5A.3

IMPACT OF STOCHASTIC CUMULUS ON THE NOGAPS ET ENSEMBLE FORECASTING SYSTEM PART II: TROPICAL CYCONE TRACK FORECAST PERFORMANCE

James S. Goerss* and Carolyn A. Reynolds
Naval Research Laboratory, Monterey, California

1. INTRODUCTION

Operational numerical weather prediction (NWP) centers have been producing global atmospheric ensemble forecasts using initial-state perturbations since the early 1990s (e.g.,Toth and Kalnay 1993; Buizza and Palmer 1998). Currently, a number of methods are used to perturb the initial state including bred vectors (Toth and Kalnay 1993), singular vectors (Molteni et al. 1996), perturbed observations (Houtekamer et al. 1996), and the ensemble Kalman Filter (Houtekamer and Mitchell 2005). The ensemble transform (ET) technique (Bishop and Toth 1999) has been applied to the global atmospheric ensemble forecasts at the National Centers for Environmental Predicition (NCEP: Wei et al. 2006, 2008) and has recently been compared to the operational bred vector method used for the Navy Operational Global Atmospheric Prediction System (NOGAPS) ensembles (McLay et al. 2007, 2008). McLay et al. (2008) found that the NOGAPS ET scheme demonstrated comparable or superior performance over the operational bred vector scheme under a variety of metrics. However, the NOGAPS ET scheme was found to produce initial ensemble perturbations that were too small in the tropics and too large in the mid-latitudes. Reynolds et al. (2008) combined the NOGAPS ET scheme with the stochastic convection model perturbation scheme described by Teixeira and Reynolds (2008) and found that the net result was larger initial perturbations in the tropics and smaller perturbations in the extratropics. They examined the performance of the NOGAPS ET ensemble under a variety of metrics, focusing on ensemble-mean errors, Brier scores, and spread-skill relationships for 250 hPa and 10 m winds. They found that, in general, the addition of stochastic convection improved ensemble performance in the tropics and had little impact upon performance in the extratropics. In this paper we examine the performance of the NOGAPS ET ensemble, with and without the addition of stochastic convection, upon tropical cyclone (TC) track forecasts.

Multi-model ensemble mean or consensus TC track forecast aids formed using TC track forecasts from regional and global NWP 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). The improvements made over the past decade in the TC track forecasts from these NWP models and from consensus forecast aids formed using these models have been well documented (Goerss et al. 2004, Sampson et al. 2005). 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 NCEP; NOGAPS (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 U. K. Meteorological Office global model (UKMI; Cullen 1993, Heming et al. 1995). One of the top forecast aids, CONU, is a consensus model that is computed when track forecasts from at least two of the aforementioned five models are available. While TC track forecasts from the NCEP Global Forecast System (GFS) ensemble system are operationally available to the NHC forecasters, they are rarely used. The forecast errors for the individual members are quite large and the forecast errors for the ensemble mean are significantly larger than those for multi-model ensembles. In this paper, we also compare the TC track forecast performance of the NOGAPS ET ensemble with that of the GFS ensemble and the multi-model forecast aid, CONU.

2. IMPACT OF STOCHASTIC CONVECTION

The NOGAPS ET ensemble system was run with and without the addition of stochastic convection over the period from July 4-October 31, 2005. The ET was run with a 6-h cycle (ie., initial-time perturbations were created using 6-h ensemble forecasts), but the extended (120h) forecasts were only run twice daily (at 00Z and 12Z). Both ET ensembles were produced using NOGAPS at a T119L30 resolution, and contain one member with no initial-time perturbations and 32 members with initial-time perturbations. Note that all of the 33 members in the stochastic convection ensemble have stochastic convection, even the member that does not have any initial-time perturbations. This period was an extremely active one covering most of the record-breaking Atlantic season. For the Atlantic, there were 12 hurricanes (including Katrina, Rita, and Wilma) and 9 tropical storms. It was an active period for the other basins as well. However, in this paper we focus our attention on the Atlantic basin.

---

*Corresponding author address: James S. Goerss, Naval Research Laboratory, 7 Grace Hopper Ave., Stop 2, Monterey, CA 93943;
e-mail: jim.goerss@nrlmry.navy.mil

TC track forecast errors were determined for 9 members (8 members with initial-time perturbations and one member with no initial-time perturbations) of the NOGAPS ET ensemble system run with and without the addition of stochastic convection for the test period and are displayed in Figs. 1 and 2, respectively. Tests indicated that there was no improvement in the ensemble mean TC track forecast error to be gained by increasing the number of ensemble members beyond this number. In Fig. 1, we can see that the TC track forecast errors for the members with initial-time perturbations (EN04, EN08, etc.) are consistently larger than those for the member with no initial-time perturbation (EN00) and that the forecast errors for the ensemble mean (ENMN) are smaller than those for any of the ensemble members. Examination of Fig. 2 indicates that the same can be said for the ensemble system run without the addition of stochastic convection. Comparing Figs. 1 and 2, we find that for all forecast lengths, forecast errors for the members with initial-time perturbations are roughly the same whether stochastic convection was added or not. The same can be said for the members with no initial-time perturbations. However, for forecast lengths greater than 48h, the ensemble means for the system run with the addition of stochastic convection (ENMN) are less than those for the system run without (EOMN). Goerss (2000) found that the ensemble mean forecast error depends on two things: 1) the mean forecast error of the individual members that make up the ensemble and 2) the degree of independence (or effective degrees of freedom) of the forecast errors of the individual members. Sampson et al. (2006) demonstrated that the effective degrees of freedom for an ensemble can be estimated by squaring the ratio of the average error of the members and the ensemble mean error. For the ensemble system run with the addition of stochastic convection, the average errors of the members were 93 nm, 153 nm, 197 nm, 254 nm, and 305 nm for the 24-h, 48-h, 72-h, 96-h, and 120-h forecasts, respectively, while the ensemble mean errors were 68 nm, 118 nm, 147 nm, 190 nm, and 222 nm. For this ensemble, the effective degrees of freedom were 1.87, 1.68, 1.8, 1.79, and 1.89 for the 24-h, 48-h, 72-h, 96-h, and 120-h forecasts, respectively. For the ensemble system run without the addition of stochastic convection, the average errors of the members were 92 nm, 148 nm, 202 nm, 266 nm, and 324 nm for the 24-h, 48-, 72-h, 96-h, and 120-h forecasts, respectively, while the ensemble mean errors were 71 nm, 118 nm, 160 nm, 213 nm, and 257 nm. For this ensemble, the effective degrees of freedom were 1.68, 1.57, 1.59, 1.56, and 1.59 for the 24-h, 48-h, 72-h, 96-, and 120-h forecasts, respectively. Thus, we see that the addition of the stochastic convection results in an increase in the independence of the ensemble member forecast errors at all forecast lengths, which, in turn, results in a reduction in the ensemble mean forecast errors.

