Specific issues related to the use of satellite data in limited area models

Reinhold Hess, Francesca Di Giuseppe,

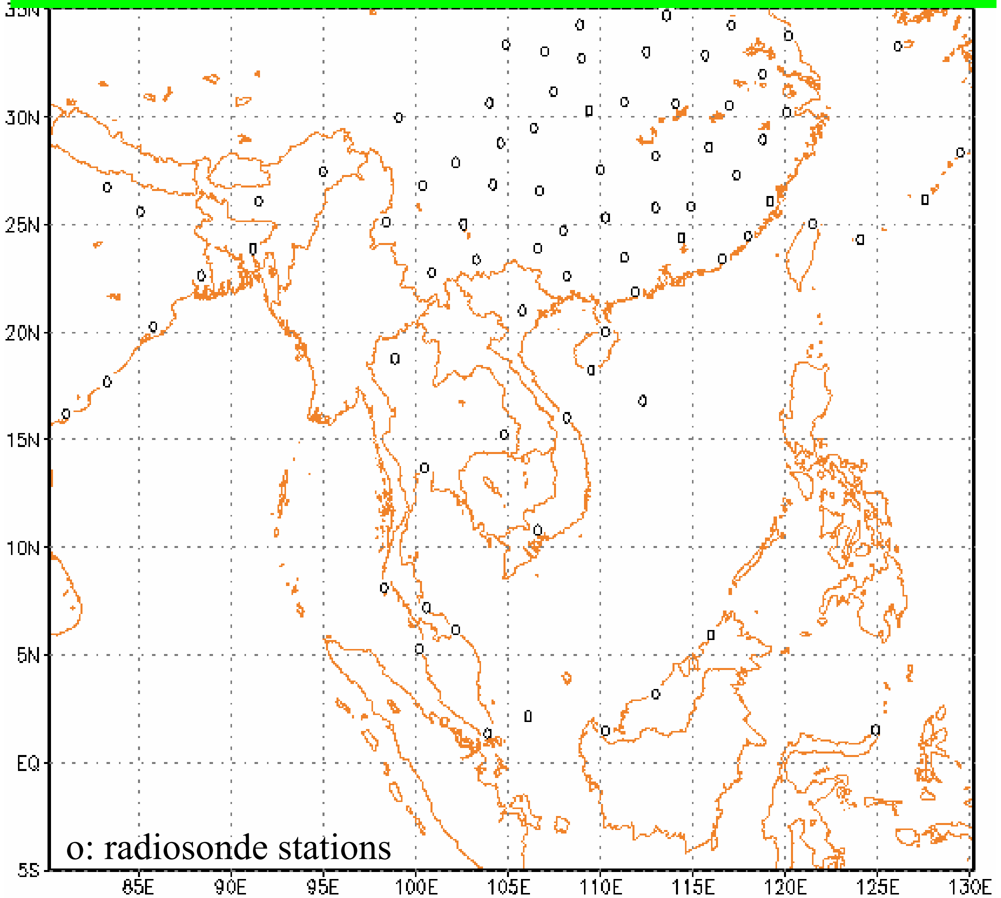
Christoph Schraff, Blazej Krzeminski

ECMWF Seminar

3-7 September 2007, Reading

Lokal Model HRM (in Vietnam)

Domain



HRM: Hydrostatic Regional Model,
resolution 28 km (14 km)
boundary and **initial** values from GME
provides operational forecasts in **Vietnam**

Project: Improvement of quantitative precipitation forecasts over Vietnam

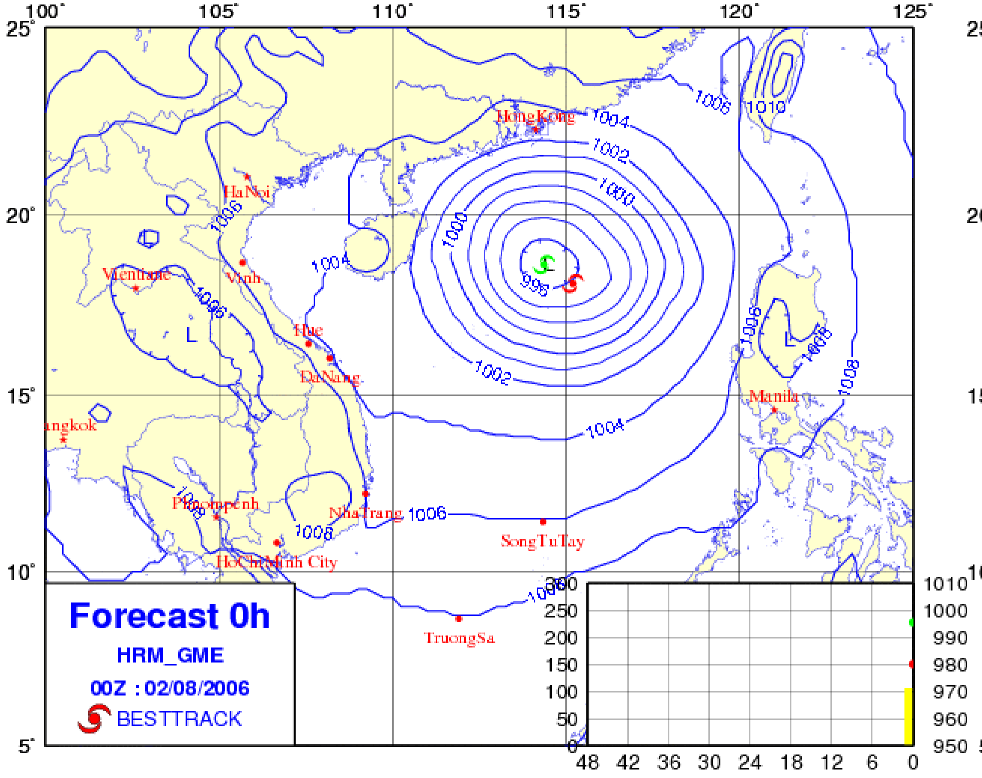
- application of **3d-Var**
- use of satellite radiances (**ATOVS**)
- adaptation and improvement of physics (convection)
- verification
- ...

Do we need own analysis for HRM?

- all data (almost) already in GME

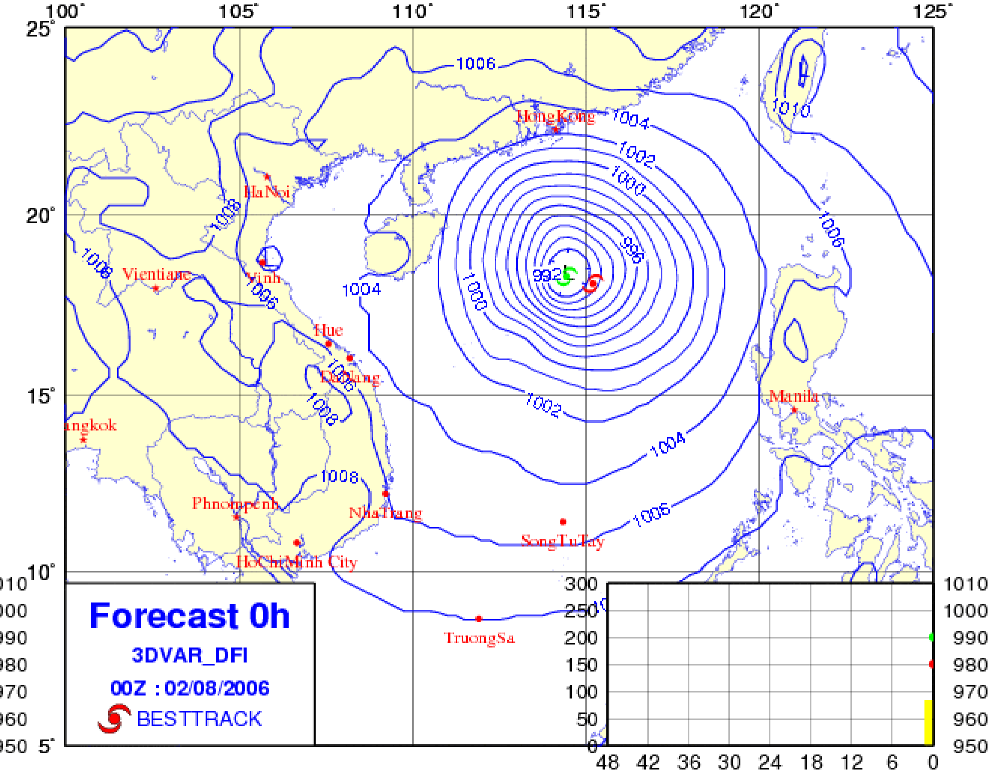
Forecast quality of regional models depend on ...

PRAPIROON storm track forecast



initial state interplated from GME

PRAPIROON storm track forecast



initial state from 3D-VAR (no satellites)

Source: Le Duc, Vietnam National University of Hanoi

Forecast quality of regional models depend on ...

- initial state
  - fit to observations (truth)
  - consistency with numerical model, small scale features in initial state (resolution, orography, vertical distribution of humidity, etc.)
- numerical model, resolution, approximations (e.g. hydrostacy), physics (e.g. convection), parameterisations
- quality of boundary values
- timeliness of forecast, (short range forecasting)

Limited area models ...

- high resolution (2-14 km)
- small scale structures in space and time (e.g. convection)
- delicate physics (e.g. steep orography, discontinuous solutions, bifurcations)
- limited predictability of small scale phenomena (computationally and physically)
- fewer constraints (e.g. hydrostacy, geostrophy)
- need to use asynchronous and high frequent observations (e.g. SEVIRI/MSG, radar)
- limited area (driven by boundary values of embedding models)
- over land (use of radiances over land)

Specific issues of limited area models for the use of satellite data (radiances)

- **assimilation scheme**
  - provide initial value fit and consistent with the limited area model
  - use temporally and spacially highly resolved observations (background and observations errors and correlations)
  - complex and situation dependent statistics for background and observation errors, flow dependence, more critical vertical structure (temp/hum/wind)
- **bias correction**
  sample size, representativity of samples, overfitting, choice of predictors
- **first guess above model top**
- **tuning** of observations (thinning)

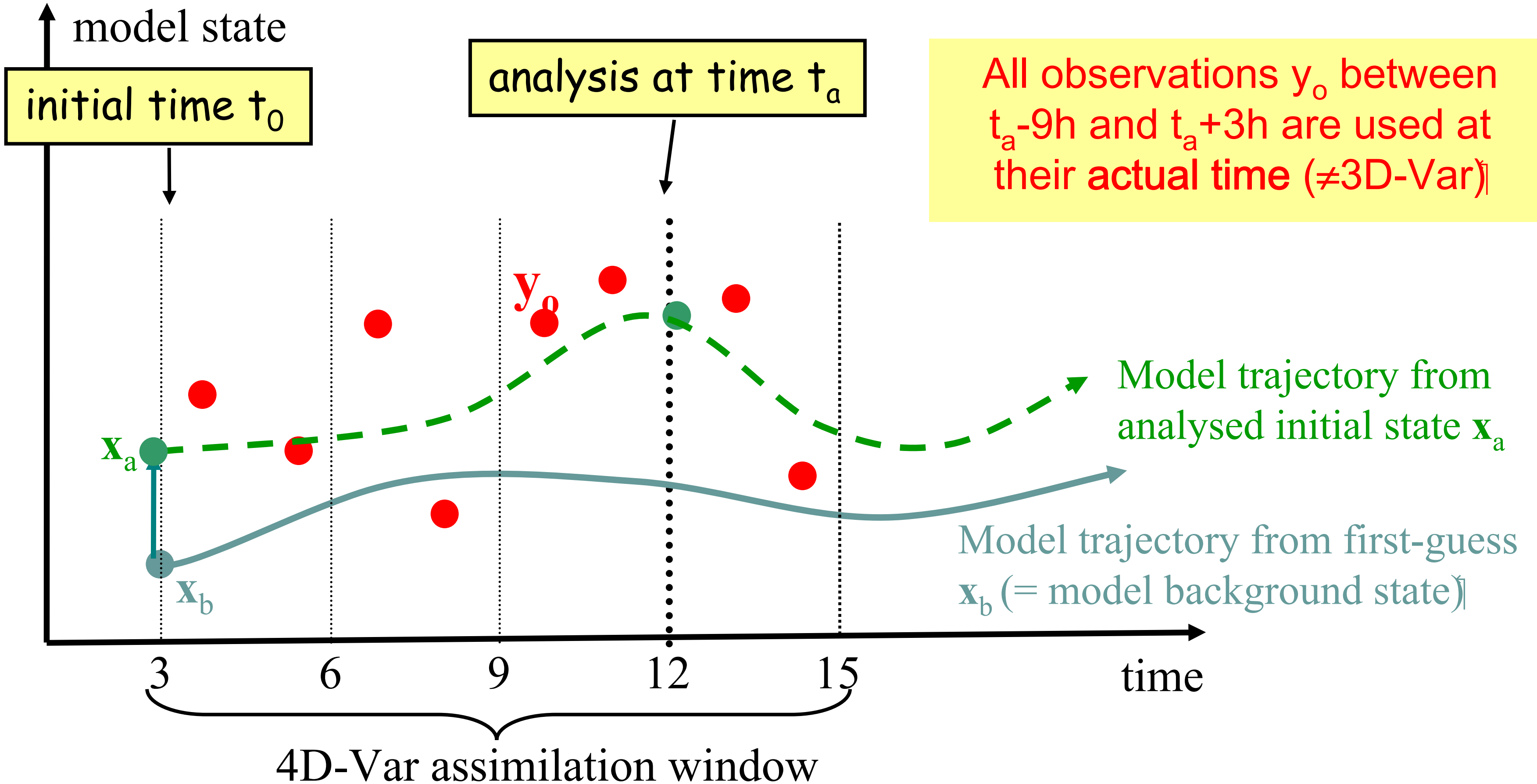
- **use of data over land** (surface emissivity, higher resolution of surface conditions)
- **verification**
  statistical representativity of results, influence of boundary values

4D-Var approach: initial state minimises misfits of model trajectory to observations and deviation from first guess.

3D-Var: as 4D-Var, but all observations valid for time of analysis, no computation of trajectory during minimisation

3D-Var-FGAT: use trajectory at observation time for first guess, but keep innovations constant

All information has to be reflected in initial state (analysis)

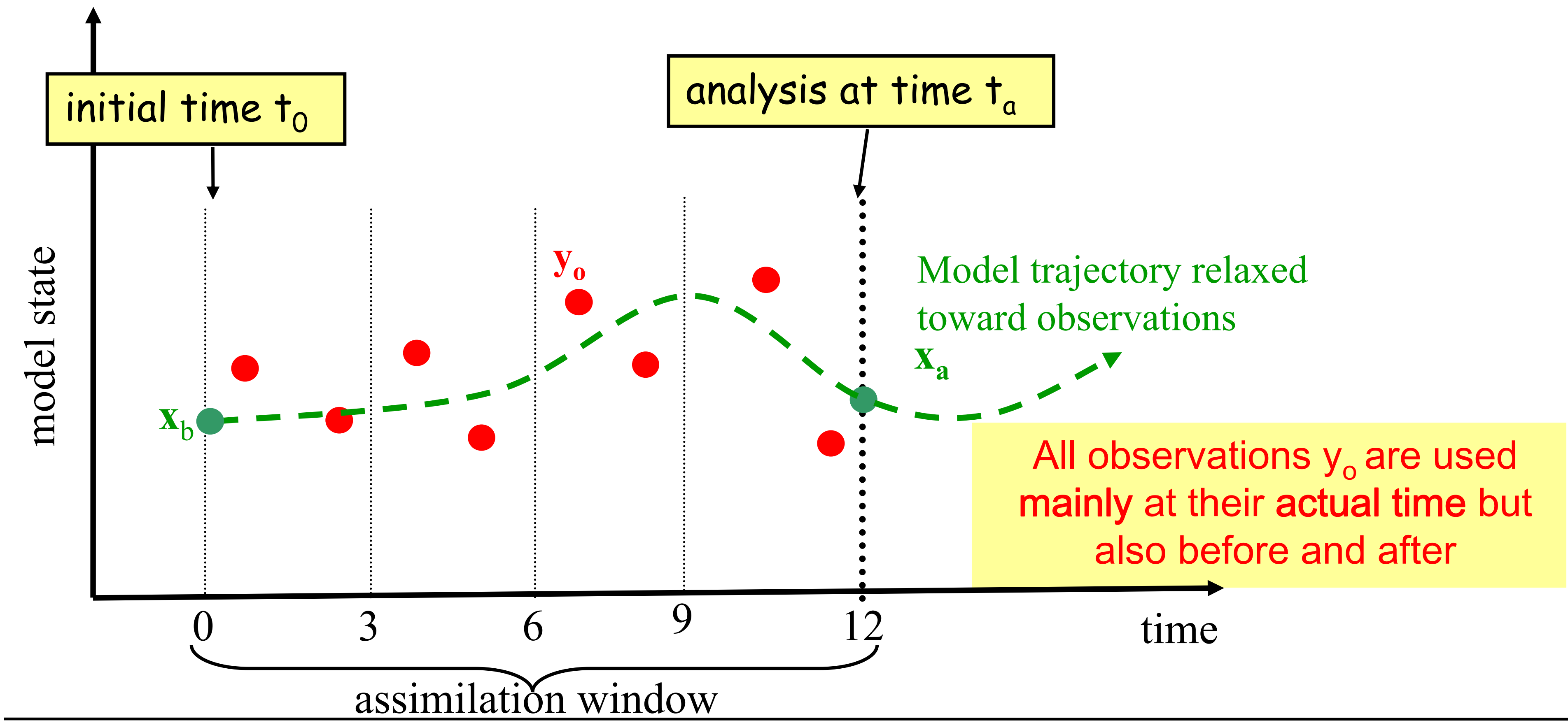


Nudging approach (Newtonian Relaxation Scheme):

The model trajectory is nudged in every time step towards the observations with special terms additional to the model dynamics (nudging towards observations during forecast).

The sizes of the terms depend on the distance to the observations and on the time difference between observation and current model time.

Difficult to use nonlinear observations



Pros and Cons for limited area models

4D-Var:

+ use of asynchronous observations
+ nonlinear observations can be used
+ consistent mathematical framework (obs and fg errors)
- solutions are less smooth and predictable (physically)
- physics are more complex (tangent linear and adjoint)
- specification of background errors is more difficult (boundary, fewer constraints)
- time consuming

3D-Var (EnKF):

+ consistent mathematical framework
+ combination with ensemble methods
- use of observations at time of analysis
- requires initialisation

Nudging:

+ unsteady solutions, complicated physics
+ use of anynchronous **and** high frequent observations
+ fast, provide timely forecast
+ no initialisation required
+ combination with ensemble methods
- use of nonlinear observation operators
- no consistent mathematical framework, lots of tuning required

COSMO-Project:

Assimilation of satellite radiances with 1D-Var and Nudging

COSMO: **Co**nsortium for **S**mall-scale **M**odeling
(Germany*, Switzerland, Italy*, Greece, Poland*, Romania)

* Member of Project

Goals of Project:

- Assimilate radiances (SEVIRI, ATOVS, AIRS/IASI) in COSMO-EU
- Explore the use of nonlinear observation operators with **Nudging**

1DVAR + Nudging

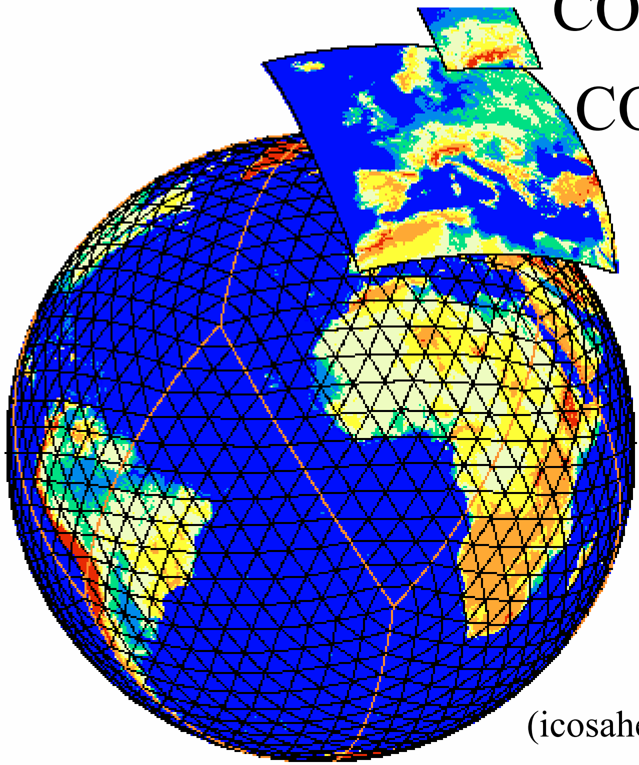
i.e. RETRIEVE temperature and humidity profiles and then nudge them as “pseudo”-observations

Lokal Model COSMO-EU (LME)

