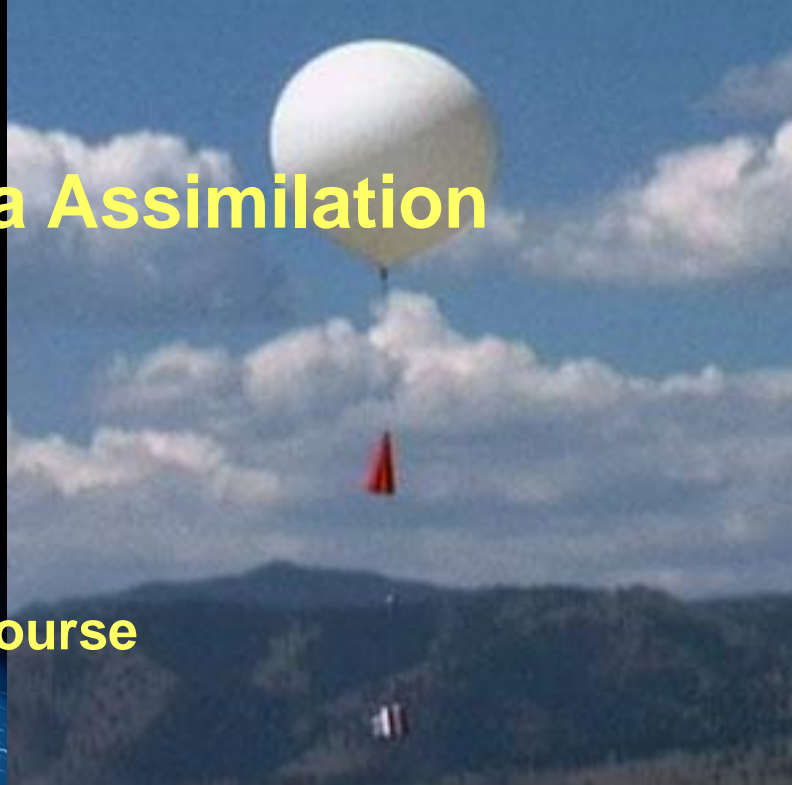# Parametrizations in Data Assimilation

ECMWF Training Course
8 March 2019

Philippe Lopez

Physical Processes Group, Research Department, ECMWF
(Room 002)

Parametrizations in Data Assimilation

Why do we need data assimilation?

- By construction, **numerical weather forecasts are imperfect**:

  ← **discrete** representation of the atmosphere in space and time (horizontal and vertical grids, spectral truncation, time step)

  ← **subgrid-scale processes** (e.g. turbulence, convective activity) **need to be parametrized** as functions of the resolved-scale variables.

  ← **errors in the initial conditions**.

- **Physical parametrizations** used in NWP models are constantly being improved:

  → more and more prognostic variables (cloud variables, precipitation, aerosols),
  → more and more processes accounted for (e.g. detailed microphysics).

- However, they remain **approximate representations of the true atmospheric behaviour**.

- Another way to improve forecasts is to **improve the initial state**.

- The goal of **data assimilation** is **to periodically constrain the initial conditions of the forecast using a set of accurate observations** that provide our best estimate of the local true atmospheric state.

General features of data assimilation

- **<u>Goal:</u>** to produce an accurate three-dimensional representation of the atmospheric state to initialize numerical weather prediction models.

- This is achieved by combining in an optimal statistical way all the information on the atmosphere, available over a selected time window (usually 6 or 12h):

  - ✓ **Observations** with their accuracies (error statistics),

  - ✓ **Short-range model forecast** (background) with associated error statistics,

  - ✓ **Atmospheric equilibria** (e.g. geostrophic balance),

  - ✓ **Physical laws** (e.g. perfect gas law, condensation, microphysics,…)

- The optimal atmospheric state found is called the **analysis**.

Which observations are assimilated?

Operationally assimilated since many years ago:

* **Surface measurements** (SYNOP, SHIPS, DRIBU,…),
* **Vertical soundings** (TEMP, PILOT, AIREP, wind profilers,…),
* **Geostationary satellites** (METEOSAT, GOES,…)
  Polar orbiting satellites (NOAA, SSM/I, AIRS, AQUA, QuikSCAT,…):
  - radiances (infrared & passive microwave in clear-sky conditions),
  - products (motion vectors, total column water vapour, ozone,…).

More recently:

* **Satellite radiances/retrievals in cloudy and rainy regions** (SSM/I, TMI,…),
* Precipitation measurements from **ground-based radars** and **rain gauges.**

Still experimental:

* Satellite cloud/precipitation radar reflectivities/products (TRMM, CloudSat),
* Lidar backscattering/products (wind vectors, water vapour) (CALIPSO),
* GPS water vapour retrievals,
* Satellite measurements of aerosols, trace gases,....
* Lightning data (TRMM-LIS, and soon GOES-R, FY-4A and MTG-LI).

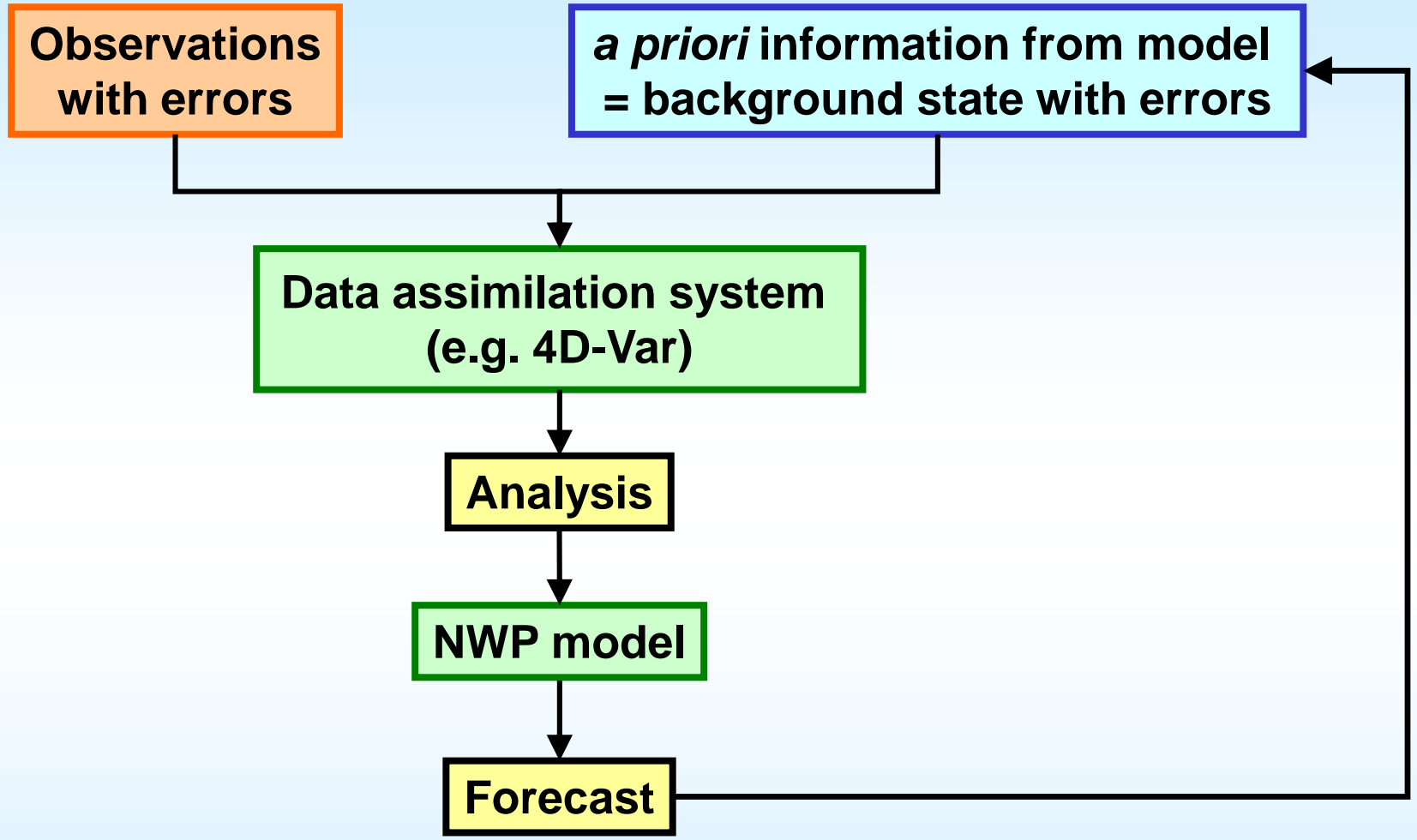# Why physical parametrizations in data assimilation?

- In current operational systems, most used observations are directly or indirectly related to **temperature, wind, surface pressure** and **humidity** outside cloudy and precipitation areas (~ 20 million observations assimilated in ECMWF 4D-Var every 12 hours).

- **Physical parametrizations are used during the assimilation to link the model's prognostic variables** (typically: T, u, v, q_v and P_s) **to more "exotic" observed quantities** (e.g. precipitation rates, radiances, radar reflectivities,…).

- Observations related to **clouds** and **precipitation** are starting to be routinely assimilated,

→ but how to convert such information into proper corrections of the model's initial state (prognostic variables T, u, v, q_v and P_s) is not so straightforward.

For instance, problems in the assimilation can arise from the discontinuous or non-linear nature of moist processes.

Improvements are still needed…

- **More observations are needed to improve the analysis and forecast of:**

  - Mesoscale phenomena (convection, frontal regions),
  - Vertical and horizontal distribution of clouds and precipitation,
  - Planetary boundary layer processes (stratocumulus/cumulus clouds),
  - Surface processes (soil moisture, snow on the ground, sea ice),
  - The tropical circulation (monsoons, squall lines, tropical cyclones).

- **Recent developments and improvements have been achieved in:**

  - **Data assimilation techniques** (OI → 3D-Var → 4D-Var → Ensemble DA),

  - **Physical parametrizations** in NWP models (prognostic schemes, detailed convection and large-scale condensation processes),

  - **Radiative transfer models** (infrared and microwave frequencies),

  - **Horizontal and vertical resolutions** of NWP models (currently at ECMWF: 9 km globally, 137 levels),

ECMWF

To summarize…

Observations with errors

a priori information from model = background state with errors

Data assimilation system (e.g. 4D-Var)

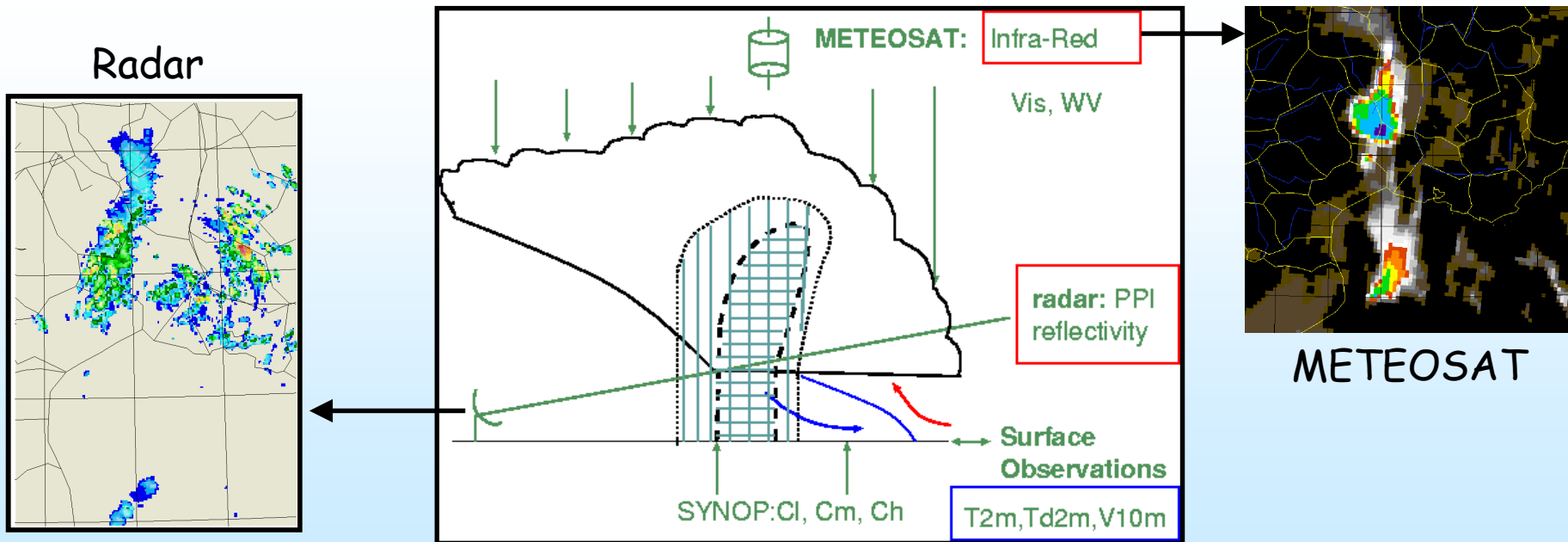**Analysis**

NWP model

Forecast

Physical parametrizations are needed in data assimilation:

- to link the model variables to the observed quantities,

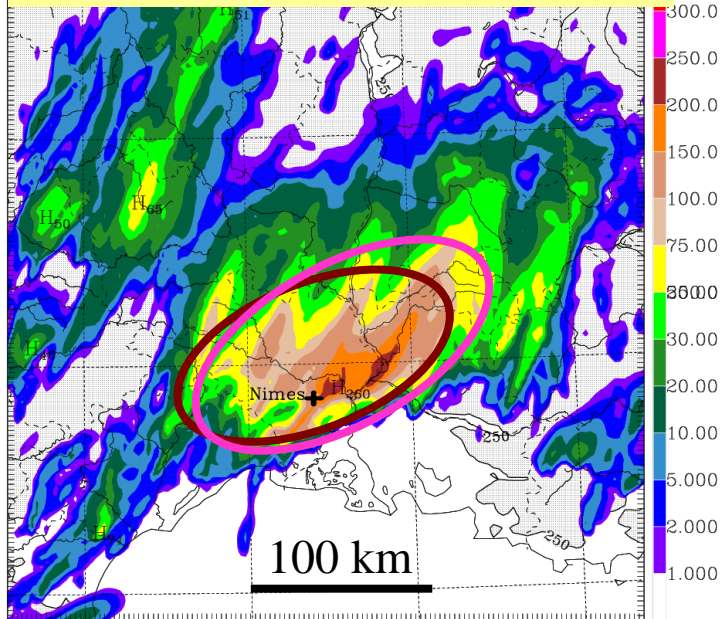- to evolve the model state in time during the assimilation (e.g. 4D-Var).

