TECHNICAL MEMORANDUM

411

A convection scheme for data assimilation: Description and initial tests

Philippe Lopez and Emmanuel Moreau

Research Department

To be submitted to Q. J. Roy. Meteor. Soc.

March 2004

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 new simplified parametrization of subgrid scale convective processes has been developed and tested in the framework of the ECMWF Integrated Forecasting System for the purpose of variational data assimilation, singular vector calculations and adjoint sensitivity experiments. Its formulation is based on Tiedtke's (1989) nonlinear convection scheme used at ECMWF, but a set of simplifications has been applied to substantially improve its linear behaviour. These include the specification of a single closure assumption based on convective available potential energy, the uncoupling of the equations for the convective mass flux and updraught characteristics and a unified formulation of the entrainment and detrainment rates. Simplified representations of downdraughts and momentum transport are also included in the new scheme.

Despite these simplifications, the forecasting ability of the new convective parametrization is shown to remain satisfactory even in seasonal integrations. A detailed study of its Jacobians and the validity of the linear hypothesis is presented. The new scheme is also tested in combination with the new simplified parametrization of large-scale clouds and precipitation developed by Tompkins and Janisková (2004). In contrast with Mahfouf's (1999) simplified convective parametrization, its tangent-linear and adjoint versions account for perturbations of all convective quantities including convective mass flux, updraught characteristics and precipitation fluxes. Therefore the new scheme is expected to be beneficial when combined with radiative calculations that are directly affected by condensation and precipitation.

Examples of applications of the new moist physics in 1D-Var retrievals using microwave brightness temperature measurements as well as in adjoint sensitivity experiments are presented.

1 Introduction

In Meteorology, variational data assimilation is used to obtain a statistically optimal representation of the atmospheric state for a given date (the *analysis*) through the combination of background information from a previous short-range forecast with useful information contained in a set of available observations. Practically, variational methods are based on the minimization of a cost function that measures the global distance of the new atmospheric state to both the background state and the observations. The major underlying assumption in these techniques is that all processes involved can be considered as linear or at least only weakly nonlinear. In four-dimensional variational data assimilation (4D-Var), the minimization of the cost function is performed over a time window of several hours, which permits the assimilation of observations at their actual time. For instance, at the European Centre for Medium-range Weather Forecasts (ECMWF), the twice-daily 4D-Var analyses (Courtier *et al.* 1994; Rabier *et al.* 2000) are currently obtained by assimilating observations over a 12-hour window.

The propagation of the 4D-Var increments through the assimilation window via adjoint integration and the conversion of the model state into equivalent observed quantities require the use of a set of parametrizations that describe the atmospheric dynamic and physical processes. In particular, a detailed enough representation of moist processes (convection and large-scale condensation) becomes essential when observations directly related to clouds and precipitation are to be assimilated. However, moist diabatic processes are by nature nonlinear and non-differentiable or discontinuous in time (e.g. saturation threshold, precipitation formation). A careless inclusion of raw parametrizations of these phenomena can quickly result in serious violation of the basic linear assumption made in variational data assimilation. This is particularly true in the case of convection as shown in Županski (1993) and Vukićević and Errico (1993) for instance. In recent years, many efforts have been dedicated to the design of simplified parametrizations of atmospheric moist processes for the specific purpose of operational data assimilation (e.g. Janisková *et al.* 1999; Mahfouf *et al.* 1996; Mahfouf 1999 [MAH99]; Mahfouf and Rabier 2000, Fillion and Mahfouf 2000). Zou (1993) successfully performed 4D-Var assimilation of pseudo-observations including the adjoint of simple large-scale condensation and convection parametrizations in the NCEP model, with no special treatment of "on-off" processes. Nonetheless, Županski and Mesinger

(1995) underlined that the importance of discontinuities in moist processes is likely to depend on the model used and on the chosen meteorological situation. These authors indicated that their convection scheme had first to be modified in order to eliminate its discontinuous behaviour. Only then was it possible to obtain a better precipitation forecast using the 4D-Var assimilation of 24-hour accumulated surface rainfall observations prior to the beginning of the forecast. In other respects, Marécal and Mahfouf (2003) demonstrated that an indirect '1D-Var + 4D-Var' assimilation of precipitation observations was more robust than a direct 4D-Var assimilation, partly because of the presence of strong nonlinearities in Tiedtke's (1989) mass-flux convection scheme.

This paper introduces the formulation, the implementation and the application of a new simplified parametrization of convection (called SIMPCV2 hereafter) suitable wherever linearized computations are needed. SIMPCV2 is much more detailed than the MAH99 simplified convection scheme (SIMPCV hereafter). In fact, the new parametrization of convection has been designed such as to retain some fundamental similarities with Tiedtke's (1989) nonlinear scheme (operationally used in forecast mode; OPNLCV hereafter) and at the same time to reduce the high degree of complexity of the latter that is a source of strong nonlinearities, detrimental to 4D-Var assimilation. Table 2 summarizes all useful acronyms defined and used throughout the paper.

Section 2 describes SIMPCV2, while section 3 presents its validation in season-long integrations. In section 4 SIMPCV2 is first compared to Tiedtke's parametrization in terms of Jacobians and in terms of linearity over a single timestep for initial perturbations of various amplitudes. Then, the linear behaviour of SIMPCV2 is assessed in 12h-long integrations for initial perturbations that are typical of analysis errors. This set-up corresponds to the one used in the ECMWF 4D-Var assimilation system. Finally, some preliminary applications of the new simplified convection scheme are presented in section 5.

2 Description of the new parametrization

Previous studies underlined the need for an improvement of the linearized version of ECMWF's parametrizations of moist processes to be used in a wide range of applications including 4D-Var assimilation, singular vector computations and 1D-Var retrievals from rain observations (Marécal and Mahfouf 2000). Tompkins and Janisková (2004) recently developed a new simplified parametrization of large-scale clouds and precipitation. A new simplified version of the operational convection scheme was also required. This new simplified mass-flux scheme (SIMPCV2) was designed such as to retain some basic similarities with OPNLCV and at the same time to reduce its degree of complexity that can be a source of unwanted strong nonlinearities. Note that modifications described by Bechtold *et al* (2004) have been recently applied to Tiedtke's original nonlinear convection scheme to improve forecast scores.

In the new simplified scheme, all types of convection (shallow, mid-level, and deep) are treated in the same way. In particular, the link between the model control variables and the cloud base mass flux (the so-called *closure assumption*) is expressed through a single formulation that depends on the release of convective available potential energy (CAPE) in time. In contrast with OPNLCV, the equations that describe the vertical evolution of the updraught mass flux, M_u (in kg m^{-2} s^{-1}), and of the updraught thermodynamic characteristics, Φ_u (dry static energy, s, and total water $q_t = q + q_c$ where q denotes specific humidity and q_c cloud condensate), are uncoupled:

$$\frac{\partial M_u}{\partial z} = (\varepsilon_u - \delta_u) M_u \tag{1}$$

$$\frac{\partial \Phi_u}{\partial z} = -\varepsilon_u (\Phi_u - \overline{\Phi}) \tag{2}$$

where ε_u and δ_u are the fractional entrainment and detrainment rates, respectively. Here, as well as in the rest of the paper, overbars denote field values in the environment, that is averaged over a model gridbox, and subscript or superscript "u" relate to the updraught characteristics. As explained below, the uncoupling allows the removal of the iterative calculations involved in OPNLCV for updating the cloud base mass flux thereby leading to an easier development of the adjoint. Based on Siebesma and Jakob (personal communication), ε_u is parametrized as $c_\varepsilon/(z-z_{st})+10^{-5}$ m^{-1} where c_ε=0.4 and z_{st} is the starting height of the updraught (in m). As for detrainment, observations (e.g. Cohen 2000) suggest that δ_u is usually smaller than ε_u in the lower part of the updraught and comparable to ε_u further up, which is taken into account by writing

$$\delta_u = \varepsilon_u\left(1-c_\delta \exp\left(-\frac{z-z_{st}}{2000}\right)\right) \tag{3}$$

where the constant c_δ is set to 0.7. Close to cloud top where updraught buoyancy becomes negative, a constant organized detrainment rate of 2×10^{-4} m^{-1} is added to δ_u computed from Eq. (3).

Convection is assumed to be activated only if the bulk convective updraught vertical velocity remains positive at cloud base. The updraught is assumed to originate from the surface only if its initial vertical velocity, w_s, is positive. w_s is calculated from the surface heat fluxes using Holtslag and Moeng (1991)

$$w_s = 1.2\left(u_*^3 - 1.5\,\frac{gz\,\kappa\,F_V}{T}\right)^{\frac{1}{3}} \tag{4}$$

where κ=0.4 (von Kármán's constant) and $F_V=-\overline{w'T_v'}$ is the surface flux of virtual temperature, T_v (primes denote perturbations). For simplicity, the friction velocity u_* is arbitrarily set equal to 0.1 m s^{-1} as in OPNLCV. The initial temperature and specific humidity departures of the updraught from the environment, δT_u and δq_u, are assumed to be dependent on the surface sensible and latent heat fluxes, F_S and F_L (in W m^{-2}; positive downward) respectively, according to

$$\delta T_u = -1.5\frac{F_S}{\rho C_p w_s} \tag{5}$$

$$\delta q_u = -1.5\frac{F_L}{\rho L_v w_s} \tag{6}$$

where ρ is the density and C_p the specific heat capacity of the air at constant pressure, and L_v is the latent heat of vaporisation of water. If convection cannot be initiated from the surface, the convective ascent may originate from higher levels provided relative humidity exceeds 80%. In this case, the initial vertical velocity of the bulk updraught is set equal to 1 m s^{-1}.

Regardless of whether the updraught originates from the surface or from higher up, the vertical evolution of its kinetic energy is computed following Simpson and Wiggert (1969), which involves the buoyancy, $\mathscr{B}(z)$, as well as the entrainment of environmental air into the updraught,

$$\frac{\partial w_u^2}{\partial z} = -\alpha\varepsilon_u w_u^2 + \beta\mathscr{B}(z) \tag{7}$$

where $\alpha=\frac{2}{3}$, β=1 and w_u is the updraught vertical velocity. The buoyancy in the updraught is defined as

$$\mathscr{B}(z) = \frac{g}{\overline{T_v}}(T_v^u-\overline{T_v}) - gq_c^u \tag{8}$$

which includes the cloud water load inside the updraught, q_c^u.

Using a similar approach, a simple parametrization of convective downdraughts that are usually associated with the strong cooling due to the evaporation of precipitation has been implemented. Uncoupled equations

describe the evolution with height of the downdraught mass flux, M_d, and of the downdraught thermodynamical characteristics, Φ_d:

$$\frac{\partial M_d}{\partial z} = -(\varepsilon_d - \delta_d) M_d \tag{9}$$

$$\frac{\partial \Phi_d}{\partial z} = \varepsilon_d (\Phi_d - \overline{\Phi}) \tag{10}$$

where the fractional entrainment and detrainment rates in the downdraughts are respectively $\varepsilon_d = 2 \times 10^{-4}$ m^{-1} and $\delta_d = c_\delta / z$, with c_δ=0.3.

In practice, Eq. (1) and Eq. (9) are solved in terms of the ratios $\mu_u = M_u / M_u^{clb}$ and $\mu_d = M_d / M_u^{clb}$ respectively, where M_u^{clb} denotes the updraught mass flux at cloud base (clb). μ_u and μ_d are respectively initialized to 1 at cloud base and to -0.3 at the level of free sinking following Tiedtke (1989). Once the vertical profiles of $\mu_u(z)$ and $\mu_d(z)$ are known, M_u^{clb} is calculated in such a way that the CAPE in the atmospheric column is removed over a characteristic resolution dependent timescale, τ (3600 s for resolutions coarser than T319, 1200 s from T319 to T511, 600 s beyond T511), according to Gregory *et al.* (2000)

$$M_u^{clb} = \frac{CAPE}{\tau} \frac{1}{\int_{cloud} \left(\frac{1+\alpha q}{C_p \overline{T_v}} \frac{\partial \overline{s}}{\partial z} + \alpha \frac{\partial \overline{q}}{\partial z} \right) \left(\mu_u(z) + \mu_d(z) \right) g \frac{dz}{\overline{p}}} \tag{11}$$

where $\alpha = \frac{R_v}{R_d} - 1$ and $CAPE = \int_{cloud} \mathscr{B}(z) dz$. Since earlier experiments suggested that the values of the cloud-base mass flux were too low in the case of shallow convection, M_u^{clb} is multiplied by 3 in such situations.

SIMPCV2 also features a simple representation of convective momentum transport. It is based on OPNLCV but in contrast with the latter the same entrainment and detrainment rates as for heat and moisture are applied to the two wind components for the sake of simplicity.

