TECHNICAL MEMORANDUM

666

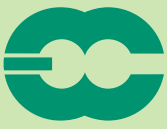
Linearized physics for data assimilation at ECMWF

M. Janisková and P. Lopez

Research Department

Submitted to Data Assimilation for Atmospheric, Oceanic and Hydrology Applications - Volume 2

January 2012



European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/publications/

Contact: library@ecmwf.int

European Centre for Medium-Range Weather Forecasts
Shinfield Park, Reading, RG2 9AX, England

Abstract

A comprehensive set of linearized physical parametrizations has been developed for the global ECMWF Integrated Forecasting System. Implications of the linearity constraint for any parametrization scheme, such as the need for simplification and regularization, are discussed. The description of the methodology to develop linearized parametrizations highlights the complexity of obtaining a physics package that can be efficiently used in practical applications. The impact of the different physical processes on the tangent-linear approximation and adjoint sensitivities, as well as their performance in data assimilation are demonstrated.

1 Introduction

Adjoint models have several applications in numerical weather prediction (NWP). In variational data assimilation (DA) for instance, they are used to efficiently determine optimal initial conditions. Another application of the adjoint technique is the computation of the fastest growing modes (i.e. singular vectors) over a finite time interval, which can be used in Ensemble Prediction Systems (EPS). Adjoint models can also be used for sensitivity studies since they enable the computation of the gradient of a selected output parameter from a numerical model with respect to all its input parameters. In practice, this is often used to obtain the sensitivity of the analysis to model parameters, sensitivities of one aspect of the forecast to initial conditions or sensitivities of the analysis to observations.

Initially, only the adiabatic linearized models were used in NWP. However, the significant role played by physical processes in various large-scale and mesoscale phenomena was soon recognized. Physical processes are particularly important in the tropics, near the surface, in the planetary boundary layer or the stratosphere, where the description of the atmospheric processes is controlled by both physics and dynamics. Therefore a lot of effort was devoted to include physical parametrizations in adjoint models. Several studies aimed at including physical parametrizations in adjoint models (Zou *et al.* 1993, Zupanski and Mesinger 1995, Tsuyuki 1996, Errico and Reader 1999, Janisková *et al.* 1999, Mahfouf 1999, Janisková *et al.* 2002, Laroche *et al.* 2002, Lopez 2002, Tompkins and Janisková 2004, Lopez and Moreau 2005, Mahfouf 2005) with encouraging results. However, these studies also showed that the linearization of physical parametrization schemes is not straightforward because of the non-linear and on/off nature of physical processes. Strong non-linearities that could lead to noise problems had to be removed from the models in order to be able to benefit from the inclusion of physical processes in the linearized model.

In recent years, four-dimensional variational (4D-Var) data assimilation became a powerful tool for exploiting information from irregularly distributed observations for initial conditions of a numerical forecast model. 4D-Var minimizes the distance between a model trajectory and observations spread over a given time interval, using the adjoint equations of the model to compute the gradient of the cost function with respect to the model state at the beginning of the assimilation period. The mismatch between model solution and observations can remain large if the imperfect adiabatic adjoint model would only be used in the minimization. Many satellite observations, such as radiances, rainfall and cloud measurements, cannot be directly assimilated with such overly simple adjoint models. Therefore it is crucial to represent physical processes in the assimilating models. Parametrization schemes for adjoint models started from very simple ones, such as Buizza (1994), which aimed at removing very strong increments produced by the adiabatic adjoint models. More complex, but still incomplete schemes were developed by Zou *et al.* (1993), Zupanski and Mesinger (1995), Janisková *et al.* (1999), Mahfouf (1999), Laroche *et al.* (2002), Mahfouf (2005). More recently, comprehensive schemes were implemented, which describe the whole set of physical processes and interactions between them al-

most as in the non-linear model, just slightly simplified and/or regularized (e.g. Janisková *et al.* 2002, Tompkins and Janisková 2004, Lopez and Moreau 2005).

In this paper, a comprehensive set of physical parametrizations developed for the linearized version of the global ECMWF model is described together with its applications in sensitivity studies and data assimilation. A description of the current package, which is unique because of its complexity, has never been published in the literature. Readers would only be able to find summaries of old parametrization schemes (Mahfouf 1999) from which hardly anything is left in the current operational model. Some information about updated versions of the schemes for shortwave radiation (Janisková *et al.* 2002) and moist processes (Tompkins and Janisková 2004, Lopez and Moreau 2005) is available, but is no longer up-to-date. In section 2, the reasons for using physics in variational data assimilation are explained. The implications of the linear constraint for any parametrization schemes, such as simplification and regularization are described in section 3. The methodology for the development of linearized simplified parametrizations is provided in section 4. To fully appreciate the achieved level of sophistication of the linearized physical parametrization schemes used at ECMWF, which can still be integrated even over 48 hours on the global scale without producing spurious noise, each of them is described in section 5. The impact of different physical processes on the tangent-linear approximation, adjoint sensitivity, as well as the performance in data assimilation are demonstrated in section 6. Finally, conclusions and perspectives are given in section 7.

2 The need for physics in variational data assimilation

Two main reasons can justify the need for linearized physical parametrizations in variational data assimilation.

The first one lies in the necessity to compute model−observation departures at a given time, so that the variational cost function can be minimized. For instance, if satellite microwave brightness temperatures are to be assimilated, one must be able to translate the model control variables (typically temperature, humidity, wind and surface pressure) into some equivalent simulated brightness temperatures. In this example, this can be achieved by applying moist physics parametrizations to simulate cloud and precipitation fields first, and then a radiative transfer model to obtain the desired microwave brightness temperatures, as seen by the model. The goal of data assimilation is to define the atmospheric state such that the mismatch between the model and observations (or cost function, J) is minimum. To minimize the cost function for obtaining the optimal increments in each model state vector component, its gradient with respect to model variables needs to be assessed. In the chosen example of microwave brightness temperatures, this would be achieved by applying the adjoint of the radiative transfer model followed by the adjoint of the moist physical parametrizations to the gradient of J in observation space. The adjoint of a given operator is simply the transpose of its Jacobian matrix with respect to its input variables.

Secondly, in the particular context of 4D-Var data assimilation, the model state needs to be compared to each available observation at the time the latter was performed. It is therefore necessary to evolve the model state from the beginning of the 4D-Var assimilation window (time 0) to the time of the observation (time i). This is achieved by integrating the full non-linear (NL) forecast model, M, from time 0 to time i. Again, the minimization of the 4D-Var cost function, J, which measures the total distance between the model and all observations available throughout the assimilation window, requires the computation of its gradient, $\nabla_{\mathbf{x}(t_0)} J$ with respect to the model state at the beginning of the 4D-Var assimilation window, $\mathbf{x}(t_0)$. To achieve this, the gradient of the observation term, J_o, of the cost function in observation space

can be first computed through simple differentiation as

$$\nabla_{\tilde{\mathbf{y}}_i} J_o \;=\; \sum_{i=0}^{n} \mathbf{R}_i^{-1}(\tilde{\mathbf{y}}_i - \mathbf{y}_i^o) \tag{1}$$

where $\tilde{\mathbf{y}}_i = H(\mathbf{x}_i)$ is the model observed equivalent, $\mathbf{y}_i^o$ is the vector of available observations and $\mathbf{R}_i$ is the observation error covariance matrix. Using the adjoint of the observation operator, $\mathbf{H}_i^T$, one can then calculate the gradient of J_o with respect to the model state at observation time, $\mathbf{x}(t_i)$,

$$\nabla_{\mathbf{x}_i} J_o \;=\; \sum_{i=0}^{n} \mathbf{H}_i^T \mathbf{R}_i^{-1}(H_i[\mathbf{x}(t_i)] - \mathbf{y}_i^o) \tag{2}$$

Finally, the gradient of J_o with respect to the model state at time 0 can be obtained by applying the adjoint (AD) of the forecast model, $\mathbf{M}^T(t_i, t_0)$,

$$\nabla_{\mathbf{x}(t_0)} J_o \;=\; \sum_{i=0}^{n} \mathbf{M}^T(t_i, t_0)\mathbf{H}_i^T \mathbf{R}_i^{-1}(H_i[\mathbf{x}(t_i)] - \mathbf{y}_i^o) \tag{3}$$

Again, since the adjoint version of the forecast model can be seen as the transpose of its Jacobian matrix, the forecast model first needs to be differentiated with respect to its inputs, yielding the so-called tangent-linear (TL) model, $\mathbf{M}$.

In contrast with the full non-linear model, the tangent-linear model works on perturbations of the input variables rather than on full model fields and is fully linear by construction. The adjoint is therefore a fully linear operator as well and, in the case of 4D-Var, its inputs are the components of $\nabla_{\mathbf{x}(t_i)} J$. As a consequence, solving the 4D-Var minimization requires the linearization of the forecast model's physical parametrizations (e.g. vertical diffusion, radiation, convection, large-scale moist processes) so that their TL and AD versions can be used to describe the (forward, respectively backward) time evolution of the model state during the minimization as seen from Eq. (3).

3 Implication of the linearity constraint

The minimization of the 4D-Var cost function is solved with an iterative algorithm and is therefore computationally rather demanding. Even though the minimization is usually performed at a much lower resolution (T159/T255[1] in current ECMWF's operations) than in the standard forecast model (T1279[2] at ECMWF), the several tens of iterations required to obtain the optimal model state means that the linearized physics package must be as cheap as possible. To reduce computational cost, it is therefore often necessary to simplify the set of linearized parameterizations by retaining only physical processes that dominate in the full forecast model. Linearity considerations can also influence this choice: if a given process is known to be highly non-linear (e.g. thresholds, switches), this process should be discarded from the linearized code since this might otherwise lead to instabilities during TL and AD integrations. However, some of those instabilities can be overcome through adequate modifications of the code. At the same time, though simplified, parametrization schemes used in the linearized model must remain realistic enough to keep the description of atmospheric processes physically sound.

[1] T159/T255 corresponding approximately to 130km/80km
[2] T1279 corresponding approximately to 16 km

3.1 Simplification

For important practical applications (incremental approach of 4D-Var - Courtier *et al.* 1994, adjoint based sensitivities, initial perturbations of EPS), adjoint based sensitivities, initial perturbations of EPS), the linearized version of the forecast model is run at a lower resolution than the non-linear model. In this case, since the dynamics is already simplified through the reduction in horizontal resolution, the linearized physics does not necessarily need to be exactly tangent to the full physics. In principle, physical parametrizations can already behave differently between non-linear and tangent-linear models due to the change in resolution. Consequently, some freedom exists in the development of a simplified physics package, as long as the parametrizations can represent general feedbacks of physical phenomena present in the atmosphere. Simplified approaches can allow the progressive inclusion of physical processes in the tangent-linear and adjoint models. This strategy has been used, for instance, in the operational 4D-Var systems of ECMWF (Mahfouf 1999, Mahfouf and Rabier 2000, Rabier *et al.* 2000, Janisková *et al.* 2002, Janisková 2003, Tompkins and Janisková 2004, Lopez and Moreau 2005) and at Météo-France (Janisková *et al.* 1999, Geleyn *et al.* 2001).

3.2 Regularization

As already mentioned, physical processes are often characterized by thresholds. These can be:

- discontinuities of some functions themselves describing the physical processes or some on/off processes (for instance produced by saturation, changes between liquid and solid phase);
- some discontinuities in the derivative of a continuous function (i.e. the derivative can go towards infinity at some points);
- some strong non-linearities (such as those created by the transition from unstable to stable regimes in the planetary boundary layer).

