BRINGING IT ALL TOGETHER:
A PROTOTYPE FOR A PROBABILISTIC NATIONAL BLEND

Acknowledgements:
Kevin Kelleher, Matt Peroutka, Yuejian Zhu,
John Wagner, Geary Layne, Jeffrey Craven

Zoltan Toth
OAR/ESRL/GSD

Mark Antolik
NWS/MDL

Malaquias Pena
NWS/NCEP/EMC & IMSG

Roman Krzysztofowicz
University of Virginia

Melissa Petty, Geary Layne
OAR/ESRL/GSD

AMS 2017 Annual Meeting
OUTLINE / SUMMARY

• Role & relevance of NWP & SPP in forecasting
  – NWP provides predictive information
  – SPP fixes NWP miscalibration - Potentially major user impact

• Proposed framework for National Blend of Models (NBM)
  – Combine & calibrate guidance for NWP prognostic variables
  – Derive user variables from calibrated prognostic variables

• A prototype for probabilistic NBM
  – Bayesian Processor of Ensemble (BPE)
    • Combines multiple ensemble / high res forecasts & obs/analyzed climate
    • Unified processing of all continuous variables
    • Comprehensive set of output variables

• BPE status
  – Algorithm, software developed
  – Ongoing transfer to MDL
  – Comparative test against EKDMOS next
STATISTICAL POST-PROCESSING (SPP) IN NWP

- **NWP is basis** of weather forecasting
  - Provides all predictive information beyond 6 hrs
    - Data assimilation, numerical model, ensemble approach

- **Systematic errors** limit utility of NWP
  - Forecasts are miscalibrated (mislabeled)

- Statistical Post-processing
  - **Re-labels** forecast information

- Traditional approach to SPP
  - **Process each NWP guidance product individually**
    - NAM-MOS, GFS-MOS, NAEFS-EKDMOS, etc
    - Multitude of products overwhelm forecasters

- Birth of NWS National Blend of Models (NBM)
  - **Merge & calibrate forecasts**
PROBLEMS WITH NWP FORECASTS

• **Drift of model solutions** from real to model world
  – 1\(^{st}\) moment calibration

• **Uncertainty not captured well** by single/ens fcsts
  – 2\(^{nd}\) moment calibration

• **Abundance of guidance products**
  – Fusing all predictive information

  *Lead-time dependent* behavior of

  *NWP Model Prog Variables (MPV)* on model grid

• **Missing variables**
  – Limited model resolution in space, processes, variables

  *Simultaneous relationships btw*

  *User Specific Variables on fine & MPVs on model grid*
OBJECTIVE OF STATISTICAL POST-PROCESSING

Uncalibrated model variables

Calibrated user variables

FORECAST DAY-1

FORECAST DAY-2

FORECAST DAY-N

Coarser model grid

Fine scale grid

Multiple lead times

USER SPECIFIC VARIABLES
COMMON APPROACH TO POST-PROCESSING

1-Step, direct approach

Uncalibrated model variables

Calibrated user variables

Cost: $L \times M \times (N / M)$; Retrain for each model update

Model gridpoints ($M$) $<<$ ($N$) Fine scale gridpoints $O(10-100)$

Multitudes of lead-time and fine-scale grid dependent relationships

USER SPECIFIC VARIABLES
CALIBRATION OF MODEL PROGNOSTIC VARIABLES

(Re)forecast sample  NWP (re)analysis

NWP ANALYSIS

Calibration Toolbox

FORECAST DAY-1

FORECAST DAY-2

FORECAST DAY-N

Multiple lead times

Cost: L x M

Model grid (M)
CALIBRATION OF MODEL PROGNOSTIC VARIABLES

(Ne)forecast sample  NWP (re)analysis

NWP ANALYSIS

FORECAST DAY-1

FORECAST DAY-2

FORECAST DAY-N

Calibration Toolbox

CALIBRATED MODEL PROG VARS

Model grid

Multiple lead times

- (Re)forecast sample
- NWP (re)analysis
- Calibration Toolbox
- Model grid

- FORECAST DAY-1
- FORECAST DAY-2
- FORECAST DAY-N

- Drift of model solutions from real to model world
- Uncertainty not captured well by single/ensembles
- 1st moment calibration
- 2nd moment calibration
- Fusing all predictive information
- Abundance of guidance products
- Lead-time dependent behavior of NWP Model Prog Variables (MPV) on model grid
CALIBRATION OF MODEL PROGNOSTIC VARIABLES

