# Prediction of Consensus Tropical Cyclone Track Forecast Error

JAMES S. GOERSS

*Naval Research Laboratory, Monterey, California*

(Manuscript received 6 October 2005, in final form 23 June 2006)

ABSTRACT

The extent to which the tropical cyclone (TC) track forecast error of a consensus model (CONU) routinely used by the forecasters at the National Hurricane Center can be predicted is determined. A number of predictors of consensus forecast error, which must be quantities that are available prior to the official forecast deadline, were examined for the Atlantic basin in 2001–03. Leading predictors were found to be consensus model spread, defined to be the average distance of the member forecasts from the consensus forecast, and initial and forecast TC intensity. Using stepwise linear regression and the full pool of predictors, regression models were found for each forecast length to predict the CONU TC track forecast error. The percent variance of CONU TC track forecast error that could be explained by these regression models ranged from just over 15% at 48 h to nearly 50% at 120 h. Using the regression models, predicted radii were determined and were used to draw circular areas around the CONU forecasts that contained the verifying TC position 73%–76% of the time. Based on the size of these circular areas, a forecaster can determine the confidence that can be placed upon the CONU forecasts. Independent data testing yielded results only slightly degraded from those of dependent data testing, highlighting the capability of these methods in practical forecasting applications.

## 1. Introduction

Consensus tropical cyclone (TC) track forecast aids formed using TC track forecasts from regional and global numerical weather prediction models have become increasingly important in recent years as guidance to TC forecasters at both the National Hurricane Center (NHC) and the Joint Typhoon Warning Center (JTWC). Goerss et al. (2004) illustrate the improvements made over the past decade in the TC track forecasts from these models and from consensus forecast aids formed using these models. Forecasters at NHC routinely use consensus forecast aids formed using the interpolated TC track forecasts from the Geophysical Fluid Dynamics Laboratory (GFDL) Hurricane Prediction System (GFDI; Kurihara et al. 1993, 1995, 1998) and the Global Forecast System (AVNI; Lord 1993) run at the National Centers for Environmental Prediction, the Navy Operational Global Atmospheric Prediction System (NGPI; Hogan and Rosmond 1991; Goerss and Jeffries 1994) and the GFDL model (GFNI; Rennick 1999) run at Fleet Numerical Meteorology and Oceanography Center, and the Met Office global model (UKMI; Cullen 1993; Heming et al. 1995). Two of these consensus forecast aids are CONU and GUNA. CONU is a consensus model that is computed when track forecasts from at least two of the following five models are available: GFDI, AVNI, NGPI, UKMI, and GFNI. GUNA is a consensus model that is computed when track forecasts from all four of the following models are available: GFDI, AVNI, NGPI, and UKMI. While the TC track forecast errors for these two consensus models are comparable, the forecast availability of CONU is decidedly superior to that for GUNA. Therefore, the focus of this study is on CONU.

The purpose of this study is to determine to what extent the TC track forecast error of the consensus model, CONU, can be predicted prior to the time when official forecasts must be issued. Predictors of consensus forecast error must be quantities that are available prior to the official forecast deadline. Goerss (2000) defined consensus model spread to be the average distance of the member forecasts from the consensus forecast and found that, in a broad sense, a forecaster could

*Corresponding author address:* James S. Goerss, NRL, 7 Grace Hopper Ave., Stop 2, Monterey, CA 93943-5502.
E-mail: goerss@nrlmry.navy.mil

DOI: 10.1175/MWR3390.1

© 2007 American Meteorological Society

use that quantity to obtain some measure of confidence to attach to the consensus forecast. Forecast displacement is defined to be the difference between the initial and forecast latitudes (or longitudes) of the TC. The possible predictors examined in this study are consensus model spread, initial and forecast TC intensity, initial TC position and forecast displacement of TC position (latitude and longitude), TC speed of motion, and the number of members available to the consensus model.

In the next section we describe how these predictors are used to estimate CONU TC track forecast error for the Atlantic basin. We also illustrate how the results of this error estimation are displayed for the NHC forecasters. In section 3 we outline the results of independent data testing of the technique. In the final section we summarize the results of this research and discuss our future research plans.

