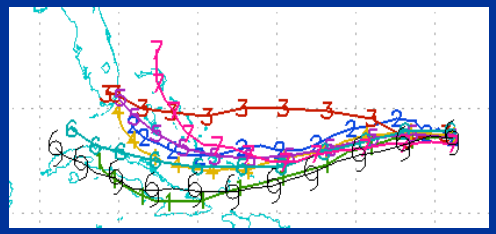
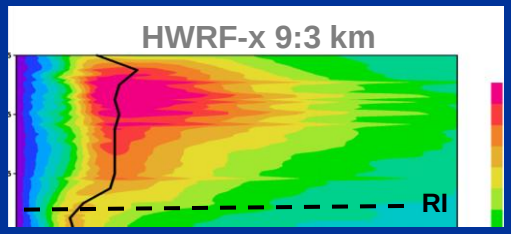
HFIP Verification Team: FY09 Accomplishments and Future Plans



HFIP Team Meeting
Miami, FL
09 November 2009



- *OAR / GFDL*
  - Tim Marchok
- *OAR / AOML / HRD*
  - Rob Rogers
- *OAR / ESRL*
  - Mike Fiorino
- *NRL*
  - Jim Goerss, Hao Jin
- *TCMT / DTC*
  - Barb Brown
- *NESDIS*
  - Mark DeMaria
- *NCEP / NHC*
  - James Franklin
- *NCEP / EMC*
  - Vijay Tallapragada

Two basic tasks of the Verification Team

Performance of Verifications

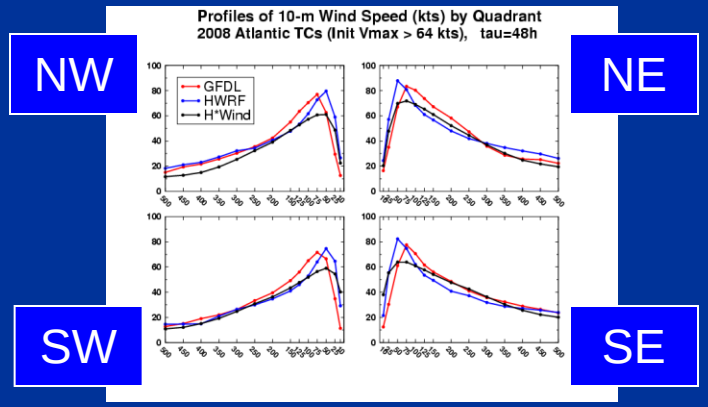
- Defining a set of verification metrics
- Performing HRH Verifications (DTC / TCMT)
- Defining a set of cases for the baseline
- Performing baseline and annual verifications for HFIP
- Reporting results to HFIP mgmt

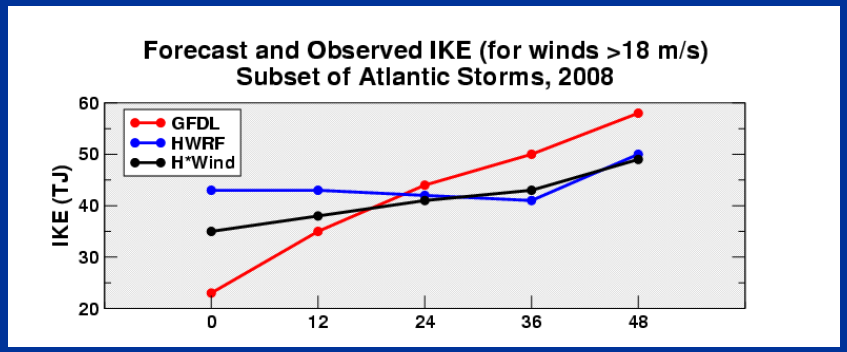
Development / assimilation of new verification techniques

- RI / RW
- Surface wind structure
- Rainfall
- Ensemble-based probabilistic guidance
- Forecast consistency
- TC genesis

Verification Team 2009 Milestones & Deliverables

- ✓ Baseline verifications (NHC):
  - ✓ Establish set of forecast metrics to be used for baseline verification
  - ✓ Establish set of test cases to be used in baseline verification
  - ✓ Perform baseline verifications
- ✓ Upgrade standard track and intensity forecast verification software to extend verifications out to 7 days (NHC)
- ✓ Begin developing software to validate wind structure forecasts (GFDL, HRD)



