

## Ensemble Held–Suarez Test with a Spectral Transform Model: Variability, Sensitivity, and Convergence

HUI WAN

*International Max Planck Research School on Earth System Modelling, and Max Planck Institute for Meteorology, Hamburg, Germany*

MARCO A. GIORGETTA

*Max Planck Institute for Meteorology, Hamburg, Germany*

LUCA BONAVENTURA

*MOX-Dipartimento di Matematica, Politecnico di Milano, Milan, Italy*

(Manuscript received 6 October 2006, in final form 12 July 2007)

### ABSTRACT

The idealized test case proposed by Held and Suarez is carried out with the atmospheric general circulation model ECHAM5 of the Max Planck Institute for Meteorology. The aim is to investigate the sensitivity of the solutions of the spectral dynamical core to spatial and temporal resolution, and to evaluate the numerical convergence of the solutions. Low-frequency fluctuations at time scales as long as thousands of days are found in ultralong integrations. To distinguish the effect of changed resolution from the fluctuations caused by the internal variability, the ensemble method is employed in experiments at resolutions ranging from T31 to T159 with 16 to 81 vertical levels. Significance of the differences between ensembles is assessed by three different statistical tests. Convergence property of the numerical solution is concisely summarized by a ratio index.

Results show that the simulated climate state in the Held–Suarez test is sensitive to spatial resolution. Increase of horizontal resolution leads to slight weakening and poleward shift of the westerly jets. Significant warming is detected in high latitudes, especially near the polar tropopause, while the tropical tropopause becomes cooler. The baroclinic wave activity intensifies considerably with increased horizontal resolution. Higher vertical resolution also leads to stronger eddy variances and cooling near the tropical tropopause, but equatorward shift of the westerly jets. The solutions show an indication of convergence at T85L31 resolution according to all the three statistical tests applied. Differences between integrations with various time steps are judged to be within the noise level induced by the inherent low-frequency variability.

### 1. Introduction

In the area of climate modeling, sensitivity of simulated climate to increased spatial resolution is a topic of both scientific interest and practical value. There have been a number of studies on this issue using various models. Examples are experiments performed by Boyle (1993) and Phillips et al. (1995) with the operational forecast model of the European Centre for Medium-Range Weather Forecasts (ECMWF), by Williamson et

al. (1995) with the climate model of the National Center for Atmospheric Research (NCAR), and by Stratton (1999), Pope et al. (2001), and Pope and Stratton (2002) with the climate model of the Hadley Centre for Climate Change. Roeckner et al. (2006) performed a series of Atmospheric Model Intercomparison Project–type (AMIP; Gates et al. 1999) experiments using the most recent version of the Max Planck Institute for Meteorology atmospheric general circulation model, ECHAM5, with resolutions ranging from T21L19 to T159L31. Among many others, these studies used full atmospheric models with various physics parameterizations included. Mixed results were obtained and some aspects of the convergence properties were found to be

---

*Corresponding author address:* Hui Wan, Max Planck Institute for Meteorology, Bundesstrasse 53, 20146 Hamburg, Germany.  
E-mail: hui.wan@zmaw.de

model dependent. Complex interactions between the dynamical part and the physics parameterizations made it difficult to untangle the intertwined effects. To reduce the complexity, it may be helpful to concentrate on the idealized dynamics. In this study, we investigate a specific issue with the ECHAM5 model: given a “perfect” parameterization, how does the dynamical core behave at different resolutions?

The problem associated with the test of dynamical core resides in the fact that no exact solutions are available to the primitive equations when realistic forcing is applied. Without the aid of analytical solutions, it is difficult to identify and quantify errors in three-dimensional numerical models. In recent years, the test case proposed by Held and Suarez (1994, hereafter HS94) has met considerable acceptance, in which the physics parameterizations of the full atmospheric general circulation models (GCMs) are replaced by prescribed forcing and dissipation. The test case is simple by design, but forces the models to produce circulations that are reasonably realistic in many aspects. Many modeling groups have been using it as the first step for validation and intercomparison of the dynamical cores of global atmospheric GCMs. Previous studies have used this test case to investigate the convergence of dynamical cores with increased horizontal resolution (e.g., Boer and Denis 1997), to explain differences between Eulerian and semi-Lagrangian dynamics (Chen et al. 1997), and to investigate the sensitivity to vertical resolution (Williamson et al. 1998). Jablonowski (1998) and Ringler et al. (2000) reported in detail the responses of new geodesic dynamical cores at different horizontal resolutions. Pope and Stratton (2002) used this idealized test to help determine the processes governing horizontal resolution sensitivities in the Hadley Centre Atmospheric Climate Model, version 3 (HadAM3). In these studies, the Held–Suarez test case has been implemented in different ways for different purposes. Usually the proposal in HS94—namely, obtaining the model climate from the last 1000 days of a 1200-day integration—is adopted, while shorter integrations are used in some other studies (e.g., Ringler et al. 2000). Boer and Denis (1997) and Jablonowski (1998) subdivided a 1200-day integration to shorter periods so as to get independent realizations for their analysis. So far, the test setup has not been analyzed in detail in view of the inherent properties of the atmospheric motions.

In the present study, the Held–Suarez test is carried out with the ECHAM5 model. Unlike many other applications, the intention here is not to use this test case to validate the spectral method, which is a well-established algorithm for solving the governing equa-

tions of the atmosphere, but to use simplified experiments to help understand the behaviors of the full model. In relation to Roeckner et al. (2006), who described the resolution sensitivity of the full ECHAM5 GCM, it makes sense to use the same model configuration and only replace the physics parameterizations by idealized forcing. Keeping all the other settings exactly the same allows for a clean comparison of our results with those obtained from the full model, and thus helps in understanding the reasons for the differences observed between simulations in Roeckner et al. (2006).

Experiments are first conducted in the traditional way (HS94) to provide results that can be directly compared with other models. Two ultralong integrations (about 60 yr) at T42L19 and T85L31 resolution are then carried out to investigate the internal variability of this test case. As will be shown later, an important impact of changed resolution on the simulated climate state appears to be the meridional shift of the westerly jets. Meridional wobbling of the jets around their centers at about 42° latitude is observed as the main feature of the long-term variation. How to distinguish these similar signals of different origin is therefore the biggest problem for assessing the convergence of the numerical solutions. This problem is avoided by the ensemble method, which is used in this study for experiments at various spatial and temporal resolutions. The sensitivity analysis and convergence assessment are performed in the sense of ensemble distributions and means.

The rest of this paper is organized as follows. Section 2 contains a brief introduction to the dynamical core of ECHAM5, as well as an overview of the model’s performance in the Held–Suarez test. Section 3 presents the analysis on the internal variability of the test case. Section 4 describes in detail the design of the ensemble experiments and the methodologies used for evaluating the convergence. The results are discussed in section 5. Conclusions are drawn in section 6.

## 2. Dynamical core of ECHAM5 and the 1000-day climate

The dynamical core of the ECHAM5 model (Roeckner et al. 2003) originates from the ECMWF operational forecast model cycle 36. It employs the spectral transform method with triangular truncation to numerically solve the hydrostatic primitive equations of the atmosphere (Hoskins and Simmons 1975; Roeckner et al. 2003). The prognostic variables are vorticity, divergence, temperature, and logarithm of surface pressure. A semi-implicit leapfrog scheme is used for time integration, with the growth of spurious computational modes inhibited by the Asselin filter. In the vertical, a

hybrid coordinate system is used that coincides with pure sigma levels near the earth's surface, pure pressure levels near the model top, and transitional levels in between. In the standard configuration, the uppermost computational level is located at 10 hPa with a total of either 19 or 31 levels. Locations of these levels are shown in appendix A. A second-order energy and angular momentum conserving scheme is used for finite differencing in the vertical (Simmons and Burridge 1981). The horizontal diffusion scheme takes the form of a scale selective hyperviscosity applied to vorticity, divergence, and temperature. To avoid spurious wave reflection at the upper boundary, the damping is enhanced in the upper layers by decreasing the order of the hyper-Laplacian operator. A detailed description of the horizontal diffusion scheme is given in appendix B.

As the first step of this study, we follow the original proposal of HS94 and run ECHAM5 with the specified forcing for 1200 days. Some of the zonal-mean 1000-day statistics at T63L19 resolution are shown in Fig. 1, which can be directly compared to previous studies with other models at similar resolutions, for example, HS94, Chen et al. (1997), Jablonowski (1998), Ringler et al. (2000), and Lin (2004), among others. The butterfly-like structure in zonal wind simulated by ECHAM5 is very similar to other models. There is a single westerly jet in each hemisphere maximizing around 250 hPa, with maximum wind of about  $30 \text{ m s}^{-1}$ . Easterlies appear in the equatorial and polar lower atmosphere, as well as in the tropics near the model top.

Results in the tropical stratosphere seem to be model dependent. Some models simulate an easterly wind band strengthening with altitude (e.g., HS94; Ringler et al. 2000), while the Goddard Earth Observing System (GEOS) GCM (Chen et al. 1997) and ECHAM5 present a closed cell near the tropopause. This discrepancy is possibly caused by horizontal diffusion. Note that ECHAM5 at T63 resolution uses eighth-order Laplacian for diffusion in the troposphere and lower orders in the upper layers (see the table in appendix B), which is different from the fourth-order diffusion employed in many other models. We have performed additional experiments with ECHAM5 using the same order of diffusion (either eighth or fourth) on all vertical levels. It is found that once the enhanced damping near model top is removed, the easterly wind maximum at the tropical tropopause disappears, and the pattern switches to what we have seen in HS94 and Ringler et al. (2000). On the other hand, the zonal wind patterns corresponding to eighth- and fourth-order diffusion are hardly distinguishable by visual comparison. In all the additional experiments we have tried, the circulation in the troposphere is never evidently affected. It should be

noted that the tropics and stratosphere are in fact inactive by design in this test case and are not the regions of interest here. Thus the intermodel differences in the tropics and in the stratosphere can be ignored. Validity of the conclusions drawn for the troposphere in this study will not be affected by the use of extra damping in the top model layers in ECHAM5.

