The Evolution and Use of Objective Forecast Guidance at NHC

James L. Franklin
Branch Chief, Hurricane Specialist Unit
National Hurricane Center

2010 EMC/MMM/DTC Workshop

Hierarchy of TC Track Models

- Statistical/Statistical dynamical
  - CLIPER: Forecasts based on established relationships between storm-specific information (i.e., location and time of year) and the behavior of previous storms
  - NHC72, NHC83, NHC90. Similar to CLIPER but add current and forecast predictors from a parent model.
- Simplified dynamical
  - LBAR: simple two-dimensional dynamical track prediction model that solves the shallow-water equations initialized with vertically averaged (850-200 mb) winds and heights from the GFS global model
  - BAMD, BAMM, BAMS: Forecasts based on simplified dynamic representation of interaction with vortex and prevailing flow (trajectory)
- Dynamical
  - GFDL, GFDN, GFS, NOGAPS, UKMET, ECMWF, NAM, HWRF: solve the physical equations of motion that govern the atmosphere.
- Consensus
  - GUNA, TCON, TVCN, TCCN, TVCC, FSSE: Combinations of other models

Early vs. Late Models

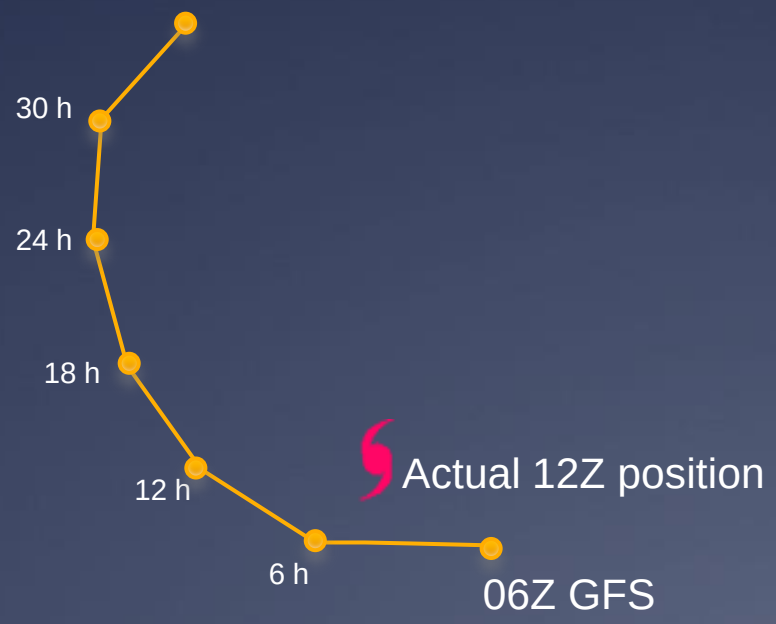
* Forecast cycle begins at synoptic time (e.g., 12Z), and forecast is released at t+3 h (15Z).

* The 12Z runs of the dynamical models GFS, UKMET, GFDL, NOGAPS, ECMWF, etc., are not available until 16Z-19Z, well after forecast is made and released. Therefore, these models are known as "late" models. Forecasts that are available in time for forecast deadlines are called "early" models (LBAR, BAMs, CLIPER).

* For the 12Z forecast cycle, the latest available run of each model is taken (from the 06Z or even 00Z cycle), and adjusted to apply at 12Z. These modified forecasts are known as "interpolated" models: GFSI, UKMI, GFDI, HWFI, NGPI, ECMI, etc.

