The Effect of Forecast Error Accumulation on Four-Dimensional Data Assimilation

DAVID L. WILLIAMSON

National Center for Atmospheric Research[1], *Boulder, Colo. 80302*

(Manuscript received 6 October 1972)

ABSTRACT

Updating experiments performed with 5° and 2.5° versions of the NCAR Global Circulation Model are described. Either wind or temperature is updated. Little difference in the asymptotic error of the induced field is found between updating the 5° model with data generated by the 5° model and updating the 2.5° model with data generated by the 2.5° model. This similarity is expected since both models are shown to have similar error growth rates. When the 5° model is updated with data generated by the 2.5° model, the error in the induced field approaches a much larger asymptote. However, this asymptotic error is still less than that of randomly chosen states. The larger asymptotic error is attributed to the rapid error accumulation in a 5° model forecast when compared to data generated by the 2.5° model.

1. Introduction

Recent numerical experiments (Charney *et al.*, 1969; Jastrow and Halem, 1970; Williamson and Kasahara, 1971) have shown that four-dimensional data assimilation using simple insertion can successfully determine the wind field from the time history of temperature, or vice versa. We refer to this specific type of four-dimensional assimilation, where some variables in a numerical forecast are regularly replaced by observed values while all other variables are left unchanged, as *updating*. The updating experiments above are all idealistic in the following way: the "observed" data in each case were determined from a simulation carried out by the same model that was used for the assimilation. Thus, given correct "observed" initial data, each model is capable of producing a perfect forecast with respect to its "observed" atmosphere. We will refer to models capable of producing a perfect forecast based on some data set as *perfect* models with respect to that data set. If errors are introduced into the "observed" data, for instance by inserting temperatures while leaving the forecast values of wind unchanged, the error at any future time in the forecast is due to the error growth of the particular model, i.e., the model predictability. We refer to the *predictability error growth* of a model as the divergence of two solutions of the same model due to processes such as nonlinear energy exchanges, frictional dissipation of energy, etc.

The degree to which one field is obtained by updating another depends on this predictability error growth of the individual models used for the updating (Williamson and Dickinson, 1972). The faster the predictability error growth, the poorer the field obtained by updating. Since the real atmosphere has a non-zero error growth rate, updating with atmospheric data a forecast model which is perfect with respect to atmospheric data will not induce the exact state. The smallest asymptotic error that could be obtained in practice with real atmospheric data and a perfect forecast model is indicated by the experiments where the model being updated also generates the updated data. This limit would be produced by models whose error growth characteristics are identical to those of the atmosphere.

In the experiments with perfect models, the only source of error is the predictability error growth. We can never hope to have a perfect forecast model especially since the atmosphere is not deterministic (Lorenz, 1969). Major sources of errors in numerical forecasts are the model imperfections such as inaccurate approximations to physical processes and numerical approximations needed to solve the governing equations. These approximations produce errors every time step, which accumulate as the integration proceeds. The error in a forecast at any time is then an accumulation of all errors caused by the model approximations since the beginning of the integration and the predictability error growth. Unless there is some unique and unlikely cancellation, this accumulated forecast error will be greater than the predictability error alone. The better the forecast model, the closer this accumulated error will be to the error resulting from just the predictability error growth.

A natural question to ask is: What is the effect of this error accumulation on four-dimensional assimilation? One would expect the updating results to depend on the rate of error accumulation in the same way as they dedepend on the predictability error growth of the experi-

[1] The National Center for Atmospheric Research is sponsored by the National Science Foundation.

ments using a perfect forecast model, i.e., the faster the error accumulation, the poorer the field induced by updating. This can be seen from the following considerations. After a field is updated, the model state consists of the correct state plus some error. This error grows by predictability error growth, but in addition the imperfect forecast model produces errors every time step, which accumulate and grow as the integration proceeds to the next update time. We will refer to the accumulated error as the total error resulting from accumulation every time step and predictability growth. Eventually, we expect to reach a balance where the improvement due to the updating matches the degradation due to error accumulation between updates. Since the accumulated error is greater than that due to the predictability growth alone, we would expect the asymptotic error level of the induced field to be greater than that found in the previous experiments where the models were "perfect." We have performed a series of update experiments at NCAR to examine the importance of this error accumulation.

