



## Data assimilation in meteorology

Loïk Berre (CNRM/GMAP/ASSIM) Météo-France

### Plan of the talk

- Numerical Weather Prediction (NWP),
   Data Assimilation (DA)
- Observations (in-situ and remote sensing)
- Error covariances : estimation and modelling

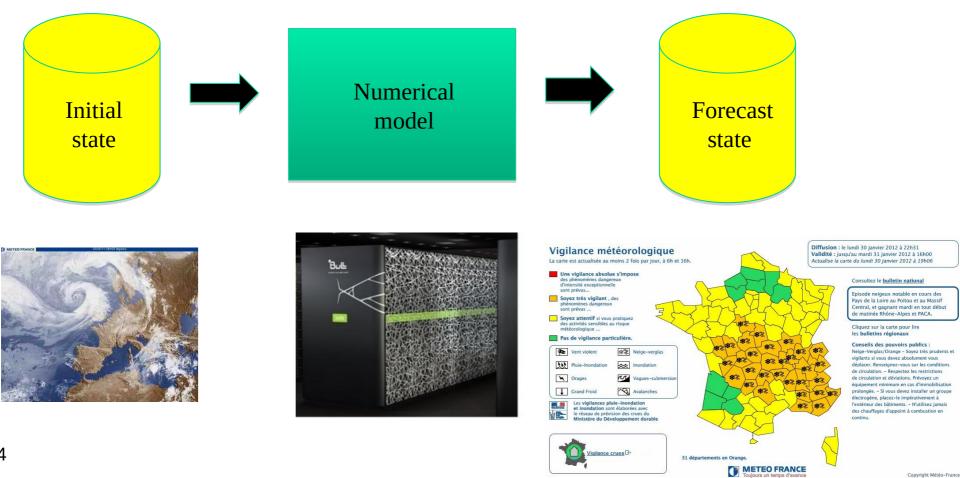




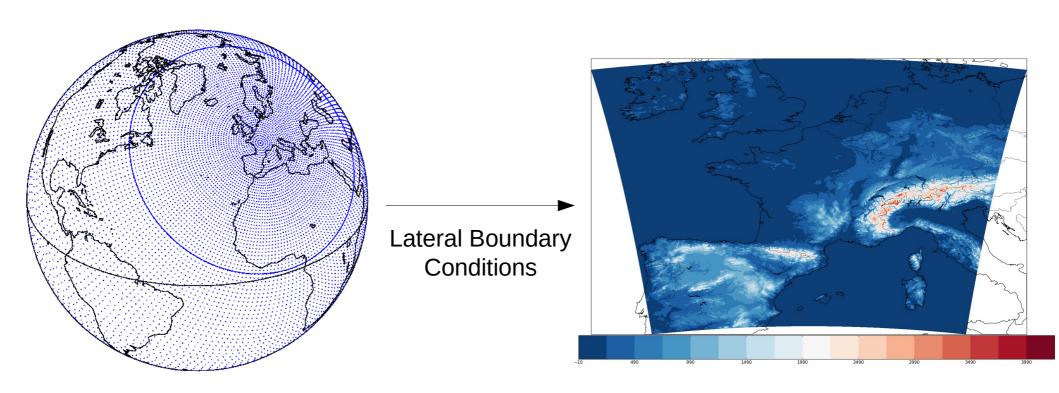
# 1. Numerical Weather Prediction and Data Assimilation

### **Numerical Weather Prediction**

Numerical resolution of fluid mechanics equations (computer code), to **forecast the atmospheric evolution** from an **estimated initial state** (which is called the « analysis »).



# NWP models at Météo-France (in collaboration with e.g. ECMWF)

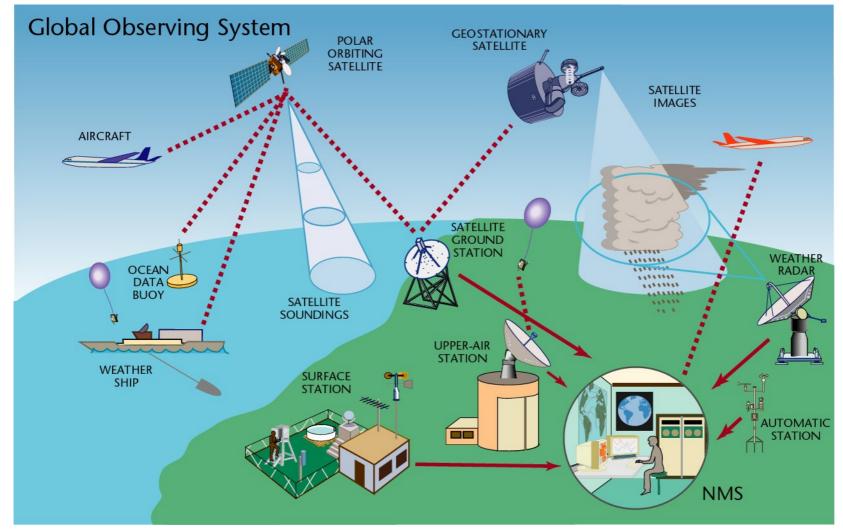


ARPEGE (5 km - 30 km) 10<sup>9</sup> model variables

AROME (1.3 km) 1,4 x 10<sup>9</sup> model variables

Equations of dynamics and physical parametrizations (radiation, shallow convection, ...) to predict the evolution of temperature, wind, humidity, etc.

### Data which are assimilated in NWP models



ARPEGE 109 model variables

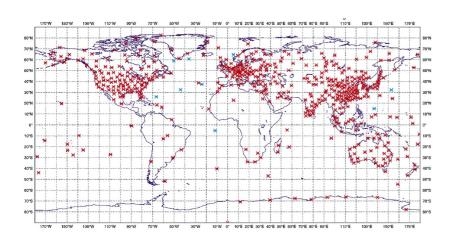
5 x 10<sup>6</sup> observations / 6h 90 % satellite Computation time : 40 min (over 6h window) AROME 1,4 x 10<sup>9</sup> model variables

2 x 10<sup>5</sup> observations / 6h Up to 75 % radar, 10 % satellite Computation time : 7 min (over 1h window)