Empirical initialization

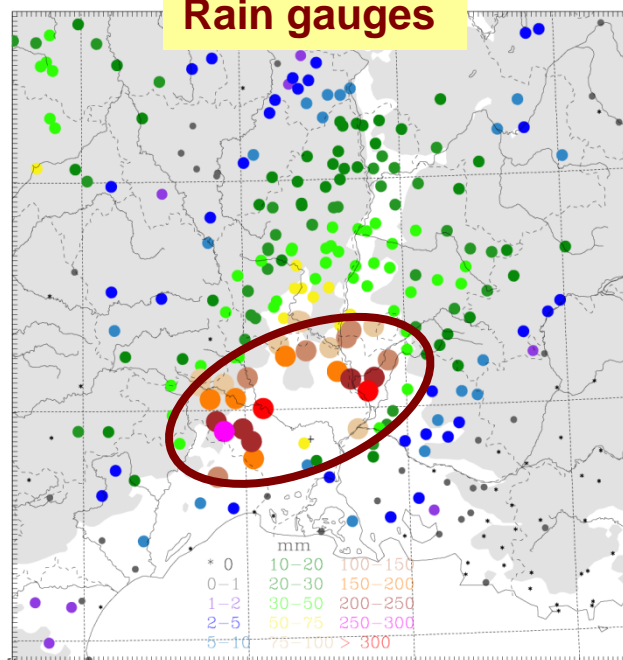## Example from Ducrocq *et al.* (2000), Météo-France:

- Using the mesoscale research model Méso-NH (prognostic clouds and precipitation).

- Particular focus on strong convective events.

- **Method:** Before running the forecast:

  1) A mesoscale surface analysis is performed (esp. to identify convective cold pools)

  2) the model humidity, cloud and precipitation fields are **empirically adjusted** to match ground-based precipitation radar observations and METEOSAT infrared brightness temperatures.
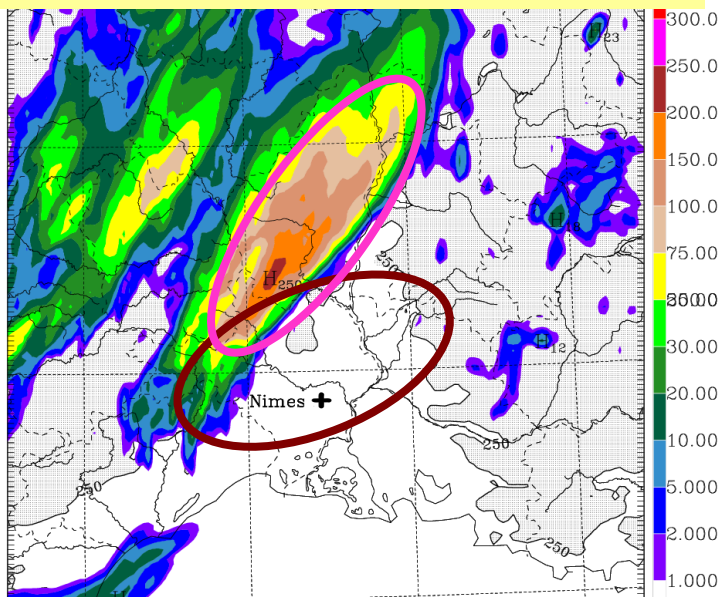

Radar

METEOSAT

12h FC from modified analysis

100 km

Rain gauges

mm
* 0
0–1
1–2
2–5
5–10
10–20
20–30
30–50
50–75
75–100
100–150
150–200
200–250
250–300
> 300

Ducrocq et al. (2004)

2.5-km resolution model Méso-NH

12h FC from operational analysis

Nîmes radar
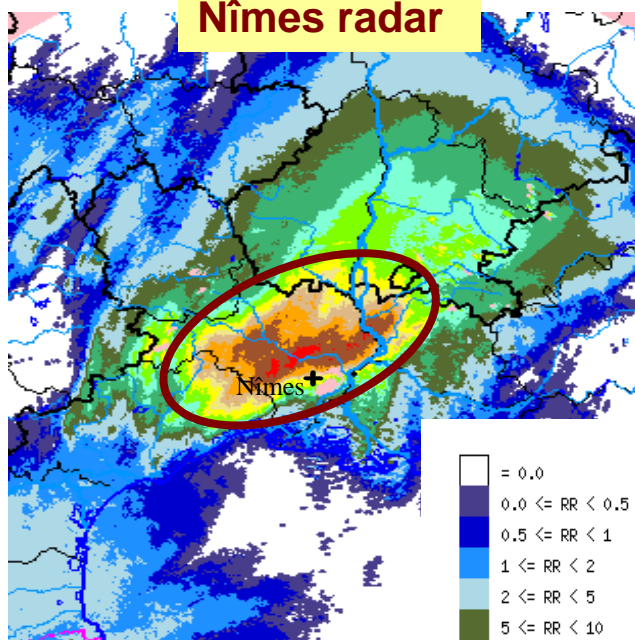
Flash flood over South of France (8-9 Sept 2002)

= 0.0
0.0 <= RR < 0.5
0.5 <= RR < 1
1 <= RR < 2
2 <= RR < 5
5 <= RR < 10
10 <= RR < 20
20 <= RR < 30
30 <= RR < 50
50 <= RR < 75
75 <= RR < 100
100 <= RR < 150
150 <= RR < 200
200 <= RR < 300
>= 300
Pas de donnees

Lame d'eau exprimee en mm.
100 km

12h accumulated precipitation: 8 Sept 12 UTC → 9 Sept 2002 00 UTC

A very simple example of variational data assimilation

- Short-range forecast (**background**) of 2m temperature from model: x_b with error σ_b.
- Simultaneous **observation** of 2m temperature: y_o with error σ_o.

The best estimate of 2m temperature (x_a=**analysis**) minimizes the following **cost function**:

$$J(x) = \underbrace{\frac{1}{2}\left(\frac{x - x_b}{\sigma_b}\right)^2}_{J_b} + \underbrace{\frac{1}{2}\left(\frac{x - y_o}{\sigma_o}\right)^2}_{J_o}$$

= quadratic distance to background and obs
(weighted by their errors)

In other words:

$$\left(\frac{dJ}{dx}\right)_{x=x_a} = \frac{(x_a - x_b)}{{\sigma_b}^2} + \frac{(x_a - y_o)}{{\sigma_o}^2} = 0 \quad\Leftrightarrow\quad x_a = x_b + \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2}(y_o - x_b)$$

And the analysis error, σ_a, verifies:

$$\frac{1}{\sigma_a^2} = \frac{1}{\sigma_b^2} + \frac{1}{\sigma_o^2} \quad\Rightarrow\quad \sigma_a^2 \leq \min(\sigma_b^2, \sigma_o^2)$$

The analysis is a linear combination of the model background and the observation weighted by their respective error statistics.

3D-Var assimilation

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

$\mathbf{B}$ is the background error covariance matrix, $\mathbf{R}$ is the observation error covariance matrix, H is the observation operator (used for converting model state vector $\mathbf{x} = (T, q_v, u, v)$ into observation space).

0D-Var

$$J = \boxed{\frac{1}{2}\left(\frac{x - x_b}{\sigma_b}\right)^2} + \boxed{\frac{1}{2}\left(\frac{x - y_o}{\sigma_o}\right)^2}$$

3D-Var

$$J = \boxed{\frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b)} + \boxed{\frac{1}{2}\left(H(\mathbf{x}) - \mathbf{y}_o\right)^T \mathbf{R}^{-1} \left(H(\mathbf{x}) - \mathbf{y}_o\right)}$$

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(H(\mathbf{x}) - \mathbf{y}_o)^T \mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y}_o)$$

$\mathbf{B}$ is the background error covariance matrix, $\mathbf{R}$ is the observation error covariance matrix, H is the observation operator (used for converting model state vector $\mathbf{x} = (T, q_v, u, v)$ into observation space).

The minimization of $\mathcal{J}$ can be performed if its gradient with respect to the atmospheric state $\mathbf{x}$ is known:

$$\nabla_{\mathbf{x}}\mathcal{J}(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \boxed{\mathbf{H}^T}\mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y}_o)$$

where $\mathbf{H}^T$ is the transpose of the tangent linear operator derived from the non-linear observation operator H.

Matrix $\mathbf{H}^T$ = Adjoint of operator H

Observation operators in data assimilation

Example of non-linear observation operator:

$$\text{model state} \quad \mathbf{x} = \begin{bmatrix} T \\ q_v \\ u \\ v \\ P_s \\ q_{liq} \\ q_{ice} \end{bmatrix} \xrightarrow{\ H\ } \tilde{\mathbf{y}}_i = \begin{bmatrix} Rad_{ch1} \\ Rad_{ch2} \\ Rad_{ch3} \end{bmatrix}$$

observation equivalent
= satellite cloudy radiances

H = radiative transfer model

Tangent-linear and adjoint operators

The tangent-linear operator is applied to perturbations:

$$\boldsymbol{\delta x} = \begin{bmatrix} \delta T \\ \delta q_v \\ \delta u \\ \delta v \\ \delta P_s \\ \delta q_{liq} \\ \delta q_{ice} \end{bmatrix} \xrightarrow{\mathbf{H}} \boldsymbol{\delta \tilde{y}}_i = \begin{bmatrix} \delta Rad_{ch1} \\ \delta Rad_{ch2} \\ \delta Rad_{ch3} \end{bmatrix}$$
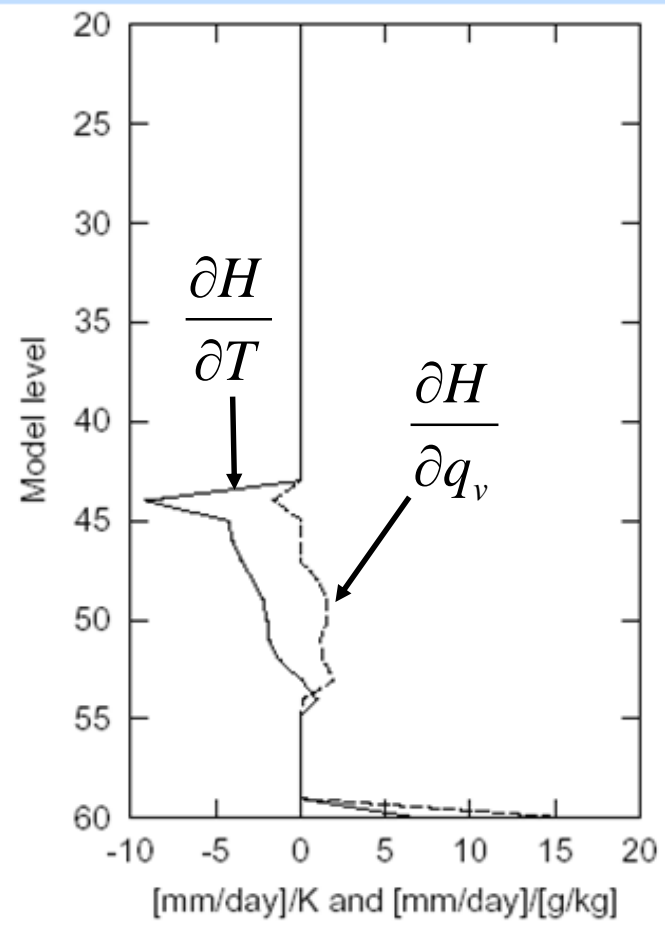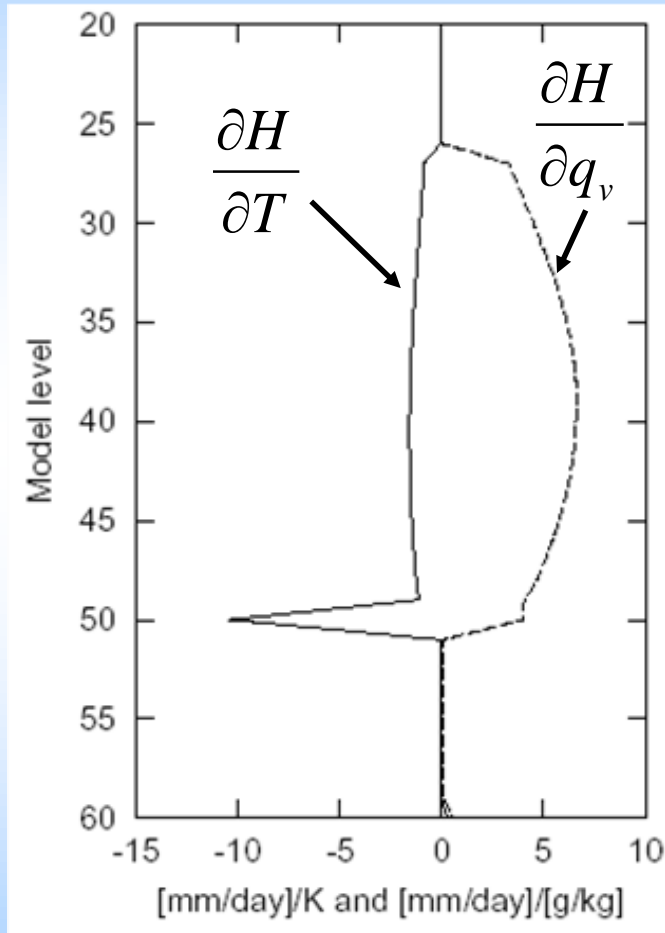
The adjoint operator is applied to the cost function gradient:

$$\nabla_{\tilde{\mathbf{y}}_i} J_o = \begin{bmatrix} \partial J_o / \partial Rad_{ch1} \\ \partial J_o / \partial Rad_{ch2} \\ \partial J_o / \partial Rad_{ch3} \end{bmatrix} \xrightarrow{\mathbf{H}^T} \nabla_{\mathbf{x}} J_o = \begin{bmatrix} \partial J_o / \partial T \\ \partial J_o / \partial q_v \\ \partial J_o / \partial u \\ \partial J_o / \partial v \\ \partial J_o / \partial P_s \\ \partial J_o / \partial q_{liq} \\ \partial J_o / \partial q_{ice} \end{bmatrix}$$