The baroclinic wave activity in this test case, as indicated by the transient eddy kinetic energy and temperature variance in Figs. 1e and 1f, is concentrated in the midlatitude regions. The single maximum of kinetic energy in each hemisphere appears at 250 hPa near  $45^\circ$  latitude, exactly where the westerly jet resides. Easterlies in the tropics show little variance. The maximum temperature variance appears in lower troposphere and extends upward and poleward, with a second maximum of smaller magnitude occurring near the tropopause. The eddy motions transport heat in the poleward direction at most grid points (Fig. 1d). The strongest fluxes coincide with the strongest temperature variance in middle latitudes, resulting in heat flux divergence in low latitudes and convergence in high latitudes. The momentum flux converges in middle latitudes and maintains the westerly jets there (Fig. 1c). These features are generally in good agreement with those shown in HS94 and many other studies. Although no analytical solution is known to the Held–Suarez test case, this agreement can be considered as an indication that the performance of ECHAM5 is *reasonable*, and the conclusions we draw from this model are possibly also true in many other models.

### 3. Internal variability

The basic idea of the Held–Suarez test case is to validate dynamical cores by evaluating the long-term statistical properties of a balanced three-dimensional global circulation. The test case does not include such external forcings as varying boundary condition at seasonal or even longer time scales. Relaxation of temperature to the prescribed radiative equilibrium takes effect on a time scale of 40 days. The  $e$ -folding time for the Rayleigh friction is one day at the earth's surface. The statistics calculated over 1000 days, as proposed in HS94, are good representatives of the basic features of the simulated climate state. However, inherent internal variability causes notable fluctuations even in the 1000-day average of, for example, the zonal-mean zonal wind, as detected in ultralong experiments of over 22 000 days (about 60 yr). We have performed the ultralong integrations at two different resolutions (T42L19 and T85L31). In the following paragraphs, we first present the analysis of the T42L19 simulation, and then briefly summarize the results of T85L31.

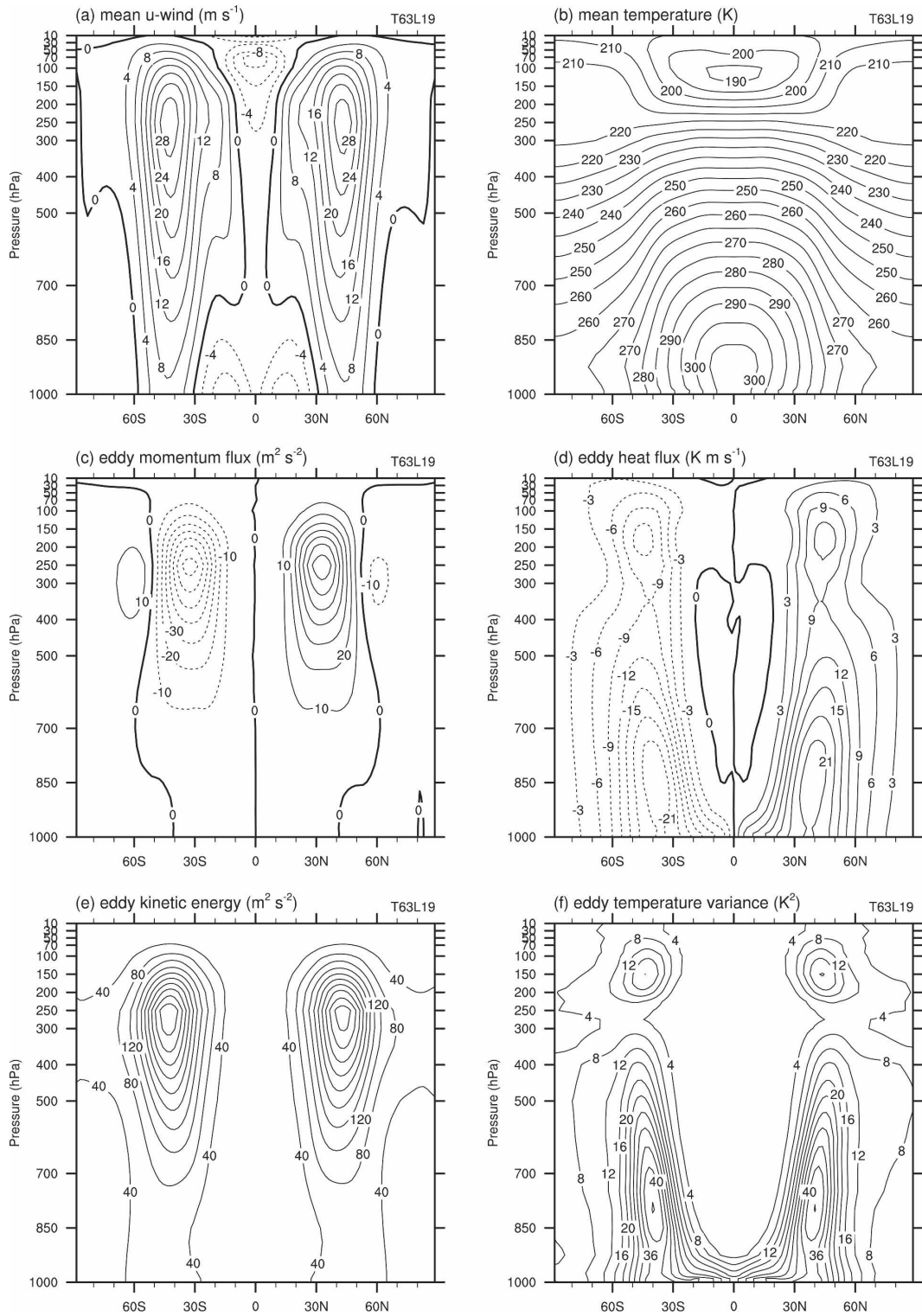


FIG. 1. The zonal mean of (a) time-mean zonal wind, (b) time-mean temperature, (c) eddy momentum flux, (d) eddy heat flux, (e) eddy kinetic energy, and (f) eddy temperature variance simulated with ECHAM5 at T63L19 resolution. Dashed contours indicate negative values. The statistics are calculated over the last 1000 days of a 1200-day integration.

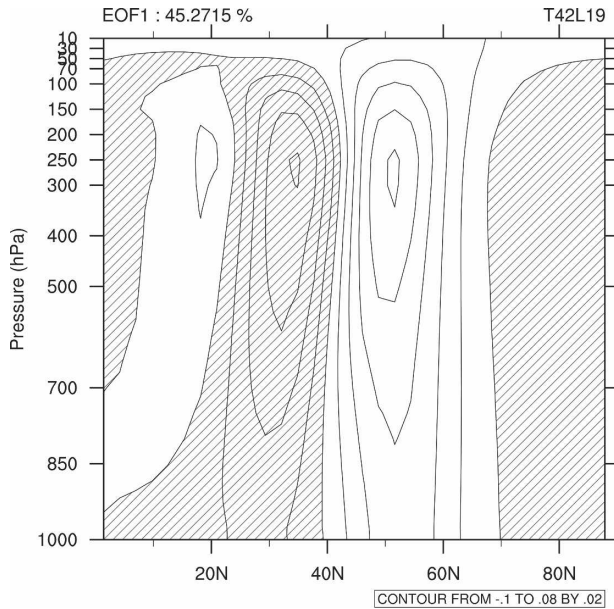


FIG. 2. First EOF of the daily zonal-mean zonal wind in the NH in an ultralong integration at T42L19 resolution, with negative values shaded.

To investigate the low-frequency variability in detail, the empirical orthogonal functions (EOF) analysis is applied to the 22 000-day daily output of the zonal-mean zonal wind. (A 200-day period has already been removed from the beginning of the whole time series to get rid of the spinup period.) Because of the feature of the prescribed forcing in this test case, symmetry is expected between the Northern (NH) and Southern Hemisphere (SH) in a statistical sense. However, the actual circulation in a specific snapshot can be quite far from symmetric. Even when the zonal-mean state is concerned, the fluctuations in the two hemispheres are not necessarily in phase. Therefore we calculate the EOFs for each hemisphere separately and present here results of the NH only. Calculation for the SH conveys essentially the same information.

The first EOF accounts for 45.27% of the total variance and indicates the meridional shift of the westerly jets (Fig. 2). The second EOF explains 13.42% and illustrates the strengthening and narrowing of the jets (not shown). The power spectrum of the first principal component (PC1) is shown in Fig. 3a, in which three different regimes can be detected. For high frequencies with periods less than 8 days, the spectrum decays rapidly with increasing frequency, showing a slope close to  $\kappa^{-4}$  where  $\kappa$  stands for frequency. The spectrum is relatively white in the period range from 8 to 40 days, and grows slowly as the period increases further, indicating a low-frequency variability. The raw time series of PC1

