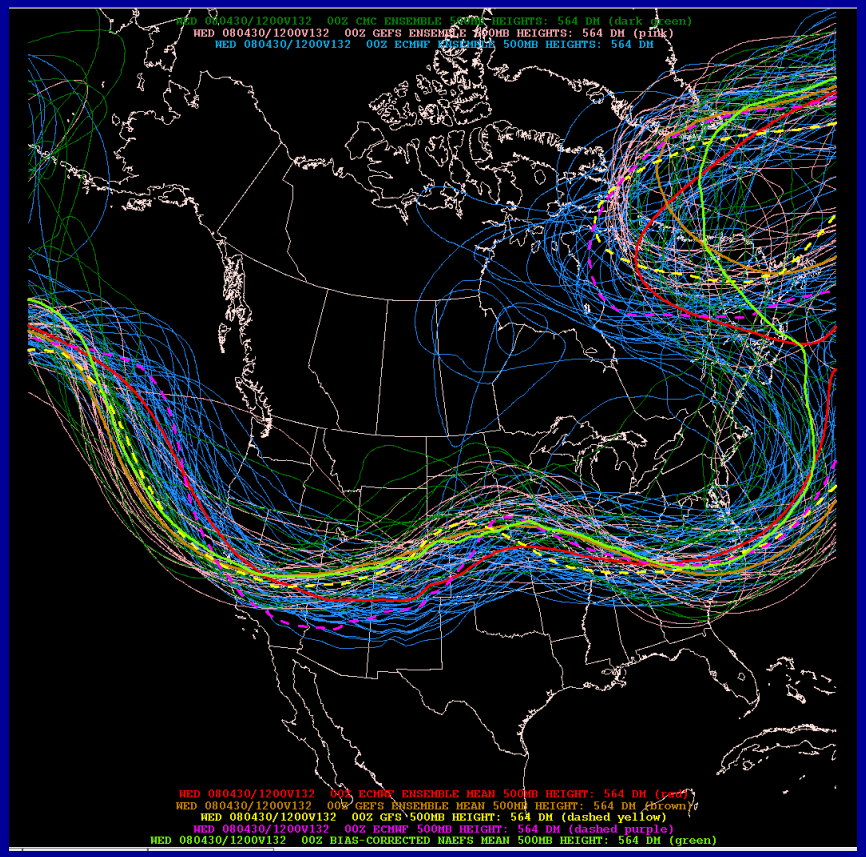
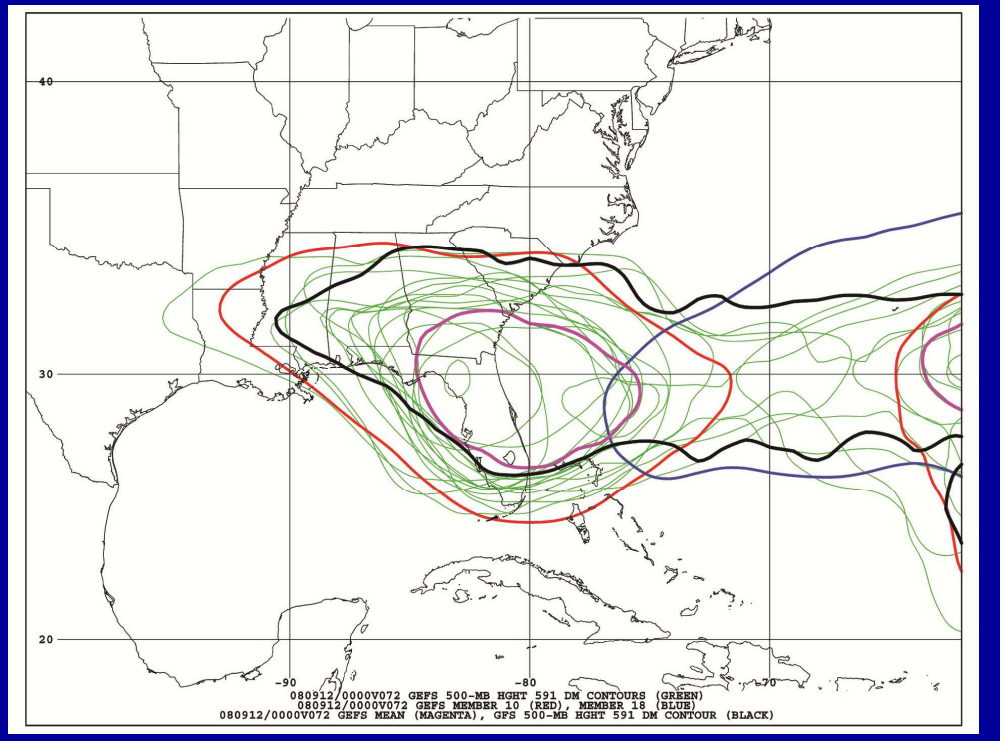
Ensemble Prediction Systems





Eric Blake
National Hurricane Center
March 6, 2018

Acknowledgements to Michael Brennan

Question 1

What are some current advantages of using single-model ensembles?

A. Estimates of uncertainty

B. TC intensity model spread

C. Alternative TC-track solutions

D. All of the above

E. A & C

Why Aren't Models Perfect?

- Atmospheric variables cannot be measured to an infinite degree of accuracy or precision (measurement error)
- Models' initial state never matches the real atmosphere (analysis error)
- Initial condition errors grow with model integration time, most rapidly at smaller scales (error growth)
- Model equations do not fully represent all of the processes in the atmosphere (model error)
- Model grid cannot explicitly resolve all features and processes in the atmosphere (model error)

Options?

- Increase our understanding of physical processes and how models represent them (research)
- More accurate and numerous observations with greater coverage (expensive)
- Improved data assimilation methods (4-D Variational Data Assimilation, Ensemble Kalman Filter)
- Faster computers and more complex models (many programs competing for resources)
- *Probabilistic forecasting with ensembles*

Definitions

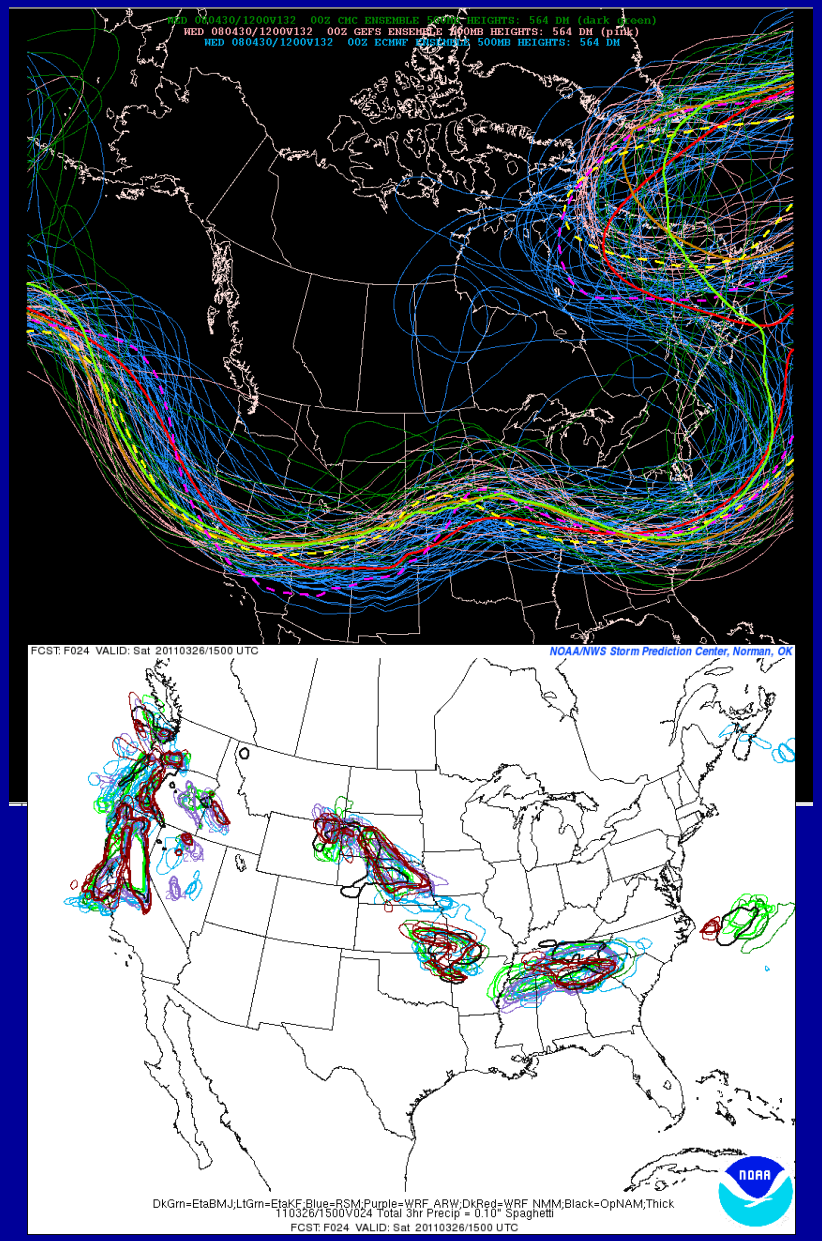
- **Deterministic Model** - single forecast from one forecast model or method using a single set of initial conditions
  - Examples: GFS, ECMWF, UKMET, GFDL, HWRF, BAMS
- **Ensemble** - collection of "member" forecasts verifying at the same time created from:
  - Different but equally viable initial conditions
  - Different forecasting methods and/or models that (ideally) statistically represent nearly all forecast possibilities

Definitions

- **Dynamical Model Ensemble** –based on perturbation of initial conditions of a single model or different models to create "member" forecasts
  - Examples: NCEP Global Ensemble Forecast System (GEFS), ECMWF Ensemble Prediction System
- **Control Run** – for dynamical model ensembles, the member of the ensemble run with the "best" initial analysis
  - The analysis used by the control run is usually perturbed to produce initial conditions for the remaining ensemble members
- **Spread** – measure of the degree of disagreement (i.e., standard deviation) between ensemble members

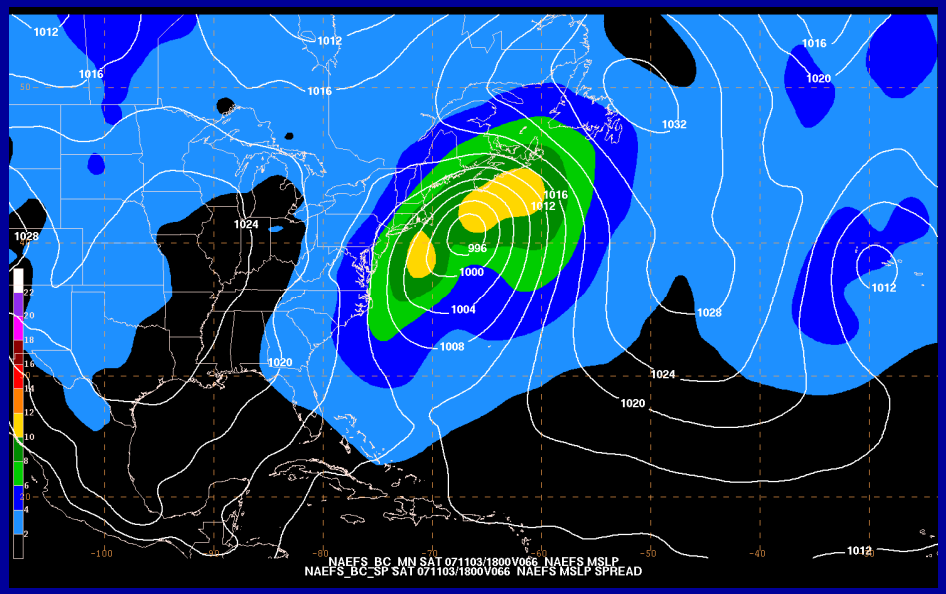
Ensemble Use

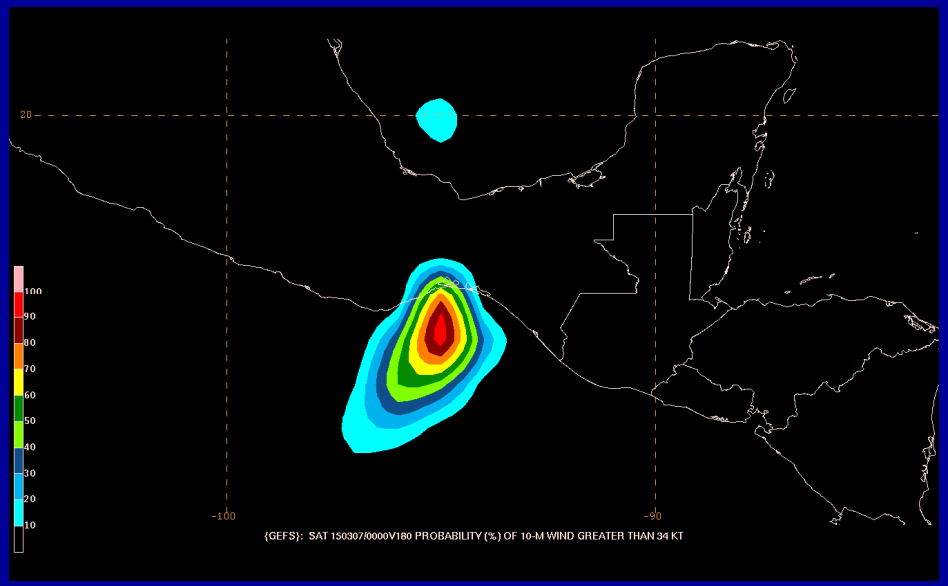
- Originally used for medium- to long-range forecasting of the large-scale pattern
- Uses have grown to encompass all temporal and spatial scales down to convective storm scale
- Address uncertainty, particularly those leading to rapidly diverging solutions
  - Initial conditions, model physics, resolution, model numerics



Ensemble Use

- Estimate rate of skill loss with time
  - Spread of solutions generally increases with time
- Compute probabilities of occurrence of a particular event or condition
  - 25 mm of precipitation, winds > 34 kt
- Identify regions where the analysis and forecast are sensitive to additional data in the analysis
  - Ensemble Kalman Filter, targeted observations



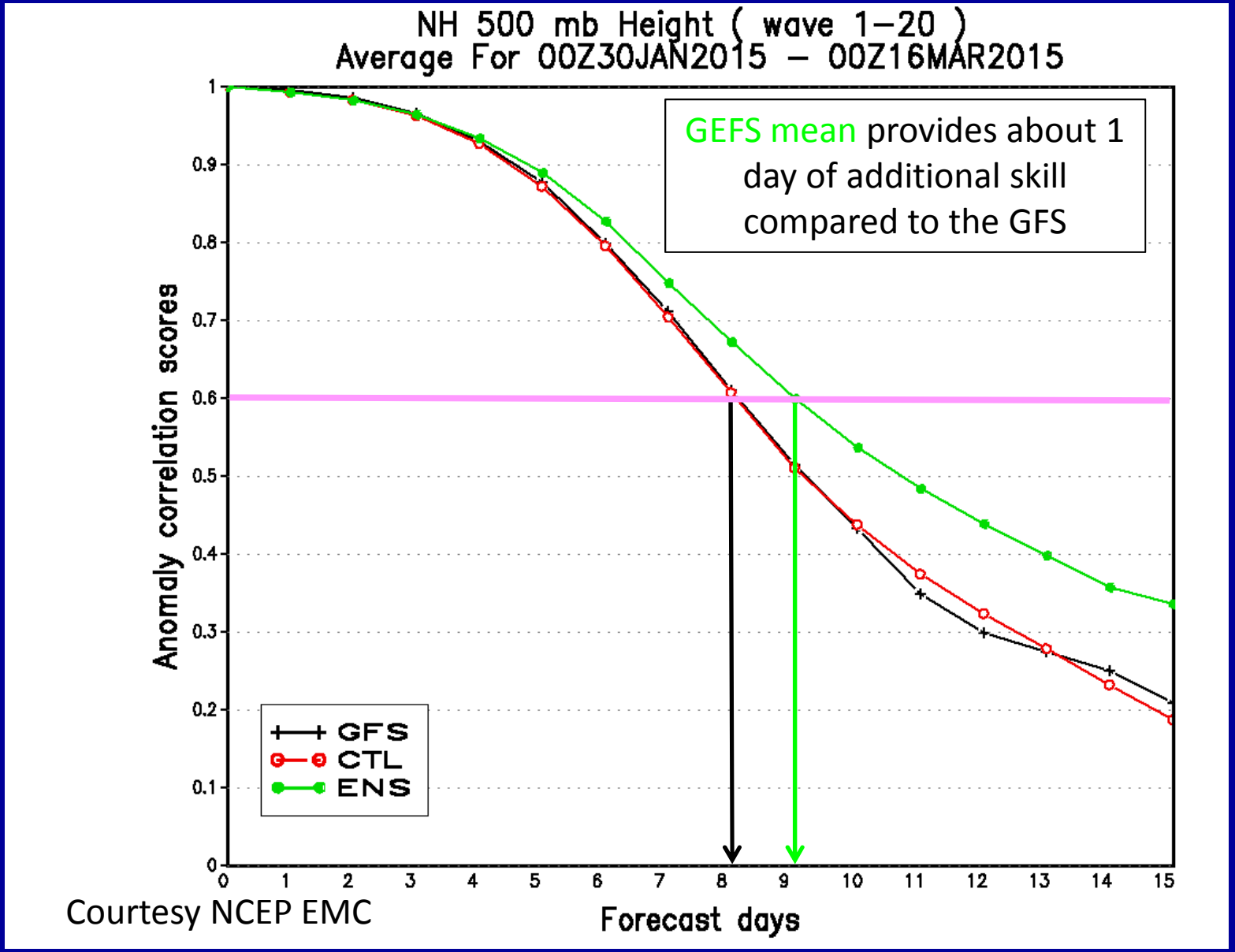


Ensemble Mean vs. Deterministic

- Deterministic runs (e.g., GFS) usually have more skill than any *individual ensemble member* due to superior resolution
- Ensemble mean usually has at least as much skill as an *equal-resolution control run*
- Ensemble mean can be more skillful than a *higher-resolution deterministic run*, especially beyond ~3 days

