Atmósfera (1992), **5**, *pp. 207-231*

Comparison of systematic errors in two forecast models with similar dynamical frameworks

ANANDU D. VERNEKAR, JIAYU ZHOU and BENJAMIN KIRTMAN

University of Maryland, Department of Meteorology. College Park, Maryland 20742, USA

(Manuscript received Oct. 21, 1991; accepted in final form March 17, 1992)

RESUMEN

En los modelos, los errores en la predicción exhiben las características de las aproximaciones en las simulaciones de procesos dinámicos y físicos. Los modelos son muy complejos, por lo que no siempre es posible identificar la aproximación responsable de un cierto patrón particular de error en los pronósticos. Una comparación entre el comportamiento pronosticado por los modelos puede ser valiosa para el aislamiento de la causa de los patrones de error. Aquí se realiza una comparación de los errores de pronóstico producidos por los modelos del AFGL (Air Force Geophysics Laboratory) y del COLA (Center for Ocean–Land–Atmosphere Interactions) con la esperanza de identificar sus causas. Los dos modelos se basan en idénticas aproximaciones de la simulación de los procesos dinámicos, y sólo tienen pequeñas diferencias en las parametrizaciones de los procesos físicos. Se realizan nueve ponósticos a diez días con el propósito de estudiar las características de los errores de los dos modelos. Los errores en los 500 mb de altura geopotencial son negativos en los trópicos y positivos fuera de ellos. Las temperaturas en el nivel de 850 mb son más frías que las observadas en los trópicos y más calientes que las observadas fuera de éstos. A 150 mb las temperaturas son más calientes que las observadas en los trópicos y fuera de éstos son más frías que las observadas. Estas características cualitativas de los errores no sólo son comunes a estos dos modelos sino que también lo son a los modelos de pronóstico del NMC (National Meteorological Center), GFDL (Geophysical Fluid Dynamics Laboratory) y del ECMWF (European Centre for Medium-Range Weather Forecast).

La diferencia en la estructura del error entre los dos modelos es la magnitud de éste en los trópicos. El error tropical en el Modelo AFGL es mayor que en el modelo COLA . Otra diferencia se da en el campo de humedad relativa a los 850 mb. En el modelo AFGL los errores en la humedad relativa son negativos, mayormente sobre los océanos y positiva sobre los continentes con excepciones menores. Esta estructura del error difiere del modelo COLA para el cual los errores son mayormente positivos en todas partes, excepto en algunas pequeñas regiones con errores negativos. Las diferencias mayores en las parametrizaciones físicas de los dos modelos están en la interacción de la radiación con nubes convectivas profundas, la manera en que se prescribe la temperatura de la superficie del mar (SST) y el transporte vertical de calor y humedad por convección somera. La magnitud de los errores tropicales en la altura geopotencial a 500 mb y temperatura a 850 mb, puede deberse a que el modelo AFGL no incluye la interacción de la radiación con las nubes de convección profunda. Los errores en la humedad relativa a 850 mb sobre los océanos se deben probablemente a la manera en que la SST (temperatura de la superficie del mar) se prescribe y a la falta de un transporte vertical adecuado de humedad por la parametrización de la convección somera.

ABSTRACT

Forecast errors exhibit the characteristics of approximations in simulating dynamical and physical processes in models. The models are very complex and hence it is not always possible to identify the approximations responsible for any particular error pattern in forecasts. A comparison between the models' forecast performances can be valuable in isolating the causes of error patterns. Here a comparison of forecast errors in the AFGL (Air Force Geophysics Laboratory) and COLA (Center for Ocean-Land-Atmosphere Interactions) models is made with the expectation of identifying the causes of forecast errors. The two models are based on identical approximations in simulating the dynamical processes and only minor differences in parameterizations of the physical processes. Nine ten-day forecasts are made to study the error characteristics in the two models. The errors in the 500 mb geopotential height are negative in tropics and positive in extratropics. The temperatures at 850 mb are colder than observed in tropics and warmer than observed in extratropics. At 150 mb the temperatures are warmer than

observed in tropics and colder than observed in extratropics. These qualitative error characteristics are not only common to these two models, but also to the NMC (National Meteorological Center), GFDL (Geophysical Fluid Dynamics Laboratory), and ECMWF (European Centre for Medium-Range Weather Forecast) forecast models. The difference in the error structure between the two models is the magnitude of the error in the tropics. The tropical error in the AFGL model is larger than that in the COLA model. Another difference is in the 850 mb relative humidity field. In the AFGL model, relative humidity errors are negative largely over the ocean and positive over land with minor exceptions. This error structure differs from that of the COLA model which consists of mostly positive errors everywhere with some small regions of negative errors. The major differences in the physical parameterizations between the two models are in the radiation interaction with deep convective clouds, the manner in which the sea surface temperature (SST) is prescribed and the vertical transport of heat and moisture by shallow convection. The magnitude of tropical errors in the geopotential height at 500 mb and temperature at 850 mb may be because the AFGL model does not include deep convective cloud-radiation interactions. The 850 mb relative humidity errors over oceans are probably due to the manner in which the SST is prescribed and the lack of proper vertical transport of moisture by the shallow convection parameterization.

1. Introduction

Accuracy of numerical weather prediction is limited by the approximations in the predictive system and the errors in determining the initial state of the atmosphere. In a perfect predictive system the accuracy of the forecast would deteriorate with time due to nonlinear interactions (Lorenz, 1969) and hydrodynamical instabilities (Leith, 1971) so long as there is a non-zero error in the initial state. The rate at which the forecast deteriorates depends on the growth rate of instabilities, the nature of nonlinear interactions and the amplitude and structure of the initial error. The total forecast error is the difference between the forecast and the observation and is commonly measured by the difference between the forecast and the analysis at the verification time. The total forecast error is therefore due to both the imperfections in the predictive system and the error in the initial state. An ensemble average of the total forecast errors based on a large number of forecasts made with synoptically independent initial conditions is referred to as the systematic error. For any particular day's forecast the difference between the mean square of the total error and the mean square of the systematic error is referred to as the mean square of the transient error. The systematic error is usually considered to be due to imperfections in the model; however, the systematic errors and the transient errors are not decoupled in any model.

A large number of studies have been made to identify the spatial structure of systematic errors (Fawcett, 1969; Hollingsworth *et al.*, 1980; Wallace and Woessner, 1981; Arpe and Klinker, 1986; among others). The systematic errors in models usually have coherent patterns in space. These geographically fixed error patterns largely change in amplitude with the increase in forecast time. The coherent structures of the systematic errors appear in zonally symmetric fields as well as in the zonally asymmetric fields. In a comprehensive study comparing wintertime systematic errors in the ECMWF (European Centre for Medium Range Weather Forecasts) and the GFDL (Geophysical Fluid Dynamics Laboratory) forecast models, Hollingsworth *et al.* (1980) estimated that the systematic error accounts for more than one third of the total error in medium-range forecasts. Hence, a significant improvement can be made in the forecast skill by reducing the systematic error. Assuming that the growth rate of systematic error was linear Miyakoda *et al.* (1986) suggested that the forecast can be improved by subtracting the systematic error patterns from the forecast fields. Schemm and Faller (1986) proposed a scheme to statistically correct the forecast during integration from empirical relationships between the model variables and the systematic errors. These procedures are only stop gap measures before we identify the deficiencies in the representations of dynamical and physical processes responsible for the systematic errors and improve the representations of these processes.

Bettge (1983) showed that the systematic errors in the 500 mb geopotential height forecasts from ECMWF and NMC (National Meteorological Center) models have similar patterns in that

negative error patterns are observed over the Rockies, the Alps and the Himalayan mountain ranges. He argued that the negative errors are due to the model orography because it is computed by averaging over the model grid which underestimates the height of high mountains. He considered the conservation of potential vorticity of the flow when passing over the model orography. As the vertical extent of the atmosphere is decreased over the top of mountain an anticyclonic vorticity is generated to conserve the potential vorticity. If the model mountains are shallow compared to that of the Earth, the model will not generate as much anticyclonic vorticity and that will result in the negative errors. Wallace *et al.* (1983) enhanced the model orography by adding a constant multiple of the orographic standard deviation to the mean orography. The standard deviation of orography was computed from high resolution orography data. They showed that the enhanced orography in the model reduces the systematic error in the vicinity of the mountain ranges.

Recently, when horizontal and vertical resolutions were increased, new systematic errors were identified which were absent in the low resolution models. The errors were positive in the mid-latitude zonally-averaged zonal wind. Also, there were negative errors in the high-latitudes zonally-averaged temperature especially in Northern Hemisphere winter. In the low resolution models both the surface drag over the mountains and the horizontal momentum flux convergence into the middle and high latitudes in the upper and lower stratosphere by eddies were underestimated. In these models, the simulated Northern Hemispheric westerly jet agreed well with observations but the Southern Hemispheric jet was slightly weak. When the horizontal resolution increased the effects of explicitly resolved eddies was to increase the horizontal momentum flux convergence. This increased the strength of the westerly jet in the Northern Hemisphere but the Southern Hemispheric westerly jet was in good agreement with observations. The cold temperature bias in high latitudes was consistent with the positive westerly bias to approximately maintain the thermal wind relation. One way of reducing the positive westerly bias is to increase the dissipative processes of zonal momentum in the Northern Hemisphere winter. Lilly (1972) proposed that the dissipative role of orographically induced gravity waves was important enough to include their effects explicitly in the numerical weather prediction models and the general circulation models. A large number of studies have now shown that the explicit representations of the effect of gravity wave drag in the model reduce the westerly bias in the zonal wind and the cold bias in the temperature (Palmer *et al.*, 1986; Helfand *et al.*, 1987; McFarlane, 1987; Iwasaki *et al.*, 1989; Kirtman *et al.*, 1992).

Determining the spatial structure of the systematic error field is a straightforward procedure; however, it is not obvious how to associate the error field with the physical or dynamical approximations in the model. A comparison between the results of different models as they relate to the real atmosphere can provide us with clues to isolate the causes of errors so that further improvements in the parameterizations can be made to reduce the forecast errors.

The purpose of this study is to compare the systematic errors of the AFGL (Air Force Geophysics Laboratory) and COLA (Center for Ocean-Land-Atmosphere Interaction) forecast models with the expectation of isolating the causes of errors so that they can be reduced by improved parameterization. The next section gives a brief comparison of the two models' dynamical and physical parameterizations. The third section explains the experiments and analysis procedure. The fourth section discusses the results. A summary and conclusions are presented in the fifth section.

2. Forecast models

Both the AFGL and COLA forecast models owe the numerical frame of their dynamics to the NMC spectral model of Sela (1980). Therefore, the simulation of atmospheric dynamical proces-

ses is identical. The large scale as well as subgrid scale orographic effects are treated in the same manner in both models. Both models include physical parameterizations of planetary boundary layer, shallow convection, deep moist convection, diurnal variations and cloud-radiation interactions. The only difference between the two models is the representations of these physical parameterizations. Here we shall describe briefly the treatment of dynamics and differences in the representations of the physical processes.

The dynamics of the model is based on the primitive equations. The prognostic equations prescribe the time changes of vorticity, divergence, thermodynamic energy, continuity of water vapor and natural log of surface pressure. Vertical σ-velocity and geopotential height are computed from diagnostic equations. The variables are represented by spherical harmonics with rhomboidal truncation at wave number 30. The horizontal transform grid is equally spaced in the east-west direction with 96 grid points along the latitude circle. In the north-south direction, there are 80 points on the Gaussian grid. The horizontal resolution is therefore approximately $2.25^{\circ} \times 3.75^{\circ}$ lat-long grid. Vertical coordinate is $\sigma = p/p_*$ where p_* is the surface pressure. There are 19 levels in the vertical. The levels are spaced in vertical as 1.00, 0.99, 0.973, 0.948, 0.893, 0.82, 0.735, 0.642, 0.546, 0.45, 0.40, 0.35, 0.30, 0.25, 0.20, 0.15, 0.10, 0.05, 0.00 in σ. The semi-implicit integration is used with 15 minutes time interval. The horizontal diffusion in the model is scale selective, that is, $k\nabla^4 F$, where F is a prognostic variable. The values of the diffusion constant, k, and the manner in which the diffusion term is used for the prognostic variables, differs in the two models. In the AFGL the model diffusion term is applied to all modes of the divergence but for the vorticity, temperature and specific humidity it is applied only to modes in the upper half of the rhomboid. The magnitude of the diffusion constant is the same in both cases, which is $k = 6 \times 10^{15}$ m^4 sec^{-1}. In the COLA model k is computed from $k = [\frac{a^2}{(N(N+1))}]^2/\tau$ where a is the radius of the Earth. N is the highest wave number resolved, that is, $N = 30$ in this model. τ is the dissipation time in seconds. For the divergence τ is 21 minutes and for the vorticity, temperature and specific humidity τ is 28 minutes. Hence, the value of k is 2.5×10^{16} for the divergence and 1.9×10^{16} for the vorticity, temperature and specific humidity.

