ASIA-PACIFIC JOURNAL OF ATMOSPHERIC SCIENCES, 44, 3, 2008, p. 259-267

Optimal Multi-model Ensemble Method in Seasonal Climate Prediction

Jong-Seong Kug[1], June-Yi Lee[2], In-Sik Kang[1], Bin Wang[2] and Chung-Kyu Park[3]

[1]*Climate Environment System Research Center, Seoul National University, Korea*
[2]*International Pacific Research Center, SOEST, University of Hawaii, USA*
[3]*APEC Climate Center, Korea*

(Manuscript received 3 April 2008; in final form 5 June 2008)

Abstract

Given a large number of dynamical model predictions, this study endeavors to improve seasonal climate prediction through optimizing multi-model ensemble (MME) method. We have developed a new MME method and evaluated it using 15 dynamical models' retrospective forecasts for the period 1981-2001 in comparison with other MME methods. The strengths of the new method lie in a statistical error correction procedure for predictions of individual model and a discreet selection procedure of reliable predictors among all possible candidates. The conspicuous improvement of the new method is achieved against other MME methods over the regions in which the average of individual models' skill is very poor, such as land area and extratropical oceans. It is demonstrated that the selection procedure for the reliable predictors allows the present method to get more effective as the number of model predictions being used increases.

Key words: Muti-model ensemble, seasonal prediction, statistical correction

1. Introduction

During the last few decades, climate scientists have made tremendous advances in understanding and modeling the Earth's climate system. Although the most advanced technologies of the coupled atmosphere-ocean modeling and multi-model ensemble (MME) technique have been applied to dynamical seasonal prediction, limitations and challenges still remain (Palmer *et al*., 2004; Wang *et al.,* 2008). It is certain that better dynamical seasonal prediction can be achieved by improving a dynamical model itself and its initialization process. In addition, given a large number of available dynamical model predictions, further improvement of prediction can be obtained from post-processing procedures such as MME techniques and statistical error correction (or downscaling) methods. During the last decade, a lot of efforts have been devoted to extracting optimal and skillful MME predictions from the available seasonal climate predictions by using a simple composite method (Peng *et al*., 2002; Palmer *et al*., 2004; Wang *et al*., 2008) or a weighted ensemble method based on the multiple regression method (Krishnamurti *et al*., 1999; Kharin and Zwiers, 2002; Yun *et al*., 2003), and a synthetic method (Chakraborty and Krishnamurti, 2006). However, it remains an open question whether the sophisticated MME schemes are any better than the simple composite method for seasonal prediction.

It has been noted that the predicted dominant modes of interannual variation in precipitation tend to have a different spatial pattern from their observed counterparts but vary in a similar way to the observed one in the time domain (Feddersen *et al*., 1999; Kang *et al*., 2004). Based on that, coupled pattern techniques have been developed to correct the systematic bias in the model by using a statistical relationship between the predicted and observed anomalies. The most commonly used methodologies of the coupled pattern technique are based on canonical correlation analysis (CCA) or singular value decomposition analysis (SVD) (Feddersen *et al*., 1999; Kang *et al*., 2004)

As a statistical post-processing, statistical down-

Corresponding Author: Jong-Seong Kug, Climate Environment System Research Center, Seoul National University, Shillim-dong, Kwanak Gu, Seoul 151-747, Korea.
Phone : +82-2-880-6726, Fax : +82-41-856-8527
E-mail: kug@hawaii.edu

scaling techniques have been also used to improve seasonal prediction. Kug *et al*., (2007) introduced a point-wise downscaling technique called a pattern projection method (PPM). The PPM method is based on the large-scale patterns of predictors correlated to local predictands. Kug *et al.,* (2008a) recently introduced an improved downscaling method named Stepwise Pattern Projection Method (SPPM), showing improved prediction after its application to predicted SST by Seoul National University (SNU) CGCM (Kug *et al*., 2008b), especially over Western North Pacific and Indian Ocean where the model produces serious systematic bias.

The objective of this study is aimed at producing optimal prediction from an existing suite of seasonal predictions. We present a new multi-model ensemble technique based on the statistical post-process. The MME scheme is verified using retrospective forecasts of 15 dynamic models for the period of 1981-2001. The data and MME methodologies used are described in section 2. In section 3, the results of MME predictions using four different MME schemes are presented. Section 4 gives a brief summary and discussion.