Since forecasts from operational NWP models are generally not available to the forecasters at NHC and JTWC, a simple technique (Sampson et al. 2006) is used to take the model forecast positions from model runs initialized 6h and/or 12h previously and adjust the forecast to apply to the current synoptic time and initial conditions. For historical reasons, these adjusted versions are known as interpolated models. The TC track forecast errors for the interpolated NOGAPS ET ensemble means with and without the addition of stochastic convection (ENMI and EOMI, respectively) are compared with those for CONU and the NHC official forecast (OFCL) in Fig. 3. For all forecast lengths, the ENMI errors are less than those for EOMI. The additon of stochastic convection resulted in improvements significant at the 95, 97, 95, and 93 percent levels for the 48-h, 72-h, 96-h, and 120-h forecasts, respectively. While not statistically significant, for forecast lengths greater than 72h, the ENMI errors are smaller than those for CONU, the multi-model forecast aid described previously. Henceforth, when we refer to the NOGAPS ET ensemble, it will be assumed that it was run with the addition of stochastic convection.

3. ENSEMBLE COMPARISON

In Fig. 4 the TC track forecast errors for the NCEP GFS ensemble members are compared with those for the ensemble mean (AEMN) and the full resolution operational GFS (AVNO). For all forecast lengths, we see that the forecast errors for the ensemble mean are smaller than those for any of the ensemble members and are considerably smaller than those for the AVNO for the 96-h and 120-h forecasts. For the GFS ensemble system, the average errors of the members were 85 nm, 139 nm, 209 nm, 315 nm, and 416 nm for the 24-h, 48-h, 72-h, 96-h, and 120-h forecasts, respectively, while the ensemble mean errors were 59 nm, 100 nm, 156 nm, 250 nm, and 313 nm. The effective degrees of freedom for the GFS ensemble were 2.07, 1.93, 1.79, 1.59, and 1.77 for the 24-h, 48-h, 72-h, 96-h, and 120-h forecasts, respectively. Comparing this ensemble with the NOGAPS ET ensemble (Fig. 1), we see that the average errors for its members were smaller and the effective degrees of freedom were larger at 24h and 48h. As a result, the ensemble mean errors were also smaller, 59 nm vice 68 nm at 24h and 100 nm vice 118 nm at 48h. While their effective degrees of freedom were almost identical at 72h, the average error for the GFS ensemble members was a bit larger than that for the NOGAPS ensemble (209 nm vice 197 nm) resulting in a larger ensemble mean error (156 nm vice 147 nm). For the 96-h and 120-h forecasts, however, the average member errors were larger and the effective degrees of freedom were smaller resulting in considerably larger ensemble mean errors (250 nm vice 190 nm at 96h and 313 nm vice 222 nm at 120h). We see the same thing when we compare the interpolated ensemble means in Fig. 5. The track forecast errors for the GFS ensemble (AEMI) are less than those for the NOGAPS ensemble (ENMI) at 24h and 48h, a little worse at 72h, and considerably worse at 96h and 120h. The AEMI improvement at 24h is significant at the 98 percent level while the ENMI improvements at 96h and 120h are significant at the 90 and 96 percent levels, respectively.

The forecast availability for the single-model ensembles (AEMI and ENMI) is compared with that for the multi-model forecast aid, CONU, in Fig. 6. We define forecast availability to be the percent of the time that forecasts were available to the forecaster when he/she was required to make a TC forecast. While not as good as that for CONU, the forecast availability for the NOGAPS ET ensemble mean is respectable, ranging from near 90 percent at 48h to just over 80 percent at 120h. Its availability is considerably better than that for the GFS ensemble mean, which ranged from about 65 percent at 24h to about 45 percent at 120h.

The relationship between ensemble spread and ensemble mean TC track forecast error is illustrated in Fig. 7 for the NOGAPS ET (ENMN), NCEP GFS (AEMN), and GUNA ensembles. GUNA is another multi-model forecast aid used by the NHC forecasters that is computed when track forecasts are available from all four of its members (GFDI, UKMI, NGPI, and AVNI). Ensemble spread is defined to be the average distance of the ensemble member forecasts from the ensemble mean forecast. With the exception of the 96-h forecast, we see that the spread-skill relationship is much stronger for the multi-model ensemble, GUNA, than for the single-model ensembles.

A new consensus track forecast aid, CON6, was formed by adding ENMI to the other 5 members of the CONU ensemble, and another, ENS3, was formed using ENMI, AEMI, and CONU as the ensemble members. Their TC track forecast errors are displayed in Fig. 8. The CON6 forecast errors were comparable to those for CONU for forecast lengths less than 96h and were less than those for CONU at 96h and 120h (significant at the 96 and 98 percent levels). The forecast errors for ENS3 were also comparable to those for CONU and CON6 for the shorter forecast lengths but were even better than those for CON6 at 96h and 120h. The respective errors for CON6 at 96h and 120h were 194 nm and 275 nm, while those for ENS3 were 184 nm and 252 nm (cf. 204 nm and 297 nm for CONU). Thus, superior TC track forecast guidance can be formed by combining single-model ensemble means with the CONU multi-model ensemble.

