HWRF + COAMPS-TC + GFDL combined ensemble retrospective test

Jon Moskaitis, Alex Reinecke, Jim Doyle: NRL-Monterey

COAMPS-TC ensemble sample:

185 cases from 10 TCs:

(1) Earl, 2010/07L
(2) Igor, 2010/11L
(3) Irene, 2011/09L
(4) Katia, 2011/12L
(5) Ernesto, 2012/05L
(6) Isaac, 2012/09L
(7) Nadine, 2012/14L
(8) Sandy, 2012/18L
(9) Gabrielle, 2013/07L
(10) Humberto, 2013/09L

HWRF ensemble sample:

520 cases from 2011-2013 HFIP stream 1.5 retro sample in Atlantic and Eastern Pacific basins

GFDL ensemble sample:

716 cases from 2011-2013 HFIP stream 1.5 retro sample in Atlantic basins

HWRF + COAMPS-TC + GFDL sample:

133 cases from 8 TCs:

(1) Irene, 2011/09L
(2) Katia, 2011/12L
(3) Ernesto, 2012/05L
(4) Isaac, 2012/09L
(5) Nadine, 2012/14L
(6) Sandy, 2012/18L
(7) Gabrielle, 2013/07L
(8) Humberto, 2013/09L

COAMPS-TC ensemble design

- 27-, 9-, 3-km horizontal grid spacing
- 1 control + 10 members with initial and boundary condition perturbations
- No physics perturbations
- No data assimilation
- Control forecast:
  - Initialized from the GFS analysis
  - Vortex initialized with a Rankine vortex based on TC vitals
- Ensemble members IC's perturbed about the control:
  - Synoptic perturbations drawn from static covariance (e.g. WRFVAR cv3)
  - Vortex IC's based on perturbed TC vitals

COAMPS-TC ensemble design

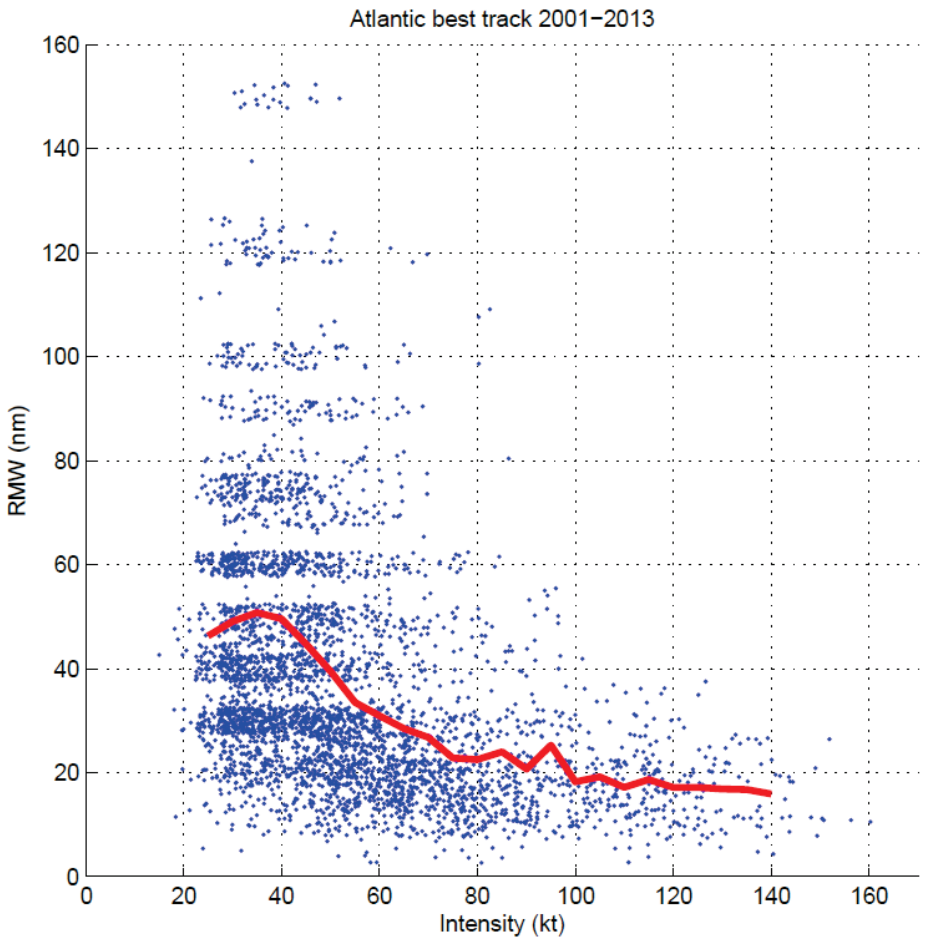
- Synoptic perturbations:
  - Perturbation drawn from a static covariance
  - Perturb the boundaries
  - Use WRFVAR cv3

- Vortex scale perturbations:
  - Vortex position, max wind, and RMW.

  - Perturbation variance from:
    - Torn and Snyder 2012
    - Landsea and Franklin 2013

  - Max wind and RMW covariance derived from 2001-2013 best track data.

  - Variances and covariances depend on TC-vital max wind speed.

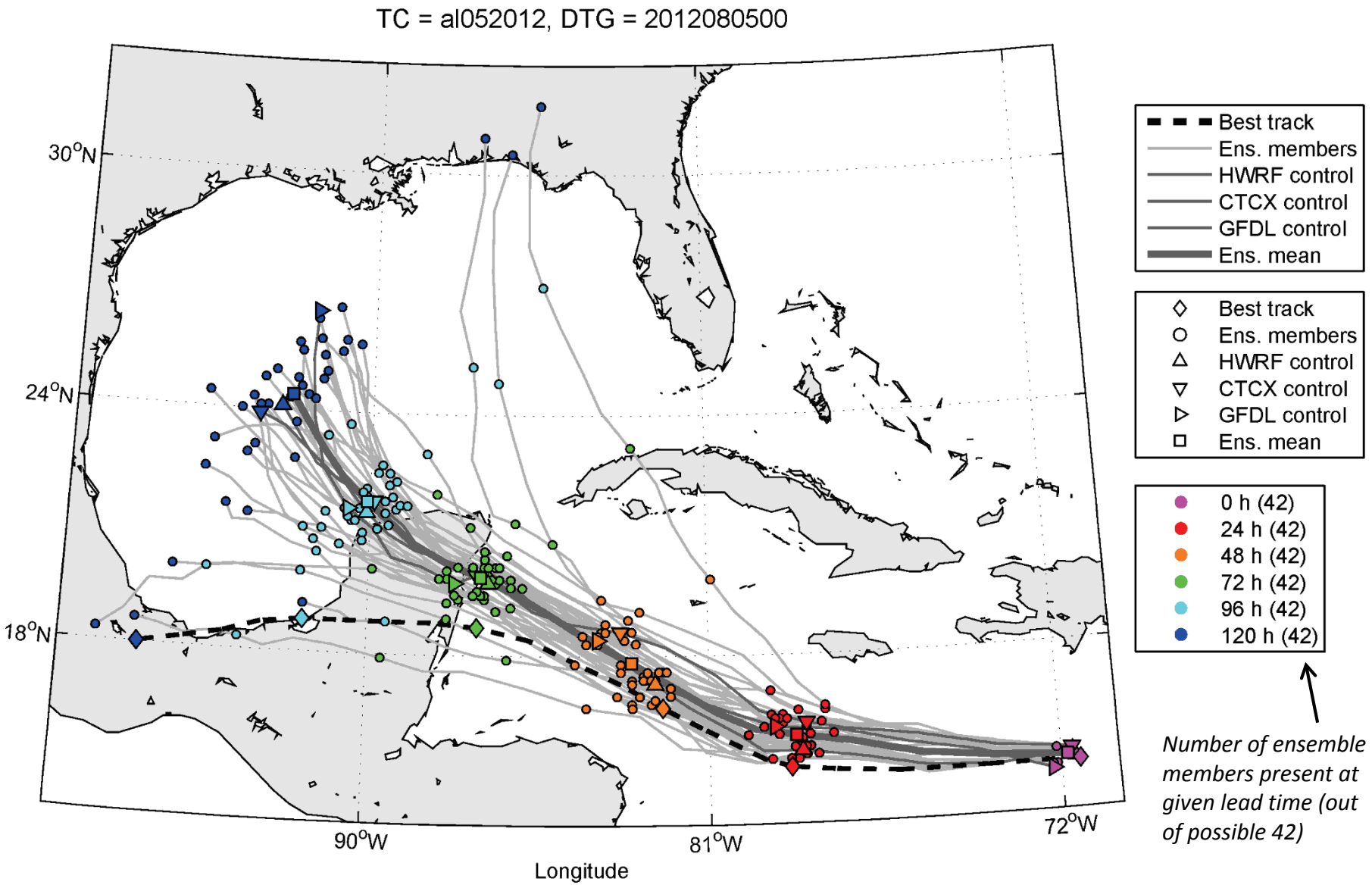


HWRF + COAMPS-TC + GFDL Ensemble products

Example shown for particular Ernesto forecast, graphics for all 133 cases are available

Track forecast plot: Lines

HWRF + COAMPS-TC + GFDL, Ernesto (05L, 2012)

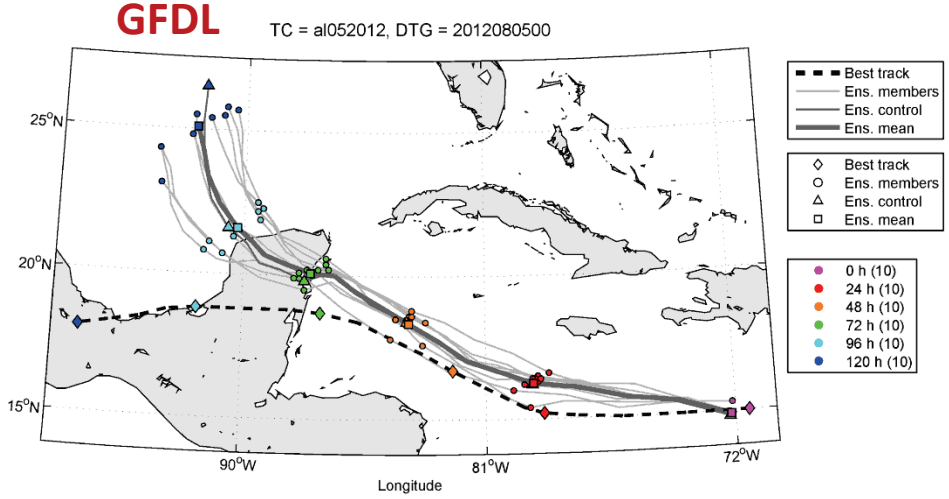


Number of ensemble members present at given lead time (out of possible 42)

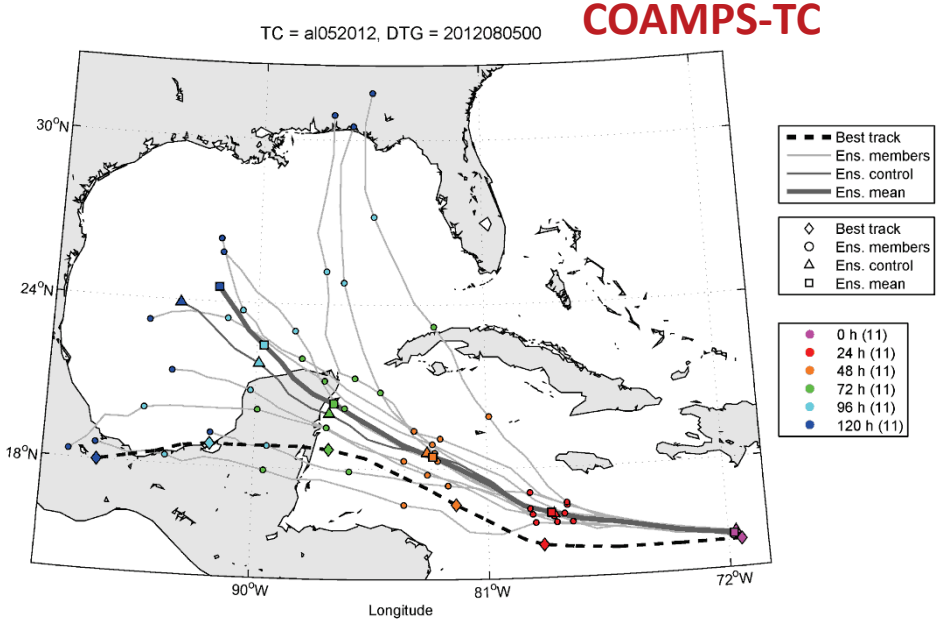
Track forecast plot: Lines

Individual model ensembles, Ernesto (05L, 2012)

