

NOAA Earth System
Research Laboratory

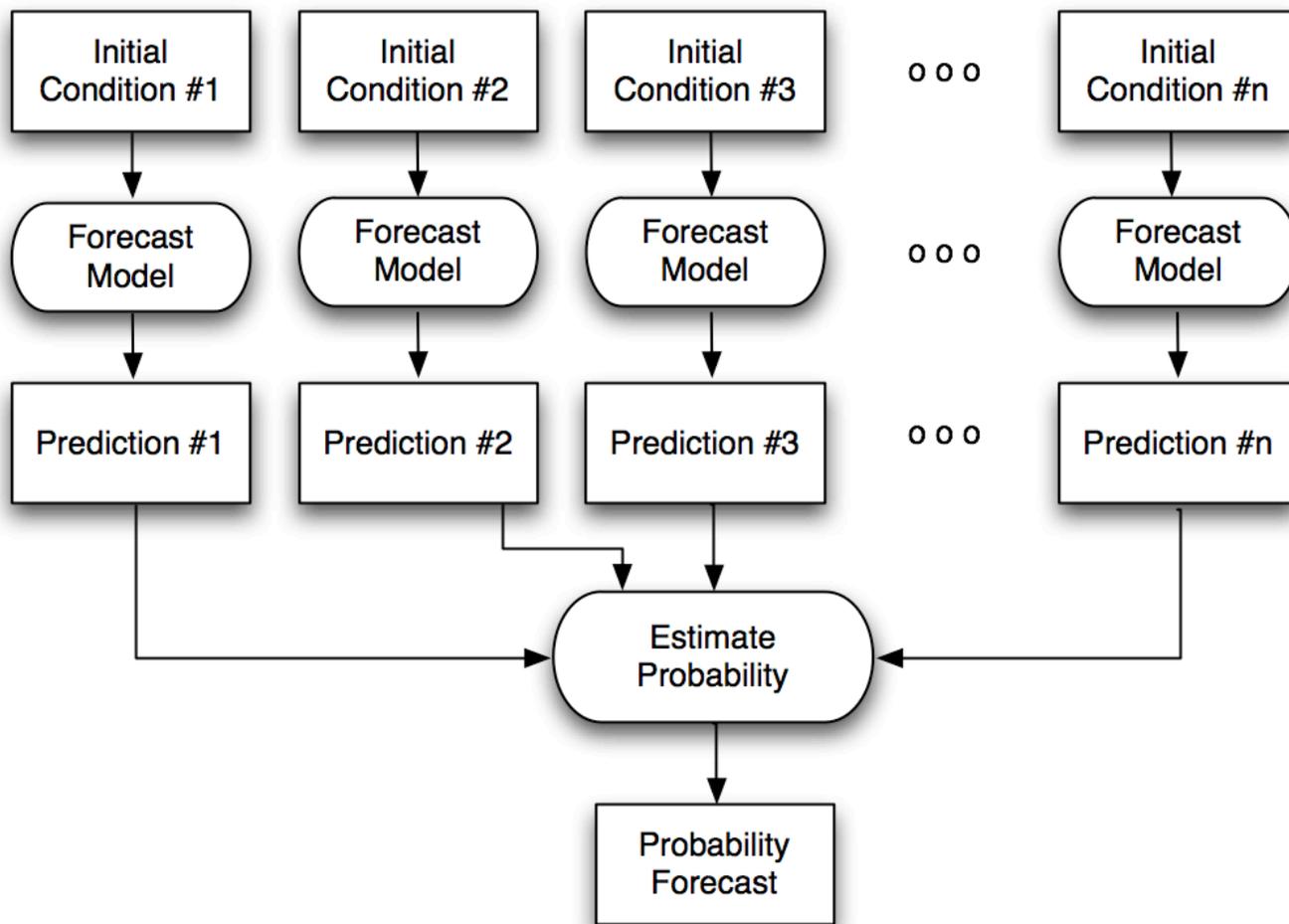
Review of ensemble prediction fundamentals

Tom Hamill

NOAA ESRL, Physical Sciences Division

tom.hamill@noaa.gov

“Ensemble weather prediction”



Topics

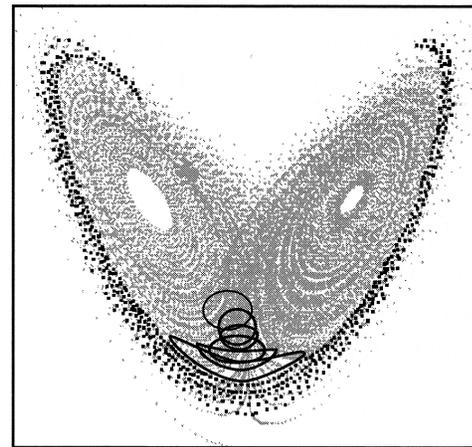
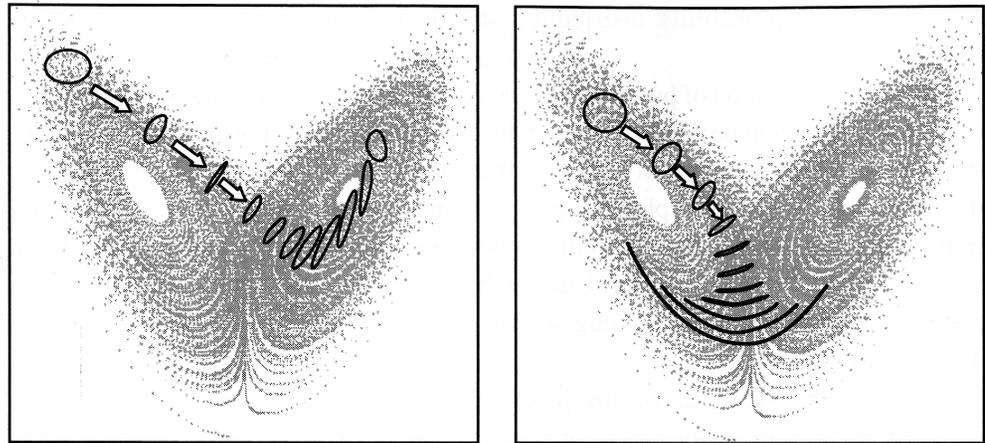
- Chaos theory & ensembles
- Desired properties of ensembles
- Initializing ensembles
- Dealing with model error
 - mostly in Carolyn Reynolds' talk
- Ensembles & hurricanes
- Some product ideas

“Chaos”

The Lorenz (1963) model

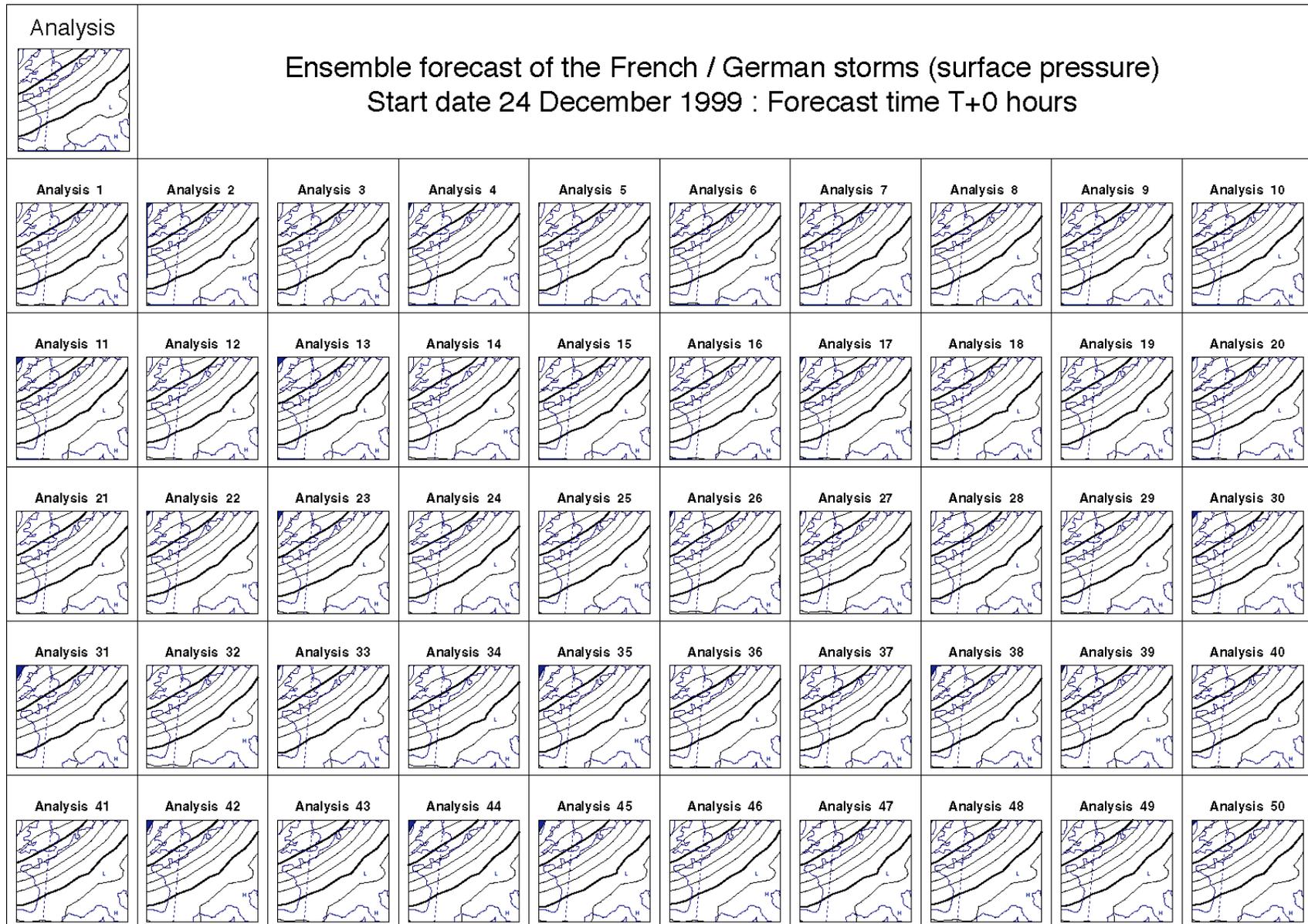
$$\begin{aligned}\frac{dx}{dt} &= \sigma(y - x) \\ \frac{dy}{dt} &= x(\rho - z) - y \\ \frac{dz}{dt} &= xy - \beta z\end{aligned}$$

σ, ρ, β are fixed.



Errors grow more quickly for some initial conditions than others.

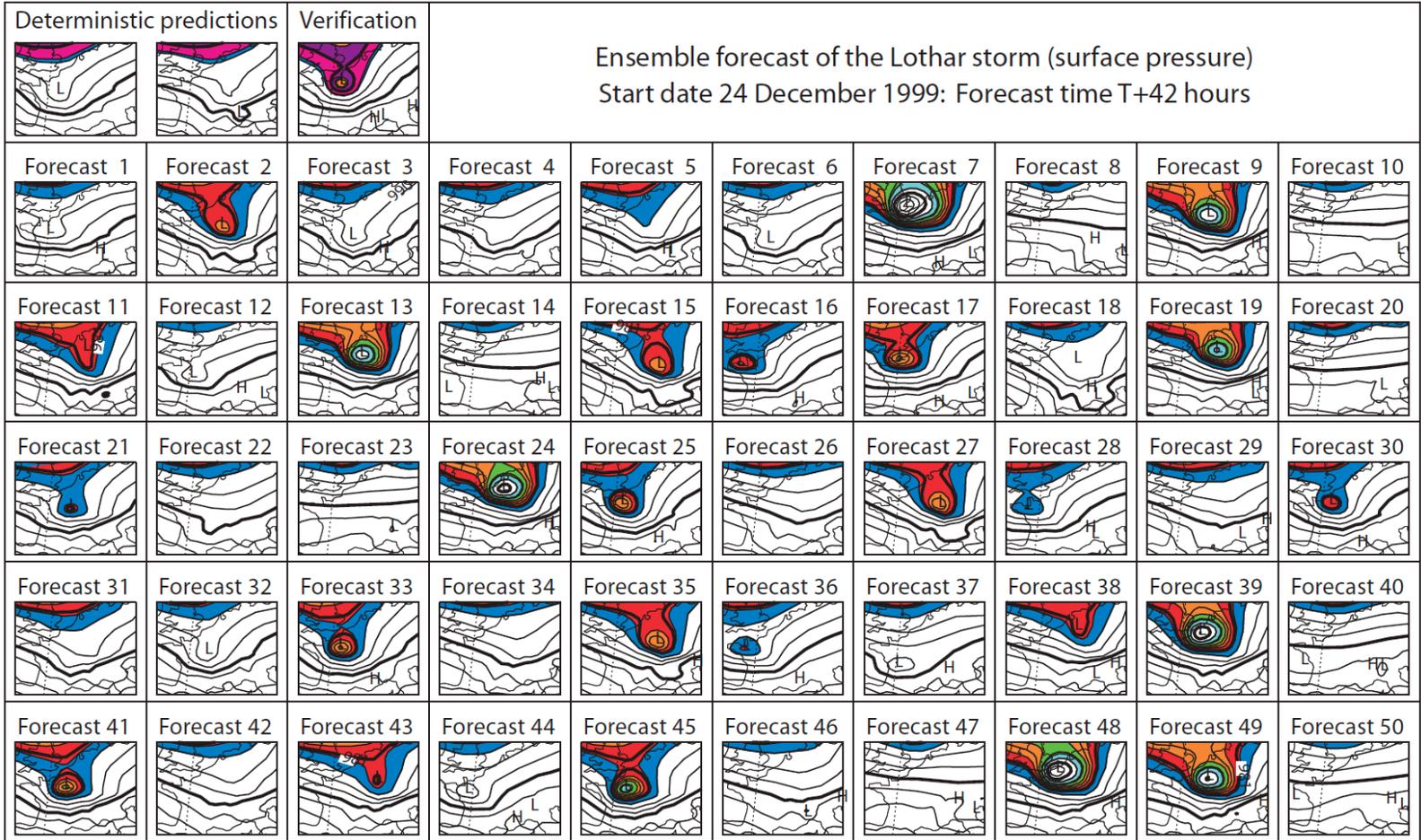
Initial conditions for “Lothar” ensemble forecasts



