Parametrizations in Data Assimilation

- Introduction
- Why are physical parametrizations needed in data assimilation?
- Tangent-linear and adjoint coding
- Issues related to physical parametrizations in assimilation
- Physical parametrizations in ECMWF’s current 4D-Var system
- Examples of applications
- Summary and conclusions
Data assimilation

- Observations with errors
- \textit{a priori} information from model = background state with errors

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

Analysis

NWP model

Forecast
4D-Var produces the analysis ($x_a$) that minimizes the distance to a set of available observations ($y_o$) under the constraint of some *a priori* background information from the model ($x_b$) and given the respective errors of observations and model background.
Incremental 4D-Var minimizes the following cost function:

$$J = \frac{1}{2} \delta x_0^T B^{-1} \delta x_0 + \frac{1}{2} \sum_{i=1}^{n} (H_i M_i \delta x_0 - d_i)^T R_i^{-1} (H_i M_i \delta x_0 - d_i)$$

where: $i =$ time index (inside 4D-Var window, typ. 12h).  
$\delta x_0 = x_0 - x^b_0$ (increment).  
$H_i =$ tangent-linear of observation operator.  
$M_i =$ tangent-linear of forecast model $(t_0 \rightarrow t_i)$.  
$d_i = y^o_i - H_i (M_i [x^b_0])$ (innovation vector).  
$B =$ background error covariance matrix.  
$R_i =$ observation error covariance matrix.

$$\nabla_{x_0} J = B^{-1} \delta x_0 + \sum_{i=1}^{n} M^T_i [t_i, t_0] H^T_i R_i^{-1} (H_i M_i \delta x_0 - d_i)$$

Adjoint of forecast model with simplified linearized physics  
(simplified: to reduce computational cost and to avoid non-linear processes)
Why do we need physical parametrizations in DA?

Physical parametrizations are needed in data assimilation:

1) To evolve the model state in time during the 4D-Var assimilation,
2) To convert the model state variables to observed equivalents, → so that obs–model differences can be computed at obs time.

For example:

\[
\begin{bmatrix}
T \\
q_v \\
u \\
v \\
P_s \\
q_{liq} \\
q_{ice}
\end{bmatrix}_{t_0} \xrightarrow{M} \begin{bmatrix}
T \\
q_v \\
u \\
v \\
P_s \\
q_{liq} \\
q_{ice}
\end{bmatrix}_{t_i} \xrightarrow{H} \begin{bmatrix}
Rad_{ch1} \\
Rad_{ch2} \\
Rad_{ch3}
\end{bmatrix}
\]

\(M\) = forecast model with physics
\(H\) = radiative transfer model
Why do we need physical parametrizations in DA?

Tangent-linear operators are applied to perturbations:

\[ \delta x_0 = \begin{bmatrix} \delta T \\ \delta q_v \\ \delta u \\ \delta v \\ \delta P_s \end{bmatrix} \rightarrow \mathbf{M}[t_0,t_i]\frac{\partial}{\partial \delta x_0} \begin{bmatrix} \delta T \\ \delta q_v \\ \delta u \\ \delta v \\ \delta P_s \end{bmatrix} \rightarrow \mathbf{H} \rightarrow \delta \tilde{y}_i = \begin{bmatrix} \delta \text{Rad}_{ch1} \\ \delta \text{Rad}_{ch2} \\ \delta \text{Rad}_{ch3} \end{bmatrix} \]

Adjoint operators are applied to cost function gradient:

\[ \nabla \tilde{y}_i J_o = \begin{bmatrix} \frac{\partial J_o}{\partial \text{Rad}_{ch1}} \\ \frac{\partial J_o}{\partial \text{Rad}_{ch2}} \\ \frac{\partial J_o}{\partial \text{Rad}_{ch3}} \end{bmatrix} \rightarrow \mathbf{H}^T \rightarrow \mathbf{M}^T[t_i,t_0] \rightarrow \nabla x_0 J_o = \begin{bmatrix} \frac{\partial J_o}{\partial T} \\ \frac{\partial J_o}{\partial q_v} \\ \frac{\partial J_o}{\partial u} \\ \frac{\partial J_o}{\partial v} \\ \frac{\partial J_o}{\partial P_s} \\ \frac{\partial J_o}{\partial q_{\text{liq}}} \\ \frac{\partial J_o}{\partial q_{\text{ice}}} \end{bmatrix} \]
The choice of physical parametrizations will affect the results of 4D-Var $M$: input = model state $(T, q_v)$ $\rightarrow$ output = surface convective rainfall rate

