ATMOSPHERIC SCIENCE LETTERS
Atmos. Sci. Let. (2013)
Published online in Wiley Online Library
(wileyonlinelibrary.com) **DOI:** 10.1002/asl2.430

Royal Meteorological Society

Improvement of grand multi-model ensemble prediction skills for the coupled models of APCC/ENSEMBLES using a climate filter

Doo Young Lee,[1,2]* Joong-Bae Ahn,[2] Karumuri Ashok[3] and Andrea Alessandri[4]

[1] *APEC Climate Center, Haeundae-gu, Busan 612-020, Republic of Korea*
[2] *Department of Atmospheric Sciences, Pusan National University, Jangjeon-dong, Busan 609-735, Republic of Korea*
[3] *Centre for Climate Change Research, Indian Institute of Tropical Meteorology, Pune 411008, India*
[4] *Centro Ricerche Casaccia, Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Roma (RM) I-00123, Italy*

**Correspondence to:*
D. Y. Lee, APEC Climate Center, 12 Centum 7-ro, Haeundae-gu, Busan 612-020, Republic of Korea.
E-mail: dylee@apcc21.net

Received: 25 July 2012
Revised: 12 February 2013
Accepted: 22 March 2013

Abstract

Twelve coupled model simulations of two multi-model ensemble (MME) systems for boreal winters from 1983 to 2005 are used to improve the climate prediction. From grading the relative capability of each simulation in reproducing the observed link between the tropical El Niño-Southern Oscillation (ENSO)-related Walker circulation and the Pacific rainfall, we find an optimal MME suite with improved prediction skills. This study demonstrates that the climate filter concept, proposed by us in a recent work, is not only useful in improving the MME prediction skills as compared to a single MME system, but also the skills of a grand MME that encompasses two well-performing MMEs. Copyright © 2013 Royal Meteorological Society

Keywords: **Grand multi-model ensemble; climate filter; prediction skills**

1. Introduction

The multi-model ensemble (MME) strategy, known as the combination of ensembles from different models, is a useful tool to reduce the uncertainty that arises from the inherent errors both in the initial conditions and in the dynamical or physical processes of the imperfect model formulations (Krishnamurti *et al*., 1999; Peng *et al*., 2002; Yun *et al*., 2005); the method generally demonstrates more skillful performance than the individual models (Barnston *et al*., 2003; Palmer *et al*., 2004; Hagedorn *et al*., 2005). Dynamical MME seasonal predictions are operationally implemented at many climate prediction centers (Saha *et al*., 2006; Lee *et al*., 2009). However, increased number of multiple model forecasts does not necessarily translate into better skills. For example, Weisheimer *et al*. (2009) investigated the prediction skill for a grand MME, which is a combination of DEMETER (Palmer *et al*., 2004) and ENSEMBLES (Weisheimer *et al*., 2009; Alessandri *et al*., 2011). They show that the prediction skills for the grand MME are yet relatively limited. In this context, in a recent paper (Lee *et al*., 2011), we showed that a few models with relatively poor forecast/hindcast skills can still limit the skills of the MME. The work also introduces the concept of using the fidelity of the model skills in reproducing a relevant climate phenomena and/or relationship to grade the models. Importantly, the study indicates that a newer MME with those models that successfully reproduce the rainfall in the tropical Pacific due to the El Niño-Southern Oscillation (ENSO)-links Walker circulations gives significantly better hindcast skills in prediction of winter climate in East Asia, parts of Australia, etc.

In this work, we apply the concept of the climate filter (Lee *et al*., 2011) for potential improvement of a grand MME, derived from a combination of APEC Climate Center (APCC) MME seasonal prediction system (Lee *et al*., 2009) and ENSEMBLES (Weisheimer *et al*., 2009; Alessandri *et al*., 2011) in order to explore whether the methodology can improve the skills of the grand MME, constituent MMEs, and individual models.

2. Data and methodology

2.1. Data used

The boreal winter (December through February, DJF) hindcast outputs for the period of 1983–2005 from seven coupled models involved in the operational 6-month MME seasonal prediction system of the APCC, and five coupled models from the European Commission FP7 project called ENSEMBLES for seasonal to annual predictions – totaling 12 coupled model hindcast sets – are used in this study (Table I). In addition, the atmospheric variables (NCEP-DOE R2; Kanamitsu *et al*., 2002), precipitation (CMAP; Xie and Arkin, 1997) and sea surface temperature (OISST V.2; Reynolds *et al*., 2002) from 1983 to 2005 are also used as observations.

Table I. Description of the coupled atmosphere–ocean general circulation models used.

