

## The Second-Generation Global Forecast System at the Central Weather Bureau in Taiwan

CHI-SANN LIOU,\* JEN-HER CHEN, CHUEN-TEYR TERNG, FENG-JU WANG, CHIN-TZU FONG, THOMAS E. ROSMOND,\* HONG-CHI KUO,+ CHIH-HUI SHIAO, AND MING-DEAN CHENG

*Central Weather Bureau, Taipei, Taiwan*

(Manuscript received 23 October 1996, in final form 2 April 1997)

### ABSTRACT

The global forecast system (GFS), which started its operation in 1988 at the Central Weather Bureau in Taiwan, has been upgraded to incorporate better numerical methods and more complete parameterization schemes. The second-generation GFS uses multivariate optimum interpolation analysis and incremental nonlinear normal-mode initialization to initialize the forecast model. The forecast model is a global primitive equation model with a resolution of 18 sigma levels in the vertical and 79 waves of triangular truncation in the horizontal. The forecast model includes a 1.5-order eddy mixing parameterization, a gravity wave drag parameterization, a shallow convection parameterization, a relaxed version of Arakawa-Schubert cumulus parameterization, grid-scale condensation calculation, and longwave and shortwave radiative transfer calculations with consideration of fractional clouds. The performance of the second-generation GFS is significantly better than the first-generation GFS. For two 3-month periods in winter 1995/96 and summer 1996, the second-generation GFS provided forecasters with 5-day forecasts where the averaged 500-mb height anomaly correlation coefficients for the Northern Hemisphere were greater than 0.6.

Observational data available to the GFS are much less than those at other numerical weather prediction centers, especially in the Tropics and Southern Hemisphere. The GRID messages of 5° resolution, ECMWF 24-h forecast 500-mb height and 850- and 200-mb wind fields available once a day on the Global Telecommunications System are used as supplemental observations to increase the data coverage for the GFS data assimilation. The supplemental data improve the GFS performance both in the analysis and forecast.

### 1. Introduction

The Central Weather Bureau (CWB) in Taiwan began developing a global numerical weather forecast system in the early 1980s. The main purpose of developing the global weather forecast system is to provide time-dependent boundary conditions for the companion limited-area model (Jeng et al. 1990) and typhoon track model (Peng et al. 1993), and to provide medium-range weather forecast guidance to forecasters. The weather at the subtropical island is influenced by a variety of different weather systems such as middle-latitude baroclinic systems, cold surges, Mei-Yu fronts, southeast summer monsoons, and tropical cyclones. The complicated weather patterns require a large-scale dynamic prediction model to provide the forecasters with guidance in changes of weather coming in from different directions. The global forecast system

(GFS) developed initially was based upon a version of the University of California, Los Angeles general circulation model. The first-generation global forecast model is a grid-point model with a resolution of 2.5° lat × 2.5° long in the horizontal and nine sigma levels in the vertical. The GFS was operational in 1988 (Liou et al. 1989), providing the CWB forecasters for the first time with comprehensive numerical weather prediction (NWP) guidance up to 5 days.

Immediately after the GFS became operational, a six-year project was started to improve the GFS. The project was designed to bring the GFS up to the proven NWP techniques at that time such as optimum interpolation (OI) analysis, nonlinear normal-mode initialization, spectral representation, and improved physical parameterization schemes. The upgrade project was completed, and the second-generation GFS started operation in January 1995 with a higher resolution of 79 waves in the horizontal and 18 sigma levels in the vertical. The purpose of this paper is to describe the new GFS and provide sample statistics on the performance of the new forecast system.

### 2. Optimum interpolation analysis

The first step in the NWP process is to transform meteorological observations into a uniform format that is

\* On leave from Naval Research Laboratory, Monterey, California.

+ On leave from Department of Atmospheric Sciences, National Taiwan University, Taipei, Taiwan.

Corresponding author address: Dr. Chi-Sann Liou, Naval Research Laboratory, 7 Grace Hopper Avenue, Monterey, CA 93943-5502.  
E-mail: liou@nrlmry.navy.mil

suitable for assimilation into a forecast model. The process involves more than simple interpolation because observational data from various data sources have widely varying error characteristics and the variables to be analyzed are dynamically correlated to each other. In the second-generation GFS, three-dimensional multivariate OI analysis is used to replace a two-dimensional univariate Barnes (1964) analysis. The three-dimensional OI analysis analyzes geopotential height and horizontal wind components at Gaussian grids on 16 constant pressure levels, 15 mandatory levels up to 10 mb plus the 925-mb level. Sea level pressure and temperature on the 16 pressure levels are computed from the analyzed geopotential height with hydrostatic balance.

The OI analysis interpolates observational data from data locations to analysis grid locations by considering the statistics of error covariance and the dynamic constraints of geostrophic coupling and hydrostatic balance. The analysis procedure can be written as

$$F_k^a = F_k^g + \sum_{i=1}^N w_{ik} f_i, \quad (1)$$

where  $F_k$  represents an analysis variable at grid point  $k$ ,  $f_i$  represents observational correction to a first-guess value by observation  $i$ ,  $w_{ik}$  is the weight of the observational correction,  $N$  is the total number of observations, and the superscripts  $a$  and  $g$  distinguish analysis and first-guess values, respectively. The theory of OI analysis developed by Gandin (1965) shows that the analysis errors are statistically minimized if the weight  $w_{ik}$  is chosen as

$$w_{ik} = \sum_j \mathbf{M}_{ij}^{-1} \mathbf{h}_{jk}, \quad (2)$$

where  $\mathbf{M}_{ij}$  is the matrix of total error covariance between the locations of observations  $i$  and  $j$ , and  $\mathbf{h}_{jk}$  is the vector of forecast error covariance between the locations of observation  $j$  and grid point  $k$ . The OI analysis in the GFS was developed closely following the schemes of Lorenc (1981) and Barker (1992). We use an overlapped-volume method to maximize the utilization of available observations by a smaller error covariance matrix. The size of each analysis volume varies from  $25.5^\circ$  lat  $\times$   $7.5^\circ$  long in data-dense areas to  $360^\circ$  lat  $\times$   $12^\circ$  long in polar areas. The dynamic constraints of hydrostatic balance and geostrophic coupling are included in the formulas defining the error covariance matrix. The forecast error correlation function  $\eta(r)$  is chosen in a form of a second-order auto-regression function,

$$\eta(r) = (1 - x_2) + x_2(1 + x_1 r)e^{-x_1 r}, \quad (3)$$

where  $r$  is the distance in units of 1000 km and  $x_1, x_2$  are two constants set to 2.7 and 0.95, respectively. This correlation function makes the covariance matrix  $\mathbf{M}$  be positive definite so that the weight  $w_{ik}$  in (2) can be solved by Cholesky's decomposition method. The constants  $x_1$  and  $x_2$  are reevaluated and adjusted to fit better the 12-h forecast error statistics after each major change made to the GFS.

In addition to a data quality control procedure, following Gandin (1988), performed prior to the OI analysis, observations are also quality checked as a part of the analysis procedure. The check includes tolerance comparison between observations and first-guess values and consistency comparison among nearby observations. In order to increase the representation of observations, highly correlated satellite observations are averaged in each  $2^\circ$  lat  $\times$   $2^\circ$  long boxes and nearby aircraft reports within a 3-h time interval are also averaged.

