# Aplikácia zjednodušenej, diferencovateľnej a realistickej fyzikálnej parametrizácie pre inkrementálnu štvorrozmernú variačnú asimiláciu dát

# Realization of a simplified, differentiable and realistic physical parametrization for incremental four-dimensional variational data assimilation

# Réalisation d'une physique simplifiée, différentiable et réaliste pour une assimilation variationnelle quadri-dimensionelle incrémentale

Marta Janisková

## POSITION OF THE PROBLEM

- IMPORTANCE OF THE ASSIMILATING MODEL

– a large part of forecasting deficiencies is connected with the imperfect assimilation of the available data

– the better the assimilating model (4D-Var consistently using the information coming from the observations and the model)

$$\Downarrow$$

the better the analysis
(and the subsequent forecast)

– the more sophisticated the model (4D-Var containing physical parametrizations)

$$\Downarrow$$

the more difficult the minimization
(on-off processes, non-linearities)

- INCREMENTAL APPROACH

the minimization can be performed at lower resolution and with a simplified linear model

$$\Downarrow$$

the linear model can be "not tangent" to the full model

(different physics, different resolution and geometry)

# STANDARD FORMULATION OF 4D-VAR

- the goal is to define the atmospheric state $\mathbf{x}(t_0)$ such that the "distance" between the model trajectory (performed from $\mathbf{x}(t_0)$) and the observations is minimum over a given time period $[t_0, t_n]$

$$\Downarrow$$

finding the model state at the initial time $t_0$ which minimizes an objective function

$$\begin{aligned}\mathcal{J}(\mathbf{x}(t_0)) &= \frac{1}{2}\sum_{i=0}^{n}[H_i(\mathbf{x}(t_i)) - \mathbf{y}_i^o]^T\mathbf{R}_i^{-1}[H_i(\mathbf{x}(t_i)) - \mathbf{y}_i^o] \\ &+ \frac{1}{2}[\mathbf{x}(t_0) - \mathbf{x}^b(t_0)]^T\mathbf{B}^{-1}[\mathbf{x}(t_0) - \mathbf{x}^b(t_0)]\end{aligned}$$

where

| | |
|---|---|
| $\mathbf{x}(t_0)$ | is the initial state of the model $M$, |
| $\mathbf{x}(t_i)$ | is the model state at time $t_i$ ($\mathbf{x}(t_i) = M(t_i, t_0)\mathbf{x}(t_0)$), |
| $\mathbf{x}^b$ | is the background vector, |
| $\mathbf{y}_i^o$ | is the observation vector at time $t_i$, |
| $\mathbf{R}_i$ | is the covariance matrix of observation error at time $t_i$, |
| $\mathbf{B}$ | is the covariance matrix of background error, |
| $H_i$ | is the observation operator at time $t_i$ |

# INCREMENTAL FORMULATION OF 4D-VAR

- model state at any time $t_i$ is defined as

$$\mathbf{x}(t_i) \;=\; \mathbf{x}^b(t_i) + \delta\mathbf{x}(t_i)$$

with

$$\begin{aligned}\mathbf{x}^b(t_i) &= M(t_i, t_0)\mathbf{x}^b(t_0) \\ \delta\mathbf{x}(t_i) &= M'(t_i, t_0)\delta\mathbf{x}(t_0) \longleftarrow \textit{tangent linear model}\end{aligned}$$

- subsequently under the tangent linear hypothesis one has:

$$\mathbf{x}(t_i) = M(t_i, t_0)\mathbf{x}^b(t_0) + M'(t_i, t_0)\delta\mathbf{x}(t_0)$$

and similarly,

$$H_i(\mathbf{x}(t_i)) \;=\; H_i(\mathbf{x}^b(t_i)) + H_i'\;\delta\mathbf{x}(t_i)$$

