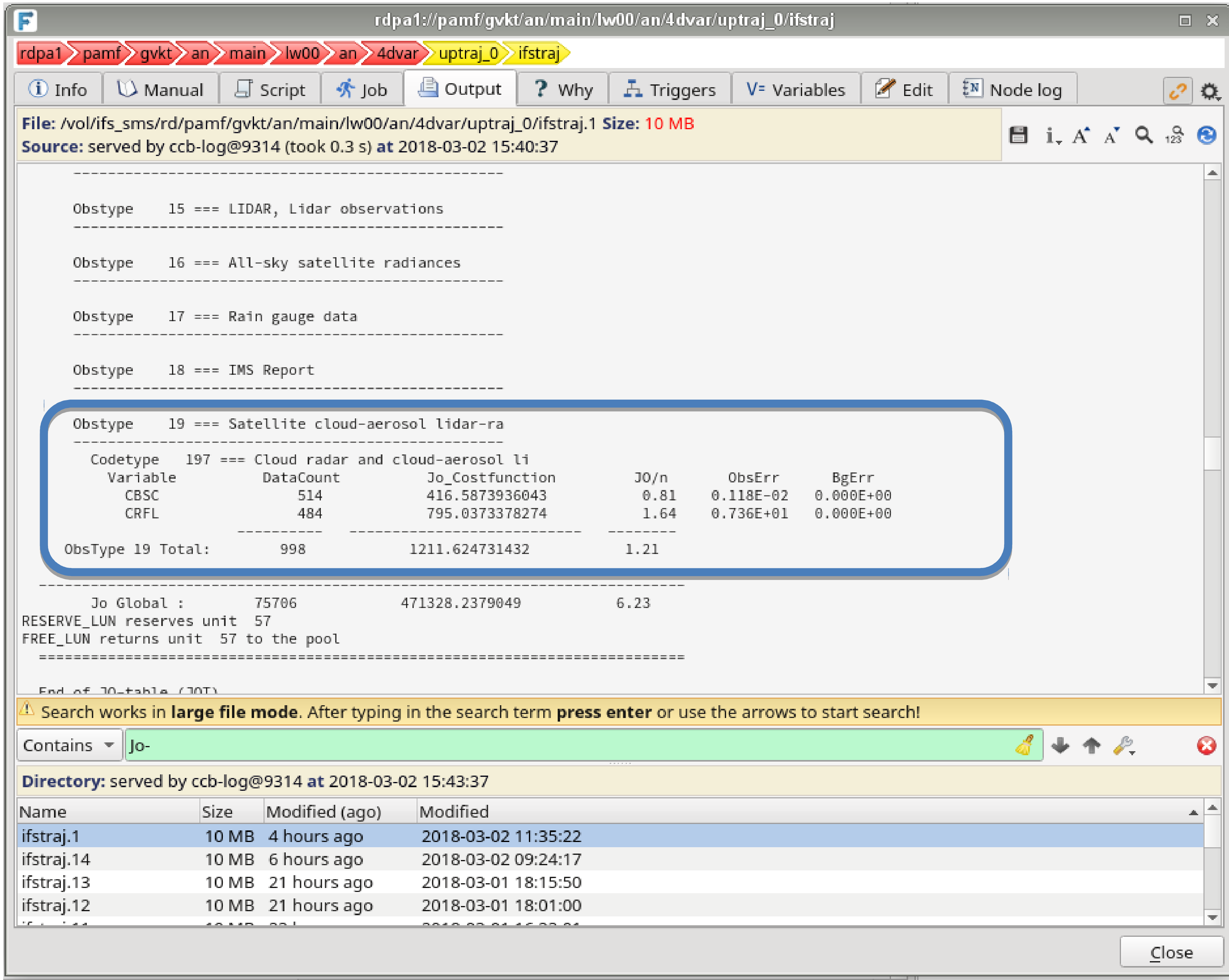
# Assimilating profiles of cloud radar and lidar observations into the ECMWF 4D-Var system

*Characterising observation errors*

Mark Fielding, Marta Janisková

European Centre for Medium-Range Weather Forecasts

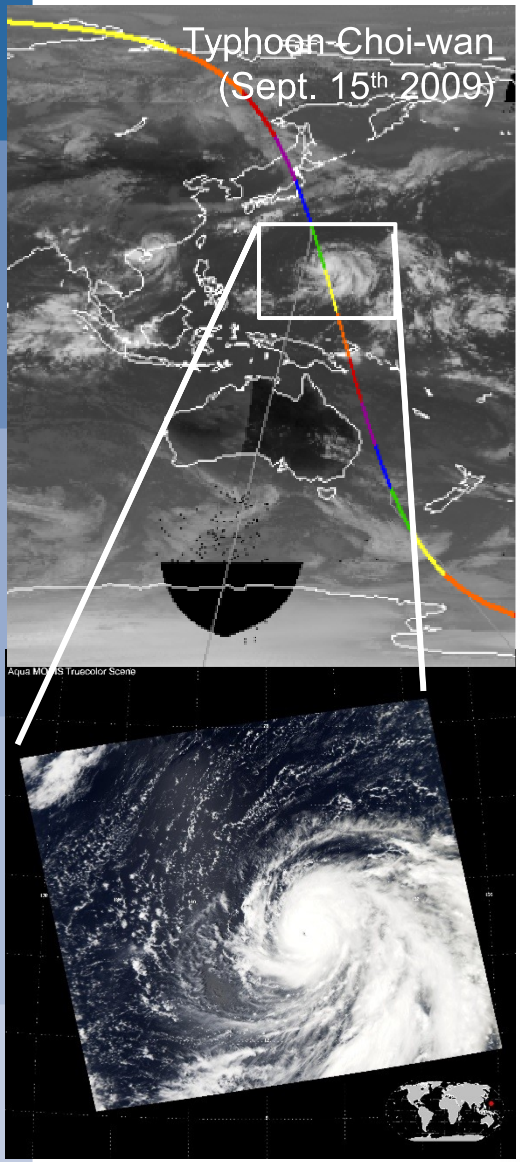
Thanks: Robin Hogan, Olaf Stiller


rdpa1://pamf/gvkt/an/main/lw00/an/4dvar/uptraj_0/ifstraj
rdpa1 pamf gvkt an main lw00 an 4dvar uptraj_0 ifstraj
Info Manual Script Job Output Why Triggers Variables Edit Node log
File: /vol/ifs_sms/rd/pamf/gvkt/an/main/lw00/an/4dvar/uptraj_0/ifstraj.1 Size: 10 MB
Source: served by ccb-log@9314 (took 0.3 s) at 2018-03-02 15:40:37
Obstype 15 === LIDAR, Lidar observations
Obstype 16 === All-sky satellite radiances
Obstype 17 === Rain gauge data
Obstype 18 === IMS Report
Obstype 19 === Satellite cloud-aerosol lidar-ra
Codetype 197 === Cloud radar and cloud-aerosol li
Variable DataCount Jo_Costfunction JO/n ObsErr BgErr
CBSC 514 416.5873936043 0.81 0.118E-02 0.000E+00
CRFL 484 795.0373378274 1.64 0.736E+01 0.000E+00
ObsType 19 Total: 998 1211.624731432 1.21
Jo Global : 75706 471328.2379049 6.23
RESERVE_LUN reserves unit 57
FREE_LUN returns unit 57 to the pool
Search works in large file mode. After typing in the search term press enter or use the arrows to start search!
Contains
Jo-
Directory: served by ccb-log@9314 at 2018-03-02 15:43:37
Name Size Modified (ago) Modified
ifstraj.1 10 MB 4 hours ago 2018-03-02 11:35:22
ifstraj.14 10 MB 6 hours ago 2018-03-02 09:24:17
ifstraj.13 10 MB 21 hours ago 2018-03-01 18:15:50
ifstraj.12 10 MB 21 hours ago 2018-03-01 18:01:00
Close