4. HURRICANE KATRINA CASE

During the early stages of Hurricane Katrina, before the storm passed over the Florida Peninsula and moved into the Gulf of Mexico, virtually all of the NWP model guidance predicted an early turn to the north taking the storm well to the east of its actual landfall south of New Orleans. As a result, the official NHC forecasts suffered as well. We now examine some of these track forecasts made about two days before almost all of the NWP model track forecasts shifted to the west facilitating the excellent guidance provided by NHC over the 60-h period leading up to landfall.

The track forecasts for ENMI, AEMI, and CONU from 00Z 25 August 2005 are displayed in Fig. 9. While the interpolated GFS ensemble mean forecast took Katrina up the Florida Peninsula and the CONU forecast landfall in the eastern Florida Panhandle, the interpolated NOGAPS ensemble mean forecast provided a much better depiction of the actual storm track. The NOGAPS ET ensemble mean and member forecasts from 12Z 24 August 2005 (from which the interpolated forecast shown in Fig. 9 was constructed) are displayed in Fig. 10. While one member turned Katrina north too soon and two members took the storm too far to the west, most of the members and, of course, the ensemble mean made excellent 120-h forecasts of the actual landfall. The GFS ensemble mean and member forecasts from 18Z 24 August 2005 (from which the interpolated forecast shown in Fig. 9 was constructed) are displayed in Fig. 11. We see that the GFS ensemble members formed two clusters, one indicating landfall in the Florida Panhandle and the other indicating an early turn to the north taking Katrina back to the Atlantic. The resulting ensemble mean forecast took Katrina up the Florida Peninsula. The track forecasts for CONU and its members are shown in Fig. 12. While the GFNI provided excellent guidance and a track very close to that for ENMI (Fig. 9), the other models all turned Katrina north too soon, and the resulting CONU forecast indicated landfall in the eastern Florida Panhandle.

5. SUMMARY

The impact of the addition of stochastic convection to the NOGAPS ET ensemble upon TC track forecasts was examined for the Atlantic basin over the period from July 4-October 31, 2005. TC track forecast errors were determined for 9 members of the ensemble system run with and without the addition of stochastic convection. We found that the addition of stochastic convection resulted in an increase in the independence of the ensemble member forecast errors at all forecast lengths resulting in a reduction in the ensemble mean forecast errors. For forecast lengths greater than 72h, TC track forecast errors for the interpolated ensemble mean (with stochastic convection), ENMI, were smaller than those for the multi-model ensemble forecast aid, CONU.

When compared with the NCEP GFS ensemble system, it was found that the ENMI forecast errors were greater than those for the GFS ensemble mean (AEMI) at 24h and 48h, a little better at 72h, and significantly better at 96h and 120h. While not as good as that for CONU, the forecast availability for ENMI was much better than that for AEMI. The relationship between ensemble spread and ensemble mean track forecast error was found to be much weaker for the single-model ensemble systems than for a simple multi-model system. By combining the single-model ensemble means with the CONU multi-model ensemble, it was found that superior TC track forecast guidance could be created.

Finally, for Hurricane Katrina, the NOGAPS ET ensemble system produced reliable track forecast guidance nearly two days before virtually all other NWP and consensus forecast aids. Five days before Katrina made landfall, while other guidance predicted landfall far to the east, the 120-h ENMI forecast position was quite close to the actual landfall south of New Orleans.

Acknowledgments. We acknowledge the support of the Office of Naval Research (ONR) through Program Element 0602435N and 0601153N, and the support of a NOAA cooperative agreement for THORPEX. The DoD High Performance Computing Program at NAVO MSRC provided part of the computing resources.

REFERENCES

Bishop, C. H., and Z. Toth, 1999: Ensemble transformation and adaptive observations. *J. Atmos. Sci.*, **56**, 1748-1765.

Buizza, R. and T. N. Palmer, 1998: Impact of ensemble size on ensemble prediction. *Mon. Wea. Rev.*, **126**, 2503-2518.

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

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

______, 2000: Tropical cyclone track forecasts using an ensemble of dynamical models. *Mon. Wea. Rev.*, **128**, 1187-1193.

______, 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. *Meteorol. 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.

Houtekamer, P.L., L. Lefaivre, J. Derome, H. Ritchie, and H. L. Mitchell, 1996: A System Simulation Approach to Ensemble Prediction. *Mon. Wea. Rev.*, **124**, 1225-1242.

______, and H. L. Mitchell, 2005: Ensemble Kalman Filtering. *Quart. J. Roy. Met. Soc.*, **131**, 3269-3289.

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.

______, M. A. Bender, 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.

Rennick, M. A., 1999: Performance of the Navy's tropical cyclone prediction model in the western North Pacific basin during 1996. **Wea. Forecasting**, **14**, 3-14.

McLay, J. G. C. H. Bishop, and C. A. Reynolds, 2008: A demonstration of the ensemble-transform analysis perturbation scheme at NRL. *Mon. Wea. Rev.*, accepted.

______, C. H. Bishop, and C. A. Reynolds, 2007: The ensemble transform scheme adapted for the generation of stochastic perturbations. *Q. J. R. Meteorol. Soc*. **133**, 1257-1266.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Q. J. R. Meteorol. Soc.*, **122**, 73-119.

Reynolds, C. A., J. Teixeira and J. G. McLay, 2008: Impact of stochastic convection on the ensemble transform. *Mon. Wea. Rev.*, accepted.

Sampson, C., J. Goerss, and A. Schrader, 2005: A consensus track forecast for southern hemisphere tropical cyclones. *Aust. Met. Mag.*, **54**, 115-119.

______, J. Goerss, and H. Weber, 2006: Operational performance of a new barotropic model (WBAR) in the western North Pacific basin. *Wea. Forecasting*, **21**, 656-662.

Teixeira, J., and C. A. Reynolds, 2008: Stochastic nature of physical parameterizations in ensemble prediction: a stochastic convection approach. *Mon. Wea. Rev.*, **136**, 483-496.

