Toward enhancement of prediction skills of multimodel ensemble seasonal prediction: A climate filter concept

Doo Young Lee,[1,2] Karumuri Ashok,[3] and Joong-Bae Ahn[2]

Received 10 June 2010; revised 5 January 2011; accepted 18 January 2011; published 29 March 2011.

[1] Using the APEC Climate Center (APCC) operational multimodel ensemble (MME) hindcasts of precipitation and temperature at 850 hPa for boreal winters for the period 1981–2003, along with those of the individual models as well as corresponding observed and reanalyzed data, we propose the use of a "climate" filter to diagnose and improve the prediction skills. The "filter" is based on the observed strong association between the El Niño–Southern Oscillation (ENSO)-associated Walker circulation and the tropical Pacific rainfall. The reproducibility of this relationship is utilized to evaluate the fidelity of the models. It is found that the retrospective forecast skill of a newer type of MME that contains only the "more skillful" models is superior to that of the all-inclusive operational MME. The difference of the prediction skills between the "more skillful" and "less skillful" MMEs varies with the region and is significant in subtropics such as East Asia, while most of the models perform well in tropics adjacent to the Pacific. Our pilot forecast with the proposed MME for two boreal winter seasons indicates that the method generally works better than the all-inclusive MME in many of the target regions.

Citation: Lee, D. Y., K. Ashok, and J.-B. Ahn (2011), Toward enhancement of prediction skills of multimodel ensemble seasonal prediction: A climate filter concept, *J. Geophys. Res.*, *116*, D06116, doi:10.1029/2010JD014610.

1. Introduction

[2] The multimodel ensemble (MME) techniques are known to be a useful and practical approach for alleviating the inherent errors contained in individual models and naturally offer better predictability and realistic features because of the ability to isolate observed outcomes from multiple models and reduce individual model-specific systematic errors [*Krishnamurti et al.*, 1999, 2000; *Palmer et al.*, 2000; *Shukla et al.*, 2000; *Barnston et al.*, 2003, *Hagedorn et al.*, 2005; *Doblas-Reyes et al.*, 2000]. Indeed, several studies [*Krishnamurti et al.*, 2000; *Palmer et al.*, 2000; *Pavan and Doblas-Reyes*, 2000; *Peng et al.*, 2002; *Shukla et al.*, 2000; *Hagedorn et al.*, 2005; *Doblas-Reyes et al.*, 2005; *Yun et al.*, 2005; *Wang et al.*, 2008, 2009; *Min et al.*, 2009; *Lee et al.*, 2010] have reported that, in general, the performance skill of the MME is higher than that of the constituent individual models. Encouraged by such results, nowadays, a number of meteorological/climate prediction centers worldwide operationally implement dynamical MME seasonal prediction, often with individual models selected through some in-house quality control, which typically looks for problems such as spurious or missing values, etc. [*Palmer et al.*, 2004; *Lee et al.*, 2009].

[3] However, it is found from operational experience that for some seasons/regions, the MME prediction skills are also relatively limited (see Figure A1 in Appendix A). This means that there is further scope to improve the MME prediction skills.

[4] In the current study we propose that reproducibility of an important observed climate signal, such as rainfall or atmospheric response to El Niño–Southern Oscillation (ENSO) in the tropical Pacific which we choose as an example, can be used as a "climate" filter to evaluate individual constituent models. We further demonstrate that an MME with constituent models, which have such fidelity, performs better as compared to the MME containing all the available models. This research aims to facilitate a new approach module for MME prediction using the relative dynamical diagnostic performance within the model itself, not directly using external forcing of model.

[5] The rest of this paper is organized as follows. Section 2 describes the observed and individual model data used, general methodology, and also introduces a "climate filter" concept to grade the individual models. In section 3 the relative performances of the various MME are presented. Concluding remarks can be found in section 4.

2. Data Used and the "Climate" Filter

2.1. Data Used

[6] In the present study we use the National Centers for Environmental Prediction (NCEP)-Department of Energy (DOE) reanalysis 2 [*Kanamitsu et al.*, 2002] and Climate Prediction Center Merged Analysis of Precipitation (CMAP) [*Xie and Arkin*, 1997] for the boreal winter seasons (December

[1]APEC Climate Center, Busan, Korea.
[2]Department of Atmospheric Sciences, Pusan National University, Busan, Korea.
[3]Centre for Climate Change Research, Indian Institute of Tropical Meteorology, Pune, India.

0148-0227/11/2010JD014610

Table 1. Description of the General Circulation Models Used

Member Economies	Acronym for Model	Organization	Model Resolution
Australia	POAMA	Bureau of Meteorology Research Centre	T47L17
Canada	MSC_GEM	Meteorological Service of Canada	2° × 2° L50
	MSC_GM2		T32L10
	MSC_GM3		T63L32
	MSC_SEF		T95L27
Chinese Taipei	CWB	Central Weather Bureau	T42L18
Korea	GCPS	Seoul National University	T63L21
	GDAPS_F	Korea Meteorological Administration	T106L21
	NIMR	National Institute of Meteorological Research	5° × 4° L17
United States	NCEP	National Centers for Environmental Prediction	T62L64

through February, hereinafter DJF) for the period 1981–2003 as observations. In order to facilitate a new approach method for MME prediction through the relative performances and error analysis of individual climate models, we adopt 10 operational seasonal prediction model hindcast outputs (see Table 1) for the same period, which, in general, constitute a major portion of the operational MME of the APEC Climate Center (APCC), which, as of today, is the most extensive operational MME setup. For this study we also adopt a simple composite method [*Peng et al.*, 2002; *Kang et al.*, 2009; *Lee et al.*, 2009] as one of several APCC operational MME techniques such as the multiple regression based on the training period [*Krishnamurti et al.*, 2000], the synthetic multimodel ensemble based on EOF-filtered data [*Yun et al.*, 2005], and the stepwise pattern projection method [*Kug et al.*, 2008]. In this technique, equal weights are assigned to the ensemble mean predictions of each of the models. Hereinafter, MME means a simple composite method, unless otherwise specified.

[7] The MME results are generated by the combination of bias corrected model forecast anomalies as a simple arithmetic mean of predictions based on individual models [*Peng et al.*, 2002; *Kang et al.*, 2009; *Lee et al.*, 2009] and the verification of the hindcast can be evaluated as an unbiased performance of models.

[8] Finally, we explore the usefulness of the method by carrying out a pilot forecast with a new approach method for improving MME prediction for two boreal winter seasons of December 2008 to February 2009 and December 2009 to February 2010, for which seasons the participant model forecast data are available.

2.2. Climate Filter

[9] It is well known that the ENSO is the most important climate phenomenon that drives climate variability in the tropics and beyond. The ENSO-related zonal circulation is the major factor for rainfall variability in the tropical Pacific, as seen from the temporal DJF correlation patterns between the local Walker circulation (calculation method follows from next paragraph) and local precipitation (Figure 1) for periods 1981–2003. Figure 1 attests to this as seen from the high correlations, in general, along 10°S–10°N and particularly with magnitudes of more than 0.8 in the central and western tropical Pacific. This implies that for any model to predict the rainfall well, simulation of the ENSO-associated Walker circulation in the tropical Pacific, and importantly, its asso-

