A Data Assimilation Testbed using KENDA and idealized COSMO Ensembles

Heiner Lange, George C. Craig

Hans-Ertel-Centre for Weather Research, Data Assimilation Branch

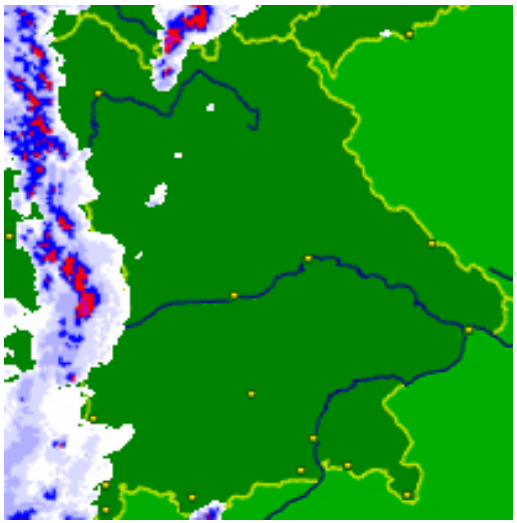
COSMO User Seminar
Data Assimilation Session
DWD Offenbach, 17.03.2014

Outline

1. idealized COSMO-KENDA
   - Convective-scale Data Assimilation
   - Nature Run and Synthetic Observations
2. Experiments: LETKF Cycling and Ensemble Forecasts
   - Analysis Ensemble Convergence
   - Forecast error growth
   - Results, Conclusions, Outlook

Thunderstorm Prediction

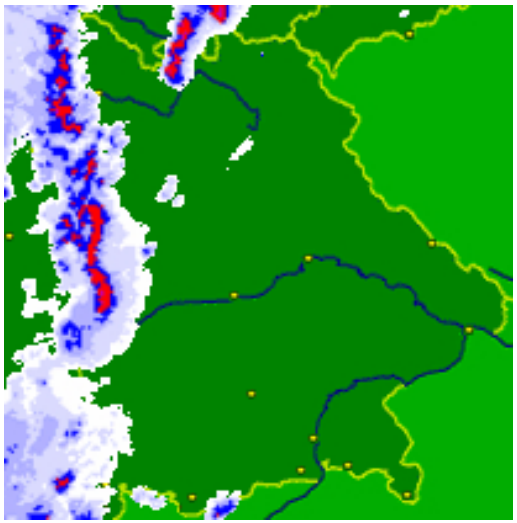
Large-scale phenomena (e.g. low-pressure systems) more predictable than smalls scales (convection, thunderstorms)



$\Delta t = 1 - 3$ h $\rightarrow$ difficult predictability for $\rightarrow \Delta x = 20km$

Thunderstorm Prediction

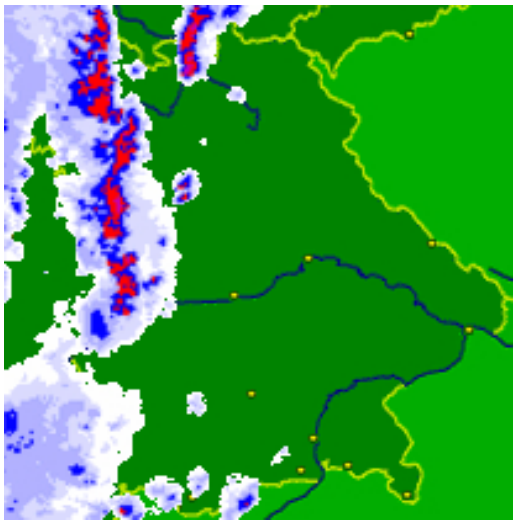
Large-scale phenomena (e.g. low-pressure systems) more predictable than smalls scales (convection, thunderstorms)



$\Delta t = 1 - 3$ h $\rightarrow$ difficult predictability for $\rightarrow \Delta x = 20km$

Thunderstorm Prediction

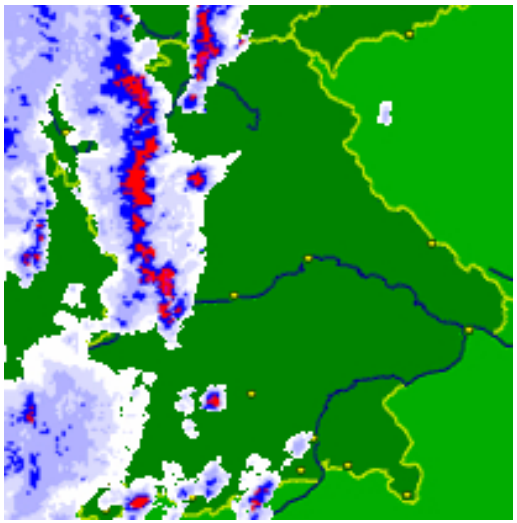
Large-scale phenomena (e.g. low-pressure systems) more predictable than smalls scales (convection, thunderstorms)



$\Delta t = 1 - 3$ h $\rightarrow$ difficult predictability for $\rightarrow \Delta x = 20km$

Thunderstorm Prediction

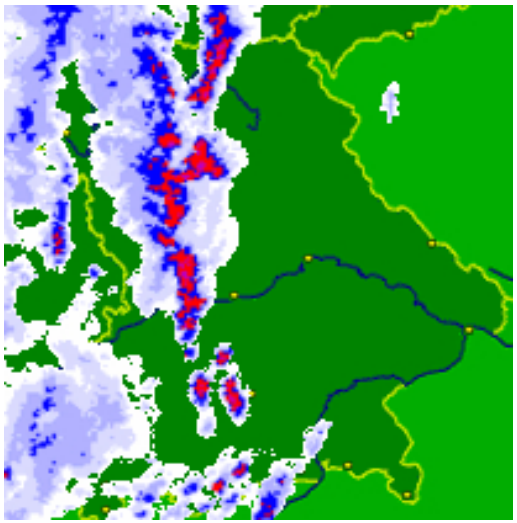
Large-scale phenomena (e.g. low-pressure systems) more predictable than smalls scales (convection, thunderstorms)



$\Delta t = 1 - 3$ h $\rightarrow$ difficult predictability for $\rightarrow \Delta x = 20km$

Thunderstorm Prediction

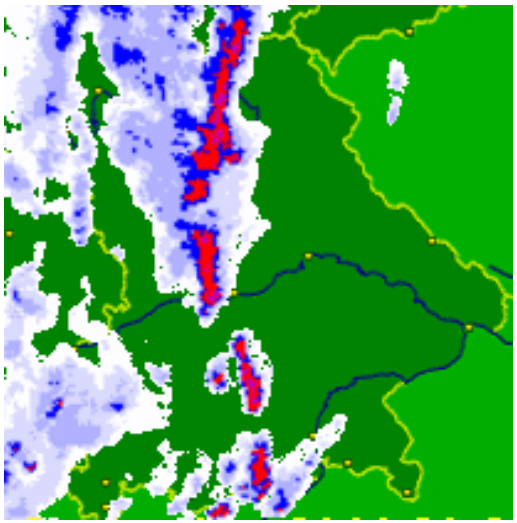
Large-scale phenomena (e.g. low-pressure systems) more predictable than smalls scales (convection, thunderstorms)



$\Delta t = 1 - 3$ h $\rightarrow$ difficult predictability for $\rightarrow \Delta x = 20km$

Thunderstorm Prediction

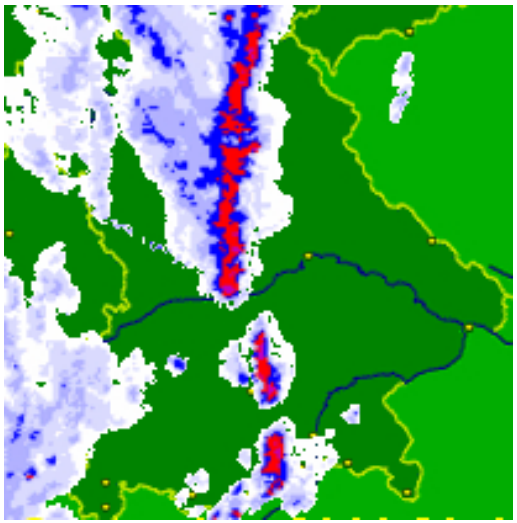
Large-scale phenomena (e.g. low-pressure systems) more predictable than smalls scales (convection, thunderstorms)



$\Delta t = 1 - 3$ h $\rightarrow$ difficult predictability for $\rightarrow \Delta x = 20km$

Thunderstorm Prediction

Large-scale phenomena (e.g. low-pressure systems) more predictable than smalls scales (convection, thunderstorms)



$\Delta t = 1 - 3$ h $\rightarrow$ difficult predictability for $\rightarrow \Delta x = 20km$

Ensemble Data Assimilation

Ensemble Forecasting and Data Assimilation:

- chaotic nature of atmosphere taken into account

Ensemble Data Assimilation

Ensemble Forecasting and Data Assimilation:

- chaotic nature of atmosphere taken into account
- ensemble forecasts more useful than deterministic forecasts:
  - ensemble mean represents state with highest probability
  - ensemble spread gives uncertainty information

Ensemble Data Assimilation

Ensemble Forecasting and Data Assimilation:

- chaotic nature of atmosphere taken into account
- ensemble forecasts more useful than deterministic forecasts:
  - ensemble mean represents state with highest probability
  - ensemble spread gives uncertainty information
- decision of DWD for operation of KENDA:
  - Kilometre-scale ENsemble Data Assimilation
  - Local Ensemble Transform Kalman Filter (LETKF) coupled with convection-permitting German forecast model COSMO-DE

Data Assimilation of Convection

Experiments for Ensemble Data Assimilation of Convection:

- **Perfect model** experiments (with nature run) to understand LETKF filter algorithms
- **Convection only**: no large scales
- **Simulated Dopper Radar** as realistic observation system

Nature Run and Ensemble

COSMO model setup (based on Daniel Leuenberger's work)

Domain: 198 x 198 x 50 gridpoints
periodic lateral boundaries conditions

Resolution: 2 km horizontally

Initial state: Horizontally homogenous sounding,
CAPE $= 2200 \frac{\text{J}}{\text{kg}}$,
random white noise on T (0.02 K) and W (0.02 $\frac{\text{m}}{\text{s}}$)
in the boundary layer

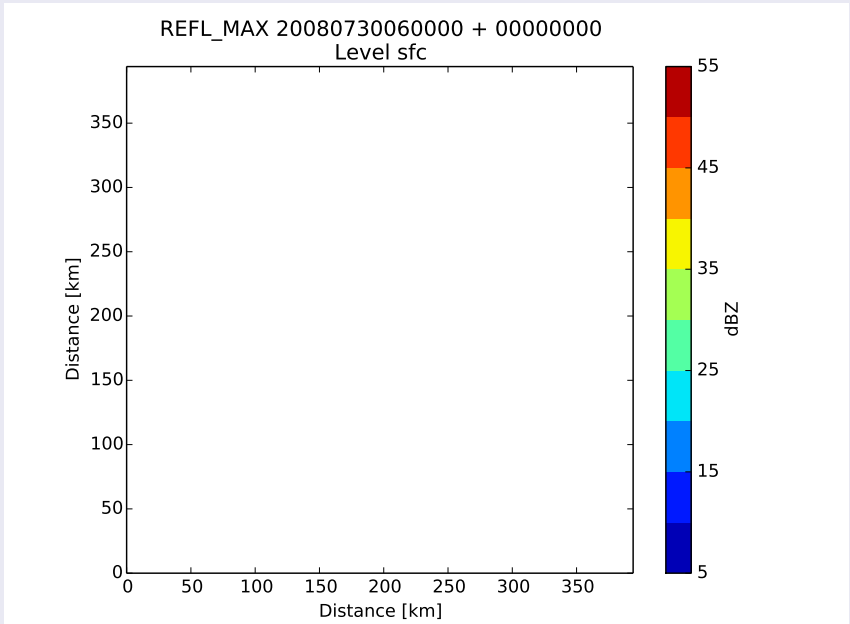
Model: Full COSMO physics with active radiation scheme

Forecast: 8 hour spinup until convection evolves:

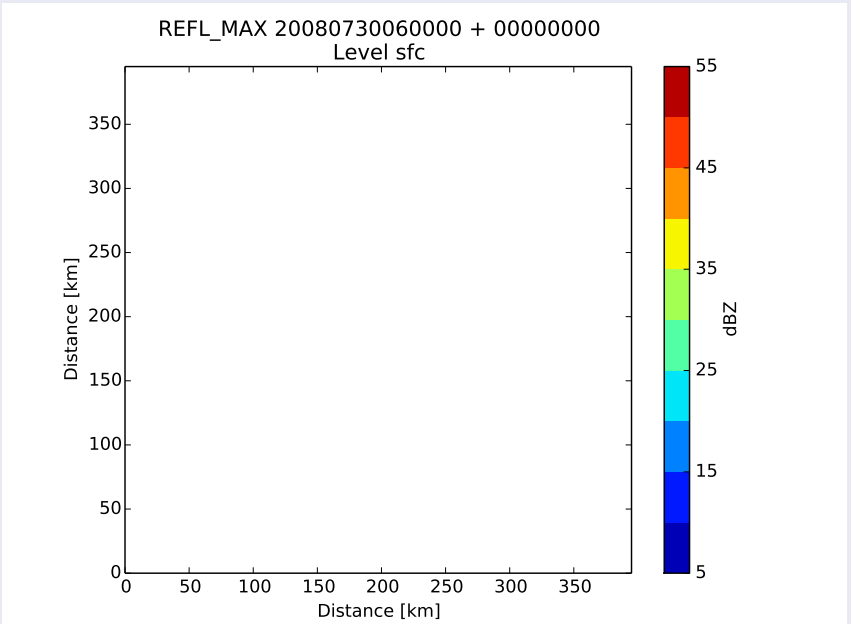
- long-lived cells, lifetime $\geq$ 6 h
- horizontal position *fully random* in ensemble