# Background

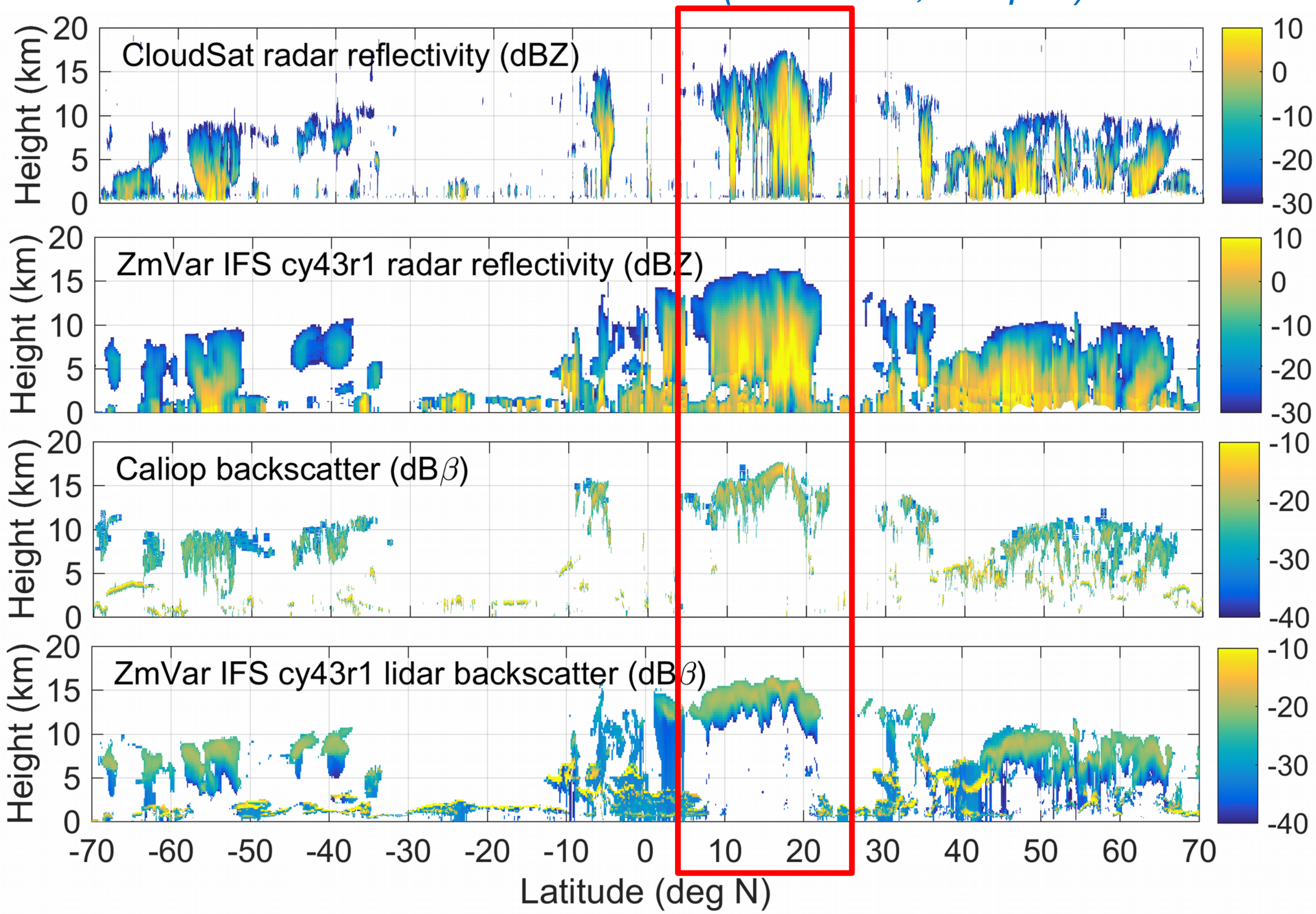
- Cloud radars have been the workhorse for understanding cloud processes and improving cloud and radiation schemes in NWP and climate models.
- Radar pulses penetrate all but deepest convection; they observe cloud structure.
- Cloud lidars have also been invaluable for NWP model evaluation by being adept at detecting cloud boundaries, ice cloud properties (particularly powerful in synergy with radar).
- Although observations of cloud are routinely assimilated at ECMWF and other NWP centres (e.g., microwave radiances), profiles of radar and lidar (e.g., CloudSAT and CALIPSO) have not been assimilated operationally.
- 1D+4D-Var experiments are encouraging; radar and lidar show positive impact in analysis and subsequent forecast (Janisková, 2015)

# Profiling measurements reveal cloud structure




•ECMWF's radar and lidar observation operators now show great agreement with CloudSat and Calipso measurements

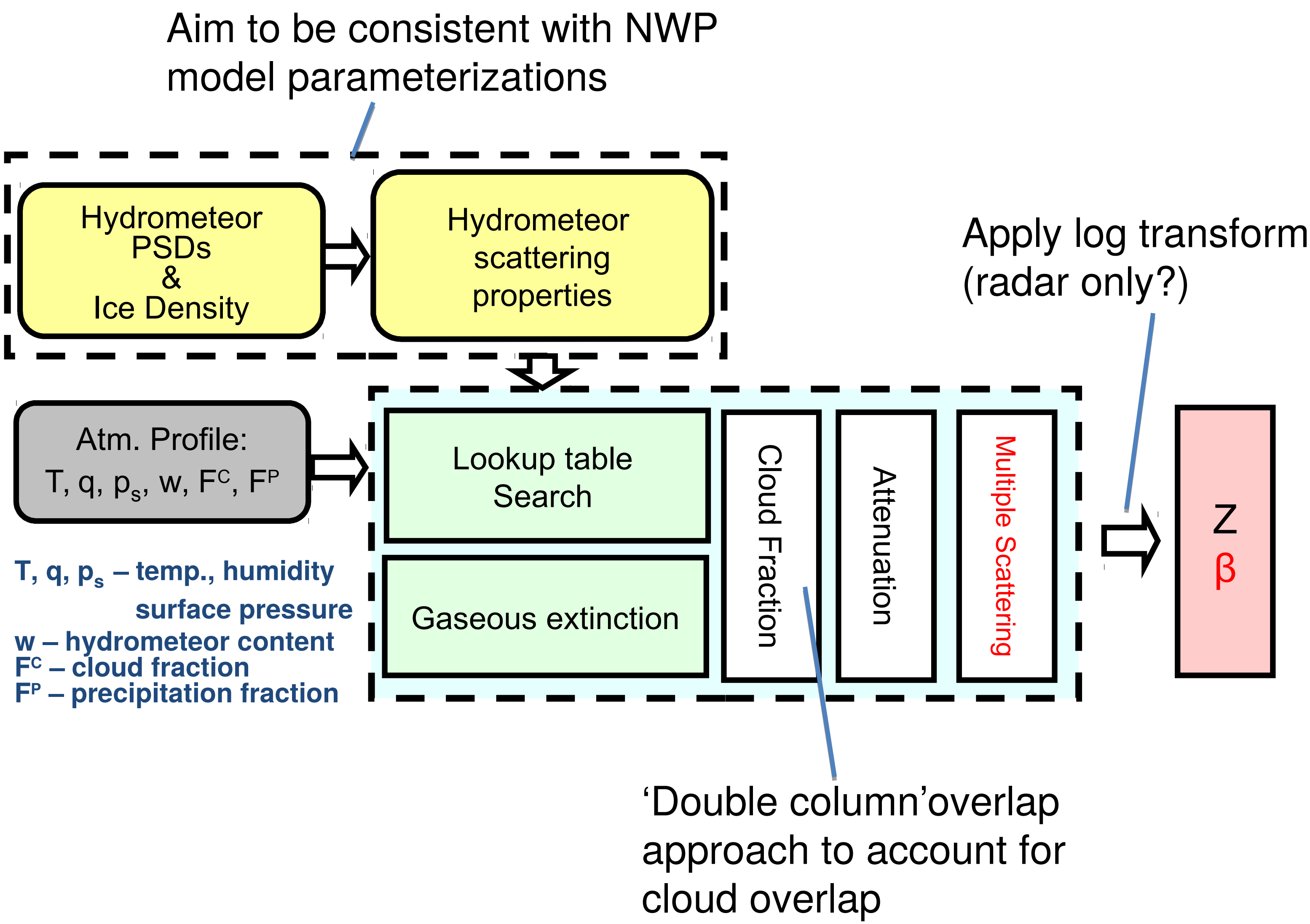
*cloud radar reflectivity (at 94 GHz, CloudSat)*
*cloud lidar backscatter (at 532 nm, Calipso)*



# Challenges of assimilating cloud radar and lidar

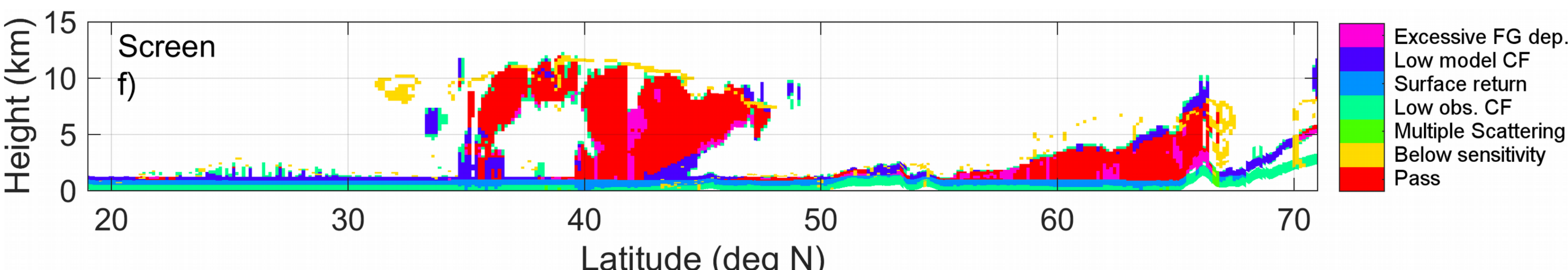
- Simplified **moist physics** (used in 4D-Var minimization) must sufficiently represent relevant cloud processes (autoconversion/accretion, evaporation, super-cooled liquid, …)
- Complex **observation operators** (non-linearity, microphysics, single scattering properties, multiple scattering, attenuation, cloud overlap, …)
- Non-trivial **bias correction** and careful **screening** are required
- Difficulties characterising **observation errors** (representativity, heteroskedacity and correlation between observations)
- **Data availability** (of satellite measurements for real-time assimilation)
  - EarthCARE (Earth Clouds, Aerosols and Radiation Explorer), due for launch in 2020, will provide opportunity for realtime assimilation.