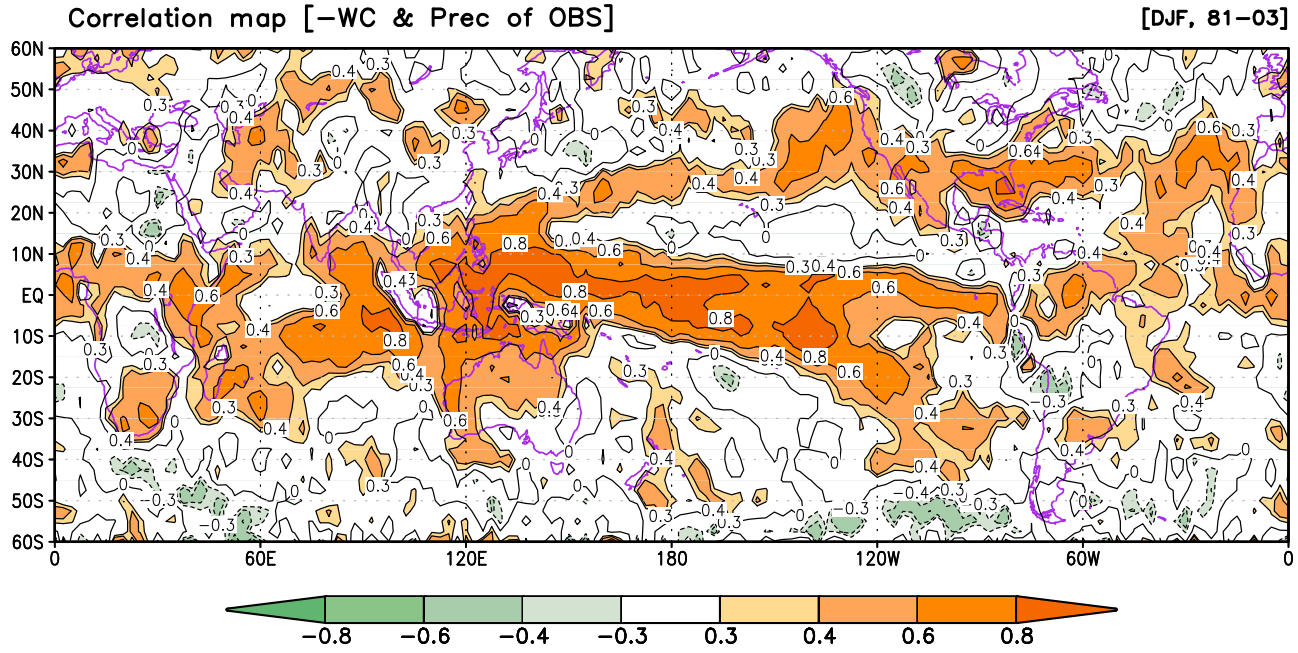

Figure 1. Observed temporal boreal winter correlation patterns between the local Walker circulation and precipitation during the period of 1981–2003.

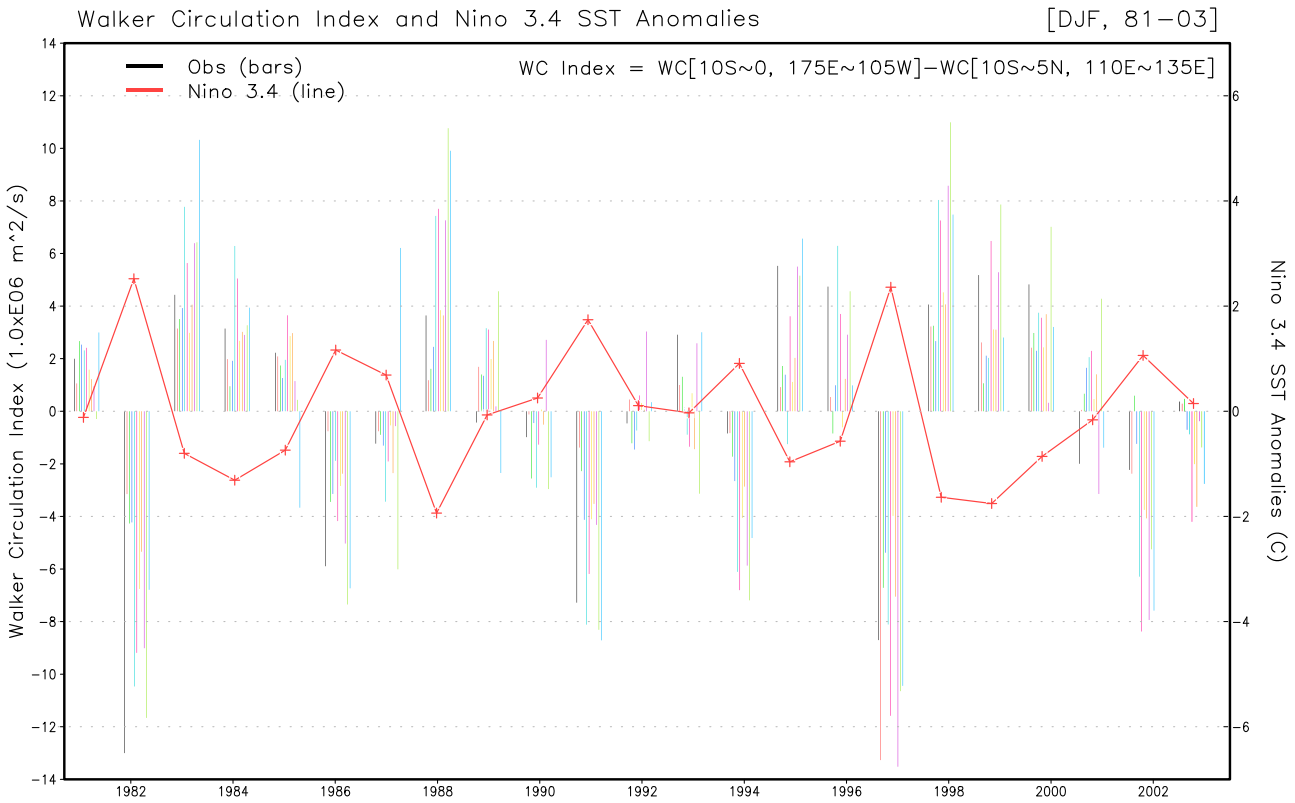

Figure 2. Comparison of the Niño 3.4 SST anomalies (solid red line) with the Walker circulation index (solid bars) of individual models, which is defined by difference between the tropical eastern Pacific (10°S–0°, 175°E–105°W) and the tropical western Pacific (10°S–5°N, 110°E–135°E) removing the zonal mean from the anomalous velocity potential at 200 hPa. The observed Walker circulation index (solid black bars) is also represented.

ciation with local rainfall, is the minimum requirement and thus an important measure of model fidelity for predicting the tropical rainfall. We utilize this concept, what we refer to as a climate filter, to grade the individual constituent models.

[10] To calculate the Walker circulation, we follow a method by *Tanaka et al.* [2004], who divide the global divergent field, represented by 200 hPa velocity potential, into contributions from the Hadley, Walker, and monsoon circulations. *Tanaka et al.* [2004] demonstrate that the time series of this Walker circulation correlates significantly well with the ENSO variations. We do not, however, consider the monsoon contribution, defined as the deviation from the 12 month running mean, as we deal with only a single season. Instead, to understand the interannual variability of Walker circulation during boreal winter better, we use seasonal anomalies obtained by subtracting the seasonal climatology based on the period 1981–2003 [*Wang*, 2002]. The distribution of our Walker circulation index is consistent with the ENSO variation represented by the Niño3.4 SST anomalies (Figure 2), and also similar to a Walker circulation index of *Wang* [2002] that uses 500 hPa vertical velocity anomalies. We calculate the difference of the Walker circulation between the tropical eastern Pacific (10°S–0°, 175°E–105°W) and the tropical western Pacific (10°S–5°N, 110°E–135°E) as an index of the Walker circulation. The correlation of the observed Walker circulation indices with the Niño3.4 SST anomalies is −0.92 and is significant at the 99% confidence level from two-tailed Student's *t* test.

