Overview of Data Assimilation Activities in COSMO

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

christoph.schraff@dwd.de
Deutscher Wetterdienst, D-63067 Offenbach, Germany

current DA method: nudging

PP KENDA for (1 – 3) km-scale EPS: LETKF

radar-derived precip:
- latent heat nudging

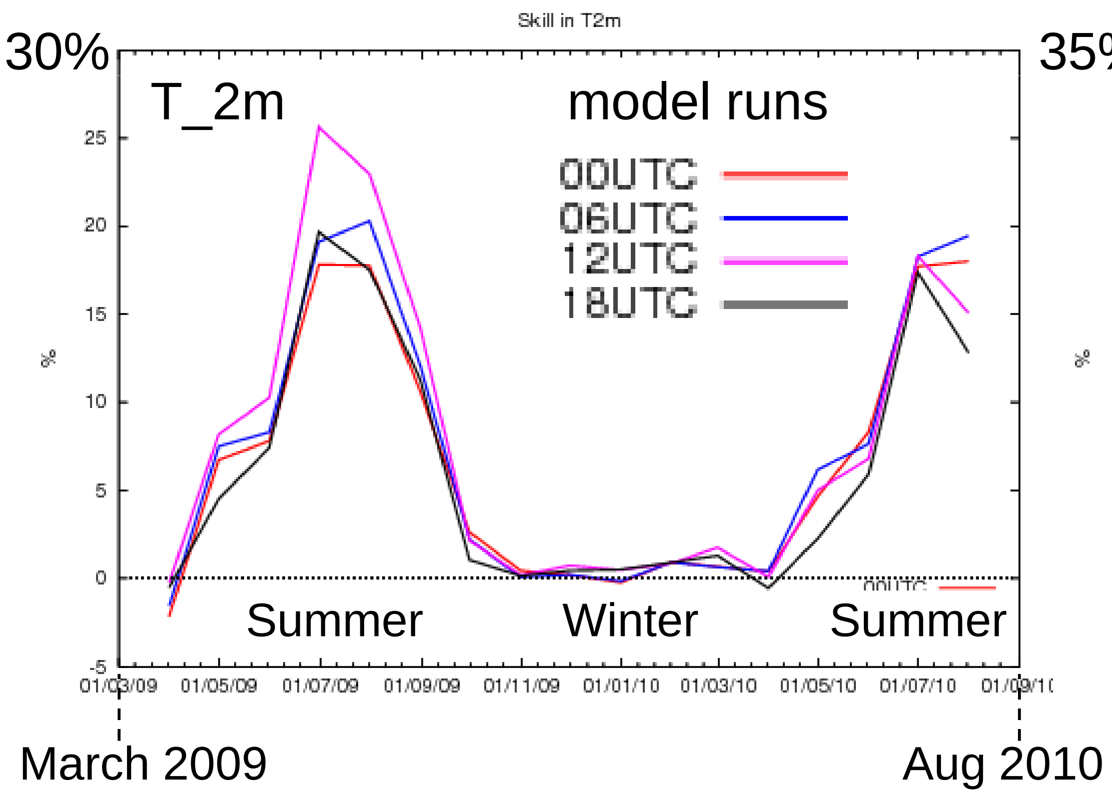
LETKF for HRM
(28 km → 10 km, opr)

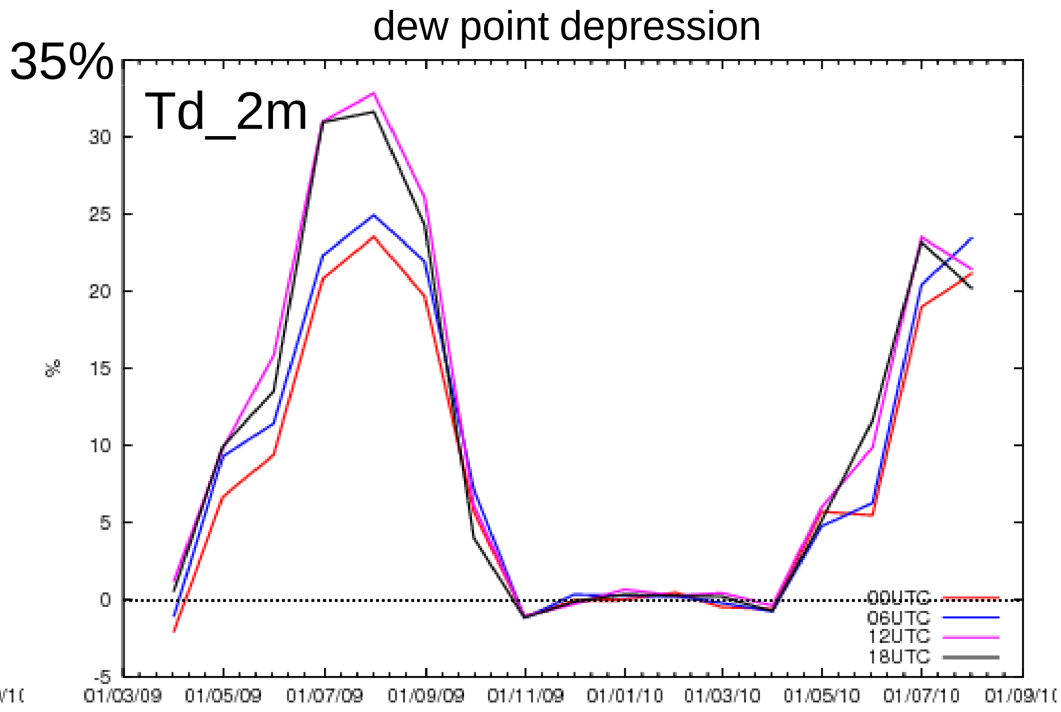
Long-term impact of LHN assimilation on soil moisture ?

Experiment: 18 months (April 2009 – August 2010) COSMO-DE without LHN
→ compare with operational COSMO-DE with LHN

(COSMO-DE: no soil moisture analysis)

monthly relative forecast skill (1 – rmse(LHN) / rmse(noLHN))





→ LHN → higher soil moisture content → improved T-2m , Td-2m forecasts,
benefit lasts over whole forecast time

Overview on Data Assimilation Activities in COSMO

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

christoph.schraff@dwd.de
Deutscher Wetterdienst, D-63067 Offenbach, Germany

current DA method: nudging

PP KENDA for (1 – 3) km-scale EPS: LETKF

radar-derived precip:
- latent heat nudging
- 1DVar (new tests: ~ neutral)

radar radial wind:
- VAD (at DWD: no benefit)
- nudge V_r (being implemented)

GPS–IWV (ceased)

screen-level:
- scatterometer wind (OSCAT, OceanSat2)
- improved T2m assimilation, adapt T_soil

Satellite radiances, 1DVAR:
- AMSU-A, SEVIRI, IASI (no benefit, ceased)

LETKF for HRM
(28 km → 10 km, opr)

daytime biases in summer:
- Vaisala RS-92 humidity obs dry bias
  → bias correction (Miloshevich et al., *J. Geophys. Res.* 2007) operational
- model warm bias (low levels)

surface (snow …) → PP COLOBOC

Status Overview of PP KENDA
Km-scale ENsemble-based Data Assimilation

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

christoph.schraff@dwd.de
Deutscher Wetterdienst, D-63067 Offenbach, Germany

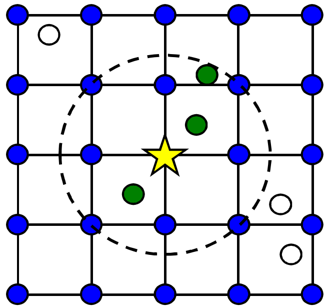
Contributions / input by:

Hendrik Reich, Andreas Rhodin, Roland Potthast, Uli Blahak, Klaus Stephan (DWD)
Annika Schomburg (DWD / Eumetat)
Yuefei Zeng, Dorit Epperlein (KIT Karlsruhe)
Daniel Leuenberger, Manuel Bischof (MeteoSwiss)
Mikhail Tsyrulnikov, Vadim Gorin (HMC)
Lucio Torrisi (CNMCA)
Amalia Iriza (NMA)

focus: development of **LETKF** (Local Ensemble Transform Kalman Filter)

LETKF (COSMO) : method

- implementation following Hunt et al., 2007
- basic idea: do analysis in the space of the ensemble perturbations
  - computationally efficient, but also restricts corrections to subspace spanned by the ensemble
  - explicit localization (doing separate analysis at every grid point, select only obs in vicinity)
  - analysis ensemble members are locally linear combinations of first guess ensemble members