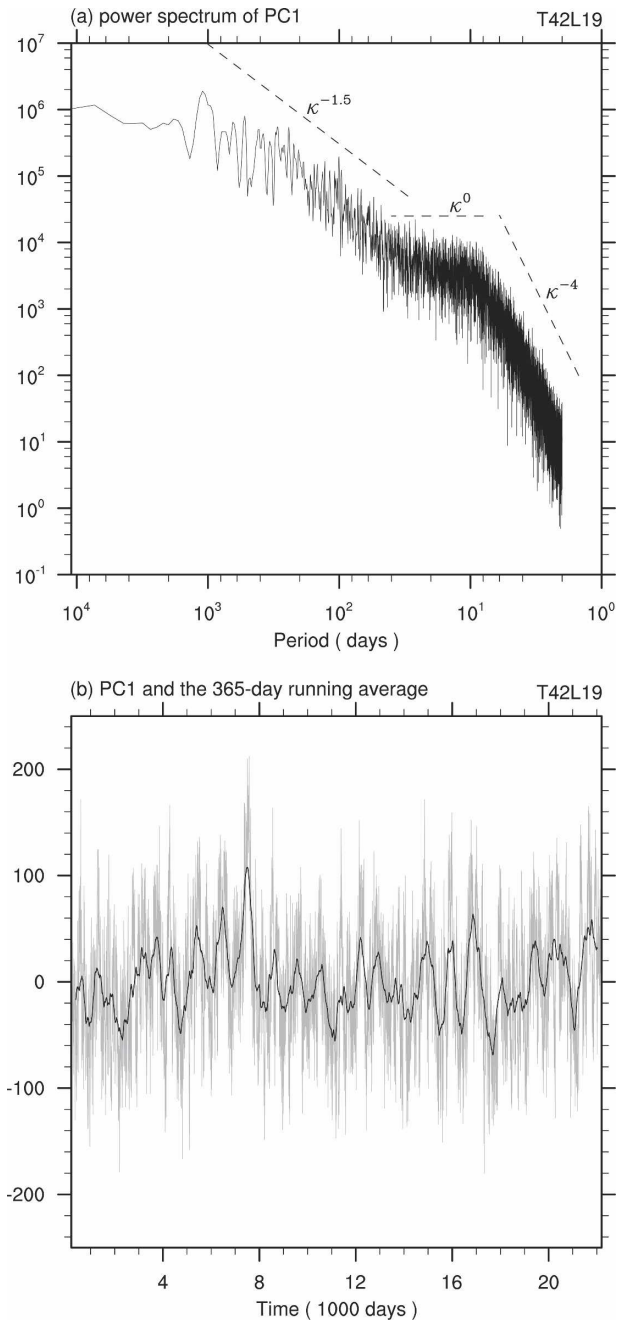


FIG. 3. (a) The power spectrum of PC1 corresponding to Fig. 2. (b) PC1 (the gray curve) and the 365-day running average (the black curve).

is shown in Fig. 3b with the thin gray curve being PC1 itself and the thick black curve being the 365-day running average. Variations at time scales of thousands of days are detectable even by eye. In fact the issue of internally generated low-frequency variability has already been addressed by several previous studies, for example, James and James (1989, 1992), James et al.

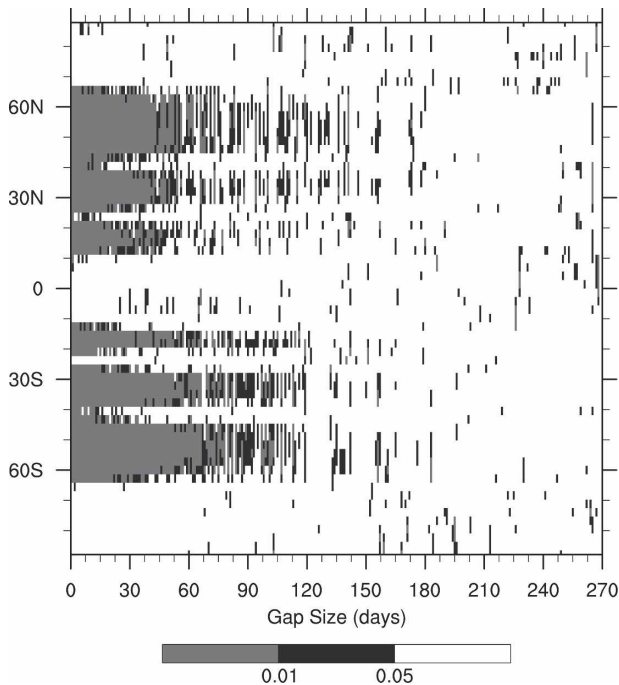


FIG. 4. Statistical significance of the lag-1 autocorrelation coefficient of the 90-day-mean zonal-mean zonal wind on the 250-hPa pressure level. The result is obtained from an ultralong integration with ECHAM5 at T42L19 resolution. See text for further details.

(1994), and Müller et al. (2002). It has been found that the inherent chaotic nature of the flow in a global atmospheric model with idealized heating and friction could lead to variabilities at time scales of several years or even longer.

The low-frequency climate fluctuations in the 1000-day statistics lead to some extent of uncertainty when the 1000-day climate is used to quantitatively assess the convergence property of the solutions by using only one realization. To deal with this problem, Boer and Denis (1997), in another idealized test with similar aim, divided a 1200-day integration into 10 periods of 120 days. Thirty days were discarded in each period and the remaining 90-day chunks were treated as independent realizations for statistical test. This approach was later adopted by Jablonowski (1998) for the Held–Suarez test with the operational global weather forecast model (GME) of the German Weather Service (Majewski et al. 2002). To get robust results from statistical test, it is necessary to check the independency of the realizations obtained in this way by analyzing the persistence of the climate state.

To do this, the ultralong integration with ECHAM5 described above is divided into 90-day periods separated by a certain number of days. Calculate the 90-day-mean zonal-mean zonal wind for each grid point in

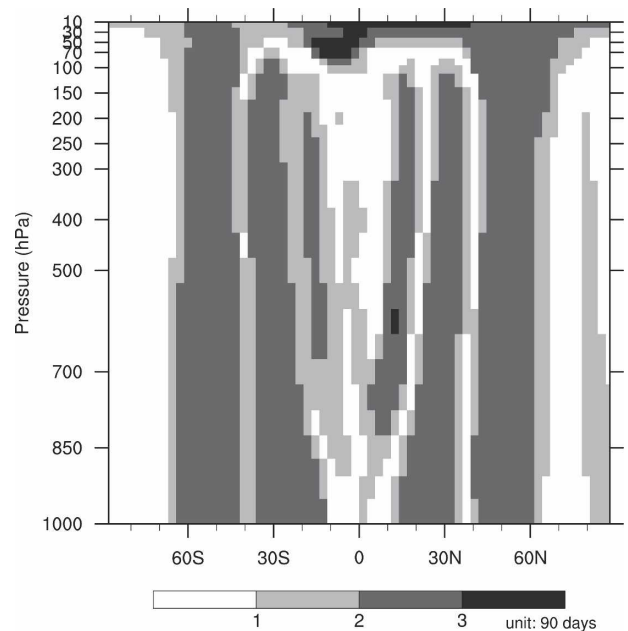


FIG. 5. Estimated decorrelation time of the 90-day-mean zonal-mean zonal wind. The result is obtained from an ultralong integration with ECHAM5 at T42L19 resolution. See text for further details.

the vertical cross section, and a time series of the mean values is obtained. Then the lag-1 autocorrelation coefficient of this new time series is computed as a function of the interval (gap) size between each two 90-day periods. Statistical significance of the autocorrelation coefficients for all latitudes at 250 hPa is displayed in Fig. 4. From this calculation, it is clear that the 90-day chunks are independent only when the gap size is larger than about 150 days. Similar estimates are obtained for all vertical levels in the troposphere, while in the stratosphere the mean state persists even longer.

The persistence of the mean state is also analyzed by estimating a “decorrelation time” of the 90-day mean. For this analysis the long integration is subdivided into 90-day periods with no gap in between. Again we calculate the time and zonal-mean zonal wind of these chunks and the autocorrelation coefficients of the resulting new time series. The decorrelation time is then estimated by

$$\tau_D = 1 + 2 \sum_{l=1}^L \left(1 - \frac{l}{L}\right) \hat{\rho}(l), \quad (1)$$

where  $\hat{\rho}(l)$  denotes the estimated autocorrelation coefficient of lag  $l$  and  $L$  the max lag (Von Storch and Navarra 1995, p. 175). In our calculation  $L = 50$  is used to avoid large biases in  $\hat{\rho}(l)$ . The result is shown in Fig. 5. Since the time series is the 90-day mean, the unit of values in this plot is in fact 90 days. For most grid points

TABLE 1. Default time step (s) of the full ECHAM5 model at various resolutions.

Resolution	T31L19	T42L19	T63L19	T85L19	T106L19	
Time step	2400	1800	1200	900	720	
Resolution		T42L31	T63L31	T85L31	T106L31	T159L31
Time step		1200	720	480	360	240

in the subtropical regions and midlatitudes, the decorrelation time ranges between 2 and 3, meaning every two to three 90-day periods of the original time series can be considered as an independent realization, which is consistent with the estimate from the former approach.

It should be noted that the information we get from Figs. 4 and 5 is not to be considered as an accurate estimate of “the memory” of the system, but only an example showing what kind of situation we may encounter in the Held–Suarez test case. As shown in Fig. 3a, the zonal-mean state has a relatively smooth spectrum without distinct periodicity. In different subsections of a very long run, the dominating variation can appear at different time scales. In fact, if plots like Figs. 4 and 5 are made from different subsections of the 22 000-day series, the computed persistence of the 90-day mean can range from less than 30 days to more than 170 days, and the two hemispheres can appear much less symmetric than in the two figures shown here.

The diagnoses described above are performed for the ultralong simulation at T85L31 resolution as well, which produces by and large the same EOF pattern as Fig. 2, shows quite a similar spectral shape to Fig. 3a, and also illustrates fluctuations at time scales up to thousands of days. The persistence of the 90-day-mean zonal wind estimated from the 22 000-day period seems only one-half as long as in Fig. 4. However, given the observations described in the former paragraph, this cannot yet be readily attributed to the higher resolution. Further investigation of the features of the long-term simulations is outside the scope of this paper. Yet the results found above underscore that subdividing time series from these experiments into separated short periods is not necessarily a safe method to obtain independent realizations. One may still think about following the idea of the decorrelation time, calculating the so-called equivalent sample size  $n'$  as a remedy and applying the Student's  $t$  test. However, as pointed out in Von Storch and Navarra (1995, 18–24), this will not work unless  $n'$  is larger than 30. Consequently, we have decided to use the ensemble technique in the following experiments with the ECHAM5 model to generate composites of statistically independent runs. Details are described in the next section.

