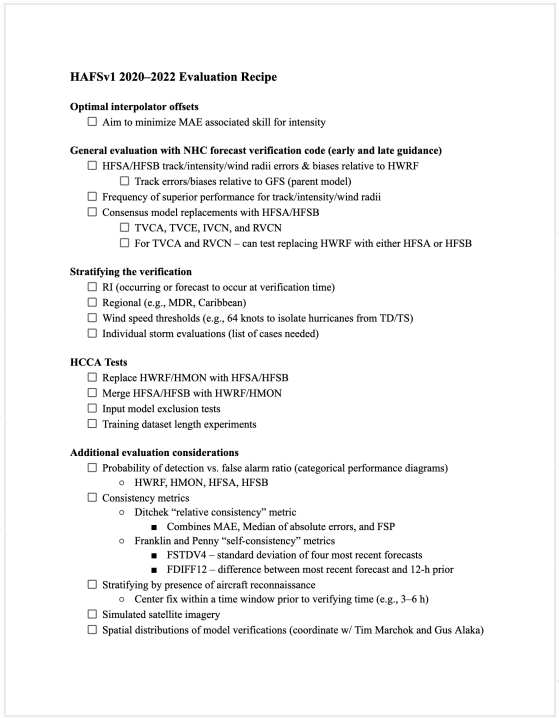
The National Hurricane Center Model Evaluation Process for the Operational Implementation of the Hurricane Analysis and Forecast System Version 1 (HAFSv1) Models

Jonathan Martinez[1,2]
David Zelinsky[1]
Wallace Hogsett[1]
Kate Musgrave[2]
John Cangialosi[1]
Benjamin Trabing[1,3]

Affiliations
[1]NOAA/NWS/NCEP National Hurricane Center
[2]Cooperative Institute for Research in the Atmosphere
[3]University Corporation for Atmospheric Research

Funding: HFIP; Cooperative Agreement NA19OAR4320073

Planning and coordinating model evaluations



HAFSv1 2020–2022 Evaluation Recipe

Optimal interpolator offsets
- ☐ Aim to minimize MAE associated skill for intensity

General evaluation with NHC forecast verification code (early and late guidance)
- ☐ HFSA/HFSB track/intensity/wind radii errors & biases relative to HWRF
  - ☐ Track errors/biases relative to GFS (parent model)
- ☐ Frequency of superior performance for track/intensity/wind radii
- ☐ Consensus model replacements with HFSA/HFSB
  - ☐ TVCA, TVCE, IVCN, and RVCN
  - ☐ For TVCA and RVCN – can test replacing HWRF with either HFSA or HFSB

Stratifying the verification
- ☐ RI (occurring or forecast to occur at verification time)
- ☐ Regional (e.g., MDR, Caribbean)
- ☐ Wind speed thresholds (e.g., 64 knots to isolate hurricanes from TD/TS)
- ☐ Individual storm evaluations (list of cases needed)

HCCA Tests
- ☐ Replace HWRF/HMON with HFSA/HFSB
- ☐ Merge HFSA/HFSB with HWRF/HMON
- ☐ Input model exclusion tests
- ☐ Training dataset length experiments

Additional evaluation considerations
- ☐ Probability of detection vs. false alarm ratio (categorical performance diagrams)
  - ○ HWRF, HMON, HFSA, HFSB
- ☐ Consistency metrics
  - ○ Ditchek "relative consistency" metric
    - ■ Combines MAE, Median of absolute errors, and FSP
  - ○ Franklin and Penny "self-consistency" metrics
    - ■ FSTDV4 – standard deviation of four most recent forecasts
    - ■ FDIFF12 – difference between most recent forecast and 12-h prior
- ☐ Stratifying by presence of aircraft reconnaissance
  - ○ Center fix within a time window prior to verifying time (e.g., 3–6 h)
- ☐ Simulated satellite imagery
- ☐ Spatial distributions of model verifications (coordinate w/ Tim Marchok and Gus Alaka)

- Formulating preliminary experiments with the Hurricane Specialist Unit (HSU) at NHC
- Active communications with the Environmental Modeling Center (EMC)
- Successive updates and adaptive integration of feedback from HSU
- *Briefings with EMC/NCEP*

UIFCW 2023
A UFS Collaboration Powered by EPIC

NHC model evaluation overview

1. Generating a fresh set of early guidance forecasts and determining the “optimal” interpolator offsets
2. Testing the addition or replacement of models in the simple consensus aids
3. Testing the addition or replacement of models in HCCA
4. Assessing the performance of individual models to be added or replaced
5. Synthesis & Recommendations

