ENSEMBLE FORECASTING AT NCEP

Zoltan Toth*[3]*,

Yuejian Zhu[4]*, Jun Du* [4]*, Richard Wobus* [4]*, Tim Marchok,*
Mozheng Wei[5]*,*

Ackn.: S. Lord [3]*, H.-L. Pan* [3]*, R. Buizza*[1]*, P. Houtekamer*[2]*, S. Tracton* [6]

[1] *: European Centre for Medium-Range Weather Forecasts, Reading UK (www.ecmwf.int)*

[2] *: Meteorological Service of Canada, Dorval, Quebec, Canada (www.msc-smc.ec.gc.ca)*

[3] *: NCEP/EMC, Washington, US (www.emc.ncep.noaa.gov)*

[4] *: SAIC at NCEP/EMC, Washington, US (www.emc.ncep.noaa.gov)*

[5] *: UCAR Visiting Scientist, NCEP/EMC, Washington, US*

[6] *: ONR*

OUTLINE

- MOTIVATION FOR ENSEMBLE/PROBABILISTIC FORECASTING
  - User Needs
  - Scientific needs
- SOURCES OF FORECAST ERRORS
  - Initial value
  - Model formulation
- ESTIMATING & SAMPLING FORECAST UNCERTAINTY
- DESCRIPTION OF NCEP ENSEMBLE FORECAST SYSTEMS
  - Global
  - Regional
  - Coupled ocean-atmosphere
- FORECAST EXAMPLES
- VERIFICATION
- ONGOING RESEARCH / OPEN QUESTIONS

MOTIVATION FOR ENSEMBLE FORECASTING

- ***FORECASTS ARE NOT PERFECT* - IMPLICATIONS FOR:**
  - **USERS:**
    - Need to know how often / by how much forecasts fail
    - Economically optimal behavior depends on
      - Forecast error characteristics
      - User specific application
        - » Cost of weather related adaptive action
        - » Expected loss if no action taken
      - EXAMPLE: Protect or not your crop against possible frost

    Cost = 10k, Potential Loss = 100k => Will protect if P(frost) > Cost/Loss=0.1

    - ***NEED FOR PROBABILISTIC FORECAST INFORMATION***

  - **DEVELOPERS:**
    - Need to improve performance - *Reduce error in estimate of first moment*
      - Traditional NWP activities (I.e., model, data assimilation development)
    - Need to account for uncertainty - *Estimate higher moments*
      - New aspect – How to do this?
    - Forecast is incomplete without information on forecast uncertainty
    - ***NEED TO USE PROBABILISTIC FORECAST FORMAT***

USER NEEDS – PROBABILISTIC FORECAST INFORMATION FOR MAXIMUM ECONOMIC BENEFIT

ECONOMIC VALUE OF FORECASTS

Given a particular forecast, a user either does or does not take action (eg, protects its crop against frost) *Mylne & Harrison, 1999*

OBSERVATION \ *FORECAST*	*YES*	*NO*
YES	H(its) ***Mitigated Loss***	M(isses) ***Loss***
NO	F(alse alarms) ***Cost***	C(orrect rejections) ***No Cost***

$$Mean\ Expense_{fc} = hML + mL + fC$$

$$Mean\ Expense_{perf} = oML$$

$$ME_{cl} = min[oL,\ oML + (1-o)C]$$

$$Value = \frac{ME_{cl} - ME_{fc}}{ME_{cl} - ME_{perf}}$$

o=climatological frequency

Optimum decision criterion for user action: P(weather event)=C/L

(Murphy 1977)

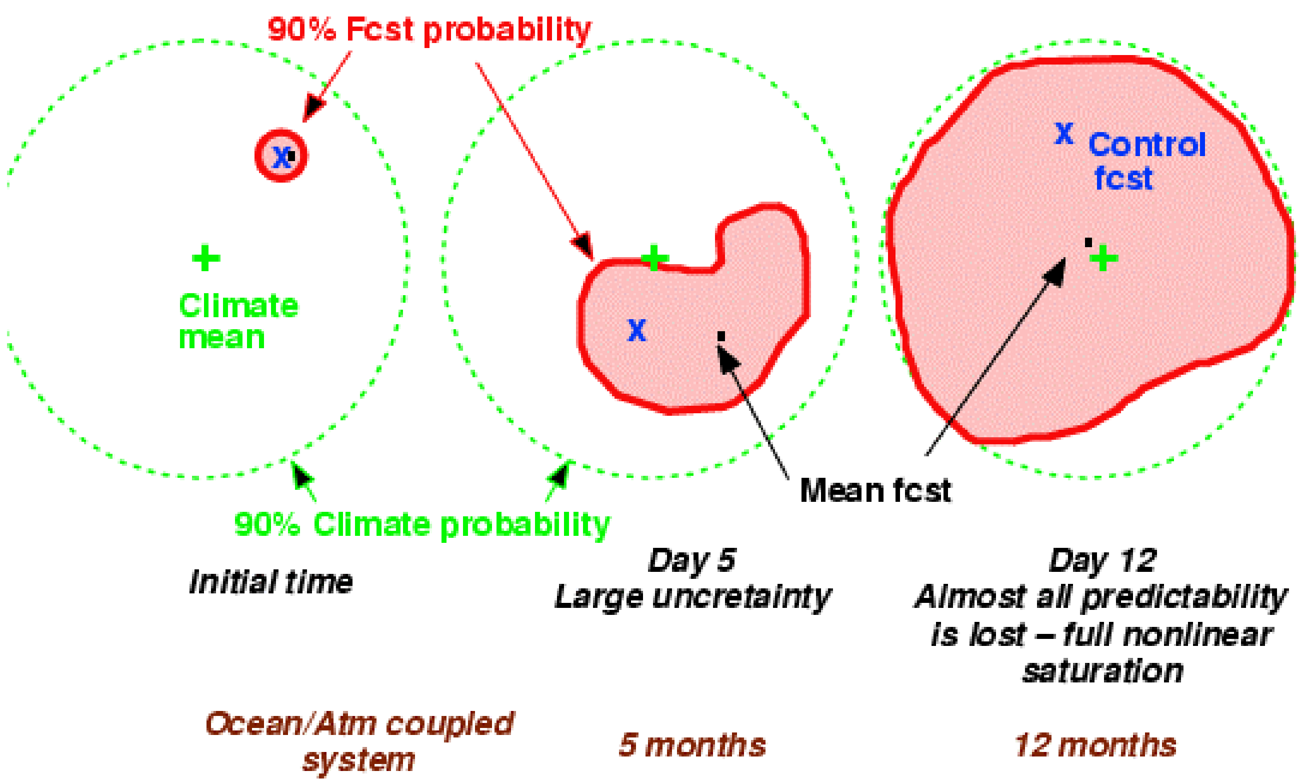
SCIENTIFIC NEEDS - DESCRIBE FORECAST UNCERTAINTY ARISING DUE TO CHAOS

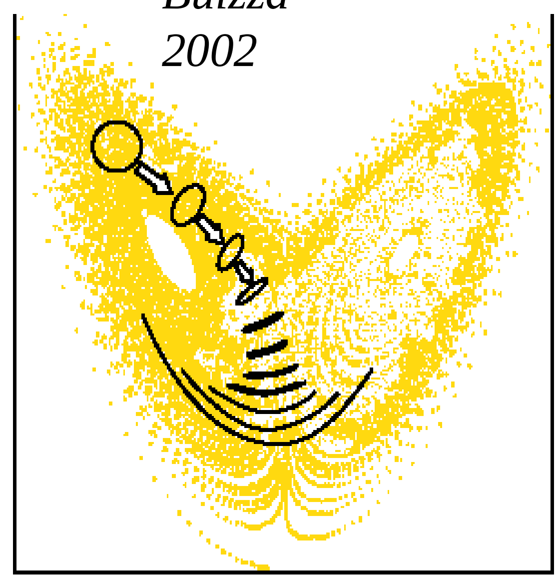
ORIGIN OF FORECAST UNCERTAINTY

1) The atmosphere is a **deterministic system** *AND* has at least one direction in which **perturbations grow**

2) **Initial** state (and model) has **error** in it ==>

Chaotic system + Initial error =(Loss of) Predictability





Buizza 2002

FORECASTING IN A CHAOTIC ENVIRONMENT

DETERMINISTIC APPROACH - PROBABILISTIC FORMAT

SINGLE FORECAST - ***One integration with an NWP model***

- Is not best estimate for future evolution of system
- Does not contain all attainable forecast information
- Can be combined with past verification statistics to form probabilistic forecast
  - Gives **no estimate of flow dependent variations in forecast uncertainty**

PROBABILISTIC FORECASTING - ***Based on Liuville Equations***

- Initialize with probability distribution function (pdf) at analysis time
- Dynamical forecast of pdf based on conservation of probability values
- **Prohibitively expensive** -
  - Very high dimensional problem (state space x probability space)
  - Separate integration for each lead time
  - Closure problems when simplified solution sought

FORECASTING IN A CHAOTIC ENVIRONMENT - 2

DETERMINISTIC APPROACH - PROBABILISTIC FORMAT

MONTE CARLO APPROACH – ***ENSEMBLE FORECASTING***

- ***IDEA*: Sample sources of forecast error**
  - Generate initial ensemble perturbations
  - Represent model related uncertainty
- ***PRACTICE*: Run multiple NWP model integrations**
  - Advantage of perfect parallelization
  - Use lower spatial resolution if short on resources
- ***USAGE*: Construct forecast pdf based on finite sample**
  - Ready to be used in real world applications
  - Verification of forecasts
  - Statistical post-processing (remove bias in 1st, 2nd, higher moments)

CAPTURES FLOW DEPENDENT VARIATIONS

IN FORECAST UNCERTAINTY

SOURCES OF FORECAST ERRORS

IMPERFECT KNOWLEDGE OF

INITIAL CONDITIONS

- Incomplete observing system (not all variables observed)
- Inaccurate observations (instrument/representativeness error)
- Imperfect data assimilation methods
  - Statistical approximations (eg, inaccurate error covariance information)
  - Use of imperfect NWP forecasts (due to initial and model errors) –
  - Effect of cycling (forecast errors "inherited" by analysis – use breeding)