Institutes	Model Name	AGCM	Resolution	OGCM	Resolution	Ensemble member
APCC – global coupled atmosphere-ocean climate models						
APCC	APCC-CCSM3	CAM3	T85L26	POP 1.3	gxlv3_L40	5
NCEP	NCEP_CFS	GFS	T62L64	MOM3	1/3°lat × 1°lon L40	15
BOM	POAMA	BAM v3.0d	T47L17	ACOM2	0.5–1.5°lat × 2°lon L25	10
FRCGC	SINTEX-F	ECHAM4	T106L19	OPA8.2	2°cos(lat) × 2°lon L31	9
Seoul National University	SNU	SNU	T42L21	MOM2.2	1/3°lat × 1°lon L32	6
University of Hawaii	UH	ECHAM4	T31L19	UH Ocean	1°lat × 2°lon L2	10
Pusan National University	PNU	CCM3	T42L18	MOM3	∼0.7 (low lat) ∼1.4 (mid lat) and ∼2.8 (high lat) L29	5
ENSEMBLES – global coupled atmosphere–ocean climate models						
ECMWF	ECMF	IFS CY31R1	T159L62	HOPE	0.3° ×1.4° L29	9
UKMO	EGRR	HadGEM2-A	N96/L38	HadGEM2-O	0.33°lat × 1°lon L20	9
MF	LFPW	ARPEGE4.6	T63L31	OPA8.2	2°lat × 2°lon L31	9
IFM/GEOMAR	IFMK	ECHAM5	T63L31	MPI-OMI	1.5°lat × 1.5°lon L40	9
CMCC-INGV	INGV	ECHAM5	T63L19	OPA8.2	2°lat × 2°lon L31	9

2.2. Basic concept of the climate filter

The rationale behind the climate filter concept is as follows. One of the most prominent and important atmospheric systems is the Walker circulation, and this zonal circulation, which drives major change in rainfall variability over the tropical Pacific, is closely linked to the ENSO (Krishnamurti *et al*., 1973; Ropelewski and Halpert, 1989). The coupled models have significant lead prediction skills in predicting ENSO indices such as the Niño3 up to 6 months in advance (Jin *et al*., 2008; Jeong *et al*., 2012). Therefore, it is expected that the ENSO-linked Walker circulation and the associated rainfall changes in the tropical Pacific should at least be predictable by decent models. This relationship is the basis of the climate filter concept that Lee *et al*. (2011) introduced to select the better performing models for seasonal prediction.

To compute the Walker circulation, we first use the seasonal 200 hPa velocity potential anomalies, obtained by removing the seasonal climatological mean from the 200 hPa velocity potential. We again remove the zonal mean field as the signal of the Hadley circulation from these anomalies, and the residual as the Walker circulation is taken.

The temporal correlation pattern between the observed Walker circulation and precipitation in the tropical region for periods 1983–2005 as a reference is calculated (figure not shown). High correlations with magnitudes of more than 0.4, significant at 95% confidence level from a Student's two-tailed t-test (henceforth t-test), are generally located along 10°S–10°N. Especially, there is a strong association of the local rainfall in the central and western tropical Pacific with the zonal circulation.

Further, to examine the relationship of the Walker circulation with ENSO, we compute the correlations of the Walker circulation index [a Walker circulation index is calculated as the difference of area-averaged Walker circulation between the tropical eastern Pacific (10°S–0°, 175°E–105°W) and the tropical western Pacific (10°S–5°N, 110–135°E)] with the Niño 3.4 index (the Niño 3.4 index is calculated with SST anomalies averaged in the box 170–120°W and 5°S–5°N). From this relationship (not shown), it can be discerned that the strong relationship between the observed Walker circulation with Niño 3.4 (−0.9, significant at 99% confidence level from a t-test) is well reproduced by hindcasts of all model Walker circulations. Also, we can find that most of model simulations well capture the general characteristics of the observed Walker circulation.

We utilize the magnitude of the squared correlation coefficients between the Walker circulation indices and the Niño 3.4 index as the weights for not only the observed Walker circulation but also the hindcast Walker circulation of each model to compute the observed and predicted ENSO-associated Walker circulation for all models. The squared correlation coefficient, known as coefficient of determination, is one of the best means for evaluating the strength of the linear association between x and y (Wilks, 1995).

On the basis of these points, we believe that it is an important measure of model fidelity to predict the tropical rainfall from model simulations of ENSO-associated Walker circulation in the tropical Pacific and also minimum requirement for any model with necessary fidelity. This concept is used as a climate filter to classify the individual models.

