



Developing NUOPC Metrics: Forecaster Surveys

NAEFS Workshop Monterey, CA

Scott Sandgathe

2 May 2012

Outline

- NUOPC Metrics Mandate
- Status of Ensemble Metrics
- 1st Annual Forecaster Survey
- Survey Results
- Summary & Discussion

Metrics Mandate

The Interim TTP Committee and the Interim UEO Committee reports established requirements for tracking NUOPC value. Specifically, metrics to:

- Assess superior GNWP performance (next slide)
- Demonstrate reduced duplication of effort
 - Track use of ESMF/NUOPC layer, common components?
- Determine accelerated technology transition
 - Not addressed yet
- Determine NUOPC value
 - Assess forecaster/customer adoption/impact?
 - Assess timeliness and reliability of ensemble data distribution

Common Performance Metrics

- Common Metrics approach agreed by ESG
- Common verification:
 - Ten parameters measured at 6-hr intervals
 - Broad verification over N/S hemispheres
 - Defined observation set, UKMO analysis, and World Wide Merged Cloud Analysis used for verification
 - Appropriate statistical methods employed (RMSE, Brier Score, CRPS, Ensemble Mean/Spread)
- USAF real-time statistics available through JAAWIN site
- NOAA seasonal statistics available publicly
- Navy beta test CAC site. Working toward public website.

User (Forecaster) Surveys

- Intended as an annual survey to assess NUOPC progress
- Target line forecasters and operations heads
- Assess organizational adoption of ensemble products
- Assess forecaster adoption of ensemble products in forecast preparation
- Assess forecaster training in ensemble prediction and ensemble products
- Assess adequacy and timeliness of ensembles

Survey Format

- “Survey Monkey” ten questions – multiple choice with space for comments. Sent via email, anonymous responses. (Questions in backup)
- Operations head questions:
 - Available, timely, used in products, preferred format, forecaster training, customer training?
- Line forecaster questions:
 - Used in forecast prep, how used, preferred format, timely, training sources, training adequate, any additional products or training needed?

Survey Response

- Operations heads:
 - NOAA 15
 - Navy 12
 - USAF 11
- Forecasters
 - NOAA 10
 - Navy 46
 - USAF 60

Ensemble Use

- Ensembles available (ops):
 - NOAA Use: 14 yes, 1 not sure; Timely 12 yes, 2 no, 1 not sure
 - Navy Use: 10 yes, 2 not sure; Timely 7 yes, 5 not sure
 - USAF Use: 11; Timely: 9 yes, 2 not sure
- Ensembles used in preparing forecasts:
 - Ops: NOAA 9 yes, 2 no, 3 not sure, Navy 7 yes, 5 not sure, USAF 9 yes, 2 not sure
 - Forecasters: NOAA 9 yes, 1 no, Navy 37 yes, 9 no, USAF 52-8
 - Bottom line: 80 to 90% of forecasters using ensembles, almost all felt they were timely.
- Goal for next year: Eliminate “not sures”

Ensemble Use (2)

- Percent of forecast products developed with ensembles (number of responses):
 - NOAA Ops: <25 (4), 26-50 (5), 51-75 (0), >75 (2)
 - NOAA Fcst: <25 (0), 26-50 (3), 51-75 (2), >75 (5)
 - Navy Ops: <25 (2), 26-50 (3), 51-75 (0), >75 (0)
 - Navy Fcst: <25 (11), 26-50 (21), 51-75 (6), >75 (1)
 - USAF Ops: <25 (4), 26-50 (3), 51-75 (0), >75 (0)
 - USAF Fcst: <25 (24), 26-50 (20), 51-75 (5), >75 (5)
- Bottom line: An average of less than 50% of products are developed using ensembles. These numbers should be NUOPC benchmark.

Ensemble Use (3)

- Forecasters used ensembles in (%):

	NOAA	Navy	USAF
Every Forecast	90	36	40
Forecast Scenarios	30	64	53
Watches, Warnings, Advisories	10	44	51
Severe Weather	30	64	66
Map Discussions	70	63	46

Preferred Format (Ops%/Fcstr%)

	NOAA	Navy	USAF
• Means	79/89	67/40	27/28
• Spaghetti	86/100	67/8	27/28
• Thumbnails	26/33	22/34	82/38
• Thresholds	35/33	11/63	64/59
• PDFs	14/67	22/34	18/34
• Joint PDFs	7/22	22/34	45/30
• Significant disconnects could be due to training or type of use (big picture vs specific forecast)			

Ensemble Training

- Is training provided to forecasters?
 - NOAA Ops: 11 yes, 2 no, 2 not sure
 - Navy Ops: 6 yes, 2 no, 4 not sure
 - USAF Ops: 6 yes, 4 no, 1 not sure
 - NOAA Fcstr: 80% yes, 20% no
 - Navy Fcstr: 74% yes, 36% no
 - USAF Fcstr: 53% yes, 47% no
- Are customers trained in use (ops)?
 - NOAA 1 yes, 6 no, 7 not sure, 1 skip
 - Navy 1 yes, 5 no, 6 not sure
 - USAF 2 yes, 8 no, 1 not sure

Ensemble Training (2)

- What training is available (Ops%/Fcstr%)?