Recently, Parrish and Derber (1992) demonstrated that three-dimensional variational analysis (3D-Var) on spectral coefficients significantly improves the forecast skills of the aviation model at the National Centers for Environmental Prediction (NCEP). The 3D-Var method removes the data selection limitation in the OI analysis and can directly assemble different types of observations without the data conversion (retrieval) beforehand. The CWB is currently working on upgrading the 3D-Var analysis.

### 3. Governing equations and spectral representation

The forecast model of the GFS is a primitive equation model on sigma coordinates. Following Simmons and Strufling (1981) and Hogan and Rosmond (1991), we choose vertical vorticity  $\zeta$ , horizontal divergence  $D$ , surface pressure  $\pi$ , virtual potential temperature  $\theta$ , and specific humidity  $q$  as prognostic variables. However, we choose a strict sigma coordinate system,  $\sigma = p/\pi$ , rather than a hybrid coordinate system as in Simmons and Strufling (1981) and Hogan and Rosmond (1991), to represent the atmosphere in the vertical. The hybrid coordinate system is a pressure-like coordinate at upper levels that reduces numerical errors in computing pressure gradient force. Since the GFS is designed for 5–7-day forecasts at this time and the calculation in the stratosphere is very primitive (e.g., no chemistry calculation), we choose the simple sigma coordinate to simplify the conversion between the sigma and pressure. The governing equations on the sigma and spherical coordinates are given as

$$\frac{\partial \zeta}{\partial t} = -\mathcal{J}(A, B), \quad (4)$$

$$\frac{\partial D}{\partial t} = \mathcal{J}(B, -A) - \nabla^2(\Phi + K), \quad (5)$$

$$\frac{\partial \theta}{\partial t} = -\nabla \cdot (\mathbf{V}\theta) + \theta D - \sigma \frac{\partial \theta}{\partial \sigma} + Q_\theta, \quad (6)$$

$$\frac{\partial q}{\partial t} = -\nabla \cdot (\mathbf{V}q) + qD - \sigma \frac{\partial q}{\partial \sigma} + Q_q, \quad (7)$$

$$\frac{\partial \pi}{\partial t} = -\nabla \cdot (\mathbf{V}\pi) - \frac{\partial(\sigma\pi)}{\partial \sigma}, \quad (8)$$

and

$$\frac{\partial \Phi}{\partial P} = -c_p \theta, \quad (9)$$

where  $J$  is an operator defined as  $J(b, c) = (1/\cos^2\varphi)[(\partial b/\partial\lambda) + (\partial c/\partial\mu)]$ ;  $\nabla$  is the horizontal gradient operator on the sigma coordinates;  $A = U(\zeta + f) + \sigma(\partial V/\partial\sigma) + (c_p/a^2)\theta(\partial P/\partial\mu) \cos^2\varphi - Q_v[(\cos\varphi)/a]$ ,  $B = V(\zeta + f) - \sigma(\partial U/\partial\sigma) - (c_p/a^2)\theta(\partial P/\partial\lambda) + Q_u[(\cos\varphi)/a]$ ,  $U = u[(\cos\varphi)/a]$ ,  $V = v[(\cos\varphi)/a]$ ,  $f = 2\Omega \sin\varphi$ ,  $\Phi = gz$ ,  $K = \frac{1}{2}(u^2 + v^2)$ ,  $P = (p/p_0)^\kappa$ ,  $\kappa = c_p/R$ ,  $\mu = \sin\varphi$ ;  $u$  and  $v$  are horizontal wind components;  $\lambda$  is longitude;  $\varphi$  is latitude;  $a$  is the earth's radius; and  $Q$ 's are diabatic forcing terms. With the definition of  $J$ ,  $U$ , and  $V$ , the vorticity and divergence can be simply related to the wind components as  $\zeta = J(V, -U)$  and  $D = J(U, V)$ . The prognostic variables on the sigma levels,  $\zeta_k$ ,  $D_k$ ,  $\theta_k$ ,  $q_k$ , and  $\pi$  are expanded in terms of spherical harmonics as  $X(\lambda, \mu, t) = \sum_{m=-M}^{+M} \sum_{n=|m|}^M X_n^m(t) P_n^m(\mu) e^{im\lambda}$ , where  $X$  is one of the above variables,  $P_n^m$  is the associated Legendre functions, and  $M$  is the total number of resolvable waves in triangular truncation. In the current operational GFS, the spectral truncation is  $M = 79$ . The triangular truncation with 79 waves (T79) gives the forecast model a horizontal resolution of  $1.5^\circ$  (about 150 km) in the corresponding Gaussian grids.

As for all operational spectral models, we calculate nonlinear tendency terms of (4)–(8) in the physical space, that is, on Gaussian grid points, and then transform them back to the spectral space. The transformation between the physical and spectral space is done by fast Fourier transforms and Legendre transforms using Gaussian quadrature (Orszag 1970). We use a vertical differencing scheme developed by Arakawa and Suarez (1983) to integrate (6), (7), and (9). The scheme ensures total energy conservation and maintains the vertically integrated pressure gradient force to be irrotational along a contour of the surface topography. The vertical differencing scheme allows us to choose  $\pi$ , rather than  $\ln(\pi)$ , as a prognostic variable to better conserve total mass without generating excessive noise near complex terrain. The time integration of the governing equations (4)–(8) is carried out by a semi-implicit scheme following Hogan and Rosmond (1991). The negative specific humidity due to spectral truncation is removed simply by borrowing moisture from levels below. At the lowest sigma level, the negative specific humidity is removed as if there is artificial surface moisture flux from the surface. The magnitude of the artificial flux is much smaller than that of the real surface moisture flux in all cases. We apply fourth-order horizontal diffusion to vorticity and divergence variables and to the perturbation of potential temperature and specific humidity that are deviated from mean reference states. The terrain field is prepared by applying a low-pass filter and Lanczos's filter once each to a silhouette profile derived from a 10-min resolution terrain dataset.

**4. Physical parameterization**

The global forecast model includes schemes to parameterize the physical processes of surface fluxes, vertical turbulence mixing, shortwave and longwave radiative transfer, cumulus convection, grid-scale condensation, and

gravity wave drag. The surface fluxes are calculated with Louis (1979) formulas that are empirical approximations of Monin and Obukhov (1954) similarity theory results. In this scheme, surface drag coefficients are computed as functions of surface bulk Richardson number and surface roughness. The surface moisture flux is also a function of soil moisture availability or so-called ground wetness, which is the ratio of current soil moisture content to total soil moisture capacity. The ground wetness and ground temperature are held constant at water points, while they are predicted at land points by budget equations that include Newtonian cooling-type forcing to account for the influence from deep soil. The deep soil temperature is computed by interpolating temperature climatology fields obtained from the National Center for Atmospheric Research (NCAR) to the terrain height at all land points.

