DATA ASSIMILATION

Norman L. Miller
Geography Department, University of California, Berkeley
Earth Sciences Division, Berkeley National Laboratory

30 October 2012

OUTLINE

PART I

- **Historical Perspective and Overview**
  - **A Brief History of Observation Networks**
  - **A Brief History of Data Assimilation**
  - **Recent Observation Networks**
- **Part II**
- **Data Assimilation and Numerical Weather Prediction**
- **Data Assimilation Basics**
- **Data Assimilation and NWP Analysis**

- **PART III**
- **Application to Land Surface and Hydrologic Modeling**
- **Application to Regional Climate Modeling**

A Brief History of Observation Networks

- Chinese - daily records of WX obs since the Ming Dynasty

- 1780s - Palatine Academy of Science and Letter, Germany organized the first international observing network, Obs collected from Massachusetts to the Ural Mountains.

- 1855 - Urbain LeVeVerrier, director of the Paris Astronomical Observatory, under Napolean III established the first national meteorological observing network.

- 1873 - International Meteorological Conference (Vienna) formed a committee for standardized measurements.

- 1900 - initial global observing network was in place, but it took another 75 years for general adoption.

A Brief History of Data Assimilation

Bjerkes (1911) - the ability to forecast the weather was posed as the *Ultimate Meteorology Problem* and required two conditions to be satisfied:

(1) The present state of the atmosphere must be characterized as accurately as possible

(2) The instrinsic laws, according to which states develop out of preceding ones, must be known.

These two conditions define weather prediction as an **initial-value problem** requiring:

- observations
- diagnostics and analysis
- prognostic modeling

Recent Observation Networks

•1950's - Upper air measurement network, but large gaps over oceans, especially the southern hemisphere.

•1960's - satellite radiometers were developed and launched with <u>international coverage</u>.

•1979 - Global Weather Experiment, Goal: Observe the atmosphere as systematically as possible.

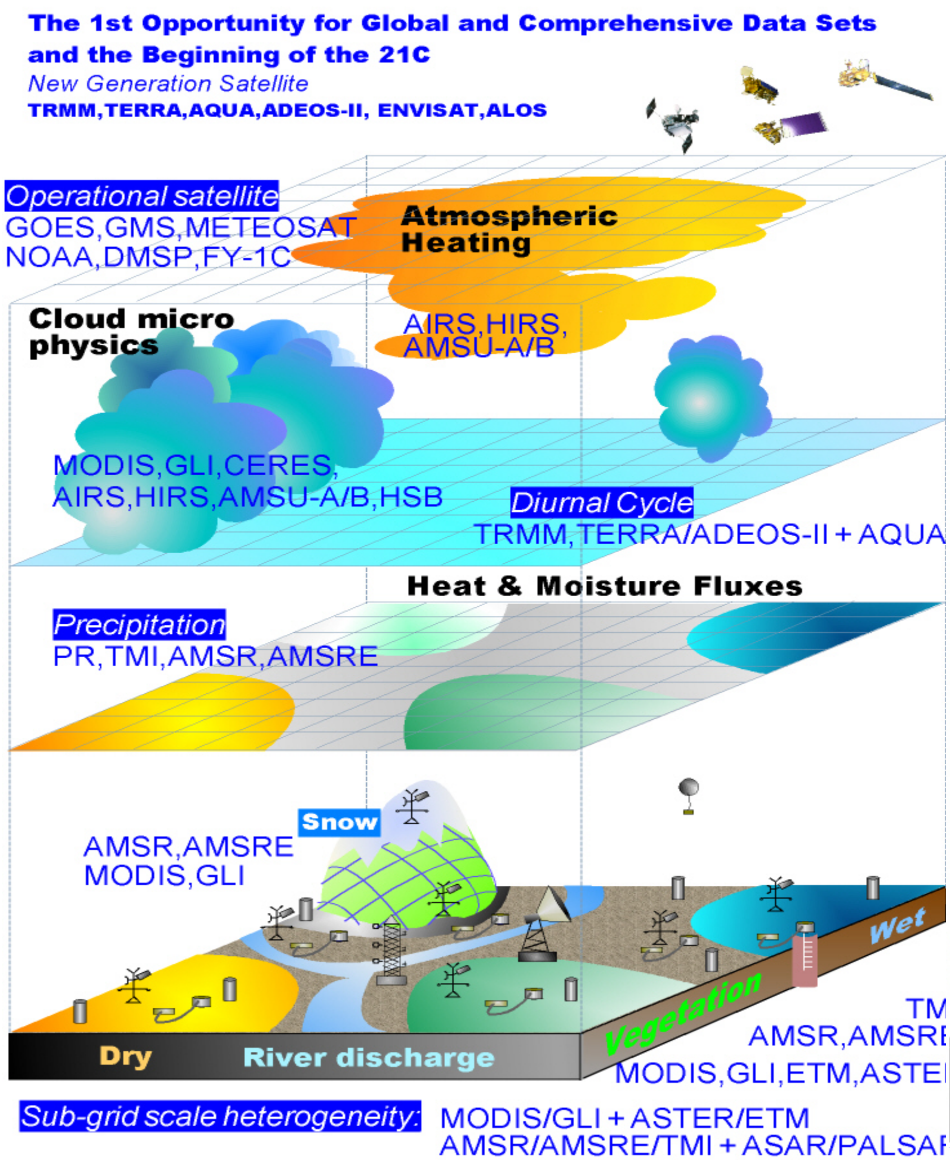
•1990's - World Meteorological Organization (WMO) has major observing programmes,

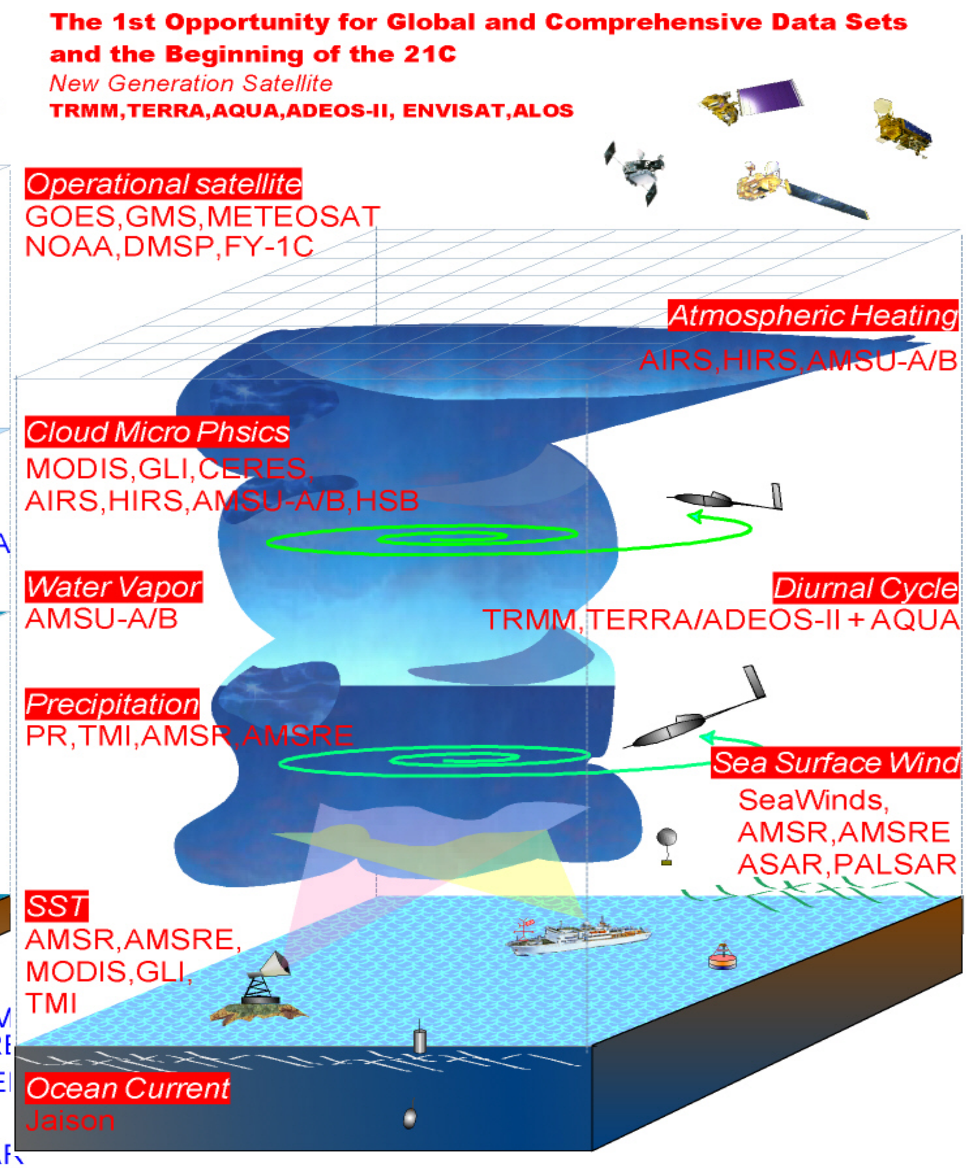
- •GOS - Global Observing System.
- •GTS - Global Telecommunications System,
- •GDPS - Global Data Processing System.

Global Earth Observing System of Systems (GEOSS)

- 2005 - Earth Observations (EO) Summit adopts a 10-year implementation plan.
- Vision - " … to realize a future wherein decisions and actions for the benefit of humankind are informed by *coordinated, comprehensive, and sustained* Earth observations and information. … and to <u>enhance prediction of the behavior of the Earth system</u>."
- GEOSS elements:
  - Reducing loss of life and property from natural or human induced disasters.
  - Understanding factors affecting human health and well-being
  - Understanding, predicting, mitigating, and adapting to climate change
  - Improving water resource management through better understanding of the water cycle
  - Improving weather forecasting, information, and warning
  - Improving management and protection of terrestrial, coastal, and marine ecosystems
  - Supporting sustainable agriculture and combating desertification
  - Understanding, monitoring, and conserving biodiversity

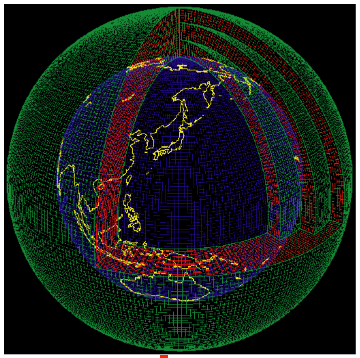
New Data Sets of the *Overall* Water Cycle by *Integrating* the Satellite Products





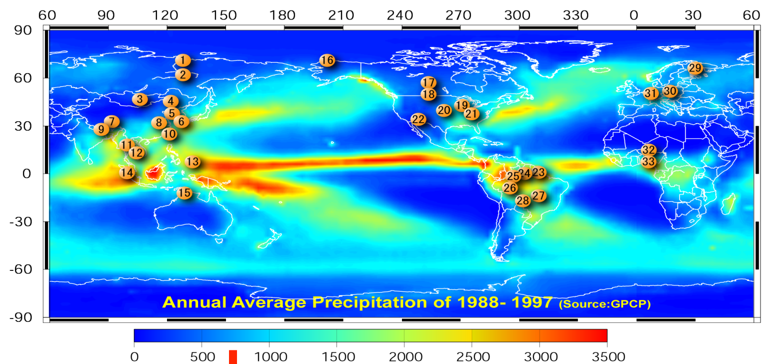
The First Global Integrated Data Sets of the Water Cycle

Model Outputs by Numerical Weather Prediction Centers



MODEL Output Data Archiving Center at Max-Planck Institute of Germany
http://www.mpg.de/