- 4D-Var can be then approximated to the first order as minimizing

$$\begin{aligned}\mathcal{J}(\delta\mathbf{x}(t_0)) = & \frac{1}{2}\sum_{i=0}^{n}\left[H_i(\mathbf{x}^b(t_i)) - \mathbf{y}_i^o + H_i'\delta\mathbf{x}(t_i)\right]^T\mathbf{R}_i^{-1} \\ & \times\left[H_i(\mathbf{x}^b(t_i)) - \mathbf{y}_i^o + H_i'\delta\mathbf{x}(t_i)\right] \qquad \mathcal{J}^o \\ + & \frac{1}{2}[\delta\mathbf{x}(t_0)]^T\mathbf{B}^{-1}[\delta\mathbf{x}(t_0)] \qquad \mathcal{J}^b\end{aligned}$$

- the analysis is obtained as

$$\mathbf{x}^a(t_0) = \mathbf{x}^b(t_0) + \delta\mathbf{x}^a(t_0)$$

- incremental approach is a cost-effective formulation of the 4D-Var
- to reduce the computational cost, a simplification operator $S$ can be introduced as

$$\delta\mathbf{w}(t_0) = S\,\delta\mathbf{x}(t_0)$$

- let $S^{-I}$ be its generalized inverse
- a simplified tangent linear dynamics $\mathbf{L}$ can also be introduced ($\mathbf{L}$ *can be a linear model at low resolution, adiabatic or containing the* **simplified physics**)

$$\delta\mathbf{w}(t_i) = \mathbf{L}(t_i, t_0)\,\delta\mathbf{w}(t_0)$$

- let $\mathbf{G}_i$ be the simplified observation operator approximating $H_i'.S^{-I}$
- the simplified minimization problem becomes

$$\begin{aligned}\mathcal{J}(\delta\mathbf{w}(t_0)) = & \frac{1}{2}\sum_{i=0}^{n}[H_i(\mathbf{x}^b(t_i)) - \mathbf{y}_i^o + \mathbf{G}_i\delta\mathbf{w}(t_i)]^T\mathbf{R}_i^{-1} \\ & \times[H_i(\mathbf{x}^b(t_i)) - \mathbf{y}_i^o + \mathbf{G}_i\delta\mathbf{w}(t_i)] \\ + & \frac{1}{2}[\delta\mathbf{w}(t_0)]^T\mathbf{B}_{(\mathbf{w})}^{-1}[\delta\mathbf{w}(t_0)]\end{aligned}$$

  $\mathbf{B}_{(\mathbf{w})}^{-1}$ is the background error covariance matrix in the simplified space

- the analysis is obtained as

$$\mathbf{x}^a(t_0) = \mathbf{x}^b(t_0) + S^{-I}\delta\mathbf{w}^a(t_0)$$

- the whole process (minimization and updating the background) can be repeated in the iterative process several times

- IMPORTANCE OF INCLUDING PHYSICS IN THE ASSIMILATING MODEL

– using an adiabatic linear model can be critical especially in the tropics, planetary boundary layer, stratosphere

– missing physical processes in assimilation systems can lead to the so-called spin-up problem

– using the adjoint of various physical processes should provide an initial atmospheric state for NWP models which is:

  ⋆ more consistent with the physical processes

  ⋆ producing a better agreement between the model and data

– inclusion of such processes is a necessary step towards the use of the observations related to the model physics (precipitation, cloudiness, etc.)

- DEVELOPMENT OF A PHYSICAL PACKAGE

  TRADE-OFF BETWEEN :

  Simple-regular and realistic physical package

  SIMPLE PHYSICAL PACKAGE FOR :

  – large-scale precipitation
  – gravity-wave drag
  – vertical diffusion
  – radiation
  – convection

**SURFACE PROCESSES ARE NOT TREATED !!!**

$$\Downarrow$$

**surface temperature**

**surface humidity**

**roughness length**

**snow depth**

**constant for any individual time step**

## STRATIFORM PRECIPITATION

- all supersaturation is immediately removed from the system
- everything is evaporated when a given layer is unsaturated
- distinction between solid and liquid type of precipitation done according temperature
  (if $T \geq 0^o C$, then the precipitation is liquid)