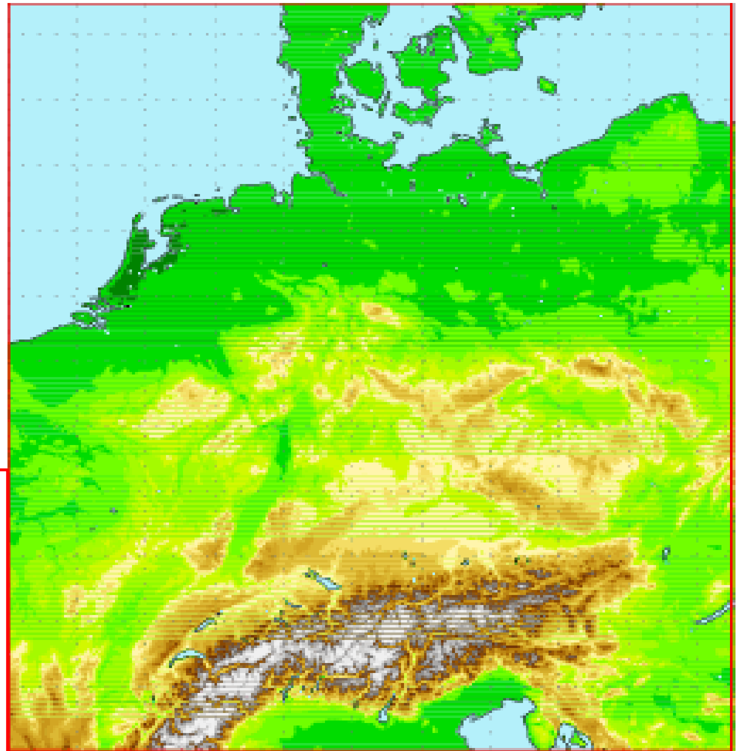


LETKF for km-scale COSMO : set-up of preliminary experiments

- use **conventional obs** (TEMP, AIREP, SYNOP, Wind Profiler)
- **3-hourly** cycles (and only up to 2 days: 7 – 8 Aug. 2009: quiet + convective day)
  $\rightarrow$ (near) future: 1-hourly / 30-min / 15-min cycles
- lateral (and upper) boundary conditions (BC)
  – **ensemble BC** from COSMO-SREPS (3 * 4 members),
  $\rightarrow$ future: ens. BC from global LETKF (GME/ICON)
- ensemble size: **32** ($\rightarrow$ near future: ~ 40)
- 'deterministic run' verified against nudging ana / obs
  (det. analysis by applying EnKF Kalman Gain
  to obs increments of det. f.g.)

COSMO-DE: Δx = 2.8 km
(deep convection explicit,
shallow convection param.)

domain size : ~ 1250 x 1150 km



christoph.schraff@dwd.de

Analysis for a deterministic forecast run : use Kalman Gain $\mathbf{K}$ of analysis mean

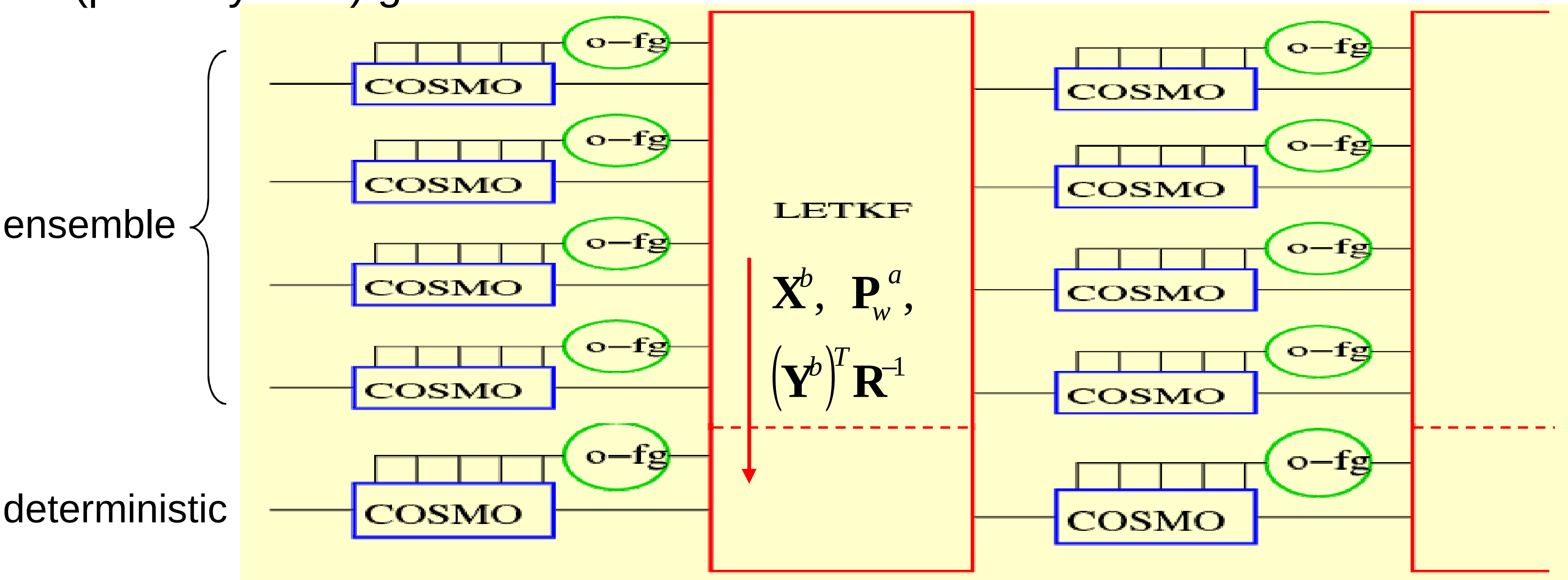
Deutscher Wetterdienst
Wetter und Klima aus einer Hand

$$\mathbf{x}^A = \mathbf{x}^B + \mathbf{K}\left[\mathbf{y}^O - H(\mathbf{x}^B)\right] \quad , \qquad \mathbf{K} = L\ \mathbf{X}^b \left(\mathbf{P}_w^{\ a}\right)^{-1} \left(\mathbf{Y}^b\right)^T \mathbf{R}^{-1}$$

$$\mathbf{P}_w^{\ a} = (k-1)\,\mathbf{I} + \left(\mathbf{Y}^b\right)^T \mathbf{R}^{-1} \mathbf{Y}^b$$

deterministic analysis recently implemented

L : interpolation of analysis increments from grid of LETKF ensemble to (possibly finer) grid of deterministic run



- deterministic run must use same set of observations as the ensemble system !
- Kalman gain / analysis increments not optimal, if deterministic background $\mathbf{x}^B$ (strongly) deviates from ensemble mean background

LETKF (km-scale COSMO) : scientific issues / refinement

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

- lack of spread: (partly) due to model error and limited ensemble size which is not accounted directly by the algorithm

  to account for it: covariance inflation, what is needed ?

  → multiplicative $X_b \to \rho \cdot X_b$ (tuning, or adaptive ($y - \mathbf{H}(x) \sim \mathbf{R} + \mathbf{H}^T\mathbf{P}_b\mathbf{H}$))

  → additive : perturbing the NWP model

  - fixed perturbations of model physics parameters : no
  - stochastic physics (being implemented, Torrisi CNMCA)
  - statistical 3DVAR-**B** → hybrid schemes !
  - additive inflation which reflects model error as estimated by statistics (comparing forecast tendencies with observed tendencies, Gorin & Tsyrulnikov)

- adaptive multiplicative covariance inflation ρ :
  - compare 'observed' quantities (errors) with 'expected' ones:
    (Desroziers et al., QJRMS 2005; Li et al., QJRMS 2009)

$$1. \quad \left\langle (y - H(x_b))(y - H(x_b))^T \right\rangle = \mathbf{R} + \rho\, \mathbf{HP}_b\mathbf{H}^T$$

$$2. \quad \left\langle (H(x_a) - H(x_b))(y - H(x_b))^T \right\rangle = \rho\, \mathbf{HP}_b\mathbf{H}^T$$