Surface Observational (*in-situ*) Data from the 36 CEOP Reference Sites

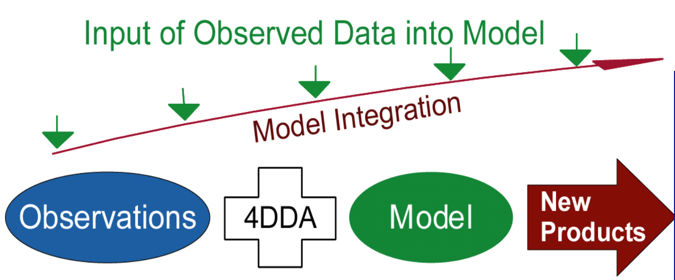


In-Situ Data Archiving Center at UCAR (Center at University Corporation for Atmospheric Research) of USA
http://www.ucar.edu/

Satellite Remote Sensing Data

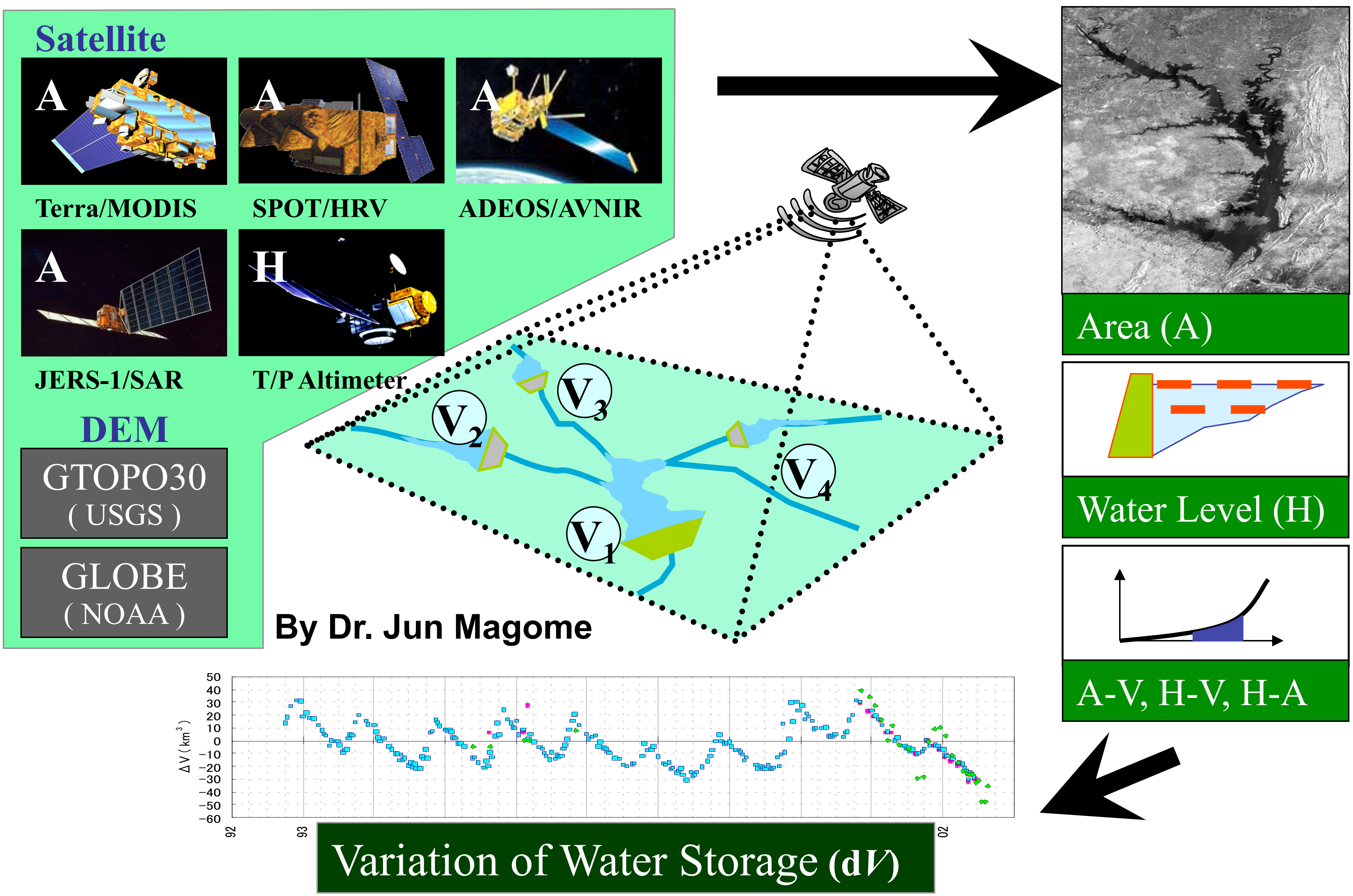
Data Integrating/Archiving Center at University of Tokyo and JAXA of Japan
http://monsoon.t.u-tokyo.ac.jp/ceop/

Data Archive Center



Global Land Data Assimilation System at NASA Goddard Space Flight Center of USA
http://ldas.gsfc.nasa.gov/

⑤ Water Storage Estimation (overview)



PART II

Data Assimilation

and

Numerical Weather Prediction

Data Assimilation

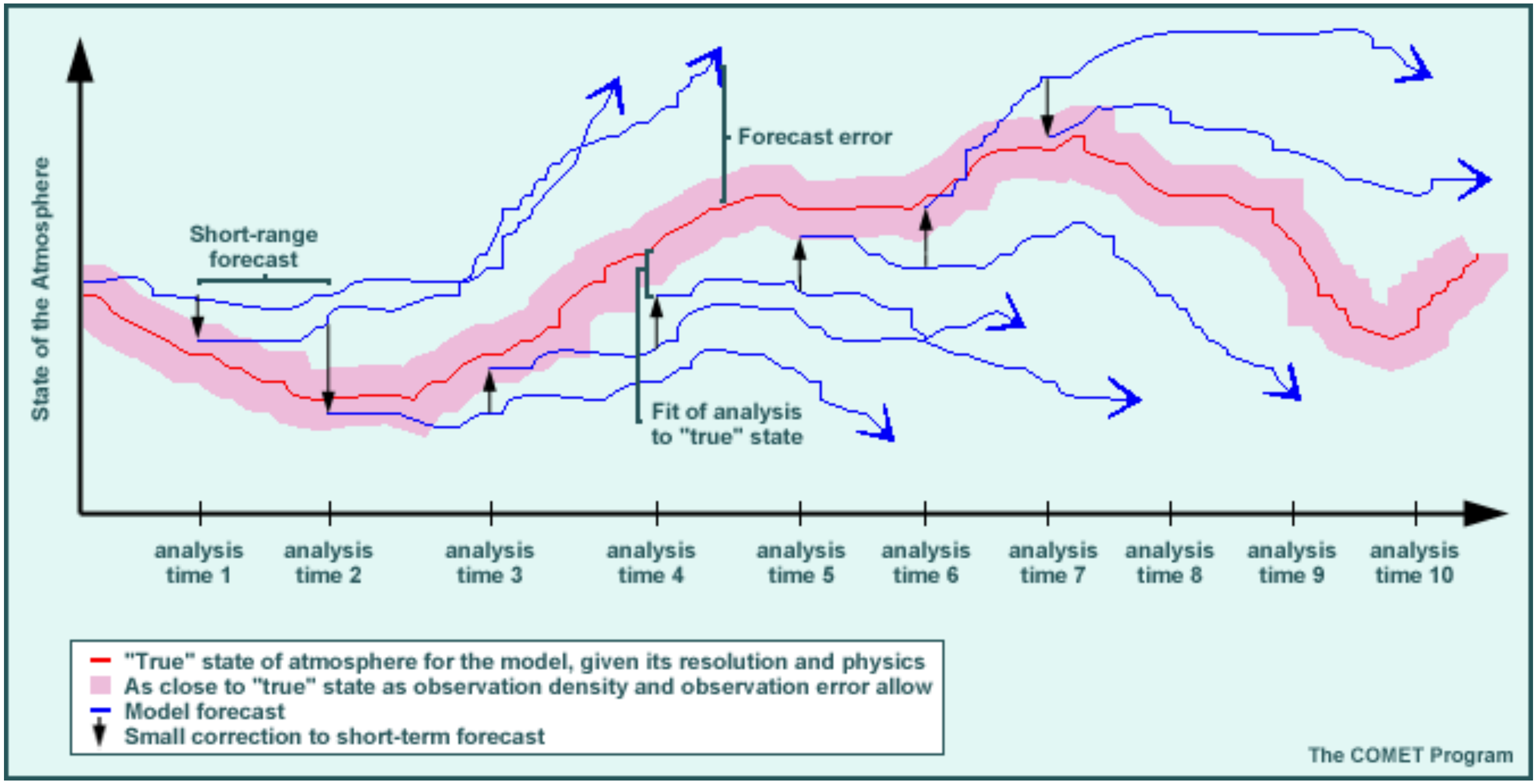
•Data assimilation provides the initial conditions that produce the best possible model forecast.

•It must create an analysis consistent with the model numerics, dynamics, physics, and resolution.

•NWP provides the short-range forecast for the analysis by making a series of small corrections to that forecast based on new information from observations.

• Analyses is different for different models and will most likely differ from the best estimate of the true state of the atmosphere produced by a hand analysis.

•Observations are assimilated to correct each short-range forecast that serves as the basis for the next analysis, resulting in a series of small corrections to the model forecast.



State of the Atmosphere
Short-range forecast
Forecast error
Fit of analysis to "true" state
analysis time 1
analysis time 2
analysis time 3
analysis time 4
analysis time 5
analysis time 6
analysis time 7
analysis time 8
analysis time 9
analysis time 10
"True" state of atmosphere for the model, given its resolution and physics
As close to "true" state as observation density and observation error allow
Model forecast
Small correction to short-term forecast
The COMET Program

Subjective Analysis

•Subjective Analysis is a representation of meteorological observation on charts and maps.

Goals:

•(1) Charts were used to deduce atmospheric flow based on the governing laws.

•(2) Diagnostic charts of past/present atmospheric states would help with the prognosis of future states.

• Synoptic (hand drawn) charts, adapted by Admiral Fitzroy in 1860 (British Meteorological Department) as a time series resulted in weather development and movement.

•Synoptic charts were used extensively during between 1860 -1920, but forecast skill was not improving.

•Bjerknes conceived weather prediction as an initial value problem and Richardson made the first weather calculation in 1922.

•Bjerknes and with Bergen developed the Norwegian Frontal Model with upper air network observations for initialization, resulting in improved predictability of storms in Norway.

•The era of model forecasting and data assimilation begins.

Objective Analysis

- An automated procedure to estimate the atmospheric dependent variables on a regular 2D and 3D grid using data available from the irregularly spaced observation network.

- Panofsky - 1949 made the first objective analysis with a polynomial fit of observation data in a small analysis domain. Observations were weighted based on presumed accuracies - an ad hoc approach.

- Gilchrist and Cressman - 1954 advanced this for a single grid point and determined what is commonly known as the *region of influence*. Automated quality control by comparison to the previous numerical forecast known as background field, first guess, prior estimate.

- Numerical Weather Prediction (NWP) was assumed to have perfect prognostics, then no corrections are made.

- Model Output Statistics (MOS) is the use of NWP output to forecast the next day.

- MOS became operational in 1972 using the technique of Glahn and Lowry.

NWP and Data Assimilation

•Numerical Weather Predictions (NWP) are assumed to make a good background for the analysis since it came from a good model, and retains information from older observations.

•Assimilated data into NWPs provide information that would otherwise be unavailable between observations, based in part on the inclusion and movement of information from previous observations.

•The forecast model fields contain numerical consistencies required for the model to make the next good forecast.

•Large discrepancies between observations and the short-term forecasts can be used to determine if the data are suspect since the forecast is assumed to be good (this is possible even if the observations are not rejected outright).

•A good forecast usually leads to a good analysis, resulting in another good forecast. However, a bad forecast may lead to a bad analysis, resulting in another bad forecast, until enough data are available to force the model to change.

The formal approach for data assimilation analysis is to solve for the minimized subspace

$$(X_a - X_b) = d_x$$

Where,

- d_x is the subspace
- Xa is the forecast
- Xb is the background (initial) state.

