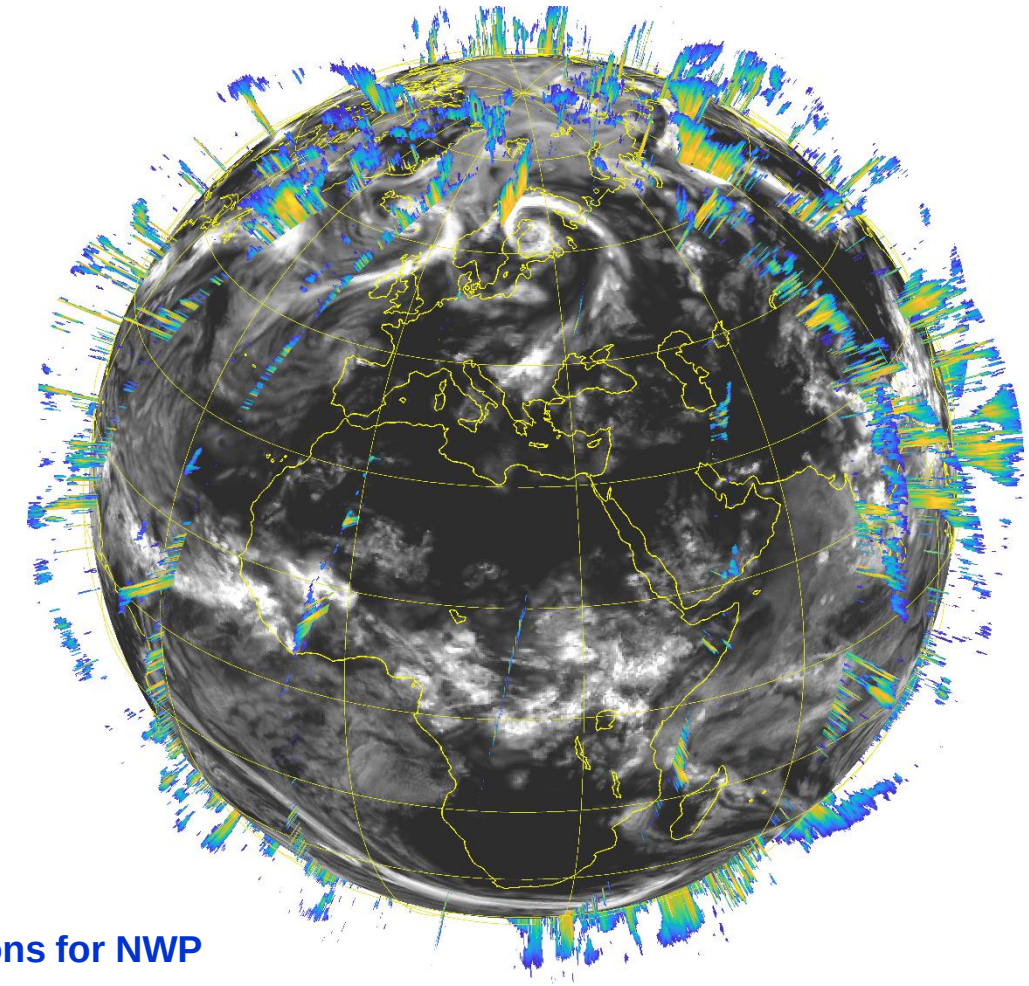
Cloud radar and lidar assimilation at ECMWF

Marta Janisková, Mark Fielding

ECMWF



Thanks to: P. Lopez, R. Hogan, A. Geer, S. English
P. Lean, G. Radnoti, T. Král, D. Vasiljević, M. Crepulja

4th workshop on assimilating satellite cloud and precipitation observations for NWP
4 February 2020

Introduction

- New possibilities for model improvement to be explored through assimilation of data related to clouds from active and passive sensors.

- Observations providing 3D-information of clouds from the space-borne active instruments on board of CloudSat & CALIPSO are already available and new ones, such as EarthCARE should appear in the near future.

- Despite the major influence of clouds and precipitation on the atmospheric water and energy balance, most cloud-affected observations are discarded in current data assimilation systems mainly because of:
  - discontinuous nature (in time and space) of clouds and precipitation
  - need to use linearized versions of these nonlinear processes (for variational assimilation)
  - spatial representativeness of satellite observations, especially from active instruments
  - non-Gaussian error characteristics of the cloud models

Overview of previous studies on assimilating space-borne radar & lidar observations

- Potential of space-borne cloud radar & lidar observations for global (NWP and climate) model improvements and for constraining cloud physical state through data assimilation:

  → *clearly demonstrated by the ESA funded projects:*
  - *QuARL – Quantitative Assessment of operational value of radar & lidar*
  - *STSE Study – EarthCARE Assimilation*
  - *GSP Study – Operational assimilation of cloud radar & lidar*

- Assimilation studies for cloud profiling observations from space-borne active instruments on board of CloudSat & CALIPSO to prepare for the EarthCARE mission carried at ECMWF

  (Janisková et al 2012, Janisková 2015, Janisková & Fielding 2018)

- In the first studies *(QuARL, STSE Study),* impact of the new observations on 4D-Var analyses and subsequent forecasts studied using a 1D+4D-Var technique:
  - *Information on T & q retrieved from 1D-Var of cloud radar / lidar data and used as pseudo-observations in 4D-Var can lead to improve initial conditions & better forecast*

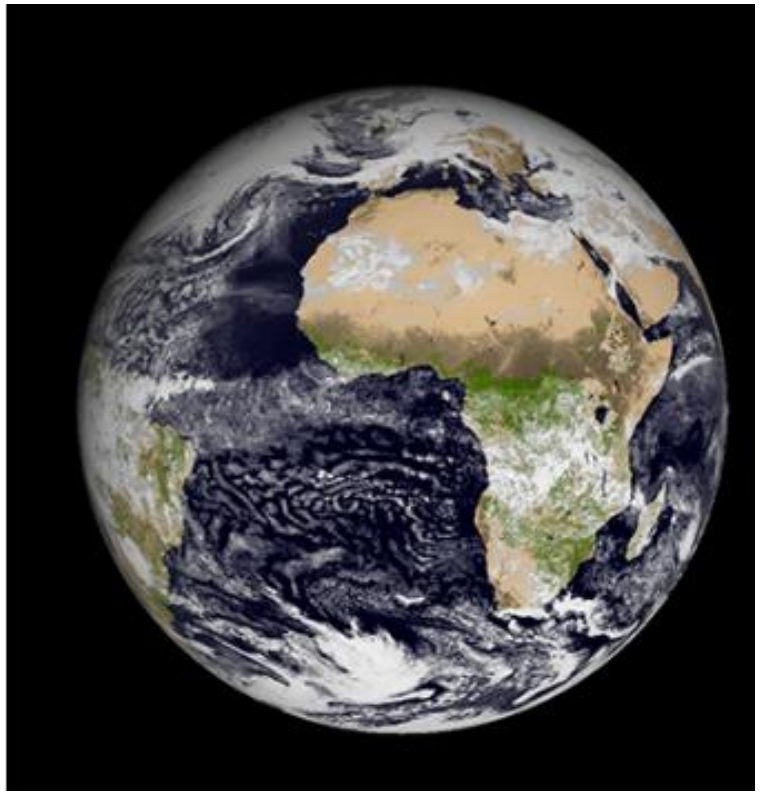
- The last finalised study, *GSP Study 'Operational Assimilation of Space-borne Radar and Lidar Cloud Profiles for NWP*' prepared the ECMWF system for:
  - *Direct 4D-Var assimilation and monitoring of these observations*

Requirements for data assimilation

- To succeed to assimilate new type of observations, especially cloud profiling, requires:
  - reasonable representation of the physical processes related to the observations
  - observation operator being able to provide realistic model equivalents to the observations
  - appropriate screening of observations
  - removal of systematic biases via a bias correction scheme
  - characterizing the components of observation error including representativity issues
  - inclusion of observations into automatic monitoring system

Accuracy of the model to represent clouds

EUMETSAT's Meteosat Second Generation (MSG-4) on 10 February 2019 at ~ 12h UTC



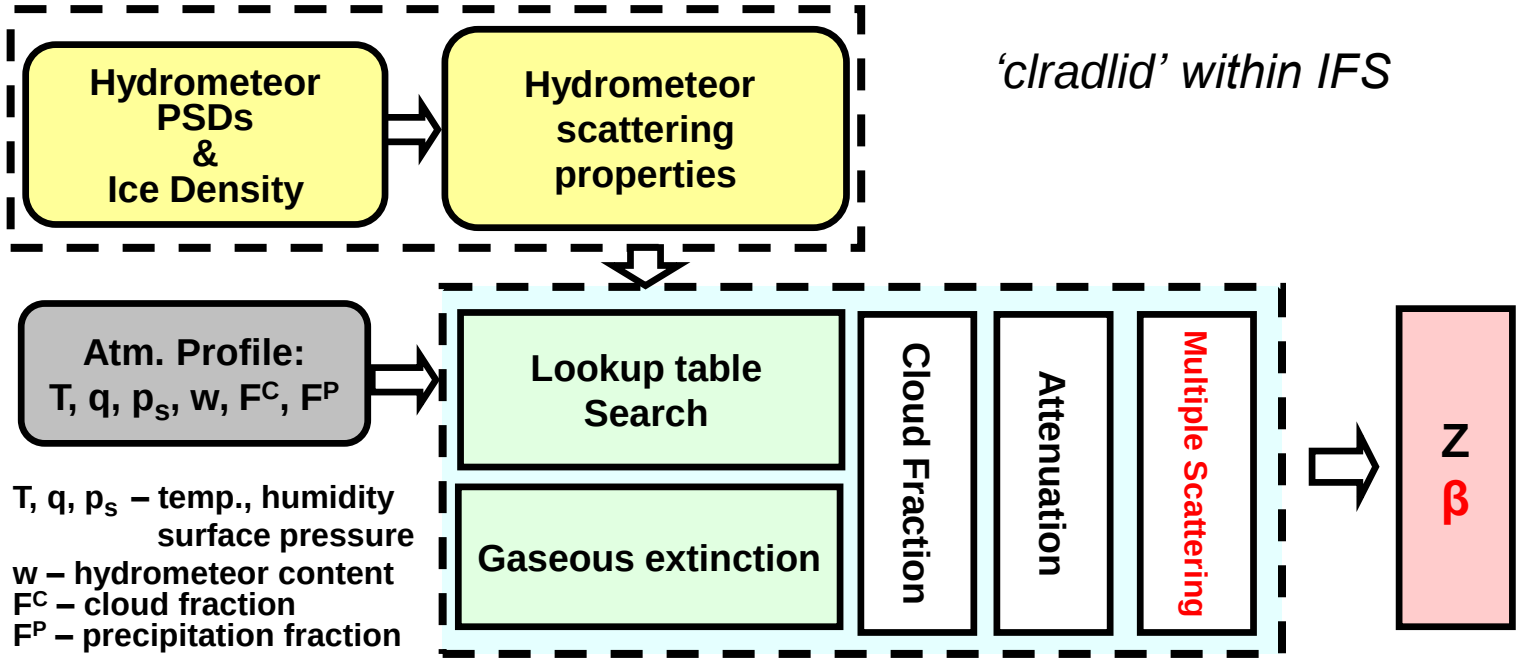
Pseudo-image generated from an 84-hour 9-km resolution ECMWF forecast initialised from operational analysis

The accuracy of both IFS analyses and forecasts has improved over the years with increasing skill further into the forecast.