# Observation operator



# Observation screening and quality control

- Observations are superobbed to (at least) model gridbox and level.
- A balance between including as much information from observations as possible whilst preventing ‘bad’ ones from degrading the analysis/forecast.
- Screening indicators:
  - height of observations
  - cloud fraction given by model and observations
  - plausible bounds for radar/lidar (observation & model equivalent to observations)
  - first guess departures
  - avoiding radar multiple scattering and lidar excessive attenuation



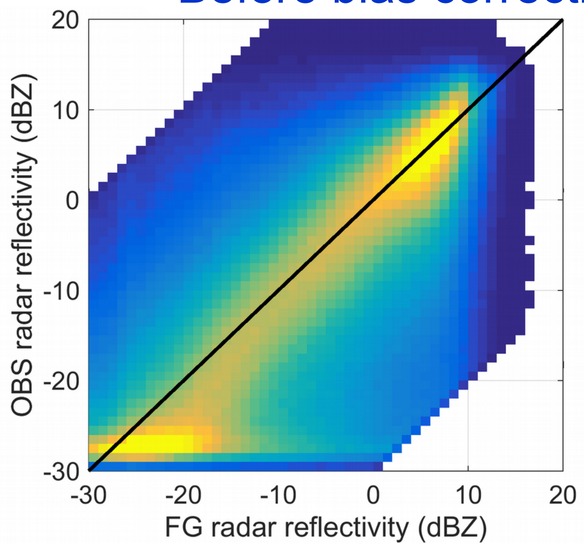
# Bias correction scheme

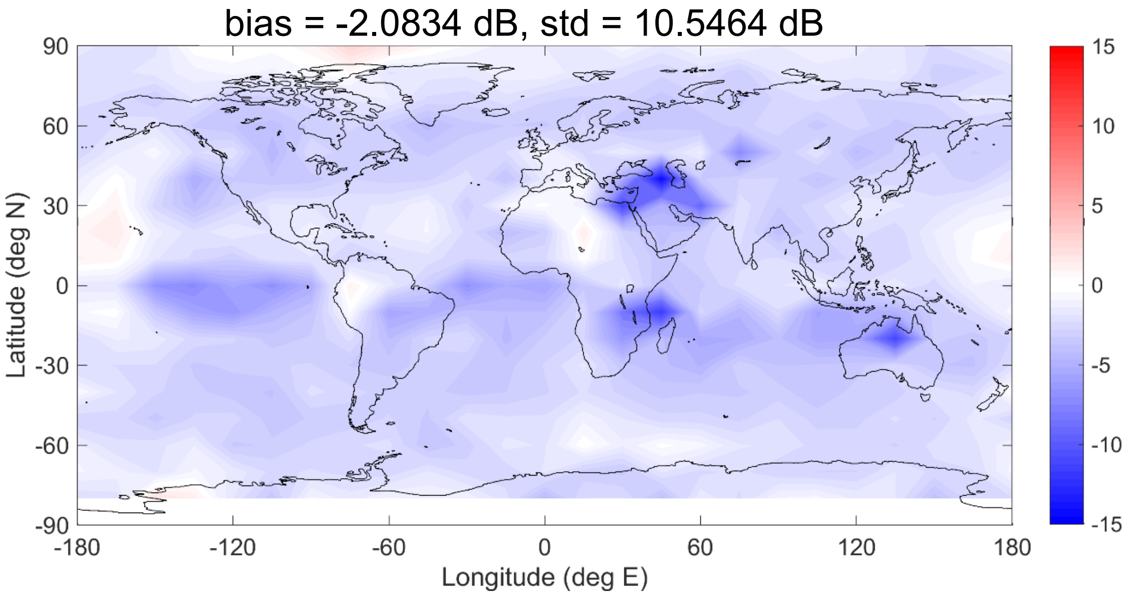
- Data assimilation systems combine a model background and observations given the errors that are inherent in both. However, any biases in either will likely degrade the subsequent analysis and forecasts.
- ECWMF uses an implicit bias correction scheme for many observation types (VarBC), but initially we will use an offline scheme
- Indicators are required to subset the data so that different biases can be accounted for. Selected bias correction indicators:
  - height
  - temperature
  - model dominant hydrometeor type
  - mean radar reflectivity/lidar backscatter ('symmetric')

FG departures based on 12 hour forecasts at TCo639 and 137 model levels using IFS cycle 43r1. Obs. superobbed to model grid.

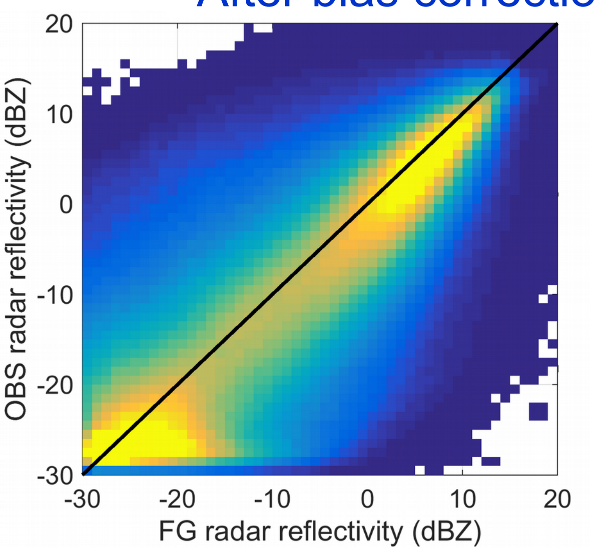
# Bias correction for Cloudsat radar (September 2007)

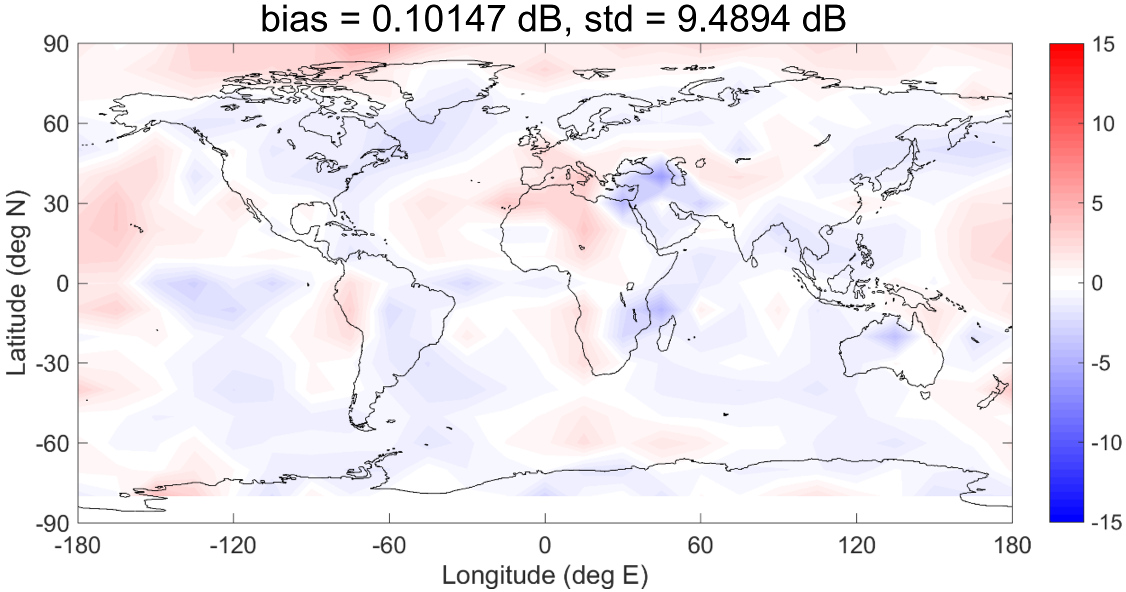
Before bias correction



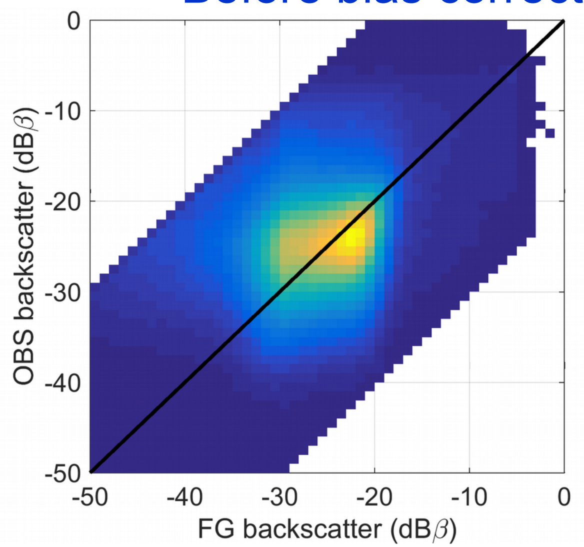

After bias correction

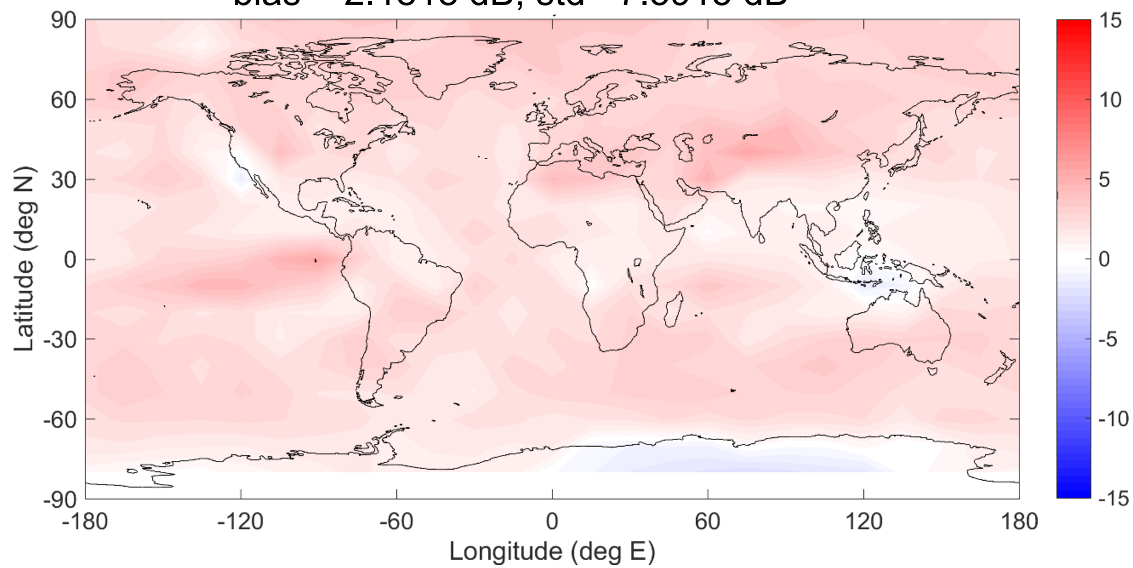


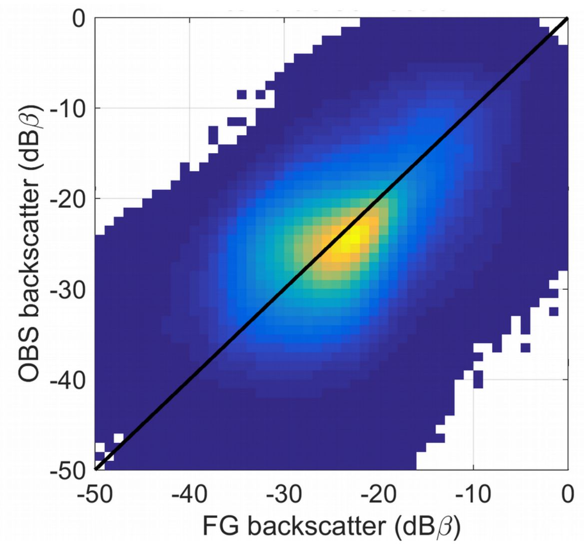
# Bias correction for Calipso lidar (September 2007)

