Meso-scale data assimilation in AROME

Pierre Brousseau,
Y. Seity, T. Montmerle, E. Wattrelot
METEO-FRANCE/CNRM

8 th Workshop on Adjoint Model Application in Dynamic Meteorology

Tannersvilles, PA

Outlines

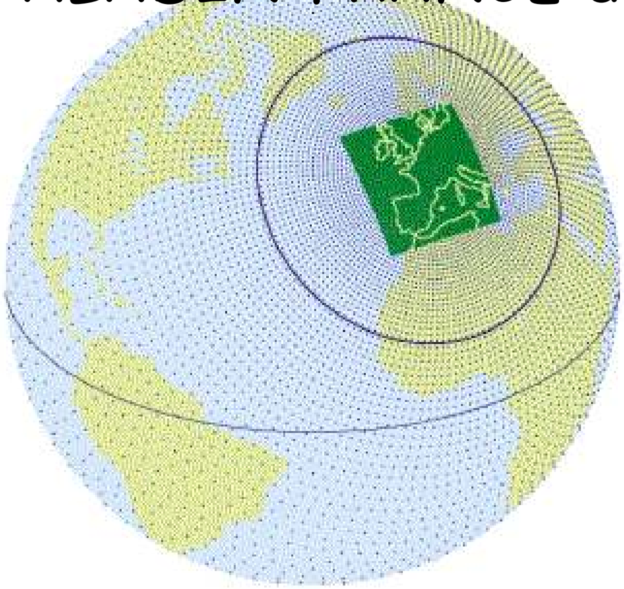
- General ideas on AROME
- The cycle strategy
- Background-error Statistics
- Observations
- Assimilation results

The AROME system

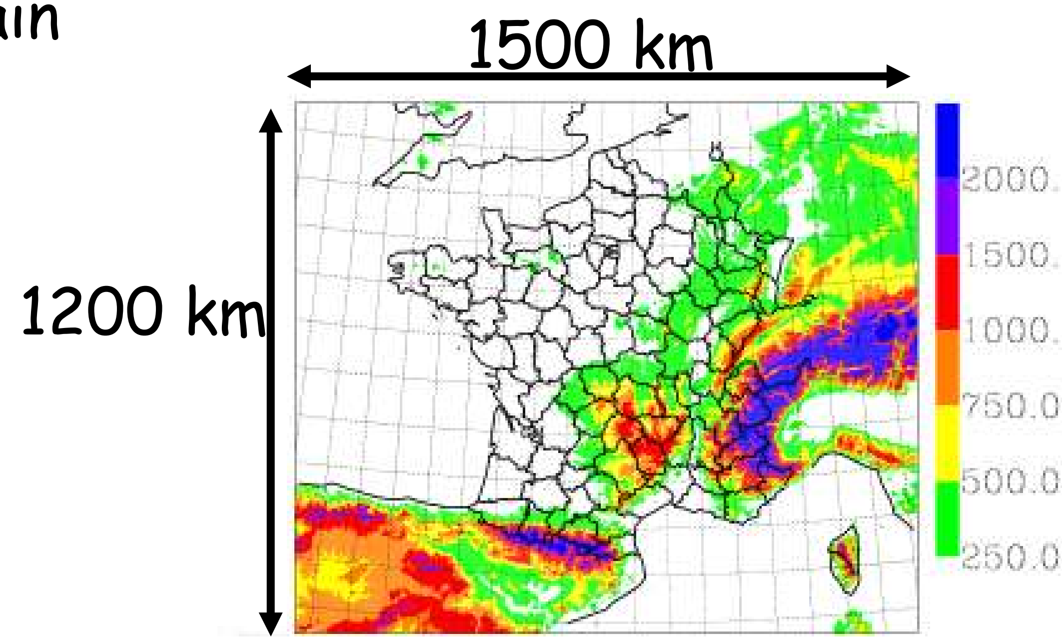
- AROME model has completed the French NWP system since the end of 2008 :
  - ARPEGE : global model (15 km over Europe)
  - ALADIN-France : regional model (10km)
  - AROME : meso scale model (2.5km)

- Aim : to improve local meteorological forecasts of potentially dangerous convective events (storms, unexpected floods, wind bursts...) and lower tropospheric phenomena (wind, temperature, turbulence, visibility...).

ARPEGE stretched grid and ALADIN-FRANCE domain



AROME France domain

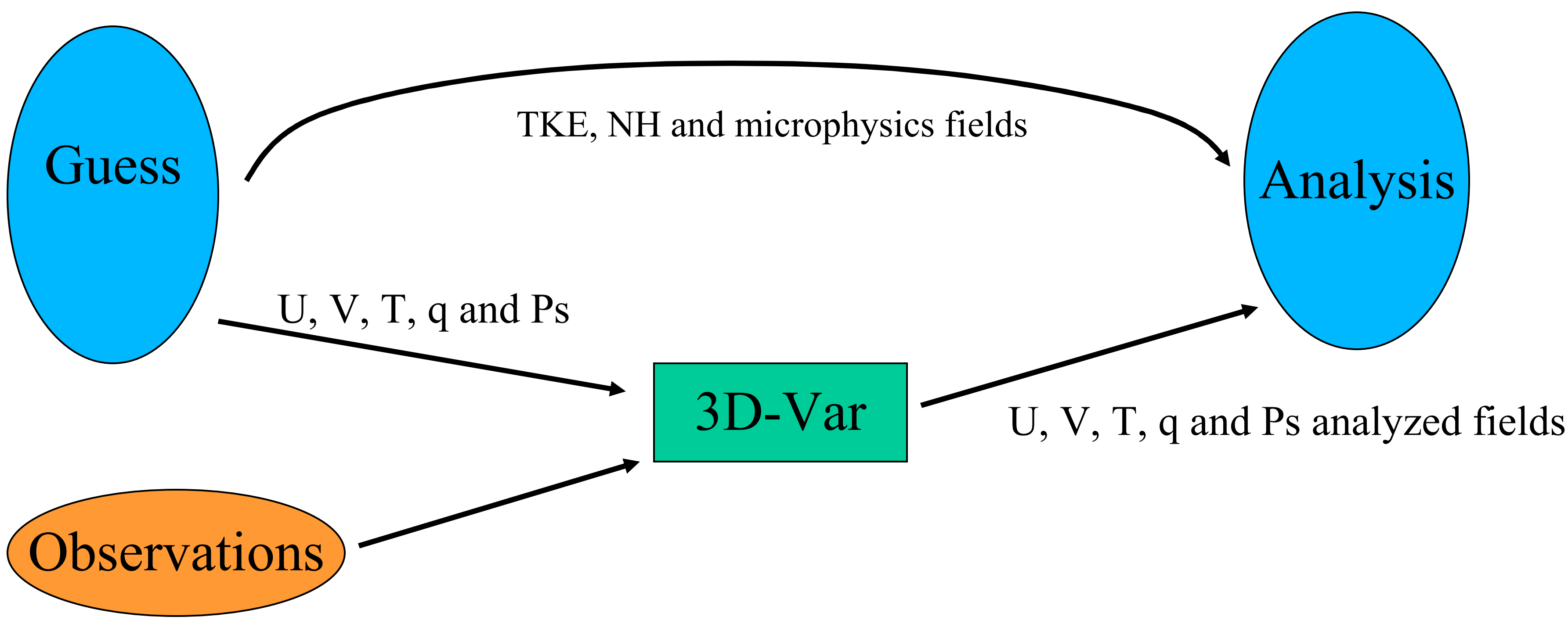


The AROME system

- Model merges research outcomes and operational progress :
  - physical package from the Meso-NH research model
  - Non-Hydrostatic version of the ALADIN software
- A complete data assimilation system derived from the ALADIN-FRANCE 3D-Var scheme (Fisher et al. 2005) operationally running at Météo-France at regional scale since the end of 2005
- Developed in the frame of ARPEGE/IFS software (Courtier et al. 1991), it inherits lots of its characteristics :
  - incremental formulation
  - observation operators
  - minimization technique
  - data flow
- Efficient also at meso-scale, after some adaptations

Assimilation scheme

- Control variable : vorticity, divergence, temperature, specific humidity and surface pressure :
  - 2 wind components, temperature, specific humidity and surface pressure are analysed at the model resolution (2.5 km).
  - Other model fields (TKE, Non-hydrostatic and microphysics fields) are cycled from the previous AROME forecast used as background



Outlines

- General ideas on AROME
- The cycle strategy
- Background-error Statistics
- Observations
- Assimilation results

Cycle strategy

- Idea :
  - Forecasts initialized with more recent observations will be more accurate
  - Using high temporal and spatial frequency observations (RADAR measurements for example) to the best possible advantage

  Use of rapid forward intermittent assimilation cycle in order to compensate the lack of temporal dimension in the 3D-Var (RUC/NCEP)

- The cycle strategy has to be investigated in order to :
  - Choose the cycle frequency allowing best performances
  - Prevent a potential drifting of such a system over a small domain due to the ignorance of a large scale analysis

Cycle strategy : frequency

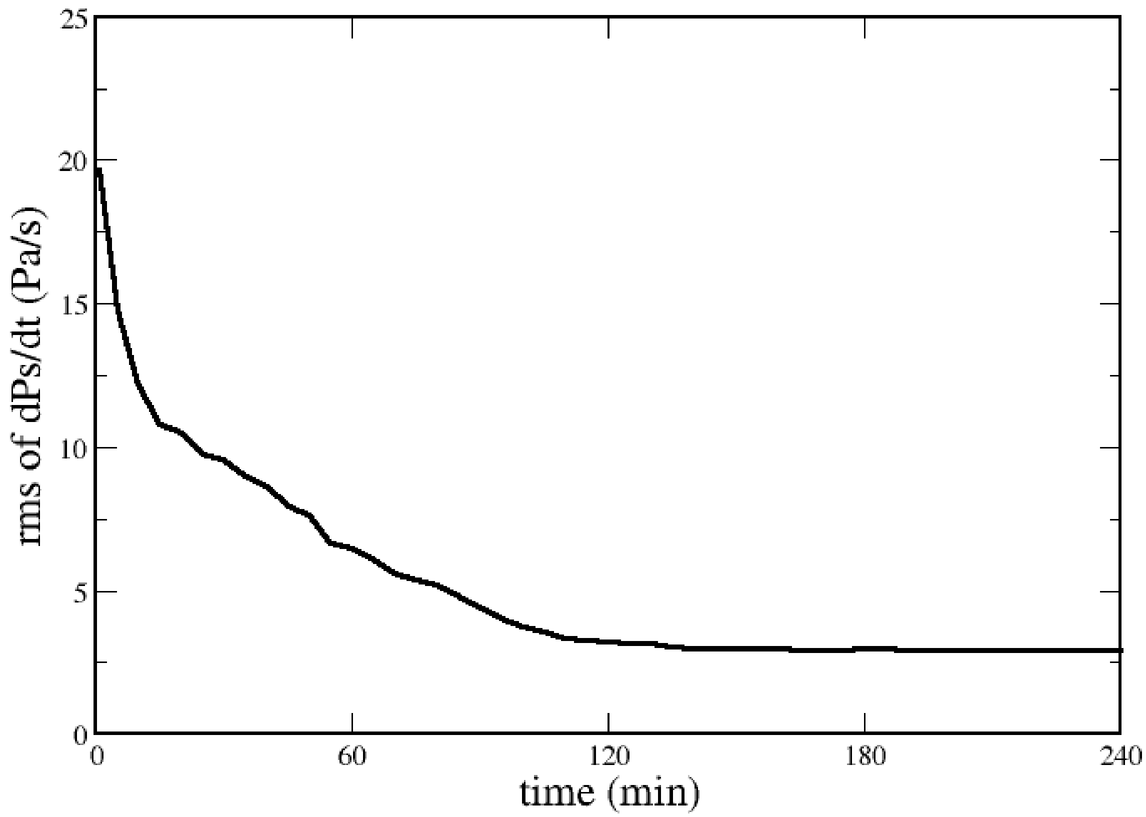
- Risk of accumulating noises and imbalances through cycle decreasing system performances.

- non meteorological values in the 2-first-hour forecast range : spurious waves are present in the model solution

- This noise is substantially reduced at the 3-hour output time

- Need to filter those spurious waves (Incremental Digital Filter Initialisation) for cycle frequency lower than 3. Not yet implemented

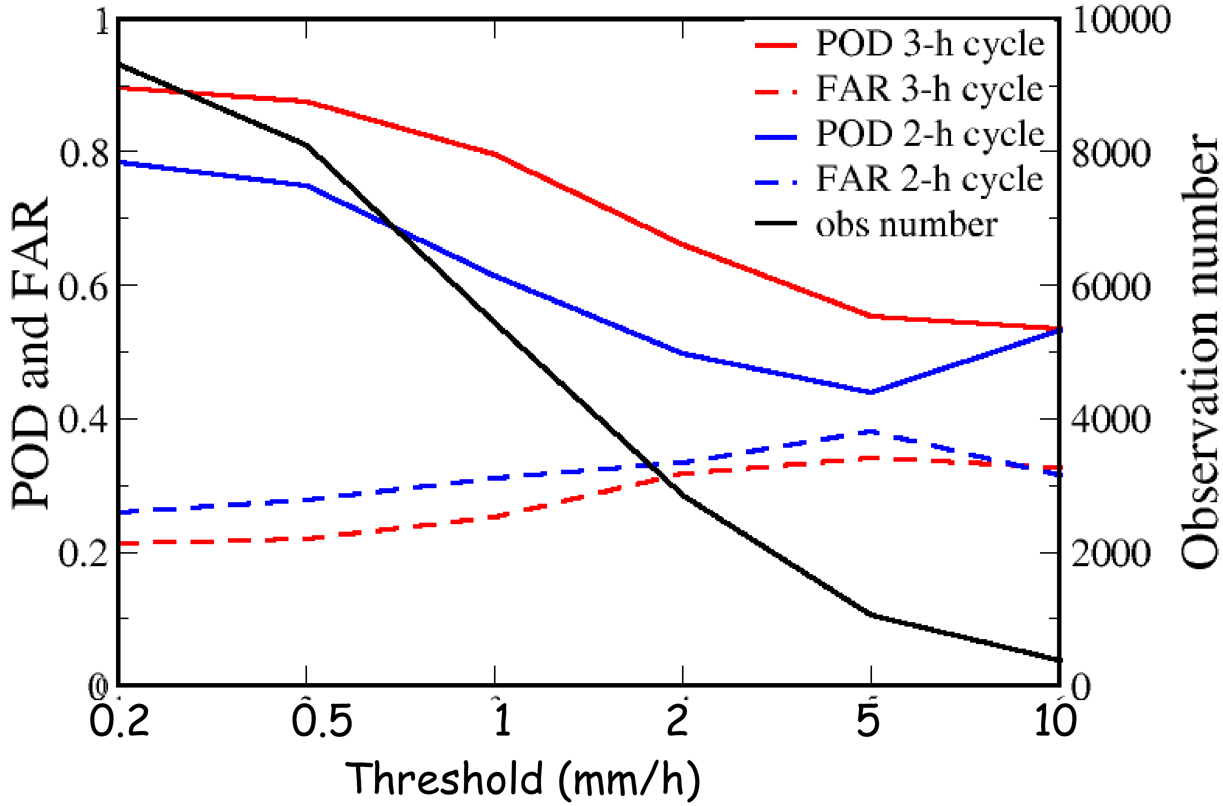
Temporal evolution of RMS of surface pressure tendency over a single time step (60s). Data point taken every 5 min of integration



