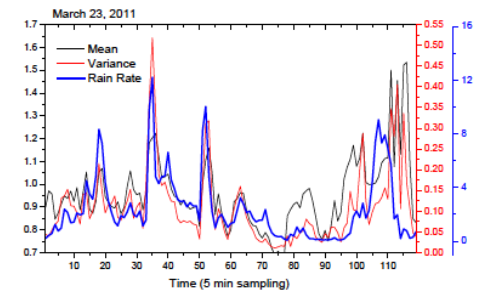
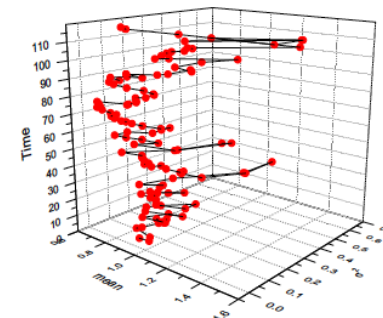
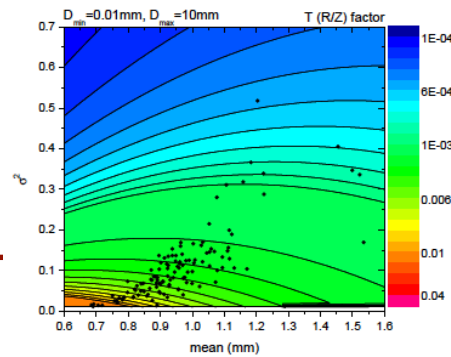
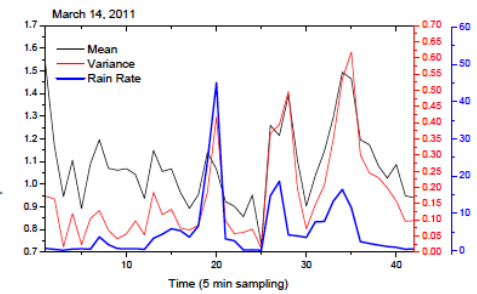
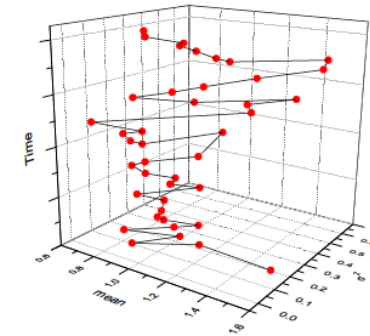
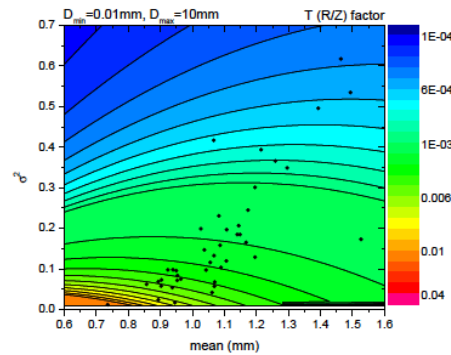
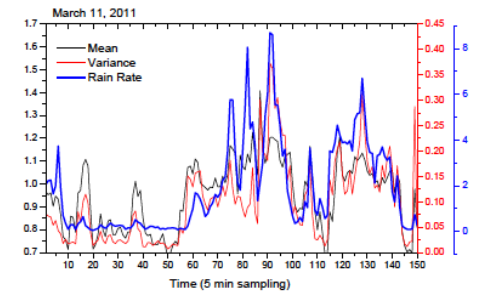