2. Data and methodology

a. Retrospective forecast data

The 15 retrospective forecast data sets used are for the period 1981-2001 and were obtained from Development of a European Multi-model Ensemble system for seasonal to inTERannual prediction (DEMETER) (Palmer *et al*., 2004) and the Asia-Pacific Economic Cooperation Climate Center/ Climate Prediction and Its Application to Society (APCC/CliPAS) (Wang *et al*., 2008) projects. The DEMETER project produced 7 one-tier predictions and the APCC/CliPAS project produced 5 one-tier and 5 two-tier predictions. Table 1 presents a brief summary of each model. For more details of the models, the reader is referred to the relevant literatures cited in the Table 1. Among those predictions, we utilized 15-model predictions integrated from May 1st for the most of models targeting one-month lead prediction. Since each prediction model has several ensemble members from slight different initial conditions, ensemble mean of each model was taken. In this study, the target variable is summer-mean (JJA) precipitation.

Before applying MME methods, the ensemble mean prediction of each model was interpolated to a common 2.5^o latitude x 2.5^o longitude grid which is comparable to the resolution of the observed data. Note that the MME methods are applied after systematic climatological biases are removed by subtracting the forecast climatology of each model.

b. Observed data

The Climate Prediction Center Merged Analysis of Precipitation (CMAP) data were used to verify the climate model's performance for the period 1981-2001. The data are produced by merging rain gauge data, five kinds of satellite estimates, and a numerical model prediction.

c. MME Schemes

In this study, we compare several possible MME methods and propose a new optimal MME scheme. The following is a brief description of the MME methods utilized here.

(1) MME-EW

This equal weighting (EW) method is a simple but powerful method for seasonal climate prediction. The MME-EW is defined by

$$Y(t) = \frac{1}{N}\sum_{i=1}^{N} F_i(t)$$

where *Y(t)* is a MME prediction for time *t*, N is the total number of models being used, $F_i(t)$ is a forecast of the *i*th model for time *t*.

(2) MME-SVD

This method is a kind of weighted ensemble scheme defined by

Table 1. Description of 15 seasonal prediction models used in this study.

Institute	AGCM	OGCM	Ensemble member	Reference
CERFACE	ARPEGE T63 L31	OPA 8.2 2.0° x2.0° L31	9	Deque (2001)
ECMWF	IFS T95 L40	HOPE--E 1.4° x 0.3°-1.4° L29	9	Gregory *et al*. (2000)
INGV	ECHAM4 T42 L19	OPA 8.2 2.0° lat x 2.0° lon L31	9	Madec *et al*. (1998)
LODYC	IFS T95 L40	OPA 8.0 182GPx152GP L31	9	Gregory *et al*. (2000)
Meteo-France	ARPEGE T63 L31	OPA 8.0 182GPx152GP L31	9	Deque (2001)
UKMO	HadAM3 2.5x3.75L19	GloSea OGCM 1.25°x0.3-1.25°L40	9	Roeckner (1996) Marsland *et al*. (2003)
MPI	ECHAM5 T42 L19	MPI-OM1 2.5° lat x 0.5°-2.5° lon L23	9	Marsland *et al*. (2002)
NCEP CFS	GFS T62 L64	MOM3 1/3° lat x 1° lon L40	15	Saha *et al*. (2006)
SNU	SNU T42 L21	MOM2.2 1/3° lat x 1° lon L32	6	Kug *et al*. (2008b)
UH	ECHAM4 T31 L19	UH Ocean 1° lat x 2° lon L2	10	Fu and Wang (2001)
FSU	FSUGSM T63 L27	SNU SST forecast	10	Cocke. and LaRow(2000)
GFDL	AM2 2° latx2.5° lon L24	SNU SST forecast	10	Anderson *et al*. (2004)
NCEP	GFS T62 L64	CFS SST forecast	15	Saha *et al*. (2006)
SNU/KMA	GCPS T63 L21	SNU SST forecast	6	Kug *et al*. (2008b)
UH	ECHAM4 T31 L19	SNU SST forecast	10	Roeckner *et al*. (1996)