1 km Resolution vs. 2 km

$\Delta x = 2$ km

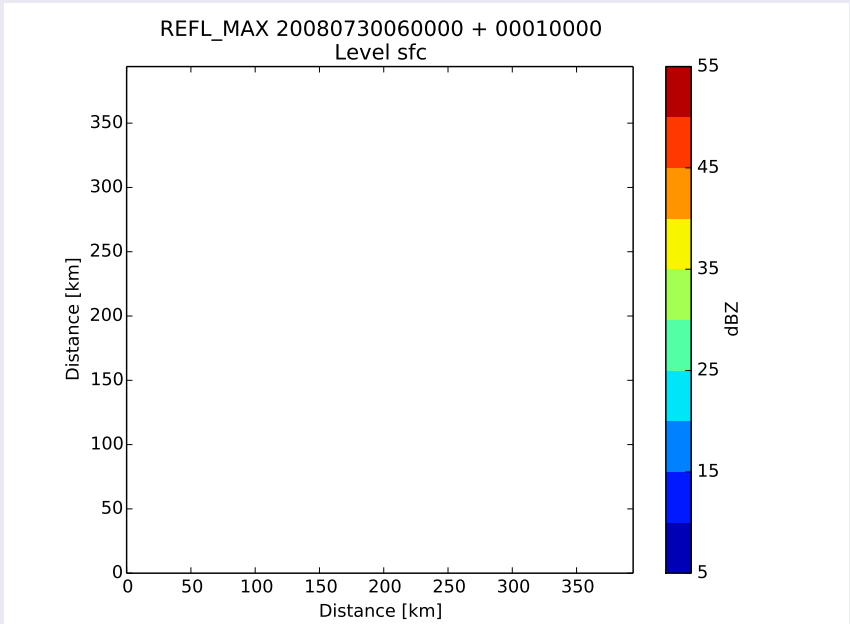


$\Delta x = 1$ km

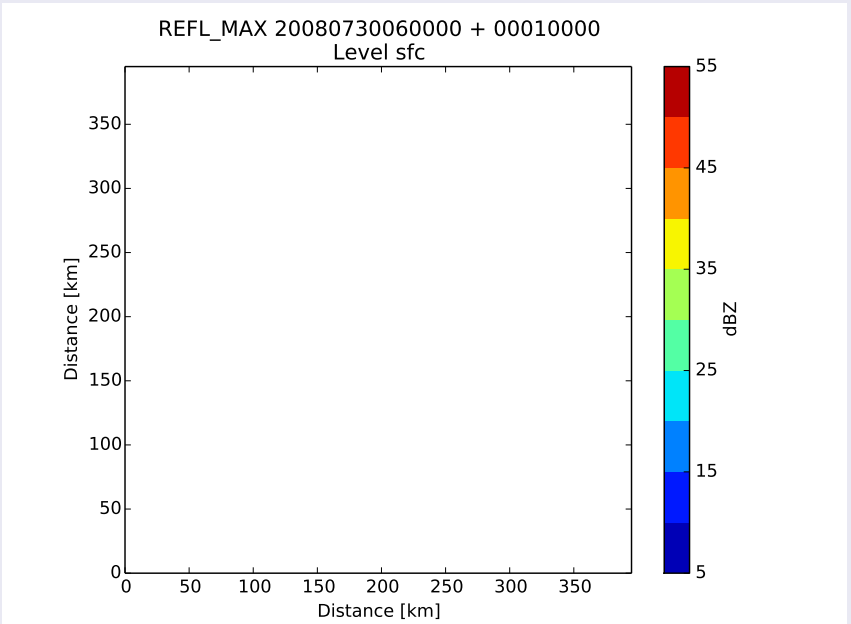


1 km Resolution vs. 2 km

$\Delta x = 2$ km

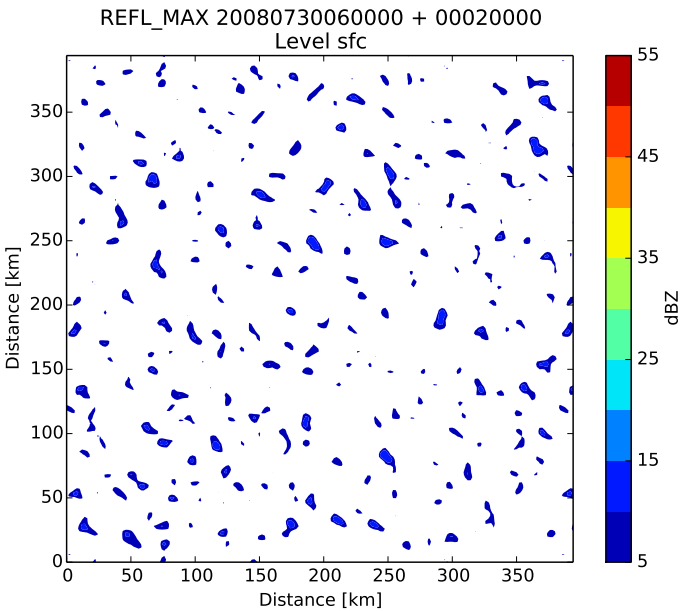


$\Delta x = 1$ km

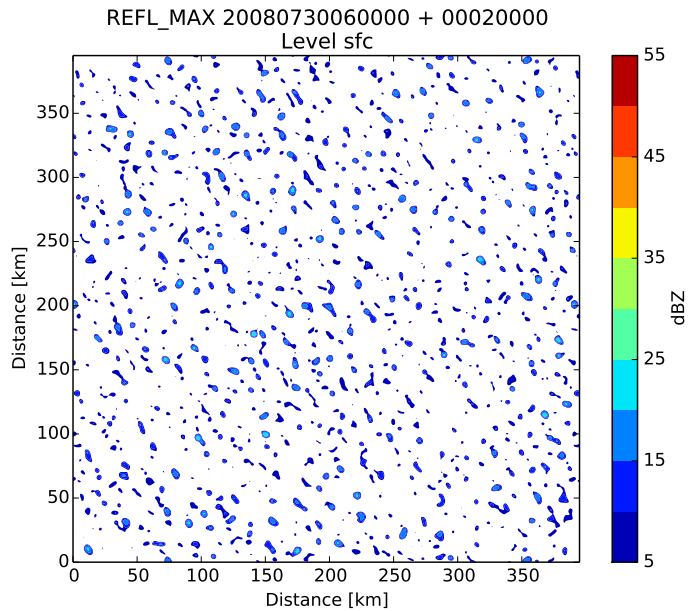


1 km Resolution vs. 2 km

$\Delta x = 2$ km

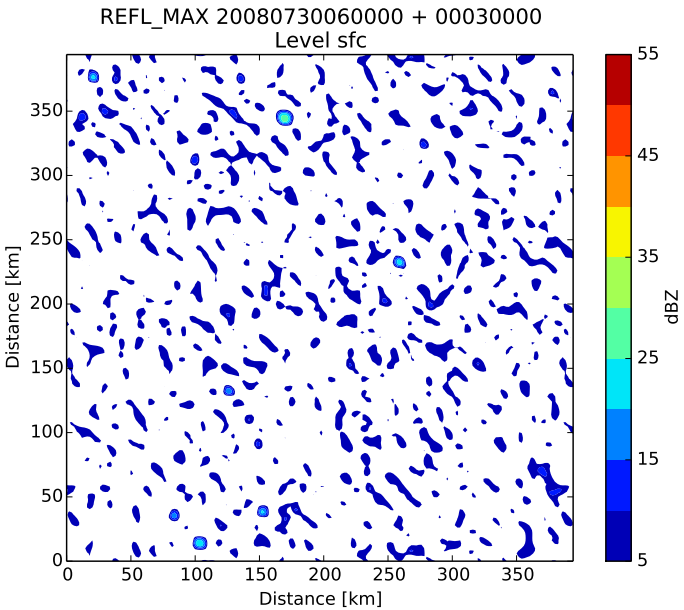


$\Delta x = 1$ km

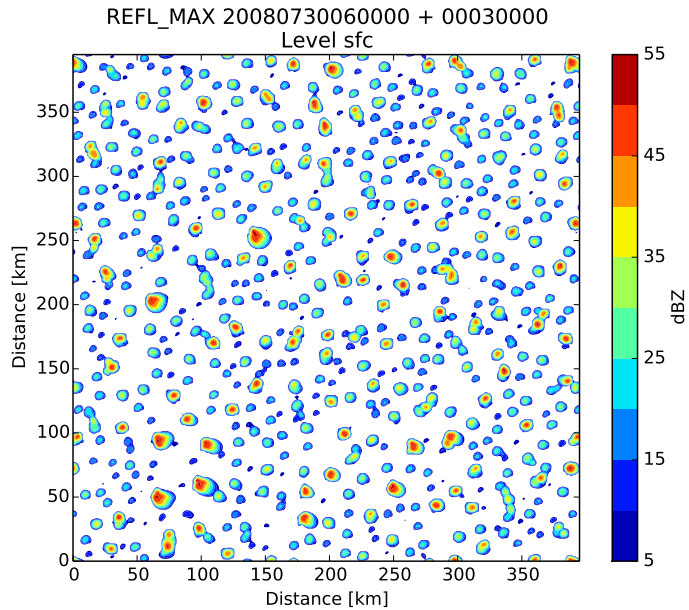


1 km Resolution vs. 2 km

$\Delta x = 2$ km

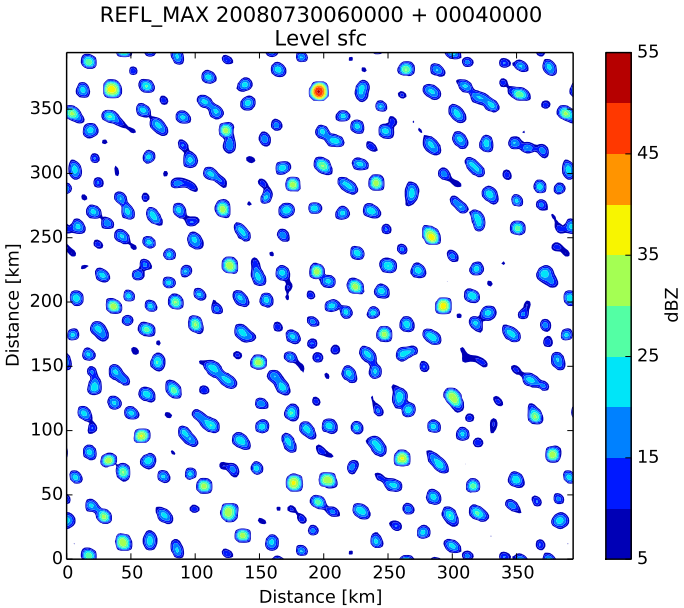


$\Delta x = 1$ km

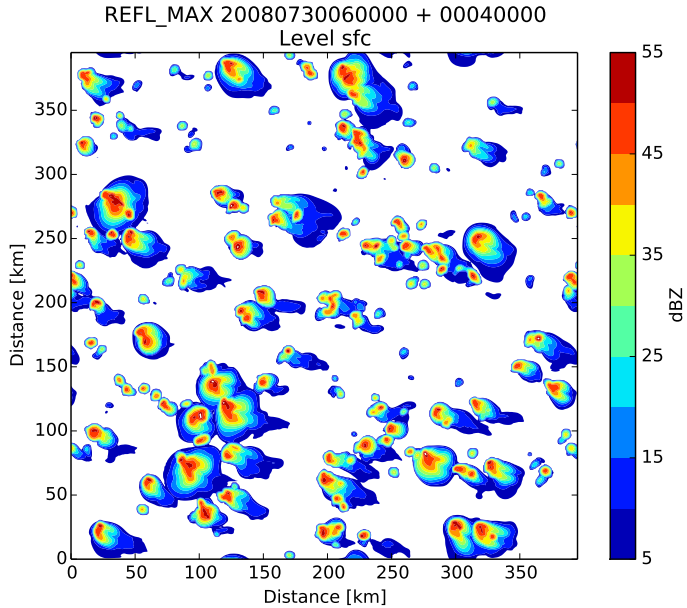


1 km Resolution vs. 2 km

$\Delta x = 2$ km

