The Design and Testing of the Navy Operational Global Atmospheric Prediction System

THOMAS E. ROSMOND

Naval Oceanographic and Atmospheric Research Laboratory, Atmospheric Directorate, Monterey, California

(Manuscript received 25 February 1991, in final form 21 January 1992)

ABSTRACT

The Navy Operational Global Atmospheric Prediction System (NOGAPS) has proven itself to be competitive with any of the large forecast models run by the large operational forecast centers around the world. The navy depends on NOGAPS for an astonishingly wide range of applications, from ballistic winds in the stratosphere to air–sea fluxes to drive ocean general circulation models. Users of these applications will benefit from a better understanding of how a system such as NOGAPS is developed, what physical assumptions and compromises have been made, and what they can reasonably expect in the future as the system continues to evolve.

The discussions will be equally relevant for users of products from other large forecast centers, e.g., National Meteorological Center, European Centre for Medium-Range Weather Forecasts. There is little difference in the scientific basis of the models and the development methodologies used for their development. However, the operational priorities of each center and their computer hardware and software environments often dictate what compromises are made and how model-based research is conducted. In this paper, NOGAPS will serve as the basis for discussing these issues and the art of numerical weather prediction model development.

1. Introduction

The U.S. Navy began development of a global numerical weather prediction (NWP) capability in the mid-1970s, when the decision was made to adopt the UCLA general circulation model (GCM) (Arakawa and Lamb 1977) as the basis for its prototype system. The Naval Environmental Prediction Research Facility (NEPRF), now part of the Naval Oceanographic and Atmospheric Research Laboratory (NOARL), was responsible for prediction system development. Fleet Numerical Oceanography Center (FNOC) runs the system operationally. The Navy Operational Global Atmospheric Prediction System (NOGAPS) became operational at FNOC in 1982. The components of this system, NOGAPS 2, described by Rosmond (1981), consisted of a somewhat modified version of the UCLA general circulation model, a successive corrections analysis, and a calculus of variations initialization scheme (Barker 1982). NOGAPS 2 ran operationally until 1988.

By the mid-1980s, it was clear that the operational performance of NOGAPS 2 was being outstripped by newer global forecast systems implemented at the other major operational centers. Both the European Centre for Medium-Range Weather Forecasts (ECMWF) and the National Meteorological Center (NMC) demonstrated that global spectral models offered advantages over finite-difference models for the global NWP problem. In particular, spectral models are superior when horizontal advection is important, such as the prediction of Rossby short waves. Moreover, spectral models easily incorporate computationally efficient semi-implicit time-integration schemes that are essential for high model resolution.

During the 1980s, optimum interpolation (OI) analysis and nonlinear normal model (NLNM) initialization also gained wide acceptance for operational NWP. Optimum interpolation (Lorenc 1981) combines a model forecast first guess with a wide variety of observation types in a mathematically rigorous way. Normal mode initialization (Machenhauer 1977) performs subtle adjustments to the analyzed fields to remove undesirable gravity-wave noise that can contaminate the first few hours of a forecast. Together, OI and NLNM have led to quite sophisticated four-dimensional data-assimilation schemes that allow exploitation of such nonconventional data as surface wind speed from the Special Sensor Microwave/Imager (SSM/I) on Defense Meteorological Satellite Program (DMSP) satellites.

Those of us responsible for development of operational navy forecast systems saw that NOGAPS had to be quickly brought up to the state of the art if we were to retain credibility in the highly competitive operational global NWP business. Because of NOGAPS 2 software design limitations, it was impractical to consider a piecemeal replacement of system components with improvements. Instead, NOGAPS 2 was scrapped

Corresponding author address: Dr. Thomas E. Rosmond, NOARL, Atmospheric Directorate, Monterey, CA 93943-5005.

and a totally new system developed in its place. The development process for this new system, NOGAPS 3, is the theme of this report.

Section 2 gives an overview description of NOGAPS 3 and discusses some NOGAPS 3 design philosophy issues. Section 3 gives several examples of the kinds of diagnostic tools and instruments that are used to develop sophisticated forecast models like NOGAPS. Section 4 gives a brief comparison of NOGAPS performance relative to some of the other major operational centers. Section 5 describes two forms of the "spinup problem" that continue to be the object of considerable research for NWP systems. Section 6 discusses another particularly nagging problem for all global systems, the systematic error or climate drift problem. This is especially important to the navy because of the important NOGAPS role of providing forcing to operational ocean models. Section 7 discusses future NOGAPS plans. Some additional discussion and a summary are given in section 8.

The emphasis of this paper is on NOGAPS design and testing methodology. NOGAPS synoptic forecast skill, while the final measure of merit of the system, is not the primary evaluation criteria used in model development. Fundamental quantities such as surface fluxes, radiation fluxes, cloud distribution, and mean meridional circulation are more closely related to model physical processes and so are more-sensitive indicators of model behavior. Readers interested in more details of NOGAPS operational performance are referred to Hogan and Rosmond (1991), Harr et al. (1991), Clune et al. (1991), and the FNOC *Quarterly Performance Summary* (1986–1990).

2. NOGAPS description and design issues

A detailed technical description of the NOGAPS 3 initialization and forecast model is given by Hogan and Rosmond (1991) and Hogan et al. (1992). The NOGAPS OI analysis (Barker et al. 1988) follows the ECMWF approach (Lorenc 1981). The major features of NOGAPS are summarized in Table 1. Perhaps the most novel feature of NOGAPS is the use of the Arakawa–Schubert (A/S) cumulus parameterization, described by Hogan and Rosmond (1991). For the most part, the other large global NWP models use simpler, more economical schemes. However, recent research results emphasizing the impact of tropical forcing on midlatitude circulations (Wallace and Gutzler 1981; Gelaro 1989) have increased the interest in improved tropical cumulus schemes. Tiedtke et al. (1988) and Tiedtke (1988) describe some of the efforts at ECMWF to improve their cumulus parameterization. Their newest scheme (Tiedtke 1989) uses a cloud-model/environmental-budgets approach somewhat similar to A/S. Improved cumulus parameterization is just an example of the overall trend toward more sophisticated diabatic process in all large NWP models.

