Operational Performance of a New Barotropic Model (WBAR) in the Western North Pacific Basin

CHARLES R. SAMPSON AND JAMES S. GOERSS

Naval Research Laboratory, Monterey, California

HARRY C. WEBER

University of Munich, Munich, Germany

(Manuscript received 30 September 2004, in final form 30 June 2005)

ABSTRACT

The Weber barotropic model (WBAR) was originally developed using predefined 850–200-hPa analyses and forecasts from the NCEP Global Forecasting System. The WBAR tropical cyclone (TC) track forecast performance was found to be competitive with that of more complex numerical weather prediction models in the North Atlantic. As a result, WBAR was revised to incorporate the Navy Operational Global Atmospheric Prediction System (NOGAPS) analyses and forecasts for use at the Joint Typhoon Warning Center (JTWC). The model was also modified to analyze its own storm-dependent deep-layer mean fields from standard NOGAPS pressure levels. Since its operational installation at the JTWC in May 2003, WBAR TC track forecast performance has been competitive with the performance of other more complex NWP models in the western North Pacific. Its TC track forecast performance combined with its high availability rate (93%–95%) has warranted its inclusion in the JTWC operational consensus. The impact of WBAR on consensus TC track forecast performance has been positive and WBAR has added to the consensus forecast availability (i.e., having at least two models to provide a consensus forecast).

1. Introduction

The Weber barotropic model (WBAR), a tropical cyclone (TC) track prediction model (Weber 2001), was originally developed using predefined 850–200-hPa deep-layer mean analyses and forecasts (DLMs) from the National Centers for Environmental Prediction's (NCEP) Global Forecasting System (GFS) as initial and boundary conditions. WBAR was found to produce more skillful TC track forecasts than other barotropic models in the North Atlantic, and its TC track forecast performance was competitive with that of more complex numerical weather prediction (NWP) models. Much of the credit for its superior performance was attributed to the careful removal of unwanted features such as mislocated weak vortices (Weber 2001).

To develop a fully operational version for the Joint Typhoon Warning Center (JTWC), WBAR was modified to use gridded data from the Navy Operational Global Atmospheric Prediction System (NOGAPS) as initial and boundary conditions. It was suspected that construction of storm-dependent DLMs (i.e., mass-weighted vertical averages of each field provided on standard pressure levels) instead of fixed DLMs as in the experimental version (Weber 2001) might improve the TC track forecast performance of the WBAR. Therefore, a method was developed to determine variable DLMs for the U.S. Navy version of WBAR on the basis of a statistical evaluation. For initial implementation in 2003, the 2002 performance of all possible realizations of WBAR (sets of runs using all NOGAPS single-level analyses and forecasts and all possible DLM combinations thereof) was assessed for different storm parameters (latitude, longitude, intensity, direction of motion, storm translation speed, radius of maximum wind speed, radius of outermost closed isobar, and date). The statistical evaluation was repeated for the 2003 season, and WBAR storm-dependent DLMs were updated in 2004.

The number of NWP models capable of producing

Corresponding author address: Charles R. Sampson, NRL, Mail Stop 2, 7 Grace Hopper Ave., Monterey, CA 93943-5502.
E-mail: sampson@nrlmry.navy.mil

high quality tropical cyclone track forecasts has grown in recent years. Currently, there are nine NWP models that routinely produce skillful track forecasts in the western North Pacific basin for operational use at JTWC. Three of these models are run operationally at the Fleet Numerical Meteorology and Oceanography Center: NOGAPS (Hogan and Rosmond 1991; Goerss and Jeffries 1994), the Geophysical Fluid Dynamics Laboratory (GFDL) Hurricane Prediction System (Kurihara et al. 1993, 1995, 1998; Rennick 1999), and the Coupled Ocean–Atmosphere Mesoscale Prediction System (COAMPS[1]; Hodur 1997). Two models are run operationally at the Japan Meteorological Agency (Kuma 1996): the global spectral model and the typhoon model. The remaining four models are the Met Office (UKMO) global model (Cullen 1993; Heming et al. 1995), the NCEP global spectral model (GFS; Lord 1993), the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5; Grell et al. 1995) run operationally by the Air Force Weather Agency (AFWA), and the TC-Limited Area Prediction System (TC-LAPS; Davidson and Weber 2000) run by the Australian Bureau of Meteorology. At the JTWC these NWP models are used to form a consensus as described in Goerss et al. (2004), which serves as a track forecast baseline. This paper discusses WBAR TC forecast performance relative to the nine NWP models and WBAR's impact on the consensus forecasts.