$$Y(t) = \frac{1}{N}\sum_{i=1}^{N} a_i F_i(t)$$

where a_i denotes a weighting coefficient for the ith model. The weighting coefficients are obtained from applying the SVD method to calculate the regression coefficients for a set of different model forecasts. In this SVD method, if the singular value is less than a minimum criterion ($1\text{x}10^{-5}$), the singular mode is not used. However, the minimum criterion is not much sensitive to the result. Refer to Yun *et al*., (2003) for details of the algorithm for obtaining the coefficients. They have shown the SVD method to be comparable to or somewhat better than the multiple regression-based scheme for seasonal climate prediction (Krishnamurti *et al*., 1999).

(3) MME-SPPM

This method involves a simple composite method after statistical downscaling of the prediction of each dynamical model based on a statistical model using the SPPM proposed by Kug *et al*., (2008a) as follows:

$$Y(t) = \frac{1}{N}\sum_{i=1}^{N} \hat{F}_i(t)$$

where $\hat{F}_i$ is a corrected forecast of the ith model obtained through the SPPM procedure. Because every dynamical model has a systematic bias, its predictive skill can be improved by reducing the systematic error. The SPPM is a statistical model which is designed to correct any systematic bias in dynamical prediction. After applying the SPPM to individual models separately, the MME prediction is obtained by the equal-weighting simple composite of the corrected individual predictions.

The SPPM is a kind of pointwise regression model. The predictor of the model is a pattern of predicted precipitation in a certain domain and the predictand is precipitation at each grid point over the global domain.

The main idea is to predict the predictand at each grid by projecting the predictor field onto the covariance pattern between the large-scale predictor field and the one-point predictand. The model equation is as follows:

$$\begin{aligned} &PRCP(t) = \alpha \cdot P(t) \\ &\alpha = \frac{\frac{1}{T}\sum_{t}^{T} PRCP(t) \cdot P(t)}{\left[\frac{1}{T}\sum_{t}^{T} P^2\right]} \\ &P(t) = \sum_{x}^{D} COV(x) \cdot \Psi(x,t) \\ &COV(x) = \frac{1}{T}\sum_{t}^{T} PRCP(t) \cdot \Psi(x,t) \end{aligned} \tag{1}$$

where x and t denotes spatial and temporal grids, respectively. The covariance pattern (*COV*) is calculated between the observed predictand, *PRCP(t)*, and predictor field, $\Psi(x,t)$, in certain domain (*D*). The covariance pattern indicates a pattern of the model prediction which is related to the observed predictand. The variable, *P,* indicates a projected time series from the covariance pattern and predictor field from model prediction, $\Psi(x,t)$. The parameter, α, is a regression coefficient of the projected time series, *P*, on the predictand during a training period, *T*. In this study, we used a training period of 20 years in the cross-validation process.

In this statistical prediction, selection of the predictor domain (*D*) plays a crucial role in the predictive skill. In general, traditional pattern projection models use a fixed geographical domain whose location and size are fixed for the predictor during the whole forecast period. This method seems appropriate for regional climate prediction, where the number of predictands is limited. However, when the prediction target covers a wide region, so that the number of predictands is large, as is the case in the present study, it is difficult to subjectively choose the predictor domain. Therefore, a method is required to facilitate the objective selection of the prediction domain. In the SPPM process, the optimal predictor domain is automatically selected with an objective criterion.

The SPPM consists of two steps to obtain the final prediction. The first step is to select the predictor domain, and the second step is a prediction by the pattern projection in Eq. (1). In the first step, to select the predictor domain, correlation coefficients between the predictand and precipitation (predictor variable) at each grid are calculated to seek out a possible predictor domain. Among them, some grid points, showing significantly higher correlation, are selected as predictor grid points. In this case, the selected grids may be split into several regions. In the traditional statistical model, only one geographical domain, including significant grids, is selected as a predictor domain, while other grids outside the selected domain are not used. However, in the SPPM, all significant grid points are gathered and a reconstructed domain is constructed by lining up the selected grid points. The reconstructed domain is regarded as a predictor domain (D). Using the selected predictors, the SPPM produces a corrected forecast based on Eq. (1). For details on the SPPM procedure, refer to Kug *et al.*, (2008a).