The vertical turbulence mixing in the forecast model is parameterized by computing eddy mixing fluxes in a form of the  $K$  theory. We choose a turbulence kinetic energy and dissipation closure parameterization to model the mixing coefficients. The choice is based upon the conclusion by Holt and Raman (1988) that the scheme performs best among all schemes they evaluated, and the experience reported by Langland and Liou (1996) that the scheme is efficient and performs well at predicting the evolution of mixed layers in an operational mesoscale model. In this scheme, following Detering and Etling (1985), the eddy mixing coefficient for momentum  $K_m$  is computed as a function of turbulence kinetic energy  $E$  and its dissipation rate  $\varepsilon$ ,  $K_m = C_2 E^2/\varepsilon$ , where  $C_2$  is a constant set to 0.026. The eddy mixing coefficient for heat  $K_h$  is computed from  $K_m$  by a Prantl number, which is in a form of surface Prantl number obtained by the similarity theory. The turbulence kinetic energy and its dissipation rate are predicted by following equations that were formulated by Detering and Etling (1985) for atmospheric models:

$$\frac{\partial E}{\partial t} = K_m \left[ \left( \frac{\partial u}{\partial z} \right)^2 + \left( \frac{\partial v}{\partial z} \right)^2 \right] - \frac{g}{\theta} K_h \frac{\partial \theta}{\partial z} + C_1 \frac{\partial}{\partial z} \left( K_m \frac{\partial E}{\partial z} \right) - \varepsilon, \tag{10}$$

$$\frac{\partial \varepsilon}{\partial t} = C_3 \frac{\varepsilon}{E} \left\{ K_m \left[ \left( \frac{\partial u}{\partial z} \right)^2 + \left( \frac{\partial v}{\partial z} \right)^2 \right] - \frac{g}{\theta} K_h \frac{\partial \theta}{\partial z} \right\} - C_4 \frac{\varepsilon^2}{E} + C_5 \frac{\partial}{\partial z} \left( K_m \frac{\partial \varepsilon}{\partial z} \right), \tag{11}$$

where the constants are chosen as  $C_1 = 1$ ,  $C_3 = 1.38$ ,  $C_4 = 1.9$ , and  $C_5 = 0.77$ . These two equations are numerically integrated by an implicit time scheme in which nonlinear terms  $\varepsilon/E$  and  $\varepsilon^2/E$  in (11) at time step  $n$  are evaluated in a form of  $C_2(E^{n-1} + E^n)/(2K_m)$  and  $\varepsilon^{n-1}[2\varepsilon^n/(E^{n-1} + E^n)]$ , respectively. Although the implicit scheme allows a large time step in time integration, we limit the maximum time step in integrating (10) and (11) to 150 s to avoid large numerical

errors. We use several consecutive smaller time steps to match the model time step, which is typically around 600 to 900 s. We include checks on upper and lower bounds of  $E$  and  $\varepsilon$  to avoid negative values and division by zero. The full prediction in  $E$  and  $\varepsilon$  makes this turbulence mixing parameterization a 1.5-order closure scheme.

A one-dimensional (1D) model based upon the surface flux and eddy mixing parameterization was constructed to test the planetary boundary layer parameterization calculated by a large time step (600 s) and a coarse vertical resolution (four levels in bottom 400 m). A marine boundary layer development case observed during the Monsoon Experiment 1979 (MONEX79, Holt and Raman 1988) was selected for the test. The 1D model simulation captures well the growth of the mixed layer and the magnitude of the mixing coefficient, as considering the coarse resolution and uncertainty in horizontal advection contribution (Fig. 1).

Radiative transfer is calculated following Harshvardhan et al. (1987). The parameterization includes a longwave and shortwave radiative transfer calculation that considers the effects of fractional clouds in radiative fluxes. The amount of stratiform clouds is diagnosed from relative humidity at different levels (Slingo and Ritter 1985), while the amount of cumulus clouds is diagnosed from precipitation rate (Slingo 1987). In the longwave calculation, the parameterization considers four broadband regions to account for absorption by  $H_2O$ ,  $CO_2$ , and  $O_3$ . The cloud effects on longwave radiative transfer are modeled by the probability of clear line of sight between different levels. The emissivity is assumed to be 1 for warm clouds ( $T > 273$  K), 0.5 for supercooled cirrus clouds ( $T < 233$  K), and linearly varied from 1 to 0.5 for clouds with temperature in between 273 and 233 K. If the cloud-layer thickness ( $CL_h$ ) is less than 20 mb, the cloud emissivity is reduced by a factor of  $CL_h/20$ . In the shortwave calculation, the scheme considers absorption by  $H_2O$  and  $O_3$  and multiple Mie scattering by clouds. The absorption by  $H_2O$  is evaluated in five broad bands for wavelengths longer than  $0.9 \mu m$ . The multiple scattering is solved by an adding method that sums up the two stream solutions for direct and diffuse fluxes. The direct fluxes are computed by the delta-Eddington two-stream approximation (Joseph et al. 1976), while the reflection and transmission of diffuse shortwave radiative transfer are computed following Sagan and Pollack (1967). To reduce computational cost in the radiation calculation, we update radiative fluxes every hour and compute longwave fluxes in every second, third, fourth, and fifth grid points of a latitude ring at the Tropics ( $20^\circ S$  to  $20^\circ N$ ), middle latitudes ( $20^\circ-60^\circ S$  and  $20^\circ-60^\circ N$ ), high latitudes ( $60^\circ-80^\circ S$  and  $60^\circ-80^\circ N$ ), and polar areas ( $80^\circ-90^\circ S$  and  $80^\circ-90^\circ N$ ), respectively. The spatial resolution for the longwave flux calculation is chosen to keep

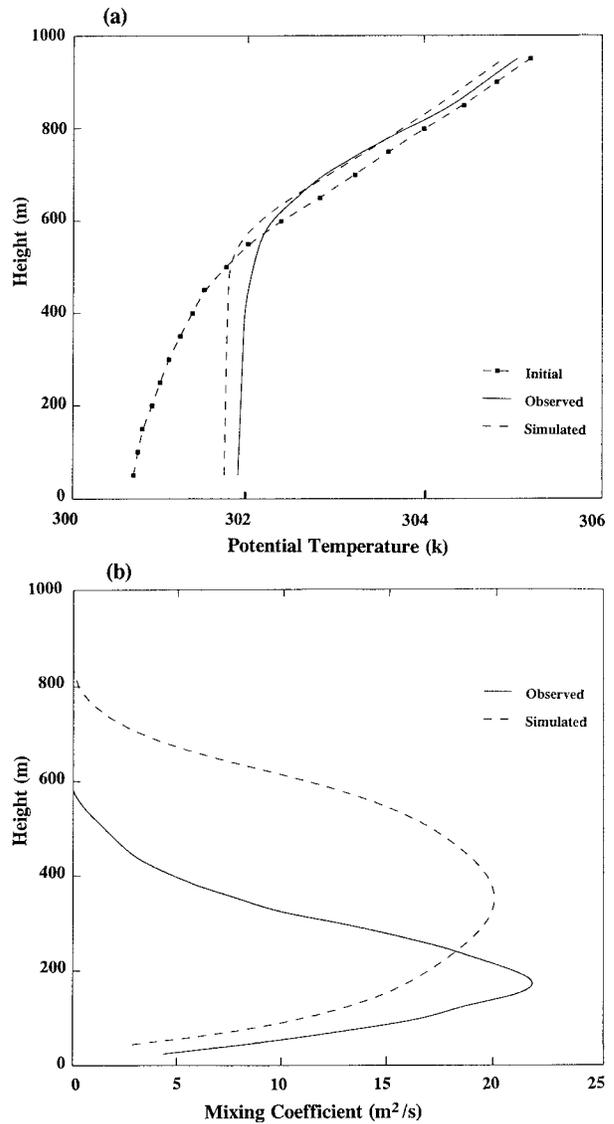


FIG. 1. Boundary layer profiles (observed (solid) on 14 July 1979 during MONEX79 and simulated (dashed) by the one-dimensional PBL model with 600-s time steps and four vertical layers in the bottom 400 m: (a) potential temperature and (b) eddy mixing coefficient for momentum. The dotted lines are initial values for the model simulation.

the meridional resolution about double the zonal resolution.

