Introduction to the HWRF-based Ensemble Prediction System

Zhan Zhang
and HWRF Team

2014 Hurricane WRF Tutorial, NCWCP, MD. January 16.

Outline

- Introduction to ensemble prediction system (EPS)
  - What and why ensemble prediction
  - Approaches to ensemble prediction
  - Hurricane ensemble prediction
- HWRF-based EPS
  - Methodology;
  - Verification I : Ensemble vs. Deterministic;
  - Statistical Characteristics of HWRF EPS;
  - Verification II: HWRF EPS vs. Top-flight models;
  - How ensemble perturbation affects storm intensity forecasts;
  - Conclusion and Future Work.

What is an Ensemble Forecast ?

An ensemble forecast is simply a collection of two or more forecasts verifying at the same time. Ensemble forecast aims to estimate the probability density function of forecast states

Why do we need ensemble forecast ?

Uncertainties, or weak noises, acting upon a numerical weather prediction (NWP) model system can have far-reaching consequences due to its chaotic and nonlinear nature (Lorenz, 1963, 1965).

What are the main source of uncertainties ?

- *IC/BC uncertainties*: observational errors, poor data coverage, and errors in DA system*;*
- *Model uncertainties:* mis-representation of model dynamics/physics, impact of sub-grid scale features.

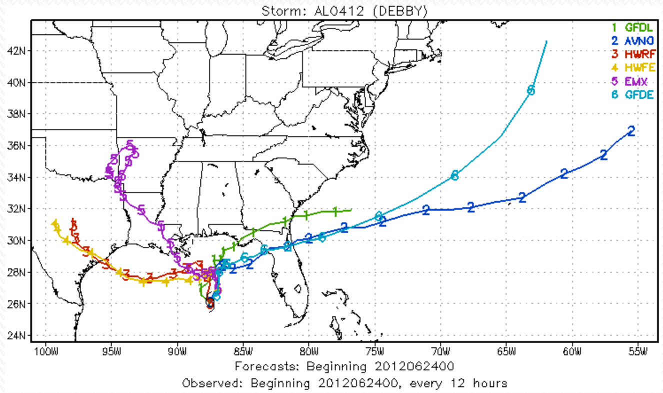
These uncertainties are inevitable. In a chaotic system like an atmospheric model, non linear errors will grow - sometimes rapidly. Eventually these growing errors cause the model forecast output to eventually become useless.

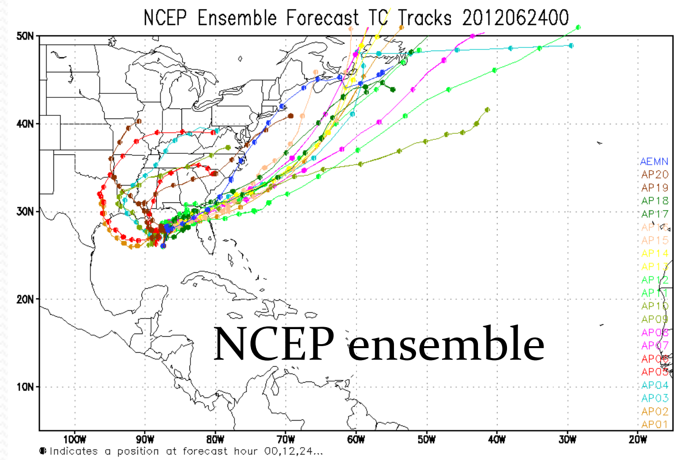
Track Prediction for Hurricane Debby, 20120624 00Z

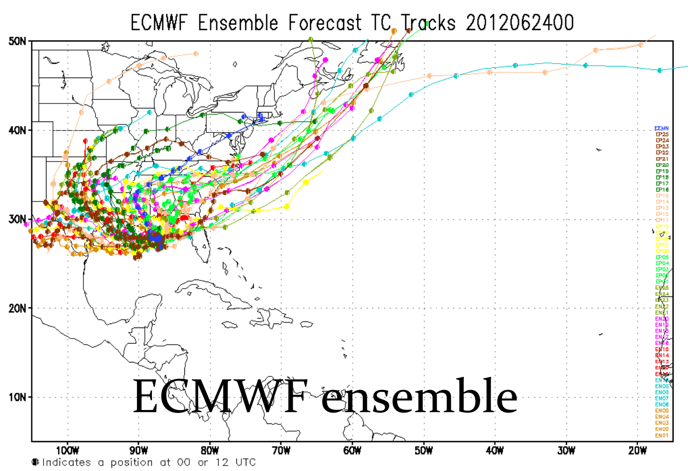
Large differences in predicted storm tracks due to:

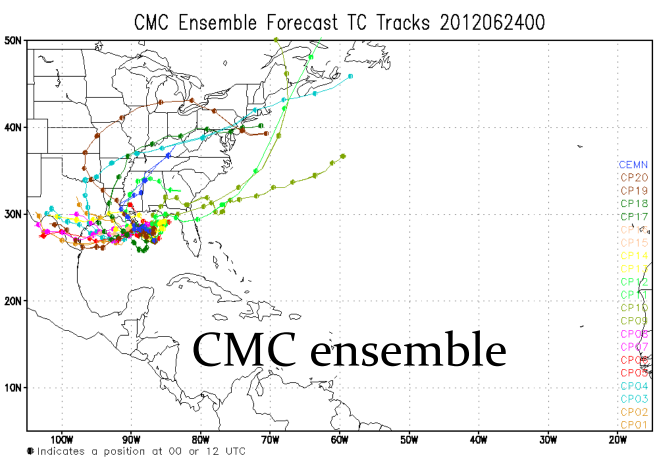
1. multi-model dynamics;
2. multi-physics;
3. multi-initial analysis.

Multi-model





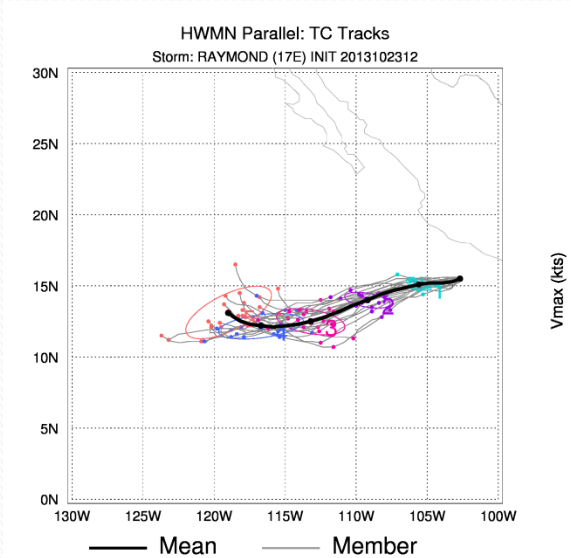


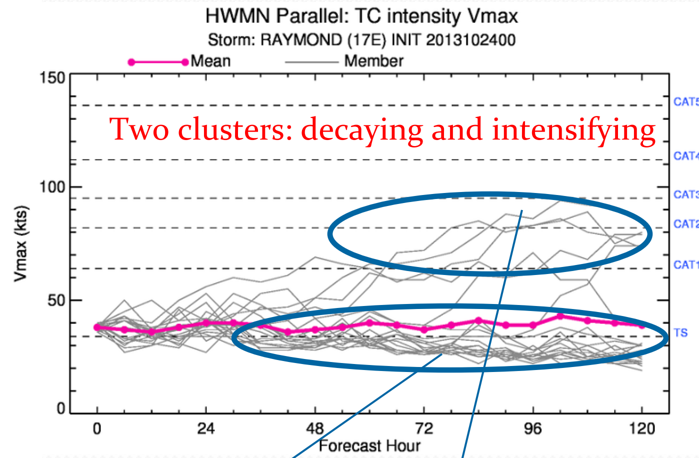


Prediction for Hurricane Raymond, 20131024 00Z

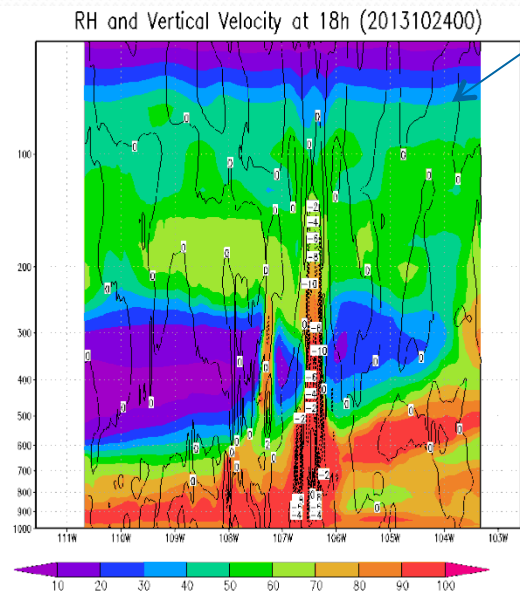
Large differences in predicted storm intensity due to sub-grid uncertainties in model physics: stochastically perturbed cumulus convection scheme in HWRF

Dry air at mid-level suppressed storm development in one member, while active convective cells overcome the dry air, storm intensified in anther member.

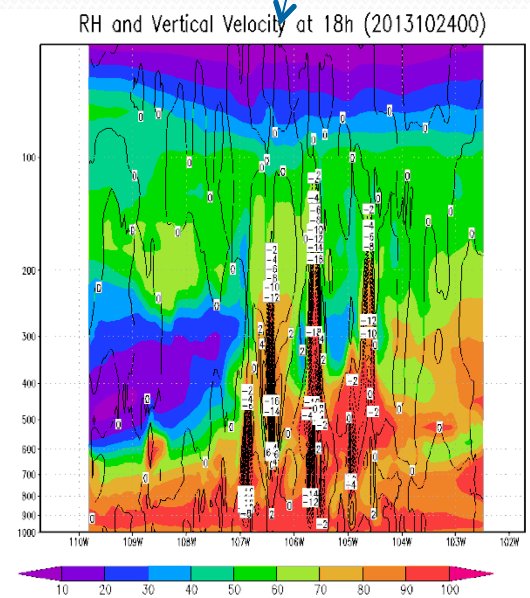




HW06



HW10



Approaches to Ensemble Prediction

- Monte Carlo Approach ---- not practically possible

  sample all sources of forecast error, perturb any input variable and any model parameter that is not perfectly known. Take into consideration as many sources as possible of forecast error.

- Reduced Sampling ----- limited resource

  Sample leading sources of forecast error. Rank error sources, prioritize, optimize sampling: growing components will dominate forecast error growth, important model physics, etc..

- Existing Methods

  - Initial uncertainties: SV-based ensemble (ECMWF), EnKF-based ensemble (MCS), BV-based ensemble (NCEP), ETKF-based ensemble (UKMet), ETR-based ensemble (NCEP), EOF-based ensemble (hurricane).
  - Model uncertainties: Multi-model ensemble; Single model with multi-physics.

- Desired Ensemble Perturbations

  Growing modes, Orthogonality, No bias, Mimic analysis errors and/or model errors.

HWRF-based Ensemble Prediction System

Considerations for Hurricane EPS:

1. Uncertainties in initial storm position, intensity, and structure;
2. Uncertainties in large scale flows (ICs/BCs);
3. Multi-scale interactions among sub-grid scales, (~0-100m), convective clouds (~100-1000m), and the large-scale environment (~100-1000km)

Background and Motivation

- Convective Trigger function in Current HWRF Cumulus Parameterization Scheme (SAS: Simplified Arakawa-Schubert)

  P_{CSL}-P_{LFC} <= DP(w)   Convection is triggered,

  P_{CSL}-P_{LFC} > DP(w)   No sub-grid convection

  P_{CSL}: Parcel pressure at Convection Starting Level,

  P_{LFC}: Parcel pressure at Level of Free Convection

  DP(w): Convective Trigger, which is function of large scale vertical velocity w.

  DP(w) is arbitrarily confined between 120hPa-180hPa

- Storm intensity (Max Wind Speed) is found very sensitive to the convective trigger function;
- Necessary to introduce fuzzy logic trigger to represent sub-grid features.

Methodology

- IC/BC Perturbations (Large scale):

20 member GEFS (ETR-based).

