Tangent-linear and adjoint models in data assimilation

Marta Janisková and Philippe Lopez

ECMWF

Thanks to: F. Váňa, M.Fielding

2018 Annual Seminar: Earth system assimilation

10 - 13 September 2018

- *Introduction*
- *Validity of the linearized model*
- *Applications of the linearized model*
- *Summary and prospects*

- **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:

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

- **Adjoint model**

The adjoint of a linear operator M' is the linear operator M^* such that, for the inner product <,> :

$$\forall\mathbf{x}, \forall\mathbf{y} \qquad \langle M'\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{M}^*\mathbf{y}\rangle$$

For the inner product in the Euclidean space:

$$\mathbf{M}^* = M'^T$$

Linearized models in NWP

- **Different well-known applications:**
  - **variational data assimilation** ⟵ like incremental 4D-Var
  - **singular vector computations** ⟵ initial perturbations for EPS
  - **sensitivity analysis** ⟵ forecast errors

- **First applications with adiabatic linearized model**

- **Nowadays, the physical processes included in the linearized model**

Including physical processes can in variational data assimilation:

- reduce spin-up
- provide a better agreement between the model and data
- produce an initial atmospheric state more consistent with physical processes
- allow the use of new observations *(rain, clouds, soil moisture, …)*

4D-Var – Four-dimensional Variational Data Assimilation
EPS – Ensemble Prediction System

Simplifications of the linearized models for practical applications

- **For important applications:**
  - incremental 4D-Var (ECMWF, Météo-*France, …),*
  - simplified gradients in 4D-Var (Zupanski 1993),
  - the initial perturbations computed for EPS (ECMWF),

linearized versions of forecast models are run at lower resolution

⇓

the linear model may not be "the exact tangent" to the full model

(different resolution and geometry, <u>different physics</u>)

⇓

simplified approaches as a way to include physical processes step-by-step in TL and AD models

- **simplifications done with the aim to have a physical package:**
  - **simple** – for the linearization of the model equations
  - **regular** – to avoid strong non-linearities and thresholds
  - **realistic enough**
  - **computationally affordable**

TL – tangent linear
AD – adjoint

General problems with adjoint models and including physics into them

- **Development** – requires substantial resources

- **Validation** – must be very thorough

  (for non-linear, tangent-linear and adjoint versions)

- **Computational cost** – may be very high when including physics or complex observation operators

- **Non-linear and discontinuous nature** of physical processes

  (affecting the range of validity of the tangent-linear approximation)

Operational constraints

Imply:

- **permanent testing of the validity of TL approximation and necessary adjustments:**
  - when the NL physics or dynamics changes significantly
  - higher horizontal and vertical resolutions, longer time-integrations

- **ensure robust stability of the linearized model:**
  - non-noisy behaviour in all situations and different model resolutions

- **code optimizations to reduce computational cost:**
  - ideally: TL is 2 times and AD is 2-3 times more expensive than the nonlinear model

- **fulfilling requirements for assimilation of observations related to the physical processes** (rain, clouds, soil moisture, ...)**:**

↓

finding best compromise between **complexity**, **linearity** and **cost**

Validity of the linearized model

Validation of tangent-linear and adjoint models

Tangent-linear (TL) and adjoint (AD) model:

- classical validation (TL - Taylor formula, AD - test of adjoint identity)
- examination of the accuracy of the linearization

Comparison:

finite differences (FD) $\leftrightarrow$ tangent-linear (TL) integration

$$\boldsymbol{M}(\mathbf{x}_{an}) - \boldsymbol{M}(\mathbf{x}_{fg}) \quad \leftrightarrow \quad \boldsymbol{M}'(\mathbf{x}_{an} - \mathbf{x}_{fg})$$

$$(an = analysis, \quad fg = first\ guess)$$

Diagnostics:

- relative errors: $\frac{\varepsilon_{\text{EXP}} - \varepsilon_{\text{REF}}}{\varepsilon_{\text{REF}}} .100\%$

  where

- mean absolute errors:

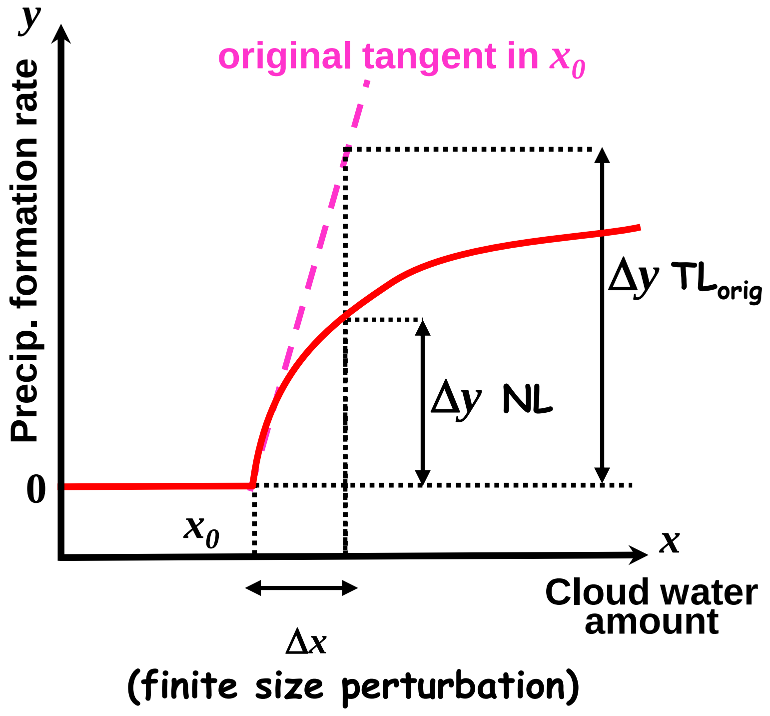
$$\varepsilon = \left| \left[M(\mathbf{x}_{an}) - M(\mathbf{x}_{fg}) \right] - M'(\mathbf{x}_{an} - \mathbf{x}_{fg}) \right|$$

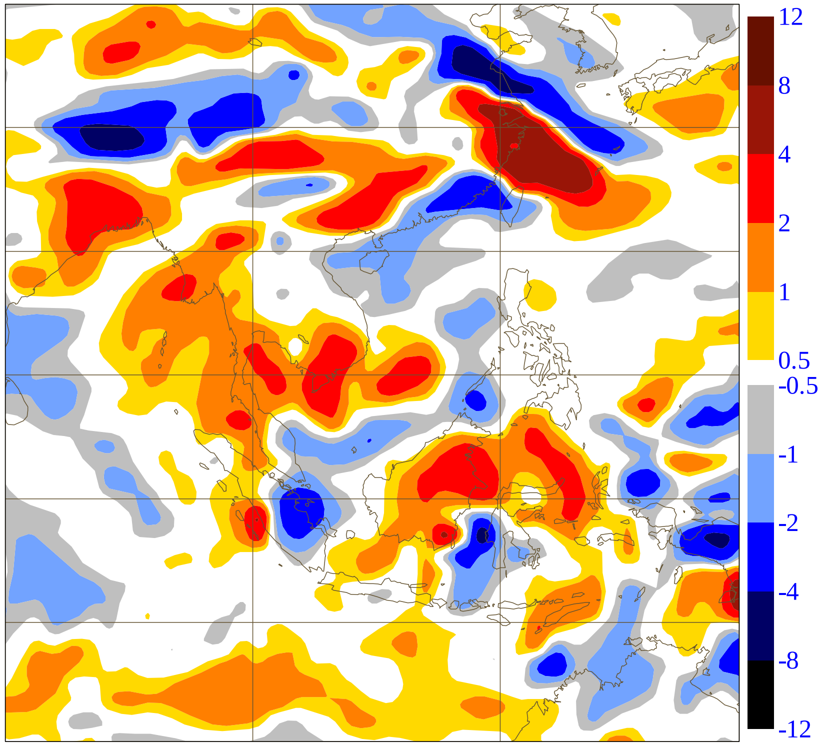
Singular vectors:

- Computation of singular vectors to find out whether the new schemes do not produce spurious unstable modes.

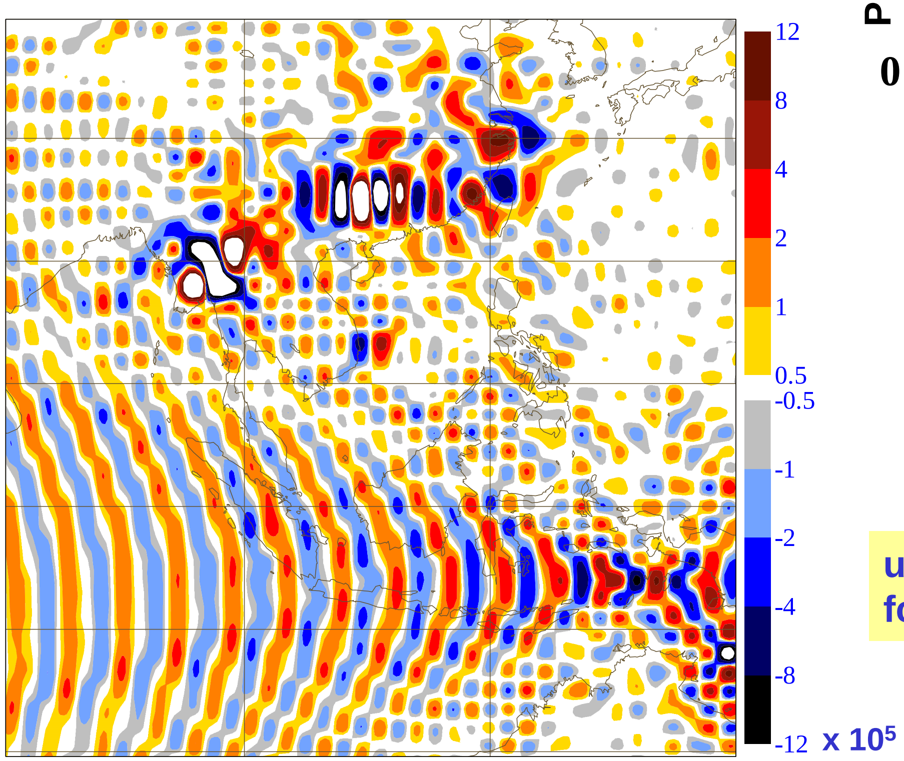
Importance of the regularization of TL model (1)

- **physical processes are characterized by:**
  - **[threshold processes:]**
    - discontinuities of some functions describing the physical processes (some on/off processes)
    - discontinuities of the derivative of a continuous function
  - **[strong nonlinearities]**





finite difference (FD)



TL integration without regularization

u-wind increments
fc t+12, ~700 hPa

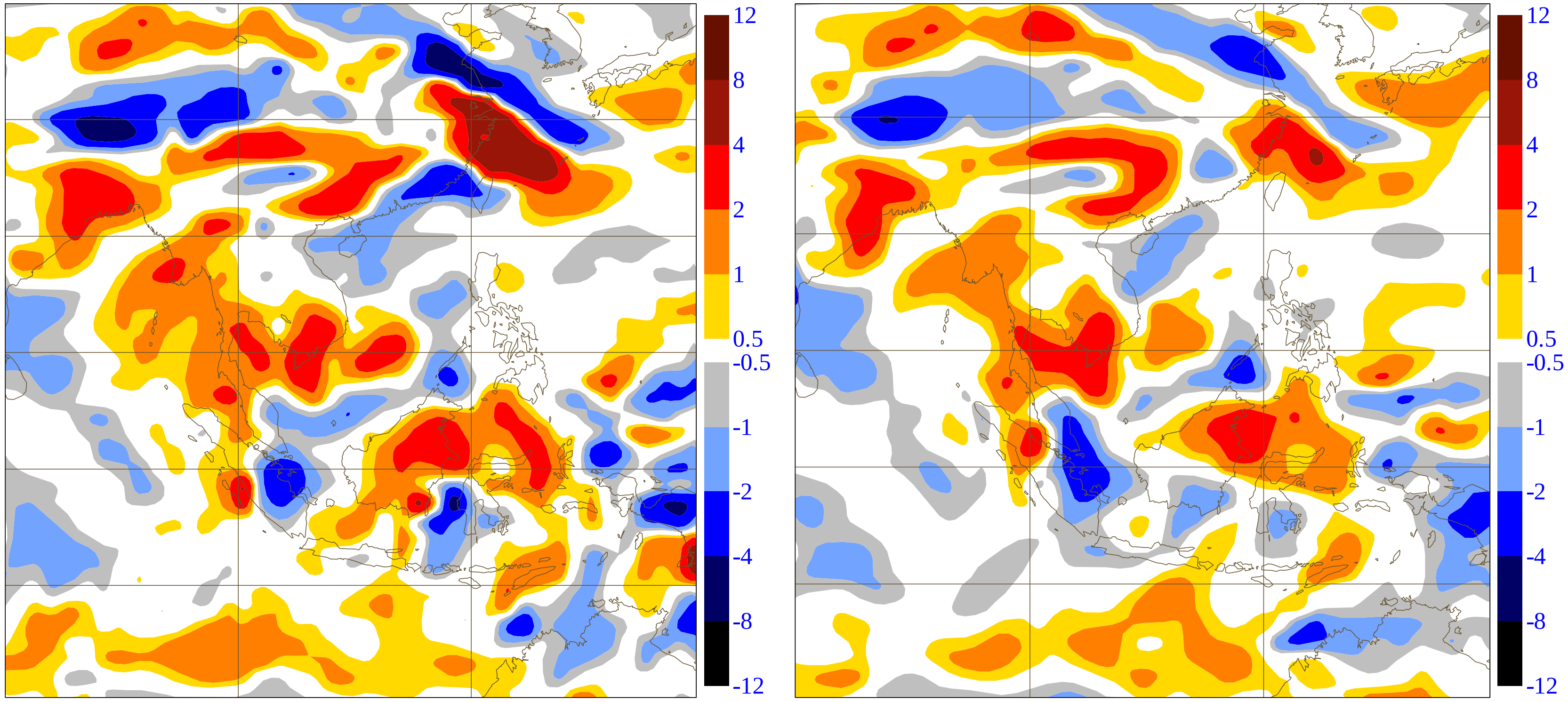
Importance of the regularization of TL model (2)