## 2. Estimation of consensus model track forecast error

First, we illustrate the relationships between some of the possible predictors and CONU TC track forecast error. For CONU in the Atlantic basin, the consensus model spread was found to be positively correlated with consensus model TC track forecast error for all forecast lengths. This relationship is illustrated in the scatterplots displayed in Fig. 1, where we see that in general, larger (smaller) CONU forecast track error is associated with larger (smaller) model spread. The weakest relationship was found for the 48-h CONU forecast (Fig. 1a) for which the correlation between spread and forecast error was 0.28. The strongest relationship was found for the 96-h CONU forecast (Fig. 1b) for which the correlation was 0.63. The correlations for the shorter forecast lengths (24–72 h) ranged from 0.28 to 0.39 while those for the longer forecast lengths (96–120 h) ranged from 0.59 to 0.63. As shown in the scatterplots displayed in Fig. 2, initial and forecast TC intensity were found to be consistently but, in general, less strongly related to track error with correlations ranging from 0.30 for the correlation between 120-h CONU track forecast error and initial TC intensity (not shown) to 0.40 for the correlation between 72-h CONU track forecast error and 72-h forecast TC intensity (Fig. 2b). Note that the forecast TC intensities come from the interpolated previous official forecast (OFCI) and are not confined to the 5-kt bins displayed in Fig. 2a for initial TC intensity. The TC intensity and CONU forecast track error were found to be negatively correlated with larger (smaller) track error associated with weaker (stronger) tropical cyclones. Other predictors were found to be reasonably well correlated with forecast error at certain forecast lengths. Scatterplots for two of the more highly correlated predictors for CONU 120-h forecast track error are displayed in Fig. 3. The negative correlation (a value of 0.52) between longitude displacement and track error is illustrated in Fig. 3a, where we see that larger (smaller) track error is associated with larger eastward (westward) forecast displacements. In Fig. 3b, we see that CONU 120-h forecast track error is positively correlated with initial TC latitude with a correlation of 0.43. Taking a closer look at this figure, it appears that this relatively high correla-

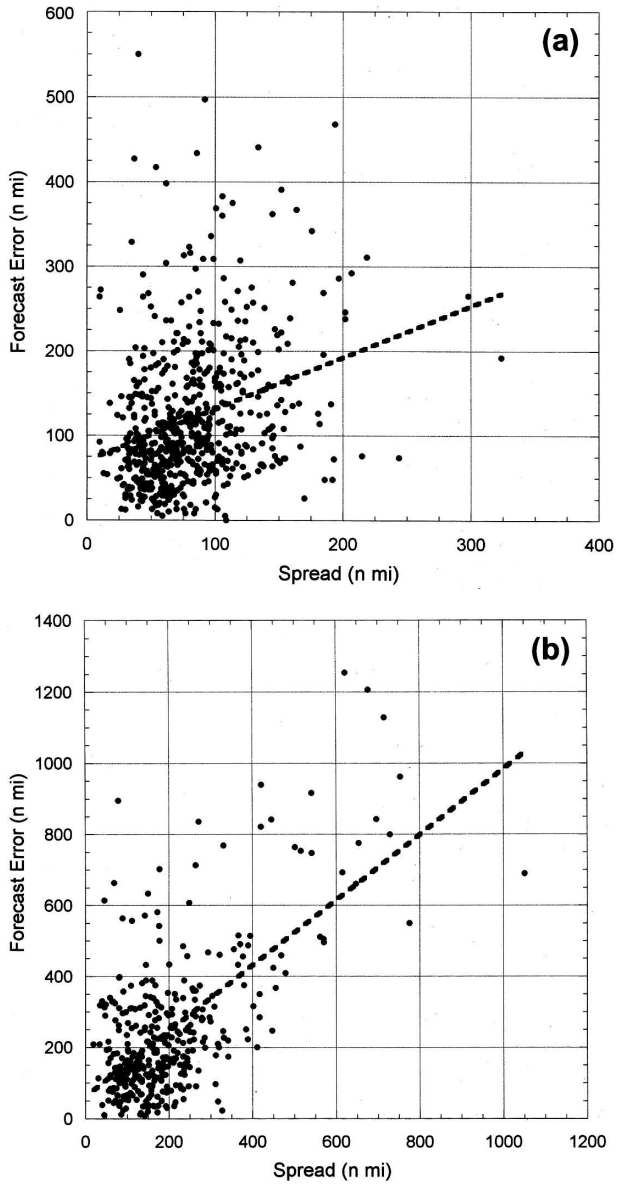


FIG. 1. CONU (a) 48- and (b) 96-h track forecast error versus spread for the Atlantic basin (2001–03).