[11] Briefly, the Walker circulations using 200 hPa velocity potential are given by the following: (1) The climatological seasonal mean field ($\overline{\chi}$) is removed from the velocity potential at 200 hPa (χ),

$$\chi'(t,x,y) = \chi(t,x,y) - \overline{\chi}(x,y), \tag{1}$$

(2) the zonal mean field ($[\chi']$) contains the information of the Hadley circulation, and (3) we obtain the Walker circulation by removing ($[\chi']$) from $\chi'(t, x, y)$,

$$\chi'^{*}(t,x,y) = \chi'(t,x,y) - [\chi'(t,y)]. \tag{2}$$

Table 2. Temporal Correlation Between Observed Niño3.4 Sea Surface Temperature Anomalies and Walker Circulation Index of Observation and Individual Models During the Period of 1981–2003

	Correlation
Observed	−0.92
Model 1	−0.76
Model 2	−0.81
Model 3	−0.91
Model 4	−0.89
Model 5	−0.95
Model 6	−0.93
Model 7	−0.92
Model 8	−0.90
Model 9	−0.94
Model 10	−0.81

Then we apply the squared correlation coefficient (r^2) between χ'* and Niño 3.4 index as the local weight for the Walker circulation χ'* field and obtain the local ENSO-associated Walker circulation χ'_{ENSO}*. (The magnitude of the squared correlation coefficient, also known as coefficient of determination, is useful because it gives the proportion of the variance (fluctuation) of one variable that is predictable from the other variable. It is the ratio of the explained variation to the total variation, and it is also one of the best means for evaluating the strength of the linear association between x and y. From this point, we believe that it is an appropriate weight to represent the variance of Walker circulation field associated with Niño 3.4 index.)

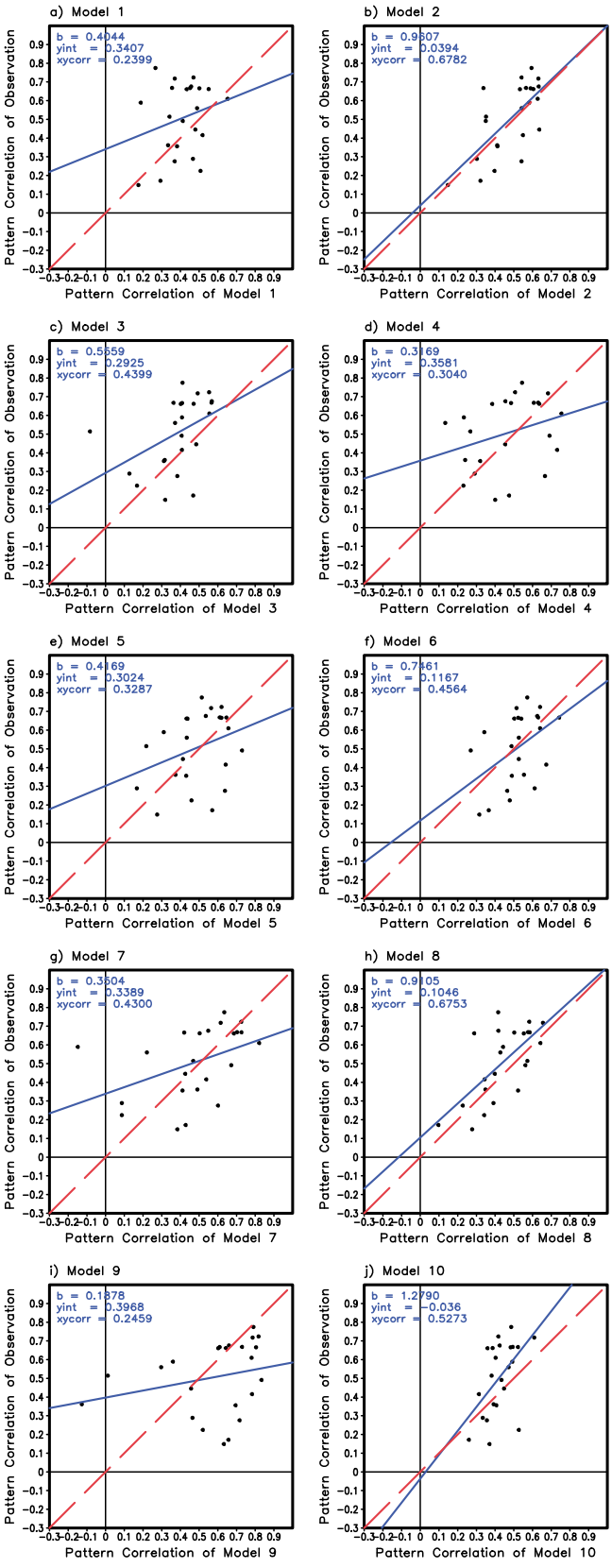

Figure 3. Observed pattern correlation between the ENSO-associated walker circulation and precipitation (y axis) over the tropical Pacific region (100°E~60°W, 10°S~10°N), for 23 boreal winter during the period 1981–2003, plotted against those from the individual models (x axis). The blue solid line is the statistical line of fit, and the red dashed line is a reference diagonal line. The slope "b" and the intercept "yint" from the regression line of fit are provided in the upper left. The "xycorr" stands for the temporal correlations between observation and each model.

[12] Further, from Figure 2 and Table 2 it is seen the fact that not only the observed Walker circulation, but also the hindcast Walker circulation of each model, are also very closely related to the observed ENSO forcing, indicating that the lower boundary SST forcings of most of the models (or predicted SSTs in coupled models) can apparently capture the ENSO-associated signals during the DJF seasons. On the basis of this interesting finding, we utilize the observed Niño3.4 index as weight to compute the ENSO-associated Walker circulation in the models.

[13] Further, we carry out a scatter diagram analysis to separate the better performing models from the rest. For this we first compute the pattern correlation [*Wilks*, 1995] between the observed ENSO-associated Walker circulation (χ'_{ENSO}*) and precipitation over the tropical Pacific (100°E–60°W, 10°S–10°N) during the DJF seasons of the study period. Similar pattern correlations are also computed for each model based on its hindcasts. The observed pattern correlations for each DJF season are plotted respectively against those from each model.

[14] The statistical significance of the correlations was computed based on the standard Student's two-tailed t test [*Spiegel and Stephens*, 2008; *Wilks*, 1995]. The simple number of degrees of freedom for the temporal correlation over the 23 year span is 21. Further effective spatial degree of freedom (ESDOF) [*Snedecor and Cochran*, 1980, *Bretherton et al.*, 1999, *Wang and Shen*, 1999] is applied to find the statistical significance for spatial pattern correlations. We follow the standard leave-one-out cross-validation method [*Michaelsen*, 1987; *Jolliffe and Stephenson*, 2003; *World Meteorological Organization*, 2006; *Kang et al.*, 2009] throughout the work while evaluating the hindcast skills. For example, we compute seasonal anomalies (of each model parameter as well as those from observations) from the corresponding climatological means that are obtained by excluding information from the target year for each year while carrying out simple composite method as well as while applying the climate filter method. In case of the climate filter, although more skillful models are not the same for all hindcast period because of different sampling year in each target year, we found that the top five most skillful models