NWP: Prediction of future atmospheric states

$$dU/dt = U(\mathcal{S}_{x,y,z}, t+1) - U(\mathcal{S}_{x,y,z}, t) = w \times f[U(\mathcal{S}_{x,y,z},t) - U_{obs}]$$

$$dT/dt = T(\mathcal{S}_{x,y,z}, t+1) - T(\mathcal{S}_{x,y,z}, t) = w \times f[U(\mathcal{S}_{x,y,z},t) - T_{obs}]$$

$$dq/dt = q(\mathcal{S}_{x,y,z}, t+1) - q(\mathcal{S}_{x,y,z}, t) = w \times f[U(\mathcal{S}_{x,y,z},t) - q_{obs}]$$

U - wind direction and magnitude

T - atmospheric temperature

Q - atmospheric water vapor

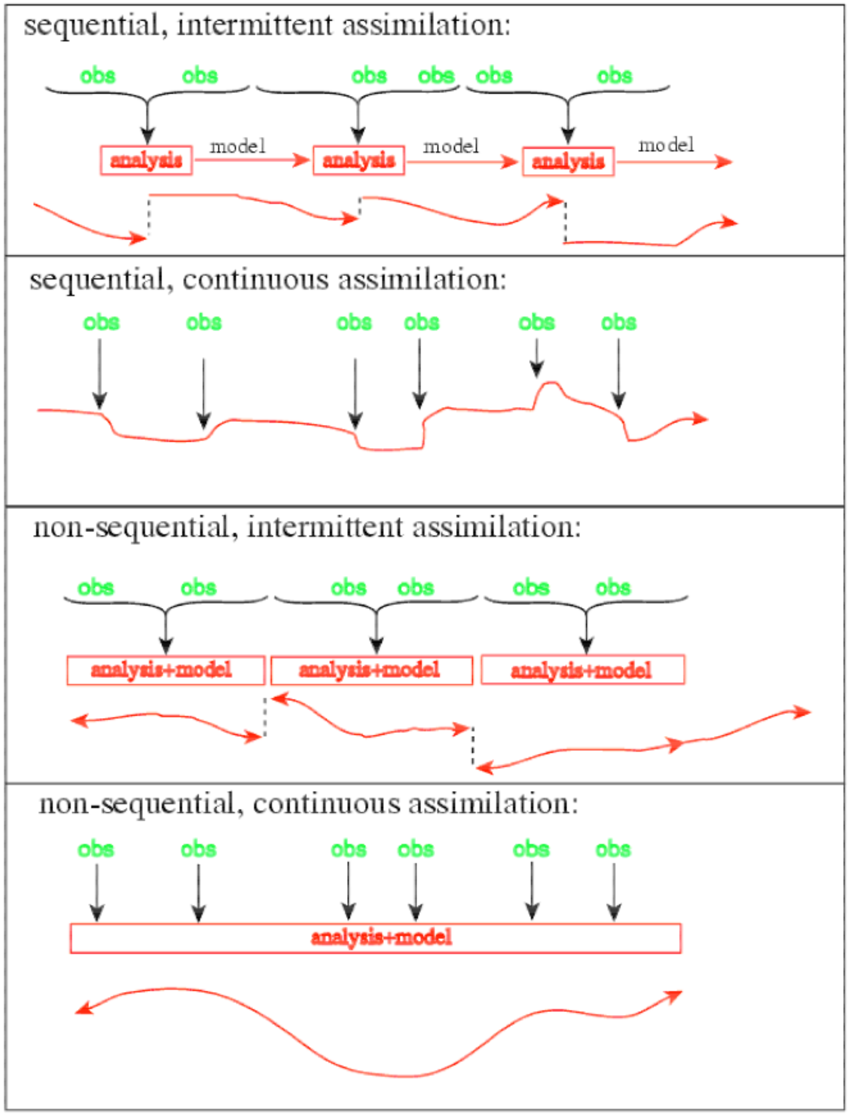
$\mathcal{S}_{x,y,z}$ - space dimensions

t - time dimension

w - assimilation weighting function

f - interpolation function

Four Types of Data Assimilation Strategies



Sequential: uses only past obs (t_{i-1})

real-time analysis

Non-Sequential: Obs from $t_i + t_{i+1}$ are used at time t in the model

retrospective analysis

Optimal Interpolation

observation

observation operator

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - H(\mathbf{x}_b))$$

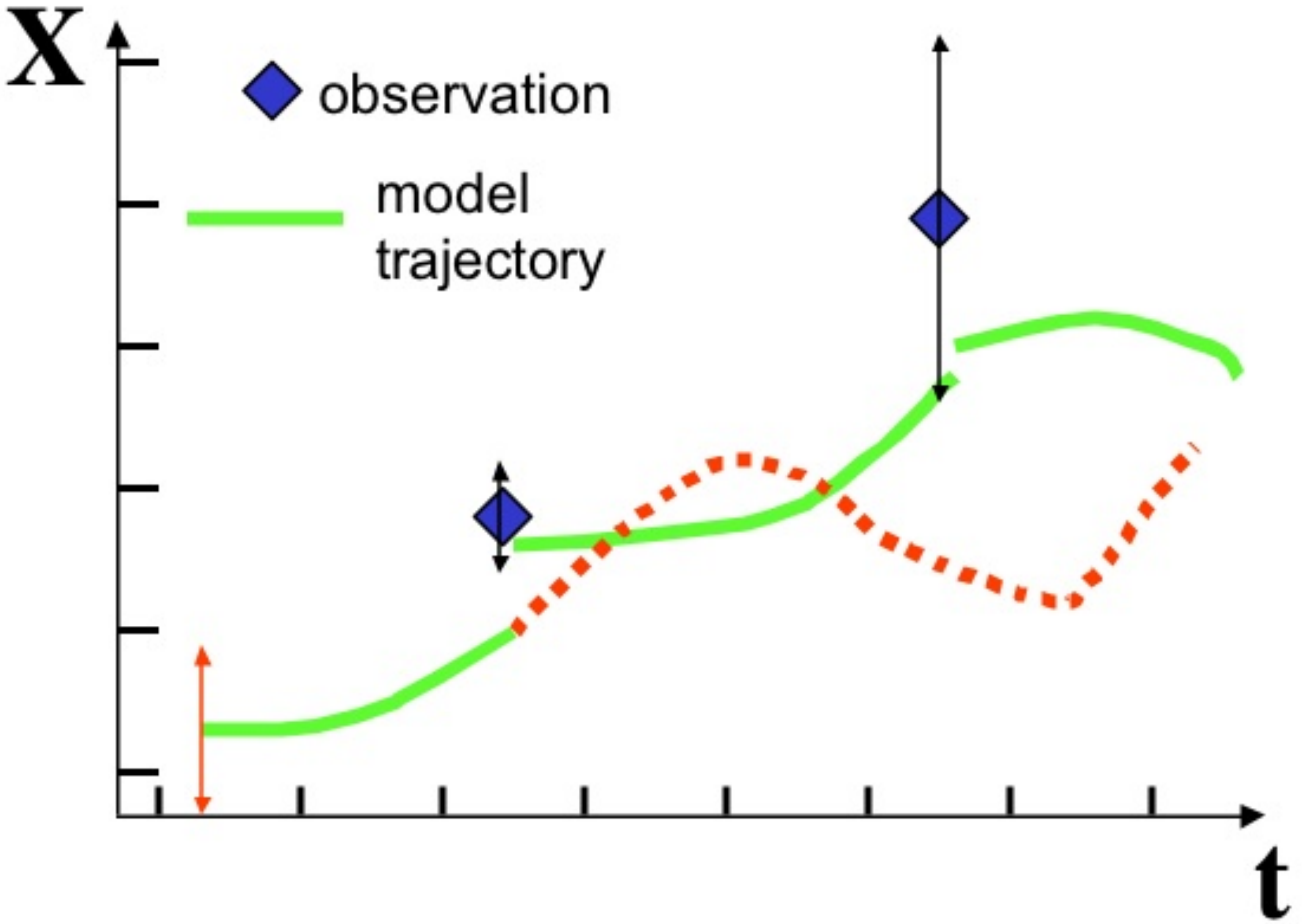
"analysis"

"background" (forecast)

- linearity $H \rightarrow \mathrm{H}$
- matrix inverse
- limited area

$$\boxed{\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{O})^{-1}}$$

DARC



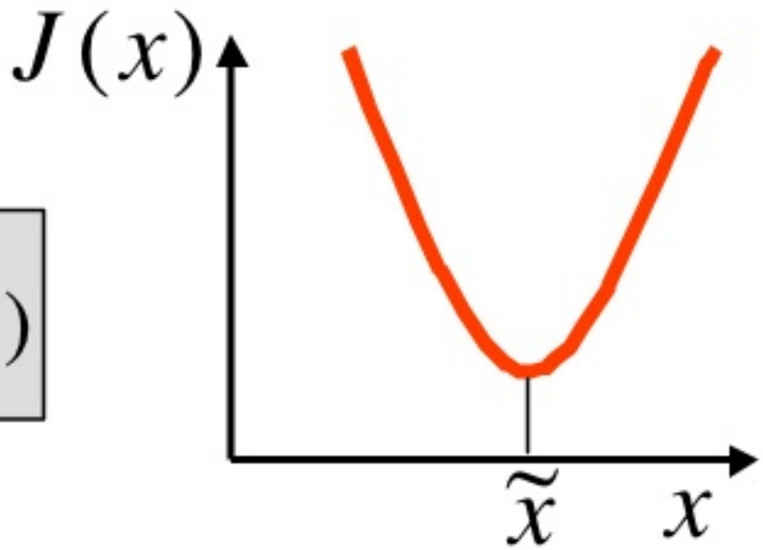
DARC

Variational Method

Vary X to minimize J(x)