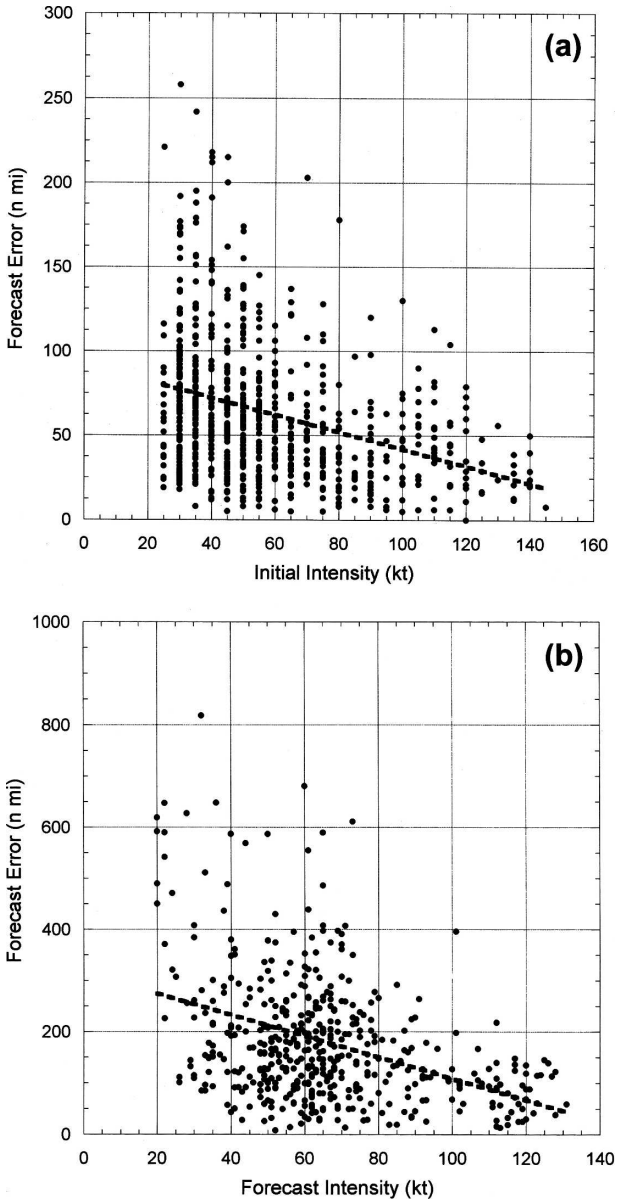


FIG. 2. CONU (a) 24- and (b) 72-h track forecast error versus (a) initial and (b) 72-h forecast TC intensity for the Atlantic basin (2001–03).

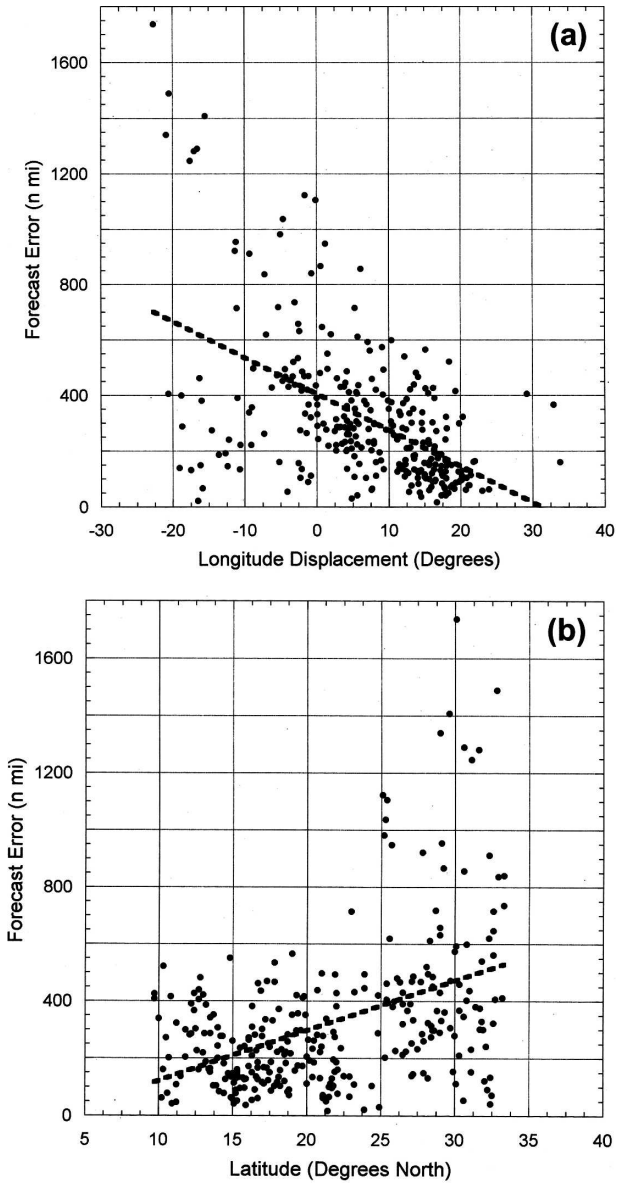


FIG. 3. CONU 120-h track forecast error versus 120-h forecast (a) lon displacement (°) and (b) initial TC lat (°N) for the Atlantic basin (2001–03). Westward (eastward) storm motion is indicated by positive (negative) lon displacement.

tion is due to a much larger range of 120-h CONU forecast track error for TCs located north of 25°N than for those located south of that latitude. Almost all of these large errors are associated with incorrect recurvature forecasts.

