Surface data assimilation

François Bouyssel (*)
many inputs from G. Balsamo (**), J.C.Calvet (*), M.Drush(**)
(*) Météo-France   (**) ECMWF

Plan

Introduction

Soil moisture and soil temperature analysis

Others analyses: snow, sea surface temperature, sea

ice, ...

Introduction

Why doing surface analysis?

Momentum, heat and moisture fluxes between the surface and the atmosphere play a key role in the evolution of meteorological fields near the ground, in the boundary layer and in the troposphere

These fluxes depend strongly on surface variables which have strong variabilities in time and space (soil temperature, soil moisture, snow, SST, ...)

Sea

Lake

Land

Town

Surface analysis

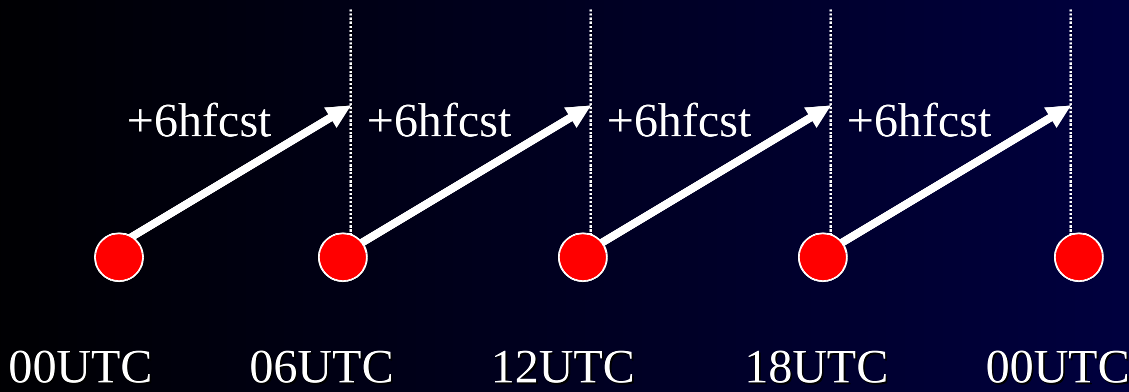
Surface analysis has same problematic than upper-air analysis:

- Which variable(s) to analyse? (state vector: X)
- Which observations are available and informative for your analysis?
- A first guess (background estimate) is used (generally a short range forecast)
- Which optimal analysis algorithm? Optimal interpolation, Variationnal method (3D,4D), Kalman filter, ...
- Observation operators are needed to produce model estimate at the observation point (vertical, horizontal and sometimes physical interpolation) : computation of innovation vector (H(X)-Y)
- Background and observation error statistics (B and R)
- Removing model and observation biaises
- Quality control

Surface analyses and upper-air analysis

For the time being surface analyses are performed separately from upper air analysis. In theory a single analysis would be better but it is much more difficult implement: 1) definition of B between upper air and surface variables, 2) time scale evolutions may be different, ...

For the time being several surface analyses are used for simplicity and because very different surface parameters (Soil temperature and Soil moisture, Snow, SST, Sea ice, ...)

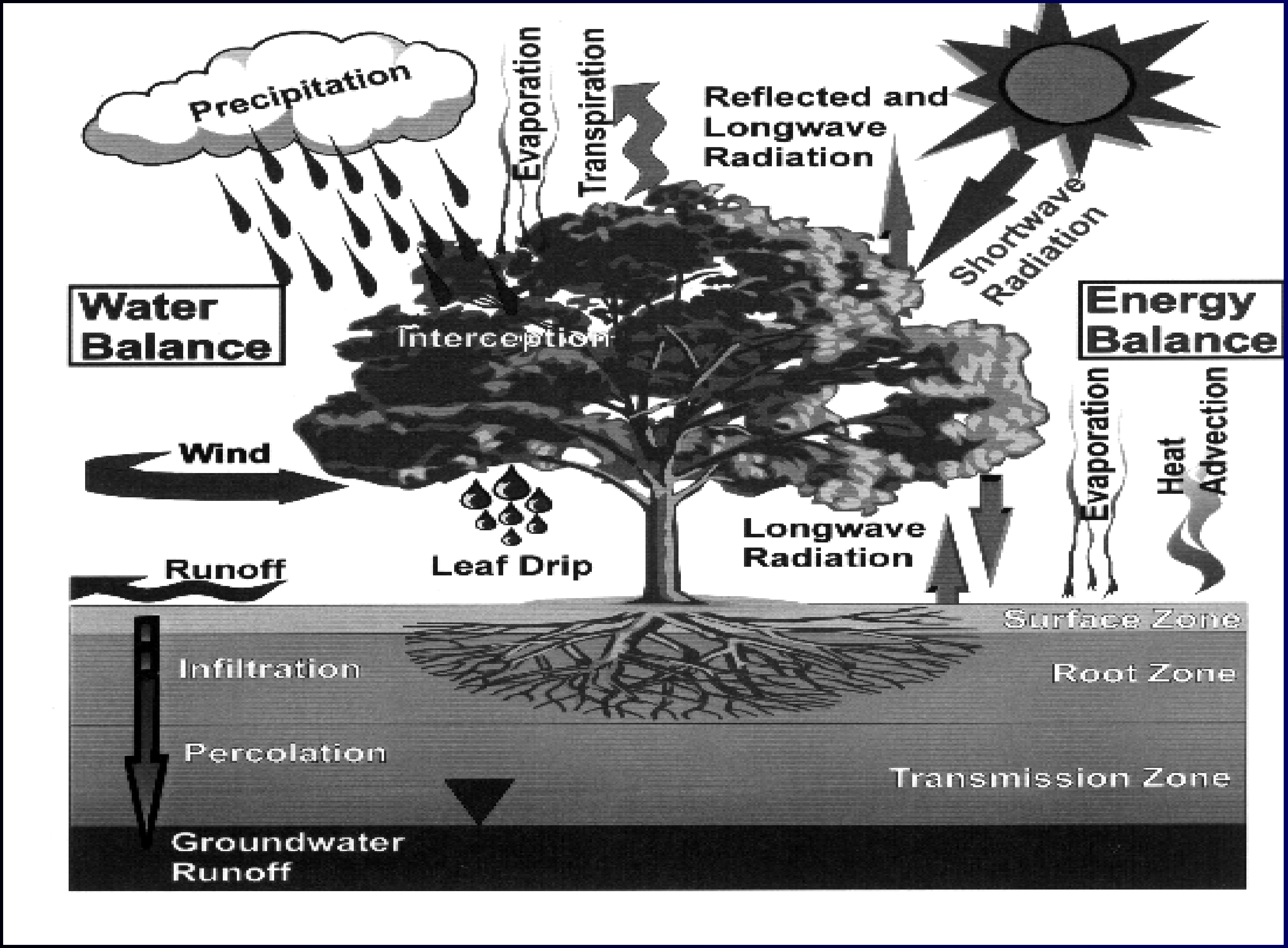


(6h sequantial analysis is just an exemple)

Atmospheric analysis and several surface analyses are done separately and combined at the end to provide the final analysis for the forecast

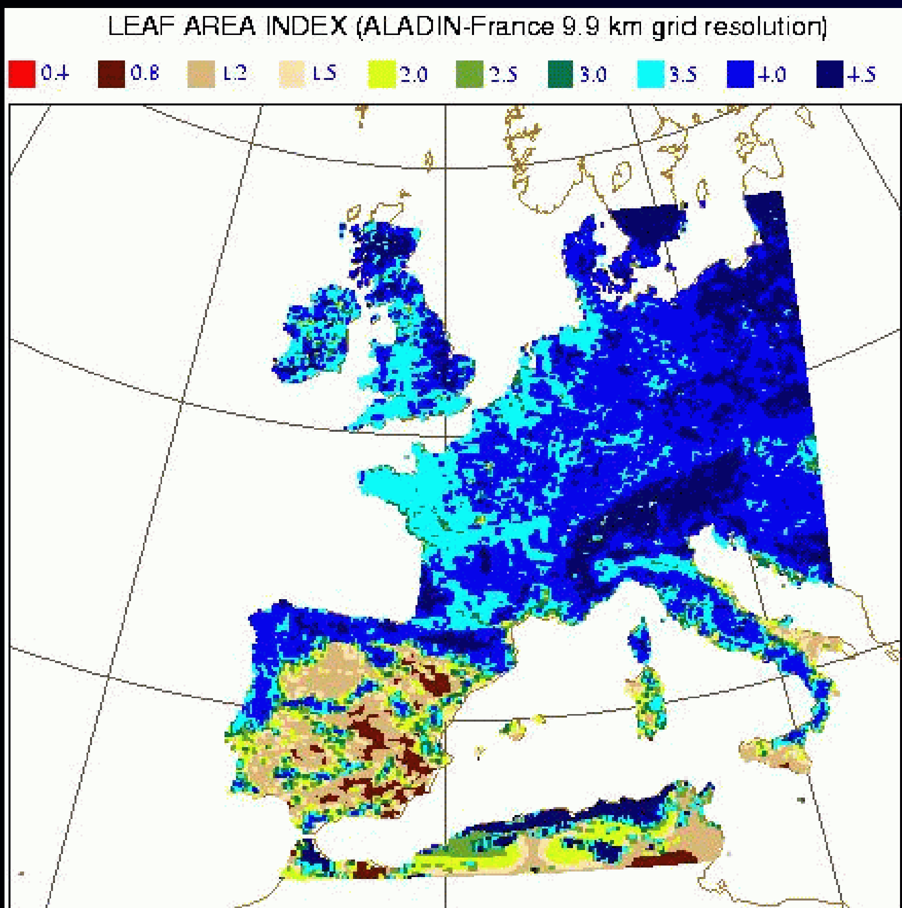
Soil moisture and soil temperature analysis

⇒ Strong improvment of surface schemes in the last 20 years (vegetation, snow, frozen soil, subgrid processes, tiling, …)

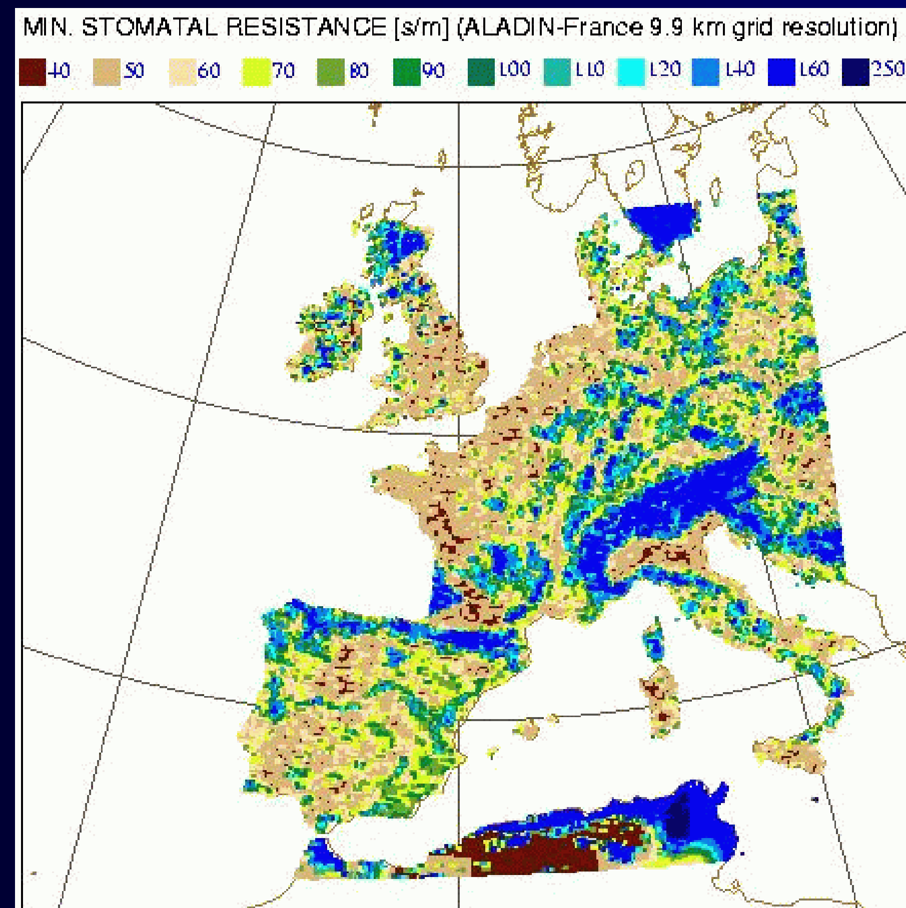


⇒ Better quality of physiographic datasets (soil and vegetation) In recent years [AVHRR, VEGETATION, MODIS, …]

Leaf Area Index (LAI)

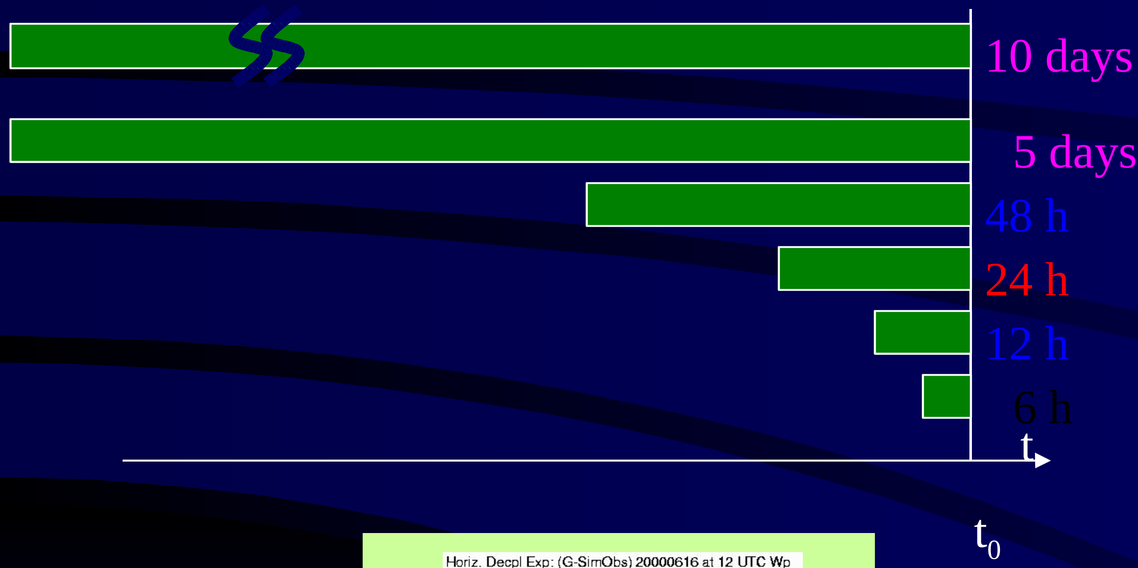


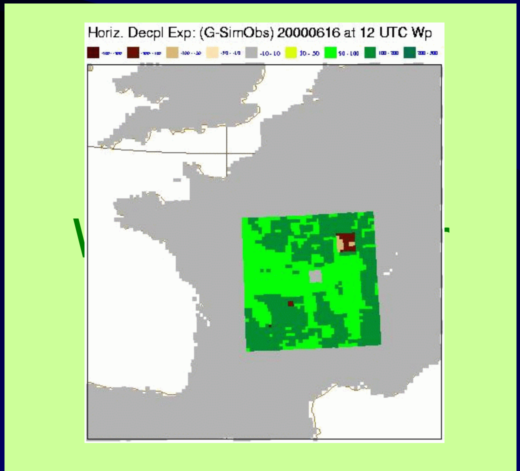
Minimal stomatal resistance

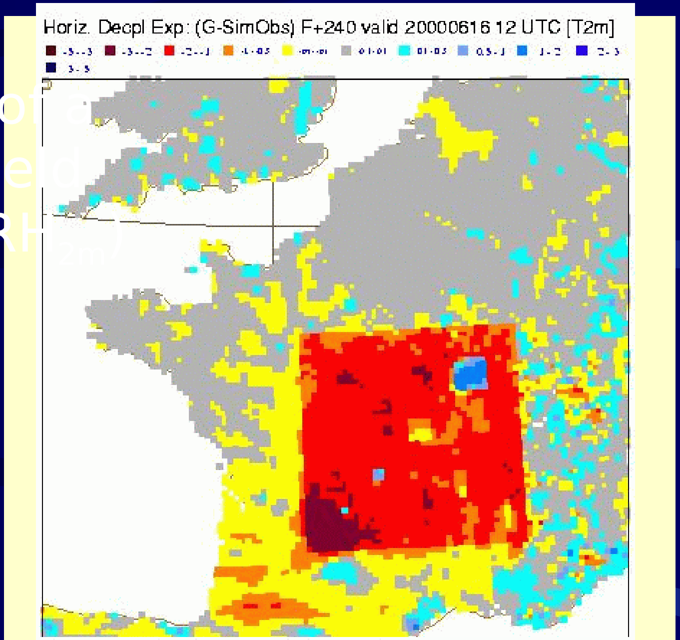


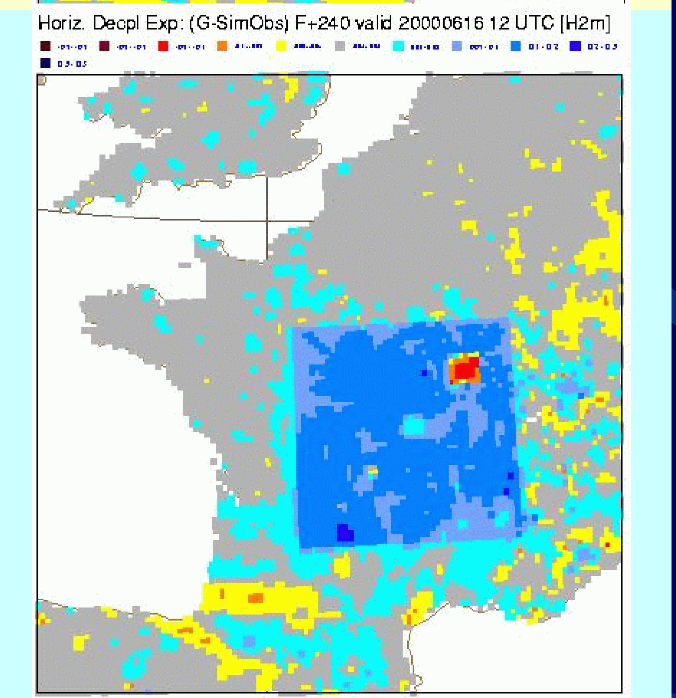
*But **strong sensibility** of surface fluxes (sensible and latent) to the soil temperature and moisture*

To illustrate the memory effect, the impact of a prescribed initial error in the soil moisture field is shown for different *forecast ranges* (T_{2m}, RH_{2m})





