Data assimilation in meteorology

Loïk Berre

Météo-France/CNRS

CNRM

Plan of the talk

- Numerical Weather Prediction (NWP) and Data Assimilation (DA)
- Observations (in-situ and remote sensing)
- Error covariance estimation

1. Numerical Weather Prediction

and Data Assimilation

The two main ingredients of weather forecasting

What will be the weather tomorrow ?

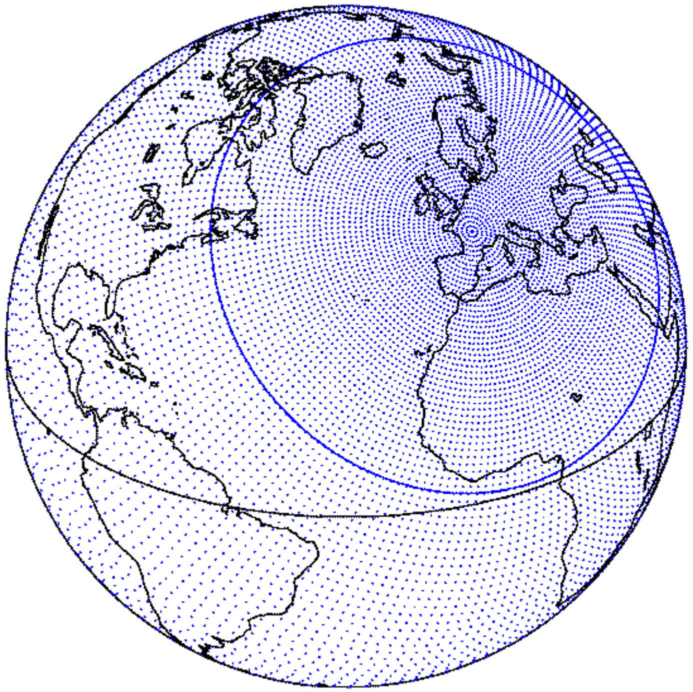
Bjerknes (1904) :

In order to do a good forecast, we need to :

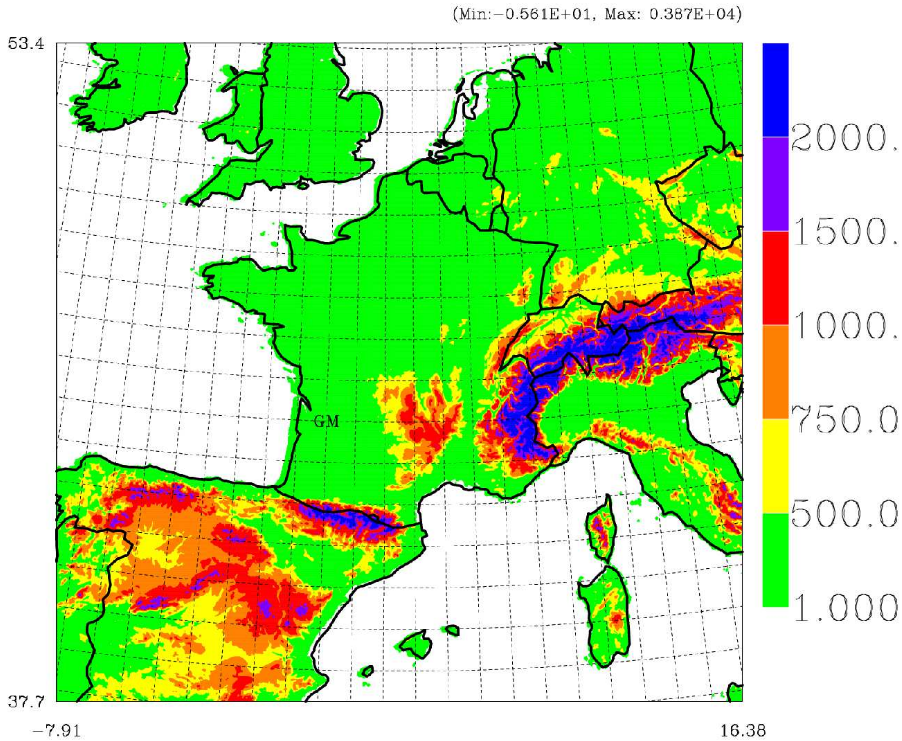
- know the atmospheric evolution laws
  (~ modeling) ;
- know the atmospheric state at initial time
  (~ data assimilation).

Numerical Weather Prediction at Météo-France
(in collaboration with e.g. ECMWF)

Global model (Arpège) : DX ~ 7-40 km

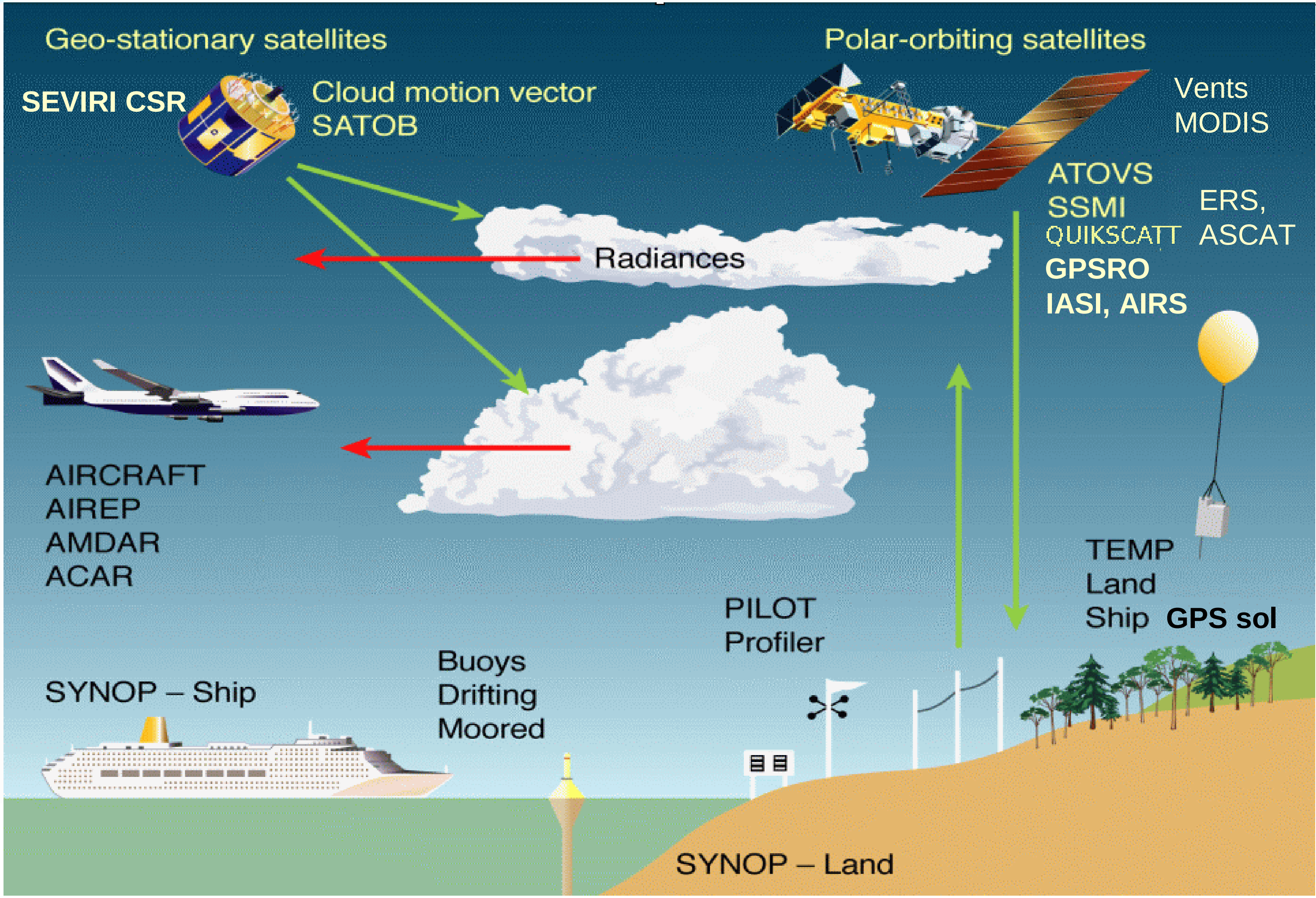


Arome : DX ~ 1.3 km



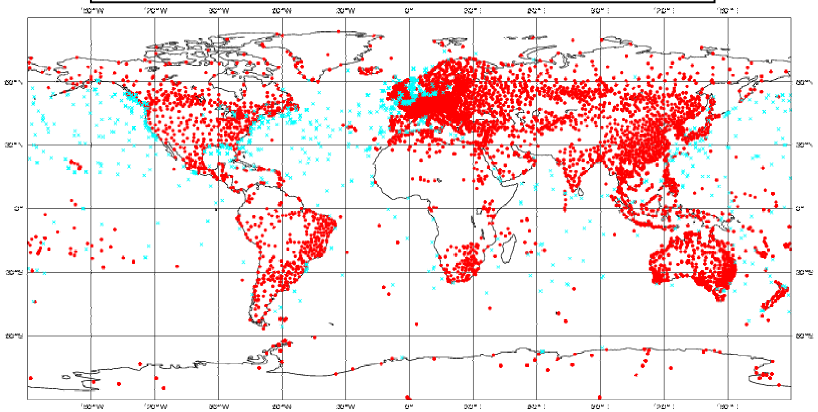
Equations of hydrodynamics and physical parametrizations (radiation, convection,…)
to predict the evolution of temperature, wind, humidity, …