#### 4. Ensemble experiments at various resolutions

The central goal of this study, as stated earlier, is to investigate the sensitivity of the simulated model climate to spatial and temporal resolution in the idealized Held–Suarez test case. The questions we are trying to answer are whether the numerical solutions converge and, if so, at which resolution the convergence is achieved within a useful tolerance for practical purposes. This section describes in detail the experiments conducted and the methods employed to analyze the results.

##### a. Experimental design

Three groups of experiments are performed. In the first group, integrations with the same vertical grid but different horizontal resolutions are compared. In the second group, we compare results obtained with the same horizontal resolution but different vertical grids. The third group of integrations has the same spatial resolution but differ in time step. Details about these experiments are given as follows:

##### 1) HORIZONTAL RESOLUTION EXPERIMENTS

Integrations are conducted at resolutions T31, T42, T63, T85, T106, and T159. Two different sets of vertical grids are used: 19 vertical layers for horizontal resolutions from T31 to T106, and 31 layers for horizontal grids from T42 to T159. These are in fact the same resolutions as used in Roeckner et al. (2006), except that we have excluded the lowest resolution (T21L19) of their study. Default time steps of the full ECHAM5 model are adopted for each resolution (Table 1). In the convergence analysis, the L19 and L31 runs are compared separately.

##### 2) VERTICAL RESOLUTION EXPERIMENTS

The second group of integrations is conducted with different vertical resolutions. T85 is chosen as the horizontal grid according to results from the first group. The vertical grid with 31 layers is used as the “control grid” from which higher and lower vertical resolutions are generated. As described in appendix A, the vertical coordinate of ECHAM5 is effectively defined by two

sets of coordinate parameters  $A$  and  $B$  that specify the interfaces between each two layers. The grid generation method employed here always uses the same  $A$  and  $B$  values as listed in the table in appendix A for the first two interfaces ( $k = 0, 1$ ) from the model top. Thus, the upper and lower boundary of the first vertical layer is kept unchanged and the uppermost computational level is fixed at 10 hPa. The other 30 layers of the L31 grid are either coarsened to 15 layers or refined into 45, 60, or 80 layers via spline interpolation of the  $A$ s and  $B$ s as functions of the normalized interface indices. The resulting grid has in total 16, 46, 61, or 81 layers, as illustrated by the figure in appendix A. All simulations in this group are performed with a 480-s time step, which is the default choice for the standard full ECHAM5 model at T85L31 resolution.

### 3) TIME STEP EXPERIMENTS

The third group of experiments is carried out at the fixed T85L31 spatial resolution with five different time steps: 120, 240, 480, 900, and 1200 s. These are meant to investigate the impact of time step on the solutions.

#### *b. Methodology*

##### 1) INITIALIZATION

The Held–Suarez test focuses on the long-term statistical properties of the simulated global circulation, which is supposed to be dominated by the baroclinic wave activity. Figure 6 shows the state ECHAM5 will evolve to, if an isothermal (300 K) state at rest is used as the initial state without perturbation to break the symmetry. In this study, the perturbation is introduced by adding random noise to the spectral coefficients of vorticity and divergence.

In the first few days after initialization, the baroclinic wave grows very slowly due to the randomness and small magnitude of the initial noise. The circulation evolves very fast toward the other equilibrium state shown in Fig. 6. After about 55 model days the rapid development stage of the baroclinic instability sets in. By day 100, a quasi-equilibrium state has already been reached in tropospheric layers, as can be detected from the zonal variance of the wind and temperature (not shown).

##### 2) ENSEMBLES

Based on the analysis in the previous sections, the ensemble method is employed in our experiments to evaluate the sensitivities of solutions. Since it is computationally expensive to run the model several times for 1200 days at each resolution, and in fact the 100-day

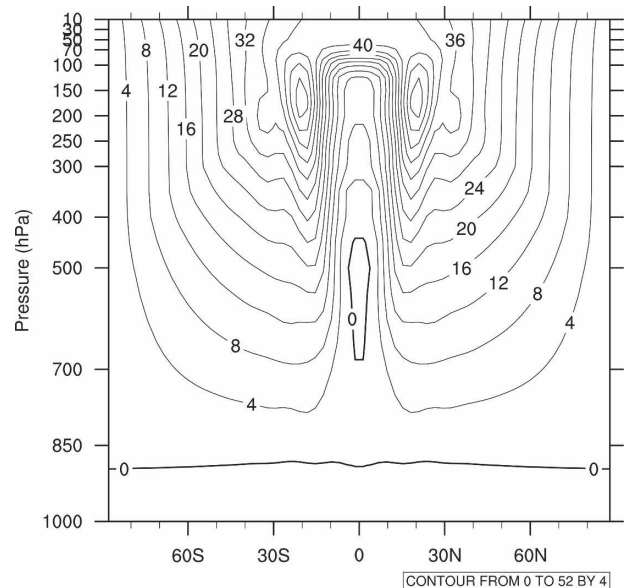


FIG. 6. Snapshot of the zonal wind ( $\text{m s}^{-1}$ ) along  $0^\circ$  longitude at day 1200 in a simulation with ECHAM5 starting from an isothermal static state.

statistics are already good representatives of the climate state, we get the ensemble by performing 10 shorter runs. As already mentioned, isothermal states with random noise are used as initial conditions. For each resolution, the NCAR Command Language random number generator is used to obtain a large amount of random values with a normal distribution. To ensure independency, nonoverlapping subsections of these random numbers are used to initialize the 10 ensemble members. Each integration proceeds 300 days. The first 200 days are discarded and the climate state is calculated over the third 100 days. (The number 100 is chosen rather than 90 or anything else just to facilitate postprocessing.) This approach is not quite efficient in the sense that we “waste” two-thirds of every integration. However, this is the most convenient way to make sure that the realizations we finally obtain are really independent and not affected by the spinup process.

##### 3) QUANTITATIVE EVALUATION OF CONVERGENCE

Given that no analytical solution is available for this test case, the highest resolution in each experiment group is taken as the reference, namely, T106L19 and T159L31 for the L19 and L31 simulations, respectively, in the first group, T85L81 for the second group, and the ensemble with 120-s time step in the third group. Significance of the differences between ensembles is assessed at each grid point using three statistical tests as described below.



The Kolmogorov–Smirnov test (K–S test) determines whether two datasets differ significantly by comparing their cumulative distribution functions. The cumulative distribution function of a dataset with  $N$  events is estimated by the cumulative *fraction* function  $S_N(x)$ : if the  $N$  events are located at values  $x_i, i = 1, \dots, N$ , then  $S_N(x)$  is the function giving the fraction of data strictly smaller than a given value  $x$ . The test statistic  $D$  is defined as the maximum value of the absolute difference between two cumulative distribution functions:

$$D = \max_{-\infty < x < \infty} |S_{N_1}(x) - S_{N_2}(x)|. \tag{2}$$

The significance of an observed value  $D_o$  under the null hypothesis that the two distributions are identical is given approximately by the formula

$$P(D > D_o) = Q_{\text{KS}} \left[ \left( \sqrt{N_e} + 0.12 + \frac{0.11}{\sqrt{N_e}} \right) D_o \right], \tag{3}$$

where  $P$  stands for probability,  $N_e$  is given by

$$N_e = \frac{N_1 N_2}{N_1 + N_2}, \tag{4}$$

and the function  $Q_{\text{KS}}$  has the form

$$Q_{\text{KS}}(\lambda) = 2 \sum_{j=1}^{\infty} (-1)^{j-1} e^{-2j^2 \lambda^2}. \tag{5}$$

The important properties that render the K–S test appropriate for utilization in this study are as follows (Press et al. 1992, 617–620):

- The test is nonparametric, meaning that it makes no assumption about the distribution of data investigated. This is advantageous since it is not known a priori how our ensemble members are distributed and the distribution is not easy to estimate given the relatively small ensemble size.
- The K–S statistic  $D$  is invariant under reparameterization of  $x$ ; in other words, if variable  $x$  is transformed to  $y$  via a monotonic function  $y = y(x)$  (e.g.,  $y = \ln x$  if applicable), the statistic  $D$ , and consequently its significance, will remain unchanged.
- The approximation (3) becomes asymptotically accurate as  $N_e$  becomes large, but is already quite good for  $N_e \geq 4$ . In our ensemble experiments we have  $N_1 = N_2 = 10$  and  $N_e = 5$ .
- The K–S test tends to be most sensitive around the median value and less sensitive at the extreme ends of the distribution. This suits very well our focus on the climate state.

To prepare data from our ensemble experiments for the K–S test, the meteorological fields are obtained at a daily interval on the Gaussian grid at hybrid vertical levels. These are then used for calculations of the zonal-mean climate state (the six quantities in Fig. 1). To compare each pair of results at different resolutions, linear interpolation is used to transform the higher-resolution result to the lower resolution to minimize the artificial difference that may be introduced by the interpolation. The interpolation is done for each ensemble member separately, resulting in a 10-member sample for each ensemble at each grid point. Then the K–S test is applied: the cumulative fraction functions  $S_{N_1}(x)$  and  $S_{N_2}(x)$  are computed using the algorithm in Press et al. (1996, p. 1274). As for the  $Q_{\text{KS}}$  function, the sum in (5) is calculated using only the first 100 terms. (We have checked the convergence of this calculation by defining the criterion as “the 100th term is either smaller than  $10^{-8}$  times the 100-term sum, or smaller than 0.001 times the 99th term, both in the sense of absolute value.” This criterion is always satisfied in our calculations.) The null hypothesis is rejected if the probability computed with (3) is less than 0.05 (or 0.01).