In each of these situations, an estimation of the derivative close to the discontinuity point will be different between the non-linear model (in terms of finite differences) and the TL model. All of this makes the tangent linear approximation less valid when the linearized model includes physical parametrizations compared to the adiabatic version only. To treat the described problems, it is important to regularize, i.e. to smooth the parametrized discontinuities in order to make the scheme as much differentiable as possible. One should recognize that it is often quite difficult to achieve a tradeoff between a physically sound description of atmospheric processes and a well-behaved linear physical parametrization. However, without a proper treatment of the most significant thresholds, the TL model can quickly become too inaccurate to be useful. Therefore a lot of effort was devoted by a number of investigators to deal with discontinuities present in parametrized physical processes (e.g. Zou *et al.* 1993, Zupanski and Mesinger 1995, Tsuyuki 1996, Errico and Reader 1999, Janisková *et al.* 1999, Mahfouf 1999, Laroche *et al.* 2002, Tompkins and Janisková 2004, Lopez and Moreau 2005).

To illustrate a potential source of problem in the linearized model, the rain production function, describing which portion of the cloud water is converted into precipitation, is shown in Fig. 1. An increase of cloud water mixing ratio by a small amount dx (Fig. 1a) leads to a small change in the precipitation amount dy_{NL} in the case of the non-linear (NL) model, but to a much larger change (dy_{TL}) in the case of the TL model. As a possible solution, one can modify the function to make it less steep (dotted line on Fig. 1b). In this case, the resulting TL increment will be significantly smaller (dy_{TL_2}). However, the required modification can be substantial and it can deteriorate the overall quality of the physical parametrization itself. Therefore one must always be careful to keep the right balance between linearity and realism of the parametrization schemes. In the future, the better the non-linear forecast model will

become, the smaller 4D-Var analysis increments and hence the hope to have less difficulties with using linearized physical processes will be.

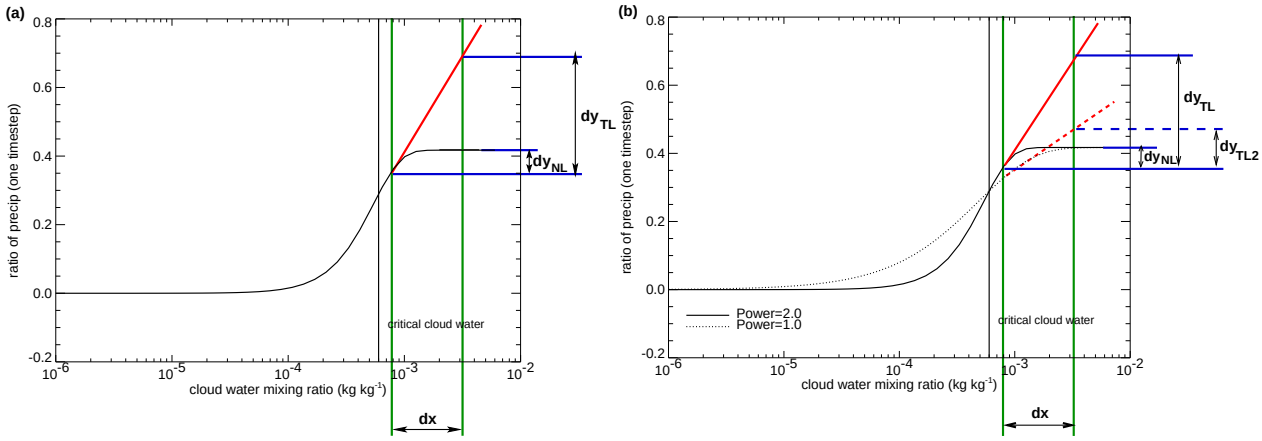


Figure 1: Autoconversion function of cloud water into precipitation (black solid line) based on Sundqvist et al. (1989). A change in the cloud water, dx, results in a change of precipitation, dy_{NL}, in the case of non-linear (NL) model. dy_{TL} is the corresponding change in precipitation given by the tangent-linear (TL) model. (b) describes the modified function which is less steep and helps to reduce the TL increments to dy_{TL_2} (closer to dy_{NL}).

4 Methodology for the development of linearized simplified parametrizations

There are several problems with including physics in adjoint models. The development requires substantial resources and it is technically very demanding. The validation must be very thorough and it must be done for the non-linear, tangent-linear and adjoint versions of the physical parametrization schemes. The computational cost of the model with physical processes can be very high despite some possible simplifications applied. One must be also very careful with the non-linear and threshold nature of physical processes which can affect the range of validity of the tangent-linear approximation as mentioned above.

The development of a new linearized physical parametrization can be divided into four main stages:

(1) Simplified non-linear forward model design, coding, tuning and validation.
(2) Tangent-linear coding and testing.
(3) Adjoint coding and testing.
(4) Performance assessment in data assimilation and other applications (see examples in section 6).

4.1 Simplified non-linear version

In the first stage, the non-linear version of the new simplified physical parametrization needs to be designed. This can be achieved through either an ”upward” or ”downward” approach, both relying on the prioritization of the various processes represented in the full NL code used in standard forecasts. With the upward technique, the simplified code is obtained by keeping only the most relevant processes found in the full NL version. In the downward approach, the simplified code is built by ignoring the least significant processes from the full NL code. Ideally, both approaches should converge to more or less

similar simplified codes, which should be computationally cheaper than the full NL code, and contain fewer discontinuities but are still able to provide realistic forecasts. Once the simplified code has been written, it is thus necessary to tune and validate it in traditional forecasts over periods at least equal to the maximum length of the expected applications. At ECMWF for instance, this period corresponds to 12 hours for 4D-Var DA or to 24 hours for singular vector computations involved in the ensemble prediction system. It is particularly crucial to ensure that the new simplified NL code does not depart too much from its full NL counterpart over this period of time. Verification in much longer integrations (up to climate timescales), although not essential, is also recommended to make sure that the new simplified scheme is stable and behaves reasonably well.

4.2 Linearization techniques

Once the NL version of the simplified scheme is deemed adequate, efforts are devoted to the development of the TL code, first, and then of the AD code. In practice, linearization can be achieved using either a manual line-by-line approach or an automatic coding software (e.g. Giering and Kaminski 1998, Araya-Polo and Hascoët 2004). However attractive automatic coding may sound, the manual technique is usually more suitable as soon as one has to deal with the large amounts of complex code used in modern NWP systems. Until now, in our own experience, the code produced through automatic differentiation and adjointing often turned out to be computationally very expensive (no optimization) and sometimes not bug-free. This is the reason why so far only manual line-by-line TL and AD coding has been applied to derive and update ECMWF's full set of linearized physical parametrizations. In the future this strategy might be revisited if automatic softwares become more efficient and reliable.

4.3 Tangent-linear version

An estimation of sensitivity of model output with respect to input required by many studies can be efficiently done by using the adjoint. For atmospheric models evolving in time, this backward integration requires to have the tangent linear model acting forward in time. To build the TL model, the linearization is performed with respect to the local tangent of the model trajectory.

If M is the model describing the time evolution of the model state $\mathbf{x}$ as:

$$\mathbf{x}(t_{i+1}) = M[\mathbf{x}(t_i)] \tag{4}$$

then the time evolution of a small perturbation $\delta\mathbf{x}$ can be estimated to the first order approximating by the tangent linear model $\mathbf{M}$ (derived from the NL model M):

$$\begin{aligned} \delta\mathbf{x}(t_{i+1}) &= \mathbf{M}[\mathbf{x}(t_i)]\delta\mathbf{x}(t_i) \\ \delta\mathbf{x}(t_{i+1}) &= \frac{\partial M[\mathbf{x}(t_i)]}{\partial \mathbf{x}}\delta\mathbf{x}(t_i) \end{aligned} \tag{5}$$

The verification of the correctness of the TL model is first performed through the classical Taylor formula:

$$\lim_{\lambda\to 0} \frac{M(\mathbf{x}+\lambda\,\delta\mathbf{x}) - M(\mathbf{x})}{\mathbf{M}(\lambda\,\delta\mathbf{x})} = 1 \tag{6}$$

This examination of asymptotic behaviour, using perturbations the size of which becomes infinitesimally small, is performed to check the numerical correctness of the TL code.

For practical applications, it is also important to investigate the accuracy of TL models for finite-amplitude perturbations (typically perturbations of the size of analysis increments). The results from applications of tangent-linear and adjoint models are only useful when the linearized approximation is valid for such perturbations. Therefore, for the validation of the tangent-linear approximation, the accuracy of the linearization of a parametrization scheme is studied with respect to pairs of non-linear results. The difference between two non-linear integrations (one starting from a background field, $\mathbf{x}^b$, and the other from an analysis, $\mathbf{x}^a$) run with the full NL model, M, is compared to time evolution of the analysis increments ($\mathbf{x}^a - \mathbf{x}^b$) obtained by integrating the TL model, $\mathbf{M}$, with the trajectory taken from the background field.

For a quantitative evaluation of the impact of linearized schemes, their relative importance is evaluated using mean absolute errors between tangent-linear and non-linear perturbations as:

$$\varepsilon = \overline{\left| \mathbf{M}(\mathbf{x}^a - \mathbf{x}^b) - \left[M(\mathbf{x}^a) - M(\mathbf{x}^b) \right] \right|} \tag{7}$$

As a reference for the comparisons, an absolute mean error for the TL model without physics, ε_{ref}, is taken. If ε_{exp} is defined as the absolute mean error of the TL model with the different physical schemes included, then an improvement coming from the inclusion of more physics in the TL model is expressed as $\varepsilon_{exp} < \varepsilon_{ref}$. The relative errors, r_{er}, and relative improvements, η, are also computed as:

$$r_{er} = \frac{\left| \mathbf{M}(\mathbf{x}^a - \mathbf{x}^b) - \left[M(\mathbf{x}^a) - M(\mathbf{x}^b) \right] \right|}{\left| M(\mathbf{x}^a) - M(\mathbf{x}^b) \right|} \tag{8}$$

$$\eta = \frac{\varepsilon_{exp} - \varepsilon_{ref}}{\varepsilon_{ref}} \tag{9}$$

Validity tests of the tangent-linear approximations are mostly performed over the time period and at the resolution at which adjoint models will be applied in practice: resolution and time length of an inner-loop integration of 4D-Var system (e.g. 12 hours, T255 and 91 vertical levels at ECMWF) or longer time periods for singular vectors and sensitivity applications (e.g. 24 hours at ECMWF). An example of the results from such TL approximation assessment will be given in section 6.1.

4.4 Adjoint version

The adjoint of a linearized operator, $\mathbf{M}$, is the linear operator, $\mathbf{M}^*$, such that:

$$\forall \mathbf{x}, \forall \mathbf{y} \qquad < \mathbf{M}.\mathbf{x}, \mathbf{y} > = < \mathbf{x}, \mathbf{M}^*.\mathbf{y} > \tag{10}$$

where $<,>$ denotes the inner product and $\mathbf{x}$ and $\mathbf{y}$ are input vectors.

Besides, the adjoint model $\mathbf{M}^*$ can provide the gradient of any objective function, J, with respect to $\mathbf{x}(t_i)$ from the gradient of the objective function with respect to $\mathbf{x}(t_{i+1})$

$$\frac{\partial J}{\partial \mathbf{x}(t_i)} = \mathbf{M}^* \left(\frac{\partial J}{\partial \mathbf{x}(t_{i+1})} \right) \tag{11}$$

The integration of the AD forecast model works backward in time. One should remember that, M being non-linear, $\mathbf{M}$ as well as $\mathbf{M}^*$ depend on the particular state of the atmosphere, $\mathbf{x}$, about which the linearization is performed. The adjoint operator, for the simplest canonical scalar product $<,>$ (Eq. 10), is actually the transpose of the tangent linear operator, $\mathbf{M}^T$ (not its inverse).