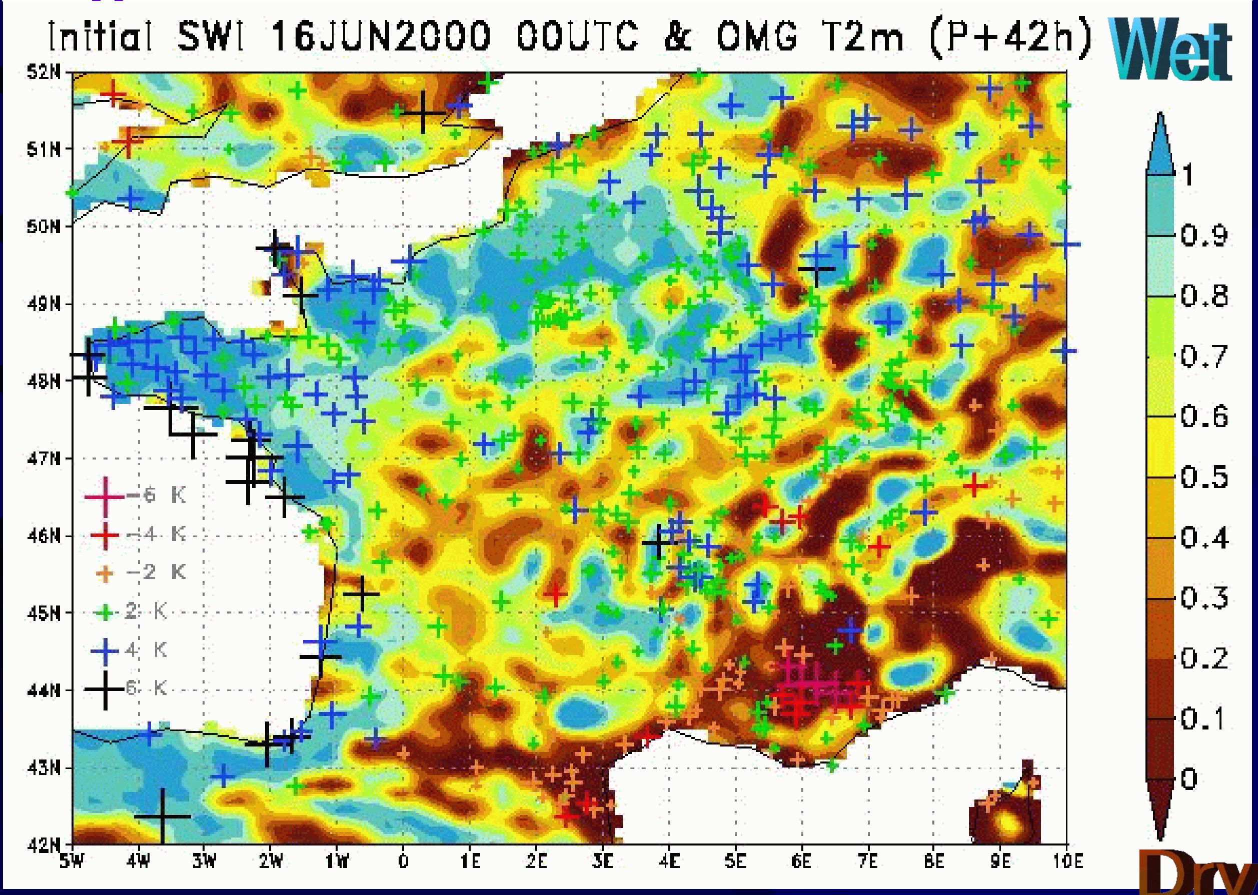




Another exemple: 2m model biases under clear sky conditions on June 2000, 13th-18th

Evaluation of 2m model errors is performed on Hi-res. observations network

Correlation of T2m errors & Soil moisture

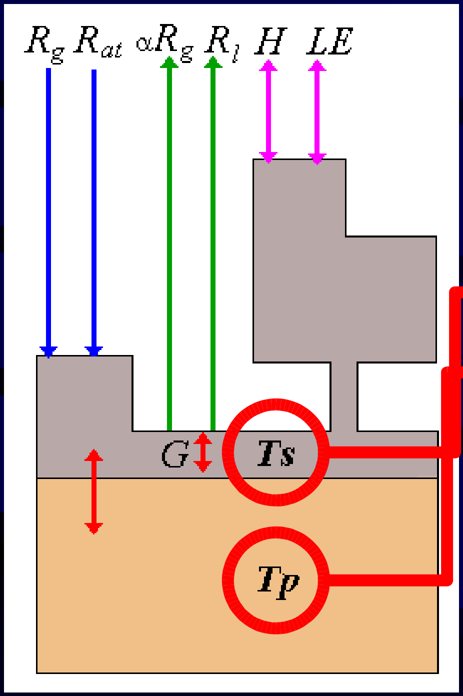


heterogeneity of soil moisture

Surface Parameterization scheme (ISBA)

Operational version : Noilhan & Planton (1989), Noilhan & Mahfouf (1996), Bazile (1999), Giard & Bazile (2000)

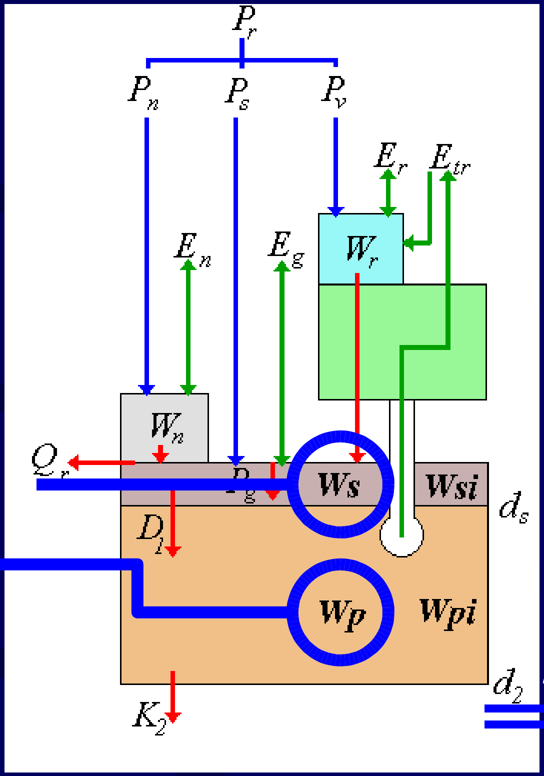
Energy



~1-2 h

~1-2 day

Water



~6-12 h

~10 days

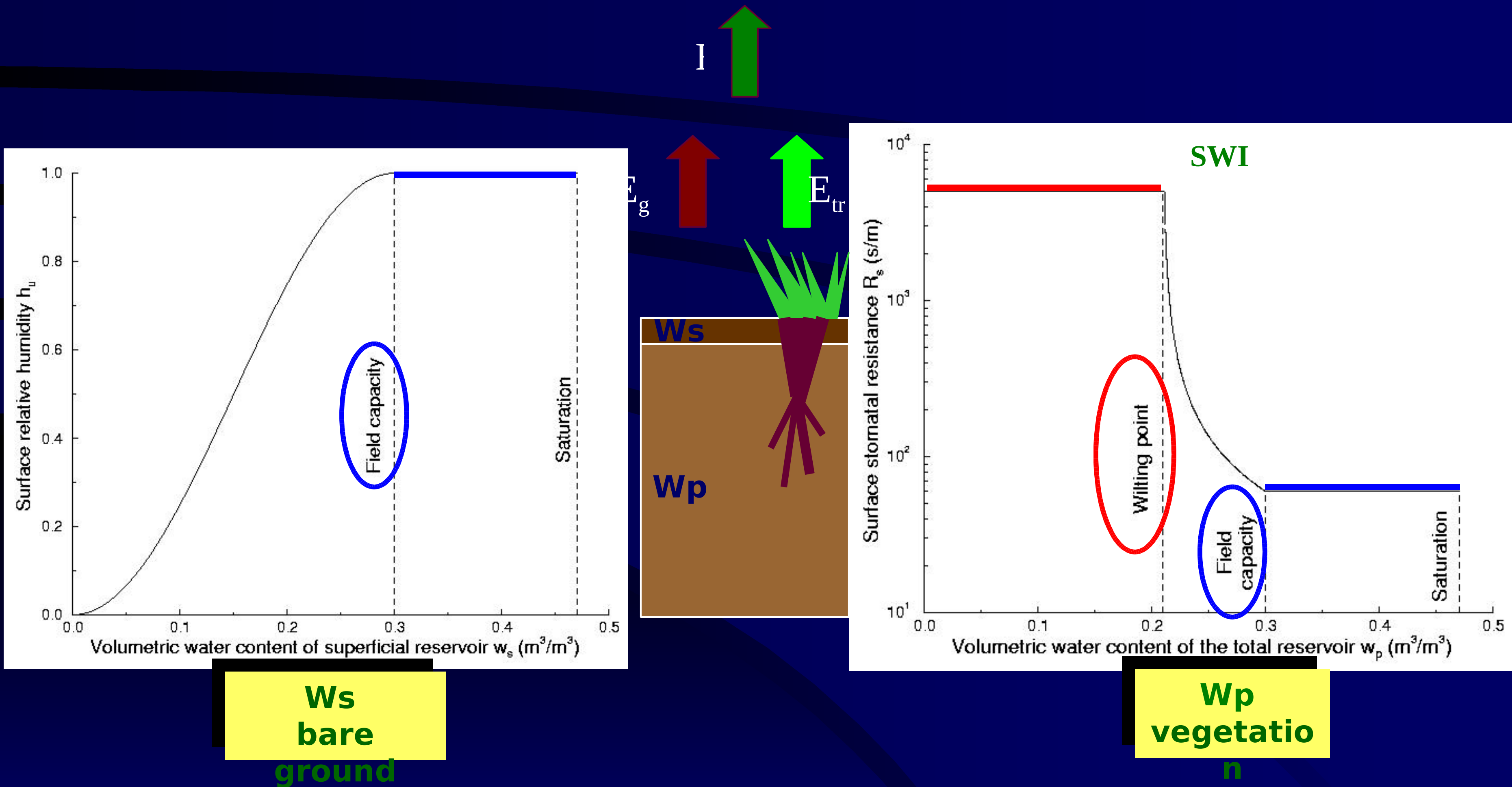
Analysis

- surface temperature
- mean soil temperature
- superficial soil water content
- **total soil water content**

Research versions : interactive vegetation module (Calvet et al. 1998), sub grid-scale runoff and sub-root layer (Boone et al 1999), explicit 3-layers snow scheme (Boone & Etchevers 2001), tiling, multi-layer soil scheme, urban scheme

The link between soil moisture and atmosphere

- The main interaction of soil moisture and atmosphere is due to evaporation and vegetation transpiration processes.



Importance of soil moisture and temperature analysis

stable surface conditions : Low surface fluxes. Influence of surface limited near the ground
instable surface conditions : Strong surface fluxes. Influence on PBL evolution and sometimes more (trigger deep convection)

Soil moisture is very important under strong solar radiation at the surface because it determines the repartition of incoming energy into sensible and latent heat fluxes.
Importance of initialization: Wr << Ws << Wp according capacity and time scale evolution. Accumulation of model error may degrade significantly the forecast during long period.

Soil temperature is important in case of stable conditions because it affects low level temperature. Importance of initialization: Ts << Tp

Necessity of same degree of sophistication between surface scheme, physiographic database, surface analysis

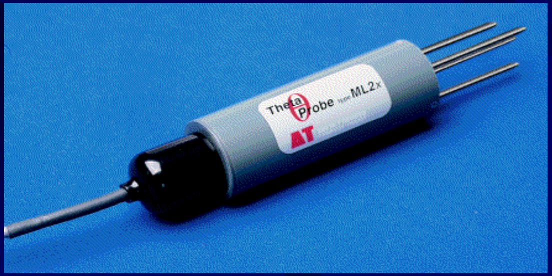
Specificities of soil moisture and temperature analysis

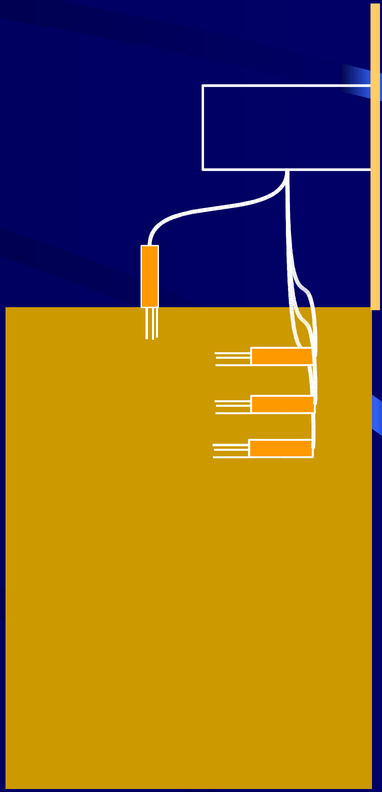
Strong soil and vegetation spatial heterogeneities (mountains, coastal regions, forest, bare ground, various cultures, towns, lakes, ...)

Strong spatial variability of soil moisture (linked with surface and soil properties and precipitations)

Lack of direct observations (very expensive and problem of representativeness)

Large variety of time scales in soil processes (up to several weeks or months)





Available observations for soil moisture analysis

Precipitations observations (rain gauges, radars) :
+ direct link with the variations of soil water content

Satellite observations:
+ global coverage
+ infrared: clear sky, low vegetation, geostationnary satellites : high temporal and spatial resolutions (energy budget), strong sensitivity to low level wind, surface roughness
+ microwave: active and passive instruments measure directly the soil moisture in the first few centimeters (scatterometer (ERS,ASCAT), passive or active radiameters (SMOS, HYDROS): resolution ~20/40km, frequency ~0.3/1 per day

2m observations (temperature et humidity):
+ good global coverage of existing network
+ close links with the fields in the ground in some meteorological conditions

Operational initialization methods

Climatological relaxation of deep soil parameters (uncertainties in these climatologies (GSWP), interannual variability not taken into account)

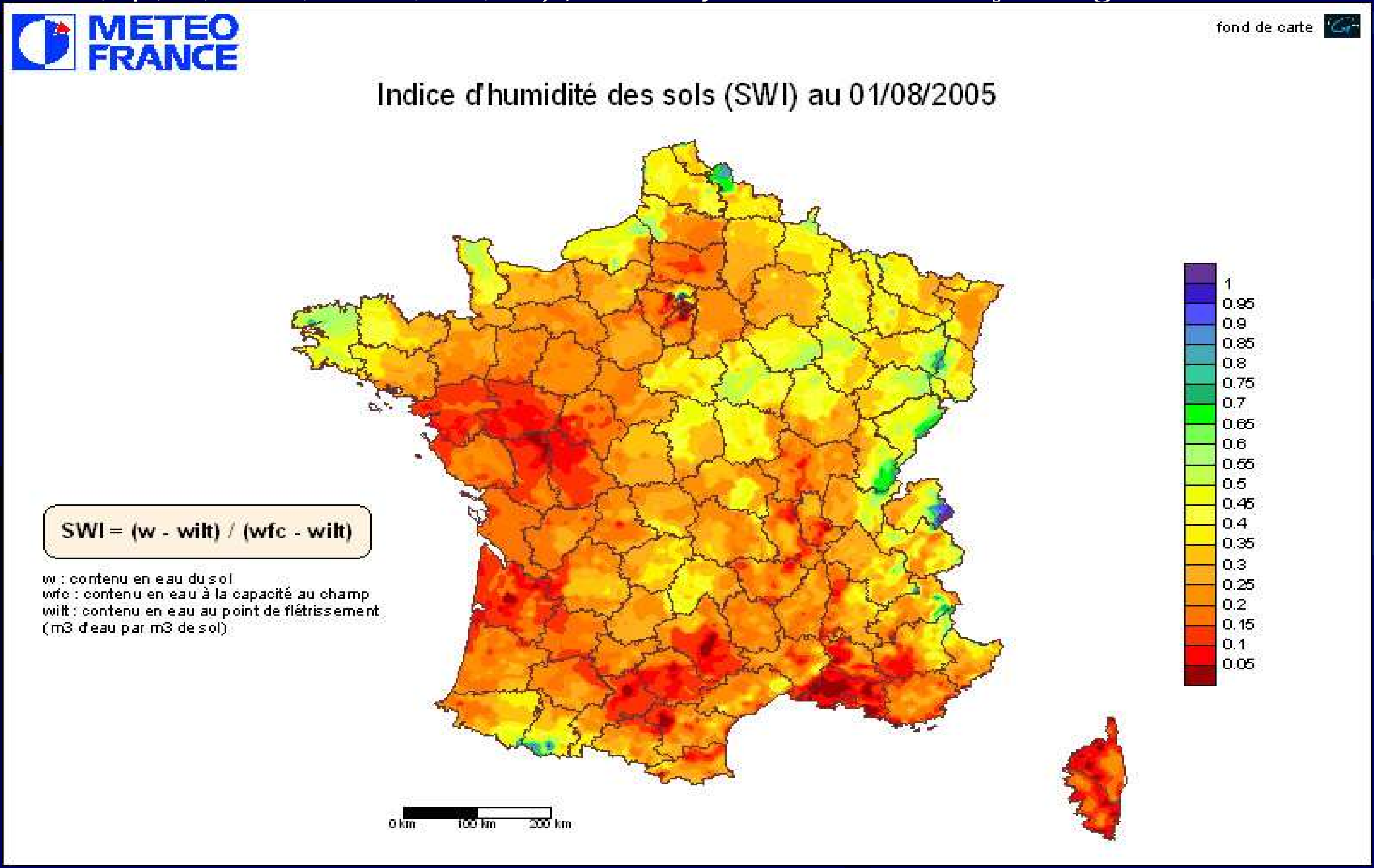
Off line surface scheme driven by forecasted or analysed fields and fluxes (flux of precipitation, of radiation, fields near the surface T_{2m}, HU_{2m}, V_{10m}, P_s)
Exemple: SAFRAN-ISBA-MODCOU

No existant utilization of satellite data for temperature and soil moisture analysis (near future)