NOGAPS 3 development began in the early 1980s, coinciding with FNOC's acquisition of a Control Data CYBER 205 supercomputer. The hardware and software properties of this machine dictate several design compromises. The 205 performance premium obtained by processing memory-contiguous data strings of several thousand elements (long vectors) makes a central memory-resident model necessary. This allows vector lengths equal to the number of grid points in the full horizontal two-dimensional domain of the model, which for the current operational resolution is nearly 30 000 points. Thirty-two-bit floating-point arithmetic is used for additional vector performance and memory conservation. The original version of NOGAPS 3 ran with 18 levels and had a horizontal resolution of 47 zonal wavenumbers with triangular spectral truncation (T47), corresponding to a 2.5° transform grid. The running time for this relatively modest resolution was about 10 min per forecast day. In 1989, NOGAPS 3 horizontal resolution was increased to T79 (NOGAPS 3.2), increasing computer resource requirements nearly threefold. Tight operational run-time schedules and computer memory size make any further resolution

TABLE 1. NOGAPS description.

Computational Details
Resolution: Triangular truncation
—Operational: T79; 18 levels
—Research: T47, T63; 18 levels
Semi-implicit time differencing
Implicit zonal advection
—Vorticity
—Moisture
Silhouette orography
Computer
—CYBER 205: 1988–1991
—CRAY YMP: 1991–Future
Diabatic Processes
Planetary boundary layer
—Explicitly resolved
—Stability-dependent mixing coeff.
—Shallow cumulus
—Predicted ground hydrology
Radiation
—NASA Goddard
—Diurnal cycle
—Predicted fractional clouds
—Ozone heating
Cumulus
—Arakawa–Schubert
—Computationally efficient
—Variational mass flux kernel solution
—Downdrafts
Large-scale precipitation
Gravity-wave drag

increases impractical until the next computer upgrade in 1992/1993.

Navy NWP priorities place some special requirements on NOGAPS 3. Oceanic midlatitude cyclogenesis is an obvious area where NOGAPS is expected to perform well, both for ship routing and tactical operations support. Special effort with sea level pressure bogusing in the NOGAPS data-assimilation system has improved cyclogenesis prediction (Goerss and Pheobus 1991). Extension of this bogusing technique to Pacific typhoons has been so successful that NOGAPS-predicted typhoon tracks are now considered the best guidance available by the Joint Typhoon Warning Center in Guam (personal communication).

An increasingly important NOGAPS role is providing surface forcing to a variety of ocean forecast models. Wind-wave, ocean mixed-layer, and ocean general circulation models all depend on NOGAPS heat, moisture, and momentum fluxes for their atmospheric forcing. The most critical problem with this forcing for any of the ocean models is the systematic error or bias. Most of the NOGAPS research effort in recent years has concentrated on reducing this bias. Surface heat and momentum budgets are quite sensitive to cloud/radiation interactions and planetary boundary-layer (PBL)/cumulus interactions. The bias error is directly related to NOGAPS's ability to accurately predict these budgets, so the design of NOGAPS diabatic processes emphasizes these interactions.

3. NOGAPS diagnostics

The diversity of output parameters available from NWP models continues to grow as the models become more sophisticated and the physics of the models more realistic. Nevertheless, the basic synoptic variables such as sea level pressure and 500-mb height are still the most popular with model users. Objective model evaluation with methods such as root-mean-square error scores or anomaly correlations (Miyakoda et al. 1972) are normally based on these traditional meteorological variables. Subjective forecast verification by operational synoptic meteorologists is likewise oriented around these variables. The reputation of an NWP model like NOGAPS is largely a function of its performance as measured by these objective and subjective means.

The importance of the traditional objective verification methods is not lost on model developers, who use them as an integral part of any model research effort. However, experience has shown that skill score methods are often inadequate for diagnosing the impact of many changes made in the course of model development research (Hogan and Rosmond 1991). Such changes have their greatest impact on the mean properties of the model's atmosphere, i.e., the model climate. Objective and subjective verification, however, usually emphasize the behavior of transient synoptic-scale features in the model, because these are of most interest to model users. Nevertheless, there is little argument between users and model developers that forecast-model systematic error or climate drift is a serious problem. Eventually, this bias error must have a negative impact on the model's synoptic-scale performance, but it can be statistically demonstrated only over a substantial number of forecast cases. The computational investment in such a demonstration is usually only justified at the end of a development effort, when a positive outcome is reasonably assured.

More economical diagnostics that reliably predict model performance are essential for productive NWP model development research. The best diagnostic parameters are those most directly related to the fundamental physical processes in the model: thermodynamic fluxes, precipitation rates, or cloud fractions, for example. A parameter that can also be verified with some form of real data is an added advantage. The earth's radiation budget at the top of the atmosphere was accurately measured by the *Nimbus 7* satellite during the Earth Radiation Budget Experiment (ERBE) (Barkstrom 1984). This data has been of enormous importance for the calibration of radiation parameterizations in global models. Assuming no spurious sources or sinks of heat, any temperature bias in a model should ultimately depend on the balance between incoming solar radiation and outgoing longwave radiation at the top of the model atmosphere. The global mean heat budget of NOGAPS compared to the real atmosphere heat budget (Ramanathan et al. 1989) is summarized in Fig. 1. The real heat budget numbers (parentheses) are in balance, but the NOGAPS numbers, computed directly from the model's physical processes, show a small imbalance. The NOGAPS numbers are averages over several multiple-week simulations, rather than an extended climate simulation, so the model is not in complete heat-budget equilibrium. Each watt per square meter imbalance in the radiation budget at the top of the atmosphere corresponds to slightly less than 0.01 K day^{-1} cooling (heating) bias. This seems trivially small, but through first-guess/analysis data-assimilation feedbacks the bias can accumulate to significant levels. NOGAPS has a global mean cold bias of about 0.1 K, most of which is concentrated in the tropical stratosphere (Fig. 2).

Not every useful diagnostic quantity has the advantage of good verification data. Surface sensible and latent heat flux are extremely difficult to measure in the field, and a global distribution is out of the question. The global mean values for sensible and latent heat flux in Fig. 1 are estimates subject to considerable uncertainty. However, an underprediction of surface fluxes in NOGAPS is consistent with a need for more heat input to correct the cold bias. Also, Fig. 3 shows a 30-day time series of insufficient model-predicted global mean precipitation. The "verifying" global mean precipitation corresponds to the 90 W m^{-2} for surface latent heat flux in Fig. 1, consistent with a balanced