Data that are assimilated in NWP models

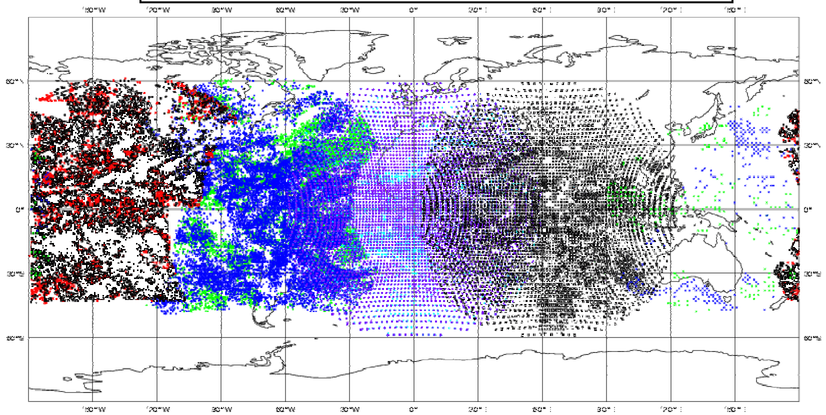


Spatial coverage and density of observations

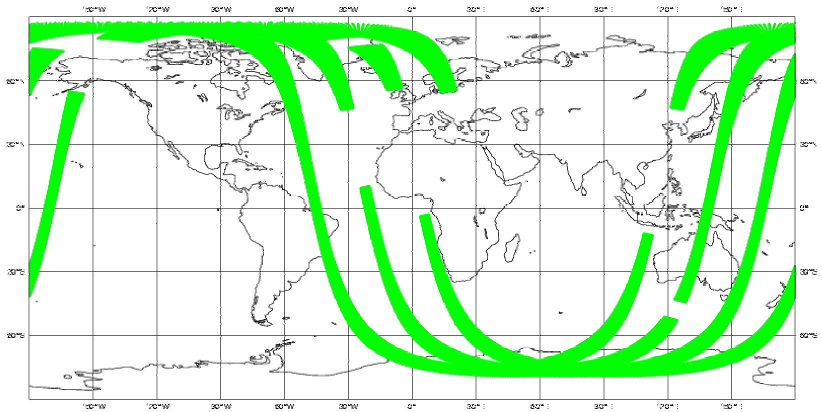
SURFACE DATA



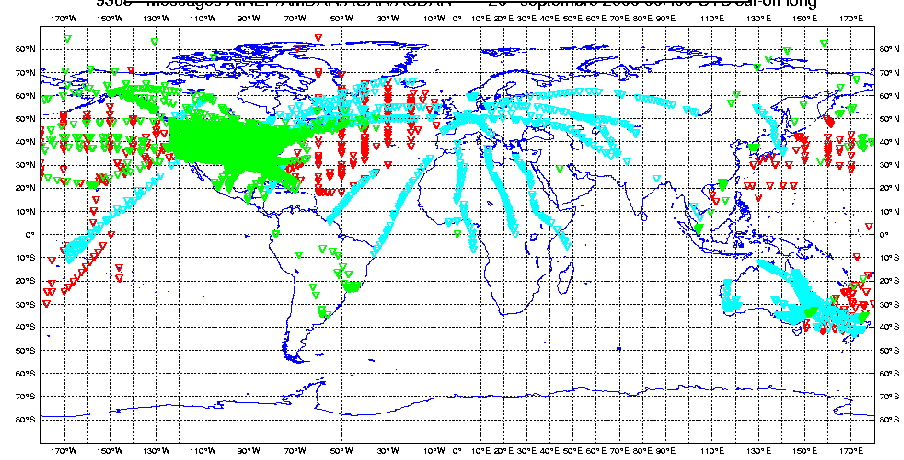
GEOSAT. WINDS



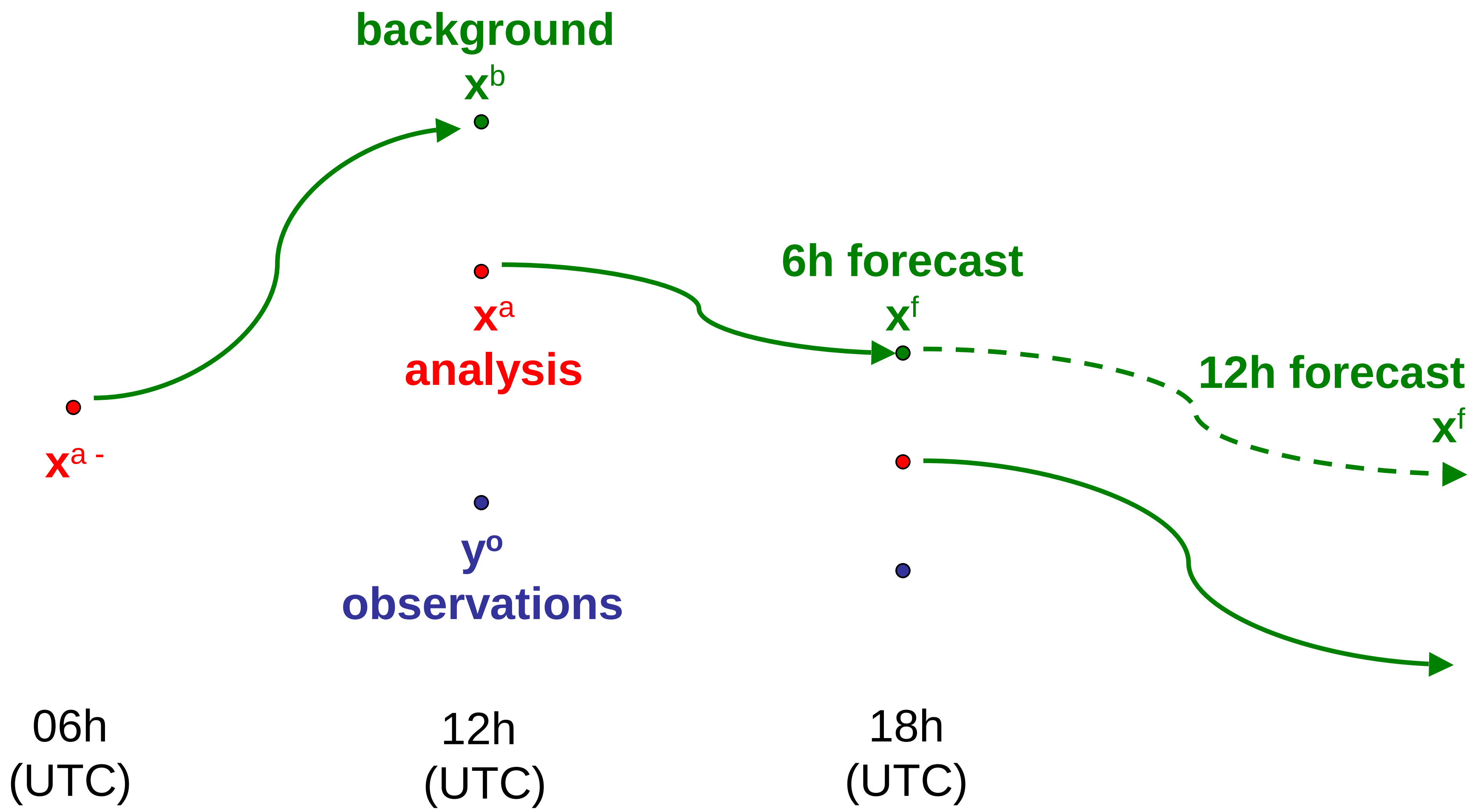
SCATTEROMETER



AIRCRAFT DATA



Temporal cycling of data assimilation : succession of analyses and forecasts



Memory of DA system is updated ~ continuously

Linear estimation of model state (1)

- BLUE analysis equation : $\mathbf{x^a = (I\text{-}KH)\, x^b + K\, y^o}$
- **H** = observation operator = projection from model to observation space (e.g. spatial interpolation, radiative transfer, NWP model).
- **K** = observation weights :

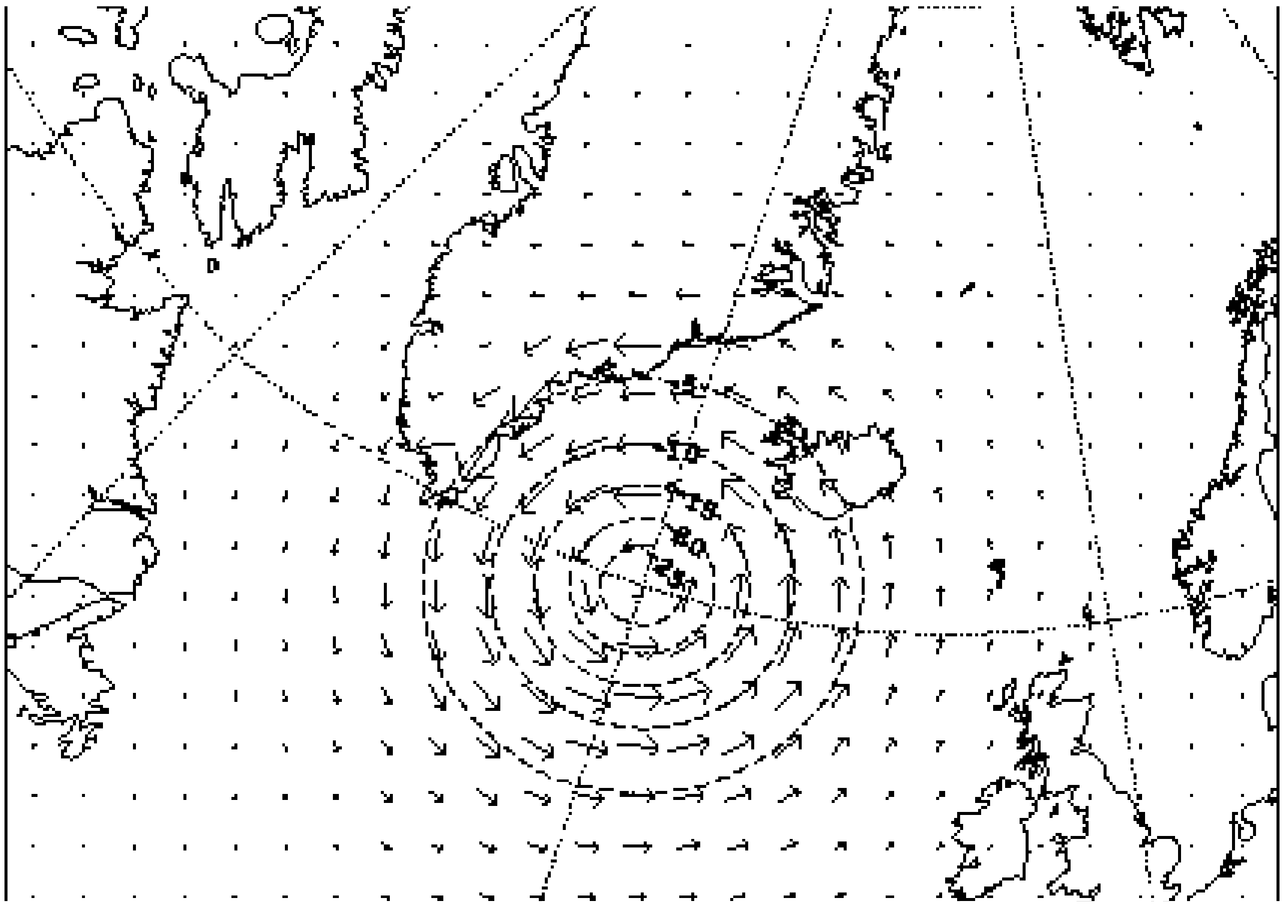
$$\mathbf{K = BH^T\,(\,HBH^T + R\,)^{-1}}$$

$$\mathbf{H\,K = (\,I + R\,(HBH^T\,)^{-1}\,)^{-1}}$$

$\Rightarrow$ ~ ratio between background error covariances (matrix **B**)
and observation error covariances (matrix **R**).

$\Rightarrow$ Accounts for relative accuracy of observations,
and for spatial structures of background errors.

Impact of one observation of temperature on the wind analysis (2D)



⇒ multivariate couplings (ex: mass/wind) are also accounted for.

Linear estimation of model state (2)

- Size of **B** is huge : square of model size ~ $(10^9)^2$ ~ 10^{18}.

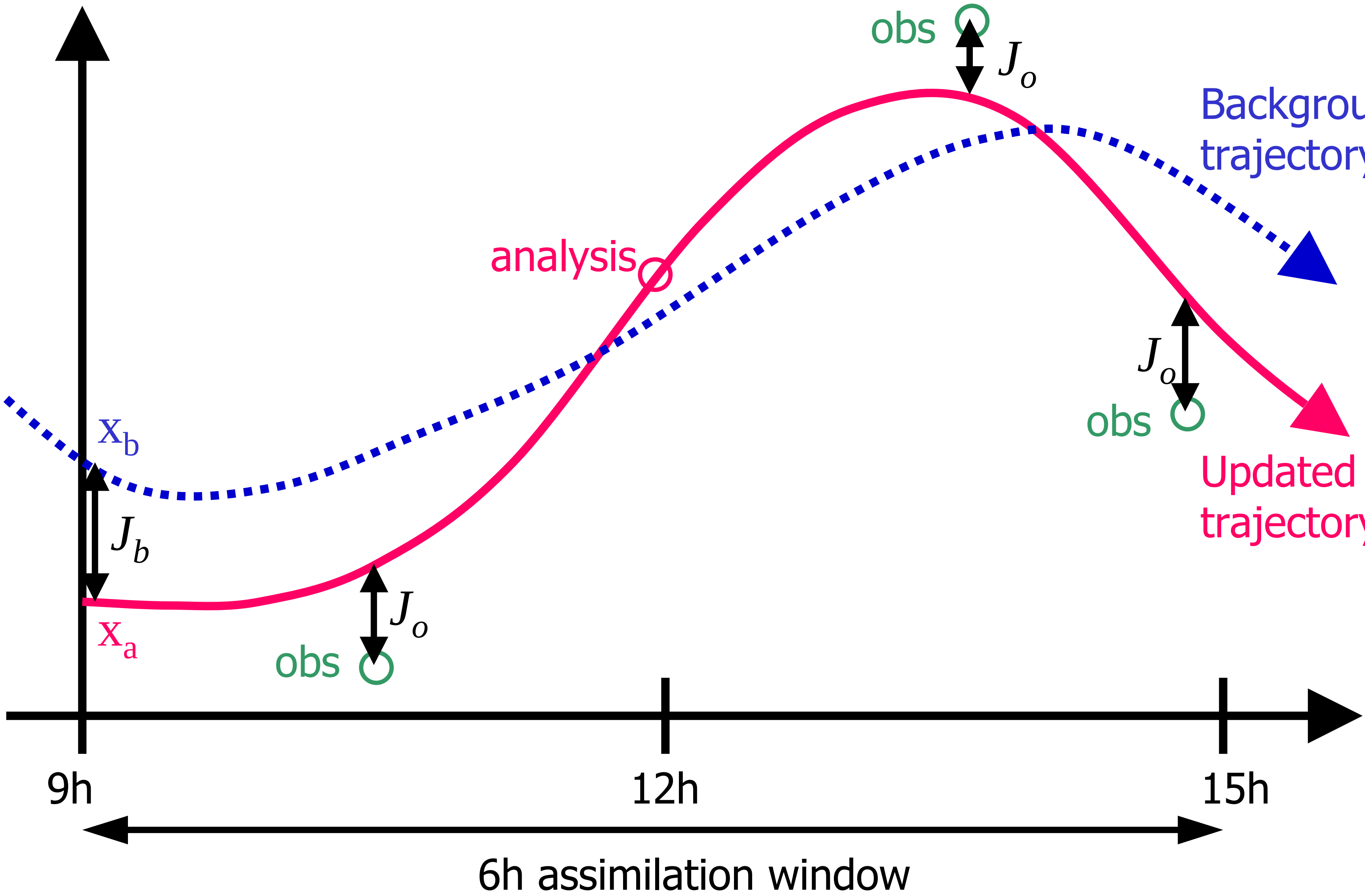
$\Rightarrow$ error covariances need to be estimated, simplified and modeled.

- Matrices too large to be inverted, but
  equivalent to minimize distance $J(\mathbf{x}^a)$ to $\mathbf{x}^b$ and y^o (4D-Var)
  without explicit matrix inversions (e.g. Talagrand and Courtier 1987).

- Non linear features accounted for
  in calculation of departures between y^o and $H(\mathbf{x}^b)$,
  and in iterative applications of 4D-Var.

Principle of 4D-VAR assimilation

(e.g. Talagrand and Courtier 1987, Rabier et al 2000)



Implementation of 4D-Var

- Analysis increment (BLUE equation) :

  $$\delta\mathbf{x} = \mathbf{x}^a - \mathbf{x}^b = \mathbf{K}\,(\,\mathbf{y}^o - \mathbf{H}\,\mathbf{x}^b\,) = \mathbf{K}\,\delta\mathbf{y}$$

  but **K** is difficult to handle explicitly in a real size system.

- Variational formulation :

  cost function : $J(\delta\mathbf{x}) = \delta\mathbf{x}^T\,\mathbf{B}^{-1}\,\delta\mathbf{x} + (\delta\mathbf{y} - \mathbf{H}\,\delta\mathbf{x})^T\,\mathbf{R}^{-1}\,(\delta\mathbf{y} - \mathbf{H}\,\delta\mathbf{x})$

  minimised when gradient $J'(\delta\mathbf{x})=0$ (equivalent to BLUE).

- Computation of J': development and use of adjoint operators (transpose).

- Generalized observation operator *H* : includes NWP model *M*.

