Ensemble Tropical Cyclone Forecast Performance And Prediction of Ensemble Forecast Error

James S. Goerss (SAIC) and Charles R. Sampson
NRL Monterey

April 4, 2019

Tropical Cyclone Forecasting at NHC and JTWC

The primary variables forecast by the U.S. tropical cyclone (TC) forecast centers, National Hurricane Center (NHC), Miami, FL and Joint Typhoon Warning Center (JTWC), Pearl Harbor, HI, are TC track, intensity, and radius of gale-force winds.

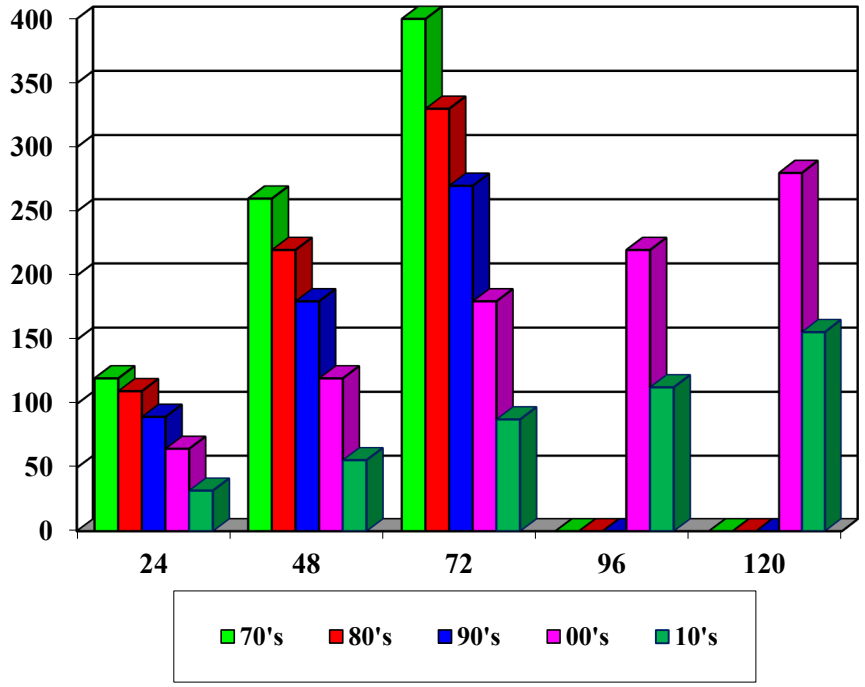
For all three variables the primary forecast guidance products used by the forecasters are multi-model ensemble mean or consensus forecasts derived using the forecasts from global and regional numerical weather prediction (NWP) models and from a number of statistical models.

The consensus forecast guidance for TC track is derived entirely from NWP model forecasts while that for TC intensity and radius of gale-force winds use a combination of NWP and statistical model forecasts.

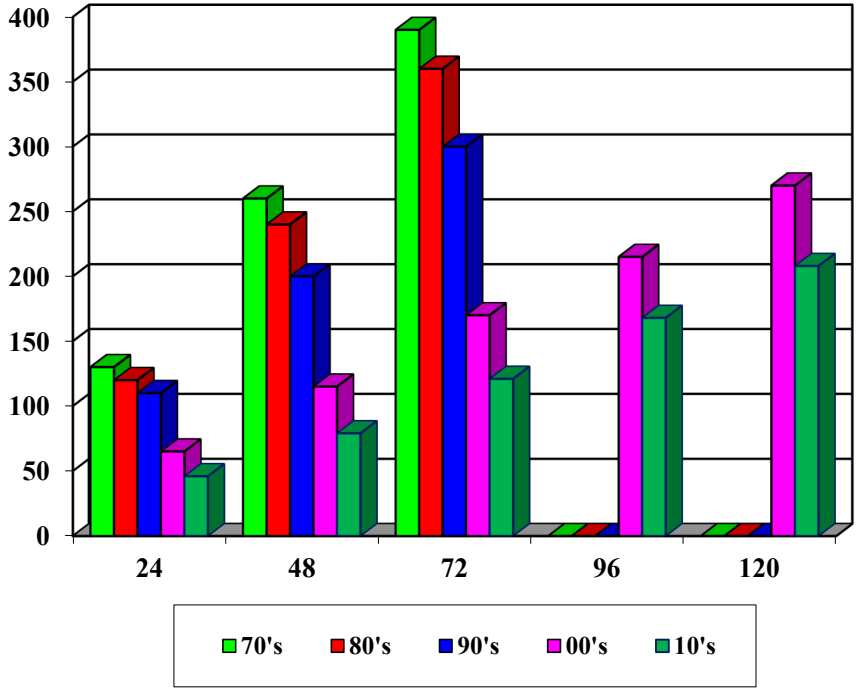
Consensus forecast guidance for radius of gale-force winds is being used operationally at JTWC but is still in its experimental stage at NHC. In this presentation we will focus on consensus forecast guidance for track and intensity.

Evolution of TC Track Forecast Error (nm) at NHC and JTWC

NHC Atlantic

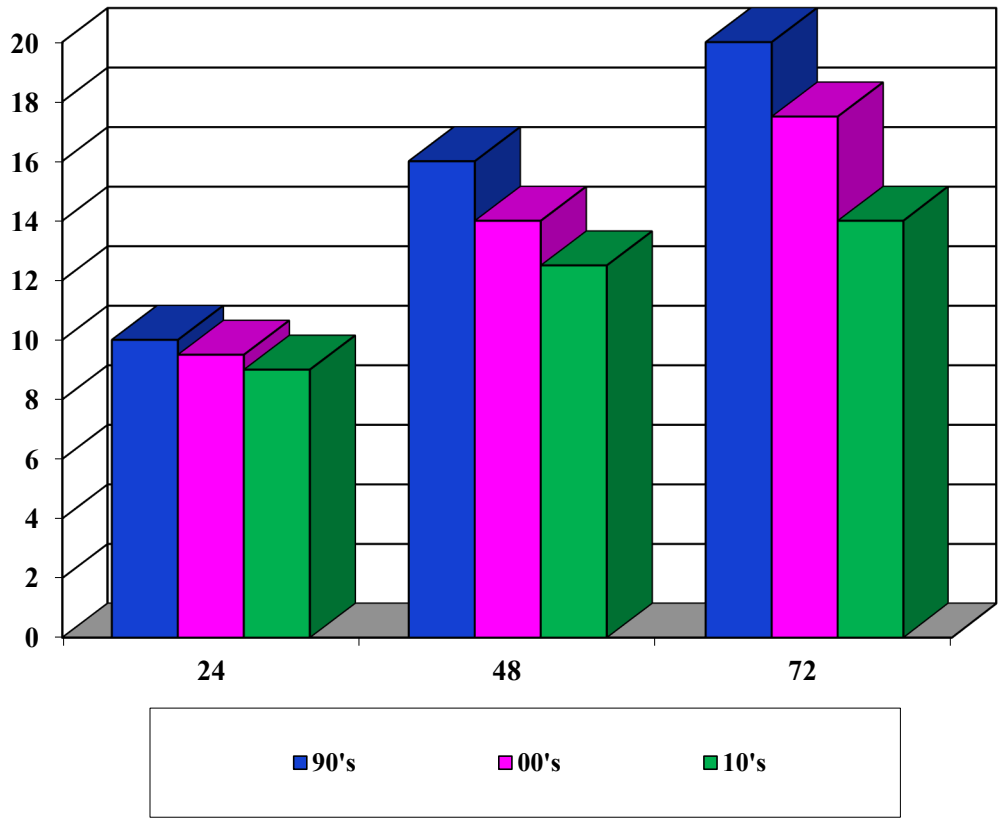


JTWC Western North Pacific



Evolution of TC Intensity Forecast Error (kt)

NHC Atlantic



NHC Track Forecast Guidance

Interpolated Model Guidance

GFSI – Global Forecast System (GFS) run at NCEP

HWFI – Hurricane WRF run at NCEP

EGRI – UK Met Office global model

EMXI – ECMWF global model

CTCI – COAMPS TC run at FNMOC

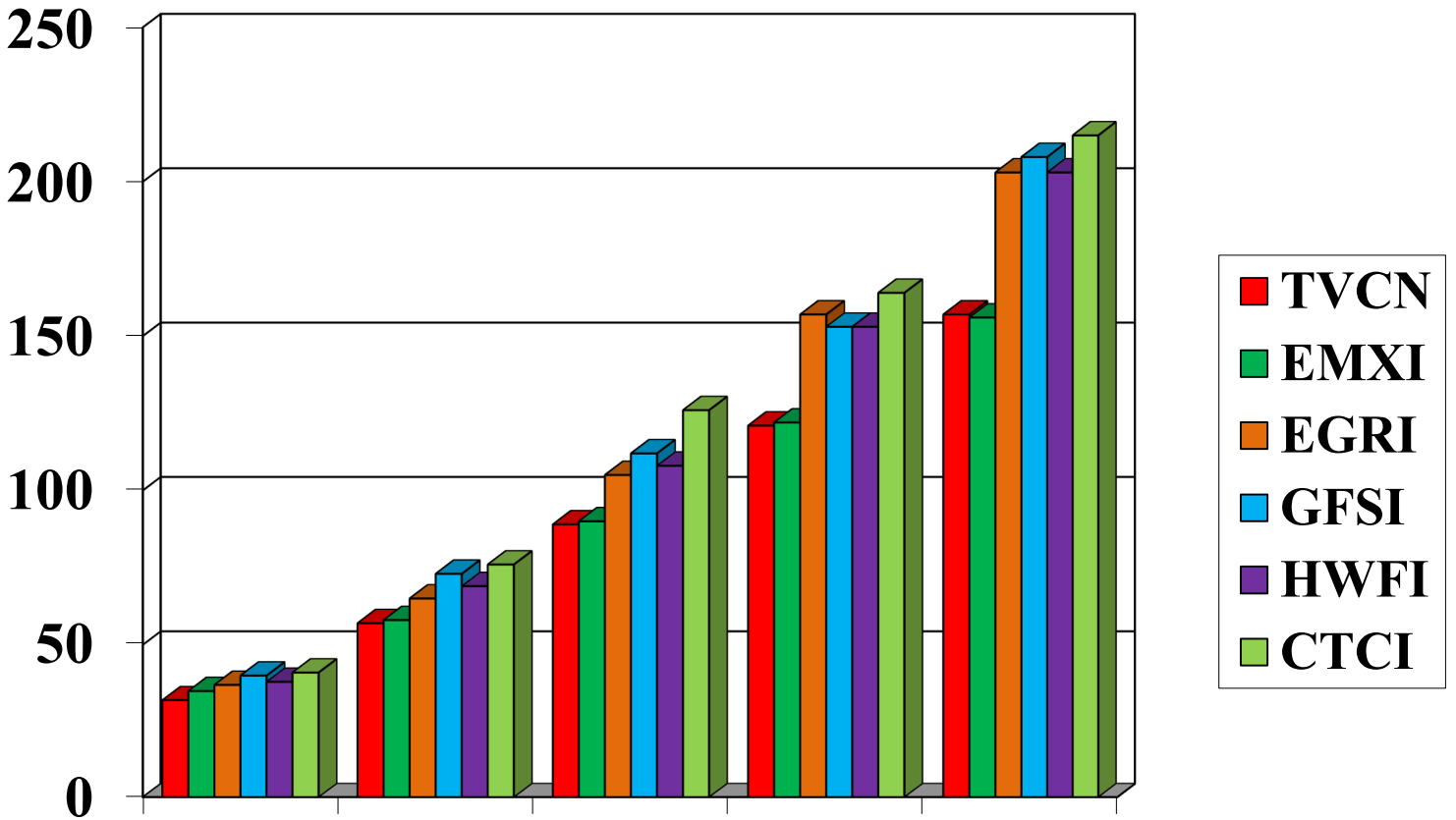
TVCN – Consensus of above models (at least two)