Courtesy of P. Lopez (2019)

Observation operator to provide realistic model equivalent to observations

- The observation operator simulates observations within a model thus converting model state variables (e.g., T, q, …) into 'model equivalents' (e.g., Z, β)
- The departure between model equivalent and the observation is used by the data assimilation scheme to increment the model variables to bring the state closer to the truth.
- For global NWP assimilation, the operator must be as accurate as possible, while being cost efficient, consistent and differentiable.



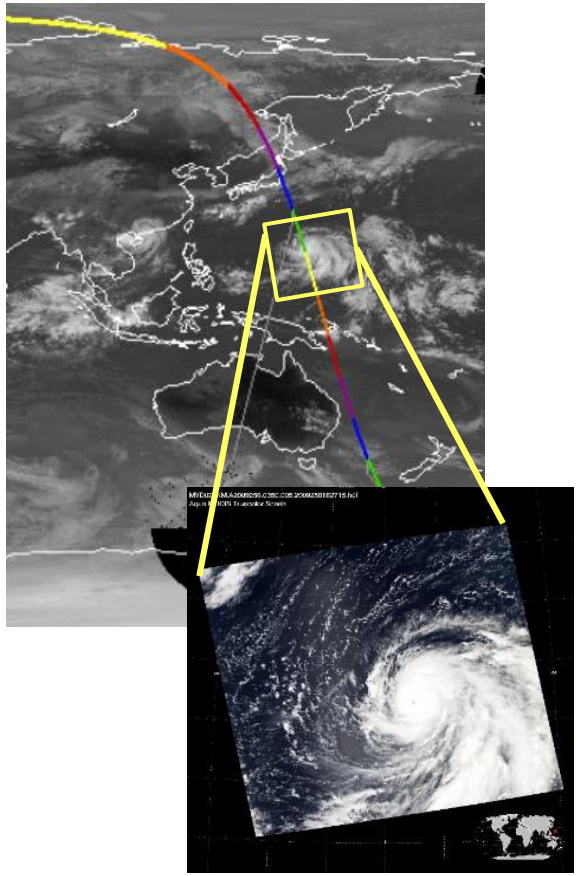
Recent modifications to obs. operator:

- Updates to microphysics and scattering assumptions
- New 'Double column' overlap approach
- Parameterized look up table
- Sensitivity and multiple scattering considerations
- Code optimization

Easily adaptable to different sensors

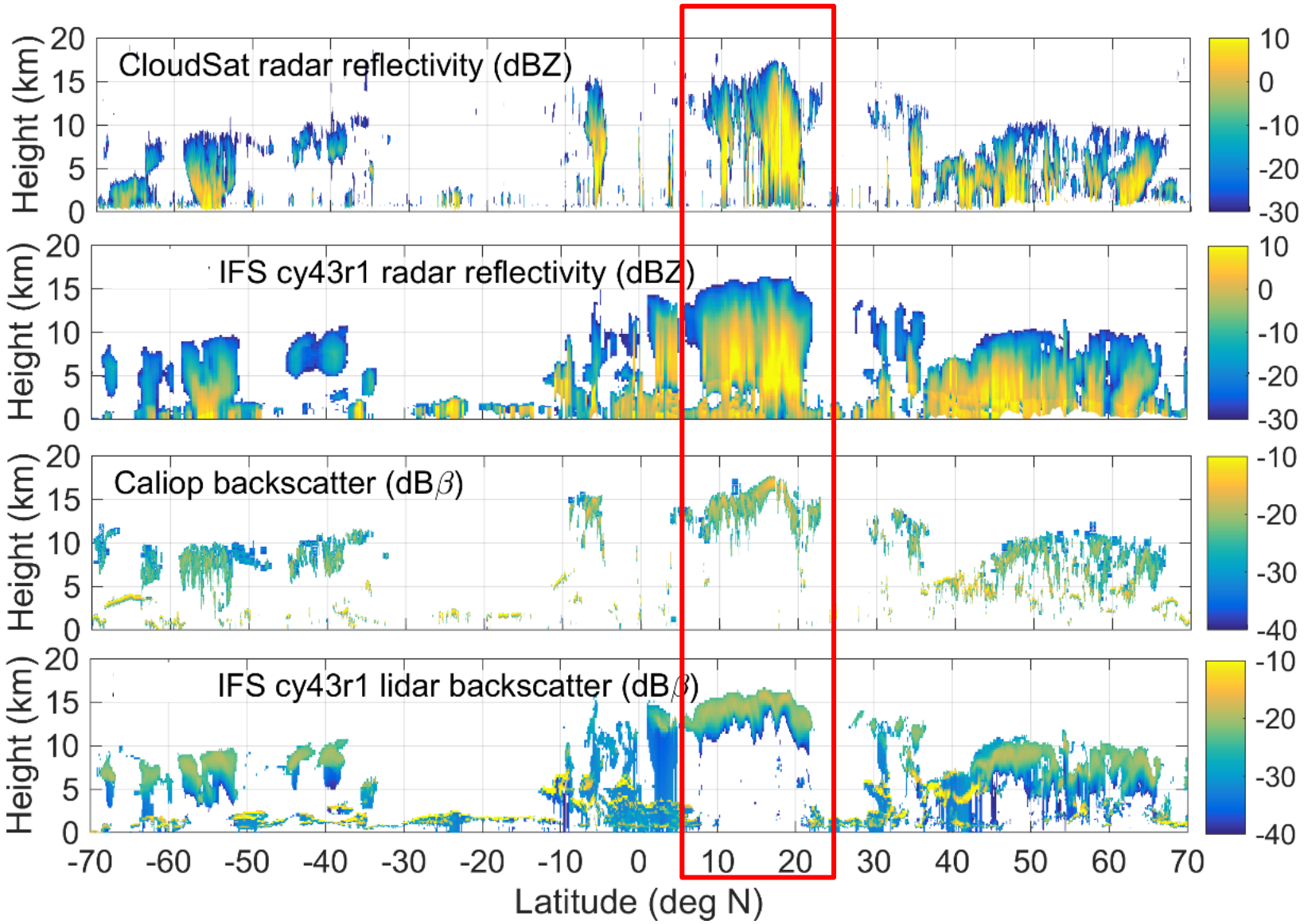
Observation operator for cloud radar and lidar – performance

Typhoon Choi-wan: 15 Sept. 2009



Radar

Lidar



Good agreement of the model equivalents of cloud radar reflectivity and lidar backscatter with observations.

Initial screening and quality control

- Screening and quality control aim to provide a balance between including as much information from observations as possible whilst preventing 'bad' observations from degrading the analysis/forecast. This is important to:
  - reduce the volume of observations used by the assimilation system
  - ensure measurements are not assimilated:
    - where they are of poor quality
    - where the forward model is not capable of representing the observations.
    - where there is an excessive non-linear relationship between perturbations in the control variables and the corresponding simulated parameters

- Observation points must pass initial screening:

Indicator	Min	Max	Reason
Height above surface (km)	1	20	Avoid ground clutter (min) & spurious signals (max)
CF_{obs}, CF_{IFS}	0.2	1.0	Non-linearity & repressentativity issues
dBZ_{obs}, dBZ_{IFS}	-30	20	Physical bounds for radar
$dB\beta_{obs}$, $dB\beta_{IFS}$	-50	10	Physical bounds for lidar
FG departures	-20	20	Remove large departures
dBZ_{int}	0.0	41.3	Radar MS not modelled by obs. operator

CF – cloud fraction int – integrated quantity

Removal of systematic biases via a 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')

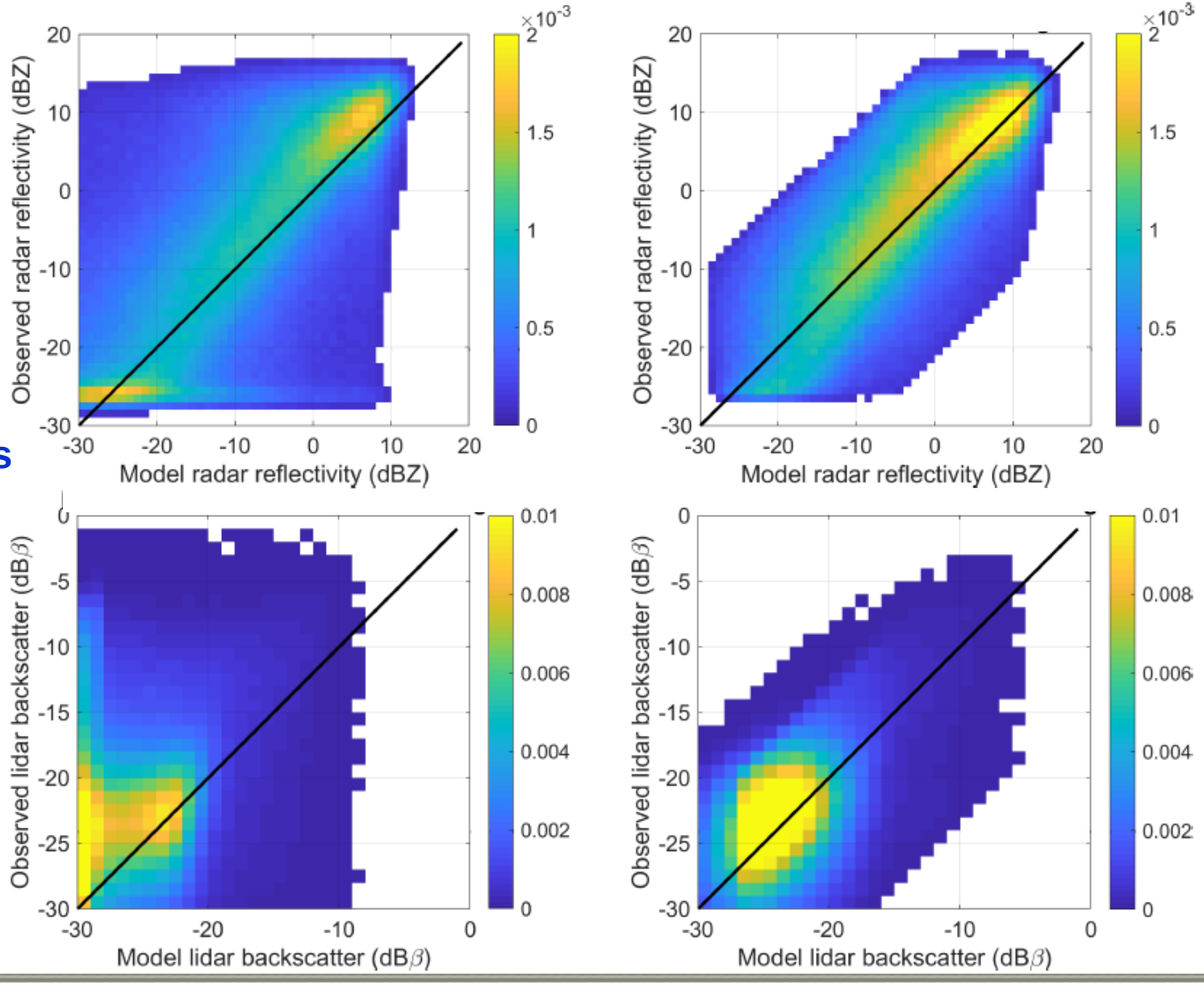
Screening and bias correction applied to CloudSat radar and CALIPSO lidar observations

Radar

Before bias correction

After bias correction

Lidar



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

Better agreement between the model first guess and corresponding observations after applying bias correction and screening.