2. Methods

The operational version of WBAR was installed as part of the suite of models run on the Automated Tropical Cyclone Forecast System (ATCF; Sampson and Schrader 2000) and consists of a set of modules: a bogus preprocessor, an initialization package, a DLM analysis package for the NOGAPS analyses and forecasts, and finally the barotropic forecast model itself. Except for the construction of variable DLMs from global 1° spherical grid analyses and forecasts from NOGAPS (zonal and meridional wind components and geopotential height) on six standard levels (850, 700, 500, 300, 250, and 200 hPa) at 12-h intervals out to 72 h discussed in section 1 and some minor modifications that address the different computational environment and the use of NOGAPS instead of GFS data, the methods used in the current model correspond with those described in Weber (2001). One extra set of analyses is used from the NOGAPS run 12 h prior to the current forecast for the computation of a smooth temporal boundary condition at the initial time of the forecast. For construction of the variable DLMs in the DLM analysis package, the model requires the initial tropical cyclone information described in section 1.

The NOGAPS data are extracted from a relational database: the Tactical Environmental Data Server (TEDS; Naval Research Laboratory 2004). The TEDS retrieval is executed once for each warning cycle. The WBAR model itself, however, must be run for each individual tropical cyclone. On a typical 2004-vintage Linux workstation, the complete 72-h WBAR forecast is executed in about 1 min. WBAR runs after the NOGAPS model run is complete and the NOGAPS grids are sent to the TEDS at JTWC. Like the NOGAPS forecast, the WBAR track forecast is then available to the forecasters 6 h later. For the WBAR track forecasts to be of the greatest use to the operational forecast centers, they must be adjusted to appear as if they are for the current synoptic time.

JTWC has procedures in place to move the time-late (6 or 12 h late) forecasts to the current time by running an "interpolator." NWP model tracks are first interpolated to intermediate times, and then interpolated positions are relocated to reflect the forecaster-analyzed (best track) position. The version of the interpolator used in this study is similar to that described in Goerss et al. (2004) with one exception—the cubic spline interpolation has been replaced by linear interpolation, which was found to yield slightly lower forecast errors. The following are the names of the interpolated tracks: WBAI for WBAR, NGPI for NOGAPS, EGRI for the UKMO global model, JAVI for the NCEP GFS, JGSI for the Japanese global spectral model, JTYI for the Japanese typhoon model, GFNI for the GFDL Hurricane Prediction System, COWI for COAMPS, AFWI for the AFWA MM5, and TCLI for the TC-LAPS.

An added advantage in producing interpolated tracks is that they can then be used to form a real-time consensus. The consensus methods described in this paper are simple averages of the members described in the previous paragraph. An attempt is made to compute a consensus forecast at each forecast period (12, 24, 36, 48, and 72 h). A consensus is computed if *two or more* members exist for a given forecast period. If less than two members exist, the consensus is not computed. In the current operational tropical cyclone forecasting climate at JTWC, one test of the utility of a model like WBAR is whether it either adds to the performance or forecast availability of a consensus.

In this paper, the results presented are from recomputed interpolations and consensus forecasts using

[1] COAMPS is a registered trademark of the Naval Research Laboratory.

methods described above and operational input. The purpose of this is to ensure that all the interpolated results are computed using the same version of the interpolator, and to produce nine-model consensus results not produced in operations. Average differences in performance between recomputed interpolations and those produced in operations are less than 1%. Nine-model consensus forecasts (consensus forecasts with one model removed) are produced to examine the effect of individual models on the operational 10-model consensus (WBAI, NGPI, EGRI, JAVI, JGSI, JTYI, GFNI, COWI, AFWI, and TCLI). For example, the first nine-model consensus, formed to examine the effect of WBAI on the 10-model consensus, includes NGPI, EGRI, JAVI, JGSI, JTYI, GFNI, COWI, AFWI, and TCLI.