AEMI – NCEP Ensemble Mean

UEMI – UK Met Office Ensemble Mean

EMNI – ECMWF Ensemble Mean

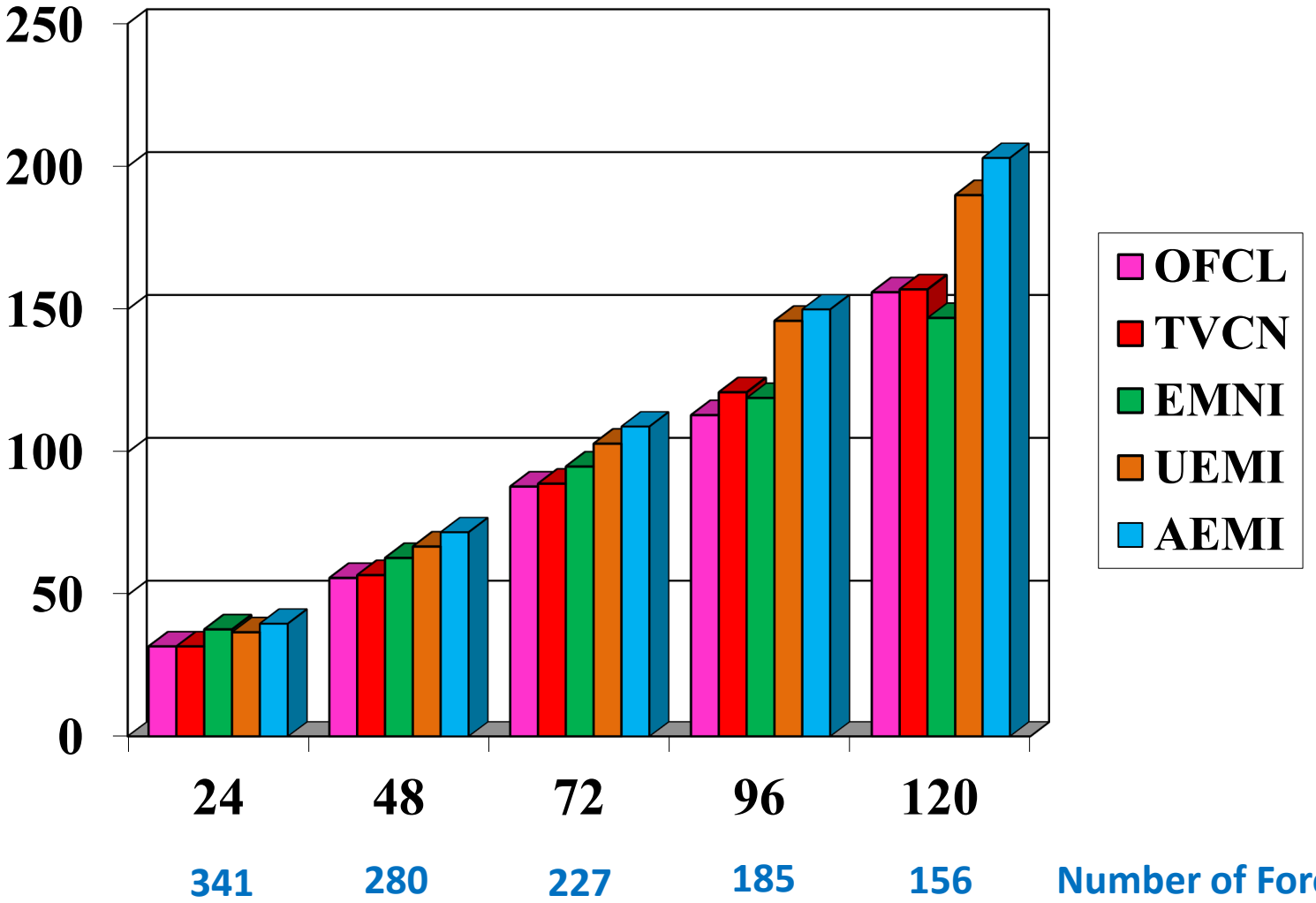
2017 Atlantic TC Forecast Error (nm)
Multi-model Consensus and Model Errors



24	48	72	96	120
341	280	227	185	156

Number of Forecasts

2017 Atlantic TC Forecast Error (nm)
Ensemble Mean Errors



NHC Intensity Forecast Guidance

Interpolated Model Guidance

HWFI – Hurricane WRF run at NCEP

HMNI – HMON non-hydrostatic model run at NCEP

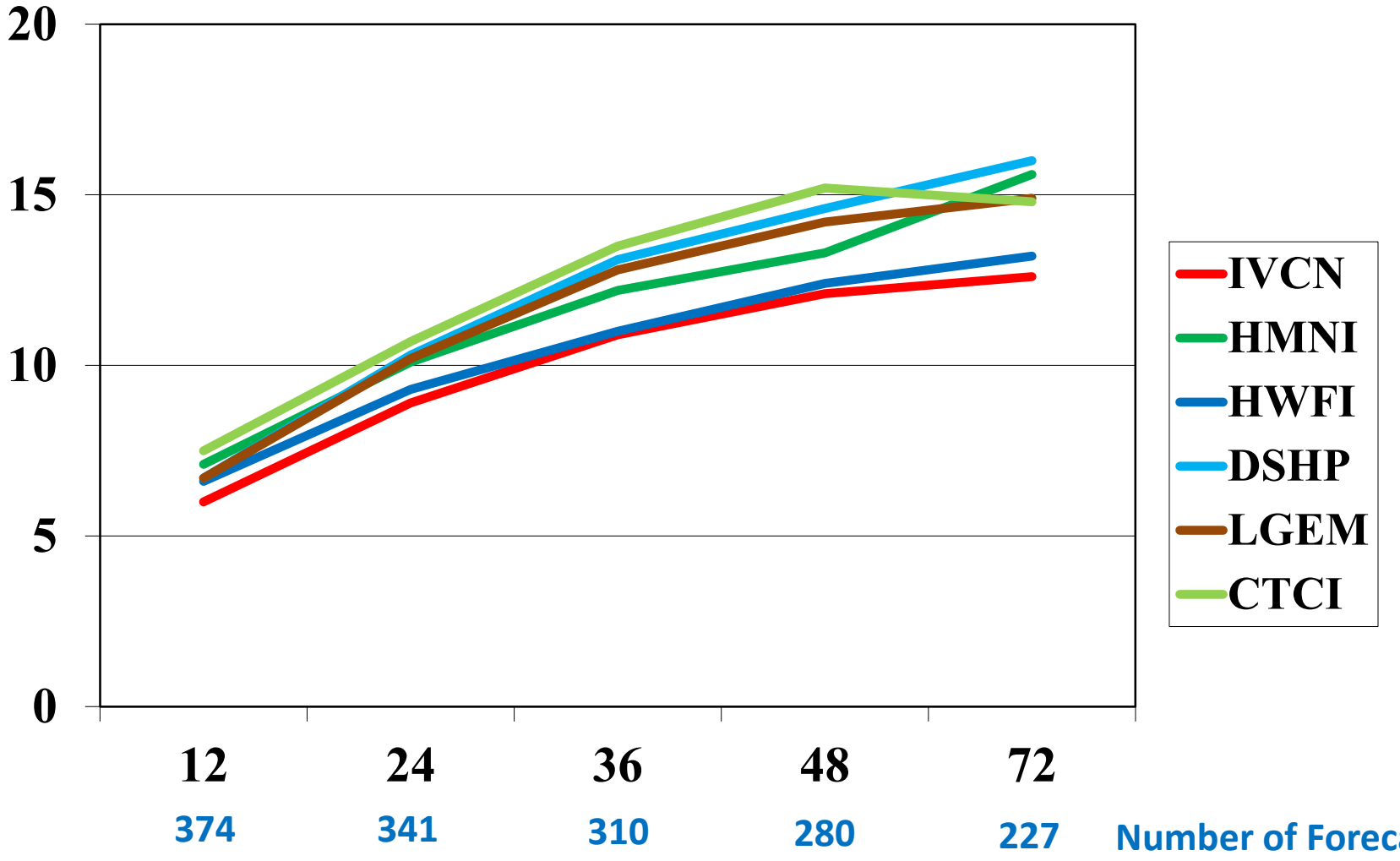
LGEM – Logistic growth equation model

DSHP – SHIPS model with inland decay

CTCI – COAMPS TC run at FNMOC

IVCN – Consensus of above models (at least two)

2017 Atlantic TC Intensity Forecast Error (kt)
Multi-model Consensus and Model Errors



TC Track Forecast Performance of Multi-model and Single-model Ensembles

The forecast position error for member i, E_i, is defined to be:

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

where C_i and A_i are the cross-track and along-track errors, respectively .

If, for simplicity, we assume that, for every i, C_i and A_i are independent and normally distributed with zero mean and standard deviation σ, then

E_i possesses a Rayleigh distribution with mean $\sigma(\pi/2)^{1/2}$.

Since each ensemble mean forecast position is merely the mean of the individual member forecast positions, one can easily show that the ensemble mean cross-track and along-track errors, denoted by C_e and A_e, are simply the means of the cross-track and along-track errors of the individual members.

Therefore, for an ensemble with n members, C_e and A_e are normally distributed with zero mean and standard deviation $\sigma/n^{1/2}$ (Central Limit Theorem), and the ensemble mean forecast error,

E_e, possesses a Rayleigh distribution with mean $\sigma(\pi/2n)^{1/2}$.