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

from Marécal and Mahfouf (2002)

<table>
<thead>
<tr>
<th>Diagram 1</th>
<th>Diagram 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{\partial M}{\partial T} )</td>
<td>( \frac{\partial M}{\partial T} )</td>
</tr>
</tbody>
</table>
| \( \frac{\partial M}{\partial q_v} \) | \( \frac{\partial M}{\partial q_v} \)

Betts-Miller (adjustment scheme)

Tiedtke (ECMWF’s oper mass-flux scheme)
TEST FOR TANGENT LINEAR MODEL

- Taylor formula:

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

<table>
<thead>
<tr>
<th>Perturbation scaling factor</th>
<th>RATIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1E-09</td>
<td>0.9994875881543574E+00</td>
</tr>
<tr>
<td>0.1E-08</td>
<td>0.9999477148855701E+00</td>
</tr>
<tr>
<td>0.1E-07</td>
<td>0.9999949234236705E+00</td>
</tr>
<tr>
<td>0.1E-06</td>
<td>0.9999993501022509E+00</td>
</tr>
<tr>
<td>0.1E-05</td>
<td>0.99999999496119013E+00</td>
</tr>
<tr>
<td>0.1E-04</td>
<td>0.9999999611338369E+00</td>
</tr>
<tr>
<td>0.1E-03</td>
<td>0.9999993179193711E+00</td>
</tr>
<tr>
<td>0.1E-02</td>
<td>0.9999724488345042E+00</td>
</tr>
<tr>
<td>0.1E-01</td>
<td>0.9998727842790062E+00</td>
</tr>
<tr>
<td>0.1E+00</td>
<td>0.9978007454264978E+00</td>
</tr>
<tr>
<td>0.1E+01</td>
<td>0.9583066504549524E+00</td>
</tr>
</tbody>
</table>
The adjoint test should be correct at the level of machine precision (i.e. 13 to 15 digits, typically). If not, something is still wrong somewhere in the code!
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.
Linearity issue

\[ \Delta y (\text{tangent-linear}) = 0 \]

Threshold

\[ \Delta y (\text{non-linear}) \]

Tangent in \( x_0 \)

\[ \Delta x (\text{finite size perturbation}) \]

Precipitation formation rate

Cloud water amount
Illustration of discontinuity effect on cost function shape:

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

Simple parametrization of rain rate:

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

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

Dry background

Saturated background

Single minimum of cost function

Several local minima of cost function
REGULARIZATION OF VERTICAL DIFFUSION SCHEME

- perturbation of the exchange coefficients is neglected, \( K' = 0 \) (Mahfouf, 1999)
  \( ( \text{exchange coefficient } K \text{ is given by } K = l^2 \left\| \frac{\partial V}{\partial z} \right\| f(Ri)) \)

- original computation of the Richardson number \( Ri \)
  \[ Ri = \frac{g}{c_p T} \frac{\frac{\partial s}{\partial z}}{\left\| \frac{\partial \bar{u}}{\partial z} \right\|^2} \] modified as
  \[ Ri' = \frac{g}{c_p T} \frac{\frac{\partial s}{\partial z}}{\left\| \frac{\partial \bar{u}}{\partial z} \right\|^2 + c} \]

  with a time constant \( c = \frac{a}{(\Delta t_{phyys})^2} \)

  where \( \Delta t_{phyys} \) 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) - (Janisková et al., 1999)
function of the Richardson number $f(R_i)$

Janisková et al. 1999
Evolution of temperature increments (24-hour forecast) with the tangent linear model using different approaches for the exchange coefficient $K$ in the vertical diffusion scheme.

