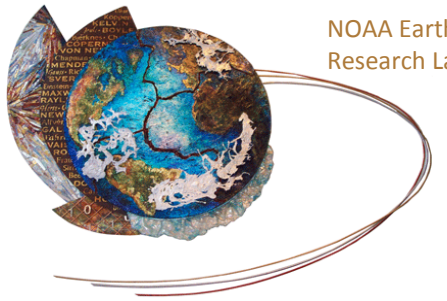


NOAA Earth System
Research Laboratory

Numerical, predictability, and observation type issues in regional hurricane modeling

Tom Hamill

NOAA Earth System Research Lab

tom.hamill@noaa.gov

Acknowledgments

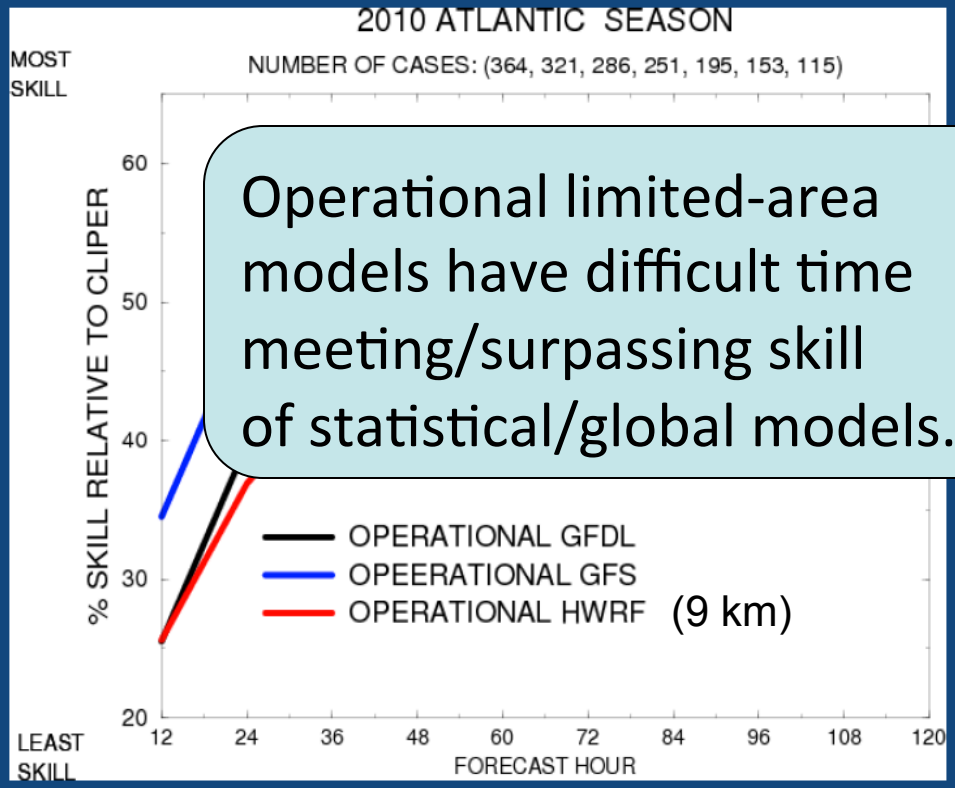
- All half-baked and overcooked ideas are, at P ~ 1.0, attributable to me.
- Properly baked ideas contributed by Jeff Whitaker; Henry Winterbottom; Tom Galarneau; Mike Fiorino; Stan Benjamin (all ESRL); Chris Davis; George Bryan; Rich Rotunno (NCAR); Greg Hakim (UW); and stolen from others, as cited.

Review: difficulties with intensity forecasting

some results from 2010 HFIP Annual Meeting

2010 Operational Verifications

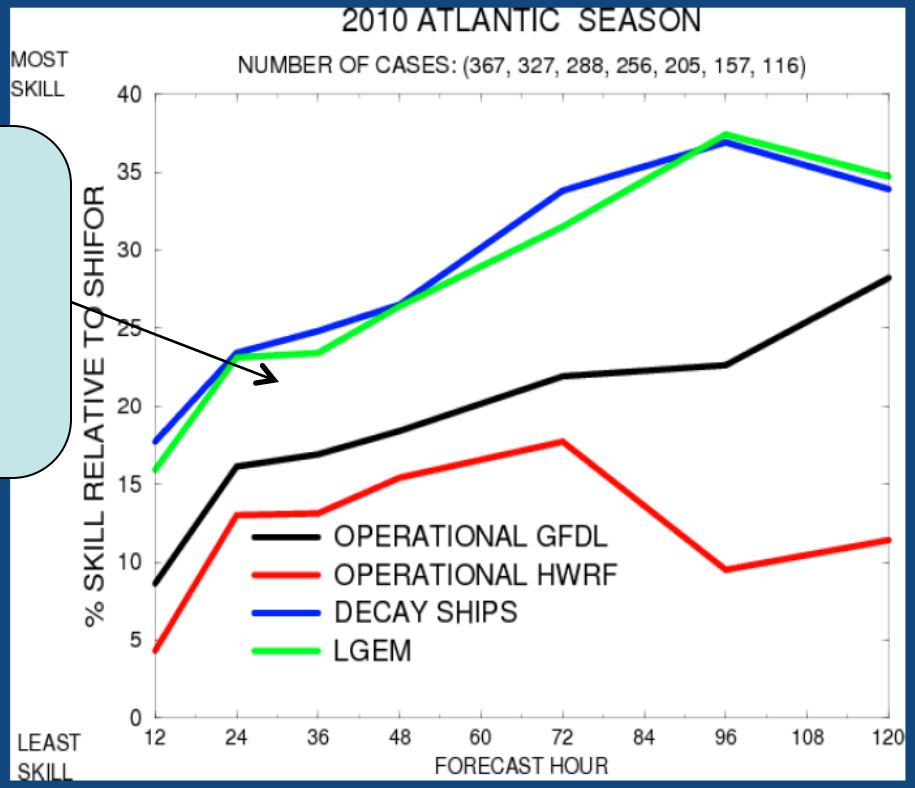
NORMALIZED TRACK ERROR



Operational limited-area models have difficult time meeting/surpassing skill of statistical/global models.

GFS global model showed more skill for Track at Every Time Level
(HWRF very poor performance in Tomas significantly affected its overall seasonal performance)

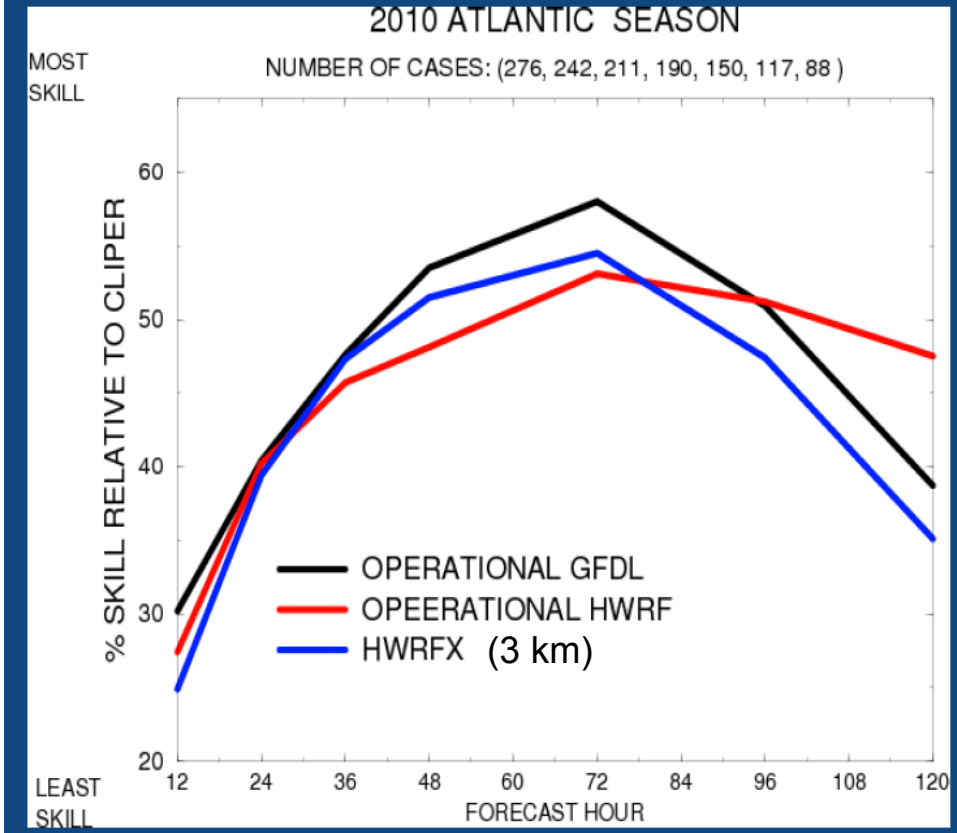
NORMALIZED INTENSITY ERROR



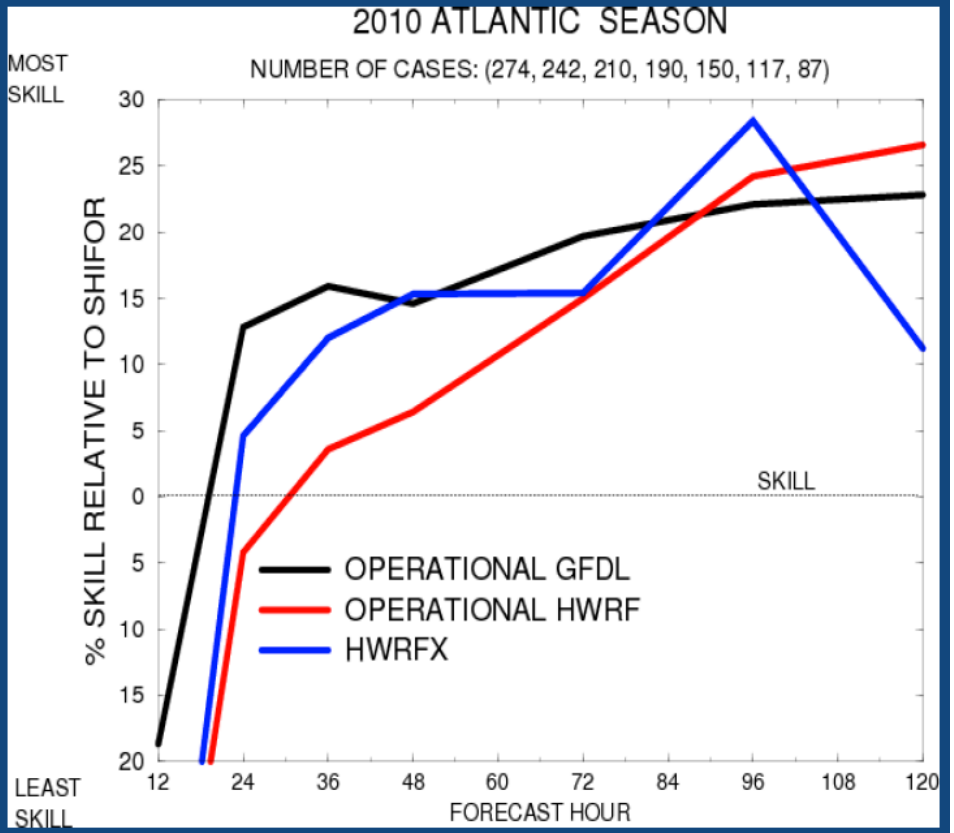
Operational Regional Models Still not as good as the statistical models for Intensity.
HWRF had significant negative bias