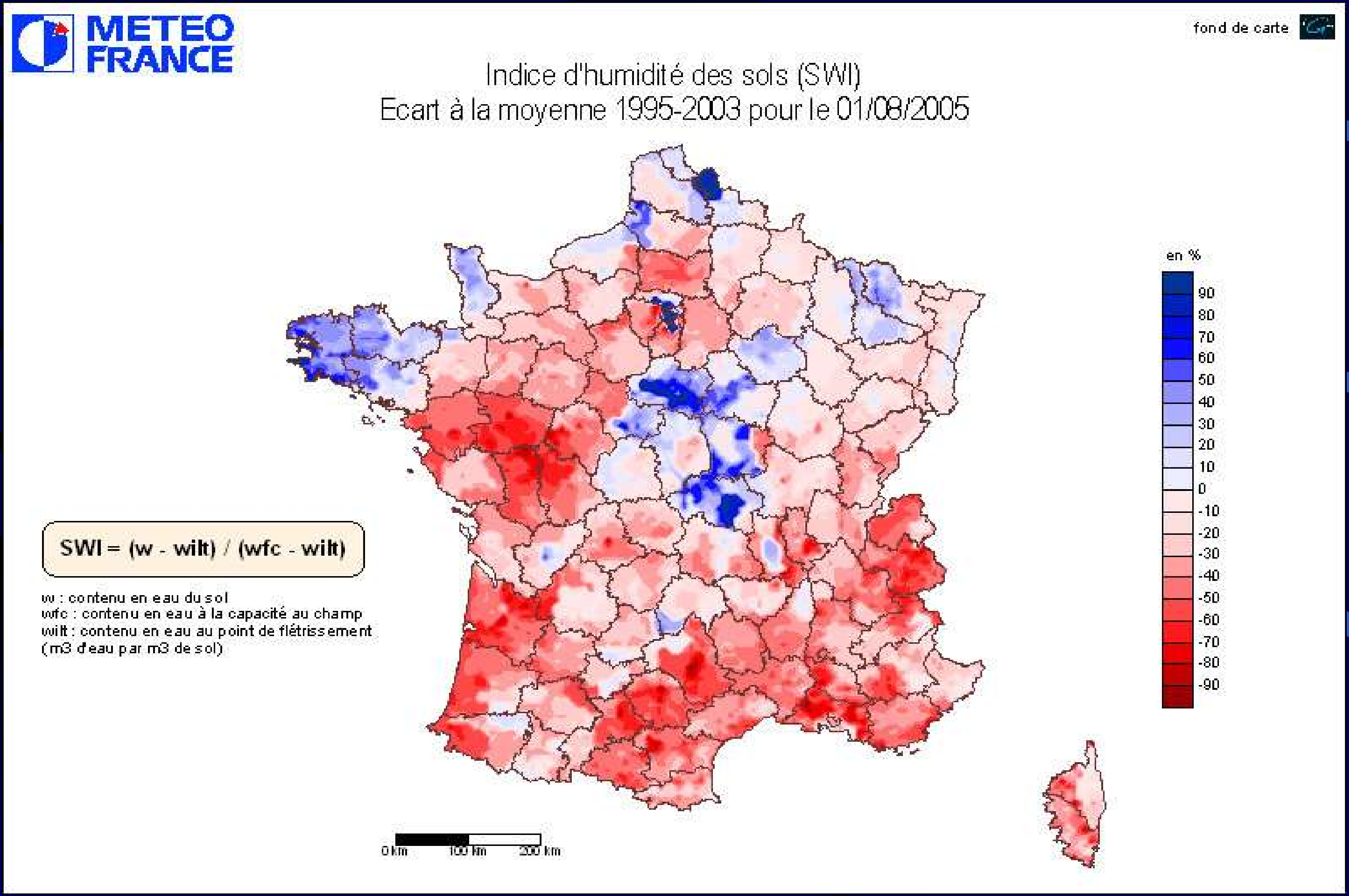
Assimilation of 2m observations of temperature and relative humidity

Off-line method (SIM exemple)

Run operationally over France at 8 km : SAFRAN (upperair analysis: Ta, qa, U, SW↓, LW ↓, RR, …) , ISBA, MODCOU Hydrological model



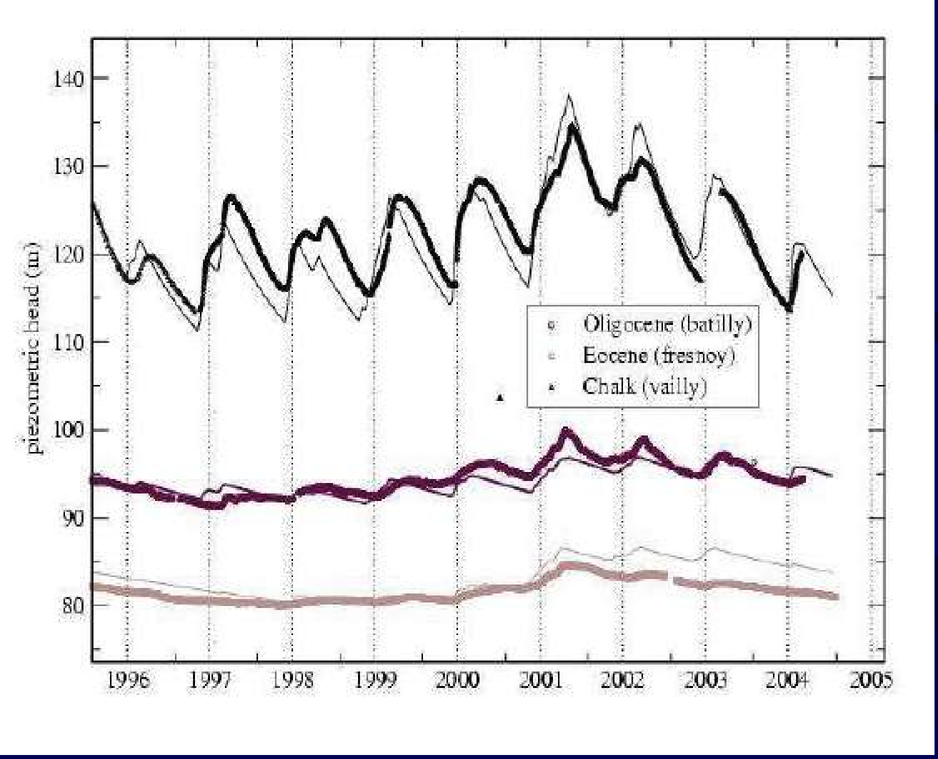
Off-line method (SIM exemple)



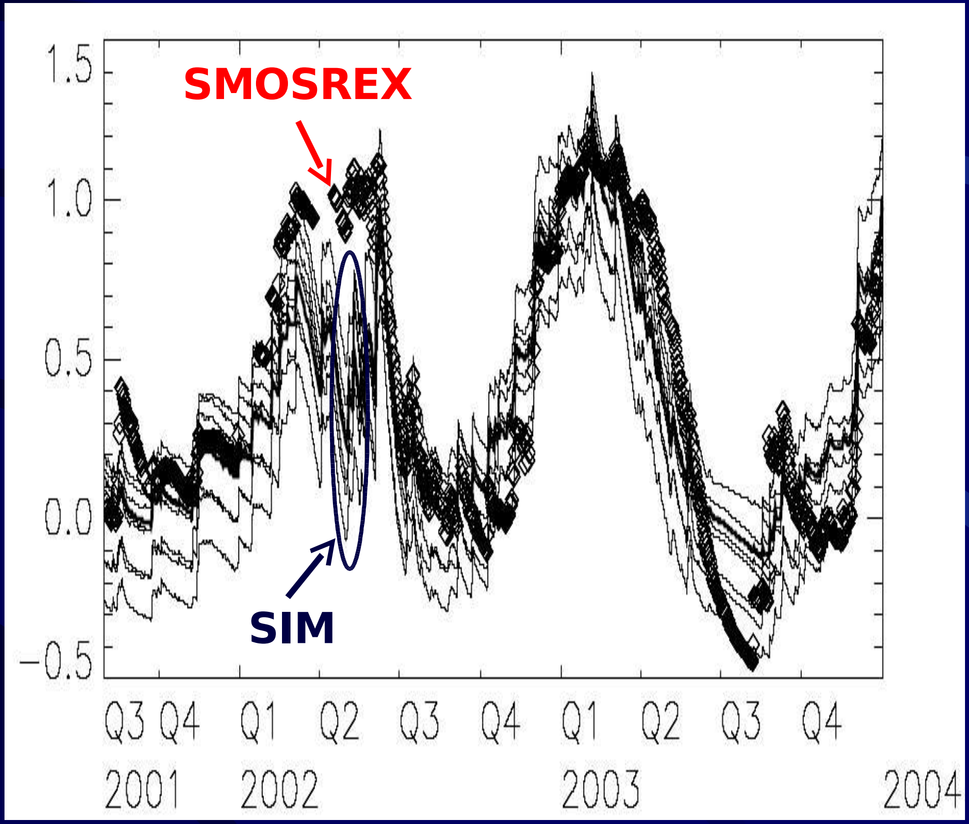
Off-line method (SIM exemple)

Validation: river flow & snow depth & measurement site

water table (Seine)



Soil Wetness Index(SMOSREX)



Off-line method (pros & cons)

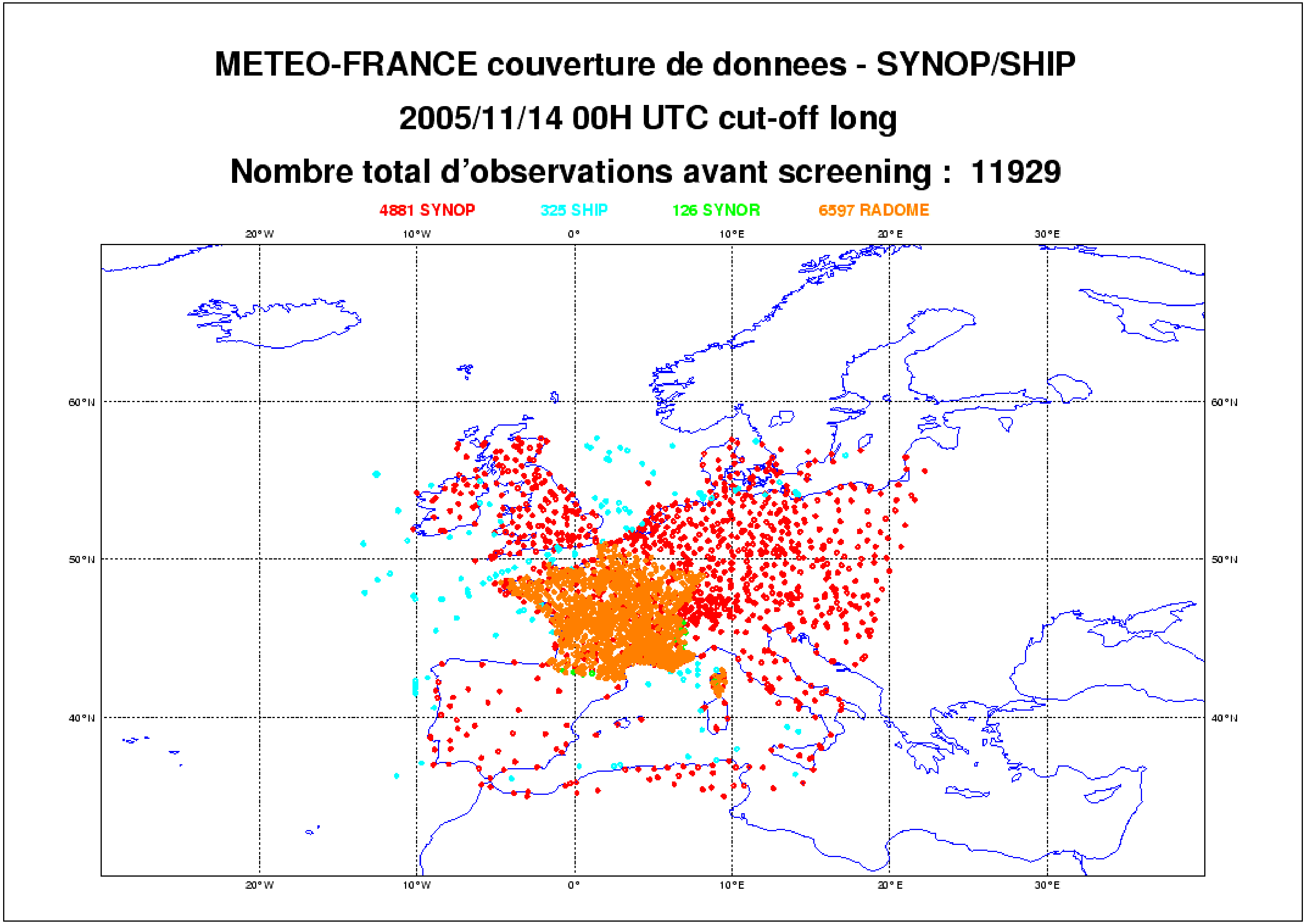
Pros:

+ with good precipitation, radiation and atmospheric forcings provides realistic soil moisture evolution even at high temporal evolution (useful for NWP, but also agriculture, water managment, ...)
+ cheap model (just the surface), work on PC, allows multi-years reanalysis
+ allows the use of complex surface model
+ high spatial resolution (RR analysis, MSG radiation fluxes)

Cons:

+ no analysis & perfect model hypothesis while surface processes are complex and physiographic database not perfect: model bias may exist on soil moisture and soil temperature and remain for a long period
+ restricted to some geographical areas (good obs coverage)

Assimilation of 2m observations



Optimum Interpolation method

Coiffier 1987, Mahfouf 1991, Bouttier 1993, Giard and Bazile 2000

1) Optimum Interpolation of T_{2m} and RH_{2m} using SYNOP observations interpolated at the model grid-point (by a 2m analysis)

$$\Delta T_{2m} = T_{2m}{}^{a} - T_{2m}{}^{b} \qquad \Delta RH_{2m} = RH_{2m}{}^{a} - RH_{2m}{}^{b}$$

2) Correction of surface parameters (T_s, T_p, W_s, W_p) using 2m increments between analysed and forecasted values

Sequential analysis (every 6h)

$$\mathbf{x}^{a} = \mathbf{x}^{b} + \mathbf{BH}^{T}(\mathbf{HBH}^{T} + \mathbf{R})^{-1}(\mathbf{y} - H(\mathbf{x}^{b}))$$

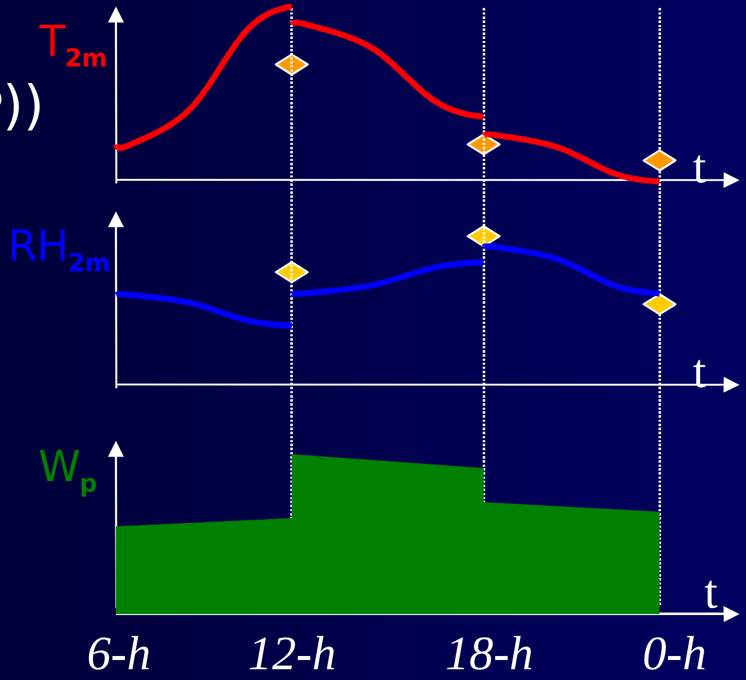
$$T_s{}^{a} - T_s{}^{b} = \Delta T_{2m}$$

$$T_p{}^{a} - T_p{}^{b} = \Delta T_{2m} / 2\pi$$

$$W_s{}^{a} - W_s{}^{b} = \alpha_{WsT} \Delta T_{2m} + \alpha_{WsRH} \Delta RH_{2m}$$

$$W_p{}^{a} - W_p{}^{b} = \alpha_{WpT} \Delta T_{2m} + \alpha_{WpRH} \Delta RH_{2m}$$

OI coefficients



Optimum Interpolation coefficients

$$\alpha_{\mathrm{Ws/pT}} = \frac{\sigma^{b}_{\mathrm{Ws/p}}}{\Phi \sigma^{b}_{\mathrm{T2m}}} \left\{ \left[1 + \left(\frac{\sigma^{a}_{\mathrm{RH2m}}}{\sigma^{b}_{\mathrm{RH2m}}} \right)^{2} \right] \rho_{\mathrm{T2m,Ws/p}} - \rho_{\mathrm{T2m,RH2m}} \rho_{\mathrm{RH2m,Ws/p}} \right\}$$

$$\alpha_{\mathrm{Ws/pRH}} = \frac{\sigma^{b}_{\mathrm{Ws/p}}}{\Phi \sigma^{b}_{\mathrm{RH2m}}} \left\{ \left[1 + \left(\frac{\sigma^{a}_{\mathrm{T2m}}}{\sigma^{b}_{\mathrm{T2m}}} \right)^{2} \right] \rho_{\mathrm{RM2m,Ws/p}} - \rho_{\mathrm{T2m,RH2m}} \rho_{\mathrm{T2m,Ws/p}} \right\}$$

$$\Phi = \left[1 + \left(\frac{\sigma^{a}_{\mathrm{T2m}}}{\sigma^{b}_{\mathrm{T2m}}} \right)^{2} \right] \left[1 + \left(\frac{\sigma^{a}_{\mathrm{RH2m}}}{\sigma^{b}_{\mathrm{RH2m}}} \right)^{2} \right] - \rho^{2}_{\mathrm{T2m,RH2m}}$$