Early vs. Late Models

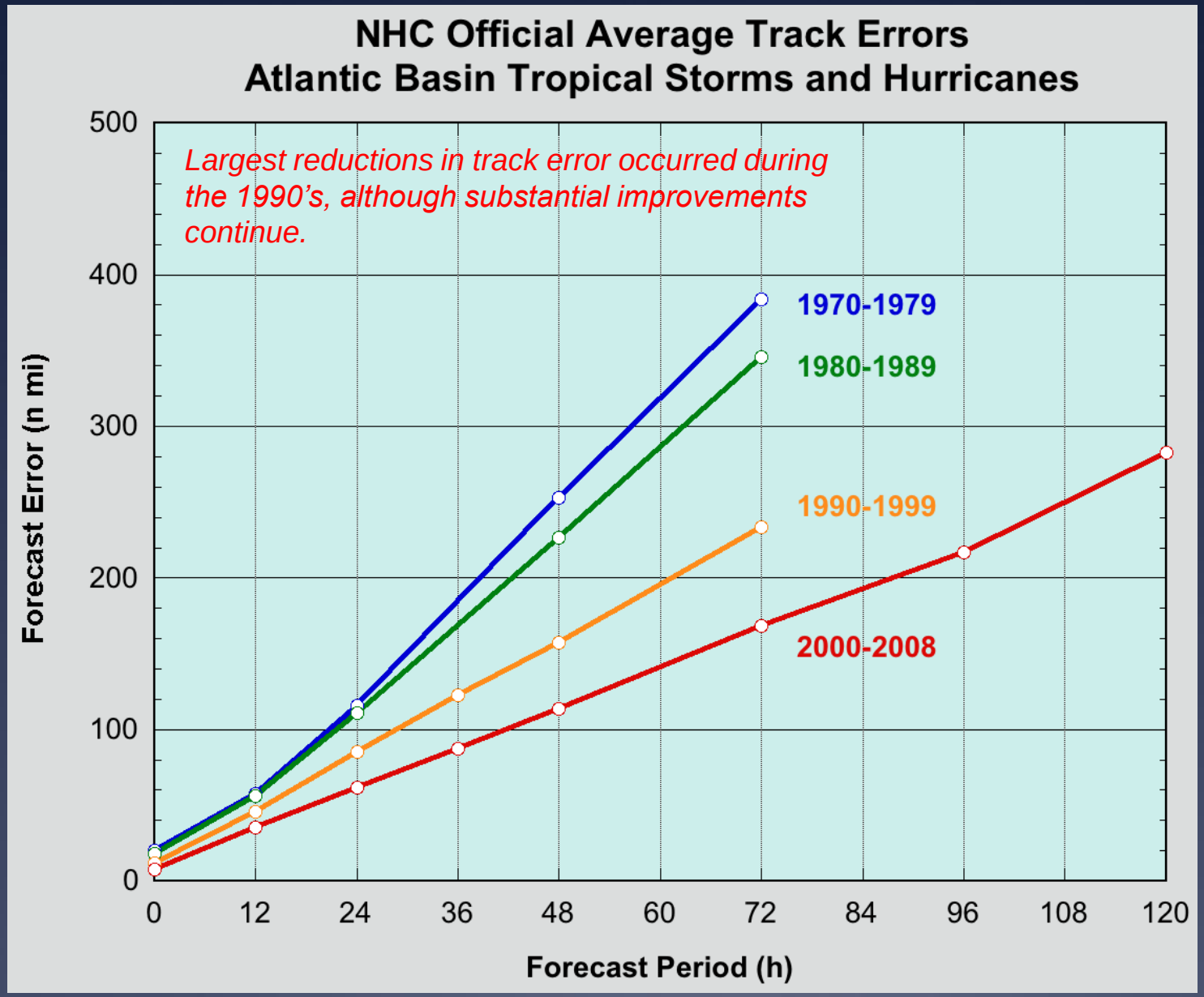
* Interpolated models are created by adjusting the previous model run such that its 6 h forecast position exactly agrees with the current storm position. Then the rest of the forecast is adjusted by the same vector.