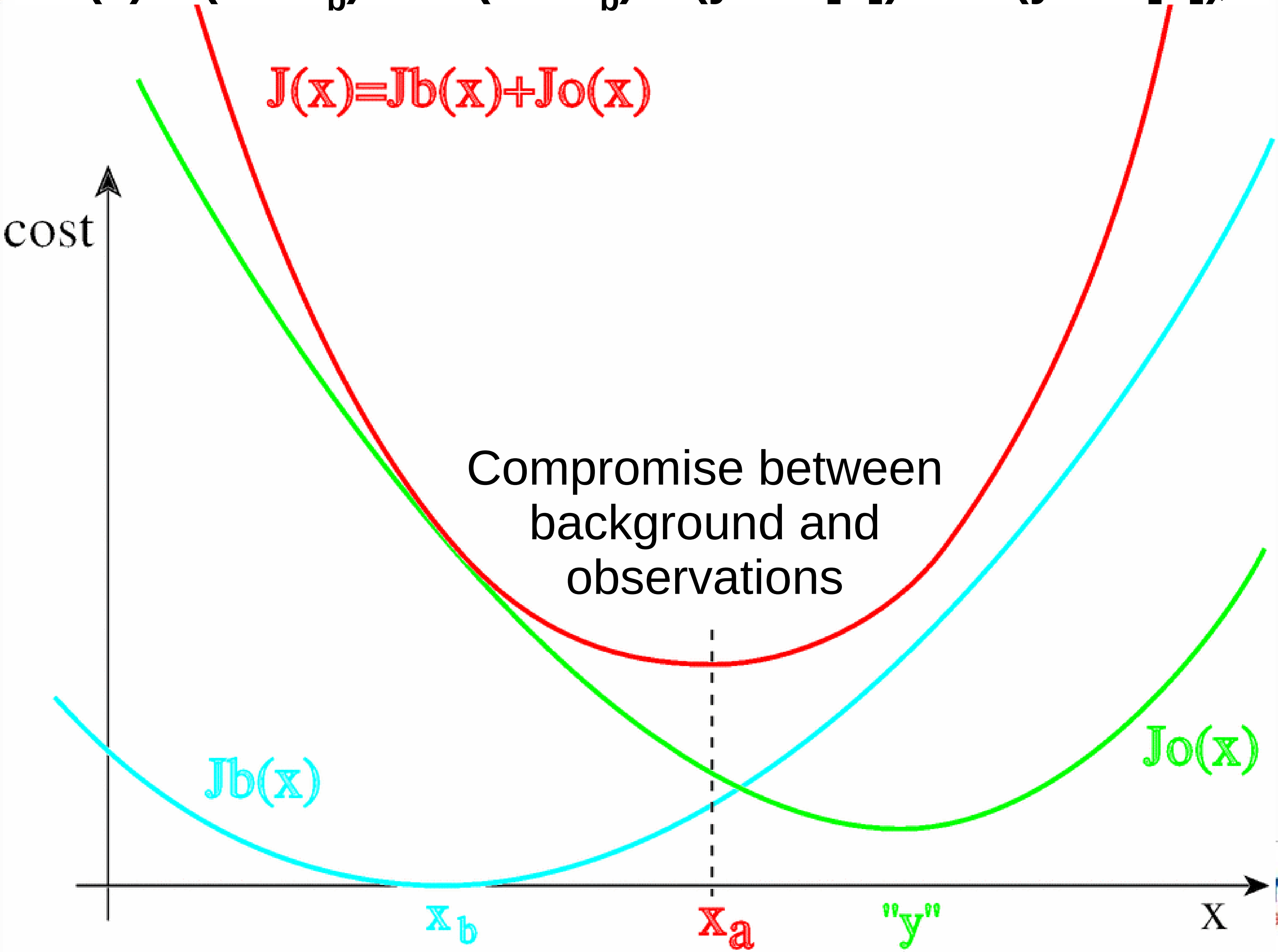
- Cost reduction :

  analysis increment $\delta\mathbf{x}$ can be computed at low resolution.

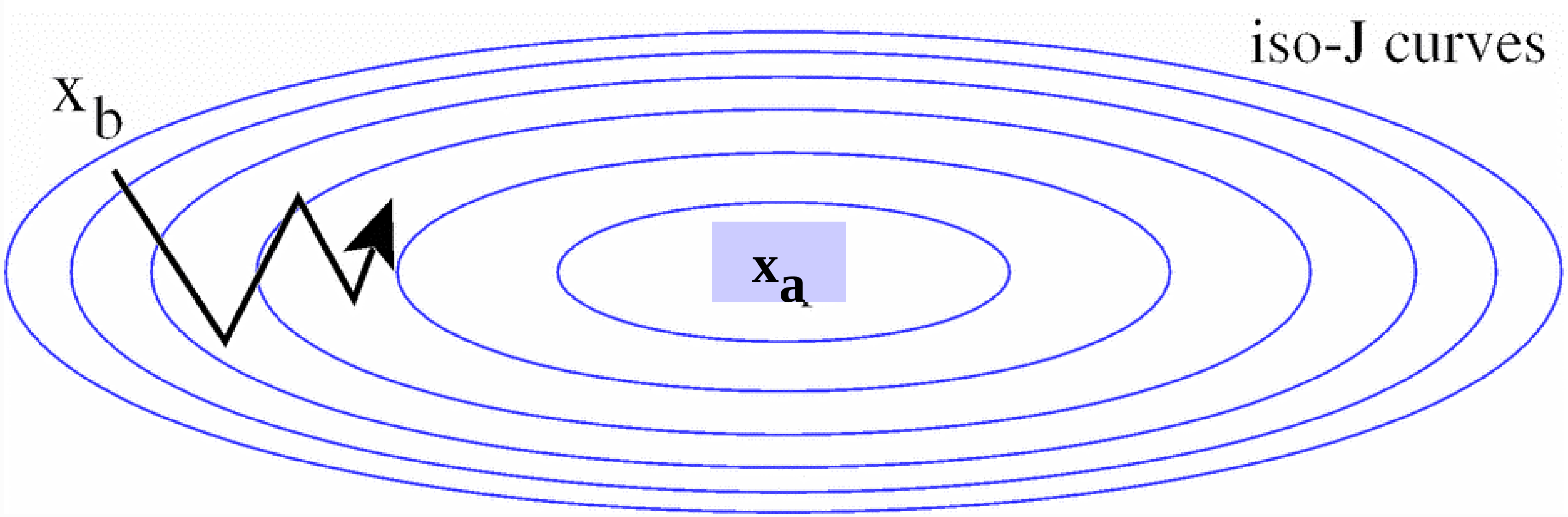
  (Courtier, Thépaut et Hollingsworth, 1994)

Schematic representation of

$$J(x) = (x - x_b)^T B^{-1} (x - x_b) + (y - H[x])^T R^{-1} (y - H[x])$$



Importance of preconditioning



• Some gradient directions have much larger amplitudes than others : problem of “narrow valley" linked to the metric of **x**.

• Use a change of variable such as J becomes nearly “circular”: much faster convergence.

2. In-situ observations and

remote sensing data

Observation networks in meteorology: in situ measurements

Observation networks in meteorology: satellite data

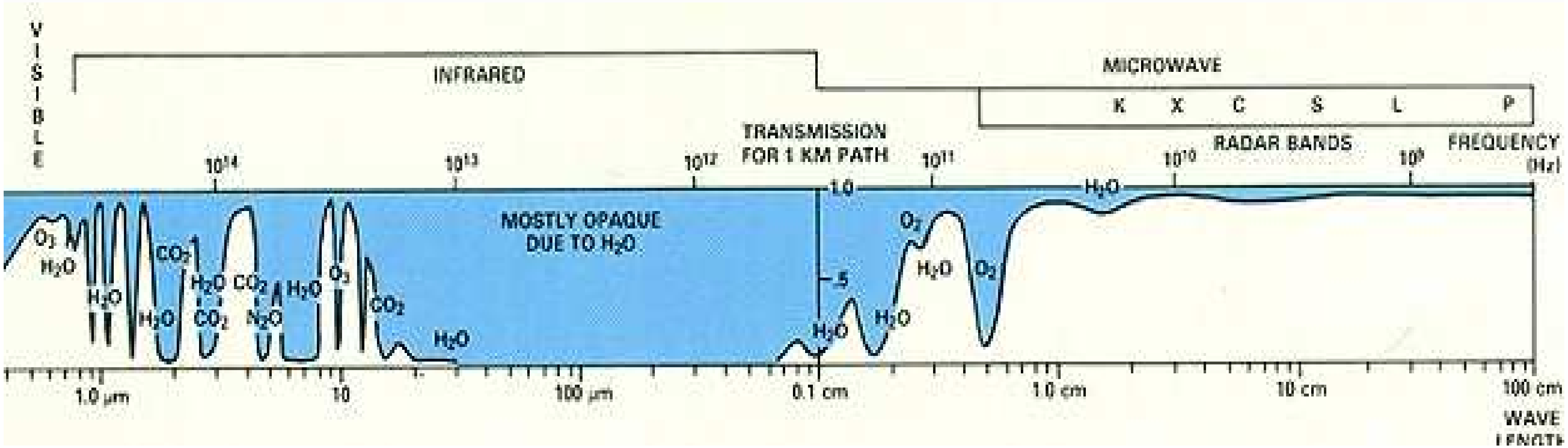
Constellation of polar orbiting or geostationary satellites

What is measured by satellite sensors ?

❑ Sensors do not measure directly atmospheric temperature and humidity, but electromagnetic radiation : brightness temperature or radiance.

❑ Depending on wave length (or frequency), information on gas concentration or physical properties (temperature or pressure or humidity) of atmosphere.

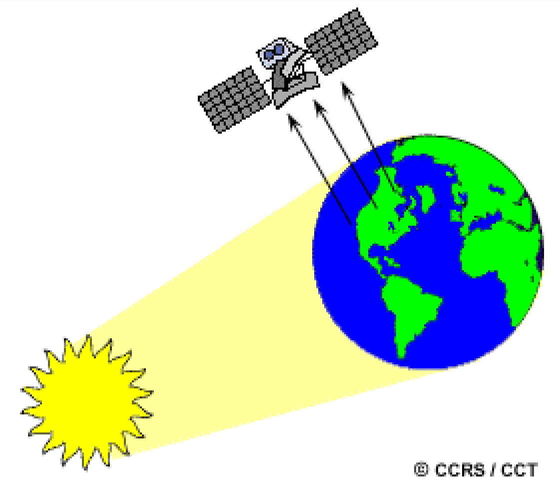
❑ Observations in atmospheric windows ➔ information on surface.



What is measured by satellite sensors ?

Passive measures

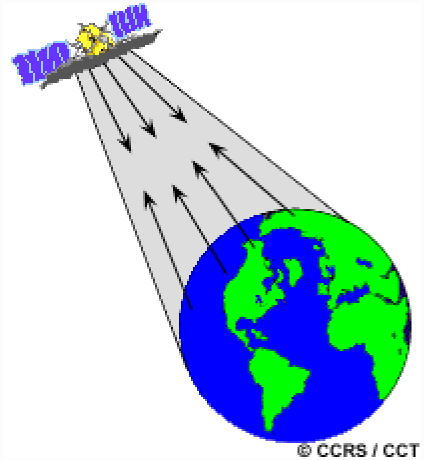
(no energy emitted from instrument)



Measures natural radiation emitted by Earth/Atmosphere from Sun origin

Active measures

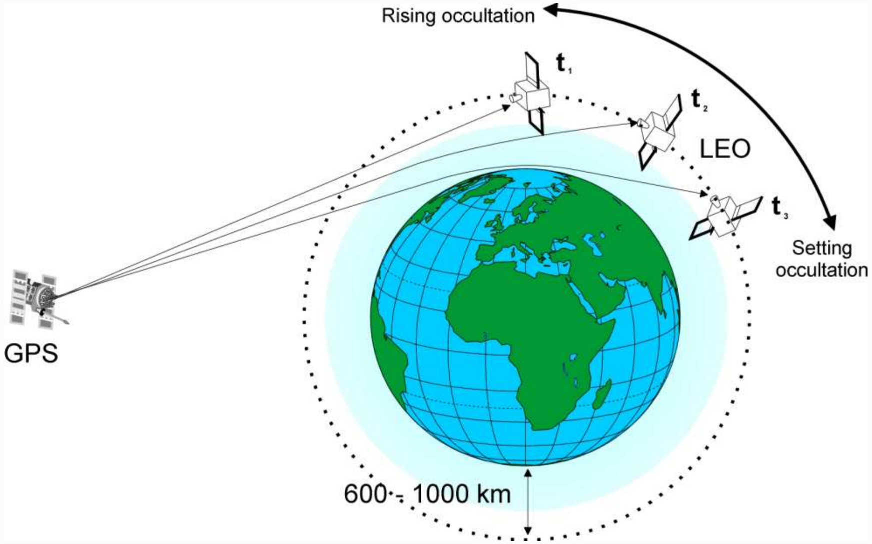
(energy emitted from instrument)



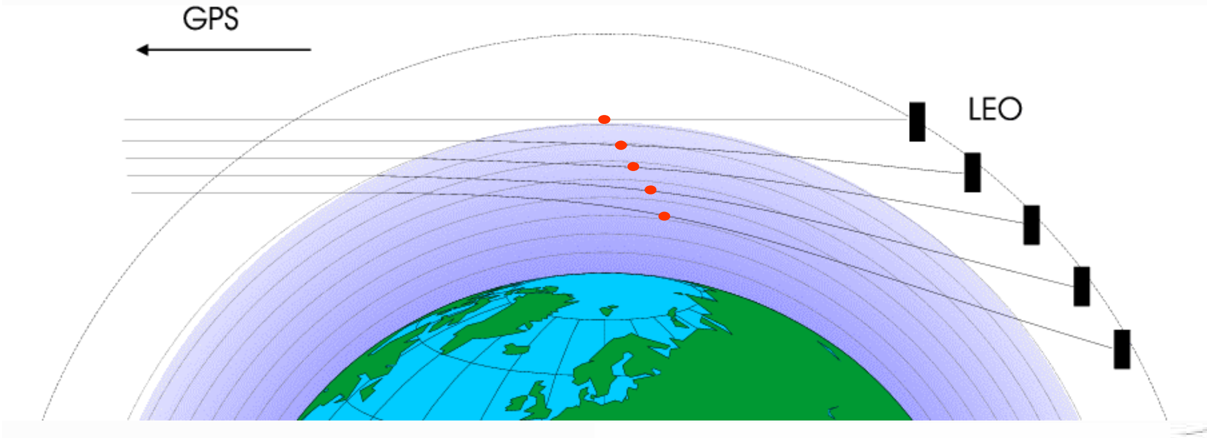
Radiation emitted by satellite and then reflected or diffused by Earth/Atmosphere

Example of active remote sensing

GPS radio occultation:



- Low-Earth Orbit satellites receive a signal from a GPS satellite.
- The signal passes through the atmosphere and gets refracted along the way.
- The magnitude of the refraction depends on temperature, moisture and pressure.
- The relative position of GPS and LEO changes over time => vertical scanning of the atmosphere.