Lothar 42-h MSLP forecasts

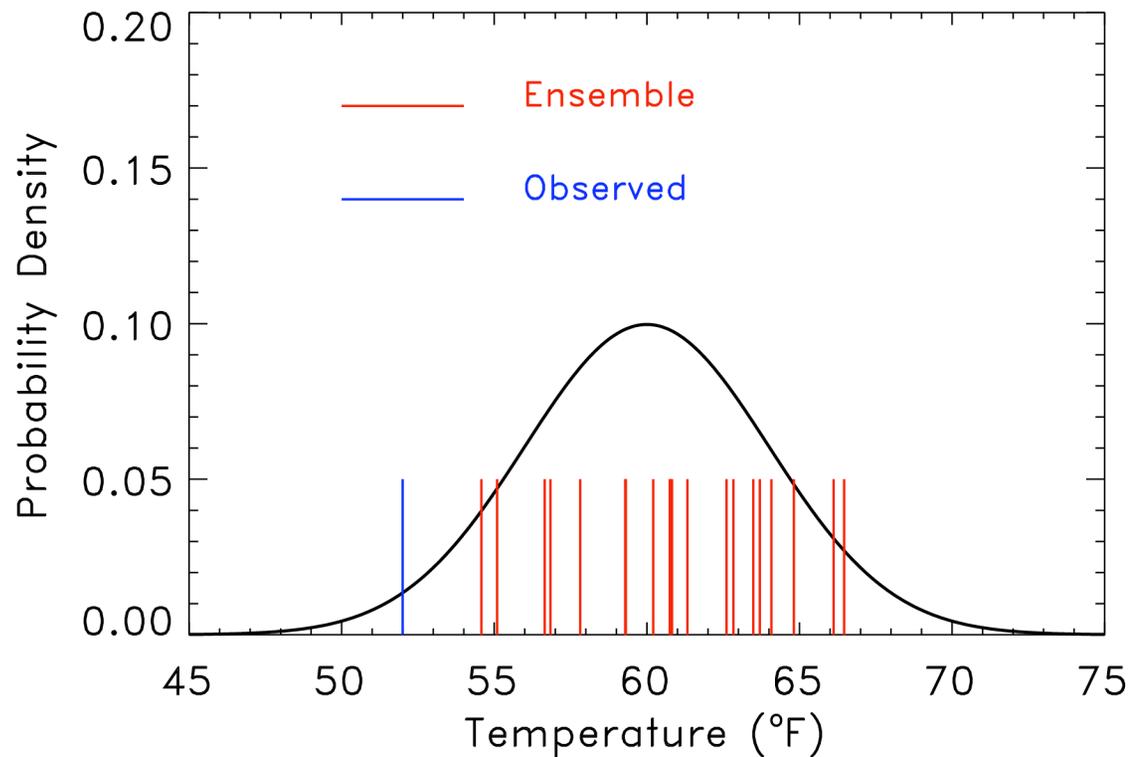
deterministic forecast → totally misses damaging storm over France; some ensemble members forecast it well.

Probabilities commonly estimated from frequency of event in the ensemble.



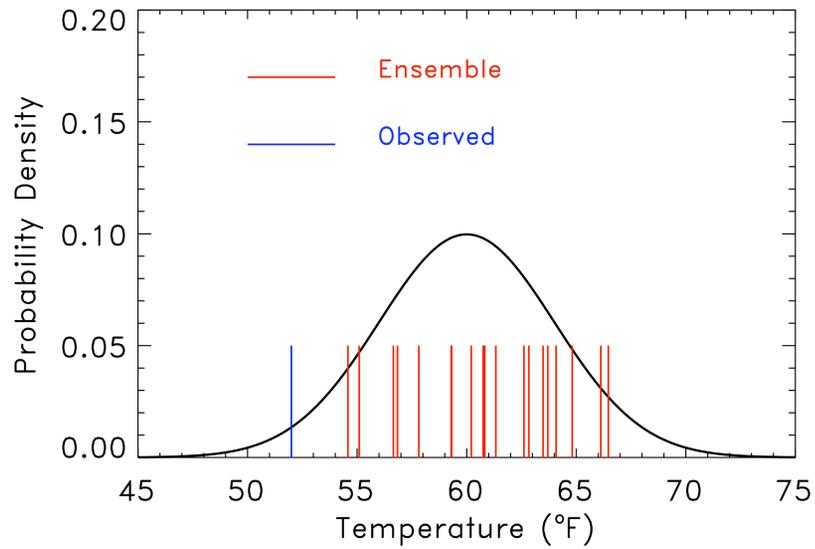
from Tim Palmer's book chapter, 2006.

Question: what constitutes a “good” ensemble forecast?

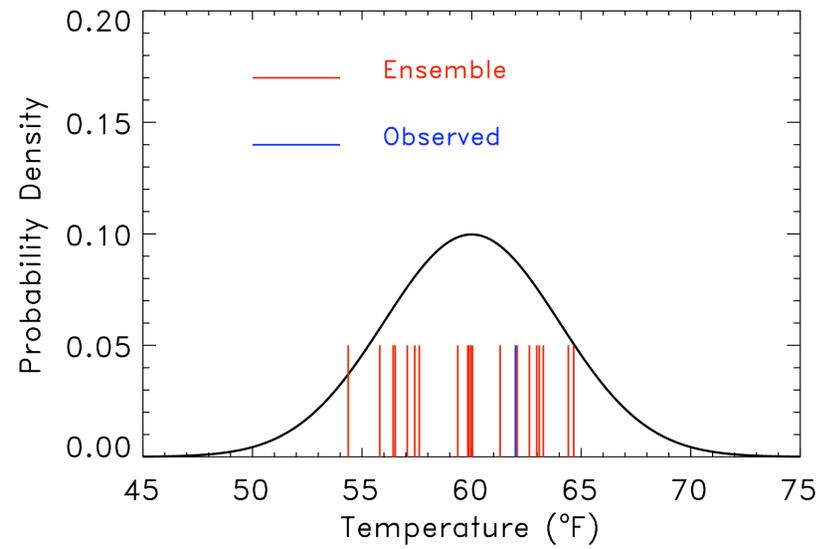


Here, the observed is outside of the range of the ensemble, which was sampled from the pdf shown. **Is this a sign of a poor ensemble forecast?**

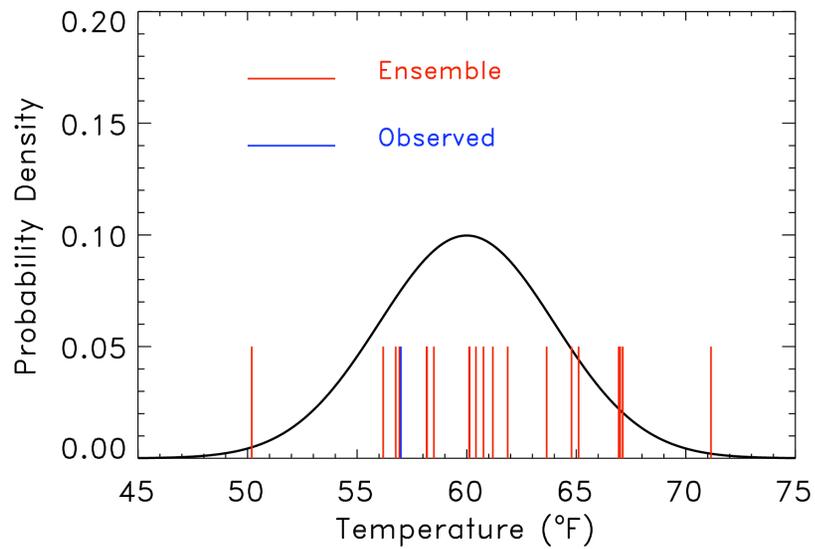
Rank 1 of 21



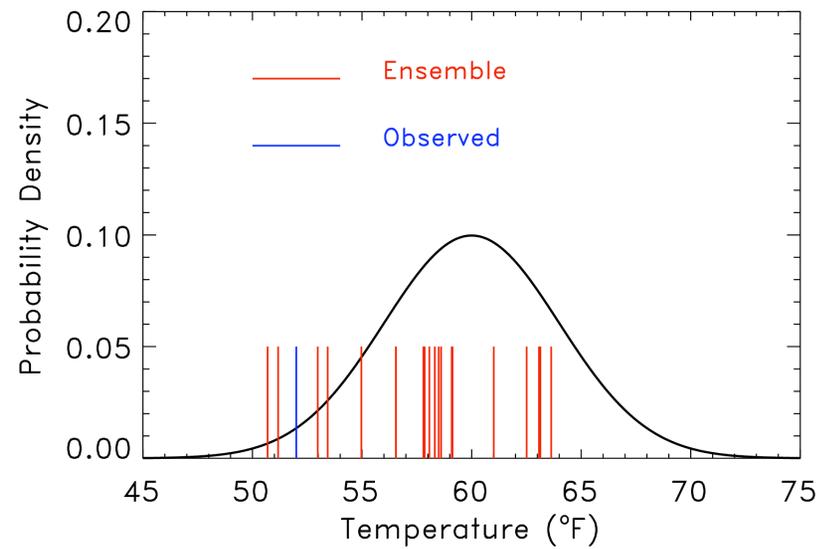
Rank 14 of 21



Rank 5 of 21

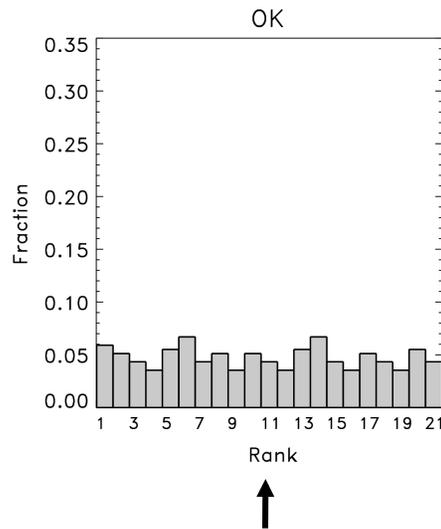


Rank 3 of 21

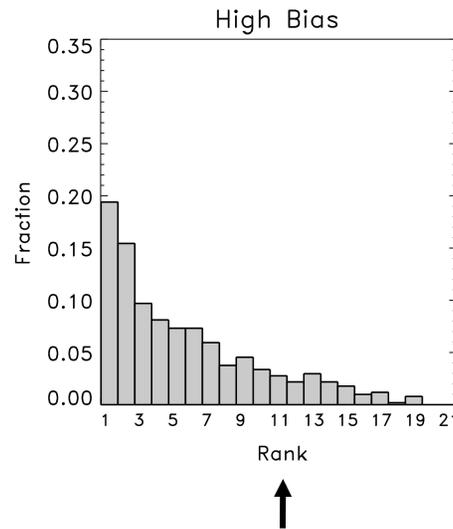


Ensembles and truth should be draws from the same distribution

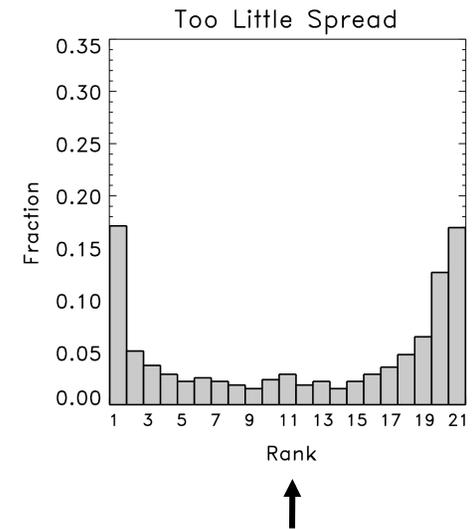
We need lots of samples from many situations to evaluate the characteristics of the ensemble.



Happens when observed is **indistinguishable** from any other member of the ensemble.