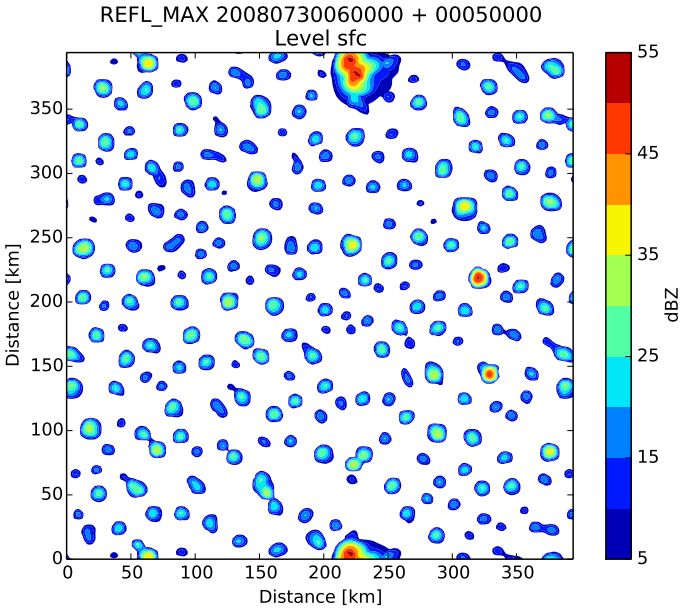


$\Delta x = 1$ km

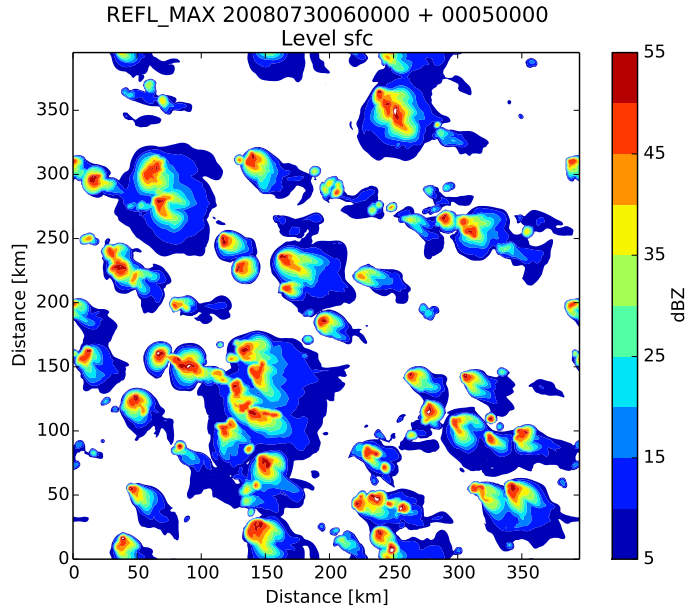


1 km Resolution vs. 2 km

$\Delta x = 2$ km

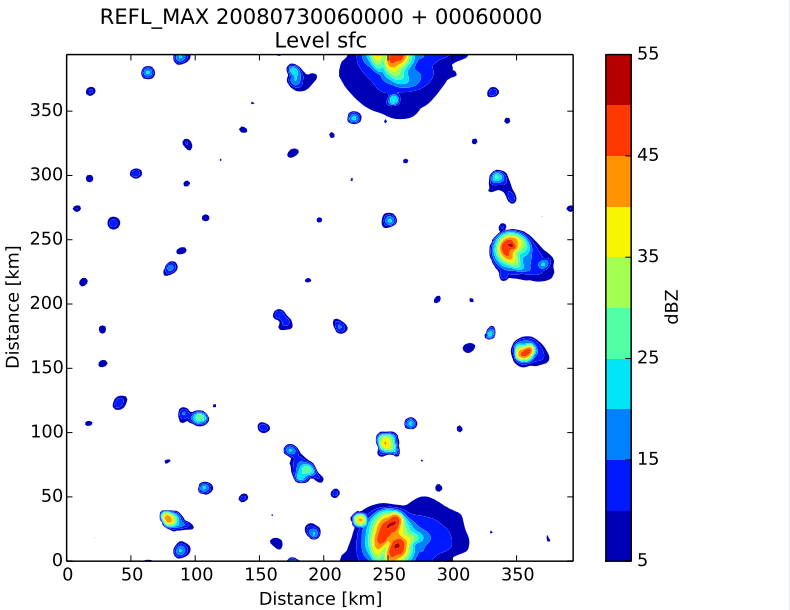


$\Delta x = 1$ km

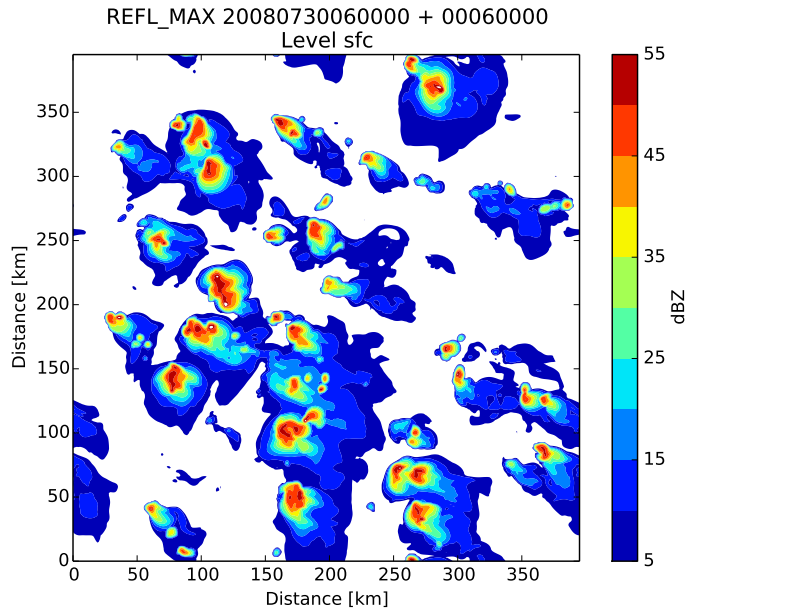


1 km Resolution vs. 2 km

$\Delta x = 2$ km

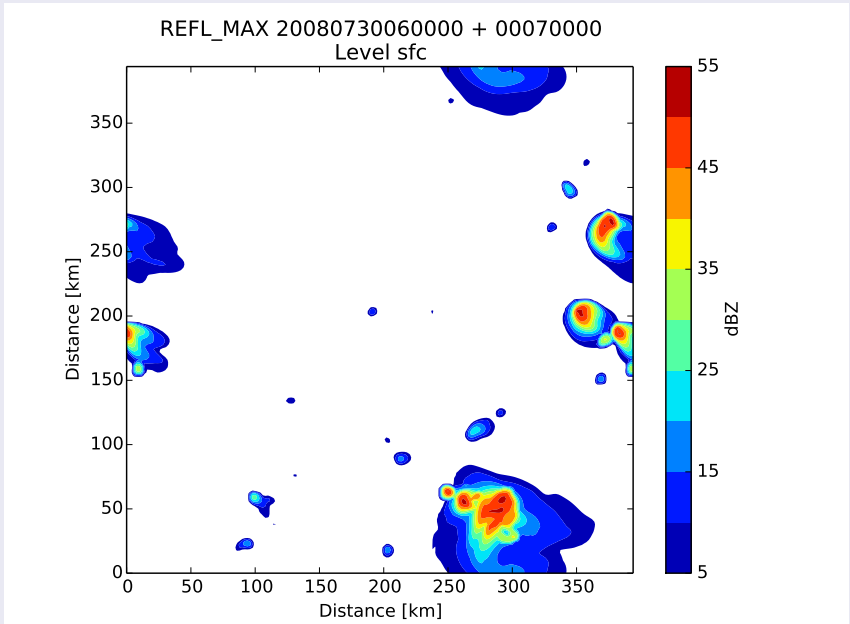


$\Delta x = 1$ km

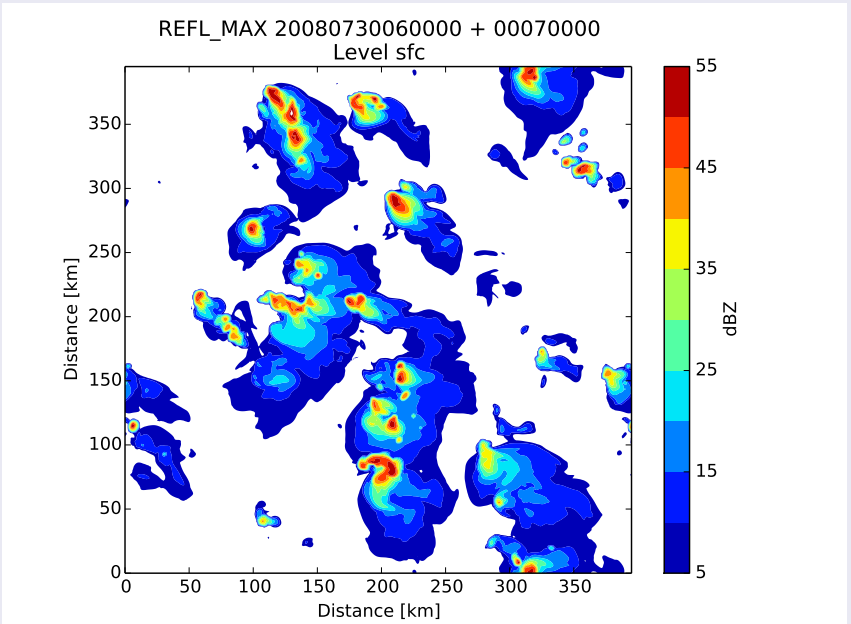


1 km Resolution vs. 2 km

$\Delta x = 2$ km

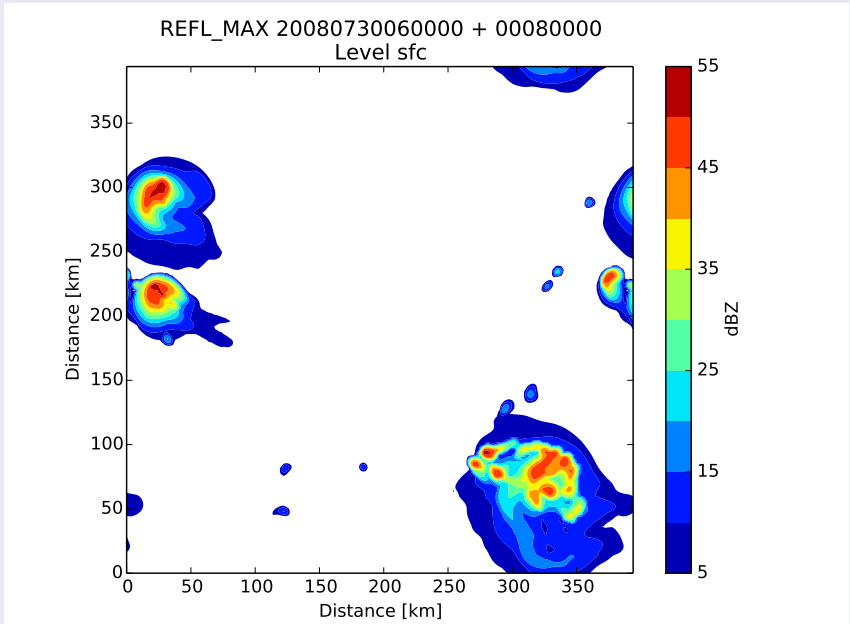


$\Delta x = 1$ km

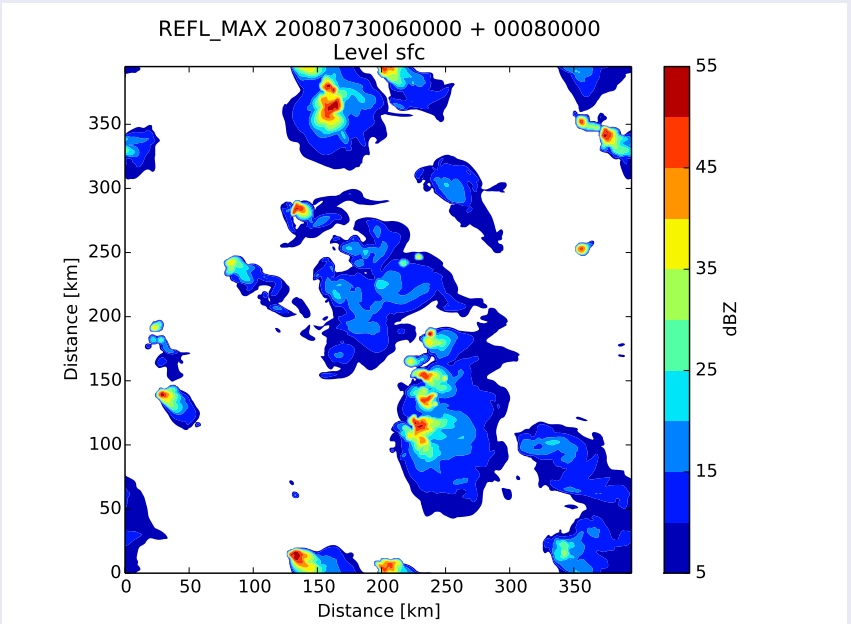


1 km Resolution vs. 2 km

$\Delta x = 2$ km

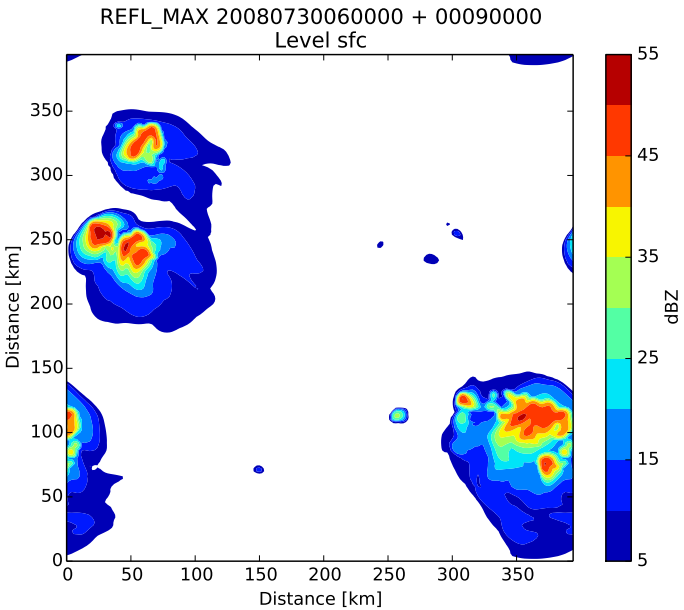