(Re)forecast sample                  NWP (re)analysis

NWP ANALYSIS

FORECAST DAY-1

FORECAST DAY-2

FORECAST DAY-N

Multiple lead times

Calibration Toolbox

CALIBRATED
MODEL PROGS

N / M savings from fixing model problems on model grid

Model grid (M)
DERIVATION OF USER VARIABLES

- Missing variables
  - Limited model resolution in space, processes, variables

Simultaneous relationships btw

Single, instantaneous relationship

User Specific Variables

NWP ANALYSES

NWP reanalysis

Fine scale (re)analysis

No need for hindcasts

Derivation Toolbox

Model grid

Fine scale grid

NWP

USER SPECIFIC VARIABLES
DERIVATION OF USER VARIABLES

NWP reanalysis

Fine scale (re)analysis

Cost: $M \times \frac{N}{M}$

Retrain only for reanalysis update

No need for hindcasts

Derivation Toolbox

Model grid

Fine scale grid

Single, instantaneous relationship

NWP analyses

User specific variables
PUTTING IT TOGETHER: 2-STAGE POST-PROCESSING

(Re)forecast sample

NWP ANALYSIS

FORECAST DAY-1
FORECAST DAY-2
FORECAST DAY-N

Multiple lead times

Calibration Toolbox

NWP reanalysis

Derivation Toolbox

Fine scale (re)analysis

No need for hindcasts

User specific variables

Single, simultaneous relationship

Model grid

Fine scale grid

Model grid

Calibrated model prog vars

NWP reanalysis

(Re)forecast sample

Multiple lead times

Calibration Toolbox

Derivation Toolbox

User specific variables

Single, simultaneous relationship

Model grid

Fine scale grid
1-STEP vs 2-STAGE APPROACH TO POST-PROCESSING

• **Disadvantages of 1-step approach**
  – Addresses coarse scale model problem on fine scale grid – **Redundancy**
  – Separate derivation of user variables for each lead time - **Redundancy**
  – Requires hindcast sample

• **Benefits of 2-stage approach**
  – Addresses independent problems separately
  – Better understanding / separate metrics to focus on issues
  – Modular design eases **community involvement**, supports collaboration
  – Shared development of derivation tool box
    • Vast array of “perfect prog” tools available – UPP, DA forward models, satellite product generation, etc
  – Calibrated model variables can be used in downstream coupled apps
    • Liberates downstream users from ever changing model version specific biases
    – Develop and use reanalysis – based relationships
    • One-way coupled user models (eg, hydro) can be **calibrated with reanalysis**
  – Simplifies forecast editing – Modify calibrated model prog variables
    • Multitude of consistent user variables automatically derived w/o addition. edits
  – Thorough **analysis of systematic model errors** for model developers
RECOMMENDED APPLICATION

- **Calibrate all model prognostic variables**
  - Lead-time dependent behavior eliminated
    - NWP (re)analysis as proxy for truth
    - Calibrated forecasts statistically behave like analyses
    - Systematic error estimates useful for model developers
  - NWP-based predictive info fused on model grid
  - Scales with number of *model gridpoints x leadtimes*

- **Derive additional user variables**
  - Fine scale analysis as proxy for truth (RUA)
  - Instantaneous relationships between analyses of
    - Model Prog Vars (MPV) & User Specific Variables (USV)
      - No lead-time dependency - No need for hindcasts
  - Toolbox with various types of routines
    - UPP, DA forward models and other physical & statistical methods
  - Scales with number of *fine scale gridpoints*
Model Prognostic Variables

REAL TIME NWP FORECAST
Ensemble, High Res. Control
Hindcast, (Re)analysis

Calibrate & Consolidate NWP Output
Derive Fine Scale User Vars.

Calibrated Model Prog Variables
Calibrated User Variables

Forecast Editor

6D-DATACUBE
Climate & Forecast Distributions
Ensemble Scenarios

Product Generator

Pregenerated Products, Interactive Services

PROBABILISTIC FORECAST SERVICES

EXTERNAL USERS

3rd Party Apps.

Coupled Models

NWS FORECASTERS

NWS FORECASTERS

Pregenerated Products, Interactive Services

15
Bayesian Processor of Ensemble

Calibrated Model Prog Variables

Calibrated User Variables

6D-DATACUBE
Climate & Forecast Distributions
Ensemble Scenarios

BPE Product Generator, other tools

Pregenerated Products, Interactive Services

Proposed iEdit

Planned Bays. Deriv., UPP, etc

Coupled Models

3rd Party Appls.

PROBABILISTIC FORECAST SERVICES

Fine Scale User Var.