Specifically, we use two empirical criteria, which evaluate the strength of hindcast relationship between ENSO-associated Walker circulation and rainfall over the tropical Pacific, applied through a scatter diagram analysis to grade the individual model skills. The criteria are: (1) the slope of the regression line fitted between the observed and simulated pattern correlations of tropical rainfall and ENSO-associated

Walker circulation should be larger than 0.5 and less than 1.5 and (2) statistically significant temporal correlation between these observed and simulated pattern correlations is more than 0.5 (significant at ~99% confidence level from a t-test).

2.3. Statistical methods used

For MME prediction, we adopt a simple composite method (Peng *et al*., 2002; Lee *et al*., 2009, 2011), known as simple arithmetic mean of bias corrected predictions, with equal weights to predictions from individual models. The bias correction in this method is performed by subtracting the model's own climatology from each model forecast as an a posteriori removal of systematic errors (Peng *et al*., 2002; Lee *et al*., 2009, 2011).

The standard t-test (Wilks, 1995) is employed to compute the statistical significance of the correlations. The degrees of freedom for the temporal correlation is estimated as N-2, where N is 23, the number of winter seasons during the study period. To find the significance levels for spatial pattern correlations, we use the effective spatial degree of freedom (ESDOF) (Wang and Shen, 1999).

Finally, to calculate seasonal anomalies of each model parameter as well as those from observations for each year, we follow the standard leave-one-out cross-validation method (Jolliffe and Stephenson, 2003). We also use this method in each target year while applying the climate filter for all hindcast periods.

3. Evaluation of the hindcast relationship through the climate filter

The scatter plot between the observed and predicted pattern correlations between tropical Pacific rainfall and ENSO-associated Walker circulation for all the 12 models is presented in Figure 1. Applying the two aforementioned empirical conditions to grade the models (see Section 2.2 for more details), we find that 4 of 12 models, namely, models 2, 5, 8 and 9 (Figure 1(b), (e),(h) and (i)) successfully represent the realistic rainfall relationship with the local ENSO-associated Walker circulation in the tropical Pacific for the boreal winter season. We further apply aforementioned cross-validation method and find that the skills of these four models are robust to the leave-one-out procedure for the study period (table not shown).

We implement three separate MME hindcast experiments, which are, for convenience, named as the M12 (essentially a grand MME involving hindcasts from all the 12 models), the A4 (means a filtered grand MME involving hindcasts from the four performing models) and the B8 (uses the rest of the model hindcasts). Figure 2 illustrates the time averages of the spatial pattern correlations between the observed and the simulated rainfall and temperature at 850 hPa from all the three MME experiments for six arbitrary regions, namely, the Globe (0°–360°E, 90°S–90°N), tropics (0°–360°E, 20°S–20°N), East Asia (90–150°E, 20–50°N), South Asia (60–120°E, 10–40°N), western North Pacific (120–160°E, 10–40°N) and Australia (110–180°E, 50–10°S). The other four regions cover only the subtropical through mid-latitude regions.

We find from Figure 2 that the hindcast skills of the A4 are in general better than, or similar to, those for M12 and B8. Even the marginally higher skills of the M12 grand MMEs for the prediction of precipitation in the global and tropical regions (0.46 and 0.51, respectively, significant at 91 and 92.6% confidence level from a t-test; ESDOF = 14.4 and 12.7) are, however, only an insignificant difference from the corresponding skills (0.45 and 0.50 significant at 91.2 and 92.7% confidence level from a t-test; ESDOF = 15.1 and 13.5, respectively) of the A4. For these two regions, these slightly better performances of the M12 are essentially due to the relatively better performances of the B8 predictions in these regions.

Meanwhile, in the four extratropical regions, it can be seen that the gap between the MME prediction skills of the A4 and those of the B8 is significantly different. Particularly, for the East Asia and South Asia, the differences of the MME prediction skills for both variables between the A4 (0.32 and 0.36 for rainfall; 0.29 and 0.21 for temperature) and the B8 (0.17 and 0.22 for rainfall; 0.14 and 0.04 for temperature) are distinctly shown as a value of about 0.15. This finding shows that the relatively poor performance of the M12 prediction is mainly due to the poor prediction skills of the B8 models.

To further investigate the area average of the spatial distribution of prediction skills of all the MME experiments for precipitation and temperature at 850 hPa, we present spatially averaged values of the temporal correlation coefficients over the eight regions for the period of 1983–2005 in Figure 3. In general, we can easily find that the high MME prediction skills for both of variables are distributed intensively throughout the tropical region. It can be discerned that the MME prediction skills of the A4 for especially temperature are considerably enhanced over the extratropical regions including the East Asia, the South Asia, the Northern Hemisphere and the Southern Hemisphere as compared with those of B8 and M12 MME. However, for precipitation prediction, it is shown that the capture of the improved hindcast skills of the A4 is a little difficult in a few regions such as the tropics, the East Asia and the Northern Hemisphere due to the relatively better performance of the B8 MME. We can also find that the MME prediction skills in the Northern Hemisphere are slightly higher than those in the Southern Hemisphere.