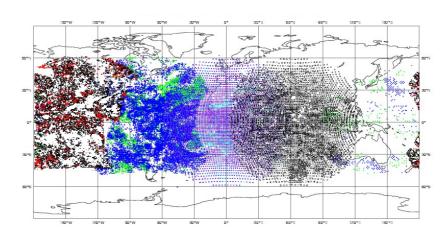


## **Spatial coverage and density of observations**

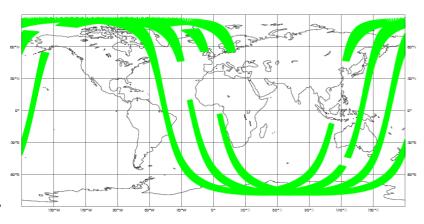
#### RADIOSONDE DATA



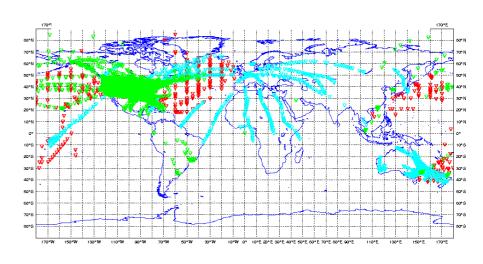
#### **GEOSAT. WINDS**



**SCATTEROMETER** 

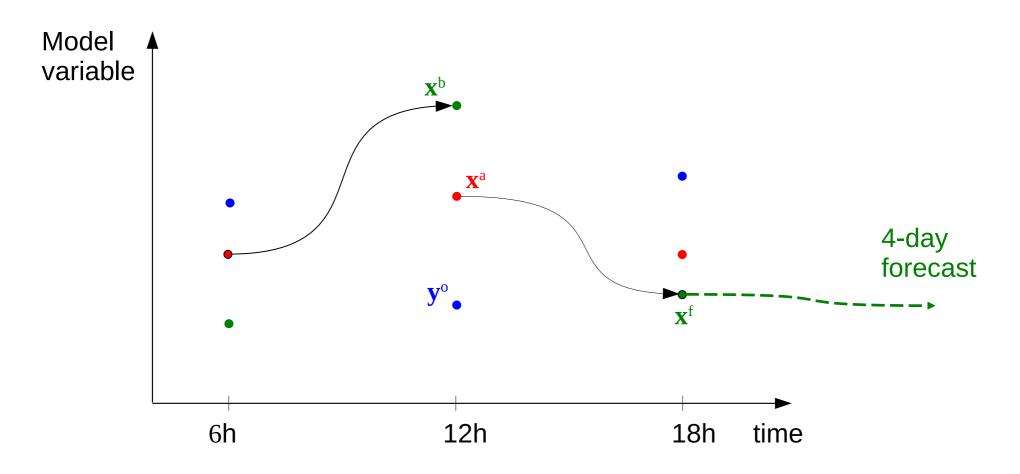


#### AIRCRAFT DATA



Page 7

## Sequential aspect of Data Assimilation: temporal succession of analysis and forecast steps

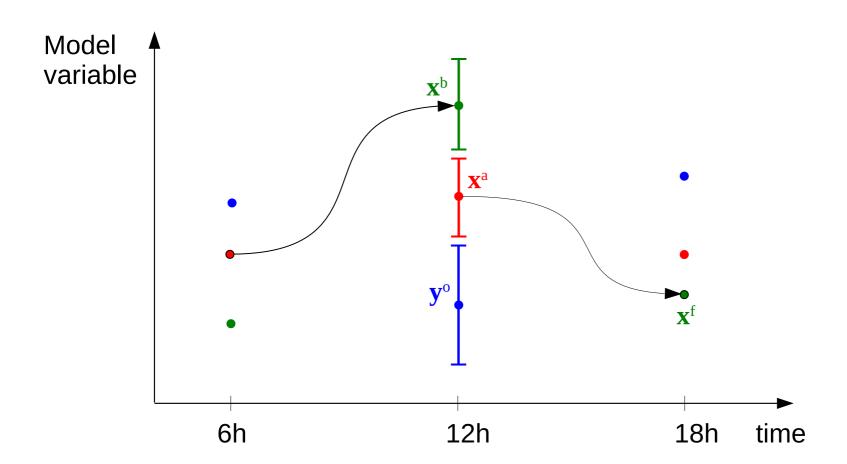


=> Every 6h, the model state is **updated** with new observations (**analysis step**) and then **propagated** by the model (**forecast step**).

The resulting 6h forecast state is called the **background**; it contains information provided by previous observations.



## Uncertainties in observations, background, analysis



Uncertainties are often measured by error variances (ex : accurate observation ↔ small observation error variance).



# Analysis equation: definition & role of observation operator *H*

- BLUE formalism :  $\mathbf{x}^{a} = \mathbf{x}^{b} + \mathbf{K} (\mathbf{y}^{o} H[\mathbf{x}^{b}])$
- Note that observation locations can differ from model gridpoint locations : some spatial interpolation is required for making a comparison.
- Note also that observed variables can be « exotic », such as satellite radiances, which are not directly represented in the NWP model!
- In order to compare observations ( $\mathbf{y}^{\circ}$ ) with the background ( $\mathbf{x}^{\circ}$ ), first step is to apply

*H* = non linear **observation operator** 

= conversion of model variables into observed variables :  $\mathbf{y} = H[\mathbf{x}]$ 

This operator H can include, for instance:

- spatial interpolation : from model gridpoints to obs locations (ex: for radiosondes);
- radiative transfer: from model temperature to simulated satellite radiances;
- **projection to observation time** (using the NWP model): for observations available at different times within 6h DA window.



## Analysis equation : definition & role of gain matrix K

- BLUE formalism:  $\mathbf{x}^a = \mathbf{x}^b + \delta \mathbf{x} = \mathbf{x}^b + \mathbf{K} (\mathbf{y}^o H[\mathbf{x}^b])$ where  $\delta \mathbf{x} = \mathbf{x}^a - \mathbf{x}^b$  is the analysis increment.
- Departures need to be filtered and propagated in space (and possibly in time) :  $\mathbf{K} \sim \text{low-pass filter} : \mathbf{K} = \mathbf{B} \mathbf{H}^{\mathsf{T}} (\mathbf{H} \mathbf{B} \mathbf{H}^{\mathsf{T}} + \mathbf{R})^{-1}$

with  $\mathbf{H}$  = tangent linear version of  $\mathbf{H}$ ,

**B** = background error covariance matrix,

**R** = observation error covariance matrix.

=> **K** accounts for relative accuracy of observations, and for amplitudes & spatial structures of background errors.



# Components in background error covariances; filtering and propagation of $y^{o}$ - $H(x^{b})$

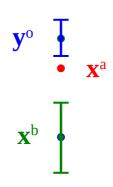
- Variances
  - Weighting/filtering of observations.

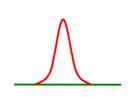


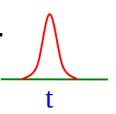
- Spatial propagation of observations.
- Spatial coherence of analysis.

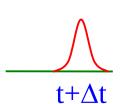


- Spatial and temporal propagation of observations.
- Spatial and temporal coherence of trajectory.

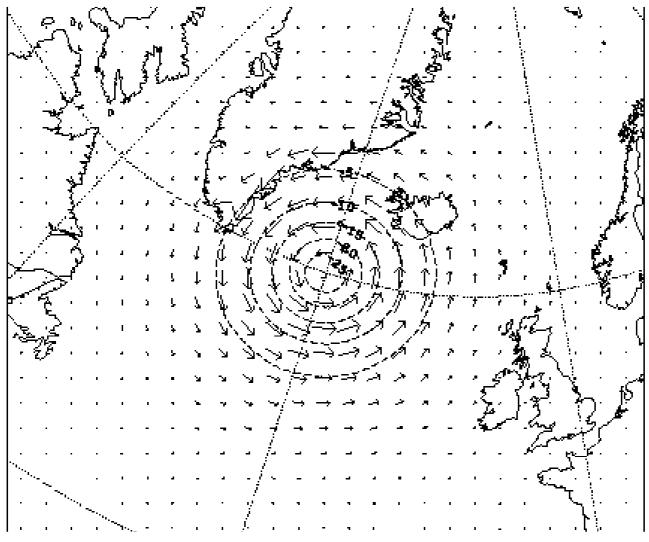








# Impact of one surface pressure observation on the pressure and wind analysis (2D)



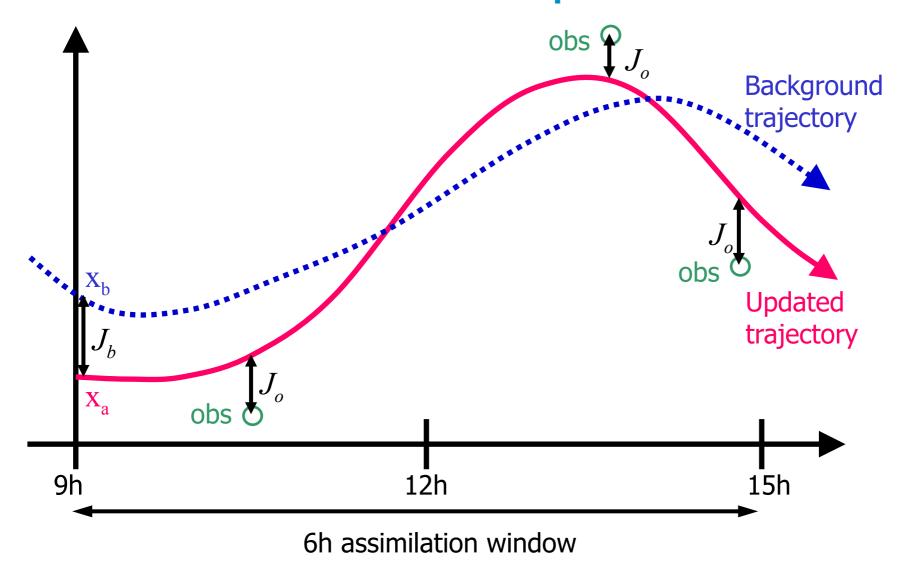
**B** contains & provides information about typical scales in the atmosphere and mass/wind couplings (such as geostrophy)



### **Variational analysis**

- Size of B is huge: square of model size ~ (109)² = 10¹8
   => B is too big to be computed explicitly or even stored in memory;
   error covariances need to be estimated, simplified and modelled.
- The matrix ( $\mathbf{H} \mathbf{B} \mathbf{H}^{\mathsf{T}} + \mathbf{R}$ ) in  $\mathbf{K} = \mathbf{B} \mathbf{H}^{\mathsf{T}} (\mathbf{H} \mathbf{B} \mathbf{H}^{\mathsf{T}} + \mathbf{R})^{-1}$  is also too big to be explicitly inverted.
  - => minimize distance  $J(\mathbf{x}^a)$  to  $\mathbf{x}^b$  and  $\mathbf{y}^o$  (variational assimilation), without explicit matrix inversions (e.g. Talagrand and Courtier 1987).
- Some (weakly) non linear features are accounted for in calculation of departures y° – H(xb) (e.g. non linear radiative transfer), and by updating the non linear trajectory in 4D-Var (non linear dynamics).