$\Delta x = 1$ km

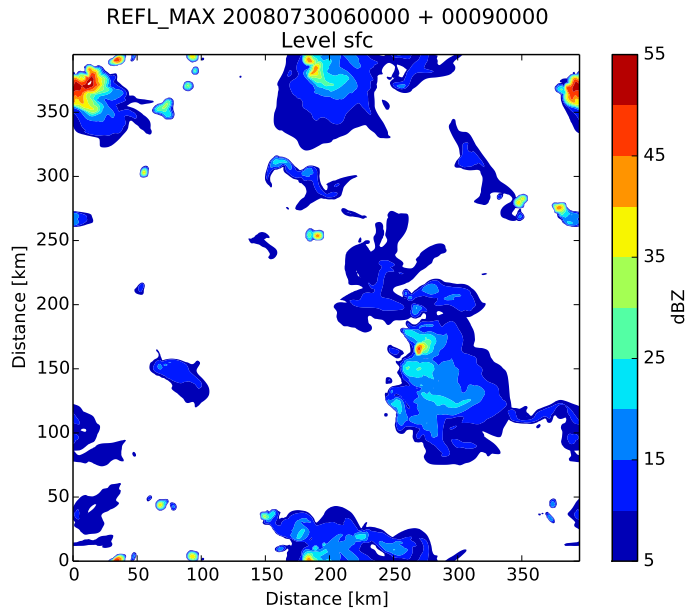


1 km Resolution vs. 2 km

$\Delta x = 2$ km

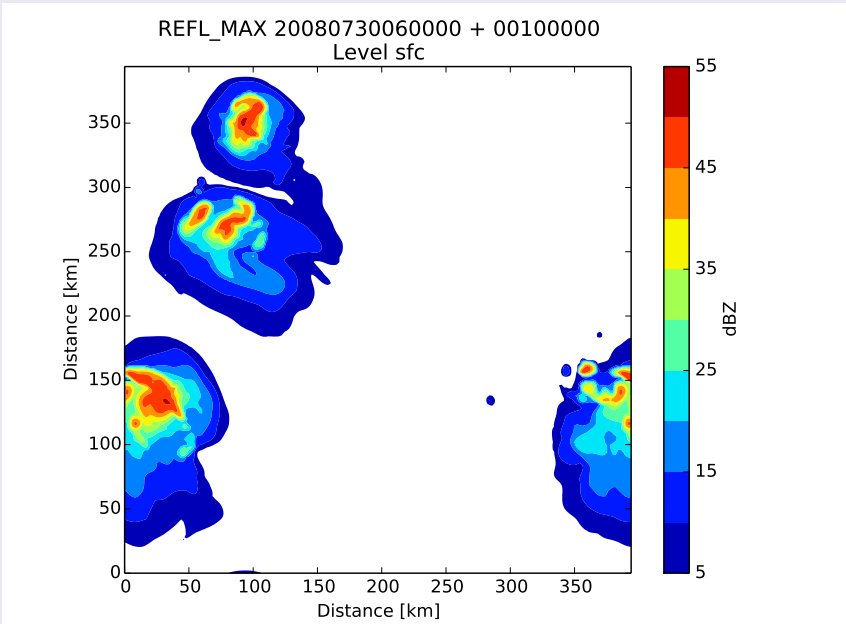


$\Delta x = 1$ km

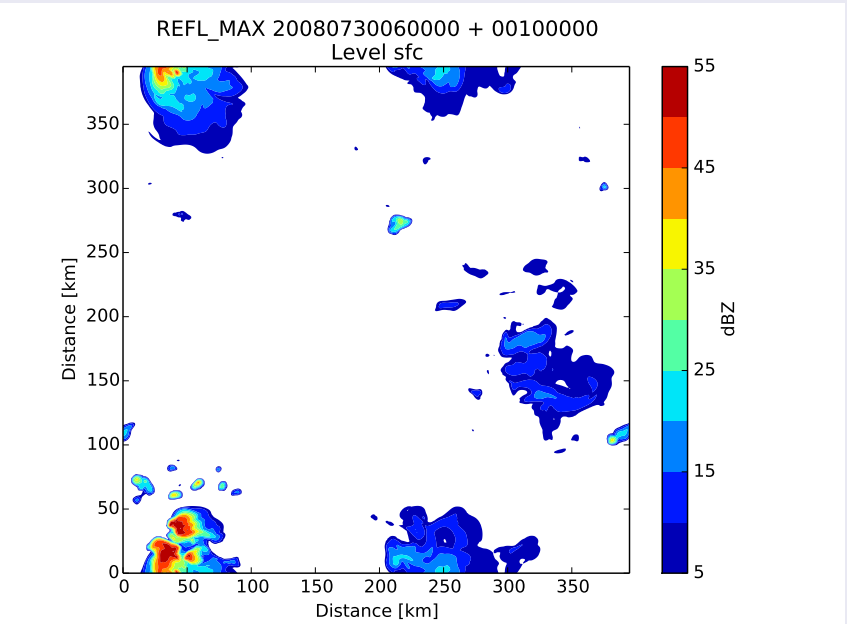


1 km Resolution vs. 2 km

$\Delta x = 2$ km

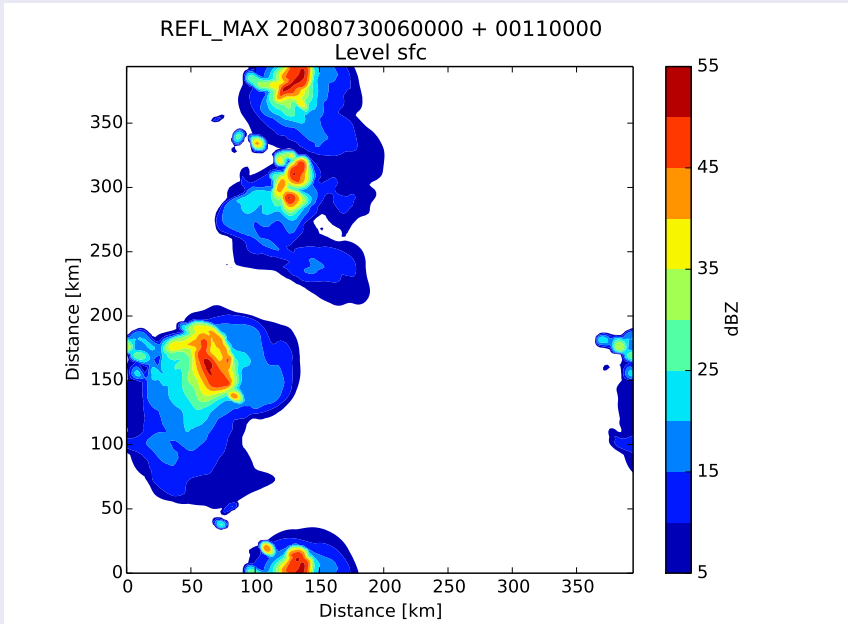


$\Delta x = 1$ km

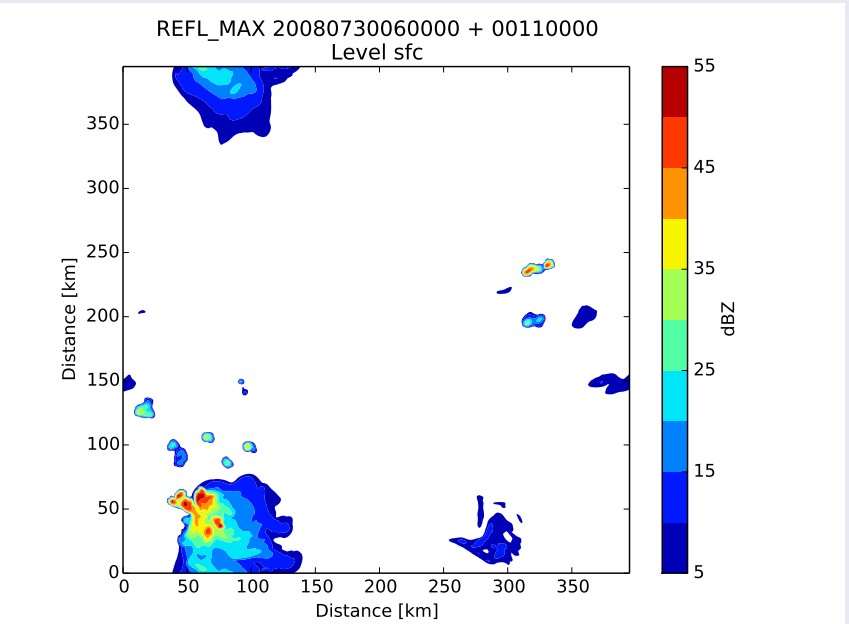


1 km Resolution vs. 2 km

$\Delta x = 2$ km

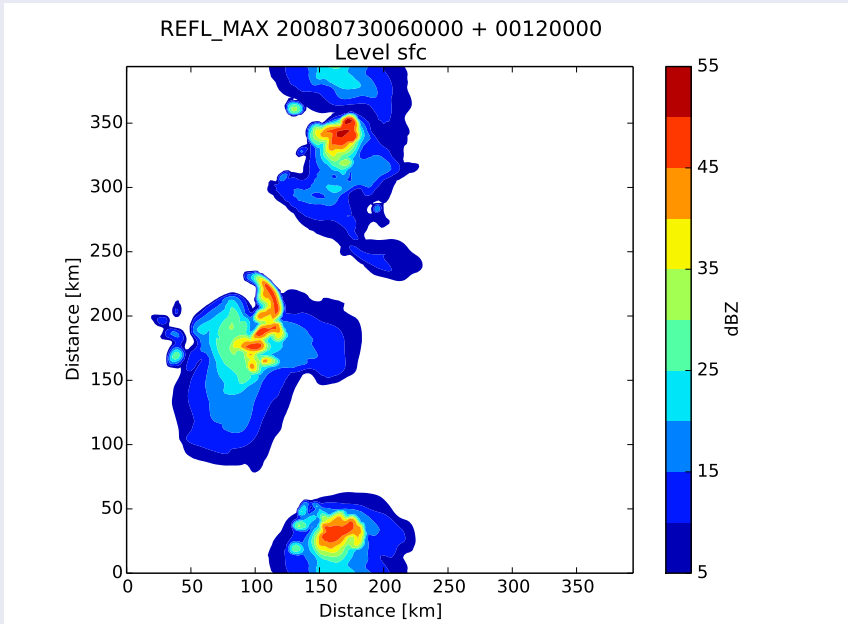


$\Delta x = 1$ km

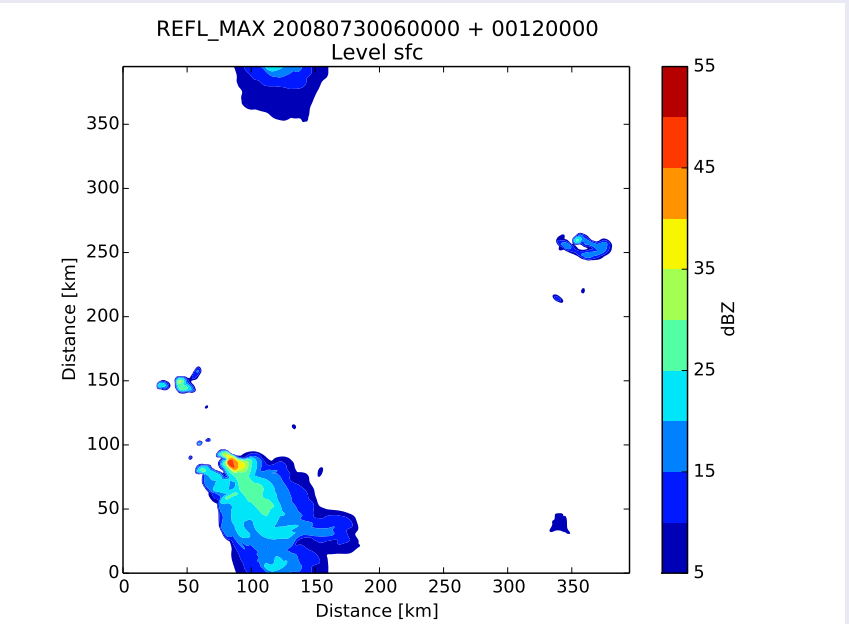


1 km Resolution vs. 2 km

$\Delta x = 2$ km

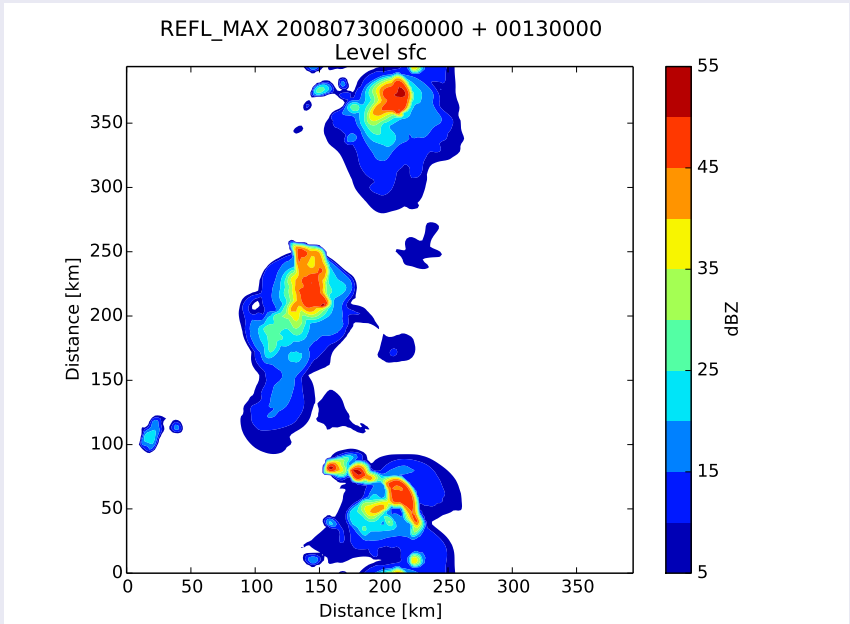


$\Delta x = 1$ km

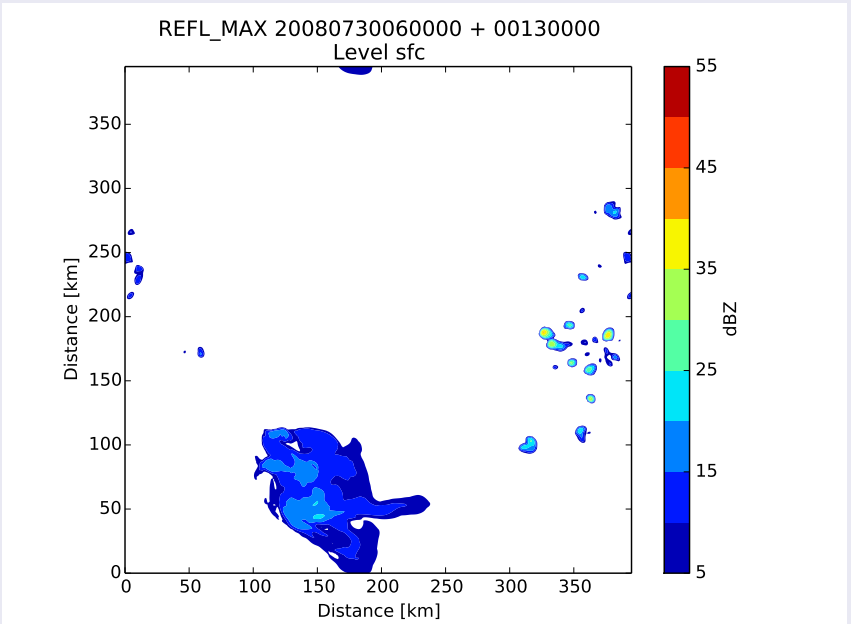