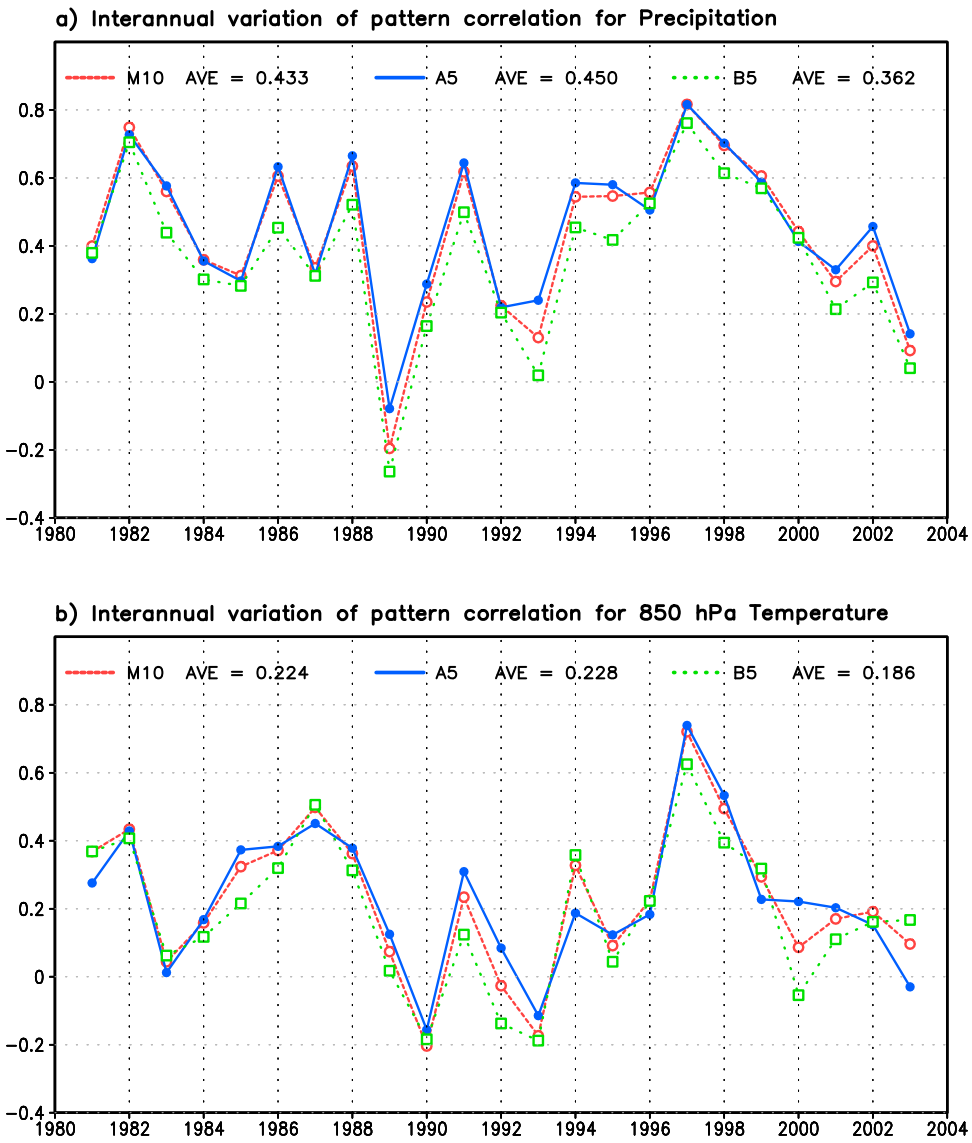

Figure 4. (a) Time series of the spatial pattern correlations between the observed and the predicted precipitation from M10 (dashed red line), A5 (solid blue line), and B5 (dotted green line) over the global region. M10, A5, and B5 are the multimodel ensemble predictions based on a simple composite method using the total of 10 models, the five more skillful models and the five less skillful models, respectively. (b) Same as Figure 4a but for the 850 hPa temperature.

cleared the criteria for the study period of 1981–2003 stably (table not shown).

3. Fidelity of the Hindcast Relationship Between ENSO-Associated Walker Circulation and Tropical Pacific Rainfall

[15] As mentioned earlier, we achieve the grading of the models through evaluation of the hindcast χ'_{ENSO}* rainfall relationship in the tropical Pacific. Scatter diagrams depicting the reproducibility of the relationship in the tropical Pacific for a "more skillful" model and a "less skillful" model are presented in Figure 3. Note that we have used two arbitrary conditions to grade the model as "more skillful" or "less skillful." The criteria are (1) that the slope of the fit between the observed and predicted pattern correlations should be larger than 0.5 and less than 1.5 and (2) that the temporal correlation coefficient between observation and model is more than 0.4, the ~90% statistical confidence level for the study period (~20 years) from a two-tailed Student's t test. On the basis of these criteria, we have found that five out of the 10 models cleared the criteria. We refer to these "more skillful" models, which successfully clear the "ENSO process" filter at least for the boreal winter season, as class "A" models. The rest of the models are referred to as class "B" models.

[16] We now conduct three types of MME hindcast experiments over the study period. In the first experiment, hereinafter referred to as the "A5," only the hindcasts from the class A models are used. In the second experiment we name as the "B5," only the class B models are used. In the third experiment, referred to as the "M10," hindcasts from

Table 3. Time Average of the Spatial Pattern Correlations Between the Observed and Various Multimodel Ensemble Hindcast Experiments for Precipitation and Those for Temperature at 850 hPa Over the Global Region[a]

Variables	Correlations		
	M10	A5	B5
Precipitation	0.433	0.450	0.362
850 hPa temperature	0.224	0.228	0.186

[a]Hindcast experiments are M10, A5, and B5. M10, A5, and B5 represent all 10 models, the five more skillful models, and the five less skillful models, respectively.

all the 10 models are used. Figures 4a and 4b show that the interannual variation of the spatial pattern correlations between the observed and predicted MME for precipitation and temperature at 850 hPa from all the three experiments over the global region, respectively. A quick look indicates that all the three MME runs result in a similar interannual variation. However, significant differences of the performance of predicted skill can be seen (see Table 3). Especially, it is noteworthy that the skill score of 0.45 (significant at 90% confidence level from two-tailed Student's *t* test; ESDOF = 15.6) for the rainfall in the A5 experiment is significantly higher than the corresponding skill score of 0.36 from the B5 (~ at 80% confidence level from two-tailed Student's *t* test; ESDOF = 16.2) in Figure 4a. The A5 skill is also superior to the skill from the all model-inclusive M10 MME experiment (significant at 85% confidence level from the two-tailed *t* test; ESDOF = 14.1). In fact, the skills for A5 are uniformly superior to those from B5 and M10 also for other variables such as velocity potential at 200 hPa and geopotential height at 500 hPa, etc. (figures not shown). Interestingly, the time averaged skills of all the MMEs for predicting the 850 hPa temperature are weaker than those for the corresponding skill score for the precipitation (Table 3), which needs further attention. It can be deduced that the skills of the M10 MME are essentially owing to those from the A5 models. It is apparent that the year-to-year variation in MME predicted skills for both of variables are sensitive to the concurrent ENSO strength, at least for the winters of 1982–1983 and 1997–1998, when the concurrent El Niño was strong (figure not shown [see *Wang et al.*, 2009, Figure 9 and 10]); the aspect of the enhanced prediction skills during the El Niño years possibly also needs more serious analysis.