Characterizing observation error

- 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
- Can be estimated directly or inferred through a statistical evaluation of FG departures and/or analysis increments

- Selected approach – defining the observation error explicitly 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 becoming 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:

$$s_{obs}^2 = s_{ins}^2 + s_{oper}^2 + s_{rep}^2$$

Components of observation error (1)

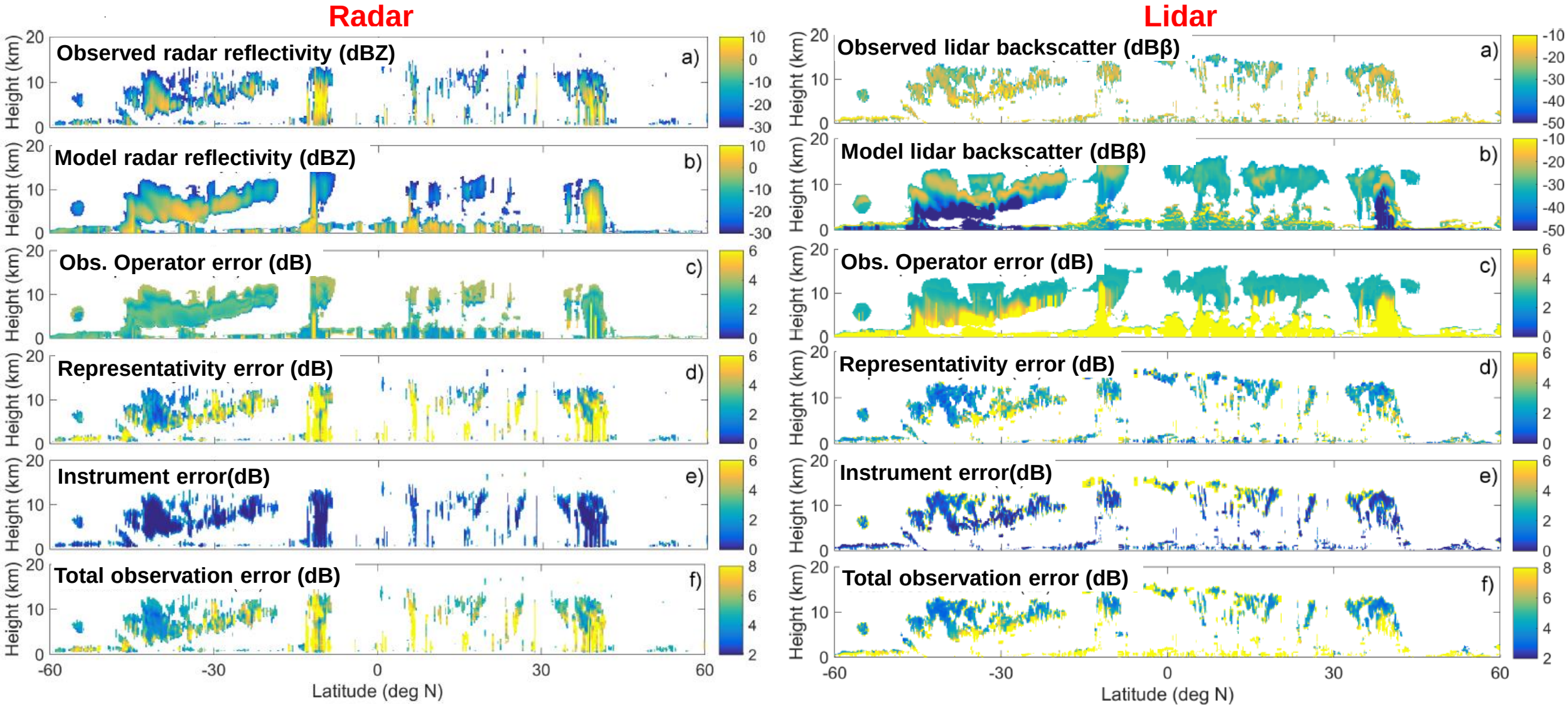
- **Instrument error:**
  - the random error in the measurement due to noise
  - if not directly provided as Level 1B product, it needs to be specified based on known instrument noise characteristics.

- **Observation operator error:**
  - based on evaluation of uncertainty in the microphysical assumptions
  - errors estimated by Monte Carlo simulation – uncertainty is standard deviation of reflectivity/backscatter given a set of random realisations of PSD variables / densities / particle shapes
  - errors are function of hydrometeor type, LWC and temperature
  - including attenuation uncertainty – errors increased as signal attenuated

- **Representativity error:**
  - the expected error due to mismatch between the model and the observational spatial and/or temporal scales
  - 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

Fielding, M. D., and O. Stiller, 2019:. Characterizing the representativity error of cloud profiling observations for data assimilation. J. Geophys. Res.:Atmospheres, 124, 4086– 4103.

Components of observation error (2)



Representativity error dominates total error.

Observation operator error tends to dominate total error.

Experimentation setup for data assimilation

- 4D-Var experimentation using a horizontal resolution of TCo639 spectral truncation (corresponding to ~ 18 km on a cubic octahedral grid) and 137 vertical levels:
  - Resolution: outer loop ~ 18 km, inner loops at T95 (~200 km), T159 (~125 km), T255 (~80 km)
  - Period: 1 August 2007 – 31 October 2007
  - Observation area: Global
  - Length of assimilation window: 12 hours
  - Observation errors follow 'inventory approach' *(Fielding and Stiller, 2019) – increased by 1.5 or 2* (**1err, 1.5err, 2err**) *Account for vertical error correlation*

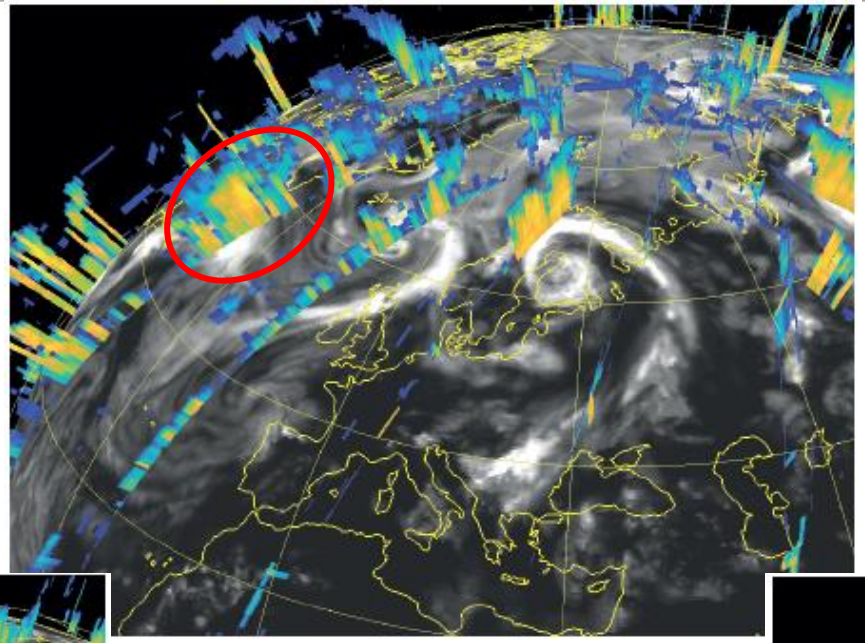
- Observations averaged to the resolution of ~ 36 km or ~72 km (**Ha**): *Account for horizontal error correlation*
  - cloud radar reflectivity (at 94 GHz, CloudSat)
  - cloud lidar backscatter (at 532 nm, CALIPSO)

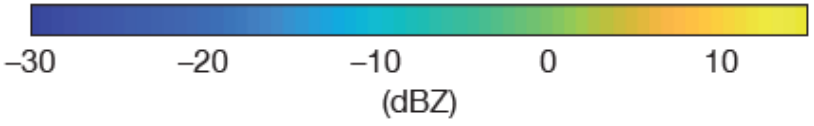
- Performed experiments:
  - **REF** – reference run, using all regularly assimilated observations
  - **EXP** – experimental run including cloud radar & lidar observations on top of other regularly assimilated observations (**RADLID**)

Fielding, M. D., and O. Stiller, 2019:. Characterizing the representativity error of cloud profiling observations for data assimilation. JGR, 124, 4086– 4103.

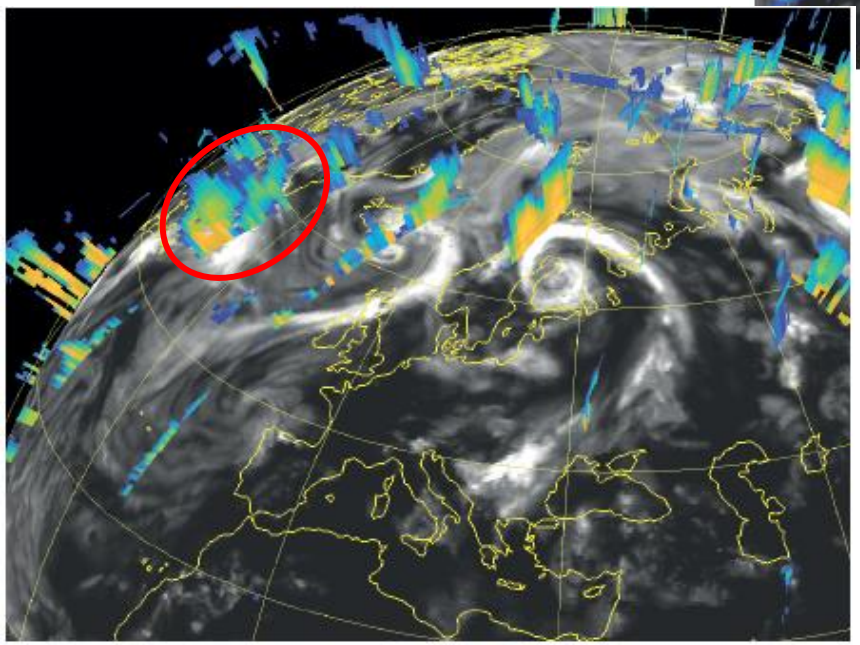
Feasibility of direct assimilation using CloudSat and CALIPSO (1)

CloudSat radar



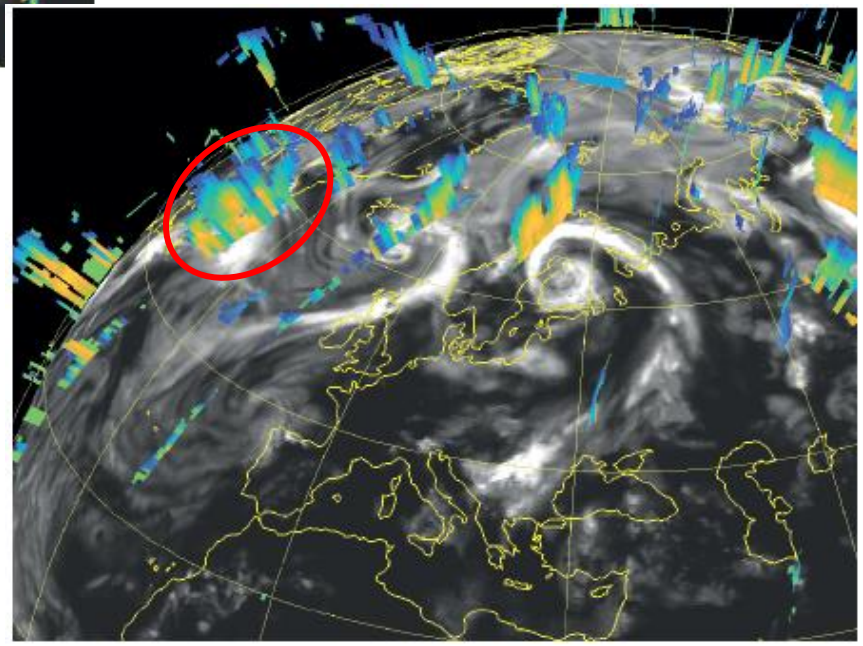


- Experiments assimilating Cloudsat radar reflectivity (94 GHz) and CALIPSO lidar backscatter (532 nm)