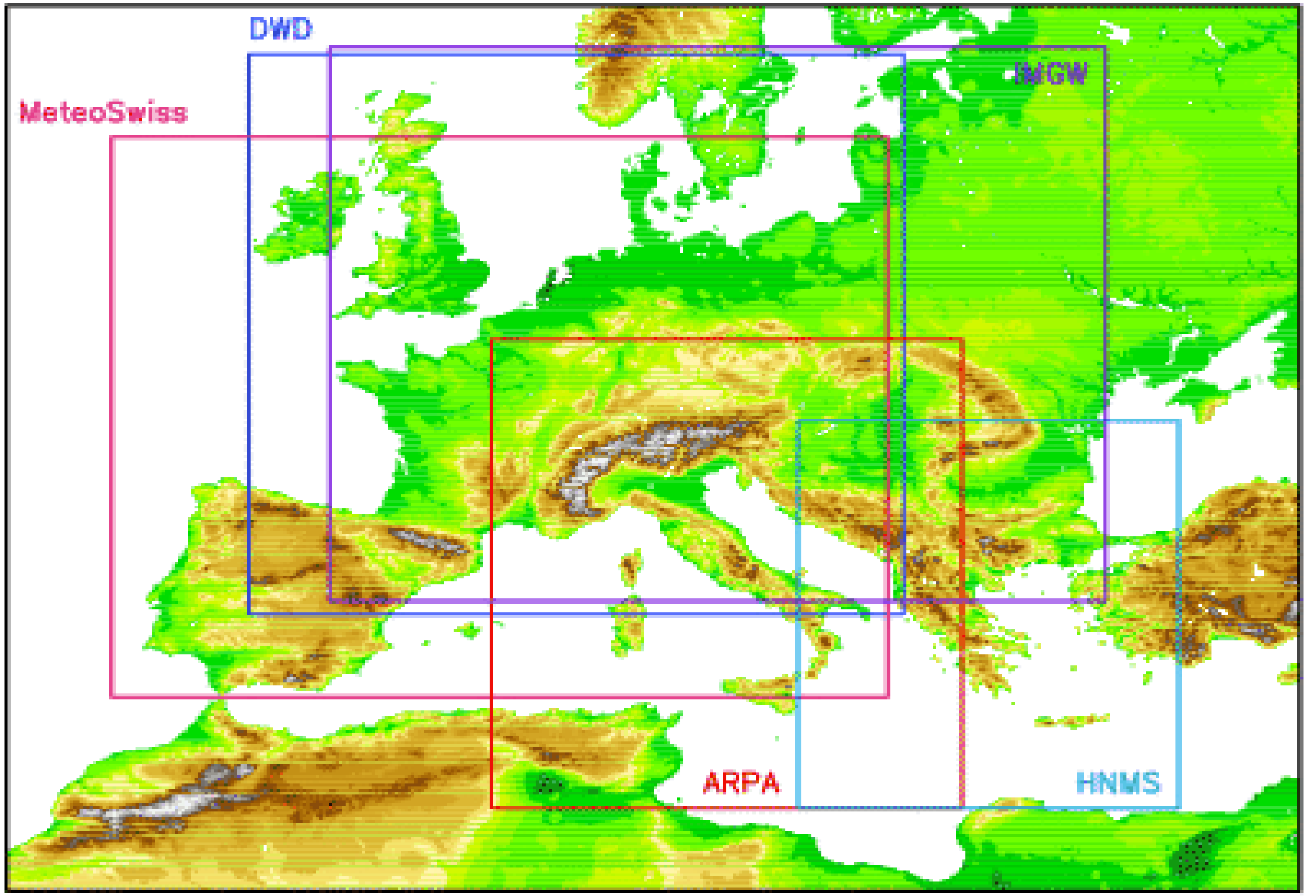
COSMO-EU (regional model): non-hydrostatic, rotated latitude-longitude grid, mesh size 7km
terrain following hyb. coordinate, 40 layers up to 20hPa
forecast range: 78 for initial dates 00, 12
prognostic cloud ice, prognostic rain
boundary values from GME

Analysis: continuous nudging
observations: conv., AIREP, AMDAR, ACARS
cutoff: 2h30
variational soil moisture analysis

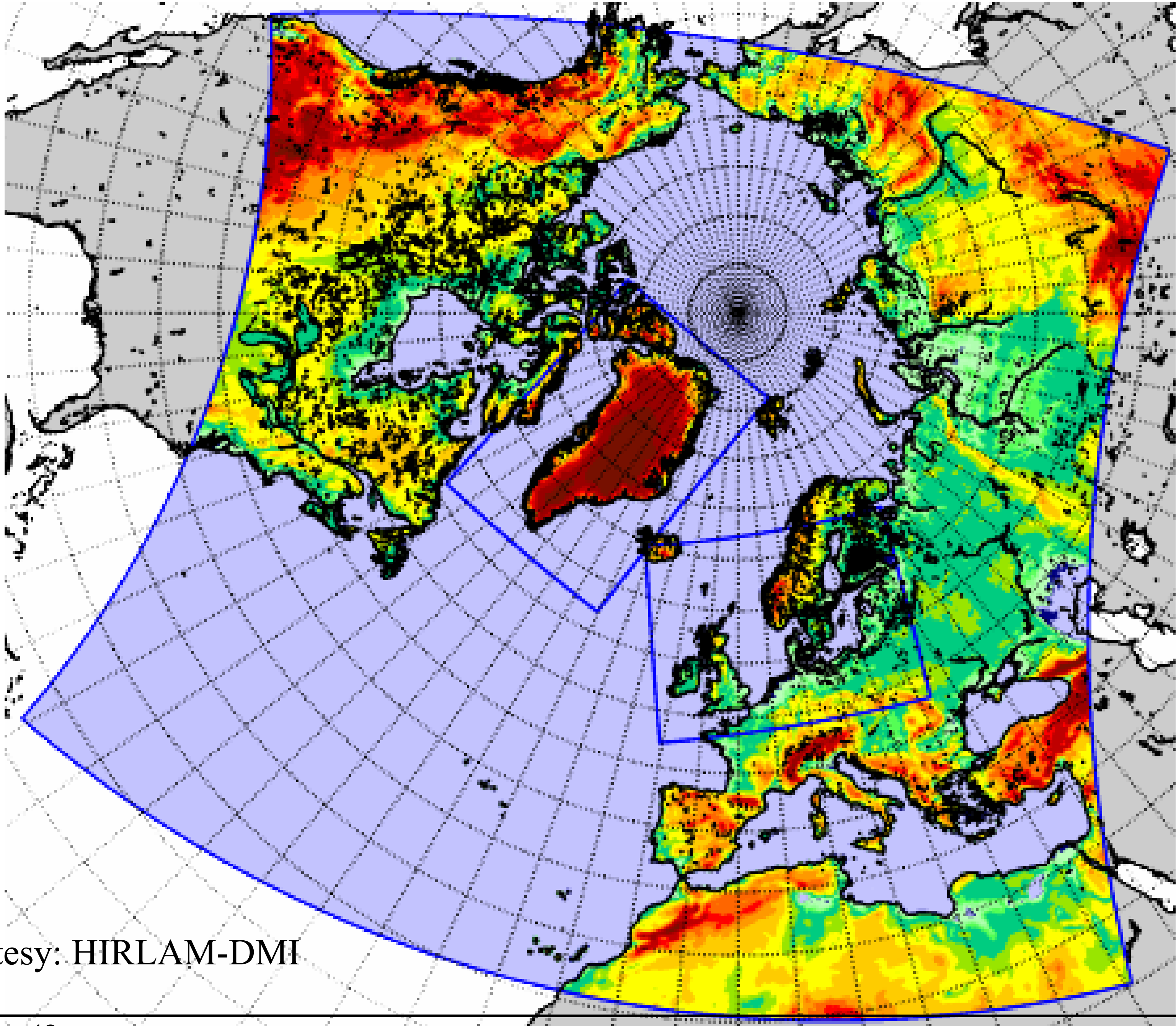
COSMO-DE (local model): similar to COSMO-EU
mesh size 2.8km, explicit convection
latent head nudging of radar reflectivities
boundary values from COSMO-EU

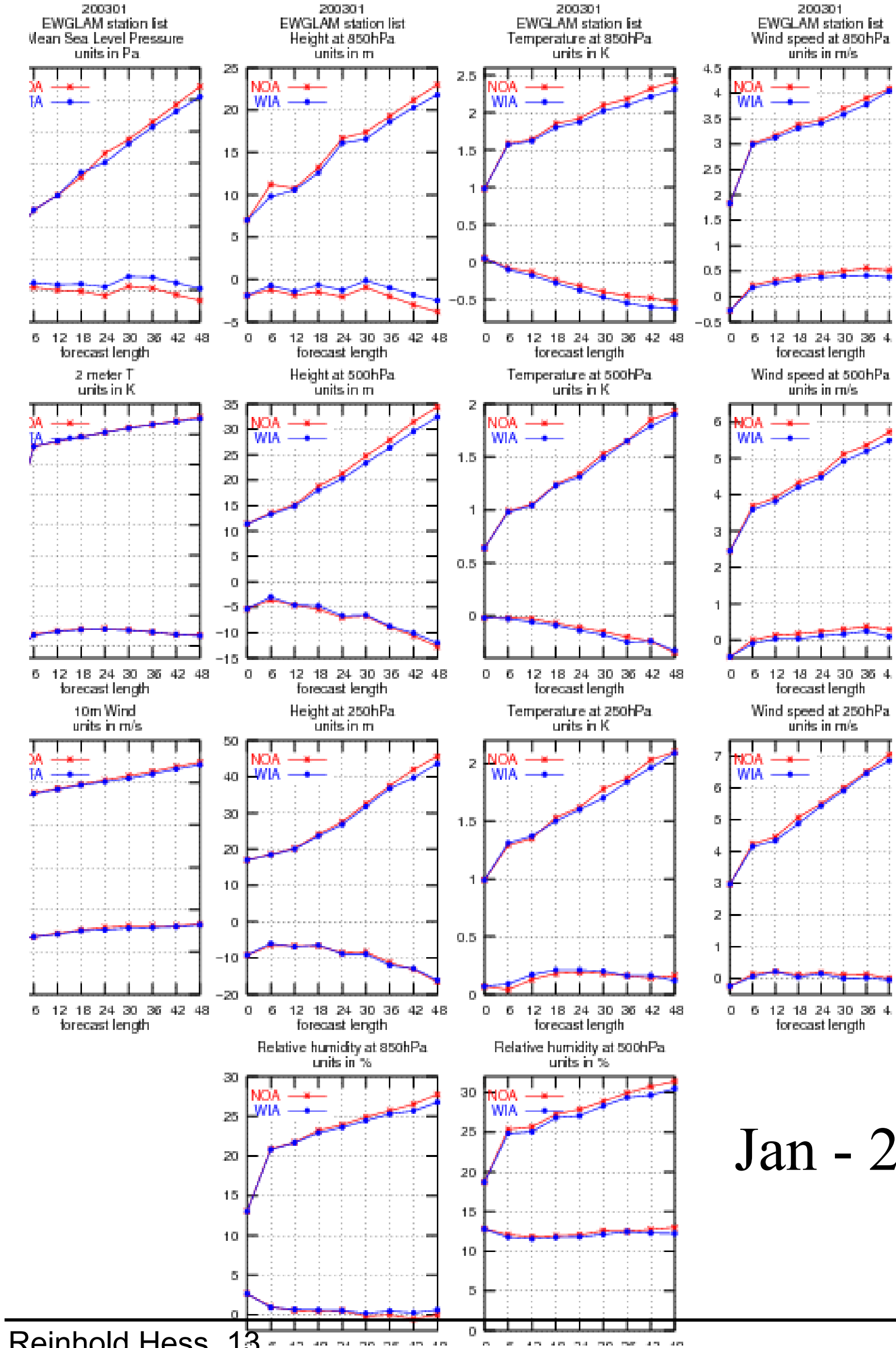


Applications of COSMO-model within COSMO-group

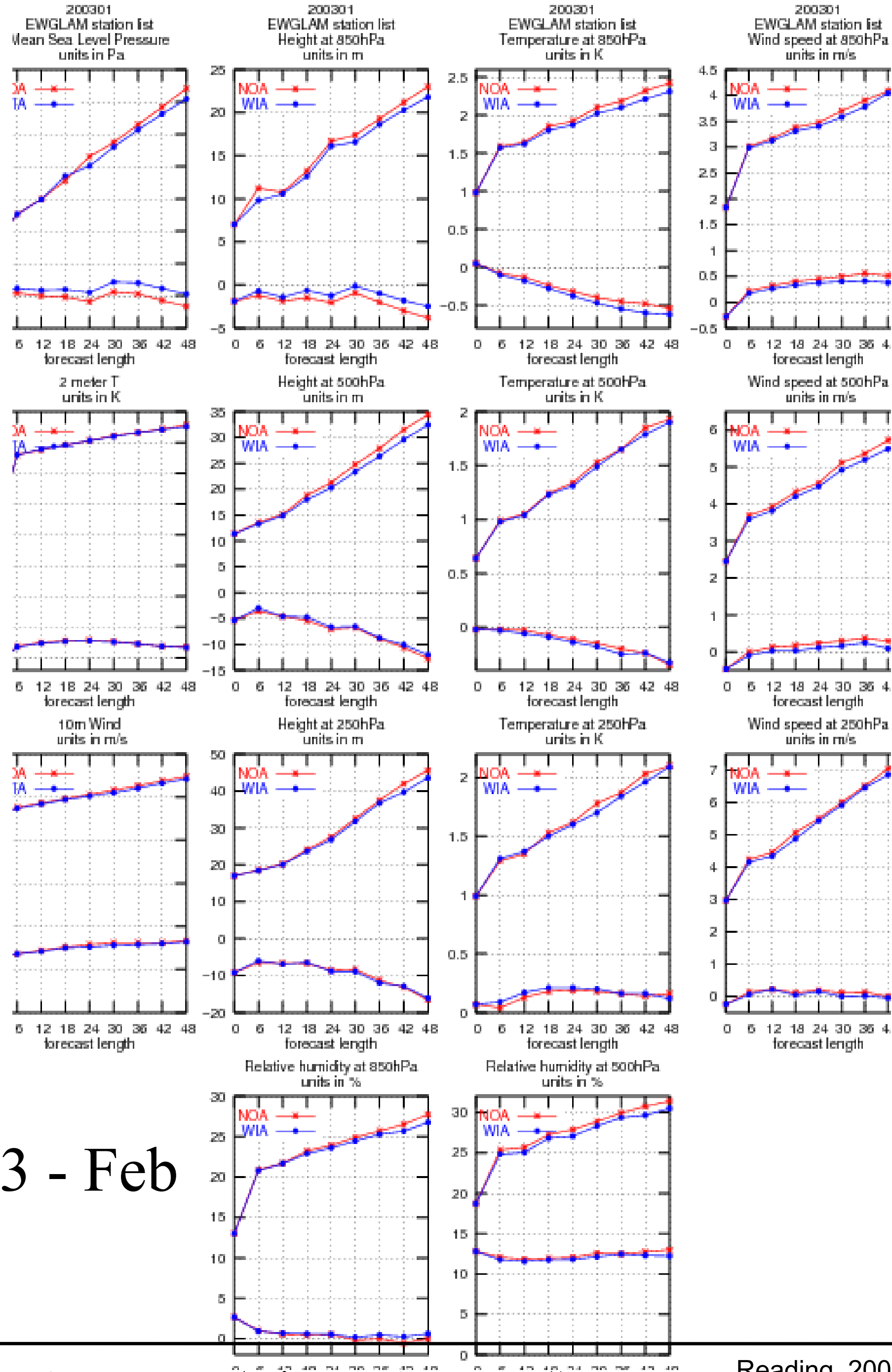


→ ensembles


courtesy: HIRLAM-DMI



Jan - 2003 - Feb



courtesy: HIRLAM-DMI (Bjarne Amstrup)

Outline of the talk

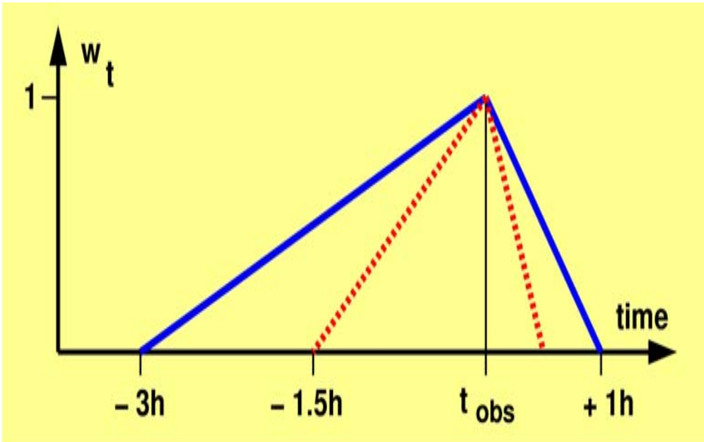
- ✓ Introduction: Forecast qualitiy of limited area models
- ✓ COSMO-Project: assimilation of radiances for COSMO-EU
  (ATOVS/AMSU-A, SEVIRI, IASI)
  - ➔ 1D-Var and nudging
  - ➔ background errors
  - ➔ bias correction
  - ➔ first guess above model top (stratosphere)
  - ➔ preliminary results for AMSU-A
  - ➔ preliminary results for SEVIRI/MSG
  - ➔ conclusions

Use of nonlinear operators with Nudging at appropriate time

conventional observations:

nudge observation 1.5 h before (and 30 min after) observation time with
temporal weighting depending on time difference to observation

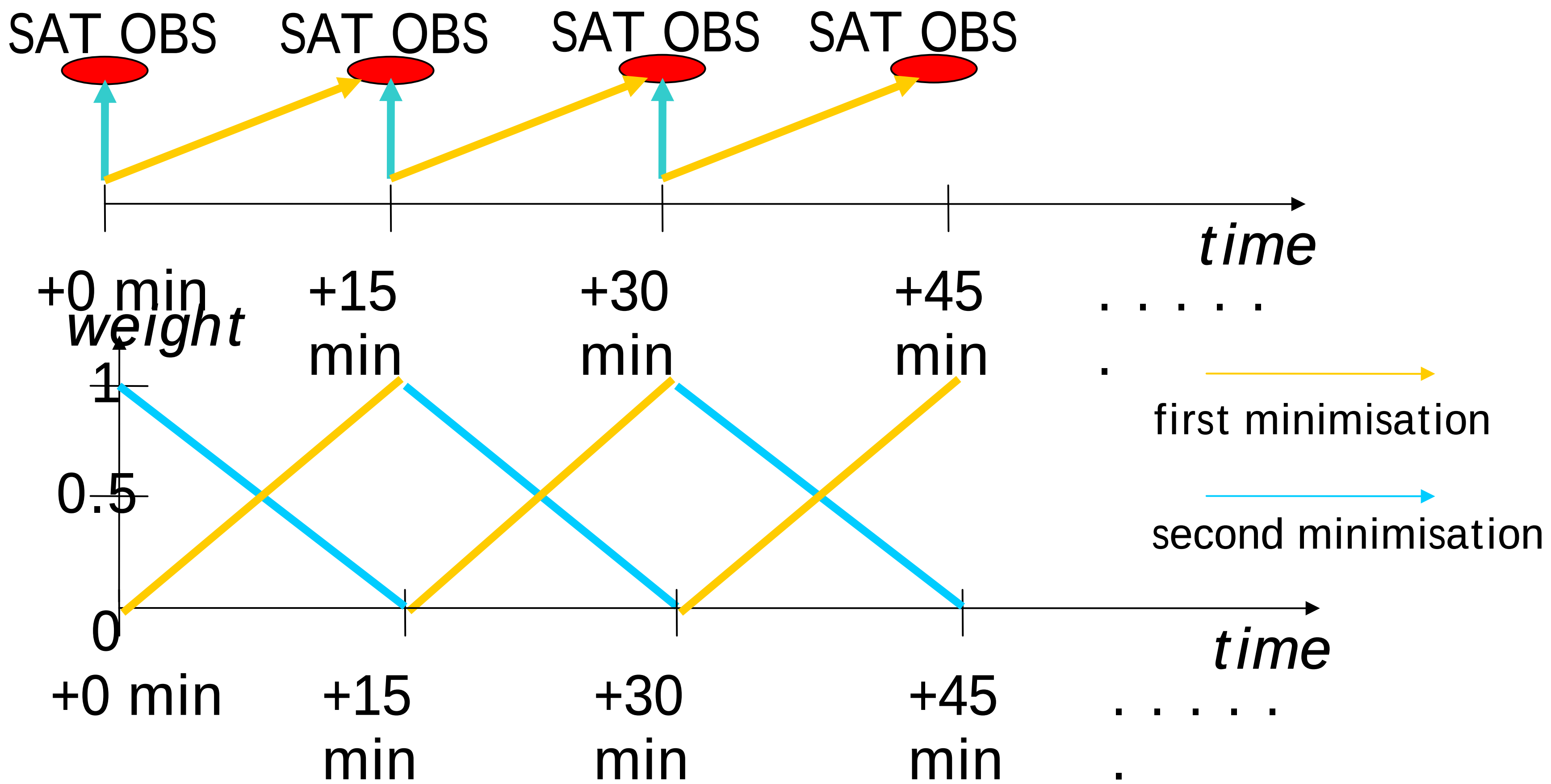
Temporal Weighting



preliminary retrievals have to be computed for nonlinear observations
use first guess available -1.5 h before obervation time
repeat retrieval every 30 min until nudging analysis reach observation time
attention: first guess and observation become correlated!!!