The computation of precipitation formation, $\mathscr{P}$ (in kg kg^{-1} m^{-1}), from the updraught cloud condensate is inspired by Tiedtke (1989)

$$\mathscr{P} = \frac{K(z)\, C_f\, q_c^u}{w_u} \tag{12}$$

where

$$K(z) = \begin{cases} 1.2 \times 10^{-3} \text{ s}^{-1} & \text{if z - z}_{clb} > 1500 \text{ m} \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

$\mathscr{P}$ represents autoconversion and collection/aggregation processes, while the coefficient C_f crudely accounts for the dependence of precipitation formation efficiency on temperature. Practically, C_f is set equal to $= 1 + \frac{4}{15}(260 - T_u)$, with the restriction $1 \leq C_f \leq 5$. It should be noted that in OPNLCV Tiedtke's original formulation of precipitation formation has been replaced by Sundqvist's (1988).

The tangent-linear and adjoint versions of SIMPCV2 are available and have been tested in different types of experiments, some of which will be presented in section 5. Simplifications and modifications that were needed to improve the scheme's linear behaviour will be summarized in section 4.4. It is worth noting that unlike SIMPCV, SIMPCV2 handles perturbations of all convective quantities including mass flux, all updraught characteristics as well as precipitation fluxes. This is expected to be beneficial when combining SIMPCV2 with a radiative transfer model for frequencies that are sensitive to cloud condensate and/or hydrometeors.

3 Season-long integrations

Although SIMPCV2 was not designed for the purpose of medium or long-range weather forecasting, it is still important to check that its results in a forecast mode do not depart too much from those obtained with OPNLCV. As an example, season-long integrations with a T95 spectral resolution (approx. 200 km) in the horizontal and 60 vertical levels are presented here for the winter 1987-1988 using either OPNLCV or SIMPCV2 in cycle 26r3 of the ECMWF forecast model. The respective experiments will be referred to as OPNLCV and SIMPCV2.

For each experiment, an ensemble of three 4-month long runs were started from 1, 2 and 3 November 1987 respectively to ensure more robust statistics. The actual validation focuses on the last three months of these integrations, namely December, January and February. Three-month averages are shown for four different fields related to clouds and precipitation: the total surface precipitation (Fig. 1), the total cloud cover (Fig. 2), the net shortwave radiation at the top of the atmosphere (SWR-TOA; Fig. 3), and the outgoing longwave radiation at the top of the atmosphere (OLR-TOA; Fig. 4). Corresponding observations from the Global Precipitation Climatology Project (GPCP), from the International Satellite Cloud Climatology Project (ISCCP), and from the Earth Radiative Budget Experiment (ERBE), are used to validate the model fields.

Each figure displays (a) the observed field, (b) the field from OPNLCV, (c) the field from SIMPCV2, (d) the difference between OPNLCV and observations, and (e) the difference between SIMPCV2 and observations. The global mean of each field appears in the title of each panel along with the root mean square (RMS) errors at the top of panels (d) and (e).

In terms of precipitation, Fig. 1.(d) and (e) indicate that OPNLCV and SIMPCV2 exhibit rather similar discrepancies with respect to the GPCP observations, except over southern Africa. In both experiments, the precipitation simulated in the Inter-Tropical Convergence Zone (ITCZ) tends to be substantially overestimated over the central and eastern Pacific, over the Atlantic and over the Indian Ocean. Local departures reach 15 mm day^{-1} in OPNLCV and 8 mm day^{-1} in SIMPCV2 slightly east of the dateline. On the other hand, precipitation is clearly underestimated over the Indonesian warm pool by up to 12 mm day^{-1} with both convection schemes. Also note the similar double structure of the ITCZ in the Indian Ocean in Fig. 1.(b) and (c), which does not appear in the observations. Overall, the average ratio of convective precipitation (not shown) on the total amount reaches 56% in OPNLCV versus 49% in SIMPCV2. The mean departure from the GPCP observations is lower in SIMPCV2 (0.12 mm day^{-1}) than in OPNLCV (0.22 mm day^{-1}), but the RMS error is slightly higher in SIMPCV2 (1.91 versus 1.86 mm day^{-1}).

Figure 2.(d) and (e) indicate that the departures from observed cloud cover exhibit similar patterns in OPNLCV and SIMPCV2: too low cloudiness over the oceanic trade-wind regions often affected by stratocumuli (west of South America, South Africa, California and Morroco), excessive cloud cover over Central America, the Andes, southwest of India and over Southeast Asia. Over the extra-tropical oceans, clouds are more widely underestimated in SIMPCV2 than in OPNLCV. In fact this corresponds to a lack of low-level clouds (not shown) that partly results from the use of the CAPE closure for shallow convection in SIMPCV2 instead of the moisture convergence closure applied in OPNLCV. Globally, cloud amount in SIMPCV2 shows a deficit of roughly 5.8% while OPNLCV overestimates it by only 1.1%. The RMS error is larger in SIMPCV2 than in OPNLCV (20.80 versus 17.73%).

Figure 3.(d) and (e) and Fig. 4.(d) and (e) exhibit SWR-TOA and OLR-TOA deficiencies that are consistent with the departures previously outlined for cloud cover. The excess of SWR-TOA in the trade-wind regions affected by stratocumuli (up to 80 W m^{-2} off the Peruvian coast) results from the lack of such clouds in the simulations. Both figures suggest that cloud depth from OPNLCV or SIMPCV2 is too low over the Indonesian warm pool, the Amazon and West Africa. On the other hand, the simulated cloud vertical extent seems to be exaggerated over the western Indian Ocean as well as off the coast of Nordeste in SIMPCV2. On the planetary scale,

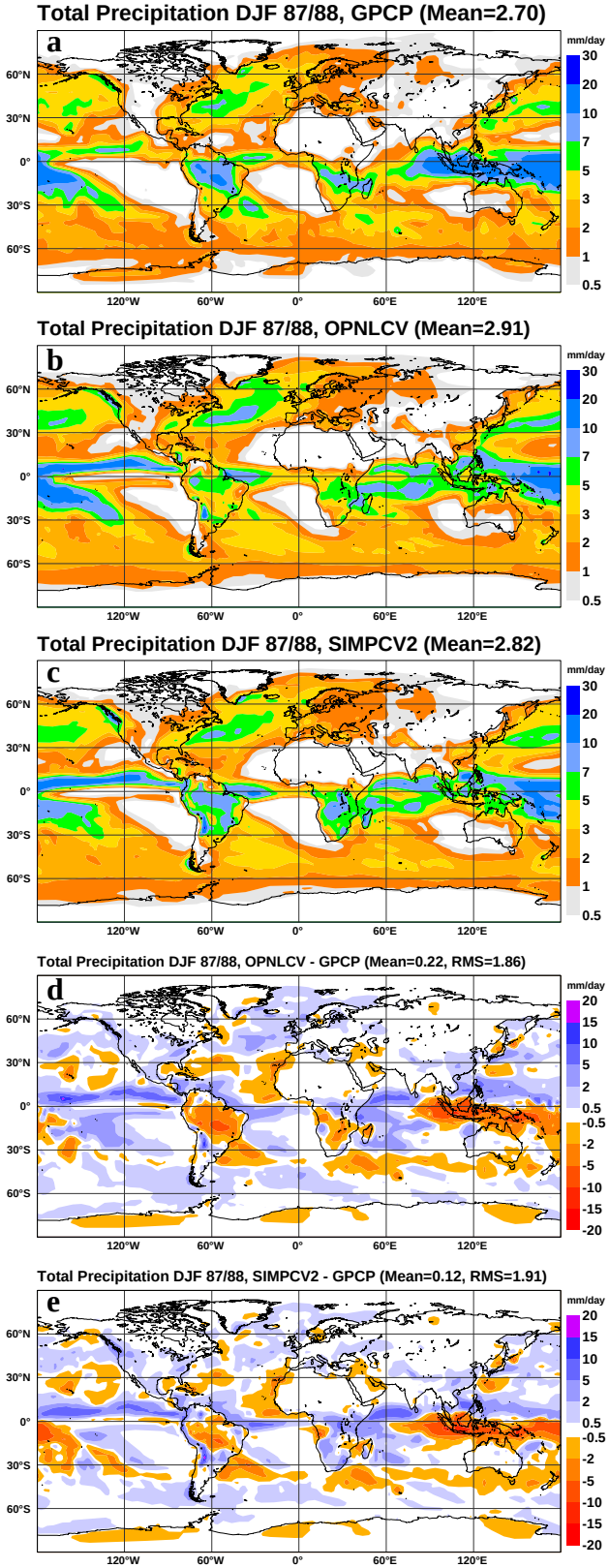

Figure 1: Total precipitation over winter 1987-88: (a) as from GPCP observations, (b) from OPNLCV and (c) from SIMPCV2 experiments (see text for more details). Differences OPNLCV-observations and SIMPCV2-observations are shown in panels (d) and (e) respectively. Mean and root mean square difference (when relevant) are mentioned at the top of each panel. Units are in mm day^{-1}.

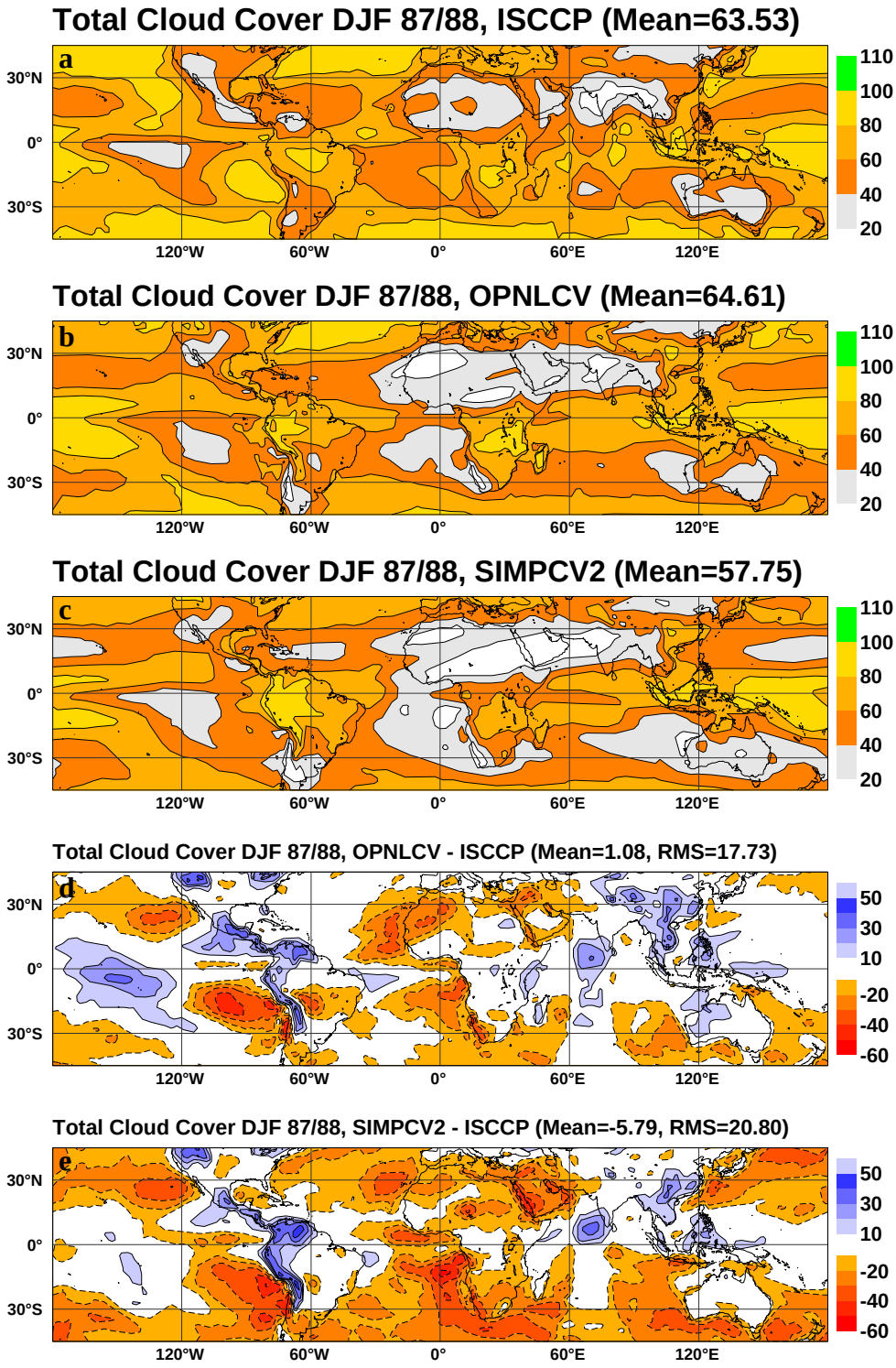

Figure 2: Same as in Fig. 1, but for total cloud cover using ISCCP observations. Units are in %.