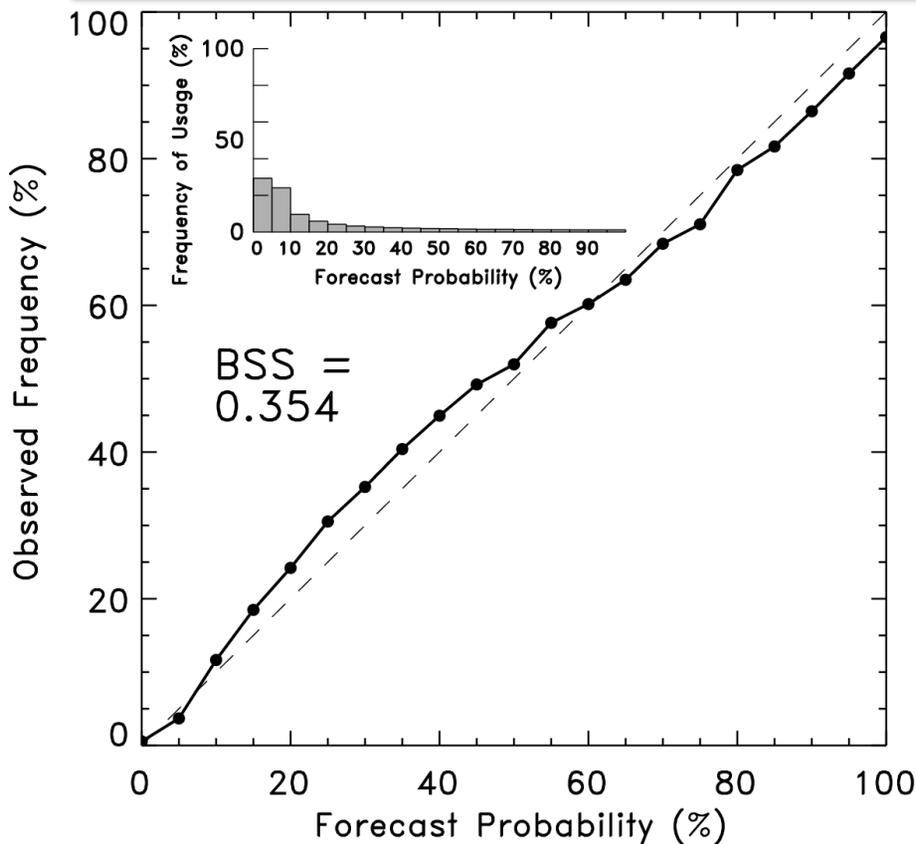


Happens when observed too commonly is lower than the ensemble members.



Happens when there are either some low and some high biases, or when the ensemble doesn't spread out enough.

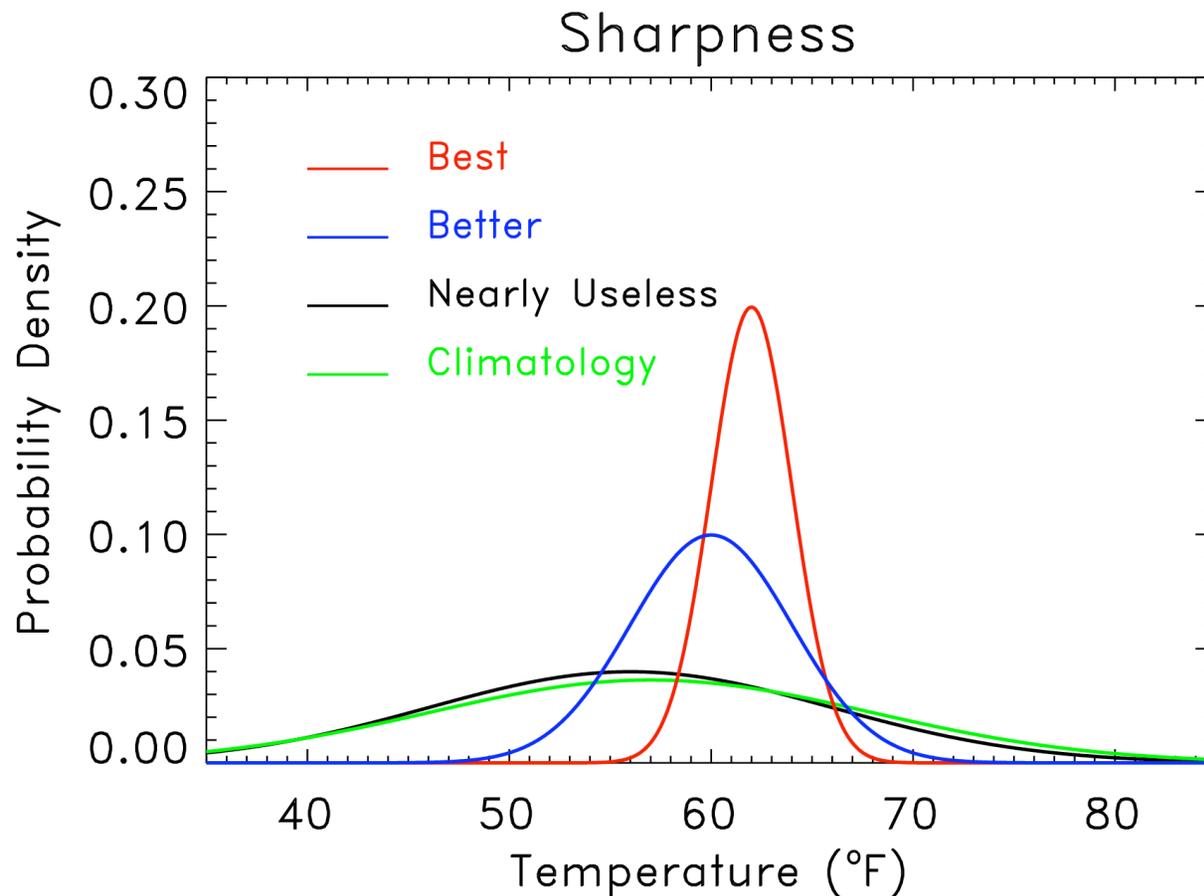
Such ensemble forecasts will be “reliable” (if there are enough members)



In a reliable forecast, the event occurs at the same relative frequency as the probability you forecast.

But even if the ensemble and the truth are drawn from the same distribution, with a small ensemble you won't get reliable probabilities due to sampling error (Richardson, QJRMS, 2001)

What other characteristics of ensemble forecasts are important?

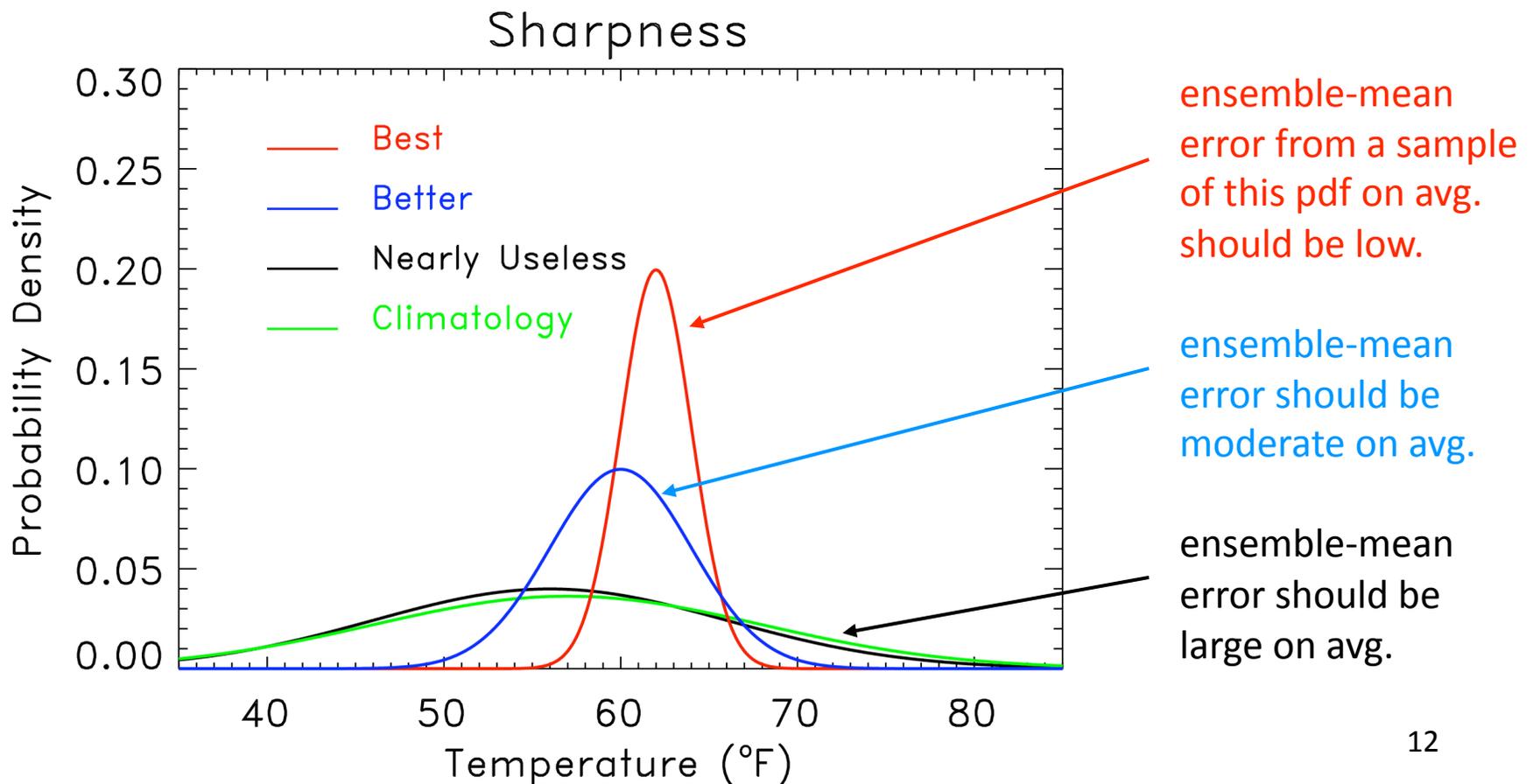


“Sharpness” measures the specificity of the probabilistic forecast. Given two reliable forecast systems, the one producing the sharper forecasts is preferable.

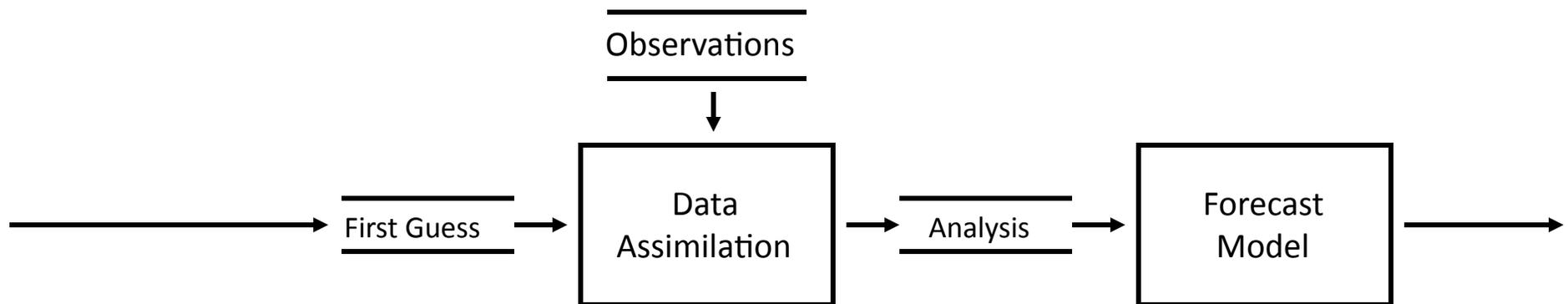
But: don’t want sharp if not reliable. Implies unrealistic confidence.

“Spread-skill” relationship

Small-spread ensemble forecasts should have less ensemble-mean error than large-spread forecasts.

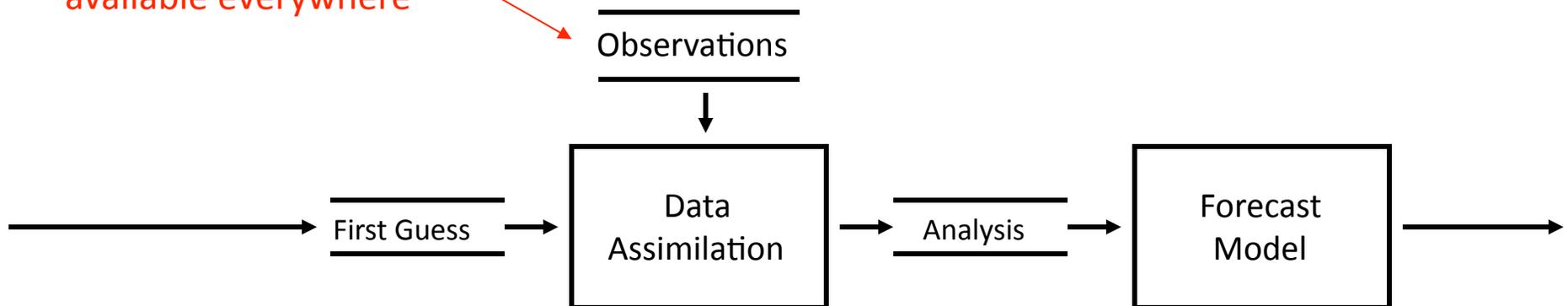


Why run forecasts from many initial conditions?



