Overview of TCMT Activities and Preliminary Results from the 2014 Demo

TCMT Team: **Paul Kucera**, **Christopher Williams**, Barbara Brown, Louisa Nance, Kathryn Newman, Mrinal Biswas, John Halley Gotway, Tatiana Burek, Laurie Carson

November 05, 2014

Outline

- This presentation will provide an overview of the recent activities in the Tropical Cyclone Modeling Team (TCMT) at NCAR
  - 2014 Retrospective Evaluation
  - HFIP Reconnaissance Data Impact Study
  - 2014 Demonstration Activities
  - Real-time display and verification system
  - NHC display and diagnostic system development
  - HFIP Database
- The second half of the presentation will highlight preliminary results from the recently completed 2014 HFIP Demo

2014 Retrospective Evaluation

- Evaluation consisted of three primary approaches:

1. Direct comparison between Stream 1.5 candidate and each of last year's top-flight models
2. Assessment of how well Stream 1.5 candidate performed relative to last year's top-flight models as a group
3. Stream 1.5 candidate impact on operational consensus forecasts
   - Track: Variable consensus (at least 3 members must be present) w/ Stream 1.5 candidate: EMXI, GFSI, EGRI, GHMI, HWFI
   - Intensity: Fixed consensus (all members must be present): DSHP, LGEM, GHMI, HWFI

Teams that participated in the 2014 HFIP retrospective evaluation and basic model information

Organization	Model	Type
EMC	HWRF coupled with HYCOM	Regional-dynamic-deterministic
EMC	HWRF	Regional-dynamic-ensemble
NRL	COAMPS-TC	Regional-dynamic-deterministic
UW-Madison	UW-NMS	Regional-dynamic-deterministic
AOML/HRD	Basinscale HWRF	Regional-dynamic-deterministic
GFDL	GFDL hurricane model	Regional-dynamic-ensemble
ESRL	FIM	Global-dynamic-deterministic
FSU	Multi-model superensemble	Weighted-consensus
NESDIS/STAR and CIRA	SPICE	Statistical-dynamical-consensus

List of modeling groups, model configurations and type of guidance selected for 2014 Stream 1.5

Organization	Model	Track	Track consensus	Intensity	Intensity Consensus
NRL	COAMPS-TC			●	●
UW-Madison	UW-NMS				●
GFDL	GFDL hurricane model	●	●	●	●
NESDIS/STAR and CIRA	SPICE			●	

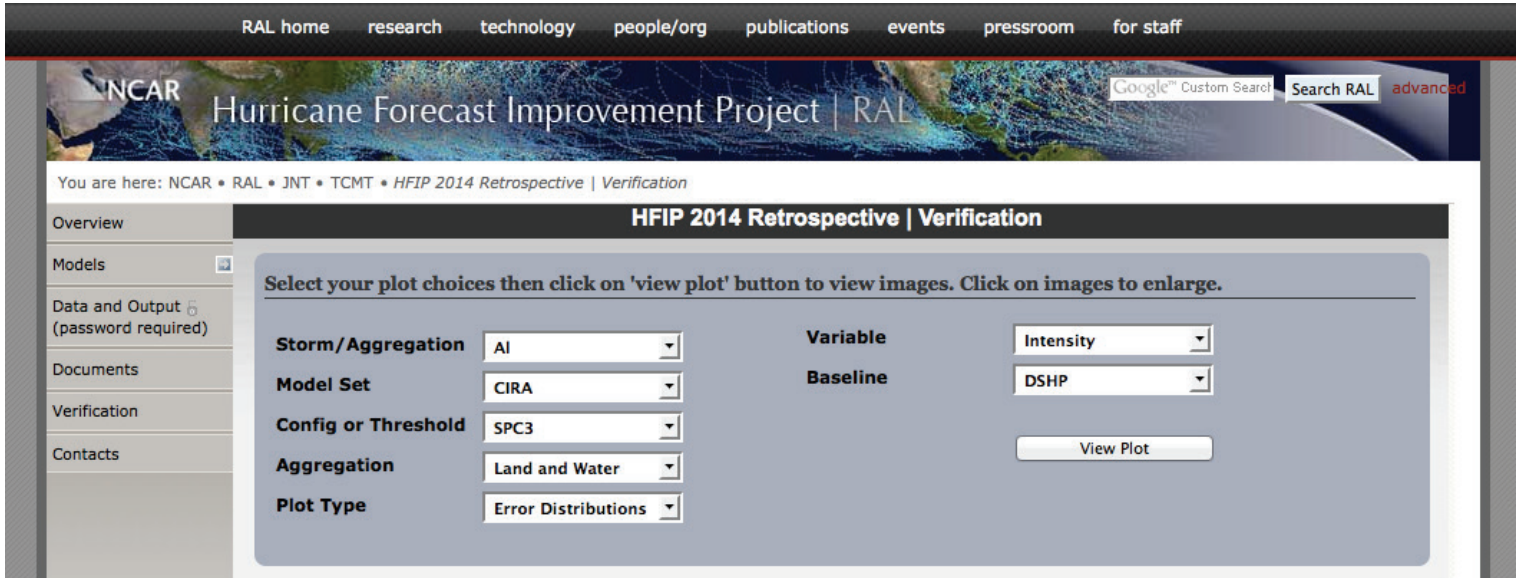
2014 Retrospective Evaluation

- All results are available on the TCMT Project Webpage:
- Full verification results:

http://www.ral.ucar.edu/projects/hfip/h2014/verify/

- Reports:

http://www.ral.ucar.edu/projects/hfip/h2014/documents/



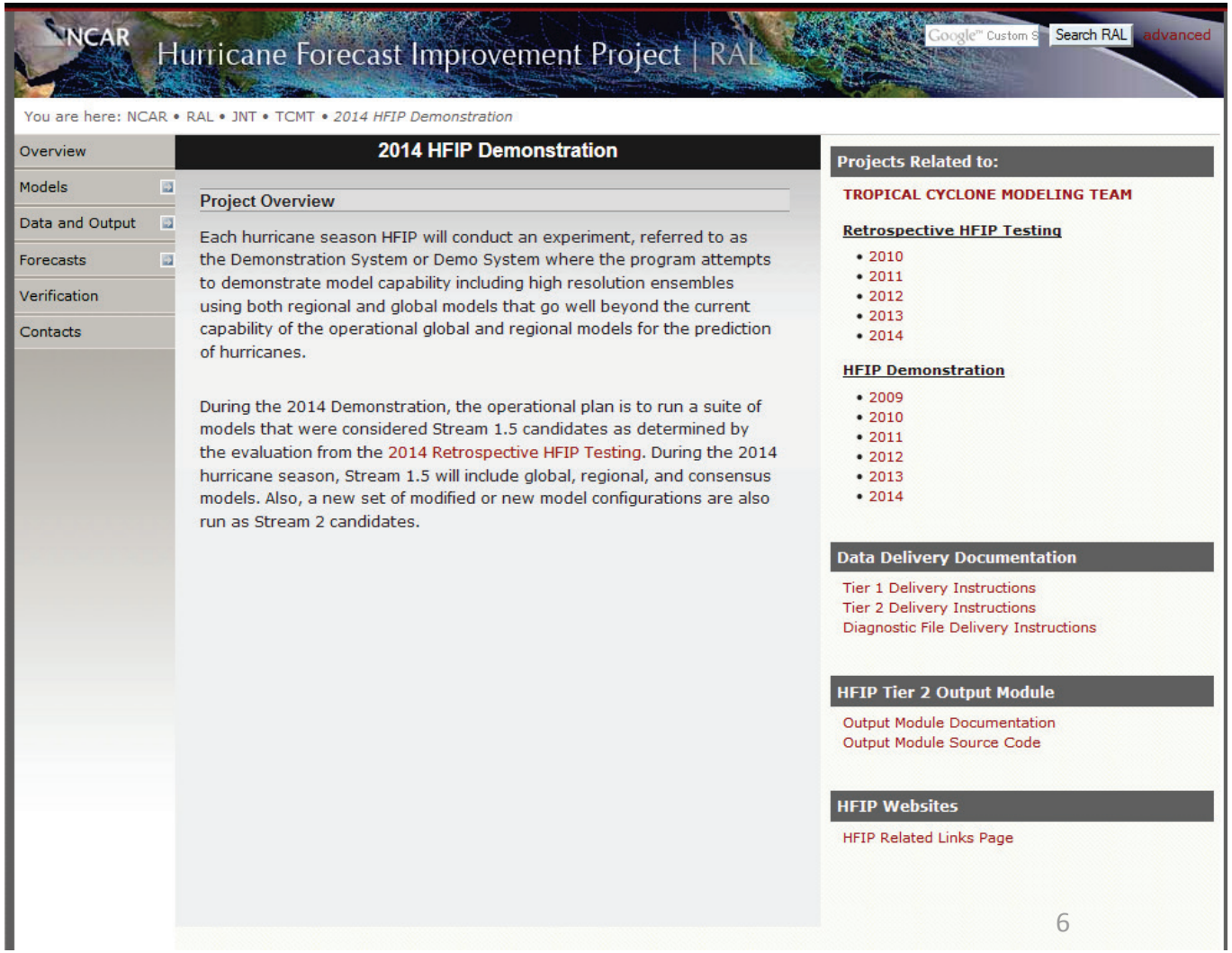
HFIP Reconnaissance Data Impact Study

- Conduct systematic investigation of the impact of aircraft reconnaissance data from the inner core of tropical cyclones on numerical guidance provided by regional tropical cyclone models
  - Provide a better understanding of the importance of tail Doppler radar (TDR) data
- Test impact of aircraft reconnaissance data in 3 regional TC models
  - PSU EnKF ARW hurricane analysis & prediction system
  - EMC operational 3-km HWRF (2013 version)
  - HRD HEDAS-HWRF (v3.2)
- A NCAR Technical Note final report will be released in the next couple of weeks
  - The report will be made available through the HFIP website (www.hfip.org)
  - Louisa Nance is the point of contact (nance@ucar.edu) for the study

2014 HFIP Demonstration Activities

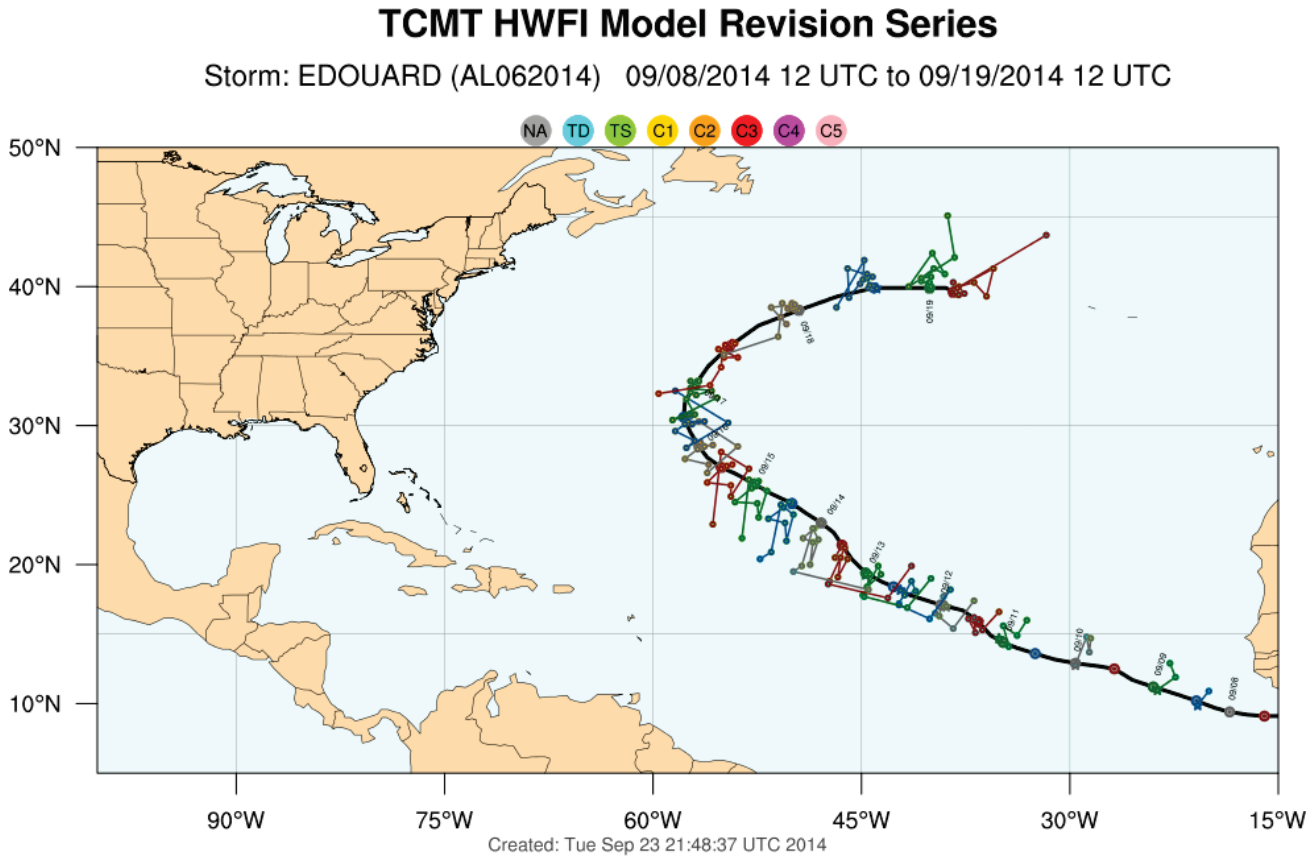
- Key Activities
  - Coordinated the archiving, dissemination, and display of Stream 1.5 and Stream 2 experimental forecasts
  - Updated forecast graphics for the TCMT and HFIP products pages
  - New evaluation tools
  - Implementation of a near real-time HFIP verification system
  - In-depth analysis of model performance (upcoming)
    - Diagnostic evaluation
    - HFIP 5 year evaluation of model performance

