NOAA HURRICANE FORECAST IMPROVEMENT PROJECT

Why Ensembles?

Bob Gall

Miami November 7, 2011

Purpose

- There is still considerable debate on the value of Ensemble systems
  - Question– Which is more valuable
    - » High resolution deterministic model
    - » Lower resolution Ensemble
  - Question– What is the value of ensembles
    - Usual Answer—you can get probability information
    - But what else?
- I'll try to explore some of these questions.

Note

- Discussion here really presents nothing new
  - However the hurricane problem makes it easier to demonstrate some of the issues
    - Main interest is position and intensity (the three numbers)
      - Unlike mid-latitude prediction where there is a plethora of things to be predicted
      - Basically we are tracking a single low and its strength.

A Basic Principle for this Discussion

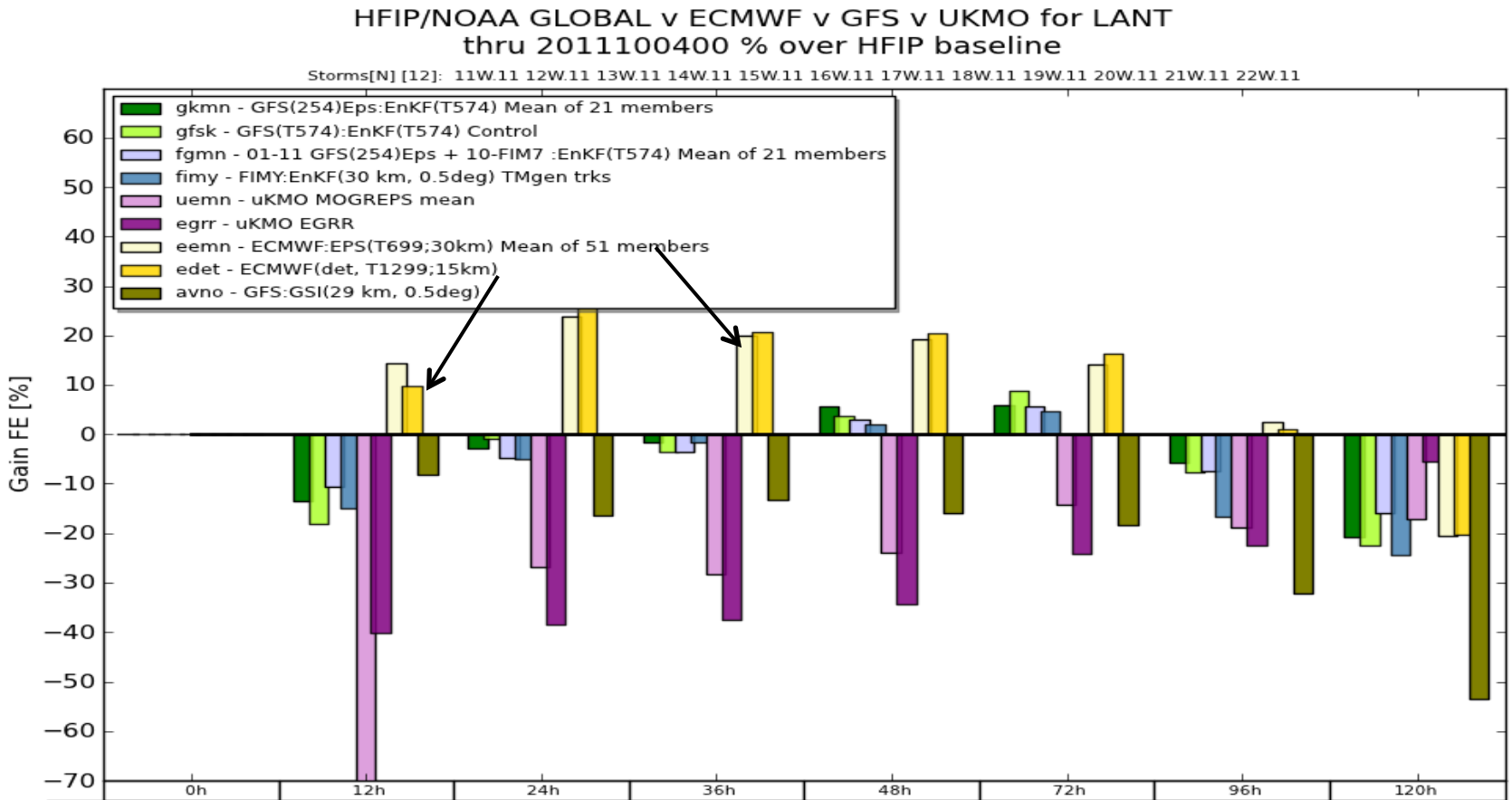
A single forecast by any model using any given initial state represents a single member of some virtual ensemble.

Comment

- This principle needs to be in the back of the mind throughout this discussion
  - It is a no brainer but tends to be forgotten.

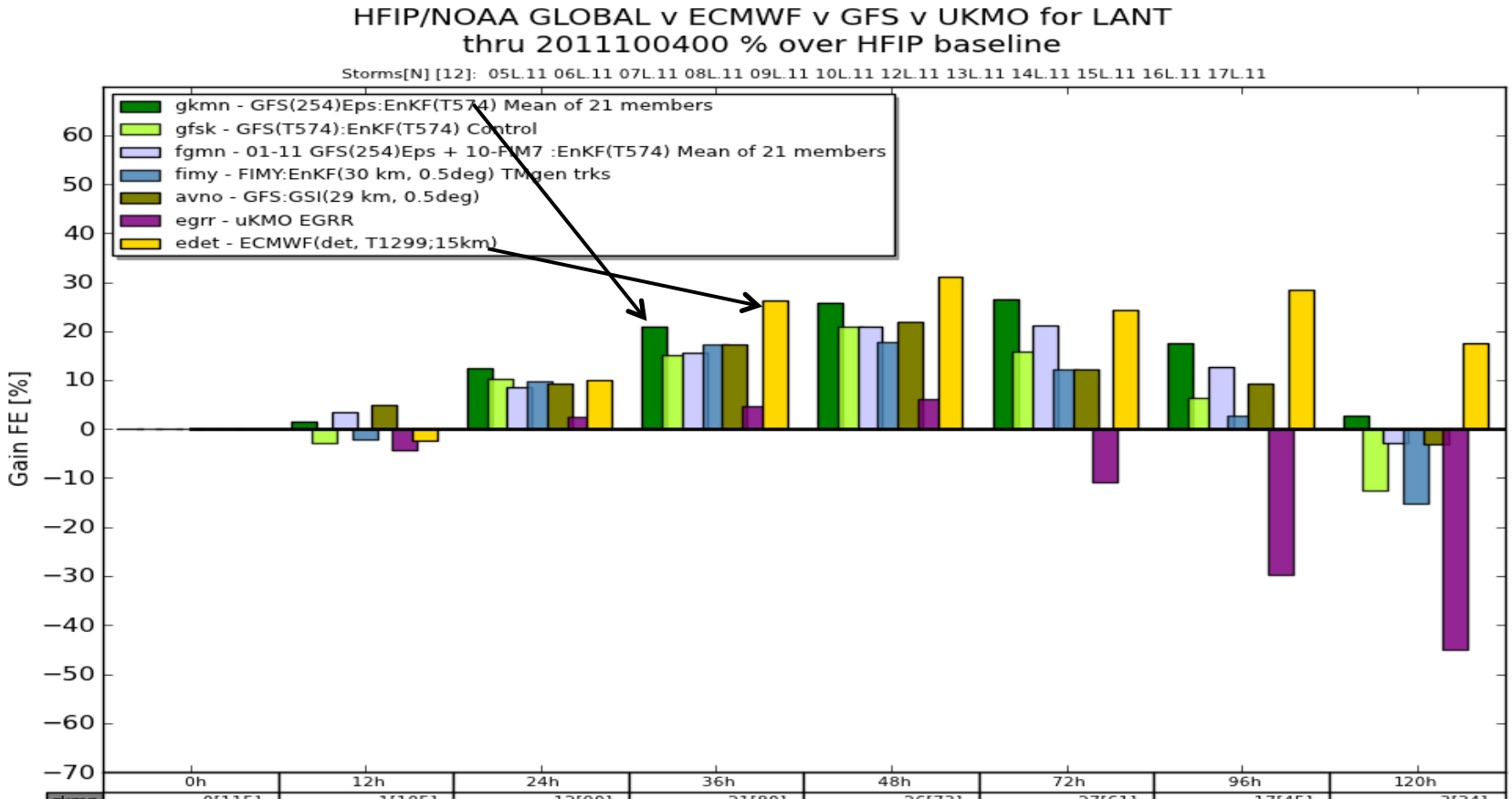
- A crude rule of thumb:
  - ***A deterministic model run at twice the resolution of an ensemble will give about the same AVERAGE error as the mean of the ensemble.***
    - (in this case the deterministic model will be cheaper if the ensemble has more than about 10 members)

Track Error of Models
(% Improvement over baseline)