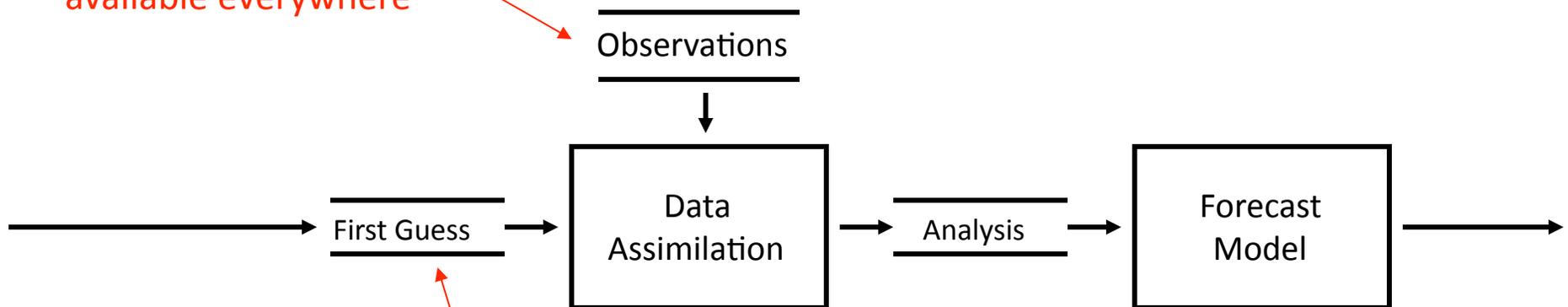
Why run forecasts from many initial conditions?

these also have errors,
and observations aren't
available everywhere



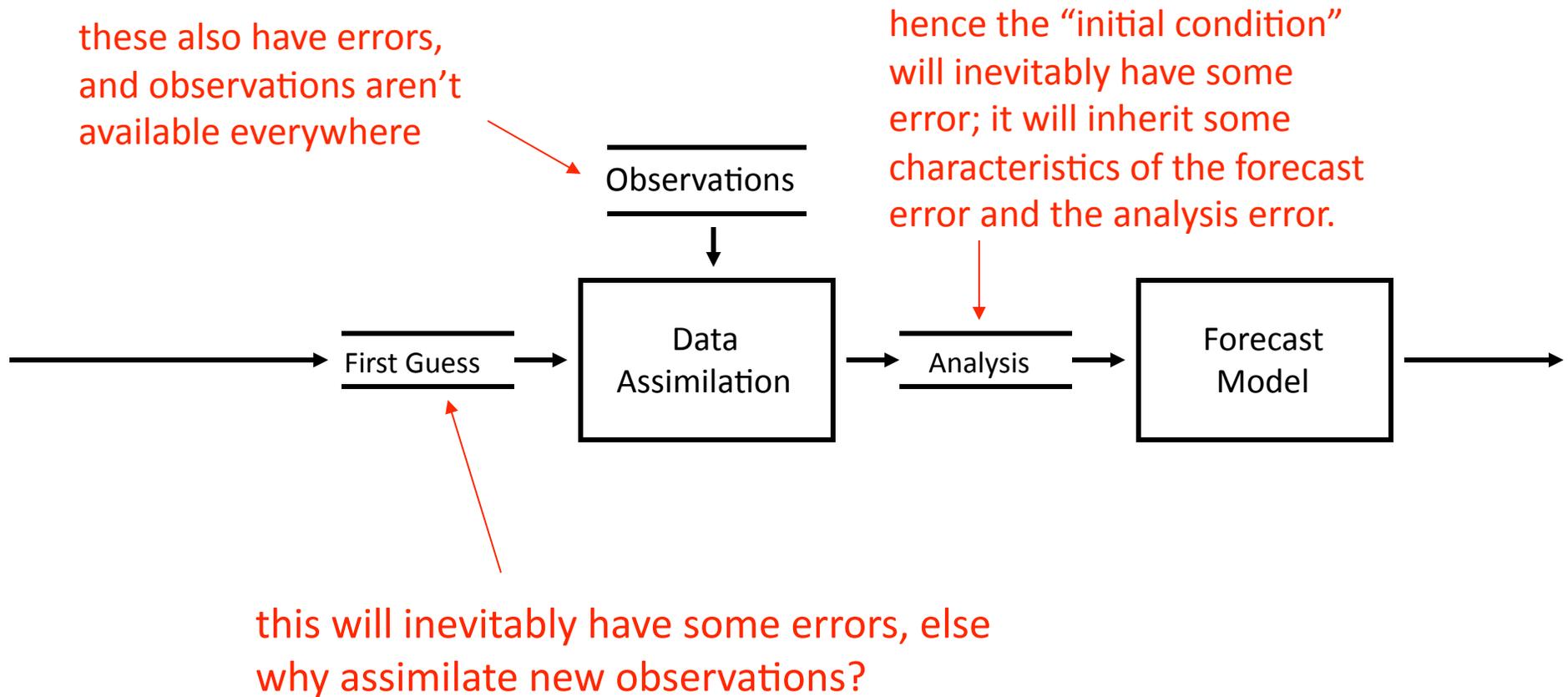
Why run forecasts from many initial conditions?

these also have errors,
and observations aren't
available everywhere

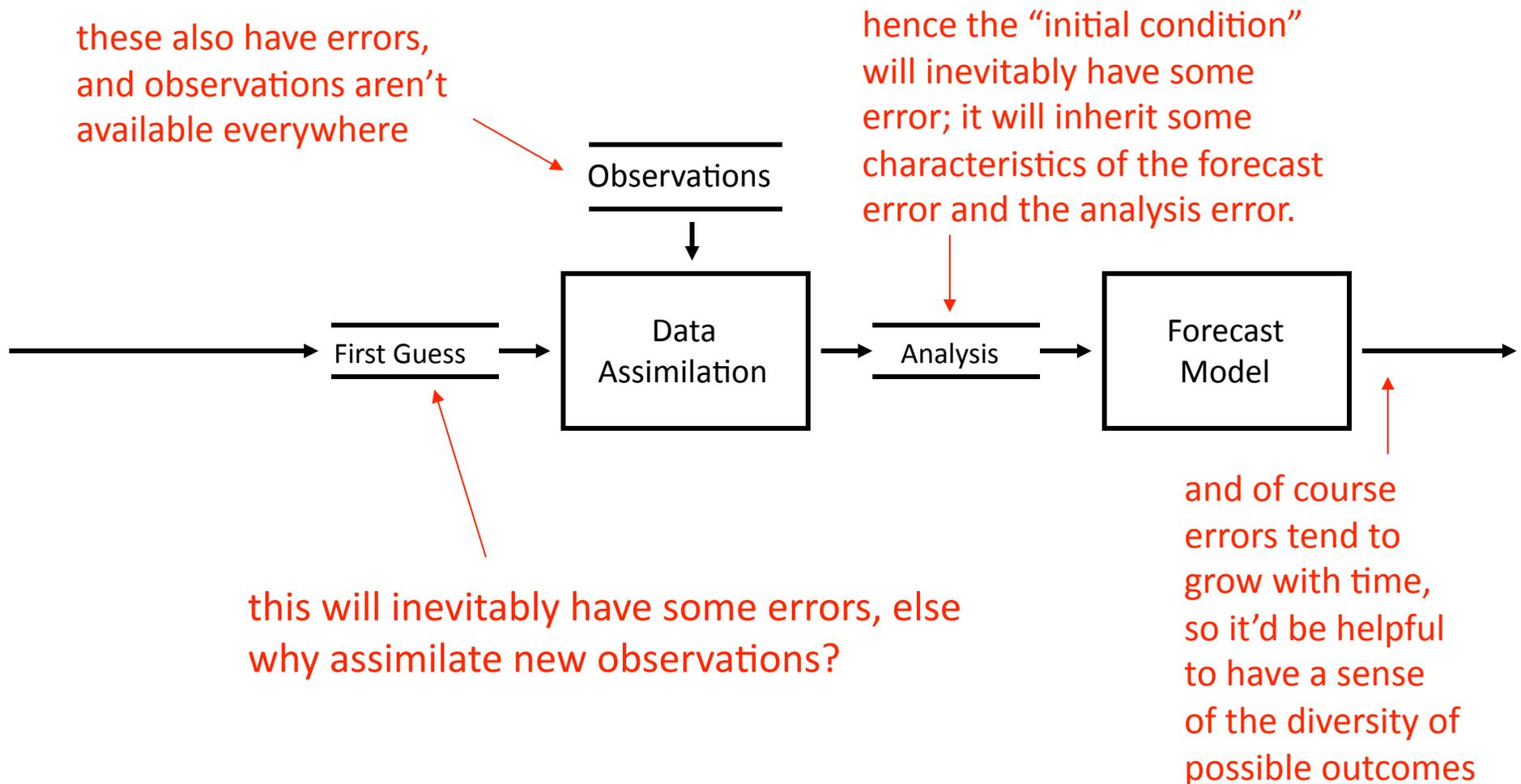


this will inevitably have some errors, else
why assimilate new observations?

Why run forecasts from many initial conditions?

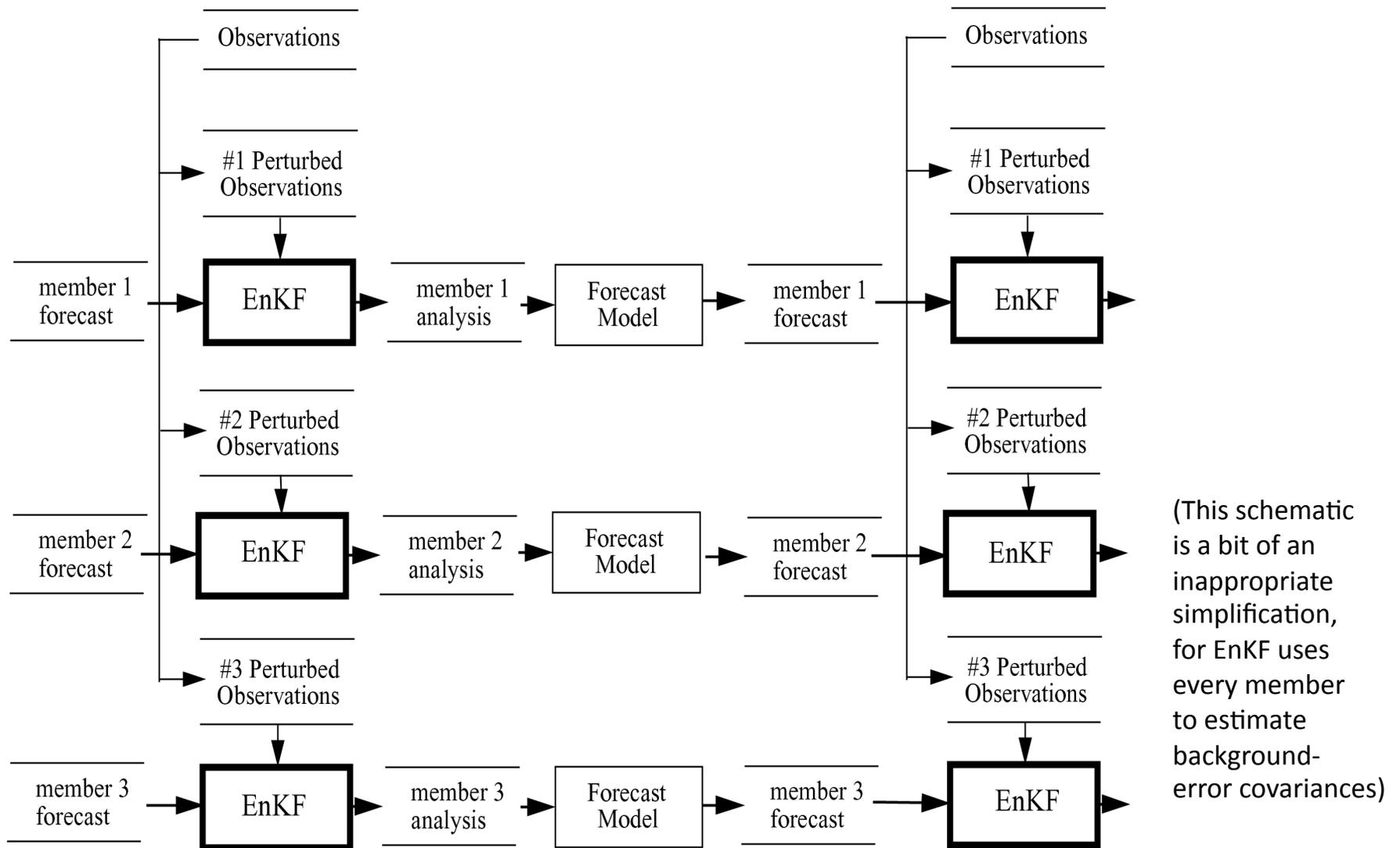


Why run forecasts from many initial conditions?