The large scale orography effects are simulated by the silhouette orography. It is computed from the global 10-minute elevation data from the U. S. Navy. At any grid point the silhouette orography is computed as an average of maximum peaks in profiles of mountains in the east-west and north-south directions over the grid box. Silhouette orography enhances the topography in a similar manner as in the envelope orography but yields a smoother surface than the envelope orography especially in regions of large subgrid scale variance in a narrow zone, along the southern edge of Tibetan plateau, for example. There is a slight difference between the silhouette orographies in the two models. In the AFGL model the spectral form of the orography was smoothed to reduce the amplitude of the Gibbs oscillations over ocean. In the COLA model no smoothing was applied.

Depending on the atmospheric stability, the gravity waves induced by the subgrid scale orography propagate vertically until they are dissipated in the critical layer or as a consequence of convective and shear instabilities. These waves transport horizontal momentum upward and deposit this momentum where the waves are ultimately dissipated. The momentum is lost by the large scale where the waves are dissipated and it is brought down to the Earth's surface where it is deposited. The parameterization of these effects consists of surface wave drag and the vertical distribution of the wave drag. The surface wave drag is parameterized by following the procedure suggested by Pierhumbert (1987) and the vertical variation is parameterized according to Palmer *et al.* (1986) and Helfand *et al.* (1987).

The atmosphere exchanges momentum, heat and water vapor with the Earth's surface in

the planetary boundary layer. This exchange depends on the surface cover. Over land, the Earth may be covered by snow, ice and a variety of different plants, sands and soils. Over the ocean, the surface characteristics are determined by sea ice and spatially varying sea surface temperature. In the AFGL model, the planetary boundary over land consists of three parts: soil layer, surface layer and turbulent mixed layer. The soil layer exchanges heat and water vapor with the surface layer. The surface layer exchanges heat, momentum and water vapor with the planetary boundary layer. In the soil layer, heat is transferred by diffusion and water by diffusion and gravitational transport (Mahrt and Pan, 1984). Surface fluxes of heat, momentum and water vapor are represented by similarity theory according to Louis (1979). The turbulent mixed layer grows in height due to the surface heating and wind shear (Troen and Mahrt, 1986). The formulation is based on bulk similarity considerations. The depth of the layer is represented in terms of modified bulk Richardson number. Over oceans, the surface layer and the turbulent mixed layer parameterizations are the same as over land but, unlike over land, there is no subsurface layer.

The COLA model planetary boundary layer over land is based on similar considerations. It also consists of three layers: soil layer, surface layer and turbulent mix layer. In this model the effects of the soil layer and the vegetation cover is treated by taking into account physiological and biological processes by a biosphere model (Sellers *et al.*, 1986). In the surface layer the aerodynamic resistances for momentum, sensible and latent heat transfer between the Earth's surface and the atmosphere are based on the Monin-Obukhov similarity theory (Miyakoda and Sirutis, 1987). The turbulent mixed layer which varies in height is based on the level 2 second-order closure model of Mellor and Yamada (1982). Over oceans, the surface layer and the turbulent mixed layer parameterizations are the same as over land. There is no subsurface layer.

The precipitation due to large-scale supersaturation and subgrid scale moist convection is simulated in both the models. The precipitation occurs in a grid box if the moist air is cooled such that the relative humidity exceeds 100%. However the precipitation occurring in a higher level may not all reach the Earth's surface. The falling precipitation may evaporate if the lower model layers are dry. The deep moist convective precipitation occurs if there is a moisture flux convergence and the atmospheric column is conditionally unstable. The shallow convection predominantly occurs in undisturbed flow in absence of large scale moisture flux convergence. The trade wind cummuli under a subsidence inversion and daytime convection over land are typical examples of shallow convection. The shallow convection transfers heat, momentum and moisture from the boundary layer to the free atmosphere. For deep moist convective precipitation the AFGL model uses Kuo's scheme (Kuo, 1965, 1974) modified by Krishnamurti *et al.* (1976) whereas the COLA model uses Kuo's scheme modified by Anthes (1977). For shallow convection the AFGL model uses a semi-empirical formula derived from GATE data (Mahrt *et al.*, 1987). The COLA model implements the scheme of Tiedtke (1983) which enhances the vertical diffusion of heat and momentum if the levels near the surface are conditionally unstable.

Both models include interaction between radiation and model-generated clouds. A broad band emissivity approach in solar and longwave radiation (in presence of clouds) by Liou and Ou (1981) is implemented in the AFGL model. According to their altitude, high, middle, and low clouds can form if the model generated relative humidity exceeds a pre-selected critical value as suggested by Geleyn (1981). Middle and low clouds are considered optically black whereas high clouds are half black. Convective clouds and their interaction with radiation are not included in the AFGL model. The COLA model implements the longwave radiation scheme of Harshvardhan *et al.* (1987) and shortwave radiation scheme of Lacis and Hansen (1974). Two types of clouds, convective clouds and supersaturation clouds, are generated in the model following a procedure similar to that of Slingo (1980, 1987) and Hou (1990). The supersaturation clouds are divided into three types, high, middle and low according to their altitude. The supersaturation cloud

amounts are determined from the model generated relative humidity and vertical velocity. Deep convective cloud amount is determined from the convective rainfall rate predicted by the Kuo (1974) convective parameterization. The optical properties of clouds are determined from liquid water content and cloud temperature.

In the AFGL model the following quantities are prescribed: the sea surface temperature (SST) from FGGE data for the month of January, roughness length, surface albedo estimated snow depth and soil moisture content for the first and second layer. Surface specific humidity is set equal to zero in this study. Canopy wetness is also set equal to zero. Soil temperature at 3 meter deep is constant in time but varies from place to place.

For the COLA model the SST is prescribed from NMC data. The biosphere model computes roughness and surface albedo. But, type of vegetation cover is prescribed from climatology for the month. Soil moisture and snow cover are prescribed initially but they are predicted in the model. Sea ice is prescribed from NMC analyses.

There is also a difference in the manner in which SST is prescribed in the two models. In the AFGL model SST is prescribed at all ocean grid points without any changes. In the COLA model the SST is modified depending on the height of the $\sigma = 1$ surface over the oceans. $\sigma = 1$ level in the model is determined by the prescribed orography in the spectral form. However due to the Gibbs oscillations there are non zero height values for $\sigma = 1$ level over the oceans. These non zero values could be as much as 1/2 km above or below the sea level. The SST values over ocean are modified assuming a uniform 6.5°K/km lapse rate everywhere. For example, if the height of $\sigma = 1$ level over ocean is $\pm$ 100m the prescribed SST is modified by adding $\mp$ 0.65°K to SST analysis from observations. This modification improves the simulation of heat and moisture fluxes at the ocean surface.

3. Experiments and analysis

We have made nine medium-range forecasts with both the AFGL and COLA models. NMC analysis was used for initialization and forecast verification. The nine dates were chosen from January 1990 data. The sample size of nine appears reasonable for the study of systematic errors for medium-range forecast (Harr *et al.*, 1983). Identical nonlinear normal mode initialization was used for the two models. As mentioned in the Introduction the purpose of the study is to evaluate the performances of the two models in comparison with observation and identify the causes of deficiencies by comparing the errors in the two models. Here we have estimated the systematic errors in the two models and also estimated the statistical significance of the systematic errors so that we consider the significant error for identifying the causes.

Let x_f and x_o represent the forecast and analysis field at the verifying time respectively. Hence the error field of the variable is

$$x = x_f - x_o. \tag{1}$$

If we have m forecasts corresponding to m initial values coming from fairly uniform climatic conditions, then the ensemble (systematic) error field is, $\overline{x} = \frac{1}{m}\Sigma x$, where $m = 9$.

The forecast field error x *for any starting date* can be expressed as,

$$x = \overline{x} + x' \tag{2}$$

where x' is the transient error. The total mean square error,

$$\frac{1}{m}\Sigma x^2 = \bar{x}^2 + \frac{1}{m}\Sigma(x')^2, \tag{3}$$

is the sum of square of the systematic error and the mean square transient error. The zonal average of the error field is,

$$x_z = \frac{1}{2\pi}\int_o^{2\pi} x d\lambda \tag{4}$$

where λ is the longitude. An average of x_z over a latitude belt between ϕ_1 and ϕ_2 is,

$$\tilde{x} = \frac{\int_{\phi_1}^{\phi_2} x_z \cos\phi d\phi}{\int_{\phi_1}^{\phi_2} \cos\phi d\phi}. \tag{5}$$

In order to determine the significance of the systematic error we have computed the students "t" statistic assuming that the forecast and observed variables are independent and normally distributed with the same population variance as:

$$t = \frac{\bar{x}}{\sqrt{v + v_o}}\sqrt{m-1} \tag{6}$$

where, $v = \frac{1}{m}\Sigma(x_f - \bar{x}_f)^2$, $v_o = \frac{1}{m}\Sigma(x_o - \bar{x}_o)^2$

and $\bar{x}_o = \frac{1}{m}\Sigma x_o$, mean of the analysis.

The t-values can be tested with a pre-selected level of significance. The t-value computed has $(2m - 2)$ or 16 degrees of freedom assuming that the forecast error is normally distributed at each grid point and the nine cases selected are statistically independent.

A t-value outside the range $\mp$ 2.1 would imply that the systematic error is significantly different from zero at a grid point from a two-tailed test at the 95% level. However, one has to be careful in interpreting the significance of this test. First, all nine cases may not be statistically independent, second this test is valid only at a grid point and not for a field. Because x is highly correlated with neighboring grid points more sophisticated tests have to be employed for field significance (Livezy and Chen, 1983). The "t" test used here should therefore be used as guidance rather than as an accurate measure of statistical field significance.

4. Results

We shall begin with a comparison of root mean square errors in the 500 mb geopotential height forecasts with the AFGL and COLA models. Figure 1 shows systematic, transient and total errors in day-1 through day-10 forecasts averaged over 22°N-86°N region for the two models. The error levels and error growth rates of systematic, transient and total errors in the two models are strikingly similar. The error growth rate for the first seven days is larger than that for the last three-day period. Most likely the error in some spatial scales is saturated by day 7. The systematic error which is less than 20m for day 1 forecast grows to about 70m by day 10. The systematic error is almost half as much as the total error. This provides significant room for improving the forecast if the systematic errors are entirely due to the approximations in simulating dynamical and physical processes. The errors in 500 mb geopotential height for

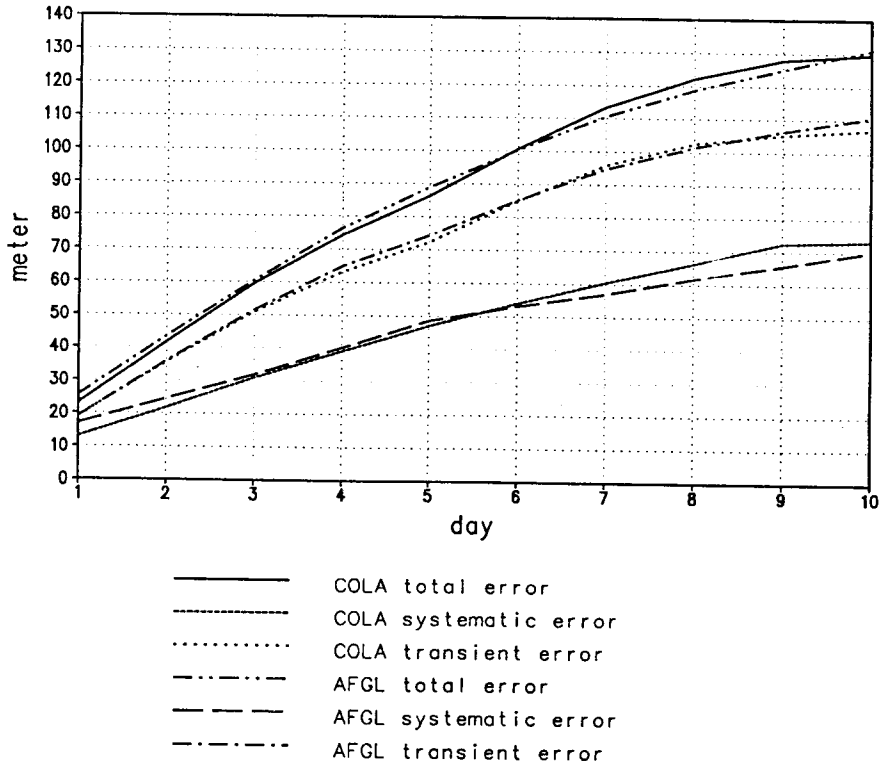

Fig. 1. The time evolution of 500 mb geopotential height errors for AFGL and COLA models averaged over extratopics (22°N-86°N). for ten days forecasts. Units: m.