# How can we handle observations distributed in time within a 6h window ? $\Rightarrow$ Principle of 4D-Var assimilation



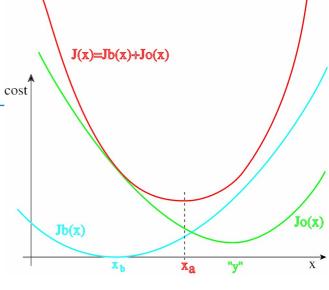
**Observations** are available at different times, while  $x_b$  is defined at beginning of window. => use model M to propagate  $x_b$  in time to compare it with observations (M is part of H); the analysis  $x_a$  can then be computed by **minimising**  $\mathbf{J} = \mathbf{J}_b + \mathbf{J}_o$ .



This allows an updated trajectory to be computed, consistent with observations at different times.

## **Implementation of 4D-Var**

Variational formulation : cost function  $J(\mathbf{x}^a) = ||\mathbf{x}^a - \mathbf{x}^b||^2_{\mathbf{B}^{-1}} + ||H(\mathbf{x}^a) - \mathbf{y}^o||^2_{\mathbf{R}^{-1}}$  minimised when gradient  $J'(\mathbf{x}^a)=0$  (equivalent to BLUE).



- Note that, if H is linear, then the cost function is quadratic,
   with a parabolic shape (see Figure).
- Computation of gradient J': development and use of adjoint operators
   (i.e. transpose of tangent-linear operators).
- Generalized observation operator H: includes NWP model M, in order to compare  $\mathbf{x}^{b}$  (valid at the beginning of the 6h window) with observations  $\mathbf{y}^{o}$  distributed in time over a 6h window.
- Reduction of computation cost : analysis increment  $δx = x^a x^b$  can be computed at low resolution (Courtier et al 1994).





# 2. In-situ observations and remote sensing data

Observation networks in meteorology: in situ measurements

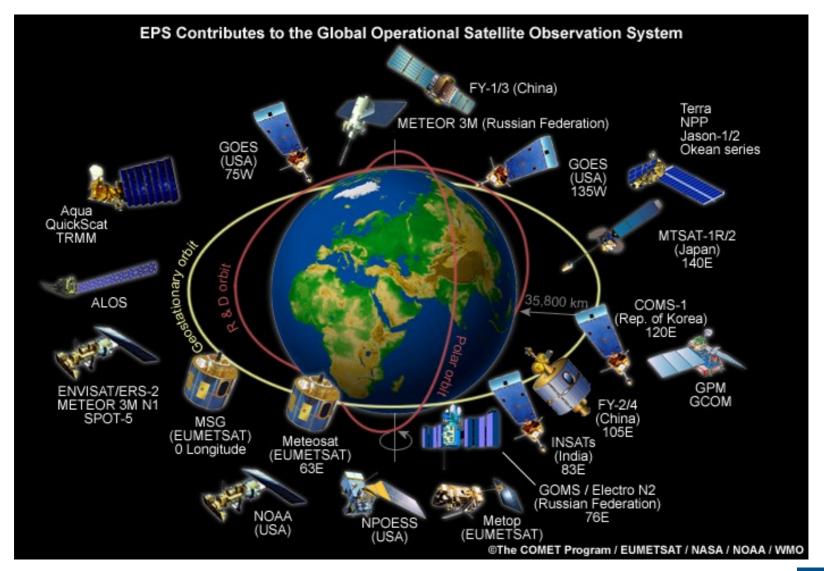
Provide direct information on the atmospheric state at the instrument location.



- \* Direct measurements of temperature, wind, humidity.
- \* Relatively easy to compare with the model, and to assimilate.
- \* High quality data, with relatively small biases.
- \* Poor horizontal coverage over the globe (ex : South Hemisphere ; oceanic areas).



## Observation networks in meteorology: satellite data





### **Geostationary satellites**

Fix position / earth, at 36 000 km height, above equator.

Same area of the globe (disk) is always observed.

### □ Advantages

Very high temporal resolution (~ 15 min).

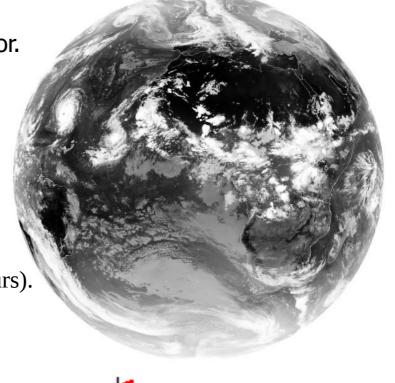
Useful for nowcasting (= very short range forecasts, e.g. within the next 2 hours).

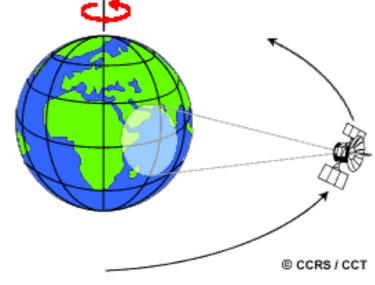
Dynamics of meteorological structures (e.g. fronts, tropical cyclones).

### □ Drawbacks

Insufficient spatial coverage of 1 satellite: several satellites are needed to cover the whole globe.

Not adapted to polar regions, due to position.





## **Polar orbiting satellites**

#### Low orbit satellites (800 km height):

Advantages

High spatial resolution ( $\sim$ 10 km).

Global spatial coverage (twice a day)

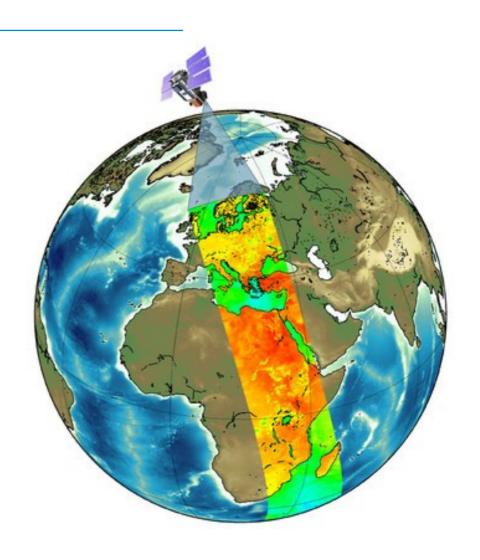
Sounding instruments (over several vertical layers):

~ vertical profiles of T at different locations

#### □ Drawbacks

Insufficient temporal resolution: a given location is only observed every 12h

(several satellites are needed, to have frequent observations over the same area)

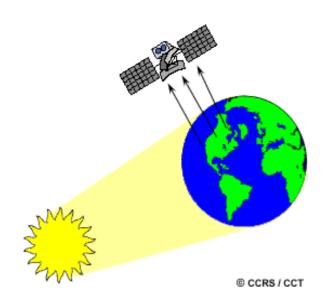




### Two types of satellite measurements

### **Passive measures**

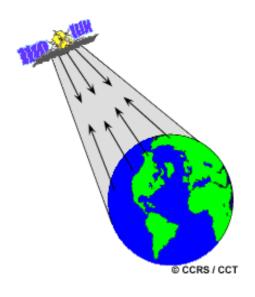
(no energy is emitted from instrument)



Measures natural radiation emitted by Earth or Atmosphere (with Sun origin)

### **Active measures**

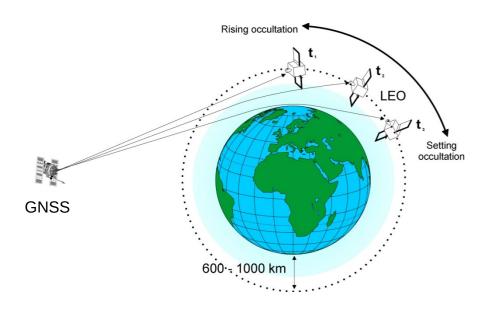
(energy is emitted from instrument)



Measures radiation emitted by satellite and then reflected or diffused by Earth or Atmosphere



# **GNSS** radio-occultation data (1st example of active remote sensing)



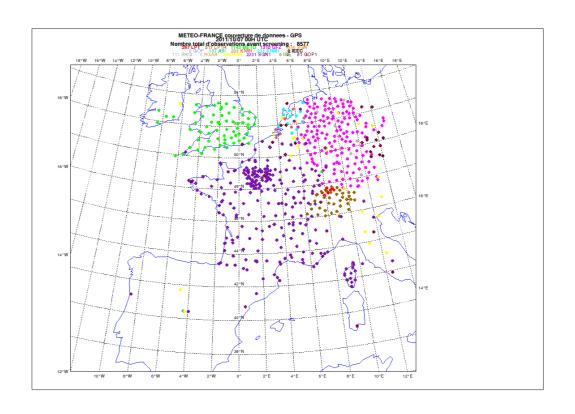
**GNSS** 

- GNSS is the Global Navigation Satellite System
   = GPS (USA) or Galileo (Europe).
- Low-Earth Orbit (LEO) satellites receive a signal emitted by a GNSS satellite.
- The GNSS signal passes through the atmosphere and it gets refracted along the way.
- The magnitude of the refraction depends on temperature, moisture and pressure near the tangent point (in red) of the path.
- The relative position of GNSS and LEO changes over time
   => vertical scanning of the atmosphere,
   with information on temperature and humidity.