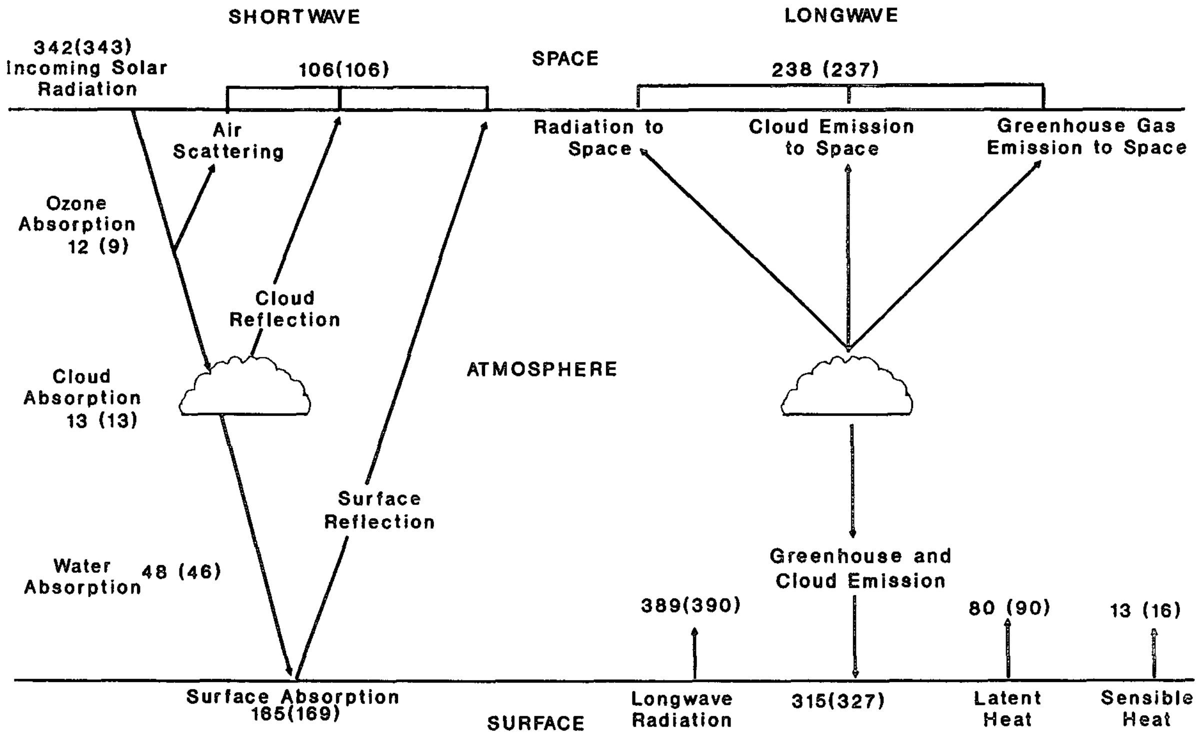

FIG. 1. Terms in the global heat balance for the NOGAPS atmosphere and the real atmosphere (in parentheses). Values are in watts per square meter.

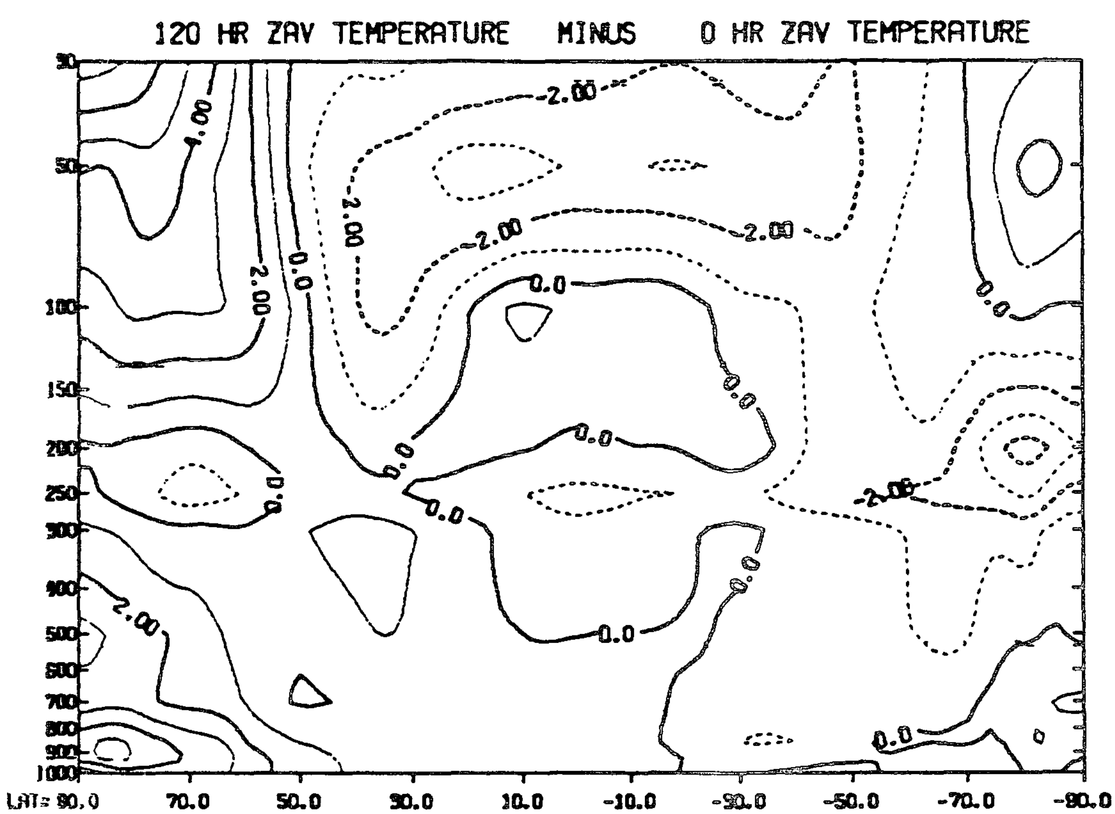

FIG. 2. December 1990 latitude–pressure cross section of NOGAPS 5-day forecast zonally averaged temperature bias. Contour interval is 1°C.

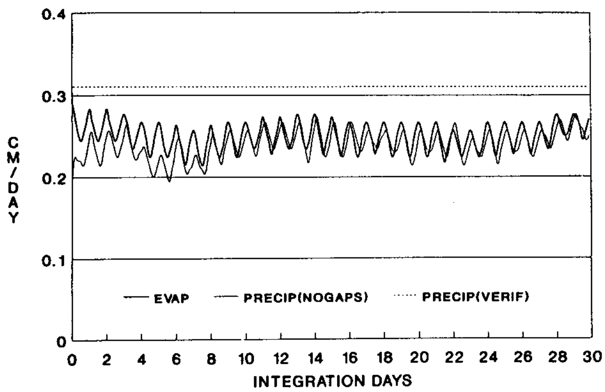

FIG. 3. Thirty-day time series of NOGAPS global mean precipitation and evaporation compared to estimated precipitation in real atmosphere.

hydrological cycle. Each time step, model precipitation and evaporation are summed over every area-weighted model grid point to produce the displayed global means. As a final indicator of the need for a more vigorous hydrologic cycle in NOGAPS, Fig. 4 shows a systematic weakening of the Hadley circulation during the early parts of a NOGAPS NWP forecast. Especially pronounced is the weakening of the ITCZ north of the equator.