UIFCW 2023
A UFS Collaboration Powered by EPIC

The National Hurricane Center Consensus Aids

Model	Description	Model(s) to Replace
TVCN	Track Variable Consensus	HWRF (+HMON in EPAC)
IVCN	Intensity Variable Consensus	HWRF + HMON
RVCN	Wind Radii Variable Consensus	HWRF
HCCA	HFIP Corrected Consensus Approach Model	HWRF + HMON
DTOPS	Deterministic to Probabilistic Statistical Model for Rapid Intensification	HWRF
NNIC	Neural Network Intensity Combination Model	HWRF

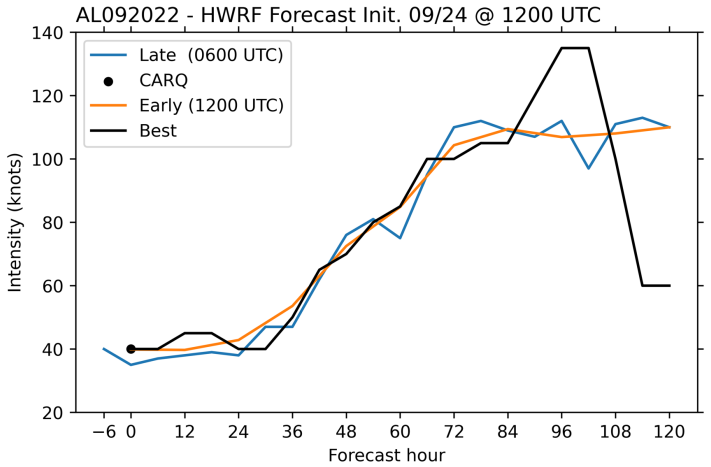
The NHC Forecast Verification Software

```
File verify_model_options.dat:  Runtime inputs for verify_model
-------------------------------------------------------------------------------------------------------
Input file locations                                                                  Normal Value
-------------------------------------------------------------------------------------------------------
data                  Subdirectory for a-decks                                        data
N                     READ_SUPA: Read data from supplemental a-decks?                 N
tc20                  Subdirectory for supplemental a-decks
wwpts_us.txt          File containing watches and warnings                            wwpts_us.txt
-------------------------------------------------------------------------------------------------------
Verification sample specifications                                                    Normal Value
-------------------------------------------------------------------------------------------------------
00000000,00000000     MINDATE,MAXDATE (yyyymmdd) to include (issuance)                00000000,00000000
YYYY                  Include forecasts from 00,06,12,18Z?                            YYYY
0,90                  RLATMIN,RLATMAX (N) for initial positions                       0,90
-140,0                RLONMIN,RLONMAX (E) for initial positions                       -140,0
N                     Include only forecasts verifying within a box?                  N
-20,90                FLATMIN,FLATMAX (N) if verifying only in a box
-140,-50              FLONMIN,FLONMAX (E) if verifying only in a box
0                     CLPR/SHFR option: 0=Calc,1=Wrt,2=Rd c-deck,3=Rd a-deck          0
0 0 0 0 0             MLAG(5): Lag (h)                                                0 0 0 0 0
Y                     Accept 12-h interpolations as 6-h interpolations?               Y
2                     Minimum number of models required for variable cons.            1
N                     WATER_ONLY: Stop verif when fcst or bt hits land?               N
A                     FCSTNEARLAND  Keep (A)ll, (N)ear or (F)ar from land?            A
A                     SIZECHECK     Keep (A)ll, (L)arge or (S)mall storms?            A
200                   DIST_LAND  Fcst tossing threshold (nmi)                         200
200                   STMSIZE  ROCI size-tossing threshold (nmi)                      200
15                    Intensity (kt) assigned to dissipation stage                    15
Y                     WS_CRIT_VT: Apply WS criteria to verification time?             Y
N                     ACPTALLST_IT: Accept any status at initial time?                N
N                     ACPTALLST_VT: Accept any status at verifying time?              N
N                     VER_TD_INIT: Only verify fcst if TD at T=0?                     N
N                     VER_NY34: Only verify prior to system reaching 34 kt?           N
N                     FDIFF_ONLY: Only verify if forecasts different?                 N
N                     VER_RI: Only verify if RI/RW occurring at VT.                   N
N                     VER_RIF: Only verify if RI occurring or fcst at VT.             N
30,24                 DELW_RI, DELT_RI: RI/RW wind,interval values (kt,hr)            30,24
N                     VICAT: Verify intensity categories, rather than speed?          N
N                     VER_OPER: Use operational track to verify with?                 N
N                     VER_SAB: Use SAB classifications to verify with?                N
Y                     SKP_ENSM: Skip over ensemble members in a-deck?                 Y
-------------------------------------------------------------------------------------------------------
Output data specifications                                                            Normal Value
-------------------------------------------------------------------------------------------------------
```

- Fortran processing and Python post-processing
- Configurable options for verification criteria
  - HFIP RI definition, Watches/Warnings, Aircraft reconnaissance, Distance to land masses, and more
- Flexible verification of single/multiple storms and seasons
  - Aggregated per storm, year, and over a range of years
- Tabulation of forecast errors
- Generates variable consensus aids on-the-fly
- CLIPER5/D-SHIFOR5 for climatology/persistence baseline in the annual Tropical Cyclone Reports