Ensemble Mean vs. Deterministic

500-mb height anomaly correlation die-off chart – 30 Jan-16 Mar 2015

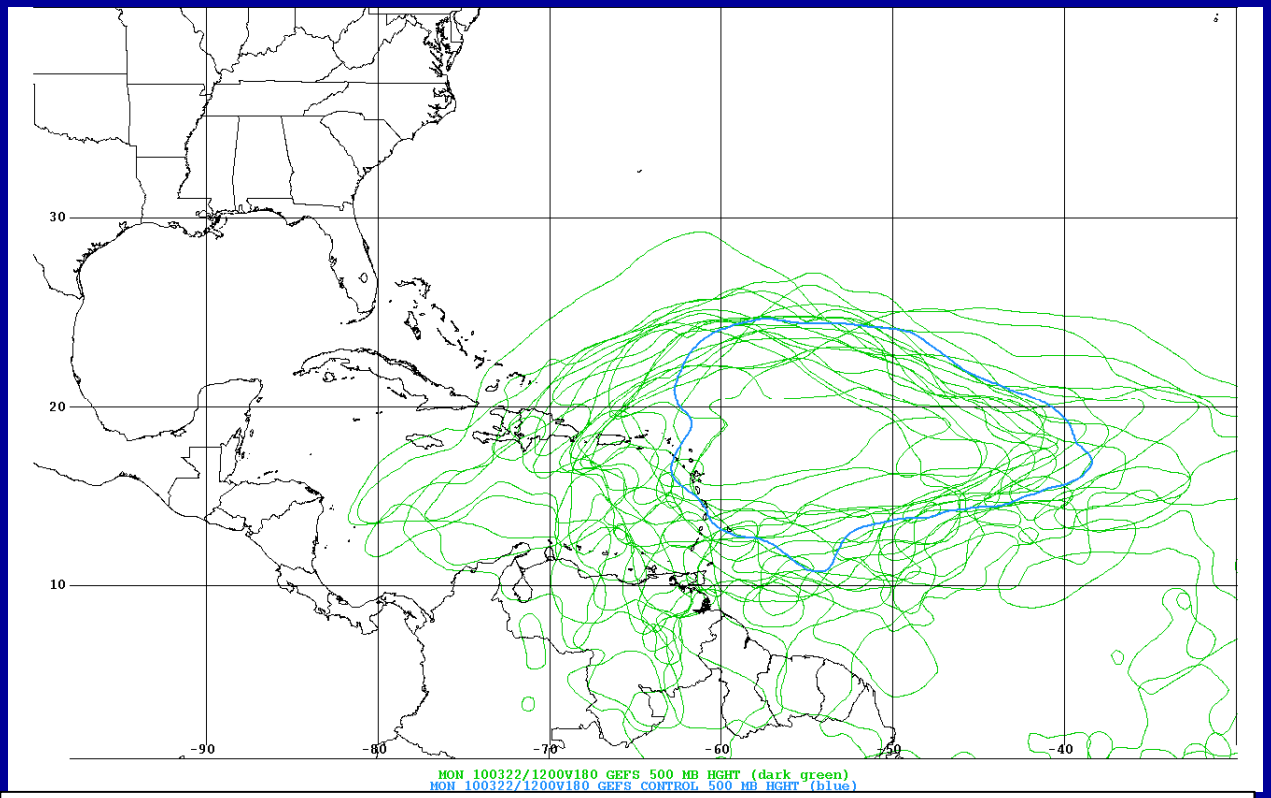


Courtesy NCEP EMC

Current Global Ensemble Systems that NHC uses most frequently

NCEP Global Ensemble Forecast System (GEFS)

- 4 cycles per day (00, 06, 12, 18 UTC)
- 21 members (1 control + 20 perturbed)
- Forecast extends out to 384 hours (16 days)



180-h forecast of 588 dm 500-mb height contour valid at 1200 UTC 22 March 2010

NCEP GEFS

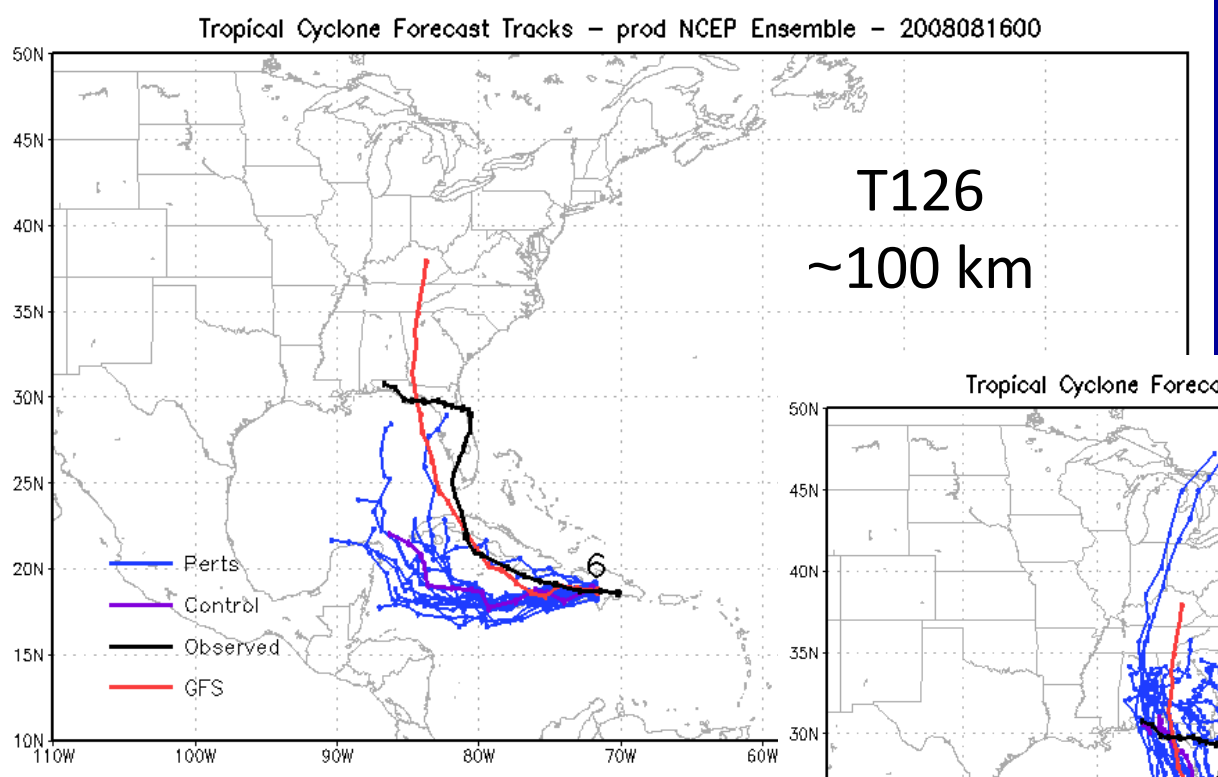
- Current Configuration (last upgrade 2015)
  - T574 (~ 34 km) through 8 days, T328 (~ 52 km) days 8-16
  - 64 vertical levels

- Ensemble members
  - 20 members generated using Bred Vector and Ensemble Transform methods to address uncertainties in the initial conditions
  - Stochastic (statistical) perturbations try to address model uncertainty
  - Includes vortex relocation to NHC/CPHC/JTWC analyzed position for tropical cyclones in each ensemble member
  - Model physics consistent with GFS

- Deterministic GFS
  - T1534 (~ 13 km) through 10 days, T574 (~ 35 km) days 10-16
  - 64 vertical levels

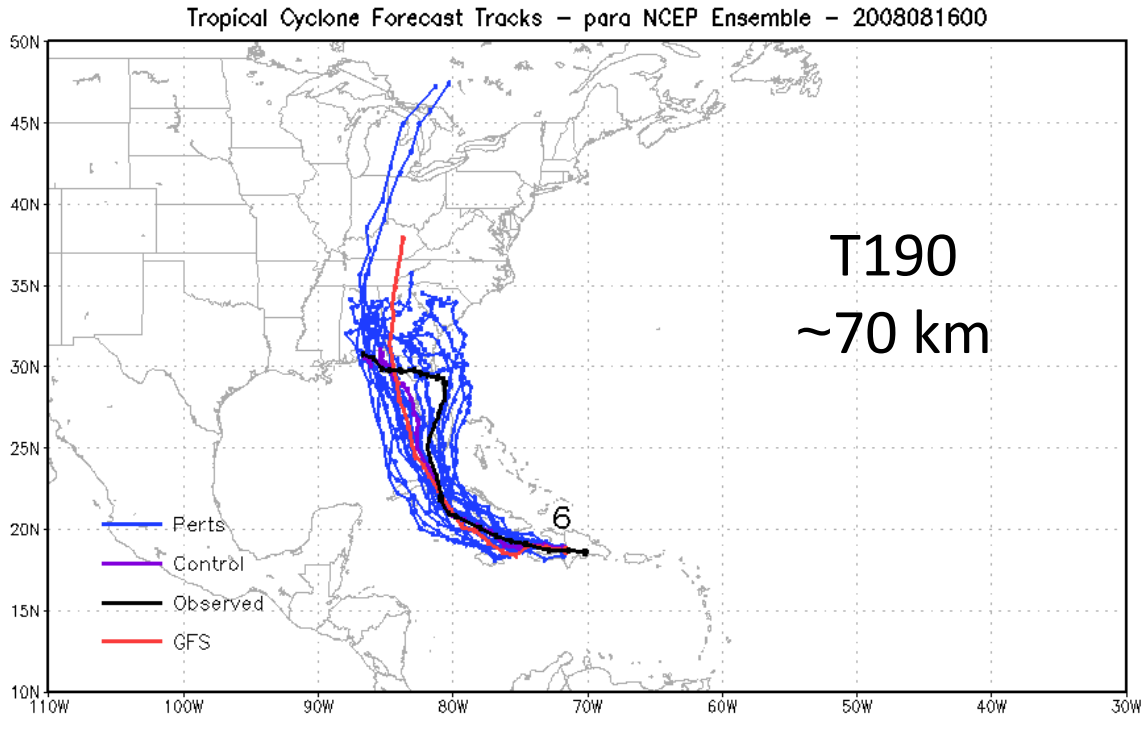
Improvements to Global Ensemble TC Track with Increasing Horizontal Resolution

Tropical Storm Fay 00Z – 16 Aug 2008



T126
~100 km

More members retain the TC and track forecasts are much improved



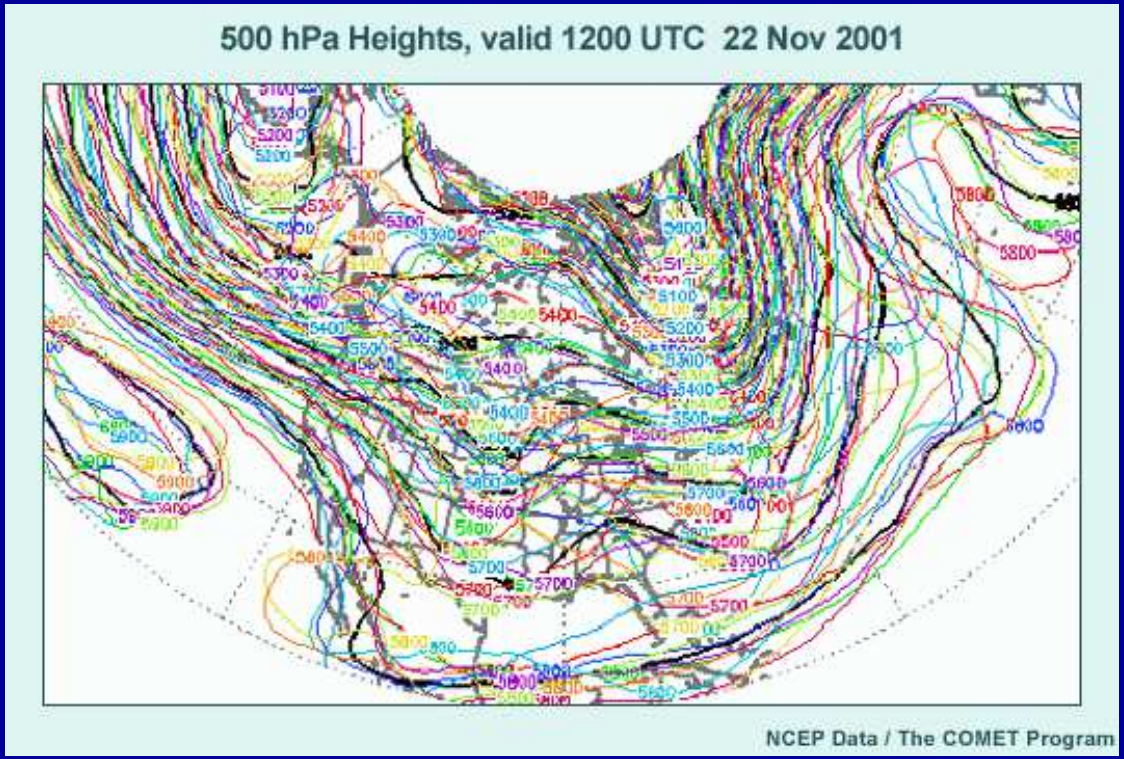
T190
~70 km

ECMWF Ensemble Prediction System

- 51 members (1 control+50 perturbed members)
- Run twice daily (00 and 12 UTC) out to 15 days
  - T639 (~ 18 km) to 15 days
  - 91 vertical levels
  - Perturbations:
  - Initial condition: generated using singular vectors and perturbations from an ensemble of data assimilations
  - Physics: generated by two stochastic parameterization schemes
- Deterministic ECMWF
  - Horizontal grid resolution T1279 (~9 km) out to 10 days with 137 vertical levels

Ensemble Display and Interpretation

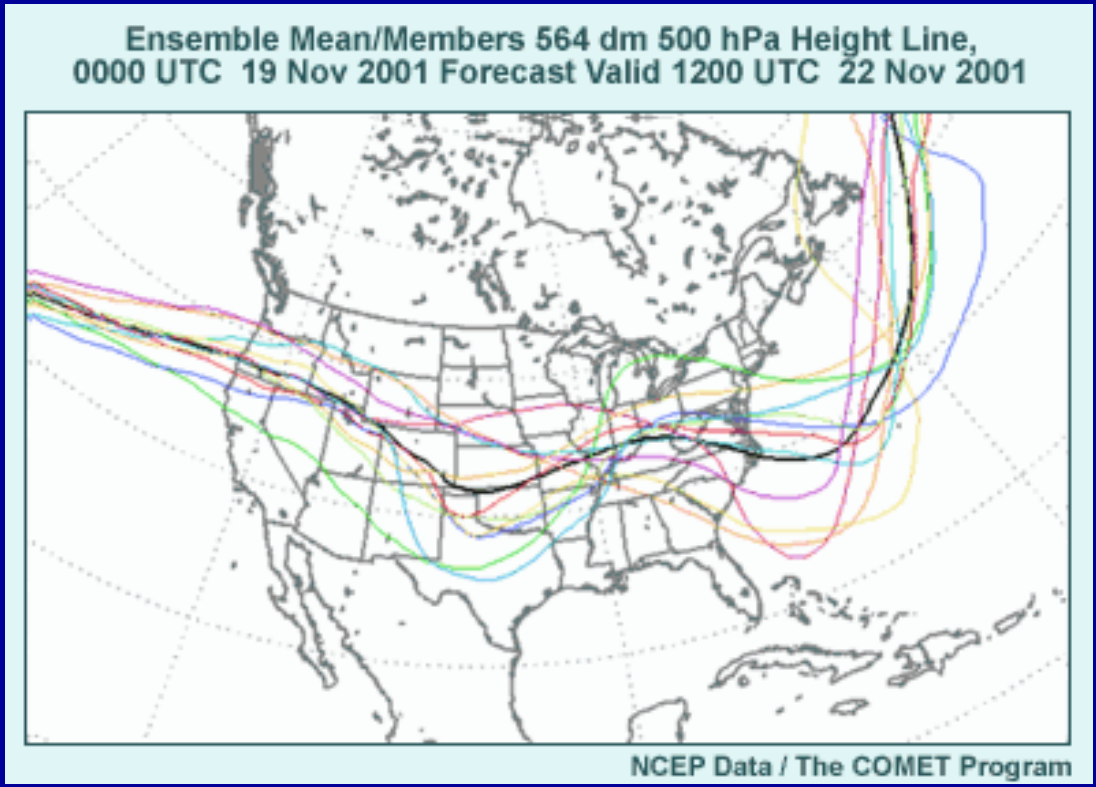
Displaying Ensembles



If we try to look at every ensemble member at once, it is messy and difficult to interpret