• regularizations help to remove the most important threshold processes in physical parametrizations effecting the range of validity of TL approximation

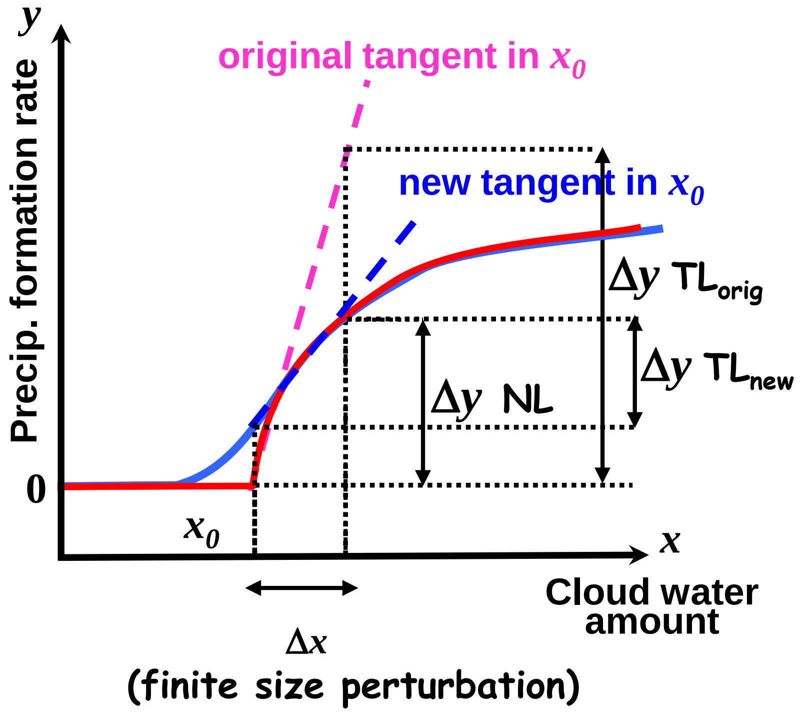
• after solving the threshold problem

⇓

clear advantage of the diabatic TL evolution of errors compared to the adiabatic evolution



finite difference (FD) **TL integration**



u-wind increments
fc t+12, ~700 hPa

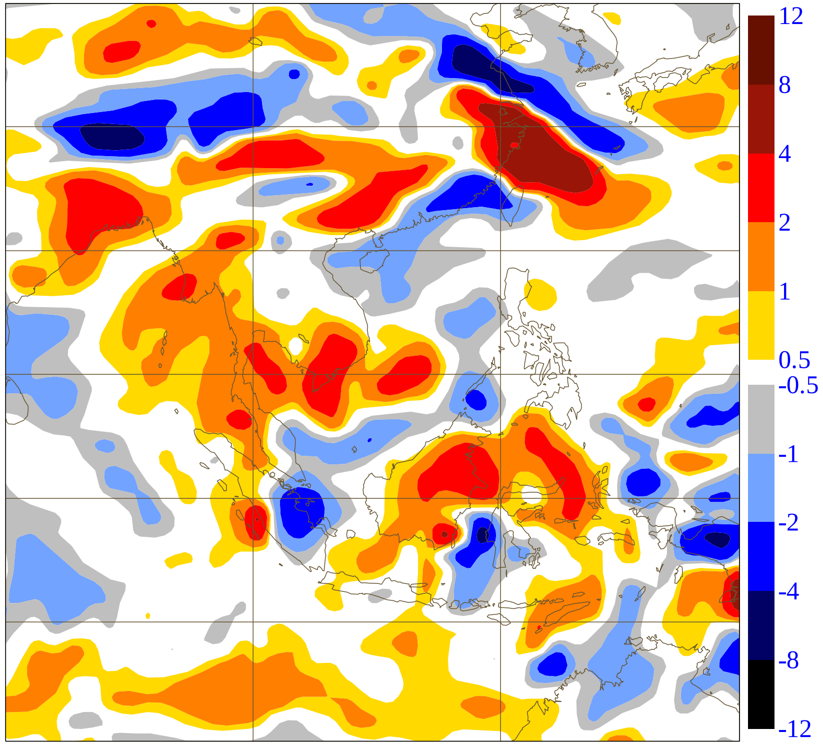
Importance of the regularization of TL model (2)

- regularizations help to remove the most important threshold processes in physical parametrizations effecting the range of validity of TL approximation

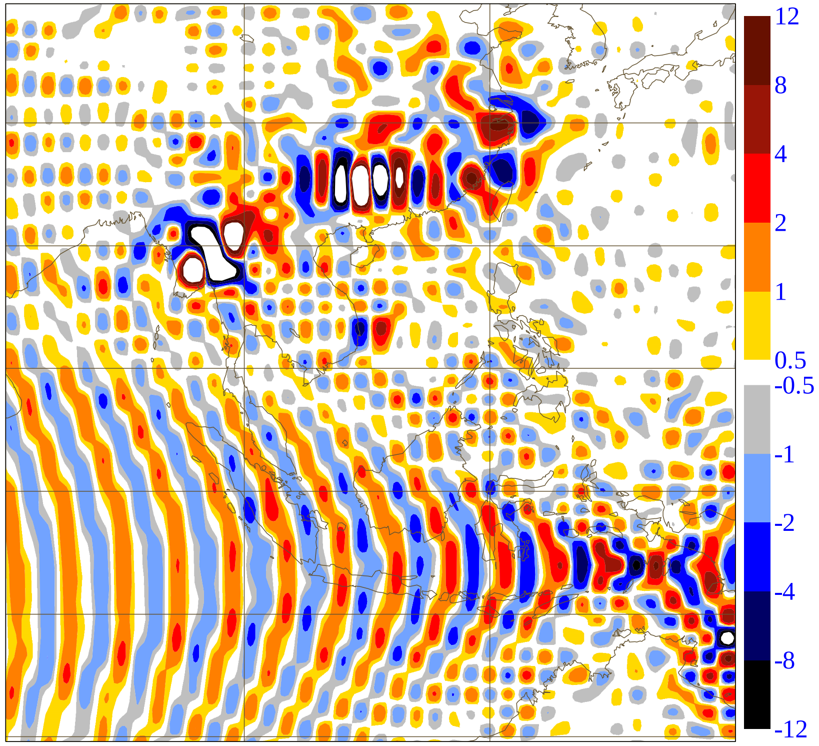
- **after solving the threshold problem**

⇓

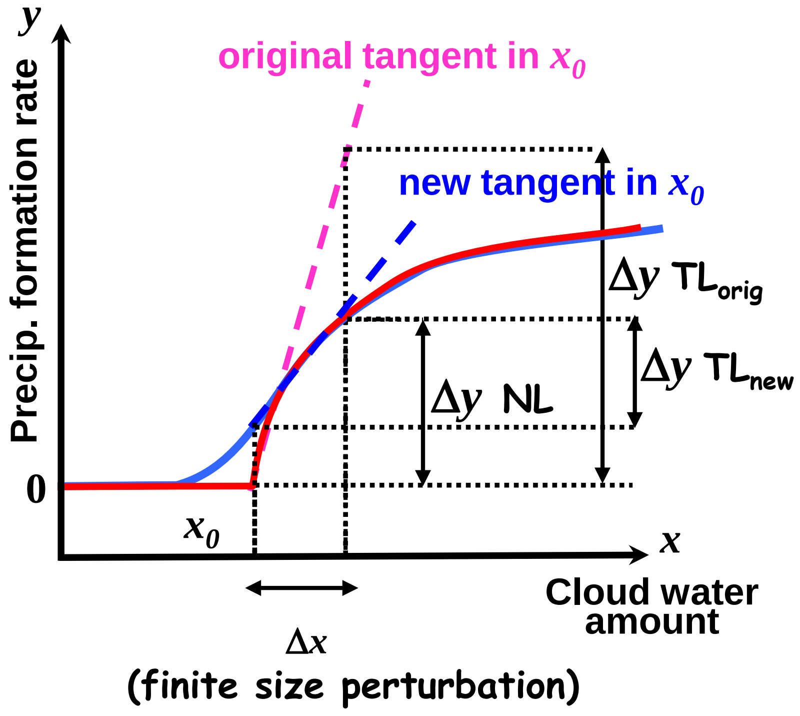
clear advantage of the diabatic TL evolution of errors compared to the adiabatic evolution



finite difference (FD)



TL integration



u-wind increments
fc t+12, ~700 hPa

Detection of problems in NL model by TL diagnostics

- Capable to discover erroneous or false sensitivities (NL model needs to be strictly deterministic – no random computational mode)

  → helping to improve often hidden problems in model and/or observation operators

- Efficient debugging tool when writing TL and AD code line-by-line from the nonlinear (NL) version

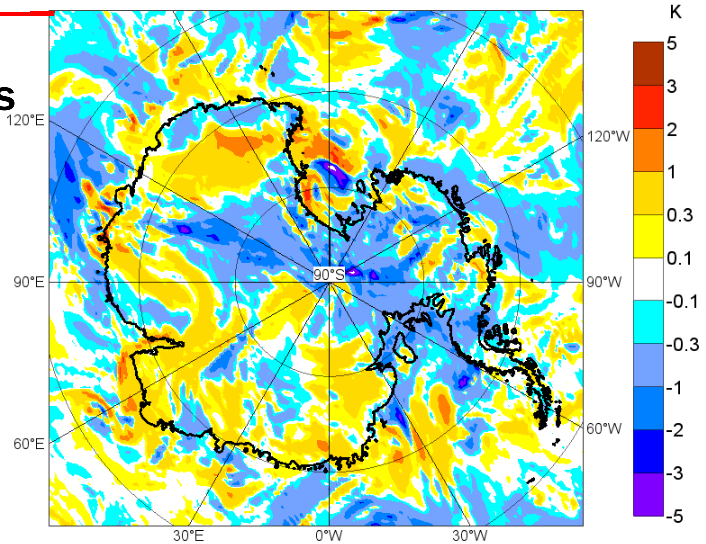
  → coding errors in NL code discovered

- Numerical inaccuracies that may look acceptable in NL models can lead to hidden problems:
  - erroneous derivatives in NL model
  - **noise in TL model**

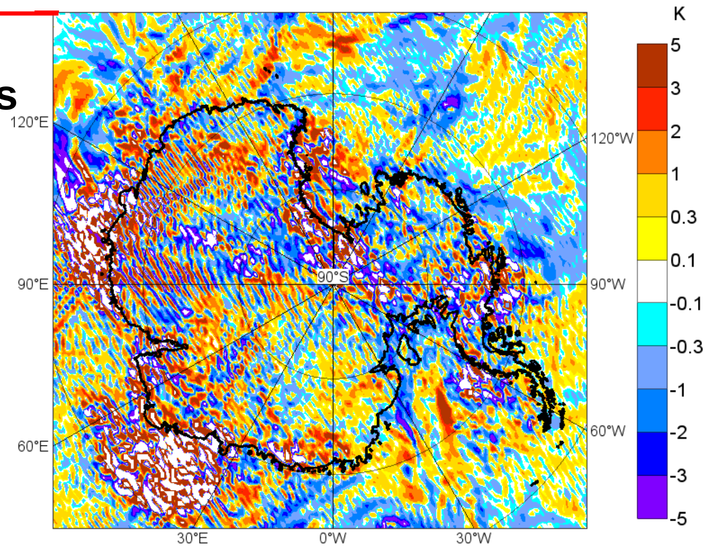
    → getting more pronounced because of increasing resolutions, *number of iterations, getting steeper orography, …*