(4) MME-SPPM2

This method is a modified version of the MME-SPPM. In the MME-SPPM, all participant models are used for a simple composite at each grid after each individual model is corrected by the SPPM. However, in the MME-SPPM2 some model predictions, which have poor skill, are excluded from the simple composite. That is, only qualified model predictions that have predictable skill are used for the multi-model ensemble at each grid point. Because poor models may degrade the skill of the multi-model ensemble, the MME-SPPM skill can be improved by removing the predictions of the relatively poor models. In order to know whether each model prediction has predictable skill or not, double-cross-validation method (Feddersen *et al.*, 1999) is used in the SPPM procedure. The double-cross-validation is outlined below for the statistical prediction at time t_0.

i) Make 20-yrs dataset by excluding t_0 data from 21-yrs data.

ii) Apply the SPPM with the cross-validation procedure. In this case, the training period will be 19 years.

iii) Calculate a hindcast skill for 20 years when the t_0 data are excluded.

iv) Repeat i)-iii) process for the other all t_0.

In the double-cross-validation process, the hindcast skill of the individual model at each grid and each prediction year is calculated. If the correlation skill during the training period is higher than a threshold value, the model prediction is selected for the multi-model ensemble. In this study, we used a correlation 0.3 as the threshold value. We checked the results with different values (0.4, 0.2, and 0.1) of the threshold, but the main results were not changed.

3. Results

The MME-SPPM2 method and another three MME schemes have been evaluated using 15 climate models' retrospective forecasts for the period 1981-2001. Figure 1 shows the mean correlation skill of 15 individual models (hereafter referred to as "single model correlation") and the correlation skill for MME with equal weighting (MME-EW), MME with the SVD method (MME-SVD) and simple MME after applying statistical correction (MME-SPPM and MME-SPPM2). All MME methods have a better correlation skill than the single model, indicating that a MME approach is an effective way to improve current climate seasonal prediction. The

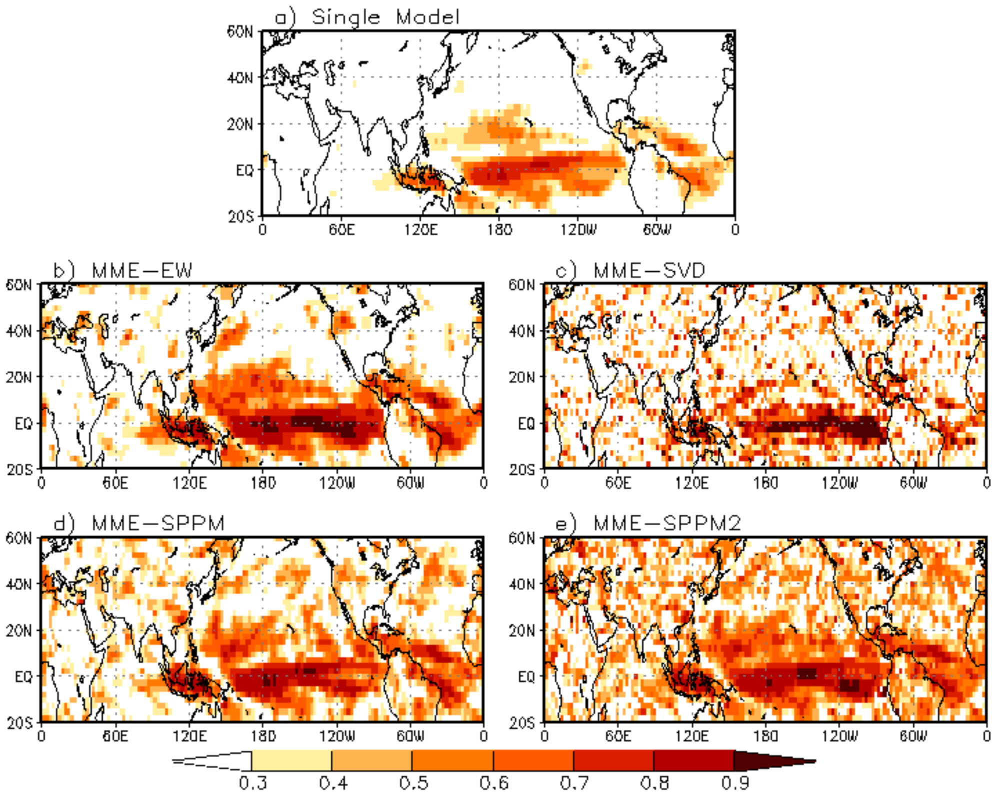

Fig. 1. a) Mean correlation skill of 15 individual models for JJA precipitation anomaly during the period 1981-2001, and correlation skill of MME by b) MME-EW, c) MME-SVD, d) MME-SPPM, and e) MME-SPPM2.

