Recent developments with the NCEP and North American Ensemble Forecast Systems

Yuejian Zhu and Zoltan Toth

Environmental Modeling Center
NOAA/NWS/NCEP
November 14 2005

Acknowledgements:
Louis Lefaivre MSC
Timothy Marchok GFDL
Richard Wobus, Qingfu Liu, Bo Cui and Steve Lord EMC
David Michaud, Brent Gordon, Maxine Brown NCO
Ed Olenic, Dave Unger CPC
OUTLINE

• Long-term performance stats for NCEP GEFS
• Recent implementation
  – Improvements in
    • Short-range statistics
    • Tropical Storm forecasts
    • Extended-range forecasting
• Planned implementation for NAEFS
• NAEFS & THORPEX (TIGGE)
Monthly Ave. Scores (NH 500hPa Height, 5-day forecasts)

NCEP ensemble mean performance for past 6-year
NCEP ensemble probabilistic performance for past 6-year
Monthly Average for NH 500hPa Height, 5-day forecasts

Ranked Probability Skill Score
RPSS = 0.372

Potential Economic Value for 1:10 cost/loss ratio
EVAL = 0.373
What is THORPEX’s goal for next 10 years?
Recent Implementation
Changes - 1

• Extend T126 portion of forecast after 180 hours (see new configuration)
  – This change is intended to improve ensemble support for 5-10 days and week-2 forecast by providing high resolution (T126) and continue (no resolution change) forecast
  – Results:
    • Increased spread for week-2 forecast
    • Improving probabilistic skill beyond 180 hours
NCEP GLOBAL ENSEMBLE FORECAST SYSTEM

CURRENT CONFIGURATION
MARCH 2003

1. Breeding method
2. 24 hours breeding cycle
3. 4 cycles per day
4. 10 m for each cycle
5. Ens. Ctl at t00z only
6. Total 45 m in 24 hours
7. 4 different resolutions
8. 16-day forecasts

NEW

NCEP GFS and Ensemble Configuration
(Will be implemented by July 2005)
Recent Implementation
Changes - 2

• Initial perturbation (breeding cycle)
  – This change is intended to enable for relocation of perturbed tropical storm. Tuning initial perturbation size is for reducing spread for short-range forecast
  – Results:
    • Decreased spread for short-rang (1-3) forecast
    • Improving forecast skill for first 3 days
    • Improving probabilistic forecast skill for short lead-time
Current breeding cycle

24 hours

T00Z 10m

T06Z 10m

T12Z 10m

T18Z 10m

24hrs | Up to 16-d

Re-scaling

New breeding cycle

6 hours

T00Z 40m

T06Z 40m

T12Z 40m

T18Z 40m

24hrs | Up to 16-d

Re-scaling

Independent vectors

Current breeding cycle

Next T00Z

Up to 16-d

Re-scaling

6hrs

Up to 16-d

Independent vectors

New breeding cycle

Next T00Z

Up to 16-d

Re-scaling

6hrs

Up to 16-d

Independent vectors

Current breeding cycle

New breeding cycle

Re-scaling

6hrs

Next T00Z

Up to 16-d

Independent vectors
Recent Implementation
Changes - 3

• Relocation of perturbed tropical storm
  – This change is intended to reduce track forecast error and uncertainty for short lead-time (1-3 days)
  – Results:
    • Reducing mean track errors by 10% for 12-48 hours
    • Reducing the ensemble track spread, that was too large, for short lead-time
    • Improving track forecast skill
GFS TS relocation

Use GFS Track information

Relocate TS to Observed position

GDAS (SANL)

FCST

Ensemble TS relocation

6hrs fcst

Use ens. Track information

Use GFS Track information

To separate into env. Flow (EF) And “storm perturbation” (SP)

Ens. Rescaling For EF (p+n)

Ens. Rescaling For SP (p+n)

Combined

FCST
Increasing spread for week-two forecast
Ensemble mean skill for D+8 and week-two forecast