GOVERNING EQUATIONS:

- Imperfect model
  - Structural uncertainty (eg, choice of structure of convective scheme)
  - Parametric uncertainty (eg, critical values in parameterization schemes)
  - Closure/truncation errors (temporal/spatial resolution; spatial coverage, etc)

NOTES:

- Two main sources of forecast errors hard to separate =>
- Very little information is available on model related errors
- Tendency to attribute all forecast errors to model problems

SAMPLING FORECAST ERRORS =

REPRESENTING ERRORS ORIGINATING FROM TWO MAIN SOURCES

INITIAL CONDITION RELATED ERRORS – *"Easy"*

- Sample initial errors
- Run ensemble of forecasts
- It works
  - Flow dependent variations in forecast uncertainty captured (show later)
  - Difficult or impossible to reproduce with statistical methods

MODEL RELATED ERRORS – *No theoretically satisfying approach*

- Change structure of model (eg, use different convective schemes, etc, MSC)
- Add stochastic noise (eg, perturb diabatic forcing, ECMWF)
- Works? Advantages of various approaches need to be carefully assessed
  - Are flow dependent variations in uncertainty captured?
  - Can statistical post-processing replicate use of various methods?
- Need for a
  - more comprehensive and
  - theoretically appealing approach

SAMPLING INITIAL CONDITION ERRORS

CAN SAMPLE ONLY WHAT'S KNOWN – FIRST NEED TO

ESTIMATE INITIAL ERROR DISTRIBUTION

THEORETICAL UNDERSTANDING – THE MORE ADVANCED A SCHEME IS
(e. g., 4DVAR, Ensemble Kalman Filter)

- The lower the overall error level is
- The more the error is concentrated in subspace of Lyapunov/Bred vectors

PRACTICAL APPROACHES **–**

ONLY SOLUTION IS MONTE CARLO (ENSEMBLE) SIMULATION

- ***Statistical approach*** (dynamically growing errors neglected)
    - Selected estimated statistical properties of analysis error reproduced
        - Baumhefner et al – Spatial distribution; wavenumber spectra
        - ECMWF – Implicite constraint with use of Total Energy norm
- ***Dynamical approach*** – Breeding cycle (NCEP)
    - Cycling of errors captured
    - Estimates subspace of dynamically fastest growing errors in analysis
- ***Stochastic-dynamic approach*** – Perturbed Observations method (MSC)
    - Perturb all observations (given their uncertainty)
    - Run multiple analysis cycles
    - Captures full space (growing + non-growing) of analysis errors

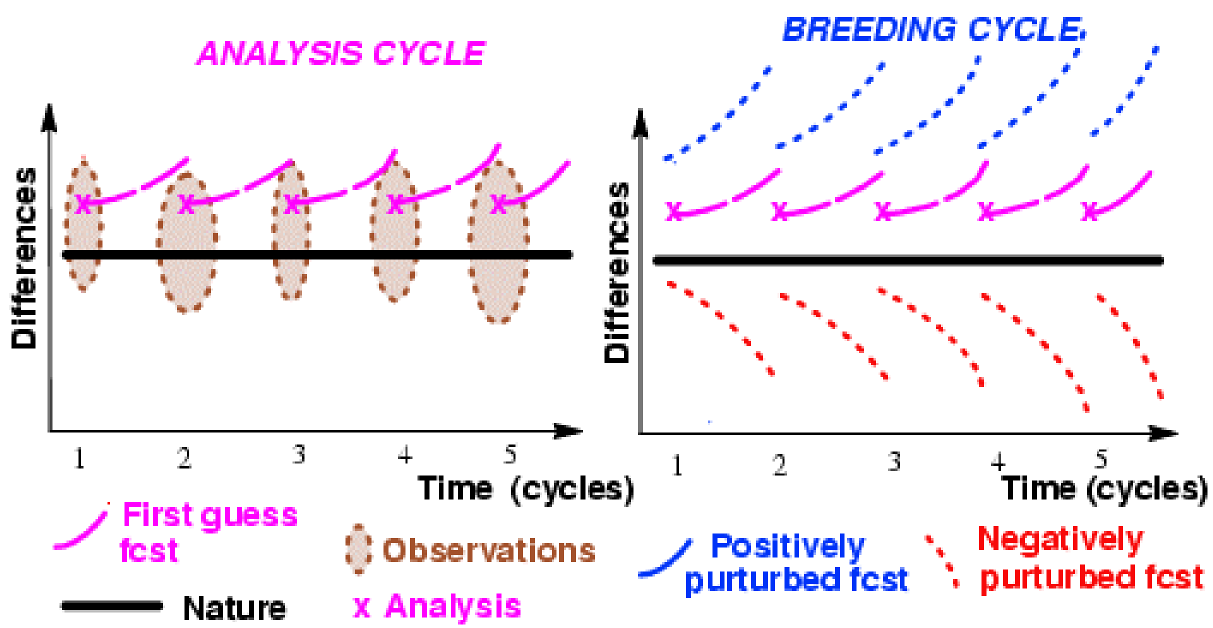
SAMPLING INITIAL CONDITION ERRORS

THREE APPROACHES – SEVERAL OPEN QUESTIONS

- **RANDOM SAMPLING** – **Perturbed observations method** (MSC)
  - Represents all potential error patterns with realistic amplitude
  - Small subspace of growing errors is well represented
  - Potential problems:
    - Much larger subspace of non-growing errors poorly sampled,
    - Yet represented with realistic amplitudes
- **SAMPLE GROWING *ANALYSIS* ERRORS** – **Breeding** (NCEP)
  - Represents dynamically growing analysis errors
  - Ignores non-growing component of error
  - Potential problems:
    - May not provide "wide enough" sample of growing perturbations
    - Statistical consistency violated due to directed sampling? Forecast consequences?
- **SAMPLE FASTEST GROWING *FORECAST* ERRORS** – **SVs** (ECMWF)
  - Represents forecast errors that would grow fastest in linear sense
  - Perturbations are optimized for maximum forecast error growth
  - Potential problems:
    - Need to optimize for each forecast application (or for none)?
    - Linear approximation used
    - Very expensive

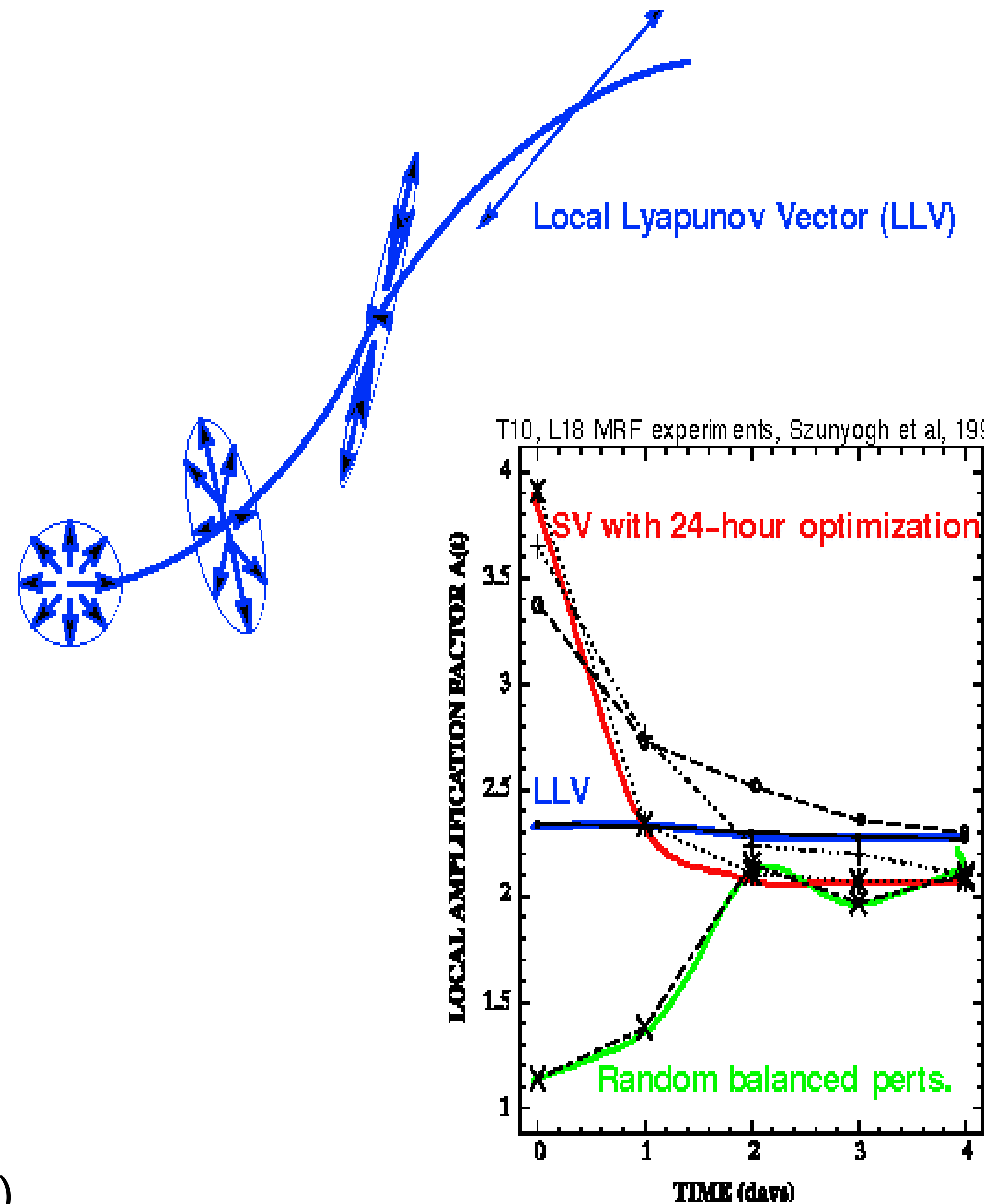
ESTIMATING AND SAMPLING INITIAL ERRORS: THE BREEDING METHOD

- **DATA ASSIM:** Growing errors due to cycling through NWP forecasts
- **BREEDING:** - Simulate effect of obs by rescaling nonlinear perturbations
  - Sample subspace of most rapidly growing analysis errors
    - Extension of linear concept of Lyapunov Vectors into nonlinear environment
    - Fastest growing nonlinear perturbations
    - Not optimized for future growth –
      - Norm independent
      - Is non-modal behavior important?