area-averaged skill over the global domain between 60°S and 60°N is 0.12 for the single model, 0.24 for the MME-EW, 0.24 for the MME-SVD, 0.31 for the MME-SPPM, and 0.38 for the MME-SPPM2.

Over the tropical Oceans, it is difficult for the sophisticated methods to beat the skill of the MME-EW, the simplest scheme. Over the continental region and extratropical oceans, however the MME-SPPM provides an improved skill to some extent. It is because the current dynamical models are capable of capturing large scale pattern related to a local variability over the continental region and extratropical oceans, though they have difficulty in predicting a local variability correctly at each grid point. It is shown that the skill of the MME-SPPM2 is significantly improved compared to that of the MME-SPPM. We found that the difference between two MME-SPPMs is significant when the MME-SPPM has poor skill (not shown). Note that the MME-SPPM uses the simple composite after the statistical correction. This implies that the simple composite method may have a problem when the skills of individual models are poor and diverse. In this sense, the MME-SVD should be the best candidate because the weighting coefficients are determined based on the skill of each model. However, the actual predictive skill of the MME-SVD seems worse than those of other MME methods, and the correlation pattern looks quite random. Because the weighting coefficients are calculated from relatively small prediction sample (20 years) and the number of models is large, the coefficients are not statistically stable in the cross-validation process. This brings about the so-called overfitting problem in the training process. Therefore, the forecast skill is degraded and correlation pattern can be localized. We have concluded that the weighted ensemble method is not appropriate to seasonal climate prediction data, unlike medium-range prediction which can use a lot of sampling data (c.f. Krishnamuti *et al*., 1999). Kharin and Zwiers (2002) also derived the similar conclusion using 10 different models.

It was also found that the major improvement of the MME-SPPM and MME-SPPM2 is achieved against MME-EW over the regions in which the single model skill is poor (See Fig. 1a). Figure 2 shows the averaged skill of each MME method according to the mean correlation skill of individual model over most of the global grids over 60°S-60°N. Note that the bin size of the x-axis is 0.04 and only the average value at each bin range of the x-axis is marked. It is shown that the MME-EW is comparable to the MME-SPPM2 and better than the other MME methods over the region where the averaged skill of individual models is over 0.2. This indicates that if the individual model has a certain level of prediction skill, the simple composite might be enough. However, the skill of the simple method is apparently degraded if the single model correlation is less than 0.2. In this case, the MME-EW becomes the worst method. Noted that the MME-EW is even worse than the single-model prediction, when the averaged correlation skill of the single models is less than zero. This indicates that this simple composite MME method works only when the single models has some predictive skill.

On the other hand, other MME methods still have

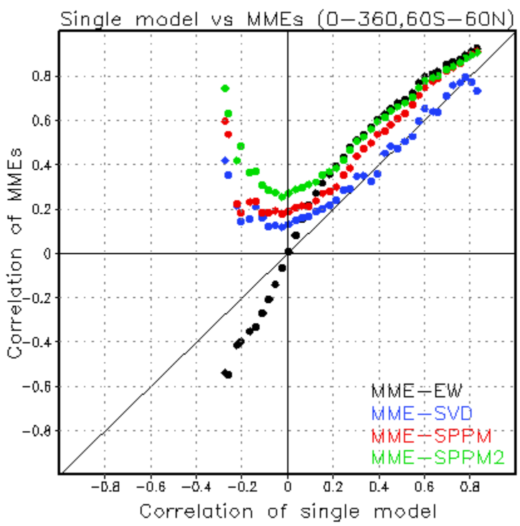

Fig. 2. Scatter diagram between averaged correlation skill of single models (x-axis), and correlation skill of each MME method (y-axis). The bin size of x-axis is 0.04 and only the average value at each bin range of x-axis is marked. Black, blue, red, and green circles denote the skill of MME-EW, MME-SVD, MME- SPPM, and MME-SPPM2 method, respectively.