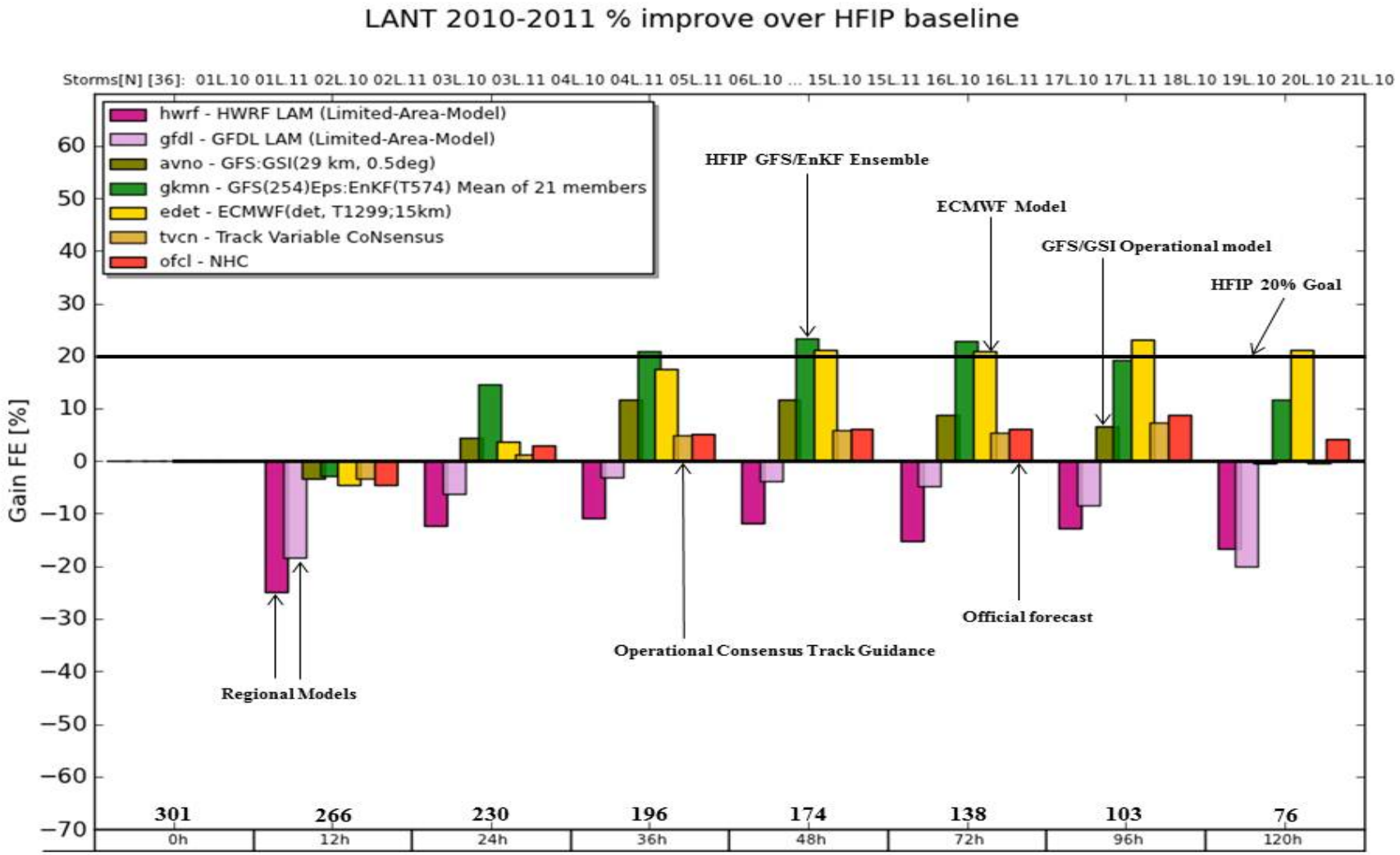
	0h	12h	24h	36h	48h	72h	96h	120h
gkmn	0[109]	-14[105]	-3[94]	-2[81]	6[72]	6[50]	-6[42]	-21[30]
gfsk	0[109]	-18[105]	-1[94]	-4[81]	4[72]	9[50]	-8[42]	-22[30]
fgmn	0[109]	-11[105]	-5[94]	-4[81]	3[72]	6[50]	-7[42]	-16[30]
fimy	0[109]	-15[105]	-5[94]	-2[81]	2[72]	5[50]	-17[42]	-24[30]
uemn	0[109]	-76[105]	-27[94]	-28[81]	-24[72]	-14[50]	-19[42]	-17[30]
egrr	0[109]	-40[105]	-38[94]	-38[81]	-34[72]	-24[50]	-22[42]	-5[30]
eemn	0[109]	14[105]	24[94]	20[81]	19[72]	14[50]	3[42]	-21[30]
edet	0[109]	10[105]	28[94]	21[81]	20[72]	16[50]	1[42]	-20[30]
avno	0[109]	-8[105]	-17[94]	-13[81]	-16[72]	-18[50]	-32[42]	-54[30]

Track Error of Models
(% Improvement over baseline cont.)



	0h	12h	24h	36h	48h	72h	96h	120h
gkmn	0[115]	1[105]	12[90]	21[80]	26[73]	27[61]	17[45]	3[34]
gfsk	0[115]	-3[105]	10[90]	15[80]	21[73]	16[61]	6[45]	-13[34]
fgmn	0[115]	3[105]	9[90]	16[80]	21[73]	21[61]	13[45]	-3[34]
fimy	0[115]	-2[105]	10[90]	17[80]	18[73]	12[61]	3[45]	-15[34]
avno	0[115]	5[105]	9[90]	17[80]	22[73]	12[61]	9[45]	-3[34]
egrr	0[115]	-4[105]	2[90]	5[80]	6[73]	-11[61]	-30[45]	-45[34]
edet	0[115]	-2[105]	10[90]	26[80]	31[73]	24[61]	29[45]	17[34]

Track Error of Models
(% Improvement over baseline cont.)

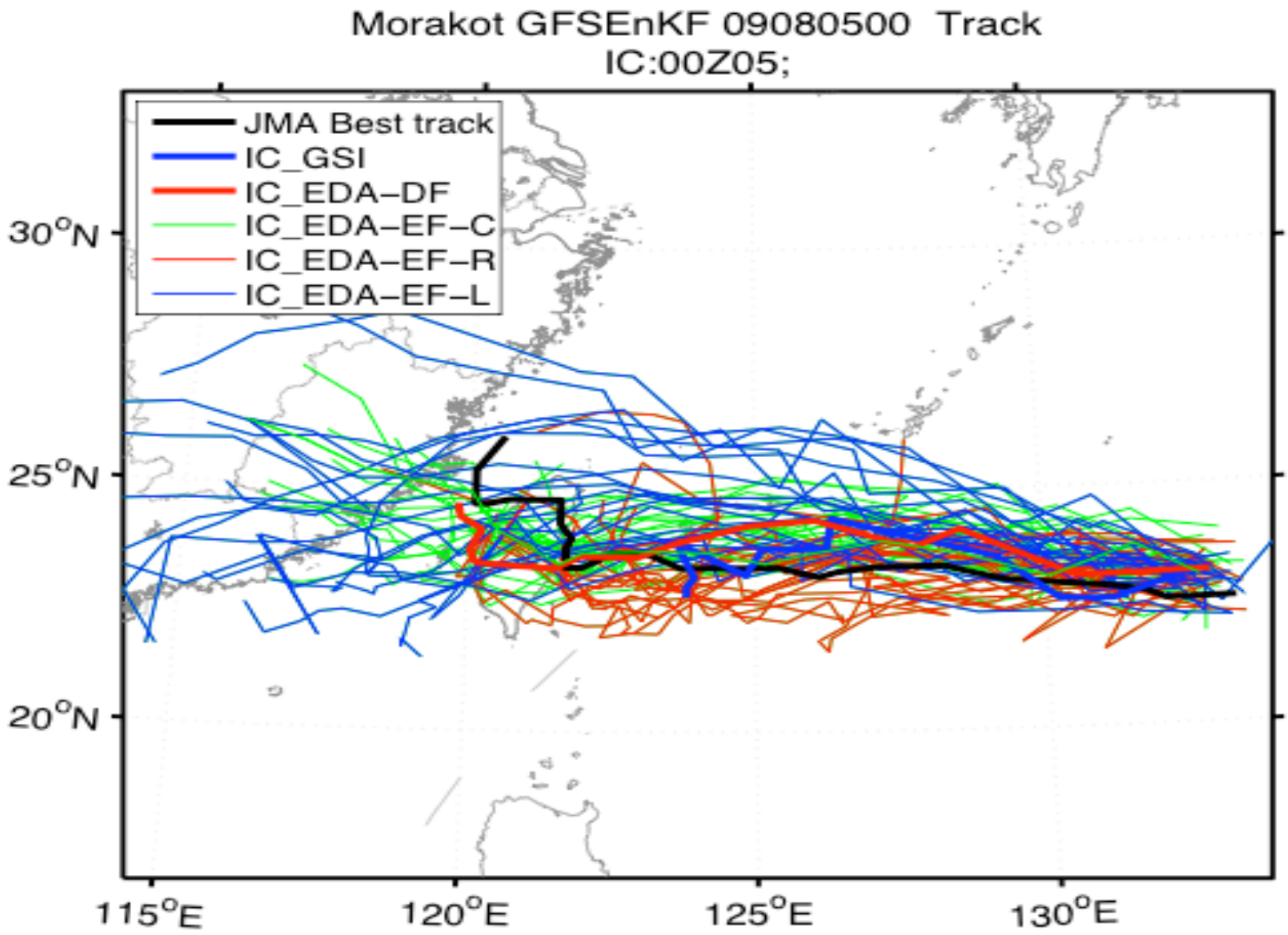


But remember the “Principle” stated above

A single forecast by any model using any given initial state represents a single member of some virtual ensemble