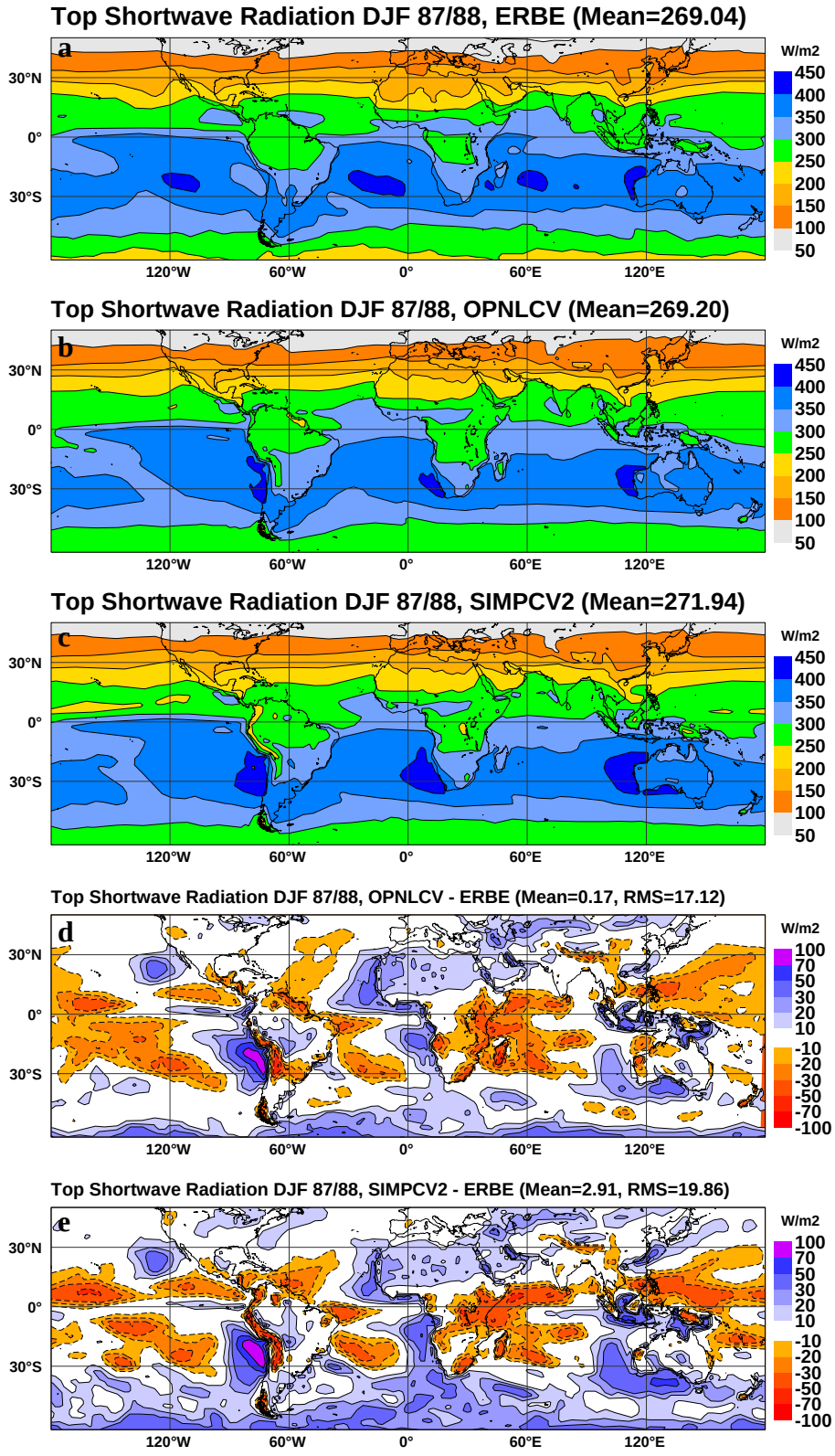

Figure 3: Same as in Fig. 1, but for top net shortwave radiation using ERBE observations. Units are in W m^{-2}.

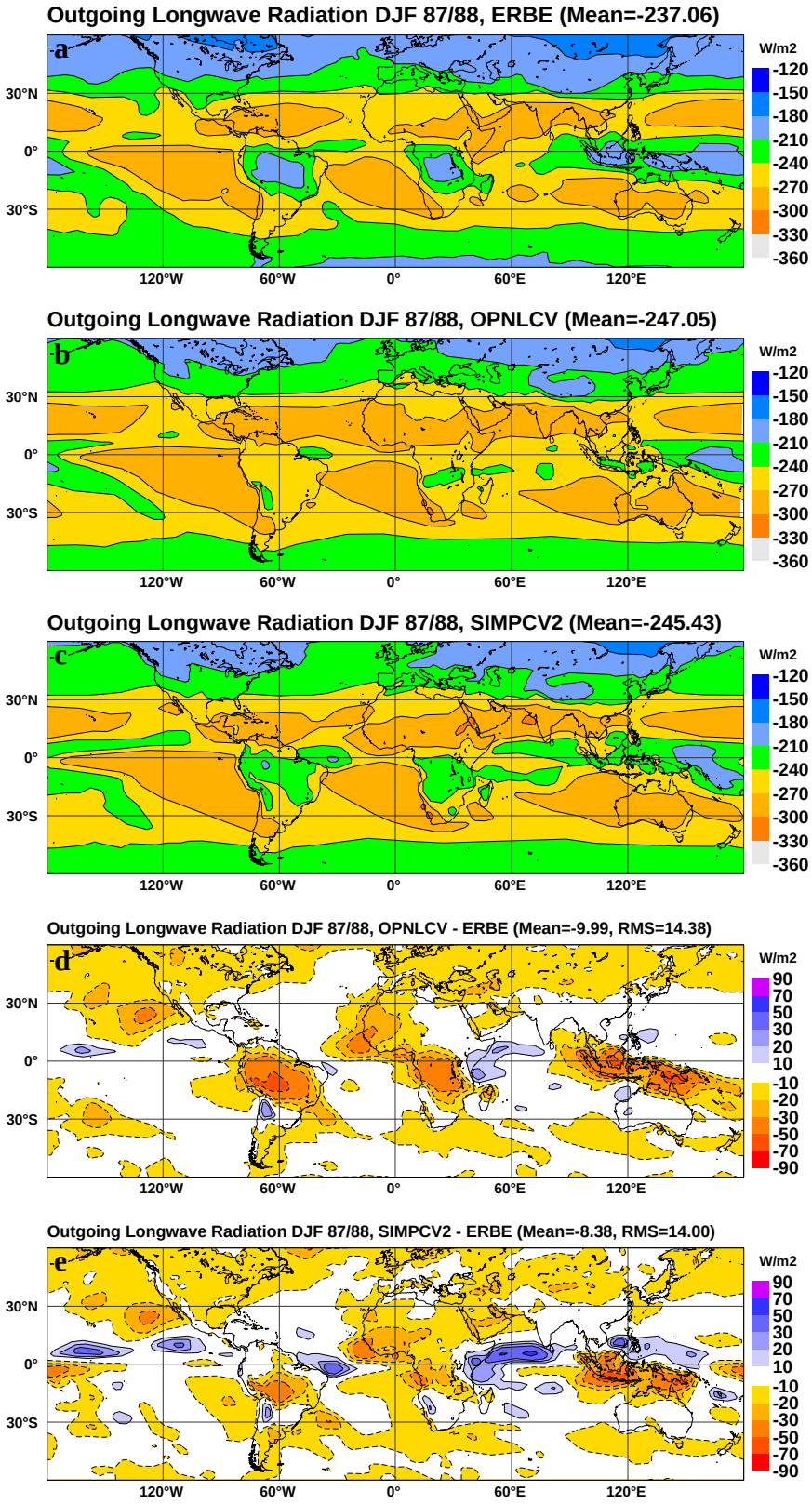

Figure 4: Same as in Fig. 1, but for outgoing longwave radiation at the top of the atmosphere using ERBE observations. Units are in W m^{-2}.

SIMPCV2 shows a SWR-TOA positive bias of 2.91 W m^{-2}, while OPNLCV is almost perfectly balanced (bias of 0.17 W m^{-2}). In terms of OLR-TOA, both experiments lead to comparable mean deviations respectively equal to -9.99 W m^{-2} and -8.38 W m^{-2}.

The results presented in this section demonstrate that, although not designed for the purpose of long simulations, SIMPCV2 behaves reasonably well with respect to observations and resembles the full operational convection scheme.

4 Linear behaviour

4.1 Jacobians

The sensitivities or Jacobians of the outputs of SIMPCV2 to the input temperature and specific humidity have been computed for several hundreds of initial atmospheric profiles using the finite-difference approach and by perturbing specific humidity and temperature on each vertical level, separately. The study of these Jacobians has permitted efficient isolation and modification of portions of the code that could lead to excessive sensitivities locally (see subsection 4.4). Sharp vertical gradients of sensitivities are not desirable since they can lead to irregular profiles of analysis increments (despite the subsequent smoothing effect of the background error covariance matrix).

Figures 5 and 6 display typical Jacobians of various output quantities from OPNLCV with respect to input specific humidity and temperature perturbations. Figure 5 focuses on the Jacobians of the convective tendencies of temperature and specific humidity and on the sensitivities of the convective updraught mass flux. Figure 6 displays the Jacobians of convective precipitation profiles and CAPE. Similar plots are shown in Fig. 7 and 8 for SIMPCV2. Note that the dynamic range of the colour scale is determined by the extreme values of the field plotted in each panel. For the selected grid point (4.39°N, 159.96°E), deep convection extends between model levels 55 (resp. 56) and 30 (resp. 28) with OPNLCV (resp. SIMPCV2), and surface convective rainfall is equal to 53.54 (resp. 85.82) mm day^{-1}. Table 1 in the appendix gives the correspondence between model and pressure levels.

Several conclusions can be drawn from these plots. First, the Jacobians of $(\frac{\partial T}{\partial t})_{conv}$ with respect to T and of $(\frac{\partial q}{\partial t})_{conv}$ with respect to q are confined around the main diagonal, which indicates that the effect of the perturbation remains localized. Secondly, the structure of the Jacobians of all other fields is almost entirely dictated by the shape of the CAPE Jacobians (bottom panels of Fig. 6 and 8). This is not surprising since the convective mass flux is directly related to CAPE through Eq. (11). In contrast, it should be noted that the denominator of Eq. (11) has a negligible influence on the Jacobians. An additional source of sensitivity of surface rainfall to temperature only appears around model level 40 with OPNLCV. Thirdly, the sign of the Jacobians can easily be explained by the changes of vertical stability induced by temperature perturbations and by the changes in the water phase caused by specific humidity perturbations. Focusing on SIMPCV2 for instance, a heating applied below model level 46 means warmer air in the convective updraught, hence stronger positive buoyancy in the ascent via Eq. (8), enhanced CAPE (Fig. 8.f) and convective mass flux (Fig. 7.f), which finally results in more convective precipitation (Fig. 8.b and d). With OPNLCV such response to low-level temperature perturbations is confined to the lowest model level. Similarly and for both convection schemes, a moistening of the environment below level 40 implies more water vapour available in the updraught through entrainment, that is a stronger latent heating through condensation in the ascent, which eventually leads to higher CAPE values (Fig. 8.e) and enhanced precipitation (Fig. 8.a and c). Note that the peak of sensitivity seen at level 57 and 58 in Fig. 6.(e) results from the recent inclusion of a mixing of updraught characteristics below cloud base in OPNLCV (Bechtold *et al.* 2004). Overall, it should be emphasized that the Jacobians