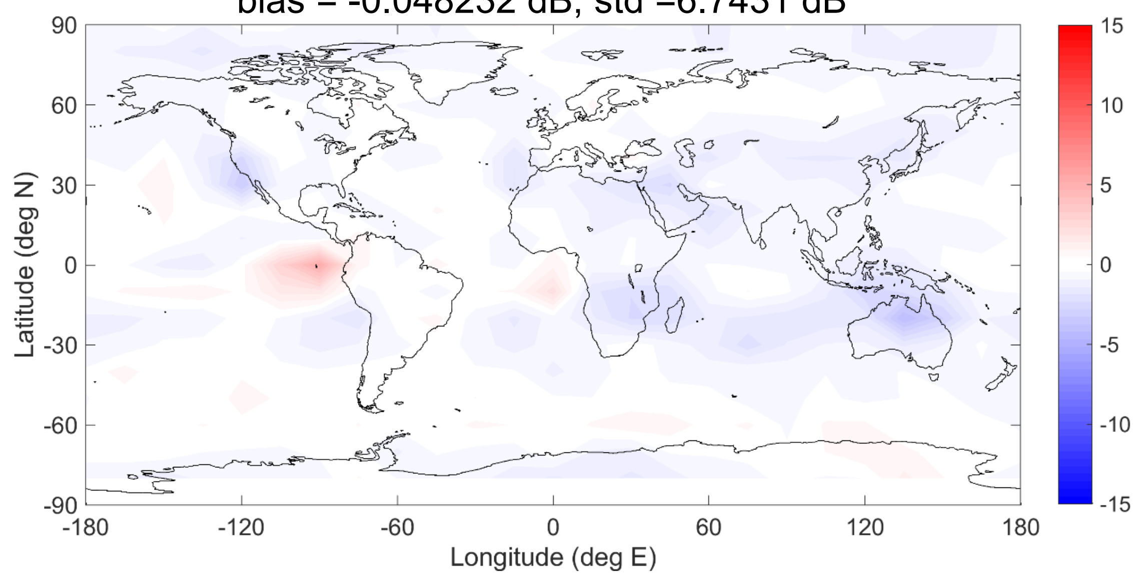
Before bias correction




After bias correction



# Observation error definition

- Observation errors are a crucial component of a data assimilation system as, coupled with the background error, control the weight each obs. is given.
- Often assumed to have no correlation & used for tuning data assim. system
- Typically inferred through a statistical evaluation of FG departures and/or analysis increments
- Selected approach – explicit specification of observation error based on physical understanding because:
  - Owing to the profiling nature of the observations, the true obs. error likely to be highly situation dependent
  - At the time EarthCARE becomes operational, no availability of long history of observations to generate a climatological obs. error covariance matrix
- Under the hypothesis of uncorrelated errors, obs. error is defined as a combination of instrument error, obs. operator error and representativity error:

$$\sigma_{obs}^2 = \sigma_{ins}^2 + \sigma_{oper}^2 + \sigma_{rep}^2$$

Collaboration with Olaf Stiller (DWD)

# A flow-dependent representativity (sampling) error

- Representativity error dominates observation error for profiling observations of cloud. It is also highly scene-dependent.
- Use 'sampling approach' based upon the assumption that:
  - the local variability of measurements along the satellite track is representative of the gridbox variability
  - the spatial variability can be approximated using a climatological correlation

$$\sigma_{RE}^2 = E[(\bar{q}_s - \bar{q}_{2D})^2]$$

$$\sigma_{RE}^2 = \underbrace{\sigma_q^2}_{\substack{\text{Variance of measured}\\ \text{variable within gridbox}}} \underbrace{(\alpha_{[2D,2D]} + \alpha_{[s,s]} - 2\alpha_{[s,2D]})}_{\text{Scaling factor}}$$

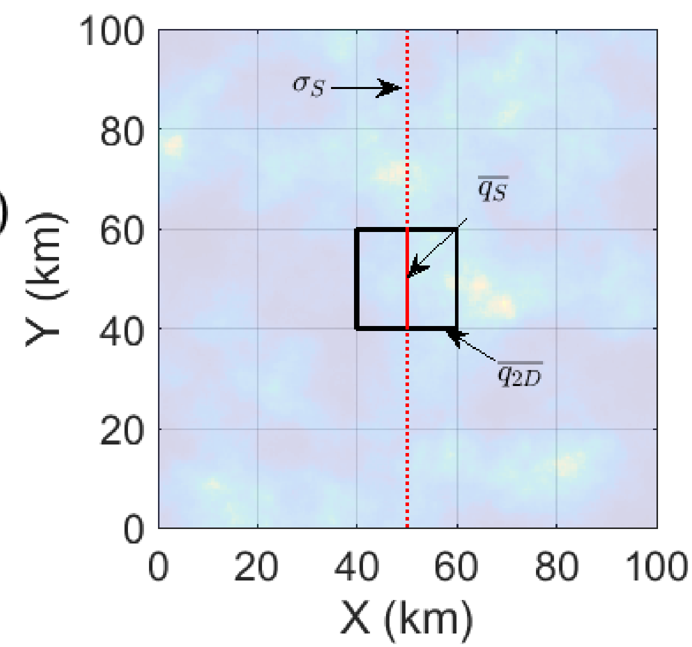
$$\alpha_{[2D,2D]} = \frac{1}{A_{2D}^2}\int_{A_{2D}}\int_{A_{2D}} \rho(\|\mathbf{x_1} - \mathbf{x_2}\|)d\mathbf{x_1}d\mathbf{x_2}$$

$$\alpha_{[s,s]} = \frac{1}{A_s^2}\int_{A_s}\int_{A_s} \rho(\|\mathbf{x_1} - \mathbf{x_2}\|)d\mathbf{x_1}d\mathbf{x_2}$$

$$\alpha_{[s,2D]} = \frac{1}{A_s A_{2D}}\int_{A_s}\int_{A_{2D}} \rho(\|\mathbf{x_1} - \mathbf{x_2}\|)d\mathbf{x_1}d\mathbf{x_2}$$

*Correlation of measured variable within gridbox*

*$A_s$ – area of superob, $A_{2D}$ – area of gridbox*

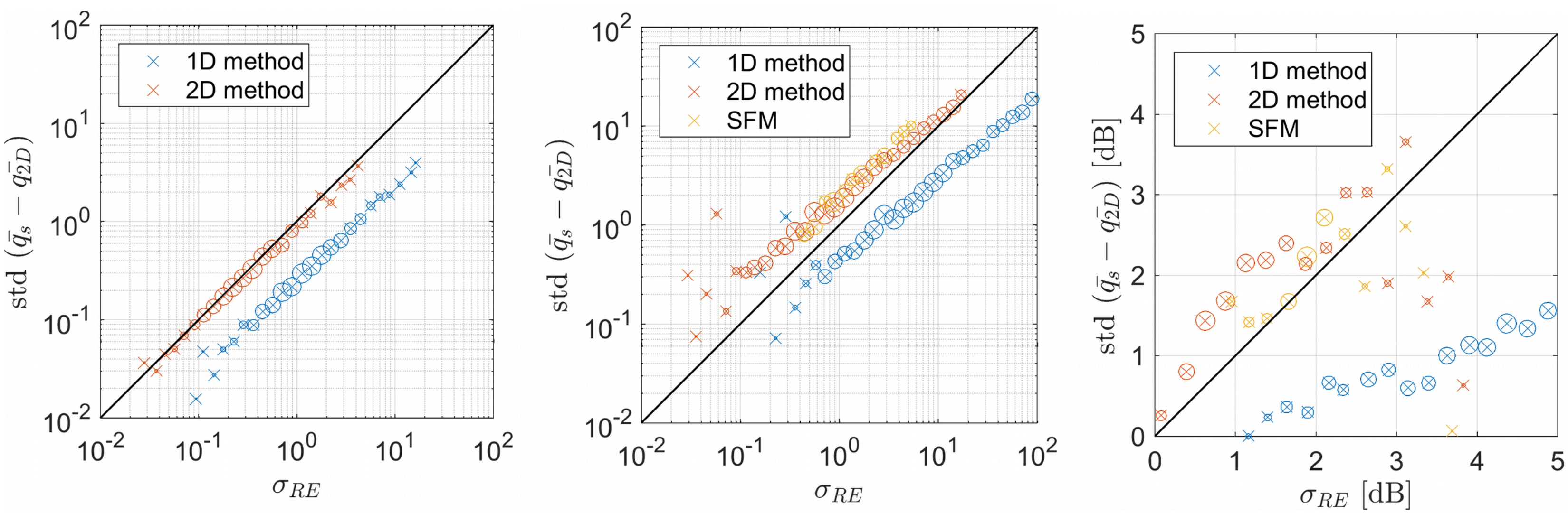



✓ Extendable to horizontal and vertical correlations

# Evaluation of representativity error methods

- Simulate profiling measurements through 'gridboxes' constructed using 2D data.
- Compare three methods:
  - '1D Method' (only accounts for correlation between measurements)
  - '2D method' (fully accounts for all correlations)
  - 'SFM', structure function maximum method (Stiller, 2010)

Synthetic data | MODIS optical depth | Scanning cloud radar




Fielding et al., (in prep)

