Subseasonal Prediction Skill from SubX

Timothy DelSole

George Mason University, Fairfax, Va and Center for Ocean-Land-Atmosphere Studies

September 12, 2017

collaborators: Abhishekh Srivastava, Laurie Trenary, Michael Tippett, Kathy Pegion, Dan Collins, Emerson LaJoie
Predictability of Week 3-4 Averages

Evidence of Sub-Seasonal Predictability

Pegion, Sardeshmukh (2011; MWR): $\psi$ and OLR (CFS, GEOS5, LIM)

Johnson et al. (2013; Wea. For.): N. America T. (empirical)

Wang et al. (2014; Climate Dyn.): MJO (CFSv2)

Vitard (2014; QJRMS): MJO and NAO (ECMWF)

Li and Roberts (2015; MWR): Summer P. (CFSv2, JMA, ECMWF)
Predictability of Week 3-4 Averages

**Evidence of Sub-Seasonal Predictability**

Pegion, Sardeshmukh (2011; MWR): $\psi$ and OLR (CFS, GEOS5, LIM)

Johnson et al. (2013; Wea. For.): N. America T. (empirical)

Wang et al. (2014; Climate Dyn.): MJO (CFSv2)

Vitard (2014; QJRMS): MJO and NAO (ECMWF)

Li and Roberts (2015; MWR): Summer P. (CFSv2, JMA, ECMWF)

No clear demonstration of skill by dynamical models for predicting week 3-4 averages of T. or P. over North America.
Correlation Skill of CFSv2 Hindcasts
Week 3–4 Prediction; Lagged Ensemble= 4 days

Jan Temp (59%) Jan Precip (41%)
Jul Temp (36%) Jul Precip (9%)

-0.65 -0.5 -0.35 -0.2 0.2 0.35 0.5 0.65
Data Details

4-Member Lagged Ensemble

- CFSv2 hindcasts initialized 0Z, 6Z, 12Z, 18Z each day 1999-2010.
- Consider only 14-day mean of weeks 3-4 (15-28d) hindcasts.
- Temperature validated with NCEP/NCAR reanalysis.
- Precipitation validated with CPC Unified Gauge-Based Analysis.
- Subtract out smoothed climatology conditioned on verification day.
Permutation Test

4-Member Lagged Ensemble

- Standard significance test is not appropriate because hindcasts initialized 6 hours apart are not independent.
- Under null hypothesis of no predictability, hindcasts are exchangeable for the same start day and lead.
Predictable Component Analysis

Determine linear combination of variables that maximizes S/N.

\[
\begin{align*}
\text{signal} & = \text{variance of ensemble means} \\
\text{noise} & = \text{variance} \ about \ the \ ensemble \ means
\end{align*}
\]
Most Predictable Component
Week 3–4 Prediction

Jan Temp (x2)  Jan Precip (x1.3)
Jul Temp (x0.7)  Jul Precip (x0.8)
Skill of Predictable Components
Week 3–4 Prediction; Lagged Ensemble= 4 days

