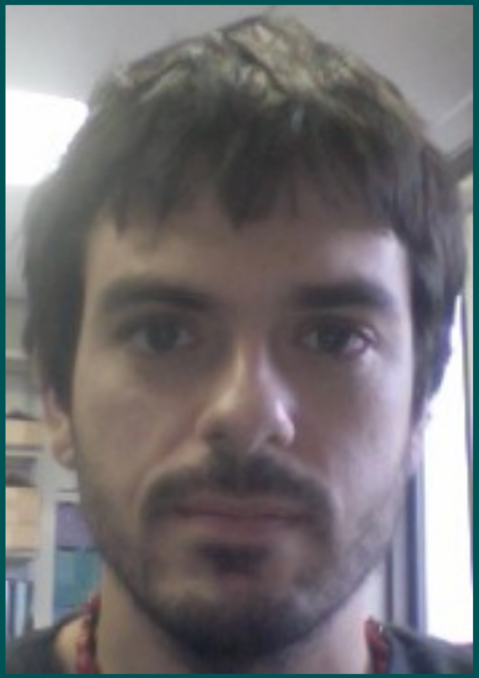
Ensemble Retrievals

Isztar Zawadzki and Bernat Puigdomeneq-Tresserras
McGill University



Background

When radar measurements are interpreted quantitatively we always use a model microphysics of the backscattering targets.

Model microphysics is an approximation to a very complex reality

There is never sufficient information in radar measurement to resolve unambiguously the complexity of microphysics.

Thus, whatever model we use there will be a good deal of uncertainty in the quantitative interpretation of radar measurements.

Radar measurements are not proper measurements without a good assessment of their uncertainties.

In radar data assimilation this is particularly fundamental.

And all this applies to satellite measurements as well

What do we call a Retrieval

Inference of atmospheric parameters from observations related but not equal to the parameters of interest.

Simplest example: retrieving LWC, M from Z

What do we call a Retrieval

Inference of atmospheric parameters from observations related but not equal to the parameters of interest.

Simplest example: retrieving LWC, *M* from *Z*

The correspondence between observations and retrieved parameters is established through physical equations. These equations are affected by uncertainties due to the natural variability of the parameters in these equations. Thus, these are approximate "models" of the physical processes.

What do we call a Retrieval

Inference of atmospheric parameters from observations related but not equal to the parameters of interest.

Simplest example: retrieving LWC, *M* from *Z*

The correspondence between observations and retrieved parameters is established through physical equations. These equations are affected by uncertainties due to the natural variability of the parameters in these equations. Thus, these are approximate "models" of the physical processes.

Simplest example: $M = \frac{\pi}{6}\rho_w \int_0^\infty D^3 N(D)dD$; $Z = \int_0^\infty D^6 N(D)dD$

$N(D)$ has a good deal of natural variability

Ensemble Retrievals

Each combination of the possible values of the parameters leads to one set of physical relationships and to one corresponding retrieval.

Simplest example:

$$N(D) = N_0 D^{\mu} \exp(-\lambda D)$$

for a given set of possible values of

N_0, μ, λ

we get one retrieval of *M* from *Z*

Ensemble Retrievals

Each combination of the possible values of the parameters leads to one set of physical relationships and to one corresponding retrieval.

Simplest example:

$$N(D) = N_0 D^{\mu} \exp(-\lambda D)$$

for a given set of possible values of

N_0, μ, λ

we get one retrieval of *M* from *Z*

By sweeping through all the possible values of the parameters we obtain the **ensemble** of possible models leading to the **ensemble** of retrievals. This **ensemble** contains all the uncertainty of the physical model.

Ensemble Retrievals

Each combination of the possible values of the parameters leads to one set of physical relationships and to one corresponding retrieval.

Simplest example:

$$N(D) = N_0 D^{\mu} \exp(-\lambda D)$$

for a given set of possible values of

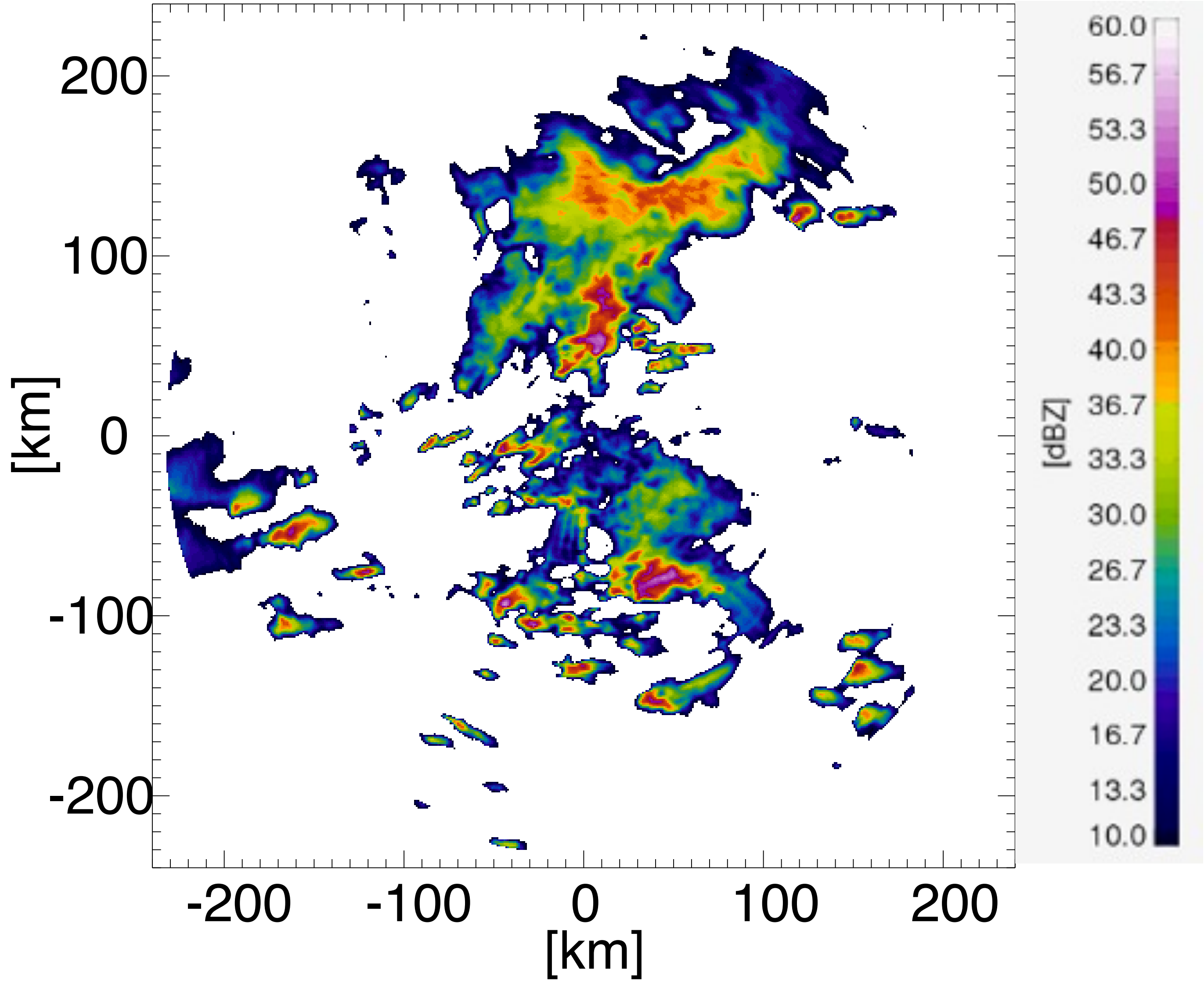
N_0, μ, λ

we get one retrieval of *M* from *Z*

By sweeping through all the possible values of the parameters we obtain the **ensemble** of possible models leading to the **ensemble** of retrievals. This **ensemble** contains all the uncertainty of the physical model.

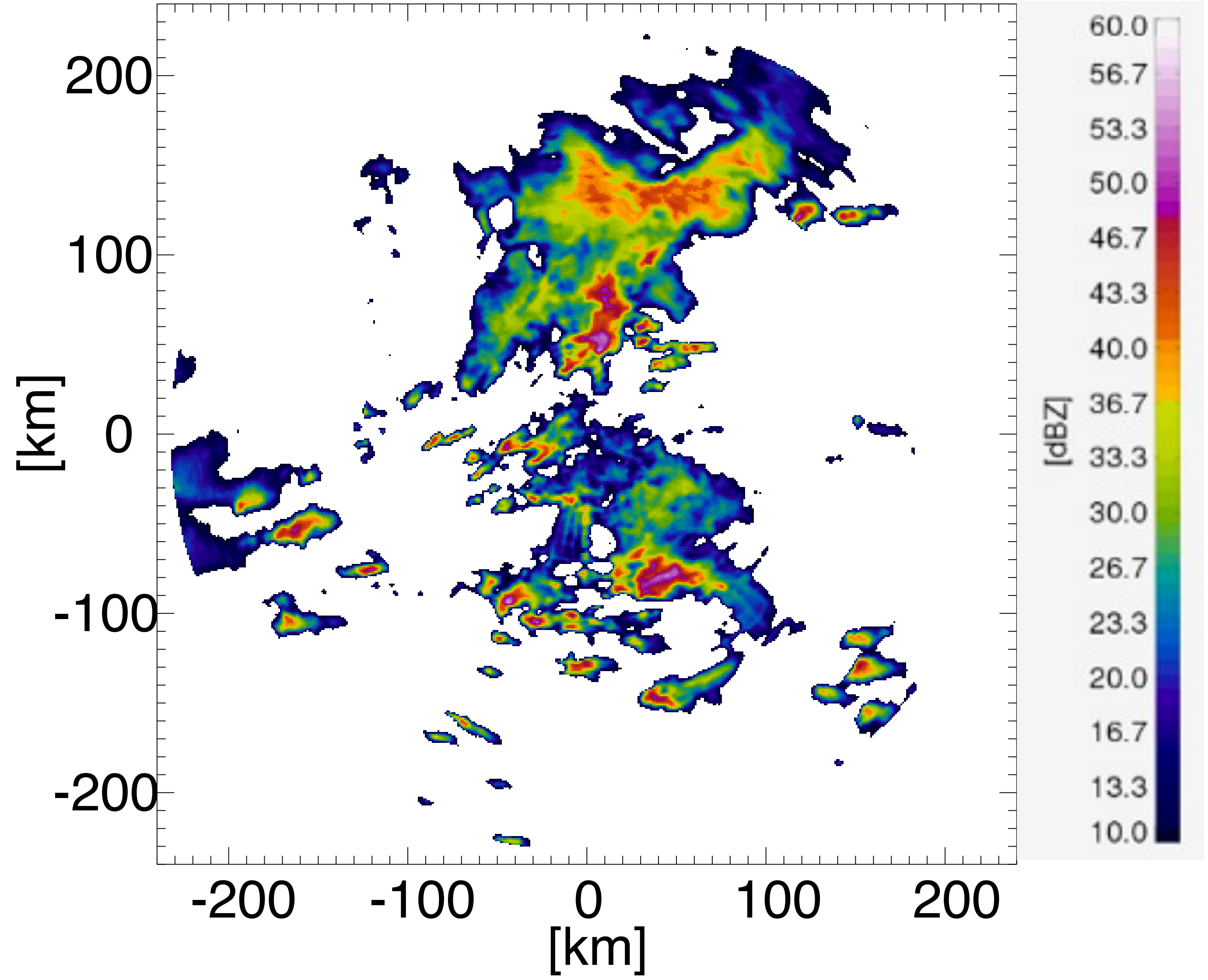
In our simplest example, the Z-M relationship, as an alternative we can use a set of observed DSDs to generate the ensemble.

A particular precipitation event



From our entire disdrometer data base we select the days which the DSDs records had maximum values of dBZ and durations compatible with this particular case, namely a mixture of convection and wide spread rain. We have ~200 days of DSD data of this type of events.

A particular precipitation event

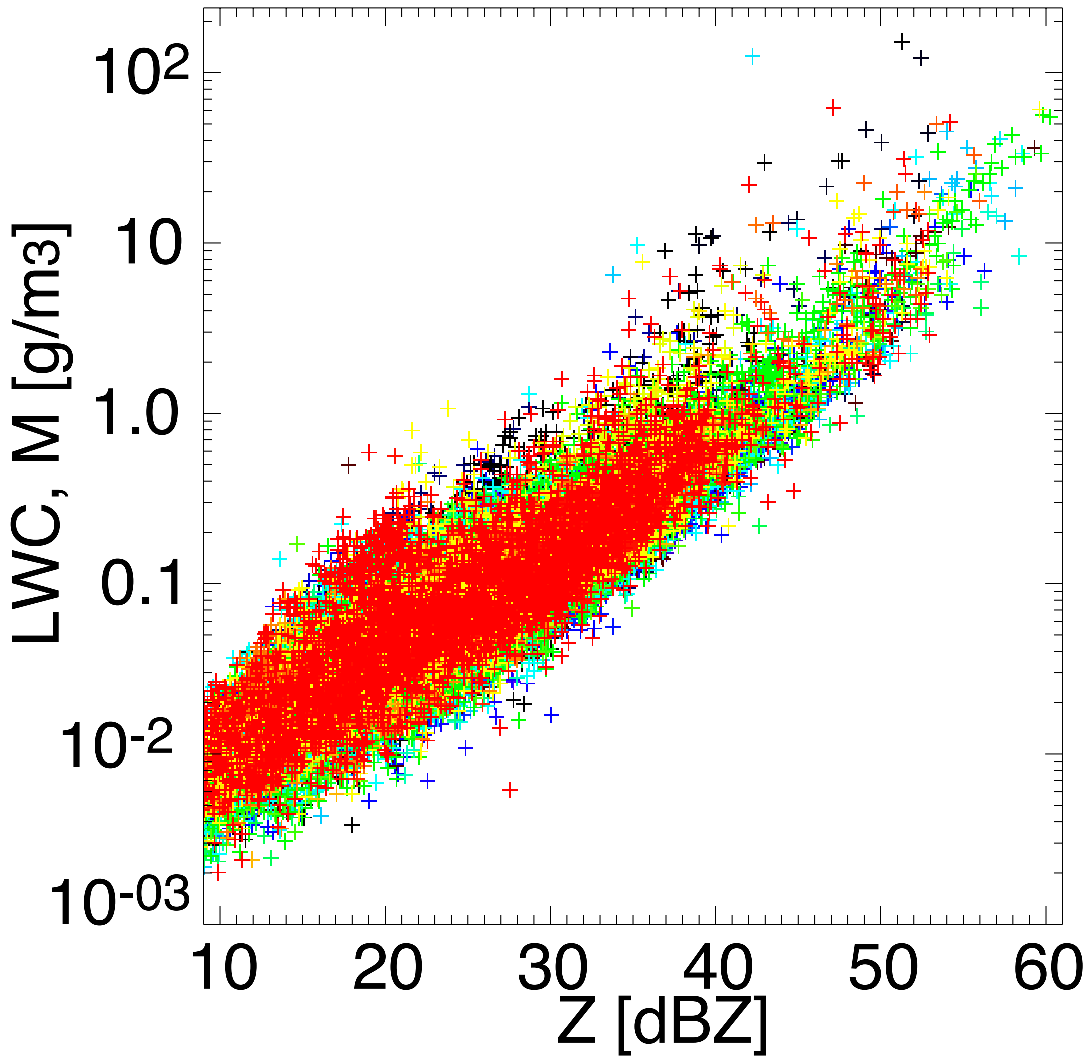


From our entire disdrometer data base we select the days which the DSDs records had maximum values of dBZ and durations compatible with this particular case, namely a mixture of convection and wide spread rain. We have ~200 days of DSD data of this type of events.

We treat the 200 days as equiprobable realizations of the same process

Data Base

From the 200 days of distrometric records we have this distribution of logM versus logZ



Typically an average Z-M relationship, obtained by some kind of regression, is used in retrieving M from Z.

Let us apply ensemble retrieval to this problem.