1 km Resolution vs. 2 km

$\Delta x = 2$ km

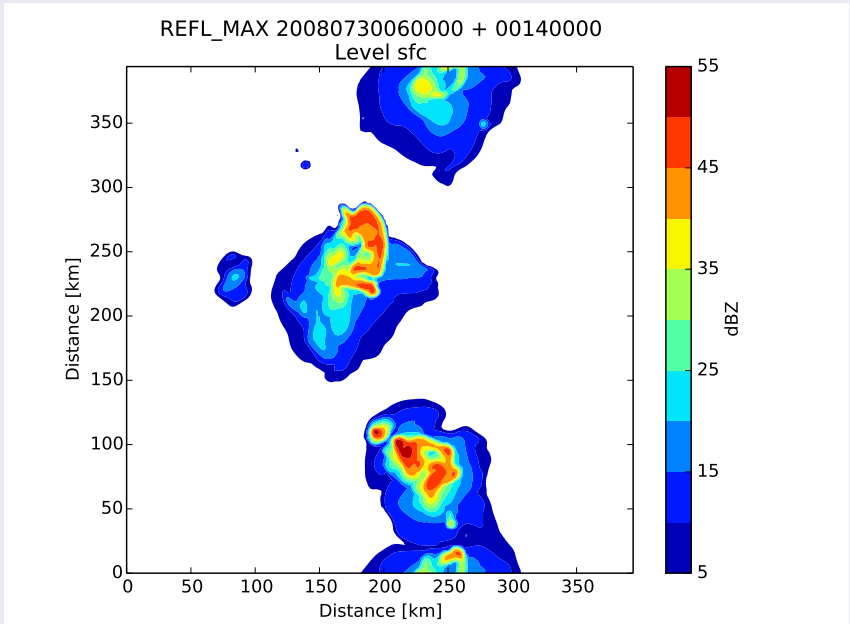


$\Delta x = 1$ km

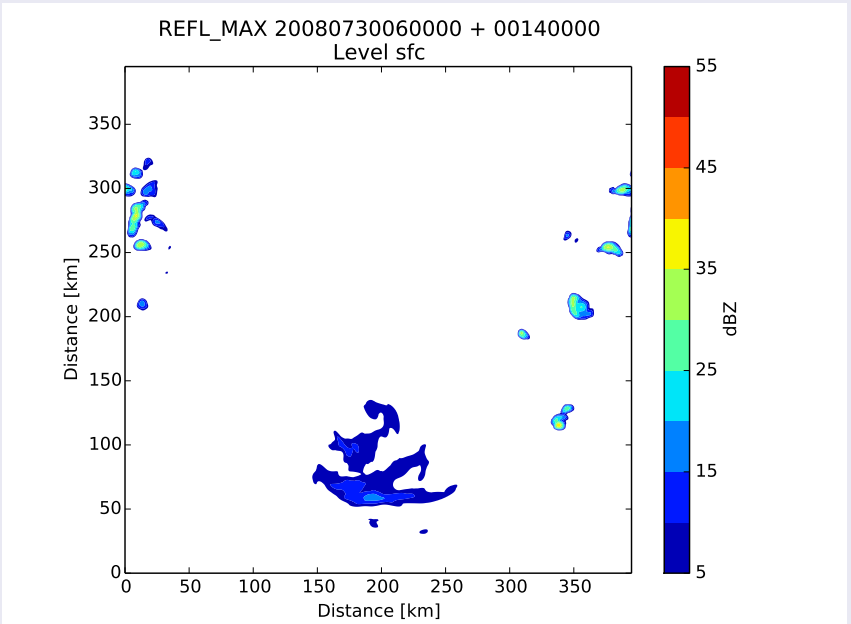


1 km Resolution vs. 2 km

$\Delta x = 2$ km

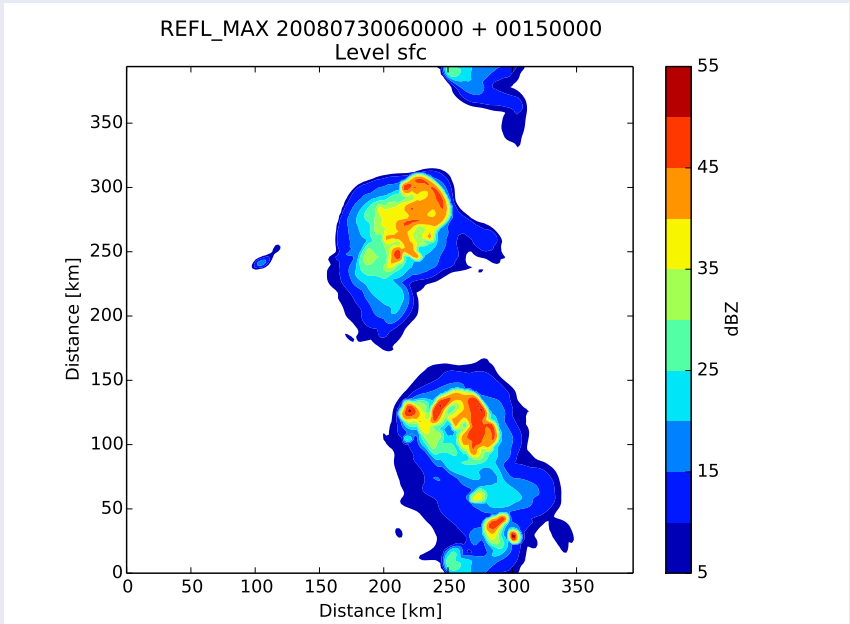


$\Delta x = 1$ km

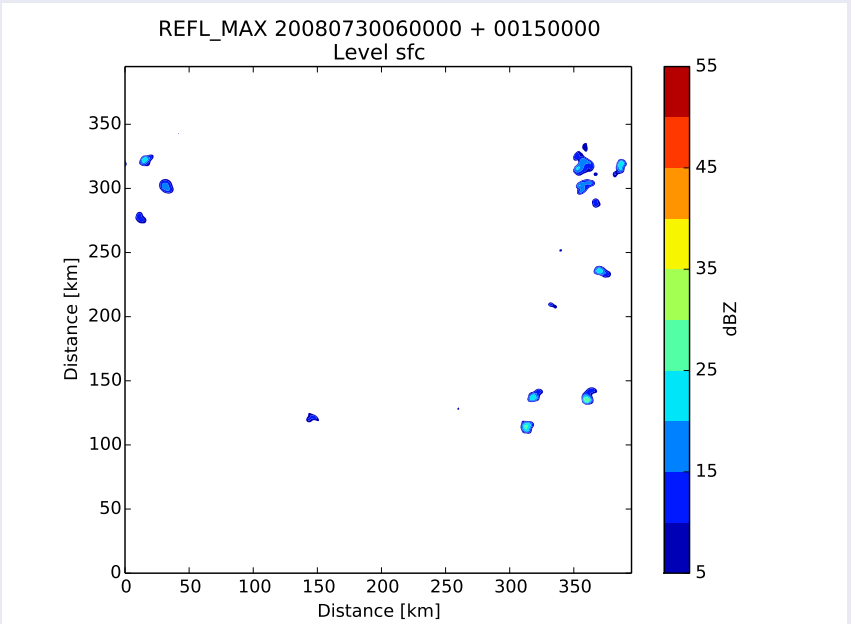


1 km Resolution vs. 2 km

$\Delta x = 2$ km

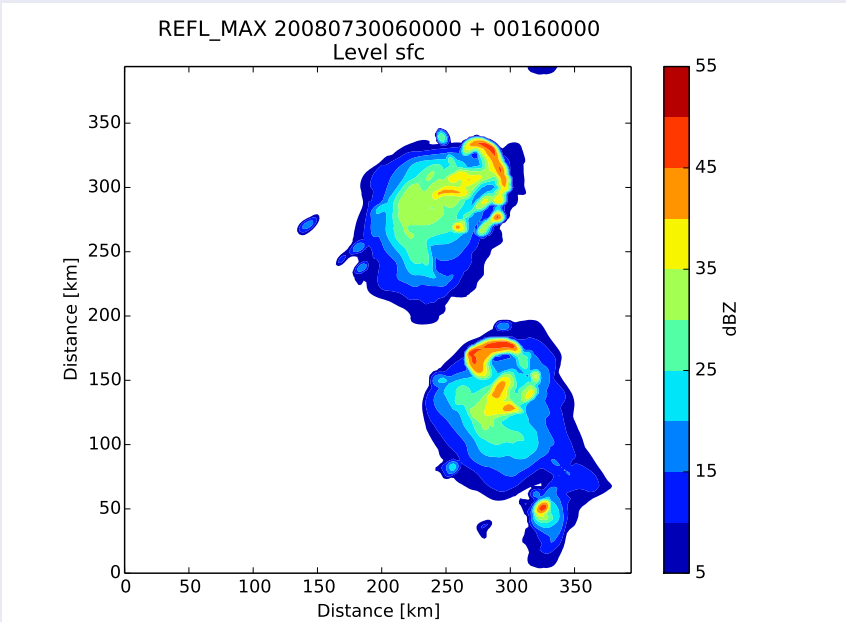


$\Delta x = 1$ km

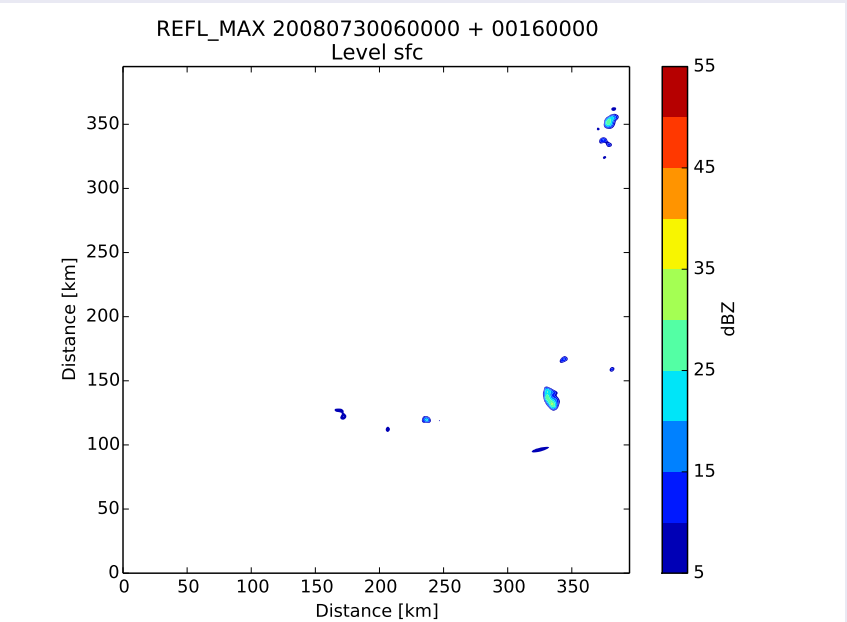


1 km Resolution vs. 2 km

$\Delta x = 2$ km

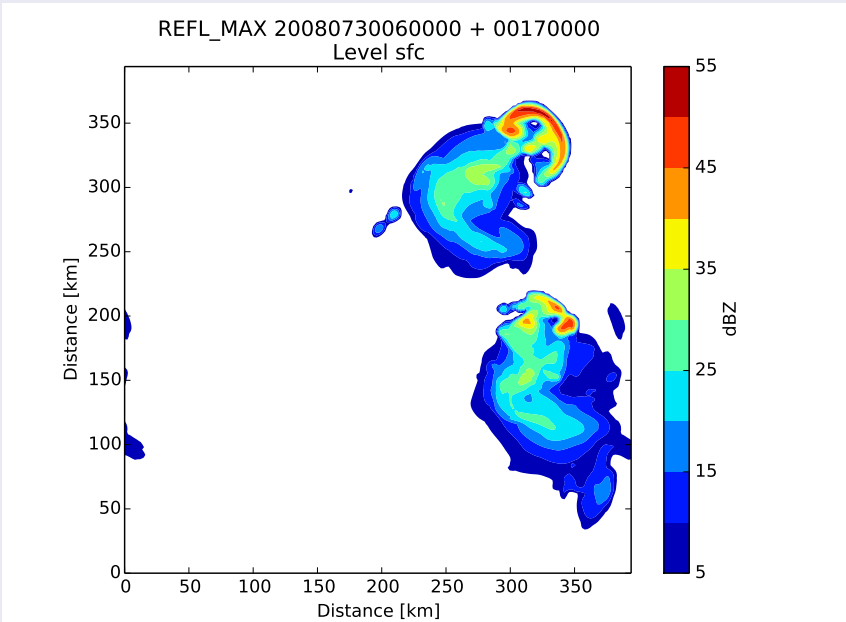


$\Delta x = 1$ km

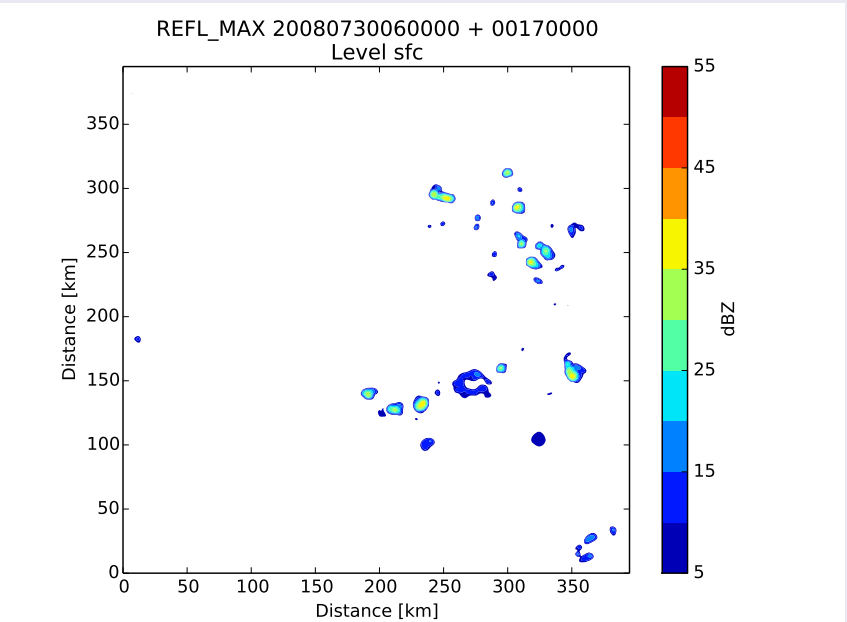


1 km Resolution vs. 2 km

$\Delta x = 2$ km