In practice, the C_i and A_i are not independent, and n denotes the

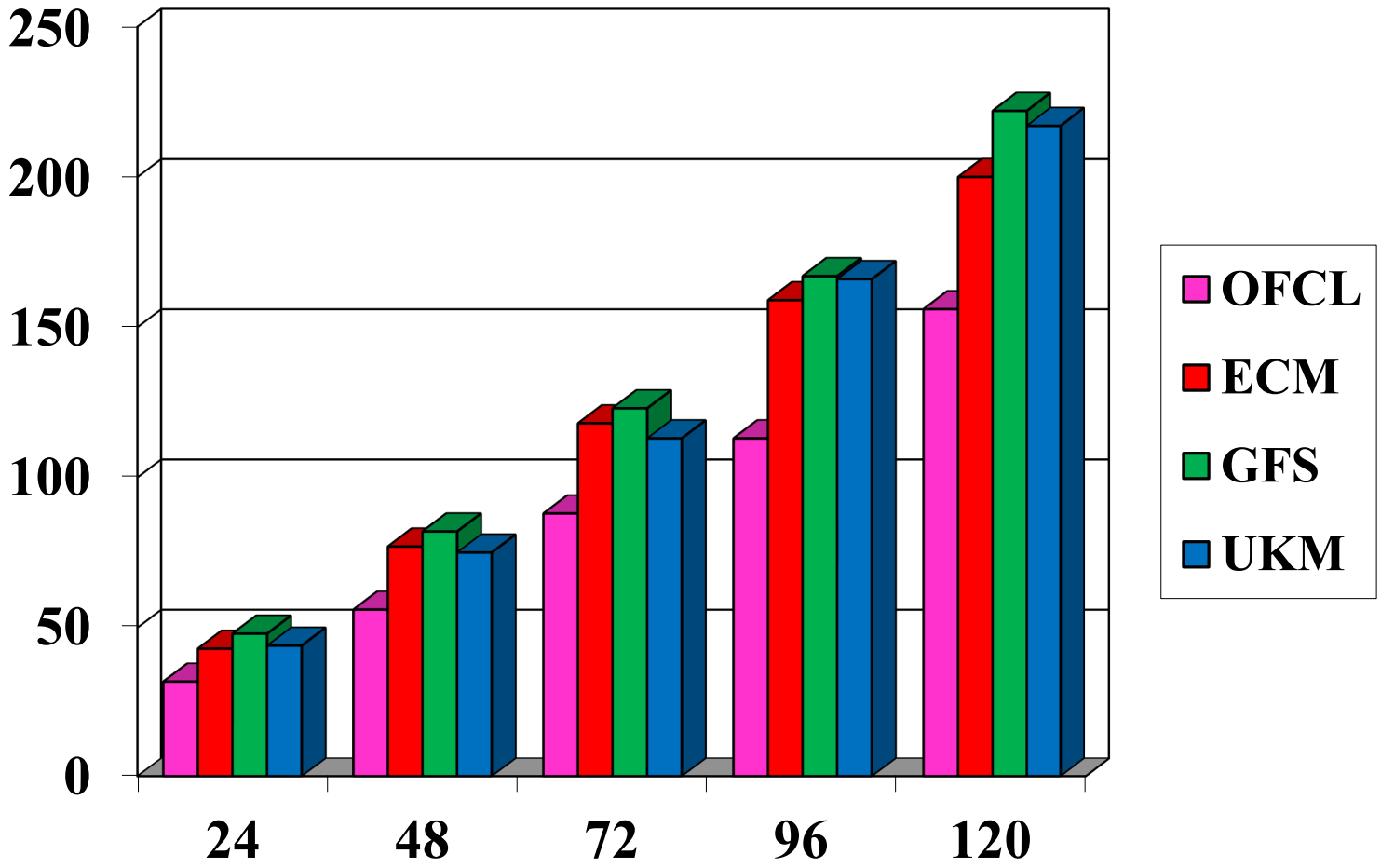
From the previous discussion we see that the ensemble mean forecast error is dependent upon two things:

(1) The mean forecast error of the individual members that make up the ensemble

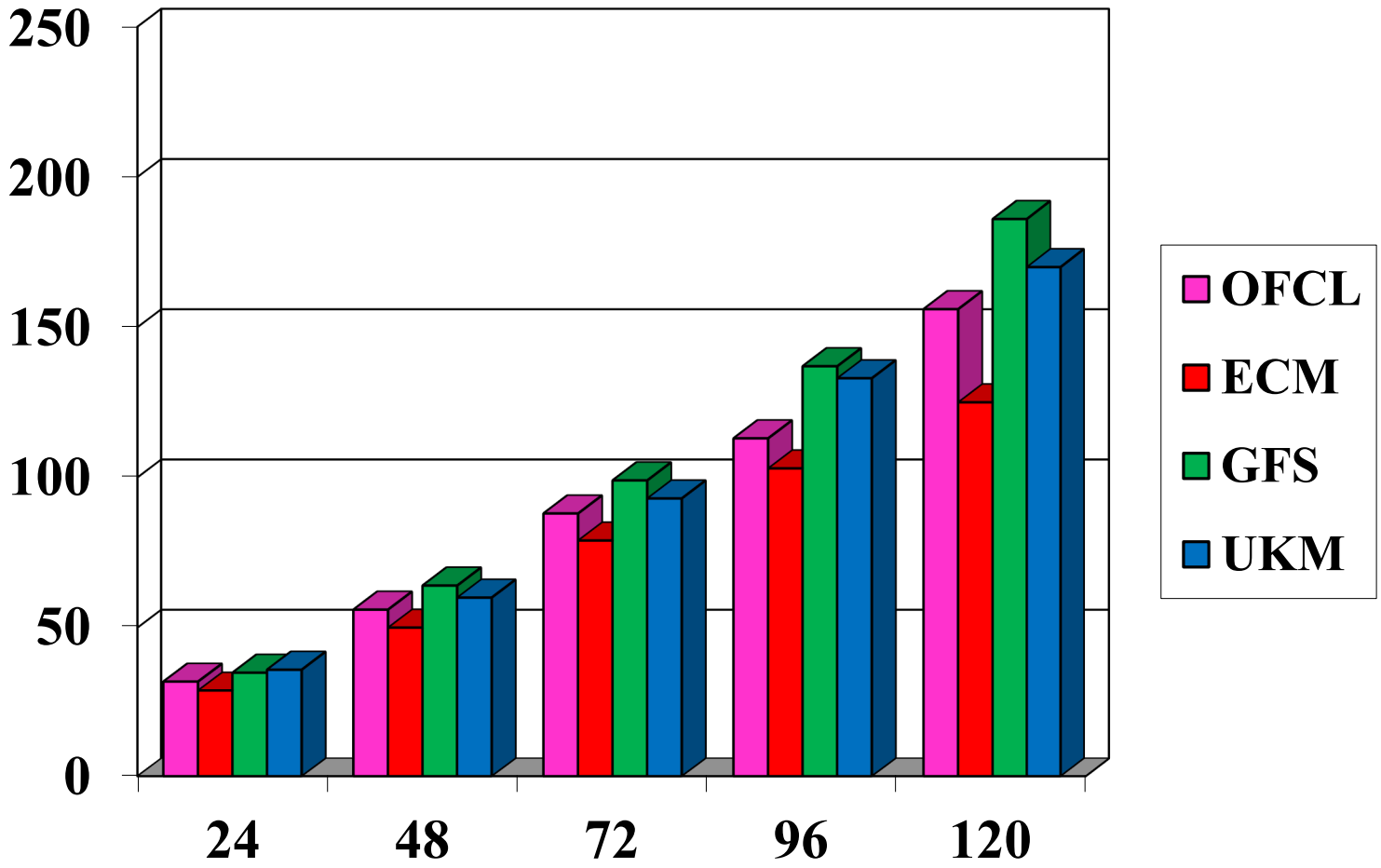
(2) The degree of independence of the forecast errors of the individual members.

Also, we see that the ratio of the mean member forecast error to the ensemble mean forecast error is $n^{1/2}$, where n is the effective degrees of freedom.

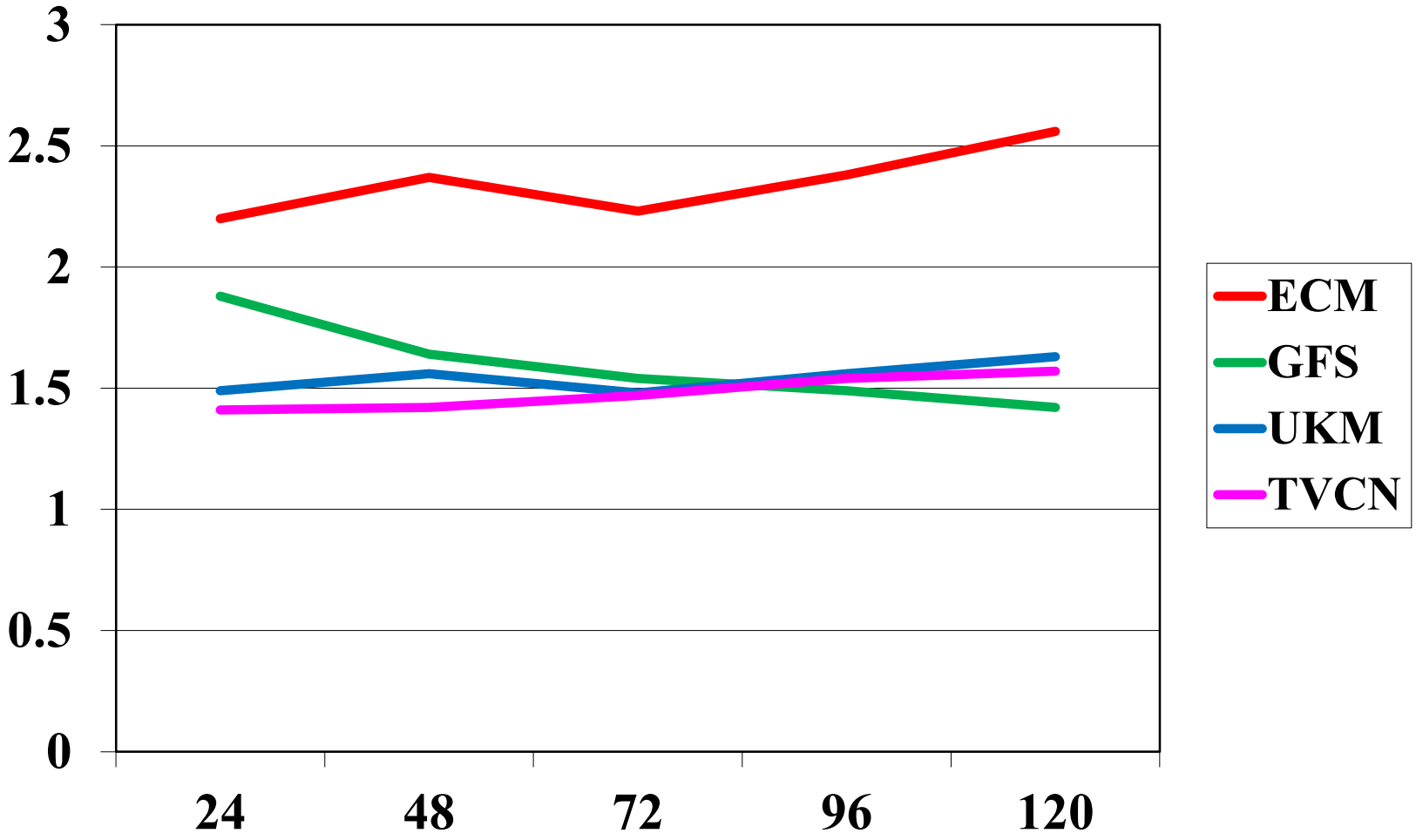
2017 Atlantic TC Forecast Error (nm)
Mean of Ensemble Member Errors