Cumulus convection is parameterized by a relaxed form of Arakawa and Schubert (1974, hereafter AS) parameterization following Moothi and Suarez (1992). We choose this scheme because it produces almost the same results as those from the standard AS parameterization, which has been proven to work well in global numerical weather prediction and climate simulation, while it removes the ill-posed integral equation problem in the standard AS and is much more efficient. The relaxed parameterization simplifies the AS parameterization by assuming the nor-

malized cloud mass flux to be a linear function of height and by adjusting temperature and moisture profiles of a grid point by the effect of one cloud type at a time. The relaxation parameter is chosen to be 0.25. Cumulus convection is triggered at grid points where the cloud base is within the bottom half of the atmosphere and a conditionally unstable layer exists above it. The cloud base is defined at a model level where the lifting condensation level (LCL) of a surface air parcel is right beneath it. In the LCL calculation, the surface air parcel is mixed with entrained environmental air by an entrainment rate yielding a mass flux linearly increasing with height. The cloud-base properties are the results of the mixing between the surface air and the entrained environmental air. Above the cloud base, the cloud air properties are computed by a moist-adiabatic process plus mixing with the environmental air by an entrainment rate defined for the cloud type. The cloud top is the level where the cloud air loses its buoyancy. The cloud type is discarded if the cloud air is not saturated at the cloud top. We speed up the cumulus parameterization calculation by first gathering all potentially unstable grid points and then computing adjustments at those points only.

A nonprecipitating-type shallow convection parameterization is also included in the forecast model to parameterize effects of surface-forced shallow convection in a conditionally unstable layer near the surface. The shallow convection parameterization, following Tiedtke (1984), vertically mixes potential temperature and specific humidity in a conditionally unstable layer with a prescribed eddy mixing coefficient of  $10 \text{ m}^2 \text{ s}^{-2}$ . The criteria to start the shallow convection are that 1) surface ground temperature is warmer than surface air temperature, 2) relative humidity at the lowest model level is greater than 70%, 3) LCL of the surface air is within 200 mb above the earth's surface, and 4) a conditionally unstable layer exists in the bottom 200 mb. The shallow convection parameterization will mix a conditionally unstable column even if there is an inversion layer underneath. The parameterization helps to smooth out strong inversions in subcloud layers in trade wind areas.

Supersaturation is removed in the forecast model through the condensation of excessive water vapor. The procedure checks supersaturation layer by layer from the model top. The saturated water vapor is condensed to liquid water and the latent heat is released to the local layer. The condensed water falls to the layer beneath and evaporates until relative humidity there reaches 98%. The remainder of the liquid water keeps falling to layers below and is subject to the same evaporation condition until it reaches the earth's surface. The condensed water reaching the surface is treated as grid-scale (large scale) rain.

To represent the effect of vertical momentum flux associated with gravity waves triggered by subgrid-

scale terrain, a gravity drag parameterization, following Palmer et al. (1986), is included in the forecast model. The gravity wave drag is formulated as

$$\frac{\partial \mathbf{V}}{\partial t} = -g \frac{\partial \boldsymbol{\tau}_{gw}}{\partial p}, \quad (12)$$

where  $\boldsymbol{\tau}_{gw}$  is the stress due to the momentum flux of gravity waves exerted by subgrid-scale terrain. The stress generated at surface is parameterized by

$$\boldsymbol{\tau}_{s,gw} = k_{gw} N \mathbf{V}_s h^2, \quad (13)$$

where  $k_{gw} = 0.000025 \text{ m}^{-1}$ ,  $N$  is the Brunt-Väisälä frequency,  $\mathbf{V}_s$  is the representative surface wind computed by mass-weighted mean wind of the lowest three layers, and  $h$  is the gravity wave amplitude. The square of the wave amplitude  $h^2$  at the surface is set to the variance of the subgrid terrain. The gravity wave generated at the surface propagates upward, and the wave stress is changed only at a critical level or at a level where the wave Richardson number,  $\text{Ri}^* = \text{Ri}[1 - (Nh/|\mathbf{V}_{gw}|)]/\{[1 + \text{Ri}^{1/2}(Nh/|\mathbf{V}_{gw}|)]^2\}$ , is less than 0.25. The wind  $\mathbf{V}_{gw}$  is the wind component along the  $\mathbf{V}_s$  direction. At the critical level, total absorption is assumed and the stress is set to zero. At the level where  $\text{Ri}^*$  is less than 0.25, partial absorption is assumed, and the amplitude of the gravity wave is recomputed with  $\text{Ri}^* = 0.25$  according to the saturation hypothesis of Lindzen (1981). The gravity wave drag parameterization makes significant improvement on model forecasts at upper levels. Figures 2 and 3 shows monthly mean results of a seasonal simulation by the T47 GFS with and without the gravity wave parameterization. The gravity wave drag parameterization increases the amplitude of monthly mean long waves at 500 mb (Fig. 2) and better defines the jets' locations and strengths at 200 mb (Fig. 3). Similar improvements are found in several case studies with T79 resolution of the GFS.

## 5. Initialization procedure

The GFS runs with 12-h intermittent data assimilation cycles to produce two forecasts a day. The forecast model is initialized by incremental nonlinear normal-mode initialization (Ballish et al. 1992). In this initialization procedure, the analyzed increments of wind, temperature, and terrain pressure are adjusted so that high-frequency gravity modes are in nonlinear balance initially. With 18 sigma levels in the vertical, we initialize two vertical modes by requiring gravity waves of these two modes with periods less than 24 h to satisfy nonlinear balance conditions.

The initial sea surface temperature (SST) is analyzed by univariate OI analysis with a first-guess field based on NCEP-published weekly mean SST and observations from weekly accumulated surface ship reports. Moisture fields are not analyzed yet in the GFS. Specific humidity is initialized by values carried over

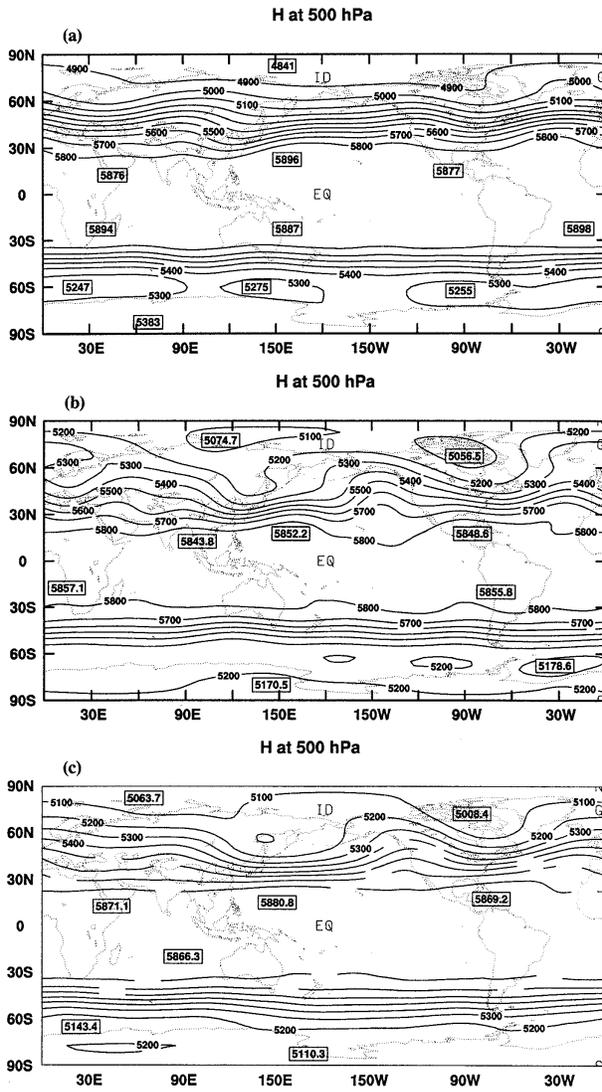


FIG. 2. Monthly mean 500-mb height in January from the T47 second-generation GFS climate simulation (a) without and (b) with the gravity wave drag parameterization, and (c) the corresponding 10-yr mean analysis from NCEP.