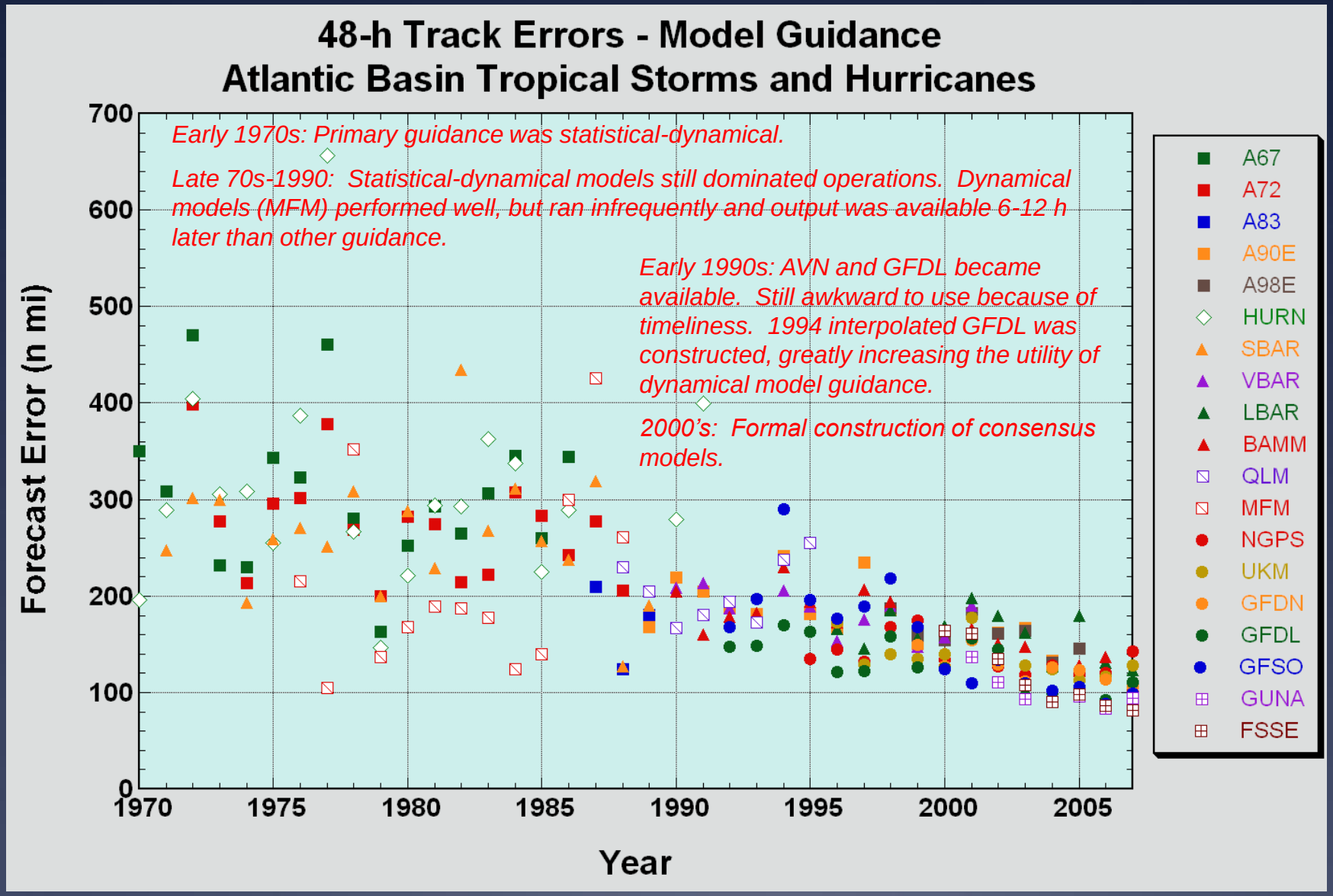
Early vs. Late Models

* Interpolated models are created by adjusting the previous model run such that its 6 h forecast position exactly agrees with the current storm position. Then the rest of the forecast is adjusted by the same vector.