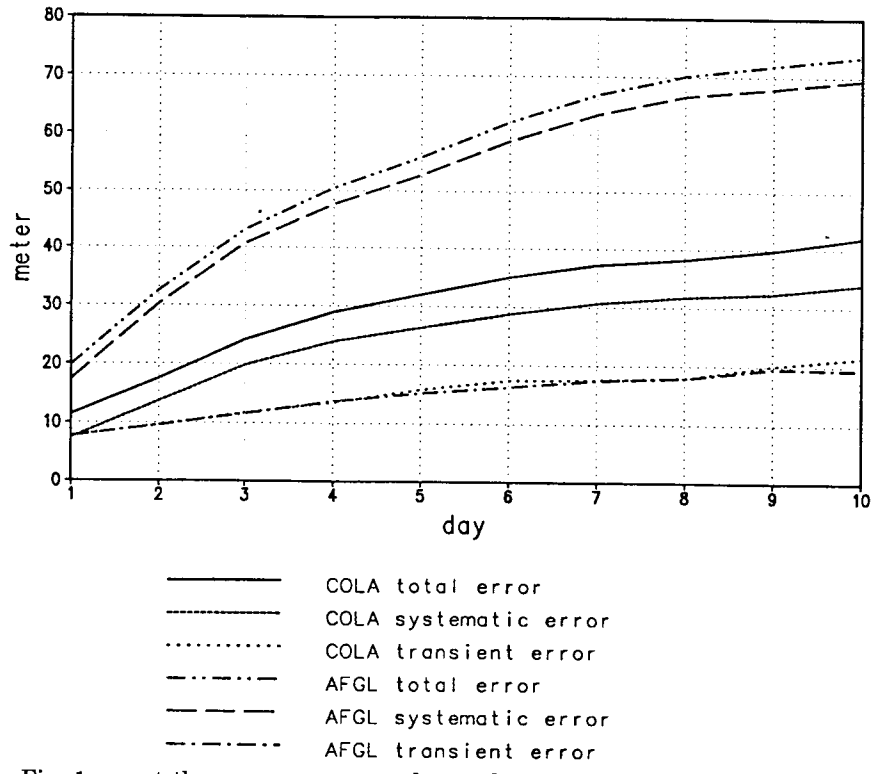

Fig. 2. Same as Fig. 1 except the errors are averaged over the tropical belt (22°N-22°S).

day 1 to day 10 forecasts averaged over the tropical belt, 22°N-22°S are shown in Figure 2. Here there is a significant difference between the systematic errors in the two models. The systematic errors for days 1 to 10 in the AFGL model is almost twice as much as in the COLA model. For day 10 the systematic errors is about 33m in the COLA model whereas in the AFGL model is close to 68m. The level of systematic errors for days 1 to 10 in tropics is about the same as in extra-tropics in the AFGL model. The transient error is about the same in the two models. In both models, the systematic error dominates the total error, in contrast to extra-tropics. Similar characteristics were noted in ECMWF forecast model (Heckley, 1985). In the tropics, error grows rapidly and saturates sooner than in extratropics (Shukla, 1981). The error growth rate for the first three days is larger than that for the remaining part of the forecast period.

The geographical structure of systematic errors of the 500 mb geopotential height for day 10 forecast is presented in Figure 3. The two models have similar error structures in that they are negative in the tropics and subropics, and positive in the polar regions except over Arctic ocean north of Eurasia. These general characteristics are also common to other models: NMC (Kanamitsu *et al.*, 1990), GFDL (Sirutis and Miyakoda, 1990) and ECMWF (Heckley, 1985).

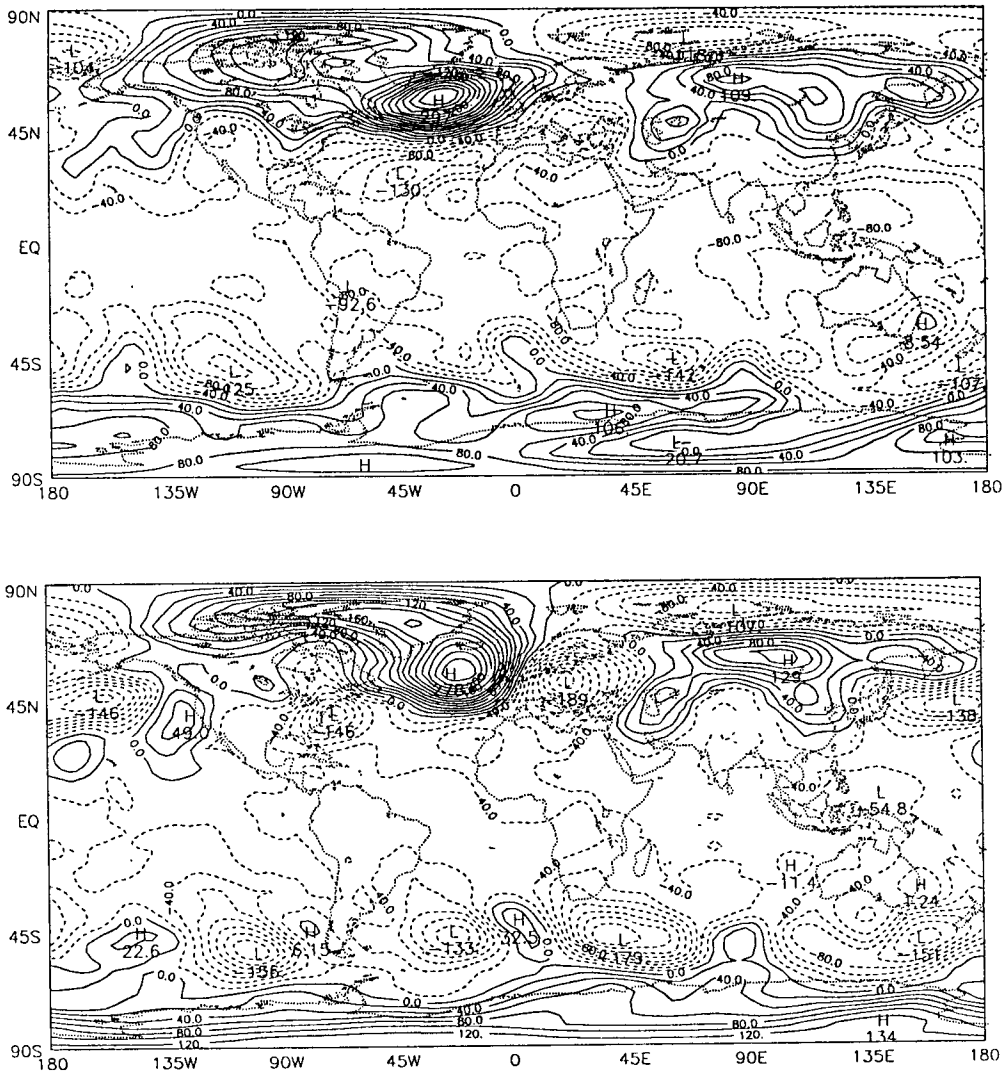

Fig. 3. The systematic errors in 500 mb geopotential height for day 10 forecast. Upper panel is for AFGL model and lower panel is for COLA model. Contour interval is 20 m.

The error character is dominated by zonal wave number four in midlatitudes and zonal wave number one in high latitudes. The phase and amplitude of the error fields of the two models are similar with minor exceptions. The magnitude of negative errors around 45°N and 45°S in the COLA model are larger than those for the AFGL model. Large tropical errors in the AFGL model seen in Figure 2 are apparent in Figure 3. Errors exceeded 80m over Indonesia, Borneo, Philippines, Southern Brazil and the Atlantic Ocean. A 60m isopleth covers a very large part of the tropics, whereas in the COLA model, the 40m isopleth covers a large region of the tropics. The errors in the tropics can be a serious problem for the predictability of extratopical flow beyond 10 days because they propagate and contaminate the predictions in the extratopics (Palmer *et al.*, 1990).The geographical distribution of the transient error in day 10 for the two models is shown in Figure 4. The transient error patterns in the two models are very similar. The error level in the extratopics is an order of magnitude larger than in the tropics. Comparing the extratropical transient error pattern with that of systematic error in Figure 3 we find the level of transient error is large where the absolute magnitude of systematic error is large. The characteristic is also common to other models, ECMWF and GFDL (Hollingsworth *et al.*, 1980). The total error for day 10 is shown in Figure 5. It reflects the combination of error structures seen in Figures 3 and 4.

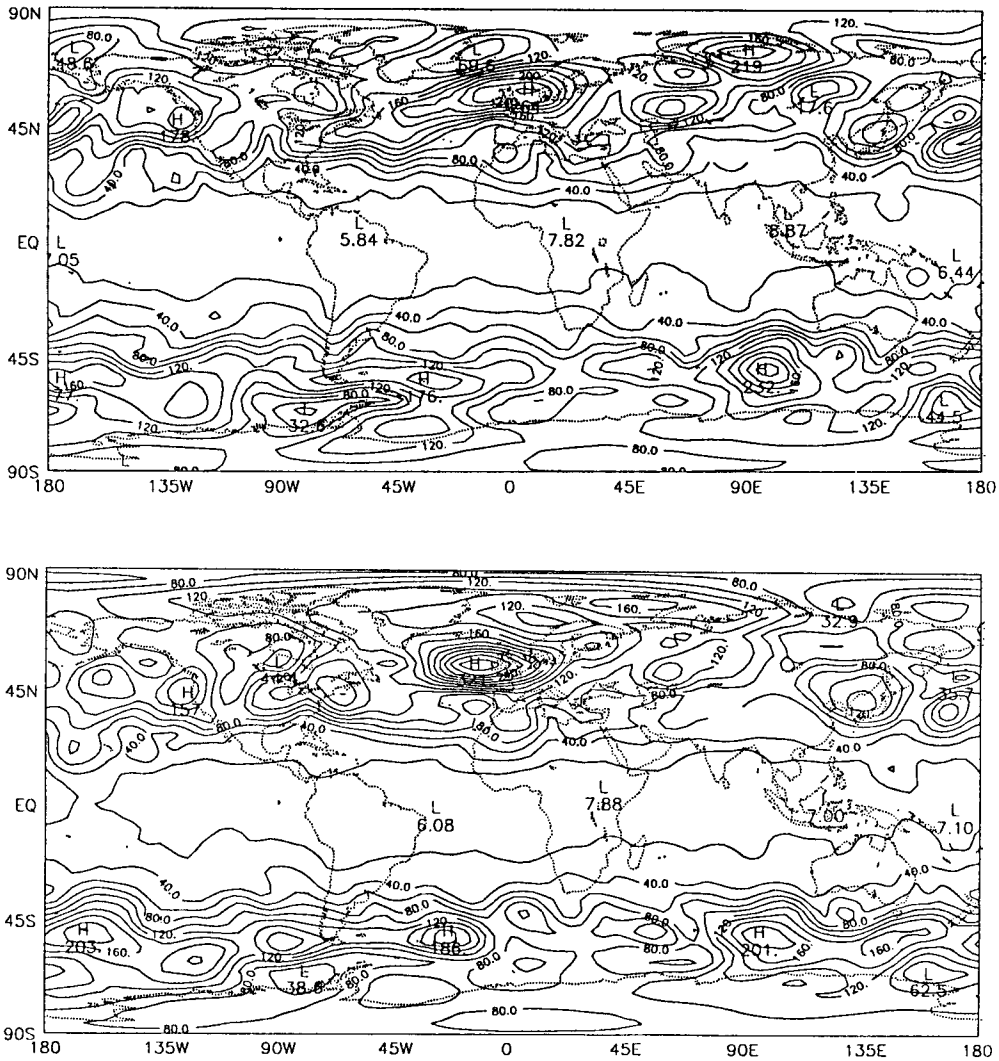

Fig. 4. Same as Fig. 3 except the transient errors.