## GRAVITY WAVE DRAG

+ linear basic model for surface drag (Boer - 1984)
+ Lindzen saturation criterium for computation of momentum deposition
+ orographic anisotropy effects and simulation of mountain blocking effects
+ blocking of the low level flow

- only the linear part of computation of orographic gravity-waves effects on momentum

## VERTICAL DIFFUSION

+ exchange and drag coefficients for heat and momentum computed according to Louis (1982)

+ shallow convection - slight modification of Geleyn version (1987)

+ computation of $F_u$ and $F_v$ using two independent tridiagonal matrices

- assumption of constant fields for surface temperature $T_s$, surface humidity $q_s \Longrightarrow$ two independent matrices also for the computation of $F_{q_v}$, $F_s$

## CONVECTION

+ mass-flux type scheme with a closure assumption of Kuo-type

+ parametrization of liquid water/ice in the cloud ascent computation

- condition of instability - the level is convective if there is conditional instability and a non-zero mass flux + non-trigerring moist convection when there is dry convection at cloud base

- detrainment depends on the variation of the mass flux

- entrainment rate is used consistently with detrainment

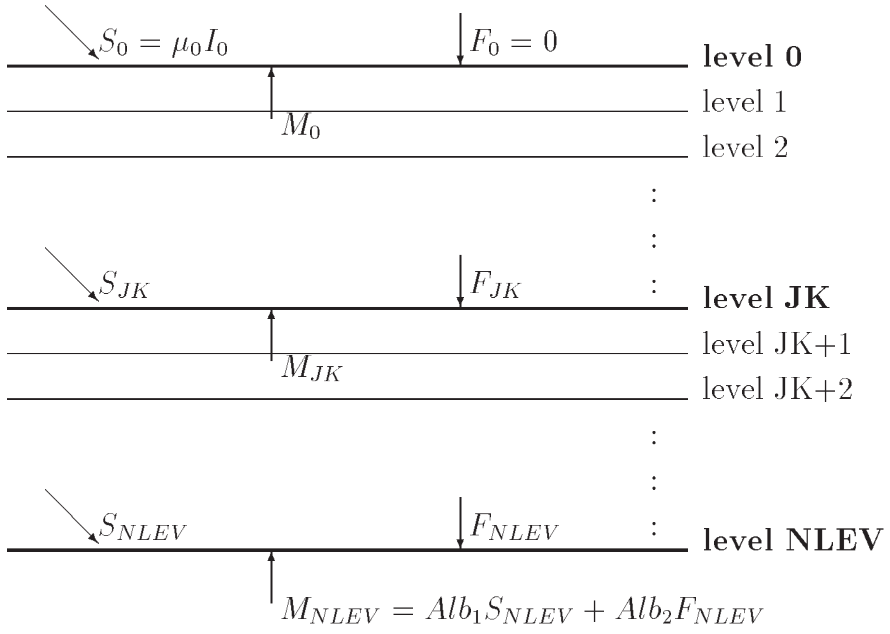
- treatment of the vertical turbulent fluxes of $s$ and $q$ separately for the diffusion and convection

RADIATION

- **assumption - the properties of gases are constant up to 24 hours**
- **coefficients for thermal radiation are recovered from the complete scheme integrated for one time step for clear sky**
- **coefficients are recomputed as a function of cloudiness in the simplified physical parametrization**
- **thermal radiation fluxes are obtained by using the Stefan-Boltzman law multiplied by the matrix of thermal radiation coefficients**
- **coefficients for solar radiation are obtained during one time step integration of full scheme for mean solar angle (for clear sky and cloudy conditions)**
- **dependency of the coefficients on actual cloudiness recomputed during the integration of the simplified model**
- **solar radiation fluxes are obtained using a two-layer simulated model**

## SOLAR RADIATION

- **representation of the fluxes with a two-layer model atmosphere**




- **solving the following set of equations:**

$$\begin{bmatrix} S_{JK} \\ F_{JK} \\ M_0 \end{bmatrix} = \begin{bmatrix} E_{a_1} & 0 & 0 \\ E_{a_2} & E_{a_4} & E_{a_5} \\ E_{a_3} & E_{a_6} & E_{a_4} \end{bmatrix} \cdot \begin{bmatrix} S_0 \\ 0 \\ M_{JK} \end{bmatrix}$$