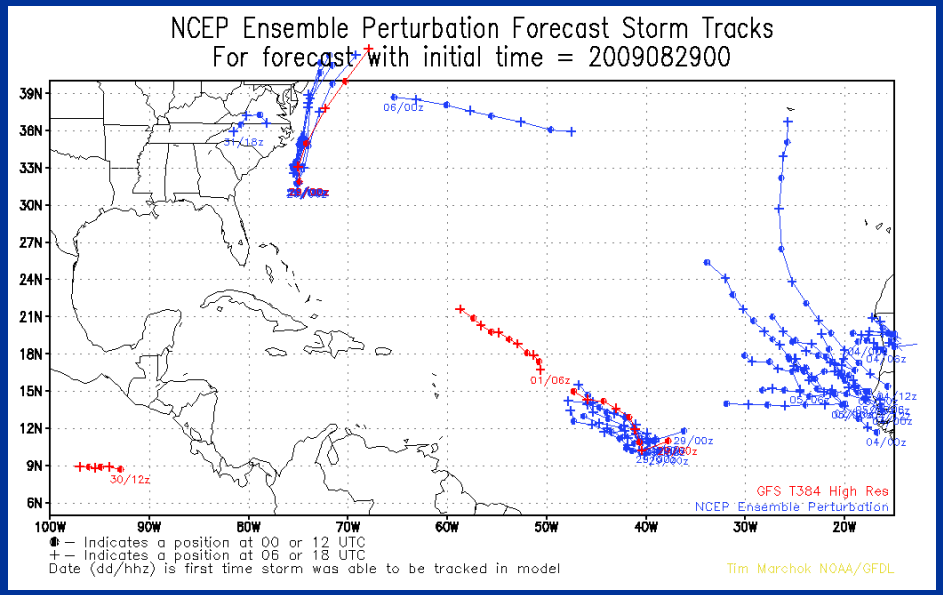


Surface wind structure diagnostics extracted from model data are compared against similar diagnostics extracted from H*Wind data

Verification Team 2009 Milestones & Deliverables

- ✓ 1st Diagnostics / Verification Workshop held at NHC, May 2009
  - ➢ http://rammb.cira.colostate.edu/research/tropical_cyclones/hfip/workshop_2009

- ✓ Coordinate with other HFIP teams to determine if TC genesis will be an area addressed by the verification team (GFDL)



Work began to modify and upgrade an automated system to detect and track new TCs within model forecasts

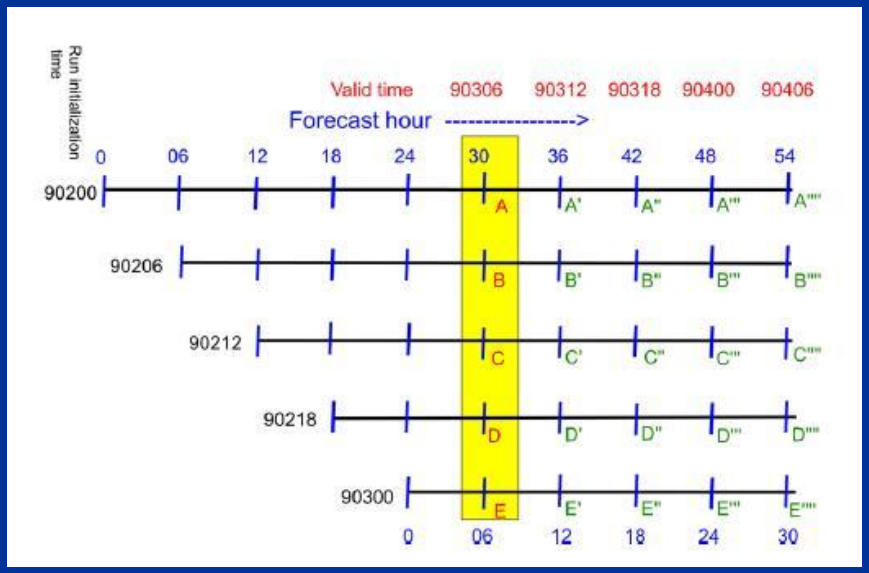
Additional FY09 Activities & Accomplishments

- High Resolution Hurricane test
  - DTC provided modeling groups with Output Module Package to output data in format needed for vortex tracker.
  - GFDL and DTC collaborated to make extensive upgrades to the vortex tracker needed for the hi-res, moveable mesh models.
  - DTC evaluated retrospective HRH forecasts and prepared report.

- Development of RI / RW verification tool (DTC)
  - Frequency of occurrence and onset timing errors
  - Event-based contingency tables based on ATCF files
  - Proportion Correct (PC), Probability of Detection (POD), Critical Success Index (CSI), and False Alarm Rate (FAR)

Additional FY09 Activities & Accomplishments

- Development of a consistency verification tool (DTC)



Differences are computed among track forecast positions from consecutive forecasts valid at the same time….

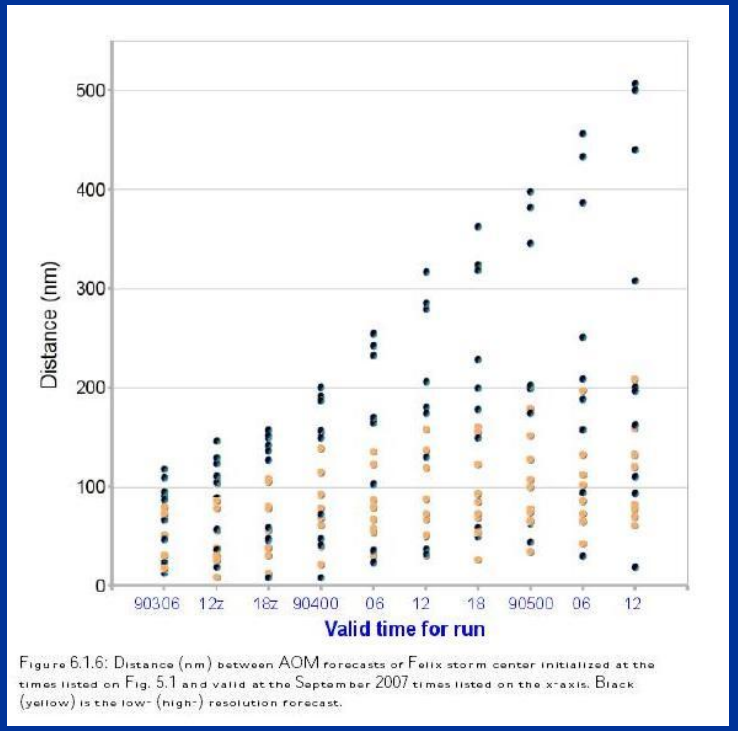


Figure 6.1.6: Distance (nm) between AOM forecasts of Felix storm center initialized at the times listed on Fig. 5.1 and valid at the September 2007 times listed on the x-axis. Black (yellow) is the low- (high-) resolution forecast.