predictive skill, even though the individual models have no predictive skill. This is because these MME methods have an algorithm to reduce any systematic bias, unlike the MME-EW. It is interesting that the skills of these methods is significantly higher as the correlation skill of a single model is negatively stronger, because the strong negative correlation means that the poor skill is contributed not by random bias but by anomalous systematic bias. It is clear that the MME-SPPM2 always has the best skill. In particular, the MME-SPPM2 has a significantly higher skill when the single models have poor skills, because it excludes some poor models' contributions to the MME. Therefore, it can be concluded that the MME- SPPM2 is the optimal MME method for at least the present hindcast data.

Why does the MME-SPPM2 have better skill? It might result from removing untrustworthy predictions before making MME. Figure 3 shows the averaged number of model which is contributed to the final MME-SPPM2 prediction for 21-years prediction. It is interesting that there is a spatial similarity between Fig. 3 and Fig. 1d, indicating that the region with high correlation skill corresponds to the region with high number of the reliable predictions. Note that the only difference between the MME-SPPM and MME-SPPM2 is either using all predictions or using just reliable predictions. Over the extratropics, only a few models show a reliable prediction, so degradation of the MME skill is inevitable if all of predictions are used for MME. However, just using only reliable predictions, the MME-SPPM2 is capable to improve skill over many regions where the number of reliable prediction is very small. In addition, we found the MME skill increases as a number of reliable prediction increases. Therefore, it seems that there are still room to improve the forecast skill over continental region and extratropical Ocean regions through adding reliable predictions more in future.

In the present study, we used a relatively large number of prediction models. Therefore, our results can change as the number of models for the MME is changed. To check this, the dependency of MME skill on the number of models being used for the MME was evaluated. To obtain the mean skill when the model number used for MME is N, N models among 15 are

[1] If all possible cases are less than 500, we used all possible combinations.

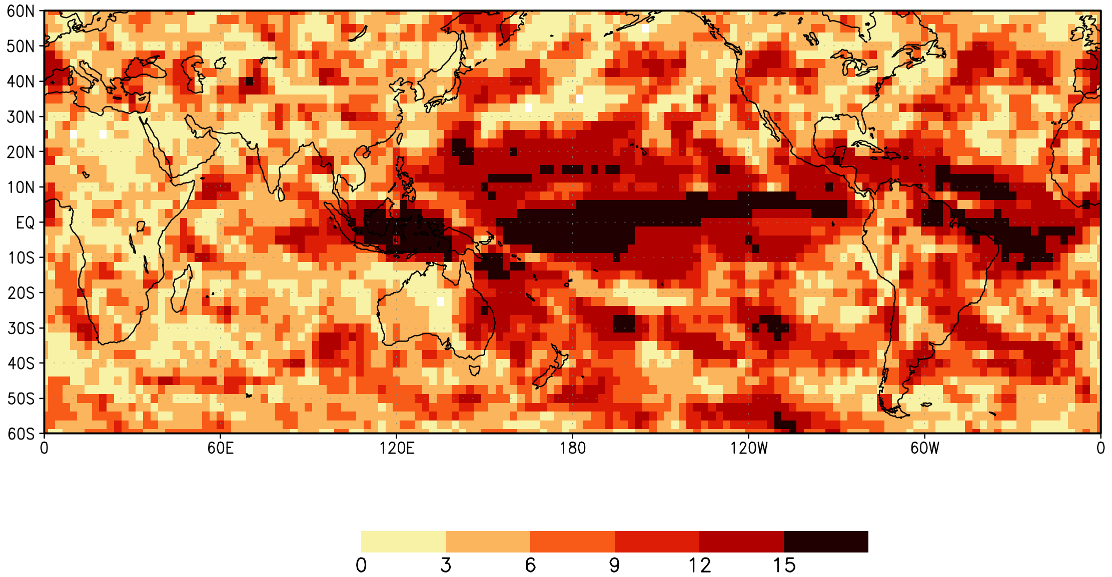

Fig. 3 The number of model which is contributed to the MME-SPPM2 prediction. The number is averaged for 21-years prediction.