Generating a fresh set of early guidance forecasts

- Prerequisite to all subsequent analyses
- Guidance from dynamical models arrives “late” to forecasters
  - Example – 1200 UTC cycle with forecast issued @ 1500 UTC, but model output arrives after forecast issuance



Interpolation Process

1. Late forecast interpolated to 3-h time steps with a linear b-spline
2. Smooth 10 times with a 1-2-1 filter
3. Adjust current intensity (t = 0 h) with CARQ entry
4. Relax intensity offset within the time window specified by the offset start (T_i) and offset end (T_f)

"Optimal" Interpolator Offsets

Baseline Definitions

1. Constant offset
2. Lagged forecast
3. Offset applied at τ <= 06 h
4. Offset applied at τ <= 12 h

For each baseline (m) and offset start/stop couplet (n^{th} model):

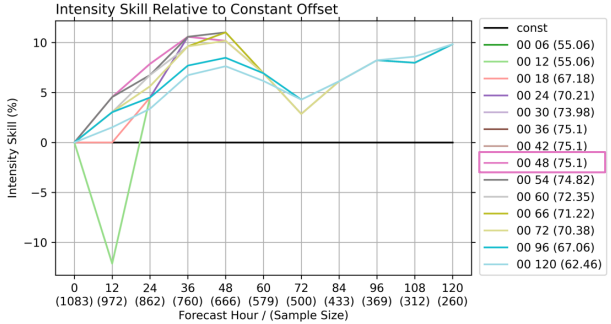
- Sum the intensity skill S over all basins (j) and verifying times (k)

$$S_{m,n} = \sum_{k=0}^{10} \sum_{j=0}^{1} S_{j,k,m,n}$$

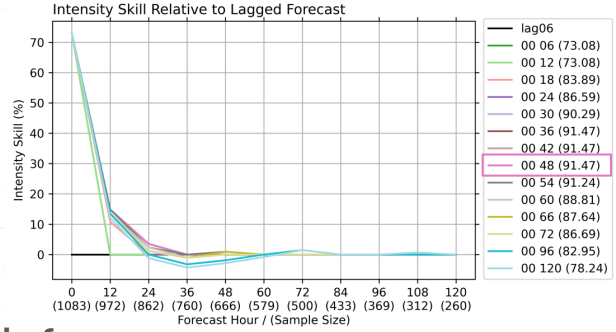
j = Basin (AL, EP)
k = Forecast verification time (0–120 h)
m = Baseline (see Slide 2)
n = Model (offset start/stop couplet)

Operational Interpolator Offsets for HFSA/HFSB

Constant Offset



Lagged Offset

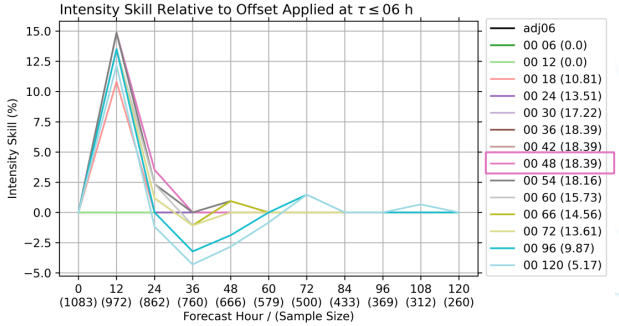


Aggregated over AL/EP basins and all verifying times τ = {00, …, 120} h

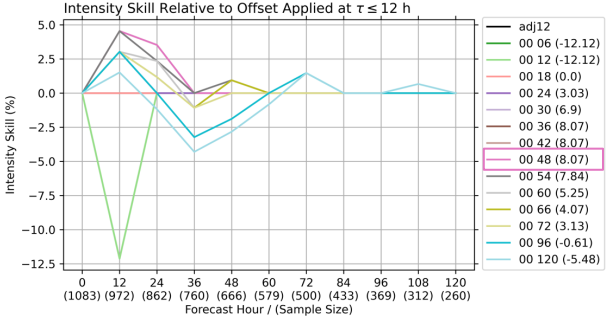
HFSA – T_i = 00; T_f = 48
HFSB – T_i = 06; T_f = 42