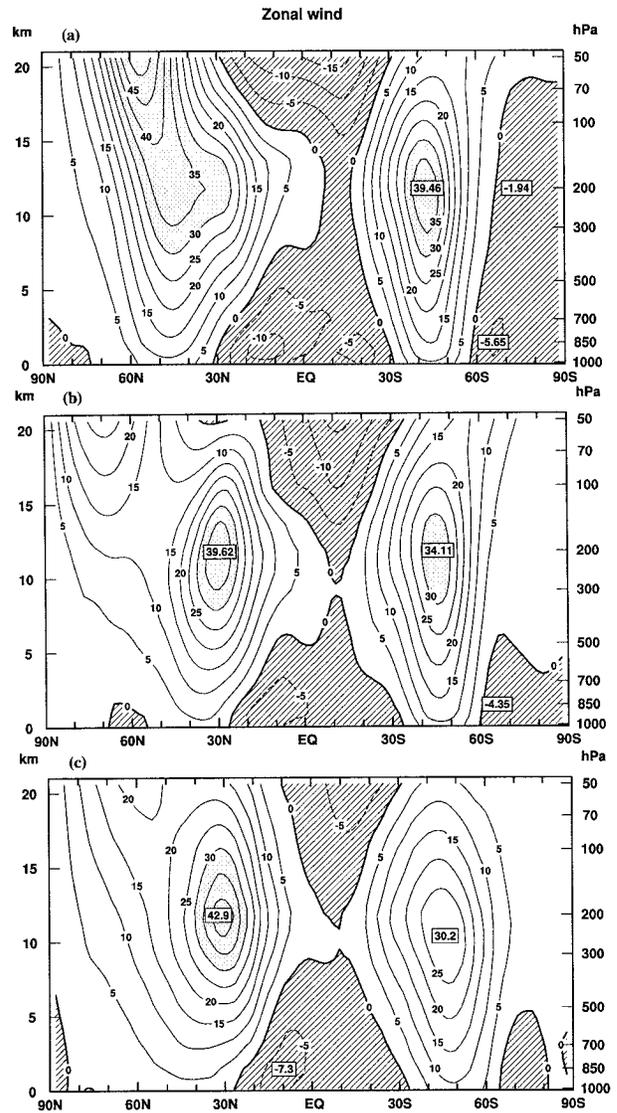


FIG. 3. Same as in Fig. 2 except for zonally averaged monthly mean wind from 1000 to 50 mb.

from previous model forecasts validated at the initial time. The ground temperature and ground wetness over land points are not analyzed either and are initialized by values carried over from previous model forecasts as well.

**6. Issues on observational data**

Observational data available from the CWB database are much less than those at other NWP centers. Satellite soundings and cloud-tracked wind data received at the CWB are about an order of magnitude less than those received at other centers. Furthermore, the CWB receives no ARNIC Communications Addressing and Reporting System or Special Sensor Mi-

crowave/Imager data. Lack of observational data certainly limits the performance of the GFS.

In an attempt to access the impact of the data shortage on the GFS performance, we conducted a numerical experiment in which nine analysis fields of  $2.5^\circ \times 2.5^\circ$  resolution at three pressure levels from the European Centre for Medium-Range Weather Forecasts (ECMWF) were used as supplemented observational data. In the experiment, the ECMWF height and wind analyses at 850, 500, and 200 mb were used as observations by the OI analysis if the total real observations for an analysis volume are less than  $\frac{1}{4}$  of the maximum allowed amount  $[(\frac{1}{4})/(360) = 90]$ . During a 15-day period from 1 to 15 July 1994, the supplemental data made profound impacts both on the analysis and forecast. For example, the 15-day

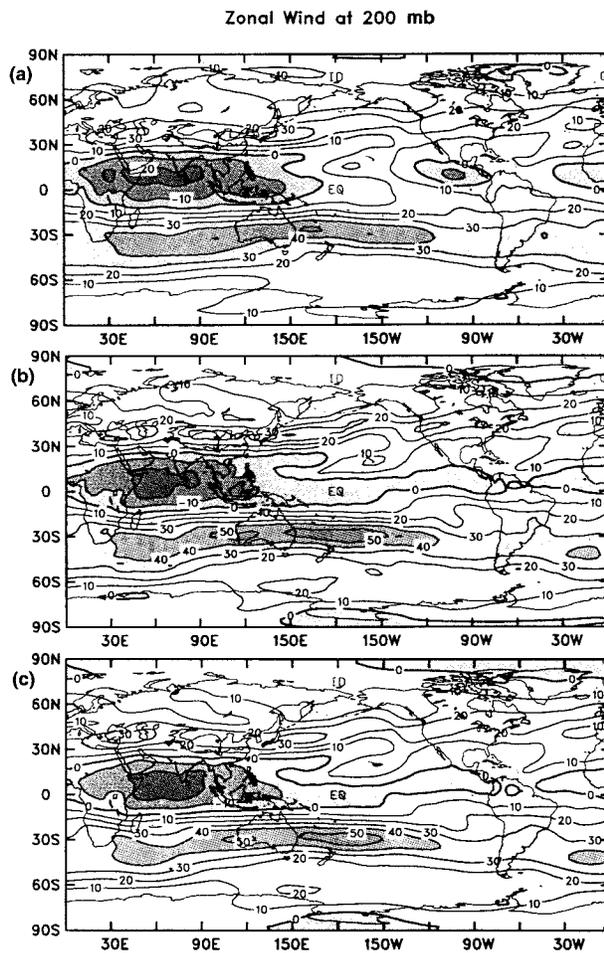


FIG. 4. Fifteen-day (1–15 July 1994) mean zonal wind at 200 mb from (a) the second-generation GFS analysis with real observations operationally received at the CWB only, (b) the ECMWF analysis, and (c) the second-generation GFS analysis with the real observations plus the supplemental observations from the 2.5° ECMWF analysis at 850, 500, and 200 mb.

mean GFS 200-mb zonal wind analysis without the supplemental data showed a distribution very different from the mean ECMWF analysis (Figs. 4a,b). The main difference was in data-sparse areas where the mean GFS analysis had a jet that was too weak in the Southern Hemisphere and a discontinued easterly wind region over the tropical central Pacific (Fig. 4a). With the supplemental data, the mean GFS 200-mb zonal wind analysis was significantly improved both in the Southern Hemisphere jet speed and the tropical easterly wind distribution (Fig. 4c). Similar impact was found in the 850-mb zonal wind analysis that the supplemental data corrected the GFS analysis problems over the Pacific where the wind was too easterly in the Tropics and less westerly in the Southern Hemisphere middle latitudes (Fig. 5). More impressive impact was found in the forecast with the supplemental data where the 15-day mean 500-mb height anomaly