First guess (FG)

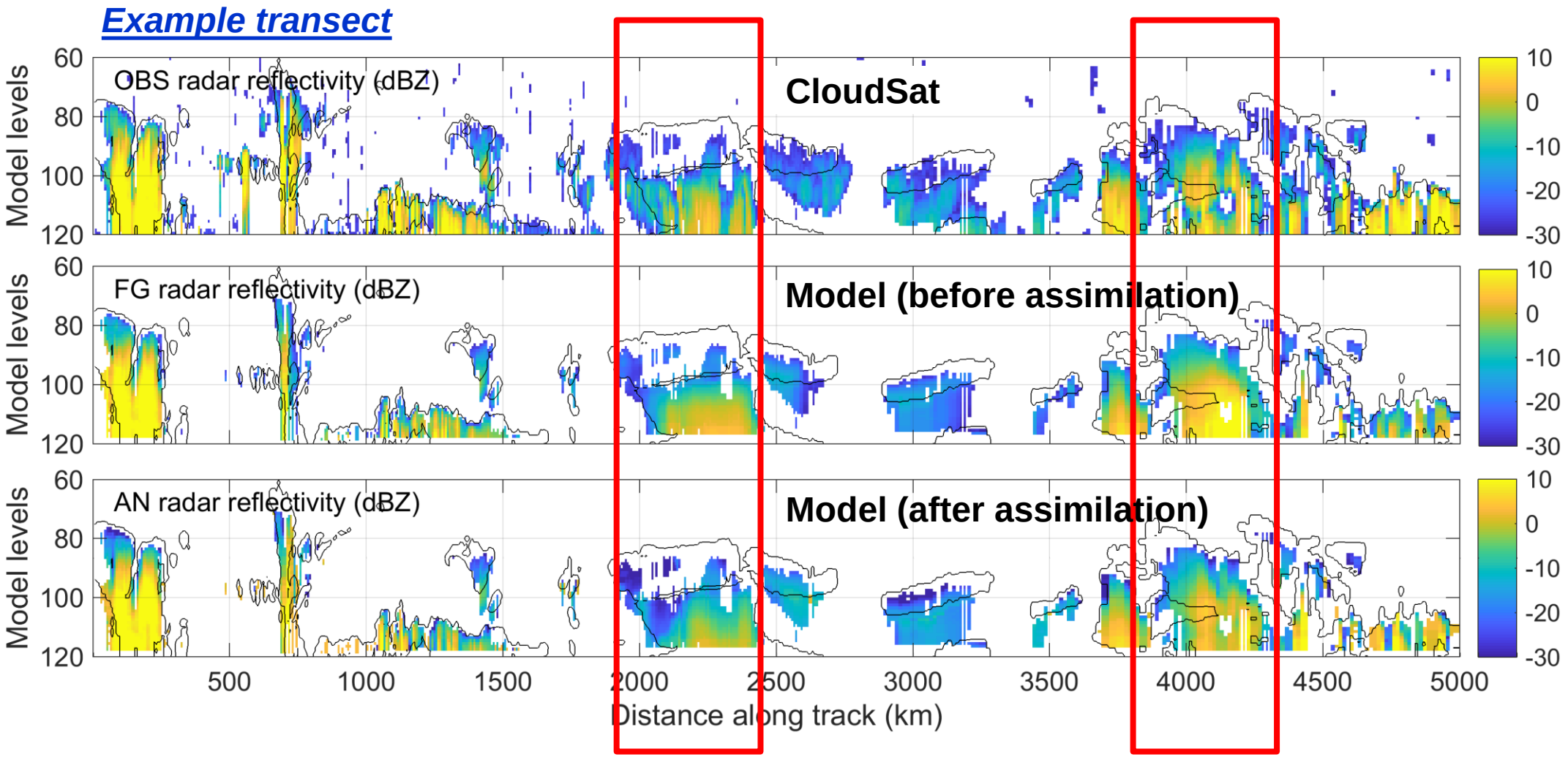
Analysis (AN)



Situation: 20070731 21:00 UTC – 20070801 09:00 UTC

Feasibility of direct assimilation using CloudSat and CALIPSO (2)

- Experiments assimilating Cloudsat radar reflectivity (94 GHz) and CALIPSO lidar backscatter (532 nm)



Situation: 20070731 21:00 UTC – 20070801 09:00 UTC

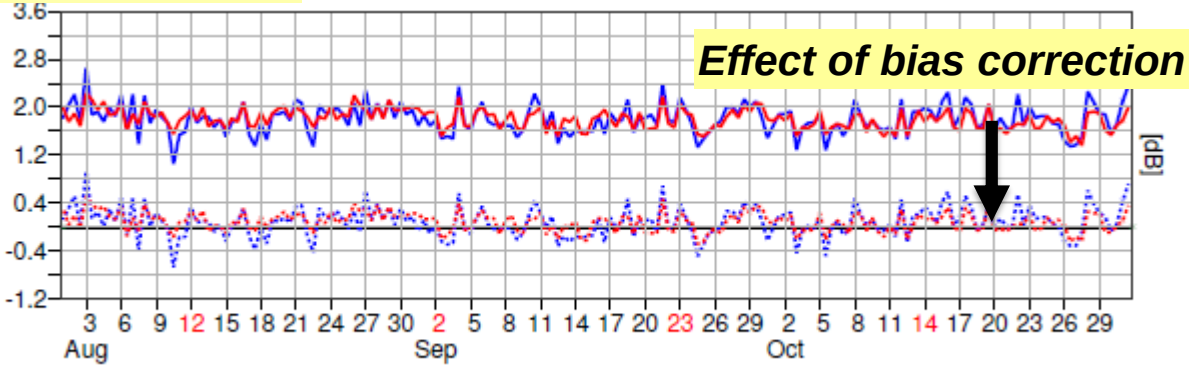
Comparison of experiments against own observations

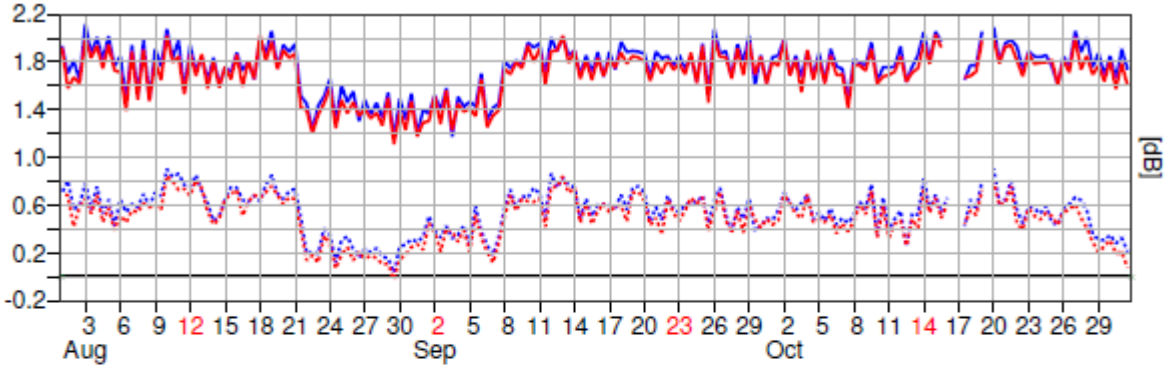
CloudSat radar reflectivity (dB)

CALIPSO lidar backscatter (dB)

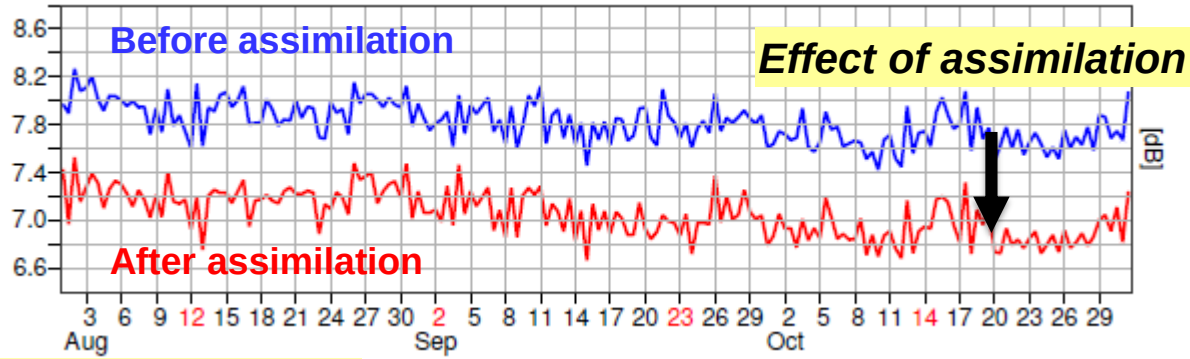
RADLID_2err
double obs. error

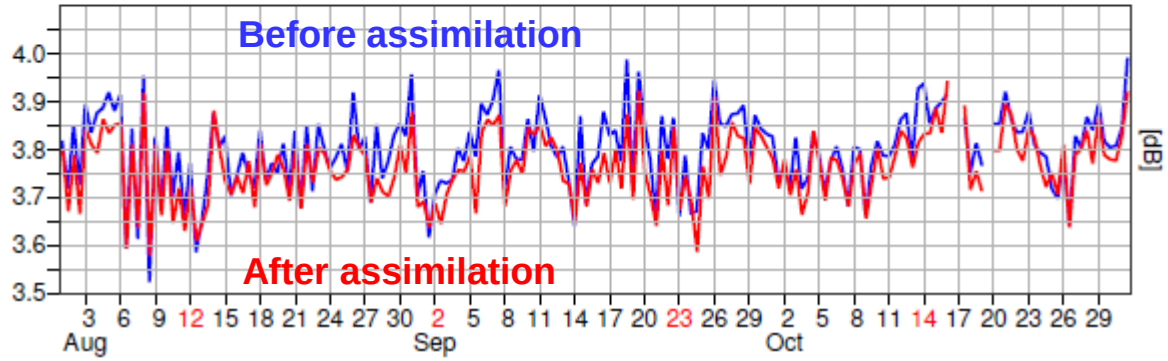
OBS–FG OBS–AN OB –FG (bcor) OBS–AN (bcor)





stdv (OBS–FG) stdv (OBS–AN)