An example of observation operator

H: input = model state (T, q_v) → output = surface convective rainfall rate

Jacobians of surface rainfall rate w.r.t. T and q_v

$\dfrac{\partial H}{\partial T}$ $\dfrac{\partial H}{\partial q_v}$

$\dfrac{\partial H}{\partial T}$ $\dfrac{\partial H}{\partial q_v}$
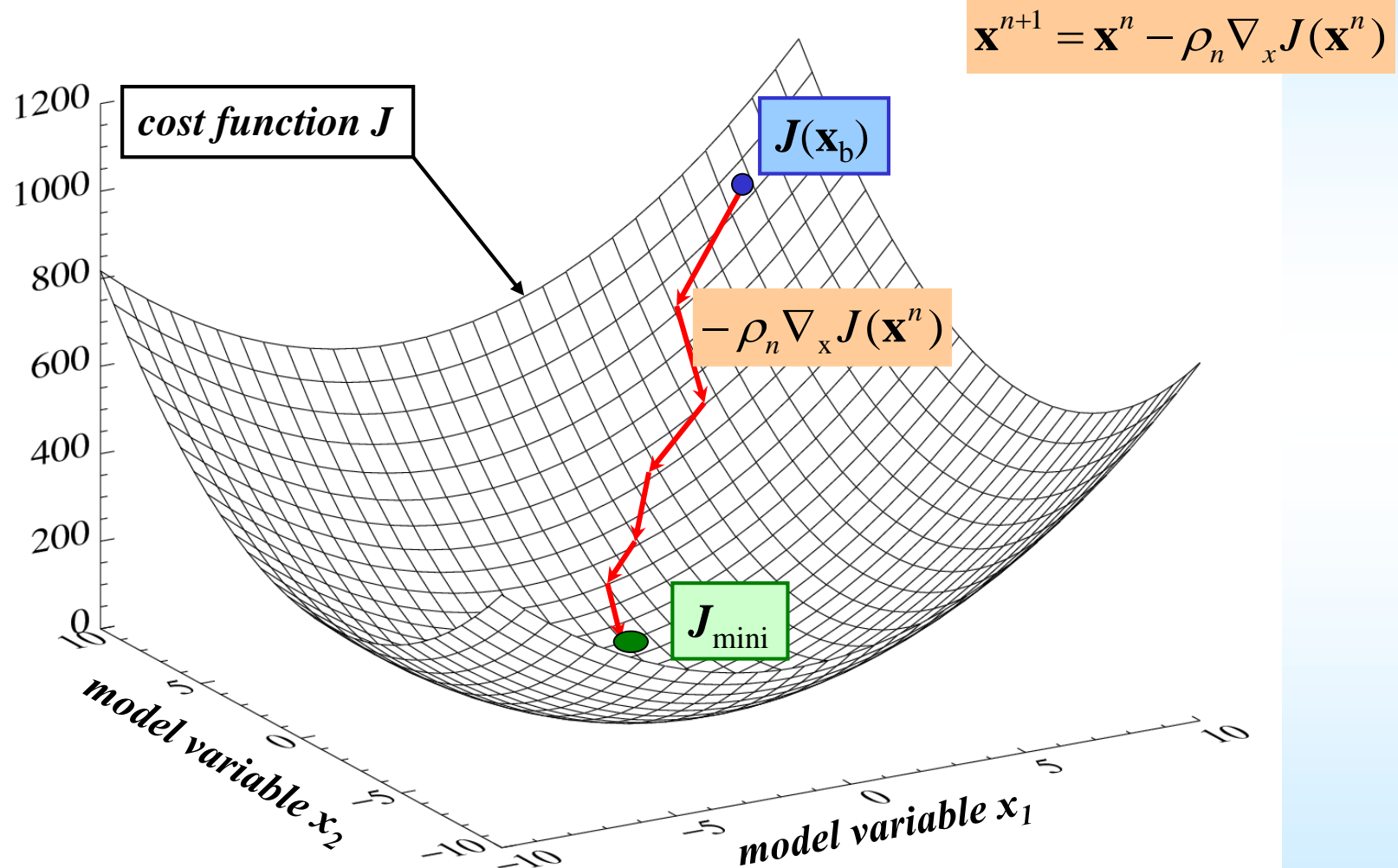
Marécal and Mahfouf (2002)

Betts-Miller (adjustment scheme)
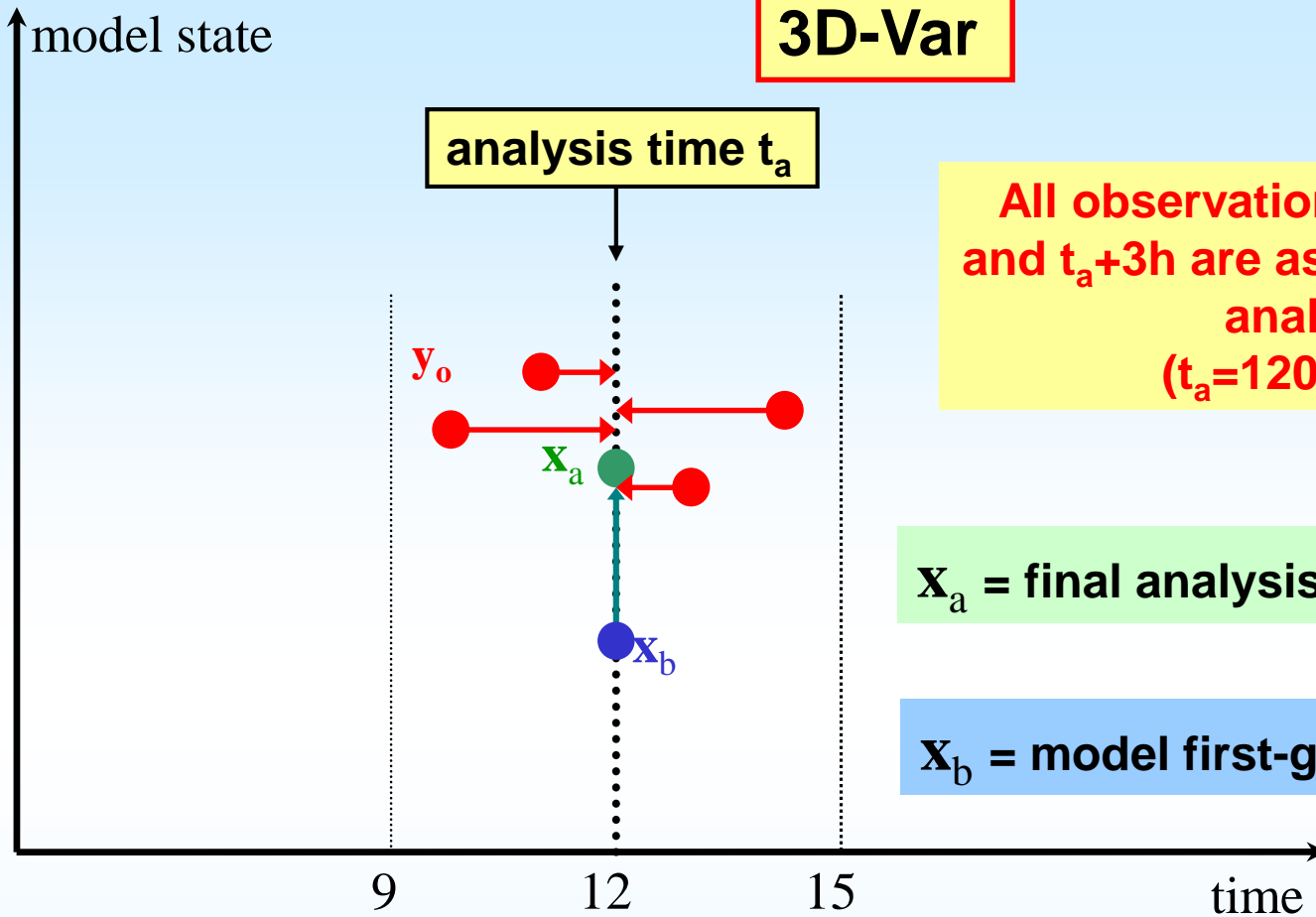
Tiedtke (ECMWF's oper mass-flux scheme)

The minimization of the cost function J is usually performed using an iterative minimization procedure

$$\mathbf{x}^{n+1} = \mathbf{x}^n - \rho_n \nabla_x J(\mathbf{x}^n)$$

cost function J

$J(\mathbf{x}_b)$

$-\rho_n \nabla_x J(\mathbf{x}^n)$

J_{mini}

model variable x_2

model variable x_1

Example with control vector $\mathbf{x} = (x_1, x_2)$

3D-Var

model state

analysis time t_a

All observations y_o between t_a-3h and t_a+3h are assumed to be valid at analysis time (t_a=1200 UTC here)

y_o

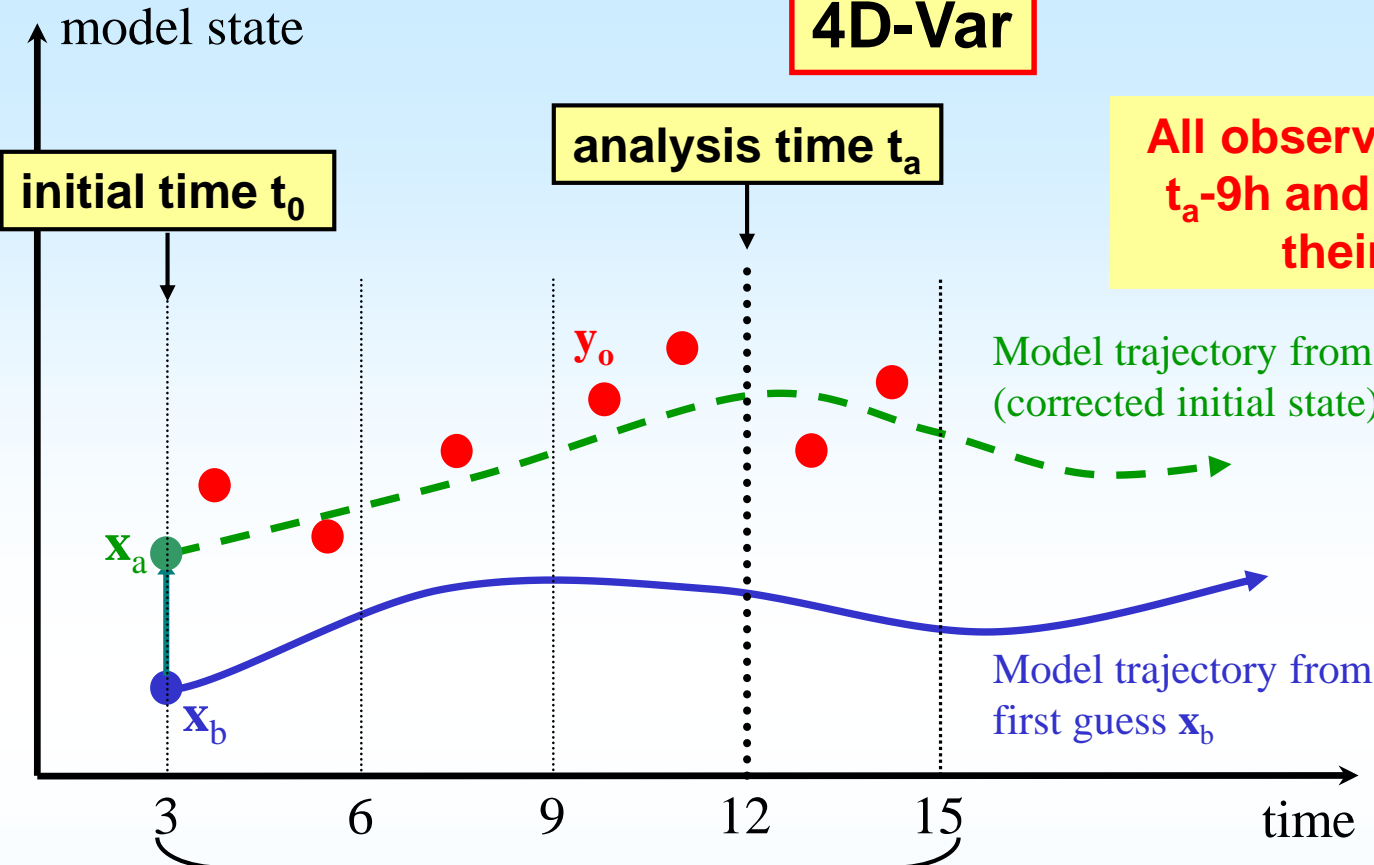$\mathbf{x}_a$

$\mathbf{x}_a$ = final analysis

$\mathbf{x}_b$

$\mathbf{x}_b$ = model first-guess

9   12   15   time

$$\min \; J = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}\left(H(\mathbf{x})-\mathbf{y}_o\right)^T \mathbf{R}^{-1}\left(H(\mathbf{x})-\mathbf{y}_o\right)$$

$$\Leftrightarrow \nabla_{\mathbf{x}} J = \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \mathbf{H}^T \mathbf{R}^{-1}\left(H(\mathbf{x})-\mathbf{y}_o\right) = 0$$

4D-Var

model state

initial time t_0

analysis time t_a

All observations y_o between t_a-9h and t_a+3h are valid at their actual time

y_o

Model trajectory from analysis (corrected initial state)

x_a

x_b

Model trajectory from first guess x_b

Forecast model is involved in minimization

3   6   9   12   15   **time**

assimilation window

Adjoint of forecast model with simplified linearized physics

$$\min \ J = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b) + \frac{1}{2}\sum_{i=0}\left(H_i(M[\mathbf{x}_0]) - \mathbf{y}_{oi}\right)^T \mathbf{R}_i^{-1}\left(H_i(M[\mathbf{x}_0]) - \mathbf{y}_{oi}\right)$$

$$\Leftrightarrow \nabla_{\mathbf{x}} J = \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b) + \sum_{i=0}^{n} \mathbf{M}^T[t_i, t_0] \mathbf{H}_i^T \mathbf{R}_i^{-1}\left(H_i(M[\mathbf{x}_0]) - \mathbf{y}_{oi}\right) = 0$$

EXAMPLE OF ADJOINT CODING

- **non-linear statement**

$$x \quad = \quad y + z^2$$

- **non-linear statement**

$$x \quad = \quad y + z^2$$

$$z \quad = \quad z$$

$$y \quad = \quad y$$

$$x \quad = \quad y + z^2$$

EXAMPLE OF ADJOINT CODING

- **non-linear statement**

$$x \;=\; y + z^2$$

$$z \;=\; z$$
$$y \;=\; y$$
$$x \;=\; y + z^2$$

- **tangent linear statement**

$$\delta z \;=\; \delta z$$
$$\delta y \;=\; \delta y$$
$$\delta x \;=\; \delta y + 2z\delta z$$

or in a matrix form:

$$\begin{pmatrix} \delta z \\ \delta y \\ \delta x \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2z & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \delta z \\ \delta y \\ \delta x \end{pmatrix}$$

- **adjoint statement**
  - transpose matrix

$$\begin{pmatrix} \delta z^* \\ \delta y^* \\ \delta x^* \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2z \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \delta z^* \\ \delta y^* \\ \delta x^* \end{pmatrix}$$

or in the form of equation set:

$$\begin{aligned} \delta z^* &= \delta z^* + 2z \delta x^* \\ \delta y^* &= \delta y^* + \delta x^* \\ \delta x^* &= 0 \end{aligned}$$

As an alternative to the matrix method, adjoint coding can be carried out using a **line-by-line** approach (what we do at ECMWF).

Automatic adjoint code generators do exist, but the output code is not optimized and not bug-free.

Testing the tangent-linear code

The correctness of the tangent-linear model, M, must be assessed by checking that the first-order Taylor approximation is valid:

$$\forall \delta \mathbf{x} \quad \lim_{\lambda \to 0} \frac{M(\mathbf{x} + \lambda\,\delta\mathbf{x}) - M(\mathbf{x})}{\lambda\,\mathbf{M}\delta\mathbf{x}} = 1$$

Example of output from a successful tangent-linear test:

λ	RATIO
0.1E-09	0.9994875881543574E+00
0.1E-08	0.9999477148855701E+00
0.1E-07	0.9999949234236705E+00
0.1E-06	0.9999993501022509E+00
0.1E-05	0.9999999496119013E+00
0.1E-04	0.9999999611338369E+00
0.1E-03	0.9999953179193711E+00
0.1E-02	0.9999724488345042E+00
0.1E-01	0.9998727842790062E+00
0.1E+00	0.9978007454264978E+00
0.1E+01	0.9683066504549524E+00

Tiny perturbations

Larger perturbations

Machine precision reached

Improvement when perturbation size decreases

ECMWF

Testing the adjoint code

The correctness of the adjoint model needs to be assessed by checking that it satisfies the mathematical relationship:

$$\forall \delta \mathbf{x}, \delta \mathbf{y} \quad \langle \mathbf{M} \delta \mathbf{x}, \delta \mathbf{y} \rangle = \langle \delta \mathbf{x}, \mathbf{M}^T \delta \mathbf{y} \rangle$$

where $\mathbf{M}$ is the tangent-linear model and $\mathbf{M}^T$ is the adjoint model.

