

# On the potential of assimilating meteorological analyses in a global climate model for the purpose of model validation

A. B. M. Jeuken, P. C. Siegmund and L. C. Heijboer <sup>1</sup>

Royal Netherlands Meteorological Institute, De Bilt, The Netherlands

J. Feichter and L. Bengtsson

Max-Planck-Institute for Meteorology, Hamburg, Germany

**Abstract.** A simple four-dimensional assimilation technique, called Newtonian relaxation, has been applied to the Hamburg climate model (ECHAM), to enable comparison of model output with observations for short periods of time. The prognostic model variables vorticity, divergence, temperature, and surface pressure have been relaxed toward European Center for Medium-Range Weather Forecasts (ECMWF) global meteorological analyses. Several experiments have been carried out, in which the values of the relaxation coefficients have been varied to find out which values are most usable for our purpose. To be able to use the method for validation of model physics or chemistry, good agreement of the model simulated mass and wind field is required. In addition, the model physics should not be disturbed too strongly by the relaxation forcing itself. Both aspects have been investigated. Good agreement with basic observed quantities, like wind, temperature, and pressure is obtained for most simulations in the extratropics. Derived variables, like precipitation and evaporation, have been compared with ECMWF forecasts and observations. Agreement for these variables is smaller than for the basic observed quantities. Nevertheless, considerable improvement is obtained relative to a control run without assimilation. Differences between tropics and extratropics are smaller than for the basic observed quantities. Results also show that precipitation and evaporation are affected by a sort of continuous spin-up which is introduced by the relaxation: the bias (ECMWF-ECHAM) is increasing with increasing relaxation forcing. In agreement with this result we found that with increasing relaxation forcing the vertical exchange of tracers by turbulent boundary layer mixing and, in a lesser extent, by convection, is reduced.

## 1. Introduction

The four-dimensional data assimilation (FDDA) based on “nudging” or Newtonian relaxation is a method which dynamically couples various observed variables using a general circulation model (GCM). Terms are added to the prognostic equations to force the model toward the observed state. FDDA was originally developed to insert synoptic data like satellite or aircraft

measurements in numerical weather prediction (NWP) systems. It is widely used by both the research and operational meteorology communities for application in data assimilation systems in weather forecast models [e.g., *Hoke and Anthes, 1976; Davis and Turner, 1977; Lyne et al., 1982; Ramamurthy and Carr, 1987; Krishnamurti et al., 1991*]. In addition, FDDA is also a powerful tool to analyse the behavior of the atmosphere and to derive variables which cannot be observed.

A new application of Newtonian relaxation could be the validation of physical parameterization schemes of a GCM, whereby the large-scale component of the modeled circulation is adjusted toward meteorological analyses. Climate studies are performed using general circulation models which are integrated over long periods, typically 30 years, to obtain statistically significant averages. Such models are evaluated by comparing the model’s multiyear averages and interannual variability

<sup>1</sup>Deceased January 1994.

with synoptic observations. But for many meteorological parameters calculated by the model, like cloud cover or cloud liquid water, no long-term observations are available. As the cloud feedback is a key issue for climate dynamics [see e.g., *Cess et al.*, 1990] there is a real need to carefully evaluate the parameters of the hydrological cycle. Moreover, due to an increasing interest in recent years on the possible impact of enhanced emissions of radiatively active gases and aerosol precursors on climate, chemistry modules have been implemented in GCMs to study feedback mechanisms between climate and atmospheric chemistry. As the database of chemical species is much smaller than for meteorological parameters, validation of the model's chemistry is only possible for short periods like during measurement campaigns where sufficient observations are available. To be able to use data covering shorter periods for model validation, one could apply the relaxation to force a GCM to simulate a specific episode.

The main objective of this study is to develop a practical and simple method which can be applied in validating climate models. Use is made of the above mentioned Newtonian relaxation technique. We have investigated which model variables should be adjusted toward the analyses and what timescale of the forcing should be applied for the different variables.

In section 2 the model used for our study is described. In section 3, a theoretical overview is given of the various aspects of the "nudging" method which are important for our application, and in chapter 4 the currently used implementation of the method and the setup of the experiments are described. In section 5 the results of the experiments are presented.

## 2. Model Description

The GCM used in the present study, the Hamburg climate model ECHAM, is a general circulation model based on the numerical weather prediction model of the European Center for Medium-Range Weather Forecasts (ECMWF) with a new physical package adapted for climate studies [*Roeckner et al.*, 1992]. The prognostic variables of ECHAM are the vorticity, the divergence, the temperature, the logarithm of surface pressure, water vapour, cloud water, and several chemical species. For its representation of the dynamic fields, the model adopts expansions in terms of spherical harmonics with triangular truncation at a specified wavenumber. Non-linear and physical terms are calculated on a Gaussian grid. Positive definite quantities, like water vapour, cloud liquid water and chemical species, are advected by a semi-Lagrangian scheme [*Rasch and Williamson*, 1990]. The model can be run with the resolutions T21 ( 500 km), T42 ( 250 km), and T106 ( 100 km). Discretization in the vertical is done by using a hybrid co-ordinate system with 19 levels with its uppermost level at about 30 km. The temporal finite difference scheme is semi-implicit with a time-step of 40 min. (T21), 24

min. (T42), or 12 min. (T106). The radiative transfer model is based on a two-stream method by *Fouquart and Bonnel* [1980] for the solar part and by *Morcrette* [1991] for the terrestrial part. The frequency spectrum is separated into two short-wave and six long-wave intervals. The stratiform cloud scheme is based on the approach of *Sundquist* [1978] who developed a method for including subgrid-scale condensation and cloud formation within the framework of a cloud liquid water transport equation. Convective clouds are parameterized by the mass flux scheme of *Tiedtke* [1989]. Turbulent exchange in the boundary layer is parameterized by using the Prandtl-Kolmogorov formulation of the eddy diffusivity as a function of the turbulent kinetic energy (TKE-closure) [*Brinkop and Roeckner*, 1995].

## 3. Newtonian Relaxation

### 3.1 Technique

The "nudging" method relaxes the model state toward observational data during the assimilation period by adding a non-physical relaxation term to the model equations:

$$\frac{\partial X}{\partial t} = F_m(X) + G(X_{obs} - X) \quad (1)$$

Here  $X$  represents any prognostic model variable and  $F_m$  is the model forcing, describing the dynamical and physical processes that determine the evolution of  $X$ . The relaxation term  $G(X_{obs} - X)$  is expressed as the product of the relaxation coefficient  $G$  ( $s^{-1}$ ) and the difference between the observational and the model calculated value of  $X$ . As observational data, both raw observations and analyzed observations can be used. When using a spectral model, a simple opportunity is offered to exclude smaller scales from relaxation, by not adjusting the higher wave number spectral components. This can be particularly helpful in high-resolution experiments when the smaller-scale variations are for a large part determined by the physical or parameterized quantities and are not represented by the observations.

Observational data are generally not available for each model time step. To obtain data for every model time step, data have to be interpolated. The simplest way to do this is linear interpolation. As mentioned by *Brill et al.* [1991], such an interpolation is allowed for data intervals up to 6 hours. It should be realized that for higher-resolution experiments, linear interpolation becomes more questionable since, as mentioned before, more smaller temporal scales are resolved. This is another reason to restrict the relaxation to low-order spectral components.