http://www.ral.ucar.edu/projects/hfip/d2014/

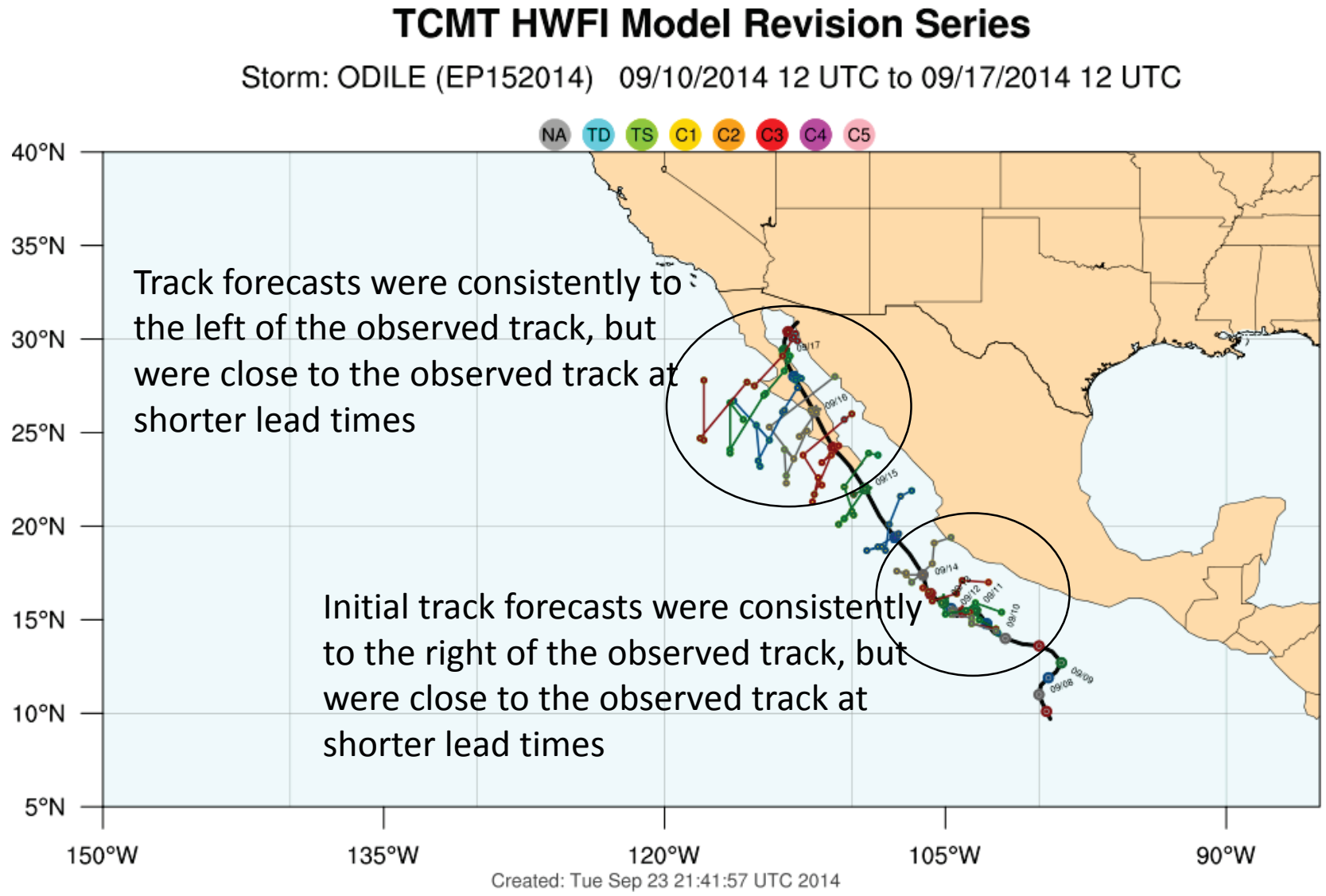


Revision Series Evaluation

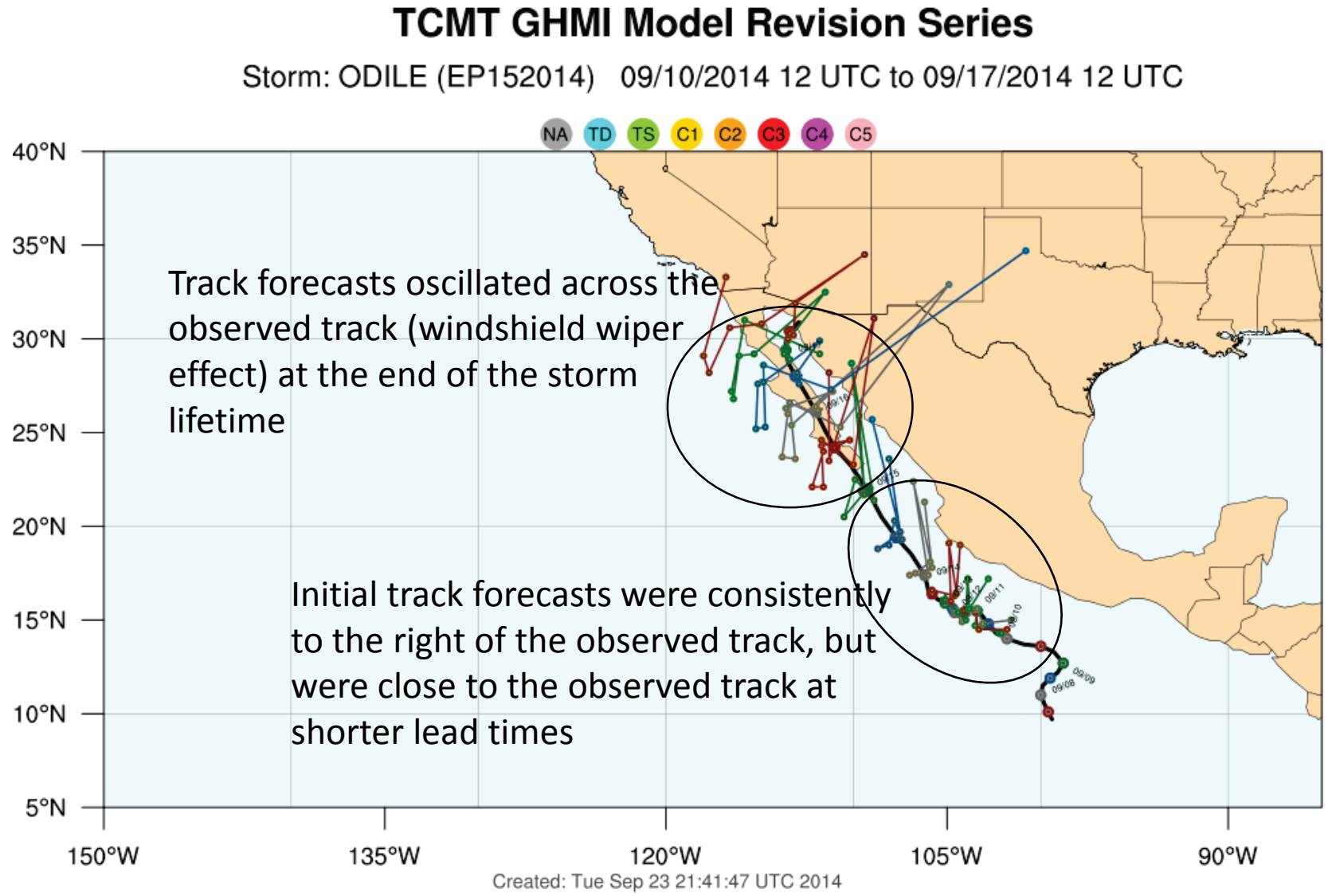
- The evaluation examines the consistency of a forecast for a given valid time.
- TCMT is currently developing metrics to quantify the consistency of forecasts
  - Path length of the revisions
  - Centroid distance from observed track
  - Area of revisions
  - …



Revision Series Evaluation



Revision Series Evaluation

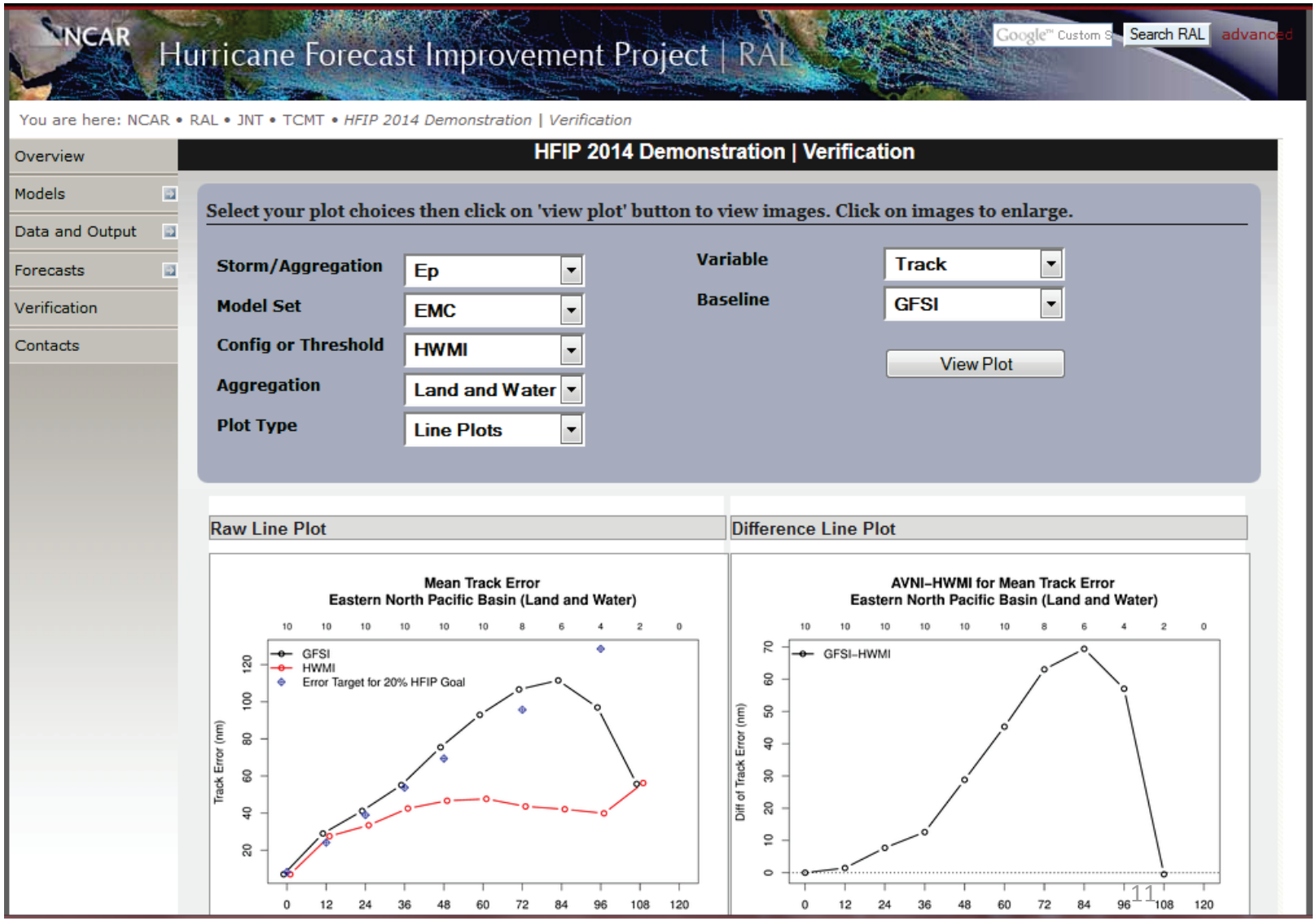


HFIP Demo Near Real-time Verification

- TCMT generated near real-time evaluation statistics for the 2014 Demonstration
- Statistics were computed from the beginning of hurricane season to present (run once per day)
  - Storm by storm statistics are also being generated
- Evaluation of Stream 1.5 and Stream 2 candidates
- A variety of evaluation products were generated (error distributions, mean errors, rank plots, skill scores) for Atlantic and Eastern Pacific Basins
- A web-based browser is available to provide access to the HFIP team

HFIP Demo Near Real-time Verification

- Link: http://www.ral.ucar.edu/projects/hfip/d2014/verify/
- Verification statistics were aggregated and updated once per day

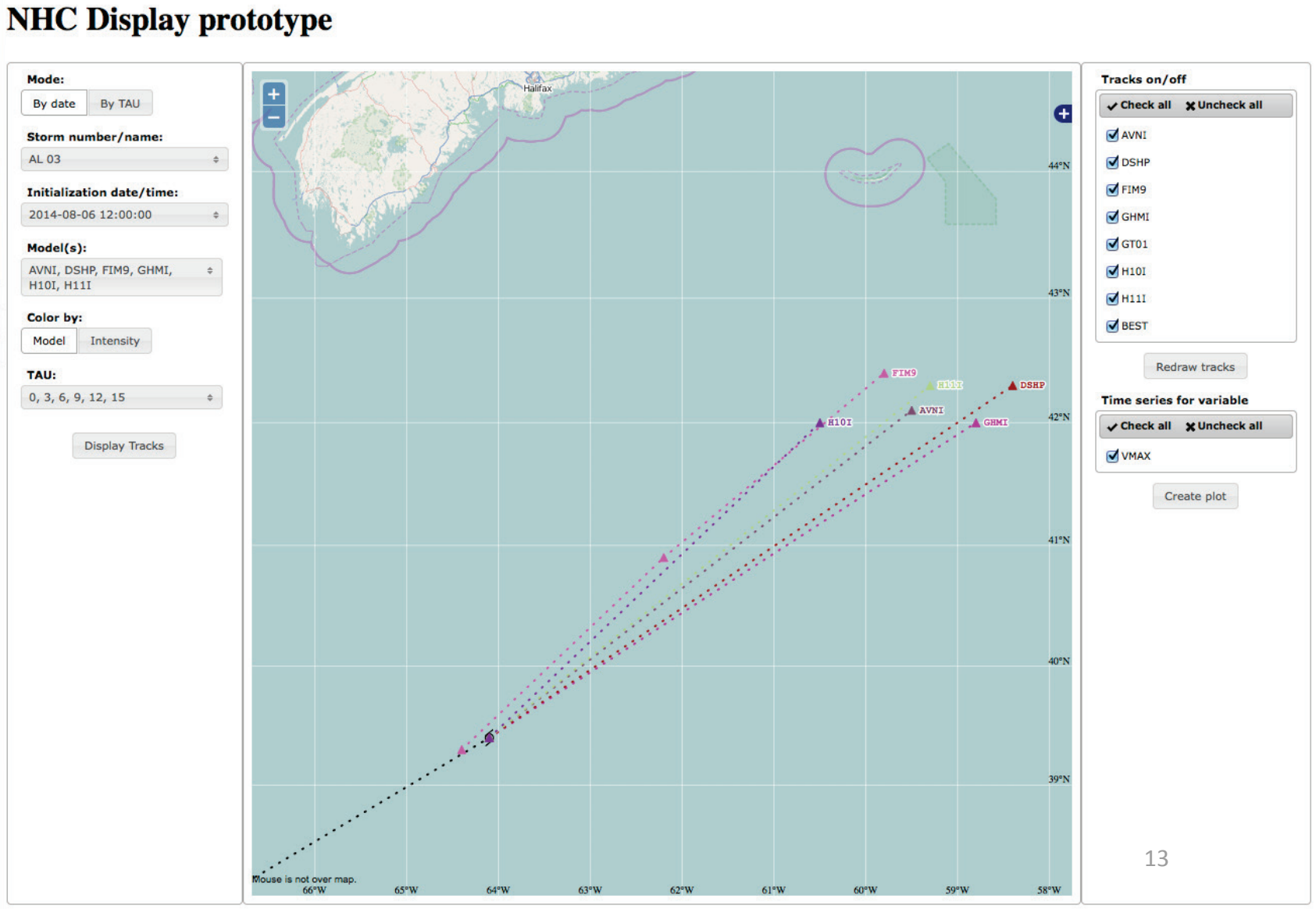


NHC Display and Diagnostic Development

- TCMT has been developing a Display and Diagnostic System for NHC
- Features:
  - Capability to color code forecast tracks by intensity (and maybe other diagnostic parameters)
  - Capability to zoom to specific region (e.g., centered on a storm)
  - Plot forecasts up to a specific lead time (e.g., 24, 48 h, etc. forecast).
  - The ability to select forecast model tracks to plot
  - The system is flexible and adaptable for future capabilities (additional databases, analysis tools, gridded products, etc.)
  - Open source that will be available to the HFIP community