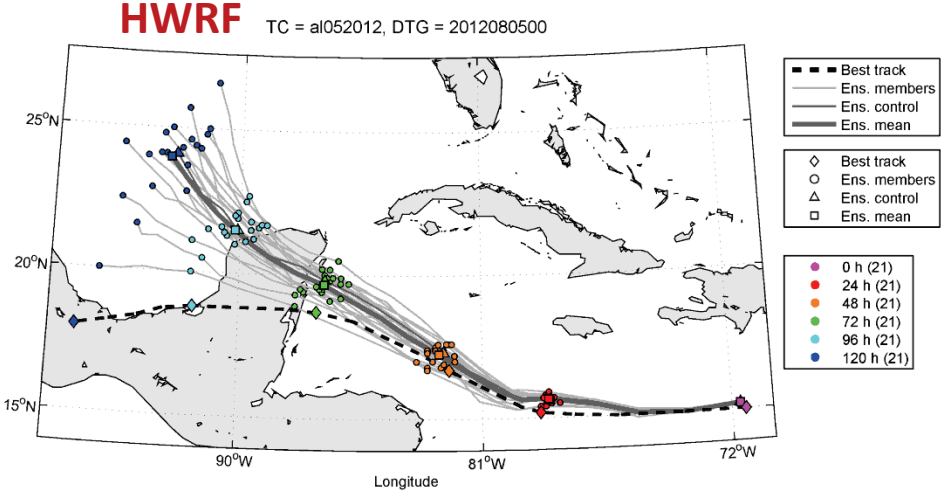
GFDL



COAMPS-TC



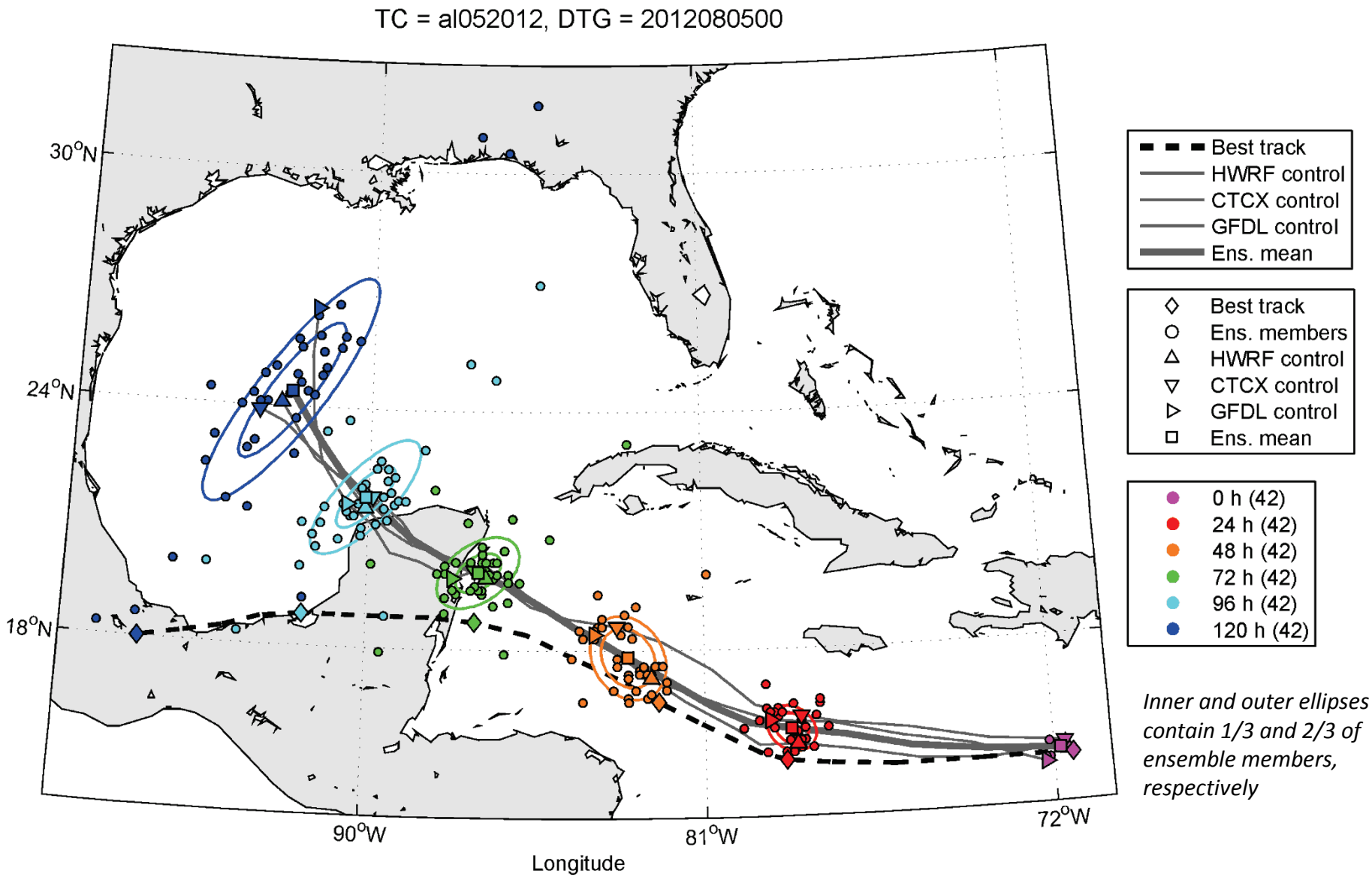
HWRF



Typically, COAMPS-TC has much greater track spread (2-3x) than GFDL or HWRF

Track forecast plot: Ellipses

HWRF + COAMPS-TC + GFDL, Ernesto (05L, 2012)



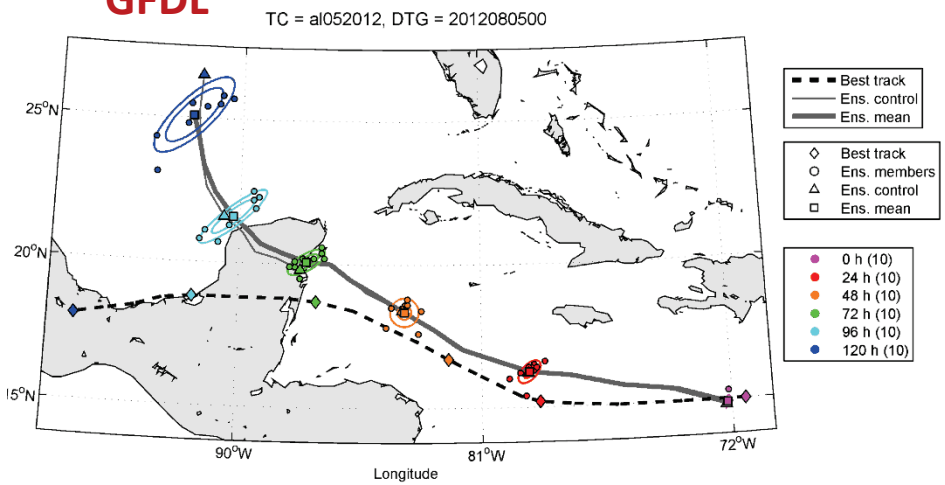
Inner and outer ellipses contain 1/3 and 2/3 of ensemble members, respectively

Ellipses are centered on ensemble mean, with length and orientation of axes according to eigenvectors of ensemble position covariance matrix

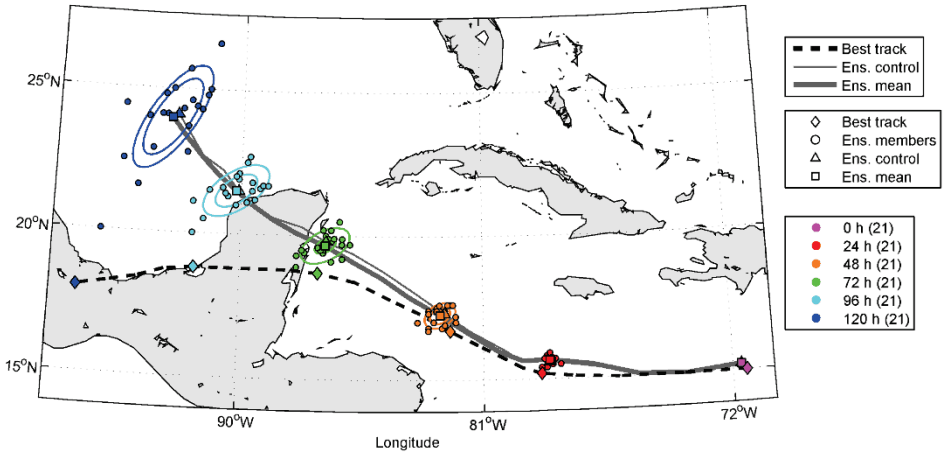
Track forecast plot: Ellipses

Individual model ensembles, Ernesto (05L, 2012)