Assessment of TL approximation revealing hidden problems in NL model

M’δx
with full physics



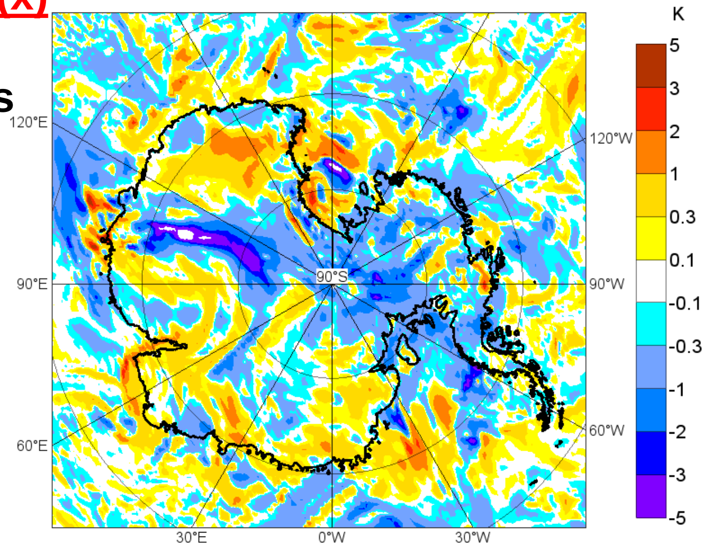
M’δx
with dynamics only



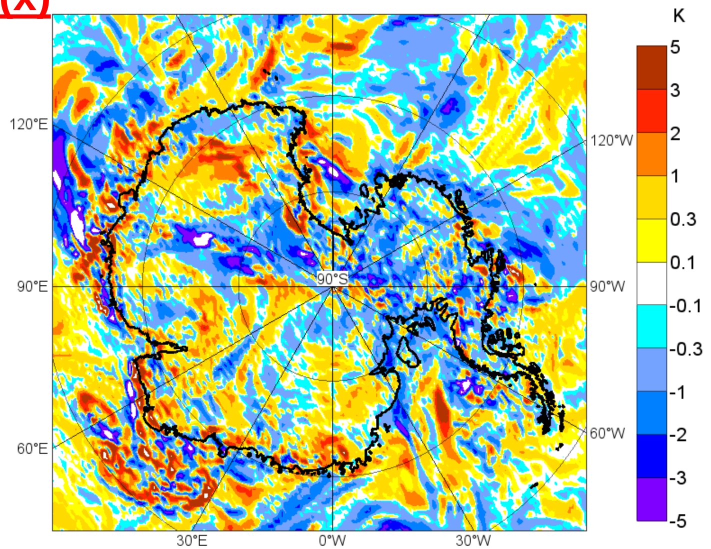
Assessment of TL approximation at very high resolution (TCo639, ~18 km, 12h).

Example: Temperature at level 129 on 20140520 at 12Z over Antarctica.

***M*(x+δx)–*M*(x)**
with full physics



***M*(x+δx)–*M*(x)**
with dynamics only

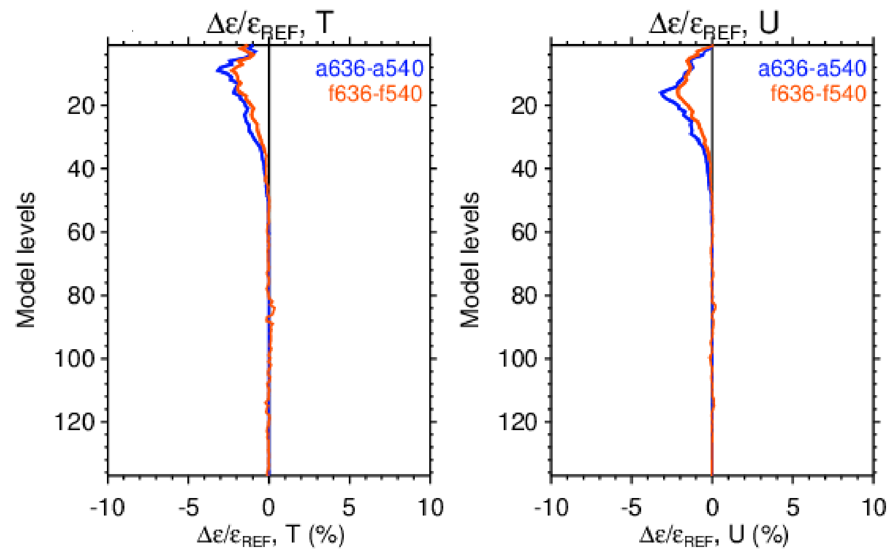


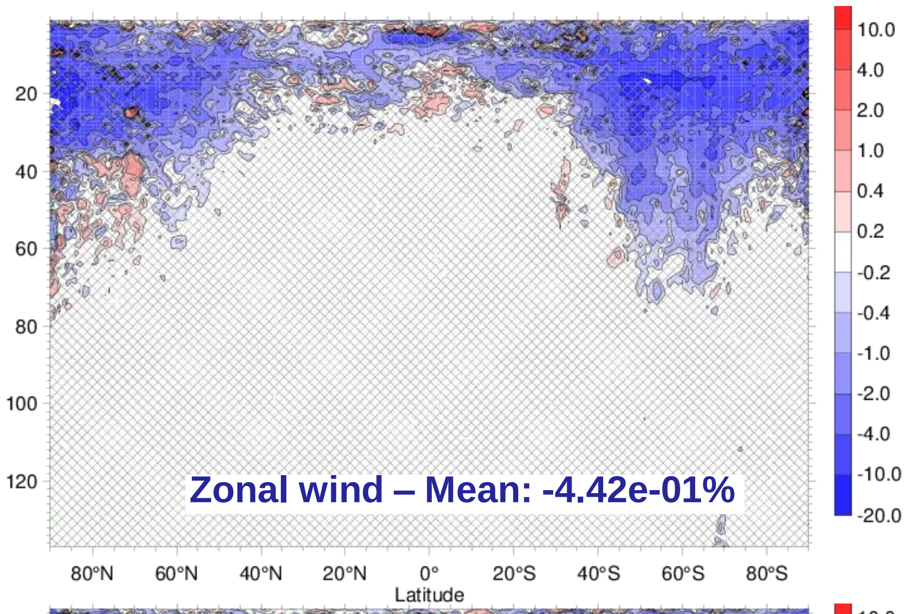
Runs with dynamics only in NL and TL are very noisy close to orography.

Example of model problems identified by TL diagnostics

- **TL diagnostics helped to identify and tune several problems in NL dynamics**, leading to modifications such as:
  - Introducing non-linear flow-dependent filter as a function of flow field deformation in the upper 20 hPa = *cure for "grid-point storms"*:
    → when the flow is laminar the filter does nothing
    → with increased flow deformation, diffusivity increased locally
    D1
  - Applying smooth transition between robust – 1st order scheme and accurate – 2nd order scheme in time for vertical velocity extrapolation above 50 hPa **D2**
  - Curvature term for vector variables computed exactly and not only interpolated
  - Introducing higher order (4th) for SL trajectory research (better respecting wind flow)
  - Increasing accuracy of wind interpolation during SL trajectory research

Improvements in TL approximation based on TL diagnostics



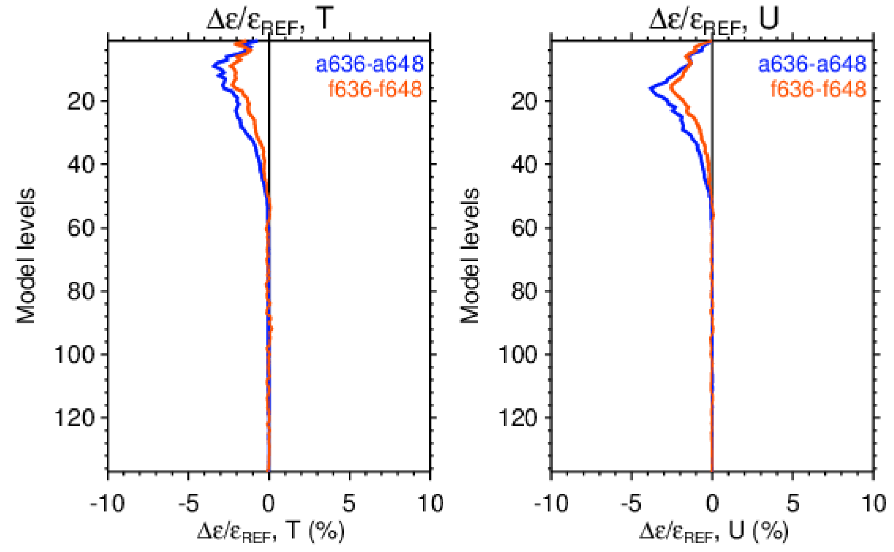


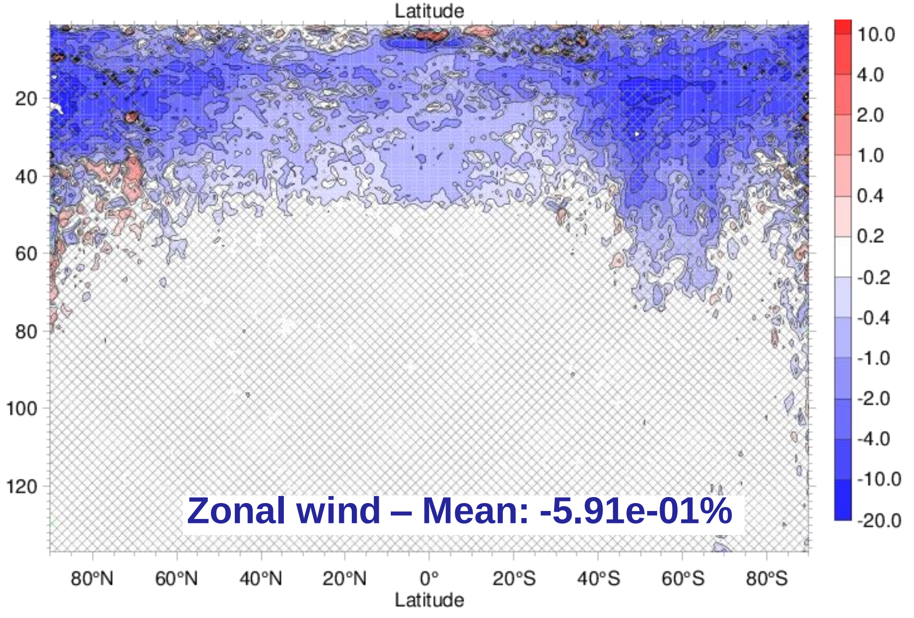
Relative change in TL error: $(\varepsilon_{EXP} - \varepsilon_{REF}) / \varepsilon_{REF}$

(50 km resolution, 20 runs; after 12h integration)

D1

Temperature **Zonal wind**





Blue = TL error reduction = ☺

D1+D2

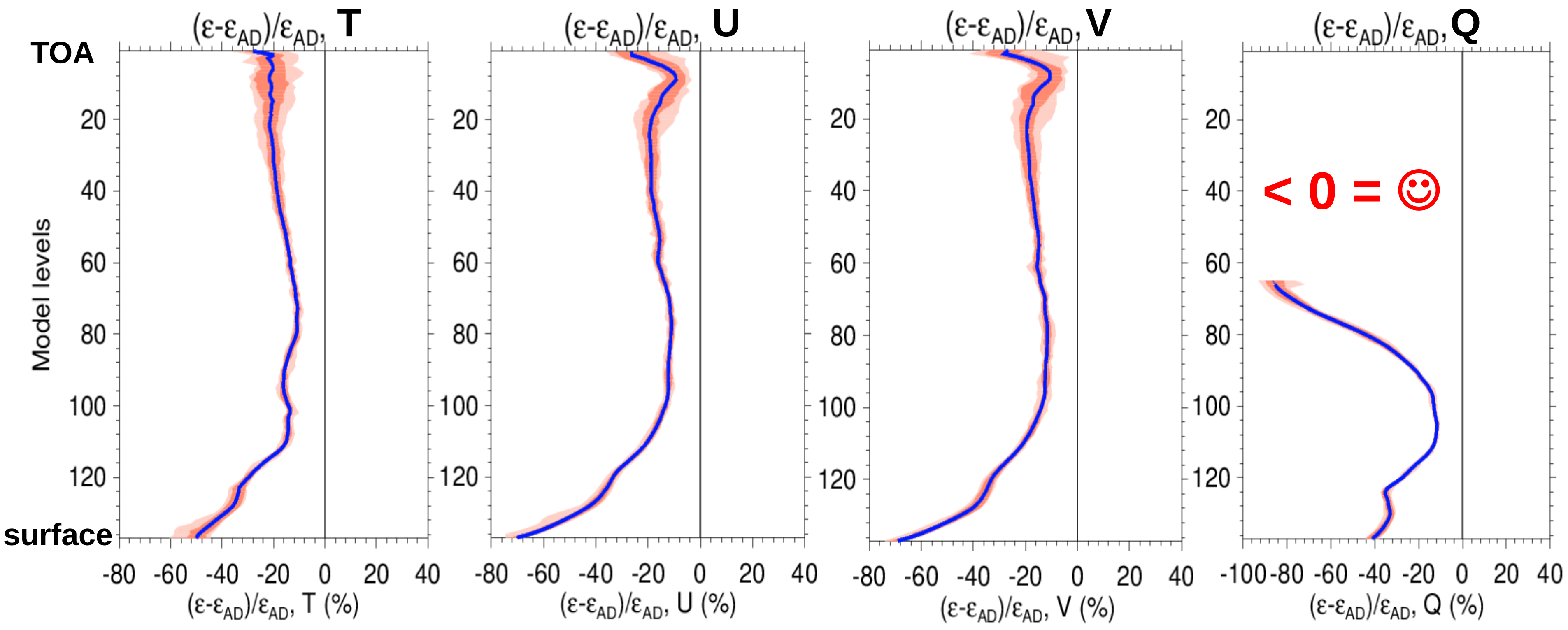
Impact of linearized physics on TL approximation

$$\frac{\varepsilon_{\mathrm{EXP}} - \varepsilon_{\mathrm{REF}}}{\varepsilon_{\mathrm{REF}}} .100\%$$

where

$$\varepsilon = \left| \left[M(\mathbf{x} + \delta\mathbf{x}) - M(\mathbf{x}) \right] - M'(\delta\mathbf{x}) \right|$$

non-linear (NL) difference ↔ tangent-linear (TL) integration



Mean vertical profile of change in TL error when full linearized physics included in TL.