Example

Here we formulate the retrieval of M as

$$M = \int m \; p(m \mid [Z \pm \delta]) \, dm$$

Let us take $\delta = 0.5 \text{ dB}$

and for each dBZ interval determine $p(m \mid Z)$

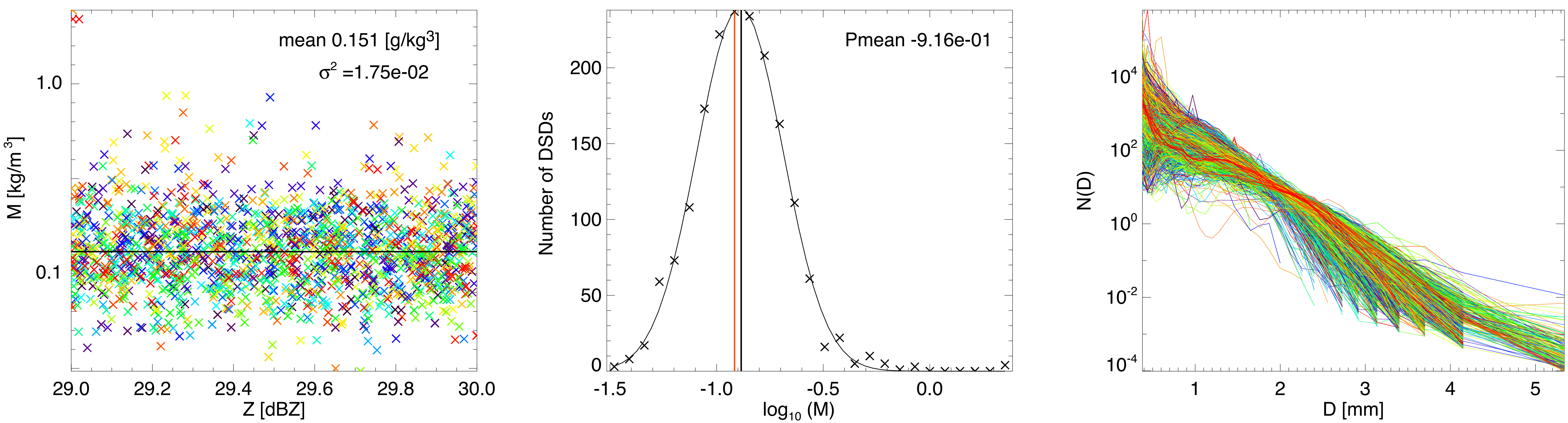
Example

Here we formulate the retrieval of M as

$$M = \int m \; p(m \,|\, [Z \pm \delta]) \, dm$$

Let us take $\delta = 0.5 \text{ dB}$

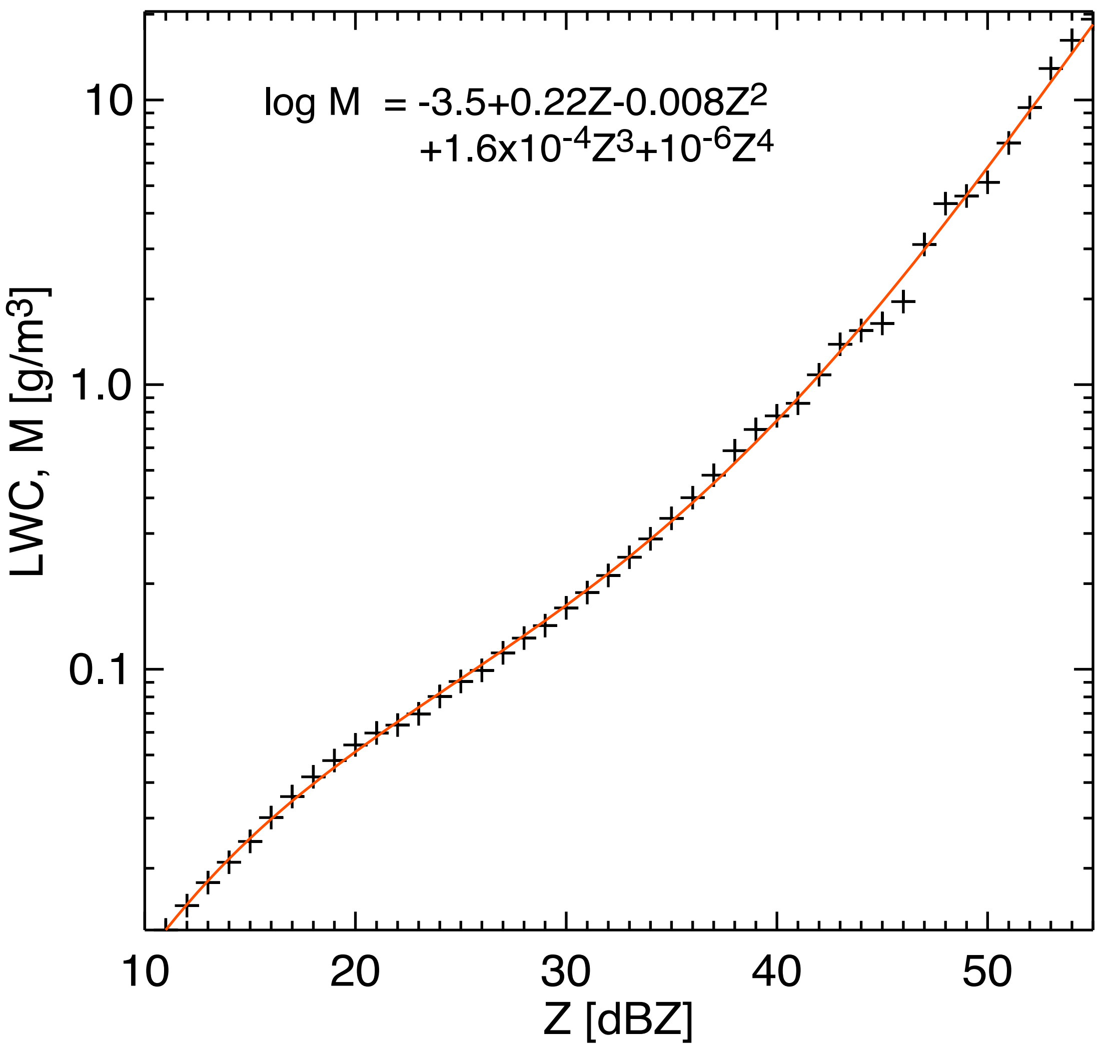
and for each dBZ interval determine $p(m \,|\, Z)$



$$M = 0.15 = \int m \; p(m \,|\, [29 - 30\, dBZ]) \, dm$$

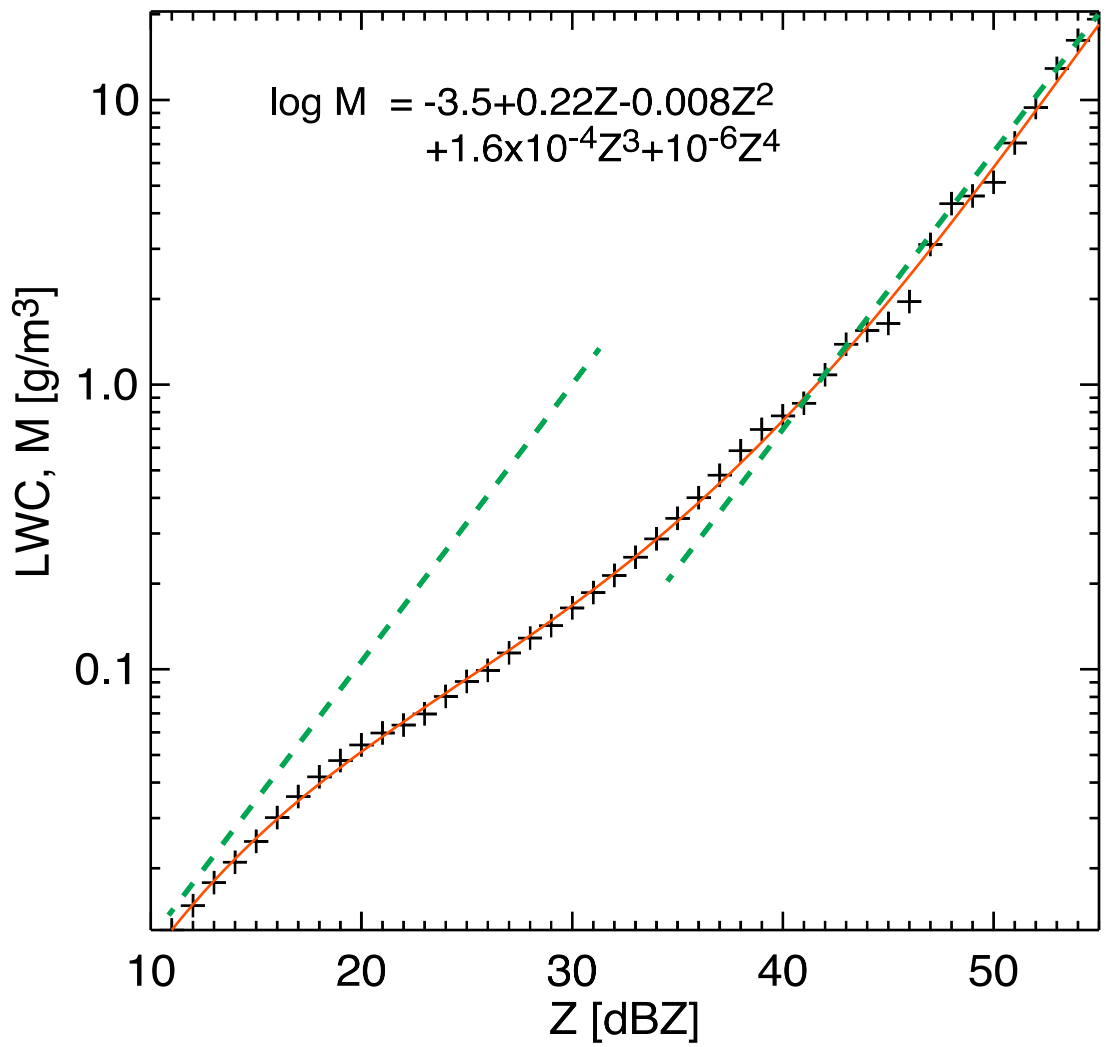
Expected Value of M

From our sample of 200 days of disdrometric records we have this expected value of logM versus logZ



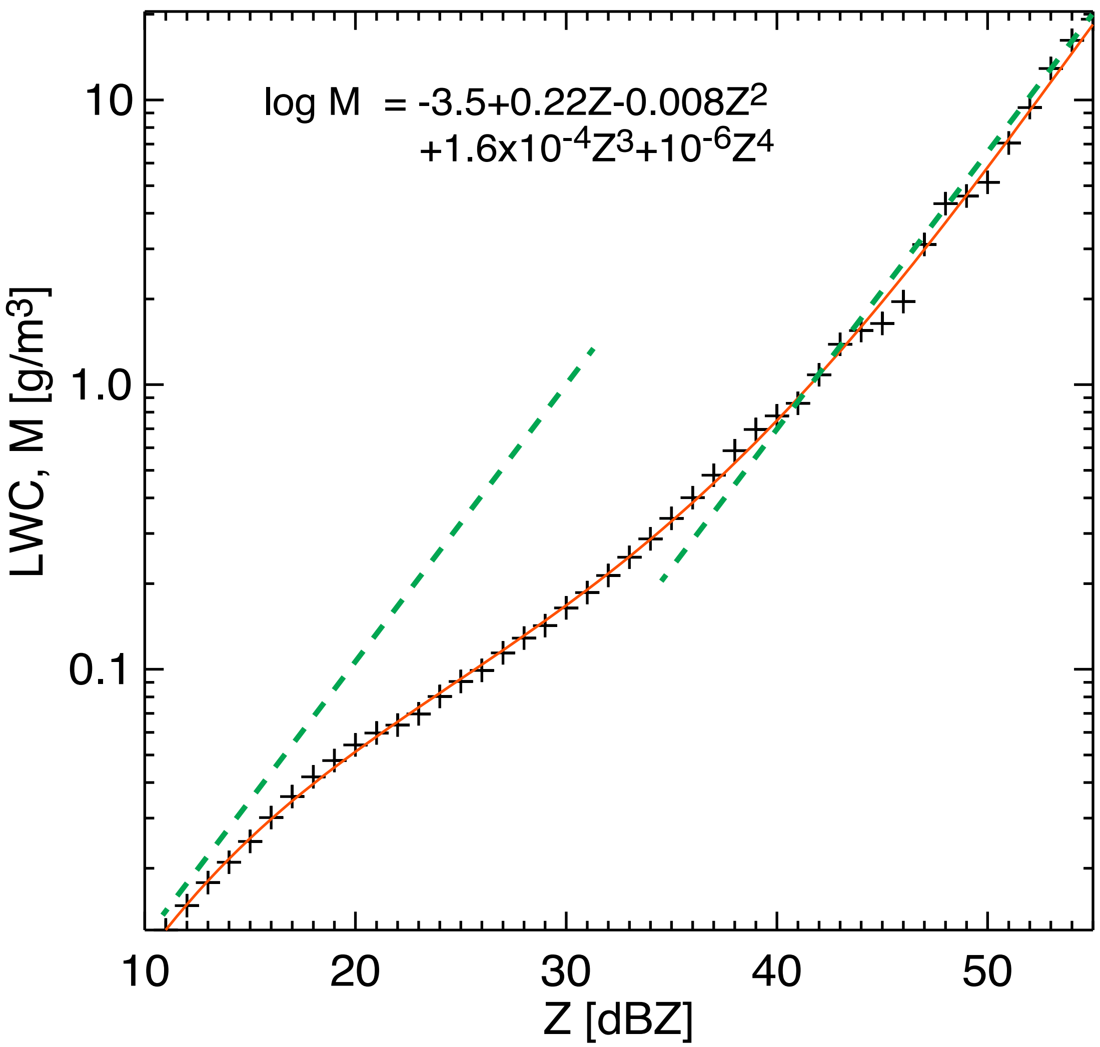
Expected Value of M

From our sample of 200 days of disdrometric records we have this expected value of logM versus logZ



Expected Value of M

From our sample of 200 days of disdrometric records we have this expected value of logM versus logZ

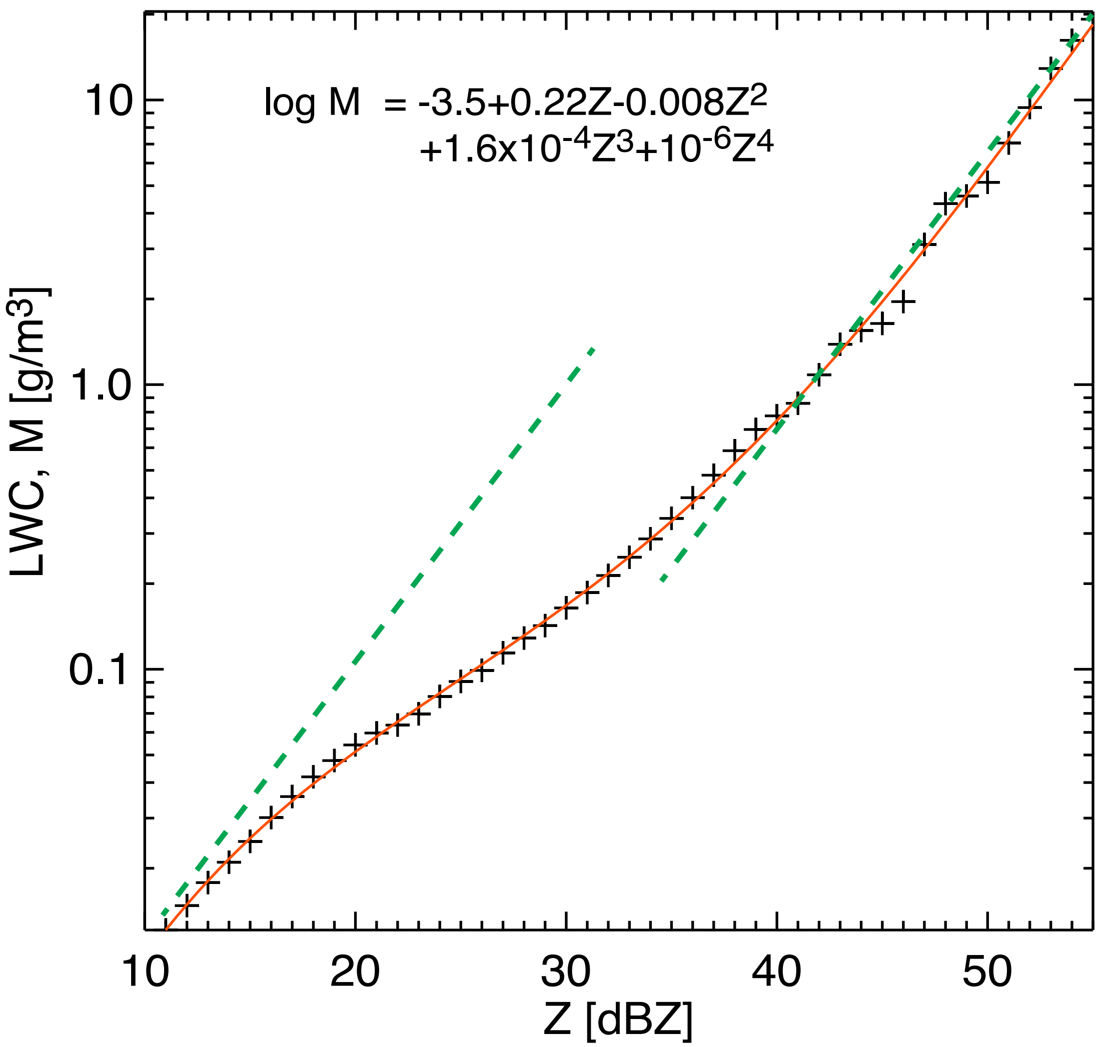


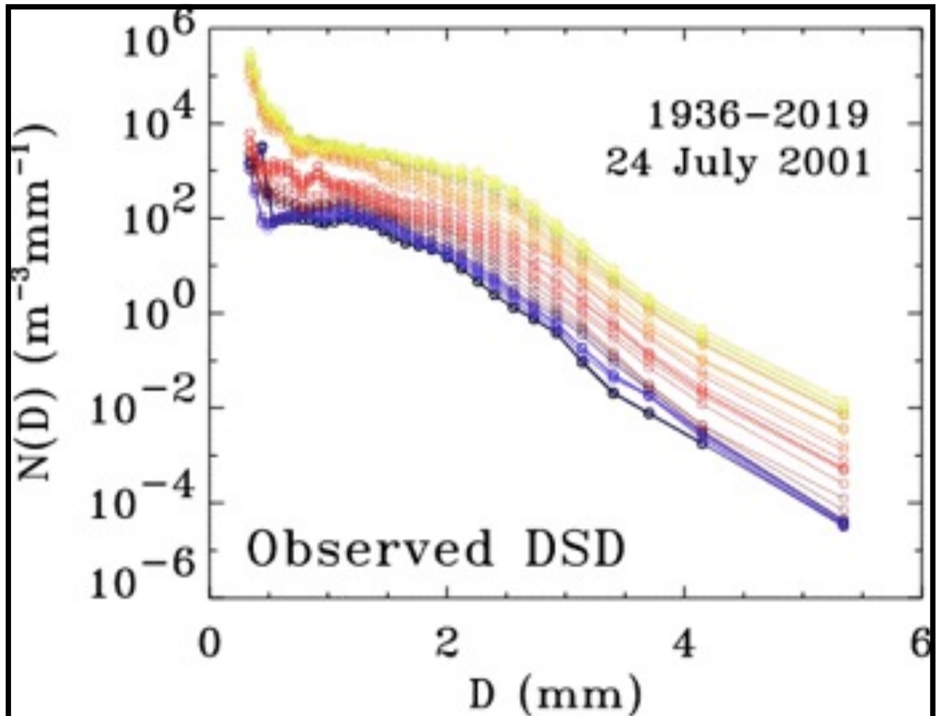
Some small progress here!