RADLID_2err
double obs. error

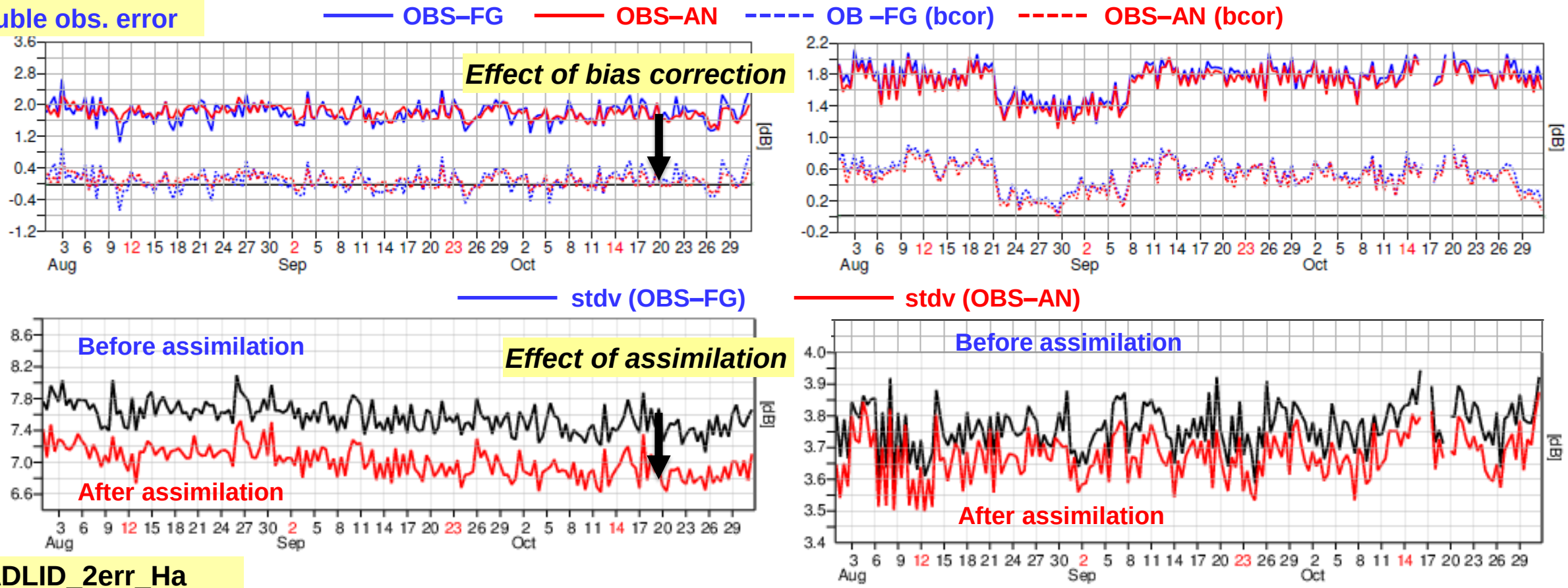
3-month period: 1 August – 31 October 2007

Comparison of experiments against own observations

CloudSat radar reflectivity (dB)

CALIPSO lidar backscatter (dB)

RADLID_2err
double obs. error



RADLID_2err_Ha
double obs. error & horizontal averaging

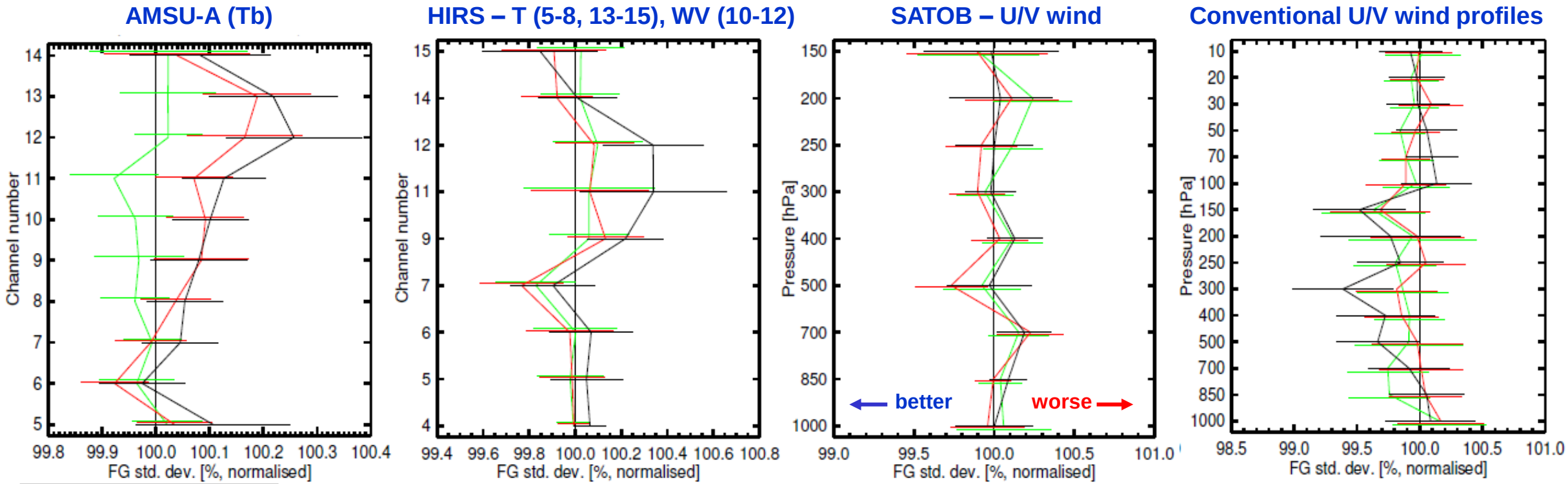
Assimilation system brings model closer to CloudSat and CALIPSO observations:

- *Applying horizontal averaging leads to additional decrease of FG and AN departures*

3-month period: 1 August – 31 October 2007

Verification against other assimilated observations – observation error impact

Impact of the size of observation error



RADLID_1err
RADLID_1.5err
RADLID_2err
100% = REF

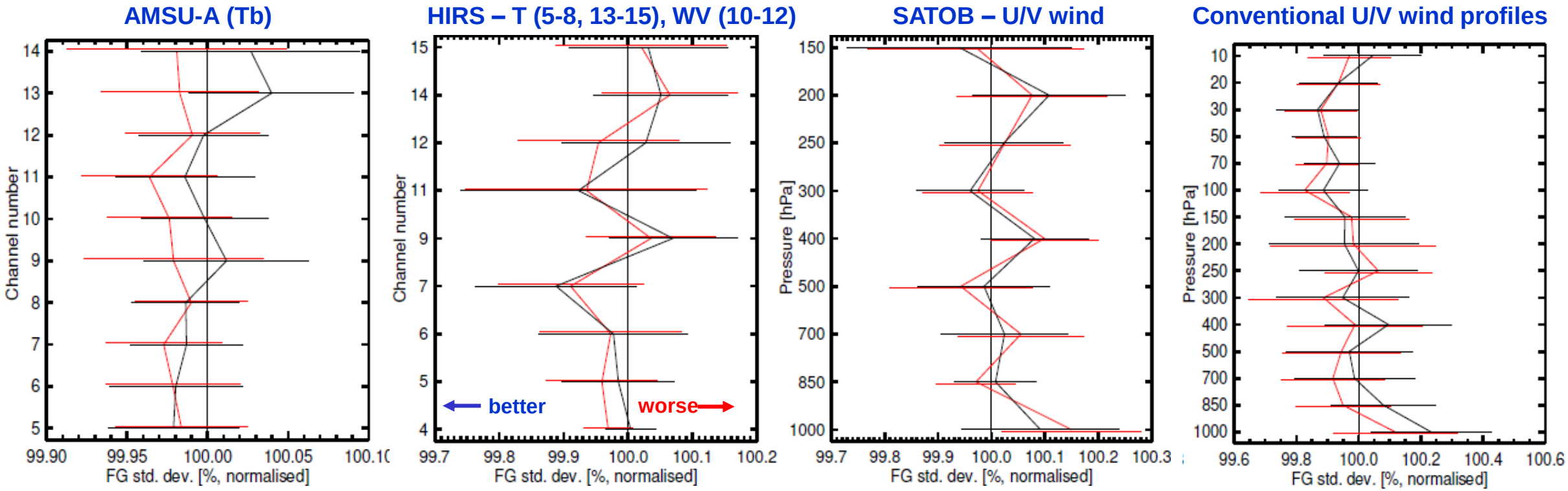
1err – single obs. errors as defined by Fielding & Stiller
1.5err –1.5 of the defined value
2err – obs. errors 2x as large as defined

- *Inflating observation errors for cloud radar & lidar observations:*
  - *better FG fits to the satellite observations sensitive to temperature and water vapour*
  - *more mixed results for satellite and conventional wind observations*
- ***Overall, increasing observation errors leads to better FG fits to other assimilated obs.***

1-month period: August 2007

Verification against other assimilated observations – impact of observation reduction

Impact of observation reduction (increased horizontal averaging)

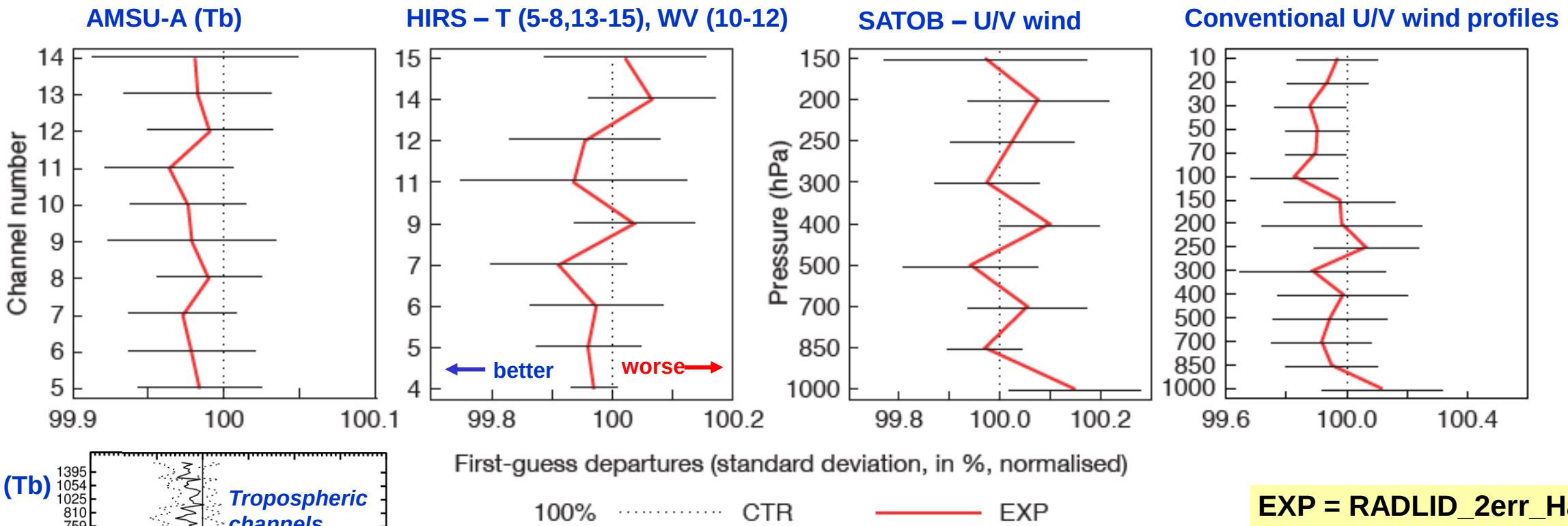


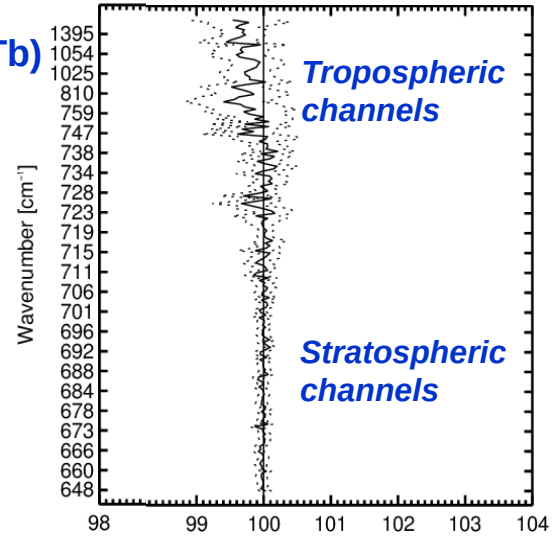
RADLID_2err
RADLID_2err_Ha
100% = REF

- *Increased horizontal averaging for cloud radar & lidar observations:*
  - *better FG fits to the satellite obs. sensitive to temperature*
  - *only rather small and mixed impact for satellite observations sensitive to water vapour & wind*
  - *improvement for conventional wind profiles*
- ***Using coarser resolution for cloud radar & lidar leads to the better FG fits to other assim.obs.***