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

Inclusion of linearized physics leads to better TL approximation.

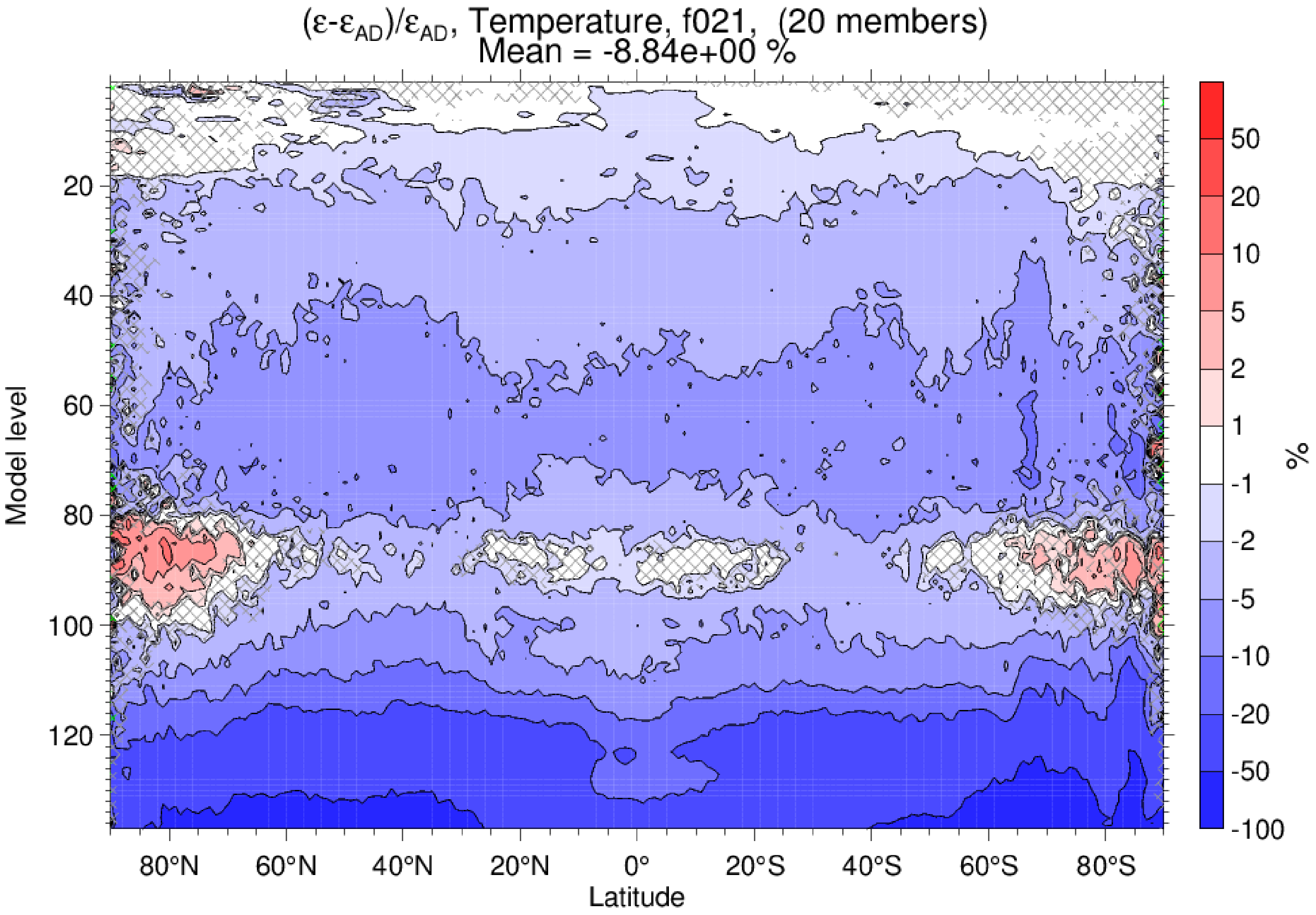
Impact of linearized physics on TL approximation – contribution from different processes (1)

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

VDIF + orog. GWD + SURF

Blue = TL error reduction = ☺

Temperature



Relative to adiabatic TL run:

- **50-km resolution**
- **20 runs**
- **after 12h integr.**

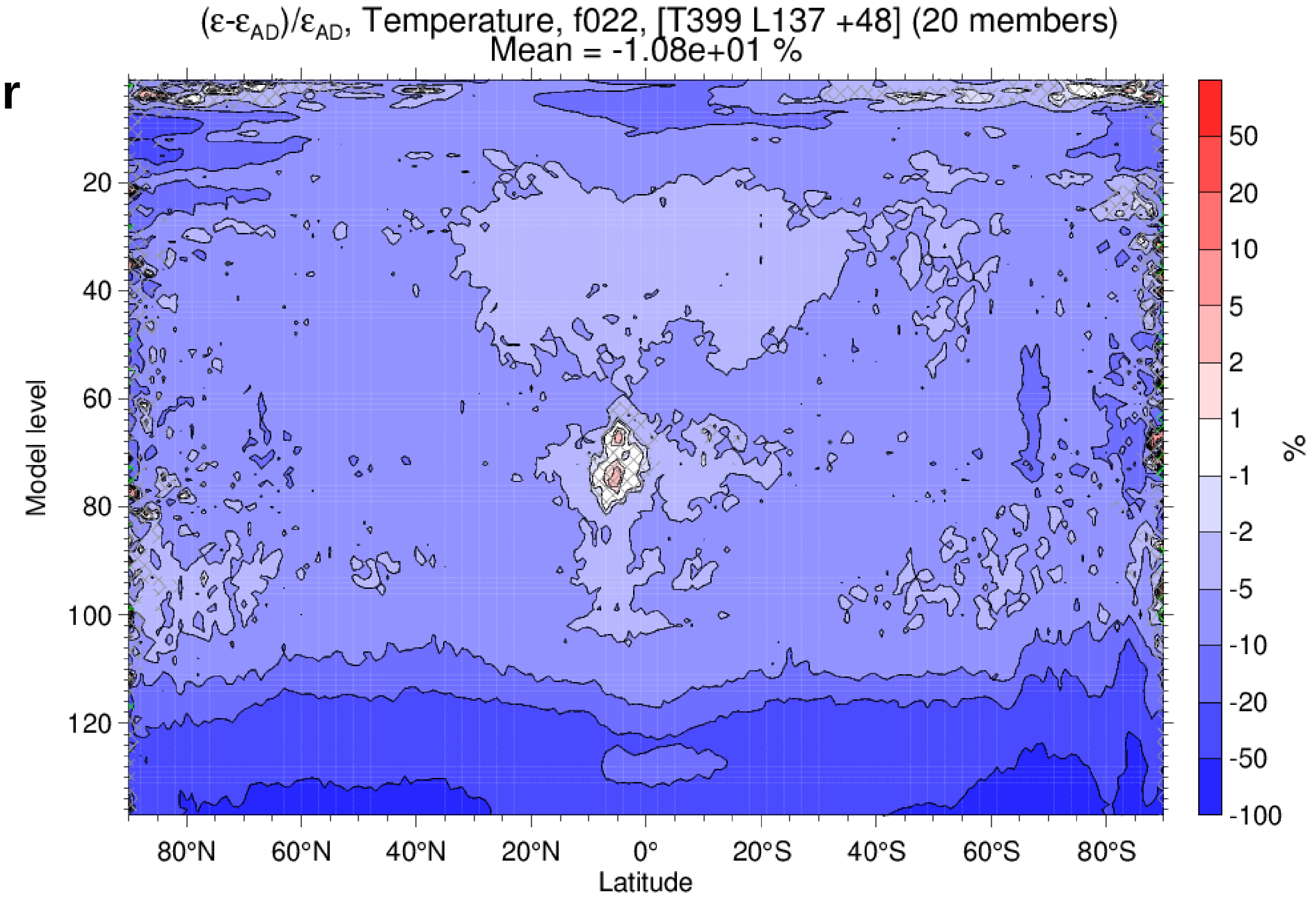
Impact of linearized physics on TL approximation – contribution from different processes (2)

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

VDIF + orog. GWD + SURF + RAD

Blue = TL error reduction = ☺

Temperature



Relative to adiabatic TL run:

- **50-km resolution**
- **20 runs**
- **after 12h integr.**

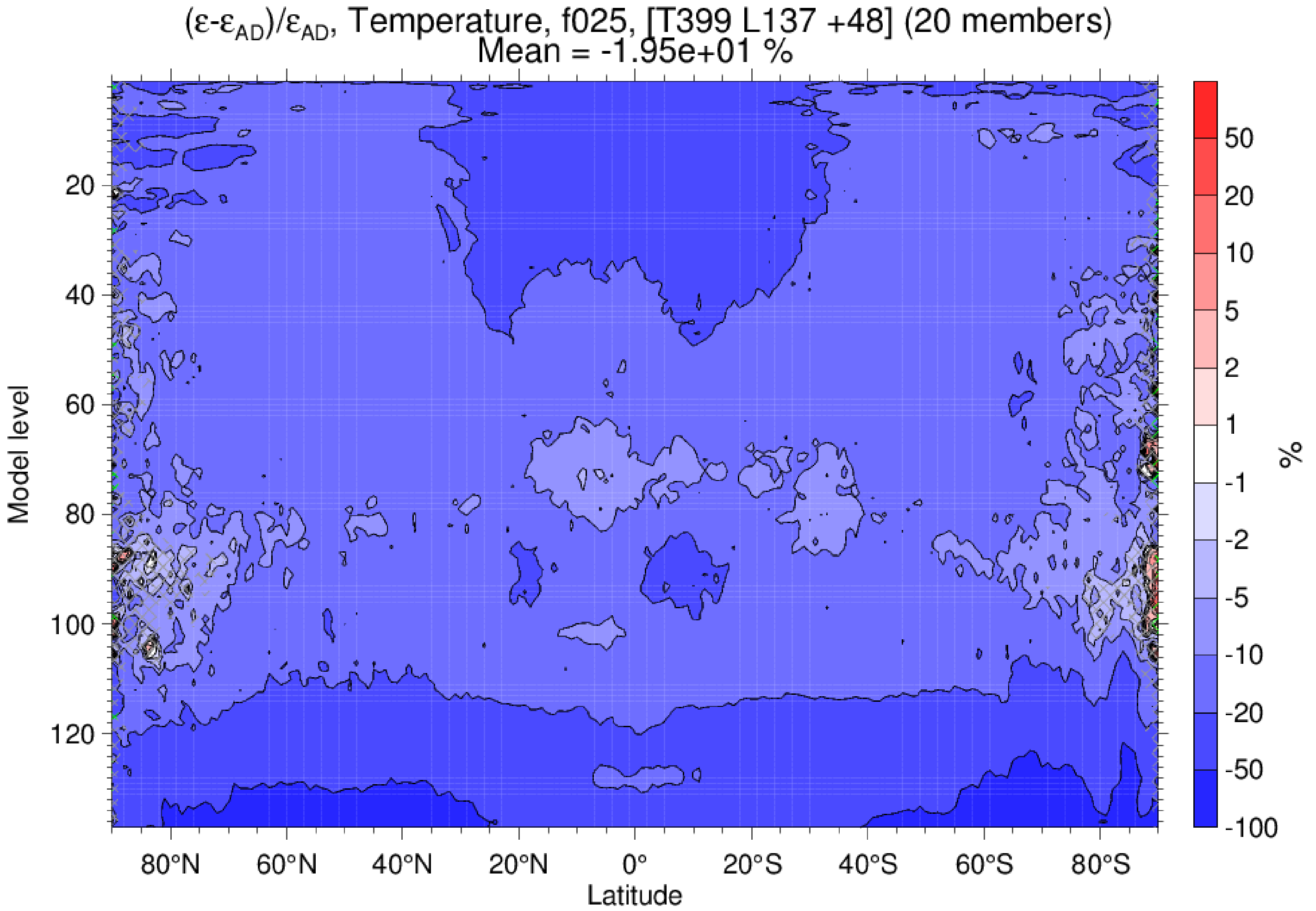
Impact of linearized physics on TL approximation – contribution from different processes (2)