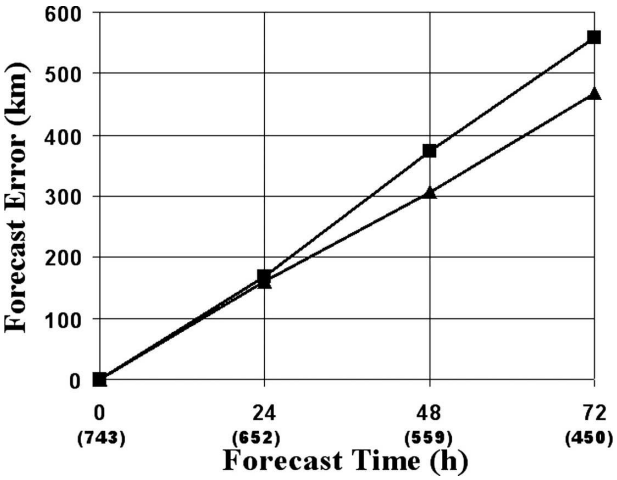

FIG. 1. The 24-, 48-, and 72-h errors (km) of WBAI (triangles) and the operational CLIPER (squares) for the period 9 Jun 2003–7 Jul 2004. The number of forecasts is shown in parentheses.

3. Results and conclusions

The WBAR model was installed, as described in the previous section, in late May 2003. However, the ATCF installation for 2003 was a major one and the WBAR model was not executed routinely until approximately 9 June 2003. Hence, the evaluation period covered in this paper is 9 June 2003–7 July 2004. Only forecasts for which there are verifying JTWC forecasts are evaluated and the tracks for the period are preliminary best tracks; the final best tracks for this period were not complete at the time this paper was submitted. Figure 1 shows the WBAI forecast errors for the period. Included in the figure is a homogeneous comparison with the Climatology and Persistence forecasting scheme (CLIPER; Neumann 1992) at JTWC, which serves as a baseline for skill in track forecasting. The CLIPER used here is the operational CLIPER. As seen in the figure, the WBAI performance shows skill at the longer forecast periods. The results for the 72-h period confirm the findings of Weber (2001) that barotropic models can produce valuable track guidance beyond forecast periods of 48 h.

Of greater interest is a comparison of WBAI performance with the nine NWP models discussed in the previous section, as shown in Fig. 2 for the 24-, 48-, and 72-h prediction times. For the period of record, WBAI is neither the top nor the bottom performer for any of the three forecast periods. The worst performer for all three periods is AFWI.

One method of determining a model's impact on the 10-model consensus is to compare results of a 9-model consensus (without the model of interest) with those of the 10-model consensus. Figure 3 shows head-to-head comparisons of the forecast errors of each consensus using 9 models with the forecast errors of the 10-model consensus. The results are limited to forecast periods for which the model of interest is available. For example, TCLI was only available 212 times for the 24-h period while WBAI was available 645 times. The impact of each model on the consensus can then be shown as the ratio of the errors of the 9-model consensus forecast errors to the 10-model consensus forecast errors.

At all forecast lengths we see that WBAI has one of the largest positive impacts upon the consensus forecasts despite the fact that its forecast errors are considerably larger than those of some of the other models (Fig. 2). Goerss (2000) found that the consensus forecast error depends on two things: 1) the mean forecast error of the individual models that constitute the consensus and 2) the degree of independence (or the effective degrees of freedom) of the forecast errors of the individual models.

To get a simple estimate of the independence of the individual models, we expand on the theoretical background described in Goerss (2000). Forecast position error E_i for model i is defined to be

$$E_i = (C_i^2 + A_i^2)^{1/2}, \quad (1)$$

where C_i and A_i are the across-track and along-track errors, respectively. For simplicity, assume that, for every i, C_i, and A_i are independent and normally distributed with zero mean and a constant standard deviation σ. Then it follows that E_i possesses a Rayleigh distribution (Lindgren 1976) with mean

$$\mu = \sigma(\pi/2)^{1/2}, \quad (2)$$

where σ is the standard deviation of the cross- and along-track error distributions. Because a consensus