EnKF (Ensemble Kalman Filter)

naturally simulates uncertainty in observations, prior forecast



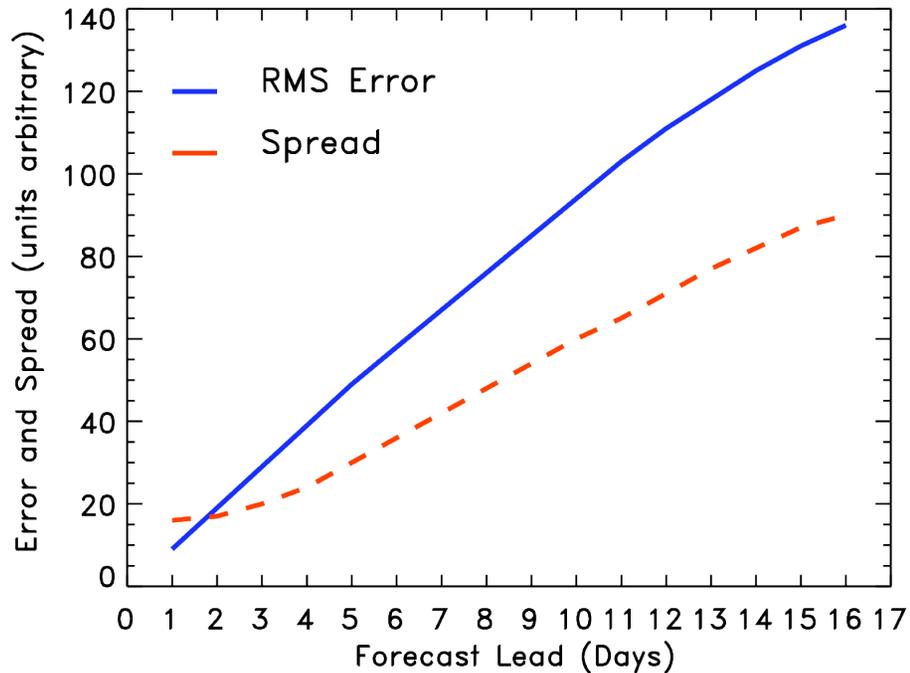
Desirable properties for ensembles of initial conditions

- (1) true model state and ensemble are random draws from the same distribution (same as before).
 - i.e., ensemble samples “analysis uncertainty.”
 - implies what you’d expect: larger differences between members in data voids, or where prior forecast differences were growing.

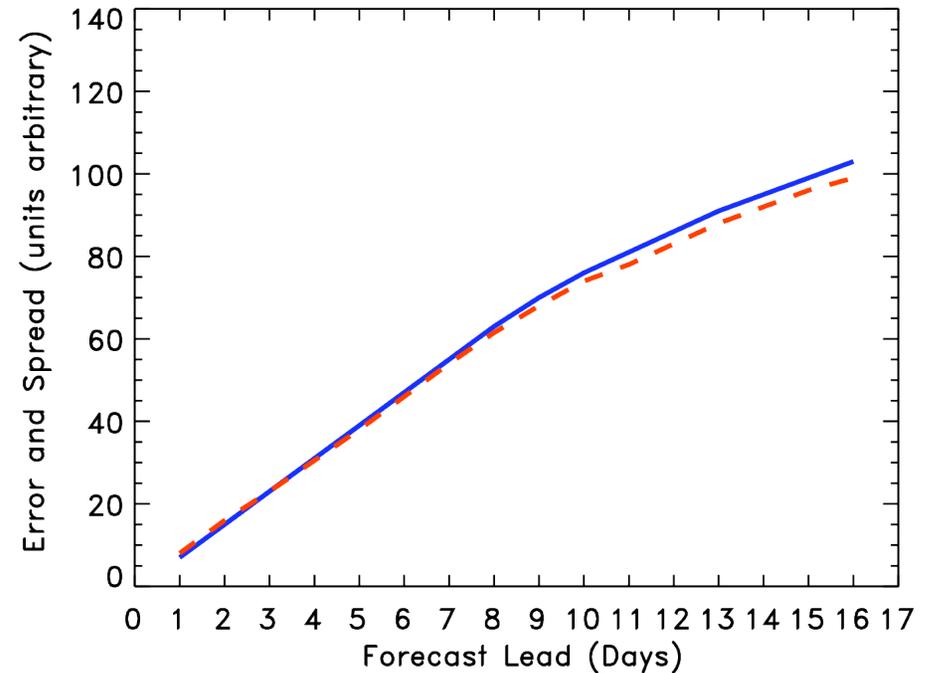
- (2) differences between subsequent forecasts ought to grow quickly enough that ensemble-spread consistent with ensemble mean error.
 - with perfect forecast model, (2) will happen naturally if you take care of (1)

Spread-error consistency

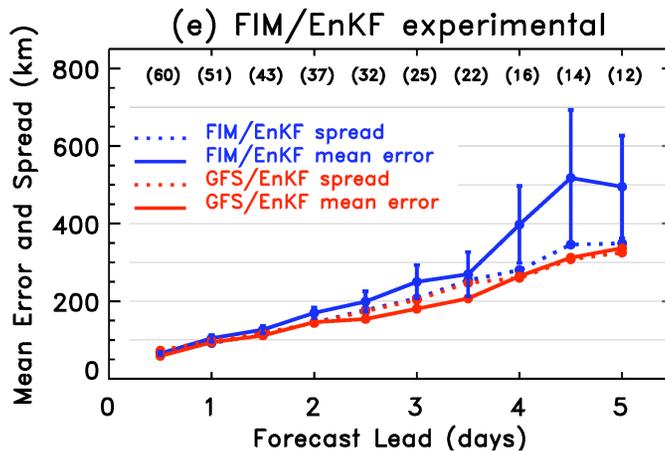
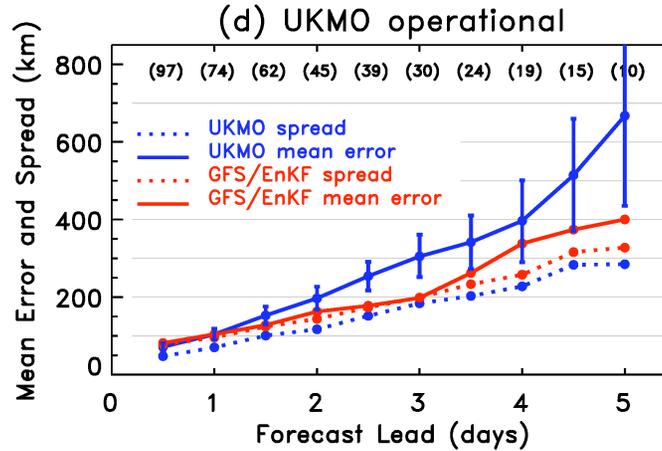
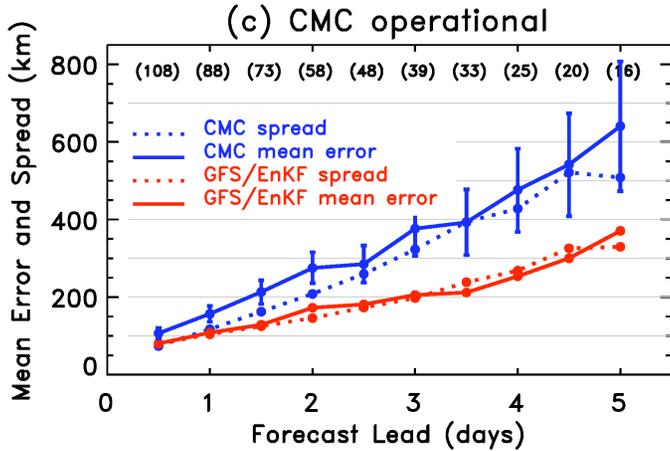
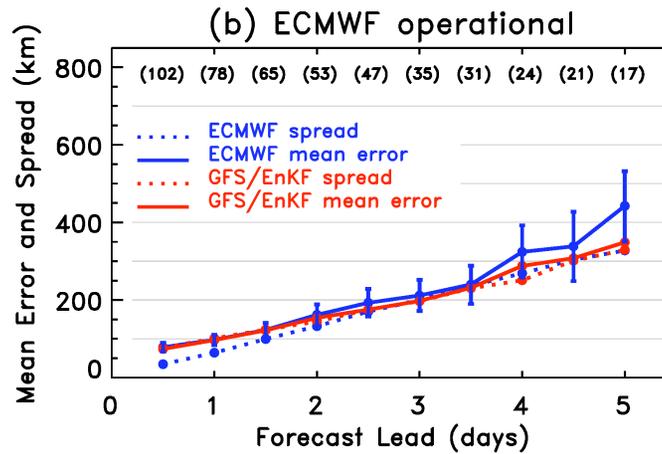
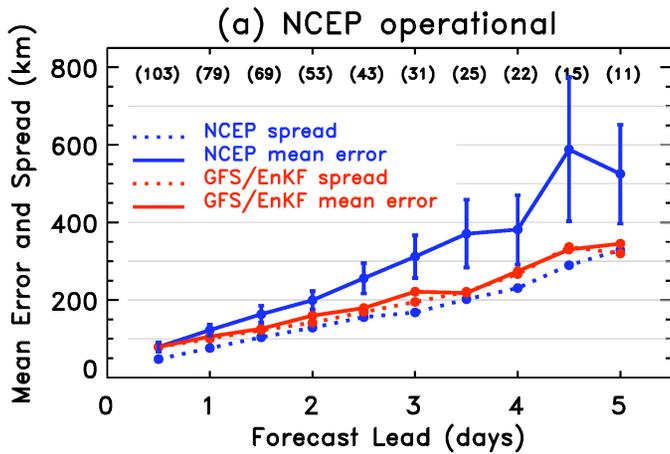
(a) Inconsistent: Large Error & Slower Spread Growth



(b) Consistent: Smaller Error & Faster Spread Growth



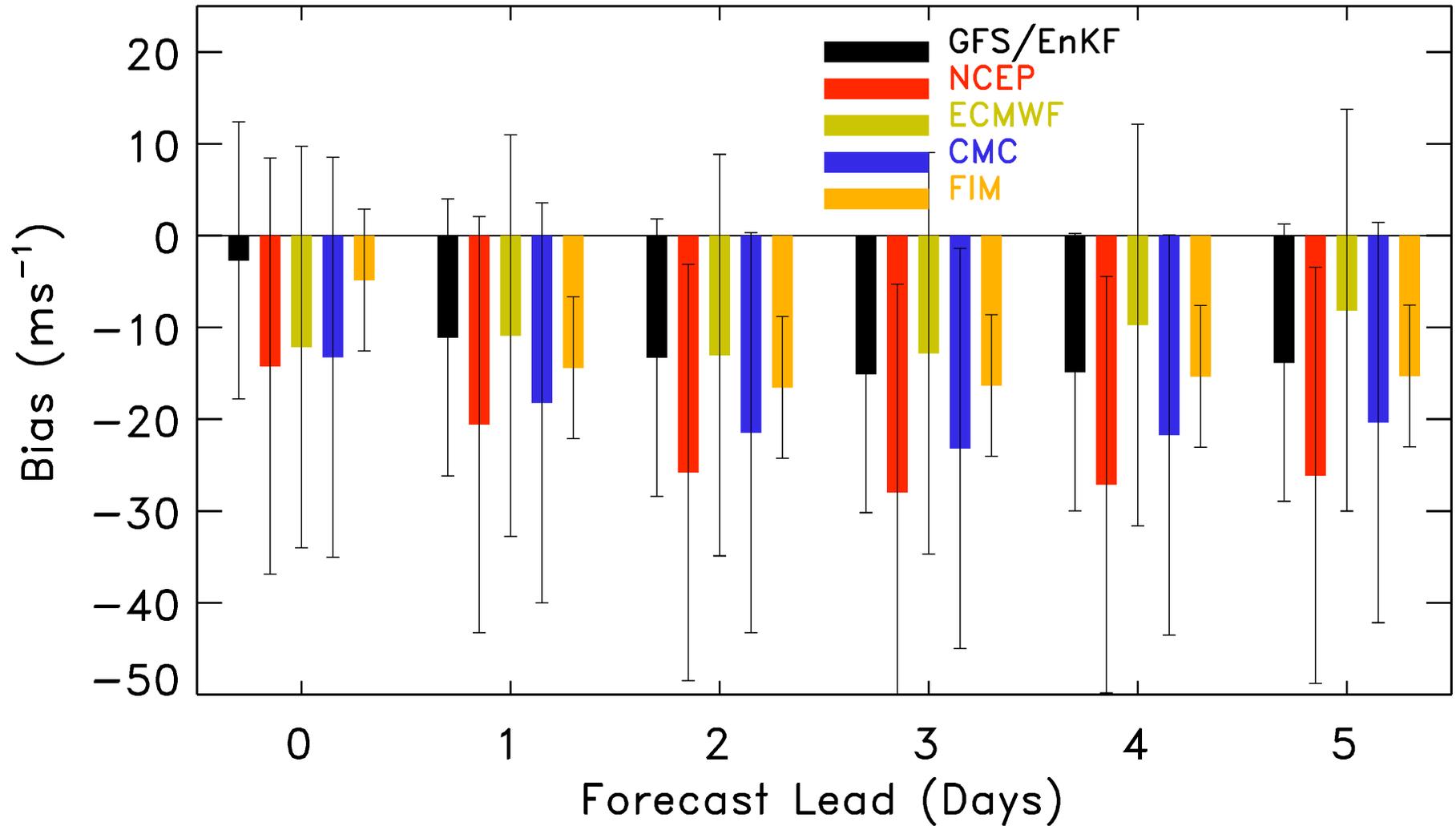
Spread should grow as quickly as error; part of spread growth from manner in which initial conditions are generated, some due to the model (e.g., stochastic physics, higher resolution increases spread growth). Focus on initial-condition aspect.



from recently submitted
Hamill et al. (2012) MWR
manuscript

For most models, spread
of hurricane tracks is
smaller than track
ensemble-mean error.