aim: satisfy statistics

problem: 1. depends heavily on $\mathbf{R}$, 2. found to give too small values ρ

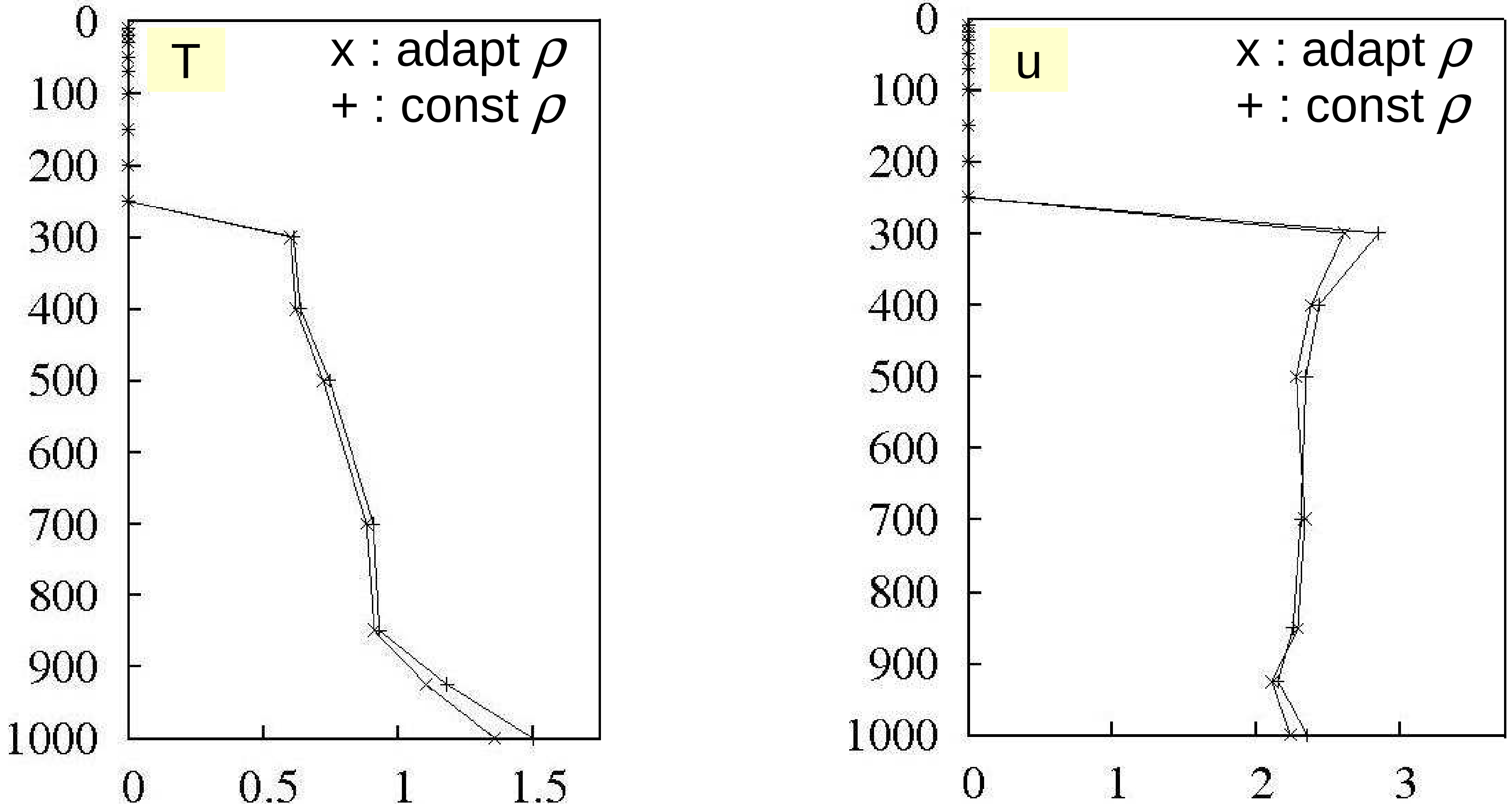
  - relaxation to prior spread (RTPS) :
    (Whitaker et al., 2010)

$$\rho = \sqrt{\alpha \frac{\sigma_b - \sigma_a}{\sigma_a} + 1} \quad , \quad \alpha \le 1$$

aim: compensate reduction of spread due to assimilation of obs

potential problem: long-term drift ; implemented in ensemble space instead of model space $\rightarrow$ ρ too small adaptive covariance inflation ρ , by statistical method

f.g. rmse against AMDAR

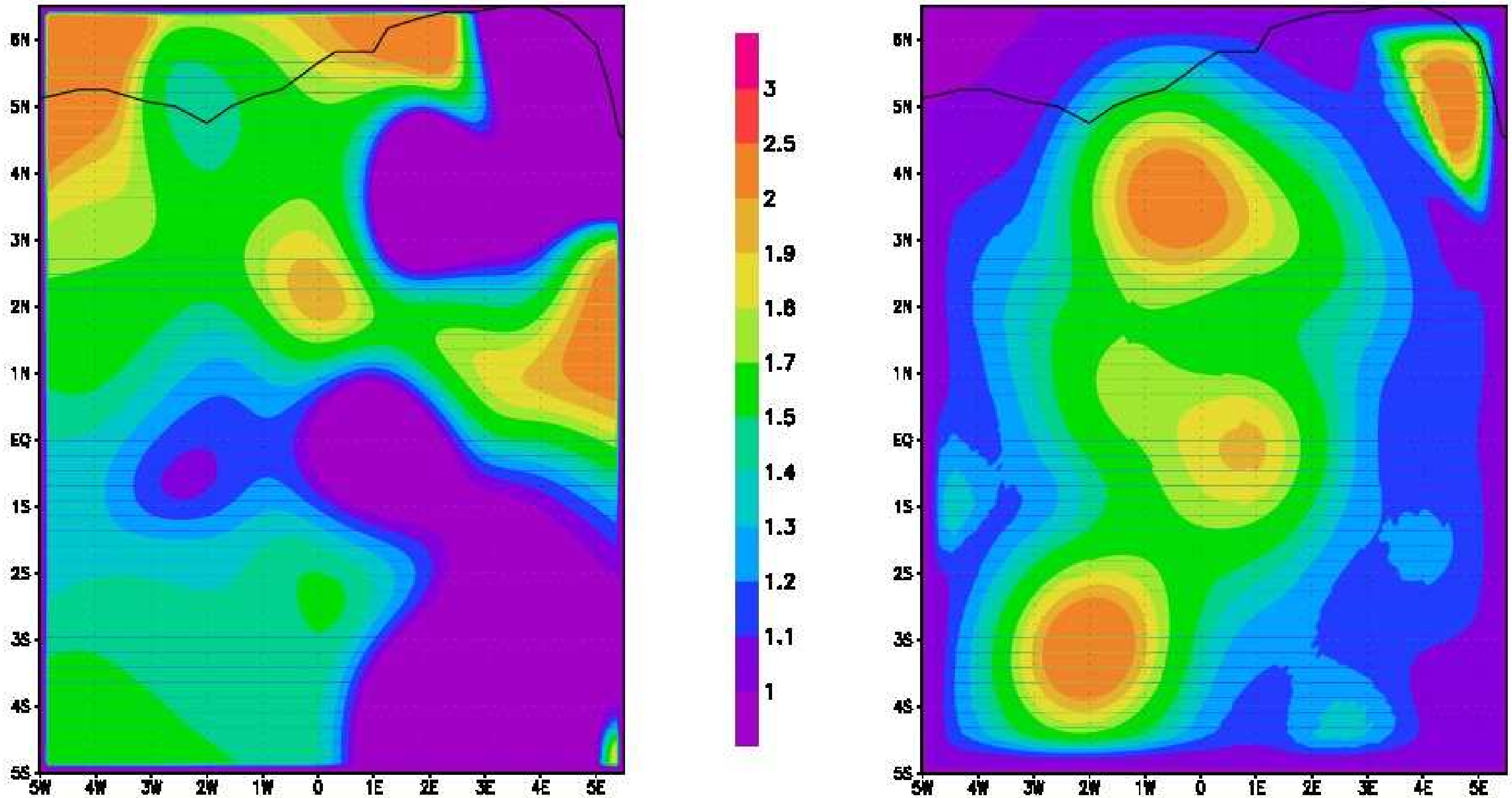


$\rightarrow$ rmse : slight positive impact

$\rightarrow$ spread : increased

adaptive covariance inflation ρ

statistical method 1. (obs space) relaxation method (ens space)



$\rightarrow$ very different spatial patterns of ρ

$\rightarrow$ similar results (area avg. rmse / spread) , preliminary !

- adaptive estimation of obs error covariance $\mathbf{R}$

(Li, Kalnay, Miyoshi, QJRMS 2009) , but our implementation: in ensemble space !

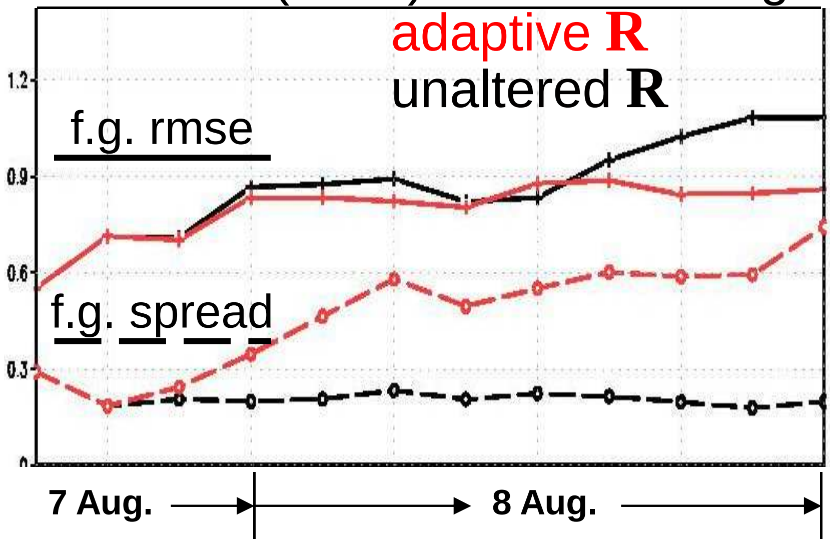
add obs, if already $N_{obs} > N_{ens}$:

- cannot be fitted well, improve analysis only slightly (rank deficiency problem)
- decreases analysis error !