Cycle strategy : frequency

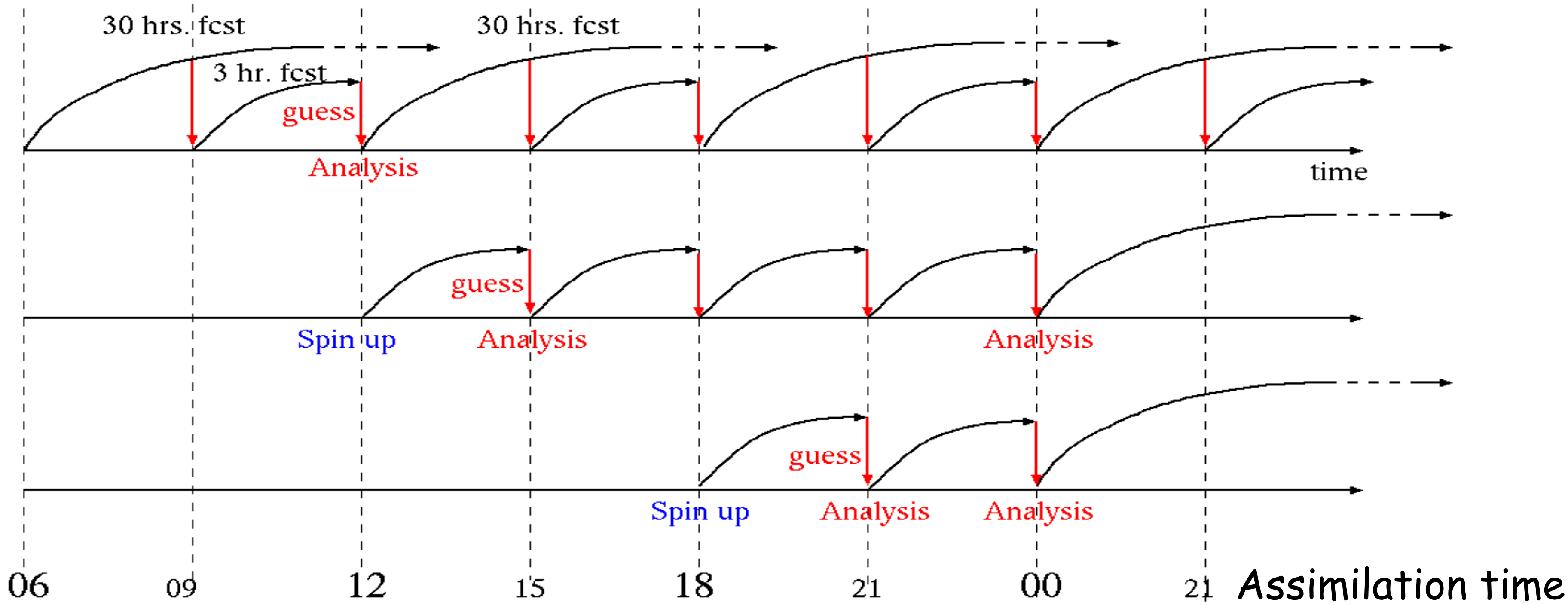
- Experiments with 1, 2 and 3-h frequency continuous cycle during a 30-day-long period
  - 1-h cycle : forecast crashed after 2 days
  - 2-h cycle : poorer performance than 3-h cycle

Quantitative Precipitation Forecats scores for different thresholds for the total rain forecast between 0- and 12-h compared to rain-gauge measurements in November 2008



Cycle strategy : drifting risk

- Problem: a continuous cycle never sees large scale analysis except through lateral boundary conditions.
- 2 experiments restarting from ALADIN analysis (spin-up mode) every day



- Objective scores : no significant differences
- QPF scores : better performances for the continuous assimilation cycle

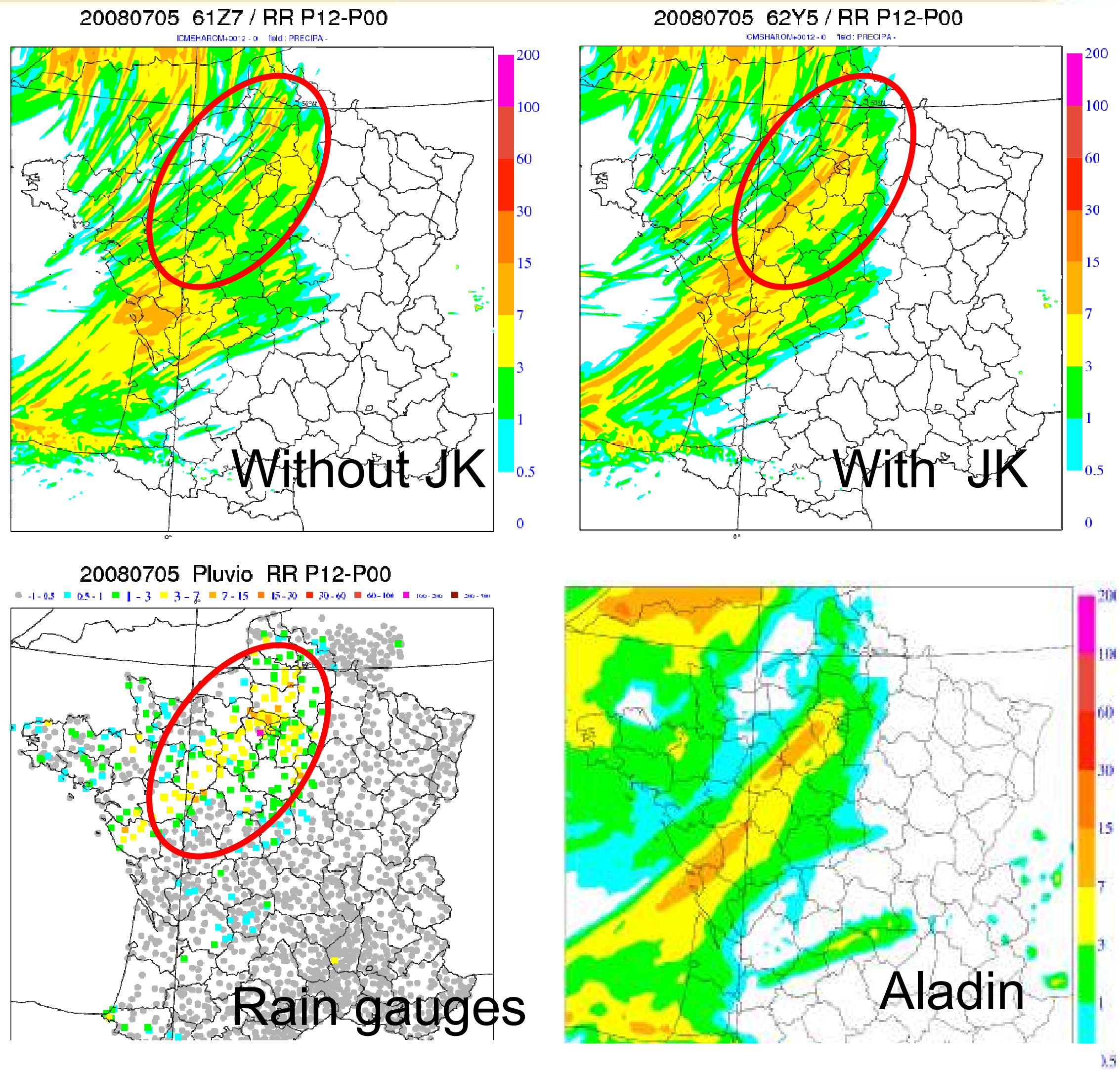
Cycle strategy : Relaxation towards larger scales

- Var (J_k) (guidard 2008) :

$$J(\delta \mathbf{x}) = Jb(\delta \mathbf{x}) + Jo(\delta \mathbf{x}) + Jk(\delta \mathbf{x})$$

25-day experiment, relaxing towards large scale (>100 km) of ALADIN analyses above 250hPa:

- neutral scores against conventional data
- small improvements in QPF scores for small precipitating amounts
- 1 case with significant improvement

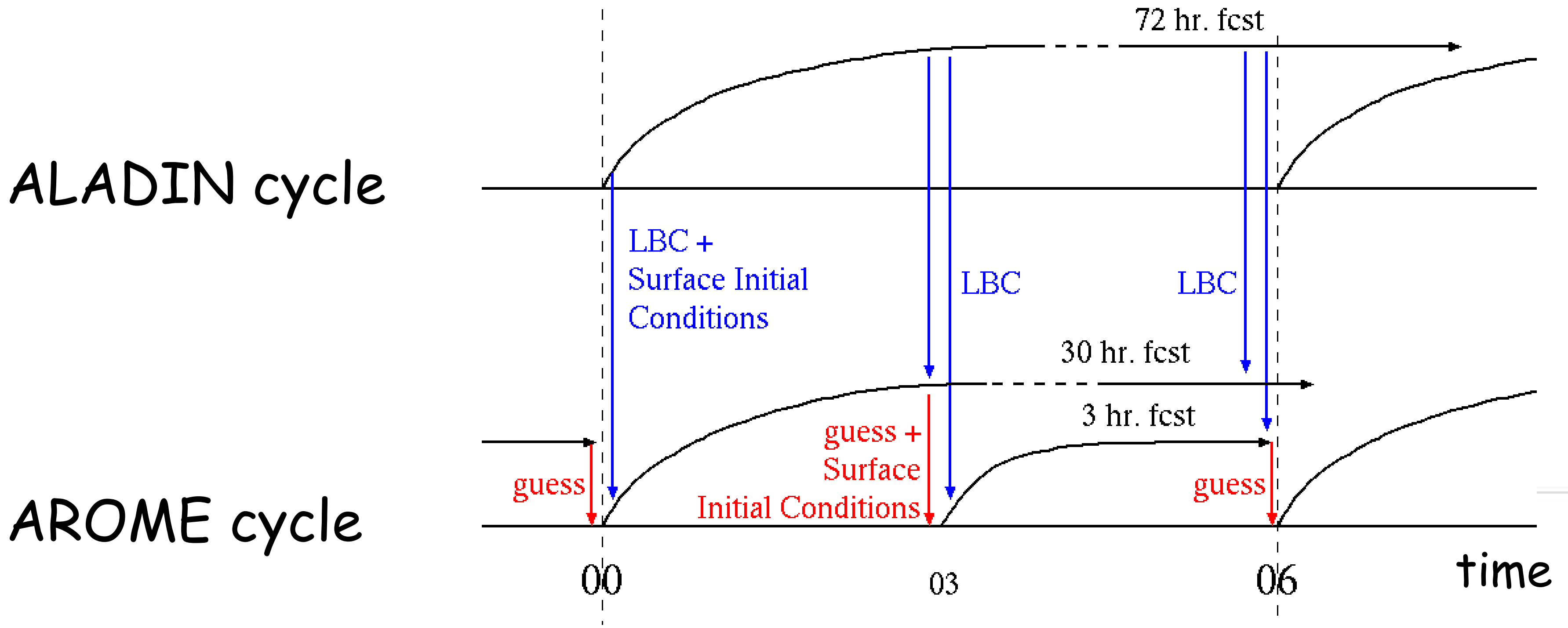


00-24 cumulative rainfalls on 5 july 2008

Operational configuration

- AROME operational configuration uses a 3-h frequency continuous assimilation cycle and performs 30-hr forecasts at synoptic time (00, 06, 12 and 18 UTC).

- the ALADIN-FRANCE operational suite provides :
  - Lateral boundary conditions
  - Surface initial conditions : OI analysis (CANARI) at 00, 06, 12 and 18 UTC (the previous AROME forecast is used otherwise).

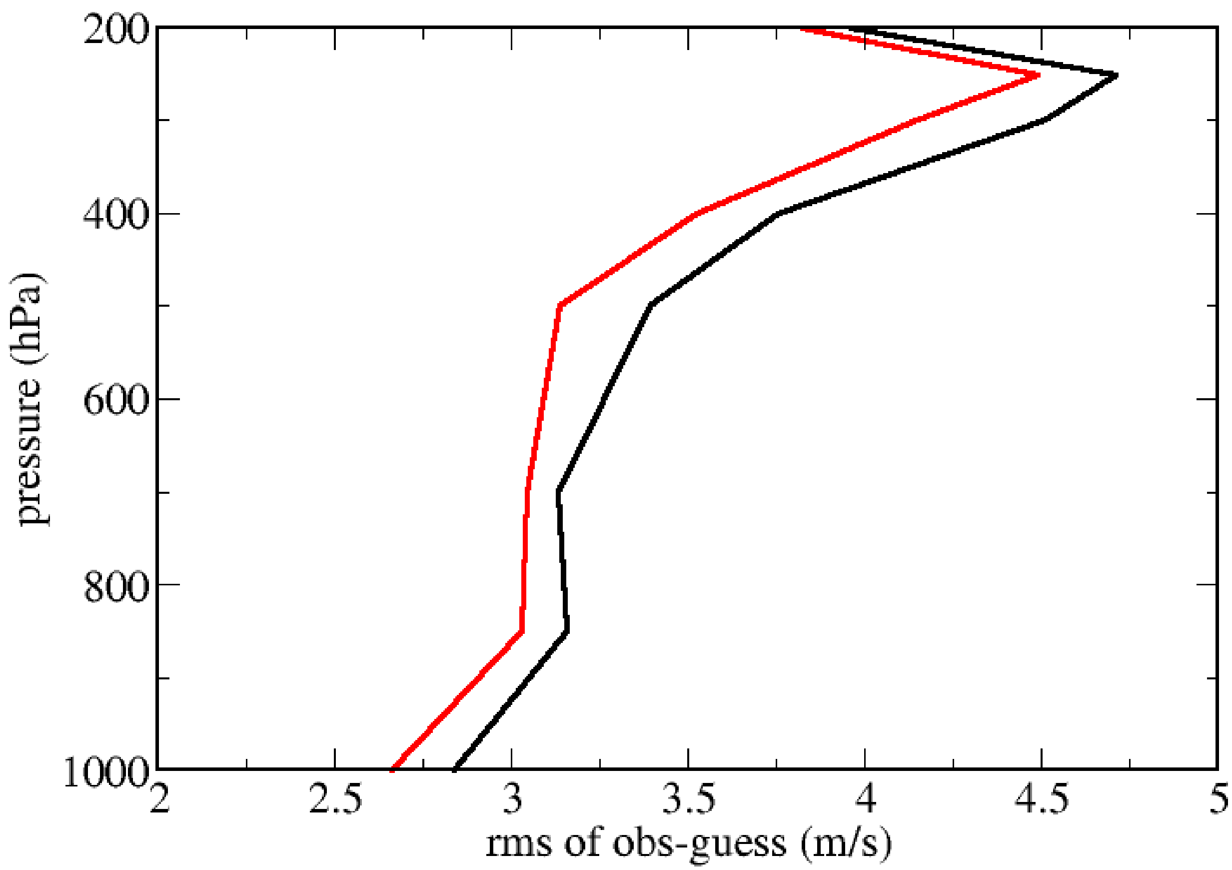


Cycle strategy : future plans

- Implementation of an Incremental Digital Filter Initialisation in AROME framework in order to use 1-h frequency cycle.
- 3D-FGAT version of the assimilation scheme : first encouraging results

Vertical profile of RMS of observation-guess for wind data from air plane measurements using a 3-h continuous cycle with :

3D-Var
3D-FGAT



Outlines

- General ideas on AROME
- The cycle strategy
- Background-error Statistics
- Observations
- Assimilation results

Background-error statistics : documentation

- Background-error statistics for AROME share the same multivariate formulation as in ALADIN-FRANCE (Berre 2000). This formalism uses errors of vorticity, divergence, temperature, surface pressure and humidity, with scale-dependant statistical regressions to represent cross-covariances.