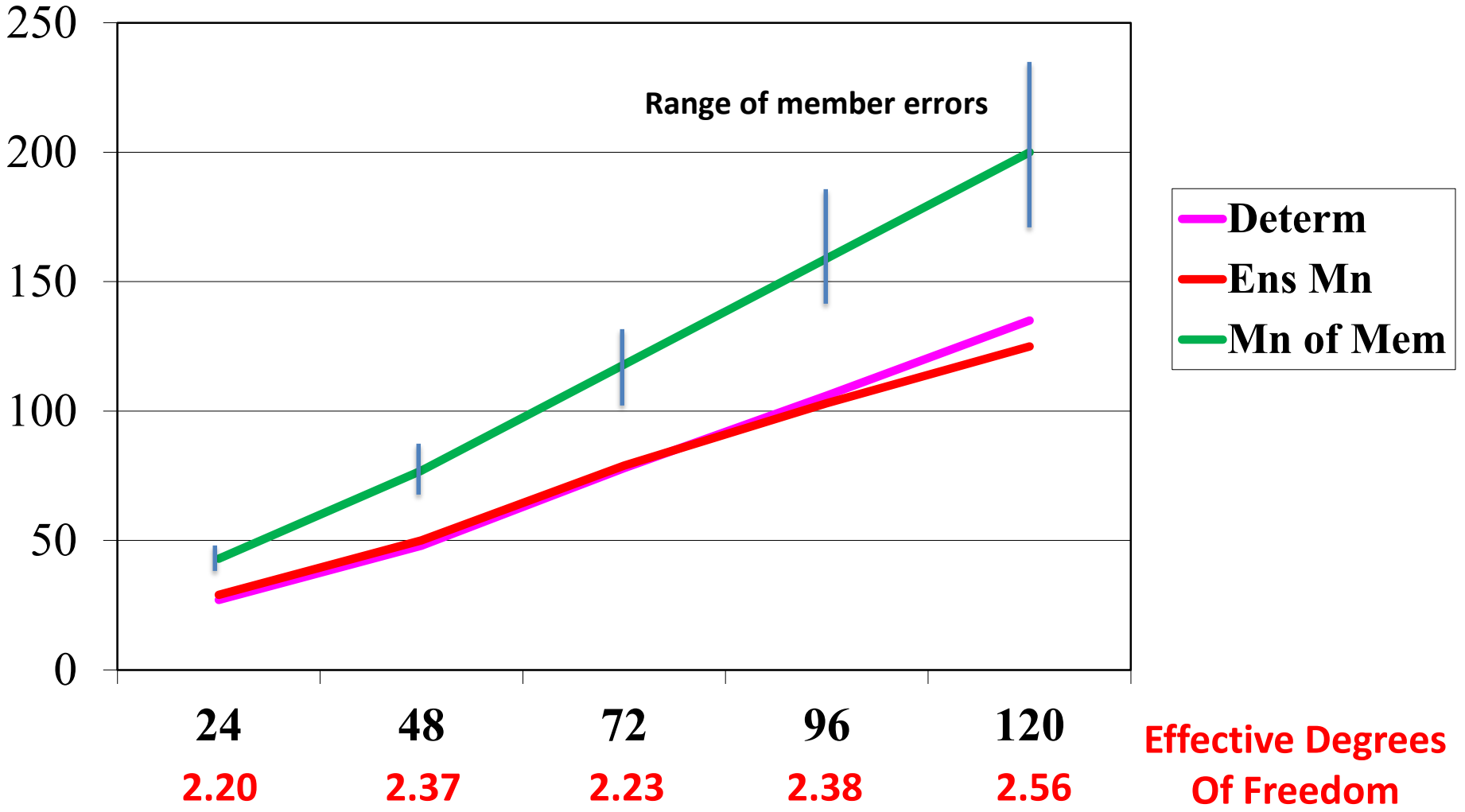
2017 Atlantic TC Forecast Error (nm)
Ensemble Mean Errors



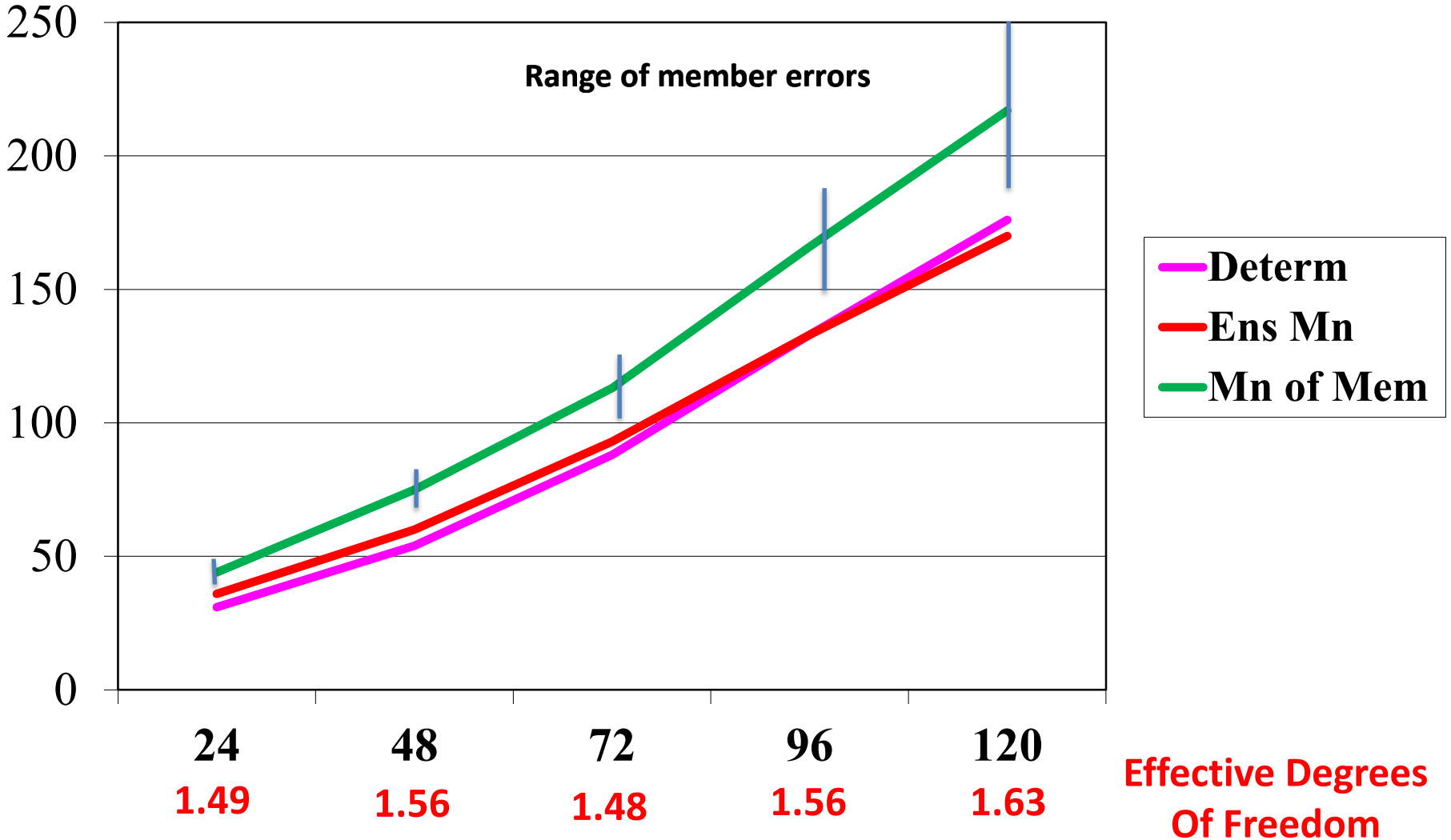
2017 Atlantic Ensembles
Effective Degrees of Freedom



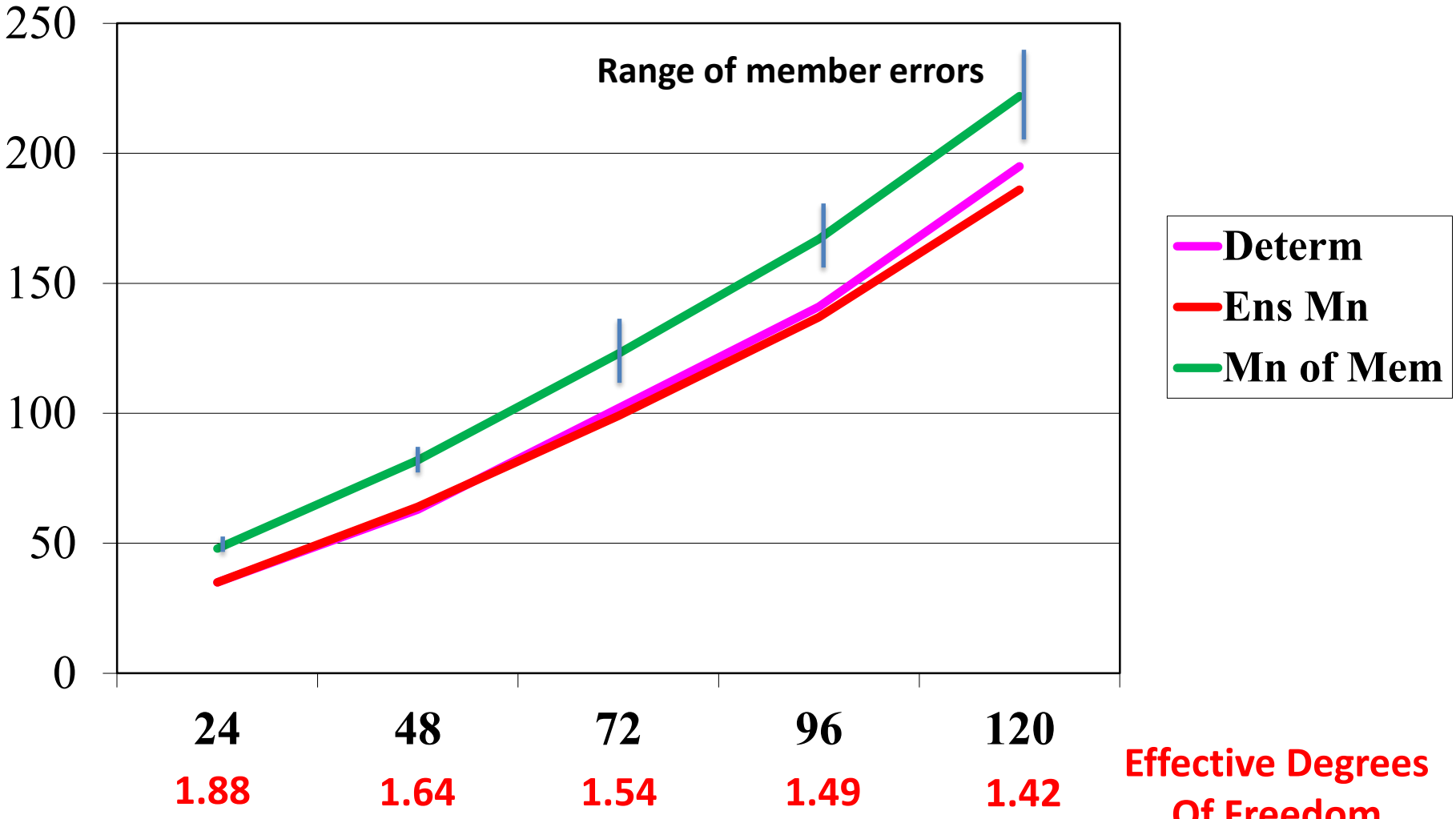
2017 Atlantic TC Forecast Error (nm)
ECMWF (50-member Ensemble)



