Methods for Post-Processing

Approach

- Define set of problems
- Provide data / tools
- Develop and run various methods
- Independent testing and intercomparison
  - Create canonical datasets for optional use in development
  - Test will be done on independent (near real time?) data
- Assessment of methods

Postprocessing Requirements

- Combine various sources of predictive info
  - Unperturbed, ensemble forecasts, latest observations, climatology
- Calibration of model variables
  - With ensemble data
  - Reanalysis - reforecast requirement - for all valid forecast times
- Derivation of user variables / downstream products
  - NWP and observations / obs-analysis requirement

Postprocessing Requirements

FACT: Post-processing requires data from deterministic and/or ensemble NWP model output paired with observations and climatology

FINDING: Analyses/Observations are not available at all forecast times of interest

RECOMMENDATION: There should be greater investment in generating more frequent analyses paired with each operational forecast model

Configuration

TWO-STAGE APPROACH

- Calibration of prognostic vars on model grid against model analysis
  - Reanalyses, reforecasts
  - Ensemble members adjusted - is balance retained? Research question
- Derivation of user variables at arbitrary points
  - E.g., obs locations, fine scale grid
  - NWP and obs-based finer scale re-analyses

ONE SWEEP APPROACH

- From raw prog vars directly to derived fine scale vars
  - Added value? Research question
    - If two stages have nonlinear connections
    - If balance is distorted by calibration on model grid

New data needs

- Fine scale reanalysis of user variables
- Canonical sample datasets for testing/comparing methods
  - Updating datasets including reforecast data as models change
- More output from real time and re-analysis/forecast
  - Frequency, more variables
  - e.g., hourly precip accumulations
  - Frequency statistic fields (e.g., hourly max, mean, standard deviation)
- Reanalysis / reforecast
  - Cost of running and benefit from it, compared to benefits from enhanced real time forecast system? - Research question
  - Method dependent need - Research question
  - Real time generation?

New Data Needs

FACT: Weather variables relevant to users are desired at finer space and time scales

FINDING: Generating more variables at finer scales requires more computing and storage power

RECOMMENDATION: NOAA should determine the cost of running and benefit from more reanalysis/reforecast data, compared to benefits from enhanced real time forecast system.

Research should determine which methods are more or less dependent reforecast data.

NOAA should provide canonical test datasets for major weather variables in conjunction with major model updates for post-processing method comparisons

Calibration Methods

- Calibration - univariate, all prog vars
  - First 2-3 moments in ensemble adjusted to match corresponding analysis distribution
  - Multiple linear regression
  - BMA, EKDMOS, BPE
- Calibration - multivariate
  - Account for correlation structure among prognostic variables
  - Copula methods, multitask methods
  - New method - too expensive? Not enough sample data?
- Combination of all forecast info
  - BMA, BPE...
- Approaches
  - Adaptive - Adjust parameters to changes in DA/model suite
    - Recursive parameter estimation, Kalman filters, etc
  - Avoid overfitting

Calibration Methods

FACT: Univariate calibration of methods corrects systematic biases in first 3 moments of ensemble distributions. This calibration can be done with multiple linear regression, BMA, BPE, EKDMOS

FINDING: Univariate calibration may affect correlations among related variables and decrease the effectiveness of calibrating user-defined variables

RECOMMENDATION: NOAA should evaluate multivariate calibration methods (Copula, Schaake Shuffle, multitask methods) against univariate methods to determine the effects on correlations across calibrated prognostic variables.

Derivation of User Variables/ Downstream Methods

- No matter method, regularize/avoid overfitting
- Perfect prog
- Traditional MOS and other linear regression methods
  - Multiple Linear Regression
  - Logistic regression
  - Ridge, lasso, elastic net regularized methods
- Machine learning methods
  - Decision Tree ensemble Methods (Random Forests, Gradient Boosting)
  - Analog/Nearest Neighbor methods
  - Neural Networks
- Bayesian methods

Derivation of User Variables/Downstream Variables

- FACT: Statistical derivation of downstream and user-relevant weather variables provides added values in many domains
- FINDING: Machine learning and post-processing methods have added predictive performance and value over linear methods in many domains both within and outside the weather community
- RECOMMENDATION: NOAA needs to be open to nonlinear machine learning and Bayesian methods in addition to perfect-prog and MOS methods and investigate how they can been implemented on a larger scale.

Evaluation

- Quality - compared to baselines (competitors, random, climate)
  - Reliability
  - Stat. resolution / info content
  - User relevant metrics - economic cost-benefit analysis
  - Preferably permutation / resampling tests (avoid parametric tests)
- Computational cost
- Maintainability
- Scientific soundness
- Extensible

Evaluation

FACT: Post-processing methods need to be judged based on statistical quality (reliability, resolution, economic value), computational costs, maintainability, scientific soundness, and extensibility

FINDING: Current statistical evaluation metrics do not fully convey all the relevant aspects of implementing current or new post-processing methods

RECOMMENDATION: When NOAA evaluates current and new methods, it should incorporating all aspects of their operation in order to determine which method should be implemented operationally. Any statistical