Credit: COMET

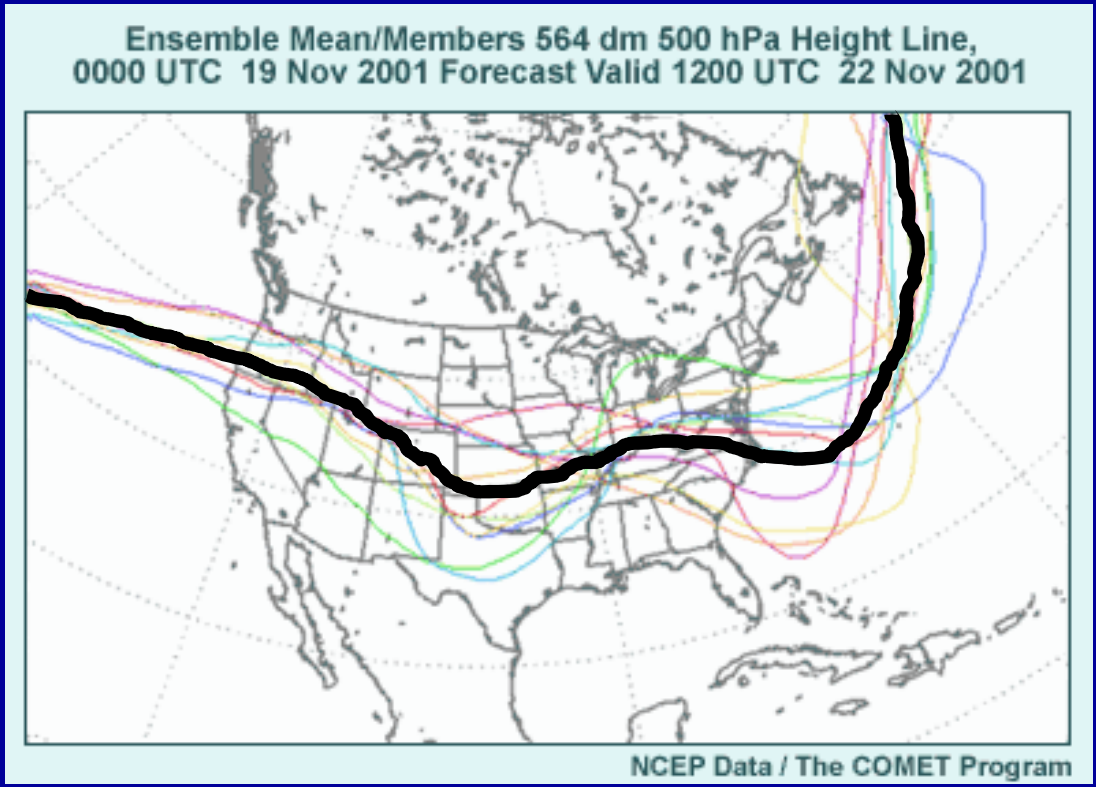
Displaying Ensembles



Spaghetti Diagram – displays one isopleth at a time from each ensemble member

Credit: COMET

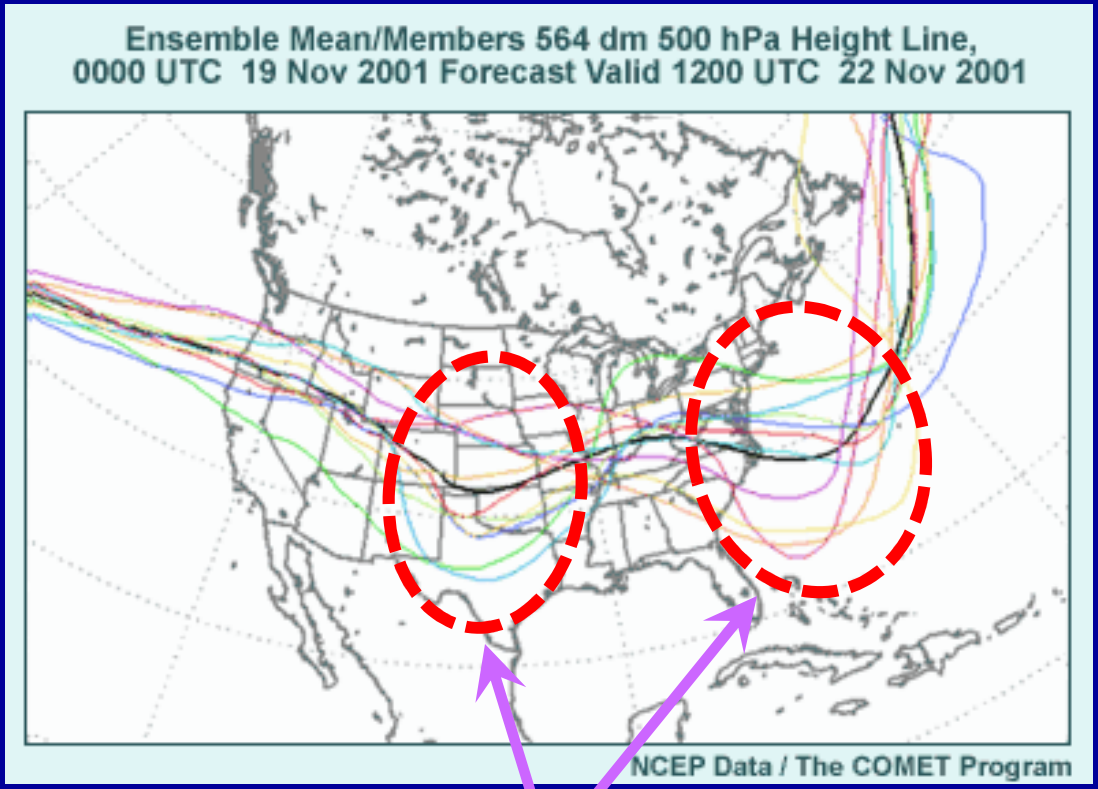
Displaying Ensembles



Ensemble Mean - average of multiple forecast members verifying at same time

Credit: COMET

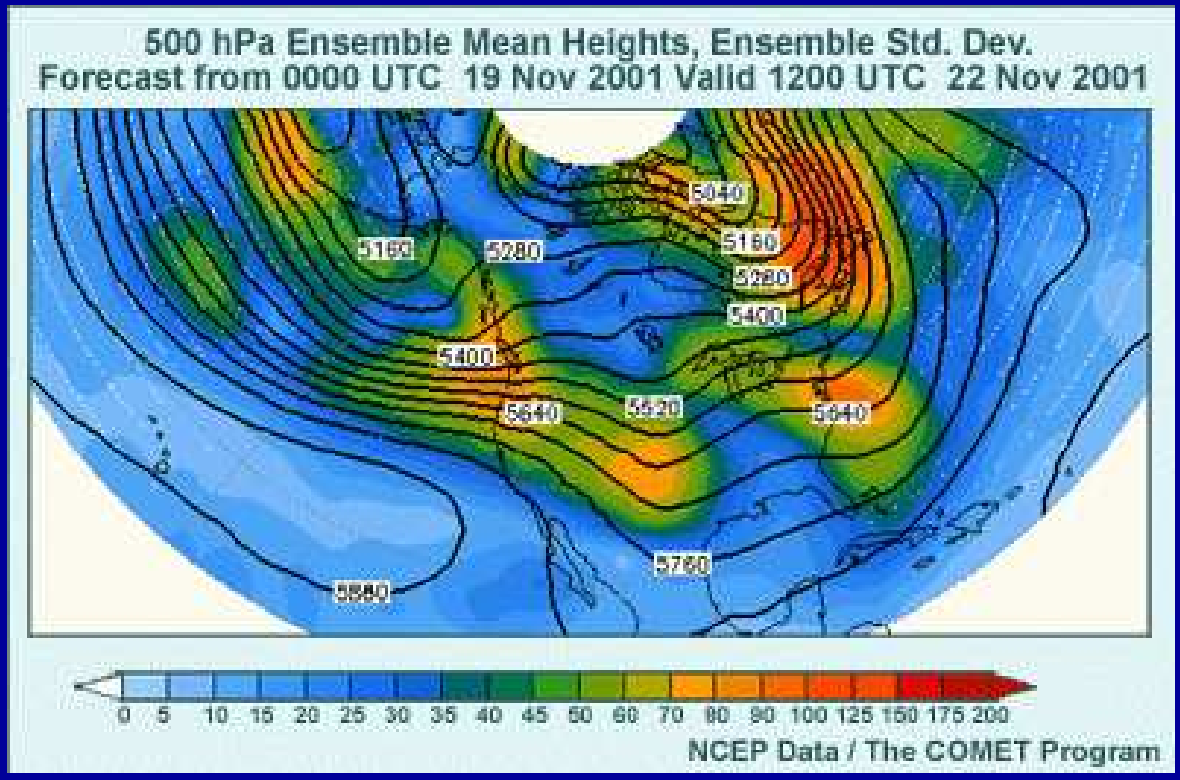
Displaying Ensembles



Disagreement, or **spread**, between ensemble members

Credit: COMET

Displaying Ensembles

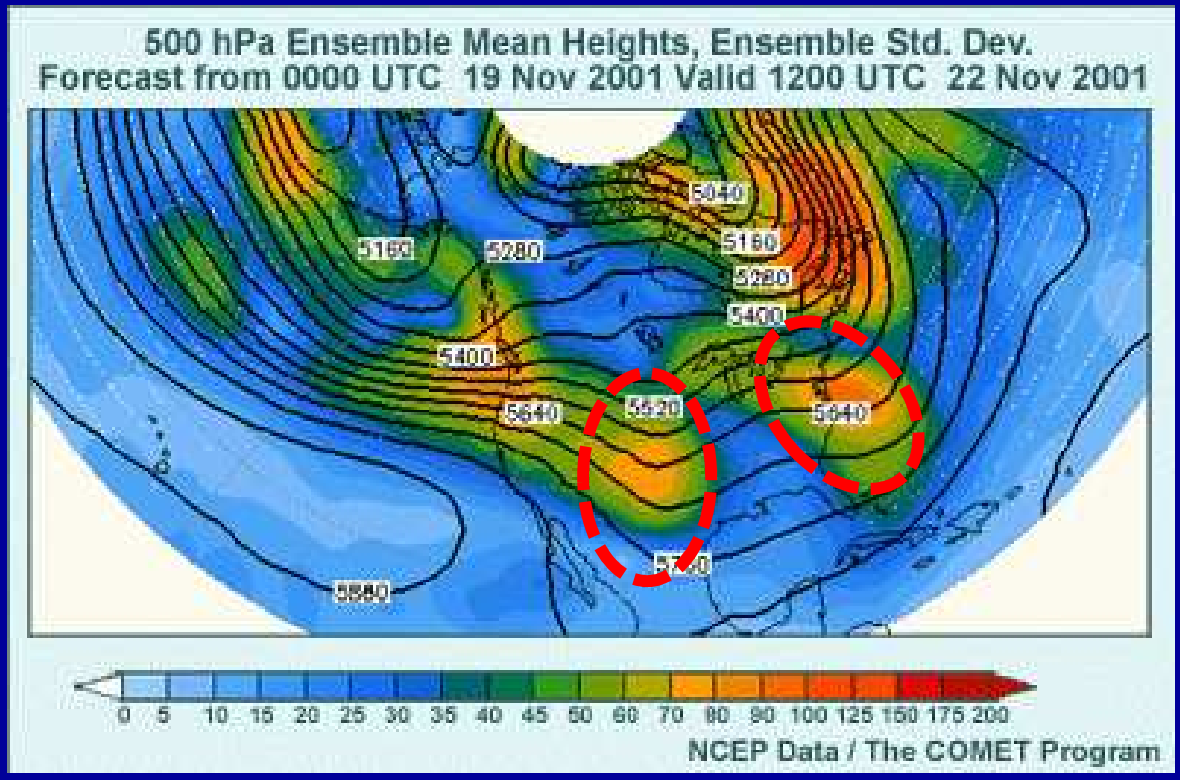


- **Black** lines = ensemble mean 500-mb height forecast
- Spread indicated by shading (meters)
  - Orange/Red – little agreement between members
  - Blue – good agreement between members

Credit: COMET

Displaying Ensembles

Ensemble Mean and Spread



- **Black** lines = ensemble mean 500-mb height forecast
- Spread indicated by shading (meters)
  - Orange/Red – little agreement between members
  - Blue – good agreement between members

Credit: COMET

Displaying Ensembles
Ensemble Mean and Spread

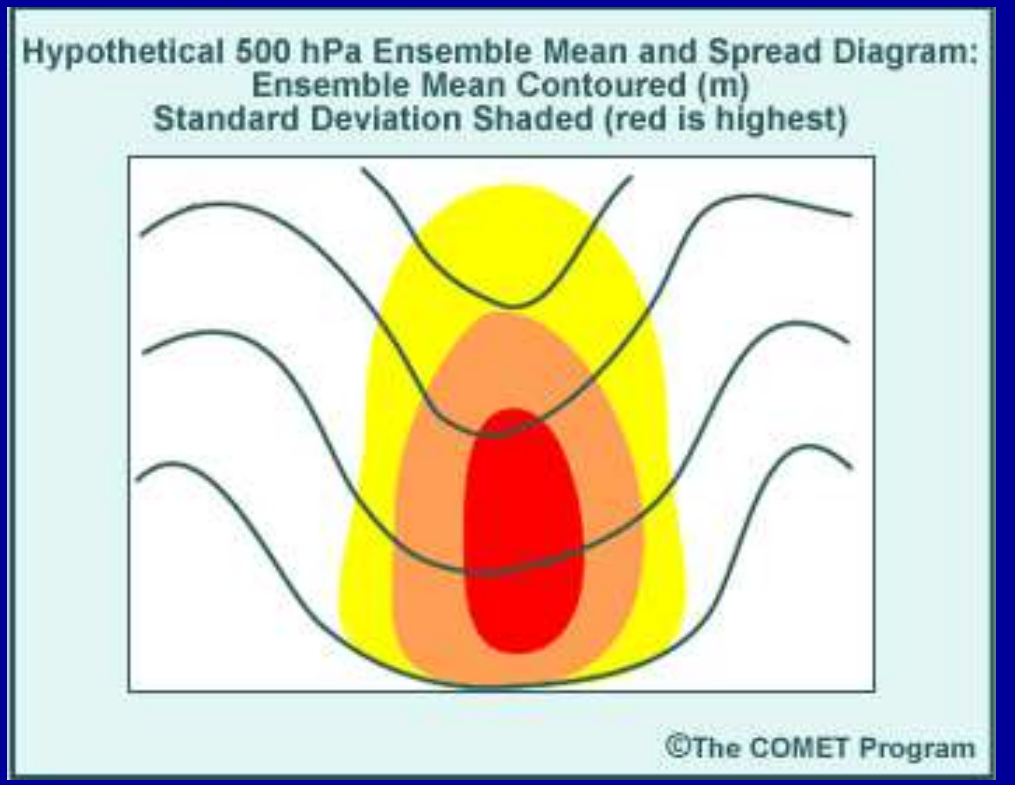
- Advantages
  - Summarizes data in easy to interpret form
  - Information provided for the entire domain
  - Low predictability features smoothed out by the ensemble mean and easily identifiable using spread
- Disadvantages
  - Ensemble mean can be misleading (and may not be the best forecast) if multiple clusters of nearly equal probability forecast outcomes exist (i.e., bi-modal distribution)
  - May not reveal extreme outlier solutions

Interpreting Mean and Spread

Large spread within the ensemble mean feature → Uncertainty in amplitude of the feature

•In this case, there is uncertainty in the **depth** (not the location) of this 500-mb trough

•If there were a tropical cyclone located southeast of this trough, would the trough be deep enough to recurve the tropical cyclone?