# Data from ground-based receiver stations of GNSS (2nd example of active remote sensing)



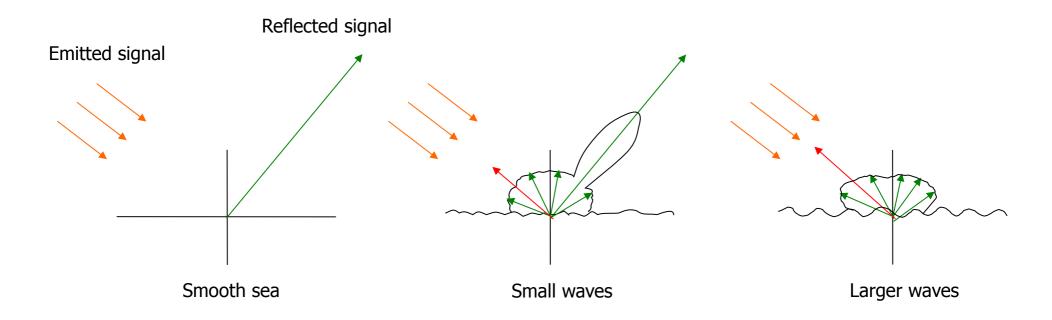


- Propagation of GNSS signal is slowed by atmosphere (dry air and water vapour):
   the propagation delay provides information about humidity in particular.
- More than 900 GNSS stations over Europe provide an estimation of Zenith Total Delay (ZTD) in real time to weather centres.
  - "All weather" instrument (e.g. for either dry or rainy conditions);
  - High temporal resolution (=> follow dynamics of convective developments).

### **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.

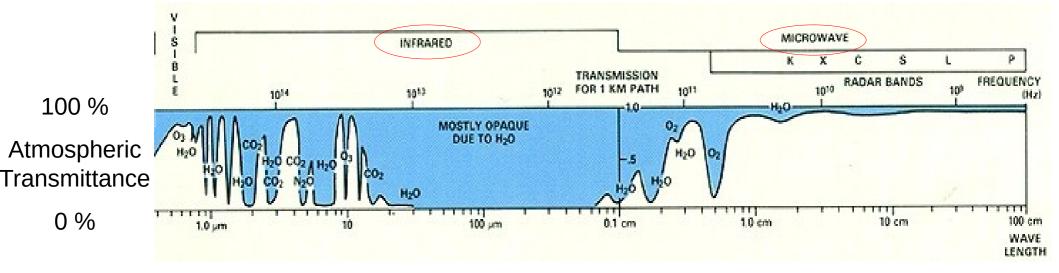


=> Measurements of near surface wind over the ocean, through backscattering of microwave signal reflected by waves.



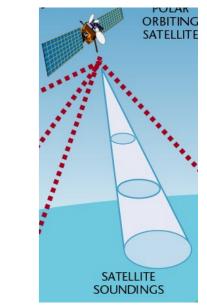
## Passive remote sensing: what is measured by satellite sensors?

- Sensors do not measure directly atmospheric temperature and humidity, but **electromagnetic radiation**: brightness temperature or radiance.
- Depending on wave length, indirect information on gas concentration (e.g. humidity) or on physical properties of atmosphere (temperature or pressure).
- Observations are often made in « atmospheric windows » (in white, below), e.g. in microwave and some infrared : frequencies with « high atmospheric transmittance » (= « low opacity » : radiation passes through the atmosphere to Earth surface, without being absorbed by gases) ; indirect info on T.



## Passive remote sensing: radiative transfer equation

What is observed is a **radiance** = quantity of energy per time unit, going through a surface, in a solid angle, and for a wave number interval of the radiation. [Unit: W/m<sup>2</sup>Sr.cm<sup>-1</sup>]



- Planck function:
  - $B_{\nu}(T)$ = radiance emitted by a black body at temperature T, for wave number  $\nu$ .
- Intensity of the radiation, emitted by the atmosphere at wave number  $\upsilon$ :

$$R_{\upsilon} = I_{0,\upsilon} \tau_{\upsilon}(z_0) + \int_{z_0} B_{\upsilon}[T(z)] \left[ \frac{d\tau_{\upsilon}(z)}{dz} \right] dz$$

 $I_{0,\upsilon}$  is the *surface emission* at altitude  $z_0$ .

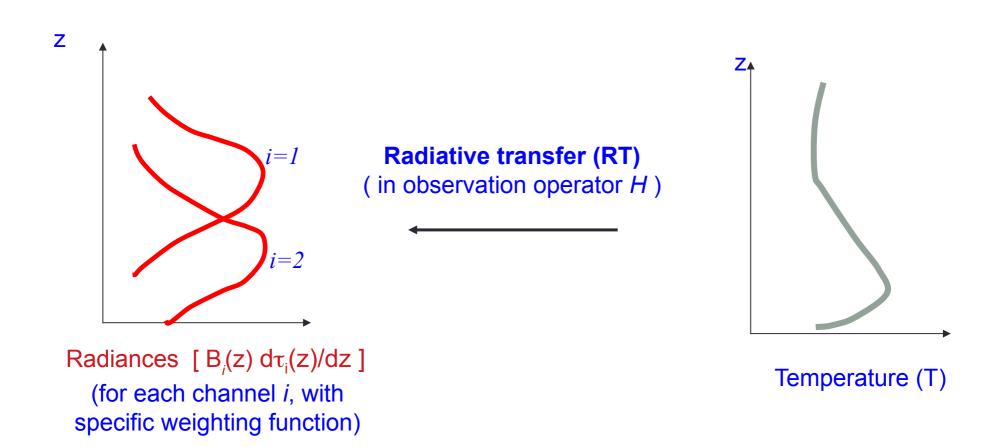
 $\tau_{\upsilon}(z)$  is the *transmittance* from z to the top of the atmosphere : it accounts for atmospheric absorption of radiation.

 $K_{\nu}(z) = \frac{d\tau_{\nu}(z)}{dz}$  is called *weighting function*:

it weights the Planck function in the radiance equation, and it determines the vertical layer of the atmosphere sounded at considered frequency v.



### Radiative Transfer: compute « simulated radiances » from temperature model profiles, wich can be compared with « observed radiances »

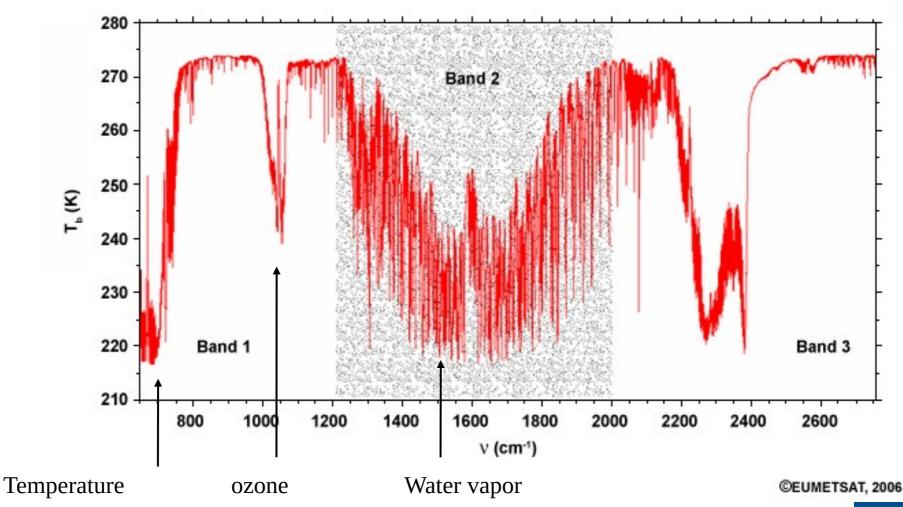


Using Radiative Transfer (in the observation operator) allows a large number of satellite radiances to be assimilated in NWP