CPC’s evaluation – by J. K. Schemm
Northern Hemisphere 500hPa height probabilistic verification:

ROC (left) and EV (right)

Improvement for short-lead time due to 6-hr breeding cycle
Improvement for extended forecast due to increased resolution
SH RMS and spread Improved outlier stats

SH ROC Improved skill for short & extended-range forecasts

SH RPSS
Hurricane Track Plots (case 1)

Frances (08/28)

With relocation

Without relocation

Large initial spread

Reduced initial spread
Hurricane Track Plots (case 2)

Ivan (09/14)

Without relocation

With relocation
Hurricane Tracks Plots (case 3)

Jeanne (09/14)

Without relocation

With relocation
Hurricane Tracks Plots (case 4)

Karl (09/18)

Without relocation

With relocation
Track error and spread
2004 Atlantic Basin (8/23-10/1)

From Timothy Marchok (GFDL)

Reduced mean track error and spread
Hurricane track errors
2 basins (Atlantic and e-Pacific)

Period: 20040824-20040930 (53-103 cases)
FORECAST PROBABILITY VS. OBSERVED FREQUENCY

When strike probability forecast is eg. 30%, storm observed in 30% of all cases

Ensemble TC track forecast reliability
2004 Atlantic Basin (8/23 – 10/1)
Hurricane track errors
Atlantic basins (up to October 2005)

Track errors (miles)
Period: for whole hurricane season of 2005 (up to October 27)
Note: ensemble TS relocation was implemented by Aug. 17th 2005

Percentage improvement to GFS

GFS  ENS

0 24 48 96

ENS better

0 24 48 96

0 50 100 150 200 250 300 350
ENSEMBLE FORECASTING - QUANTIFYING UNCERTAINTY

STRIKE PROBABILITY

At any point, how many members of ensemble had a storm within 65 nm radius

NCEP Ensemble-based probability (%) of storm 24L passing within 65 nm during the 180h period beginning 2005101612
NCEP Ensemble Mean track (solid line, marker every 24h)

- Forecast track
- Observed track

Strike probability =>
HURRICANE BETA
STRIKE PROBABILITY

Probability of storm within 65 nm vicinity of any point on map

NCEP Ensemble–based probability (%) of storm 26L passing within 65 nm during the 180h period beginning 2005102706
NCEP Ensemble Mean track (solid line, marker every 24h)

- Forecast track
- Observed track

Strike probability =>
NAEFS

- PARTICIPANTS
- PROJECT DESCRIPTION
- TIMELINE
- IMPLEMENTATION SCHEDULE
- CONCEPT OF OPERATIONS
- NAEFS & THORPEX
- BASIC PRODUCTS
- END PRODUCTS
- DETAILS – RESOURCE ISSUES
- FUTURE EXPANSION
- NEW NWP PARADIGM
- Visit: http://wwwt.emc.ncep.noaa.gov/gmb/ens/NAEFS.htm
CONCEPT OF OPERATIONS

1. Exchange ~50 selected variables
   - Use GRIB2 to reduce volume of data

2. Generate basic products using same algorithms/codes
   - Reduce systematic error
     - Bias estimation
   - Combine two ensembles
     - Determine weights
   - Express forecast in terms of climatological anomalies
     - Prepare & compare forecast with reanalysis climate distribution

3. Generate center-specific end products

4. Evaluate & provide feedback for improvements
   - Verification using same algorithms
   - User feedback

2. MSC-NCEP basic production suite
   - Same algorithms/codes used at both centers
     - Duplicate procedures provide full backup in case of problems at either end
     - If one component of ensemble missing, products based on rest of ensemble
   - Basis for different sets of center-specific end products
     - Ensures consistency between end products even if their format is different
   - All basic products to be made available via ftp to user and research community
NAEFS & THORPEX

• Expands international collaboration
  - Mexico joined in November 2004
  - UK Met Office to join in 2006

