Data assimilation in meteorology

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

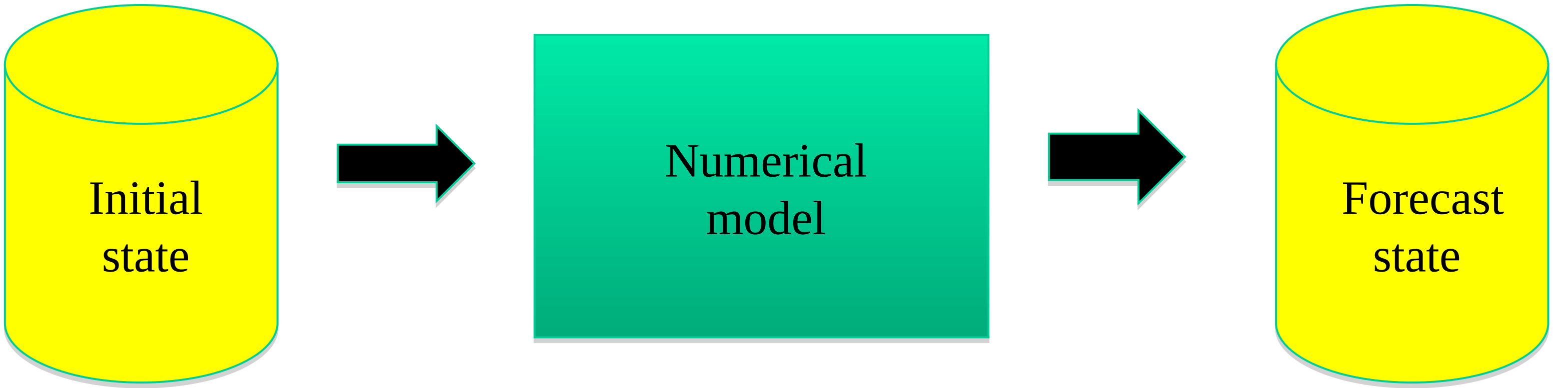
Plan of the talk

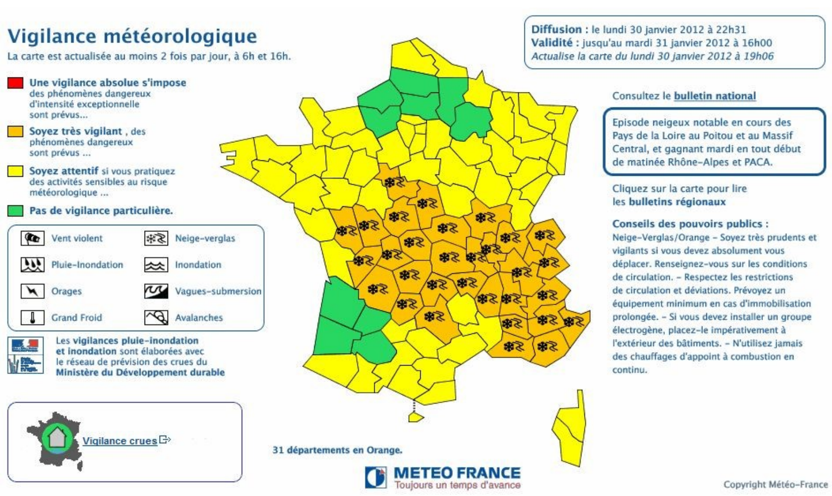
- Numerical Weather Prediction (NWP),
  Data Assimilation (DA)
- Observations (in-situ and remote sensing)
- Error covariances : estimation and modelling

1. Numerical Weather Prediction and Data Assimilation

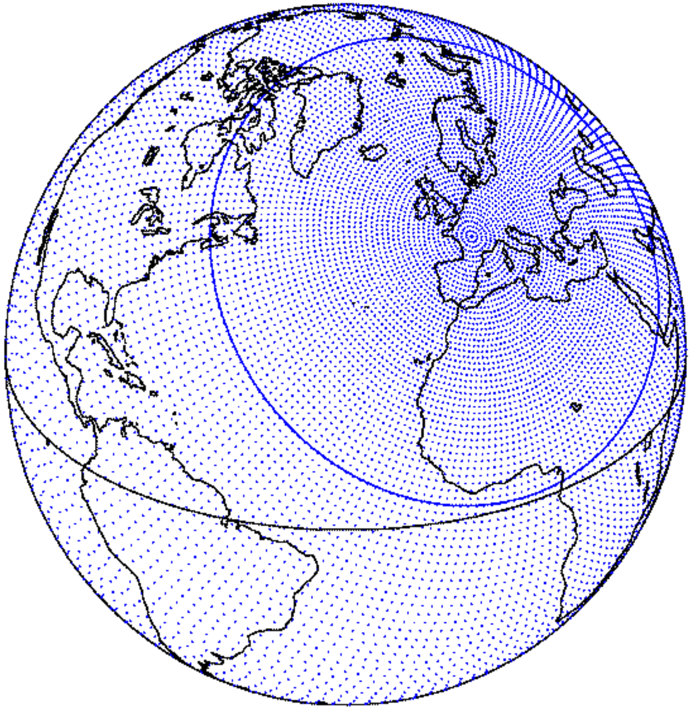
Numerical Weather Prediction

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

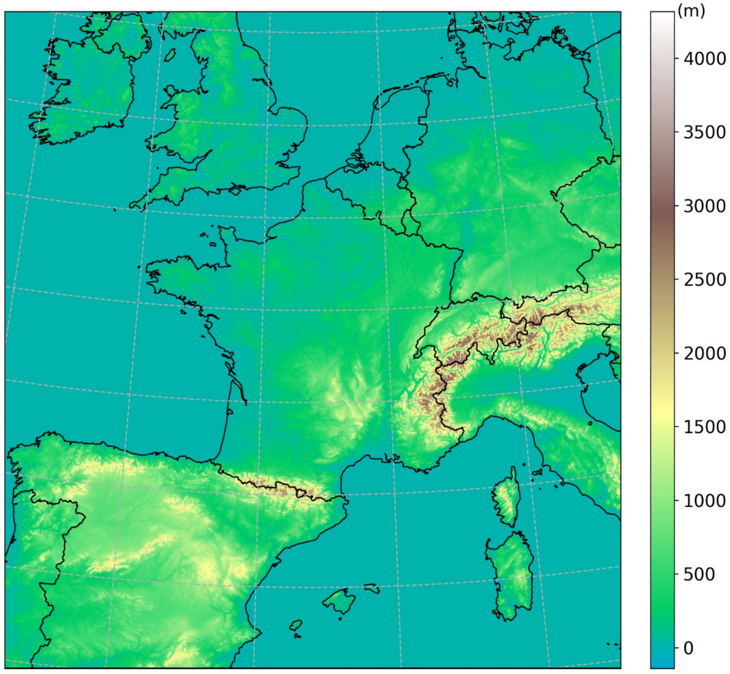




NWP models at Météo-France (in collaboration with ECMWF + other national institutes in Europe & N. Africa)



Lateral Boundary Conditions

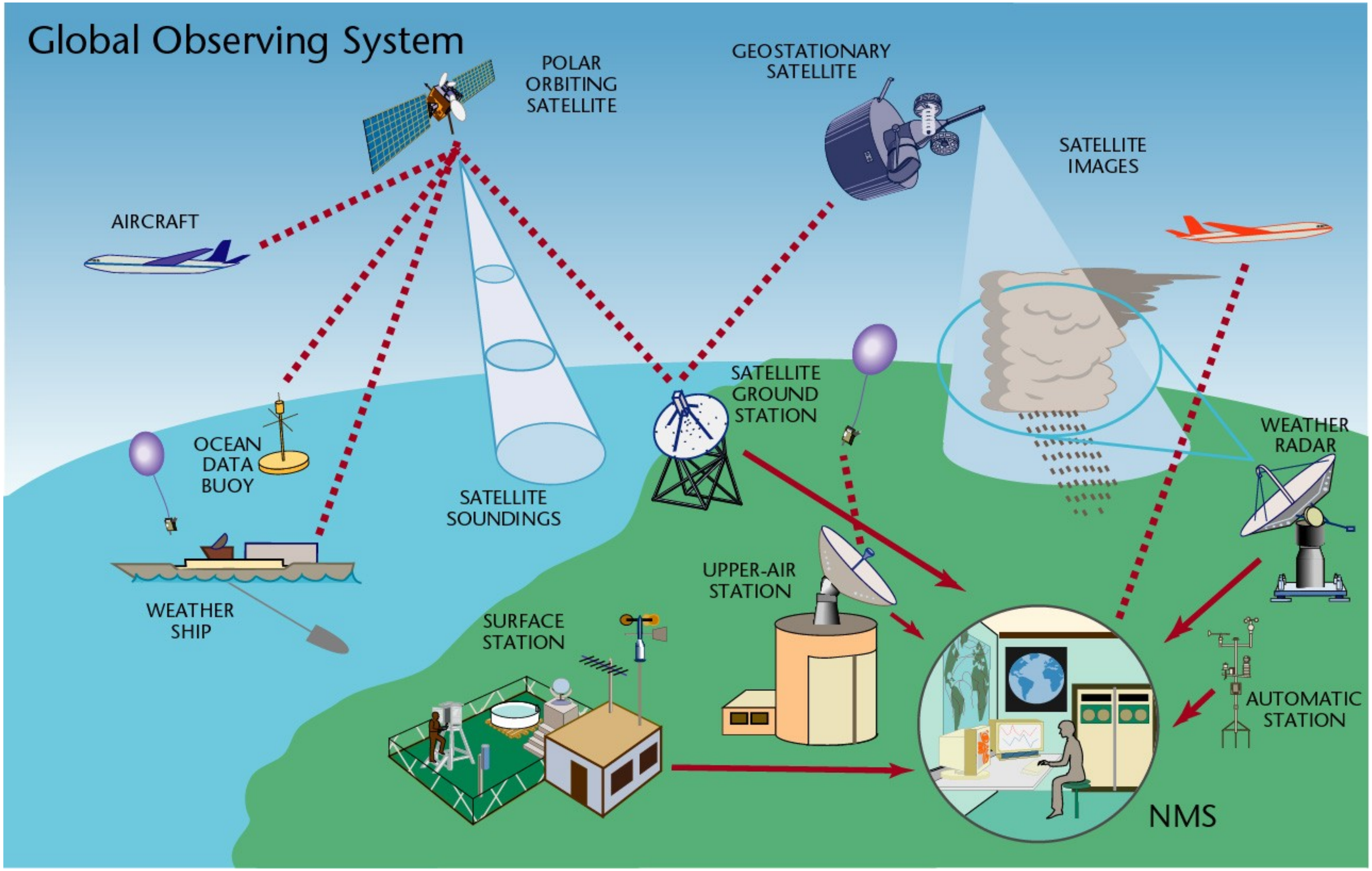


ARPEGE (5 km - 30 km)
10^9 model variables

AROME (1.3 km)
1,4 x 10^9 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
10^9 model variables

5×10^6 observations / 6h
90 % satellite
Computation time :
40 min (over 6h window)

