The challenges of Linearized Physics (LP) in Data Assimilation (DA)

Philippe Lopez, *ECMWF*

with thanks to Marta Janisková (ECMWF).

Outline

- **Brief reminder of what DA is about.**
- **Physics in (4D-Var) DA: Why, where and how?**
- **Constraints specific to LP.**
- **Benefits of LP.**
- **Summary and prospects.**

The purpose and ingredients of Data Assimilation

* **The purpose of DA** is to merge information coming from observations with *a priori* information coming from a forecast model to obtain an optimal 3D representation of the atmospheric state at a given time (= the "**analysis**").

This 3D analysis then provides initial conditions to the numerical forecast model.

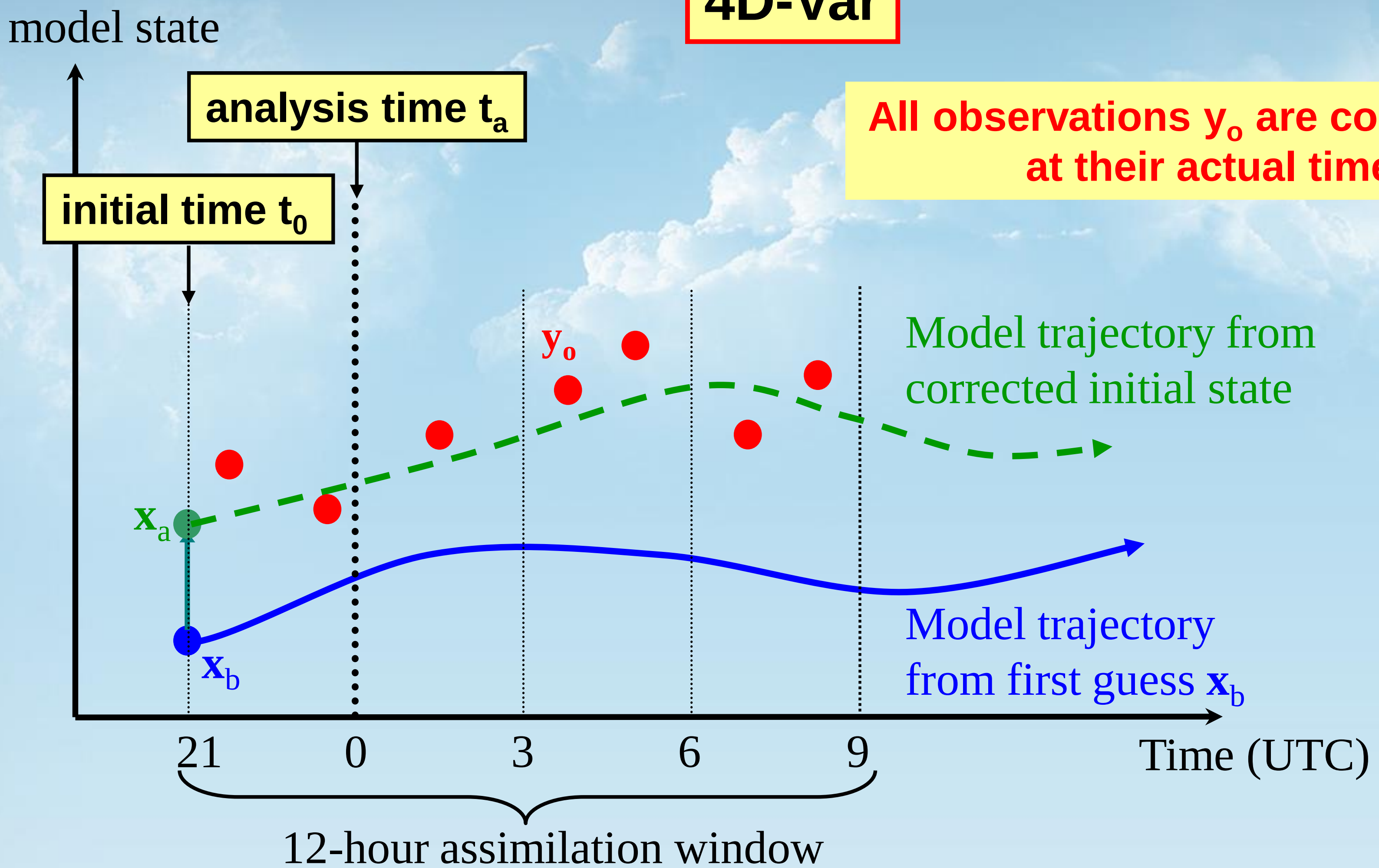
* **The main ingredients of DA are:**

- a set of **observations** available over a period of typically a few hours.
- a previous short-range forecast from the NWP model ("**background**" information).
- some statistical description of the **errors** of both observations and model background.
- a **data assimilation method** (e.g. variational DA: 3D-Var, 4D-Var; EnKF,…).

☞ **This presentation will focus on the use of physical parameterizations in 4D-Var DA.**

4D-Var

All observations y_o are considered at their actual time



4D-Var produces the analysis (x_a) which minimizes the distance to a set of available observations (y_o) and to some *a priori* background information from the model (x_b), given the respective errors of observations and model background.

ECMWF

4D-Var

Incremental 4D-Var aims at minimizing the following cost function:

$$J(\delta \mathbf{x}_0) = \frac{1}{2}\delta \mathbf{x}_0^T \mathbf{B}^{-1}\delta \mathbf{x}_0 + \frac{1}{2}\sum_{i=1}^{n}(\mathbf{H}\mathbf{M}_i\delta \mathbf{x}_0 - \mathbf{d}_i)^T \mathbf{R}_i^{-1}(\mathbf{H}\mathbf{M}_i\delta \mathbf{x}_0 - \mathbf{d}_i)$$

where: $\delta \mathbf{x}_0 = \mathbf{x}_0 - \mathbf{x}^b_0$ (increment; at lower resolution).

$\mathbf{B}$ = background error covariance matrix.

$\mathbf{d}_i = \mathbf{y}^o_i - H(M_i[\mathbf{x}^b_0])$ (innovation vector; at high resolution).

$\mathbf{M}_i$ = tangent-linear of forecast model ($t_0 \rightarrow t_i$).

$\mathbf{H}$ = tangent-linear of observation operator.

i = time index (4D-Var window is split in n intervals).

$\mathbf{R}_i$ = observation error covariance matrix.

$$\Rightarrow \quad \nabla_{\delta \mathbf{x}_0} J = \mathbf{B}^{-1}\delta \mathbf{x}_0 + \sum_{i=1}^{n} \mathbf{M}_i^T[t_i,t_0]\mathbf{H}^T \mathbf{R}_i^{-1}\left(\mathbf{H}\mathbf{M}_i\delta \mathbf{x}_0 - \mathbf{d}_i\right)$$

Adjoint of forecast model with <u>simplified</u> linearized physics
(simplified: to reduce computational cost and to avoid nonlinear processes)

In other words, physical parameterizations are used in 4D-Var DA:

* to describe the time evolution of the model state over the assimilation window as accurately as possible.

* to convert the model state variables (typ. T, wind, humidity, Psurf) into observed equivalents so that the differences obs–model can be computed (at the time of each observation).