• Provides framework for transitioning research into operations
  - Prototype for ensemble component of THORPEX legacy forecast system:
    *Global Interactive Forecast System (GIFS)*
**LIST OF VARIABLES IDENTIFIED FOR ENSEMBLE EXCHANGE BETWEEN MSC - NCEP**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>CMC</th>
<th>NCEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ensemble</td>
<td>8 SEF, 8 GEM</td>
<td>10 paired</td>
</tr>
<tr>
<td>Grid</td>
<td>2.5x2.5 deg (144x73) &amp; 1.2x1.2 deg (300x151)</td>
<td>2.5x2.5 deg (144x73) &amp; 1.0x1.0 deg (360x181)</td>
</tr>
<tr>
<td>Domain</td>
<td>Global</td>
<td>Global</td>
</tr>
<tr>
<td>Format</td>
<td>WMO GRIB Format</td>
<td>WMO GRIB Format</td>
</tr>
<tr>
<td>Hours</td>
<td>0, 12, 24, 36, 48, …, 216, 228, 240</td>
<td>0, 6, 12, 18, 24, …, 360, 366, 372, 378, 384</td>
</tr>
<tr>
<td>GZ</td>
<td>200, 250, 500, 700, 850, 925, 1000</td>
<td>200, 250, 500, 700, 850, 925, 1000</td>
</tr>
<tr>
<td>TT</td>
<td>200, 250, 500, 700, 850, 925, 1000</td>
<td>200, 250, 500, 700, 850, 925, 1000</td>
</tr>
<tr>
<td>E</td>
<td>Tdd at 200, 250, 500, 700, 850, 925, 1000</td>
<td>RH at 200, 250, 500, 700, 850, 925, 1000</td>
</tr>
<tr>
<td>U, V</td>
<td>200, 250, 500, 700, 850, 925, 1000</td>
<td>200, 250, 500, 700, 850, 925, 1000</td>
</tr>
<tr>
<td>TT Sfc</td>
<td>12000, redefined in GRIB file as 2m AGL</td>
<td>2m</td>
</tr>
<tr>
<td>U, V Sfc</td>
<td>Redefined in GRIB file as 10m AGL</td>
<td>10m</td>
</tr>
<tr>
<td>ES</td>
<td>Tdd at 12000, redefined in GRIB file as 2m AGL</td>
<td>RH at 2M</td>
</tr>
<tr>
<td>MSLP</td>
<td>(PN) level 0</td>
<td>PRMSL</td>
</tr>
<tr>
<td>PR (total precip)</td>
<td>Level 0, i.e. at surface</td>
<td>Level 0, i.e. at surface</td>
</tr>
<tr>
<td>NT (total cloud cover)</td>
<td>Level 0</td>
<td>Column</td>
</tr>
<tr>
<td>IH (total precipitable cover)</td>
<td>Level 0</td>
<td>Column</td>
</tr>
<tr>
<td>Sfc Pres</td>
<td>(SEF) (P0) level 0 at surface</td>
<td>Sfc Pressure</td>
</tr>
<tr>
<td>Model Topography</td>
<td>Model Topography</td>
<td>Model Topography at t=0 and t=192</td>
</tr>
<tr>
<td>CAPE</td>
<td>Most unstable layer</td>
<td>Most unstable layer</td>
</tr>
<tr>
<td>Precip Type</td>
<td>4 accumulations processed into 4 bitmaps</td>
<td>4 bitmap variables for 4 types</td>
</tr>
<tr>
<td>Tmax</td>
<td>2m derived from hourly</td>
<td>2m</td>
</tr>
<tr>
<td>Tmin</td>
<td>2m derived from hourly</td>
<td>2m</td>
</tr>
<tr>
<td>WAM</td>
<td>Later</td>
<td>Later</td>
</tr>
</tbody>
</table>

Black: data presently exchanged  
Blue: data exchanged & processed by NCEP June 2004  
Red: data added in September 2004  
Green: data to be exchanged later