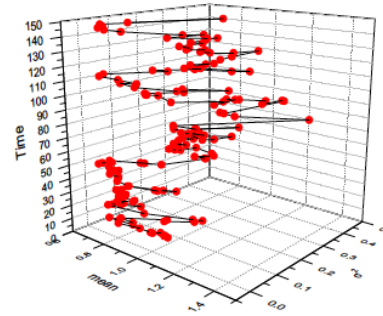
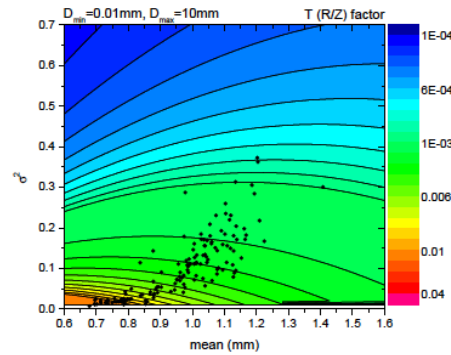
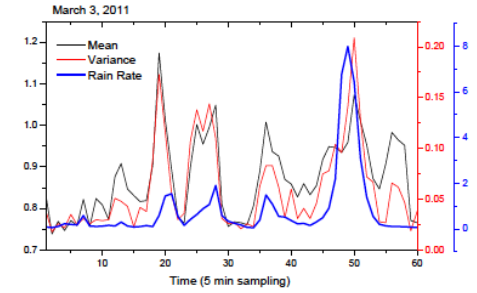
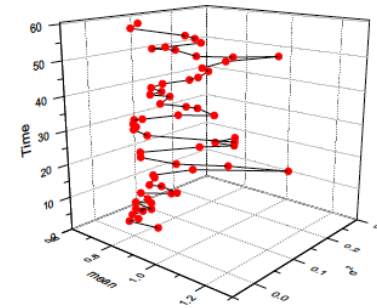
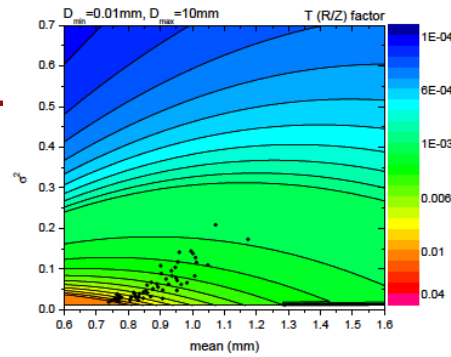
Measured rain  
(dots) is restricted  
to a narrow band:



# Time-evolution of $m$ and $\sigma^2$ (and thus time-evolution of $T=R/Z$ )



# Evolution for 4 episodes





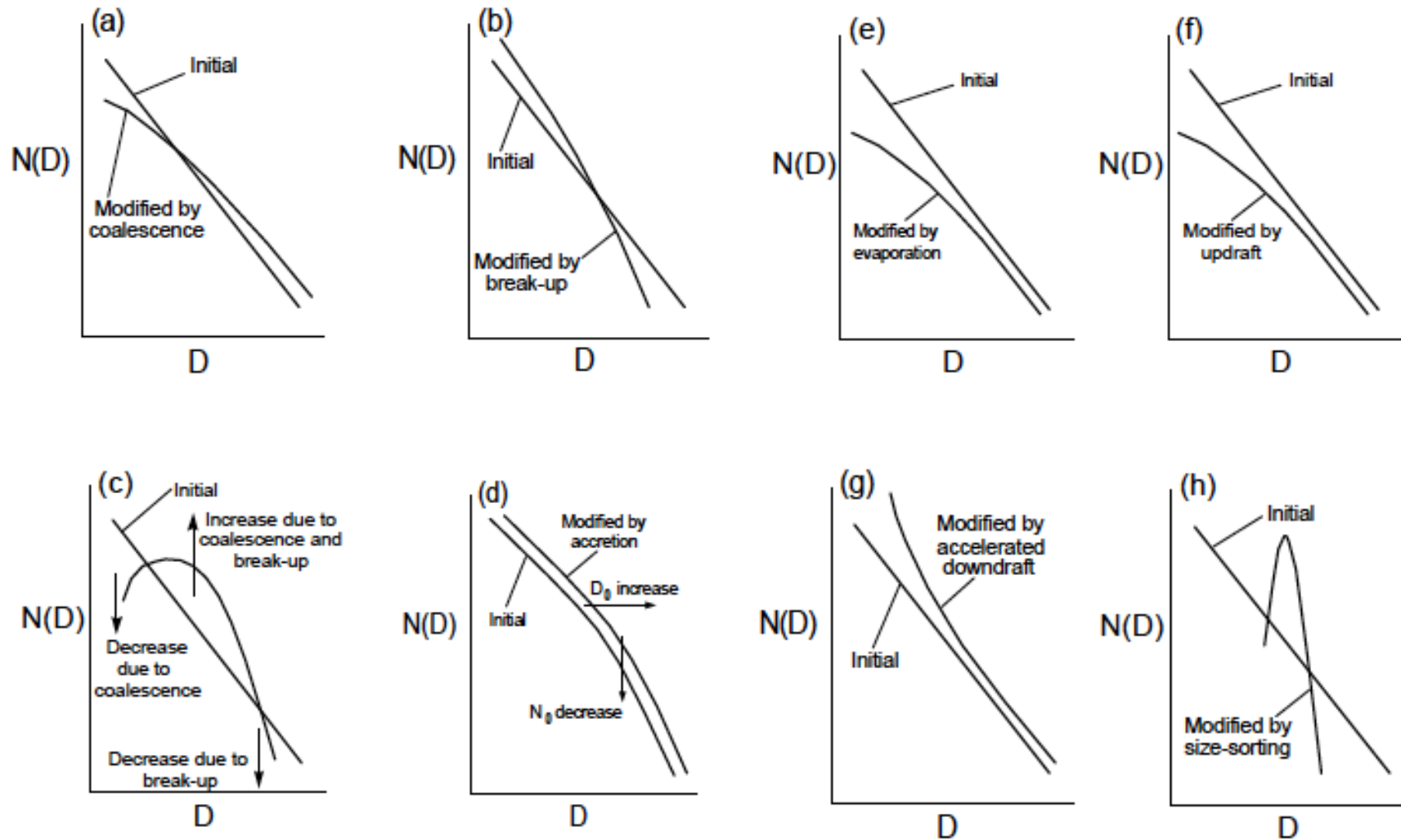
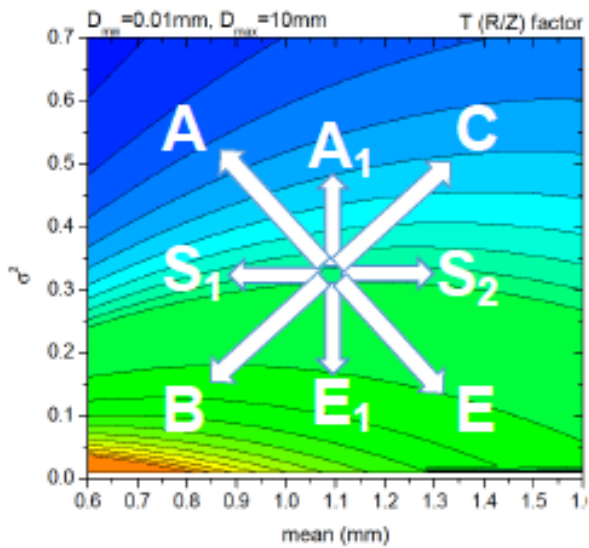