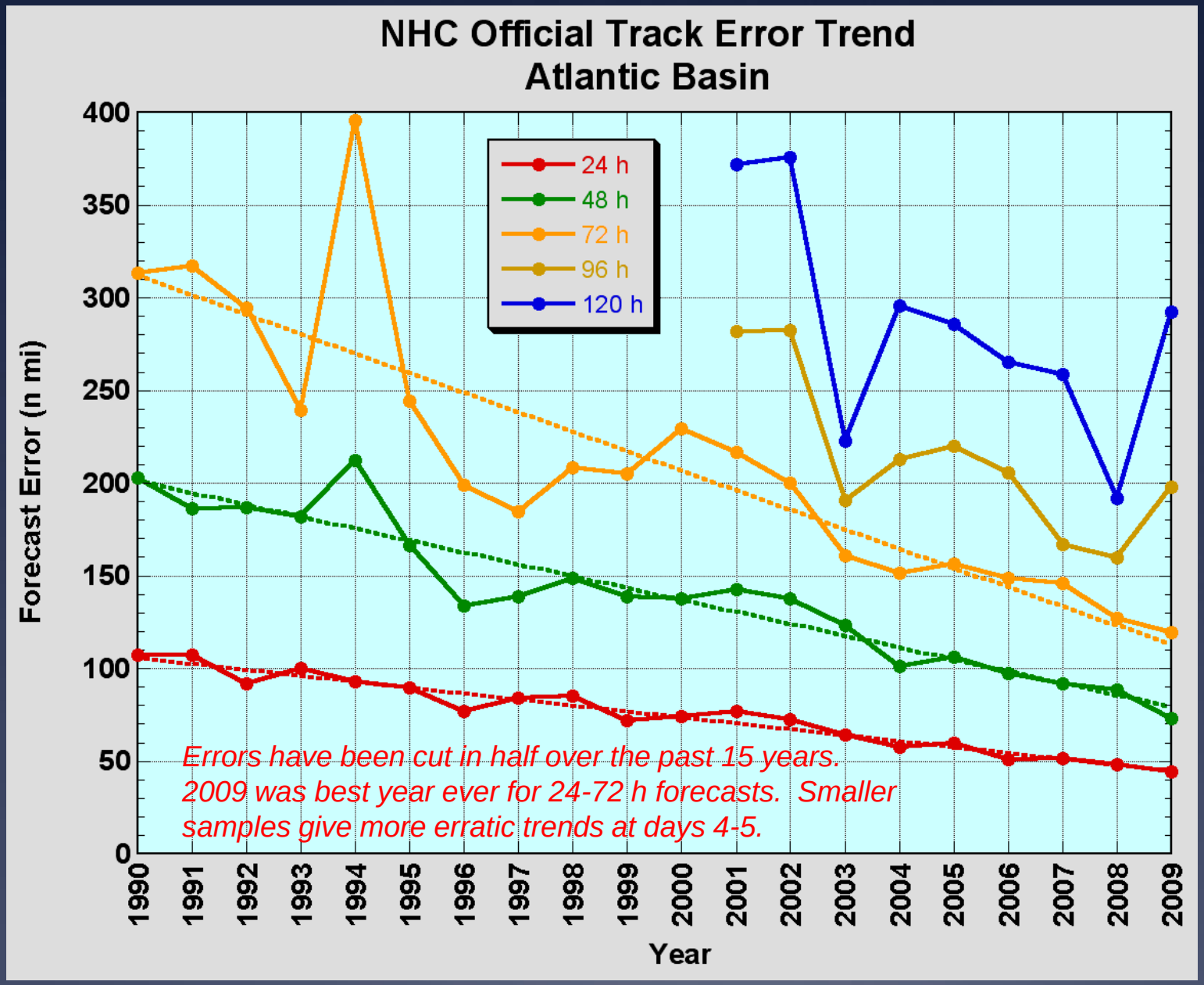
Atlantic Track Error Trends



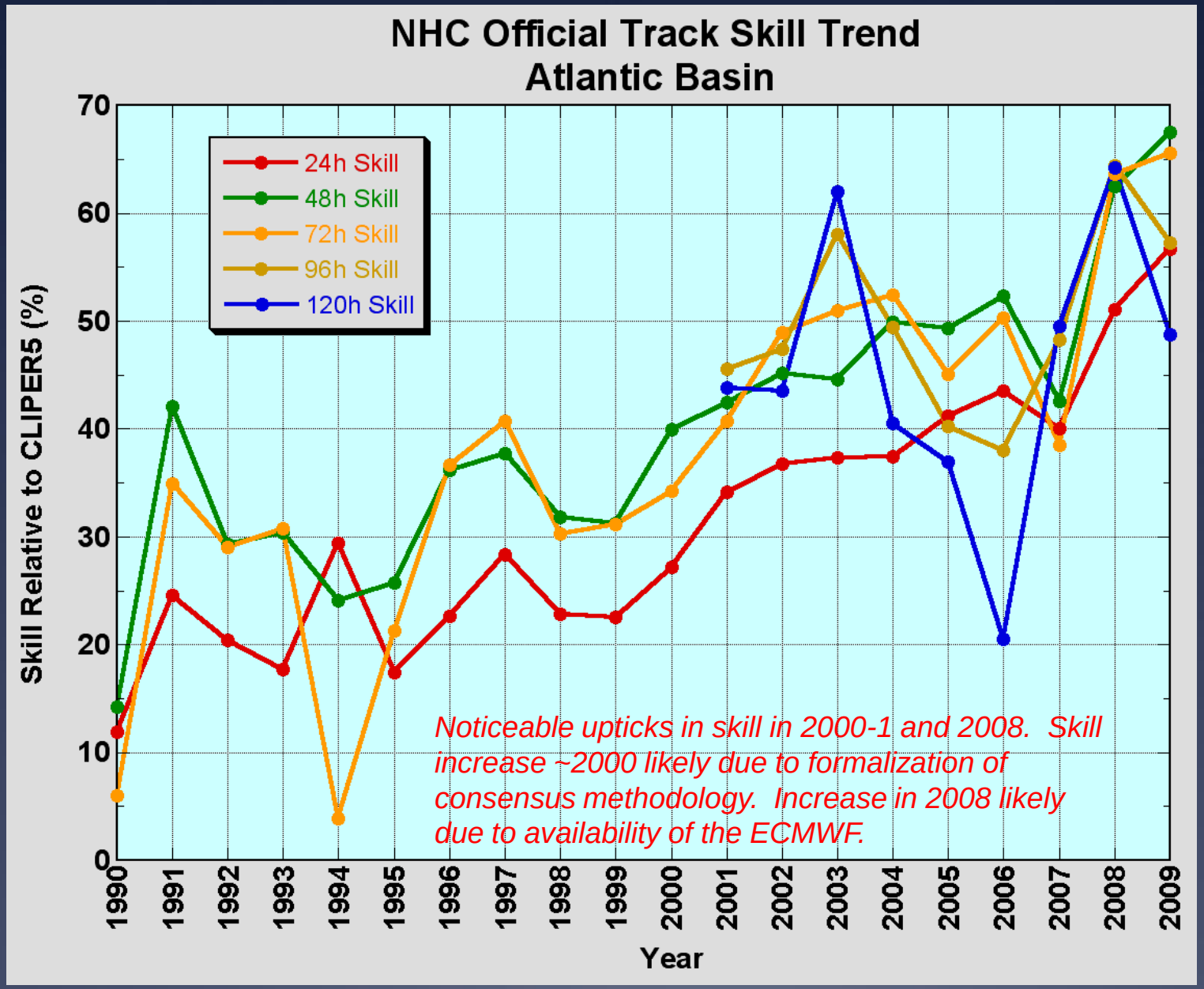
Atlantic Track Model Error Trends
Non-homogeneous Sample