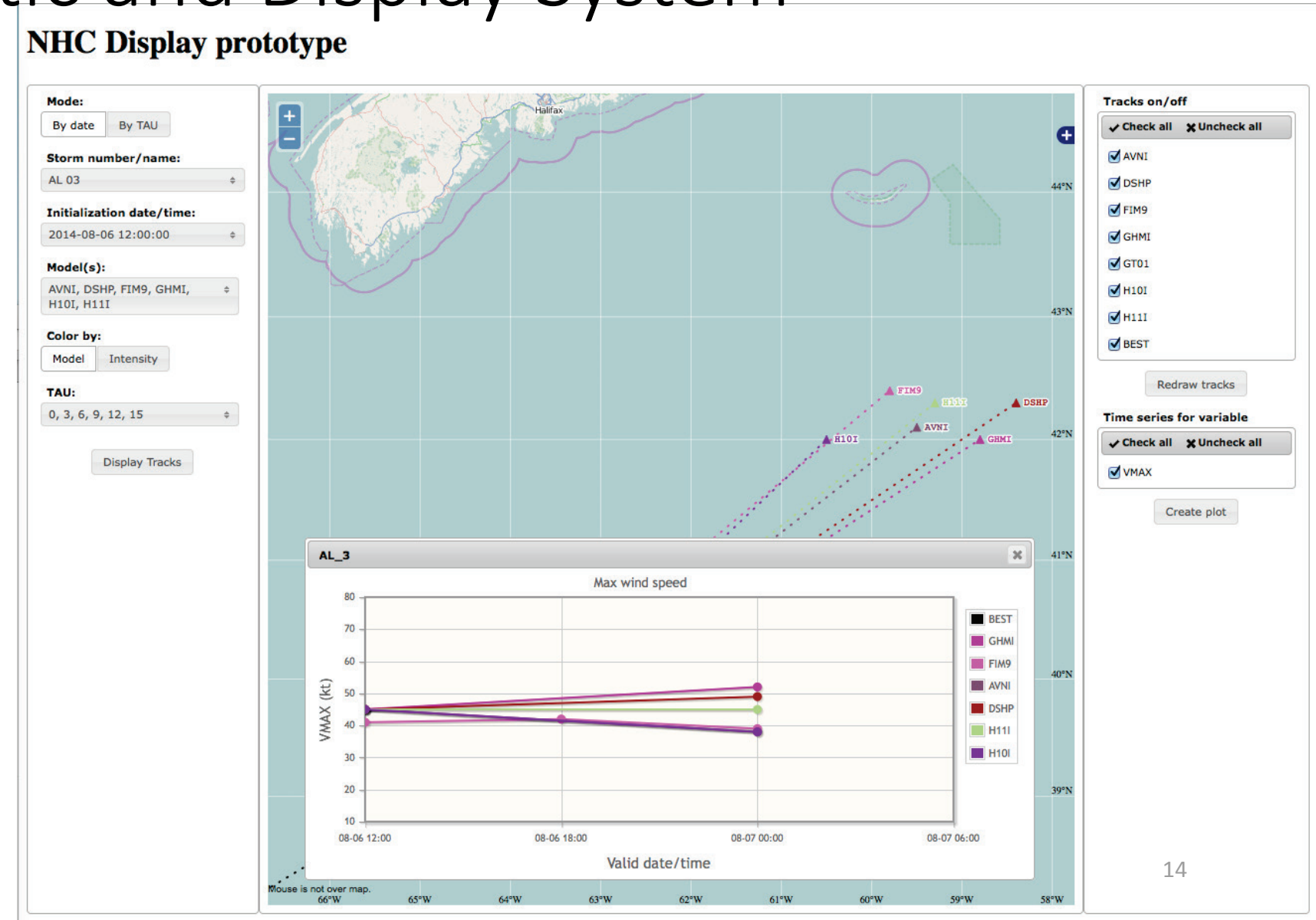
NHC Diagnostic and Display System

- Display is developed using:
  - OpenLayers Mapping tools
    - Platform independent, no license requirements
  - MySQL database
- Prototype will be implemented at NHC in the spring
- The plan is to implement an initial version on the HFIP website for community feedback



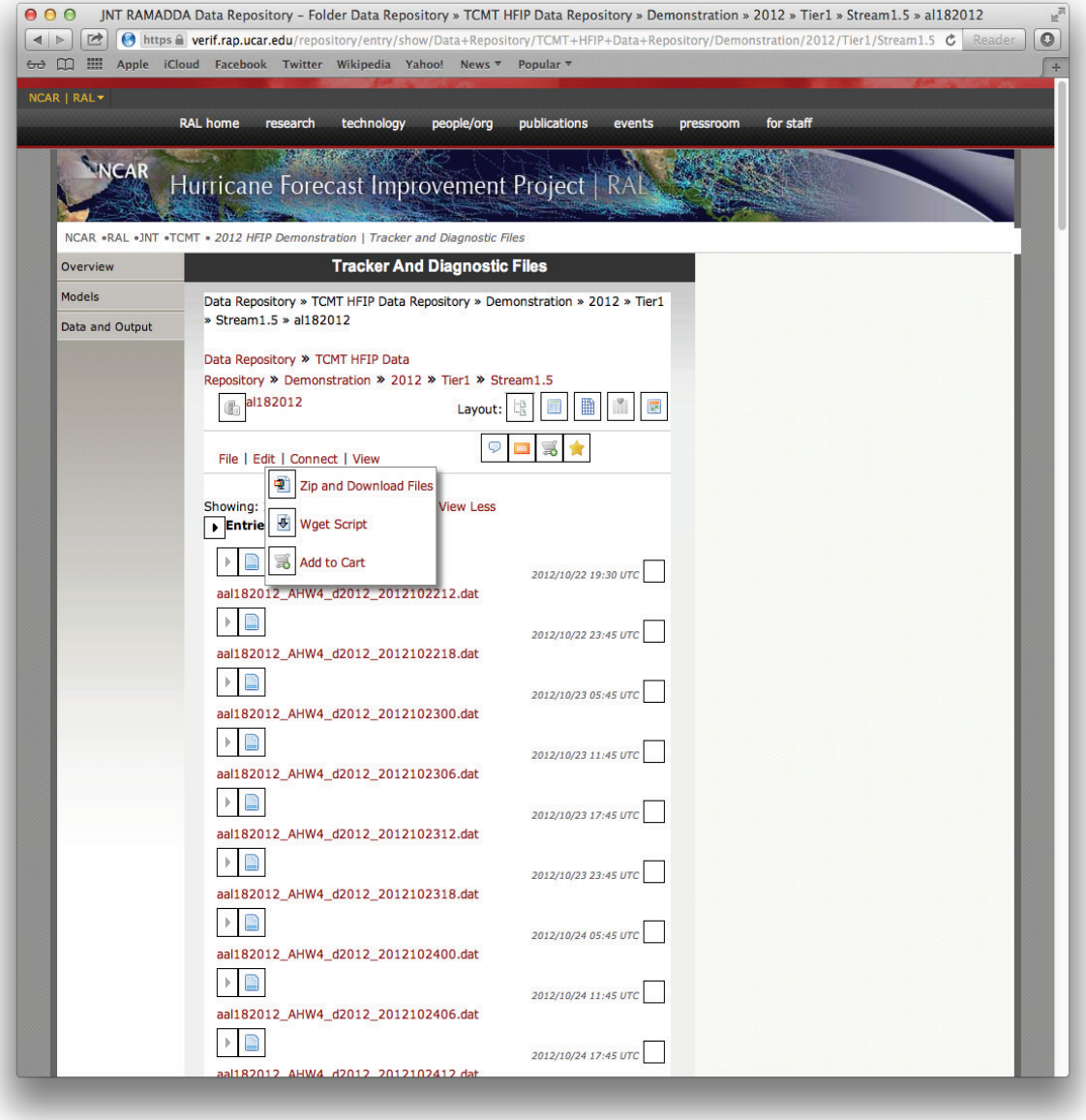
NHC Diagnostic and Display System

- Display both track and intensity information
- Diagnostic evaluation (intensity)
- Consensus forecasts
- Future capabilities
  - Gridded fields
    - Forecast products
    - Satellite observations



TCMT HFIP Data Service

- Available data include:
  - Tier 1 (Stream 1.5 and 2) and diagnostic files
  - 2011-2014 demonstration periods
  - 2010-2014 retrospective periods
- Password protected interface:
  - Username & password: ***hfipteam***
- Future Enhancements:
  - Access to available Tier 2 gridded products
  - Interactive graphical user interface (GUI) for improved accessibility, quick look graphics, download capabilities



- https://verif.rap.ucar.edu/repository

Preliminary Results from the 2014 Demo

- Christopher will highlight some of the initial results from the 2014 Demo

Three Streams

HFIP Research to Operations (R2O) Development Streams:

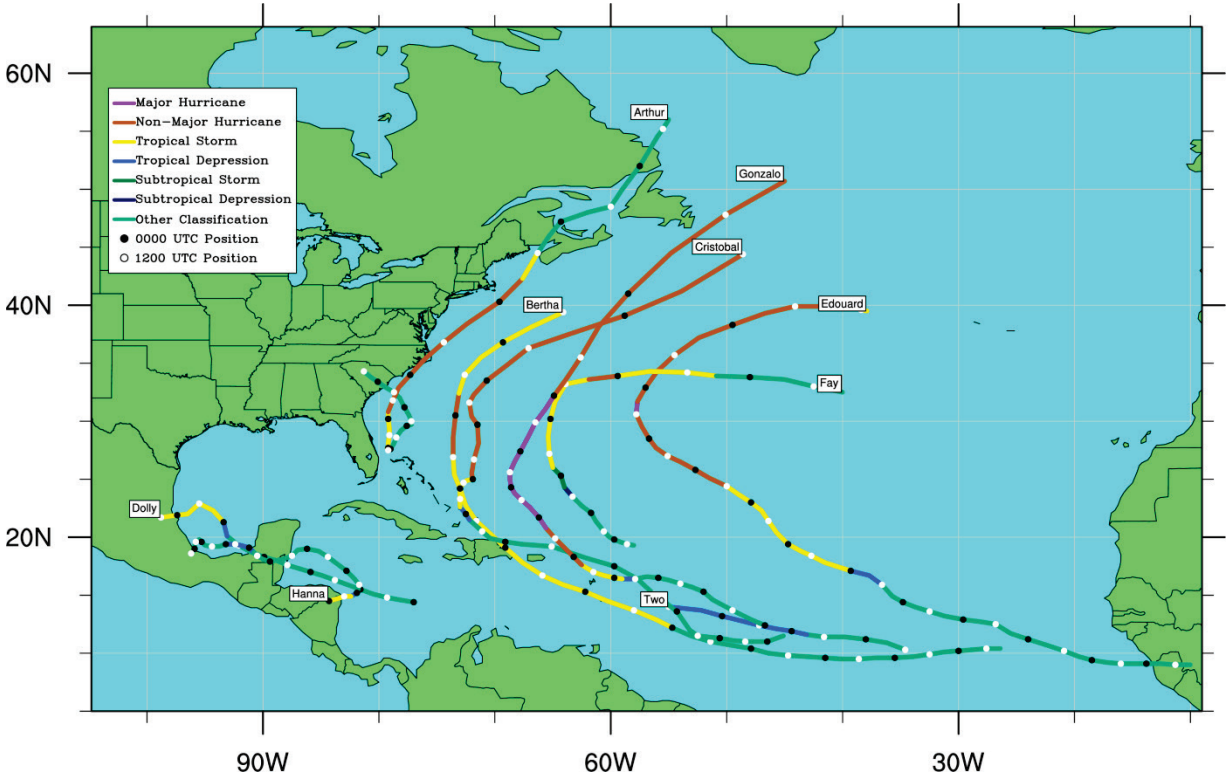
- ***Stream 2.0*** - Enhancements requiring multiple years of applied research, development and transition-to-operations work
- ***Stream 1.5*** - Experimental models or techniques that NHC, based on prior assessment, may access in real-time during a particular hurricane season
- ***Stream 1.0*** - Yearly upgrades to operational numerical weather prediction capabilities

Goals for objective evaluation

- Provide HFIP and modeling teams with in-depth statistical evaluations of the participating models/techniques
- Explore new approaches that provide added insight into the performance of the participants

2014 Hurricane Season

*working best tracks as of November

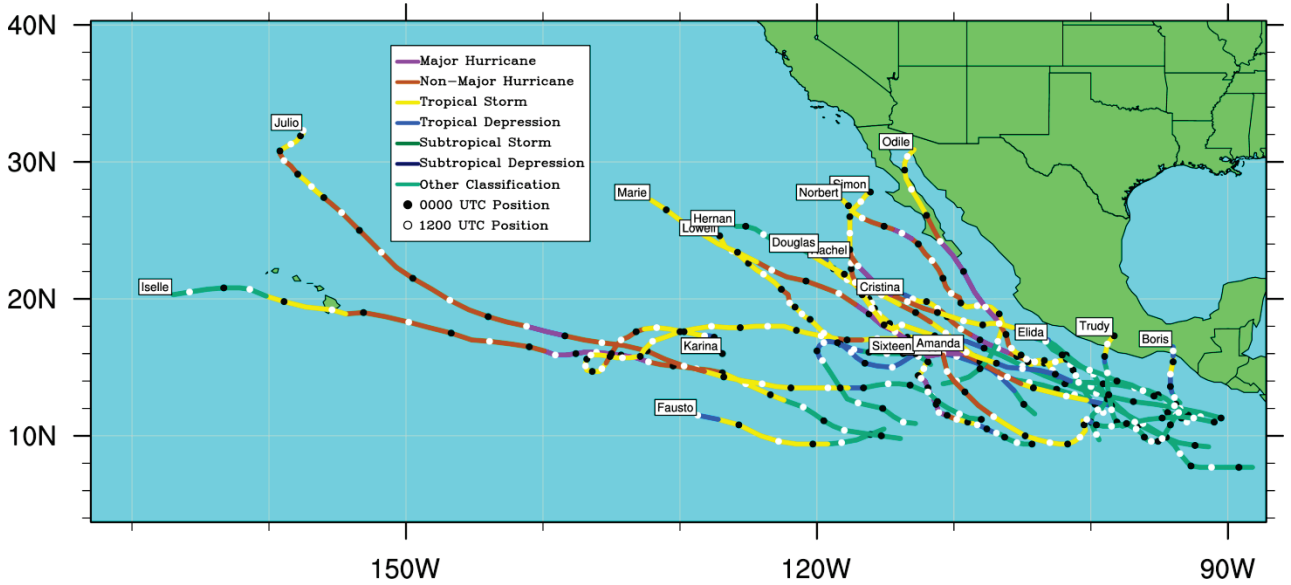


Atlantic Basin

9 storms

Eastern North Pacific Basin

21 storms



2014 Demo Participants

Stream 1.5

Organization	Early	Type	Configurations
NRL	***CXTI***	**Regional/Dynamic/Deterministic**	**1**
GFDL	***GPMI***	**Regional/Dynamic/Ensemble Mean**	**11**
UW - Madison	***UW4I***	**Regional/Dynamic/Deterministic**	**1**
CIRA	***SPC3***	**Statistical/Dynamic/Consensus**	**1**

2014 Demo Participants

Stream 2.0

Organization	Early	Type	Configurations
GSD	*FM9I*	Global/Dynamic/Deterministic	1
U Utah	*A3UI*	Regional/Dynamic/Deterministic	1
PSU	*APSI*	Regional/Dynamic/Deterministic	1
	PnnI, where nn=01-10	Regional/Dynamic/Ensemble	10
GFDL	*GTMI*	Regional/Dynamic/Ensemble	11
HRD	*H3WI*	Regional/Dynamic/Deterministic	1
	HECI		1
	HEDI		1
EMC	*HWMI*	Regional/Dynamic/Ensemble	22
UW - Madison	*UWMI*	Regional/Dynamic/Ensemble	9
FSU	*MMSI*	Multimodel Consensus	1

Near End-of-Demo Data Snapshot

Tier 1 data inventory adjusted per model based on following expectations:

d2014	A3UT	APSU	CTCX	FIM9	GPMN	GPNN	GTMN	GTNN	H3HW	HECT	HEDS	HWMN	HWNN	MMSE	PSNN	SPC3	UWMN	UWN4	UWNN
Runs/day:	1	4	4	4	4	4	4	4	4	4	4	4	4	2		4	4	4	4
Basin(s):	AL	AL	AL EP	AL EP	AL EP	AL EP	AL EP	AL EP	AL EP	AL EP	AL EP	AL EP	AL EP	AL		AL EP	AL EP	AL EP	AL EP
TY Stage:	limited to TD, TS, HU, SD, and SS development levels																		
Time period:	no restrictions																		
Total possible cases:	based only on NHC best track files excluding invests and CP storms																		