We noted above that there are two main sources of error in numerical forecasts in addition to predictability error growth: numerical truncation error and physical approximation error. Only numerical truncation error is considered in the following. The "observed" data are generated by a simulation with the NCAR global circulation model. These data are used to update another model with coarser grid resolution, but with the same physical approximations. This type of experiment is one step in understanding the results of updating a model with real data. Another step, not taken into account here, would be to consider only physical approximation errors. In this type of experiment, the model generating the data and the model being updated would have the same resolution and comparable numerical approximations but different physical approximations.

2. Prediction models and generation of "observed" data

Two versions of the NCAR Global Circulation Model are used for the experiments. The versions are basically the same as used by Williamson and Kasahara (1971) but with the addition of soil moisture and snow cover. The only difference in the two versions is the grid resolution. One version uses a 5° grid and the other, 2.5°. The 5° grid is a subset of the 2.5° grid, i.e., all points of the 5° grid correspond to points of the 2.5° grid. The physical bases of the two models are identical and are described in Kasahara and Washington (1967, 1971) and Washington and Kasahara (1970). The addition of soil moisture and snow cover is described by Washington (1971).

Two sets of historical data were generated by the models. The first was generated by the 2.5° version starting from an isothermal atmosphere of no motion, with fixed declination of the sun and ocean temperatures for simulation of January. The "2.5° observed" data were collected from this simulation starting on day 30. The second set, referred to as "5° observed" data, was obtained from a simulation with the 5° version of the model. This 5° control run started from day 30 of the 2.5° simulation and used the same external parameters as the 2.5° run.

3. Error accumulation and predictability experiments

We performed several error growth experiments with the 5° and 2.5° versions to determine the predictability of each version. We measure the difference between two states of the models by the root mean square (rms) error defined for a scalar field ψ by

$$\mathrm{rms}(\psi)=\Big[\sum_i (\psi_i-\psi_i^*)^2\cos\phi_i/\sum_i \cos\phi_i\Big]^{\frac{1}{2}},$$

where ϕ is geographical latitude and the sum is taken over some subset yet to be defined of grid points of the 5° grid. The starred and unstarred variables simply distinguish the two states. We refer to the *global* rms error when the sum is taken over all 5° grid points and to the *zonal* rms error when the sum is over all 5° grid points at a given latitude and height.

The global rms temperature (T) error and global rms zonal wind (u) error from the predictability experiments are shown in Figs. 1 and 2, respectively. Curve 1 in each figure shows the error of the 5° model when a 1 m sec^{-1} random error is added initially to the zonal wind velocity at 13.5 km in the 5° control case. The error is the difference between the 5° perturbed case and the 5° control case and thus represents the predictability error growth of the 5° model. Curve 2 shows the error of the 2.5° model when a 1 m sec^{-1} random error is added initially to the zonal wind velocity at 13.5 km in the 2.5° control case. Curve 3 shows the error of the 2.5° model when a very small change, less than 1%, is made in the surface pressure of the 2.5° control case while the temperature is left unchanged. Curves 2 and 3 represent the predictability error growth of the 2.5° model.