Maximum Wind Speed Bias



Global ensemble forecast models have systematic under-estimation of maximum wind speed. Lesson: we're far from conquering model error in NWP and ensembles.

“Model error”

- Imperfections in the forecast model, due to:
 - inadequate resolution
 - unduly simple physical parameterizations
 - deterministic may be inappropriate
 - coding bugs
 - lack of coupling, e.g., ocean-atmosphere
 - use of limited-area nested model
 - boundary-condition imperfections
 - one-way nesting of outer domain, lack of ability for resolved scales to interact with planetary scales
 - etc.

Treating model error

- Improve your model
- Incorporate stochastic parameterizations where appropriate
- Multi-parameterization
- Multi-model

(Carolyn Reynolds will review further)

Ensemble products

(what we're here to discuss)

- For the fields where we are starting to have some confidence in ensemble guidance, what can we do to convey that information in useful ways to the forecaster and to the public?

Initialized 00 UTC
5 August 2009.

* indicates observed best-track position.

Bi-variate normal distribution fit to ensemble member positions; contour encloses 90% of fitted probability.

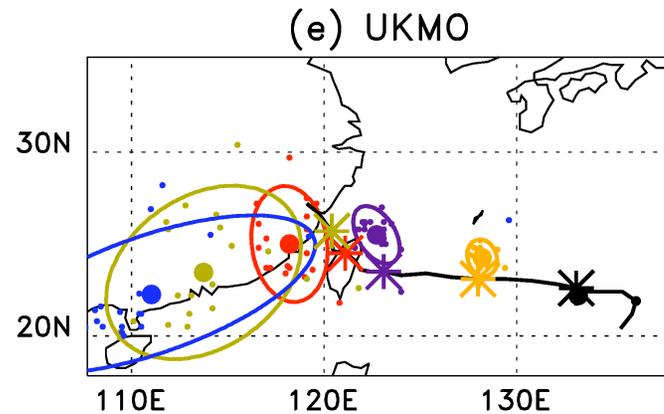
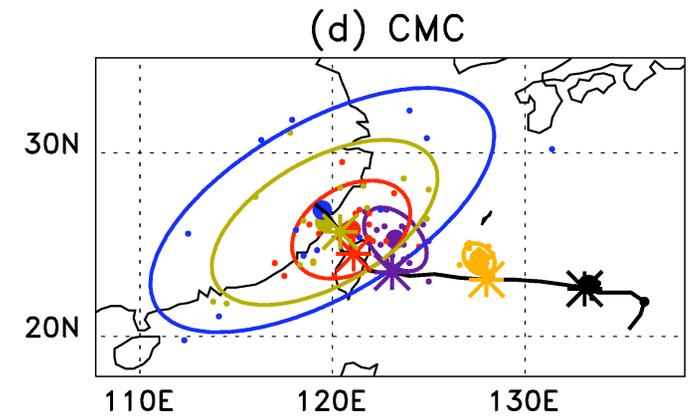
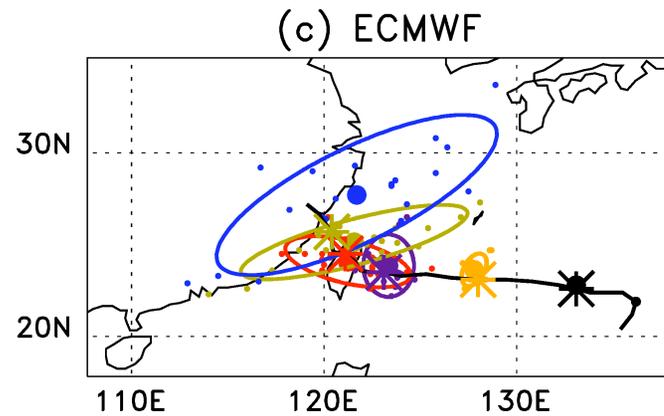
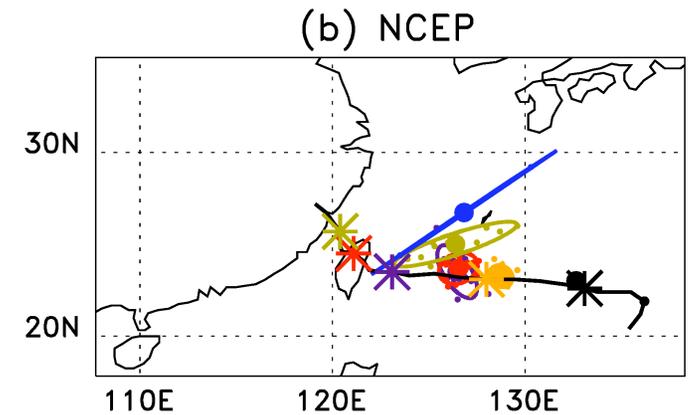
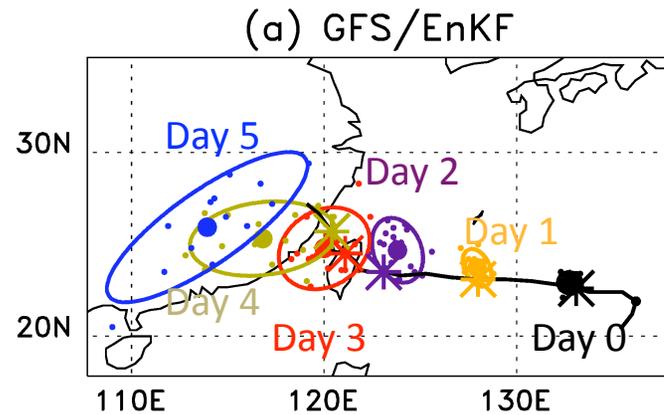
GEFS/EnKF a bit north and too fast.

NCEP has northward & westward bias, few members track.

ECMWF tracks decent up to Taiwan landfall

CMC has very large spread, esp. after landfall.

UKMO too north, too fast.

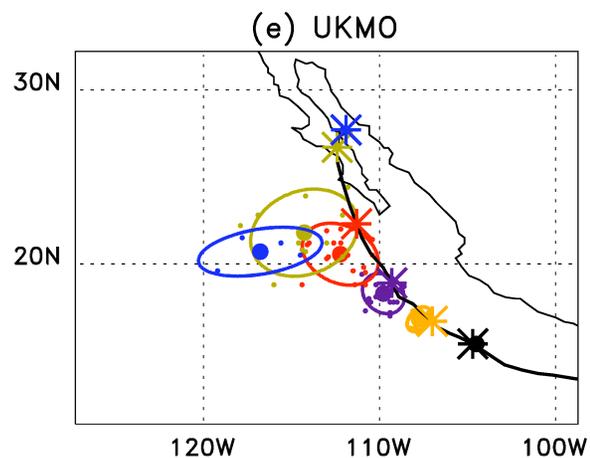
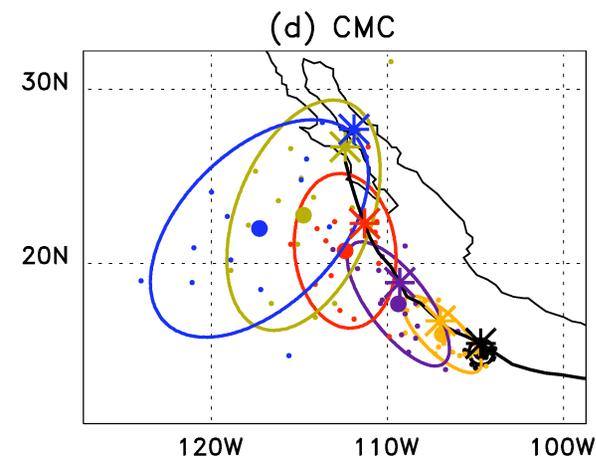
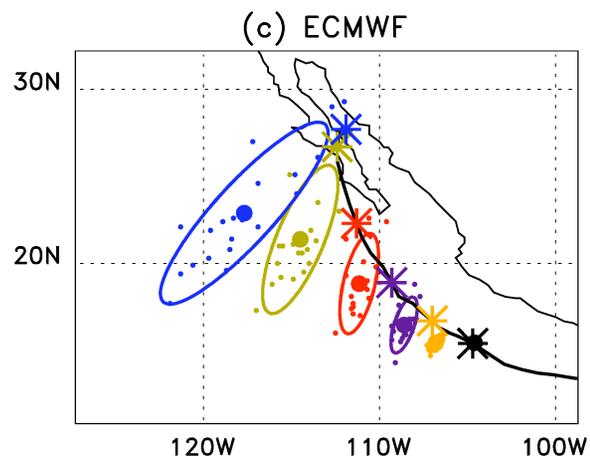
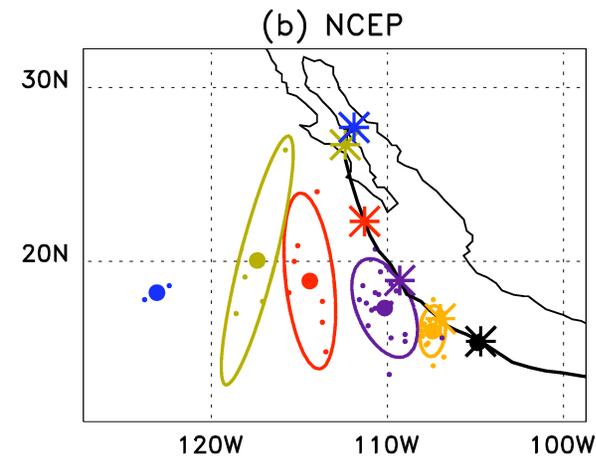
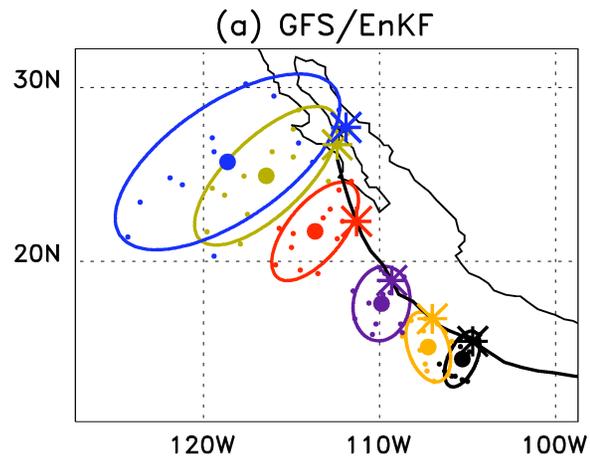


Example: Typhoon Morakot

Example: hurricane Jimena

Initialized 00 UTC 30 Aug 2009

all models have westward
bias; none of the forecasts
particularly good.