For the practical verification of the adjoint code, one must test the identity described in Eq. (10). It should be emphasized that it is absolutely essential to ensure that the TL and AD codes verify Eq. (10) to the level of machine precision, even when vectors **x** and **y** are global 3D atmospheric states and even for time integrations up to 12 or 24 hours. Note that a correct adjoint test does not imply the correctness of tangent-linear code.

4.5 Singular vectors

Besides the verification of the numerical correctness of TL and AD versions of the model and the examination of the validity of TL approximation, singular vectors can be used to find out whether the new schemes do not lead to a growth of spurious unstable modes. Such modes would indicate the existence of strong non-linearities and threshold processes in the model and would have a negative impact on the usefulness of the linearized model.

5 ECMWF's linearized physics package

5.1 Description

The set of ECMWF physical parametrizations used in the linearized model (called simplified or linearized parametrizations) comprises six different schemes for: radiation, vertical diffusion, orographic gravity wave drag, moist convection, large-scale condensation/precipitation and non-orographic gravity wave activity , sequentially called in this order. The current linearized physics package is therefore quite sophisticated and is believed to be the most comprehensive one used in operational global data assimilation. Each physical parametrization scheme of this package is described below starting with dry processes.

5.1.1 Radiation

The radiation scheme solves the radiative transfer equation in two distinct spectral regions. The computations for the longwave (LW) radiation are performed over the spectrum from 0 to 2820 cm^{-1} ($\sim$ 100 μm to 3.5 μm). The shortwave (SW) part of the scheme integrates the fluxes over the whole shortwave spectrum between 0.2 and 4.0 μm. The scheme used for data assimilation purposes must be computationally efficient to be called at full spatial resolution to improve the description of cloud-radiation interactions during the assimilation period (Janisková *et al.* 2002).

The shortwave radiation scheme:

The linearized code for the SW radiation scheme has been derived from ECMWF's original non-linear scheme developed by Fouquart and Bonnel (1980) and revised by Morcrette (1991). In this scheme, previously used in the operational forecast model, the photon-path-distribution method is applied to separate the parametrization of scattering processes from that of molecular absorption. Upward $F^{\uparrow}_{\rm sw}$ and downward $F^{\downarrow}_{\rm sw}$ fluxes at a given level j are obtained from the reflectance and transmittance of the atmospheric layers as

$$F^{\downarrow}_{\rm sw}(j) = F_0 \prod_{k=j}^{N} T_{\rm bot}(k) \tag{12}$$

$$F^{\uparrow}_{\rm sw}(j) = F^{\downarrow}_{\rm sw}(j) R_{\rm top}(j-1) \tag{13}$$

Computations of the transmittance at the bottom of a layer, $T_{\rm bot}$, start at the top of atmosphere and work downwards. Those of the reflectance at the top of the same layer, $R_{\rm top}$, start at the surface and work upwards. In the presence of cloud in the layer, the final fluxes are computed as a weighted average of the fluxes in the clear-sky and in the cloudy fractions of the column (depending on the cloud-overlap assumption).

The non-linear scheme is reasonably fast for application in 4D-Var and has therefore been linearized without a-priori changes (Janisková *et al.* 2002). The only modification with respect to the non-linear version (used operationaly until June 2007; since then Rapid Radiation Transfer model for SW radiation is used - Mlawer and Clough 1997), is the use of two spectral intervals, instead of six intervals. This is meant to reduce the computational cost.

The longwave radiation scheme:

The LW radiation scheme, used in the ECMWF full NL forecast model is the Rapid Radiation Transfer Model (RRTM; Mlawer *et al.* 1997, Morcrette *et al.* 2001). The complexity of the RRTM scheme for the LW part of the spectrum makes accurate computations expensive. In the variational assimilation framework, the older operational scheme of Morcrette (1989) was linearized. In this scheme, the LW spectrum from 0 to 2820 cm^{-1} is divided into six spectral regions. The transmission functions for water vapour and carbon dioxide over those spectral intervals are fitted using Padé approximations on narrow-band transmissions obtained with statistical band models (Morcrette *et al.* 1986). Integration of the radiation transfer equation over wavenumber ν within the particular spectral regions yields the upward and downward fluxes.

The inclusion of cloud effects on the LW fluxes follows the treatment discussed by Washington and Williamson (1997). The scheme first calculates upward and downward fluxes ($F_0^{\uparrow}(i)$ and $F_0^{\downarrow}(i)$) for a clear-sky atmosphere. In any cloudy layer, the scheme evaluates the fluxes assuming a unique overcast cloud of emissivity unity, i.e. $F_n^{\uparrow}(i)$ and $F_n^{\downarrow}(i)$ for a cloud present in the nth layer of the atmosphere. The fluxes for the actual atmosphere are derived from a linear combination of the fluxes calculated in the previous steps with some cloud overlap assumption (see below) in the case of clouds spreading over several layers. If N is the number of model layers starting from the top of atmosphere to the bottom, C_i the fractional cloud cover in layer i, the cloudy upward $F_{lw}^{\uparrow}$ and downward $F_{lw}^{\downarrow}$ fluxes are then expressed as:

$$F_{lw}^{\uparrow}(i) = (1-CC_{N,i})F_0^{\uparrow}(i) + \sum_{k=i}^{N}(CC_{i,k+1}-CC_{i,k})F_k^{\uparrow}(i) \tag{14}$$

$$F_{lw}^{\downarrow}(i) = (1-CC_{i-1,0})F_0^{\downarrow}(i) + \sum_{k=1}^{i-1}(CC_{i,k+1}-CC_{i,k})F_k^{\downarrow}(i) \tag{15}$$

where $CC_{i,j}$ is the cloudiness encountered between any two levels i and j in the atmosphere computed using the overlap assumption described below.

In the case of semi-transparent clouds, the fractional cloudiness entering the calculations is an effective cloud cover equal to the product of the emissivity due to condensed water and gases in the layer by the horizontal coverage of the cloud cover. This is the so called effective emissivity approach.

To reduce a computational cost of the linearized LW radiation for data assimilation, the scheme is not called at each time step. Furthermore, the transmission functions are only computed for H_2O and CO_2 absorbers (though the version taking into account the whole spectrum of absorbers is also coded for aerosols and other gases. The cloud effects on LW radiation are only computed up to cloud top.

Cloud overlap assumptions:

Cloud overlap assumptions must be made in atmospheric models in order to organize the cloud distribution used for radiation. This is is necessary to account for the fact that clouds often do not fill the whole grid box. The maximum-random overlap assumption (originally introduced in Geleyn and Hollingsworth 1997) is used operationally (Morcrette and Jakob 2000). Adjacent cloudy layers are combined by assuming maximum overlap to form a contiguous cloud and discrete layers separated by clear-sky are combined randomly.

Cloud optical properties:

When one considers cloud-radiation interactions, it is not only the cloud fraction or cloud volume, but also cloud optical properties that matter. In the case of SW radiation, cloud radiative calculations depend on three different parameters: the optical thickness, the asymmetry factor and the single scattering albedo. They are derived from Fouquart (1987) for water clouds, and Ebert and Curry (1992) for ice clouds. They are functions of cloud condensate and a specified effective radius.

Cloud LW optical properties are represented by the emissivity, related to the condensed water amount, and by the condensed-water mass absorption coefficient obtained from Smith and Shi (1992) for water clouds and Ebert and Curry (1992) for ice clouds.

5.1.2 Vertical diffusion

Vertical diffusion applies to wind components, dry static energy and specific humidity. The exchange coefficients in the planetary boundary layer and the drag coefficients in the surface layer are expressed as functions of the local Richardson number, Ri, (Louis *et al.* 1982). They differ from the formulation of the operational forecast model (i.e. full non-linear scheme) where for the unstable regime ($Ri < 0$) the Monin-Obukhov (M-O) formulation is used, together with a K-profile approach for convectively mixed layer in the case of unstable surface conditions. In the linearized model, the exchange coefficients are computed according to Louis *et al.* (1982). For the stable regime ($Ri > 0$), diffusion coefficients according to the Louis scheme are used close to the surface and above 300 m, then they tend asymptotically to the M-O formulation. A mixed layer parametrization is also included. This is consistent with the full non-linear model.

Analytical expressions are generalized for the situation with different roughness lengths for heat and momentum transfer. For any conservative variable ψ (wind vector components, u and v; dry static energy, s; specific humidity, q), the tendency produced by vertical diffusion is

$$\frac{\partial \psi}{\partial t} = \frac{1}{\rho}\frac{\partial}{\partial z}\left(K(Ri)\frac{\partial \psi}{\partial z}\right) \tag{16}$$

where ρ is the air density. The exchange coefficient K for heat and momentum transfer is given by

$$K = l^2 \left\| \frac{\partial \mathbf{U}}{\partial z} \right\| f(Ri) \tag{17}$$

where $\mathbf{U}$ is the wind vector and $f(Ri)$ represents the coefficient accounting for the dependence of vertical turbulent diffusion on the local Richardson number, either computed according to Louis *et al.* (1982), $f_L(Ri)$, or to the Monin-Obukhov formulation, $f_{MO}(Ri)$. l is the mixing length profile based on the formulation of Blackadar (1962) with a reduction in the free atmosphere.

A continuous transition between Louis coefficients near the surface to about 300 m and M-O coefficients above is computed as

$$\frac{1}{l\sqrt{f(Ri)}} = \frac{1}{\kappa z\sqrt{f_L(Ri)}} + \frac{1}{\lambda\sqrt{f_{MO}(Ri)}} \tag{18}$$

where κ is the Von Karman's constant, z is the height and λ is the asymptotic mixing length.

To parametrize turbulent fluxes at the surface, the drag coefficient, C_{sf}, (i.e. the exchange coefficient between the surface and the lowest model level) is computed as

$$C_{sf} = g_{sf}(Ri)\, C_{\mathrm{N}} \tag{19}$$

where C_{N} is the neutral drag coefficient, which is a function of the roughness length, and $g_{sf}(Ri)$ is a function of the local Richardson number. Different formulations of C_{N} and $g_{sf}(Ri)$ are used for momentum and heat, according to Louis *et al.* (1982).

Regularization:

In earlier versions of the model, perturbations of the exchange coefficients were simply neglected ($K' = 0$), in order to prevent spurious unstable perturbations from growing in the linearized version of the scheme (Mahfouf 1999). Later, some regularization of exchange coefficients was introduced at upper model levels to allow partial perturbations of these coefficients. This consists in the perturbations being more significantly reduced around the neutral state (i.e. Ri close to zero) where both the function of Ri itself and its derivative exhibit a significant rate of change. The reduction is eased exponentially away from the neutral state.

5.1.3 Subgrid scale orographic effects

Only the low-level blocking part of the operational non-linear scheme developed by Lott and Miller (1997) is taken into account in TL and AD calculations. The deflection of the low-level flow around orographic obstacles is supposed to occur below an altitude Z_{blk} such that

$$\int_{Z_{\mathrm{blk}}}^{3\mu} \frac{N}{|\mathbf{U}|}\,\mathrm{d}z \geq H_{n_{\mathrm{crit}}} \tag{20}$$

where $H_{n_{\mathrm{crit}}}$ is a critical non-dimensional mountain height, μ is the standard deviation of subgrid-scale orography and N is the Brunt-Väisälä frequency.