	NOAA	Navy	USAF
On line training	93/89	100/42	90/50
School courses	9/33	41/36	30/48
On the job training	83/66	67/74	80/74
No training	18/27	0/8	0/8

- Where did you receive training?

	NOAA	Navy	USAF
On line training	89	36	38
School courses	67	51	48
On the job training	67	74	73
No training	0	8	8

Additional Products

- NOAA
 - Box & Whisker plots (range, spread, mean), Standard deviations
 - Data with parameters for workstations ingest
 - Probability of exceedance (survival plots) diagrams
- USAF
 - Mean field grids, numerical vice graphic data, higher resolution
 - Data for ingest into LEADS
 - more Aerostat areas, Red/yellow/green plots
 - Probability plots for severe weather, thunderstorms, wind, snow..
 - All significant weather on one map display
 - Ensemble data on weekends (several requests)
- Navy
 - Probability tables for ICAOs (individual stations)

Desired Training

- NOAA
 - Refresher courses would be beneficial
 - Training for emergency managers & end users
 - Training on interpretation of ensembles
 - Best practices for use of specific ensemble products
- Navy/USAF
 - Curriculum for regular in house training updates, SME visits
 - Training on specific products
 - Better understanding of underlying models and ensembling process, how to apply forecaster input
 - Better understanding of when to use, not use ensembles
 - Better understanding/interpretation of likelihood
 - Better understanding of probabilities on high winds, etc., actual probability much greater than indicated
- Training not uniform. Some said training excellent, others none!

Summary

- Implementing metrics required by Interim NUOPC reports
- Skill metrics mostly in place
- First of annual surveys to track NUOPC progress and impact on forecasters and products?
- Operations heads need to be more aware of ensemble use and training
- Forecasters need to be made aware of training sources available, need in house curriculum with regular updates
- Training required on interpreting and using specific products
- “Not sures” need to decrease!
- **How do we cause change? Tiger teams? Traveling show?**

Questions & Discussion

User Workshop Actions - Training

- Socialize available COMET and NCEP ensemble training material
- Archive significant events as lessons for training
- Ensure ensemble/probabilistic material is part of routine training

User Workshop Actions - Development

- Develop a GUI/Viewer for viewing complete ensemble (NUE/NAEFS) with capability to:
 - Pull data and products from NUE site
 - Handle all ensemble data/fields
 - Create joint PDF products
 - Apply user-defined thresholds to PDFs
 - Perform clustering, plan view, and other data manipulation technologies
 - Provide tailored specific events forecasts
 - Identify long-term events (flooding, heat waves, etc.)
- Develop techniques to increase use of covariance in ensemble interpretation
- Provide more products tailored to decision makers
- Invest in social science for communicating uncertainty

User Workshop Actions - Operations

- Use ensemble information to identify and automate benign/certain forecast cases
- Produce ensemble bulletin or map discussion (or add ensembles to existing map discussions)
- Provide sensitivity information to forecasters
- Review operational products for inclusion of probabilistic information

User Workshop Actions – Additional NUOPC Items

- Advertise on weather blogs (clear statement of data/product availability and how to access, training sites)
- Develop ensemble website containing:
 - Links to training
 - Examples of use
 - Product description/availability/access
 - Discussion bulletin board for comments, lessons learned
- Create multi-agency post processing code repository including format/coding standards
- Review ensemble management (number members, optimum frequency, forecast period) for NUE production

Common Metrics Table

Parameter	Thresholds	Statistics	Regions	Ground Truth	Fcst Hour Interval	Comments
500 mb HGT	N/A	Anomaly Correlation, RMSE/Bias Ensemble Mean, Spread, CRPS	20N-20S,20N-80N,20S-80S, Northern Hemisphere, Southern Hemisphere, CONUS	UKMO, Observations (After IOC)	12hr	NCEP provides gridded climatology
10 Meter Winds	>25,35,50kts	Brier Score, CRPS, RMSE/Bias Ensemble Mean, Spread, Conditional Metrics (Forecast RMSE > 35kts/50kts)	20N-20S,20N-80N,20S-80S, Northern Hemisphere, Southern Hemisphere, CONUS	UKMO, Observations (After IOC)	12hr	
24hr Accumulated Precipitation	>.25,.5,1,2"	Brier Score, CRPS	20N-20S,20N-80N,20S-80S, Northern Hemisphere, Southern Hemisphere, CONUS	Observations	12hr	
700mb RH	N/A	RMSE/Bias Ensemble Mean, Spread, CRPS	20N-20S,20N-80N,20S-80S, Northern Hemisphere, Southern Hemisphere, CONUS	Observations (After IOC)	12hr	
2 Meter Temperature	N/A	RMSE/Bias Ensemble Mean, Spread, CRPS	20N-20S,20N-80N,20S-80S, Northern Hemisphere, Southern Hemisphere, CONUS	UKMO, Observations (After IOC)	12hr	
850mb Temperature	N/A	RMSE/Bias Ensemble Mean, Spread, CRPS	20N-20S,20N-80N,20S-80S, Northern Hemisphere, Southern Hemisphere, CONUS	UKMO, Observations (After IOC)	12hr	
Significant Wave Height	>12,18,24ft	RMSE/Bias Ensemble Mean, Spread, CRPS	20N-20S,20N-80N,20S-80S, Northern Hemisphere, Southern Hemisphere	Observations	12hr	

