STATISTICAL POST-PROCESSING FOR HURRICANE ENSEMBLE FORECASTING

Zoltan Toth, Roman Krzysztofowicz[1], and Tom Hamill

ESRL/OAR/NOAA

[1] Univeristy of Virginia

Acknowledgements:
Paul Schultz, Huiling Yuan, Isidora Jankov, Brad Beechler, Yuejian Zhu, Andre Methot, Kathy Gilbert et al.

HFIP / THORPEX Ensemble Product Development Workshop
April 20-21, 2010, Boulder, CO

OUTLINE / SUMMARY

- User requirements
  - Joint probabilities / co-variances – use ensembles
  - Perfect reliability & highest possible resolution

- Current NWP systems
  - Capabilities
    - Forecast skill out to ~20 days
  - Limitations
    - Lead-time dependent biases in mean, spread, etc
    - Coarse resolution, limited set of variables
    - Time consuming execution – need to fuse with info form latest observations

- Statistical post-processing
  - Positional and amplitude corrections
    - Use generic methods when possible
    - Expand to enhance TC-specific performance

WHY DO WE NEED PROBABILISTIC FORECASTS?

An example - House in hurricane prone area

- Weather situation – Heightened threat of hit by a TC
  - Climatological probability of hit by major hurricane on any day in season
    - 0.01%
  - Forecast probability of hit at 7 day lead
    - 10% (1000 times above climatological chance)
- User situation – Owner set to leave on cruise next day
  - Cancel trip at loss of $1k so if needed, can protect house
  - Protect house at cost of $3k then leave
  - Go on trip without protecting house and risk potential damage up to 100k

WHY DO WE NEED PROBABILISTIC FORECASTS? 2

Another example - Air travel (after Ken Mylne)

- Air travel safety situation
  - "Climatological" chance of aircraft crash
    - 0.00025%
  - Pilot says on today's flight, chance of crash due to hurricane caused storms
    - 0.25% (1000 times climatological chance)
- User situation – Passenger must reach destination today
  - Take small chance of an accident – Most likely (99.75%) gets there but risks losing life
  - Cancel trip and hence miss very important appointment to avoid accident for sure

Moral

- Humans make everyday decisions in uncertain situations all the time
  - We are capable of and have strong need for assessing and using information with uncertainty (evolution)

TRADITIONAL WAY OF GENERATING PROBABILISTIC FORECASTS

- Based on single NWP forecast
  - Statistical analysis of errors in forecast
  - E.g., single variable probabilistic forecasts from MOS – PoP, etc

WHY DO WE NEED ENSEMBLE FORECASTS?

- Atmosphere is nonlinear system
  - Error evolution can be highly non-linear
  - Probabilistic forecast based on single integration can have serious limitations

- Users often affected by numerous weather variables
  - Need to know co-variance across time, space, variables
    - E.g., probability of wind above X, temp below T, and precip above P
  - Traditional statistical methods cannot practically deliver such info

- Real life user situations can be complicated & highly nonlinear
  - Users have myriad factors to consider beside weather
  - Decision making via study of hypothetical scenarios is imperative, e.g.
    - Utility company designs optimal strategy before hurricane hits
      - Must have model of operations, incorporating effect of weather
    - “Downstream” applications
      - E.g., Flood forecasting via hydrologic ensemble forced by atmospheric ensemble

AVIATION EXAMPLE

- Recovery of a carrier from weather related disruptions
  - Operational decisions depend on multitude of factors
    - Based on United / Hemispheres March 2009 article, p. 11-12

- Factors affecting operations
  - *Weather* – multiple parameters
    - *Over large region / CONUS during coming few days*
  - Federal regulations / aircraft limitations
    - Dispatchers / load planners
  - Aircraft availability
    - Scheduling / flight planning
  - Maintenance
    - Pre-location of spare parts & other assets where needed
  - Reservations
    - Rebooking of passengers
  - Customer service
    - Compensation of severely affected customers

- How to design economically most viable operations?
  - Given goals / requirements / metrics / constraints

SELECTION OF OPTIMAL USER PROCEDURES

- Generate ensemble weather scenarios e_i, i = 1, n
- Assume weather is e_i, define optimal operation procedures o_i
- Assess cost/loss c_{ij} using o_i over all weather scenarios e_j
- Select o_i with minimum expected (mean) cost/loss $\underline{c_i}$ over $e_1,\ldots e_n$ as optimum operation