randomly selected and its MME skill is calculated. After this selection is repeated 500 times[1)], the mean skill for the 500 cases is determined at each grid point. Figure 4 shows the mean correlation skill with various model numbers. For simplicity, the correlation skill is area-averaged over the globe (0^{o}-360^{o}E, 60^{o}S-60^{o}N). It is clear that the MME skills are higher as the number of models increases, because the reduction of systematic and random errors is more effective. It is interesting that the MME-EW and MME-SPPM show very similar sensitivities which are characterized by rapid saturation of skill after more than 8 models are used. It turns out that the MME-SPPM2 scheme is the most effective for MME prediction with a large number of models. It is quite interesting that the skill of the MME-SPPM2 seems to be still increasing even when all 15 models are used. Therefore it is expected that the skill of the MME-SPPM2 can be further improved by increasing the model number. However, if the number of models for MME prediction is smaller than 5, the skill of the MME-SPPM is better than that of MME-SPPM2. This indicates the MME-SPPM2 is the optimal method when the number of participant models for MME is large.

4. Summary and Discussion

During the last decade, a lot of effort has been devoted to extracting optimal and skillful MME prediction from available seasonal climate predictions by using either simple or sophisticated statistical methods. While sophisticated MME schemes tend to produce better forecast skills than simple MME with equal weighting (MME-EW) during training period, they have difficulty in beating MME-EW for independent or real-time forecast (over-fitting problem). In this study, we propose a new optimal MME method and identify the strength and disadvantage of existing and the newly invented MME schemes.

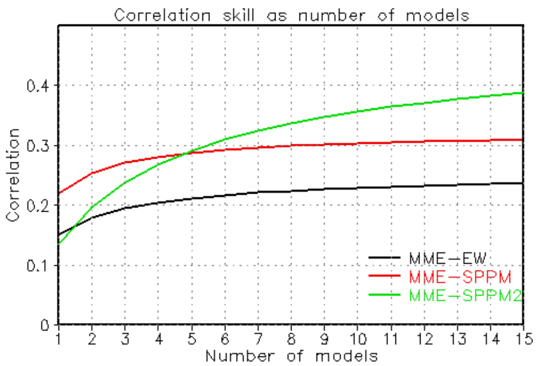

Fig. 4. The dependence of MME prediction skill on the number of model averaged over the globe (0-360E, 60S-60N). Black, red, and green lines indicate the skill of MME-EW, MME-MME-SPPM, and MME-SPPM2, respectively.

A new MME technique, named the MME-SPPM2 has been proposed and evaluated for seasonal climate predictions of 15 models. To assess the value of the MME-SPPM2 in use, it has been compared with two existing methods, the MME-EW and MME-SVD, and pre-version of the new MME scheme named the MME-SPPM. When it comes to the seasonal prediction for the summer mean precipitation, the MME-SPPM2 provides better skill than other MME methods over the continental regions and extra-tropical ocean. However, it is hard for the MME-SPPM2 to defeat the skill of the MME-EW over the equatorial Pacific region. It is noted that if the averaged skill of individual models is higher than 0.2, the MME-EW is comparable to the MME- SPPM2 and better than the other MME methods. Otherwise, the prediction skill is conspicuously improved by using the MME-SPPM2. Also, it turns out that the MME-SPPM2 works most effectively when a larger number of models are used for MME.

From the aforementioned results, we demonstrate that the MME-SPPM2 is the optimal MME method, especially when the number of models being used is large. Nevertheless, its forecast skill for the summer mean precipitation is still inadequate to provide useful information to the public, especially over the continental region, and further improvement is necessary.

How can we improve the current seasonal prediction and MME-SPPM2? We propose three possible ways. Firstly, further improvement of a dynamical model itself is essential. The resolution and physical parameterization are crucial factors in any dynamic model's performance. Also, initialization in the atmosphere, land, and ocean coupled system will be big challenge in seasonal climate prediction.