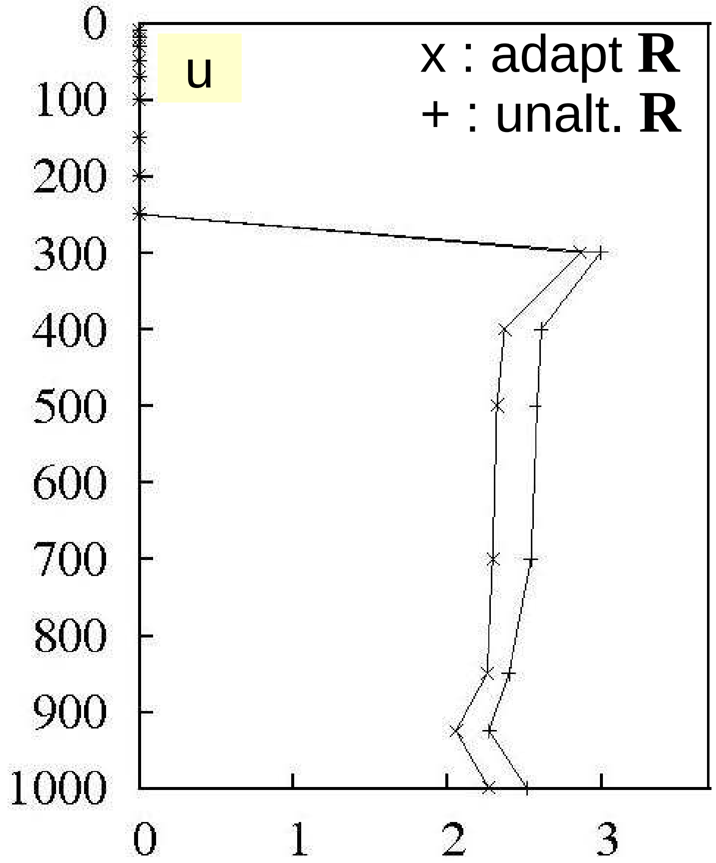
$$\mathbf{P}_w^{\ a} = \left[(k-1)\,\mathbf{I} + \left(\mathbf{Y}^b\right)^T \mathbf{R}^{-1}\mathbf{Y}^b\right]^{-1}$$

$\rightarrow$ adaptive $\mathbf{R}$ in ensemble space takes rank deficiency problem into account and increases $\mathbf{R}$

adaptive obs error covariance **R**



(exp. done with 'old setting' for pre-specified **R**)



→ rmse : positive impact in most levels / variables

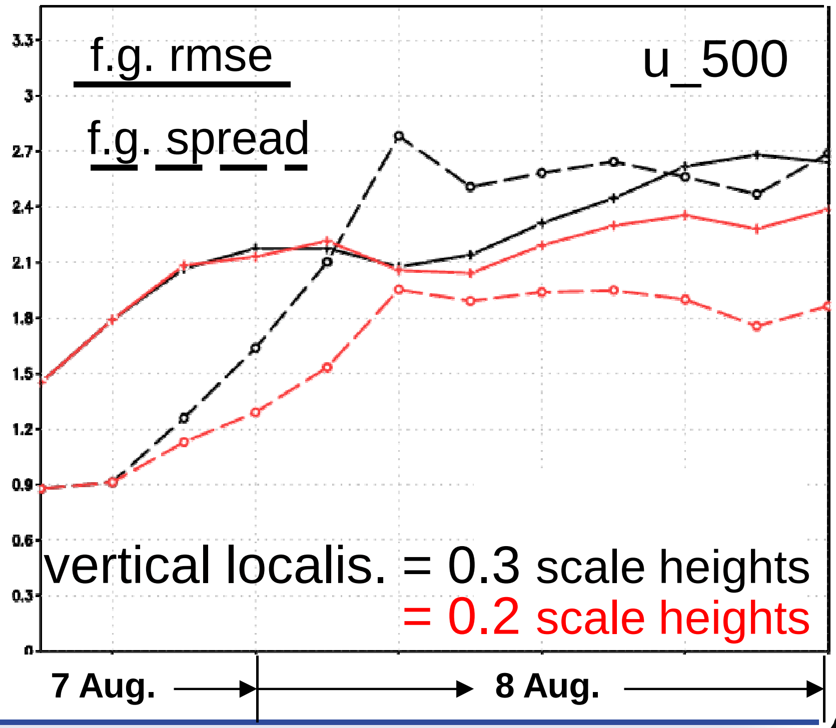
→ spread : strongly increased

LETKF (km-scale COSMO) : scientific issues / refinement

- adaptive methods important when N_{obs} large
  (strongly increasing N_{obs} degraded results, if previous setting used !)

- need to optimise setup of adaptive methods :
  - \+ need for careful specification / tuning of obs errors
  - – large N_{obs} : adaptive increase of $\mathbf{R}$ indicates non-optimal use of obs

- localisation !



LETKF (km-scale COSMO) : scientific issues / refinement

- **localisation** (multi-scale data assimilation,
  successive LETKF steps with different obs / localisation ?
  adaptive , dep. on obs density ?)

- **update frequency** $\Delta_a t$? 1 hr $\geq$ $\Delta_a t$ $\geq$ 15 min
  non-linearity vs. noise / lack of spread / 4D property ?

- perturbed lateral BC , noise control ?

- non-linear aspects, convection initiation (outer loop , (latent heat nudging) ?)

- (technical aspects: efficiency, system robustness)

LETKF (km-scale COSMO) : some important observations at km scale

- **radar : direct 3-D radial velocity & 3-D reflectivity**
  sufficiently accurate and efficient observation operators soon available

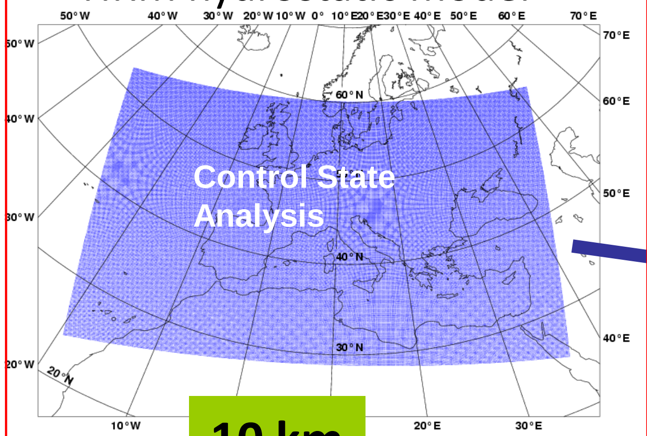
- **ground-based GPS slant path delay** (start Jan. 2012)
  direct use in LETKF, or tomography

- **cloud information based on SEVIRI and conventional data**
  (Annika.Schomburg@dwd.de , DWD / Eumetsat, start 2011-03)

L. Torrisi, Marcucci et al. (CNMCA, Rome) M. Bonavita (ECMWF)

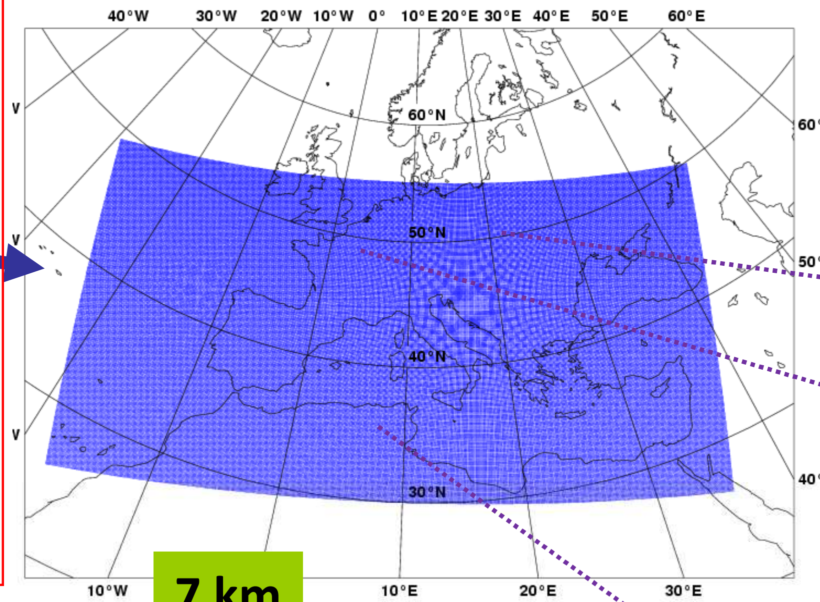
LETKF analysis ensemble (40 members) every 6h
using TEMP, PILOT, SYNOP, SHIP, BUOY, Wind Profiler, AMDAR-ACAR-AIREP, MSG AMV, METOP/ERS2 scatt. winds (soon: NOAA/METOP AMSUA radiances)
+ Land SAF snow mask, IFS SST analysis once a day