- Model Physics Perturbations (Sub-grid scale):

Stochastic Convective Trigger

$P_{CSL} - P_{LFC} <= DP(w) + R_r\ (n)$

R_r is white noise, ranging from -50hPa to +50hPa, n is nth ensemble member, used as random seed. No spatial and temporal correlations

HWRF EPS vs. its Deterministic versions

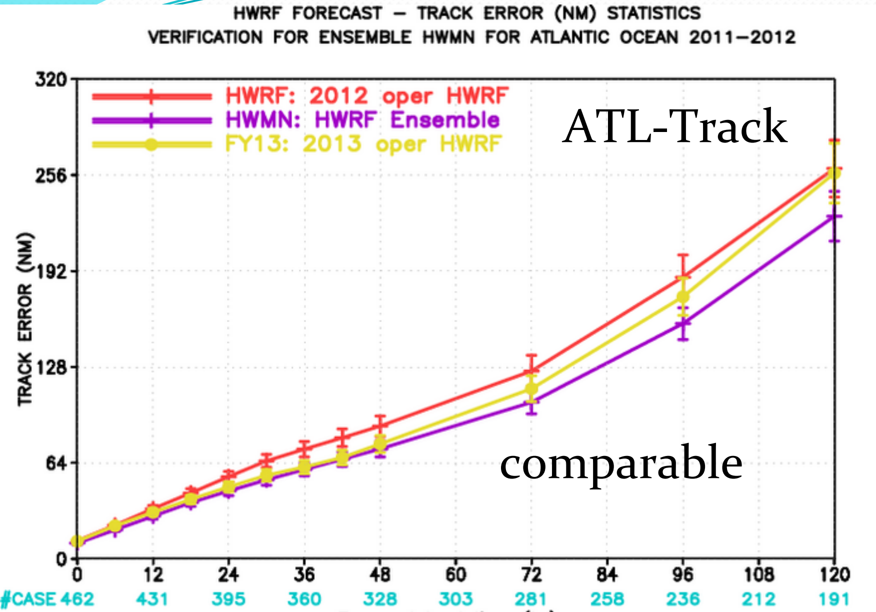
HW01 – HW20:
Individual Ensemble Members Generated from Stochastic Convective Trigger + GEFS IC/BC;

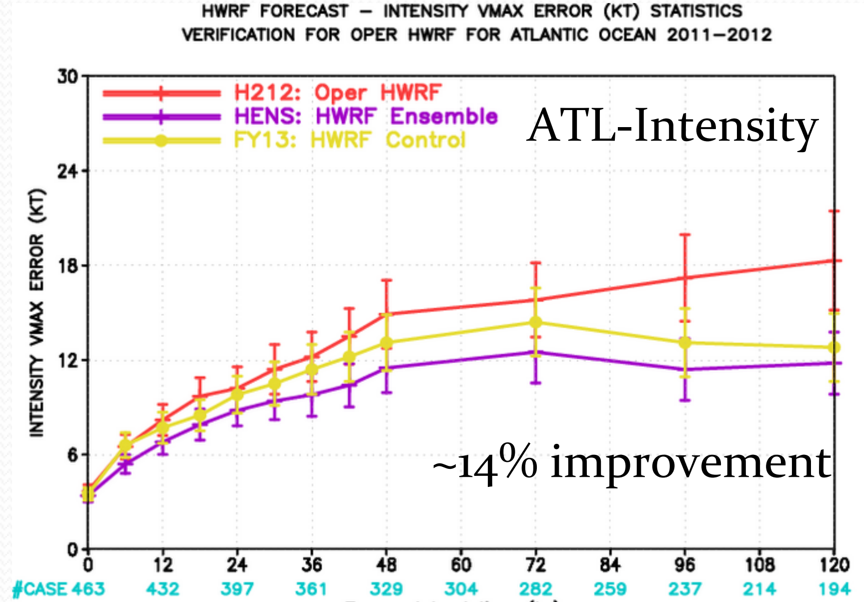
HWMN:
Mean of HW01-HW20 ensembles;

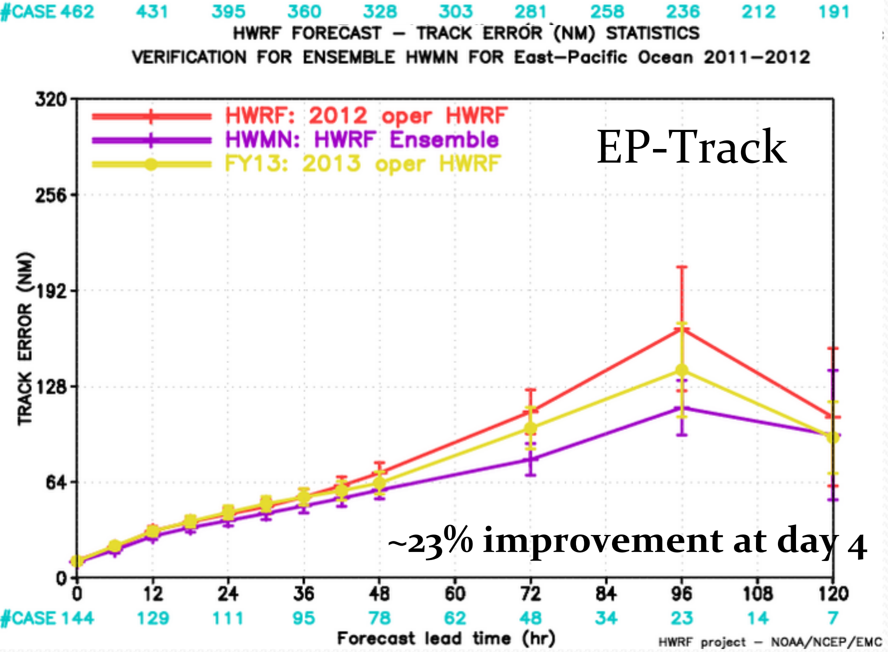
FY13:
2013 Operational HWRF, Deterministic version of HWRF;

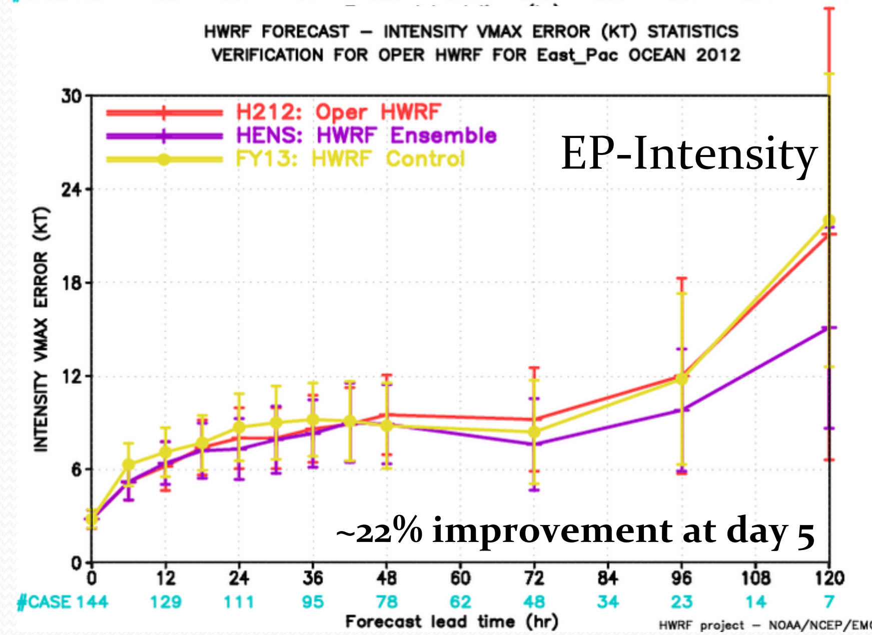
H212:
2012 Operational HWRF;

2012: all storms during Aug. 1 to Oct. 31; 463 cycles
2011: 09L, 12L, 14L and 16L; 144 cycle
Total: 607 cycles.

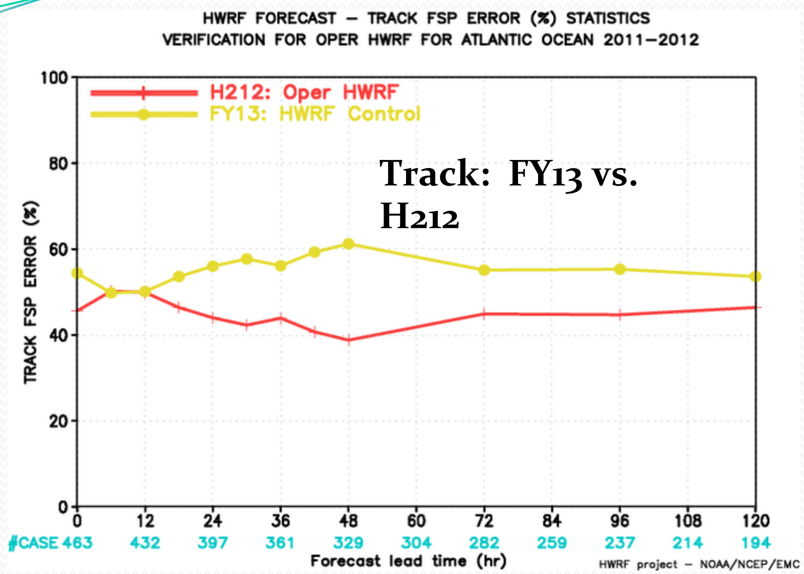

HWRF FORECAST – TRACK ERROR (NM) STATISTICS
VERIFICATION FOR ENSEMBLE HWMN FOR ATLANTIC OCEAN 2011–2012
HWRF: 2012 oper HWRF
HWMN: HWRF Ensemble
FY13: 2013 oper HWRF
ATL-Track
comparable
TRACK ERROR (NM)
#CASE 462 431 395 360 328 303 281 258 236 212 191

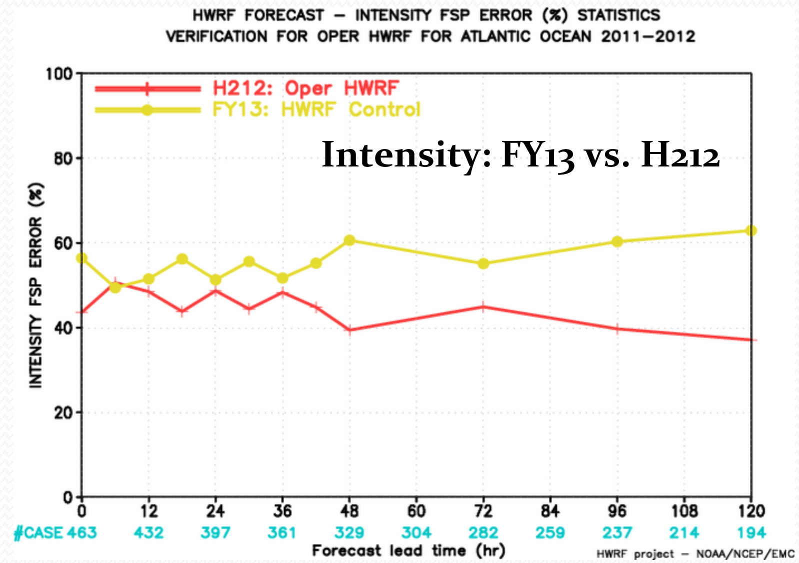

HWRF FORECAST – INTENSITY VMAX ERROR (KT) STATISTICS
VERIFICATION FOR OPER HWRF FOR ATLANTIC OCEAN 2011–2012
H212: Oper HWRF
HENS: HWRF Ensemble
FY13: HWRF Control
ATL-Intensity
~14% improvement
INTENSITY VMAX ERROR (KT)
#CASE 463 432 397 361 329 304 282 259 237 214 194