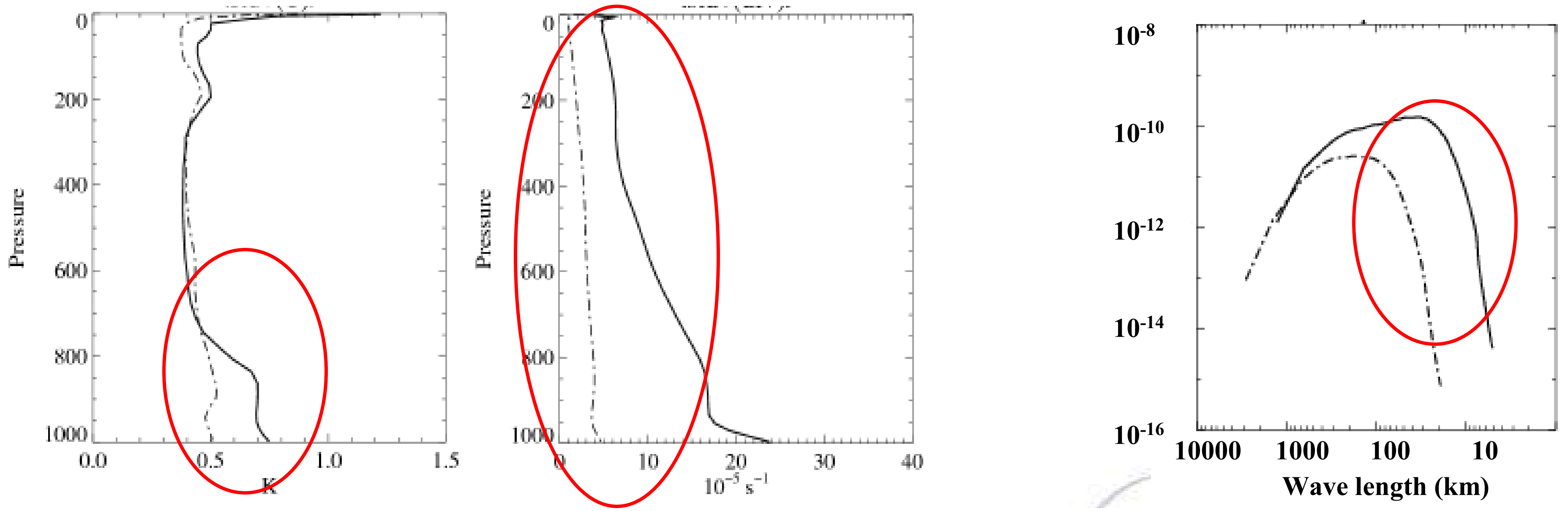
- calculated using an ensemble-based method (Berre et al. 2006), with a six-member ensemble of AROME forecasts in spin-up mode carried out over two 15-day periods :
  - Anticyclonic winter
  - Convective summer

- Initial and lateral conditions are provided by an ARPEGE/ALADIN-FRANCE assimilation ensemble

Background error statistics : AROME VS ALADIN (1)

- Greater σ_b for AROME than for ALADIN : The background is less to be trusted, mostly in the boundary layers, for small scales and for variables that are representative for these small scales
- Consistent with the explicit representation of small scale structures in AROME, which are either unrepresented or numerically dissipated by ALADIN

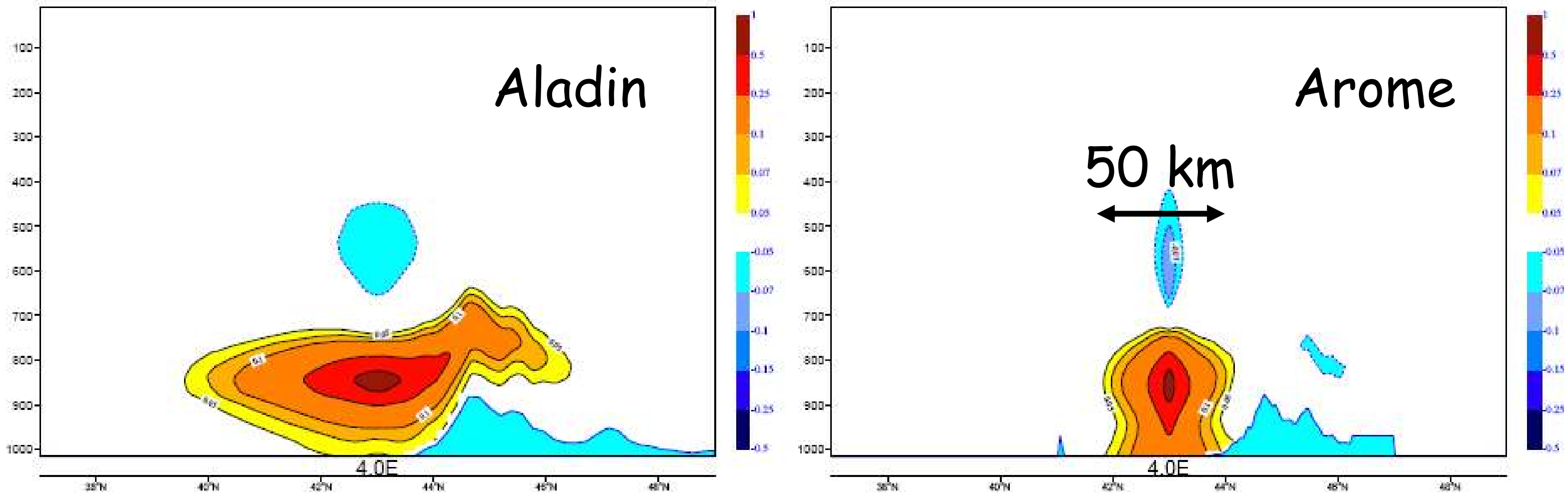
Spectral average of σ_b for T | Spectral average of σ_b for div | Variance spectra for div at 500 hPa



Arome -·-·· Aladin

Background error statistics : AROME VS ALADIN (2)
Horizontal correlation lengthscales

- Shorter correlation lengths for AROME than for ALADIN, which is coherent with the smaller domain and smaller horizontal resolution.



One obs experiment: 2K temperature innovation at 850 hPa.
Vertical section of the analysis increment

- The assimilation of one observation leads to a more localized increment.
- Dense observation networks (ground measurements, geostationary satellites, GPS, radars…) can be used with a higher horizontal resolution (by paying attention to correlations between observation errors).

Background error statistics : winter/summer

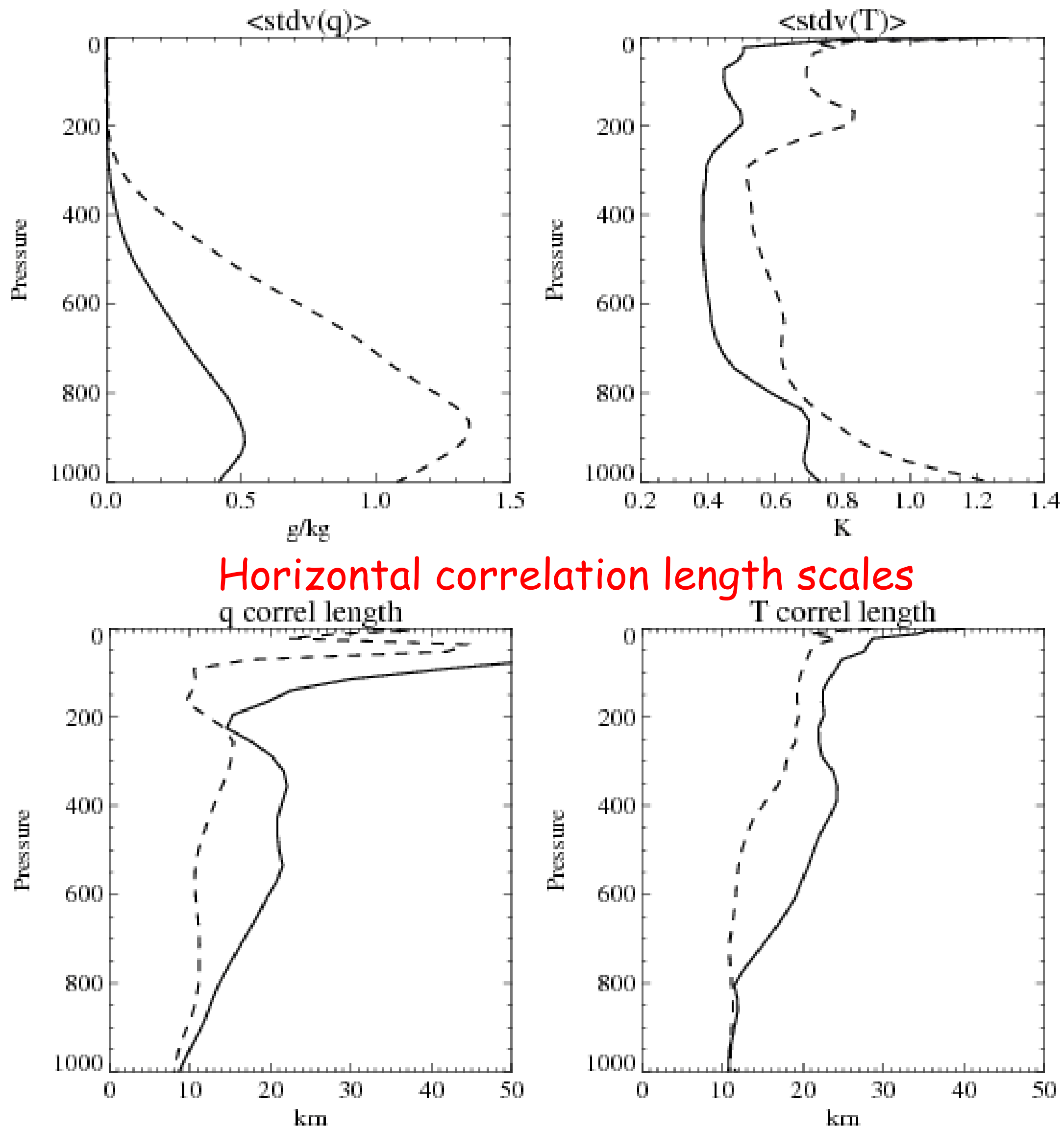
- Background error statistics strongly depend on the meteorological situation . In summer :
  - Higher standard deviation
  - Shorter correlation lenghtscales
- Coherent with convective phenomena
- limitation of a "climatological" B matrix : use of flow dependant statistics (Loïk Berre's talk in session 9)

  ____ winter

  ---- summer

Background error standard deviation

Horizontal correlation length scales



Background error statistics : heterogeneous B matrix

Thibault Montmerle (Following an idea by Philippe Courtier (1998), and used by Mark Buenher 2008).

- To use more suitable background error statistics in clear air and precipitating areas :

$\mathbf{B} = \alpha \mathbf{B}_r + \beta \mathbf{B}_{nr}$ With: $\alpha = \mathbf{FMF}^{-1}$ and $\beta = \mathbf{F(1\text{-}M)F}^{-1}$

$\mathbf{M}$: grid point mask deduced from observed radar reflectivity.

$\mathbf{B}_{r}$ and $\mathbf{B}_{nr}$ are separately computed by performing statistics on an assimilation ensemble of precipitating cases, considering a mask based on simulated precipitations.

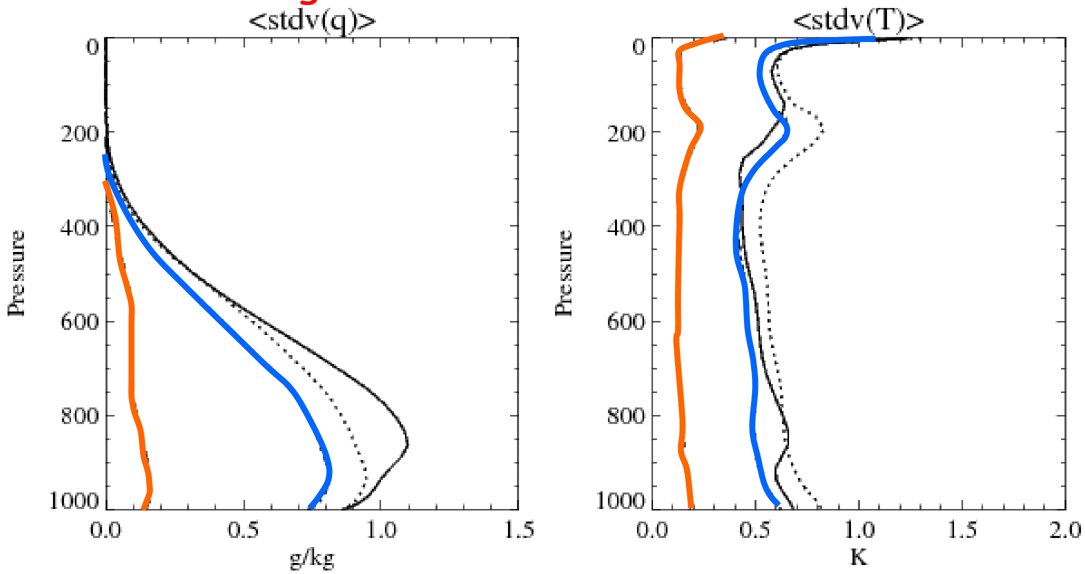
- The increment is written:

$$\delta x = \mathbf{B}^{1/2}\chi = \left(\alpha^{1/2}\mathbf{B}_r^{1/2} + \beta^{1/2}\mathbf{B}_{nr}^{1/2}\right)\begin{pmatrix}\chi_1 \\ \chi_2\end{pmatrix}$$

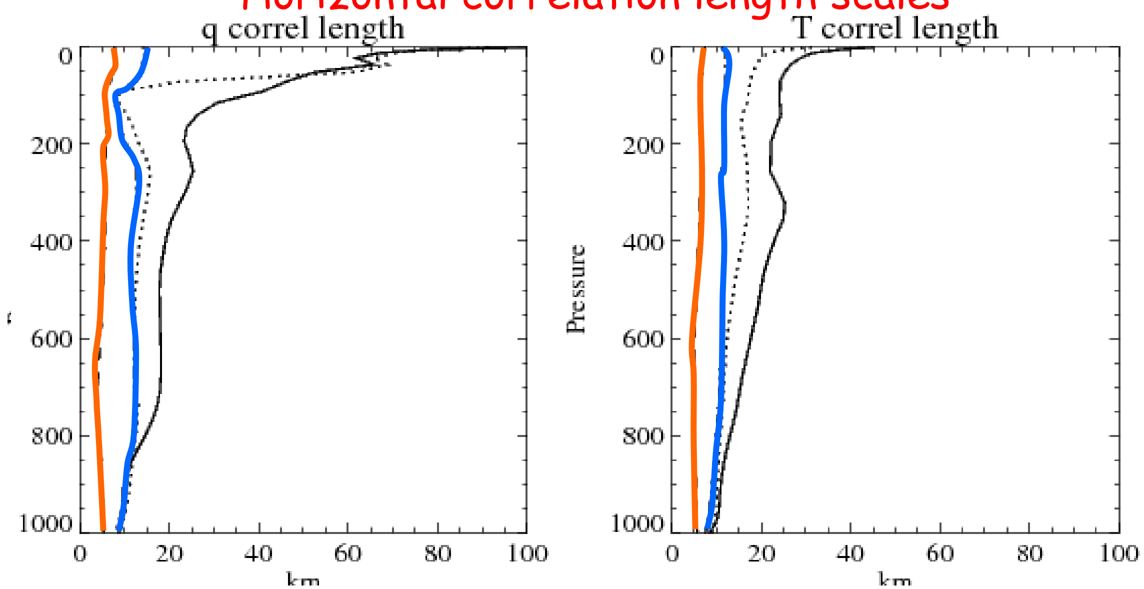
$\Rightarrow$ Which implies doubling the control variable χ and the gradient $\nabla_\chi J$