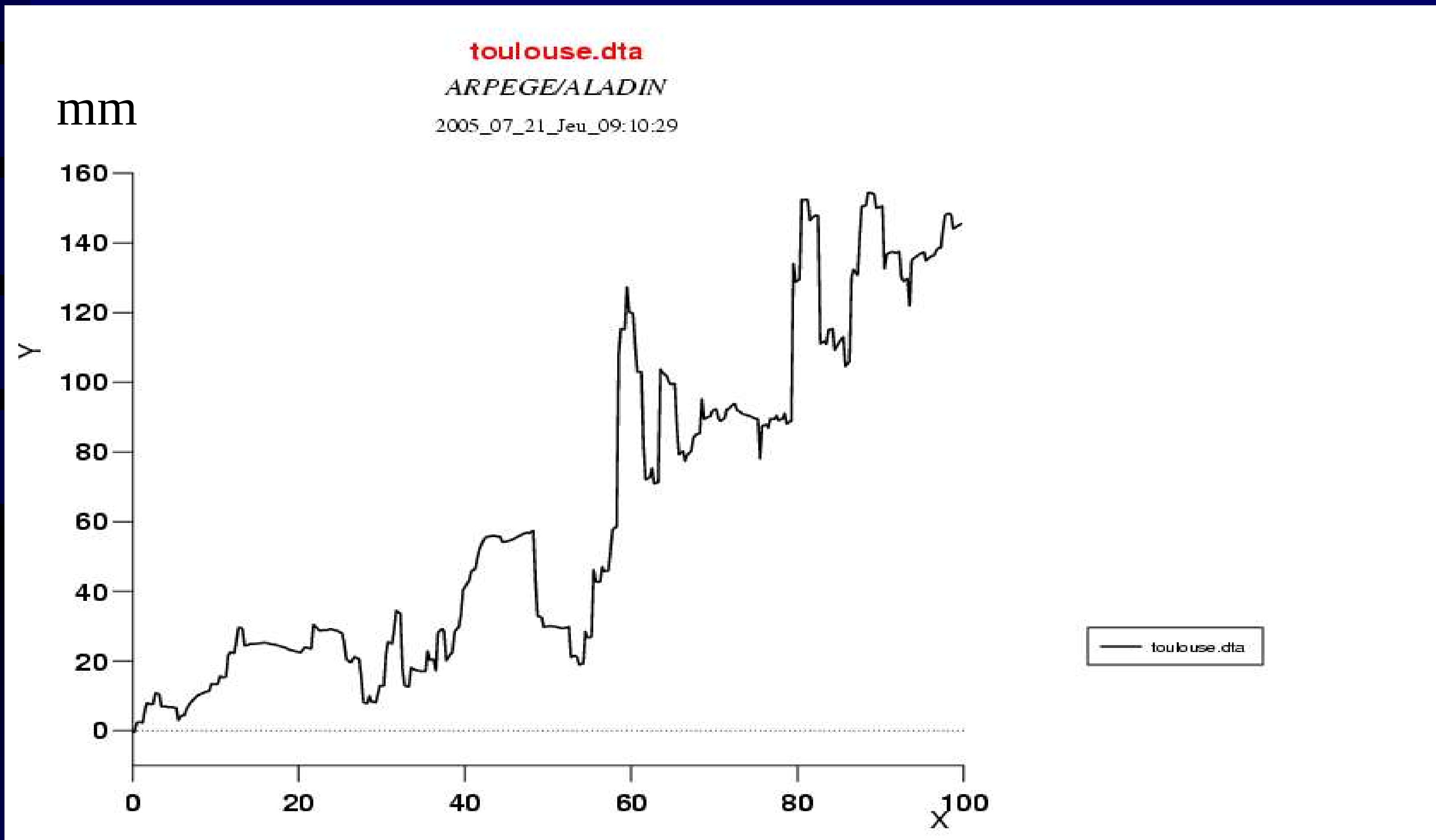
Very strong dependency of these backgroung error statistics to physiographic properties and meteorological conditions

MonteCarlo method under summer anticyclonic conditions to get the dependency to physiography (deriving analytical formulation of OI coefficients) + empirical additional dependency to meteorological conditions

$$\alpha_{Wp/sT/RH} = f\,(t, veg, LAI/Rs_{min}, texture, atmospheric\ conditions)$$

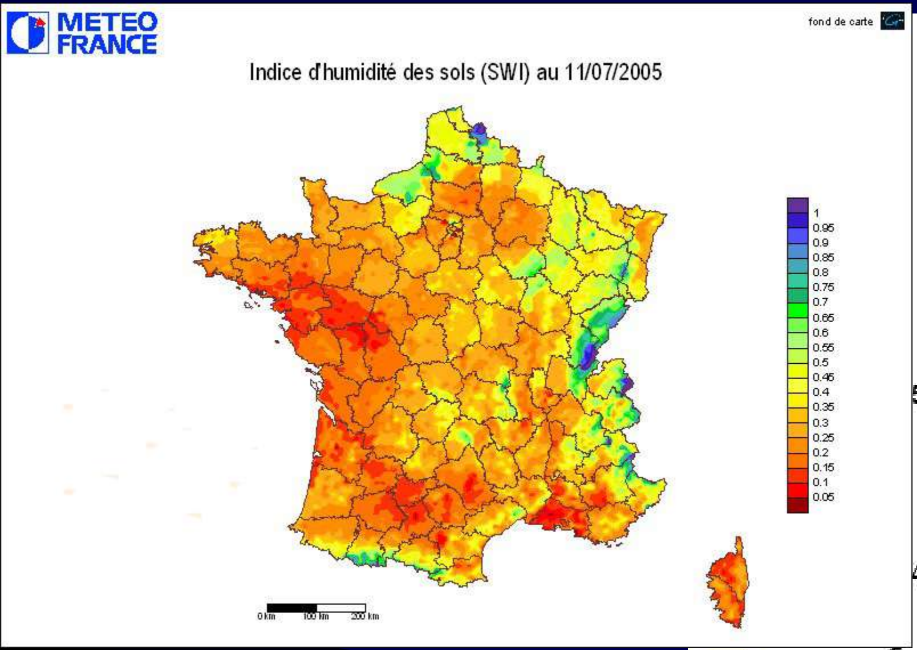
Long and difficult work (in principle should be redo with model or physiography evolutions!)

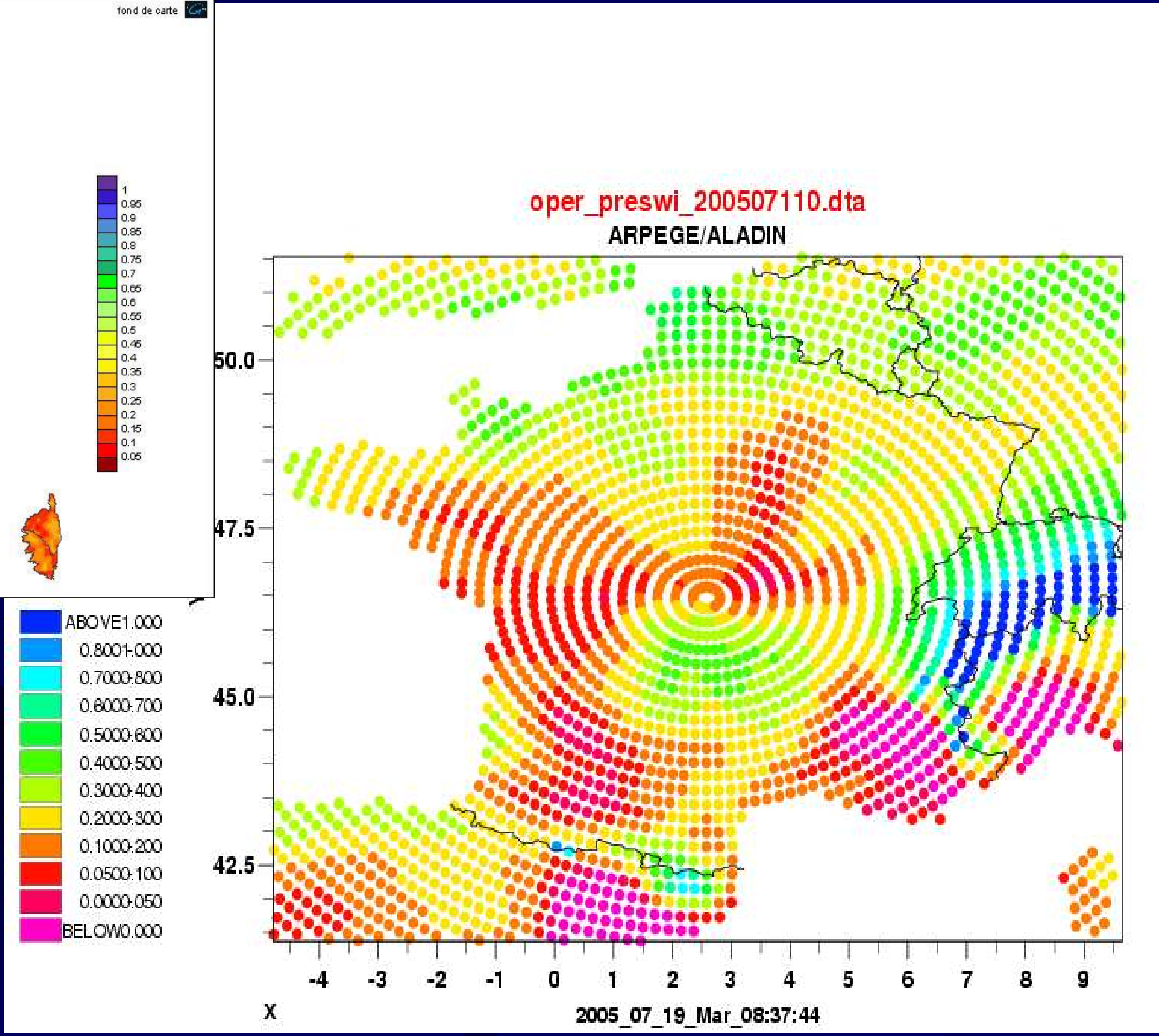
Cumulated analysis increments for Toulouse's (south of France) nearest gridpoint from 1st April to 11 July 2005



days

Soil Wetness Index in SIM (left) et in ARPEGE (right) 11 July 2005





Optimal interpolation with 2m obs (pros & cons)

Pros:

+ suitable in most area in the world, quite cheap analysis
+ work for soil moiture and soil temperature
+ take into account model errors (surface model, physiographic database) to provide suitable soil moisture for fitting 2m observations (if no model error sensible and latent heat fluxes are correct). Said differently: « surface processes are too complex to be represented exactly with simple land surface model»

Cons:

+ requires good T2m & RH2m analyses (B not really homogeneous and isotropic)
+ difficulty to distinguish model biais from observation biais (mainly representativeness error)
+ remouving these biases is particularly difficult (strong spatial and temporal variability)
+ soil moisture may be on some area not realistic because of these bias

Variational surface analysis

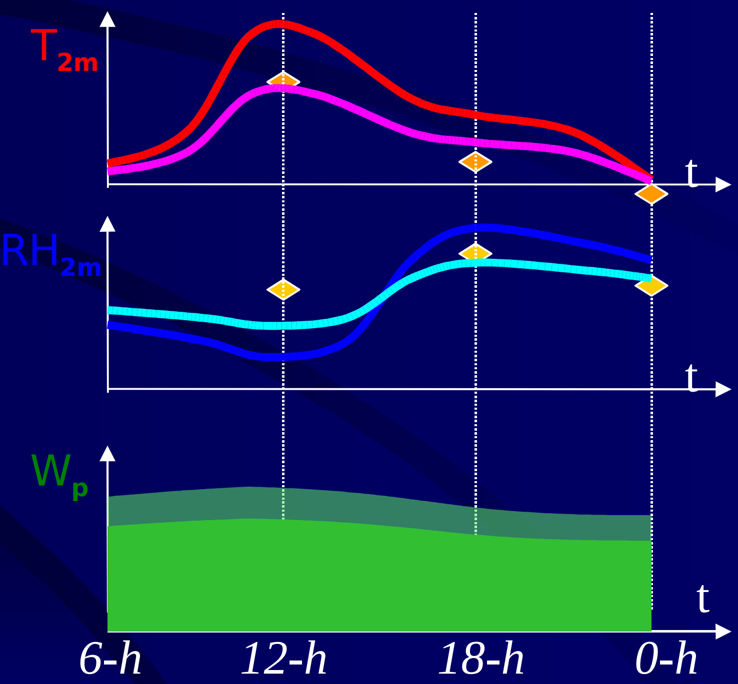
Mahfouf (1991), Callies et al. (1998), Rhodin et al. (1999), Bouyssel et al. (2000)

Formalism:

$$J(\mathbf{x}) = J^{b}(\mathbf{x}) + J^{o}(\mathbf{x}) = \tfrac{1}{2}(\mathbf{x} - \mathbf{x}^{b})^{T}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^{b}) + \tfrac{1}{2}(\mathbf{y} - H(\mathbf{x}))^{T}\mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{x}))$$

x is the control variables vector
y is the observation vector
H is the observation operator
B is the background error covariance matrix
R is the observation error covariance matrix

Continuous analysis



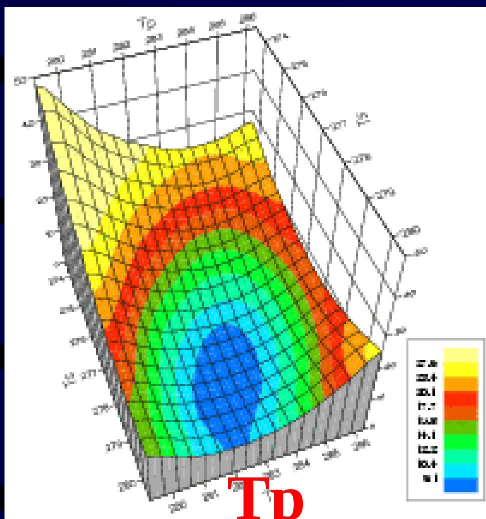
The analysis is obtained by the minimization of the cost function $J(\mathbf{x})$

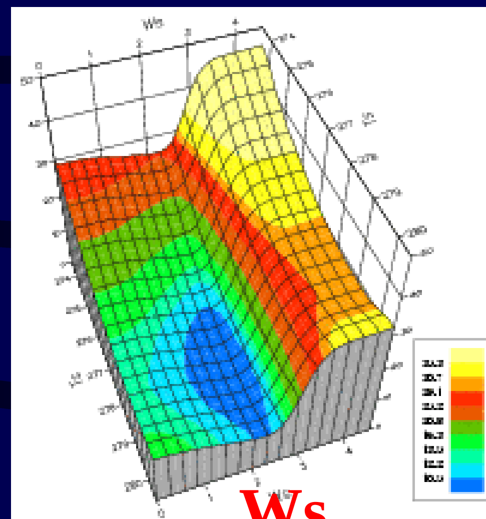
For high dimensional problems: TL/AD models

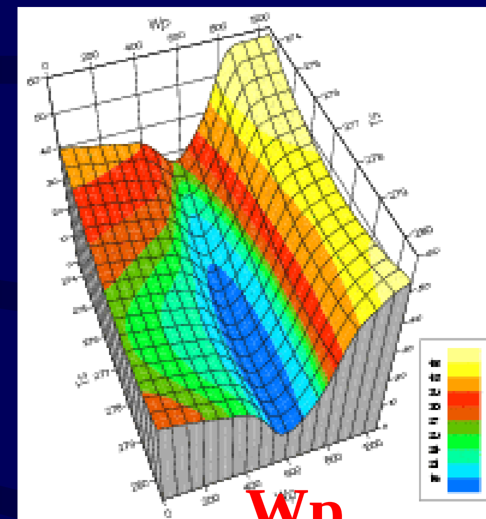
For low dimensional problems: finite differences

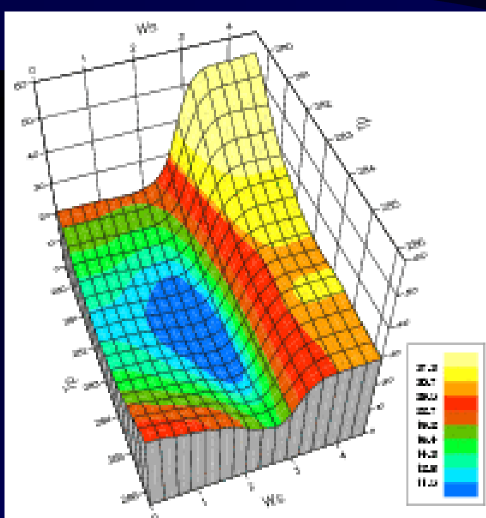
Anticyclonic conditions
Simulated and noisy obs (1°, 10%)
Frequency of observations = 3h
Assimilation window = 24h