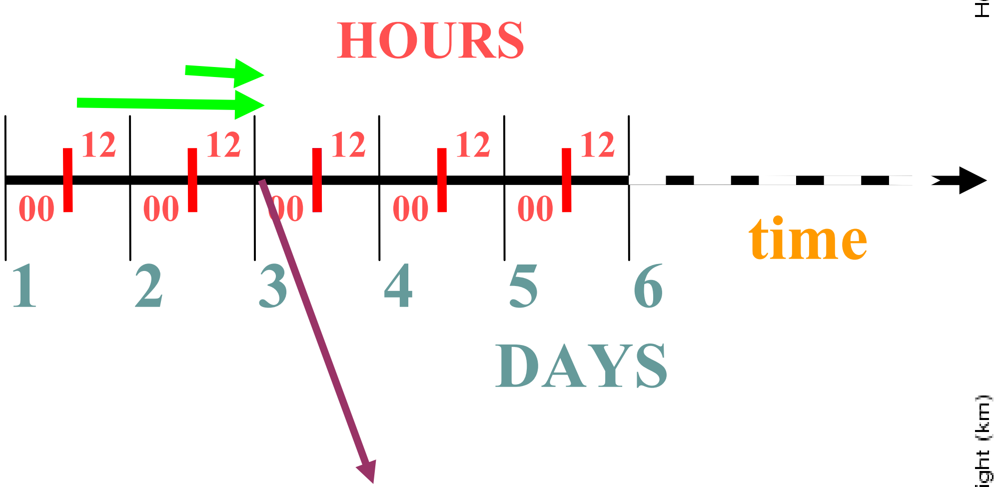
Use of nonlinear operators with Nudging for observations with high temporal resolution

e.g. MSG/SEVIRI with time resolution of 15 min

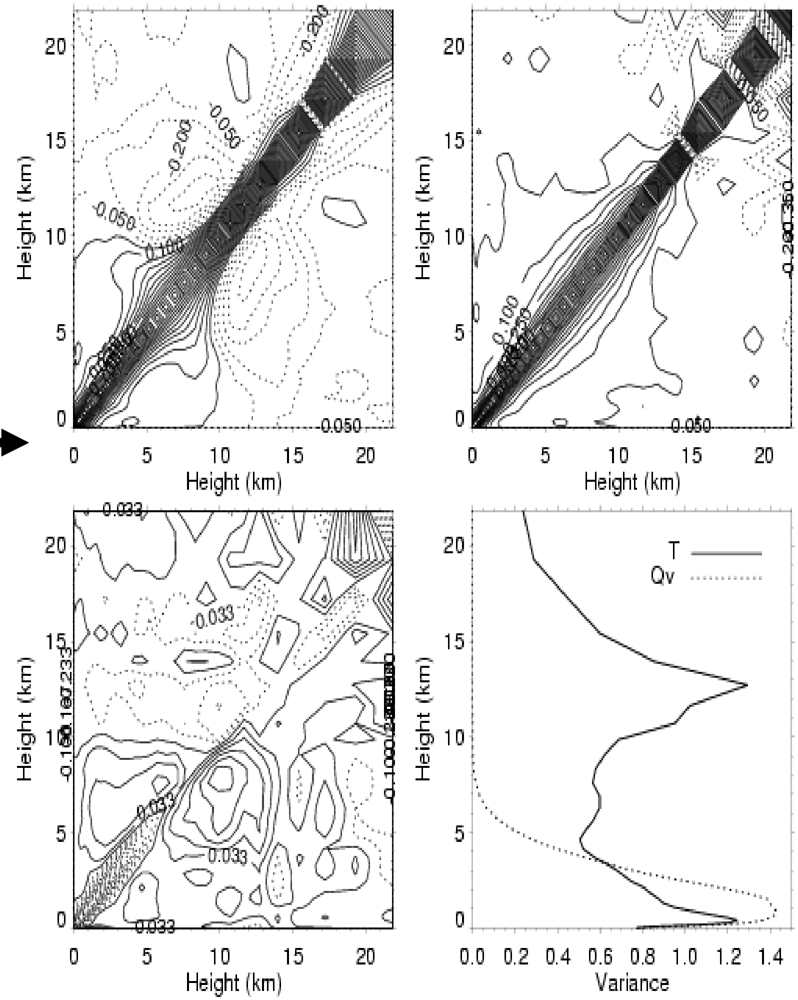


Error covariance matrix B

The B matrix is calculated using forecast comparisons at +12h and +36h averaged over three months worth of data.



Hour of comparison of the two forecasts



Background error covariance matrix B

complication: error structures from boundary values

- standard NMC-method
  (large error structures in statistics do not reflect small scale errors)
- lagged NMC-method (ALADIN): use identical boundary values
  (use boundary values from the same run of the embedding model)
  no boundary errors lead to error statistics of smaller scale
- ensemble B: pertubated observations or physis
  error statistics somehow in between standard and lagged NMC-method

However:

- less constraints on B (geostrophy, hydrostacy)
- high variation in error structures, more motivation for
  situation dependent, flow dependent or adaptive error structures

Bias Correction for ATOVS

scanline and airs mass dependent correction (Eyre, Harris & Kelly)

scanline correction: what is reasonable sampling size?

→ variance of a mean variable $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is $\sigma_{\bar{x}}^2 = \frac{1}{n^2}\sum_{i=1}^{n}\sigma_x^2 = \frac{\sigma_x^2}{n}$

for given error variance the required sample size is $n \geq \frac{\sigma_x^2}{\sigma_{\bar{x}}^2}$

→ application to obs – fg brightness temperatures: $\Delta y_i = (y_i^o - y_i^{fg})$

(statistics are for each individual fov)

examples:

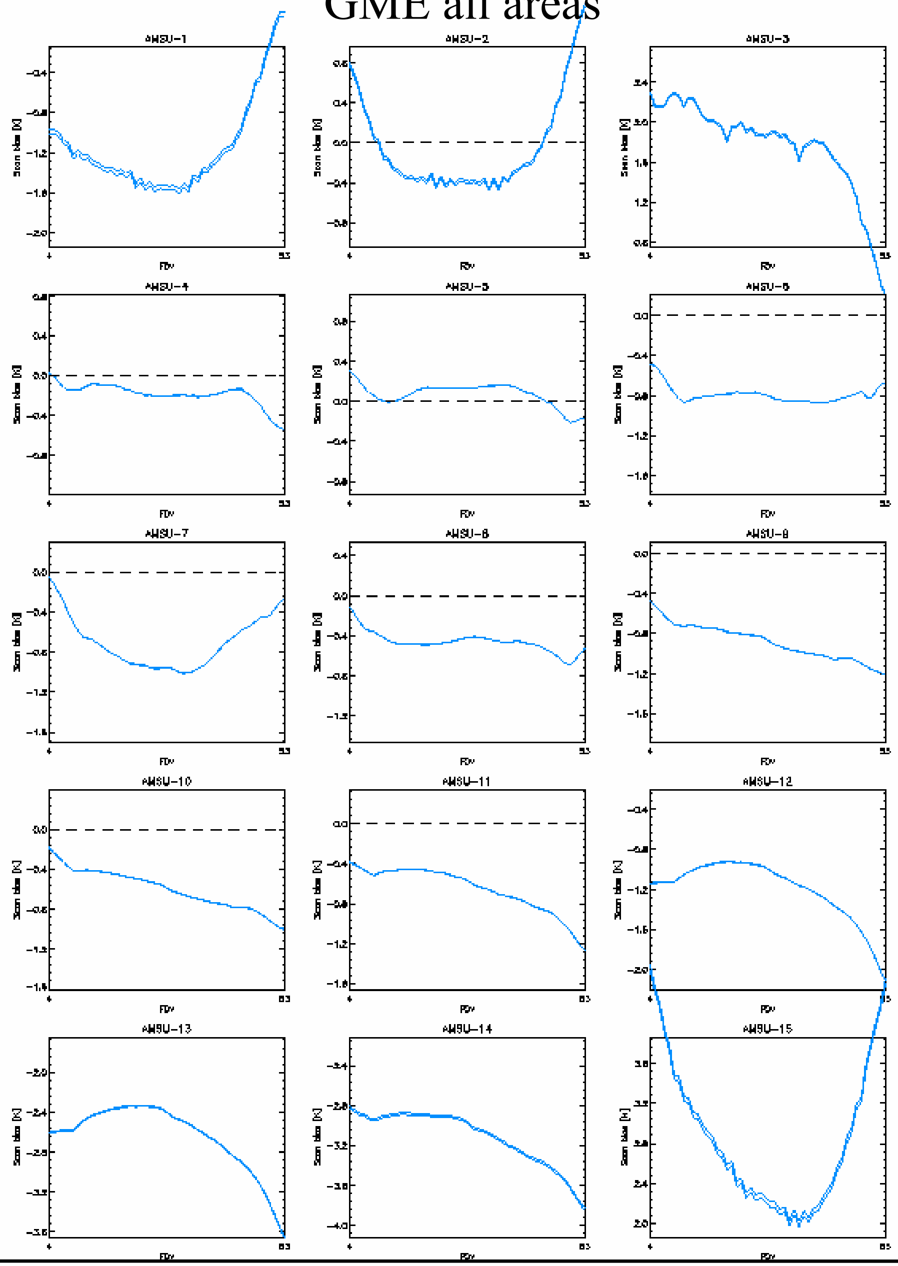
channel	AMSU-A 6	AMSU-A 4	AMSU-A 3
σ_x =	0.2 K	0.4 K	2.1 K
n =	400	1600	21000

for $\sigma_{\bar{x}} \leq 0.01\,\mathrm{K}$

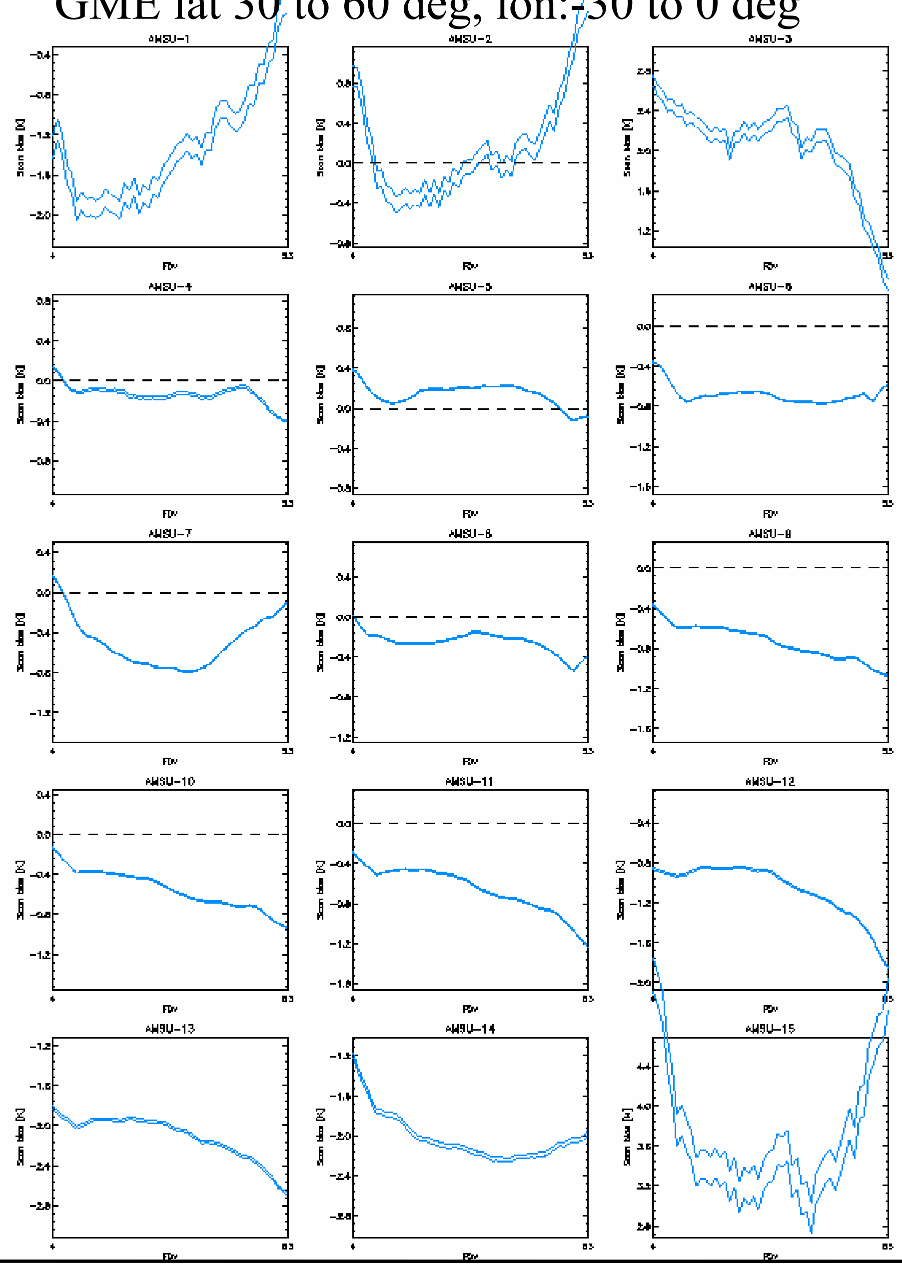
- time to obtain required sample sizes for individual fovs depend on model area (size over sea)
- for COSMO-EU two weeks for most relevant temperature sounding channels
- what about **representativity** (synoptic scenarious, seasonal changes) ?

scanline biases AMSU/NOAA 18 (15 to 25 June 2007)

GME all areas



GME lat 30 to 60 deg, lon:-30 to 0 deg

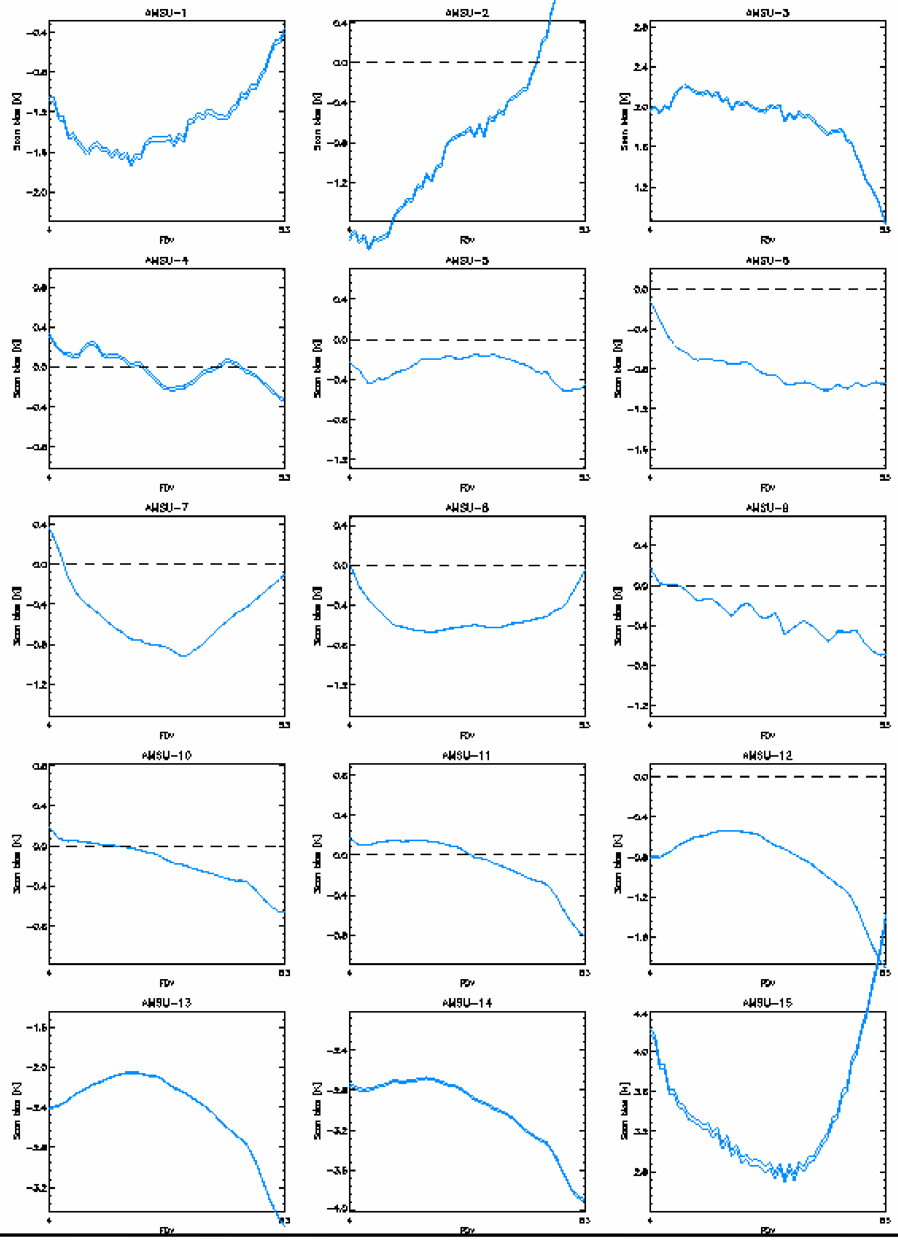


approx 50000 obs/fov

approx 1500 obs/fov

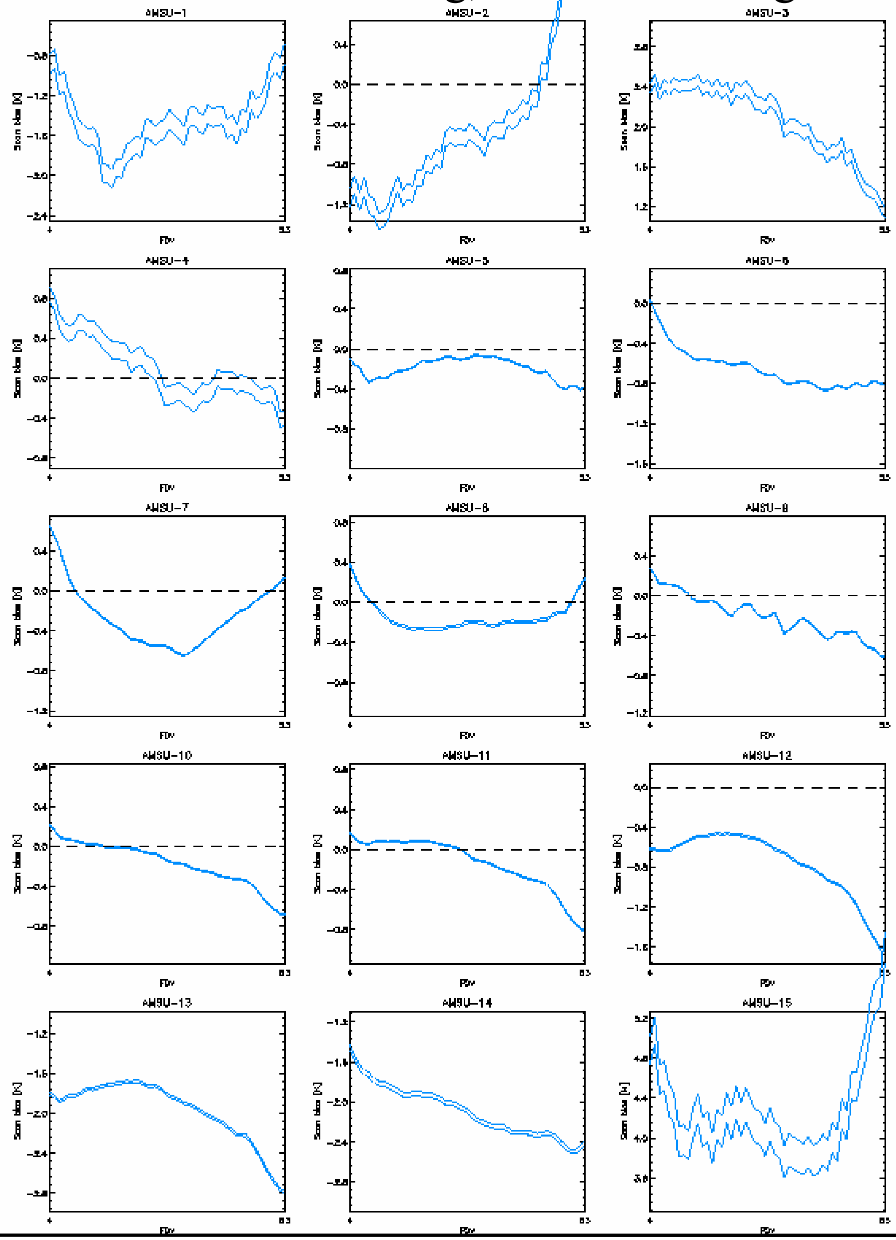
scanline biases AMSU/NOAA 16 (15 to 25 June 2007)