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

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

Blue = TL erro reduction = ☺

Temperature



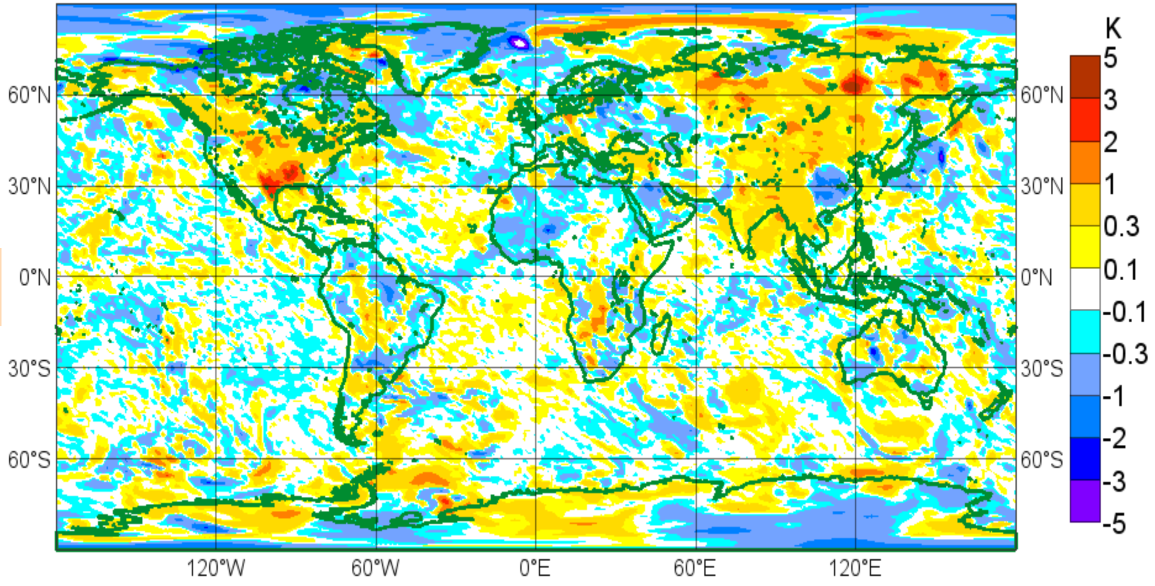
Relative to adiabatic TL run:

- **50-km resolution**
- **20 runs**
- **after 12h integr.**

TL approximation at <u>high</u> resolution (~ 18 km)

Temperature at level 125 (~950 hPa) on 20140105 at 12Z.

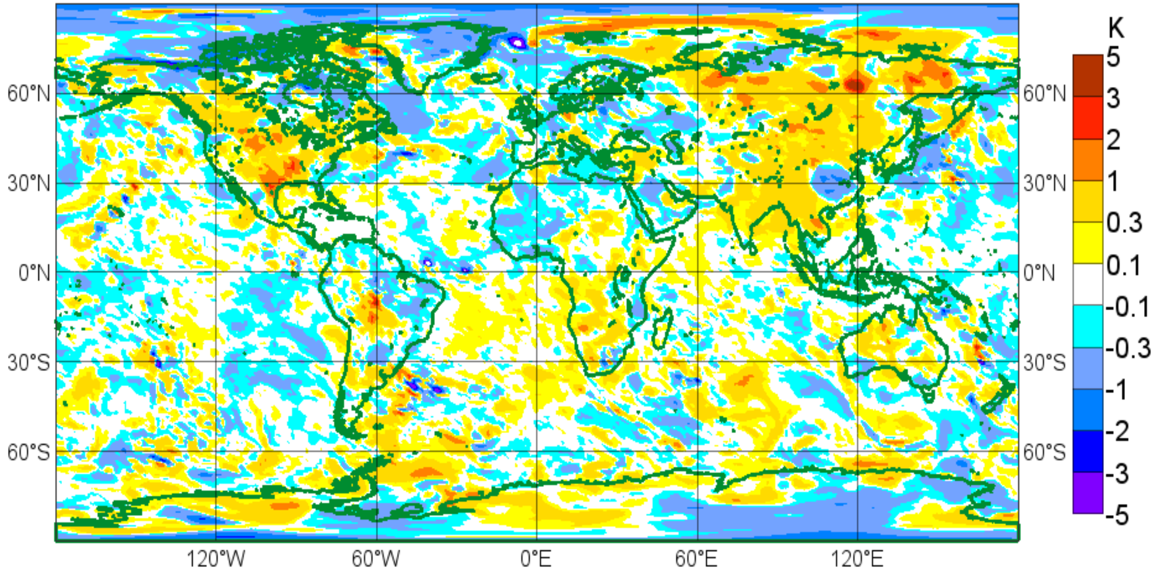
M(**x**+δ**x**)–*M*(**x**)



Comparison of NL difference M(x+δx)–M(x) with perturbation evolved using the TL model M'δx after <u>12h of integration</u>.

TCo639
~ 18 km

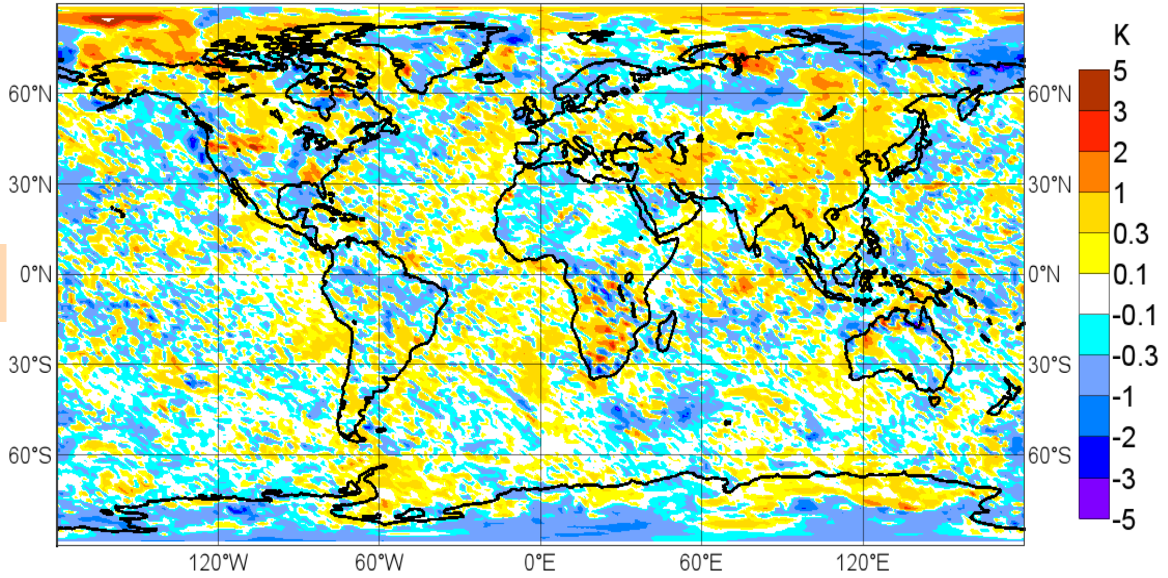
M'δx



Thanks to stabilization of both the dynamics and the physics in the TL model, resolutions as fine as 18 km might be considered in 4D-Var minimizations, provided some (minor) sources of noise can be eliminated.

TL approximation at even higher resolution (~ 9 km)

M(**x**+δ**x**)–*M*(**x**)

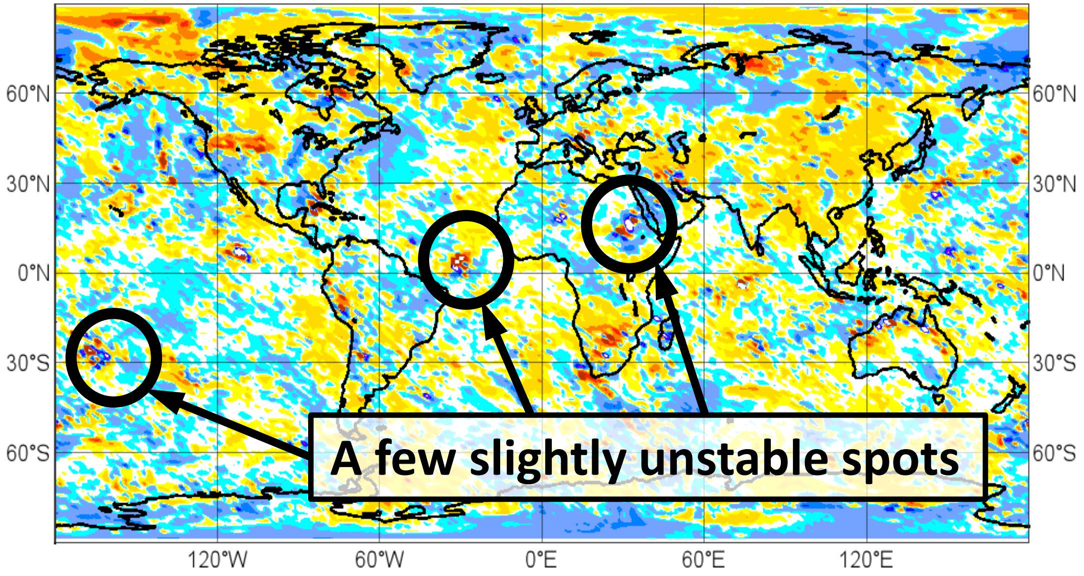


Temperature at level 129 (~980 hPa) on 20140105 at 12Z.

Comparison of NL difference M(x+δx)–M(x) with perturbation evolved using the TL model M'δx after 12h of integration.

TCo1279 ~ 9 km

M'δx



First time our TL model tested at such high resolution and the results surprisingly encouraging.

(Note: this single run required 320 nodes)

Applications of the linearized model

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
  (~ 10 million observations assimilated in ECMWF 4D-Var every 12 hours).

- Over the last several years, observations related to clouds and precipitation started to be routinely assimilated.

- Physical parametrizations are used during the assimilation:
  - to link the model's prognostic variables (typically: T, u, v, q_v and P_s) to the observed quantities (e.g. radiances, reflectivities, precipitation, …),
  - to evolve the model state in time during the assimilation (e.g. 4D-Var).

- But beware: problems in the assimilation can arise due to the discontinuous or non-linear nature of physical processes.

Example: Physics (full & simplified) in incremental 4D-Var system

4D-Var →

$$\min J(\delta\mathbf{x}_0) = \frac{1}{2}\delta\mathbf{x}_0^T\mathbf{B}^{-1}\delta\mathbf{x}_0 + \frac{1}{2}\sum_{i=0}^{n}\left(\mathbf{H}_i(\delta\mathbf{x}_i) - \mathbf{d}_i\right)^T\mathbf{R}_i^{-1}\left(\mathbf{H}_i(\delta\mathbf{x}_i) - \mathbf{d}_i\right)$$

$$\Leftrightarrow \nabla_{\delta\mathbf{x}_0} J = \mathbf{B}^{-1}\delta\mathbf{x}_0 + \frac{1}{2}\sum_{i=0}^{n}\mathbf{M}^T(t_i,t_0)\mathbf{H}_i^T\mathbf{R}_i^{-1}\left(\mathrm{H}_i(\delta\mathbf{x}_i) - \mathbf{d}_i\right) = 0$$

$\mathbf{d}_i = y_i^o - H_i(\mathbf{x}_i^b)$ - innovation vector

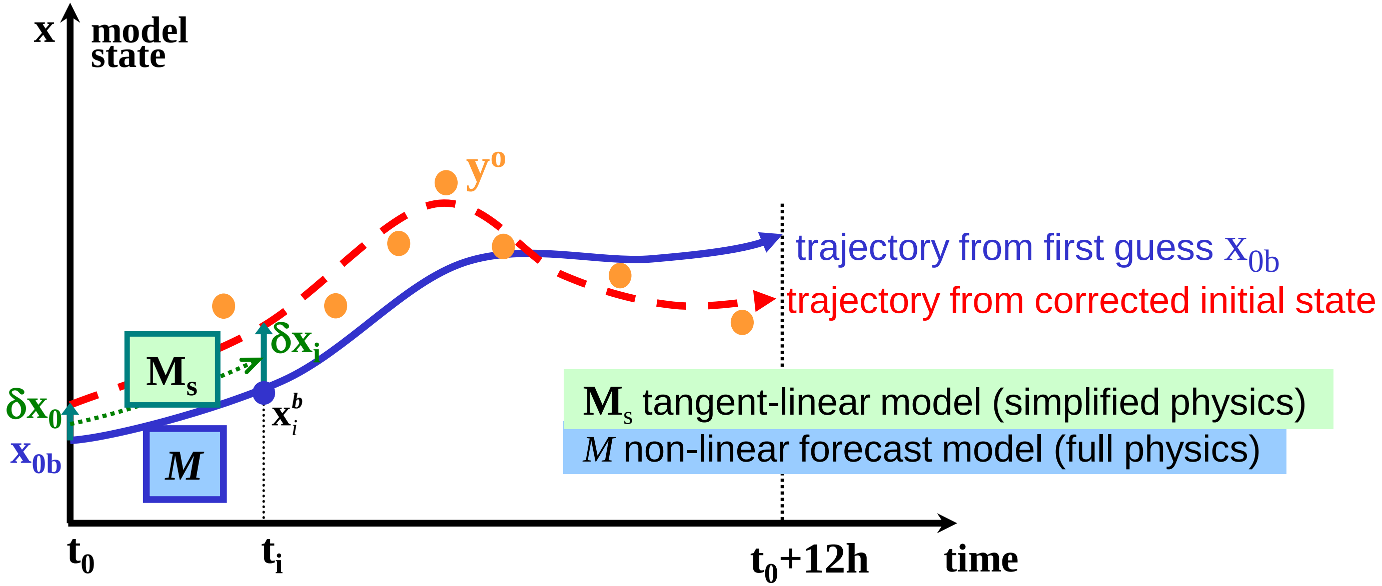
H_i non-linear observation operator

$\mathbf{H}_i$ tangent-linear observation operator

$\mathbf{d}_i$ ← using non-linear model M at high resolution & full physics

$\delta\mathbf{x}_i$ ← using tangent-linear model $\mathbf{M}$ at low resolution & simplified physics

