Advance of Dynamical Seasonal Prediction: Assessment of the APEC Climate Center (APCC) / Climate Prediction and its Application to Society (CliPAS) 14-Model Ensemble Retrospective Seasonal Prediction (1980-2004)

Submitted by: University of Hawai‘i


1. School of Ocean & Earth Science & Technology, University of Hawaii, USA
2. School of Earth and Environmental Sciences, Seoul National University, Korea
3. Climate Dynamics Program, George Mason University, USA
4. School of Earth and Environmental Sciences, Seoul National University, Korea
5. APCC, Busan, Korea
6. NCEP/CPC, Maryland, USA
7. NASA/GSFC, Maryland, USA
8. NOAA GFDL, Princeton University, USA
9. Florida State University, USA
10. FRSGC, Tokyo, Japan
11. University of Tokyo, Japan
12. IAP/CAS, China
13. Bureau of Meteorology Research Center, Australia
14. Korean Meteorological Administration, Korea

Corresponding Author: Bin Wang, Department of Meteorology and International Pacific Research Center, University of Hawaii, Honolulu Hawaii 96822, USA. Email: wangbin@hawaii.edu

Abstract

This study assesses current status of multi-model ensemble (MME) deterministic and probabilistic seasonal prediction based on 25-year (1980–2004) retrospective forecasts performed by 14 climate model systems (7 one-tier and 7 two-tier systems) that participate in the Climate Prediction and its Application to Society (CliPAS) project sponsored by the Asian-Pacific Economic Cooperation Climate Center. Based on the assessment, future direction for improvement of seasonal prediction is discussed.

We found that two measures of probabilistic forecast skill, the Brier Skill Score (BSS) and the Area under the Relative Operating Characteristic curve (AROC), display similar spatial patterns as that of the Temporal Correlation Coefficient (TCC) score of the deterministic MME forecast. An AROC score of 0.7 corresponds approximately to a BSS of 0.1 and a TCC of 0.6 in the CliPAS MME system. The MME method is demonstrated to be a useful and practical approach for reducing errors and quantifying forecast uncertainty due to model formulation. The MME prediction skill is substantially better than the averaged skill of all individual models. For instance, the TCC score of CliPAS one-tier MME forecast of Niño 3.4 index at a six-month lead initiated from May 1 is 0.77, which is significantly higher than the corresponding averaged skill of 7 individual coupled models (0.63). The MME made by using 14 coupled models from both DEMETER and CliPAS shows an even higher TCC score of 0.87. For probabilistic forecast the CliPAS MME gains considerable skill from increased forecast reliability; the forecast resolution also increases for 2 m temperature but slightly decreases for precipitation forecast. The effectiveness of the MME depends on the averaged skill of individual models and their mutual independency.

Equatorial sea surface temperature (SST) anomalies are primary sources of climate predictability worldwide. The MME one-month lead hindcast can predict, with high fidelity, the spatial-temporal structures of the first two leading empirical orthogonal modes of the equatorial SST anomalies for both the boreal summer (JJA) and winter (DJF), which account for about 80% to 90% of the total variance. The major bias is a westward shift of the SST anomaly between the dateline and
120°E, which may potentially degrade global teleconnection associated with it. The TCC score for SST predictions over the equatorial eastern Indian Ocean reaches about 0.68 with a six-month lead forecast. However, the TCC score for the Indian Ocean Dipole (IOD) index drops below 0.40 at a three-month lead for both the May and November initiations due to prediction barriers across January and July.

The MME prediction skills are well correlated with the amplitude of Niño 3.4 SST variation. The forecasts are better in El Niño years than in La Niña years. The precipitation and circulation are predicted better in ENSO-decaying JJA than in ENSO-developing JJA. There is virtually no skill in ENSO-neutral years. Thus, precipitation forecast depends on accurate forecast of the amplitude, spatial patterns, and detailed temporal evolution of ENSO cycle. Therefore, continuing improvement of the one-tier climate model’s slow coupled dynamics in reproducing a realistic ENSO mode is a key for long-lead seasonal forecast. Forecast of monsoon precipitation remains a major challenge. The seasonal rainfall predictions over land and during the local summer have little skill, especially over the tropical Africa. The differences in the forecast skills over land areas between CliPAS and DEMETER MMEs indicate potentials for further improvement of predictability over land. There is an urgent need to assess the impact of the land surface initialization on the skill of the seasonal and monthly forecast using a multi-model framework.