Future NOGAPS research will concentrate on changes to model physics to increase the fluxes. Most importantly, the emphasis is on changes to the physics as a whole, without concentrating on the planetary boundary-layer (PBL) parameterization that is responsible for producing the fluxes. Surface fluxes are most sensitive to the temperature difference between the earth's surface and the near-surface air. In NOGAPS, this difference is most sensitive to cloud/radiation interactions, rather than the details of the PBL formulation. This underscores an increasingly prevalent modeling-community view that better cloud/radiation parameterizations are the key to continued model improvements.

4. NOGAPS verification

Comparing the performance of different forecast systems is never a simple procedure. Verification of forecast fields against analysis fields biases the results toward the system that has input into the analysis, because of the incestuous relationship between analysis and forecast models in four-dimensional data assimilation. This is particularly true for large verification domains, such as the Northern Hemisphere, with large, data-sparse oceanic areas. Direct verifications against observations such as rawinsondes eliminates this source of bias, but the results are then geographically biased toward land areas. Nevertheless, the higher quality of land data makes verification there essential, since reduced statistical uncertainty makes the scores more reliable. NOGAPS navy users must ask to what degree land-biased performance measures translate to oceanic areas. Generally, they translate well, especially off the east coasts of North America and Asia, where the data-rich continents upstream positively influence model performance over the oceans, where local data is sparse.

Whatever approach is used, the results are always subject to qualification, depending on the user's interests and prejudices. Nevertheless, there is considerable curiosity in the user community about the relative performance of the large operational forecast systems. In this spirit, Fig. 5 compares the performance of the 48-h/500-mb height forecasts for some of the major operational prediction systems for August–December 1990. The anomaly correlation score is

$$AC = \frac{\sum_n (\phi_n^f - \phi_n^c)(\phi_n^o - \phi_n^c)}{\sum_n ((\phi_n^f - \phi_n^c)^2(\phi_n^o - \phi_c^c)^2)^{1/2}}, \qquad (1)$$

where

n = number of verification points,
ϕ^f = forecast height,
ϕ^o = observed height, and
ϕ^c = climatological height.

It is a measure of the ability of a forecast system to predict departures from climatology; this has become one of the most widely used objective measures of NWP model skill.

A five-point filter is applied to the daily scores to yield smoother time series. Each system is verified against its own analysis to minimize any bias in favor of one system. Forecasts for 120 h are also frequently verified internally by the major operational centers, but either they are not run on the same "watch" at each center or the forecast fields are not available on the global telecommunications system, making similar cross-center comparison impractical.

Examination of the time series shows the consistent superiority of the ECMWF system over the other operational centers. NOGAPS shows some advantage over NMC during the summer and early fall, but then loses this advantage by winter. This systematic behavior has been observed during other years, where by late winter NOGAPS falls below the other models' performance (Jim Goerss, personal communication). Aggressive research to correct this troubling behavior is being conducted, with several changes ready for evaluation during the 1991–1992 winter.

To the author's knowledge, no significant changes were made to any of the operational models during the period shown. However, all major centers routinely make small changes and adjustments that may impact verification statistics, so some caution is advised in drawing conclusions, since by the time this report ap-