The figures indicate that the error growth rates of the 5° and 2.5° models are basically the same. Smagorinsky (1969) reports a similar agreement between the GFDL N40 (~4.5°) and N20 (~2.25°) models. The flat sections of the curves which occur for a short time during the growth are probably due to the details of the particular control case and would probably not appear in an ensemble average. The temperature and wind errors of the 5° model approach asymptotic values that are less than the corresponding asymptotic values of the 2.5° version. These asymptotic levels represent the difference between randomly chosen states of the models. The difference in asymptotic levels, which is more noticeable in the winds than in the temperatures,

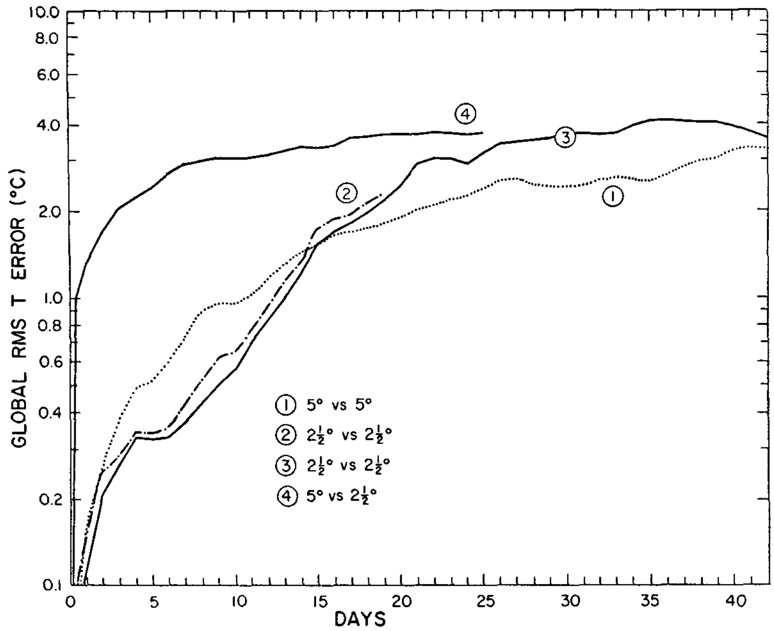

FIG. 1. Growth of the global rms temperature error.

is probably due to deficiencies of the 5° version which tend to truncate the synoptic-scale systems to some extent (Baumhefner, 1970). Baumhefner and Julian (1972) point out that the rms differences of randomly chosen states of the 5° NCAR model are less than the corresponding differences of the atmosphere.

Curve 4 in Figs. 1 and 2 shows the results of an error accumulation experiment. The 5° model started from "2.5° observed" data with a 1 m sec^{-1} random error in the zonal component of wind at 13.5 km. The curves show the difference between the 5° model forecast and the "2.5° observed" data. The difference in resolution of the two models results in relative phase and amplitude errors in the two numerical solutions. Such phase and amplitude errors are a significant part of the error in a numerical forecast of the real atmosphere, although other errors are also present. Thus, the error accumulation represented by Curve 4 is similar to what would

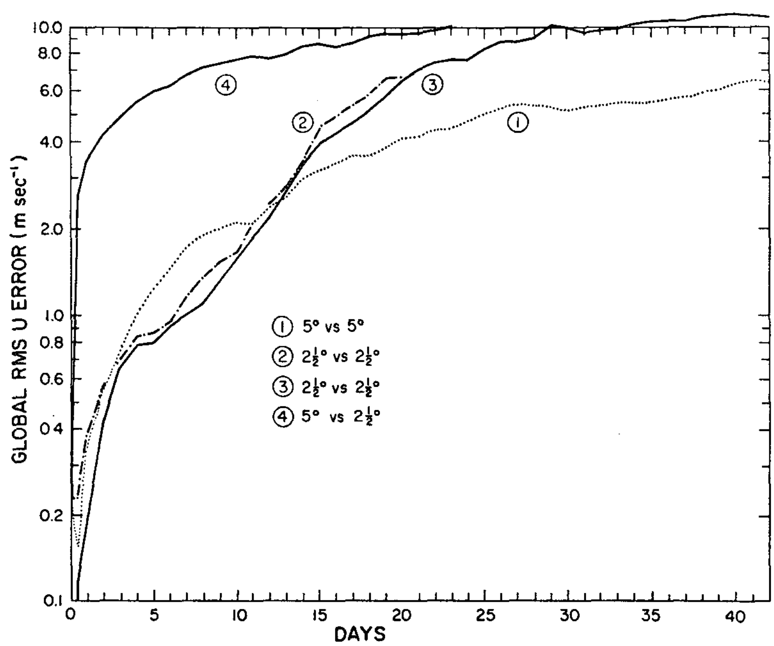

FIG. 2. Growth of the global rms zonal wind component error.