$$J(x) = \frac{(x - x_1)^2}{\sigma_1^2} + \frac{(x - x_2)^2}{\sigma_2^2}$$

$\tilde{x}$ at $\min J(x)$



DARC

Variational Data Assimilation

$$J(\mathbf{x}) =$$

$$(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) +$$

$$(\mathbf{y} - H(\mathbf{x}))^T \mathbf{O}^{-1} (\mathbf{y} - H(\mathbf{x}))$$

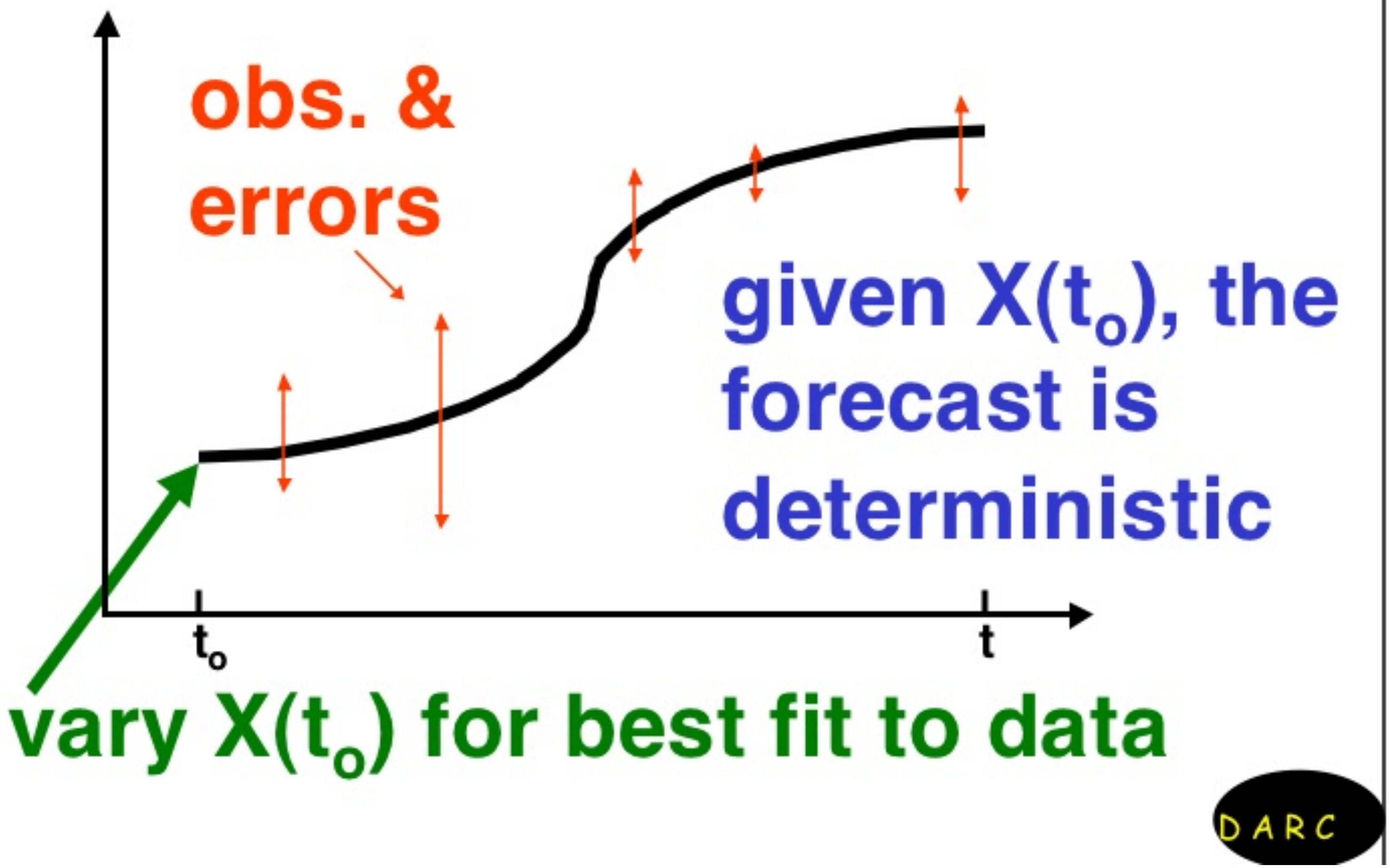
nonlinear operator

assimilate y directly

global analysis

DARC

4D Variational Data Assimilation



Variational Methods

Classical setting: finding the extremum of an integral involving a function and its derivatives

Application: finding the trajectory of a particle under external field.

The key idea here is that the problem of interest is formulated as an optimization problem

Variational surface analysis

Mahfouf (1991), Callies et al. (1998), Rhodin et al. (1999), Bouyssel et al. (2000), Hess (2001), Balsamo et al. (2003)

$$J(\mathbf{x}) = J^{\mathrm{b}}(\mathbf{x}) + J^{\mathrm{o}}(\mathbf{x})$$

- Formalism:

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

x is the control variables vector
y is the observation vector
H is the observation operator

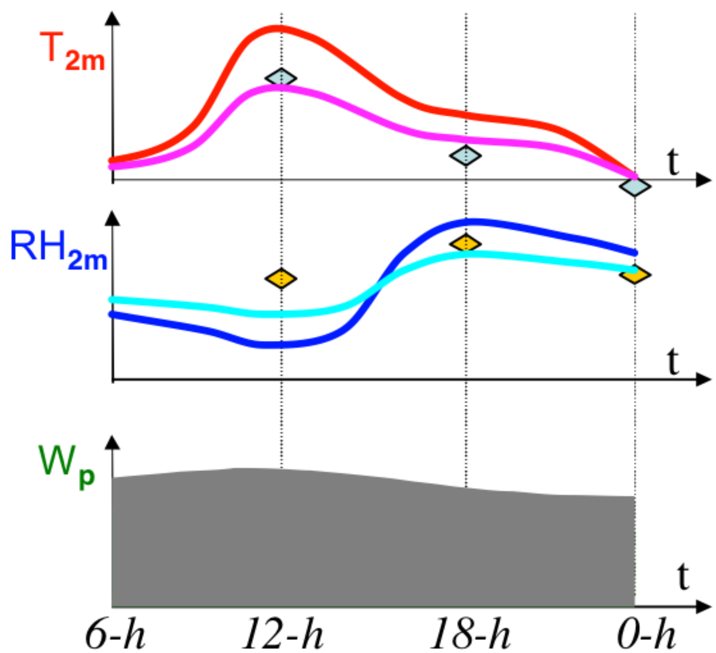
The analysis is obtained by the minimization of the cost function $J(\mathbf{x})$

B is the background error covariance matrix

R is the observation error covariance matrix

- Advantages: Easier assim. asynop. obs. Extension on longer assim. Window (24-h)

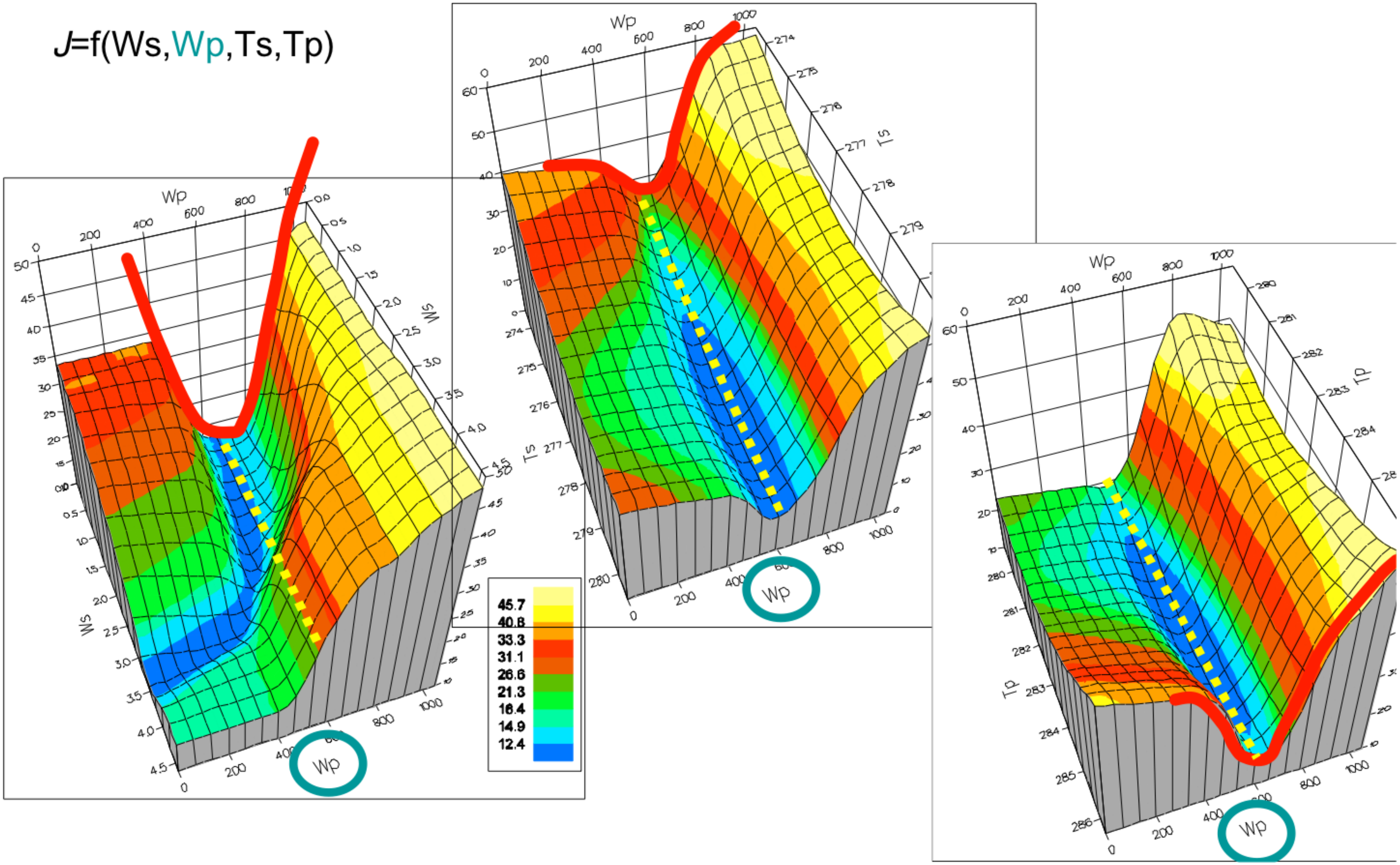
Continuous analysis



The shape of the cost function J for ISBA

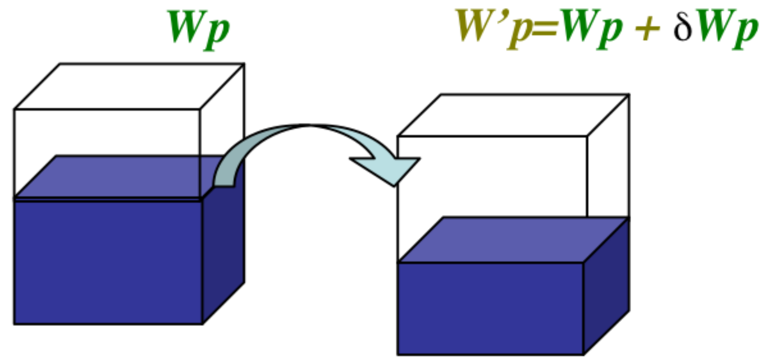
(Bouyssel et al. 1999, Bouyssel 2001)

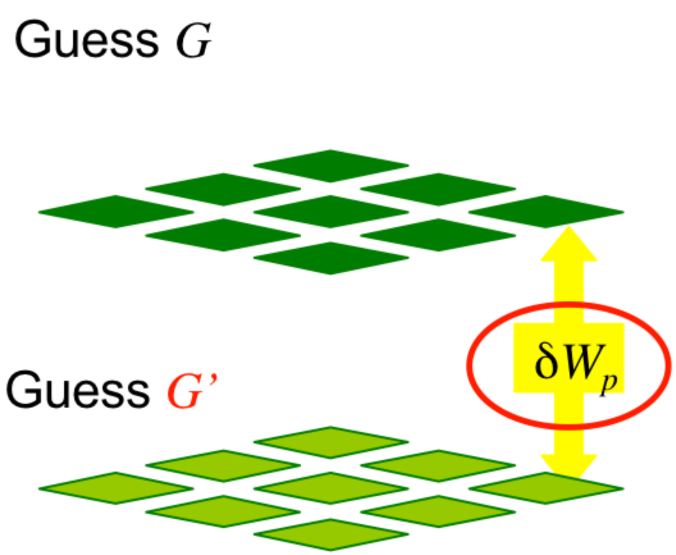
J=f(Ws,Wp,Ts,Tp)