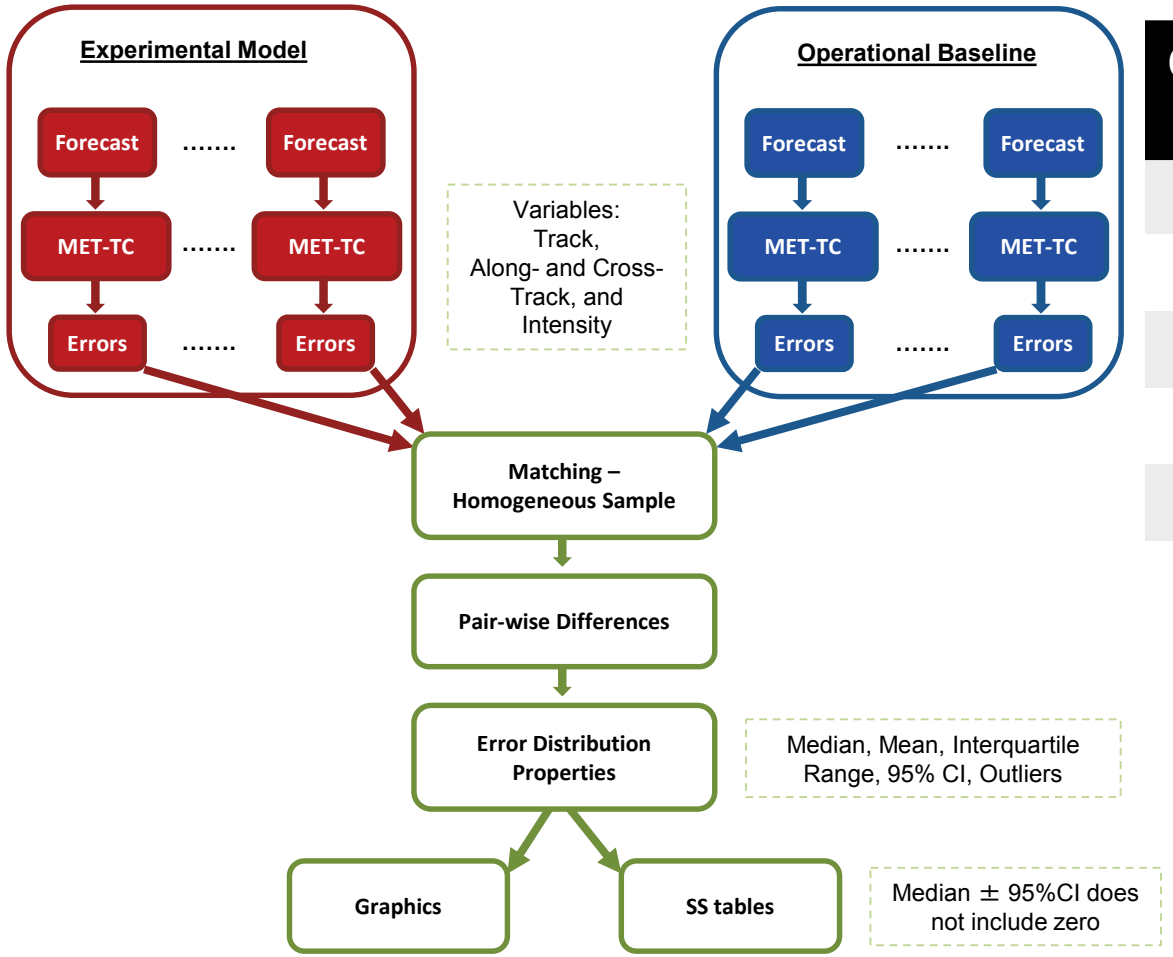
d2014	A3UT	APSU	CTCX	FIM9	GPMN	GPNN	GTMN	GTNN	H3HW	HECT	HEDS	HWMN	HWNN	MMSE	PSNN	SPC3	UWMN	UWN4	UWNN
al012014	---	---	27	19	17	17	17	17	15	---	---	19	14	1	---	11	5	26	5
al022014	---	---	7	9	9	9	9	9	5	---	---	---	---	---	---	3	9	9	9
al032014	2	---	35	36	37	37	37	37	19	---	---	31	35	11	---	30	34	37	34
al042014	6	31	31	40	39	39	39	39	16	---	---	38	38	11	3	36	40	40	40
al052014	3	16	18	17	12	12	12	12	10	---	---	4	4	3	2	11	18	18	18
al062014	8	43	48	47	45	45	45	45	26	---	---	44	45	16	2	38	48	48	48
al072014	---	14	16	19	17	17	17	17	9	---	---	---	---	6	2	15	17	17	17
al082014	---	30	33	36	37	37	37	37	19	---	---	27	28	13	---	35	38	37	38
al092014	---	21	15	12	13	13	13	13	8	---	---	2	2	---	3	11	13	13	13
ep012014	n/a	n/a	---	---	29	29	29	29	12	---	---	---	---	n/a	---	---	---	---	---
ep022014	n/a	n/a	7	4	8	8	8	8	4	---	---	---	---	n/a	---	---	---	---	---
ep032014	n/a	n/a	26	7	16	18	18	18	10	---	---	---	---	n/a	---	---	---	---	---
ep042014	n/a	n/a	27	9	25	25	25	25	13	---	---	---	---	n/a	---	8	5	27	5
ep052014	n/a	n/a	8	10	8	8	7	8	4	---	---	---	---	n/a	---	1	---	11	---
ep062014	n/a	n/a	7	9	7	7	7	7	5	---	---	---	---	n/a	---	2	9	9	9
ep072014	n/a	n/a	65	88	71	71	71	71	31	---	---	---	---	n/a	---	26	47	53	48
ep082014	n/a	n/a	15	16	15	15	15	15	9	---	---	---	---	n/a	---	5	16	16	16
ep092014	n/a	n/a	34	39	36	36	36	36	12	---	---	---	---	n/a	---	32	35	36	35
ep102014	n/a	n/a	47	55	53	53	53	53	17	---	---	---	---	n/a	---	48	55	56	55
ep112014	n/a	n/a	51	58	62	62	62	62	23	---	---	---	---	n/a	---	60	65	65	65
ep122014	n/a	n/a	20	29	28	28	28	28	16	---	---	---	---	n/a	---	28	31	31	31
ep132014	n/a	n/a	28	36	35	35	35	35	14	---	---	---	---	n/a	---	32	38	38	38
ep142014	n/a	n/a	28	27	24	24	24	24	14	---	---	10	12	n/a	---	24	29	27	29
ep152014	n/a	n/a	35	36	34	34	34	34	20	---	---	---	---	n/a	---	29	35	34	35
ep162014	n/a	n/a	21	21	20	20	20	20	11	---	---	---	---	n/a	---	15	21	20	21
ep172014	n/a	n/a	31	31	30	30	30	30	17	---	---	---	---	n/a	---	24	31	30	31
ep182014	n/a	n/a	36	38	35	35	35	35	19	---	---	---	---	n/a	---	31	37	35	37
ep192014	n/a	n/a	35	36	32	32	32	32	18	---	---	---	---	n/a	---	31	35	34	35
ep202014	n/a	n/a	18	18	9	9	9	9	9	---	---	---	---	n/a	---	8	18	18	18
ep212014	n/a	n/a	18	19	15	15	15	15	9	---	---	---	---	n/a	---	15	---	18	18
TotalNum	*19*	*155*	*787*	*841*	*818*	*820*	*819*	*820*	*416*	*0*	*0*	*175*	*178*	*61*	*12*	*609*	*729*	*803*	*748*

Approx. Sample Range

850 - 0

Verification Setup

Methodology



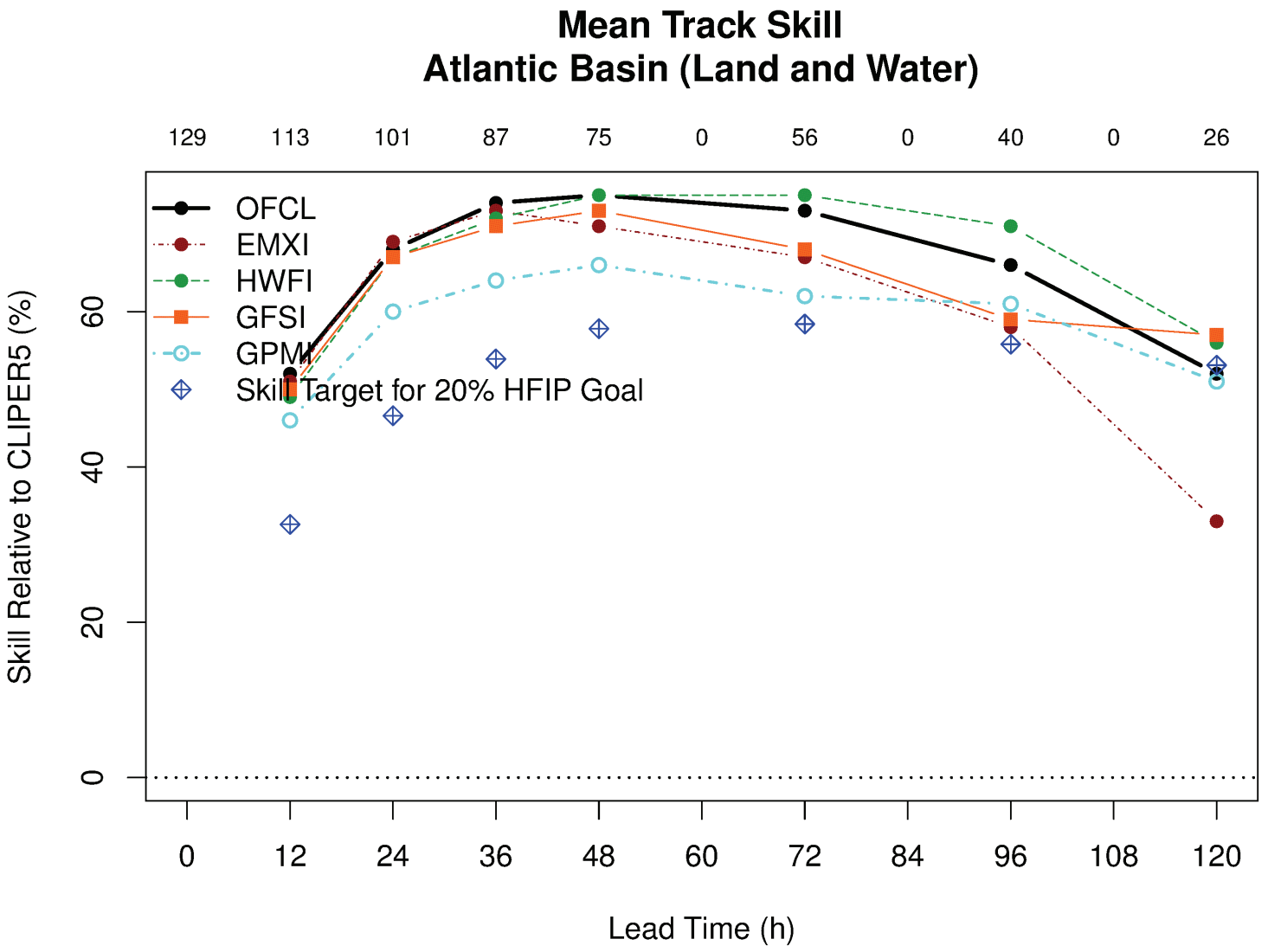
Baseline Comparisons

Operational Baseline	Track	Intensity
DSHP		✓
EMXI	✓	
GFSI	✓	
HWFI	✓	✓
LGEM		✓

CASES:
AL and EP storms during Demo (Aug 1 to Oct 31)

Preliminary Results
&
Working Best Tracks

Stream 1.5 Skill (AL Track)

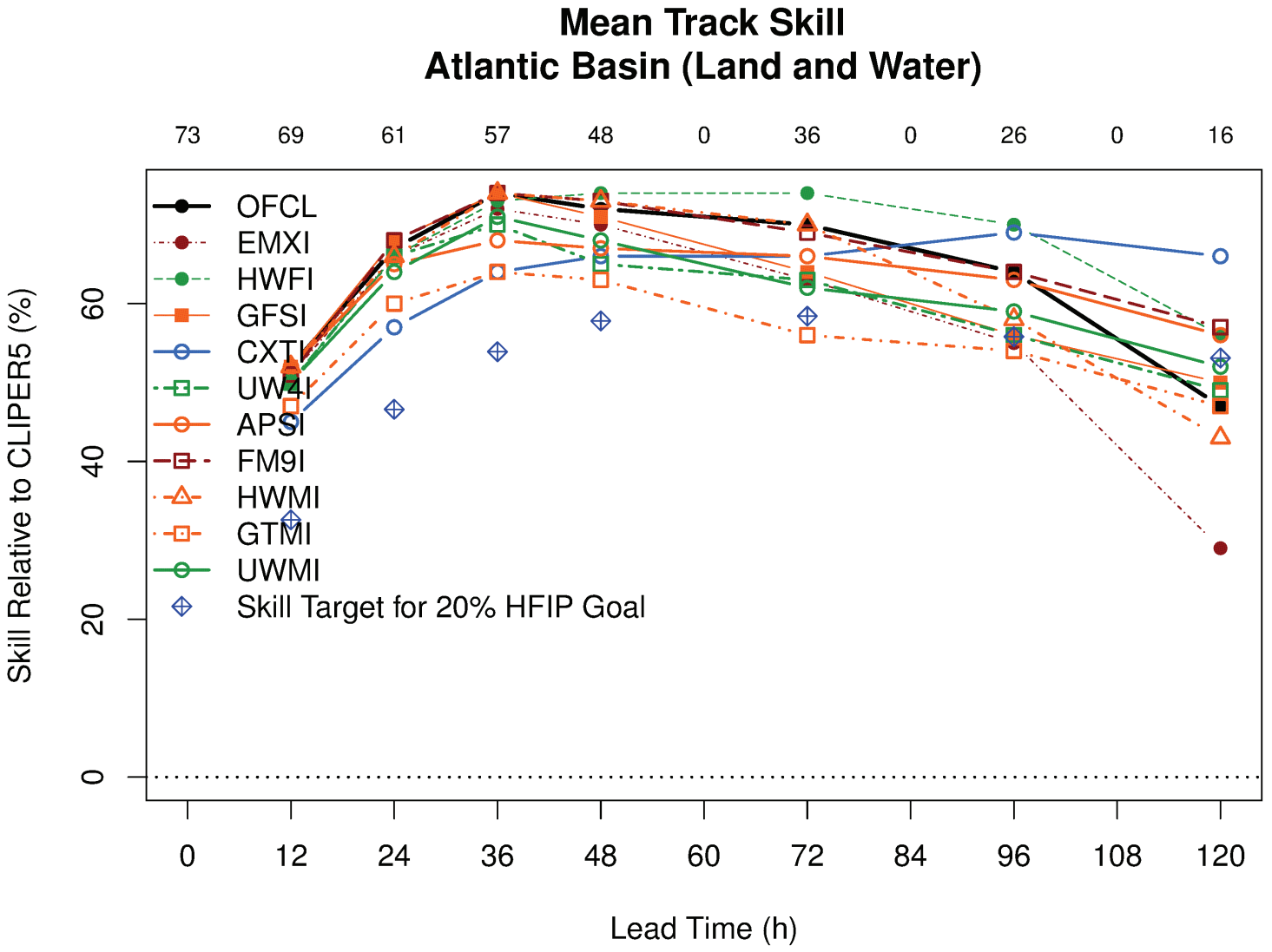


GPMI lowest skill for most lead times, but still higher than HFIP 5-yr skill goal

---

HWRF highest skill among operational models

Stream 2.0 Skill (AL Track)