And so could be any one of those members

Example of an Ensemble Forecast



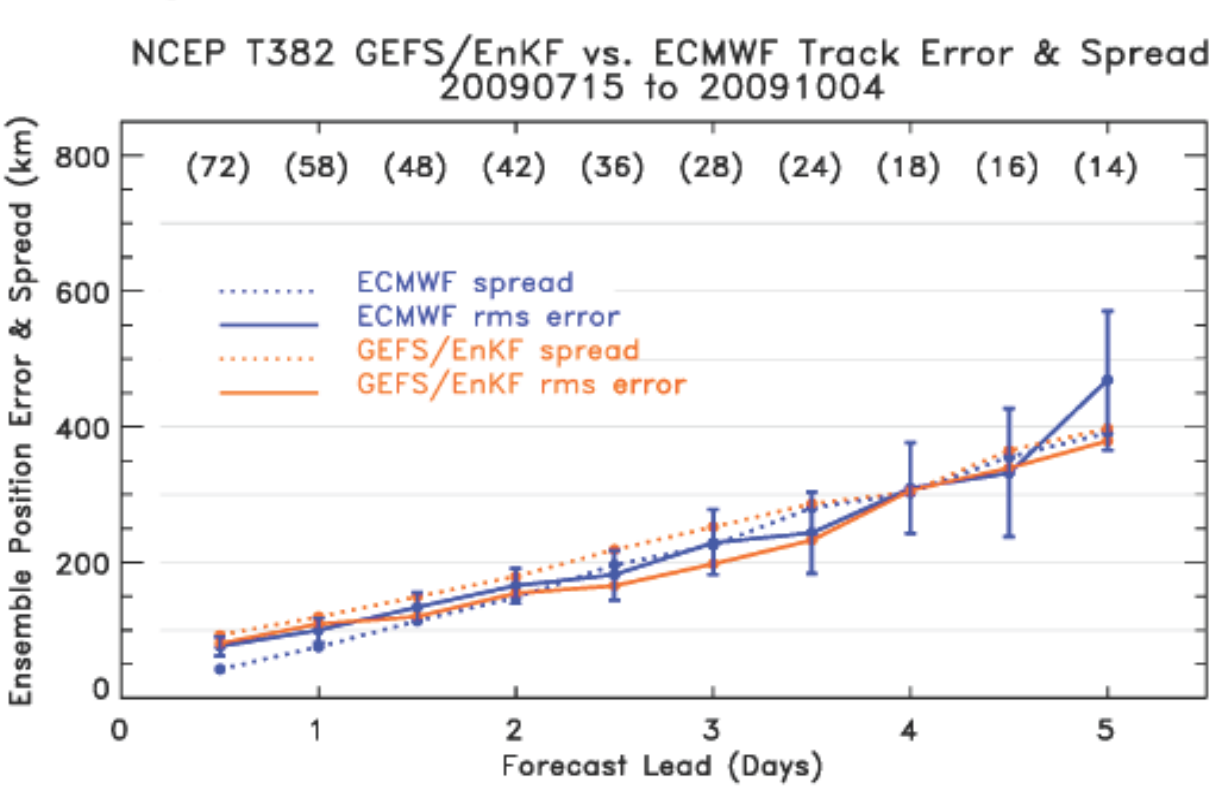
In fact a rule of thumb for an ensemble is:

The "spread" about the mean should be about the same as the error of the mean

- The "Spread" is of course similar to what is shown as the "cone of uncertainty" on the NHC forecasts

- Even though the average error of a deterministic model may be similar to a lower resolution ensemble mean, the actual individual forecast is likely to be anywhere within this spread.

T382 GEFS/EnKF vs. operational T399 ECMWF



competitive with ECMWF in position error

10

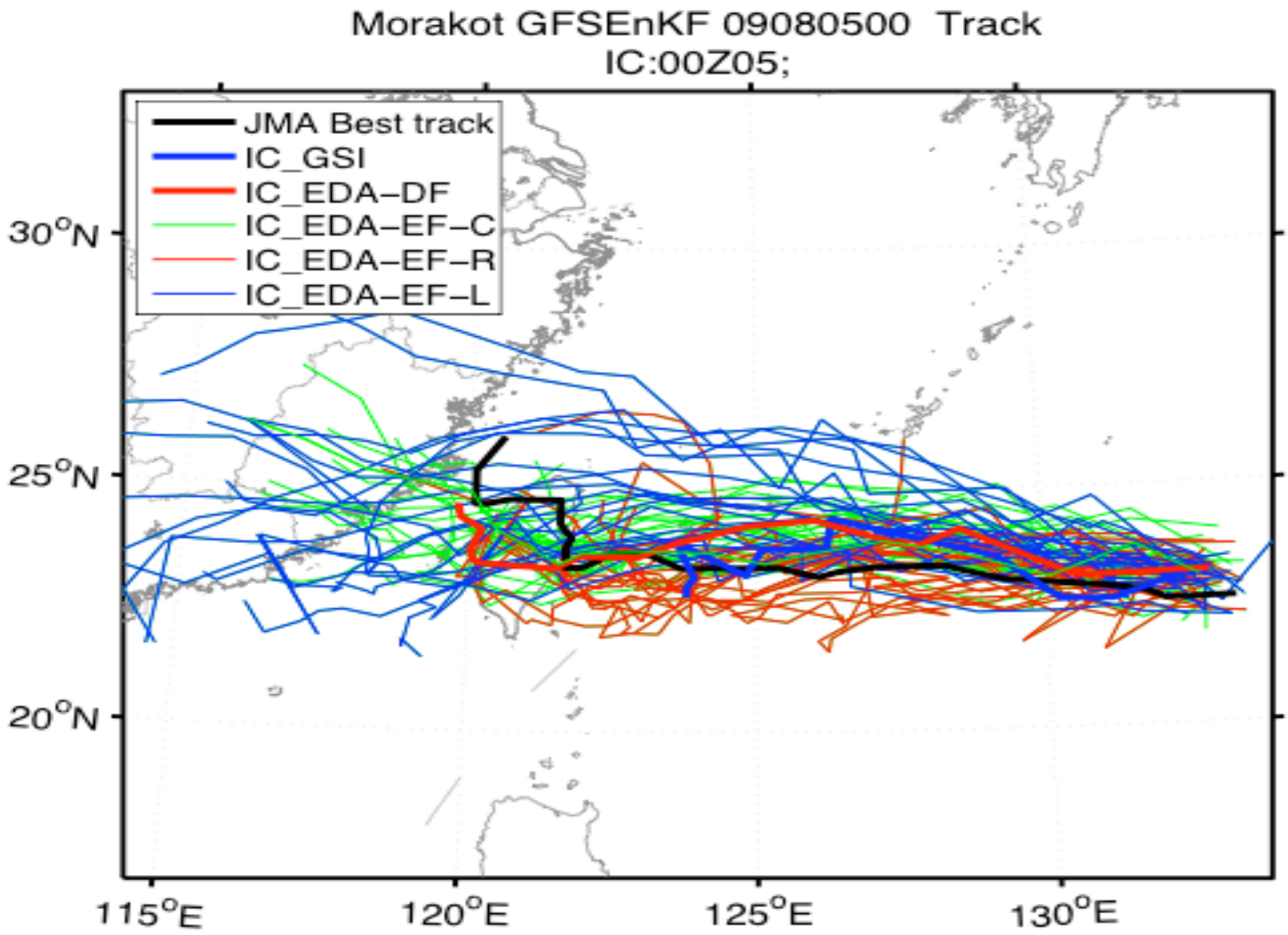
ECMWF Ensemble (50 members) for Rina 1200Z
October 25, 2011

Other Uses

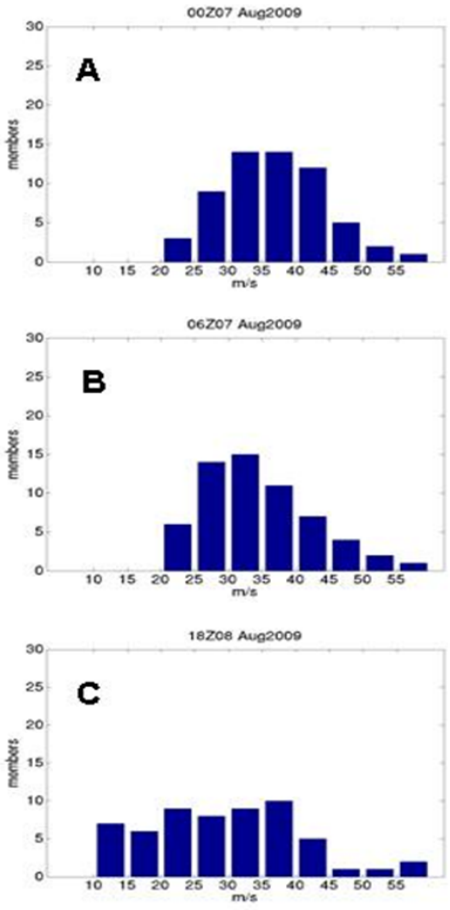
Examination of behavior of individual ensemble members can help identify causes of uncertainty