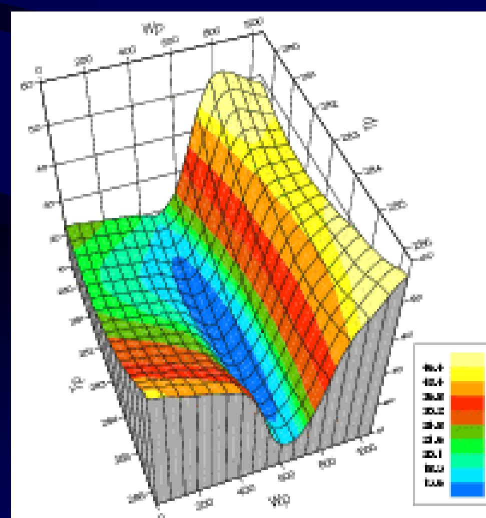
Topology of Jo

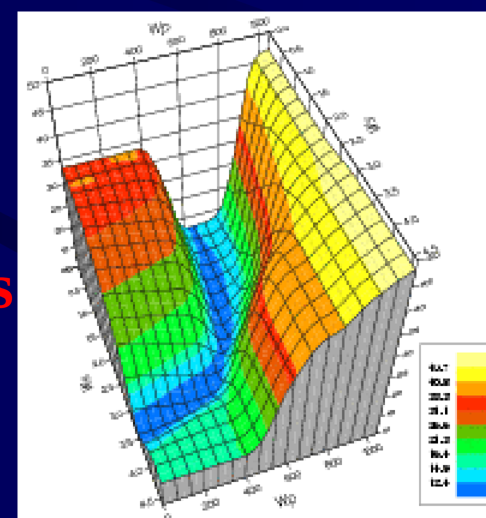




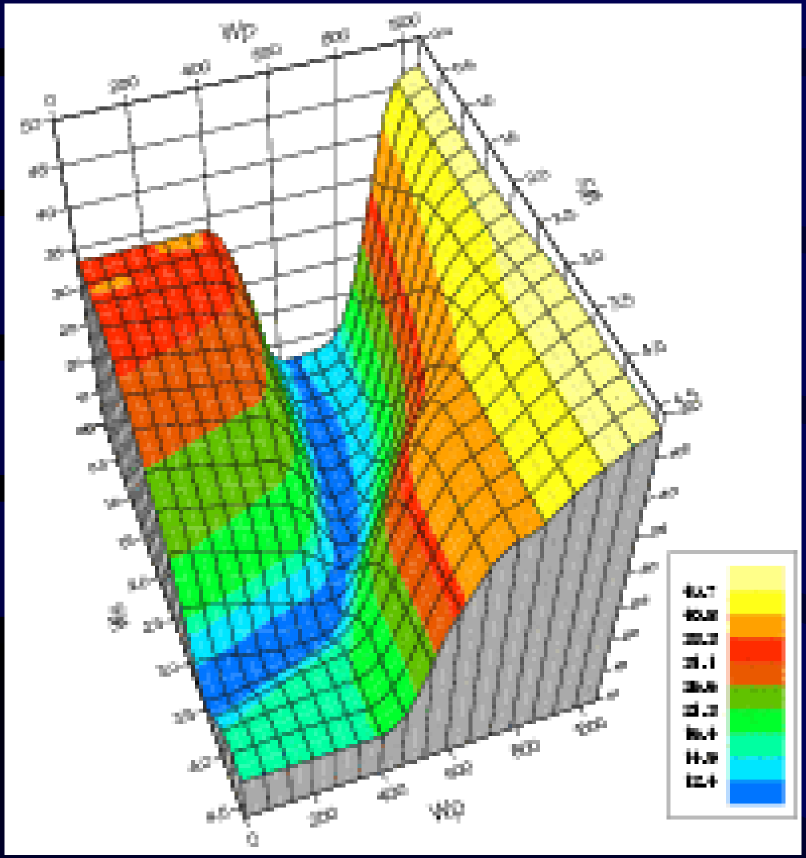


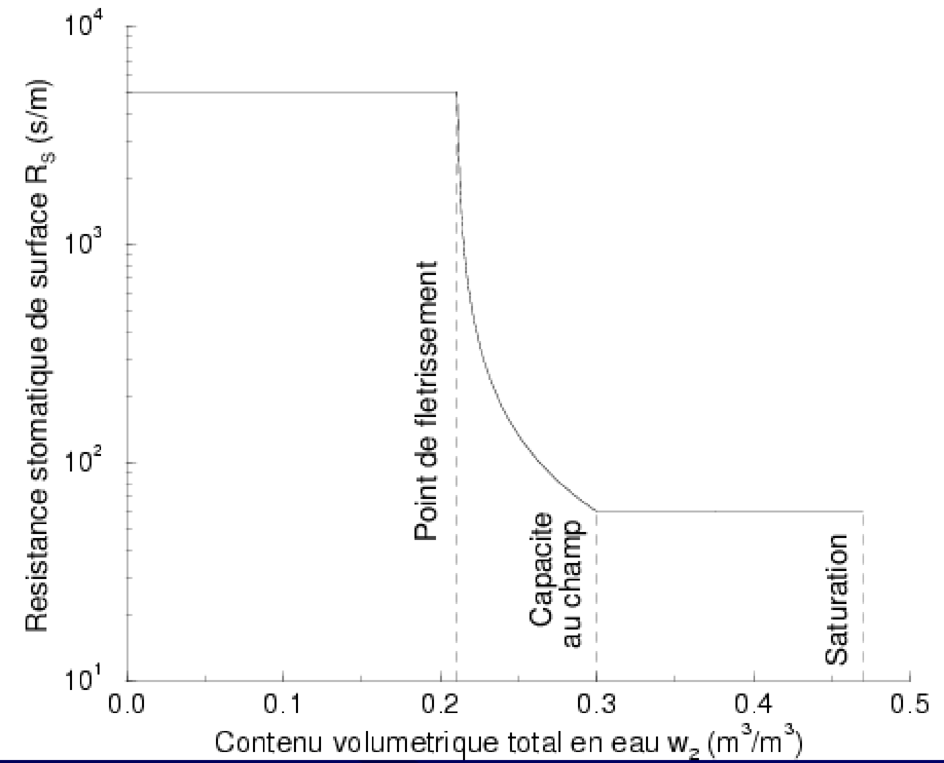






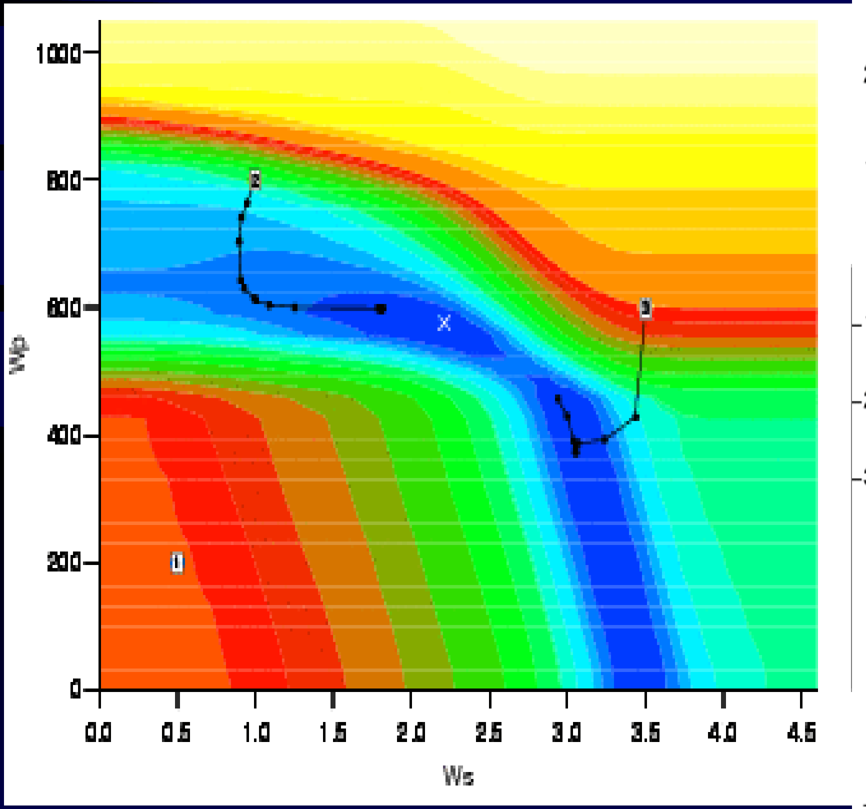
Topology of Jo

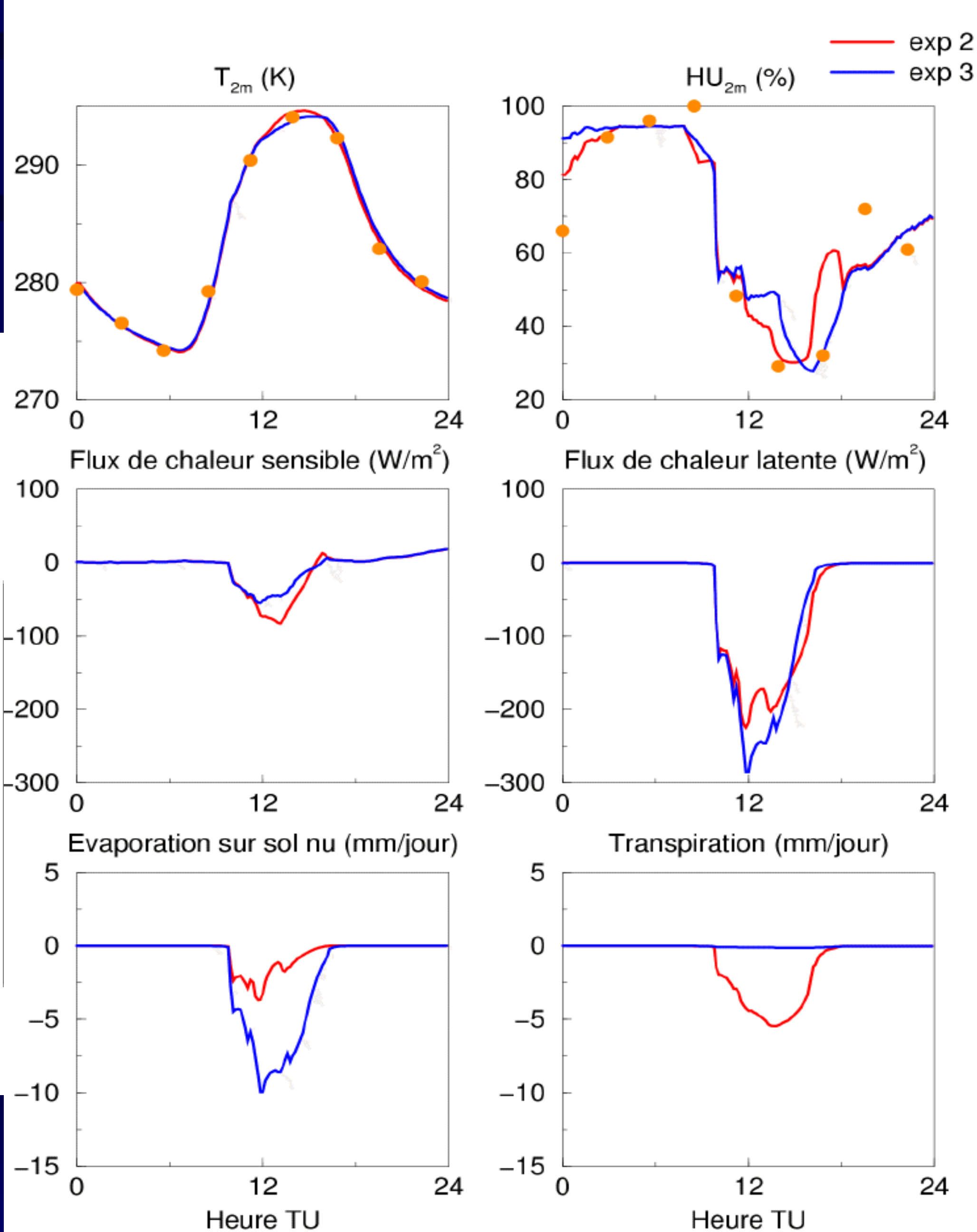




Not a linear problem when we analyse Ws & Wp!

Problem of convergence

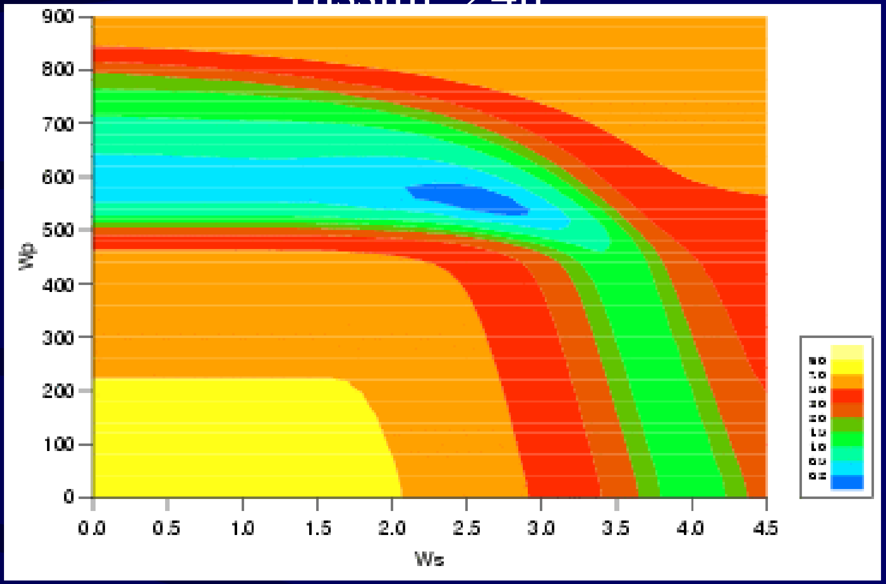


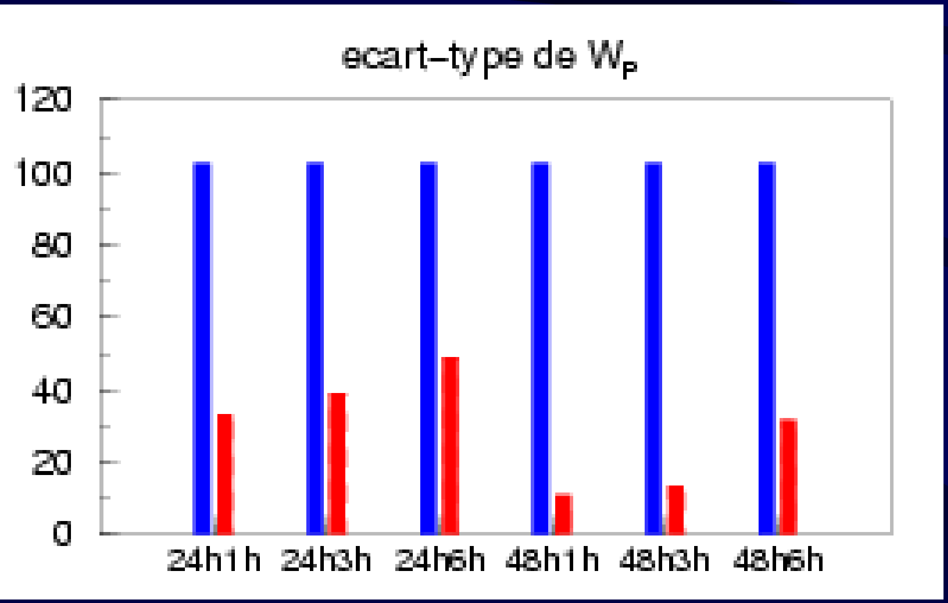


Longer assimilation window

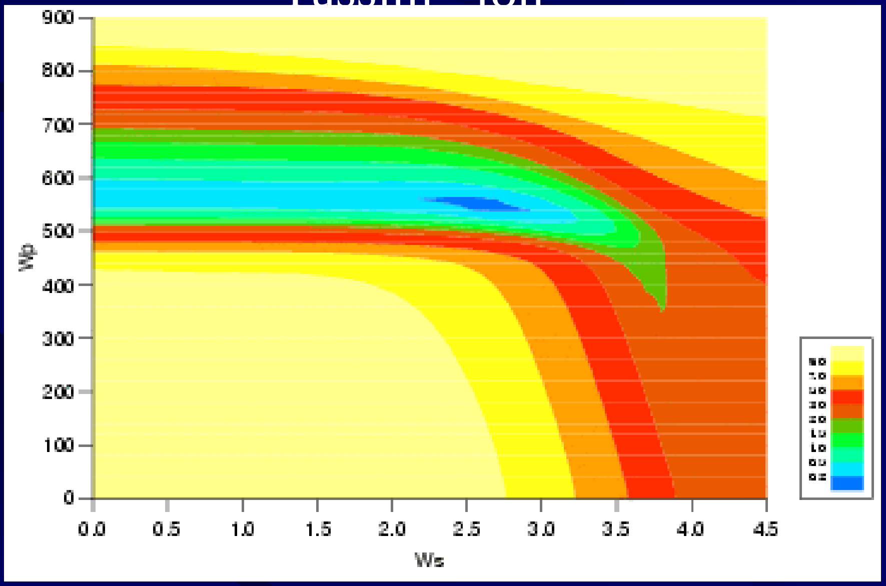
Simulated and noisy obs (1°, 10%)
Fobs = 3h
J=Jb+Jo   100 analyses
$X^G = X^T + \varepsilon_G(0,\sigma)$ with Ts=2°
Tp=1° Ws=1mm Wp=100mm

Tassim=24h





Tassim=48h

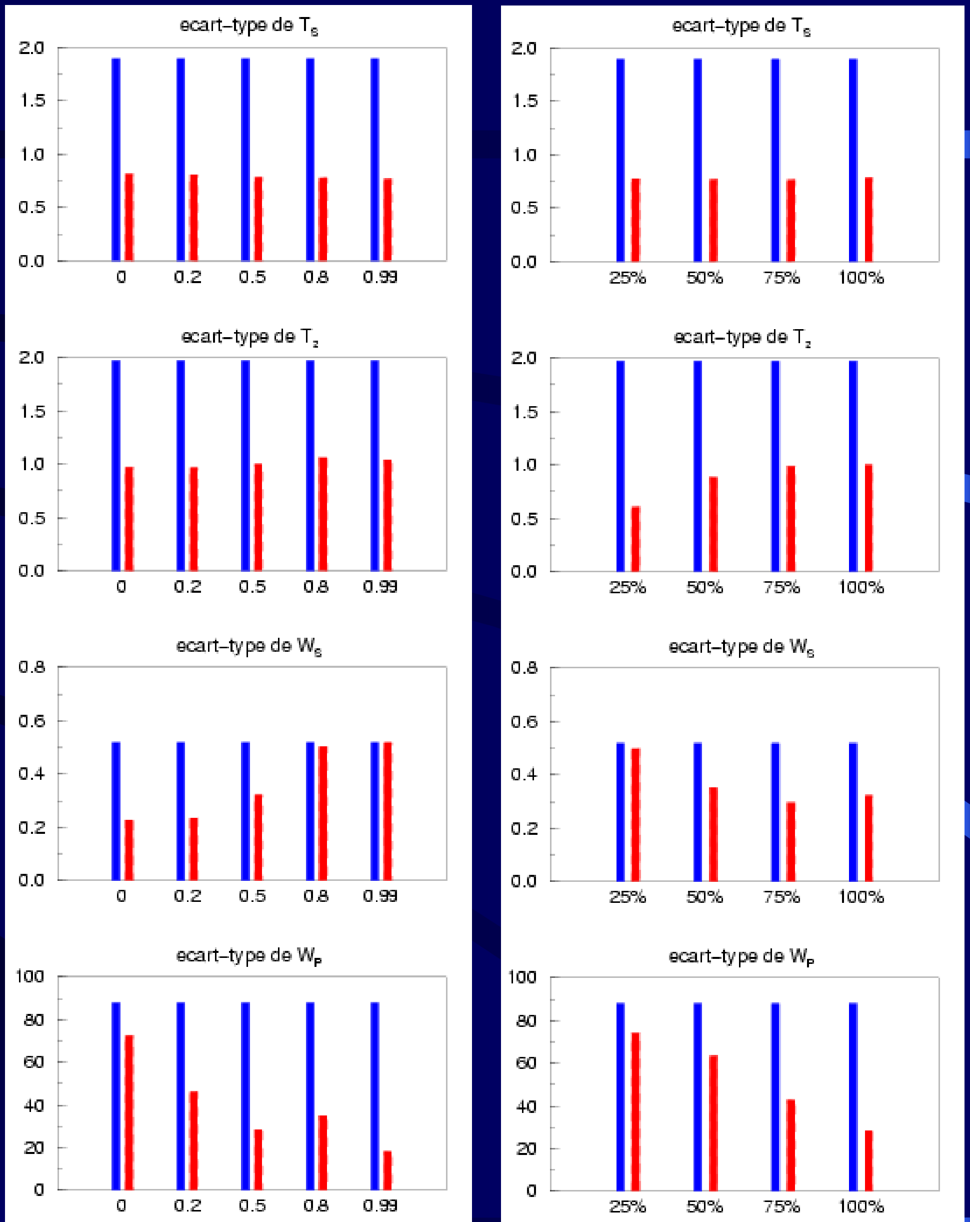


Informativity of 2m observations on surface fields

Variability of the relations between surface and 2m fields

Simulated and noisy obs (1°, 10%)
Fobs = 3h Tassim=24h
J=Jb+Jo 100 analyses
$X^G = X^T + \varepsilon_G(0,\sigma)$ with Ts=2°
Tp=1° Ws=1mm Wp=100mm