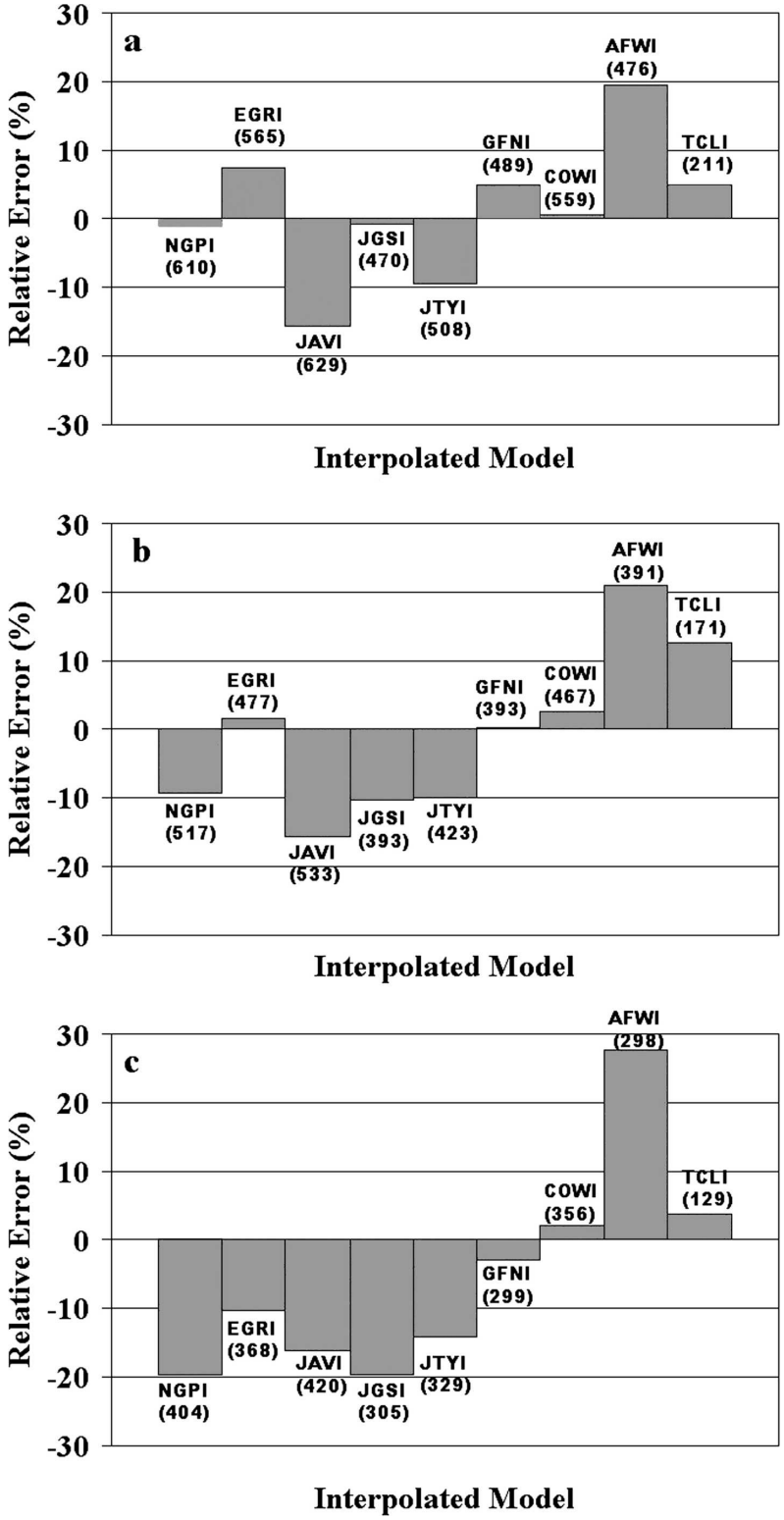

FIG. 2. Average forecast errors relative to the WBAI (%) for each interpolated NWP model forecast and for the period 9 Jun 2003–7 Jul 2004. Forecast errors shown are at (a) 24, (b) 48, and (c) 72 h. Interpolated NWP models are as defined in the text. Positive values indicate errors larger than WBAI. The number of forecasts is shown in parentheses.