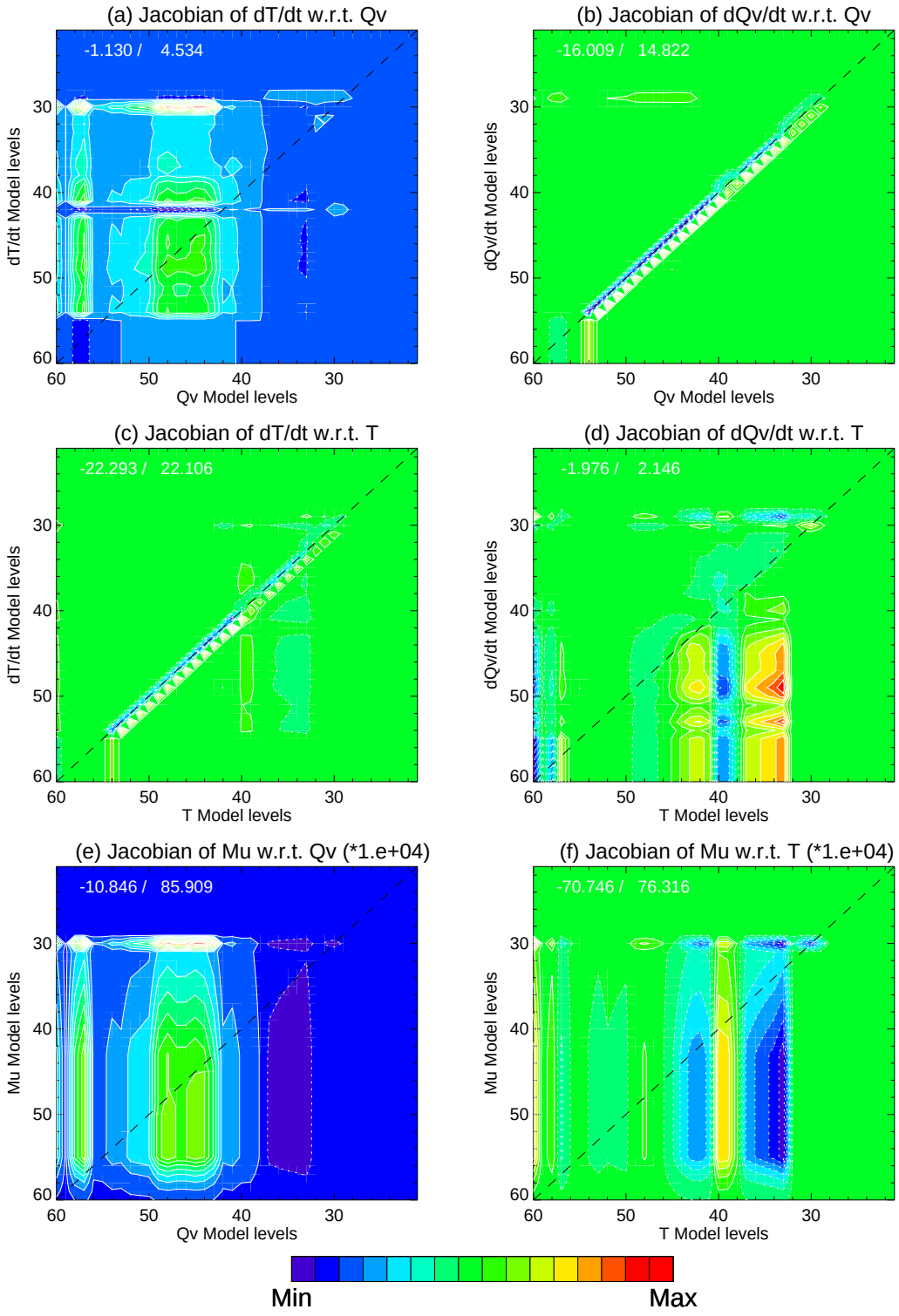

Figure 5: Jacobians of output temperature (a,c) and specific humidity (b,d) tendencies and convective updraught mass flux (e,f) with respect to input specific humidity and temperature for OPNLCV. The x-axis (resp. y-axis) shows the model levels at which the input (resp. output) variable is considered. Jacobians of the mass flux (e,f) have been scaled by 10^4. Extreme values of the Jacobians are given at the top left of each panel (SI units). Solid (resp. dashed) white isolines correspond to positive (resp. negative) sensitivities.

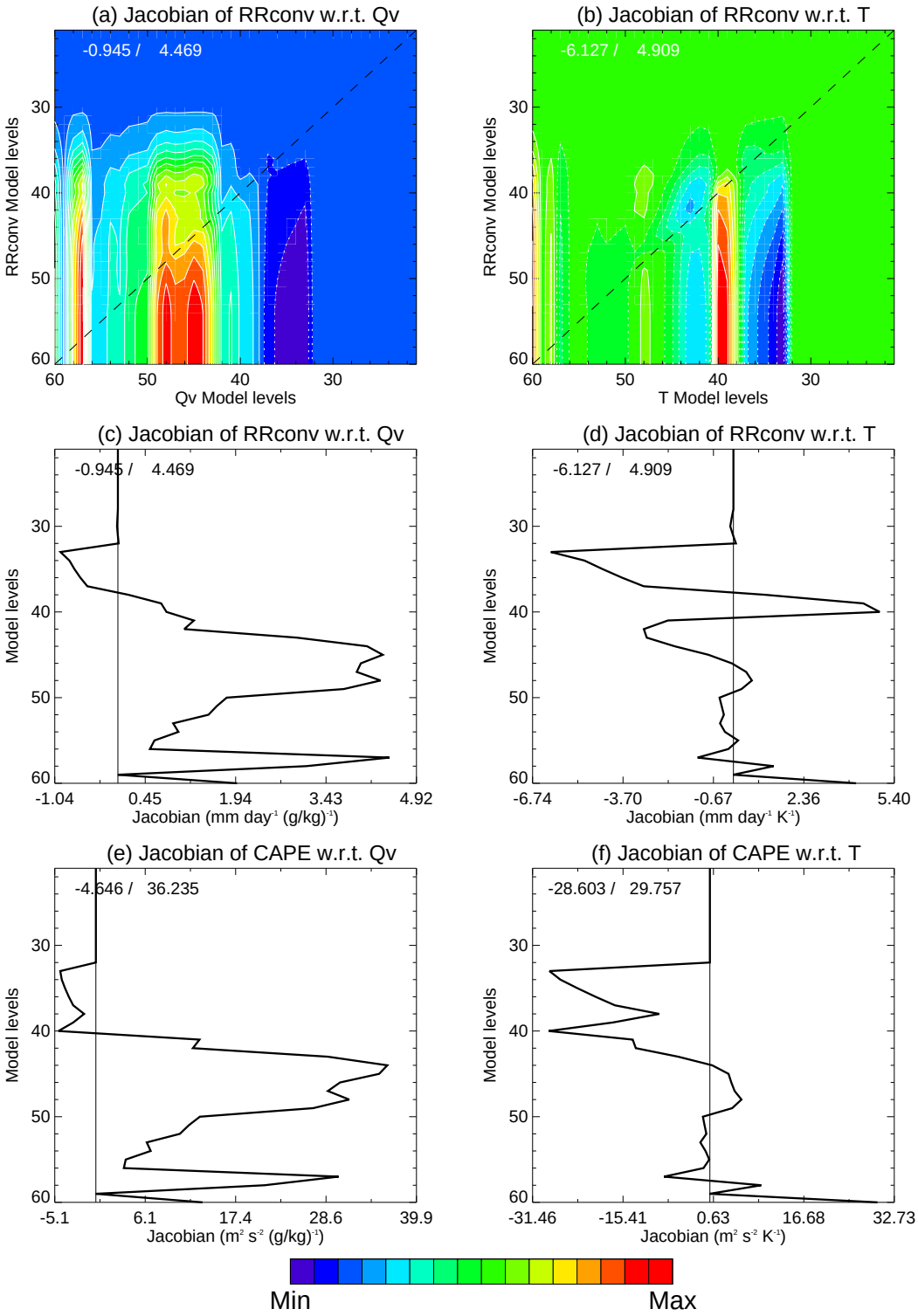

Figure 6: Jacobians of convective precipitation profiles (a,b), surface convective precipitation (c,d) and CAPE (e,f) with respect to input specific humidity and temperature for OPNLCV. First row: same layout as in Fig. 5; other panels: the x-axis shows the values of the Jacobians while the y-axis shows the model levels at which the input variable is perturbed.

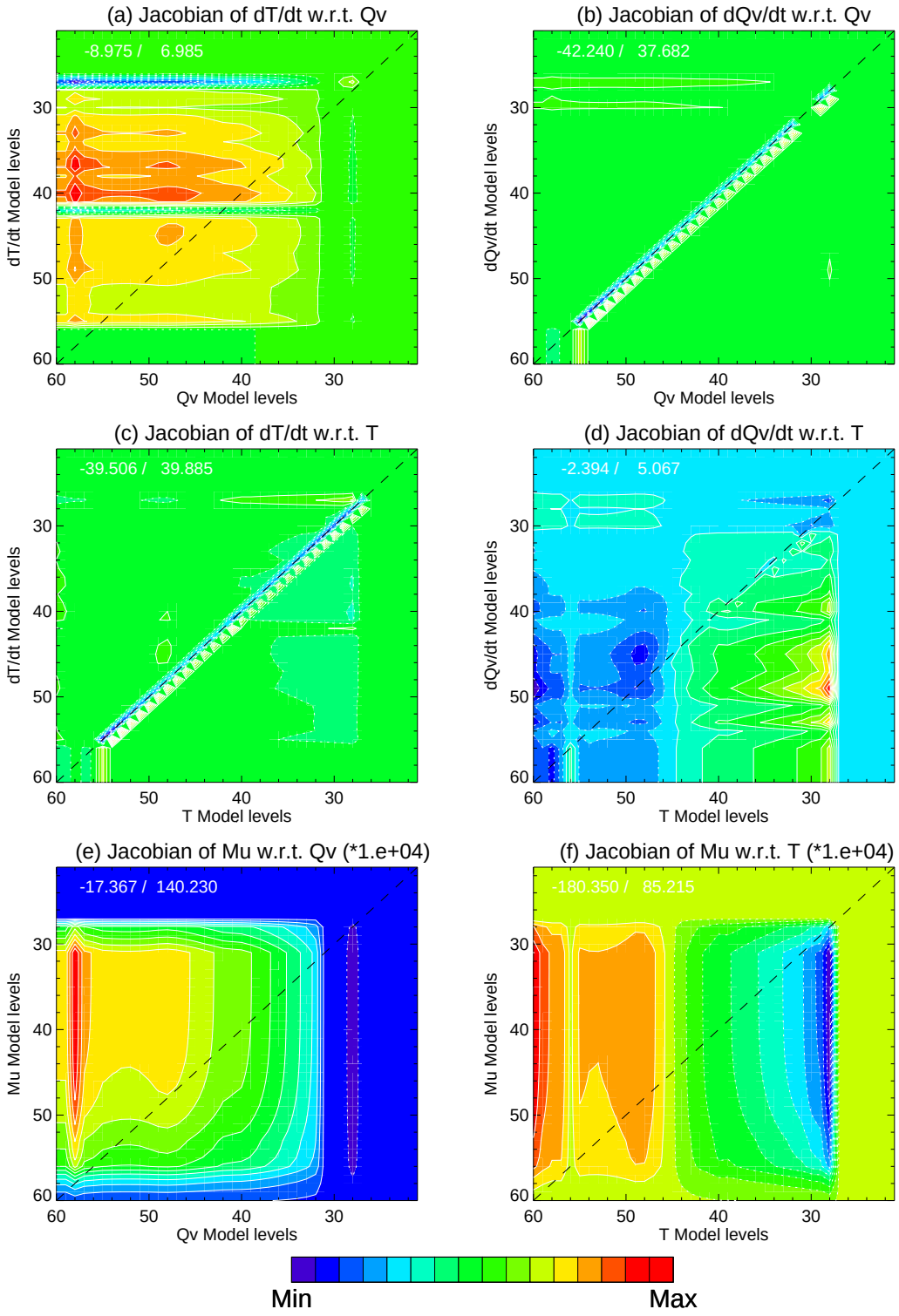

Figure 7: Same as in Fig. 5 but for SIMPCV2.