R. Wobus  
R. Hogue
Basic Products
Post-Processing

• Bias corrected forecasts
  – Consider 35 variables in the first phase
• Statistical weights
  – Consider 35 variables in the first phase
• Anomaly forecasts
  – Consider 19 variables in the first phase
• GRIB2
• NAWIPS grids and graphics
• NDGD grids
## List of Variables for Bias Correction/weights for CMC & NCEP Ensembles

<table>
<thead>
<tr>
<th>Ensemble</th>
<th>CMC &amp; NCEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRID</td>
<td>CMC (8 SEF, 8 GEM), NCEP (GFS)</td>
</tr>
<tr>
<td>DOMAIN</td>
<td>1x1 deg (360x180 lat-lon)</td>
</tr>
<tr>
<td>FORMAT</td>
<td>Global</td>
</tr>
<tr>
<td>HOURS</td>
<td>WMO Grib Format</td>
</tr>
<tr>
<td>GZ</td>
<td>6 hourly out of 384 hours</td>
</tr>
<tr>
<td>TT</td>
<td>(current 240 hours for CMC Ensemble)</td>
</tr>
<tr>
<td>U, V</td>
<td>200, 250, 500, 700, 850, 925, 1000</td>
</tr>
<tr>
<td>TT</td>
<td>200, 250, 500, 700, 850, 925, 1000</td>
</tr>
<tr>
<td>U, V</td>
<td>200, 250, 500, 700, 850, 925, 1000</td>
</tr>
<tr>
<td>MSLP</td>
<td>2m</td>
</tr>
<tr>
<td>Sfc Pres</td>
<td>10m</td>
</tr>
<tr>
<td>Tmax</td>
<td>Sea Level Pressure</td>
</tr>
<tr>
<td>Tmin</td>
<td>Surface Pressure</td>
</tr>
<tr>
<td></td>
<td>2m</td>
</tr>
<tr>
<td></td>
<td>2m</td>
</tr>
</tbody>
</table>
**Review of Recent Results**

- **First moment correction:**
  - Previous results: kept reinitializing the prior, based on 40-day flat average difference.
  - Current system: keeps cycling the bias estimate after initializing the prior, which starts from July 1, 2003. Choose decaying weight 10%, 5%, 2%, 1%, 0.5% and 0.25%, respectively, and apply on 500 mb height of NCEP & CMC ensemble.
    - Northern and Southern Hemisphere: the smaller weight is better for longer lead time, and larger weight is better for shorter lead time.
    - Tropical region: 2% is the best one among the six weight factors.

- **Bias correct CMC member individually & bias correct CMC member in 2 groups (8 SEF member & 8 GEM member)** due to CMC multi-model ensemble and each model & member has its own physics parameterization.
  - Applying the bias correction scheme on each member is the better approach though the differences are small between the two methods.

- **Combined ensemble, use equal weight for all members (5 NCEP & 5 CMC member)**.
Ongoing Development & Testing Plans for Implementation

- **Bias correction**
  - First moment correction
    - choose a fixed weight factor (2% as a default), or vary it as a function of lead time and location (how to determine variations?)
    - apply bias correction scheme
      - 35 variables (NCEP & CMC)
      - on 1x1 degree ensemble data (NCEP & CMC)
      - on 00Z and 12Z (NCEP & CMC, 06 & 18Z for NCEP)
  - Second moment correction
    - may not be included in next spring operational implementation

- **Weighting**
  1. BMA method: only tested for surface temperature
  2. Use frequency of “best member of ensemble” statistics
NCEP RPSS: 500mb Height, Northern Hemisphere
2004 Annual Mean

Northern Hemisphere 500hPa Height
Ranked Probability Skill Scores (RPSS)
Average For 20040301 – 20050228

http://www.emc.ncep.noaa.gov/gmb/wx20cb/Bias_Correction_Algorithm/1st_2nd_Moments/Training_1month/Plot_Comb_Post/z500_2004_ncep_annual/
NCEP RMS: 500mb Height, Northern Hemisphere
2004 Annual Mean

NH 500hPa Height
Average For 20040301 - 20050228

RMS errors vs. Forecast days graph
NCEP RPSS: 500mb Height, Southern Hemisphere
2004 Annual Mean