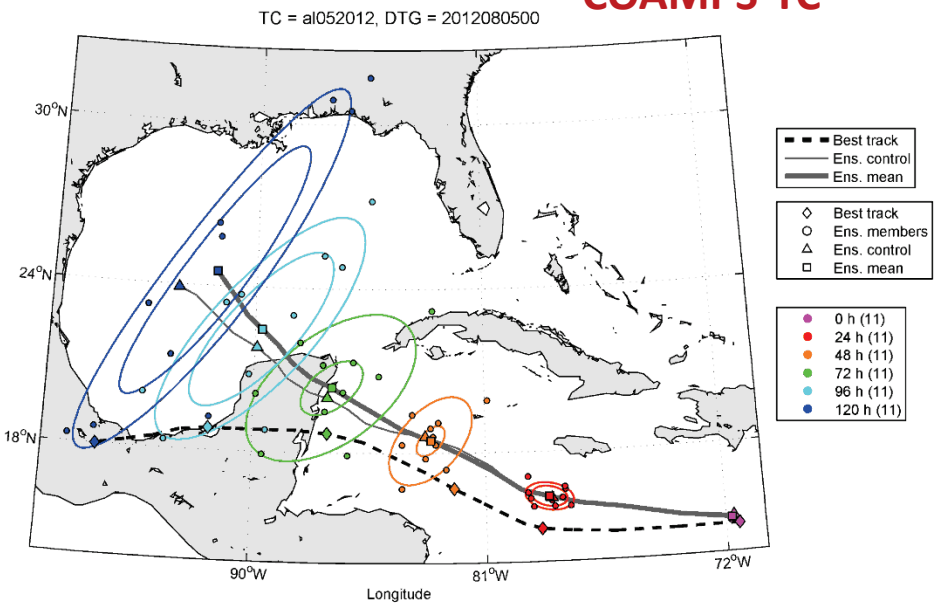
GFDL



HWRF



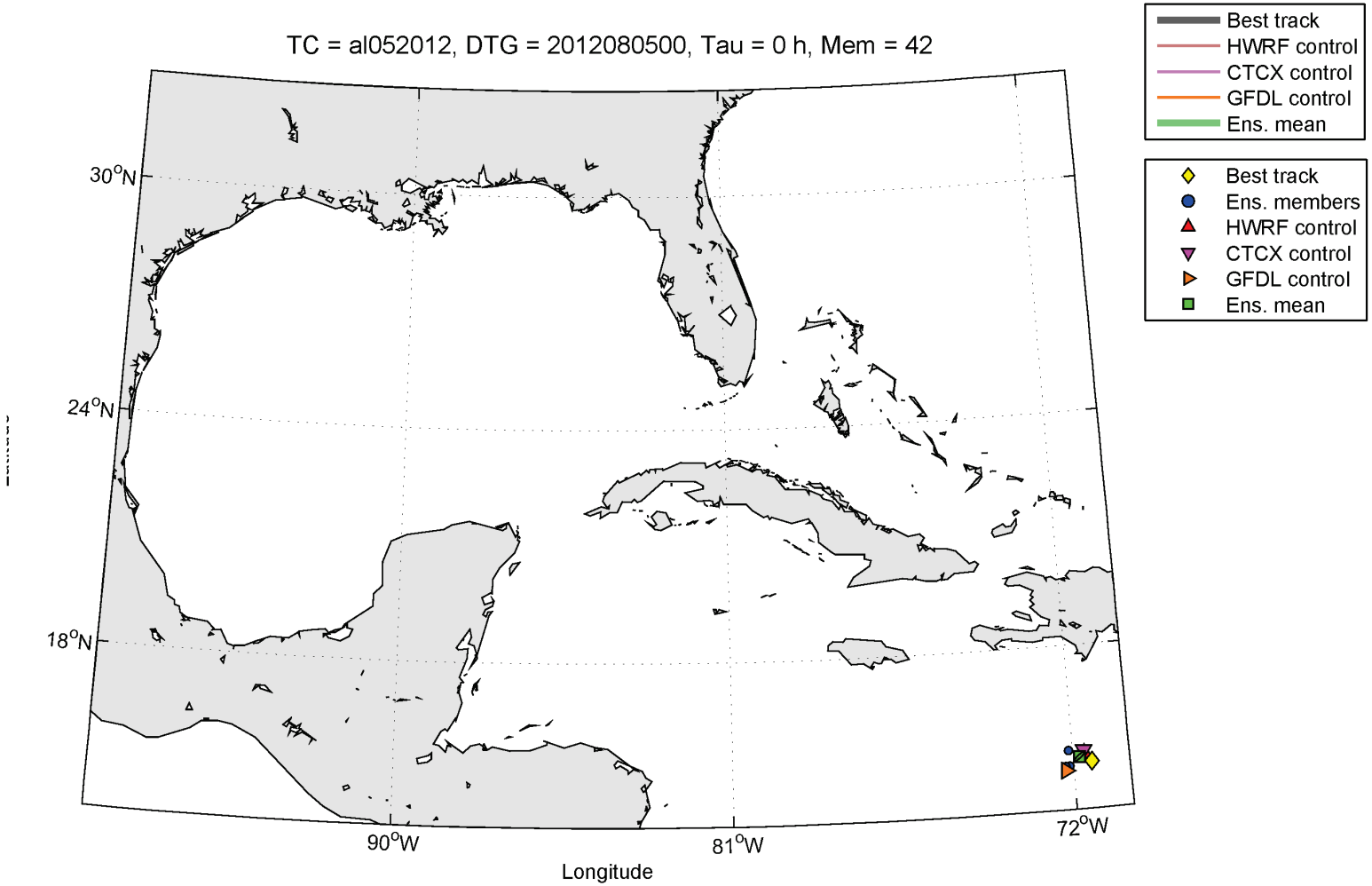
COAMPS-TC



Typically, COAMPS-TC has much greater track spread (2-3x) than GFDL or HWRF

Track forecast plot: Movie

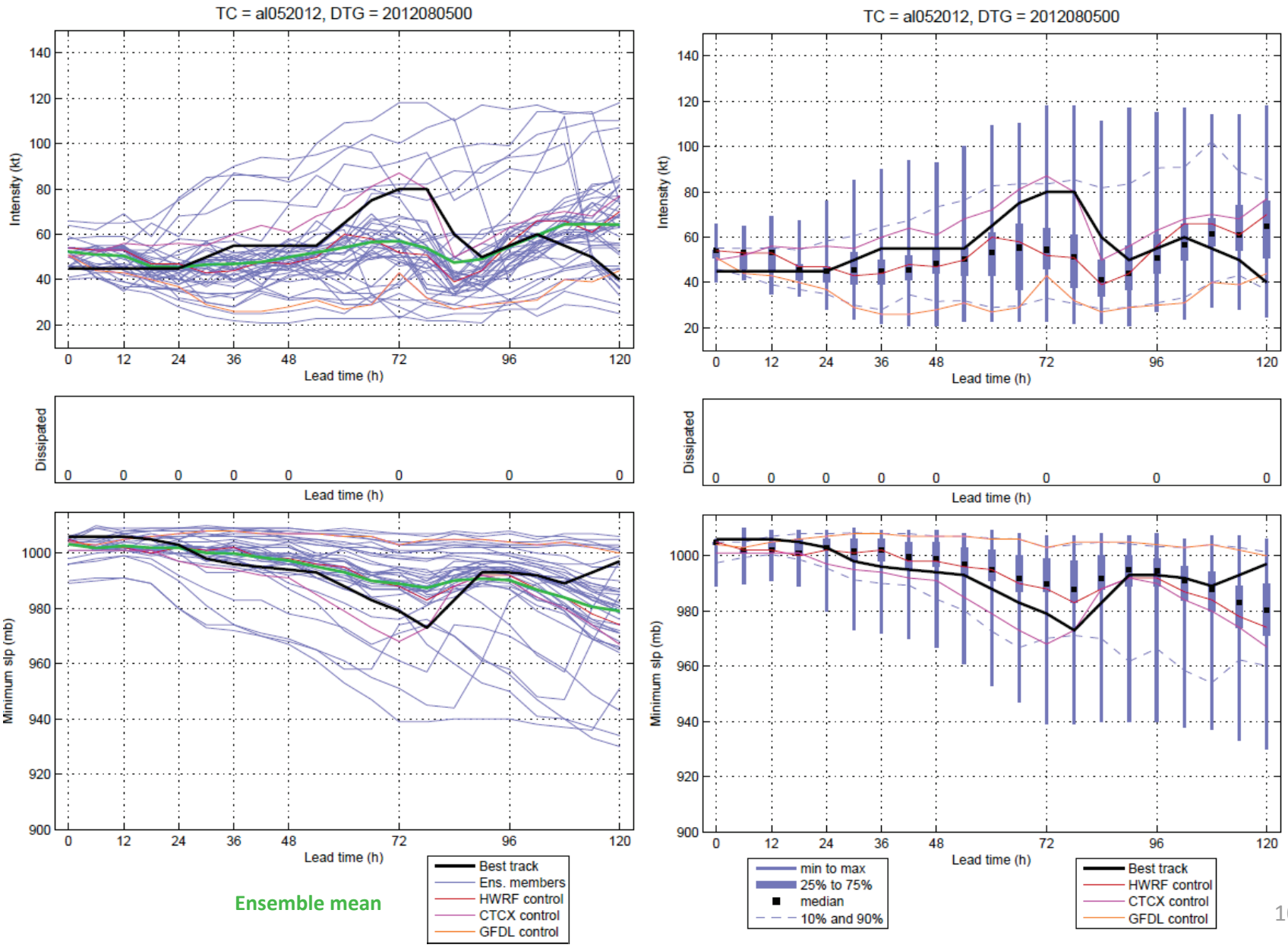
HWRF + COAMPS-TC + GFDL, Ernesto (05L, 2012)



Ellipse contains 2/3 of ensemble members, current lead time in black and older lead times in gray

Intensity and Min SLP

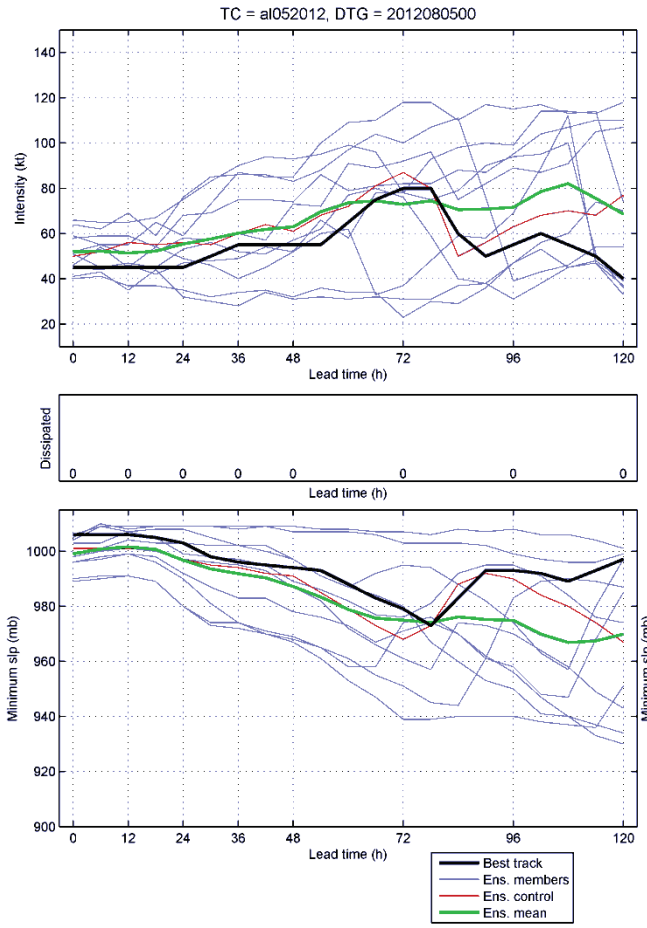
HWRF + COAMPS-TC + GFDL, Ernesto (05L, 2012)



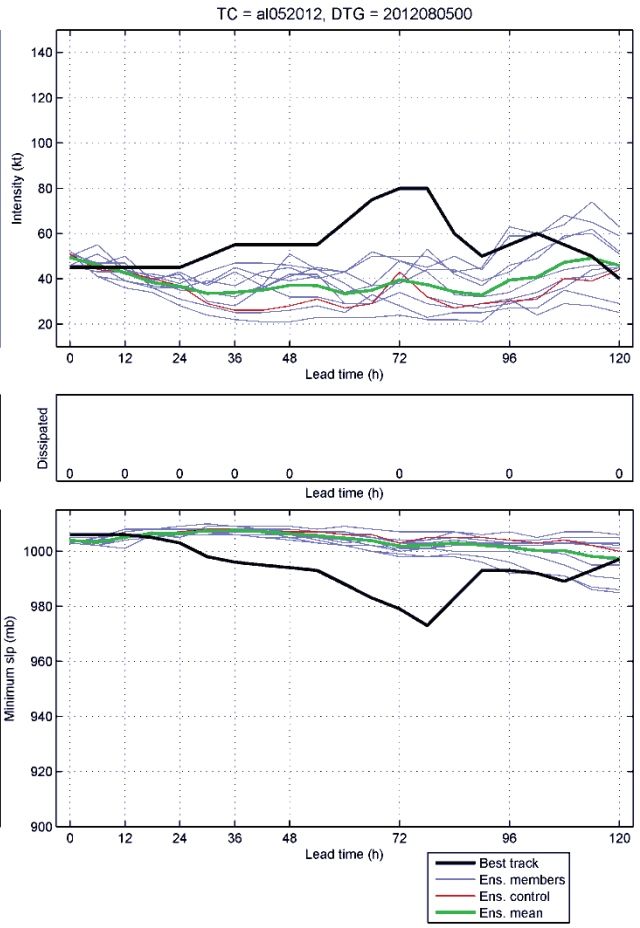
Intensity and Min SLP

Individual model ensembles, Ernesto (05L, 2012)