### 3.2. Choice of Relaxation Coefficients

The choice of the right value of  $G$  forms a main difficulty of the "nudging" technique [*Stauffer and Seaman*, 1990]. If  $G$  is chosen too large the relaxation term will

dominate the model forcing. In this case, possible dynamical imbalances in the observations may be spuriously amplified. On the other hand, if  $G$  is chosen too small, the observations will have little effect on the solution. Various considerations concerning the choice of  $G$  can be found in literature. *Hoke and Anthes* [1976] suggest that the optimal value should depend on the observational accuracy, the kind of variable being adjusted, and the typical magnitude of the model forcing. This means that  $G$  may vary both in space and time since observational accuracy as well as the magnitude of model forcing vary in space and time. Nevertheless, in most applications the variation of  $G$  in time and space is not taken into account. In some recent studies the values of  $G$  are calculated for each grid point separately using a parameter estimation procedure which makes use of the adjoint equations of the numerical model [ *Zou et al.*, 1992; *Stauffer and Bao*, 1993].

Different approaches exist with respect to the choice of the variables to be adjusted. One question is whether to adjust the wind field (and let the mass field adjust to it by the model equations) or to adjust the mass field (and let the wind field adjust to it) or both. *Kuo and Guo* [1989] conclude that simultaneous assimilation of both mass and momentum yields the best results. A second question is whether to adjust the rotational part of the windfield alone or both the rotational and divergent part. The rotational part of the wind is considered to be more important to adjust than the divergent part, for several reasons. The divergent part is relatively poorly observed [*Holopainen*, 1987] and is closely associated with gravity waves. The rotational part of the wind, on the other hand, represents the slow manifold of quasi-geostrophic motion [*Daley and Puri*, 1980]. Quasi-geostrophic theory is capable to describe the evolution of Rossby wave type flow for timescales up to a few days. Adjusting the vorticity field of the model is, therefore, an efficient way to adjust the large-scale flow in the extratropics. In many applications the relaxation coefficient for vorticity is therefore larger than the coefficient for divergence.

*Brill et al.* [1990] also adjusted the surface pressure. They argue that adjusting the predicted variables only above the surface can lead to an accumulation of errors at the surface. Adjusting the moisture field in the surface layer to improve the diabatic forcing has been applied by a few investigators [*Douglas and Warner*, 1987; *Krishnamurti et al.*, 1988, 1991; *Brill et al.*, 1990].

Derivation of mathematically optimal relaxation coefficients would require a method similar to the variational approach as used by *Zou et al.*, [1992]. Variational methods require, however, huge computer memory and power, which make them impractical to use in a large global model.

### 3.3. Influence on Model Physics

As an example to illustrate the influence of relaxation on the parameterized part of the model, we examine the

thermodynamic energy equation including a relaxation term:

$$\frac{dT}{dt} = \frac{Q}{c_p} + \omega \frac{\kappa T}{p} + G(T_{obs} - T), \quad (2)$$

where the symbols have their conventional meteorological meaning. The symbol  $Q$  denotes the diabatic heating, which mainly consists of the release of latent heat due to the condensation of water vapour, the flux of sensible heat from the Earth's surface into the atmosphere, and the radiative heating. In the model these heating components are parameterized in terms of the large-scale dynamical quantities, like wind and temperature. Therefore, the relaxation term not only modifies the temperature, but also indirectly modifies the diabatic heating in the model. Without relaxation, a physically consistent balance between the parameterized and the dynamical quantities exists in the model. The relaxation in (2), therefore, not only forces the modelled temperature toward the observations, but also hinders the parameterized quantities to reach a balanced state. Similar arguments apply to the equations for momentum and moisture.

This problem is analogous to the spin-up problem in weather prediction models, in which assimilation of observations causes an initial imbalance in the model physics. This imbalance disappears a short time after the assimilation. However, in a setup in which the model is relaxed toward new observational data every time step, there will be a continuous spin-up.

## 4. Description of Method and Experiments

### 4.1. Implementation in ECHAM

A model integration step (1) is solved using a semi-implicit finite difference scheme and is carried out in two steps. First, the model forcing  $F_m$  is calculated yielding a new value for prognostic variable  $X_{t+\Delta t}^m$ . Next the relaxation term is calculated:

$$\frac{X_{t+\Delta t}^m - X_{t+\Delta t}^{an}}{2\Delta t} = G(X)[X_{t+\Delta t}^{an} - X_{t+\Delta t}] \quad (3)$$

in which  $X_{t+\Delta t}^{an}$  are the meteorological analyses of prognostic variable  $X$  to which the model is relaxed.  $G(X)$  is the relaxation coefficient which in the present study depends only on the variable which is relaxed.  $X_{t+\Delta t}$  finally is the value of  $X$  after the assimilation step and can be easily calculated from the above equation.

Since ECHAM is a spectral model assimilation is done in the spectral space. The resolution used in all of the experiments described in this paper is T21. For this coarse T21 model version, all spectral components are equally adjusted. However, one exception has been made: to conserve mass and momentum, the first spectral component, which represents the global average, is not adjusted.

## 4.2. ECMWF Data

In our study we have used ECMWF data for two purposes: the assimilation itself and for comparison afterwards with ECHAM model results.

**Data for Assimilation.** As input data for the assimilation we have used spectral ECMWF first guess fields, which is the 6-hour forecast, of vorticity, divergence, temperature and surface pressure. We have simulated 2 months, January and July 1989, by relaxing the ECHAM model toward the ECMWF analyses for these periods. The model runs have been started from initial files from a January and a July month. The input fields for the relaxation are only available every 6 hours. To obtain data for each model time step, we have linearly interpolated the data in time. We found that cubic spline interpolation did not show any improvement for our application.

Before applying the relaxation, the ECMWF data, which originally are in T106 resolution, are truncated to T21. In 1989 the ECMWF model used the same vertical hybrid coordinate parameters as the ECHAM model uses. The orographies used in the two models, however, are different. ECHAM uses a mean orography, whereas the ECMWF model uses an envelope orography in which the subgrid-scale variance is added to the mean orography of a grid box [Simmons, 1987]. Differences between the two may amount to a few hundreds of meters. Since close to the surface the hybrid levels are terrain-following, this means that when data are used over mountainous areas, the levels on which they are valid are different for the ECHAM and the ECMWF model. To avoid errors, caused by these non-matching orographies, we vertically interpolate.

The following procedure has been applied (M. Hortal, personal communication, 1995): the input spectral ECMWF data are truncated to the desired output resolution and transformed to grid point space, where instead of vorticity and divergence, the wind components are calculated. The input (envelope) orography (grid point data) is interpolated to the desired output resolution, by transforming it to spectral space, next truncating it to the new resolution and then transforming it back to grid point space. Next the desired output (mean) orography is constructed from the high-resolution original U.S. Navy data set. In the new orography we compute the logarithm of the surface pressure from the interpolated data. From this we compute the pressures at the model levels corresponding to the output resolution. Then we interpolate (or extrapolate) the wind components and the temperature linearly in pressure. Finally, we transform the wind components to vorticity and divergence in the spectral space.

Owing to this vertical interpolation, one can expect inaccuracies in areas where the difference between the surface pressure in the envelope and mean orography is large. One should realize that the ECMWF data used

for assimilation can be seen as a good approximation for the observed atmospheric state in data-rich areas. In large parts of the tropics, however, meteorological observations are sparse and the ECMWF data for these areas rather reflect the model prediction than the observations and, therefore, may not be very accurate.