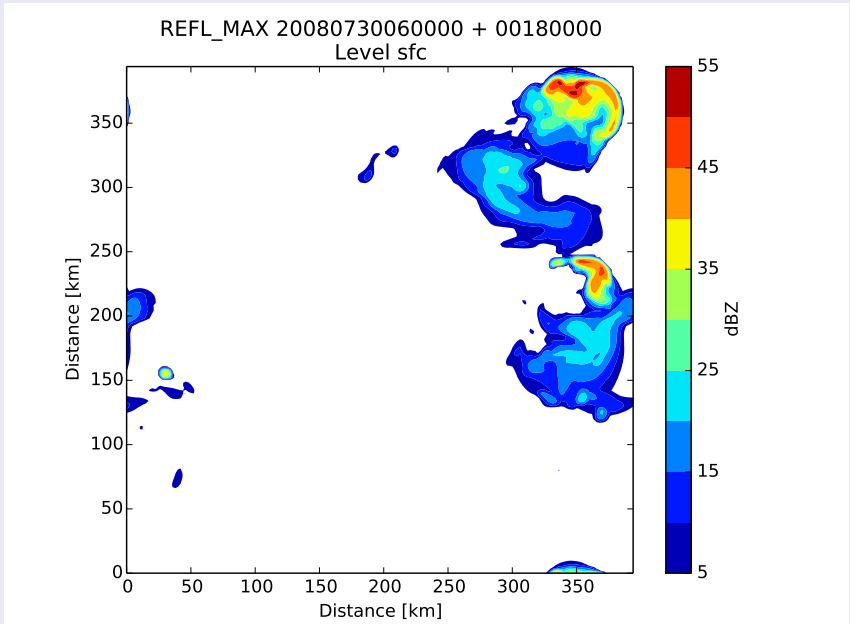


$\Delta x = 1$ km

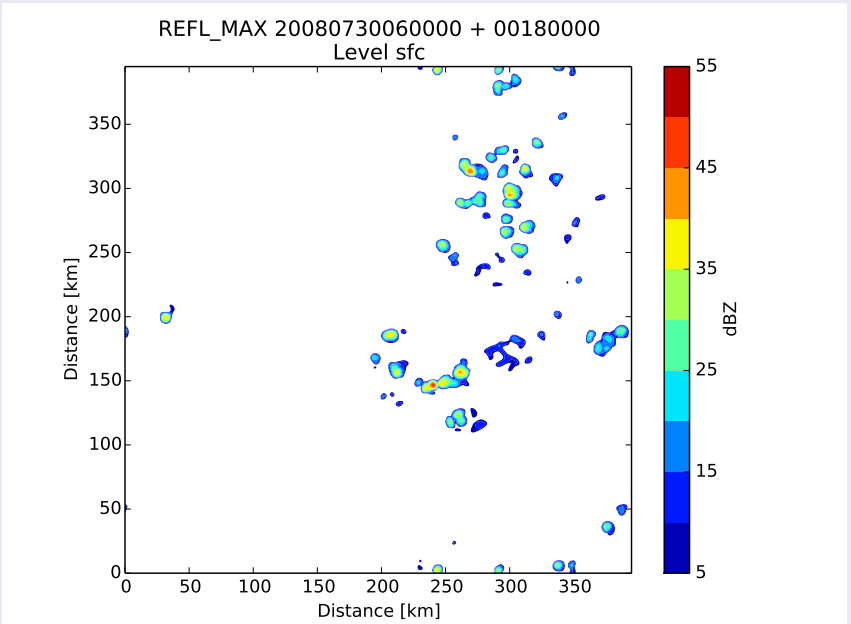


1 km Resolution vs. 2 km

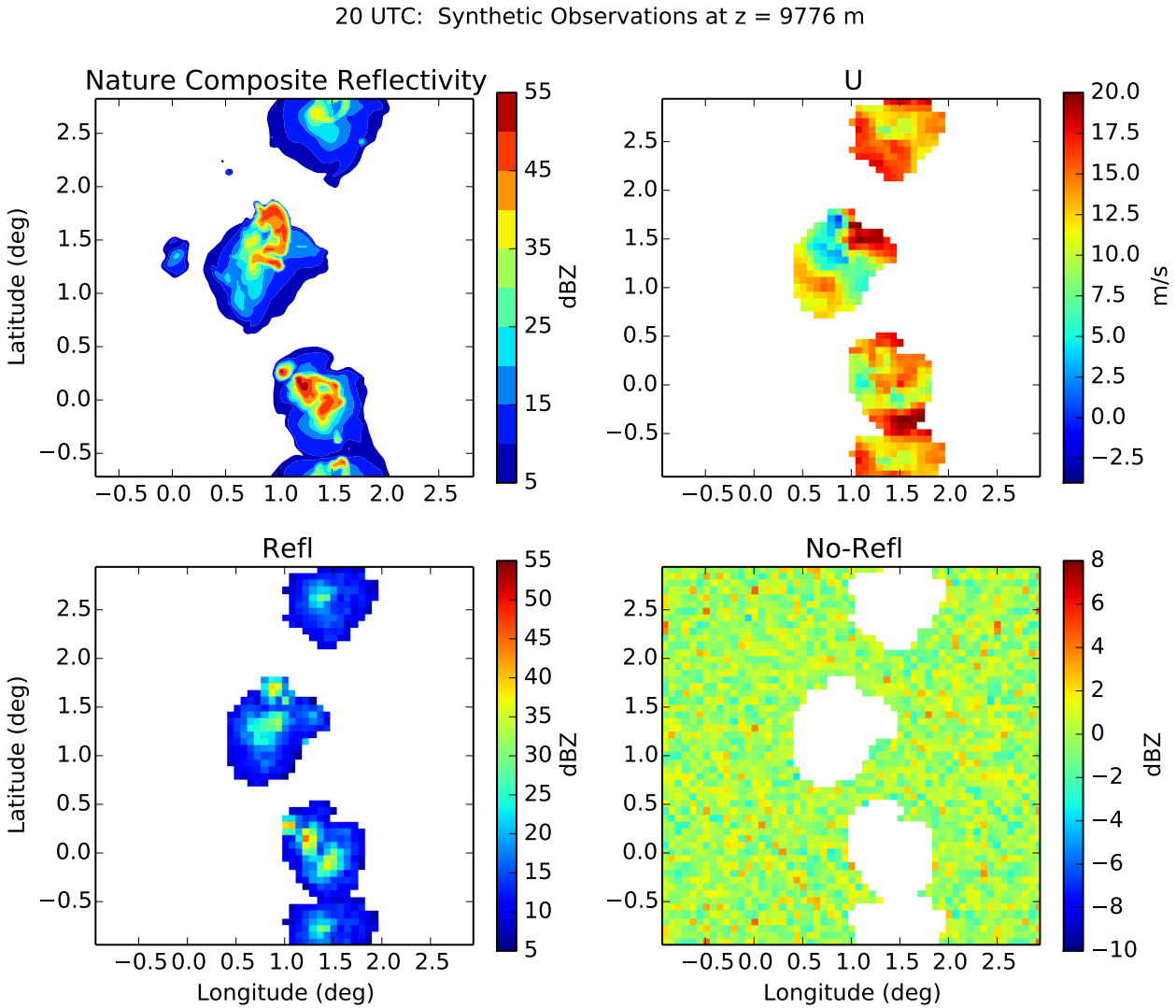
$\Delta x = 2$ km



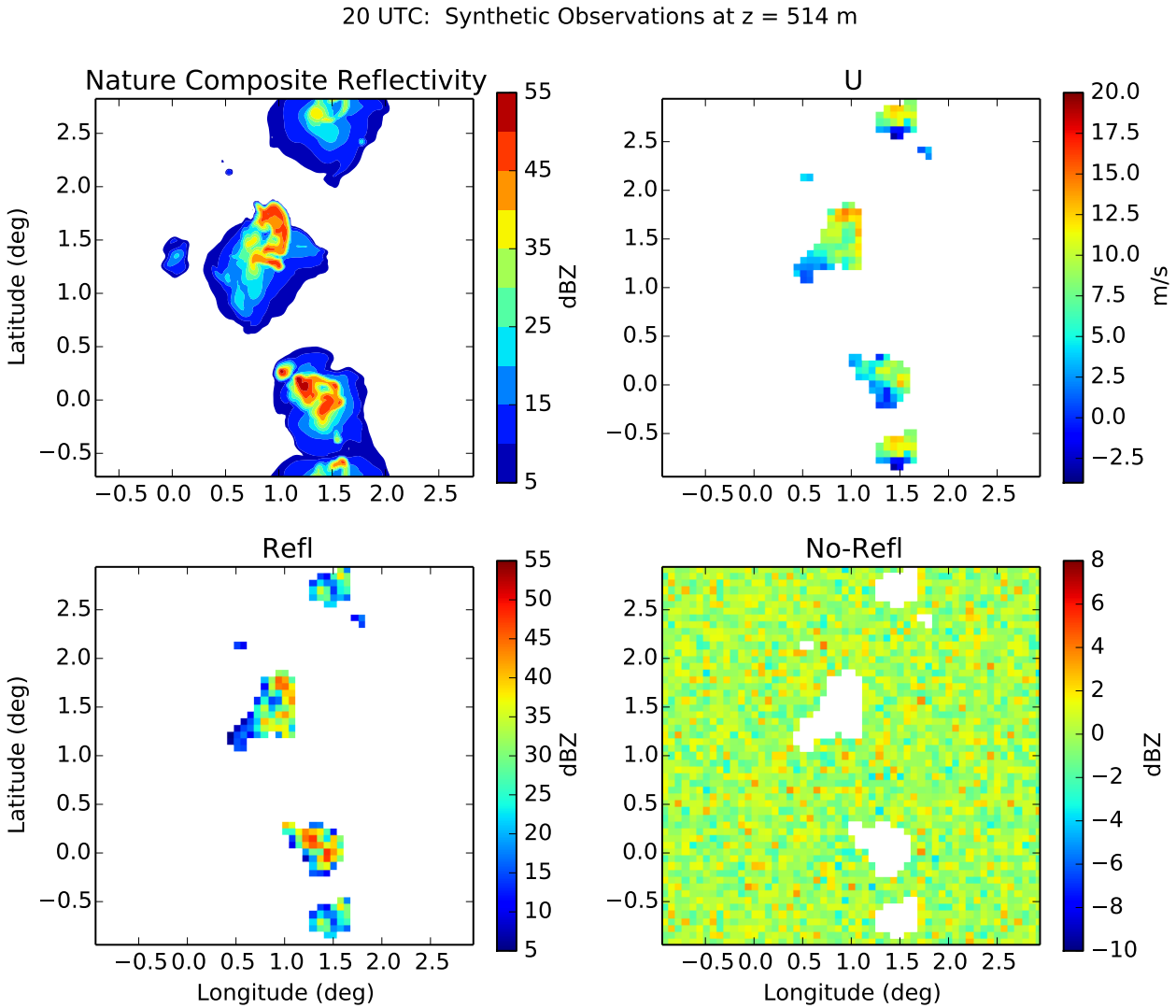
$\Delta x = 1$ km



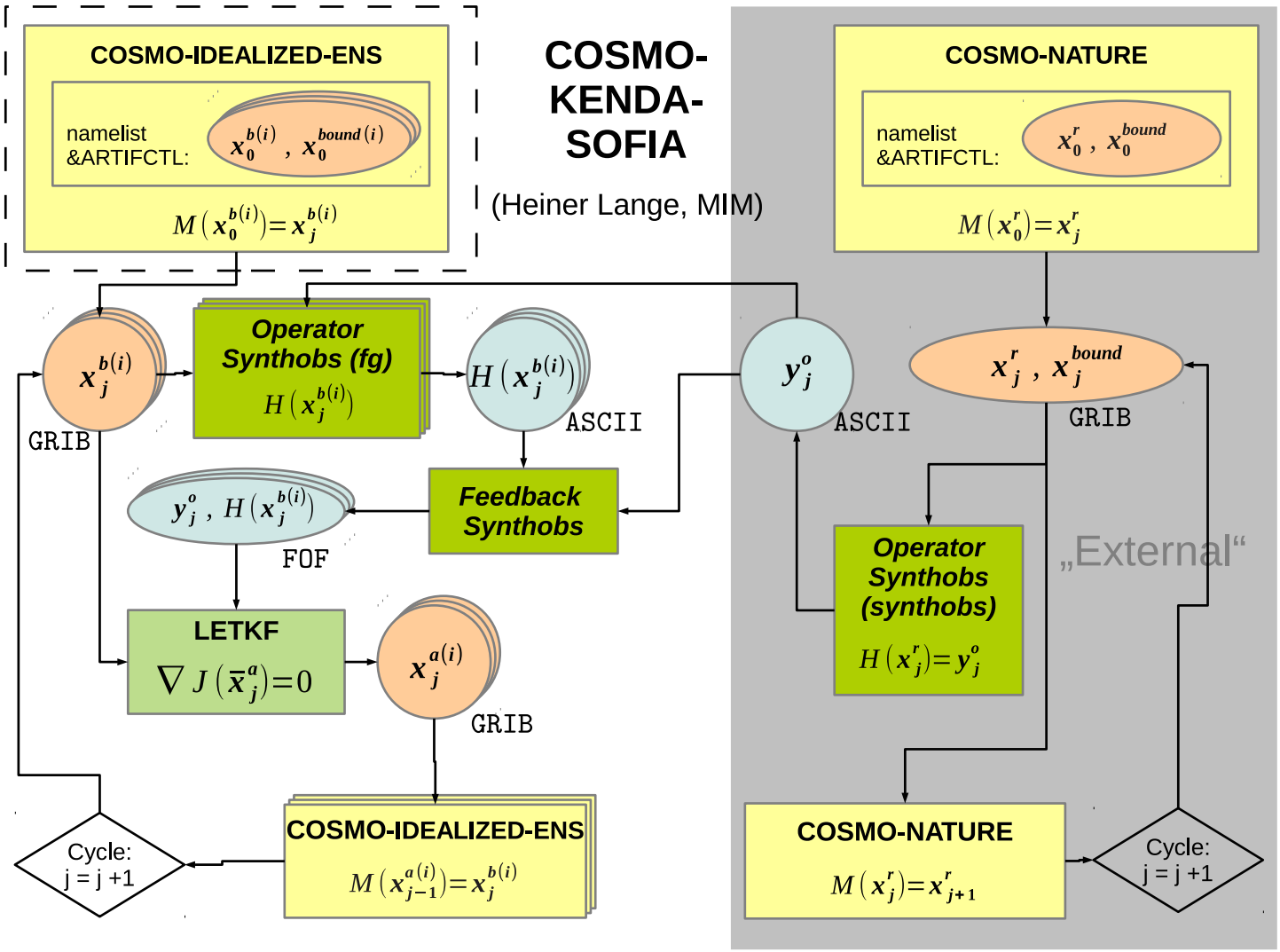
Synthetic Observations



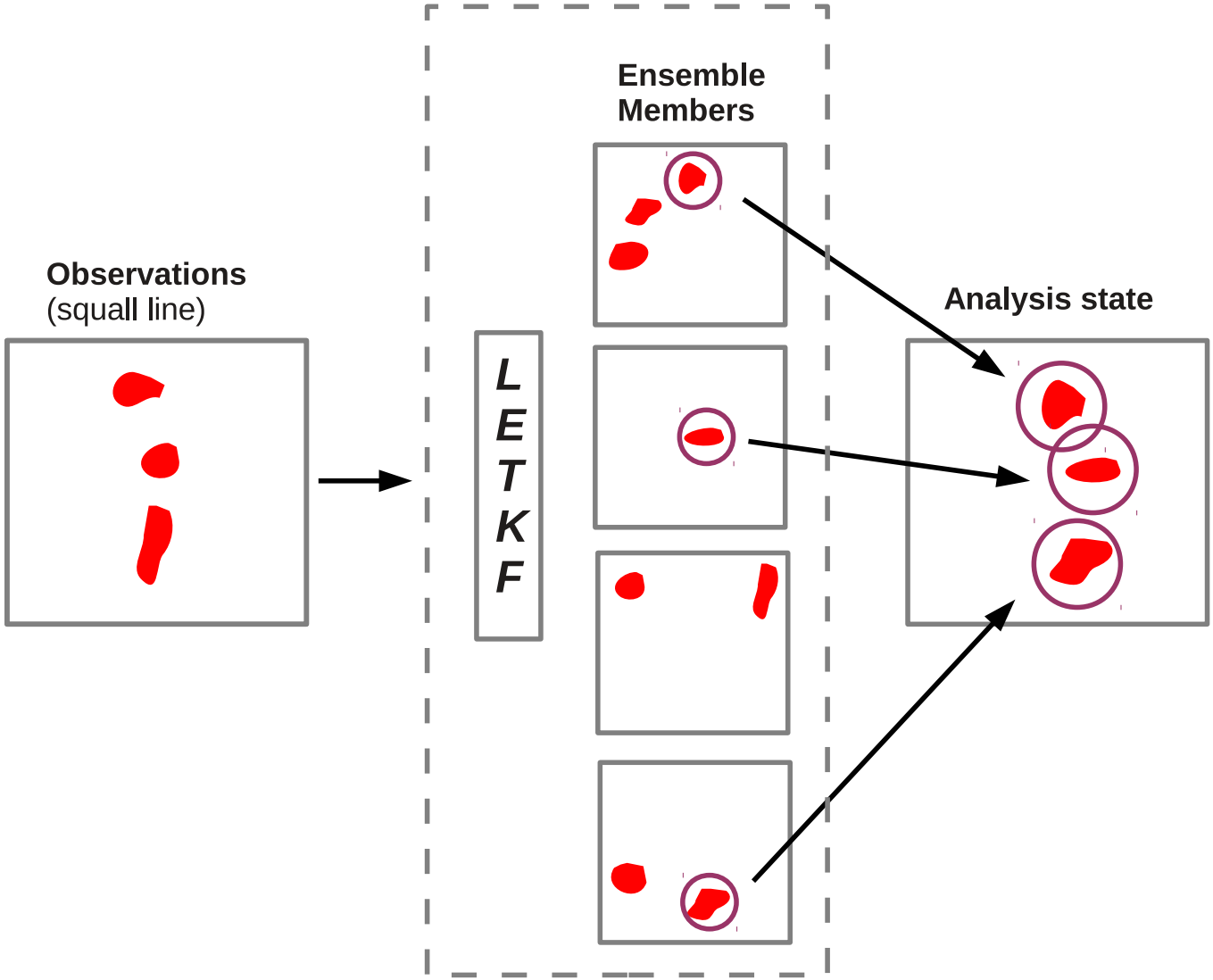
Synthetic Observations



LETKF-Flowchart



LETKF - How it works



LETKF-Setup 1/2

Assimilation setup

- 50 member ensemble (perfect model)
- `Python`-simulated observations of *radial wind* and *(no)-reflectivity*
- periodic LETKF-solution
- `Python`-cycling environment