GME all areas



approx 50000 obs/fov

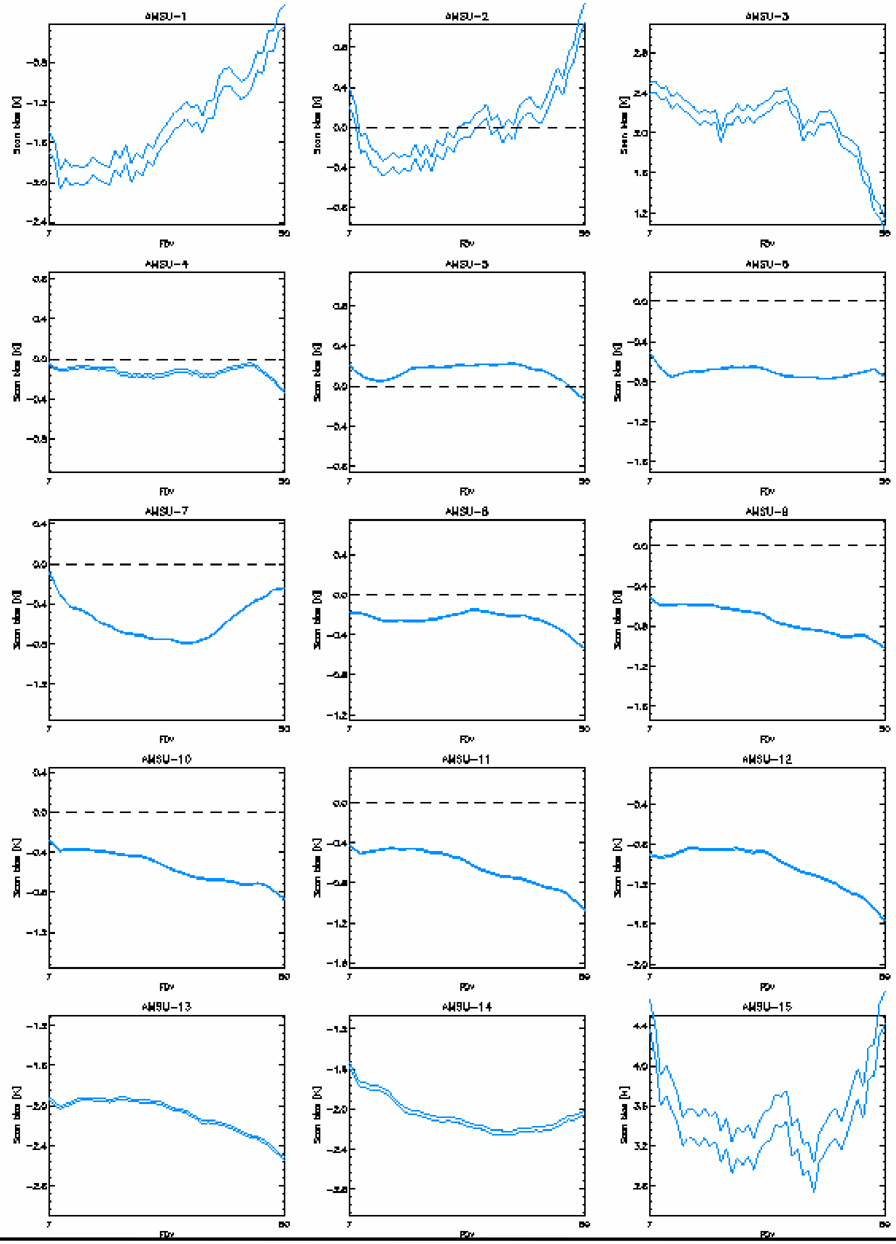
GME lat 30 to 60 deg, lon: -30 to 0 deg



approx 1500 obs/fov

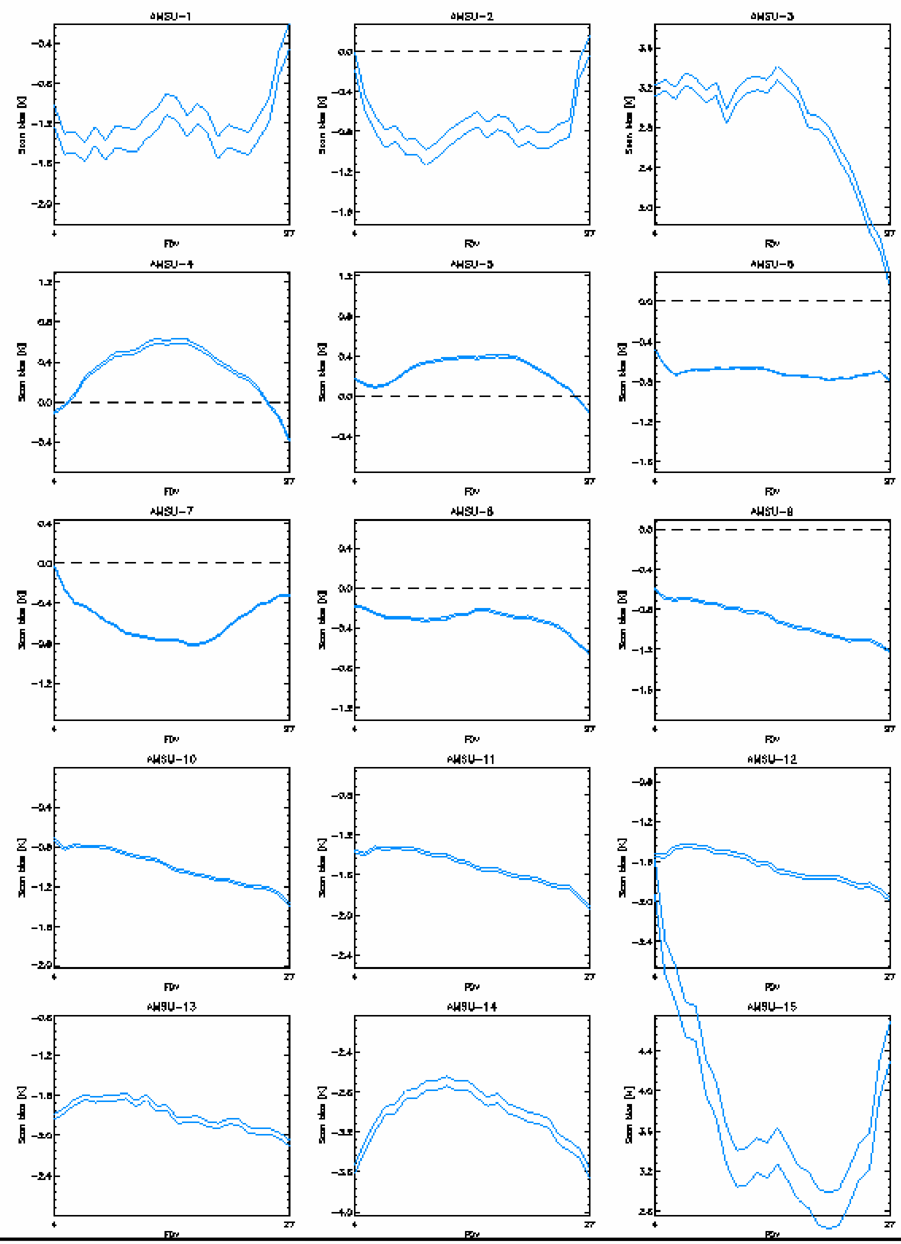
scanline biases AMSU/NOAA 18 (15 to 25 June 2007)

GME lat 30 to 60 deg, lon:-30 to 0 deg



approx 1200 obs/fov

COSMO-EU: approx 1200-1500 fovs

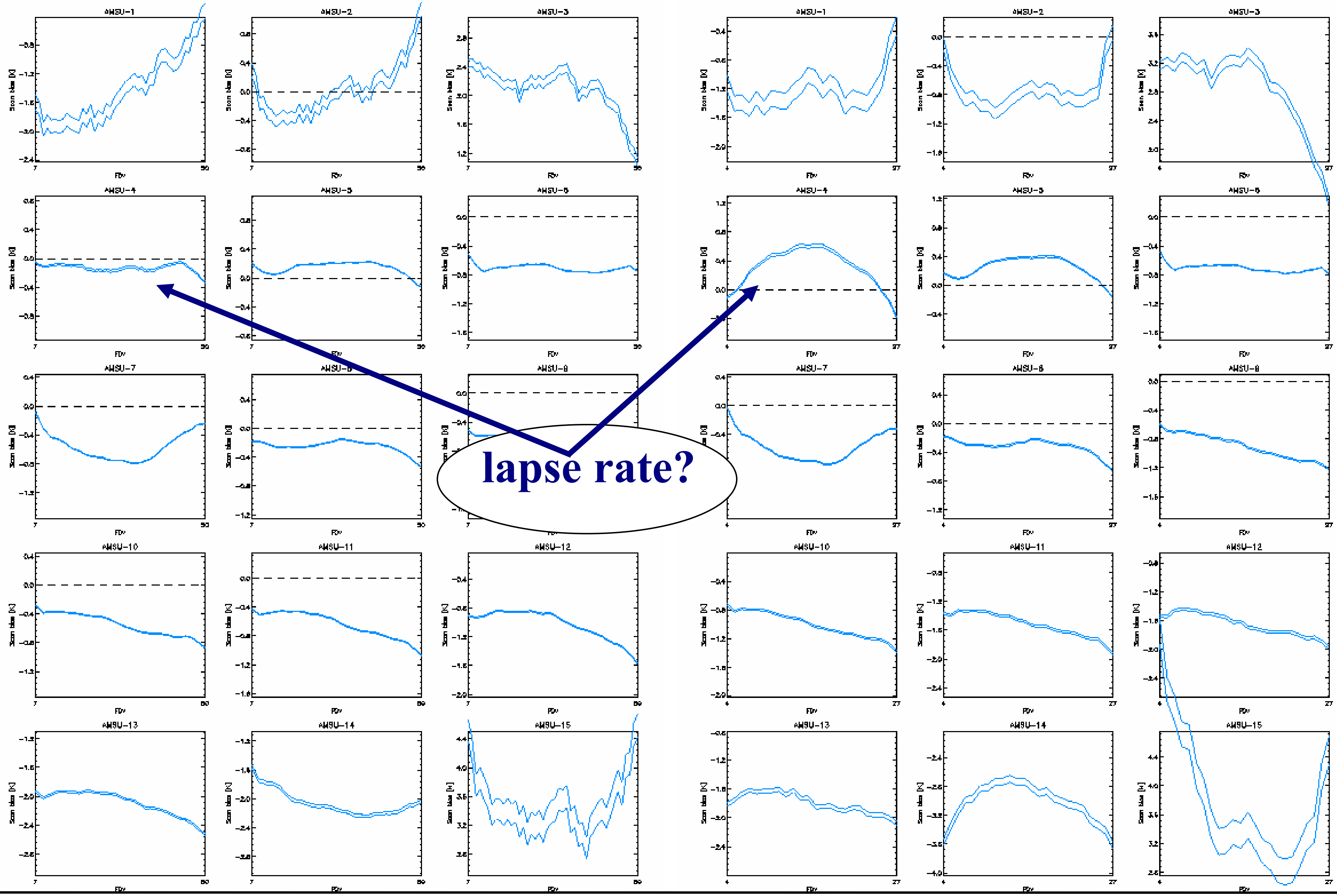


approx 1000-1500 obs/fov

scanline biases AMSU/NOAA 18 (15 to 25 June 2007)

GME lat 30 to 60 deg, lon:-30 to 0 deg

COSMO-EU: approx 1200-1500 fovs



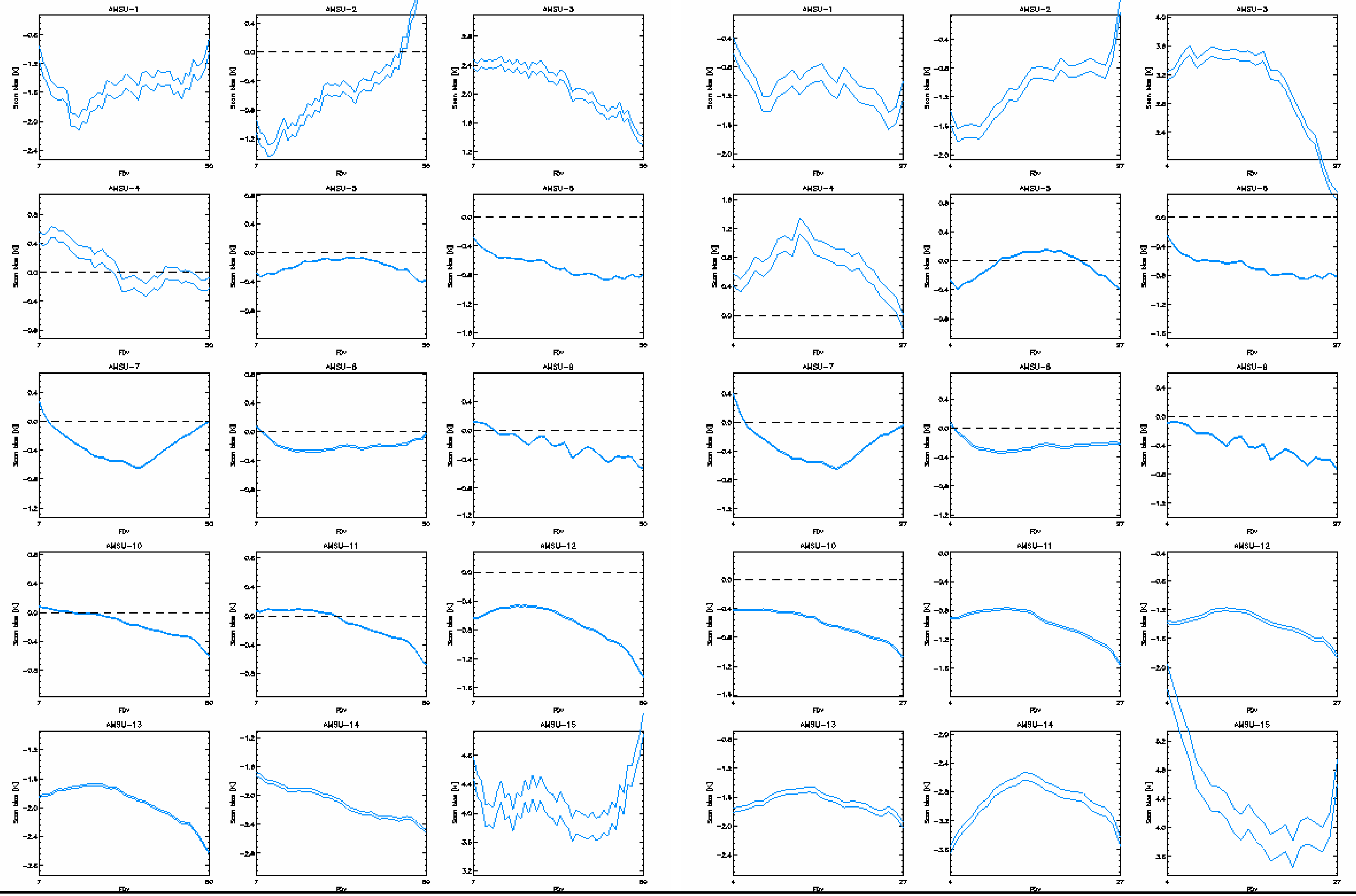
approx 1200 obs/fov

approx 1000-1500 obs/fov

scanline biases AMSU/NOAA 16 (15 to 25 June 2007)

GME lat 30 to 60 deg, lon:-30 to 0 deg

COSMO-EU: approx 1200-1500 fovs



approx 1200 obs/fov

approx 1200-1500 obs/fov

Bias Correction for ATOVS

scanline correction

- irregular shape
- constant with area (for most channels, no lattitude dependency)
- sample size not too critical
- representativity seems no issue
- GME and COSMO-EU show significant differences only for surface (and humidity channels)
- no significant influence of interplation with ATOVPP/AAPP (6 resp. 4 side fovs removed)

air mass correction

- idea: air mass bias is situation dependent, model these biases using meteorological predictors
- choices of predictors:
  - AMSU-A 5 and 9 (observed or simulated)
  - mean temperatures (50-200 hPa and 300-1000 hPa), SST, IWV
  - predictors for scanline correction: zenith angle, sqare of zenith angle (or remove scanline correction before air mass correction)
  - latitude as predictor/coefficients variable with latitude band

What is a good choice of predictors?

air mass correction, choice of predictors

example: using two predicors p1 and p2 and a constant
fit parameters c1, c2 and c3 to obs-fg differences:

minimise: $F(c_1, c_2, c_3) = \frac{1}{2} \sum_{i=1,\dots,n} \left(y_i^o - y_i^{fg} - c_1 p_{1,i} - c_2 p_{2,i} - c_3 \right)^2$

solution of $\frac{\partial F}{\partial c_1} = \frac{\partial F}{\partial c_2} = \frac{\partial F}{\partial c_3} = 0$ provides $c_1, c_2, c_3 \, (\Delta y_{i,i=1,\dots,n})$ with $\Delta y_i = (y_i^o - y_i^{fg})$

Gaussian error analysis:

What are the errors of c1, c2 and c3 for given $\Delta y_i = (y_i^o - y_i^{fg})$ with errors σ_o ?

derivatives of $c_1, c_2, c_3 \, (\Delta y_{i,i=1,\dots,n})$ are
(chain rule for implicit functions)

$$\begin{pmatrix} \frac{\partial c_1}{\partial \Delta y_i} \\ \frac{\partial c_2}{\partial \Delta y_i} \\ \frac{\partial c_3}{\partial \Delta y_i} \end{pmatrix} = -(\mathrm{Hess}\, F)^{-1} \begin{pmatrix} \frac{\partial^2 F}{\partial c_1 \partial \Delta y_i} \\ \frac{\partial^2 F}{\partial c_2 \partial \Delta y_i} \\ \frac{\partial^2 F}{\partial c_3 \partial \Delta y_i} \end{pmatrix}$$

approximated errors:

$$\begin{pmatrix} \Delta c_1 \\ \Delta c_2 \\ \Delta c_3 \end{pmatrix} = \begin{pmatrix} \sum_{i=1,\dots,n} \frac{\partial c_1}{\partial \Delta y_i} \Delta y_i \\ \sum_{i=1,\dots,n} \frac{\partial c_2}{\partial \Delta y_i} \Delta y_i \\ \sum_{i=1,\dots,n} \frac{\partial c_3}{\partial \Delta y_i} \Delta y_i \end{pmatrix}$$

covariance:

$$\mathrm{Kov} \begin{pmatrix} \Delta c_1 \\ \Delta c_2 \\ \Delta c_3 \end{pmatrix} = -(\mathrm{Hess}\, F)^{-1} \sigma_o^2 \qquad \mathrm{Hess}\, F = \begin{pmatrix} \sum p_{1,i}^2 & \sum p_{1,i} p_{2,i} & \sum p_{1,i} \\ \sum p_{1,i} p_{2,i} & \sum p_{2,i}^2 & \sum p_{2,i} \\ \sum p_{1,i} & \sum p_{2,i} & n \end{pmatrix}$$

air mass correction, choice of predictors

Hessian and error correlation matrix only depend on predictors
- same for all channels

using AMSU-A 4, 9 and a constant as predictors:

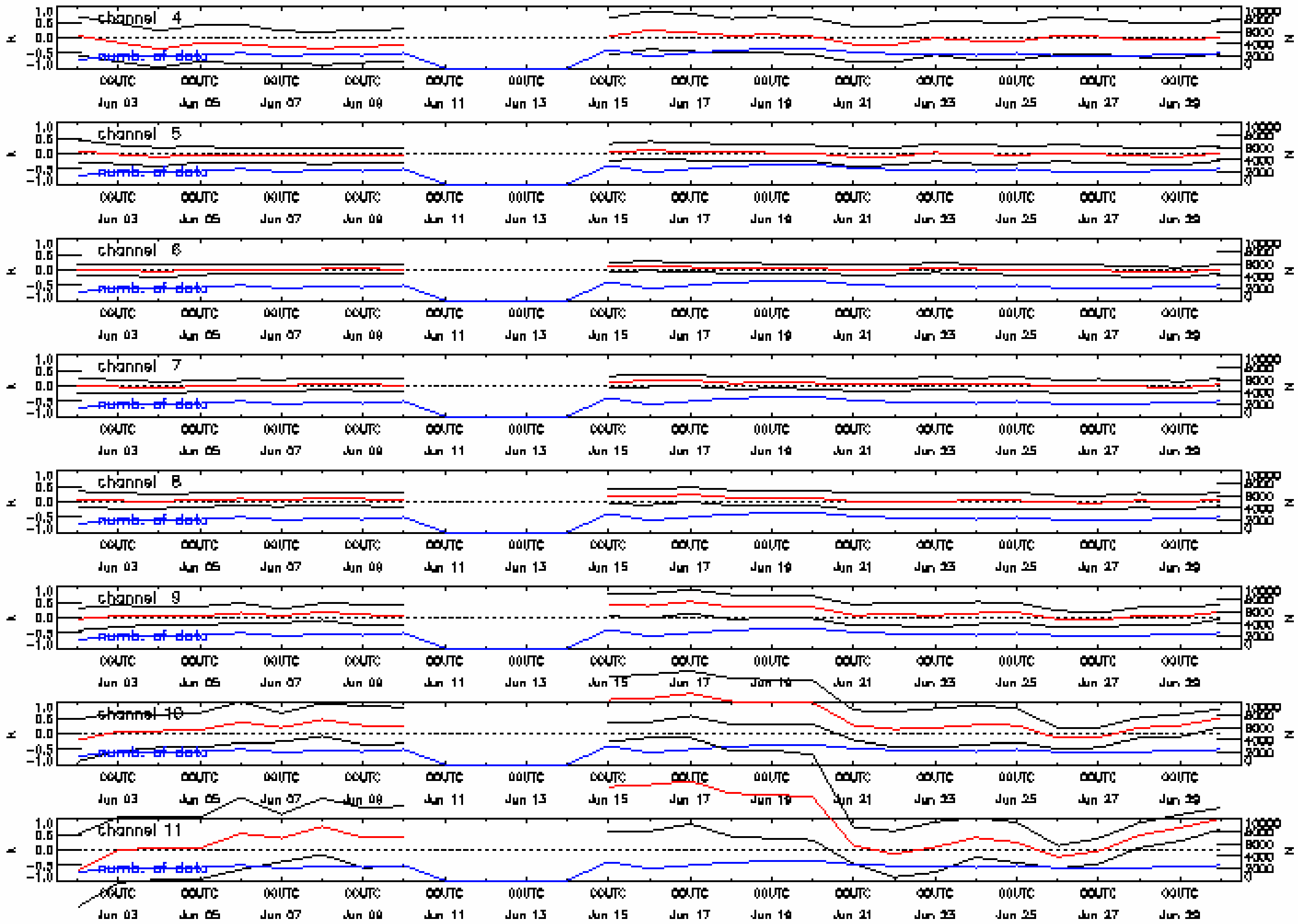
- COSMO-EU (15-25.6.07): $\text{Cor}=\begin{pmatrix} 1{,}00 & 0{,}94 & -0{,}99 \\ 0{,}94 & 1{,}00 & -0{,}98 \\ -0{,}99 & -0{,}98 & 1{,}00 \end{pmatrix}$
- GME (15-25.6.07), lim area: $\text{Cor}=\begin{pmatrix} 1{,}00 & 0{,}88 & -0{,}98 \\ 0{,}88 & 1{,}00 & -0{,}95 \\ -0{,}98 & -0{,}95 & 1{,}00 \end{pmatrix}$
- GME (15-25.6.07), all area: $\text{Cor}=\begin{pmatrix} 1{,}00 & 0{,}60 & -0{,}91 \\ 0{,}60 & 1{,}00 & -0{,}87 \\ -0{,}91 & -0{,}87 & 1{,}00 \end{pmatrix}$

errors of predictors are highly correlated (in this example)
overfitting – useless, loose information,by removing signal
other choice of predictors
longer time series for higher representativity
usually: **several months** of data