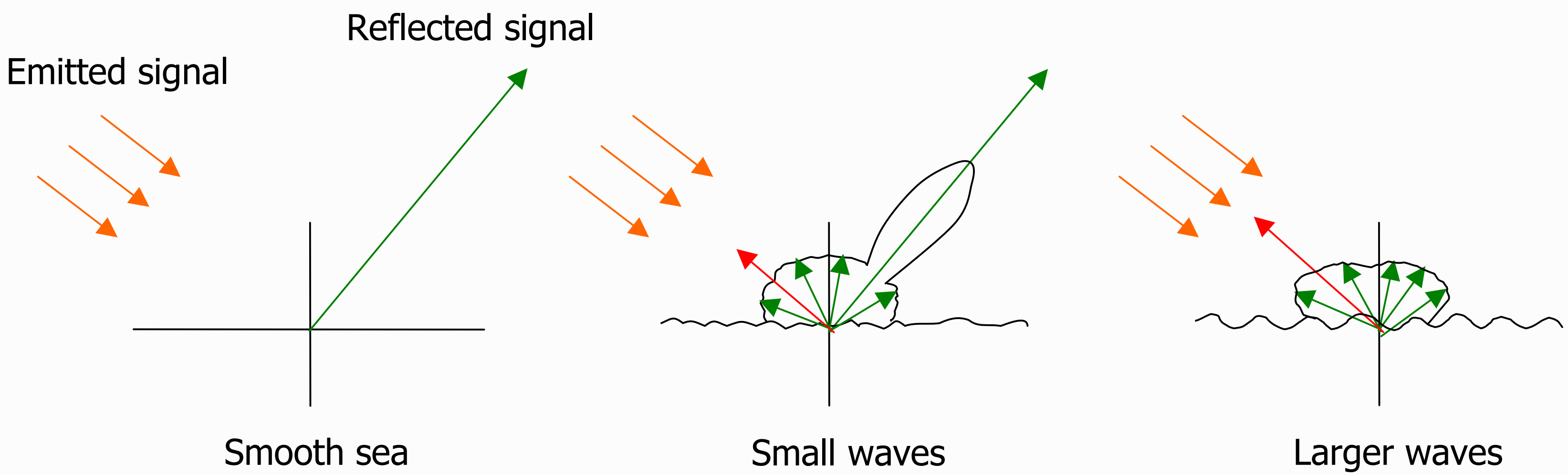


GPS stations of Météo France: Toulouse and Brest

- Propagation of GPS signal is slowed by atmosphere (dry air and water vapour).
- More than 500 GPS stations over Europe provide an estimation of Zenith Total Delay (ZTD) in real time to weather centres.
  - All weather instrument
  - High temporal resolution

Scatterometers

They send out a microwave signal towards a sea target.
The fraction of energy returned to the satellite depends on wind speed and direction.



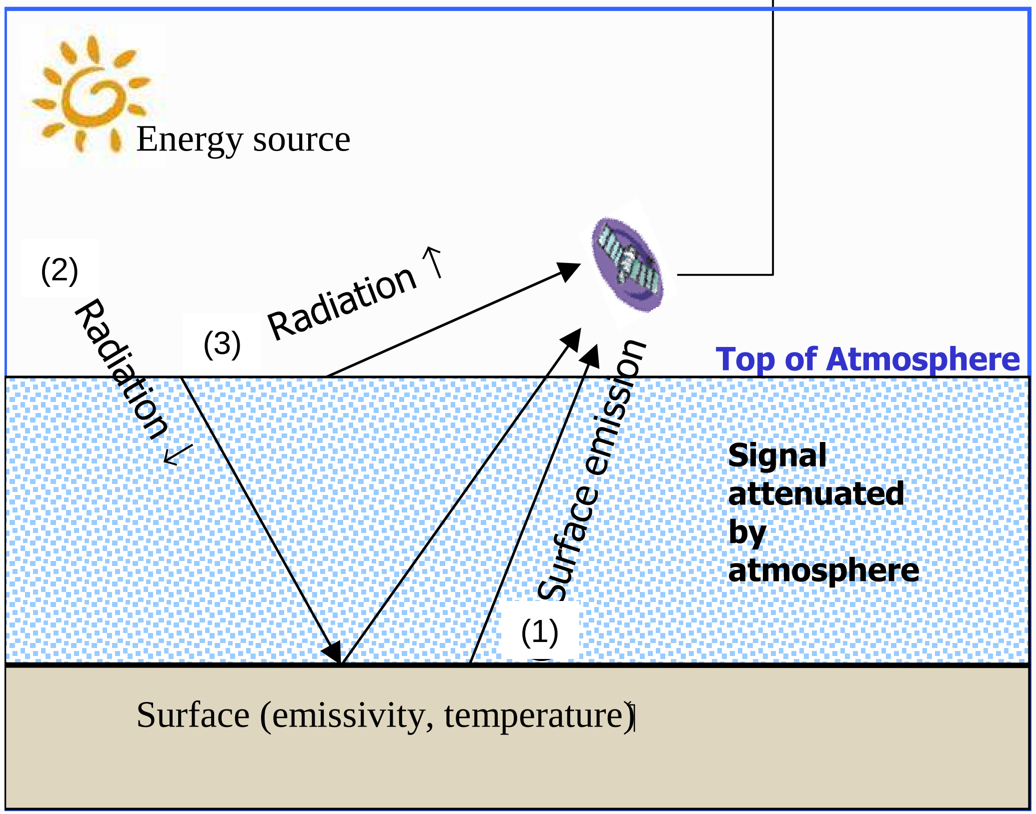
=> Measurements of near surface wind over the ocean,
through backscattering of microwave signal reflected by waves.

Passive remote sensing

Only natural sources of radiation (sun, earth...) are involved, and the sensor is a simple receiver, « passive ».

Atmosphere in Parallel Plan, no diffusion, specular surface

$$T(p,\upsilon)=\varepsilon(p,\upsilon)Ts\tau+(1-\varepsilon(p,\upsilon))\tau T(\upsilon,\downarrow)+T(\upsilon,\uparrow)$$



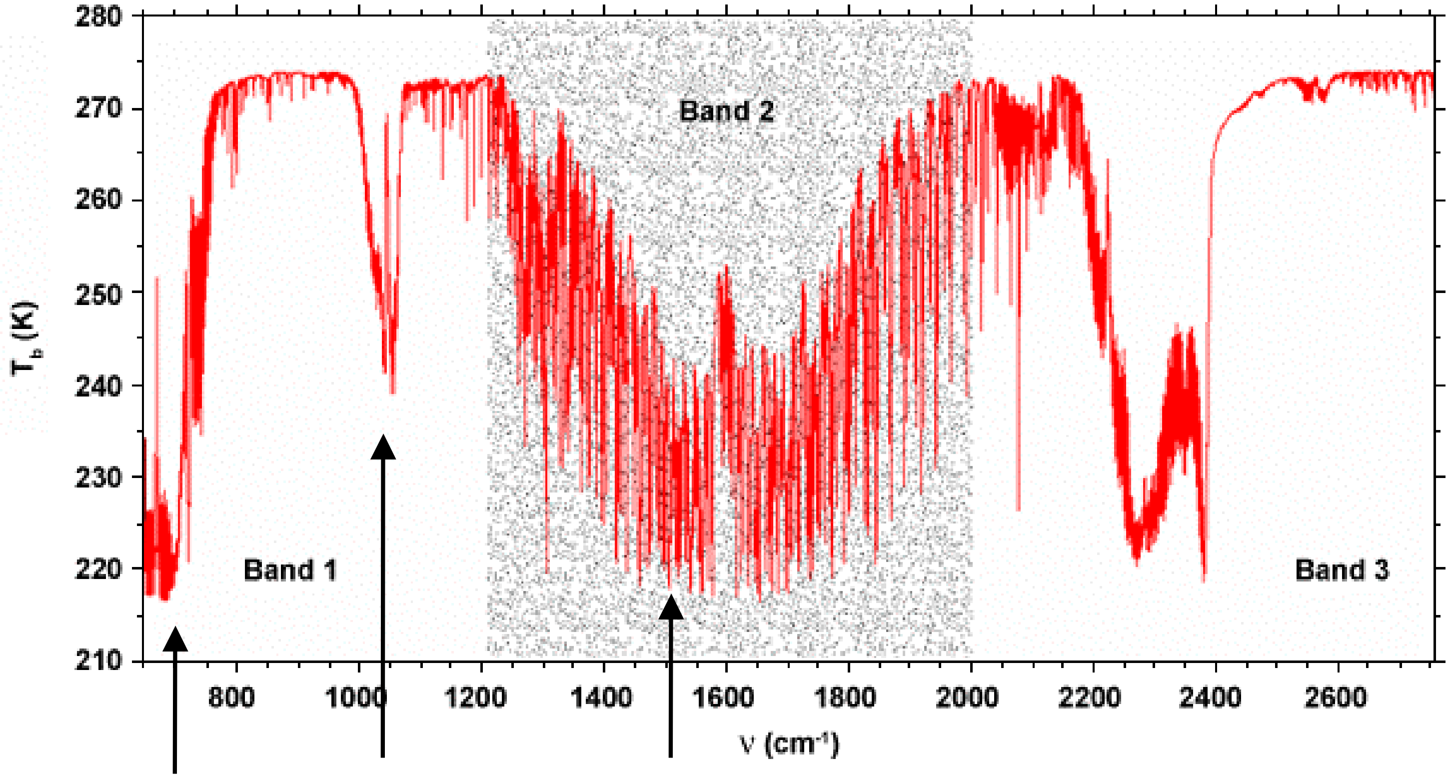
Emissivity

Radiative transfer equation dependent on T, Q :

Observation operator for satellite radiances.

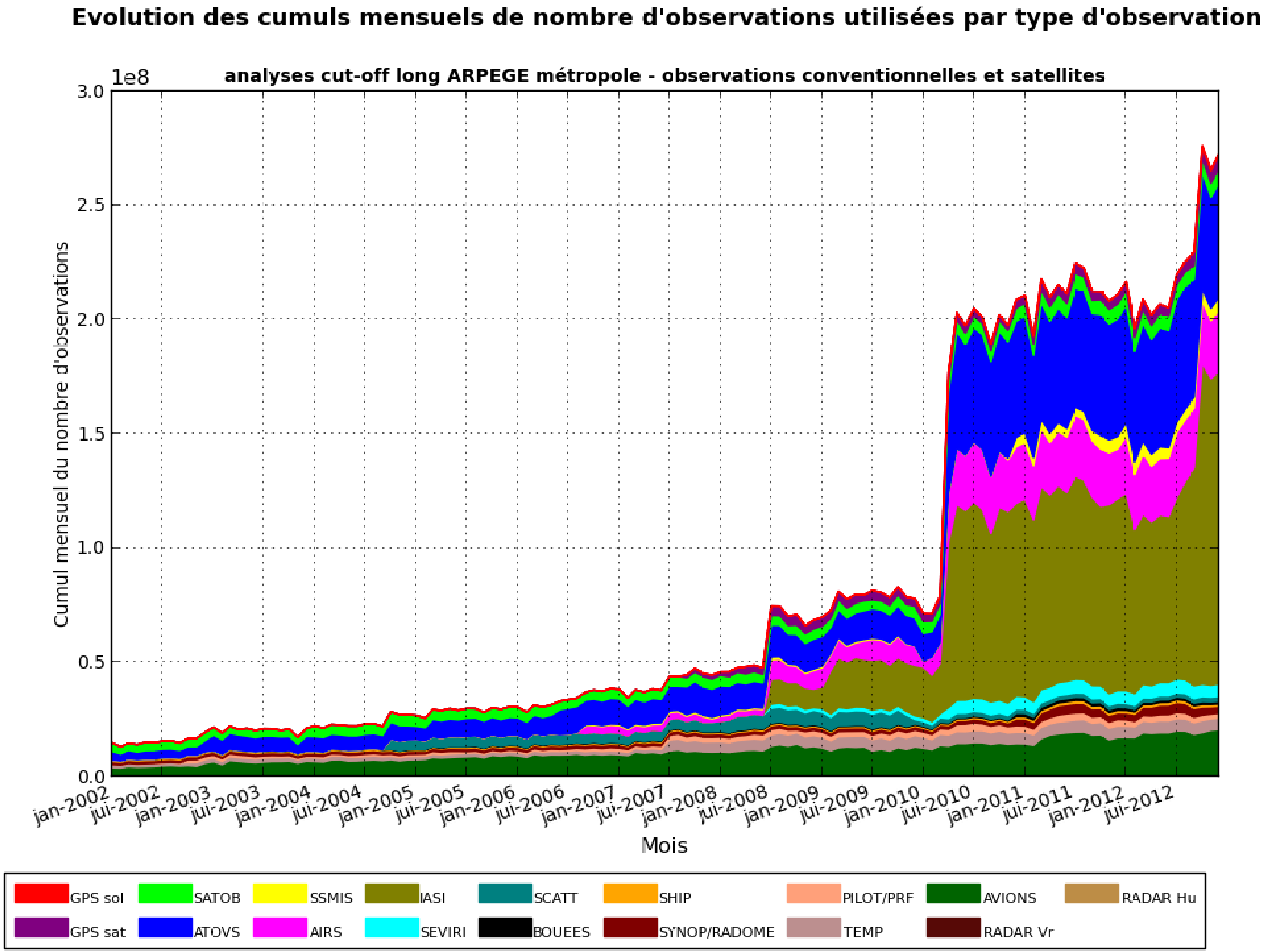
IASI, infra-red interferometer developed by CNES and EUMETSAT

