Super-Ensemble Statistical Forecasting of Monthly Precipitation over the Contiguous US, with Improvements from Ocean-Area Precipitation Predictors

Thomas Smith¹, Sam Shen², and Ralph Ferraro¹

1. NOAA/NESDIS/STAR and CICS/ESSIC/U. Maryland
2. San Diego State University

The contents of this presentation are solely the opinions of the authors and do not constitute a statement of policy, decision, or position on behalf of NOAA or the U. S. Government.
Definitions

• Ensemble: A weighted mean of multiple estimates
  – Traditionally used for GCM forecast runs with different initial conditions

• Statistical Ensemble: A weighted mean of different statistical estimates
  – Ensemble members may have different predictors, different predictor regions, or use different statistical models to give different estimates

• Super Ensemble: Use ensemble-averaging weights that reflect the accuracy of each member
Regions for predictors: OI SST and GPCP P

4 Ocean predictor areas with 20°N-23°N overlap

Regions likely to influence $P_{US}$, similar to Lau et al. (2002) areas

Predictors for ensemble:
• Ocean area $SST_k(t-1)$
• US area $P_{US}(t-1)$
• Ocean area $P_k(t-1)$

Always predict $P_{US}(t)$ anoms
Two Models: CCA and JEOF

- **CCA**
  - Decomposes predictor and predictand fields using EOFs

- **JEOF**
  - Simultaneous EOF of normalized predictor and predictand fields

- Predictors are leading SST and P, predictand is US P

- Super-ensemble weights use cross-validation skill of each forecast
Data & Evaluations

• GPCP precipitation and OI SST
  – 1997-2014 1dd GPCP averaged to monthly, compute anomalies

• Cross-validation testing of 0-lead monthly forecasts
  – Omit all data for the year of analysis and 3 months on either side of the year
  – Data from month t-1 to predict month t

• Correlations used to evaluate skill and improvements
Annual Cycle of US Average Skill

Ensemble CCA using SST(t-1) regions better than CCA using the same SST(t-1) combined (upper panel)

Ensemble improved more when including prediction from \( P_{US}(t-1) \)

Using SST(t-1) and \( P_{US}(t-1) \) predictors, JEOF better than CCA and using both is best (lower panel)

More models and super ensemble method gives improvements
Cross-Validation Precipitation Anomaly Correlation: June, no oceanic precipitation

JEOF and CCA skill patterns similar, but not identical

Regions of high skill different in different models

Super ensemble using both takes the best of each
Cross-Validation Precipitation Anomaly Correlation:
December, no oceanic precipitation

Both JEOF and CCA show skill gaps but in different regions

Using both expands the region of good skill

Methods Conclusions:
1) Ensembles dividing predictors into regions improves skill
2) Using ensemble members from multiple models also improves skill
Including Oceanic Precipitation in 4 Regions

Skill increases when including members with ocean area $P(t-1)$ predictors

JEOF better than CCA, using both is best
Cross-Validation Precipitation Anomaly Correlation: June, with oceanic precipitation

Ocean P ensemble members improve both JEOF and CCA

JEOF still better, and combining them still gives higher skill
Cross-Validation Precipitation Anomaly Correlation: December, with oceanic precipitation

More regions with higher skill than the case with no oceanic precipitation: satellite-based P improves the forecast

Best skill apparently from ENSO

Low-skill regions for both JEOF and CCA not improved by combining them
Skill from more than ENSO

- Skill from Tropical Pacific area SST or Precip important but not the whole story
- Combining with forecasts using SST and Precip from other regions doubles average correlation
- All averages omit no-skill regions (correlations < 0)

Temporal cross-validation correlations against GPCP computed for each month (1997-2014), averaged over the contiguous US and annually.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>CCA</th>
<th>JEOF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{T_Pac}$</td>
<td>0.20</td>
<td>0.18</td>
</tr>
<tr>
<td>$P_{T_Pac}$</td>
<td>0.21</td>
<td>0.23</td>
</tr>
<tr>
<td>$E[T_i,P_{US}]$</td>
<td>0.31</td>
<td>0.35</td>
</tr>
<tr>
<td>$E[T_i,P_{i},P_{US}]$</td>
<td>0.39</td>
<td>0.45</td>
</tr>
</tbody>
</table>
Overall Improvements from oceanic precipitation

- Adding satellite-based $P_i(t-1)$ predictors improves ensembles

Temporal cross-validation correlations against GPCP computed for each month (1997-2014), averaged over the contiguous US and annually.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>CCA</th>
<th>JEOF</th>
<th>JEOF+CCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E[T_i, P_{US}]$</td>
<td>0.31</td>
<td>0.35</td>
<td>0.42</td>
</tr>
<tr>
<td>$E[T_i, P_i, P_{US}]$</td>
<td>0.39</td>
<td>0.45</td>
<td>0.50</td>
</tr>
</tbody>
</table>
Comparisons to Similar NAMME Tests
Similar Skill Levels but in Different Regions

X-Val Ens Corr \[P(m),F(m-1)\]

From Mo and Lettenmaier (2014, J. Hydromet.)
Conclusions

• Super-ensemble-statistical forecast are better than non-ensemble forecasts
  – Method improvements include using multiple statistical models and super-ensemble averaging weights

• Ocean-area precipitation predictors improve US-area precipitation forecasts
  – Additional predictors add skillful members to the ensemble and give higher ensemble skill
  – Many other predictors may give skill and improve the forecast, including different statistical predictors and estimates from numerical models; more testing is needed