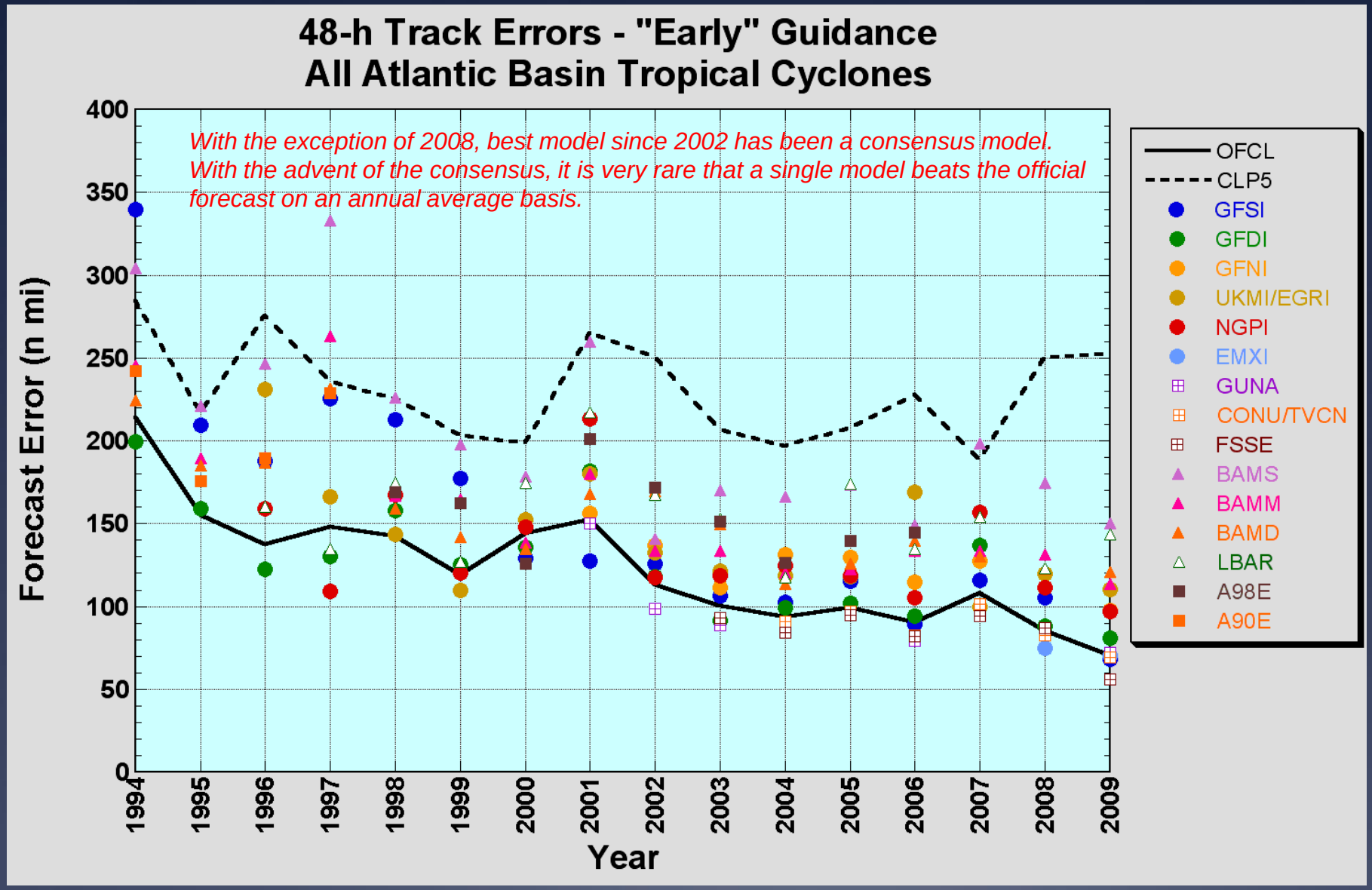
Atlantic Track Error Trends



Atlantic Track Skill Trends

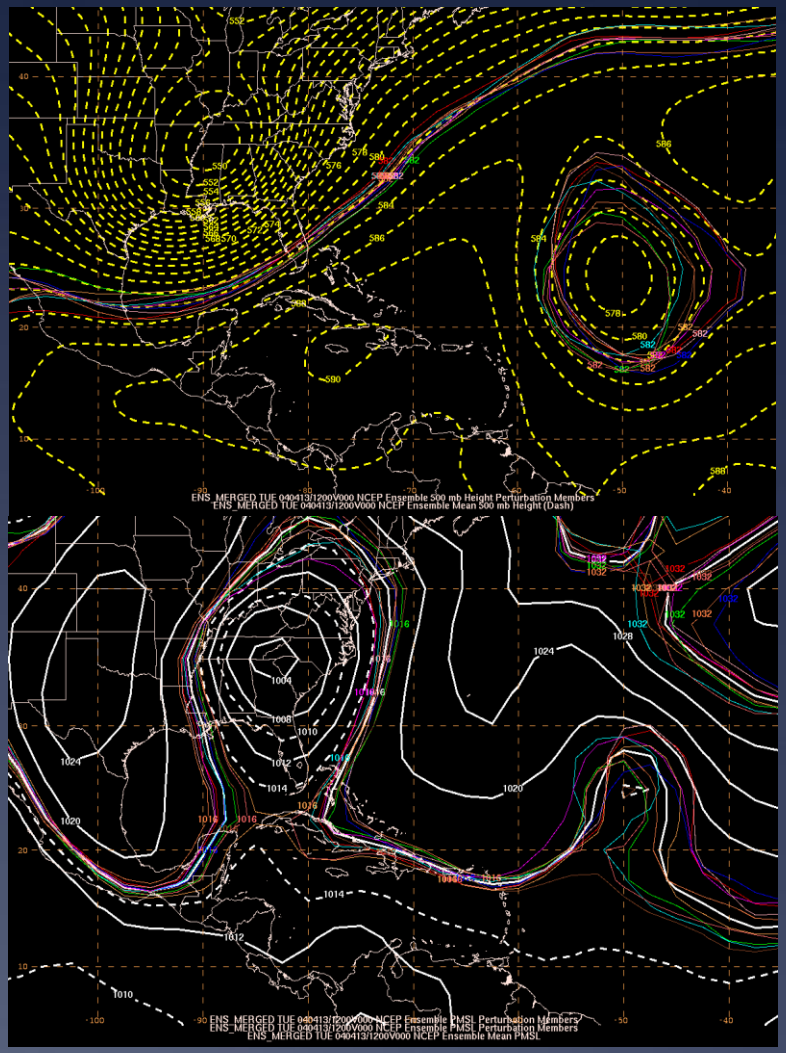


Atlantic Track Model Error Trends
Homogeneous Sample



Ensembles and Consensus

- An ensemble is a collection of forecasts all valid at the same forecast time.
- Can be formed from a single model (e.g., the GFS) by making multiple runs of the model with slightly different (perturbed) initial conditions, or from completely independent models.
- At some forecast time, the average of all the ensemble member's forecasts is the ensemble mean or consensus. The average distance of each member's forecast from the ensemble mean is the ensemble spread.