We get a more complex relationship that in fact has some physical sense: it is consistent with the tendency to equilibrium DSDs at Z>40dBZ and the expected behaviour at very low rates where cloud collection is the prevailing mechanism of precipitation growth

Expected Value of M

From our sample of 200 days of disdrometric records we have this expected value of logM versus logZ

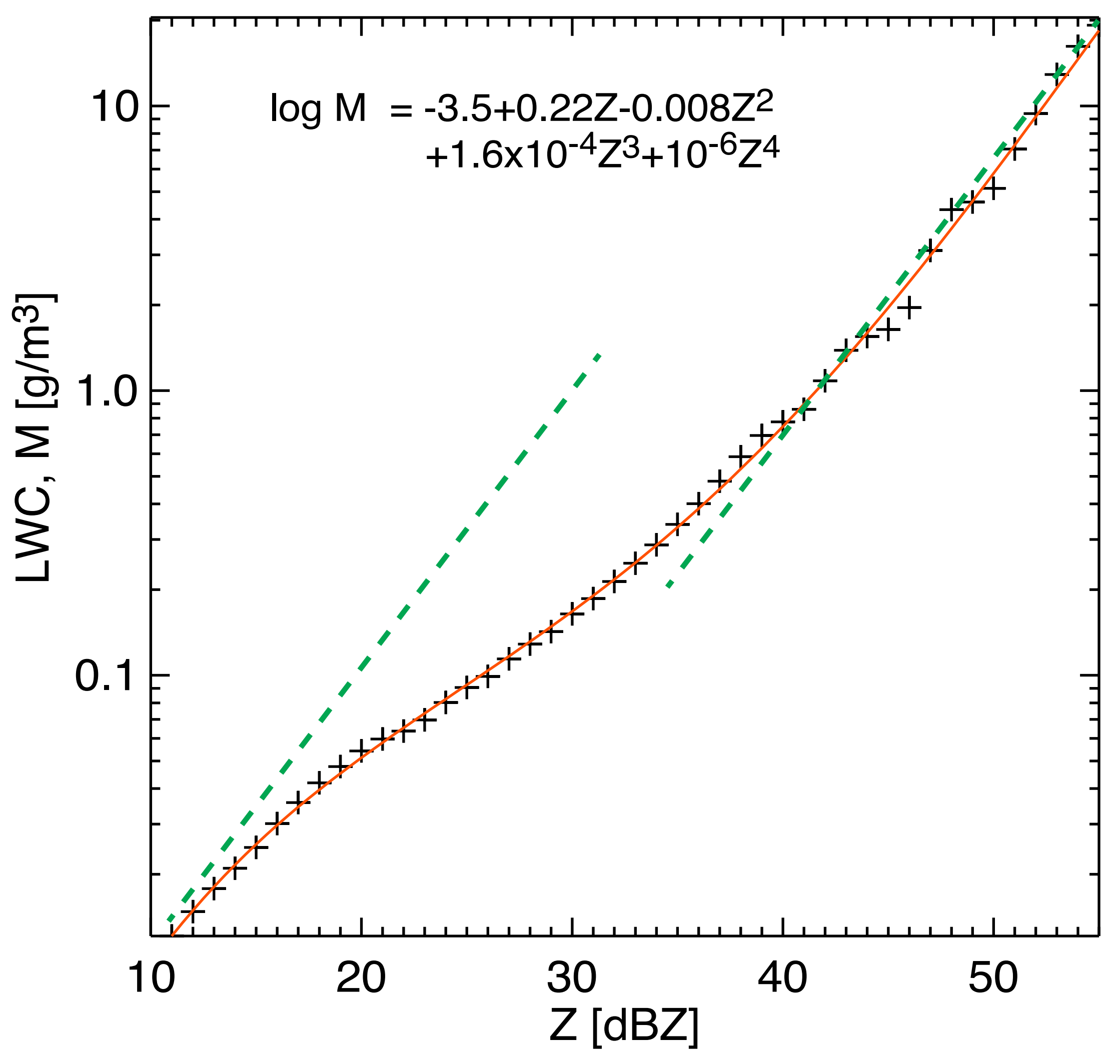


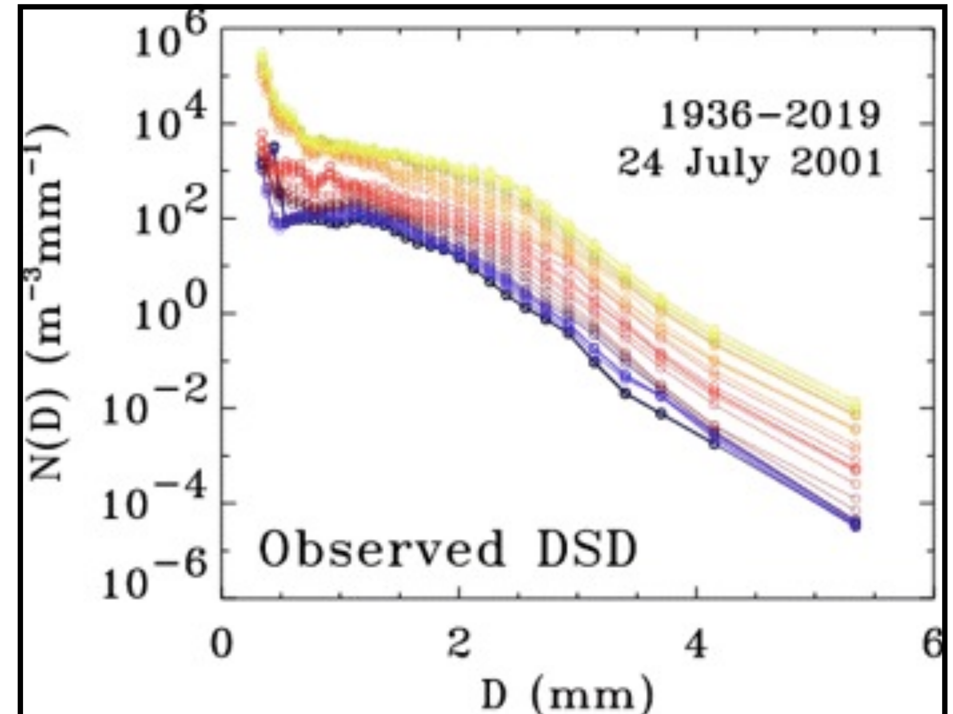


It is consistent with the tendency to equilibrium DSDs at Z>40dBZ and the expected behaviour at very low rates where cloud collection is the prevailing mechanism of precipitation growth

Expected Value of M

From our sample of 200 days of disdrometric records we have this expected value of logM versus logZ