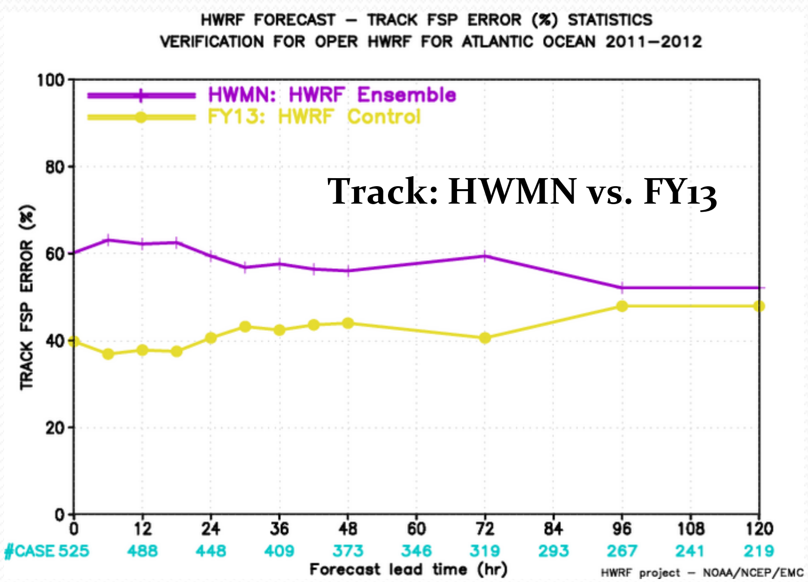

HWRF FORECAST – TRACK ERROR (NM) STATISTICS
VERIFICATION FOR ENSEMBLE HWMN FOR East–Pacific Ocean 2011–2012
HWRF: 2012 oper HWRF
HWMN: HWRF Ensemble
FY13: 2013 oper HWRF
EP-Track
~23% improvement at day 4
TRACK ERROR (NM)
#CASE 144 129 111 95 78 62 48 34 23 14 7
Forecast lead time (hr)
HWRF project – NOAA/NCEP/EMC

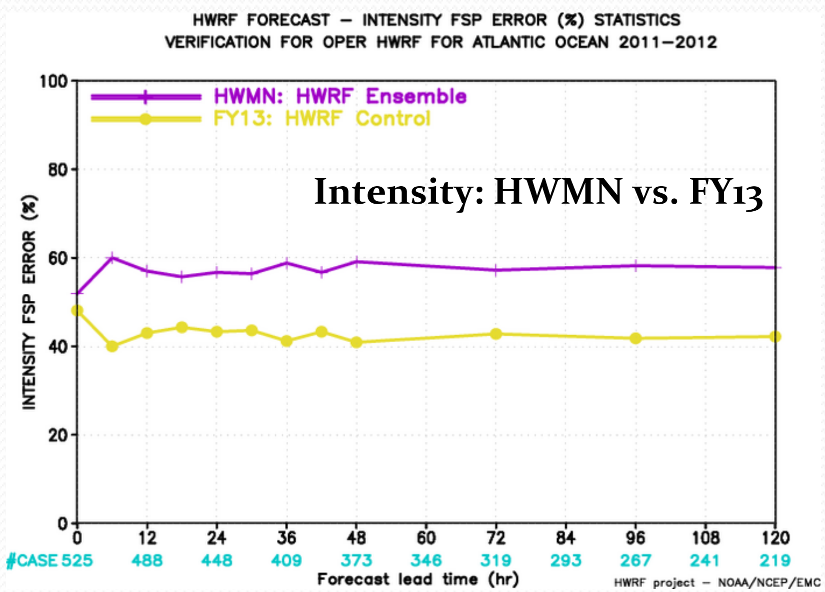

HWRF FORECAST – INTENSITY VMAX ERROR (KT) STATISTICS
VERIFICATION FOR OPER HWRF FOR East_Pac OCEAN 2012
H212: Oper HWRF
HENS: HWRF Ensemble
FY13: HWRF Control
EP-Intensity
~22% improvement at day 5
INTENSITY VMAX ERROR (KT)
#CASE 144 129 111 95 78 62 48 34 23 14 7
Forecast lead time (hr)
HWRF project – NOAA/NCEP/EMC

Frequency Superior Performance Verifications (Atlantic Basin)

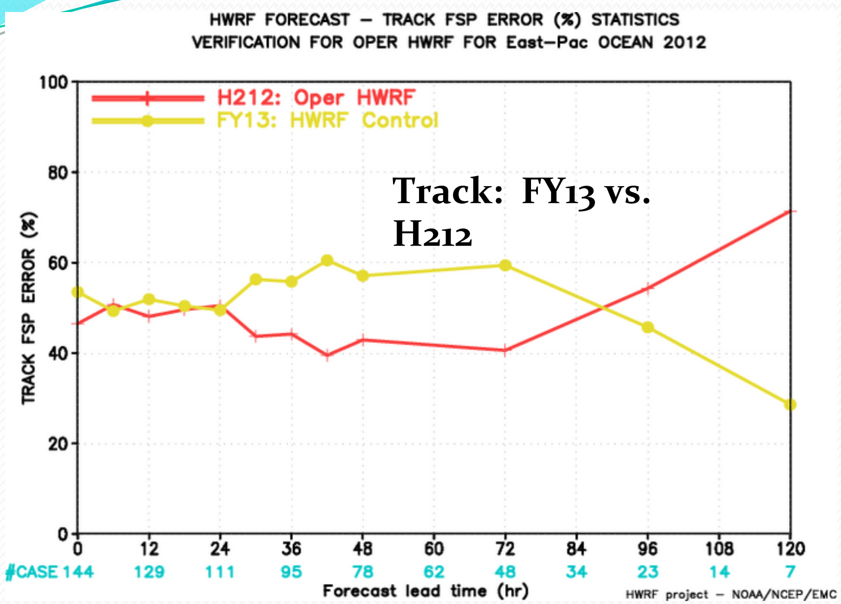


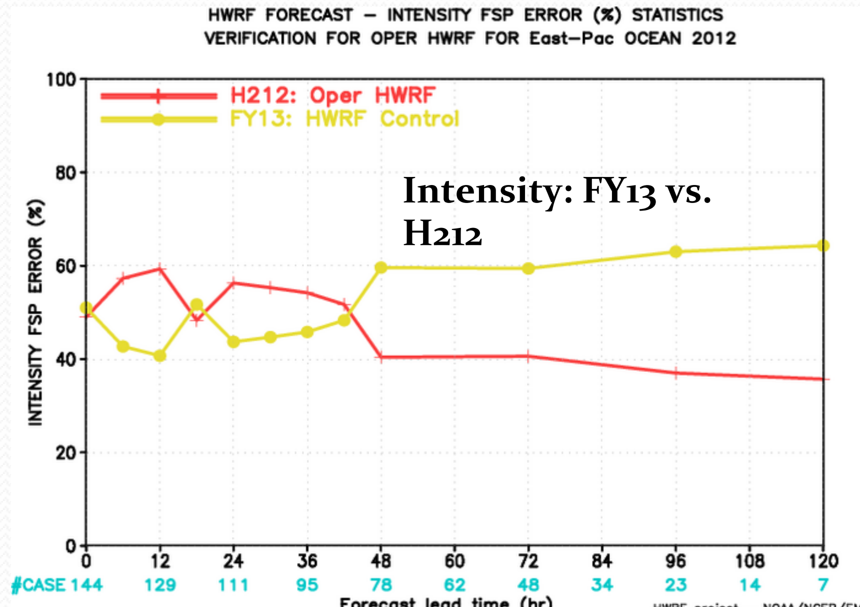


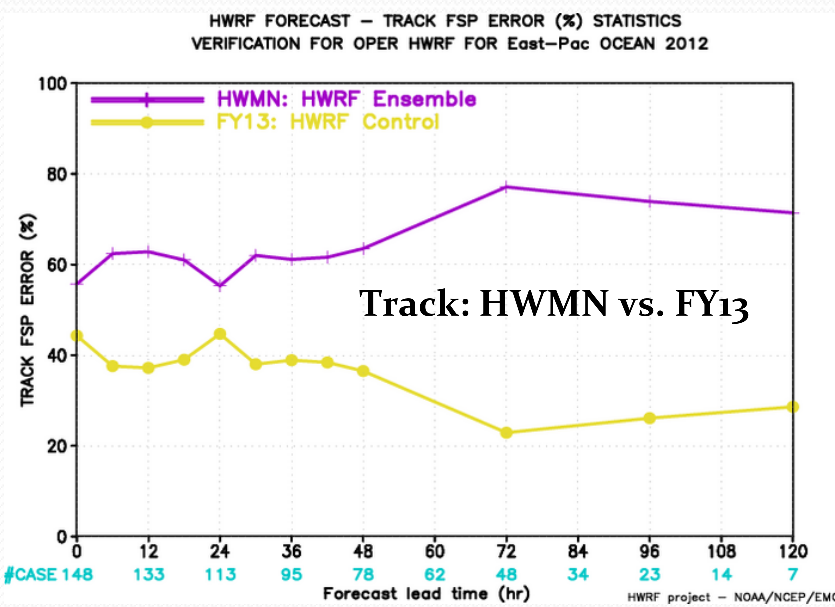


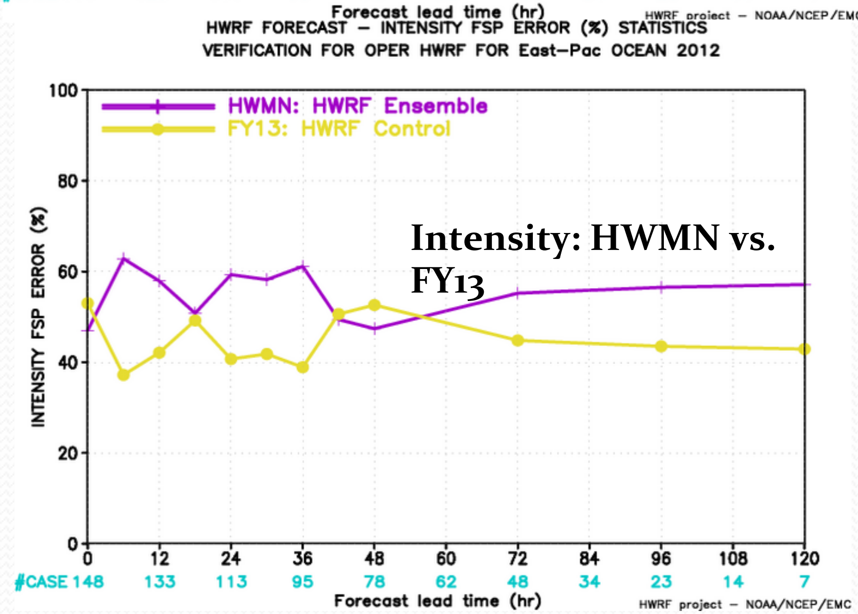


Frequency Superior Performance Verifications (E-Pac Basin)



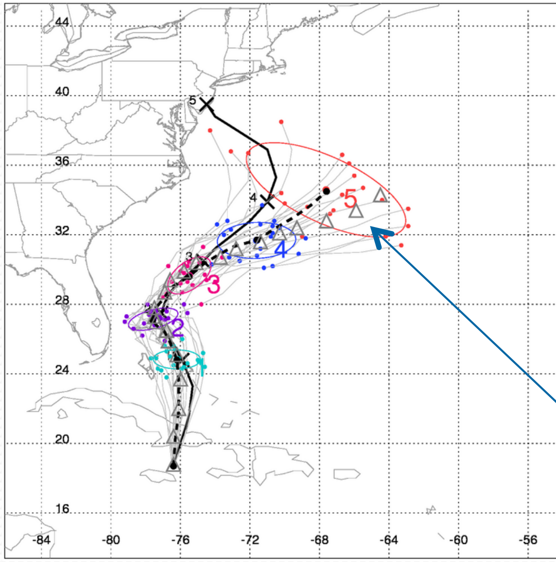






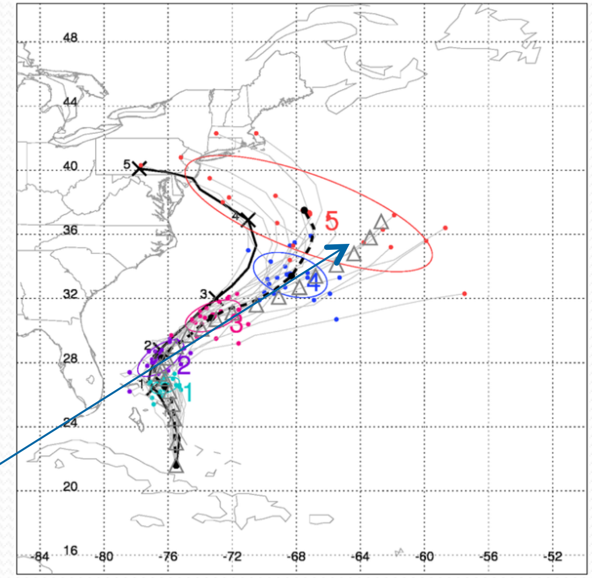
Track Probability Forecasts for Hurricane Sandy