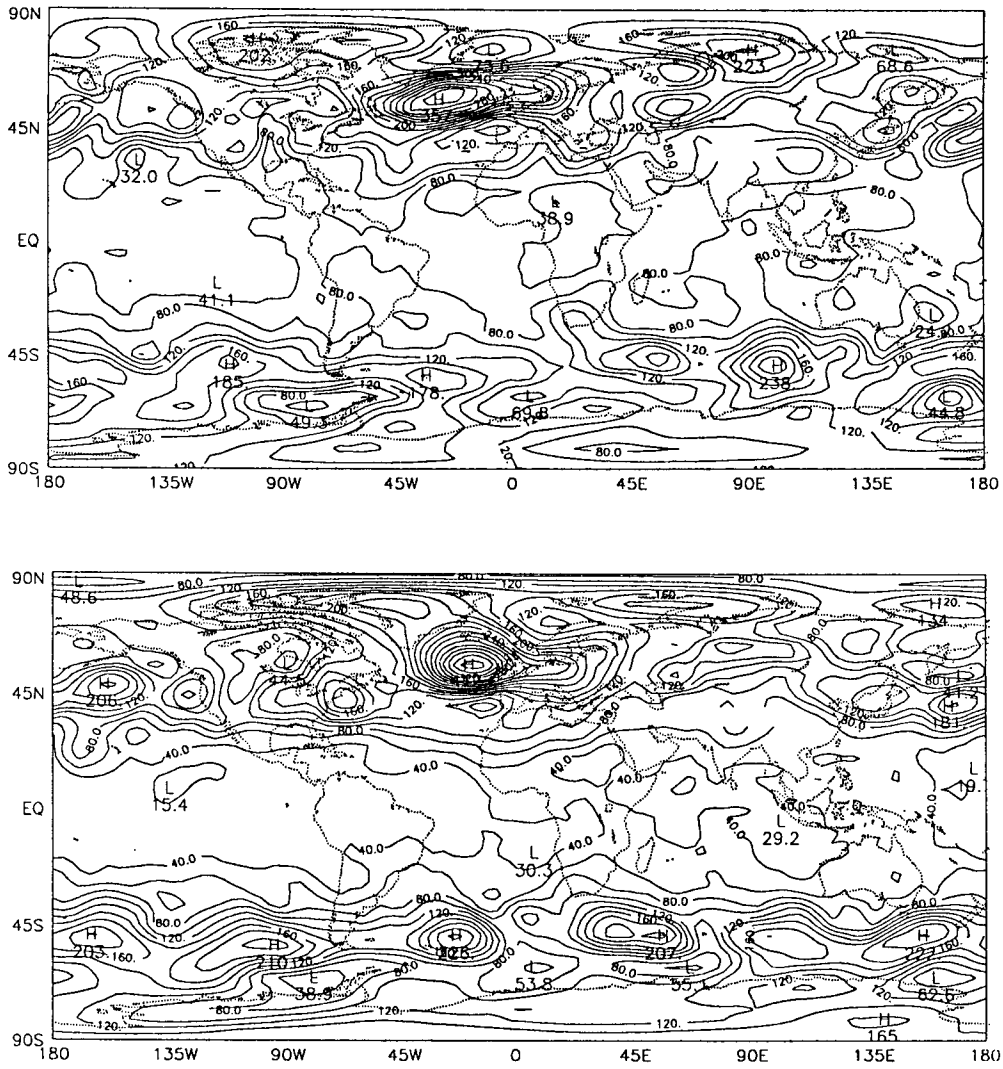

Fig. 5. Same as Fig. 3 except the total errors.

Finally the distribution of t-values for day 10, which is measured of significance of the systematic error is shown in Figure 6. The systematic error may be considered significantly different from zero in a region where absolute value of t exceeds 2.1. As noted earlier, this is not an accurate test of field significance, therefore we should interpret the results with some caution. Both models show similar characteristics, that is, the systematic error in the tropics is significantly different from zero whereas for large regions of the extratropics they are not. There are some regions over the North Atlantic ocean and Siberia where the systematic errors are probably different from zero. In tropics, the level of significance is higher for the AFGL model compared to that for the COLA model.

Next we shall examine global error statistics for 850 mb temperature averaged over 22°N-86°N (Fig. 7). The general characteristics of systematic, transient and total errors for the two models are very similar to those for the 500 mb geopotential height in Figure 1. The levels and growth rate of errors in the two models are about the same. The level of systematic error in the two models is about 1.5°K for day 1 and by day 10 it is about 3.4°K. The systematic error

for the first two days of the forecast is larger than the transient error. As in the case of 500 mb geopotential height error, the rate of error growth in 850 mb temperature for the first seven days is larger than that for the remaining part of the forecast period.

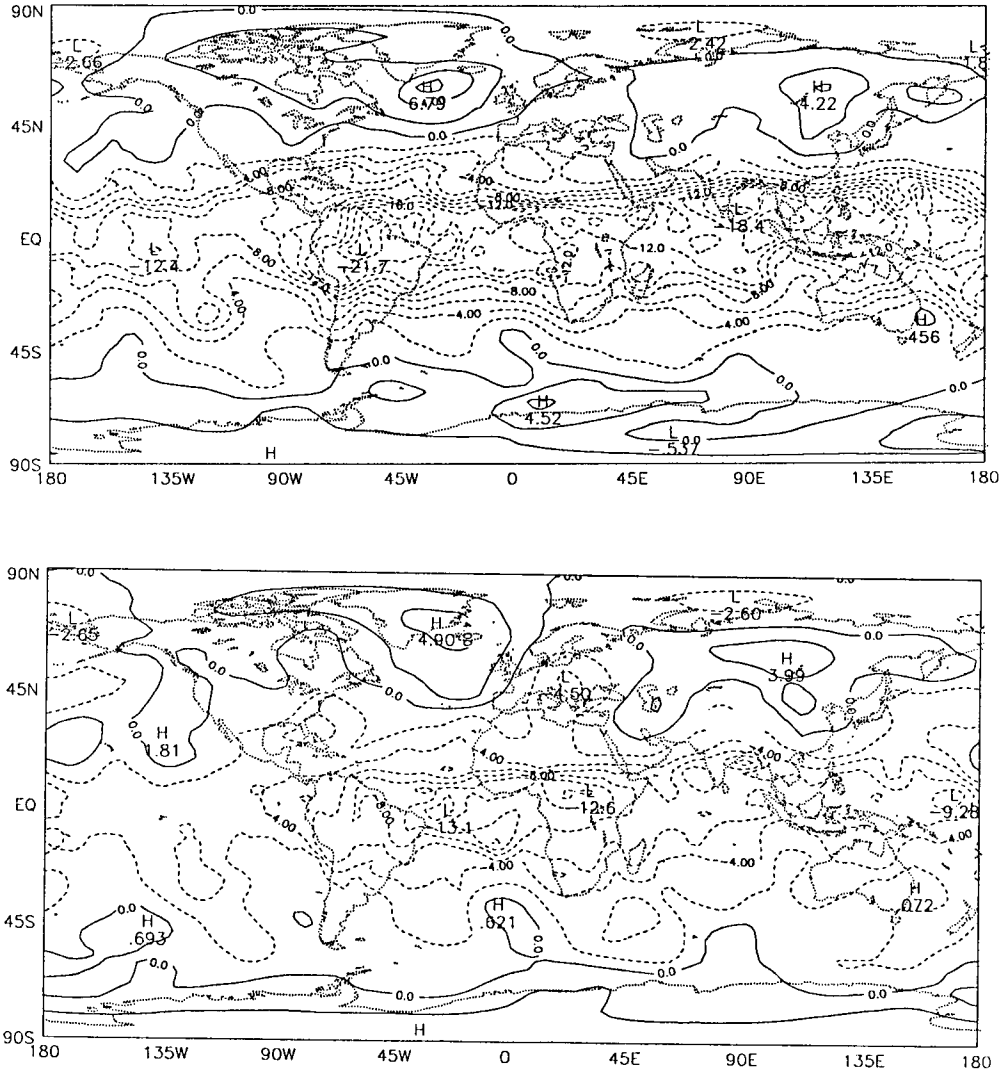

Fig. 6. Same as Fig. 3 except *t*-values. Contour interval is 2.

The 850 mb temperature error statistics for the tropical belt between 22°N and 22°S are shown in Figure 8. The transient error level in the two models are about the same, slightly lower than 1°K for day 1 and slightly higher than 1°K for 10 day. The systematic errors in the two models differ significantly. For day 1 the systematic error is about 2.8°K in the COLA model and 3.4°K in the AFGL model. For day 10 it is about 3.8°K in the COLA model whereas it is about 5.2°K in the AFGL model. The error growth rate of systematic error is large for the first two days of forecast. By day 4 it appears that the error is probably saturated in most of the spatial scales. The error characteristics in 850 mb temperature are similar to those in 500 mb geopotential height field. This is more clearly seen in Figure 9 which shows the geographical

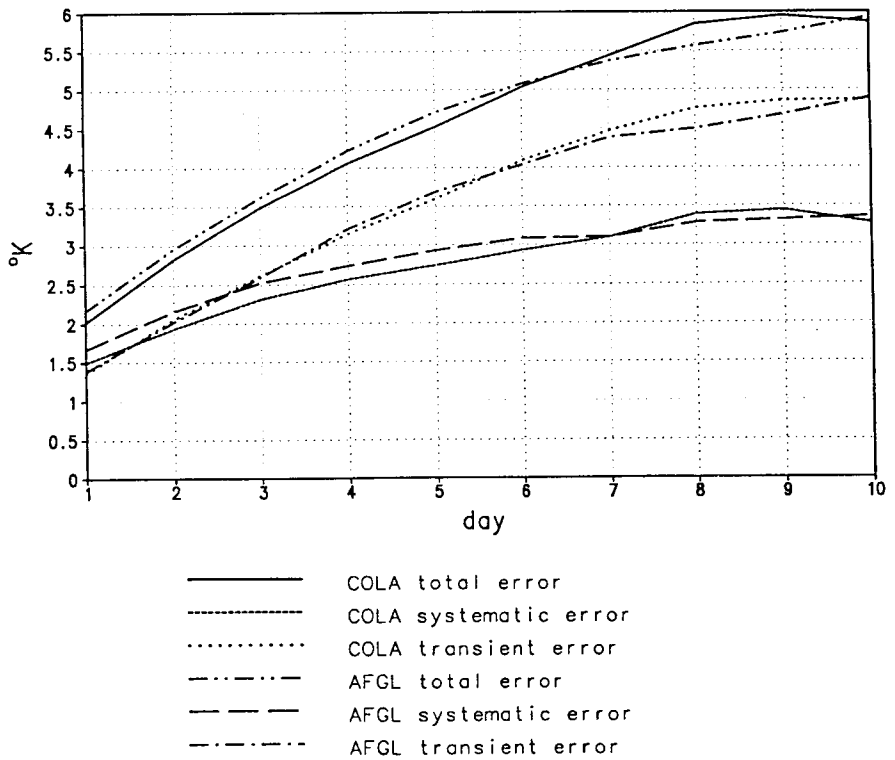

Fig. 7. The time evolution of 850 mb temperature error for AFGL and COLA models averaged over extratropics, (20°N-86°N). Units: °K.