To illustrate results of the statistical test in a concise way, a “ratio index”

$$r_{\text{KS}} = \frac{m}{M} \tag{6}$$

is defined. Here  $m$  is the number of grid points where two ensembles of solutions are judged to be significantly different at 0.05 significance level, and  $M$  is the total number of grid points in the vertical cross section.

In addition to the K–S test, the Mann–Whitney test and Student’s  $t$  test are also applied to compare ensemble means. The Mann–Whitney test is also nonparametric although it requires that the probability functions of the two samples have the same shape. This test involves sorting the combined sample of  $N_1 + N_2$  realizations and calculating the test statistic by summing up the rank of all realizations in one original sample. The details can be found in Von Storch and Zwiers (2001, p. 117). For implementation in this study, the same interpolated data as in the K–S test are used. The critical values of the test statistic are from appendix I in Von Storch and Zwiers (2001).

The  $t$  test is the parametric counterpart of the Mann–Whitney test with a more restrictive requirement that the underlying distributions are normal. We do not detail the algorithm here since it is very widely used in climate research. The only point to note is that for data preprocessing, the ensemble mean and standard deviation are first computed on the original model grid and then interpolated to a common grid for comparison.

The subsequent calculations are done with the  $t$ -test function in the NCAR Command Language. The same ratio index as (6) is defined for the  $t$  test and for the Mann–Whitney test as well.

In the sensitivity analysis presented in the next section, we have found that all three tests agree quite well with one another. Regions of the vertical cross section in which differences are judged to be significant by the three tests show only marginal differences. This is possibly because the additional assumptions made by the Mann–Whitney and  $t$  test are actually satisfied, given the way in which the ensemble experiments are conducted and the property of the quantities investigated. Because of space constraint, only results of the K–S test will be presented in the following sections.

## 5. Sensitivity and convergence of the numerical solutions

In this section we compare the responses of the ECHAM5 dynamical core to the idealized forcing at different resolutions by analyzing results of the ensemble experiments. The climate statistics investigated are the same as in Fig. 1.

### *a. Sensitivity to horizontal resolution*

The L19 and L31 simulations are similar regarding changes with horizontal resolution. The variables most sensitive to resolution are the eddy kinetic energy and eddy temperature variance, which show dramatic increases when the grid size gets smaller (Figs. 7e,f). This is fully expected since the dynamical core at higher horizontal resolution is able to resolve motions at smaller scales and hence allow for stronger wave activity. Besides the enhancement, the core regions of temperature variance in lower troposphere move upward. Large values at the earth's surface in the T31 and T42 runs disappear when the grid gets finer, while the high-resolution simulations show a well-defined maximum near 800 hPa (not shown).

All the second-order statistics investigated also exhibit an evident poleward displacement with increased resolution. The displacement, together with the enhancement in eddy heat flux, leads directly to warming in middle and high latitudes throughout the atmosphere, with the most notable signal occurring in the upper troposphere (Fig. 7b). It is worth noting that similar changes have been found by Roeckner et al. (2006) in realistic climate simulations with the full ECHAM5 model. Moreover, the pattern of temperature difference in Fig. 7b resembles closely the result from a dynamical core experiment with the finite-dif-

ference model HadAM3 (Fig. 5d in Pope and Stratton 2002). Locations of the strongest warming near the poles are impressively similar although these two models use completely different numerical schemes. This implies that the commonly found reduction of cold biases in high-resolution climate simulations with full models may be reasonably attributed to the sensitivity to the resolution of the dynamic core. Major changes in zonal wind include the poleward shift of the westerly wind zones as well as downward movement and weakening of the core regions (Fig. 7a). The near-surface easterlies in the tropics remain the same at all the resolutions on the other hand.

Although the changes with increased resolution are evident, the solutions do converge when the horizontal grid is sufficiently fine. Differences between the T106L19 and T85L19 simulation are shown in Fig. 8, where no systematic structure can be detected at all. Simulations at T159L31 and T85L31 differ moderately in terms of eddy kinetic energy. The other climate statistics are not significantly different at most grid points. (The dramatic decrease of differences illustrated by the ratio index in the first two panels in Fig. 10 indicates a clear trend of convergence.)

### *b. Sensitivity to vertical resolution*

Having found indication of convergence at T85 in the previous group of experiments, we proceed by fixing the horizontal resolution at T85 to investigate the impact of vertical resolution. As described in section 4a, a series of vertical grids with the same structure are generated based on the L31 grid. Increasing the resolution from L16 to L31 is approximately equivalent to reducing the layer thickness from 1.2 to 0.6 km in the middle part of the troposphere. Statistical tests of the ensembles reveal that this change leads to moderate enhancement of the baroclinic wave activity (Figs. 9e,f) and cooling near the tropical tropopause (Fig. 9b), which are similar to the effect of higher horizontal resolution. On the other hand, a slight equatorward shift of the westerly jets is detected (Fig. 9a), as being noticed earlier in the AMIP experiment with the full ECHAM5 model by Roeckner et al. (2006). In this idealized test case, centers of the wave activity also shift accordingly (Figs. 9c–f). Further reduction of the grid size does not lead to much change in the climate state, as can be seen in Fig. 10c. The L31 grid seems adequate for the horizontal resolution T85 in this idealized test case.

It should be noted that due to the specification of the Held–Suarez test case, we only concentrate on the troposphere in this study. The conclusion that 31 layers are adequate is valid merely for this highly idealized test for

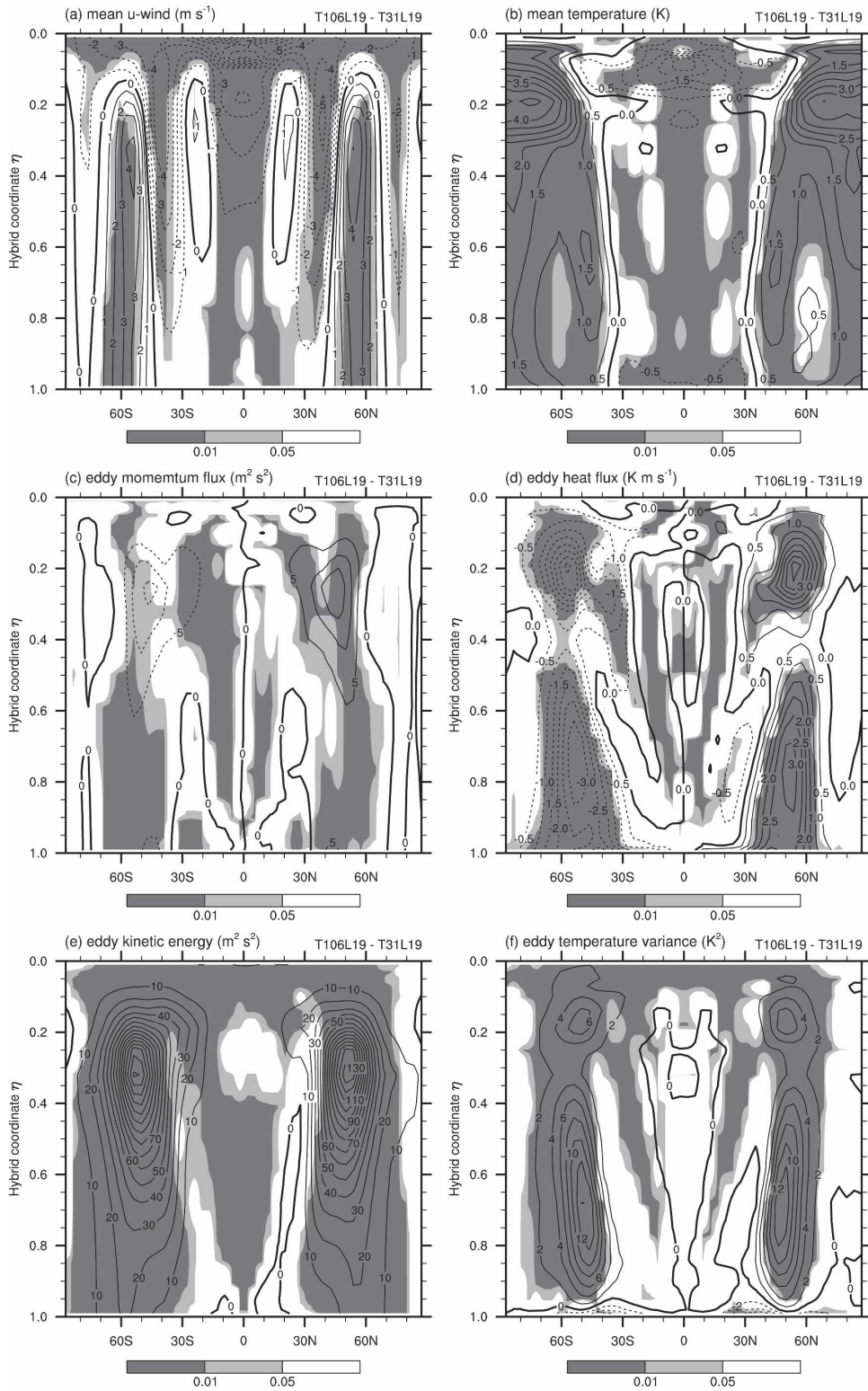


FIG. 7. Differences between the ensemble mean solutions at T106L19 and T31L19 resolution. Dashed contours indicate negative values. Light and dark shaded areas are judged to be significantly different by the Kolmogorov–Smirnov test at 0.05 and 0.01 significance levels, respectively.