**Importance of regularization to prevent instabilities in tangent-linear model**

- Perturbations of $K$ included in TL
- Perturbations of $K$ set to zero in TL
Corresponding perturbations evolved with tangent-linear model

12-hour ECMWF model integration (T159 L60)

Temperature on level 48 (approx. 850 hPa)

Importance of regularization to prevent instabilities in tangent-linear model

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

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.)
\[ M(x + \lambda \delta x) - M(x) \approx \lambda M' \delta x \]

\[ <M' \delta x, \delta y> = <\delta x, M* \delta y> \]

**Timing:**
- Pure coding
- Debugging and testing (incl. regularization)

**4D-Var (minim)**

**Singular Vectors (EPS)**
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).
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]
  - 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.
  - mixed-layer parametrization and PBL top entrainment.
  - 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 ↔ TL evolution of initial perturbations

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

\[ M(x_{an}) - M(x_{bg}) \leftrightarrow M'(x_{an} - x_{bg}) \]

typical size of 4D-Var analysis increments

**Diagnostics:**

- mean absolute errors:
  \[ \varepsilon = \left| \frac{1}{n} \sum_{i=1}^{n} \left[ M(x_{an}) - M(x_{bg}) \right] - M'(x_{an} - x_{bg}) \right| \]

- relative error change:
  \[ \frac{\varepsilon_{EXP} - \varepsilon_{REF}}{\varepsilon_{REF}} \times 100\% \] (improvement if < 0)

- here: REF = adiabatic run (i.e. no physical parametrizations in tangent-linear)
Impact of operational vertical diffusion scheme

$$\varepsilon_{\text{EXP}} - \varepsilon_{\text{REF}}$$

REF = ADIAB

12-hour T159 L60 integration

Relative improvement [%]

Adiab simp vdf | vdf

EXP
Impact of dry + moist physical processes

$$\varepsilon_{\text{EXP}} - \varepsilon_{\text{REF}}$$

REF = ADIAB

12-hour T159 L60 integration

Relative improvement [%]

Adiab simp vdif | vdif + gwd + radold + lsp + conv

Temperature

Adiab simp vdif | vdif + gwd + radold + lsp + conv
Impact of all physical processes (including new moist physics & radiation)

\[ \varepsilon_{\text{EXP}} - \varepsilon_{\text{REF}} \]

REF = ADIAB

12-hour T159 L60 integration

Adiab simp vdif | vdif + gwd + radnew + cl_new + conv_new
Applications
1D-Var with radar reflectivity profiles

**Background**

\[ x_b = (T_b, q_b, \ldots) \]

\[ J_b = \frac{1}{2} (x - x_b)^T B^{-1} (x - x_b) \]

\[ \nabla_x J_b = B^{-1} (x - x_b) \]

**Initialize**

\[ x = x_b \]

**Moist Physics**

\[ x = (T, q, \ldots) \]

\[ \nabla_x J \approx 0? \]

**MINIM**

**Analysis**

\[ x_a = x \]

**Reflectivity Model**

\[ Z_{mod} \]

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

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

**Reflectivity observations**

\[ Z_{obs} \] with errors \[ \sigma_{obs} \]

**K** = number of model vertical levels
1D-Var with TRMM/Precipitation Radar data

Tropical Cyclone Zoe (26 December 2002 @1200 UTC; Southwest Pacific)
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.
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.

NHem: 700hPa temperature - Anomaly correlation

NHem: 200hPa vector wind - Anomaly correlation

> 0 = 😊

SHem: 700hPa temperature - Anomaly correlation

SHem: 200hPa vector wind - Anomaly correlation
**Idea:** It might be feasible to optimize the values of parameters used in the physical schemes with the variational data assimilation approach. This would require to include the parameter(s) in the control vector of the data assimilation system (4D-Var, for instance):

\[
J = \frac{1}{2} (x - x_b)^T B^{-1} (x - x_b) + \frac{1}{2} (p - p_b)^T B_p^{-1} (p - p_b)
\]

\[
+ \frac{1}{2} (H(x, p) - y_o)^T R^{-1} (H(x, p) - y_o)
\]

**Limitations:** Only parameters that are present in both the forecast model and the linearized simplified physics (TL & AD) can be treated in this way. Discrepancies between the full non-linear physics and the TL & AD physics (used in the minimization of \(J\)) might lead to sub-optimal results.
The validity of the linear assumption for precipitation quickly drops in the first hours of the forecast, especially for smaller scales.