**Data for Comparison.** The following ECHAM model variables have been compared with the corresponding ECMWF fields (see also Table 2): surface pressure, the temperature at 850 hPa, the zonal wind at 200 hPa, and the 500 hPa geopotential height. Further, we compare the specific humidity at 700 hPa. Although this is a prognostic variable and therefore transported by the model resolved large-scale flow, it is mainly determined by parameterized subgrid-scale processes. Therefore we cannot expect perfect agreement between the ECHAM humidity field and the ECMWF data.

When using ECMWF model fields of precipitation and evaporation, one should realize that there is a spin-up in the short-range forecast of these fields. The duration of the spin-up depends on the model and on the area for which the variables are calculated [Krishnamurti *et al.*, 1988, Arpe, 1991]. Arpe [1991] argues that ECMWF forecasts of daily mean precipitation and evaporation for day 0.5-1.5 provide the best compromise between spin-up and forecast error. He further concludes that for the northern hemispheric extratropics the short range forecasts of evaporation over the oceans and precipitation are probably accurate estimates of the truth, not only for monthly means but also for the day to day variability. In the tropics and southern hemisphere, the distribution seems realistic but there is a strong spin-up in the data. Therefore we have used the day 0.5-1.5 forecast period when comparing time series at individual grid points. For the comparison of global averages, however, fields should be free of bias and for this purpose we compare the ECHAM fields with a range of ECMWF forecasts up to 3 days.

The grid point values of precipitation and 700 hPa specific humidity and evaporation have been smoothed in space (3x3 grid points) and 700 hPa specific humidity has been also smoothed in time (running average over three intervals of 6 hours each). This filters out much of the quasi-random small-scale variations that otherwise might prevent a useful comparison of the two fields obtained from different parameterizations.

## 4.3. Experiments

Like others [e.g., Krishnamurti *et al.*, 1988] we decide by "trial and error" which relaxation coefficients give the best results for our application. A series of assimilation experiments (shown in Table 1) has been performed, in which the relaxation coefficients are varied according to physical arguments as described in section 3.2 and within a range observed in literature. Moisture has not been included in our experiments because

**Table 1.** Specification of the Assimilation Experiments

Exp.	$G_T$	$G_D$	$G_{V_o}$	$G_{Insp}$
0	0.0	0.0	0.0	0.0
1	$1.10^{-4}$	$0.5.10^{-4}$	$1.10^{-4}$	$1.10^{-4}$
2	0.0	$0.5.10^{-4}$	$1.10^{-4}$	$1.10^{-4}$
3	$1.10^{-4}$	0.0	$1.10^{-4}$	$1.10^{-4}$
4	0.0	$0.5.10^{-4}$	$1.10^{-4}$	0.0
5	$1.10^{-5}$	$0.5.10^{-4}$	$1.10^{-4}$	$1.10^{-4}$
6	$1.10^{-5}$	$0.5.10^{-5}$	$1.10^{-5}$	$1.10^{-5}$
7	$1.10^{-3}$	$0.5.10^{-3}$	$1.10^{-3}$	$1.10^{-3}$

$G_D$ ,  $G_{V_o}$ ,  $G_T$  and  $G_{Insp}$  are the relaxation coefficients ( $s^{-1}$ ) for divergence, vorticity, temperature and surface pressure respectively.

the moisture field is determined by diabatic processes rather than dynamical processes. An assimilation technique based on dynamical adjustment would therefore not be very beneficial. As outlined in section 3.2, the rotational part of the wind is considered most important to adjust. For this reason vorticity has been included in each of the assimilation experiments.

In judging which set of relaxation coefficients would be most useful, we look upon two different aspects: Validating model parameterizations or chemistry modules with measurements requires in first instance that the basic flow, temperature, and moisture pattern from which most subgrid-scale parameterizations are derived and which determine transport and reaction rates of chemical tracers, is represented well by the relaxed model.

A second requirement is that the model physics work without too much restrictions imposed by the relaxation. It is desired that the model physics adapt to the, by the relaxation imposed, flow and temperature distribution without being too much suppressed by the relaxation. This aspect is more complicated to investigate. Three different approaches are used: First, the precipitation (convective and large-scale) and evaporation are compared with the corresponding ECMWF fields. Since the ECHAM and ECMWF model use different parameterization schemes, again we may not expect good quantitative agreement. A comparison can, however, be useful to see if the relaxed model produces realistic fields from parameterized subgrid-scale processes. A second approach is to directly compare the forcing of the temperature by parameterized processes with the forcing by relaxation. The distribution of the forcing in time and space is also considered. In this way some qualitative understanding can be gained in where and when the forcing by relaxation could dominate the normal model forcing. A third, more quantitative approach, is the comparison of the vertical exchange of tracers in an undisturbed experiment with a series of relaxation experiments in which the relaxation forcing is systematically increased. The bulk of the vertical exchange of longer-lived tracers from the surface to the free tro-

posphere and from the lower to the upper troposphere is in the model controlled by parameterized processes. The systematic change in the amount of vertically exchanged tracer mass in this approach therefore can be used as a measure for the influence of the relaxation on the model physics.

Comparison between ECHAM produced fields and ECMWF fields has been done using various statistical measures like the anomaly pattern correlation, the tendency correlation, the bias and the root mean squared difference/error (RMSE). The anomaly pattern correlation is calculated as the spatial correlation between the ECMWF and ECHAM anomaly fields. The anomaly is calculated with respect to an ECHAM monthly mean field from a climatological run. The tendency correlation is calculated as the correlation between ECHAM and ECMWF timeseries at individual grid points and gives us some insight in how well ECHAM adjusts in time to the ECMWF data. Also, the RMSE is calculated for time series of individual grid points and provides insight in the magnitude of differences between the two models.

## 5. Results

In this section we will address three subjects. First, the agreement between the model simulations and the ECMWF data set for the basic observed quantities (the first five variables in Table 2) will be quantified. Second, some ECHAM-derived, parameterized quantities (precipitation and evaporation) will be compared with ECMWF forecasts. Finally, we will investigate the impact of the assimilation on the model's parameterized quantities.

### 5.1. Agreement with Basic Observed Quantities

As a first indication of what agreement can be obtained, in Table 3 the global averaged statistics are presented for the comparison of ECHAM output fields (assimilation experiment 5 from Table 1) with corresponding ECMWF fields for January 1989. From these global mean results it appears that ECHAM, when relaxed toward ECMWF analyses, is well capable to reproduce these analyses. As can be seen in Table 3, the anomaly pattern correlations are extremely high. This means that the spatial patterns of the ECMWF fields are almost exactly reproduced by the ECHAM simulation. The agreement for the specific humidity is somewhat less; the magnitude of the RMSE is about 25 % of the value of the globally averaged specific humidity. Since, as explained earlier, the specific humidity has not been adjusted in the relaxation experiments and for a large part is determined by the model parameterized moist processes which are different between the two models, this lesser agreement is not surprising.