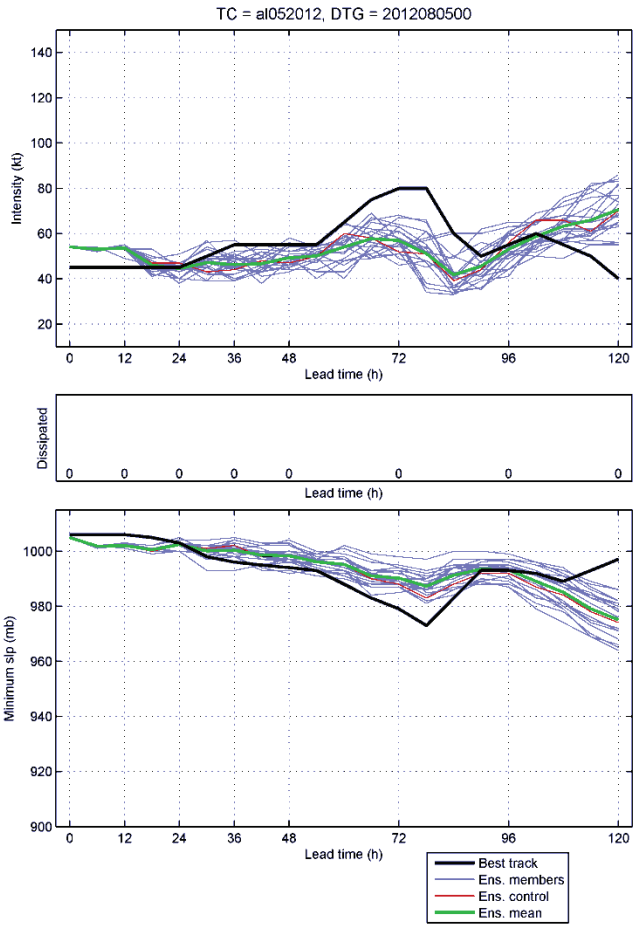
COAMPS-TC



GFDL



HWRF

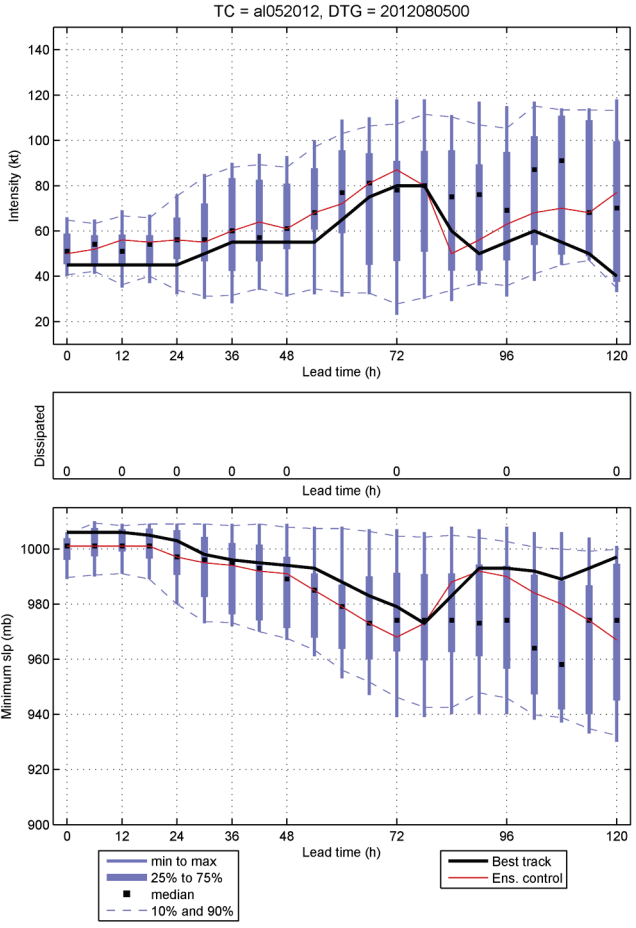


Typically, COAMPS-TC has greater intensity spread than GFDL or HWRF

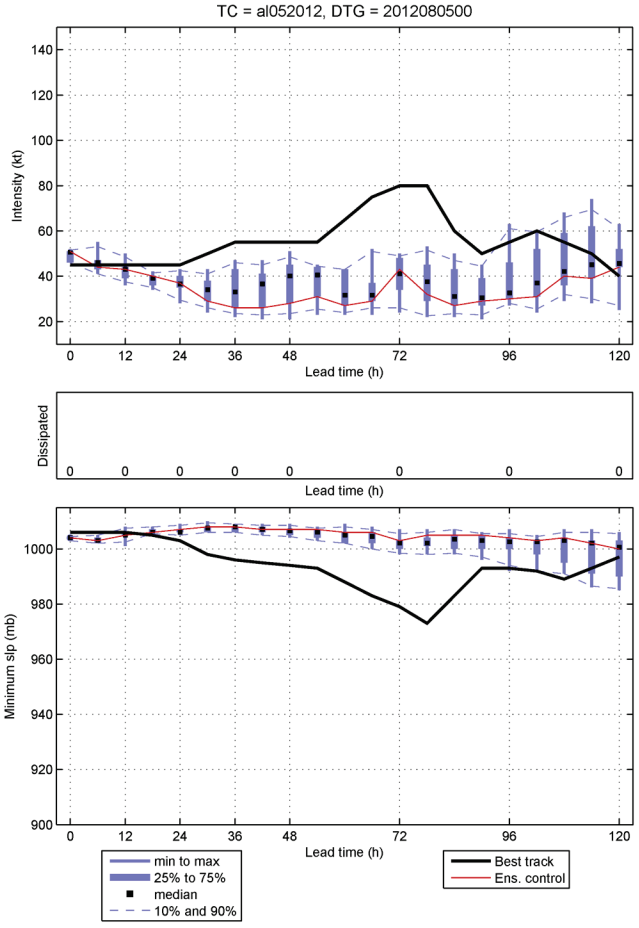
Intensity and Min SLP

Individual model ensembles, Ernesto (05L, 2012)

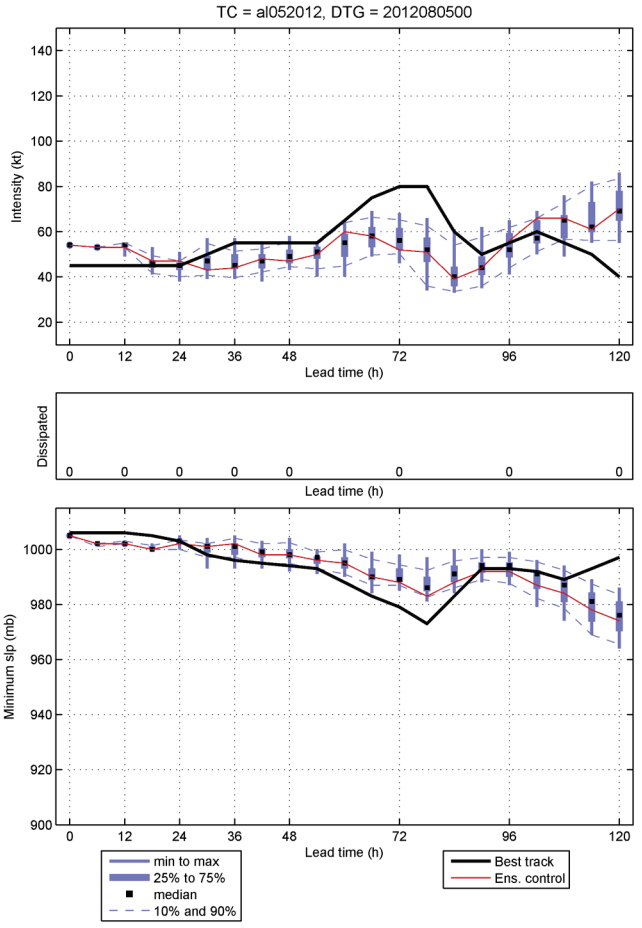
COAMPS-TC



GFDL



HWRF



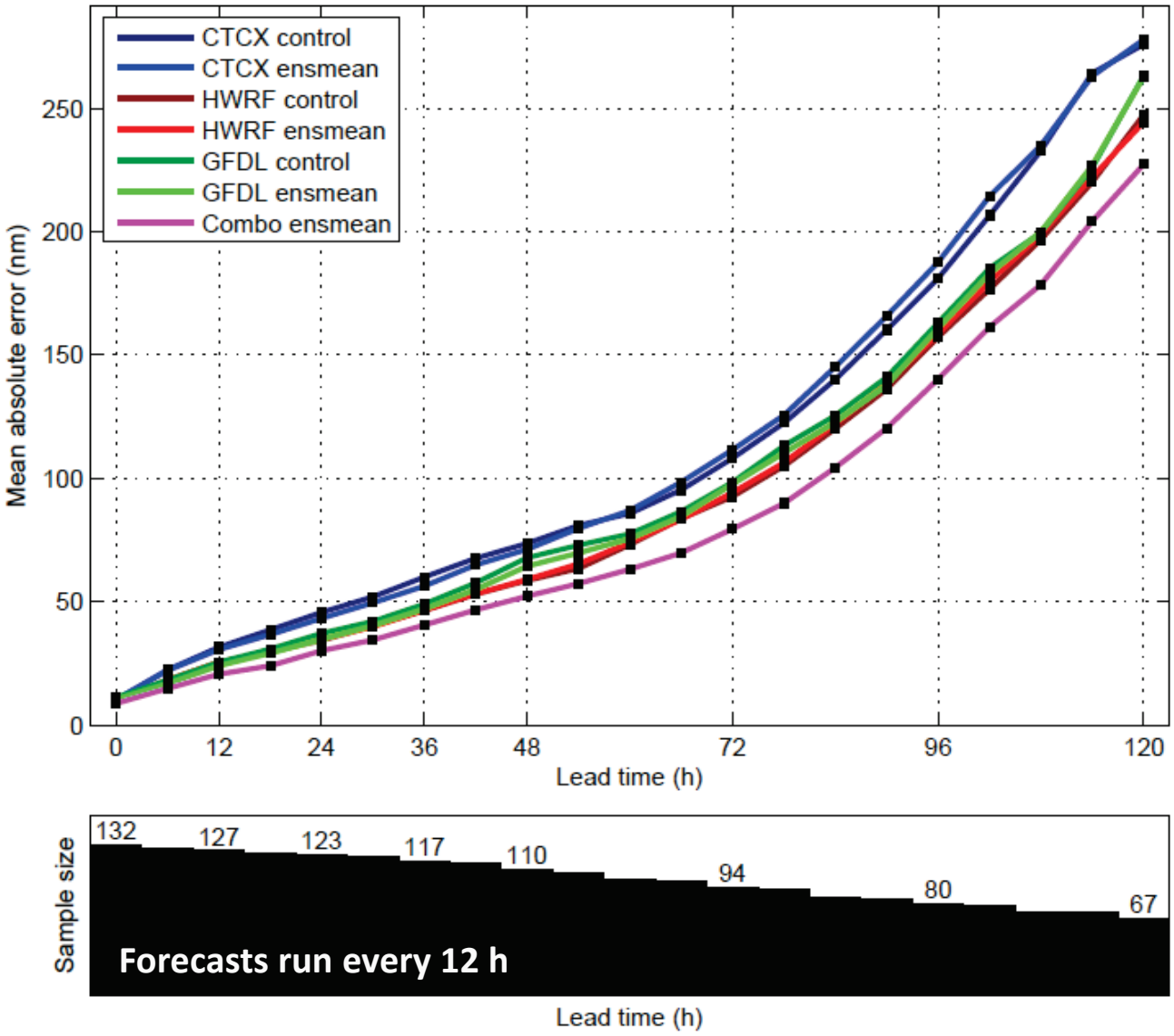
Typically, COAMPS-TC has greater intensity spread than GFDL or HWRF

Deterministic forecast validation

Homogeneous comparison of:

- *COAMPS-TC control*
- *COAMPS-TC ensemble mean (9+ members)*
- *HWRF control*
- *HWRF ensemble mean (17+ members)*
- *GFDL control*
- *GFDL ensemble mean (8+ members)*
- *Combined ensemble (34+ members)*

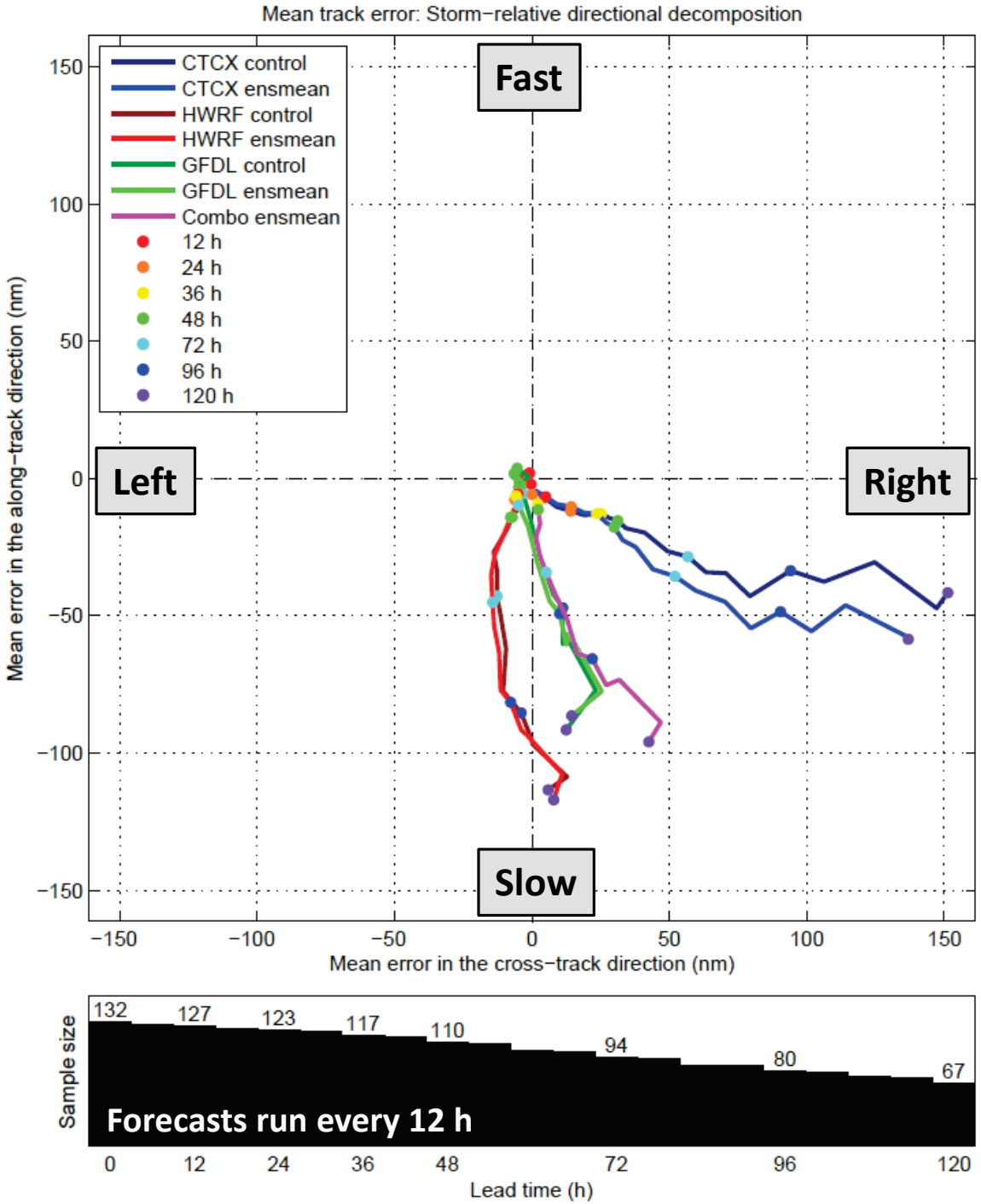
Track mean absolute error



HWRF, COAMPS-TC, and GFDL ensemble means perform similarly to their respective controls