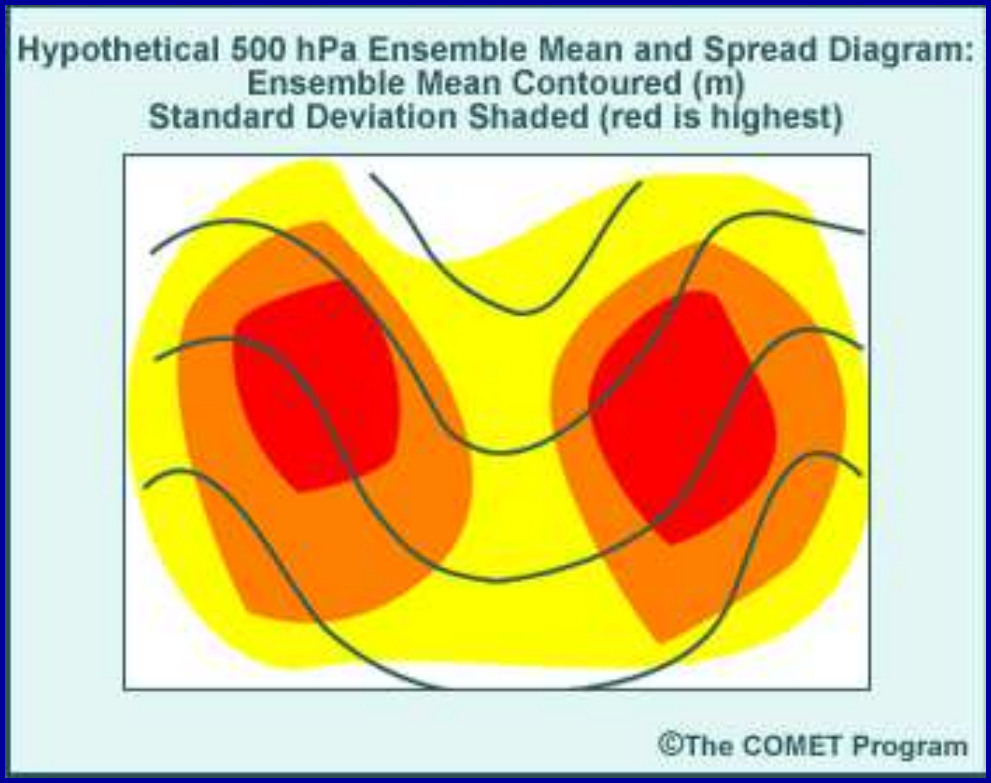
Credit: COMET

Interpreting Mean and Spread

Large spread upstream or downstream of an ensemble mean feature → Uncertainty in the location of the feature

•In this case, there are nearly equal chances that the 500-mb trough will be east or west of the position shown by the ensemble mean trough

•If a tropical cyclone was located southeast of this trough, at what time will the tropical cyclone begin to be influenced by this trough?



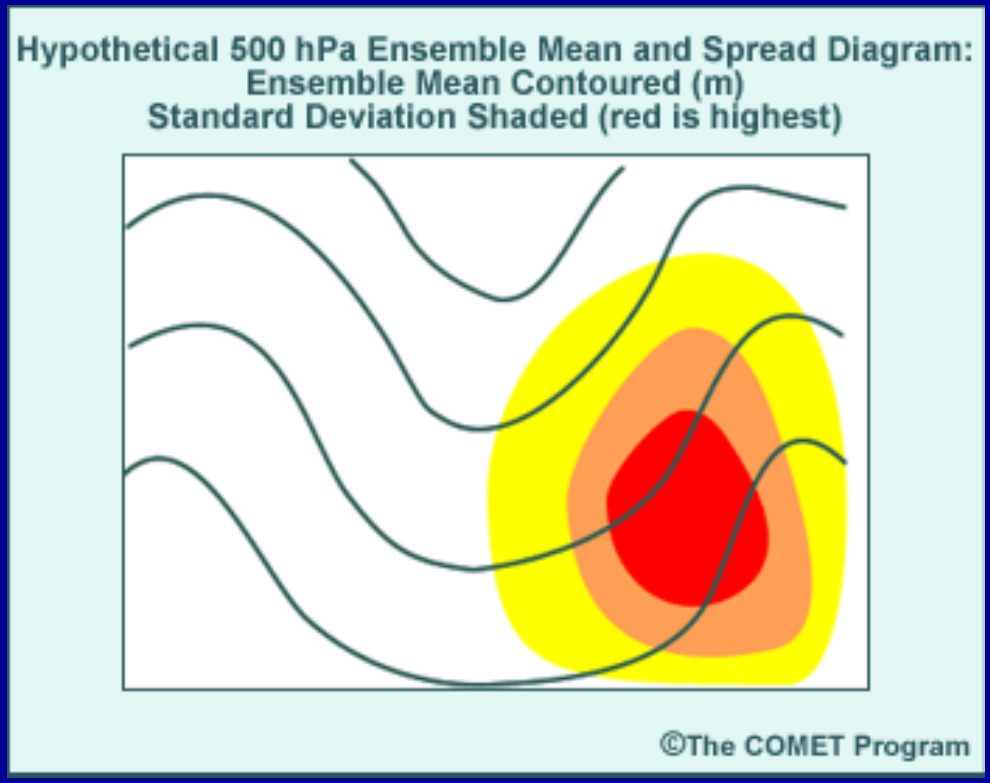
Credit: COMET

Interpreting Mean and Spread

Large spread on one side of an ensemble mean feature → A cluster of ensemble members different from the ensemble mean

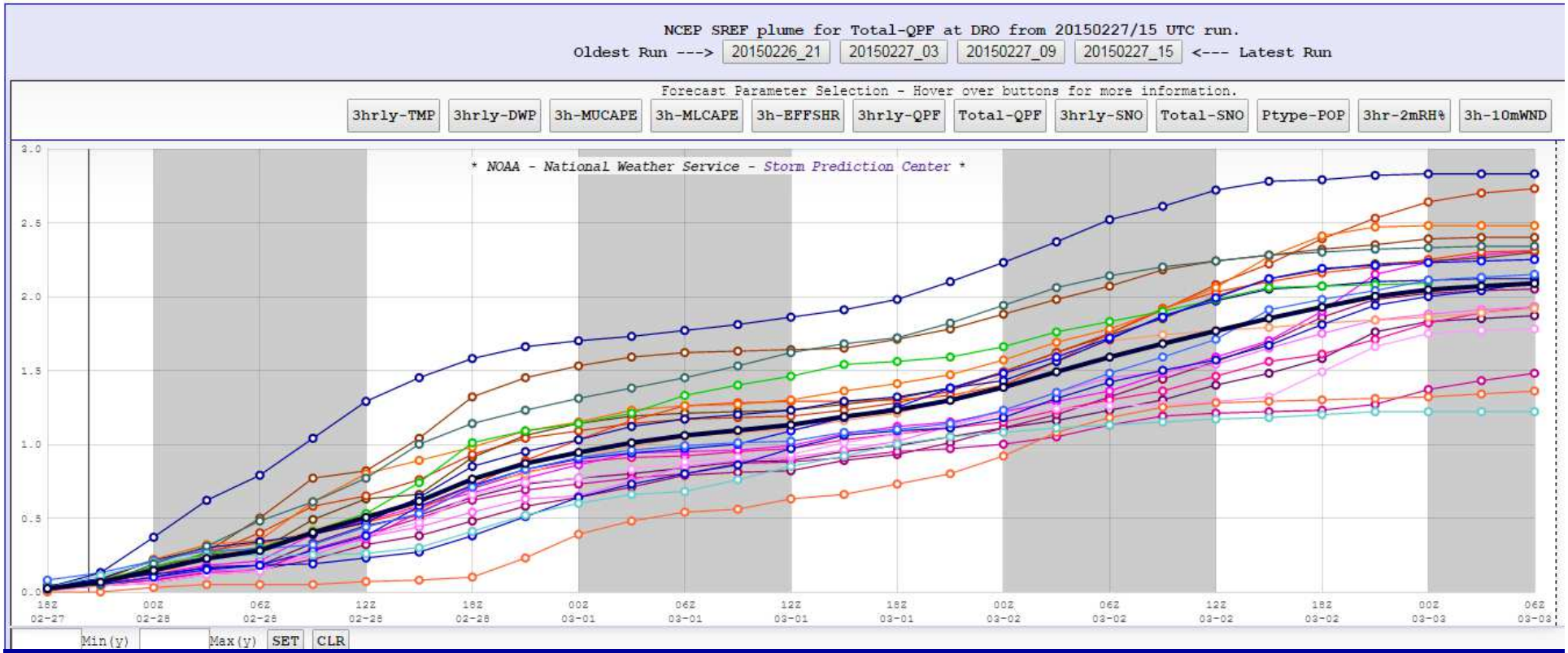
•In this case, the spread indicates greater potential for the trough axis to be east of the ensemble mean trough than to the west

•If there was a tropical cyclone located southeast of this trough, at what time will the tropical cyclone begin to be influenced by this trough?



Credit: COMET

Plume Diagrams



NCEP Short Range Ensemble Forecast System (SREF) plume diagram for total precipitation at Durango, Colorado, starting at 15Z 27 Feb 2015 (courtesy NWS SPC)

Genesis Guidance

Little objective guidance is seen with ensembles now, though they help subjectively.

In-house product→

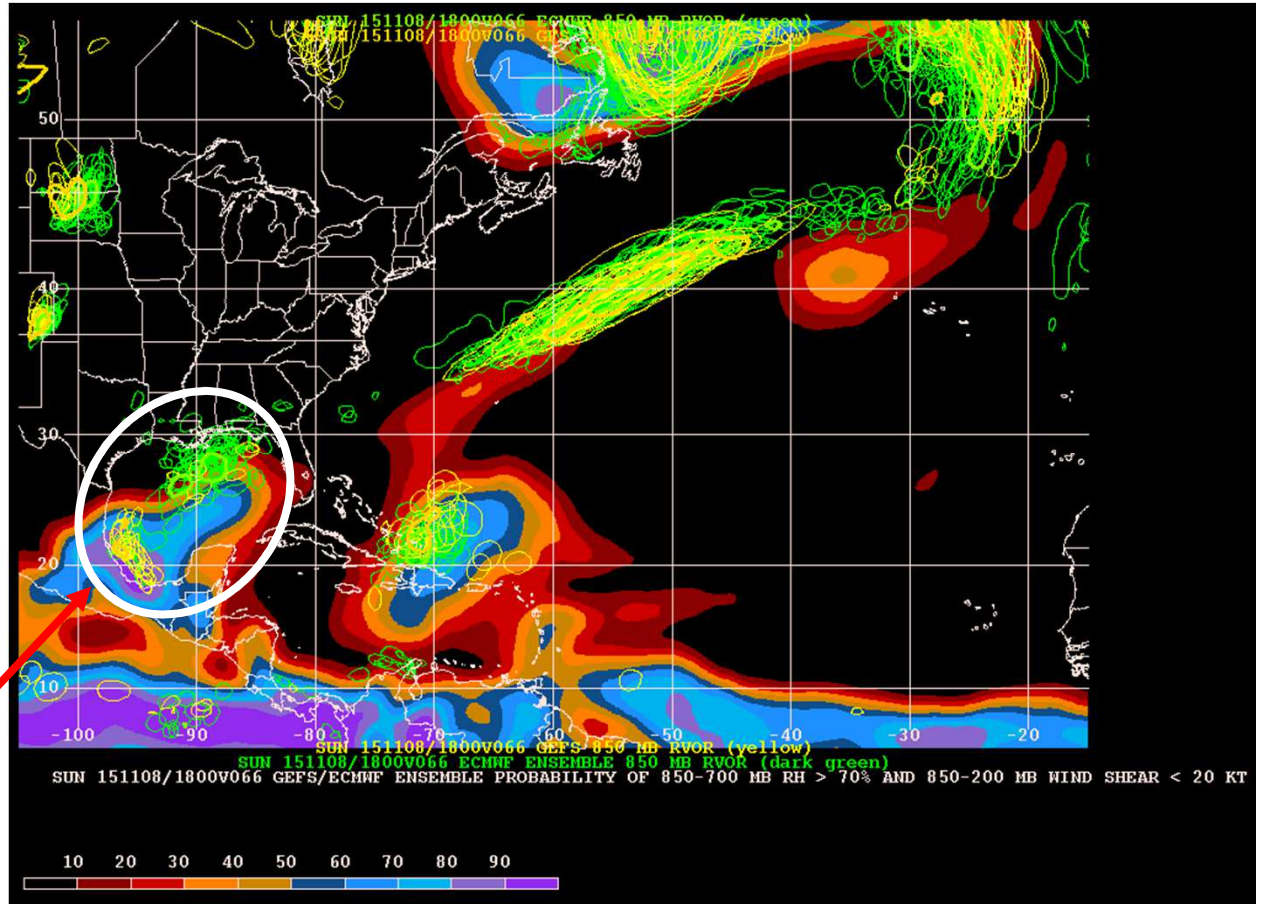
shading: combined probability of 70 ensemble members (GEFS + ECENS):

- 850 – 700 hPa RH > 70%
- 200 – 850 hPa vertical wind shear < 20 kt

contours: 850 hPa relative vorticity (8×10^{-5} s^{-1} intervals)
thin green: ECENS members
thick green: **ECMWF deterministic**
thin yellow: GEFS members
thick yellow: **GFS deterministic**

Invest AL93

0000 UTC November 6, 2015 **+ 66 h**



Ensemble Problems

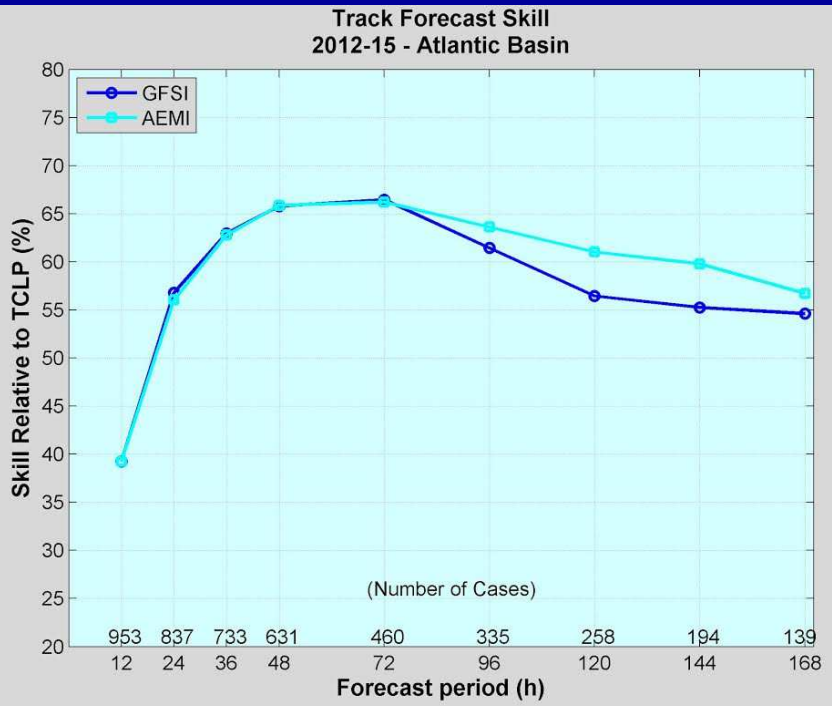
- Need a properly calibrated system
  - GEFS is currently underdispersive
  - This problem results in an overconfident forecast
  - Lower resolution can also hinder a more accurate track forecast (i.e. when track especially dependent on intensity)
- Other issues
  - Ensemble mean can be misleading (and may not be the best forecast) if multiple clusters of nearly equal probability forecast outcomes exist (i.e., bi-modal distribution)
  - May not reveal extreme outlier solutions

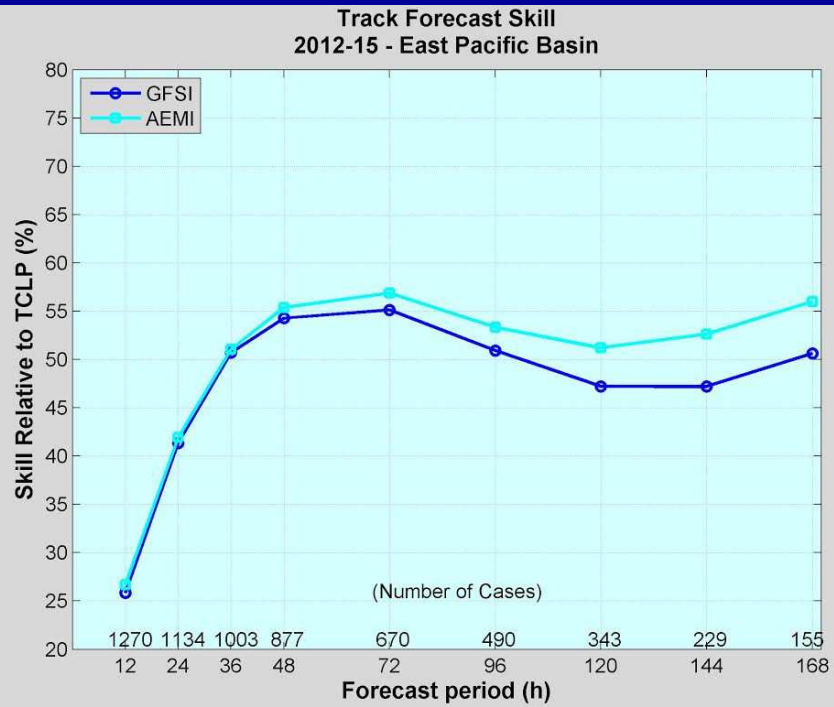
Single-Model Ensembles for TC Track Forecasting