We should, however, consider the spatial variability of the statistical parameters as well. Figure 1 gives information about the latitudinal dependence of the ten-

**Table 2.** ECMWF Data Used for Comparison

Variable	Type	Period	Interval	Averaging	Interp.
200 hPa zonal wind	FG	6h	6h	no	no
500 hPa geopotential height	FG	6h	6h	no	no
700 hPa specific humidity	FG	6h	6h	3x3 gridcells 3 time steps	no
850 hPa temperature	FG	6h	6h	no	yes
Surface pressure	FG	6h	6h	no	yes
Precipitation	FC	12h-36h	12h	3x3 gridcells	no
Evaporation	FC	12h-36h	12h	3x3 gridcells	no

Type FG (first guess) or FC (forecast), period of forecast, output interval, averaging in time and/or space and vertical interpolation to mean orography have been indicated.

dency correlation and the RMSE. It can be clearly seen that the relaxation technique yields less satisfying results in the tropics than in the extratropics. Tendency correlations generally are much lower in the tropics, particularly for the 850 hPa temperature and surface pressure. The RMSE is not higher in the tropics. In fact for all variables, except the zonal wind at 200 hPa, it is lower in the tropics. Relative to the low temporal variability of the presented variables, the error is larger in the tropics.

Figure 2 shows the dependence of tendency correlation and the RMSE on the choice of relaxation coefficient. The tropics (between 22.5°N and 22.5°S) and extratropics are analyzed separately. Experiment 0 is the control run in which no relaxation has been applied. It can be seen in Figure 2 that the results are not very sensitive to the choice of the relaxation coefficients for the range of values applied. The variability among the experiments is low except for the 850 hPa temperature.

Most of this variability for the 850 hPa temperature can be explained by whether or not temperature has been relaxed. In experiment 2 and 4 the temperature has not been included and the relaxation coefficient for temperature has a low value in experiment 5 and 6. These experiments differ significantly from experiments 1, 3, and 7 in which temperature has been relaxed with the same coefficient as vorticity.

For the 500 hPa geopotential height the values of the RMSE and the tendency correlation have about the same value for almost all experiments, except for experiment 6. This experiment has low values for the relaxation coefficients. The only important variable for a realistic simulation of 500 hPa geopotential height seems to be the vorticity. Excluding other variables from assimilation does not influence the results. The same behavior with respect to assimilation experiment is observed for the 200 hPa zonal wind (not shown), with the only exception that the correlations are slightly lower.

Almost the same can be concluded for the surface pressure. Not adjusting the surface pressure (experiment 4) worsens the statistics only slightly. A negative effect of the same order of magnitude can be detected for experiment 6.

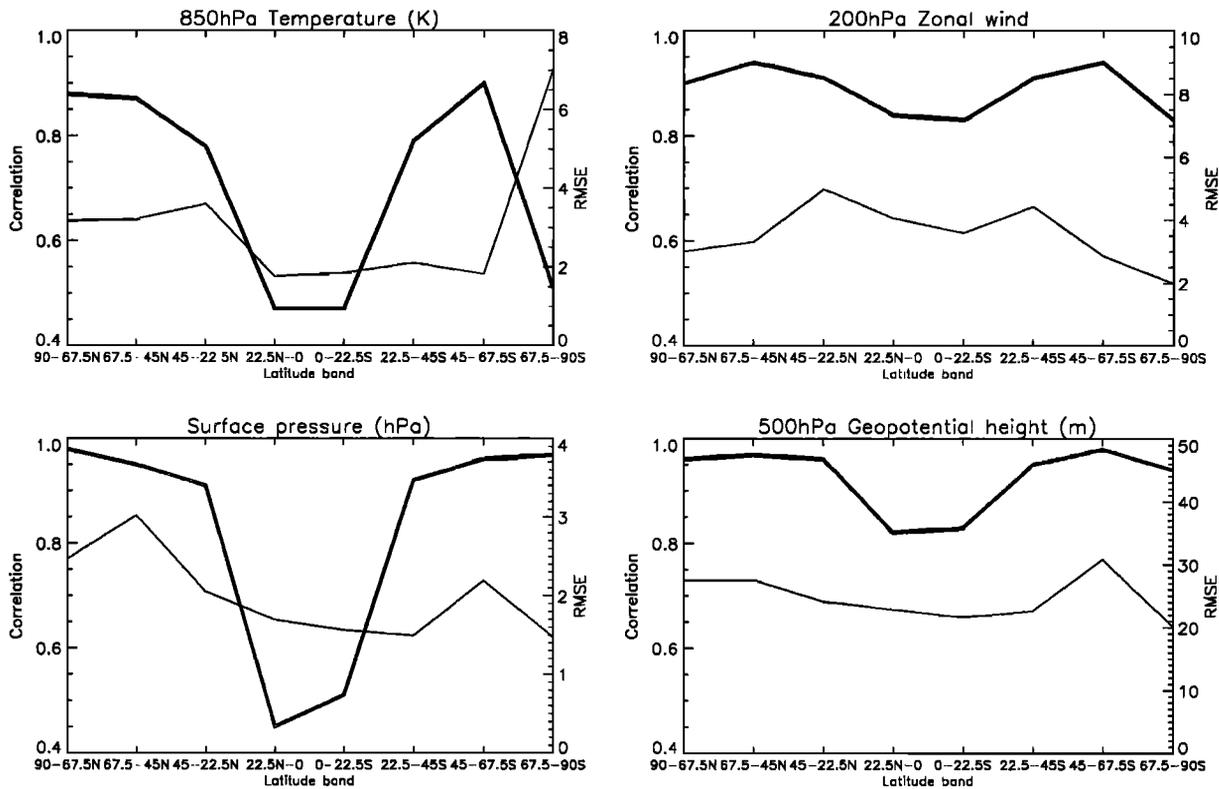
The agreement of the ECHAM specific humidity field at 700 hPa with ECMWF data is considerably less than for the other variables shown in Figure 2. Adjusting the divergence appears to be important for the humidity field, because not including the divergence in the assimilation (experiment 3) results in a significant increase in RMSE and a decrease of correlation in the tropics.

As was already shown by Figure 1, Figure 2 again shows that the relaxation technique yields the best re-