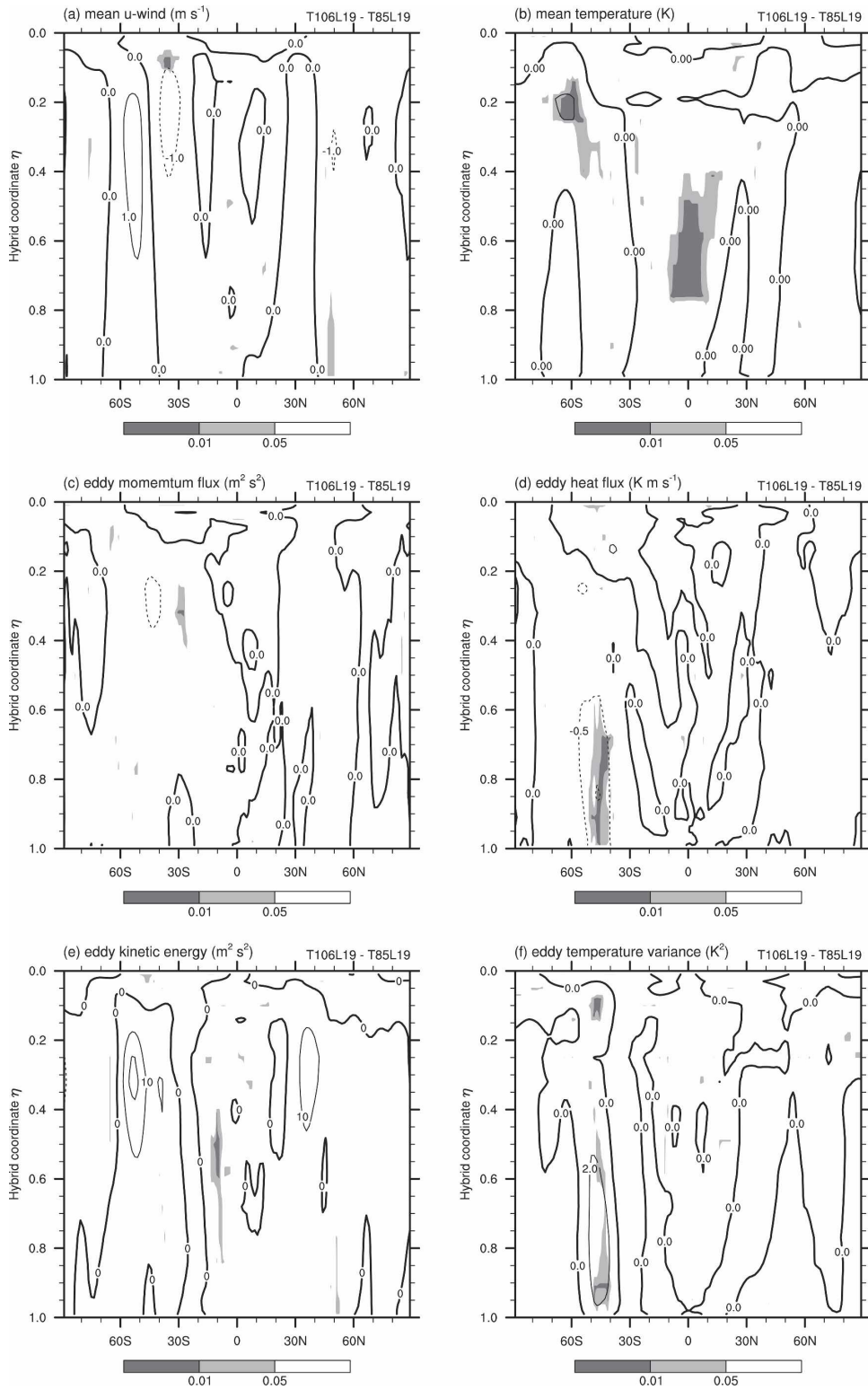


FIG. 8. As in Fig. 7, but for the differences between the T106L19 and T85L19 simulation.

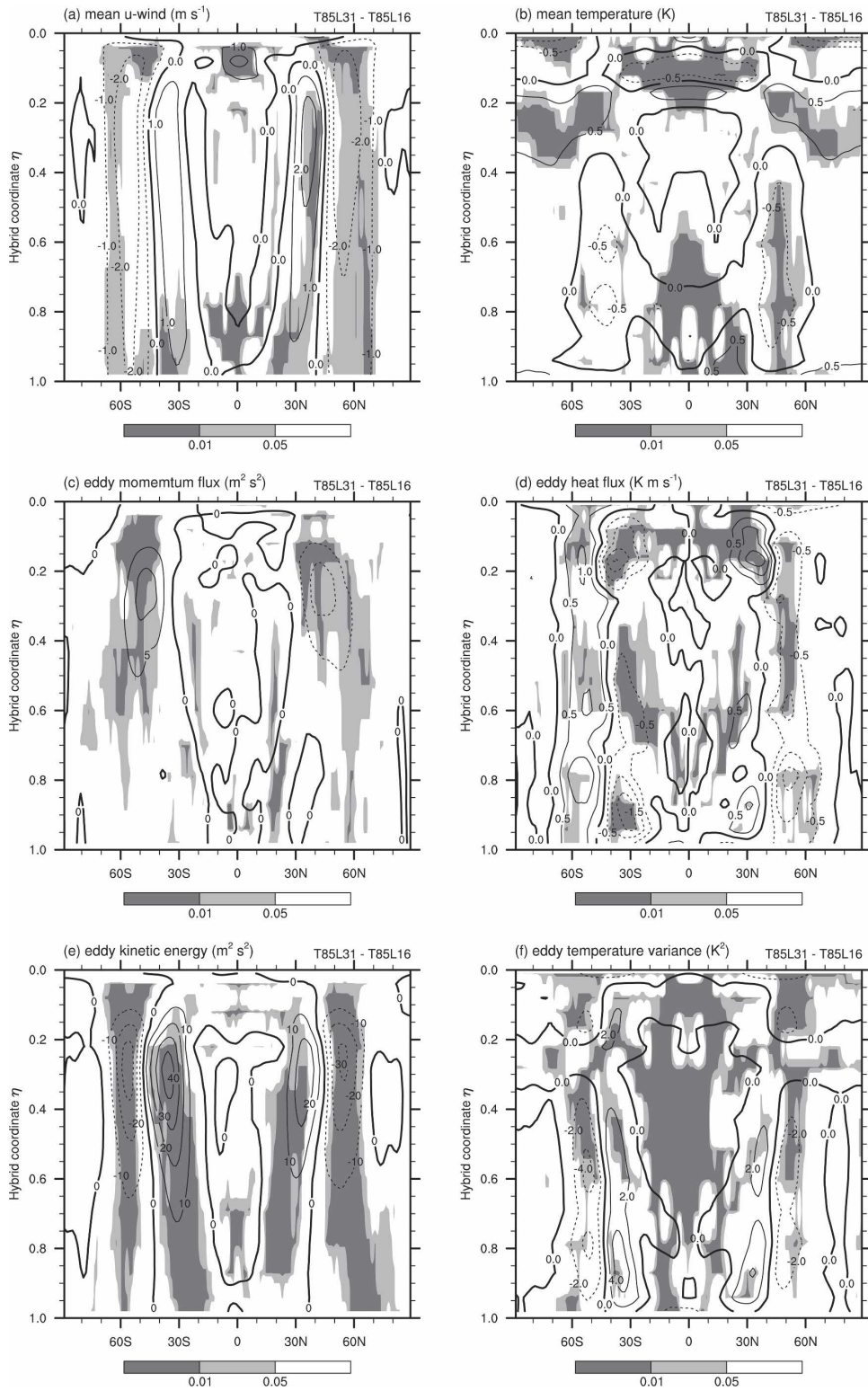


FIG. 9. As in Fig. 7, but for the differences between the T85L31 and T85L16 simulation.

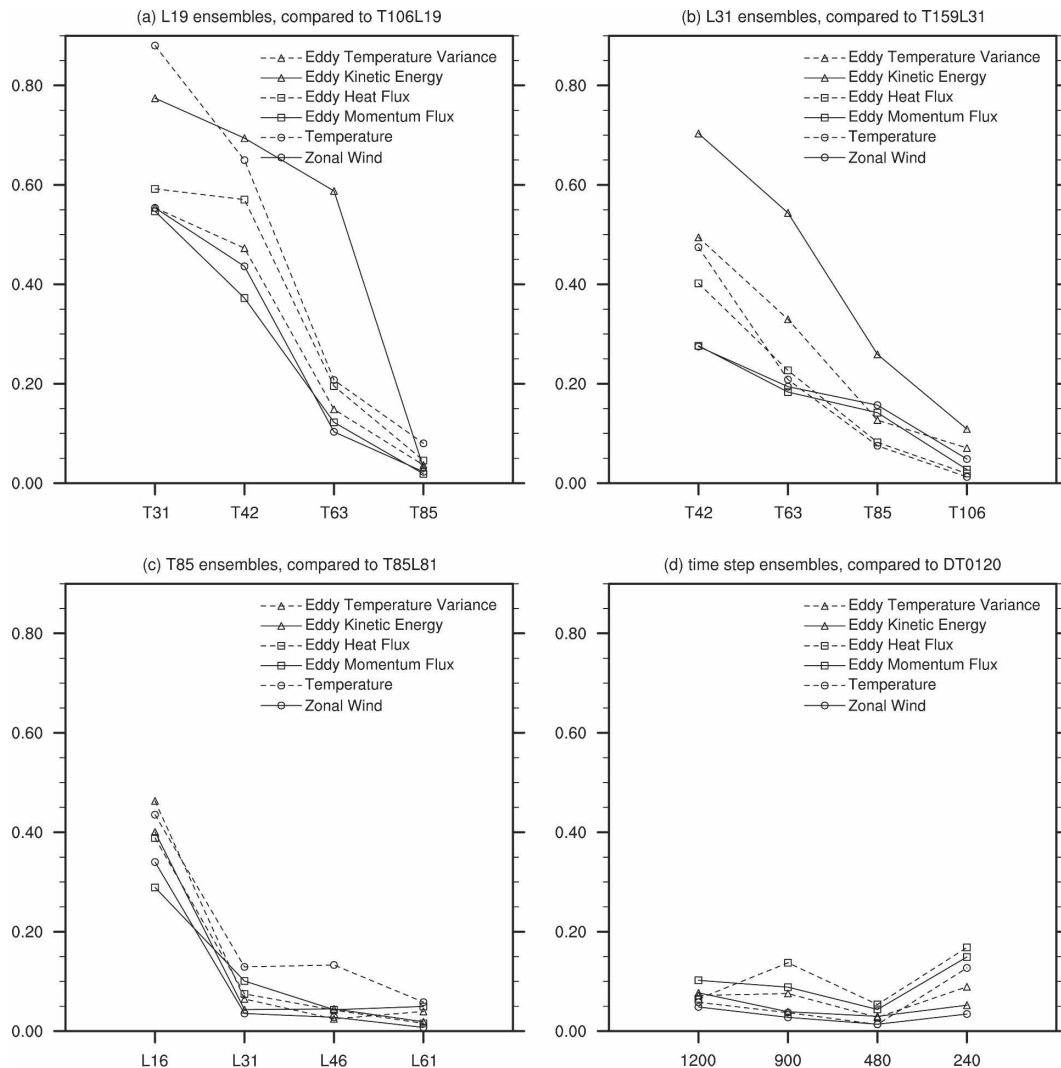


FIG. 10. The ratio index for the Kolmogorov–Smirnov test in (a), (b) horizontal resolution experiments; (c) vertical resolution experiments; and (d) time step experiments.