Example of output from a successful adjoint test:

$$\langle \mathbf{M}\, \delta\mathbf{x}, \quad \delta\mathbf{y} \rangle = -0.13765102625164E\text{-}01$$
$$\langle \quad \delta\mathbf{x}, \mathbf{M}^T\, \delta\mathbf{y} \rangle = -0.13765102625168E\text{-}01$$
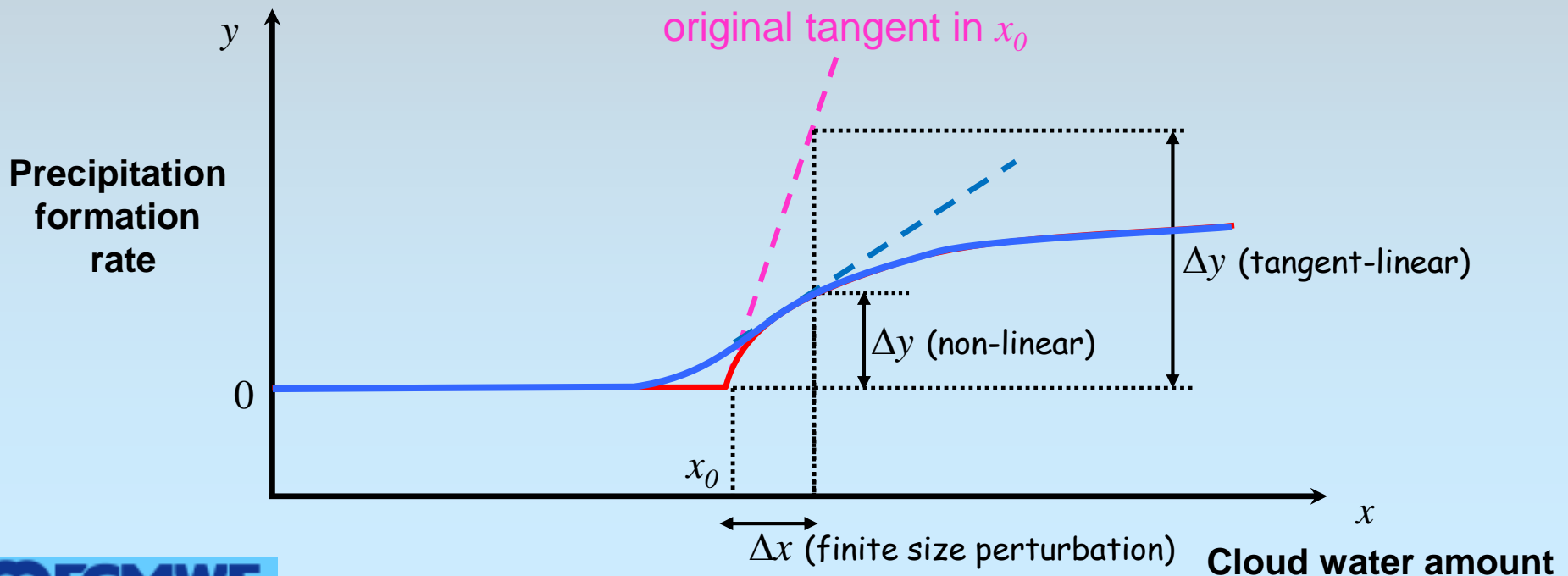
The difference is 11.351 times the zero of the machine

The adjoint test should be correct at the level of machine precision (typ. 13 to 15 digits for the entire model).
Otherwise there must be a bug in the code!

ECMWF

Linearity assumption

- Variational assimilation is based on the strong assumption that the analysis is performed in a (**quasi-)linear** framework.

- However, in the case of physical processes, strong non-linearities can occur in the presence of discontinuous/non-differentiable processes
(e.g. **switches or thresholds** in cloud water and precipitation formation).

→ "Regularization" needs to be applied: **smoothing of functions, reduction of some perturbations.**

y

original tangent in x_0

Precipitation formation rate

Δy (tangent-linear)

Δy (non-linear)

0

x_0

x

Δx (finite size perturbation)

Cloud water amount

ECMWF

An example of spurious TL noise caused by a threshold in the autoconversion formulation of the large-scale cloud scheme.

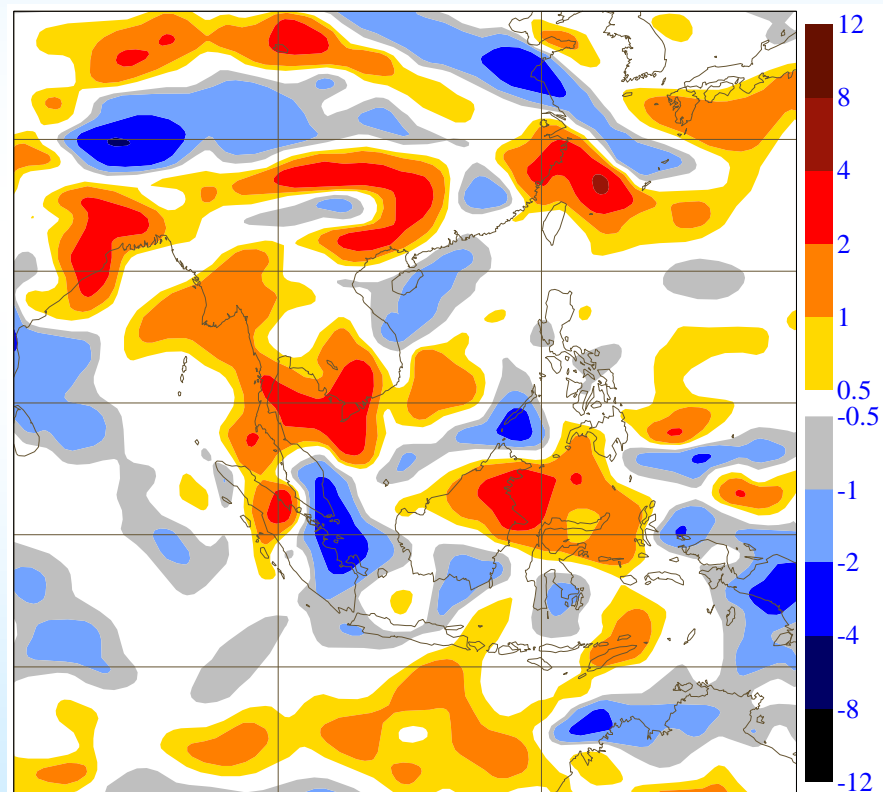**~700 hPa zonal wind increments [m/s] from 12h model integration.**

Nonlinear finite difference:
$$M(\mathbf{x}+\delta\mathbf{x}) - M(\mathbf{x})$$
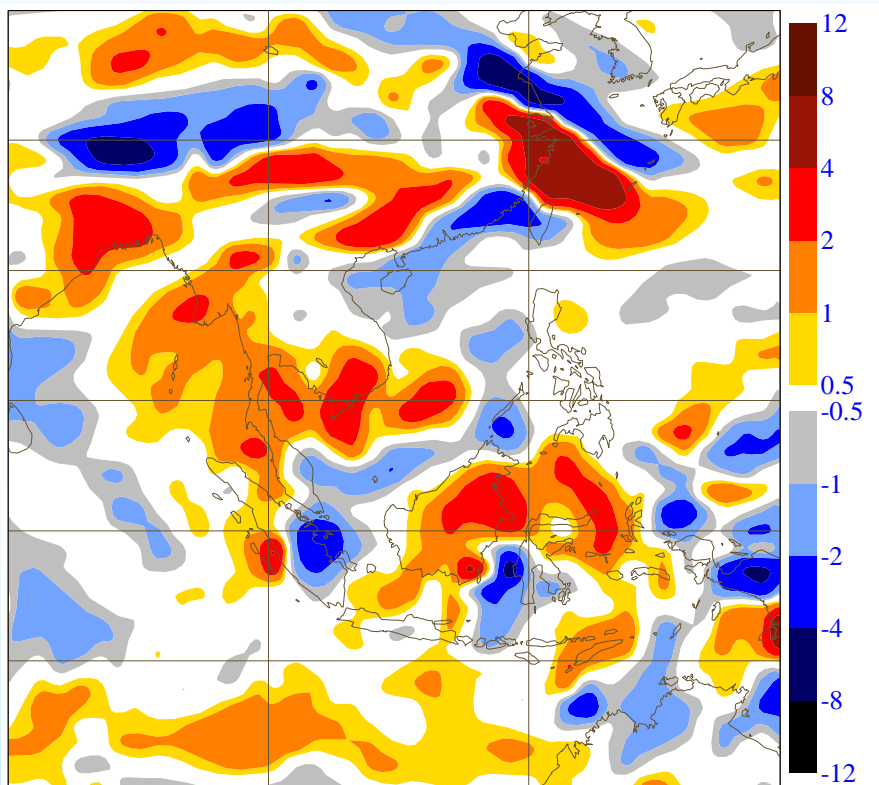
Tangent-linear integration: $\mathbf{M}\delta\mathbf{x}$

from M. Janisková

with perturbation reduction in autoconversion

ECMWF

ECMWF operational LP package (operational 4D-Var)

Currently used in ECMWF operational 4D-Var minimizations (main simplifications with respect to the full non-linear versions are highlighted **in red**):

- **Large-scale condensation scheme:** *[Tompkins and Janisková 2004]*
    - based on a uniform PDF to describe subgrid-scale fluctuations of total water,
    - melting of snow included,
    - precipitation evaporation included,
    - **reduction of cloud fraction perturbation and in autoconversion of cloud into rain.**

- **Convection scheme:** *[Lopez and Moreau 2005]*
    - mass-flux approach *[Tiedtke 1989]*,
    - deep convection (CAPE closure) and shallow convection (q-convergence) are treated,
    - perturbations of all convective quantities are included,
    - coupling with cloud scheme through detrainment of liquid water from updraught,
    - **some perturbations (buoyancy, initial updraught vertical velocity) are reduced.**

- **Radiation:** TL and AD of longwave and shortwave radiation available *[Janisková et al. 2002]*
    - **shortwave:** based on *Morcrette (1991)*, **only 2 spectral intervals** (instead of 6 in non-linear version),
    - **longwave:** based on *Morcrette (1989)*, **called every 2 hours only.**

ECMWF operational LP package (operational 4D-Var)

- **Vertical diffusion:**
    - mixing in the surface and planetary boundary layers,
    - based on K-theory and Blackadar mixing length,
    - exchange coefficients based on *Louis et al. [1982]*, near surface,
    - Monin-Obukhov higher up,
    - mixed layer parametrization and PBL top entrainment recently added.
    - **Perturbations of exchange coefficients are smoothed (esp. near the surface).**

- **Orographic gravity wave drag:** *[Mahfouf 1999]*
    - subgrid-scale orographic effects *[Lott and Miller 1997]*,
    - **only low-level blocking part is used.**

- **Non-orographic gravity wave drag:** *[Oor et al. 2010]*
    - isotropic spectrum of non-orographic gravity waves *[Scinocca 2003]*,
    - **Perturbations of output wind tendencies below 200 hPa reset to zero.**

- **RTTOV** is employed to simulate radiances at individual frequencies (infrared, longwave and microwave, with cloud and precipitation effects included) to compute model–satellite departures in observation space.

Impact of linearized physics on tangent-linear approximation

Comparison:

Finite difference of two NL integrations $\leftrightarrow$ TL evolution of initial perturbations

Examination of the accuracy of the linearization for typical analysis increments:

$$M(\mathbf{x}_{an}) - M(\mathbf{x}_{bg}) \leftrightarrow \mathbf{M}'(\underbrace{\mathbf{x}_{an} - \mathbf{x}_{bg}})$$

typical size of 4D-Var
analysis increments

Diagnostics:

- **Mean absolute errors:** $\varepsilon = \overline{\left| \left[M(\mathbf{x}_{an}) - M(\mathbf{x}_{bg}) \right] - \mathbf{M}'(\mathbf{x}_{an} - \mathbf{x}_{bg}) \right|}$

- **Relative error change:** $\dfrac{\varepsilon_{\text{EXP}} - \varepsilon_{\text{REF}}}{\varepsilon_{\text{REF}}} \times 100\%$ **(improvement if < 0)**

- **Here: REF = adiabatic run (i.e. no physical parametrizations in tangent-linear)**

ECMWF

Zonal mean cross-section of change in TL error when TL includes:

   VDIF + orog. GWD + SURF

Relative to adiabatic TL run (50-km resol.; 20 runs; after 12h integr.).

Blue = TL error

reduction = ☺

Temperature

$(\varepsilon - \varepsilon_{AD})/\varepsilon_{AD}$, Temperature, f021, (20 members)
Mean = -8.84e+00 %
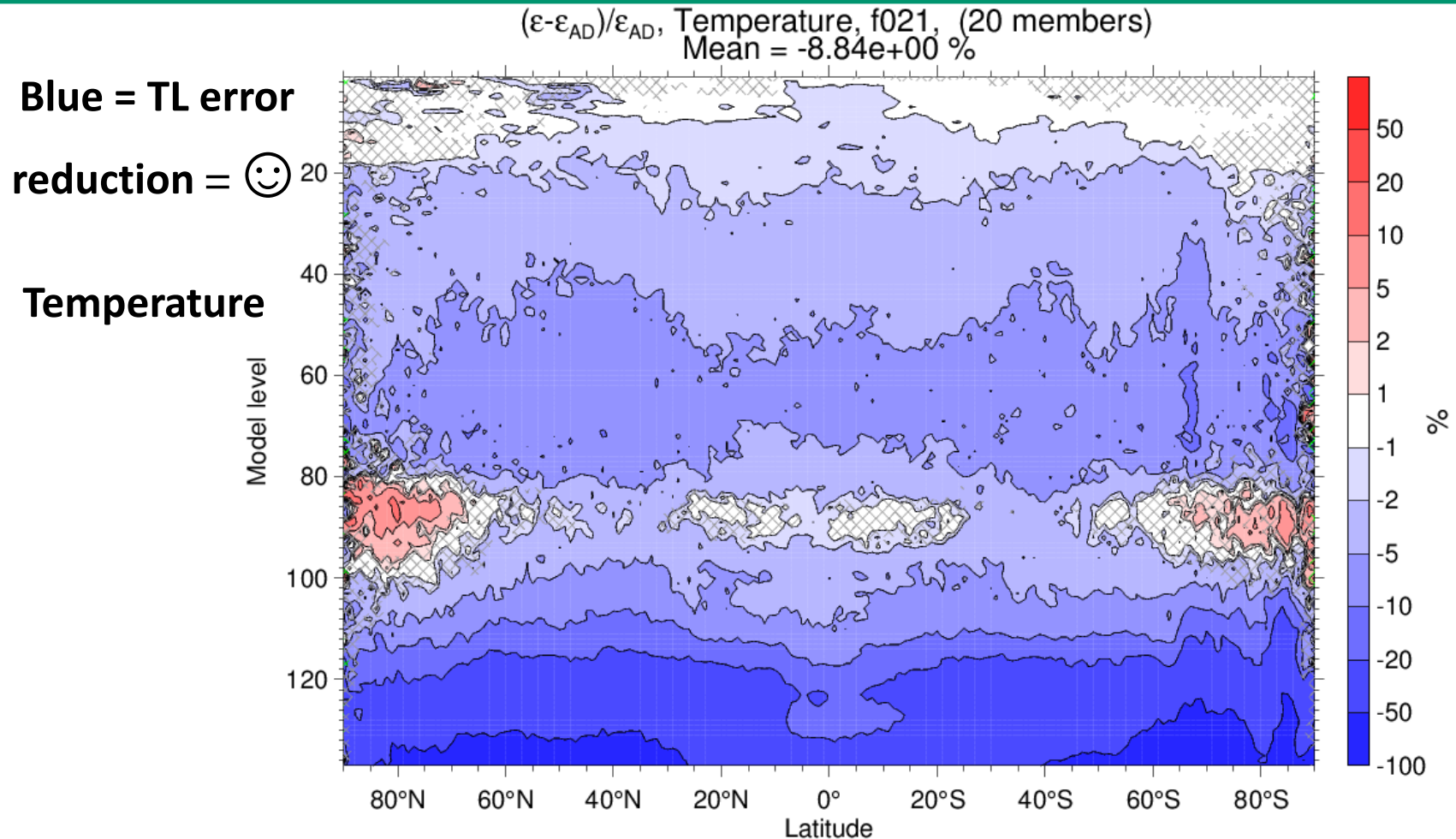
Impact of linearized physics on TL approximation (2)