LETKF-Setup 1/2

Assimilation setup

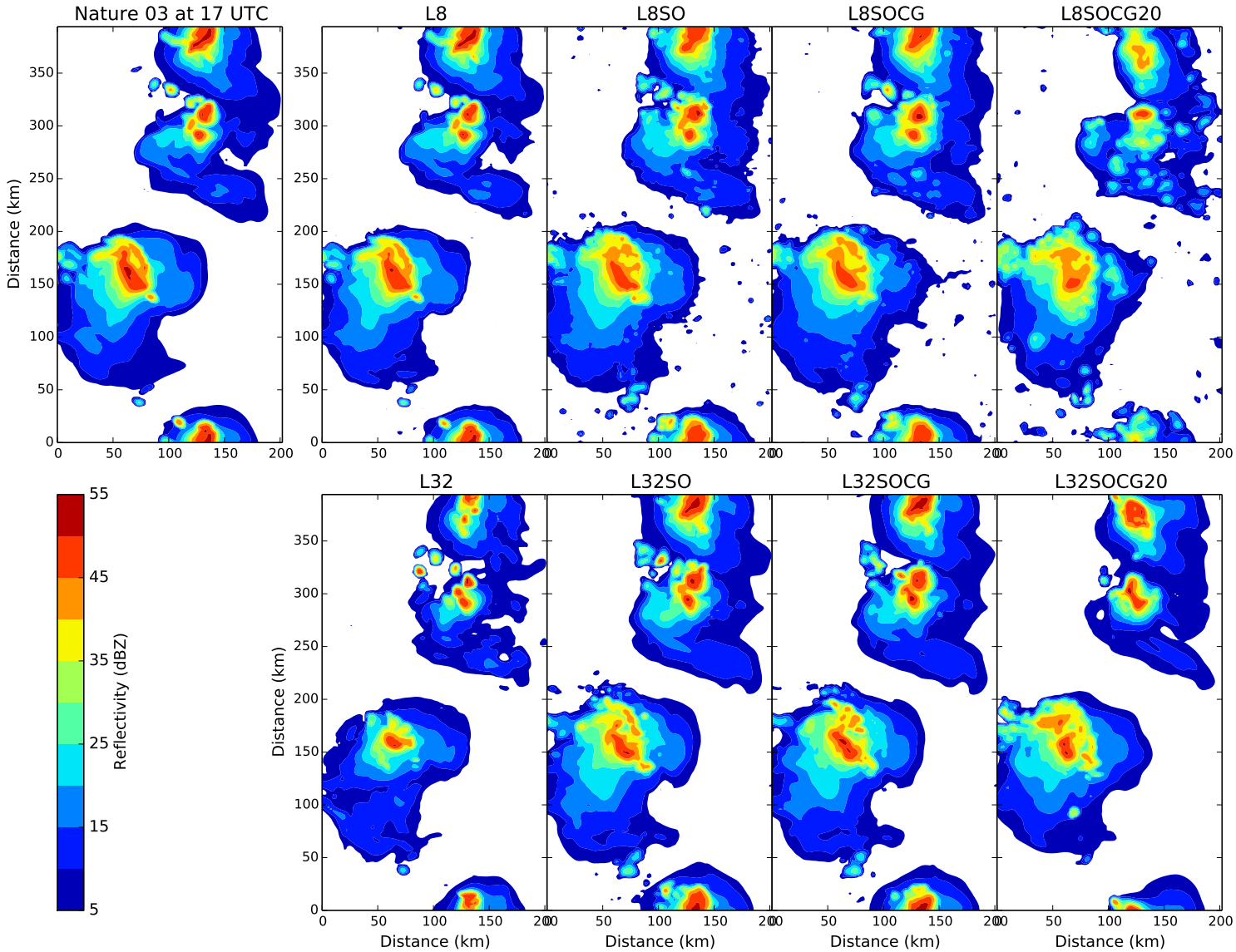
- 50 member ensemble (perfect model)
- `Python`-simulated observations of *radial wind* and *(no)-reflectivity*
- periodic LETKF-solution
- `Python`-cycling environment
- 3 hours cycled assimilation
- 3 hours ensemble forecast

LETKF-Setup 2/2

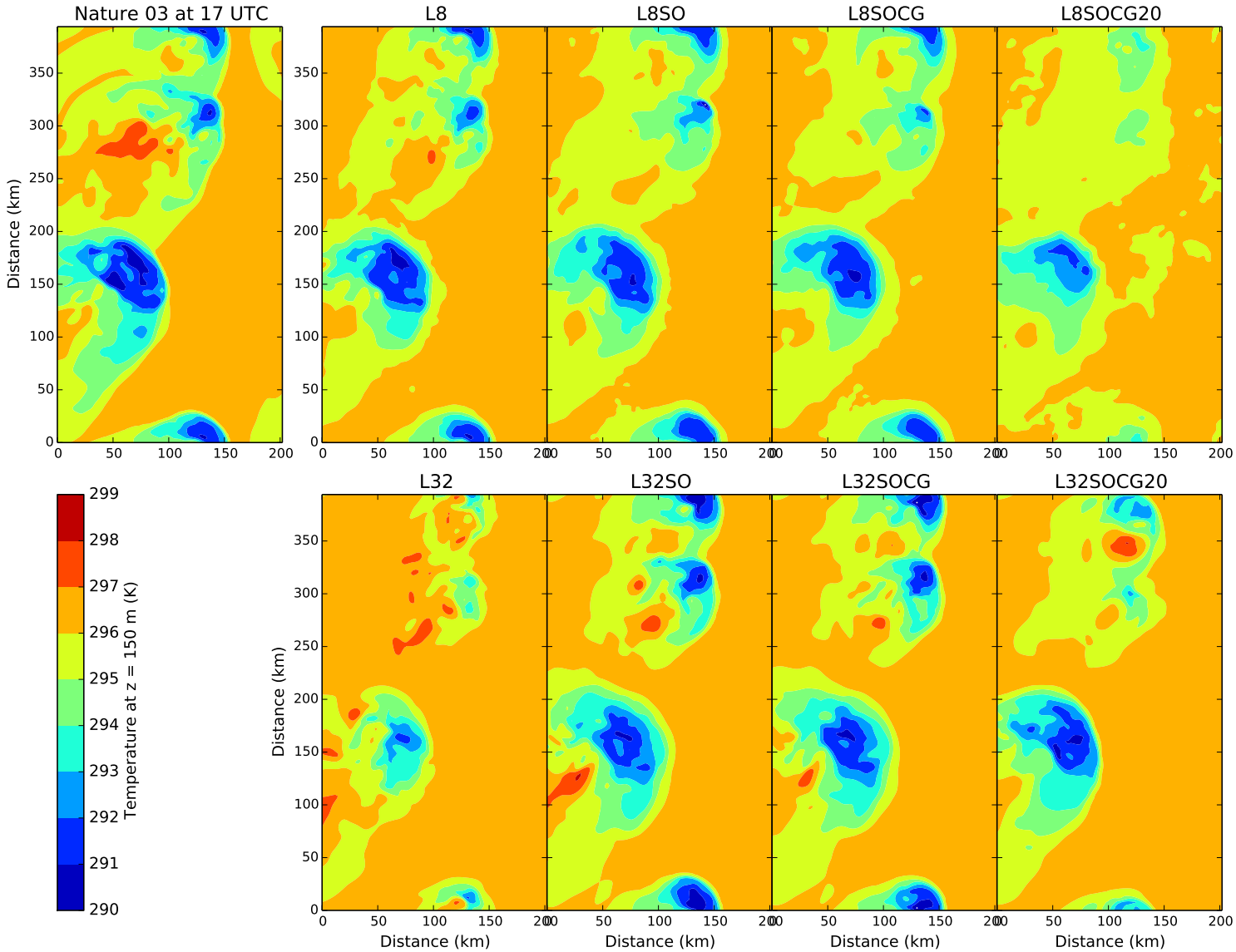
Assimilation settings

- Observation Resolution 2 and 8 km (superobservations SO)
- LETKF coarse grid on 2 and 8 km (coarse grid CG)
- Assimilation interval 5 and 20 minutes (20)
- Horizontal localization 8 and 32 km (L8 and L32)

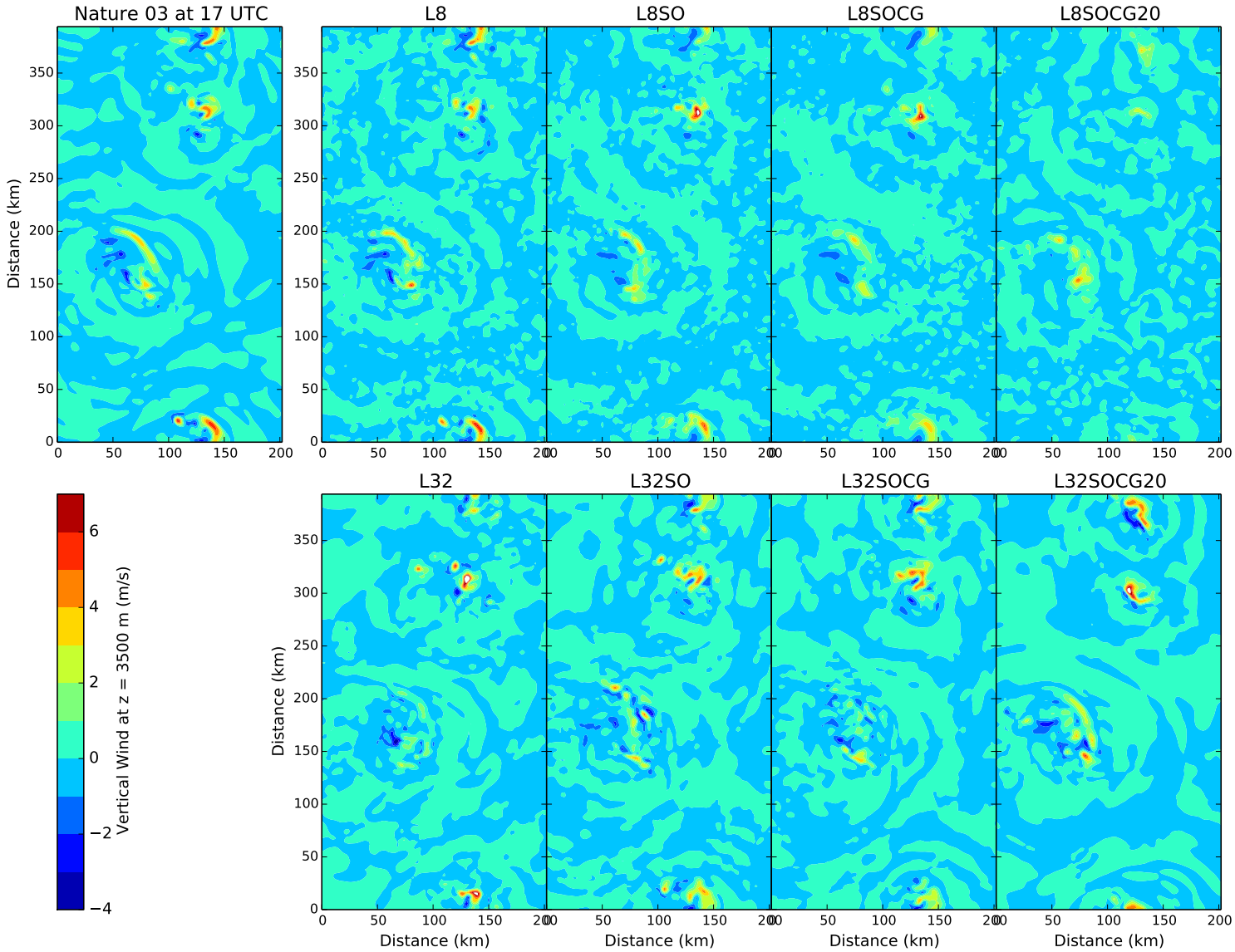
Assimilation Results: Nature vs. Analysis Ensemble Means