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.

The validity of the linear assumption for precipitation quickly drops in the first hours of the forecast, especially for smaller scales.
Developing new simplified parametrizations for data assimilation requires:

- Compromise between realism, linearity and computational cost,
- Validation of simplified forward code against observations,
- Comparison to full non-linear code used in forecast mode (4D-Var trajectory),
- Numerical tests of tangent-linear and adjoint codes for small perturbations,
- Validity of the linear hypothesis for perturbations with larger size (typical of analysis increments) → no spurious perturbation growth.
- Successful convergence of 4D-Var minimizations.

Physical parametrizations have become important components in variational data assimilation systems because they allow:
- the comparison of the model state with observations (forward model),
- the minimization of the cost function (tangent-linear and adjoint models).

However, their linearized versions (tangent-linear and adjoint) require some special attention (regularizations/simplifications) in order to eliminate or smooth possible discontinuities and non-differentiability of the physical processes they represent.

This is particularly true for the assimilation of observations related to precipitation, clouds and soil moisture.

General conclusions

- Developing new simplified parametrizations for data assimilation requires:
  - Compromise between realism, linearity and computational cost,
  - Validation of simplified forward code against observations,
  - Comparison to full non-linear code used in forecast mode (4D-Var trajectory),
  - Numerical tests of tangent-linear and adjoint codes for small perturbations,
  - Validity of the linear hypothesis for perturbations with larger size (typical of analysis increments) → no spurious perturbation growth.
  - Successful convergence of 4D-Var minimizations.
Thank you!
Example of observation operator $H$ (radiative transfer model):

$$
\mathbf{x} = \begin{bmatrix}
T_1 \\
\vdots \\
T_n \\
q_1 \\
\vdots \\
q_n
\end{bmatrix}
\quad
\xrightarrow{H}
\quad
\mathbf{y} = \begin{bmatrix}
\text{Rad}_{ch1} \\
\text{Rad}_{ch2} \\
\text{Rad}_{ch3}
\end{bmatrix}
$$

and its tangent-linear operator $\mathbf{H}$:

$$
\mathbf{H} = \begin{bmatrix}
\frac{\partial \text{Rad}_{ch1}}{\partial T_1} & \ldots & \frac{\partial \text{Rad}_{ch1}}{\partial T_n} & \frac{\partial \text{Rad}_{ch1}}{\partial q_1} & \ldots & \frac{\partial \text{Rad}_{ch1}}{\partial q_n} \\
\frac{\partial \text{Rad}_{ch2}}{\partial T_1} & \ldots & \frac{\partial \text{Rad}_{ch2}}{\partial T_n} & \frac{\partial \text{Rad}_{ch2}}{\partial q_1} & \ldots & \frac{\partial \text{Rad}_{ch2}}{\partial q_n} \\
\frac{\partial \text{Rad}_{ch3}}{\partial T_1} & \ldots & \frac{\partial \text{Rad}_{ch3}}{\partial T_n} & \frac{\partial \text{Rad}_{ch3}}{\partial q_1} & \ldots & \frac{\partial \text{Rad}_{ch3}}{\partial q_n}
\end{bmatrix}
$$
• **Variational data assimilation:**
  Bouthier, F. and P. Courtier, 1999: Data assimilation concepts and methods, ECMWF lecture notes, available at [http://www.ecmwf.int/newsevents/training/lecture_notes/LN_DA.html](http://www.ecmwf.int/newsevents/training/lecture_notes/LN_DA.html).

• **The adjoint technique:**

• **Tangent-linear approximation:**

• **Physical parameterizations for data assimilation:**
• **Microwave Radiative Transfer Model for data assimilation:**

• **Assimilation of observations affected by precipitation and/or clouds:**