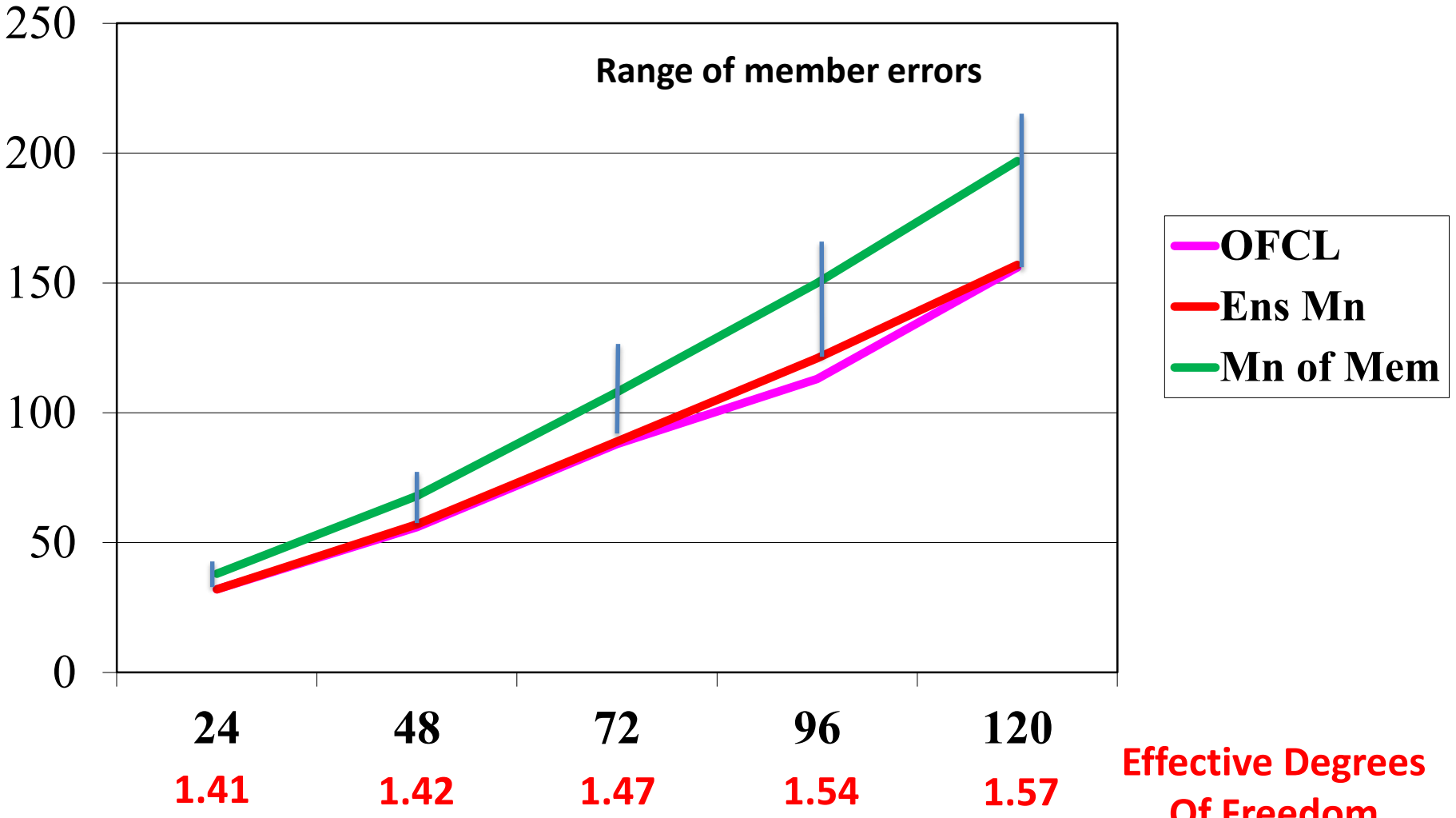
2017 Atlantic TC Forecast Error (nm)
UK Met (35-member Ensemble)



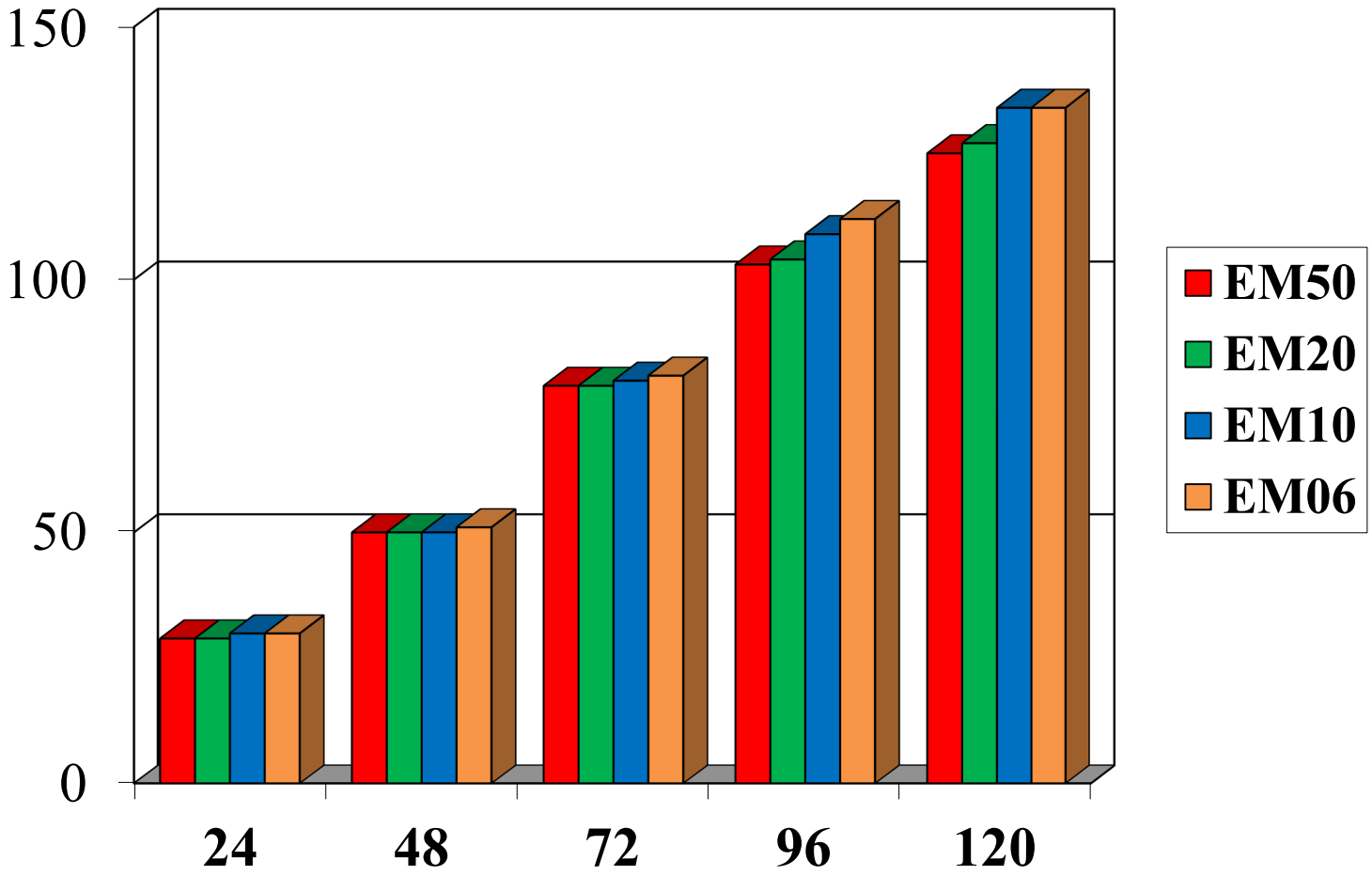
2017 Atlantic TC Forecast Error (nm)
GFS (20-member Ensemble)