Toth, Z, and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Am. Meteor. Soc.*, **74**, 2317-2330.

Wei, M., Z. Toth, R. Wobus, and Y. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operation forecast system. *Tellus*, **60A**, 62-79.

______, Z. Toth, R. Wobus, Y. Zhu, C. H. Bishop, and X. Wang, 2006: Ensemble transform Kalman filter-based ensemble perturbations in an operational global prediction system at NCEP. *Tellus*, **58A**, 28-44.

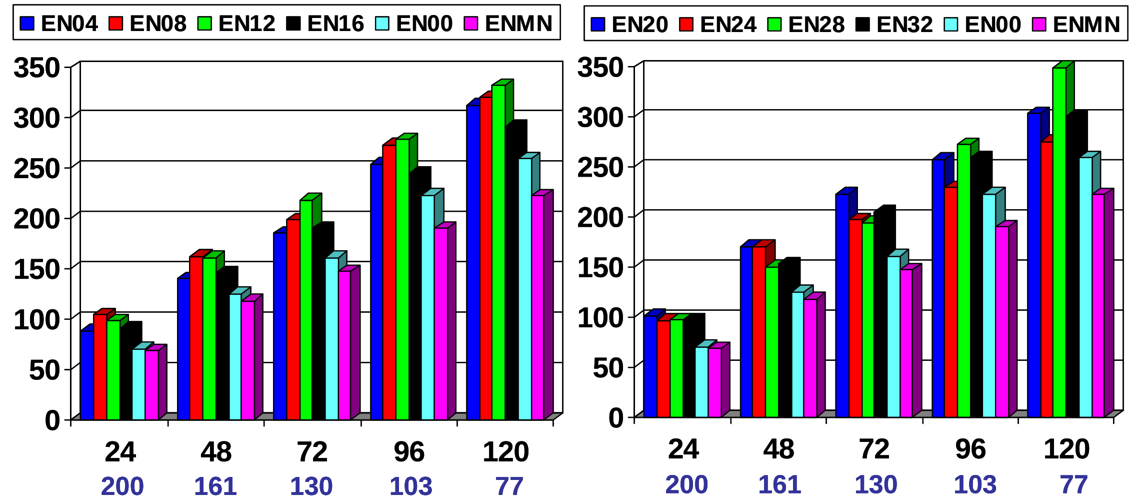

Fig. 1. Comparison of the NOGAPS ET ensemble (with the addition of stochastic convection) TC track forecast error (nm) for July 4-October 31, 2005 for the Atlantic basin. The number of forecasts verified is listed below the forecast length.

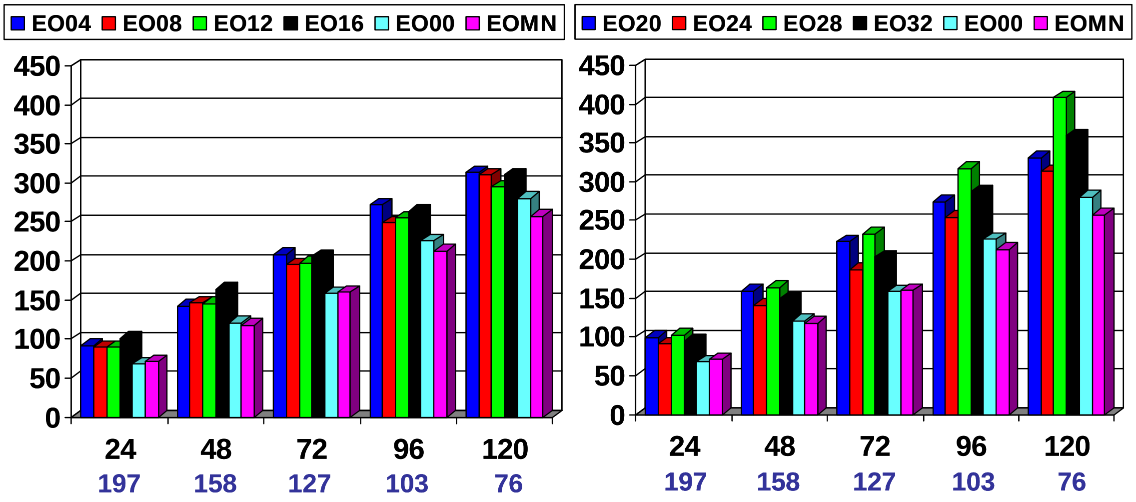

Fig. 2. Comparison of the NOGAPS ET ensemble (without the addition of stochastic convection) TC track forecast error (nm) for July 4-October 31, 2005 for the Atlantic basin. The number of forecasts verified is listed below the forecast length.