biases depend on synoptic situation, season

timeserie of bias corrected observations minus first guess

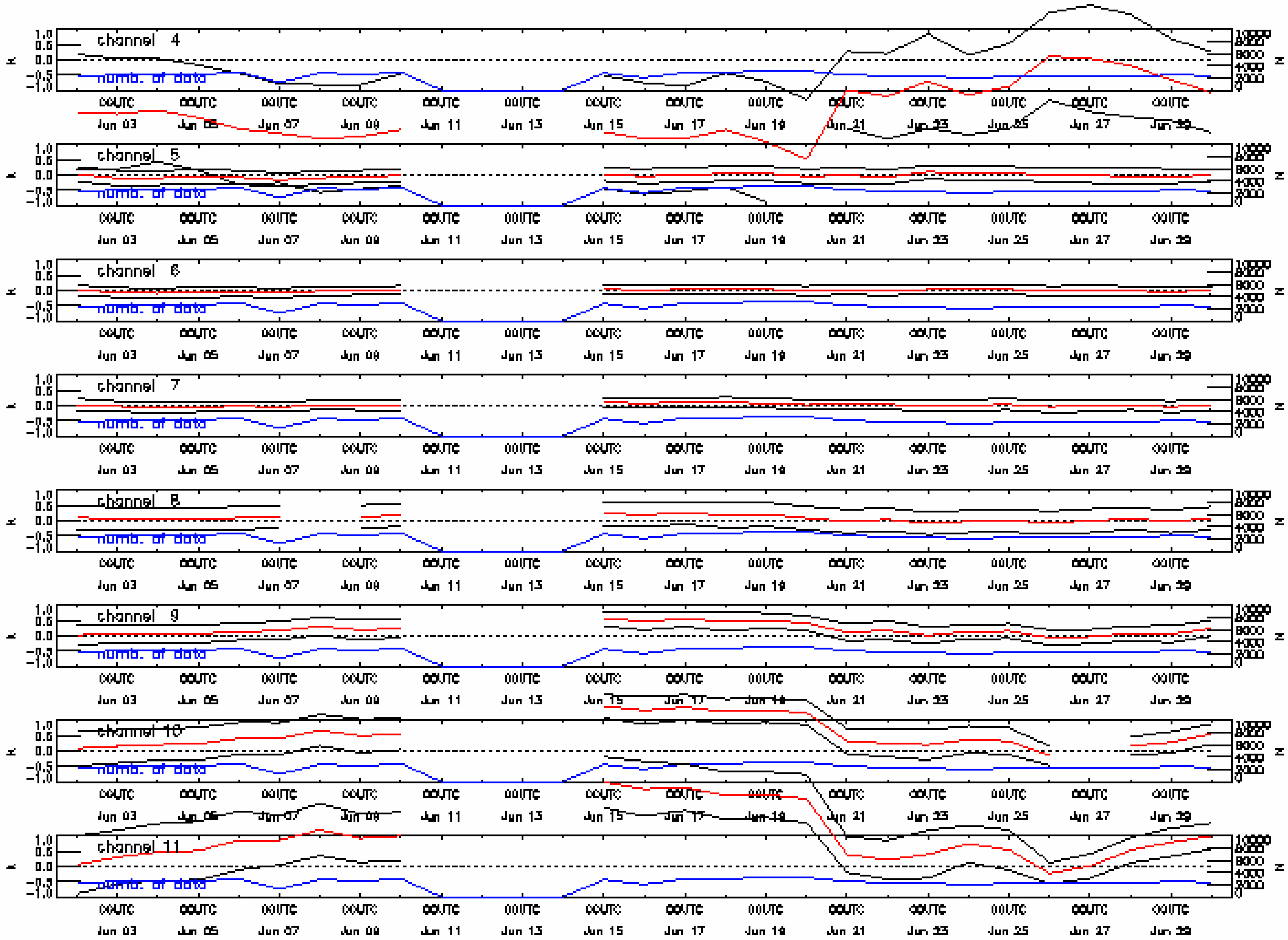
AMSU-A channels 4-11, NOAA-18



stable in the troposphere, however large variations for high sounding channels

timeserie of bias corrected observations minus first guess

AMSU-A channels 4-11, NOAA-16



stable in the stratosphere, however large variations for high sounding channels

provide first guess values above model top

limited area models usually have lower model top than required by RTTOV (0.1hPa - 0.05hPa)

COSMO-EU: 30hPa
HIRLAM: 10hPa
ALADIN: 1hPa

- increase height of model top
- use climatological values (inaccurate, use only lower peaking channels)
- use forecasts of global model IFS (accurate, but timely receive of IFS forecasts required)
- linear regression of high peaking channels to model levels (Met Office)

$y = W x$ x: high peaking channels
y: temperatures on RTTOV levels
W: regression matrix

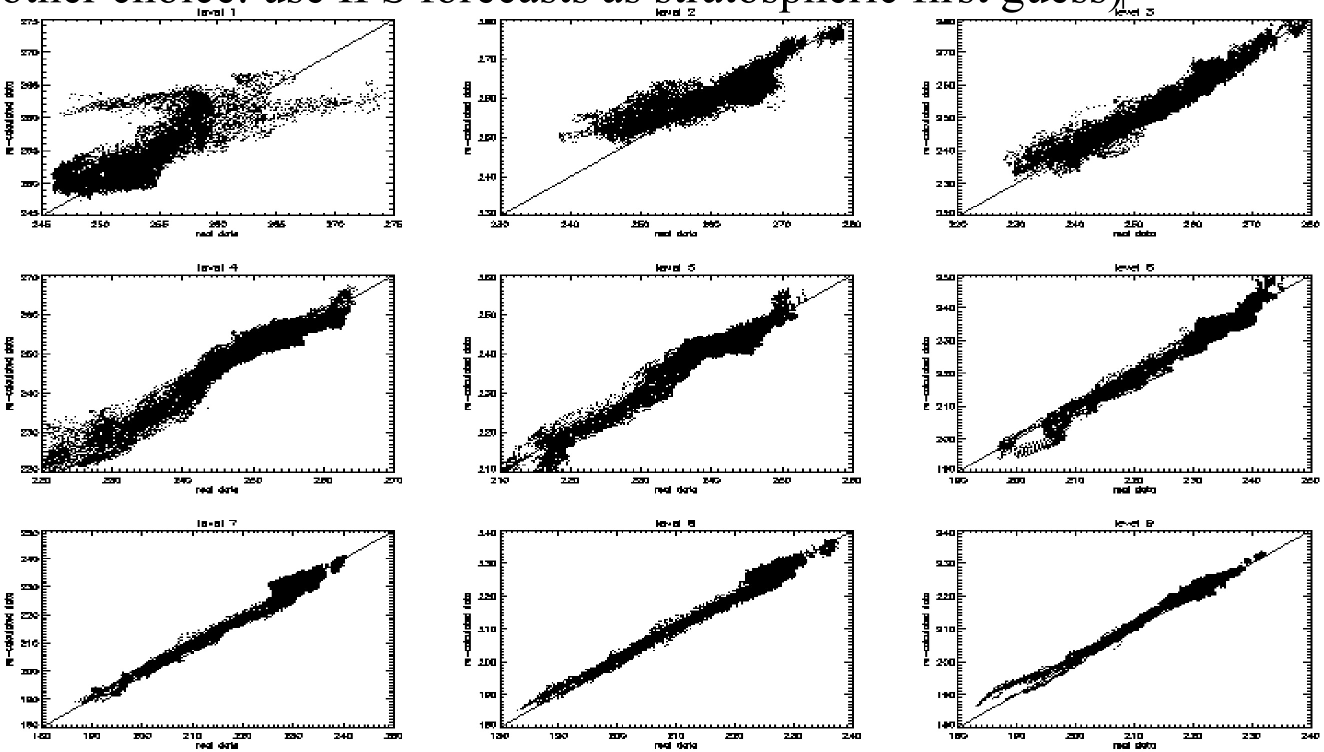
compute W with training data set

reasonable:
- no humidity
- more or less linear relation between, high peaking channels and level temperature (no clouds)

provide first guess values above model top (COSMO-EU: 30hpa)

- → use of climatological values (ERA40) seems not sufficient

- → linear regression of top RTTOV levels from stratospheric channels
  (other choice: use IFS forecasts as stratospheric first guess)



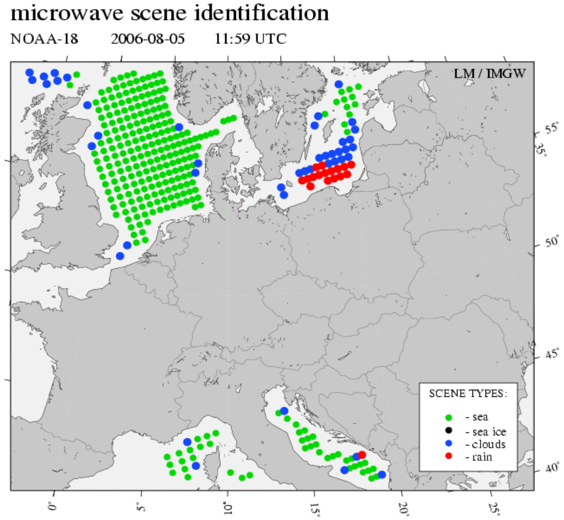
ECMWF profiles versus estimated profiles, top GME levels
accuracy about 5K for lower levels, but ECMWF may have errors in stratosphere too

levels: 0.10, 0.29, 0.69, 1.42, 2.611, 4.407, 6.95, 10.37, 14.81 hPa

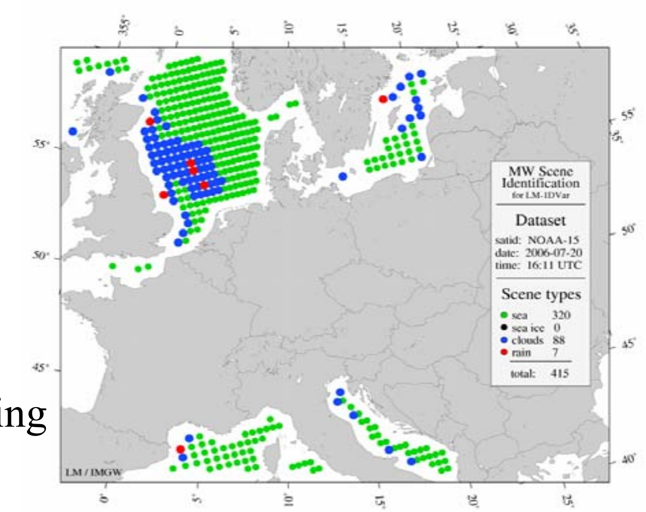
1D-Var for LME – Cloud and Rain detection

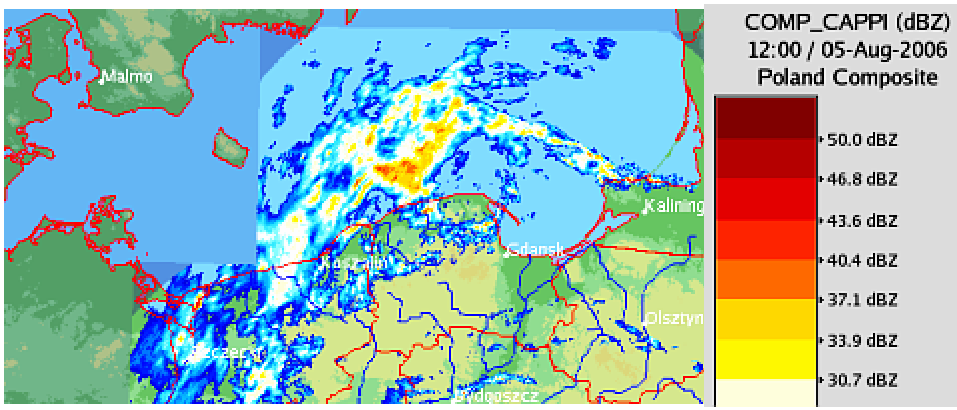
Microwave surface emissivity model: rain and cloud detection (Kelly & Bauer)

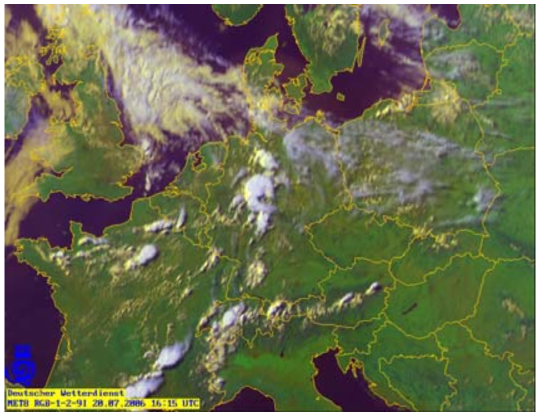
Validation with radar data



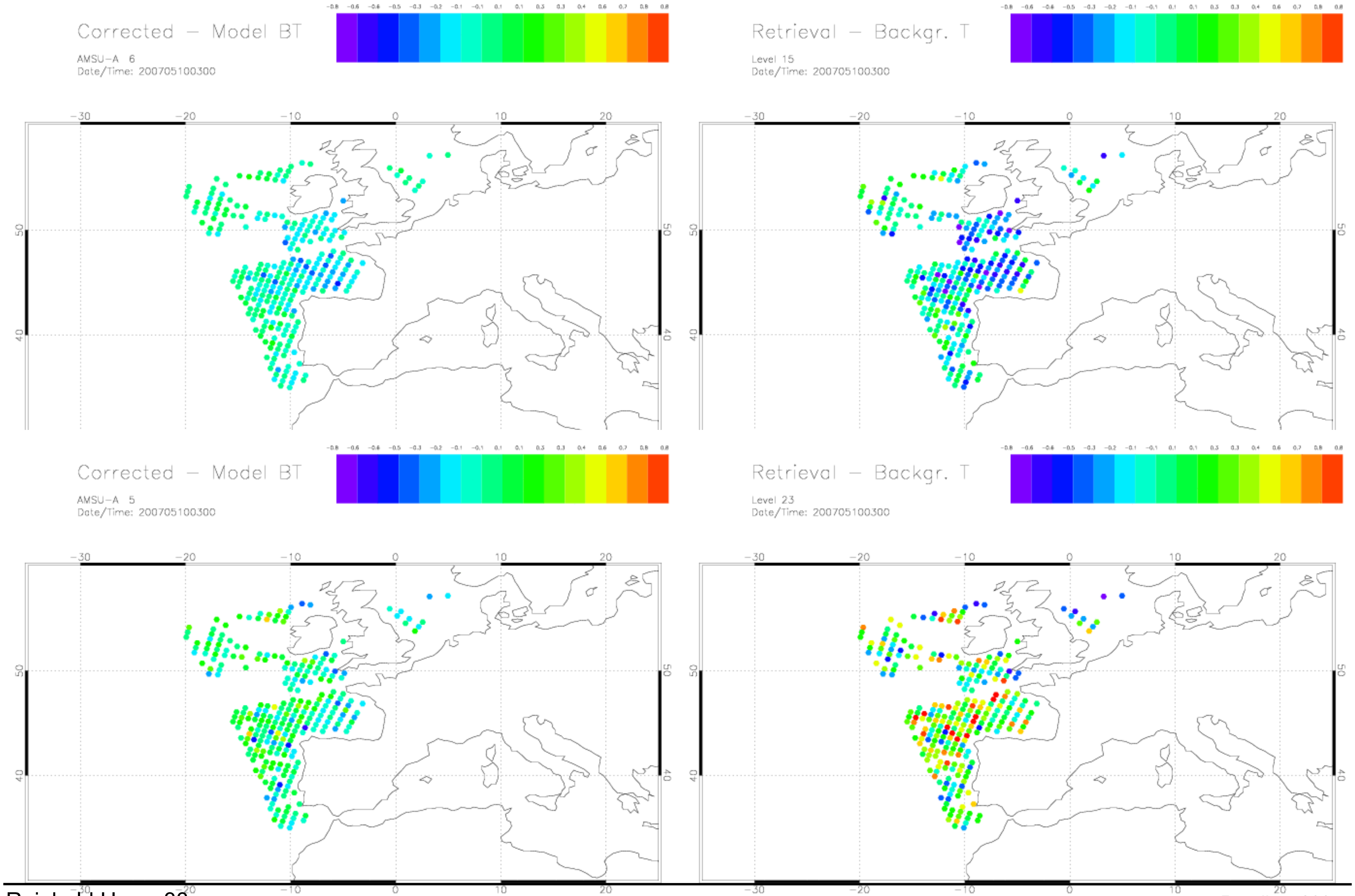
Validation with MSG imaging





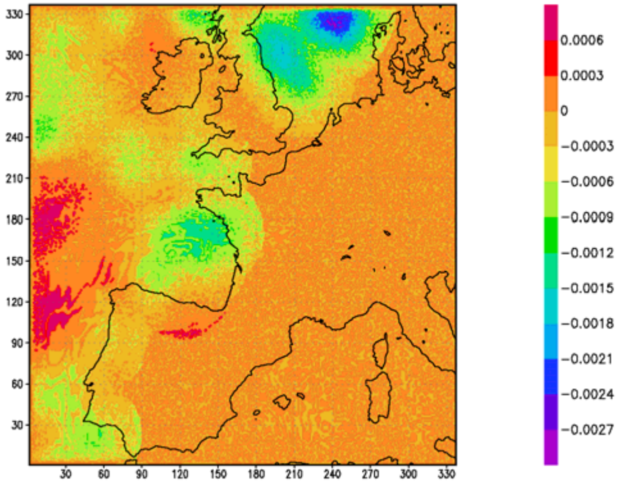


observation increments and resulting 1D-Var increments

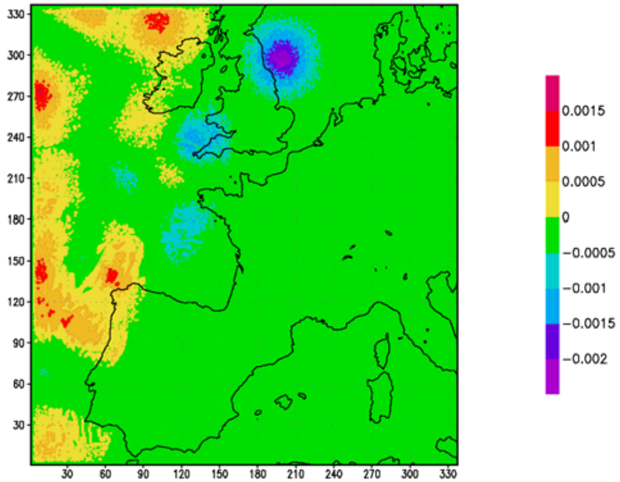


Δ*T*-‘analysis increments’ from ATOVS, after 1 timestep (sat only), $k = 20$

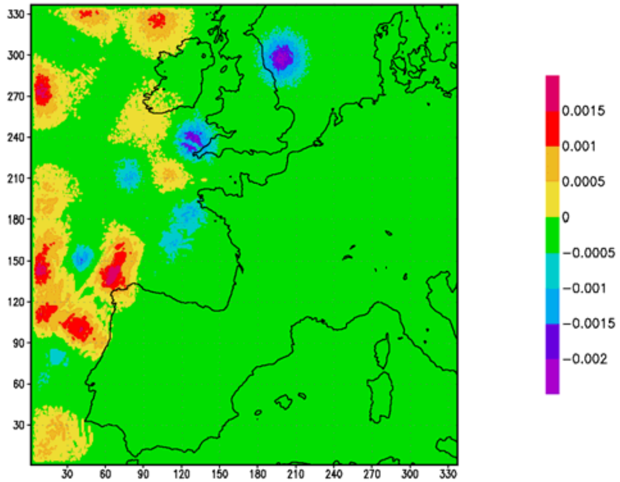
no thinning of 298 ATOVS



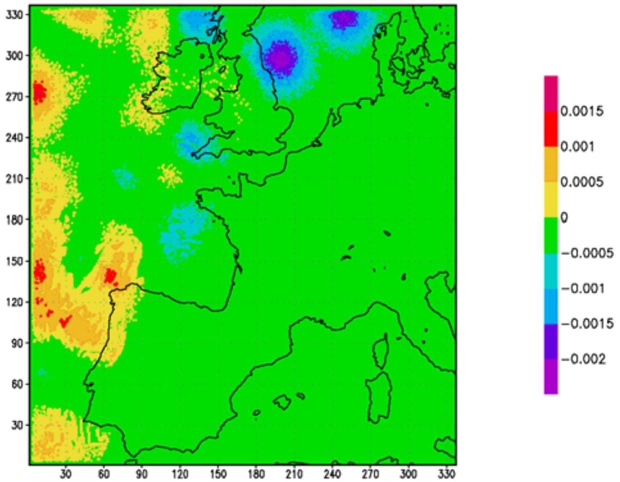
30 ATOVS by old thinning (3)