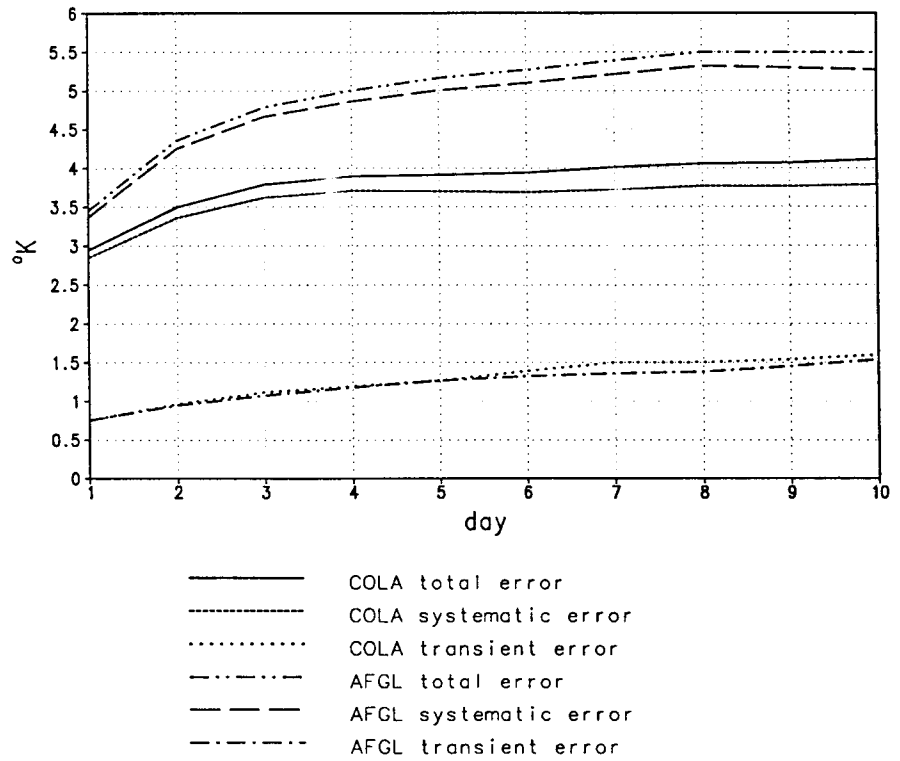

Fig. 8. Same as Fig. 7 except the errors are averaged over the tropical belt (22°N-22°S).

pattern of the systematic errors for the two models. There are negative errors in the tropics and positive errors in extratropics except over Arctic Ocean north of Eurasia. However there is a difference in the error structure between 850 mb temperature and 500 mb geopotential height. Unlike the error patterns in 500 mb height, the magnitude of 850 mb temperature error in the tropics is about the same as in the extratropics except in the Northern Territories of Canada. A 4°K isotherm covers a large region of tropical oceans in both the models, but in the AFGL model forecast there are large regions of 6°K isotherms over the Pacific and Indian Oceans and a small region over the Atlantic Ocean.

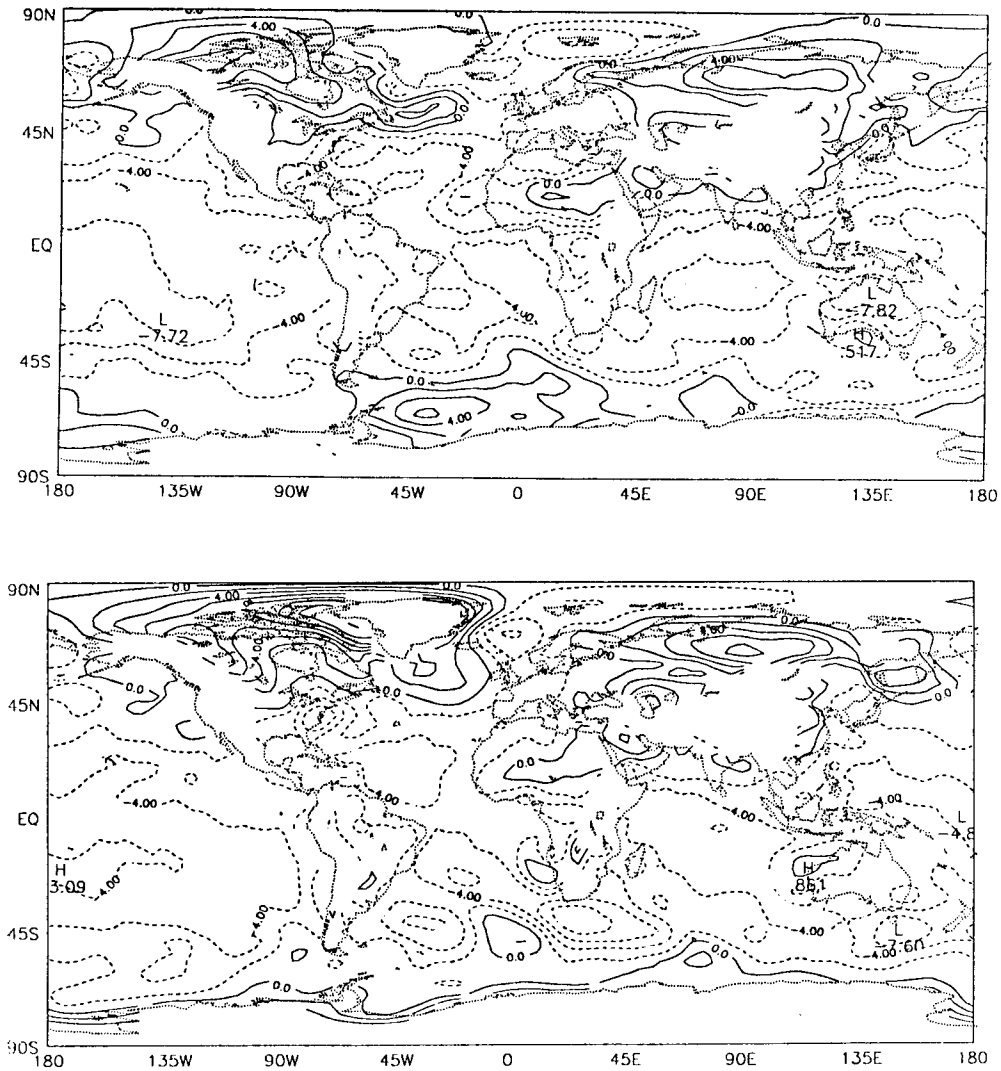

Fig. 9. The systematic errors in 850 mb temperature for day 10 forecast. Upper panel is for AFGL model and lower panel is for COLA model. Contour interval is 2°K.

Similarities in Figure 3, the 500 mb geopotential systematic error, and Figure 9, the 850 mb temperature systematic error, are expected because of the hydrostatic relation. The general pattern of these two figures, that is, too cold in the tropics and too warm in the extratopics, suggest that the baroclinic properties of the models are going to differ from that of the atmosphere because of weaker north-south temperature gradient. Some of the error structure in the

extratropics in Figure 3 may be consequence of growth rate and propagation speed of some waves in the model differing from those of the atmosphere (Lambert and Merilees, 1978). Again the characteristic is also common to other forecast models and GCMs (Kanamitsu *et al.*, 1990; Heckley, 1985; Sirutis and Miyakoda, 1990). The significance of 850 mb temperature systematic error pattern is shown in Figure 10. The errors in the tropics appear highly significant in both the models. There are also some regions in the extratropics such as over Siberia, Greenland and North Atlantic Ocean where the errors are probably different from zero.

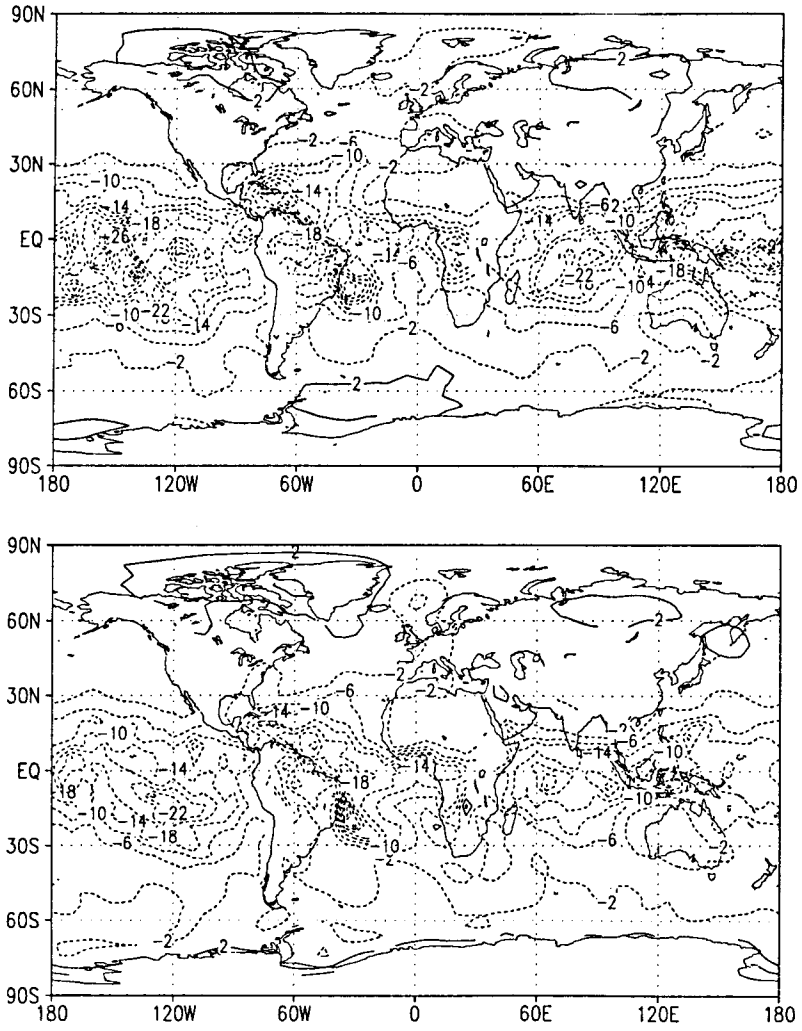

Fig. 10. Same as Fig. 9 except for t-values and contour interval is 4.