Using stepwise linear regression (Draper and Smith 1966) and the pool of predictors from the Atlantic basin for 2001–03, regression models were found for each forecast length to predict the CONU TC track forecast error. To avoid overfitting the dependent dataset, we required that a predictor explain at least 3% of the variance before allowing it to be used by the final regression equation. All of the final regression coefficients were found to be significantly different from zero at well above the 99% level. The regression equations and scatterplots displaying the relationship between CONU track forecast error and predicted error are shown in Fig. 4. The model spread (SPR) was found to be the leading predictor at 96 and 120 h and the second

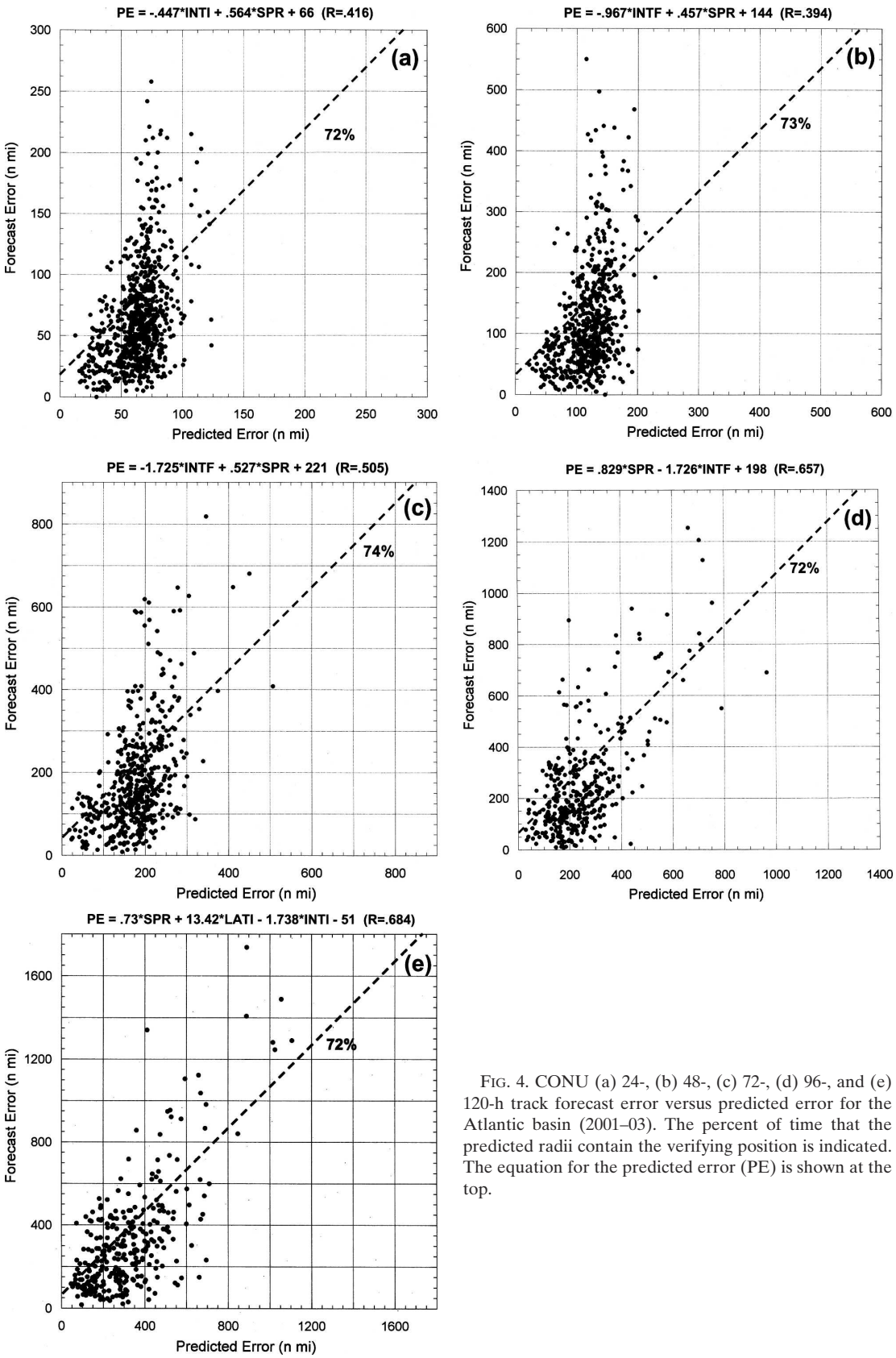


FIG. 4. CONU (a) 24-, (b) 48-, (c) 72-, (d) 96-, and (e) 120-h track forecast error versus predicted error for the Atlantic basin (2001–03). The percent of time that the predicted radii contain the verifying position is indicated. The equation for the predicted error (PE) is shown at the top.

leading predictor at 24, 48, and 72 h. The percent variance of CONU track forecast error that could be explained by spread alone ranged from about 10%–15% for the early forecast lengths to 35%–40% at 96 and 120 h. Initial TC intensity (INTI) was found to be the leading predictor at 24 h and the third leading predictor at 120 h. Forecast TC intensity (INTF) was found to be the leading predictor at 48 and 72 h and the second leading predictor at 96 h. For all forecast lengths, the percent variance of CONU track forecast error explained by intensity alone ranged from about 10% to 15%. Possibly because of the enhanced correlation due to recurvature errors, initial TC latitude (LATI) was found to be the second leading predictor at 120 h. Using these linear regression models, the percent variance of CONU TC track forecast error that could be explained for the 2001–03 Atlantic seasons ranged from just over 15% at 48 h to nearly 50% at 120 h. This increase in variance explained for the longer forecast lengths is consistent with the increase in variance explained by spread, the leading predictor at 96 and 120 h.