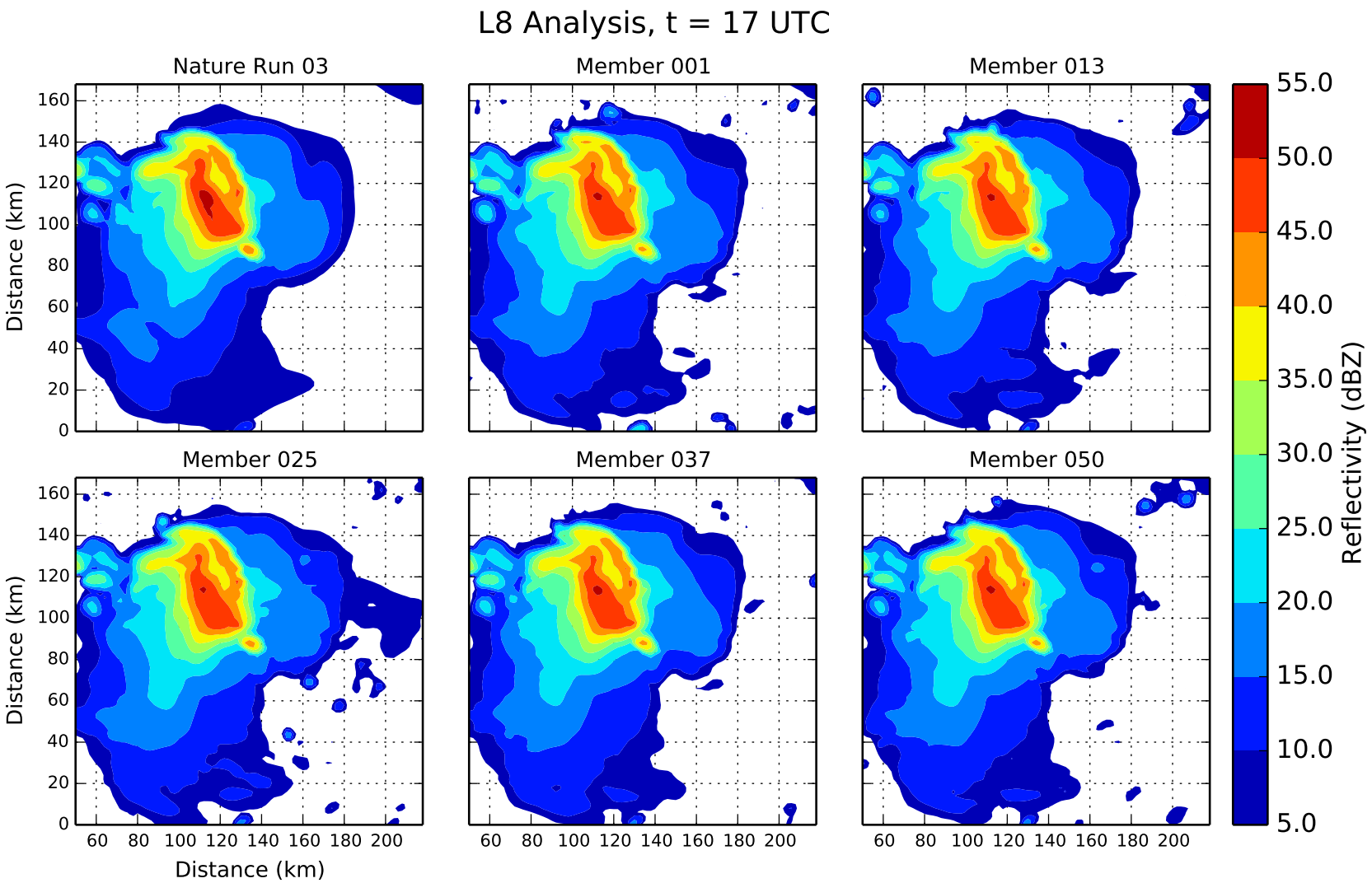
Assimilation Results: Nature vs. Analysis Ensemble Means



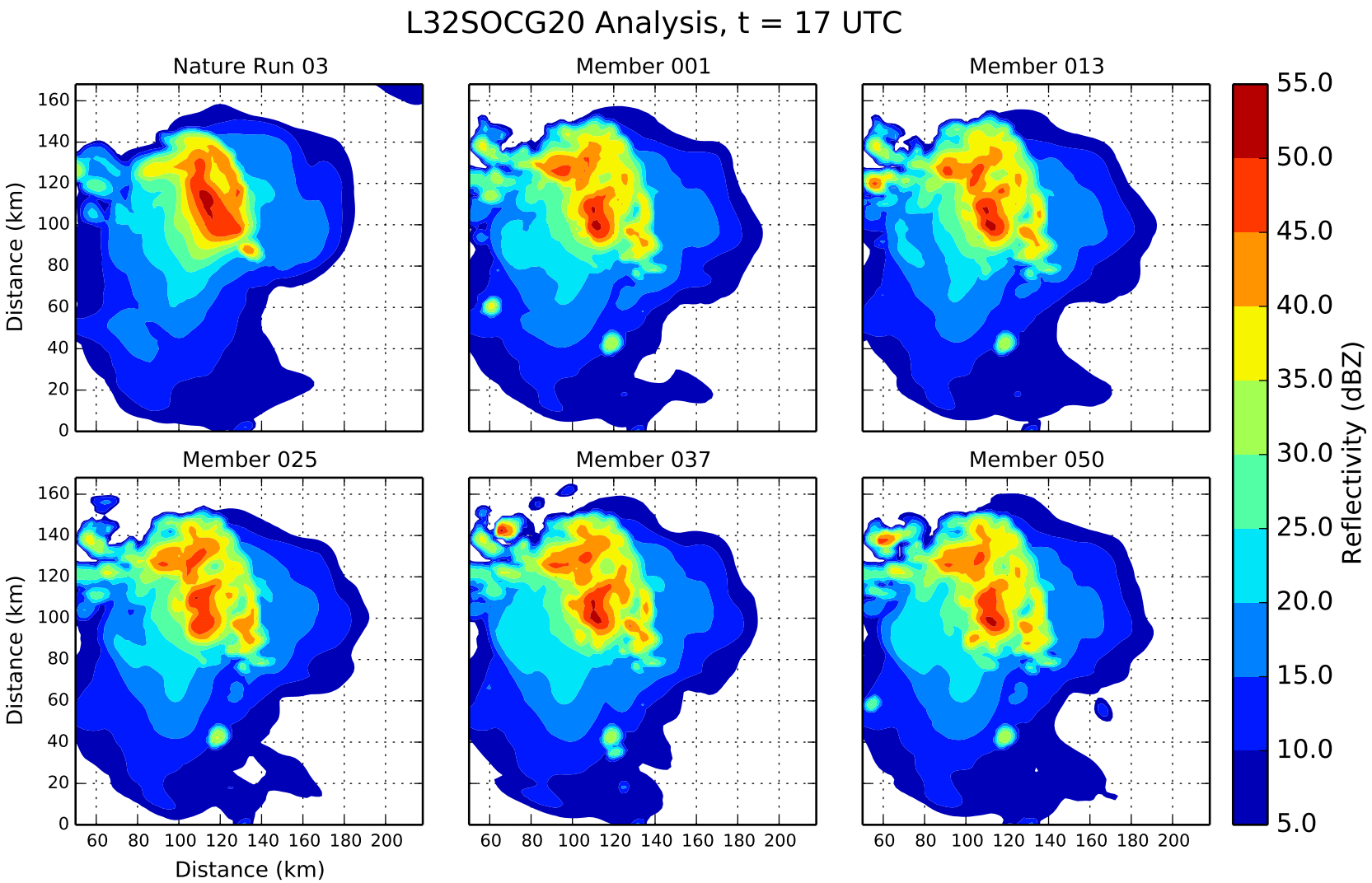
Assimilation Results: Nature vs. Analysis Ensemble Means



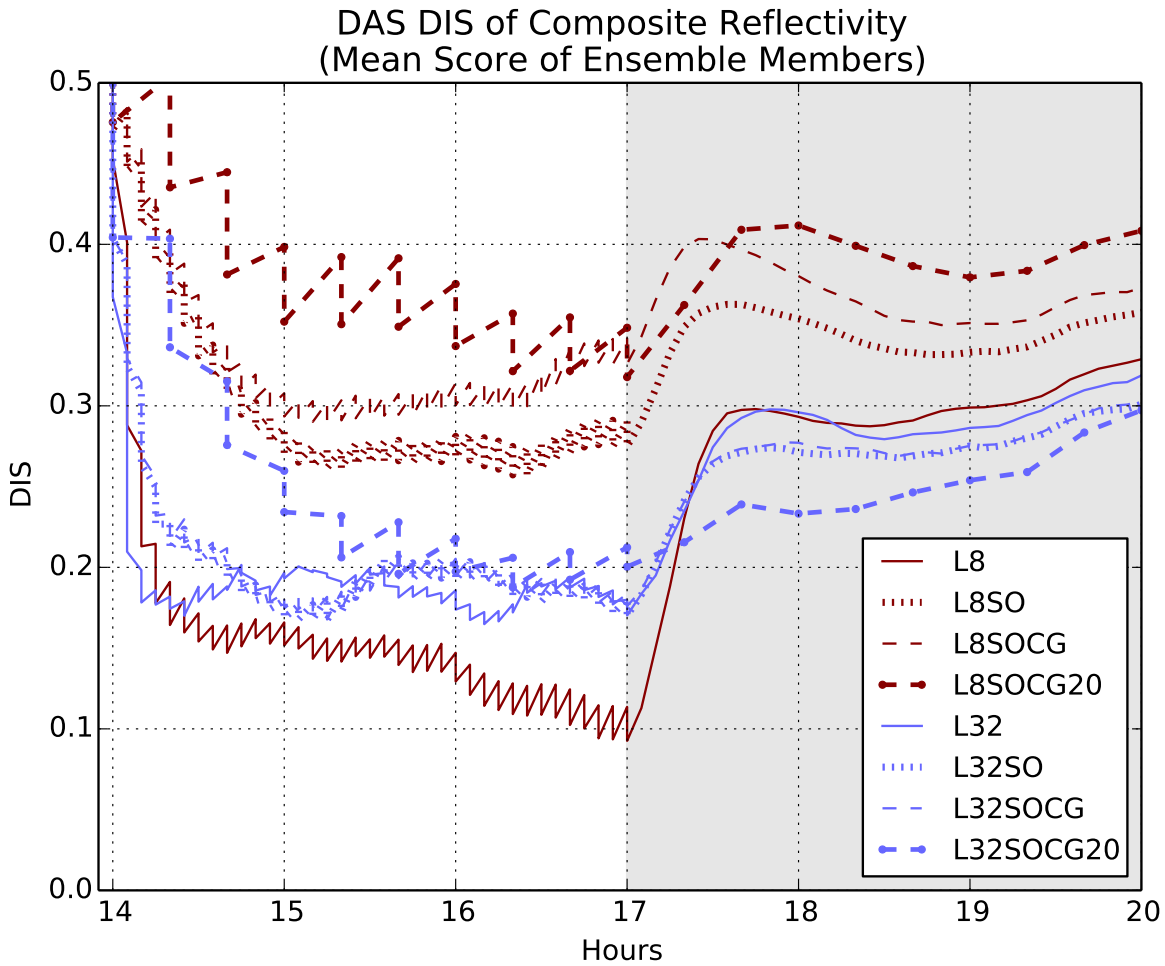
Analysis Members L8



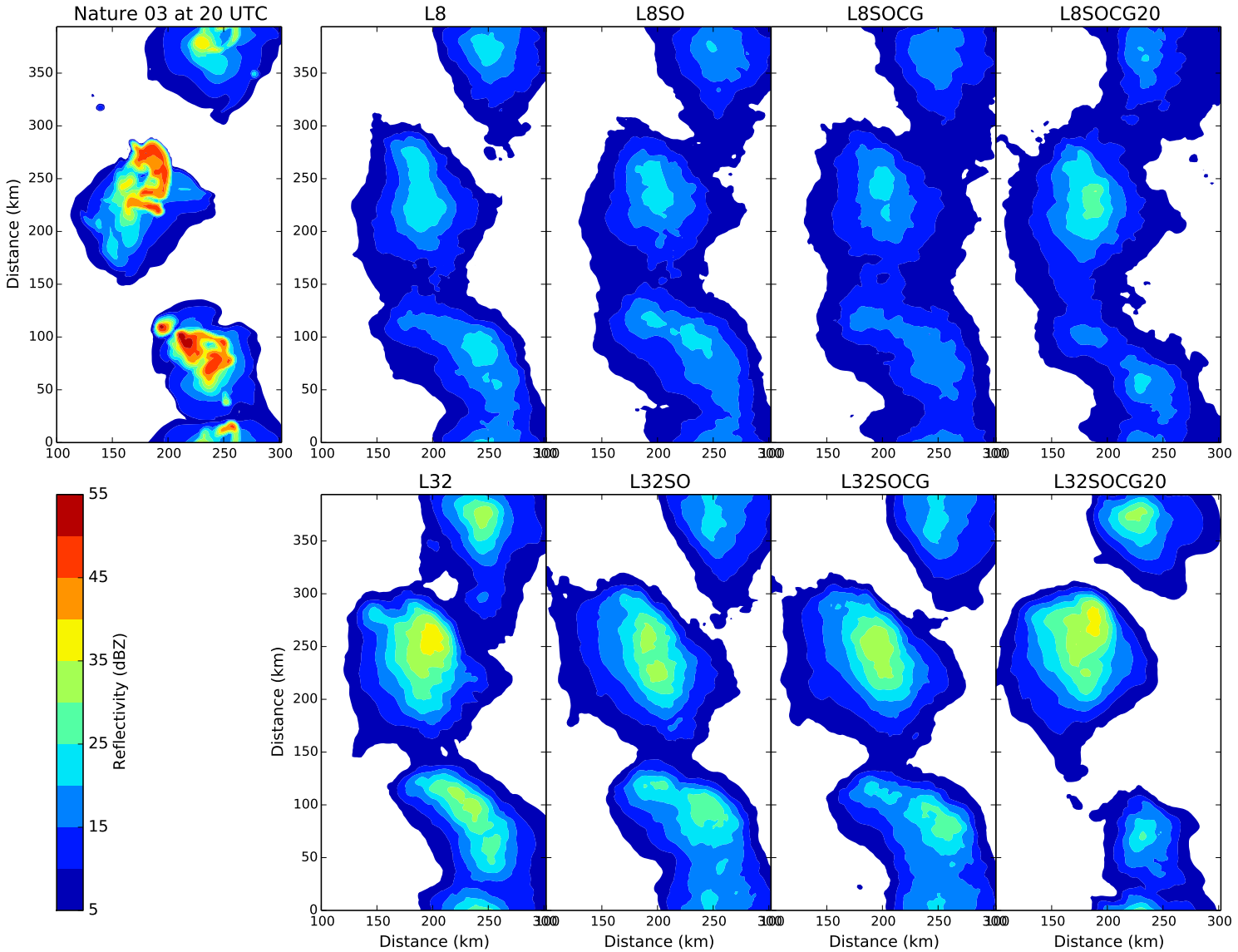
Analysis Members L32SOCG20



Forecast error growth (displacement of storms)



Forecast Results: Nature vs. Forecast Ensemble Means



Results and Conclusions

Assimilation Results

- Observation resolution of 4 x Δx_{model} sufficient for analysis
- Update of temperature through covariances: problematic!
- Less noise introduced with $\Delta t_{ass} = 20$ minutes

Results and Conclusions

Assimilation Results

- Observation resolution of 4 x Δx_{model} sufficient for analysis
- Update of temperature through covariances: problematic!
- Less noise introduced with $\Delta t_{ass} = 20$ minutes

Forecast Results

- Very fast displacement error growth
- Bias issues with spurious convection
- Forecast quality limited by:
  - Predictability of convection
  - Dynamical imbalance of initial states (= LETKF analyses)

Outlook

Further idealized studies

- Stability and imbalance in convective DA due to
  - strong ensemble convergence (observation errors and scale)
  - biases from strong localization
  - nonlinear observation operators
- Assimilation of convection in bad environmental soundings

"Advertisement"

Idealized KENDA

- Creative sandbox to study assimilation properties:
  - arbitrary types of convection (idealized COSMO)
  - free choice of ensemble spread / biases / model error
  - free choice of observation types, coverage and errors (very simple `Python` operator)
- KENDA-in-a-box: One `Python` cycling script (namelist-controlled) steers everything