result from a numerical forecast of the real atmosphere, although the details might be different. The figures show that the error accumulation is much faster than the predictability error growth. The accumulated error after one day is comparable to the error resulting from two weeks of predictability error growth alone. The asymptotic accumulated error levels are the same as the asymptotic levels approached by the predictability growth alone.

4. Updating experiments

Results from updating the temperature and from updating the winds are shown in Figs. 3 and 4, respectively. Fig. 3 shows the global rms u-error when the temperature and surface pressure are updated everywhere every 12 hr in the different models, and Fig. 4 the global rms T-error when the winds are updated everywhere every 12 hr. Curve 1 in each figure is from the 5° model being updated with the "5° observed" data. These curves can be compared with similar curves in Williamson and Kasahara (1971).

Curve 2 in each figure is from the 2.5° model being updated with the "2.5° observed" data. The figures show that the 2.5° data updated in the 2.5° model result in the same asymptotic error levels as produced by 5° data updated in the 5° model. This result is not surprising since the two models have approximately the same error growth rate. However, the rate at which the asymptotic level is approached is different in the two cases. Temperature updating in the 5° version approaches the asymptote faster than in the 2.5° version, but wind updating in the 5° version approaches the asymptote slower than in the 2.5° version.

Curves 3 and 4 in Figs. 3 and 4 are from cases where the 5° model is updated with data from the 2.5°

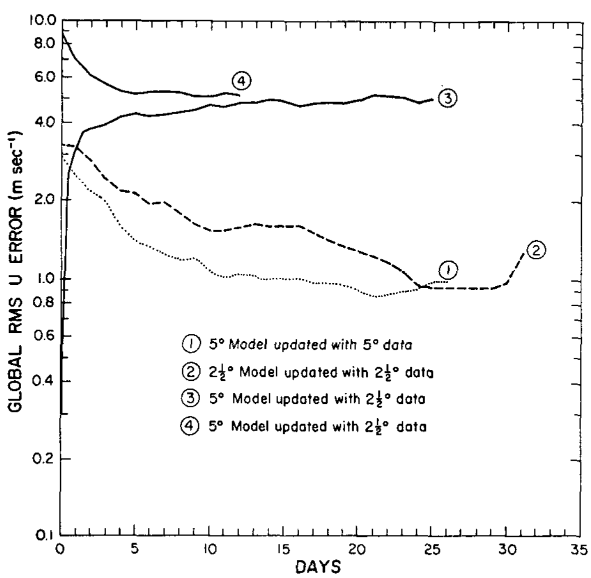

FIG. 3. Global rms wind error resulting from updating temperature and surface pressure every 12 hr.

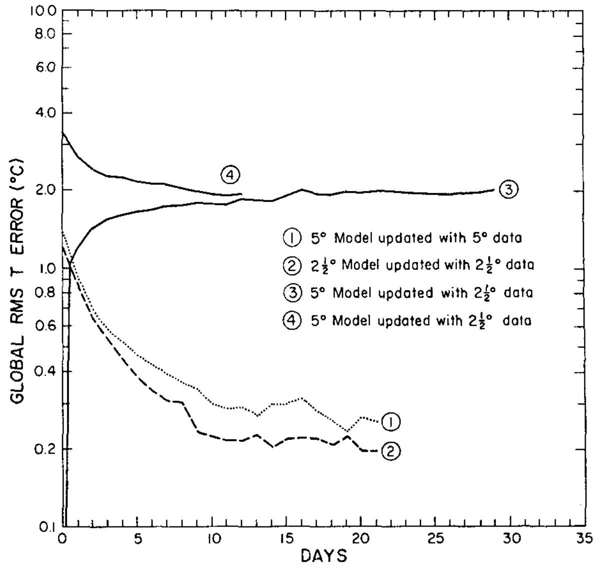

FIG. 4. Global rms temperature error resulting from updating wind every 12 hr.