Late model Verification of HWRFx vs. Regional Operational Models

NORMALIZED TRACK ERROR



NORMALIZED INTENSITY ERROR

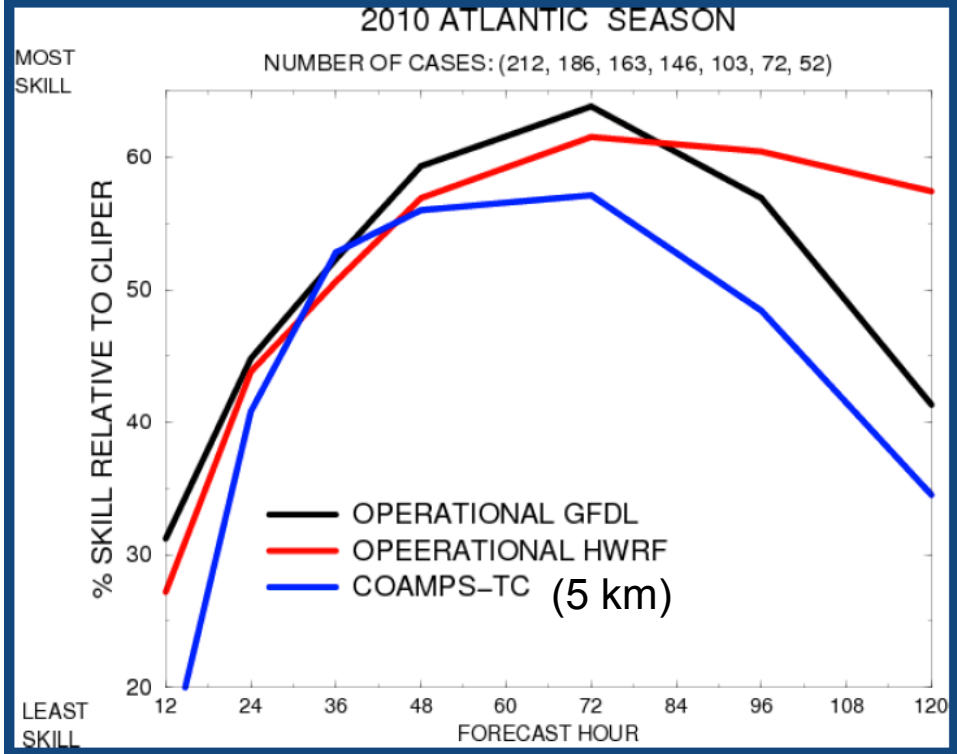


HWRFx Neutral or Inferior for Track

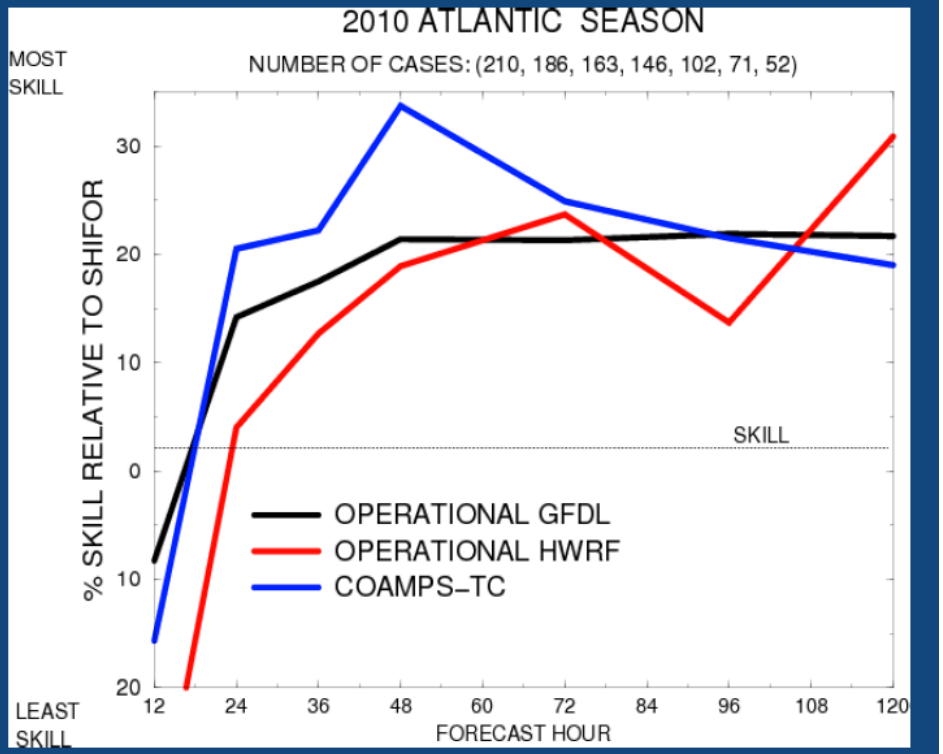
HWRFx Slightly Improved Intensity

Late model Verification of COAMPS-TC vs. Regional Operational Models

NORMALIZED TRACK ERROR



NORMALIZED INTENSITY ERROR

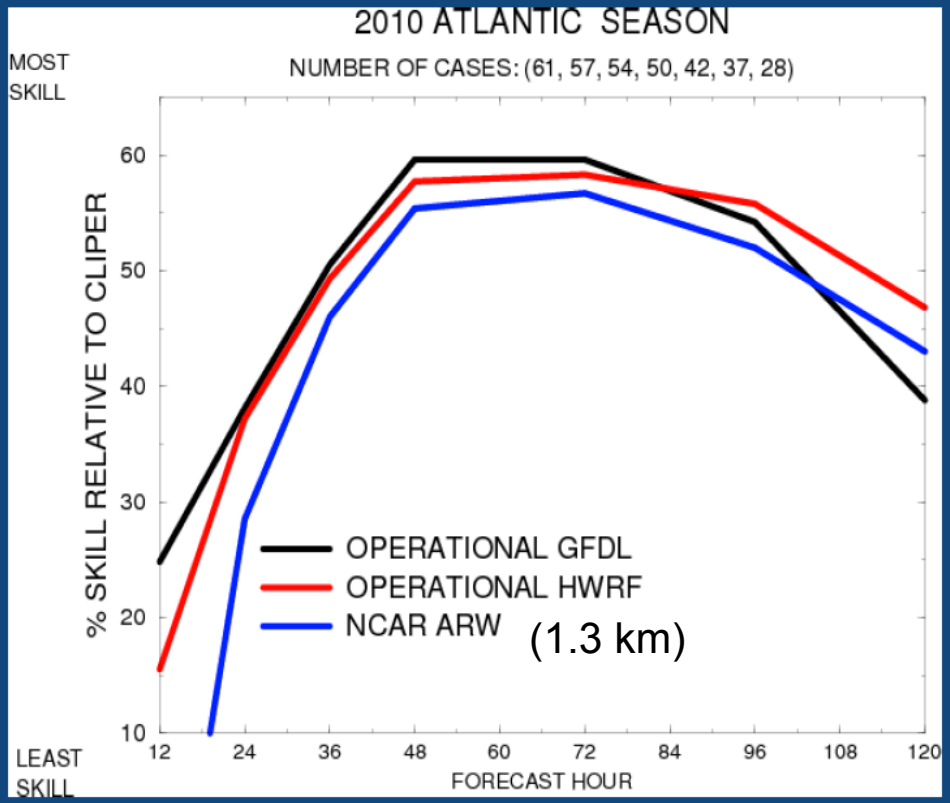


Most Skillful of the Regional Models for Intensity through 2 Days

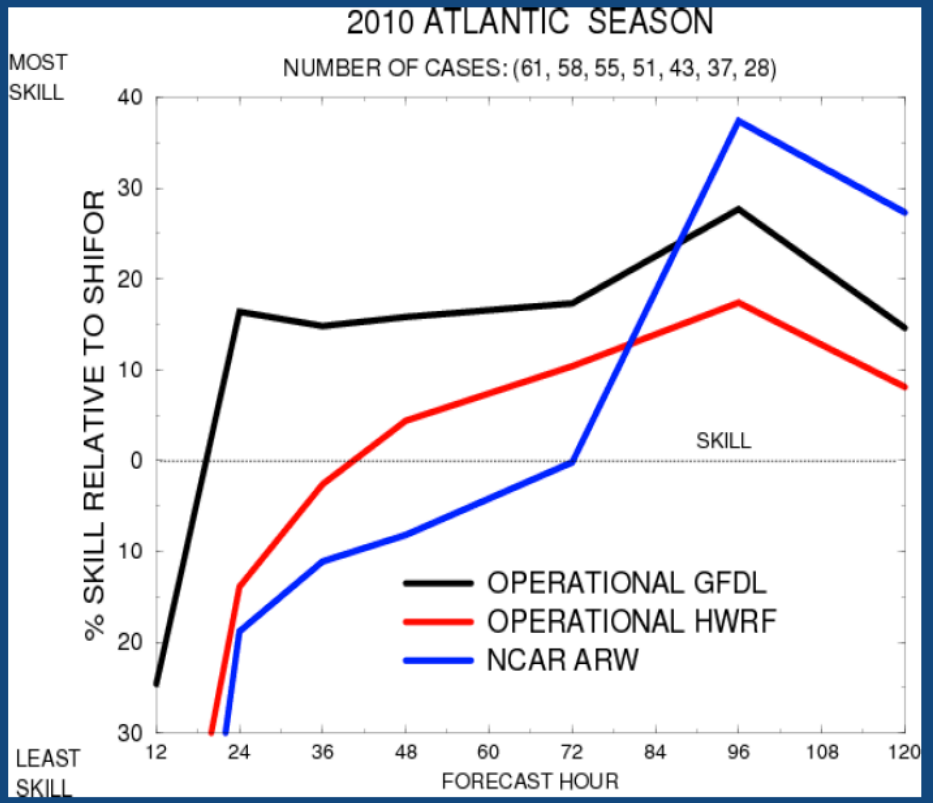
Late model Verification of NCAR ARW vs. Regional Operational Models

(very limited data set for meaningful comparison)

NORMALIZED TRACK ERROR



NORMALIZED INTENSITY ERROR



With transition of HWRFx to HWRFV3.2 System HFIP Regional Model Development is at a Crucial Crosswords to achieve its goals within 5 years to significantly improve intensity skill.