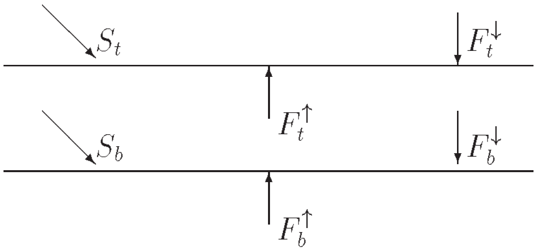
$$\begin{bmatrix} S_{NLEV} \\ F_{NLEV} \\ M_{JK} \end{bmatrix} = \begin{bmatrix} E'_{a_1} & 0 & 0 \\ E'_{a_2} & E'_{a_4} & E'_{a_5} \\ E'_{a_3} & E'_{a_6} & E'_{a_4} \end{bmatrix} \cdot \begin{bmatrix} S_{JK} \\ F_{JK} \\ M_{NLEV} \end{bmatrix}$$

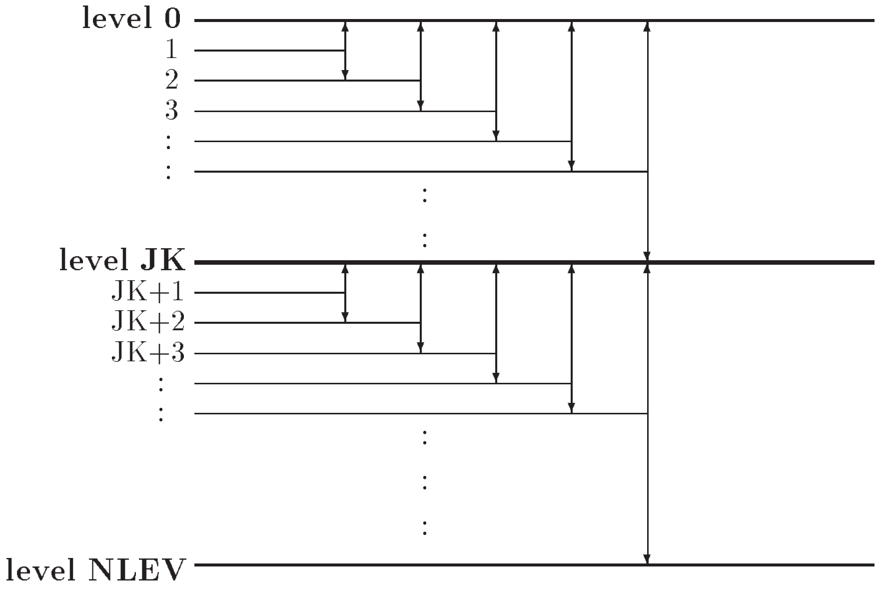
- **with boundary conditions:**
  at the top: $S_0 = \mu_0 I_0$, $F_0 = 0$
  at the bottom: $M_{NLEV} = Alb_1 S_{NLEV} + Alb_2 F_{NLEV}$

- **computation of $M_{NLEV}$ using one-layer model created by the two extreme levels (0 and NLEV)**

- **equivalent coefficients** $E_{a_1}, \ldots, E_{a_6}$ are obtained by solving the following system of matrices several times:

$$\begin{bmatrix} S_b \\ F_b^{\downarrow} \\ F_t^{\uparrow} \end{bmatrix} = \begin{bmatrix} a_1 & 0 & 0 \\ a_2 & a_4 & a_5 \\ a_3 & a_6 & a_4 \end{bmatrix} \cdot \begin{bmatrix} S_t \\ F_t^{\downarrow} \\ F_b^{\uparrow} \end{bmatrix}$$