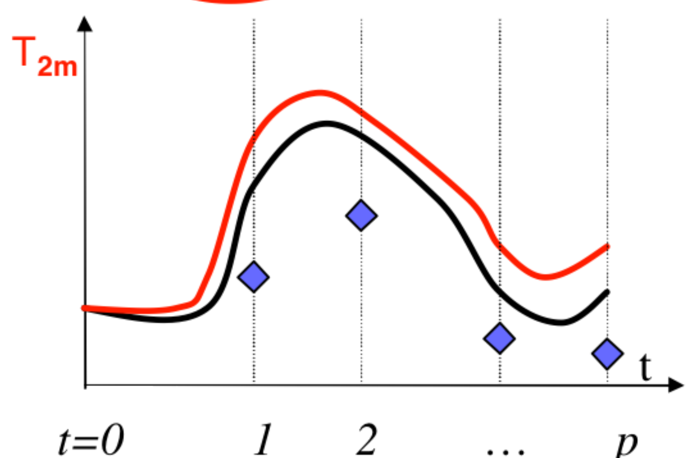
3D study of 2D-VAR: method

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





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



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

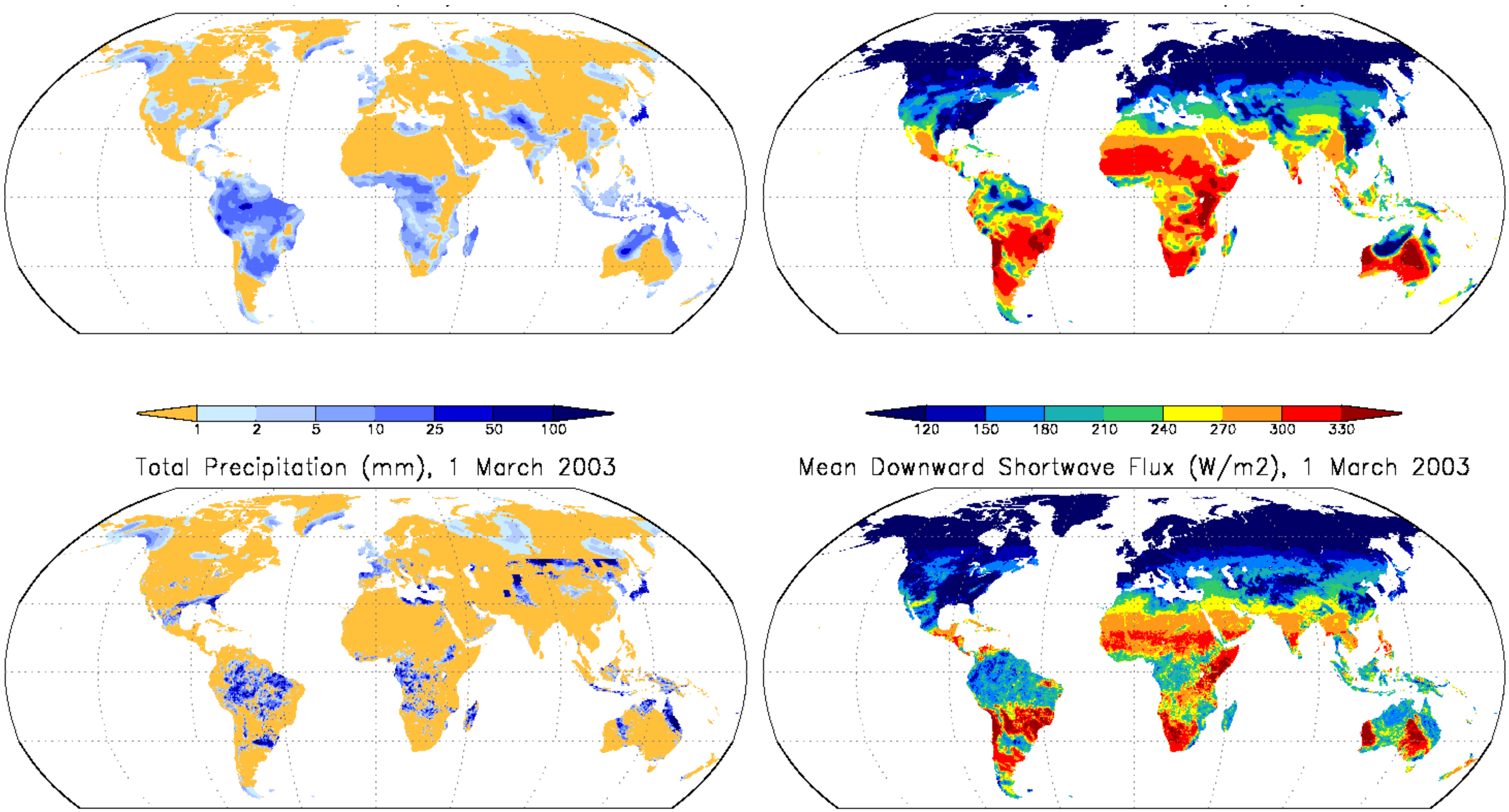
$$\boldsymbol{y} - H(\boldsymbol{x}_b) = \left(\Delta T_{2m}^{(1)}, \Delta RH_{2m}^{(1)}, \dots, \Delta T_{2m}^{(p)}, \Delta RH_{2m}^{(p)}\right)$$

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

PART III

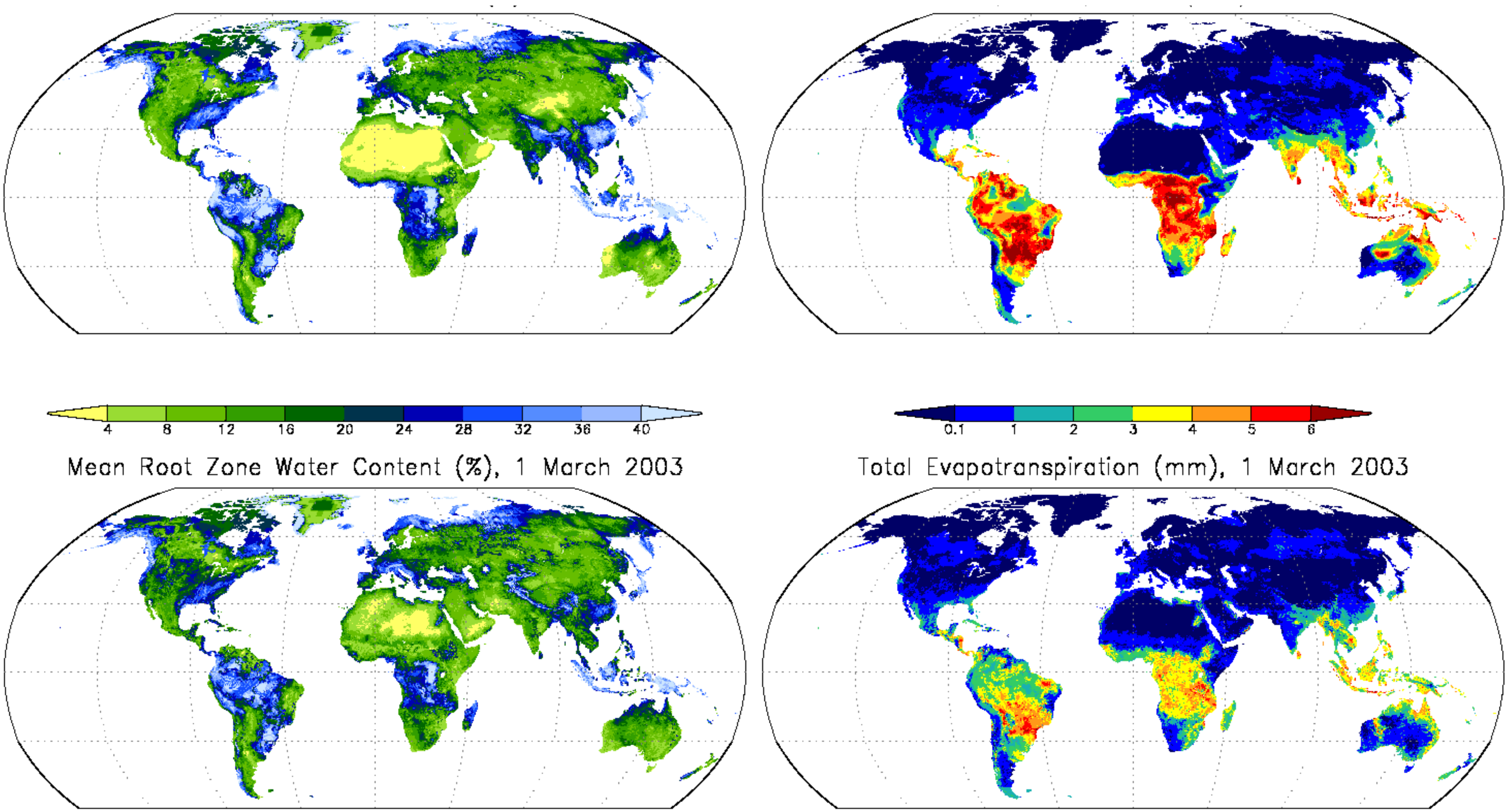
Applications

GLDAS Forcing



atmospheric data assimilation output (top)
vs.
observation based (bottom)

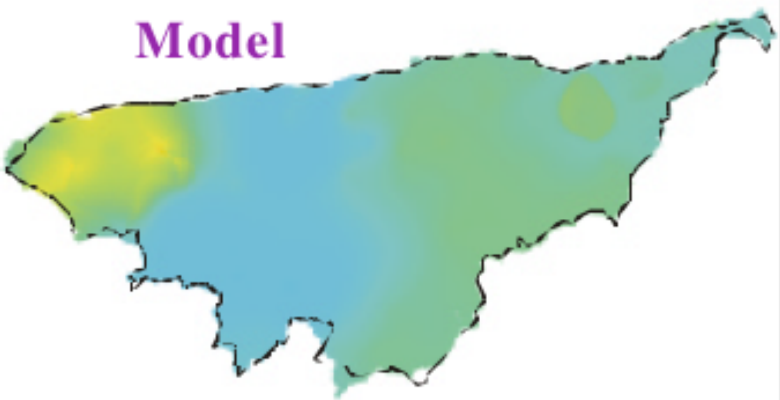
GLDAS Output



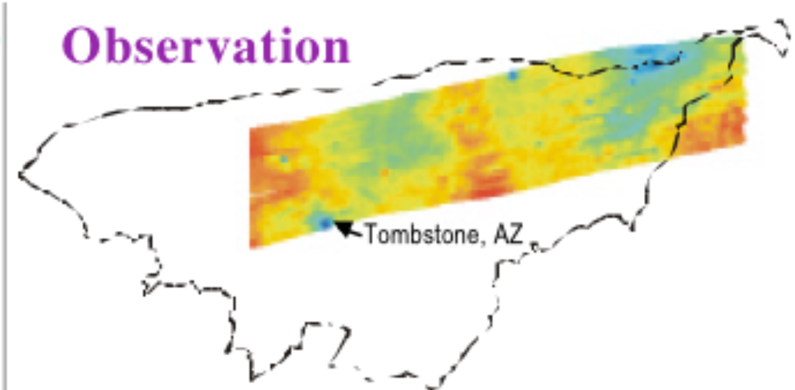
atmospheric data assimilation output forced (top)
vs.
observation forced (bottom)