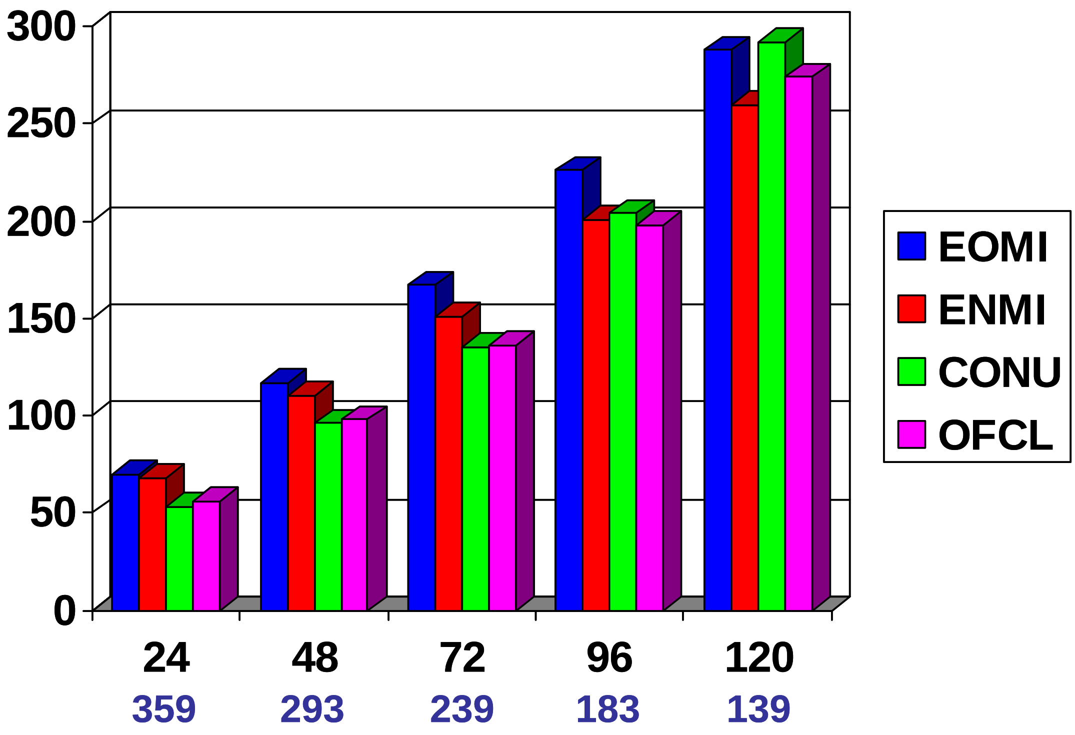

Fig. 3. Comparison of TC track forecast error (nm) for July 4-October 31, 2005 for the Atlantic basin. The number of forecasts verified is listed below the forecast length.

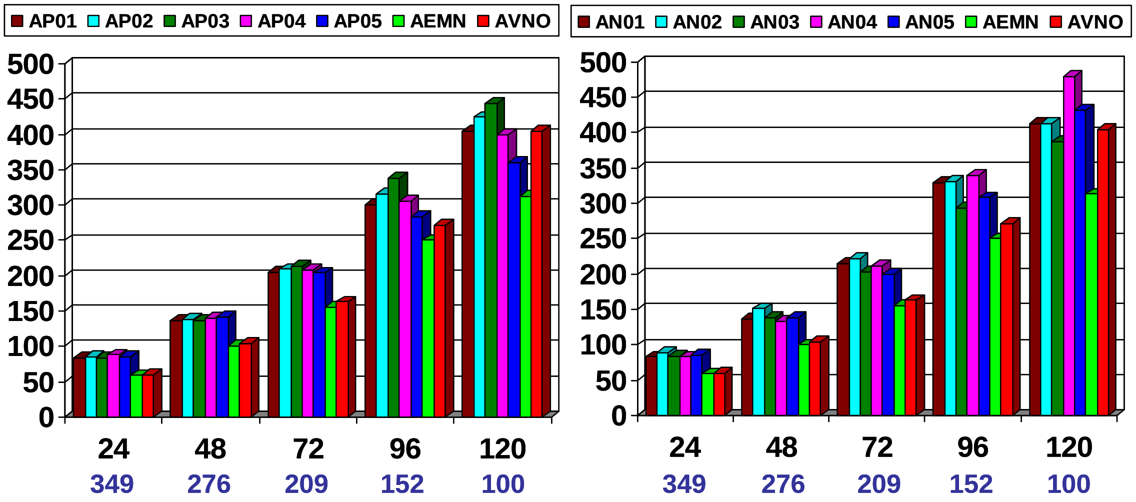

Fig. 4. Comparison of the NCEP GFS ensemble TC track forecast error (nm) for July 4-October 31, 2005 for the Atlantic basin. The number of forecasts verified is listed below the forecast length.

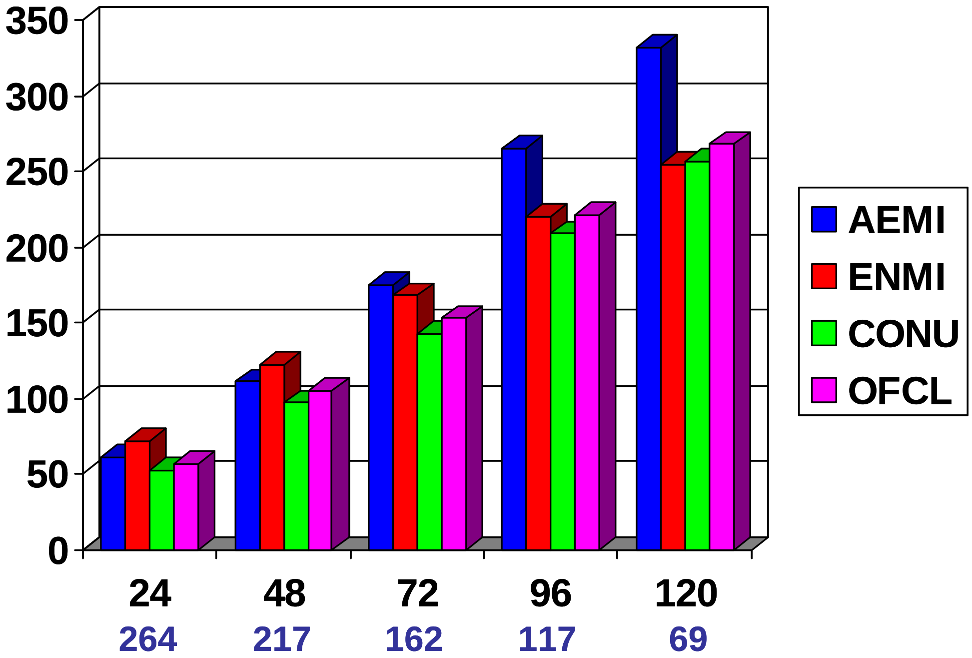

Fig. 5. Comparison of TC track forecast error (nm) for July 4-October 31, 2005 for the Atlantic basin. The number of forecasts verified is listed below the forecast length.