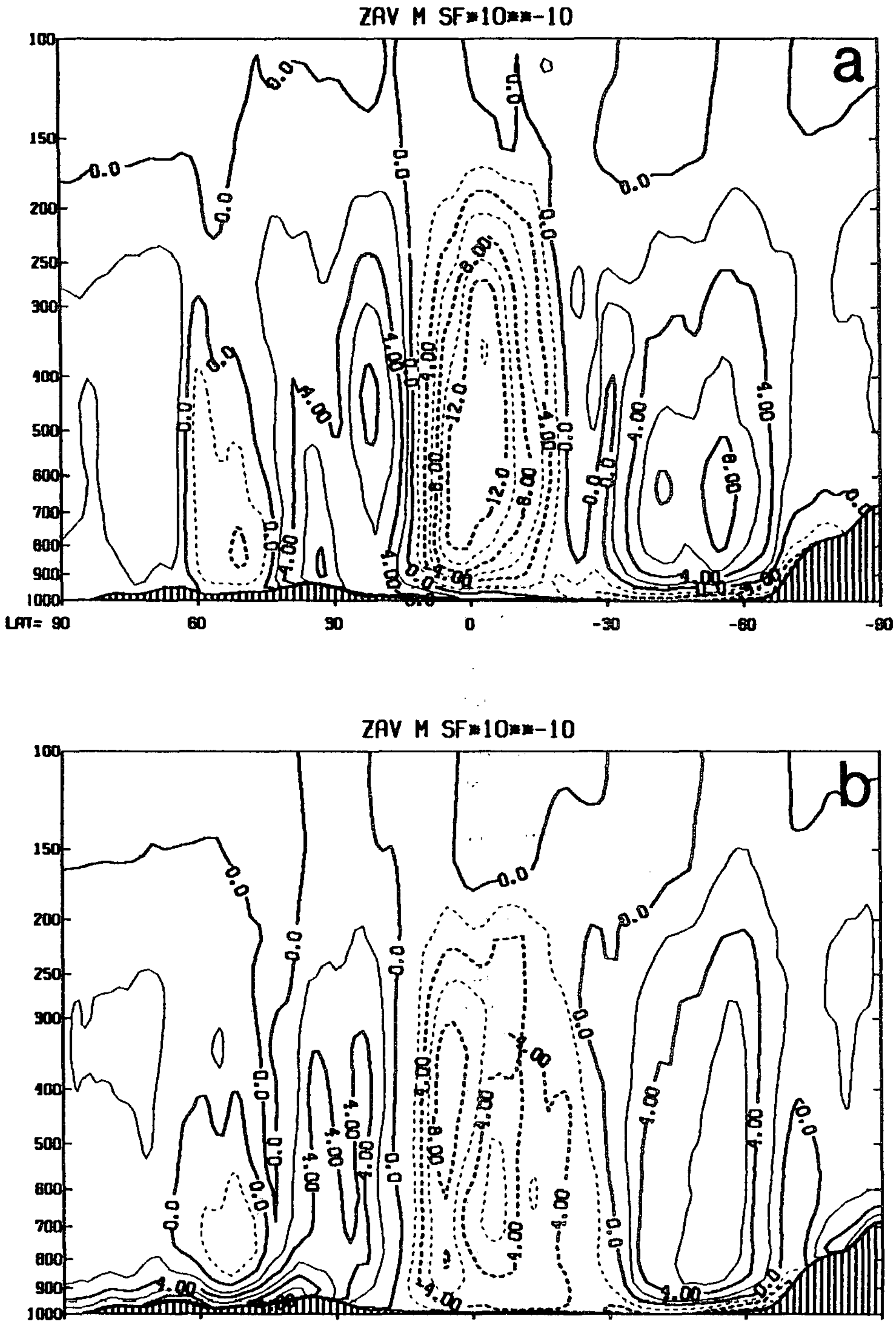

FIG. 4. NOGAPS zonally averaged meridional streamfunction for (a) initial conditions, and (b) 36-h forecast. Contour interval is $2 \times 10^{10}\ s^{-1}$.

pears in print the comparisons could be quite different. Nevertheless, NOGAPS is quite competitive with the other models and ranks among the top four or five in the world.

Readers interested in some of the special verification methods used for NOGAPS verification over the oceans are directed to Hogan and Rosmond (1991), Clune et al. (1991), and the FNOC *Quarterly Performance*

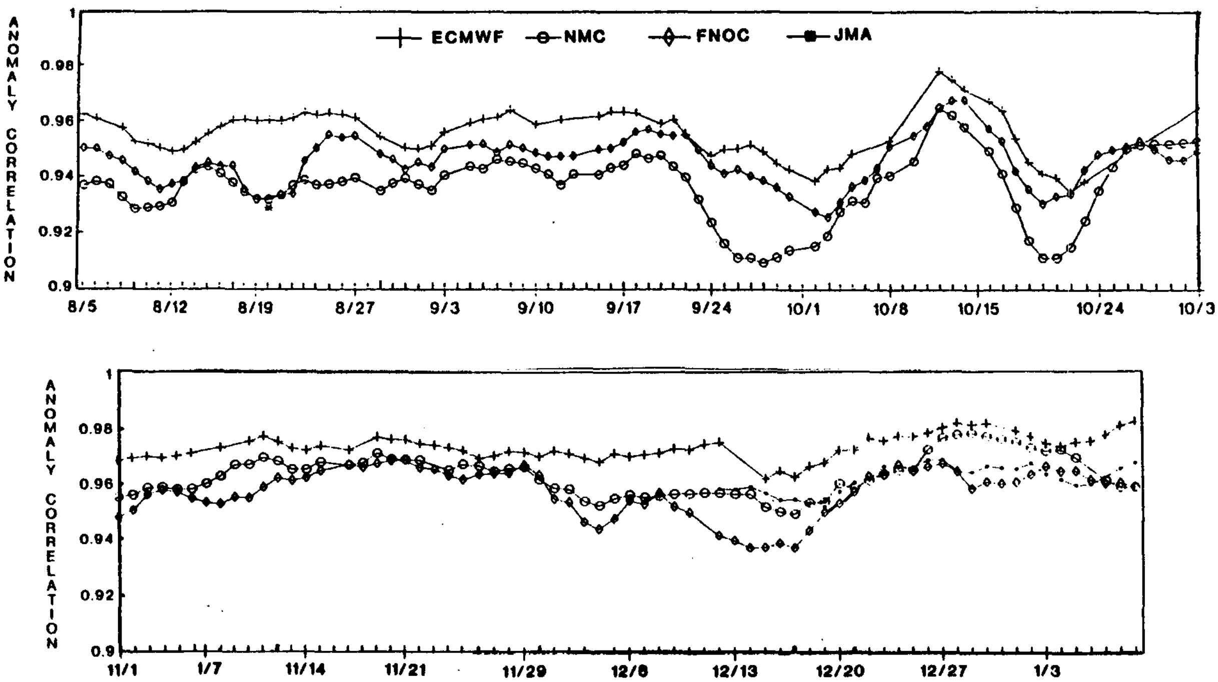

FIG. 5. Time series of five-point averaged 48-h Northern Hemisphere anomaly correlation scores for major operational centers from early August 1990 to early January 1991. JMA (Japanese) scores became available late in period.

Summaries. NOGAPS storm track climatology and cyclone deepening and filling behavior are emphasized.

5. The spinup problem

A largely misunderstood characteristic of a large global forecast system like NOGAPS is that it is more than the sum of its parts. The individual components of the system are 1) a sophisticated observation-data quality control, 2) the optimum interpolation analysis, 3) the normal mode initialization, and 4) the forecast model. In a carefully designed system, the performance of any individual component is enhanced by its interactions with the other components. For example, a major part of the NOGAPS development process involves running the system in a 4D data-assimilation mode for several days to test the impact of system changes. These experiments usually require that the system be "cold started"; that is, initialized from a set of archived fields that may be from a previous NOGAPS analysis or perhaps from a research dataset such as from the First GARP Global Experiment (FGGE). As an experiment proceeds, the error level of each first-guess forecast, which is usually either 6 or 12 h, will decrease for several cycles until it reaches an equilibrium level significantly less than the forecast error for the first forecast in the sequence. Figure 6 shows an example of the decrease in error growth rate of the first-guess forecasts as a data-assimilation experiment proceeds. It is related to the discoveries of Charney et al. (1969), who first observed the ability of a forecast model to assimilate "observations" and become conditioned to these data in such a way that the model becomes more skillful at predicting future states of the atmosphere from which more observations are taken. Bengtsson (1975) describes in considerable detail this update-cycle process, which is now well known to all who develop 4D data-assimilation systems. Update-cycle spinup is quite different from the spinup of forecast-model moist processes discussed below. Update-cycle spinup has important implications for the evaluation of forecast system performance and intercom-

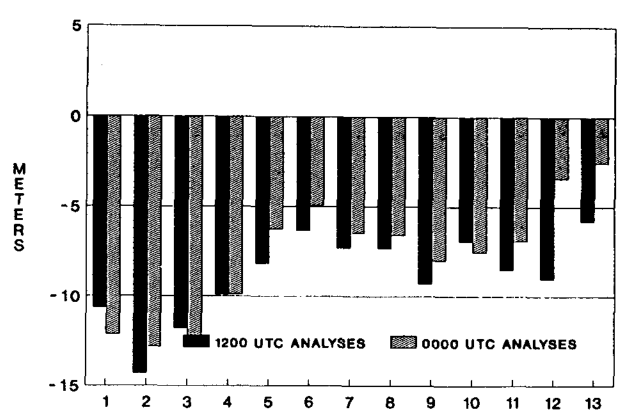

FIG. 6. Example of update-cycle spinup, showing reduction in 500-mb mean error (cold bias) of 12-h first-guess forecasts after data-assimilation "cold start."

parisons between systems. NOGAPS users frequently request that interesting forecast situations be rerun from historical data. Usually these requests occur after a significant system upgrade, and users want to see how the new system performs in a situation where a previous version fared poorly. The results of such a rerun can have considerable impact on the NOGAPS image to its users and navy sponsors.