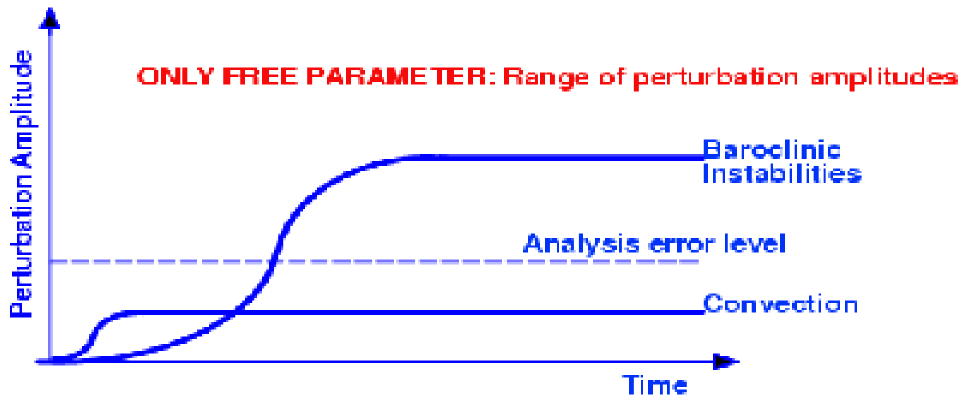
LYAPUNOV, SINGULAR, AND BRED VECTORS

- **LYAPUNOV VECTORS (LLV):**
  - Linear perturbation evolution
  - Fast growth
  - Sustainable
  - Norm independent
  - Spectrum of LLVs
- **SINGULAR VECTORS (SV):**
  - Linear perturbation evolution
  - Fastest growth
  - Transitional (optimized)
  - Norm dependent
  - Spectrum of SVs
- **BRED VECTORS (BV):**
  - Nonlinear perturbation evolution
  - Fast growth
  - Sustainable
  - Norm independent
  - Can orthogonalize (Boffeta et al)



PERTURBATION EVOLUTION

- **PERTURBATION GROWTH**
  - Due to effect of instabilities
  - Linked with atmospheric phenomena (e.g, frontal system)
- **LIFE CYCLE OF PERTURBATIONS**
  - Associated with phenomena
  - Nonlinear interactions limit perturbation growth
  - Eg, convective instabilities grow fast but are limited by availability of moisture etc
- **LINEAR DESCRIPTION**
  - May be valid at beginning stage only
  - If linear models used, need to reflect nonlinear effects at given perturb. Amplitude
- **BREEDING**
  - Full nonlinear description
  - Range of typical perturbation amplitudes is only free parameter



DESCRIPTION OF NCEP ENSEMBLE FORECAST SYSTEMS

OPERATIONAL

- Global ensemble forecast system (based on MRF/GFS system)
- Limited Area Ensemble Forecast System (SREF, over NA)

PLANNED

- Seasonal Ensemble Forecast System (Planned, coupled model)

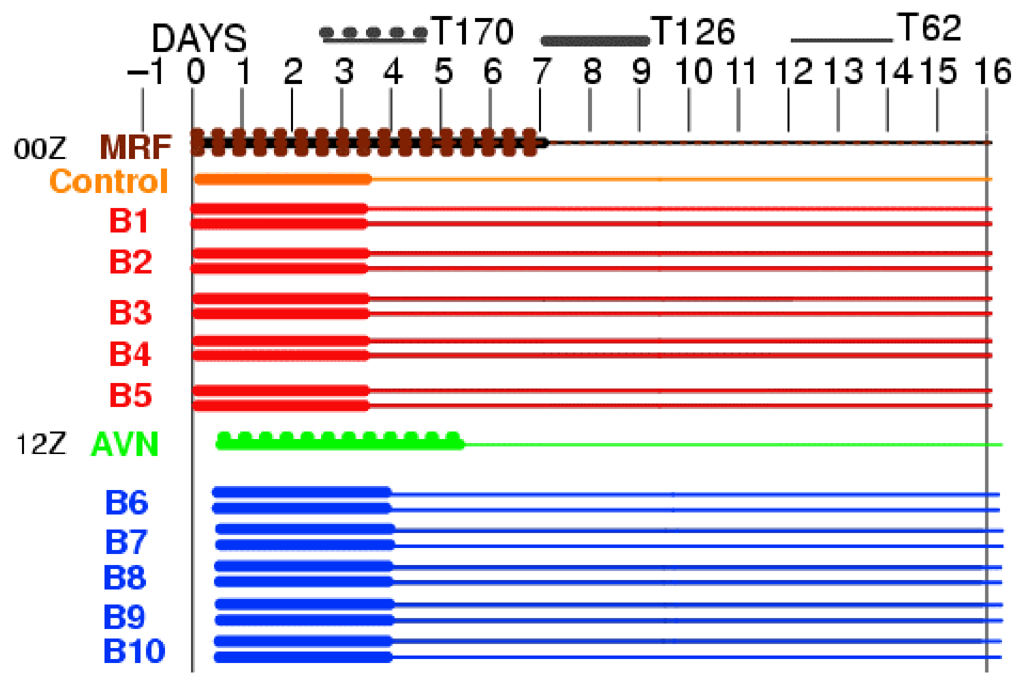
FOR EACH SYSTEM:

- Configuration
- Initial perturbations
- Model perturbations
- Main users
- Applications
- Examples
- Discussion/Conclusion

NCEP GLOBAL ENSEMBLE FORECAST SYSTEM

CURRENT *(APRIL 2003)* SYSTEM

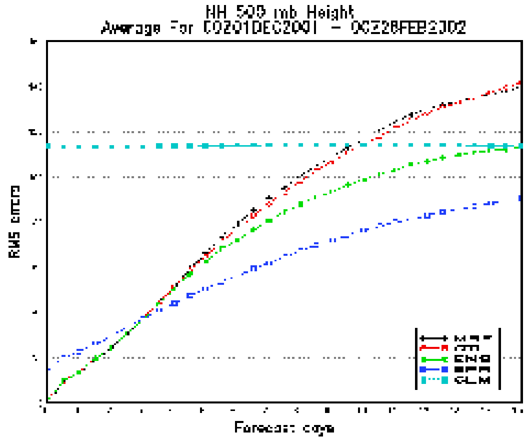
- 10 members out to 16 days
- 2 *(4)* times daily
- T126 out to 3.5 *(7.5)* days
- Model error not yet represented



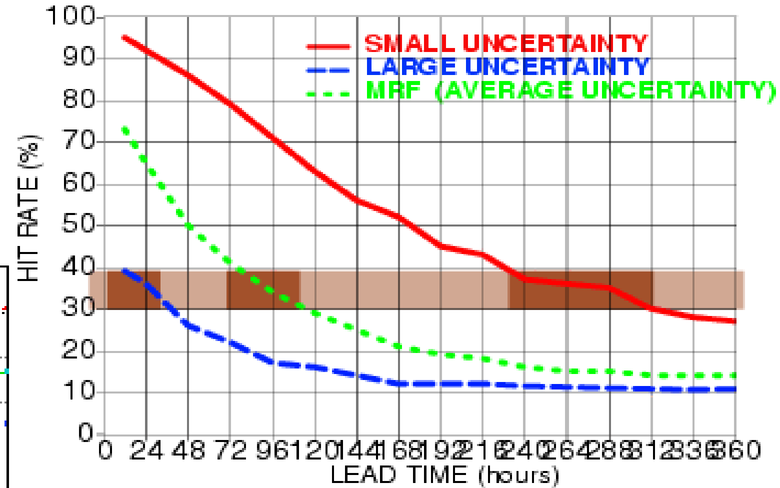
- **PLANS**
- ***Initial perturbations***
  - Rescale bred vectors via ETKF
  - Perturb surface conditions
- ***Model errors***
  - Push members apart
  - Multiple physics (combinations)
  - Change model to reflect uncertainties
- ***Post-processing***
  - Multi-center ensembles
  - Calibrate 1st & 2nd moment of pdf
  - Multi-modal behavior?