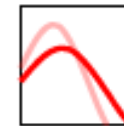
Fig. 3 (a)-(d): Schematic diagrams illustrating the effects on the raindrop size distribution of (a) raindrop coalescence, (b) raindrop break-up, (c) coalescence and break-up acting simultaneously and (d) accretion of cloud droplets.

Fig. 3(e)-(g): Schematic diagrams illustrating the effects on the raindrop size distribution of (e) evaporation, (f) an updraft, (g) an accelerated downdraft, and (h) size-sorting.

# Microphysics

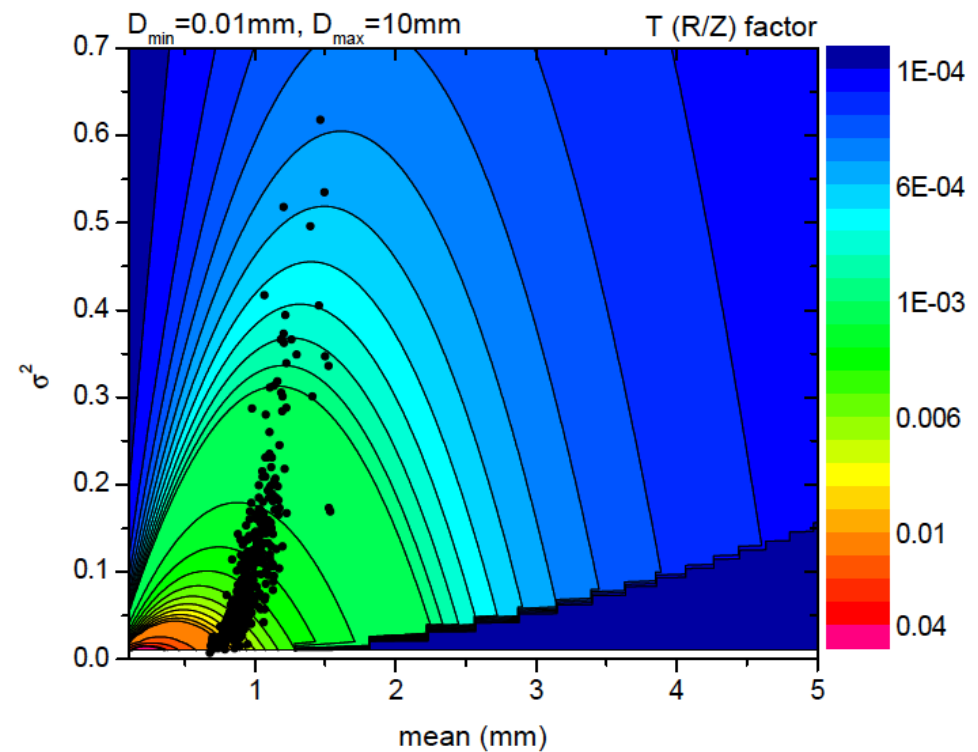
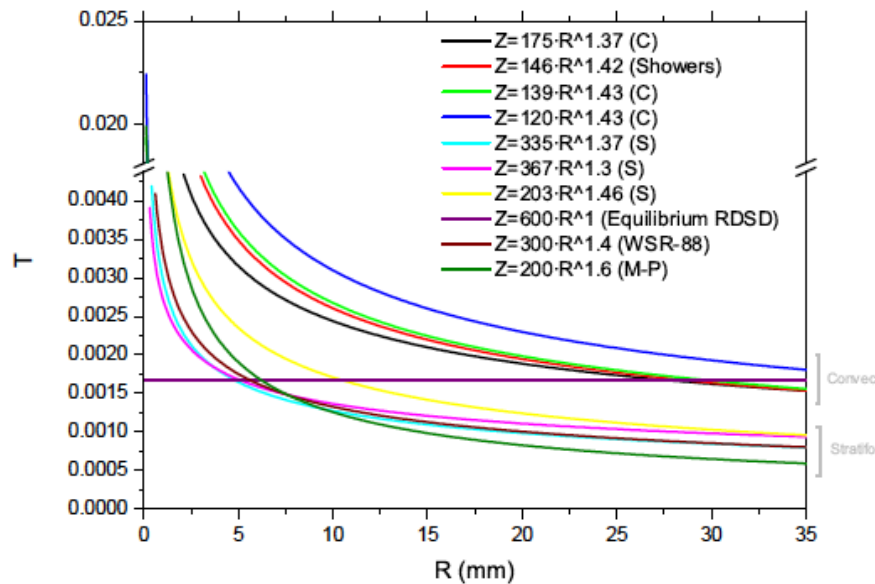


<b>A</b>	$m$ $\sigma^2$	↓ ↑	small drops large drops	↑ ↑		Accretion
<b>B</b>	$m$ $\sigma^2$	↓ ↓	small drops large drops	↑ ↓		Break-up
<b>C</b>	$m$ $\sigma^2$	↑ ↑	small drops large drops	↓ ↑		Coalescence
<b>E</b>	$m$ $\sigma^2$	↑ ↓	small drops large drops	↓ ↓		Evaporation
<b>S<sub>1</sub></b>	$m$ $\sigma^2$	↓ =	small drops large drops	↓ ↓		Size sorting
<b>S<sub>2</sub></b>	$m$ $\sigma^2$	↑ =	small drops large drops	↑ ↑		Size sorting
<b>A<sub>1</sub></b>	$m$ $\sigma^2$	= ↑	small drops large drops	↑ ↑		Accretion
<b>E<sub>1</sub></b>	$m$ $\sigma^2$	= ↓	small drops large drops	↓ ↓		Evaporation