The deceleration of wind due to low-level blocking is given by

$$\left(\frac{\partial \mathbf{U}}{\partial t}\right)_{\mathrm{blk}} = -C_{\mathrm{d}} \max\left(2-\frac{1}{r},0\right)\frac{\sigma}{2\mu}\sqrt{\frac{Z_{\mathrm{blk}}-z}{z+\mu}}\,(B\cos^2\alpha + C\sin^2\alpha)\,\frac{\mathbf{U}|\mathbf{U}|}{2} \tag{21}$$

where C_{d} is the low-level drag coefficient, σ is the mean slope of the subgrid-scale orography, and α is the angle between the low-level wind and the principal axis of orography. r is determined as $r = (\cos^2\alpha + \gamma\sin^2\alpha)/(\gamma\cos^2\alpha + \sin^2\alpha)$, where γ is the anisotropy of the subgrid-scale orography. The functions B, C are written as (Phillips 1984)

$$B = 1 - 0.18\gamma - 0.04\gamma^2 \qquad \text{and} \qquad C = 0.48\gamma + 0.3\gamma^2.$$

The final wind tendency produced by the low-level drag parametrization is then obtained from the following partially implicit discretization of Eq. (21)

$$\left(\frac{\partial \mathbf{U}}{\partial t}\right)_{\mathrm{blk}} = \frac{\mathbf{U}^{n+1}-\mathbf{U}^n}{\Delta t} = -A|\mathbf{U}^n|\mathbf{U}^{n+1} = -\frac{\beta}{1+\beta}\frac{\mathbf{U}^n}{\Delta t} \tag{22}$$

where $\beta = A|\mathbf{U}^n|\Delta t$ and $\mathbf{U}^{n+1} = \mathbf{U}^n/(1+\beta)$.

5.1.4 Non-orographic gravity wave drag

The tangent-linear and adjoint versions of the non-linear scheme for non-orographic gravity waves (in details described by Orr *et al.* 2010) were developed in order to reduce discrepancies between the full NL and linearized versions of the model, especially in the stratosphere. The parametrization scheme used in the NL model is based on Scinocca (2003). This is a spectral scheme that follows from the Warner and McIntyre (1996) scheme representing the three basic mechanisms that are conservative propagation, critical level filtering, and non-linear dissipation. Since the full nonhydrostatic and rotational wave dynamics considered by Warner and McIntyre (1996) is too costly for operational models, only hydrostatic and non-rotational wave dynamics are employed.

The dispersion relation for a hydrostatic gravity wave in the absence of rotation is

$$m^2 = \frac{k^2 N^2}{\tilde{\omega}^2} = \frac{N^2}{\tilde{c}^2} \tag{23}$$

where k, m are horizontal and vertical wavenumbers, while $\tilde{\omega} = \omega - kU$ and $\tilde{c} = c - U$ are the intrinsic frequency and phase speed (with c being the ground based phase speed and U the background wind speed in the direction of propagation), respectively.

The launch spectrum, which is globally uniform and constant, is given by the total wave energy per unit mass in each azimuth angle ϕ following Fritts and VanZandt (1993) as

$$\tilde{E}(m,\tilde{\omega},\phi) = B\Big(\frac{m}{m_*}\Big)^s \frac{N^2\tilde{\omega}^{-d}}{1-\big(\frac{m}{m_*}\big)^{s+3}} \tag{24}$$

where B, s and d are constants, and $m_* = 2\pi L$ is a transitional wavelength. Instead of the total wave energy $\tilde{E}(m,\tilde{\omega},\phi)$, the momentum flux spectral density $\rho\tilde{F}(m,\tilde{\omega},\phi)$ is required. This is obtained through the group velocity rule. In order to have the momentum flux conserved in the absence of dissipative processes as the spectrum propagates vertically through height-varying background wind and buoyancy frequency, the coordinate framework (k,ω) is used instead of $(m,\tilde{\omega})$ as shown in Scinocca (2003).

The dissipative mechanisms applied to the wave field in each azimuthal direction and on each model level are critical level filtering and non-linear dissipation. After application of them, the resulting momentum flux profiles are used to derive the net eastward $\rho\bar{F}_E$ and northward $\rho\bar{F}_N$ fluxes. The tendencies for the (u, v) wind components are then given by the divergence of those fluxes obtained through summation of the total momentum flux (i.e. integrated over all phase speed bins) in each azimuth ϕ_i projected onto the east and north directions, respectively:

$$\frac{\partial u}{\partial t} = g\frac{\partial(\rho\bar{F}_E)}{\partial p} \tag{25}$$

$$\frac{\partial v}{\partial t} = g\frac{\partial(\rho\bar{F}_N)}{\partial p} \tag{26}$$

where g is the gravitational acceleration and p is pressure.

Regularization:

In order to eliminate the spurious noise in TL computations caused by the introduction of this scheme, it was necessary to implement some regularizations. These include rewriting buoyancy frequency (N)

computations in the NL scheme in height coordinates instead of pressure coordinates (as used in the original code) and setting increments for momentum flux in the last three spectral elements (high phase speed) of the launch spectrum to zero.

5.1.5 Moist convection

The original version of the simplified mass-flux convection scheme currently used in the minimization of 4D-Var is described in Lopez and Moreau (2005). Through time, the original scheme has been updated so as to gradually converge towards the full convection scheme used in high-resolution 10-day forecasts (Bechtold *et al.* 2008). The transport of tracers by convection is also implemented.

The physical tendencies produced by convection on any conservative variable ψ (dry static energy, wind components, specific humidity, cloud liquid water) can be written in mass-flux form following Betts (1997)

$$\frac{\partial \psi}{\partial t} = \frac{1}{\rho}\left[(M_u + M_d)\frac{\partial \psi}{\partial z} + D_u(\psi_u - \psi) + D_d(\psi_d - \psi)\right] \tag{27}$$

The first term on the right hand side represents the compensating subsidence induced by cumulus convection on the environment through the mass flux, M. The other terms account for the detrainment of cloud properties in the environment with a detrainment rate, D. Subscripts u and d refer to the updraughts and downdraughts properties, respectively. Evaporation of cloud water and precipitation should also be added in Eq. (27) for dry static energy, $s = c_p T + gz$, and specific humidity.

Equations for updraught and downdraught:

The equations describing the evolution with height of the convective updraught and downdraught mass fluxes, M_u and M_d, respectively, are

$$\frac{\partial M_u}{\partial z} = (\varepsilon_u - \delta_u)M_u \tag{28}$$

$$\frac{\partial M_d}{\partial z} = -(\varepsilon_d - \delta_d)M_d \tag{29}$$

where ε and δ respectively denote the entrainment and detrainment rates. A second set of equations is used to describe the evolution with height of any other characteristic, ψ, of the updraught or downdraught, namely

$$\frac{\partial \psi_u}{\partial z} = -\varepsilon_u(\psi_u - \overline{\psi}) \tag{30}$$

$$\frac{\partial \psi_d}{\partial z} = \varepsilon_d(\psi_d - \overline{\psi}) \tag{31}$$

where $\overline{\psi}$ is the value of ψ in the large-scale environment.

In practice, Eqs. (28)-(29) are solved in terms of $\mu = M/M_u^{\text{base}}$, where M_u^{base} is the mass flux at cloud base (determined from the closure assumption as described further down).

Triggering of moist convection:

The determination of the occurrence of moist convection in the model is based on whether a positively buoyant test parcel starting at each model level (iteratively from the surface and upwards) can rise high enough to produce a convective cloud and possibly precipitation. For an updraught starting from the

lowest model level, its initial temperature and moisture departures with respect to the environment and its initial vertical velocity depend on surface sensible and latent heat fluxes, following Jakob and Siebesma (2003). When starting from higher model levels, the ascent is initially set to 1 m s^{-1} and its initial thermodynamic characteristics are assumed to be representative of a few hundred metre deep mixed-layer, with typical excesses of 0.2 K for temperature and 1×10^{-4} kg kg^{-1} for moisture. A 200 hPa threshold for cloud depth is prescribed to distinguish between shallow and deep convection. Mid-level convection is treated as deep convection.

Entrainment and detrainment:

Entrainment rate in the updraught (ε_u) is split into turbulent and organized components, which are both modulated by humidity conditions in the environment. Detrainment in the updraught (δ_u) is assumed to occur inside the convective cloud only where the updraught vertical gradient of kinetic energy and buoyancy are negative, that is usually in the upper part of the convective cloud.

Entrainment in downdraughts (ε_d) is assumed to occur only between the level of free sinking and the top of the 60 hPa atmospheric layer just above the surface. Inside this layer, it is set to a constant value. Detrainment (δ_d) is defined such as to ensure a downward linear decrease of downdraught mass flux to zero at the surface.

Precipitation processes:

The formation of precipitation from the cloud water contained in the updraught is parametrized according to Sundqvist *et al.* (1989) and a simple representation of precipitation evaporation is included. Precipitation formed from cloud liquid water at temperatures below the freezing point is assumed to freeze instantly, which affects the dry static energy tendency.

Closure assumptions:

One needs to formulate so-called closure assumptions to relate the convective updraught mass-flux at cloud base, M_u^{base}, to quantities that are explicitly resolved by the model. For deep convection, the closure is based on the balance between the convective available potential energy in the subgrid-scale updraught and the total heat release in the resolved larger-scale environment. The cloud base mass flux is expressed as the ratio of the latter two quantities, modulated with an adjustment timescale. This timescale depends on the updraught vertical velocity averaged over its depth and on spectral truncation. For shallow convection, the closure assumption links the moist energy excess at cloud base to the moist energy convergence inside the sub-cloud layer. The ratio of these two quantities yields the cloud base mass flux for shallow convection.

Regularization:

Various regularizations need to be applied in the TL and AD code of the convection scheme to avoid spurious instabilities. These include reducing or setting to zero the perturbations of some terms that directly depend on the updraught vertical velocity as well as reducing updraught buoyancy and cloud base mass flux perturbations.

5.1.6 Large-scale condensation and precipitation

The original version of the simplified diagnostic large-scale clouds and precipitation scheme currently used in the minimization of 4D-Var is described in Tompkins and Janisková (2004). Only a summary of

its main features is given here.

The physical tendencies of temperature and specific humidity produced by moist processes on the large-scale can be written as

$$\frac{\partial q}{\partial t} = -C_{\rm ce} + E_{\rm prec} + D_{\rm conv} \tag{32}$$

$$\frac{\partial T}{\partial t} = L(C_{\rm ce} - E_{\rm prec} - D_{\rm conv}) + L_f(F_{\rm rain} - M_{\rm snow}) \tag{33}$$

where $C_{\rm ce}$ denotes large-scale condensation/evaporation, $E_{\rm prec}$ is the moistening due to the evaporation of precipitation and $D_{\rm conv}$ is the detrainment of cloud water from convective clouds. $F_{\rm rain}$ and $M_{\rm snow}$ correspond to the freezing of rain and melting of snow, respectively. L and $L_{\rm f}$ are the latent heats of vaporisation/sublimation and fusion, respectively.

Condensation:

The subgrid-scale variability of humidity is assumed to be represented by a uniform distribution. This allows the estimation of the stratiform cloud fraction, $C_{\rm strat}$, and cloud condensate amount, $q_{\rm c}^{\rm strat}$, from the grid-box relative humidity, RH, as

$$C_{\rm strat} = 1 - \sqrt{\frac{1-RH}{1-RH_{\rm crit}-\kappa(RH-RH_{\rm crit})}} \tag{34}$$

$$q_{\rm c}^{\rm strat} = q_{\rm sat} C_{\rm strat}^2 \left\{ \kappa(1-RH) + (1-\kappa)(1-RH_{\rm crit}) \right\} \tag{35}$$

where $q_{\rm sat}$ is the saturation specific humidity. The critical relative humidity threshold, $RH_{\rm crit}$, and the coefficient κ are specified as in Tompkins and Janisková (2004). A simple diagnostic partitioning based on temperature is used to separate cloud condensate into liquid and ice.