Example: Hurricane Bill

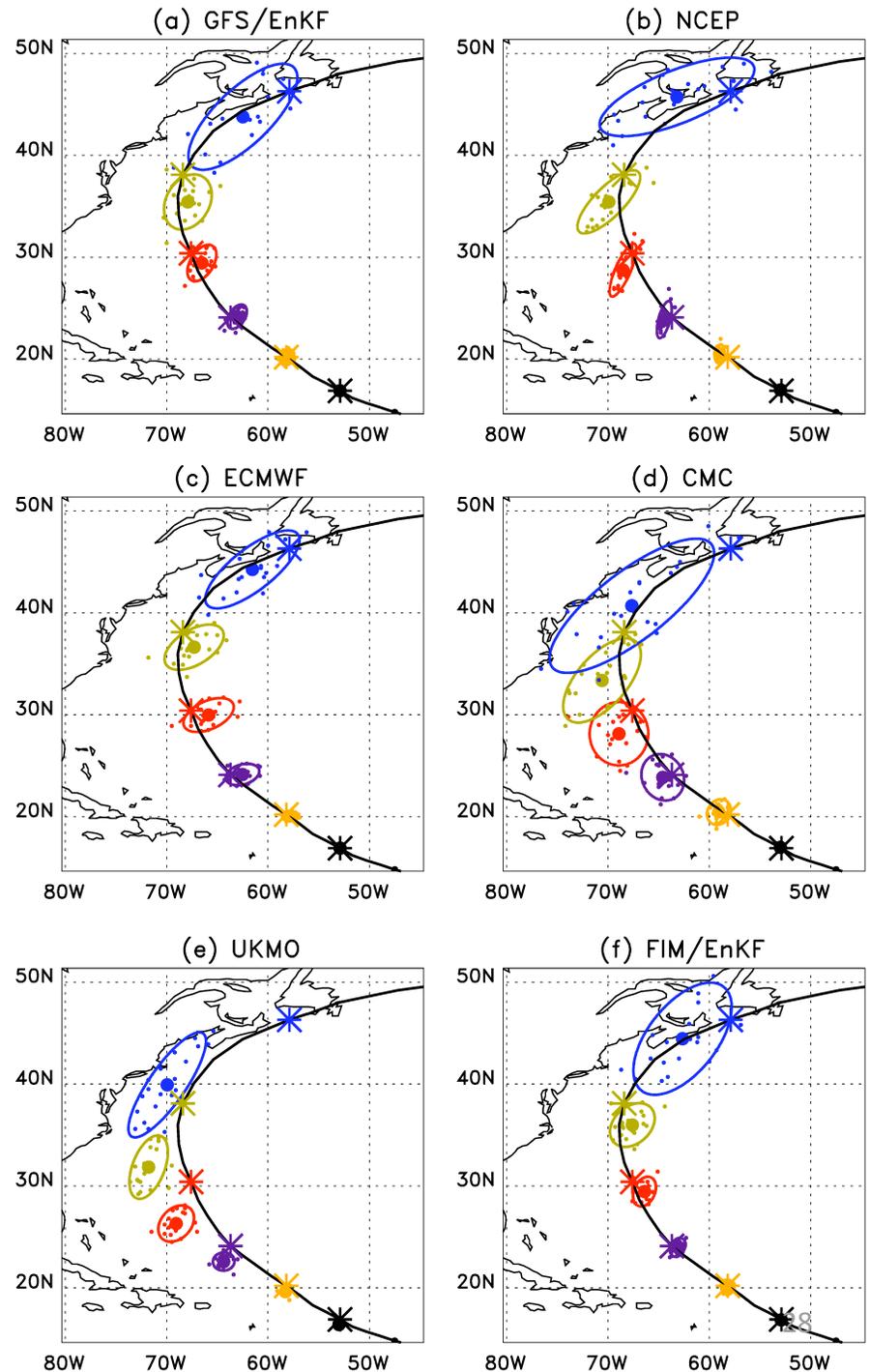
Initialized 00 UTC 19 August 2009.

All models slow, to varying extents.

GEFS/EnKF and ECMWF tracks decent.

UKMO, CMC have westward bias.

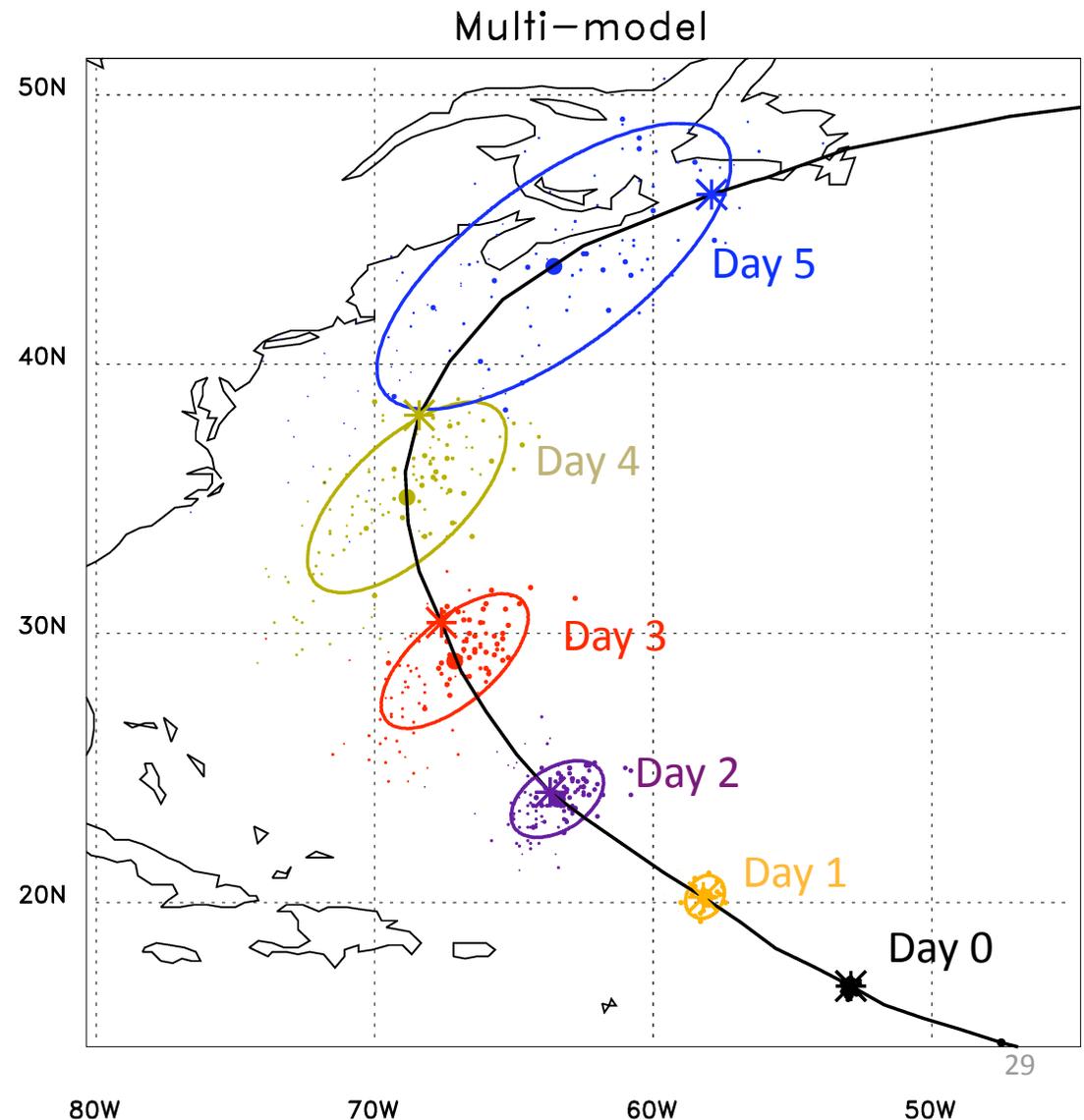
NCEP, FIM decent.



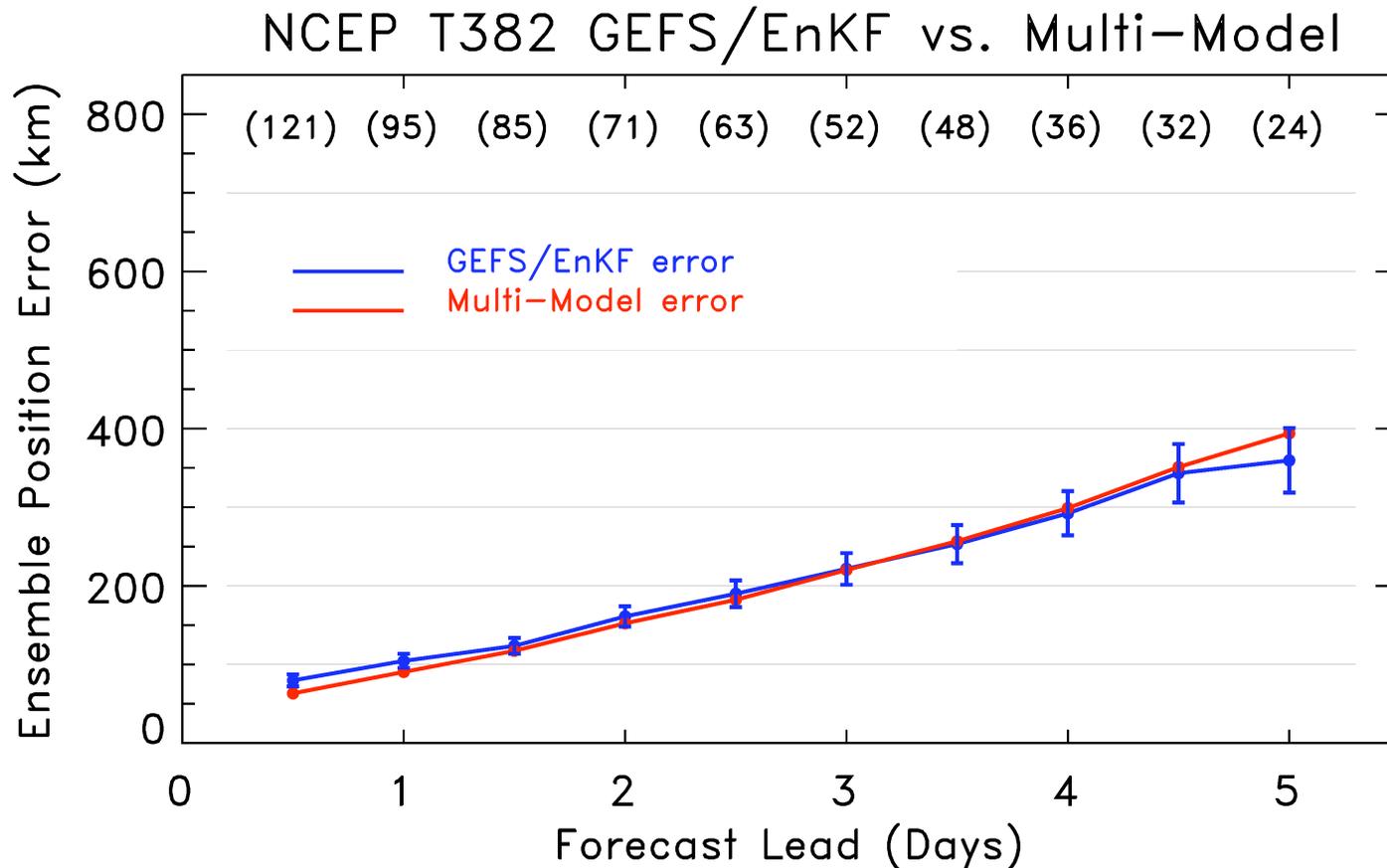
An experimental multi-model product

Dot area is proportional to the weighting applied to that member

- = ens. mean position
- * = observed position

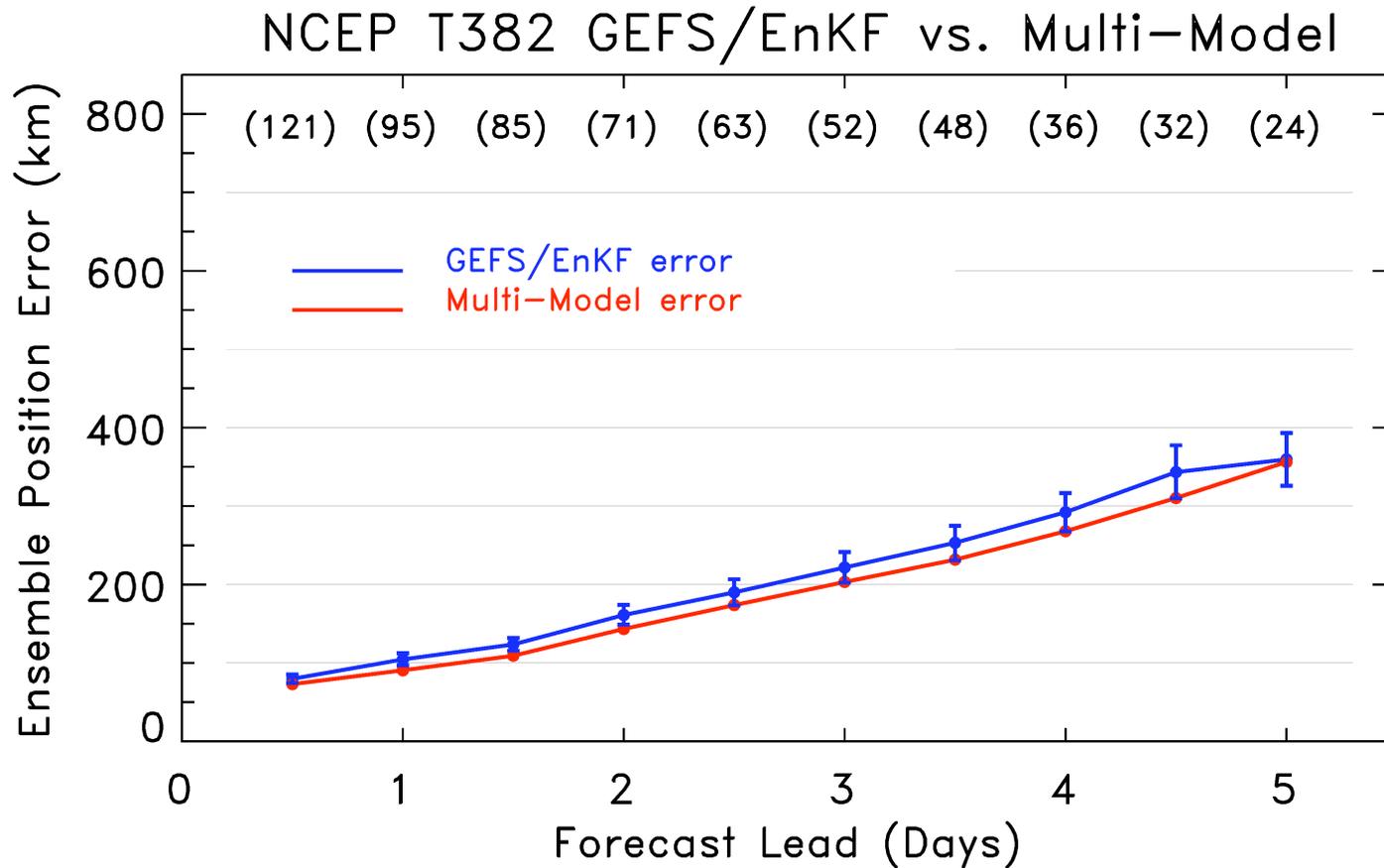


Multi-model error (GEFS/EnKF, ECMWF, FIM, UKMO, CMC, NCEP)



Not much improvement from multi-model. Why?