Combined ensemble mean has lower MAE than HWRF ensemble mean, COAMPS-TC ensemble mean and GFDL ensemble mean at all lead times

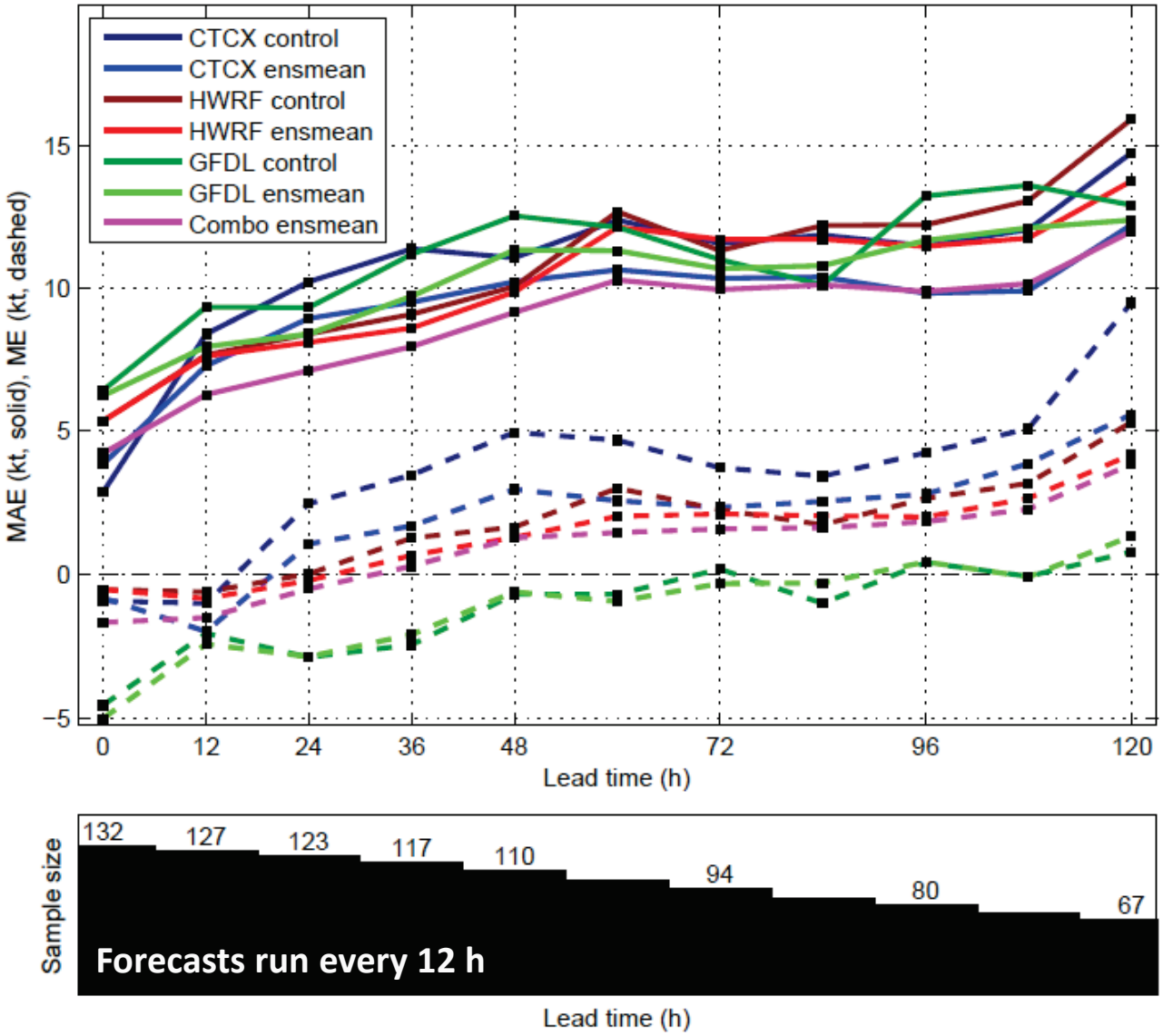
Track mean error



Ensemble means have basically the same biases as the control members

It appears that the ensemble members are primarily making the same type of errors as the control.

Intensity MAE (solid) and ME (dashed)



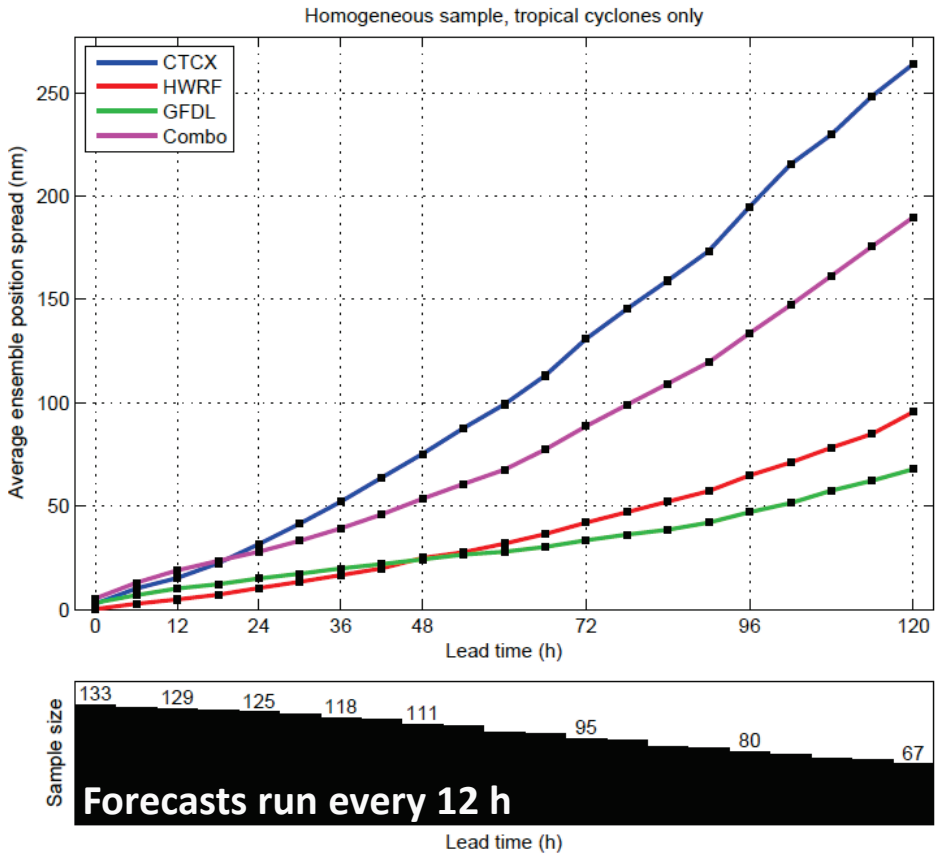
The MAE and ME of the ensemble mean is generally lower than that of the control run

The MAE of the combined ensemble mean is superior to the HWRF, COAMPS-TC, and GFDL ensemble means at nearly all lead times.