REAL TIME NWP FORECAST
Ensemble, High Res. Control
Hindcast, (Re)analysis

HISTORICAL DATA
Reanalysis

NWS FORECASTERS

EXTERNAL USERS
1) Estimate climate distribution

2) Assess skill in & combine all predictors

3) Cast forecast in climate context

4) Offer multiple output formats

Krzysztofowicz & Toth 2008, Krzysztofowicz 2010
BAYESIAN PROCESSOR OF ENSEMBLE (BPE) BASICS

• Use **climatology as basis** – link w climate prediction
  – Fit **parametric distribution** to large climate sample
    • Calibration ensured, extremes handled gracefully w/o large hindcast
  – Express forecasts as anomalies from climate
    • No need for large hindcast sample to define obs. Climatology

• Transform each variable into **normal space**
  – Unified approach for ALL continuous variables

• Assess **predictive information** in guidance products
  – Multilinear regression, 1\(^{st}\) & 2\(^{nd}\) moments as predictors
    • Well established technology (MOS); Forecast information maximized

• Process **current forecast**
  – Calibrated forecast as modified climate distribution (percentiles)
    • Most economic storage – 2 modified pars of climate distribution
  – Prob. dist., quantiles, posterior ensemble (scenarios / joint prob)
6D NDFD – CENTRAL TENET OF FORECAST PROCESS

Connects NWP output w. users via value added steps

- **Role**
  - Repository of authoritative probabilistic guidance
    - Forecaster modification via *editor tool*

- **Function**
  - Holds *answer to all questions* about future environ.
    - *Access info via interrogator tool*

- **Content**
  - Calibrated *prog. vars.* from *post-processing tool*
  - Calibrated *user variables* – incl. climate percentiles

- **Format**
  - 6D - Space (3), time, variables, uncertainty

- **Implementation**
  - **Phase 1** – *Only 2-fold expans. of current NDFD*
    - Climatological distributions – negligible cost
    - 2 BPE posterior parameters to describe fcst distrib
  - **Phase 2**
    - Ensemble members for joint probabilities & scenarios
NATIONAL BLEND OF MODELS –
BRINGING ALL THE PIECES TOGETHER

• Modular/comprehensive/expandable **framework**
• Academic **research & NOAA operations**
• Array of theory based **techniques**
• End-to-end **workflow** w up-to-date software
• Variety of **forecast products** & climate info
• **All model prog vars** processed in unified way
• Calibrated and informative **guidance**
• Comprehensive array of **output formats**
• Serving diverse **applications**
BACKGROUND
OUTLINE / SUMMARY

• Role & relevance of NWP & SPP in forecasting
  – NWP provides predictive information
  – SPP fixes NWP miscalibration - Potentially major user impact

• Proposed framework for National Blend of Models (NBM)
  – Combine & calibrate guidance for NWP prognostic variables
  – Derive user variables from calibrated prognostic variables

• A prototype for probabilistic NBM
  – Bayesian Processor of Ensemble (BPE)
    • Combines multiple ensemble / high res forecasts & obs/analyzed climate
    • Unified processing of all continuous variables
    • Comprehensive set of output variables

• BPE status
  – Algorithm, software developed
  – Ongoing transfer to MDL
  – Comparative test against EKDMOS next
CURRENT GFS CONFIGURATION

GFS COMPONENTS

GFS Model

OUTPUT

Raw Model Prognostic Variables

APPLICATIONS

Coupled Models & Other Downstream User Applications

Biased User Variables
NGGPS SYSTEM DESIGN CONSIDERATIONS
Component functionalities, links, software infrastructure

NGGPS COMPONENTS

NWP Model
Calibration Stat. Post-Pr.
Derivation UPP, SPP, etc

Raw Model Prognostic Variables
Calibrated Model Prog Variables
Calibrated Fine Scale User Vars

APPLICATIONS
Coupled Models & Other Downstream User Applications

OUTPUT

Two-way Coupling
One-way Coupling
BPE PROJECT STATUS

- **OMM version** - One ensemble, Multiple control & Multiple auxiliary predictors
  - Algorithm & documentation complete (U. Va)
  - Codes complete (GSD)
  - Off-line testing (GSD, ongoing)
  - Transition to MDL (Oct16)
  - Testing at MDL (Nov16)

- **MMM version** – Multiple ensembles added
  - Algorithm & documentation (U. Va, Sep16)
  - Codes complete (GSD, Nov16)
  - Off-line testing (GSD, Jan17)
  - Comparison w EKDMOS in operational environment (MDL, Mar17)
    - Assess quality & computational speed at observation sites

- **Final report** (All, May 2017)