Background error statistics : heterogeneous B matrix

Background error standard deviation



Horizontal correlation length scales



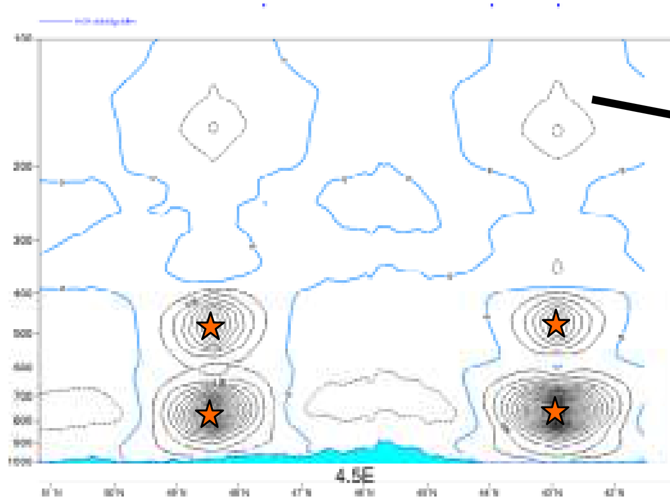
- In precipitating areas :
  - Smaller σ_b for q and T
  - Smaller horizontal correlations
- Precipitating observations can be used with a greater density

Bnr

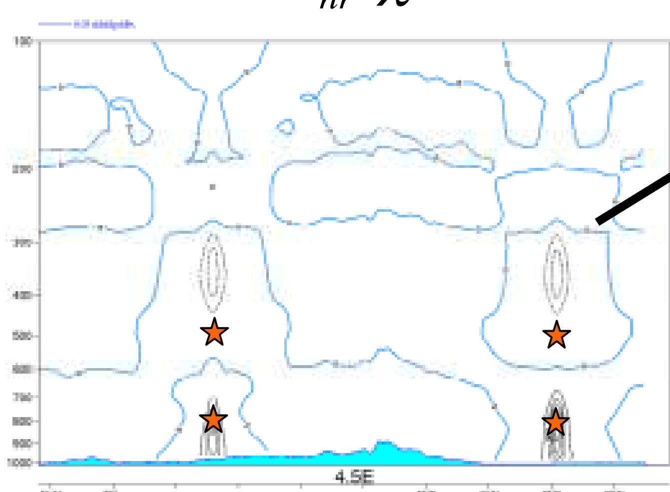
Br

Background error statistics : heterogeneous B matrix (3)

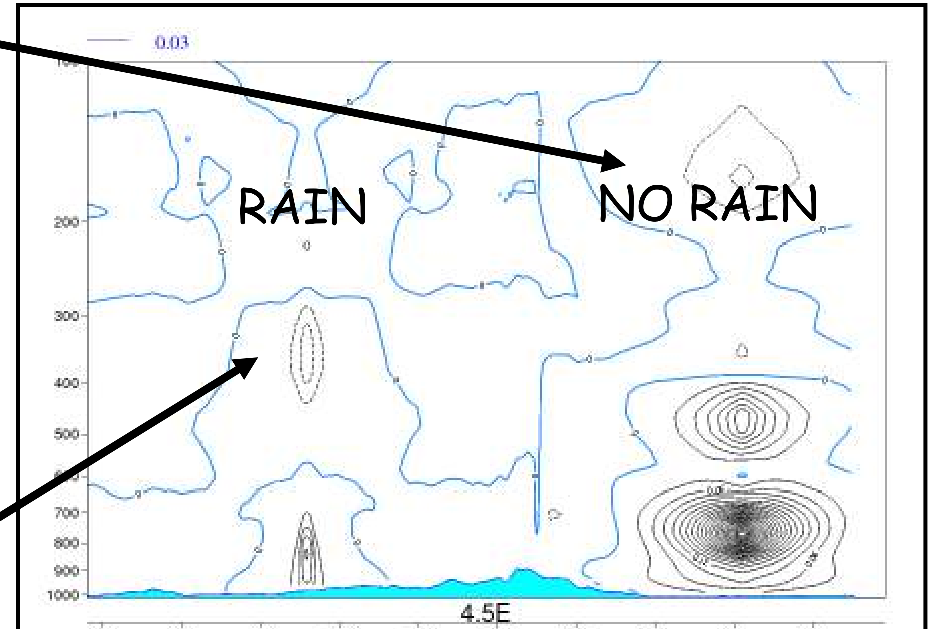
Two observations experiments : vertical section of humidity increment



$$\delta x = \mathbf{B}_{nr}^{1/2}\chi$$



$$\delta x = \mathbf{B}_{r}^{1/2}\chi$$



$$\delta x = \mathbf{B}^{1/2}\chi = \left(\alpha^{1/2}\mathbf{B}_{r}^{1/2} + \beta^{1/2}\mathbf{B}_{nr}^{1/2}\right)\begin{pmatrix}\chi_1 \\ \chi_2\end{pmatrix}$$

Innovations of - 30% RH
At 800 and 500 hPa

Outlines

- General ideas on AROME
- The cycle strategy
- Background-error Statistics
- Observations
- Assimilation results

Observations

- Same observations such as in ALADIN-France operational suite : conventional observations, CLS observations (T2m, HU2m, V10m), IR radiances from ATOVS and SEVIRI instruments, winds from AMV and scatterometers, ground based GPS.

- No specific spatial selection (thinning) appropriate to AROME resolution except for ground GPS. Studies on this topic are still ongoing (plane measurements, IR radiances…)

- Operational use of Doppler RADAR wind data in AROME since December 2008 (Montmerle and Faccani, 2009)

Observations

- Same observations such as in ALADIN-France operational suite + Doppler RADAR wind data
- The number of observation depends on the assimilation time
- SYNOP, RADAR doppler winds, plane measurements and SEVIRI radiances are of great interest to supply the data assimilation system.

Observation kind	Number per analysis
SYNOP+SHIP/BUOY	1500/8
RADAR doppler winds	0<Nobs<1000
Ground GPS	170
Radiosondes	0<Nobs<20
Plane	0<Nobs<1600
Cloud motion winds	15
Scatterometers winds	0<Nobs<80
ATOVS	0<Nobs<55
SEVIRI	150<Nobs<300

Plans : RADAR reflectivities assimilation

(Wattrelot & al., 2008)

- Reflectivity observation operator needs a complete description of warm and cold hydrometeors : realistic simulation can be obtained with AROME.

- Reflectivities can provide useful information about the atmosphere water cycle (rain, snow, graupel, primary ice), but in the context of variational assimilation, assimilation of rain is very difficult because:
  - The direct observation operator involves physical processes which are characterized by discontinuities and nonlinearities, and there is need of simplification in the linearized versions to get some good results…
  - "rain" is not a variable which is in the "control variable" of the analysis

- Rainfalls have a short shelf life in the atmosphere. Therefore, it's better to try to modify only the humidity field : need of a 1D method to get some relative humidity retrievals from reflectivities before using the 3D-Var scheme

Radar reflectivities assimilation: inversion method

The « best » estimate of atmosphere x given the observation y_0 and using Bayses's theorem (Lorenc, 1986)

$$E(x) = \iiint x.P(x = x_{true} \mid y = y_0)dx$$

Bayses's

$$E(x) = \iiint x.P(y = y_0 \mid x = x_{true}).P(x = x_{true})dx$$

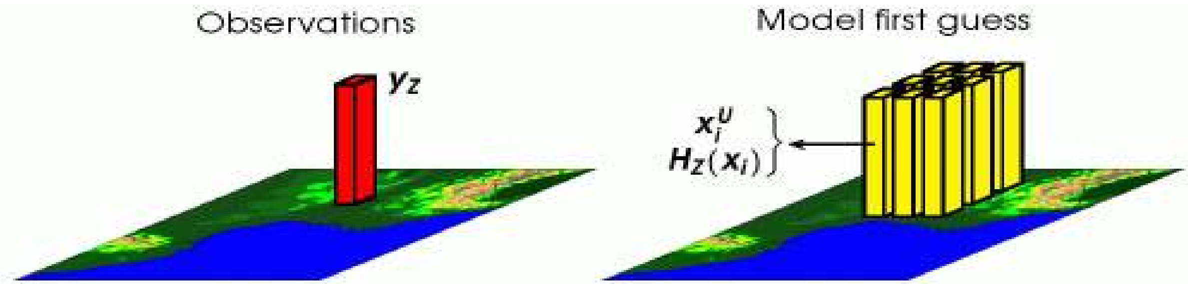
Olson, 1996 (Gaussian and uncorrelated errors) and x_j database of atmospheric profiles

$$E(x) = \frac{\sum_j x_j \exp\frac{-1}{2}\cdot\|y_0 - y_s(x_j)\|^2}{\sum_j \exp\frac{-1}{2}\cdot\|y_0 - y_s(x_j)\|^2}$$

with $\|y_0 - y_s(x_j)\|^2 = [y_0 - y_s(x_j)]^T \cdot (O + S)^{-1} \cdot [y_0 - y_s(x_j)]$

Radar data assimilation : Inversion method of reflectivities profiles

Caumont, 2006: use of model profiles in the vicinity of the observation as representative database



$$y_{po}^U = \sum_{i \in \text{neighbours}} x_i^U \frac{\exp\left(-\frac{1}{2}\|y_Z - H_Z(x_i)\|^2\right)}{\sum_{j \in \text{neighbours}} \exp\left(-\frac{1}{2}\|y_Z - H_Z(x_j)\|^2\right)}$$

y_{po}^U: column of pseudo-observed relative humidity,
y_Z: column of observed reflectivities,
x_i^U: column of relative humidity,
$H_Z(x_i)$: column of simulated reflectivities.

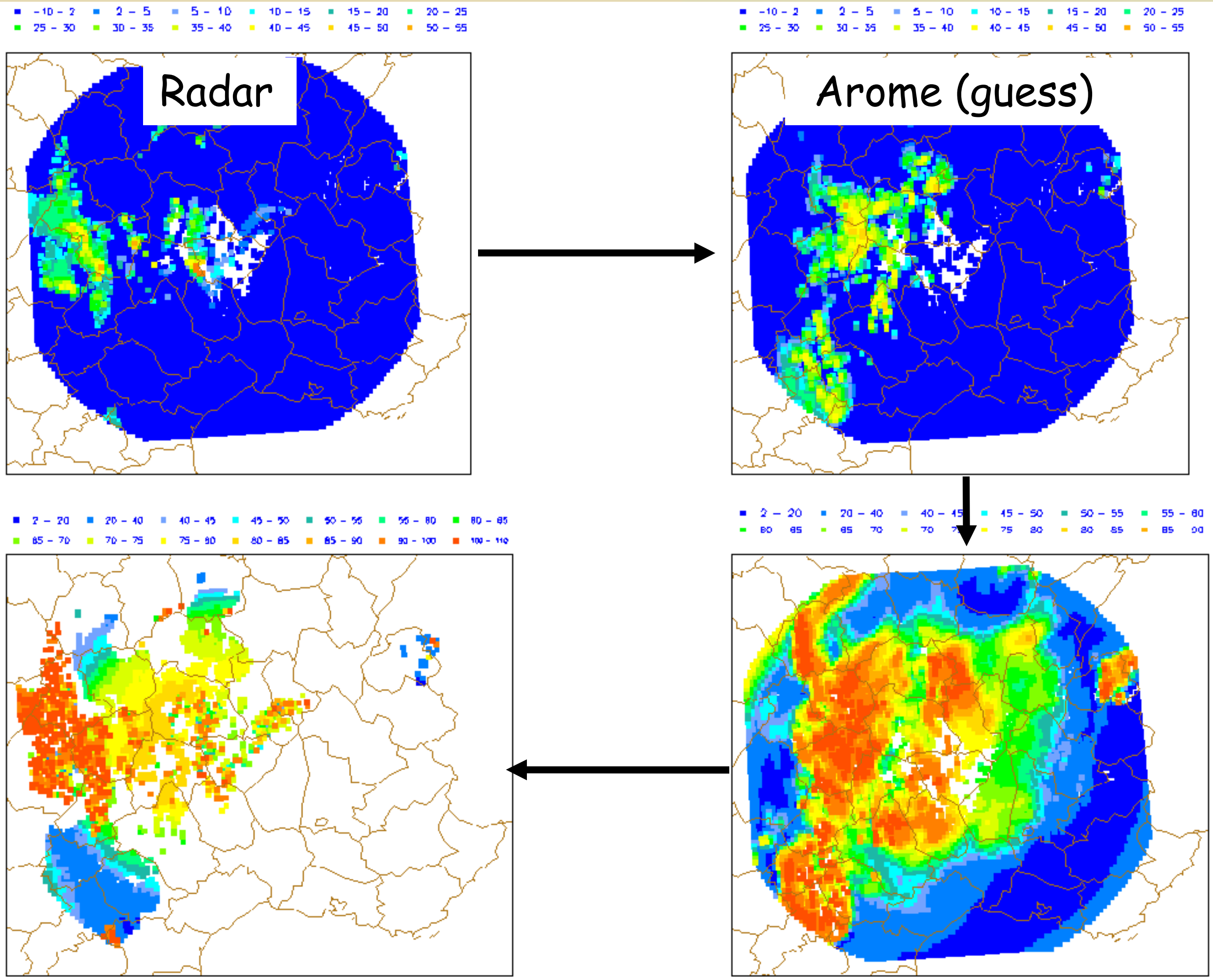
- Consistency between the retrieved profile and clouds/precipitations that the model is able to create
- Possibility of wrong solution if model too far from the reality… needs checking

Radar data assimilation : one radar assimilated

Reflectivities

Elevation 0.44°

Relative humidity

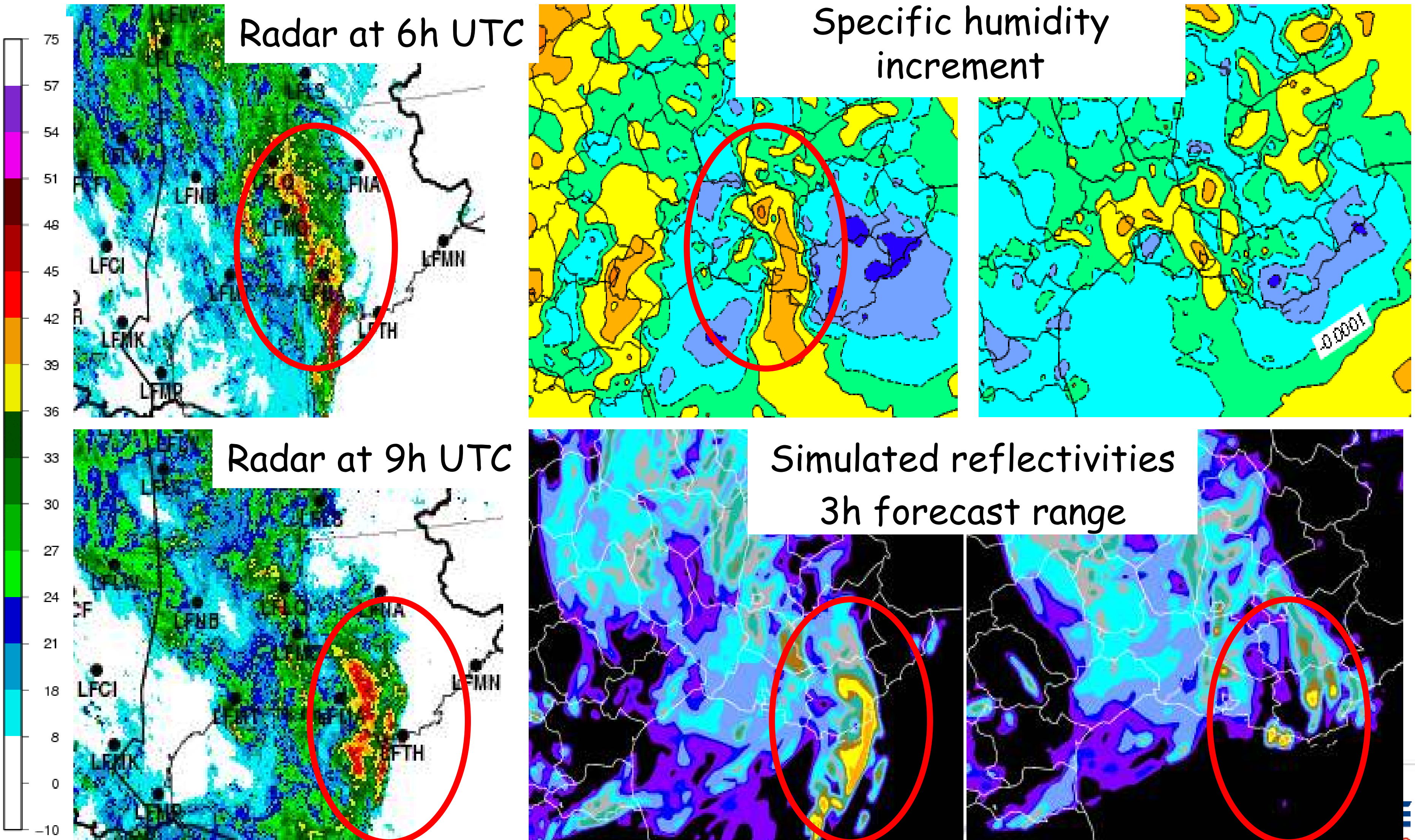


Pseudo-observations

Arome (guess)