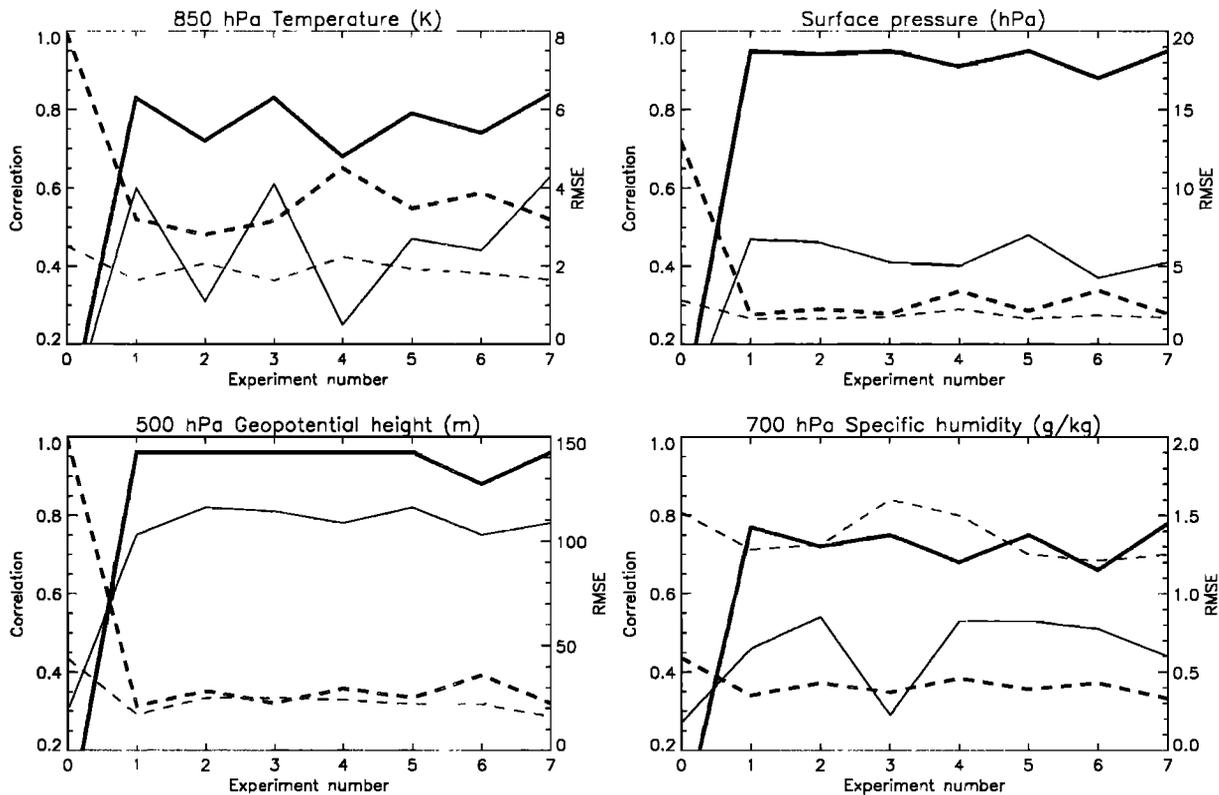
**Table 3.** Statistics for Experiment 5, Global Averages for January 1989

Variable	Mean	Bias	RMSE	TC	APC
850 hPa temperature (K)	279.183	-0.260	1.203	0.801	0.975
200 hPa zonal wind (m/s)	16.053	-0.004	3.862	0.883	0.975
500 hPa geopotential height (m)	5549.37	18.56	24.18	0.909	0.985
700 hPa specific humidity (g/kg)	3.160	0.238	0.794	0.657	0.680
Surface pressure (hPa)	985.565	0.583	1.911	0.759	0.974

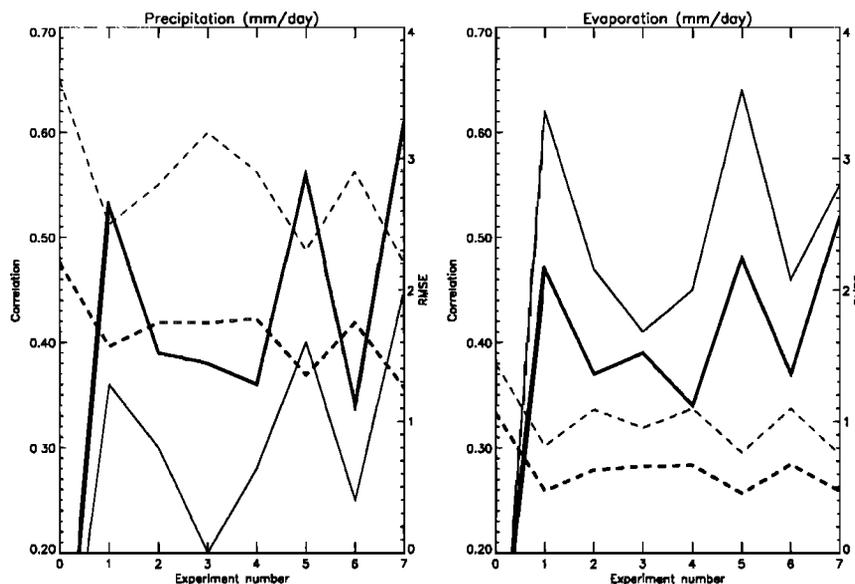
The Bias is calculated as ECHAM-ECMWF and mean values are taken from ECHAM. TC stands for tendency correlation and APC for anomaly pattern correlation



**Figure 1.** Comparison of assimilation experiment 5 and European Center for Medium-Range Weather Forecast (ECMWF) first guess fields: Zonal mean tendency correlation (thick line) and root mean squared error (RMSE) (thin line) for January 1989.



**Figure 2.** Correlation (solid line) and RMSE (dashed line) between ECMWF first guess fields and the assimilation experiments outlined in Table 1 for January 1989. Thick lines represent averages over the extratropics and thin lines over the tropics.



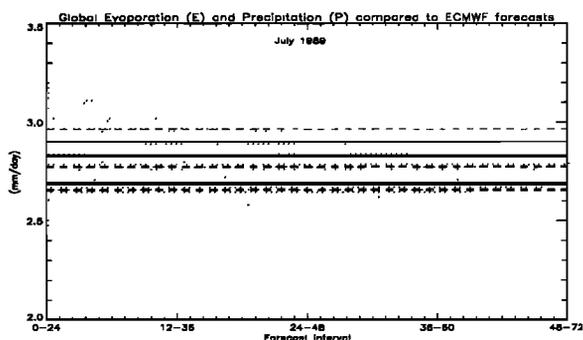
**Figure 3.** Correlation (solid line) and RMSE (dashed line) between ECMWF 12-36h forecast fields and the assimilation experiments outlined in Table 1 for July 1989. Thick lines represent averages over the extratropics and thin lines over the tropics.

sults in the extratropics. It also shows that there is hardly any improvement in the RMSE to be observed compared to the control run in the tropics.

For July 1989 (not shown) the results for the basic observed quantities are similar. The moisture field, however, has improved considerably in the tropics. RMSE's have been reduced to values below 1.0 g/kg and the tendency correlation now is similar to the correlation for the extratropics. The change to the Tiedtke scheme [Tiedtke, 1989] as convection scheme of the ECMWF model in May 1989 is probably the reason for this improvement since ECHAM also uses the Tiedtke scheme.

## 5.2. Comparison of Derived Quantities

In a similar manner as has been done in previous section, the ECHAM precipitation and evaporation from



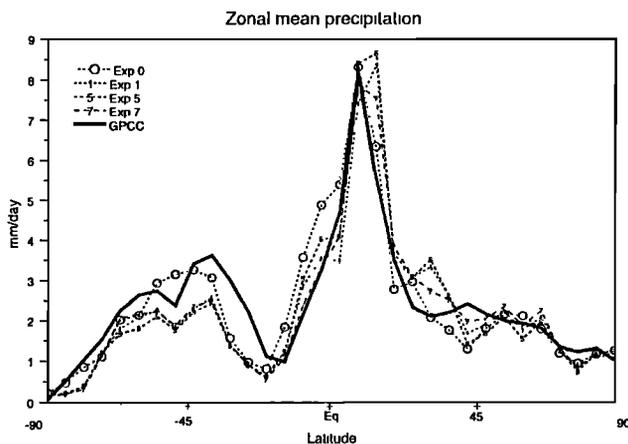
**Figure 4.** ECMWF global mean precipitation (upper boundary of shaded area) and evaporation (lower boundary of shaded area) as a function of forecast interval. The horizontal lines represent in order of increasing thickness the precipitation (solid) and evaporation (dashed) for assimilation experiments 0, 5, and 7.