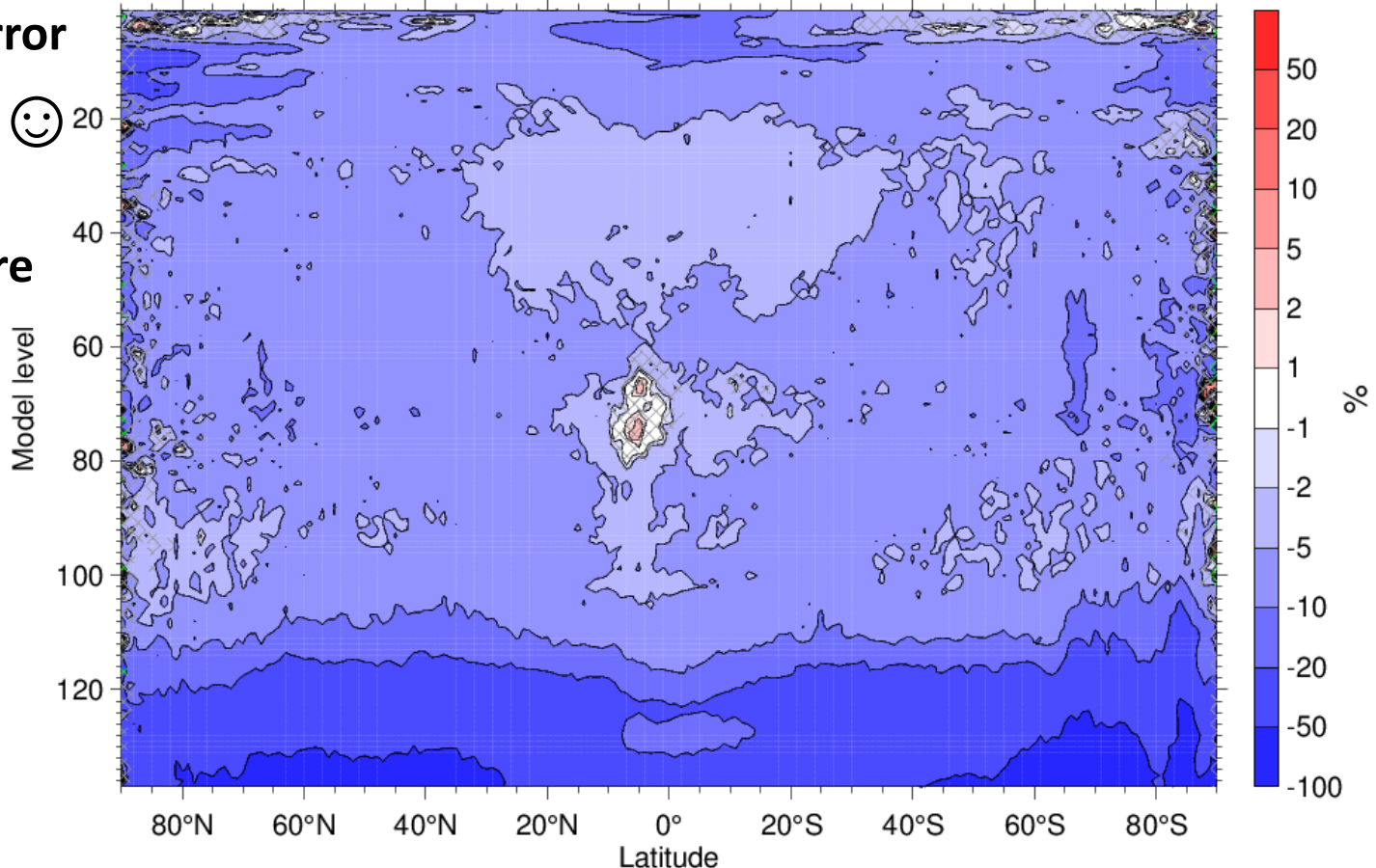Zonal mean cross-section of change in TL error when TL includes:

VDIF + orog. GWD + SURF + RAD

Relative to adiabatic TL run (50-km resol.; 20 runs; after 12h integr.).

Blue = TL error

reduction = ☺

Temperature

$(\varepsilon - \varepsilon_{AD})/\varepsilon_{AD}$, Temperature, f022, [T399 L137 +48] (20 members)
Mean = -1.08e+01 %

Zonal mean cross-section of change in TL error when TL includes:

VDIF + orog. GWD + SURF + RAD + non-orog GWD + moist physics

Relative to adiabatic TL run (50-km resol.; 20 runs; after 12h integr.).

Blue = TL error

reduction = ☺

Temperature

$(\varepsilon - \varepsilon_{AD})/\varepsilon_{AD}$, Temperature, f025, [T399 L137 +48] (20 members)
Mean = -1.95e+01 %
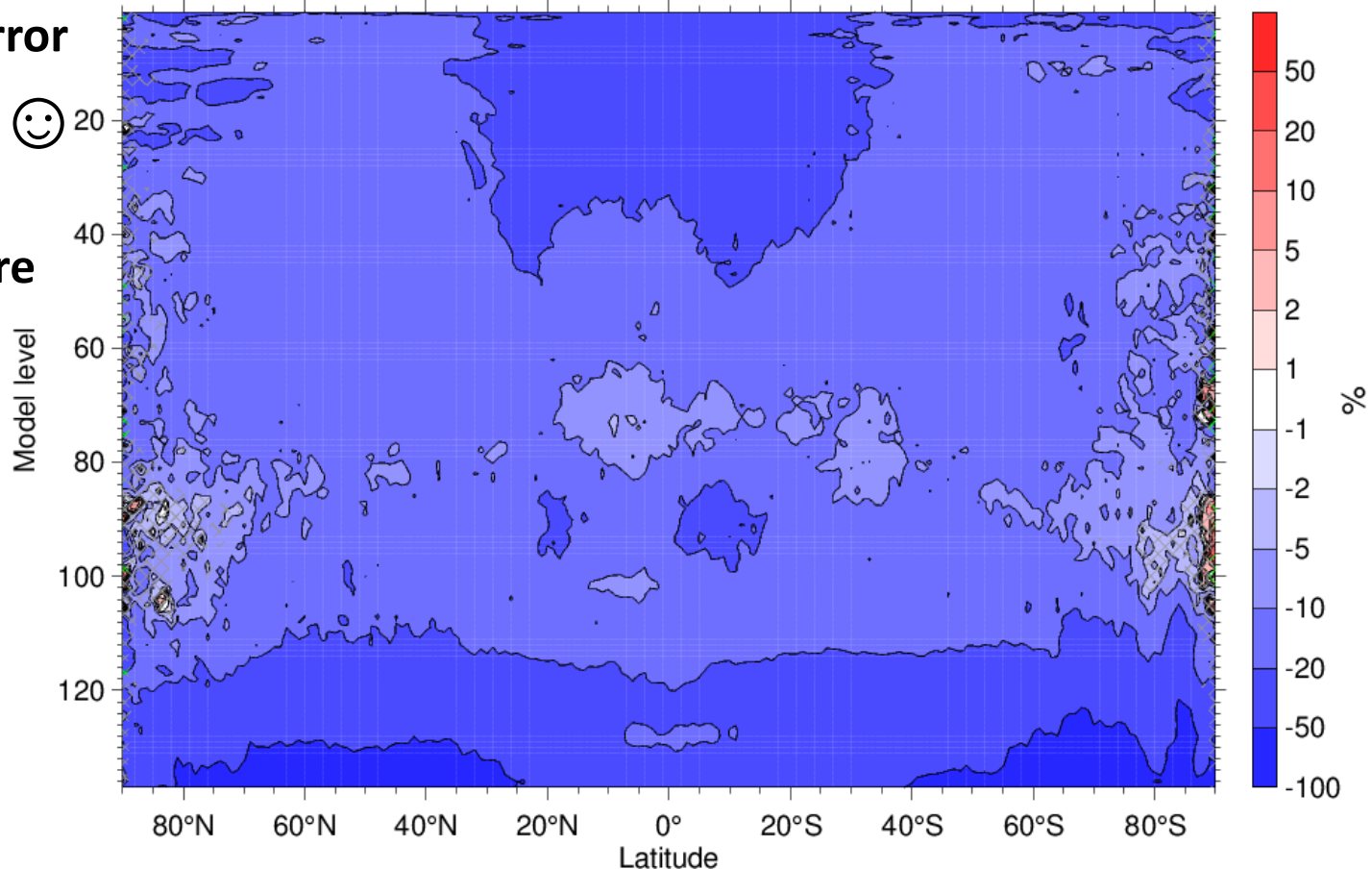
Applications

1D-Var with radar reflectivity profiles

Background
$\mathbf{x}_b = (T_b, q_b, \ldots)$

$$J_b = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b)$$

$$\nabla_{\mathbf{x}} J_b = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b)$$

Initialize

$\mathbf{x} = \mathbf{x}_b$

$\mathbf{x} = (T, q, \ldots)$

no

$\nabla_{\mathbf{x}} J \approx 0 ?$

MINIM

$\nabla_{\mathbf{x}} J$

$\nabla_{\mathbf{x}} J_o$

yes

Moist Physics

Analysis

$\mathbf{x}_a = \mathbf{x}$

Moist Physics (ADJ)

Reflectivity Model

Reflectivity Model (ADJ)

Z_{mod}

$$J_o = \frac{1}{2} \sum_{k=1}^{K} \left(\frac{Z_{mod}^k - Z_{obs}^k}{\sigma_{obs}^k} \right)^2$$

$$\nabla_{\mathbf{z}} J_o = \left[\frac{Z_{mod}^k - Z_{obs}^k}{\left(\sigma_{obs}^k\right)^2} \right]_{k=1,K}$$

Reflectivity observations
Z_{obs} **with errors** σ_{obs}

K = number of model vertical levels
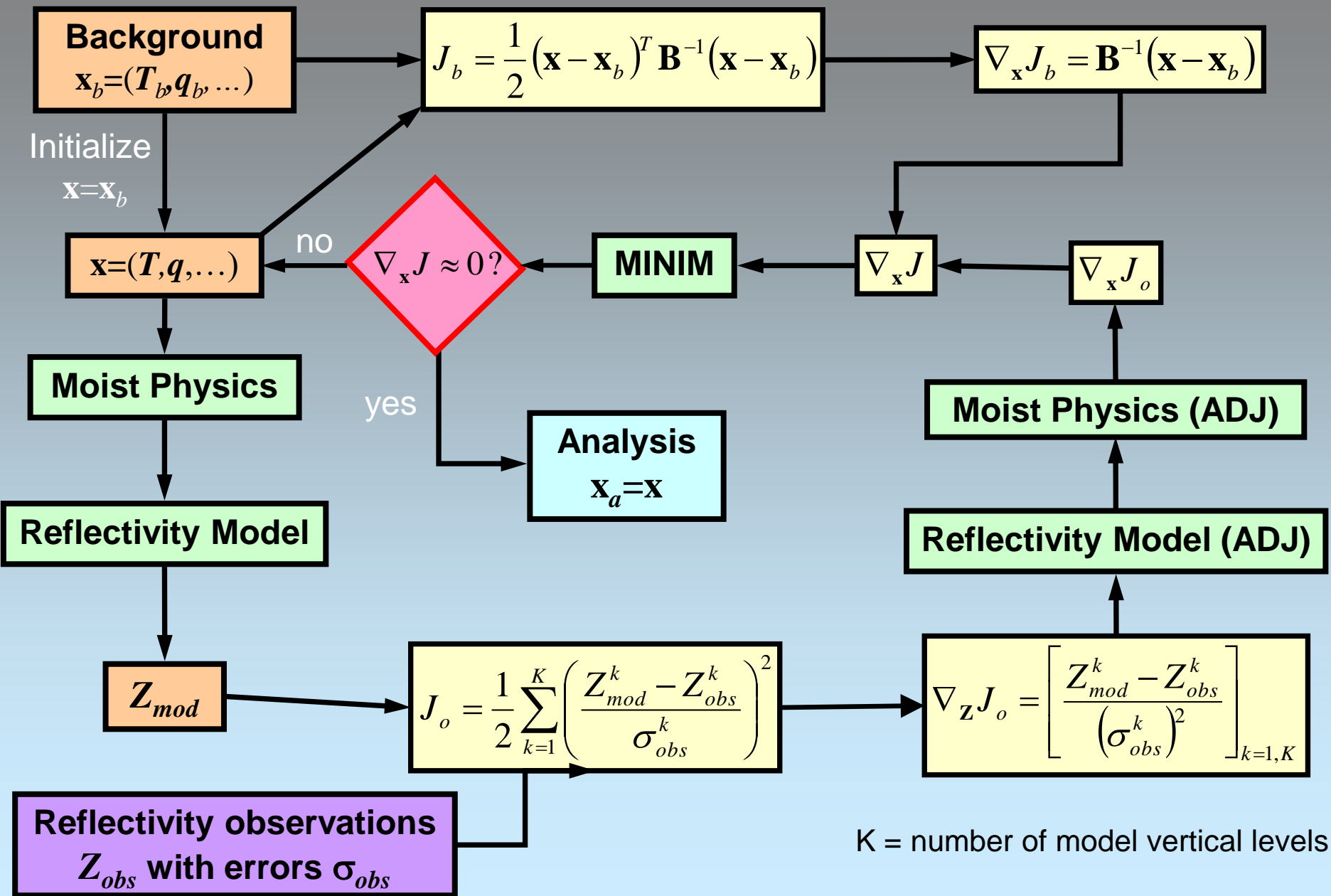
Impact of ECMWF linearized physics on forecast scores