Radar data assimilation : case study

With reflectivities — Without reflectivities



Toujours un temps d'avance

Outlines

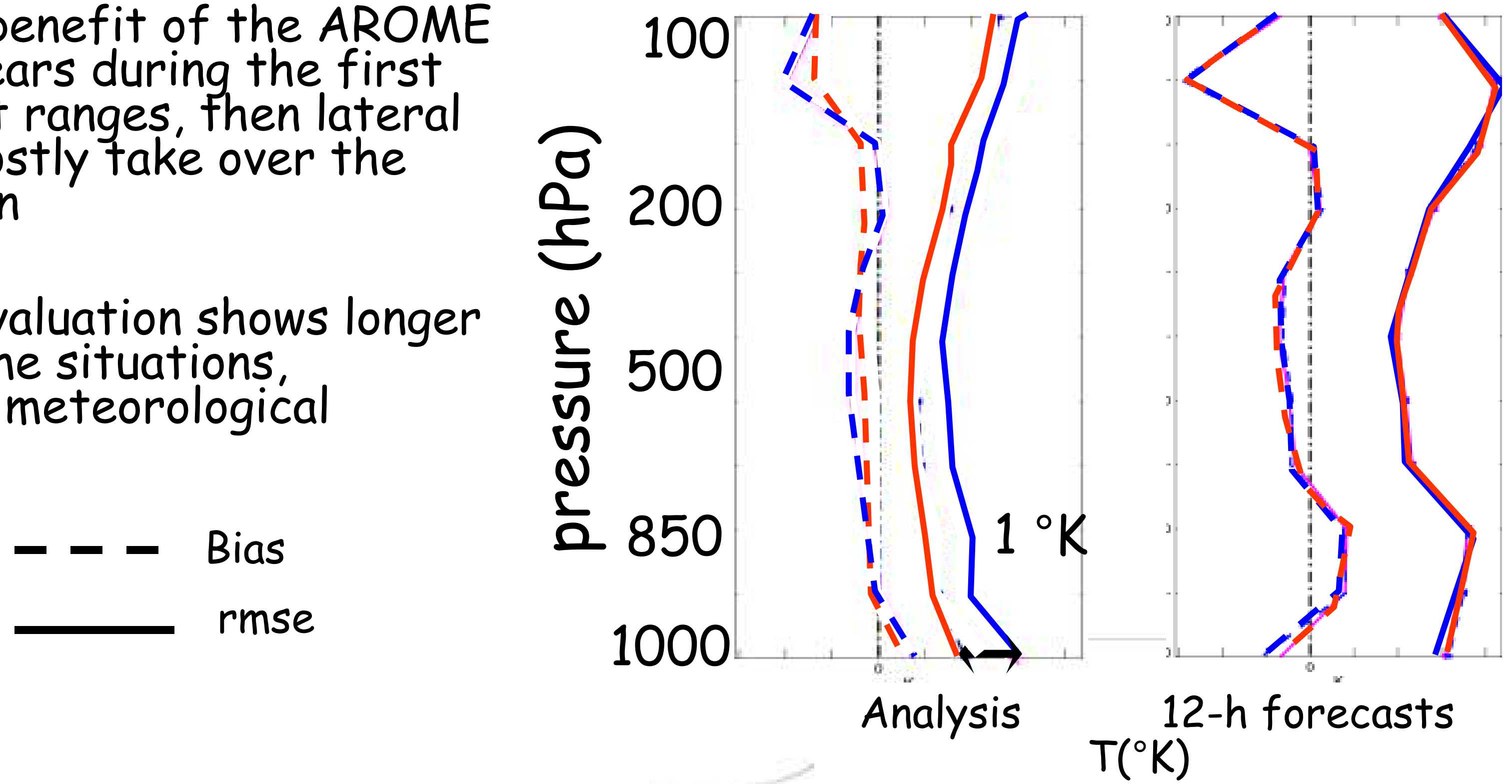
- General ideas on AROME
- The cycle strategy
- Background-error Statistics
- Observations
- Assimilation results

Objective scores : analysis and forecast compared to radiosondes

- Analysis from the AROME RUC compared to ALADIN analysis show an important reduction of Root Mean Square Error and Bias for all parameters all over the troposphere
- AROME 12-h forecasts initialized with an analysis from the AROME RUC and an ALADIN analysis (spin-up mode) seem very close compared to radiosonde.

Vertical profiles of rmse and biais for temperature compared to radiosonde measurements

- The general benefit of the AROME analysis appears during the first 12-h forecast ranges, then lateral conditions mostly take over the model solution
- Subjective evaluation shows longer impact on some situations, depending on meteorological conditions



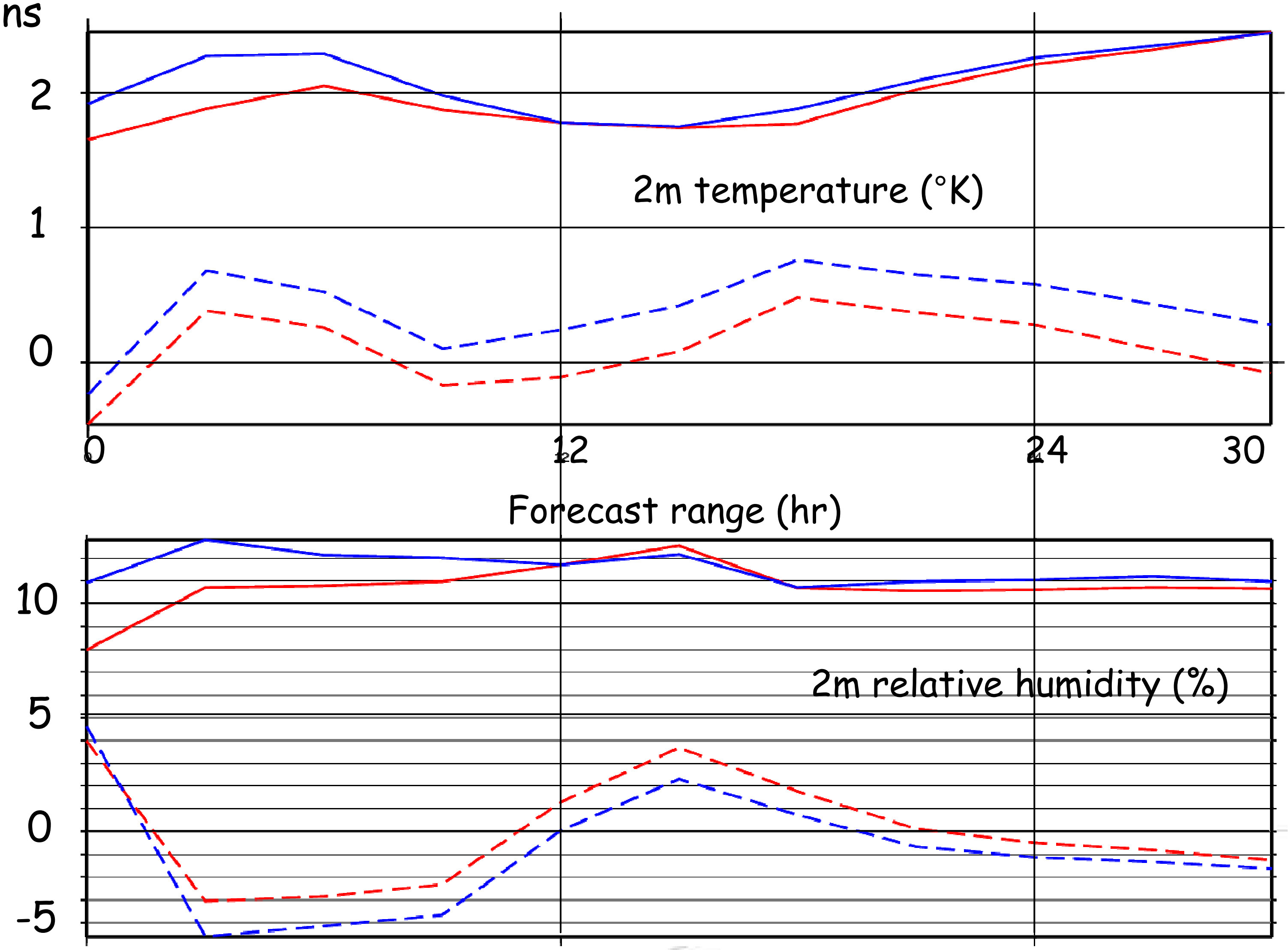
Objective scores : forecast compared to synop

- The same feature is observed regarding scores compared to SYNOP observations

Temporal evolution of rmse and biais compared to SYNOP measurements

- - - Bias

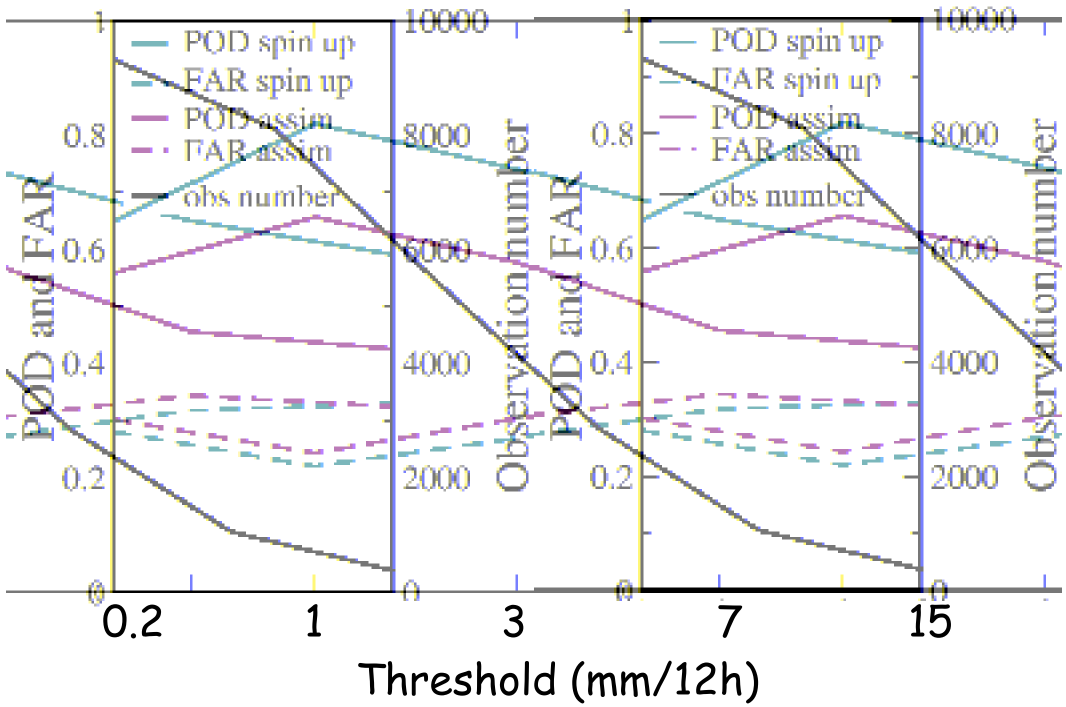
——— rms



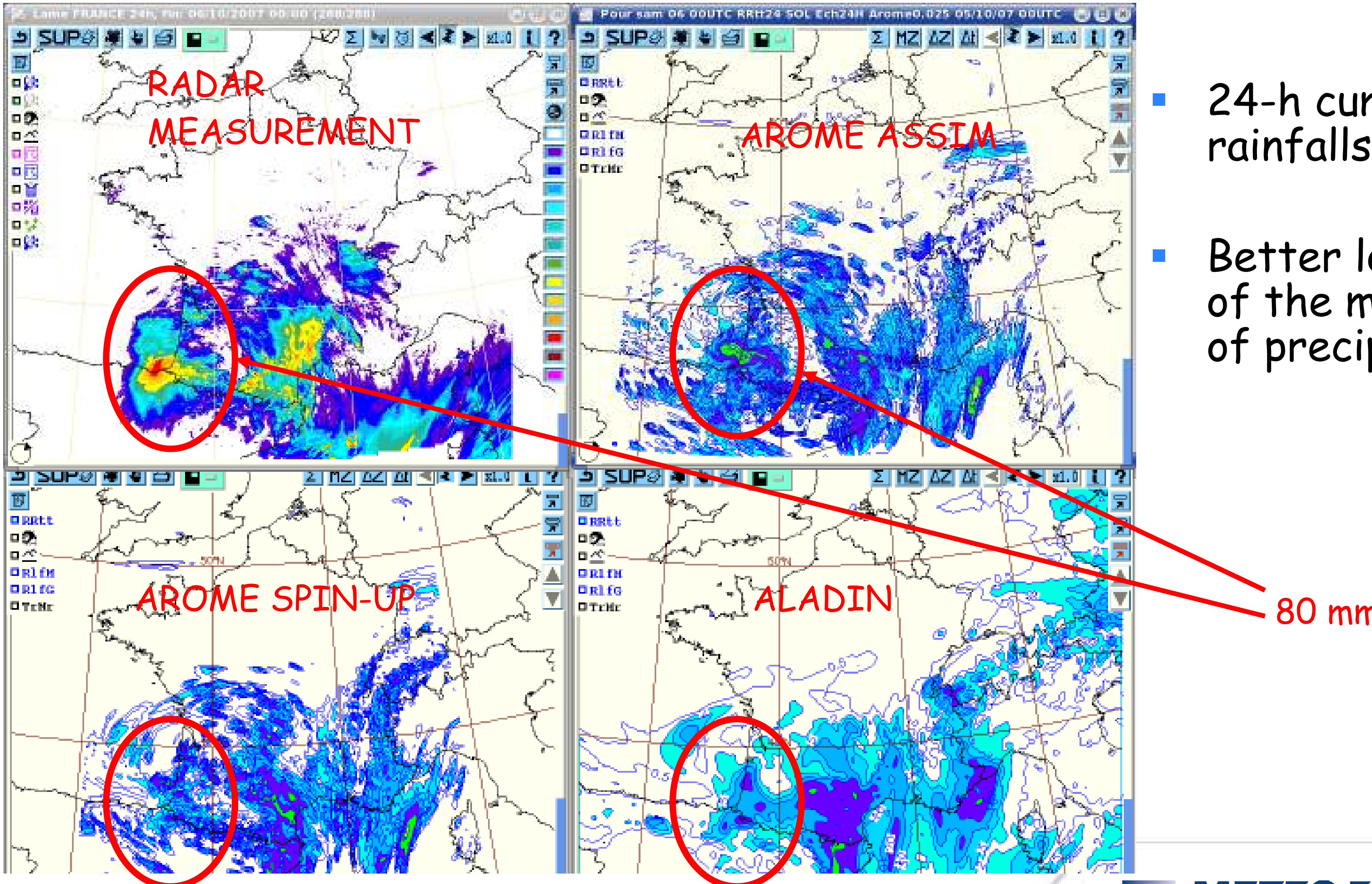
assimilation
spin-up mode

Quantitative Precipitation Forecast scores

- QPF scores for different thresholds for the total rain forecast between 0- and 12-h compared to raingauge measurements in november 2007. With AROME analysis :
  - 20 % increase of POD
  - 2 % increase of FAR