HRM hydrostatic model



10 km
40 L

COSMO-ME (7km) ITALIAN MET SERVICE



7 km
40 L

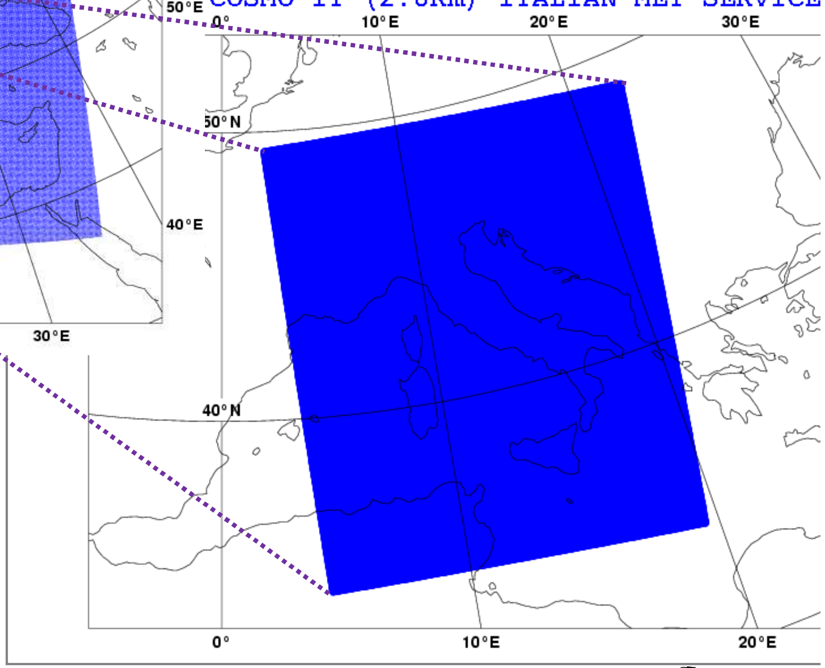
COSMO

- compressible equations
- parameterized convection

2.8 km
50 L

- compressible equations
- explicit convection

COSMO-IT (2.8Km) ITALIAN MET SERVICE



In the CNMCA LETKF implementation, model errors and sampling errors are taken into account using:

- state dependent **multiplicative** inflation according to Whitaker et al (2010)

ana. perturb. $$\mathbf{x}_{\mathrm{a}}' = \mathbf{x}_{\mathrm{a}}' \sqrt{\alpha \frac{\sigma_{\mathrm{b}}^2 - \sigma_{\mathrm{a}}^2}{\sigma_{\mathrm{a}}^2} + 1}$$

α = 0.95
σ^2 = variance

- climatological **additive** noise

ana. member $$\mathbf{x}_i^a \leftarrow \mathbf{x}_i^a + \alpha \mathbf{x}_i^n, \quad \alpha \mathbf{x}_i^n \sim N(0, \mathbf{Q})$$

α : scale factor

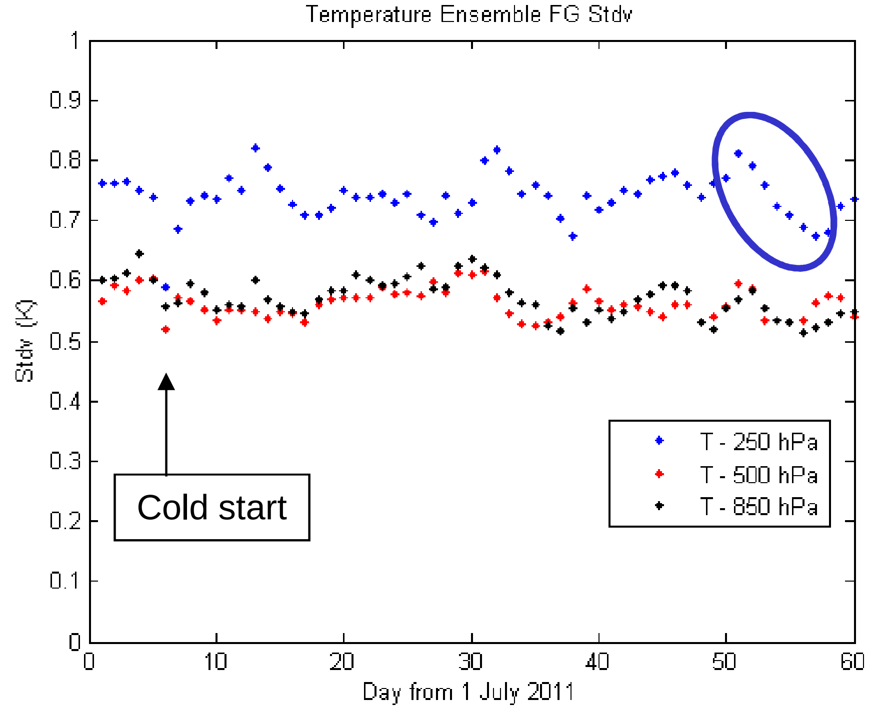
$\mathbf{x}_i^n$ randomly selected, 48-24h forecast differences

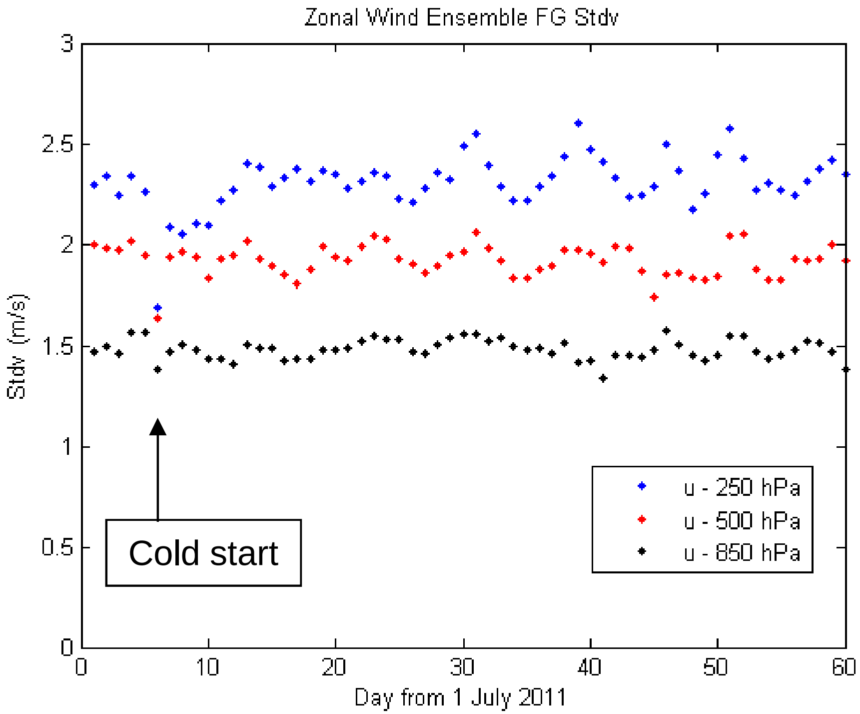
- lateral boundary condition perturbation using EPS
- climatologic perturbed SST

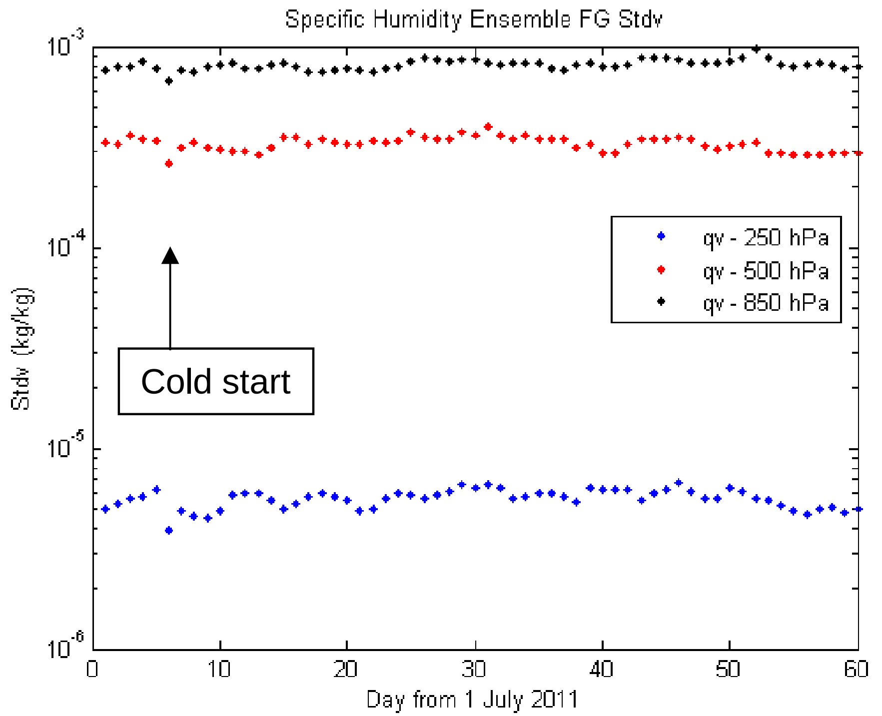
CNMCA LETKF:

time series (60 days, 0 UTC) of ensemble background spread

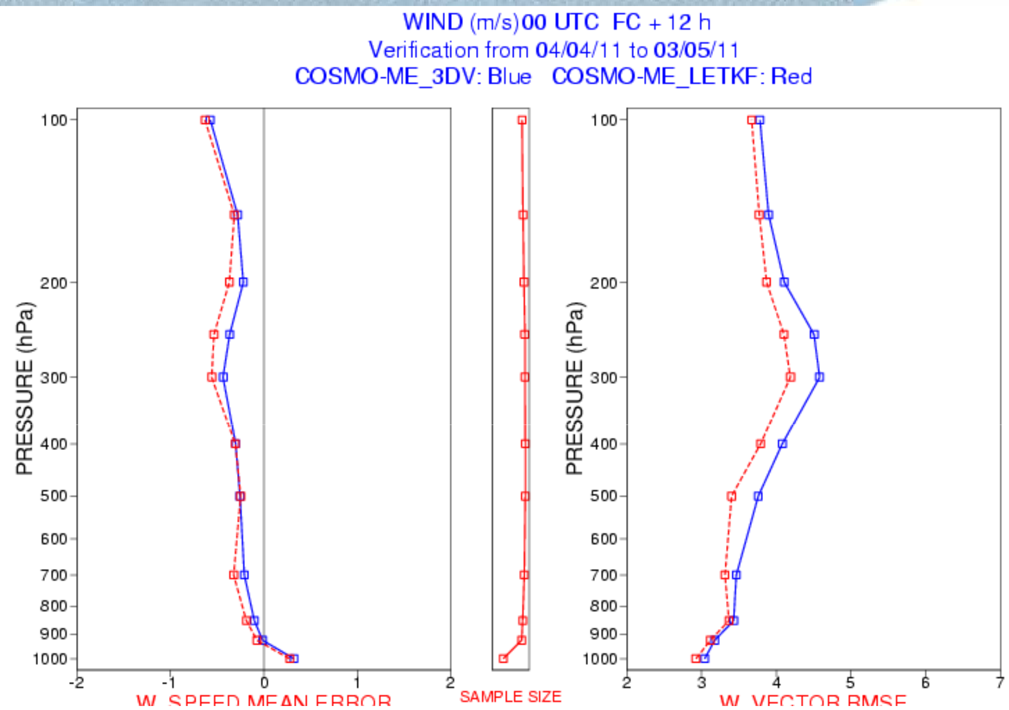
→ spread remains stable







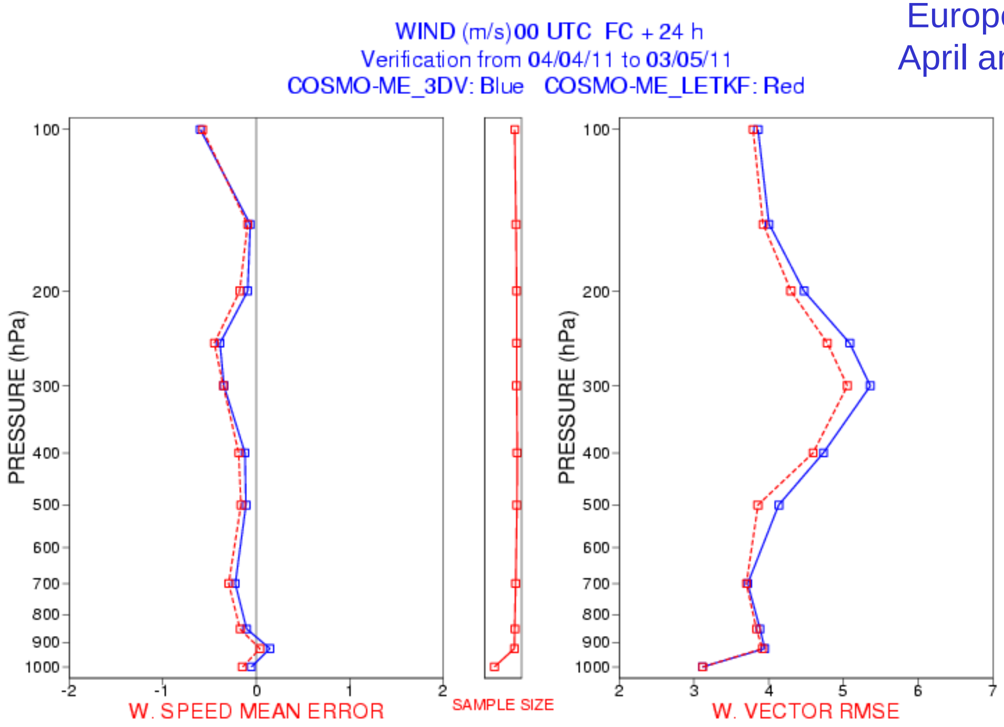
CNMCA LETKF

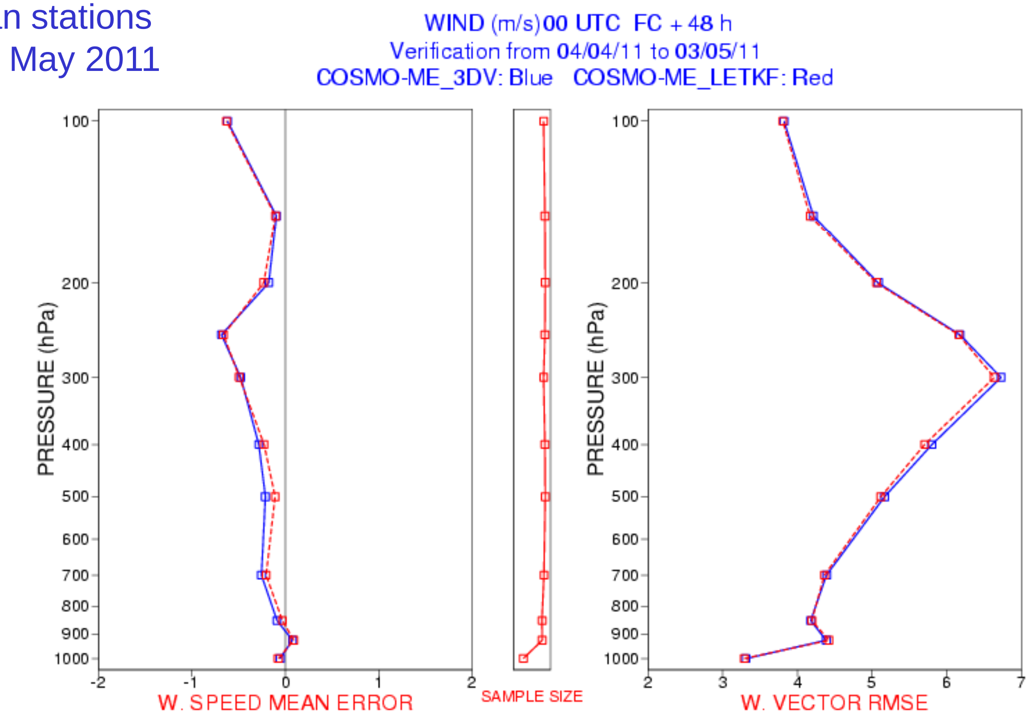


comparison of forecasts from interpolated
LETKF (10 km) vs. 3DVAR (14 km)
analyses

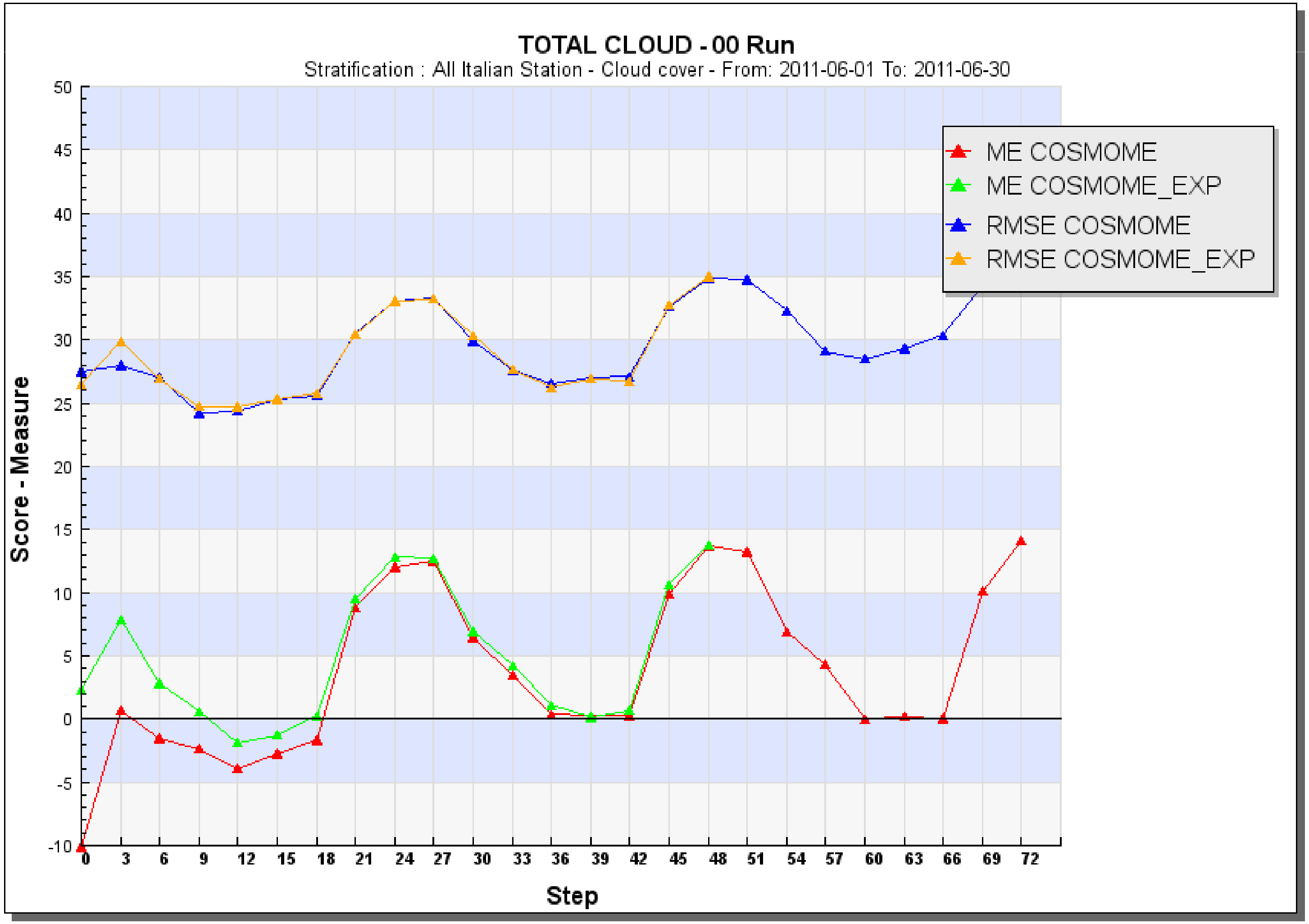
can we improve on our operational 3DVar without the complications of 4DVar ?
→ yes !