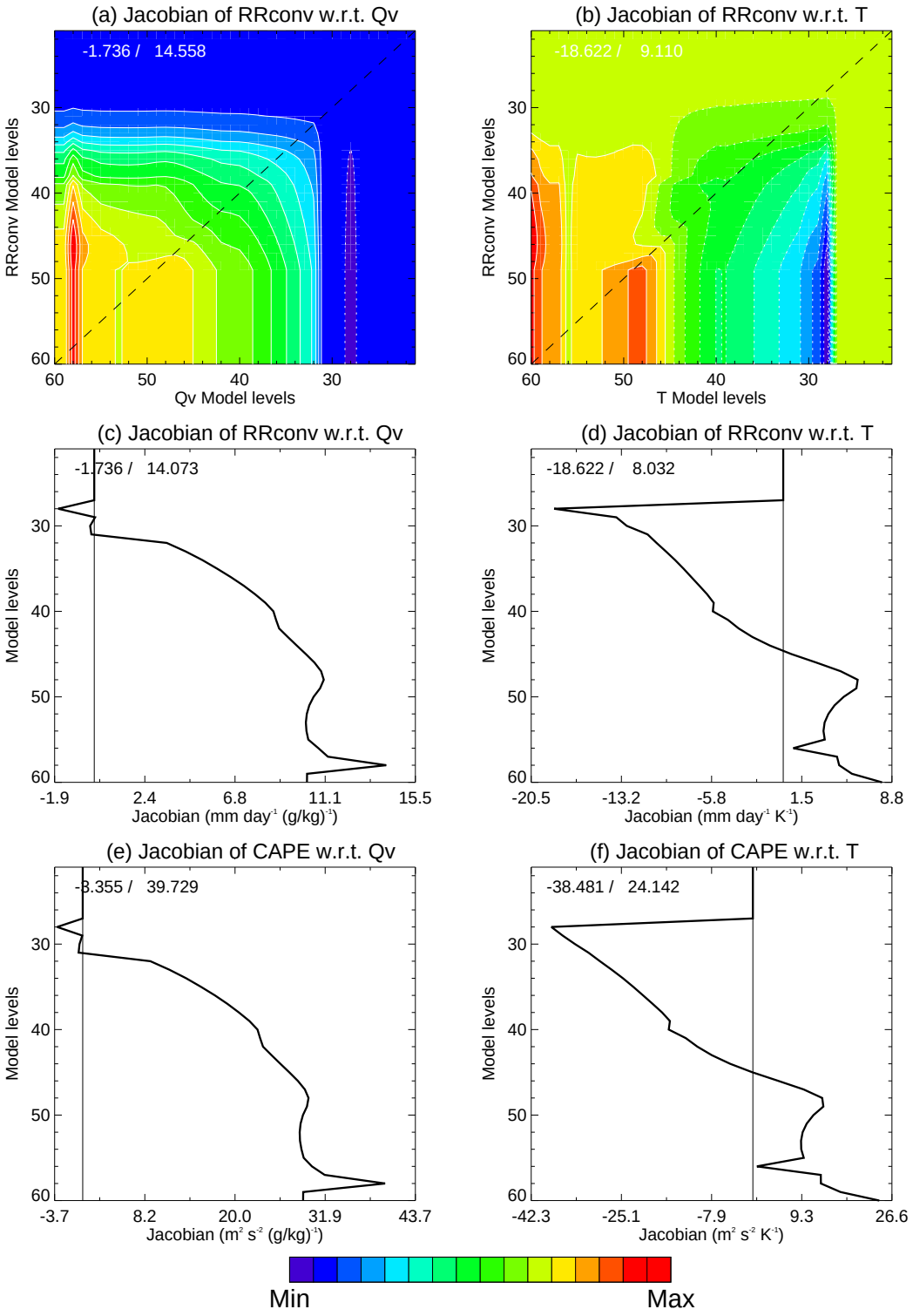

Figure 8: Same as in Fig. 6 but for SIMPCV2.

exhibit a smoother structure and a larger amplitude with SIMPCV2 than with OPNLCV.

4.2 Linearity over a single time-step

Reasonably well-behaved physical parametrizations in terms of linearity are necessary to ensure a successful convergence of the minimization in variational data assimilation. Therefore, the linear behaviour of the modified convection scheme has been assessed through the calculation of nonlinear residuals and compared to that of OPNLCV. For a given nonlinear operator M and its tangent-linear (TL) version $\mathbf{M}^{l}$, the nonlinear residuals are defined as

$$RES_{NL} = M(\mathbf{x}+\lambda\,\delta\mathbf{x}) - M(\mathbf{x}) - \mathbf{M}^{l}[\mathbf{x}]\lambda\,\delta\mathbf{x} \tag{14}$$

where $\mathbf{x}$ denotes the input temperature and specific humidity in the present case. The residuals have been computed for the convective tendencies $\partial T/\partial t$ and $\partial q/\partial t$ that are produced by the convective parametrization M. The calculations of Eq. (14) have been repeated for values of the scaling factor λ ranging from 10^{-5} to 1 and from -1 to -10^{-5}. The reference vector of perturbations $\delta\mathbf{x}$ was set equal to typical values of the standard deviation of the background model errors. An example of vertical profiles of such errors is shown in Fig. 9.

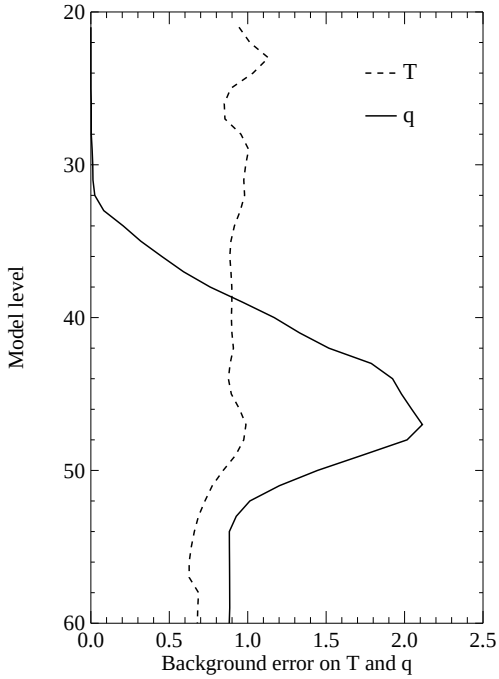

Figure 9: Vertical profiles of typical values of the standard deviation of the model background errors on temperature (dashed line; in K) and specific humidity (solid line; in g kg^{-1}).

Figures 10 and 11 show nonlinear residuals, averaged over 100 atmospheric profiles with deep convection, that are obtained when respectively temperature and specific humidity perturbations of varying size are applied at the lowest model level (60). Results are shown for this level because input perturbations imposed there usually bring large modifications of convective activity but similar conclusions can be drawn for other levels. The field actually plotted in Fig. 10 and 11 is $\log_{10}(RES_{NL}/RES_{NL}^{max})$ where RES_{NL}^{max} denotes the maximum value of the residual for a given plot and for both OPNLCV and SIMPCV2. This normalization permits the comparison of the residuals obtained with the two parametrizations and the smaller the values of the plotted field,

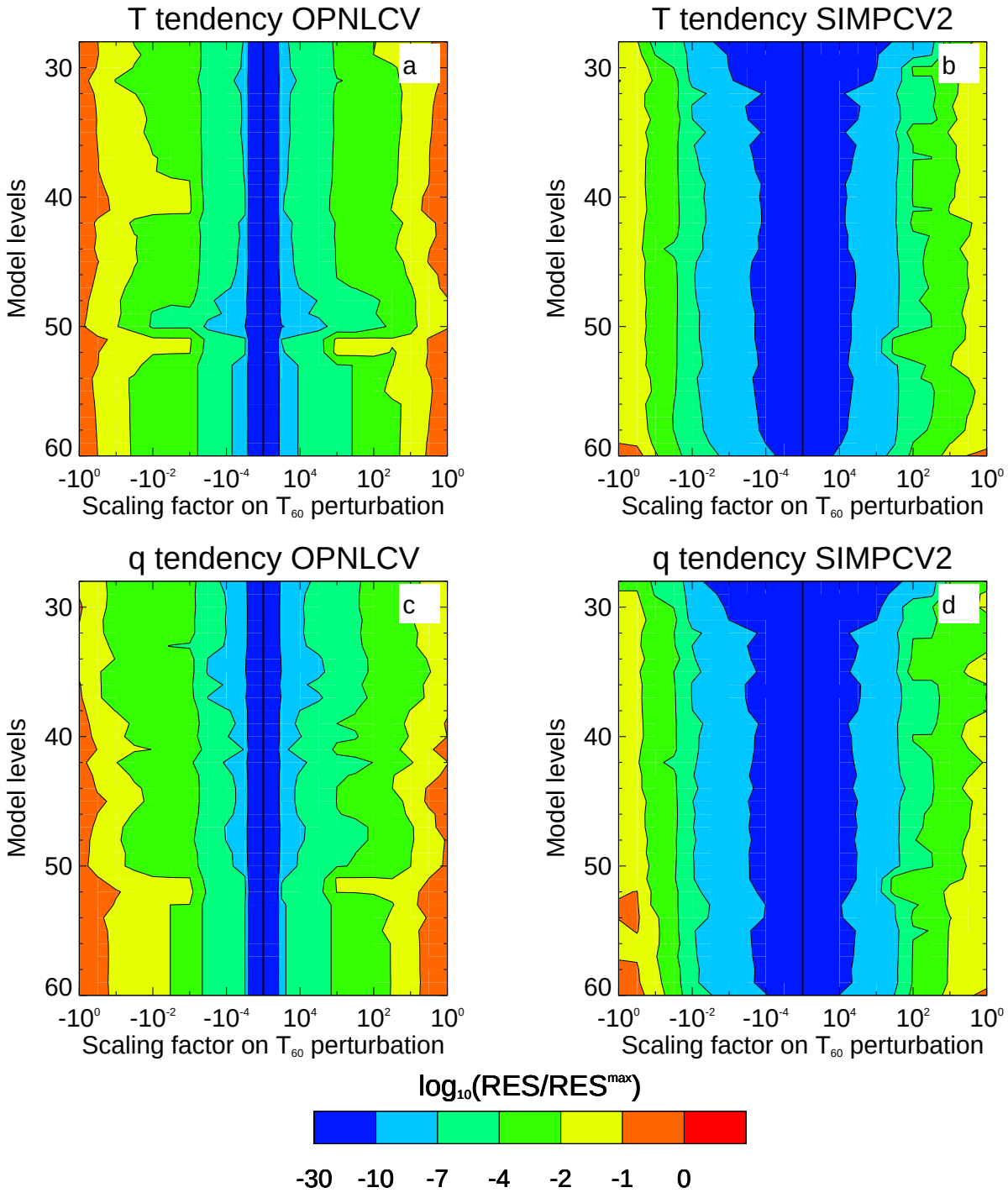

Figure 10: Nonlinear residuals of convective temperature and specific humidity tendencies as functions of parameter λ that determines the size of the input temperature perturbations imposed on model level 60. The displayed residuals have been averaged over one hundred convective profiles. The field actually plotted is $\log_{10}(RES_{NL}/RES_{NL}^{max})$ where RES_{NL}^{max} denotes the maximum value of the residual for a given pair of input/output parameters and for both studied convection schemes (*i.e.* for a given row of the figure). Left: OPNLCV, right: SIMPCV2.

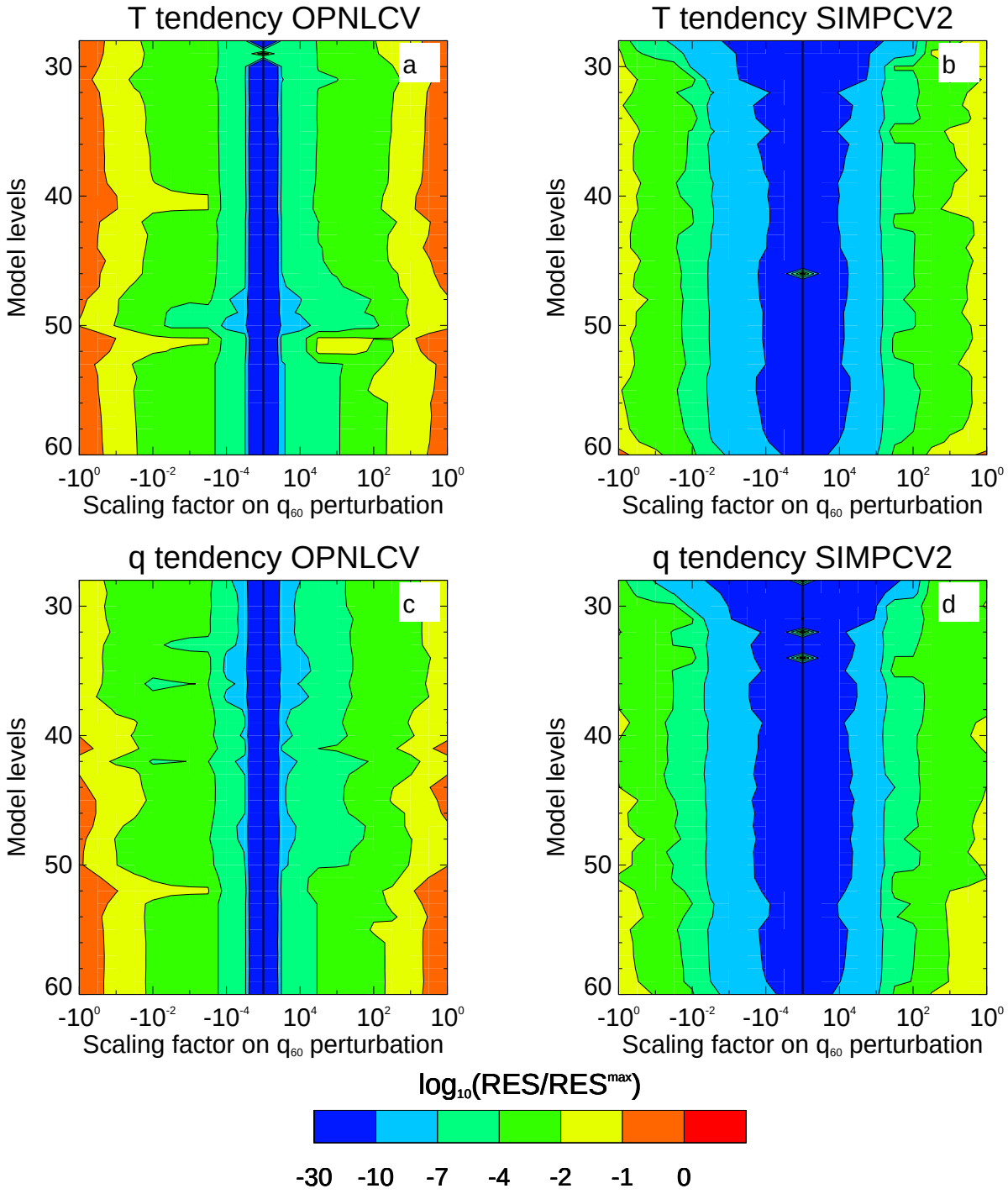

Figure 11: Same as in Fig. 10, but when input specific humidity perturbations are imposed on model level 60.

the more valid the linear assumption. Figures 10 and 11 demonstrate that the nonlinear residuals are almost systematically one order of magnitude smaller with the new scheme than with the operational one. The weak asymmetries about the y-axis result from the fact that large negative temperature or moisture perturbations at the surface can turn off convection thereby leading to a less linear behaviour of the scheme.