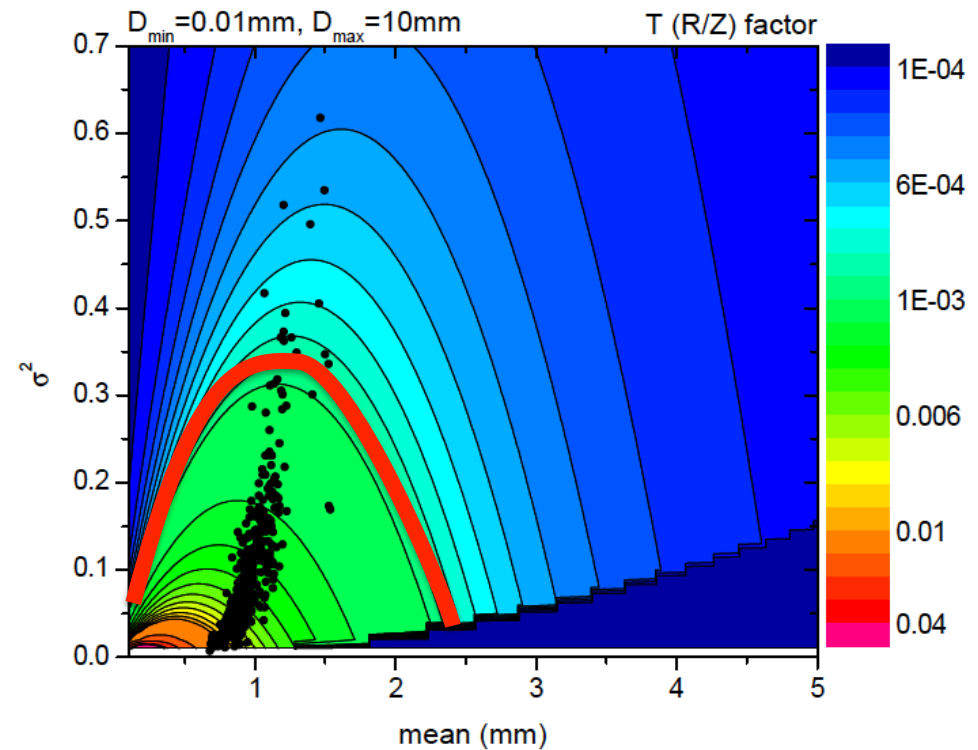
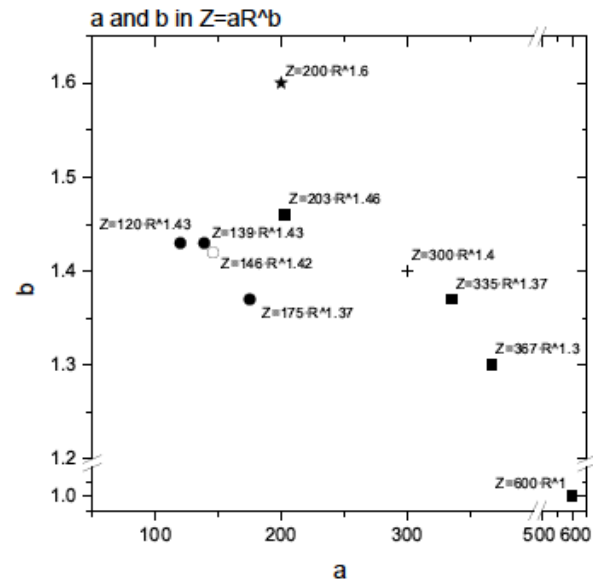
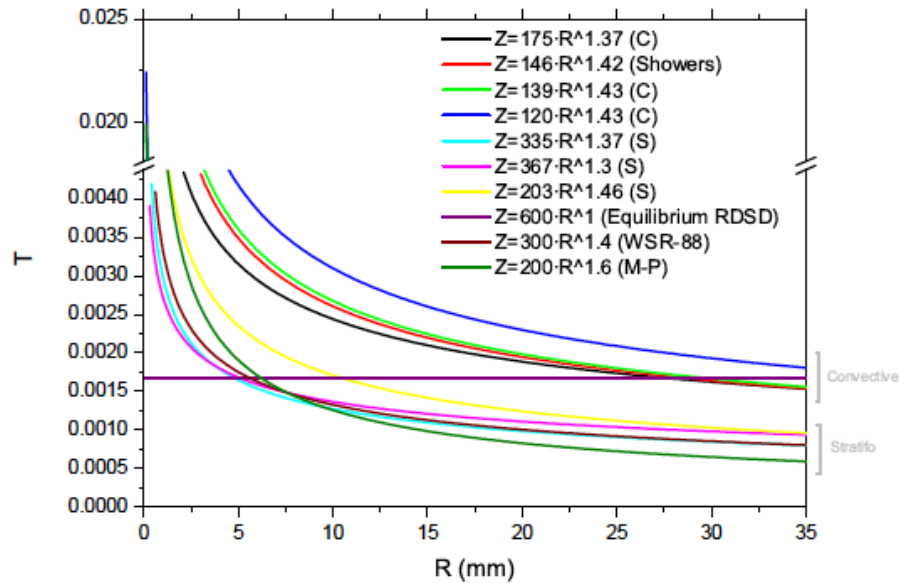