European stations
April and May 2011



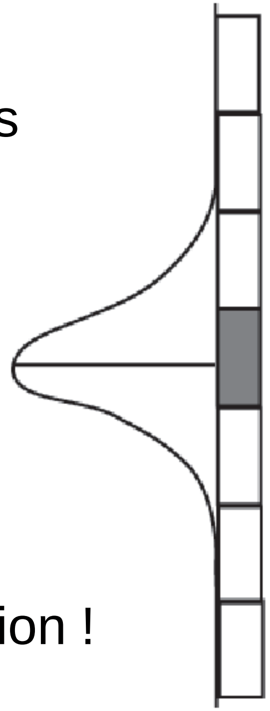


comparison of forecasts starting from
deterministic analysis (control) vs. ensemble mean analysis
→ spin-up



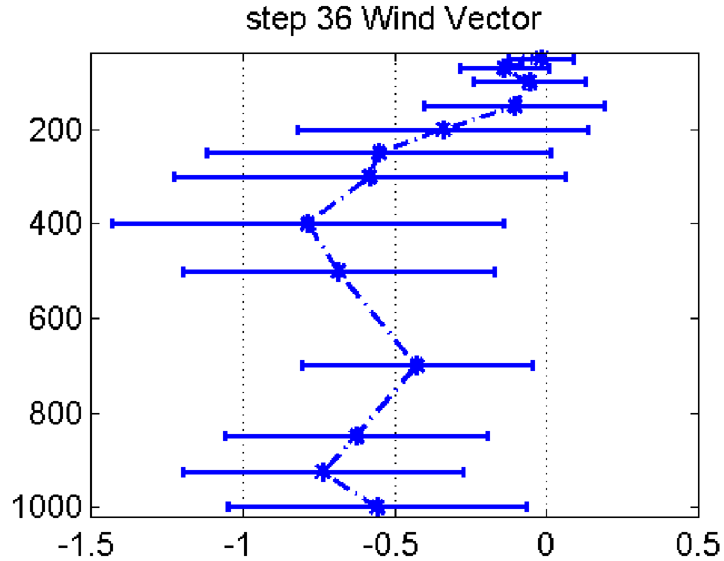
CNMCA LETKF :
impact of AMSU-A radiances (EURO-HRM, 28 km)

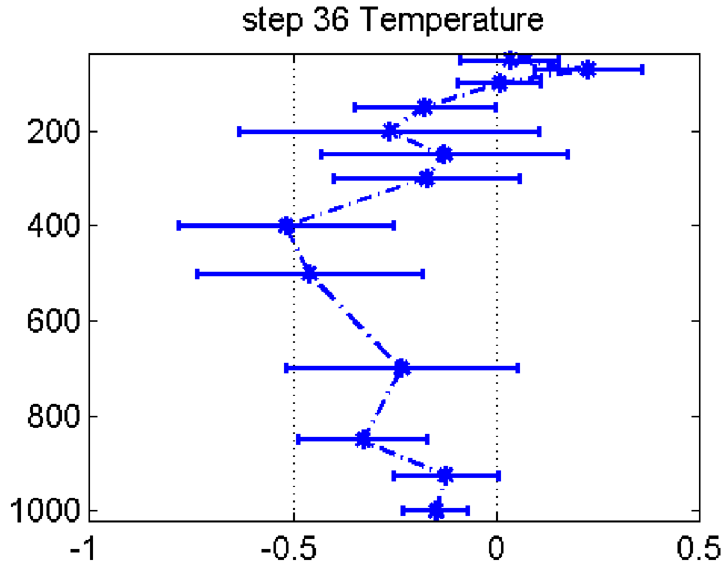
- AMSU-A treated as "single-level" obs
- assign radiance obs model level for which magnitude of weighting function (*wf*) is largest (maximum-based selection, Fertig et al., 2007
- Use *wf* shape as vertical covariance localization function !



(a) Maximum-Based Selection

relative difference (%) in RMSE computed against IFS analysis





00 UTC
EuroHRM runs
(28km)
11-10 to 10-11-2009

CNMCA LETKF

- as far as we know, CNMCA is the first meteorological centre which uses operationally a pure ensemble DA (LETKF) to initialize a deterministic NWP model (COSMO-ME)
- deterministic forecast started from ‘det. analysis’ (not ensemble mean ana.)
- AMSU-A improves LETKF analysis (28 km, large domain !)

outlook

- AMSU-B/MHS and IASI retrievals: investigate soon
- balancing and non-linearities to be addressed (outer loop / RIP not beneficial)
- further tuning of model error representation
  (tuning of cov. localization, evolved additive noise, bias correction, etc.)
- tests with COSMO
- implement short-range EPS based on LETKF

thank you for your attention

LETKF scientific issues / refinements: model error estimation

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

- **objective estimation and modelling of model (tendency) errors**
  (not to be confused with forecast errors)
  M. Tsyrulnikov, V. Gorin (HMC Russia)

1. set up a (revised) stochastic model (parameterisation) for model error (ME)

   MEM : $\underline{e = \mu * F_{phys}(\mathbf{x}) + e_{add}}$

   – involves stochastic physics ($\mu * F_{phys}(\mathbf{x})$) and additive components e_{add} and includes multi-variate and spatio-temporal aspects
   – model error model parameters: (E(μ), D(μ), D(e_{add}))

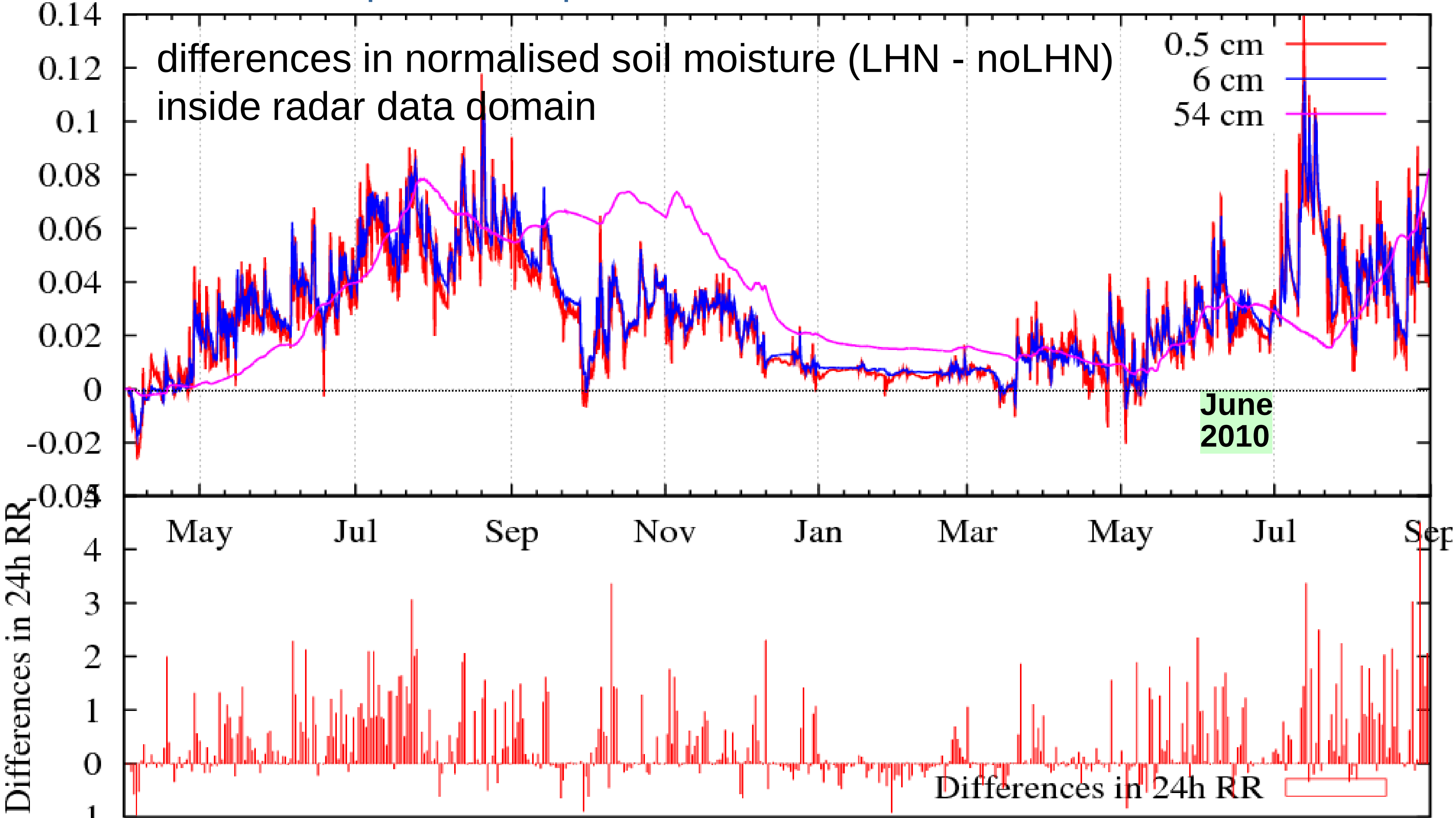
2. develop MEM **Estimator** : estimate parameters by fitting to statistics from real forecast (COSMO-RU-14) and observation (radiosonde) tendency data, using a revised (maximum likelihood based) method (tested using bootstrap)