The errors in the 850 mb relative humidity averaged over the globe (86°N-86°S) for the two models is shown in Figure 11. The systematic error in the COLA model is slightly lower than that in the AFGL model for the two days forecast but then it increases and by day 10 it is about 2% higher than that in the AFGL model. The transient error in the AFGL model is larger than that in the COLA model and as a result the total error in the AFGL model is larger than that in the COLA model. The error growth rate is much higher for the first four days of the forecast compared to the remaining period, a characteristic of error growth in the tropics. Geographical

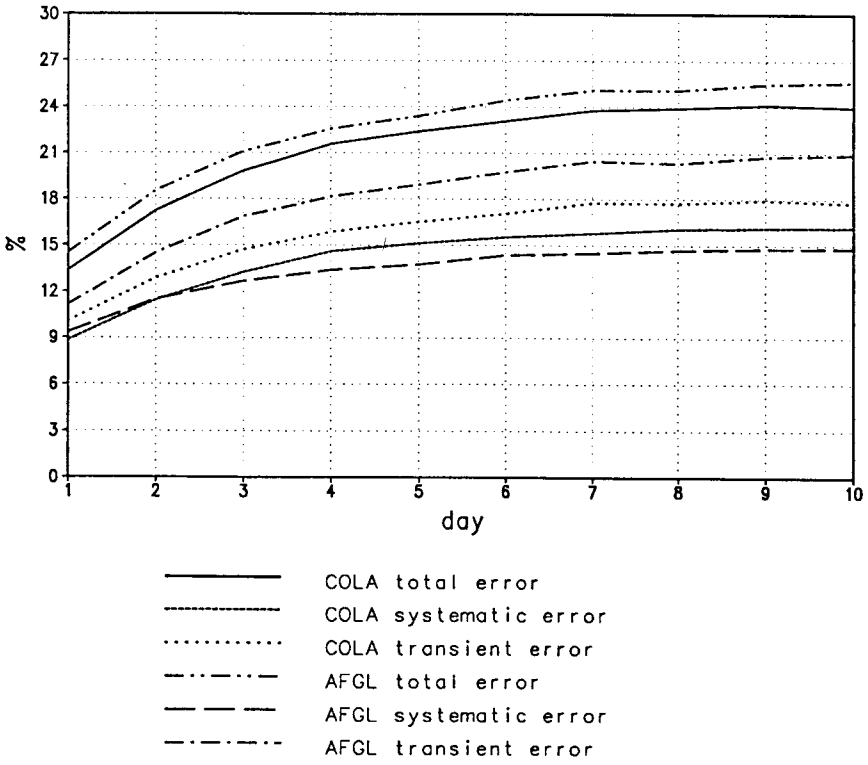

Fig. 11. The time evolution of 850 mb relative humidity errors in per cent for AFGL and COLA models averaged over the globe (86°S-86°N).

distributions of the systematic errors for day 10 in the two models are shown in Figure 12. In the AFGL model, negative errors are largely over the oceans. Over land, errors are largely positive. In the COLA model, the errors are mostly positive. If the relative humidity errors are largely due to error in the 850 mb temperature one would expect mostly positive errors in the tropics and negative errors in the extratropics because the saturation vapor pressure decreases with decreasing temperature. In the COLA model the temperature errors may be responsible for the relative humidity errors to some extent. The errors in the AFGL model may be caused by the fact that the vertical flux of moisture over oceans is not properly simulated by the boundary layer physics or by the shallow convection. A comparison of tropical relative humidity errors in the two models suggests that the AFGL model is drier than the COLA model whereas the former is colder in tropics than the latter. Hence the specific humidity negative errors in the AFGL model are probably larger because the saturation vapor pressure decreases with temperature.

Errors in relative humidity or specific humidity can have serious consequences on prediction beyond 10 days because of their influence on other physical processes such as cloud-radiation interaction and precipitation. The drier atmosphere (or negative specific humidity error) is less opaque to the infrared radiation and as a result it will cool further. Supersaturation clouds are simulated from model generated relative humidity. Positive relative humidity errors are likely to generate more clouds. Assuming that the albedo effect (cooling) of clouds dominates the greenhouse effect (warming), more clouds will cool the atmosphere. But the negative relative humidity errors will warm the atmosphere. The moisture convergence and conditional instability are necessary conditions for convective precipitation. Precipitation due to supersaturation is determined from the relative humidity. The errors in relative humidity (specific humidity) can affect the amount of precipitation and hence that of latent heat of condensation.

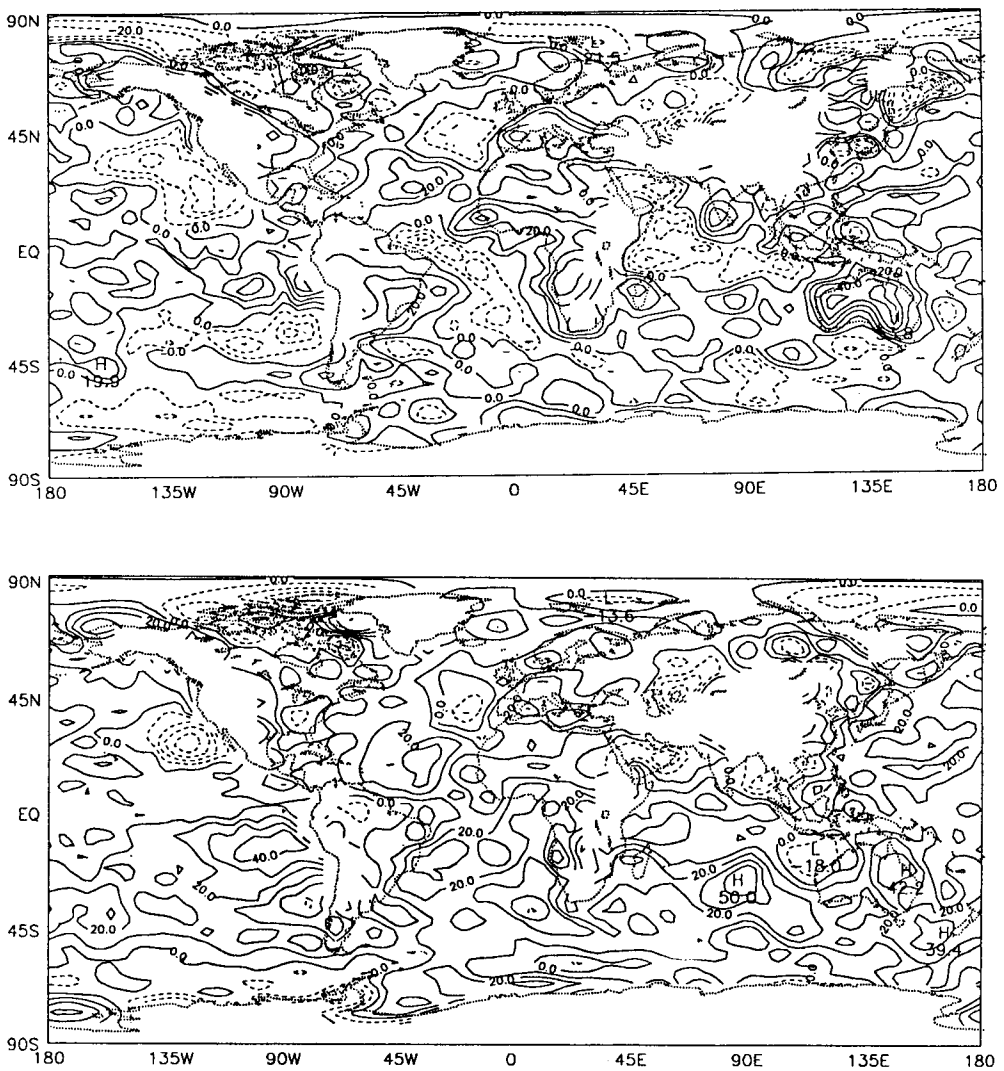

Fig. 12. The systematic errors in 850 mb relative humidity for day 10 forecast. Upper panel is for AFGL model and the lower panel is for COLA model. Contour interval is 10%.

The errors in 150 mb temperature averaged over the extratropics, 22°N-86°N, are shown in Figure 13. The systematic error in the AFGL model is about the same as that of the COLA model for the first two days of the forecast, but then it decreases and by day 10 it is about 0.4°K less than that of the COLA model. However the transient error in the AFGL model is larger than that of the COLA model. The error growth rate characteristics of 150 mb temperature are similar to that of 850 mb temperature. The growth rate is larger for the first seven days compared to that of the remaining period. The geographical patterns of the systematic error in the two models are shown in Figure 14. Extratropical error patterns in the two models are similar. They are mostly negative except over northern Eurasia. The major differences are in the tropical error. In the AFGL model they are mostly positive and large. In the COLA model they are more negative.

Now we shall consider the error structure in zonally averaged temperature and zonal winds. The systematic errors of zonally averaged temperature for the AFGL and COLA day 10 forecasts

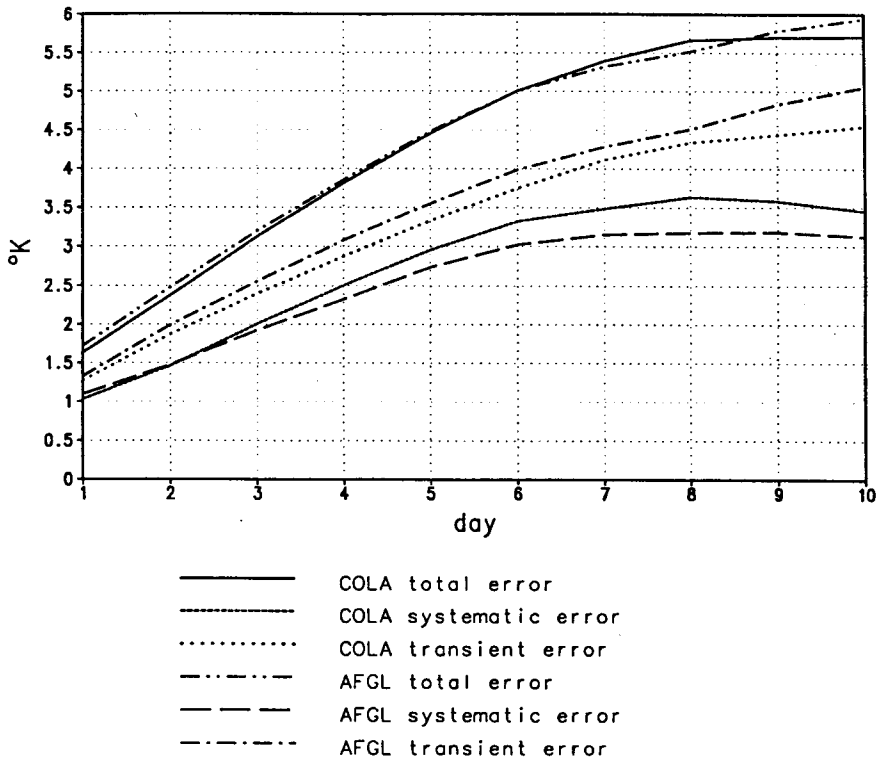

Fig. 13. The time evolution of 150 mb temperature error for AFGL and COLA models averaged over extratropics (22°N-86°N). Units: °K.

are shown respectively in Figures 15a and 15b. Both models have similar error structure in that the troposphere is cold except in high latitudes of lower troposphere and equatorial lower stratosphere. The only major difference in the two figures is the magnitude of tropical stratospheric error which is larger in the AFGL model than in the COLA model. The characteristics of error structure in 850 mb temperature in Figure 9 with cold tropics and warm extratropics can be seen in the zonally averaged error field. As mentioned earlier, such error structure will reduce the baroclinicity of the model atmosphere. The vertical error structure, too cold in lower troposphere and too warm in lower stratosphere, will reduce the convective instability of the tropical model atmosphere. However, these characteristics are not only common to these two models but also in other forecast and GCM models (Kanamitsu *et al.*, 1990; Sirutis and Miyakoda, 1990; Heckley, 1985).

The systematic error in zonally averaged zonal wind for the AFGL and COLA models is shown in Figures 16a and 16b respectively. The error structures in the two models are similar in that the errors have a barotropic structure. Both models have easterly bias in tropics and westerly bias in Southern Hemispheric subtropics and again easterly bias in high latitude of the Southern Hemsiphere. Both models have positive errors in high latitudes of the Northern Hemisphere and negative errors around 45°N. The COLA model has westerly bias in subtropics of the Northern Hemisphere but in the AFGL model the westerly bias is only in the stratosphere and in the troposphere the bias is easterly. Prior to the inclusion of explicit parameterization of gravity

wave drag the magnitude of these errors were much larger (Kirtman *et al.*, 1992). The error structure of the zonal winds are consistent with errors in the zonally averaged temperature in that they approximately satisfy the thermal wind equation.