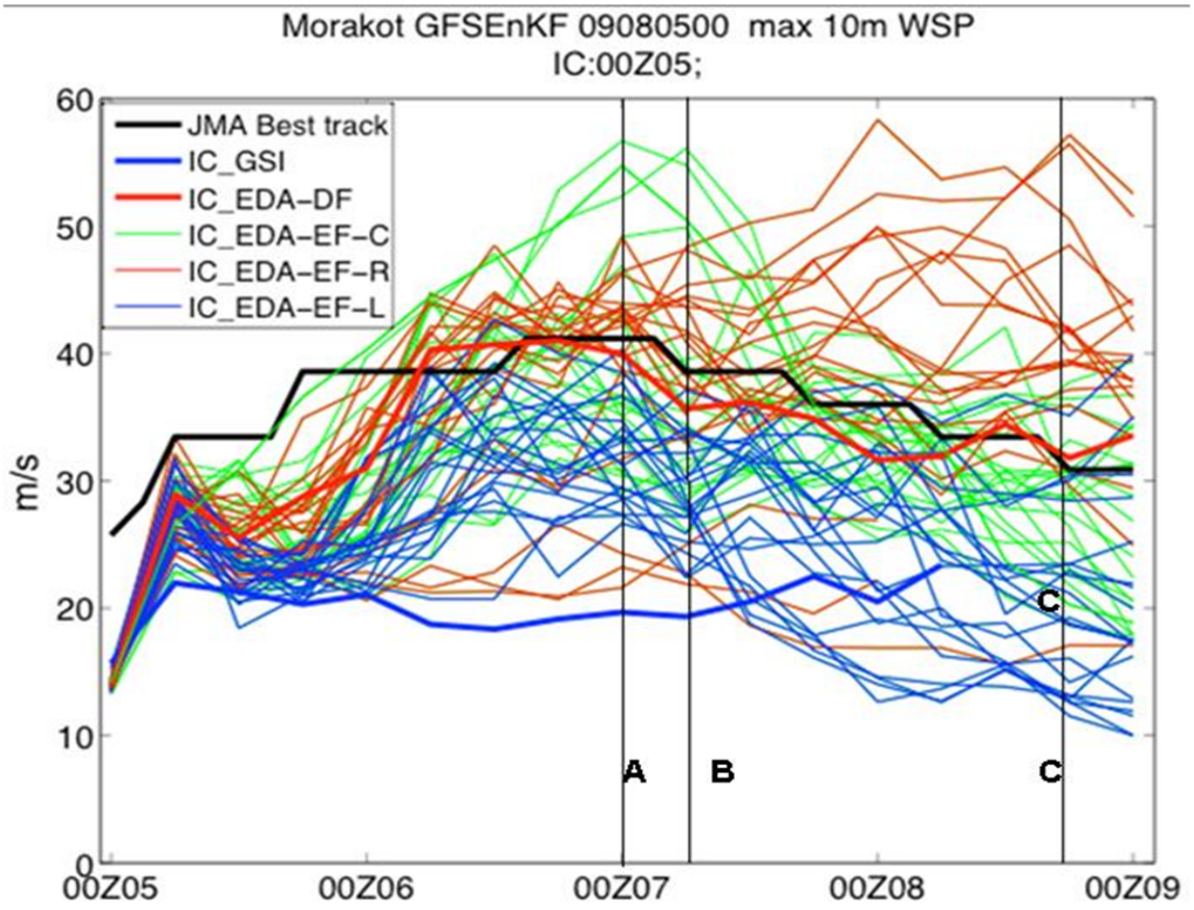
- For example comparison with other aspects of a global model forecast such as the location and speed of fronts
- Refer to the previous slide
- Dependency of intensity on track (even over open ocean)

Example of an Ensemble Forecast



Example of an Ensemble Forecast (Cont.)



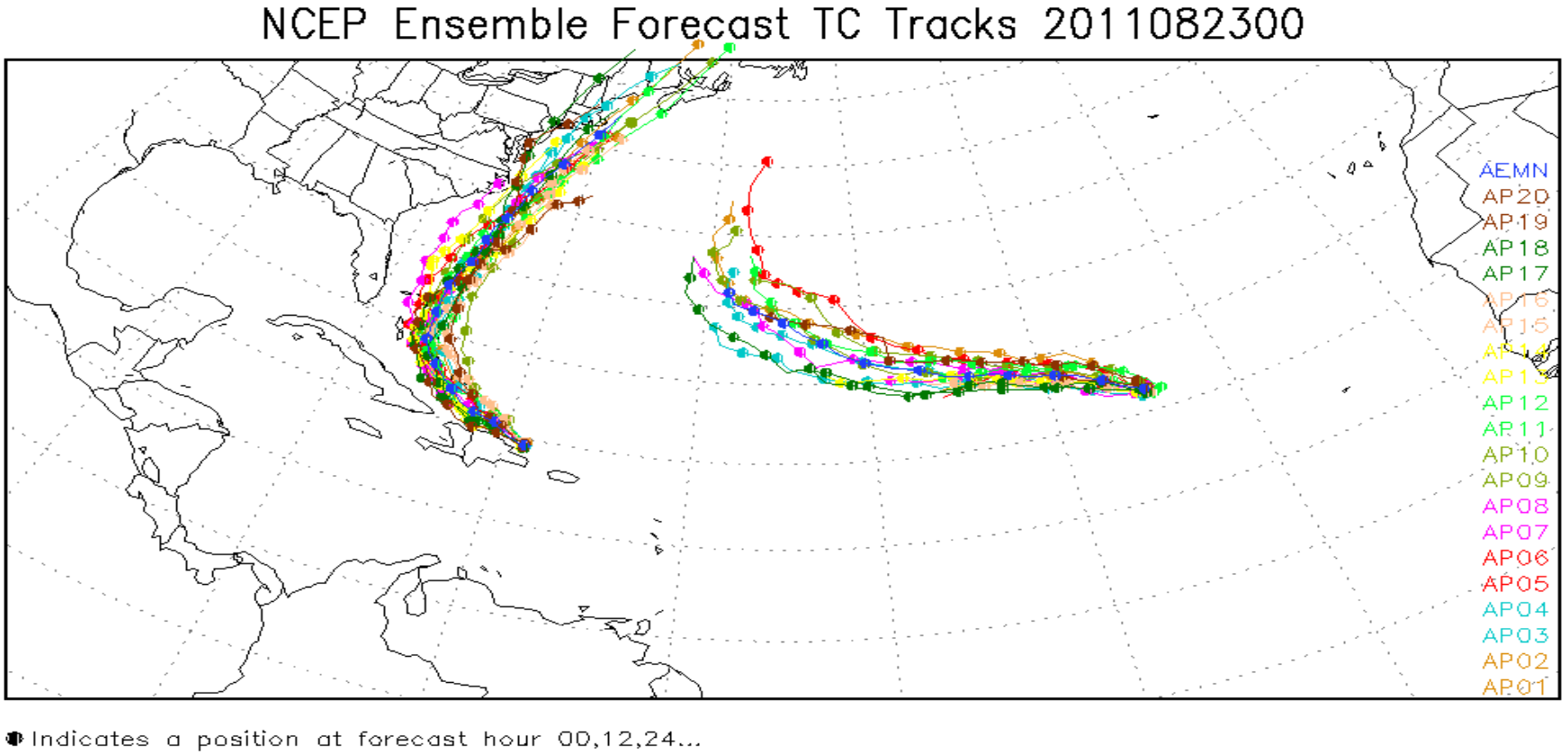


Multi model ensembles

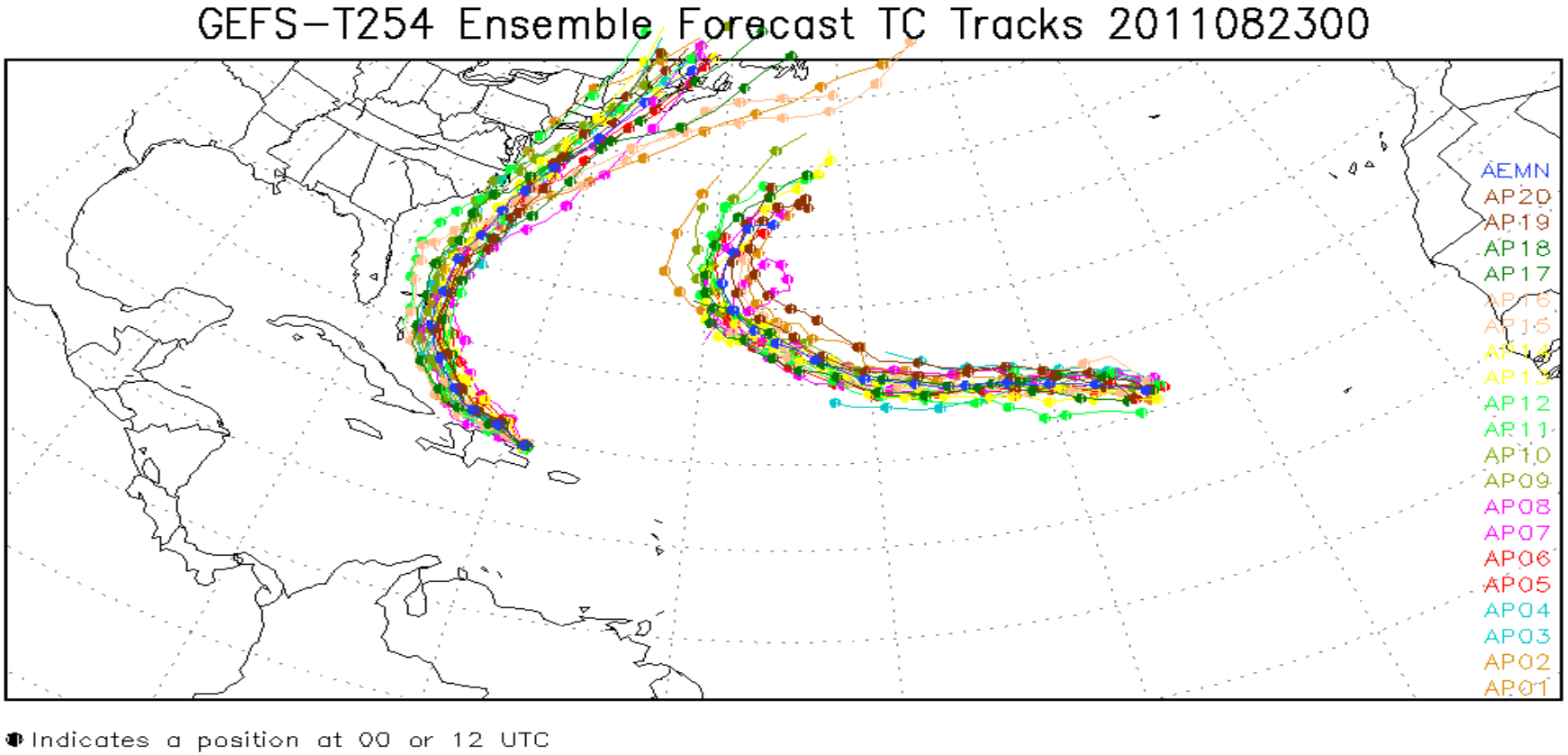
Generally known to provide a more accurate ensemble mean than the mean of ensembles the individual member models.

- But comparing the variation in the individual model ensembles with each other gives some estimate of the various model's errors for each case.