# Practical implementation of method

1. Using CloudSAT and CALIPSO, gather correlation by parametrizing the observed correlation function, i.e.:

$$\rho(\Delta x) = \exp\left[-\left(\frac{\Delta x}{a}\right)^{b}\right]$$

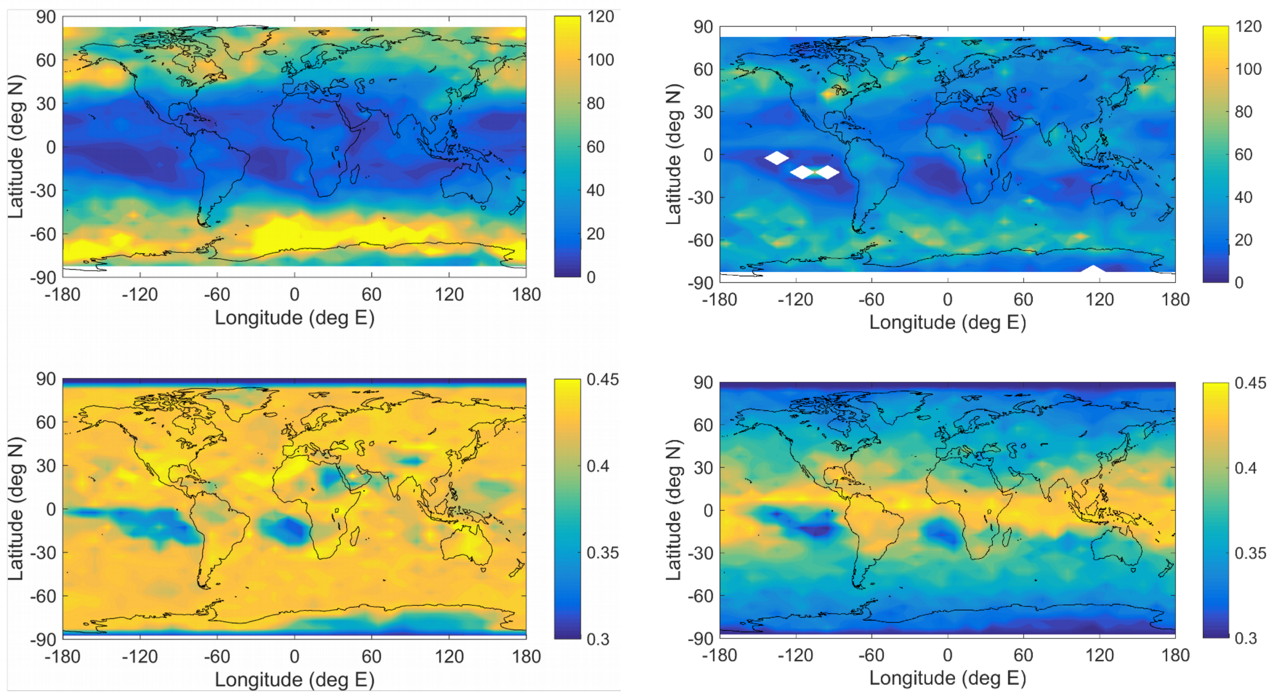
2. Create lookup table of 'scaling factors' with height, latitude and longitude as indicators
3. For each observation use standard deviation along track as estimate of gridbox population standard deviation
4. Multiply corrected sample standard deviation by scaling factor

Low cloud (< 4 km)

High cloud (> 4 km)

Corr. Length (km)

Scaling factor



# Other sources of observation error

Instrument error:

- The random error in the measurement due to noise

- Typically small compared to other errors, but straightforward to estimate

Observation operator error:

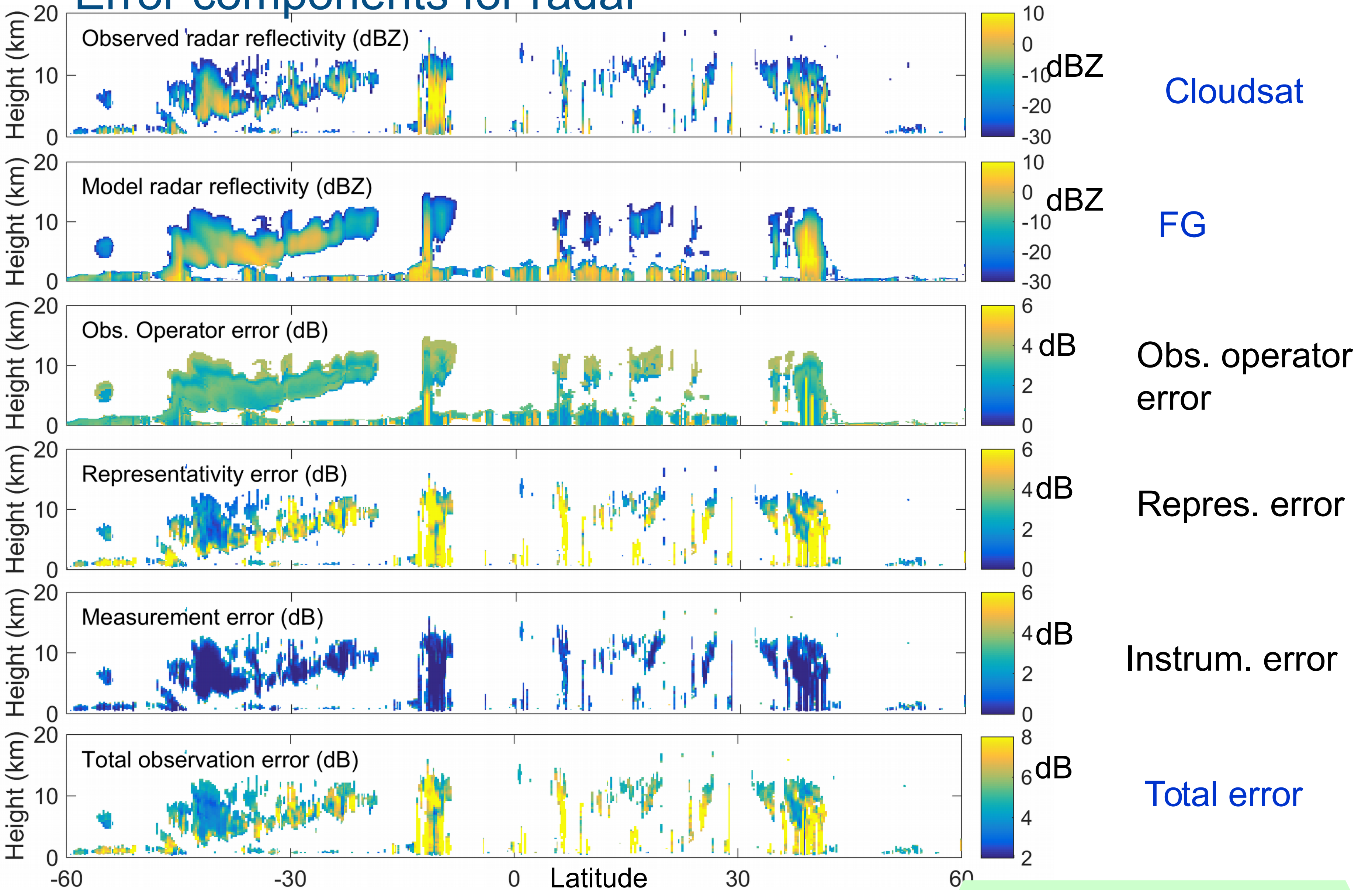
- To convert model hydrometeor content into radar reflectivity/lidar backscatter, many assumptions made with the potential to introduce error in:
  - Radiative transfer of scattering models
  - Hydrometeor shape
  - Particle size distribution
  - Multiple scattering
  - Subgrid assumptions (overlap, inhomogeneity & convective precip.fraction)

- To characterize errors:
  - perturbing parameters with plausible bounds
  - using Monte Carlo simulation – PSD uncertainty is st.dev. of reflectivity/ backscatter given a set of random realisations of PSD variables / densities / particle shapes