Example for HFSA

Offset at $\tau \leq 06$ h



Offset at $\tau \leq 12$ h



UIFCW 2023
A UFS Collaboration Powered by EPIC

Track Variable Consensus (TVCN) Verification

TVCN – Variable, equally-weighted consensus with minimum 2 members

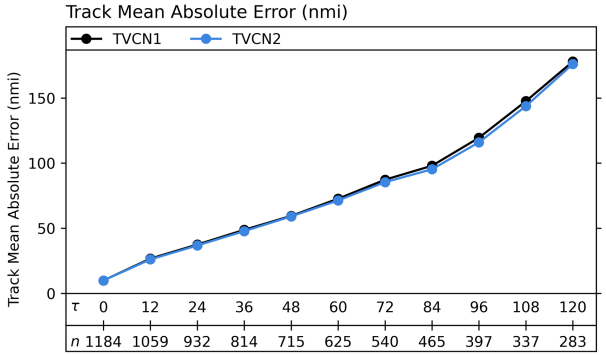
North Atlantic (AL) – 6 models with only HWRF to be replaced

- TVCN1 – AVNI, CTCI, EGRI, EMNI, EMXI, HWFI
- TVCN2 – Replace HWFI with HFAI

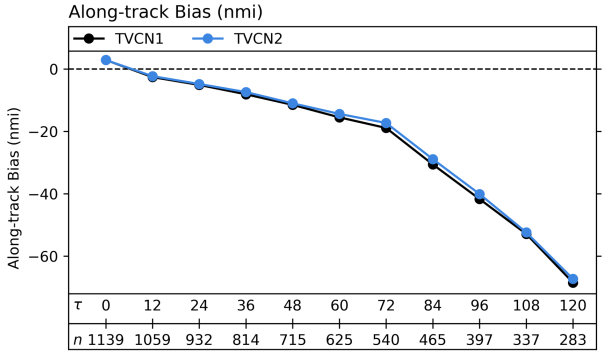
** Note –TVCN with HWRF, HMON, HFSA, HFSB in operations for 2023

Track Variable Consensus (TVCN) Verification

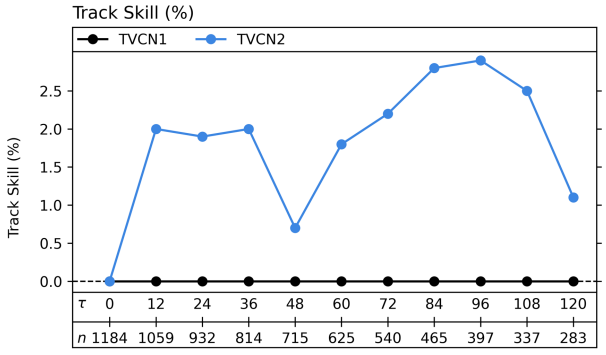
Mean Absolute Error



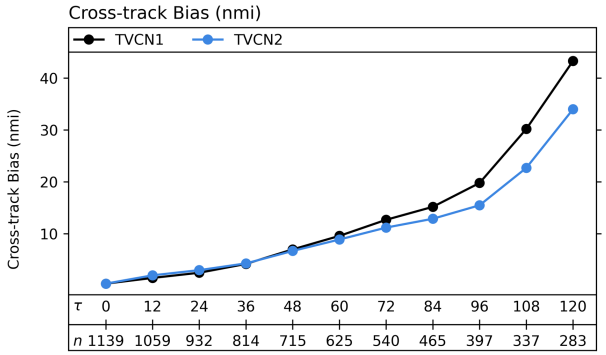
Along-track Bias



Relative Skill



Cross-track Bias



- HAFSv1 track forecasts consistently outperform HWFI and HMNI
- Reduced along- and cross-track biases

UIFCW 2023
A UFS Collaboration Powered by EPIC

Intensity Variable Consensus (IVCN) Verification

IVCN – Variable, equally-weighted consensus with minimum 2 members

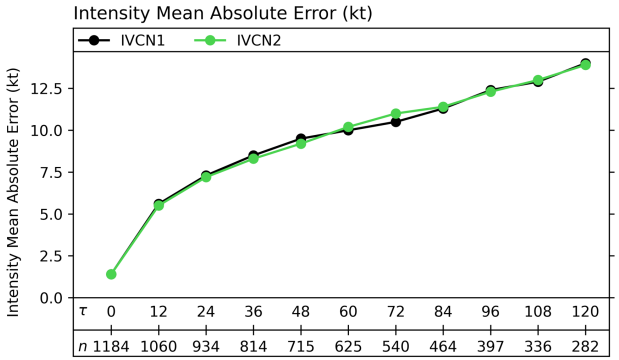
North Atlantic (AL) – 5 models with HWRF + HMON to be replaced

- IVCN1 – DSHP, CTCI, HMNI, HWFI, LGEM
- IVCN2 – Replace HWFI + HMNI with HFAI + HFBI

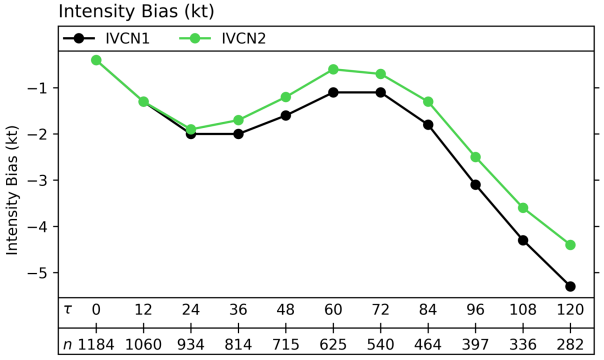
** Note –IVCN with HWRF, HMON, HFSA, HFSB in operations for 2023