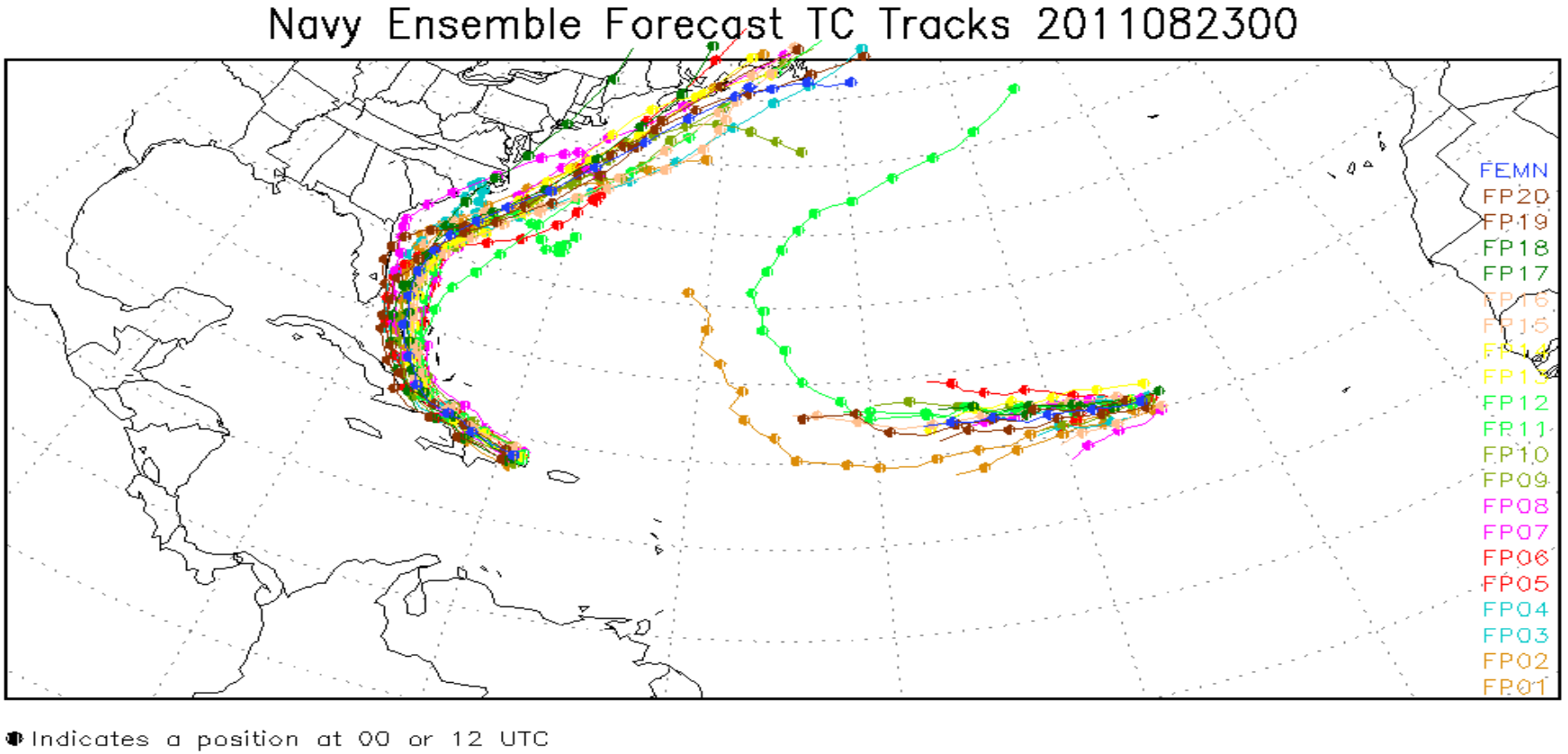
Various Ensembles for Irene: 000Z August 23, 2011



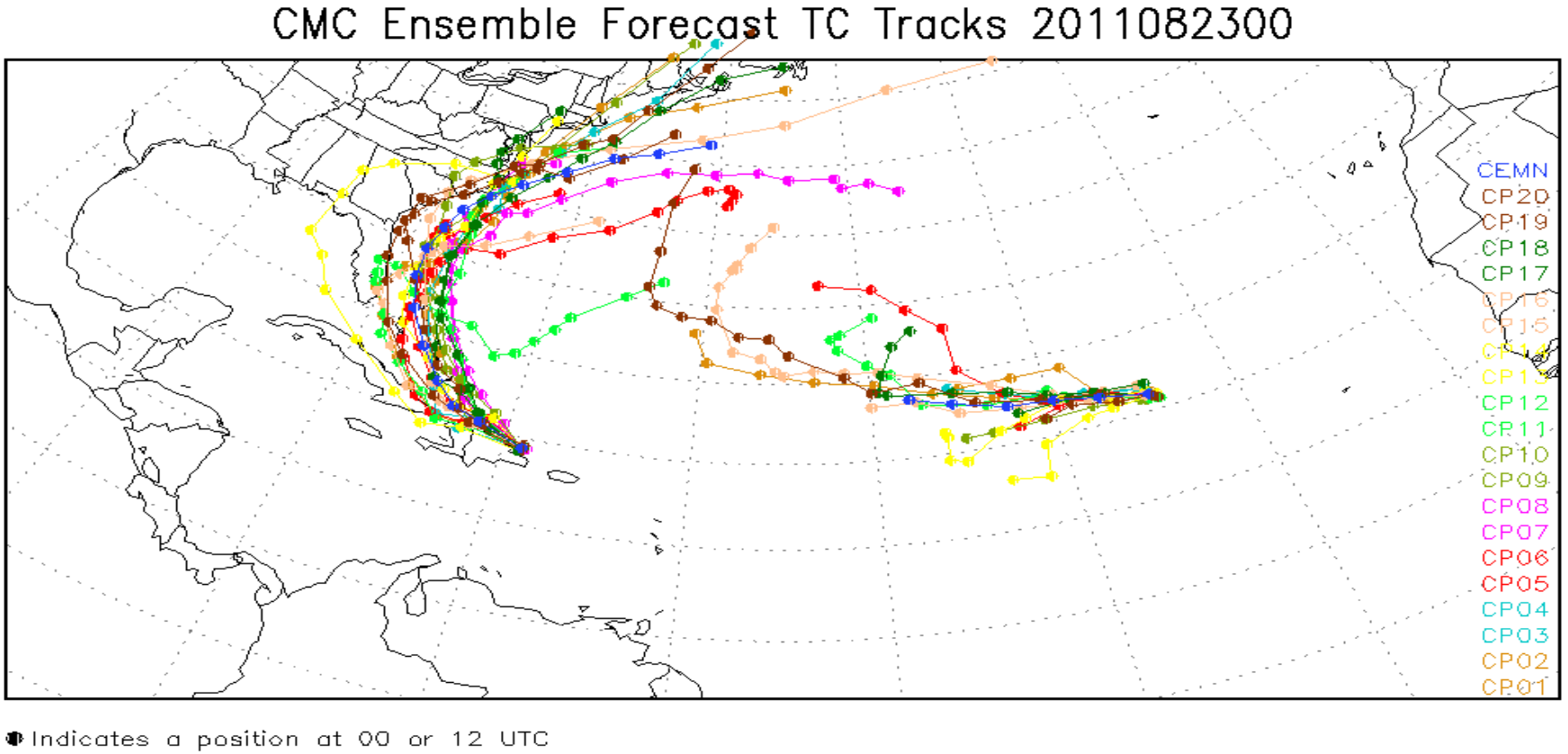
Various Ensembles for Irene: 000Z August 23, 2011



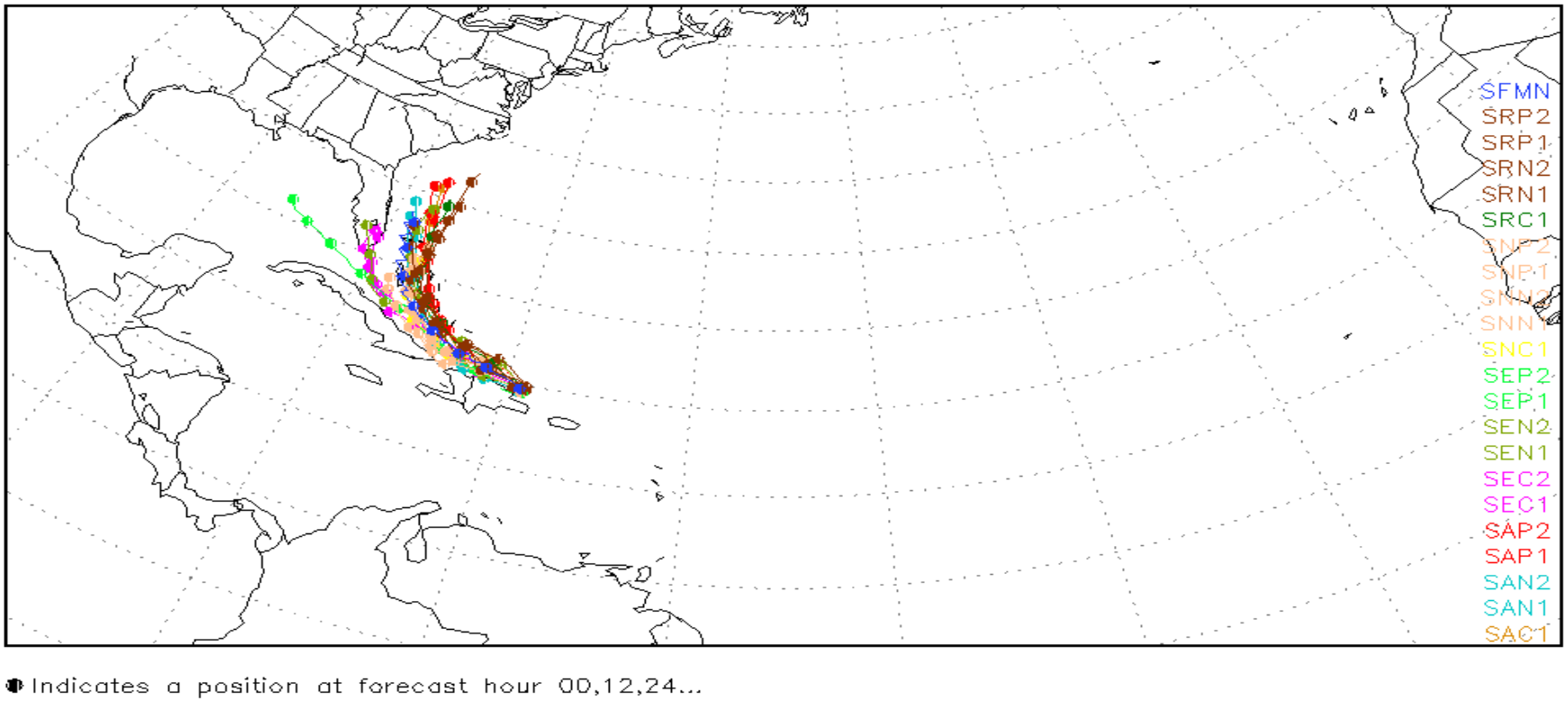
Various Ensembles for Irene: 000Z August 23, 2011