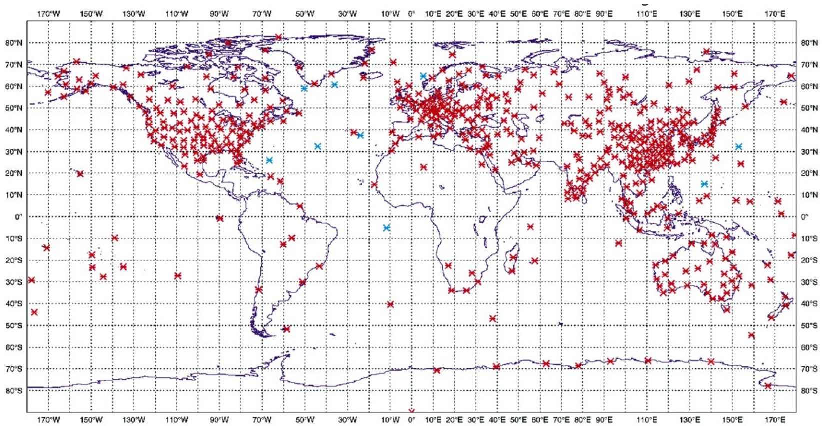
AROME
1,4 x 10^9 model variables

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

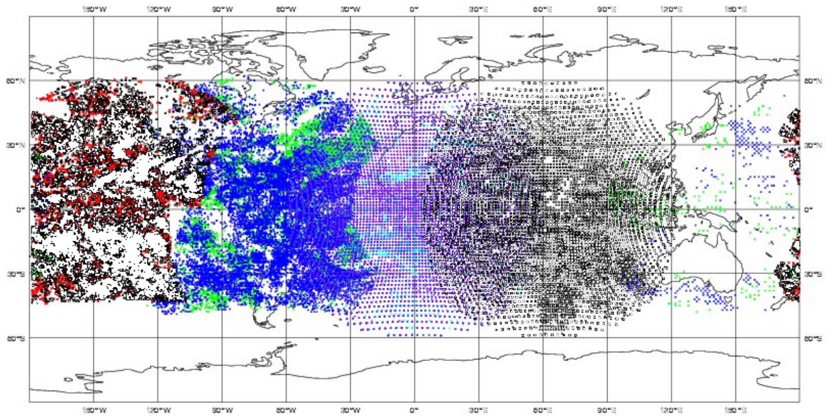
METEO FRANCE

Spatial coverage and density of observations

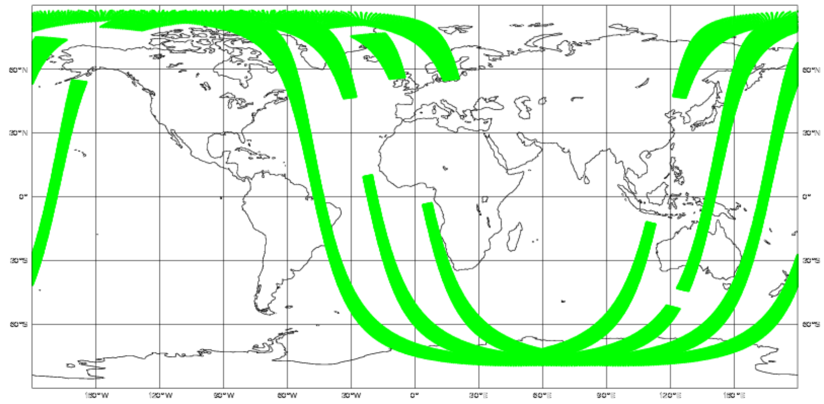
RADIOSONDE DATA



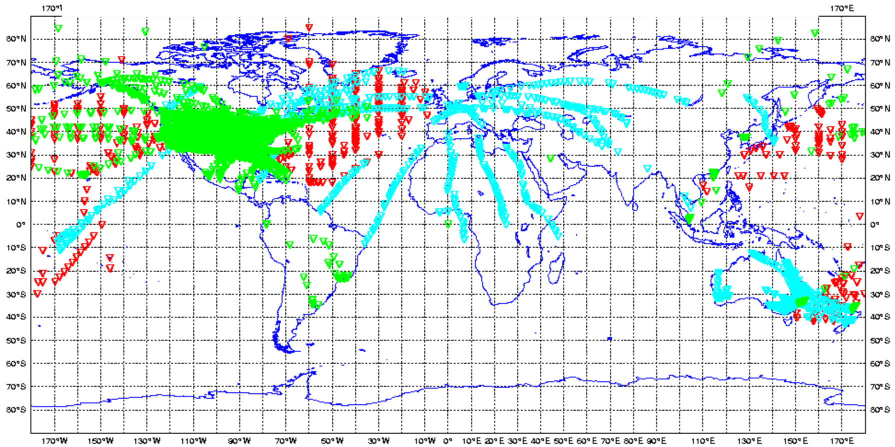
GEOSAT. WINDS



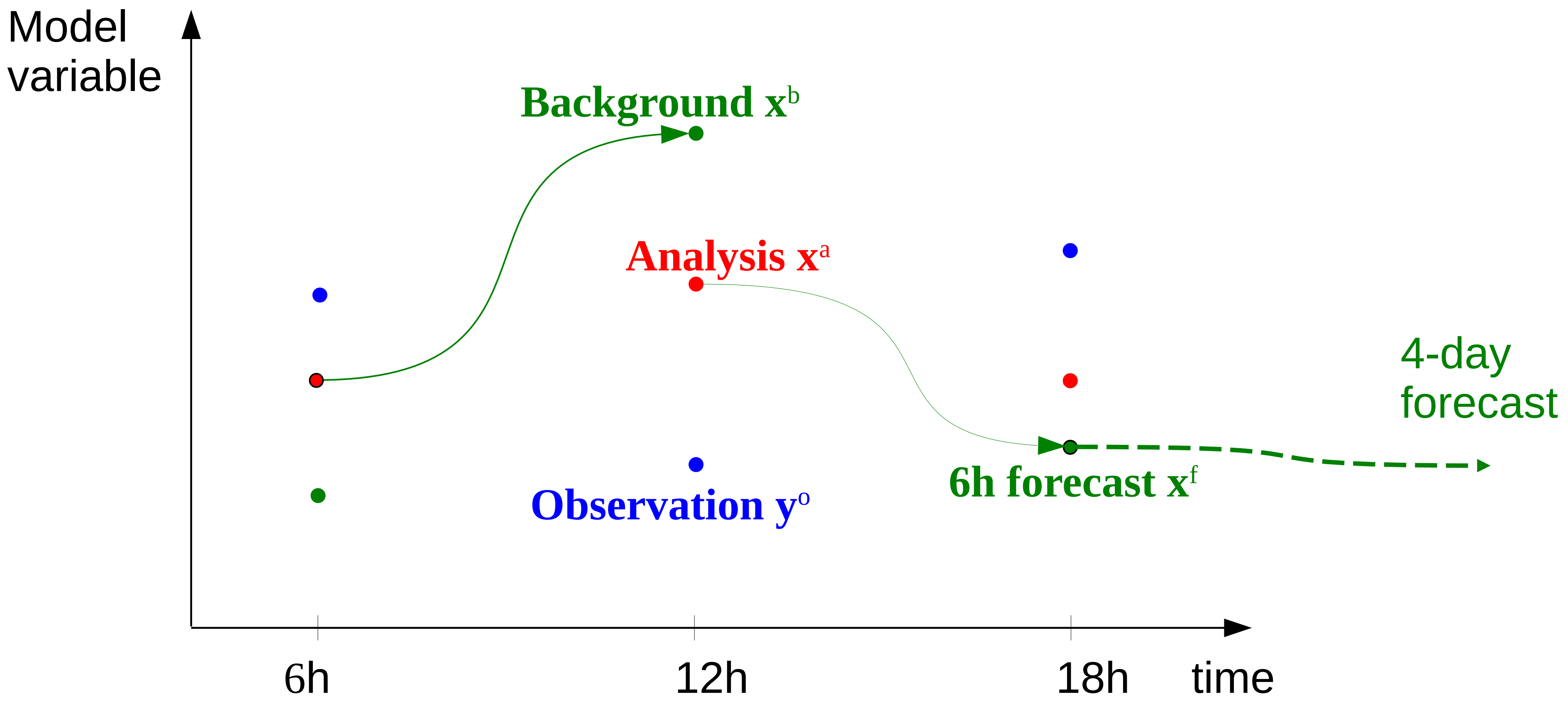
SCATTEROMETER



AIRCRAFT DATA



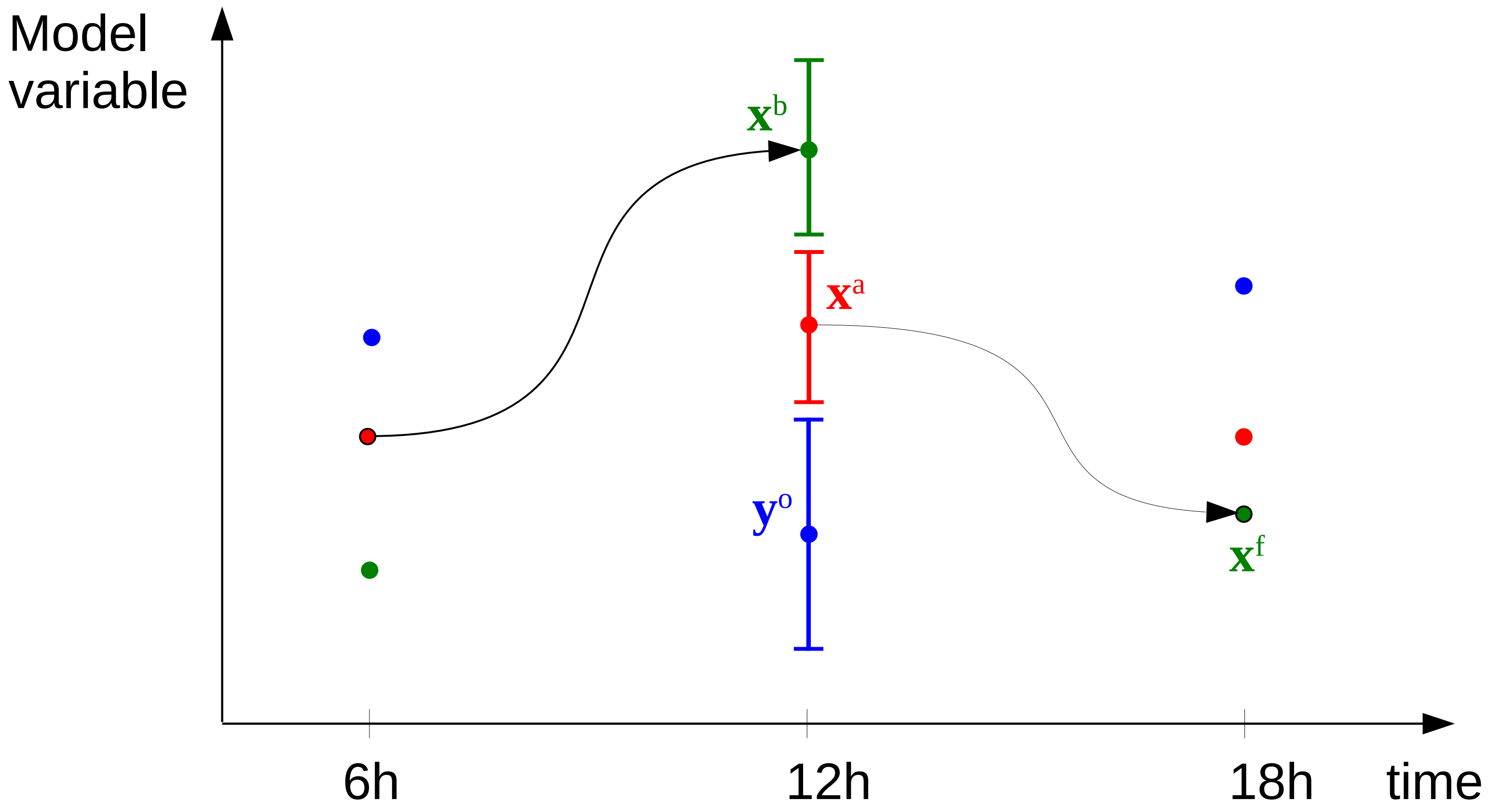
Sequential aspect of Data Assimilation :
temporal succession of analysis and forecast steps



=> 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** for the next analysis step ; 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).

=> Use linear estimation theory to account for errors.

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 directlty represented as such in the NWP model !

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

In order to compare observations ($\mathbf{y}^o$) with the background ($\mathbf{x}^b$), 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 time interval (= DA window).

METEO FRANCE

Analysis equation : definition & role of gain matrix K

- BLUE formalism : $\mathbf{x}^a = \mathbf{x}^b + \boldsymbol{\delta}\mathbf{x} = \mathbf{x}^b + \mathbf{K} \, (\mathbf{y}^o - H[\mathbf{x}^b])$

  where $\boldsymbol{\delta}\mathbf{x} = \mathbf{x}^a - \mathbf{x}^b$ is the analysis increment.

- Because observations are generally local & affected by small scale errors,

  $\mathbf{y}^o - H[\mathbf{x}^b]$ needs to be filtered and propagated in space (and possibly in time) :

  K ~ low-pass filter : $\mathbf{K} = \mathbf{B}\,\mathbf{H}^T (\mathbf{H}\,\mathbf{B}\,\mathbf{H}^T + \mathbf{R})^{-1}$

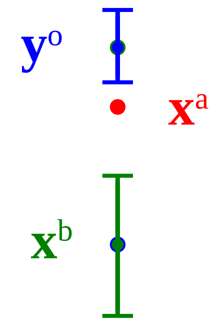
  with **H** = tangent linear version of ***H***,

  B = background error covariance matrix,

  R = observation error covariance matrix.

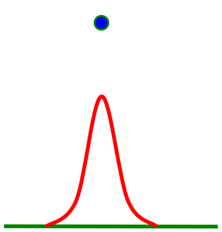
  => **K** accounts for the accuracy of observations (described in **R**), and for amplitudes & spatial structures of background errors (described in **B**).

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



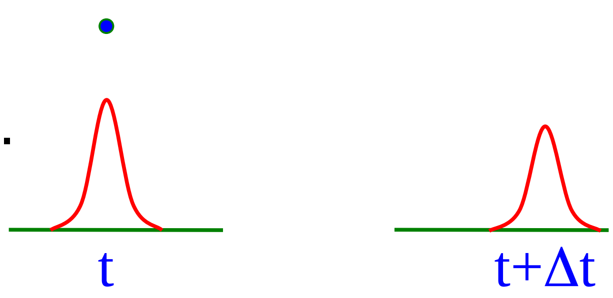
- Variances

  - Weighting/filtering of observations.



- 3D spatial correlations

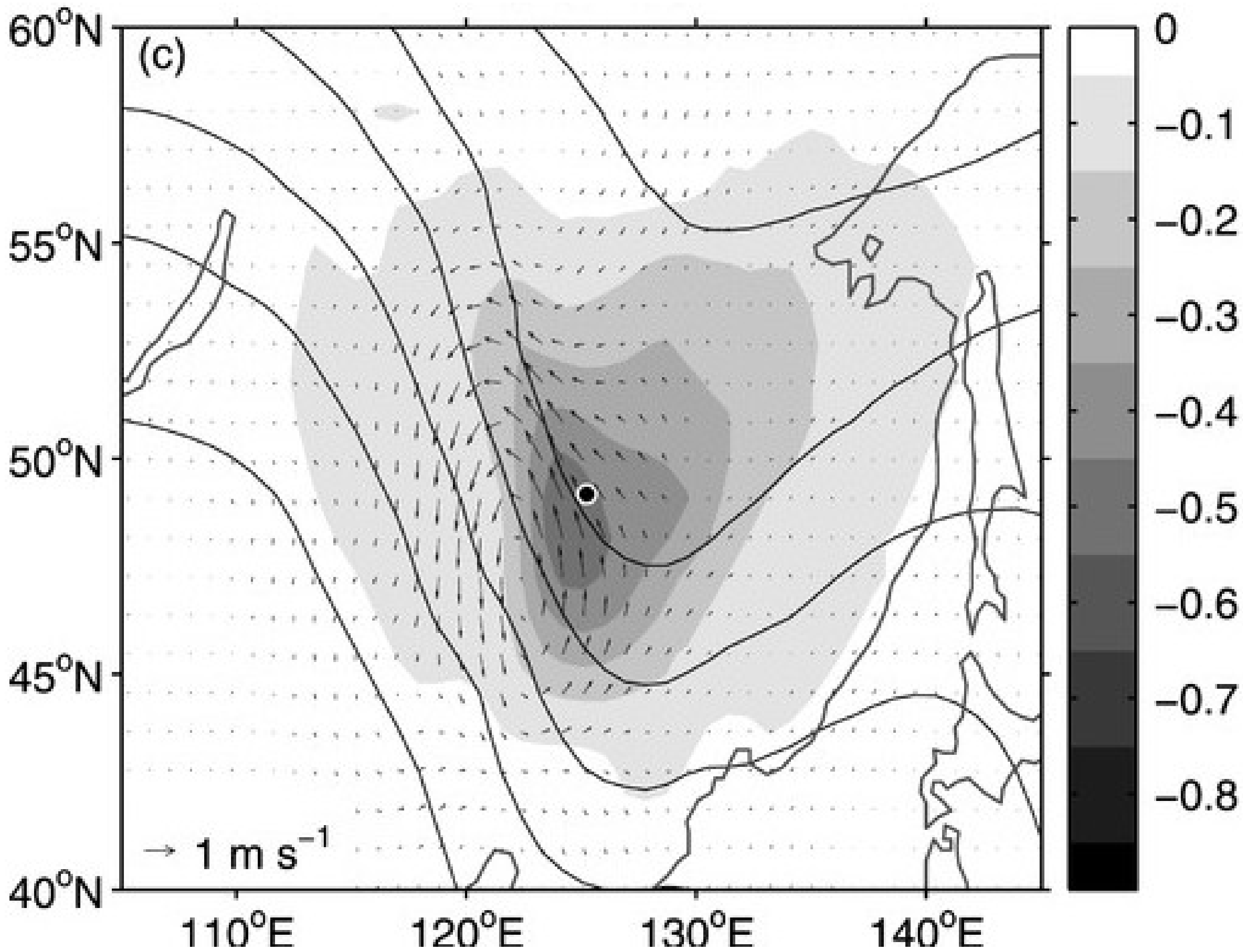
  - Spatial propagation of observations.
  - Spatial coherence of analysis.



- 4D spatio-temporal correlations

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

Impact of one temperature observation on temperature and wind analysis



B contains & provides information about typical scales in the atmosphere and mass/wind couplings (thermal wind law = geostrophy + hydrostatism)

(Buehner et al 2010)

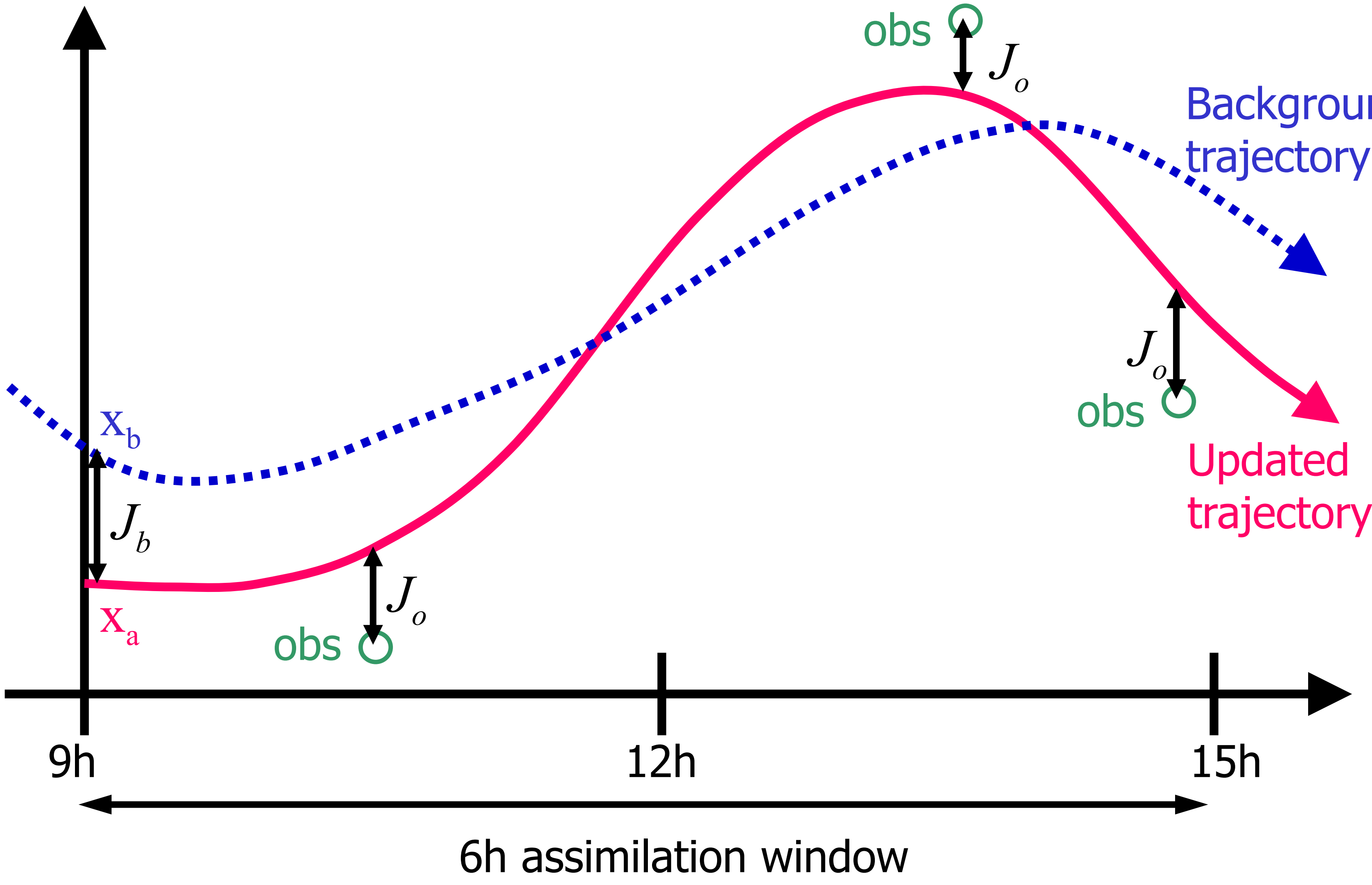
Variational analysis

- Size of **B** is huge : square of model size ~ $(10^9)^2 = 10^{18}$
  => **B** is too big to be computed explicitly or even stored in memory ;
  error covariances need to be estimated, simplified and modelled.

- The matrix ($\mathbf{H}\,\mathbf{B}\,\mathbf{H}^T + \mathbf{R}$) in $\mathbf{K} = \mathbf{B}\,\mathbf{H}^T\,(\mathbf{H}\,\mathbf{B}\,\mathbf{H}^T + \mathbf{R})^{-1}$
  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 $\mathbf{y}^o - H(\mathbf{x}^b)$ *(e.g. non linear radiative transfer),*
  and by updating the non linear trajectory in 4D-Var (non linear dynamics).]

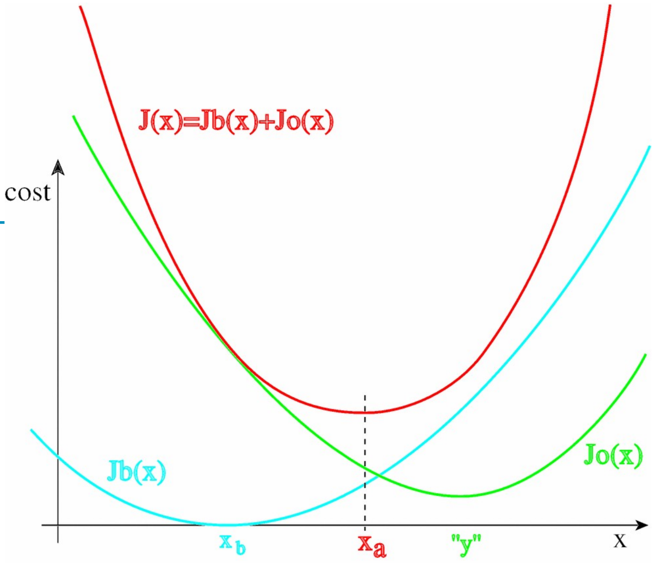
How can we handle observations distributed in time within a 6h window ? ⇒ Principle of 4D-Var



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** $\mathbf{J = J_b + J_o}$.
This allows an **updated trajectory** to be computed, consistent with observations at different times.

METEO FRANCE

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 top right 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 $\delta\mathbf{x} = \mathbf{x}^a - \mathbf{x}^b$ can be computed at low resolution (Courtier et al 1994).

4DEnVar for high resolution DA (e.g. km-scale)

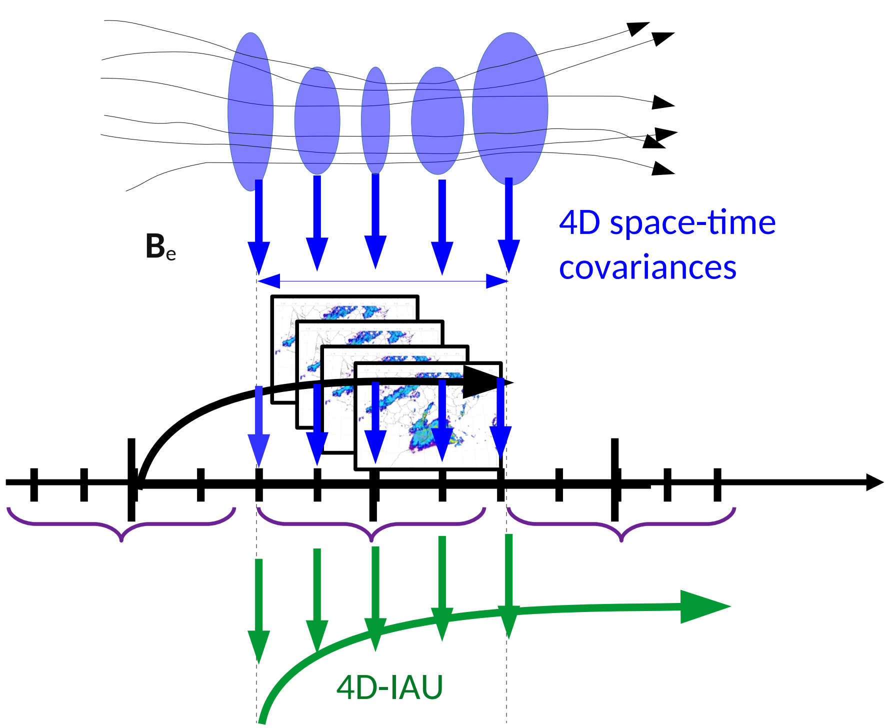
4DEnVar :

$$J(\underline{\delta\mathbf{x}}) = \frac{1}{2}(\underline{\delta\mathbf{x}})^T\underline{\mathbf{B}}^{-1}(\underline{\delta\mathbf{x}}) + \frac{1}{2}(\underline{\mathbf{d}} - \underline{\mathbf{H}}\underline{\delta\mathbf{x}})^T\underline{\mathbf{R}}^{-1}(\underline{\mathbf{d}} - \underline{\mathbf{H}}\underline{\delta\mathbf{x}})$$

$$\underline{\delta\mathbf{x}} = \begin{pmatrix} \delta\mathbf{x}_0 \\ \delta\mathbf{x}_1 \\ \vdots \\ \delta\mathbf{x}_K \end{pmatrix} \qquad \underline{\mathbf{B}} = \underline{\widetilde{\mathbf{B}}}^{\mathrm{e}} = \begin{pmatrix} \widetilde{\boldsymbol{B}}^{\mathrm{e}}_{0,0} & \widetilde{\boldsymbol{B}}^{\mathrm{e}}_{0,1} & \cdots & \widetilde{\boldsymbol{B}}^{\mathrm{e}}_{0,K} \\ \widetilde{\boldsymbol{B}}^{\mathrm{e}}_{1,0} & \widetilde{\boldsymbol{B}}^{\mathrm{e}}_{1,1} & & \widetilde{\boldsymbol{B}}^{\mathrm{e}}_{1,K} \\ \vdots & & \ddots & \\ \widetilde{\boldsymbol{B}}^{\mathrm{e}}_{K,0} & \cdots & & \widetilde{\boldsymbol{B}}^{\mathrm{e}}_{K,K} \end{pmatrix}$$

(e.g. Desroziers et al 2014)

EDA



From large-scale 4D-Var (currently operational for ARPEGE) **to 4DEnVar at high resolution** (soon operational for AROME ; 4DEnVar is also experimented & considered for ARPEGE) :

- Severe **simplifications in TL/AD physics** of 4D-Var become more prominent at high resolution ;

- **Scalability issues** of TL/AD model integrations ;

- **4DEnVar** allows 4D error covariances to be derived at **high resolution,** from an EDA, without TL/AD models ;

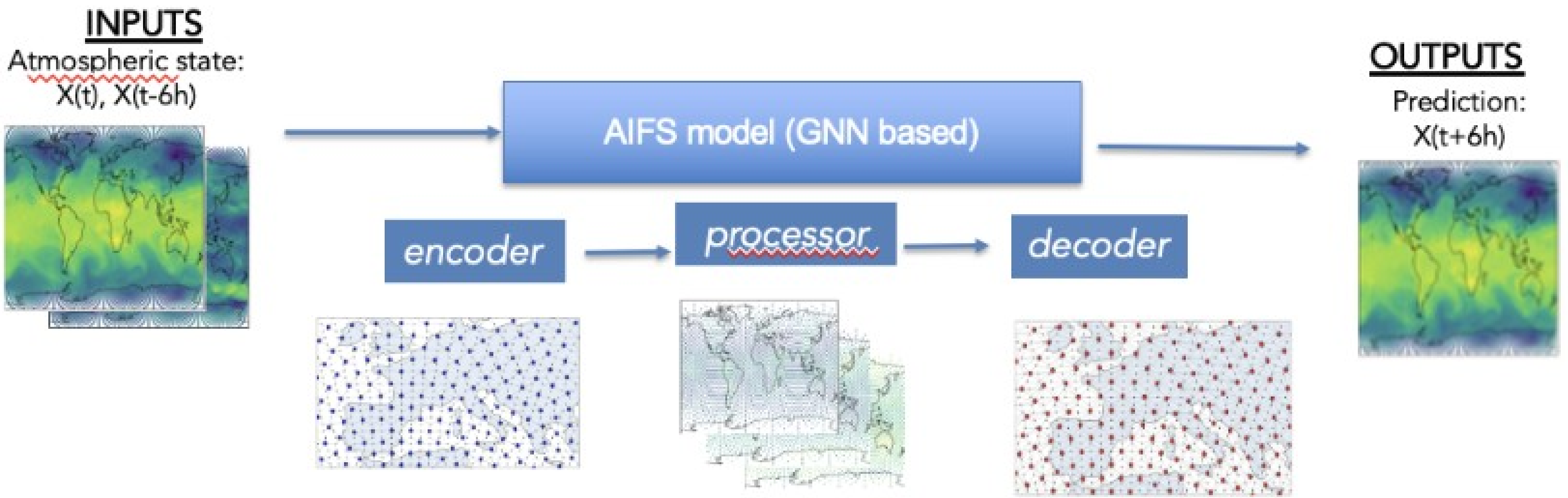
- 4DEnVar also easens extension of DA to **hydrometeors, surface DA and coupled DA**.