Even though GFNI track forecasts were only available out to 72 h for the 2001–03 Atlantic seasons, this study was done in anticipation of their availability out to 120 h for the 2004 season. As was done for CONU, linear regression models were also found for GUNA, a consensus model that does not use GFNI. For all forecast lengths, these regression models were quite similar to those found for CONU. The similarity of these models for forecast lengths out to 72 h gave us confidence that the regression models derived for CONU for 96 and 120 h (without GFNI) could be utilized during the 2004 season, when the longer-range GFNI forecasts would be available.

Next we want to translate these results into a form that has meaning for the NHC forecasters. For the 2001–03 Atlantic hurricane seasons, circular areas with static radii based on NHC's official forecast error for the last 10 yr of 81 n mi at 24 h, 150 n mi at 48 h, 225 n mi at 72 h, 282 n mi at 96 h, and 374 n mi at 120 h drawn around the official forecasts contained the verifying TC position 67%–71% of the time. These circular areas form the basis for the potential day 1–5 track area graphic routinely disseminated by NHC. We would like to construct similar radii to be placed about the CONU forecasts that will contain the verifying TC position roughly 70% of the time. For the 2001–03 Atlantic hurricane seasons, radii were computed by adding a constant varying with forecast length to the predicted CONU TC forecast error derived using the linear regression models. The constants (15 n mi at 24 h, 30 n mi at 48 h, 45 n mi at 72 h, 60 n mi at 96 h, and 75 n mi at 120 h) were chosen so that the verifying TC position would be contained within the circular area surrounding the CONU forecast position roughly 70% of the time. These predicted radii, which varied from approximately 30 to 140 n mi at 24 h, 55 to 260 n mi at 48 h, 65 to 550 n mi at 72 h, 90 to 1000 n mi at 96 h, and 125 to 1175 n mi at 120 h, were used to draw circular areas around each of the CONU forecast positions. For the 2001–03 Atlantic hurricane seasons, these areas were found to contain the verifying TC position 72%–74% of the time. These radii are represented in Fig. 4 by the dashed lines. Note that the scatterplots in Fig. 4 display the relationship between CONU track forecast error and predicted error (not predicted radii). The points below the dashed lines represent the cases where the CONU forecast error is less than the predicted radius and the verifying TC position would be contained within the circular area surrounding the CONU forecast position. The points above the lines represent the cases where the CONU forecast error is greater than the predicted radius and the verifying TC position would be located outside the circular area surrounding the CONU forecast position.