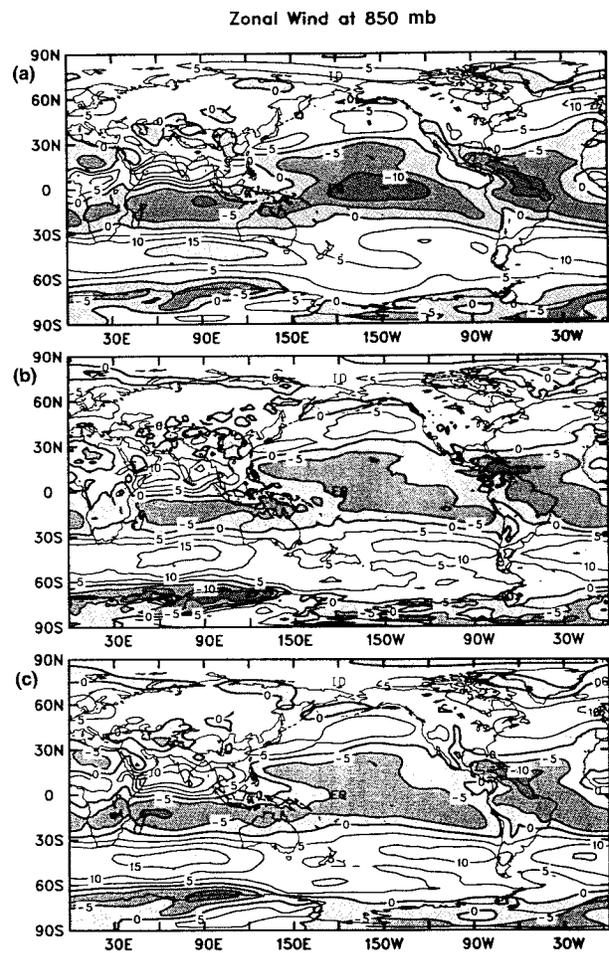


FIG. 5. Same as in Fig. 4 except for zonal wind at 850 mb.

correlation for the Northern Hemisphere (20°–80°N) was improved by 0.03 in the analysis and by up to 0.13 in the 5-day forecast (Fig. 6). The anomaly correlation is computed by correlating the deviation of the GFS analysis and forecast from the NCAR climatology to the deviation of the ECMWF analysis from the same NCAR climatology. The results of the experiment clearly demonstrate the potential skills of the GFS and the necessity of adding more data in the Tropics and Southern Hemisphere for the GFS data assimilation.

In real-time daily operation, the ECMWF analysis fields, as well as analysis fields from any other major NWP centers, are not available to the CWB. The best and most reliable supplemental observations we can obtain are those from GRID messages on the Global Telecommunications System (GTS). We use the GRID messages of ECMWF 24-h forecast 500-mb height and 850- and 200-mb winds as the supplemental observations in the GFS daily operation. The GRID messages are available only once a day at 1200 UTC with 5° × 5° resolution covering only two zones of 90°–20°N and 90°–20°S for 500-mb height, and one 30°N–

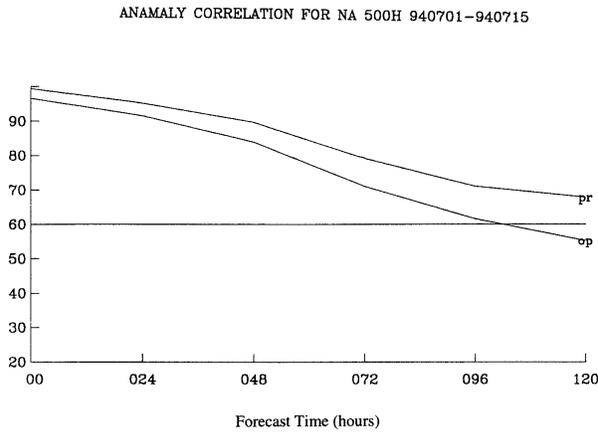


FIG. 6. Fifteen-day (1–15 July 1994) mean 500-mb height anomaly correlation for the Northern Hemisphere (20°–80°N) from the second-generation GFS forecasts with (marked “pr”) and without (marked “op”) the supplemental observations from the 2.5° ECMWF analysis at 850, 500, and 200 mb. The anomaly correlation is computed by correlating the GFS forecast anomaly to the ECMWF analyzed anomaly.

30°S tropical band for 850- and 200-mb winds. Although the supplemental data for the real-time operational GFS are much less than those in the above experiment (24-h forecasts of five fields once a day at 5° resolution nonglobal coverage versus analyses of nine fields twice a day at 2.5° resolution global coverage), the improvement in the GFS performance is still significant (not shown). The supplemental data assimilation procedure replaces the old procedure used in the first-generation GFS where the climatological values were blended with the first-guess fields for the analysis in the Tropics and the Southern Hemisphere.

7. Operational status

The GFS codes are fully vectorized and designed for multitasking on a CRAY YMP computer. For the resolution of T79 truncation and 18 sigma levels, with six processors, it takes about 4 min of wall time to complete the OI analysis and 10 min of wall time to