3. develop ME **Generator** embedded in COSMO code

4. develop MEM **Validator**: exactly known ME in OSSE set-up, retrieve with MEM Estimator

$\rightarrow$ The reproducibility of the MEM parameters is not yet established !

Experiment: 18 months (April 2009 – August 2010) COSMO-DE without LHN
→ compare with operational COSMO-DE with LHN



impact of LHN on soil moisture:
impact on surface parameters

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

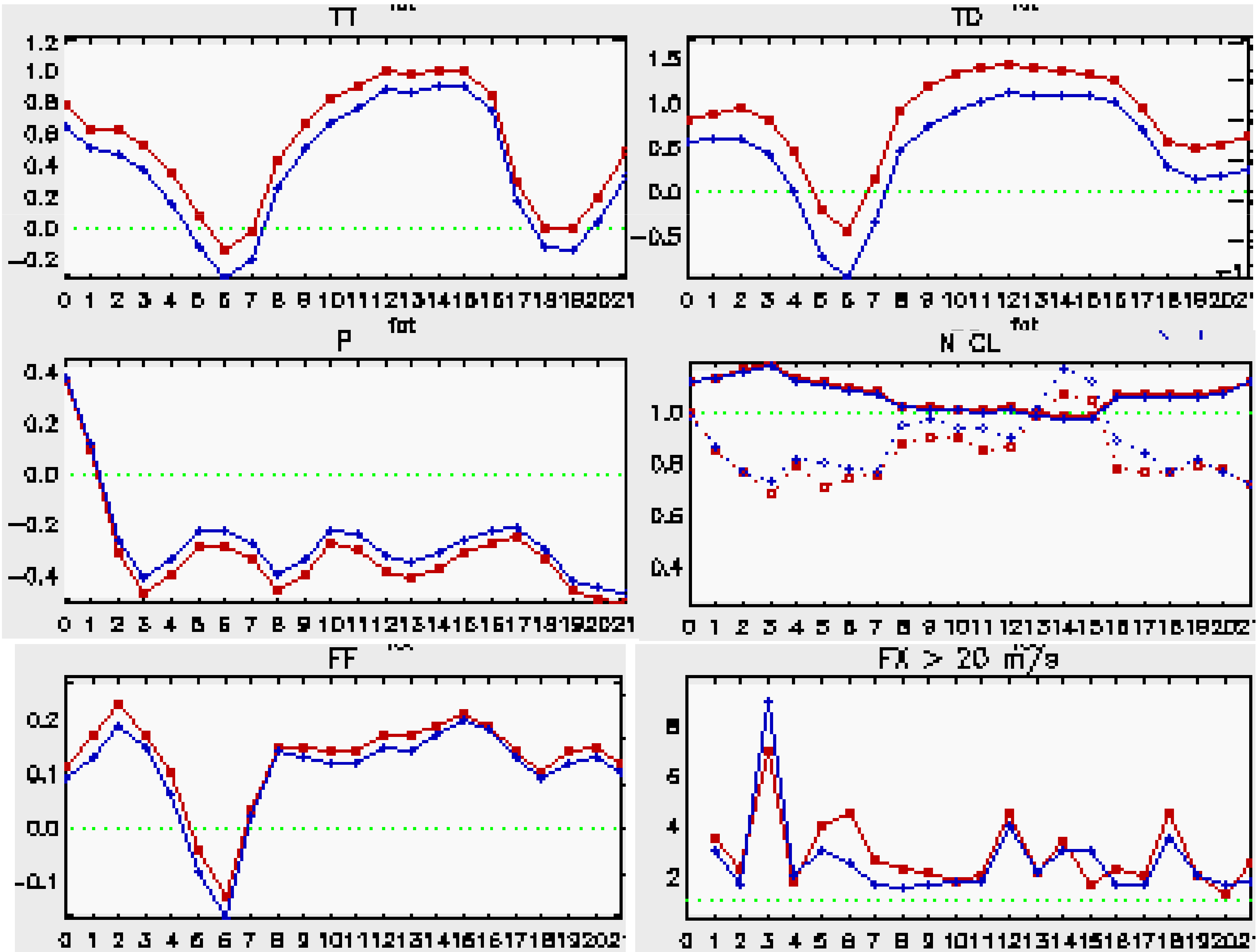
June 2010,

12 UTC runs

BIAS

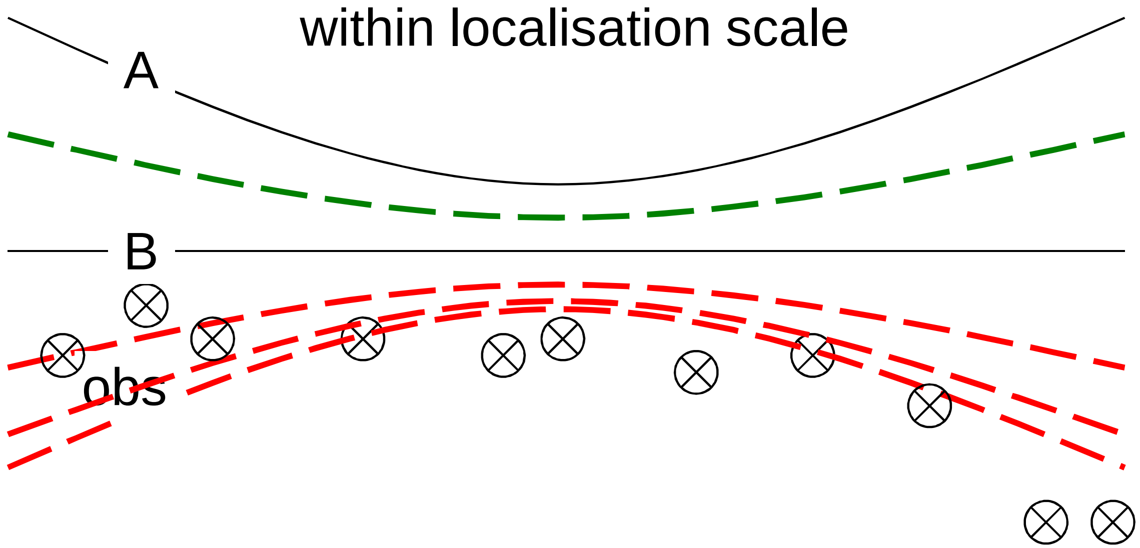
LHN
noLHN

bias reduced for different parameters



- adaptive estimation of obs error covariance $\mathbf{R}$

  (Li, Kalnay, Miyoshi, QJRMS 2009) , but our implementation: in ensemble space !



f.g. mean = 0.5 A + 0.5 B

ana = f.g.mean + (B – A)
= 1.5 B – 0.5 A
(1 perfect obs)

2 obs → least square fit

add obs, if already $N_{obs} > N_{ens}$:

- cannot be fitted well, improve analysis only slightly
- decrease analysis error !

$$\mathbf{P}_w^{\ a} = \left[(k-1)\,\mathbf{I} + \left(\mathbf{Y}^b\right)^T \mathbf{R}^{-1} \mathbf{Y}^b \right]^{-1}$$

→ adaptive $\mathbf{R}$ takes rank deficiency into account and increases $\mathbf{R}$