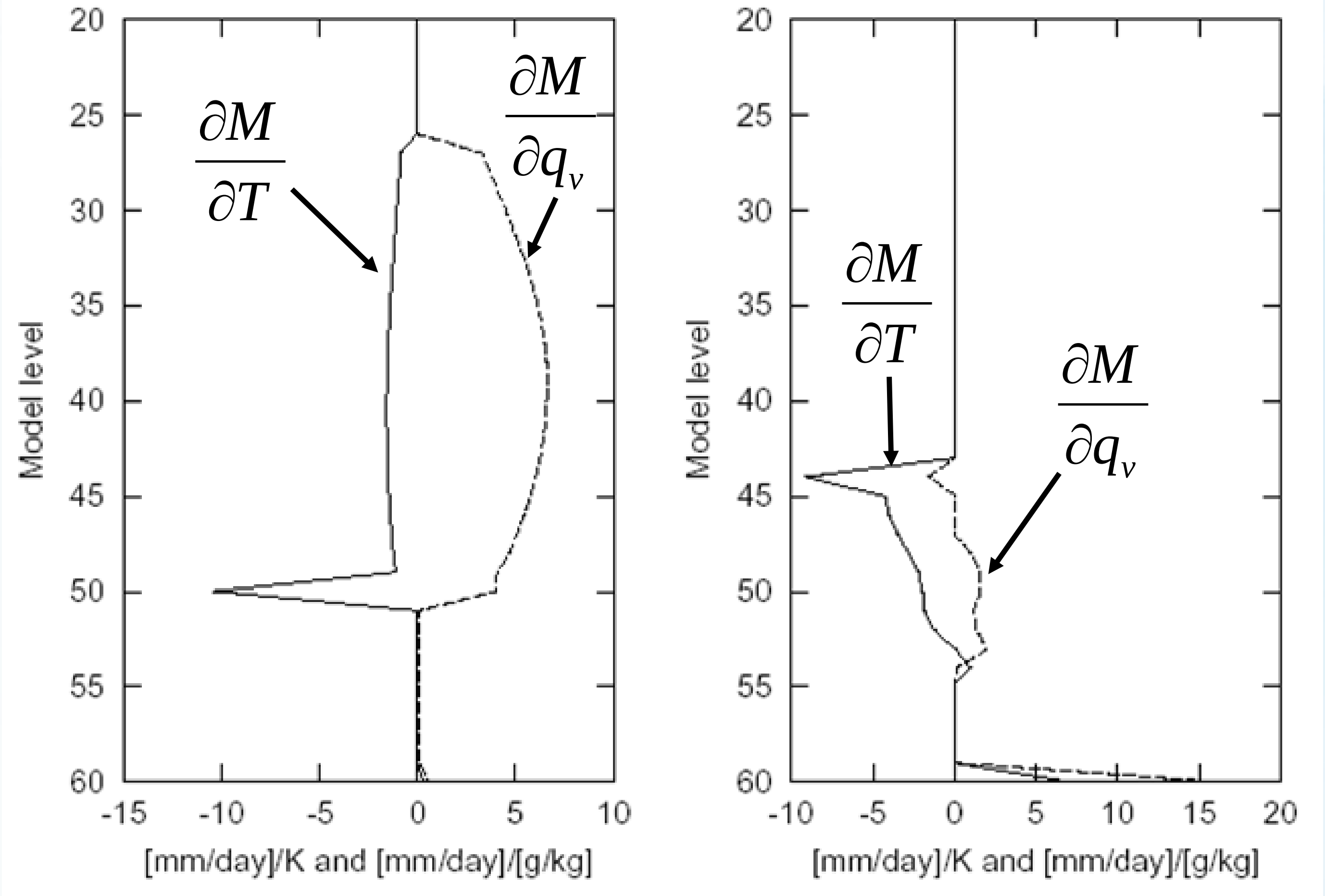
Side remark:

One advantage of the incremental approach is that the minimization of the cost function can be run at lower resolution than the trajectory computations (at ECMWF: 80 km versus 16 km).

The choice of physical parameterizations will affect the 4D-Var results.

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

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



from Marécal and Mahfouf (2002)

Betts-Miller (adjustment scheme)

Tiedtke (ECMWF's oper mass-flux scheme)

Testing the tangent-linear code

The correctness of the tangent-linear model 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	
Tiny perturbations	0.1E-09	0.99948758815435574E+00	**Machine precision reached**
	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	0.9999996111338369E+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	
Larger perturbations	0.1E+01	0.9683066504549524E+00	

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:

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

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

Example of output from a successful adjoint test:

$$<\mathbf{M}\,\delta\mathbf{x}, \quad \delta\mathbf{y}> = -0.13765102625164\text{E-}01$$

$$< \quad \delta\mathbf{x}, \mathbf{M}^T\delta\mathbf{y}> = -0.13765102625168\text{E-}01$$

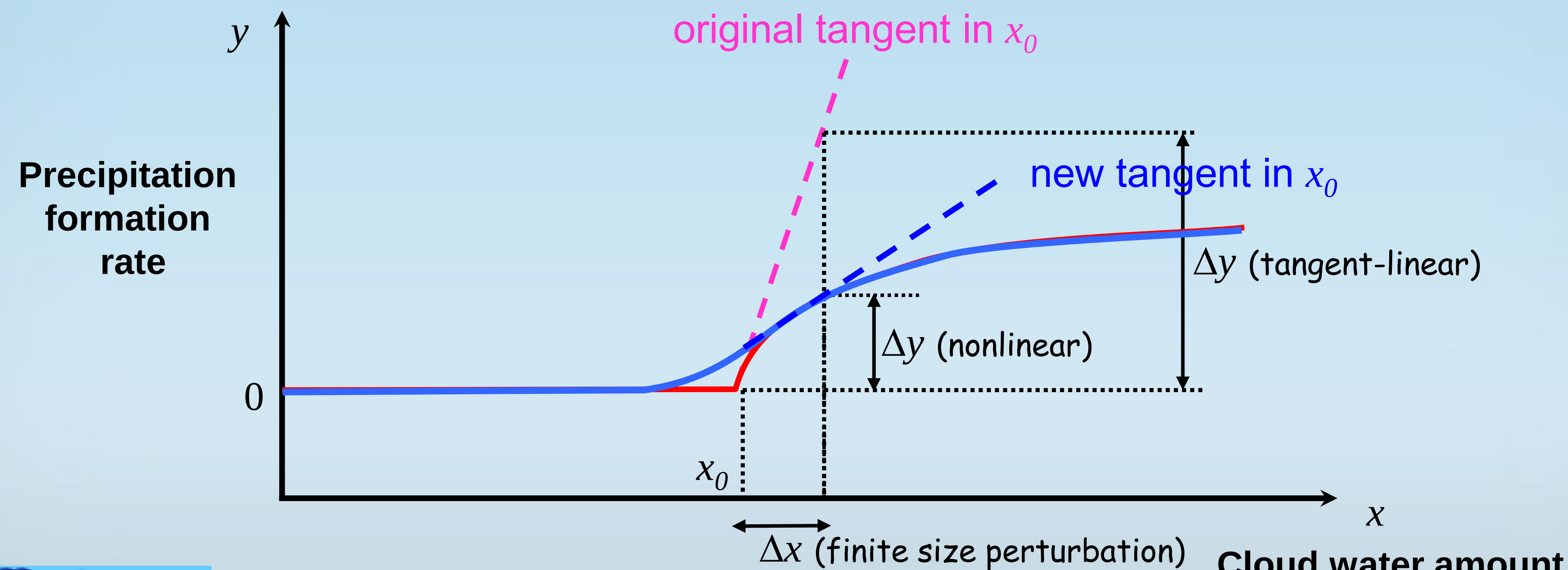
The difference is 11.351 times the zero of the machine

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!

Linearity assumption

- Variational assimilation is based on the strong assumption that the analysis is performed in a **(quasi-)linear** framework.
- However, physical processes are often nonlinear and their parameterizations involve discontinuities or non-differentiable functions (e.g. **switches or thresholds**). The trickiest parameterizations are convection, large-scale cloud processes and vertical diffusion.