ADVANTAGES OF USING ENSEMBLE (VS. CONTROL) FCSTS

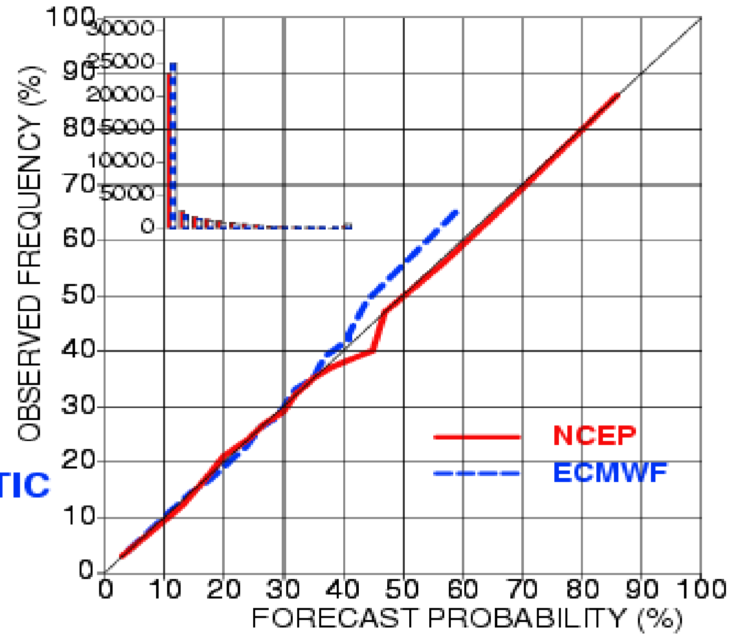
1) IMPROVED EXPECTED VALUE FORECAST



2) CASE DEPENDENT UNCERTAINTY ESTIMATE



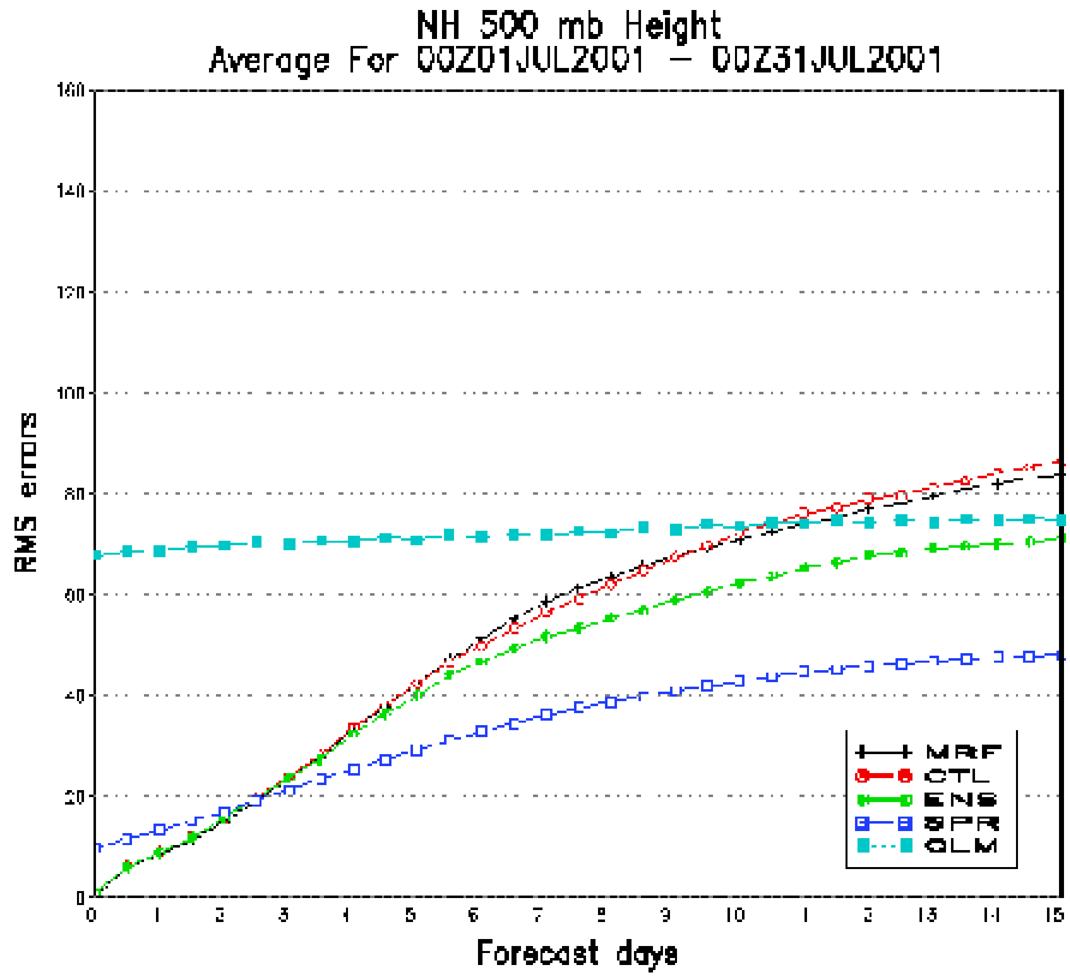
3) DETAILED PROBABILISTIC FORECAST



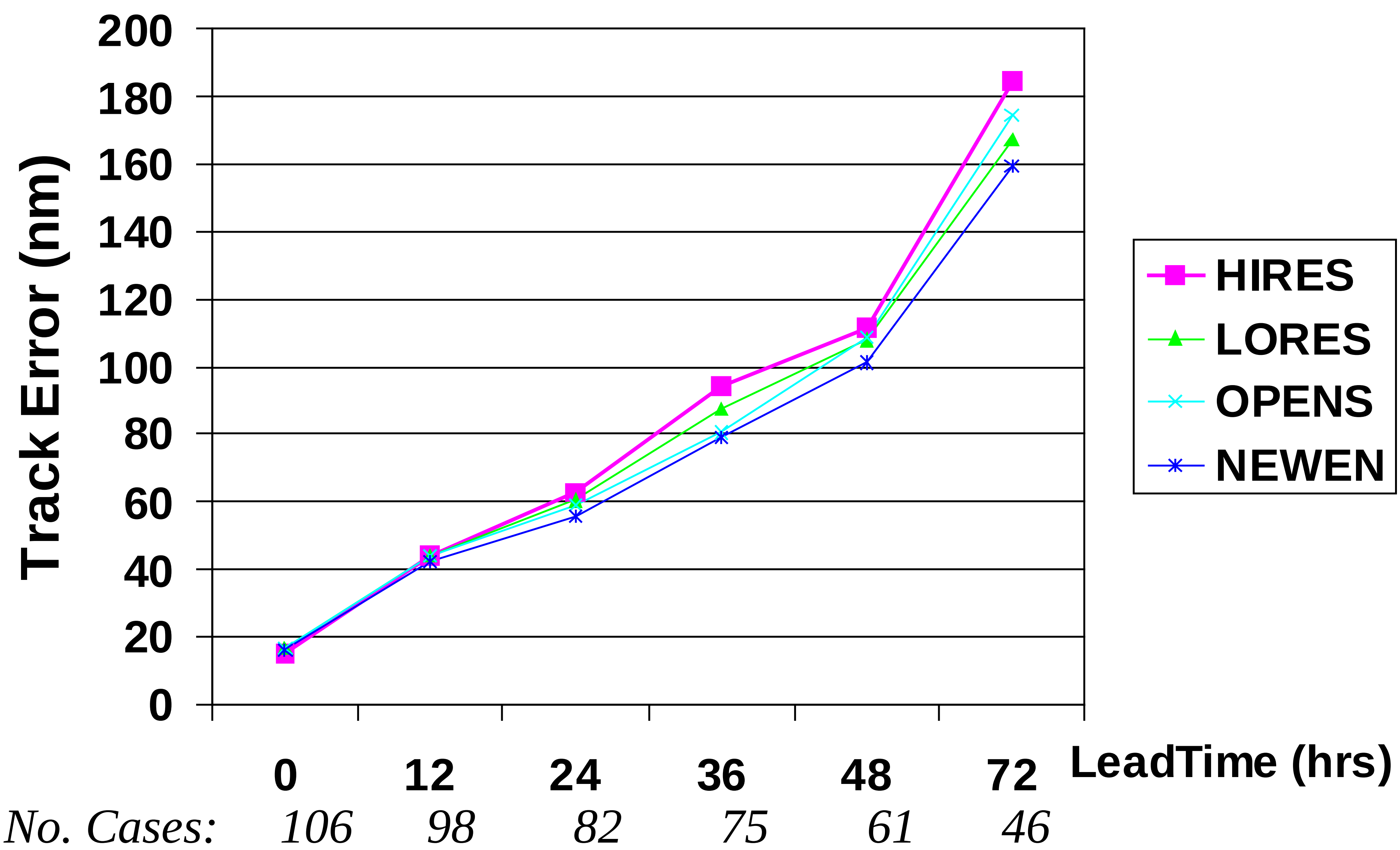
BEST ESTIMATE OF FUTURE STATE

- RMS error
  - Ensemble mean beats control
  - Skill above climatology even in summer, out to 16 days
  - Low resolution control beats hires control

- Ensemble spread
  - Lower than ensemble mean error
    - Due to lack of model perturbations



Aug-Sept. 2002 3-basin Tropical torm track errors



RESOLUTION OF ENSEMBLE BASED PROB. FCSTS

QUESTION:

What are the typical **variations in foreseeable forecast uncertainty?**

What variations in predictability can the ensemble resolve?

METHOD:

Ensemble mode value to distinguish high/low predictability cases

Stratify cases according to ensemble mode value –

Use 10–15% of cases when ensemble is highest/loewest

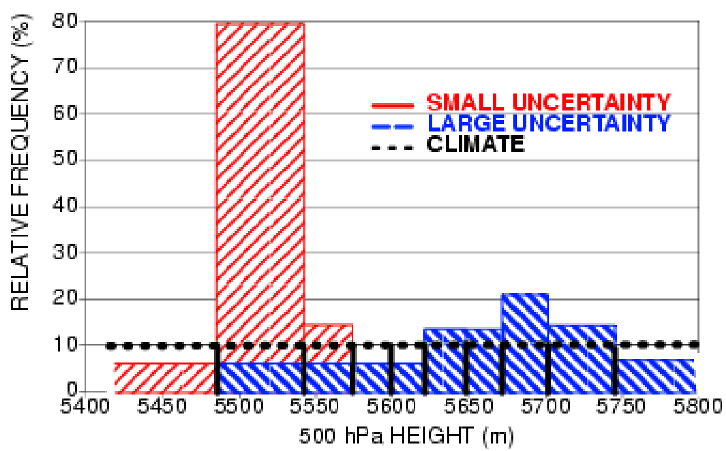
DATA:

NCEP **500 hPa NH extratropical ensemble fcsts** for March–May 1997

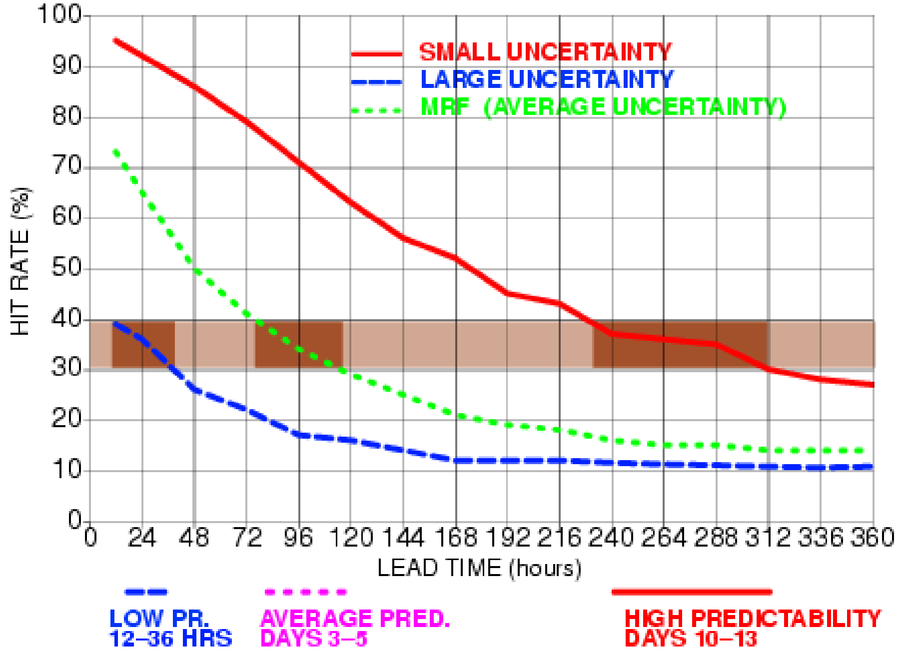
14 perturbed fcsts and high resolution control

VERIFICATION:

Hit rate for ensemble mode and hires control fcst



SEPARATING HIGH VS. LOW UNCERTAINTY FCSTS



*THE **UNCERTAINTY OF FCSTS** CAN BE **QUANTIFIED IN ADVANCE***

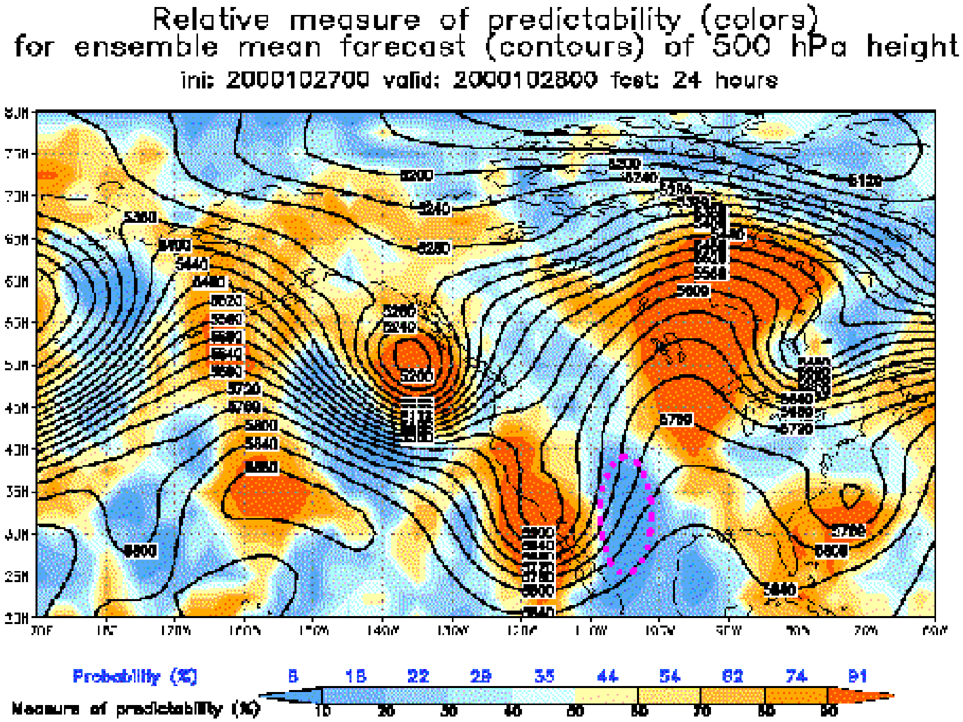
HIT RATES FOR 1–DAY FCSTS

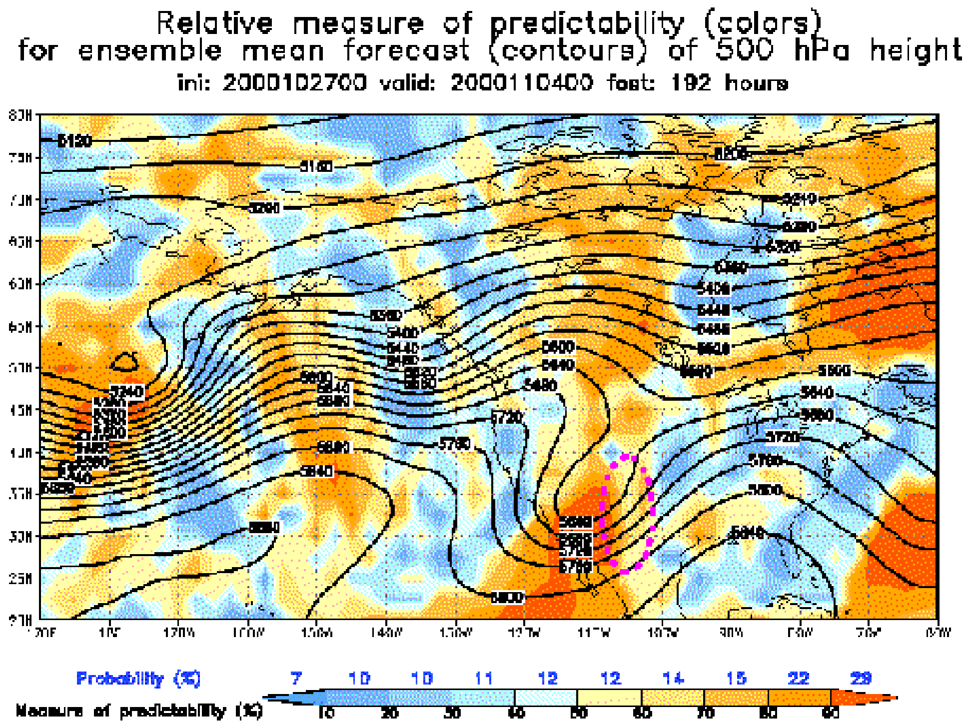
*CAN BE **AS LOW AS 36%**, OR **AS HIGH AS 92%***

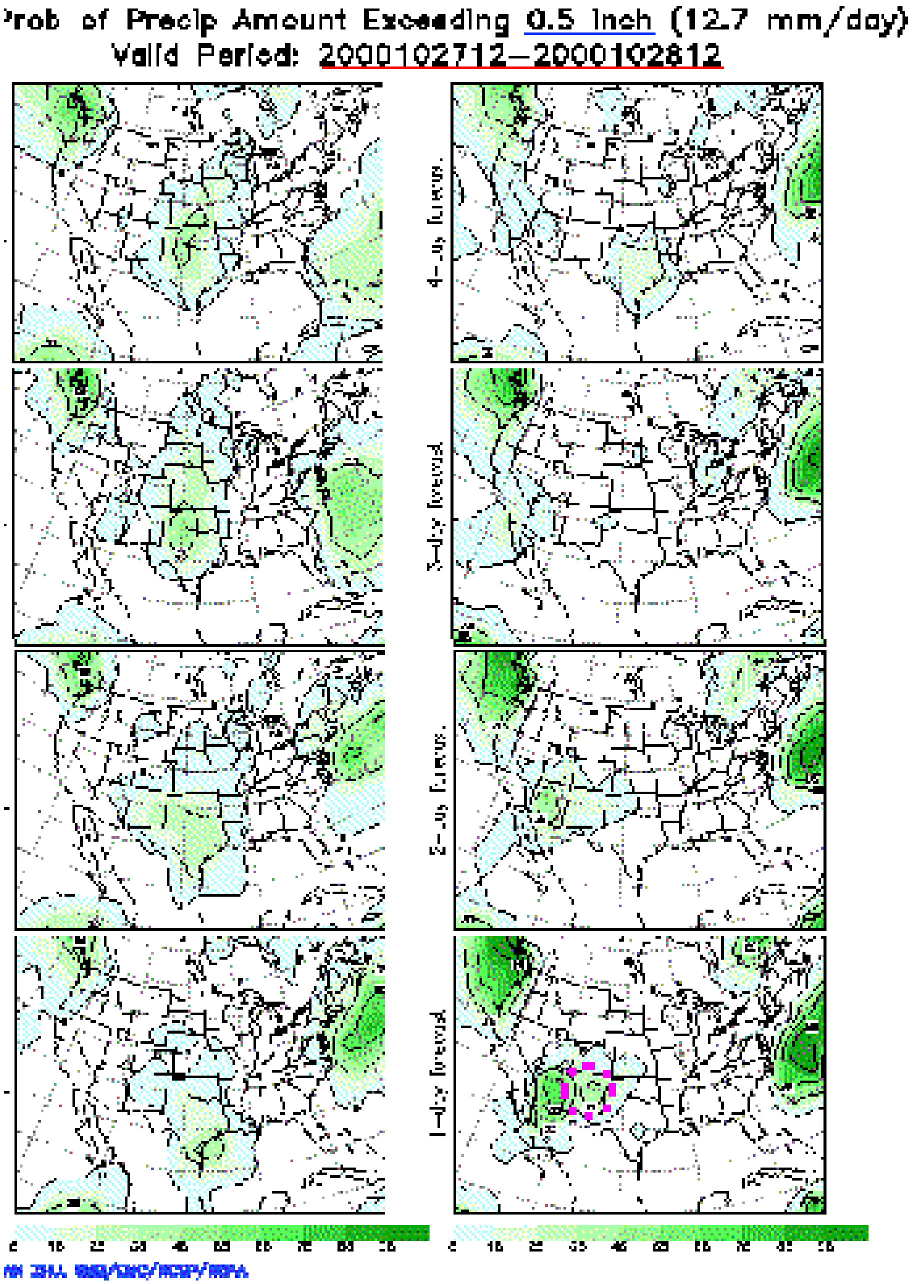
10–15% OF THE TIME A 12–DAY FCST CAN BE AS GOOD, OR A 1–DAY FCST CAN BE AS POOR AS AN AVERAGE 4–DAY FCAST