$\nabla_{\delta\mathbf{x}_0} J$ ← using a low resolution adjoint model $\mathbf{M}^T$ & simplified physics



$\mathbf{M}_s$ tangent-linear model (simplified physics)

M non-linear forecast model (full physics)

$\mathbf{x}_i$ – model state at t_i

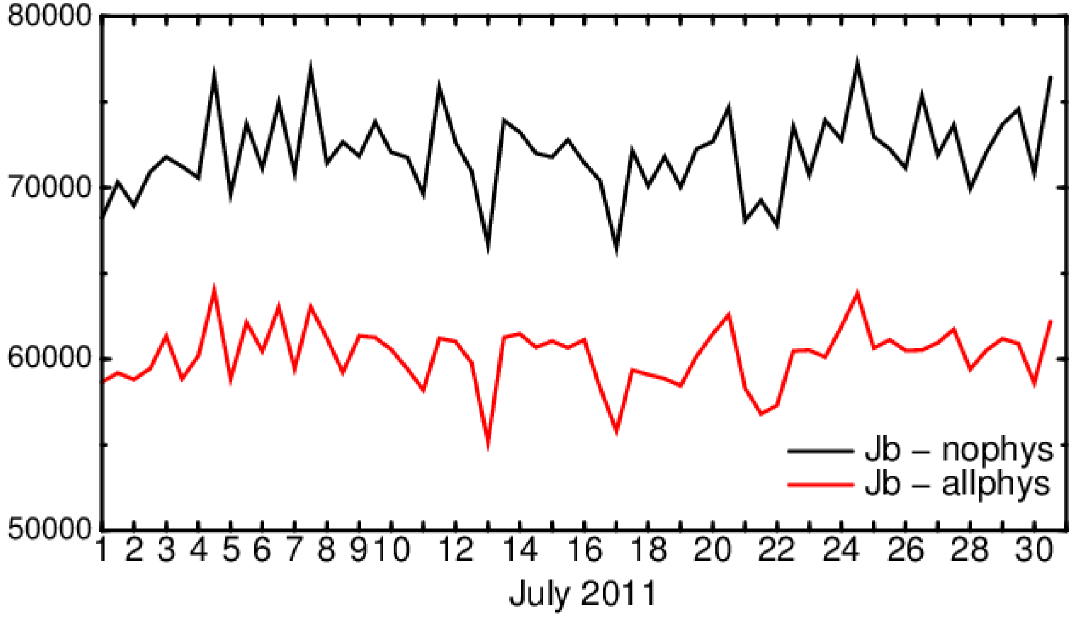
y_i^o – observations at t_i

Impact of linearized physics on analysis

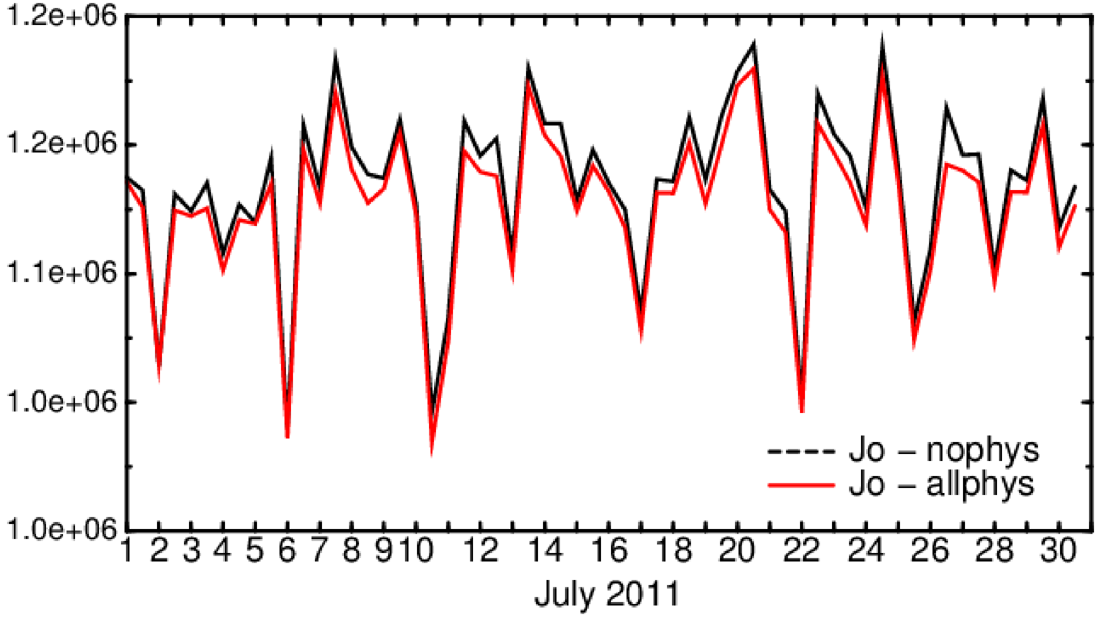
Coming just from including the ECMWF linearized physics in 4D-Var (Janisková & Lopez, 2013)

$$J = J_o + J_b$$

background cost function



observation cost function



4D-Var experiment – July-Sept.2011

nophys = only very simple vertical diffusion and surf.drag of Buizza (1994)

allphys = all linear. phys.parametrization:

- vertical diffusion
- gravity wave drag
- radiation
- nonorog. gravity wave drag
- large scale cond. & precip.
- convection

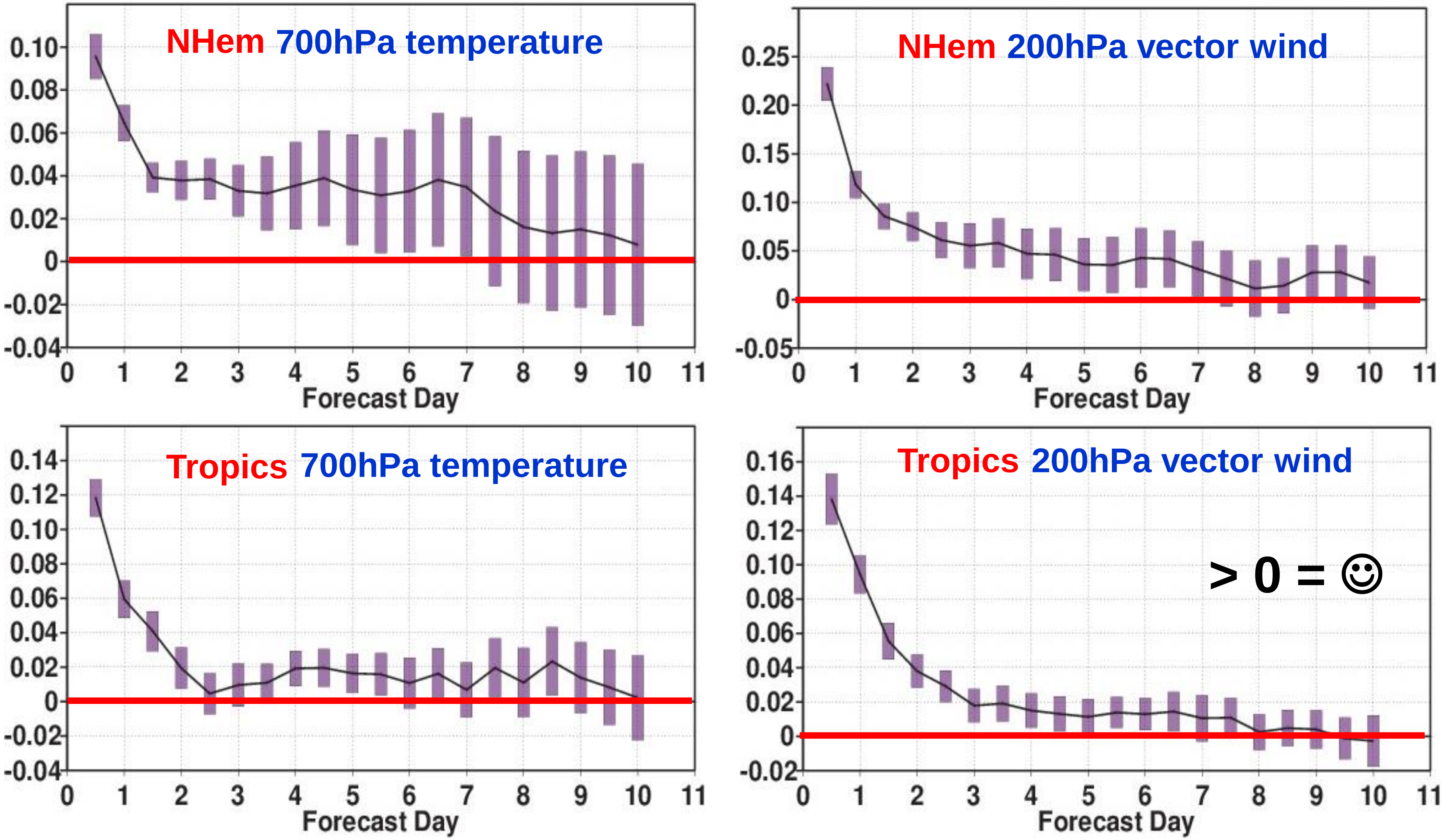
Direct relative improvement of forecast scores from linearized physics (1)

Coming just from including the ECMWF linearized physics in 4D-Var (Janisková & Lopez, 2013)

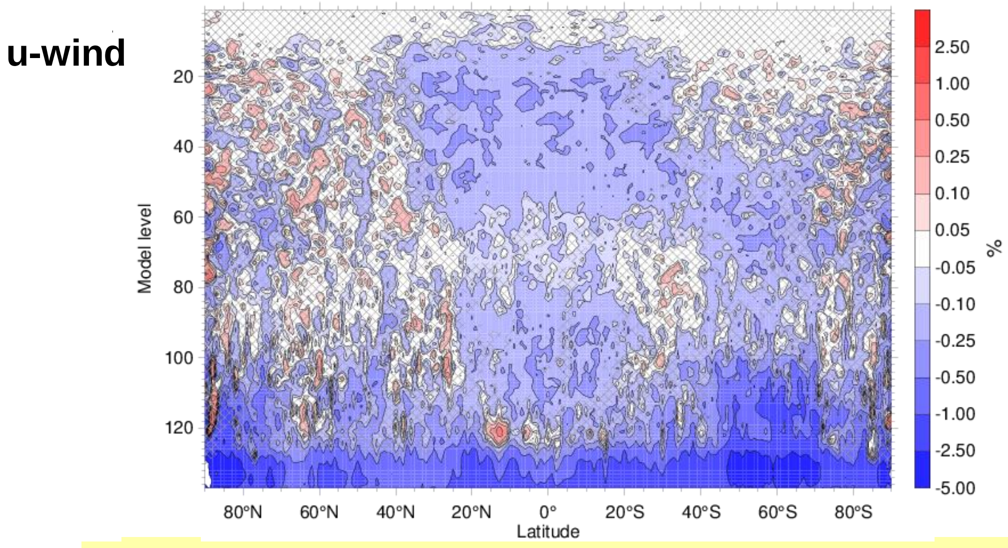
Anomaly correlation – 3 month experiment:

bars indicate significance at 95% confidence level

T511L91 FC run: Forecast scores against operational analysis



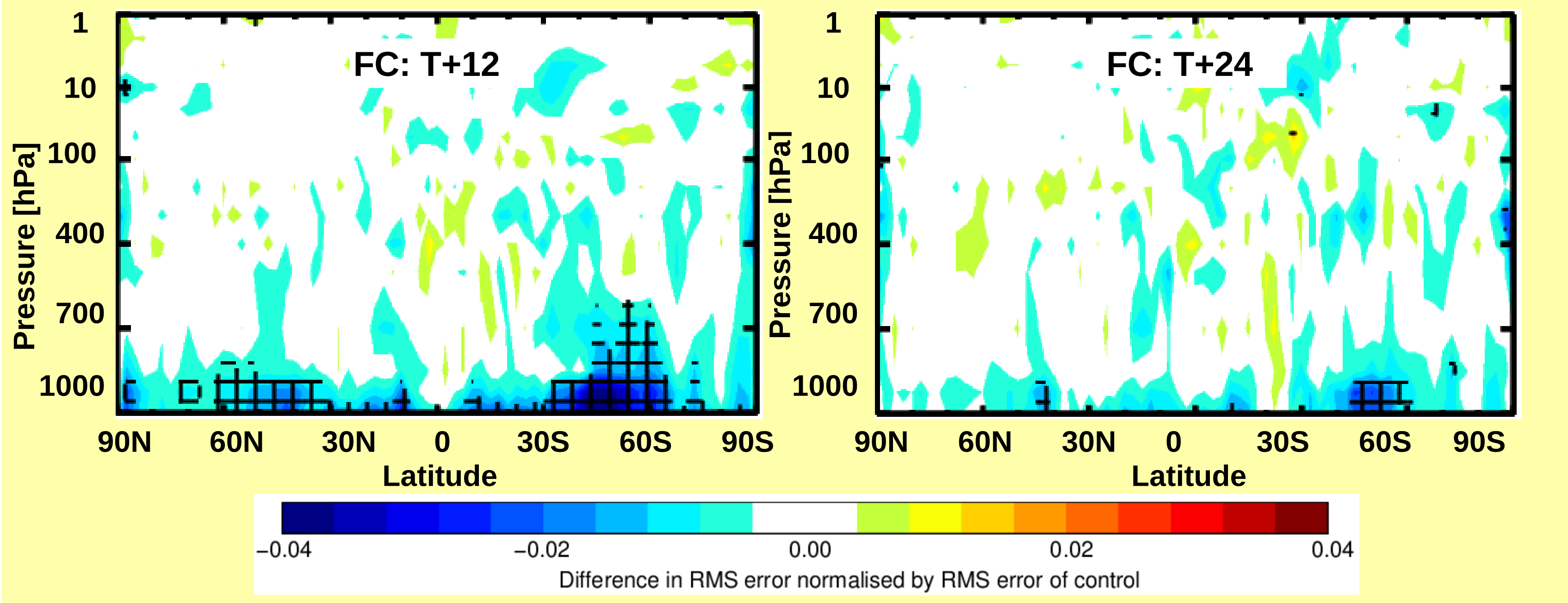
Direct relative improvement of forecast scores from linearized physics (2)



Relative impact [%] of the surface related modifications in the tangent-linear model (TL) on TL approximation:

12-hour evolution of zonal wind increments

Negative values (blue)
↓
improvement

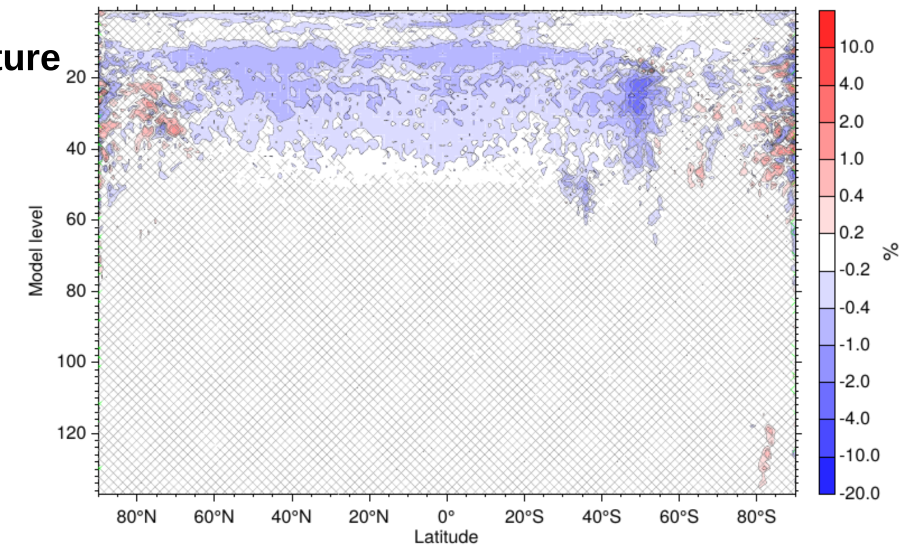


Forecast:
Wind vector

Difference in RMS error normalized by RMS error of control:
February-March 2014

Relative improvement of forecast scores from dynamics modifications

Temperature

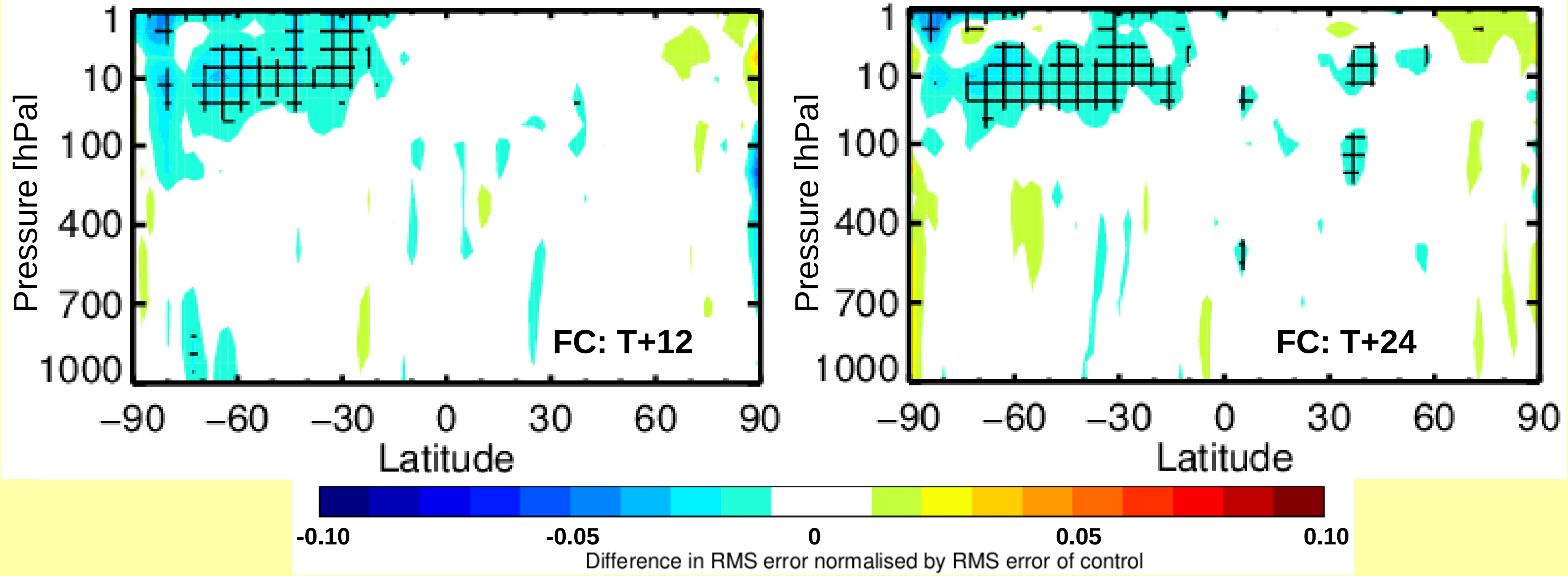


Relative impact [%] of the dynamics related modifications in the tangent-linear model (TL) on TL approximation:

12-hour evolution of <u>temperature increments</u>

D2

Negative values (blue)
↓
improvement



Forecast:

<u>Geopotential</u>

Difference in RMS error normalized by RMS error of control:
November-January 2017

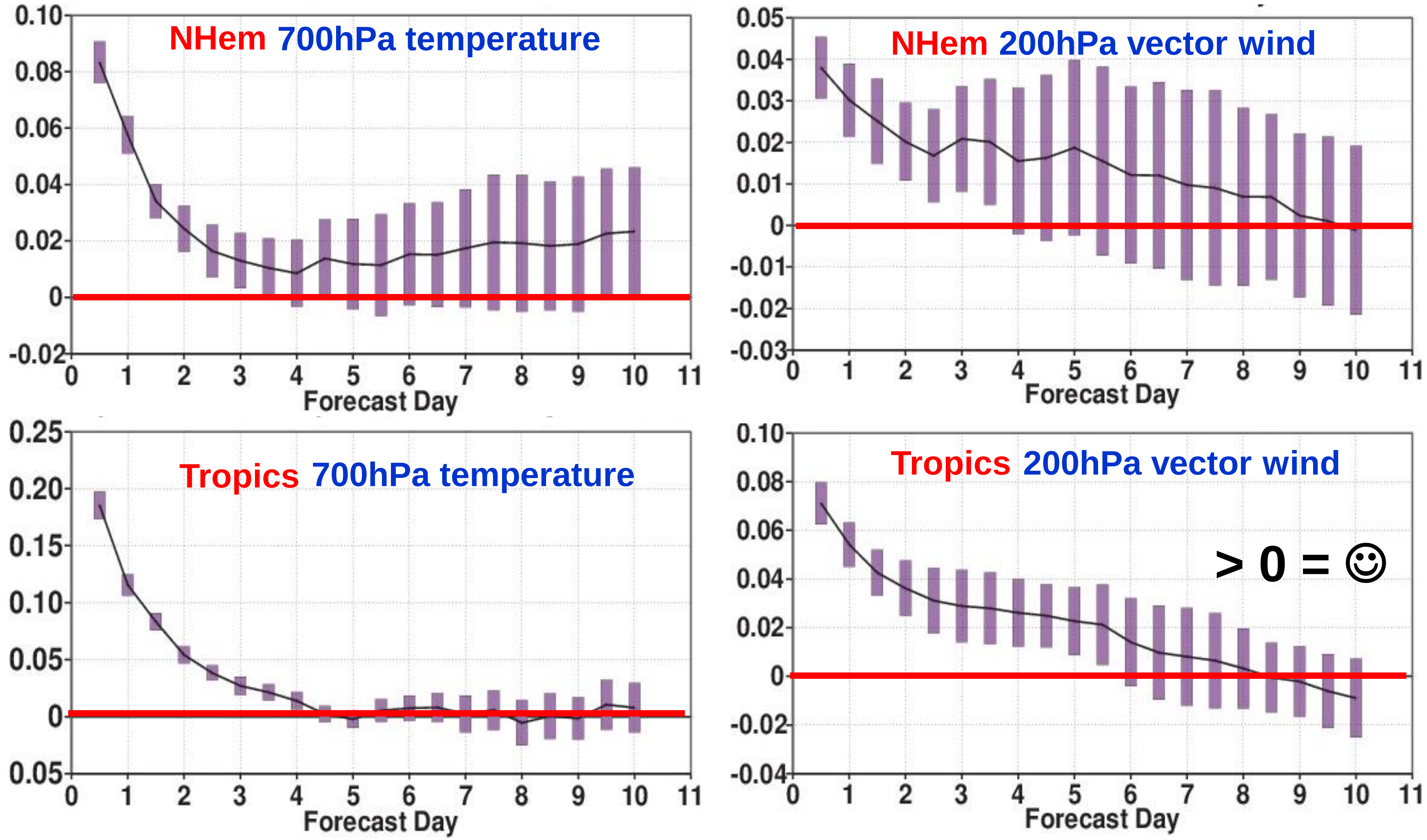
D2 – **Applying smoother transition between 1st and 2nd order schemes in time for vertical velocity extrapolation above 50 hPa**

Indirect relative improvement of forecast scores from linearized physics

Using observations directly related to physical processes *(e.g. rain, clouds, …)*

Anomaly correlation – 3 month experiment:
bars indicate significance at 95% confidence level

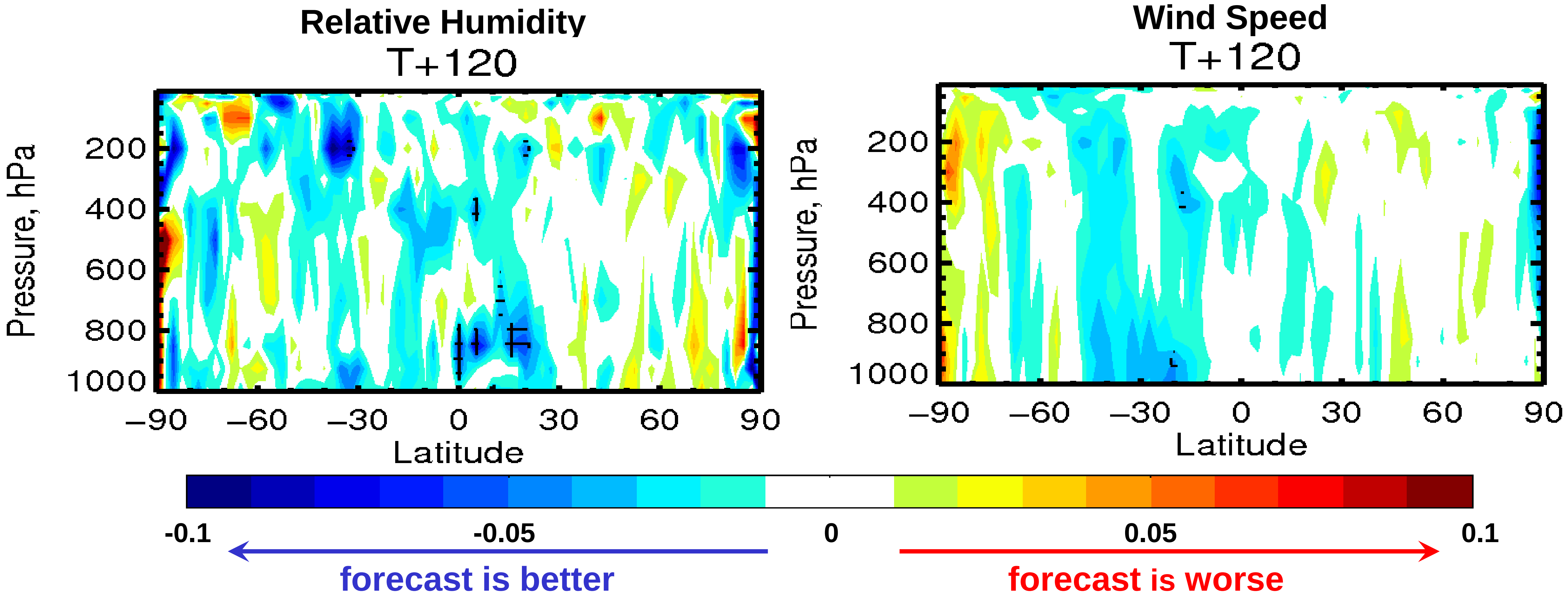
T799L137 FC run: Forecast scores against operational analysis



4D-Var assimilation of SSM/I brightness temperatures

Impact of the direct 4D-Var assimilation of SSM/I all-skies TBs on the <u>relative change in 5-day forecast RMS errors</u> (zonal means).