Comparison of two T511 L91 4D-Var 3-month experiments with & without full linearized physics: Relative change in forecast anomaly correlation.
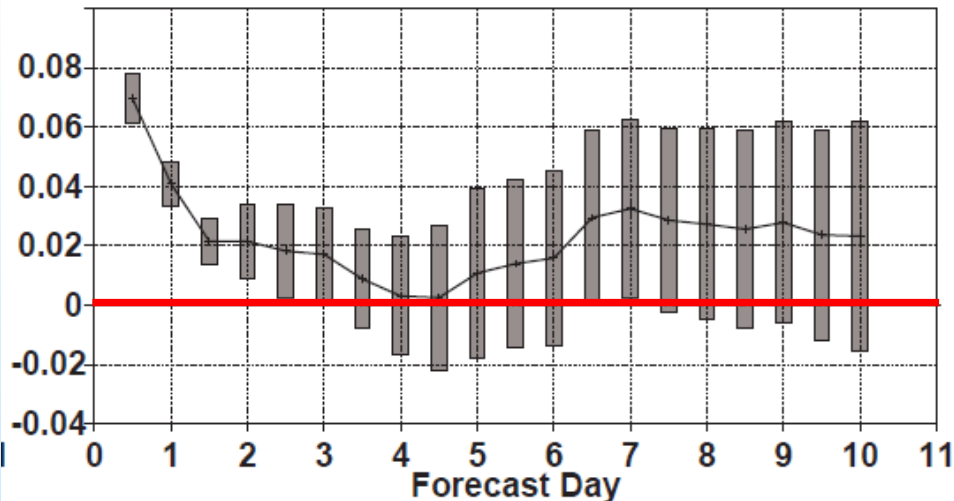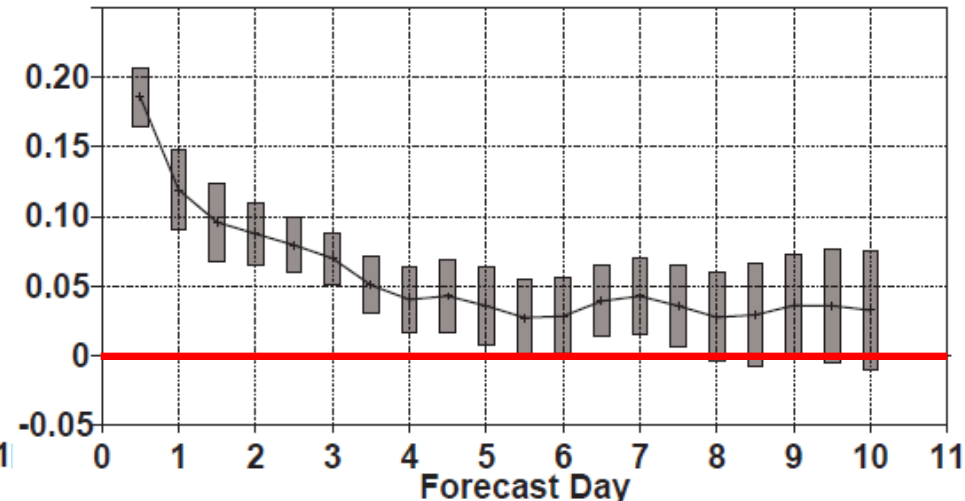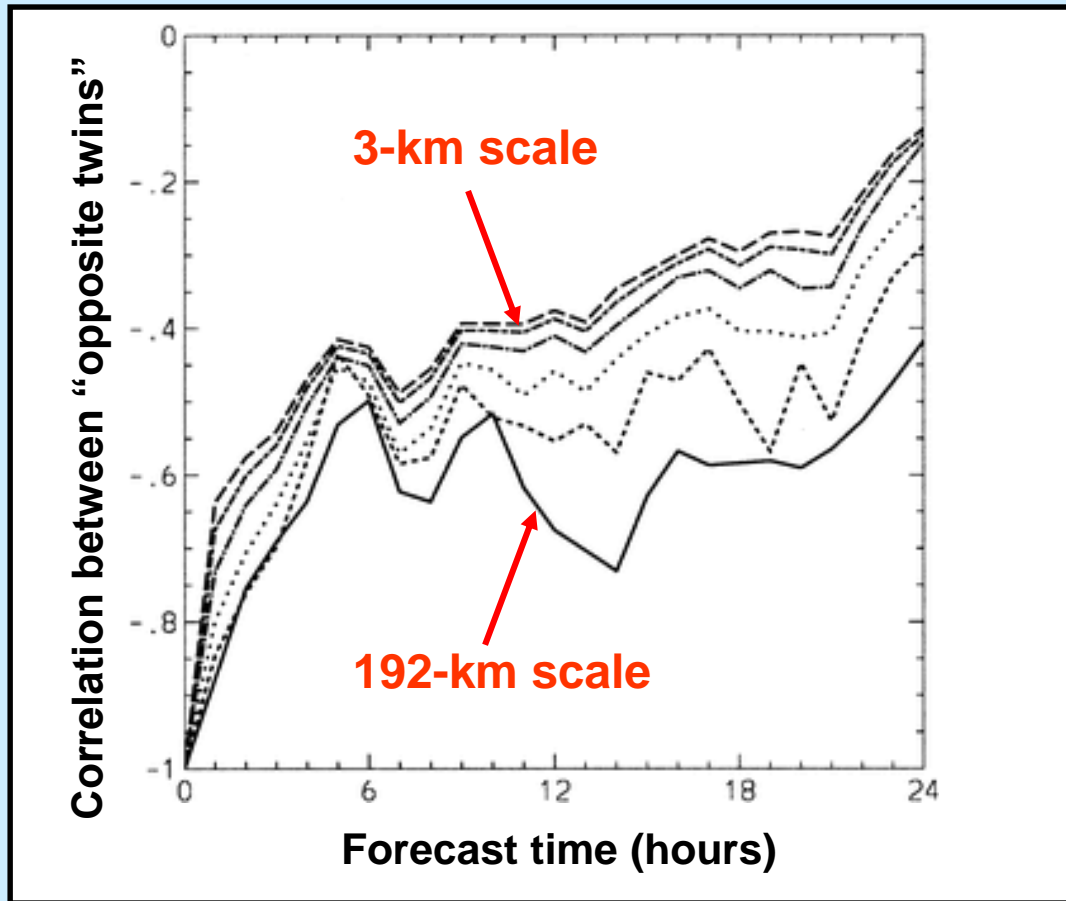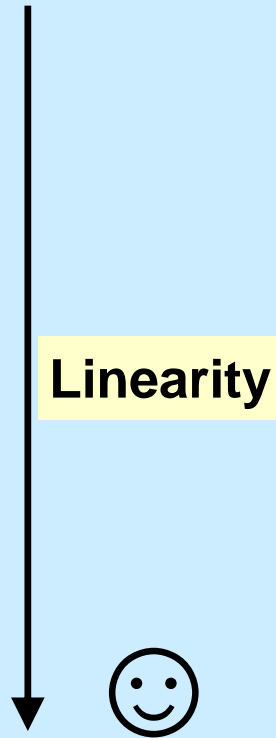
Influence of time and resolution on linearity assumption in physics

Results from ensemble runs with the MC2 model (3 km resolution) over the Alps, from *Walser et al. (2004)*.
Comparison of a pair of "opposite twin" experiments.

Linearity

☺

→ **The validity of the linear assumption for precipitation quickly drops in the first hours of the forecast, especially for smaller scales.**

Summary and prospects (1)

- **Linearized physical parameterizations have become essential components of variational data assimilation systems (4D-Var):**

→ **Better representation of the evolution of the atmospheric state during the minimization of the cost function (via the adjoint model integration).**

→ **Extraction of information from observations that are strongly affected by physical processes (e.g. by clouds or precipitation).**

- **However, there are some <u>limitations</u> to the LP approach:**

  1) <u>Theoretical</u>:

  The domain of validity of the linear hypothesis shrinks with increasing resolution and integration length.

  2) <u>Technical</u>:

  Linearized models require sustained & time-consuming attention:

  → **Testing tangent-linear approximation and adjoint code.**
  → **Regularizations / simplifications to eliminate any source of instability.**
  → **Revisions to ensure good match with reference non-linear forecast model.**

- **In practice, it all comes down to achieving the best compromise between:**

<div align="center">

Realism

Cost   **Linearity**

</div>
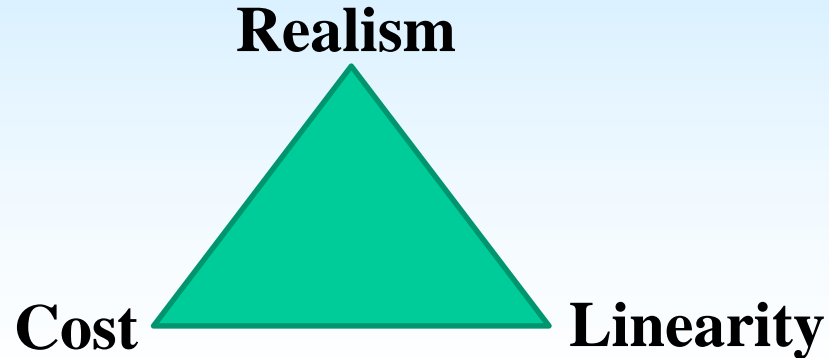
- **Alternative data assimilation methods exist that do not require the development of linearized code, but so far none of them has been able to outperform 4D-Var, especially in global models:**

  → **Ensemble Kalman Filter (EnKF; still relies on the linearity assumption),**

  → **Particle filters (difficult to implement for high-dimensional problems).**

- **So what is the future of LP?**

From a small challenge…

… to a much bigger challenge…

- **Eventually, it might become impractical or even impossible to make LP work efficiently at resolutions of a few kilometres, even if the linearity constraint can be relaxed (e.g. by using shorter 4D-Var window or weak-constraint 4D-Var).**

?

Thank you!

Adjoint technique

- Non-linear observation operator:

$$\mathbf{y} = H(\mathbf{x})$$

- Tangent linear operator:

$$\delta\mathbf{y} = \mathbf{H}(\delta\mathbf{x})$$

- **H** is the Jacobian matrix derived from H:

$$\mathbf{H}_{ij} \quad = \quad \frac{\partial y_i}{\partial x_j}$$

$$\delta y_i \quad = \quad \sum_{j=1}^{N} \frac{\partial y_i}{\partial x_j} \, \delta x_j$$

Adjoint technique

- **Observation term of the cost-function:**

$$\mathcal{J}_o = \frac{1}{2}(\mathbf{y} - \mathbf{y}_o)^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{y}_o)$$

- **Gradient with respect to y:**

$$\nabla_y \mathcal{J}_o = \mathbf{R}^{-1}(\mathbf{y} - \mathbf{y}_o)$$

- **Gradient with respect to x:**

$$\frac{\partial \mathcal{J}_o}{\partial x_i} = \sum_{j=1}^{M} \frac{\partial \mathcal{J}_o}{\partial y_j} \underbrace{\frac{\partial y_j}{\partial x_i}}_{\mathbf{H}_{ij}^T}$$

which involves the adjoint (transpose) of the tangent-linear operator.

- **Finally:**

$$\nabla_x \mathcal{J}_o = \mathbf{H}^T(\nabla_y \mathcal{J}_o) = \mathbf{H}^T \mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y}_o)$$

TL AND AD MODELS

- **TANGENT LINEAR MODEL**

  If M is a model such as:

  $$\mathbf{x}(t_{i+1}) = M[\mathbf{x}(t_i)]$$

  then the tangent linear model of M, called M', is:

  $$\delta\mathbf{x}(t_{i+1}) = M'[\mathbf{x}(t_i)]\delta\mathbf{x}(t_i) = \frac{\partial M[\mathbf{x}(t_i)]}{\partial\mathbf{x}}\delta\mathbf{x}(t_i)$$

- **ADJOINT MODEL**

  The adjoint of a linear operator M' is the linear operator $\mathbf{M}^*$ such that, for the inner product $<,>$,
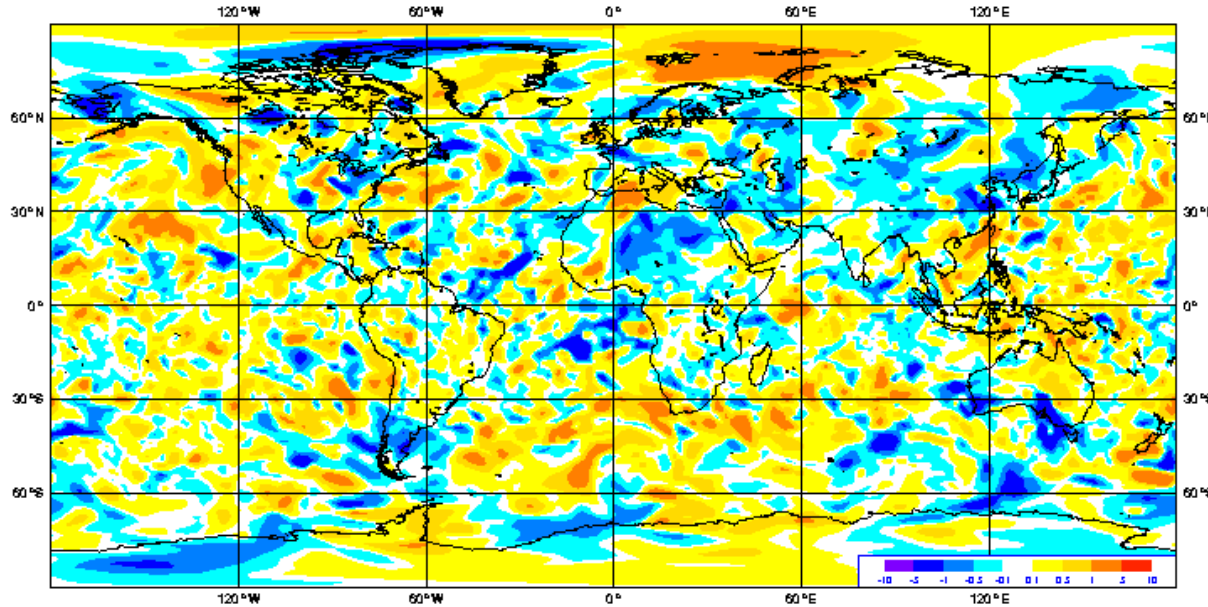
  $$\forall\mathbf{x}, \forall\mathbf{y} \qquad < M'\mathbf{x}, \mathbf{y} >=< \mathbf{x}, \mathbf{M}^*\mathbf{y} >$$

  Remarks:

  – with the euclidian inner product, $\mathbf{M}^* = M'^T$.

  – in variational assimilation, $\nabla_x\mathcal{J} = \mathbf{M}^*\nabla_y\mathcal{J}$, where $\mathcal{J}$ is the cost function.