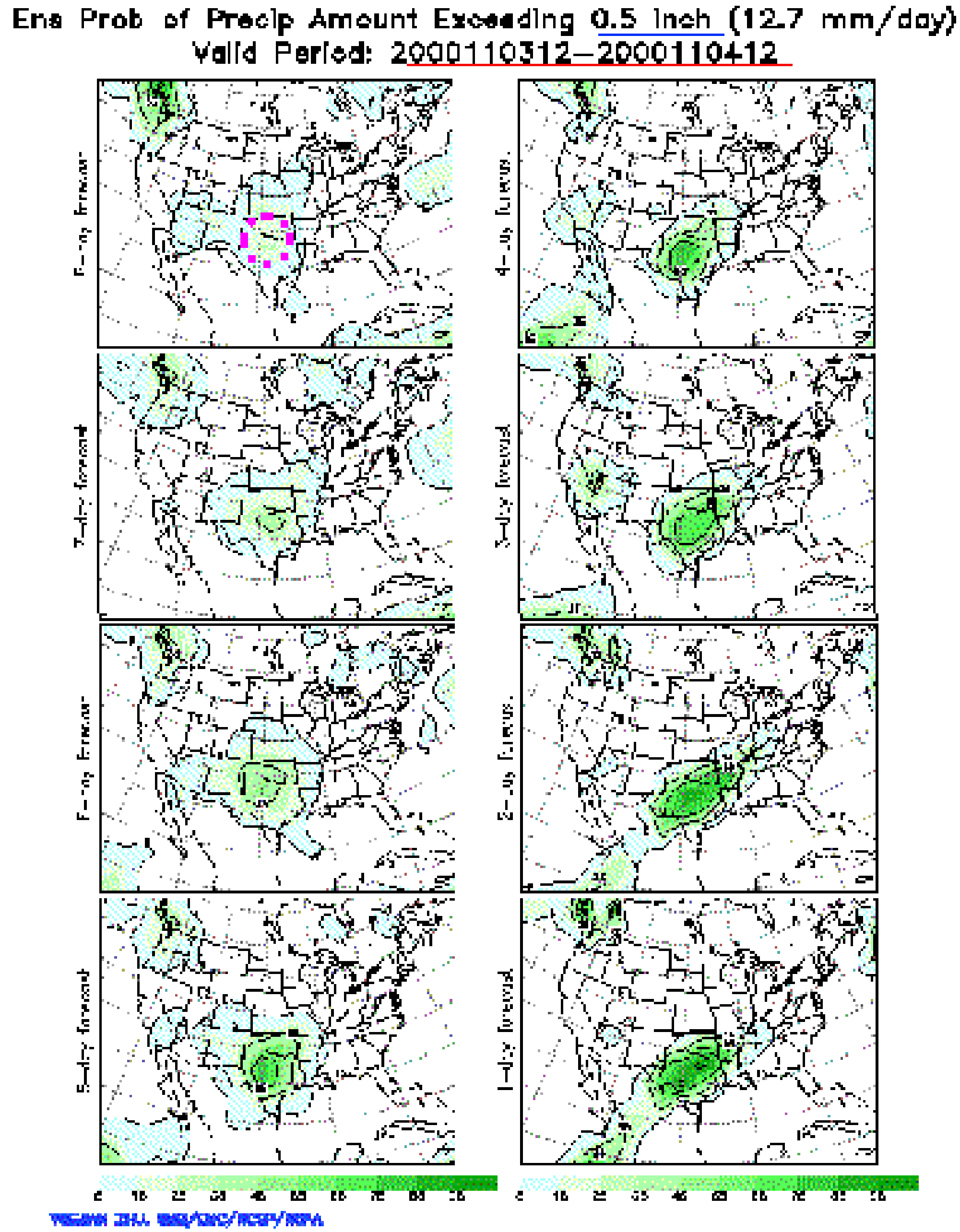
1–2% OF ALL DAYS THE 12–DAY FCST CAN BE MADE WITH MORE CONFIDENCE THAN THE 1–DAY FCST

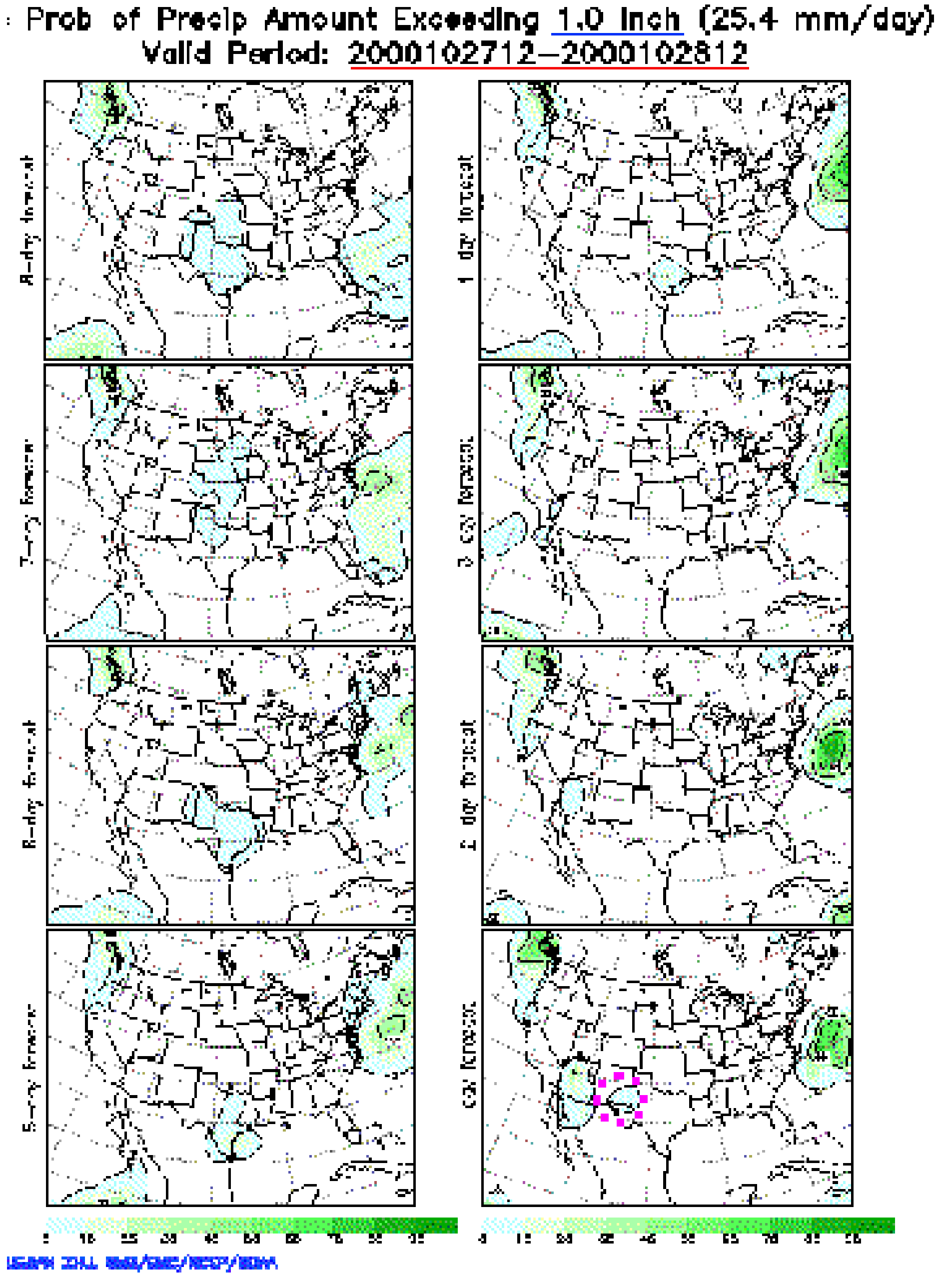
AVERAGE HIT RATE FOR EXTENDED–RANGE FCSTS IS LOW –
VALUE IS IN KNOWING WHEN FCST IS RELIABLE