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

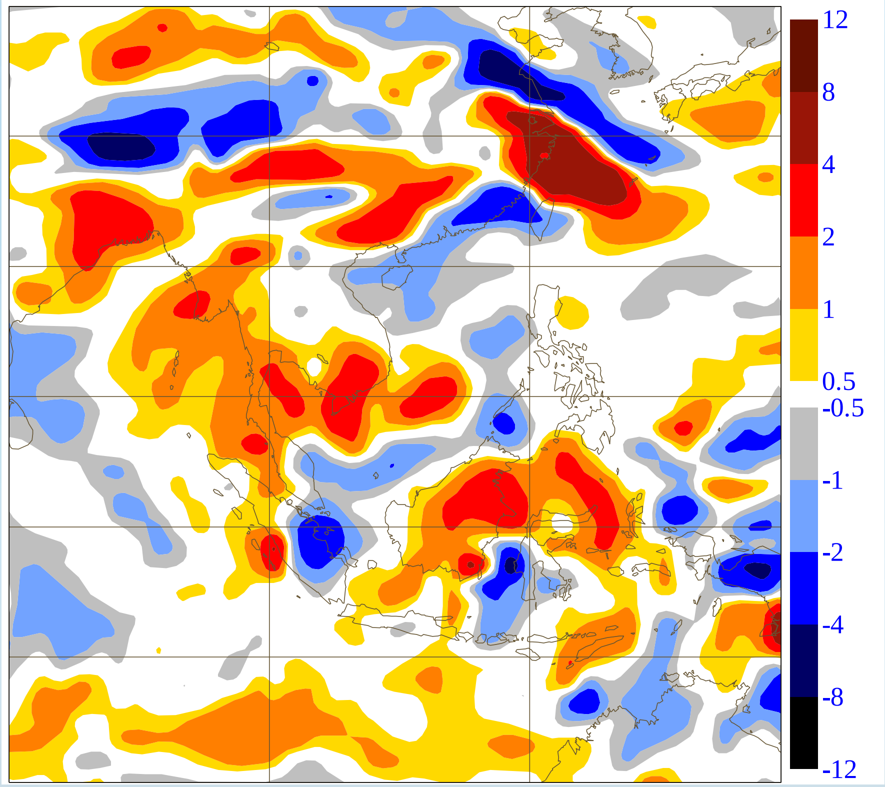


ECMWF

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

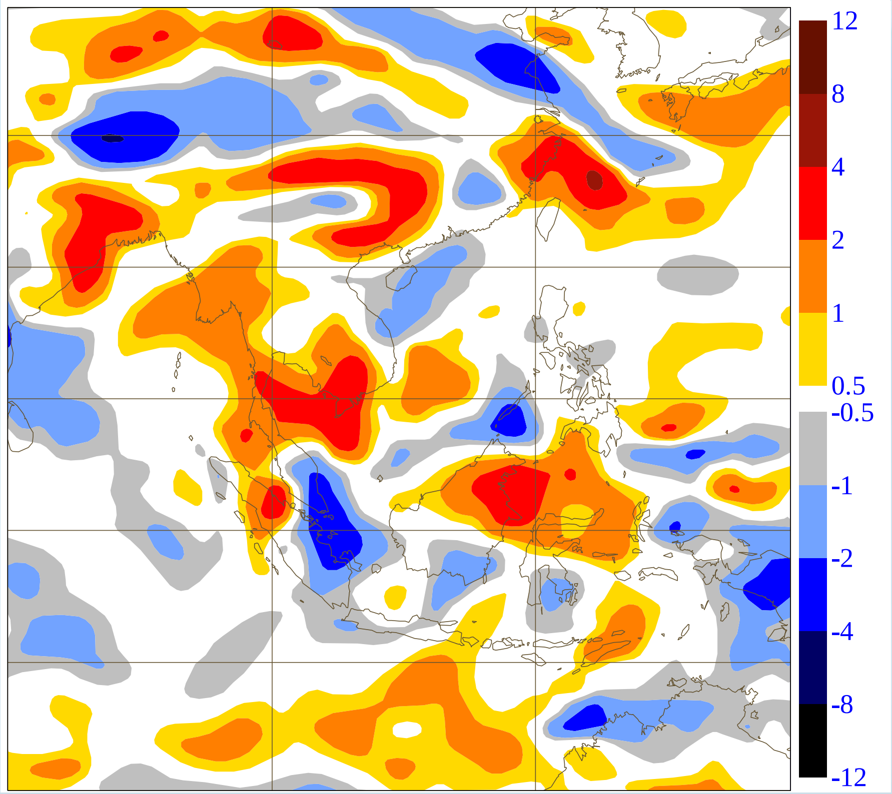
~700 hPa zonal wind increments [m/s] from 12h model integration.

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



from M. Janisková

Tangent-linear integration: $\mathbf{M}\delta\mathbf{x}$



with perturbation reduction in autoconversion

Linearized physics package used in ECMWF's operational 4D-Var (1)

Main simplifications/regularizations with respect to full nonlinear model 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 nonlinear version).
  - **longwave:** based on *Morcrette (1989)*, **called every 2 hours only.**

Linearized physics package used in ECMWF's operational 4D-Var (2)

- Vertical diffusion: [*Janisková and Lopez 2013]*
  - mixing in the surface and planetary boundary layers.
  - based on K-theory and Blackadar mixing length.
  - exchange coefficients based on: *Louis et al. [1982]* in stable conditions,
    Monin-Obukhov in unstable conditions.
  - mixed-layer parametrization and PBL top entrainment.
  - **Perturbations of exchange coefficients are smoothed (esp. near the surface).**

- Surface scheme: [*Janisková 2014, pers. comm.]*
  - evolution of soil, snow and sea-ice temperature.
  - **perturbation reduction in snow phase change formulation.**

- 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.**

Impact of linearized physics on TL approximation (1)

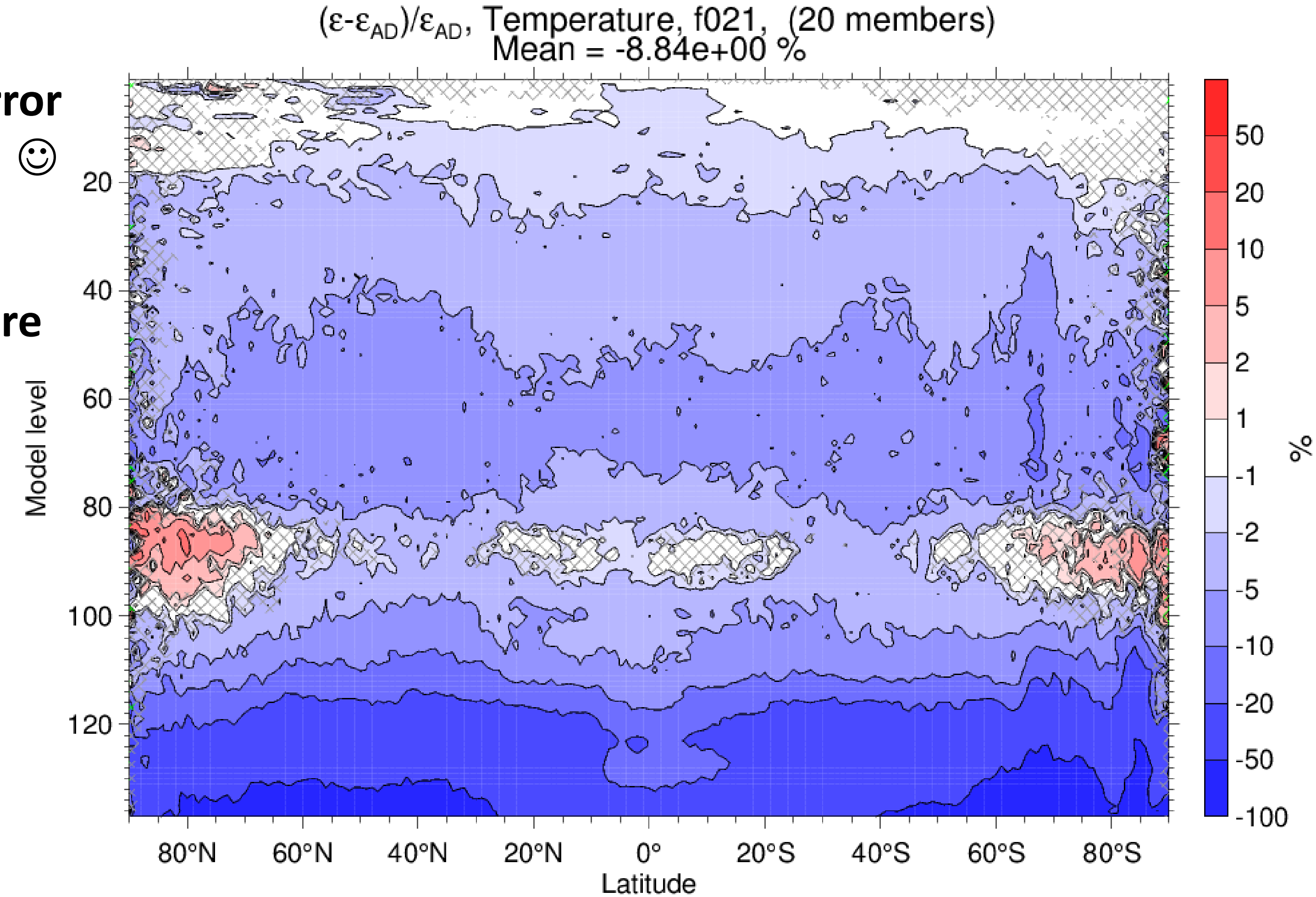
Zonal mean cross-section of change in TL error when TL includes:

VDIF + orog. GWD + SURF

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

Blue = TL error reduction = ☺

Temperature



Impact of linearized physics on TL approximation (2)

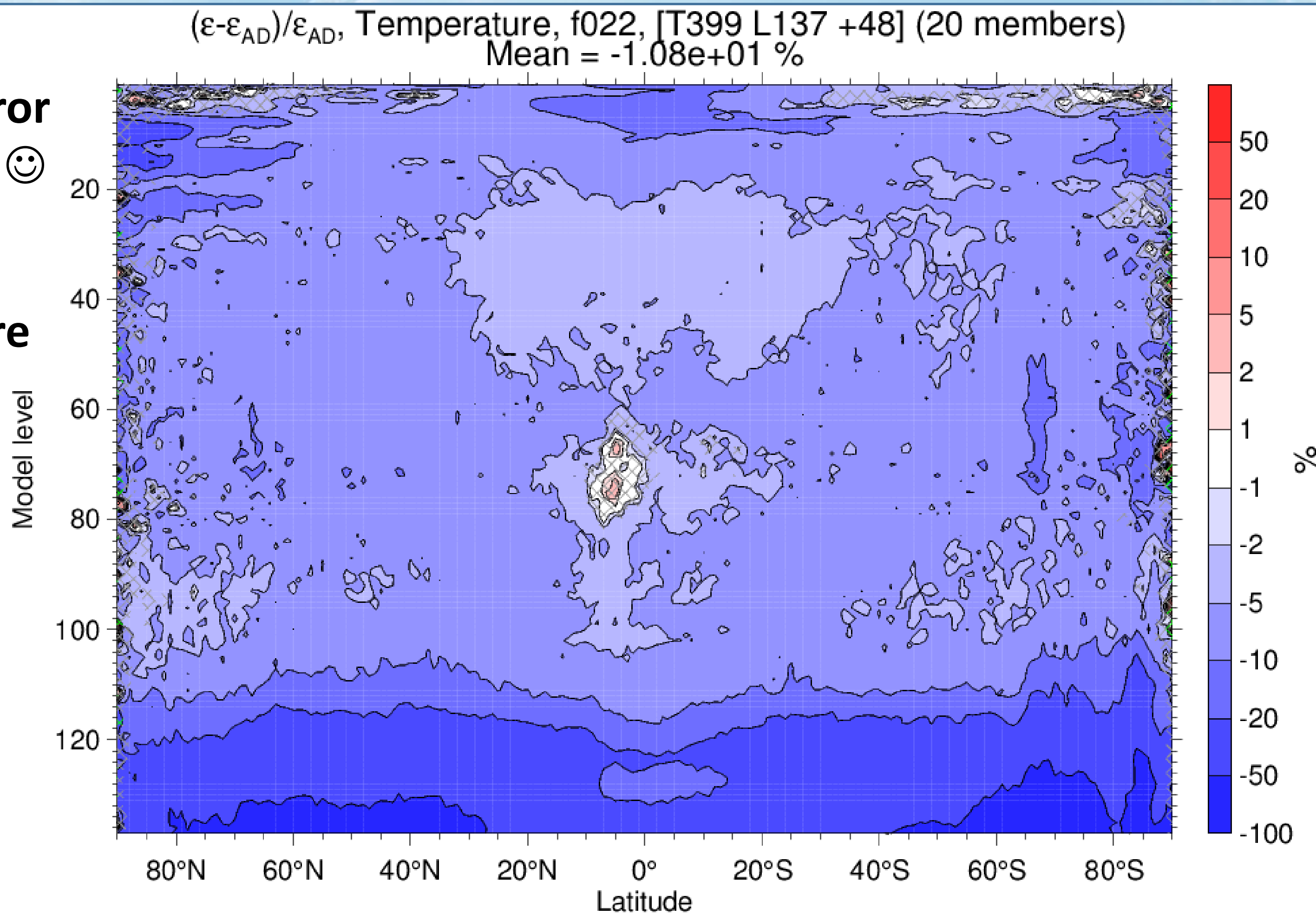
Zonal mean cross-section of change in TL error when TL includes:

VDIF + orog. GWD + SURF + RAD

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

Blue = TL error reduction = ☺

Temperature



Impact of linearized physics on TL approximation (3)

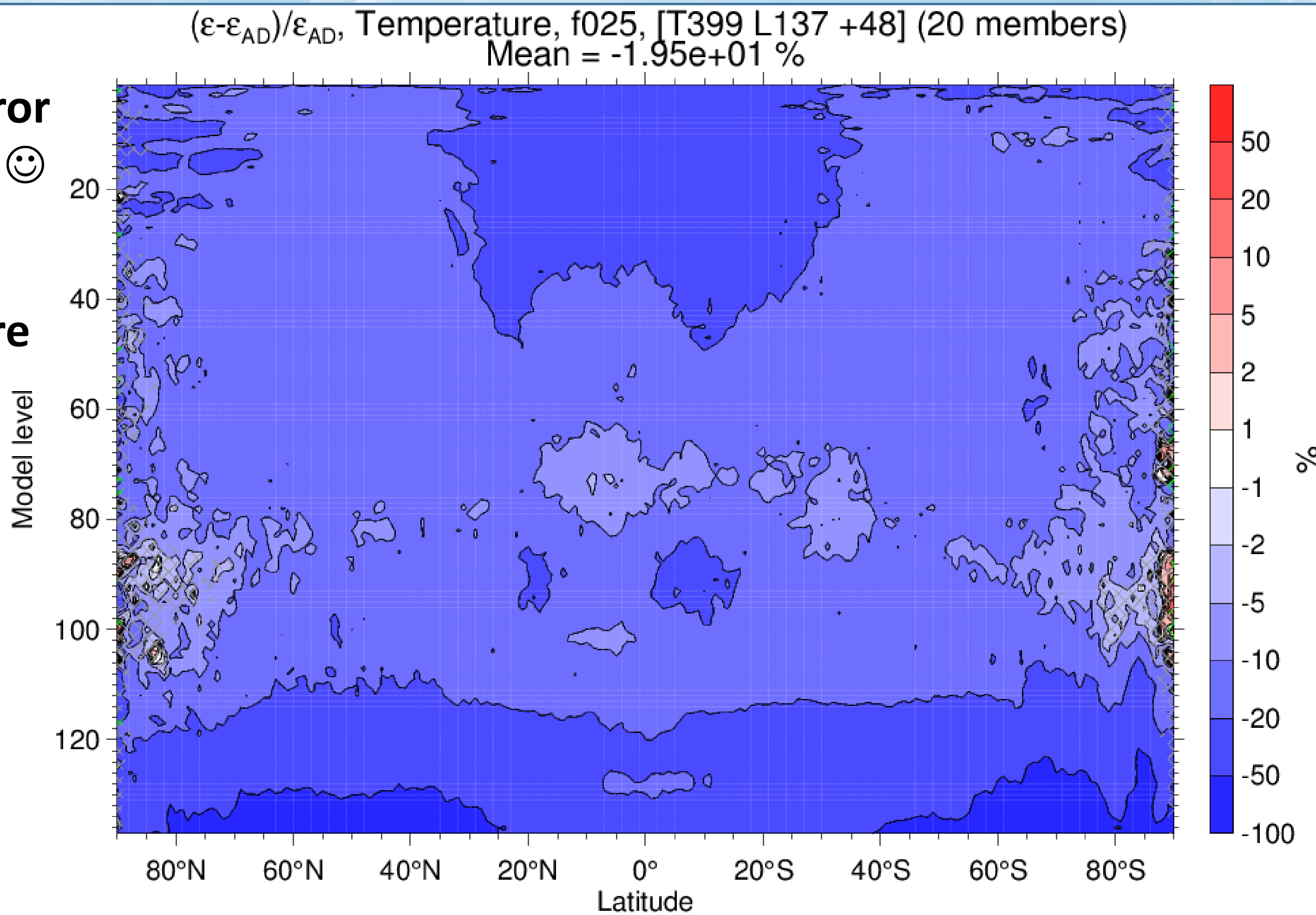
Zonal mean cross-section of change in TL error when TL includes:

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

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

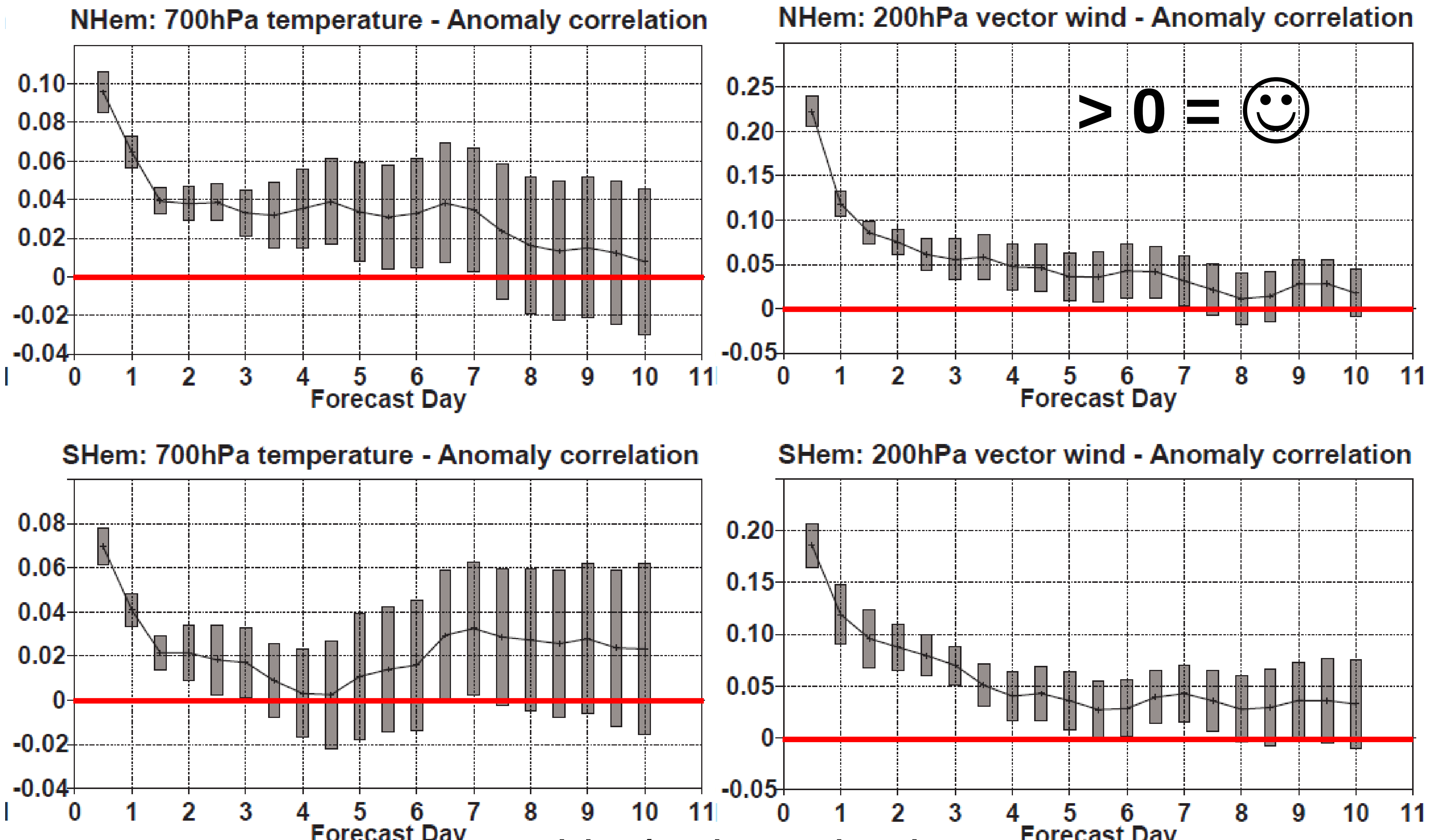
Blue = TL error reduction = ☺

Temperature



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.



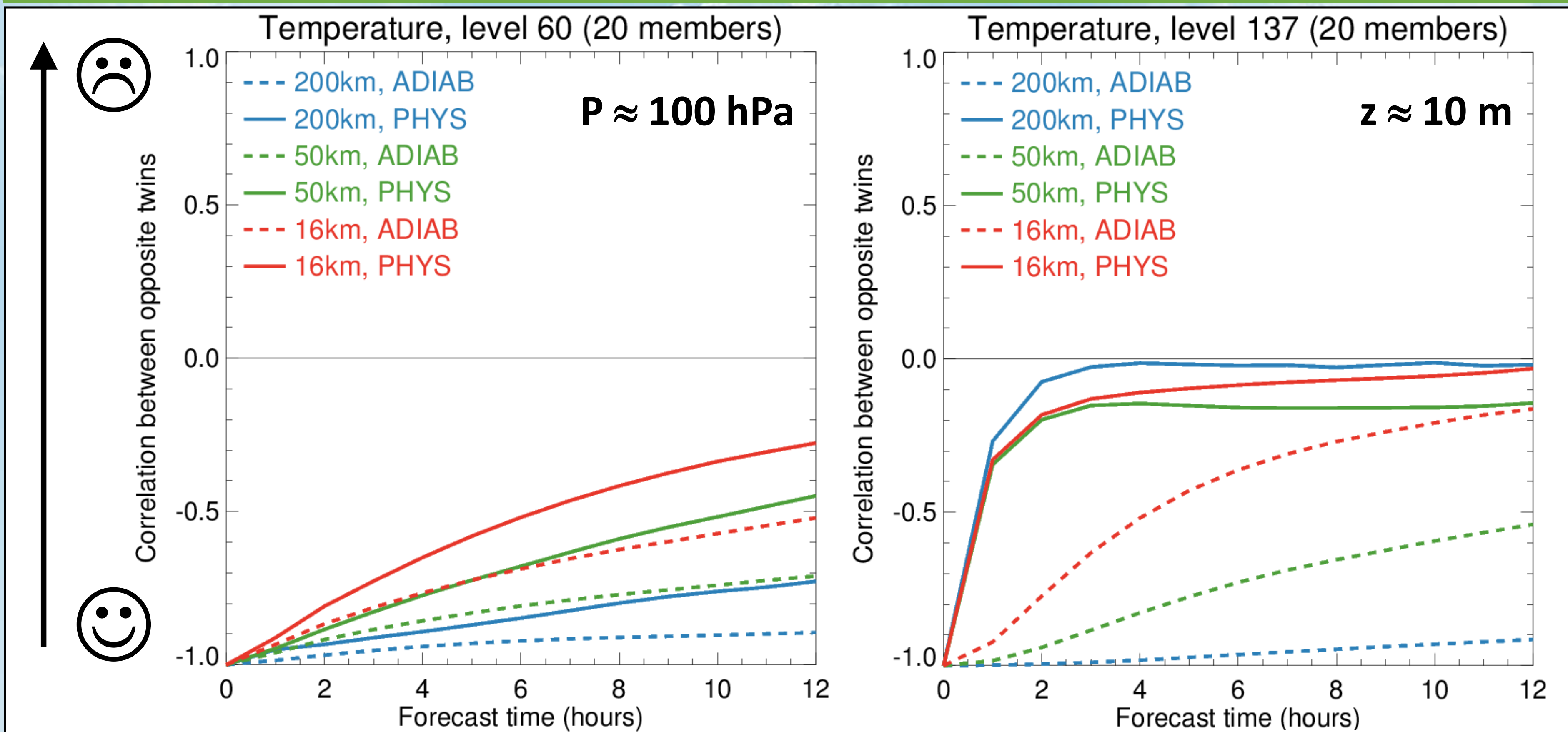
Janisková and Lopez (2013)

ECMWF

Influence of integration length and resolution on linearity assumption

Comparison of pairs of “opposite twin” experiments using ECMWF’s <u>nonlinear model</u>.
Time evolution of correlation between $M(\mathbf{x}+\delta\mathbf{x})-M(\mathbf{x})$ and $M(\mathbf{x}-\delta\mathbf{x})-M(\mathbf{x})$.