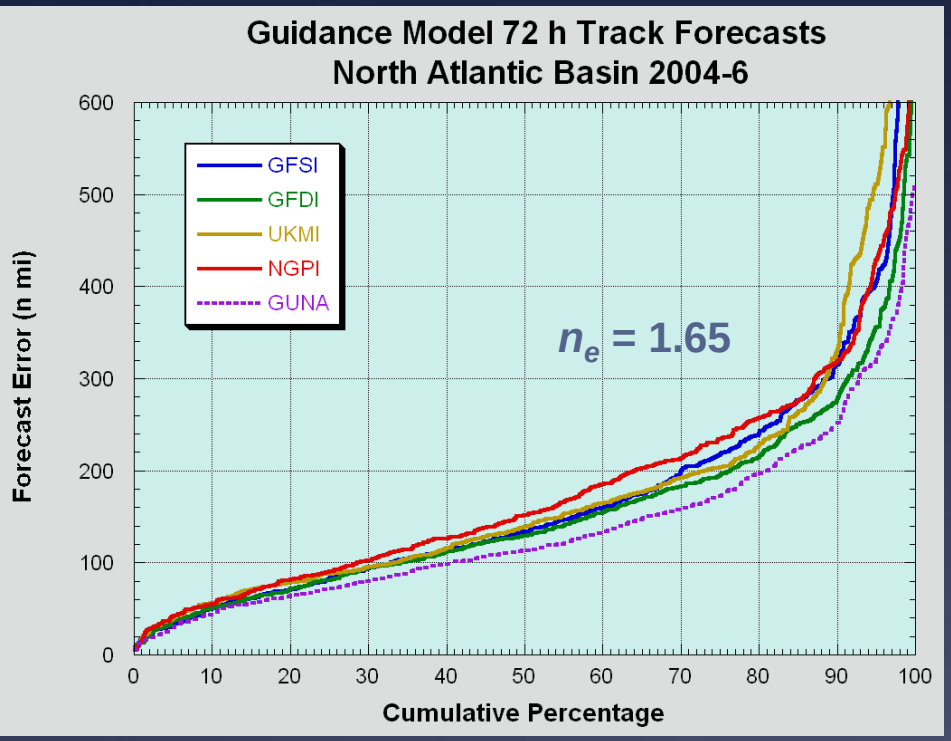
Ensembles and Consensus

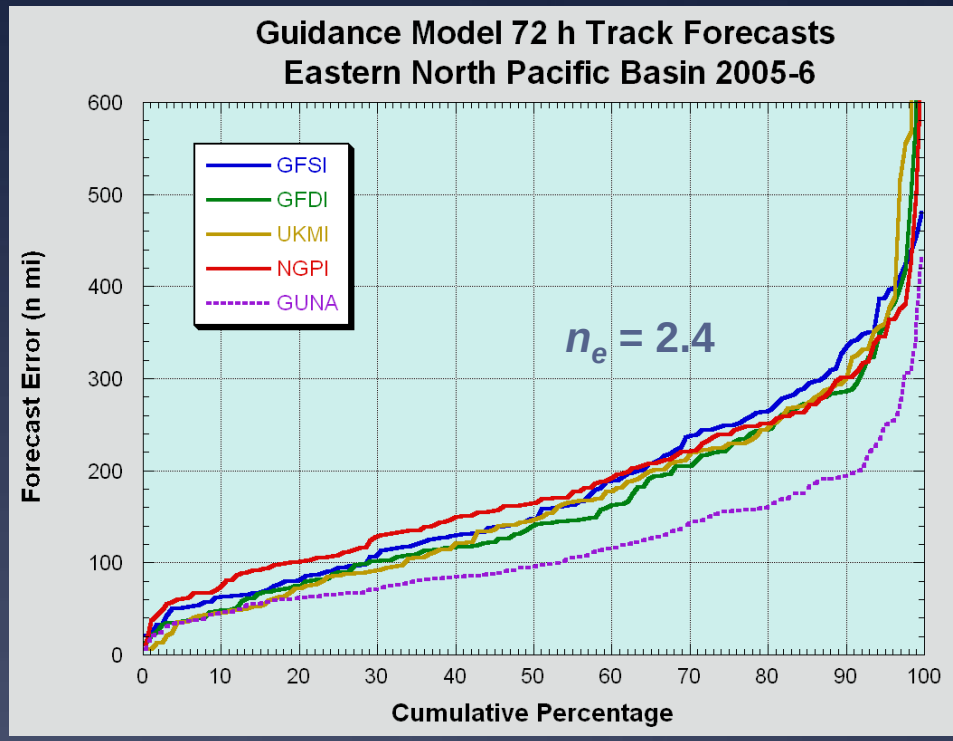
- For tropical cyclones, the most successful consensus models have been those formed from an ensemble of good performing models with a high degree of independence.
- Goerss (2000) shows that if a model's along and cross track errors are normally distributed with zero mean and standard deviation σ, then the model's mean error
  - $ERR_i = \sigma(\pi/2)^{1/2}$
- If all the *N* members of the ensemble were independent and had the same mean error, then the ensemble mean error would be
  - $ERR_{ens} = \sigma(\pi/2N)^{1/2}$

Ensembles and Consensus

- In practice, the member models are not independent, so N is replaced by n_e, the effective degrees of freedom, where $n_e <= N$.
- In this case, the ensemble mean error is
  - $ERR_{ens} = \sigma(\pi/2n_e)^{1/2} = ERR_i / n_e^{1/2}$
- We see that the mean ensemble forecast error is dependent on
  - The mean forecast error of the individual models that make up the ensemble, and
  - The degree of independence of the forecast errors of the individual models.
- Goerss shows that $n_e = (s_i/s_{ens})^2$, where s_i is the average standard deviation of the individual model errors, and s_{ens} is the st dev of the consensus error.