The impact of convective activity on large-scale clouds, which is particularly important in the tropics and mid-latitude summers, is accounted for through the detrainment rate produced by the convection scheme. This detrainment term is used to compute the additional cloud cover and cloud condensate resulting from convection (i.e. convection called before condensation).

Precipitation processes:

The formation of precipitation from cloud condensate, $q_{\rm c}$, is parametrized according to Sundqvist *et al.* (1989), but the Bergeron-Findeisen mechanism and collection processes are currently disregarded. Precipitation formed from cloud liquid water at temperatures below the freezing point is assumed to freeze instantly, which corresponds to term $F_{\rm rain}$ in Eq. (33). On the other hand, precipitation evaporation is estimated from the overlap of precipitation with the uniformly distributed subgrid fluctuations of humidity inside the clear-sky fraction of the grid box.

Regularization:

Perturbations of $C_{\rm strat}$ were found to cause spurious instabilities in TL and AD integrations and are therefore artificially reduced according to the value of $C_{\rm strat}$ in the trajectory. A reduction of perturbations in the autoconversion of cloud condensate to precipitation is also needed.

5.2 A few remarks

The set of physical parametrization schemes developed for the ECMWF linearized model was described in section 5.1. Although there are some simplifications and regularizations applied in the different parametrization schemes, the whole package is comprehensive and its non-linear form is able to provide 2-3 days forecasts that show a degree of realism which does not depart too much from that of the non-linear physics. Different levels of simplification of the schemes have been driven either by the requirement to decrease computational cost for operational applications or the necessity to avoid unrealistic perturbations in the linearized version of the scheme. The applied regularizations and simplifications allow global integrations of the linearized model with elaborated physical parametrization schemes even up to 48 hours without producing spurious noise.

Overall, the presented package is a result of compromise between realism, linearity and computational cost while at the same time the level of complexity for the parametrization schemes is also influenced by the required applications. It is a constant challenge to maintain the best tradeoff between all those requirements.

5.3 Benefits of regularization

The validity of the tangent-linear approximation can be highly degraded due to the non-linear and discontinuous nature of physical processes (see section 3.2). If the derivatives of model outputs with respect to the model state variables become very large, the linearization will become useless.

As an illustration of how strong nonlinearities can lead to erroneous behaviour of the tangent linear model, Fig. 2b shows the evolution of zonal wind increments at the model level around 700 hPa when using the TL model without any regularization in the cloud parametrization scheme. When compared to the finite differences (differences between two non-linear integrations) in Fig. 2a, one can notice that strong spurious noise develops in the tangent linear model. This noise comes from the autoconversion function (Fig. 1) describing the conversion of cloud water to precipitation. When regularization is applied to this function, TL increments (Fig. 2c) agree well with the finite differences (Fig. 2a).

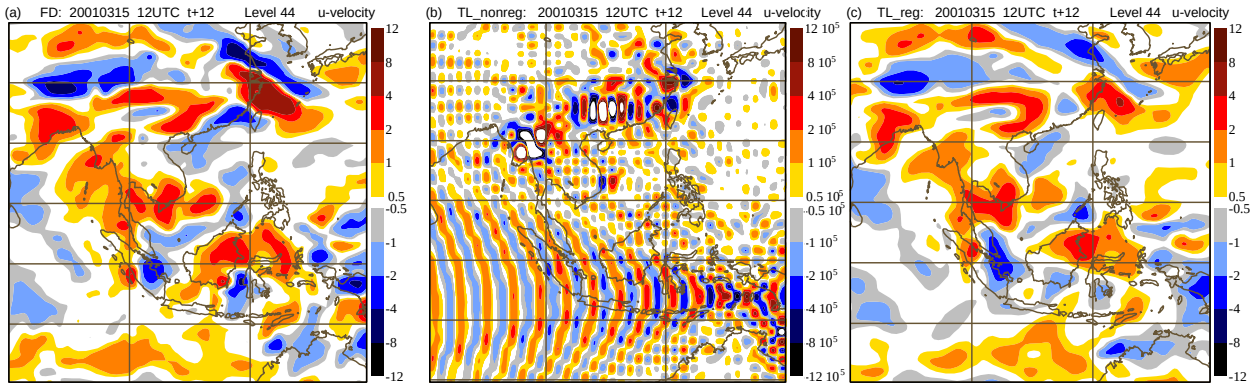


Figure 2: Zonal wind increments around 700 hPa after 12-hour evolution: (a) finite-differences (FD), (b) tangent-linear (TL) model without any regularization and (c) TL model with applied regularization in the cloud parametrization scheme.

6 Performance of the linearized physics

6.1 TL approximation

Once discontinuities in the different physical parametrization schemes have been properly smoothed, the TL model describing the evolution of perturbations with the simplified physical parametrizations generally fits the finite differences between two non-linear forecasts much better than an adiabatic TL model.

To demonstrate the impact of the different physical processes in the TL model, experiments have been performed at a horizontal resolution equal to T255 and 91 levels in vertical (L91) using the linearized physics of ECMWF described in section 5. The impact of the different physical processes on the tangent-linear evolution of temperature and zonal wind increments is shown in Fig. 3. Results are presented in terms of zonal mean of error difference as in Eq.(7) (i.e. the fit to the non-linear model with full physics) between the TL model including the whole set of linearized parametrization schemes and the adiabatic TL model (i.e. $\varepsilon_{exp} - \varepsilon_{adiab}$). Negative values are associated with an improvement of the model using the parametrization schemes with respect to the adiabatic TL model since they correspond to a reduction of the errors. The improvement is observed over most of the atmosphere, and is maximum in the lower troposphere.

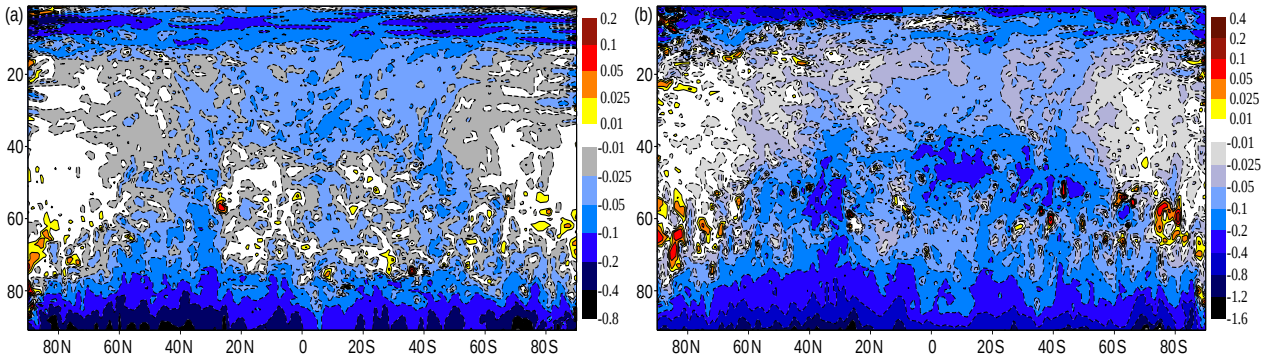


Figure 3: Impact of the ECMWF operational linearized physics on the evaluation of (a) temperature and (b) u-wind increments in zonal mean. Results are presented as the error differences (in terms of fit to the non-linear model with full physics) between the TL model with full linearized physics and the purely adiabatic TL model.

Figure 4 shows the global relative improvement (see Eq. (9)) coming from including (a) dry physical parametrization schemes (i.e. vertical diffusion, gravity wave drag, non-orographic gravity wave and radiation) alone and (b) in combination with the moist processes (i.e. convection and cloud with large-scale parametrization schemes) in the linearized model compared to adiabatic tangent linear model for temperature, wind and specific humidity. The additional improvement due to the inclusion of the moist parametrization schemes is not only coming from these schemes, but also from cloud-radiation interactions.

The relative error of the TL model with respect to the finite differences using the full non-linear physical parametrizations is also used for evaluation. Figure 5 shows vertical profiles of global relative error reduction (therefore negative values) of the TL model using different physical parametrization schemes with respect to the adiabatic TL model. The error reduction becomes larger by gradually including parametrization schemes, e.g. by including moist processes on top of the dry physical parametrization schemes.

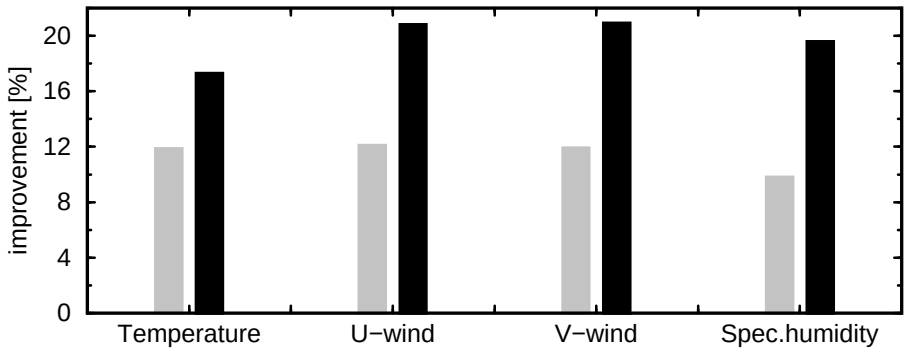


Figure 4: Global relative improvement [%] of the tangent-linear approximation for temperature, wind and specific humidity coming from including: (a) dry physical parametrization schemes (i.e. vertical diffusion, gravity wave drag, non-orographic gravity wave and radiation) alone (grey bars) and (b) in combination with moist processes (i.e. convection and large-scale cloud parametrization schemes - black bars) into the linearized model compared to the purely adiabatic tangent linear model.

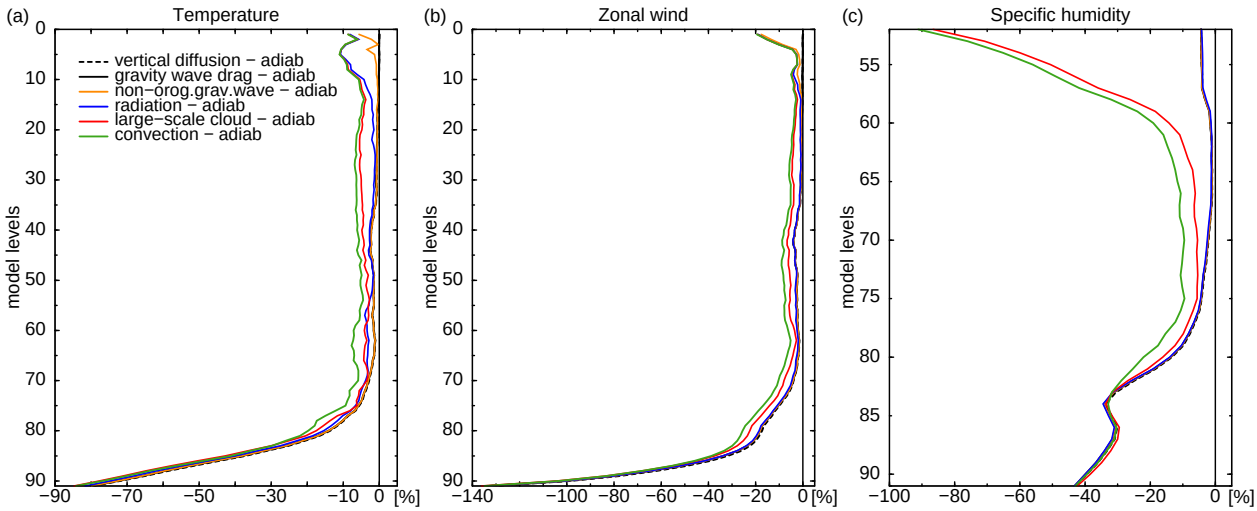


Figure 5: Global relative error reduction of the tangent-linear (TL) model obtained from the gradual inclusion of physical parametrizations as shown in legend with respect to adiabatic TL model for (a) temperature, (b) zonal wind and (c) specific humidity. Relative errors of TL model are computed with respect to finite differences using the full non-linear physics.