Development of Machine Learning emulators for NWP

- Data Assimilation produces **a large time series of consistent atmospheric states** on a regular geographical grid.

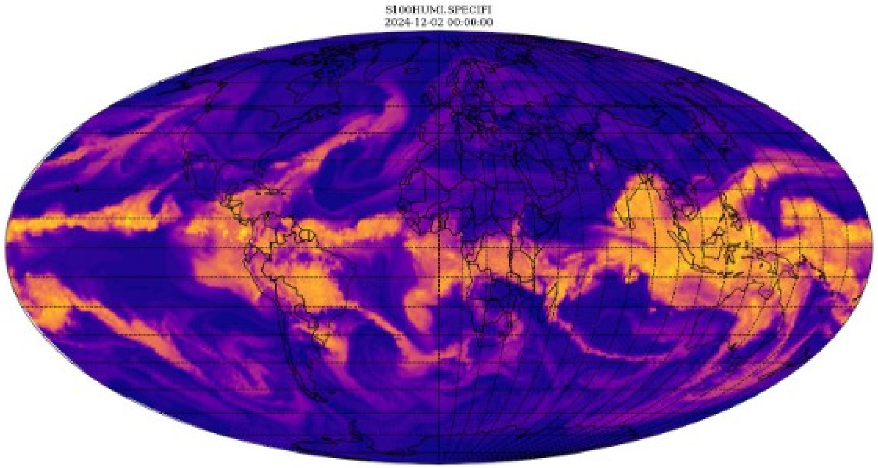
- Such data sets provide an ideal framework for Machine Learning : neural networks (emulators) can be trained to **learn how the atmosphere evolves** from one timestep (t) to the next one (e.g. t+6h).

- Examples : AROME-IA to emulate AROME, and ARPEGE-IA for ARPEGE, at Météo-France ; AIFS for the IFS physics-based model at ECMWF.



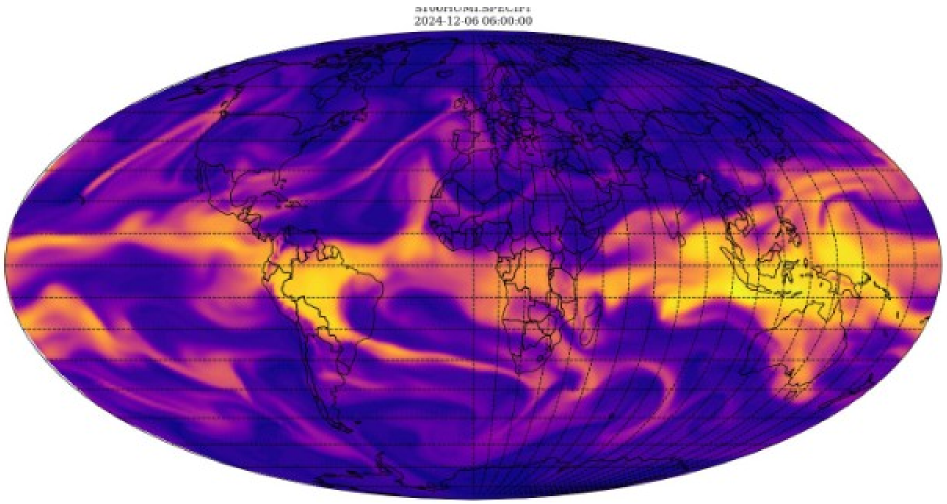
Use of Machine Learning emulators for NWP

- Such model emulators are also developed for AROME and ARPEGE, taking advantage of data sets of high resolution DA fields.
- Emulators tend to have **better accuracy in predicting large-scale features**, but tend to be **too smooth** and they predict a rather limited number of variables.



ARPEGE analysis (initial state)

of specific humidity near the surface



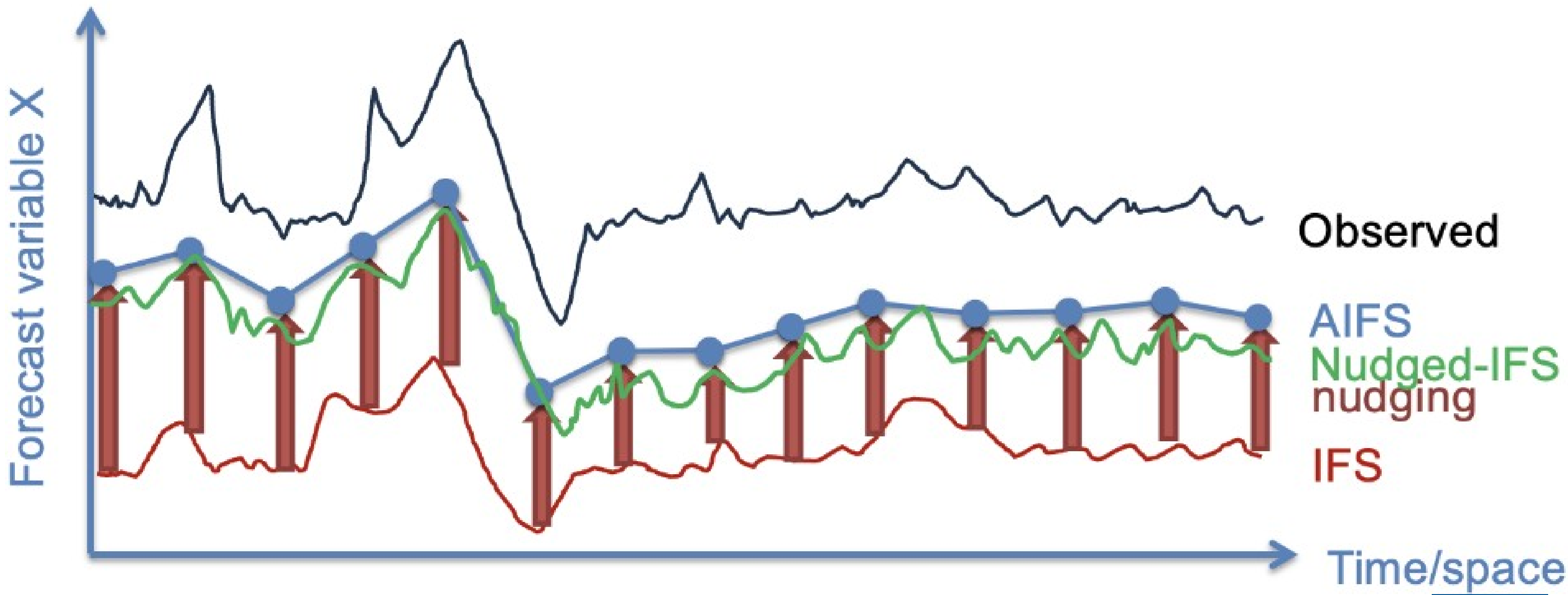
AIFS 3-day forecast

of specific humidity near the surface

METEO FRANCE

Use of Machine Learning emulators for NWP

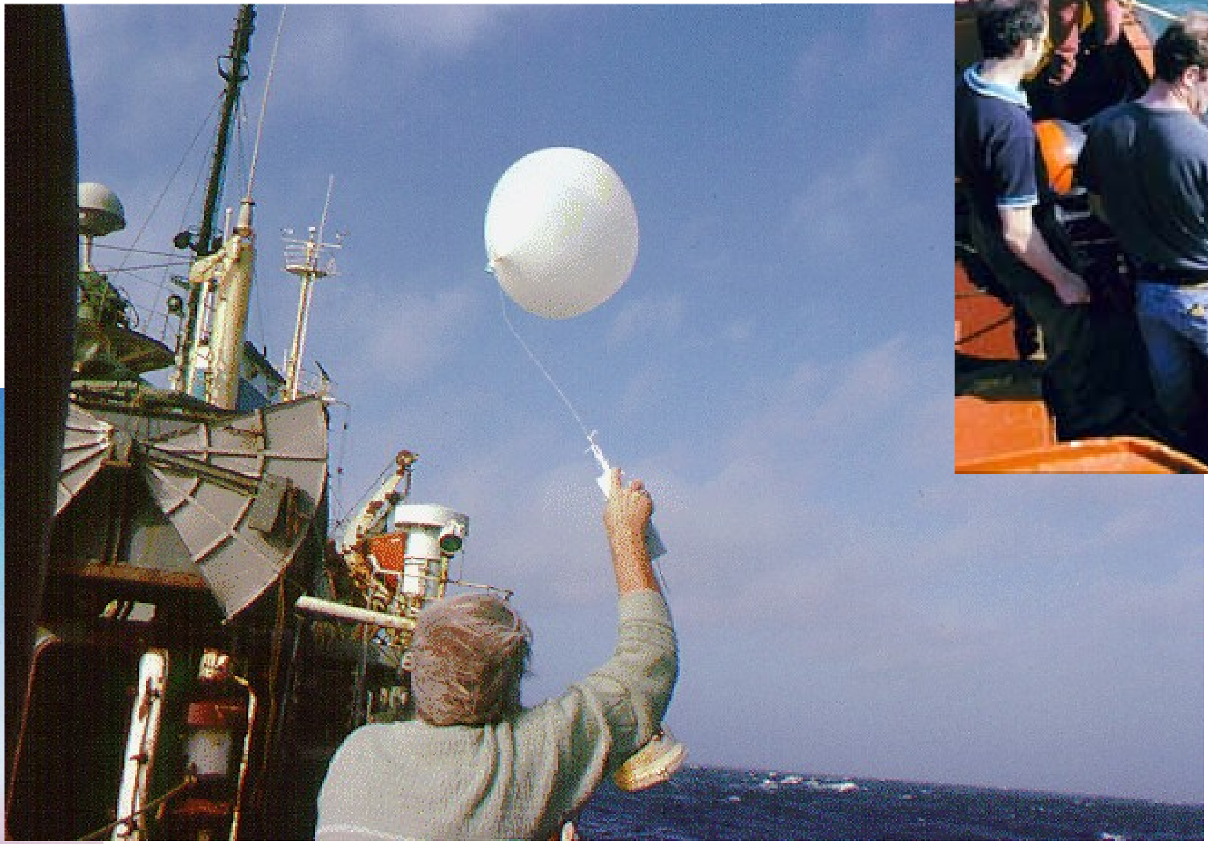
- Emulators can be combined with physics-based NWP models through **large-scale nudging** (i.e. as an additional forcing term in the model equations), in order to take advantage of both IA- and physics-based approaches.



2. In-situ observations and remote sensing data

Observation networks in meteorology : in situ measurements

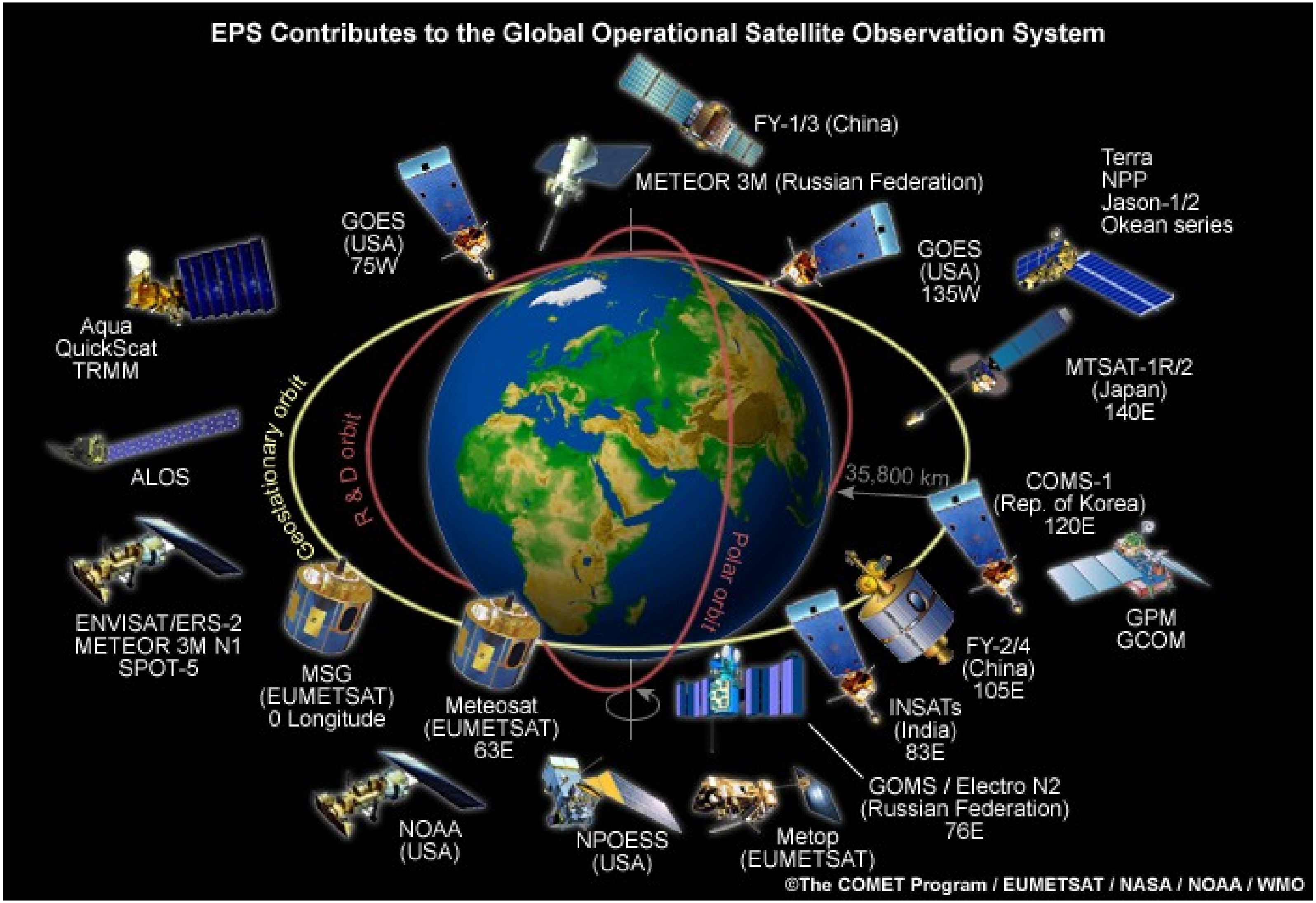
Provide direct information on the atmospheric state
at the instrument location.



* 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).

METEO FRANCE

Observation networks in meteorology : satellite data

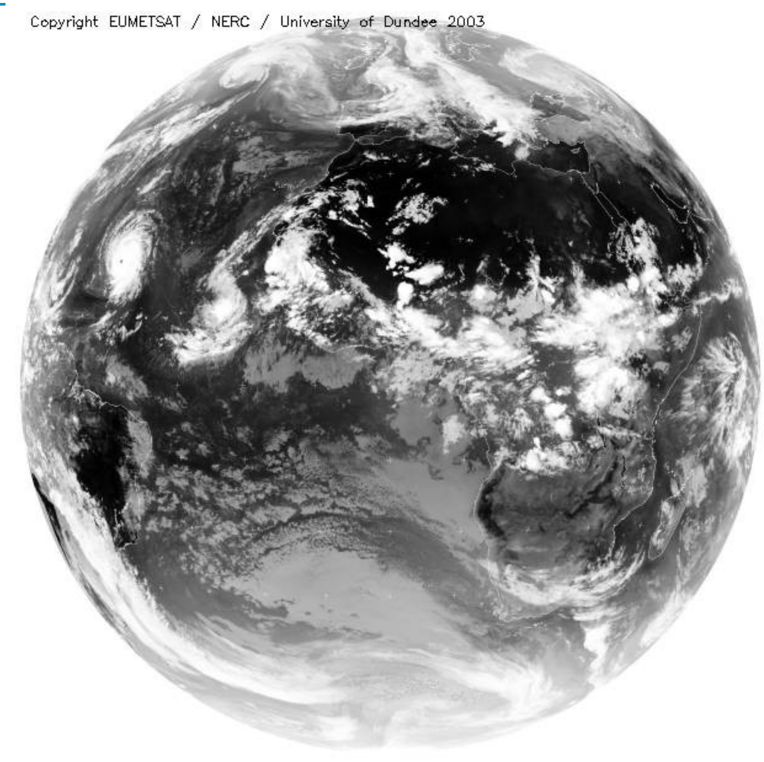


Constellation of polar orbiting or geostationary satellites

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).

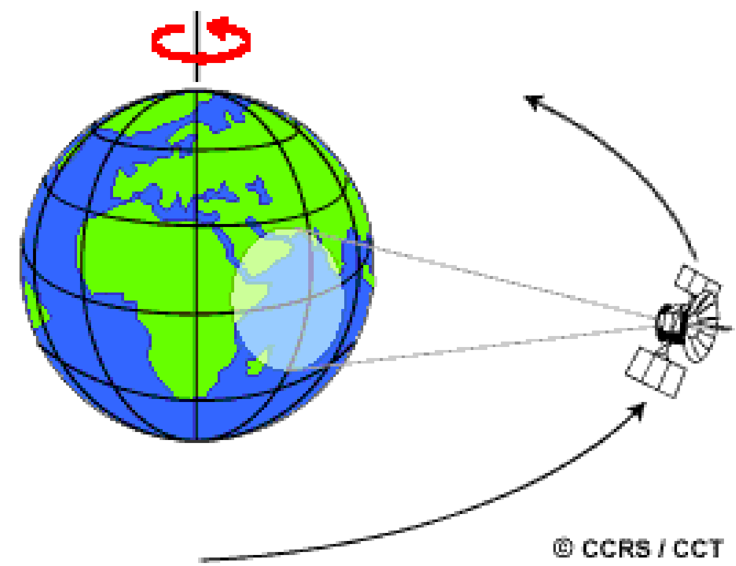
  Useful for nowcasting
  (= very short range forecasts, e.g. within the next 2 hours).