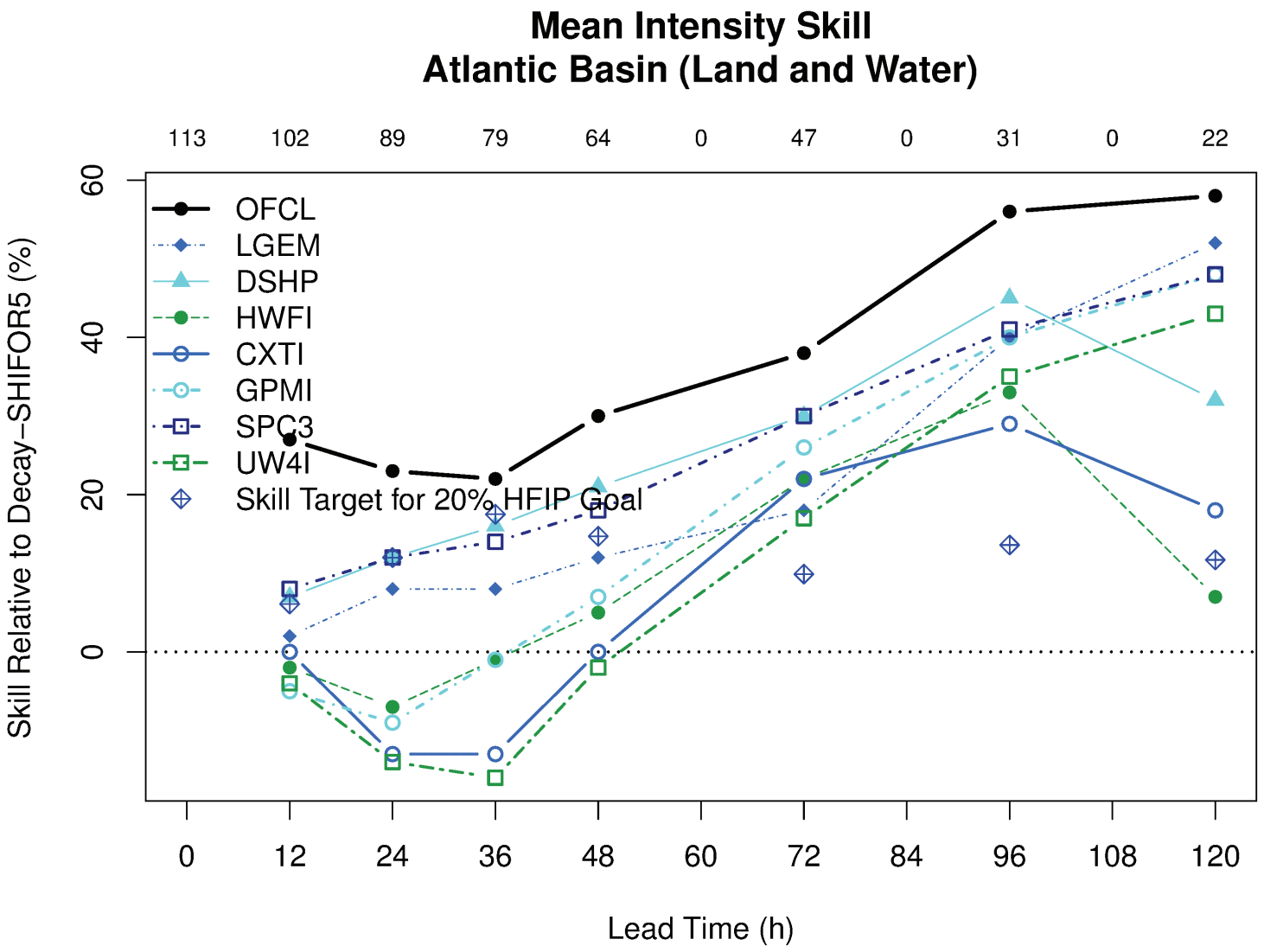
Smaller sample size

HWRF and FIM configurations show highest skill most lead times

---

HFIP 5-yr skill goal surpassed for most lead times

Stream 1.5 Skill (AL Intensity)



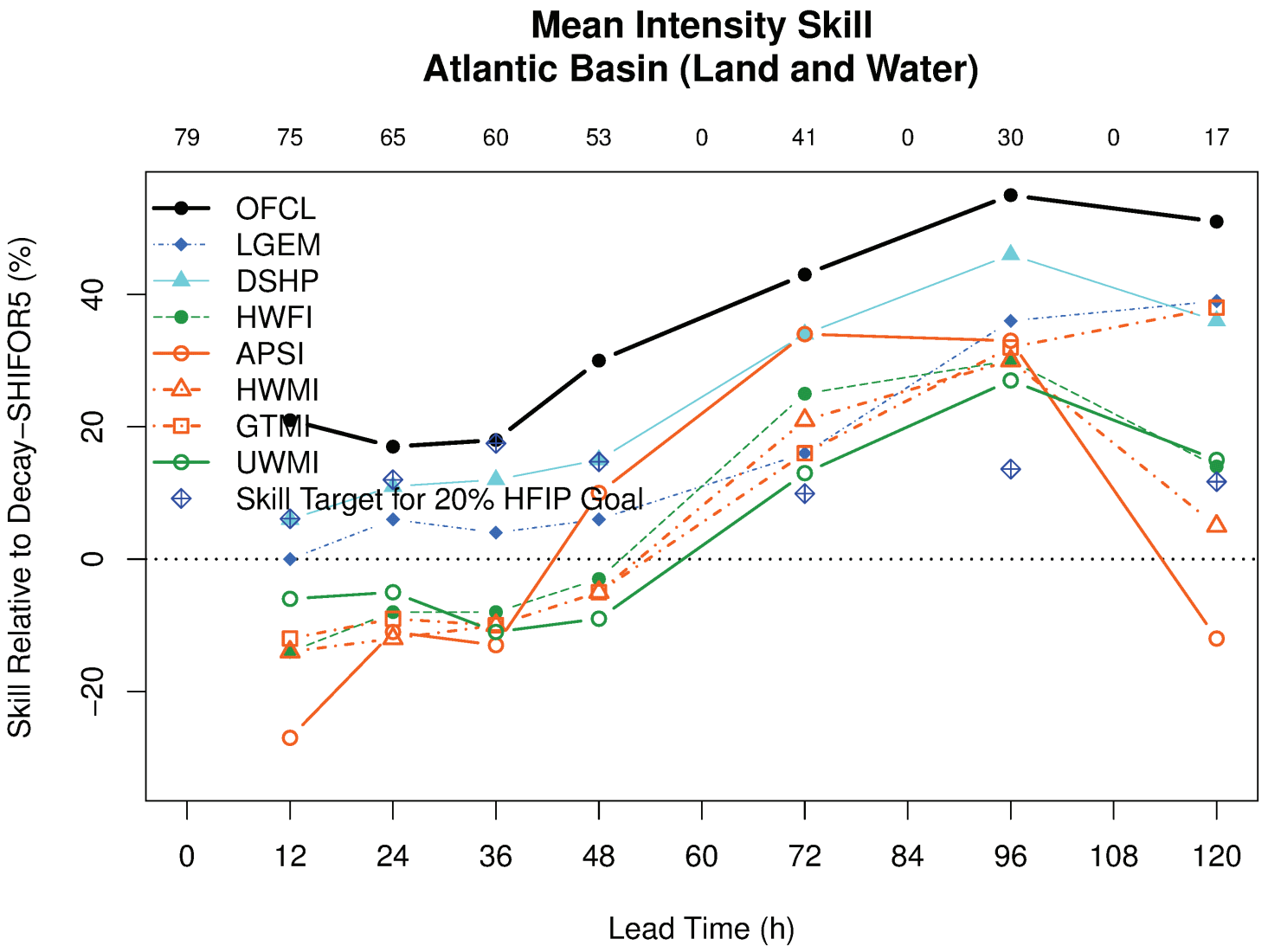
Statistical-dynamical configurations show highest skill including SPC3

Dynamical models transition from (-) to (+) skill with lead time

---

CXTI and UW4I show lowest skill most lead times

Stream 2.0 Skill (AL Intensity)

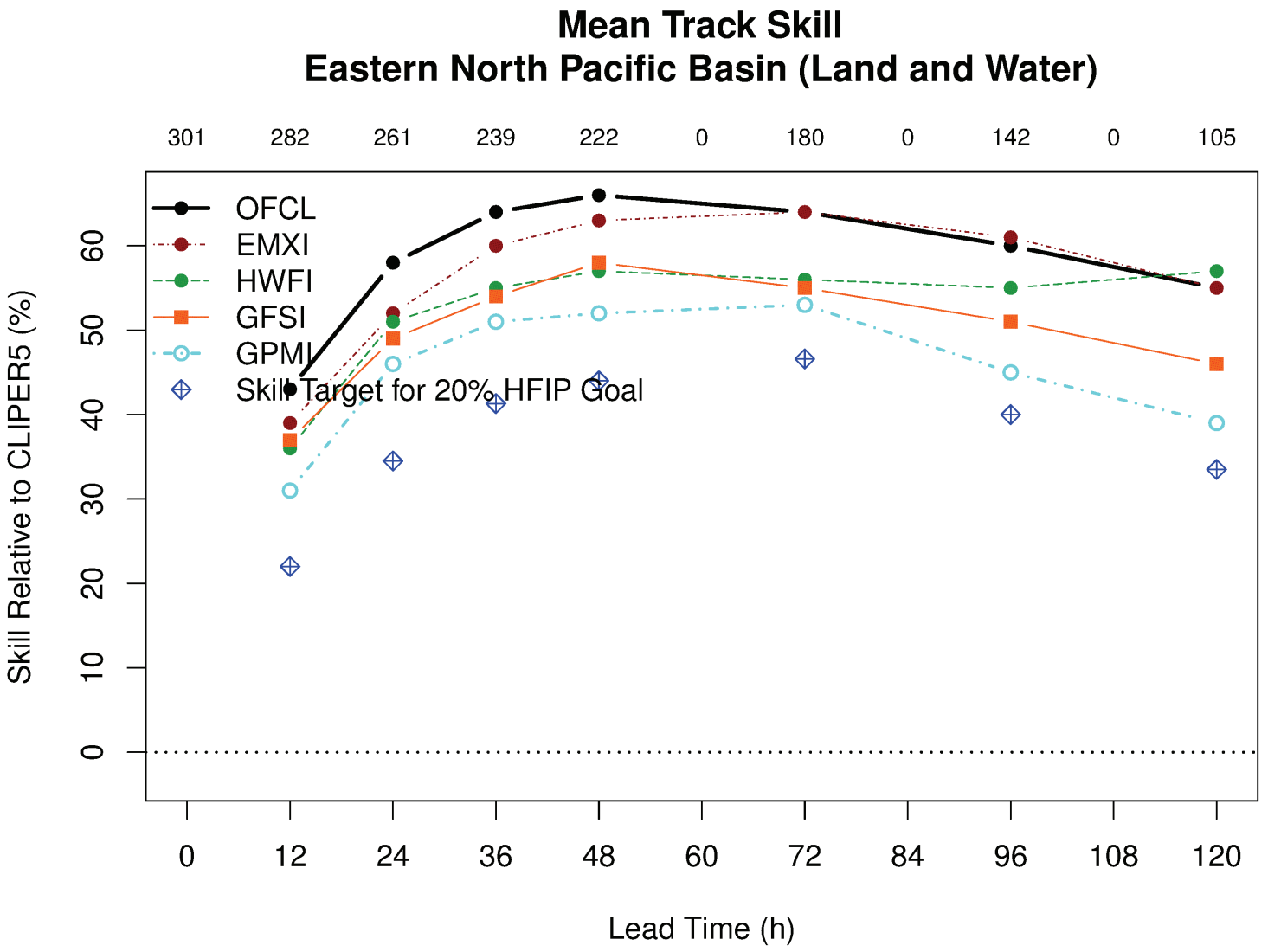


Smaller sample size

Decay SHIPS shows highest model skill

APSI skill gain is largest through 72 h

Stream 1.5 Skill (EP Track)

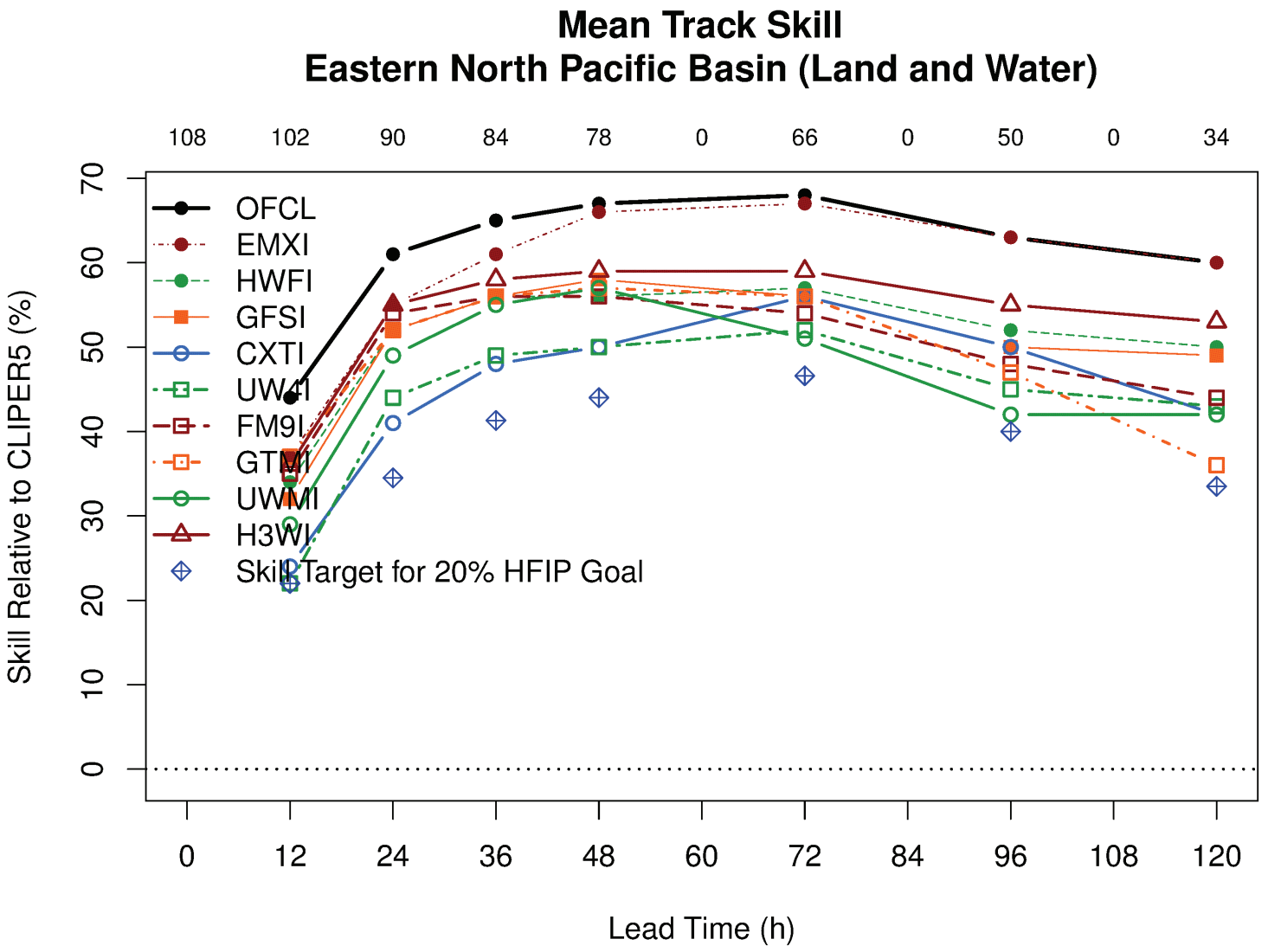


GPMI lowest skill for most lead times, but still higher than HFIP 5-yr skill goal

---

ECMWF highest skill among operational models

Stream 2.0 Skill (EP Track)