6.2 Adjoint sensitivities

Adjoint models can also be used for sensitivity studies since they allow to compute the gradient of one output parameter of a numerical model with respect to all input parameters. This property of adjoint allows to study, for instance, the sensitivity of a physical parametrization scheme to its input parameters (e.g. Li and Navon 1998, Janisková and Morcrette 2005). It is a more effective method compared with other standard approaches of repetitively using the direct schemes to obtain the sensitivity of all outputs by modifying in turn each input variables. More generally, an adjoint can be applied to analyze the sensitivity of a forecast aspect to initial conditions as proposed, for instance, by Errico and Vukicevic (1992) or Rabier *et al.* (1996). The adjoint method can also be used to measure the sensitivity with respect to any parameter of importance of the data assimilation system. In recent years, adjoint-based observation sensitivity techniques have been used as a diagnostic tool to monitor the observation impact on short-range forecasts (e.g. Langland and Baker 2004, Cardinali and Buizza 2004, Zhu and Gelaro 2008,

Cardinali 2009). Such technique is restricted by the tangent-linear assumption and its validity. The better the tangent-linear approximation, the more realistic and useful the sensitivity patterns. Results obtained through the adjoint integration when using a too simplified adjoint model with large inaccuracies or adjoint models without a proper treatment of nonlinearities and discontinuities, can be incorrect.

The adjoint($\mathbf{F}^T$) of the linear operator $\mathbf{F}$ can provide the gradient of any objective function, J, with respect to $\mathbf{x}$ (input variables) given the gradient of J with respect to $\mathbf{y}$ (output variables) as:

$$\frac{\partial J}{\partial \mathbf{x}} = \mathbf{F}^T . \frac{\partial J}{\partial \mathbf{y}} \tag{36}$$

As an example, Fig. 6 displays the adjoint sensitivity of the 24-hour forecast error to the initial conditions, i.e. to the analysis $\frac{\partial J}{\partial \mathbf{x}}$, where J is a measure of the forecast error (Rabier *et al.* 1996, Cardinali 2009). The sensitivity with respect to specific humidity and temperature at the lowest model level are shown for the situation on 28 August 2010 at 21:00 UTC from the run at T255L91 resolution. The results are presented for two different experiments, the first one run with only the dry parametrization schemes (i.e. vertical diffusion, gravity wave drag, non-orographic gravity wave and radiation) included in the adjoint model (Fig. 6a,b) and the second one with moist processes also taken into account (Fig. 6c,d). With only dry parametrization schemes, sensitivity to specific humidity is quite small and localized in areas of strongest dynamical activity. Even for temperature, it is obvious that some sensitivities are quite weak, especially in convective regions. Adding moist processes in the adjoint model brings additional structures to the sensitivity in areas affected by large-scale condensation/evaporation and convection. Therefore, using a more sophisticated adjoint model also provides more flow-dependent and more realistic sensitivities.

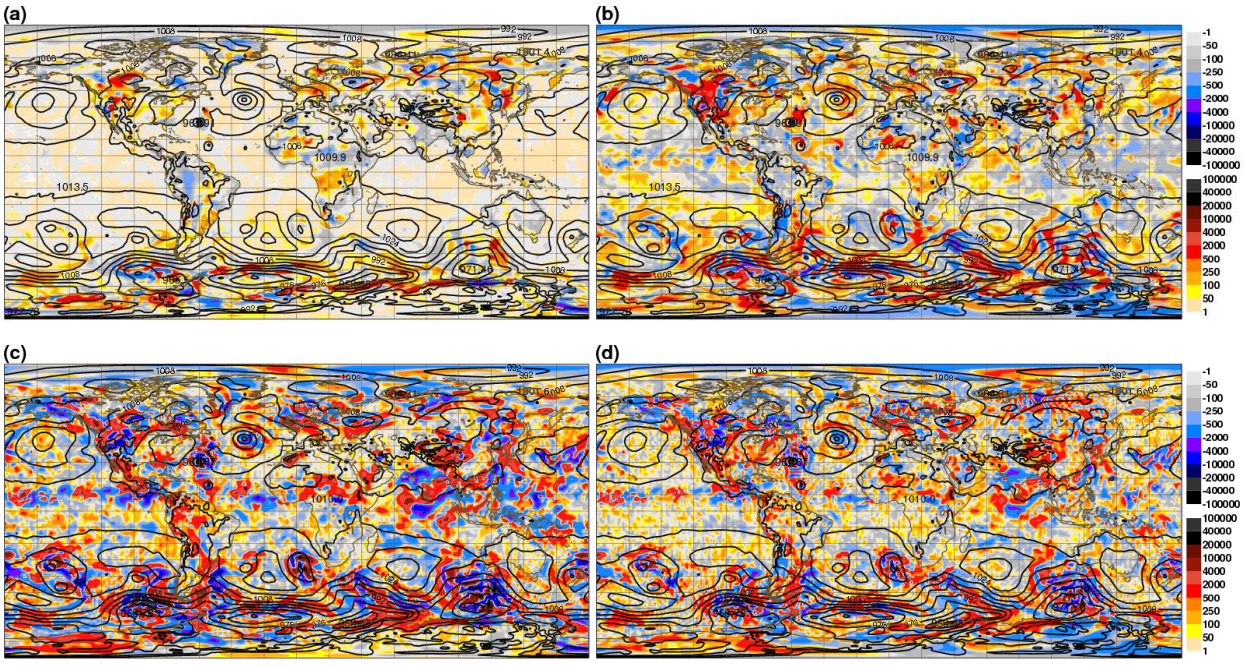


Figure 6: Adjoint sensitivity of the 24-hour forecast error to initial conditions in (a,c) specific humidity (J kg^{-1}/(g kg^{-1})) and (b,d) temperature (J kg^{-1}/K) at the lowest model level for the situation on 28 August2010 at 21:00 UTC. The results are presented for (a,b) an experiment with dry parametrization schemes (i.e. vertical diffusion, gravity wave drag, non-orographic gravity wave and radiation) used in the adjoint model and (c,d) with moist processes also included. Sensitivities are shown with colour shading. Black isolines represent mean-sea-level pressure (hPa).

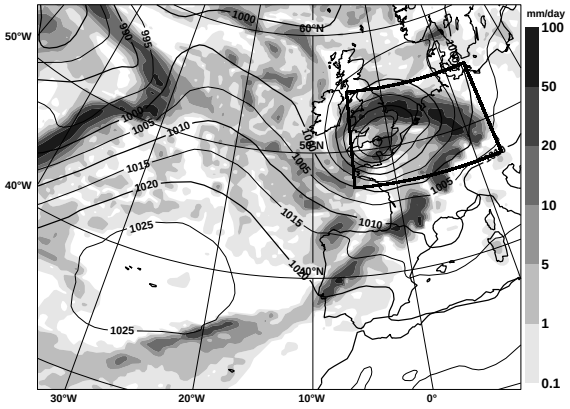


Figure 7: Map of 3-hour precipitation accumulations ending at 0000 UTC 10 February 2009 and used for computing the precipitation cost function inside the target black box over northwestern Europe. Grey shading shows precipitation (in mm day^{-1}), while black isolines of mean-sea-level pressure are also plotted (in hPa).

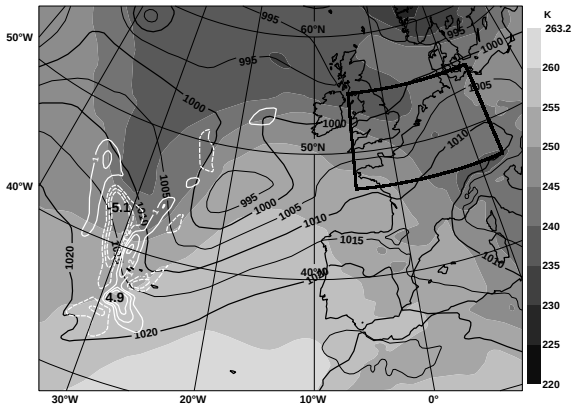


Figure 8: Adjoint sensitivities of the precipitation cost function defined over the black box (see Fig. 7) with respect to 500-hPa temperature at 0000 UTC 9 February 2009, i.e. with a lead time of 24 hours. Sensitivities are shown with white isolines (solid for positive, dash for negative) and are expressed in units of 10^{-4} (mm day^{-1}) K^{-1}. The background 500 hPa temperature field valid at 0000 UTC 9 February 2009 is displayed using grey shading and black isolines of mean-sea-level pressure are also plotted (in hPa).

Another example of adjoint sensitivity computations using the adjoint version of the linearized physics package is given here, where the cost function was defined as the 3-hour precipitation averaged over the core of a mid-latitude winter storm over northwestern Europe. One should emphasize that this kind of computation is only possible if the adjoint of moist physics parametrizations is available. Figure 7 shows the field of 3-hour precipitation accumulation used for the evaluation of the precipitation cost function inside the black box at 0000 UTC 10 February 2009. As an illustration, Fig. 8 displays the adjoint sensitivities of the precipitation cost function with respect to 500 hPa temperature at 0000 UTC 9 February 2009 (i.e. 24 hours beforehand and computed at T159L91 resolution). In other words, Fig. 8 points out the regions where temperature ought to be modified in order to change precipitation inside the target box, 24 hours later. In Fig. 8, the region of maximum sensitivity is found in the vicinity of the cold front associated with the 990 hPa low pressure system located at 19°W/47°N. The dipolar pattern of sensitivities indicates that a strengthening of the cross-frontal temperature gradient would result in a precipitation increase inside the black box, 24 hours later.

Of course, it would also be possible to plot sensitivities with respect to moisture, wind and surface pressure fields for this case (not shown). In fact, sensitivities can be computed with respect to any variable which is part of the control vector of the adjoint model. However, one should also keep in mind that the relevance and usefulness of adjoint sensitivities can be limited by the degradation of the linearity assumption over time.

6.3 Data assimilation

Experiments have been performed over July-September 2011 in order to compare two versions of the ECMWF 4D-Var system at resolution T511[3]L91: the first one including the linearized physics described

[3]T511 corresponding approximately to 40 km

above and the second one without it. Actually, in the version without the described linearized physics, a simple linear vertical diffusion (dry and acting mainly close to surface) and surface drag scheme (Buizza 1994) had to be used to avoid strong wind increments close to the surface. Precipitation and cloud related observations have not been taken assimilated in order to use the same type of observations in both experiments. Indeed, without the linearized moist physics in 4D-Var, cloud and precipitation observations cannot be assimilated since no observation equivalent can be produced from the model.

Including physical processes in the linear model of 4D-Var not only decreases the background cost function (measuring the distance between the initial state of the model and the background), but also brings model closer to observations as indicated by the general decreased observation cost function (measuring the distance between the model trajectory and corresponding observations) as seen in Fig. 9. Thus the distance between the model and the observations is better optimized when the linearized physics is used in the 4D-Var minimization.