  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 (~10 km).

  Global spatial coverage (twice a day)

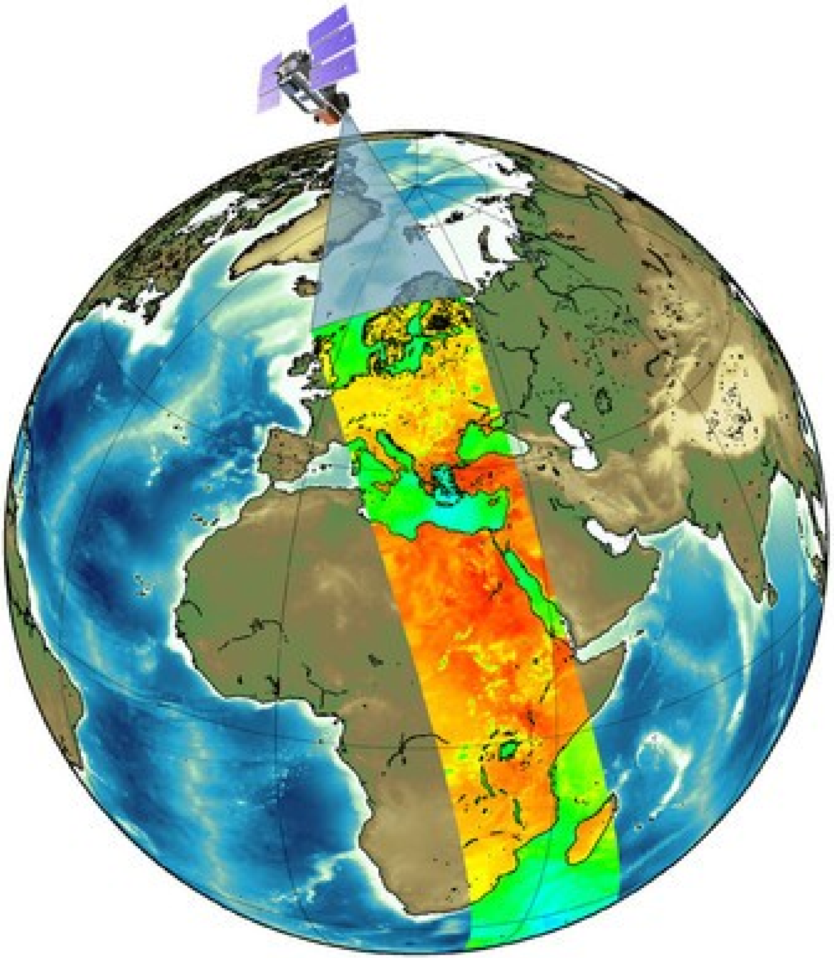
  Sounding instruments
  (over several vertical layers) :

  ~ vertical profiles of T at different locations

- Drawbacks

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

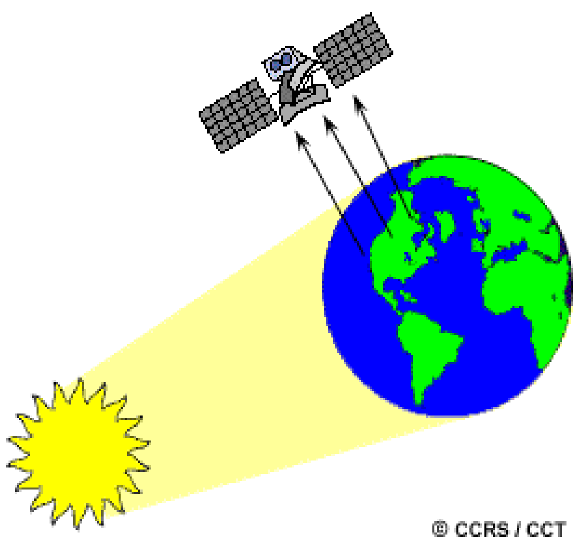
  (several satellites are needed, to have
  frequent observations over the same area)



Two types of satellite measurements

Passive measures

(no energy is emitted from instrument)

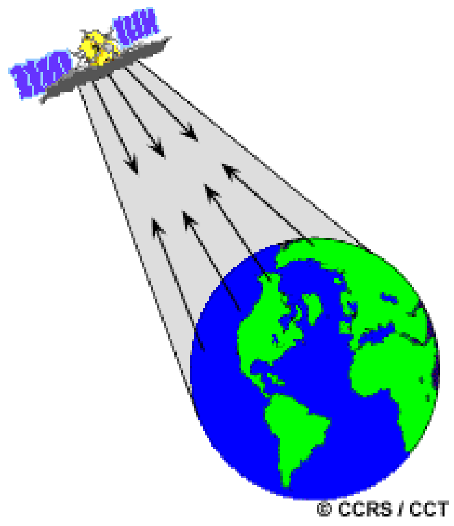


Measures natural radiation emitted

by Earth or Atmosphere (with Sun origin)

Active measures

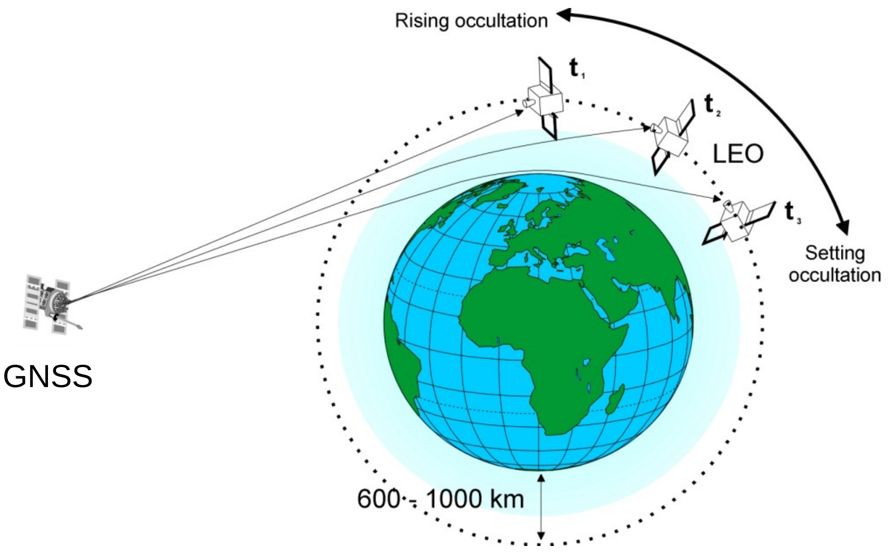
(energy is emitted from instrument)



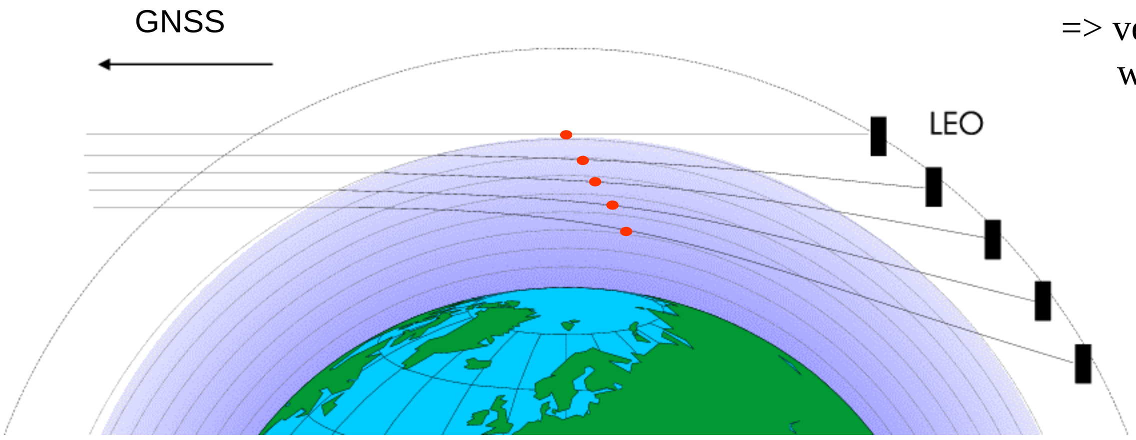
Measures radiation emitted by satellite and then

reflected or diffused by Earth or Atmosphere

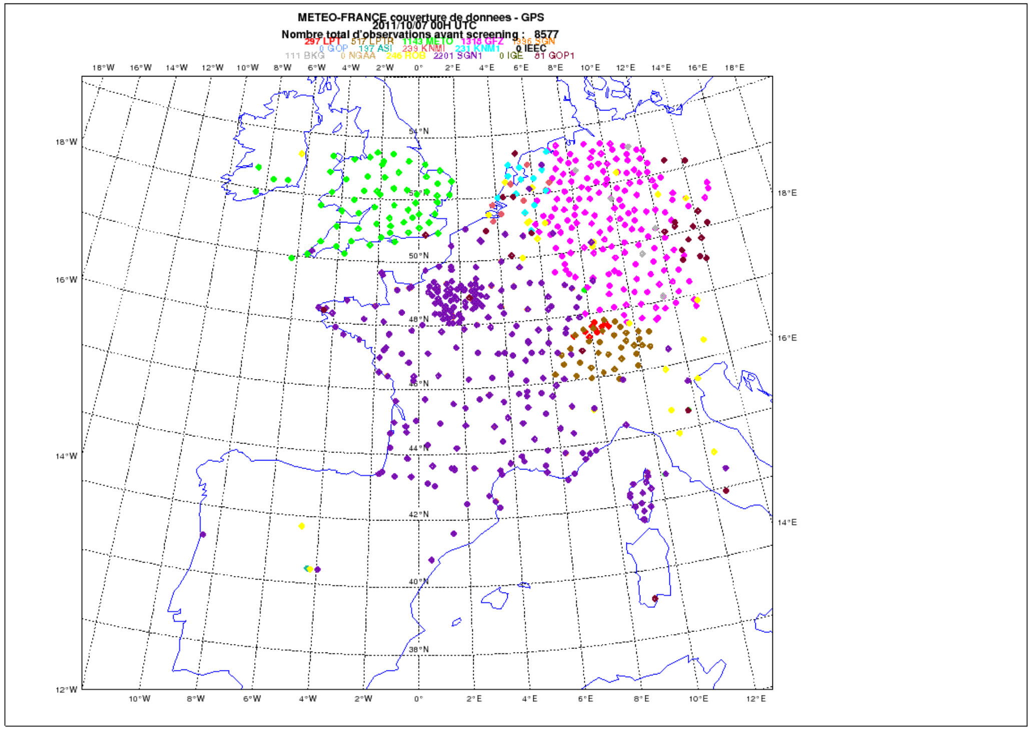
GNSS radio-occultation data
(1st example of active remote sensing)



- GNSS 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)

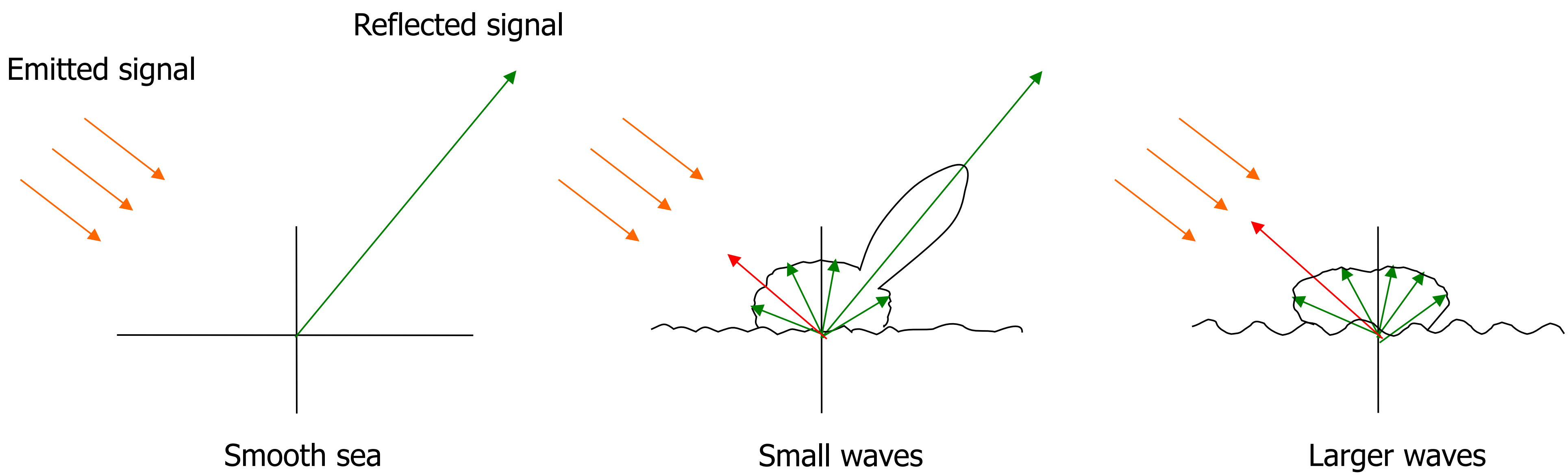


- Propagation of GNSS signal is slowed by atmosphere (dry air and water vapour) : the propagation delay provides information about humidity in particular.

- 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.

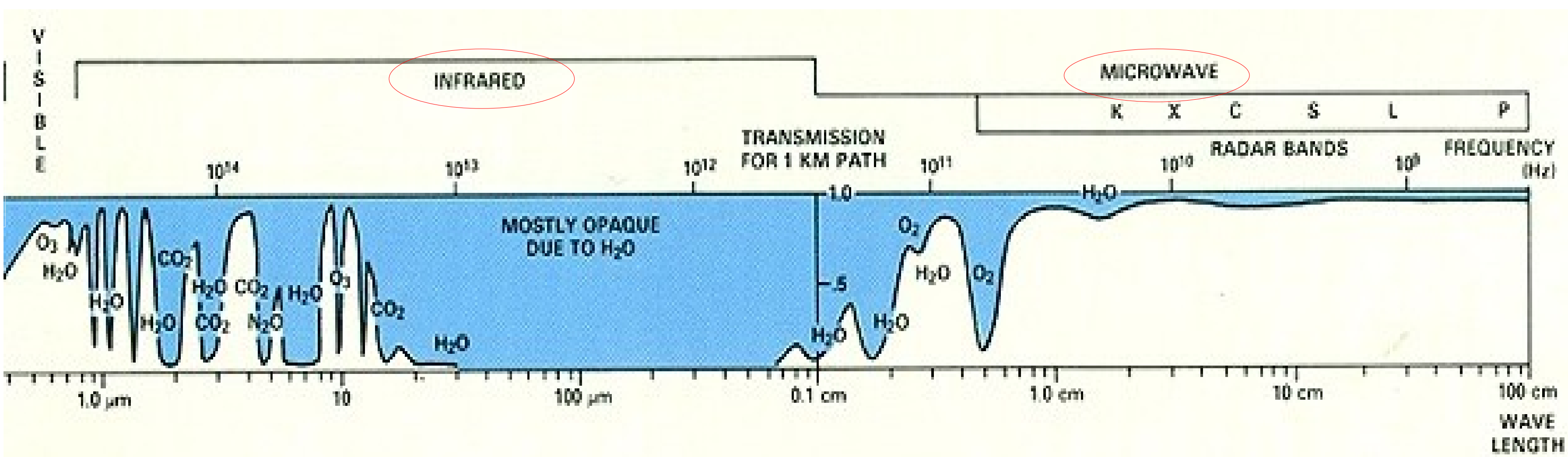
- Depending on **wave length**, indirect information on gas concentration (e.g. humidity) or on physical properties of atmosphere (temperature or pressure).

- 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.