Relative Power of Multi-model Consensus

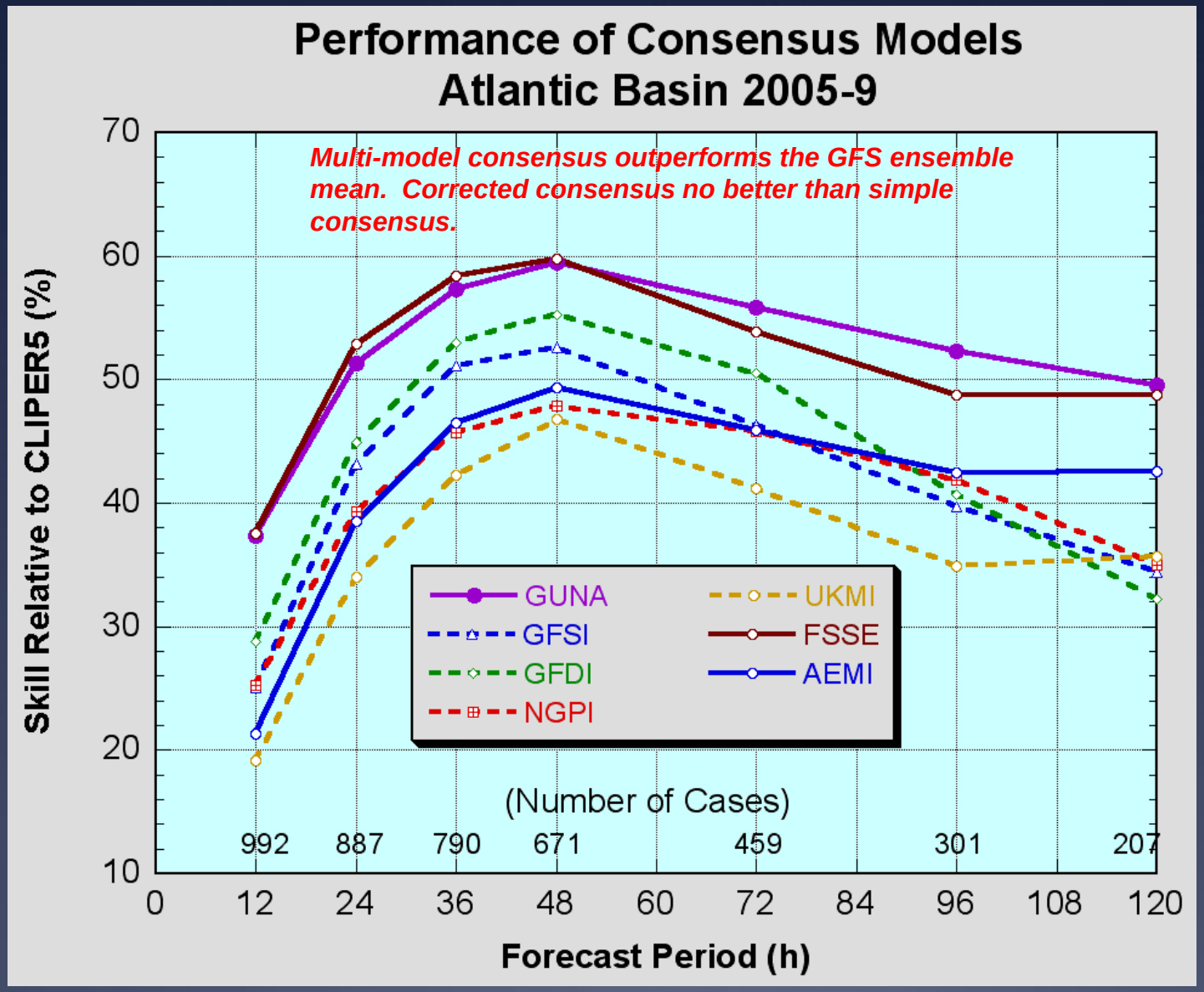




Consensus produces lower mean errors mainly by reducing the number of really large errors, i.e., when any single model "goes bad" the other independent models mitigate the damage to the consensus.

One can speculate that Atlantic basin track forecasts are dominated by relatively strong, well-measured, and varying steering currents that are largely similar from model to model, while in the eastern North Pacific, more uniform steering flows and lack of upstream data tend to make the forecasts more dependent on the individual model initializations and physics.

Consensus


Performance of Consensus Models
Atlantic Basin 2005-9
Multi-model consensus outperforms the GFS ensemble mean. Corrected consensus no better than simple consensus.
Skill Relative to CLIPER5 (%)
70
60
50
40
30
20
10
0
12
24
36
48
60
72
84
96
108
120
Forecast Period (h)
(Number of Cases)
992
887
790
671
459
301
207
GUNA
GFSI
GFDI
NGPI
UKMI
FSSE
AEMI

TC Intensity Models

- **Statistical**
  - **Decay-SHIFOR (Statistical Hurricane Intensity FORecast with Decay over land)**
    - The climatology/persistence analog to CLIPER. Knows NOTHING about the current state of the atmosphere. Tells you what is a typical intensity change for similar storms, based on history.
    - Used as skill benchmark. Is (unfortunately) competitive with the more sophisticated models.