ten

at Z

be

wh

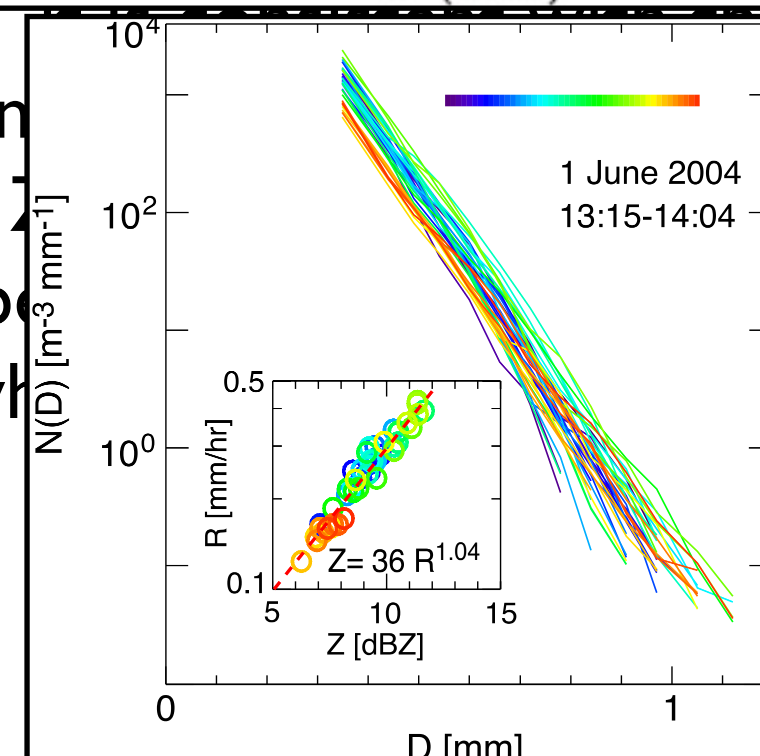
SDs

cted

tes

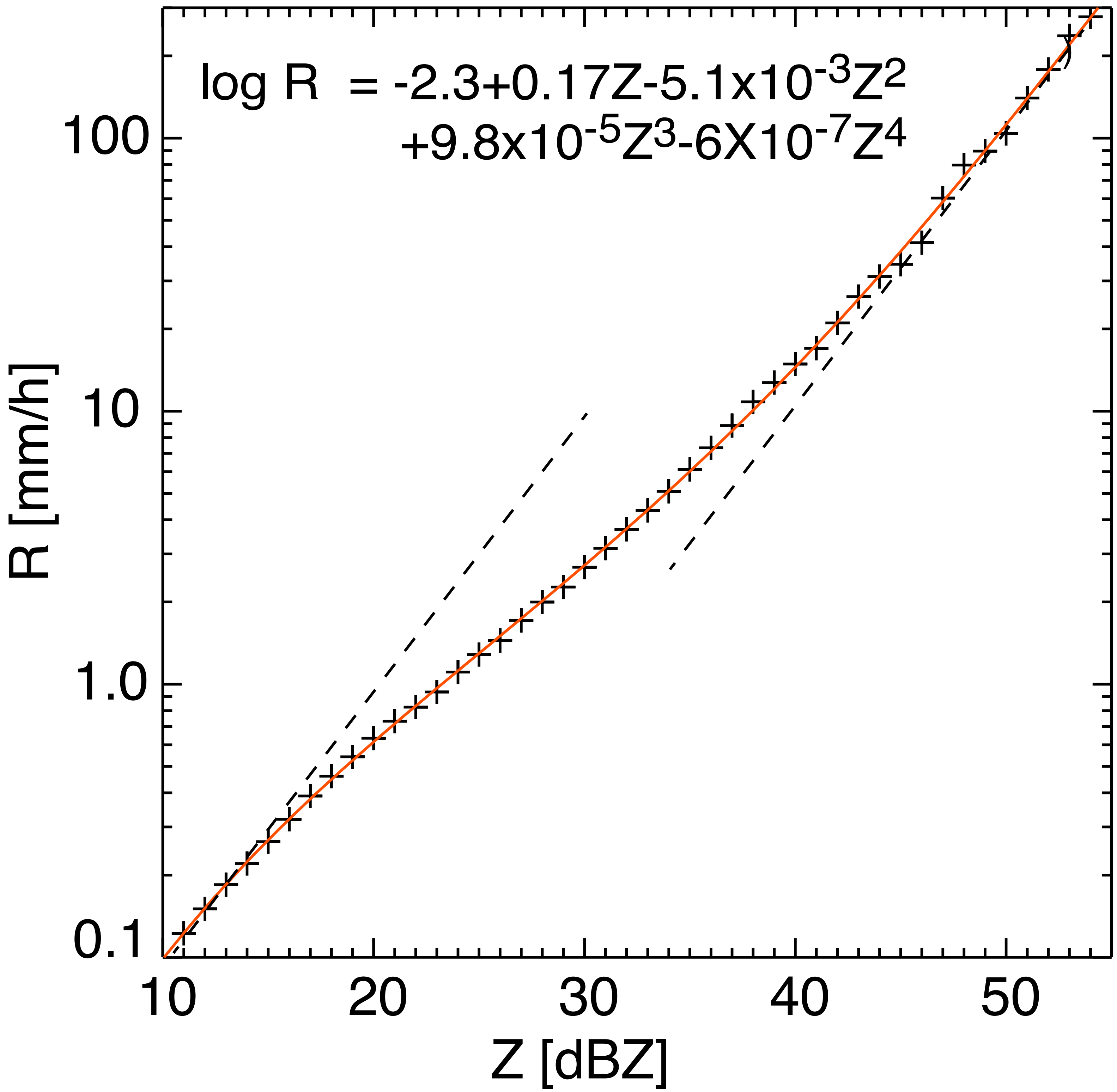
the

f

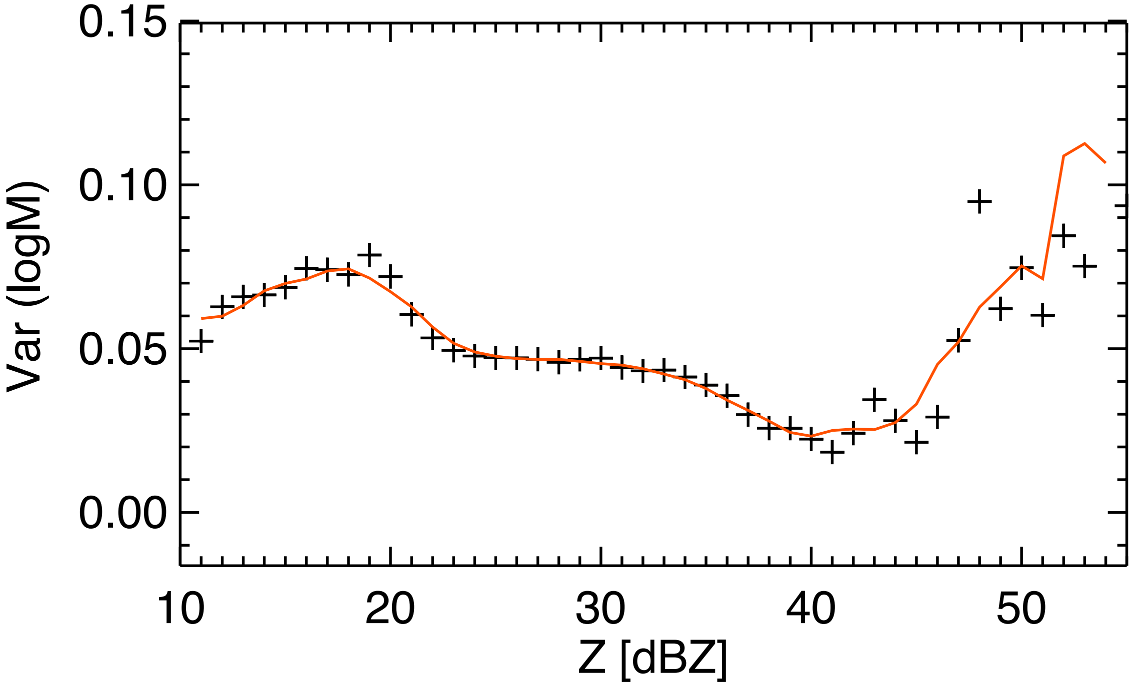


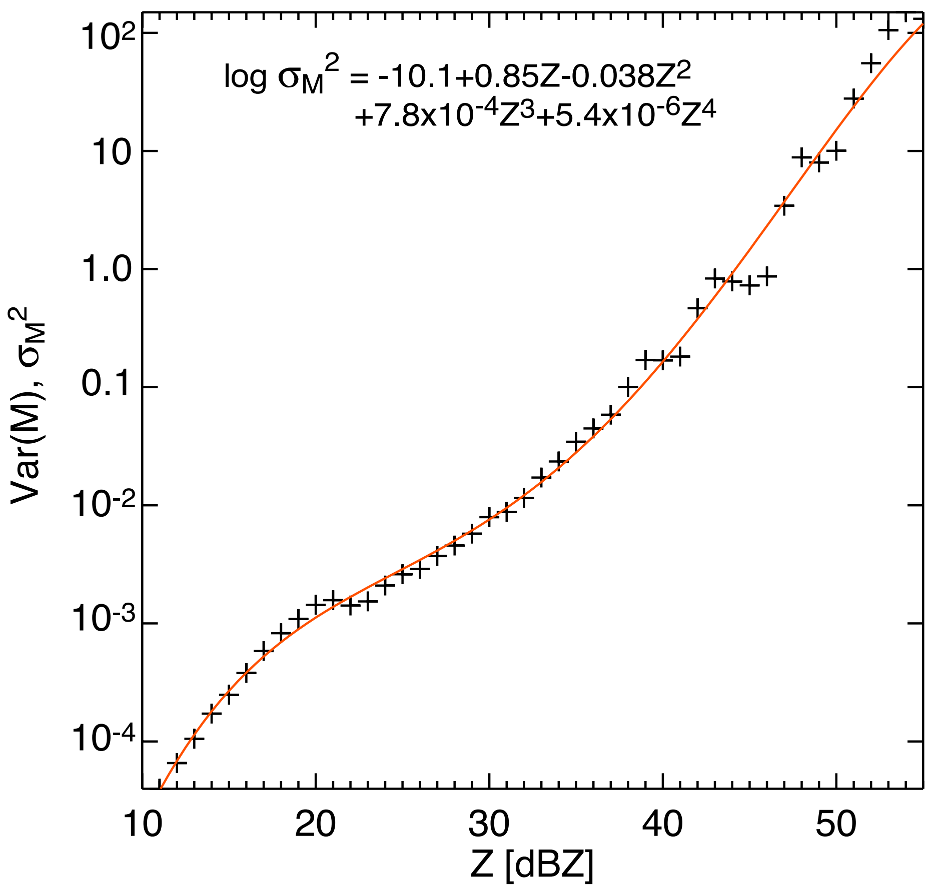
Expected Value of R

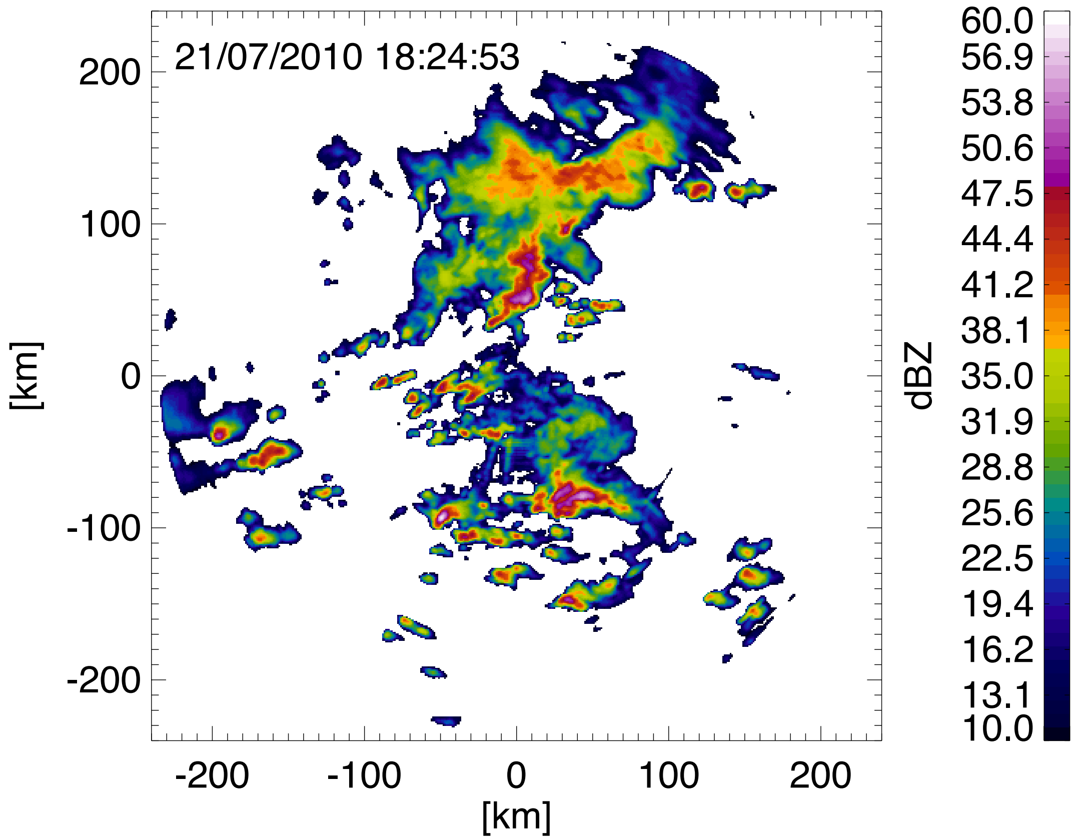
From our sample of 200 days of disdrometric records we have this expected value of logR versus logZ