….resulting in a plot that allows for comparison of consistency between two different models.

Assessing performance: Baseline development

- James Franklin offered the following baseline proposals in a draft from 05 May 2009:

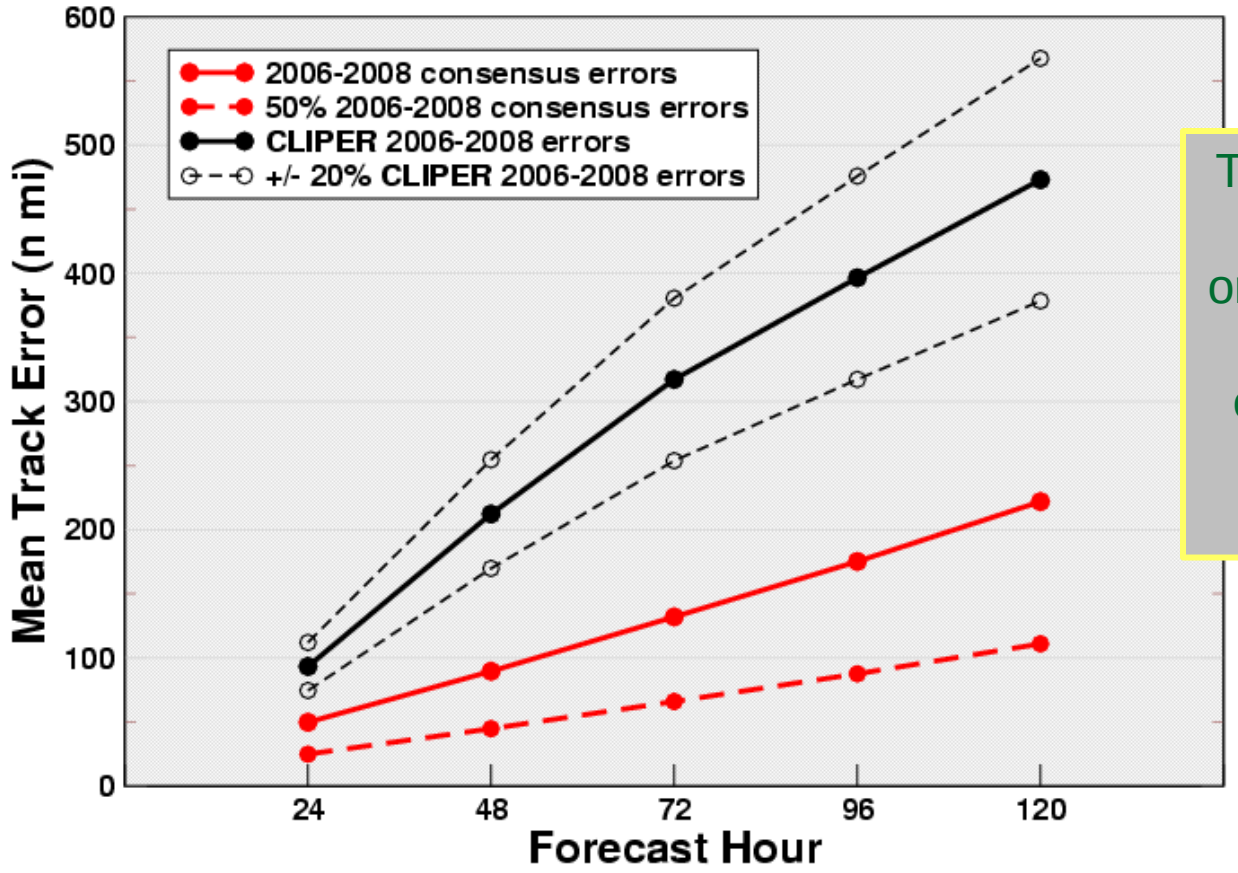
  - Track and intensity goals: Use a consensus of operational guidance models, evaluated for the Atlantic basin over the period 2006-2008.

    - Track membership: GFSI, GFDI, UKMI, NGPI, HWFI, GFNI, EMXI.
    - Intensity membership: GHMI, HWFI, DSHP, and LGEM.

  - Goal for 7-day lead time: Accuracy at least as good as the 5-day official forecast in 2003.