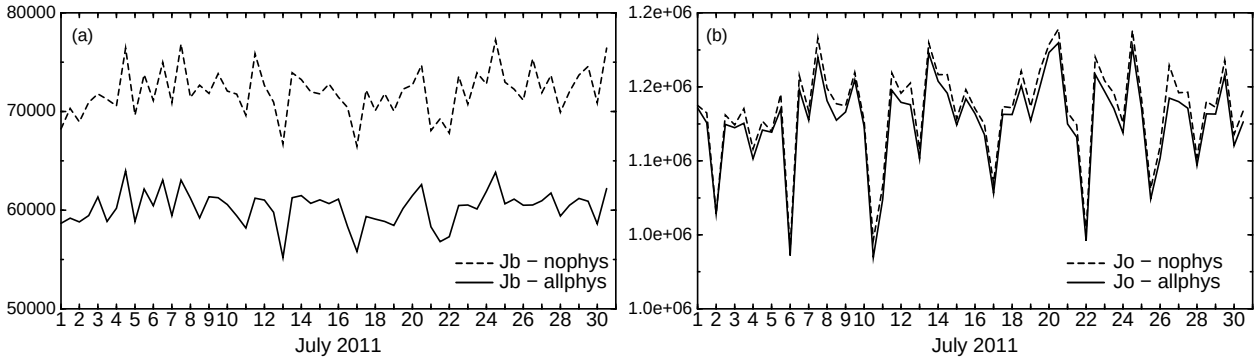


Figure 9: Global values of (a) background cost function and (b) observation cost function for 4D-Var assimilation experiments run with all linearized physical parametrization schemes included (solid line) and using only very simple vertical diffusion of Buizza (1994) (dashed line). Statistics are shown over July 2011.

The significance of the impact coming from including the linearized physical parametrization schemes in the 4D-Var system on the subsequent forecast is illustrated in Fig. 10 for the period of July - September 2011. The forecasts are scored against operational analyses in terms of anomaly correlation. A systematic and significant improvement for all plotted parameters, levels and regions is clearly obvious. Close to analysis time (where obviously the impact of the linearized physics in 4D-Var should be the largest), the biggest improvement is found in the middle and upper troposphere (e.g. 200 hPa wind vector scores) and overall in the Tropics. The positive impact is also generally remarkable in the lower troposphere (e.g. 700 hPa temperature scores or 700 hPa relative humidity scores).

The results presented above only show which impact the linearized physical parametrizations have on the evolution of the model state from the beginning of the 4D-Var assimilation window to the time of observations. However, including physical processes in the linearized model also allows to assimilate observations that are directly related to the physical processes, such as cloud and precipitation observations. Therefore further improvement in producing more realistic initial atmospheric states can be achieved. Since the late 1990s, significant efforts have been devoted to the assimilation of such observations. This is also the case at ECMWF, where a 1D+4D-Var technique has been first used operationally for the assimilation of precipitation-related observations using microwave brightness temperatures from SSM/I (Bauer *et al.* 2006) from June 2005 until March 2009. This was then replaced by direct 4D-Var assimilation unifying the treatment of clear-sky, cloudy and precipitation situations, leading to an all-sky approach (Bauer *et al.* 2010, Geer *et al.* 2010). Direct 4D-Var of rain- and cloud-affected observations

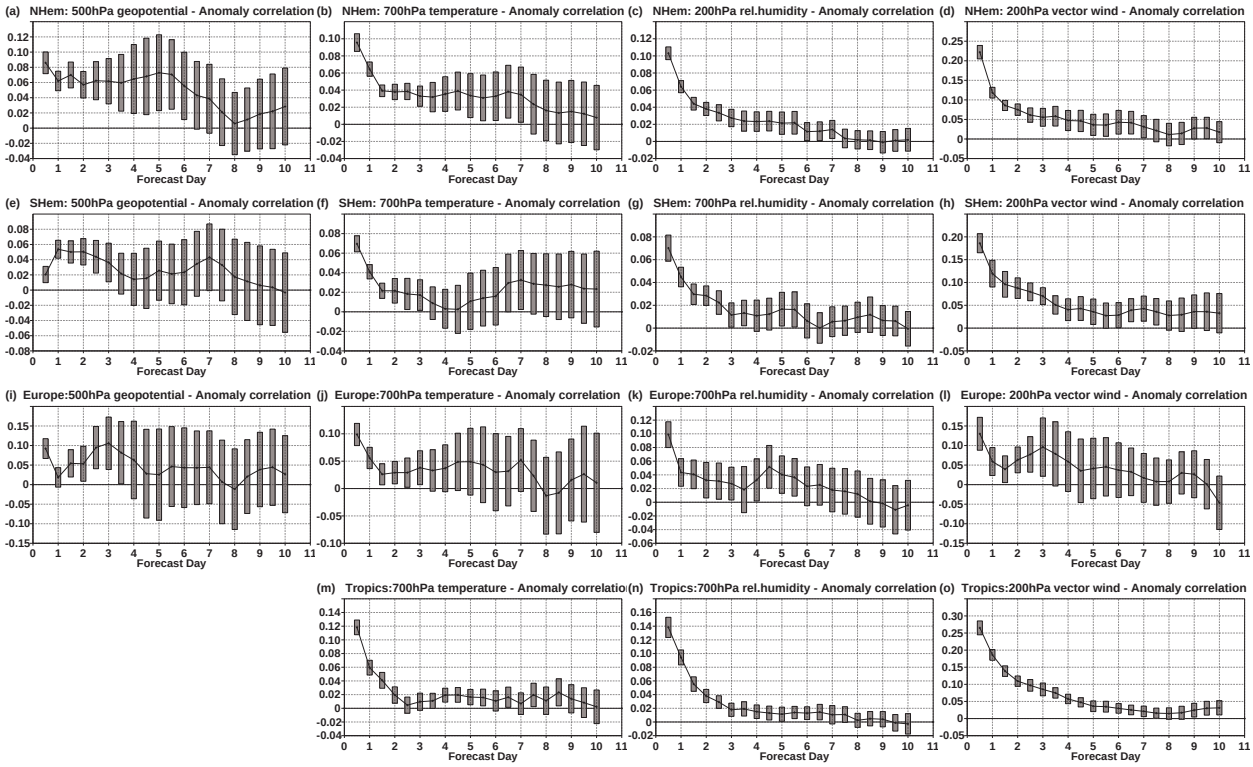


Figure 10: Relative impact from the inclusion of the linearized physical parametrization schemes in ECMWF's 4D-Var system. Forecast scores against operational analysis are shown in terms of anomaly correlations for ranges up to 10 days. Score change is normalized by the control and positive values correspond to an improvement. Grey bars indicate significance at the 95 % confidence level. Results are shown (from left to right) for: 500 hPa geopotential, 700 hPa temperature, 700 hPa relative humidity, 200 hPa vector wind and for the different regions: (a-d) Northern extratropics, (e-h) Southern extratropics, (i-l) Europe and (m-o) tropics. Statistics are valid for the period of July - September 2011.

allows a physically consistent adjustment of model dynamics with temperature and humidity increments, due to the sensitivity of radiance observations to the atmospheric state through the combined radiative transfer model and the moist-physics parametrization. Furthermore, direct 4D-Var of surface rain data from ground-based NCEP Stage IV rain radars and gauges over the Eastern USA recently became operational in ECMWF global forecasting system (Lopez 2011) providing the clear improvement of short-range precipitation forecasts over the region. In the longer term, one could consider the assimilation of more radar networks (e.g. Europe, China, Canada, ...) once problems of data availability and quality are solved.

Experimental studies for assimilation of other observations related to the physical processes which may be considered for the future operational assimilation and therefore requiring parametrization schemes being able to provide a realistic counterpart to these observations were also performed at ECMWF. Experiments were conducted to assimilate spaceborne cloud optical depths (from MODIS, Benedetti and Janisková 2008), precipitation radar reflectivities (from TRMM precipitation radar, Benedetti *et al.* 2005) and cloud radar data (from CloudSat, Janisková *et al.* 2011). More recently, the potential benefits of directly assimilating synoptic station (SYNOP) rain gauge observations in 4D-Var was investigated (Lopez 2012) in both, a high resolution operations-like context and a lower-resolution data-sparse reanalysis-like framework.

The results from all above mentioned studies are not shown here, since they are well documented in the literature.

7 Conclusions and prospects

Past experimentation and operational implementation in ECMWF's Integrated Forecasting System have clearly demonstrated the benefits of including linearized physical parameterization schemes in the data assimilation process. Linearized physics can also be beneficial to singular vector computations for the Ensemble Prediction System, leading to more realistic initial perturbations. It can be useful to diagnose short-range forecast sensitivities to observations. Furthermore, employing linearized moist physics parametrizations in the 4D-Var minimizations has permitted the assimilation of the ever-increasing number of satellite and ground-based observations that are sensitive to clouds and/or precipitation.

However, the development of efficient and well-behaved TL and AD codes is made difficult by many obstacles and is therefore time consuming and often tedious, if not sometimes rather frustrating. In particular, a substantial amount of work is required to simplify and regularize the code or, in other words, to eliminate or smooth out the discontinuities and non-linearities that often characterize physical processes. The behaviour of the linearized physics package also needs to be constantly and thoroughly monitored in a wide range of potential applications (e.g. data assimilation, singular vectors computations, sensitivity experiments). In particular, every time one of the physical parameterizations is modified in the non-linear forecast model (which in practice occurs at every new model release), it is necessary to verify that the tangent-linear approximation is not degraded. If it is, appropriate updates have to be made to the TL and AD code so as to avoid a likely degradation of the 4D-Var operational performance. Eventually, a delicate compromise must constantly be achieved between linearity, computational efficiency and realism, to ensure that the best analysis and (above all) forecast performance are obtained.

With the continual trend towards higher and higher resolutions (both in the horizontal and the vertical), maintaining a well-behaved linearized physics package is bound to become more and more challenging. Currently, the minimizations involved in ECMWF's 4D-Var are still run at a relatively coarse resolution of roughly 80 km, even though trajectories and final analyses are computed at 16 km resolution. When minimization resolution is increased, the ability to represent smaller-scale and often noisier processes (such as convection) is likely to make it more difficult to fulfil the TL hypothesis. However it should be mentioned that preliminary TL approximation tests were recently performed with a global resolution of 25 km and over 12 hours, with no sign of a degradation. One of the major uncertainty for the future is whether it will remain possible to make linearized physics to work when the resolution of the non-linear forecast model reaches a few kilometres, while the resolution remains well above 10 km in the 4D-Var minimizations. At this stage, the paradox of explicitly resolving convection in the trajectory but still needing to parametrize it in the minimization could be very challenging, and the current 4D-Var approach might need to be modified so as not to include the smaller scales in the entire analysis process (e.g. through trajectory smoothing).

There should nonetheless be some even greater concern about the growing complexity of the physical parametrizations used in the non-linear forecast model. Over the years, the increasing level of detail added to the representation of physical processes has been synonymous for enhanced and more numerous sources of non-linearity, which by construction cannot be included in the linearized physics package. There is a risk that if nothing is done to keep this trend under control, it will become impossible to make the linearized physics follow its non-linear counterpart closely enough, in which case 4D-Var as we know it may not be sustainable. Even though there is some hope that future configurations of data assimilation

based on weak-constraint 4D-Var might provide some ways to slightly relax the linearity constraint in time, it is paramount that non-linear model developers always remember that 4D-Var can only deliver good analyses if the linearity assumption remains valid over the entire assimilation window.

Acknowledgements

The authors would like to thank Peter Bauer, Anton Beljaars and Jean-Noël Thépaut for their useful reviews of this paper.