Coalescence

# C/S for several standard Z-R relationships



Microphysically, C/S  
 transients into having or not a  
 large mean diameter  
 (and thus large variance)



But the nicest thing about the probabilistic approach is that you don't have to limit yourself to the gamma distribution or to two parameters

In probability theory,  
there is a general PDF form:

$$p(D) = \exp\left(-\sum_{i=0}^n \lambda_i g_i(D)\right)$$

$$p(D) = \exp\left(-\sum_{i=0}^n \lambda_i g_i(D)\right)$$

comes from

Maximize this function:  $S = -k \int p(D) \log(p(D)) dD$

Subject to:  $E\{g_i(D)\} = \int p(D) g_i(D) dD \quad i = 1, \dots, n$

## Thus for instance

Maximize this function:  $S = -k \int p(D) \log(p(D)) dD$

$$\text{Subject to: } \left\{ \begin{array}{l} \int p(D) D dD = E\{D\} \\ \int p(D) \log(D) dD = E\{\log(D)\} \end{array} \right.$$

$$\text{Yields: } p(D) = D^\mu \Lambda^{\mu+1} \frac{e^{-\Lambda D}}{\Gamma(\mu + 1)}$$



# Therein, the gamma distribution is the PDF arising if you assume that:

1. Drop diameters are monotonically growing values, i.e.

$$\int p(D)DdD = E\{D\}$$

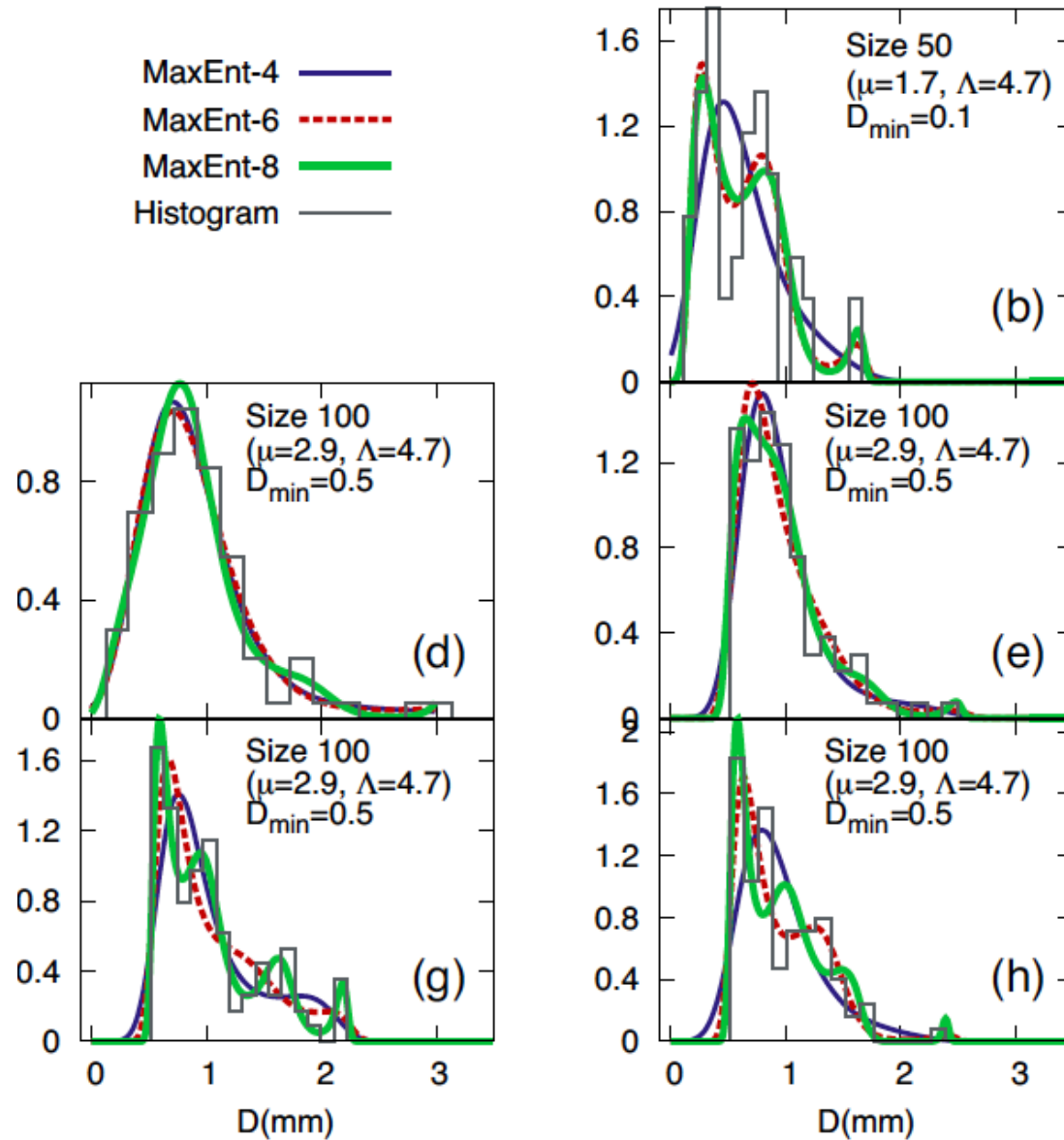
2. Drop diameters follow a scaling law, i.e.