100 %

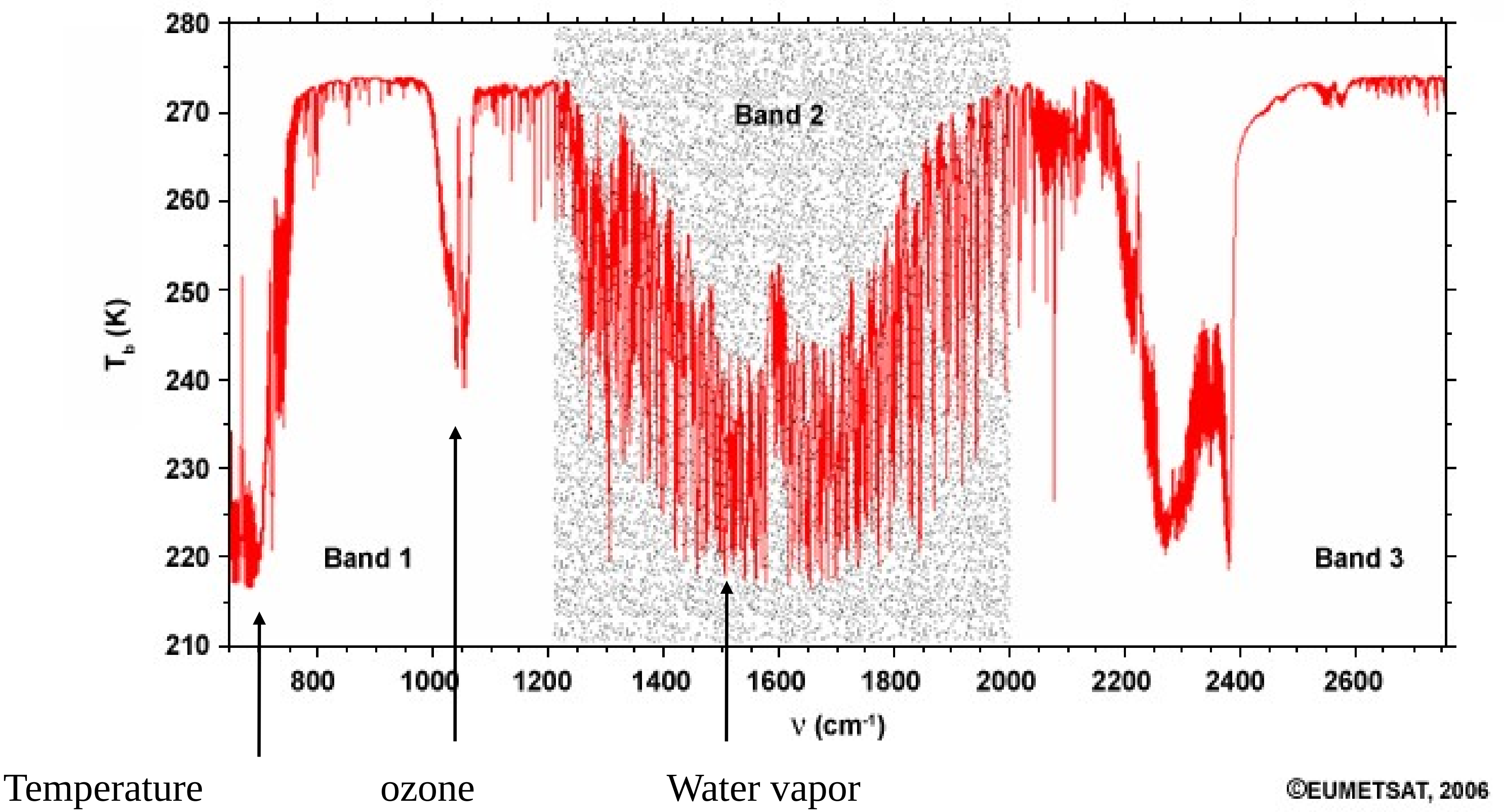
Atmospheric Transmittance

0 %



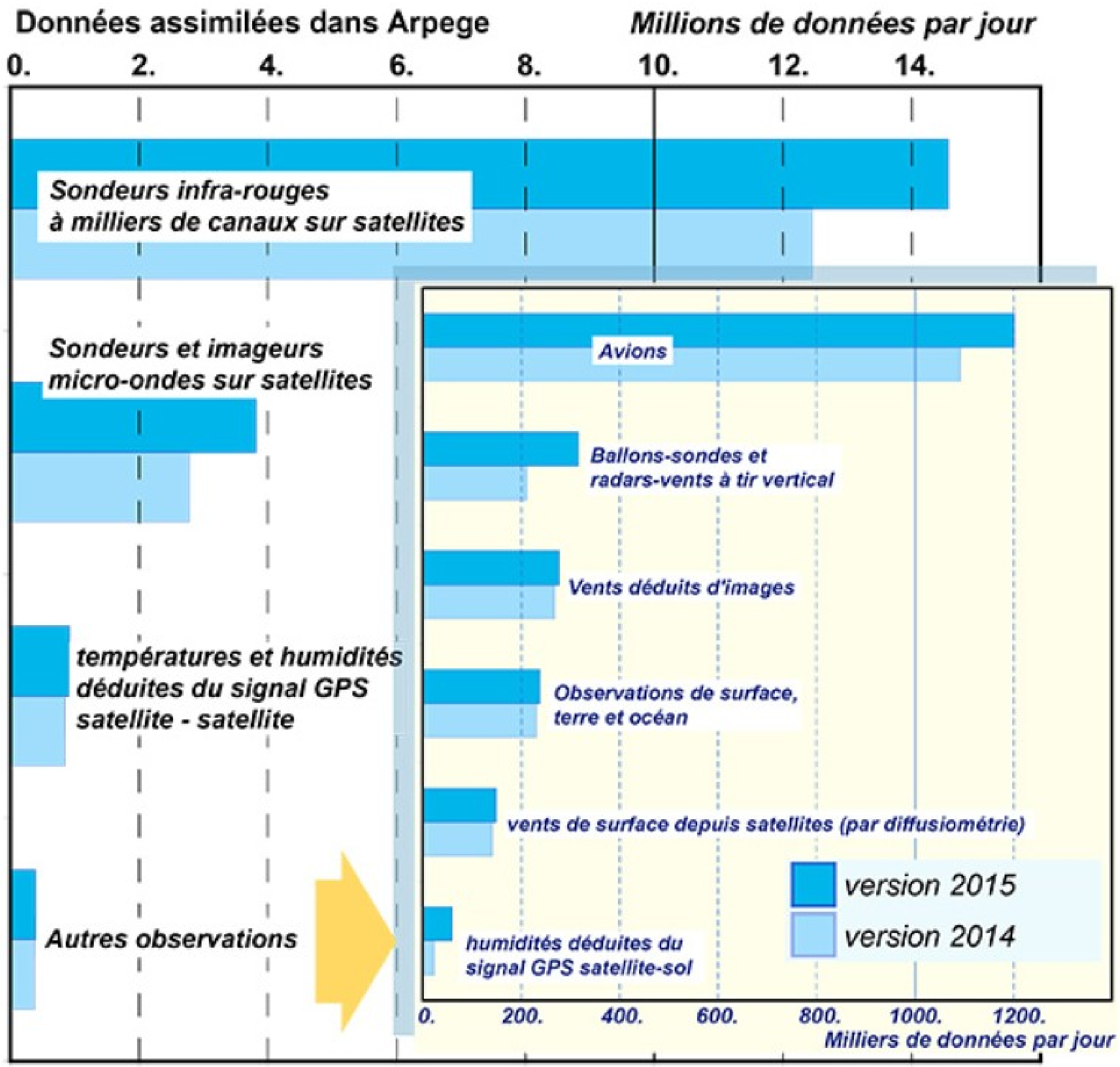
IASI : infra-red interferometer developed by CNES and EUMETSAT

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



METEO FRANCE

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



Total ~ 20 million obs per day

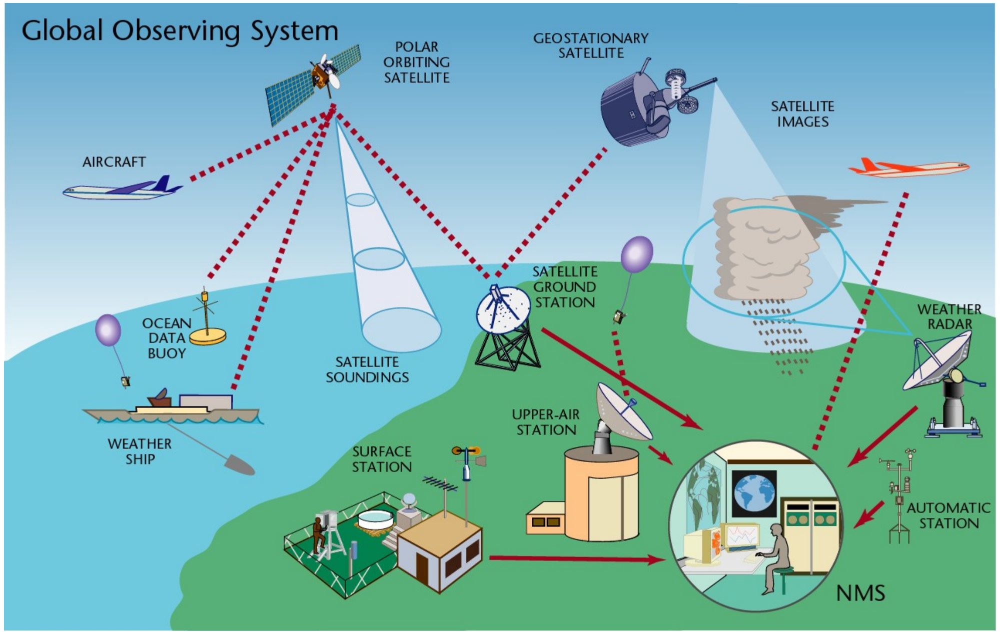
How do observations meet global NWP requirements ?

- **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.

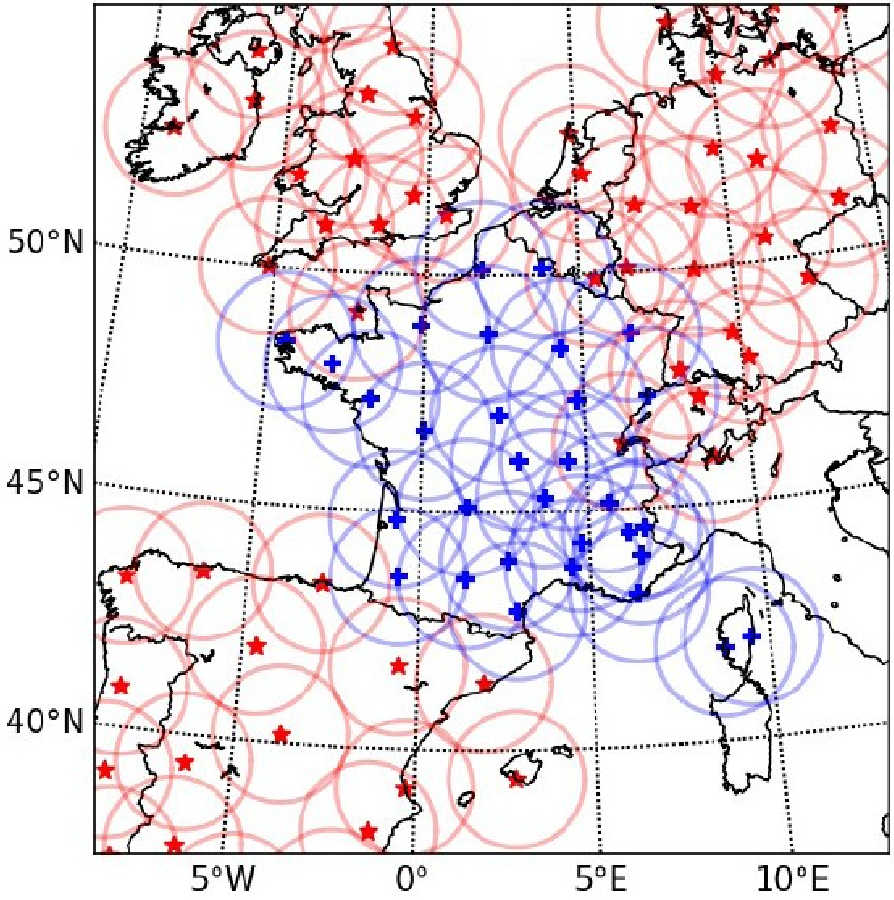
- Observations :

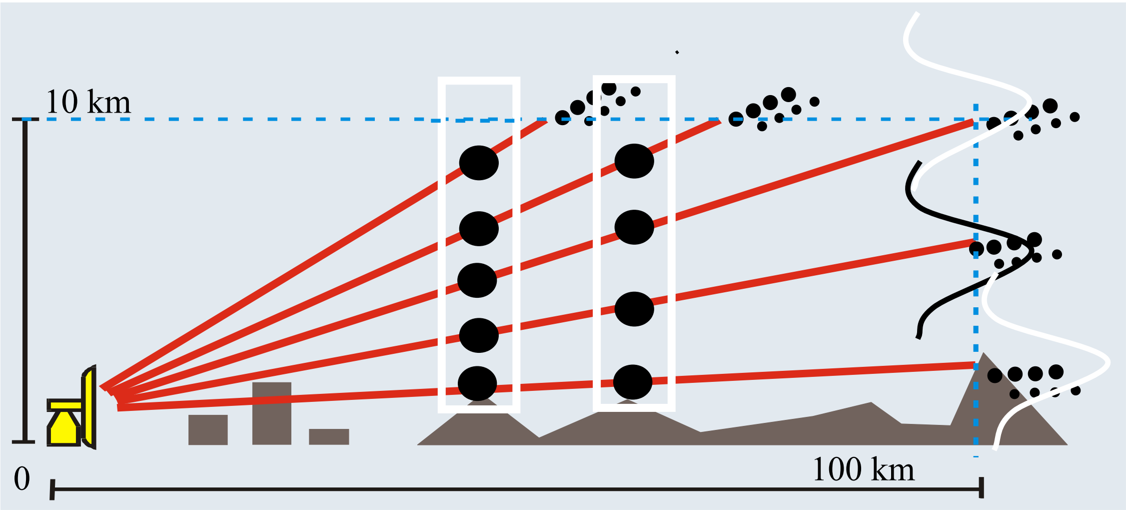
reflectivities Z (related to precipitation) ;

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.





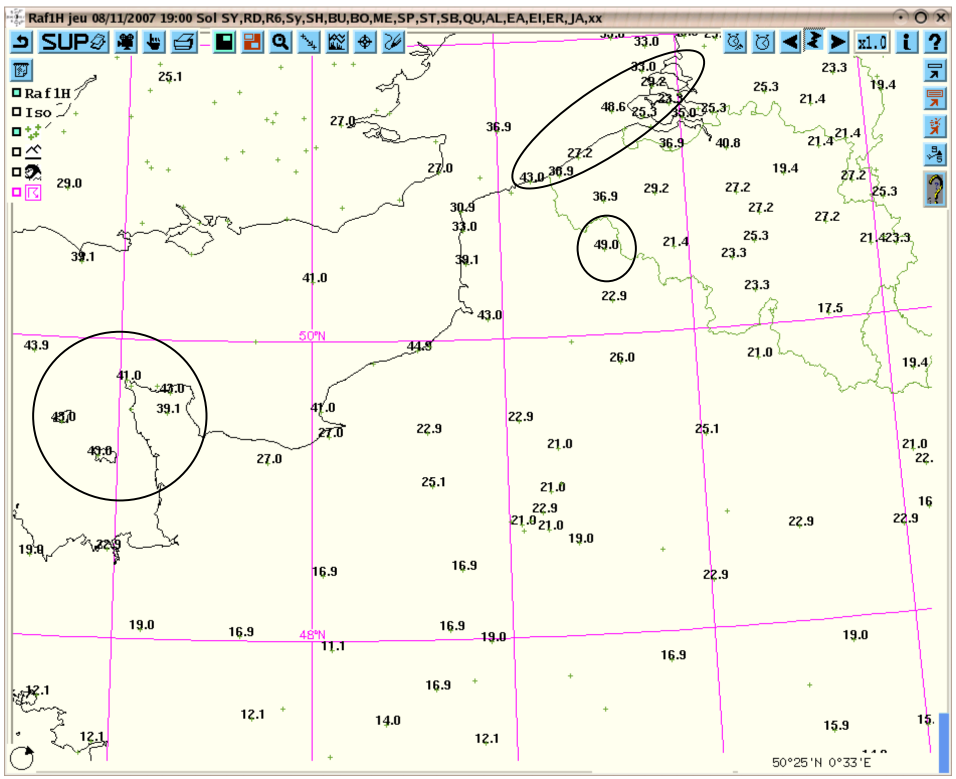
Observations assimilated as vertical profiles, after estimating the pixel altitudes

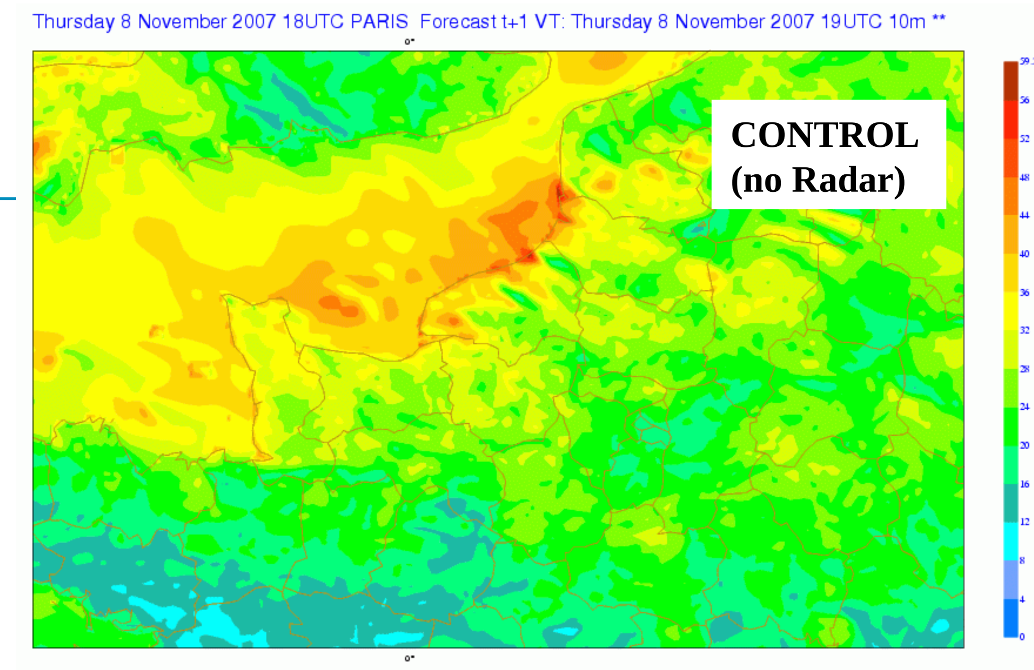
Assimilation of radar radial winds

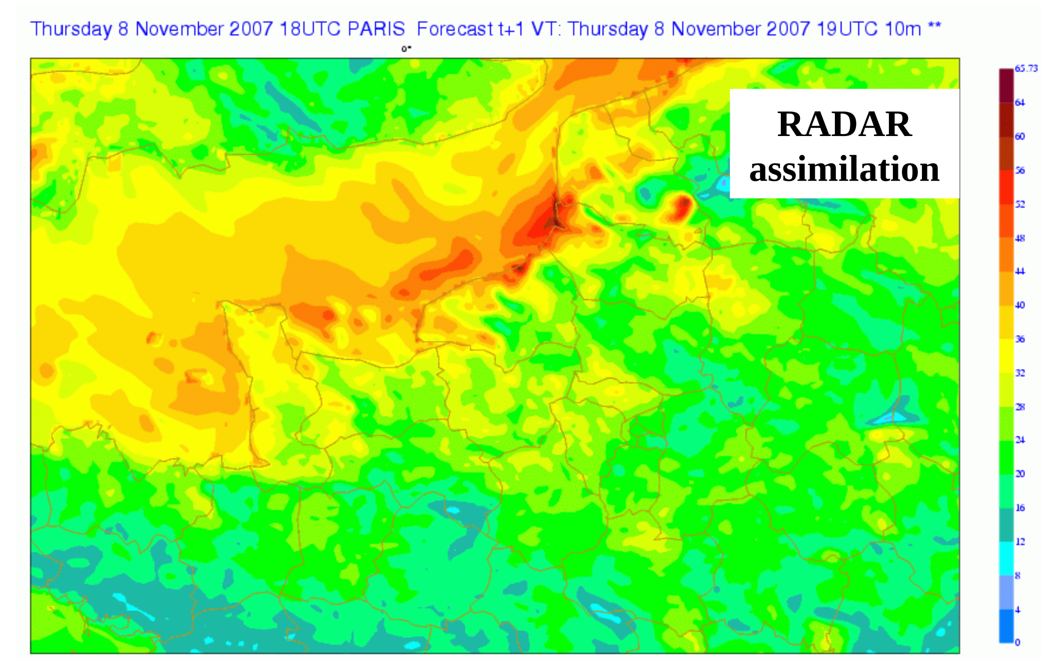
Wind gust at 10 m (kt)

Forecast +1h (19 UTC)

OBS



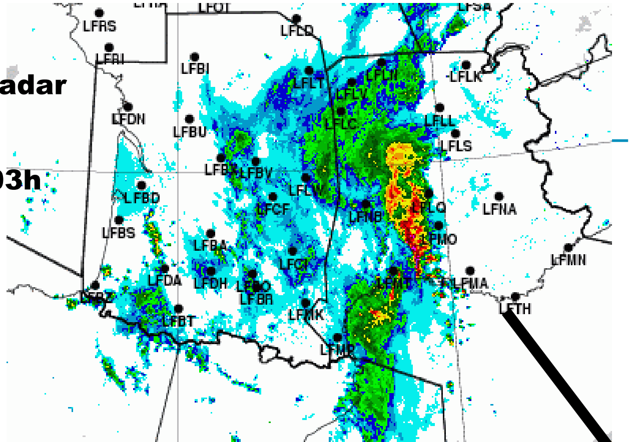




Impact of radar reflectivities

Radar

03h

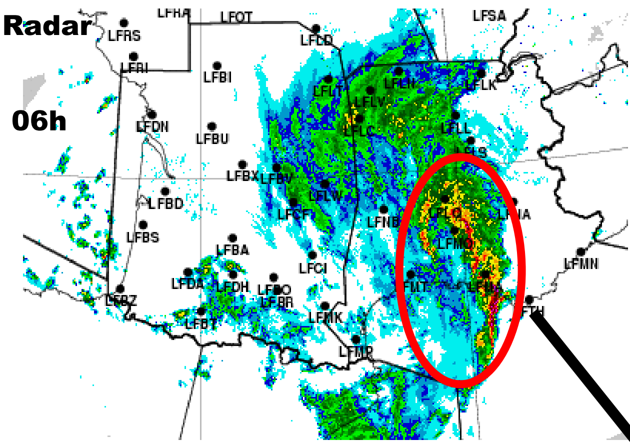


Example of Heavy Precipitation Event (South-East of France) :

REFL (assimilation of reflectivities) vs **CONTROL** (no reflectivities in DA)