IASI offers a very high spectral resolution



Temperature ozone Water vapor

Number of observations used in ARPEGE
(global DA at Météo-France)



Radar network in France

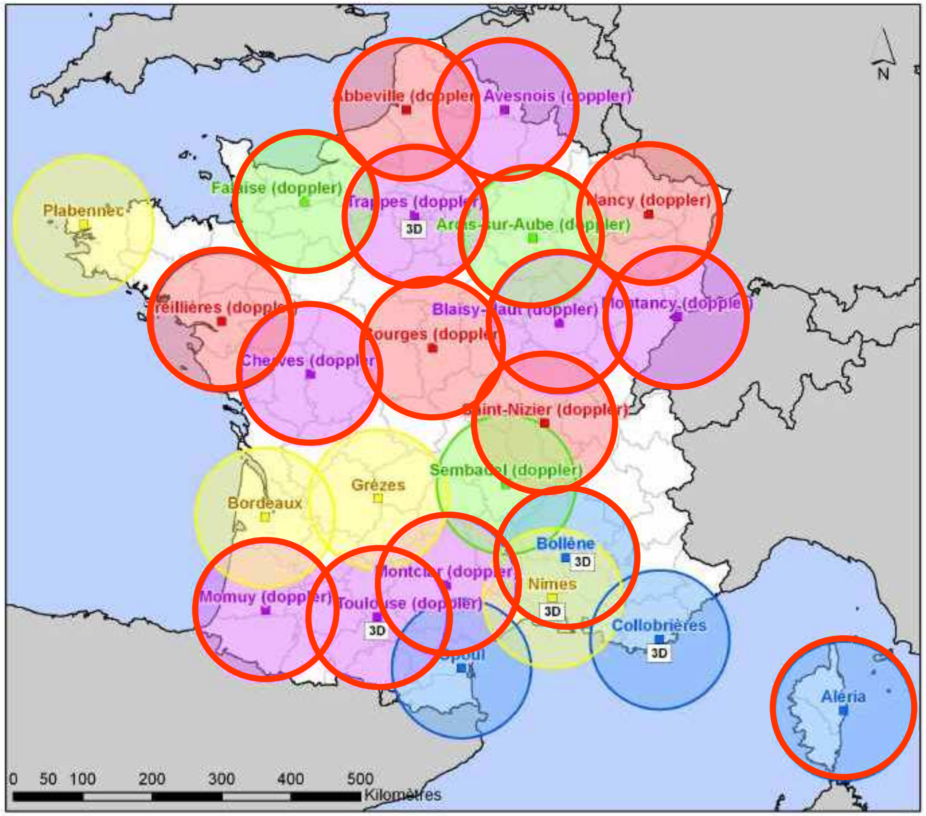
• 24 radars (17 Doppler C-Band,
every 15 minutes, at 1 km resolution).

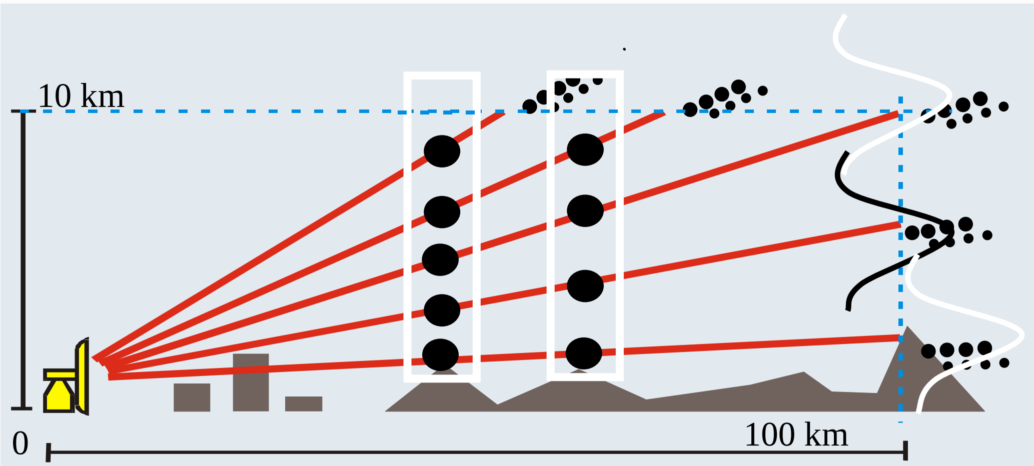
Doppler Radar

• Observations :

reflectivities Z (related to precipitation),

radial winds Vr (doppler effect : modified frequency of signal, when the target is moving => wind observation).





Observations assimilated as profiles in the model

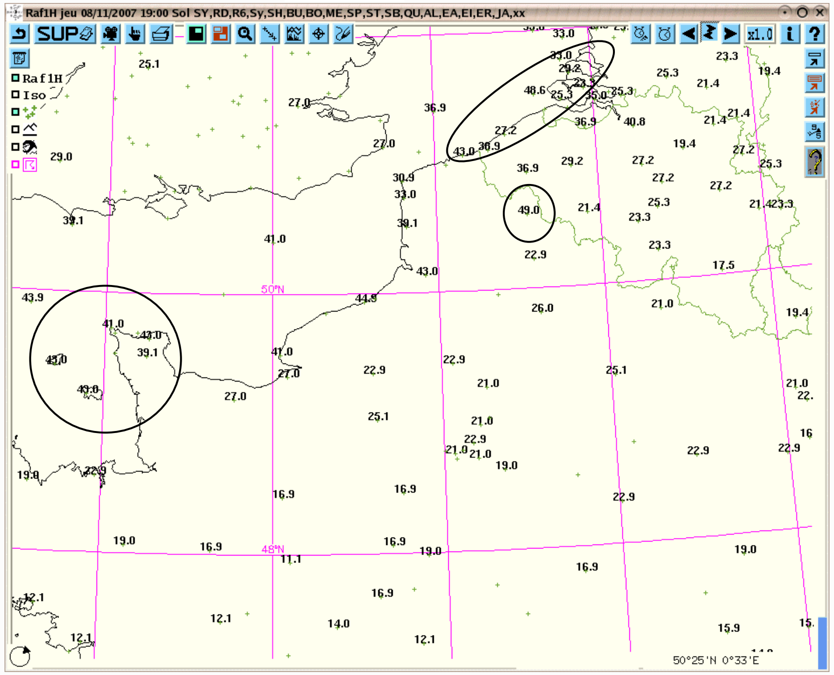
Pixel altitude is computed using a constant refractivity index along the path (effective radius approximation)

Assimilation of radar radial winds

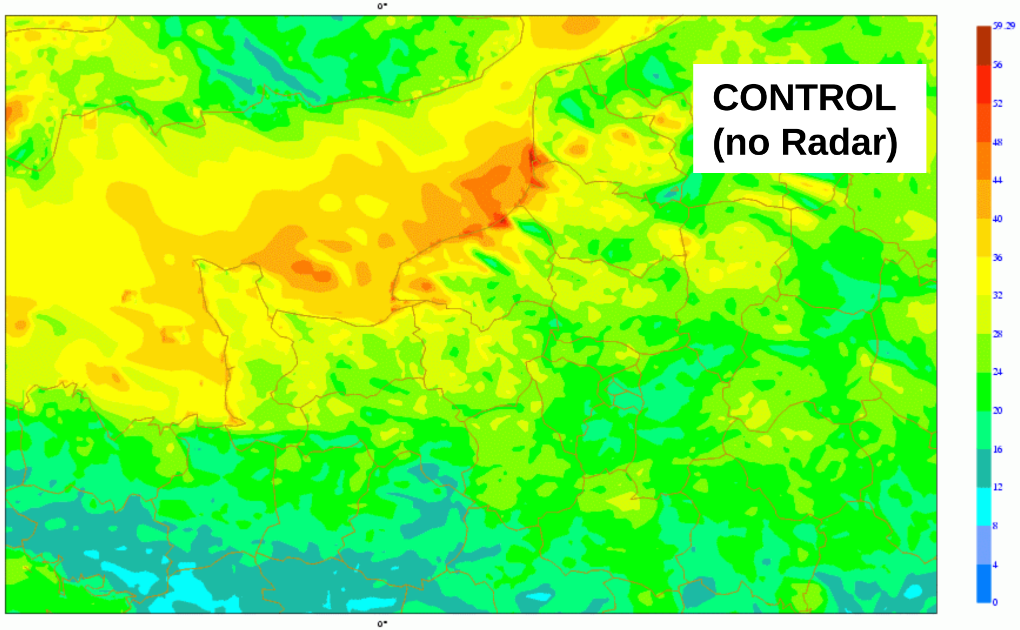
Wind gust at 10 m (kt)

Forecast +1h (19 UTC)

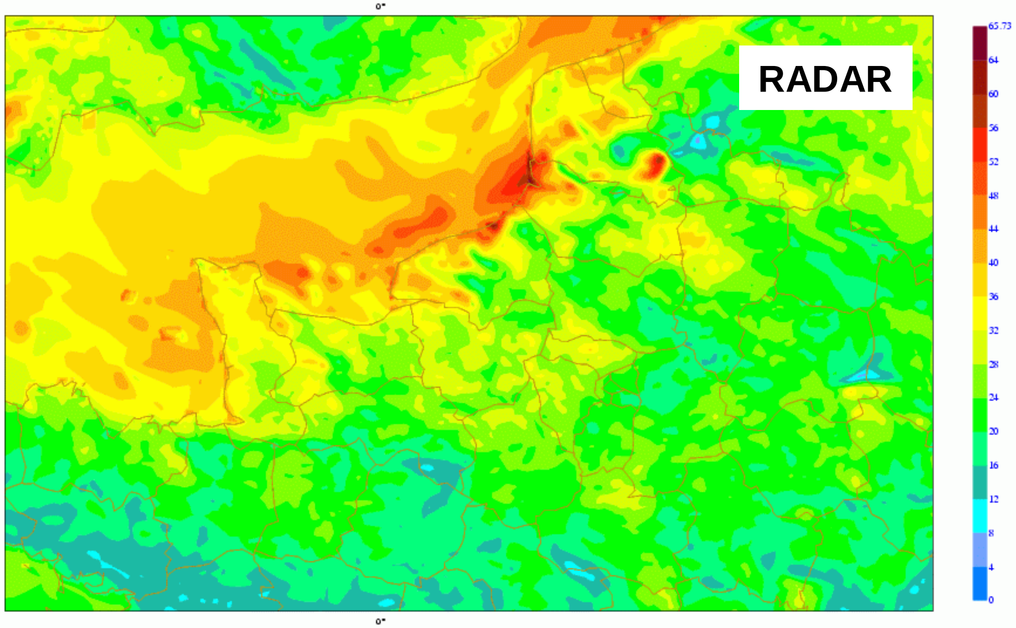
OBS



Thursday 8 November 2007 18UTC PARIS Forecast t+1 VT: Thursday 8 November 2007 19UTC 10m **

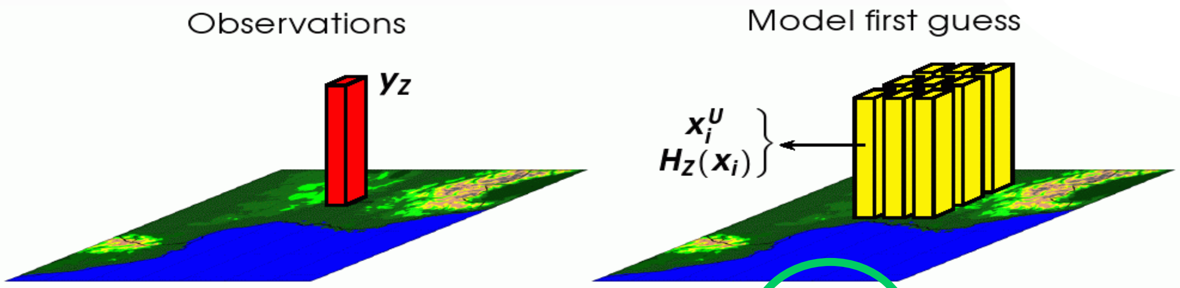


Thursday 8 November 2007 18UTC PARIS Forecast t+1 VT: Thursday 8 November 2007 19UTC 10m **



Inversion method of reflectivity profiles

Caumont, 2006: use model profiles in the neighborhood of observations



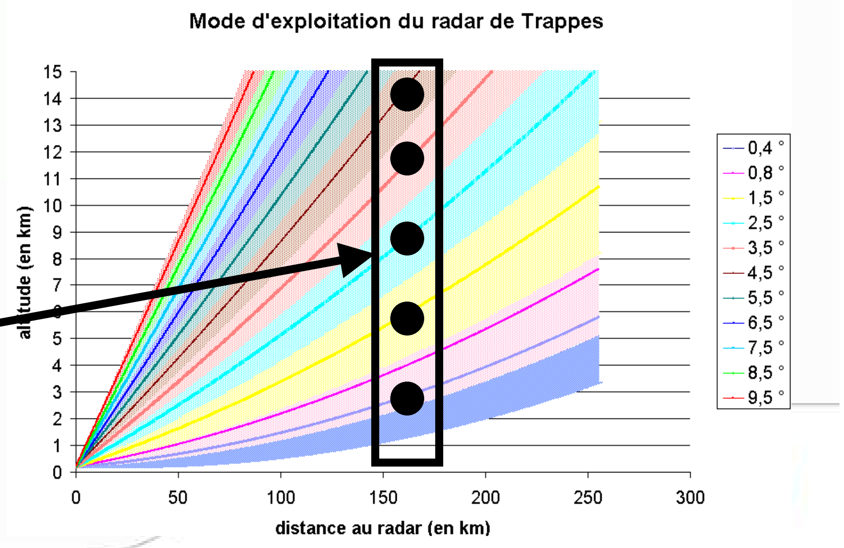
$$y_{po}^U = \sum_{i \in \text{neighbours}} x_i^U \frac{\exp\left(-\frac{1}{2}\|y_Z - H_Z(x_i)\|^2\right)}{\sum_{j \in \text{neighbours}} \exp\left(-\frac{1}{2}\|y_Z - H_Z(x_j)\|^2\right)}$$

Observation operator for reflectivities

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.