2017 Atlantic TC Forecast Error (nm)
TVCN (5-member Ensemble)



2017 Atlantic TC Forecast Error (nm)
ECMWF



Prediction of Consensus Forecast Error

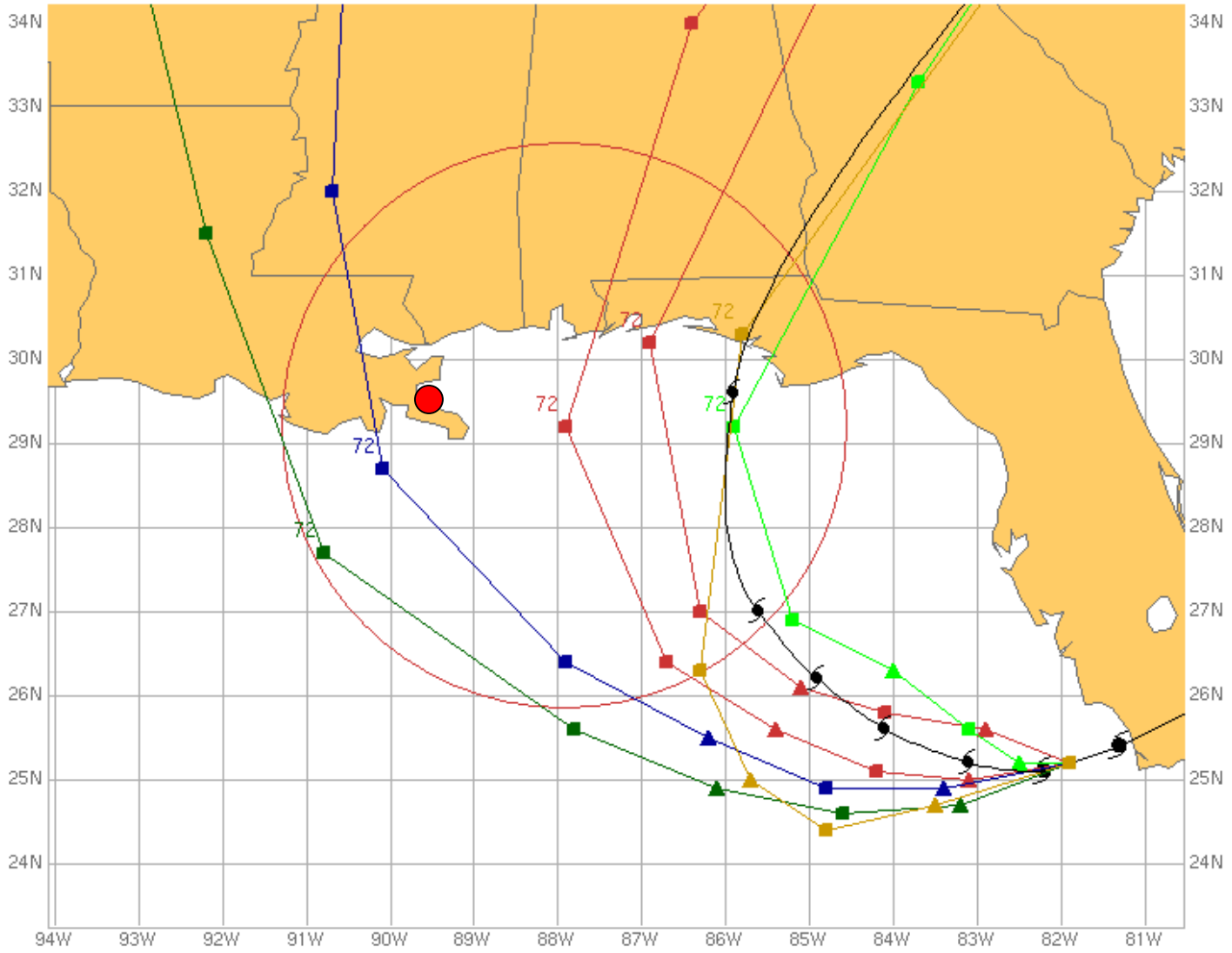
Prediction of Consensus TC Track Forecast Error

- Predictors of consensus forecast error must be quantities that are available prior to the time when official forecasts must be issued.
- Consensus model spread is defined to be the average distance of the member forecasts from the consensus forecast.
- The possible predictors are consensus model spread; initial and forecast TC intensity; initial TC location and forecast displacement of TC location (latitude and longitude); and TC speed of motion.

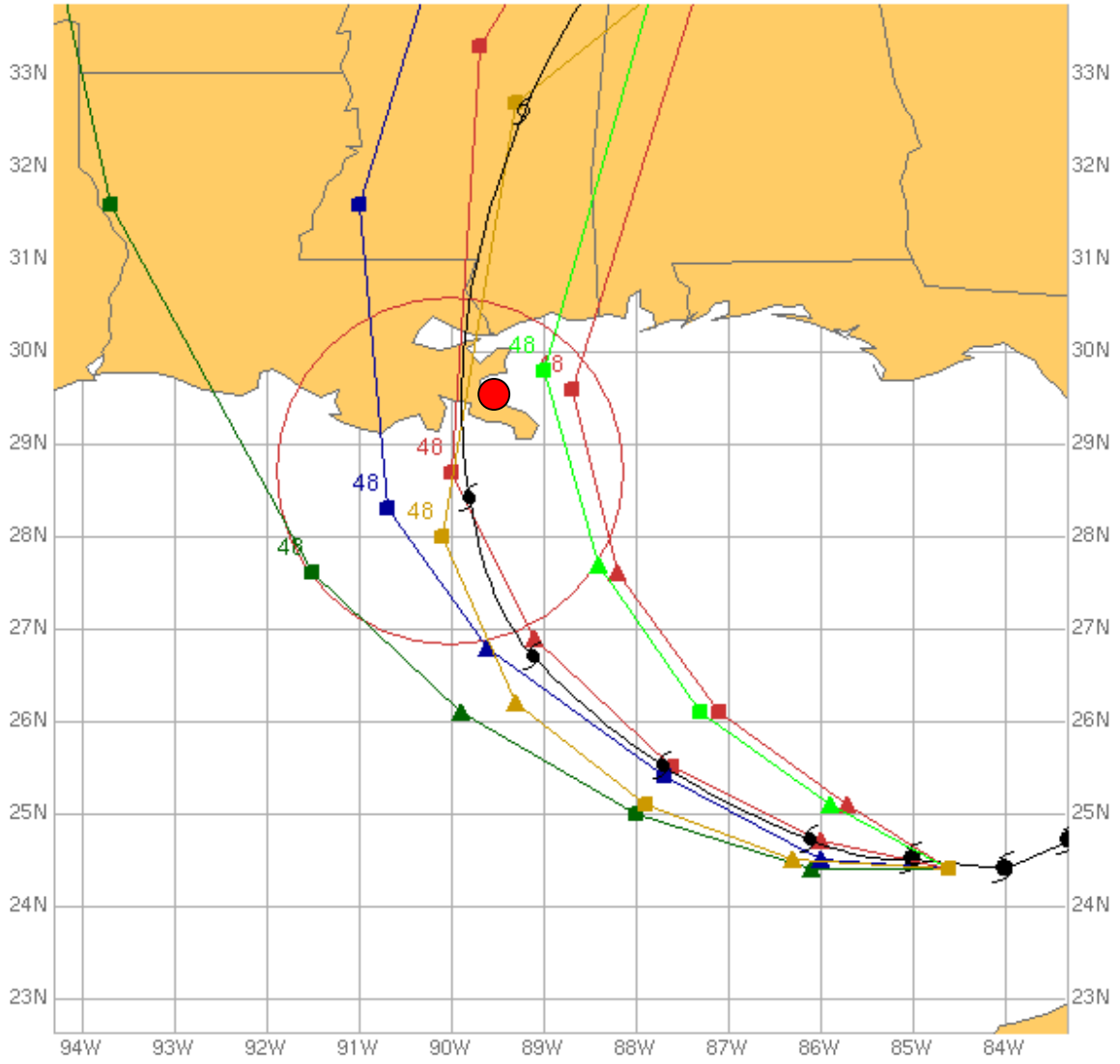
Prediction of Consensus TC Track Forecast Error

- Using stepwise linear regression and the aforementioned pool of predictors for previous seasons, regression models are found to predict consensus TC track forecast error for each combination of forecast length, consensus model, and basin.
- The regression models are then used to determine the radii of circular areas drawn around the consensus model forecast positions within which the verifying TC position is expected to be contained approximately 67% of the time.
- These circular areas are graphically displayed on the ATCF for use by the forecasters at NHC and JTWC. This graphical predicted consensus error product is referred to as GPCE (“gypsy”).

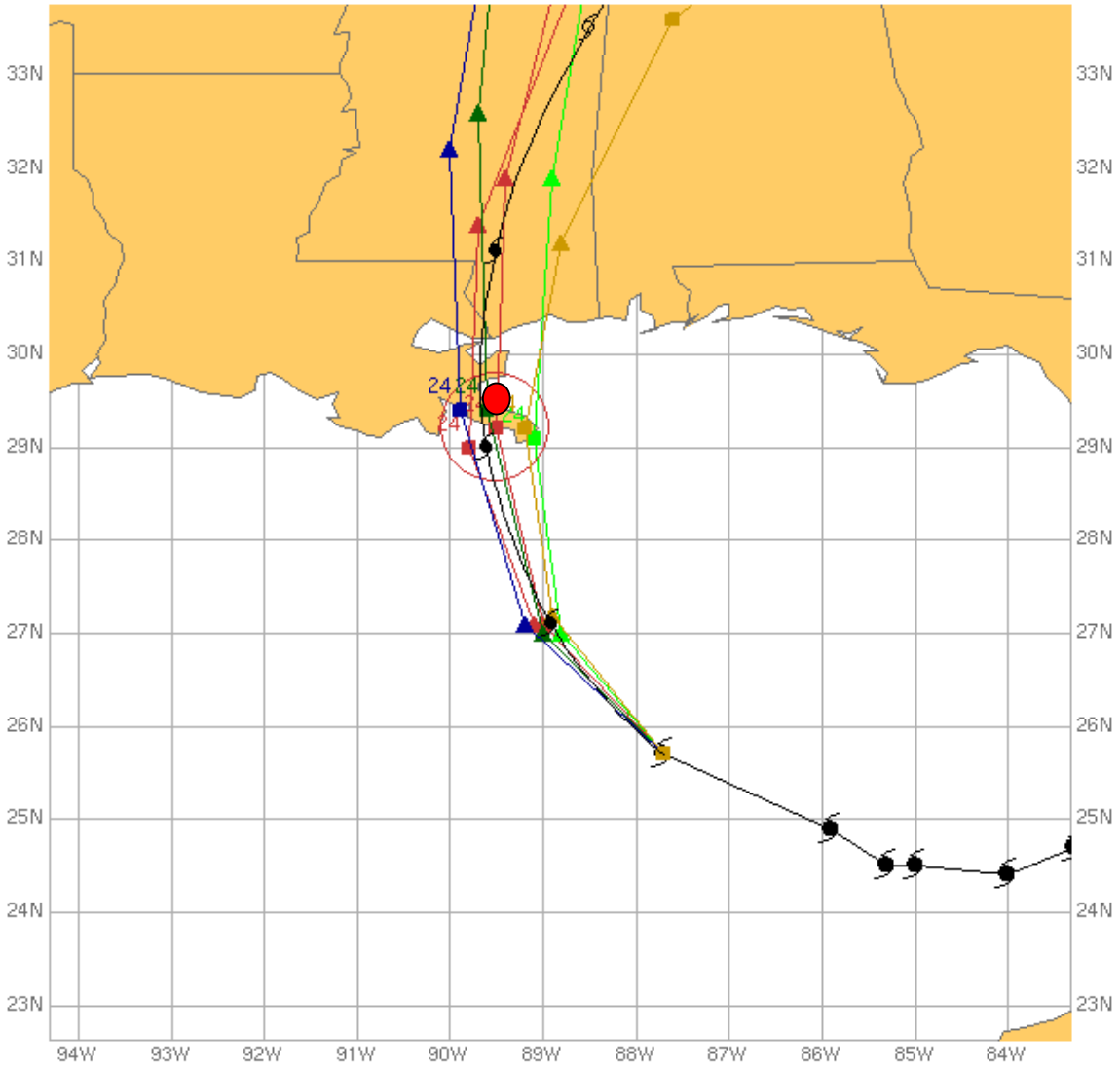
72-h Predicted Consensus Error
Hurricane Katrina - 12Z 26 August 2005



48-h Predicted Consensus Error
Hurricane Katrina - 12Z 27 August 2005



24-h Predicted Consensus Error
Hurricane Katrina - 12Z 28 August 2005



Track GPCE Performance
2017 Atlantic

- **For the 2017 Atlantic season, the circular areas displayed by GPCE contained the verifying TC position 67%, 67%, 67%, 75%, and 71% of the time at 24 h, 48 h, 72 h, 96 h, and 120 h, respectively.**

- **For the 2017 Atlantic season, the GPCE predicted radii for TVCN varied from 10-75 nm at 24 h cf an average of 38 nm, 21-142 nm at 48 h cf an average of 67 nm, 40-243 nm at 72 h cf an average of 103 nm, 96-490 nm at 96 h cf an average of 154 nm, and 124-365 nm at 120 h cf an average of 198 nm.**

Prediction of Consensus TC Intensity Forecast Error

- Consensus model spread is defined to be the average of the absolute intensity differences between the member forecasts and the consensus forecast.
- The possible predictors are consensus model spread; initial and forecast TC intensity; forecast TC intensity change; initial TC position; and TC speed of motion.
- As for TC track, regression models are found to predict consensus TC intensity forecast error for each combination of forecast length, consensus model, and basin.
- The regression models are then used to determine the half-widths of intervals centered on the consensus model forecast intensities within which the verifying TC intensity is expected to be contained approximately 67% of the time.
- These intervals are graphically displayed on the ATCF for use by the forecasters at NHC and JTWC.