the different relaxation experiments are compared with the appropriate ECMWF forecasts (see section 4). In Figure 3 the dependence of the correlation and RMSE on the choice of relaxation coefficients is shown for evaporation and precipitation. The improvements obtained by assimilation in comparison with the reference experiment are smaller than the improvements seen in section 5.1 for the basic observed quantities. The maximum improvements obtained in RMSE are 35 and 50 % for precipitation and evaporation, respectively. Correlations are considerably lower. Differences in statistics between the tropics and extratropics are smaller than diagnosed for the observed quantities. Some important differences among the experiments can be seen. Not adjusting the temperature (experiment 2 and 4) worsens the statistics for both evaporation and precipitation. The difference between the RMSE of experiment 2 and 5, for example, is about 0.4 mm/d and 0.2 mm/d for precipitation and evaporation, respectively. Not adjusting the divergent part of the windfield (experiment 3) has a negative effect on the statistics for both precipitation and evaporation. This is shown most clearly for the precipitation in the tropics. The strong relation between large-scale convergence and convection used in the convection parameterization is probably the reason for this effect. Provided that the ECMWF model output is realistic, we learn from this comparison two things: First, the influence of relaxation on the derived variables in the tropics is beneficial. Second, temperature and divergence are important variables to assimilate in order to improve the precipitation and evaporation field.

Figure 4 shows a comparison between the total precipitation and evaporation fields from assimilation experiments 0, 5, and 7 and several ECMWF forecast in-

tervals up to 3 days. It can be seen that a forecast period of more than 2 days is needed to reduce the spin-up to an acceptable level. Further, Figure 4 shows that the global hydrological balance for the relaxation experiments is well simulated. The difference between evaporation and precipitation is less than 0.1 mm/d. The global mean evaporation and precipitation simulated by experiment 5 and 7 also compare both well with the ECMWF 2-3 day forecast. Whereas the bias (ECHAM-ECMWF) for the basic observed quantities (not shown) does not vary much with the relaxation coefficients chosen, there seems to be a systematic change in both the bias for precipitation and evaporation. It can be seen that by increasing the strength of the relaxation coefficients (from experiment 0 via 5 to 7) the globally averaged values of precipitation and evaporation decrease with about 0.25 mm/d, which is about 10 % of the global mean values.

Figure 5 shows the latitudinal dependence of ECHAM precipitation from relaxation experiments 5, 1, and 7 compared to the control run and to precipitation data from the Global Precipitation Climatology Center (GPCC). These data are a mix of land-based observations and microwave, radar, and satellite measurements over the sea and where no data are available, the ECMWF short-range forecast is used [Rudolf, 1995]. In general, the ECHAM produced precipitation patterns look realistic. Especially in the northern hemispheric extratropics the differences with the observational GPCC data are small. Large differences of more than 1 mm/d can be observed in the southern hemispheric subtropics, near the tropical maximum in the northern hemisphere and near 35° N where the ECHAM data show another precipitation maximum and the GPCC data do not. Since the satellite measurements are not available south of 30° S and since there are not many land-based observations in the southern



**Figure 5.** Zonal and monthly mean precipitation (millimeters per day) from assimilation experiments 0, 1, 5, and 7 compared to the Global Precipitation Climatology Center (GPCC) precipitation data [Rudolf, 1995] for July 1989.

hemisphere, the GPCC curve south of 30° S is mainly determined by the ECMWF short-range forecast. It can be seen that the ECHAM relaxation experiments follow the same pattern for this area but that there is a clear offset. This might be due to spin-up in the ECMWF shortrange forecast which is shown by Figure 4 for the global mean field. It appears that in the tropics and subtropics, the precipitation is reduced with increasing relaxation strength but in the extratropics, especially in the northern hemisphere, the opposite is true. However, due to the larger area of the tropical gridcells, which is taken into account when calculating global averages, the global value is reduced, which is shown in figure 4. This aspect will be further discussed in the following section.

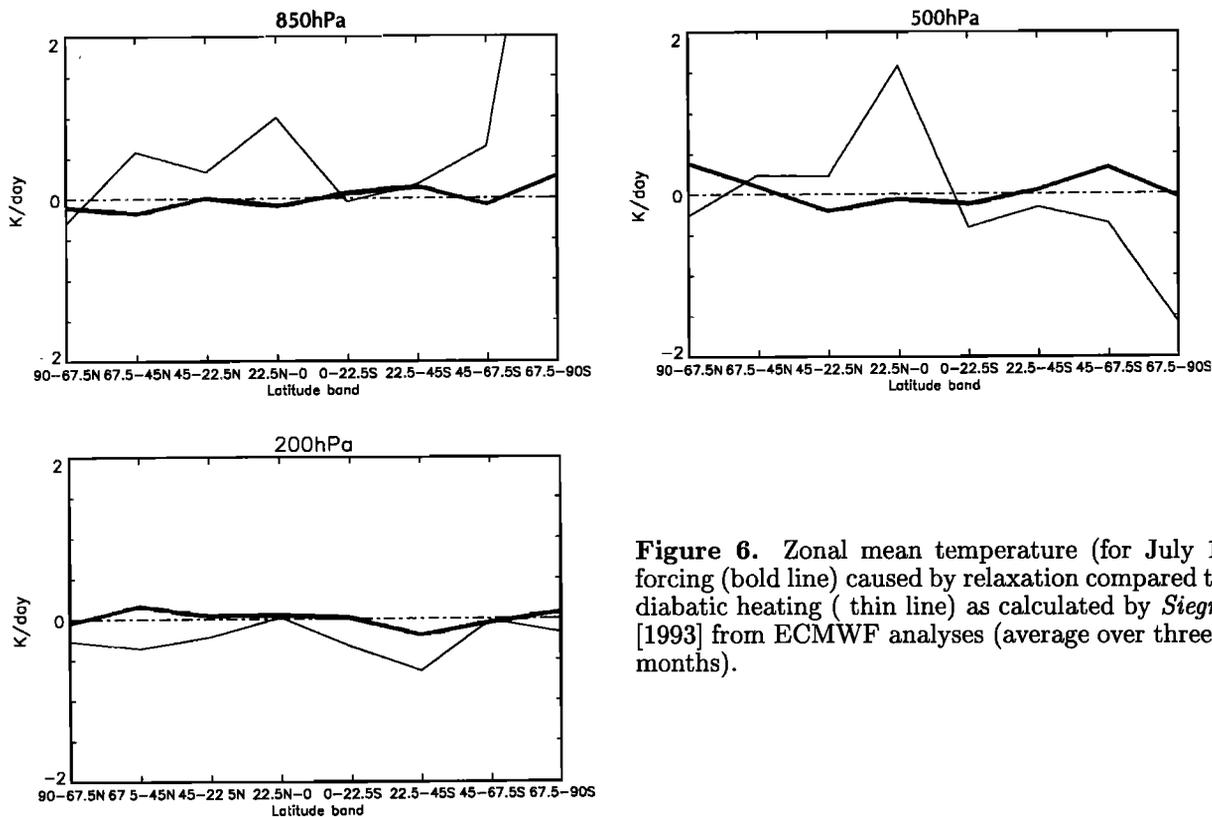
### 5.3. Influence on the Model Physics

As outlined in section 3.3, the addition of an extra term to the model equations can influence the performance of the physical parameterizations. We have shown in section 5.2 that when using strong relaxation coefficients, the global averaged precipitation and evaporation is slightly reduced. The most serious effects of relaxation can be expected when the relaxation term is large compared to the other terms in the equations. Figure 6 shows the magnitude of the relaxation term for temperature and its geographical distribution compared to the magnitude of the diabatic heating term, that is the third right-hand side term of (2) compared to the first right-hand side term. The diabatic heating rates are calculated from ECMWF initialized analyses by Siegmund [1993]. The values for the relaxation term are obtained from relaxation experiment 5.