- Errors are function of hydrometeor type, LWC and temperature

**PSD – particle size distribution**
**LWC – liquid water content**

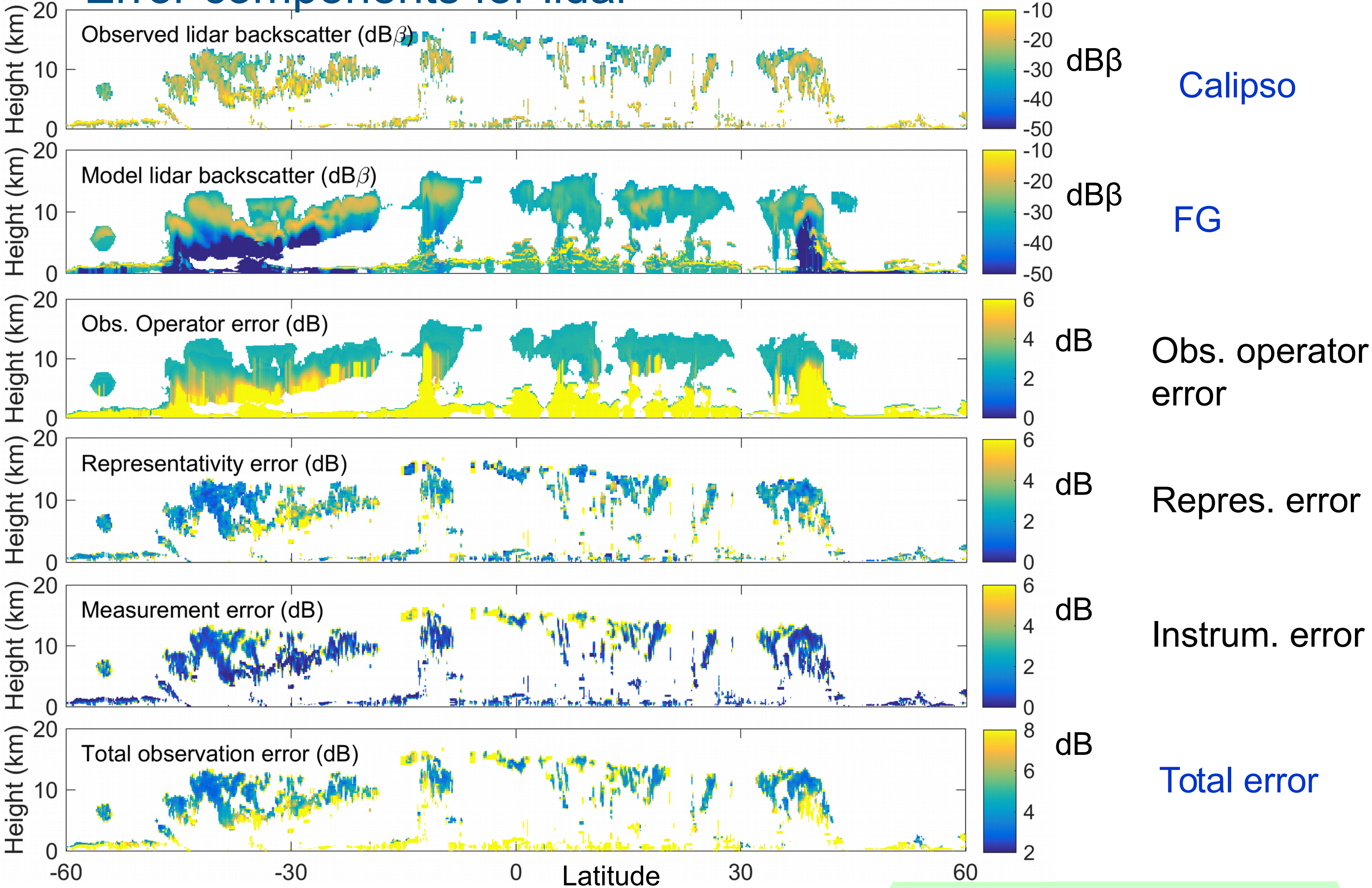
# Error components for radar






Representativity error dominates total error

# Error components for lidar

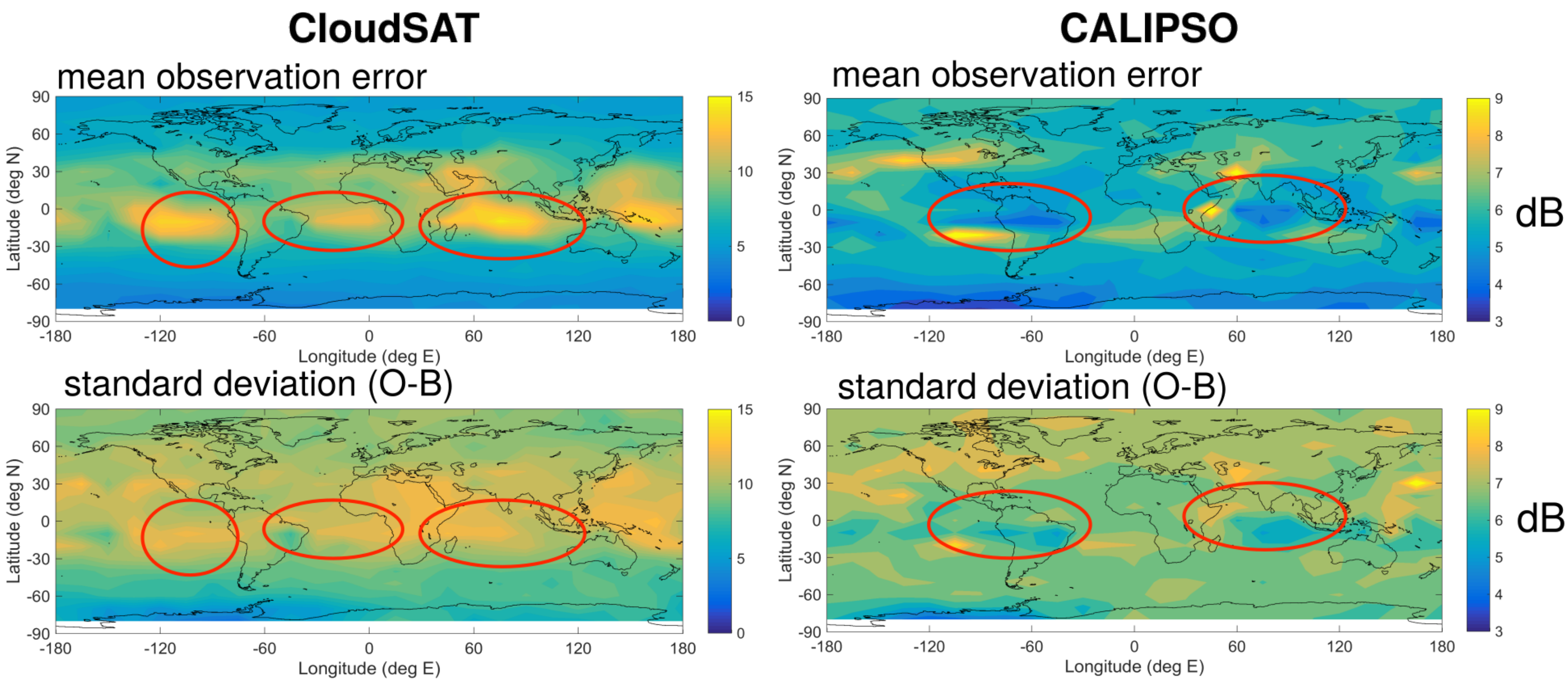



Obs. operator error tends to dominate total error

# Global statistics of observation error

$$\underbrace{< \mathbf{d}, \mathbf{d}^T >}_{\text{FG departures}} = \underbrace{\mathbf{HPH}^T}_{\text{Forecast error}} + \underbrace{\mathbf{R}}_{\text{Observation error}}$$

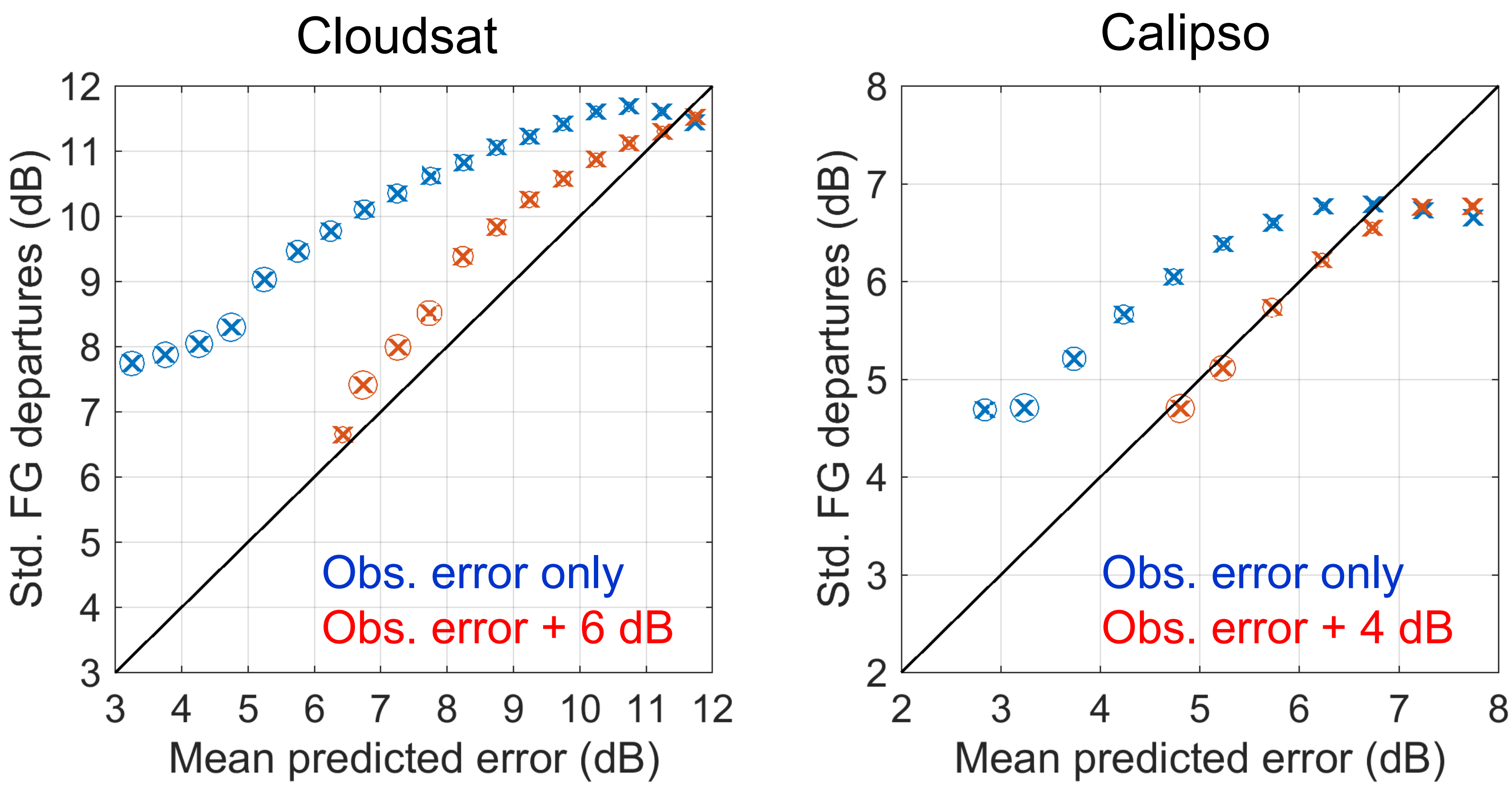
Variance of FG dep. is first order approximation of observation error



# Observation errors are a good predictor of FG departures when a fixed offset is applied

$$< \mathbf{d}, \mathbf{d}^T > = \mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R}$$