The final step is to effectively convey this information to the NHC forecasters. We illustrate (Fig. 5) the application on the Automated Tropical Cyclone Forecasting System (Sampson and Schrader 2000) of these predicted circular areas for two specific cases. In Fig. 5a, the 120-h CONU forecast for Hurricane Isabel made at 0000 UTC 13 September 2003 is shown along with the predicted circular area. For this case, the verifying position of Hurricane Isabel (indicated by the large black dot) falls just within the boundary of this unusually small area (radius of approximately 150 n mi), which is consistent with the small spread of the 120-h model forecasts for AVNI, GFDI, NGPI, and UKMI. On the other hand, in Fig. 5b, the predicted circular area surrounding the 120-h CONU forecast for Hurricane Kate made at 0000 UTC 30 September 2003 is quite large (radius of approximately 800 n mi), which is consistent with the large spread of the 120-h model forecasts. This large spread is primarily due to the UKMI forecast, which is over 1000 n mi east-southeast of the CONU forecast position and even farther from the forecast positions of the other models. While the verifying position of Kate is contained within the circular area, it is approximately 600 n mi from the CONU forecast position. Thus, based on the size of these circular areas, a forecaster can determine the confidence that can be placed upon the CONU forecasts and use that information in the process of producing the official forecast.

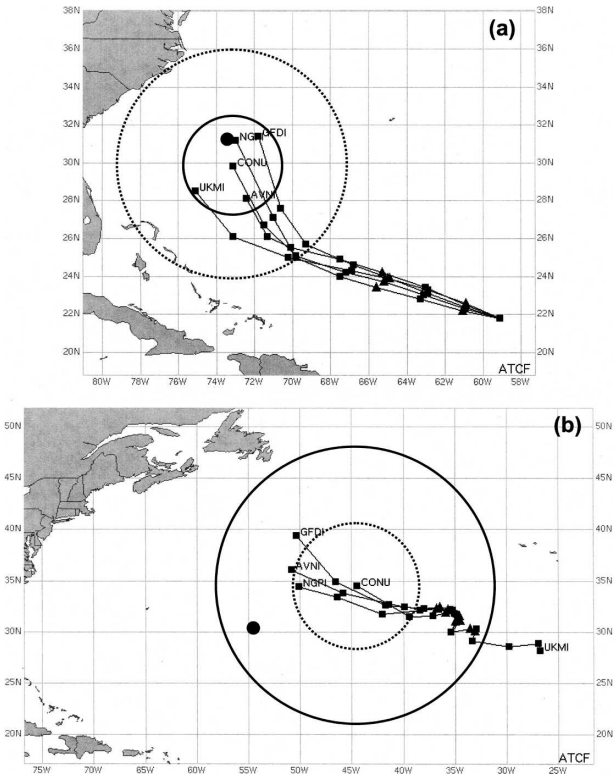


FIG. 5. Predicted 72% confidence radius (solid circle) surrounding the 120-h CONU forecast for (a) Hurricane Isabel, 0000 UTC 13 Sep 2003 and (b) Hurricane Kate, 0000 UTC 30 Sep 2003. The individual model tracks used to create the CONU track are shown along with the 120-h radius (dotted circle) used by the NHC potential day 1–5 track area graphic.

## 3. Independent data testing

We now outline the results of an independent data test that was performed for CONU for the Atlantic basin. For each forecast length, regression equations were computed using stepwise linear regression and the pool of predictors from the Atlantic basin for 2001–02. The predicted TC forecast errors were then computed for the 2003 Atlantic season. Scatterplots displaying the relationship between CONU track forecast error and the independent predicted error are shown in Fig. 6. While the values of the various regression coefficients vary, we found that the predictors chosen to be used at each forecast length were the same as those chosen when the 2001–03 Atlantic seasons were used as the dependent dataset. Using these linear regression models, the percent variance of the CONU track forecast error that could be explained for the 2003 Atlantic season ranged from 23% at 24 h to 46% at 96 h, quite similar to what we found from the dependent testing outlined in section 2. Just as was done previously, predicted radii were computed to be used to form circular areas around the CONU forecast positions. For this independent sample, these predicted radii varied from approximately 30 to 140 n mi at 24 h, 30 to 330 n mi at 48 h, 45 to 545 n mi at 72 h, 60 to 650 n mi at 96 h, and 75 to 825 n mi at 120 h. These radii are represented in Fig. 6 by the dashed lines. Note that the scatterplots in Fig. 6 display the relationship between CONU track forecast error and predicted error (not predicted radii). The points below the dashed lines in Fig. 6 represent the cases where the CONU forecast error is less than the predicted radius and the verifying TC position would be contained within the circular area surrounding the CONU forecast position. These areas were found to contain the verifying TC position 68%–83% of the time and, except for 120 h, the percentages were actually higher than those found for the dependent sample (Fig. 4). In all areas of comparison (percent variance of consensus model track forecast error explained, range of the predicted radii, and percent of verifying TC positions contained within the circular areas), we see that the results found from independent testing compare quite favorably with those found from dependent testing. We conclude from our independent testing that the regression equations derived from previous seasons appear to be stable and can be effectively applied to the next season.