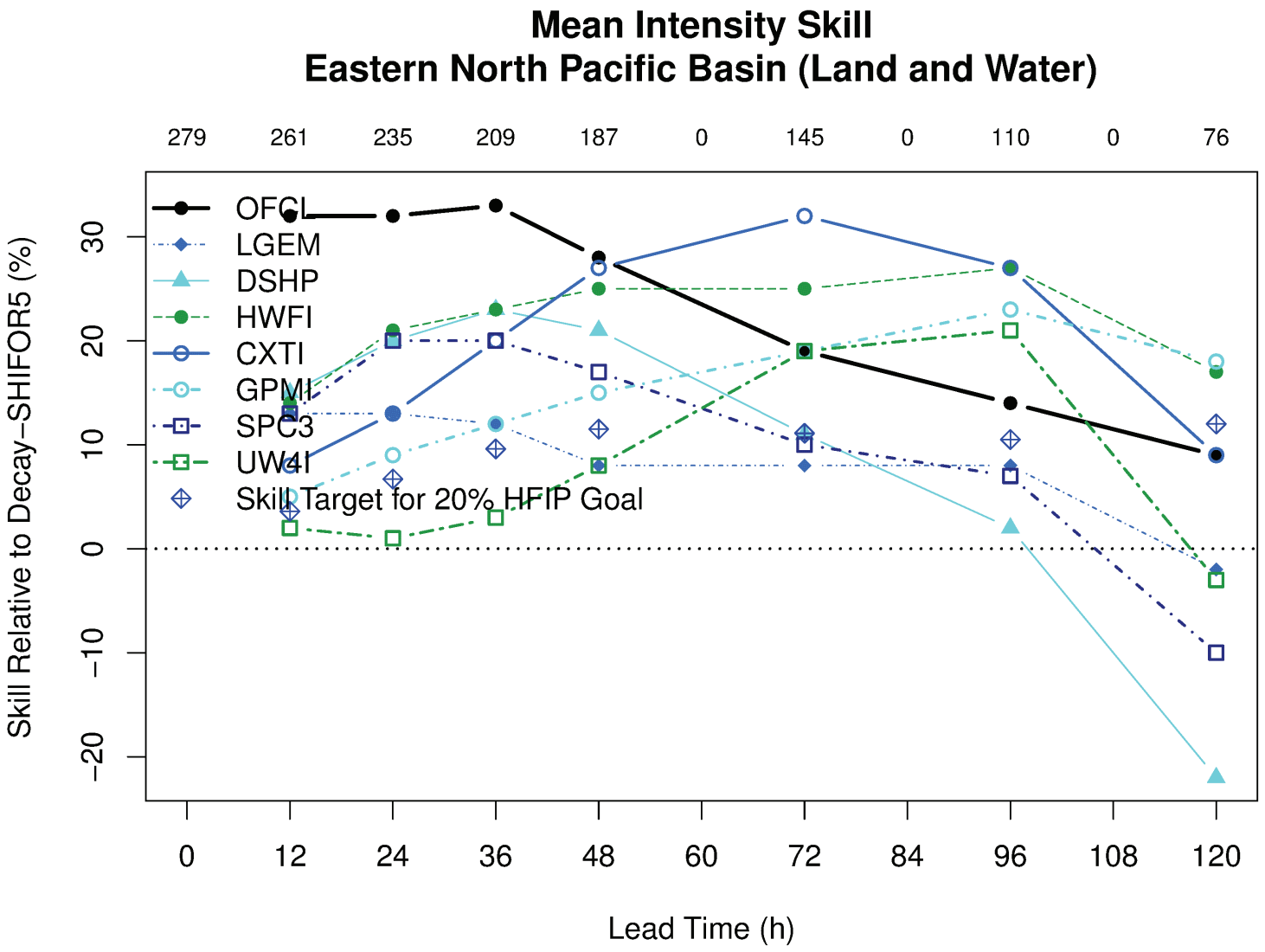
Smaller sample size

HWRF Basin Scale shows highest skill after ECMWF

UW4I and CXTI have lowest skill

HFIP 5-yr skill goal surpassed

Stream 1.5 Skill (EP Intensity)



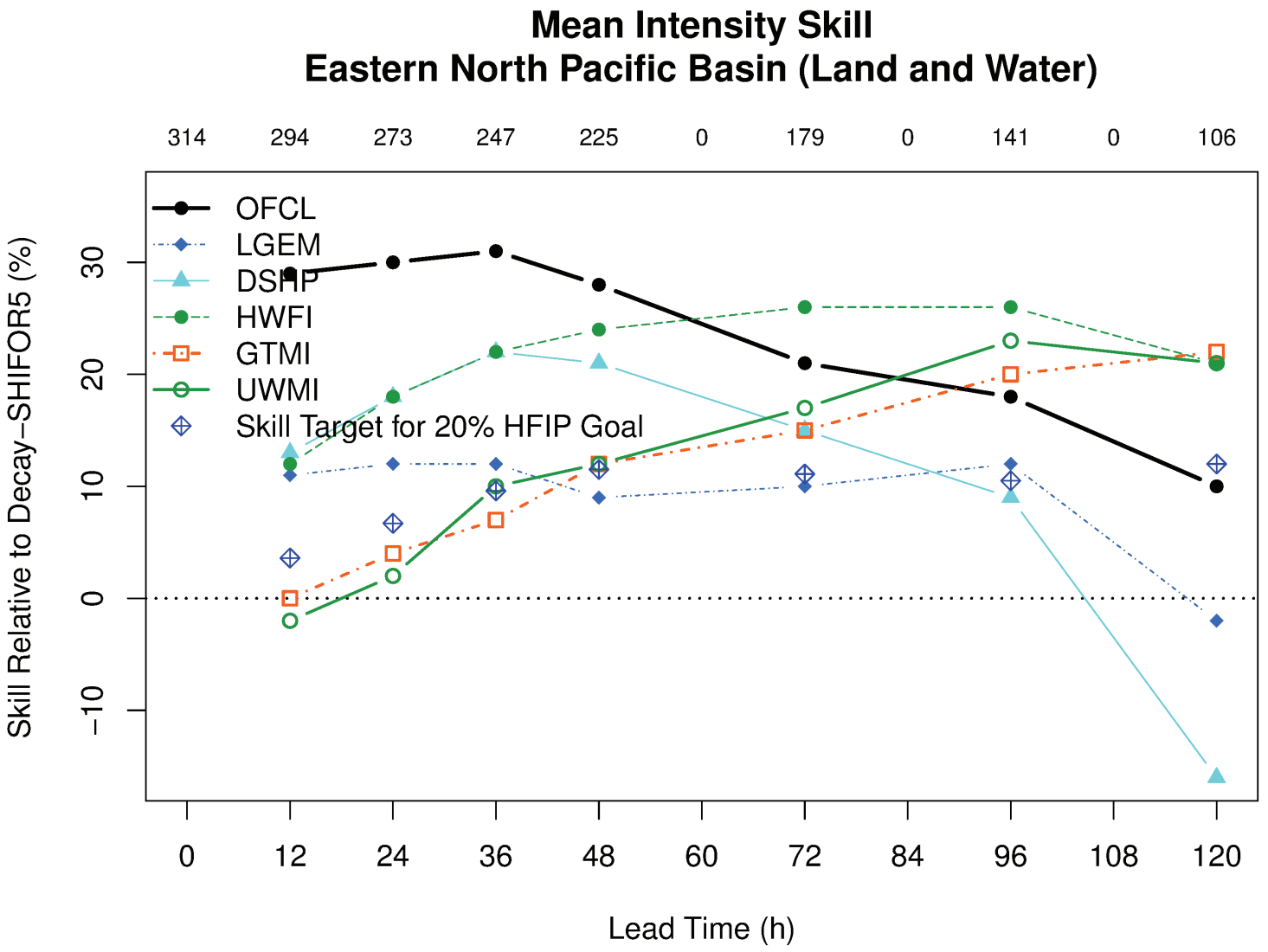
CXTI and HWFI show highest skill

Statistical-dynamical configurations generally lose skill with lead time

---

HFIP 5-yr skill goal met intermittently

Stream 2.0 Skill (EP Intensity)



GTMI and UWMI skill slowly increases with lead time

HWRF displays highest model skill

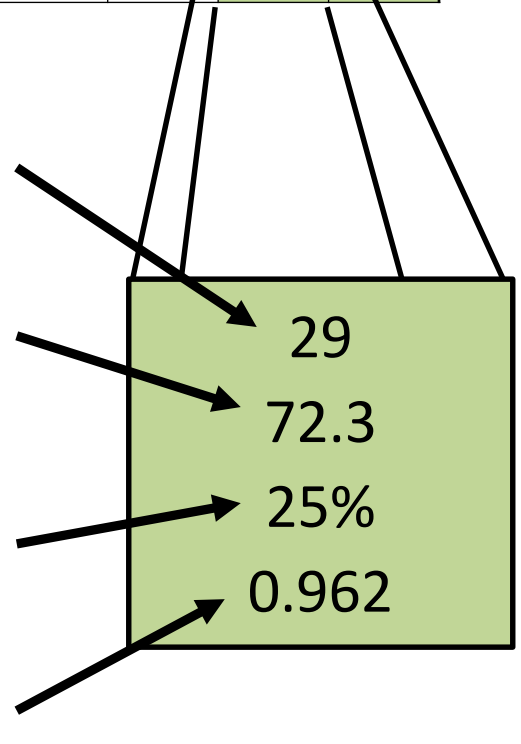
Statistical Significance (SS) Pairwise Differences

Summary Tables

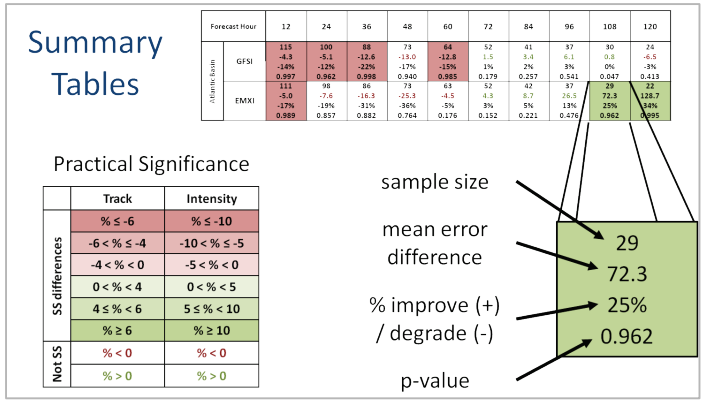
Forecast Hour		12	24	36	48	60	72	84	96	108	120
Atlantic Basin	GFSI	115 -4.3 -14% 0.997	100 -5.1 -12% 0.962	88 -12.6 -22% 0.998	73 -13.0 -17% 0.940	64 -12.8 -15% 0.985	52 1.5 1% 0.179	41 3.4 2% 0.257	37 6.1 3% 0.541	30 0.8 0% 0.047	24 -6.5 -3% 0.413
	EMXI	111 -5.0 -17% 0.989	98 -7.6 -19% 0.857	86 -16.3 -31% 0.882	73 -25.3 -36% 0.764	63 -4.5 -5% 0.176	52 4.3 3% 0.152	42 8.7 5% 0.221	37 26.5 13% 0.476	29 72.3 25% 0.962	22 128.7 34% 0.995

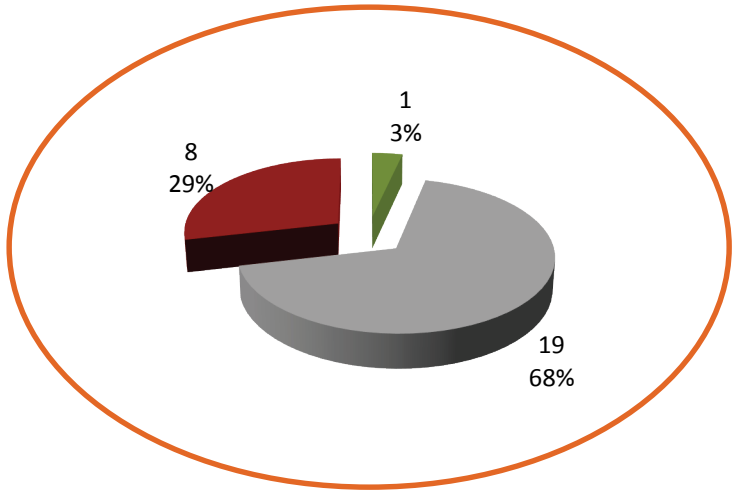
Practical Significance

	Track	Intensity
SS differences	% ≤ -6	% ≤ -10
	-6 < % ≤ -4	-10 < % ≤ -5
	-4 < % < 0	-5 < % < 0
	0 < % < 4	0 < % < 5
	4 ≤ % < 6	5 ≤ % < 10
	% ≥ 6	% ≥ 10
Not SS	% < 0	% < 0
	% > 0	% > 0