*FG departures* *Forecast error* *Observation error*




*Circle size proportional to sample size*

# Summary

- Given overview of current developments towards real-time assimilation of cloud radar and lidar at ECMWF
- Outlined flow-dependent observation error, including a new simple approach to characterising the sampling error
- Explored O-B statistics for CloudSat and Calipso, and their relationship with the expected observation error

# Future work

- Feasibility studies in full 4D-Var system
- Optimising superob size
- Consideration of observation error correlations (particularly important in the vertical)

# Additional slides

# Microphysical assumptions

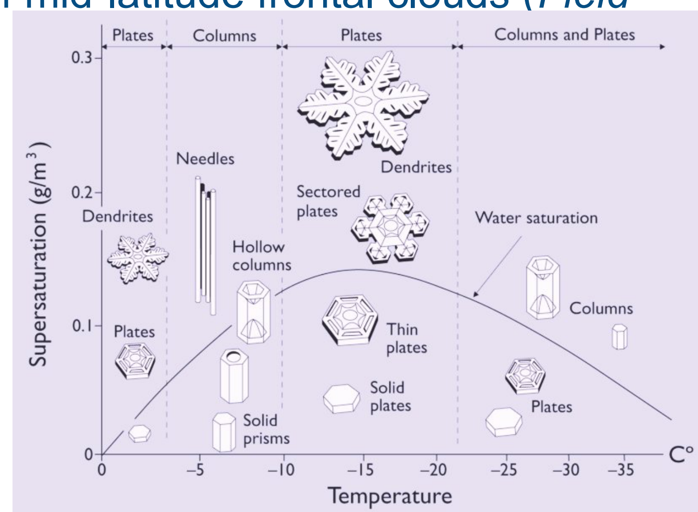
## PSD

Highlighted boxes show developments

- Warm phase
  - Lognormal for cloud (*Miles et al.*, 2000)
  - Exponential distribution with empirical fit for rain (*Abel and Boutle*, 2012)
- Cold phase
  - Temperature dependent, based on observations of mid-latitude frontal clouds (*Field et al.*, 2007) (Cloud ice has constant T = –70 °C)

## Scattering properties

- Radar
  - Ice cloud: 5 bullet rosettes (*Liu*, 2008)
  - Strat. and conv. snow: Aggregates (*Hong*, 2007)
- Lidar
  - Ice: 5 bullet rosettes
  - Snow: Aggregates (*Yang et al.*, 2000)

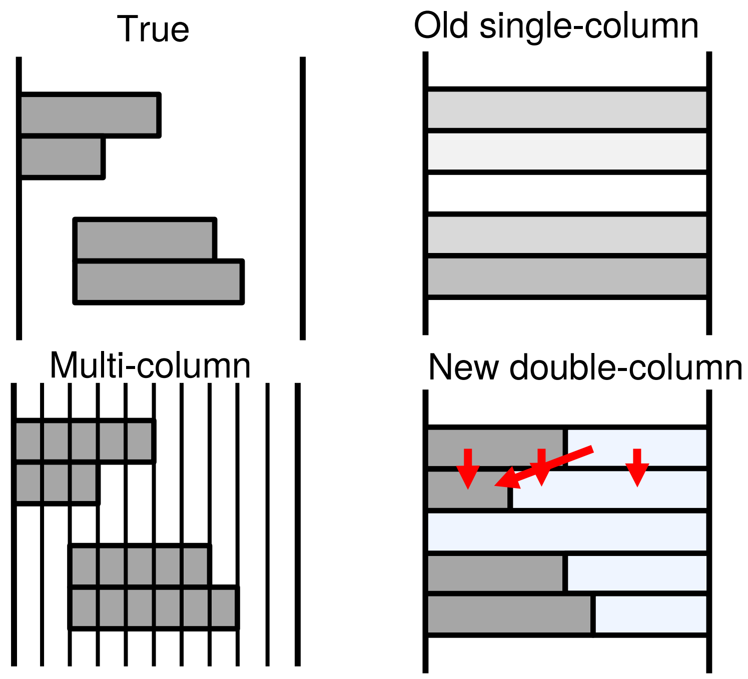


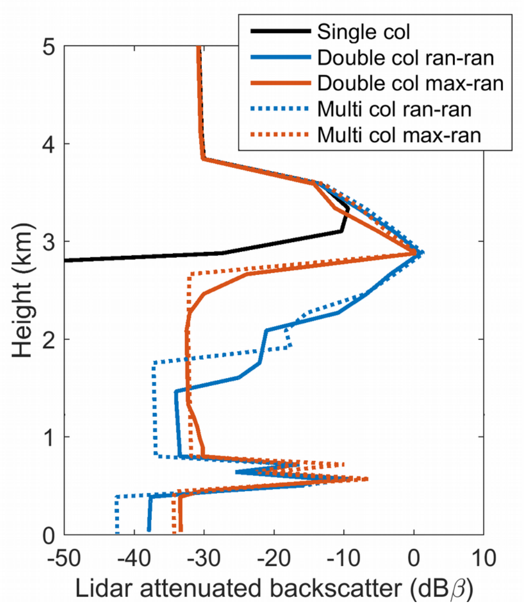

Nakaya snow crystal morphology diagram (Libbrecht 2005)

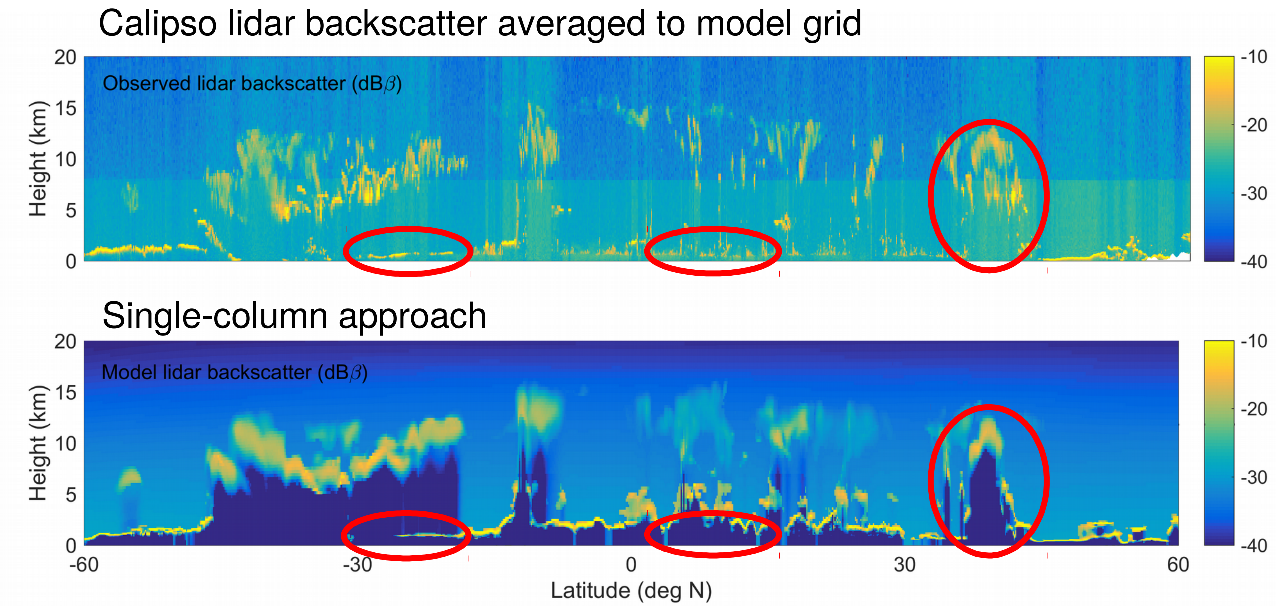

Miles et al., 2000: Cloud droplet size distributions in low-level stratiform clouds
Abel and Boutle, 2012: An improved representation of the raindrop size distribution for single-moment physics
Liu, G., 2008: A database of microwave single-scattering properties for nonspherical ice particles
Hong, G., 2007: Radar backscattering properties of nonspherical ice crystals at 94 GHz
Yang et al., 2000: Parameterization of the scattering and absorption properties of individual ice crystals

# Improving the impact of lidar observations

- Initial testing (e.g., 1D+4DVar) showed radar dominated information content
- New double-column approach mimics accurate but expensive multi-column method

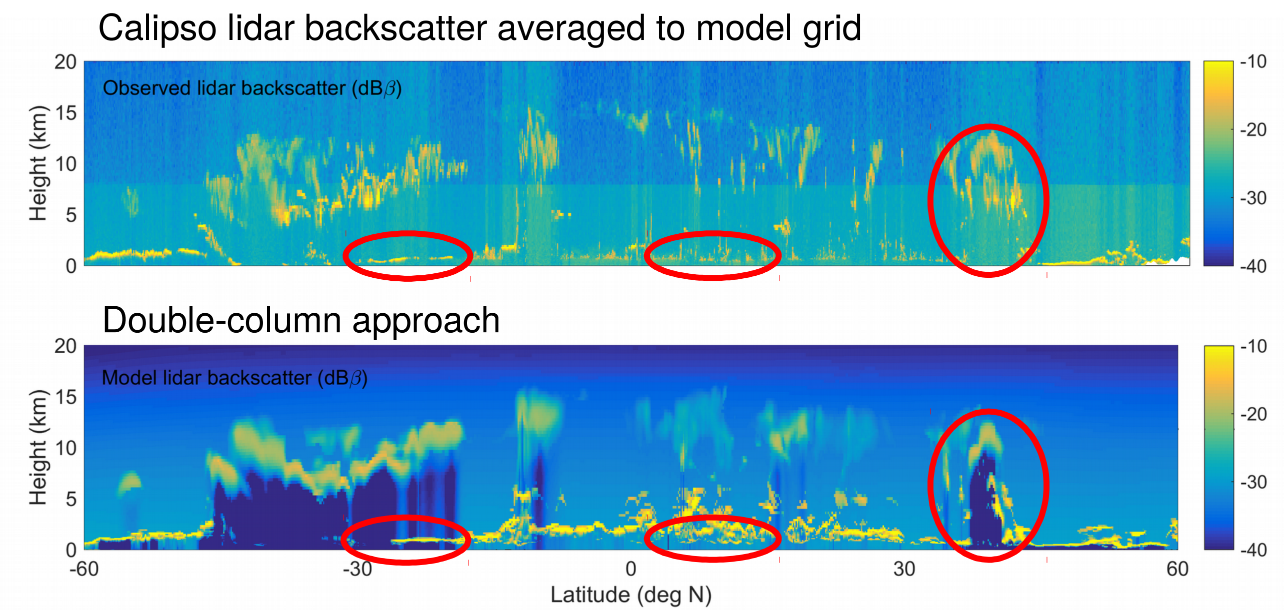




Calipso lidar backscatter averaged to model grid
Observed lidar backscatter (dBβ)
Height (km)
Single-column approach
Model lidar backscatter (dBβ)
Latitude (deg N)