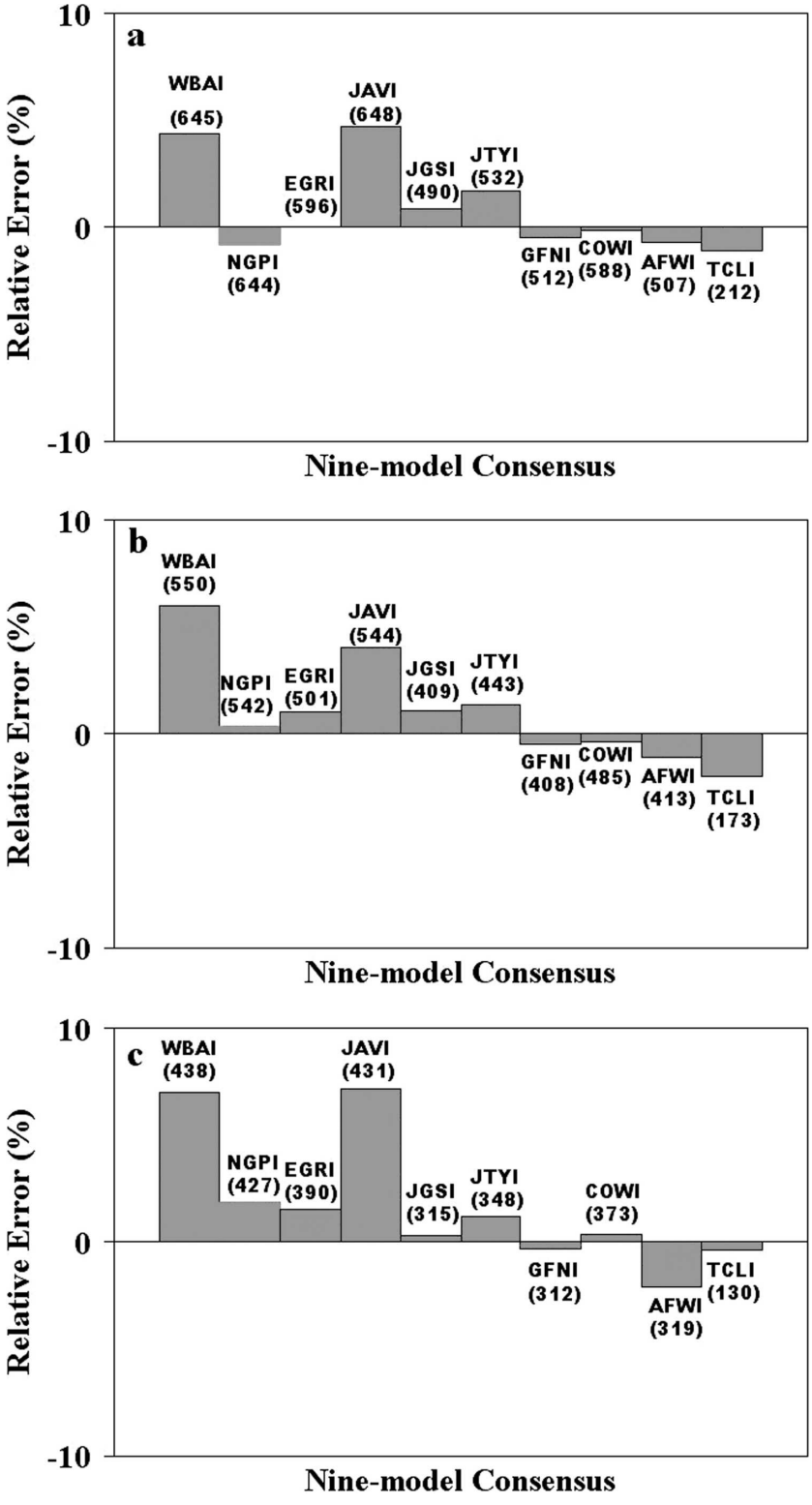

FIG. 3. Average forecast errors relative to the 10-model consensus (%) for each consensus consisting of nine models. Nine-model consensus errors are labeled with the model that was removed from the consensus. Forecast errors shown are at (a) 24, (b) 48, and (c) 72 h for the period 9 Jun 2003–7 Jul 2004. Positive values show the positive impact of the labeled model on the 10-model consensus. The number of forecasts is shown in parentheses.