There must be a careful and focused effort over next 3-4 years to develop physics that will correctly represented hurricane inner core physics processes that can be adequately resolved at high resolutions.

Still need to devote resources to find out what is the minimum horizontal resolution needed to adequately resolve the hurricane inner core (1km, 2km, 3km ??)

Collaboration between the academic community and NOAA will be essential to help achieve needed physics improvements.

Closer collaboration is also needed between observational and model development teams to improve physics (e.g., surface, micro-physics)

Perceived HFIP foci in regional modeling and data assimilation

- highest-possible resolution over core
- inner-core physics
- inner-core radar data assimilation (w. EnKF)

No argument that these are worth exploring, but **what else** may also be limiting the improvement of intensity forecasts? That's the focus of this talk.

Two other issues that may be affecting our ability to predict intensity

(1) Inherent limitations of regional modeling systems, e.g., lateral boundary conditions, initialization from global models.

(2) Predictability & observation type.

(a) How do chaos and model uncertainty limit our ability to predict intensity? To what extent does predictability depend on where the errors are, in environment, in inner core?

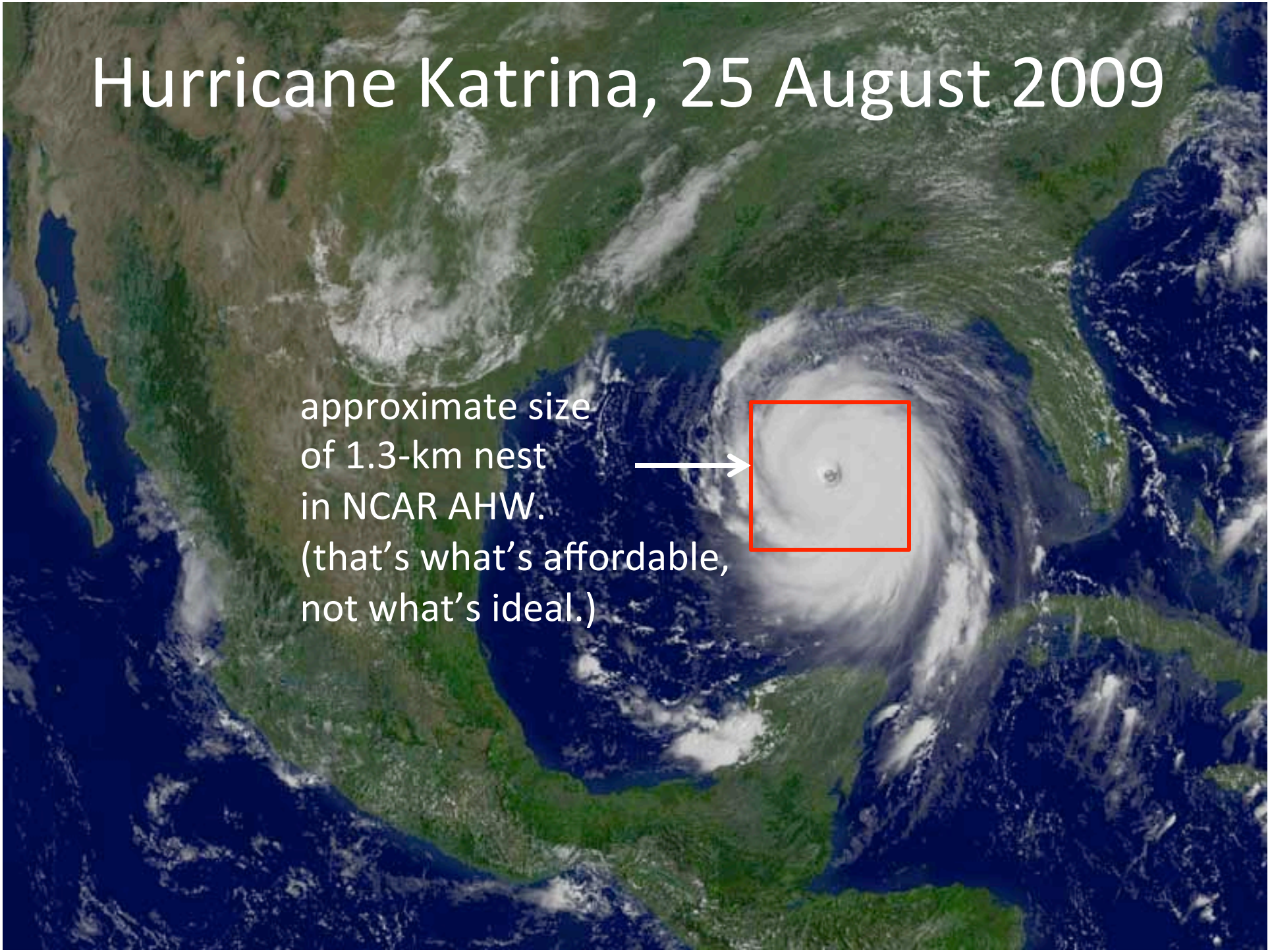
(b) Are we gathering the most crucial observations to improve intensity predictions?

ISSUE 1
Limited-area models and lateral boundary condition issues

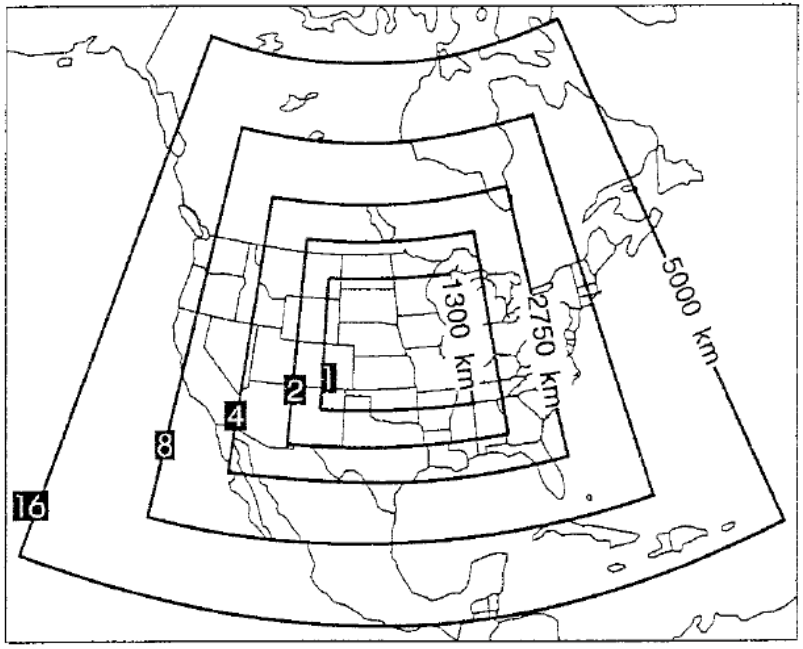
- LBCs generally provided by coarser-resolution forecast models, and this "sweeps" in low-resolution information, sweeps out developing high-resolution information.
- Physical process parameterizations for model driving LBCs may be different than for interior. Can affect veracity of interior grid calculations.
- With 1-way LBCs, resolvable scales for the limited-area domain cannot interact with scales larger than some limit defined by domain size.
- LBC info may introduce erroneous information for other reasons, e.g., wave reflection and refraction at grid interfaces.
- LBC initialization can produce transient gravity-inertia modes.

Ref: Warner et al. review article, *BAMS*, November 1997

Hurricane Katrina, 25 August 2009



Influence of domain size



An old study, but principles still relevant.

Here, a T-126 global model (~140-km grid spacing) provided lateral boundary conditions for nests with 40-km grid spacing of limited-area model.

from Warner et al. Nov 1997 *BAMS*, and Treadon and Peterson (1993), *Preprints, 13th Conf. on Weather Analysis and Forecasting*