4.3 Validity of the linear assumption in time

In addition to the previous linearity tests performed on convection alone, it is crucial to verify that the inclusion of SIMPCV2 in TL integrations can lead to a better approximation of the finite difference of two nonlinear integrations when other processes such as vertical diffusion, large-scale condensation and radiation are also activated and when the perturbations involved are not necessarily limited in size. Indeed, in 4D-Var assimilation for instance, the magnitude of such perturbations may reach several K for temperature, several g kg^{-1} for specific humidity and a several m s^{-1} for wind.

The mean absolute error, E, between a TL run and the corresponding pair of nonlinear integrations is defined as

$$E \quad = \quad \overline{|M(\mathbf{x}+\delta\mathbf{x}) - M(\mathbf{x}) - \mathbf{M}'[\mathbf{x}]\delta\mathbf{x}|} \tag{15}$$

where $\mathbf{x}$ is equal to a given background model state and the perturbation $\delta\mathbf{x}$ is set to the difference between this background and the operational verifying analysis. The size of the initial perturbations imposed to the TL model is therefore that of analysis−background departures.

In the 2004 version of the ECMWF 4D-Var system, the model trajectory is calculated using OPNLCV, while the minimization and the evolution of analysis increments involve the simplified linearized convective parametrization SIMPCV. Since SIMPCV2 is an improved approximation to OPNLCV, its use in the TL integration should lead to a reduction of the linearity error. It is also important to check how SIMPCV2 behaves when combined with the new simplified parametrization of large-scale clouds and precipitation (CLOUDST, hereafter; Tompkins and Janisková 2004).

The integrations presented here are run over 12 hours with a spectral resolution of T159 and 60 vertical levels in order to match the current assimilation window and inner-loop horizontal and vertical resolutions of the ECMWF 4D-Var system. The full simplified physics is activated in all TL integrations but one (adiabatic run).

Figures 12 and 13 respectively display zonal mean error differences for temperature (in K) and specific humidity (in g kg^{-1}) when the TL model is integrated:

(a) with the operational simplified physical package (SIMPOP: vertical diffusion, gravity wave drag, deep convection, large-scale moist processes from MA99 and radiation from Janisková *et al.* (2002)) instead of no simplified physics at all (ADIAB),

(b) with SIMPCV2 as a substitute for SIMPCV, and

(c) with both SIMPCV2 and CLOUDST instead of SIMPOP moist physics.

The global relative change in the linearity error is mentioned at the top of each panel and negative values of the plotted field indicate improvements of the linearity.

First of all, panel (a) shows that the inclusion of the operational simplified physics improves the linearity error by 13.96% for temperature and 18.4% for specific humidity relative to the adiabatic TL run. Statistics computed for panel (b) of Fig. 12 and 13 indicate that SIMPCV2 further reduces the global mean TL error by 0.49% and

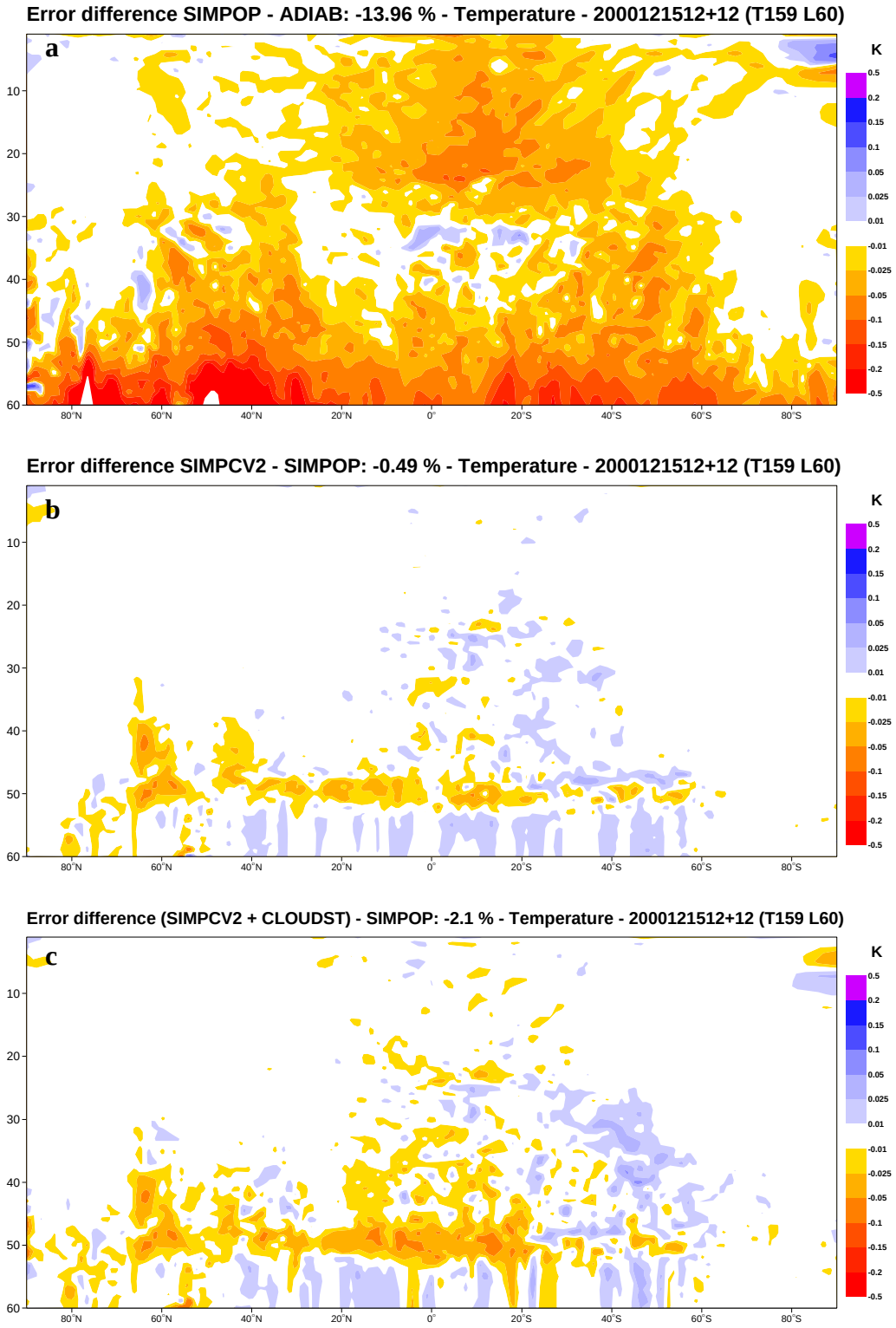

Figure 12: Impact of the new moist physics on the tangent-linear approximation error as defined in the text. Panel (a) shows the temperature error difference (in K) when SIMPOP is used in the tangent-linear calculations instead of no physics at all (adiabatic run). The two other panels show the further changes in the error due to the use (b) of SIMPCV2 and (c) of SIMPCV2 combined with CLOUDST. Error changes are displayed as zonal means and model levels are shown on the vertical axis. The global relative error reduction appears at the top of each panel.

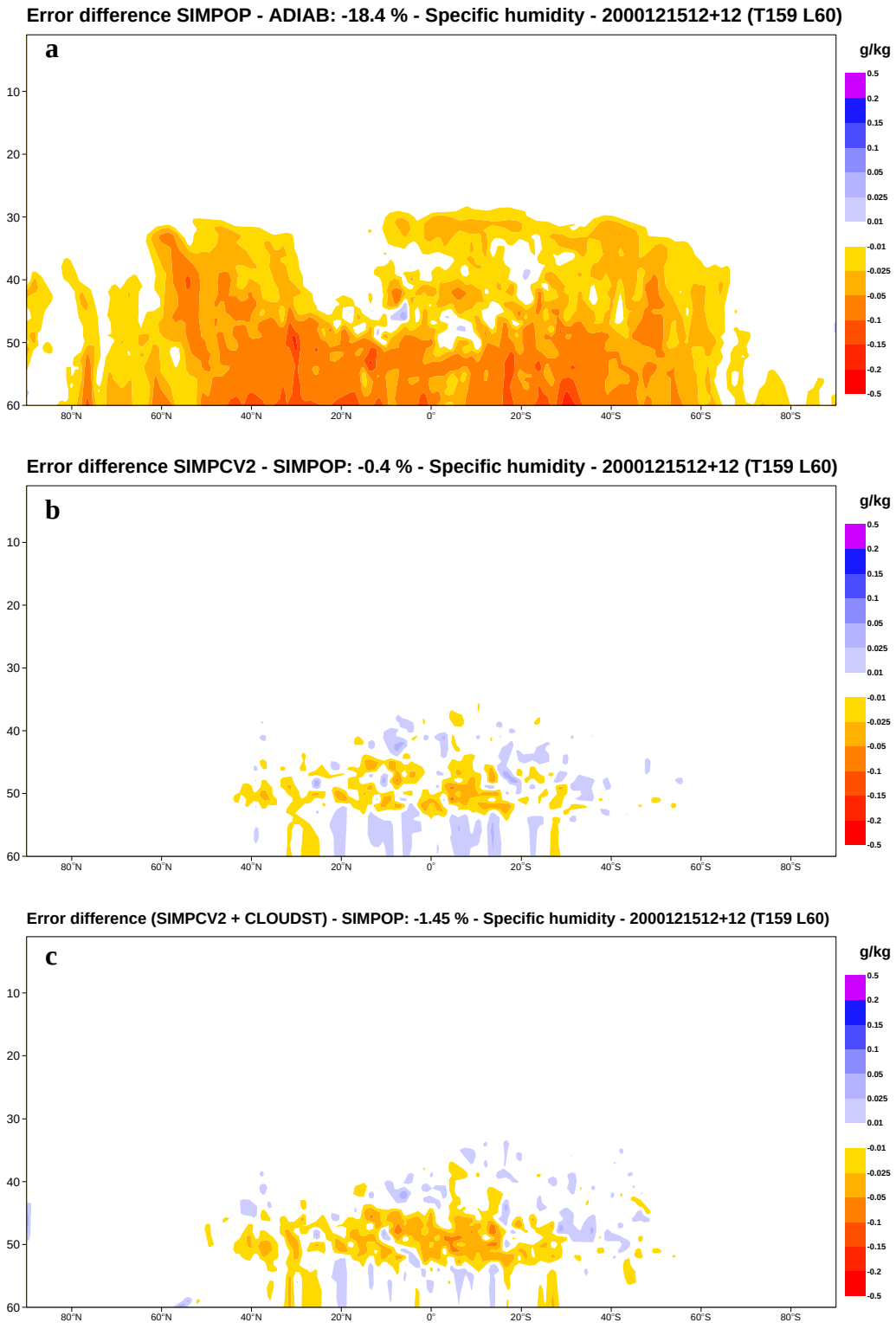

Figure 13: Same as in Fig. 12 but for specific humidity.

0.40% respectively. The strongest reduction of the error occurs between 60°S and 70°N and is confined to model levels 39 to 50. Between the surface and model level 50 the error is slightly increased. Above level 40, error changes remain small and limited to the Tropics. Figures 12.(c) and 13.(c) show that the combination of SIMPCV2 with CLOUDST brings an overall reduction of E equal to 2.10% for temperature and 1.45% for specific humidity with respect to the experiment with the current operational simplified moist physics.

4.4 Elimination of instability and nonlinearity sources

In some earlier Jacobian computations, some sources of excessive sensitivity were isolated and removed by adequate modifications of the code. These code adjustments consisted of

- the removal of the dependence of organized detrainment on the updraught vertical velocity,
- the specification of a constant updraught vertical velocity (3 m s^{-1}) in the calculation of precipitation formation (Eq. (12)) whenever the buoyancy becomes negative at levels below 7000 m.