Not surprisingly, NOGAPS developers are reluctant to engage in a case rerun unless the system is going to perform under optimum conditions. This is certainly not going to be true if the test case is run in isolation, since it would correspond to the first forecast in a sequence like that of Fig. 6. The alternative is to assemble the necessary observational data from several days prior to the time of the case of interest. Then, enough data-assimilation cycles can be run to ensure model error growth rate is minimized when the initial conditions for the case of interest are produced. For a global forecast system such as NOGAPS, the computational cost and data-management logistics for conducting such an effort are far greater than just rerunning a single forecast from a set of initial analysis fields. Only if an important research question can be combined with the forecast case study is such a large investment of researchers' time and computer time justified.

Paradoxically, update-cycle spinup also complicates model intercomparisons when each model is run from the same initial fields. The degree of incompatibility between the initial fields and each model varies considerably. The greater the incompatibility, the poorer a model will perform relative to its potential when it is able to spin up with its own update cycle. The original FGGE IIIB fields produced by ECMWF are a case in point. The ECMWF forecast model that produced these fields used a version of the Kuo (1974) cumulus-parameterization scheme that introduced a moist bias, particularly in the tropical mixed layer (Tiedtke et al. 1979). NOGAPS, on the other hand, uses the A/S cumulus scheme, which produces dry tropical biases (Hogan and Rosmond 1991). NOGAPS performed quite poorly even in the middle latitudes when run from these FGGE IIIB fields because of an excessively strong Hadley circulation as the model vigorously rained out the excess moisture sensed by the A/S scheme. Running the NOGAPS data-assimilation system from FGGE IIB observations, however, allowed the update cycle to spin up and the model moisture field to come into equilibrium with its cumulus scheme. These NOGAPS FGGE forecasts were significantly better than those run from the IIIB fields.

A more familiar spinup problem is the recovery from initialization "shock" that characterizes every NWP model during the first few hours of a forecast (Krishnamurti et al. 1991). Moist physics are extremely sensitive to small imbalances and inconsistencies in the initial conditions. Figure 7 shows the typical evolution of global mean precipitation and evaporation in a NOGAPS forecast. This behavior is characteristic of many global forecast models. In recent years, advances such as the analysis and initialization of increments and more realistic model physics have reduced the problem considerably. From an operational point of view, it is largely a cosmetic problem, except for very short-term precipitation forecasts, normally not a global forecast system requirement. However, the continued presence of moist process spinup suggests there is still room for improvement in model initialization and model physics. Spinup is one indicator of the quality of a forecast system, and if our model research efforts can reduce it, progress is being made.

6. The climate drift problem

Section 3 discussed the NOGAPS global heat balance in the context of model temperature bias. More accurate prediction of the various terms of the global heat budget, and the geographical variation of these terms, is certainly one of the great challenges for modelers in the future. However, the climate drift problem has other equally important and challenging components. All global models suffer to some degree from lack of amplitude in planetary-scale and synoptic-scale waves (i.e., an excessively high zonal index). Envelope topography (Wallace et al. 1983) and gravity-wave drag (Palmer et al. 1986) are mechanisms used in many models to counter this bias. In spite of legitimate physical arguments in favor of such methods, there seem to be more fundamental deficiencies in the formulation of all models that cause the problem. For example, there is evidence that vertical discretization methods may contribute (Arakawa and Moorthi 1988) by allowing computational-scale baroclinic instability. Whatever the cause, progress toward improved medium-range forecasts is dependent on better solutions to this climate drift problem.

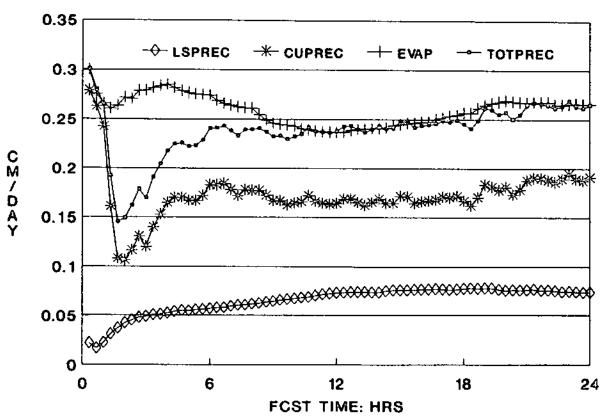

FIG. 7. Moist physics spinup after initialization, showing typical behavior of global mean evaporation, cumulus precipitation, and large-scale precipitation for early hours of a forecast.

Interest within the scientific community in global climate change has stimulated development of coupled atmosphere–ocean models. There is no doubt that understanding atmosphere–ocean interaction, and our ability to successfully model these interactions, are critical to predicting any changes in global climate in the next century. Likewise, prospects for extended-range deterministic forecasting for a few weeks or months depend on successful atmosphere–ocean models. In addition, the navy has an interest in coupled atmosphere–ocean models, because of its interest in deterministic ocean prediction (i.e., NWP for the ocean) (Clancy and Sadler 1992). NOGAPS is a critical component in the development of the navy's coupled global atmosphere–ocean model.

To a much greater degree than the atmosphere, ocean NWP depends on accurate boundary conditions at the air–sea interface. These boundary conditions are the fluxes of heat, moisture, and momentum derived from the NOGAPS PBL parameterization. Ocean models are extremely sensitive to biases in these fluxes. The systematic underprediction of surface heat fluxes diagnosed from the NOGAPS global heat budget has also been confirmed by the response of the navy's thermodynamic ocean prediction system (TOPS) (Clancy and Martin 1981) to these fluxes. NOGAPS and TOPS have been coupled as a prototype coupled atmosphere–ocean. Figure 8 shows the 10-day, mid-January, western-Pacific sea surface temperature change predicted by the coupled system, compared to the observed change. Overall, the predicted cooling north of the tropics is about double what it should be. These biases limit the potential of coupled models as indicators of any climate-change signal that may be present in the real atmosphere. To date, the only recourse for controlling these biases has been ad hoc methods such as adding flux-correction terms to the sea surface heat balance (Gallimore and Houghton 1987). This is clearly not a satisfactory solution to the climate drift problem. Ironically, however, the obvious nature of the sea surface temperature biases gives modelers reason for optimism that the climate drift problem can be solved. Global sea surface temperature is relatively well observed on a real-time basis. This, combined with its sensitivity to the surface heat balance, makes it an ideal diagnostic indicator of coupled model performance. No other physical parameter in the atmosphere or ocean combines these two properties as well. Therefore, coupled atmosphere–ocean models predicting SST change will be essential for research into the climate drift problem. Also, extended-range (e.g., 30-day) deterministic forecasting is now being proposed (Rosati and Miyakoda 1988), and interactive atmosphere/mixed-layer models will likewise be critical to progress in this area. Operational application of such models may still be far off, but the research certainly can begin now.