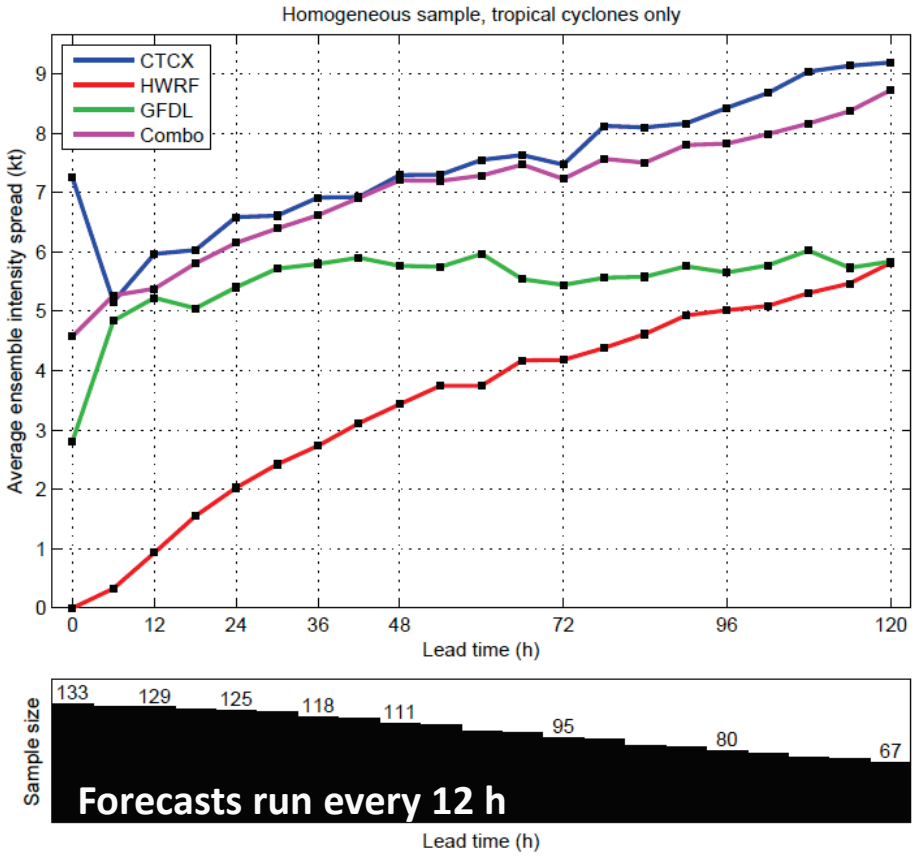
Probabilistic forecast diagnostics

Average forecast spread

Track



Intensity

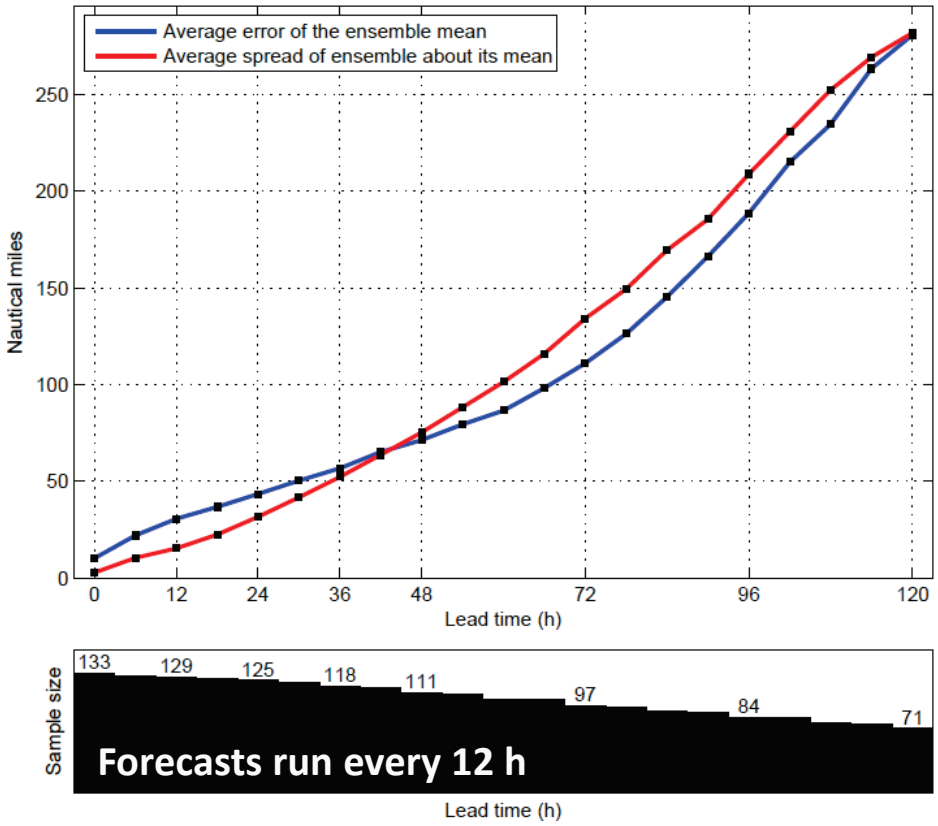


Number of members required for ensemble mean calculation →

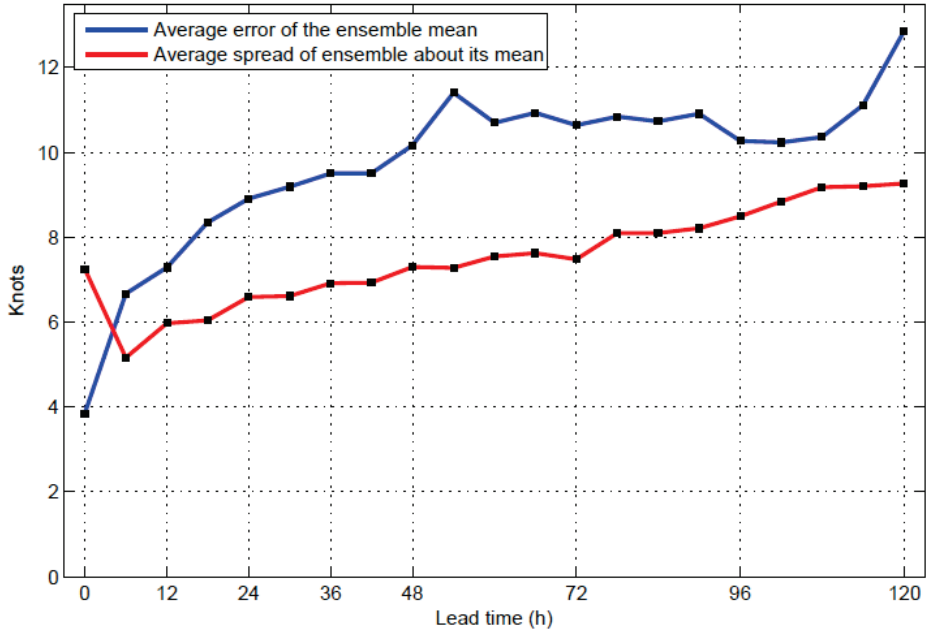
COAMPS-TC	HWRF	GFDL	Combination
9 of 11	17 of 21	8 of 10	34 of 42

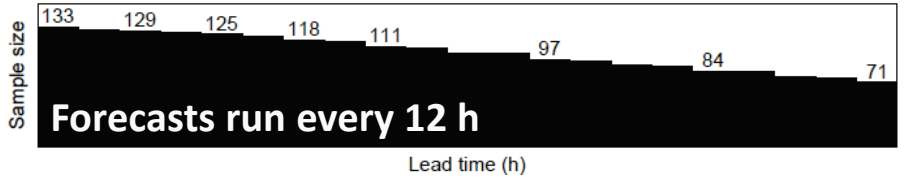
Ensemble spread vs. Error of ensemble mean