Influence of domain size, continued

(a) **large nested domain**

(b) **small nested domain**

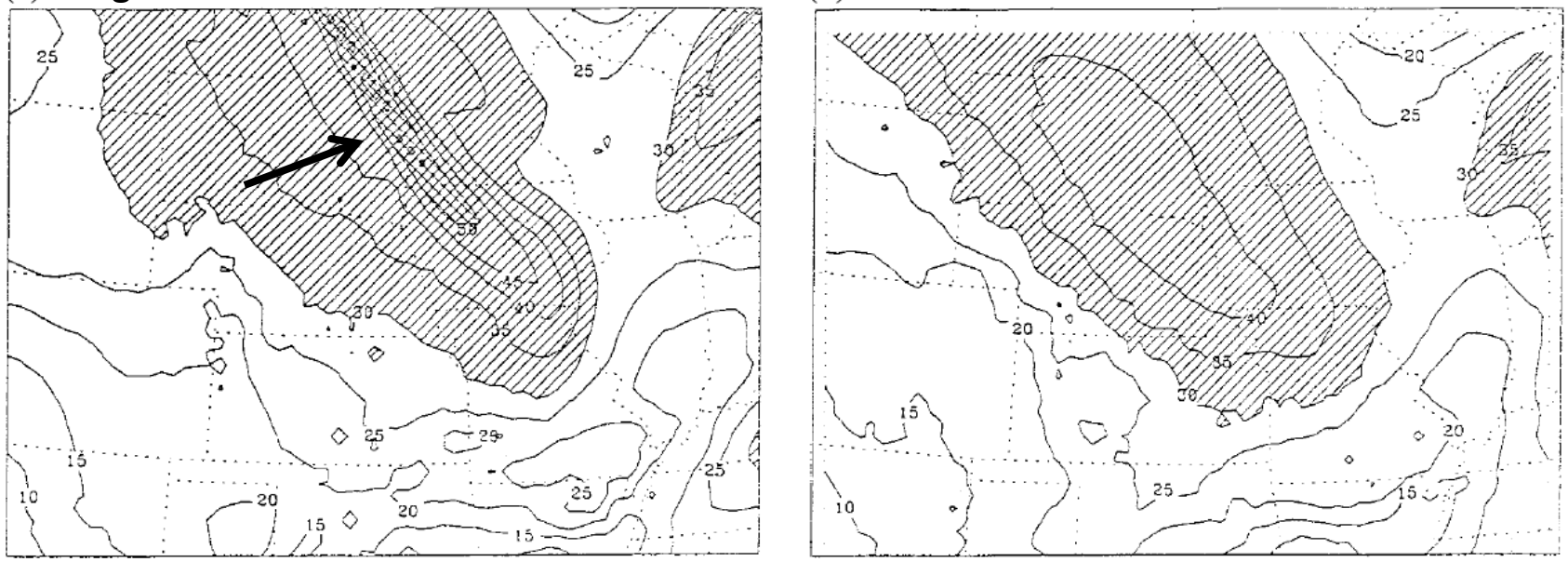
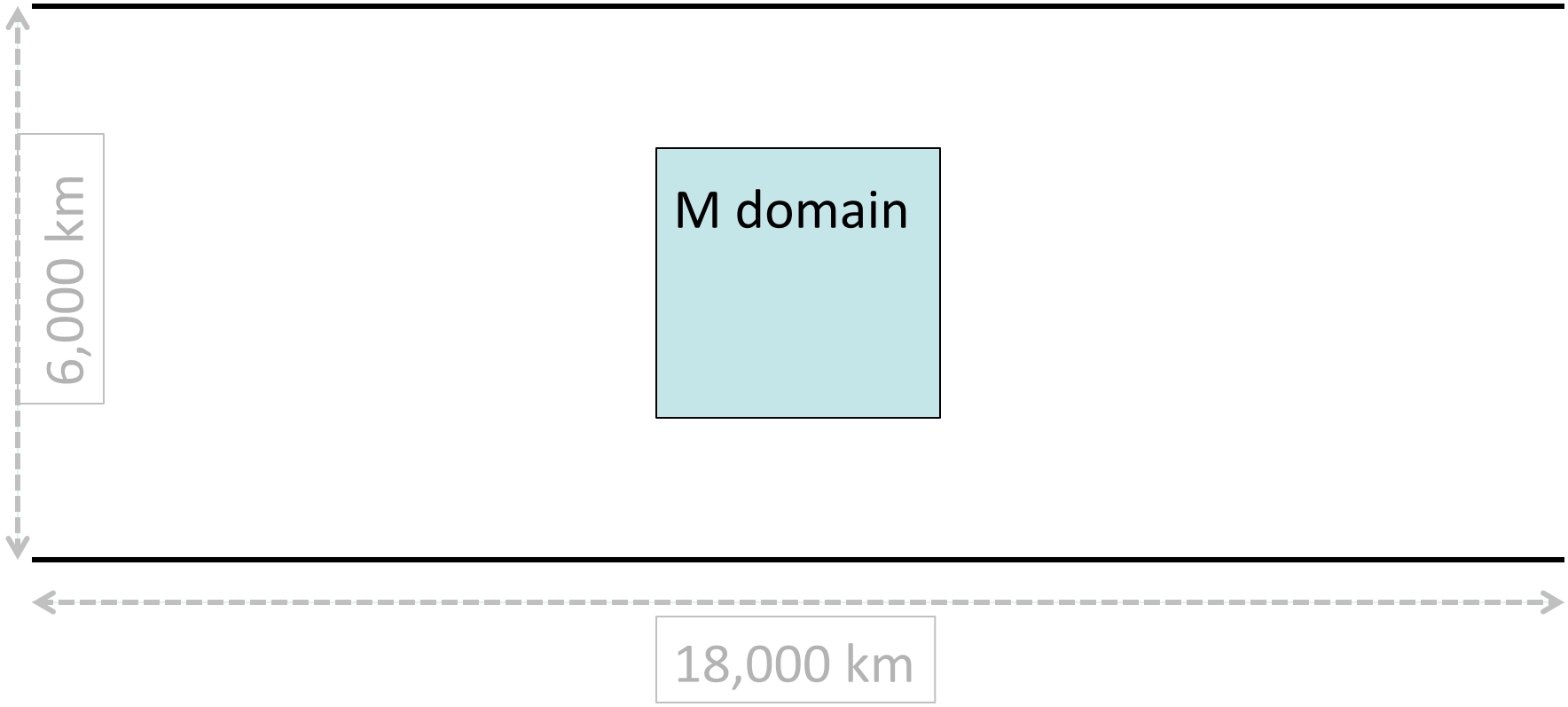


FIG. 6. Simulated 250-hPa isotachs (m s^{-1}) from the 40-km grid increment Eta Model initialized at 1200 UTC 3 August 1992 for the largest computational domain (a) and the smallest (b). The isotach interval is 5 m s^{-1}. From Treadon and Peterson (1993).

40-km nested domain in global model had thin, realistic jet streak using large domain (left) and smeared-out, unrealistic jet streak using small domain (right). High resolution of interior domain not useful here because of "sweeping" in of low-resolution information.

Ref: ibid

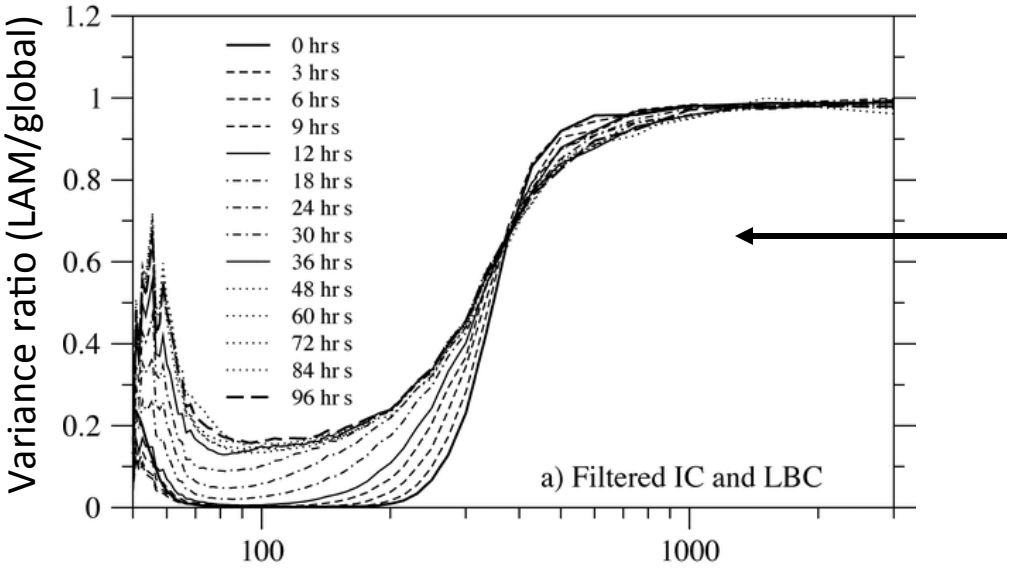
“Sweeping” : Paul Nutter et al.’s experiments



25-km grid spacing in channel, barotropic model, then smoothed to remove scales below 150 km, which supplies 1-way LBCs to M domain. What are effects on variance spectra in box (“M domain”) with 25-km grid spacing?

ref: Nutter et al. MWR, Oct 2004.

Nutter et al.'s experiments, cont'd.

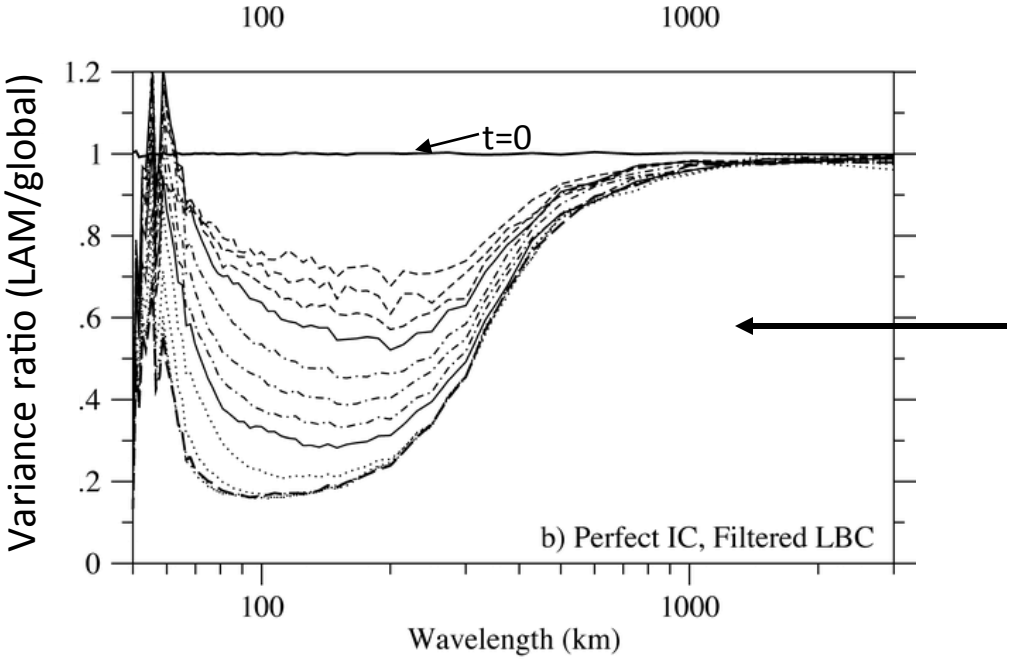


Simulating the effects of initializing high-resolution interior domain with coarse resolution analysis.