Intensity Variable Consensus (IVCN) Verification

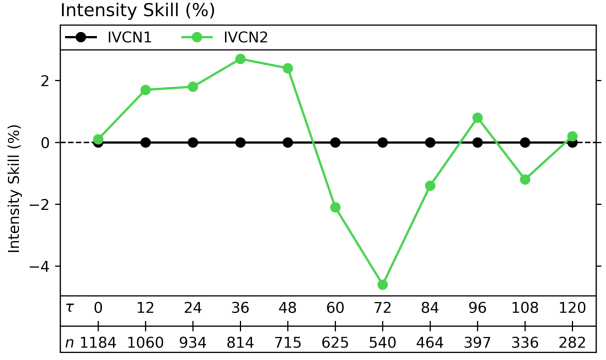
Mean Absolute Error



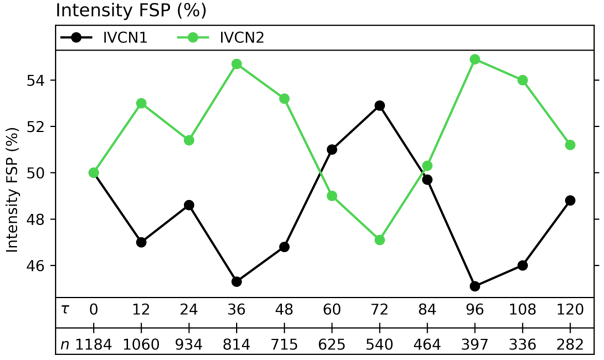
Bias



Relative Skill



FSP



- HAFSv1 improves intensity forecast skill of IVCN through day 2
- Slight degradation of skill between days 2–4
- Reduced intensity bias

UIFCW 2023
A UFS Collaboration Powered by EPIC

Wind Radii Variable Consensus (RVCN) Verification

RVCN – Variable, equally-weighted consensus with minimum 1 member

Different techs given uniform interpolator offsets (06/18) compared to track/intensity

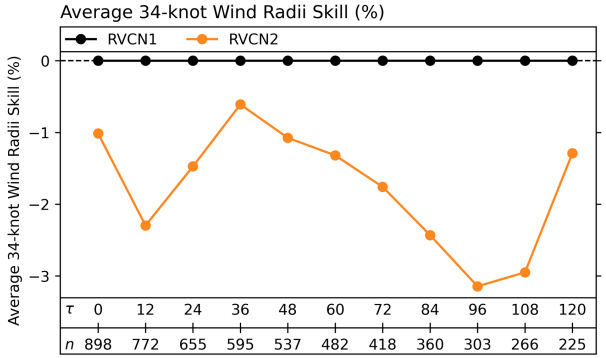
North Atlantic (AL) – 4 models with HWRF to be replaced

- RVCN1 – AHNI (AVNI), CHCI (CTCI), EHXI (EMXI), HHFI (HWFI)
- RVCN2 – Replace HHFI (HWFI) with HHAI (HFAI)

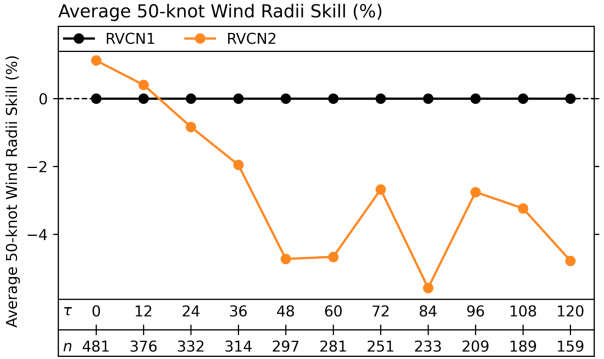
** Note –RVCN with HWRF, HFSA, HFSB in operations for 2023