Secondly, we suggest that a retrospective forecast period should be extended. Most seasonal prediction studies and operational prediction systems have been based on retrospective hindcast data for the recent two decades. However, these short hindcast data cannot produce a stable statistical relation in statistical post-processing because the statistical model cannot separate systematic bias from random bias correctly (Kug *et al.*, 2008a). Therefore, the post-processing cannot be effective. This problem is easily solved by just increasing the hindcast sample (see Kug *et al.*, 2008a). Thirdly, it is suggested to increase the number of participant models in MME prediction. Increasing trend in Fig. 4 indicates that the forecast skill of the MME-SPPM2 will be possibly improved with more predictions in addition to 15 predictions especially over the extratropical continental regions where few predictions are reliable among 15 predictions (Fig. 3). Including more participant predictions is expected to improve the MME skill further.

Acknowledgements. This work is supported by APEC Climate Center as a part of APCC international project (CliPAS). I.-S. Kang was supported by the SRC program of Korea Science and Engineering Foundation, and Brain Korea 21 Project.

REFERENCE

Anderson, D. L., and Coauthors, 2003: *Comparison of the ECMWF seasonal forecast system 1 and 2, including the relative performance for the 1997/8 El Nino.* ECMWF Technical Memorandum, 404, 93 pp.

Chakraborty, A., and T. N. Krishnamurti, 2006: Improved seasonal climate forecasts of the South Asian summer monsoon using a suite of 13 coupled ocean-atmosphere models. *Mon. Wea. Rev.,* **134**, 1697-1721.

Cocke, S., and T. E. LaRow, 2000: Seasonal predictions using a coupled ocean atmospheric regional spectral model. *Mon. Wea. Rev.*, **128**, 689-708.

Feddersen, H., A. Navarra, and M. N. Ward, 1999: Reduction of model systematic error by statistical correction for dynamical seasonal prediction. *J. Climate*, **12**, 1974-1989.

Fu, X., and B. Wang, 2004: The boreal-summer intraseasonal oscillations simulated in a hybrid coupled atmosphere-ocean model. *Mon. Wea. Rev.*, **132**, 2628-2649.

Kang, I.-S., J.-Y. Lee, and C.-K. Park, 2004: Potential predictability of summer mean precipitation in a dynamical seasonal prediction system with systematic error correction. *J. Climate,* **17**, 834-844.

Kharin V., and F. Zwiers, 2002: Climate predictions with multimodel ensembles. *J. Climate,* **15**, 793-799.

Krishnamurti, T. N., and Coauthors, 1999: Improved weather and seasonal climate forecasts from multi-model superensemble. *Science,* **285**, 1548-1550.

Kug, J.-S., J.-Y. Lee, and I.-S. Kang, 2007: Global sea surface temperature prediction using a multi-model ensemble. *Mon. Wea. Rev.*, **135**, 3239-3247.

_____, I.-S. Kang, and J.-Y. Lee, 2008a: Systematic bias correction of dynamical seasonal prediction using a stepwise pattern projection method. *Mon. Wea. Rev.*, in press.

_____, _____, and D. H. Choi, 2008b: Seasonal climate predictability with Tier-one and Tier-two prediction system. *Climate Dyn.*, doi:10.1007/s00382-007-0264-7.

Palmer, T. N., and Coauthors 2004: Development of a european multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bull. Amer. Meteor. Soc*, **85**, 853–872.

Peng, P., A. Kumar, A. H. Van den Dool, and A. G. Barnston 2002: An analysis of multi-model ensemble predictions for seasonal climate anomalies. *J. Geophys. Res.*, **107**, 4710 doi:10.1029/2002JD002712.

Roeckner, E., and Coauthors, 1996: *The atmospheric general circulation model ECHAM4: Model description and simulation of present-day climate.* Technical Report, 218, Max Planck Institute for Meteorology, Hamburg, Germany, 90 pp.

Saha, S., and Coauthors, 2006: The NCEP climate forecast system. *J. Climate*, **19**, 3483-3517.

Wang, B., and Coauthors, 2008a: How accurately do coupled climate models predict the Asian-Australian monsoon interannual variability? *Climate Dyn.*, doi:10.1007/s00382 -007-0310-5.

_____, and Coauthors, 2008b: Advance and prospectus of seasonal prediction: Assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980-2004). *Climate Dyn.*, in press.

Yun, W.-T., L. Stefanova, and T. N. Krishnamurti 2003: Improvement of the superensemble technique for seasonal forecasts. *J. Climate,* **16**, 3834-3840.