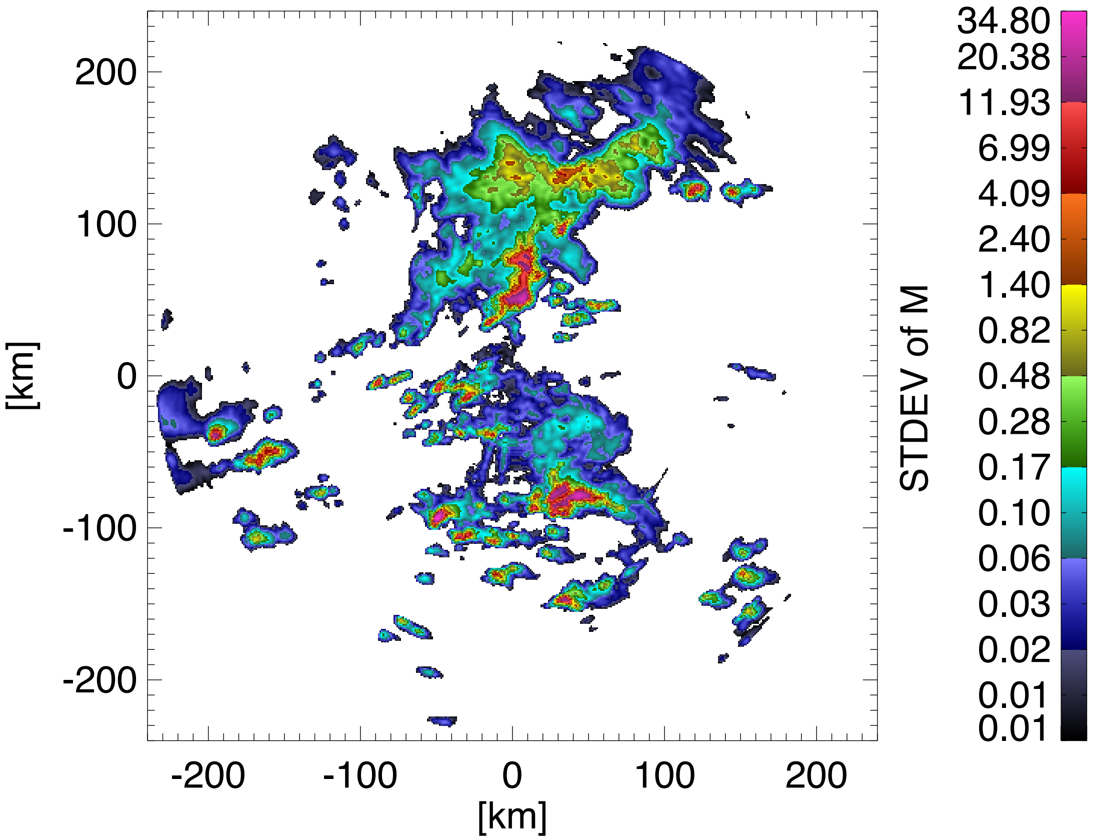


Variance is a function of intensity

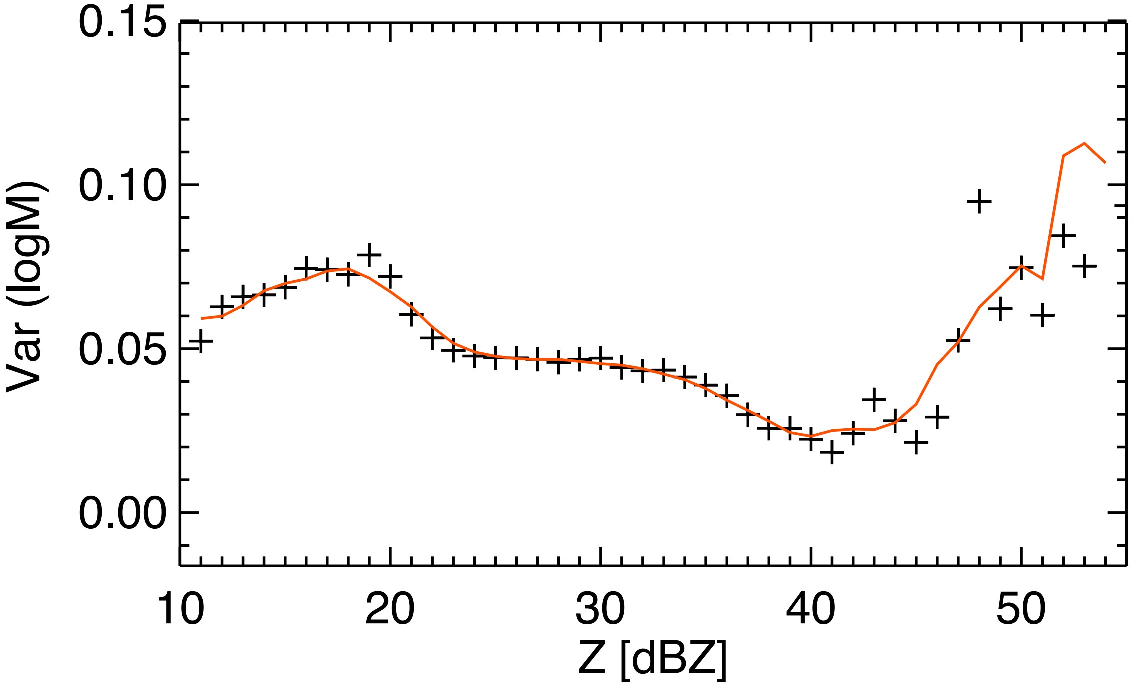


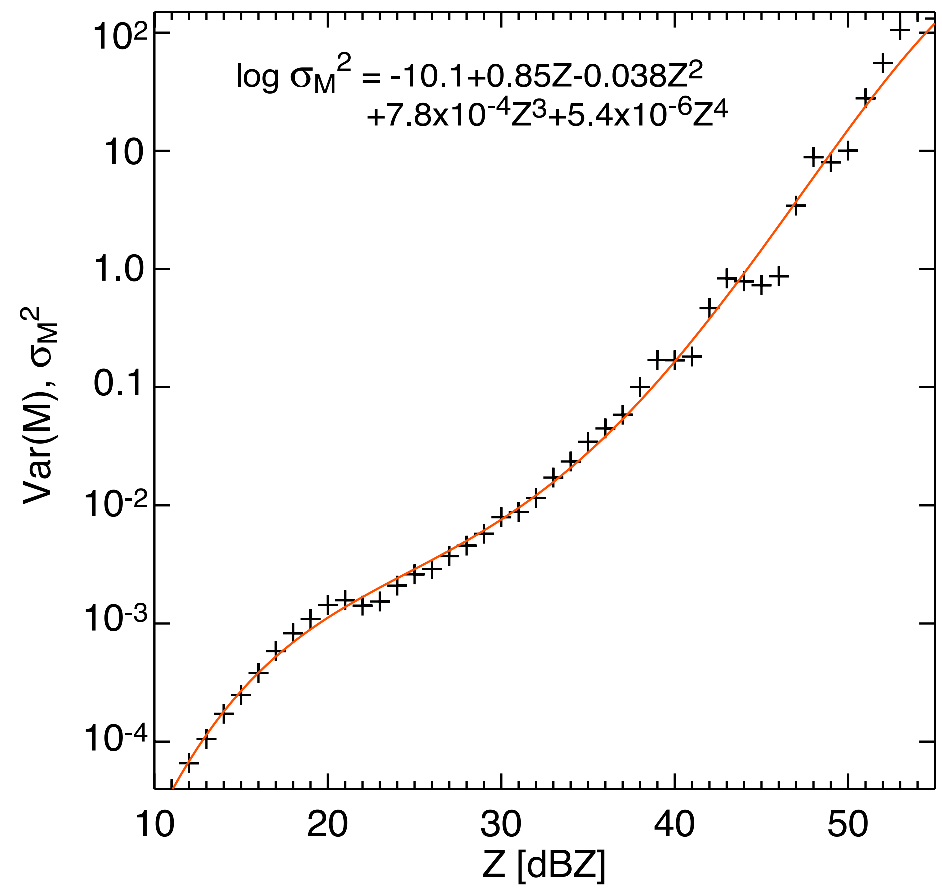




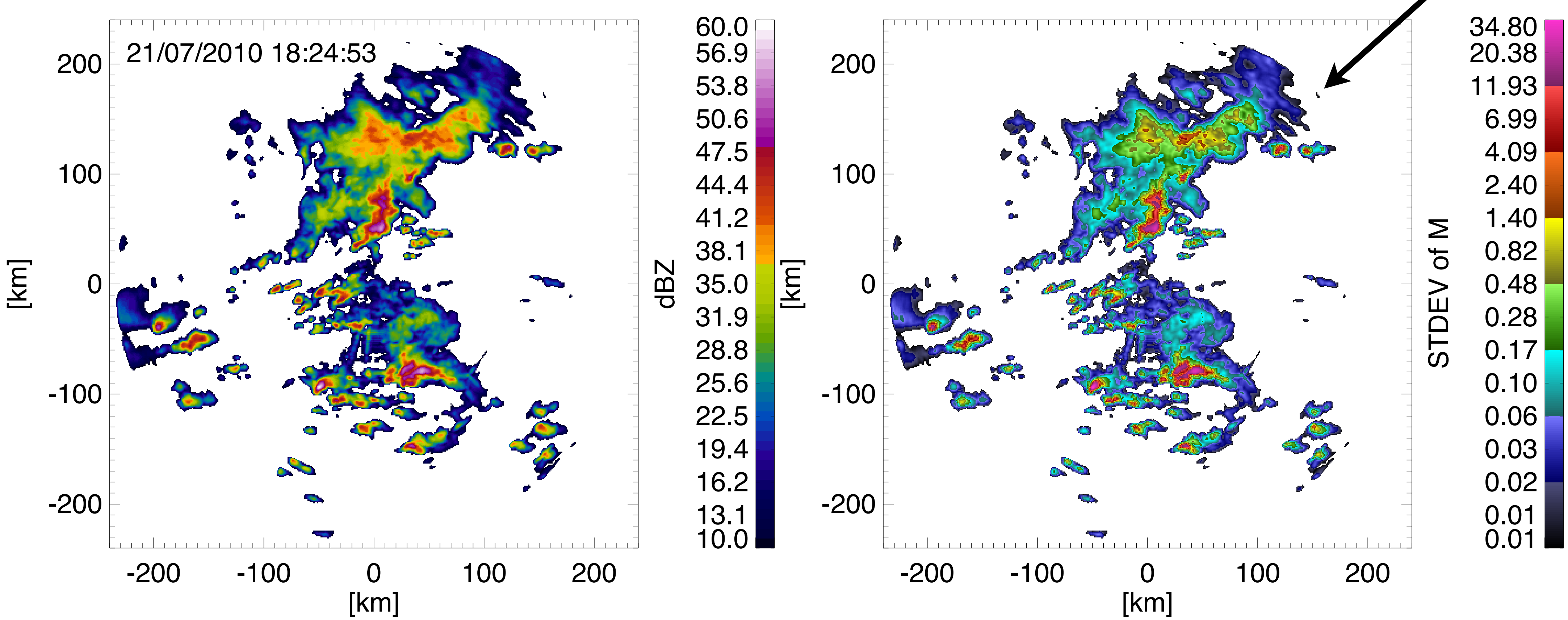


Variance is a function of intensity



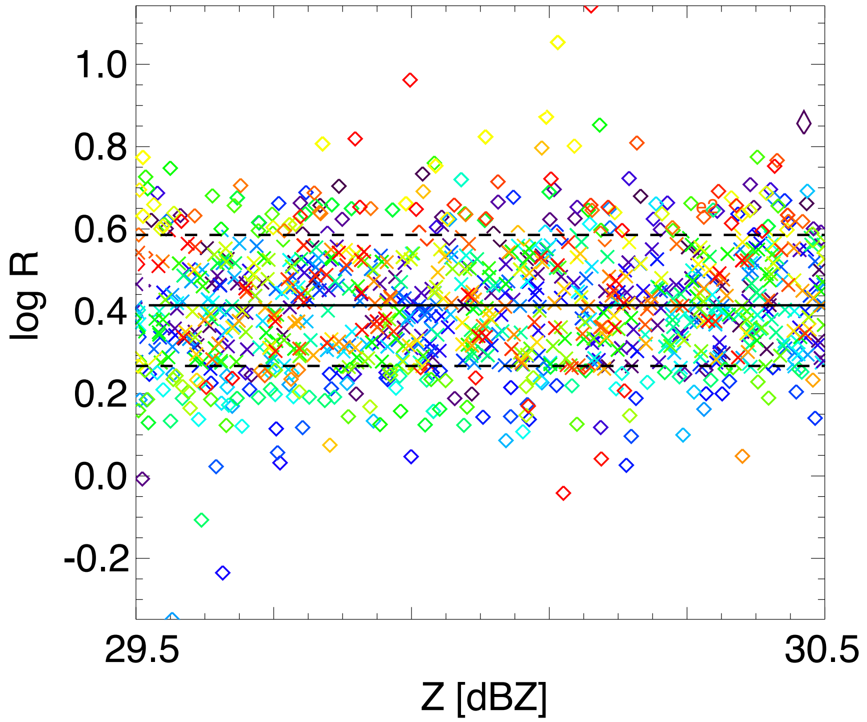


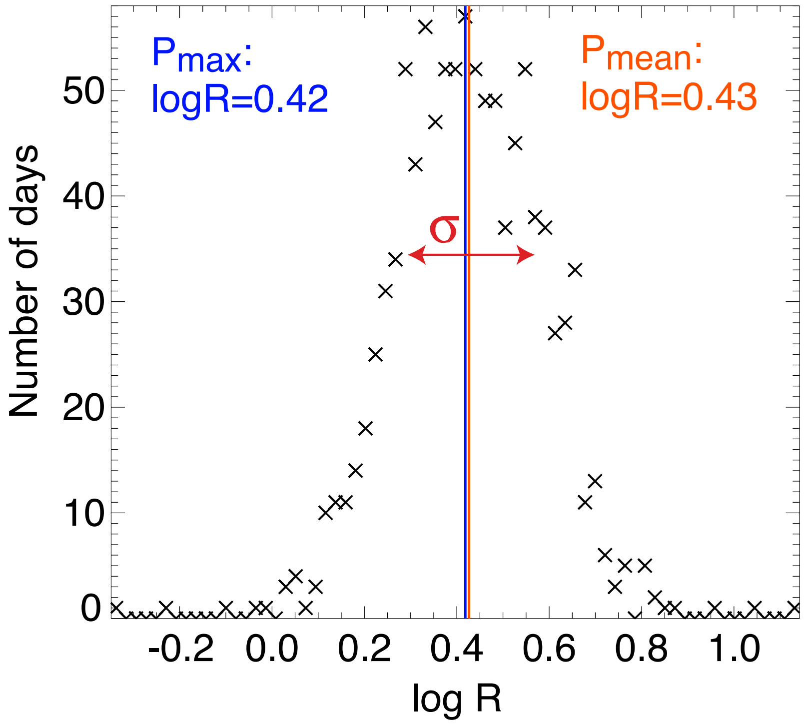
Here we have the diagonal terms of the error covariance matrix

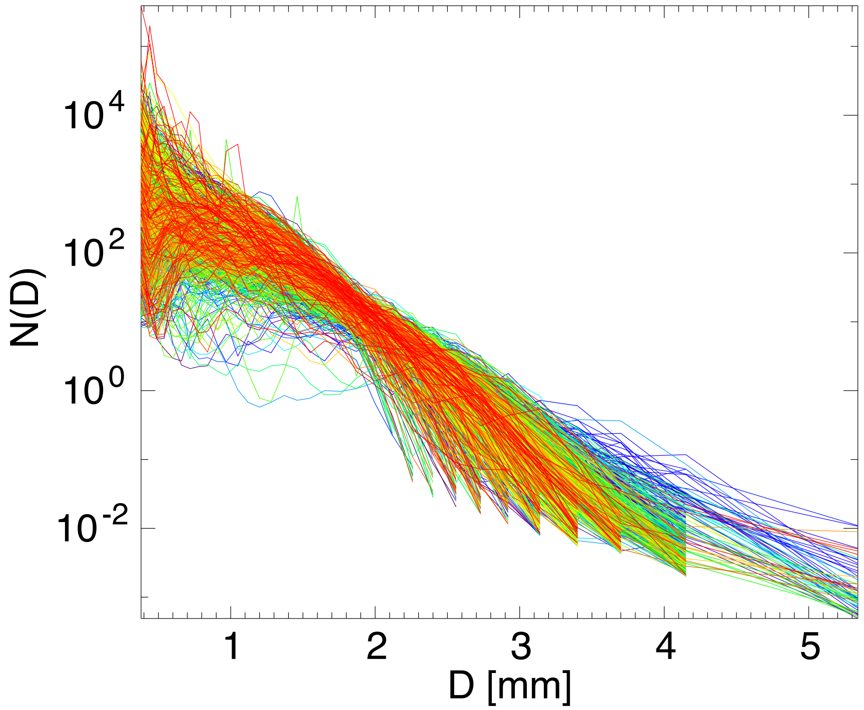


Extremes (shown here for logR)

$$\log R = \int \log r \; p(\log r \mid dBZ \pm \delta)\, dr$$

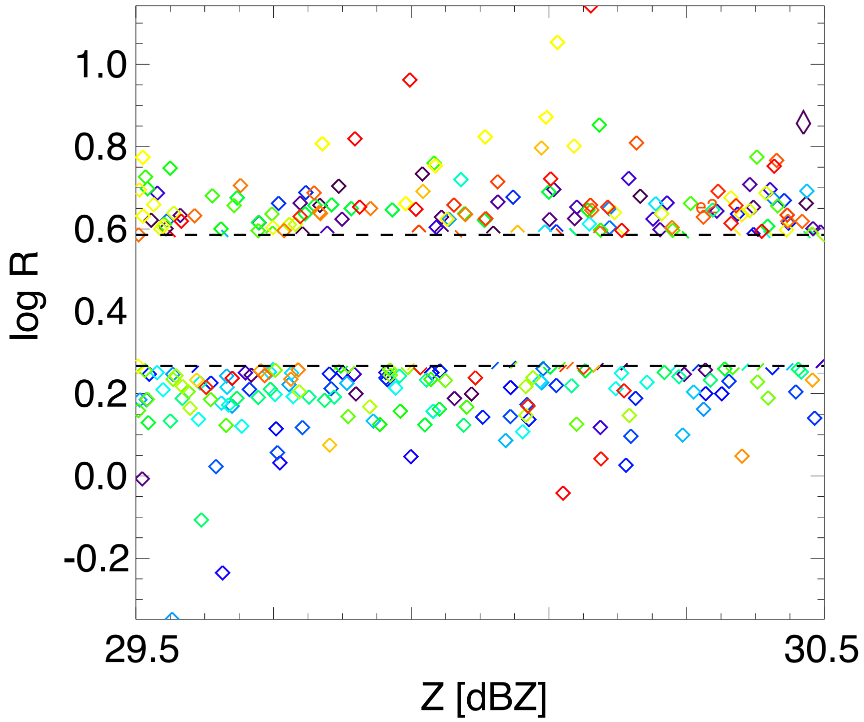


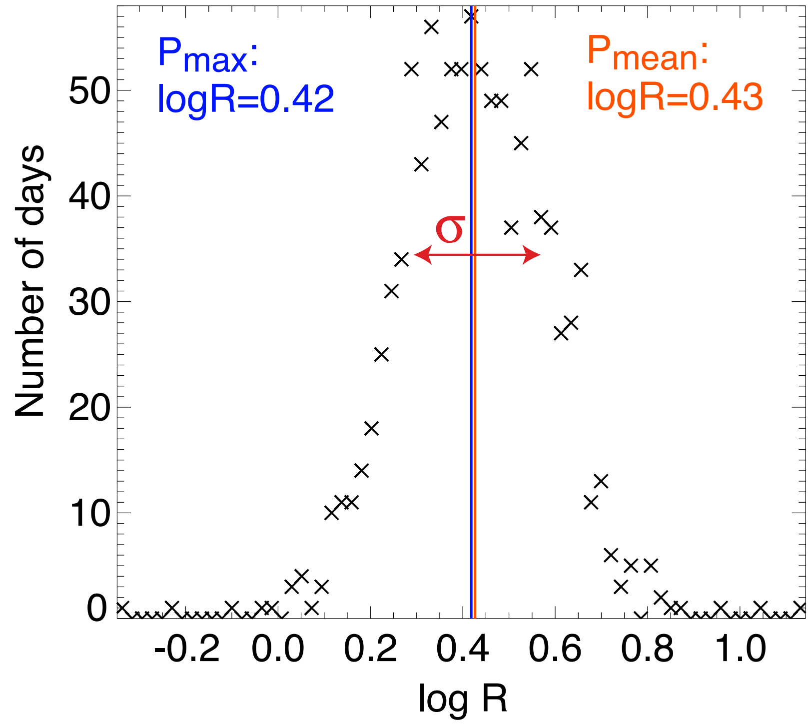


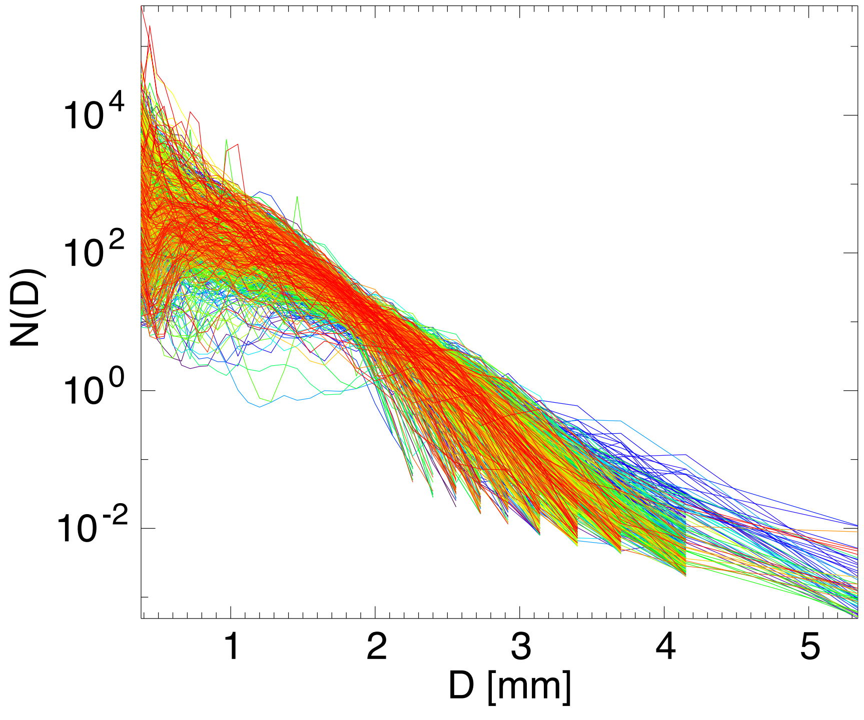


Extremes (shown here for logR)

$$\log R = \int \log r \ p(\log r \mid dBZ \pm \delta)\, dr$$

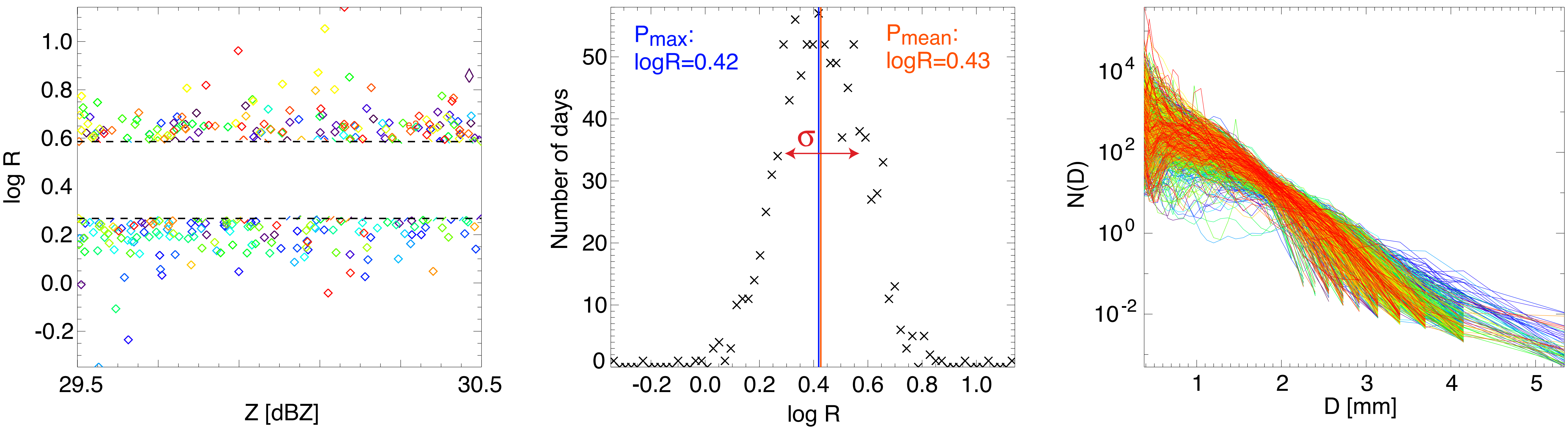




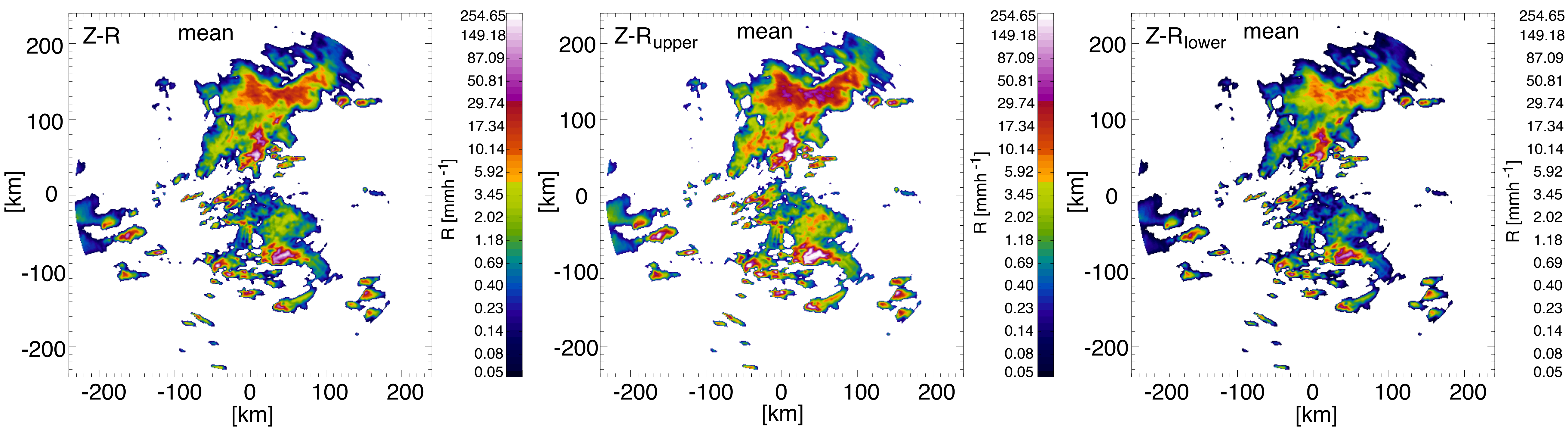


Extremes (shown here for logR)