SANDY18L.2012102500



Few members turned west

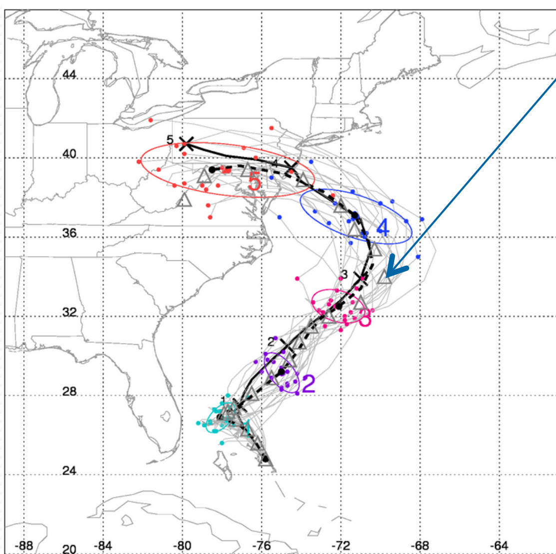
SANDY18L.2012102512



More members turned west

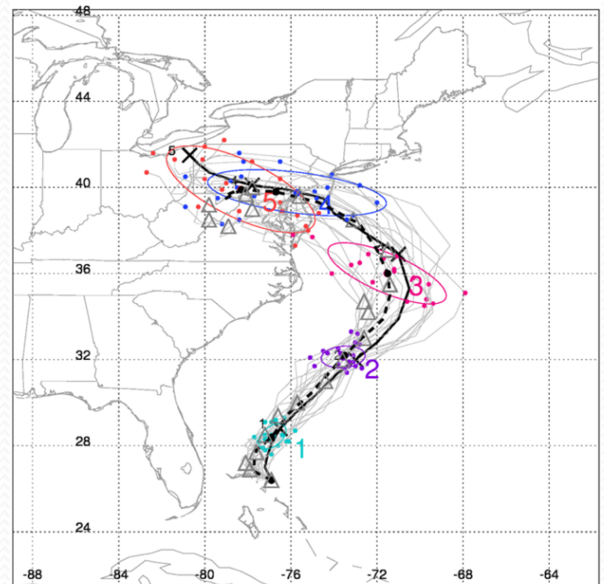
FY13

SANDY18L.2012102600

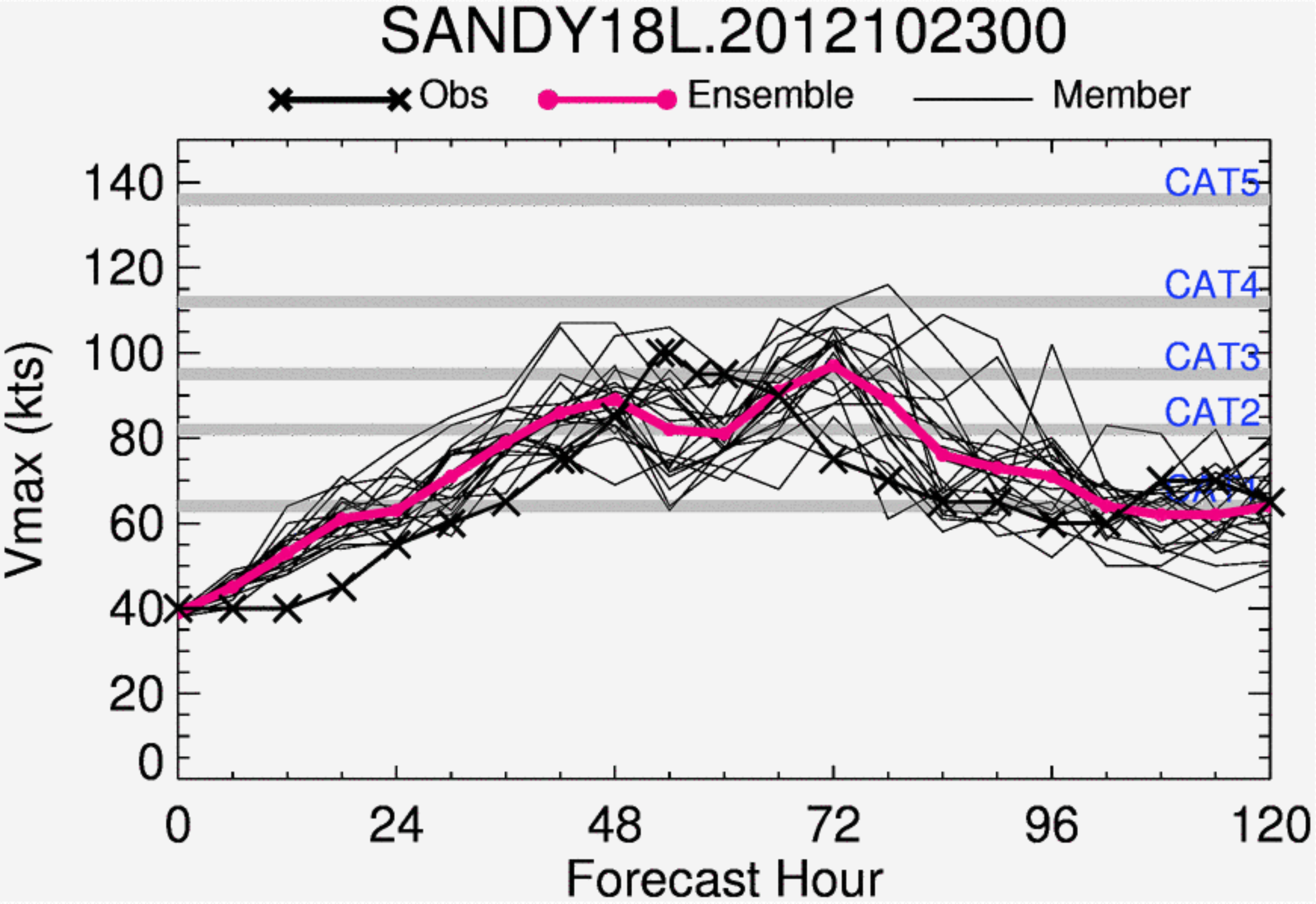


All members turned west

SANDY18L.2012102612

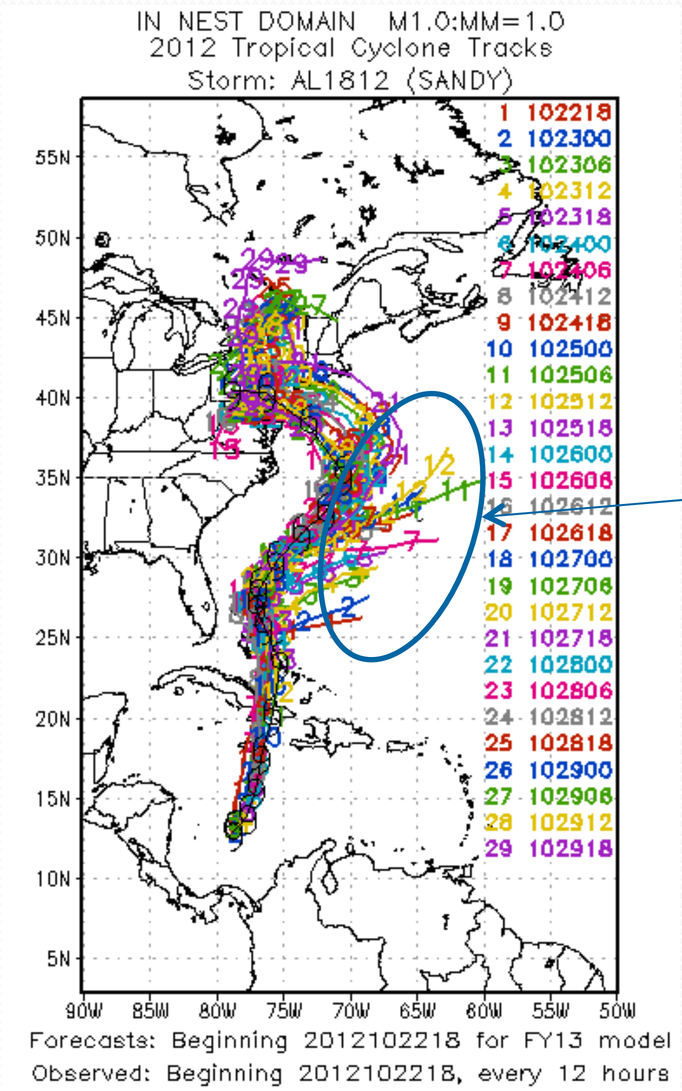


Ensemble Intensity Forecasts for Hurricane Sandy

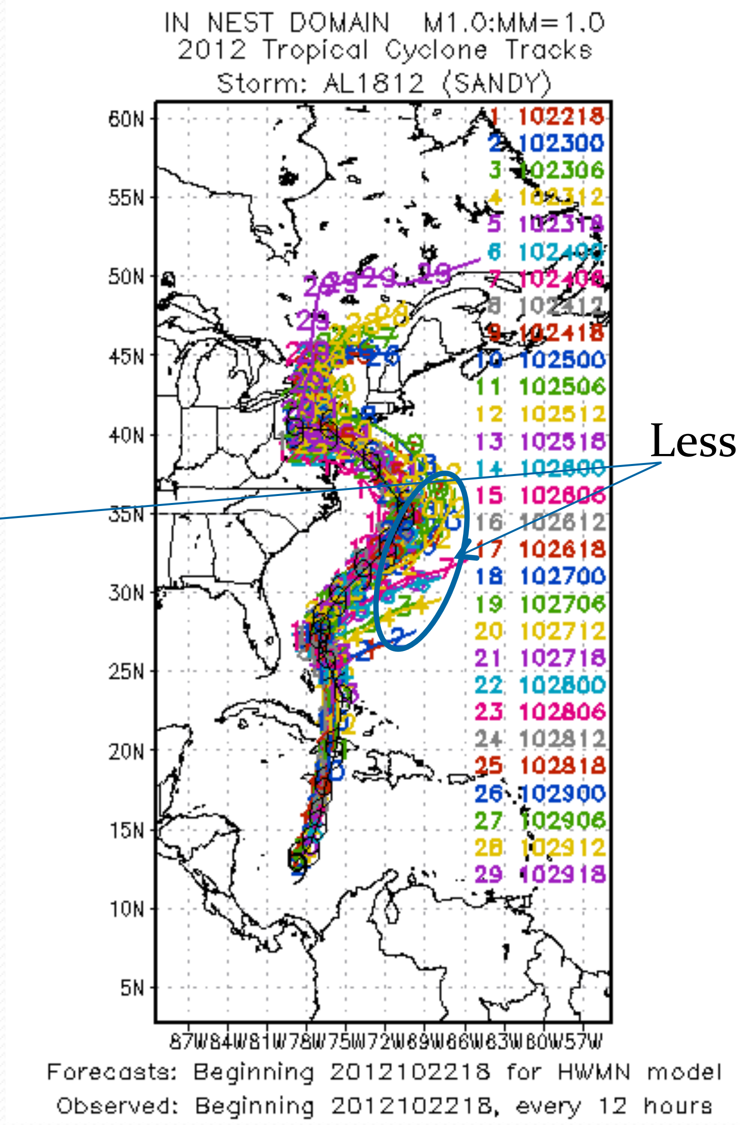


Spaghetti Plots Comparison for Track Forecasts Sandy, 2012

FY13



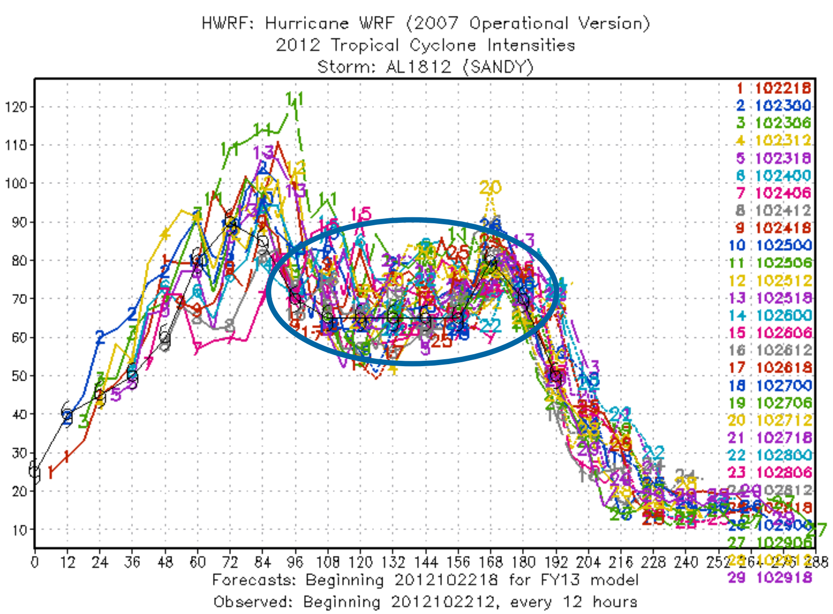
HWMN



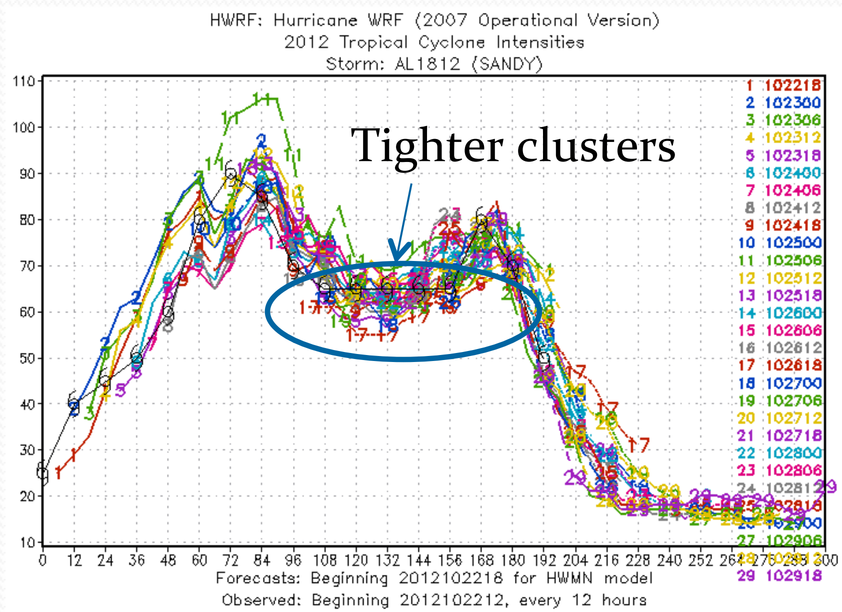
Less outliers

Spaghetti Plot Comparison for Intensity Forecasts Sandy, 2012

FY13



HWMN

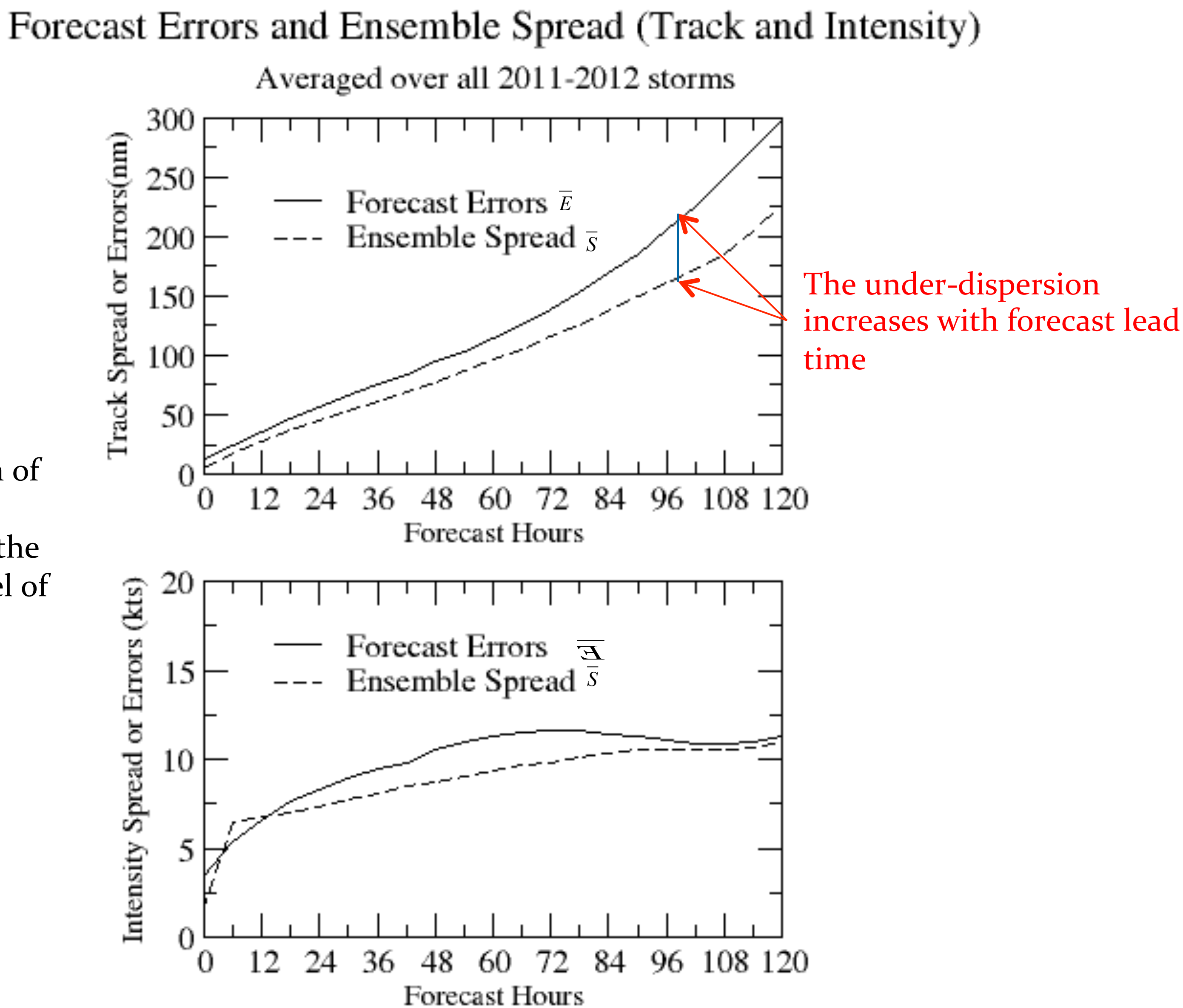


Statistical Characteristics of HWRF EPS