SS Table-to-High Level Chart





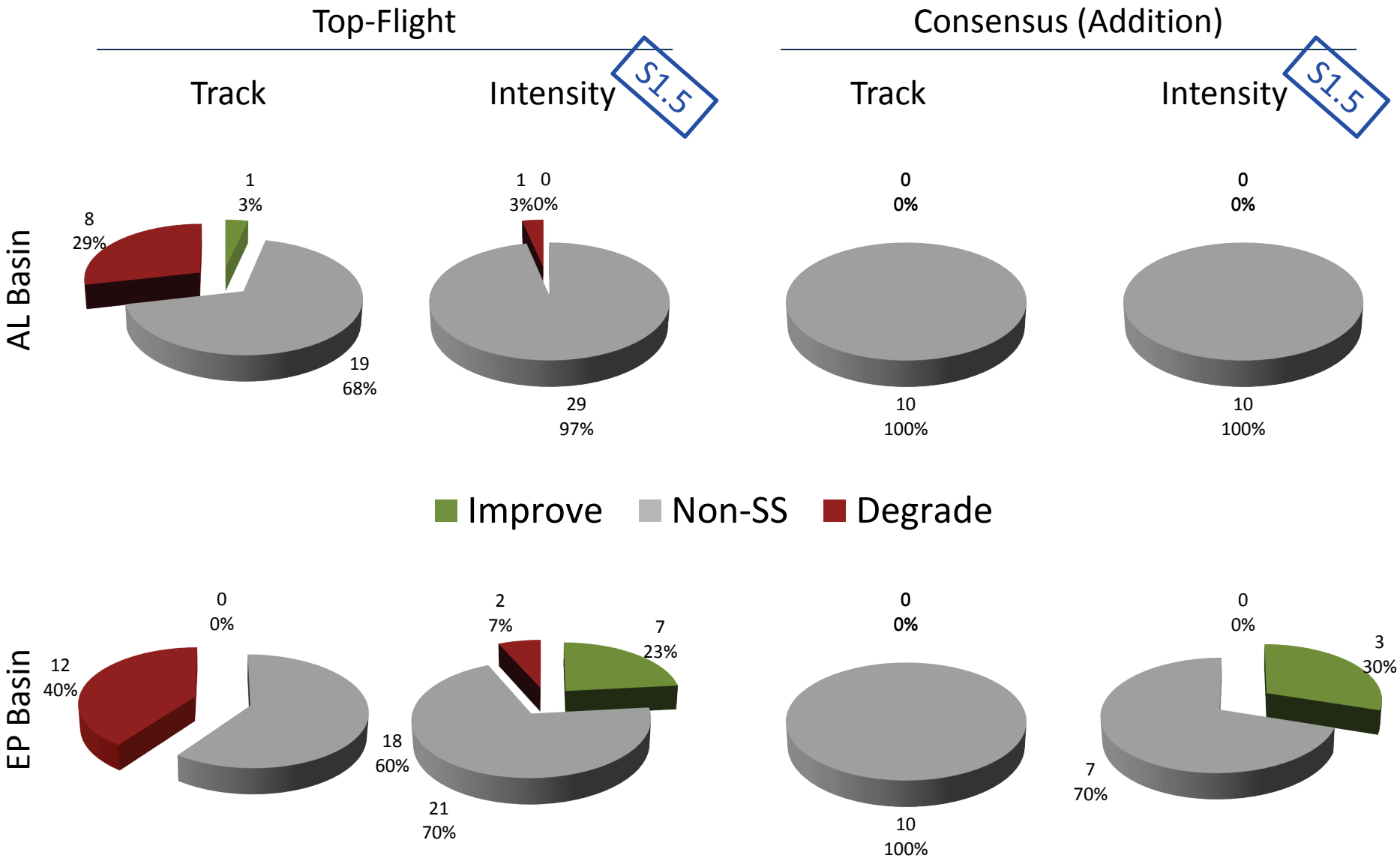
Detailed Info

Summary Info

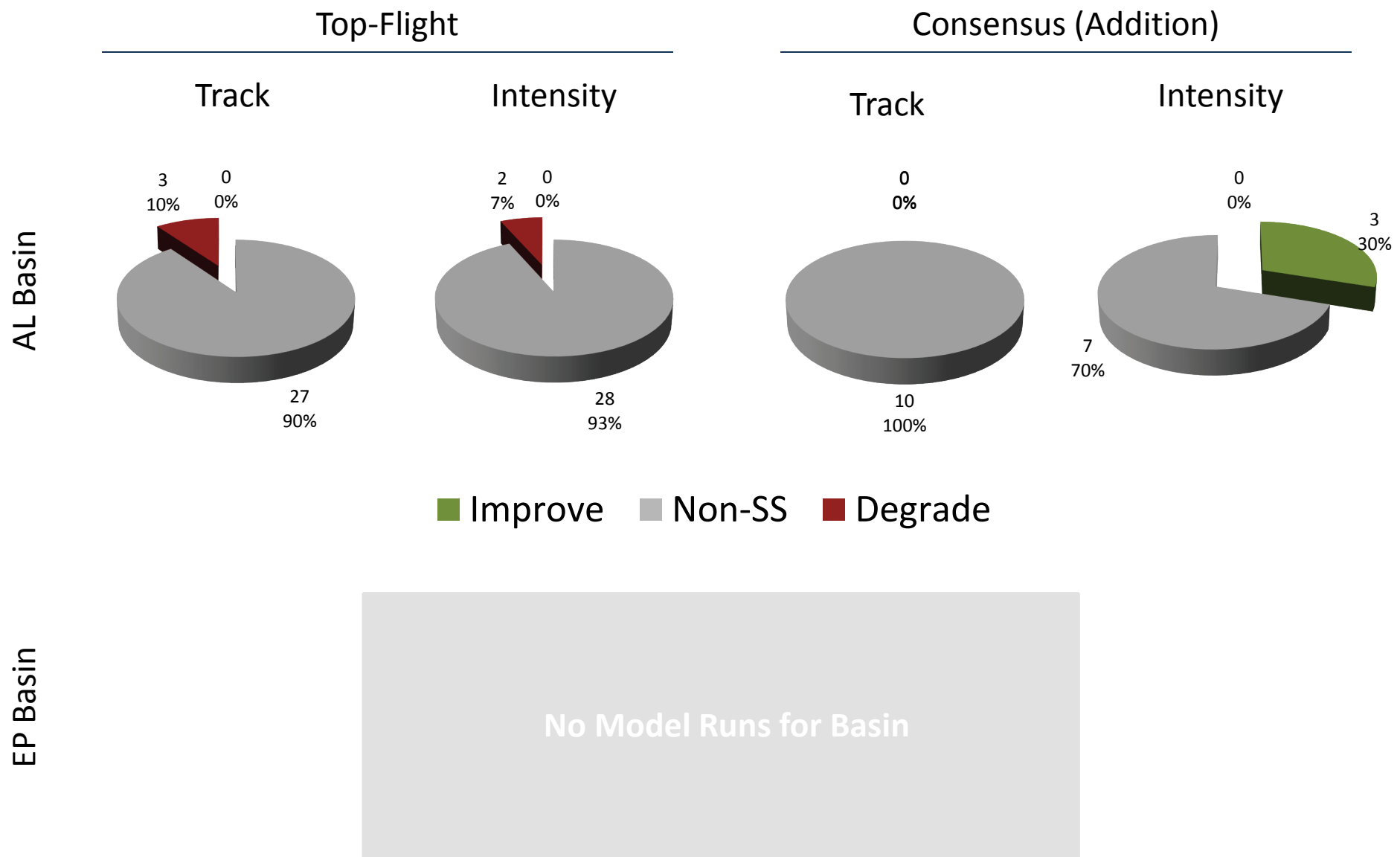
Improve Non-SS Degrade

Focus on cumulative fraction of lead times for which HFIP experimental model pairwise difference *improves* upon, is *not statistically significantly* different from, or *degrades* upon operational baseline model

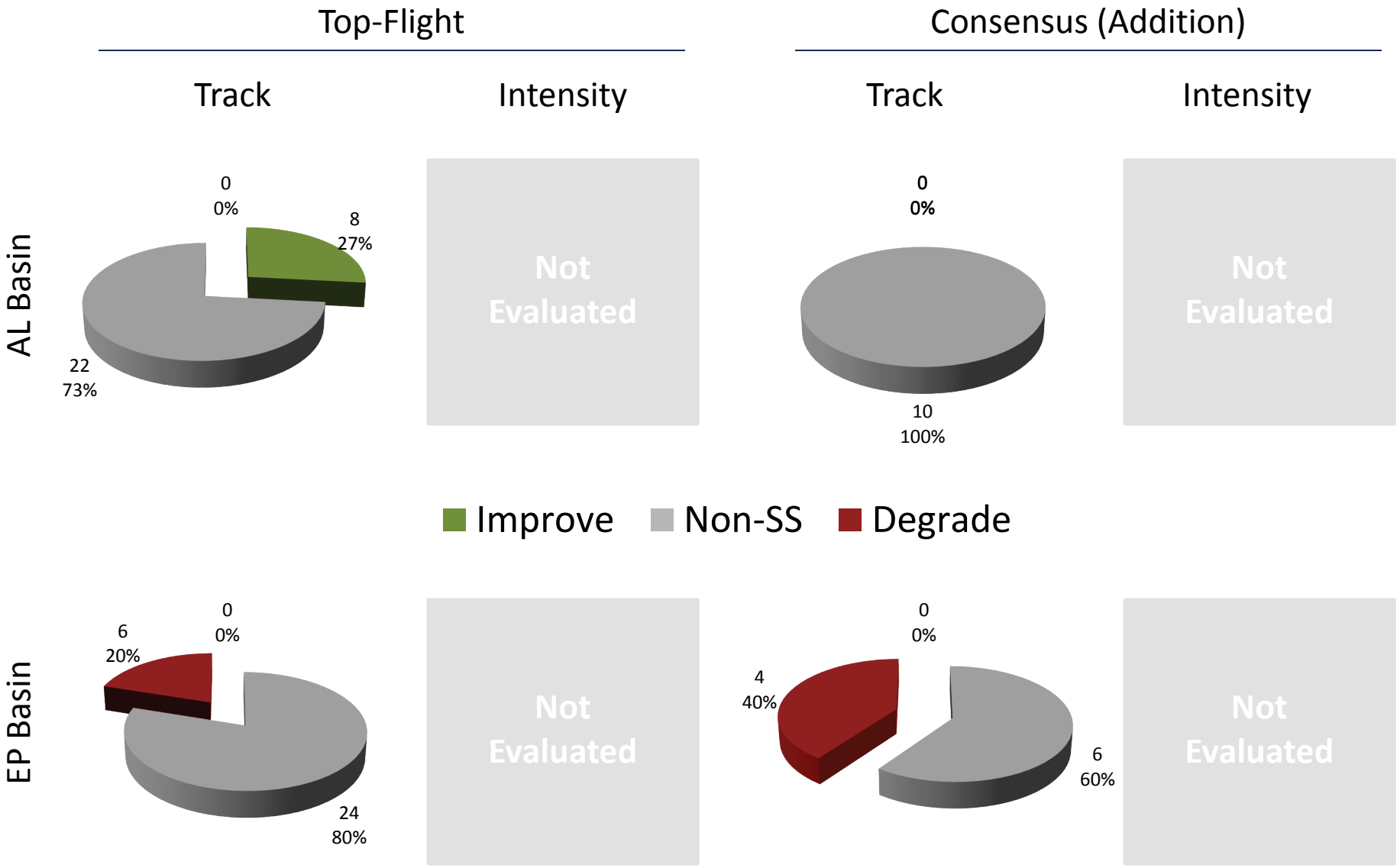
5 km COAMPS-TC



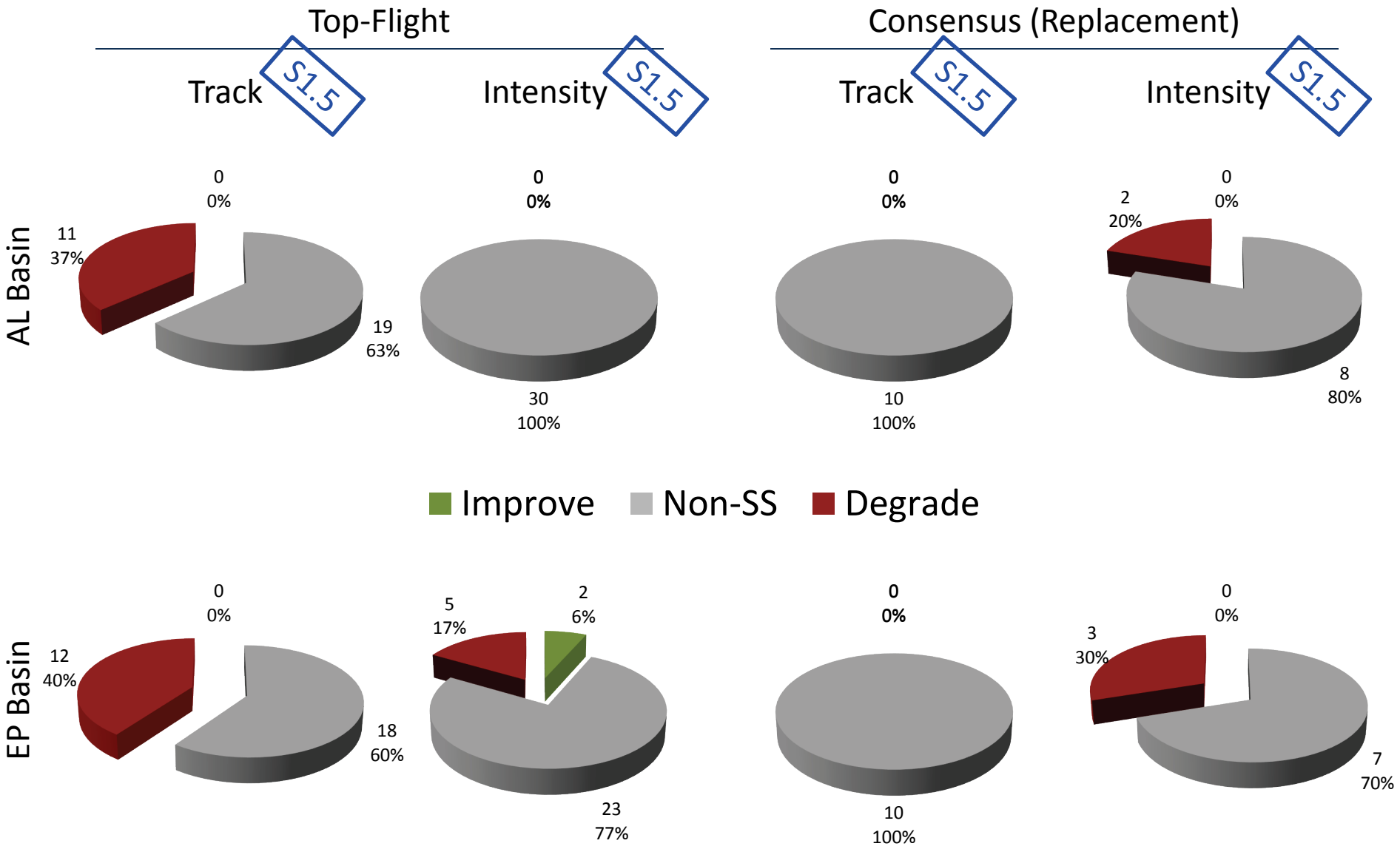
3 km WRF/ARW/PSU (All Recon)



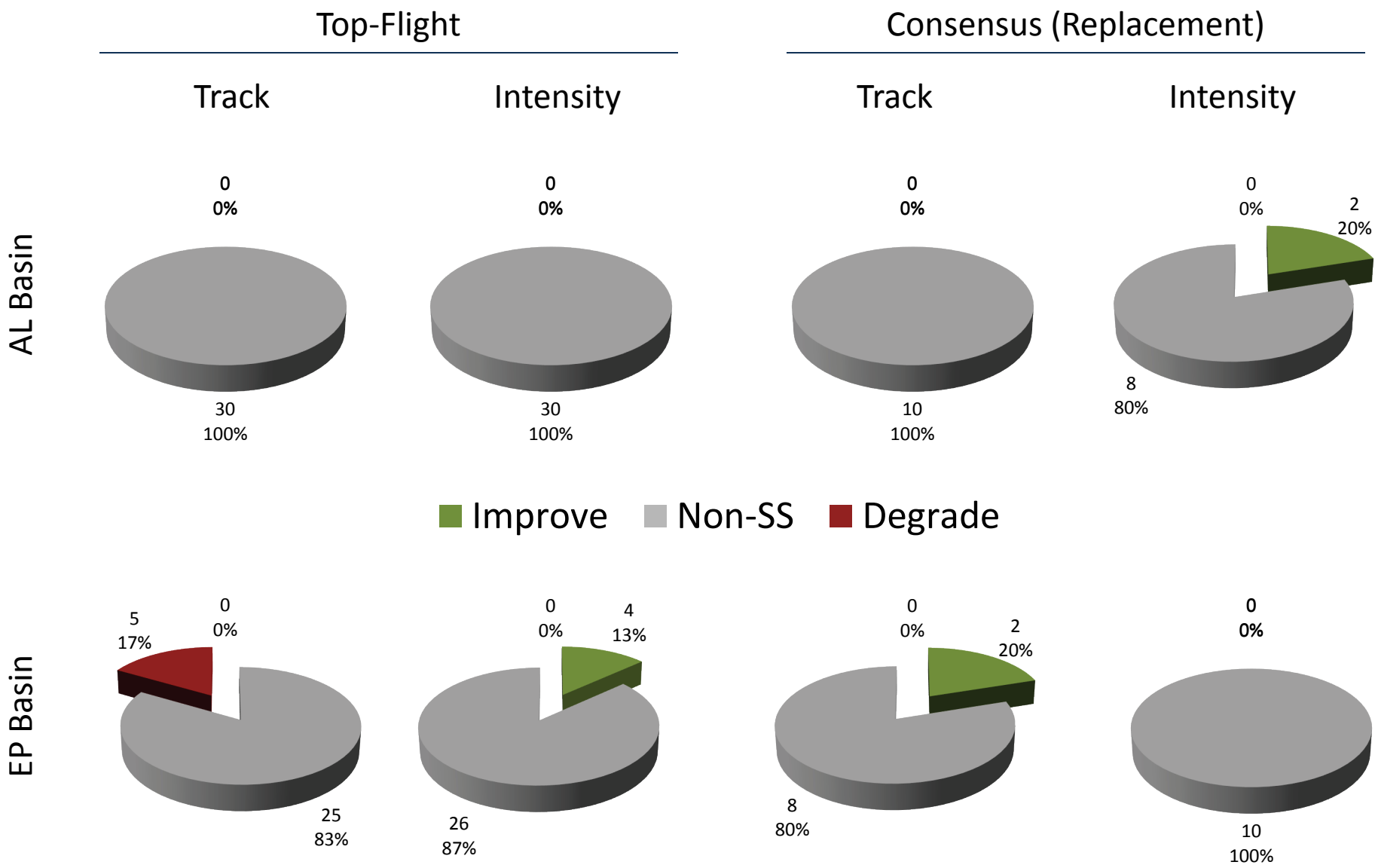
15 km FIM


Top-Flight
Consensus (Addition)
Track
Intensity
Track
Intensity
AL Basin
0
0%
8
27%
22
73%
Not Evaluated
0
0%
10
100%
Not Evaluated
Improve
Non-SS
Degrade
EP Basin
0
0%
6
20%
24
80%
Not Evaluated
0
0%
4
40%
6
60%
Not Evaluated