Various Ensembles for Irene: 000Z August 23, 2011



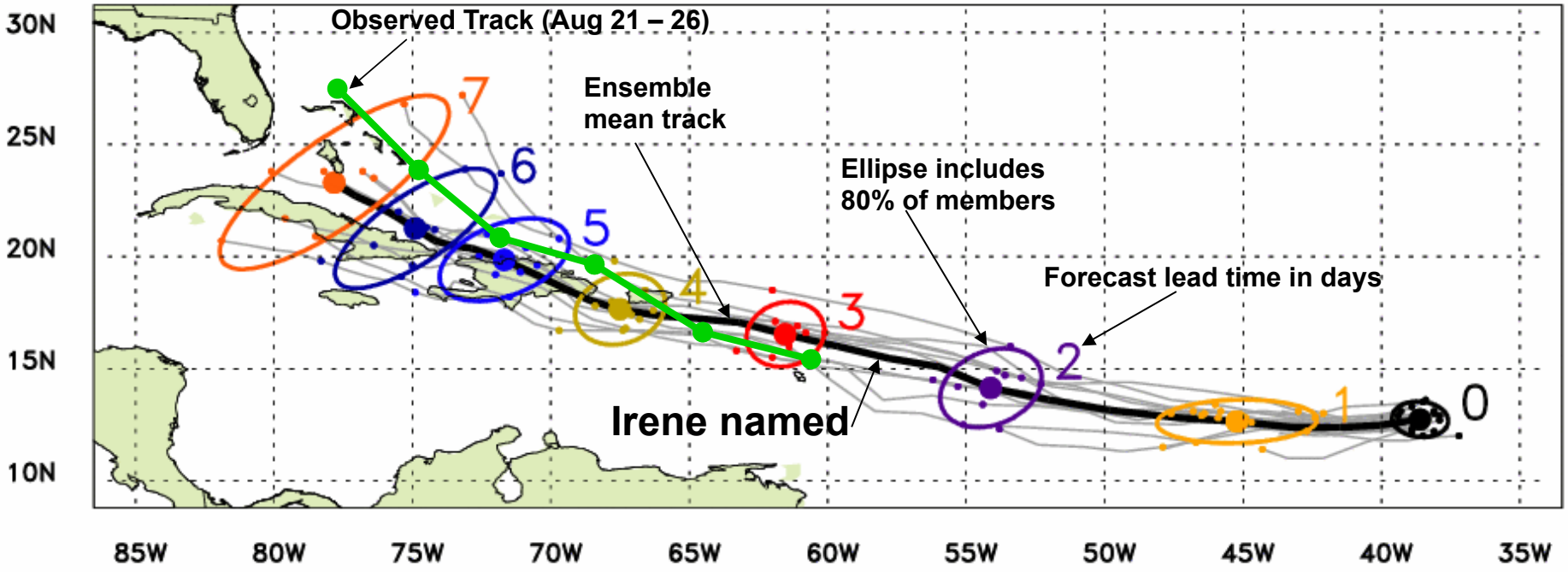
Various Ensembles for Irene: 000Z August 23. 2011



Genesis

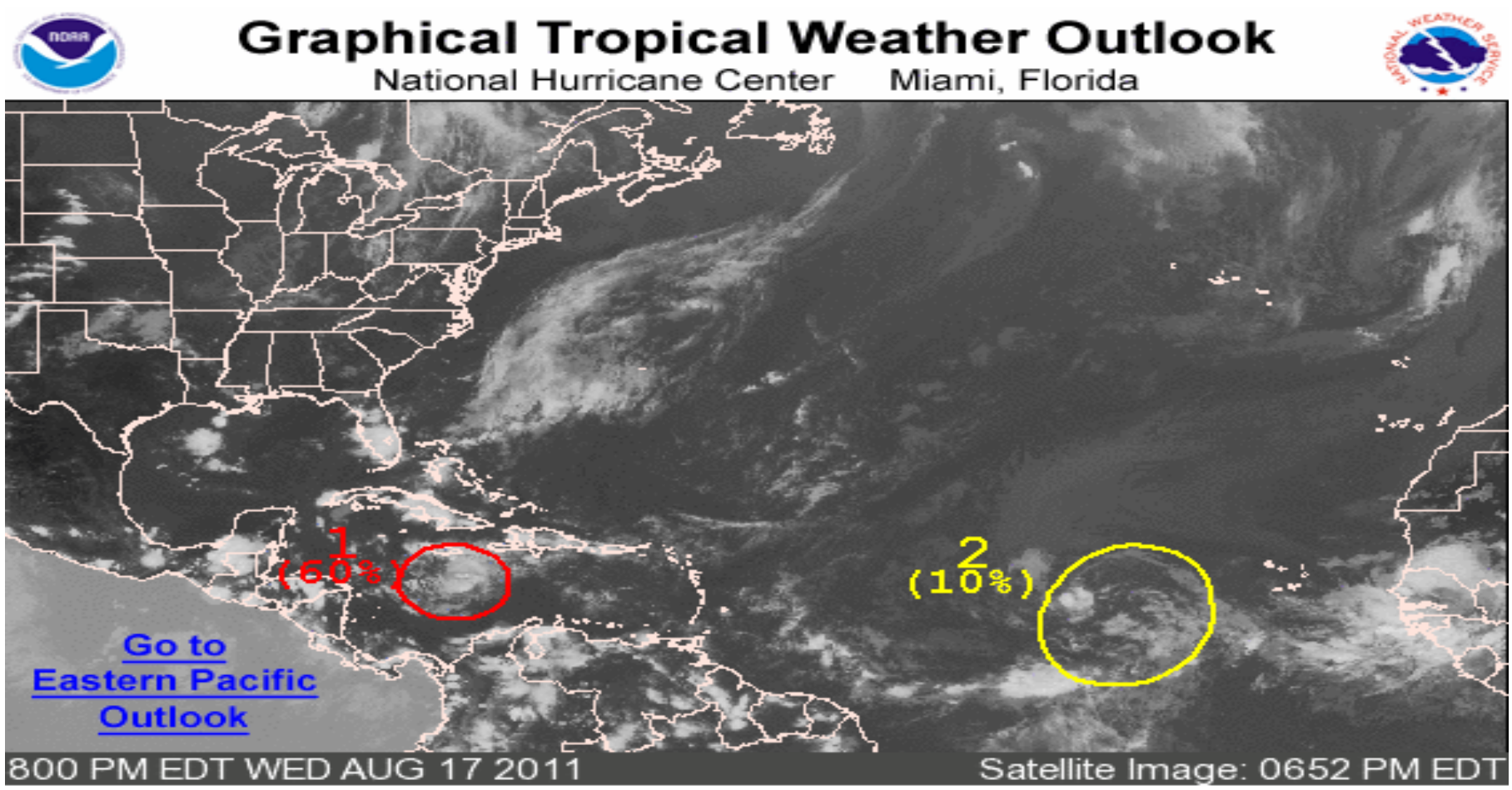
- This is really something that can only be done with an ensemble:
  - Can be constructed in many ways
    - Multi model
    - From individual models
  - Probably best done with a global ensemble since genesis can often be may days after model initialization
- Global models this summer—even the relatively low resolution GFS/EnKF system (T256)—appeared to do very well
  - This needs to be verified

HFIP Global Ensemble Forecast for Irene Starting at 1200Z **August 18, 2011**



- Irene declared an investigation area at 1200Z on August 18, 2011
- Irene named at 0000Z August 21, 2011
- Initial indication of the formation of Irene from ensemble at **00Z August 16, 2011**
  - 2 days before it was declared an investigation area
  - 5 days before it was named