Here, as a baseline, the model forecasts throughout the full channel are cycled for a while at high resolution. The variance spectra in the M domain is calculated ("global"). The simulation is then repeated, but initial and LBC information provided to the M domain is filtered to remove scales below 150 km, simulating initialization with a coarse-resolution analysis and coarse-resolution information from LBCs. Variance spectra in M domain is recalculated ("LAM"). Ratio of the filtered/unfiltered in LAM is plotted.

The small grid spacing on the interior is useless at first, inheriting global ensembles without small scales. Even after a long time, there is not much variance at the small scales that develops due to the higher interior resolution. The extra resolution is largely wasted.

Nutter et al.'s experiments, cont'd.

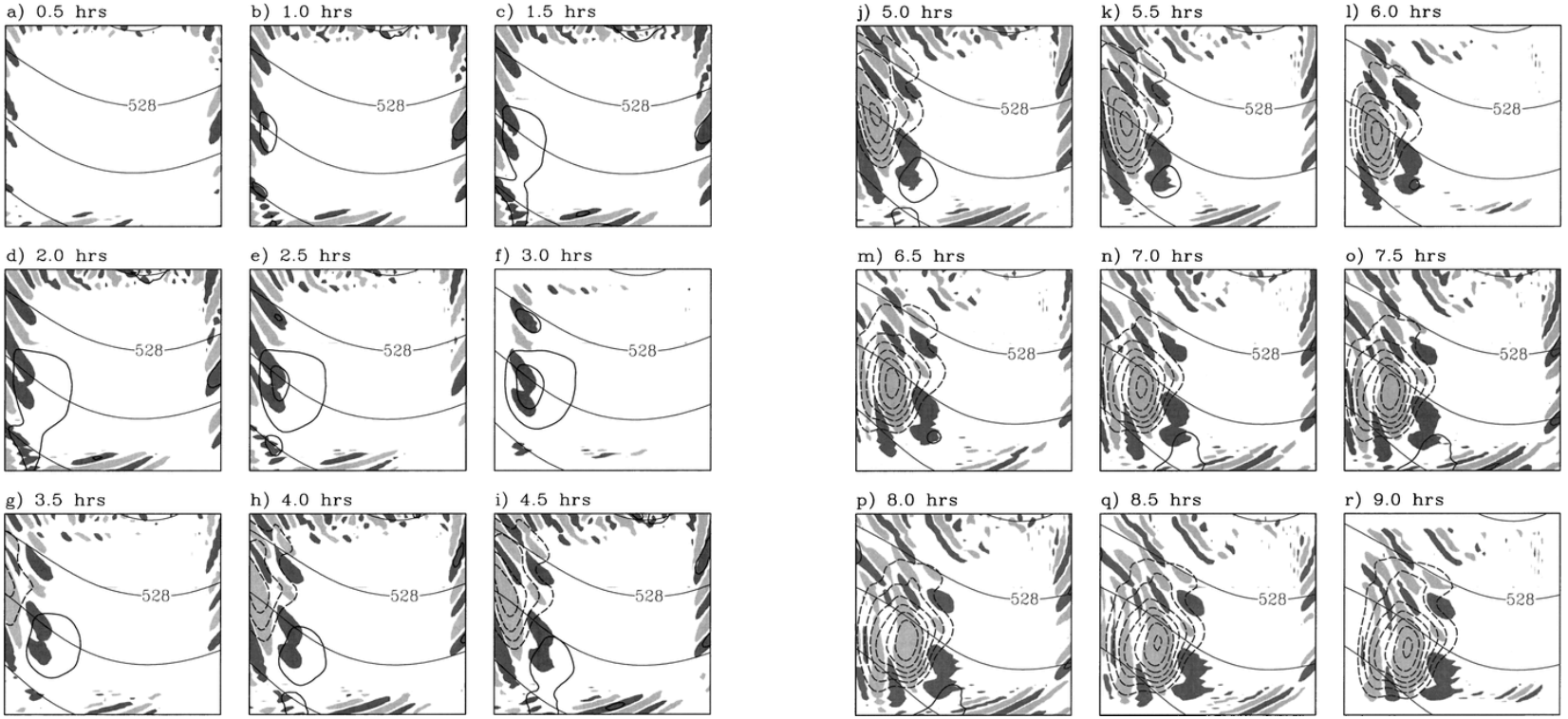


Here the initial condition does initially contain all scales of motion, but M domain thereafter receives filtered lateral boundary conditions.

Even with a quality initialization, the small scales are "swept" away with time by the lower-resolution information coming in from the LBC's.

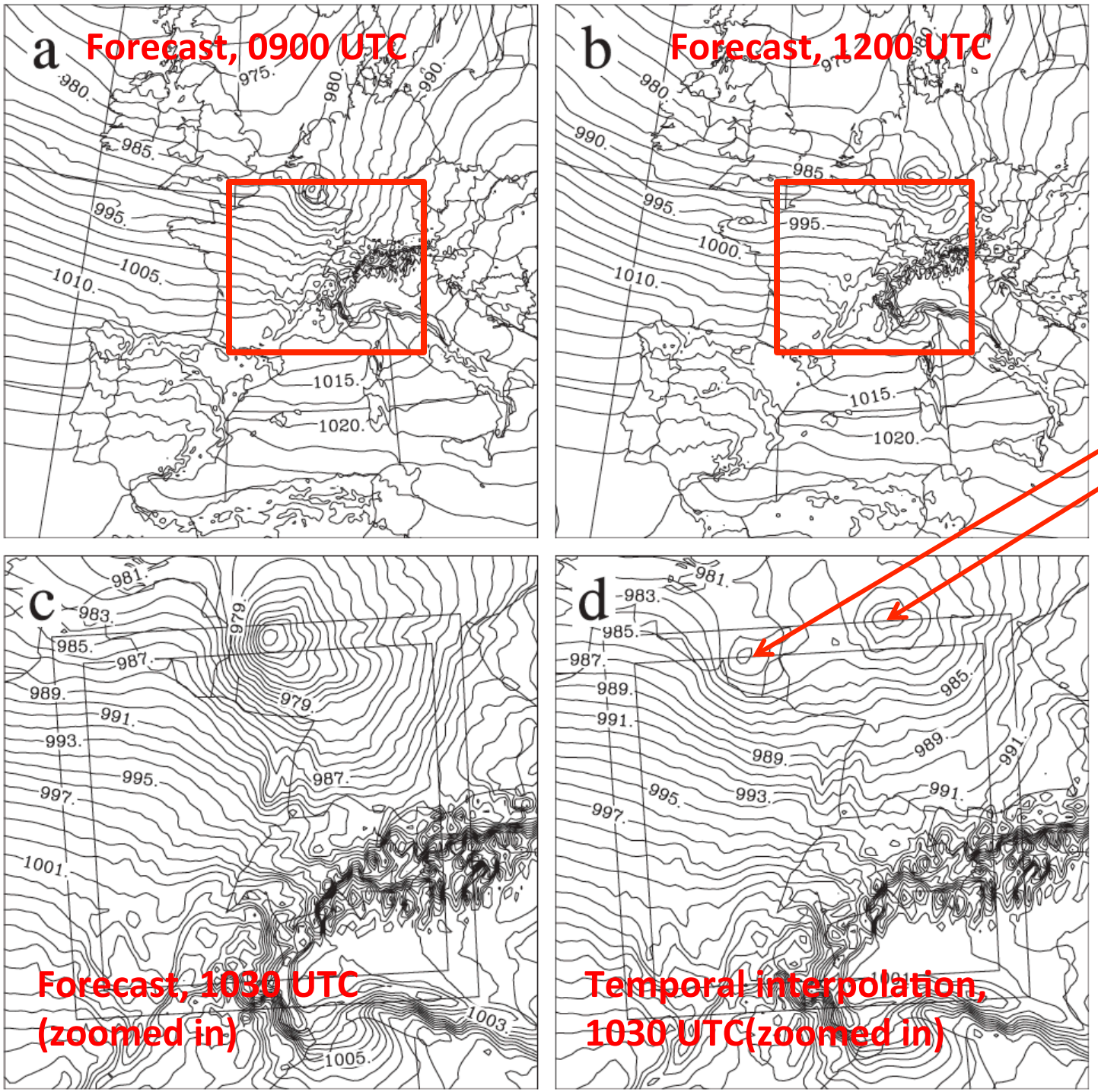
Nutter et al.'s experiments, cont'd.

Effects of using linear temporal interpolation with 3-hourly boundary conditions



Here interior, exterior resolutions are the same, but correct LBC's are used only every third hour, and otherwise interpolated, as is commonly practiced. Shading is vorticity error, contours are streamfunction error.

Notice pulsing of errors, reduced on boundaries at 0, 3, 6, 9, but larger at in-between times. However, errors generally grow as a result of temporal interpolation.