- **dependency of the coefficients on cloudiness** is expressed as:

$$\left.\begin{aligned} E_{a_i} &= (1 - N_{JK})E_{a_i}^{C} + N_{JK}E_{a_i}^{N} \\ E'_{a_i} &= (1 - N'_{JK})E_{a_i}^{\prime C} + N'_{JK}E_{a_i}^{\prime N} \end{aligned}\right. ,i=1,\ldots,6$$

- **equivalent cloudiness** is computed using empirical formula:

$$N_{JK}/N'_{JK} = \alpha[1 - \prod_{K=1/JK+1}^{JK/NLEV}(1 - N_K)] + (1 - \alpha)\frac{\sum_{K=1/JK+1}^{JK/NLEV}(N_K \Delta p_K)}{\sum_{K=1/JK+1}^{JK/NLEV}\Delta p_K}$$

# 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:**

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

- **ADJOINT MODEL**

**The adjoint of a linear operator $M'$ is the linear operator $M^*$ such that, for the inner product $<,>$ and given vectors x and y it holds:**

$$\forall \mathbf{x}, \forall \mathbf{y} \qquad < M'.\mathbf{x}, \mathbf{y} >=< \mathbf{x}, M^*.\mathbf{y} >$$

## VALIDATION OF TL MODEL

– using the Taylor formula for the verification of a correctly prepared TL model:

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

– comparing TL simplified evolution of a perturbation $M'(\delta\mathbf{x})$ (of the order of magnitude of analysis increments) against two nonlinear runs runs using full physical parametrizations (one from from a basic state $M(\mathbf{x})$ and the other from a perturbed state $M(\mathbf{x} + \delta\mathbf{x})$)

# IMPORTANCE OF THE REGULARIZATION OF TL MODEL

- **physical processes are characterized by:**
  - **+ threshold processes:**
    - **– discontinuities of some functions describing the physical processes (some on/off processes)**
      (for instance, produced by supersaturation, changes between liquid and solid state)
    - **– discontinuities of the derivative of a continuous function**
      (i.e. the derivative can go towards infinity in some points)
  - **+ some strong non-linearities**
    (for instance, created by the transition from unstable to stable regimes in the PBL)

- **regularizations help to remove the most important threshold processes in physical parametrizations which can affect the range of validity of the tangent linear approximation**