3-month period: 1 August – 31 October 2007

Verification of assimilation runs against other assimilated observations - summary





EXP = RADLID_2err_Ha
double obs. error & horizontal averaging

Reduction in first guess standard deviation indicate better model fit to observations when assimilating space-borne cloud radar and lidar observations

3-month period: 1 August – 31 October 2007

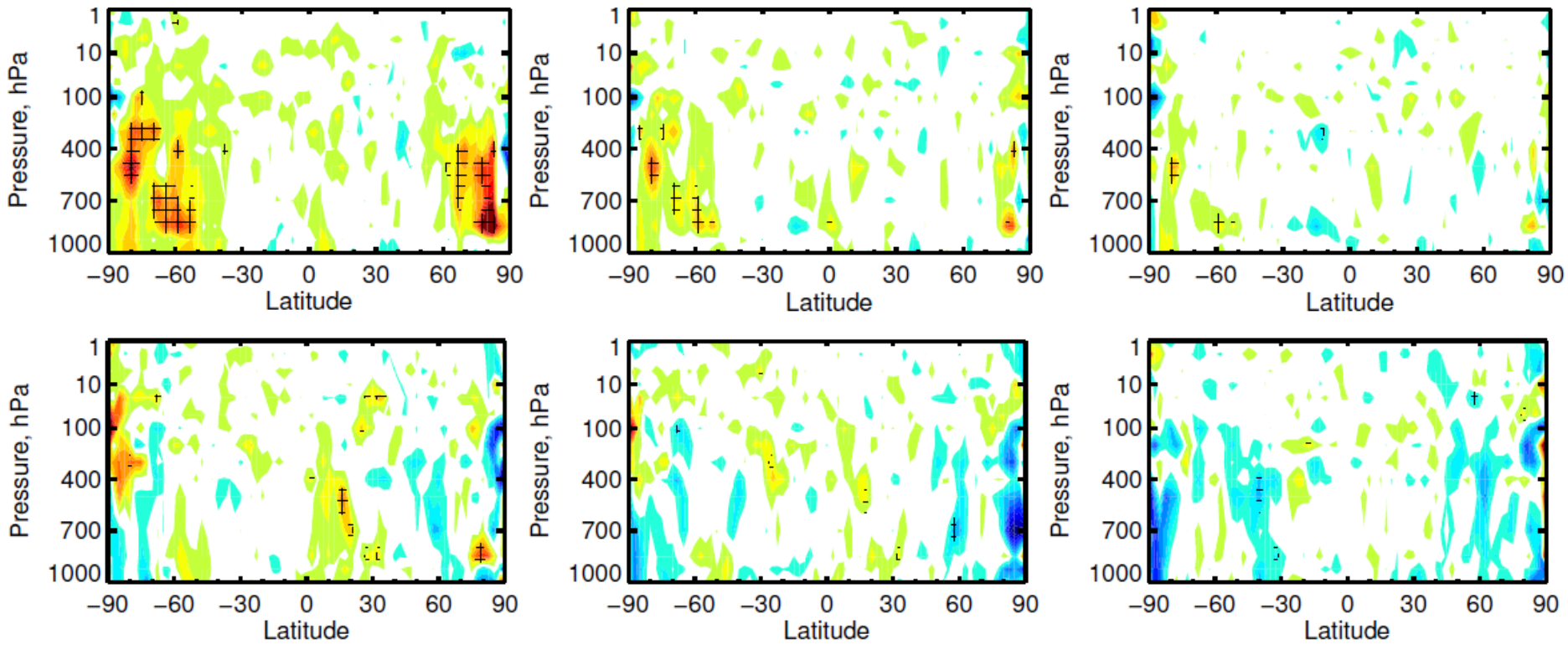
Verification of forecast against experiment's own analysis

Impact of the size of observation error

1err: single obs. error | **1.5 err: 1.5 obs. error** | **2err: double obs. error**

Temperature T+12

Temperature T+72



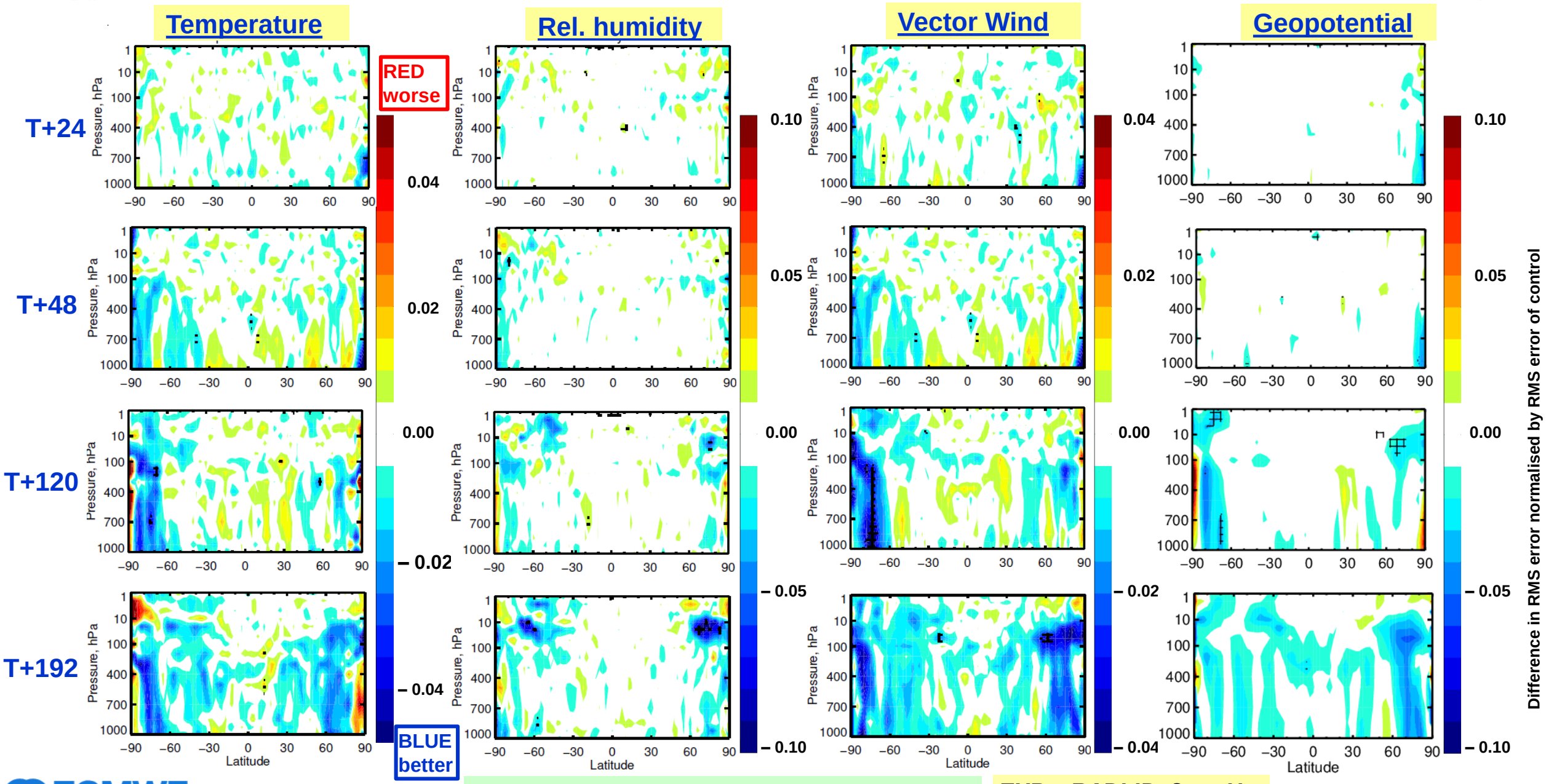
BLUE better

RED worse

- *Clear reduction in the forecast errors when increasing observation errors for cloud radar & lidar observations.*
- *The best results achieved when observation errors twice as large as defined for cloud radar reflectivity & lidar backscatter.*

3-month period: 1 August – 31 October 2007

4D-Var using new observations – impact on subsequent forecast (1)

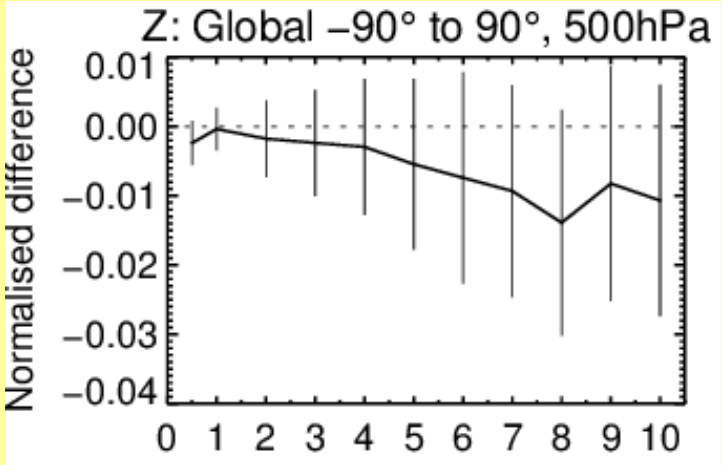


3-month period: 1 August – 31 October 2007

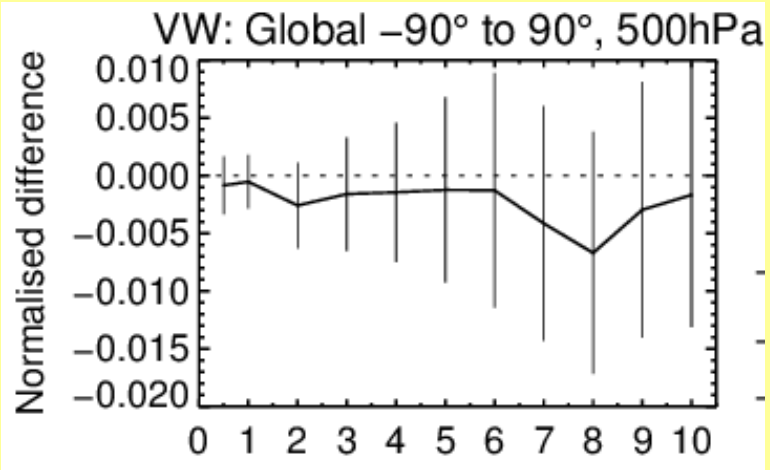
EXP = RADLID_2err_Ha

4D-Var using new observations – impact on subsequent forecast (2)