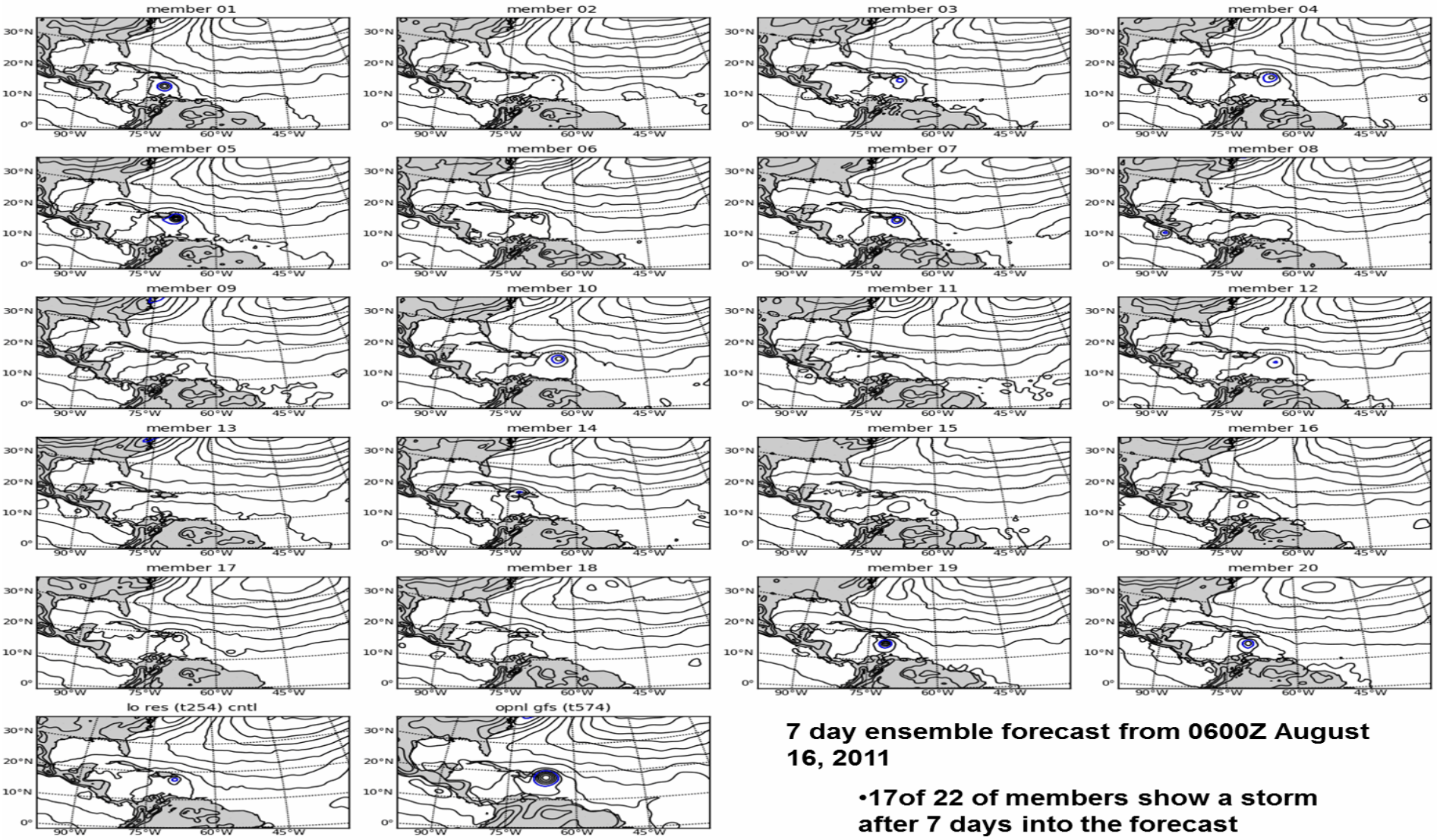
Graphical Tropical Weather Outlook



Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

Low <30% Medium 30-50% High >50%

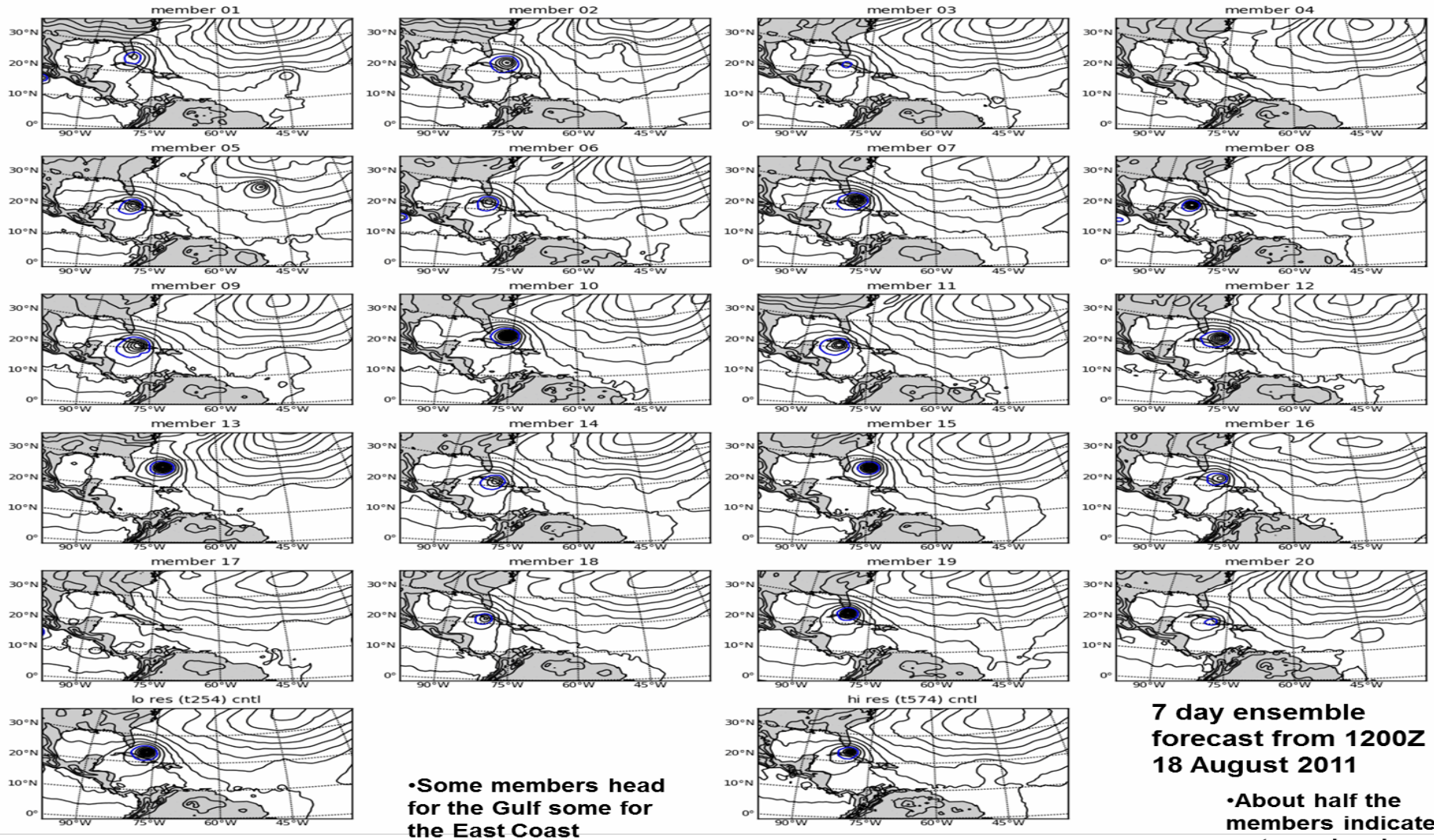
7 Day Ensemble Forecast
(Irene August 16, 2011)



7 day ensemble forecast from 0600Z August 16, 2011

- **17of 22 of members show a storm after 7 days into the forecast**

7 Day Ensemble Forecast
(Irene August 18, 2011)



- Some members head for the Gulf some for the East Coast

7 day ensemble forecast from 1200Z 18 August 2011

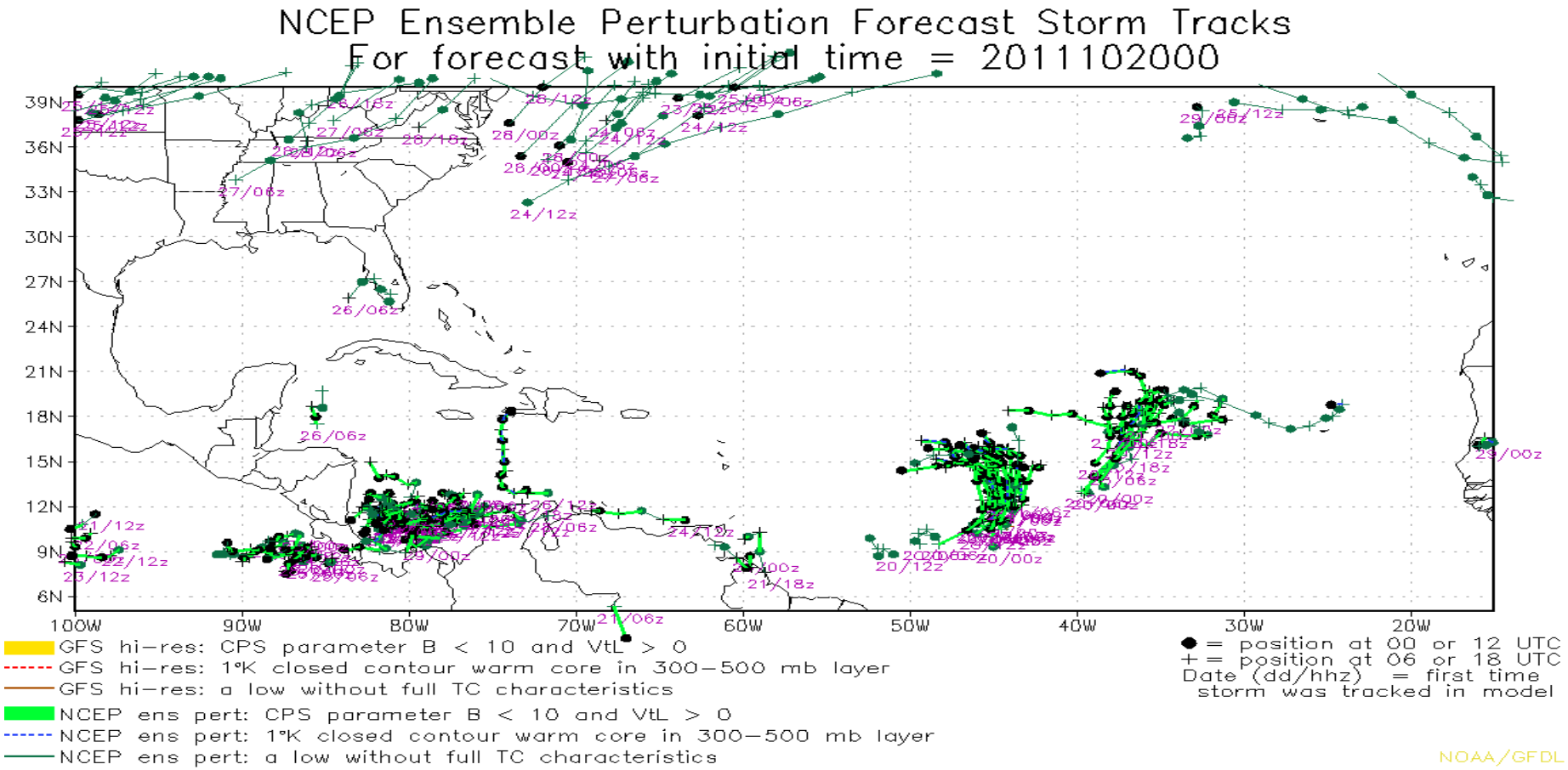
- About half the members indicate a strong hurricane