The absolute values of the diabatic heating are generally larger than the relaxation term. The difference is largest at the 850 and 500 hPa levels, where the heating is an order of magnitude larger than the relaxation term. The relaxation term hardly depends on latitude. The diabatic heating, on the other hand, is largest in the intertropical convergence zone and decreases toward the poles. The peak over Antarctica is probably due to errors in the ECMWF model [see Siegmund, 1993].

To gain a more quantitative insight in how much the model physics are affected by the relaxation process, we also considered the dependence of the vertical motions on the strength of relaxation. For this purpose we released a passive tracer at the model's surface with a constant source strength and a lifetime of 10 days. Since the source is distributed uniformly over the entire globe, changes in the vertical distribution of the tracer are mainly due to vertical transport. Figure 7 shows the ratios between the tracer concentrations at 250 hPa level and the surface, and the 750 hPa level and the surface, respectively. The values are averages over 15 day-simulations for July 1989 (A-F on the horizontal axis are six assimilation experiments described in Table 4).

The ratio of the tracer concentrations between 750 hPa and the surface is a measure of the vertical ex-

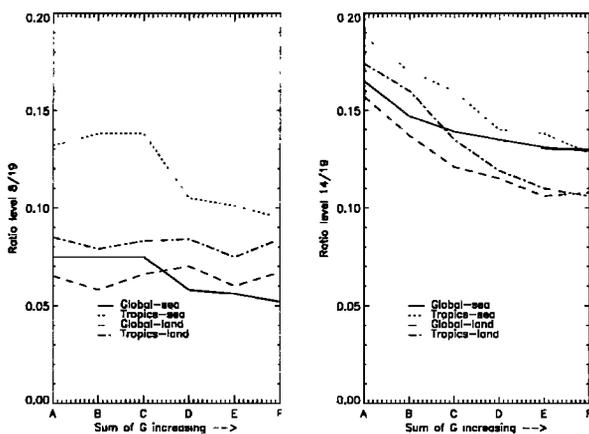


**Figure 6.** Zonal mean temperature (for July 1989) forcing (bold line) caused by relaxation compared to the diabatic heating ( thin line) as calculated by *Siegmund* [1993] from ECMWF analyses (average over three July months).

change of tracer by boundary layer processes, and the ratio of tracer concentrations between 250 hPa and the surface is a measure for the vertical exchange through convective processes. Figure 7 shows that the convective exchange of tracers over sea is quite sensitive to the relaxation coefficient for temperature since there is a large drop in ratio when going from experiment C to D. The only difference between experiment C and D is a factor 10 in the relaxation coefficient for temperature.

The convective exchange over land is hardly influenced by the strength of the relaxation. Apart from the magnitude of the ratio, the behavior is similar for the tropics and the extratropics.

Figure 6 also shows that the boundary layer exchange decreases when increasing the strength of relaxation. This decrease is largest for the tropics over land where it is almost 40 % between experiment A and F. A difficulty with interpreting Figure 7 is that we do not know the real vertical exchange. It is, however, unlikely that there is such a large difference between the vertical exchange in ECHAM and the ECMWF model since both models apply similar boundary layer diffusion and convection schemes. We therefore think that a main reason of the observed systematic decrease of vertical tracer



**Figure 7.** (left) The ratio of tracer concentrations between model level 8 ( 200 hPa) and the surface and (right) model level 14 ( 750 hPa) and the surface after a 15 day-run for July 1989. The ratios are calculated for sea and land and tropics and extratropics, separately.

**Table 4.** Specification of the relaxation coefficients used for the experiments plotted in Figure 7

Exp.	$G_T$	$G_D$	$G_{V_o}$	$G_{l_{ns}p}$
A(=0)	0.0	0.0	0.0	0.0
B	$1.10^{-6}$	$0.5.10^{-5}$	$1.10^{-5}$	$1.10^{-5}$
C(=5)	$1.10^{-5}$	$0.5.10^{-4}$	$1.10^{-4}$	$1.10^{-4}$
D(=1)	$1.10^{-4}$	$0.5.10^{-4}$	$1.10^{-4}$	$1.10^{-4}$
E	$1.10^{-4}$	$0.5.10^{-3}$	$1.10^{-3}$	$1.10^{-3}$
F(=7)	$1.10^{-3}$	$0.5.10^{-3}$	$1.10^{-3}$	$1.10^{-3}$

$G_D, G_{V_o}, G_T$  and  $G_{l_{ns}p}$  are the relaxation coefficients ( $s^{-1}$ ) for divergence, vorticity, temperature and surface pressure respectively. The numbers between the brackets refer to Table 1

exchange is the spin-up effect which becomes stronger with increasing relaxation strength. This effect apparently has a stronger influence on the boundary layer exchange than on the convective exchange. Since we want to maintain the intensity of parameterized processes in the ECHAM model, it implies that we should use as weak relaxation coefficients as possible.

## 6. Discussion and Conclusions

We have shown that in the extratropics, good agreement between ECHAM fields from model simulations with relaxation and ECMWF analyses is obtained for the basic observed quantities. For example, tendency correlations between ECMWF analyses and ECHAM simulation are higher than 90 %. The agreement in the tropics is poor. This is a major drawback of the proposed method. As model results in the tropics are strongly determined by the model physics we have to expect differences due to different parameterizations. Additionally, the relaxation method which is based on geostrophic adjustment of dynamical processes, will have little effect in tropical regions.

Since ECHAM and ECMWF have different physical parameterizations, the agreement of the derived variables is not as good as for the observed ones. Moreover, the ECMWF model output is, due to the long spin-up time, quite sensitive to the choice of the forecast interval. We found that only after 2.5 days the spin-up error in the ECMWF precipitation forecast is reduced to an acceptable level. The bias of the precipitation and evaporation between ECHAM and ECMWF as simulated in the assimilation experiments is within the range of the ECMWF spin-up error. However, validation of derived variables can only be done properly by comparison with measurements. This has been done by comparing with the GPCP data set [Rudolf, 1995], which is one of the most reliable precipitation data sets. The zonal averaged precipitation pattern is well simulated by the ECHAM relaxation experiments, although locally some large differences exist.

The problem of selecting satisfying relaxation coefficients for assimilation is a key issue. We have shown that too strong relaxation coefficients reduce the global averaged precipitation and evaporation and the vertical exchange of tracers. This implies that the model physics should not be too much dominated by the relaxation process. On the other hand, too weak relaxation, as in experiment 6, will worsen the agreement with the observational data. We consider the relaxation coefficients as used in experiment 5 to be a good compromise. In this experiment we used a value of  $1.10^{-5} s^{-1}$  for temperature,  $0.5.10^{-4} s^{-1}$  for divergence and  $1.10^{-4} s^{-1}$  for vorticity and surface pressure. The model performs better if the relaxation coefficient for temperature is relatively small, since the precipitation and the vertical exchange in convective clouds appear to be sensitive to the magnitude of this coefficient. It has been shown in

Figure 6 that with a temperature relaxation coefficient of  $1.10^{-5} s^{-1}$ , the relaxation forcing does not dominate the diabatic heating term.