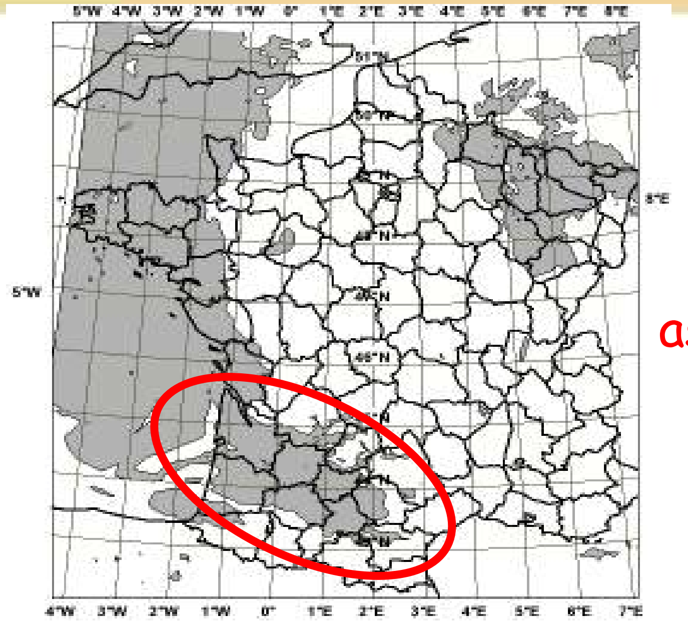
Precipitating event, 5 october 2007



- 24-h cumulative rainfalls
- Better location of the maximum of precipitation

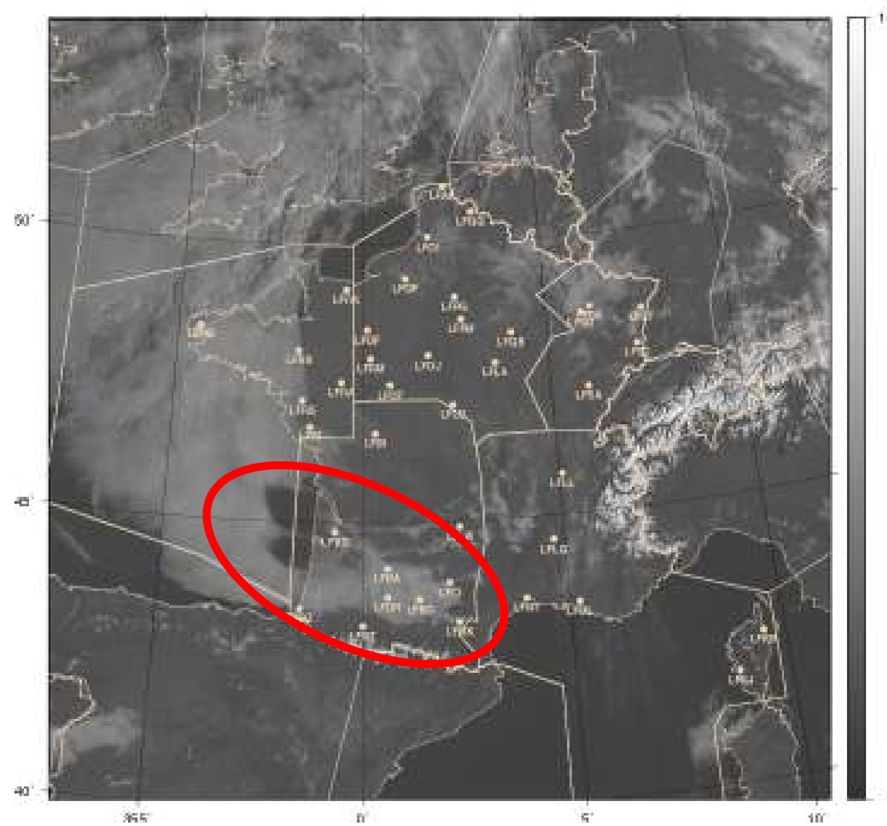
Fog event, 7 february 2008

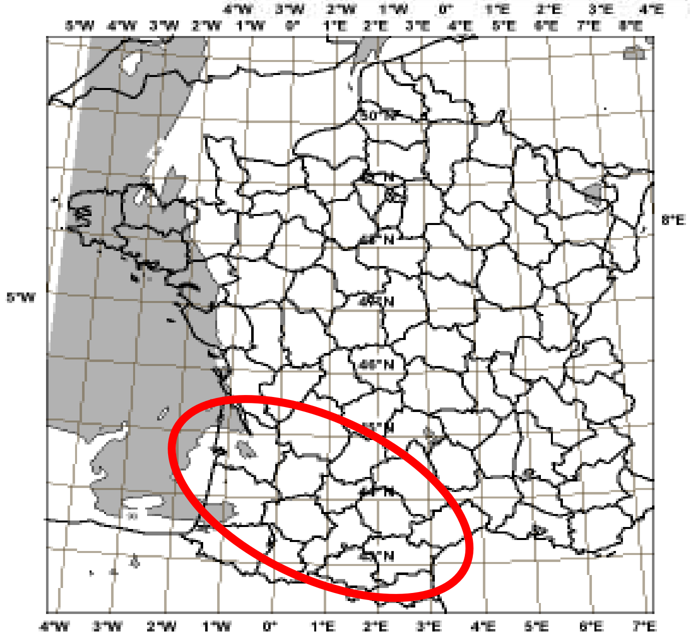
- AROME low cloud cover at 9-h UTC
- Fog is not simulated in spin-up mode



assimilation

METEOSAT VIS 07 02 2008 09h15Z





Spin-up

Conclusion

- AROME and its data assimilation system is operationally running at Météo-France on the meso scale since the end of 2008 using a 3-h continuous assimilation cycle.

- This system is supplied by the same kind of observation as the ALADIN-FRANCE operational suite and radial velocities from doppler RADAR of the french network

- The background-error statistics for AROME have been calculated by an ensemble method using the same multivariate formulation as in ALADIN-France. Compared to the ALADIN-FRANCE ones
  - background-error standard deviations are increased
  - horizontal correlation lenghtscales are much shorter

  Analysis increments are stronger and narrower.

- This system shows its ability to improve analysis and forecasts, giving a more realistic depiction of initial conditions.

- The general benefit of the analysis appears during the first 12-hour forecast ranges, then lateral conditions mostly take over the model solution.

Outlook

- Works currently in progress on :
  - the assimilation of RADAR reflectivities (in pre-operational suite this summer).
  - the use of observations at a higher spatial resolution (airep, IR radiances,…).
  - A surface assimilation coherent with the model's surface scheme and resolution.

- Works are planned in order to
  - take a better advantage of high-frequency observations using :
    - 3D-FGAT (First Guess at Appropriate Time) assimilation scheme
    - Incremental Digital Filter Initialization allowing 1-h cycle
  - introduce flow dependence in forecast error statistics using an ensemble assimilation

Bibliography

- Benjamin S. G., et al., 2004 : An hourly assimilation-forecast cycle : The RUC, *Mon. Wea. Rev*., **132**, 4959-518.
- Berre L., 2000 : Estimation of synoptic and mesoscale forecast error cavariances in a limited area model, *Mon. Wea. Rev*., **128**, 644-667.
- Berre et al., 2006 : The representative of the analysis effect in three error simulation techniques. *Tellus*, **58A**, 196-209.
- Caumont, O., 2007 : Simulation et assimilation de données radar pour la prévision de la convection profonde à fine échelle. PhD of Université de Toulouse, 252 p
- Ficher et al., 2005 : An overview of the variational assimilation in the ALADIN/France numerical weather-prediction system, *Quart. J. Roy. Meteo. Soc*., **131**, 3477-3492.
- Montmerle, T and C. Faccani, 2008: Mesoscale assimilation of radial velocities from Doppler radar in a pre-operational framework. *Mon. Wea. Rev, in press*
- Wattrelot E.et al. 2008 : 1D+3Dvar assimilation of radar reflectivities in the pre-operational AROME model at Météo-France *Erad2008, Helsinki (Finland)*

Thank you for your attention…

Assimilation scheme

- Derived from the ALADIN-FRANCE 3D-Var scheme (Fisher et al. 2005) operationally running at Météo-France on the regional scale since the end of 2005

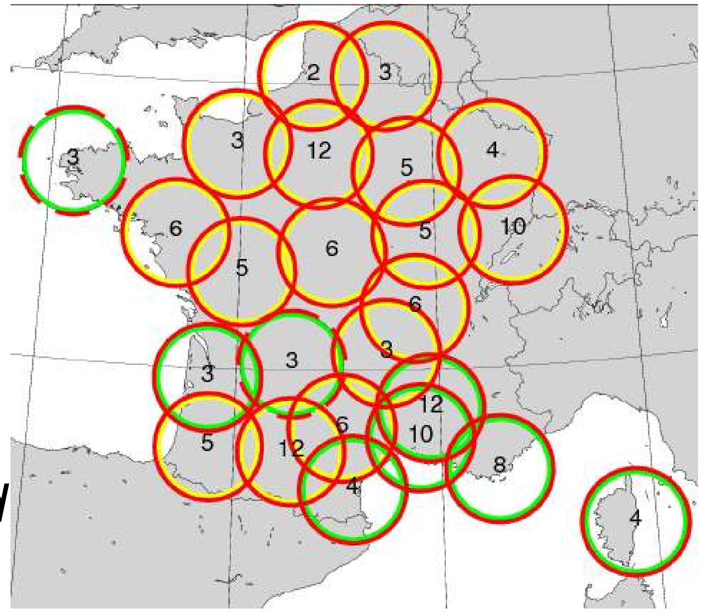
- Developed in the frame of ARPEGE/IFS software (Courtier et al. 1991), it inherits lots of its characteristics :
  - incremental formulation,
  - observation operators,
  - minimization technique,
  - data flow …

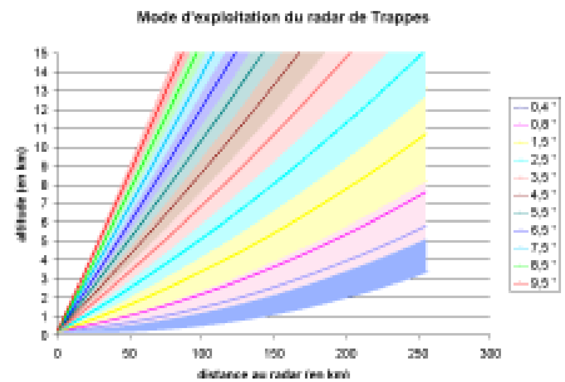
- Efficient also on the meso-scale, after some adaptations

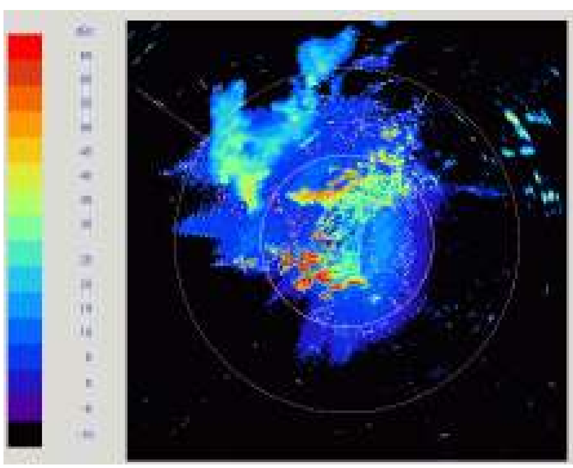
Radar data assimilation : French network

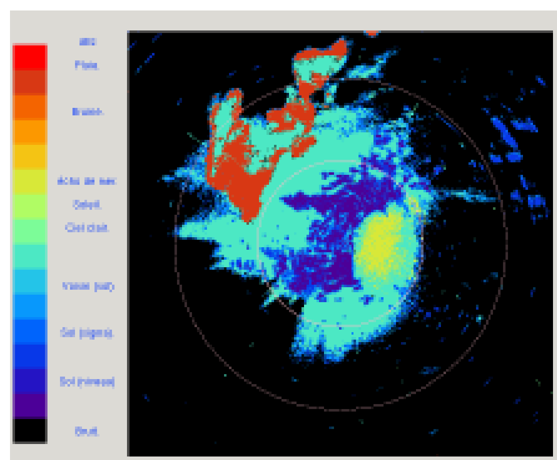
24 radars: 16 in C band (yellow circles) + 8 in S band (green circles). Volumes reflectivities (from 2 to 13 elevations).

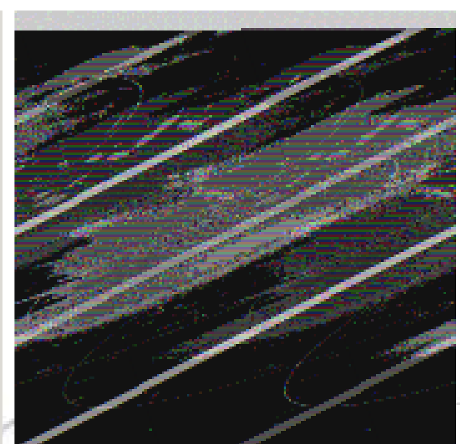
22 Doppler radars (red circles), 2 planned (dashed red circles)









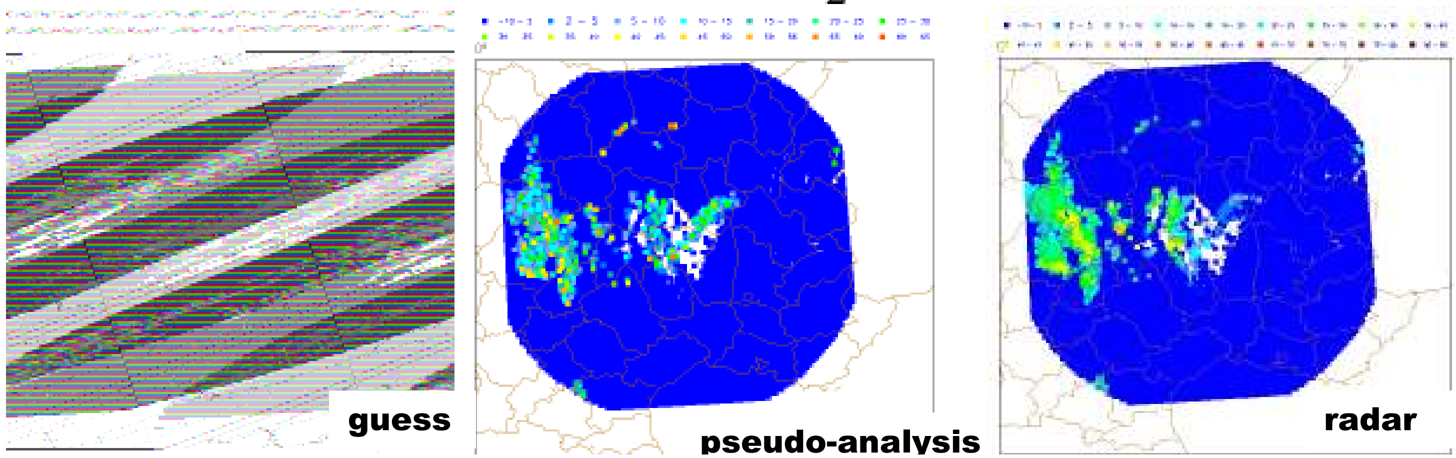


To evaluate the 1D-method, use of a « pseudo-analysis of reflectivity »