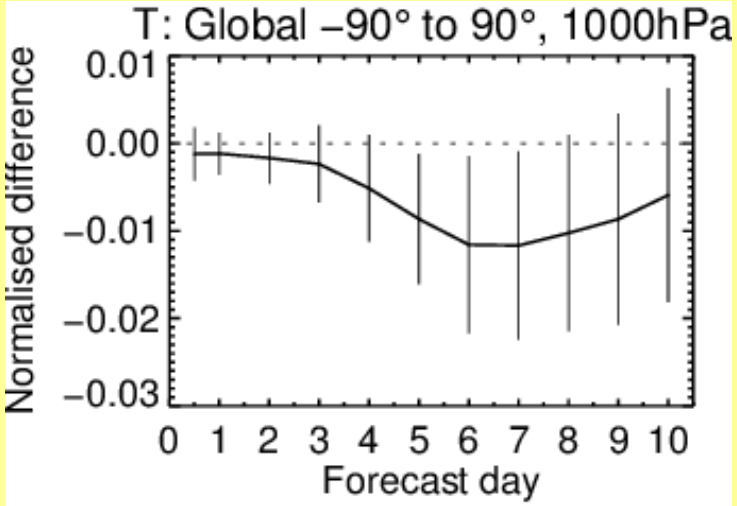
Geopotential – 500 hPa



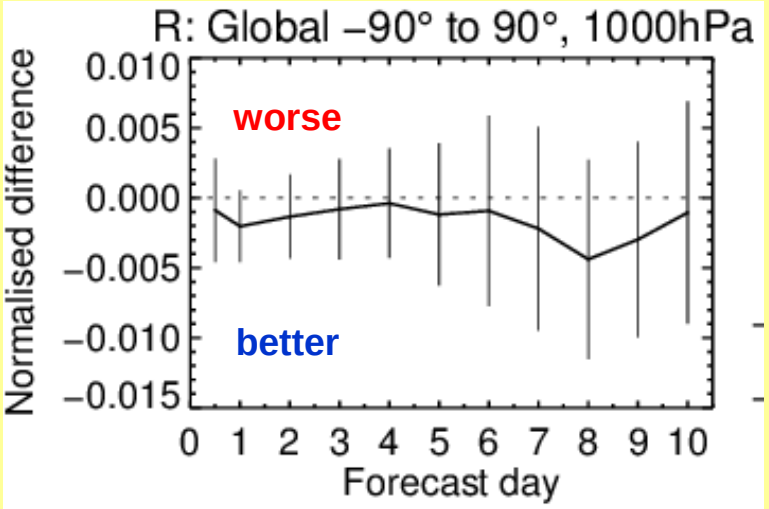
Wind – 500 hPa



Temperature – 1000 hPa



Humidity – 1000 hPa



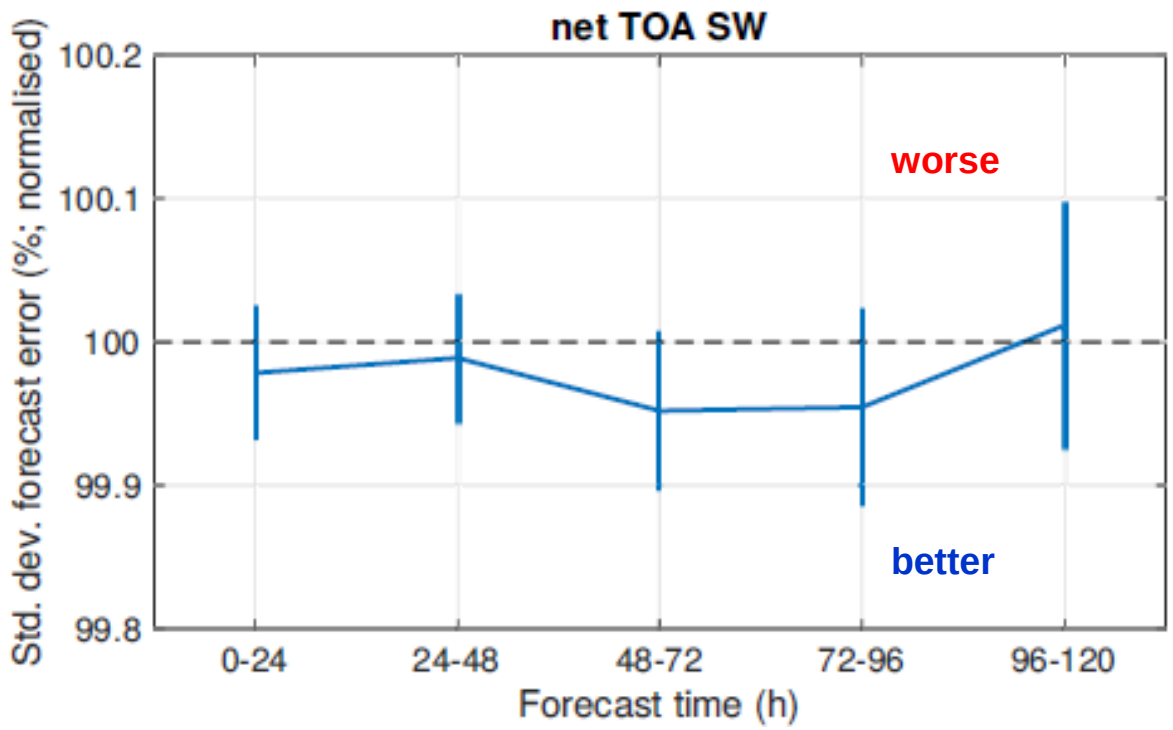
Negative values indicate reduction in forecast error in the experimental run

Assimilating space-borne cloud radar and lidar observations in combination with all other observations has positive impact on the subsequent forecast.

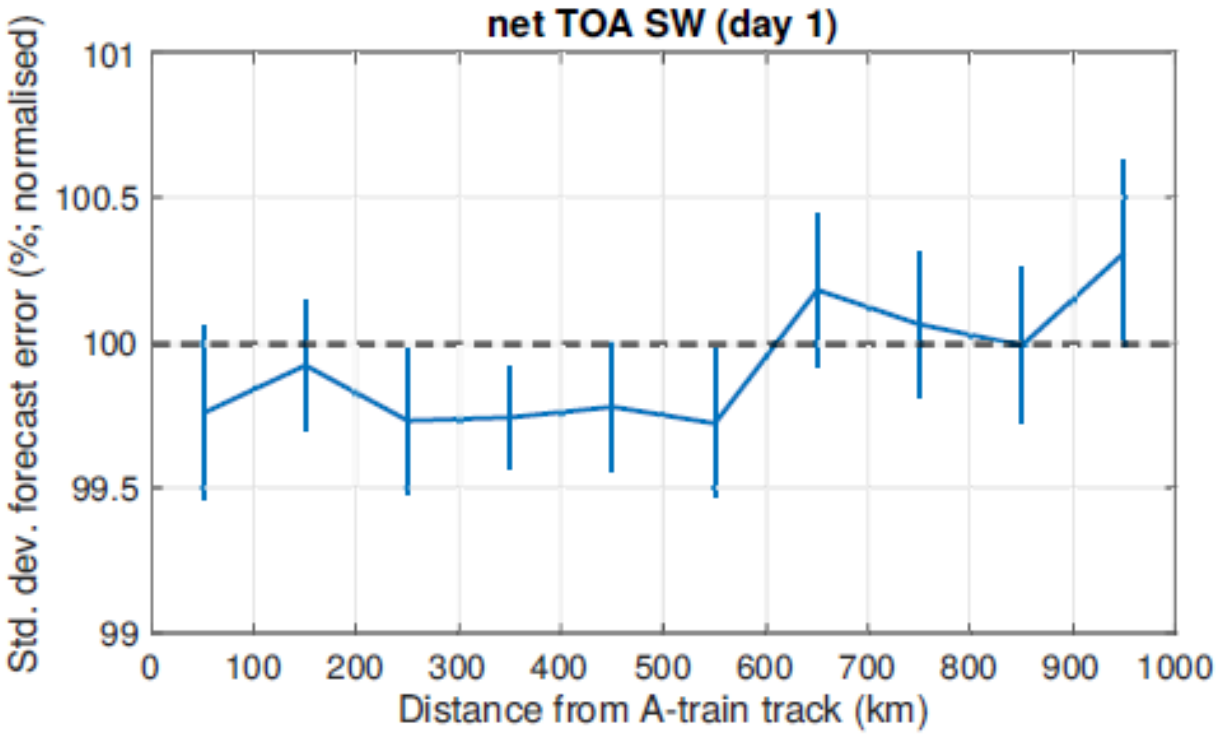
3-month period: 1 August – 31 October 2007

EXP = RADLID_2err_Ha

Impact on subsequent forecast – comparison against independent observations



Verification of forecast against net TOA short wave (SW) radiation from CERES for 3 months of 4D-Var cycling



CloudSat and CALIPSO improve forecast of TOA SW radiation:

- *up to 96 hours*
  → then decreased in time from analysis
- *significantly up to ~500 km from A-train track*
  → impact diminishing with distance

CERES – Clouds and the Earth's Radiant Energy System

3-month period: 1 August – 31 October 2007

EXP = RADLID_2err_Ha

Summary & perspectives

- Assimilation studies using observations of clouds from space-borne active instruments on board of CloudSat & CALIPSO have been performed to prepare for the EarthCARE mission

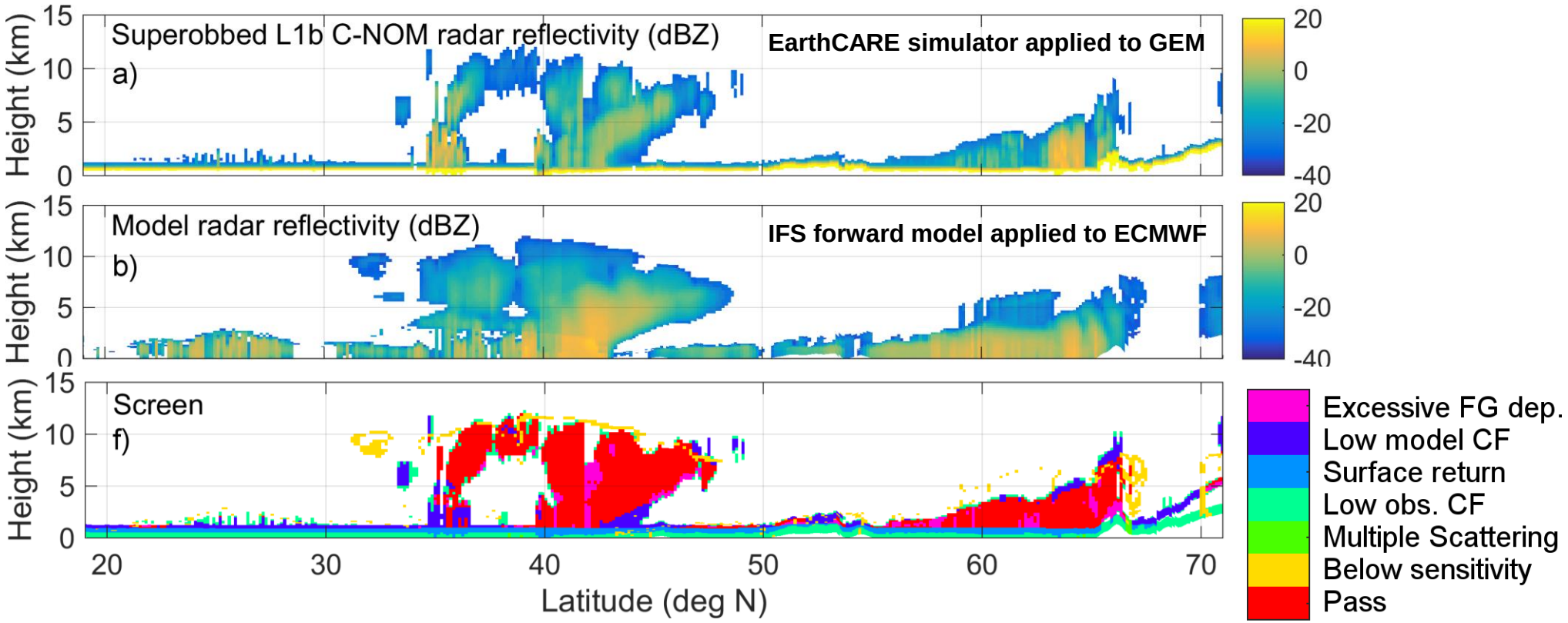
- Impact of the new observations on analyses and subsequent forecasts were studied using a 4D-Var technique in a **global NWP model**