Southern Hemisphere 500hPa Height
Ranked Probability Skill Scores (RPSS)
Average For 20040301 – 20050228

Skill Scores vs Forecast days
Climate anomalies

Express bias-corrected forecasts (each member) in terms of climate percentile

- Forecasts bias corrected wrt NCEP & CMC oper. analysis
  - 1.0*1.0 (lat/lon) grid
- Climate based on NCEP/NCAR reanalysis data
  - 4 cycles (00UTC, 06UTC, 12UTC and 18UTC) per day
  - 40 years (Jan. 1st 1959 – Dec. 31th 1998)
  - 2.5*2.5 (lat/lon) grid
- Need to consider the systematic difference between reanalysis and oper. analysis (NCEP & CMC respectively)
- Variables (possible to add more)
  - Height: 1000hPa, 700hPa, 500hPa, 250hPa
  - Temperature: 2m, 850hPa, 500hPa, 250hPa
  - Wind: 10m, 850hPa, 500hPa, 250hPa
  - PRMSL, max/min temperature
Climate anomalies

**PROCEDURE**

- Determine climatological distribution for each day using reanalysis data
  - Use first few harmonics to describe annual variations
  - Compute all stats for 4 times per day
  - Estimate climate mean (first moment)
  - Estimate distribution around mean
  - Archive data to be used on daily basis

- Determine systematic difference between reanalysis and operational analysis fields
  - Use standard NAEFS “bias estimation” method

- Adjust bias corrected NAEFS forecasts by systematic difference between reanalysis & oper. analysis

- Compare bias corrected & adjusted NAEFS forecasts to reanalysis distribution
  - Express each forecast as percentile of climate distribution
Estimating the climate mean

• To consider monthly mean (tested)
  – Monthly mean (large data samples – 1240)
    • Interpolate to daily (shifted from season)

• To consider daily mean (tested)
  – 5-day running mean for daily climatology
    • Data samples – 200
  – 5-day center weighted mean for monthly climatology
    • Data samples – 200
      – \((d-2)*0.12+(d-1)*0.22+d*0.32+(d+1)*0.22+(d+2)*0.12\)

• To consider annual cycle (tested)
  – Fits the first 1-4 Fourier annual modes to daily data to obtain annual cycle.
Higher moments (estimation)
- work on the anomalies from mean

• To consider monthly (tested):
  – Data size of 40 (year) * 31 (dom for Jan) = 1240
  – Fitting distributions (three parameters)
    • Gamma, Pearson type-III, GE3 (generalized extreme-value)

• Compute a smooth standard deviation (tested)
  – Based on annual cycle

• Discussions and questions
2 meter temperature for January (30N, 60E)

- Monthly mean
- 5-day weighted mean

GEV
10 meter U–wind for January (30N, 60E)

- Monthly mean
- 5-day weighted mean

GEV

FIT–1
FIT–2
Data
Monthly mean

5-day weighted mean
ENSEMBLE 10-, 50- (MEDIAN) & 90-PERCENTILE FORECAST VALUES (BLACK CONTOURS) AND CORRESPONDING CLIMATE PERCENTILES (SHADES OF COLOR).

2–meter temperature 5–day forecast (valid at 06/15/2005)

Example of probabilistic forecast in terms of climatology.
Based on raw forecasts, no climate and current analysis correction
NDGD FORECAST UNCERTAINTY RECOMMENDATION

- Provide 3 ensemble-based guidance products for inclusion in NDGD:
  - 10, 50, and 90 percentile values
    - SREF guidance out to day 3
    - NAEFS guidance out to 16 days
  - Use NDGD grid (5x5 km), with GRIB2 packing, minimal space overhead

- Approach
  - Solicit comments on specific proposal from NCEP Service Centers and regions/field
  - Use NAWIPS software (available soon?) to generate products
    - Work with NAWIPS group to provide algorithm:
      - Simple counting of members with linear interpolation now
      - Gaussian Kernel method in later implementation
  - Factor of 3 increase in disc space
    - D. Ruth positively inclined (WG member at NDFD Workshop)
NDGD FORECAST UNCERTAINTY - DOWNSCALING

• Ensemble uncertainty information
  – Sent on NDGD grid for convenience (if no big overhead)
  – Valid on model grids (32km for regional, 110 km for global ensemble)
  – How to bridge gap between model and NDGD grids?