7. Future NOGAPS plans

The preceding section describes the navy's motivation for coupled atmosphere–ocean model research. NOGAPS development over the next decade will certainly concentrate on the goal of useful coupled atmosphere–ocean forecasts, regardless of the obvious difficulties. Any progress toward reduction of systematic model error will directly benefit NOGAPS in its traditional role of global NWP guidance for fleet users.

A natural complement to NOGAPS development for its operational mission is the navy's role in the national global climate-change program. The navy has been designated the lead Department of Defense agency for climate-change modeling, and plans are being made for an active NOGAPS rôle. The emphasis will be on the extended-range forecast and interseasonal time scales, rather than long-term climate simulation. In addition to using the coupled NOGAPS as a platform for improved understanding of atmosphere–ocean interactions, the climate drift problem will be studied as a critically important research problem in its own right. This is a natural niche for NOGAPS because of its access to the navy's environmental database, which, particularly for oceanographic data, may be the most complete in the world. Multiweek and multimonth integrations of a coupled atmosphere–ocean NOGAPS and a fixed SST NOGAPS will assess the sensitivity of the model climate to changes in model physics, boundary conditions, resolution, and numerical methods. The diagnostic tools demonstrated above will allow us to dissect the physical processes of NOGAPS, particularly at the air–sea interface, where model bias error is condensed to a relatively simple set of parameters. The history of numerical model development tells us that there will be no dramatic breakthrough in research, just incremental progress toward better models. This is sure to be the case with future NOGAPS research.

8. Summary and final thoughts

This report gives readers a flavor of how large NWP models such as NOGAPS are developed, what the priorities are, and what compromises must be made to satisfy operational constraints. An equally important message is that a model such as NOGAPS is far more than a group of algorithms extracted from textbooks and journal articles, linked together in some arbitrary way. Just as the interactions between physical processes in the real atmosphere are often too complicated to understand, the interactions of the modeled physical processes are equally complicated. The "art" of numerical modeling is combining understanding of what happens in the real atmosphere with what happens in the numerical model. If this were not the case, model development would be no more challenging than a graduate-level homework problem.

The initial model development is only the beginning

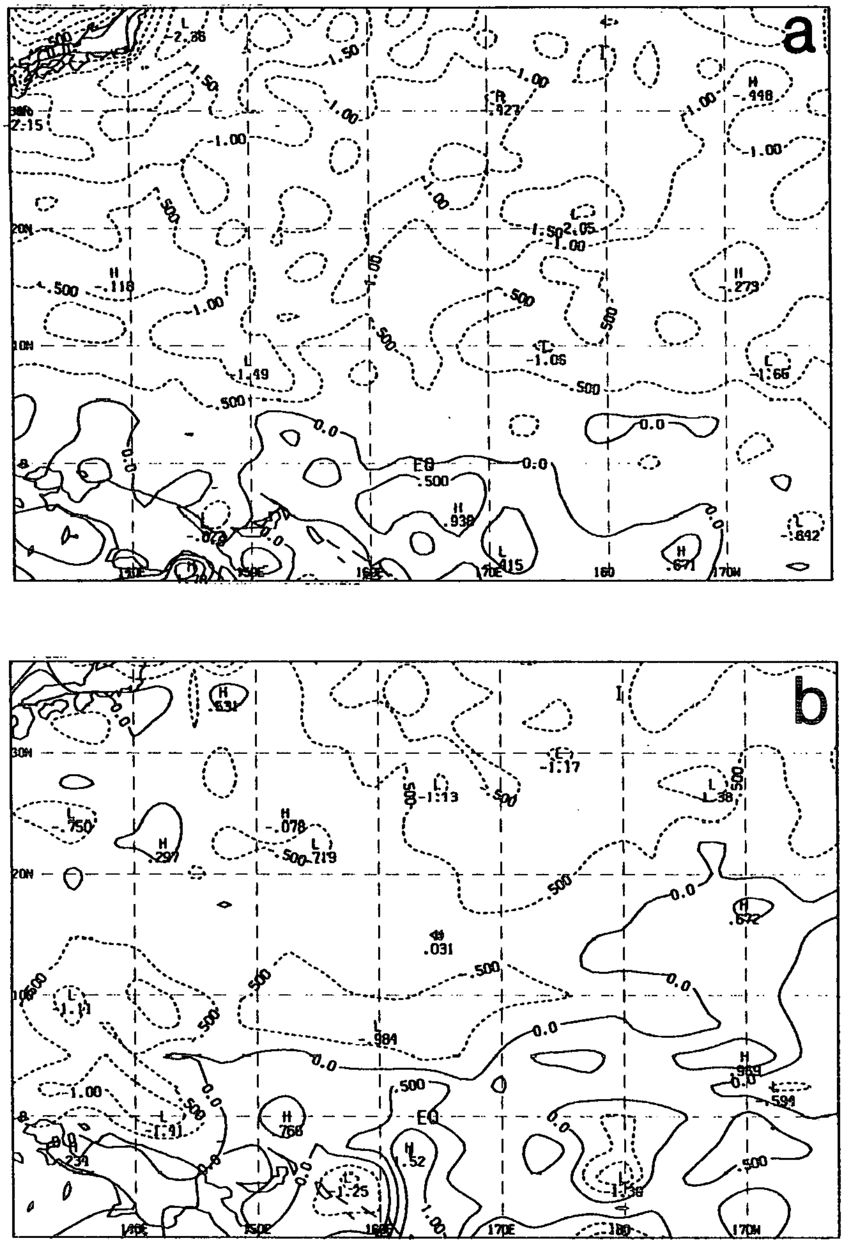

FIG. 8. Mid-January, 10-day SST change for western Pacific: (a) predicted by coupled NOGAPS/TOPS, (b) observed change. Contour interval 0.5°C.

of this subtle combination of the art and science of modeling. It is a common experience at the large NWP centers for operational models to improve with age. This occurs because, over their life, the models are under intense scrutiny, their physical parameterizations constantly being refined and "tuned" based on physical process diagnostics. Taken in isolation, each change may have insignificant impact, but the cumulative effect of many such small improvements, averaged over the enormous data sample size an operational NWP model provides, eventually pays off in significant improvement. The true challenge of model development is understanding a model well enough to make these improvements. The future role of NOGAPS in support of navy requirements for meteorological and oceanographic research dictates that we continue to meet this challenge.