To ensure no bias introduced, and adequate ensemble spread.

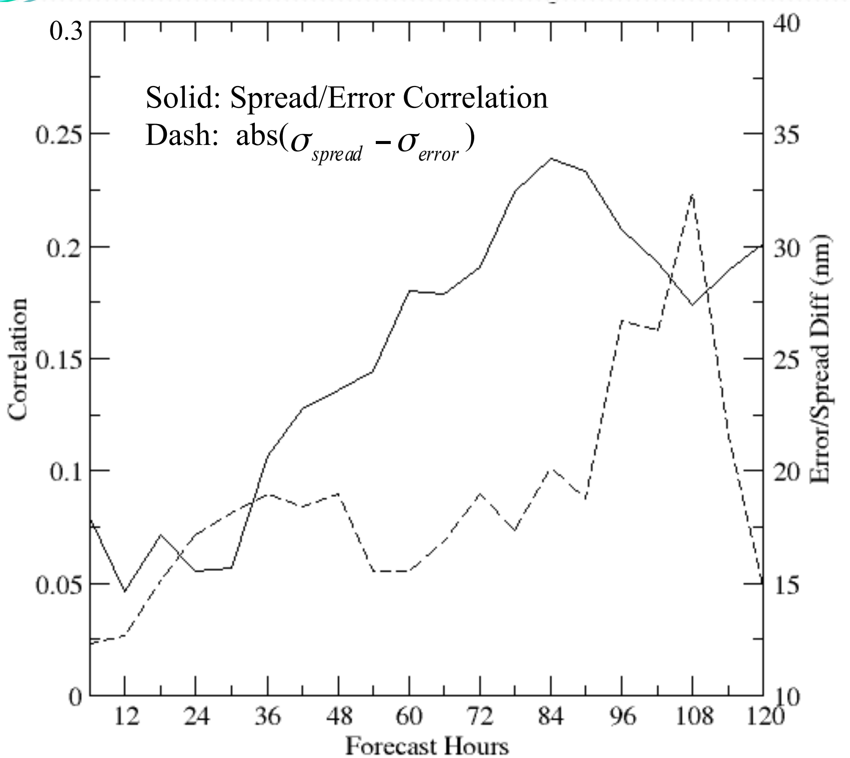
- Ensemble spread and forecast error relationship:
  - Assume the ensemble spread and forecast error are two random variables: S(t)=N($\bar{S}, \sigma_{spread}$), E(t)=N($\bar{E}, \sigma_{error}$). Evaluate "climatological" mean and standard Deviations;
  - Spread/error correlation: researches have shown that the spread/error correlation is generally not as high as expected, less than 0.5 (Baker 1991, Houtekamer 1993, Whitaker and Loughe 1997)
- Spread evaluation: rank histogram for ensemble forecasts

1. The mean of ensemble spread is close to the mean of the forecast errors;
2. The difference between the two lines indicates the level of ensemble dispersion;