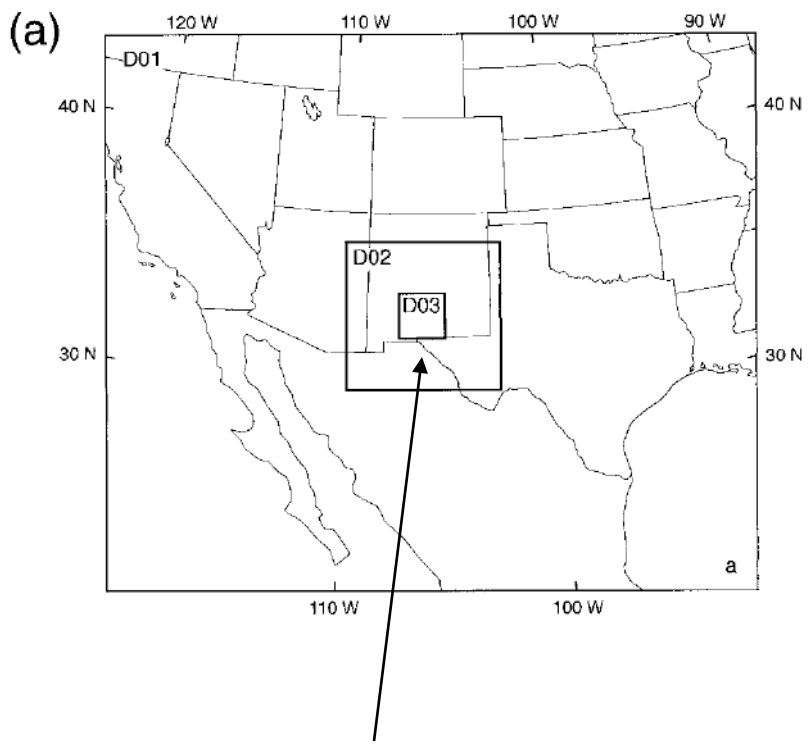


Example: interpolation errors in Aladin simulation of 1999 Lothar storm

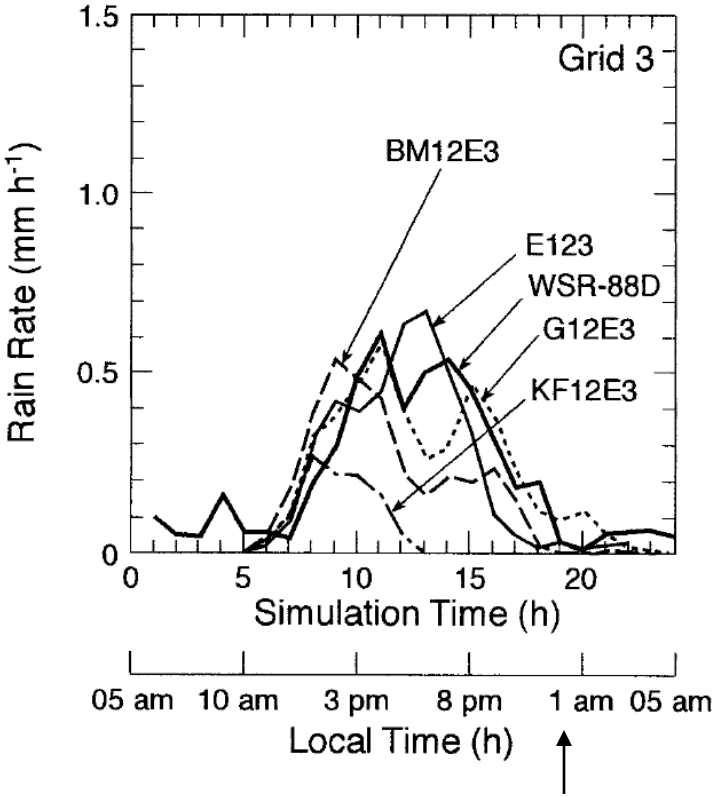
temporal interpolation creates two lows from one.

Ref: Tudor & Termonia, MWR, July 2010

Problems caused by using outer domain convective parameterization with explicit convection in nest



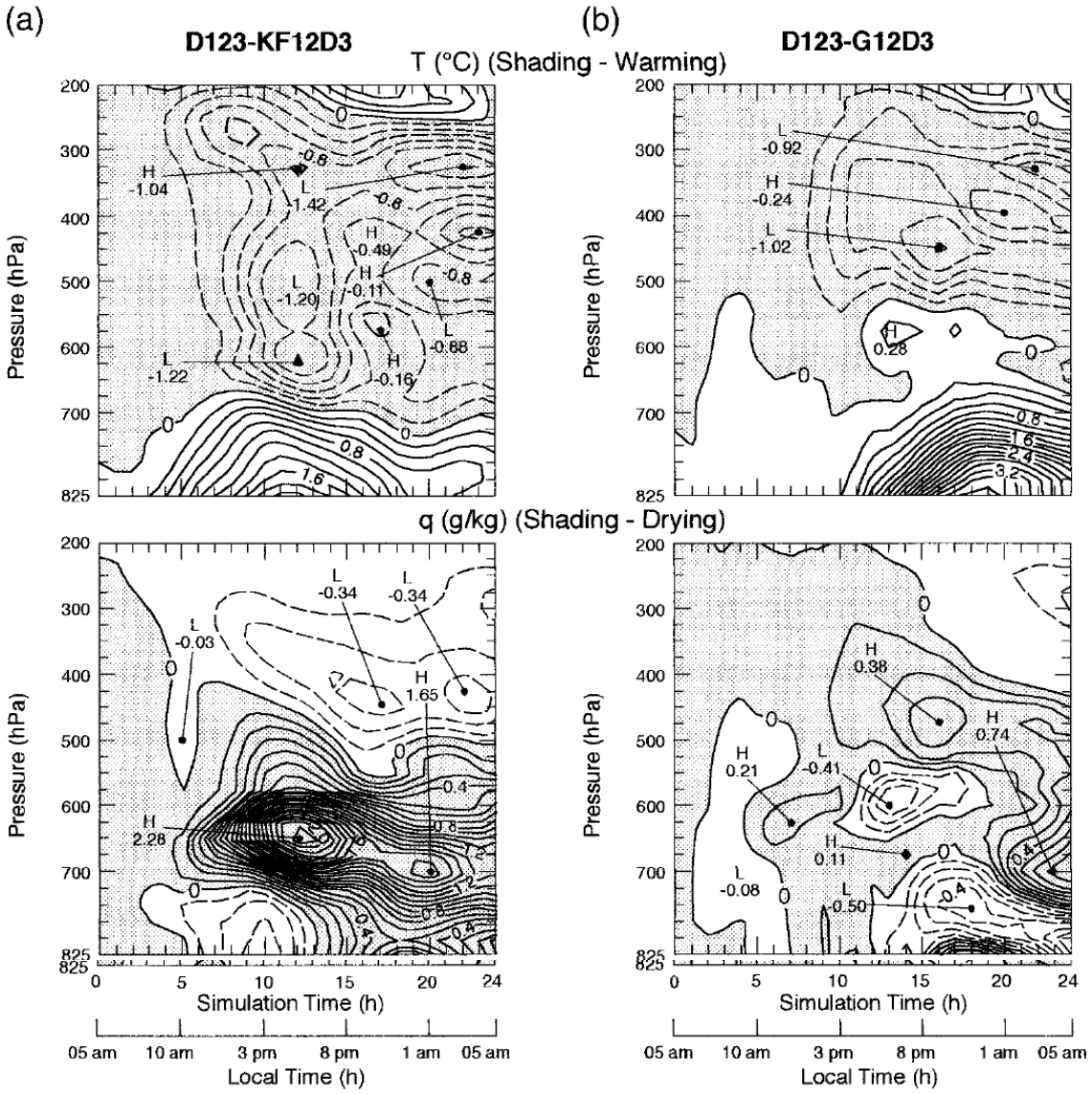
Simulation of nested domains, explicitly resolved convection on inner (3.3 km grid spacing, various parameterized convection on outer
(10, 30 km).



Rainfall on inner domain affected by choice of what is done on outer domain. E123 is explicit on each domain, KF12E3 is Kain-Fritsch on 1&2, explicit on 3.

Warner and Hsu, *MWR*, July 2000

Problems caused by using outer domain convective parameterization with explicit convection in nest



Domain 3 forced to be dry simulation. In comparison with a dry simulation everywhere (D123) to one where convection permitted on outer domains (KF12D3, G12D3), the effect of parameterized convection on the thermodynamics of the interior grid is to heat the upper troposphere and dry the middle troposphere.

Possible lessons:

(1) If possible, use **large** convectively resolving domains to spread the effect out over a larger area.
(2) Develop/utilize convective parameterizations with physically reasonable mass-field responses.
(3) Develop ways of tuning convective parameterizations to minimize nonphysical competition between explicit and parameterized convection.

Issue 2: predictability and observation type

- How do chaos and model uncertainty limit our ability to predict intensity?

- Are we gathering the most crucial observations to improve intensity predictions?

How rapidly will errors grow in an intensity forecast as a result of:

- meso-γ errors, e.g., position of updrafts and downdrafts in eyewall (occasionally somewhat observable through tail Doppler);
- meso-β changes, e.g., eyewall size, asymmetries (observable in many situations)
- meso-α changes, e.g., size of hurricane itself, radial structure, SST patterns (relatively observable)
- synoptic scales (environmental shear, steering flow)?

Scant literature on TC predictability

- Sang et al., QJRMS, 2008: Small moisture perturbations can change eye wall asymmetries dramatically; these are essentially unpredictable. Model physics (e.g., inclusion of warm rain) affects intensification rate.
- Sippel and Zhang, JAS, June 2010: Hurricane Humberto's forecast intensity varied among EnKF members; variations in moisture, frontal position, low-level instability affected precip & storm strength.
- Sippel and Zhang, 2008, Zhang and Sippel 2009: Variations in initial CAPE, moisture affect genesis forecast of nondeveloping TC in Gulf of Mexico.
- Holland, JAS, 1997: MPI sensitive to relative humidity under eye wall, height of the warm core, i.e., to some inner-core details rather than synoptic conditions.
- Hendricks et al. 2004; Krishnamurti et al. 2005; Montgomery et al. 2006: Upscale growth of vortical hot towers (limits predictability?).
- Reynolds et al. 2011 cyclone workshop: Adjoint sensitivity calculations show sensitivity to synoptic environment, mid-level Q in T. Lupit.

Scant literature on TC predictability

- Sang et al., QJRMS, 2008: Small moisture perturbations can change eye wall asymmetries dramatically; these are essentially unpredictable. Model physics (e.g., inclusion of warm rain) affects intensification rate.
- Sippel and Zhang, JAS, intensity varied among position, low-level inst
- Sippel and Zhang, 200 CAPE, moisture affect genesis forecast of nondeveloping TC in Gulf of Mexico.
- Holland, JAS, 1997: MPI sensitive to relative humidity under eye wall, height of the warm core, i.e., to some inner-core details rather than synoptic conditions.
- Hendricks et al. 2004; Krishnamurti et al. 2005; Montgomery et al. 2006: Upscale growth of vortical hot towers (limits predictability?).
- Reynolds et al. 2011 cyclone workshop: Adjoint sensitivity calculations show sensitivity to synoptic environment, mid-level Q in T. Lupit.

To my mind, these articles don't necessarily contradict the notion that larger scales may be important, for the larger scales may determine the environment for the eye wall & its vortical hot towers.

Mid-latitude predictability theory

- Lorenz (1969, 1984), based on assumed -5/3 slope to energy spectrum, predicted rapid upscale cascade of errors, finite predictability timescale. Hard time limit to predictability.