COST/LOSS c_{ij} GIVEN e_j WEATHER & o_i OPERATIONS		ENSEMBLE SCENARIOS				EXPECTED COST
		e_1	e_2	.	e_n	
OPERATION PROCEDURES	o_1	c_{11}	c_{12}	.	c_n	$\underline{\mathbf{c}}_{\underline{\mathbf{1}}}$
	o_2	c_{21}	c_{22}	.	c_{2n}	$\underline{\mathbf{c}}_{\underline{\mathbf{2}}}$
	.	.	.	.	.	.
	o_n	c_{n1}	c_{n2}	.	c_{nn}	$\underline{\mathbf{c}}_{\underline{\mathbf{n}}}$

USER REQUIREMENTS FOR FORMAT

- Ensemble of forecasts for
  - Complex decision making situations
  - “Downstream” applications

- Probabilistic forecasts
  - For simpler applications
  - Can be derived from / must be consistent with ensembles

Consider ensemble as basic format from which all other forecast info can be derived

USER REQUIREMENTS FOR QUALITY

- Statistical resolution ("predictive skill")
  - Seek highest possible skill in ensemble of forecasts
  - Need to extract and fuse all predictive information
    - Ensembles, high resolution unperturbed forecasts, observations, etc
  - Smallest possible spread that allows reliability

- Statistical reliability
  - Need to make ensemble members statistically indistinguishable from reality
    - Correct systematic errors (first moment correction)
    - Assess error statistics (higher moment corrections)
    - Use climatology as background information

REALITY?

- Any forecast (single or ensemble member)
  - Has lead-time dependent systematic errors, e.g.,
    - Central pressure
    - Wind speed
    - Precipitation
- Collection of forecasts (ensembles by opportunity or design)
  - Spread and higher moments have lead-time dependent systematic errors
- Observations taken after NWP initial time
    - May offer valuable additional forecast information
      - E.g., “Interpolated” forecast storm positions
- NWP forecast variables different from user relevant variables
  - Information needed
    - At finer resolution
    - For additional variable

Need for statistical post-processing

STATISTICAL POST-PROCESSING

- Problem
  - Relate coarse resolution biased forecasts to user relevant fine resolution information
- Tasks broken up to facilitate collaboration / transition to operations
  - Bias correct coarse resolution ensemble grid wrt NWP analysis
    - Cheap
    - Sample of forecasts / hind-casts needed
  - Merge various guidance
    - Fuse all predictive info into “unified ensemble”
  - Create observationally based fine resolution analysis
    - Estimate of truth
  - Downscale bias-corrected ensemble forecast
    - Relate coarse resolution NWP and fine resolution observationally based analyses
      - Perfect prog approach - No need for hind-casts
  - Derive additional variables – AIVs
    - Based on bias corrected & downscaled ensemble
- Outcome
  - Skillful and statistically reliable ensemble of hurricane impact variables on fine grid

PRACTICAL APPROACHES

- Can general stat. post-processing be used for TC forecasting?
  - Consider TC specific extensions only when needed
    - Use tools that can serve multiple applications if possible

- Consider NAEFS system as baseline
  - Current state
    - NCEP, Canadian plus FNMOC ensembles & hires forecasts
    - Wind speed, etc on 5 (2.5) km NDFD grid
  - Next phase – Bayesian Processor
    - “Homogenization” – bias removal
    - Fusing – Combine info from ensemble/hires fcsts, obs, climate into posterior cdf
    - Adjust ensemble members to be consistent with posterior cdf
    - Downscale to fine resolution grid / Derive user relevant variables

- NUOPC interagency ensemble development linked with HFIP
  - TC specific enhancements
  - International link to GIFS-TIGGE community
    - Contributions to / benefits from GIFS development

TC SPECIFIC ENHANCEMENTS

- How to address systematic positional errors?
  - Generalize concept of “track errors”
    - Define positional error for entire grid (not only TC)
    - Link with verification WG

- Decompose errors into positional and amplitude components
  - Test use of “Field adjustment” approach of S. Ravela
  - Assess random and systematic errors for both components
  - Correct systematic positional error before Bayesian amplitude corrections
    - Include positional error correction for second & higher moments
      - I.e., adjust forecast spread for tracks

TC SPECIFIC ENHANCEMENTS - 2

- How to improve intensity forecasts?
  - Consider regime dependent processing
    - Natural extension of Bayesian framework
    - Consider testing use of additional predictors

- Issue of sample size required for statistical post-processing
  - TC specific and regime dependent processing may require large sample
    - Analysis of cost of hind-cast generation vs. benefit from larger sample