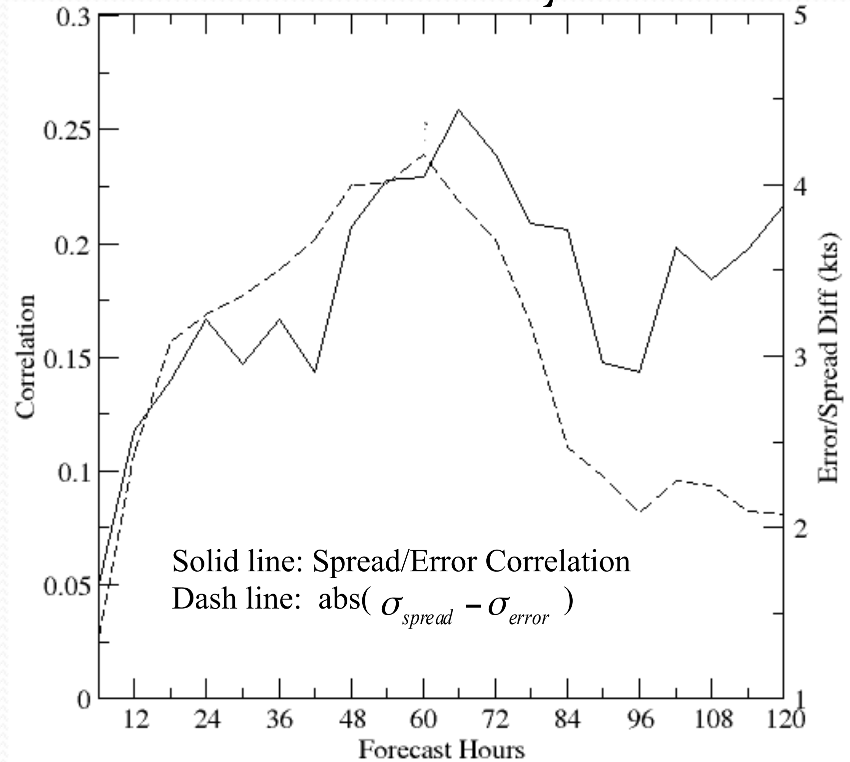


Relationship between Ensemble Spread and Forecast Error

Track



Intensity

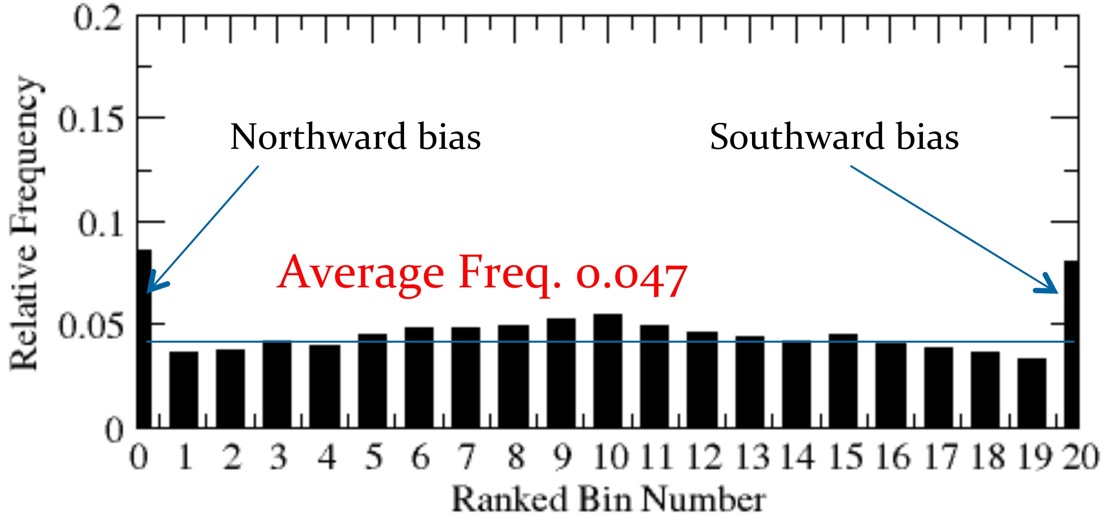


* σ_{error} : standard deviation of ensemble spread, σ_{spread} : standard deviation of forecast error;

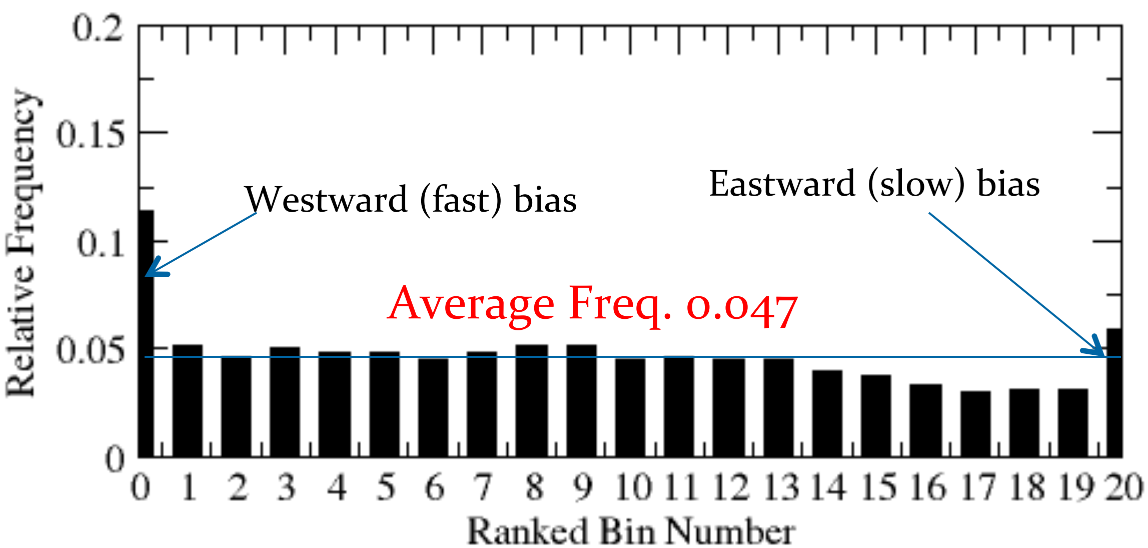
1. Spread/error correlations reach to the peak at day 3 for both track and intensity;
2. The difference $\sigma_{spread} - \sigma_{error}$ measures the closeness of the 'climatological' values of ensemble spread and forecast errors; the ensemble spread has less predictive value (of the forecast skill) when $| \sigma_{spread} - \sigma_{error} |$ is small;
3. In HWRF EPS, intensity spread is a good indicator of forecast skill up to 72h, track spread is not.

Analysis of Rank Histogram for Track Forecast

Latitude



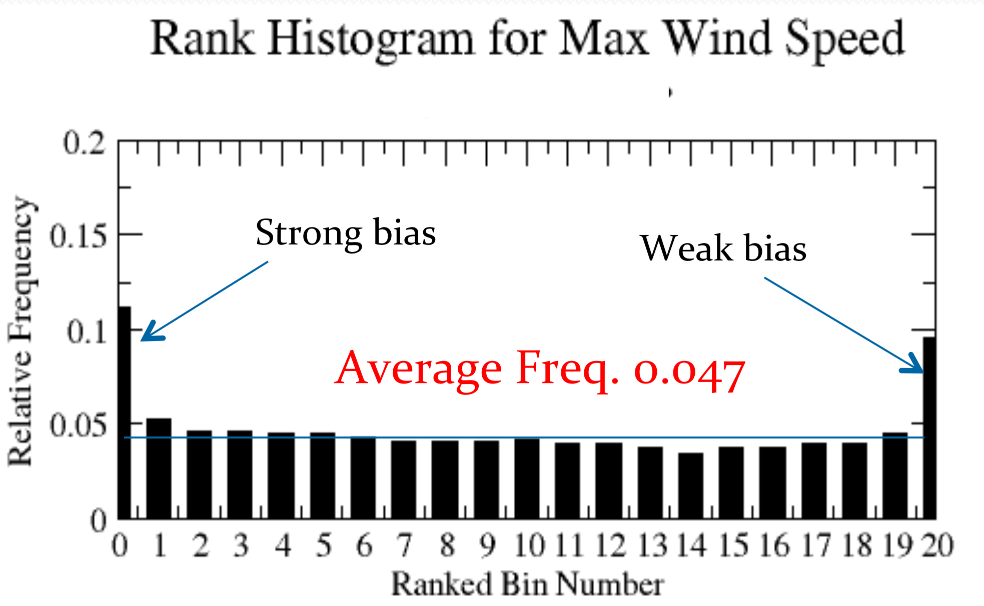
Longitude

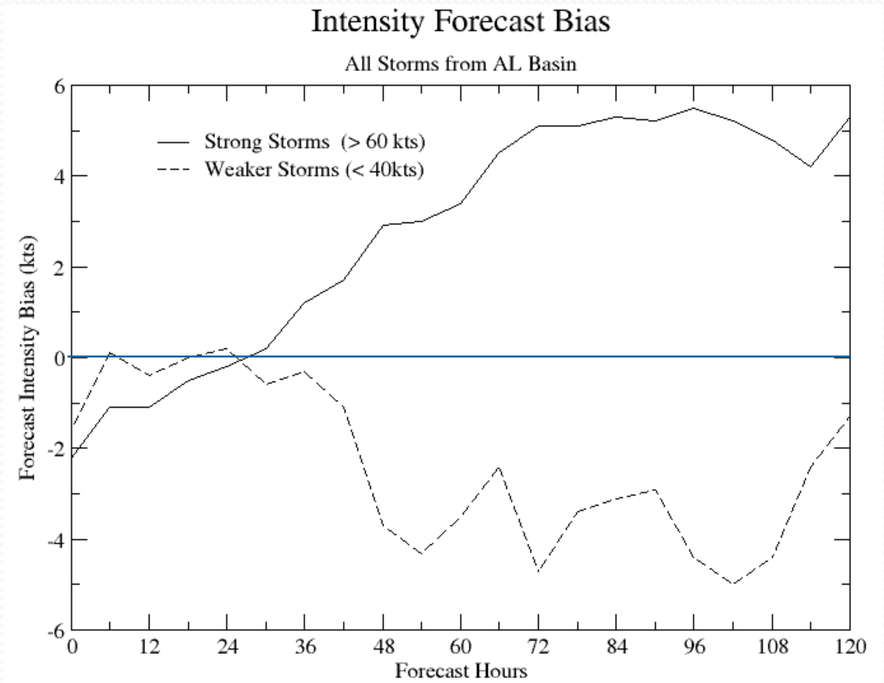


- Ranks the ensemble member forecasts from lowest value to highest value for each specific forecast;
- Forms bins using the interval between each pair of ranked values;
- Count the % of the Obs. that fall into each bins;
- Evaluate whether an ensemble system is biased, under- or over-dispersed.

1. Only AL basin data is used;
2. The observed values are distributed evenly in inner bins;
3. The N/S bias is mostly due to longer lead time;
4. W (fast) track bias.