⇒Influence of the fraction of vegetation
⇒Influence of the solar radiation



1D experiement on MUREX observation site

MUREX : Météo-France, CESBIO, LTHE

South of France (alt:240m) during 3 years : 85, 86, 87

⇒ Precip, Ra, Rg, H, G, P_s, T_{2m}, q_{2m}, V_{10m} 30 min

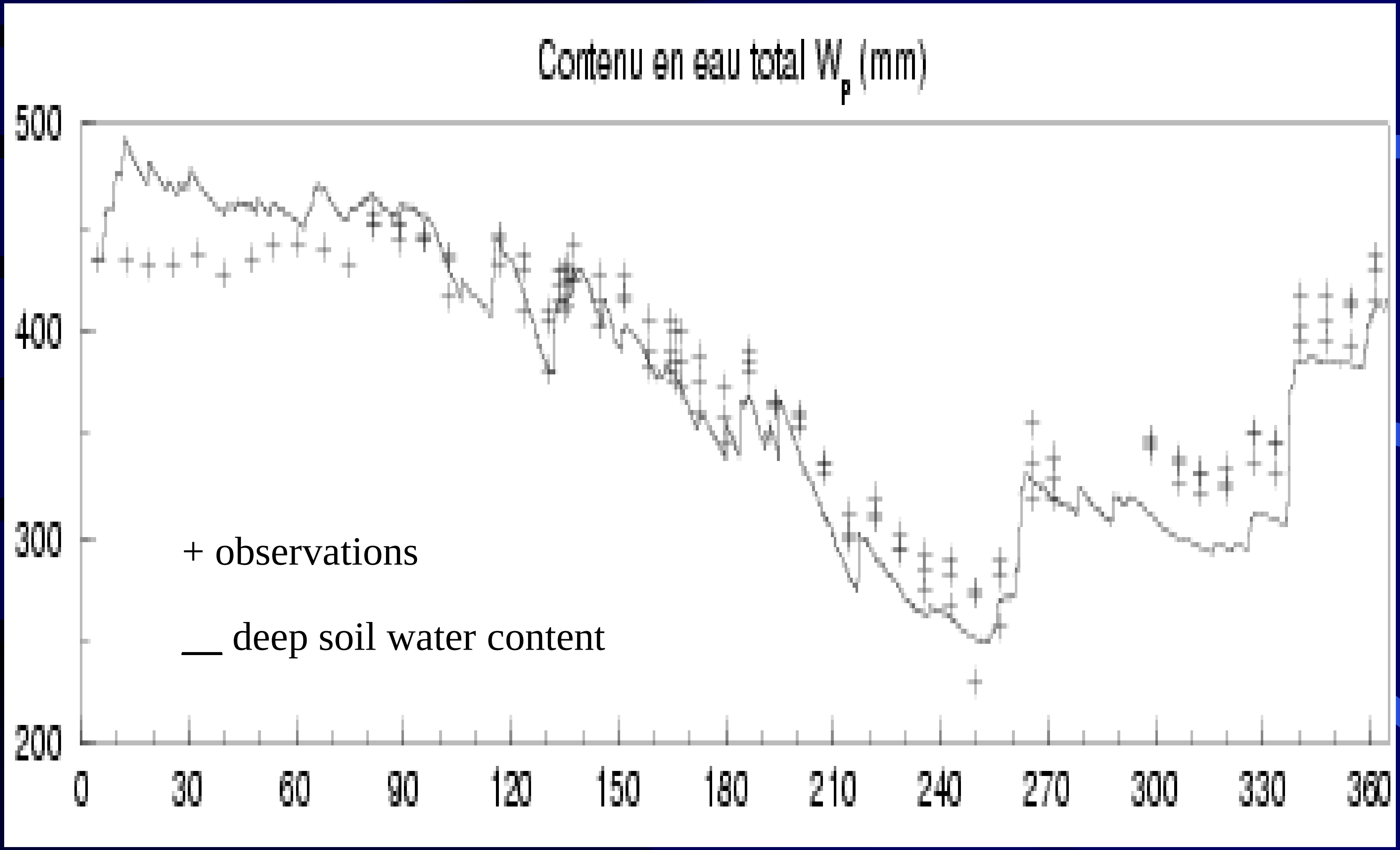
⇒ Deep soil water content (1.30m / 10cm) hebdomadaire

⇒ Superficial water content (5cm / 1cm) 2 POI / 30 min

⇒ Surface temperature at 1cm and Ts infrared 30 min

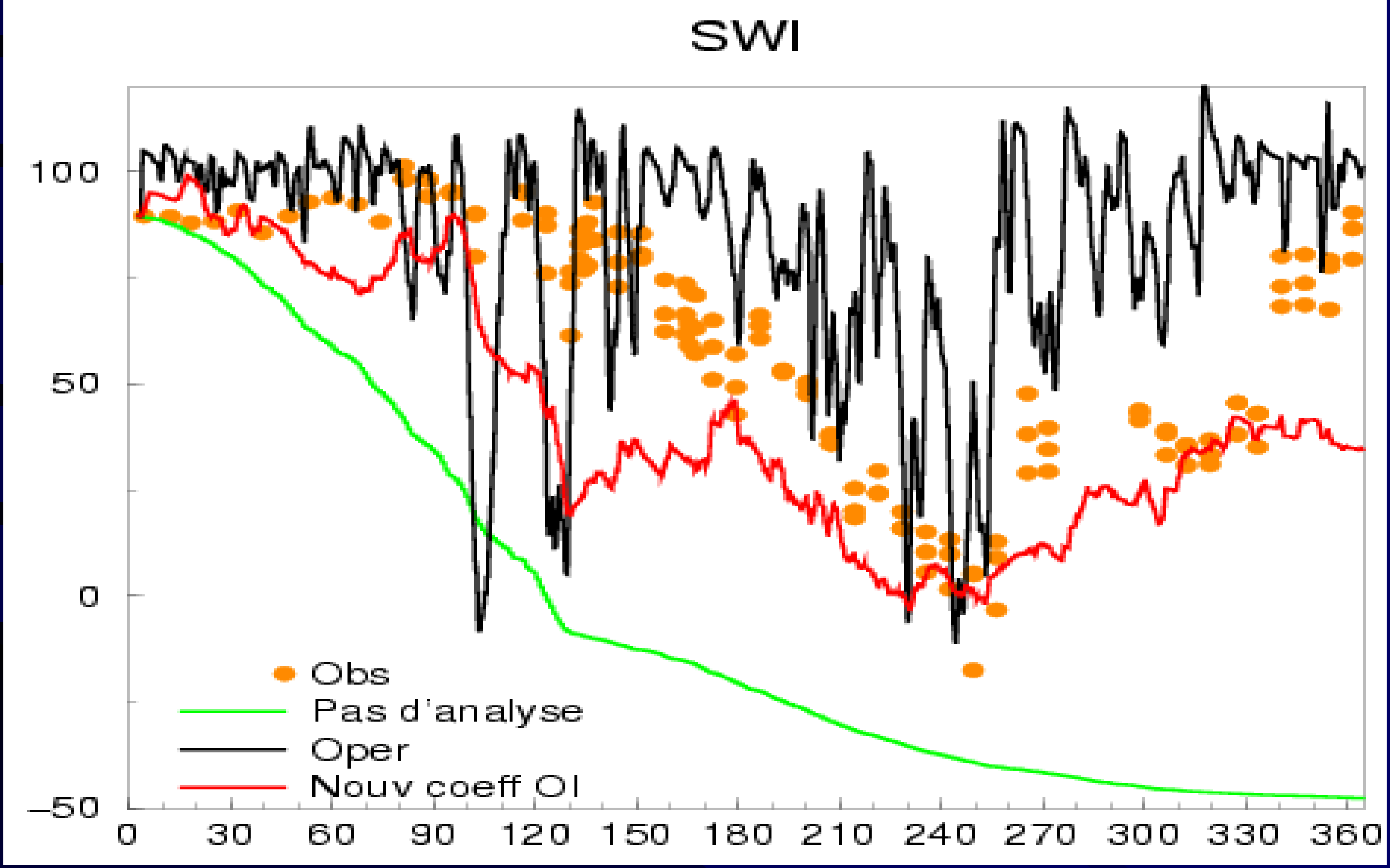
ISBA run with observed fluxes (RR,Rad,CLS)

Wp



Julian days

Optimal Interpolation

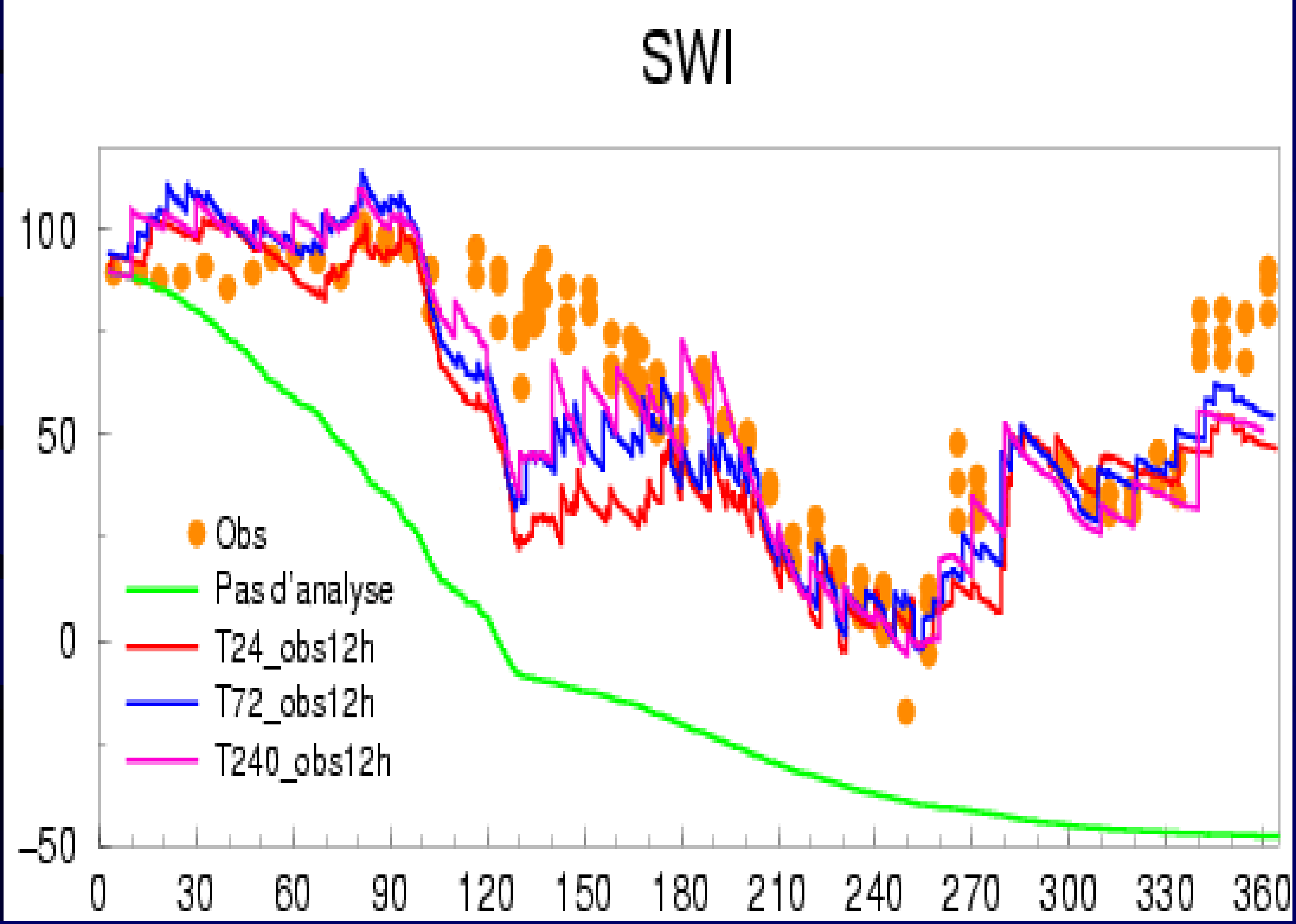


SWI

Julian days

Real observations : Variational Analysis

SWI



Julian days

Variational analysis with 2m obs (pros & cons)

Pros/Cons identical with Optimal interpolation except:

+ Take optimally and implicitely into account physiography and meteorological conditions dependency. No need of Montecarlo method to estimate background error correlation errors between surface variable and T2m, HU2m
+ Assimilation of asynchroneous observations
+ Long assimilation window improve the analysis of slow evolving analysed variables (Tp, Wp): reduce the importance of B

- Much more costly
- Requires the adjoint if many surface variables to analyse (many tiles)

Dynamical optimal interpolation

Hess (2001), Balsamo et al. (2002)

TL hypothesis : $H(\mathbf{x}+d\mathbf{x}) \cong H(\mathbf{x}) + \mathbf{H}.d\mathbf{x}$ (acceptable for Wp)

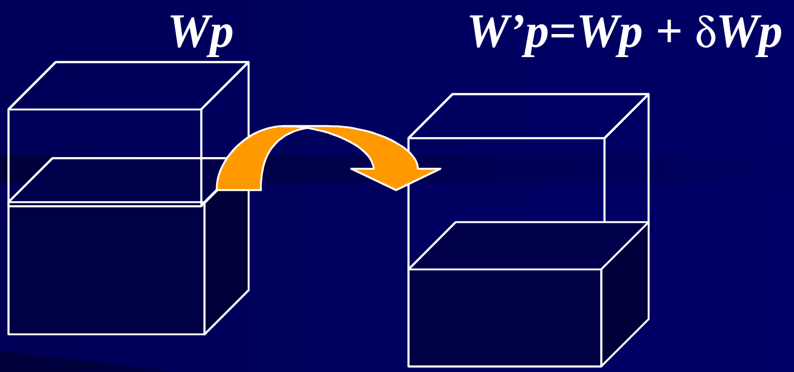
$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{BH}^T(\mathbf{HBH}^T + \mathbf{R})^{-1}(\mathbf{y} - H(\mathbf{x}^b))$$

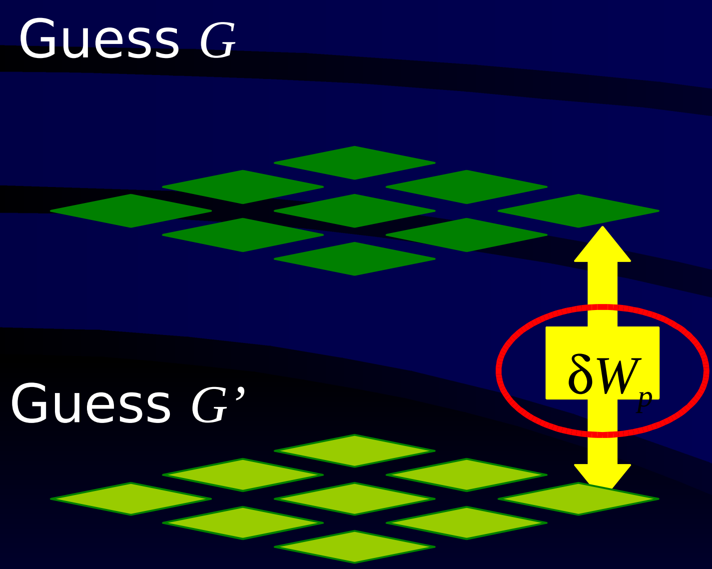
$$W_p^a - W_p^b = \alpha_{WpT}\,\Delta T_{2m} + \alpha_{WpRH}\,\Delta RH_{2m}$$

- “Normal” OI coefficient α_{WpT} and α_{WpRH} are evaluated statistically (once)
- Dynamical OI coefficients α_{WpT} and α_{WpRH} are evaluated dynamically (each time)

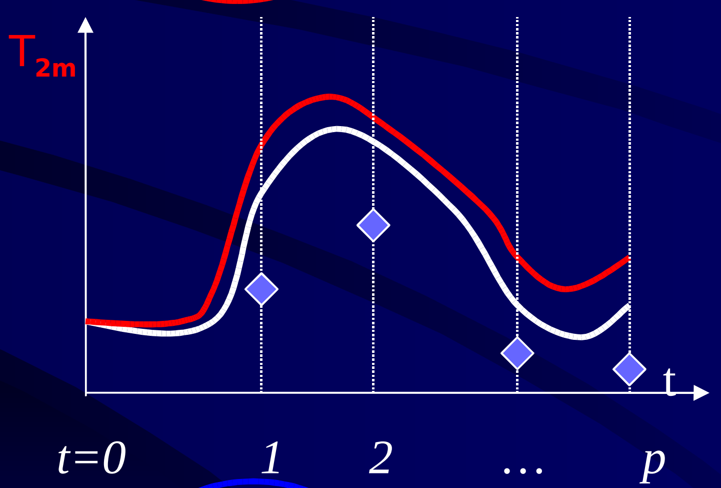
3D study of dynamical OI : method

From a perturbation of the initial total soil moisture δ ***Wp*** applied on each model land grid-point.





$$\delta T_{2m}^{(i)} = T_{2m}^{G\,(i)} - T_{2m}^{G'\,(i)}$$



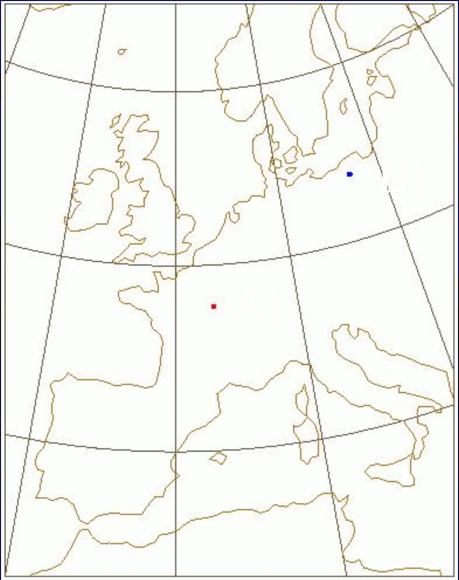
$$\mathbf{H}^T = \begin{pmatrix} \frac{\delta T_{2m}^{(1)}}{\delta W_p} \\ \frac{\delta RH_{2m}^{(1)}}{\delta W_p} \\ \dots \\ \frac{\delta T_{2m}^{(p)}}{\delta W_p} \\ \frac{\delta RH_{2m}^{(p)}}{\delta W_p} \end{pmatrix}$$