Coherence between the inverted profile and the precipitating cloud that the model is able to create

y_{po}^U



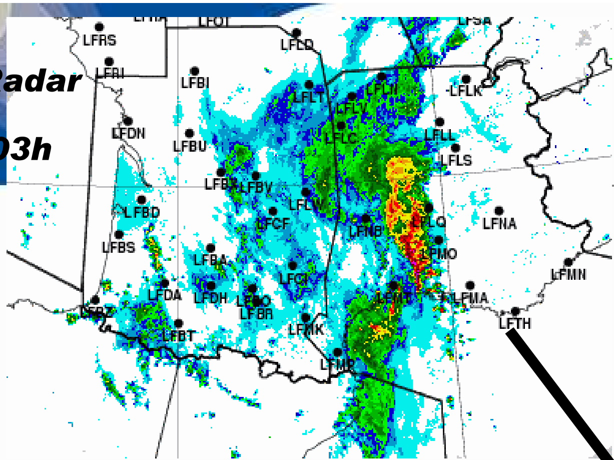
Case 7/8 october: South-East

Comparison of 3h FORECASTS
between REFL runs and CONTROL runs:

line of heavy precipitation is well analysed in REFL.

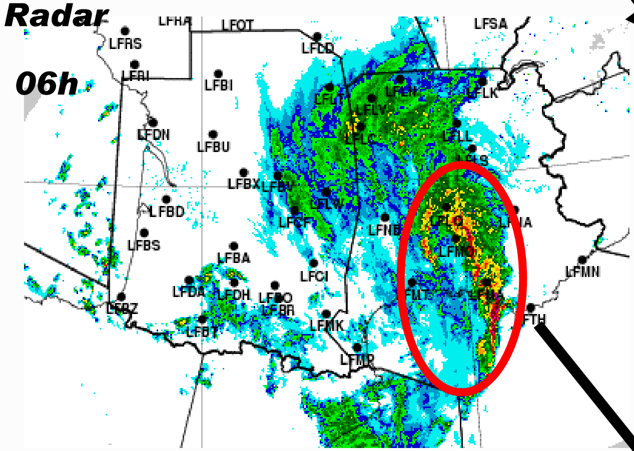
Radar

03h



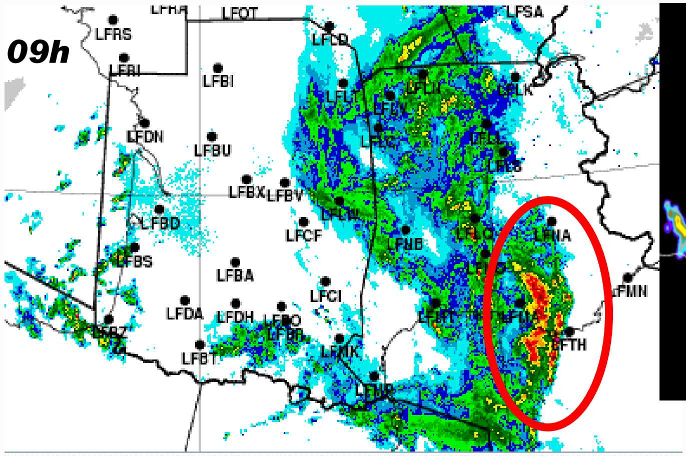
Radar

06h

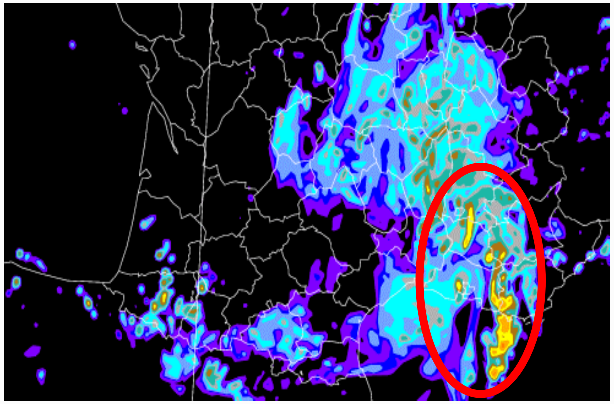


Radar

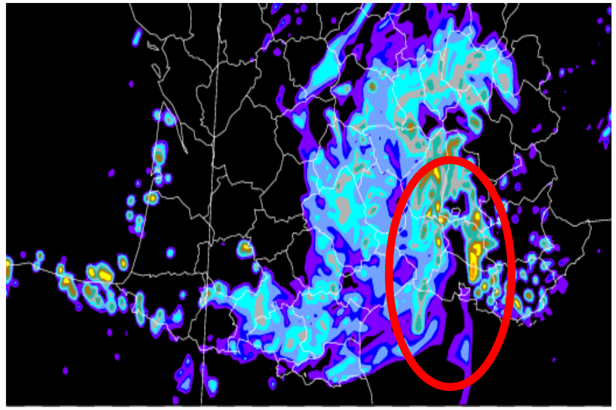
09h



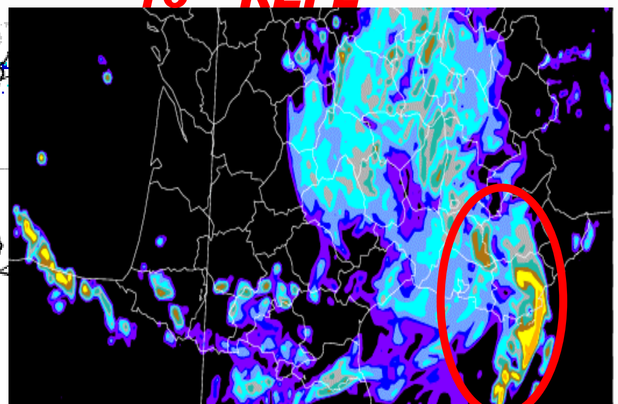
r3 – REFL



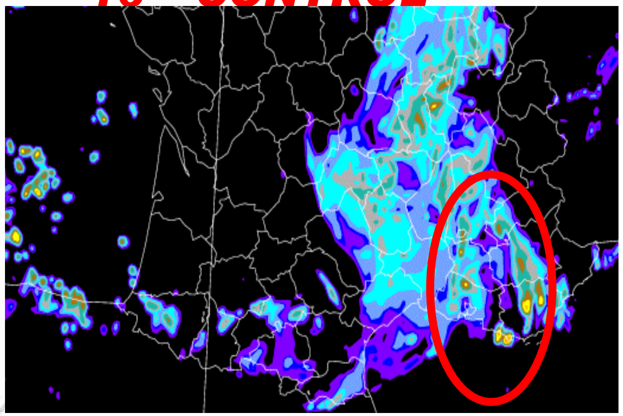
r3 – CONTROL



r6 – REFL



r6 – CONTROL



3. Error covariance estimation

Observation weights and error covariances

- BLUE analysis equation :

$$\mathbf{x^a = (I\text{-}KH)\ x^b + K\ y^o}$$

- **K** = observation weights :

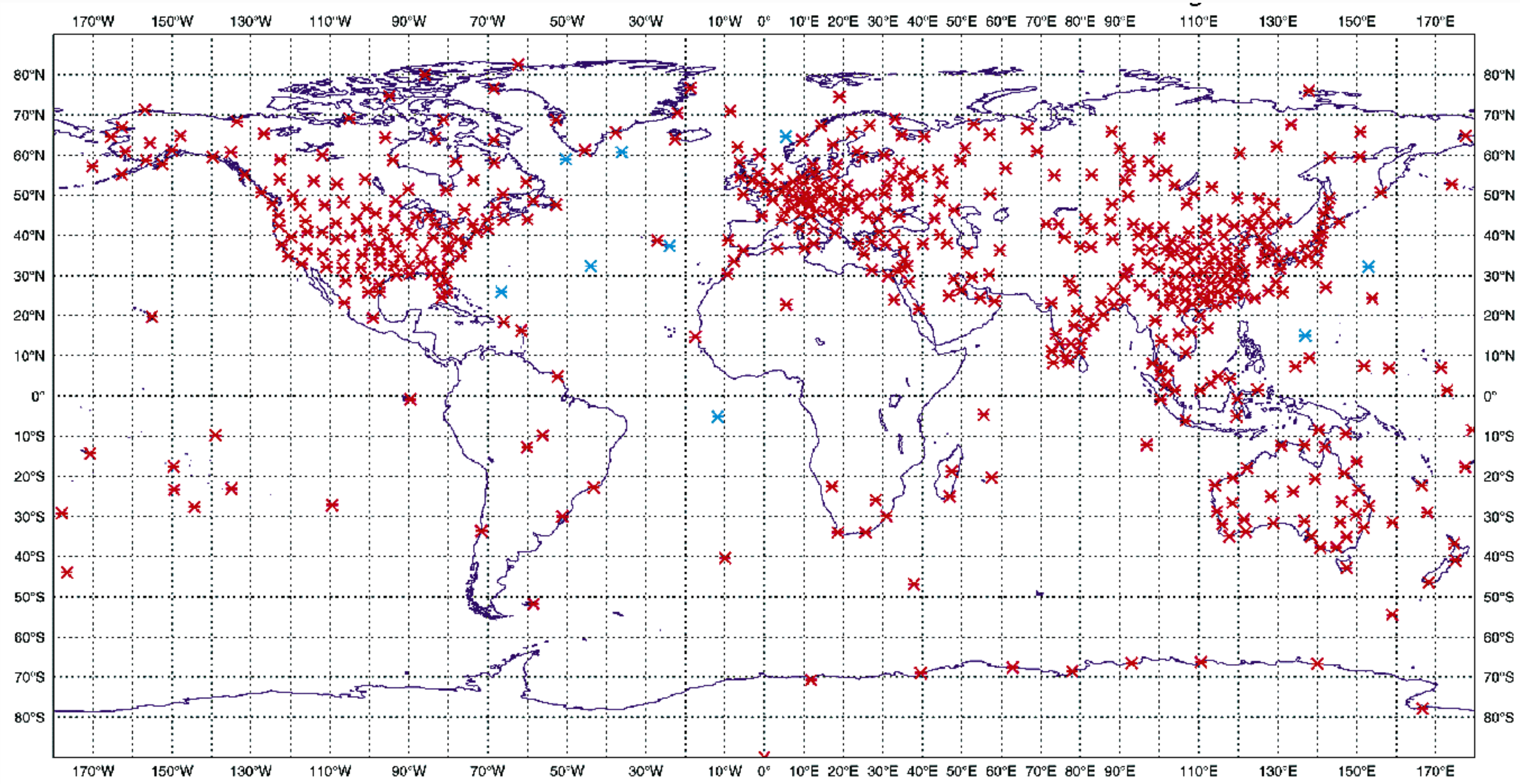
$$\mathbf{K = BH^T\ (\ HBH^T + R\ )^{-1}}$$

$\Rightarrow$ Need to estimate **B** and **R**, before specifying them.

How can we estimate error covariances ?

- The true atmospheric state is never exactly known.
- Use observation-minus-background departures
  to estimate some average features (e.g. variances, correlations)
  of **R** and **B**, using assumptions on spatial structures of errors.
- Use ensemble to simulate the error evolution and
  to estimate complex background error structures.