$$\int p(D)\log(D)dD = E\{\log(D)\}$$

# In terms of stochastic processes...

If we hypothesize that rain is a stochastic, Poisson process, theory dictates that the waiting time for a drop of a given diameter to be found follows a Gamma distribution

# Some other choices



# Summary

Modeling the RDSD with 3  
fundamental variables:

$$N_T, \{m, \sigma^2\}$$

Z/R depends on:

$$m, \sigma^2$$

Microphysics can be transparently  
described in terms of:

$$m, \sigma^2$$

# Thanks

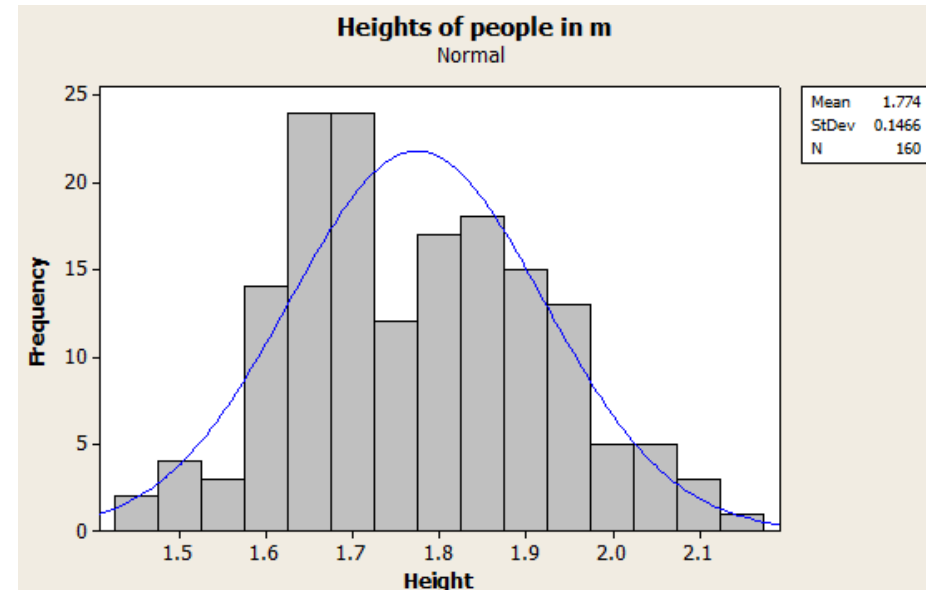
Rain Drop Size Distribution
$n(D) = N_T D^\mu \Lambda^{\mu+1} \frac{e^{-\Lambda D}}{\Gamma(\mu+1)}$
Independent variables $N_T, m, \sigma^2$
$\Lambda = \frac{m}{\sigma^2}$
$\mu = \frac{m^2 - \sigma^2}{\sigma^2}$
Parameters
$k_1 = 1.0$ for $[Z] = mm^6 \cdot m^{-3}$
$k_2 = \frac{\pi}{6000}$ for $[W] = mm^3 \cdot m^{-3}$
$k_3 = \frac{6\pi}{10^4}$ for $[R] = mm \cdot h^{-1}$
$D_{min}$ diameter of the smallest measurable drop, in mm
$D_{max}$ diameter of the largest measurable drop, in mm
$v_1 = 3.78, v_2 = 0.67$ (Atlas and Ulbrich 2000)
Main moments (depend on $N_t, m$ and $\sigma^2$ )
$Z = k_1 N_T \Lambda^{-6} \frac{\gamma(\mu+7, D_{max}\Lambda) - \gamma(\mu+7, D_{min}\Lambda)}{\Gamma(\mu+1)}$
$W = k_2 N_T \Lambda^{-3} \frac{\gamma(\mu+4, D_{max}\Lambda) - \gamma(\mu+4, D_{min}\Lambda)}{\Gamma(\mu+1)}$
$R = k_3 v_1 N_T \Lambda^{-(3+v_2)} \frac{\gamma(\mu+4+v_2, D_{max}\Lambda) - \gamma(\mu+4+v_2, D_{min}\Lambda)}{\Gamma(\mu+1)}$
$D_m = \Lambda^{-1} \frac{\gamma(\mu+5, D_{max}\Lambda) - \gamma(\mu+5, D_{min}\Lambda)}{\gamma(\mu+4, D_{max}\Lambda) - \gamma(\mu+4, D_{min}\Lambda)}$ , does not depend on $N_T$
Z-R relationship (depend on $m$ and $\sigma^2$ )
$R = T \cdot Z$
$T \equiv \frac{k_3}{k_1} \cdot v_1 \cdot \Lambda^{(3-v_2)} \frac{\gamma(\mu+4+v_2, D_{max}\Lambda) - \gamma(\mu+4+v_2, D_{min}\Lambda)}{\gamma(\mu+7, D_{max}\Lambda) - \gamma(\mu+7, D_{min}\Lambda)}$
Scaled intercept parameter $N_w$ (depends on $N_t, m$ and $\sigma^2$ )
$N_w = \frac{4^4 k_2}{\pi \rho_w} \cdot N_T \cdot \Lambda \frac{[\gamma(\mu+4+v_2, D_{max}\Lambda) - \gamma(\mu+4+v_2, D_{min}\Lambda)]^b}{\Gamma(\mu+1) [\gamma(\mu+5, D_{max}\Lambda) - \gamma(\mu+5, D_{min}\Lambda)]^4}$