Period: 22 August 2007 – 30 September 2007



Geer et al. 2010, Bauer et al. 2010

Assimilation of NCEP Stage IV hourly precipitation data over the U.S.A.

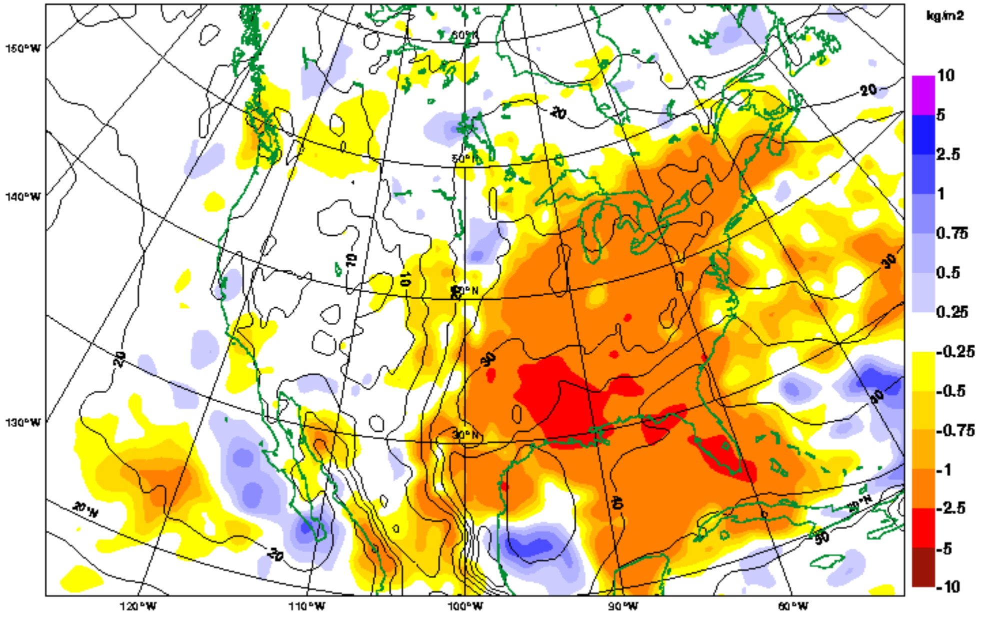
Own impact of combined ground-based radar & rain gauge observations

Three 4D-Var assimilation experiments (20 May - 15 June 2005):

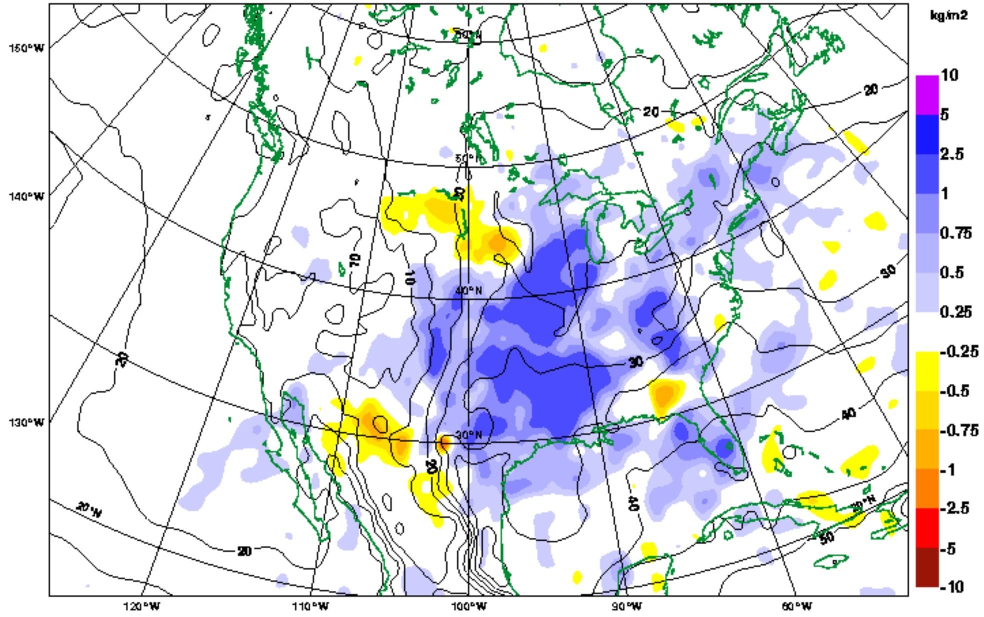
CTRL = all standard observations.
CTRL_noqUS = all obs except no moisture obs over US (surface & satellite).
NEW_noqUS = CTRL_noqUS + NEXRAD hourly rain rates over US ("1D+4D-Var").

Mean differences of TCWV analyses at 00UTC

CTRL_noqUS – CTRL

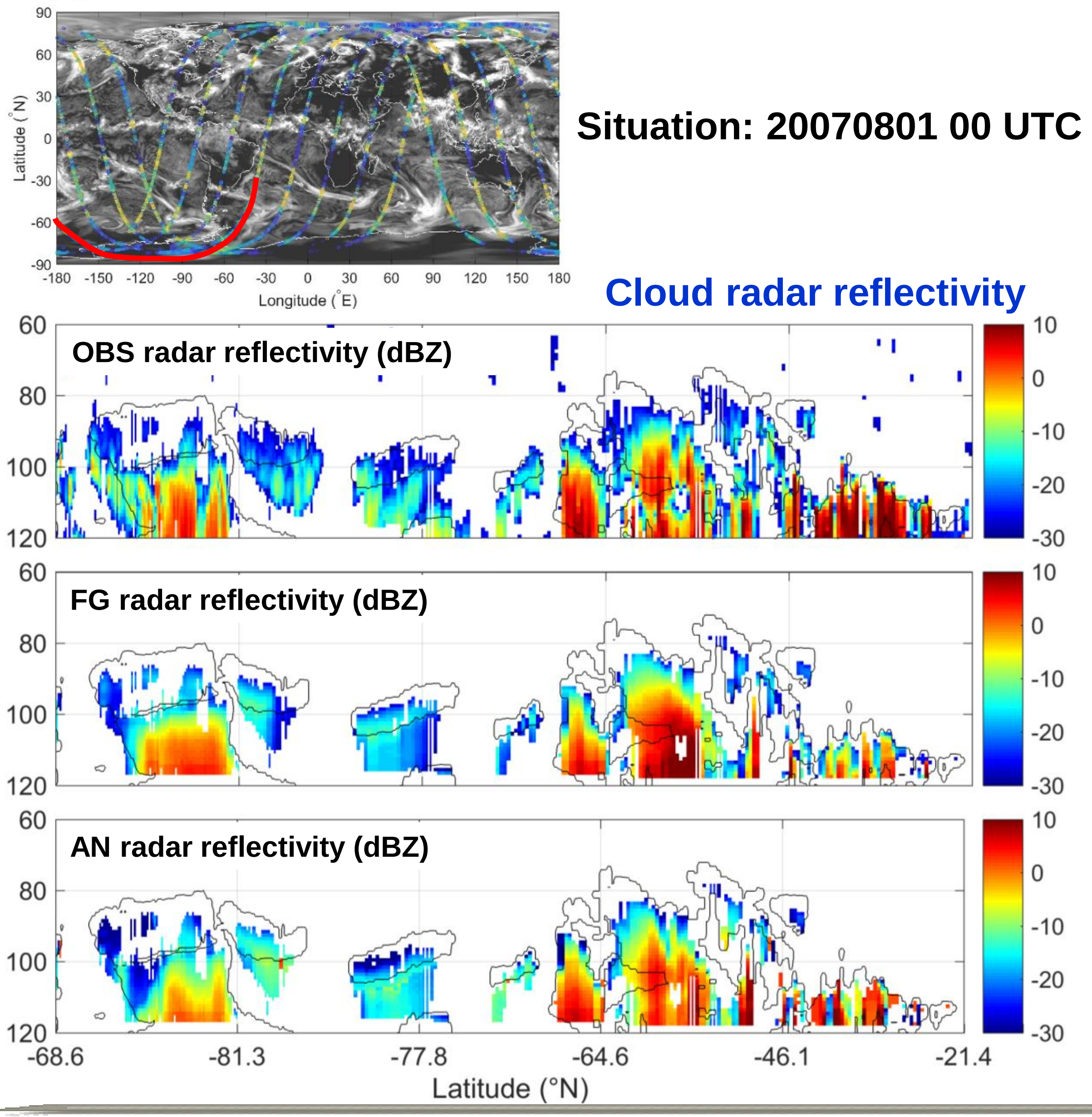


NEW_noqUS – CTRL_noqUS



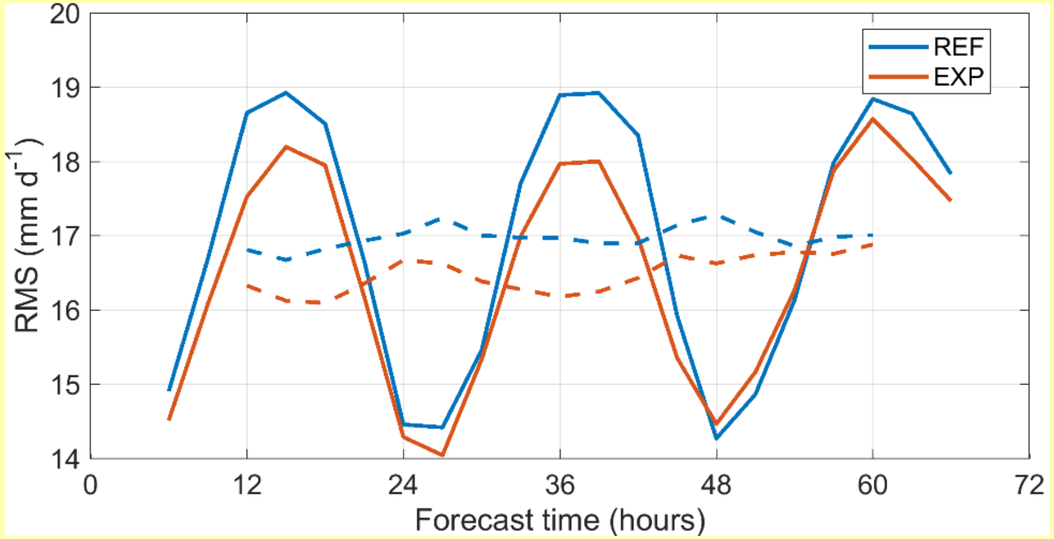
Lopez and Bauer (Monthly Weather Review, 2007)

4D-Var assimilation of CloudSat cloud radar reflectivity and CALIPSO cloud lidar backscatter



Verification of forecast against TRMM data for 7 days of 4D-Var cycling:

REF – all standard observations
EXP – all standard obs + cloud radar&lidar



- Assimilation of CloudSat and CALIPSO obs:
  - a positive impact on analysis fit to obs and subsequent short-term forecast
  - improving forecast of rain rates in Tropics

Janisková and Fielding, 2018

Summary and prospects (1)

- **Physical parametrizations become important 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;
  - positive impact on analysis and subsequent forecast.

- **However, there are some limitations to the approach using linearized models:**

1) **Theoretical:**
   The domain of validity of the linear hypothesis shrinks with increasing resolution & integration lenght.

2) **Technical:**
   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, for the linearized model it is important to achieve the best compromise between:**

  → linearity

  → realism

  → cost

- **Alternative data assimilation methods not requiring the development of linearized code exist, 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).

- **The good TL approximation obtained at global high resolution up to 9 km is encouraging** as current minimizations are run at 50 km at best.