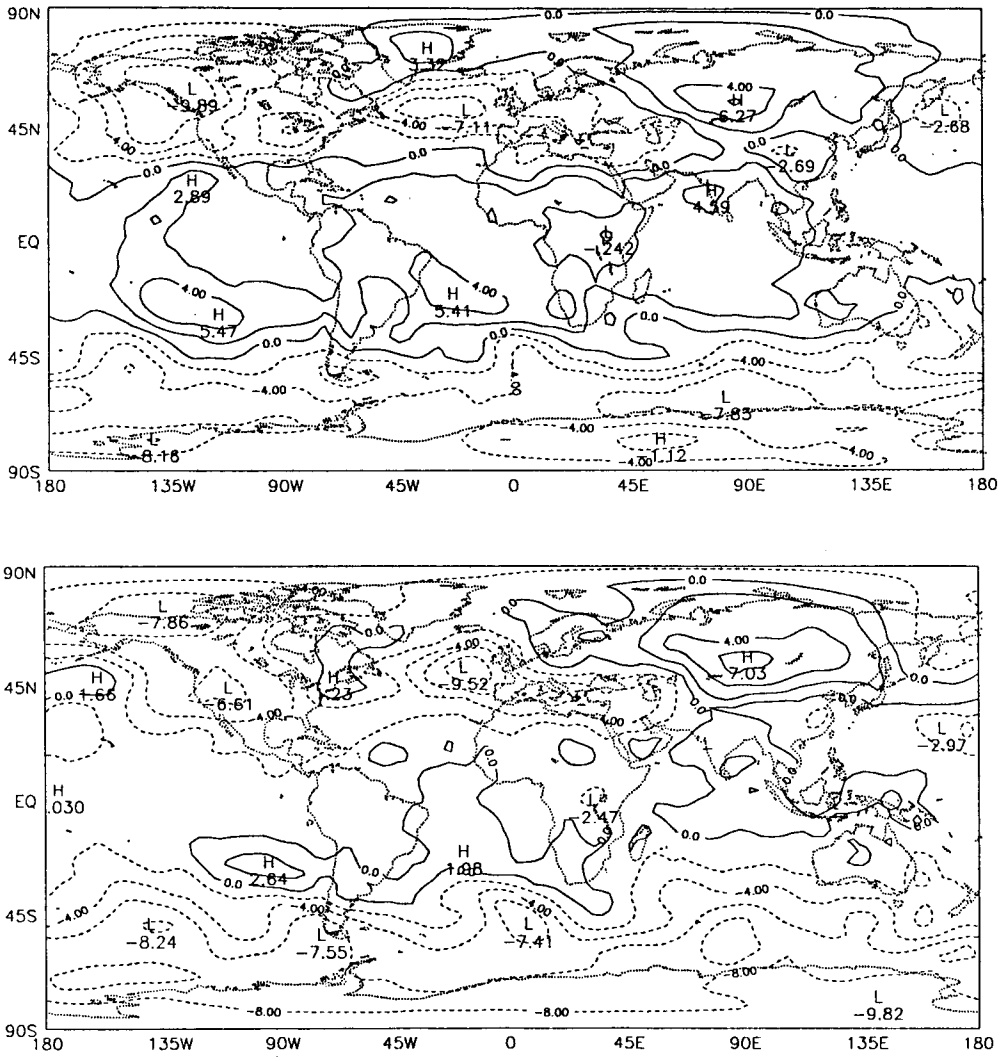

Fig. 14. The systematic errors in 150 mb temperature for day 10 forecast. Upper panel is for AFGL model and lower panel is for COLA model. Contour interval is 2°K.

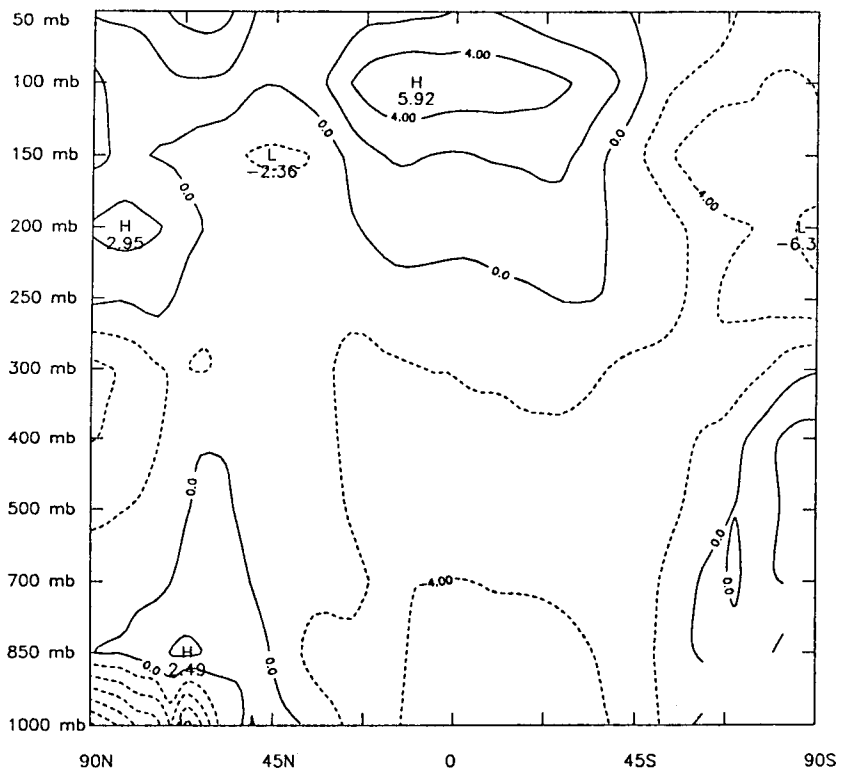

Fig. 15a. The systematic errors in zonally averaged temperature for day 10 forecast with AFGL model. Contour interval is 2°K.

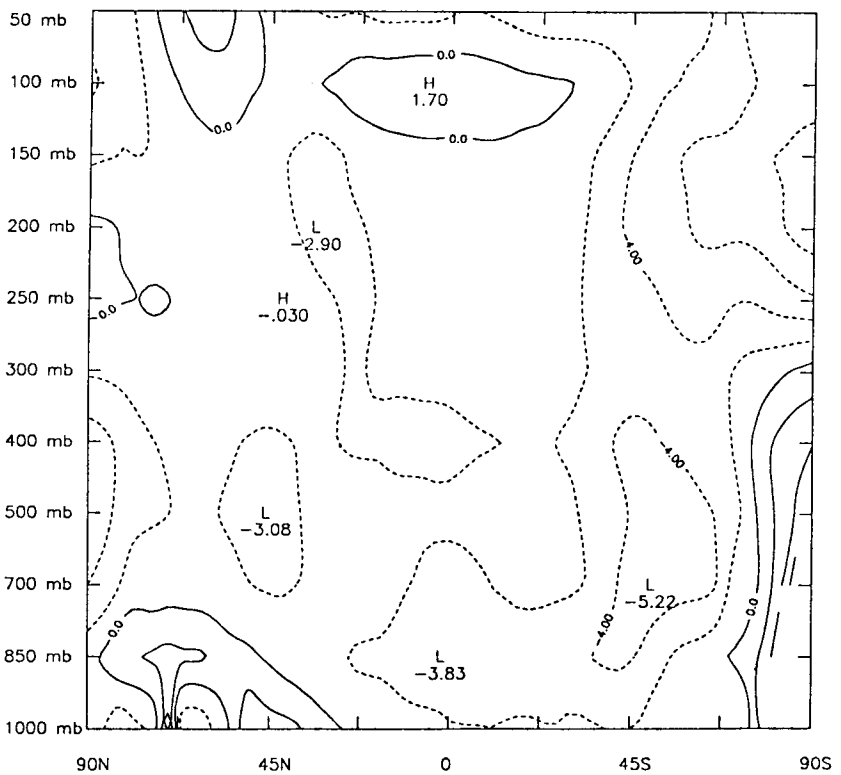

Fig. 15b. Same as Fig. 15a except for COLA model.

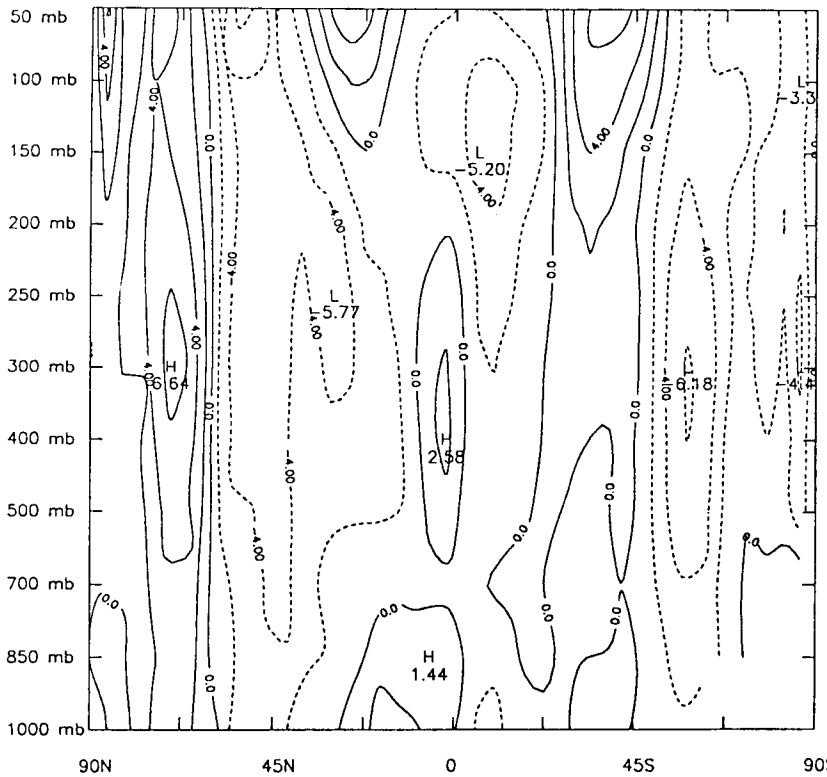

Fig. 16a. the systematic errors in zonally averaged zonal wind for day 10 forecast with AFGL model. Contour interval is 2m/sec.

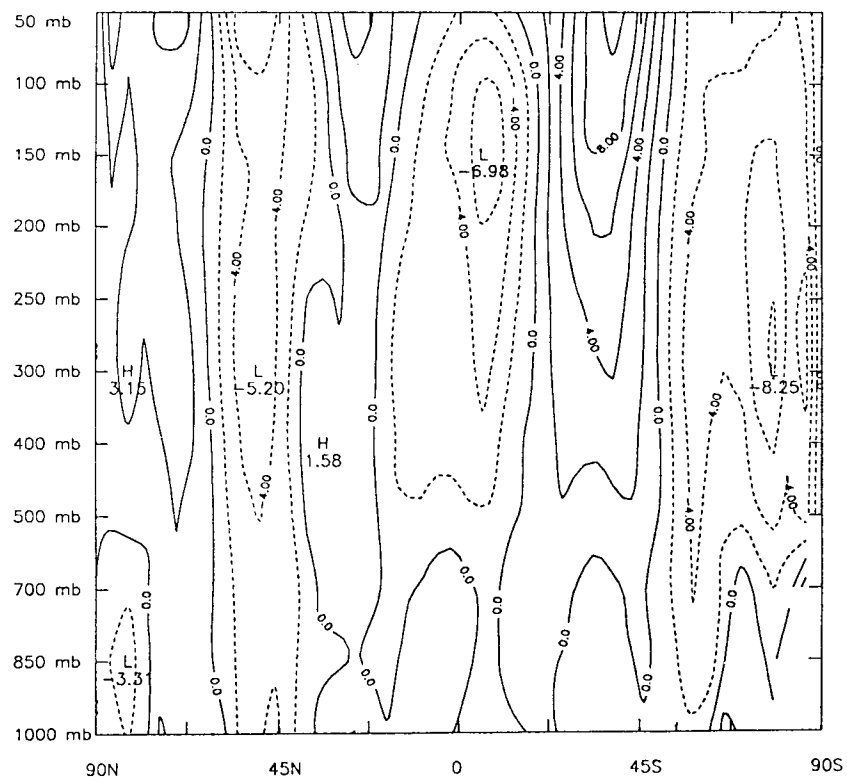

Fig. 16b. Same as Fig. 16a except for COLA model.