Regional Scale: *Walnut Gulch (Monsoon 90)*

Model

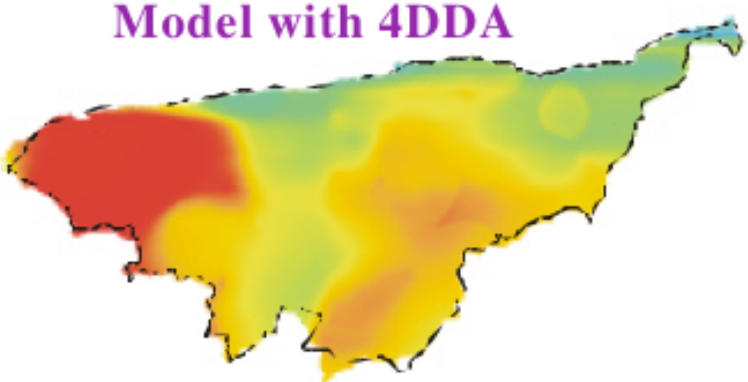


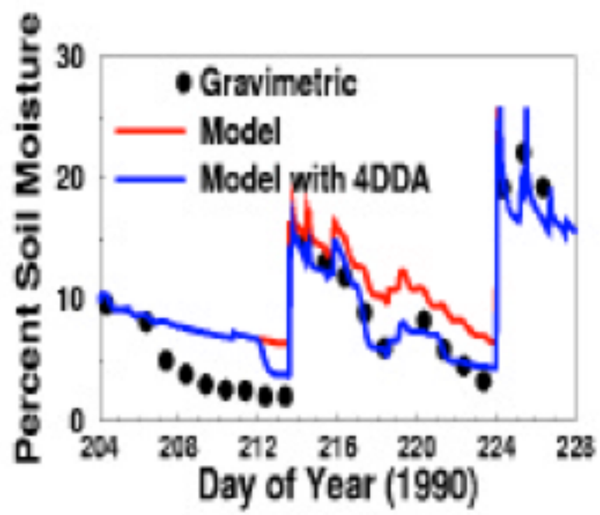
Observation



0%  20%

Model with 4DDA

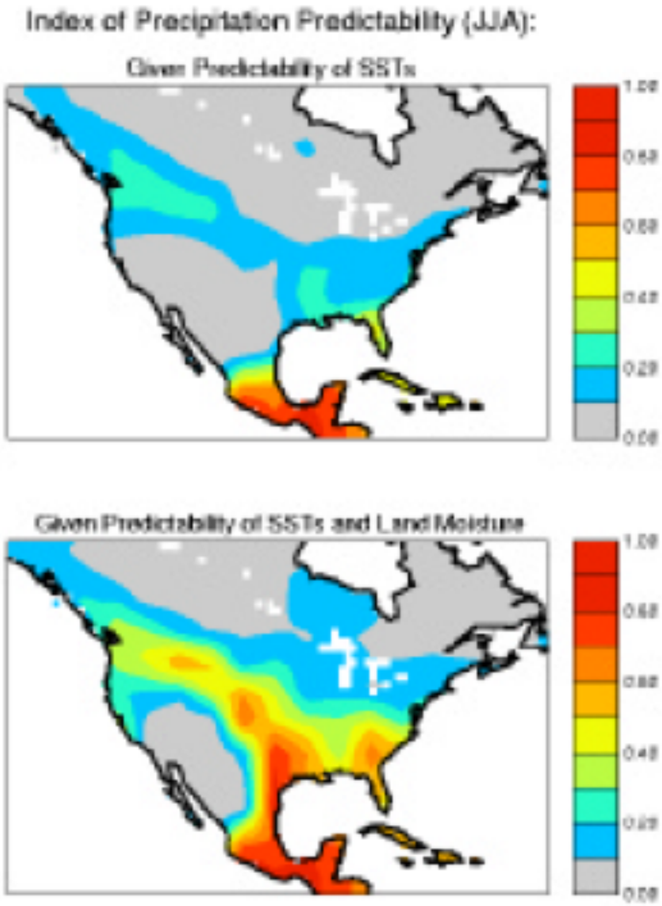


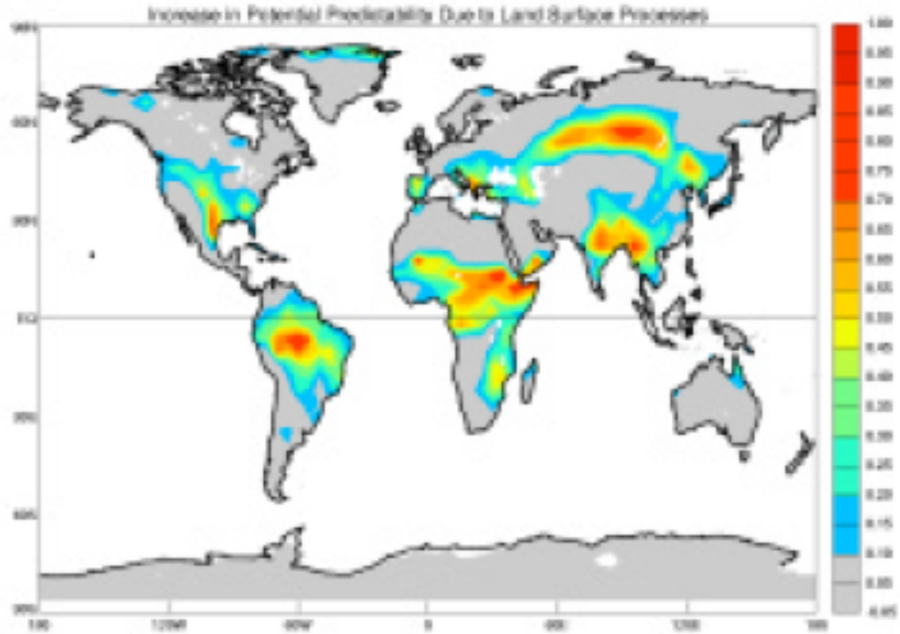


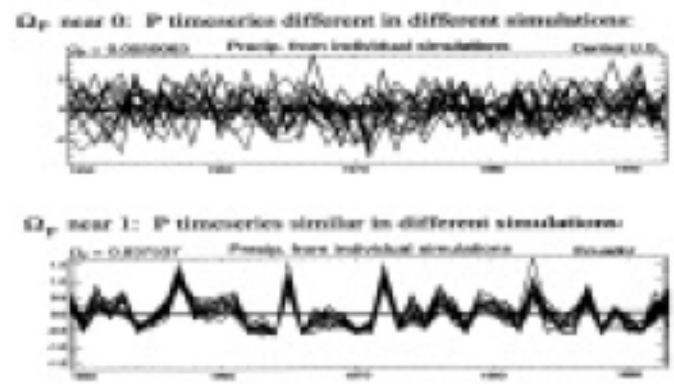
Houser et al., 1998

Land Initialization: Motivation

- Knowledge of soil moisture has a greater impact on the predictability of summertime precipitation over land at mid-latitudes than Sea Surface Temperature (SST).







Importance of Soil Moisture in mesoscale NWP

- Mean soil moisture largely determines the ratio between sensible (H) and latent (LE) heat fluxes at the surface (Bowen ratio)
- It affects the atmospheric boundary layer evolution
- Mean soil moisture has a long time-scale memory (several weeks)

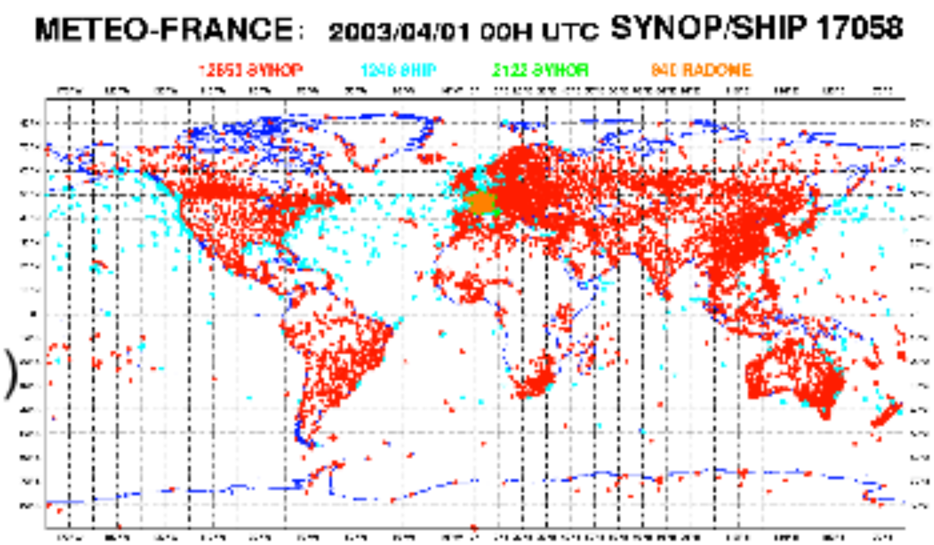
Problems:

- No regular observations of soil moisture on extensive domain
- High spatial variability of surface parameters

Idea:

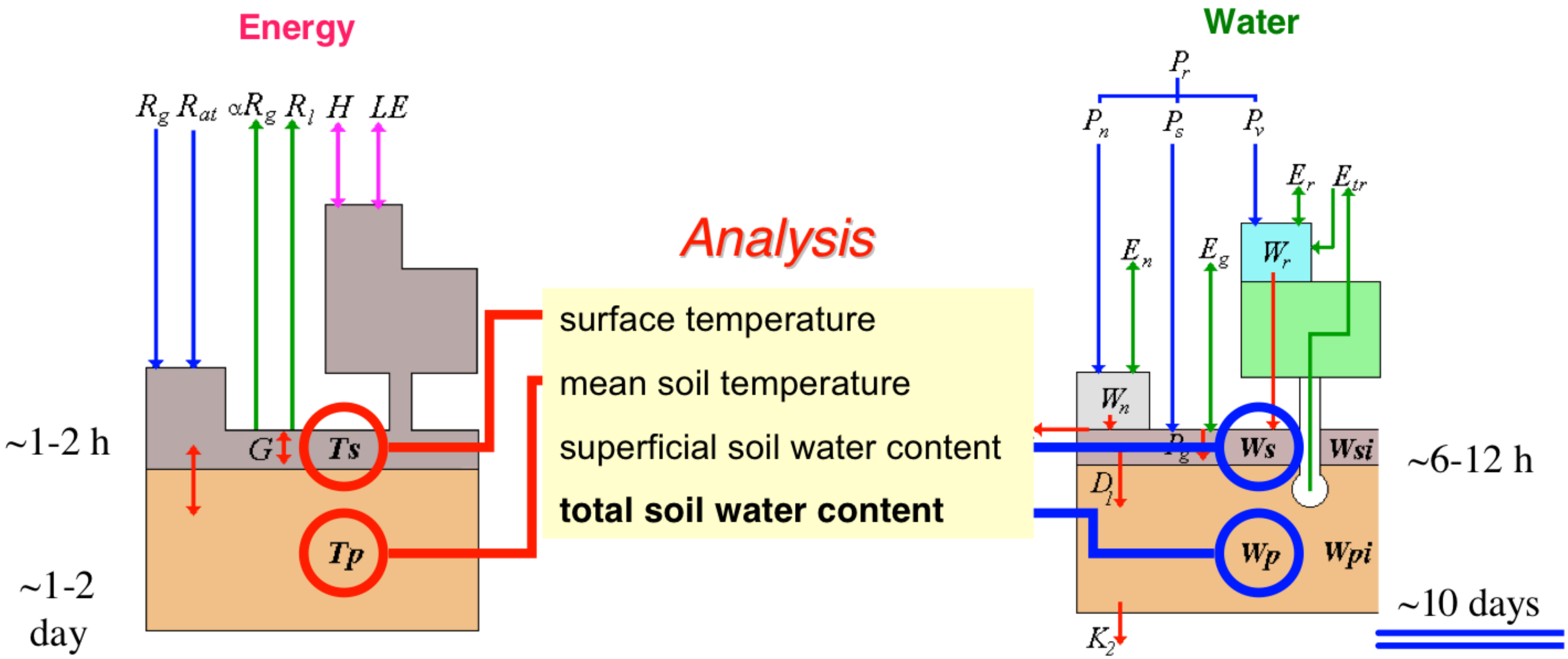
- Use indirect observations (prec., 2m obs., satellite obs.)

- In the present study: T_{2m}, RH_{2m} (Coiffier et al. 1987, Mahfouf 1991).