Analysis of Rank Histogram for Intensity Forecast

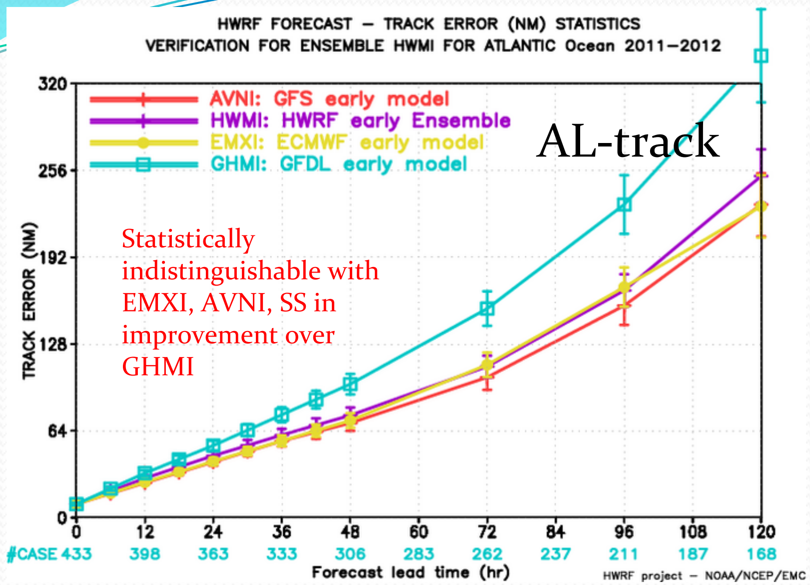


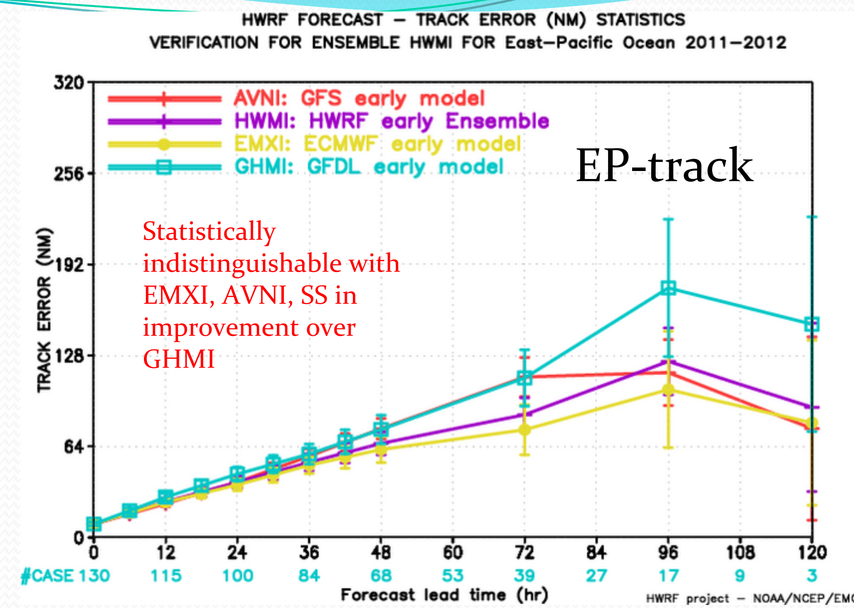


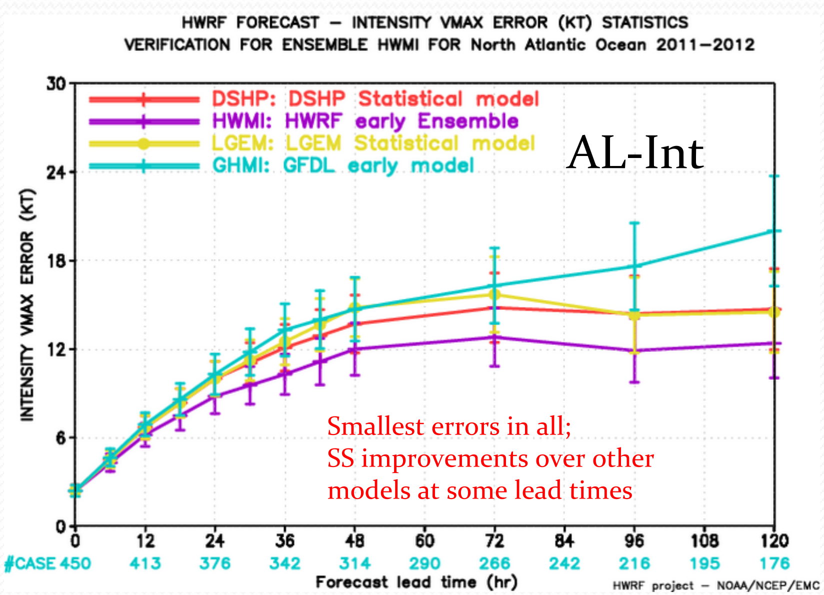
1. The observed values are distributed evenly in inner bins;
2.The deterministic model tends to have positive bias for strong storms and negative bias for weaker storms.

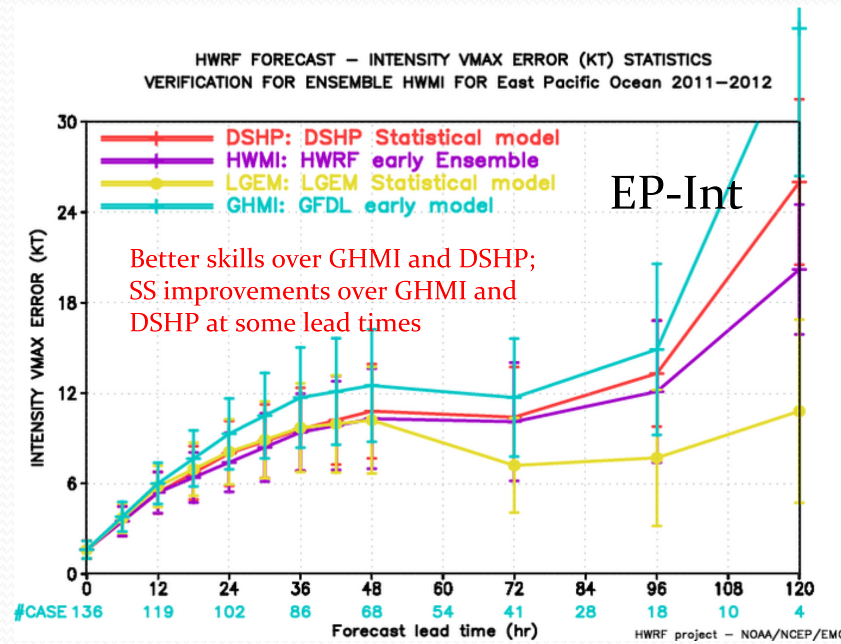
Comparisons of HWRF EPS with the Best Models in 2012

Track: AVNI, EMXI, GHMI
Intensity: DSHP, LGEM, GHMI


HWRF FORECAST – TRACK ERROR (NM) STATISTICS
VERIFICATION FOR ENSEMBLE HWMI FOR ATLANTIC Ocean 2011–2012
AVNI: GFS early model
HWMI: HWRF early Ensemble
EMXI: ECMWF early model
GHMI: GFDL early model
AL-track
Statistically indistinguishable with EMXI, AVNI, SS in improvement over GHMI
TRACK ERROR (NM)
#CASE 433 398 363 333 306 283 262 237 211 187 168
Forecast lead time (hr)
HWRF project – NOAA/NCEP/EMC


HWRF FORECAST – TRACK ERROR (NM) STATISTICS
VERIFICATION FOR ENSEMBLE HWMI FOR East–Pacific Ocean 2011–2012
AVNI: GFS early model
HWMI: HWRF early Ensemble
EMXI: ECMWF early model
GHMI: GFDL early model
EP-track
Statistically indistinguishable with EMXI, AVNI, SS in improvement over GHMI
TRACK ERROR (NM)
#CASE 130 115 100 84 68 53 39 27 17 9 3
Forecast lead time (hr)
HWRF project – NOAA/NCEP/EMC


HWRF FORECAST – INTENSITY VMAX ERROR (KT) STATISTICS
VERIFICATION FOR ENSEMBLE HWMI FOR North Atlantic Ocean 2011–2012
DSHP: DSHP Statistical model
HWMI: HWRF early Ensemble
LGEM: LGEM Statistical model
GHMI: GFDL early model
AL-Int
Smallest errors in all;
SS improvements over other models at some lead times
INTENSITY VMAX ERROR (KT)
#CASE 450 413 376 342 314 290 266 242 216 195 176
Forecast lead time (hr)
HWRF project – NOAA/NCEP/EMC


HWRF FORECAST – INTENSITY VMAX ERROR (KT) STATISTICS
VERIFICATION FOR ENSEMBLE HWMI FOR East Pacific Ocean 2011–2012
DSHP: DSHP Statistical model
HWMI: HWRF early Ensemble
LGEM: LGEM Statistical model
GHMI: GFDL early model
EP-Int
Better skills over GHMI and DSHP;
SS improvements over GHMI and DSHP at some lead times
INTENSITY VMAX ERROR (KT)
#CASE 136 119 102 86 68 54 41 28 18 10 4
Forecast lead time (hr)
HWRF project – NOAA/NCEP/EMC

Pairwise Comparison Between HWMI and Top-flight Models Intensity For Atlantic Basin

Atlantic Basin	Forecast Hour	12	24	36	48	60	72	84	96	108	120
	GHMI *Land/Water*	0.7 10% 0.989	1.5 15% 0.989	3.1 24% 0.999	3.0 21% 0.997	3.2 21% 0.988	3.8 23% 0.990	4.3 26% 0.990	5.3 30% 0.995	5.7 31% 0.994	6.5 34% 0.996
	GHMI *Water Only*	0.6 9% 0.972	1.5 15% 0.985	3.2 25% 0.999	3.0 21% 0.991	3.0 19% 0.952	4.0 24% 0.979	4.5 26% 0.961	5.4 30% 0.978	4.8 27% 0.949	5.7 30% 0.955
	LGEM *Land/Water*	0.5 8% 0.944	1.1 12% 0.976	2.2 18% 0.997	2.7 19% 0.998	2.9 19% 0.992	2.6 17% 0.954	2.2 15% 0.898	1.9 13% 0.814	0.8 6% 0.456	0.8 6% 0.376
	LGEM *Water Only*	0.3 5% 0.790	1.0 11% 0.968	2.2 18% 0.998	2.6 18% 0.994	2.7 18% 0.985	2.4 16% 0.904	1.7 12% 0.730	0.8 6% 0.391	-0.8 -7% 0.441	-0.7 -6% 0.388
	DSHP *Land/Water*	0.5 7% 0.897	1.2 12% 0.977	2.0 16% 0.996	2.1 15% 0.986	2.2 15% 0.977	2.3 15% 0.961	2.5 17% 0.969	2.4 16% 0.946	2.0 14% 0.814	1.2 9% 0.433
	DSHP *Water Only*	0.3 4% 0.663	1.2 12% 0.977	2.0 17% 0.996	1.9 14% 0.969	2.2 15% 0.976	2.5 16% 0.967	2.1 14% 0.903	1.5 10% 0.728	0.4 3% 0.191	-0.3 -3% 0.170

From TCMT verification

Pairwise Comparison Between HWMI and Top-flight Models Intensity For East-Pacific Basin

	Forecast Hour	12	24	36	48	60	72	84	96	108	120
Eastern North Pacific Basin	GHMI *Land/Water*	0.5 8% 0.711	2.0 21% 0.999	2.5 21% 0.970	1.9 15% 0.954	2.2 18% 0.969	1.5 13% 0.818	0.3 3% 0.142	0.2 2% 0.077	4.9 27%	9.3 30%
	GHMI *Water Only*	0.5 9% 0.763	2.0 21% 0.999	2.5 21% 0.970	1.9 15% 0.954	2.3 19% 0.974	1.5 13% 0.795	0.0 0% 0.016	0.3 2% 0.102	4.9 27%	9.3 30%
	LGEM *Land/Water*	0.5 8% 0.712	1.1 13% 0.862	0.6 6% 0.401	0.0 0% 0.011	-0.5 -6% 0.130	-3.0 -41% 0.567	-4.4 -65% 0.673	-4.0 -51% 0.476	-4.7 -54%	-9.5 -88%
	LGEM *Water Only*	0.4 8% 0.681	1.1 13% 0.862	0.6 6% 0.401	0.0 0% 0.011	-0.5 -6% 0.130	-3.0 -41% 0.567	-4.4 -65% 0.673	-4.0 -51% 0.476	-4.7 -54%	-9.5 -88%
	DSHP *Land/Water*	0.0 1% 0.089	0.7 9% 0.896	0.2 2% 0.223	0.5 4% 0.465	-0.3 -3% 0.216	0.1 1% 0.039	-0.1 -1% 0.034	2.2 16% 0.728	5.3 28%	5.8 22%
	DSHP *Water Only*	0.0 1% 0.074	0.7 9% 0.896	0.2 2% 0.223	0.5 4% 0.465	-0.3 -3% 0.216	0.1 1% 0.039	-0.1 -1% 0.034	2.2 16% 0.728	5.3 28%	5.8 22%

Sample size is not big enough for EP

From TCMT verification

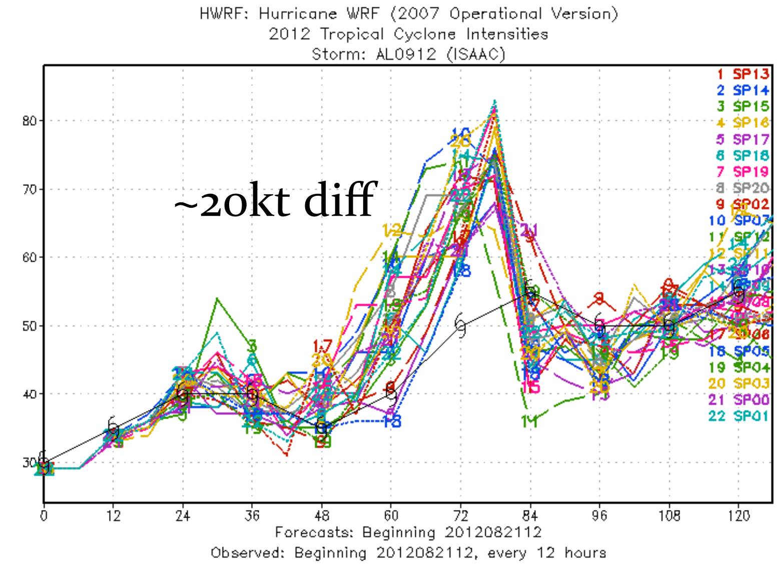
How stochastic convective trigger function perturbation affects storm intensity forecasts

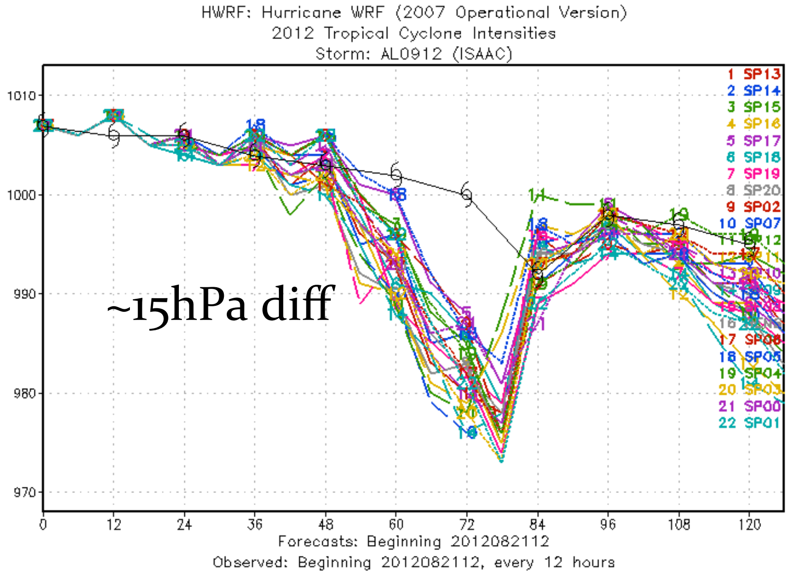
1. What is the cause of storm intensity ensemble spread? Convective trigger perturbation or initial storm cycling?