5. Summary and conclusions

We have made nine ten-day forecasts with the two models, AFGL and COLA, to study the systematic errors. Nine initial dates were chosen from January 1990 such that they are synoptically independent. The geographical patterns of systematic errors and their statistical significance were computed for 500 mb geopotential height, 850 mb temperature and relative humidity and 150 mb temperature. Additionally, vertical and meridional distributions of systematic errors were computed for zonally averaged temperature and zonal wind. The systematic errors in the two models have some common features. For example, the errors in 500 mb geopotential height are negative in tropics and positive in extratropics. Consistent with 500 mb height errors, the 850 mb temperatures are cold in tropics and warm in extratropics compared to NMC analysis. The zonally averaged temperatures in tropics are cold in the troposphere and warm in the lower stratosphere. In the extratropics the temperatures are warm in lower troposphere and cold in the upper troposphere compared to the observations. The systematic errors of zonally averaged zonal winds have a barotropic character. By and large, the errors are negative (easterly bias) in the tropics, middle latitudes and high latitudes of the Southern Hemisphere. The positive errors (westerly bias) appear in the subtropics and high latitudes of the Northern Hemisphere. The errors in the zonally averaged zonal wind are consistent with those of temperature errors to apporximately satisfy the thermal wind relation. Extratropical systematic errors in day 10 forecast of 500 mb geopotential height and 850 mb temperatures in both the models are larger than those in tropics but they are not all significantly different from zero because the transient errors are also large. The tropical systematic errors in both the models appear significantly different from zero. One of the differences between the error characteristics of the two models is the magnitude of error in the tropics. The AFGL model tropical error magnitudes in 500 mb geopotential height, and 850 mb and 150 mb temperature are larger than the corresponding errors in the COLA model, in some cases twice as large. There are large differences in the systematic error structure in 850 mb relative humidity between the two models. In the AFGL model negative errors are largely over ocean and positive errors are over land. In the COLA model the error are largely positive with some small regions of negative errors. In the AFGL model the relative humidity errors appear rather serious. The model's tropical tropospheric temperature has a cold bias. Because the saturation vapor pressure decreases with temperature, one would expect positive error in the relative humidity as in the COLA model. The negative errors imply that lower troposphere over tropical oceans is very dry.

There are two major differences in the manner in which the physics is treated in these models. First is that the AFGL model does not include radiation interaction with deep convective clouds. Second is the manner in which the SST is prescribed. The $\sigma = 1$ surface in the model is determined by the prescribed orography in the spectral form. Due to the Gibbs oscillations there are some non-zero height values over the ocean. The SST prescribed form observations create fictitious horizontal temperature gradients on $\sigma = 1$ surface because it is lower or higher than the sea level. This can have an adverse effect on heat and moisture fluxes between the surface and the atmosphere. To minimize this effect the SST in the COLA model is interpolated to $\sigma = 1$ surface assuming uniform 6.5°K/km lapse rate. This modification in SST is not included in the AFGL model, but its effect is to some extent minimized by the smoothed orography. The tropical temperature errors in the AFGL could be reduced by including the convective cloud-radiation interaction. Heating at the cloud base in the lower troposphere and cooling at the cloud top in the lower stratosphere will improve the temperature forecast in the tropics. Interpolating the prescribed SST to $\sigma = 1$ level may improve the error in relative humidity and temperature over oceans. The AFGL model includes a shallow convection scheme based on an empirical relation derived from the GATE data (Mahrt *et al.*, 1987). It is possible that this parameterization is not as effective as the atmosphere in transporting heat and moisture from the surface to the atmosphere. The lack of vertical transport of heat and moisture may be responsible for the large errors in 850 mb temperature and relative humidity.

Acknowledgement

We would like to thank Dr. Chien-Hsiung Yang and Doland Aiken for providing the model code and necessary data for this study. We are indebted to their kind and prompt responses to numerous inquires on the model code in the initial stage of this study. We are grateful to Dr. Yang for his suggestions to improve the clarity of the manuscript. This research was supported by Air Force Geophysics Laboratory, Hanscom Air Force Base, Massachusetts, under the contract F19628-88K0015 and NASA grant NAGW-558.

We thank Dr. David Straus for providing the input data for R30 version of the COLA model. We are grateful to Charlene Mann and Corinne Preston for diligently preparing the manuscript.

REFERENCES

Anthes, R. A., 1977. A cumulus parameterization scheme utilizing a one dimensional cloud model. *Mon. Wea. Rev.*, **105**, 270-286.

Arpe, K. and E. Klinker, 1986. Systematic errors in the ECMWF operational model in mid-latitudes. *Quart. J. Roy. Meteor. Soc.*, **112**, 181-202.

Bettge, T. N., 1983. A systematic error comparison between the ECMWF and NMC prediction models. *Mon. Wea. Rev.*, **111**, 2385-2389.

Epstein, E. S., 1988. How systematic are systematic errors? *Preprints of Eighth Conference on Numerical Weather Prediction,* February 22-26, 1988, Baltimore, MD.

Fawcett, E. B., 1969. Systematic errors in operational baroclinic prognoses at the National Meteorological Center. *Mon. Wea. Rev.*, **97**, 670-682.

Geleyn, J. F., 1981. Some diagnostics of the cloud radiation interaction on ECMWF forecating model. *ECMWF workshop on radiation and cloud-radiation-interaction in numerical modeling*, pp. 135-162.

Harr, P. A., T. L. Tsui and L. R. Brody, 1983. Identification of systematic errors in a numerical weather forecast, *Mon. Wea. Rev.*, **111**, 1219-1227.

Harshvardhan, R. Davis, D. A. Randall and T. G. Corsetti, 1987. A fast radiation parameterization for general circulation models. *J. Geophys. Res.*, **92**, 1009-1016.

Heckley, W. A., 1985. Systematic errors of ECMWF operational forecasting model in tropical regions. *Quart. J. Roy. Meteor. Soc.*, **111**, 709-738.

Helfand, H. M., J. C. Jusem, S. Pfaendtner, J. Tenenbaum and E. Kalnay, 1987. The effect of a gravity wave drag parameterization on GLA fourth order GCM forecasts. *Short and Medium-Range Numerical Weather Prediction Collection of Papers Presented at the WMO/IUGG NWP Symposium,* Tokyo, 4-8 August 1986.

Hollingsworth, A., K. Arpe, M. Tiedtke, M. Capaldo and H. Savijärvi, 1980. The performance of a medium-range forecast model in winter – Impact of physical parameterization. *Mon. Wea. Rev.*, **108**, 1736-1773.

Hoskins, B. J., 1980. Representation of the Earth topography using spherical harmonics. *Mon. Wea. Rev.*, **108**, 111-115.

Hou, Y.-T., 1990. Cloud-Radiation-Dynamics Interaction. Ph.D. Dissertation, Department of Meteorology, University of Maryland, College Park, MD.

Iwasaki, T., S. Yamada and K. Tada, 1989. A parameterization scheme of orographic gravity wave drag with two different vertical partitionings. Part I: Impact on medium-range forecasts. *J. Meteor. Soc. Japan,* **67**, 11-27.

Kanamitsu, M., K. C. Mo and E. Kalnay, 1990. Annual cycle integration of the NMC Medium-Range Forecasting (MRF) model. *Mon. Wea. Rev.*, **118**, 2543-2567.

Kirtman, B., A. Vernekar, D. DeWitt and J. Zhou, 1992. Impact of orographic gravity wave drag on extended-range forecasts with the COLA-GCM. *Atmósfera.* Accepted for publication.

Krishnamurti, T. N., M. Kanamitsu, R. Godbole, C.-B. Chang, F. Carr and J. H. Chow, 1976. Study of monsoon depression (II), Dynamical Structure. *J. Meteor. Soc. Japan*, **54**, 208-225.

Kuo, H. L., 1965. On formation and intensification of tropical cyclones through latent heat release by cumulus convection. *J. Atmos. Sci.*, **22**, 40-63.

Kuo, H. L., 1974. Further studies of the parameterization of the influence of cumulus convection on large-scale flow. *J. Atmos. Sci.*, **31**, 1232-1240.

Lacis, A. and J. E. Hansen, 1974. A parameterization of absorption of solar radiation in the Earth's atmosphere. *J. Atmos. Sci.*, **31**, 118-133.

Lambert, S. J. and P. E. Merilees, 1978. A study of planetary wave errors in a spectral numerical weather prediction model. *Atmosphere-Ocean,* **16**, 197-211.

Leith, C. E., 1971. Atmospheric predictability and two-dimensional turbulence. *J. Atmos. Sci.*, **28**, 145-161.

Lilly, D. K., 1972. Wave momentum flux: A GARP problem. *Bull. Amer. Meteor. Soc.*, **20**, 17.

Liou K.-N. and S. C. Ou, 1981. Parameterization of infrared radiative transfer in cloudy atmospheres. *J. Atmos. Sci.*, **38**, 2707-2716.

Livezy, R. and W. Y. Chen, 1983. Statistical field significance and its determination by Monte Carlo techniques. *Mon. Wea. Rev.*, **111**, 46-59.

Lorenz, E. N., 1969. The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289-307.

Louis, J.-F., 1979. A parametric model of vertical eddy fluxes in the atmosphere. *Bound. Layer Meteorol.*, **17**, 187-202.

Mahrt, L. and H.-L. Pan, 1984. A two-layer model of soil hydrology. *Bound. Layer Meteorol.*, **29**, 1-20.

Mahrt, L., H.-L Pan, P. Ruscher and C.-T. Chu, 1987. Boundary layer parameterization for a global spectral model. *AFGL-TR-87-0246. ADA 199440.*

McFarlane, N. A., 1987. The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere. *J. Atmos. Sci.*, **44**, 1775-1800.

Mellor, G. L. and T. Yamada, 1982. Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys.*, **20**, 851-875.

Miyakoda, K. and J. Sirutis, 1977. Comparative integrations of global models with various parameterization processes of subgrid scale vertical transport: Description of the parameterizations. *Beitr. Phys. Atmos.*, **50**, 445-487.

Miyakoda, K. J. Sirutis and J. Ploshay, 1986. One-month forecast experiments - without anomaly boundary forcings. *Mon. Wea. Rev.*, **114**, 2363-2401.

Palmer, T. N., G. J. Shutts and R. Swinbank, 1986. Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parameterization. *Quart. J. Roy. Meteor. Soc.*, **112**, 1001-1039.

Palmer, T. N., C. Brankovic, F. Molteni, S. Tibaldi, L. Ferranti, A. Hollingsworth, V. Cubasch and E. Klinker, 1990. The European Centre for Medium-Range Weather Forecasts (ECMWF) program on extended-range prediction. *Bull. Amer. Meteor. Soc.*, **71**, 1317-1330.

Pierrehumbert, R. T., 1987. An essay on the parameterization of orographic gravity wave drag. Geophysical Fluid Dynamics Laboratory/NOAA/Princeton University, Princeton, NJ 08542.

Schemm, J. K. E. and A. J. Faller, 1986. Statistical corrections to numerical prediction, Part IV. *Mon. Wea. Rev.*, **114**, 2402-2417.

Sela, J. G., 1980. Spectral modeling at NMC. *Mon. Wea. Rev.*, **108**, 1279-1292.

Sellers, P. J., Y. Mintz, Y. C. Sud and A. Dalcher, 1986. A simple biosphere model (SIB) for use within general circulation models *J. Atmos. Sci.*, **43**, 505-531.

Shukla J., 1981. 'Predictability of the tropical atmosphere' in report on Workshop on Tropical Meteorology and its Effects on medium Range Medium-Range Weather Prediction at Middle Latitudes. ECMWF, 11-13 March 1981, pp. 21-51.

Sirutis, J. and K. Miyakoda, 1990. Subgrid scale physics in 1-month forecasts. Part I: Experiment with four parameterization packages. *Mon. Wea. Rev.*, **118**, 1043-1064.

Slingo, J. M., 1980. A cloud parameterization scheme derived from GATE data for use with a numerical model. *Quart. J. Roy. Meteor. Soc.*, **106**, 747-770.

Slingo, J. M., 1987. The development and verification of a cloud prediction scheme for the ECMWF model. *Quart. J. Roy. Meteor. Soc.*, **113**, 899-927.

Tiedtke, M., 1983. The sensitivity of the time mean large scale flow to cumulus convection in the ECMWF model. Workshop on convection in large scale numerical models. ECMWF, 297-316.

Troen, I. and L. Mahrt, 1986. A simple model of the atmospheric boundary layer: Sensitivity to surface evaporation. *Bound. Layer Meteor.*, **37**, 129-148.

Wallace, J. M. and J. K. Woessner, 1981. An analysis of forecast error in the NMC hemispheric primitive equation model. *Mon. Wea. Rev.*, **109**, 2444-2449.

Wallace, J. M., S. Tibaldi, and A. J. Simmons, 1983. Reduction of systematic forecast errors in the ECMWF model through the introduction of envelope orography. *Quart. J. Roy. Meteor. Soc.*, **109**, 683-717.