• Anomaly uncertainty information – proposed methodology
  – Establish reanalysis climatology
    • In progress for global (NAEFS), methods can be transferred to regional reanalysis
  – Bias correct ensemble forecasts (wrt operational analysis)
  – Take 10-50-90 percentile values from bias corrected ensemble
  – (For establishing anomaly forecasts, adjust 10-50-90 percentile values to look like re-analysis)
  – Check climatological percentile corresponding to 10-50-90 forecast percentiles

• Provide climatological percentiles corresponding to 10-50-90 percentile forecast values as second set of guidance products
**ENSEMBLE-BASED PRODUCTS FOR NDGD**

- **National Digital Forecast Database (NDFD)**
  - Official NWS forecast, prepared by WFO offices (central guidance, coordination)
  - 5x5 (2.5x2.5) km grid, out to 7 days
  - Selected parameters (~15)
  - Available in digital format, query tools, etc
  - No (minimal) provision for information on forecast uncertainty
    - Recommendations from an NDFD workshop, Salt Lake City, 2003
    - Interactive Forecast Preparation System (IFPS) offers tools to work with NDFD grids (forecasters can manipulate gridded data, etc)

- **National Digital Guidance Database (NDGD)**
  - For posting numerical guidance products same way as NDFD
  - New system, possibility to complement NDFD with forecast uncertainty info
    - Based on global (NAEFS) and regional ensemble forecasts

- **What forecast uncertainty info to post in NDGD?**
NAWIPS grids data, graphic and GIF images

• Mean of selected members
  – Z500, z700, z850

• Spread of selected members
  – U10m, V10m

• Exceeding probabilities for selected threshold values
  – 10m wind speed: thresh 20, 34, 50, and 64 kts
  – Significant wave height at various values

• Spaghetti plots
  – Height: 200hPa, 300hPa and 500hPa
  – Psml
  – T2m: 0c isotherm
  – QPF: 0.01”, 0.25”, 0.5”, 1.0”, 2.0”, 3.0” and 4.0” (for 6-h and 24-h)
  – Snow: 1”, 2”, 4”, 6”, 8”, 12”, 18”, 24” (for 6-h and 24-h)
  – Freezing rain: 0.01”, 0.1”, 0.25”, 0.5”, 1.0” (for 6-h and 24-h)
THANKS!!!
Products (plan)

- Based on 4 different considerations
- Assuming the normal distributions of the 40 years climate data
  - PDF will be presented by first two moments (mean and standard deviation)
- Considering the differences between reanalysis and current GDAS
- Using bias corrected forecasts
- To calculate climate anomaly:
  - For 1x1 degree grid point globally.
  - For all 19 variables.
  - For each ensemble member.
  - In percentile (0-100%, 50%=normal).
Discussion

• How many modes we need to consider?
  – In general, more modes will be better
  – First two modes are enough for the heights
  – Surface variables and winds are challenging

• Are all variables normally distributed?
  – Depends on variables and geographical locations (?)
  – Most of them are quasi-normally distributed
    • Examples of 2-meter temperature and 10-meter u
    • Monthly distribution of 500hPa height has a little seasonal tilt

• Examples of time series for daily mean and standard deviation
  – Two physical locations (near Washington DC and Ottawa)
  – Are these plots enough to evaluate methods?
Ensemble Mean and Spread from Wave Ensemble
- For Significant Wave Height

2005/11/11 00z, 084 fcst_hr

Ensemble Mean (contour, m) & Spread of Hs Valid 2005/11/14 12z

Spread (m)

0 0.3 0.6 0.9 1.2 1.5 1.8 2.1 2.4