simulation. As was pointed out in the preceding section, the 5° model has a rapid error accumulation when compared to the "2.5° observed" data. As seen in the figure, this accumulation results in a much larger asymptotic error from updating than is obtained when the model is updated with data produced by the same model. However, updating the 5° model with 2.5° data does result in an improvement over a randomly chosen state as can be seen by comparing Figs. 3 and 4 with 1 and 2.

For the cases of the 5° model updated with 2.5° data we expanded the temperature error and the zonal wind component error into spherical harmonics of the form

$$\psi(\lambda,\theta)=\sum_{m=0}^{M}\sum_{n=m}^{N}(A_n^m\cos m\lambda+B_n^m\sin m\lambda)P_n^m(\sin\phi).$$

The coefficients A_n^m and B_n^m were computed using the method of weight functions to orthogonalyze the associated Legendre polynomials over the discrete grid (Ellsaesser, 1966). The data were interpolated to a 2.5° grid in latitude to obtain expansion coefficients equivalent to the 5° model grid. The results are squared, summed over m, and plotted as a function of the two-dimensional wavenumber n as suggested by Baer (1972):

$$\psi_n{}^2=(A_n^0)^2+\sum_{m=1}^{n}\tfrac{1}{2}[(A_n^m)^2+(B_n^m)^2].$$

A sum over n of $\psi_n{}^2$ then equals the grid point average of the square of the error weighted by the weight function used in the expansion.

The results, averaged from day 14 to 21, for the zonal wind error at 4.5 km resulting from updating the 5° model with "2.5° observed" temperatures are shown by the crosses in Fig. 5. The wind error averaged for the same period from the error accumulation experiment (5° vs 2.5°) is shown by the dots for reference. Fig. 6 shows the temperature error at 4.5 km resulting from updating the 5° model with "2.5° observed" winds for the same period by crosses and the corresponding error from the error accumulation experiment (5° vs 2.5°) by dots. The difference in each figure between the updating experiment and the accumulation experiment represents the improvement resulting from updating as a function of scale.

Fig. 5 shows that temperature updating induces the wind field best for the large scales (small wavenumber). This scale dependence agrees with geostrophic adjustment theory. However, Fig. 6 shows that wind updating induces the temperature field best for the intermediate scales rather than the smallest scales as expected from geostrophic adjustment theory. This apparent discrepancy is probably caused by large errors in the numerical models when forecasting the smallest resolvable scales in the model. Thus, the error accumulation destroys the improvement from updating more rapidly for the smallest scales.

Fig. 7 illustrates the latitudinal dependence of the results of updating the 5° model with 2.5° data. Parts (a) and (b) are from updating temperature and parts (c) and (d) from updating winds. Curves 1 and 2 are from the updating experiments, and curve 3 is from the error accumulation experiment for reference. The improvement resulting from updating is determined by comparing the asymptotic levels approached by the errors. Figs. 7a and 7b are examples that the updating of temperature improves the wind field more at mid-latitudes than at low latitudes. This relationship is expected from geostrophic adjustment theory. Geostrophic adjustment theory also predicts that wind updating would improve the temperature field more