$$\log R = \int \log r \ p(\log r \mid dBZ \pm \delta)\, dr$$

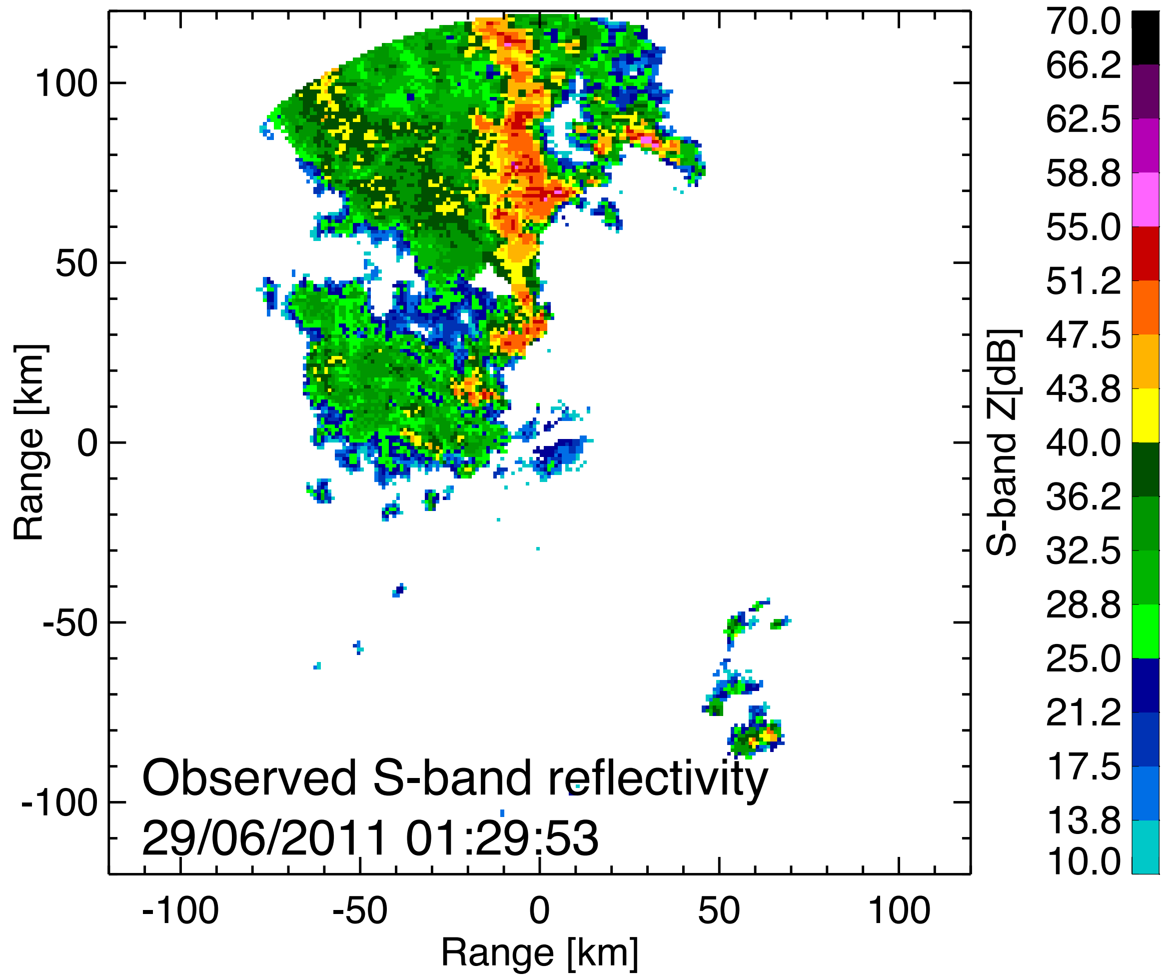


Here are some extreme members of the ensemble. The last two are averages of all possibilities outside the SD of the entire population



A more complete ensemble member

A single Z-R relationship smoothes-out existing variability in rain rate and eliminates high values

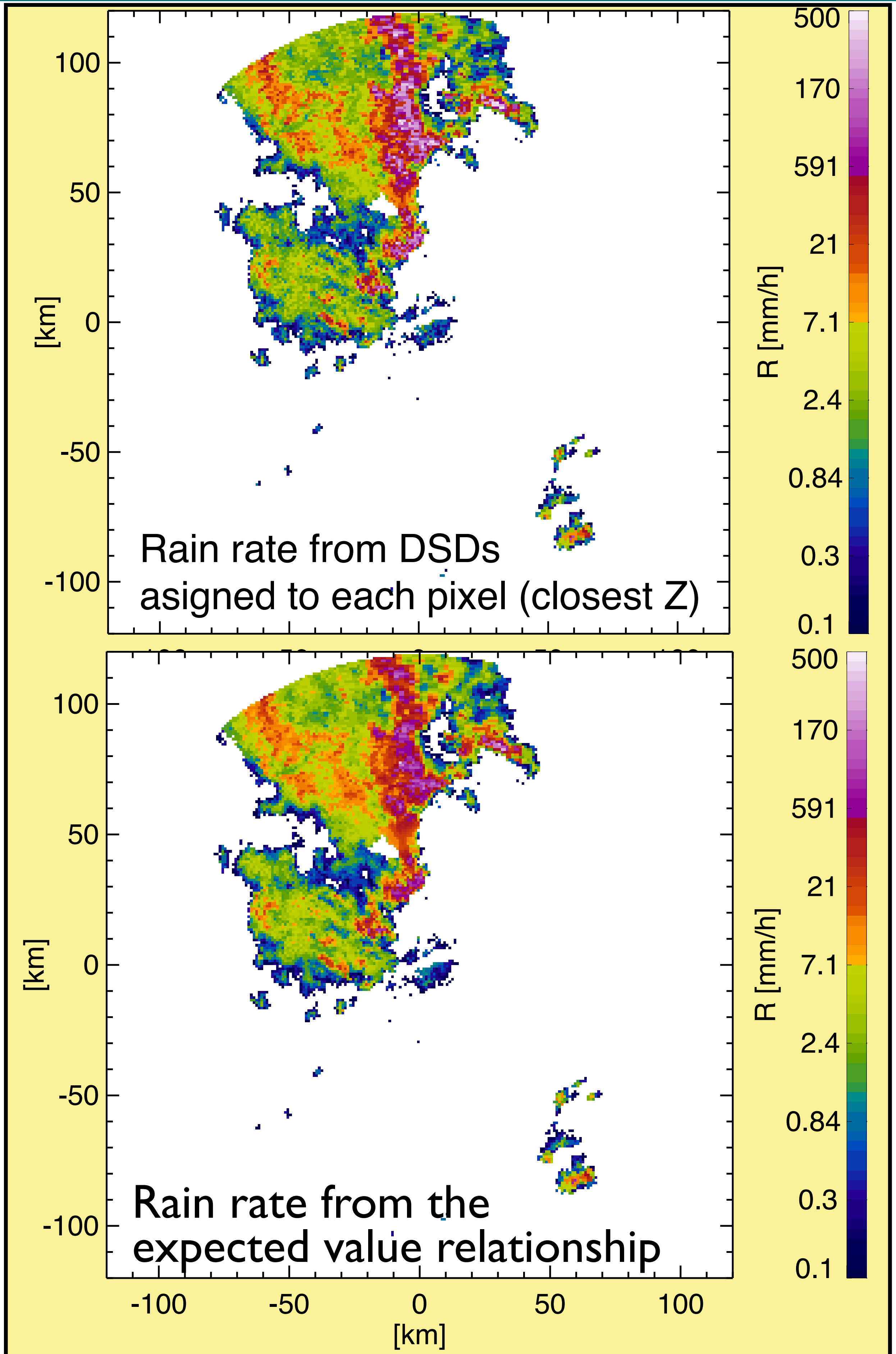


For each pixel take a DSD with the closest Z, selected from a sample of 34000 of spectra taken during convective situations (max. Z > 42dBZ).

For the selected DSD compute R for that pixel.

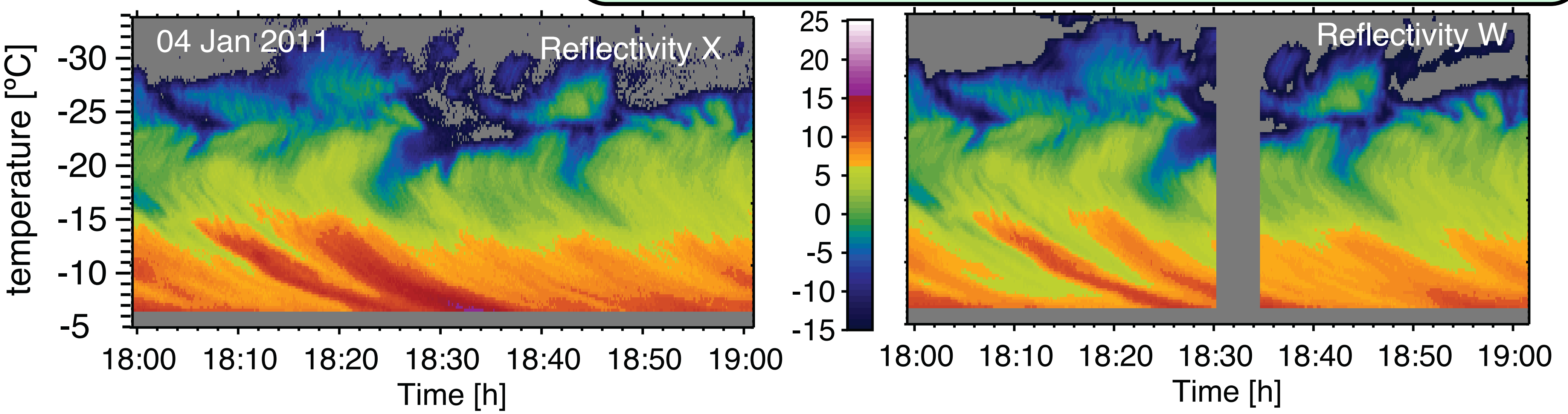
A more complete ensemble member

A single Z-R relationship smoothes-out existing variability in rain rate and eliminates high values



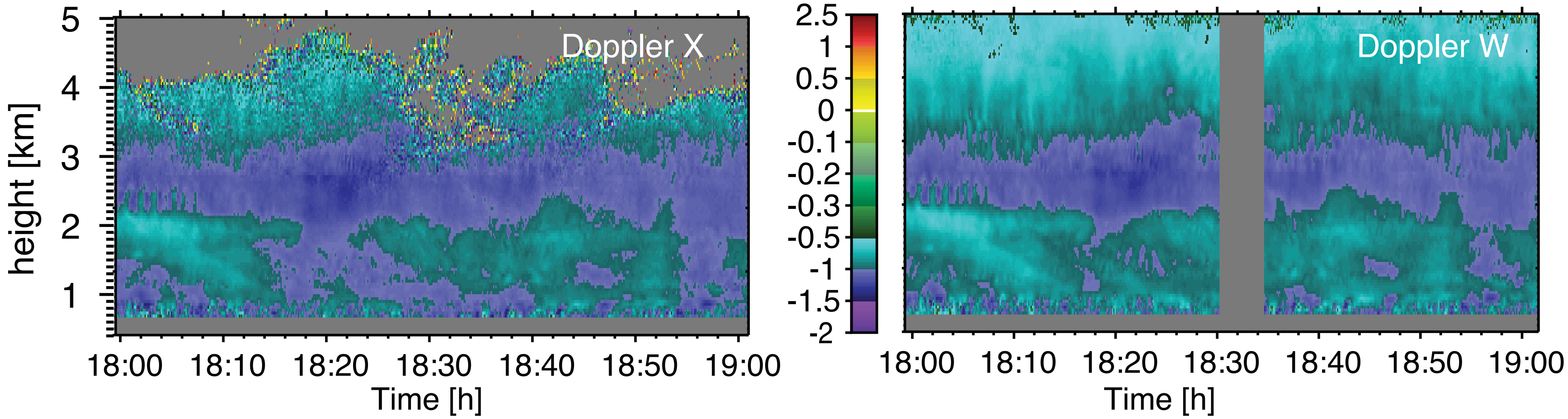
Observations

and their representation



$$Z_e^{(x)} = 3.2^4 \Big/ \left(0.93\pi^5\right) \int \sigma_x(D)\, n(D)\, dD$$

$$Z_e^{(w)} = 0.32^4 \Big/ \left(0.93\pi^5\right) \int \sigma_w\left(D, \rho_s(D)\right) n(D)\, dD$$



$$U^{(x)} = w - \frac{1}{\int \sigma_x(D)\ n(D)\, dD} \int u(D)\, \sigma_x(D)\ n(D)\, dD$$

$$U^{(w)} = w - \frac{1}{\int \sigma_w(D, \rho_s)\ n(D)\, dD} \int u(D)\, \sigma_w\left(D, \rho_s(D)\right)\ n(D)\, dD$$

four measurements and four equations, but...

Observations

$$DWR = dBZ_e^{(X)} - dBZ_e^{(W)} \qquad DDV = V^{(X)} - V^{(W)}$$

Observation are considered with their uncertainty interval. Thus, a microphysical model descriptor satisfies an observation if it falls within its uncertainty interval.

Here we formulate the retrievals as

$$N_0^* = \int n_0^* \; p(n_0^* \mid [DWR \pm \delta_{\Delta Z}], [DDV \pm \delta_{\Delta U}], [U_D^{(X)} \pm \delta_{Ux}], [dBZ^{(X)} \pm \delta_{Zx}])\, dn_0^*$$

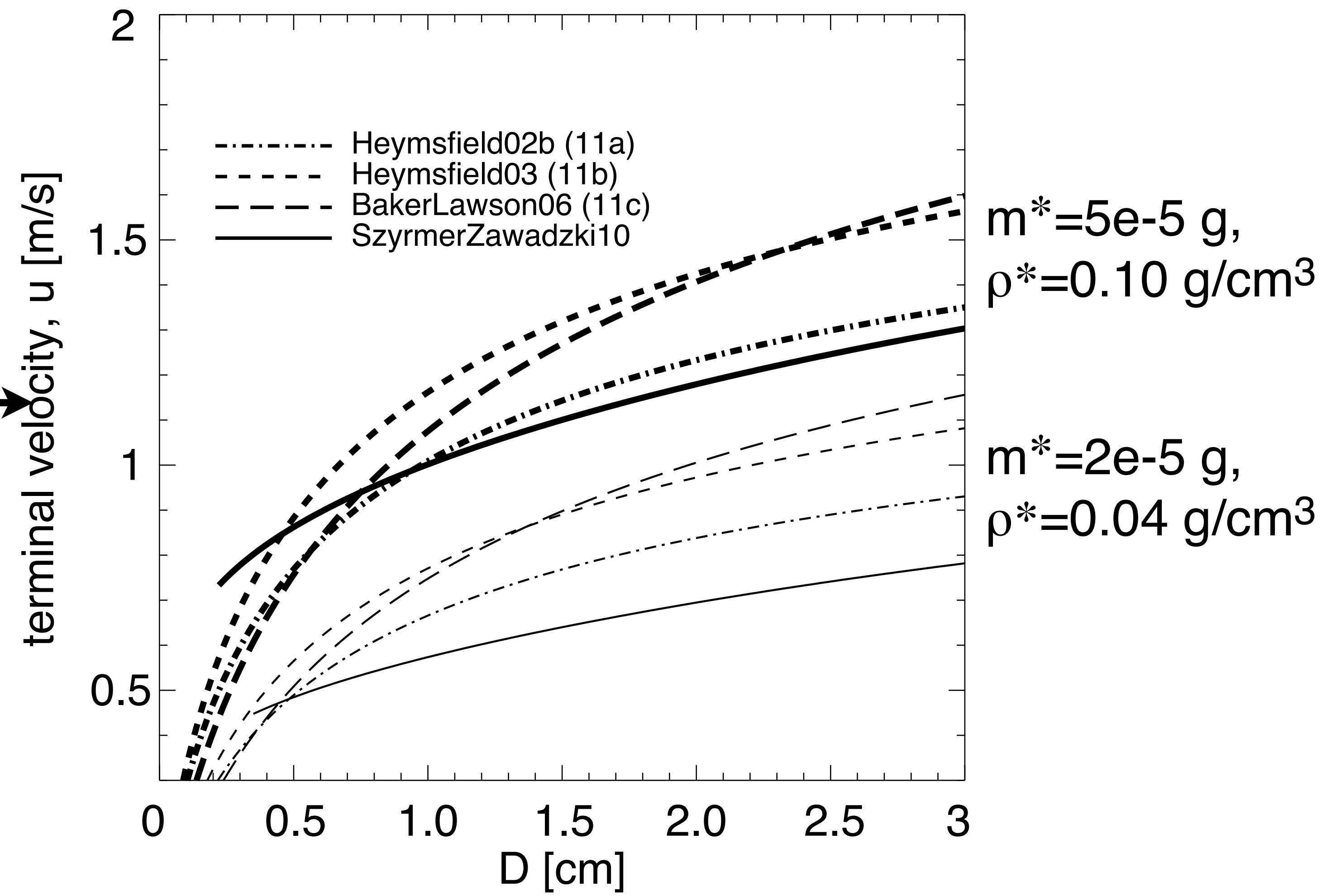
for every combination of model descriptors.