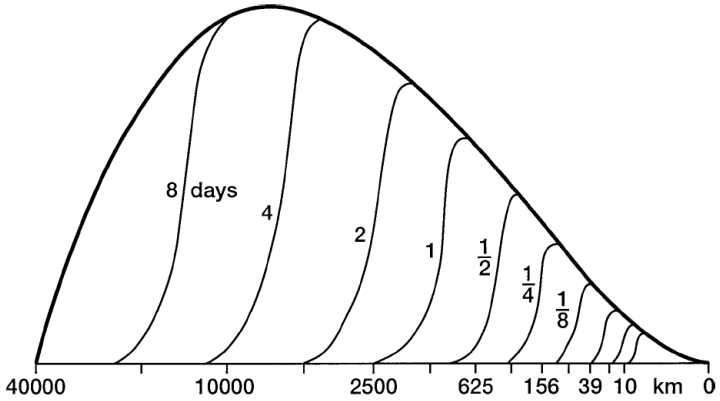


FIG. 1. Growth of errors initially confined to smallest scales, according to a theoretical model Lorenz (1984). Horizontal scales are on the bottom, and the upper curve is the full atmospheric motion spectrum.

- Tribbia and Baumhefner (subsequent slides) suggest more nuanced view; rapid growth of small scales; at larger scales, -3 slope more appropriate, with more exponential, modal growth of baroclinic scales. Reducing analysis errors may increase time range of usable predictive skill.

Mid-latitude predictability theory

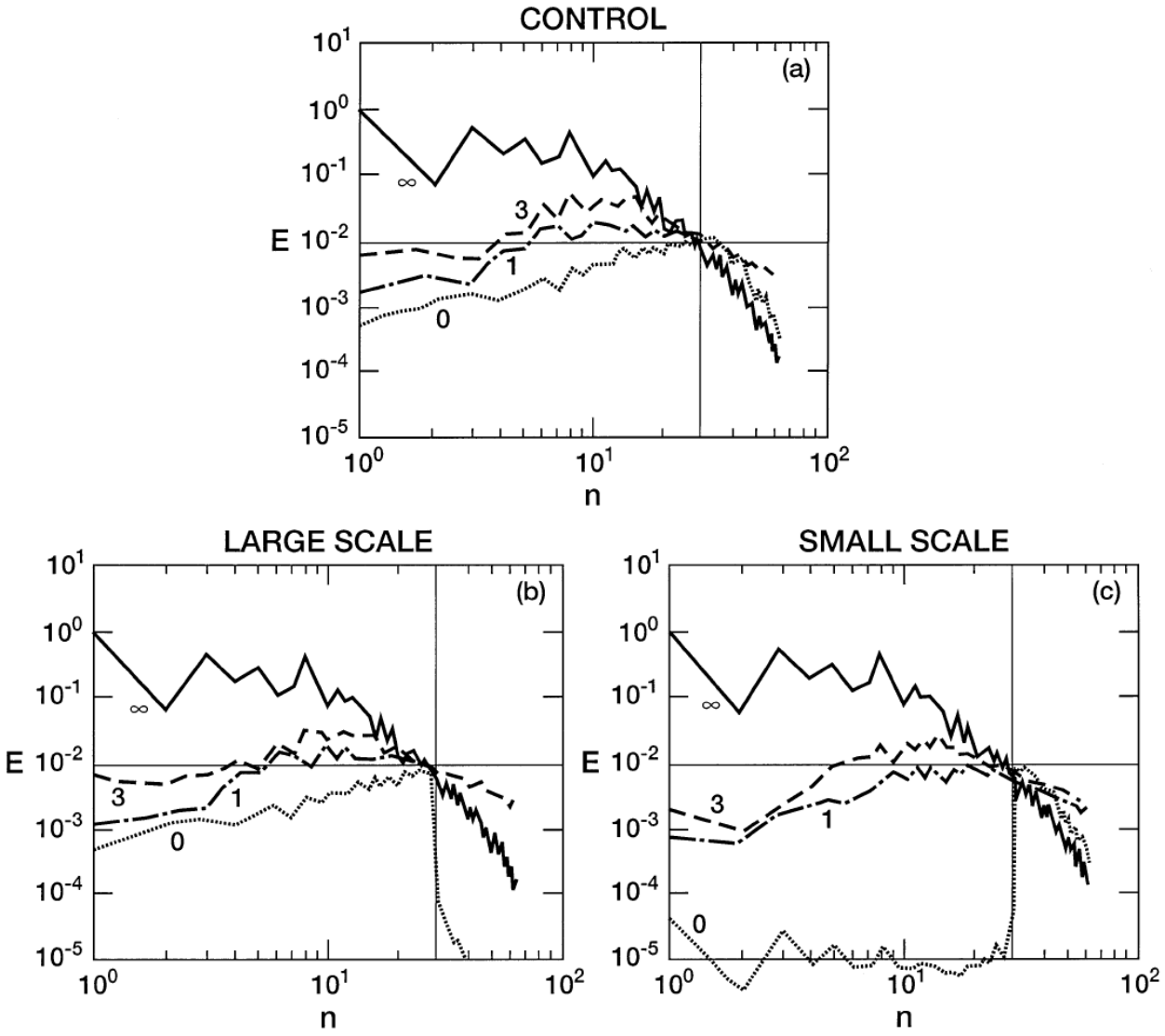


FIG. 2. Kinetic energy spectrum for 500 mb plotted for 2D wavenumber n indicated on abscissa. Amplitude of each wavenumber is on ordinate. Solid line is the spectrum of full field, and dotted (day 0), dashed-dotted (day 1), and dashed lines (day 3) are the spectrum of the difference fields from ensemble members (a) full perturbation in all scales, (b) perturbations only in wavenumber ≤ 30, and (c) perturbations only in wavenumber >30.

Experiment with T63 NCAR CCM3. Nature run + perturbed initial conditions generally consistent with analysis error. Consider the growth of the error in kinetic energy. 0, 1, and 3 lines denote spectrum of the error in kinetic energy for 0, 1, 3 - day lead. Solid black line is kinetic energy spectrum of full field.

Ref: Tribbia and Baumhefner, MWR, March 2004

Mid-latitude predictability theory

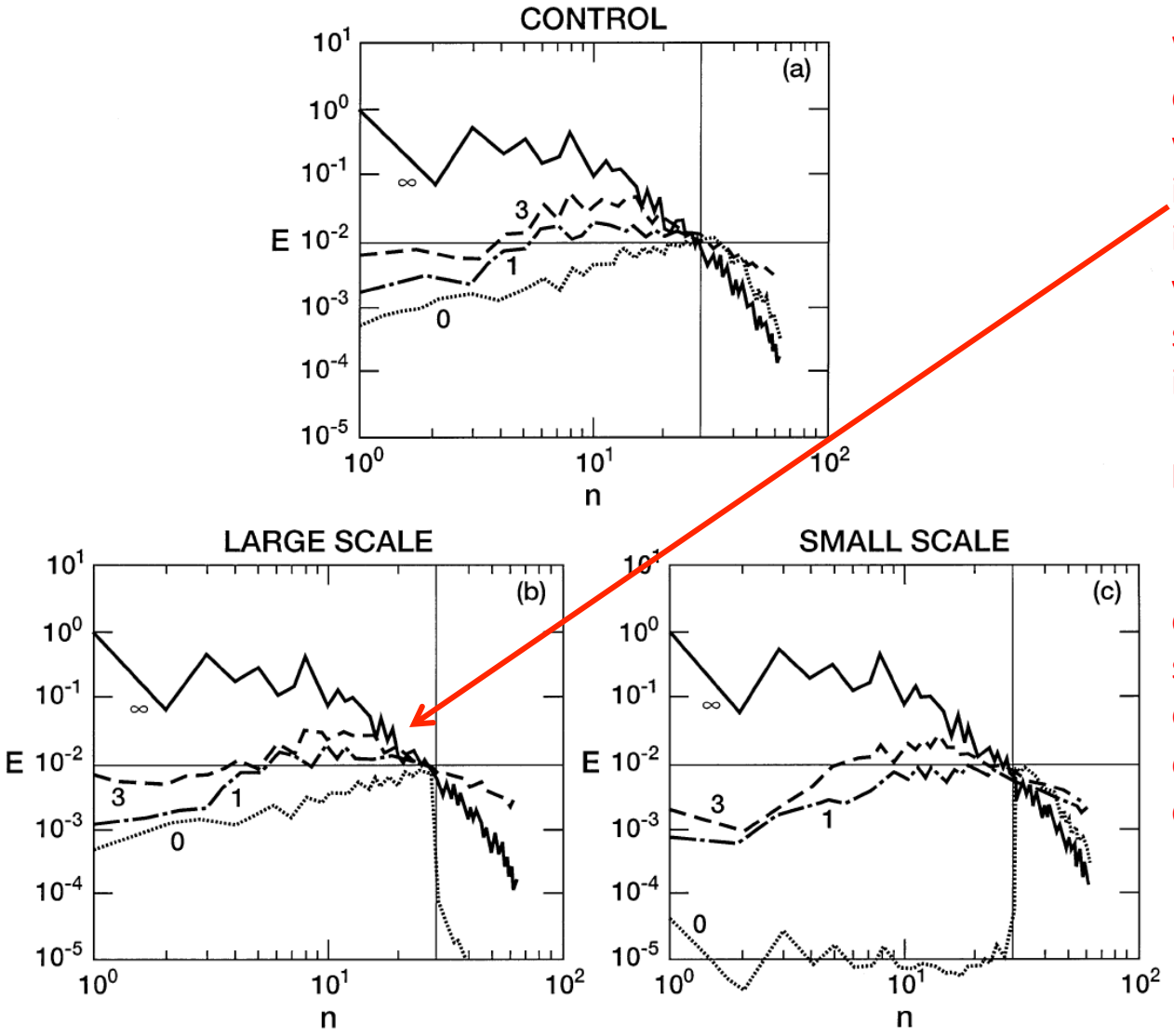


FIG. 2. Kinetic energy spectrum for 500 mb plotted for 2D wavenumber n indicated on abscissa. Amplitude of each wavenumber is on ordinate. Solid line is the spectrum of full field, and dotted (day 0), dashed-dotted (day 1), and dashed lines (day 3) are the spectrum of the difference fields from ensemble members (a) full perturbation in all scales, (b) perturbations only in wavenumber ≤ 30, and (c) perturbations only in wavenumber >30.

When errors are initially confined to the large scales but with perfect small-scale information, the errors in the small scales grow very, very quickly; the small-scale detail and error is determined by the large scales.

Possible importance to hurricanes?

Is it possible that no matter how carefully the inner-core detail is specified, if there are significant errors in the synoptic scales, this error will contaminate the inner core detail quickly.