Easy to compute by use of the weights of the 1D inversion…

$$
\begin{aligned}
Z_{ps} &= \sum_i Z_i \frac{exp(-\frac{1}{2} J_z(x_i))}{\sum_j exp(-\frac{1}{2} J_z(x_j))} \\
J_z(\mathbf{x}) &= \frac{1}{2}{}^t (\mathbf{z} - H_z(\mathbf{x})) \mathbf{R}_z^{-1} (\mathbf{z} - H_z(\mathbf{x}))
\end{aligned}
$$

Reflectivity field



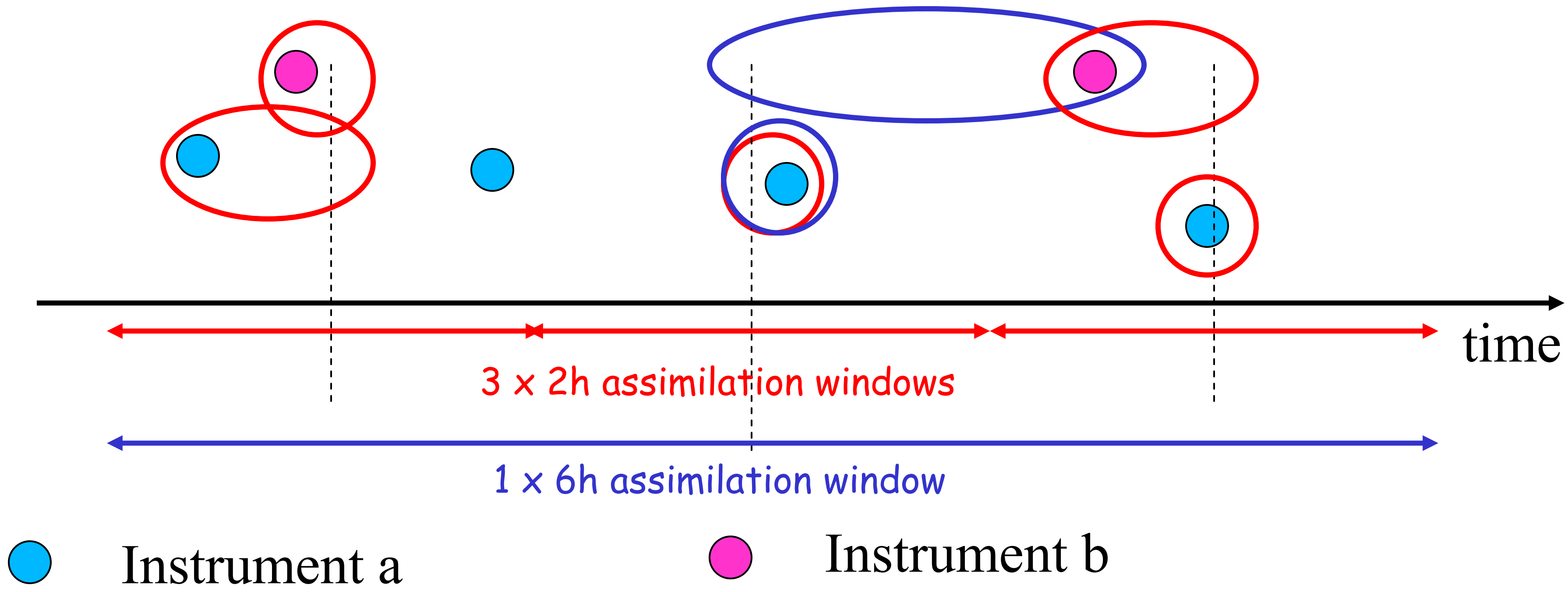
If $||Z_{ps} - Z_{obs}||$ low =>

1. good convergence of the 1D method. (RMS deviation can be a measurement of the quality of the retrieved profiles: useful for monitoring)

2. good consistency between the pseudo-observations and the model (because of use of model information in the 1D-inversion). Used for the quality control in the screening

3. Possibility to take into account this value in the choice made in the thinning of the observations (tests are underway)

Assimilation window



- 3D-Var : short and numerous assimilation windows = more observations assimilated in a more realistic way

Background error statistics : heterogeneous B matrix

Thibault Montmerle (Following an idea of Philippe Courtier (1998), and used by Mark Buenher 2008).

- To use more suitable background error statistics in clear air and precipitating areas : $\mathbf{B} = \alpha \mathbf{B}_p + \beta \mathbf{B}_{np}$

$$\mathbf{B} = \underbrace{\begin{pmatrix} \alpha^{1/2}\mathbf{B}_p^{1/2} & \beta^{1/2}\mathbf{B}_{np}^{1/2} \end{pmatrix}}_{\mathbf{B}^{1/2}} \underbrace{\begin{pmatrix} \mathbf{B}_p^{T/2}\alpha^{T/2} \\ \mathbf{B}_{np}^{T/2}\beta^{T/2} \end{pmatrix}}_{\mathbf{B}^{T/2}}$$

With: $\alpha = \mathbf{FMF}^{-1}$ and $\beta = \mathbf{F(1\text{-}M)F}^{-1}$

$\mathbf{M}$: grid point mask deduced from observed radar reflectivity.

$\mathbf{B_p}$ and $\mathbf{B_{np}}$ are separately computed by performing statistics on an assimilation ensemble of precipitating cases, considering a mask based on simulated precipitations.

The increment writes:

$$\delta x = \mathbf{B}^{1/2}\chi = \begin{pmatrix} \alpha^{1/2}\mathbf{B}_1^{1/2} & \beta^{1/2}\mathbf{B}_2^{1/2} \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}$$

$\Rightarrow$ Which implies to double the control variable χ and the gradient $\nabla_\chi J$

Background error statistics : heterogeneous B matrix (2)

Multivariate formulation of errors:

$$\zeta = \zeta \qquad \zeta = \zeta$$

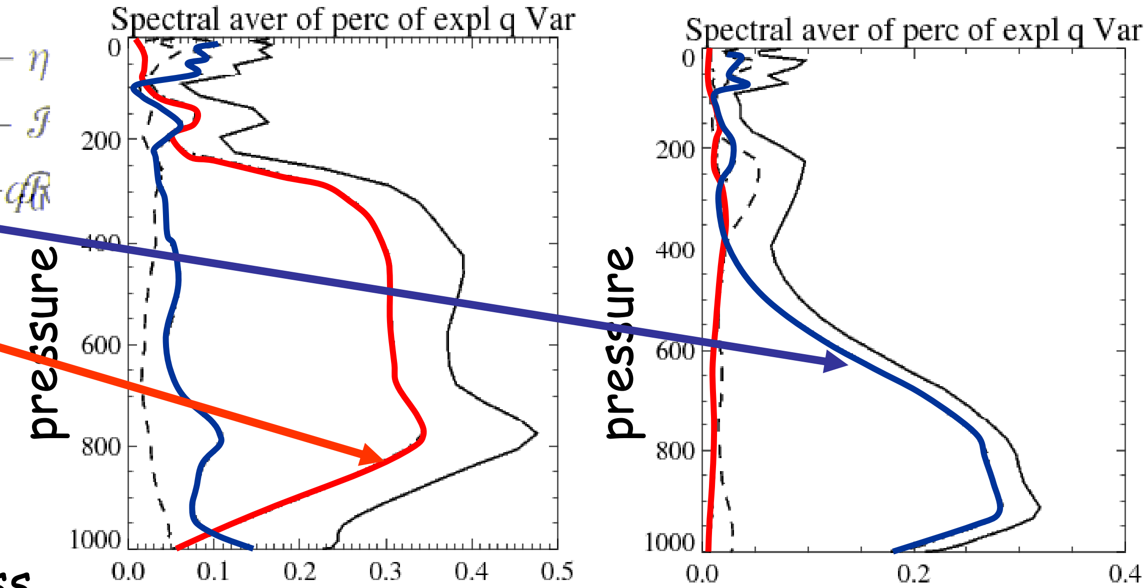
$$\eta = \mathcal{M}\mathcal{H}\zeta + \eta_u \qquad \eta = \mathcal{M}\mathcal{H}\zeta + \eta$$

$$(T, P_s) = \mathcal{N}\mathcal{H}\zeta + \mathcal{P}\eta_u + (T, P_s)_u \qquad (T, P_s) = \mathcal{N}\mathcal{H}\zeta + \mathcal{P}$$

$$q = \mathcal{Q}\mathcal{H}\zeta + \mathcal{R}\eta_u + \mathcal{S}(T, P_s)_u + q_u \qquad q = \mathcal{Q}\mathcal{H}\zeta + \mathcal{R}$$

In precipitating areas, $\sigma_b(q)$ is mostly explained by η_u at mesoscale, whereas it is almost univariate and linked to the mass field in clear air

Vertical profile of spectral averages of the percentage of explained humidity variance



B_{rain} $B_{no\ rain}$

total

unbalanced divergence

balanced geopotential

unbalanced mass field

$\Rightarrow$ B_r and B_{nr} are characterized by very different structure functions, which is coherent with the model's physics in both precipitating and non-precipitating areas.

Comparisons between structure functions

Multivariate formulation of errors:

$$\zeta = \zeta \qquad \zeta = \zeta$$

$$\eta = \mathcal{M}\mathcal{H}\zeta + \eta_u \qquad \eta = \mathcal{M}\mathcal{H}\zeta + \eta$$

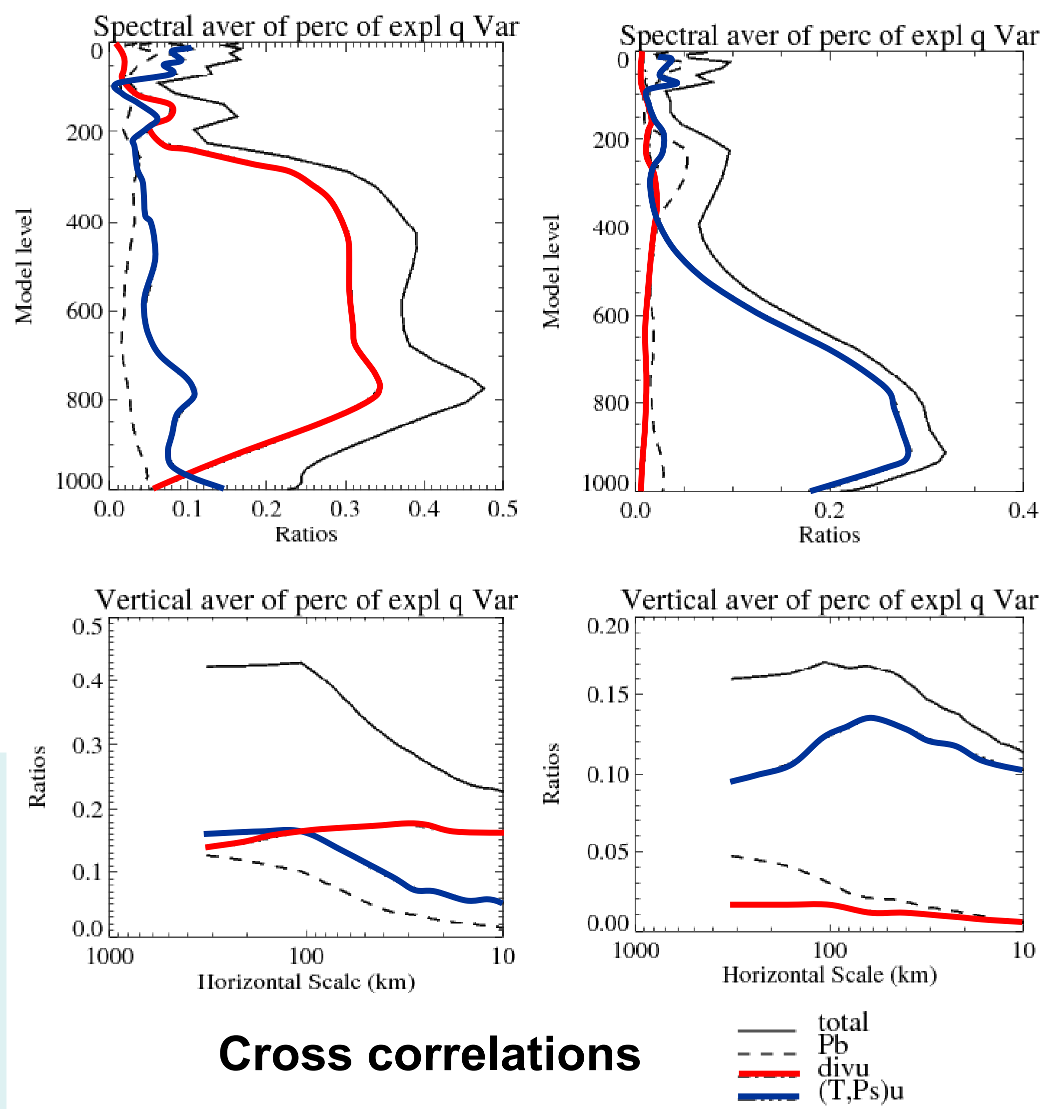
[illegible]

[illegible]

In precipitating areas, σ_b(q) is mostly explained by η_u at mesoscale, whereas it is almost univariate and linked to the mass field in clear air

$\Rightarrow$ B_p et B_{np} are characterized by very different structure functions that are coherent with the model's physic in precipitating and non-precipitating areas respectivelly

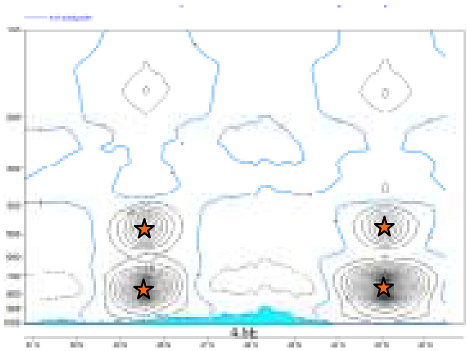
$\bar{\sigma}_b(q)$ $\mathbf{B}_p$ $\mathbf{B}_{np}$