Surface Parameterization scheme (ISBA)

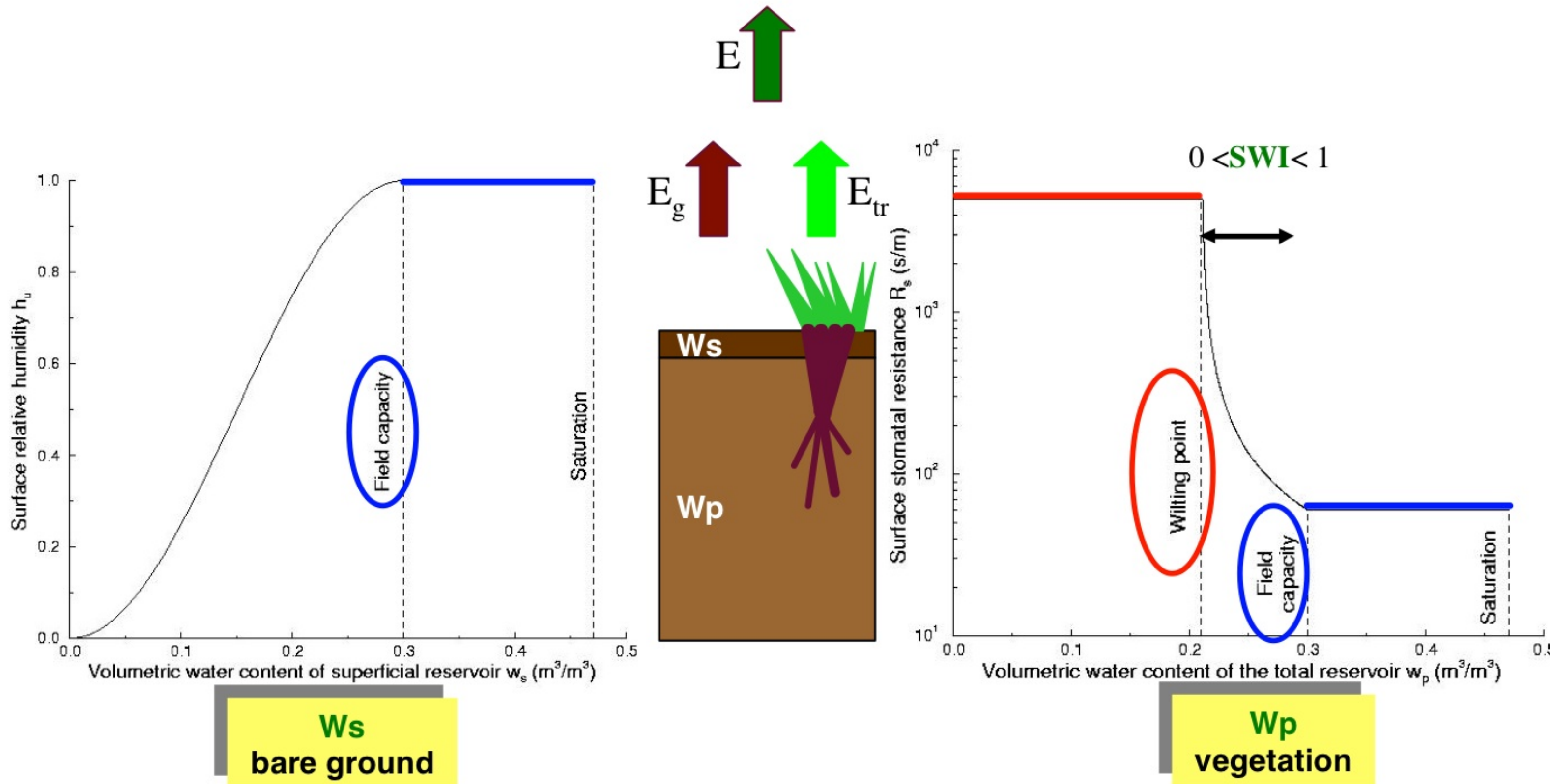
Operational version : Noilhan & Planton (1989), Noilhan & Mahfouf (1996), Bazile (1999), Giard & Bazile (2000)



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

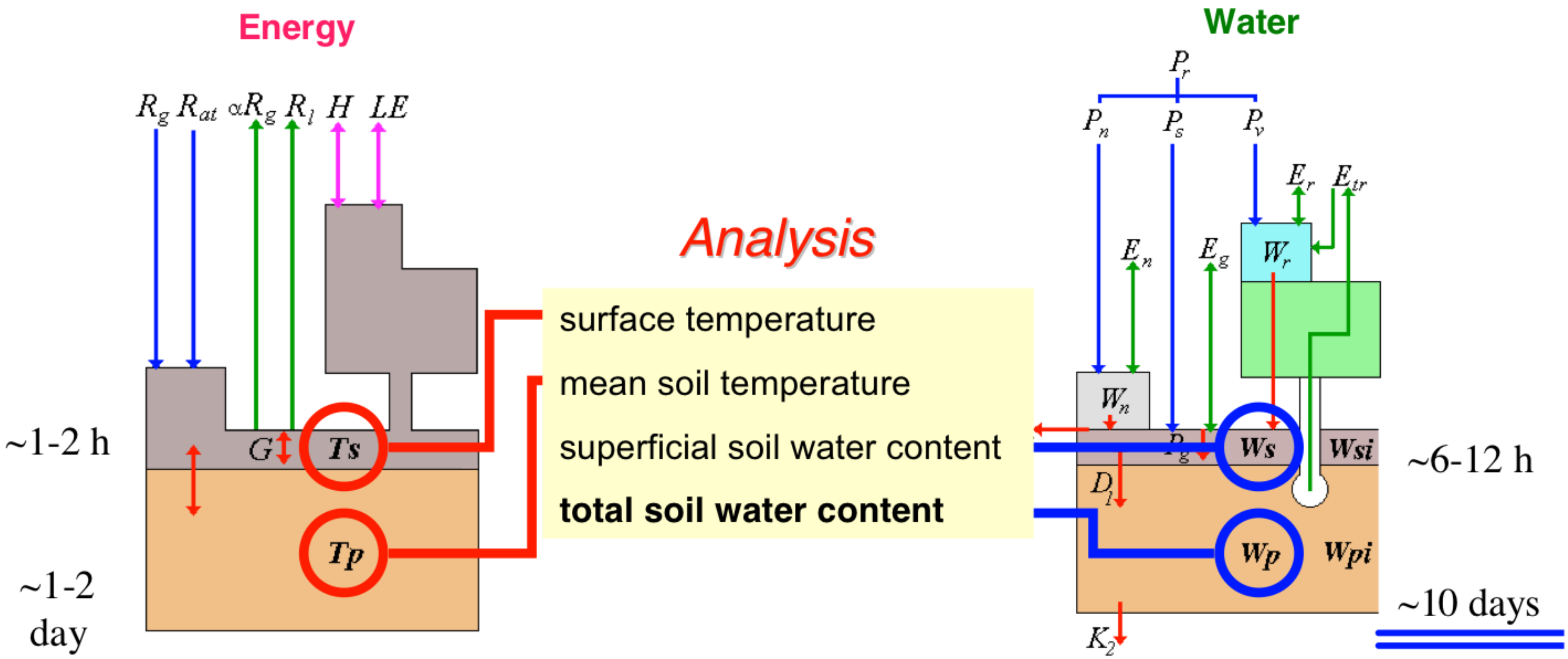
The link between soil moisture and atmosphere

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



Surface Parameterization scheme (ISBA)

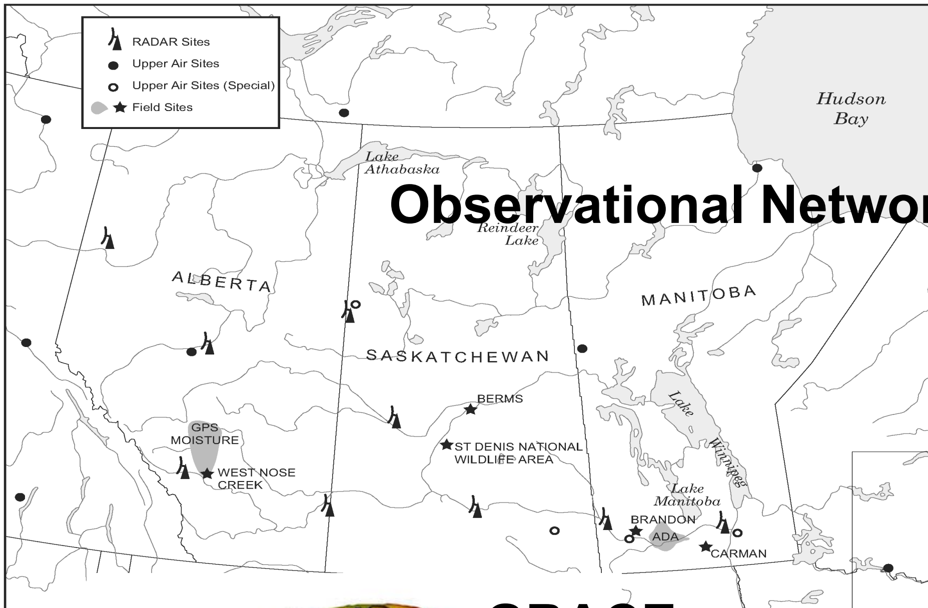
Operational version : Noilhan & Planton (1989), Noilhan & Mahfouf (1996), Bazile (1999), Giard & Bazile (2000)



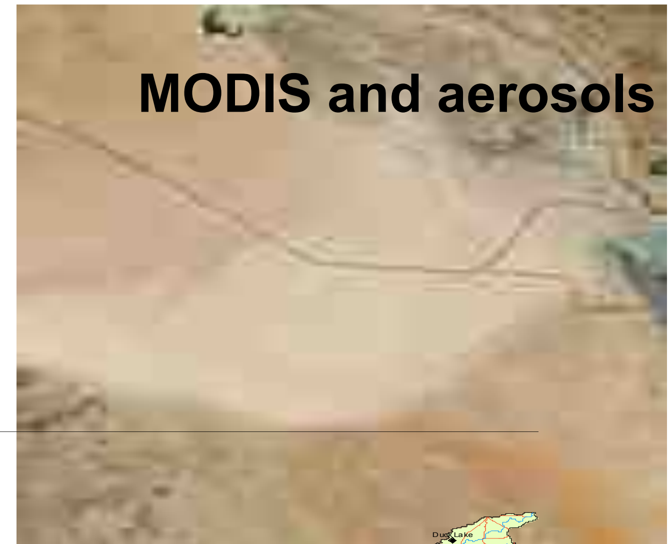
Research versions : interactive vegetation module (Calvet et al. 1998),
sub grid-scale runoff and sub-root layer (Boone et al 1999),
explicit 3-layers snow scheme (Boone & Etchevers 2001), tiling,
multi-layer soil scheme, urban scheme