Ref: Tribbia and Baumhefner, MWR, March 2004

Mid-latitude predictability theory

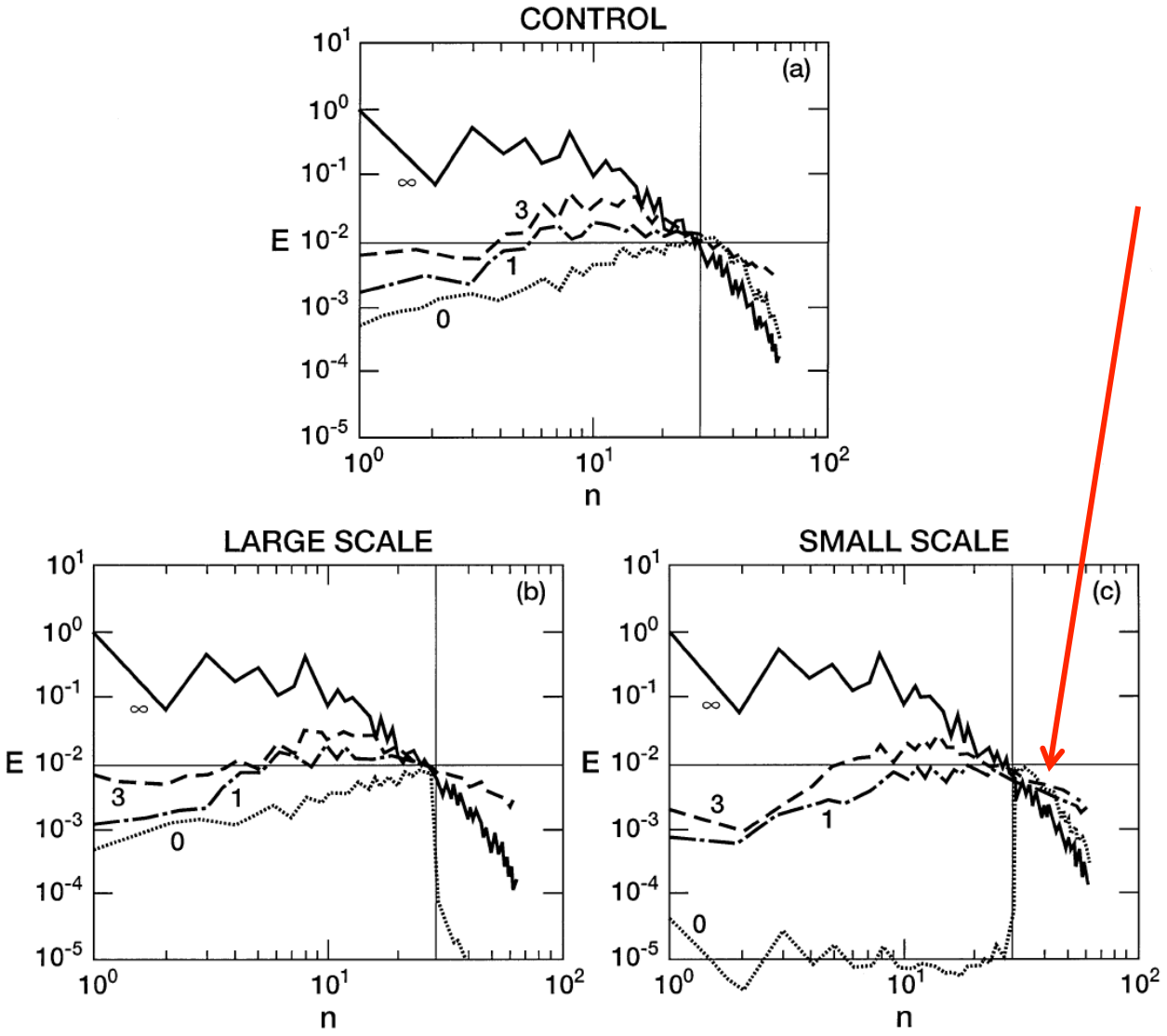


FIG. 2. Kinetic energy spectrum for 500 mb plotted for 2D wavenumber n indicated on abscissa. Amplitude of each wavenumber is on ordinate. Solid line is the spectrum of full field, and dotted (day 0), dashed-dotted (day 1), and dashed lines (day 3) are the spectrum of the difference fields from ensemble members (a) full perturbation in all scales, (b) perturbations only in wavenumber ≤ 30, and (c) perturbations only in wavenumber > 30.

When errors are initially confined to the small scales but with perfect large-scale information, the errors in the larger scales are generally reduced; it takes a while for errors to cascade upscale, whence they grow baroclinically (exponential growth till saturation).

Possible importance to hurricanes?

Is defining the synoptic environment and large scales of motion potentially more critical to TC intensity prediction? But is this analogy to mid-latitudes imperfect, since much of hurricane energy in inner core?

Ref: Tribbia and Baumhefner, MWR, March 2004

Tropical cyclone predictability theory?

There is no unifying theory now.

Relevance of mid-latitude theory unknown.

LGEM experiments suggest some *larger-scale* control of intensity

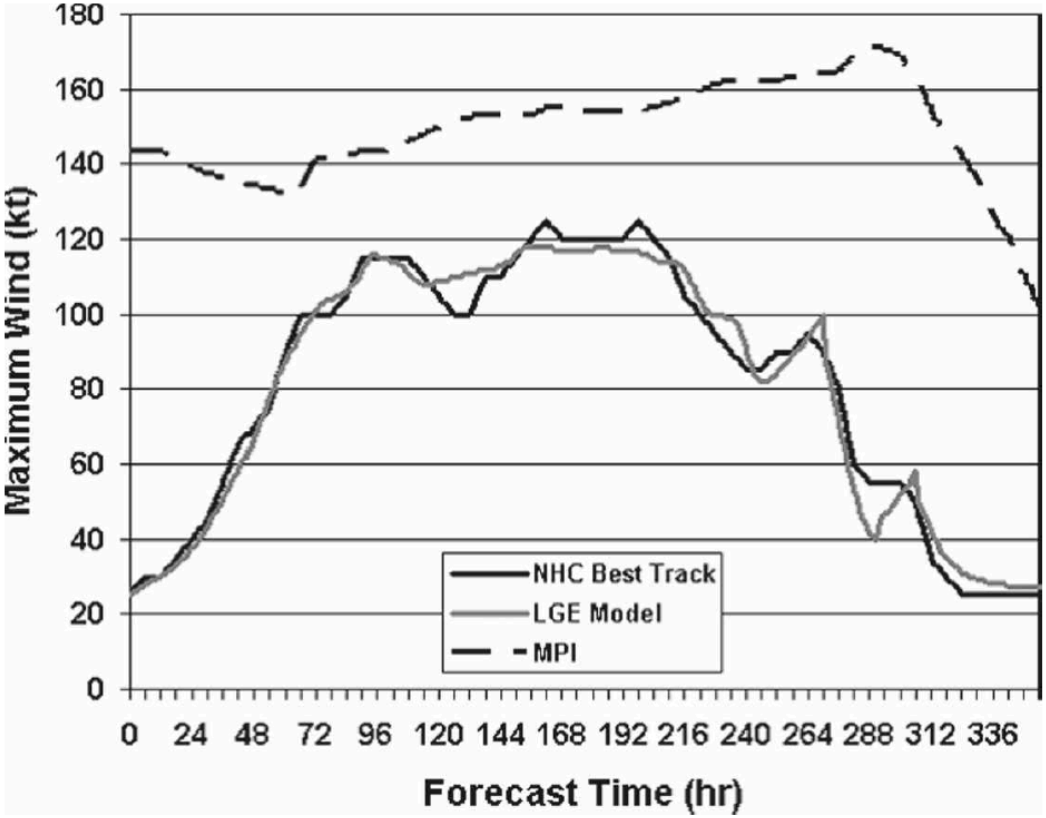


FIG. 6. The 14.75-day LGEM forecast of the intensity of Hurricane Frances (2004) and the corresponding NHC best-track intensity. The MPI estimated from the SST is also shown.

from DeMaria, MWR, Jan 2009

LGEM model of hurricane Frances. SST, vertical shear (S), and instability (C) are estimated from Reynolds SST and GFS analyses. The predictors **S** and **C** are normalized.

Remarkably good intensity forecasts of Frances were possible when LGEM model was driven with analysis data, with storm position, shear, CAPE known (as opposed to values being estimated from forecast data).

Suggests improvement in track, larger-scale conditions may improve intensity forecasts when processed through a statistical model like LGEM.

Possible HFIP experiments to assess predictability?

- Adjoint experiments, e.g., Reynolds et al.'s recent work.
- Perfect- or imperfect model sensitivity experiments
  - Generate "nature run" at high resolution, enough where hurricane details appear realistic. Example: doubly periodic HWRF, f-plane, warm SSTs, < 3-km grid spacing.
  - Generate perturbed initial conditions, magnitudes & structures consistent with analysis errors.
  - Experiments such as examining perturbations to inner core only, or storm environment only.
  - Examine rate of growth of differences between perturbed and nature run, effects on the prediction of intensity.
- Who would do this work? How funded? Leave to NSF?

Proposal: an HFIP "Predictability and Observability" subgroup

- Results here suggest it may be time to take a step back, do some basic science and use this inform HFIP long-term plans.
- Possible tasks of a new group:
  - Estimate the predictability of important hurricane phenomena such as intensity changes through classical predictability experimental methods, adjoint methods.
  - Experimentally determine what types of observations, in what regions (inner core, outer storm, environment, etc.) are most crucial to sample to improve predictions.
  - Feed the knowledge gained to the data assimilation, modeling, observing systems, ensemble groups to test proposed optimal observation and modeling techniques.
- Needed: personnel, computer time mostly.

Other work for subgroup

- To what extent are simplest experiments (uniform, f-plane) realistic? How are predictability estimates changed by more real-world conditions?
- What are the important *dynamical processes* governing TC intensity for inner core, whole hurricane, environment?

Conclusions

- HFIP is concentrating resources in high-resolution regional modeling, model physics, assimilation of inner-core data to improve intensity predictions.
- Regional modeling as we currently practice it may have numerical issues that hinder the ability of the model to render accurate forecasts.
- HFIP's may achieve greater long-term progress by clarifying basic issues of predictability and what is critical to observe.