Cross correlations

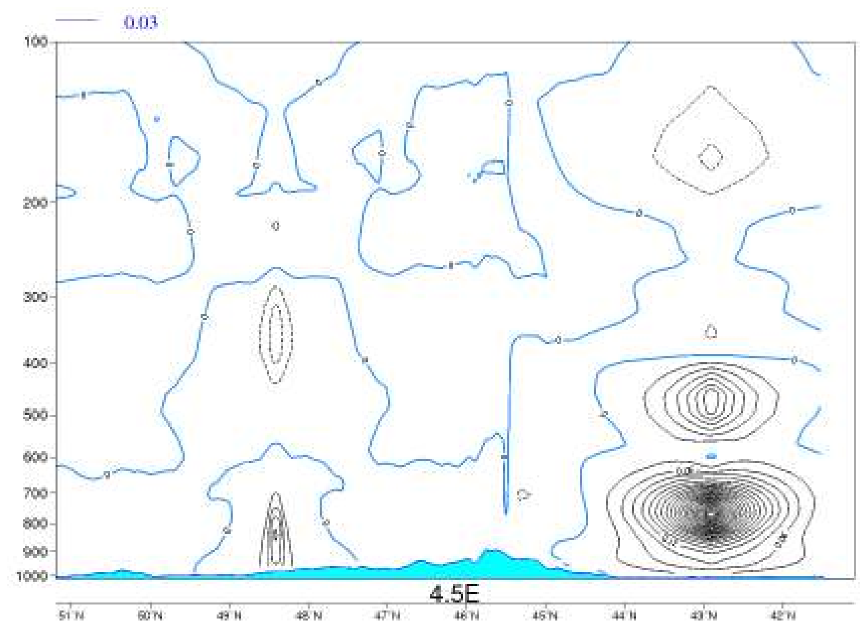
2 obs experiment

Innovations of – 30% RH
At 800 and 500 hPa

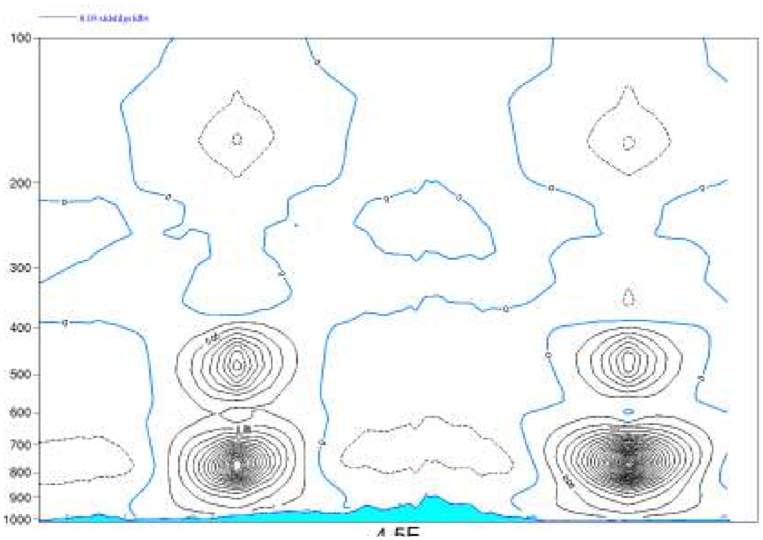


$\delta x = B_{np}^{1/2}\chi$

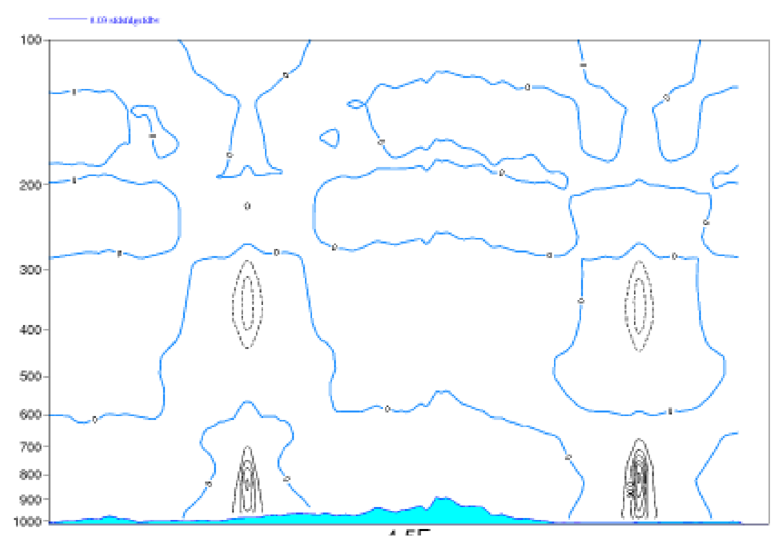
$\delta x = B_{p}^{1/2}\chi$



$\delta x = \alpha^{1/2}B_{np}^{1/2}\chi_1 + \beta^{1/2}B_{p}^{1/2}\chi_2$



$\delta x = \beta^{1/2}B_{np}^{1/2}\chi_1 + \beta^{1/2}B_{np}^{1/2}\chi_2$



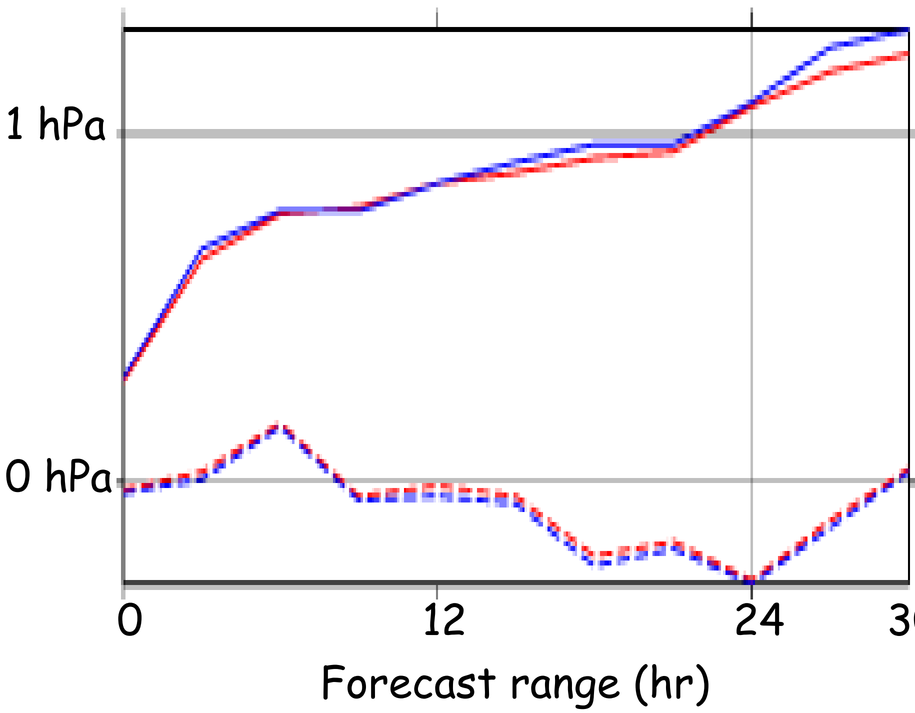
$\delta x = \alpha^{1/2}B_{p}^{1/2}\chi_1 + \alpha^{1/2}B_{p}^{1/2}\chi_2$

Cycle strategy : frequency

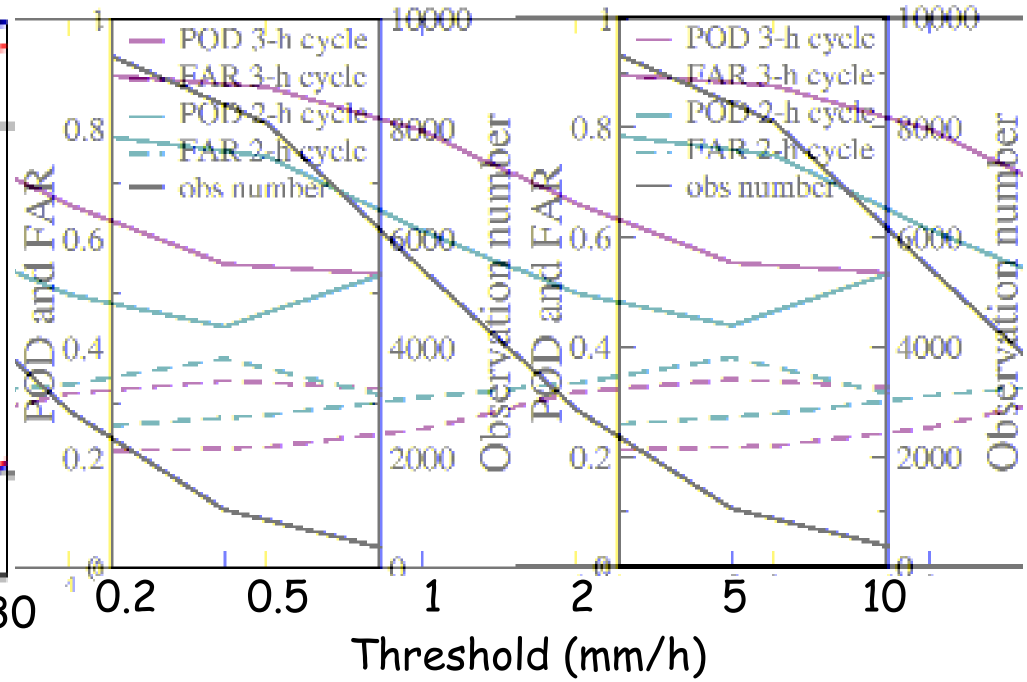
- Experiments with 1, 2 and 3-h frequency continuous cycle during a 30-day-long period
  - 1-h cycle : forecast crashed after 2 days
  - 2-h cycle : poorer performance than 3-h cycle

Scores against surface pressure observations :

—— rmse ---- biais

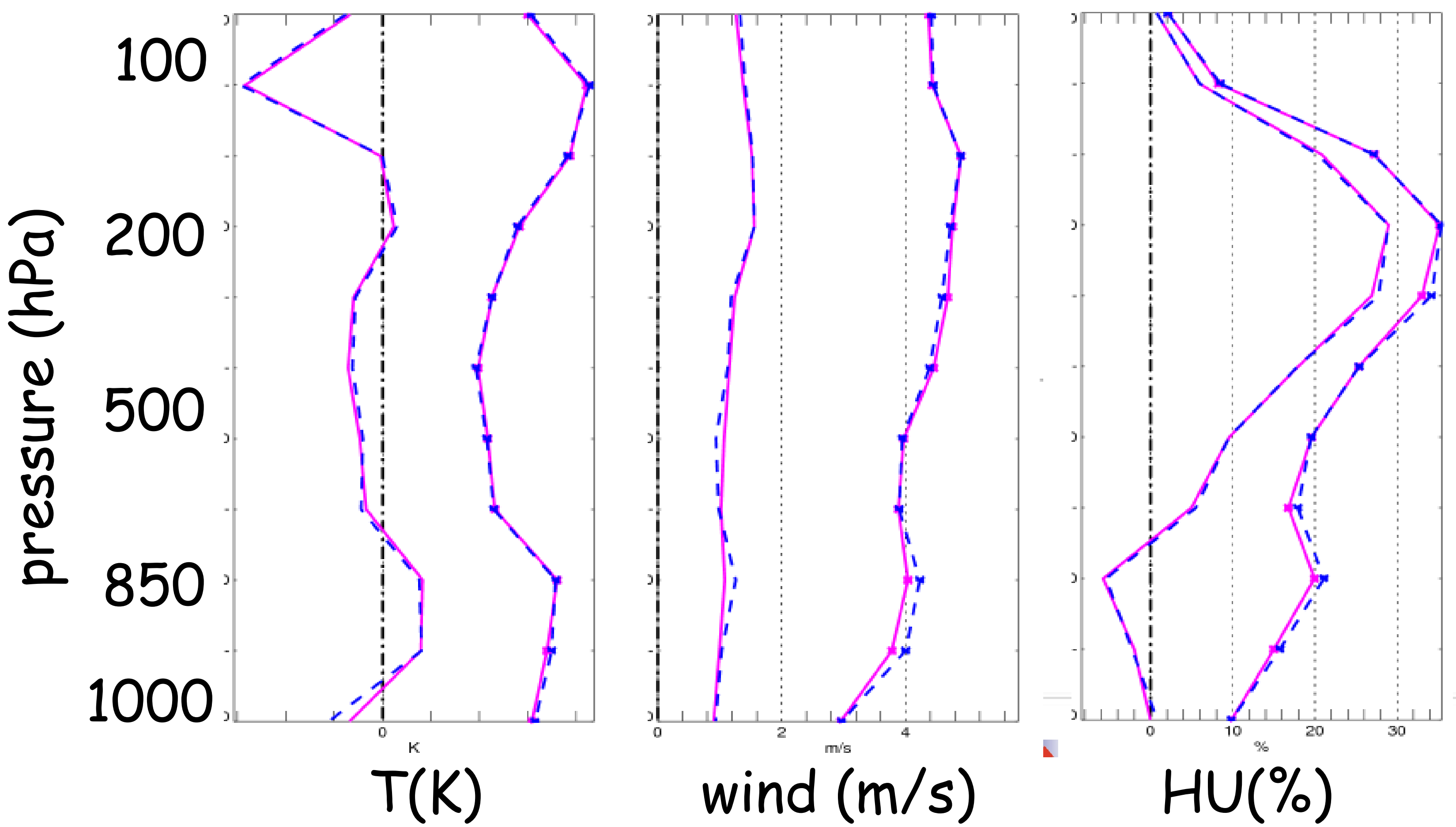


QPF scores for different thresholds for the total rain forecast between 0- and 12-h compared to rain-gauge measurements in November 2008



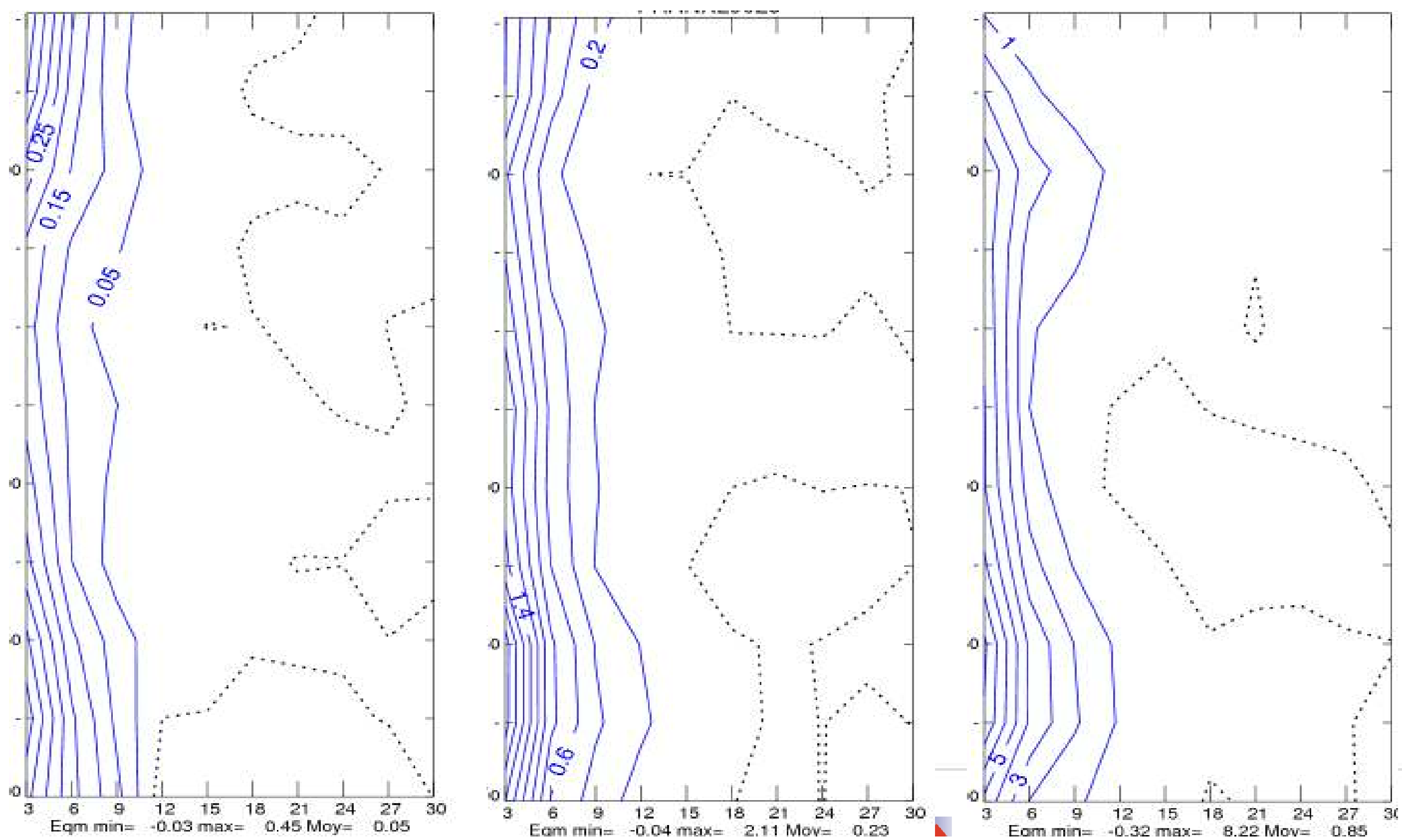
Objective scores : 12-h forecast scores compared to radiosonde

- AROME 12-h forecasts initialized with an analysis from the AROME RUC and an ALADIN analysis (spin-up mode) seem very close compared to radiosonde.



Objective scores : forecast scores compared to AROME analysis

- Differences of score compared to AROME analysis (spin-up minus assim) show forecast differences up to 12-h forecast. For longer forecast ranges, the two forecasts are very close.



T(K) wind (m/s) HU(%)