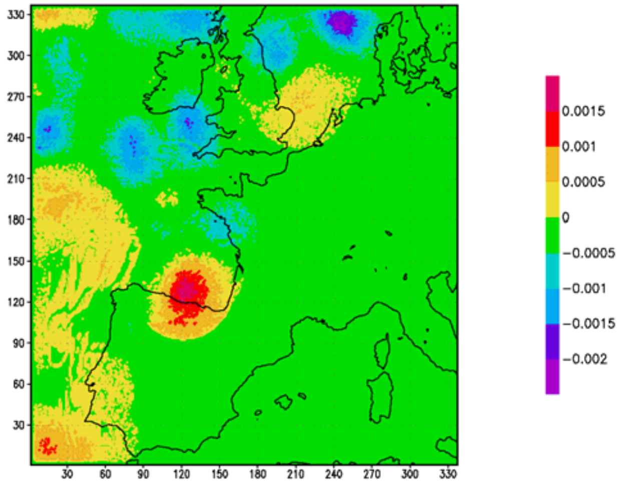
30 ATOVS, correl. scale 70%



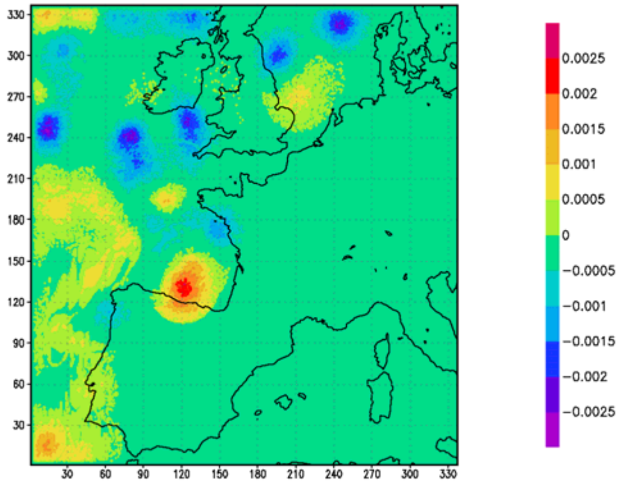
40 ATOVS by thinning (3)



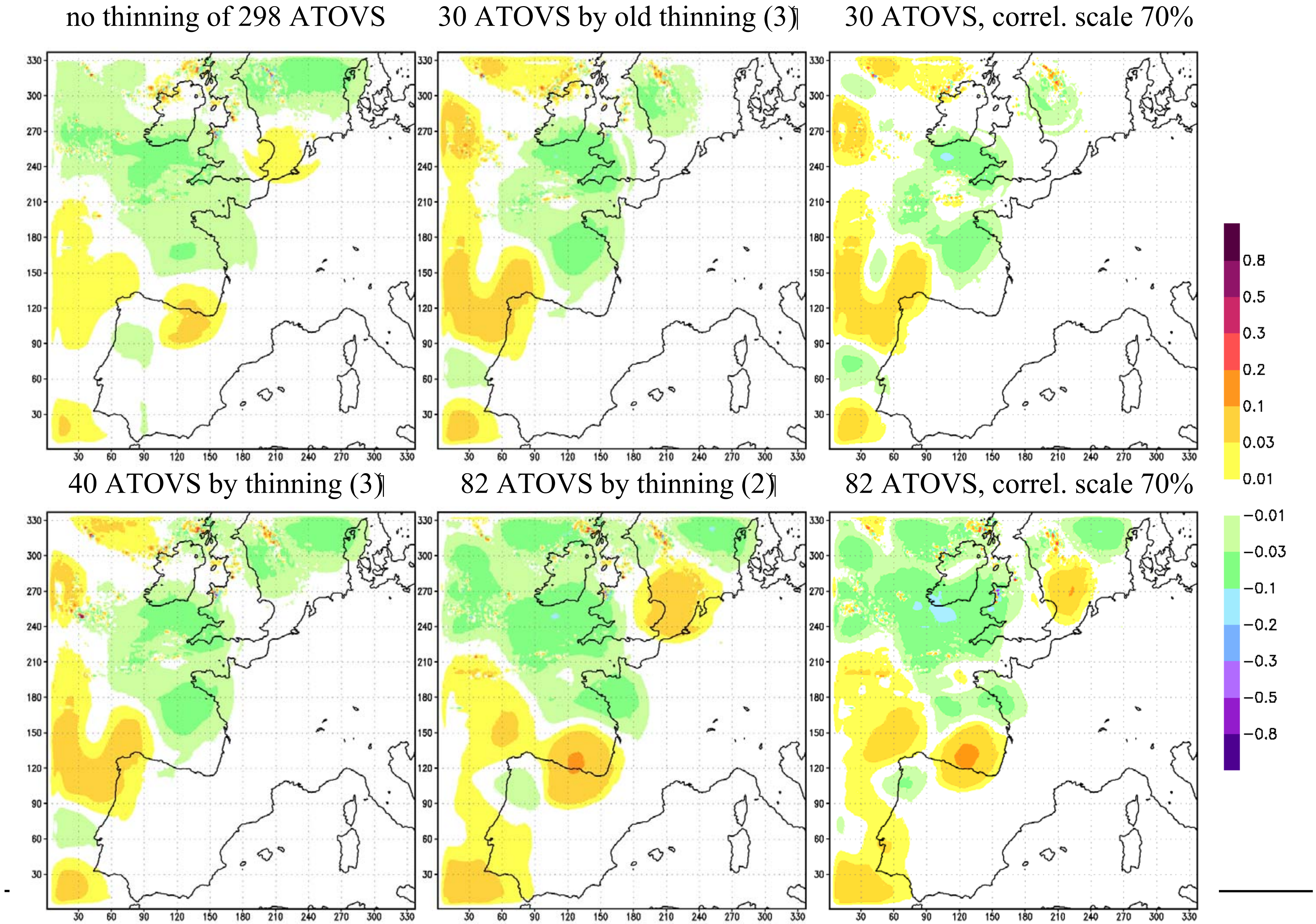
82 ATOVS by thinning (2)



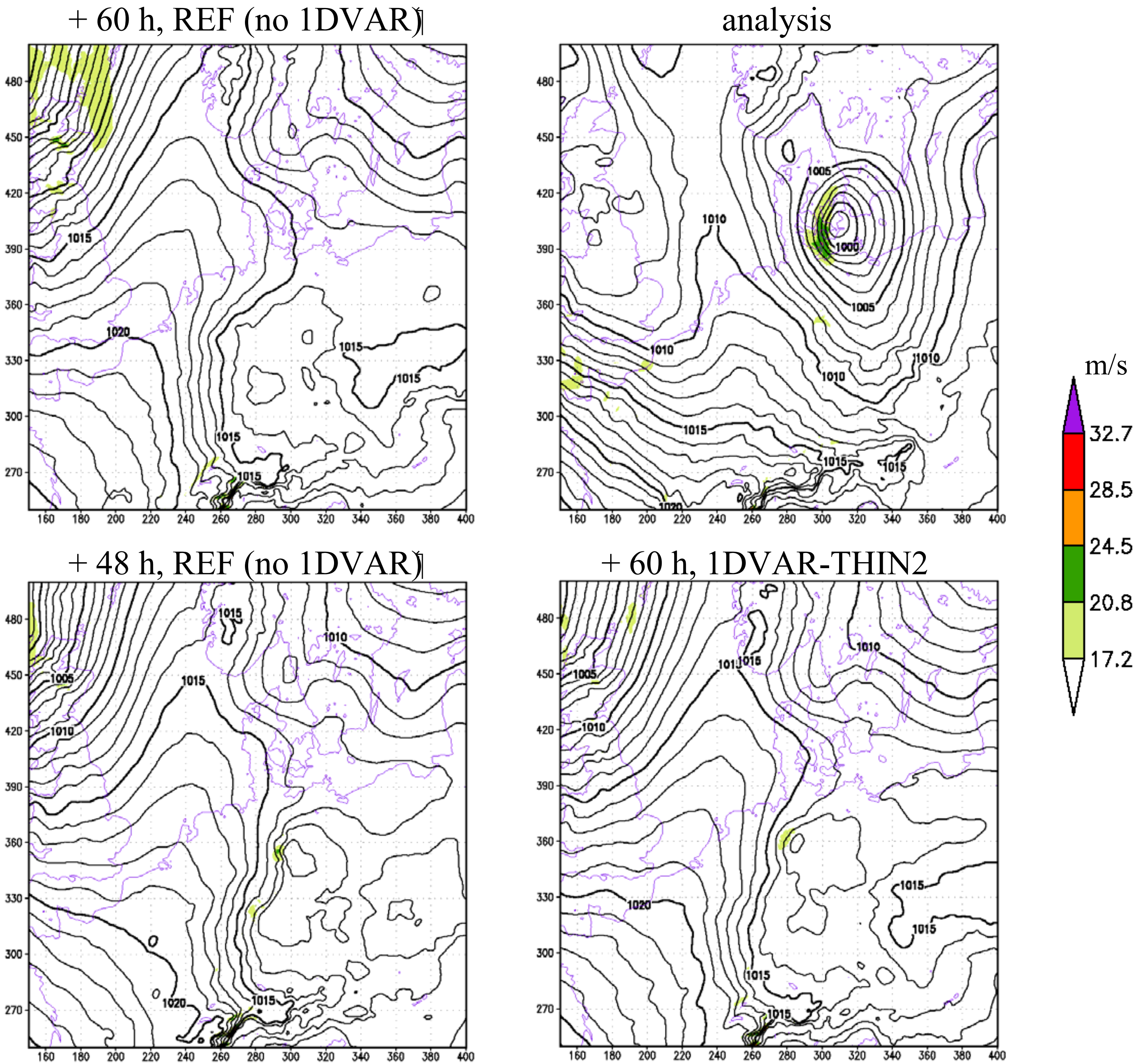
82 ATOVS, correl. scale 70%



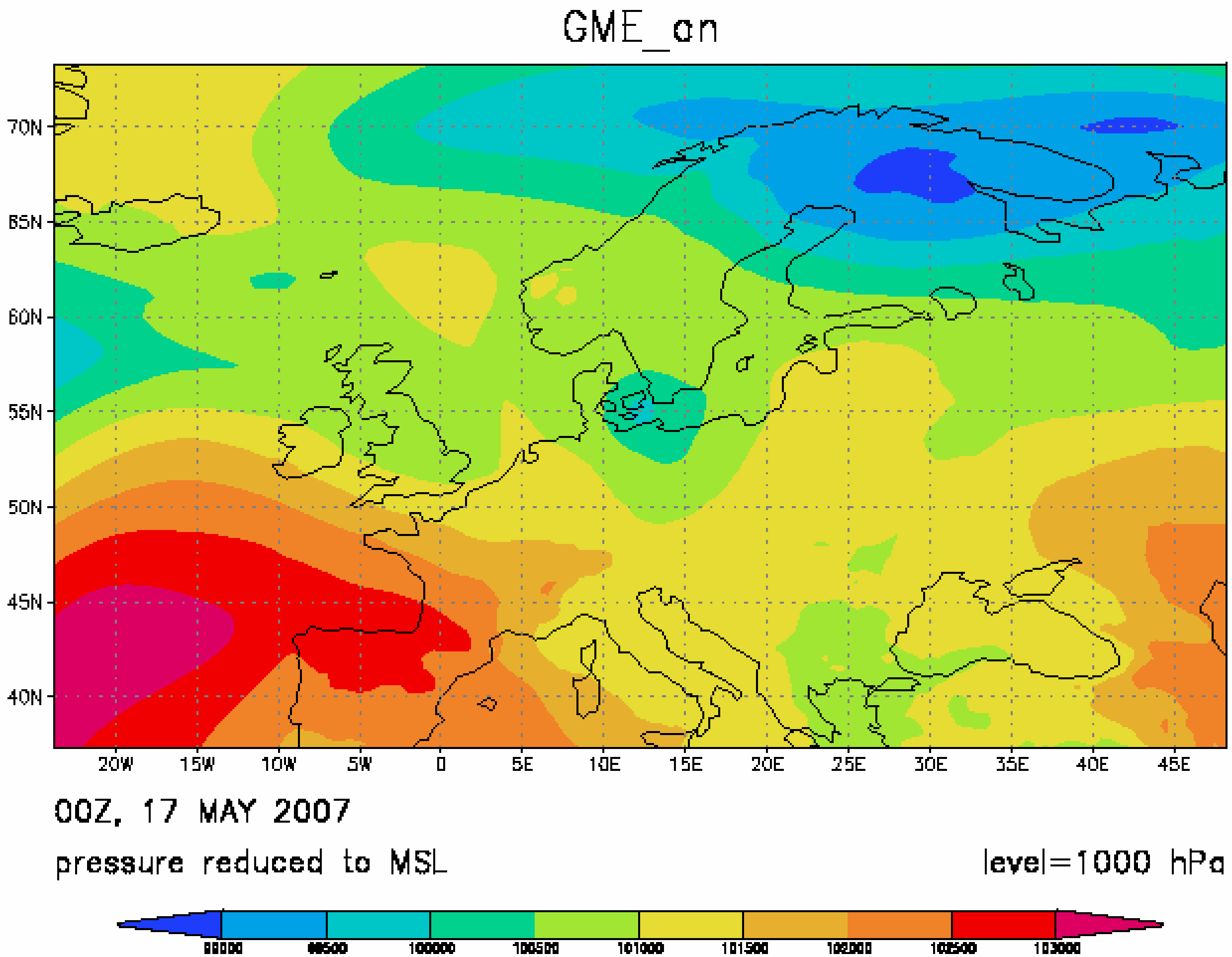
ΔT-‘analysis increments’ from ATOVS, after 30 minutes (sat only), $k = 20$



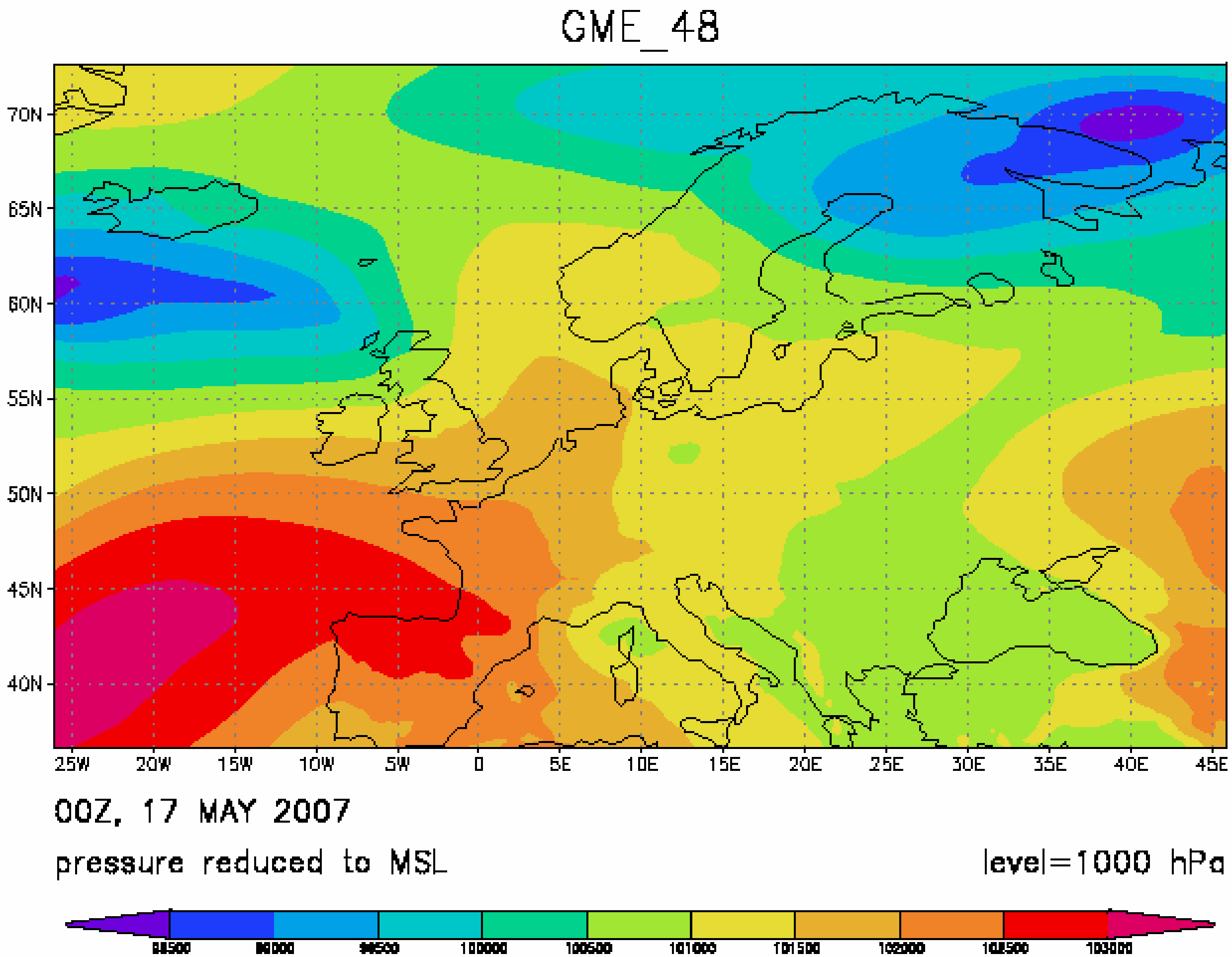
mean sea level pressure & max. 10-m wind gusts valid for 17 May 2007 , 0 UTC