APCC/CliPAS team
Presented by Bin Wang
Department of Meteorology and IPRC
University of Hawaii

APCC Symposium August 19-21 2008 Lima
International Cooperation for Climate Prediction

APCC/CliPAS Team

COLA/GMU: J. Shukla (Co-PI), B. Kirtman, J. Kinter, K. Jin
SNU: I.-S. Kang (Co-PI), J.-S. Kug
NASA/GSFC: S. Schubert, M. Suarez, W. Lau
NOAA/GFDL: N.-C. Lau, T. Rosati, W. Stern
NOAA/NCEP: A. Kumar, J. Schemm
FSU: T. Krishnamurti, S. Cocke
FRCGC/JAMSTEC: J. Luo, T. Yamagata (UT)
IAP/CAS: T. Zhou, B. Wang
BMRC: O. Alves, H. H. Hendon
KMA: W.-T. Yun
CCSR: M. Kimoto
CMC
APCC/CliPAS Tier-1 Models

<table>
<thead>
<tr>
<th>Institute</th>
<th>AGCM</th>
<th>Resolution</th>
<th>Ensemble</th>
<th>Member</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMRC</td>
<td>BAM v3.0d</td>
<td>T47L17</td>
<td>ACOM2</td>
<td>0.5-1.5° lat x 2° lon L25</td>
<td>10</td>
</tr>
<tr>
<td>FRCGC</td>
<td>ECHAM4</td>
<td>T106 L19</td>
<td>OPA 8.2</td>
<td>2° cos(lat) x 2° lon L31</td>
<td>9</td>
</tr>
<tr>
<td>GFDL</td>
<td>R30</td>
<td>R30L14</td>
<td>R30</td>
<td>R30 L18</td>
<td>10</td>
</tr>
<tr>
<td>NASA</td>
<td>NSIPP1</td>
<td>2° lat x 2.5° lon L34</td>
<td>Poseidon V4</td>
<td>1/3° lat x 5/8° lon L27</td>
<td>3</td>
</tr>
<tr>
<td>NCEP</td>
<td>GFS</td>
<td>T62 L64</td>
<td>MOM3</td>
<td>1/2° lat x 1° lon L40</td>
<td>15</td>
</tr>
<tr>
<td>SNU</td>
<td>SNU</td>
<td>T42 L21</td>
<td>MOM2.2</td>
<td>1/2° lat x 1° lon L32</td>
<td>6</td>
</tr>
<tr>
<td>UH</td>
<td>ECHAM4</td>
<td>T31 L19</td>
<td>UH Ocean</td>
<td>1° lat x 2° lon L2</td>
<td>10</td>
</tr>
</tbody>
</table>

APCC/CliPAS Tier-2 Models

<table>
<thead>
<tr>
<th>Institute</th>
<th>AGCM</th>
<th>Resolution</th>
<th>Ensemble</th>
<th>Member</th>
<th>SST BC</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFDL</td>
<td>AM2</td>
<td>2° lat x 2.5° lon L24</td>
<td>10</td>
<td>SNU SST forecast</td>
<td>Anderson et al. (2004)</td>
<td></td>
</tr>
<tr>
<td>IAP</td>
<td>LASG</td>
<td>2.8° lat x 2.8° lon L28</td>
<td>6</td>
<td>SNU SST forecast</td>
<td>Wang et al. (2004)</td>
<td></td>
</tr>
<tr>
<td>NCEP</td>
<td>GFS</td>
<td>T62 L64</td>
<td>15</td>
<td>CFS SST forecast</td>
<td>Kanamitsu et al. (2002)</td>
<td></td>
</tr>
<tr>
<td>SNU/KMA</td>
<td>GCPS</td>
<td>T63 L21</td>
<td>6</td>
<td>SNU SST forecast</td>
<td>Kang et al. (2004)</td>
<td></td>
</tr>
<tr>
<td>UH</td>
<td>CAM2</td>
<td>T42 L26</td>
<td>10</td>
<td>SNU SST forecast</td>
<td>Liu et al. (2005)</td>
<td></td>
</tr>
<tr>
<td>UH</td>
<td>ECHAM4</td>
<td>T31 L19</td>
<td>10</td>
<td>SNU SST forecast</td>
<td>Roecker et al. (1996)</td>
<td></td>
</tr>
</tbody>
</table>