TC SPECIFIC ENHANCEMENTS

- How to address systematic positional errors?
  - Generalize concept of “track errors”
    - Define positional error for entire grid (not only TC)
    - Link with verification WG
- Decompose errors into positional and amplitude components
  - Test use of “Field adjustment” approach of S. Ravela
  - Assess random and systematic errors for both components
  - Correct systematic positional error before Bayesian amplitude corrections
    - Include positional error correction for second & higher moments
      - I.e., adjust forecast spread for tracks
- How to improve intensity forecasts?
  - Consider regime dependent processing
    - Natural extension of Bayesian framework
    - Consider testing use of additional predictors
- Issue of sample size required for statistical post-processing
  - TC specific and regime dependent processing may require large sample
    - Analysis of cost of hind-cast generation vs. benefit from larger sample

OUTLINE / SUMMARY

- User requirements
  - Joint probabilities / co-variances – use ensembles
  - Perfect reliability & highest possible resolution

- Current NWP systems
  - Capabilities
    - Forecast skill out to ~20 days
  - Limitations
    - Lead-time dependent biases in mean, spread, etc
    - Coarse resolution, limited set of variables
    - Time consuming execution – need to fuse with info form latest observations

- Statistical post-processing
  - Positional and amplitude corrections
    - Use generic methods when possible
    - Expand to enhance TC-specific performance

BACKGROUND

4-D DATACUBE

- **Ensemble data depository / access**
  - Create NOAA digital forecast database
    - Summary statistics from ensemble
      - E.g., 10/50/90 percentile forecasts - Pase 1
    - All ensemble members
      - E.g., 20-100 members - Phase 2
  - Provide easy access to internal / external users
    - NOMADS, etc?
    - Link with multi-center ensembles
      - NAEFS – NUOPC – GIFS

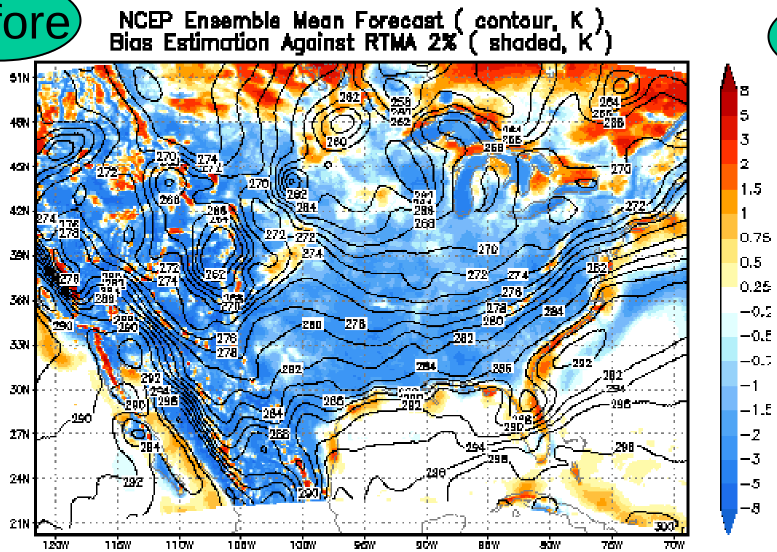
- **Database interrogation / forecaster tools**
  - Modify summary statistics
  - Back-propagate modified information into ensemble
  - Derive any information from summary statistics / ensembles