Importance of regularization to prevent instabilities in tangent-linear model

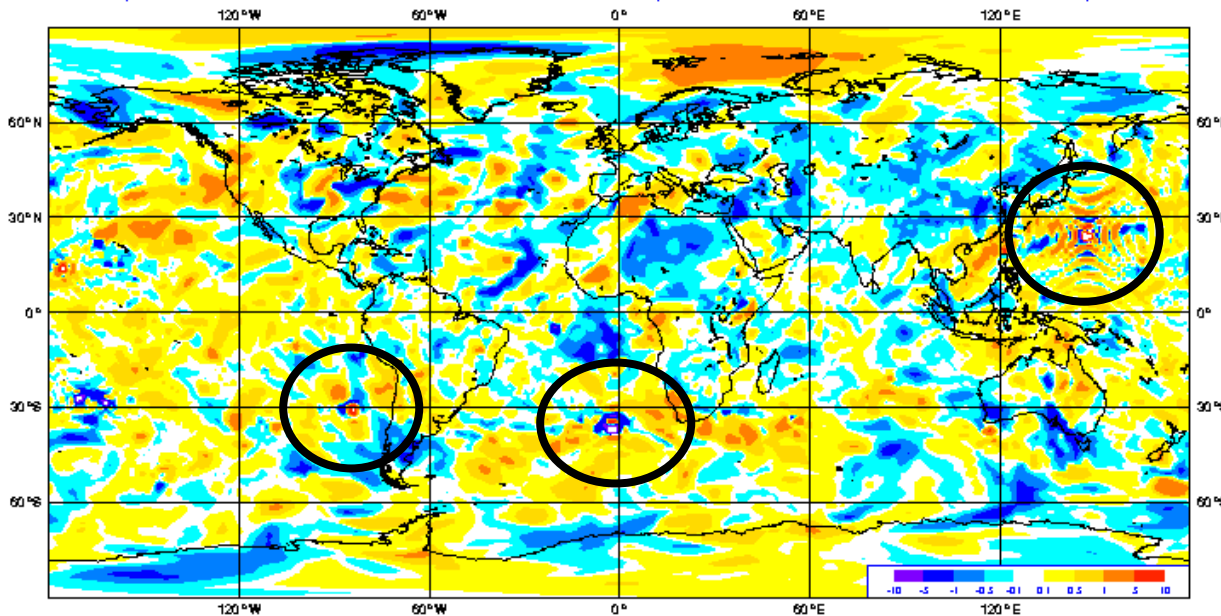Thursday 15 March 2001 12UTC ECMWF Forecast t+12 VT: Friday 16 March 2001 00UTC Model Level 48 **Temperature

Thursday 15 March 2001 12UTC ECMWF Forecast t+12 VT: Friday 16 March 2001 00UTC Model Level 48 **Temperature

12-hour ECMWF model integration (T159 L60)

Temperature on level 48 (approx. 850 hPa)

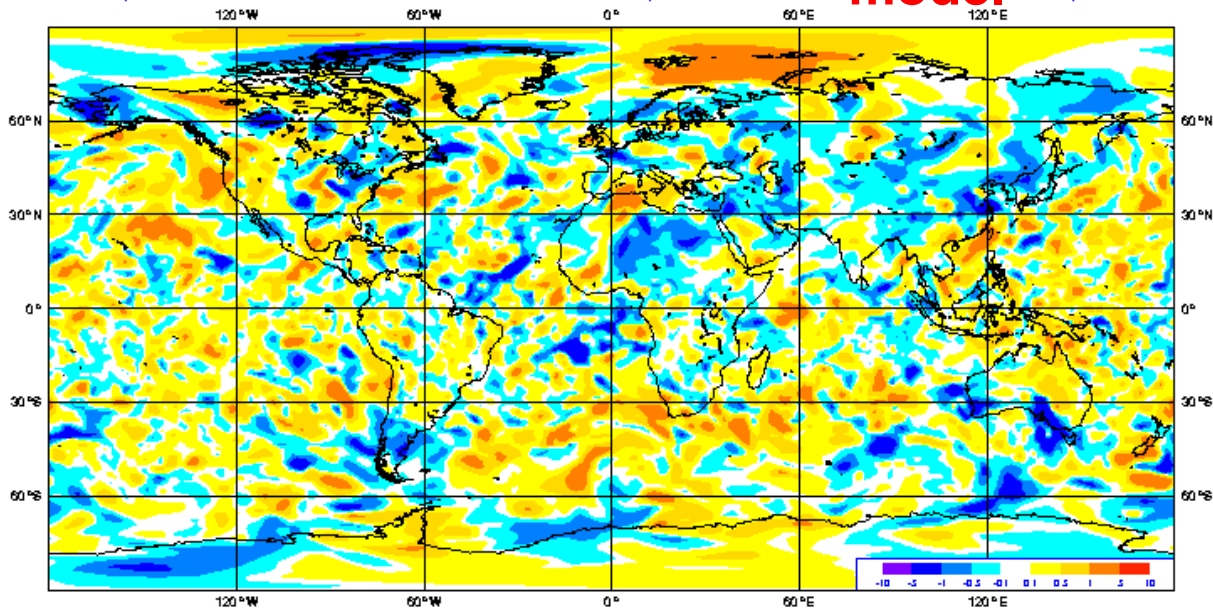**Finite difference between two non-linear model integrations**

Corresponding perturbations evolved with tangent-linear model

No regularization in convection scheme

Importance of regularization to prevent instabilities in tangent-linear model

Thursday 15 March 2001 12UTC ECMWF Forecast t+12 VT: Friday 16 March 2001 00UTC Model Level 48 **Temperature

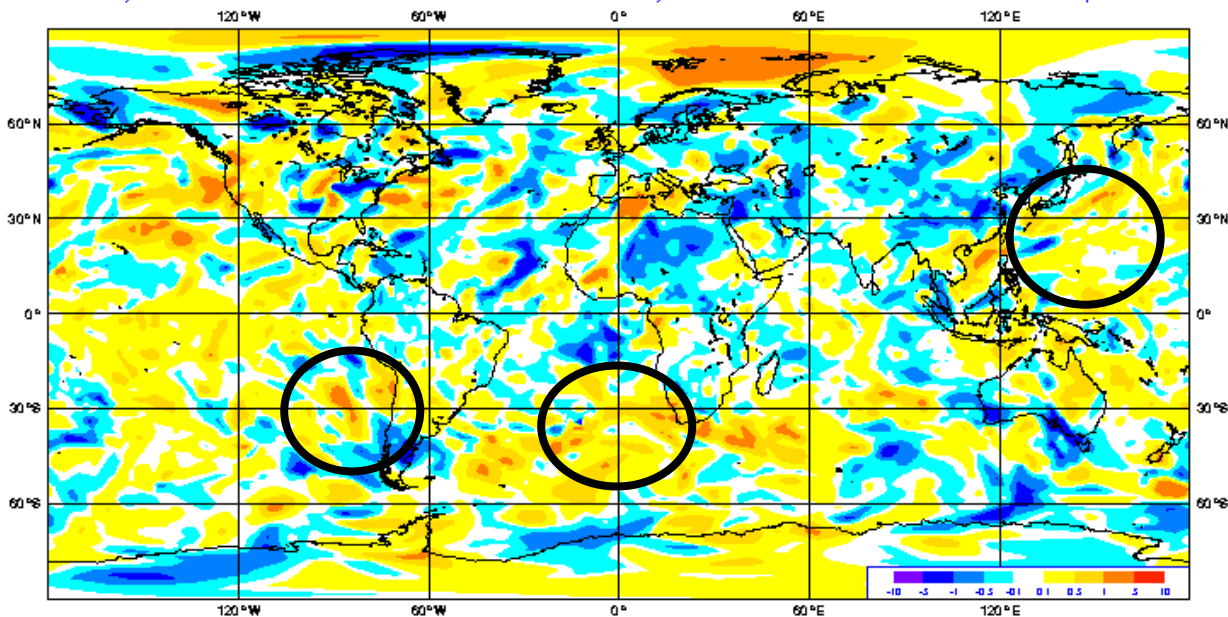Thursday 15 March 2001 12UTC ECMWF Forecast t+12 VT: Friday 16 March 2001 00UTC Model Level 48 **Temperature

12-hour ECMWF model integration (T159 L60)

Temperature on level 48 (approx. 850 hPa)

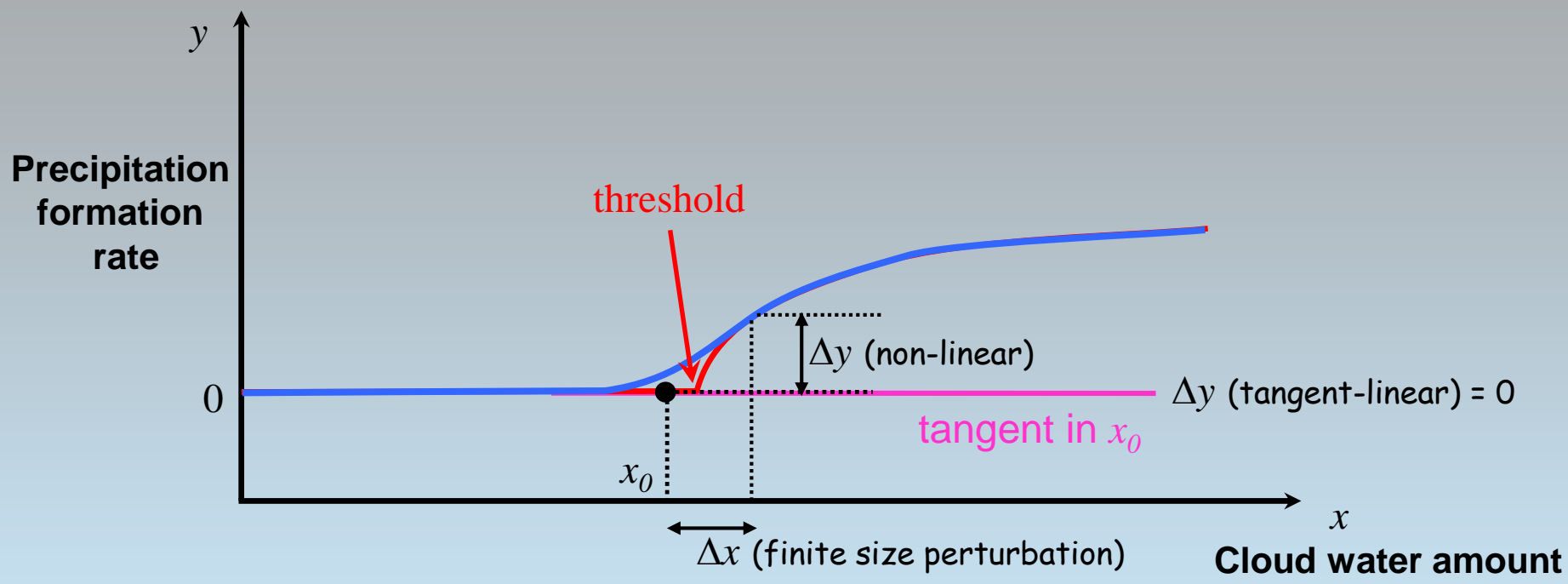**Finite difference between two non-linear model integrations**

Corresponding perturbations evolved with tangent-linear model

Regularization in convection scheme (buoyancy & updraught velocity reduced perturb.)

Linearity issue

Precipitation formation rate

threshold

Δy (non-linear)

0

Δy (tangent-linear) = 0

tangent in x_0

x_0

Δx (finite size perturbation)

x

Cloud water amount

ECMWF

Basic rules for line-by-line adjoint coding (1)

Adjoint statements are derived from tangent linear ones in a reversed order

Tangent linear code	Adjoint code
$\delta x = 0$	$\delta x^* = 0$
$\delta x = A\,\delta y + B\,\delta z$	$\delta y^* = \delta y^* + A\,\delta x^*$ $\delta z^* = \delta z^* + B\,\delta x^*$ $\delta x^* = 0$
$\delta x = A\,\delta x + B\,\delta z$	$\delta z^* = \delta z^* + B\,\delta x^*$ $\delta x^* = A\,\delta x^*$

do k = 1, N
     $\delta x(k) = A\,\delta x(k-1) + B\,\delta y(k)$
 end do
	do k = N, 1, – 1 (Reverse the loop!)
	     $\delta x^*(k-1) = \delta x^*(k-1) + A\,\delta x^*(k)$
	     $\delta y^*(k) = \delta y^*(k) + B\,\delta x^*(k)$
	     $\delta x^*(k) = 0$
	 end do

if (*condition*) *tangent linear code*	if (*condition*) *adjoint code*

And do not forget to initialize local adjoint variables to zero !

Basic rules for line-by-line adjoint coding (2)

To save memory, the trajectory can be recomputed just before the adjoint calculations.

The most common sources of error in adjoint coding are:
1) Pure coding errors (often: confusion trajectory/perturbation variables),
2) Forgotten initialization of local adjoint variables to zero,
3) Mismatching trajectories in tangent linear and adjoint (even slightly),
4) Bad identification of trajectory updates:

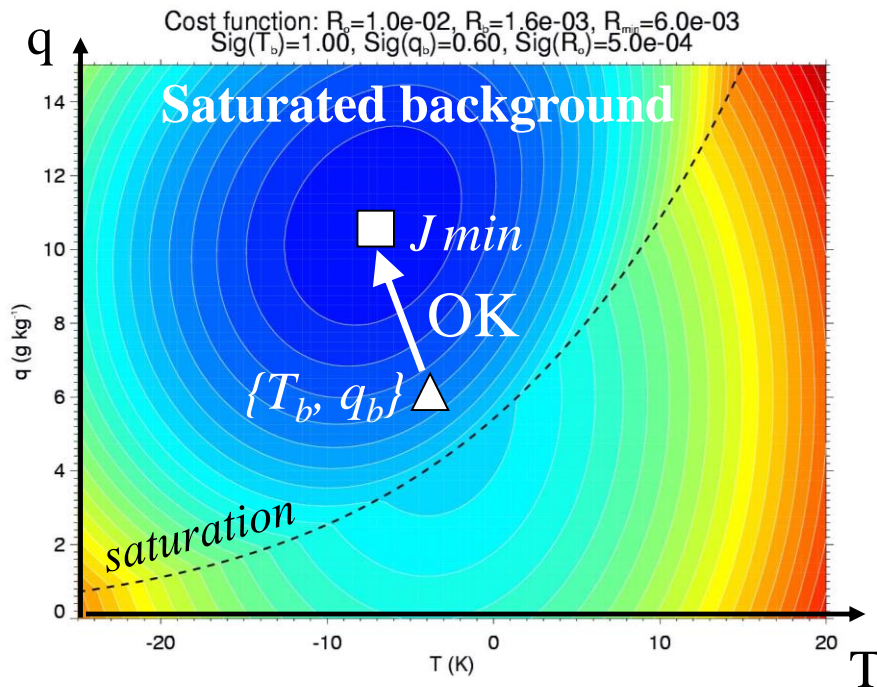| Tangent linear code	Trajectory and adjoint code

if $(x > x0)$ then
 	
 $\delta x = A \, \delta x \, / \, x$
 	
 $x = A \, Log(x)$
 	
 end if
	------------ Trajectory ---------------
	
	 $x_{store} = x$ (storage for use in adjoint)
	 if $(x > x0)$ then
	   $x = A \, Log(x)$
	 end if
	
	 -------------- Adjoint -----------------
	 if $(x_{store} > x0)$ then
	   $\delta x^* = A \, \delta x^* \, / \, x_{store}$
	 end if