[17] To visualize the sensitivity of the fidelity of various MME to the location of the region where the rainfall or temperature is predicted, in Figure 5 we present the time averages of pattern correlation coefficients between the observed and the predicted rainfall and temperature at 850 hPa for the five regions of East Asia (90°E–150°E, 20°N–50°N), Australia (110°E–160°E, 45°S–10°S), North America (50°W–140°W, 10°N–70°N), western North Pacific (120°E–180°E, 10°N–30°N), and tropical South America (30°W–90°W, 10°S–15°N), as examples. The first four regions also cover the subtropical through midlatitude regions while the last one is essentially in the tropics. In general, these results indicate the predicted skills by the A5 are better than those for M10 and B5 except for the precipitation and temperature at 850 hPa over the tropical South America region, where the M10 gives marginally higher skills but with an insignificant improvement. This better performance of M10 for tropical South America is essentially due to the relatively better performance of the B5 MME here as compared to its performance in the other regions. On the other hand, for the East Asia and western North Pacific, the difference of prediction skill for rainfall between A5 (0.29 and 0.44) and B5 (0.073 and 0.199) reaches a high value of more than 0.2. This fact indicates that the relatively poor performance of the M10 for these areas is mainly due to the poor prediction skills of the B5 models. The relative performances of the various MME in the other extratropical regions of Australia and North America are also similar to those for East Asia or western North Pacific. This result also supports the findings of *J.-Y. Lee et al.* [2011] and *S.-S. Lee et al.* [2011] that the forecast skill for seasonal precipitation over East Asia and the WNP region mainly

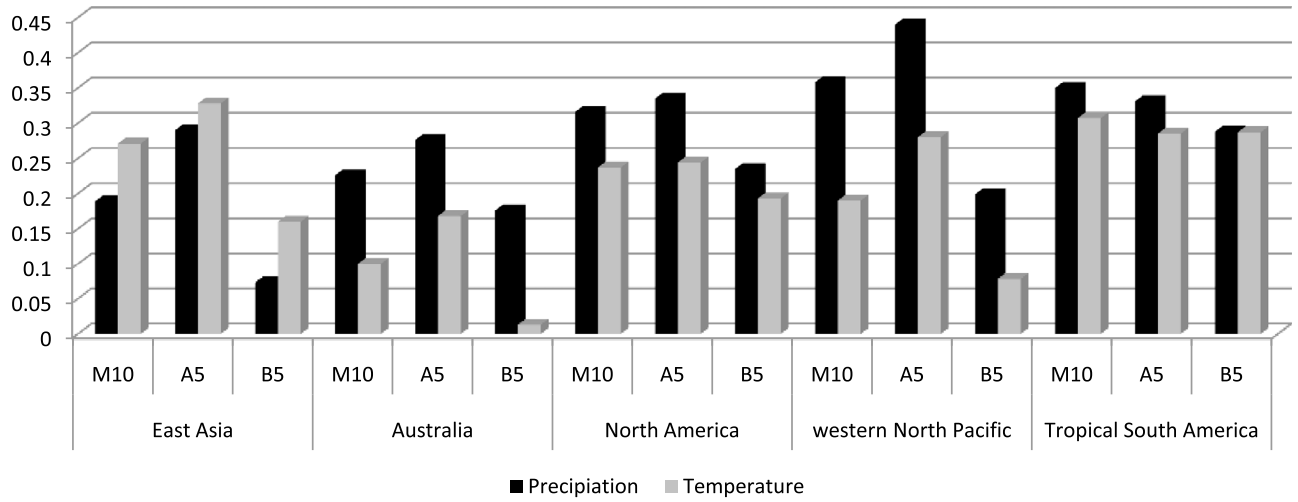

Figure 5. Time average of spatial pattern correlations between the observed and the predicted precipitations (solid bars) and those for the temperature at 850 hPa (light-shaded bars) over the five regions of the East Asia (90°E–150°E, 20°N–50°N), Australia (110°E–160°E, 45°S–10°S), North America (50°W–140°W, 10°N–70°N), western North Pacific (120°E–180°E, 10°N–30°N), and tropical South America (30°W–90°W, 10°S–15°N).

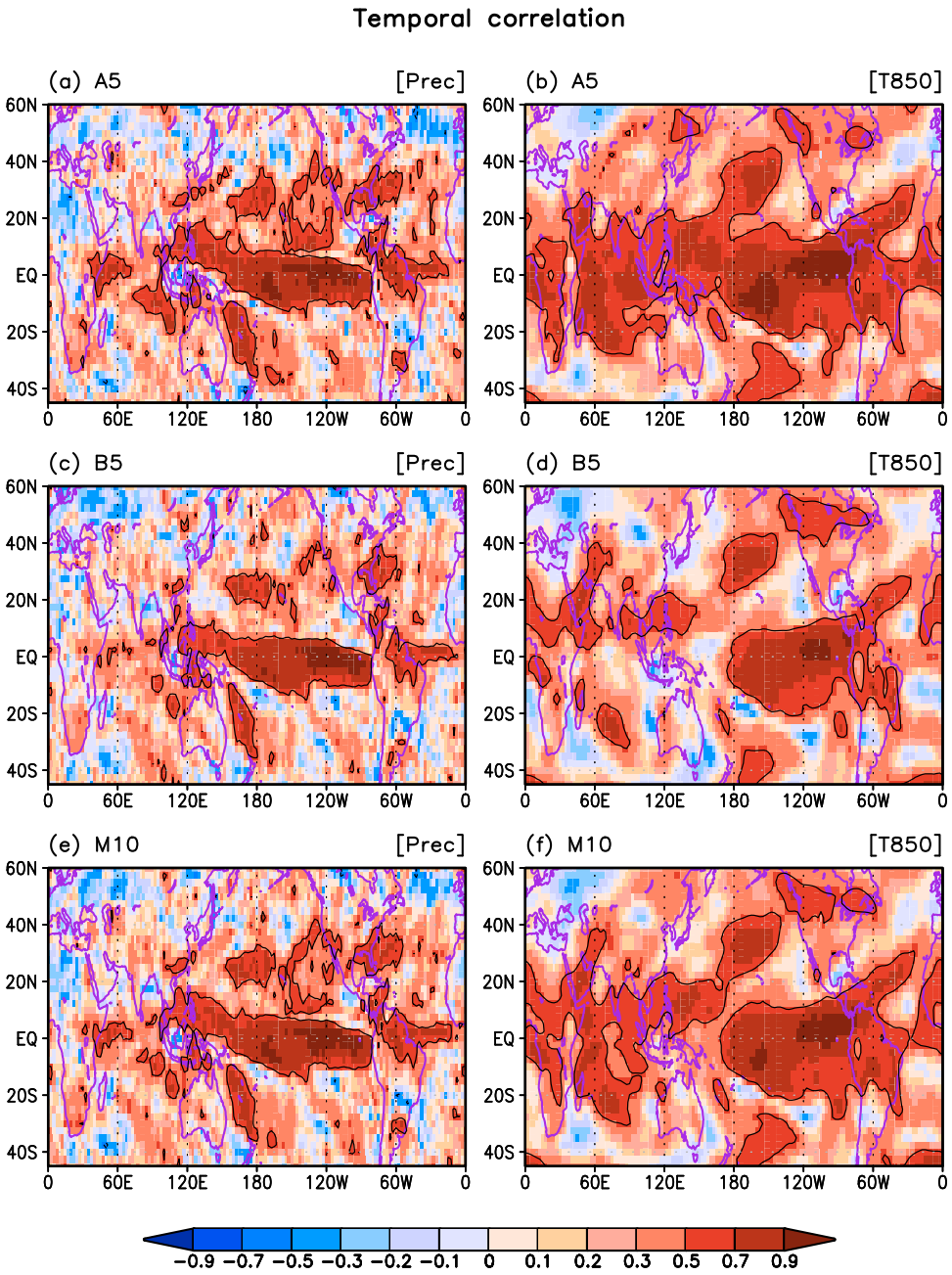

Figure 6. Temporal correlation coefficients between (a, c, and e) the observed and predicted precipitations and (b, d, and f) those for the temperatures at 850 hPa from A5 (Figures 6a and 6b), B5 (Figures 6c and 6d), and M10 (Figures 6e and 6f). The contour lines represent significant correlation coefficients at the 99% confidence level. The value is 0.526 from a two-tailed Student's *t* test.

comes from ENSO teleconnection. The decent performance of the B5 models for the tropical South America is probably due to a better fidelity in prediction of rainfall related to ENSO in this tropical region sitting next to the eastern Pacific. A similar result is also noted for the tropical Pacific (figures not shown).