the dry dynamical core. In a full atmospheric model, on the other hand, the tracer distributions and the physics parameterization schemes will have additional impact on the convergence with the vertical resolution.

### c. Impact of time step

Besides the spatial resolution addressed above, it is also of interest to study the response of the ensemble solutions to different temporal resolutions, and to find out if the merit of higher spatial resolution can be achieved by using shorter time steps. This is the motivation for the third group of our experiments with various time steps. Results show that generally the ensemble solutions do not differ significantly in statistical sense. Differences between the integration with the shortest time step (120 s) and all the other ones are of

the same magnitude. No systematic change is detected when the temporal resolution increases. In fact, differences between one ensemble and the other with immediately shorter or longer time step are often of the opposite sign in the vertical cross section, the distributions resembling those of the ensemble standard deviation (not shown). It is reasonable to regard these differences as sampling errors and conclude that time step changes within the selected range have little impact on the climate state of this test case. Again we need to keep in mind that this conclusion is drawn specifically for this test case with the dynamical core only. It is possible that in an atmospheric model with full physics, some parameterized processes will lead to a stronger sensitivity with respect to time step.

Furthermore, it is worth noting that although the ra-

TABLE A1. Vertical coordinate parameter  $A$  and  $B$  of the 19- and 31-layer ECHAM5 model (from Roeckner et al. 2003) and the corresponding pressure values (in hPa) when  $p_s = 1000$  hPa.

$k$	$A_{k+1/2}$	$B_{k+1/2}$	$P_{k+1/2}$	$A_{k+1/2}$	$B_{k+1/2}$	$P_{k+1/2}$
0	0.000000	0.0000000000	0.00	0.00000000	0.00000000	0.00
1	2000.000000	0.0000000000	20.00	2000.00000000	0.00000000	20.00
2	4000.000000	0.0000000000	40.00	4000.00000000	0.00000000	40.00
3	6046.110595	0.0003389933	60.80	6000.00000000	0.00000000	60.00
4	8267.927560	0.0033571866	86.04	8000.00000000	0.00000000	80.00
5	10609.513232	0.0130700434	119.17	9976.13671875	0.00039086	100.15
6	12851.100169	0.0340771467	162.59	11820.53906250	0.00291970	121.13
7	14698.498086	0.0706498323	217.63	13431.39453125	0.00919413	143.51
8	15861.125180	0.1259166826	284.53	14736.35546875	0.02031916	167.68
9	16116.236610	0.2011954093	362.36	15689.20703125	0.03697486	193.87
10	15356.924115	0.2955196487	449.09	16266.60937500	0.05948764	222.15
11	13621.460403	0.4054091989	541.62	16465.00390625	0.08789498	252.55
12	11101.561987	0.5249322235	635.95	16297.62109375	0.12200361	284.98
13	8127.144155	0.6461079479	727.38	15791.59765625	0.16144150	319.36
14	5125.141747	0.7596983769	810.95	14985.26953125	0.20570326	355.56
15	2549.969411	0.8564375573	881.94	13925.51953125	0.25418860	393.44
16	783.195032	0.9287469142	936.58	12665.29296875	0.30623537	432.89
17	0.000000	0.9729851852	972.99	11261.23046875	0.36114502	473.76
18	0.000000	0.9922814815	992.28	9771.40625000	0.41820228	515.92
19	0.000000	1.0000000000	1000.00	8253.21093750	0.47668815	559.22
20				6761.33984375	0.53588659	603.50
21				5345.91406250	0.59508425	648.54
22				4050.71777344	0.65356457	694.07
23				2911.56933594	0.71059442	739.71
24				1954.80517578	0.76540524	784.95
25				1195.88989258	0.81716698	829.13
26				638.14892578	0.86495584	871.34
27				271.62646484	0.90771586	910.43
28				72.06358337	0.94421321	944.93
29				0.00000000	0.97298521	972.99
30				0.00000000	0.99228150	992.28
31				0.00000000	1.00000000	1000.00

ratio index defined by (6) is expected to converge to 0.05 in theory, the spatial correlation between grid points can result in deviations from this value. The asymptotic property of this index is not easy to evaluate. Nevertheless, Fig. 10d gives us a rough idea about the associated uncertainty. If we take the maximum value of the ratio index in this panel as an estimate and compare it to the other panels, we will come to the same judgments on the numerical convergence as already made earlier in this section.

## 6. Concluding remarks

In this study, the test case proposed by HS94 is carried out with the spectral dynamical core of the ECHAM5 model. From ultralong integrations at T42L19 and T85L31 resolution, it is found that the internal variability of this idealized test case leads to low-frequency variations at time scales as long as thousands of days. The persistence of the climate state is investi-

gated and found to vary considerably among different subsections of the time series. It is not yet clear whether the persistence is affected by resolution of the dynamical core.

Some aspects of the implementation strategy of the Held–Suarez test case are discussed. In view of the low-frequency variability, the ensemble method is employed in experiments with different horizontal and vertical resolutions as well as time steps to assess the sensitivity and convergence of the numerical solutions. The simulated climate state, as represented by ensemble distributions and ensemble means of 100-day statistics, is found to be sensitive to horizontal resolution. When the horizontal grid gets finer, westerly jets in the middle latitudes slightly decrease in strength and shift poleward. Temperature increases in the high latitudes. Baroclinic wave activity steadily intensifies with horizontal resolution. The significance of the differences between results at different resolutions is assessed by the Kolmogorov–Smirnov test, Mann–

Whitney test, and the Student's  $t$  test. A clear trend of convergence is observed.

Increase of vertical resolution leads to stronger eddy variances and equatorward shift of the westerly jets. Evidence of convergence is found in a group of integrations at T85 combined with different vertical resolutions. The L31 grid used by the ECHAM5 model, which has a layer thickness equivalent to 0.6 km in the middle troposphere, seems adequate for simulating the climate state with the Held–Suarez forcing.

Experiments also indicate that time steps within the selected range have little impact on the simulated climate in the absence of realistic parameterizations for physics processes. Differences between integrations with various time steps are no larger than the noise level induced by the inherent variability, and thus can be considered as sampling errors. From all the experiments mentioned above, we come to the conclusion that convergence of numerical solutions by the dynamical core of ECHAM5 has been detected in the Held–Suarez test case. Results at the T85L31 with a 1200-s time step can be considered as a sufficiently good solution.

It is worth noticing that some common features have been observed in the idealized test case with the dynamical core and the climate simulations with the full ECHAM5 model, for example, polar warming in high latitudes at increased horizontal resolution and equatorward shift of the westerly jets with finer vertical grid. The idealized test not only helps us understand the dynamical part of the GCM, but also provides useful hints on the tuning of parameterization schemes when more realistic climate simulations are expected from higher resolutions.

*Acknowledgments.* The authors are grateful to Erich Roeckner, Jin-song von Storch, Christiane Jablonowski, and Rita Seiffert for helpful comments and discussions. We would also like to thank the two anonymous reviewers whose useful and constructive comments helped improve the original manuscript. All experiments in this study were performed on the NEC SX-6 supercomputer at the German Climate Computing Centre (DKRZ) in Hamburg. HW is recipient of a fellowship from the ZEIT Foundation “Ebelin and Gerd Bucerius” in Hamburg, Germany.