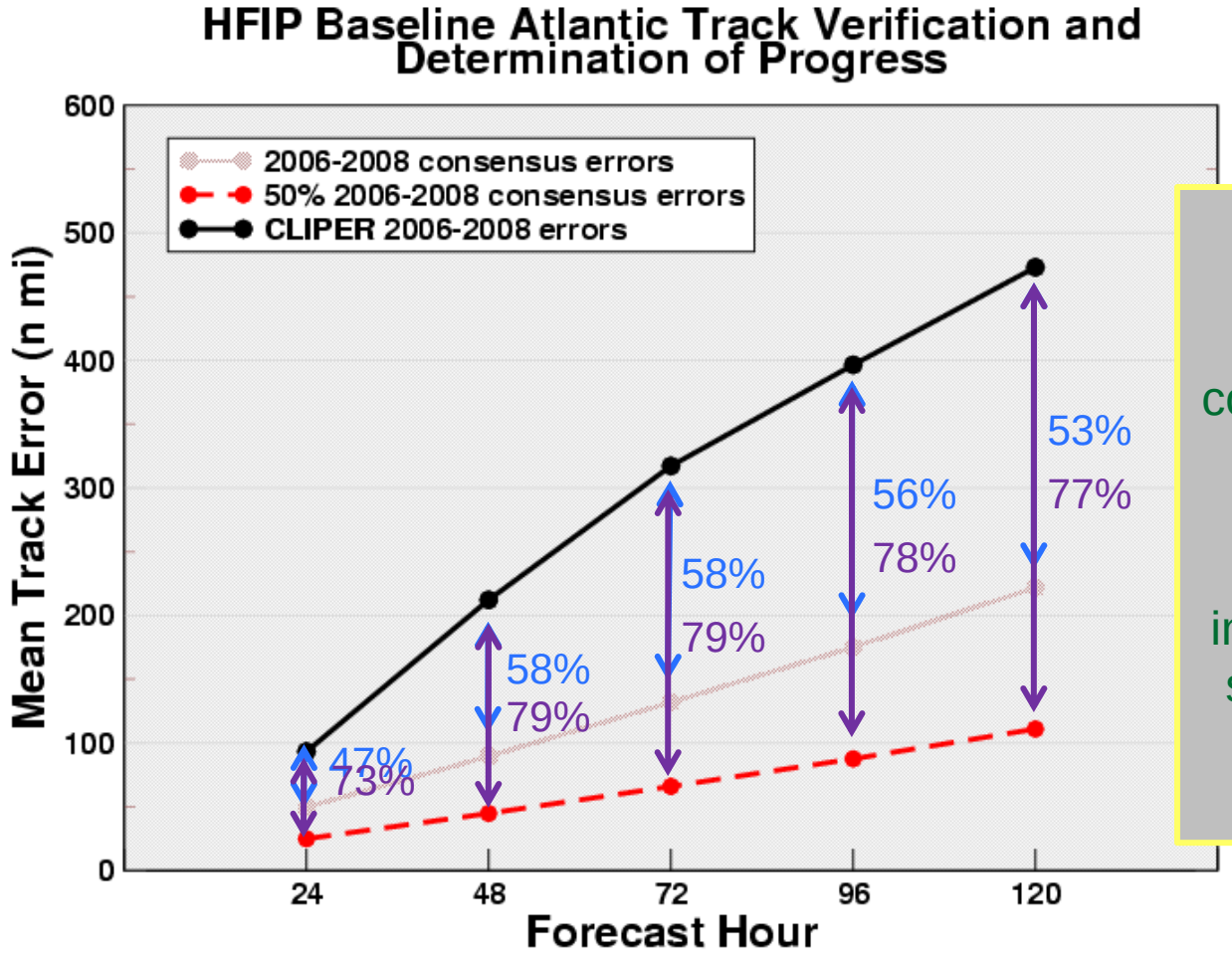
  - RI / RW: No official forecasts of RI are given. Could use GFDL model evaluated over the period, 2006-2008.

Baseline development: Track Example



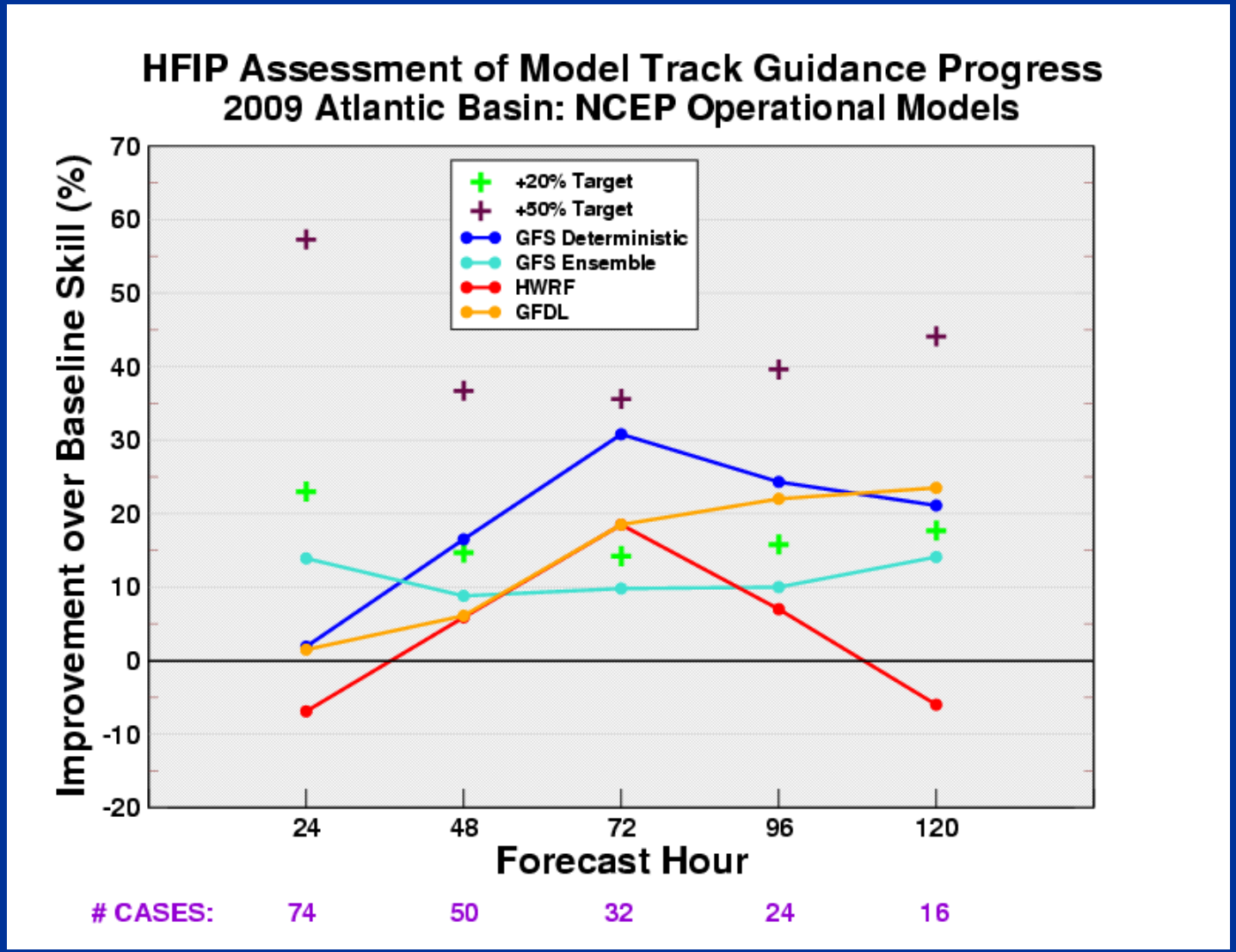
That 50% error reduction is only meaningful if CLIPER errors remain the same or increase.