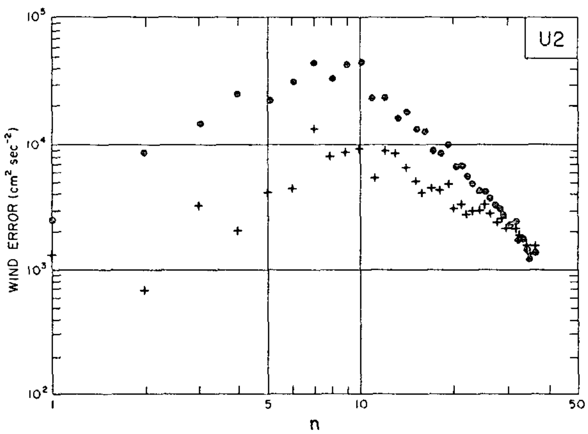

FIG. 5. Spherical harmonic expansion of the zonal wind error at 4.5 km averaged for days 14 to 21. The crosses are from updating the 5° model with 2.5° temperature data. The dots are from the error accumulation experiment (5° model vs 2.5° data).

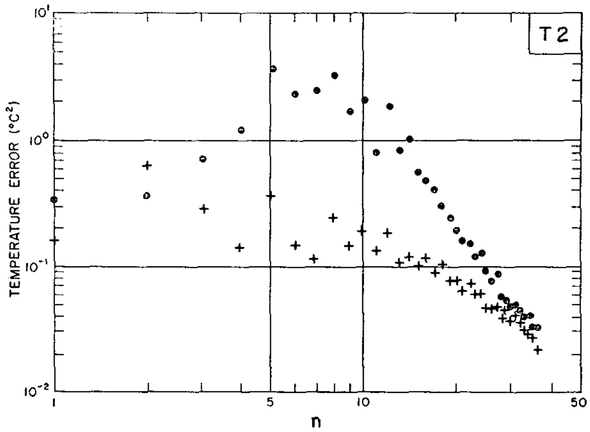

FIG. 6. As in Fig. 5 except for the temperature error.

at low latitudes than mid-latitudes. As is seen in Figs. 7c and 7d, this is not the case since the temperature is improved more at mid-latitudes. In addition, at the equator the temperature error resulting from wind updating is greater than that produced by just error accumulation in the model. This unexpected behavior was also present at other levels at the equator and to some extent at 15S, but not 15N. Williamson and Kasahara (1971) also noted in their experiments that the wind updating produced results whose latitude dependence did not correspond to adjustment theory.

5. Conclusions

A series of experiments was performed to investigate the effect of grid resolution in numerical models and error accumulation on four-dimensional data assimilation. Experiments were run updating a 2.5° NCAR global circulation model with data from a simulation using the same 2.5° model, and were also run updating a 5° model with data from a simulation using the 5° model. In one set of experiments temperature and surface pressure were updated every 12 hr; in another, winds were updated every 12 hr. The asymptotic error levels resulting from the two sets of update experiments showed little difference between the 5° and 2.5° cases. This similarity is not surprising since the error growth rate of the two models is similar. In these experiments a perfect forecast model was used, i.e., the only source of error in a forecast is growth of initial errors since the model being updated also produced the data used for the updating.

Another series of experiments was performed to investigate updating when the model being updated did not produce the updated data, which will be the case in actual practice. In this series, the 5° model was updated using data generated by the 2.5° model so that truncation error provides an additional source of forecast error. The error accumulation of a forecast with the 5° model compared to the data from the 2.5° model is much greater than just the predictability error