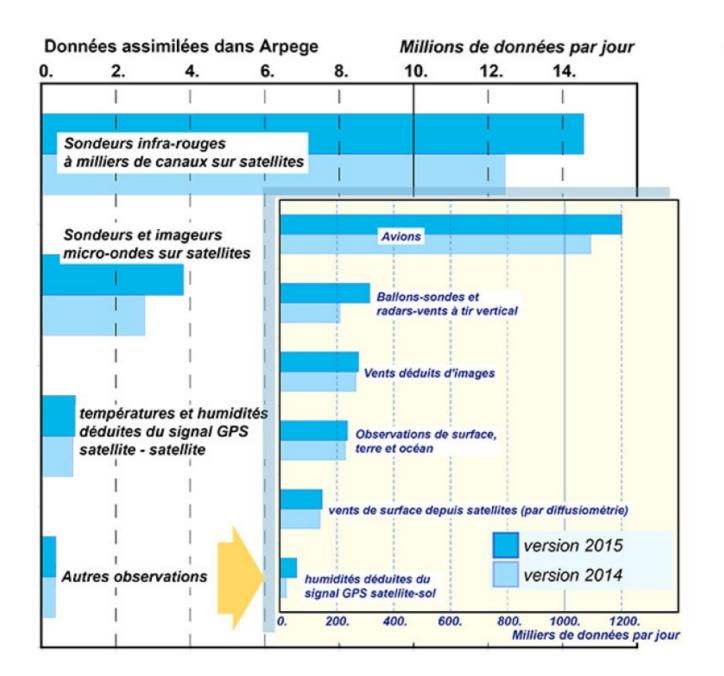


# IASI: infra-red interferometer developed by CNES and EUMETSAT

### IASI offers a very high spectral resolution (~ 8000 channels)



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



Total ~ 20 million obs per day



## How do observations meet global NWP requirements ?

#### Surface observations

good coverage over land, sparse coverage over sea; observations not suited to describe upper levels.

#### Aircraft observations

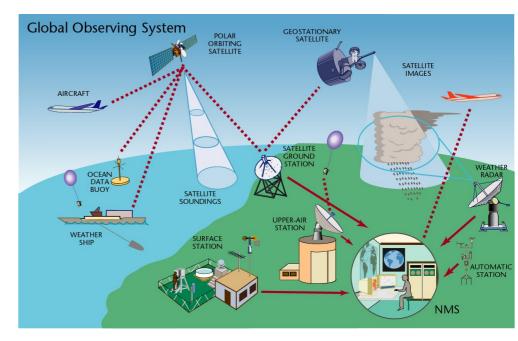
good accuracy, but do not describe the 3D state of the atmosphere (except near airports, during takeoff and landing).

#### Radiosonde data

good accuracy, good vertical resolution, but poor horizontal coverage over the globe.

#### Satellite data

good horizontal coverage over the globe, but poor vertical resolution (reduced to 1 level for satellite winds or imagers).





### Radar network in Arome-France

31 radars in France,62 radars in neighbouring countries;every 15 minutes, at 1 km resolution.

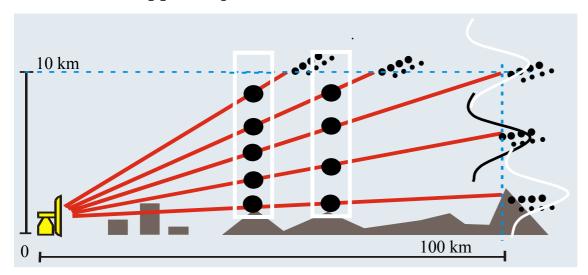
#### • Observations :

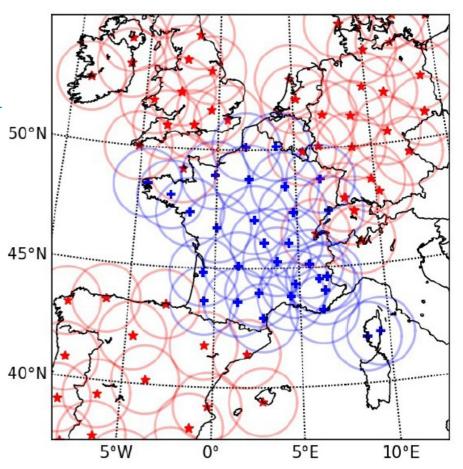
reflectivities Z (related to precipitation);

radial winds Vr (doppler effect):

the emitted microwave signal returns to the radar with a modified frequency, when the target is moving (wind).

=> invert Doppler equation to obtain a wind observation.





Observations assimilated as vertical profiles, after estimating the pixel altitudes

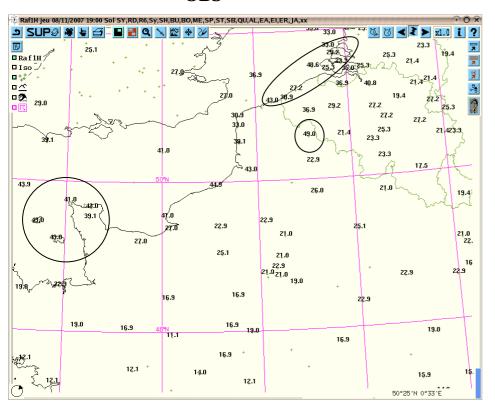
( Pixel altitude is computed using a constant refractivity index along the path )(= effective radius approximation)

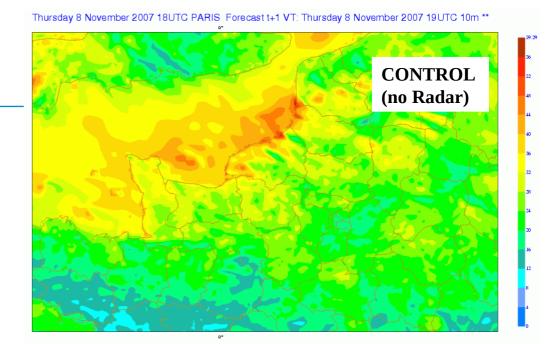
## **Assimilation of radar** radial winds

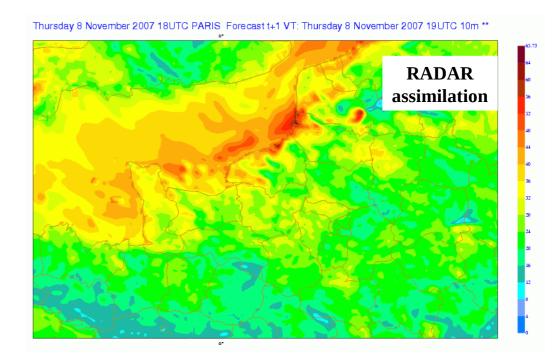
### Wind gust at 10 m (kt)

Forecast +1h (19 UTC)

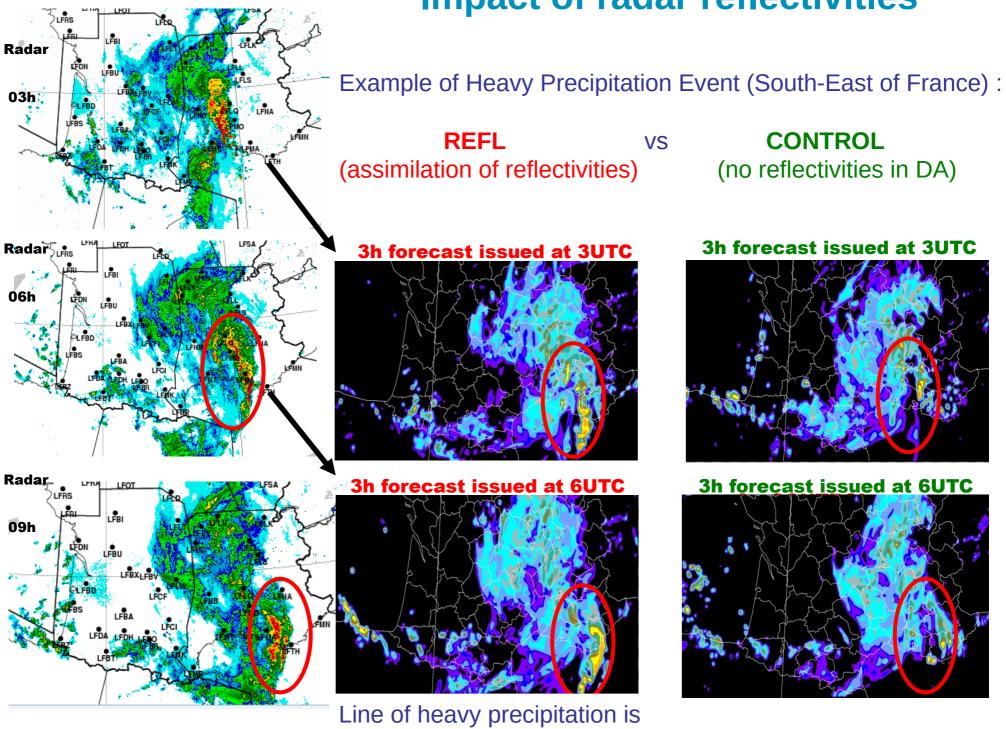
#### **OBS**







### Impact of radar reflectivities



well simulated in REFL run.