QUANTIFY THE DROUGHT

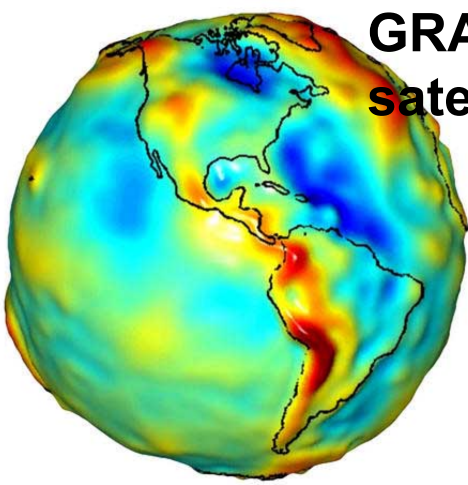
Observational Networks



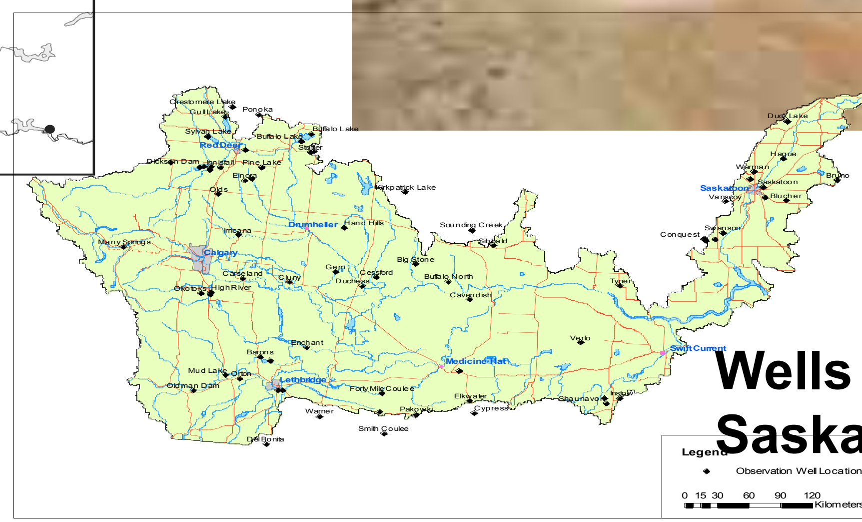
MODIS and aerosols



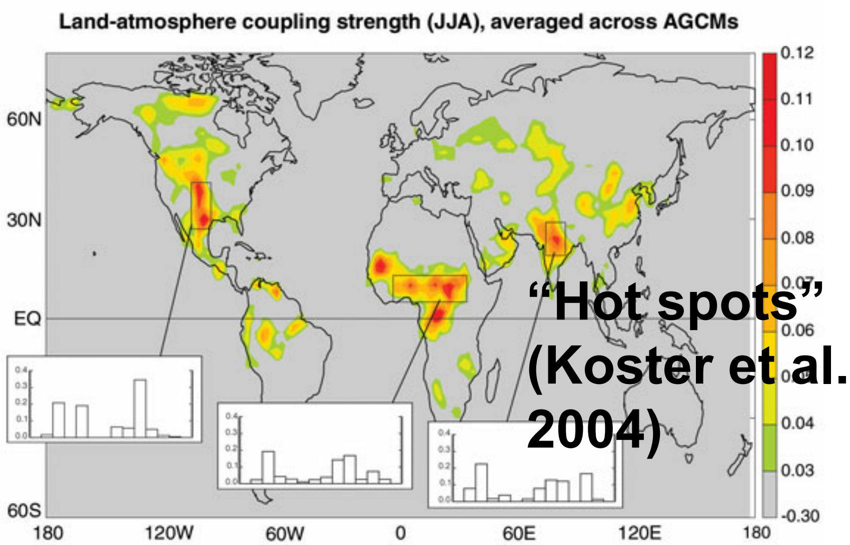
GRACE satellite



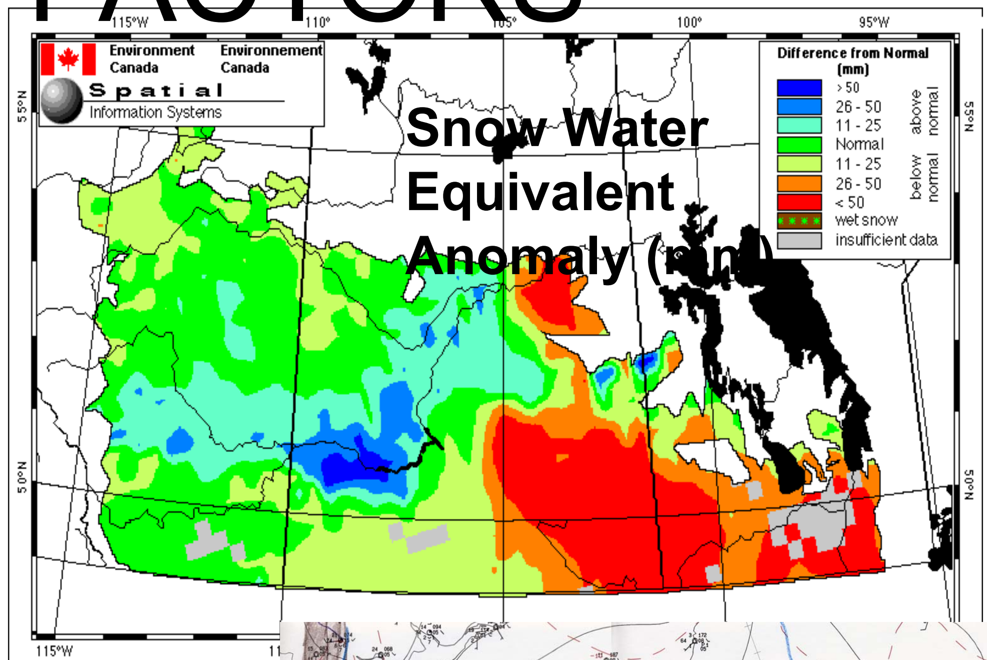
Wells in South Saskatchewan



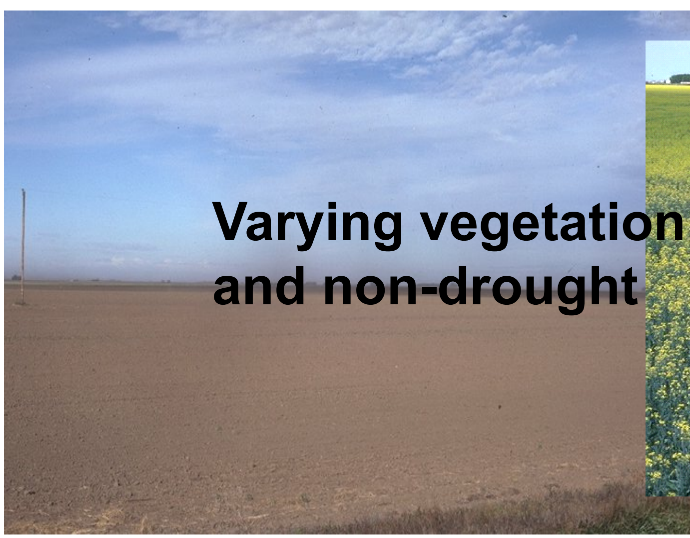
CRITICAL FACTORS



"Hot spots" (Koster et al. 2004)



Snow Water Equivalent Anomaly (mm)

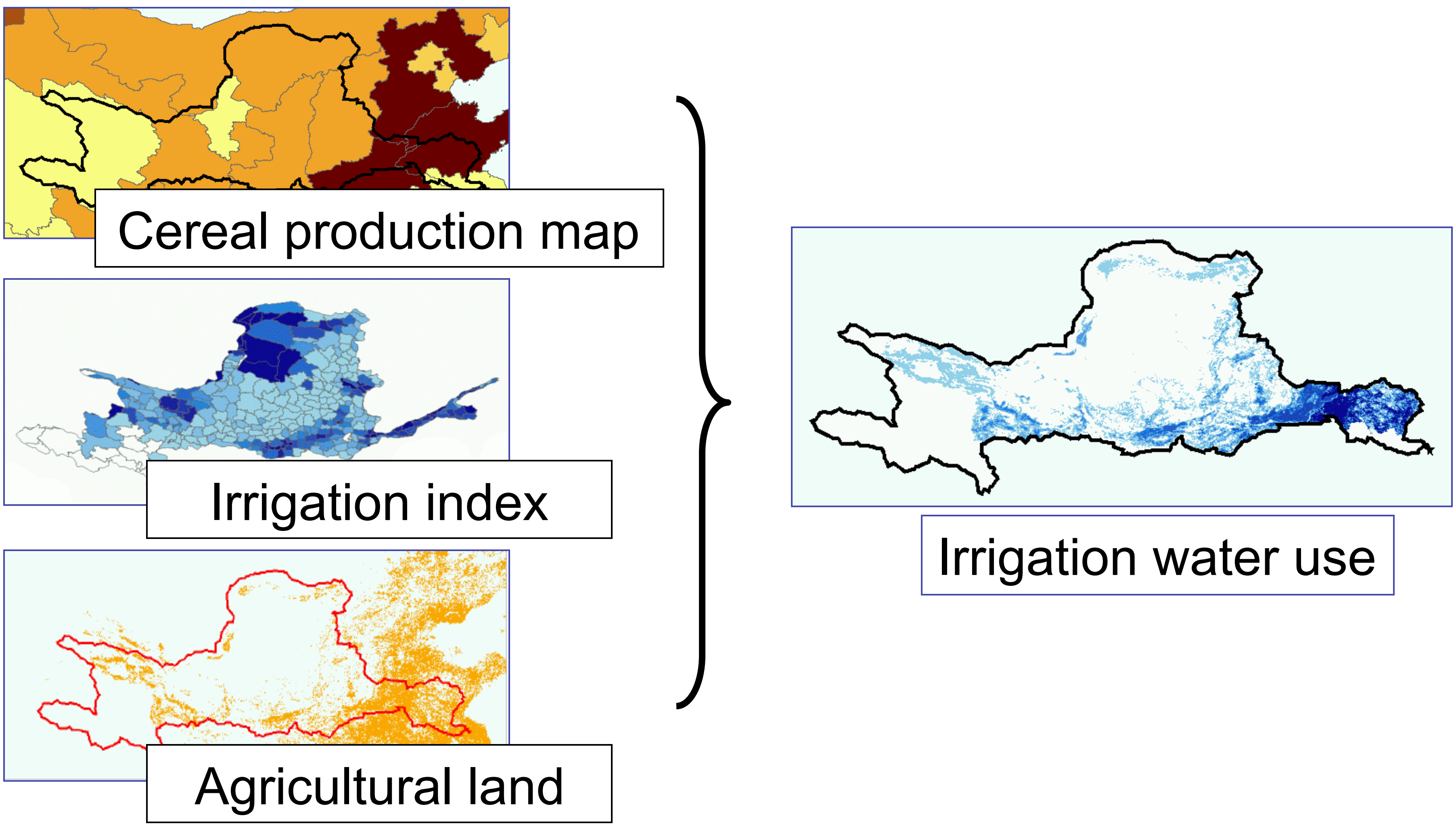



Varying vegetation between drought and non-drought

Individual storm events

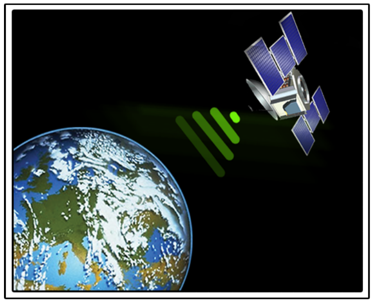
③ Agricultural water use

Estimation of distribution of irrigation demand



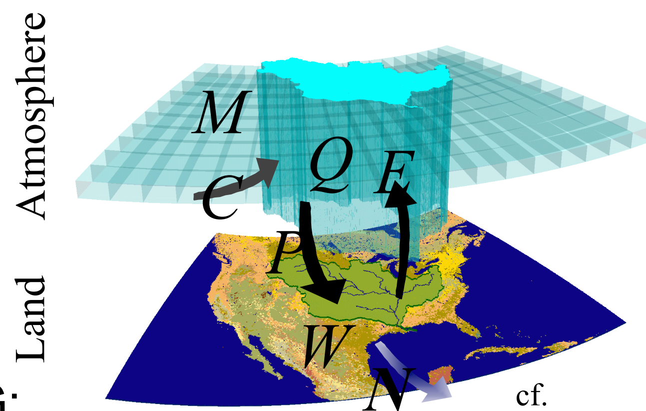
THE GLOBAL WATER CYCLE VISION: A NEW EPOCH OF WATER MANAGEMENT IN OUR LIFETIMES THAT IS BUILT ON FULL EXPLOITATION OF ENHANCED OBSERVATIONS.

OBSERVATIONS



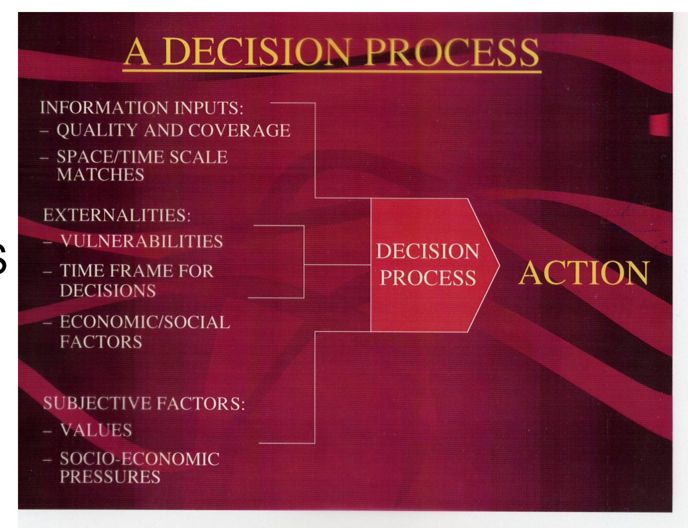
IMPROVED CAPABILITY TO ASSIMILATE AND PREDICT

WHAT WILL THIS NEW EPOCH LOOK LIKE IN TERMS OF SUPPORTING SYSTEMS?



cf. Sorooshian

INTEGRATED DECISION SUPPORT SYSTEMS



WHAT'S MISSING:

- A FULL RANGE OF PRODUCTS.
- ACCURATE MEASURMENTS AT REQUIRED RESOLUTIONS
- ADEQUATE ASSIMILATION AND PREDICTION MODELS.
- BETTER WAYS OF HANDLING UNCERTAINTY.
- FLEXIBLE DECISION SUPPORT SYSTEMS THAT EFFECTIVELY INCORPORATE WC INFORMATION.

THANK YOU