- Jan Temp
- Jan Precip
- Jul Temp
- Jul Precip

component
correlation
SubX BY THE NUMBERS

7 Global Models

1 Year of Real-time Forecasts

17 Years of Retrospective Forecasts

3-4 week guidance for Climate Prediction Center Outlooks

courtesy of Kathy Pegion
<table>
<thead>
<tr>
<th>Model</th>
<th>Lead (d)</th>
<th>E</th>
<th>Time Steps</th>
<th>Initial Conds</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECCC – GEM</td>
<td>32</td>
<td>4</td>
<td>6976</td>
<td>only 4 IC dates every year: 10Aug, 17Aug, 24Aug, 31Aug, 7Sep</td>
</tr>
<tr>
<td>EMC – GEFS</td>
<td>35</td>
<td>11</td>
<td>914</td>
<td>every 7th day starting from 02June and ending at 30Nov of each year</td>
</tr>
<tr>
<td>ESRL – FIMr1p1</td>
<td>32</td>
<td>4</td>
<td>835</td>
<td>every 7th day starting from 06Jan1999 and ending at 31Dec2014</td>
</tr>
<tr>
<td>GMAO – GEOSv2p1</td>
<td>45</td>
<td>4</td>
<td>5990</td>
<td>every 5th day starting from 05July and ending at 27Nov of each year</td>
</tr>
<tr>
<td>NRL – NESM</td>
<td>45</td>
<td>1</td>
<td>5995</td>
<td>each set of 4 consecutive ICs starting from 03Jul1999 separated by 3 days</td>
</tr>
<tr>
<td>RSMAS – CCSM4</td>
<td>45</td>
<td>3</td>
<td>6569</td>
<td>every 7th day starting from 07Jan and ending at 29Apr of each year</td>
</tr>
</tbody>
</table>
ESRL Start Dates (Ensemble Size = 4)
every 7th day starting 6Jan1999

2-3 samples per day
GMAO Start Dates (Ensemble Size = 4)
every 5th day starting 5July/ending 27Nov of each year

17 samples 1 day/week, 0 samples on other days of the week
16 samples 1 day/week, 0 samples on other days of the week
Local Linear Regression

Local linear regression result

\( \hat{Y}(X_0) \)

\( Y(X) \)

data points
Monte Carlo Experiment

1. Synthetically generate data from the model

\[ Obs(t) = \sum_{k=1}^{2} \left( a_k \cos \left( \frac{2\pi t k}{365} \right) + b_k \left( \frac{2\pi t k}{365} \right) \right) + \text{noise} \]

2. Subsample in a way similar to SubX (e.g., every 7 days for 16 years)

3. Test different methods for estimating climatology:
   - **sample mean**: average of each calendar day
   - **harmonic**: estimate parameters in (1) using least squares.
   - **local (28)**: local linear regression with 28-day window
16 years; IC every 7 days

- true climo
- sample mean
- harmonic
- local (p=1, 28)
- local (p=2, 28)
ESRL Re-forecasts of Temperature
lon=290; lat=70; lead=28; E=4
Correlation Skill (masked) for ESRL Week 3–4 Hindcasts

month=Jan; 1999–2015; E=4
ESRL Climatology for Leads 15–28 days
lon=290; lat=70; N/T=0.013; E=4

temperature

calendar day
Measure of Lead-Time Dependence of Climatology

\[
\frac{N}{T} = \frac{\text{standard deviation about each calendar-day mean}}{\text{standard deviation over all calendar days}}
\]
ESRL Climatology for Leads 15−28 days
lon=240; lat=32.5; N/T=0.063; E=4

calendar day
temperature

0 100 200 300
287 288 289 290 291 292

0 100 200 300
287 288 289 290 291 292

calendar day
ESRL Climatology vs Lead

lon=240; lat=32.5; day=200; N/T=0.063; E=4
ESRL re-forecasts of precipitation
lon=290; lat=70; lead=28; E=4
ESRL Climatology for Leads 15–28 days
 precipitation; lon=260; lat=17.5; N/T=0.06; E=4

![Graph showing precipitation over calendar days from January to January, with peaks in July and September.](image-url)
ESRL Climatology for Leads 15–28 days
precipitation; lon=292.5; lat=45; N/T=0.8; E=4
ESRL Climatology vs Lead

precipitation; lon=292.5; lat=45; day=200; N/T=0.8; E=4
Correlation Skill (masked) for ESRL Week 3–4 Hindcasts

month=All; 1999–2015; E=4
Summary

Predictability of Week 3-4 CFSv2 Forecasts over CONUS

- CFSv2 skillfully predicts week 3-4 temperature and precipitation.
- Significance of the skill determined by rigorous permutation test.
- Skill detected also using Predictable Component Analysis.
- Most predictable patterns are related to ENSO.
- Some predictability of winter precipitation related to MJO.

SubX over CONUS

- Data inhomogeneities complicate estimation of model climatology.
- Local linear regression (LOESS) appears very promising.
- Including lead-time dependence in climatology will be critical.
- ESRL has statistically significant skill for week 3-4 temperature.
- Subtracting climatology may not be best way to remove model precipitation biases.