Common Metrics Table

250mb Wind Speed	N/A	RMSE/Bias Ensemble Mean, Spread, CRPS	20N-20S,20N-80N,20S-80S, Northern Hemisphere, Southern Hemisphere, CONUS	UKMO, Observations (After IOC)	12hr	
Total Cloud Cover (percent)	N/A	RMSE/Bias Ensemble Mean, Spread, CRPS	Northern Hemisphere, Southern Hemisphere	Hemispheric World Wide Merged Cloud Analysis (WWMCA)	12hr	AFWA provides WWMCA
Hurricane Tracks	N/A	Spread, Mean Track Error	As Required	Best Track	Best Track Interval	Verified after the fact once best track is available
		CONUS - Continental Boundaries between Canada and Mexico and the east and west coasts				
		Bias - average difference between model and observation value				
		CRPS - Continuously Ranked Probability Score - measurement of how well a probabilistic ensemble forecast predicted a categorical observation (e.g., rain/no rain)				
		RMSE - Root Mean Square Error - measure of square root of the average squared error between model and observation				
		Spread - measure of the distance of an ensemble forecast distribution from the ensemble mean				
		Ensemble Mean - average of all ensemble forecasts				
		Anomaly Correlation - measure of similarity of departures of model and observation values from a climatological mean field				

Operations Heads Survey

- Are ensemble products available for forecaster use?
- Is ensemble information received in a timely manner for use in development of forecasts/products?
- Is ensemble information used for providing probabilistic forecasts to any operational decision makers?
- What percent of products used for forecast production are ensembles?
 - a. 1-25% b. 26-50% c. 51-75% d. over 75%
- What is the preferred delivery format of ensemble products for your forecasters (choose as many as apply)?
 - a. Means b. Thumbnails c. Probability Density Functions d. Joint probabilities
 - e. Spaghetti Plots f. Thresholded Percentages
- What customer products developed by your organization could (or do) benefit the most from incorporation of ensemble probabilistic information and certainty?

Operations Head Survey (2)

- Is training provided by your Agency to your office/forecasters on proper use of ensembles? a. Yes b. No c. Not Sure
- Is training provided to end-customers and decision makers on how to interpret probabilistic forecasts? a. Yes b. No c. Not Sure
- What training sources are available to the forecasters?
a. On-line websites b. School-house course c. On-the-job training curriculum d. None e. Other (please specify)
- Training is sufficient for effective use of ensemble products.
a. Agree b. Slightly Agree c. Neutral d. Slightly Disagree e. Disagree
- What training areas need improvement?

Forecaster Survey (1)

- Are you using ensemble products to prepare any of your forecasts?
a. Yes b. No c. Not Sure
- What percent of your forecast products include input from ensemble information?
a. 1-25% b. 26-50% c. 51-75% d. over 75%
- How are ensemble products and information used?
a. Every forecast b. Specific forecast scenarios c. Watches/warnings /advisories d. To focus forecast effort on highest chance of impacting weather
- Does your Agency's synoptic discussion, map discussion, or equivalent product describe the model performance (initialization and verification) of ensembles? i.e. probability, uncertainty, or sensitivity of regimes a. Yes b. No c. Not Sure

Forecaster Survey (2)

- In what format do you prefer ensemble product delivery (choose as many as apply)? a. Means b. Thumbnails c. Probability Density Functions d. Joint probabilities e. Spaghetti Plots f. Thresholded Percentages
- Is ensemble information delivered in a timely manner and available for use in development of forecasts/products? a. Yes b. No c. Not Sure
- If you could request any ensemble products that you feel would be useful, they would be....
- Is training provided by your Agency on proper use of ensembles?
- What training sources are available to you? a. On-line websites b. In person instructor c. On-the-job training curriculum
- Which forms of training have you completed? a. On-line websites b. In person instructor c. On-the-job training curriculum d. Other e. None
- The training was sufficient to enable your effective use of ensemble products. a. agree b. slightly agree c. neutral d. slightly disagree e. disagree