Wind Radii Variable Consensus (RVCN) Verification

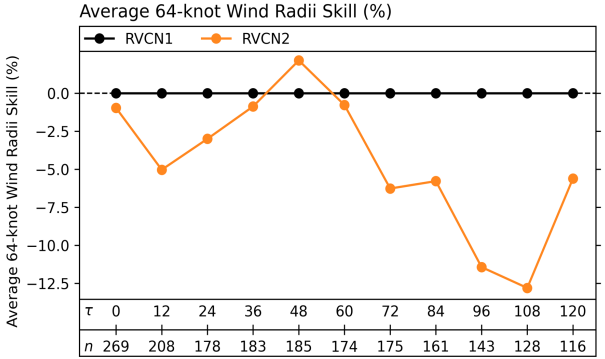
R34 Relative Skill



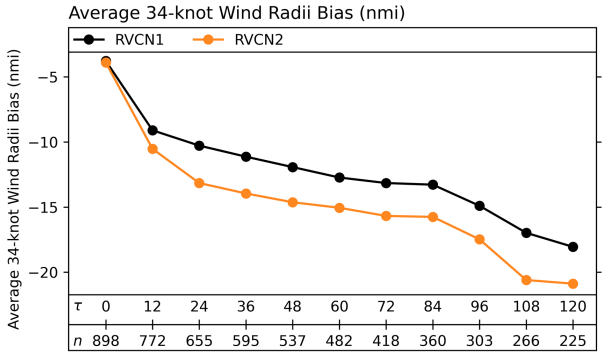
R50 Relative Skill



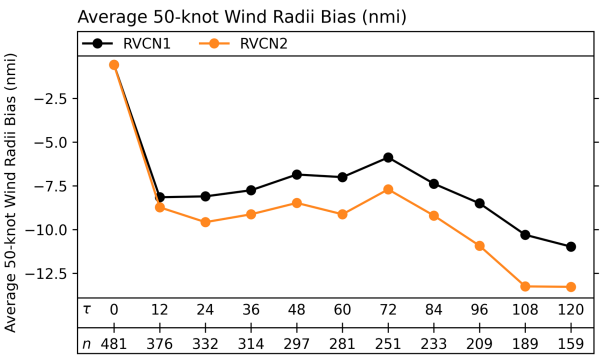
R64 Relative Skill



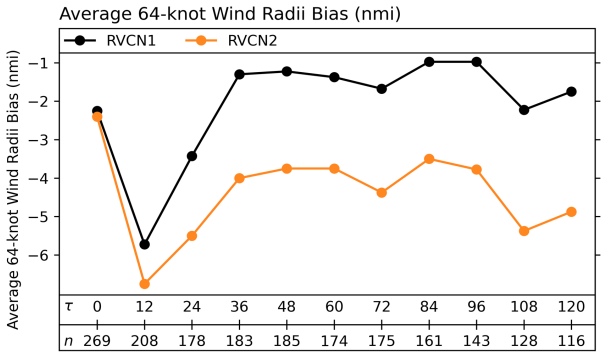
R34 Bias



R50 Bias



R64 Bias



HCCA Verification

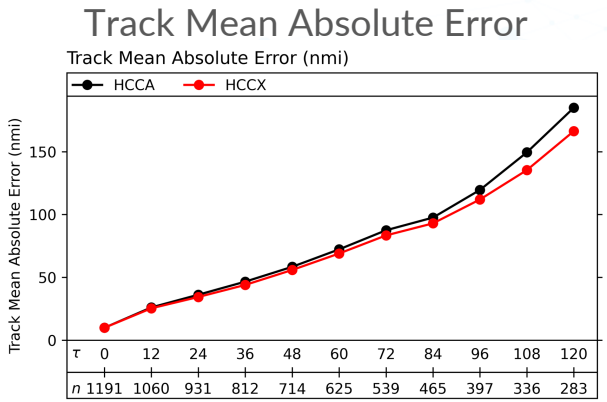
HCCA – Variable, unequally-weighted consensus using a “super ensemble” approach

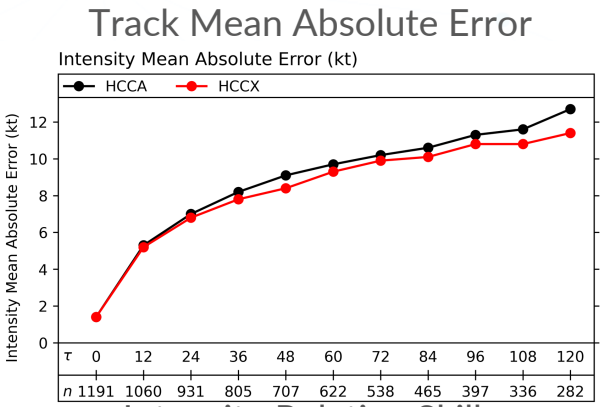
Leave-one-out training – All TCs included in training data except the one being forecast