# REGULARIZATION OF THE VERTICAL DIFFUSION SCHEME

- **original computation of the Richardson number** $Ri$

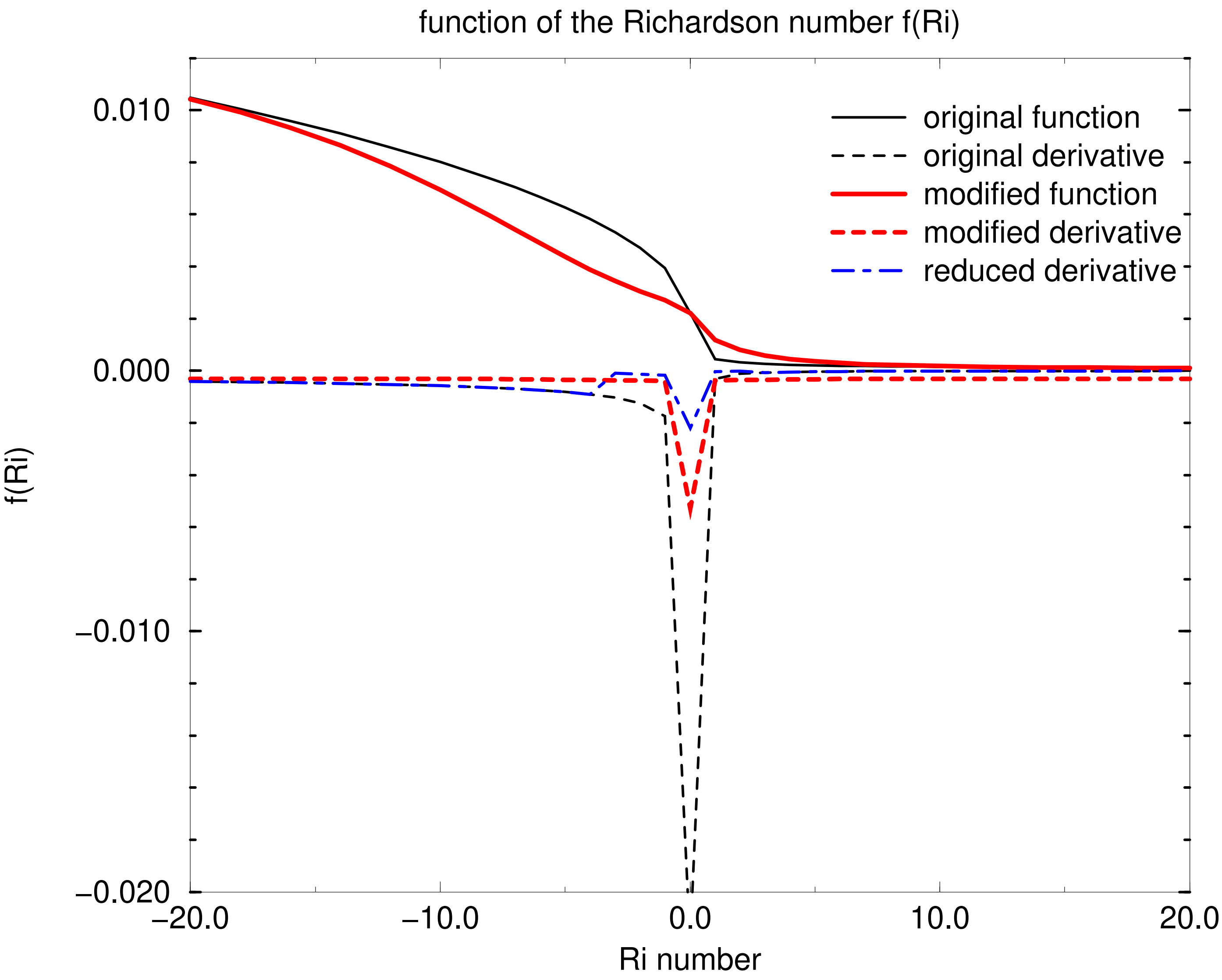
$$Ri = \frac{g}{c_p T} \frac{\frac{\partial s}{\partial z}}{\left\| \frac{\partial \vec{v}}{\partial z} \right\|^2}$$

**modified as**

$$Ri' = \frac{g}{c_p T} \frac{\frac{\partial s}{\partial z}}{\left\| \frac{\partial \vec{v}}{\partial z} \right\|^2 + c} \quad \textbf{with a time constant } c = \frac{a}{(\Delta t_{phys})^2}$$

where $\Delta t_{phys}$ is the physical time step
$a$ is a tuning parameter of the regularization step

- **reducing a derivative** $f(Ri)$ **by a factor 10 in the central part (around the point of singularity)**




## REGULARIZATION OF THE RADIATION SCHEME

**Problem:**

- strong noise due to plenty of possible thresholds in the cloudiness parametrization for radiation scheme

**Solution:**

- using zero perturbation of cloudiness (i.e., the constant cloudiness in a particular time step)

# REGULARIZATION OF THE CONVECTION SCHEME

**Problem:**

**- noise created by the unrealistic behaviour of the scheme when**

$$\mathbf{F}_{conv}=0 \not\Longleftrightarrow \mathbf{F}_{precip\_surf}=0$$

**Solution:**

**- re-introducing the condition**

$$\mathbf{F}_{conv}=0 \Longleftrightarrow \mathbf{F}_{precip\_surf}=0$$

**in order to remove the noise in the perturbation evolution using TL convection scheme compared to the finite differences between two nonlinear forecasts using the full physical parametrization**

$$\Downarrow$$

**adopted compromise:**

**removing the noise for the price of deteriorating the quality of the convergence in the Taylor test**

## 4D-VAR EXPERIMENTS

- **incremental 4D-Var assimilation**

  **full model:**

  - high resolution model (T95L27), stretched (c=3.5) and rotated
  - containing a complete physical parametrization