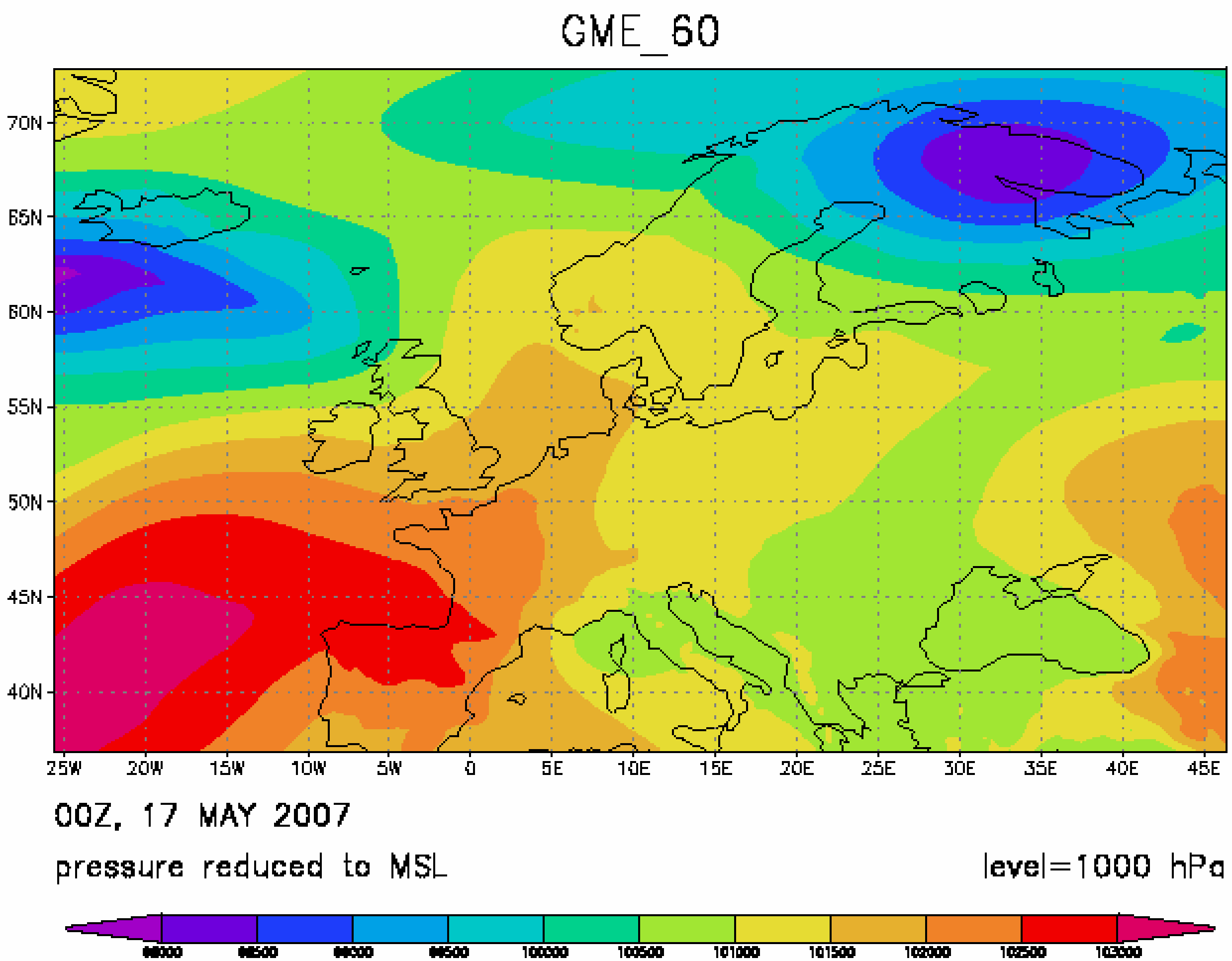
GME analysis



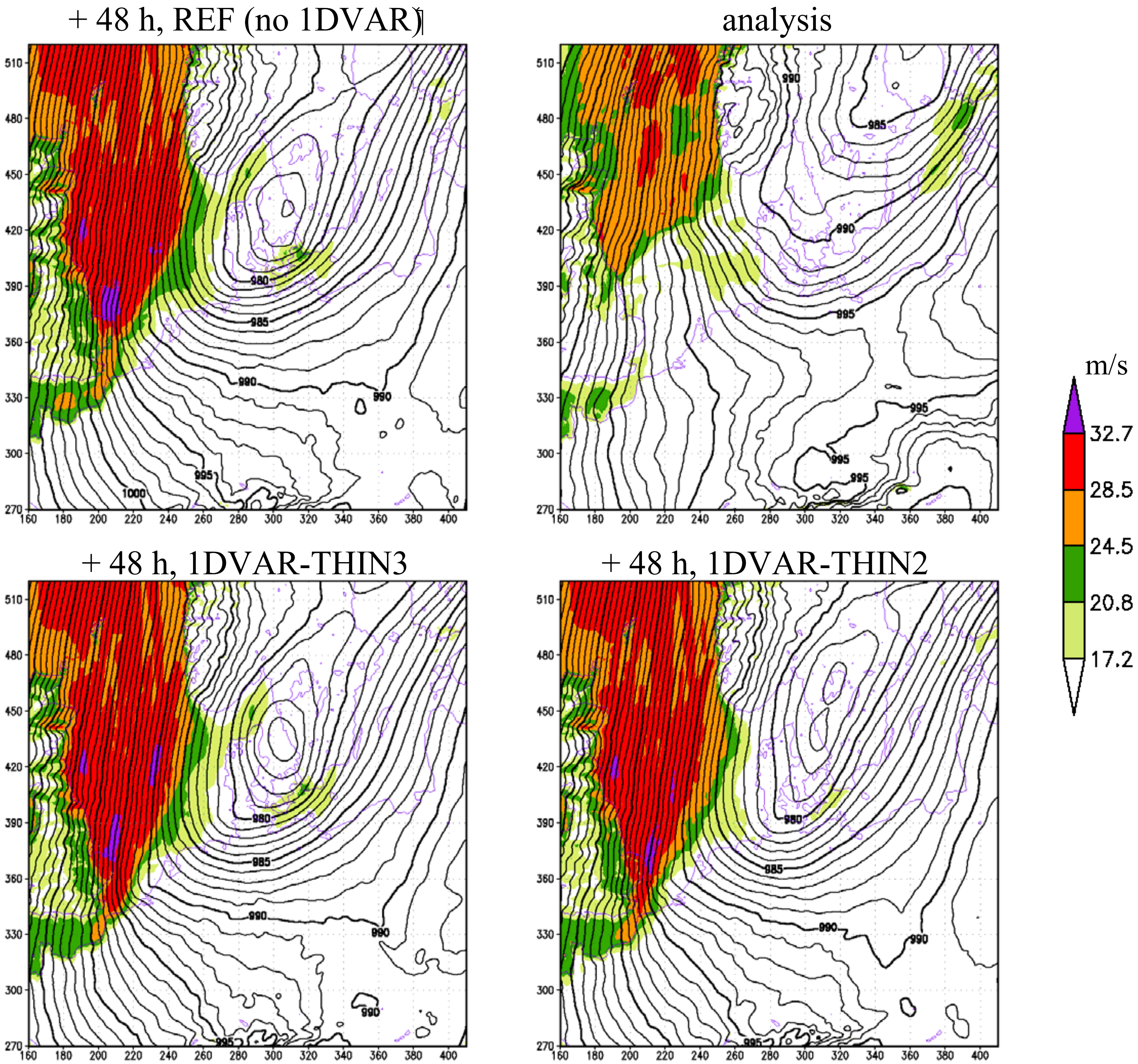
GME forecast for 48 hours



GME forecast for 60 hours

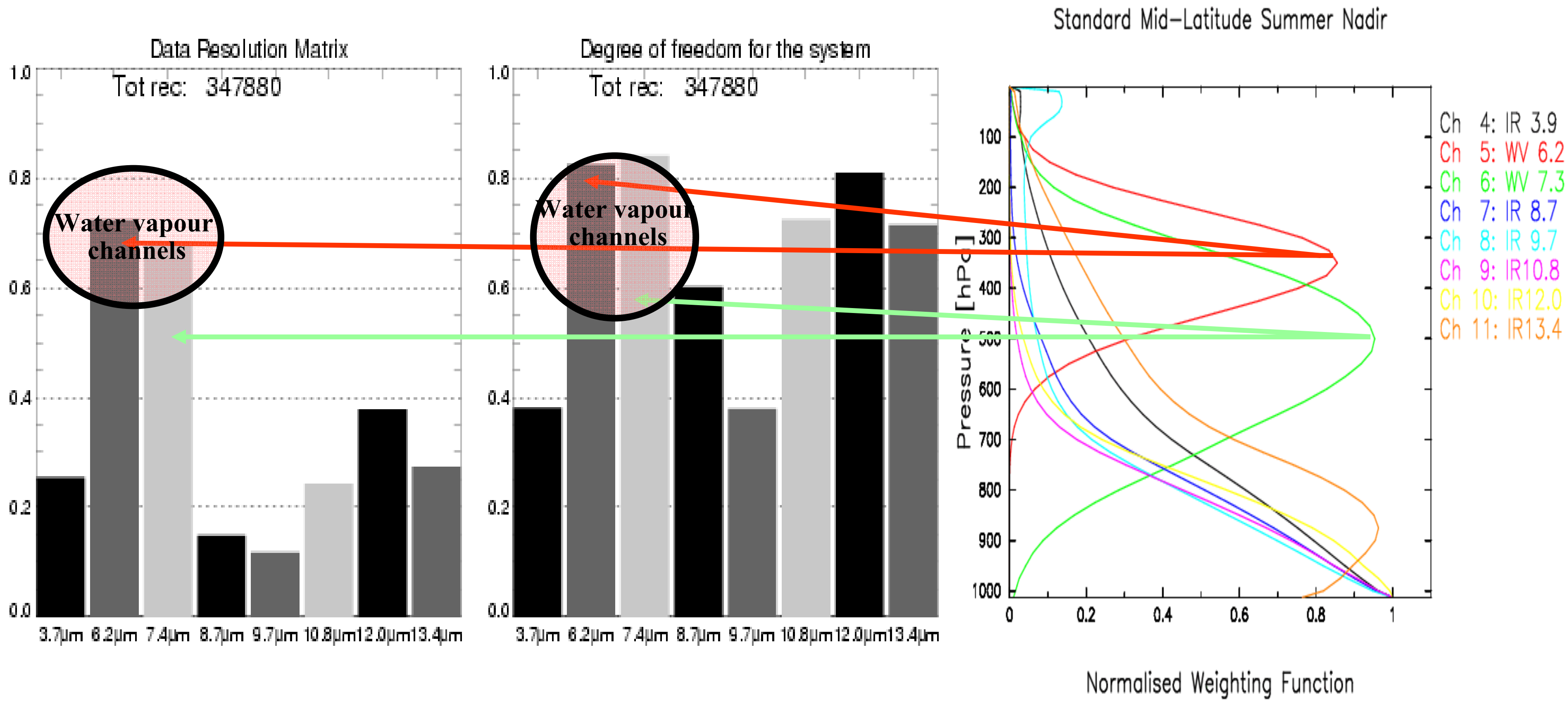


mean sea level pressure & max. 10-m wind gusts valid for 20 March 2007 , 0 UTC



Assimilation of SEVIRI/MSG

DRM e DFS: period 1st -21st of September 2006

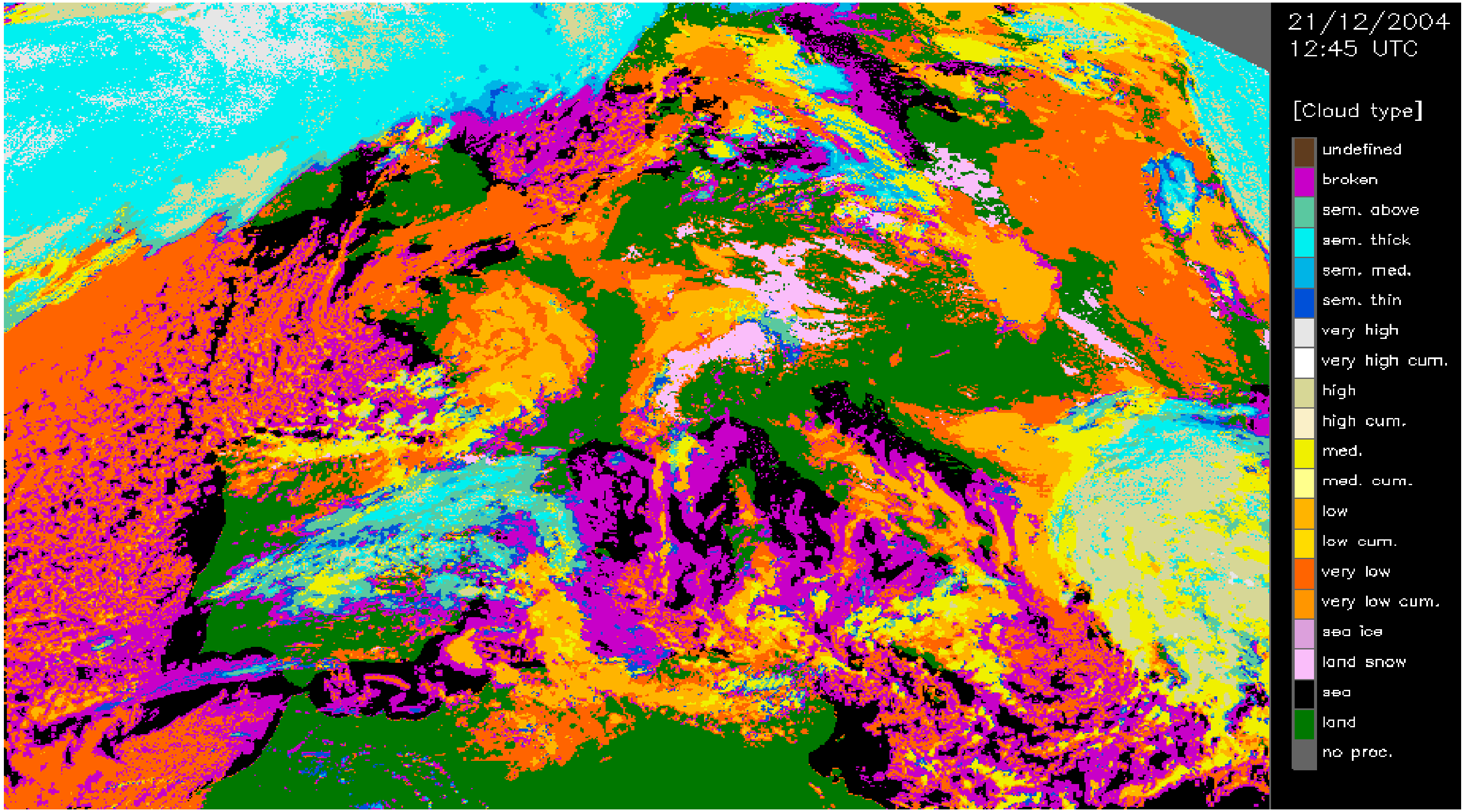


DRM tells the most useful channels in between the ones used

DFS instead is a measure of how much a channel in isolation is able to reduce the background error

Cloud Clearing for MSG

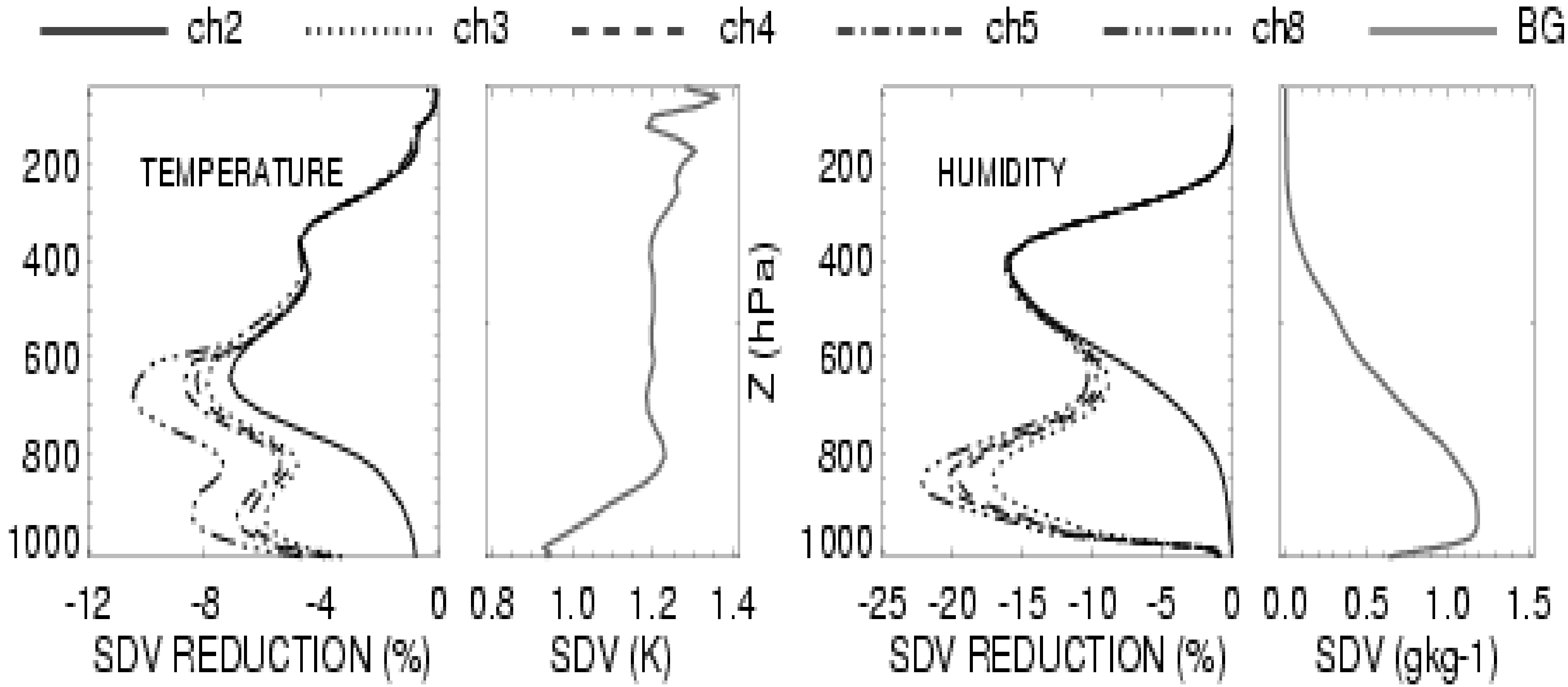
MSG: Cloud clearing using SAF-NWC software for MSG1 and MSG2



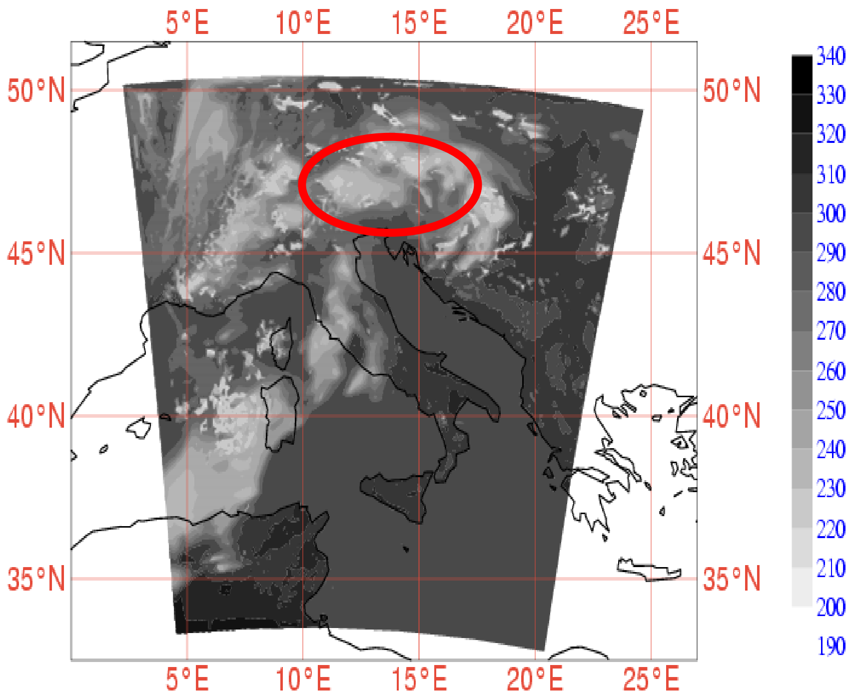
Cloud Type from SAF-NWC-PGE02

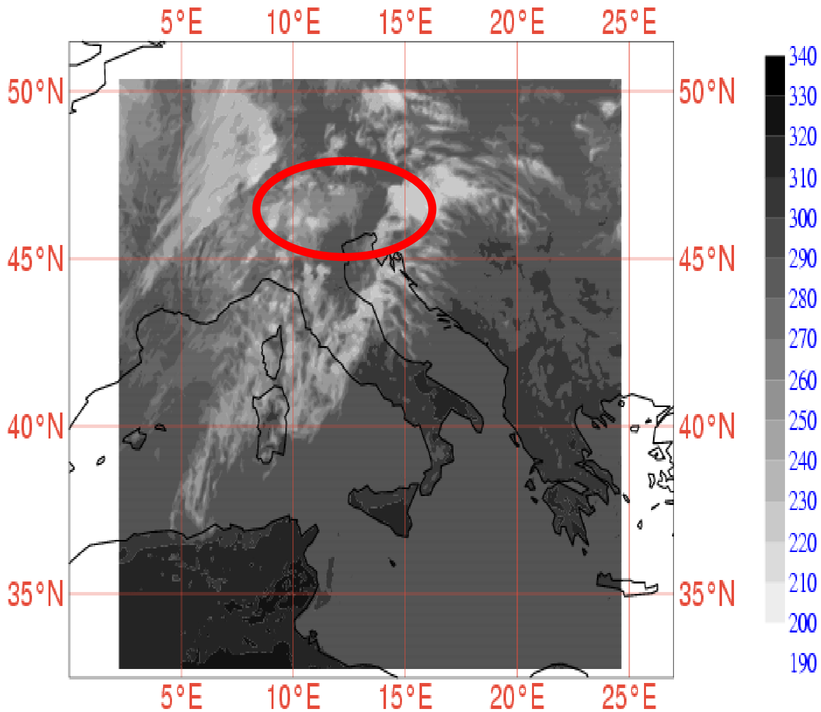
Analysis error reduction (Rogers)
period 1st -21st of September 2006

Set 2		6.2	7.3					
Set 3		6.2	7.3			10.8		
Set 4		6.2	7.3			10.8	12.0	
Set 5		6.2	7.3	8.7		10.8	12.0	
Set 8		6.2	7.3	8.7	9.7	10.8	12.0	13.4



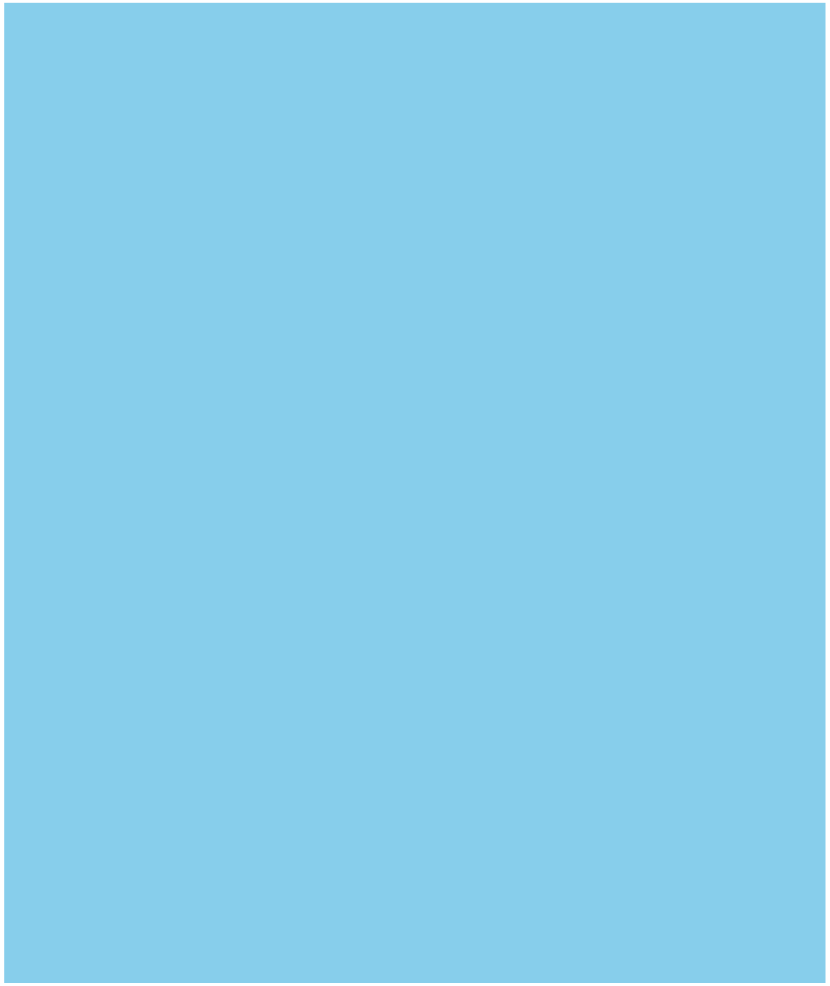
FIRST CASE STUDY: 8 July 2004

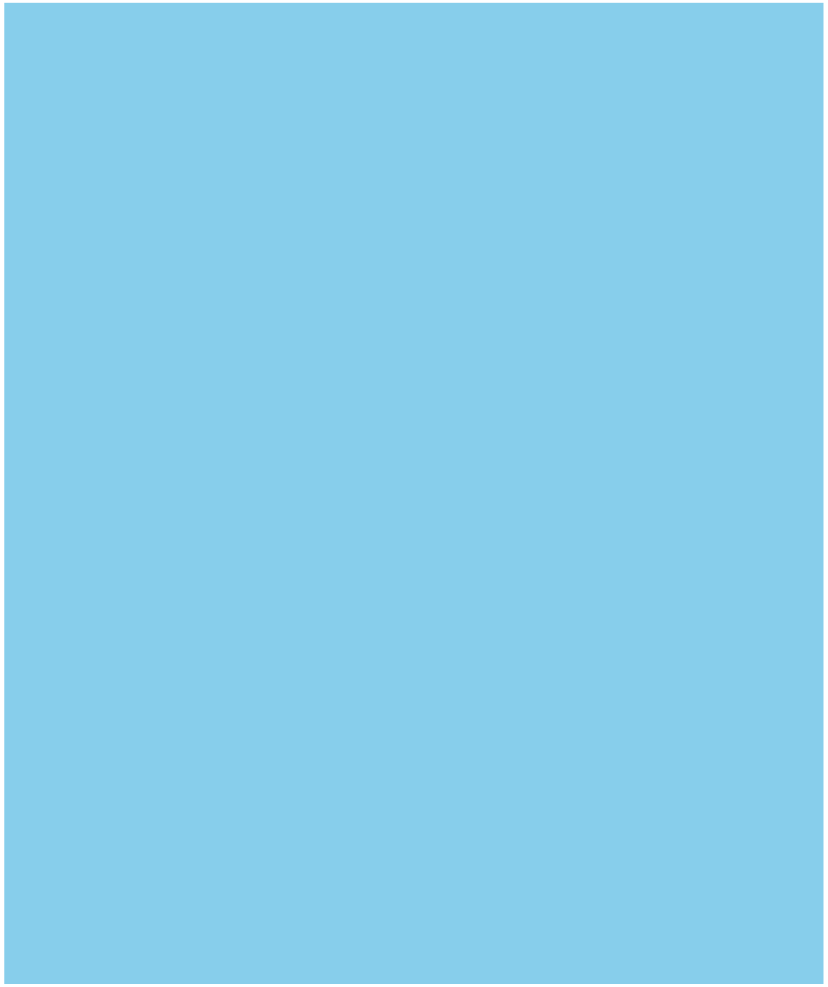




The selected case study is a false alarm occurred in North North-Eastern Italy, Trentino Alto Adige and Friuli-Venezia-Giulia, on the 8th of July 2004. A risk scenario was diagnosed by LM outputs. In particular a large atmospheric instability and convection events were forecasted. In reality the event was of minor intensity and drier winds with associated scattered thunderstorms were recorded only on the early morning of the 9th July.