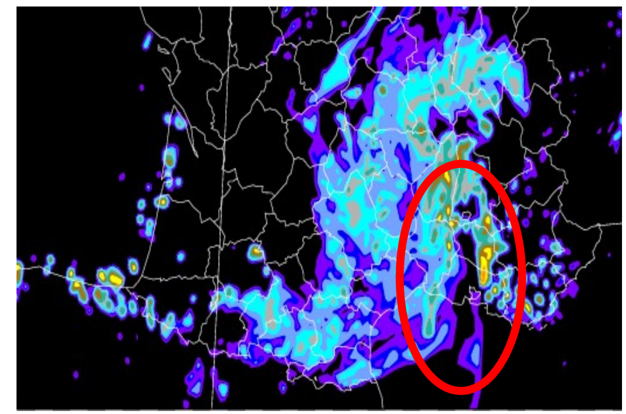
Radar

06h



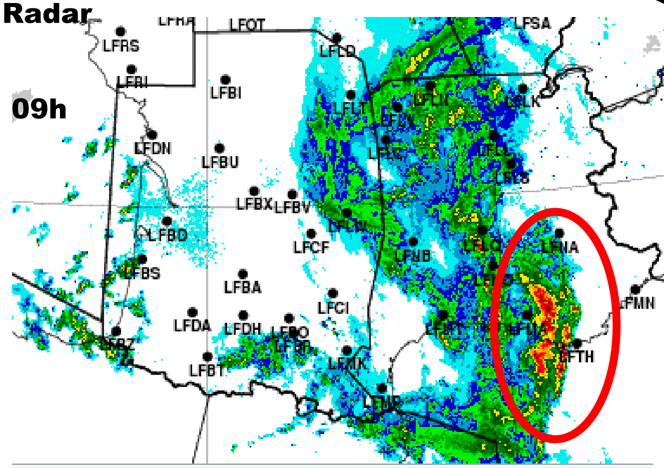
3h forecast issued at 3UTC

3h forecast issued at 3UTC

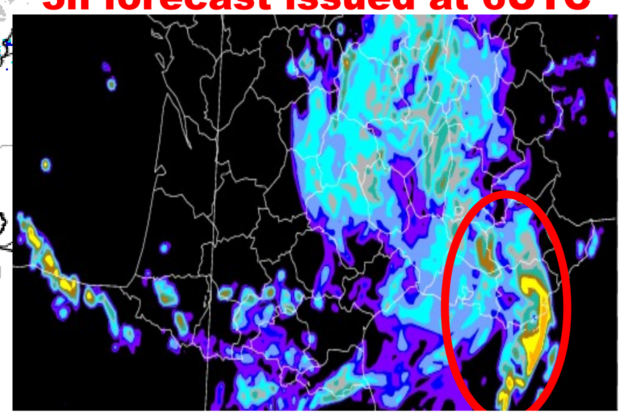


Radar

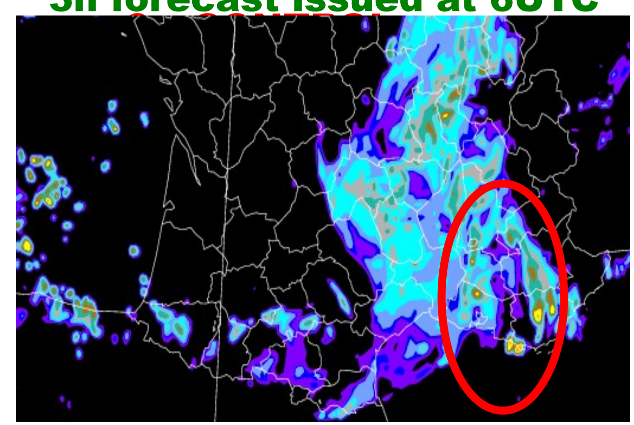
09h



3h forecast issued at 6UTC

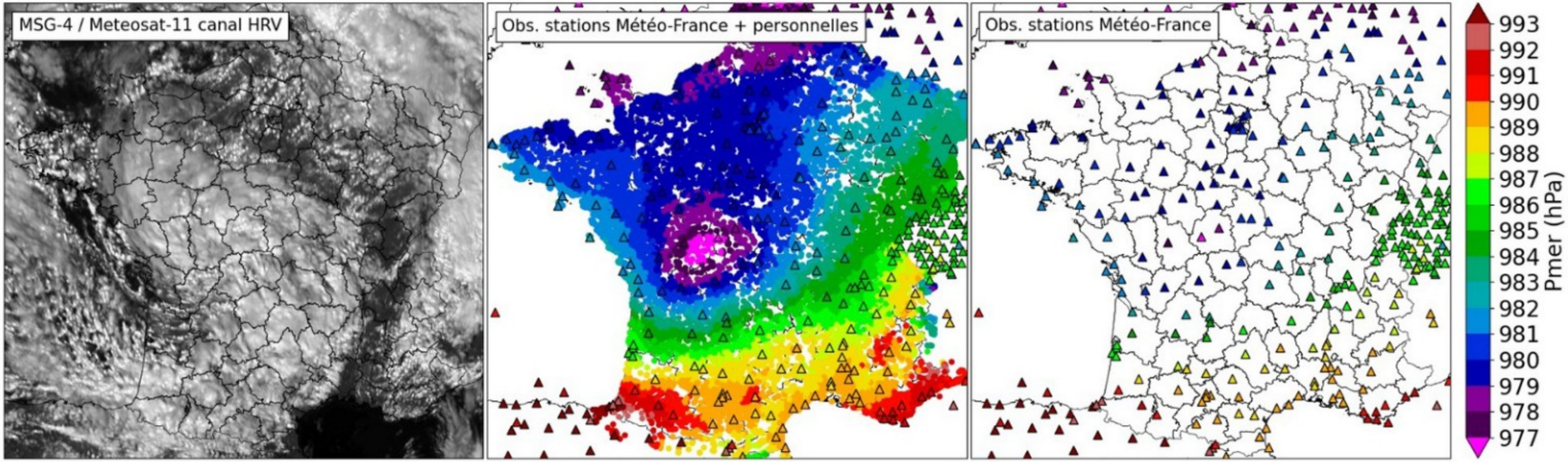


3h forecast issued at 6UTC



Line of heavy precipitation is well simulated in REFL run.

Personal Weather Stations



Storm Aline passing over mainland France on 20 October 2023 at 14:00 UTC. (a) Cloud roll observed by the visible 0.6 μm channel of MSG4 (b, c) Observations pressure from (b) personal stations combined with the network of conventional and (c) conventional stations alone. By combining the networks, it is possible to distinguish very precisely the location, shape and pressure gradients around the low-pressure minimum.

- Growing number of personal or agricultural weather stations connected to the Internet, transmitting observations in real time and at high frequency (< 10 min)
- Variable quality: strict automated quality control is essential, using WMO- standard observations
- Previously inaccessible meteorological structures are revealed (Mandement and Caumont, 2020).

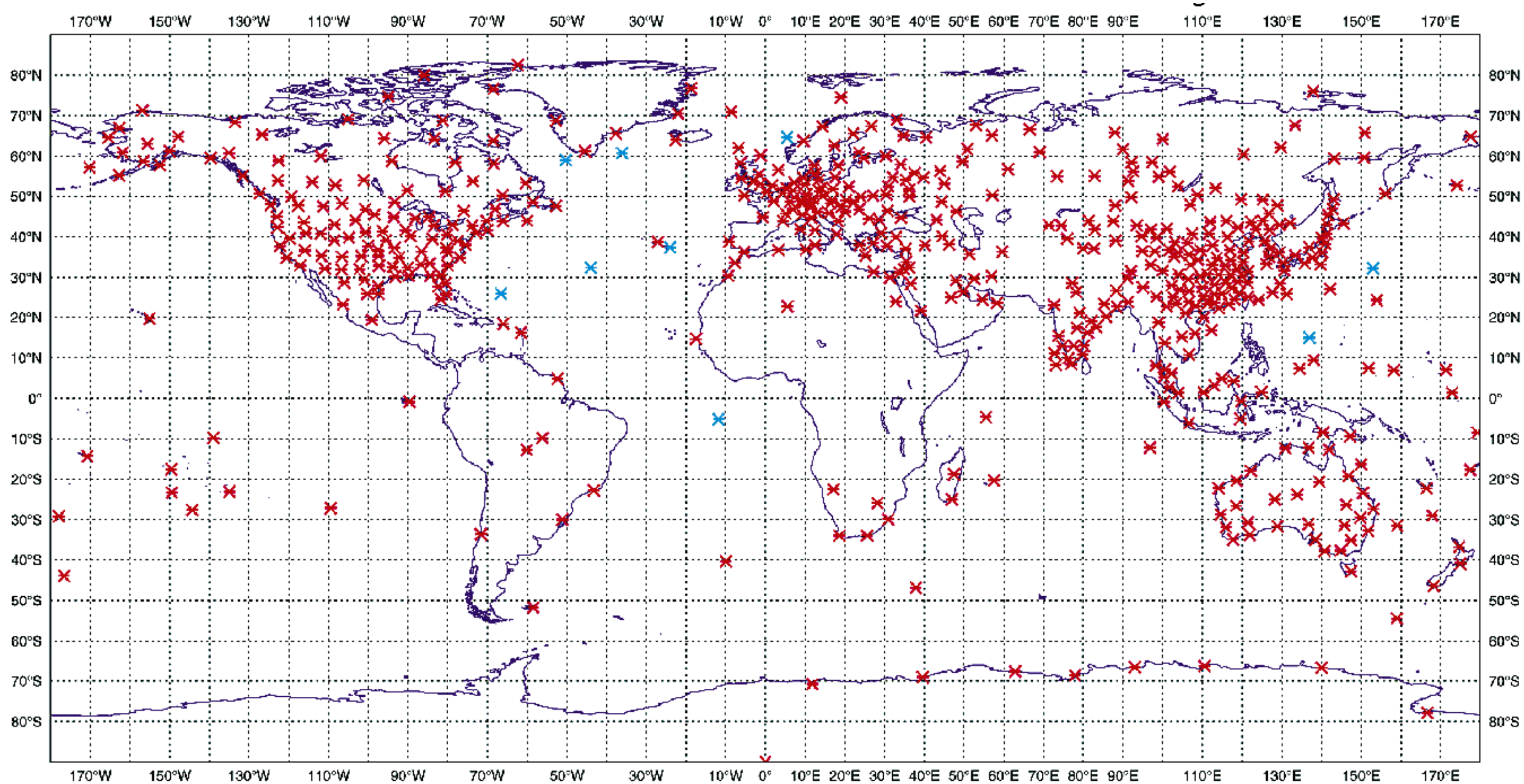
Error covariances : estimation and modelling (to weight and spatialise observed information)

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** approach 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.

METEO FRANCE

Radiosonde observation network



Covariances of innovations

- Innovations = observation-background departures :

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

$$\approx \mathbf{e}^o - \mathbf{He}^b$$

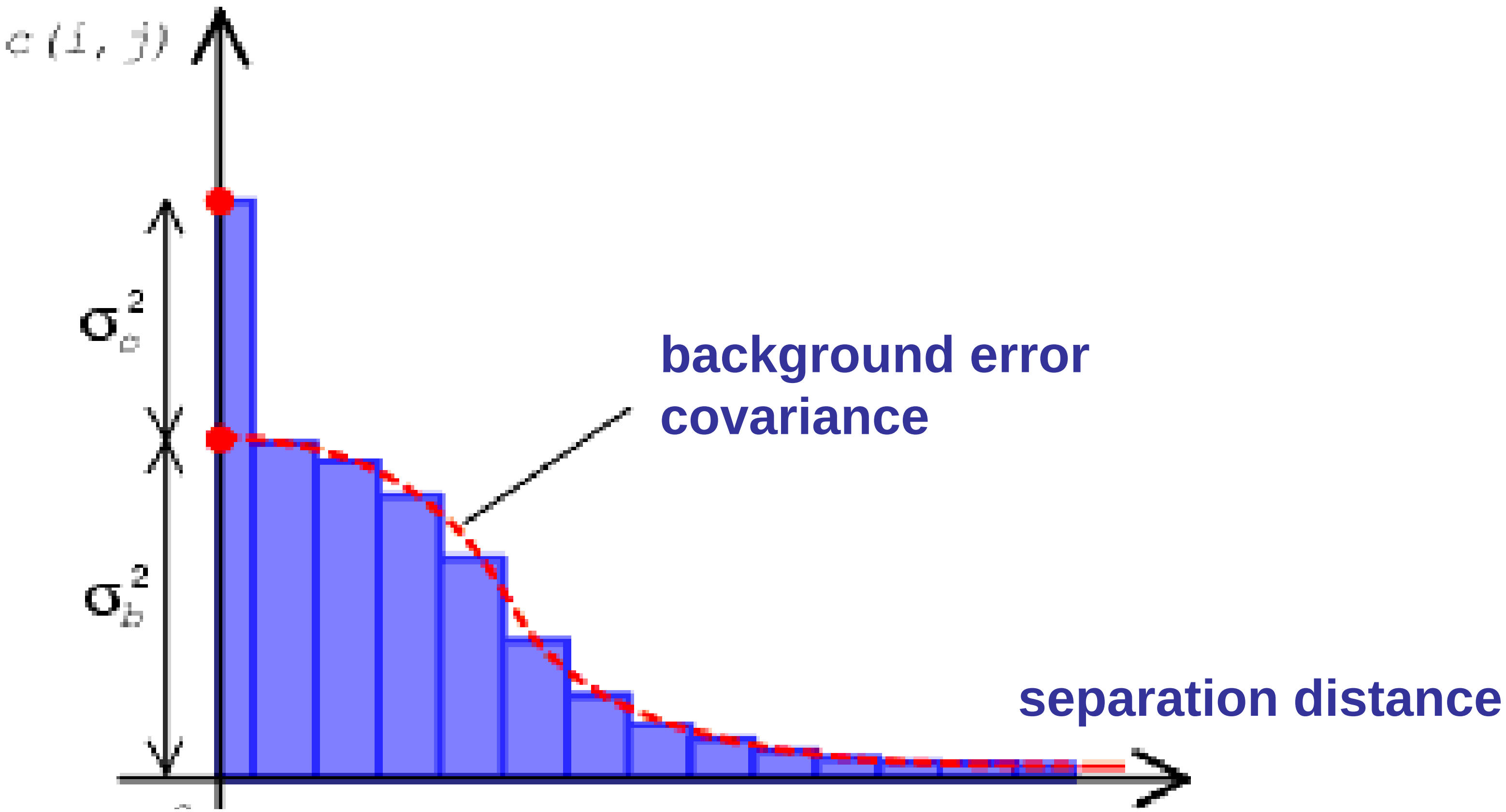
- Innovation covariances :

$$\mathrm{E}[(\mathbf{y}^o - H(\mathbf{x}^b))(\mathbf{y}^o - H(\mathbf{x}^b))^T] = \mathbf{R} + \mathbf{H}\,\mathbf{B}\,\mathbf{H}^\mathbf{T}$$

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

(e.g. Hollingsworth and Lönnberg 1986).

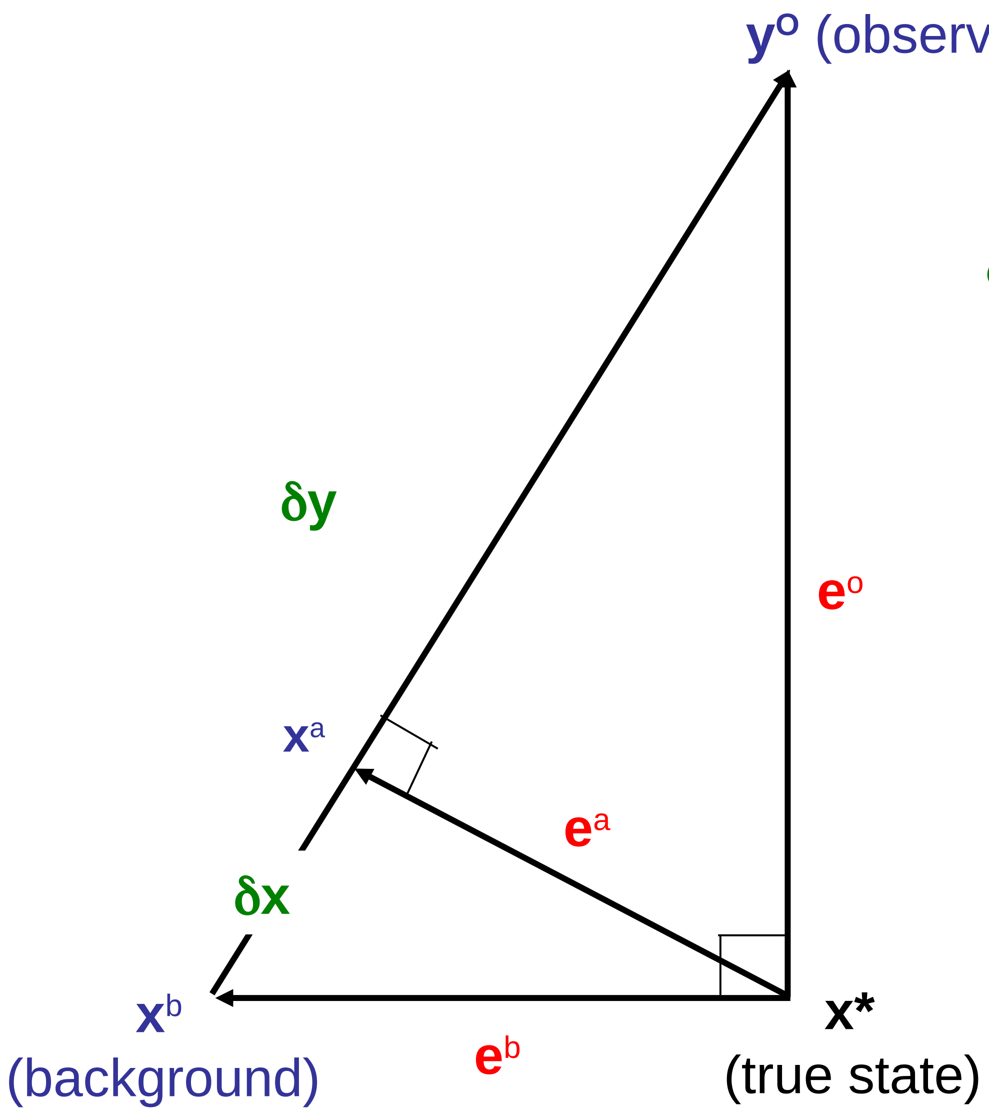
Covariances of innovations
(with extrapolation to zero separation distance)



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

Extrapolation to zero separation distance
allows different contributions to be estimated.