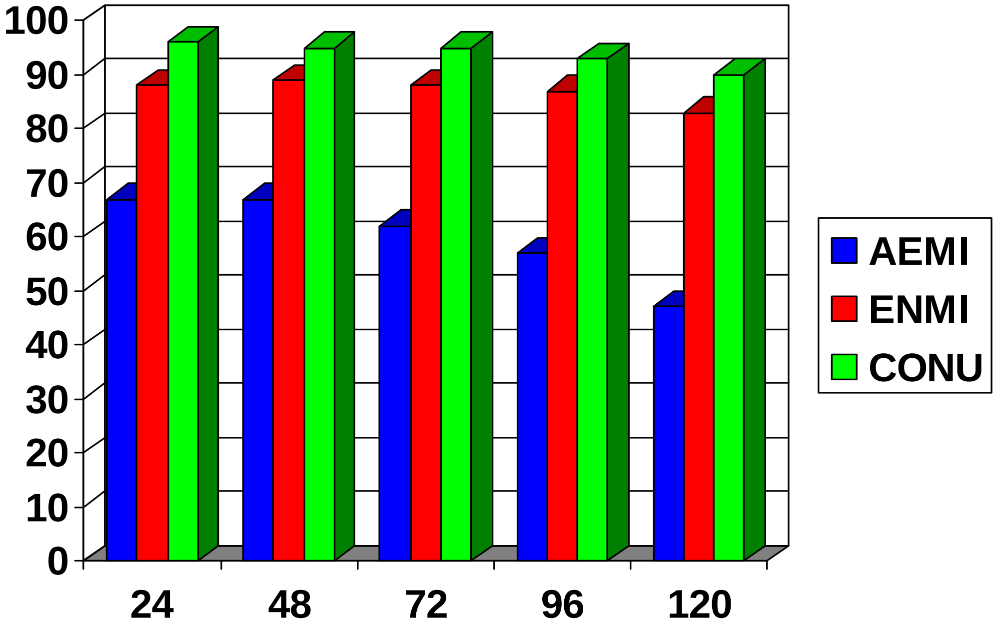

Fig. 6. Comparison of forecast availability for July 4-October 31, 2005 for the Atlantic basin.

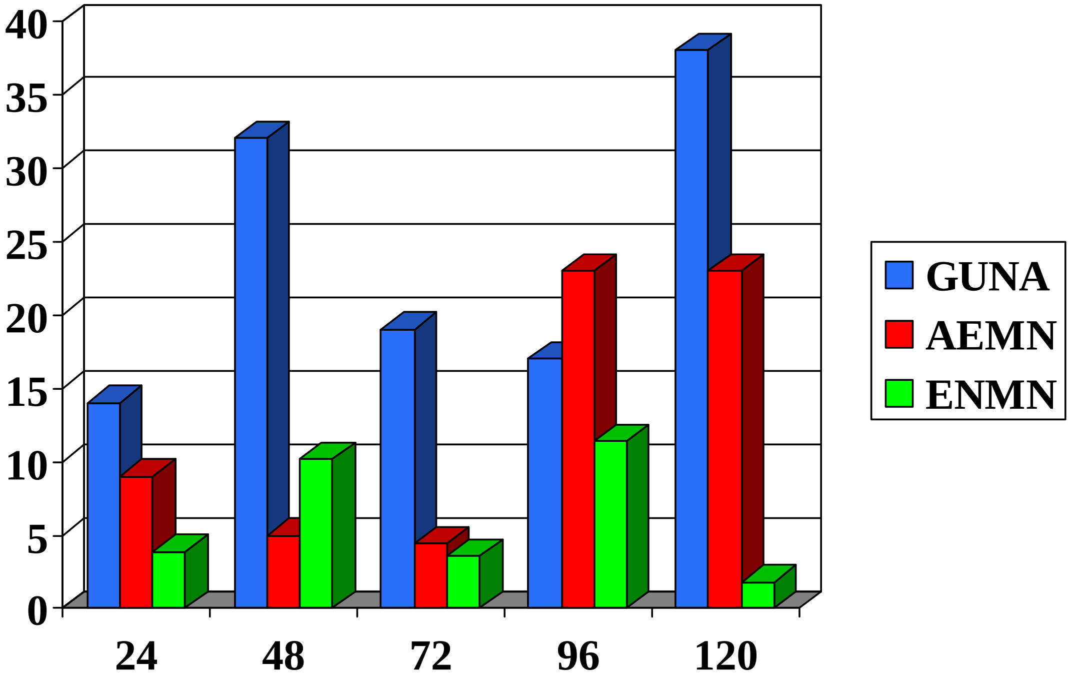

Fig. 7. Percent variance of TC track forecast error explained by spread for the GUNA, NCEP GFS, and NOGAPS ET ensembles.