CLOUD TABLE AND PROCESSING FLAG



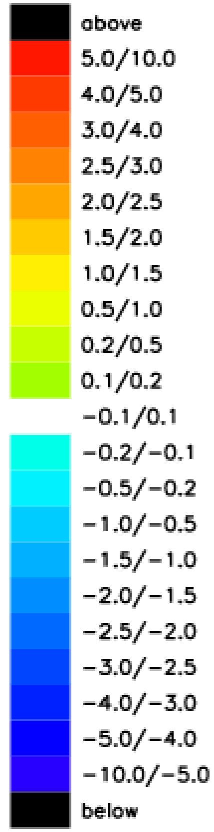


INTEGRATED WATER VAPOUR AND INTEGRATED SATURATION WATER VAPOUR INCREMENTS

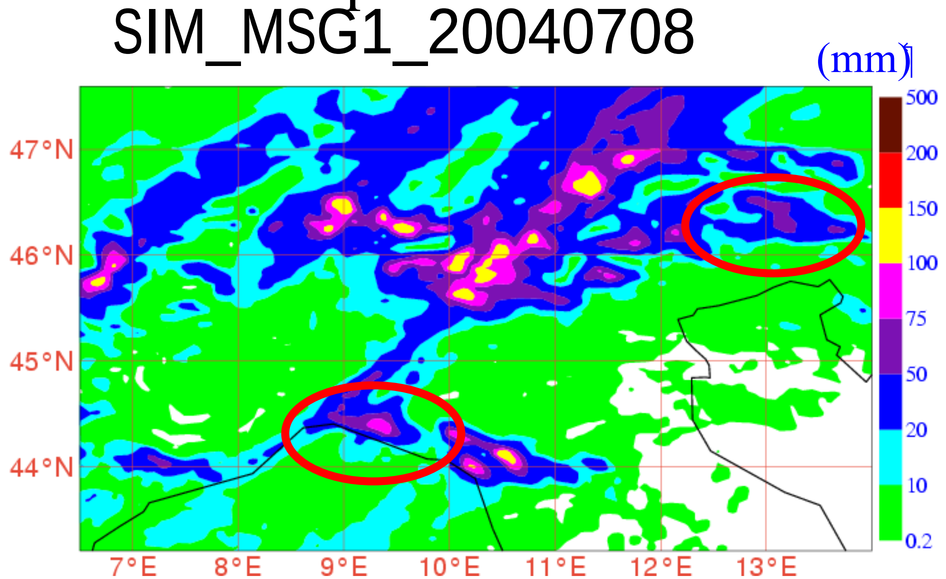
INT WV

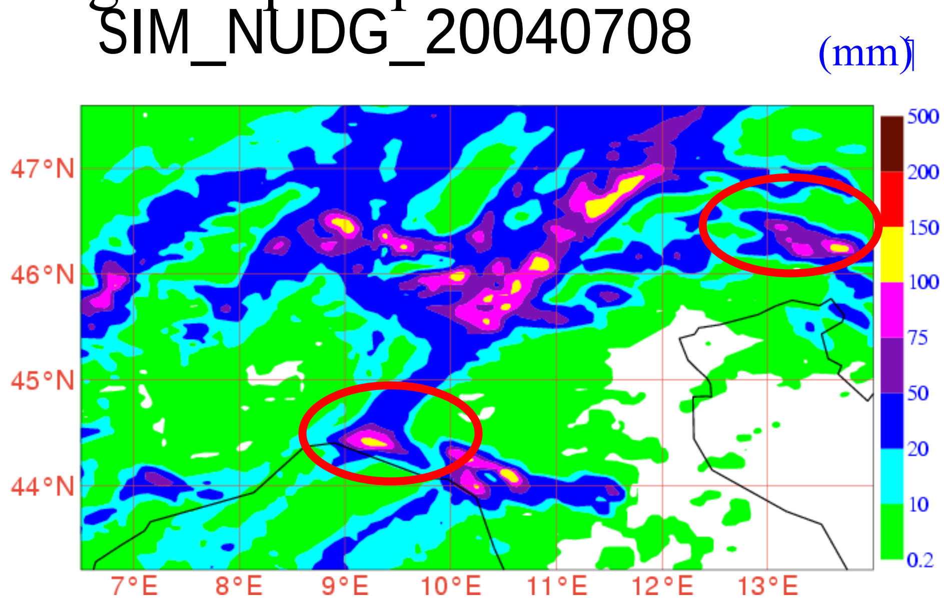
INT SAT WV (T)

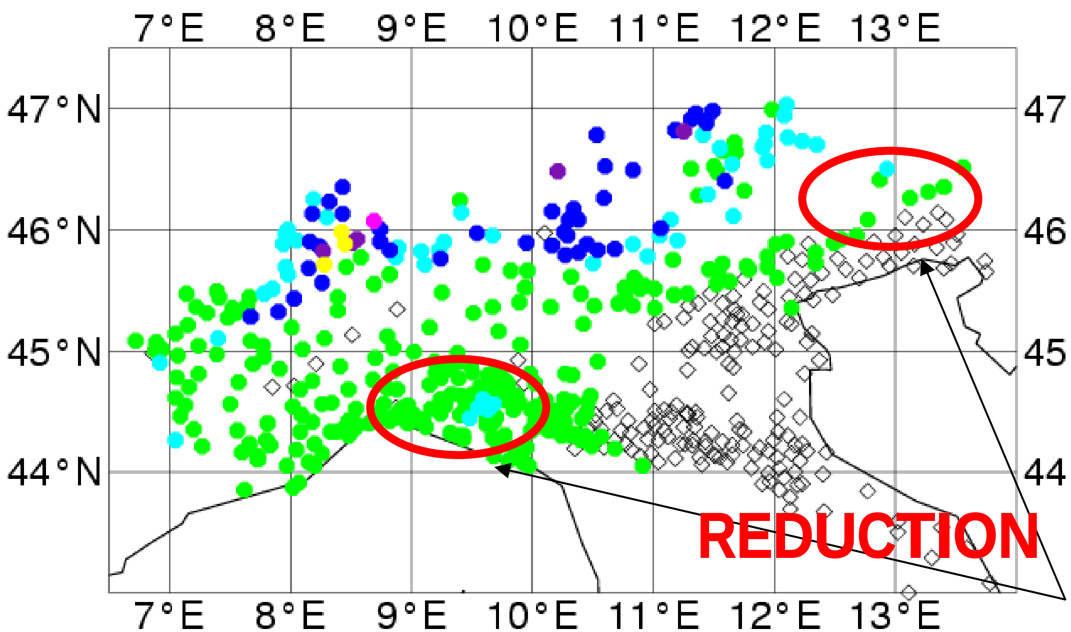
kg/m^2



Precipitation forecast: 24 hrs integrated precipitation

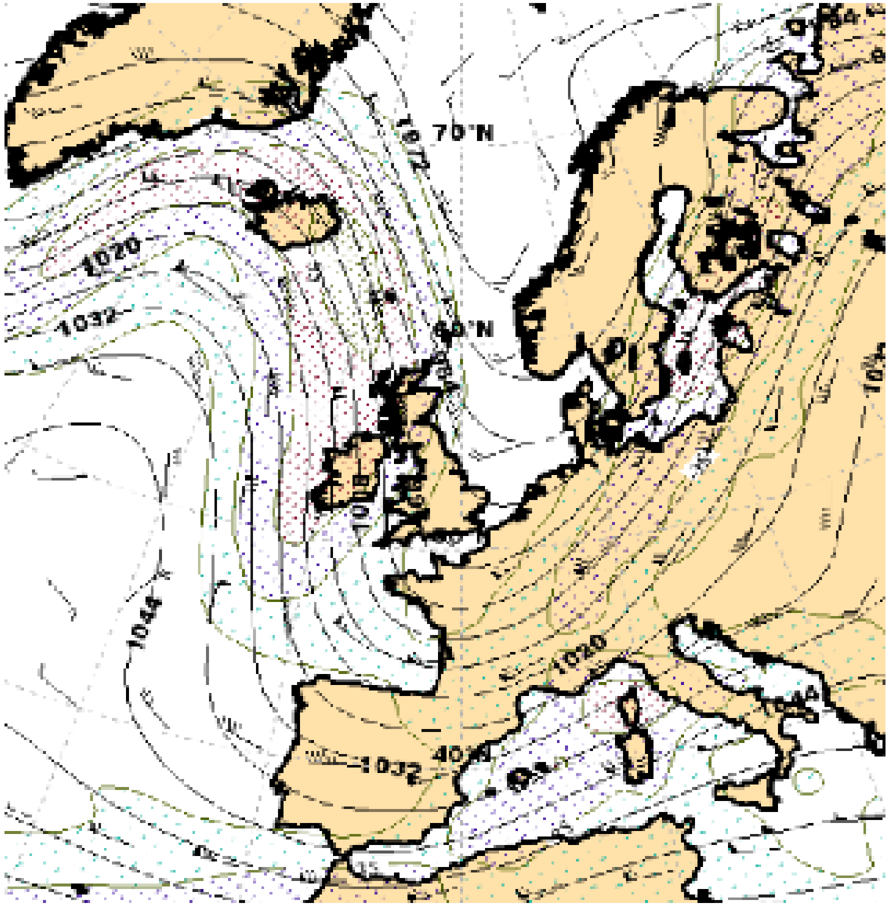




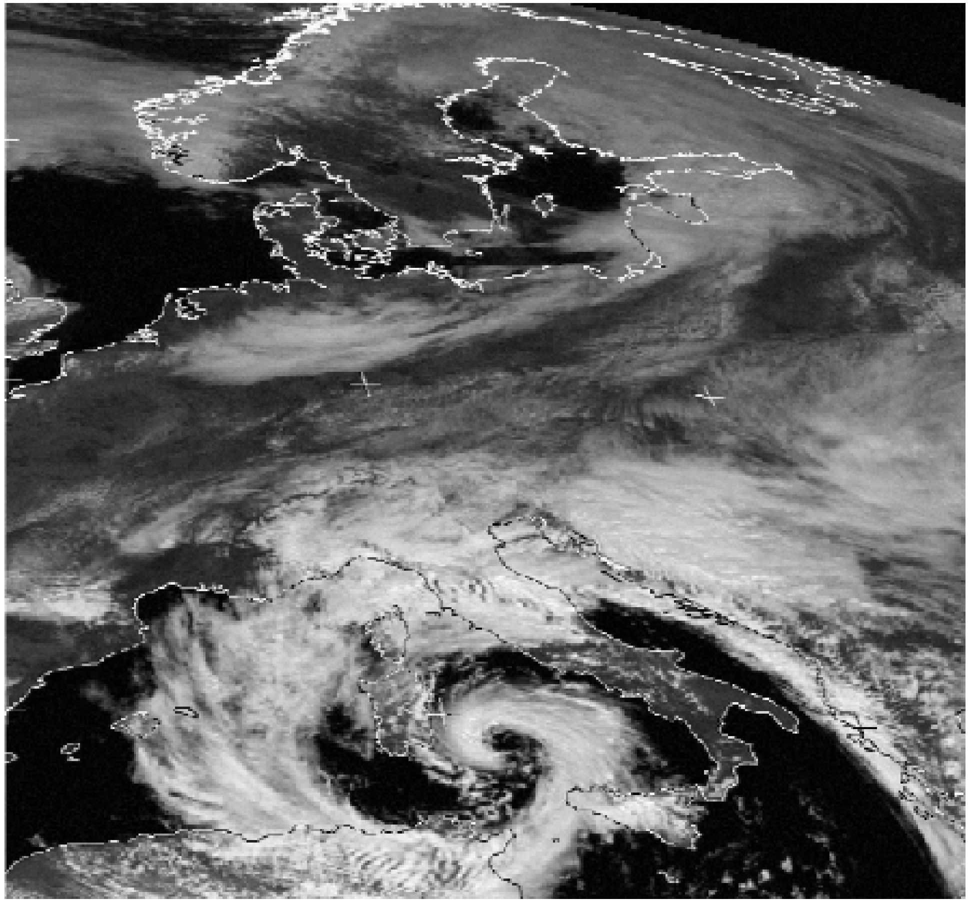


1. Both the forecasts do not correctly forecast the maximun of precipitation between 8 and 9 degrees of longitude and 46 degrees of latitude;
2. overestimation of the precipitation on Friuli-Venezia-Giulia region remains
3. the use of satellite data is able to reduce most of the overestimation of precipitation in the central Alpine area.

2nd case study : 9th of April 2005



250 hPa, 12UTC 8th

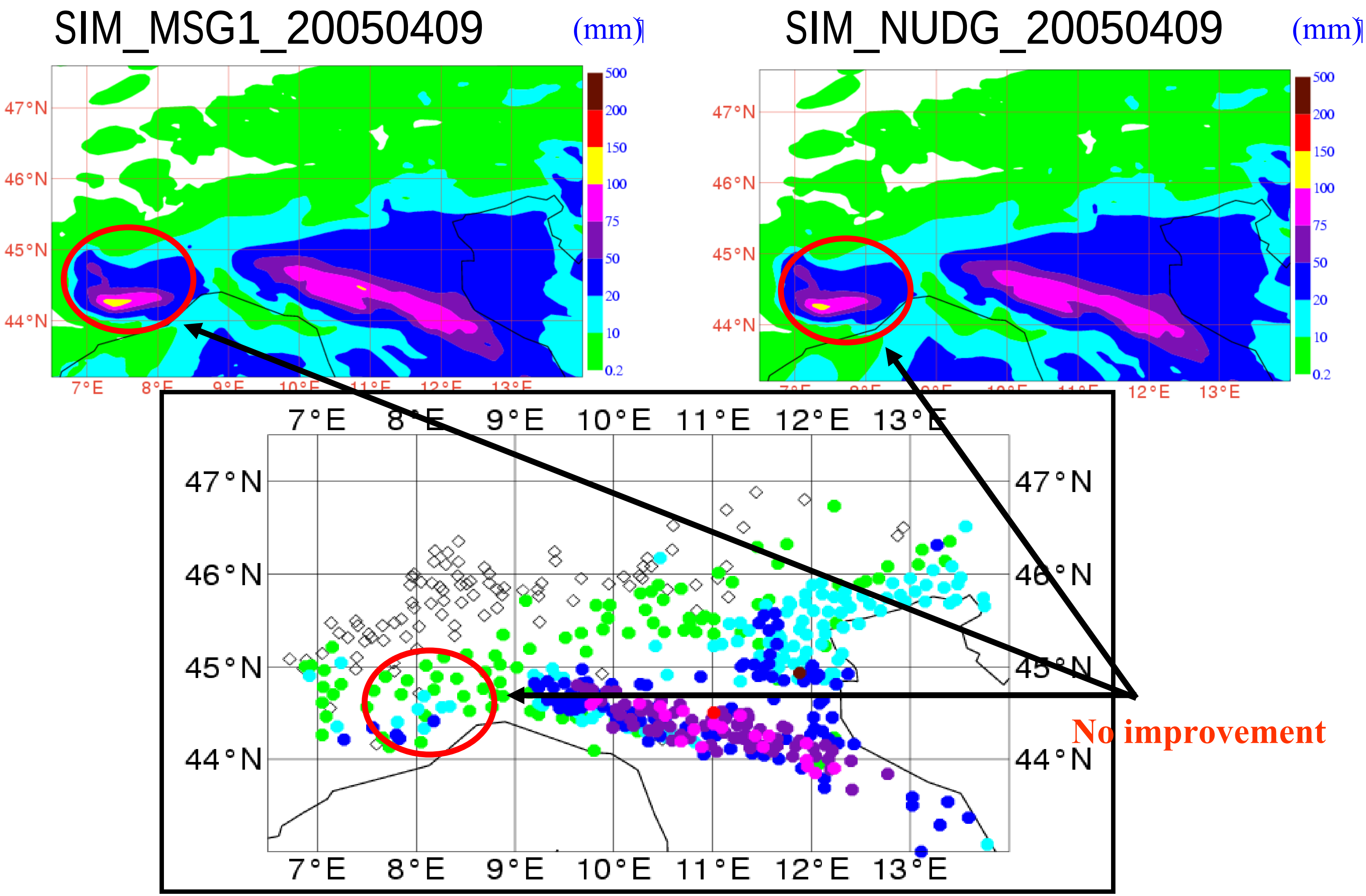


METEOSAT 7, 12 UTC 11th

Forecast of heavy precipitation in the liguria region. Typically produced by south-westerly up-stream flow due to orographic forcing

CLOUD TABLE AND PROCESSING FLAG

Precipitation forecast



Conclusions for MSG:

- Wealthy information for temperature and humidity
- Use information of clouds and rain
  - to improve data coverage
  - to avoid spreading „dry“ (cloud free) increments into a „wet“ (cloudy) region

Conclusions for AMSU:

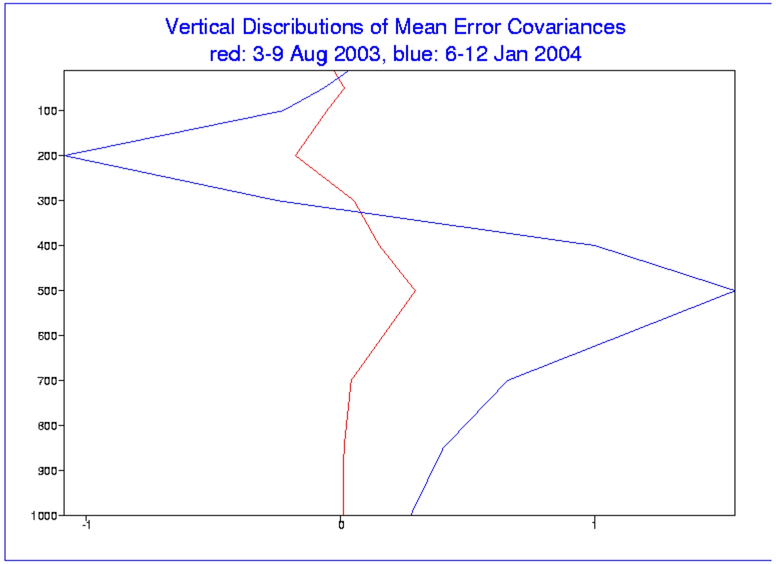
- similar approaches for limited area models, some complications: (stratosphere, bias correction, background errors)
- use information over land
- expect less impact on forecast quality than for global models

Conclusions for the assimilation of satellite data in limited area models:

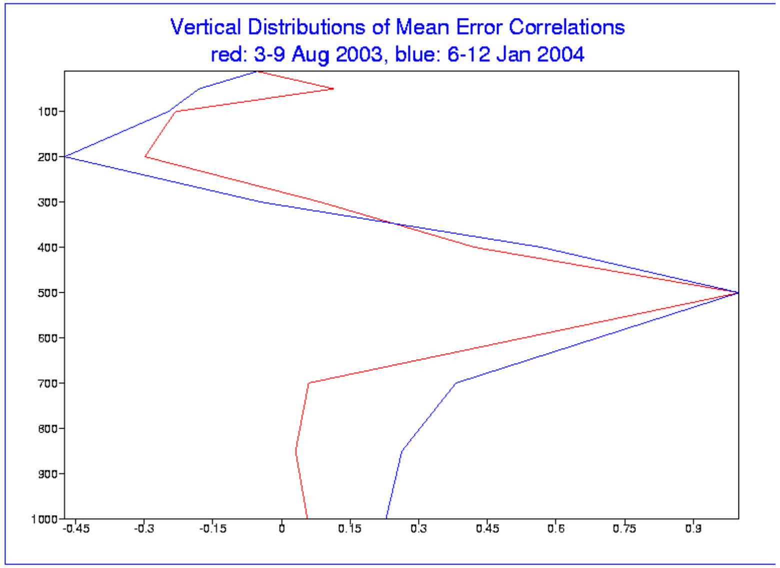
- with increasing resolution and more sophisticated physics
  more flexible background error structures are required
  (especially for sounders with high vertical resolution)
- variational methods: increasing interest for adaptive, situation or flow dependent background error covariances
- nudging: fit for high resolution observations
  tuning without mathematical framework required
- combine idea of nudging with variational developments
  (framework for nudging weights)
- interest for ensemble methods in order to prescribe adaptive background errors

Background error covariance matrix B

vertical error stuctures derived from IFS
blue: westerly winds, red: stable high pressure



covariances with 500hPa



correlations with 500hPa

Thank You for attending