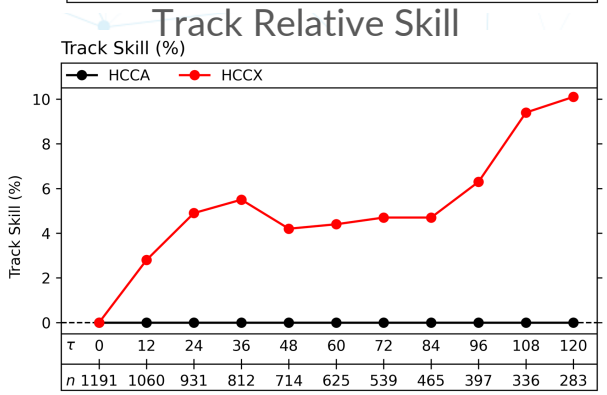
- HCCA – baseline 2022 configuration
- HCCX (track) – add HFAI + HFBI (keep HWFI)
- HCCX (intensity) – add HWFI + HMNI (keep HWFI + HMNI)

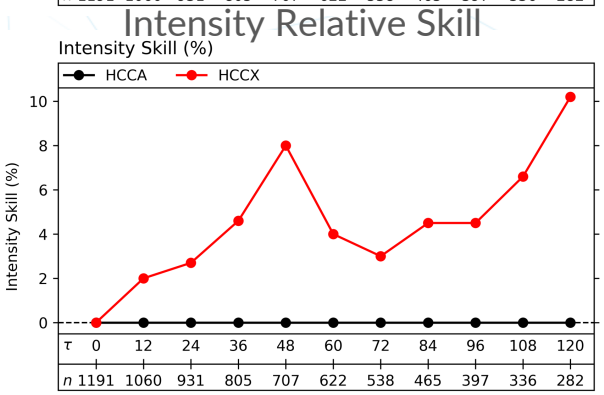
** Note – HCCA has HWRF, HMON, HFSA, HFSB and additional updates for 2023

HCCA Verification



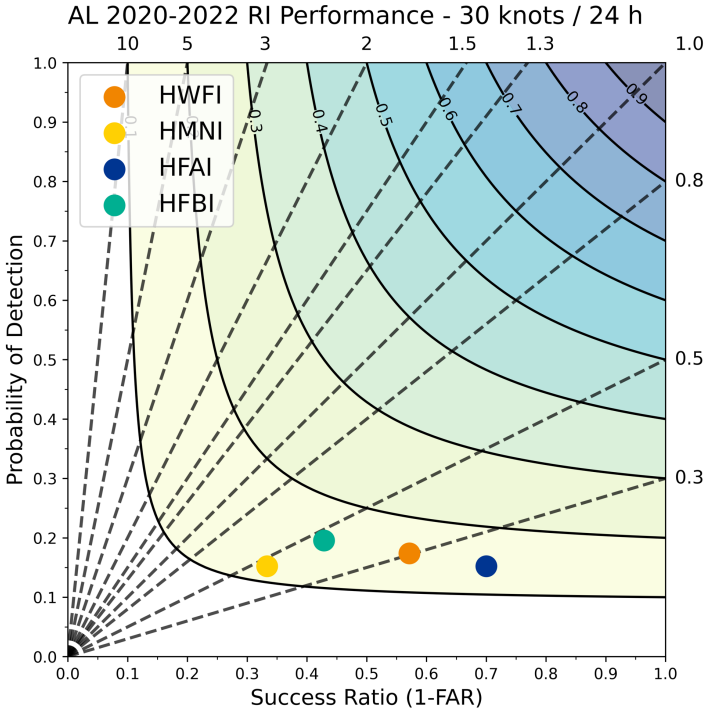






- Track and intensity forecast skill improve up to ~10%

Rapid Intensification – 25 kt/24 h (occurring or forecast)



HWFI — RI Observed

RI Forecast	Yes	No	Total
Yes	8	6	14
No	38	767	805
Total	46	773	819

HMNI — RI Observed

RI Forecast	Yes	No	Total
Yes	7	14	21
No	39	759	798
Total	46	773	819

HFAI — RI Observed

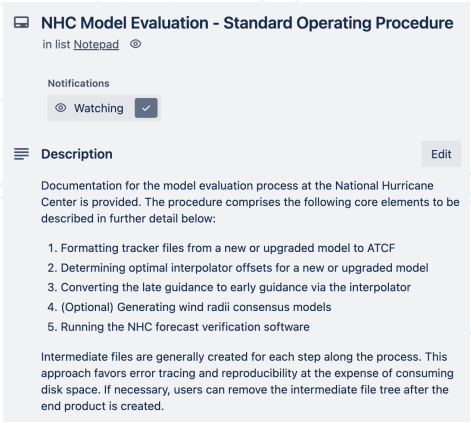
RI Forecast	Yes	No	Total
Yes	7	3	10
No	39	770	809
Total	46	773	819