And 3 similar equations for $\rho^*, D_{2,3}, IWC$

Uncertainties in fall velocity representation

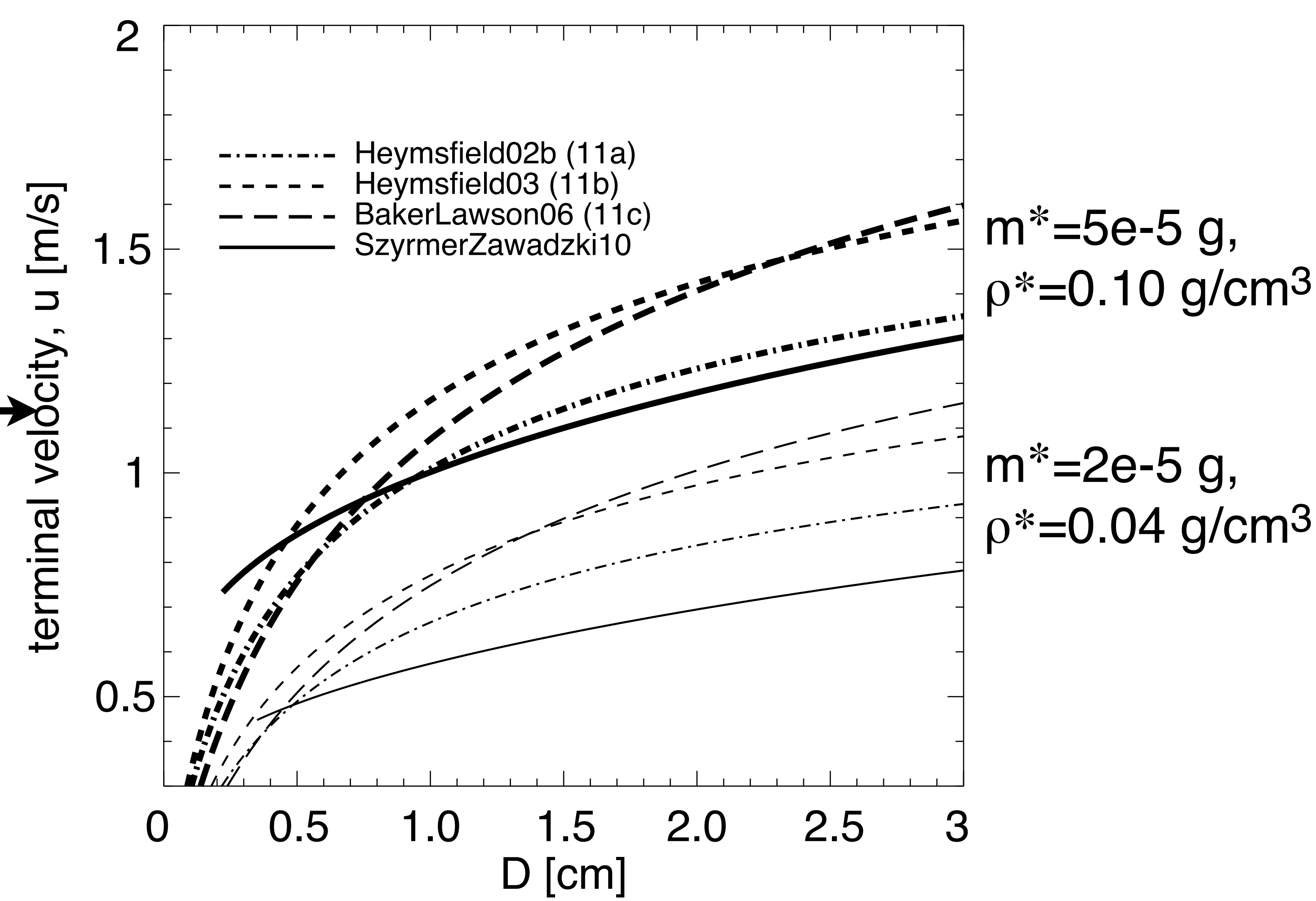
Uncertainties in fall velocity representation

We use relationships between density and fall velocity:



Uncertainties in fall velocity representation

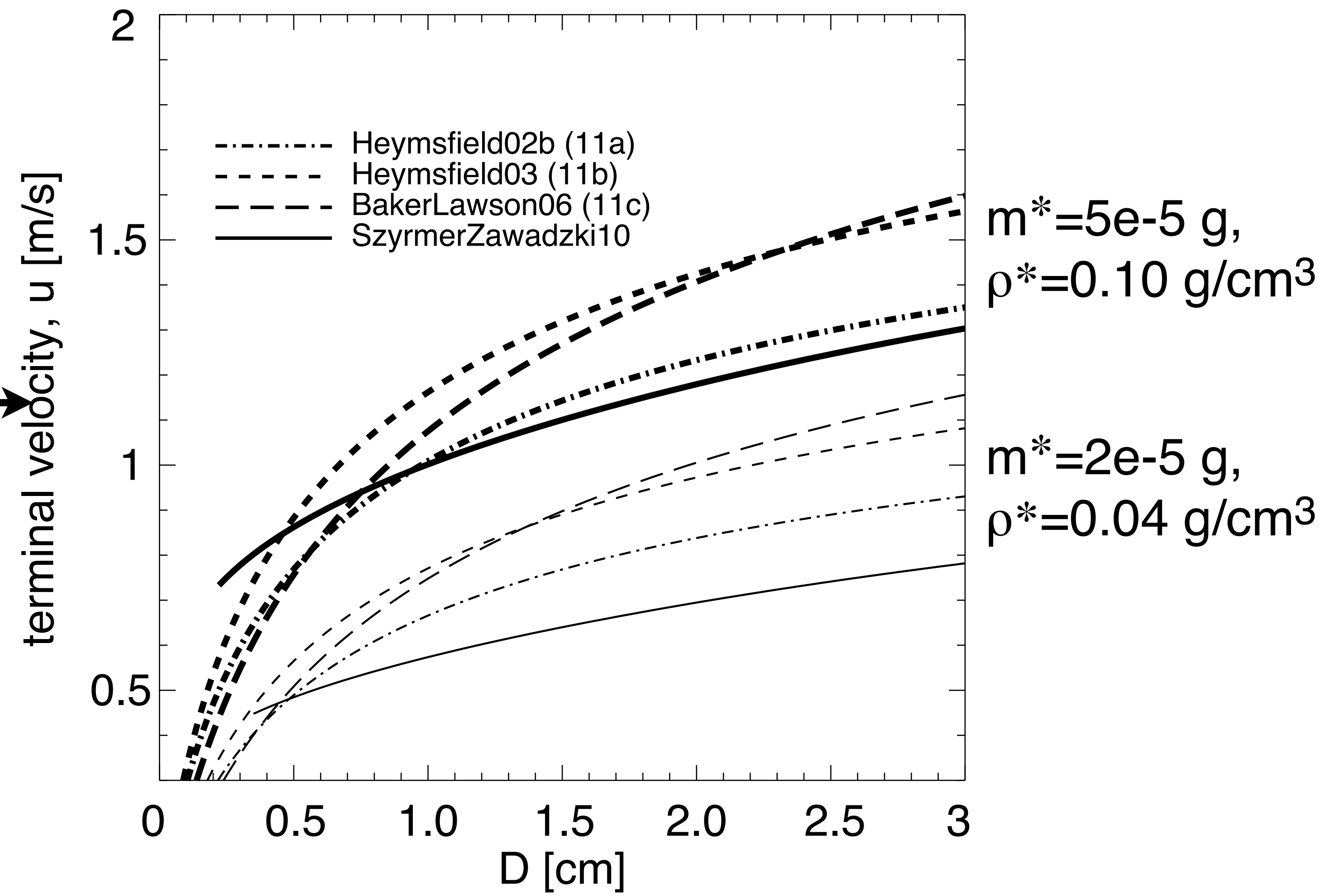
We use relationships between density and fall velocity:



The uncertainty of the model is represented by all combinations of various functional size distributions and velocity-mass relationships.

Uncertainties in fall velocity representation

We use relationships between density and fall velocity:

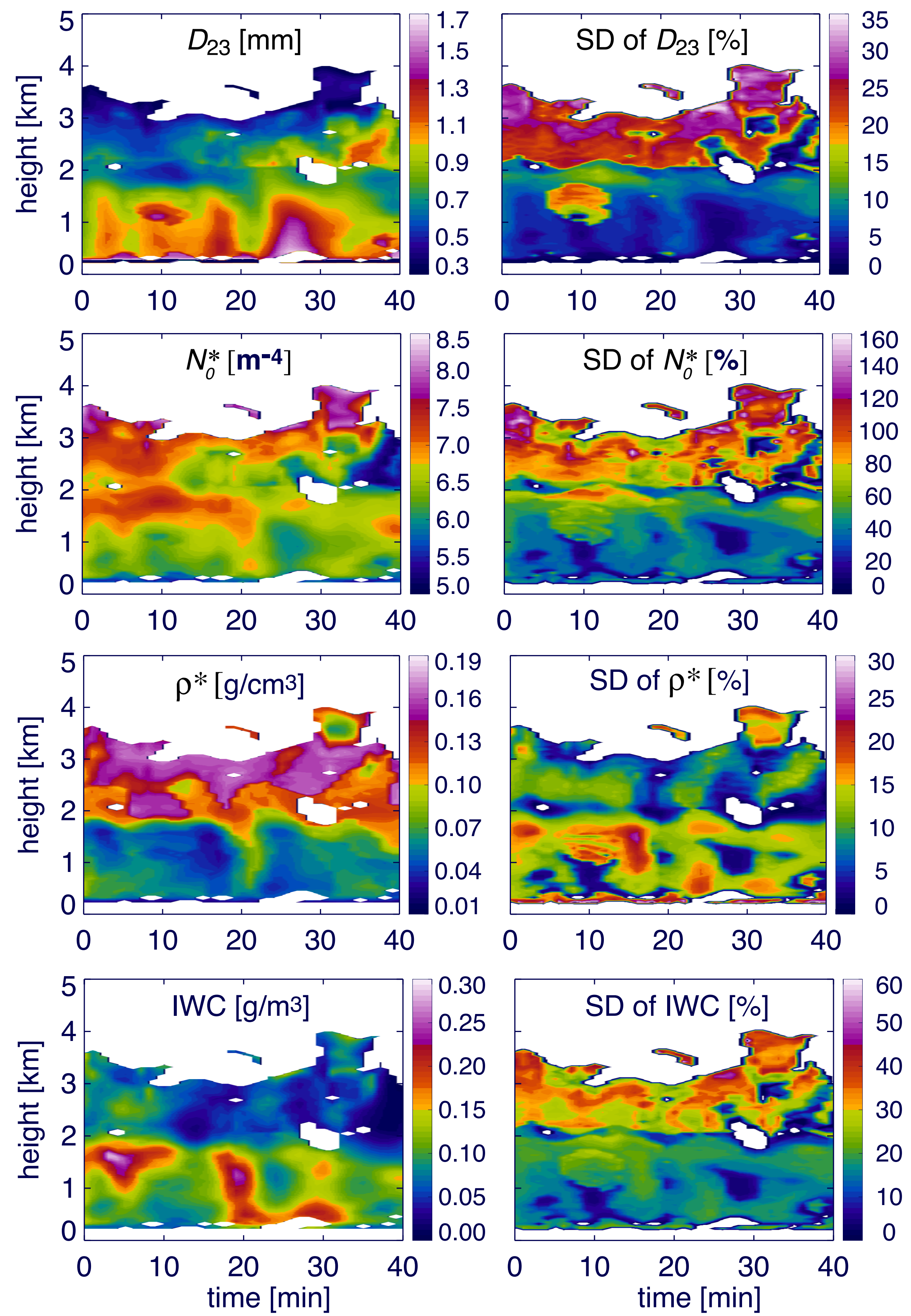


The uncertainty of the model is represented by all combinations of various functional size distributions and velocity-mass relationships.

All of the possible combinations of relationships are used as perturbations of the model and are assumed here to be equiprobable

Results: Expected values and SDs of retrieved snow descriptors

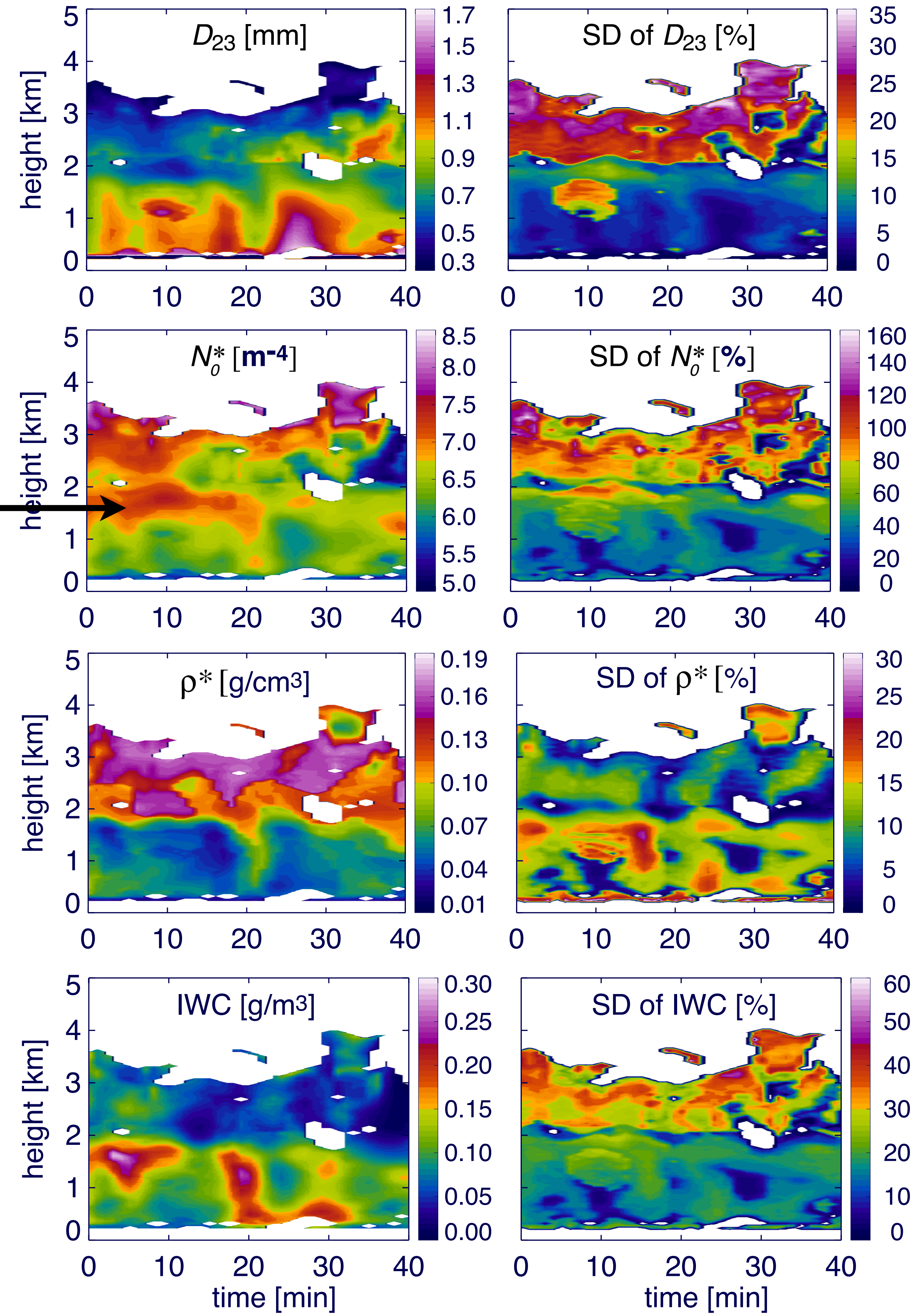
SDs here are due to model uncertainty only