References

Araya-Polo M. and Hascoët L., 2004: Data Flow Algorithms in the Tapenade tool for Automatic Differentiation. In *Proceedings of 4th European Congresson Computational Methods, ECCOMAS'2004, Jyvaskyla, Finland*

Bauer P., Lopez P., Salmond D., Benedetti A., Saarinen S. and Bonazzola M., 2006: Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. II: 4D-Var. *Q. J. R. Meteorol. Soc.*, **132**, 2307–2332

Bauer P., Geer A.J., Lopez P. and Salmond D., 2010: Direct 4D-Var assimilation of all-sky radiances. I: Implementation. *Q. J. R. Meteorol. Soc.*, **136**, 1868–1885

Blackadar A. K., 1962: The vertical distribution of wind and turbulent exchange in a neutral atmosphere. *J. Geophys. Res.*, **67**, 3095–3102

Bechtold P., Köhler M., Jung T., Leutbecher M., Rodwell M., Vitart F. and Balsamo G., 2008: Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Q. J. R. Meteorol. Soc.*, **134**, 1337–1351

Benedetti A., Lopez P., Bauer P. and Moreau E., 2005: Experimental use of TRMM precipitation radar observations in 1D+4D-Var assimilation. *Q. J. R. Meteorol. Soc.*, **131**, 2473-2495

Benedetti A. and Janisková M., 2008: Assimilation of MODIS cloud optical depths in the ECMWF model. *Mon. Weather Rev.*, **136**, 1727–1746

Betts A., 1997: The parametrization of deep convection: A review. In *Proc. ECMWF Workshop on New Insights and Approaches to Convective Parametrization, Reading, 4–7 November 1996*, 166–188

Buizza R., 1994: Impact of simple vertical diffusion and of the optimisation time on optimal unstable structures. ECMWF Technical Memorandum **192**, ECMWF, Reading, UK, 25 pp

Cardinali C. and Buizza R., 2004: Observation sensitivity to the analysis and the forecast: A case study during ATreC targeting campaign. In *Proceedings of the first THORPEX international science symposium, 6-10 December 2004, Montreal, Canada. WMO/TD-1237, WWRP/THORPEX No. 6*

Cardinali C., 2009: Monitoring the observation impact on the short-range forecast. *Q. J. R. Meteorol. Soc.*, **135**, 239–250

Courtier P., Thépaut J.N. and Hollingsworth A., 1994: A strategy for operational implementation of 4D-Var using an incremental approach. *Q. J. R. Meteorol. Soc.*, **120**, 1367–1387

Errico R.M. and Reader K.D., 1999: An examination of the accuracy of the linearization of a mesoscale model with moist physics. *Q. J. R. Meteorol. Soc.*, **125**, 169-196

Errico R.M. and Vukicevic T., 1992: Sensitivity analysis using an adjoint of the PSU/NCAR mesoscale model. *Mon. Weather Rev.*, **120**, 1644–1660

Ebert E. E. and Curry J. A., 1992: A parametrization of ice optical properties for climate models. *J. Geophys. Res.*, **97D**, 3831–3836

Fouquart Y. and Bonnel B., 1980: Computations of solar heating of the earth's atmosphere: A new parametrization. *Beitr. Phys. Atmos.*, **53**, 35–62

Fouquart Y., 1987: Radiative transfer in climate models. In Physically Based Modelling and Simulation of Climate and Climate Changes, M. E. Schlesinger, Kluwer Acad. Publ., 223–284

Fritts D.C. and VanZandt T.E., 1993: Spectral estimates of gravity wave energy and momentum fluxes. Part I: Energy dissipation, acceleration, and constraints. *J. Atmos. Sci.*, **50**, 3685–3694

Geer A. J., Bauer P. and Lopez P., 2010: Direct 4D-Var assimilation of all-sky radiances. Part II: Assessment. *Q. J. R. Meteorol. Soc.*, **136**, 1886-1905

Geleyn, J.-F. and Hollingsworth, A., 1997: An economical and analytical method for the interactions between scattering and line absorption of radiation. *Contrib. to Atmos. Phys.*, **52**, 1–16.

Geleyn, J.-F., Banciu, D., Bellus, M., El Khatib, R., Moll, P., Saez, P. and Thépaut, J.-N., 2001: The operational 4D-Var data assimilation system of Météo-France: Specific characteristics and behaviour in the special case of the 99 Xmas

storms over France. In Proceedings of 18th Conference on Weather Analysis and Forecasting and 14th conference on Numerical Weather Prediction.

Giering R. and Kaminski T., 1998: Recipes for Adjoint Code Construction. *ACM Trans. Math. Software*, **24(4)**, 437–474

Jakob C. and Siebesma A. P., 2003: A new subcloud model for mass flux convection schemes. Influence on triggering, updraught properties and model climate. *Mon. Weather Rev.*, **131**, 2765–2778

Janisková M., Thépaut J.–N. and Geleyn J.–F., 1999: Simplified and regular physical parameterizations for incremental four-dimensional variational assimilation. *Mon. Weather Rev.*, **127**, 26–45

Janisková M., Mahfouf J.-F., Morcrette J.-J. and Chevallier F., 2002: Linearized radiation and cloud schemes in the ECMWF model: Development and evaluation. *Q. J. R. Meteorol. Soc.*, **128**, 1505–1527

Janisková M., 2003: Physical processes in adjoint models: potential pitfalls and benefits. In Proceedings of ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean, 8-12 September 2003, Reading, U.K., 179-191

Janisková M. and Morcrette J.-J., 2005: Investigation of the sensitivity of the ECMWF radiation scheme to input parameters using adjoint technique. *Q. J. R. Meteorol. Soc.*, **131**, 1975-1995

Janisková M., Lopez P. and Bauer P., 2011: Experimental 1D+4D-Var assimilation of CloudSat observations. *Q. J. R. Meteorol. Soc.*, DOI:10.1002/qj.988

Langland R. H. and Baker N. L., 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, **56A**, 189-201

Laroche, S., Tanguay, M. and Delage, Y., 2002: Linearization of a simplified planetary boundary layer parametrization, *Mon. Weather Rev.*, **130**, 2074-2087

Li Z. and Navon I. M., 1998: Adjoin sensitivity of the Earth's radiation budget in the NCEP medium-range forecasting model. *J. Geophys. Res.*, **103 (D4)**, 3801-3814

Lopez P., 2002: Implementation and validation of a new prognostic large-scale cloud and precipitation scheme for climate and data-assimilation purposes. *Q. J. R. Meteorol. Soc.*, **128**, 229-257

Lopez P. and Moreau E., 2005: A convection scheme for data assimilation: Description and initial tests. *Q. J. R. Meteorol. Soc.*, **131**, 409–436

Lopez P. and Bauer P., 2007: "1D+4D-Var" assimilation of NCEP Stage IV Radar and gauge hourly precipitation data at ECMWF. *Mon. Weather Rev.*, **135**, 2506–2524

Lopez P., 2011: Direct 4D-Var Assimilation of NCEP Stage IV Radar and Gauge Precipitation Data at ECMWF. *Mon. Weather Rev.*, **139**. 2098-2116

Lopez P., 2012: Experimental 4D-Var assimilation of SYNOP rain gauge data at ECMWF. ECMWF Technical Memorandum **661**, 25 pp.

Lott F. and Miller M. J., 1997: A new subgrid-scale orographic drag parametrization: Its formulation and testing. *Q. J. R. Meteorol. Soc.*, **123**, 101–127

Louis J.-F., Tiedtke M. and Geleyn J.-F., 1982: A short history of the PBL parametrization at ECMWF. In *Proc. ECMWF Workshop on Planetary Boundary Layer Parameterization, Reading, 25–27 November, 1981*, 59–80

Mahfouf J.-F., 1999: Influence of physical processes on the tangent-linear approximation. *Tellus*, **51**, 147–166

Mahfouf J.-F. and Rabier F., 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: Part II: Experimental results with improved physics. *Q. J. R. Meteorol. Soc.*, **126**, 1171–1190

Mahfouf J.-F., 2005: Linearization of a simple moist convection for large-scale NWP models. *Mon. Weather Rev.*, **133**, 1655–1670

Mlawer E. and Clough S. A., 1997: Shortwave and longwave enhancements in the rapid radiative transfer model. In *Proceedings of 7th Atmospheric Radiation Measurement (ARM) Science Team Meeting, U.S. Department of Energy, CONF-9603149*, available from http://www.arm.gov/publications/proceedings/conf07/title.stm/mlaw-97.pdf

Mlawer E., Taubman S. J., Brown P. D., Ianoco M. and Clough S. A., 1997: Radiative transfer for inhomogeneous atmospheres: RRTM a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16663–16682

Morcrette J.-J., Smith L. and Fouquart Y., 1986: Pressure and temperature dependence of absorption in longwave radiation parametrizations. *Beitr. Phys. Atmos.*, **59**, 455–469

Morcrette J.-J., 1989: Description of the radiation scheme in the ECMWF operational weather forecast model. ECMWF Tech. Memo. No. 165

Morcrette J.-J., 1991: Radiation and cloud radiative properties in the ECMWF operational forecast model. *J. Geophys. Res.*, **96D**, 9121–9132

Morcrette J.-J. and Jakob C., 2000: The response of the ECMWF model changes in the cloud overlap assumption. *Mon. Weather Rev.*, **128**, 876-887.

Morcrette J.-J., Mlawer E., Iacono M. and Clough S., 2001: Impact of the radiation transfer scheme RRTM in the ECMWF forecating system. *ECMWF Newsletter No. 91*, 2–9

Orr A., Bechtold P., Scinoccia J., Ern M. and Janisková M., 2010: Improved middle atmosphere climate and analysis in the

ECMWF forecasting system through a non-orographic gravity wave parametrization. *J. Climate*, **23**, 5905–5926

Ou S.C. and Liou K.-N., 1995: Ice microphysics and climate temperature feedback. *Atmos. Research*, **35**, 127–138

Phillips S. P., 1984: Analytical surface pressure and drag for linear hydrostatic flow over three-dimensional elliptical mountains. *J. Atmos. Sci.*, **41**, 1073–1084

Rabier, F., Klinker, E., Courtier P. and Hollingsworth A., 1996: Sensitivity of forecast errors to initial conditions. *Q. J. R. Meteorol. Soc.*, **122**, 121–150

Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.-F. and Simmons, A., 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.*, **126**, 1143-1170

Scinocca J. F., 2003: An accurate spectral nonorographic gravity wave drag parameterization for general circulation models. *J. Atmos. Sci.*, **60**, 667–682

Smith E. A. and Shi L., 1992: Surface forcing of the infrared cooling profile over the Tibetan plateau. Part I: Influence of relative longwave radiative heating at high altitude. *J. Atmos. Sci.*, **49**, 805–822

Sundqvist, H., Berge, E. and Kristjansson, J.E., 1989: Condensation and cloud parametrization studies with mesoscale numerical weather prediction model. *Mon. Weather Rev.*, **117**, 1641-1657

Tompkins A.M. and Janisková M., 2004: A cloud scheme for data assimilation: Description and initial tests. *Q. J. R. Meteorol. Soc.*, **130**, 2495–2517

Tsuyuki, T., 1996: Variational data assimilation in the tropics using precipitation data. Part II: 3-D model. *Mon. Weather Rev.*, **124**, 2545-2561

Warner C. D. and McIntyre M. E., 1996: An ultra-simple spectral parametrization for non-orographic gravity waves. *J. Atmos. Sci.*, **58**, 1837–1857

Washington W. M. and Williamson D. L., 1977: A description of the NCAR GCMs. GCMs of the atmosphere. J. Chang, Ed. *Methods in Computational Physics*, **17**, 111–172

Zhu Y. and Gelaro R., 2008: Observation sensitivity calculations using the adjoint of the Gridpoint Statistical Interpolation (GSI) analysis system. *Mon. Weather Rev.*, **136**, 335-351

Zou, X., Navon, I.M. and Sela, J.G., 1993: Variational data assimilation with moist threshold processes using the NMC spectral model. *Tellus*, **45A**, 370-387

Zupanski, D. and Mesinger, F., 1995: Four-dimensional variational assimilation of precipitation data. *Mon. Weather Rev.*, **123**, 1112-1127.