Illustration of discontinuity effect on cost function shape:
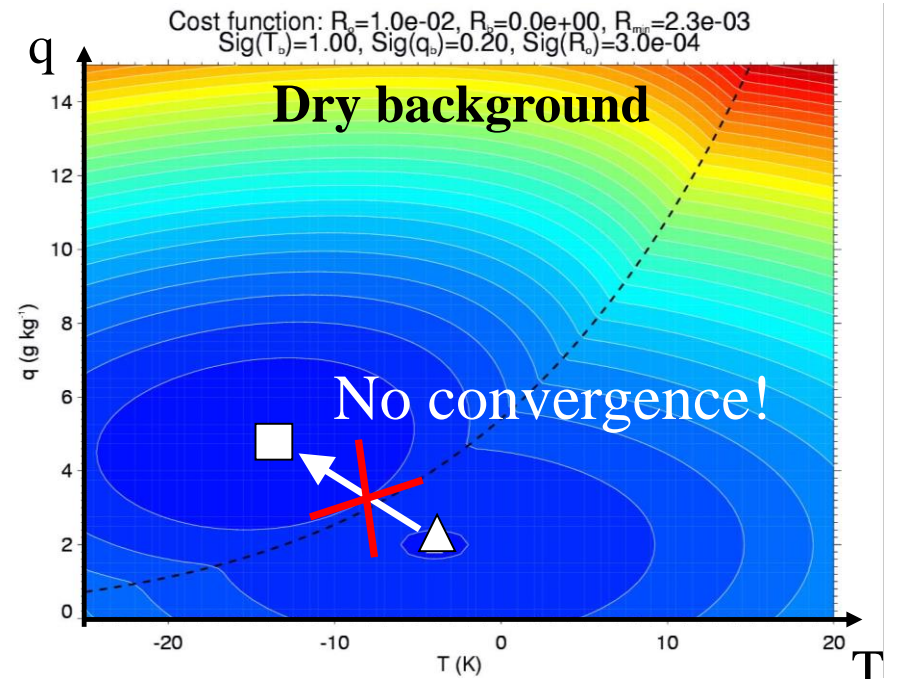Model background = $\{T_b, q_b\}$; Observation = RR_{obs}
Simple parametrization of rain rate:

$$RR = \begin{cases} \alpha\{q - q_{sat}(T)\} & \text{if } q > q_{sat}(T), \\ 0 & \text{otherwise} \end{cases}$$

$$J = \frac{1}{2}\left(\frac{T - T_b}{\sigma_T}\right)^2 + \frac{1}{2}\left(\frac{q - q_b}{\sigma_q}\right)^2 + \frac{1}{2}\left(\frac{\alpha[q - q_{sat}(T)] - RR_{obs}}{\sigma_{RR_{obs}}}\right)^2$$

Cost function: R₀=1.0e-02, Rₐ=1.6e-03, Rₘᵢₙ=6.0e-03
Sig(Tb)=1.00, Sig(qb)=0.60, Sig(Ro)=5.0e-04

Saturated background

$J\,min$

OK

$\{T_b, q_b\}$

saturation

Cost function: R₀=1.0e-02, Rₐ=0.0e+00, Rₘᵢₙ=2.3e-03
Sig(Tb)=1.00, Sig(qb)=0.20, Sig(Ro)=3.0e-04

Dry background

No convergence!

Single minimum of cost function   Several local minima of cost function

A short list of existing LP packages used in operational DA

- Tsuyuki (1996): Kuo-type convection and large-scale condensation schemes (FSU 4D-Var).

- Mahfouf (1999): full set of simplified physical parametrizations (gravity wave drag currently used in ECMWF operational 4D-Var and EPS).

- Janisková et al. (1999): full set of simplified physical parametrizations (Météo-France operational 4D-Var).

- Janisková et al. (2002): linearized radiation (ECMWF 4D-Var).

- Lopez (2002): simplified large-scale condensation and precipitation scheme (Météo-France).

- Tompkins and Janisková (2004): simplified large-scale condensation and precipitation scheme (ECMWF).

- Lopez and Moreau (2005): simplified mass-flux convection scheme (ECMWF).

- Mahfouf (2005): simplified Kuo-type convection scheme (Environment Canada).

1D-Var with TRMM/Precipitation Radar data

Tropical Cyclone Zoe (26 December 2002 @1200 UTC; Southwest Pacific)
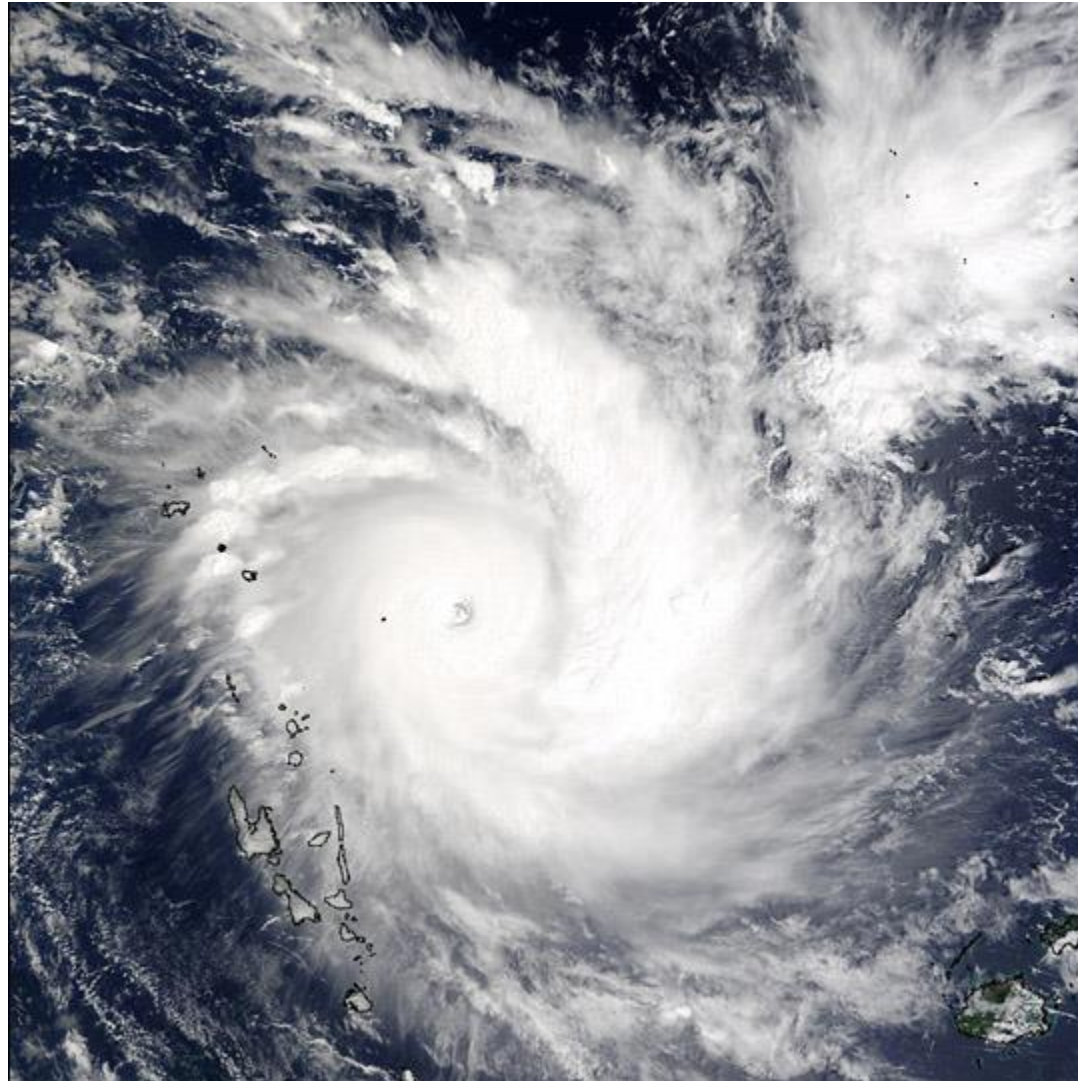
AQUA MODIS image

TRMM Precipitation Radar
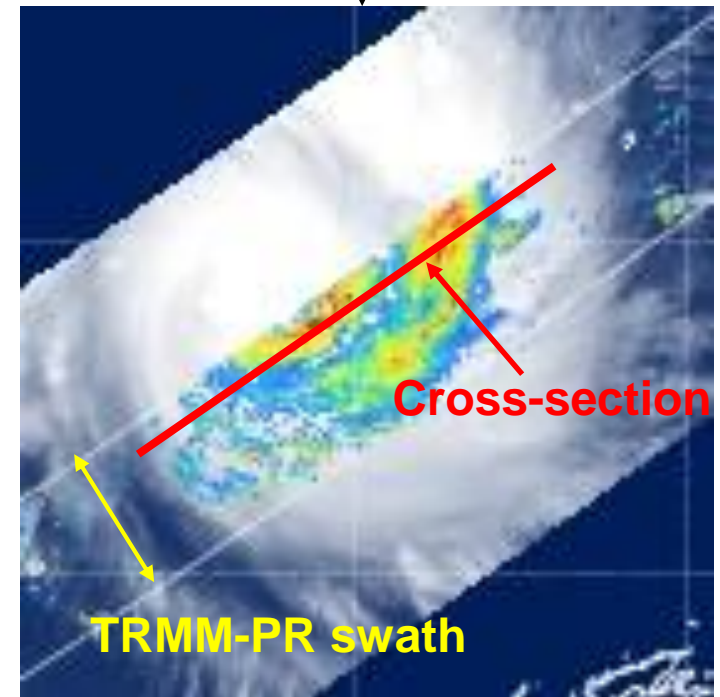
Cross-section

TRMM-PR swath

1D-Var with TRMM/Precipitation Radar data
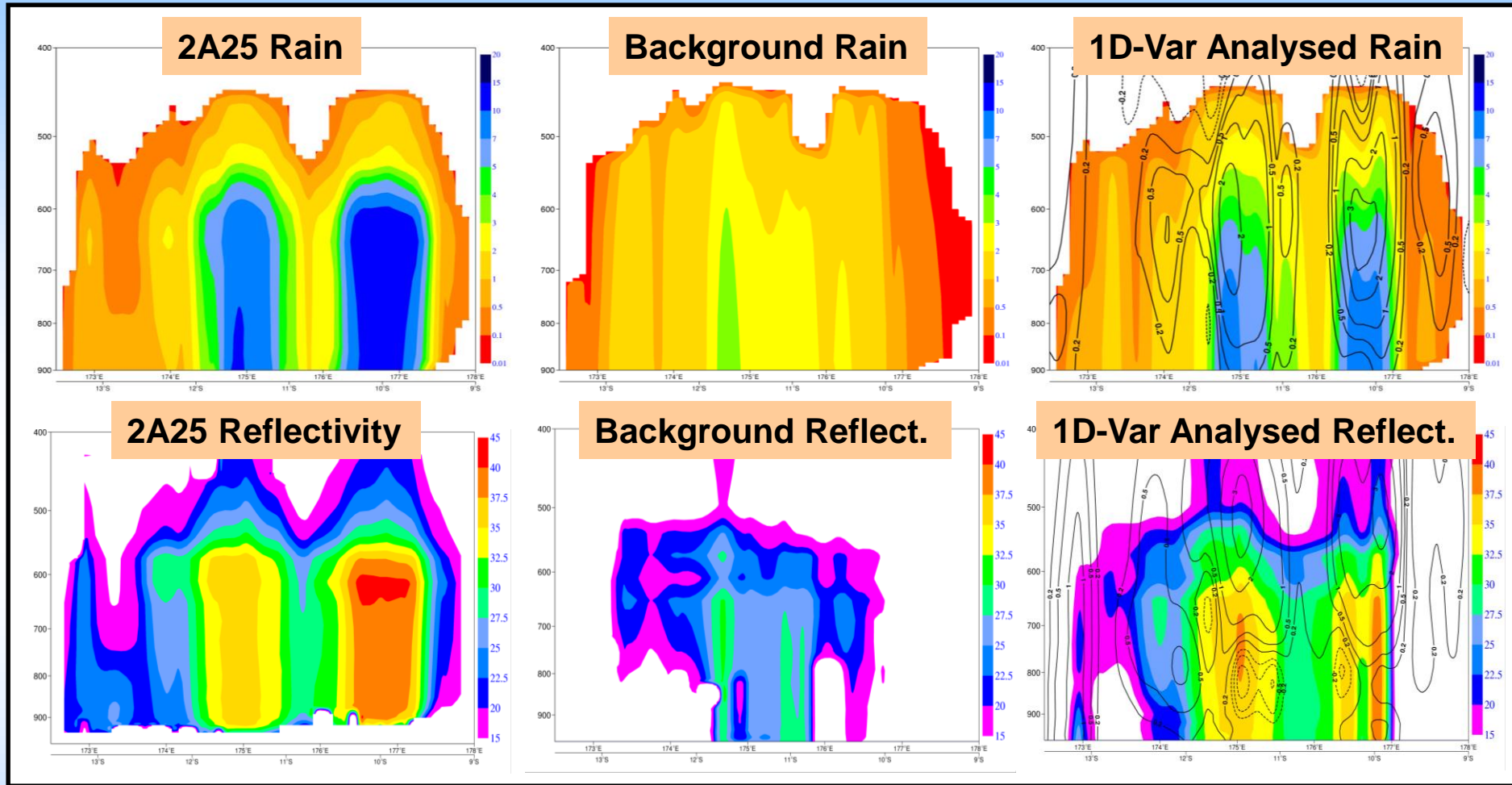
Tropical Cyclone Zoe (26 December 2002 @1200 UTC)

Vertical cross-section of rain rates (top, mm h^{-1}) and reflectivities (bottom, dBZ): observed (left), background (middle), and analyzed (right).

Black isolines on right panels = 1D-Var specific humidity increments.

Own impact of NCEP Stage IV hourly precipitation data over the U.S.A. (combined ground-based radar & rain gauge observations)

Three 4D-Var assimilation experiments (20 May - 15 June 2005):
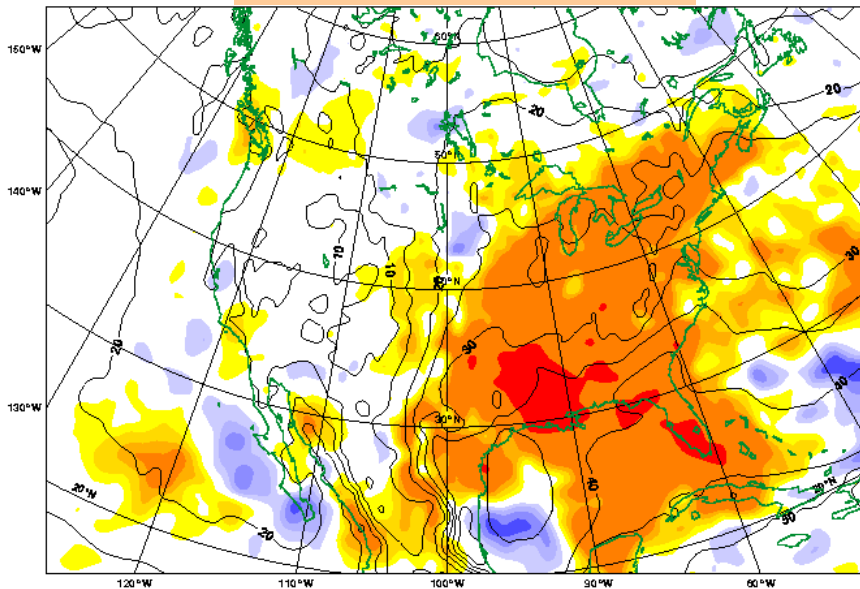
CTRL = all standard observations.

CTRL_noqUS = all obs except <u>no moisture obs over US</u> (surface & satellite).

NEW_noqUS = CTRL_noqUS + <u>NEXRAD hourly rain rates over US</u> ("1D+4D-Var").

Mean differences of TCWV analyses at 00UTC

CTRL_noqUS – CTRL   **NEW_noqUS – CTRL_noqUS**

Lopez and Bauer (*Monthly Weather Review*, 2007)