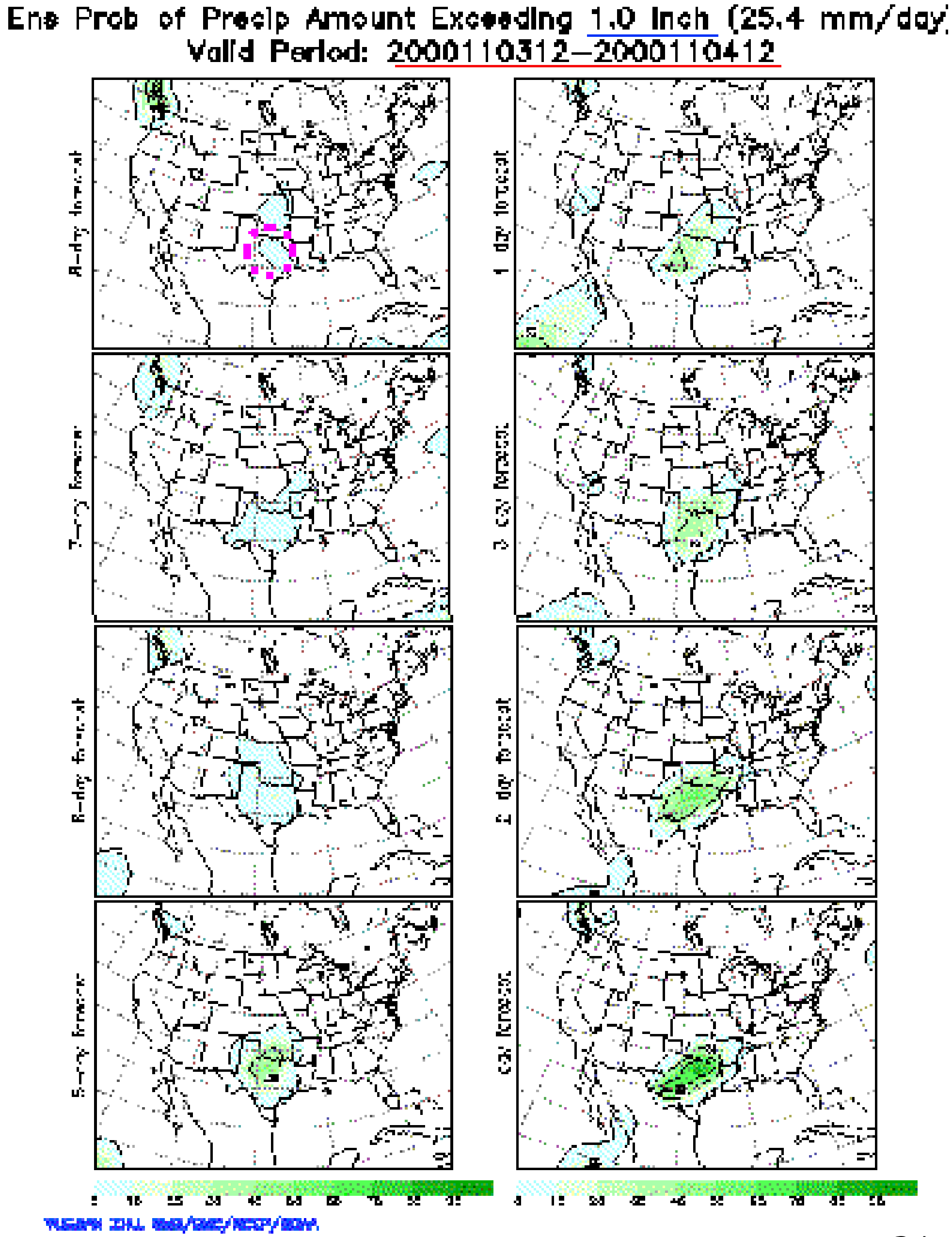

Relative measure of predictability (colors)
for ensemble mean forecast (contours) of 500 hPa height
ini: 2000102700 valid: 2000102800 fcst: 24 hours
Probability (%) 8 16 22 28 35 44 54 62 74 91
Measure of predictability (%) 10 20 30 40 50 60 70 80 90


Relative measure of predictability (colors)
for ensemble mean forecast (contours) of 500 hPa height
ini: 2000102700 valid: 2000110400 fcst: 192 hours
Probability (%) 7 10 10 11 12 12 14 15 22 29
Measure of predictability (%) 10 20 30 40 50 60 70 80 90

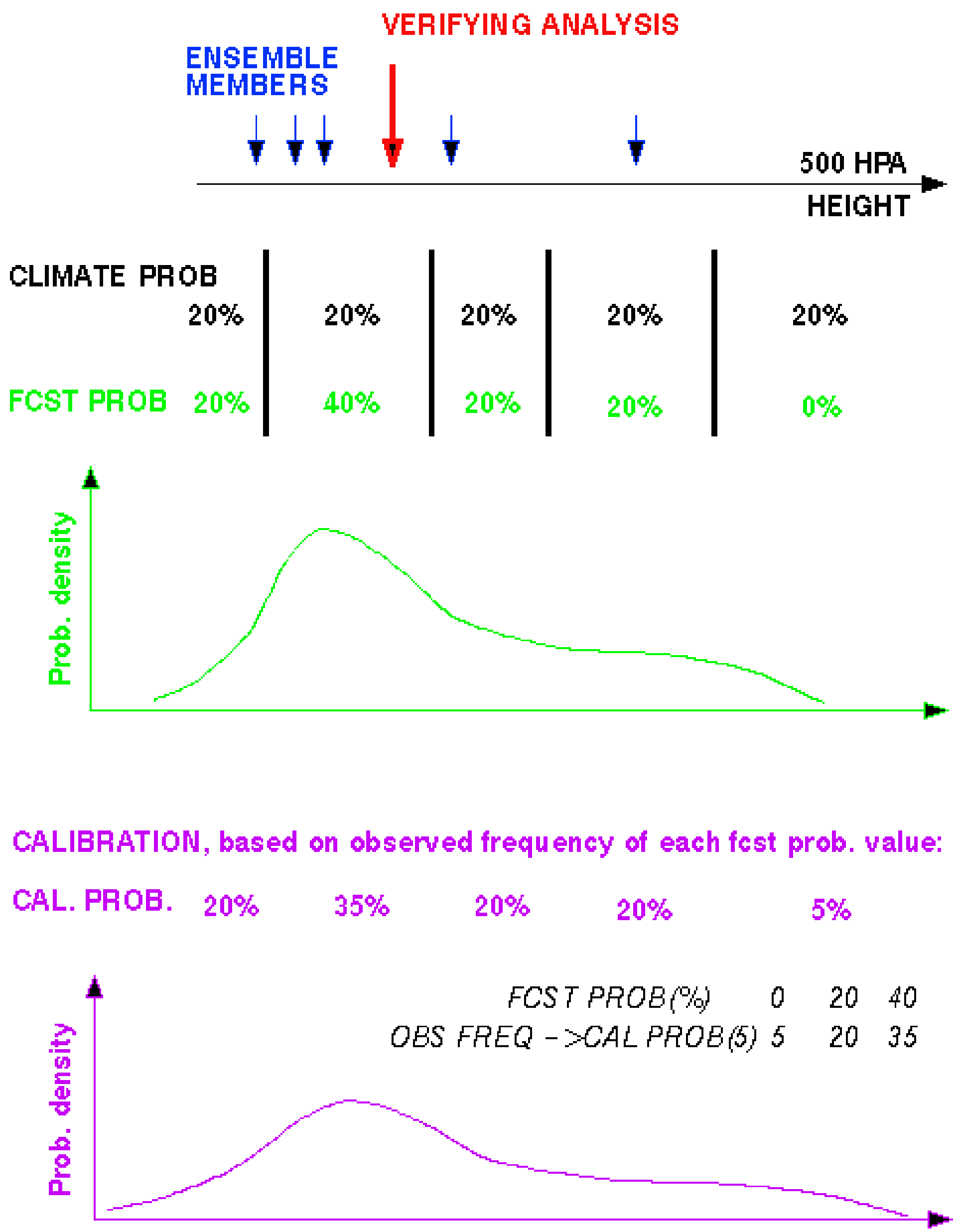

'rob of Precip Amount Exceeding 0.5 inch (12.7 mm/day)
Valid Period: 2000102712–2000102812

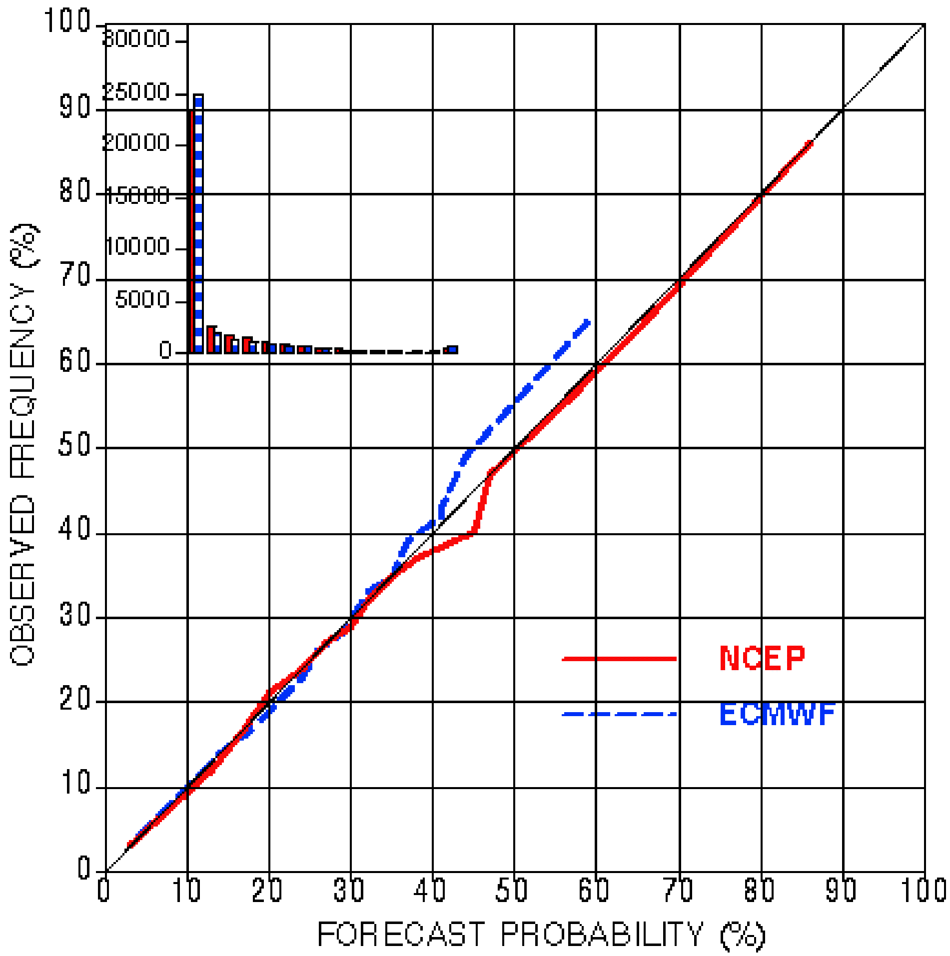

Ens Prob of Precip Amount Exceeding 0.5 inch (12.7 mm/day)
Valid Period: 2000110312–2000110412


Prob of Precip Amount Exceeding 1.0 Inch (25.4 mm/day)
Valid Period: 2000102712–2000102812

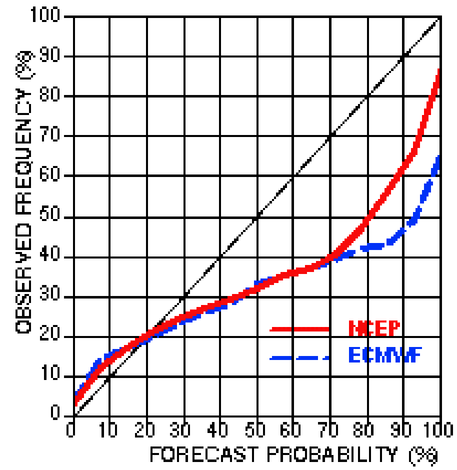

Ens Prob of Precip Amount Exceeding 1.0 Inch (25.4 mm/day
Valid Period: 2000110312–2000110412