# OLD




Double-column approach will allow more scenes to be assimilated

**NEW**

# Estimation of particle size and shape uncertainty

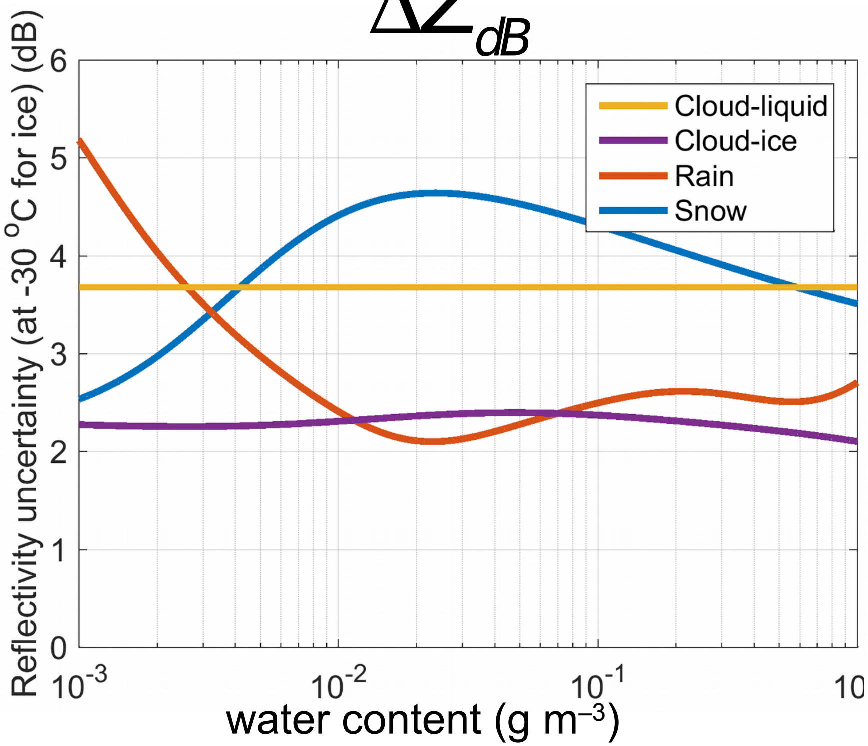
$$Z' = Z\exp(-2\tau) \Rightarrow \ln(Z') = \ln(Z) - 2\tau$$

$$\Delta Z'_{dB} = 4.343\sqrt{(\Delta \ln Z)^2 + 4(\Delta\tau)^2}$$

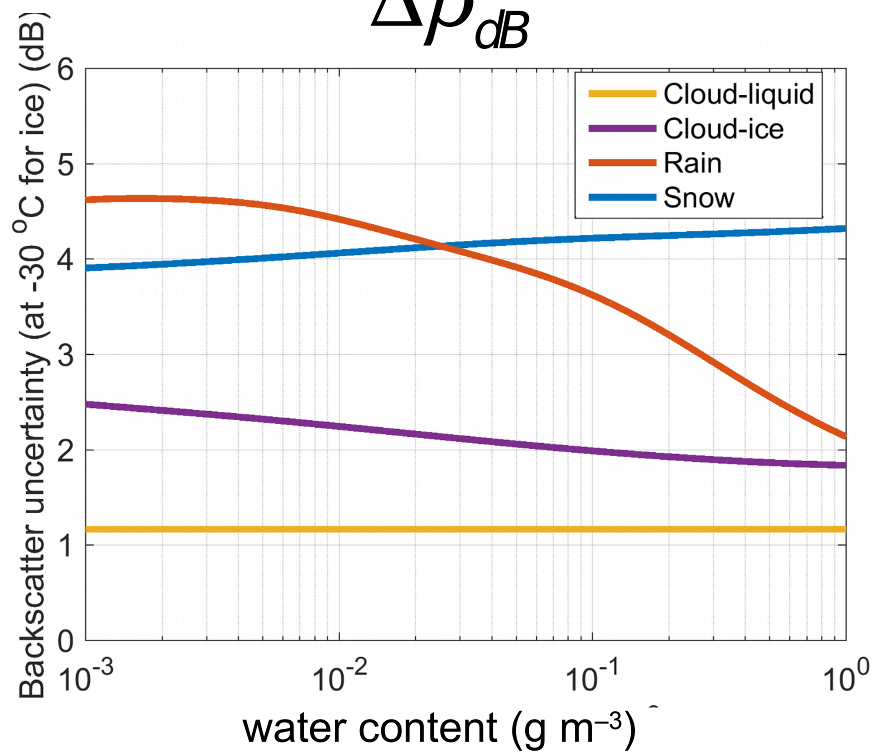
Uncertainty in backscatter

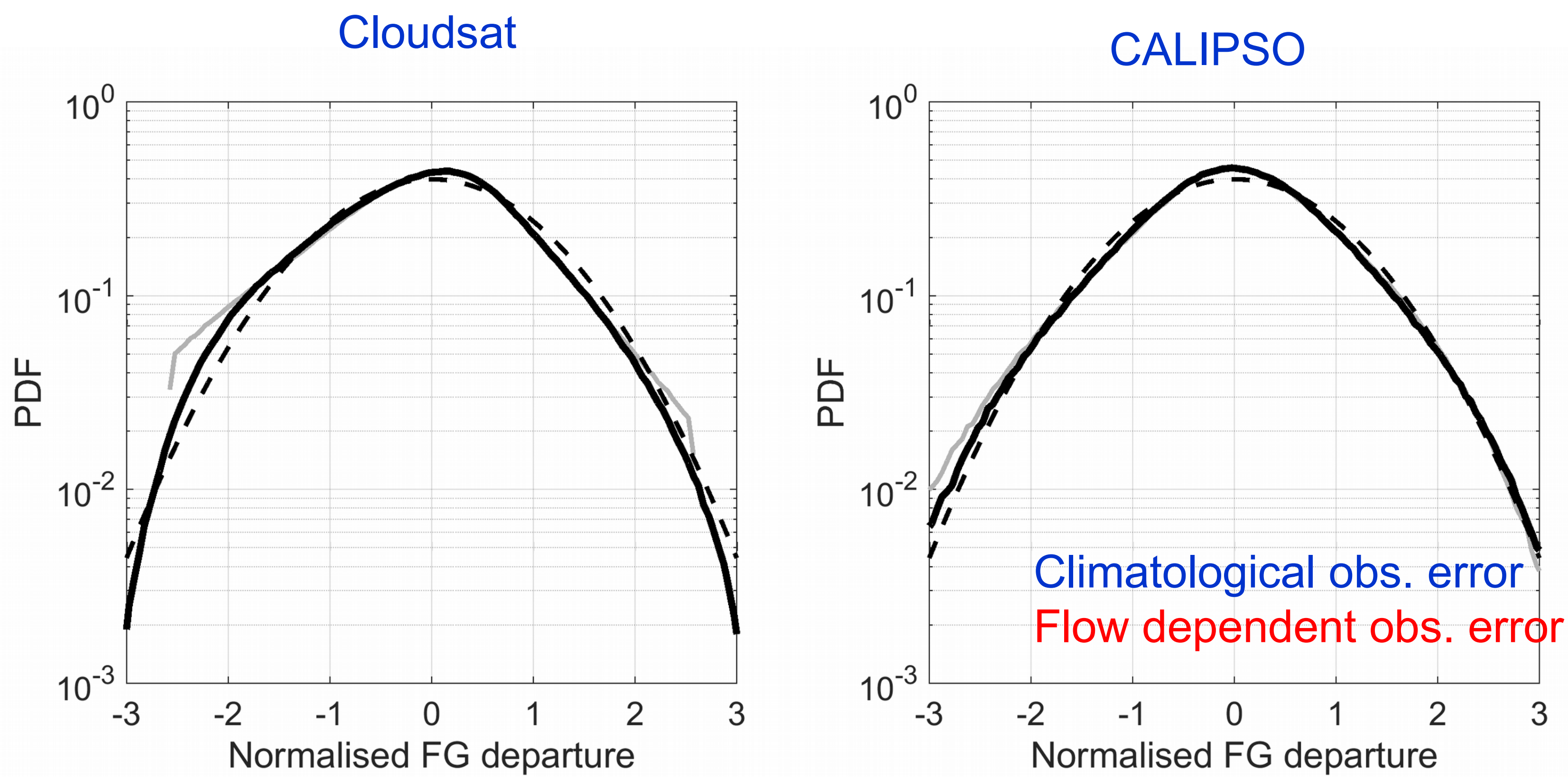
Uncertainty in attenuation

$\Delta Z_{dB}$




$\Delta\beta_{dB}$




Cloudsat
CALIPSO
PDF
$10^{0}$
$10^{-1}$
$10^{-2}$
$10^{-3}$
-3
-2
-1
0
1
2
3
Normalised FG departure
Climatological obs. error
Flow dependent obs. error