Baseline and Assessment of Progress

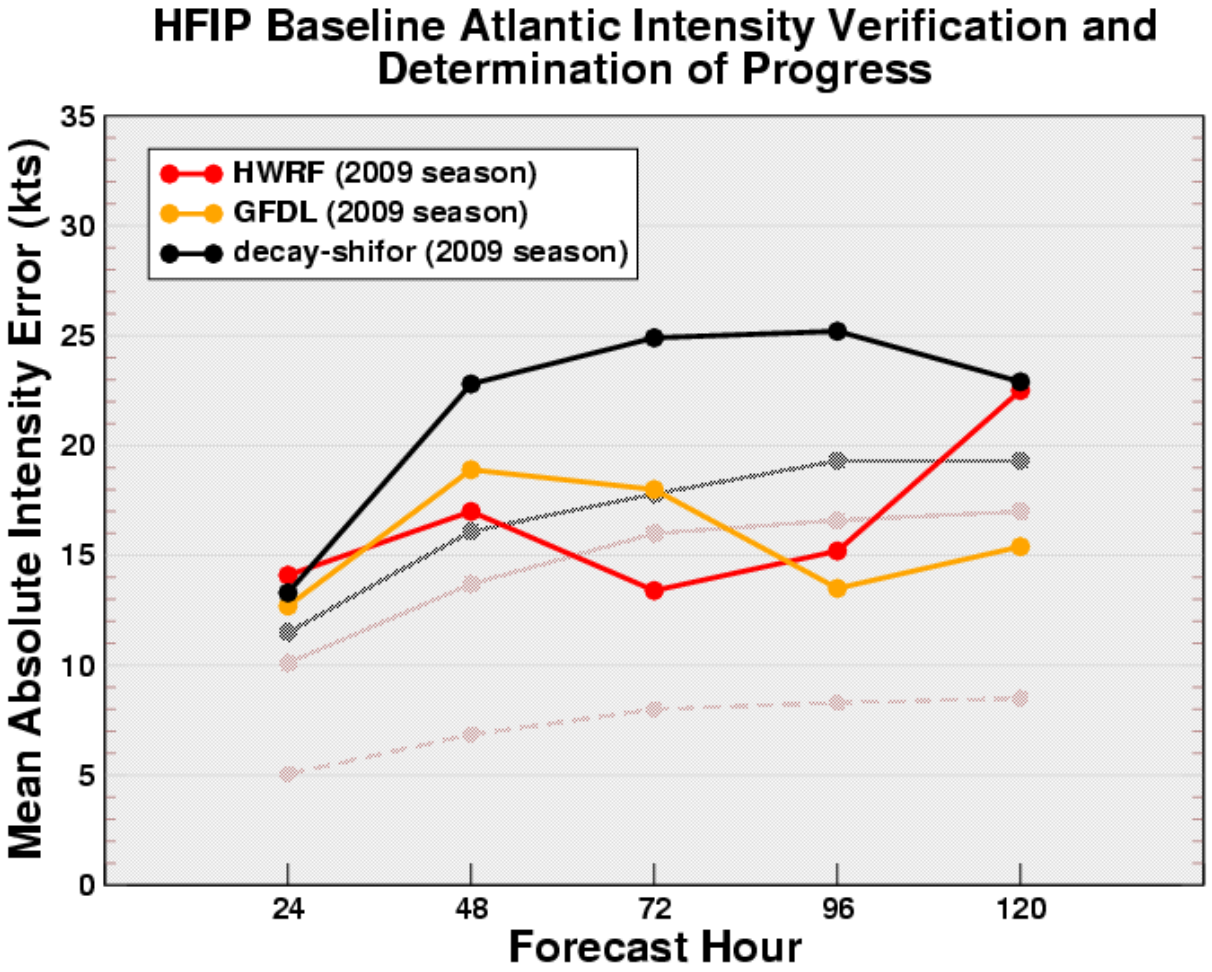


The 50% error reduction, assuming constant CLIPER errors, represents a significant increase in track skill that will be used as the target.