COAMPS-TC Track



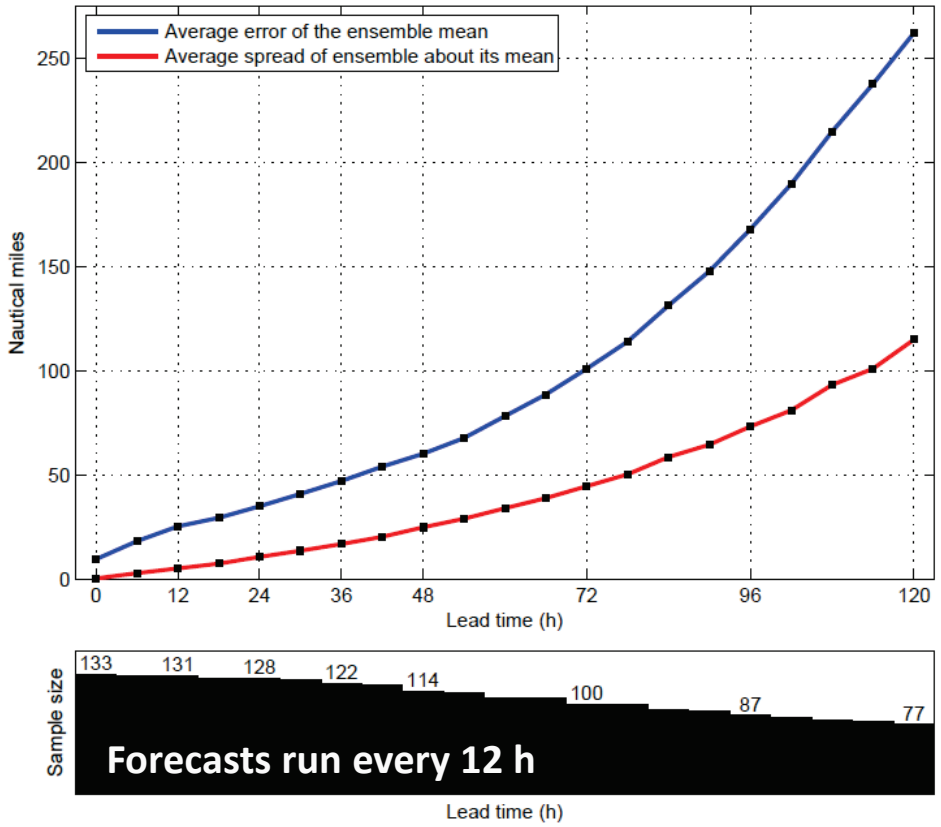
COAMPS-TC Intensity



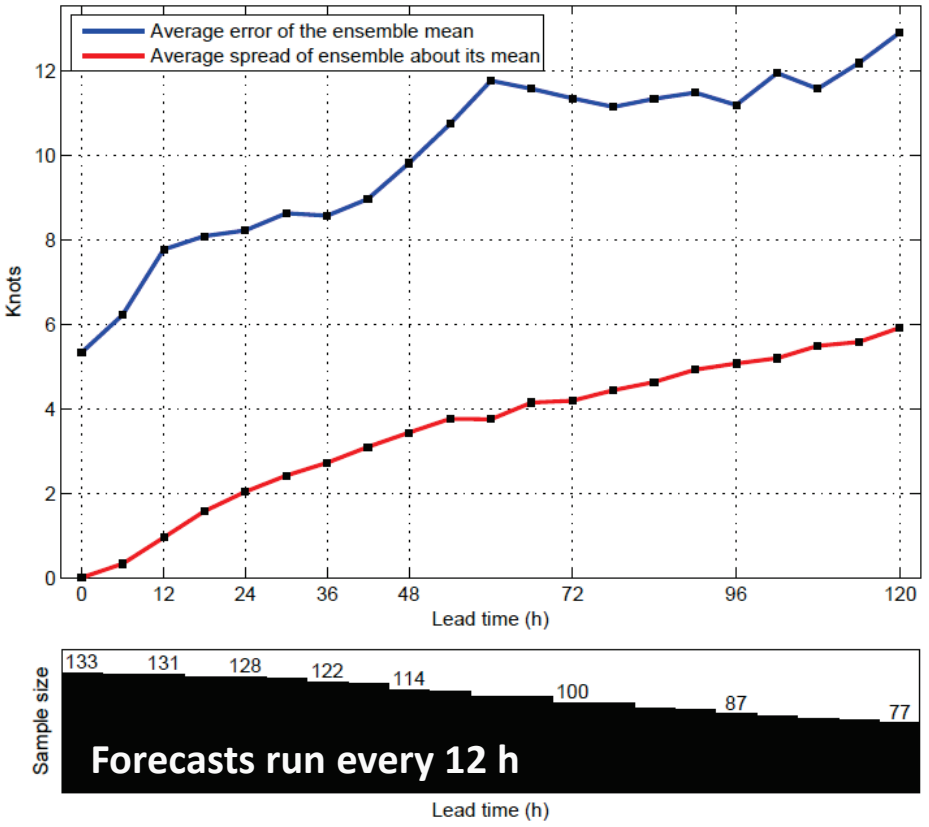


Ensemble spread vs. Error of ensemble mean

HWRF Track

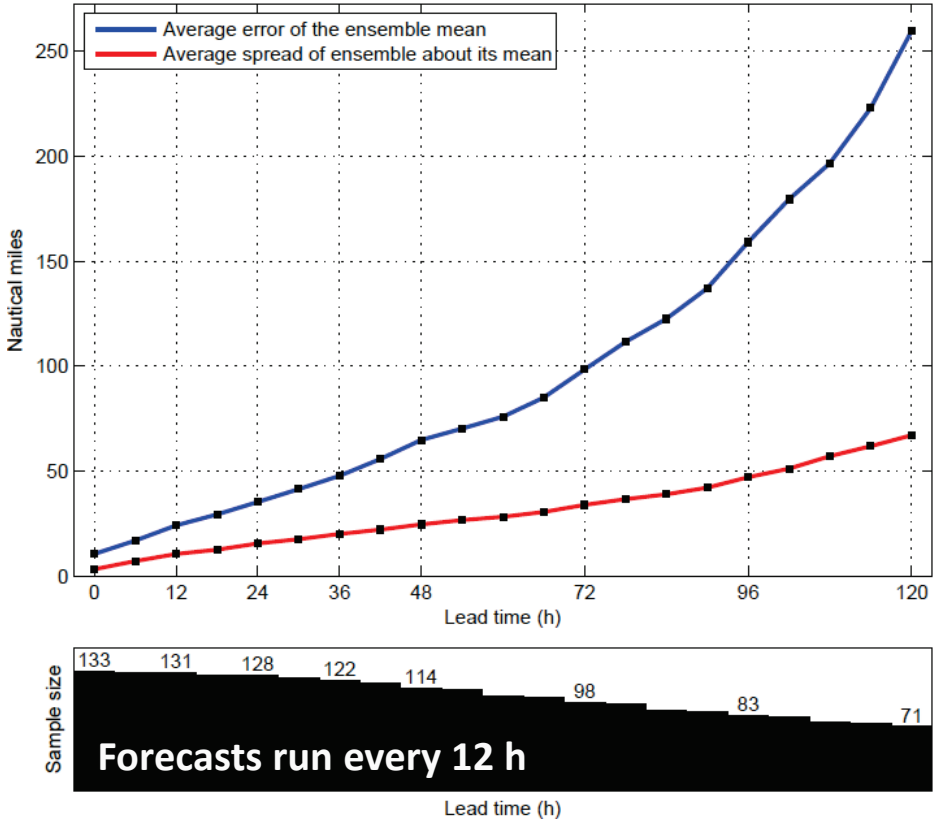


HWRF Intensity

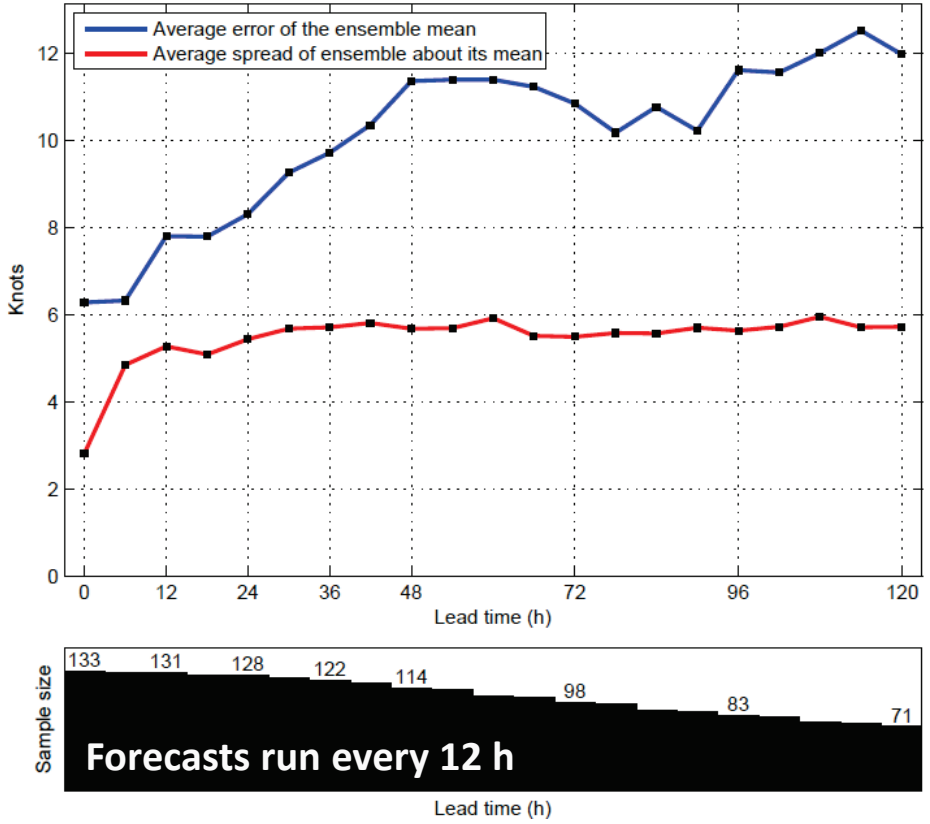


Ensemble spread vs. Error of ensemble mean

GFDL Track

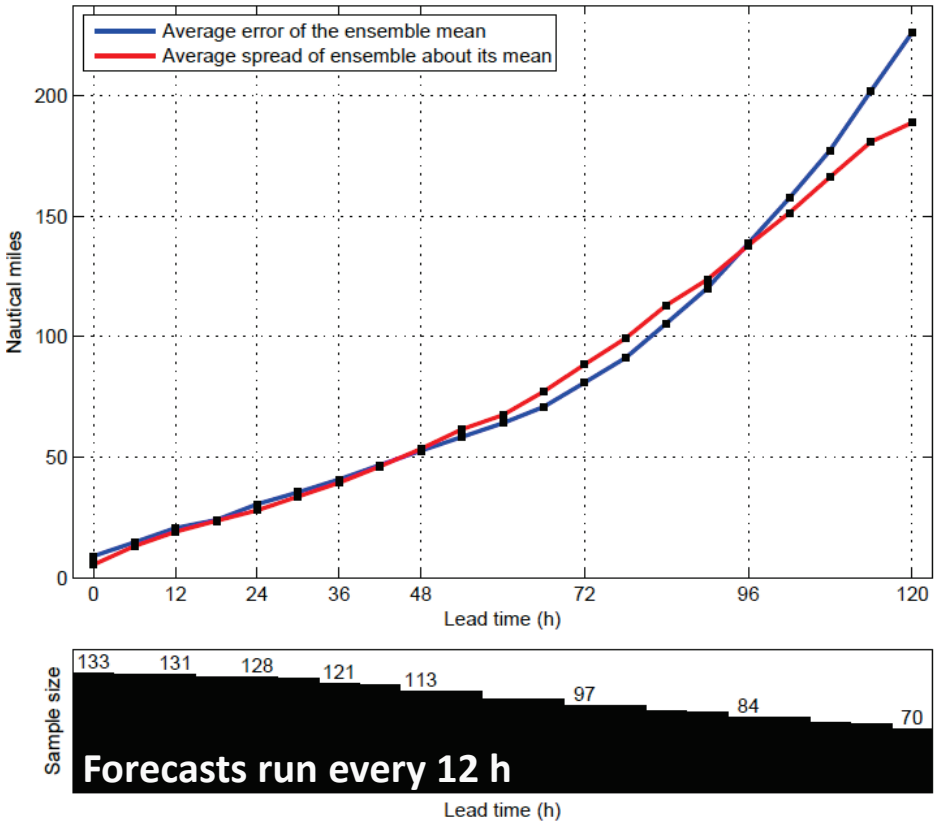


GFDL Intensity

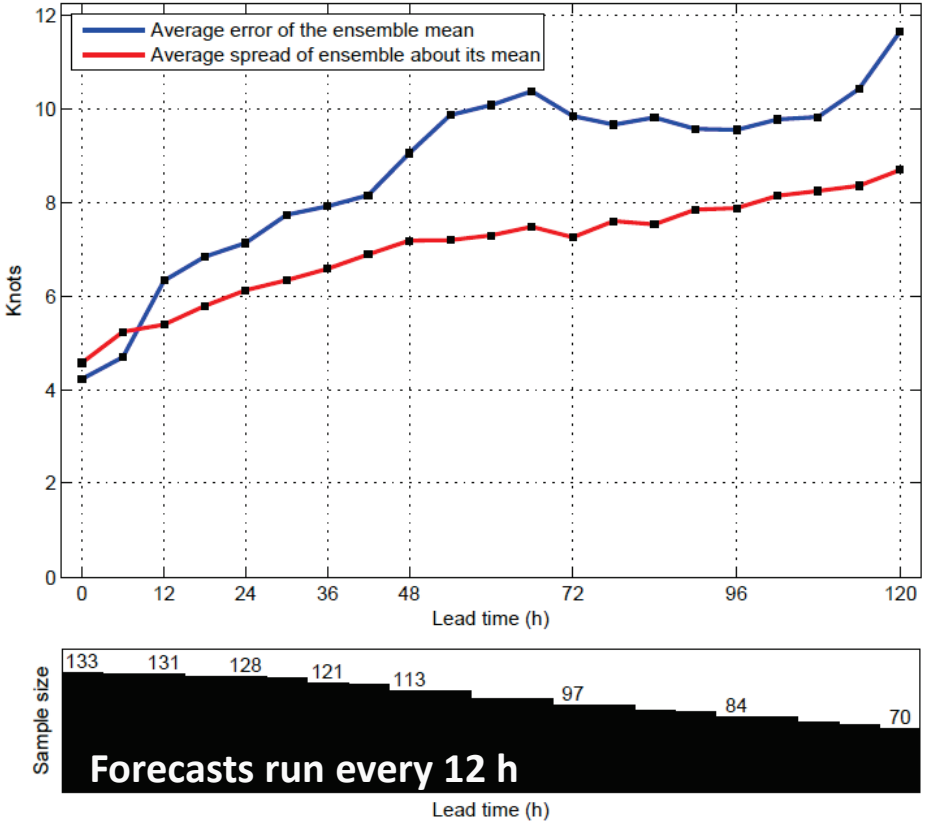


Ensemble spread vs. Error of ensemble mean

Combined Track

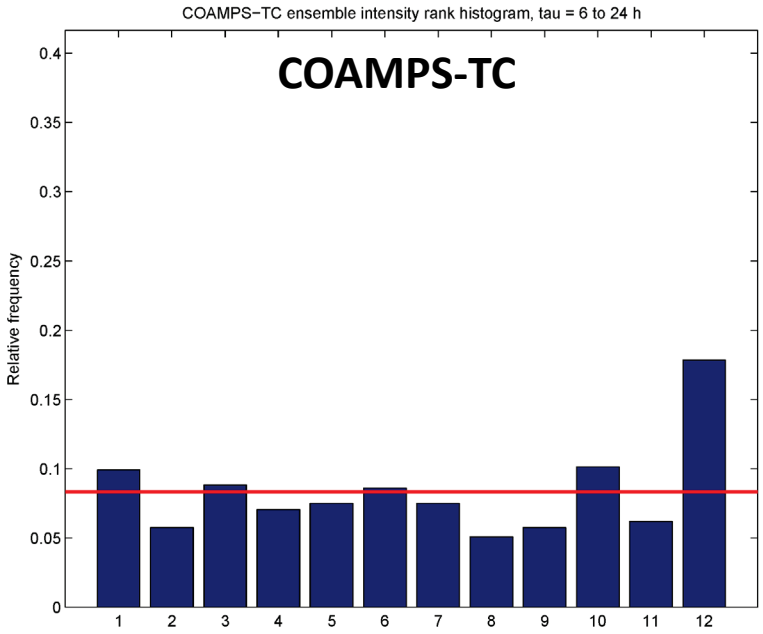


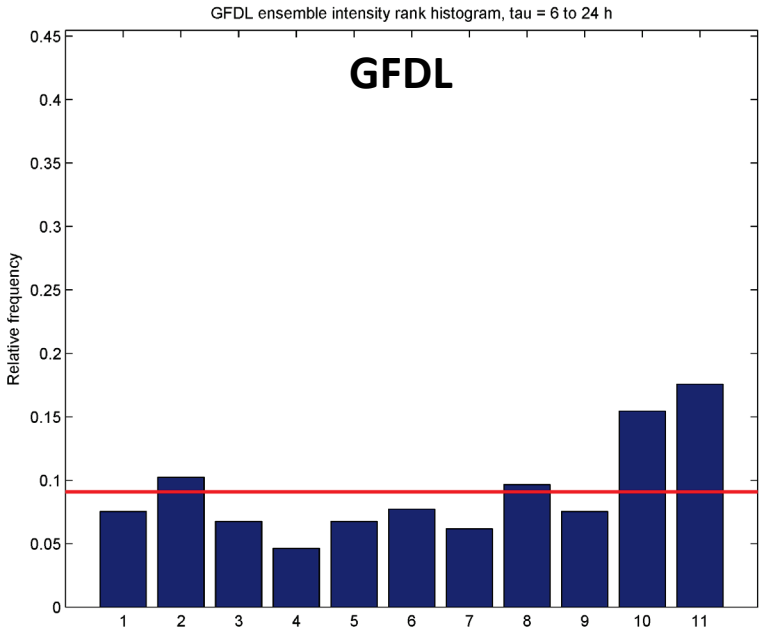
Combined Intensity

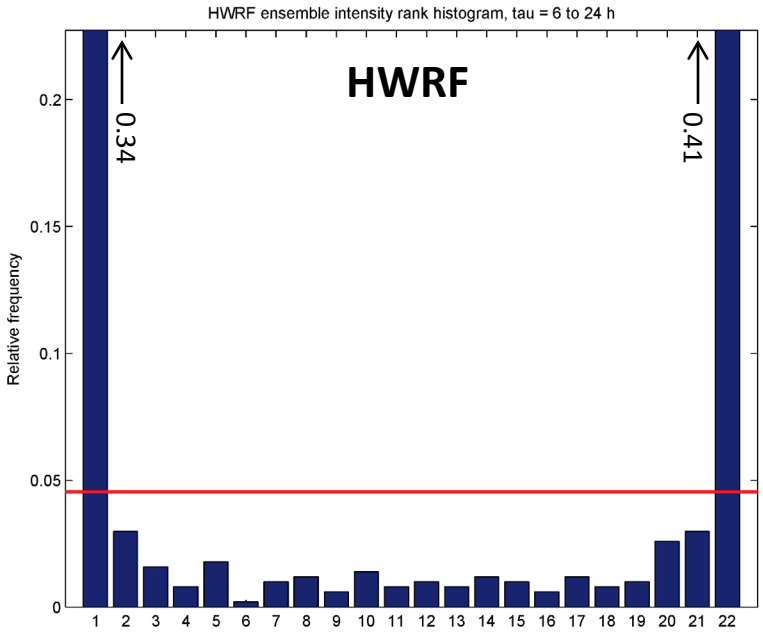


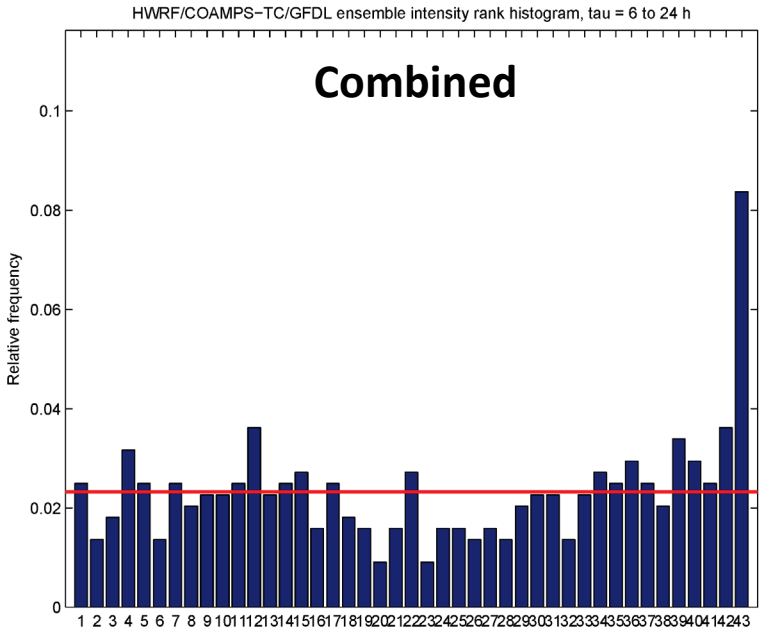
Intensity rank histograms

Tau = 6 to 24



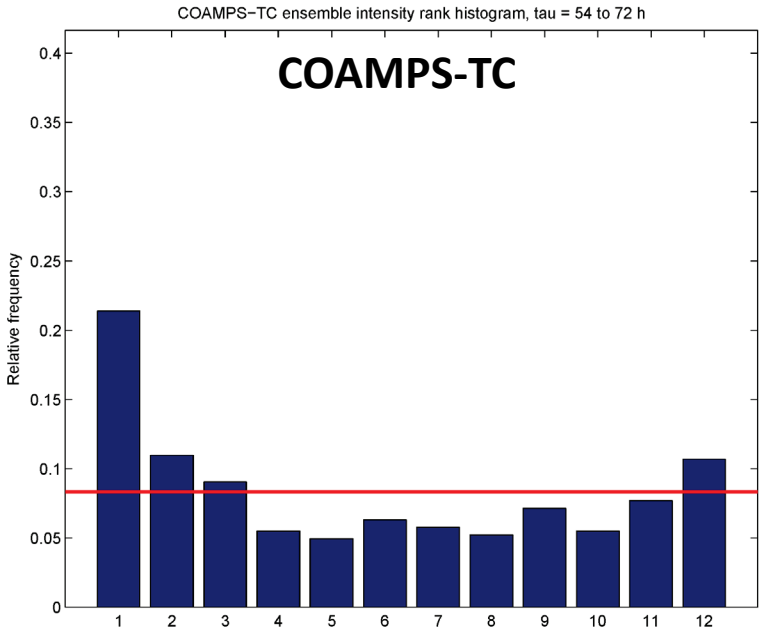


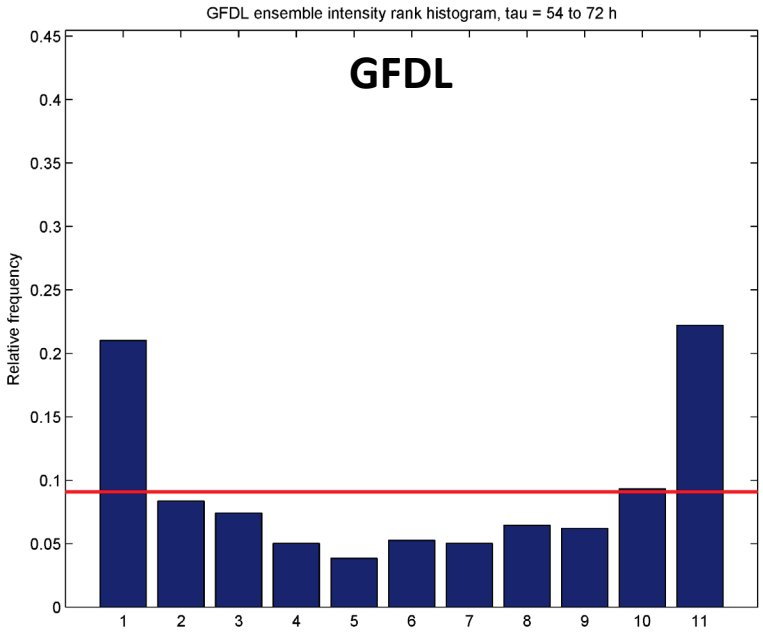


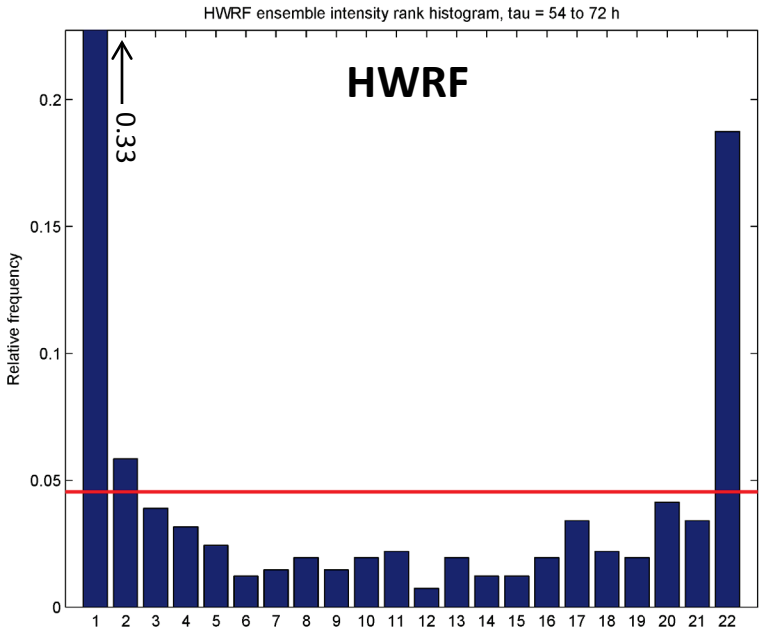


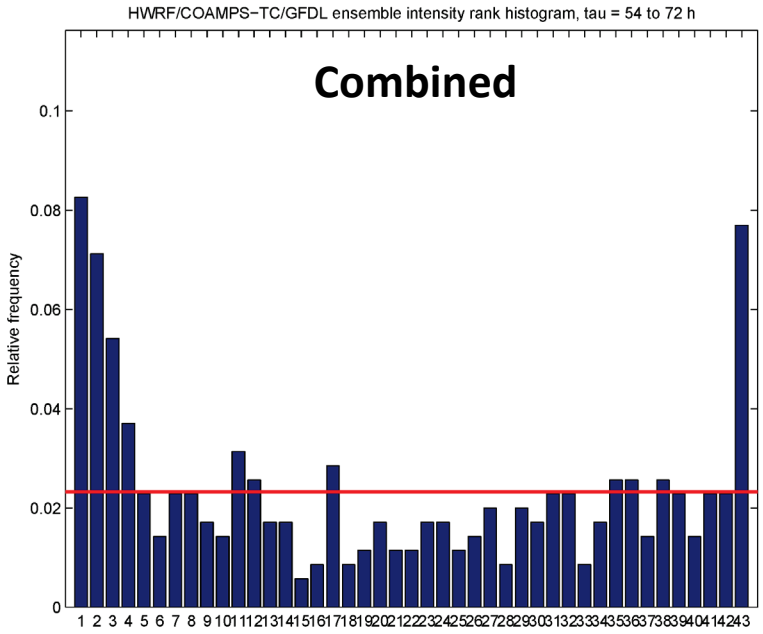
Intensity rank histograms