NCEP Global Ensemble Forecast System

Tropical Cyclone Track Forecast Guidance

GEFS Mean vs. GFS (2012-2015)



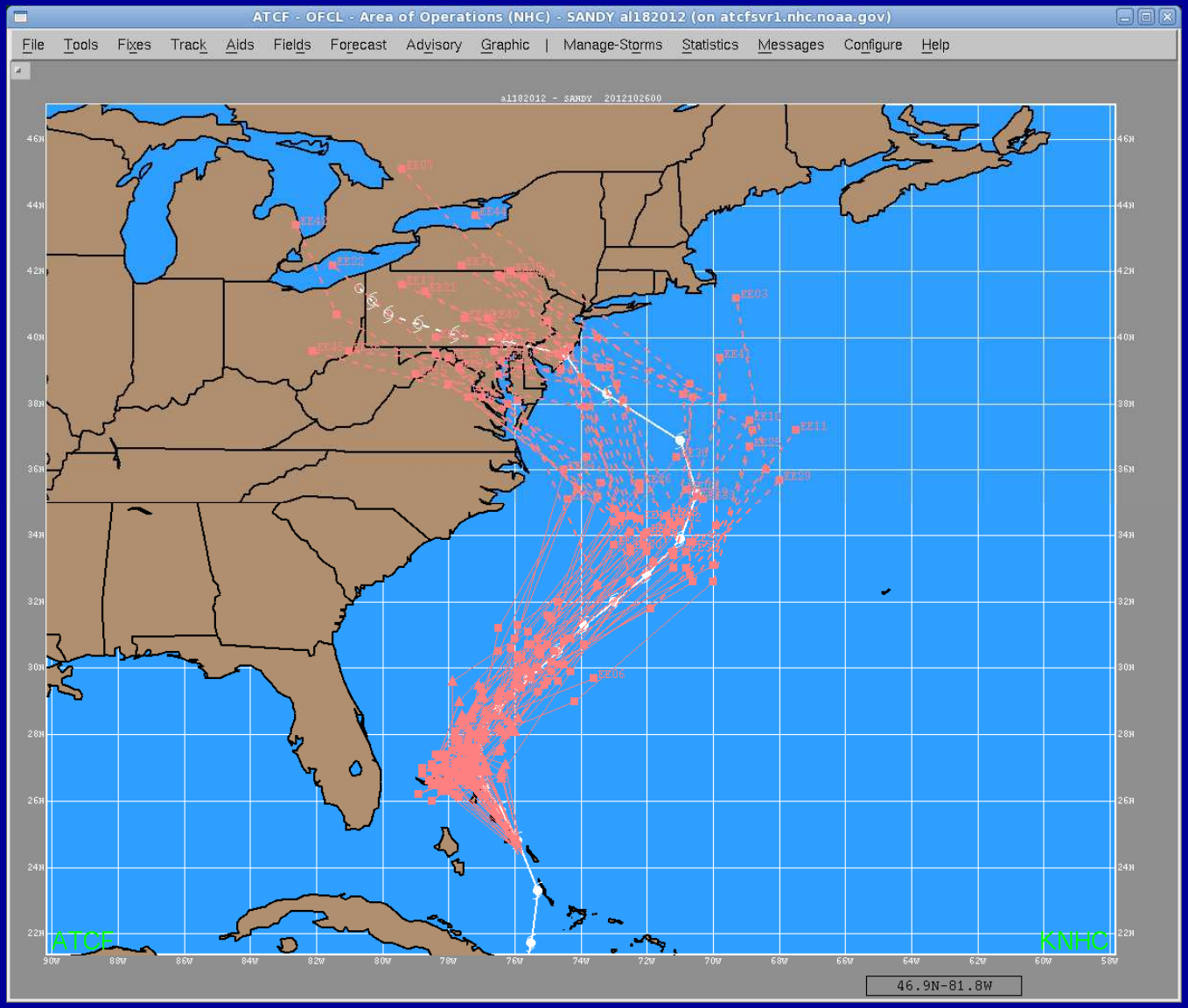


In the Atlantic, the GEFS ensemble mean track forecast (AEMI) is competitive with the deterministic GFS (GFSI) through day 3 and better afterward

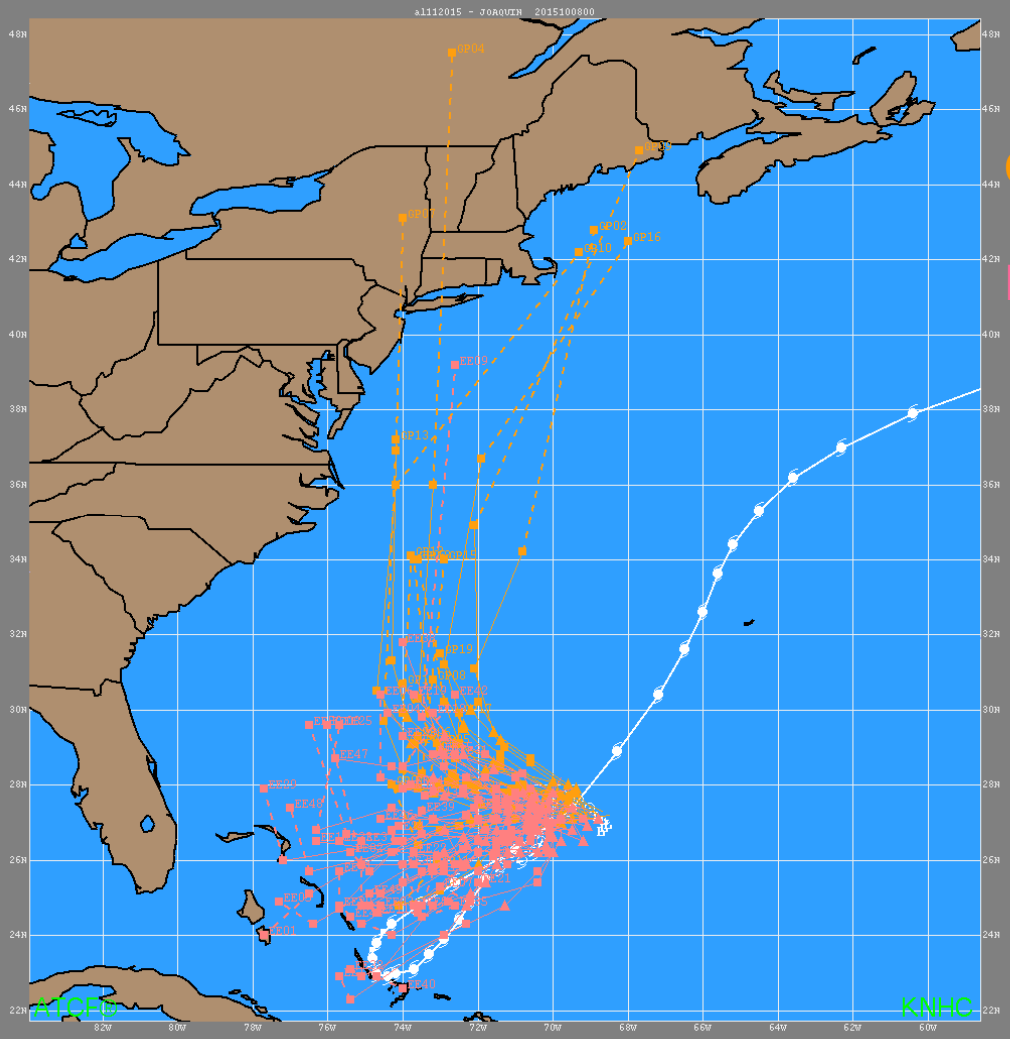
In the east Pacific, AEMI beats GFSI at 48 h and beyond