00hr GEFS Ensemble Mean & Bias Before/After Downscaling 10%

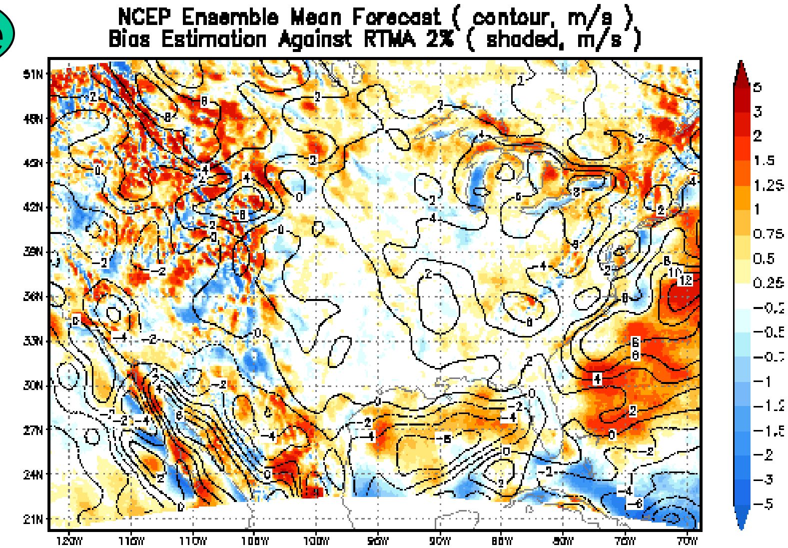
2m Temperature

10m U Wind

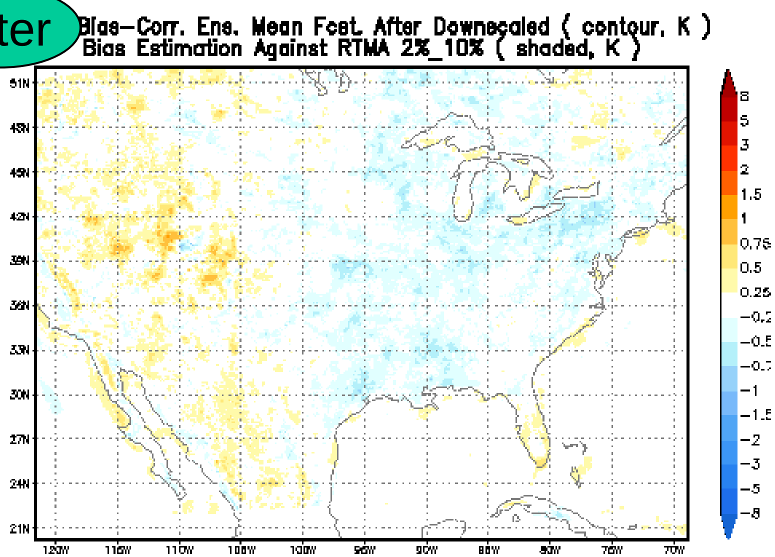
Before



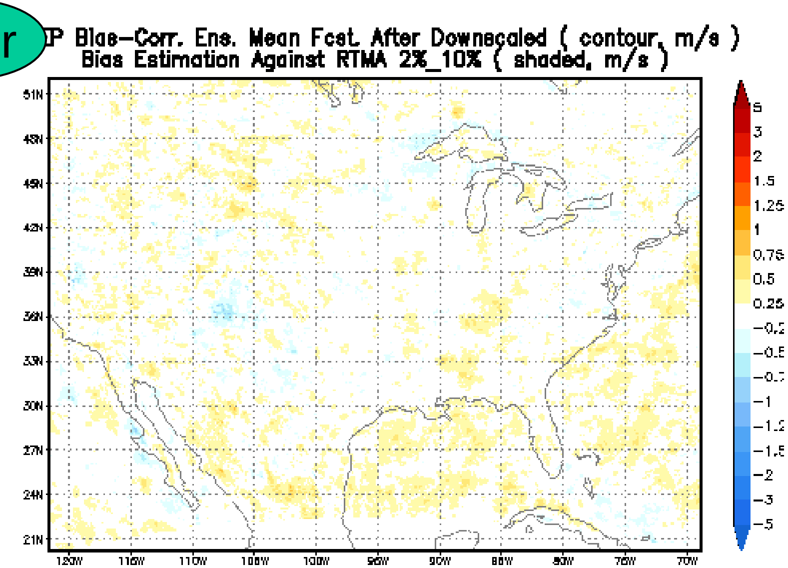
Before



After



After

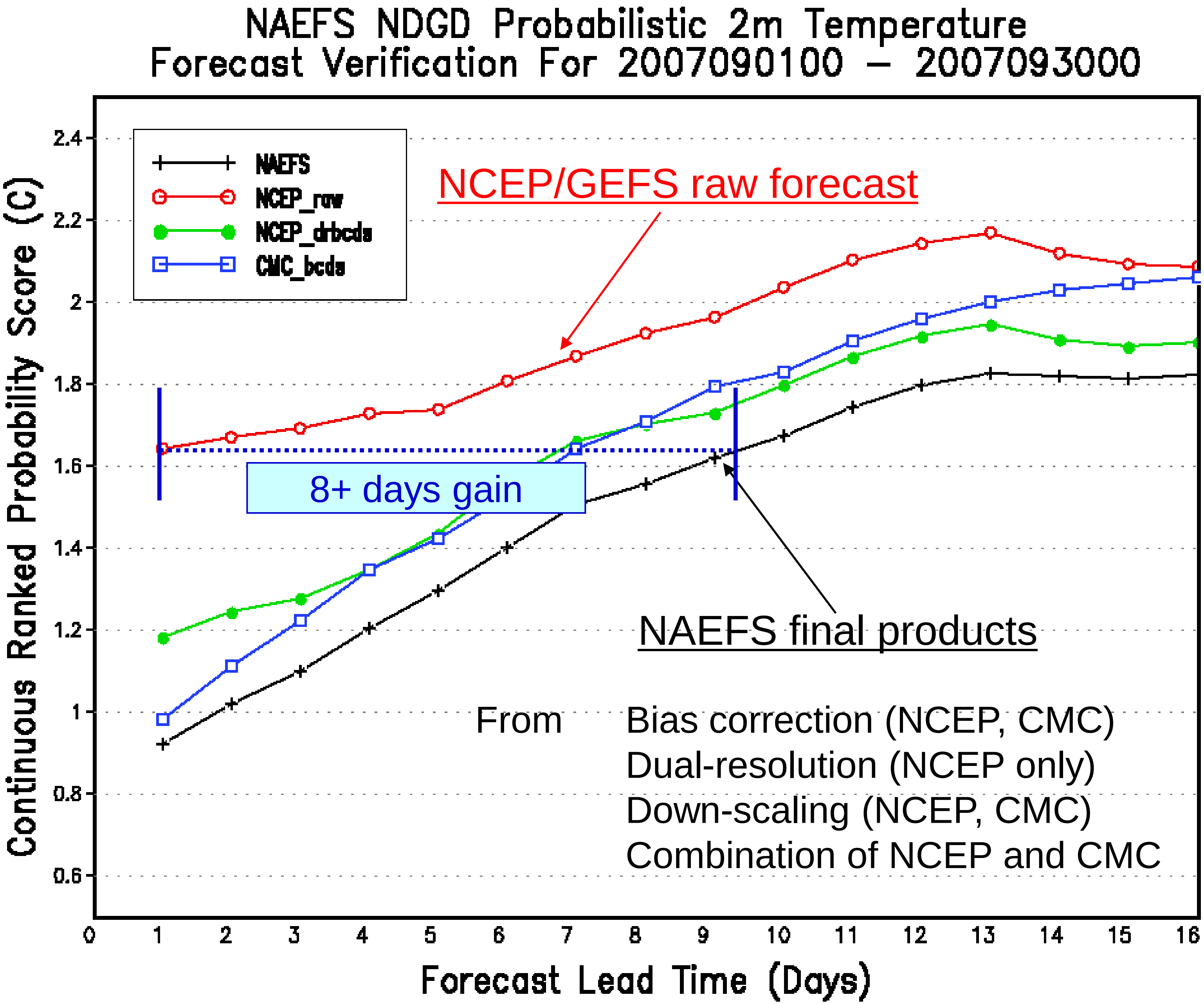


BO CUI, GCWMB/EMC/NCEP/NOAA

BO CUI, GCWMB/EMC/NCEP/NOAA

CONTINUOUS RANKED PROBABILITY SCORE

RAW / BIAS CORR. & DOWNSCALED & HIRES MERGED / NAEFS



High resolution control & Canadian ensemble adds significant value

=>

8-day total gain in skill

NUMERICAL WEATHER PREDICTION (NWP) BASICS

COMPONENTS OF NWP

- Create initial condition reflecting state of the atmosphere, land, ocean
- Create numerical model of atmosphere, land, ocean

ANALYSIS OF ERRORS

- Errors present in both initial conditions and numerical models
- Coupled atmosphere / land / ocean dynamical system is chaotic
  - Any error amplifies exponentially until nonlinearly saturated
  - Error behavior is complex & depends on
    - Nature of instabilities
    - Nonlinear saturation

IMPACT ON USERS

- Analysis / forecast errors negatively impact users
  - Impact is user specific (user cost / loss situation)
- Information on expected forecast errors needed for rational decision making
  - Spatial/temporal/cross-variable error covariance needed for many real life applications
  - How can we provide information on expected forecast errors?

HOW CAN WE REDUCE & ESTIMATE EXPECTED FORECAST ERRORS?

STATISTICAL APPROACH

- Statistically assess errors in past unperturbed forecasts (eg, RUC)
  - Can correct for systematic errors in expected value
  - Can create probabilistic forecast information – Eg, MOS PoP
- Limitation
  - Case dependent variations in skill not captured
  - Error covariance information practically not attainable

DYNAMICAL APPROACH – Ensemble forecasting

- Sample initial & model error space - Monte Carlo approach
  - Leverage DTC Ensemble Testbed (DET) efforts
- Prepare multiple analyses / forecasts –
  - Case dependent error estimates
  - Error covariance estimates
- Limitation
  - Ensemble formation imperfect – not all initial / model errors represented

DYNAMICAL-STATISTICAL APPROACH

- Statistically post-process ensemble forecasts
  - Good of both worlds
  - How can we do that?