Results: Expected values and SDs of retrieved snow descriptors

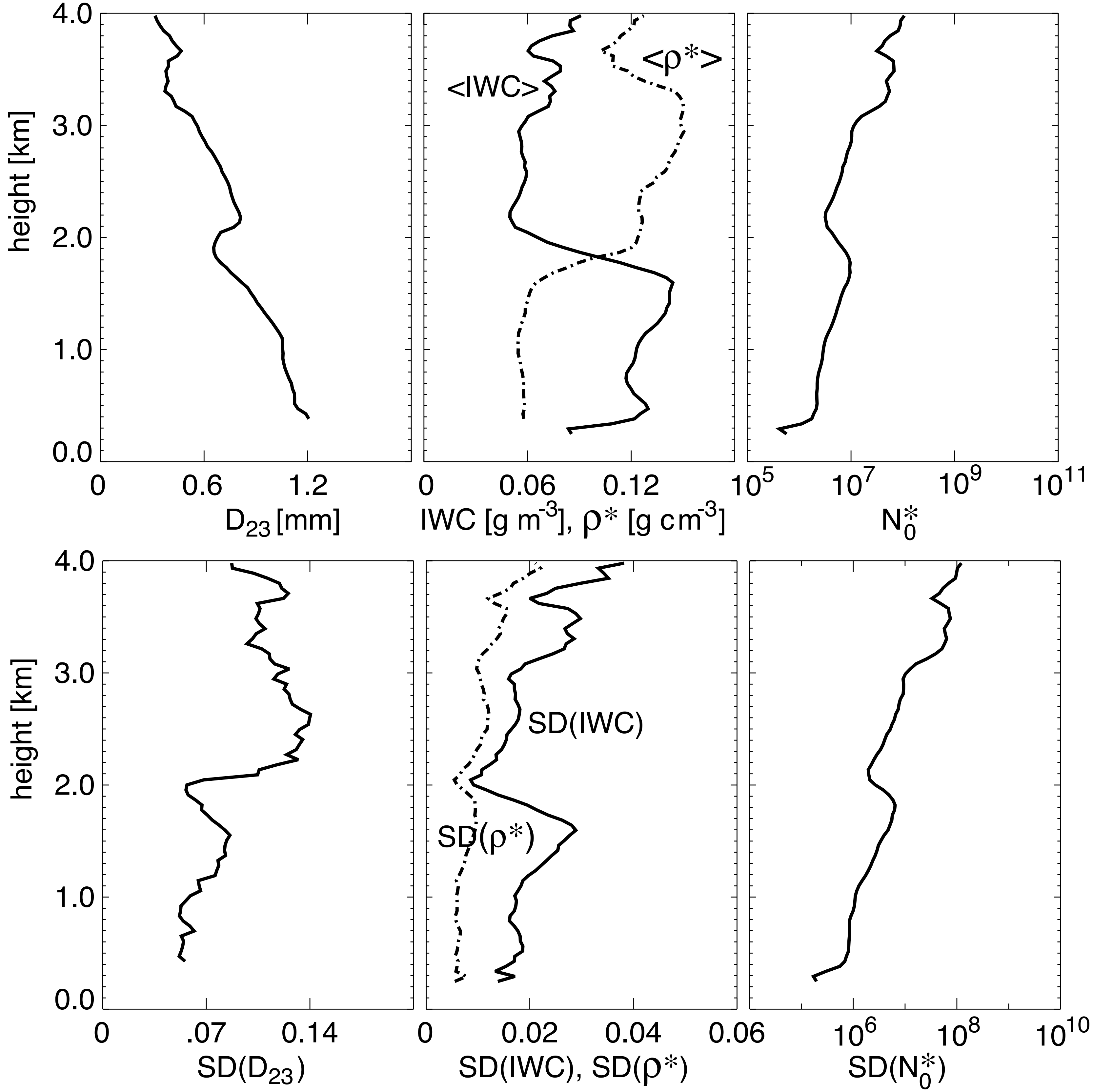
SDs here are due to model uncertainty only

Note the increase of particle concentration at ~ 2 km (-15°C)

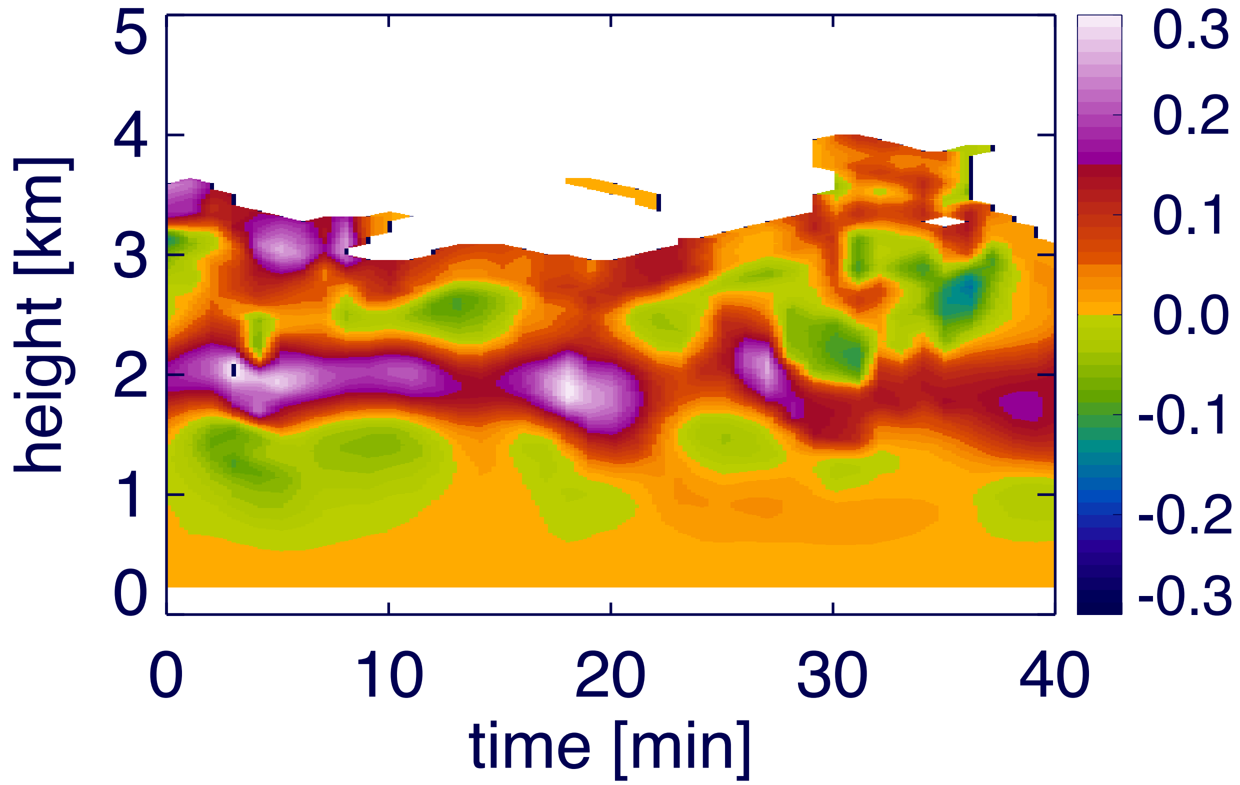


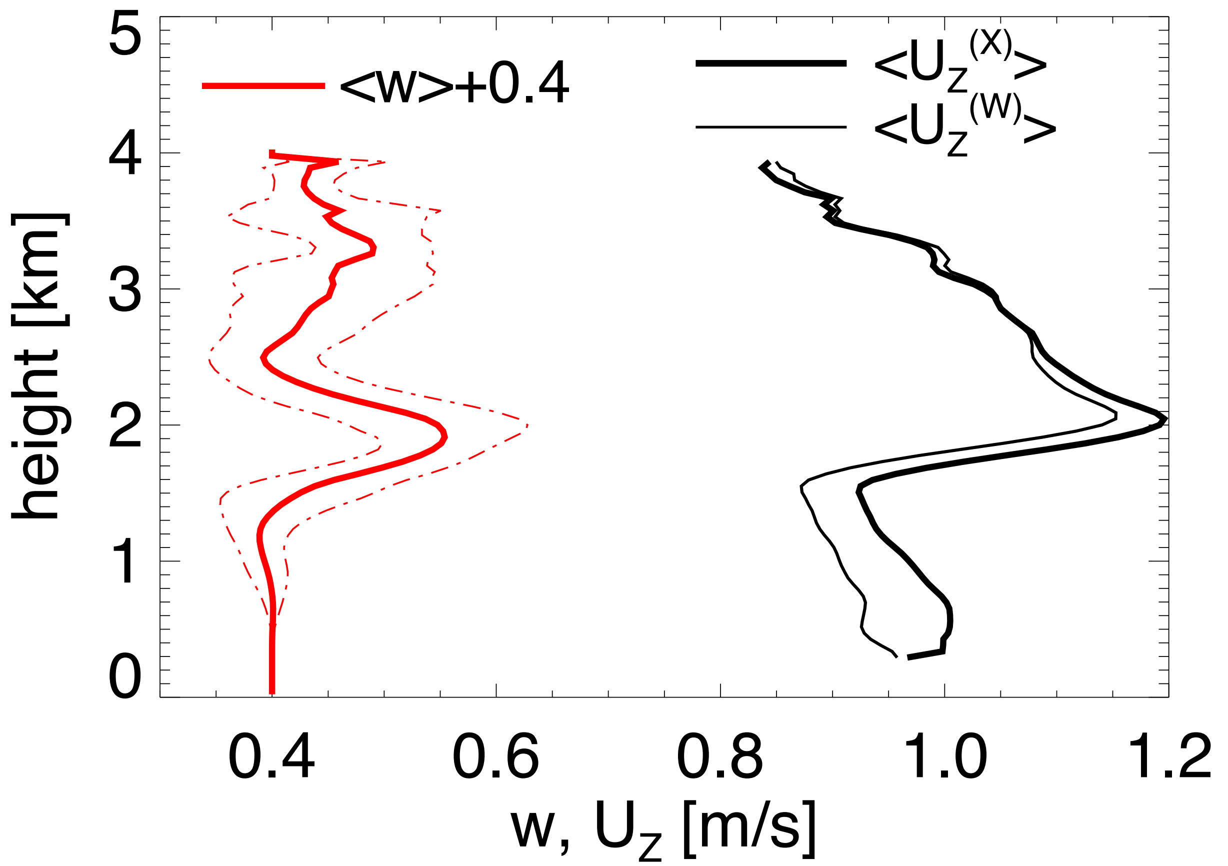
Results: Expected values and SDs of retrieved snow descriptors

SDs here are due to model uncertainty only

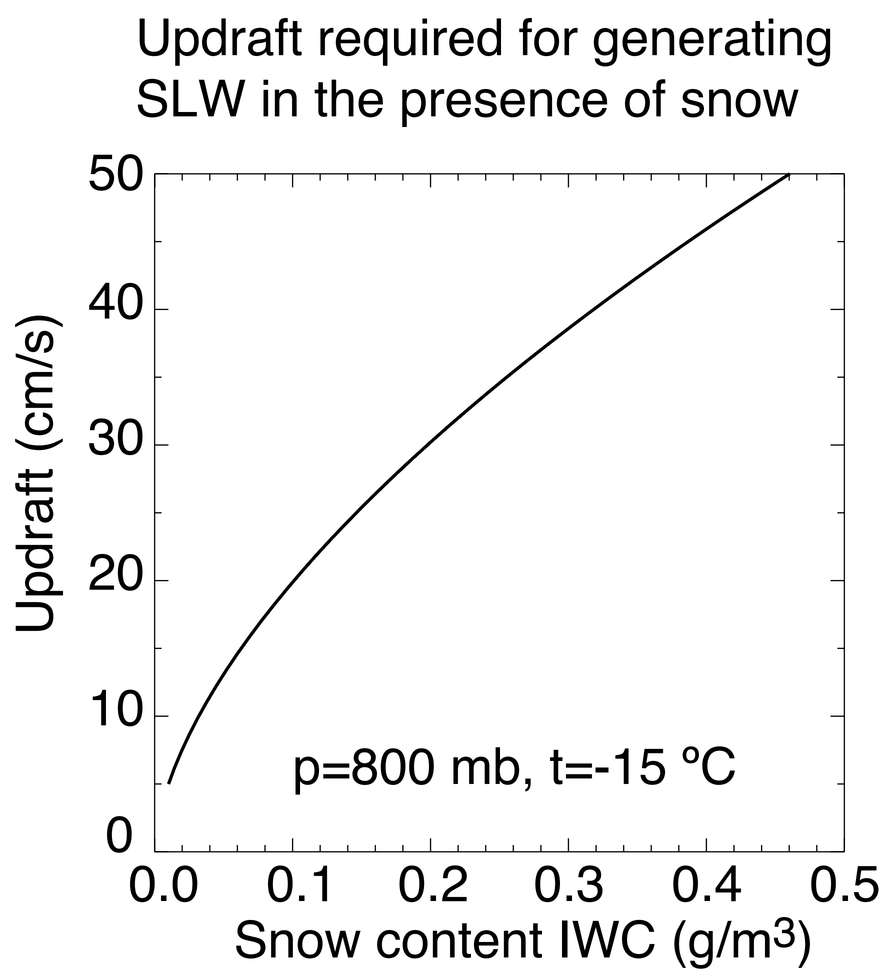


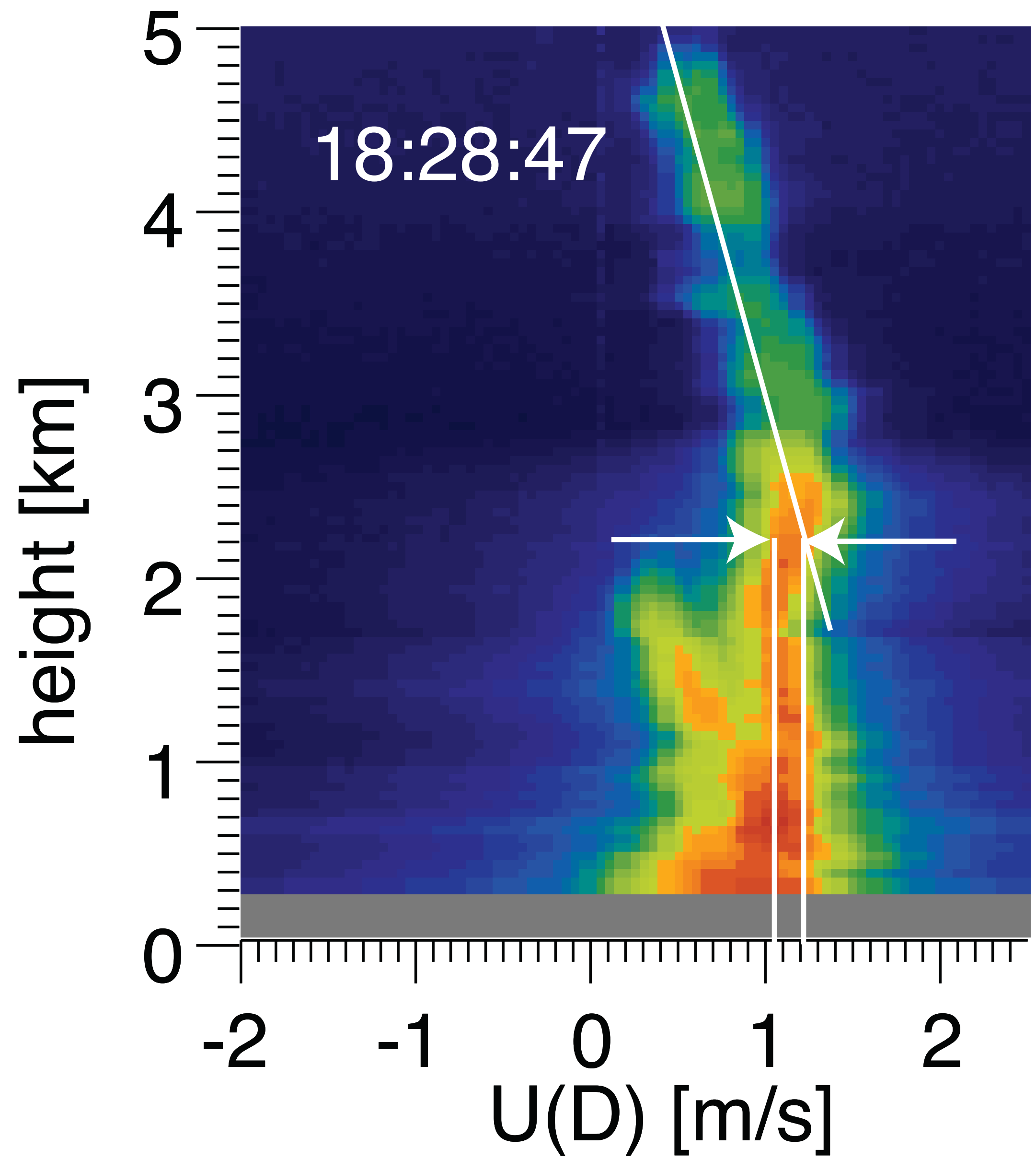
Results: Vertical air motions





Verification





Conclusions

Ensemble retrievals reduce biases

Conclusions

Ensemble retrievals reduce biases

Give information on the error covariance matrix

Conclusions

Ensemble retrievals reduce biases

Give information on the error covariance matrix

Lead to new insight

Conclusions

Ensemble retrievals reduce biases

Give information on the error covariance matrix

Lead to new insight

Are as simple to perform as common retrievals

Conclusions

Ensemble retrievals reduce biases

Give information on the error covariance matrix

Lead to new insight

Are as simple to perform as common retrievals

Help to establish a common language between observational and numerical teams of researchers

Conclusions

Ensemble retrievals reduce biases

Give information on the error covariance matrix

Lead to new insight

Are as simple to perform as common retrievals

Help to establish a common language between observational and numerical teams of researchers

There is no reason for not doing things this way