ECMWF Ensemble

Sandy example of desirable spread/verification



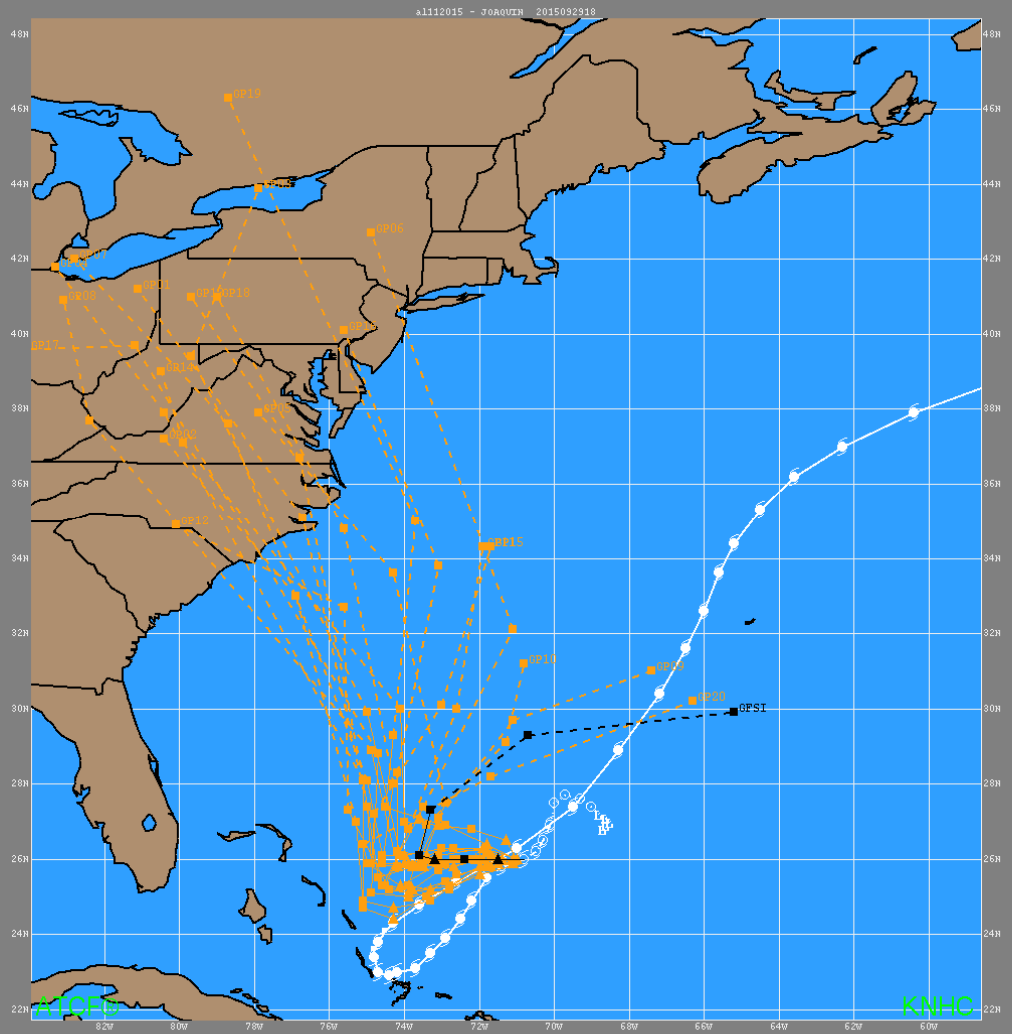
Joaquin ensemble guidance



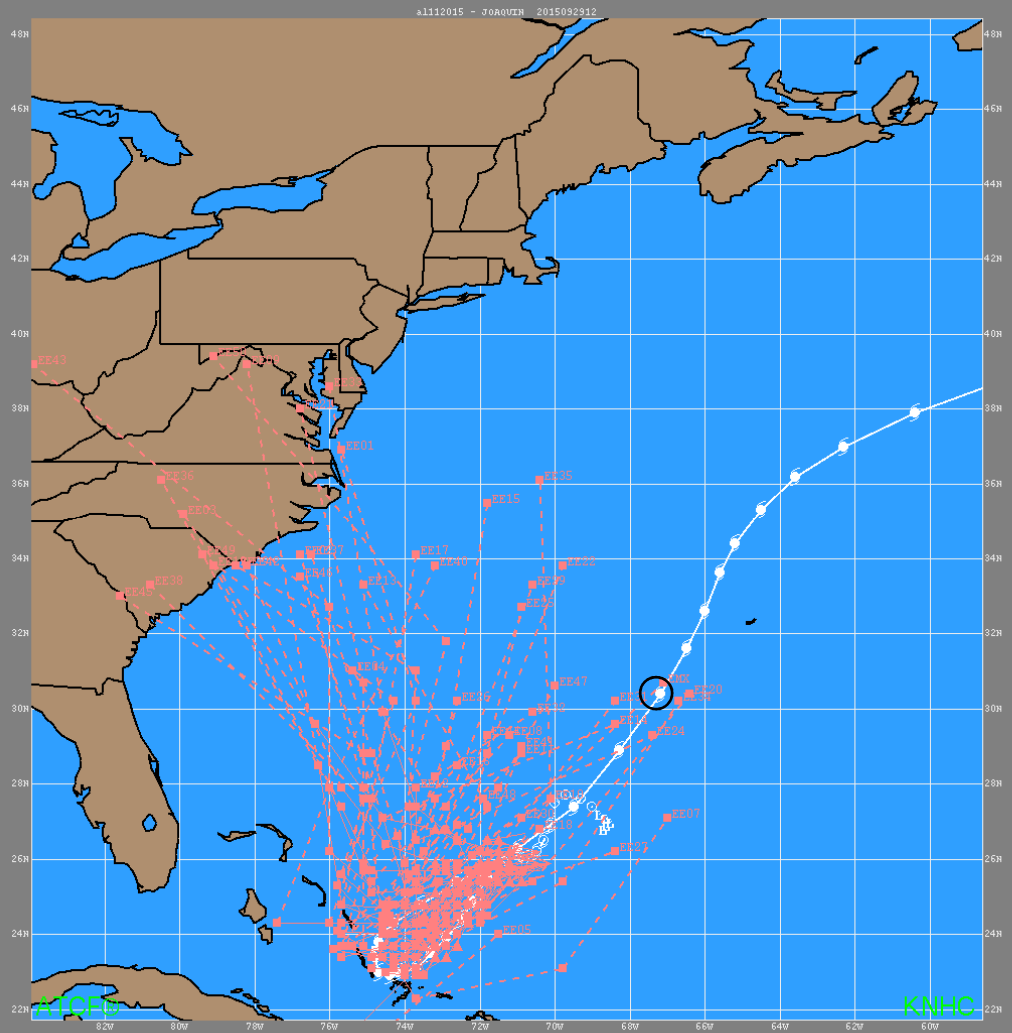
GEFS

EC Ensemble

GFS Joaquin ensembles 29 Sep 1200 UTC

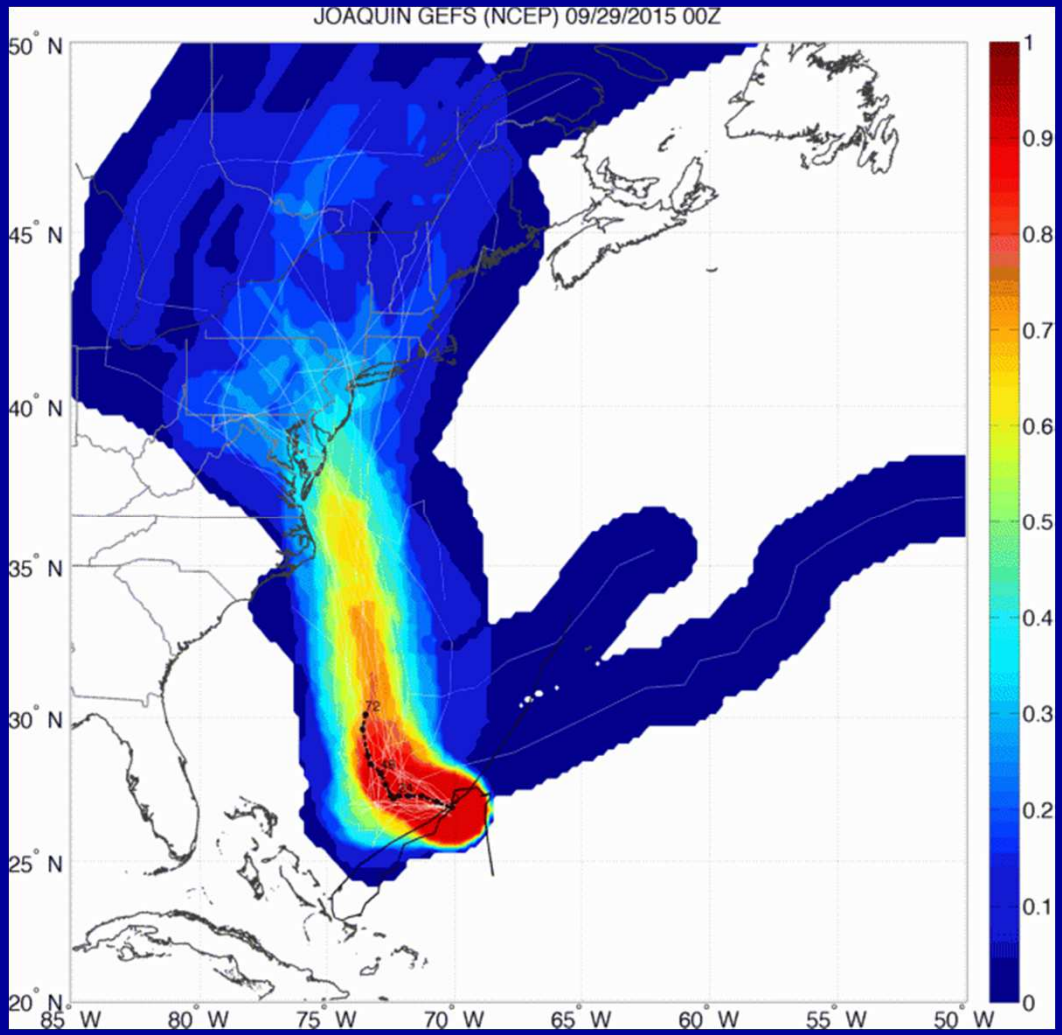


ECMWF Joaquin ensembles 29 Sep 1200 UTC

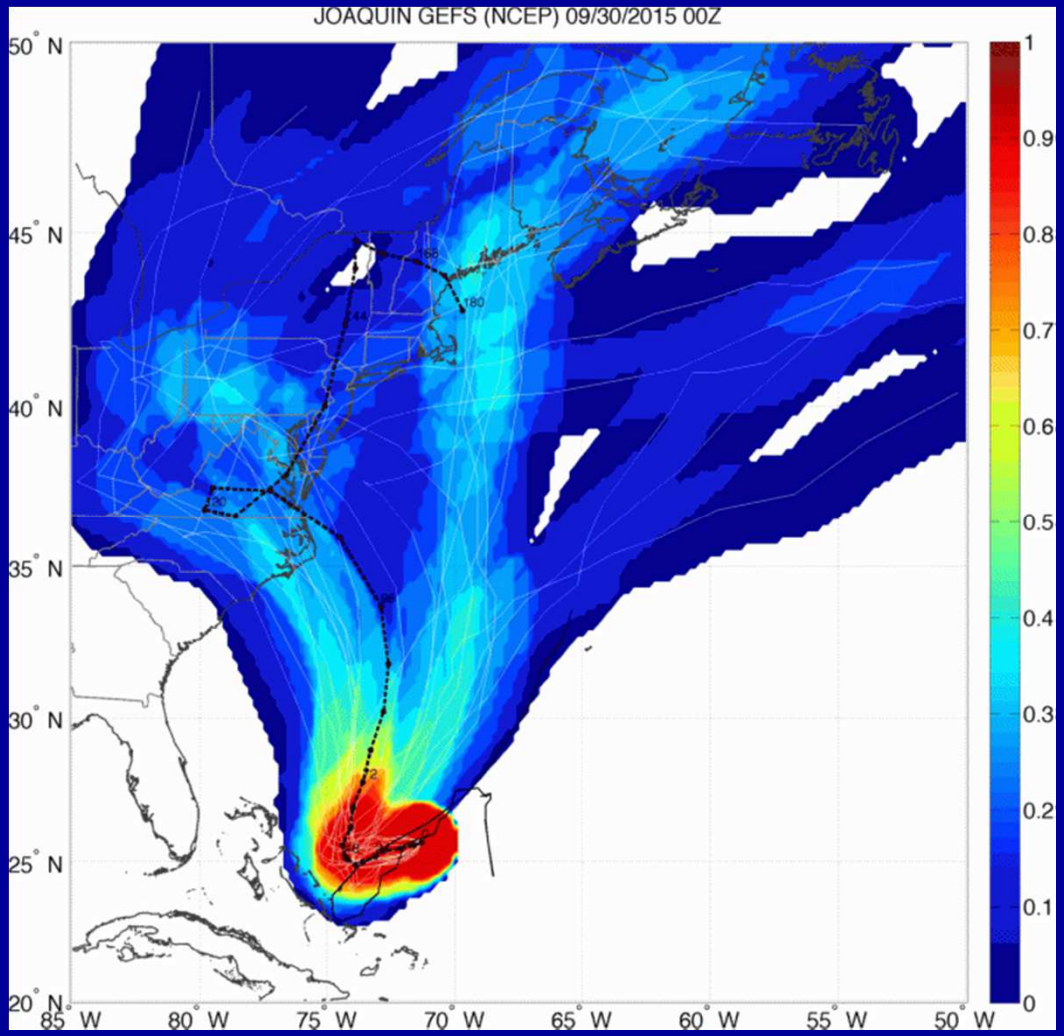


A different way to view the data using probabilities

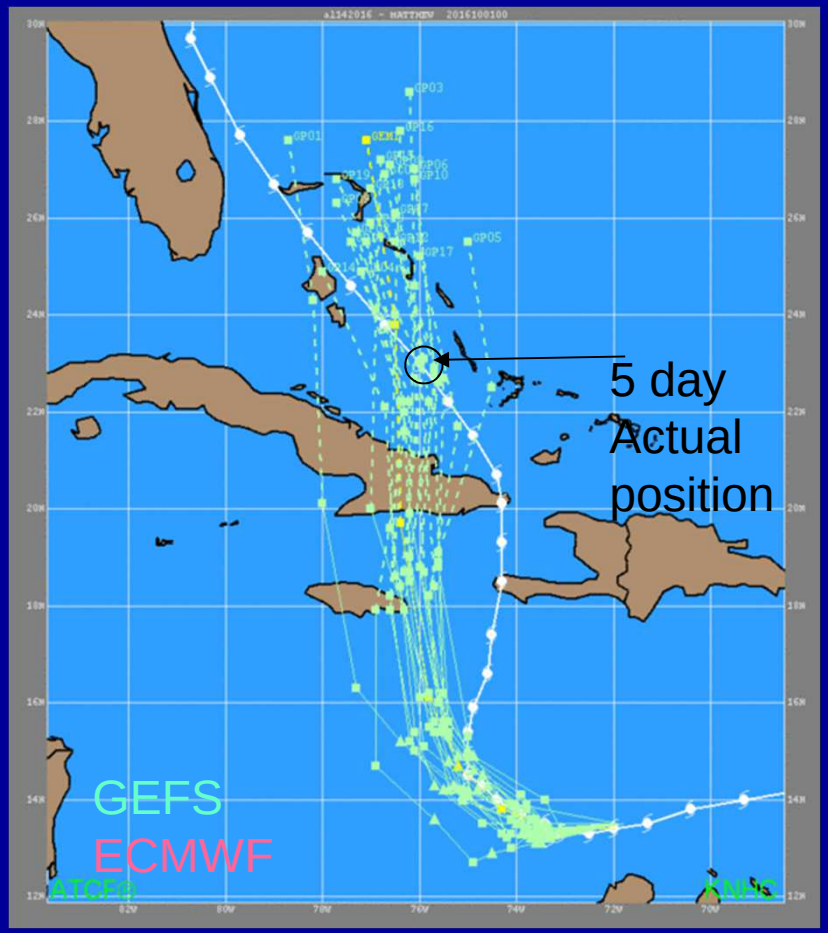
GEFS vs EC Ensemble 29 Sep 0000 UTC



GEFS vs EC Ensemble 30 Sep 0000 UTC



Matthew ensemble guidance 1 Oct 00 UTC

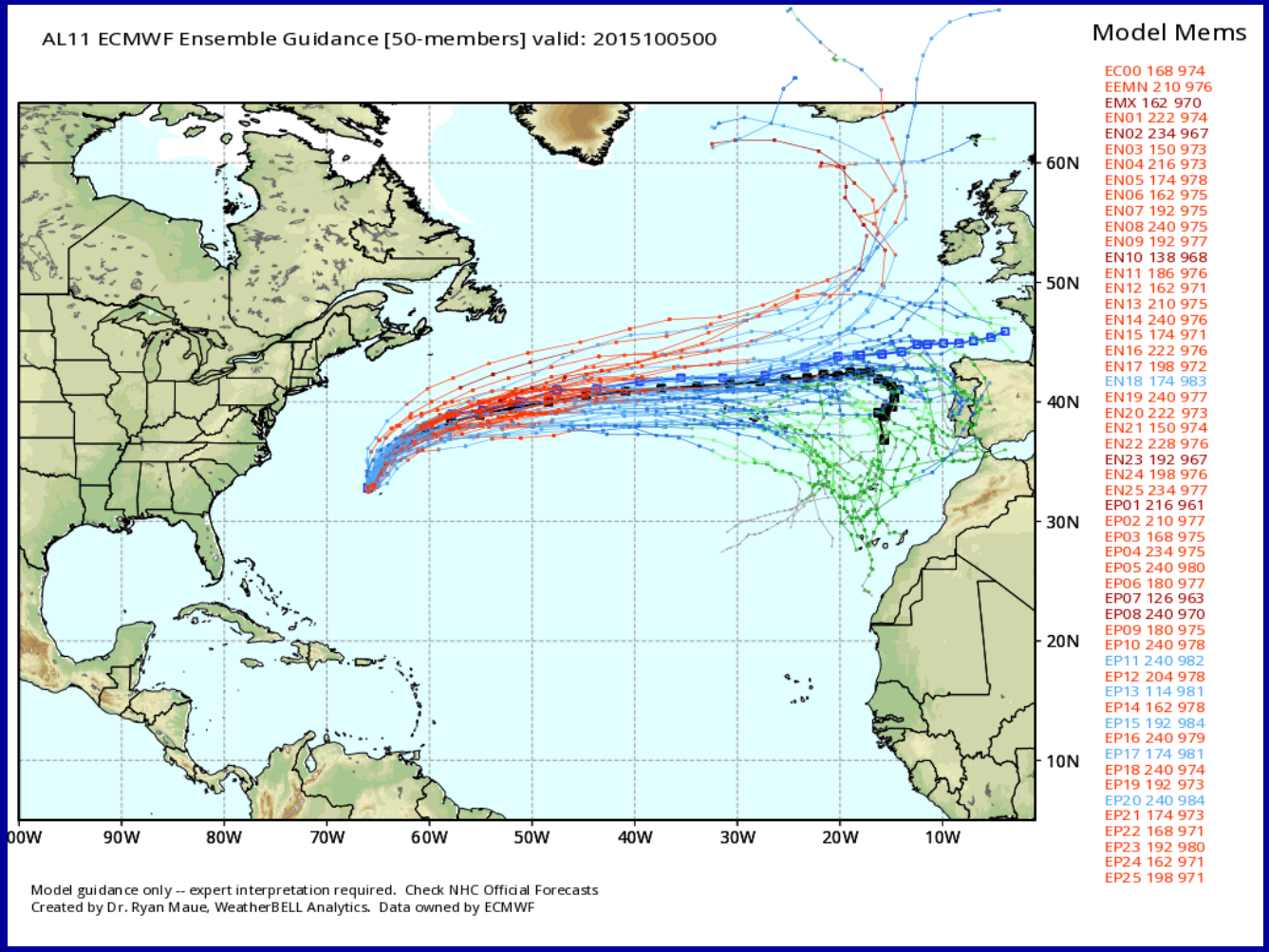


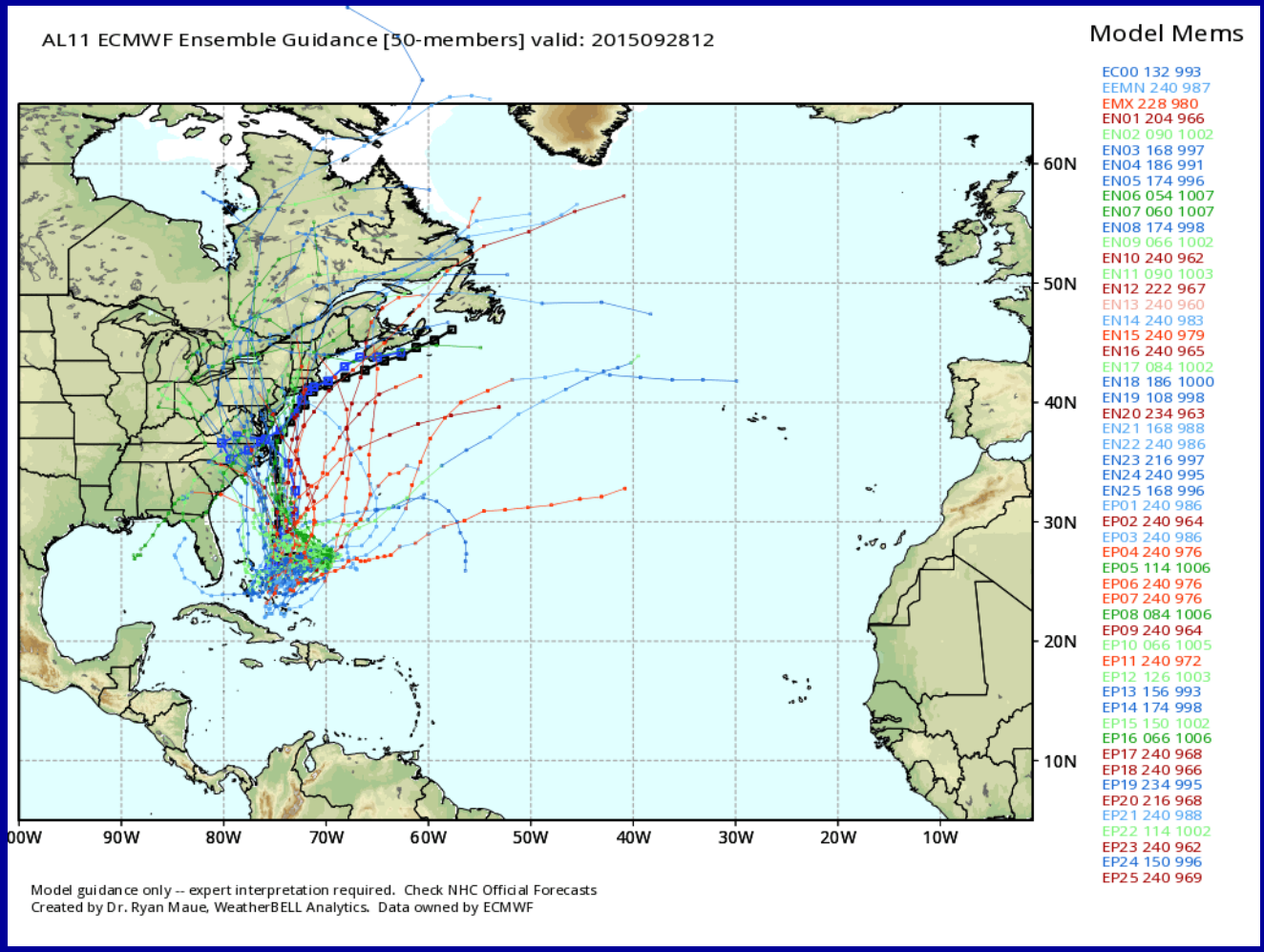
GEFS (blue) too underdispersive, especially in Caribbean

Every single GEFS member also too fast at 5 days

ECMWF (red) has more realistic spreads, albeit potentially too large

ECMWF ensemble colored by intensity





-Stronger members farther right

-Weaker members farther north

Question 2

In which situation(s) is a well-calibrated ensemble system likely to fail?