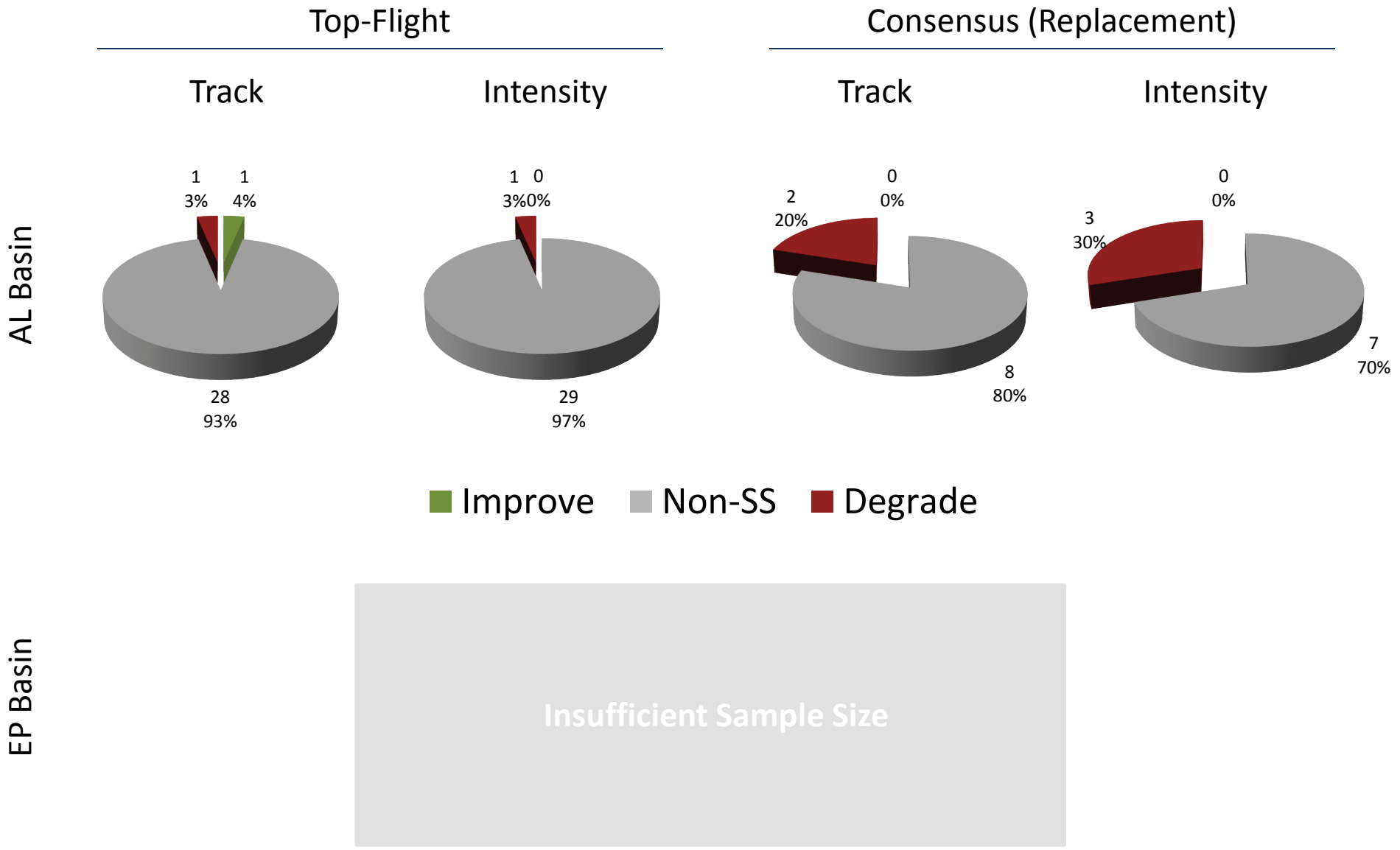
GFDL Ensemble Mean (external vortex tracker)



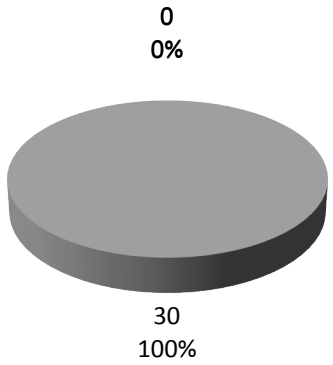
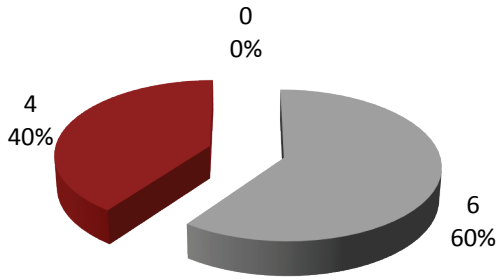
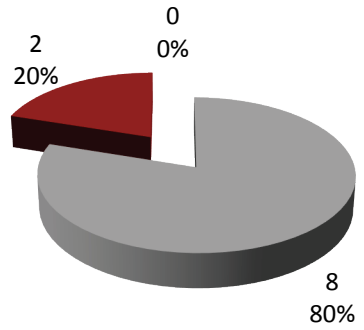
HWRF Basin Scale



HWRF Regional Hurricane Ensemble Mean



FSU Superensemble

	Top-Flight		Consensus (Comparison)	
	Track	Intensity	Track	Intensity
AL Basin	Insufficient Sample Size			
EP Basin	No Model Runs for Basin			

SPICE

S1.5

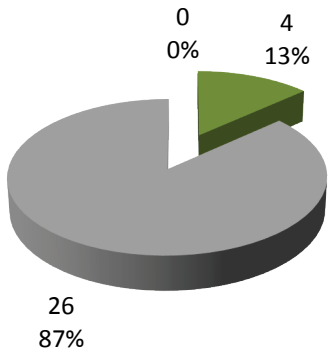
Top-Flight

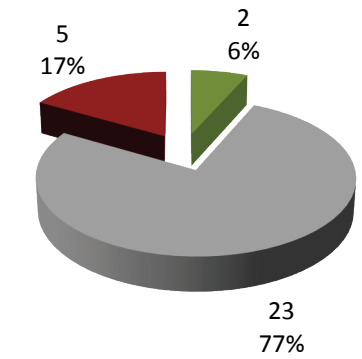
Track

AL Basin: Not Applicable

EP Basin: Not Applicable

Intensity





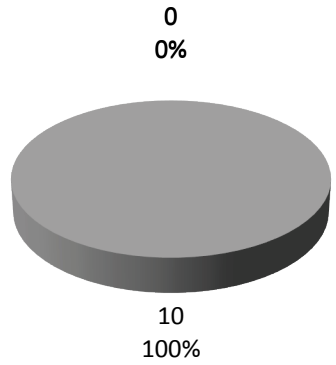
Consensus (Comparison)

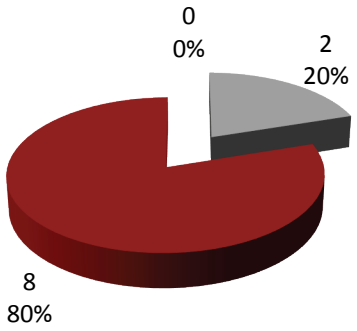
Track

AL Basin: Not Applicable

EP Basin: Not Applicable

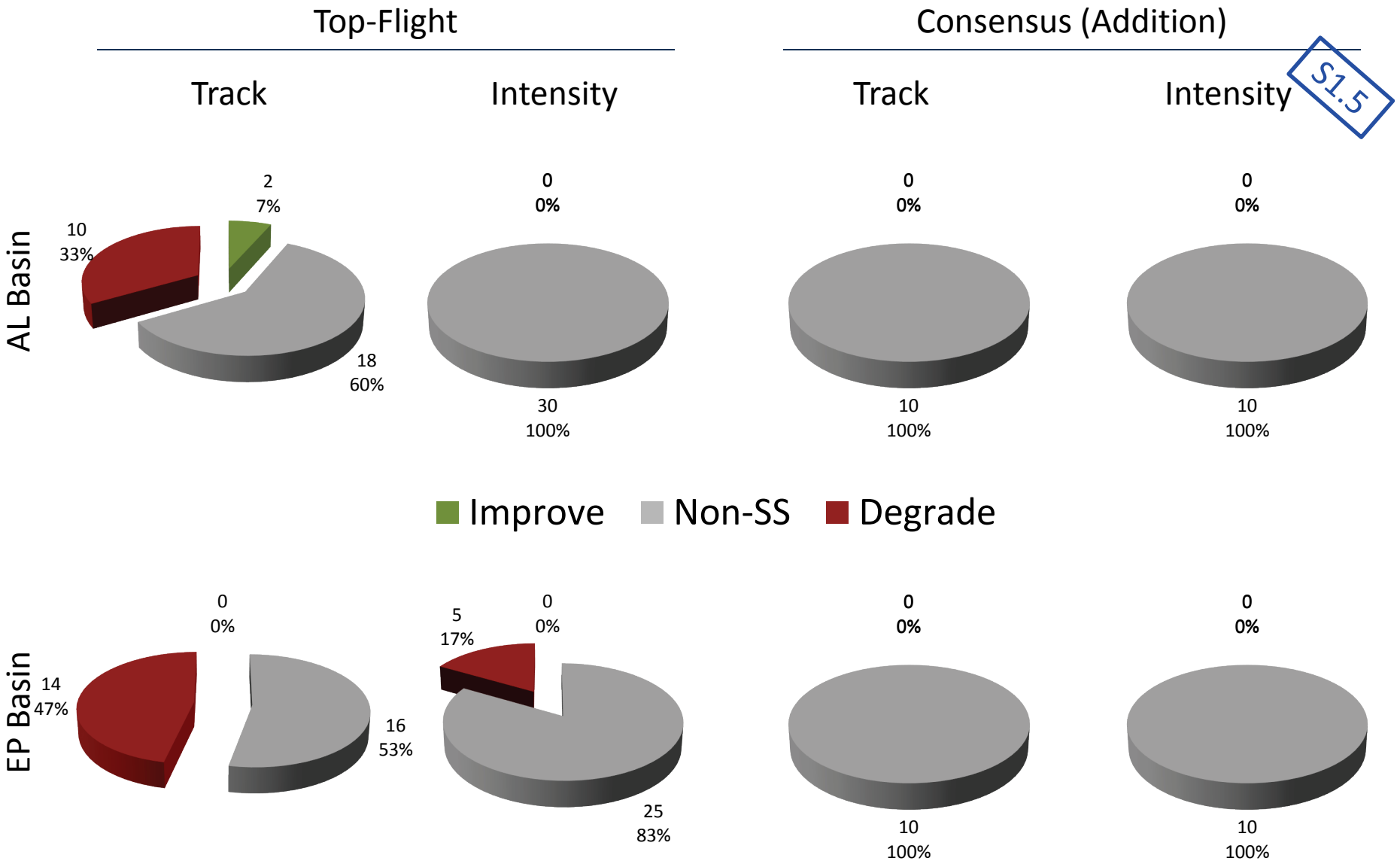
Intensity



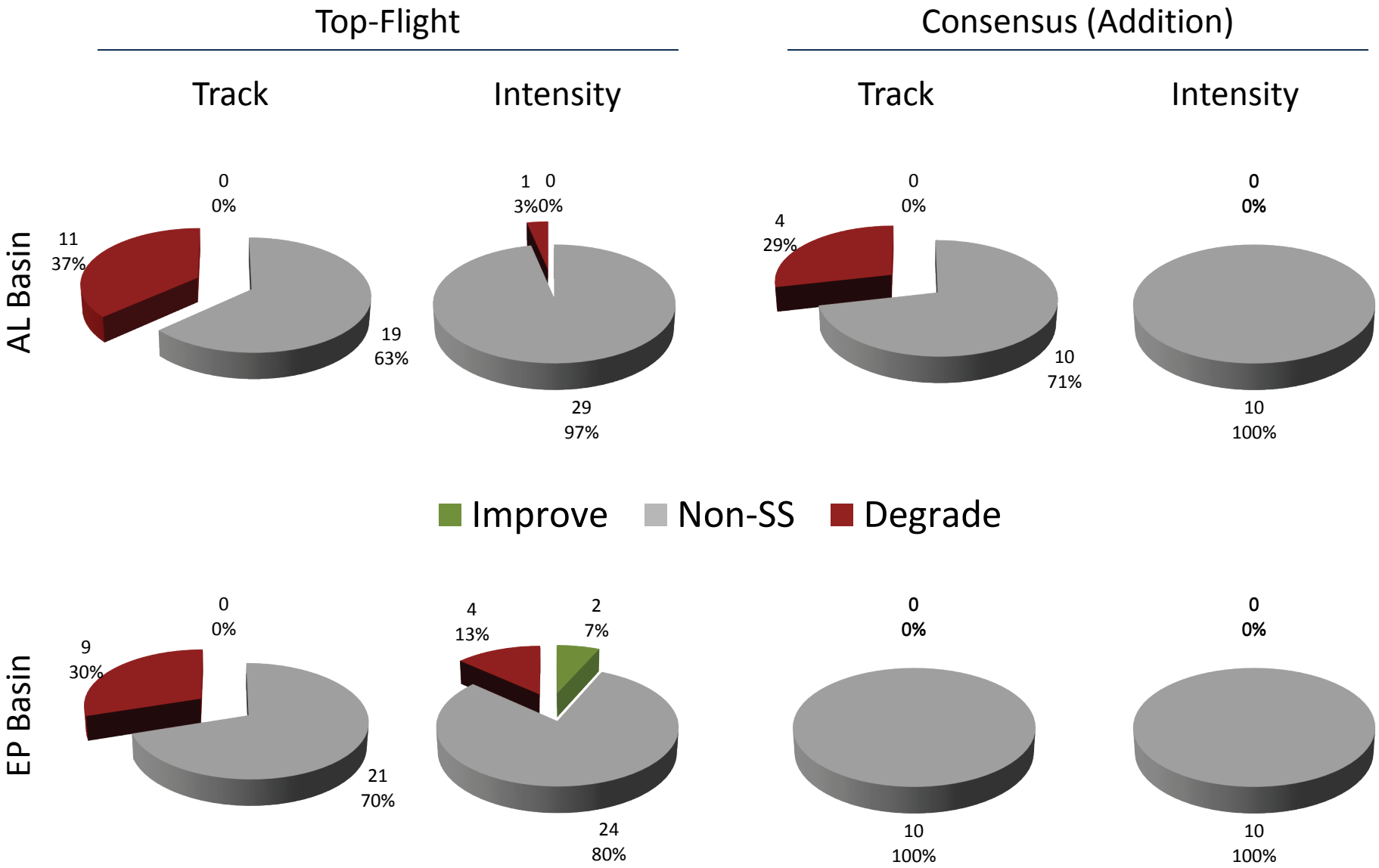


Improve Non-SS Degrade

4 km UW-NMS


Top-Flight
Consensus (Addition)
Track
Intensity
Track
Intensity
S1.5
AL Basin
2
7%
10
33%
18
60%
0
0%
30
100%
0
0%
10
100%
0
0%
10
100%
Improve
Non-SS
Degrade
EP Basin
0
0%
14
47%
16
53%
5
17%
0
0%
25
83%
0
0%
10
100%
0
0%
10
100%

5.6 km UW-NMS Ensemble Mean



Next Steps in Demo Evaluation

- Evaluate the experimental forecasts for the Demo with additional forecast provided by the modeling groups
- Conduct an in-depth diagnostic evaluation to improve the understanding of forecast errors in the models:
  - Stratification based on the amount of shear
  - Stratification based on location (function of latitude/location in basin)
  - Two dimensional structure evaluation
  - Etc.
- Retrospective evaluation of model improvements for the first five years of HFIP

Questions?