Inspired from Walser et al. (2004).



→The linearity assumption becomes less valid with integration time and resolution, especially close to the surface and when physics is activated.

Summary and prospects (1)

- Linearized physical parameterizations have become essential components of variational data assimilation systems:

→ 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:

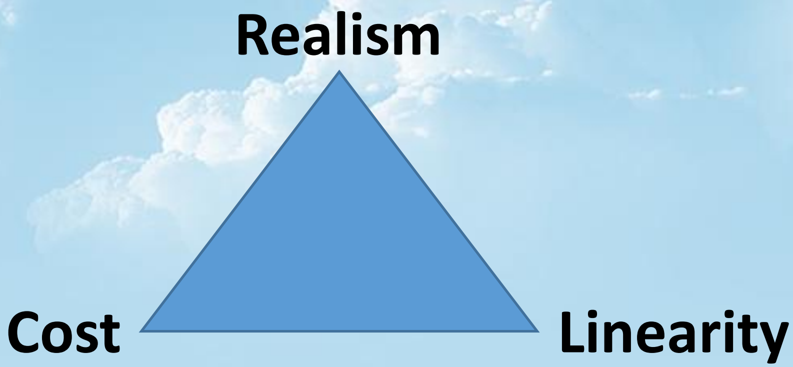
1) <u>Theoretical:</u>
The domain of validity of the linear hypothesis shrinks with increasing resolution and integration length.

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.

Summary and prospects (2)

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



- 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…
na Wildlife Park / Barcroft I

Summary and prospects (3)

- **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).**

?

Simple example of tangent-linear and adjoint codes.

- <u>**Simplified nonlinear code:**</u> $Z = a / X^2 + b\,Y \log(W)$

- <u>**Tangent-linear code:**</u> $\delta Z = -(2\,a / X^3)\,\delta X + b \log(W)\,\delta Y + (b\,Y / W)\,\delta W$

- <u>**Adjoint code:**</u>

$$\delta X^* = 0$$
$$\delta Y^* = 0$$
$$\delta W^* = 0$$

$$\delta X^* = \delta X^* - (2\,a / X^3)\,\delta Z^*$$
$$\delta Y^* = \delta Y^* + b \log(W)\,\delta Z^*$$
$$\delta W = \delta W^* + (b\,Y / W)\,\delta Z^*$$

$$\delta Z^* = 0$$

Example of nonlinear, tangent-linear and adjoint operators in the context of 4D-Var.

<u>Nonlinear operators are applied to full fields:</u>

$$\text{model initial state } \mathbf{x}_0 = \underset{\text{time } t_0}{\begin{bmatrix} T \\ q_v \\ u \\ v \\ P_s \end{bmatrix}} \xrightarrow{\;M\;} \underset{\text{time } t_i}{\begin{bmatrix} T \\ q_v \\ u \\ v \\ P_s \\ q_{liq} \\ q_{ice} \end{bmatrix}} \xrightarrow{\;H\;} \tilde{\mathbf{y}}_i = \begin{bmatrix} Rad_{ch1} \\ Rad_{ch2} \\ Rad_{ch3} \end{bmatrix}$$

observation equivalent = satellite cloudy radiances time t_i

M = forecast model with physics

H = radiative transfer model

Tangent-linear operators are applied to perturbations:

$$\text{time } t_0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{time } t_i$$

$$\boldsymbol{\delta}\mathbf{x}_0 = \begin{bmatrix} \delta T \\ \delta q_v \\ \delta u \\ \delta v \\ \delta P_s \end{bmatrix} \xrightarrow{\mathbf{M}[t_0,t_i]} \begin{bmatrix} \delta T \\ \delta q_v \\ \delta u \\ \delta v \\ \delta P_s \\ \delta q_{liq} \\ \delta q_{ice} \end{bmatrix} \xrightarrow{\mathbf{H}} \boldsymbol{\delta}\tilde{\mathbf{y}}_i = \begin{bmatrix} \delta Rad_{ch1} \\ \delta Rad_{ch2} \\ \delta Rad_{ch3} \end{bmatrix}$$

Adjoint operators are applied to the cost function gradient:

$$\text{time } t_i \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{time } t_0$$

$$\nabla_{\tilde{\mathbf{y}}_i} J_o = \begin{bmatrix} \partial J_o / \partial Rad_{ch1} \\ \partial J_o / \partial Rad_{ch2} \\ \partial J_o / \partial Rad_{ch3} \end{bmatrix} \xrightarrow{\mathbf{H}^T} \begin{bmatrix} \partial J_o / \partial T \\ \partial J_o / \partial q_v \\ \partial J_o / \partial u \\ \partial J_o / \partial v \\ \partial J_o / \partial P_s \\ \partial J_o / \partial q_{liq} \\ \partial J_o / \partial q_{ice} \end{bmatrix} \xrightarrow{\mathbf{M}^T[t_i,t_0]} \nabla_{\mathbf{x}_0} J_o = \begin{bmatrix} \partial J_o / \partial T \\ \partial J_o / \partial q_v \\ \partial J_o / \partial u \\ \partial J_o / \partial v \\ \partial J_o / \partial P_s \end{bmatrix}$$

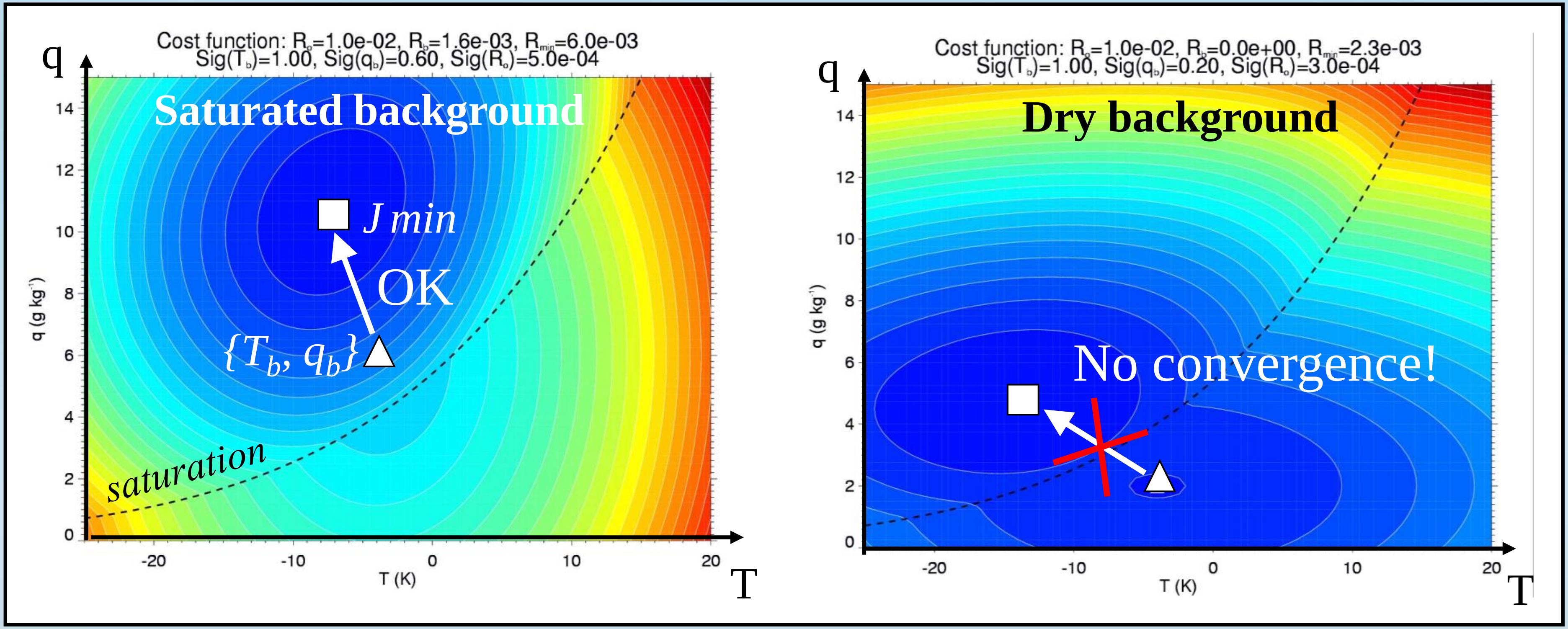
Illustration of discontinuity effect on cost function shape:

Model background = $\{T_b, q_b\}$; Observation = RR_{obs}

Simple parametrization of rain rate:

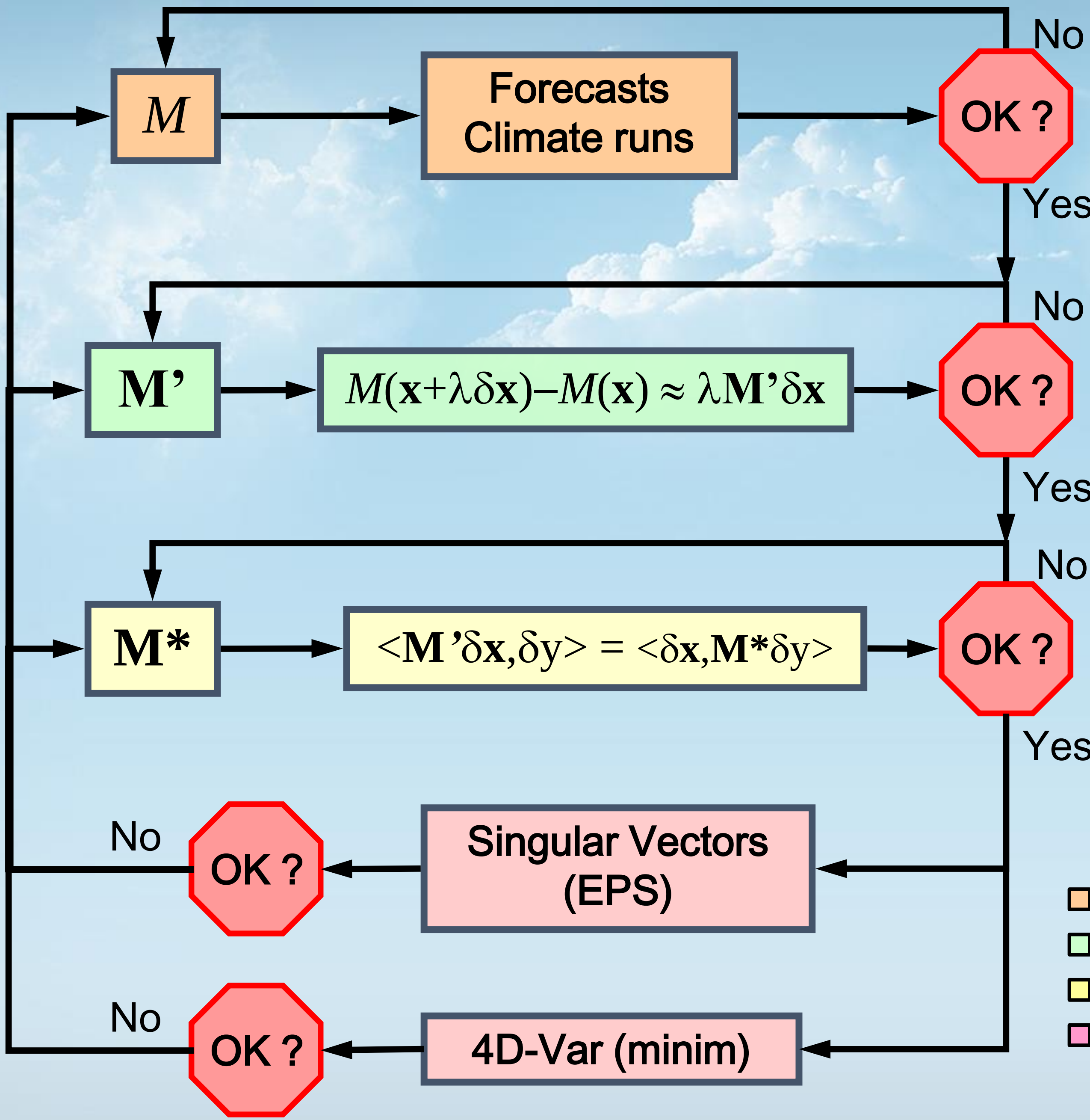
$$RR = \begin{cases} \alpha\{q - q_{sat}(T)\} & \text{if } q > q_{sat}(T), \\ 0 & \text{otherwise} \end{cases}$$

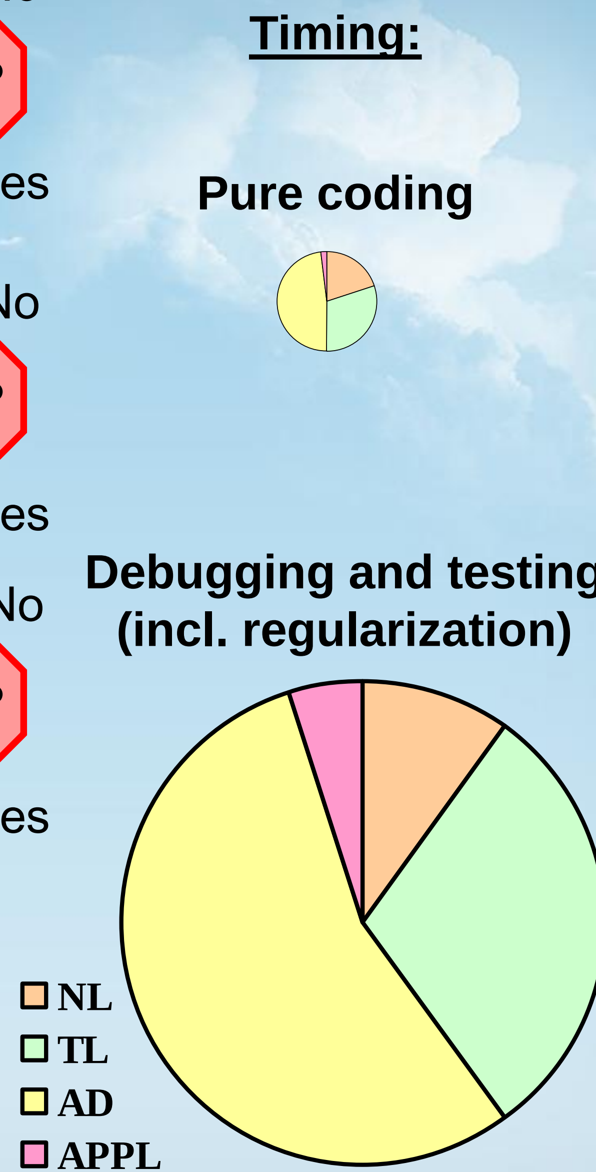
$$J = \frac{1}{2}\left(\frac{T - T_b}{\sigma_T}\right)^2 + \frac{1}{2}\left(\frac{q - q_b}{\sigma_q}\right)^2 + \frac{1}{2}\left(\frac{\alpha[q - q_{sat}(T)] - RR_{obs}}{\sigma_{RR_{obs}}}\right)^2$$



Single minimum of cost function

Several local minima of cost function





Stages in the development of a new linearized parameterization

Tangent-linear approximation and associated error

- **The tangent-linear approximation is assessed by comparing the difference between two integrations of the full non-linear model:**

$$M(\mathbf{x}+\delta\mathbf{x}) - M(\mathbf{x})$$

with an integration of the tangent-linear model from the same initial perturbation ($\delta\mathbf{x} = \mathbf{x}_{\mathbf{ana}} - \mathbf{x}_{\mathbf{bg}}$):

$$\mathbf{M}\delta\mathbf{x}$$

- **The TL error is then defined as:**

$$\varepsilon = \langle |\, M(\mathbf{x}+\delta\mathbf{x}) - M(\mathbf{x}) - \mathbf{M}\delta\mathbf{x}\, | \rangle$$