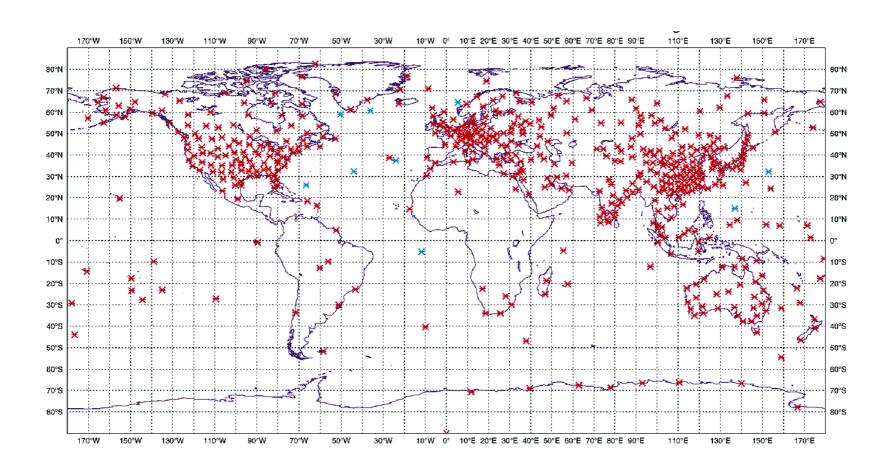
Error covariances: estimation and modelling (to weight and spatialise observed information)

### How can we estimate error covariances?

- The true atmospheric state is never (exactly) known.
- Use observation-minus-background departures
   to estimate some average variances and correlations of R and B,
   using assumptions on spatial structures of errors.
- Use an ensemble to simulate the error evolution and to estimate space- and time-dependent background error structures.
- Use covariance modelling to filter out sampling noise and other uncertainties in the ensemble.



#### Radiosonde observation network





#### **Covariances of innovations**

Innovations = observation-background departures :

$$\mathbf{y}^{o} - H(\mathbf{x}^{b}) = \mathbf{y}^{o} - H(\mathbf{x}^{t}) + H(\mathbf{x}^{t}) - H(\mathbf{x}^{b})$$
  
 $\approx \mathbf{e}^{o} - \mathbf{H}\mathbf{e}^{b}$ 

Innovation covariances :

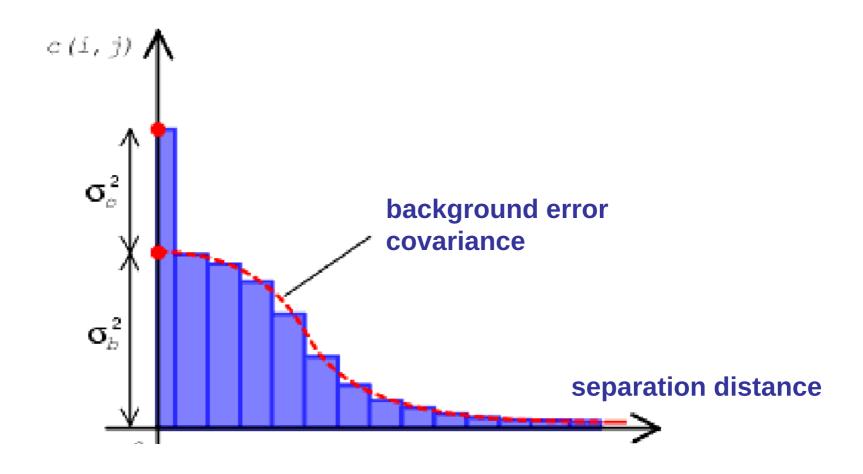
$$E[(\mathbf{y}^{\circ} - H(\mathbf{x}^{\circ}))(\mathbf{y}^{\circ} - H(\mathbf{x}^{\circ}))^{\mathsf{T}}] = \mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^{\mathsf{T}}$$

assuming that  $E[e^{o}(He^{b})^{T}] = 0$ .

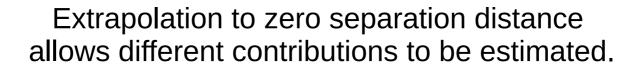
(e.g. Hollingsworth and Lönnberg 1986).



# Covariances of innovations (with extrapolation to zero separation distance)

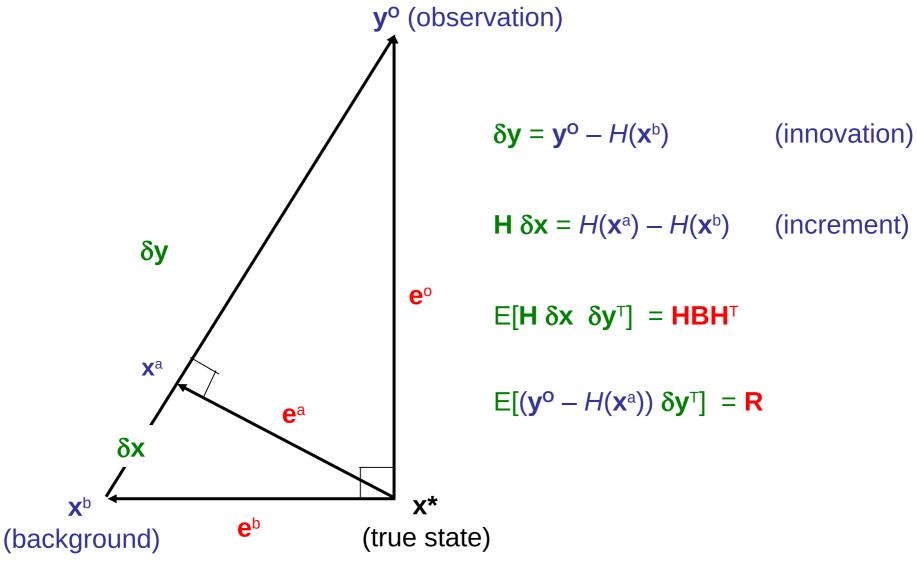


$$E[(\mathbf{y}^{\circ} - H(\mathbf{x}^{\circ}))(\mathbf{y}^{\circ} - H(\mathbf{x}^{\circ}))^{\mathsf{T}}] = \mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^{\mathsf{T}}$$



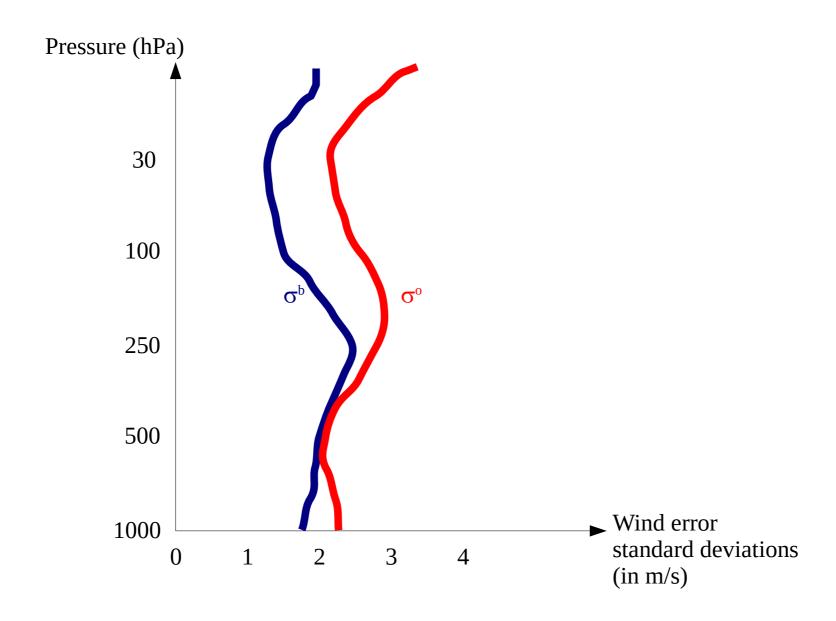


#### **Covariances of analysis residuals**



(Desroziers et al 2005)

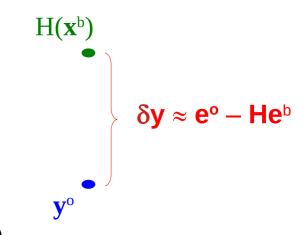
# Vertical profiles of standard deviations of background errors and observation errors





# Space & time averages of innovation-based covariances

At a given location and time, there is only 1 innovation value  $\delta y$ : a single error realization is available locally (e.g. for estimating background error variance).