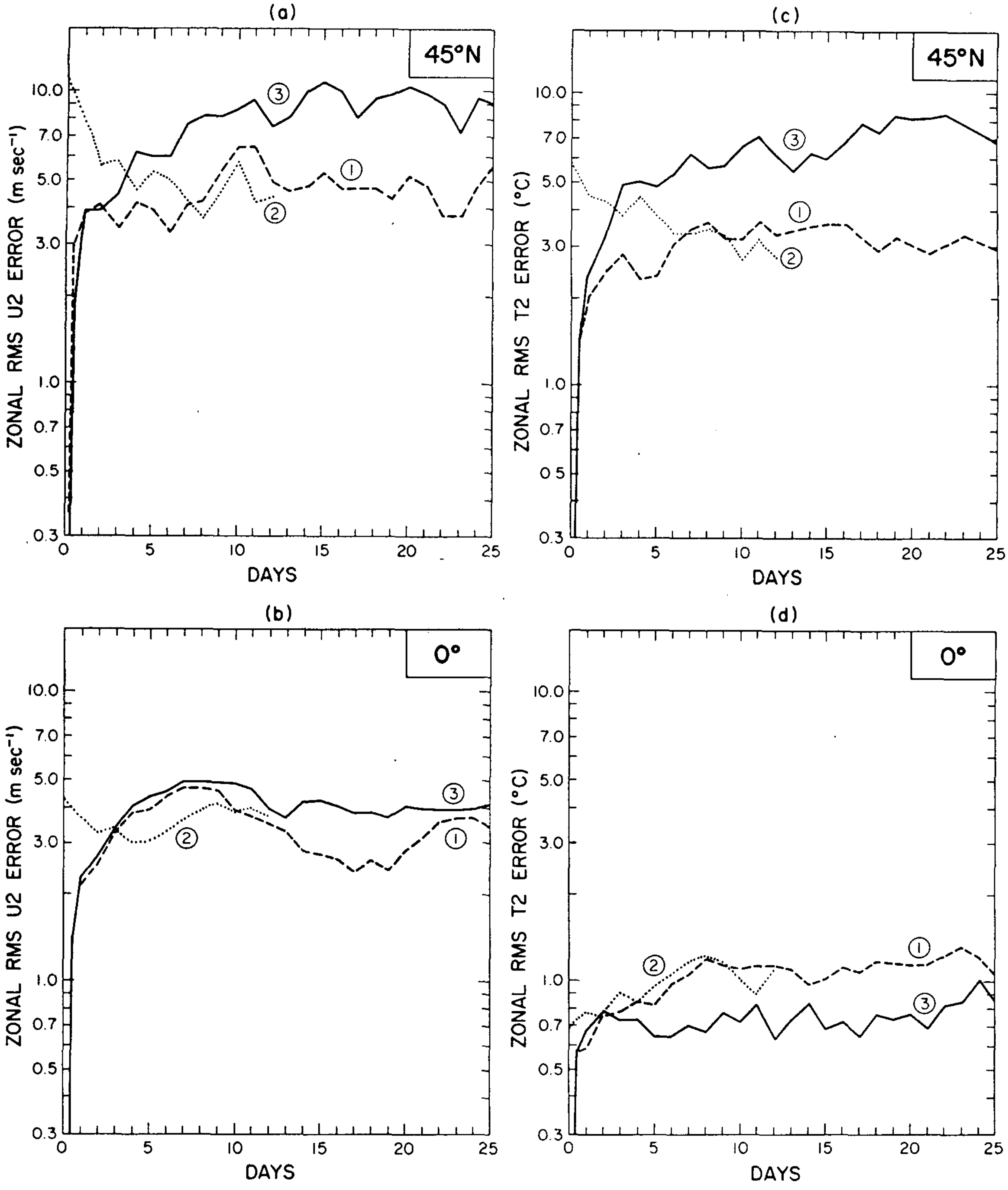

FIG. 7. Zonal rms wind (a and b) and temperature (c and d) errors. Curves 1 and 2 are from updating the 5° model with 2.5° temperature (a and b) or 2.5° wind (c and d) data. Curve 3 is from the error accumulation experiment (5° model vs 2.5° data).

growth alone of the 5° model. This accumulation results in a much larger (factor of 5–10) asymptotic error in a field induced by updating than is obtained when the model is updated with data produced by the same model. However, updating the 5° model with 2.5° data does result in an induced atmospheric state which is a slight improvement (factor of 2) over a randomly chosen one.