A. Unusual forecast track cases

B. When TC track is dependent on intensity

C. If deterministic models are in poor agreement

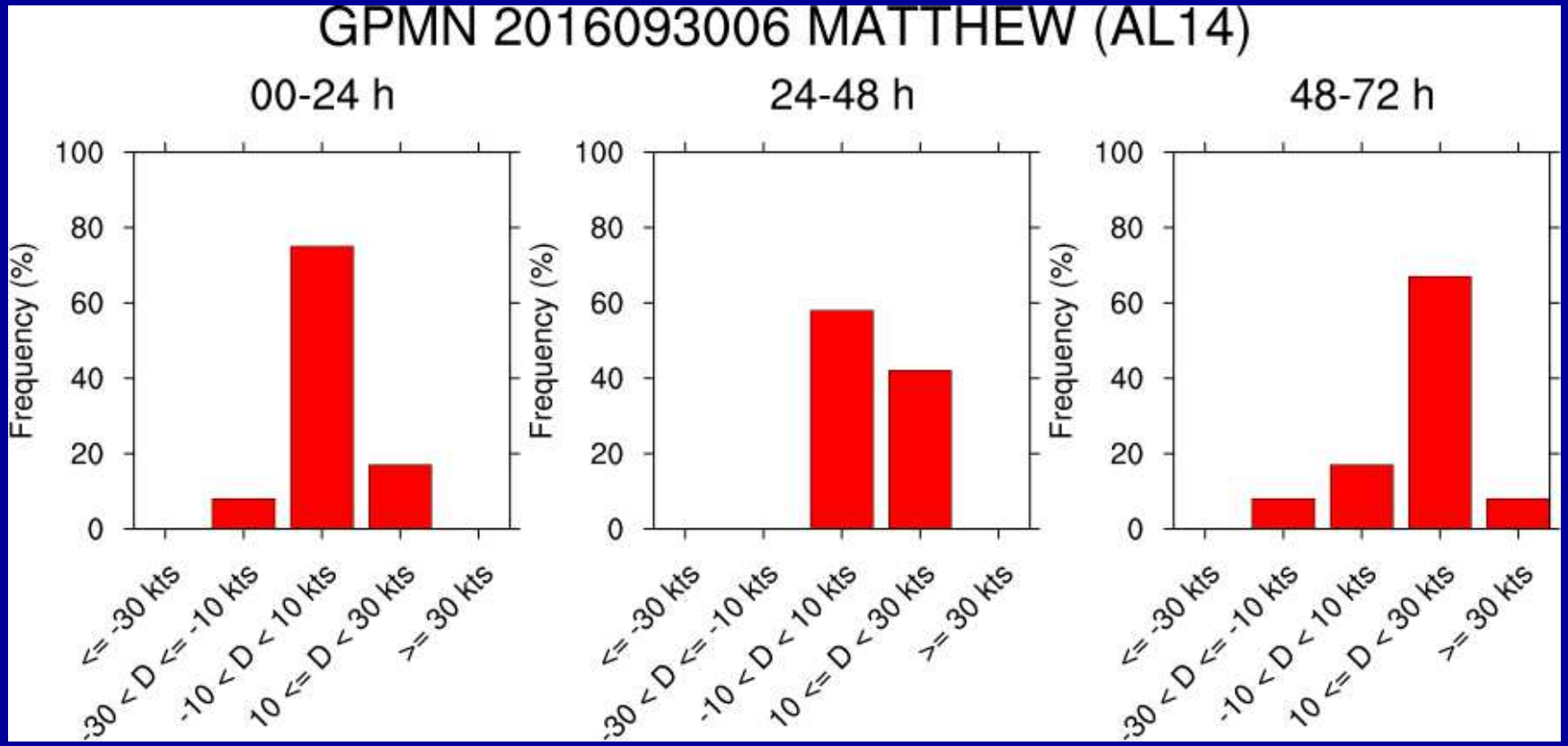
D. All of the above

E. B & C

TC Intensity Ensemble Forecasting

- Little skill above single-model deterministic at present
- Very computational expensive to run high-resolution (<3 km) intensity ensembles
- HFIP is funding efforts to find products that could be operationally useful

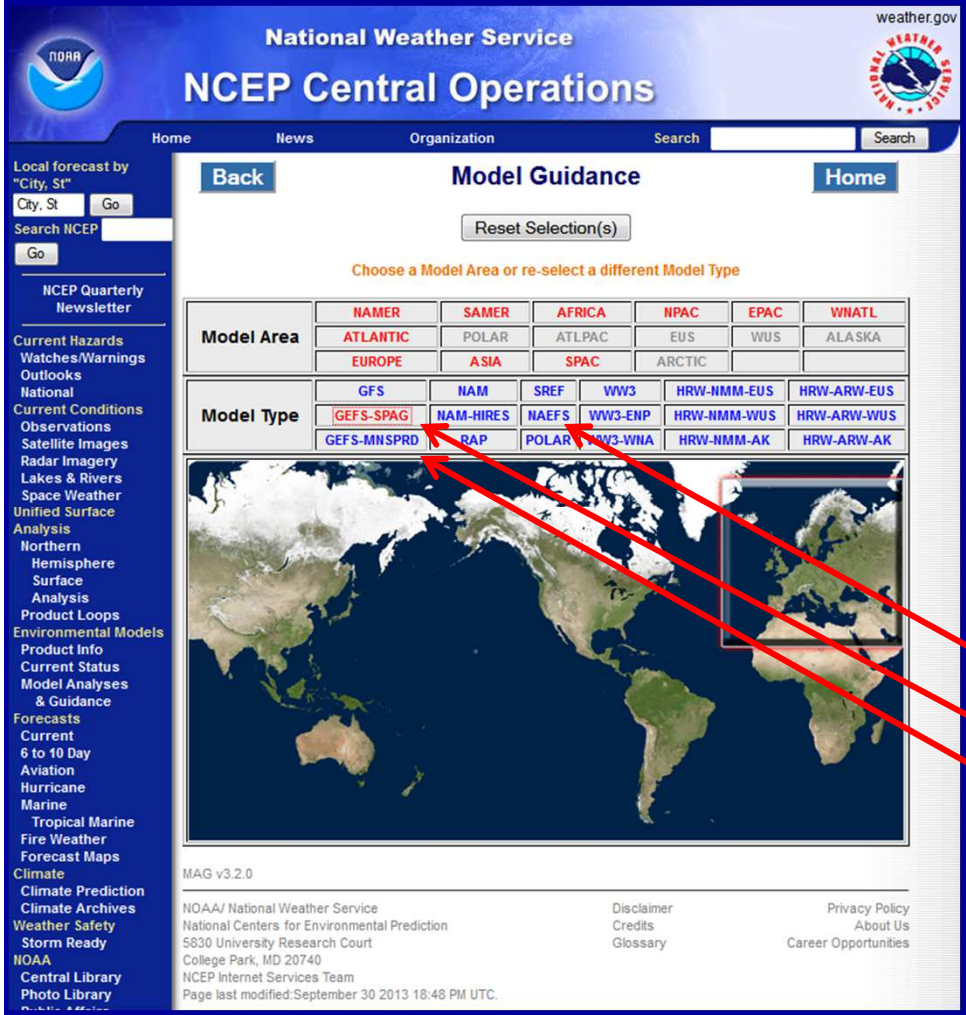
Intensity Change Probability Distributions

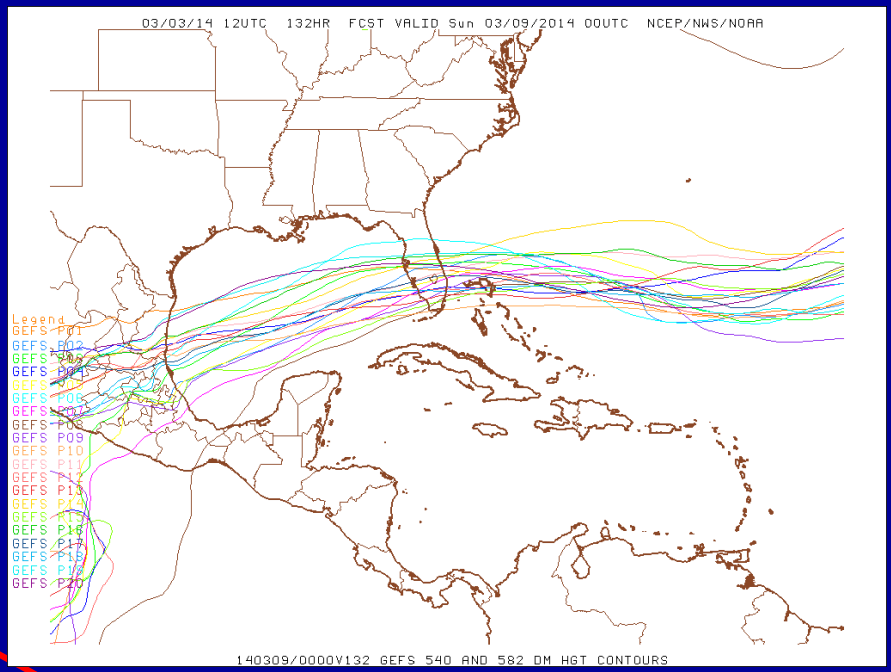


Online Access to Ensemble Output and Training Resources

Access to Ensemble Output

NCEP GEFS and NAEFS: http://mag.ncep.noaa.gov/

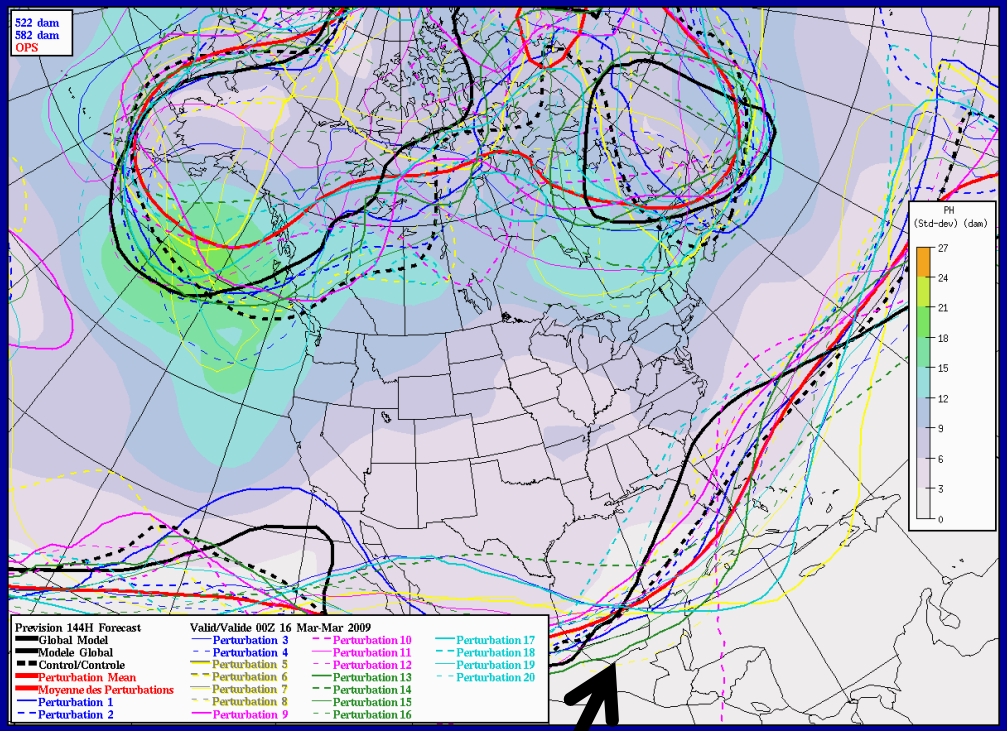




Access to ensemble mean, spread, and spaghetti plots

Canadian Ensembles

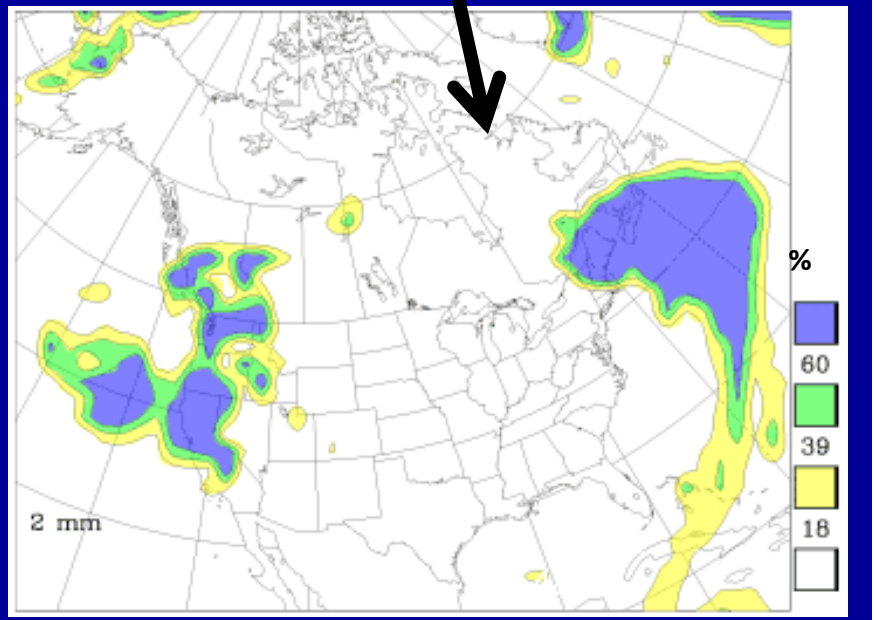
http://weather.gc.ca/ensemble/index_e.html



Spaghetti diagram of 500-mb 522 and 582 dm height contours

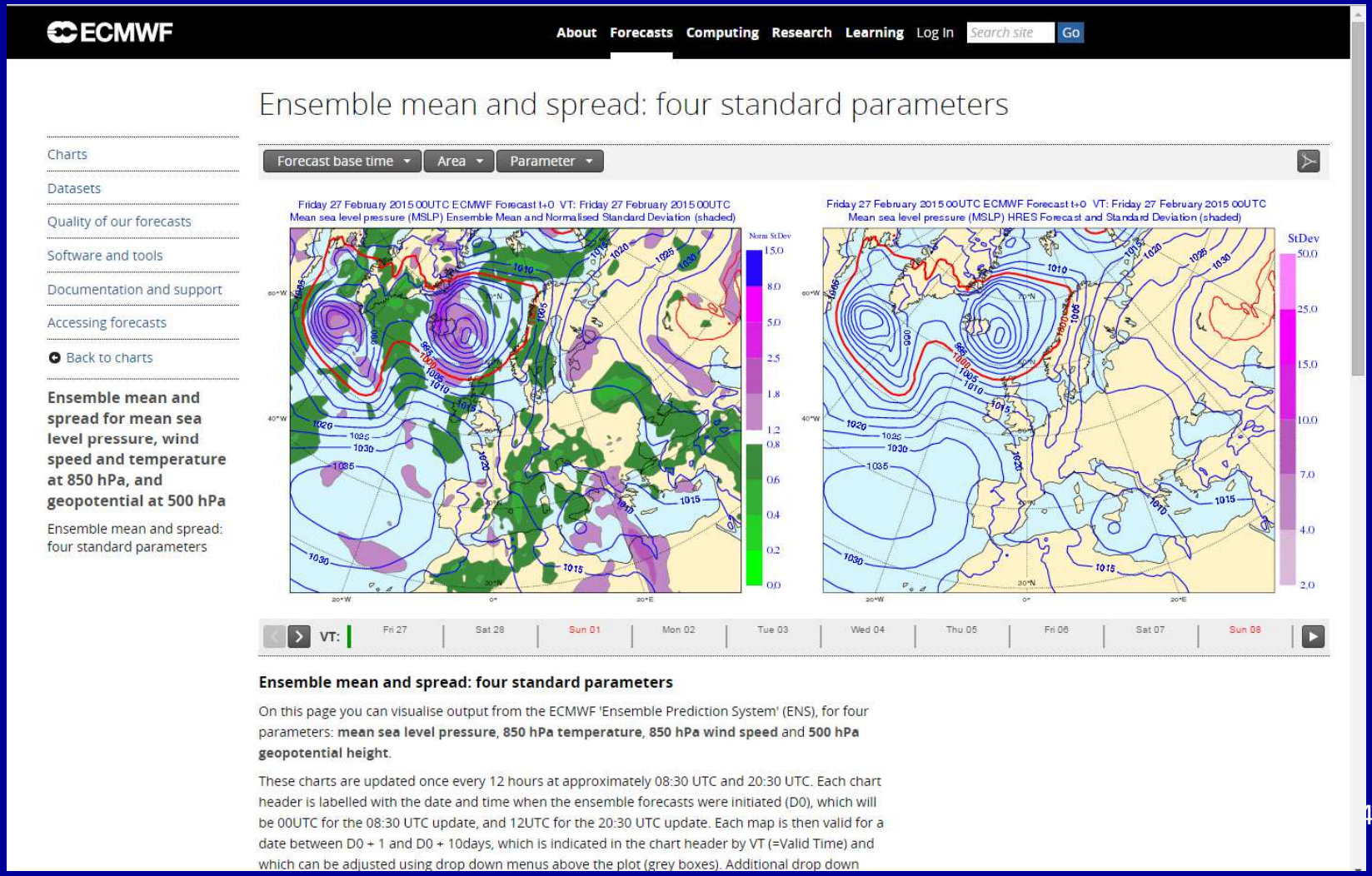
FHR 72 forecast of the probability that the 12 hour accumulation exceeds 2 mm

(The 12-h accumulation period immediately precedes the valid time)



Access to Ensemble Output

- ECMWF Ensembles:
  http://www.ecmwf.int/en/forecasts/charts/medium/ensemble-mean-and-spread-four-standard-parameters