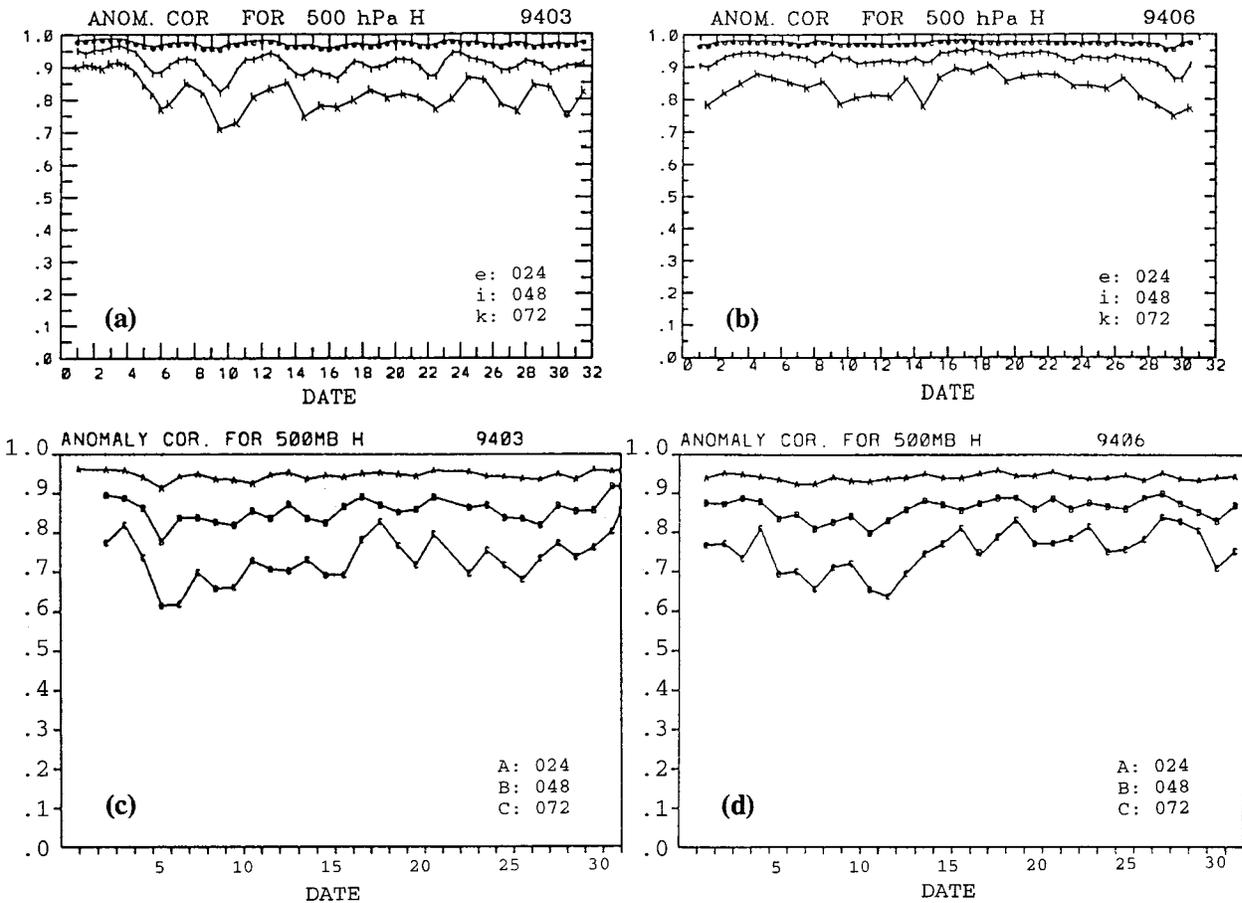


FIG. 7. The 500-mb height anomaly correlation for the Northern Hemisphere (20°–80°N) of 24-, 48-, and 72-h operational forecasts by (a) the second-generation GFS in March 1994, (b) the second-generation GFS in June 1994, (c) the first-generation GFS in March 1994, and (d) the first-generation GFS in June 1994. The anomaly correlation is computed by correlating the forecast anomaly to their own analyzed anomaly.

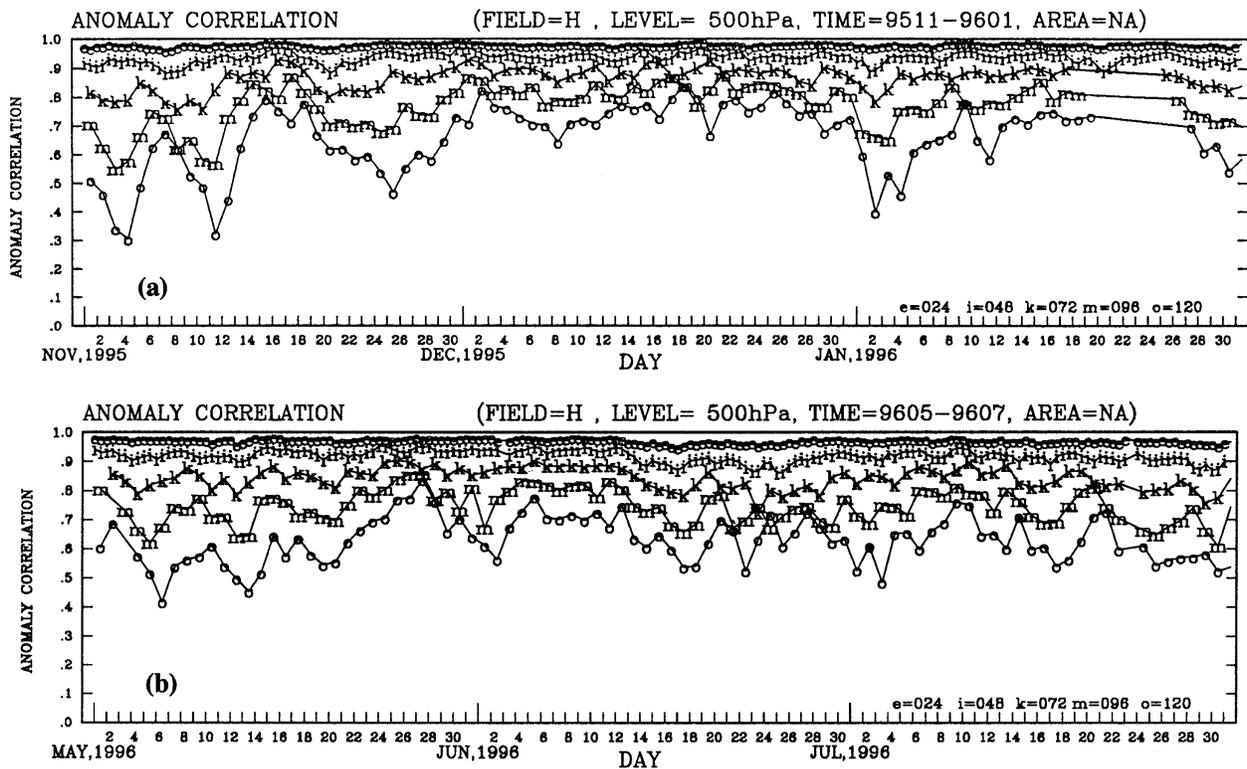


FIG. 8. The 500-mb height anomaly correlation for the Northern Hemisphere (20°–80°N of 24-, 48-, 72-, 96-, and 120-h operational forecasts by the second-generation GFS for (a) November 1995–January 1996 and (b) May 1996–July 1996. The anomaly correlation is computed by correlating the forecast anomaly to its own analyzed anomaly.

complete a 24-h forecast. The GFS is executed to produce 3-day forecasts in daytime and 7-day forecasts in nighttime. The model forecasts provide not only forecast guidance to the CWB forecasters but also time-dependent boundary conditions to the limited area model and typhoon track model at the CWB. The second-generation GFS is significantly better than its earlier version. During a 4-month parallel run period from March to June 1994, the forecast scores of anomaly correlation, S1 scores, mean bias errors, and standard deviation errors were computed and compared between the two versions of the GFS. All scores of the second-generation GFS are significantly better than those of the first-generation GFS. Figure 7 shows the comparison of 500-mb height anomaly correlation for the Northern Hemisphere (from 20° to 80°N) between the two versions of the GFS in March and June 1994. The anomaly correlation is computed by correlating the forecast anomaly, deviation from the NCAR climatology, to the its own analyzed anomaly. For the second-generation GFS, the monthly mean anomaly correlation coefficients in March 1994 are 0.97, 0.91, and 0.82 for forecast periods of 24, 48, and 72 h, and the similar mean coefficients in June are 0.96, 0.93, and 0.85. For the first-generation GFS, the corresponding mean coefficients are 0.95, 0.86,

and 0.73 in March 1994 and 0.93, 0.86, and 0.76 in June 1994. Figure 8 shows the 500-mb height anomaly correlation for the Northern Hemisphere from the operational forecasts of the second-generation GFS in a 3-month period of winter 1995/96 and a 3-month period of summer 1996. The averaged anomaly correlation coefficients for 5-day forecasts were greater than 0.6 in both winter and summer seasons.

Besides the above score statistics, we have evaluated the GFS skills in predicting changes of the weather systems important to Taiwan area. The strength and location of the subtropical high over the west Pacific in summer are better predicted by the second-generation GFS even without the supplemental data discussed in section 6. In the first-generation GFS, the subtropical high was weakened and retreated eastward during the late period of the 5-day forecast. The second-generation GFS maintains the strength and location of the high center much better. The improvement is attributed to better parameterization schemes used in the eddy mixing and radiative transfer calculation. The forecast of cold surges (cold air suddenly moves in from the northwest) in winter is also improved in the second-generation GFS. The problems of too fast and too weak cold surge forecasts are reduced in the second-generation GFS. The increase

in the model resolution is identified as the main reason for the improvement of the cold surge forecasts.