It was also found that the activation of the mass flux limiter that is required to prevent numerical instability triggered by the violation of the Courant-Friedrich-Levy criterion leads to rather unrealistic Jacobians. For instance, the sensitivity of the surface rainfall rate with respect to temperature at lower levels becomes at least one order of magnitude smaller and opposite to that shown in the bottom right panel of Fig. 7. However, only about 5% of the total number of convective grid points are affected by the mass flux limiter for timesteps smaller than 1800 s and for the current model resolution (T511 L60).

Furthermore, in some TL evolution experiments run with SIMPCV2 a spurious amplification of TL perturbations occurred at a few geographical points after a few hours of integration, which then contaminated the entire Globe. In fact, such problems can arise when the perturbation of a given field can become as large as the field itself, which was found to be the case for

- the cloud base mass flux, M_u^{clb} from Eq. (11),
- the buoyancy, $\mathscr{B}(z)$, as defined in Eq. (8),
- the initial vertical velocity of the updraught, w_s, from Eq. (4).

To avoid these instabilities, the corresponding perturbations are reduced by a factor of 0.25 for M_u^{clb} and 0.35 for $\mathscr{B}(z)$ and w_s. For the latter field, the limitation is applied only when w_s is lower than 0.5 m s^{-1}. These factors were tuned so as to provide a good compromise between stability and validity of the linear assumption, and are required when running SIMPCV2 in all tangent-linear and adjoint calculations involving time integration.

5 Applications

5.1 1D-Var in rainy areas

SIMPCV2 has been extensively tested in 1D-Var retrievals of temperature and specific humidity profiles from observed rainfall rate (RRs) and microwave brightness temperatures (TBs). The two corresponding methods will be abbreviated as 1D/RR and 1D/TB hereafter. In such experiments, vertical profiles of increments of temperature and specific humidity are obtained by minimizing a cost function that measures the fit of the model state both to the observations and to a background state. Error statistics on model and observations need to be properly specified. The new moist physical package (SIMPCV2+CLOUDST) is called to convert the model state (T,q) into precipitation fields that can be used in turn to simulate microwave brightness temperatures

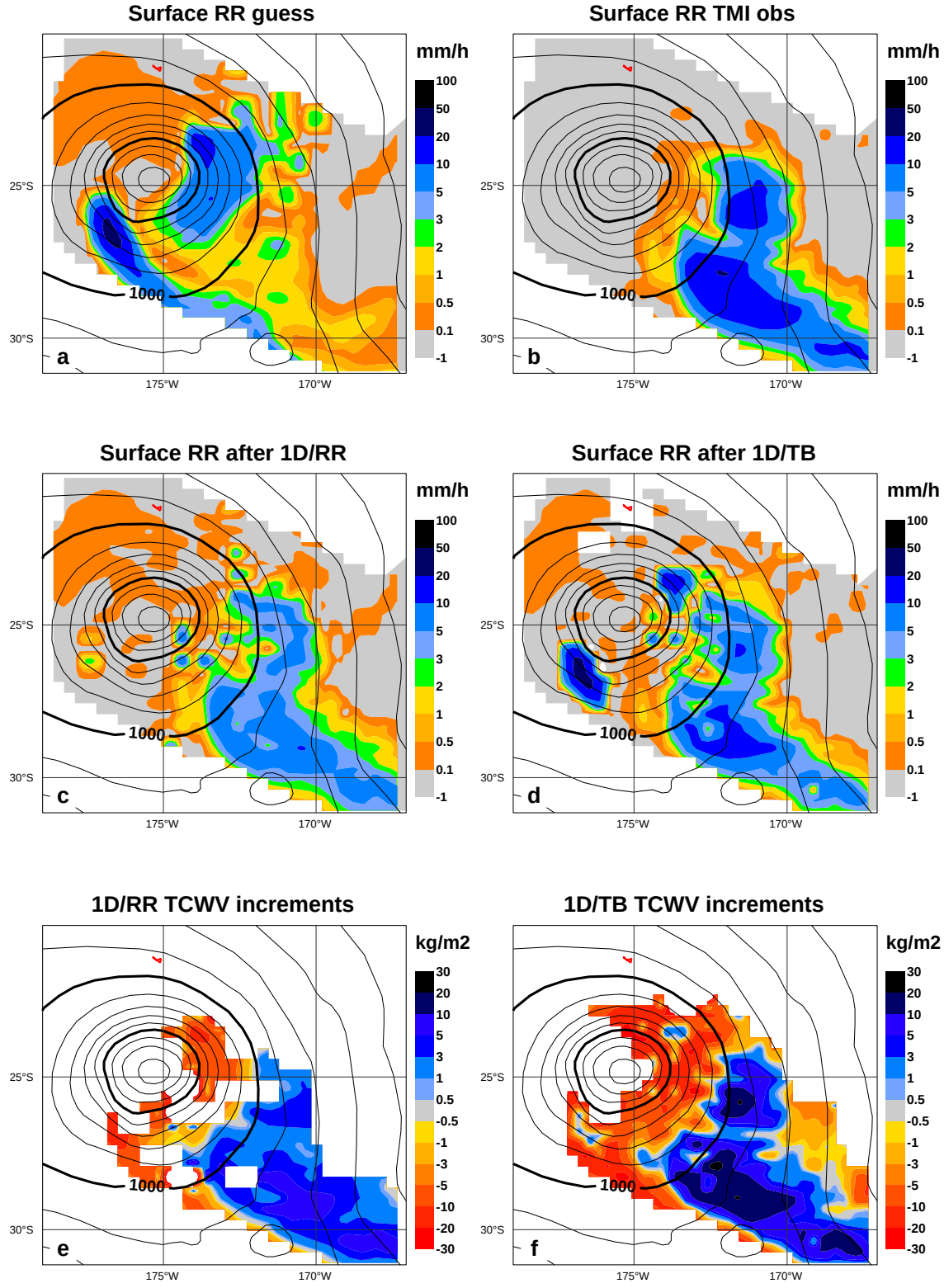

Figure 14: 1D-Var experiments using SIMPCV2 and CLOUDST: (a) surface RRs as simulated from the background state of (T,q), (b) as derived from TRMM multi-channel microwave TBs using the PATER algorithm, and as recomputed from the modified model (T,q) profiles after (c) 1D/RR and (d) 1D/TB (using the 10, 19, 22 and 37 GHz channels in vertical and horizontal polarizations). 1D-Var Total Column Water Vapour (TCWV) increments from (e) 1D/RR and (f) 1D/TB are also plotted. In all panels background mean-sea-level isobars are drawn every 2 hPa.

using the radiative transfer model of Bauer (2002). The iterative minimization of the cost function involves the adjoint version of the moist physics. The adjoint of the radiative transfer model is also required in 1D/TB. More technical details on 1D/RR and 1D/TB in precipitating areas can be found in Moreau *et al.* (2004).

An example of 1D-Var retrieval using observations from the Tropical Rainfall Measuring Mission (TRMM) polar-orbiting satellite is shown in Fig. 14 for tropical cyclone Ami at 1800 UTC 14 January 2003 in the western Pacific. Note that the surface rainfall rates used as observations in 1D/RR are obtained by applying the retrieval algorithm PATER developed by Bauer *et al.* (2001). Figure 14 demonstrates that both 1D-Var approaches are able to produce adjustments of the model's temperature and specific humidity that improve surface rainfall rates with respect to the PATER observations. This is even more remarked in the case of 1D/TB since this method works directly on the TBs and has therefore no explicit information on precipitation. In the proposed example, 1D/TB has been applied to the 10, 19, 22 and 37 GHz channels in vertical and horizontal polarizations. A rather systematic and common feature in 1D/RR and 1D/TB using SIMPCV2 and CLOUDST is the predominance of specific humidity increments relative to those in temperature (after conversion into equivalent increments of saturation specific humidity). This supports the choice made by Marécal and Mahfouf (2003) to define TCWV pseudo-observations that are obtained by vertical integration of the 1D-Var specific humidity increments. These TCWV pseudo-observations can in turn be fed into the 4D-Var assimilation system. In other respects, one should also note that the increments are usually systematically twice as large with 1D/TB as with 1D/RR. This is the result of the sensitivity of the selected microwave frequencies not only to rain but also to cloud liquid water content and water vapour, but also to a lesser extent to the specification of the observation error statistics. 1D/TB retrievals only using the 22 GHz channel lead to TCWV increments that are more comparable in magnitude to those from 1D/RR (not shown).

In the near future, SIMPCV2 and CLOUDST will be used operationally in the 1D-Var calculations involved in the '1D-Var + 4D-Var' method proposed by Marécal and Mahfouf (2003) to assimilate TBs from the Special Sensor Microwave Imager (SSM/I). Eventually, the new simplified moist physics will be tested within 4D-Var itself.

5.2 Sensitivity experiments

The new moist physics has also been compared with the operational one to compute the sensitivity to initial conditions of a given characteristic, J, of the forecast at the final time over a selected geographical domain, $\mathscr{D}$. The gradient of J with respect to the model's initial state, $\nabla_0 J$, is obtained from the gradient of J at final time, $\nabla_f J$, through the integration of the adjoint model with simplified moist physics, M^*. Formally this relation can be written

$$\nabla_0 J \quad = \quad \mathbf{B}^{-\frac{1}{2}} \, M^* \, P^* \, P \, \nabla_f J \tag{16}$$

where the norm based on the ECMWF 4D-Var background error covariance matrix, **B**, (Derber and Bouttier 1999) is used at the initial time, P is the geographical projector and the asterisk denotes the adjoint operator.

In the example described here, the forecast characteristic of interest, J, is chosen to be the quadratic error of specific humidity

$$J \quad = \quad \iint_{\mathscr{D}} \int_P \left(q_{fc}(P) - q_{an}(P)\right)^2 \frac{dP}{g} \, d\mathscr{D} \tag{17}$$

where q_{fc} is the specific humidity from a 48-hour forecast valid at 1200 UTC 12 November 1999 and q_{an} the same field from the verifying analysis. J involves both vertical integration and horizontal accumulation over domain $\mathscr{D}$ with longitudes ranging from 0° to 5°E and latitudes between 40°N and 45°N. The main patterns of

mean-sea-level pressure (MSLP) and 500 hPa geopotential height at the beginning and at the end of 48-hour window are summarized in Fig. 15. By 1200 UTC 12 November 1999, a low had developed over Spain, which led to torrential rainfall over southern France (up to 620 mm in 24 hours locally; Bechtold and Bazile 2001) due to the enhancing effect of the intense moisture supply through surface evaporation over the Gulf of Lyons.

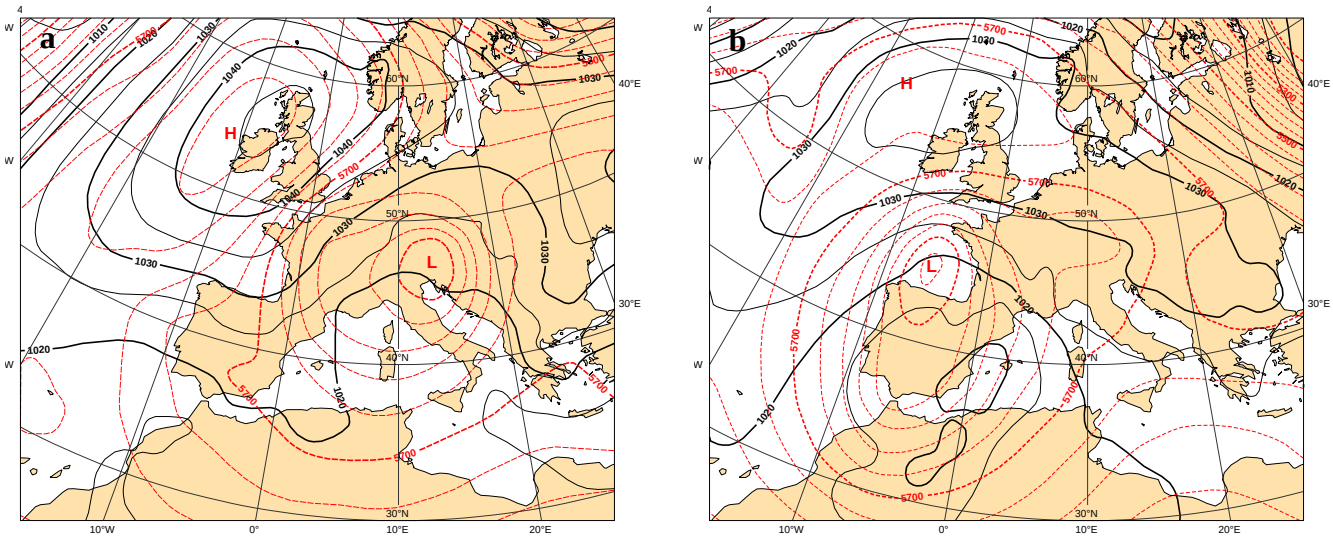

Figure 15: MSLP (black solid line; every 5 hPa) and 500 hPa geopotential height (red dashed line: every 40 m) fields from the operational analysis at 1200 UTC (a) 10 November and (b) 12 November 1999.