Multi-model error (GEFS/EnKF, ECMWF only)



Now some improvement, ~ 6 - 9 hours lead.

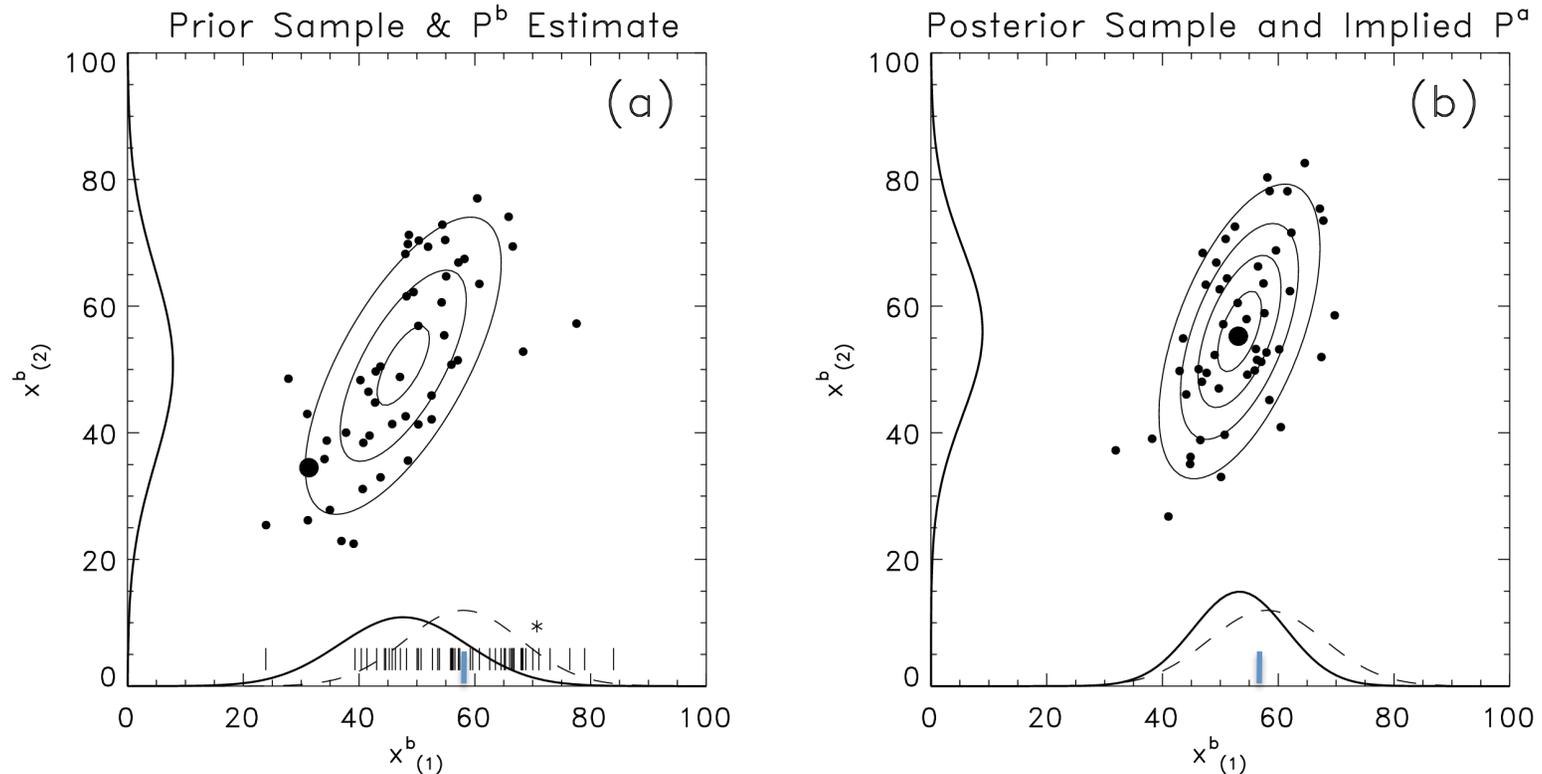
Questions

- Mine:
 - Are ellipses, colors useful way of conveying ensemble information?
 - Are products like the multi-model synthesis shown here potentially useful to forecasters?
- Yours

Acknowledgments

- NOAA HFIP program for 2009 summer test support
- NOAA THORPEX, ESRL base for EnKF development
- U. Texas NSF supercomputer for CPU cycles
- Phil Pegion for programming support, analysis.
- Mike Fiorino (tracker files, software sanity check)
- Other ESRLites, e.g., Chris Harrop for scripting, Stan Benjamin for FIM development and management, Paula McCaslin for display software
- Tim Marchok (tracker files)
- Daryl Kleist (GFS parallel runs)
- Chris Snyder, Jeff Anderson for ensemble bull sessions
- and you, for your constructive feedback

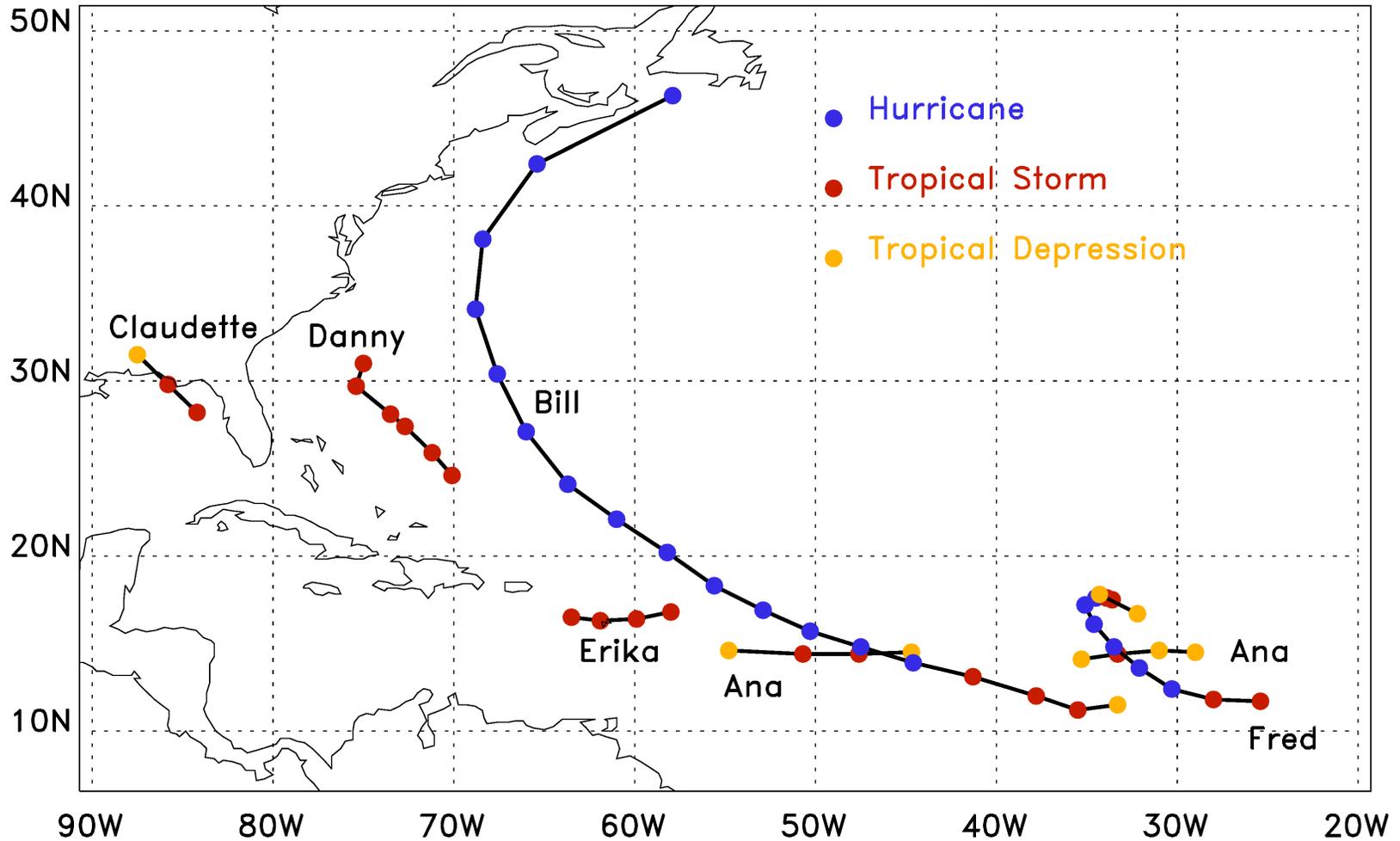
How the EnKF works: 2-D example



Start with a random sample from bimodal distribution used in previous Bayesian data assimilation example. Contours reflect the Gaussian distribution fitted to ensemble data.

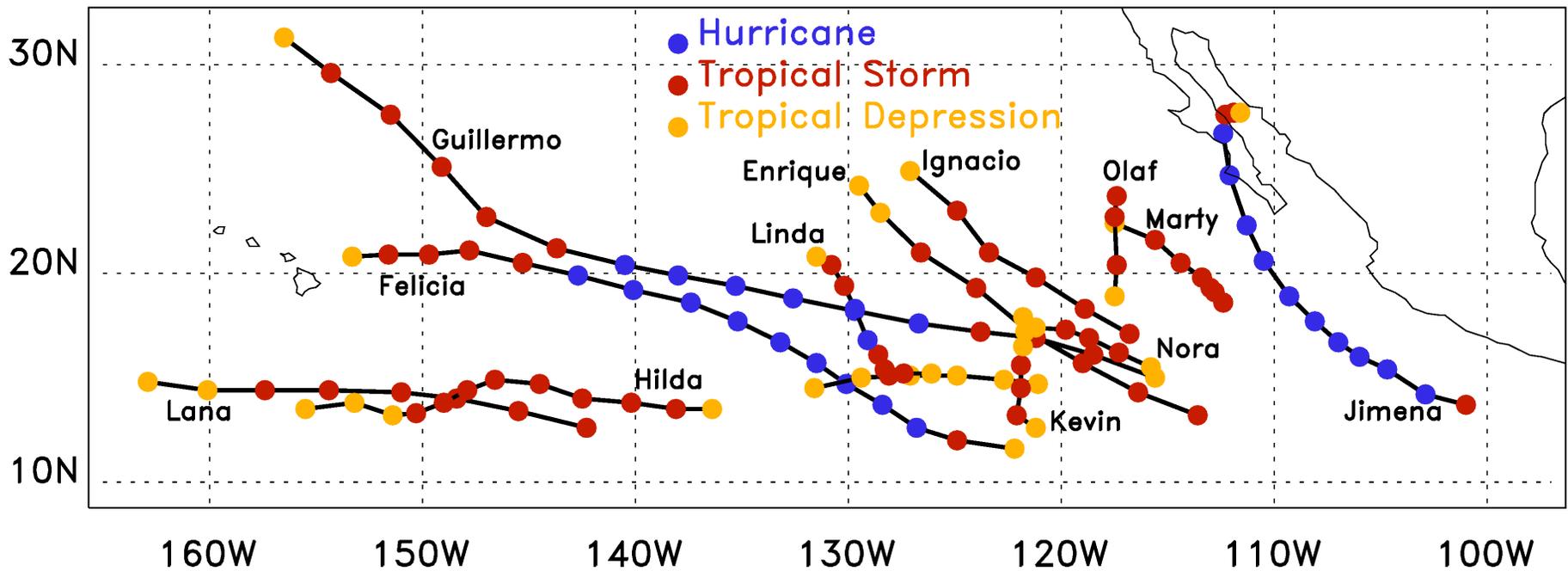
Review of Atlantic Basin activity

Atlantic Basin Storms, 31 Jul 2009 to 03 Oct 2009



Review of Eastern-Pacific activity

Eastern Pacific Storms, 31 Jul 2009 to 03 Oct 2009



Review of Western-Pacific activity