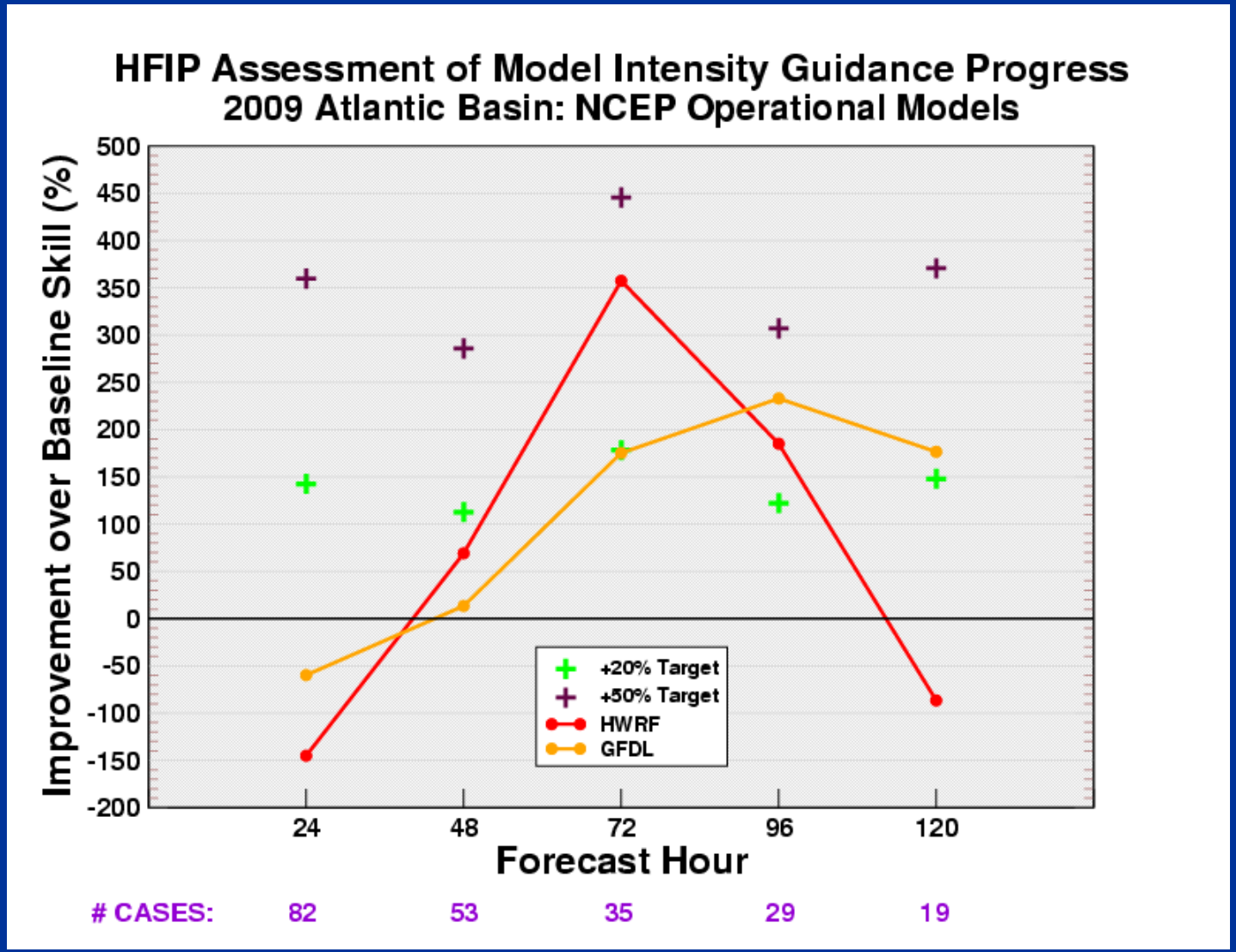
Baseline & NCEP 2009 operational track models



Baseline development: Intensity



Baseline development: Intensity



Verification Team: Future Plans

Efforts over the next 1-3 years will focus on these 3 primary areas:

- Performance of model verifications
- Development / transition of new verification techniques
- Organization and distribution of verification software

Future Plans: Model Verifications

- Evaluation of 2009 demo forecasts (TCMT/DTC)
  - Initial overview of HFIP demo forecasts using contributed track forecasts (Goal: 15 January 2010)
  - Run GFDL tracker on Tier 2 data; apply verification software and summarize “final” results (Goal: 31 March 2010)
- Extended evaluation of HRH sample (TCMT/DTC)
  - e.g., examine skill of poor-man’s ensemble from member models
- Development of skill baseline for 6/7-day forecasts (NHC)
  - Current plan is just to use the 5-day Official errors from the first year 5-day forecasts were used (2003)

Future Plan

- Design a testing methodo
  - Goal: Define experimer in predicting HFIP-relev
  - Recommend retrospect
- Design a testing methodo
  - Design specific compre be relevant for pre-impl
  - Investigate new statistical approaches for evaluating & comparing performance of multiple models (e.g., How should a representative sample be chosen? Should correlations between variables such as track & intensity be taken into account?)
- Set up testing & evaluation infrastructures (TCMT/DTC)
  - Extension of testing setup used for the HRH project
  - Extend for evaluation on njet of Demo and HWRF forecasts