## 8. Summary and future plans

The GFS that entered operational use in 1988 has been upgraded to include better numerical methods, improved physical parameterization, and supplemental observations in data assimilation. The second-generation GFS uses spectral representation in the numerical integration of the governing equations. The major improvements in the GFS are to use multivariate OI analysis in place of univariate Barnes analysis, to use nonlinear normal model initialization in place of linear balance of all modes, to use 1.5-order eddy mixing parameterization in place of a well-mixed boundary layer at the bottom, to use a more completed radiative transfer calculation, and to use a relaxed Arakawa-Schubert parameterization to speed up the computation for cumulus convection. The supplemental observations from the GRID messages on the GFS are used to increase the data coverage that replaces the old procedure of blending the climatological values with the first-guess fields in the Tropics and Southern Hemisphere. The forecast skills of the second-generation GFS are significantly improved over those of the first-generation GFS. With relatively less observational data and T79 resolution, the GFS can provide more than 5 days of useful forecast guidance (anomaly correlation coefficient of 500-mb height > 0.6) to forecasters. The forecast skills of the second-generation GFS are, however, still behind those of other major NWP centers. As for winter 1996/97, the 500-mb height anomaly correlation coefficient for the Northern Hemisphere was about 0.02 to 0.1 below the scores of the ECMWF for 24 to 120-h forecasts, respectively.

Researchers at the CWB are currently working on several projects to keep improving the GFS. The GFS has been tested to run with 120 wave truncation. Preliminary results show that the increase of resolution improved the score of the 500-mb height anomaly correlation in January 1997 by about 0.005 to 0.04 for 24 to 120-h forecasts, respectively. A new version of cumulus parameterization that includes the downdraft effects following Cheng and Arakawa (1993) is being tested for implementation to the GFS. A 3D-Var analysis method is in development to replace the OI analysis. An adjoint model of the global forecast model is in development for 4D-Var analysis and ensemble forecasts in the future. Other efforts such as tuning physical parameterization schemes, better use of available data, and evaluating results to identify systematic problems of the GFS are carried out daily by researchers in the CWB.

*Acknowledgments.* We thank Drs. C.-P. Chang and

C.-Y. Tsai, and Director S.-L. Shieh for their leadership and support in this NWP project.

## REFERENCES

- Arakawa, A., and W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment, Part I. *J. Atmos. Sci.*, **31**, 674–701.
- , and M. Suarez, 1983: Vertical differencing of the primitive equations in sigma coordinates. *Mon. Wea. Rev.*, **111**, 34–45.
- Ballish, B., X. Cao, E. Kalnay, and M. Kanamitsu, 1992: Incremental nonlinear normal mode initialization. *Mon. Wea. Rev.*, **120**, 1723–1734.
- Barker, E. H., 1992: Design of the navy's multivariate optimum interpolation analysis system. *Wea. Forecasting*, **7**, 220–231.
- Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.*, **3**, 396–409.
- Cheng, M.-D., and A. Arakawa, 1993: A cumulus parameterization with rainwater budget and convective downdrafts. Preprints, *10th Conf. on Hurricane and Tropical Meteorology*, San Antonio, TX, Amer. Meteor. Soc., 313–316.
- Detering, H. W., and D. Etling, 1985: Application of the E-e turbulence model to the atmospheric boundary layer. *Bound.-Layer Meteor.*, **33**, 113–133.
- Gandin, L. S., 1965: *Objective Analysis of Meteorological Fields*. Israel Program of Scientific Translation, 242 pp.
- , 1988: Complex quality control of meteorological observations. *Mon. Wea. Rev.*, **116**, 1138–1156.
- Harshvardhan, R. Davies, D. Randall, and T. Corsetti, 1987: A fast radiation parameterization for atmospheric circulation models. *J. Geophys. Res.*, **92**, 1009–1016.
- Hogan, T. F., and T. E. Rosmond, 1991: The description of Navy Operational Global Atmospheric Prediction System's spectral forecast model. *Mon. Wea. Rev.*, **119**, 1786–1815.
- Holt, T., and S. Raman, 1988: A review and comparative evaluation of multilevel boundary layer parameterizations for first-order and turbulent kinetic energy closure schemes. *Rev. Geophys.*, **26**, 761–780.
- Jeng, B.-F., H.-J. Chen, S.-C. Lin, T.-M. Leou, M. S. Peng, S. W. Chang, W.-R. Hsu, and C.-P. Chang, 1991: The limited-area forecast systems at the Central Weather Bureau in Taiwan. *Wea. Forecasting*, **6**, 155–178.
- Joseph, J. H., W. J. Wiscombe, and J. A. Weinman, 1976: The delta-Eddington approximation for radiative flux transfer. *J. Atmos. Sci.*, **33**, 2452–2459.
- Langland, R. H., and C.-S. Liou, 1996: Implementation of an E-ε parameterization of vertical subgrid-scale mixing in a regional model. *Mon. Wea. Rev.*, **124**, 905–918.
- Lindzen, R. S., 1981: Turbulence and stress due to gravity wave and tidal breakdown. *J. Geophys. Res.*, **86**, 9707–9714.
- Liou, C.-S., C.-T. Terng, W.-S. Kau, T. Rosmond, C.-S. Chen, J.-H. Chen, and C.-Y. Tsai, 1989: Global weather forecast system at Central Weather Bureau. *Pap. Meteor. Res.*, **12**, 205–228.
- Lorenz, A. C., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109**, 701–721.
- Louis, J. F., 1979: A parametric model of vertical eddy fluxes in the atmosphere. *Bound.-Layer Meteor.*, **17**, 187–202.
- Monin, A. S., and A. M. Obukhov, 1954: Basic regularity in turbulent mixing in the surface layer of the atmosphere. *Akad. Nauk. SSSR Trud. Geofiz. Inst. Tr.*, **24**, 163–187.
- Moothi, S., and M. J. Suarez, 1992: Relaxed Arakawa-Schubert: A parameterization of moist convection for general circulation models. *Mon. Wea. Rev.*, **120**, 978–1002.
- Orszag, S. A., 1970: Transform method for the calculation of vector-coupled sums: Application to the spectral form of the vorticity equation. *J. Atmos. Sci.*, **27**, 890–895.
- Palmer, T. N., G. Shutts, and R. Swinbank, 1986: Alleviation of a systematic westerly bias in general circulation and numer-

- ical weather prediction models through an orographic gravity wave drag parameterization. *Quart. J. Roy. Meteor. Soc.*, **112**, 1001–1039.
- Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.
- Peng, M. S., B.-F. Jeng, and C.-P. Chang, 1993: Forecast of typhoon motion in the vicinity of Taiwan during 1989–1990 using a dynamic model. *Wea. Forecasting*, **8**, 309–325.
- Sagan, C., and J. B. Pollack, 1967: Anisotropic nonconservative scattering and clouds of Venus. *J. Geophys. Res.*, **72**, 466–477.
- Simmons, A. J., and R. Strufing, 1981: An energy and angular momentum conserving finite-difference scheme, hybrid coordinates, and medium-range weather prediction. ECMWF Tech. Rep. 28, 68 pp. [Available from European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading RG2 9AX, United Kingdom.]
- Slingo, J., 1987: The development and verification of a cloud prediction scheme for the ECMWF model. *Quart. J. Roy. Meteor. Soc.*, **113**, 899–927.
- , and B. Ritter, 1985: Cloud prediction in the ECMWF model. ECMWF Tech. Rep. 46, 48 pp. [Available from European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading RG2 9AX, United Kingdom.]
- Tiedtke, M., 1984: The sensitivity of the time-scale flow to cumulus convection in the ECMWF model. *ECMWF's Workshop on Convection in Large-Scale Numerical Models*, ECMWF, Reading, United Kingdom, 297–316. [Available from European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading RG2 9AX, United Kingdom.]