## 4. Summary and conclusions

The purpose of this study was to determine to what extent the TC track forecast error of the consensus model, CONU, can be predicted prior to the forecast deadline. The possible predictors examined in this study were consensus model spread, initial and forecast TC intensity, initial TC position and forecast displacement of TC position (latitude and longitude), TC speed of motion, and the number of members available to the consensus model.

For CONU in the Atlantic basin, it was found that consensus model spread was the most important predictor followed by TC intensity (either initial or forecast). Consensus model spread was found to be positively correlated with consensus model TC track forecast error while intensity was found to be negatively correlated.

Using stepwise linear regression and the pool of predictors for the 2001–03 Atlantic seasons, regression models were found to predict CONU TC track forecast error for each forecast length. It was found that the regression models explained 15%–20% of the track forecast error variance for the shorter forecast lengths

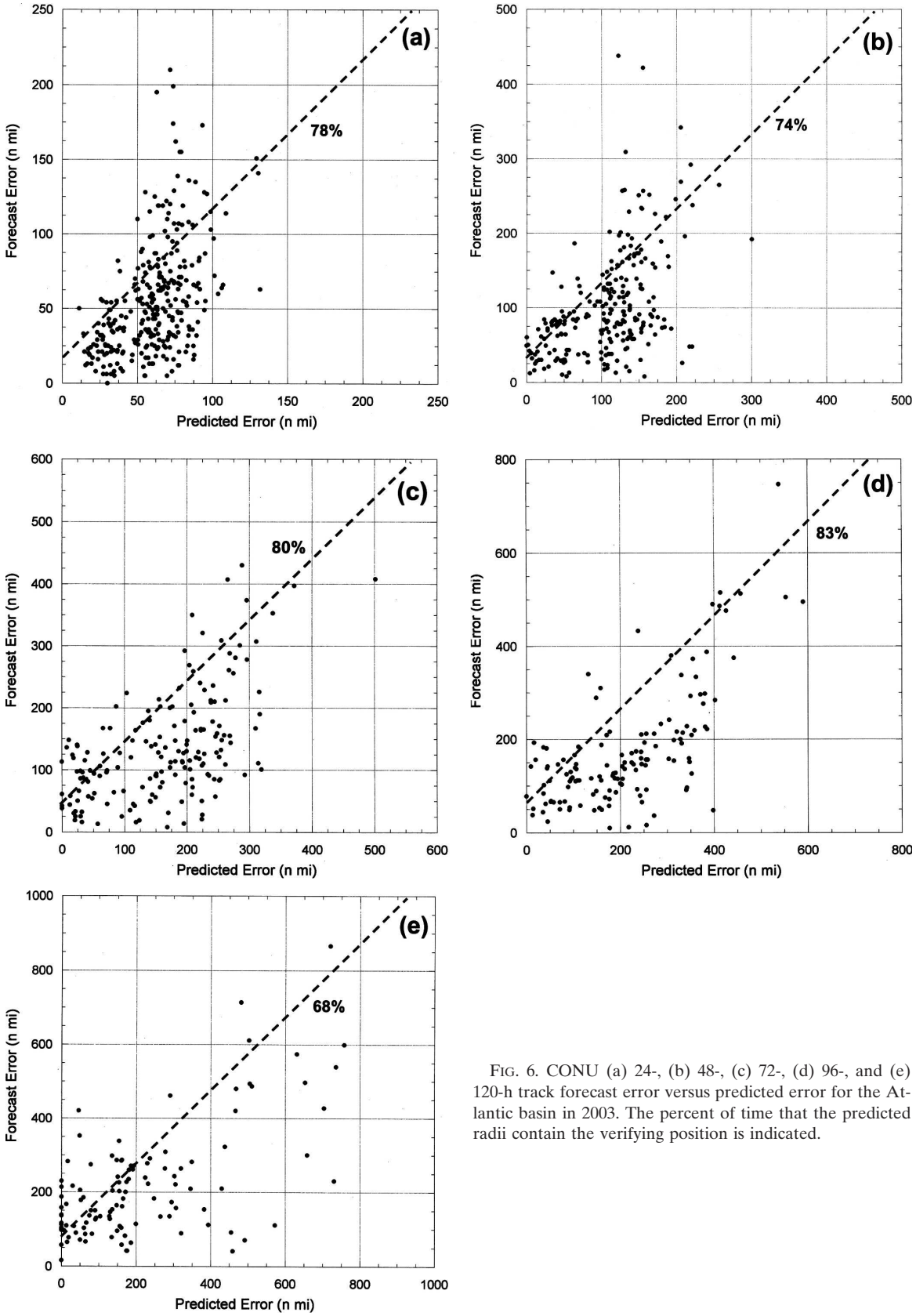


FIG. 6. CONU (a) 24-, (b) 48-, (c) 72-, (d) 96-, and (e) 120-h track forecast error versus predicted error for the Atlantic basin in 2003. The percent of time that the predicted radii contain the verifying position is indicated.

(24–72 h) and 45%–50% of the track forecast error variance for the longer forecast lengths (96 and 120 h). Predicted radii were derived by adding a constant (which varied with respect to forecast length) to the track forecast error predicted by the regression models. These radii were used to draw circular areas around each of the CONU forecast positions for each forecast length. The additive constants were chosen so that the verifying TC position was contained within the circular area surrounding the CONU forecast position 72%–74% of the time for the 2001–03 Atlantic seasons. These predicted radii varied from approximately 30 to 140 n mi at 24 h, 55 to 260 n mi at 48 h, 65 to 550 n mi at 72 h, 90 to 1000 n mi at 96 h, and 125 to 1175 n mi at 120 h. Based on the size of these radii, a forecaster can determine how much (or little) confidence can be ascribed to the CONU forecast position.

Independent data testing was performed. The 2001–02 Atlantic seasons were used as the dependent dataset and regression models were derived. These regression models were then applied to the 2003 Atlantic season. In all areas of comparison (percent variance of consensus model track forecast error explained, range of the predicted radii, and percent of verifying TC positions contained within the circular areas), we found that the results from the independent testing compared quite favorably with those found from dependent testing. Thus, we concluded that we should be able to effectively use the regression models determined from the 2001–03 Atlantic seasons to produce guidance to be used during the 2004 Atlantic season.