Forecast quality measures

- For the deterministic forecast, MME prediction was made using simple arithmetic average of 14 models’ ensemble means.
  Multiple regression and SVD method
  SPPM method (more sophisticated statistical post-processes

- Probabilistic forecast was derived by simple democratic counting using 109 individual realizations from the ten models (5 one-tier and 5-two tier models) after normalizing each simulation with respect to its own mean and standard deviation.
**Evaluation of deterministic forecast skills**

- **Metrics** used to measure prediction skill of MME mean forecast include anomaly pattern correlation coefficient (PCC) and root mean square error (RMSE) normalized by the corresponding observed standard deviation.
- **Temporal correlation coefficient (TCC)** was used for a specific time series of a predictand. For convenience of comparison, we also calculated the time-averaged anomaly PCC and RMSE over a specific region.

**Evaluation of Probabilistic Forecast Skill**

- **Brier Skill Score (BSS)**
  
  The BSS can be computed as the sum of following three terms (Murphy 1973, Wilks 1995, Stanski et al. 1989)

  \[ b = b_{rel} - b_{res} + b_{unc} \]

  Reliability measures the bias in predicted probabilities relative to the verified frequency for given event.
  Resolution measures the intrinsic ability of the forecast system to detect (or resolve) situations in which the observed frequency of the event is different from the climatological frequency.

- **Area under ROC curve (AROC)**
  
  Relative operating characteristics (ROC) is a graph of hit rate against false alarm rate within a range of probability threshold.
  AROC is a way to quantify the ROC by calculating the area beneath the ROC curve (Green and Swets 1966). If the area is less than 0.5 of the whole (unit area), then the model is less skillful than a random or constant forecast.
1. Relationship between Probabilistic and Deterministic prediction skills

Two probabilistic forecast skills, the Brier Skill Score (BSS) and the Area under the Relative Operating Characteristic curve (AROC), display similar spatial patterns as the Temporal Correlation Coefficient (TCC) score of the deterministic MME forecast.

While the relationship is nonlinear, AROC 0.7 ~ BSS 0.1 ~ TCC 0.6 in the CliPAS MME system.
2. Effectiveness of MME: ENSO prediction

- The 6-month lead forecast initiated from May 1
- Averaged skill of 7 individual coupled models: 0.63.
- The MME of 7 CliPAS model: 0.77.
- The MME made by 14 CliPAS + DEMETER models: 0.87.

2a Effectiveness of MME: Precipitation

- The effectiveness of the MME depends on the averaged skill of individual models and their mutual independency.
2b Probabilistic forecast: greater advantages

The reliability and resolution term of BSS

The reliability of BSS score (rain) increases from 0.48 to 0.73 when 10 models are used.

The resolution score also increases for 2 m temperature but slightly decreases for precipitation forecast.

3. Hindcast Equatorial SST [10S-5N]

The MME one-month lead hindcast can predict, with high fidelity, the spatial-temporal structures of the first two leading empirical orthogonal modes of the equatorial SST anomalies for both the boreal summer (JJA) and winter (DJF), which account for about 80% to 90% of the total variance. The major bias is a westward shift of the SST anomaly between the dateline and 120°E.

Wang et al. 2008
4. Hindcast NINO 3.4 SST

Overall Skill

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Tier-1 MME</th>
<th>Dynamic-Statistical Model</th>
<th>Persistence</th>
</tr>
</thead>
</table>

Tier-1 MME Forecast

Seasonal dependence

ENSO Phase-dependence

5. Indian Ocean SSTA Hindcast

Correlation Skill for Indian Ocean SSTA

(a) WIO [50°-70°E, 105°-100°W] Initiated from May 1

(b) EIO [90°-110°E, 105°-90°W] Initiated from May 1

(c) IOD Initiated from May 1

The TCC score over the equatorial EIO reaches about 0.68. The TCC score for IOD index drops below 0.40 at a three-month lead for both the May and November initiations due to prediction barriers across January and July.
6. Precipitation Prediction: TCC 13 Coupled Model MME

MME prediction skills depend highly on the amplitude of Niño 3.4 SST variation. Forecasts are better in El Niño years than in La Niña years. Prediction is better in ENSO-decaying JJA than in ENSO-developing JJA. Virtually no skill in ENSO-neutral years.
6b Longitudinal dependence


JJA skill (0.53) is better than DJF skill (0.50). MME prediction of air temperature is considerably superior to the persistence skill in the warm pool oceans, but not necessarily over the land.
8. Skill for circulation fields are high

200 hPa streamfunction shows very good correlation skill almost everywhere. The high-skill prediction of 850 hPa streamfunction shifts eastward from JJA to DJF.