COMET Courses
http://www.meted.ucar.edu

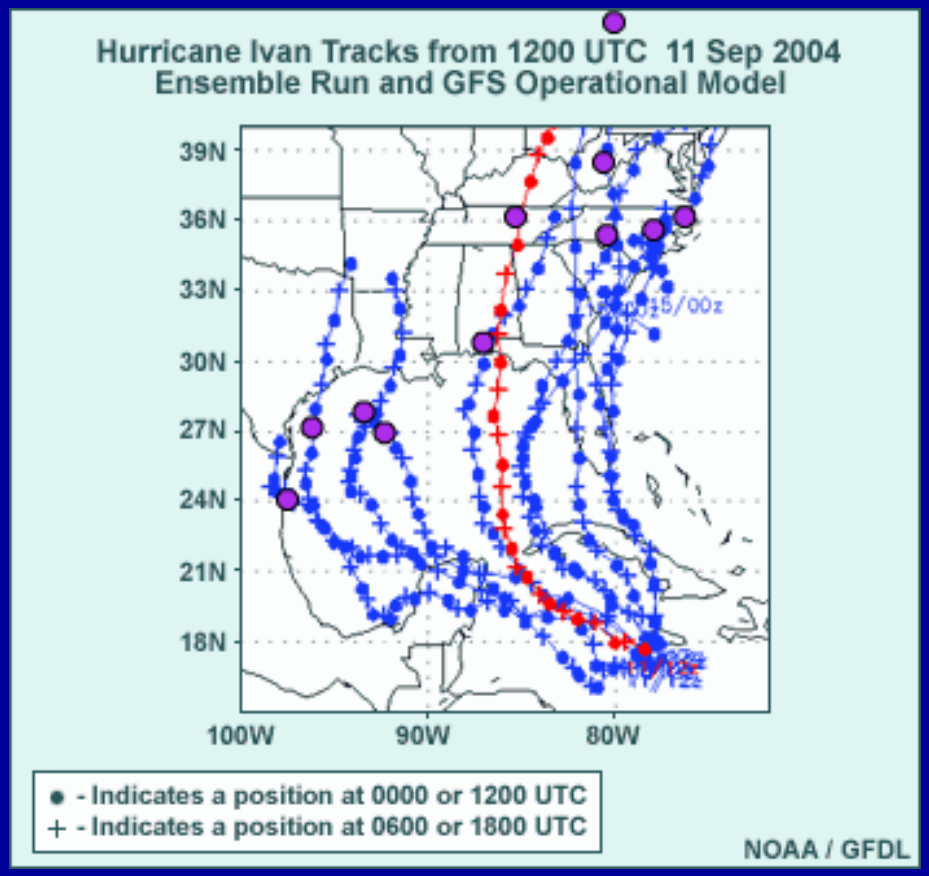
- Introduction to Ensemble Prediction: http://www.meted.ucar.edu/nwp/pcu1/ensemble_webcast/
- Ensemble Forecasting Explained: http://www.meted.ucar.edu/nwp/pcu1/ensemble/
- Ensemble Prediction System Matrix: Characteristics of Operational Ensemble Prediction Systems (EPS): http://www.meted.ucar.edu/nwp/pcu2/ens_matrix/
- Wave Ensembles in the Marine Forecast Process: http://www.meted.ucar.edu/nwp/WaveEnsembles/
- NWP Workshop on WRF and NAEFS: http://www.meted.ucar.edu/s_africa_work/

Thank you

Questions?

Case Example

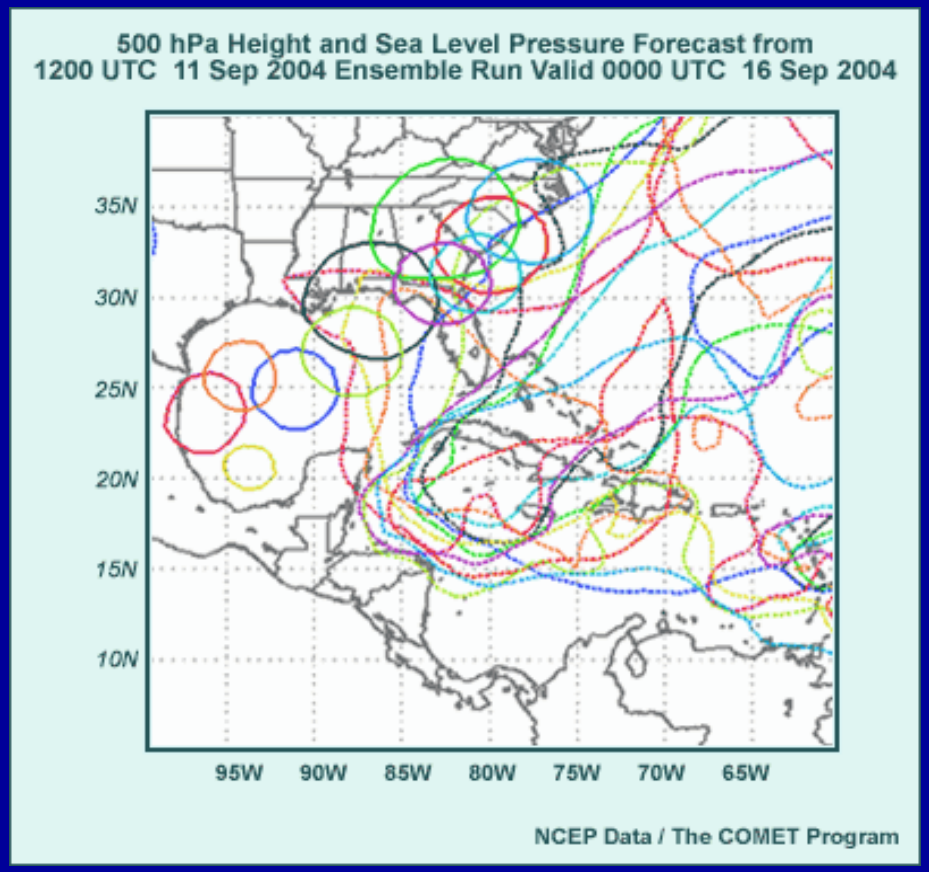
Ensemble Forecast Example



- **Initial time: 1200 UTC 11 Sep 2004 NCEP Ensemble members + and operational GFS +**
- **Purple dots = forecast position at 0300 UTC 17 Sep 2004 (FHR135)**
- **Ensemble forecast shows large uncertainty in ultimate path of Hurricane Ivan**
- **Tendency for clustering of tracks**
  - 5 members east of the GFS track and faster than GFS at 0300 UTC 17 Sep 2004
  - 4 members west of GFS
  - Operational GFS and 1 member in the middle of the ensemble solutions

Credit: COMET

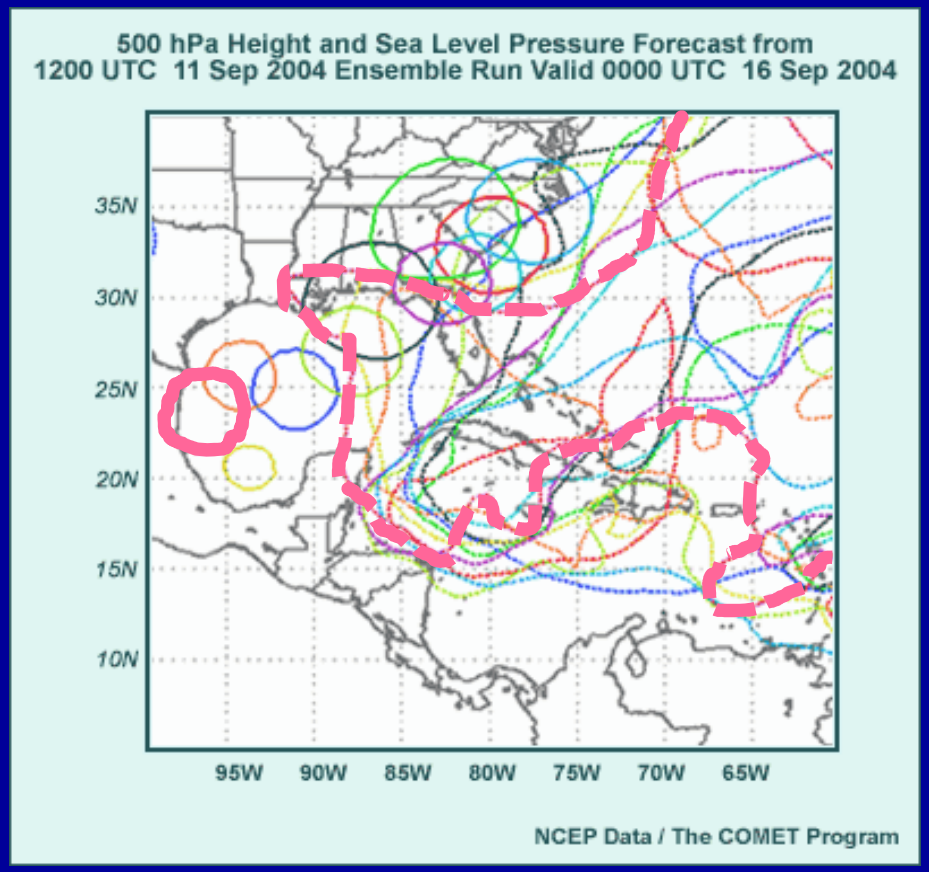
Ensemble Forecast Example



- **Forecast: 0000 UTC 16 Sept 2004 108-hour NCEP ensemble forecast**
- **500-mb 589-dm height (dashed) and 1000-mb PMSL (solid), color coded by ensemble member**
- **Degree of weakening of western Atlantic ridge over the northeast Gulf of Mexico determines position of Hurricane Ivan**
  - Ridge strongest in pink: Ivan near northeastern Mexico, 589-dm height contour in mid-Gulf
  - Ridge weakest in light blue: Ivan over the Georgia coast, 589-dm height contour over the western Atlantic/northwest Caribbean

Credit: COMET

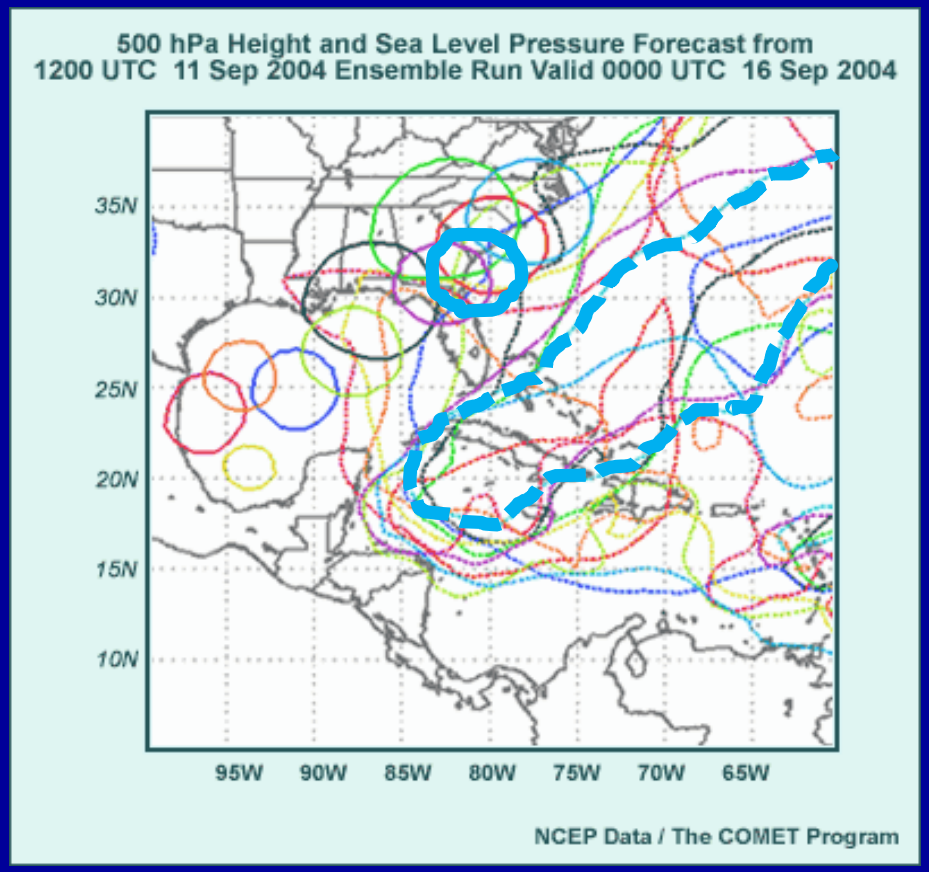
Ensemble Forecast Example



- Forecast: 0000 UTC 16 Sept 2004 108-hour NCEP ensemble forecast
- 500-mb 589-dm height (dashed) and 1000-mb PMSL (solid), color coded by ensemble member
- Degree of weakening of western Atlantic ridge over the northeast Gulf of Mexico determines position of Hurricane Ivan
  - Ridge strongest in pink: Ivan near northeastern Mexico, 589-dm height contour in mid-Gulf
  - Ridge weakest in light blue: Ivan over the Georgia coast, 589-dm height contour over the western Atlantic/northwest Caribbean

Credit: COMET

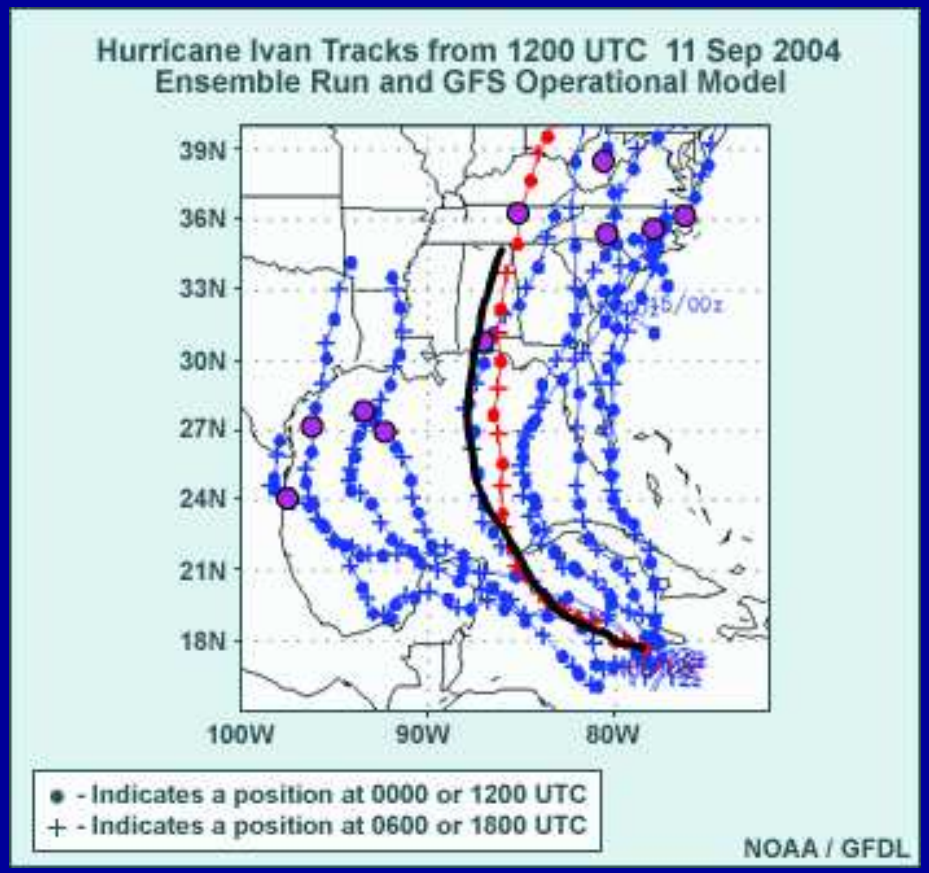
Ensemble Forecast Example



- Forecast: 0000 UTC 16 Sept 2004 108-hour NCEP ensemble forecast
- 500-mb 589-dm height (dashed) and 1000-mb PMSL (solid), color coded by ensemble member
- Degree of weakening of western Atlantic ridge over the northeast Gulf of Mexico determines position of Hurricane Ivan
  - Ridge strongest in pink: Ivan near northeastern Mexico, 589-dm height contour in mid-Gulf
  - Ridge weakest in light blue: Ivan over the Georgia coast, 589-dm height contour over the western Atlantic/northwest Caribbean

Credit: COMET

Ensemble Forecast Example



- **Ultimate path for Hurricane Ivan (black) - not too far from GFS and in the middle of the ensemble envelope of solutions**
  - Wide envelope of possible tracks
  - Because of uncertainty in the weakening of the Atlantic ridge, it turned out to be the best solution
  - Typically, one would be wary of using the ensemble mean forecast when there is clustering of the solutions
  - Look at the handling of the ridge by the other dynamical models to determine which "cluster" to lean toward

Credit: COMET