[18] In Figure 6 we further examine the spatial distribution of all the MMEs' prediction skills for precipitation and temperature at 850 hPa in terms of temporal correlation coefficient at each gird point for the period of 1981–2003. The statistical significance of the correlation coefficients was computed using Student's two-tailed *t* test. The region of significant correlation at the 99% confidence level is outlined. In general, high prediction skills are confined to the global tropical regions, which are especially central and eastern tropical Pacific. Though the prediction skills of the A5 are significantly superior to those of B5 and M10 MME for both of variables over the whole global region, it is shown that the capture of the good skills is a little difficult in a few regions such as the Maritime continent, the northern part of Africa and the North Atlantic Ocean. We also find that the A5 prediction skills are considerably improved as compared to those

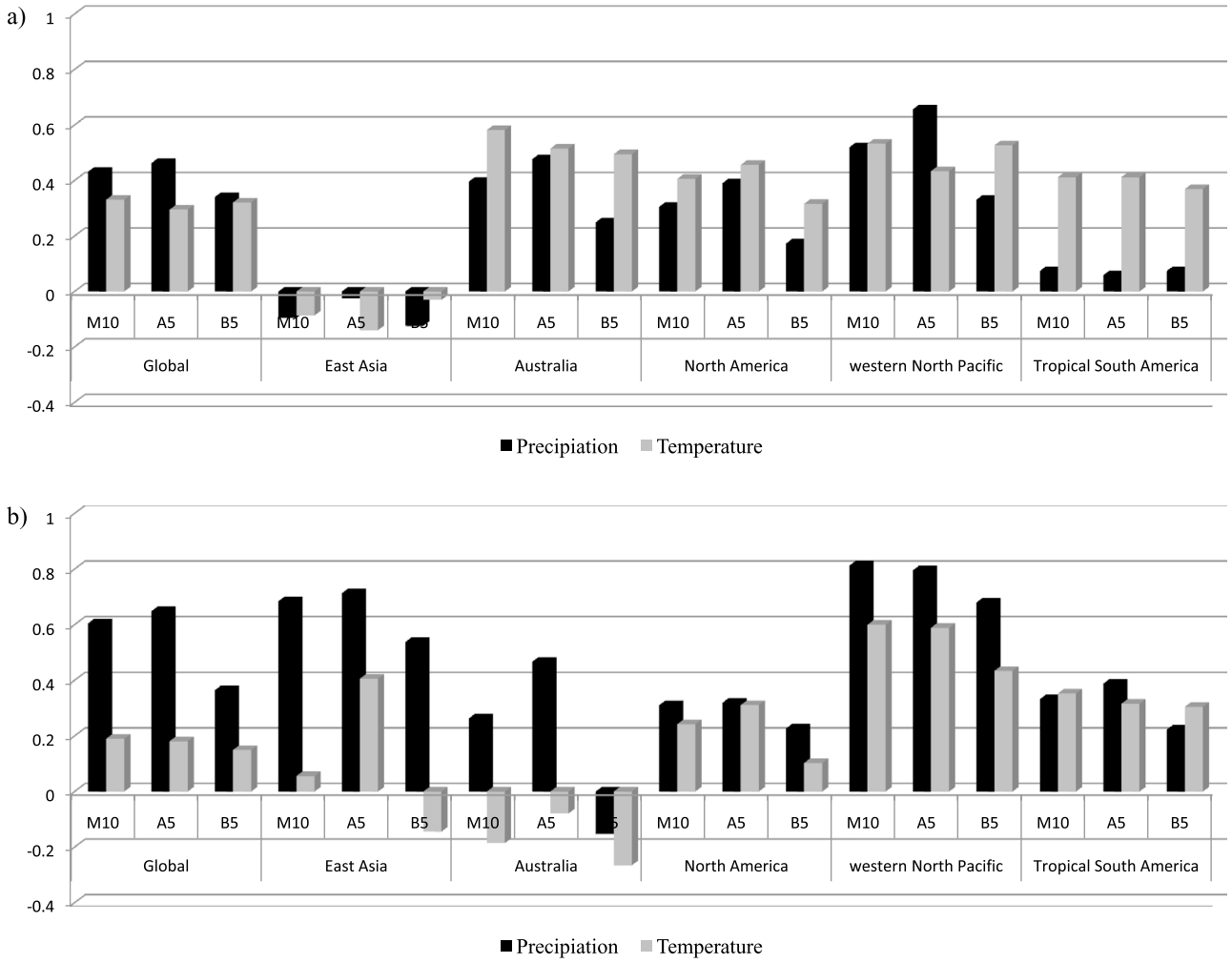

Figure 7. Spatial pattern correlations between the observed and the predicted precipitations (solid bars) and those for the temperature at 850 hPa (light-shaded bars) for boreal winter (DJF) of (a) December 2008 to February 2009 and (b) December 2009 to February 2010 over the six regions of the global region (0°–360°E, 90°S–90°N), East Asia (90°E–150°E, 20°N–50°N), Australia (110°E–160°E, 45°S–10°S), North America (50°W–140°W, 10°N–70°N), western North Pacific (120°E–180°E, 10°N–30°N), and tropical South America (30°W–90°W, 10°S–15°N).

of B5 and M10 for two variables over the western North Pacific including the East Asia and the Indian Ocean.

[19] In short, as shown above, although the contribution from the nonperforming models is not considered in the MME, the skills for a few tropical regions are unaffected. However, in midlatitudes, such omission will actually improve the prediction skills.

[20] In order to explore the usefulness of the climate filter method for real-time seasonal prediction, we have applied the current method to MME prediction for boreal winter seasons of December 2008 to February 2009 and December 2009 to February 2010. Figure 7 shows the spatial pattern correlations between the observed and the predicted rainfall and temperature at 850 hPa over the six regions, namely, the Global region (0°–360°E, 90°S–90°N), East Asia (90°E–150°E, 20°N–50°N), Australia (110°E–160°E, 45°S–10°S), North America (50°W–140°W, 10°N–70°N), western North Pacific (120°E–180°E, 10°N–30°N), and tropical South America (30°W–90°W, 10°S–15°N) for the two winter seasons. In general, the performances of predicted skills for precipitation by A5 are better than those by all model-inclusive M10 and B5. It is to be noted, however, that the skills over East Asia for all the three MME suites for 2008 are particularly poor, the reason for which needs to be examined. Notwithstanding this, from these results it is noteworthy that the A5 actually gives, in general, more skillful predictions than M10 for the examined cases and hence is potentially useful for real time seasonal prediction.