Inconsistencies between the ECMWF first-guess data and the ECHAM model may have produced some artificial forcings. Close to mountainous terrain the flow will be different for an envelope orography model than for one using mean orography. These differences are not entirely canceled out by simple vertical interpolation. The ECMWF data used for the assimilation have been generated by T106 model runs. Although truncated to T21, these data will contain more information about smaller scales than ECHAM which has been run at T21 resolution. Therefore by using these data for assimilation, ECHAM is forced to adjust to scales which it in principle cannot reproduce. Better agreement can be expected if the resolution of the model and the analyses are the same. Results of higher-resolution experiments will be discussed in forthcoming papers.

Despite the fact that the proposed method has some major drawbacks, we think that it is promising comparing specific episodes during measurement campaigns with climate models. The method is very simple and practical to apply. Only little additional computer time is required. Additional research should be done to improve results for the tropics. In its current setup the method is appropriate to validate ECHAM data with measurement in the northern hemispheric extratropics.

**Acknowledgments.** The authors thank Mariano Hortal, Klaus Arpe, and Ullrich Schlese for their scientific advice and technical support. Comments of two anonymous reviewers helped improve the manuscript and were greatly appreciated.

## References

- Arpe, K., The hydrological cycle in the ECMWF short range forecast, *Dyn. Atmos. Oceans*, 16, 33, 1991.
- Brill, K.F., L.W. Uccellini, J. Manobianco, P.J. Kocin, and J.H. Homan, The use of successive dynamic initialization by nudging to simulate cyclogenesis during GALE IOP, 1, *Meteorol. Atmos. Phys.*, 45, 15, 1991.
- Brinkop, S., and E. Roeckner, Cloud turbulence interactions: Sensitivity of a general circulation model to closure assumptions, *Tellus*, 47A, 197, 1995.
- Cess, R.D., et al., Intercomparison and interpretation of climate feedback processes in seventeen atmospheric general circulation models, *J. Geophys. Res.*, 95D, 16,601, 1990.
- Daley, R., and K. Puri, Four-dimensional data assimilation and the slow manifold, *Mon. Weather Rev.*, 108, 88, 1980.
- Davies, H.C., and R.E. Turner, Updating prediction models by dynamical relaxation: An examination of the technique, *Q. J. R. Meteorol. Soc.*, 103, 225, 1977.
- DKRZ, Modellbetreuungsgruppe, The ECHAM3 atmospheric general circulation model, *Tech. Rep. 6*, Hamburg, Germany, 1992.
- Douglas, S. G., and T.T. Warner, Utilization of VAS satellite data in the initialization of an oceanic cyclogenesis simulation., *Mon. Weather Rev.*, 115, 2996, 1987.

- ECMWF, ECMWF forecast model adiabatic part, *Research Manual 2, 2nd. ed.* Eur. Cent. Meteorol. Weather Forecasting, Reading, England, 1988.
- Fouquart, Y., and B. Bonnel, Computations of solar heating of the Earth's atmosphere: A new parameterization, *Beitr. Phys. Atmos.*, *53*, 35, 1980.
- Hoke, J.E., and A. Anthes, The initialization of numerical models by a dynamic-initialization technique, *Mon. Weather Rev.*, *104*, 1551, 1976.
- Holopainen, E.O., Recent estimates of diabatic forcing on the planetary scale : A review, *paper presented at the ECMWF Workshop on Diabatic Forcing*, Eur. Cent. Meteorol. Weather Forecasting, Reading, England, 1987.
- Krishnamurti, T.N., H.S. Bedi, W. Heckley, and K. Ingles, Reduction of the spin-up time for evaporation and precipitation in a spectral model, *Mon. Weather Rev.*, *116*, 907, 1988.
- Krishnamurti, T.N., Jishan Xue, H.S. Bedi, K. Ingles and D. Oosterhof, Physical initialization for numerical weather prediction over the tropics, *Tellus*, *43AB*, 53, 1991.
- Kuo, Y. H., and Y. R. Guo, Dynamic initialization using observations from a network of profilers and its impact on short-range numerical weather prediction, *Mon. Weather Rev.*, *117*, 1975, 1989.
- Lyne, W. H., R. Swinbank, and T. N. Birch, A data assimilation experiment and the global circulation during FGGE special observing periods, *Q. J. R. Meteorol. Soc.*, *108*, 575, 1982.
- Morcrette, J.-J., Radiation and cloud radiative properties in the ECMWF operational weather forecast model, *J. Geophys. Res.*, *96D*, 9121, 1991.
- Ramamurthy, M.K., and F.H. Carr, Four-Dimensional Data Assimilation in the monsoon region, I: Experiments with wind data, *Mon. Weather Rev.*, *115*, 1679, 1987.
- Rasch, P. J., and D. L. Williamson, Computational aspects of moisture transport in global models of the atmosphere, *Q. J. R. Meteorol. Soc.*, *116*, 1071, 1990.
- Roeckner, et al., Simulations of the present-day climate with the ECHAM model: Impact of model physics and resolution, *Rep. 93*, Max-Planck-Institute for Meteorology, Hamburg, Germany, 1992.
- Rudolf, B., Die Bestimmung der zeitlich-räumlichen Struktur des globalen Niederschlags, *Ber. Deutsch. Wetterdien.*, 196, Offenbach, Germany, 1995.
- Siegmund, P.C., Cloud diabatic forcing of the atmosphere, estimated from simultaneous ECMWF diabatic heating and ISCCP cloud amount observations, *J. Clim.*, *6*, 2419, 1993.
- Simmons, A.J., Orography and the development of the ECMWF forecast model, in *Proceedings of 1986 ECMWF Seminar on Observation, Theory and Modelling of orographic effects, vol. 2*, Eur. Cent. Meteorol. Weather Forecasting, Reading, England, 1987.
- Stauffer, D.R., and J-W. Bao, Optimal determination of FDDA coefficients using the adjoint equations, *Tellus*, *45A*, 358, 1993.
- Stauffer, D.R., and N.L. Seaman, Use of Four-Dimensional Data Assimilation in a limited area mesoscale model part II: Effects of data assimilation within the planetary boundary layer, *Mon. Weather Rev.*, *119*, 734, 1991.
- Sundquist, H., A parameterization scheme for non-convective condensation including prediction of cloud water content, *Q. J. R. Meteorol. Soc.*, *104*, 677, 1978.
- Tiedtke, M., A comprehensive mass flux scheme for cumulus parameterization in large-scale models, *Mon. Weather Rev.*, *117*, 1779, 1989.
- Zou, X., I.M. Navon and F.X. Ledimet, An optimal nudging data assimilation scheme using parameter estimation, *Q. J. R. Meteorol. Soc.*, *118*, 1163, 1992.

---

L. Bengtsson and J. Feichter, Max-Planck-Institute for Meteorology, Bundesstrasse 55, 20146 Hamburg, Germany (e-mail: bengtsson@dkrz.d400.de; feichter@dkrz.d400.de)  
 A. B. M. Jeuken and P. C. Siegmund, Royal Netherlands Meteorological Institute, Postbus 201, 3730 AE De Bilt, Netherlands (e-mail: jeuken@knmi.nl; siegmund@knmi.nl)

(Received October 12, 1995; revised April 9, 1996; accepted April 9, 1996.)