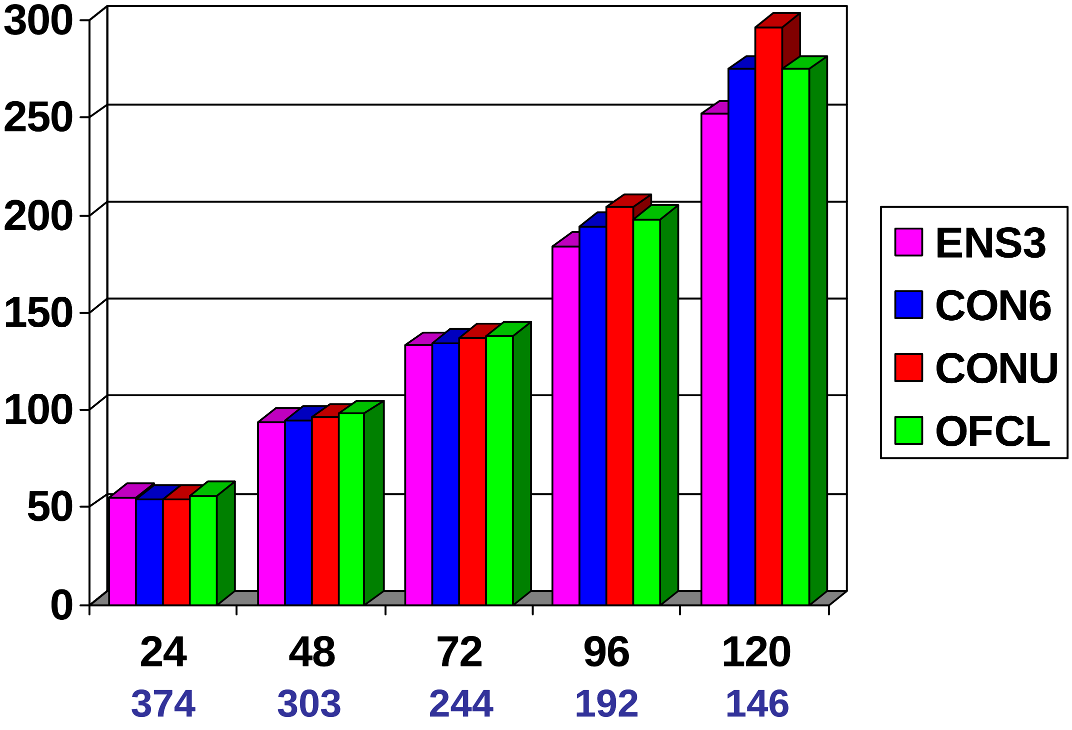

Fig. 8. Comparison of TC track forecast error (nm) for July 4-October 31, 2005 for the Atlantic basin. The number of forecasts verified is listed below the forecast length.

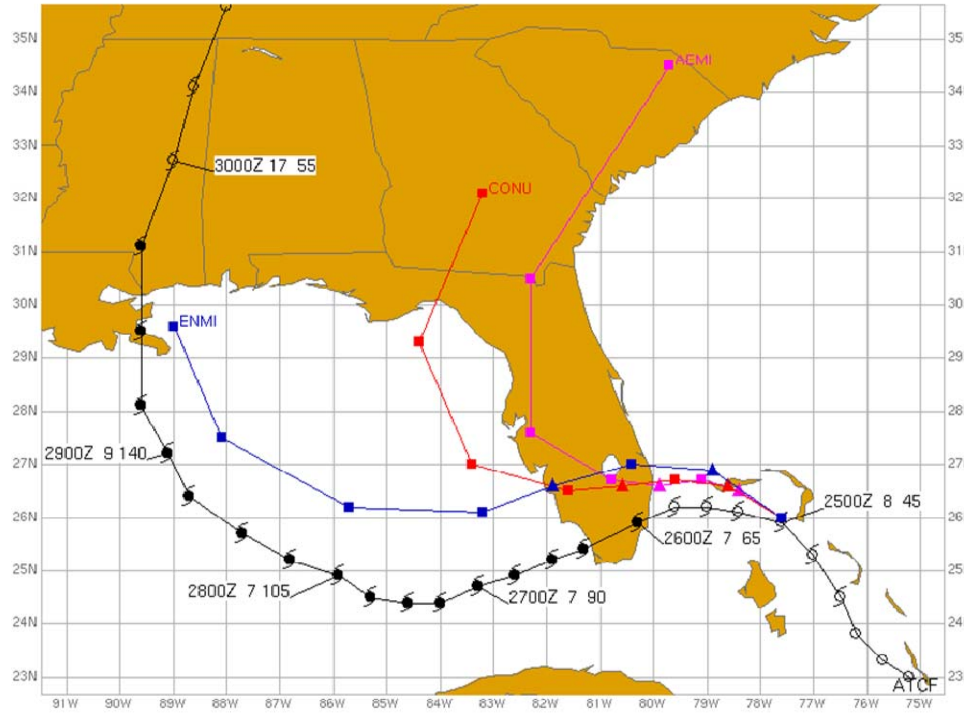

Fig. 9. Track forecasts for Hurricane Katrina for 00Z 25 August 2005. The 12-h and 36-h forecast positions are denoted by triangles while those for the 24-h, 48-h, 72-h, 96-h, and 120-h forecasts are denoted by squares.

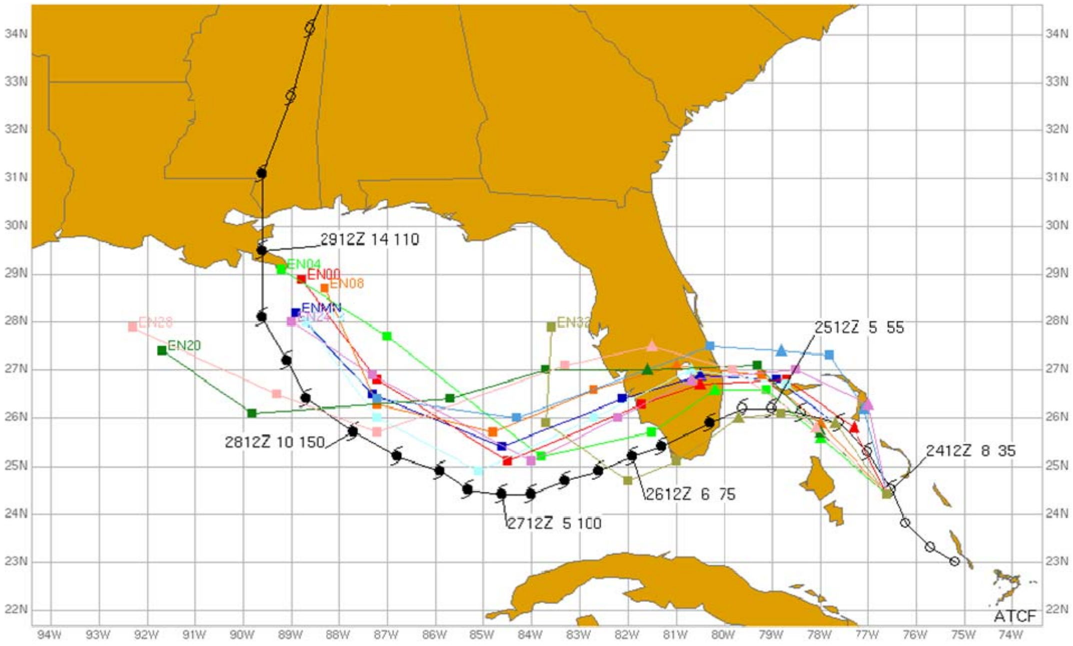

Fig. 10. NOGAPS ET ensemble (with the addition of stochastic convection) track forecasts for Hurricane Katrina for 12Z 24 August 2005. The 12-h and 36-h forecast positions are denoted by triangles while those for the 24-h, 48-h, 72-h, 96-h, and 120-h forecasts are denoted by squares.