4. Concluding Remarks

[21] Using the NCEP-DOE reanalysis [*Kanamitsu et al.*, 2002], CMAP rainfall data sets [*Xie and Arkin*, 1997] and operational MME hindcast data sets of the APEC climate center, we propose and demonstrate a new approach to further improve the MME hindcasts for the boreal winter seasons (DJF) during the period 1981–2003. We find that the observed rainfall and local ENSO-associated Walker circulation in the

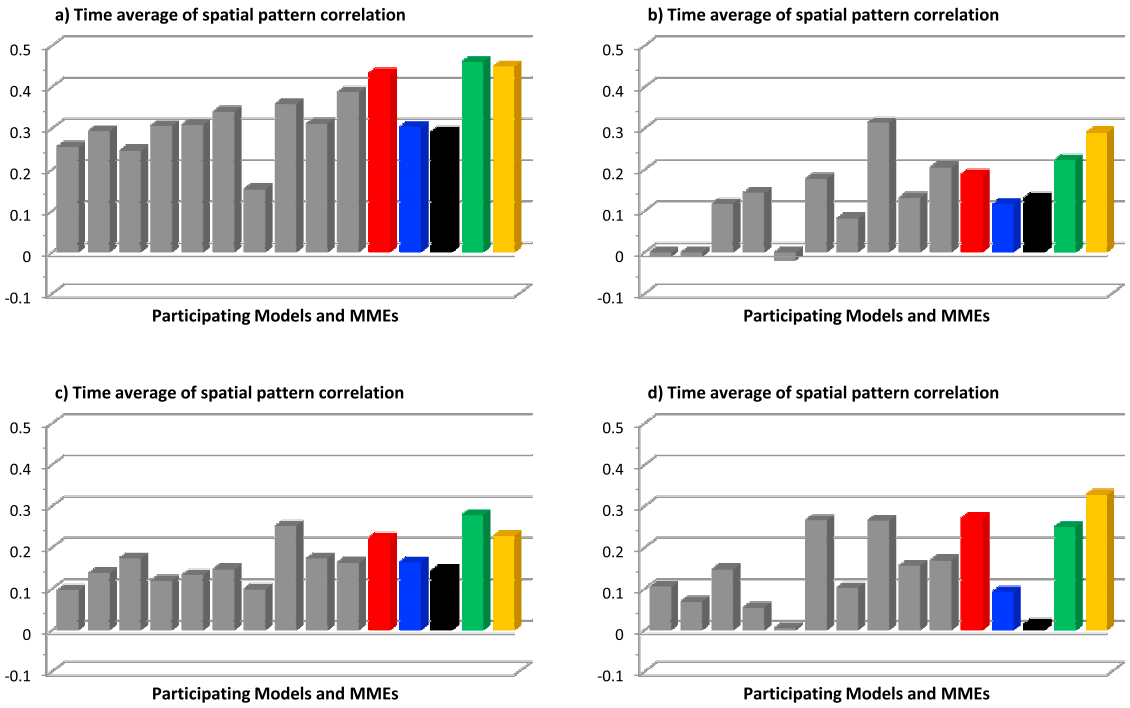

Figure A1. Time average of spatial pattern correlation coefficients between the observed and the predicted precipitation for 23 boreal winters (DJF) during the period of 1981–2003 over the (a) global and (b) East Asian (90°E–150°E, 20°N–50°N) regions. The gray bars are the hindcast skill scores of individual models involved in routine operational MME-based climate predictions at APCC. The colored bars indicate hindcast skills of the four operational deterministic multimodel ensemble methods of APEC Climate Center for the same season; red (M10) and yellow (A5) bars are for the simple composite of bias corrected model ensemble means, the blue ones represent the multiple regression based blend of model ensemble means, the black ones are for the synthetic multimodel ensemble based on multiple regression on leading PCs of EOF; the green ones present the stepwise pattern projection methods based on the pointwise regression method. (c and d) The same as Figures A1a and A1b, except for the temperature at 850 hPa.

tropical Pacific are strongly correlated. Accordingly, we apply a "climate filter" that examines the degree of reproducibility of the above observed feature to grade the fidelity of the model performance. We also explore the possible use of the above filter to devise an improved MME suite for seasonal prediction.

[22] It can be seen that prediction skills of our sample MME, involving all the available 10 models, only comes from the five "better" models that successfully reproduce the tropical Pacific ENSO-rainfall relationship. Further, the distribution of the prediction for the MMEs with "more skillful" and "less skillful" models is region-sensitive. It can be discerned that the gap between skills of MMEs comprising of only the "more skillful" models and those with only the "less skillful" models significantly widens over the extratropics and consequently drags down the skills of a comprehensive MME which contains all the available models. The possible reason is that the "more skillful" models which clear through the climate filter may provide appropriate heat sources in the tropics which facilitate better teleconnection to the extratropics and beyond, though detailed diagnosis to verify this hypothesis is beyond the scope of the current study. There is very little difference in the prediction skills for the MME carried out with all the models as well as those carried out with the "more skillful" models in the regions such as tropical South America that are adjacent to the eastern tropical Pacific, where almost all the models perform relatively well. Further, this research also supports the results of *Yoo and Kang* [2005] multimodel composite provides a better skill if models that are more skillful; we have ascertained this fact by computing the correlation skills of MME, composite variance, etc. (figures not shown), following *Yoo and Kang* [2005]. The study, however, is subject to the limitations of data length, which is unavoidable due to the lack of longer hindcasts. The fact that the ENSO alone is not a major climate driver for the climate everywhere is also an issue. Nonetheless, our work shows that selection of the models that represent realistic climate features and associated responses enhance the chances of better prediction. Improvement of model fidelity of climate processes [e.g., *Iizuka et al.*, 2003; *Lee et al.*, 2008] and use of data assimilation and initialization [e.g., *Hudson et al.*, 2011] with judicious use of "climate filters" such as presented here will translate into better MME-based seasonal prediction.

Appendix A: Hindcast Skill Scores of Operational MME-Based Climate Prediction at APEC Climate Center

[23] In APEC Climate Center (APCC), dynamical seasonal prediction information is produced through a state-of-the art multimodel ensemble prediction system utilizing the

model predictions from APEC member economies every month. The climate prediction information and performance, produced by APCC, are disseminated to the APEC regions.

[24] Figure A1 shows the prediction skills of the routine operational MME (colored bars) and individual models (gray bars) for precipitation and 850 hPa temperature for 23 boreal winters (DJF) during the period of 1981–2003. Four kinds of deterministic operational MME techniques are used. First, red bars indicate the simple arithmetic mean of bias corrected predictions based on individual member models [*Peng et al.*, 2002], the blue ones are for the pointwise multiple regression technique based on the training period [*Krishnamurti et al.*, 2000], the black ones represent synthetic multimodel ensemble based on EOF-filtered data as minimizing the residual error variance [*Yun et al.*, 2005], the last green ones present the pointwise regression method for predicting 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 [*Kug et al.*, 2008]. Finally, yellow bars are the same as red bars, except for the simple arithmetic mean carried out with the class A5 models which clear through the climate filter.

[25] **Acknowledgments.** The authors appreciate the participating institutes of the APCC multimodel ensemble operational system for providing the hindcast experiment data. The authors acknowledge constructive comments from three anonymous reviewers that helped improve the original manuscript. Discussions with B. Wang, N. H. Saji, N. Deshpande, and T. P. Sabin are acknowledged. Most of the figures in this paper have been carried out using the GrADs/COLA software. The NCEP-DOE reanalysis 2 data and CMAP precipitation data have been provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/.

References

Barnston, A. G., S. J. Mason, L. Goddard, D. G. Dewitt, and S. E. Zebiak (2003), Multimodel ensembling in seasonal climate forecasting at IRI, *Bull. Am. Meteorol. Soc.*, *84*, 1783–1796, doi:10.1175/BAMS-84-12-1783.

Bretherton, C., M. Widmann, V. Dymnikov, J. Wallace, and I. Blade (1999), The effective number of spatial degrees of freedom of a time-varying field, *J. Clim.*, *12*, 1990–2009, doi:10.1175/1520-0442(1999)012<1990:TENOSD>2.0.CO;2.

Doblas-Reyes, F. J., M. Deque, and J.-P. Piedelievre (2000), Multi-model spread and probabilistic seasonal forecasts in PROVOST, *Q. J. R. Meteorol. Soc.*, *126*, 2069–2088, doi:10.1256/smsqj.56704.

Doblas-Reyes, F. J., R. Hagedorn, and T. N. Palmer (2005), The rationale behind the success of multi-model ensembles in seasonal forecasting: II. Calibration and combination, *Tellus, Ser. A*, *57*, 234–252.