Sensitivity fields are compared using either the SIMPOP moist physics package or the revised one (SIMPCV2 and CLOUDST) in the adjoint computations. Figure 16 displays the sensitivities of J computed at 1200 UTC 12 November 1999 to specific humidity on model levels 46 (left) and 55 (right): panels (a) and (b) show the sensitivities at the beginning of the adjoint integration while the four other panels display the sensitivities at the end of the 48-hour backward adjoint integration using either the SIMPOP moist physics (c and d) or SIMPCV2+CLOUDST (e and f). Note that sensitivities have been normalized by the pressure thickness of the corresponding model layer.

Figure 16.(a) and (b) exhibit expected patterns of $\frac{\partial J}{\partial q}$ at the beginning of the adjoint integration. Cross-sections (not shown) suggest that the largest sensitivities at the end of the adjoint integration are confined between model level 43 and the surface, with maximum values at level 46. Furthermore, sensitivity patterns obtained with SIMPCV2+CLOUDST are usually similar to those obtained with the SIMPOP moist physics as illustrated by panels (c) and (e). Some significant differences only appear below model level 50 over the Gulf of Lyons as indicated by panels (d) and (f), which could be attributed to the direct link between surface heat fluxes and convective activity in SIMPCV2 through Eq. (4)-(6). These preliminary sensitivity experiments demonstrate that the new simplified moist physics is able to remain stable during the 48 hours of adjoint integration (*i.e.* no spurious amplification of perturbations) and that it produces sensitivity patterns that are consistent with the operational ones. There seems to be a potential additional impact of physical processes that are explicitly represented in SIMPCV2 and CLOUDST adjoint versions but not in SIMPOP. Whether this impact is actually beneficial will be determined in the context of forecast sensitivity experiments as well as in singular vector calculations (used for instance to generate initial perturbations in an ensemble prediction system) and of course in 4D-Var data assimilation.

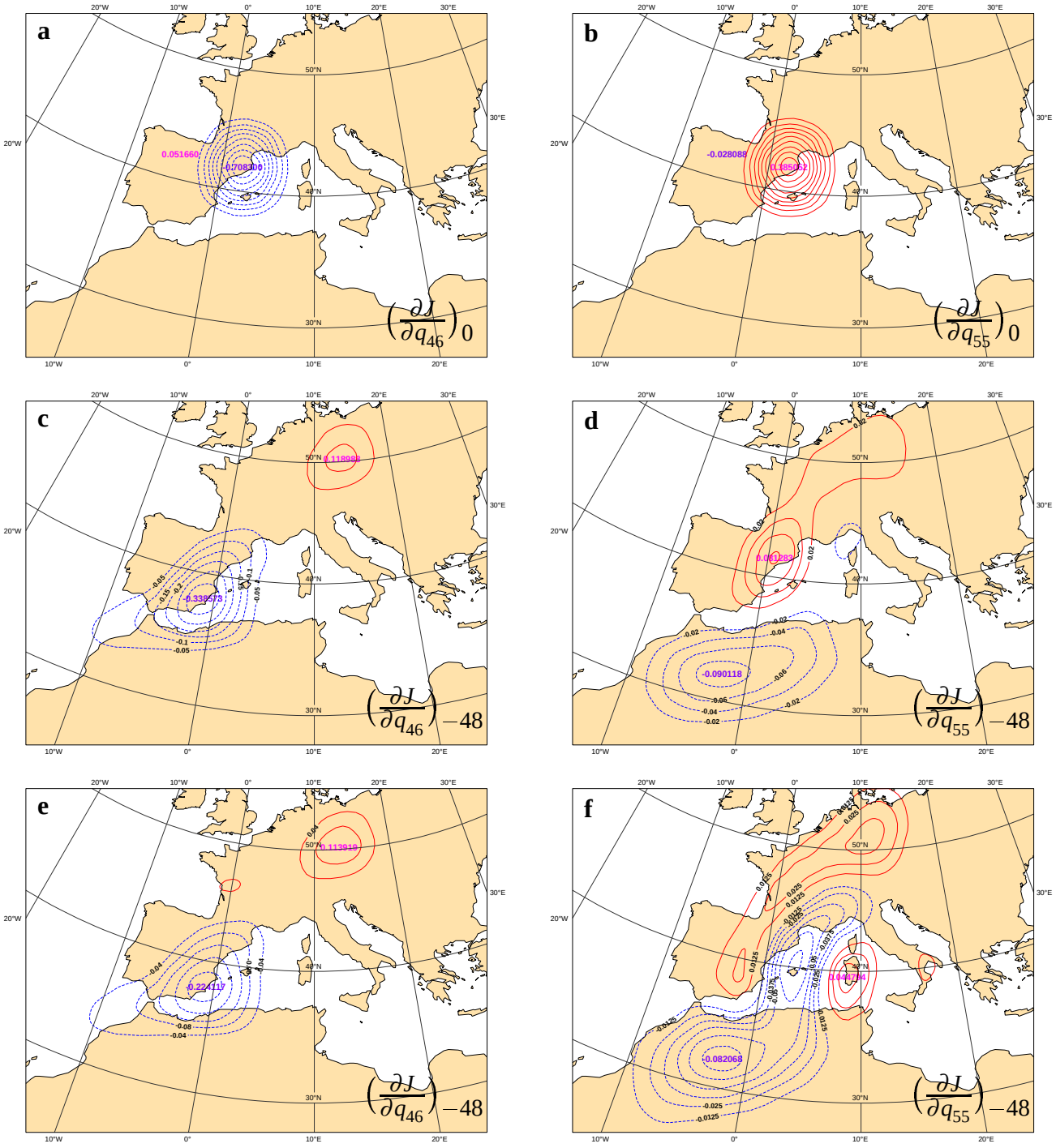

Figure 16: Sensitivities of specific humidity forecast error at 1200 UTC 12 November 1999 to specific humidity on model levels 46 (left panels) and 55 (right panels) at the same time (a and b) and 48 hours beforehand when using the SIMPOP moist physics (c and d) or SIMPCV2+CLOUDST (e and f) in adjoint calculations. Positive and negative values of sensitivity are shown with red solid and blue dashed isolines respectively. Sensitivities are rendered dimensionless by the normalization applied (see text).

6 Conclusions

A new simplified parametrization of subgrid scale convective processes has been developed and tested in the framework of the ECMWF Integrated Forecasting System for the purpose of variational data assimilation, singular vector calculations and adjoint sensitivity experiments. It is based on the full nonlinear scheme that is used in the ECMWF operational forecasts, but it was demonstrated that a set of simplifications could lead to a substantial improvement of its linear behaviour without significantly degrading its forecasting ability even in seasonal integrations. The new scheme has also been successfully combined with the new simplified parametrization of large-scale clouds and precipitation developed by Tompkins and Janisková (2004). In contrast with the MAH99 simplified convective parametrization, its tangent-linear and adjoint versions account for perturbations of all convective quantities including convective mass flux, updraught characteristics and precipitation fluxes. Therefore the new scheme is expected to be beneficial in the 4D-Var context when combined with radiative calculations that are directly affected by condensation and precipitation. In particular, this is the reason why SIMPCV2 has been successfully used in 1D-Var retrievals of TCWV from microwave brightness temperatures in precipitating areas (Moreau *et al.* 2004). In the near future, the impact of the new simplified moist physics in 4D-Var data assimilation itself will be tested.

Acknowledgements

We would like to thank Jean-François Mahfouf for his key-role in initiating this work and helping us in its early stage. Marta Janisková, Adrian Tompkins, Martin Miller, Anton Beljaars, Peter Bechtold, Peter Bauer and Jean-Noël Thépaut should be acknowlegded for their useful reviews of this paper. We are grateful to INRIA (Institut National de Recherche en Informatique et en Automatique) for providing the M1QN3 mimization code. This study was supported by the European Space Agency projects EuroTRMM-2 (#12651/97/NL/NB) and EGPM (#3-10600/02/NL/GS).

References

Bauer, P. (2002). Microwave radiative transfer modelling in clouds and precipitation. Part I: Model description. Technical report. Satellite Application Facility for Numerical Weather Prediction, NWPSAF-EC-TR-005, version 1.0.

Bauer, P., Amayenc, P., Kummerow, C. D., and Smith, E. A. (2001). Over-ocean rainfall retrieval from multisensor data of the Tropical Rainfall Measuring Mission (TRMM). Part II: Algorithm implementation. *J. Ocean. Atmos. Tech.*, 18:1838–1855.

Bechtold, P. and Bazile, E. (2001). The 12-13 november 1999 flash flood in southern france. *Atmos. Research*, 56:171–189.

Bechtold, P., Chaboureau, J.-P., Beljaars, A., Betts, A. K., Köhler, M., Miller, M., and Redelsperger, J.-L. (2004). The simulation of the diurnal cycle of convective precipitation over land in a global model. *Q. J. R. Meteorol. Soc.*, 130. in press.

Cohen, C. (2000). A quantitative investigation of entrainment and detrainment in numerically simulated cumulonimbus clouds. *J. Atmos. Sci.*, 57:1657–1674.

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.*, 120:1367–1388.

Derber, J. and Bouttier, F. (1999). A reformulation of the background error covariance in the ECMWF global assimilation system. *Tellus*, 51A:195–221.

Fillion, L. and Mahfouf, J.-F. (2000). Coupling of moist-convective and stratiform precipitation processes for variational data assimilation. *Mon. Weather Rev.*, 128:109–124.

Gregory, D., Morcrette, J.-J., Jakob, C., Beljaars, A. C. M., and Stockdale, T. (2000). Revision of convection, radiation and cloud schemes in the ECMWF Integrated Forecasting System. *Q. J. R. Meteorol. Soc.*, 126:1685–1710.

Holtslag, A. A. M. and Moeng, C.-H. (1991). Eddy diffusivity and countergradient transport in the convective atmospheric boundary layer. *J. Atmos. Sci.*, 48:1690–1698.

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

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

Mahfouf, J.-F., Buizza, R., and Errico, R. M. (1996). Strategy for including physical processes in the ECMWF variational data assimilation system. In *Proceedings of the ECMWF seminar on data assimilation, 2–6 September*, pages 595–632. Available from ECMWF, Shinfield, Reading, UK.

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

Marécal, V. and Mahfouf, J.-F. (2000). Variational retrieval of temperature and humidity profiles from TRMM precipitation data. *Mon. Weather Rev.*, 128:3853–3866.

Marécal, V. and Mahfouf, J.-F. (2003). Experiments on 4D-Var assimilation of rainfall data using an incremental formulation. *Q. J. R. Meteorol. Soc.*, 129:3137–3160.

Moreau, E., Lopez, P., Bauer, P., Tompkins, A. M., Janisková, M., and Chevallier, F. (2004). Variational retrieval of temperature and humidity profiles using rain rates versus microwave brightness temperatures. *Q. J. R. Meteorol. Soc.*, 130:827–852.

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

Simpson, J. and Wiggert, V. (1969). Models of precipitating cumulus towers. *Mon. Weather Rev.*, 97:471–489.

Tiedtke, M. (1989). A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Weather Rev.*, 117:1779–1800.

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

Vukićević, T. and Errico, R. M. (1993). Linearization and adjoint of parameterized moist diabatic processes. *Tellus*, 45A:493–510.

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.

Županski, D. (1993). The effects of discontinuities in the Betts-Miller cumulus convection scheme on four-dimensional variational data assimilation. *Tellus*, 45A:511–524.

Županski, D. and Mesinger, F. (1995). Four-Dimensional Variational Assimilation of Precipitation Data. *Mon. Weather Rev.*, 123:1112–1127.

Model level number	Pressure (hPa)
5	1
10	4
15	12
20	36
25	96
30	202
35	353
40	539
45	728
50	884
55	979
60	1012

Table 1: Pressure (in hPa) on every five model levels between levels 5 and 60. Values are given assuming a surface pressure of 1013.25 hPa.

Acronym	Description
OPNLCV	Nonlinear convection scheme as used in operational forecasts and 4D-Var trajectory (cycle 26r3, Tiedtke 1989)
SIMPCV	Simplified convection scheme as used in operational tangent-linear and adjoint calculations (cycle 26r3; Mahfouf 1999)
SIMPOP	All simplified physical parametrizations as used in operational tangent-linear and adjoint calculations (cycle 26r3; Mahfouf 1999; Janisková *et al.* 2002)
CLOUDST	New simplified parametrization of clouds and precipitation (Tompkins and Janisková 2004)
SIMPCV2	New simplified convection scheme (topic of this paper)

Table 2: List of acronyms used in the text.