2. How stochastic convective trigger perturbations in D01 and D02 impact storm intensity? (Cumulus convection is explicitly expressed in domain 3, ~3km, in current configuration)?

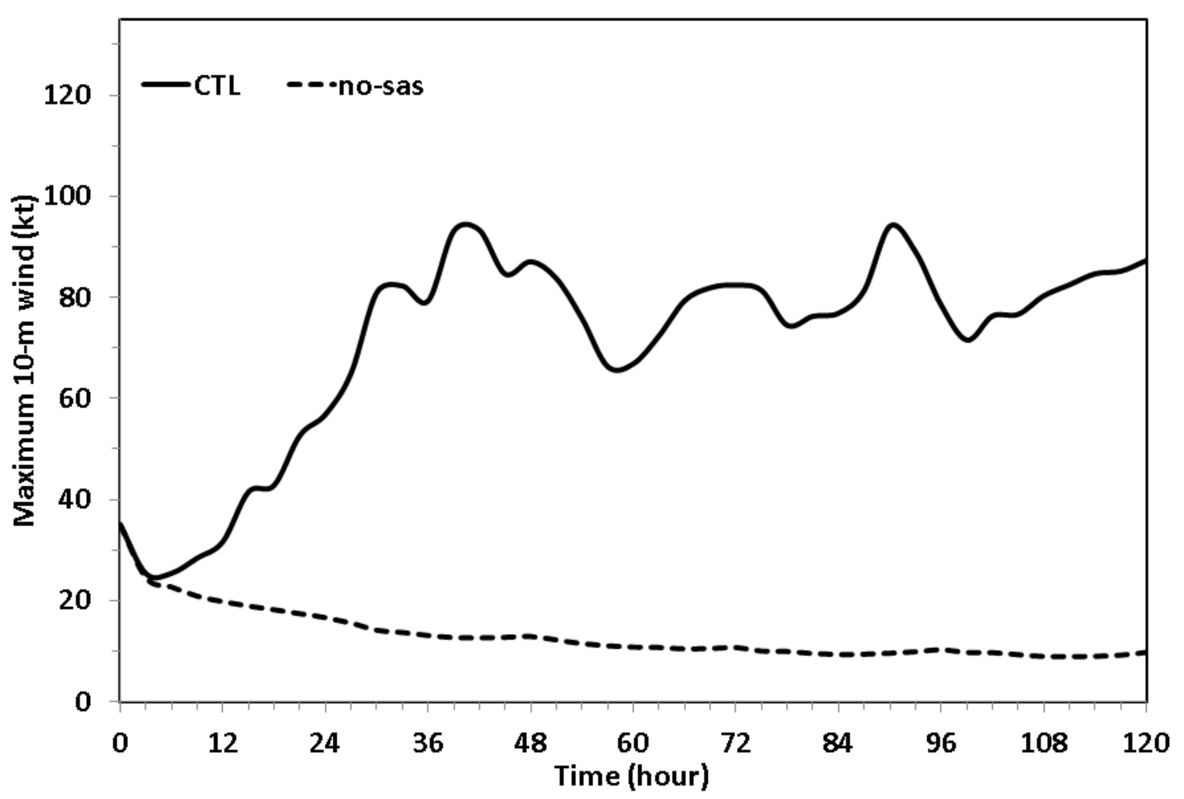
Cold-start Experiment With GFS

Several cold-start experiments indicate there are large intensity forecast spread even without model cycling.



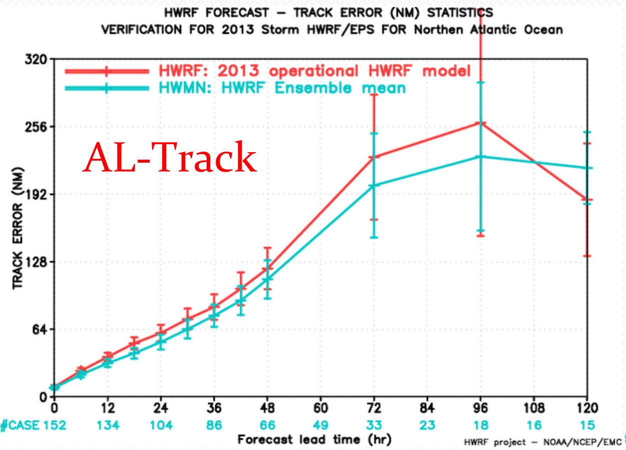


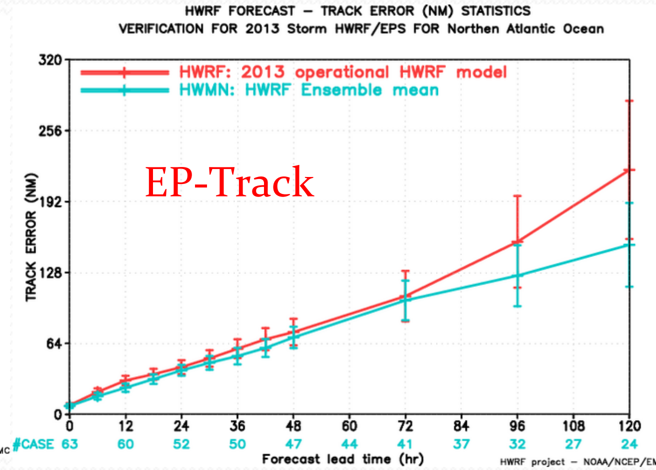
Impact of Cumulus Convection on Large Scale Flow
Idealized Experiment

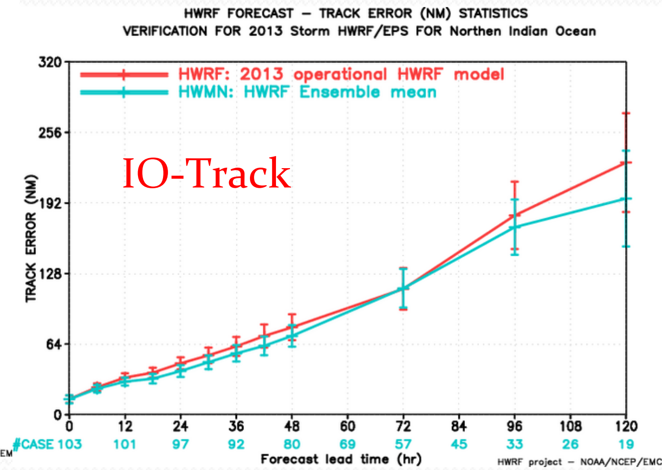


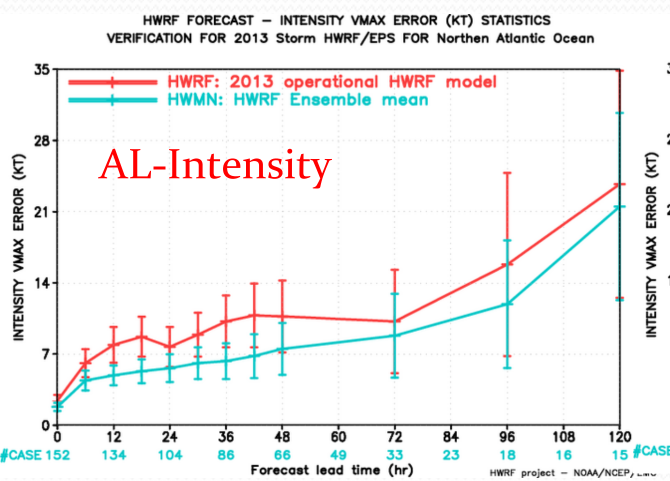
Solid line: SAS in D01 and D02 (27km and 9km)
Dash line: no SAS in D01 and D02.
Sub-grid convection is explicitly expressed in D03 in both exps.
The model storm will not develop when SAS scheme is turned off in the 27- and 9-km domains even if the domain 3 (3km) resolution is high enough to resolve the convection scheme.

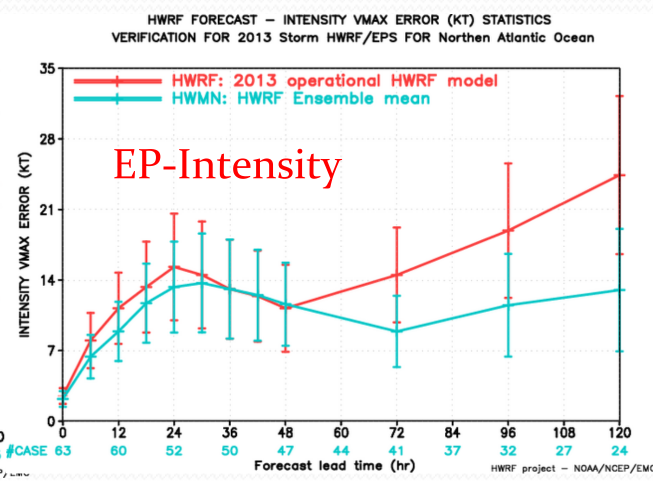
HWRF/EPS Verifications for 2013 Storms

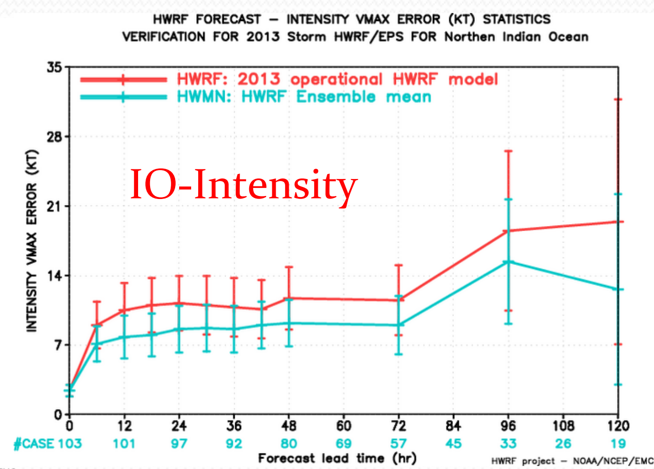












Summary

- Ensemble forecast is necessary due to the chaotic and non-linear nature of the atmosphere;
- Additional uncertainties, e.g. sub-grid scales, initial storm positions, needs to be take into account in EPS;
- HWRF-based EPS includes perturbations from large scale flows (GEFS) and sub-grid scales (physics-based SCT);
- HWRF EPS introduces no bias but inherits some biases from the deterministic model in terms of track/intensity forecasts;
- Both HWMI track and intensity forecast skills are improved over its deterministic versions (H212 and FY13), with more improvements in intensity forecasts;
- The HWMI forecasts are statistically undistinguishable from that of the top-flight models; The HWMI produces SS improvements over GHMI in both track and intensity forecasts at all lead times.

Summary (Continued)

- In HWRF EPS, the intensity spread may be a good predictor for intensity forecast skill up to 72h, not the track spread;

Future work:

- Experiment with SAS scheme turned on in third domain;
- Develop an ensemble ranking system that identifies the ensemble member closest to the ensemble mean based on their predicted storm centers, intensities, and storm structures.