TC Intensity Models

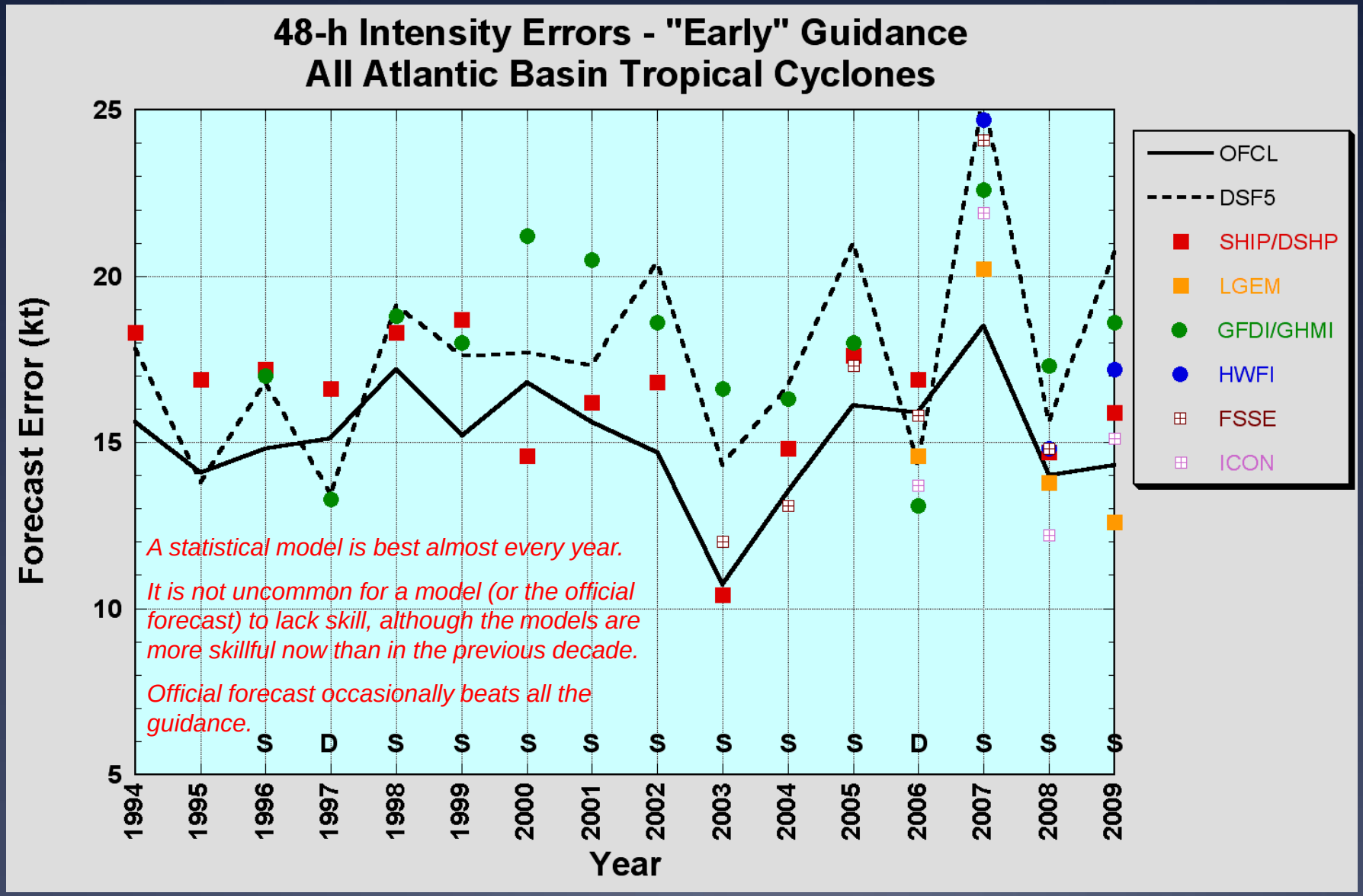
- **Statistical-Dynamical**
  - **Decay-SHIPS (Statistical Hurricane Intensity Prediction Scheme w/Decay)**
    - Similar to Decay-SHIFOR but adds current and forecast environmental factors as predictors (such as sea-surface temperature, environmental wind shear, etc.). Forecast parameters come from the GFS global model.
  - **LGEM (Logistic Growth Equation Model)**
    - Uses similar input to Decay-SHIPS, but the data are used in a more sophisticated way.

- ***Because rapid intensification is relatively rare, and statistical models by design predict "average" behavior, none of the statistical models are capable of forecasting rapid intensity change.***

TC Intensity Models

- **Dynamical (Global)**
  - **GFS, ECMWF, UKMET, NOGAPS**
    - Because the global models have relatively coarse resolution, they cannot represent the hurricane core and are not used for intensity prediction.
- **Dynamical (Regional)**
  - **GFDL, GFDN, HWRF**
    - Have fine-enough resolution to simulate the hurricane core, and are capable of predicting rapid changes in intensity.
    - Unfortunately, they do not do so reliably! Skill is approaching that of the statistical models.
    - One limiting factor is that we do not make the necessary measurements of wind, temperature, and humidity throughout the core and near environment to initialize the models properly.

Atlantic Intensity Model Error Trends
Homogeneous Sample



Intensity Guidance