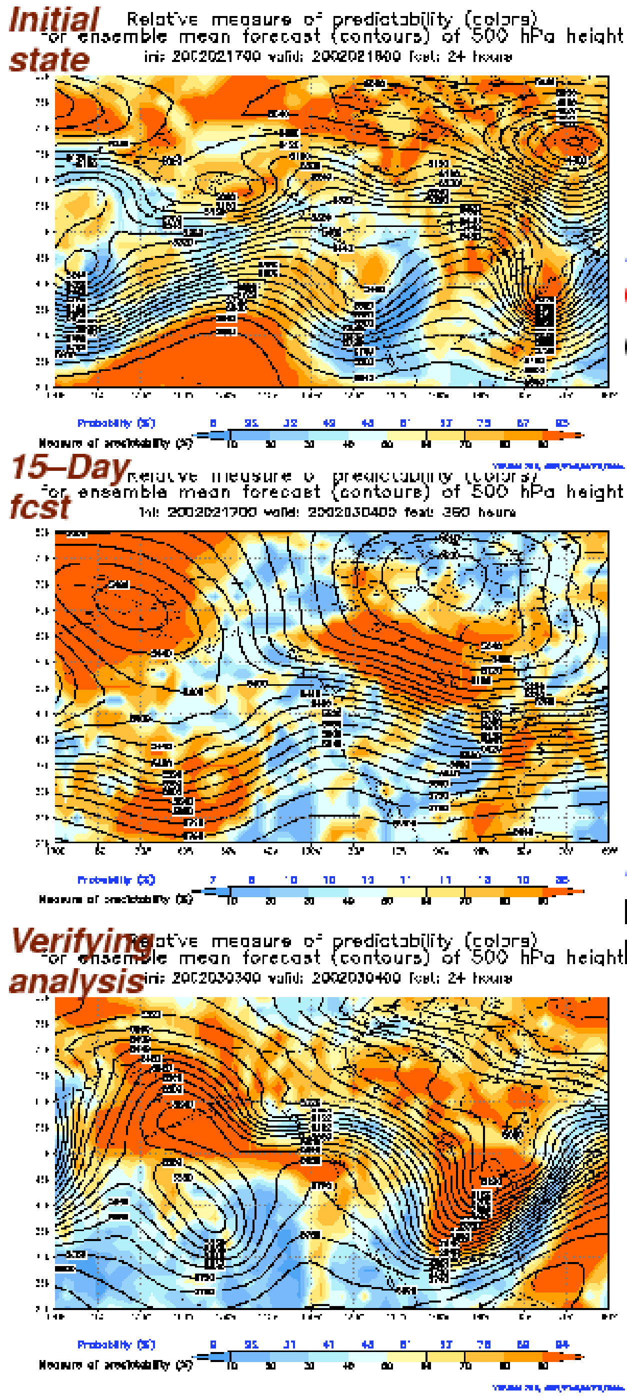
ENSEMBLE BASED PROBABILISTIC FORECASTS AND THEIR VERIFICATION





Reliability diagram for 3-day lead time ensembles for January 1996. Forecast probabilities are based on observed frequencies associated with the same number of ensemble members falling in a particular bin during December 1-20, 1995. The diagram for uncalibrated forecasts is shown on the right.





GREAT CHALLANGE:
PREDICTING REGIME TRANSITION AT EXTENDED RANGE

ENSEMBLE APPROACH:
CAN BE DONE SUCCESSFULLY
(FROM TIME TO TIME)

LOW LEVEL OF SKILL:
USE PROBABILISTIC FCSTS

+/– 24 HR TIMING ACCURACY:
USE DAILY FCST DATA
(NOT TIME MEAN)

FALSE ALARMS:
NEED TO SIMULATE NATURE'S DIVERSITY IN NWP MODEL

MISSED EVENTS:
1) LIMIT OF PREDICTABILITY – NO PROBLEM
2) MODEL FAILURE – NEED TO INCREASE MODEL DIVERSITY (NOT "ACCURACY")

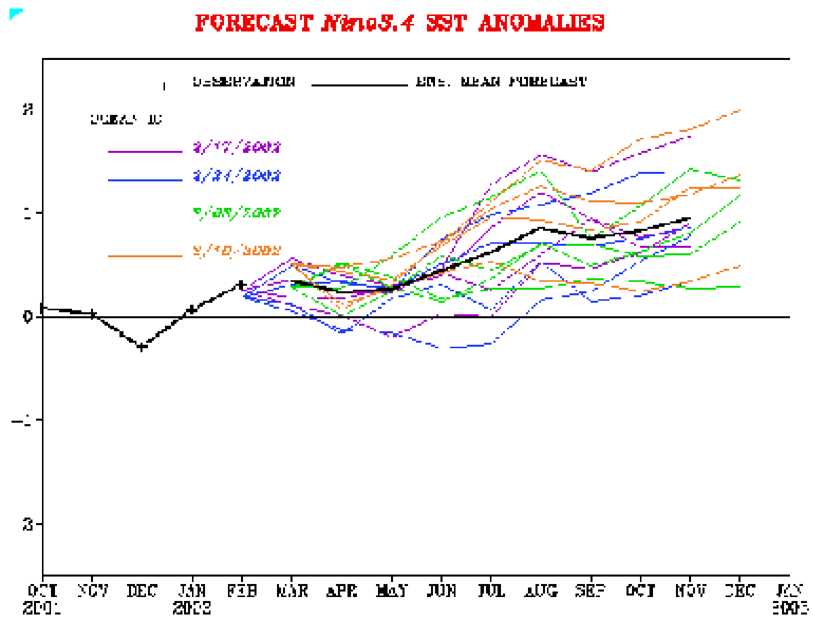
BRIDGING THE GAP BETWEEN WEATHER AND CLIMATE

CURRENT NWS PRACTICE

2) "CLIMATE" ENSEMBLE:

a) 12–months coupled ocean–atm fcsts

b) Average the SST fcsts



c) Run AGCM ensemble forced by average SST fcst

STRENGTH:

Ensemble approach used both for coupled and AGCM model fcsts for enhancing (weak) signal

SHORTCOMINGS:

a) Coupled ensemble (lagged fcst) perturbations not optimal

b) Uncertainty information related to SST fcst is discarded

c) Initial condition information from atmosphere not used

BRIDGING THE GAP BETWEEN WEATHER AND CLIMATE

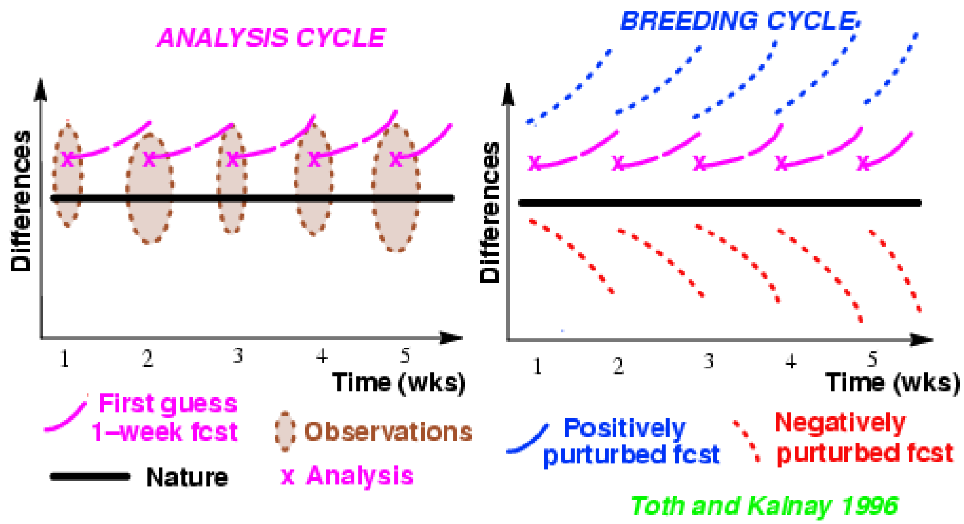
PLANS

3) POSSIBLE FUTURE SYSTEM:

"WEATHER *AND* CLIMATE" ENSEMBLE?

COUPLED MODEL ENSEMBLE –

Use dynamically constructed perturbations



a) *Nonlinear bred perturbations capture dominant ENSO instability*

b) *Initial error present in analysis dominated by same instability*

c) *Symmetrically placed perturbed fcsts provide optimal ensemble*

AGCM ENSEMBLE – PART OF COUPLED SYSTEM?

i) Use ensemble SST fcsts as various boundary scenarios

ii) Single set of AGCM fcsts for all time ranges *(D1–climate)*

ONE–TIER SYSTEM – If possible, with coupled ocean model

NCEP SHORT-RANGE ENSEMBLE FORECAST SYSTEM (SREF)

OPERATIONAL SYSTEM

- 10 Members out to 63 hrs
- 2 Models used:ETA & RSM
- 09 & 21 UTC initialization
- NA domain
- 48 km resolution
- Bred initial perturbations
- Products (on web):
  - Ens. Mean & spread
  - Spaghetti
  - Probabilities
  - Aviation specific
- Ongoing training

PLANS

- *5 more members*
- *More model diversity*
- *4 cycles per day (3&15 UTC)*
- *32 km resolution*
- *New products*
  - *Aviation specific*
  - *AWIPS*
- *Transition to WRF*

PERTURBATION VS. ERROR CORRELATION ANALYSIS (PECA)

METHOD: Compute correlation between ens perturbtns and error in control fcst fc

- Individual members
- Optimal combination of members
- Each ensemble
- Various areas, all lead time

EVALUATION: Large correlation indicates ens captures error in control forecast

- Caveat – errors defined by analysis

RESULTS:

- **Canadian** best on large scales
  - Benefit of model diversity?
- **ECMWF** gains most from combinations
  - Benefit of orthogonalization?
- **NCEP** best on small scale, short term
  - Benefit of breeding (best estimate initial error)?
- PECA increases with lead time
  - Lyapunov convergence
  - Nonlilnear saturation
- Higher values on small scales