Hagedorn, R., F. J. Doblas-Reyes, and T. N. Palmer (2005), The rationale behind the success of multi-model ensembles in seasonal forecasting: I. Basic concept, *Tellus, Ser. A*, *57*, 219–233.

Hudson, D., O. Alves, H. H. Hendon, and G. Wang (2011), The impact of atmospheric initialisation on seasonal prediction of tropical Pacific SST, *Clim. Dyn.*, *36*, 1155–1171, doi:10.1007/s00382-010-0763-9.

Iizuka, S., K. Orito, T. Matsuura, and M. Chiba (2003), Influence of cumulus convection schemes on the ENSO-like phenomena simulated in a CGCM, *J. Meteorol. Soc. Jpn.*, *81*, 805–827, doi:10.2151/jmsj.81.805.

Jolliffe, I. T., and D. B. Stephenson (2003), *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, John Wiley, Hoboken, N. J.

Kanamitsu, M., et al. (2002), NCEP-DOE AMIP-II Reanalysis (R-2), *Bull. Am. Meteorol. Soc.*, *83*, 1631–1643, doi:10.1175/BAMS-83-11-1631(2002)083<1631:NAR>2.3.CO;2.

Kang, H., et al. (2009), Statistical downscaling of precipitation in Korea using multimodel output variables as predictors, *Mon. Weather Rev.*, *137*, 1928–1938, doi:10.1175/2008MWR2706.1.

Krishnamurti, T. N., et al. (1999), Improved weather and seasonal climate forecasts from multimodel superensemble, *Science*, *285*, 1548–1550, doi:10.1126/science.285.5433.1548.

Krishnamurti, T. N., C. M. Kishtawal, D. W. Shin, and C. E. Williford (2000), Multi-model superensemble forecasts for weather and seasonal climate, *J. Clim.*, *13*, 4196–4216, doi:10.1175/1520-0442(2000)013<4196:MEFFWA>2.0.CO;2.

Kug, J.-S., J.-Y. Lee, and I.-S. Kang (2008), Systematic error correction of dynamical seasonal prediction using a stepwise pattern projection method, *Mon. Weather Rev.*, *136*, 3501–3512, doi:10.1175/2008MWR2272.1.

Lee, D. Y., C.-Y. Tam, and C.-K. Park (2008), Effects of multicumulus convective ensemble on East Asian summer monsoon rainfall simulation, *J. Geophys. Res.*, *113*, D24111, doi:10.1029/2008JD009847.

Lee, J.-Y., et al. (2010), How are seasonal prediction skills related to models' performance on mean state and annual cycle?, *Clim. Dyn.*, *35*, 267–283, doi:10.1007/s00382-010-0857-4.

Lee, J.-Y., et al. (2011), How predictable is the Northern Hemisphere summer upper-tropospheric circulation?, *Clim. Dyn.*, doi:10.1007/s00382-010-0909-9, in press.

Lee, S.-S., J.-Y. Lee, K.-J. Ha, B. Wang, and J. K. E. Schemm (2011), Deficiencies and possibilities for long-lead coupled climate prediction of the Western North Pacific-East Asian summer monsoon, *Clim. Dyn.*, *36*, 1173–1188, doi:10.1007/s00382-010-0832-0.

Lee, W.-J., et al. (2009), APCC 2009 final report, APEC Clim. Cent., Busan, Korea. (Available at http://www.apcc21.net/activities/activities 03_01. php).

Michaelsen, J. (1987), Cross-validation in statistical climate forecast models, *J. Clim. Appl. Meteorol.*, *26*, 1589–1600, doi:10.1175/1520-0450(1987)026<1589:CVISCF>2.0.CO;2.

Min, Y.-M., V. N. Kryjov, and C.-K. Park (2009), A probabilistic multimodel ensemble approach to seasonal prediction, *Weather Forecast.*, *24*, 812–828, doi:10.1175/2008WAF2222140.1.

Palmer, T. N., C. Brankovic, and D. S. Richardson (2000), A probability and decision-model analysis of PROBOST seasonal multi-model ensemble integrations, *Q. J. R. Meteorol. Soc.*, *126*, 2013–2033, doi:10.1256/smsqj.56702.

Palmer, T. N., et al. (2004), Development of a European multi-model ensemble system for seasonal to inter-annual prediction (DEMETER), *Bull. Am. Meteorol. Soc.*, *85*, 853–872, doi:10.1175/BAMS-85-6-853.

Pavan, V., and J. Doblas-Reyes (2000), Multimodel seasonal hindcasts over the Euro-Atlantic: Skill scores and dynamic features, *Clim. Dyn.*, *16*, 611–625, doi:10.1007/s003820000063.

Peng, P., A. Kumar, H. van den Dool, and A. G. Barnston (2002), An analysis of multimodel ensemble predictions for seasonal climate anomalies, *J. Geophys. Res.*, *107*(D23), 4710, doi:10.1029/2002JD002712.

Shukla, J., et al. (2000), Dynamical seasonal prediction, *Bull. Am. Meteorol. Soc.*, *81*, 2593–2606, doi:10.1175/1520-0477(2000)081<2593:DSP>2.3.CO;2.

Snedecor, G. W., and W. G. Cochran (1980), *Statistical Methods*, 7th ed., 507 pp., Iowa State Univ. Press, Ames.

Spiegel, M. R., and L. J. Stephens (2008), *Schaum's Outline of Theory and Problems of Statistics*, 4th ed., 577 pp., McGraw-Hill, New York.

Tanaka, H. L., I. Noriko, and A. Kitoh (2004), Trend and interannual variability of Walker, monsoon and Hadley circulations defined by velocity potential in the upper troposphere, *Tellus, Ser. A*, *56*, 250–269.

Wang, B., et al. (2008), How accurately do coupled climate models predict the leading modes of Asian-Australian monsoon interannual variability?, *Clim. Dyn.*, *30*, 605–619, doi:10.1007/s00382-007-0310-5.

Wang, B., et al. (2009), Advance and prospectus of seasonal prediction: Assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980–2004), *Clim. Dyn.*, *33*, 93–117, doi:10.1007/s00382-008-0460-0.

Wang, C. (2002), Atmospheric circulation cells associated with the El Niño–Southern Oscillation, *J. Clim.*, *15*, 399–419, doi:10.1175/1520-0442(2002)015<0399:ACCAWT>2.0.CO;2.

Wang, X., and S. S. Shen (1999), Estimation of spatial degrees of freedom of a climate field, *J. Clim.*, *12*, 1280–1291, doi:10.1175/1520-0442(1999)012<1280:EOSDOF>2.0.CO;2.

Wilks, D. S. (1995), *Statistical Methods in the Atmospheric Sciences: An Introduction*, 467 pp., Academic, San Diego, Calif.

World Meteorological Organization (2006), Standardised Verification System (SVS) for Long-Range Forecasts (LRF): New attachment II-8 to the manual on the GDPFS, *WMO 485*, vol. I, 83 pp., Geneva, Switzerland.

Xie, P., and P. A. Arkin (1997), Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, *Bull. Am. Meteorol. Soc.*, *78*, 2539–2558, doi:10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2.

Yoo, J. H., and I.-S. Kang (2005), Theoretical examination of a multimodel composite for seasonal prediction, *Geophys. Res. Lett.*, *32*, L18707, doi:10.1029/2005GL023513.

Yun, W.-T., et al. (2005), A multi-model superensemble algorithm for seasonal climate prediction using DEMETER forecasts, *Tellus, Ser. A*, *57*, 280–289.

J.-B. Ahn and D. Y. Lee, Department of Atmospheric Sciences, Pusan National University, Busan, 609-735, Korea.

K. Ashok, Centre for Climate Change Research, Indian Institute of Tropical Meteorology, Pashan, Pune 411 008, India. (ashok@tropmet.res.in)