Acknowledgments. The author would like to thank Dr. Tim Hogan and Dr. Ron Gelaro of NOARL for providing several of the figures shown, Mr. Robin Brody of NOARL for assisting with some of the verification data, and Prof. Chuck Wash of the Naval Postgraduate School for his encouragement in getting the manuscript prepared.

The support of the sponsors, Office of Naval Research, Mr. R. Peloquin, program manager, program element 63704N; and Office of Naval Technology, Mr. J. Cauffman, program manager, program element 62435N, is gratefully acknowledged. NOARL contribution no. 432.023.91. Approved for public release; distribution is unlimited.

REFERENCES

Arakawa, A., and V. R. Lamb, 1977: Computational design of the basic dynamical processes of the UCLA general circulation model. *Methods in Computational Physics, Vol. 17,* Academic Press, Inc., 173–265.

——, and S. Moorthi, 1988: Baroclinic instability in vertically discrete systems. *J. Atmos. Sci.,* **45,** 1688–1707.

Barker, E. H., 1982: A comparison of two initialization methods in data assimilation. Naval Environmental Prediction Research Facility TR 82-06, 159 pp. [Available from National Technical Information Service, U.S. Dept. of Commerce, Springfield, VA 22161.]

——, J. Goerss, and N. Baker, 1988: The navy's operational multivariate optimum interpolation analysis method. Preprints, *Eighth Conf. on Numerical Weather Prediction,* Baltimore, Amer. Meteor. Soc., 161–163.

Barkstrom, B. R., 1984: The Earth Radiation Budget Experiment (ERBE). *Bull. Amer. Meteor. Soc.,* **65,** 1170–1185.

Bengtsson, L., 1975: Four-dimensional assimilation of meteorological observations. GARP publication series no. 15, Geneva, World Meteorological Organization, 76 pp.

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.

Clancy, R. M., and P. J. Martin, 1981: Synoptic forecasting of the oceanic mixed layer using the navy's operational environmental database: Present capabilities and future applications. *Bull. Amer. Meteor. Soc.,* **62,** 770–784.

——, and W. D. Sadler, 1992: The Fleet Numerical Oceanography Center suite of oceanographic models and products. *Wea. Forecasting,* **7,** 307–327.

Clune, W. M., P. A. Harr, and L. R. Brody, 1991: The man–machine mix in operational product quality control and verification at Fleet Numerical Oceanography Center. *Wea. Forecasting,* **7,** 280–287.

FNOC, 1986–1990: *Quarterly Performance Summary.* Volumes 1–5, Nos. 1–4. [Available upon request from Commanding Officer, Fleet Numerical Oceanographic Center, Monterey, CA 93943-5005.]

Gallimore, R. G., and D. D. Houghton, 1987: Approximation of ocean heat storage by ocean–atmosphere energy exchange: Implications for seasonal cycle mixed layer ocean formulations. *J. Phys. Oceanogr.,* **17,** 1214–1231.

Gelaro, R., 1989: The structure and dynamics of tropical–midlatitude interactions. Ph.D. thesis, The Pennsylvania State University, 227 pp. [Available from The Pennsylvania State University, Dept. of Meteorology, 503 Walker Building, University Park, PA 16802.]

Goerss, J. S., and P. A. Phoebus, 1991: The navy's operational atmospheric analysis. *Wea. Forecasting,* **7,** 232–249.

Harr, P. A., R. L. Elsberry, T. F. Hogan, and W. M. Clune, 1991: North Pacific cyclone sea-level pressure errors with the Navy Operational Global Atmospheric Prediction System. Preprints, *Ninth Conf. on Numerical Weather Prediction,* Denver, Amer. Meteor. Soc., 746–749.

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.

——, ——, and R. Gelaro, 1992: The NOGAPS forecast model: A technical description. NOARL Report 013, 212 pp. [Available from National Technical Information Service, U.S. Dept. of Commerce, Springfield, VA 22151.]

Krishnamurti, T. N., J. Xue, H. S. Bedi, K. Ingles, and D. Oosterhof, 1991: Physical initialization for numerical weather prediction over the tropics. *Tellus,* **43A–B,** 53–81.

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.

Lorenc, A. C., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.,* **109,** 701–721.

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model, with application to normal model initializations. *Contrib. Atmos. Phys.,* **50,** 253–271.

Miyakoda, K., G. D. Hembree, R. F. Strickler, and I. Shulman, 1972: Cumulative results of extended forecast experiments. Part I: Model performance for winter cases. *Mon. Wea. Rev.,* **100,** 836–855.

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.

Ramanathan, V., B. R. Barkstrom, and E. F. Harrison, 1989: Climate and the earth's radiation budget. *Phys. Today,* **42,** 22–32.

Rosati, A., and K. Miyakoda, 1988: A general circulation model for upper ocean circulation. *J. Phys. Oceanogr.,* **18,** 1601–1626.

Rosmond, T. E., 1981: NOGAPS: Navy Operational Global Atmospheric Prediction System. Preprints, *Fifth Conf. on Numerical Weather Prediction,* Monterey, Amer. Meteor. Soc., 74–79.

Tiedtke, M., 1988: Parameterization of cumulus convection in large-scale models. *Physically-Based Modeling and Simulation of Climate and Climate Change,* M. Schlesinger, Ed., D. Reidel, 375–431.

——, 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Wea. Rev.,* **117,** 1779–1800.

——, J. F. Geleyn, A. Hollingsworth, and J. F. Louis, 1979: ECMWF ymodel-parameterization of sub-grid scale processes. ECMWF Tech. Report No. 10, 45 pp. [Available from the European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading RG2 9AX, U.K.]

——, W. A. Heckley, and J. Slingo, 1988: Tropical forecasting at ECMWF: On the influence of physical parameterization on the mean structure of forecasts and analyses. *Quart. J. Roy. Meteor. Soc.,* **114,** 639–664.

Wallace, J. M., and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.,* **109,** 785–812.

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