Covariances of analysis residuals



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

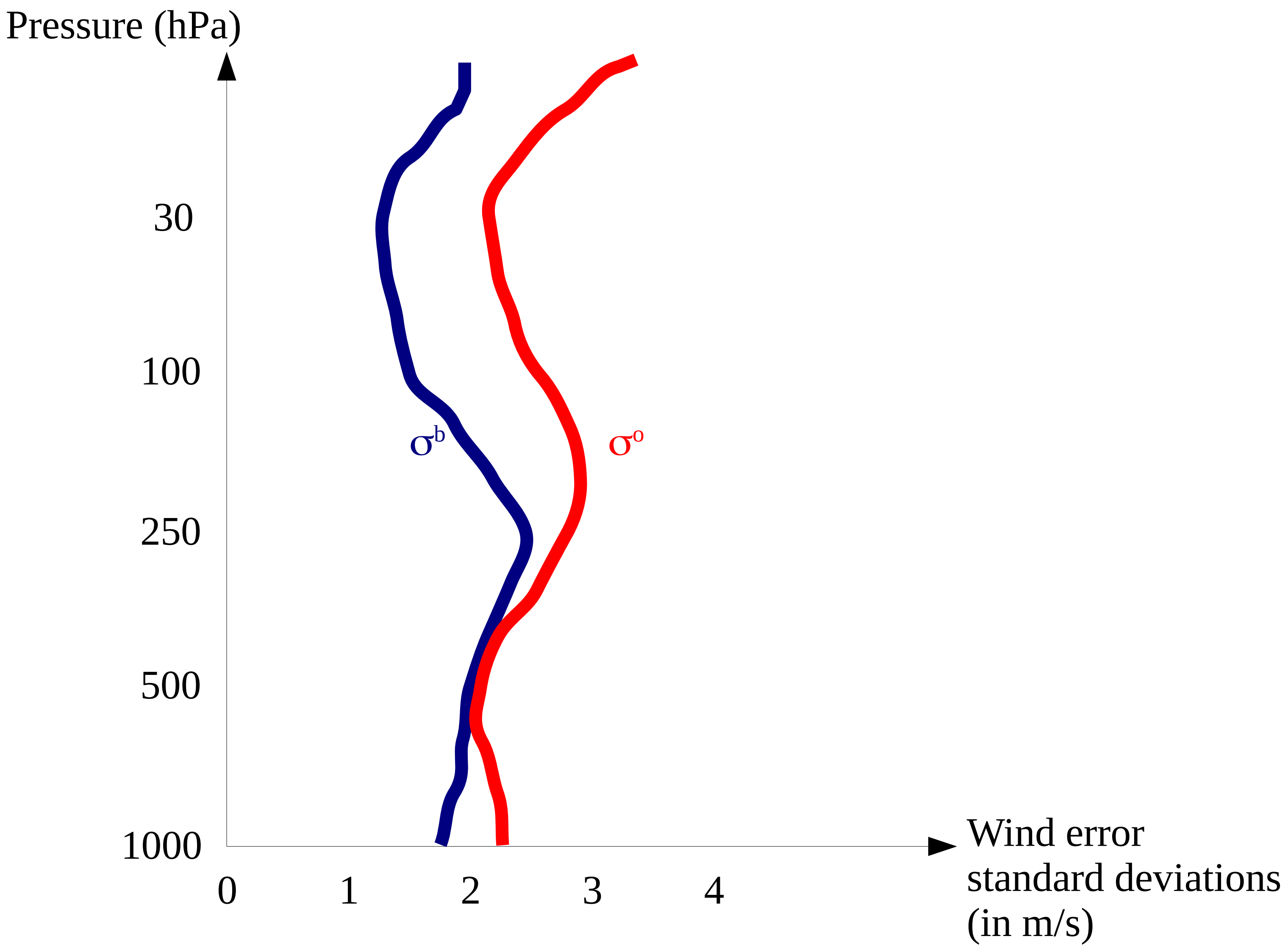
$\mathbf{H}\,\boldsymbol{\delta}\mathbf{x} = H(\mathbf{x^a}) - H(\mathbf{x^b})$ (increment)

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

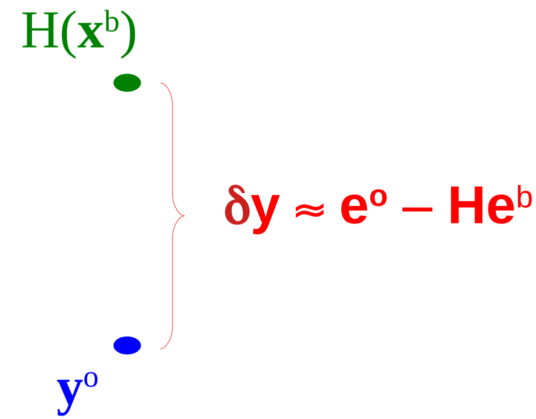
$E[(\mathbf{y^o} - H(\mathbf{x^a}))\ \boldsymbol{\delta}\mathbf{y}^T] = \mathbf{R}$

(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 $\delta \mathbf{y}$:

a single error realization is available locally

(e.g. for estimating background error variance).

=> Statistical averages (mathematical expectations)

need to be replaced by

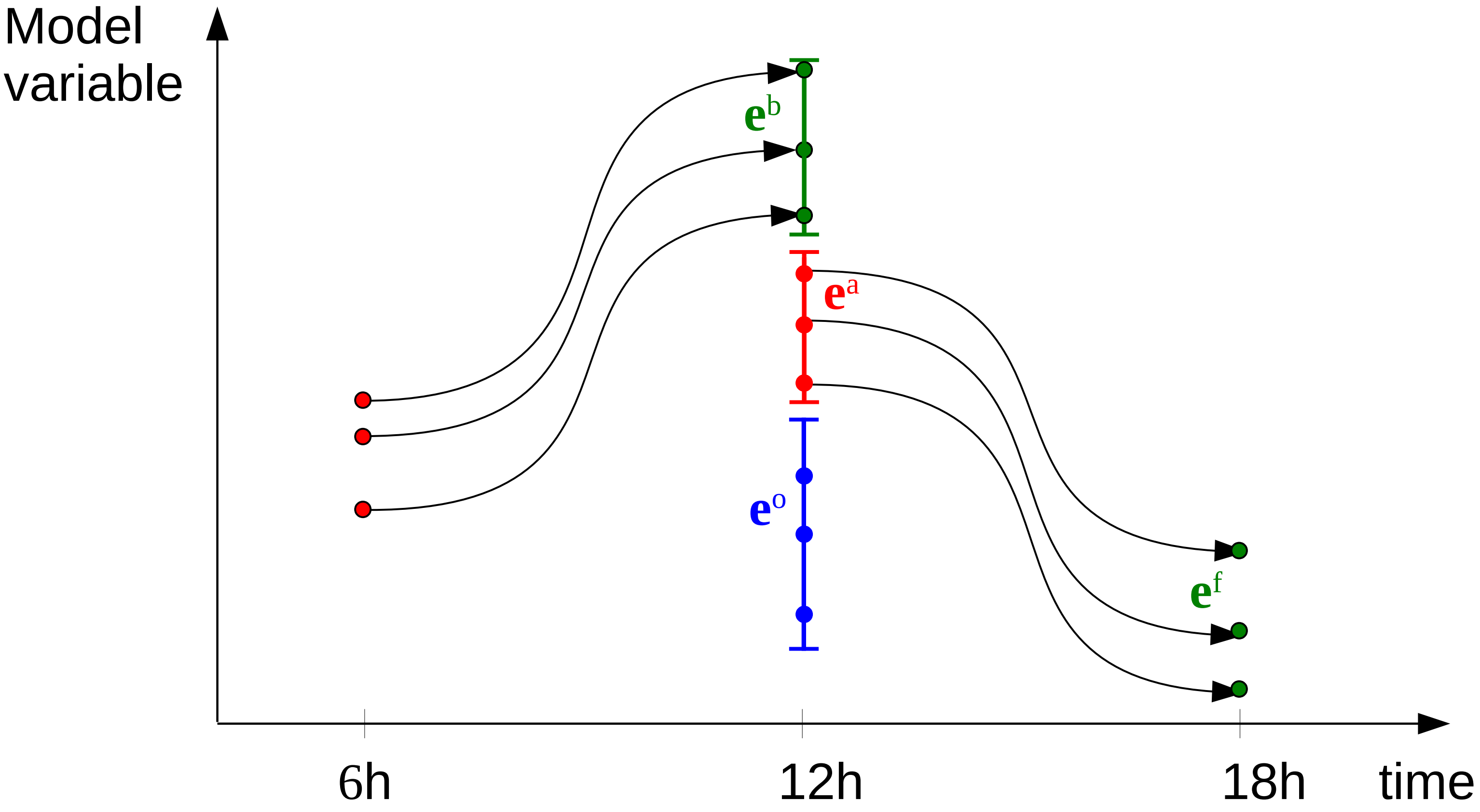
space and time averages (ergodic assumption).

=> 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



$e^a = (I\text{-}KH)e^b + Ke^o$

$e^f = Me^a + e^m$

with $e^b = e^{f-}$

and $e^o = R^{1/2}\,\eta$ (random draws of **R**)

(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** : $\mathbf{e}^o = \mathbf{R}^{1/2}\, \eta^o$.

- Model errors can be simulated in different ways, e.g. by :

  adding random draws of **Q** : $\mathbf{e}^m = \mathbf{Q}^{1/2}\, \eta^m$ (additive or mult. inflation);

  using a multi-model approach (or multi-physics) ;

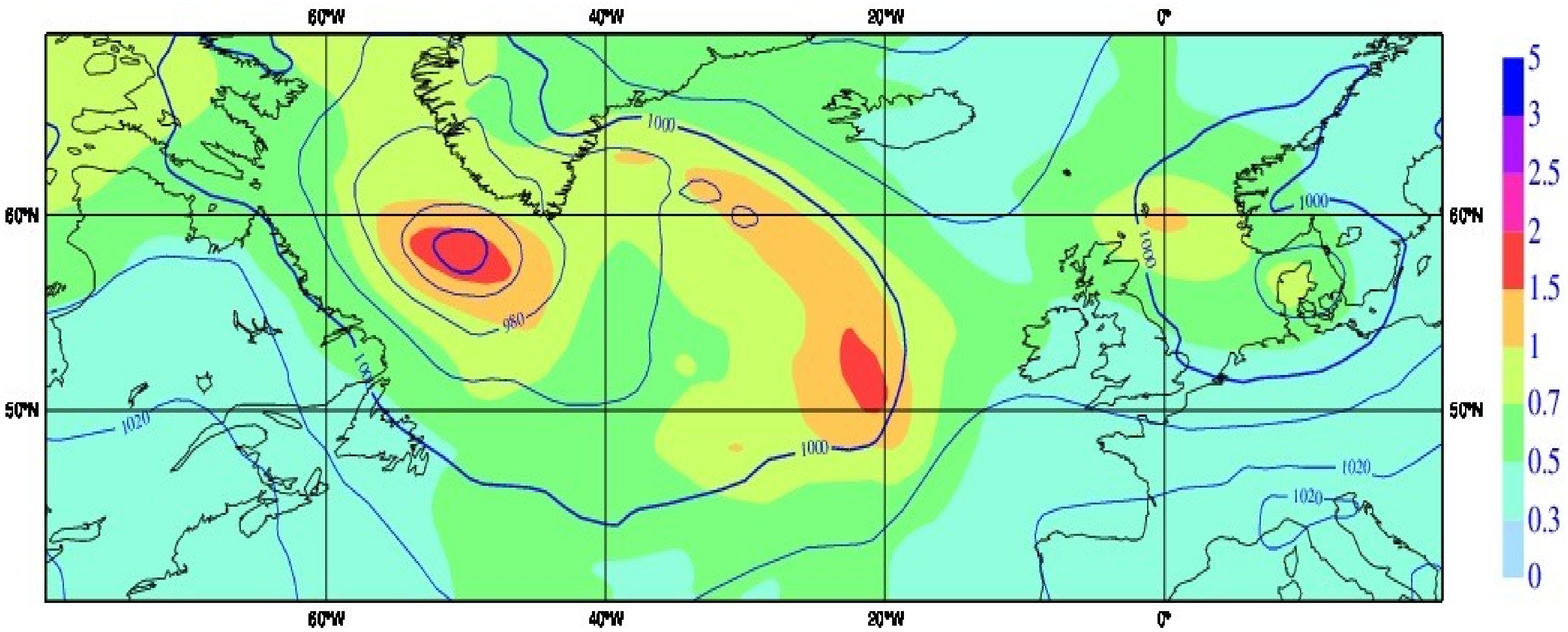
  perturbing physical tendencies of the model ;

  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 : $\mathbf{B} = \mathbf{U}\,\mathbf{U}^T$

  $\mathbf{U} = \mathbf{L}\,\mathbf{S}\,\mathbf{C_f}$

L ~ mass/wind cross-covariances (related to geostrophy),

including flow-dependence (non linear balances).

S flow-dependent standard deviations (~ expected error amplitudes),

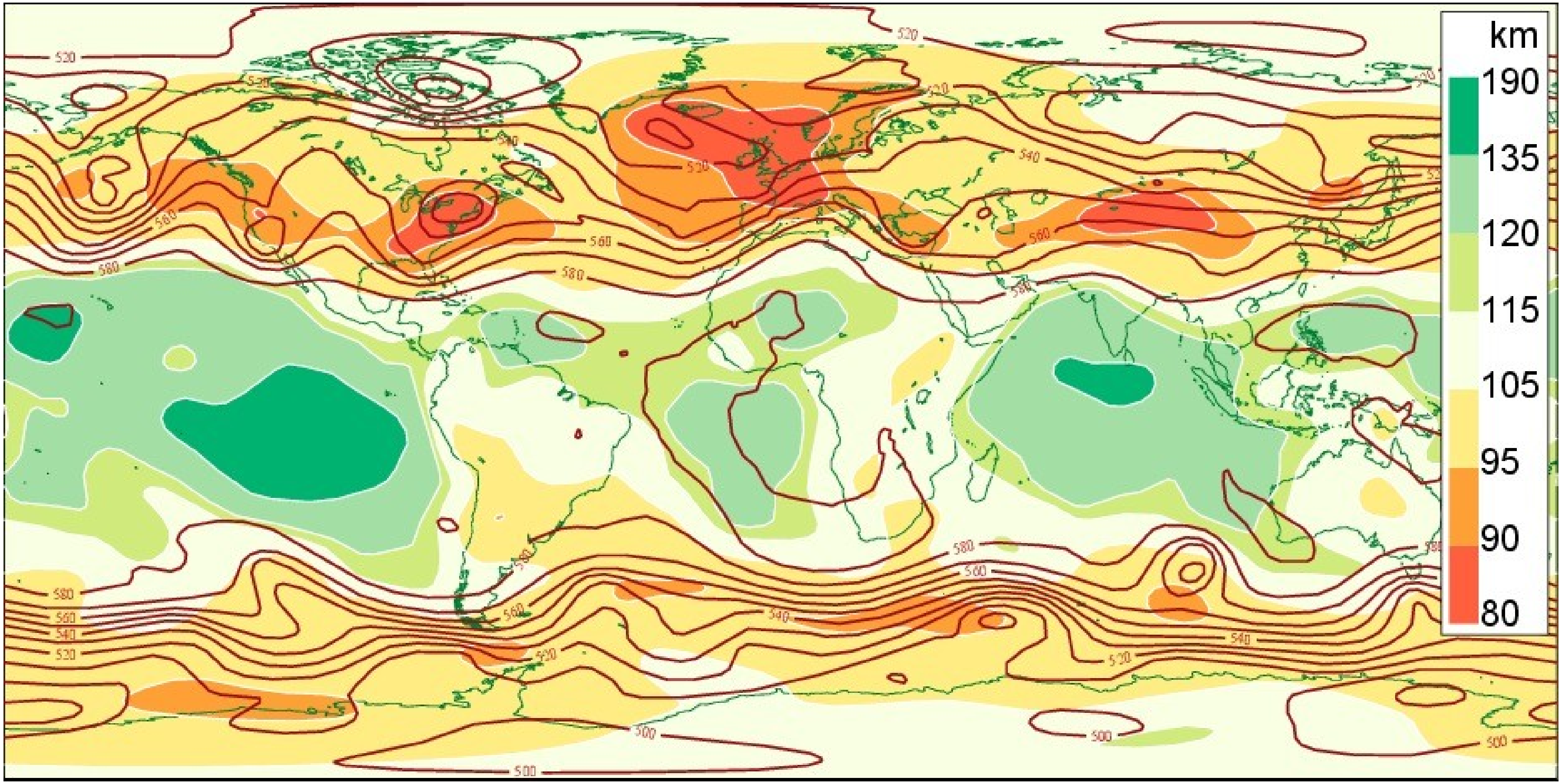
filtered spatially.

$\mathbf{C} = \mathbf{C_f}\,\mathbf{C_f}^T$ matrix of 3D spatial correlations (~ spatial structures of errors),

filtered in wavelet space (block-diagonal model).

METEO FRANCE

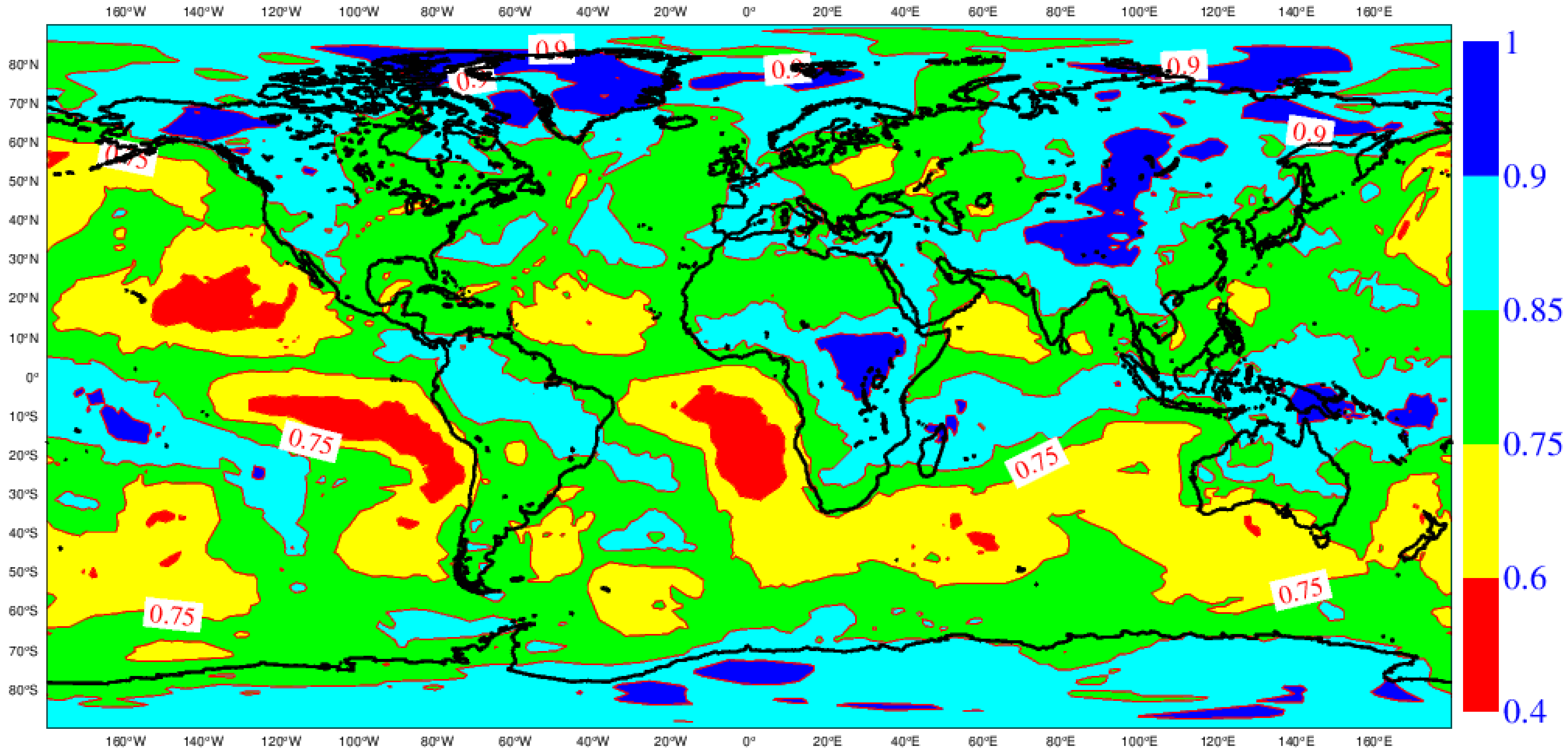
Dynamics of horizontal correlations



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

=> observations spatialised more locally in areas with large spatial gradients.