forecast position is the mean of the individual model forecast positions, the consensus across- and along-track errors, denoted by C_c and A_c, respectively, are simply the means of the across- and along-track errors of the individual models. Therefore, for a consensus with n members, C_c and A_c are normally distributed (Hoel 1962) with zero mean and a standard deviation σ_c defined as

$$\sigma_c = \sigma/n^{1/2}. \quad (3)$$

The consensus error E_c also possesses a Rayleigh distribution with mean

$$\mu_c = \sigma_c(\pi/2)^{1/2}. \quad (4)$$

Substituting the definition of σ_c in Eq. (3) into Eq. (4), we get

$$\mu_c = \sigma(\pi/2n)^{1/2}. \quad (5)$$

In practice, the along- and cross-track errors of the models are not independent, so we replace n in Eqs. (3) and (5) with n_e, the effective degrees of freedom. If the forecast errors of the individual models were totally uncorrelated, then n_e would equal n; otherwise, n_e would be less than n. Solving Eq. (5) for σ, substituting into Eq. (2), and solving for n_e, we obtain

$$n_e = (\mu/\mu_c)^2. \quad (6)$$

We can now use Eq. (6) to estimate the model independence using only mean model and consensus forecast errors.

For example, suppose we have two models with respective TC track forecast errors of 380 and 420 km and a consensus forecast error of 320 km. The effective degrees of freedom for this two-model consensus can be estimated by squaring the ratio of the average error of the models (400 km) and the consensus error giving a value of 1.56. If the consensus error had been 400 km, the effective degrees of freedom would be 1.0, indicating that the forecast errors are completely dependent. If the consensus error had been 283 km, the effective degrees of freedom would be 2.0, indicating that the forecast errors are completely independent. When we examined every possible 2-model consensus that could be formed using the 10 models, we found that the average effective degrees of freedom for those that included WBAI as a member were 1.40, 1.54, and 1.61 at 24, 48, and 72 h, respectively. The average effective degrees of freedom for two-model consensuses that did not include WBAI ranged from 1.25 to 1.35, 1.30 to 1.40, and 1.34 to 1.47 at 24, 48, and 72 h, respectively. Thus, the forecast errors for WBAI displayed more independence from those of the other nine models than the forecast errors of any of the other nine models did from each other. It is the greater independence of the WBAI forecast errors compared with the other models that results in WBAI having such a large positive impact upon the consensus error despite its relatively large forecast error. We suspect that this increased independence stems from WBAI being the only barotropic model used in the consensus. We hypothesize that the forecast errors for the other models, all of which are baroclinic models with full physical parameterization schemes, are more correlated with each other than they are with the errors of the more simple barotropic model.

For the forecast periods shown in Fig. 3, the impact of adding any single model to the consensus is less than 10%. In some cases, the forecast performance is actually degraded slightly by adding a model to the consensus. For example, the AFWI degrades the consensus slightly at all three forecast times. However, inclusion of AFWI for the entire period of record only degrades the 10-model consensus by approximately 2% at 72 h. Including 10 models in a consensus limits performance degradation of any of its members due to bugs, vortex tracker problems, upgrades, and data entry issues.

Increased forecast availability is an equally important reason to include many members in a consensus. The WBAI itself was available for 93%–95% of the verifying JTWC forecasts. Consequently, the 10-model consensus was available for approximately 99% of the official forecasts compared with 98% for the consensus without WBAI. Ideally, WBAI would be available for 100% of the JTWC forecasts, but problems with software and NOGAPS input data used to run WBAR have, so far, prevented that from occurring in operations. The end goal is to produce high quality consensus forecasts for 100% of the JTWC forecasts so that there is always a baseline for track forecasting. High quality track forecasts are also useful in other situations. For example, high quality track forecasts can be used in the development of Tropical Cyclone Formation Alerts and in forecasting extratropical transitions. The authors suspect that inclusion of another skillful barotropic model such as WBAR run with GFS input may add value to these situations.