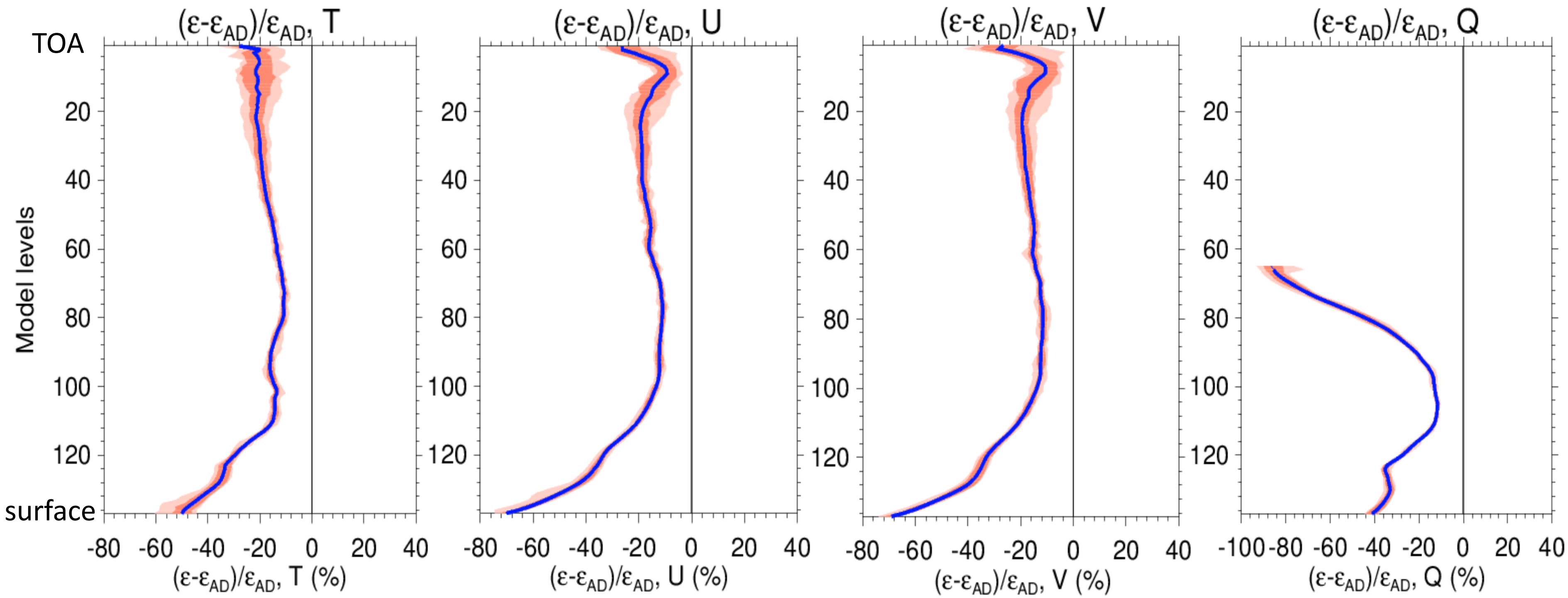
where $\langle \ldots \rangle$ denotes spatial averaging (e.g. zonally or globally).

Impact of linearized physics on TL approximation (4)

Mean vertical profile of change in TL error for T, U, V and Q when full linearized physics is included in TL computations.

Relative to adiabatic TL run (50-km resol.; twenty runs, 12h integ.)

< 0 = ☺

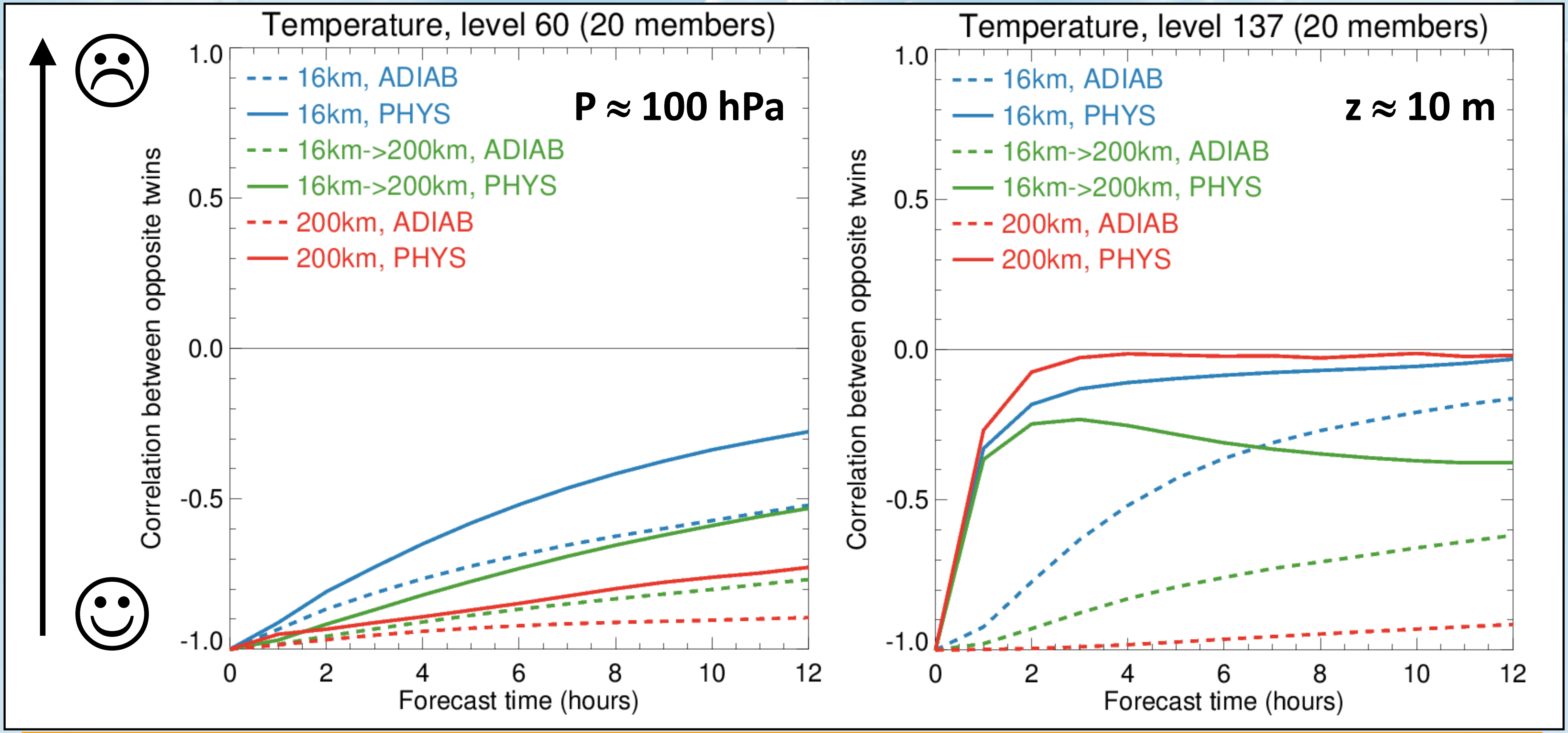


Inclusion of linearized physics leads to better TL approximation.

Effects of model nonlinearities versus effects of resolution

Comparison of pairs of "opposite twin" experiments using ECMWF's nonlinear model. Time evolution of correlation between $M(\mathbf{x}+\delta\mathbf{x})-M(\mathbf{x})$ and $M(\mathbf{x}-\delta\mathbf{x})-M(\mathbf{x})$.

Inspired from Walser et al. (2004).



→The effects of nonlinearities are of comparable magnitude with those of resolution, except near the surface when physics is activated (the former effects then dominate).