## APPENDIX A

### The Vertical Grids

In ECHAM5, the atmosphere is divided into vertical layers by defining the pressure of the interfaces (known as the “half levels”) between them using the formula

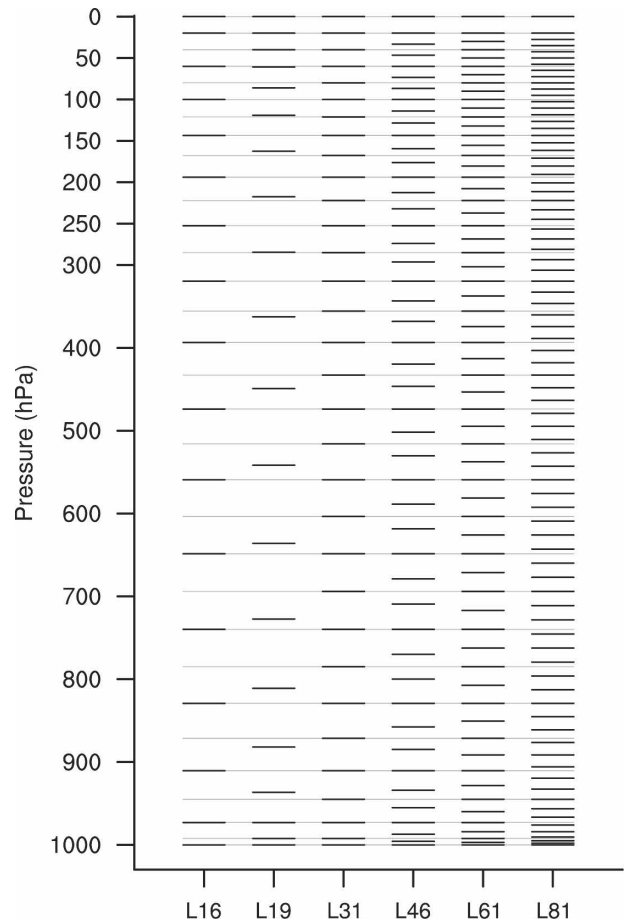


FIG. A1. Locations of the vertical layer interfaces (with  $p_s = 1000$  hPa) at all vertical resolutions used in this study. The thin gray reference lines are the layer interfaces of the L31 grid.

$$p_{k+1/2} = A_{k+1/2} + B_{k+1/2}p_s \quad (\text{A1})$$

for  $k = 0, 1, 2, \dots, K$ , where  $K$  is the total number of vertical layers (i.e., 19 or 31 for the standard configuration). Here  $p_s$  denotes the time- and horizontal-location-dependent surface pressure (Roeckner et al. 2003). Values of parameters  $A$  and  $B$  for the L19 and L31 grid are listed in Table A1. The corresponding pressure values assuming  $p_s = 1000$  hPa (truncated at two digits after the decimal point) are listed in the same table and plotted in Fig. A1 to give the readers a direct picture of the layer distributions. Other vertical resolutions employed in this study are generated from the L31 grid using the algorithm described in section 4a(2). These grids are also plotted in Fig. A1.

## APPENDIX B

### Horizontal Diffusion in ECHAM5

As in many other models, the horizontal diffusion in ECHAM5 is expressed in the form of hyper-Laplacian



TABLE B1. Default  $e$ -folding damping time  $\tau_0$  (h) for the largest wavenumber as used in the horizontal diffusion scheme for vorticity in ECHAM5 (from Roeckner et al. 2003).

Resolution	T31	T42	T63	T85	T106	T159
$\tau_{0\text{vor}}$	12	9	7	5	3	2

$$\frac{\partial \chi}{\partial t} = -(-1)^q K_\chi \nabla^{2q} \chi, \quad (\text{B1})$$

where  $\chi$  is either vorticity, divergence, or temperature;  $q$  is a positive integer. The diffusion coefficient  $K_\chi$  is given by

$$K_\chi = \tau_{0\chi}^{-1} [n_0(n_0 + 1)a^{-2}]^{-q}. \quad (\text{B2})$$

Here  $a$  is the mean radius of the earth;  $\tau_{0\chi}$  is the empirically determined  $e$ -folding damping time of the highest resolvable wave component  $n_0$  and is independent of vertical level but changes with horizontal resolution (Roeckner et al. 2003). For historical reasons, different values of  $\tau_0$  are used for vorticity, divergence, and temperature. The values for vorticity are listed in Table B1, while those for divergence and temperature are given by

$$\tau_{0\text{div}} = 0.2\tau_{0\text{vor}}, \quad (\text{B3})$$

$$\tau_{0T} = 2.5\tau_{0\text{vor}}. \quad (\text{B4})$$

The order of diffusion  $2q$  varies with both horizontal resolution and vertical level (Table B2) and is used equally for vorticity, divergence, and temperature.

Specific values of damping time and order have been chosen based on two main concerns. First, a realistic kinetic energy spectrum should be retained in the troposphere. Very weak diffusion is desired for components with wavenumber less than 16, so as to disturb the large-scale motions as little as possible; on the other hand, relatively strong diffusion is needed for the highest resolved wavenumbers to prevent spectral blocking and remove energy from the shortest spatial scales. Therefore high-order Laplacian is used in the troposphere for the middle- and low-resolution simulations to provide strong scale-selectivity (see last row in Table B2). With more wavenumbers resolved at higher resolution, the damping order is reduced (see last row in Table B2) given that such strong contrast in the diffusion strength between large and small scales is no longer crucial.

The second consideration is wave propagation. The standard version of ECHAM5 has the center of its first vertical layer located at 10 hPa. Spurious reflection of planetary waves have been observed at the upper boundary in climate simulations when the same damp-

TABLE B2. Default order ( $2q$ ) of horizontal diffusion in ECHAM5 at various resolutions for different vertical layers (from Roeckner et al. 2003).

Vertical layer from model top	T31	T42	T63	T85	T106	T159
1	2	2	2	2	2	2
2	2	2	2	2	2	2
3	2	2	2	2	2	2
4	4	4	4	4	4	4
5	6	6	6	6	6	6
6	8	8	8	8	8	6
$\geq 7$	10	10	8	8	8	6

ing order is used for all vertical levels, and is attributed to the lack of dissipation in the upper layers. The problem is solved by enhancing the damping via second-order Laplacian at the top three vertical levels. To form a smooth transition, the order is decreased gradually from the middle and lower part of the troposphere to the model top (Table B2).

#### REFERENCES

- Boer, G. J., and B. Denis, 1997: Numerical convergence of the dynamics of a GCM. *Climate Dyn.*, **13**, 359–374.
- Boyle, J. S., 1993: Sensitivity of dynamical quantities to horizontal resolution for a climate simulation using the ECMWF (cycle 33) model. *J. Climate*, **6**, 796–815.
- Chen, M., R. B. Rood, and L. L. Takacs, 1997: Impact of a semi-Lagrangian and an Eulerian dynamical core on climate simulation. *J. Climate*, **10**, 2374–2389.
- Gates, W. L., and Coauthors, 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Amer. Meteor. Soc.*, **80**, 29–55.
- Held, I. M., and M. J. Suarez, 1994: A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *Bull. Amer. Meteor. Soc.*, **75**, 1825–1830.
- Hoskins, B. J., and A. J. Simmons, 1975: A multi-layer spectral model and the semi-implicit method. *Quart. J. Roy. Meteor. Soc.*, **101**, 637–655.
- Jablonowski, C., 1998: Test of the dynamics of two global weather prediction models of the German weather service: The Held-Suarez test (in German). M.S. thesis, Meteorological Institute of the University of Bonn, Germany, 151 pp.
- James, I. N., and P. M. James, 1989: Ultra-low-frequency variability in a simple atmospheric circulation model. *Nature*, **342**, 53–55.
- , and —, 1992: Spatial structure of ultra-low-frequency variability of the flow in a simple atmospheric circulation model. *Quart. J. Roy. Meteor. Soc.*, **118**, 1211–1233.
- James, P. M., K. Fraedrich, and I. N. James, 1994: Wave-zonal-flow interaction and ultra-low-frequency variability in a simplified global circulation model. *Quart. J. Roy. Meteor. Soc.*, **120**, 1045–1067.
- Lin, S.-J., 2004: A “vertically Lagrangian” finite-volume dynamical core for global models. *Mon. Wea. Rev.*, **132**, 2293–2307.
- Majewski, D., and Coauthors, 2002: The operational global icosahedral-hexagonal gridpoint model GME: Description and high-resolution tests. *Mon. Wea. Rev.*, **130**, 319–338.

- Müller, W., R. Blender, and K. Fraedrich, 2002: Low-frequency variability in idealised GCM experiments with circumpolar and localised storm tracks. *Nonlinear Processes Geophys.*, **9**, 37–49.
- Phillips, T. J., L. C. Corsetti, and S. L. Grotch, 1995: The impact of horizontal resolution on moist processes in the ECMWF model. *Climate Dyn.*, **11**, 85–102.
- Pope, V. D., and R. A. Stratton, 2002: The processes governing horizontal resolution sensitivity in a climate model. *Climate Dyn.*, **19**, 211–236.
- , J. A. Pamment, D. R. Jackson, and A. Slingo, 2001: The representation of water vapour and its dependence on vertical resolution in the Hadley Centre climate model. *J. Climate*, **14**, 3065–3085.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992: *Numerical Recipes in Fortran 77: The Art of Scientific Computing*. 2nd ed. Cambridge University Press, 992 pp.
- , —, —, and —, 1996: *Numerical Recipes in Fortran 90: The Art of Parallel Scientific Computing*. 2nd ed. Cambridge University Press, 500 pp.
- Ringler, T., R. Heikes, and D. Randall, 2000: Modeling the atmospheric general circulation using a spherical geodesic grid: A new class of dynamical cores. *Mon. Wea. Rev.*, **128**, 2471–2490.
- Roeckner, E., and Coauthors, 2003: The atmospheric general circulation model ECHAM5. Part I: Model description. MPI-M Tech. Rep. 349, Max Planck Institute for Meteorology, 127 pp.
- , and Coauthors, 2006: Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model. *J. Climate*, **19**, 3771–3791.
- Simmons, A. J., and D. M. Burridge, 1981: An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. *Mon. Wea. Rev.*, **109**, 758–766.
- Stratton, R. A., 1999: A high resolution AMIP integration using the Hadley Centre model HadAM2b. *Climate Dyn.*, **15**, 9–28.
- Von Storch, H., and A. Navarra, Eds., 1995: *Analysis of Climate Variability: Applications of Statistical Techniques*. Springer, 334 pp.
- , and F. W. Zwiers, 2001: *Statistical Analysis in Climate Research*. Cambridge University Press, 484 pp.
- Williamson, D. L., J. T. Kiehl, and J. J. Hack, 1995: Climate sensitivity of the NCAR Community Climate Model (CCM2) to horizontal resolution. *Climate Dyn.*, **11**, 377–397.
- , J. G. Olson, and B. A. Boville, 1998: A comparison of semi-Lagrangian and Eulerian tropical climate simulations. *Mon. Wea. Rev.*, **126**, 1001–1012.