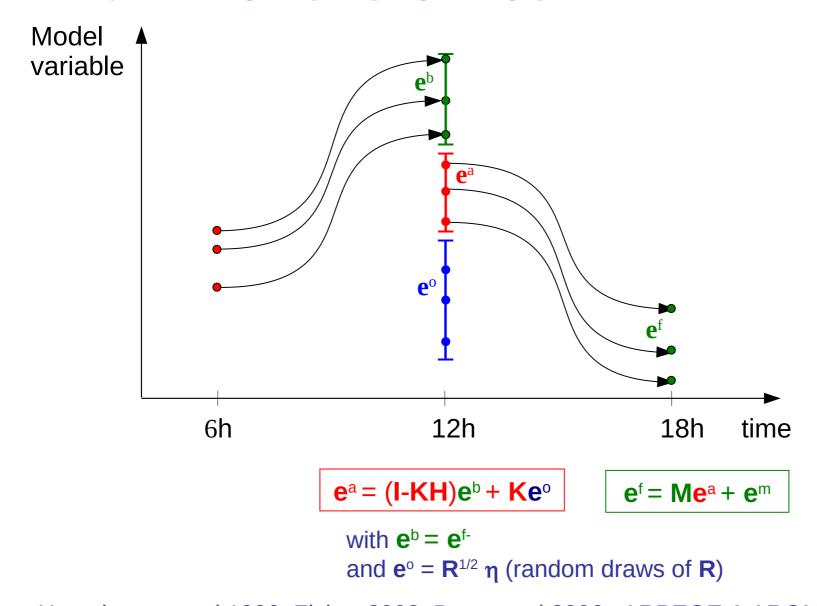


- => Statistical averages (mathematical expectations) need to be replaced by space and time averages (ergodic assumption).
- => only space or time averages of **B** and **R** can be estimated from innovation data.



=> consider other approaches, such as ensemble methods.

# Ensemble Data Assimilation (EDA): simulation of error cycling, by adding & propagating perturbations

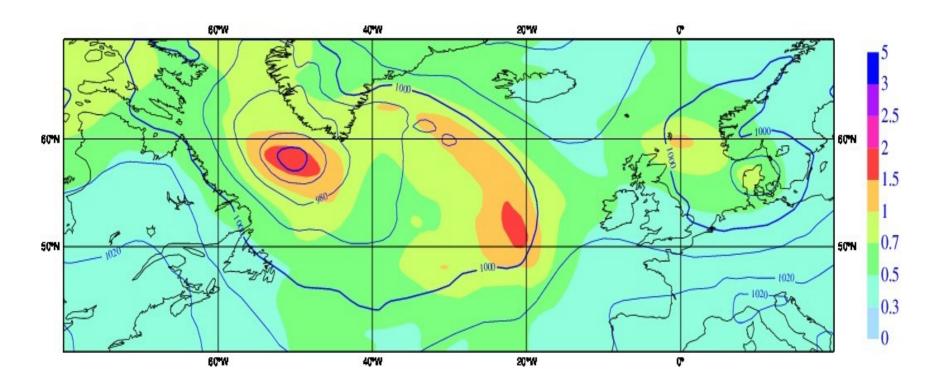


(e.g. Houtekamer et al 1996, Fisher 2003, Berre et al 2006 ; ARPEGE & AROME EDA : 50 members to estimate flow-dependent **B** and to initialise ensemble predictions)

# Simulation and propagation of observation errors and model errors

- Observation errors can be simulated by adding random draws of R: e° = R<sup>1/2</sup> η°.
- Model errors can be simulated in different ways, e.g. by:
   adding random draws of Q: e<sup>m</sup> = Q<sup>1/2</sup> η<sup>m</sup> (additive or mult. inflation);
   using a multi-model approach (or multi-physics);
   perturbing physical tendencies of the model;
   perturbing model parameters. (...)
- Observation and model perturbations are propagated during the successive analysis/forecast steps of DA cycling.
- Flow-dependent background error covariances can be estimated from the ensemble spread.

### **Dynamics of background error variances**



Standard deviations of surface pressure errors (hPa) (superimposed with MSLP analysis (hPa)).

=> larger weight given to observations in regions where the background is particularly uncertain (intense weather events)

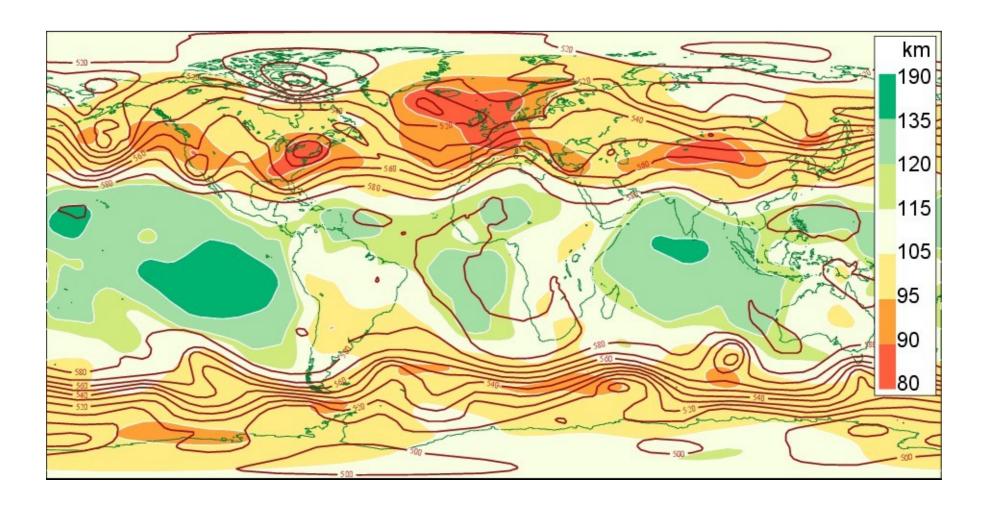


### Modelling and filtering covariances

- Huge size of B: model it with operators which are sparse and/or of small size.
- Sampling noise, and other uncertainties. => Spatio-temporal filtering.
- Factorisation :  $\mathbf{B} = \mathbf{U} \mathbf{U}^{\mathsf{T}}$  $\mathbf{U} = \mathbf{L} \mathbf{S} \mathbf{C}_{\mathsf{f}}$
- L ~ mass/wind cross-covariances (related to geostrophy), including flow-dependence (non linear balances).
- **S** flow-dependent standard deviations (~ expected error amplitudes), filtered spatially.
- $C = C_f C_f^T$  matrix of 3D spatial correlations (~ spatial structures of errors), filtered in wavelet space (block-diagonal model).