Dynamics of vertical correlations

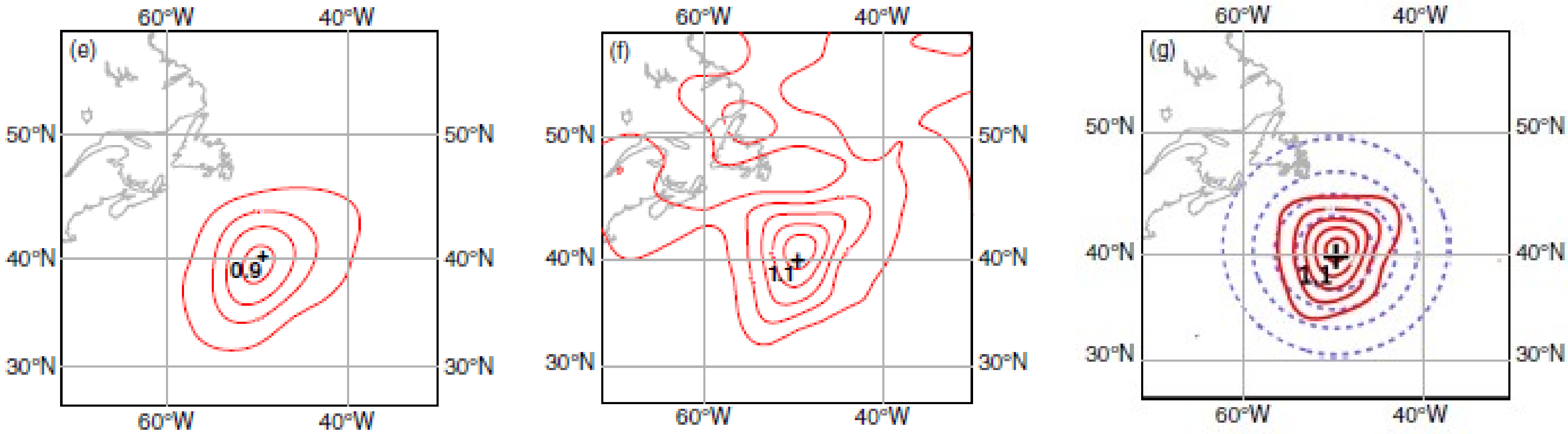


Vertical correlations of temperature errors
between 850 & 870 hPa

=> lesser amount of vertical propagation in stable areas.

Covariance anisotropy and localisation

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



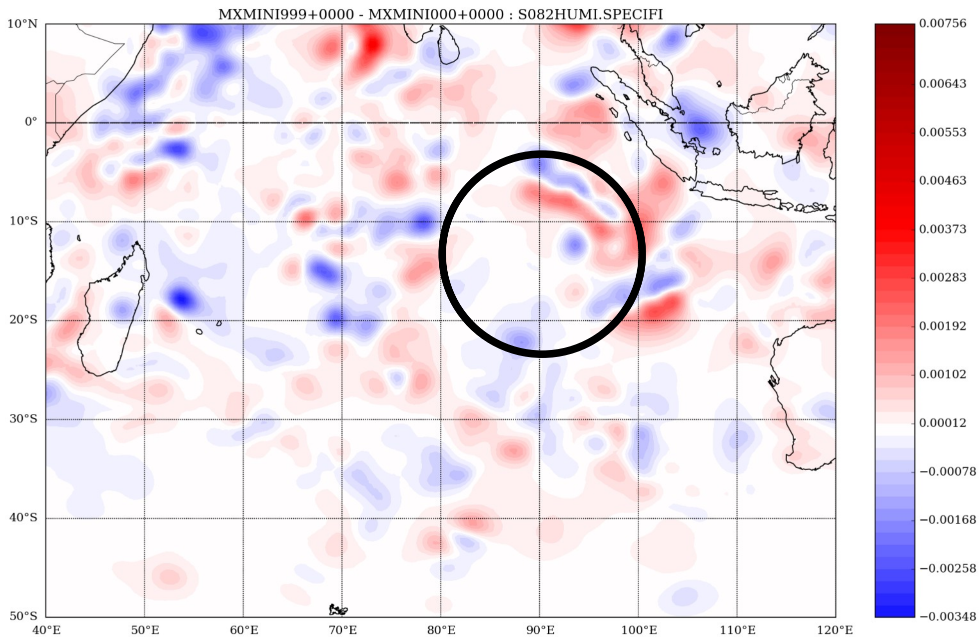
« Exact » covariances

Raw covariances (200 members)

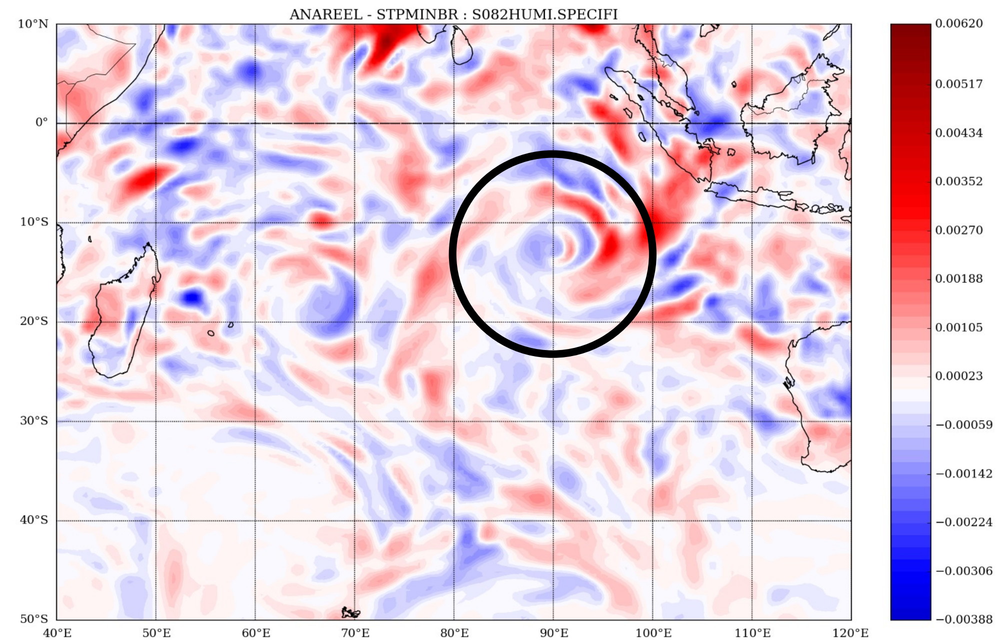
Localised covariances (200 members)

Flow-dependent anisotropic covariances

Humidity analysis increments (near 850 hPa)



With isotropic wavelet correlations



With anisotropic correlations, filtered by localisation

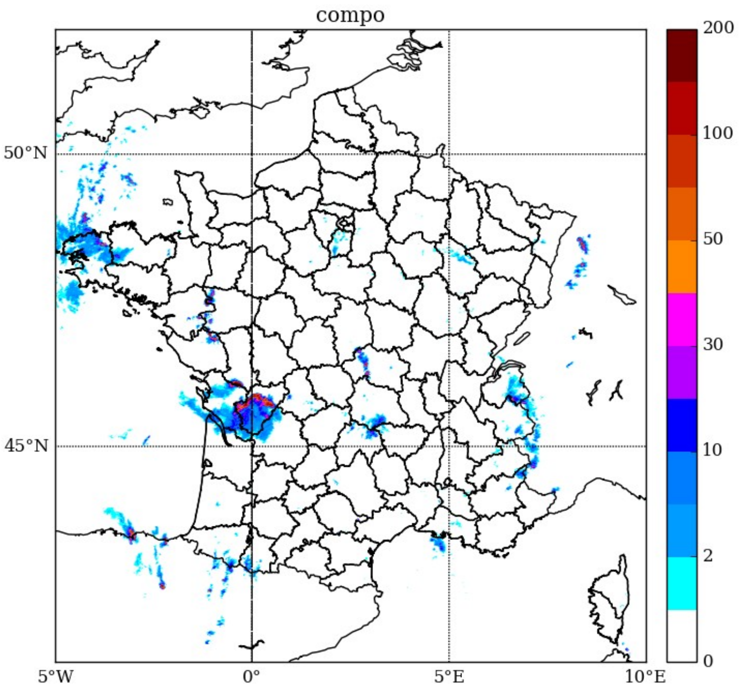
$$B^{1/2} = L\ S\ C^{1/2}$$

$$B_0^e = X_0^{b'} X_0^{b'T} \circ L$$

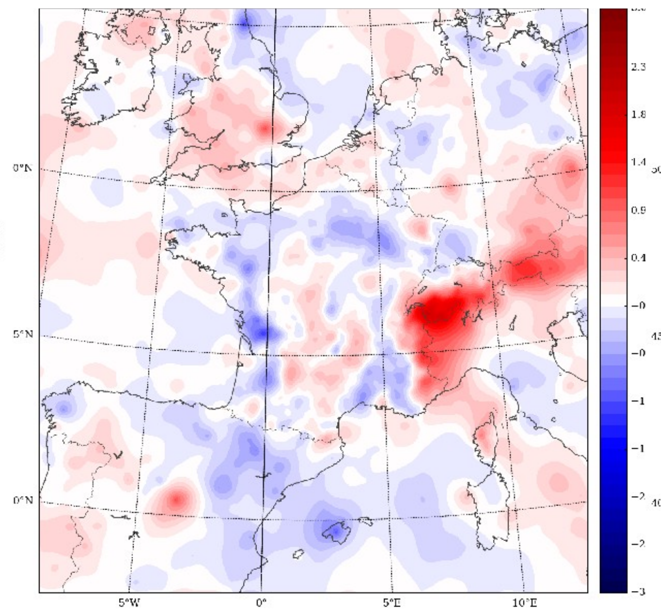
=> flow-dependent anisotropic covariances
in ARPEGE and in AROME (operational since 2024)

METEO FRANCE

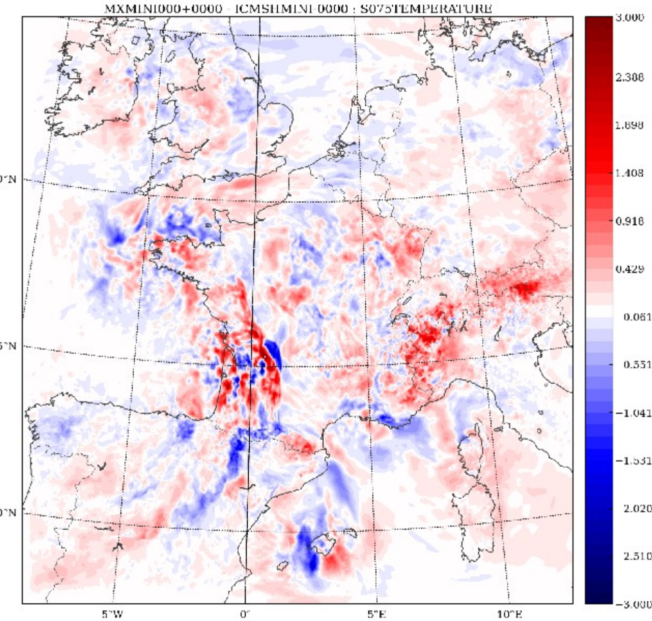
4DEnVar increments resulting from 4D ensemble covariances



3D-Var increment
14UTC



3DEnVar increment
14UTC



4DEnVar increments
13h30-14h30

4DEnVar allows 4D analysis increments to be provided consistently at different times,

in accordance with ongoing processes
in such a convective situation at HR.

(Plots are temperature increments (in K) at 850 hPa.)

METEO FRANCE

(P. Brousseau)

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).
- AI for emulating the model and/or DA (component-wise, or as a whole).

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).

METEO FRANCE

4DEnVar
Variational analysis based on a 4D Ensemble

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

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

with $\underline{\mathbf{B}} = \underline{\mathbf{X}}^{b'} \underline{\mathbf{X}}^{b'T} \circ \underline{\mathbf{L}}$, where $\mathbf{L}$ is a localization matrix,

$$\underline{\mathbf{X}}^{b'} = (\underline{\mathbf{x}}^{b'}_1, \ldots, \underline{\mathbf{x}}^{b'}_{Ne}),$$

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

$\underline{\mathbf{x}}^{b'}$ 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).

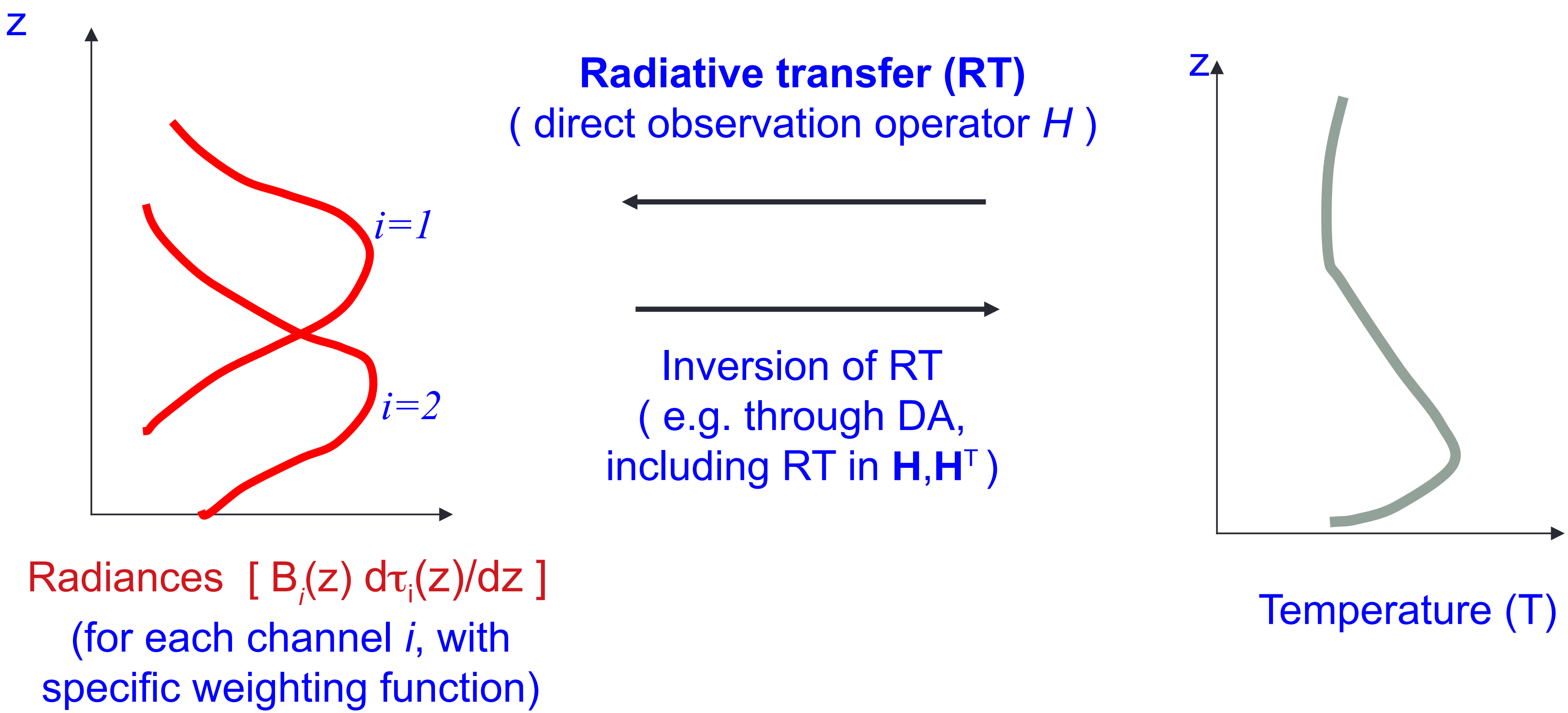
METEO FRANCE

4DEnVar
Variational analysis based on a **4D En**semble

- 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).

METEO FRANCE

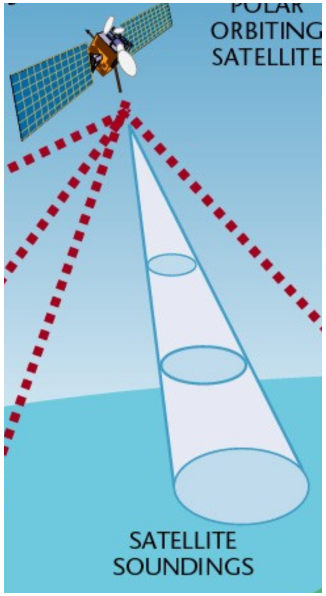
Radiative Transfer : compute « simulated radiances » from temperature model profiles, wich can be compared with « observed radiances »



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

METEO FRANCE

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 : $W/m^2 Sr.cm^{-1}$]

- Planck function:
  $B_\upsilon(T)$= radiance emitted *by a black body* at temperature T, for wave number υ.

- Intensity of the radiation, emitted *by the atmosphere* at wave number υ:

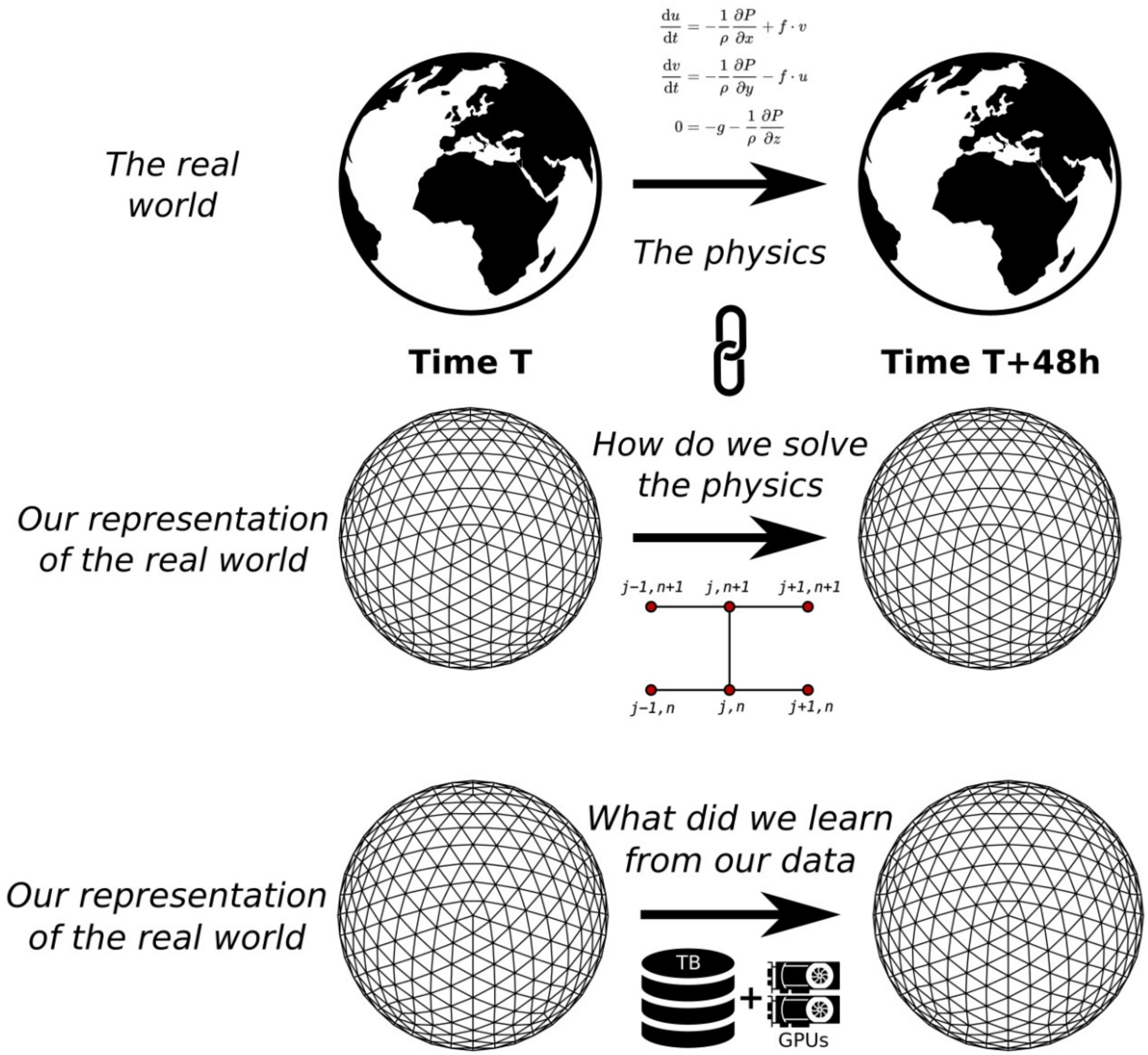
$$R_\upsilon = I_{0,\upsilon}\,\tau_\upsilon(z_0) + \int_{z_0} B_\upsilon[T(z)]\,[d\tau_\upsilon(z)/dz]\,dz$$

  $I_{0,\upsilon}$ 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.

  $K_\upsilon(z) = d\tau_\upsilon(z)/dz$ 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 υ.

Development of Machine Learning emulators for NWP



METEO FRANCE