RADIOSONDE OBSERVATIONS



Covariances of innovations

- Innovation = observation-minus-background :

$$\mathbf{y_o - H\,x_b = y_o - H\,x_t + H\,x_t - H\,x_b}$$
$$\mathbf{= e_o - H\,e_b}$$

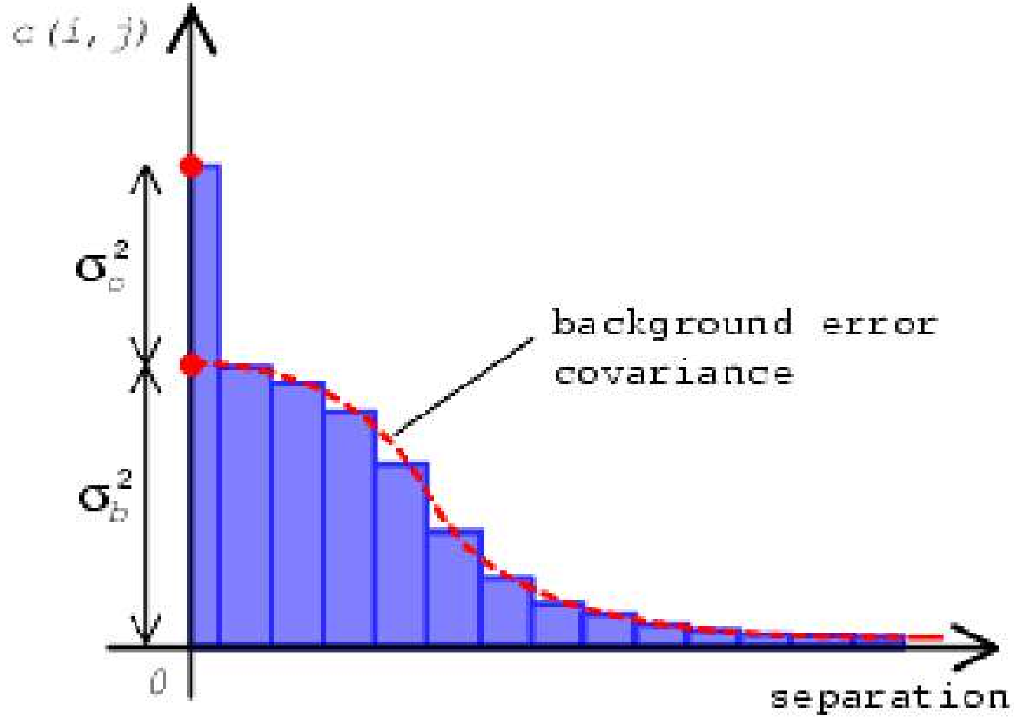
- Innovation covariances :

$$\mathbf{E[(y_o{-}Hx_b)(y_o{-}Hx_b)^T] = R + HBH^T}$$

assuming that $\mathbf{E[(e_o)(He_b)^T]=0.}$

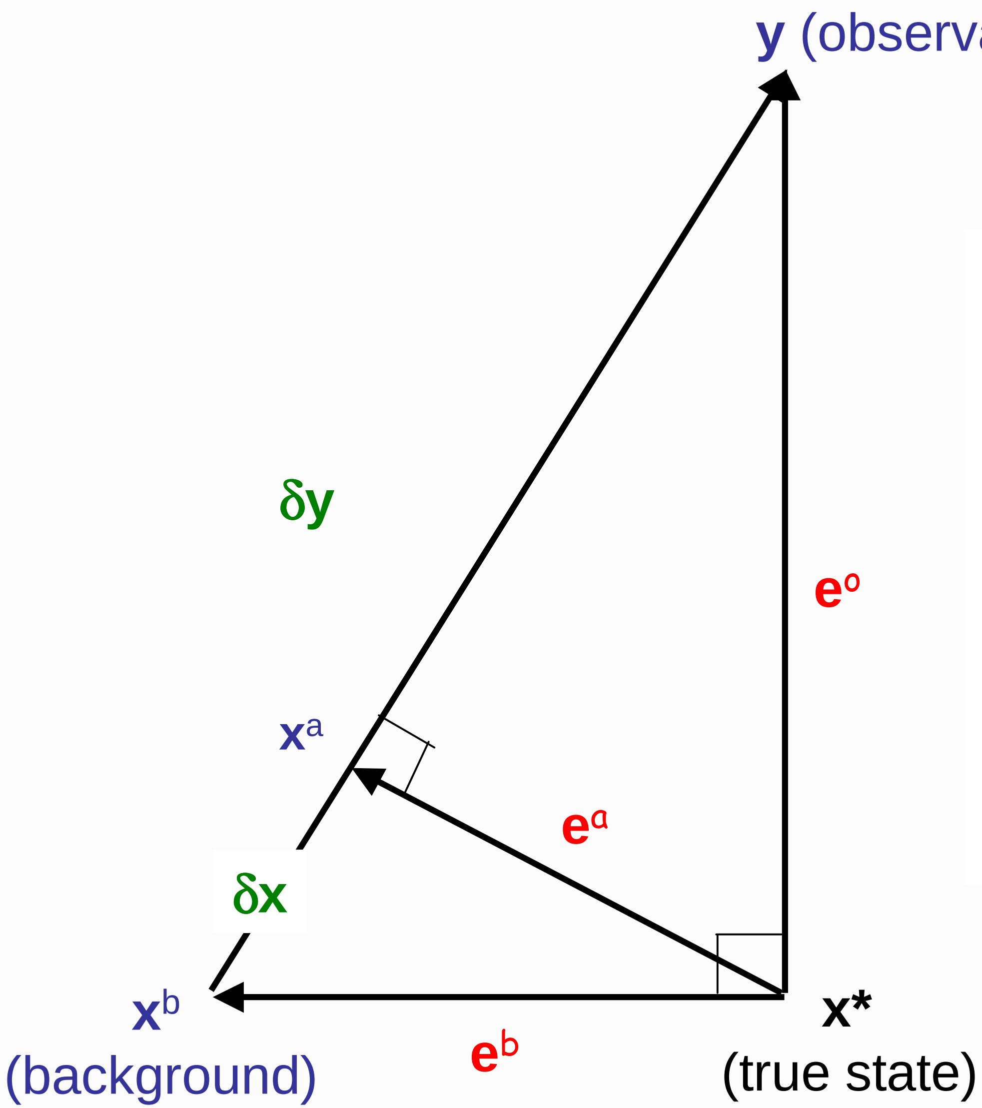
(e.g. Hollingsworth and Lönnberg 1986)

Hollingsworth and Lönnberg method



(From Bouttier and Courtier, ECMWF)

Covariances of analysis residuals



$\delta \mathbf{y} = \mathbf{y} - H(\mathbf{x}^b)$ (innovation)

$\mathbf{H}\,\delta \mathbf{x} = H(\mathbf{x}^b) - H(\mathbf{x}^a)$ (increment)

$E[\mathbf{H}\,\delta \mathbf{x}\,\delta \mathbf{y}^T] = \mathbf{HBH}^T$

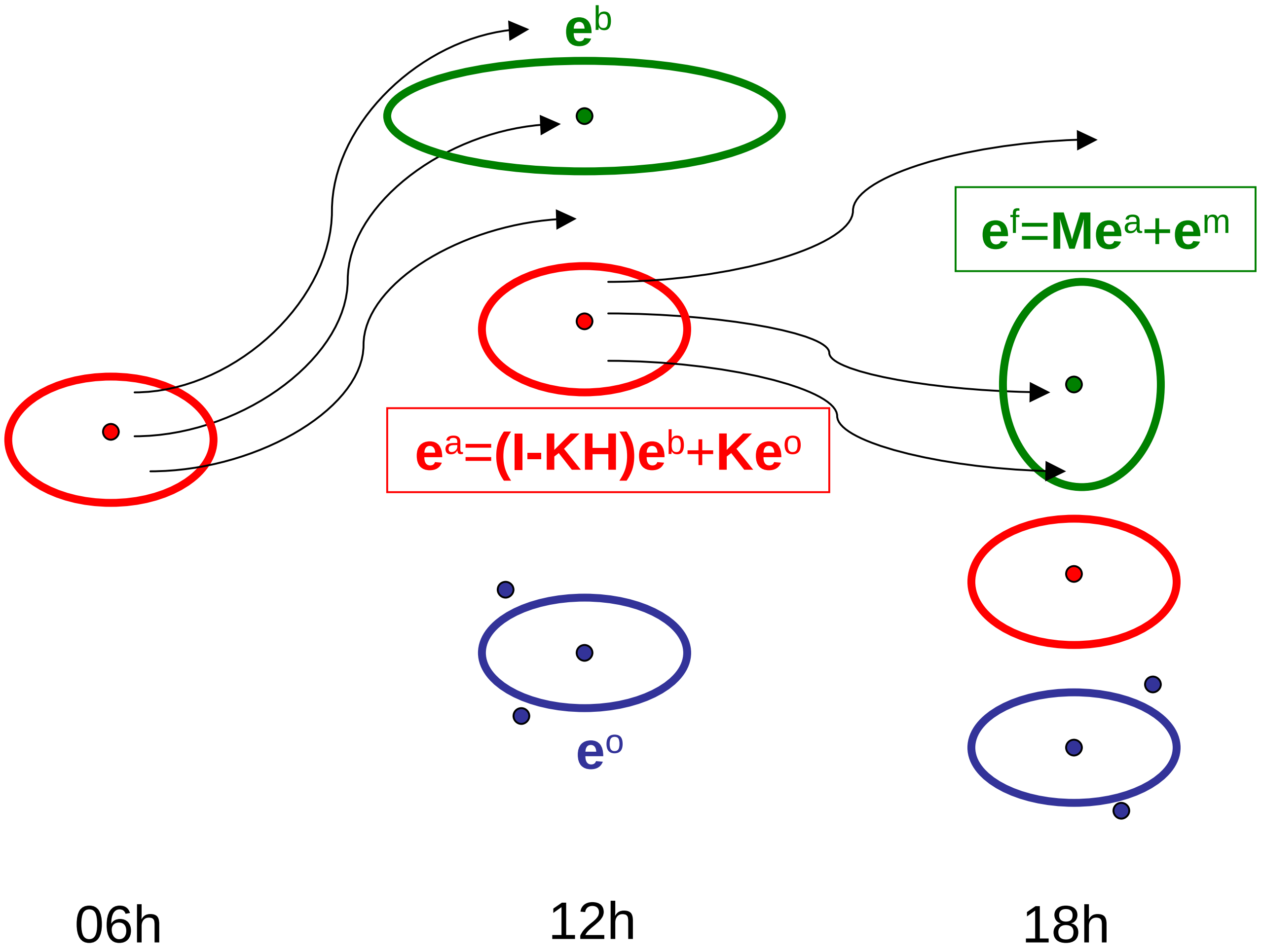
$E[(\mathbf{y} - H(\mathbf{x}^a))\,\delta \mathbf{y}^T] = \mathbf{R}$

(Desroziers, Berre et al 2005)

Innovation method : properties

- Provides estimates in observation space only.
- A good quality data dense network is needed.
- Assumption that observation errors are « white ».
- An objective source of information on **B** and **R**.

Ensemble Data Assimilation : simulation of error cycling

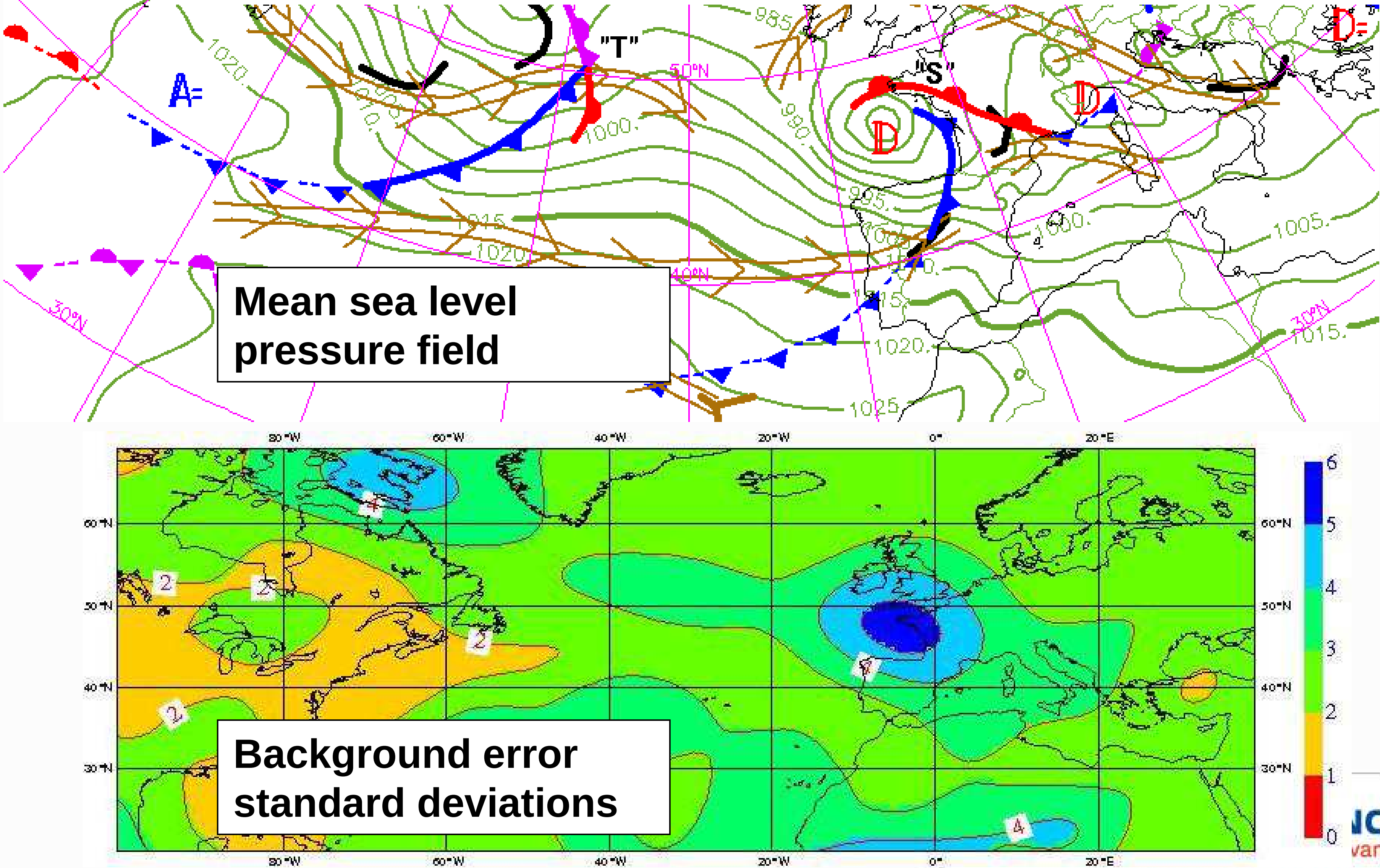


(e.g. Houtekamer et al 1996, Fisher 2003, Berre et al 2006)