REGIONAL MODEL AVAILABILITY: ATLANTIC 2009

	00	12	24	36	48	72	96	120
GFDL	116	100	83	69	56	39	32	21
HWRF	115	103	84	70	58	39	33	20
H3HW	54	49	38	30	25	18	15	10
MMEN	37	32	25	20	15	8	4	1
COTC	17	13	9	7	6	4	2	0
ARRE	17	15	12	9	6	1	0	0
AHW1	16	13	9	7	6	4	2	0
ARFS	6	6	6	6	6	4	2	0
HOMOG(*)	15	11	9	7	6	4	2	0
HOMOG(**)	3	3	3	3	3	1	0	0

(*) = All models except ARRE and ARFS
(**) = All models

Future Plans: New Verification Techniques

- Complete transition of verification packages (GFDL, HRD)
  - Wind structure verification
  - TC rainfall verification
- Complete development of verification & other packages (GFDL)
  - Modifications to vortex tracker for genesis
  - Software to validate model genesis forecasts
- Improve techniques (GFDL, TCMT/DTC)
  - Diagnose and improve vortex tracker

Future Plans: New Verification Techniques

- Investigate new techniques
  - Spatial verification of TC rainfall forecasts (TCMT/DTC)
  - Vortex-scale, 3-D wind structure validation (HRD, GFDL)
    - Use airborne doppler radar data to verify V_r, V_T, omega
  - Evaluation of ensemble forecasts (TCMT, others…?)
    - Focus on distributions as opposed to just the mean
    - Increased emphasis on ensemble-based probabilistic guidance
    - Collaborate with ensemble team on this issue

- Explore new areas (All)
  - Coordinate with observations team and coupled ocean / wave model development team – What should be evaluated for the ocean?