These results indicate in a qualitative way the effect of updating real atmospheric observations into a numerical model by simple insertion. The smallest asymptotic error that could be obtained in practice with real atmospheric data and a perfect forecast model is indicated by the experiments where the model being updated also generates the updated data. In these cases the only source of error is the predictability error growth. This asymptotic error limit would be produced by models whose error growth characteristics are identical to those of the atmosphere.

In reality, any model will produce errors every time step due to the approximations, both numerical and physical, in the model. These errors, combined with the predictability error growth, will accumulate faster than just the error growth alone. The experiments where the 5° model is updated with data generated by the 2.5° model approximate this situation. The experiments indicate that updating forecasting models by simple insertion of real atmsopheric data will result in much larger asymptotic errors than the ideal case of a perfect forecast model.

These results do not imply that updating with atmospheric data *must* result in large asymptotic errors. One obvious method to improve the results of updating is to improve the numerical model used. The closer the model approximates the atmosphere, the better the results of updating should be. The method of updating used in these experiments was simple insertion of the updated values, leaving all other fields unchanged. More refined procedures involving, in some way, the dynamical relations between fields, such as locally balanced winds and pressures, might also reduce the asymptotic error levels. Work should continue in both these areas to achieve successful four-dimensional data assimilation with real atmospheric data.

Acknowledgments. I would like to thank David Kennison, Judith Slater and Gloria Williamson for running the general circulation model (GCM) and making the modifications to the GCM codes for the present experiments. I am very grateful to Robert Dickinson and Akira Kasahara for many comments which improved the clarity of the exposition.

REFERENCES

Baer, F., 1972: An alternate scale representation of atmospheric energy spectra. *J. Atmos. Sci.*, **29**, 649–664.

Baumhefner, D. P., 1970: Global real-data forecasts with the NCAR two-layer general circulation model. *Mon. Wea. Rev.*, **98**, 92–99.

——, and P. R. Julian, 1972: The reference-level problem: Its location and use in numerical weather predictions. *J. Atmos. Sci.*, **29**, 285–299.

Charney, J., M. Halem and R. Jastrow, 1969: Use of incomplete historical data to infer the present state of the atmosphere. *J. Atmos. Sci.*, **26**, 1160–1163.

Ellsaesser, H. W., 1966: Expansion of hemispheric meteorological data in antisymmetric surface spherical harmonic (Laplace) series. *J. Appl. Meteor.*, **5**, 263–276.

Jastrow, R., and M. Halem, 1970: Simulation studies related to GARP. *Bull. Amer. Meteor. Soc.*, **51**, 490–513.

Kasahara, A., and W. M. Washington, 1967: NCAR global general circulation model of the atmosphere. *Mon. Wea. Rev.*, **95**, 389–402.

——, and ——, 1971: General circulation experiments with a six-layer NCAR model including orography, cloudiness, and surface temperature calculations. *J. Atmos. Sci.*, **28**, 657–701.

Lorenz, E. N., 1969: Three approaches to atmospheric predictability. *Bull. Amer. Meteor. Soc.*, **50**, 345–349.

Smagorinsky, J., 1969: Problems and promises of deterministic extended range forecasting. *Bull. Amer. Meteor. Soc.*, **50**, 286–311.

Washington, W. M., 1971: On the role of radiation in dynamical climate simulation and numerical weather prediction. *Proc. Miami Workshop on Remote Sensing*, Atlantic Oceanogr. Meteor. Labs., NOAA, 39–67.

——, and A. Kasahara, 1970: A January simulation experiment with the two-layer version of the NCAR global circulation model. *Mon. Wea. Rev.*, **98**, 559–580.

Williamson, D. L., and A. Kasahara, 1971: Adaptation of meteorological variables forced by updating. *J. Atmos. Sci.*, **28**, 1313–1324.

——, and R. E. Dickinson, 1972: Periodic updating of meteorological variables. *J. Atmos. Sci.*, **29**, 190–193.