  **simplified linear model:**

  - lower horizontal resolution (T63) and non-stretched
  - adiabatic or containing simplified physical parametrizations

- **structure of the assimilation cycle:**

  - 6-hour assimilation cycles
  - the same quality control as in the former operational OI system
  - background term $\mathcal{J}^b$ identical to the operational 3D-Var one
  - minimization done using three passes in the outer loop with 2 times 25 iterations adiabatically and the last 20 iterations performed with:
    - ⋆ **adiab** - adiabatic model with a crude vertical diffusion and surface drag
    - ⋆ **rsphys** - reduced set of the simplified physics (vertical diffusion, gravity wave drag, large scale precipitation)
    - ⋆ **wsphys** - reduced set + convection + radiation
  - digital filter used only after the last update

.

- **comparison also with OI and 3D-Var**
- **validation on FASTEX type situations:**

IOP3 : assimilation from 13/1/97 to 14/1/97 + 72 hour forecast

IOP17 : assimilation from 17/2/97 to 18/2/97 + 72 hour forecast

one week assimilation cycle from 13 to 20 January 1997

## CONCLUSION

- a set of simplified, regular and realistic physical parametrization has been developed

- the validation of the simplified physical package has been done

$$\Downarrow$$

  although the main tendencies of the atmospheric evolution were caught, it is still necessary to devote some time to tune the developed simplified parametrizations (mainly convection and radiation)

- the validation of TL model following a linear evolution of a perturbation using the model with the simplified physical parametrizations has been done

- regularizations removing the most of important threshold processes in the physical parametrizations have been developed

$$\Downarrow$$

  the results indicate a clear advantage of the diabatic tangent linear evolution of errors compared to the adiabatic one

- **first 4D-Var assimilation experiments using simplified physical parametrization schemes have been performed**

$$\Downarrow$$

**4D-Var with simplified physics is able to represent better some cyclogenesis over North Atlantic, but generally the impact of the physics in 4D-Var is weak (maybe due to the short assimilation window)**

**in the tropics the results are not satisfactory when using 4D-Var**

- **combination of the incremental method and the inclusion of the simplified and regular physics in the assimilating model seems to be the way to improve the future operational data assimilation system at Météo-France**

# PERSPECTIVES

- **additional tunings and developments in the two more complex physical parametrizations (convection, radiation) should be done**
  *[weaknesses identified in the simplified physical parametrizations]*

- **some further investigation is required for the definition of the "optimal" 4D-Var set-up, which is likely to be meteorological situation dependent**
  *[a scenario has been defined with a limited number of experiments]*

- **impact of the initialization of the increments should also be studied**
  *[almost a priori choice has been made so far]*

- **future experiments should be done in order to study how the effects of physics become more important as the duration of the assimilation window is enlarged**
  *[experiments done up to now only with 6-hour assimilation window]*

## more generally:

- impact of using additional data related to the model physics (cloudiness, precipitation data, liquid water contents) in 4D-Var with simplified physics will be studied

- possible help of tuning of the complete physical parametrizations using parameter estimation will be investigated

- progressive inclusion of the tangent linear and adjoint of the simplified physical parametrizations should be extended to the limited area model ALADIN

# ONE-WEEK ASSIMILATION CYCLE

- experiments from 13 to 20 January 1997 using 3D-Var, 4D-Var_adiab and 4D-Var_wsphys
- forecasts performed up to 72 hours twice per day, from 00UTC and 12UTC
- comparisons of 72-hour forecasts in 12-hour intervals with OI analysis and some verifications against observations
- the verification of the forecast against analyses showed a general improvement of the forecast in Europe-Atlantic domain and a degradation in the low troposphere of the tropics when using 4D-Var
- the inclusion of the physics marginally improved the scores
- the study of how the assimilation systems are able to deal with the spin-up effect showed that this is a common problem of both 4D-Var systems (adiabatic or with the simplified physics)
- in the tropics some improvement was achieved when using 4D-Var_wsphys compared to 4D-Var_adiab