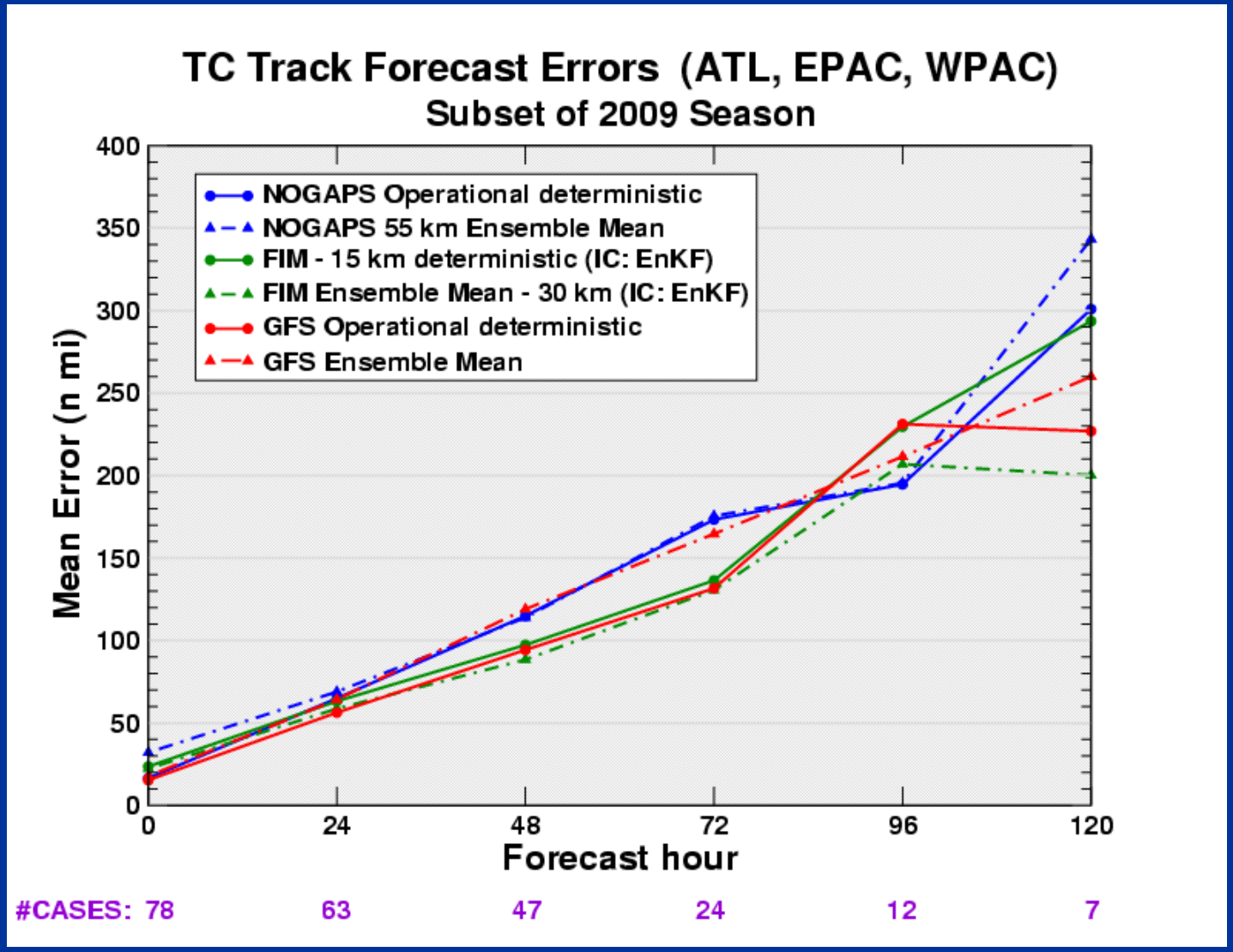
Future Plans: Distribution of Software

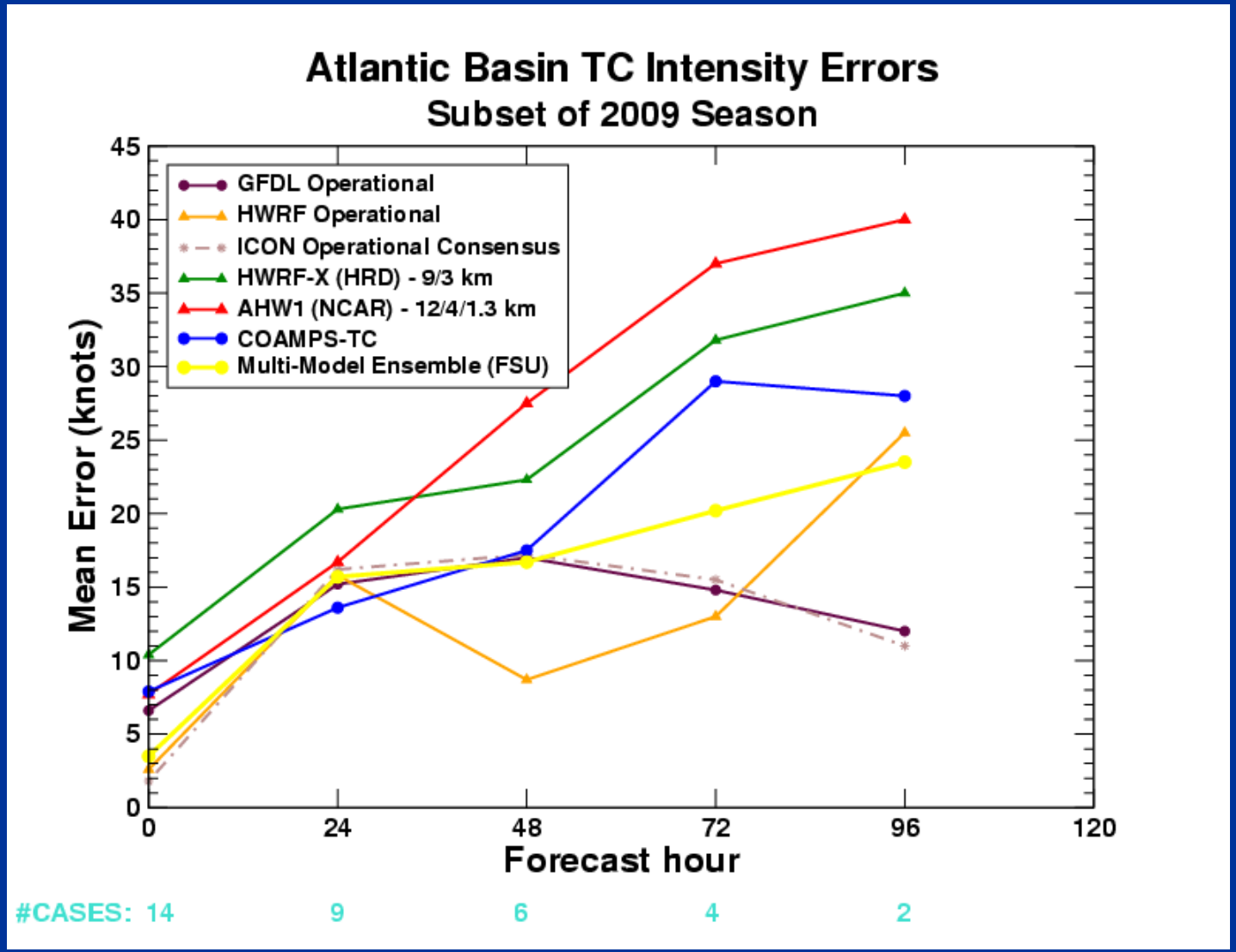
- GFDL vortex tracker (GFDL, TCMT/DTC)
  - Prepare code for public release in early 2010
  - Prepare documentation
  - Begin support of tracker at DTC

- TC verification toolkit (TCMT/DTC)
  - Assess available tools and offer feedback to the developers
  - Develop toolkit that will be available to the community
  - Consider expanding DTC's Model Evaluation Tools (MET) to have a TC verification component, or create a separate toolkit
  - Begin development of a tutorial for these tools

Questions?

Extra slides


TC Track Forecast Errors (ATL, EPAC, WPAC)
Subset of 2009 Season
NOGAPS Operational deterministic
NOGAPS 55 km Ensemble Mean
FIM - 15 km deterministic (IC: EnKF)
FIM Ensemble Mean - 30 km (IC: EnKF)
GFS Operational deterministic
GFS Ensemble Mean
Mean Error (n mi)
0
50
100
150
200
250
300
350
400
0
24
48
72
96
120
Forecast hour
#CASES: 78 63 47 24 12 7


Atlantic Basin TC Intensity Errors
Subset of 2009 Season
GFDL Operational
HWRF Operational
ICON Operational Consensus
HWRF-X (HRD) - 9/3 km
AHW1 (NCAR) - 12/4/1.3 km
COAMPS-TC
Multi-Model Ensemble (FSU)
Mean Error (knots)
0
5
10
15
20
25
30
35
40
45
0
24
48
72
96
120
Forecast hour
#CASES: 14
9
6
4
2