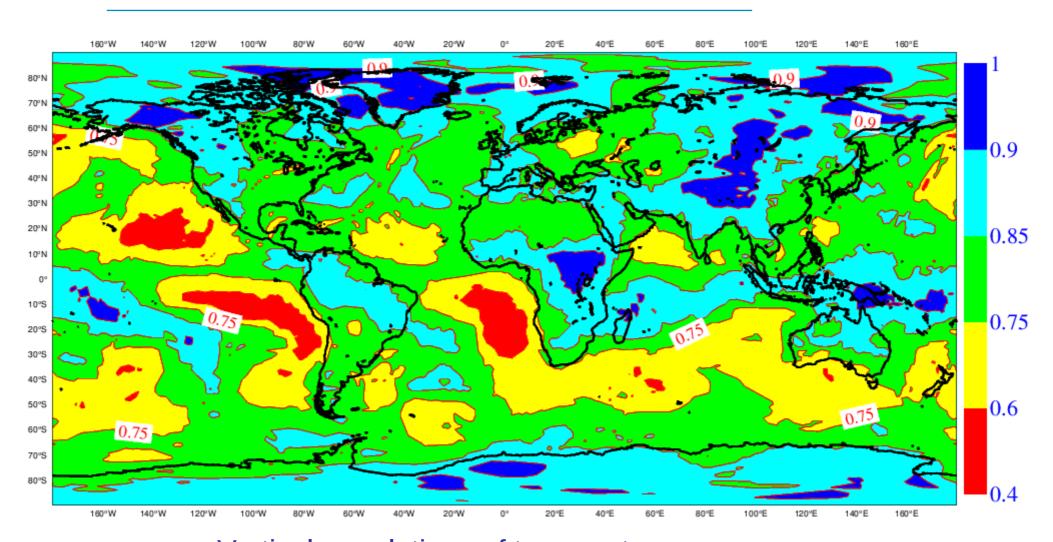
## **Dynamics of horizontal correlations**



Horizontal length-scales (in km) of wind errors near 500 hPa, superimposed with geopotential



#### **Dynamics of vertical correlations**

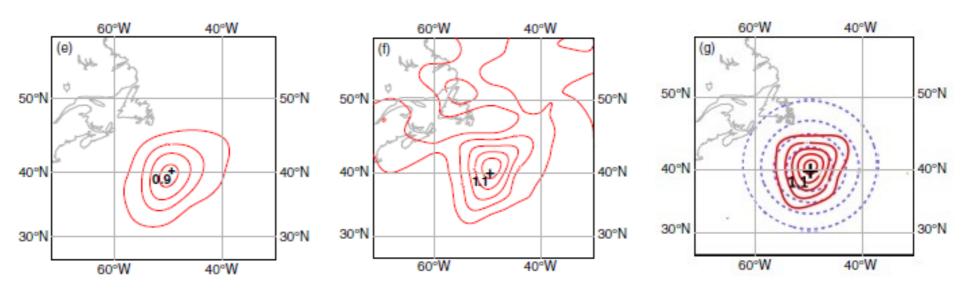


Vertical correlations of temperature errors between 850 & 870 hPa



#### **Covariance anisotropy and localisation**

Use ensemble to get information on anisotropy, but it requires filtering = localisation.



« Exact » covariances

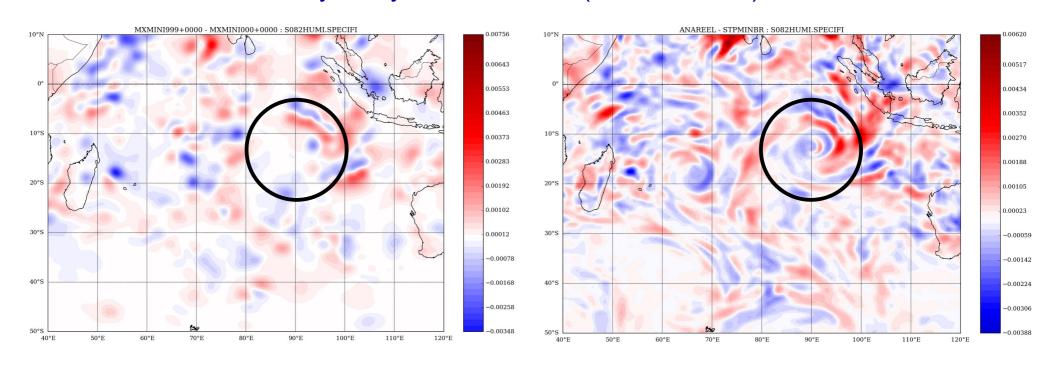
Raw covariances (200 members)

Localised covariances (200 members)



## Flow-dependent anisotropic increments

#### Humidity analysis increments (near 850 hPa)



With isotropic correlations

With anisotropic correlations, filtered by localisation





## **Conclusions**

#### **Conclusions**

- Data assimilation is vital for weather forecasting.
- Observations are very diverse in type, density and quality.
- 4D schemes for temporal and non linear aspects.
- Observation-background departures for estimation of average variances and correlations in R and B.
- Ensemble DA for error simulation and for covariance dynamics.
- Sampling noise issues and filtering methods.
- Towards 4DEnVar (variational assimilation based on a 4D ensemble).





# Thanks for your attention



#### Liens principaux entre thématiques ensemblistes

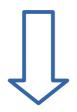


#### Spécification des incertitudes

(observations, modèle) du système d'analyse/prévision



Assimilation d'ensemble : simulation de la propagation des erreurs au cours du "cyclage" de l'assimilation



Spécification des covariances d'erreur d'ébauche, formulations EnVar de l'assimilation

Prévision d'ensemble : simulation de la propagation / amplification des erreurs au cours de la prévision



Prévision probabiliste : traitement statistique des prévisions de l'ensemble

#### **Properties of innovation methods**

- Provides estimates in observation space.
- A good quality data dense network is needed.
- Assumption that observation errors are spatially uncorrelated.
- An objective source of information on B and R.
- At a given location and time, only 1 innovation value :
   only a single error realization is available.
  - => Statistical averages (expectations) are replaced by space and time averages (ergodic assumption).



### 4DEnVar Variational analysis based on a 4D Ensemble

Minimisation of  $J(\underline{\delta x})$  where  $\underline{\delta x}$  is a 4D analysis increment :

$$J(\underline{\delta}\mathbf{x}) = \underline{\delta}\mathbf{x}^{\mathsf{T}}\underline{\mathbf{B}}^{\mathsf{-1}}\underline{\delta}\mathbf{x} + (\underline{\mathbf{d}}\underline{\mathbf{H}}\underline{\delta}\mathbf{x})^{\mathsf{T}}\underline{\mathbf{R}}^{\mathsf{-1}}(\underline{\mathbf{d}}\underline{\mathbf{H}}\underline{\delta}\mathbf{x})$$

with  $\underline{\mathbf{B}} = \underline{\mathbf{X}}^{b'} \underline{\mathbf{X}}^{b'T}$  o  $\underline{\mathbf{L}}$ , where  $\mathbf{L}$  is a localization matrix,  $\underline{\mathbf{X}}^{b'} = (\underline{\mathbf{x}}^{b'}_{1}, \dots, \underline{\mathbf{x}}^{b'}_{Ne})$ ,

$$\underline{\mathbf{x}}^{b'}_{ne} = \underline{\mathbf{x}}^{b}_{ne} - \langle \underline{\mathbf{x}}^{b} \rangle / (N^{e}-1)^{1/2}, \text{ ne = 1, } N^{e}.$$

 $\underline{\mathbf{x}}^{b'}$  of dimension K+1 (time) x M (3D variables) x N (dim 3D).

(Liu et al, 2008, 2009; Buehner et al, 2010; Lorenc, 2012;



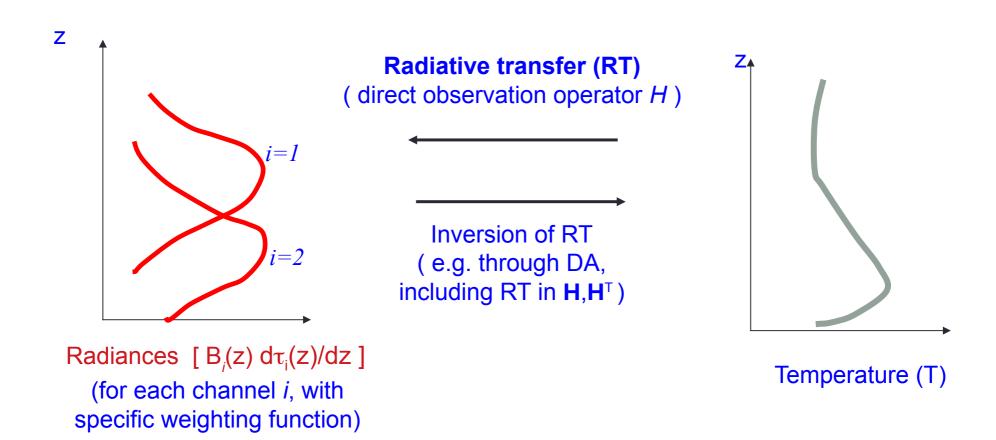
Desroziers et al 2014).

## 4DEnVar Variational analysis based on a 4D Ensemble

- 4D covariances from an ensemble of trajectories.
- Improved realism of 4D background error covariances
   (anisotropies, non linear evolution including all physical processes).
- Lesser need to develop and maintain an adjoint model in this case.
  - Especially important for AROME.
- Pursue within the variational framework
  - Global assimilation of all available observations, distributed in space and in time.
- Introduces additional levels of parallelism (space, time, ensemble).



#### Radiative Transfer: compute « simulated radiances » from temperature model profiles, wich can be compared with « observed radiances »



Using Radiative Transfer (in the observation operator) allows a large number of satellite radiances to be assimilated in NWP