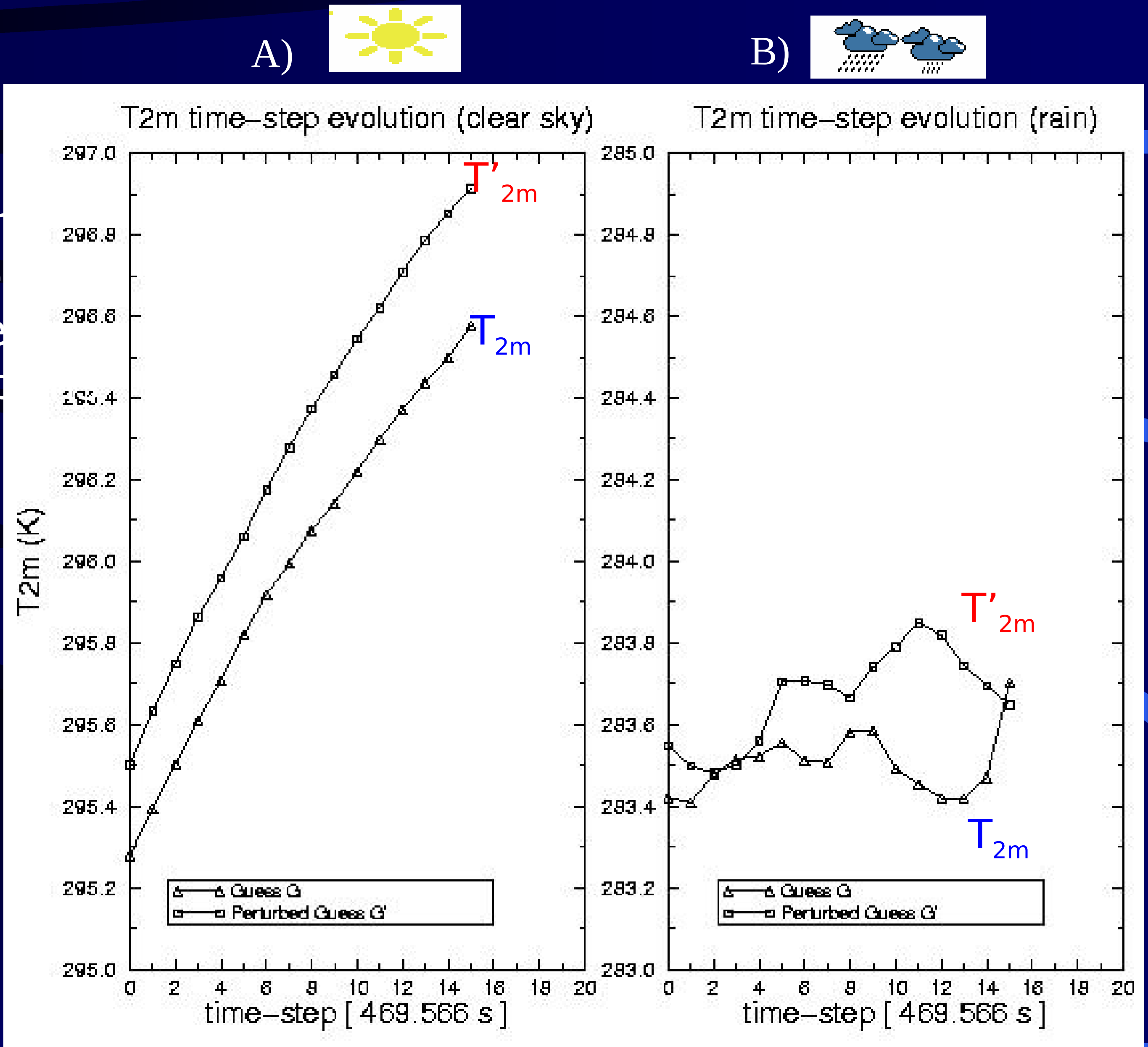
$$y - H(x_b) = \left(\Delta T_{2m}^{(1)}, \Delta RH_{2m}^{(1)}, \dots, \Delta T_{2m}^{(p)}, \Delta RH_{2m}^{(p)}\right)$$

$$\Delta T_{2m}^{(i)} = T_{2m}^{G\,(i)} - T_{2m}^{O\,(i)}$$

Tangent Linear hypothesis

Considering a real situation (16th June 2000 at 12UTC), a sensitivity test or initial soil moisture is run under differe atmospheric condit

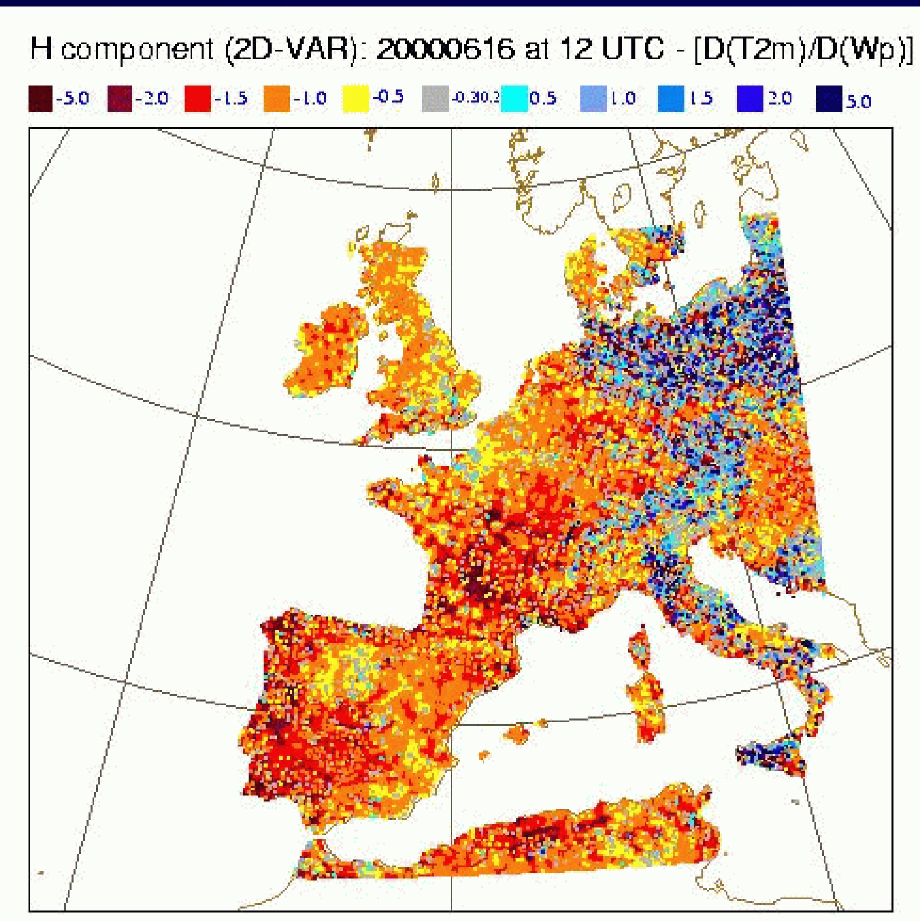


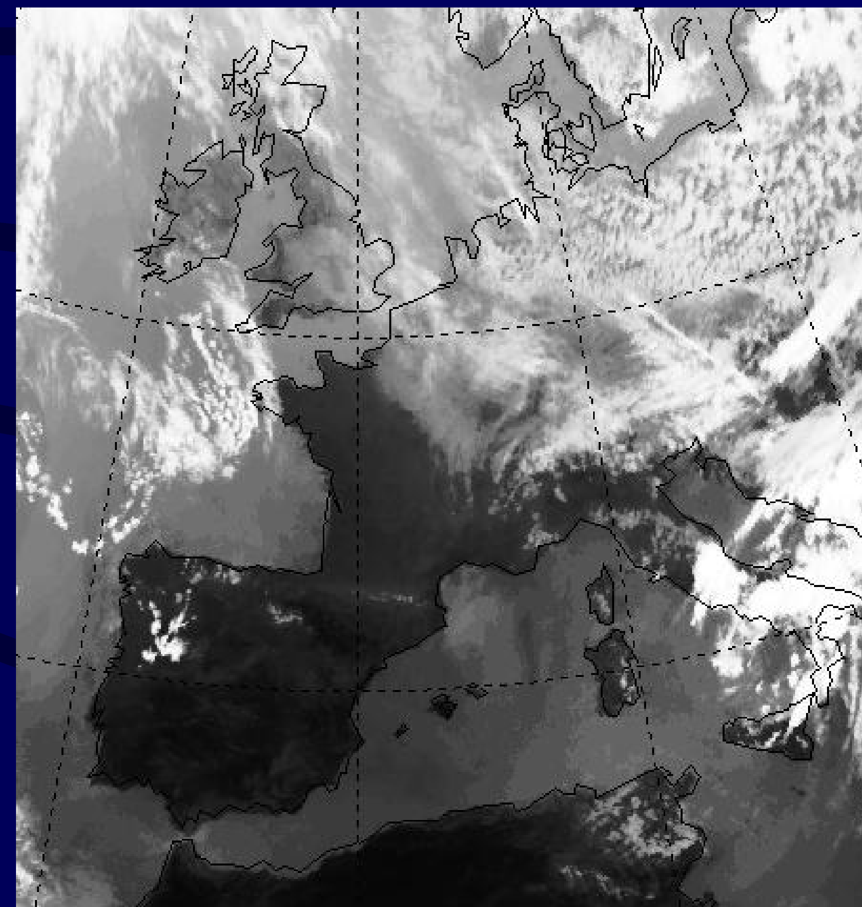


Tangent Linear hypothesis

Calculate the **H** matrix t $\frac{\delta T_{2m}^{(1)}}{\delta W_p}$ [K/m] and compare it with the satellite image

20000616 at 12 UTC





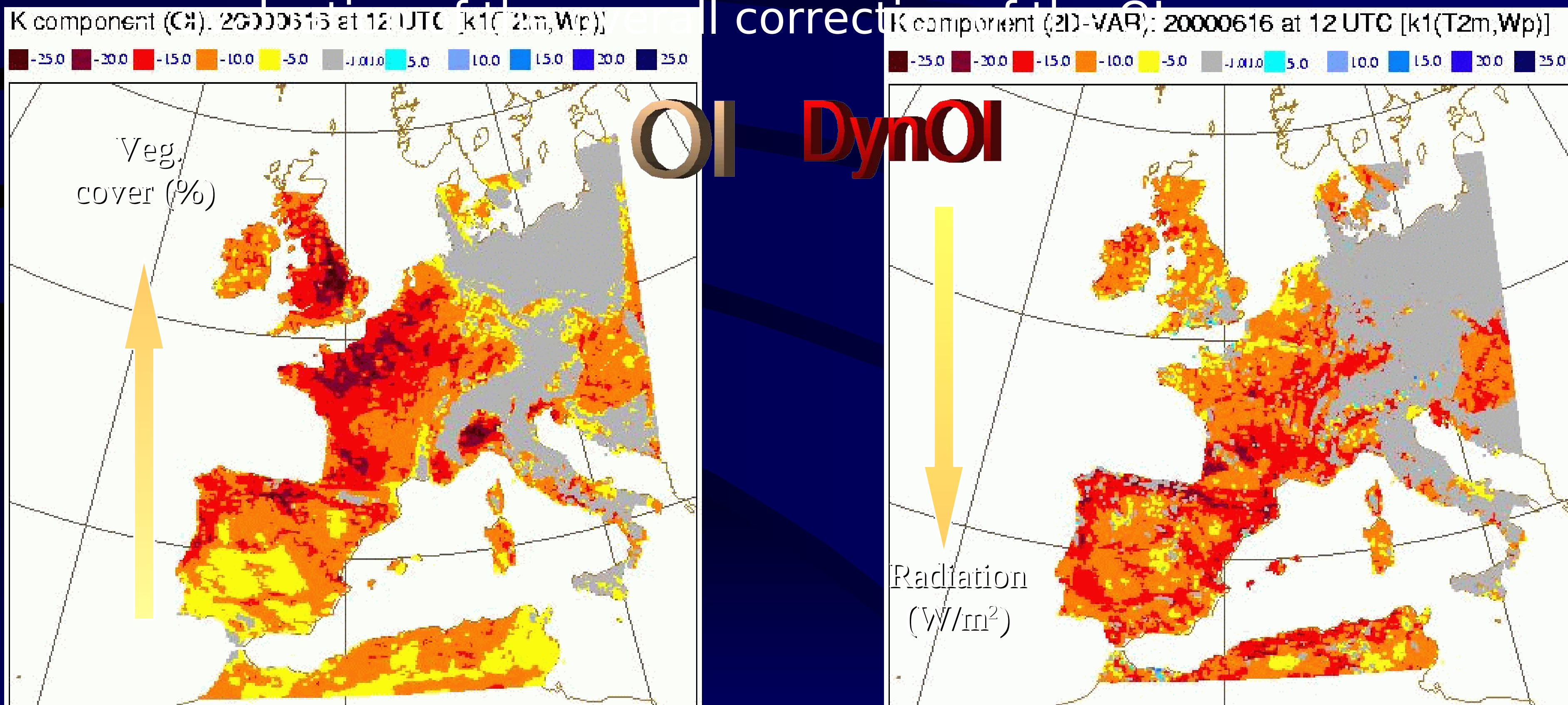
- Switch off analysis on some meteorological conditions or ensemble estimation by running several perturbations

Comparison of statistical and dynamical OI

A comparison with OI (Gain Matrix and OI coefficients) is useful to point out

some properties of the variational approach

- masking of low sensitivity grid-points (coherence of masked areas)
- dependency from radiation rather than vegetation
- ...l correcti...



DynOI analysis with 2m obs (pros & cons)

Pros/Cons identical with Optimal interpolation except:

+ Take optimally and implicitely into account physiography and meteorological conditions dependency. No need of Montecarlo method to estimate background error correlation errors between surface variable and T2m, HU2m
+ Assimilation of asynchroneous observations
+ Long assimilation window improve the analysis of slow evolving analysed variables (Tp, Wp): reduce the importance of B
+ Similar to extended Kalman filter if we provide an evolution of B and prescribe model error statistics (Q)
- Much costly

Pros/Cons identical with Variational method except:

+ Much less costly
- Assume linearity of obs operator which is not perfectly true, particularly for analysing variables with smaller time scale evolution (Ws for instance)

ERS-1/2 scatterometer derived soil moisture (ASCAT coming soon)

Data set produced by:

Institute of Photogrammetry
and Remote Sensing,
Vienna University of Technology

Basis:

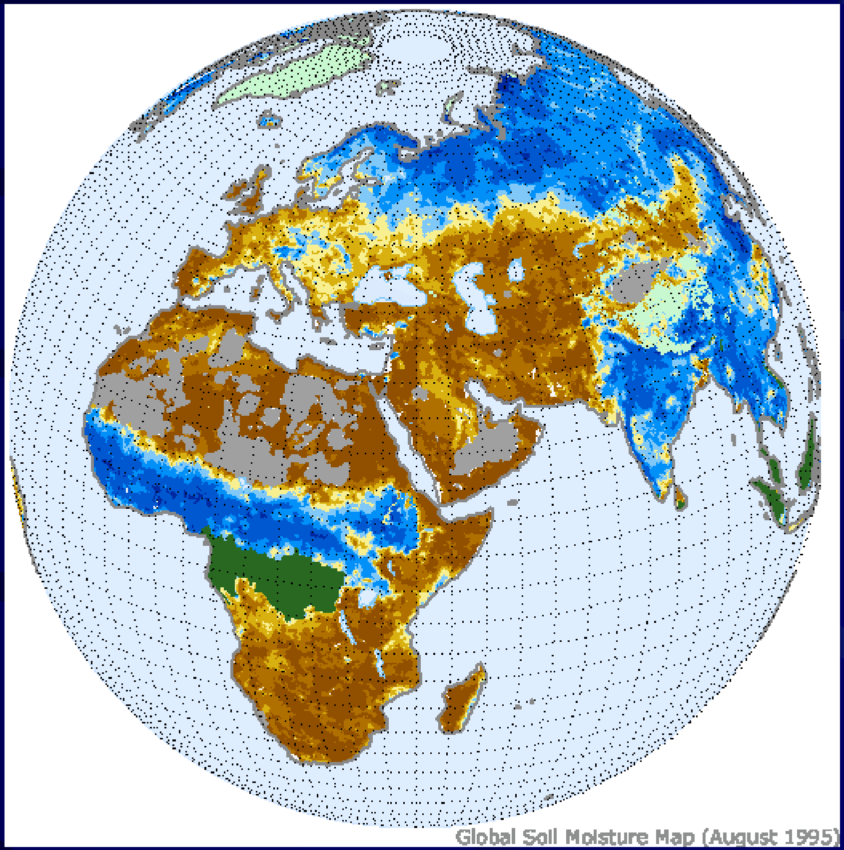
ERS scatterometer backscatter measurements

Method:

change detection method for extrapolated backscatter at 40° reference incidence angle

Output:

topsoil moisture content in relative units (0 [dry] to 1 [wet])

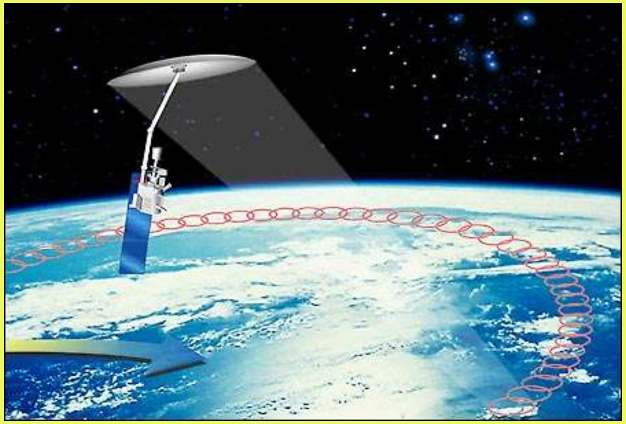


http://ipf.tuwien.ac.at/radar/ers-scat/home.htm

Use satellite observations over land

L-band Tb

HYDROS NASA mission (2010)



SMOS ESA mission (2007)

C-band Tb

AQUA AMSR-E instrument

(05/2002

IR Ts

Geostationnary satellites

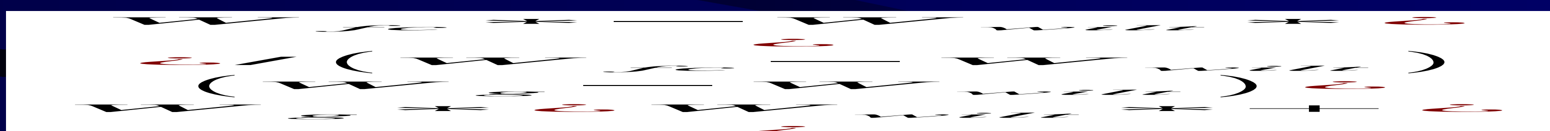
MUREX

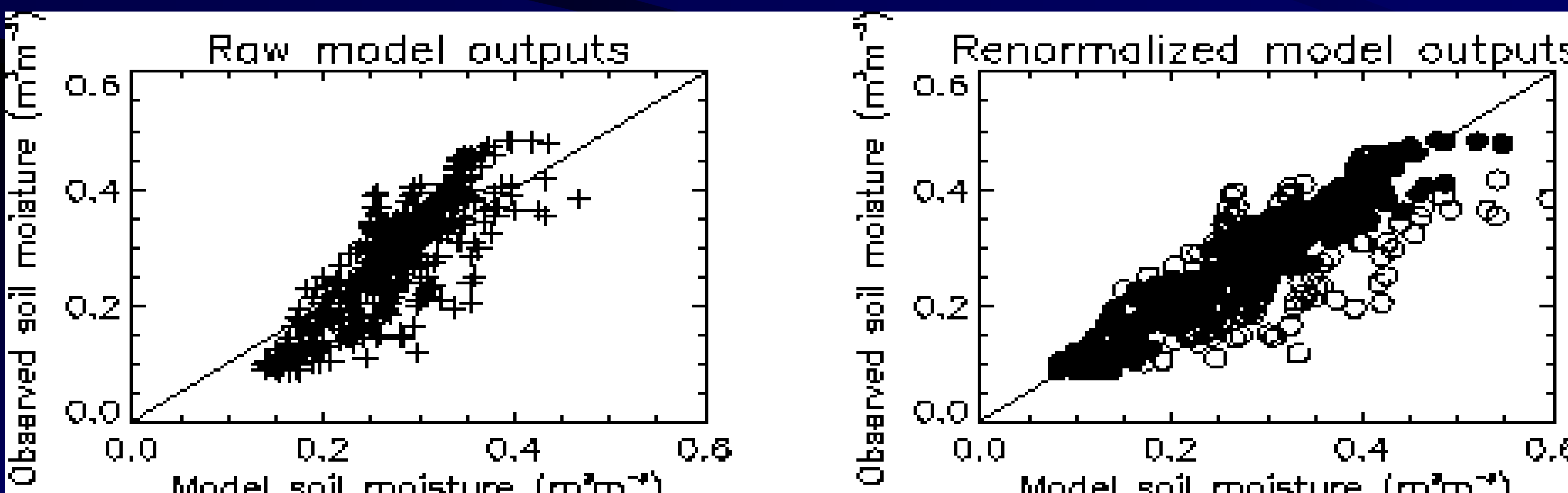
BIAS PROBLEM WILL BE VERY IMPORTANT AND DIFFICULT FOR SATELLITE OBS:

Necessite to normalize superficial soil moisture over MUREX:

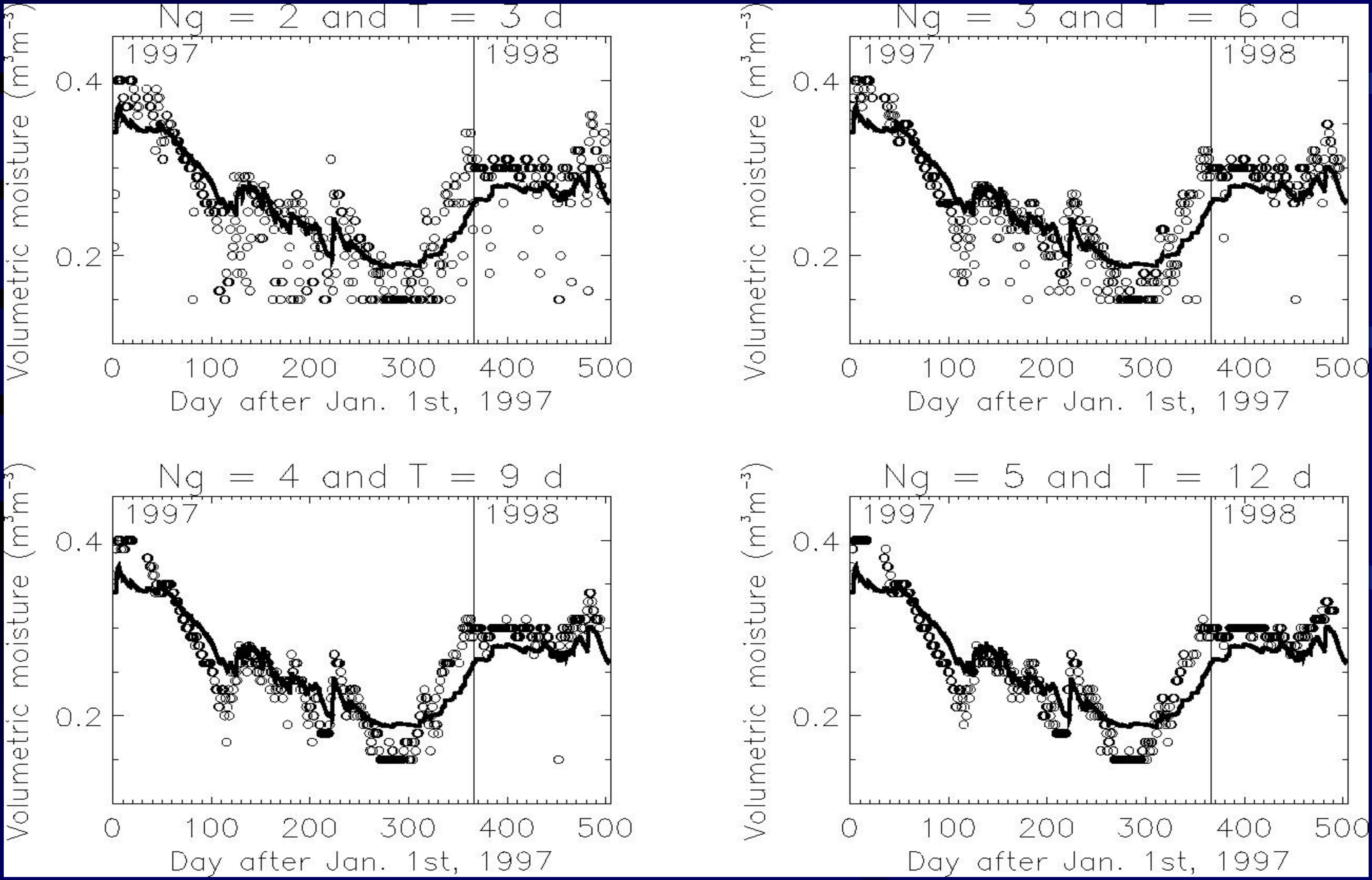
Some errors in surface processes:

- Vertical gradients of soil texture/structure?
- Depth of superficial layer
- Etc.





Variational assimilation of observed superficial soil moisture with no first guess over MUREX (Calvet & Noilhan 2000)



Others analyses:

Sea surface temperature

Sea ice

Snow

…

SST analysis

Current NWP models consider SST constant during the forecast (will change in the near future)

Observations : Buoys, Ship, Satellite radiances (infrared)

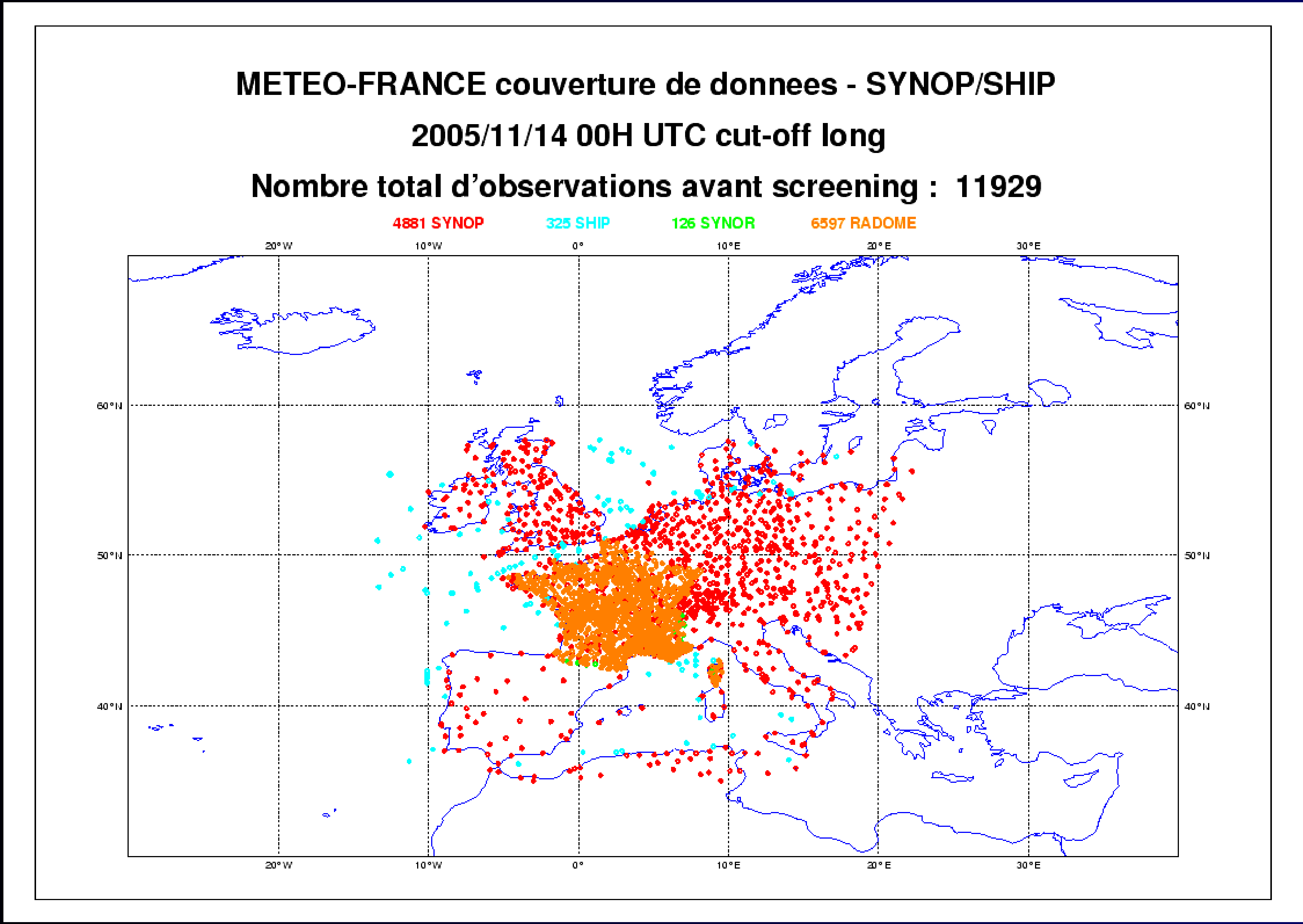
Satelitte IR observations provide very high spatial and temporal resolution with global coverage under clear sky conditions. Better quality during night. Lower quality if no wind.

Available SST analysis by non NWP centers: NESDIS analysis (0.5°*0.5° replaced soon by 0.125°*0.125°), SAF-OSI analysis over Europe

Sea ice analysis

Satellite microwave observation to determine sea ice fraction (SSMI used operationnally)

Buoys and ship SST observations in blue



Snow analysis

Snow is generally represented by snow cover and snow water content

Observations : SYNOP (kg/m2), Satellite observations (infrared & microwave)

Satellite IR observations provide snow cover very high spatial and temporal resolution with global coverage under clear sky conditions

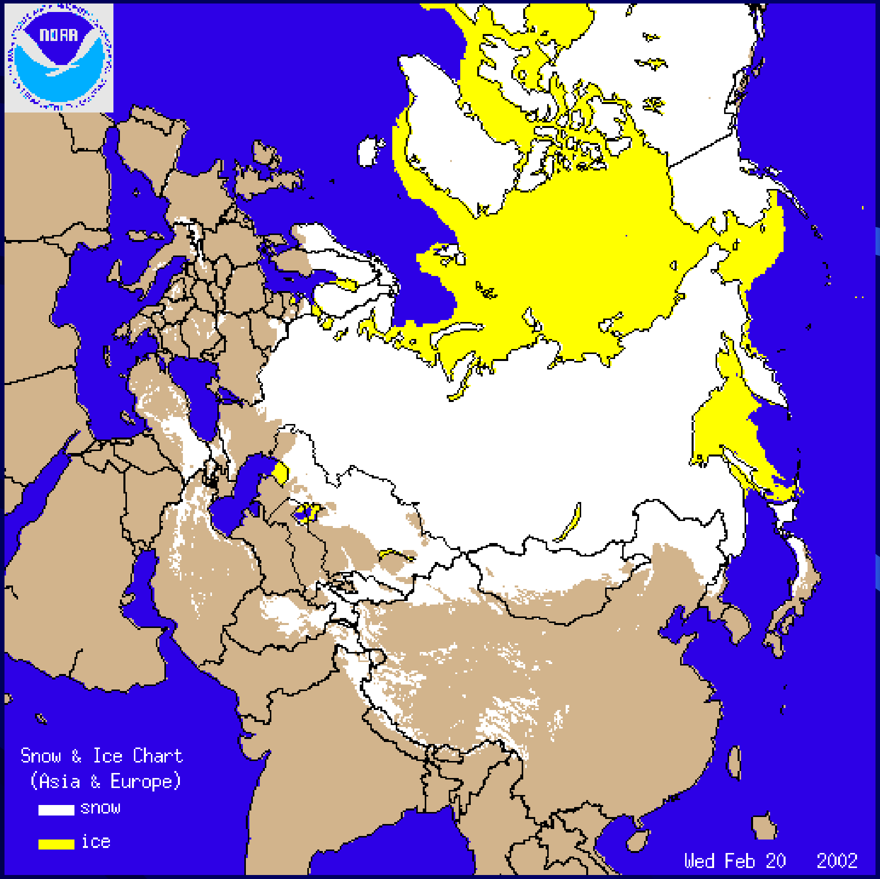
Satellite MW observations provide snow cover 1-2/day and low resolution (several tens of kms) with global coverage. The retrieve of snow water content is very difficult because of the snow granulosity.

The combination of satellite and SYNOP observations is interesting to obtained both snow cover analysis and snow water content. However it is very difficult to specify the B and R statistics statistics and H observation operators in mountains.

NOAA / NESDIS Snow Extent

Interactive Multisensor Snow and Ice Mapping System:

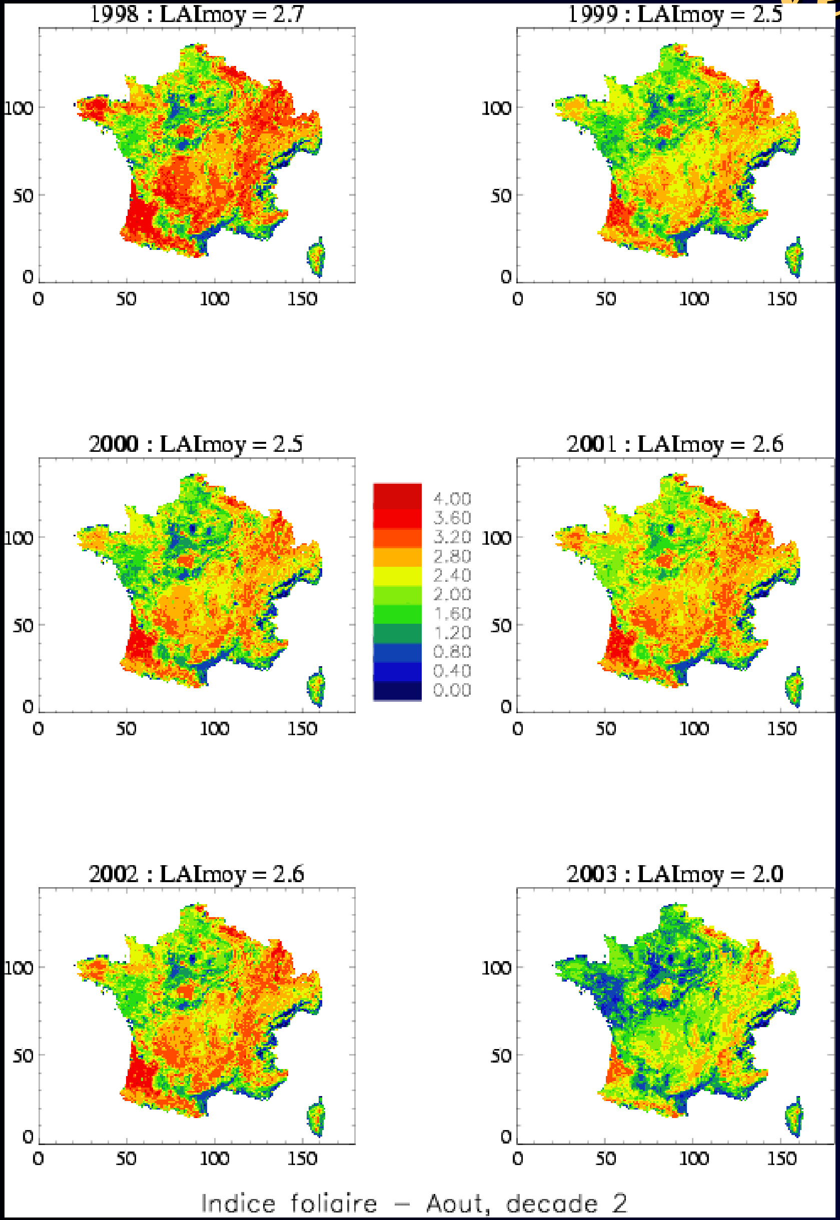
- time sequenced imagery from geostationary satellites,
- AVHRR,
- SSM/I,
- station data,
- previous day's analysis



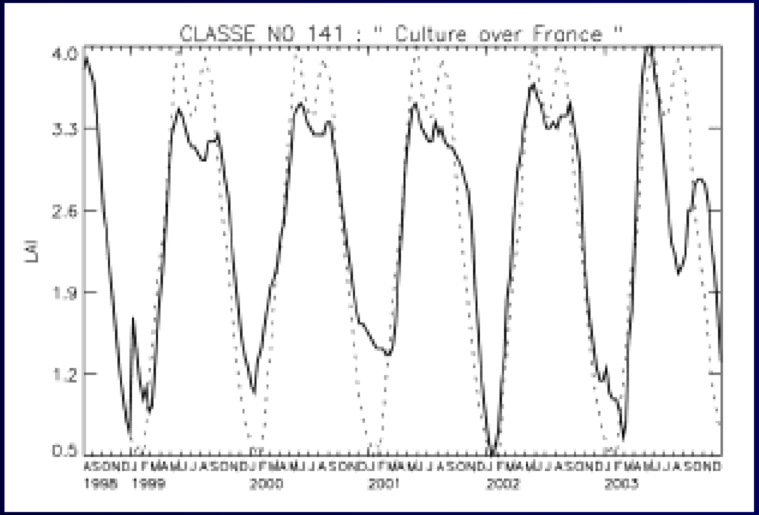
Northern Hemisphere product

- real time
- polar stereographic projection
- 1024 × 1024 elements

Assimilation of vegetation parameters (LAI, Veg, ...)



Intervariability of LAI over France



Conclusions and Perspectives

Surface analysis is very important in NWP models (importance is increasing for higher resolution models)

Separate surface analyses are for the time being preferred for simplicity. But will it continue like it?

Soil moisture is currently the most difficult parameter to analyse and very important since it determines ratio between sensible and latent heat fluxes

Use of « dynamical OI » or « EKF » with analysed atmospheric forcings is attractive (investigated during ELDAS project) since it allows asynchroneous observations and uses model physics to compute observation operator

Soil moisture and soil temperature analysis is treated for each grid-point separately. It is very convenient and allows « dynamical OI » or « EKF » but requires first a spatialization of observations.

The use of a full 3D surface variational analysis requires adjoint model which is not obvious for land surface model (non linearities, more and more surface processes)

Conclusions and Perspectives

T2m, H2m observations are very powerful to correct model errors but an atmospheric model or a PBL model is till now necessary which is very costly compared with off-line system. Research are currently for solving this problem.

Combination of in situ and satellite observations

Satellite observations will be more and more used (infrared, microwave) for surface analyses

New satellite observations for soil moisture analysis (SMOS, HYDROS)

Biais correction and dealing with surface heterogeneity will be very important for the assimilation of satellite observation

Analysis of vegetation parameters (LAI, Veg, ...) combined with soil moisture analysis

Lake surface temperature analysis