Tau = 54 to 72



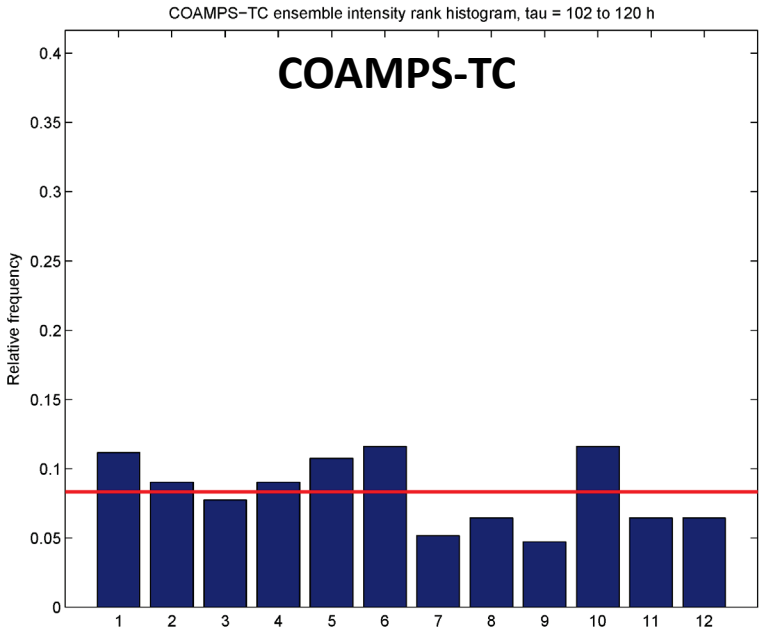


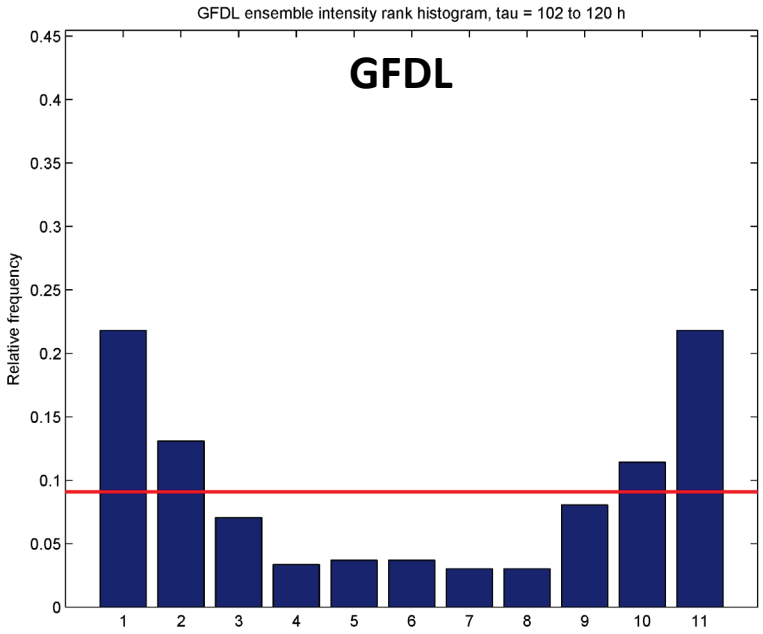


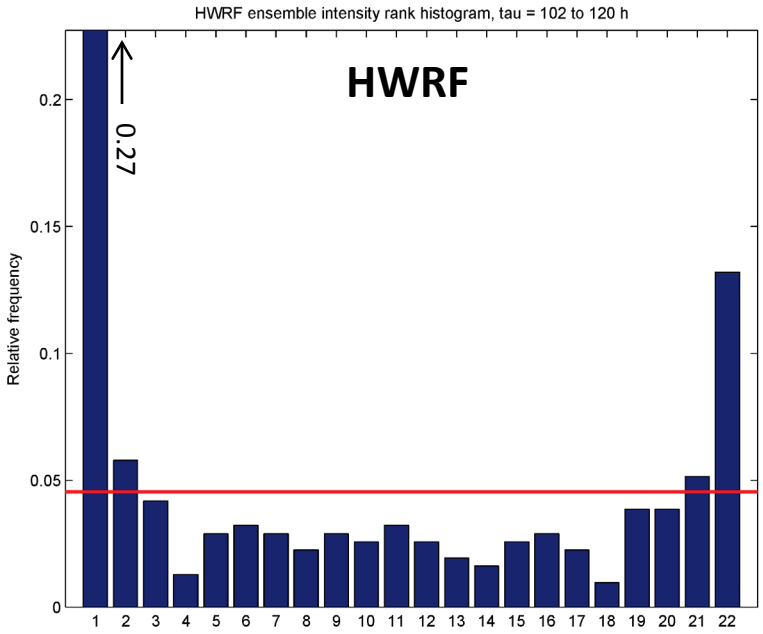


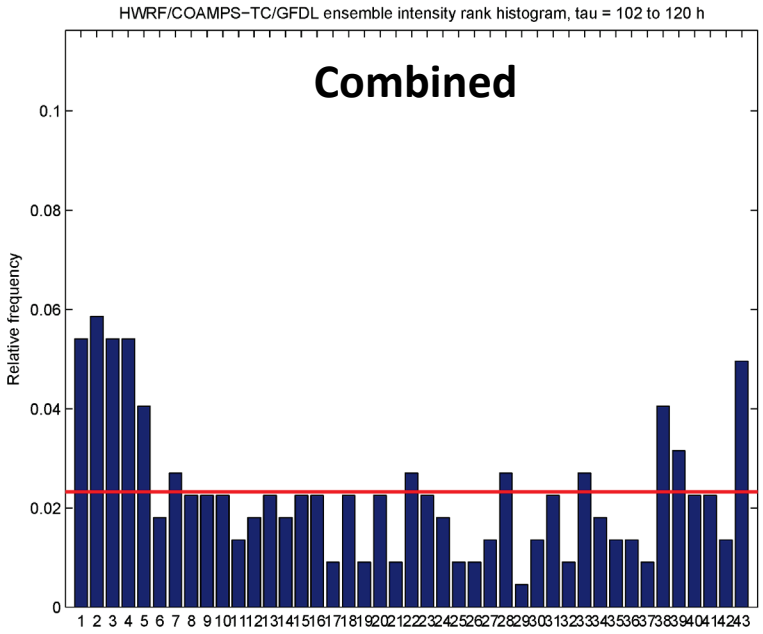
Intensity rank histograms

Tau = 102 to 120









Summary

- HWRF/COAMPS-TC/GFDL combined ensemble mean outperforms its component single model ensemble means in track MAE and intensity MAE

- For track, combined ensemble spread is similar to the MAE of the ensemble mean. However, for intensity, spread-skill plots and rank histograms show underdispersion

- Work is ongoing at NRL to apply additional probabilistic verification approaches to the individual and combined ensemble forecasts:

  Track reliability: Does the observed position fall within the 2/3 probability ellipse for 2/3 of the forecast cases?

  Summary scores for intensity: Continuous Ranked Probability Score, which facilitates comparison of probabilistic and deterministic forecasts