Genesis Products

We now have trackers that can detect genesis of a low system

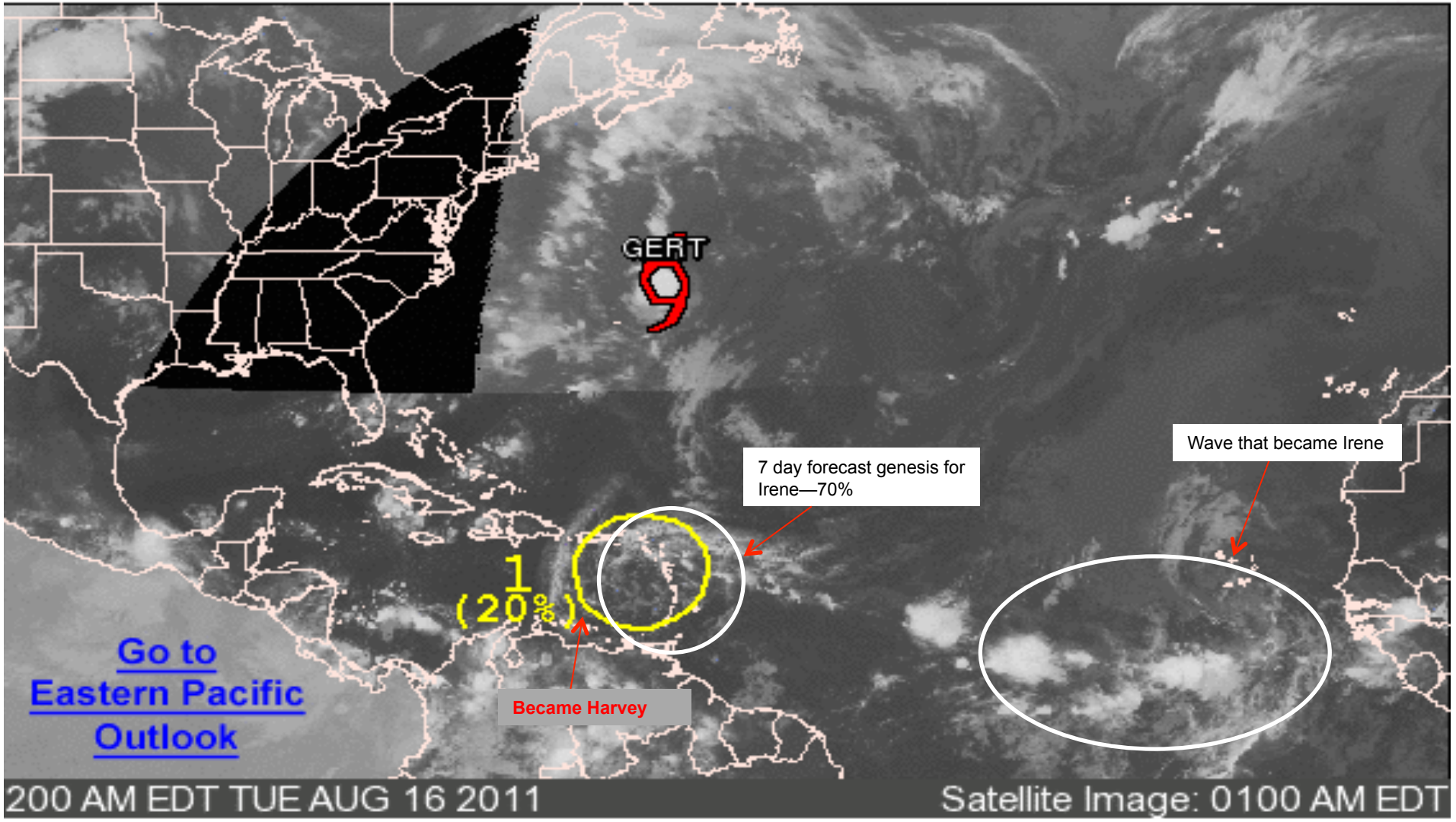
- Show example from Marchok's analysis
- It should be possible to isolate a particular forecast lead time (3, 5, 7 days for example) and construct ellipses similar to those shown earlier for tracked storms
- These ellipses could also indicate the percentage of ensemble members with a system in ellipse to give a probability of genesis.
  - Note that such a chart should be fairly easy to verify

NCEP Ensemble Perturbation



Graphical Tropical Weather Outlook

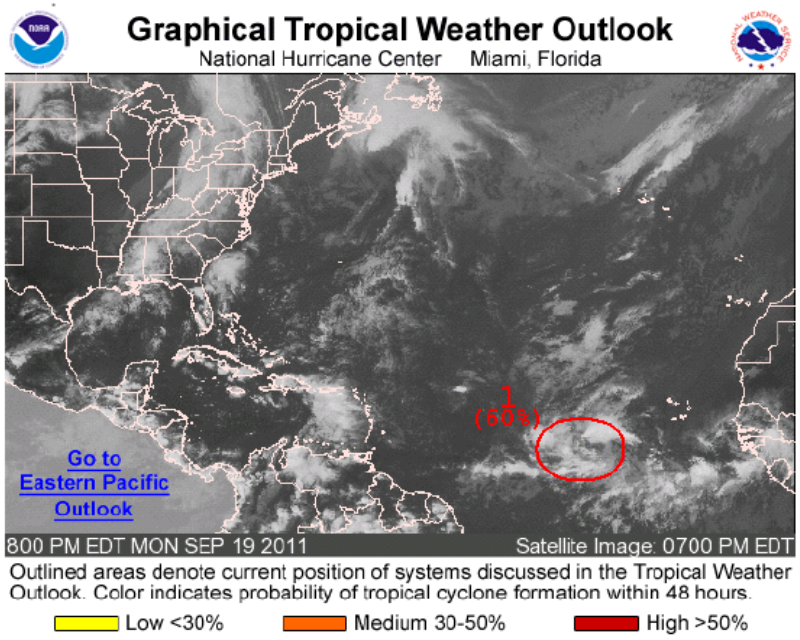
National Hurricane Center Miami, Florida

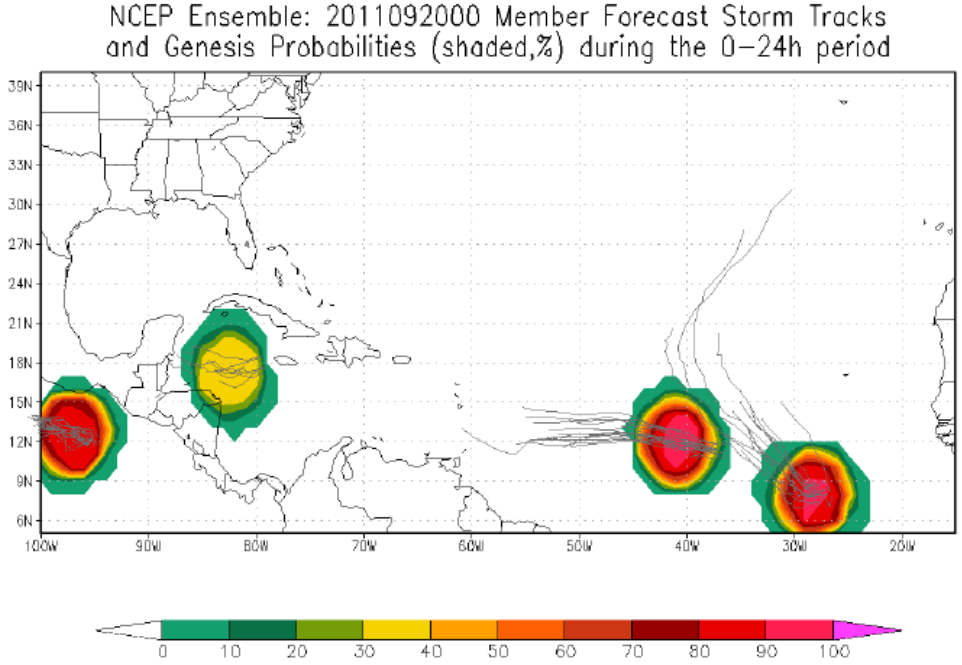


Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

Low <30% Medium 30-50% High >50%

Ensemble track-based probabilistic genesis guidance





Probability is simply the percentage of members indicating genesis in a given lead time window (here, 0-24h).

Summary

- Note that if the deterministic model has three times the resolution of the ensemble then the cost would be about the same the ensemble with 30 members.
  - The accuracy of the deterministic may then be better than the ensemble average
  - ***But the earlier discussion with respect to track still applies!***
- For HFIP I believe the best strategy is:
  - Run the global models as ensembles (multi-model, 10-20 members each) at as high a resolution as computational resources allow
    - Provides the track forecasts out to 7 days.
  - Run the regional models as ensembles (multi-model, 10-20 members each) at as high a resolution as possible
  - Fix the initialization problem for the regional models and switch the global initialization to the GSI Hybrid system
    - For regional models add radar and satellite data