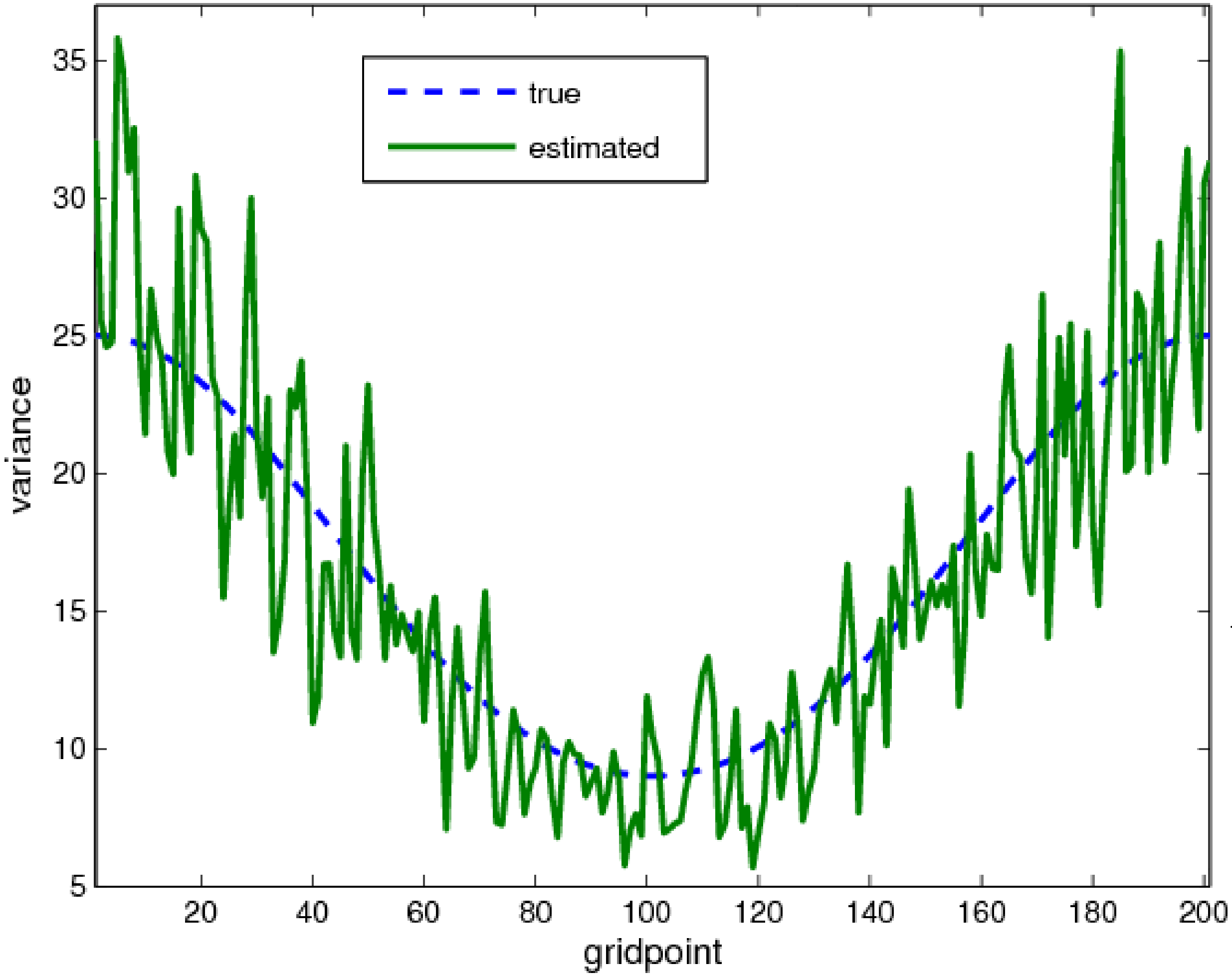
Estimation of background error variances from ensemble spread

$$\mathrm{Var}(e_b) = 1/(N-1)\, \Sigma_n \left[\, x'_b(n) - \overline{x'_b}\, \right]^2$$

Connexion between large errors and intense weather
(Klaus storm, 24/01/2009, 00/03 UTC)



Spatial structure of sampling noise for variances



$\varepsilon_b = \mathbf{B}^{1/2}\,\eta$

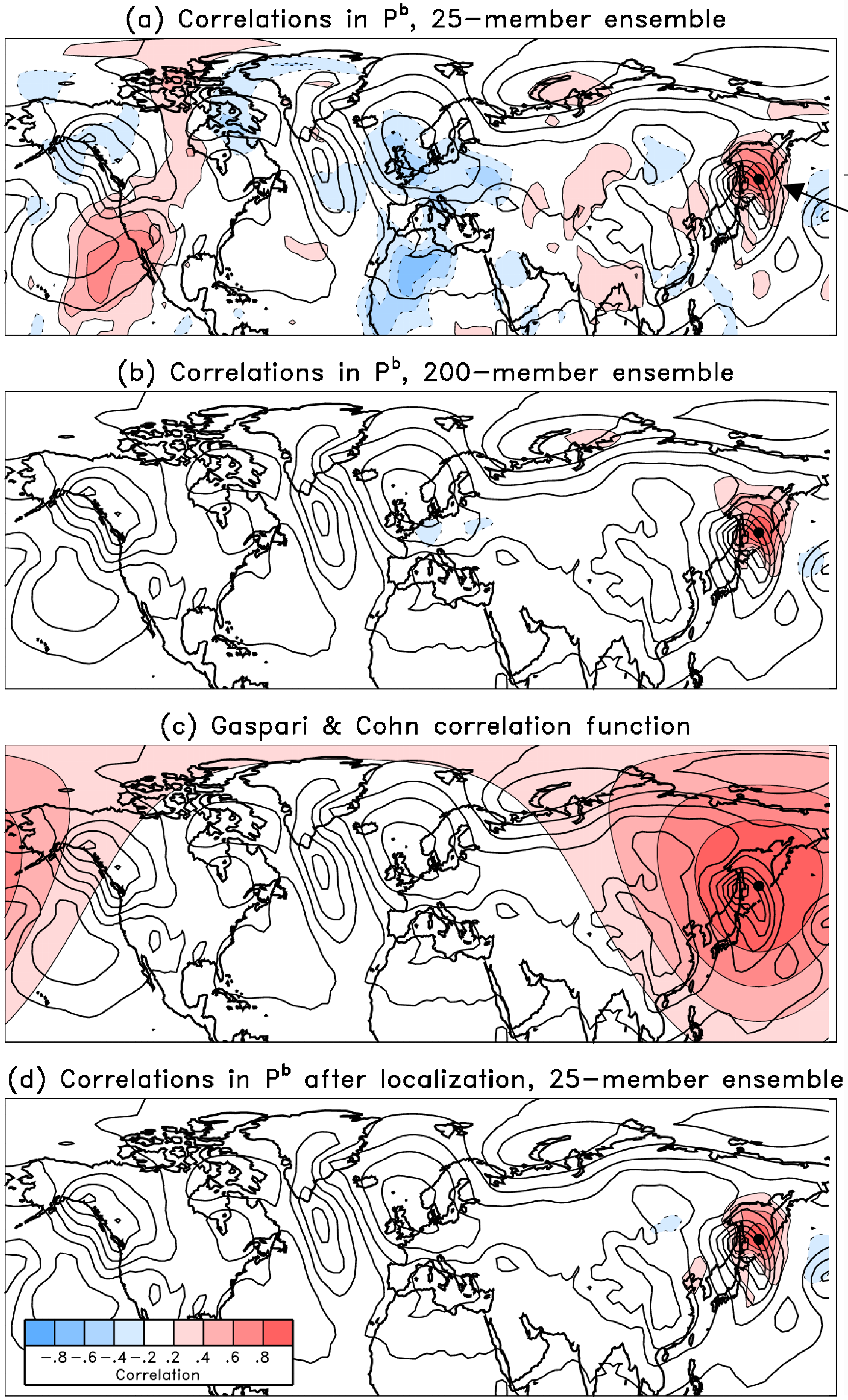
$\eta \sim \mathcal{N}(0,\mathbf{I})$

N = 50 members

L(ε_b) = 200 km

$\overline{V^e\,(V^e)^T} = 2/(N-1)\ \mathbf{B}^* \circ \mathbf{B}^*$

⇒ Spatial filtering in order to extract large scale signal, and remove small scale sampling noise. (e.g. Raynaud et al 2009)

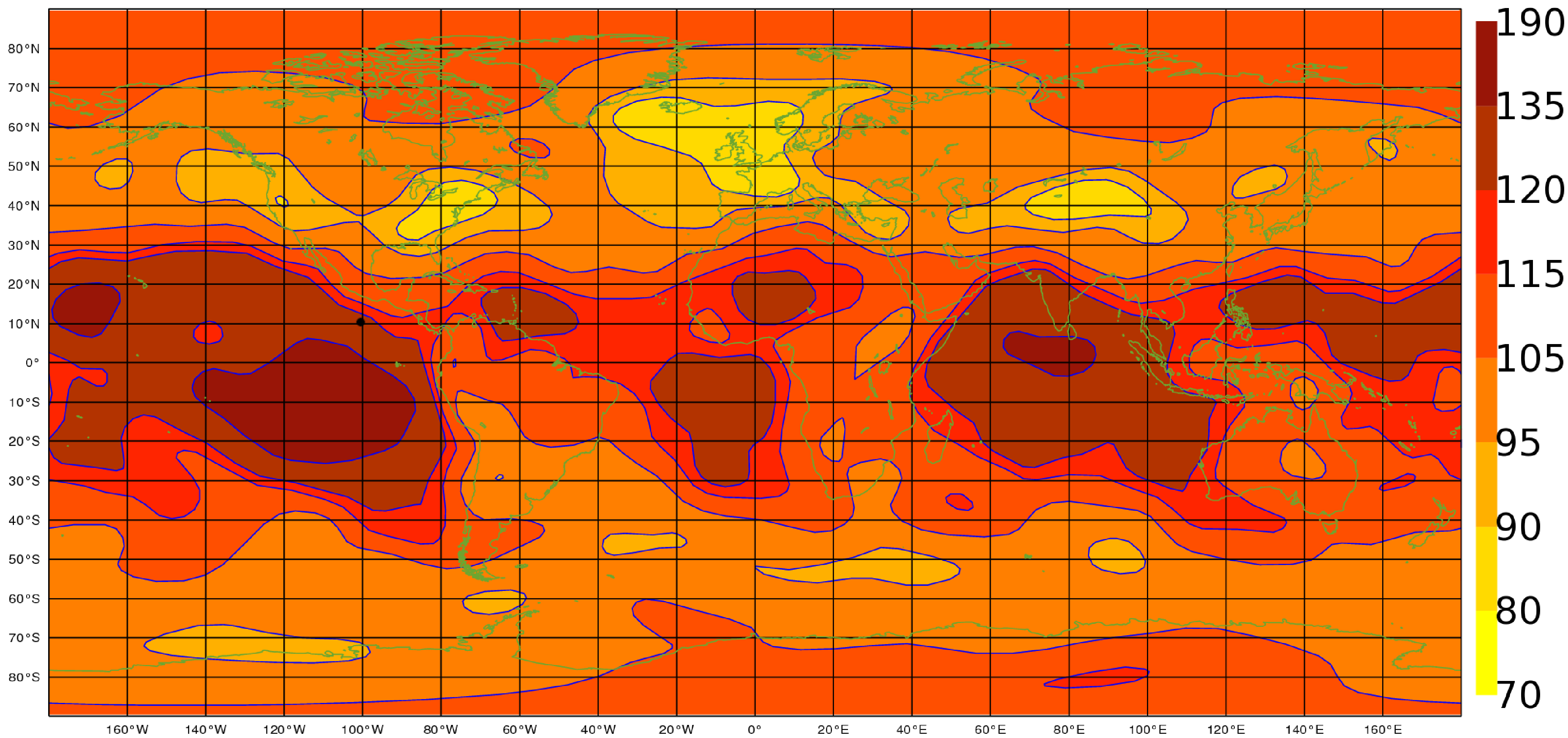


Observation location

Schur filtering of long-distance correlations

from Hamill, Chapter 6 of "*Predictability of Weather and Climate*"

Flow-dependent background error correlations using EnDA and wavelets



Wavelet-implied horizontal length-scales (in km), for wind near 500 hPa, averaged over a 4-day period.

(e.g. Fisher 2003, Varella et al 2011)

Conclusions

- Data Assimilation (DA) is vital for weather forecasting (NWP).
- Observations are very diverse in type, density and quality.
- 4D-Var for temporal and non linear aspects.
- Ensemble DA methods for error simulation and covariance estimation.
- Sampling noise issues and filtering techniques.
- Observation-background departures for validation of error covariances, and for estimation of model errors.

Some references

- Desroziers, G., Berre, L., Chapnik, B. and Poli, P. (2005), Diagnosis of observation, background and analysis-error statistics in observation space. Q.J.R. Meteorol. Soc., 131: 3385–3396.
- Fisher, M., 2003: Background error covariance modeling. Proc. ECMWF Seminar on "Recent Developments in Data Assimilation for Atmosphere and Ocean", 8-12 Sept 2003, Reading, U.K., 45-63.
- Hollingsworth, A. and Lönnberg, P., 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. Tellus, 38A, 111-136
- Houtekamer, P. L., Louis Lefaivre, Jacques Derome, Harold Ritchie, Herschel L. Mitchell, 1996: A System Simulation Approach to Ensemble Prediction. Mon. Wea. Rev., 124, 1225–1242.
- Houtekamer, P. L., Herschel L. Mitchell, Xingxiu Deng, 2009: Model Error Representation in an Operational Ensemble Kalman Filter. Mon. Wea. Rev., 137, 2126–2143.
- Rabier et al 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: Experimental results with simplified physics. Q. J. R. Meteorol. Soc., 126, 1143–1170.
- Talagrand, O. and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. Quart. J. Roy. Meteor. Soc., 113, 1311-1328.
- Berre, L., Ştefănescu, S., Belo Pereira, M.. The representation of the analysis effect in three error simulation techniques. Tellus A, 58A, pp 196-209.

Thank you

for your attention