We close with a discussion of our future research plans. The techniques described in this paper will be extended to the eastern and western North Pacific and the Southern Hemisphere basins in JTWC's area of responsibility. The predicted consensus error guidance will be verified for all basins and forecast lengths. We will examine the question of what length training period is optimal to produce the regression models to be used for an upcoming season. In this study we used simple stepwise linear regression. The use of more sophisticated regression methods along with the transformation and combination of selected predictors will be investigated. As illustrated by Goerss et al. (2004), the TC track forecasting skill of the operational numerical weather prediction models is steadily improving as the various operational centers make upgrades to their forecast systems. The impact of changes made to the individual models upon the predicted consensus error guidance will be examined. Finally, we would like to take a hard look at consensus track error in much the same way as Neumann and Pelissier (1981) examined TC track forecast error for the Atlantic in the 1970s. The change in the relationship between the CONU 120-h track forecast error and initial TC latitude at approximately 25°N displayed in Fig. 3b suggests the same stratification scheme used by Neumann and Pelissier. They used a latitude of 24.5°N to divide storms that remain in the easterlies (or which are just beginning to recurve) from storms that have already recurved into the westerlies (or which are well into recurvature). We plan to investigate the impact of this stratification on the determination of the regression models to predict consensus track error and to also look into the prediction of zonal and meridional consensus track errors.

*Acknowledgments.* Special thanks are given to James Franklin and Colin McAdie of NHC for their invaluable assistance in graphical display and regression analysis and to Jim Gross of NHC and Buck Sampson of NRL Monterey for making possible the implementation of this research on the ATCF. This research was performed on Project A8R2WRP entitled "Quantifying Tropical Cyclone Track Forecast Uncertainty and Improving Extended-range Tropical Cyclone Track Forecasts Using an Ensemble of Dynamical Models," funded by the National Oceanic and Atmospheric Administration Joint Hurricane Testbed administered by the U.S. Weather Research Program.

## REFERENCES

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

Draper, N. R., and H. Smith, 1966: *Applied Regression Analysis*. Wiley, 407 pp.

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. Gross, 2004: A history of western North Pacific tropical cyclone track forecast skill. *Wea. Forecasting,* **19,** 633–638.

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

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.

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.

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.

Neumann, C. S., and J. M. Pelissier, 1981: An analysis of Atlantic tropical cyclone forecast errors, 1970–1979. *Mon. Wea. Rev.,* **109,** 1248–1266.

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.