- The feasibility of assimilating space-borne radar & lidar cloud observations demonstrated:
  - *positive impact on both the analysis fit to observations and the subsequent forecast*
  - *radar likely to provide the largest impact on forecast, but a combination of radar and lidar giving the greatest total impact*
  - *improved analysis and forecast of rain rates in the tropics and TOA SW radiation globally*

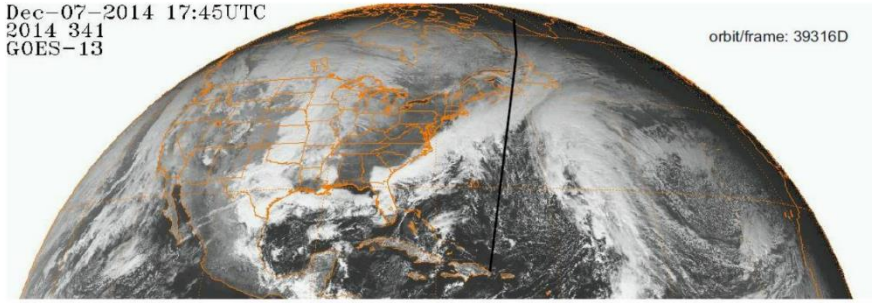
- The results are really promising and open many avenues of further research:
  - *careful tuning of observation error definition for gains in forecast skill*
  - *improvement in the forward operator assumptions & screening criteria in the different weather regimes*
  - *optimisation of the observation pre-processing (averaging scale, thinning, ...)*
  - *investigation how cloud radar&lidar obs can support the assimilation of other obs sensitive to clouds*

EXTRA SLIDES

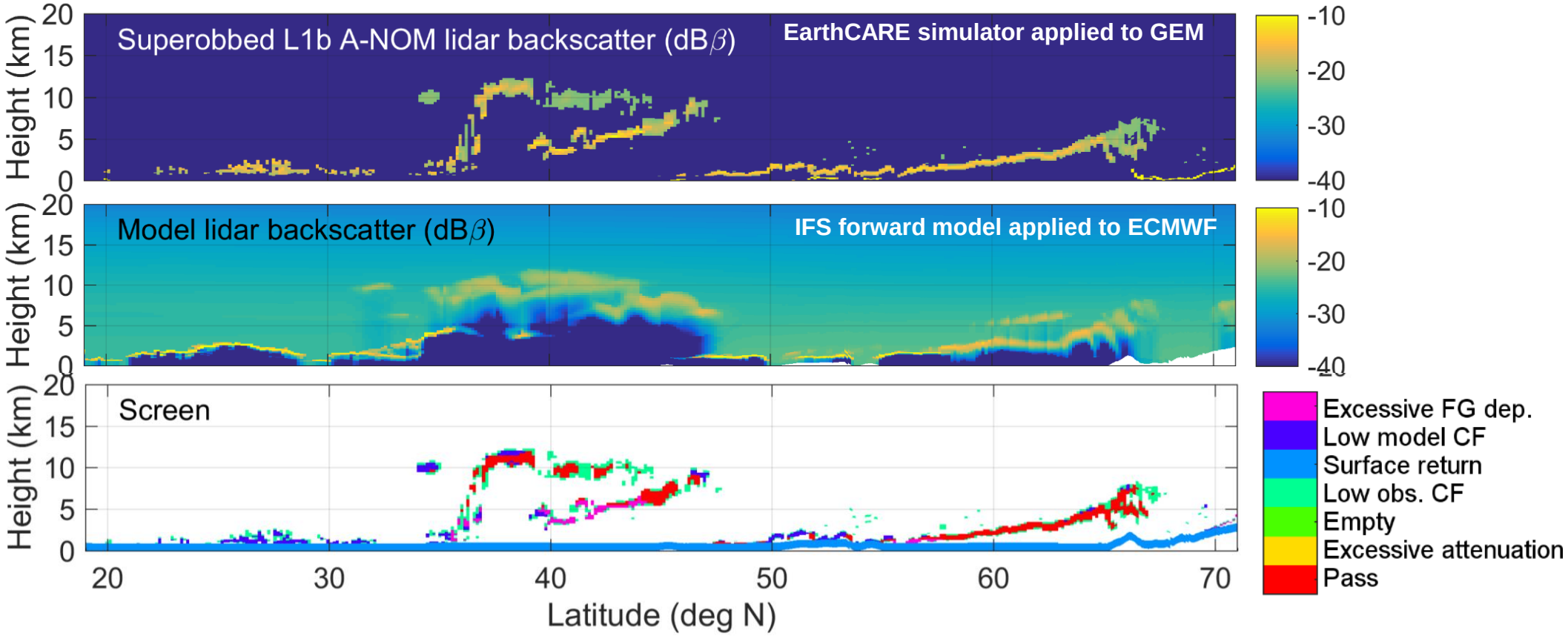
Initial screening: radar example



- Screening example based on C-NOM test data from Halifax scene (*courtesy of Aleksandra Tatarevic*)
- C-NOM data was also used to test obs. Pre-processing (e.g., conversion to BUFR and ODB)



Initial screening: lidar example



Instrument error

- The random error in the measurement due to noise
- An estimate of the lidar backscatter uncertainty is expected to be included in the L1B ATLID. For the CPR, the noise will be estimated from empty gates.

Radar

$$\Delta Z_{dB} = \frac{4.343}{\sqrt{N}}\left(1+\frac{1}{SNR}\right)$$

Hogan et al., 2005

SNR – Signal to noise ratio

N – Number of independent samples

Lidar

$$(\Delta b)^2 = NSF^2 b + \left(\frac{b}{V}\right)^2\left[(\Delta V_b)^2 + (\Delta \overline{V_b})^2\right]$$

Liu et al., 2006

$$\Delta b_{dB} = 4.343\left[\frac{NSF^2}{b} + \left(\frac{1}{SNR} + \frac{1}{\sqrt{N}SNR}\right)^2\right]^{\frac{1}{2}}$$

NSF – Noise scale factor

ΔV_b – Signal power

$\Delta \overline{V_b}$ – st.dev. mean background signal

Observation operator error

- To convert model hydrometeor content into radar reflectivity/lidar backscatter, many assumptions made with the potential to introduce error
- Possible sources of error:
  - Radiative transfer in scattering models
  - Hydrometeor shape
  - Particle size distribution
  - Multiple scattering
  - Subgrid assumptions (overlap, inhomogeneity & convective precip.fr.)
- Perturbing parameters within plausible bounds to characterize errors *(careful when parameters correlated or not having physical constraints...)*
- Errors estimated by Monte Carlo simulation – uncertainty is standard deviation of reflectivity/backscatter given a set of random realisations of PSD variables / densities / particle shapes
- Careful selection of PSD parameters to perturb to avoid spurious results
- Errors are function of hydrometeor type, LWC and temperature
- Including attenuation uncertainty – errors increased as signal attenuated

Representativity error

- Representativity error is the expected error due to mismatch between the model and the observational spatial and/or temporal scales.
- 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

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

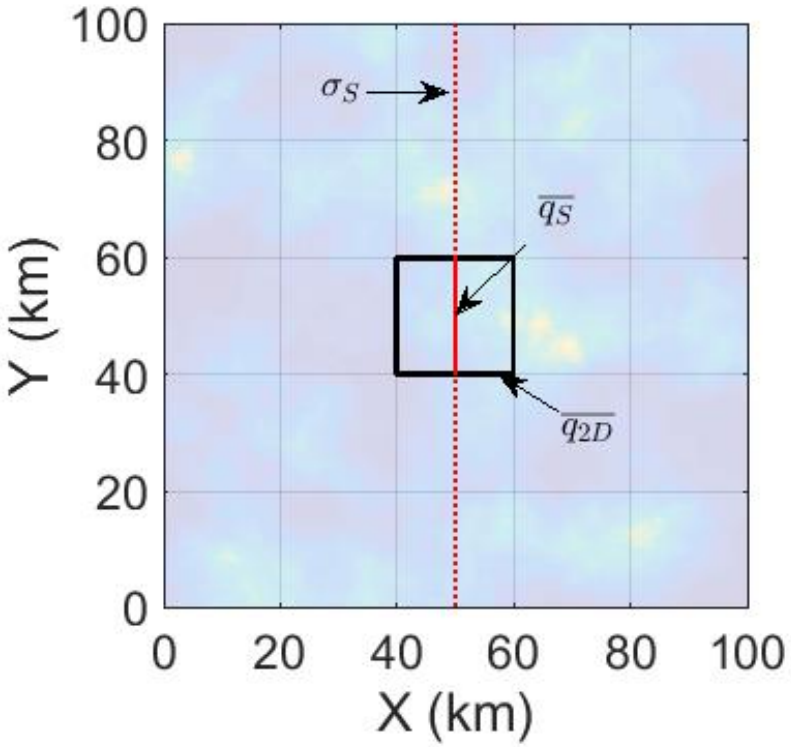
$$\sigma_{RE}^2 = \sigma_q^2 \underbrace{(\alpha_{[2D,2D]} + \alpha_{[s,s]} - 2\alpha_{[s,2D]})}_{\text{Scaling factor}}$$

Variance of measured variable within gridbox

$$\left.\begin{aligned}
\alpha_{[2D,2D]} &= \frac{1}{A_{2D}^2}\int_{A_{2D}}\int_{A_{2D}} \rho(\|\mathbf{x_1}-\mathbf{x_2}\|)d\mathbf{x_1}d\mathbf{x_2} \\
\alpha_{[s,s]} &= \frac{1}{A_s^2}\int_{A_s}\int_{A_s} \rho(\|\mathbf{x_1}-\mathbf{x_2}\|)d\mathbf{x_1}d\mathbf{x_2} \\
\alpha_{[s,2D]} &= \frac{1}{A_s A_{2D}}\int_{A_s}\int_{A_{2D}} \rho(\|\mathbf{x_1}-\mathbf{x_2}\|)d\mathbf{x_1}d\mathbf{x_2}
\end{aligned}\right\}$$

Correlation of measured variable within gridbox

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

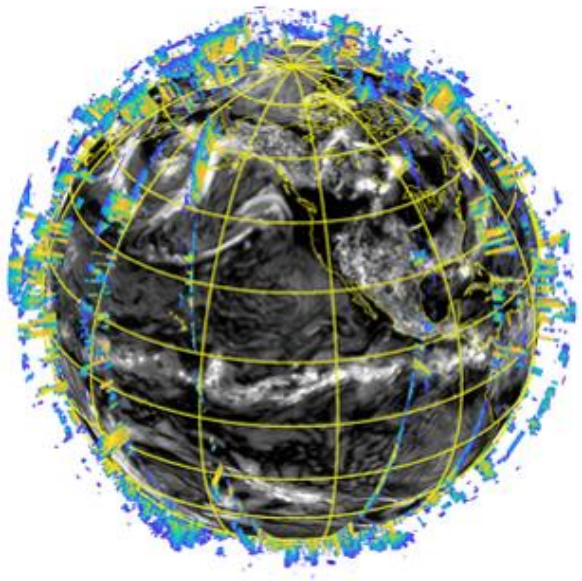


Progress towards assimilation of EarthCARE cloud radar and lidar observations

The recent GSP study '*Operational Assimilation of Space-borne Radar and Lidar Cloud Profiles for NWP*' developed the capability for assimilating EarthCARE at ECMWF, facilitating:

1. Monitoring of data quality in near real-time against the operational model.
2. Model-to-observation evaluation of cloud and precipitation and subsequent model development.
3. Potential for improving the model analysis leading to direct improvements of forecast skill.

Expected data coverage for 12 hours



Expected data coverage for one month