Intensity GPCE Performance 2017 Atlantic

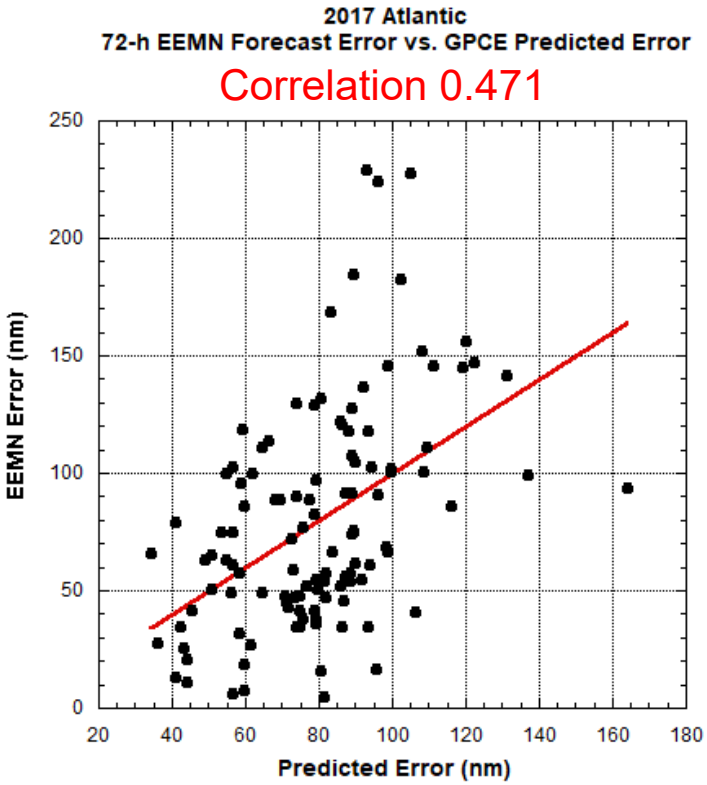
- For the 2017 Atlantic season, the intervals displayed by GPCE centered on the IVCN forecast intensity contained the verifying TC intensity 69%, 68%, and 66% of the time at 24 h, 48 h, and 72 h, respectively.

- For the 2017 Atlantic season, the GPCE predicted half-widths for IVCN varied from 6-17 kt at 24 h cf an average of 10 kt, 3-23 kt at 48 h cf an average of 13 kt, and 2-28 kt at 72 h cf an average of 15 kt.

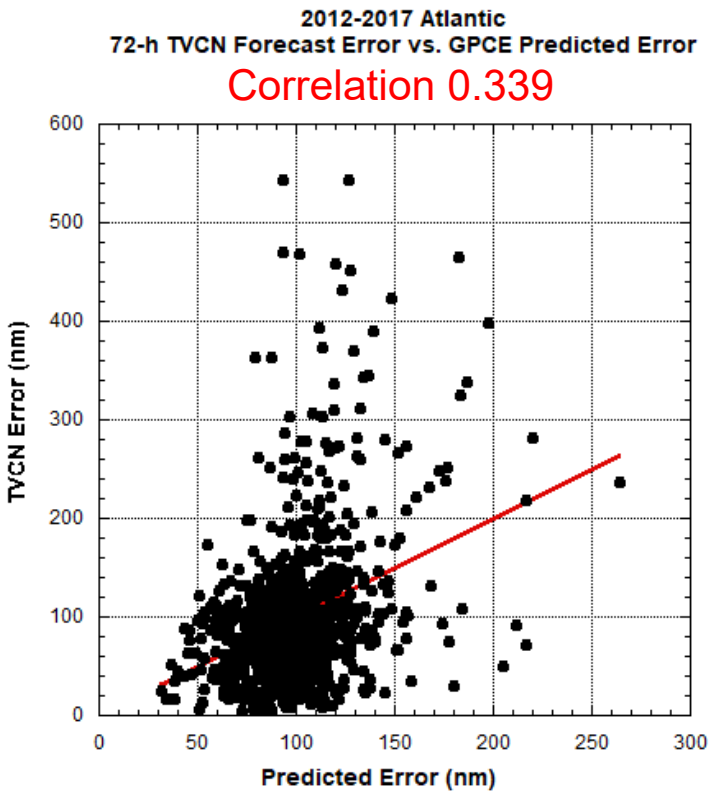
Application of GPCE to ECMWF Ensemble

- The GPCE technique was applied to the ECMWF ensemble for the 2017 Atlantic season. For every forecast length, the leading predictor was found to be ensemble spread with initial intensity as the second leading predictor. For the dependent dataset, the variance of the ensemble mean forecast error explained by the predicted error ranged from 13-22 percent.
- When applied to TVCN for the 2012-2017 Atlantic seasons, the leading predictor was also found to be ensemble spread for every forecast length, and initial intensity was the second leading predictor for the 12-72 h forecasts. For the dependent dataset, the variance of the TVCN forecast error explained by the predicted error ranged from 11-22 percent.

Application of GPCE to ECMWF Ensemble



PE = .540*SPR - .324*INTI + 57



PE = .547*SPR - .438*INTI + 86

Summary

- There have been huge improvements in TC track forecasts from the 1990's to the 2010's due to individual NWP model improvements and the use of multi-model consensus forecasts.
- Over the same period there have been only modest improvements in TC intensity forecasts. While the best model guidance is now from an NWP model (HWRF), statistical models still make up a large part of multi-model consensus guidance.
- For both TC track and intensity, multi-model consensus forecasts provide the primary guidance used by the forecasters at NHC and JTWC. For the 2017 Atlantic season, the TC track forecasts for the single-model ECMWF ensemble mean were comparable to those for the multi-model consensus TVCN.

Summary

- The TC track forecast performance of single-model and multi-model ensembles was examined for the 2017 Atlantic season. Ensemble mean forecast error is dependent on the mean forecast error of the ensemble members and the degree of independence of the forecast error of the members. Primarily, because of the degree of independence of its members, the TC track forecast performance of the ECMWF was found to be the best of the single-model ensembles.
- The performance of GPCE for TC track and intensity was illustrated for the 2017 Atlantic season.
- The GPCE technique was applied to the ECMWF ensemble. The results were strikingly similar to those obtained for the multi-model consensus TVCN.

Questions?

References

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

Goerss, J. S., 2007: Prediction of consensus tropical cyclone track forecast error. *Mon. Wea. Rev.*, **135**, 1985–1993.

Goerss, J. S., and C. R. Sampson, 2014: Prediction of consensus tropical cyclone intensity forecast error. *Wea. Forecasting*, **29**, 750–762.

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

Sampson, C. R., and J. A. Knaff, 2015: A consensus forecast for tropical cyclone gale wind radii. *Wea. Forecasting*, **30**, 1397–1403.

Sampson, C. R., E. M. Fukada, J. A. Knaff, B. R. Strahl, M. J. Brennan, and T. Marchok, 2017: Tropical cyclone gale wind radii estimates for the western North Pacific. *Wea. Forecasting*, **32**, 1007–1028.

Sampson, C. R., J. S. Goerss, J. A. Knaff, B. R. Strahl, E. M. Fukada, and E. A. Serra, 2018: Tropical cyclone gale wind radii estimates, forecasts, and error forecasts for the western North Pacific. *Wea. Forecasting*, **33**, 1081–1092.