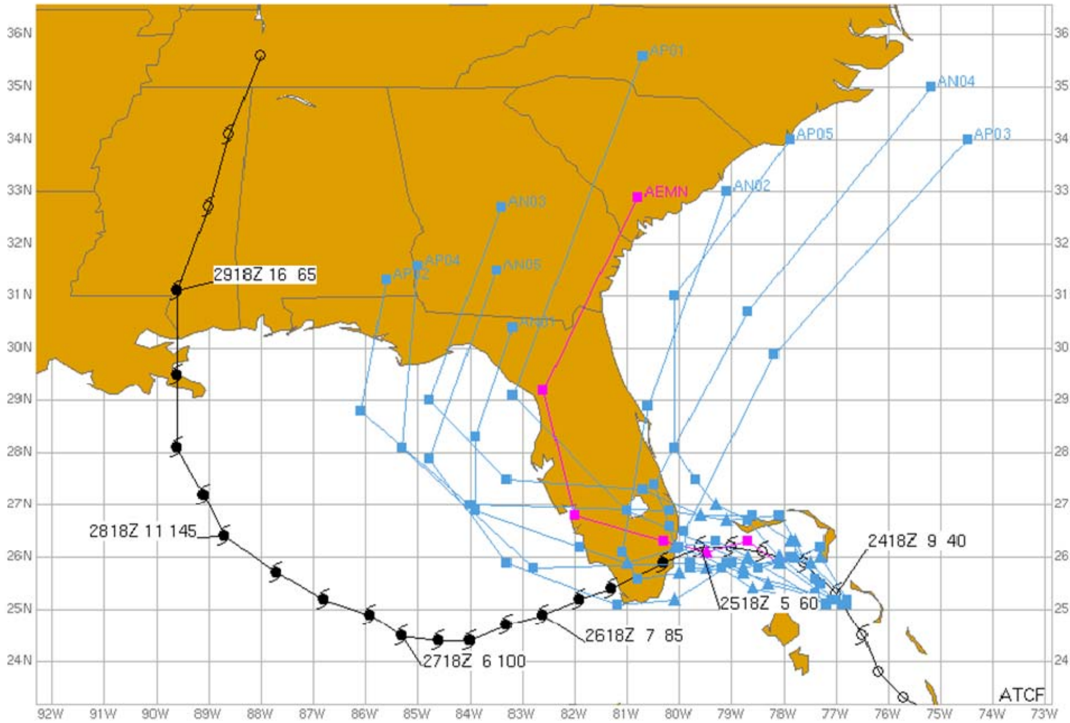

Fig. 11. NCEP GFS ensemble track forecasts for Hurricane Katrina for 18Z 24 August 2005. The 12-h and 36-h forecast positions are denoted by triangles while those for the 24-h, 48-h, 72-h, 96-h, and 120-h forecasts are denoted by squares.

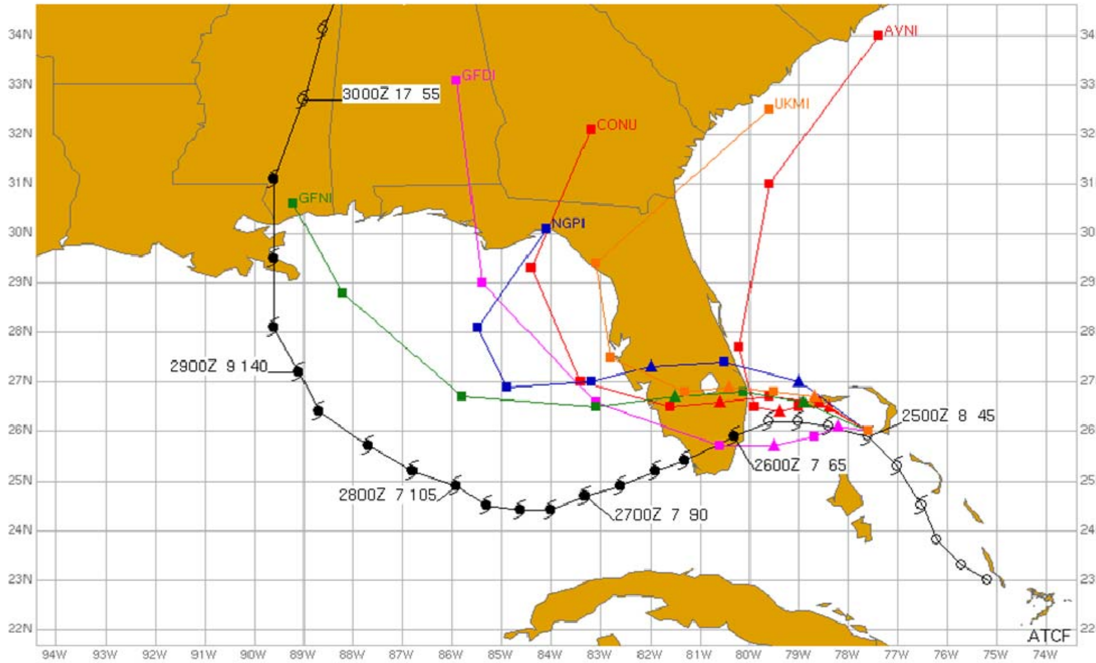

Fig. 12. CONU ensemble track forecasts for Hurricane Katrina for 00Z 25 August 2005. The 12-h and 36-h forecast positions are denoted by triangles while those for the 24-h, 48-h, 72-h, 96-h, and 120-h forecasts are denoted by squares.