Acknowledgments. The authors acknowledge Ann Schrader for her work with the ATCF, Jim Gross and Mark DeMaria for development of the interpolator, Mike Fiorino for his vortex trackers, and the entire staff at JTWC for their diligent efforts with the ATCF. Without them, none of this would work. The authors also acknowledge Sam Brand, Ted Tsui, and two anonymous reviewers for their thoughtful comments. This

project is supported by the Oceanographer of the Navy through the program office at the PEO C4I&Space/PMW-180, Program Element 0603207N, and the Office of Naval Research through Grant N00014-03-1-0185.

REFERENCES

Cullen, M. J. P., 1993: The Unified Forecast/Climate Model. *Meteor. Mag.,* **122,** 81–122.

Davidson, N. E., and H. C. Weber, 2000: The BMRC high-resolution tropical cyclone prediction system: TC-LAPS. *Mon. Wea. Rev.,* **128,** 1245–1265.

Goerss, J. S., 2000: Tropical cyclone track forecasts using an ensemble of dynamical models. *Mon. Wea. Rev.,* **128,** 1187–1193.

——, and R. A. Jeffries, 1994: Assimilation of synthetic tropical cyclone observations into the Navy Operational Global Atmospheric Prediction System. *Wea. Forecasting,* **9,** 557–576.

——, C. R. Sampson, and J. M. Gross, 2004: A history of western North Pacific tropical cyclone track forecast skill. *Wea. Forecasting,* **19,** 633–638.

Grell, G. A., J. Dudhia, and D. R. Stauffer, 1995: A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note NCAR/TN-398+STR, 122 pp.

Heming, J. T., J. C. L. Chan, and A. M. Radford, 1995: A new scheme for the initialization of tropical cyclones in the UK Meteorological Office global model. *Meteor. Appl.,* **2,** 171–184.

Hodur, R. M., 1997: The Naval Research Laboratory's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS). *Mon. Wea. Rev.,* **125,** 1414–1430.

Hoel, P. G., 1962: *Introduction to Mathematical Statistics*. Wiley, 427 pp.

Hogan, T. F., and T. E. Rosmond, 1991: The description of the Navy Operational Global Atmospheric Prediction System's spectral forecast model. *Mon. Wea. Rev.,* **119,** 1786–1815.

Kuma, K., 1996: NWP activities at Japan Meteorological Agency. Preprints, *11th Conf. on Numerical Weather Prediction,* Norfolk, VA, Amer. Meteor. Soc., J15–J16.

Kurihara, Y., M. A. Bender, and R. J. Ross, 1993: An initialization scheme of hurricane models by vortex specification. *Mon. Wea. Rev.,* **121,** 2030–2045.

——, ——, R. E. Tuleya, and R. J. Ross, 1995: Improvements in the GFDL hurricane prediction system. *Mon. Wea. Rev.,* **123,** 2791–2801.

——, R. E. Tuleya, and M. A. Bender, 1998: The GFDL hurricane prediction system and its performance in the 1995 hurricane season. *Mon. Wea. Rev.,* **126,** 1306–1322.

Lindgren, B. W., 1976: *Statistical Theory*. Macmillan, 614 pp.

Lord, S. J., 1993: Recent developments in tropical cyclone track forecasting with the NMC global analysis and forecast system. Preprints, *20th Conf. on Hurricanes and Tropical Meteorology,* San Antonio, TX, Amer. Meteor. Soc., 290–291.

Naval Research Laboratory, cited 2004: NITES Tactical Environmental Data Server (TEDS). [Available online at http://www.nrlmry.navy.mil/~lande/TEDS/.]

Neumann, C., 1992: Final report: Joint Typhoon Warning Center (JTWC92) model. SAIC Contract Rep. N00014-90-C-6042 (Part 2), 41 pp. [Available from Naval Research Laboratory, 7 Grace Hopper Ave., Monterey, CA 93943-5502.]

Rennick, M. A., 1999: Performance of the navy's tropical cyclone prediction model in the western North Pacific basin during 1996. *Wea. Forecasting,* **14,** 297–305.

Sampson, C. R., and A. J. Schrader, 2000: The Automated Tropical Cyclone Forecasting System (version 3.2). *Bull. Amer. Meteor. Soc.,* **81,** 1231–1240.

Weber, H. C., 2001: Hurricane track prediction with a new barotropic model. *Mon. Wea. Rev.,* **129,** 1834–1858.