Tapiador, FJ, Haddad, ZS and Turk, FJ, 2013. *A Probabilistic View on the RSD* (submitted to J. Hydromet.)

# Probability vs. Statistics



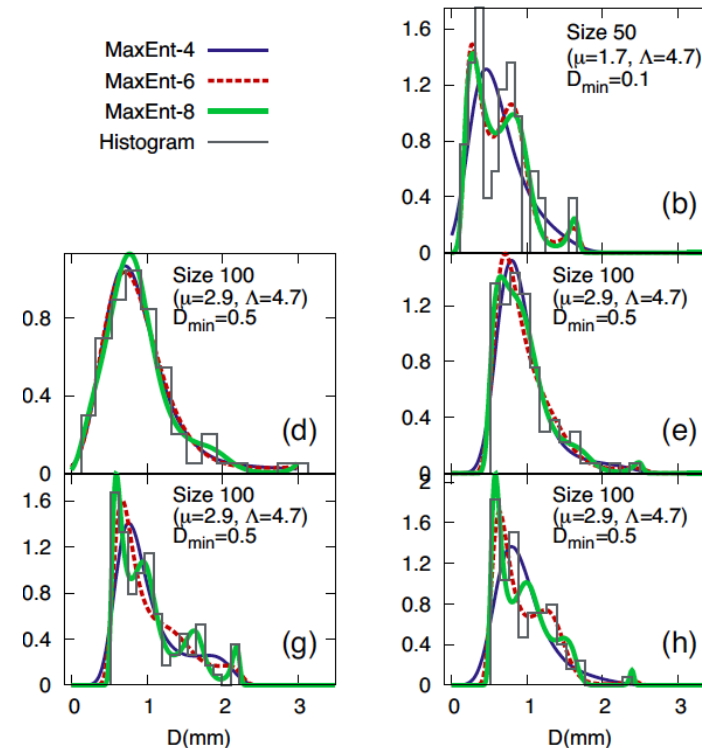
We don't need to carry out any measurement to know that the chances of drawing a King are  $4/52=1/13$



We cannot know the average size of people without performing measurements on a representative sample of the population.

[a large number of draws confirms this]

# Probability vs. Statistics



We **hypothesize** that rain is a stochastic, Poisson process. Then, by crushing numbers we get that the waiting time for a drop of a given diameter to be found follows a Gamma distribution

We **measure** drops and **then** we fit the histogram to a curve. We chose a Gamma distribution because it minimizes the RMSE.

But that creates fitting artifacts