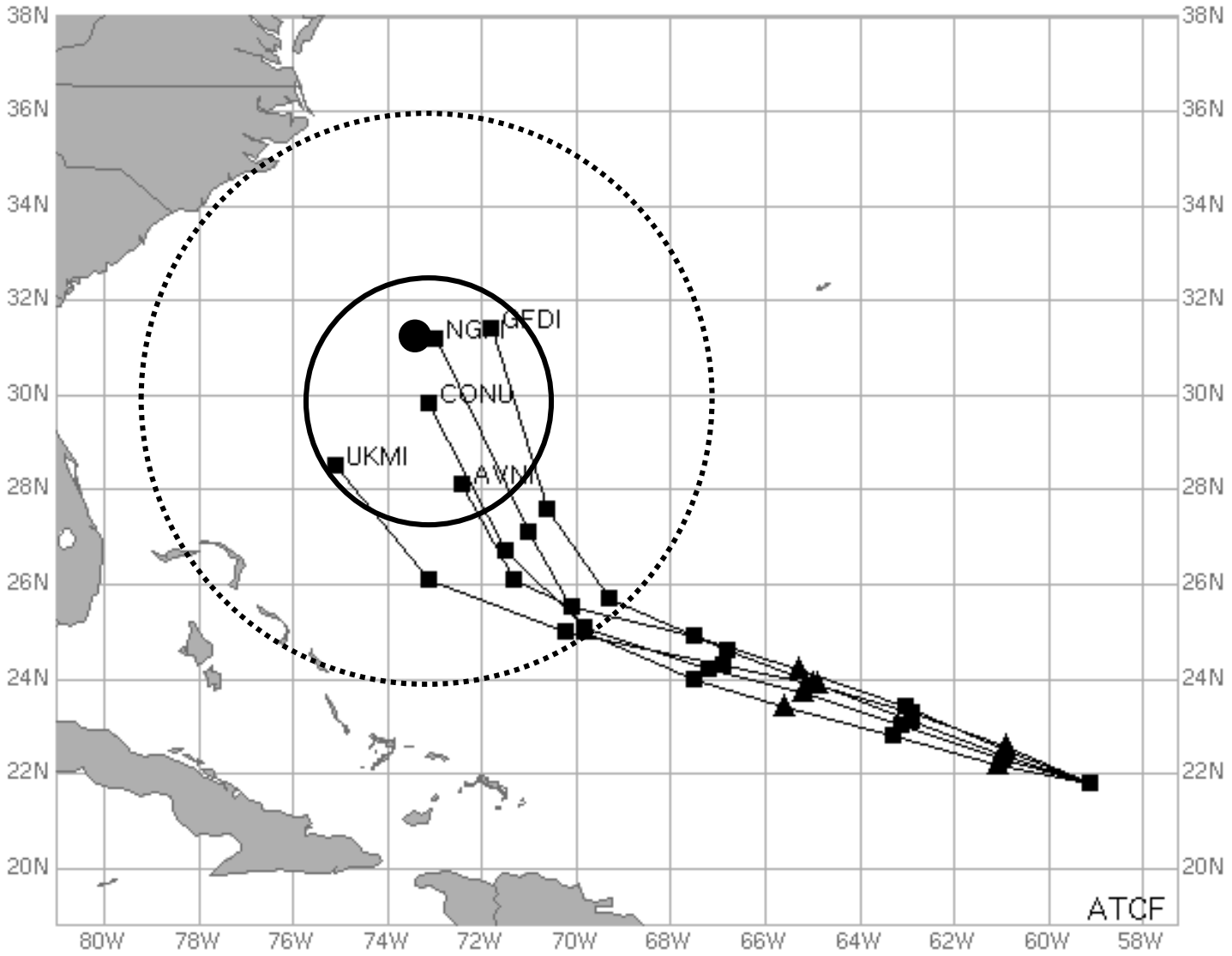


Fig. 5a. Predicted 72% confidence radius (solid) surrounding the 120-h CONU forecast for Hurricane Isabel, 00Z September 13, 2003. The individual model tracks used to create the CONU track are shown along with the 120-h radius (dotted) used by the NHC Potential Day 1-5 Track Area graphic.

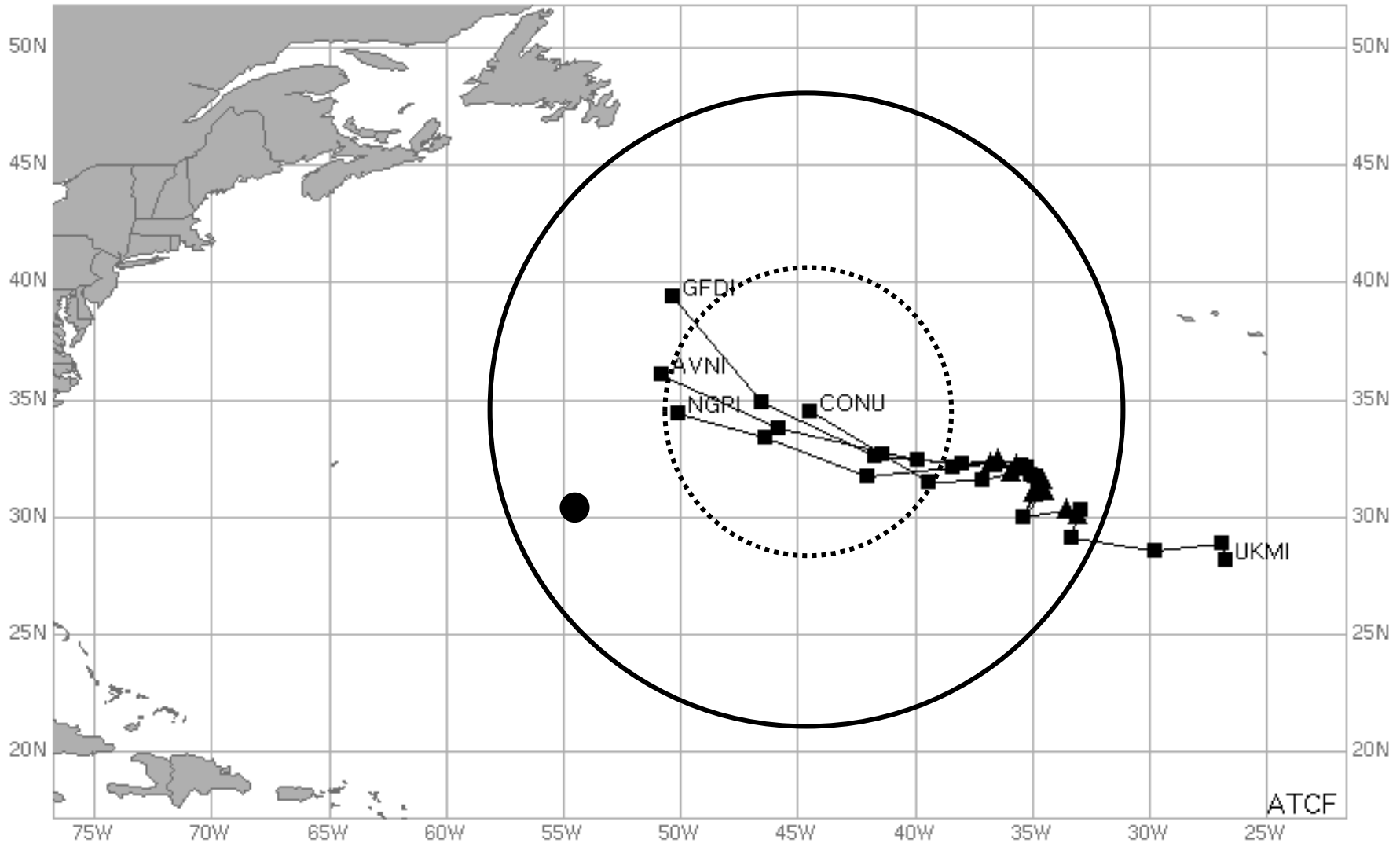


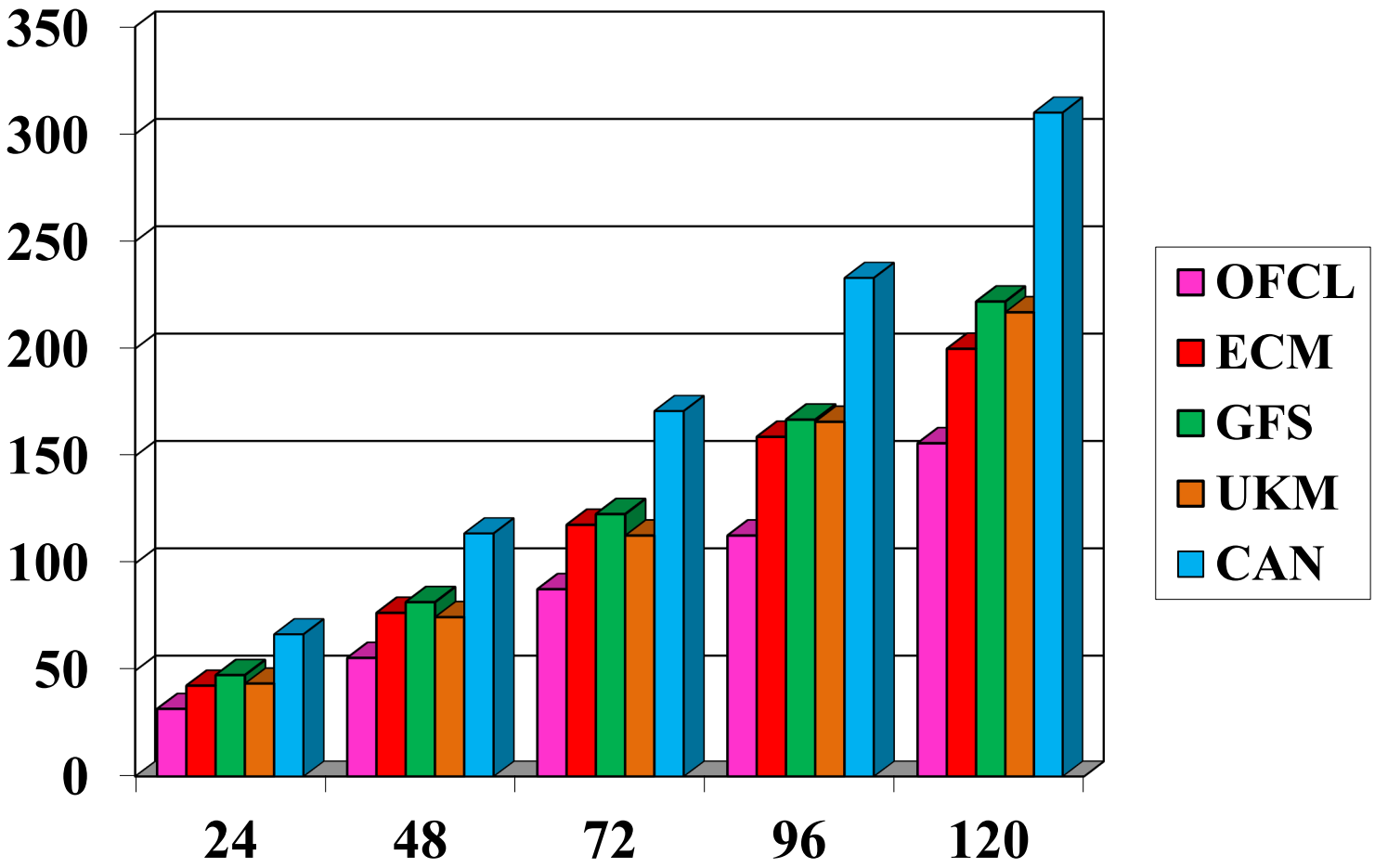
Fig. 5b. Predicted 72% confidence radius (solid) surrounding the 120-h CONU forecast for Hurricane Kate, 00Z September 30, 2003. The individual model tracks used to create the CONU track are shown along with the 120-h radius (dotted) used by the NHC Potential Day 1-5 Track Area graphic.

HMON

HMON which stands for Hurricanes in a Multi-scale Ocean-coupled Non-hydrostatic model is a new Hurricane forecast system running operationally at NCEP. This version, HMON v1.0.0, is the first version for the system. This release has been fully tested and compared with the discontinued Geophysical Fluid Dynamics Laboratory (GFDL) Hurricane Model (GHM) results and is considered its replacement. It has shown significant skill improvements in terms of storm track and intensity forecasts in Northern Atlantic (NATL), Eastern Pacific (EPAC) and the Central Pacific (CPAC) basins. HMON also provides a first step for the NCEP Environmental Modeling Centers' (EMC) efforts towards unification of operational models within the NOAA Environmental Modeling System (NEMS) framework.

The dynamical core of HMON is the Non-hydrostatic Multi-scale Model on a B grid (NMMB). It has 43 vertical levels with the model top fixed at 50 hPa. It includes vortex relocation, but has no data assimilation. It has been coupled to the HYCOM ocean model for EPAC and CPAC basins but runs uncoupled for the NATL basin. The HMON model runs on-demand with input provided by the National Hurricane Center (NHC) and consists of multiple movable two-way interactive nested grids that follow the projected path of a tropical system.

2017 Atlantic TC Forecast Error (nm)
Mean of Ensemble Member Errors



2017 Atlantic TC Forecast Error (nm)
Ensemble Mean Errors