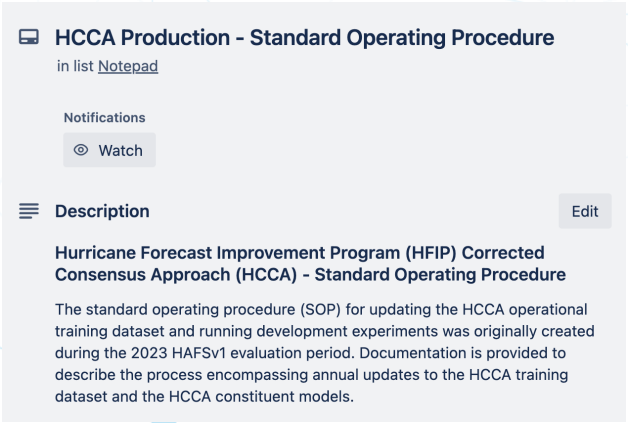
HFBI — RI Observed

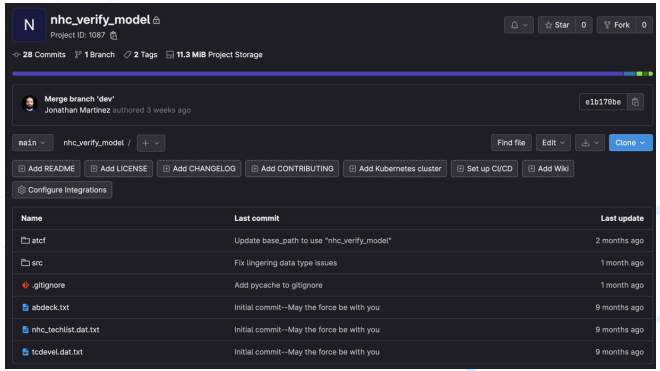
RI Forecast	Yes	No	Total
Yes	9	12	21
No	37	761	798
Total	46	773	819

Summary of NHC Model Evaluation Process

- Multifaceted and multiorganizational evaluation for HAFSv1 transition to operations
  - Coordination between NHC, CIRA, EMC, HRD, DTC, and others
- Moving forward– standardizing the model evaluation process
- Looking ahead – parallel evaluations with METplus
  - Synthesizing NHC Verification Software features

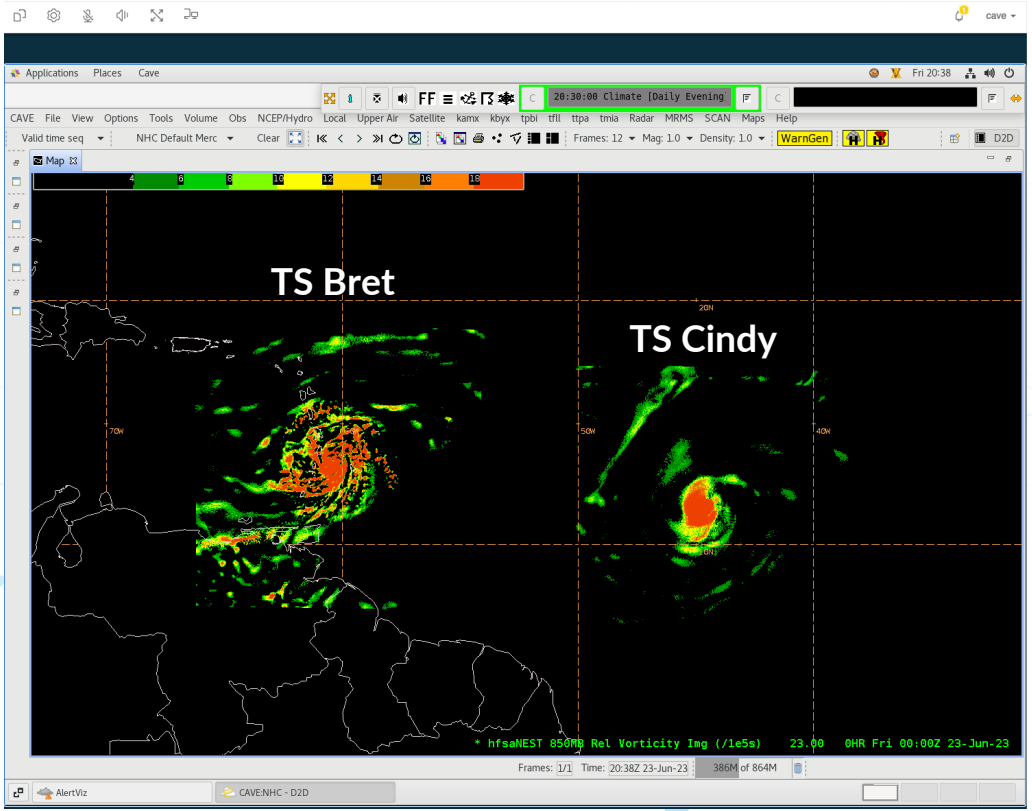






UIFCW 2023
A UFS Collaboration Powered by EPIC

A sneak peak of HAFS in Cloud AWIPS



- Coordinating with Technology and Science Branch (TSB)
- Cloud AWIPS as a testbed/on-ramp for operations

UIFCW 2023
A UFS Collaboration Powered by EPIC