Linear and Nonlinear Postprocessing of Ensemble Forecasts

Ranran Wang\textsuperscript{1} and Caren Marzban\textsuperscript{1,2}

\textsuperscript{1}Department of Statistics
\textsuperscript{2}Applied Physics Laboratory
Univ. of Washington, Seattle, Washington
Generalities

NWP forecasts have problems.

Correction of forecasts = post-processing = MOS.

Long history:
- Glahn, Lowry 1972
- Wilson, Vallee 2002
- Marzban 2003
- Pasini et al.
- Marzban, Sandgathe, Kalnay 2005
- Marzban, Leyton, Colman 2007
- Gneiting, Raftery 2005
- Berrocal, Raftery, Gneiting 2007
- Wilson, Beauregard, Raftery, Verret 2007
- Sloughter, Raftery, Gneiting 2007
- Gneiting, Kleiber, Schlather 2009
- Bao, Gneiting, Raftery, Grimit, Guttorp 2009
Method

Here, we postprocess ensemble forecasts.

At a given station, at a given time, one has
- M forecasts (from an M-member ensemble),
- of meteorological variables (temp, wind, etc), and
- surface observations.

Goal: Develop a map from forecasts to obs.

Statistical Models:

Regression: temperature and wind (U and V)
Classification: yes/no prcp.

Raw: ensemble mean and ensemble proportion.
Linear: multiple regression (LM) and linear discriminant analysis (LDA)
Nonlinear: Neural Network (NN), minimizing MSE or cross-entropy.
Method ...

Reduce dimensionality of input data.

2-level hierarchical approach:

Step 1:
\[
T_{\text{fit}} = \text{regress}(T_{\text{obs}} \sim T_1 + T_2 + \cdots + T_{10})
\]
\[
U_{\text{fit}} = \text{regress}(U_{\text{obs}} \sim U_1 + U_2 + \cdots + U_{10})
\]
\[
V_{\text{fit}} = \text{regress}(V_{\text{obs}} \sim V_1 + V_2 + \cdots + V_{10})
\]

Step 2:

LM1
\[
T_{\text{obs}} \sim T_1 + T_2 + \cdots + T_{10}, \quad T_{\text{obs}} \sim T_1 + T_2 + \cdots + T_{10} + U_{\text{fit}} + V_{\text{fit}},
\]
\[
U_{\text{obs}} \sim U_{\text{fit}} + V_{\text{fit}}, \quad U_{\text{obs}} \sim U_{\text{fit}} + V_{\text{fit}} + T_{\text{fit}},
\]
\[
V_{\text{obs}} \sim U_{\text{fit}} + V_{\text{fit}}. \quad V_{\text{obs}} \sim U_{\text{fit}} + V_{\text{fit}} + T_{\text{fit}}.
\]
LM2

LDA1
\[
P_{\text{obs}} \sim P_1 + P_2 + \cdots + P_{10} \quad P_{\text{obs}} \sim P_1 + P_2 + \cdots + P_{10} + T_{\text{obs}} + U_{\text{obs}} + V_{\text{obs}}.
\]

LDA2

Similarly, for NNs.

Input data = residuals (to filter out seasonality).

residuals = observed - conditional median.

Input data = residuals (to filter out seasonality).

residuals = observed - conditional median.
Performance

Measures:
Mean Squared Error (MSE) for regression,
Area under ROC curve (AUC) for classification.

Assessment:
Cross-validation (on years) → sampling distribution of training and validation performance measures.

Compare ensemble mean, LM, and NN.
Compare ensemble proportion, LDA, and NN.

In cross-validation, models are paired.
Must look at pair-wise difference of models.
E.g. Where is zero on the distribution of \( \text{MSE}_{vld}(\text{LM}) - \text{MSE}_{vld}(\text{NN}) \) ?

Vary \( H = \) number of hidden nodes.
Data (from Fanyou Kong and Steve Leyton)

Regional fine-scale ensemble based on WRF-ARW Version 3.0.1

Ten members, each with unique initial perturbations and varying physics options and land-use tables.

Initial perturbations = bred vectors produced by 6-hourly breeding cycle, using the North America Regional Reanalysis (NARR) as background fields.

Three-domain, two-way nesting, innermost domain = 15 km grid spacing.

19 years (1987-2005), 48-hr re-forecasts of temperature, wind, and prcp, at 90 stations across the continental US.

5-day intervals, leading to 1387 cases.

Plus surface obs (analysis, for prcp) at each station.
Surface obs at KSEA

Temperature

Wind U

Wind V

Precipitation
Conditional Medians

Temperature

Wind U

Wind V
NNs Overfit!

KOKC MSE Temp vs. ensemble mean, LM, and NN with $H = 2, 4, 8, 16, 32, 64$.

Training

Validation
KSEA

**Training**

**Validation**
Conclusion I

Temperature, Wind U, Wind V:

- LM improves T, U and V predictions over ensemble means in terms of MSE.

- For wind U/V, adding T as additional covariate improves on MSE significantly for only a few stations. Forecasts of T are improved by adding wind U/V for most stations.

- For wind U/V, NN does not improve over LM in terms of MSE; forecasts of T from NN are worse than LM for all stations.
Conclusion II

Precipitation:

• LDA improves over ensemble proportion, except for a few stations in the Northwestern region.

• NN makes no improvement over LDA.

• A clear East-West spatial pattern observed when we compare LDA w/ additional covariates (LDA2) with LDA w/o additional covariates (LDA1).

We thank Fanyou Kong and Steve Leyton for the data.