Conclusion

1. Relationship between two types of forecast measures
2. Effectiveness of MME approach
3. Equatorial SST: Principal modes
4. ENSO forecast: Season- and Phase-dependent
5. Indian Ocean SST forecast: winter and summer prediction barrier
6. Precipitation: season-, longitude-, and ENSO-dependent
7. Temperature: gains over persistence
8. High circulation prediction
Conclusion

- MME is an extremely valuable approach for reducing errors and quantifying forecast uncertainty due to model formulation.
How do we move forward with seasonal prediction? Two aspects need to be considered.

- Given the current levels of the climate models, how do we get the best forecast through MME?
- From a long-run, what are the priorities we should take in improving our climate models’ physics?

**Given the current levels of the climate models, how do we get the best forecast through MME?**

- Improvement of MME approach for better down scaling
- Determine predictability of ISO and make statistical and dynamical forecast ISO and monthly prediction.
- Statistical-dynamical approach for climate prediction of extreme events.
- Urgent need is to determine the role of land-atmosphere interaction in monsoon predictability.
Directions

- Improvement of MME approach for better down scaling
- Determine predictability of ISO and make statistical and dynamical forecast ISO and **monthly prediction**.
- Statistical-dynamical approach for climate prediction of extreme events.
- Urgent need is to determine the role of land-atmosphere interaction in monsoon predictability.
- Continuously improving **slow coupled physics** is a key for long-lead seasonal forecast.
- Improve initialization of **one-tier system**, including coupled data assimilation and reanalysis.
- Development of **High resolution global models** for prediction of TC and other extreme events.
- Improvement of models' physics representation and correcting systematic mean errors.

CliPAS Publication

- Published:
  - Overall assessment of CliPAS MME: Advance and prospectus (Wang et al. 2008);
  - Predictability of ENSO (Jin et al. 2008),
  - Predictability of the A-AM (Wang et al. 2008),
  - Predictability of ISV (Kim et al. 2008, Fu and Wang 2007),
  - Performance of coupled models on annual cycles and the relation to seasonal prediction skills (Lee et al. 2008), and
  - Optimal MME method for seasonal climate prediction (Kug et al 2008).
- A number of manuscripts to be submitted this year
Physical basis

(1) Two-Tier System

Forced problem:
*Predictability comes from anomalous lower boundary forcing*
Charney and Shukla (1977, 1981),
Lorenz (1982)

Two-tier system:
*AGCM forced by predicted SST*
Bengtsson et al. (1993), Barnett et al. (1994), Levezey et al. (1996)


(2) One-Tier System

Initial value problem:
*Predictability comes from Initial memory and slow coupled dynamics.*
Palmer (1993), Palmer and Shukla (2000),

ENSO forecast: Cane and Zebiak (1985)

One-Tier system:
*Coupled A-OGCM*
Ji et al. (1996), Stockdale et al. (1998)

Systematic biases in coupled mean states and coupled modes
MME prediction: Optimal Selection of Models

Optimal Selection of a Subgroup of Models

Example: East Asian Domain [105-145E, 20-45N]

The best MME skill is obtained using 4 models.

Evaluation of Deterministic forecast

(1) Simple composite
- Equal weighting method \( P = \frac{1}{M} \sum F_i \)

(2) Multiple regression and SVD method
- Superensemble method \( P = \sum a_i F_i \)

(3) SPPM method (more sophisticated statistical post-processes)
Precipitation skill in DJF (0.57) is significantly higher than that in JJA (0.46) over the global tropics. The high DJF skill is in NH between 0 and 40°N, and especially in the northern subtropics between 20°N and 40°N from 40°E to 140°E in the A-AM sector and from 60°W and 90°W in the American sector.