To further explore a practical use of the climate filter method for a grand MME prediction of APCC/ENSEMBLES, we carry out the MME predictions for the seven APCC coupled models (named as the APCC) and the five ENSEMBLES models (named as the ENS), respectively. In Figure 4, the four MME

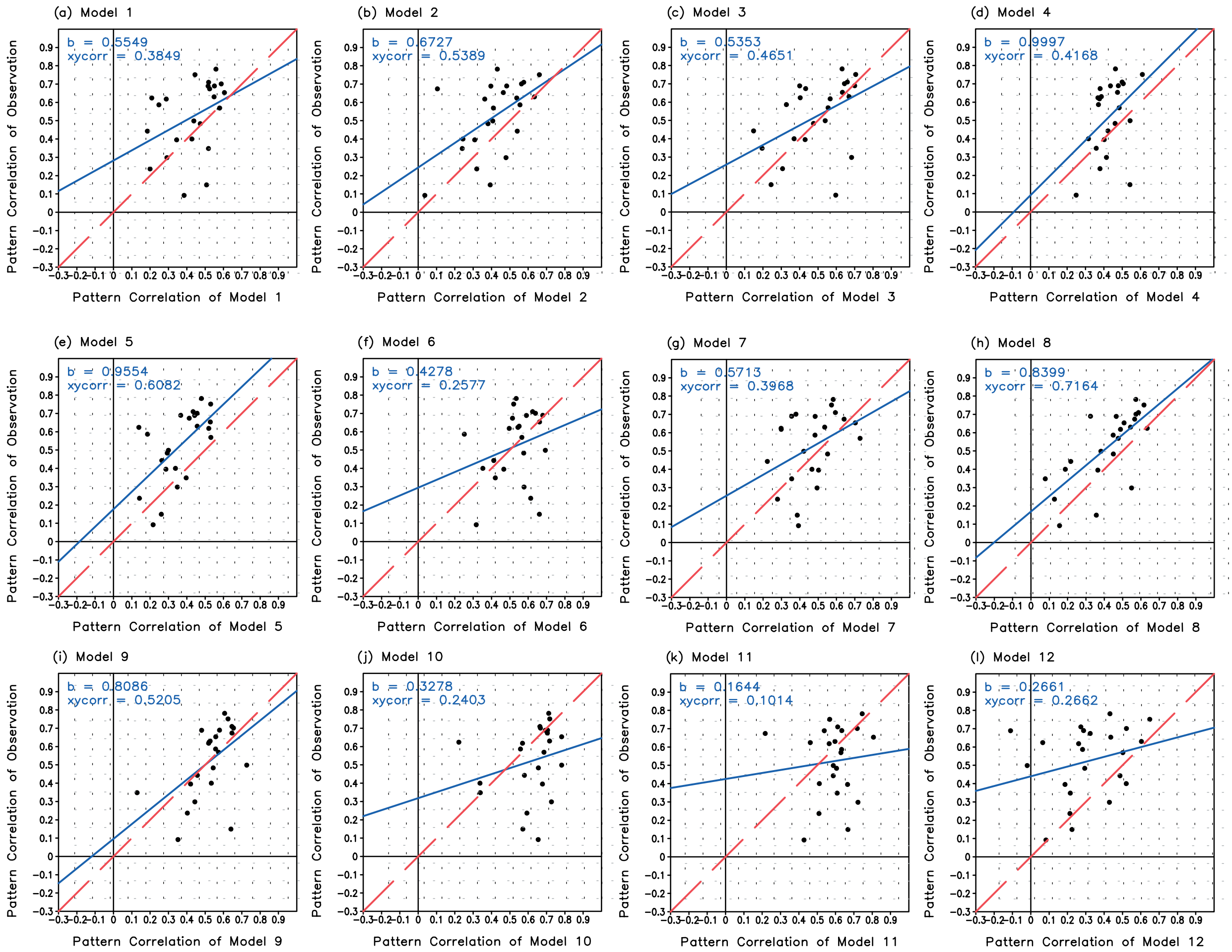

Figure 1. Scatter diagrams depicting spatial pattern correlation between the ENSO-associated Walker circulation and precipitation from observation (*y* axis) over the tropical Pacific region (100°E–60°W, 10°S–10°N), for 23 boreal winter during the period of 1983–2005, plotted against those from the individual models (*x* axis). The solid blue line is the regression line of fit and the dashed red line is a reference diagonal line. The slope 'b' from the fitted regression line is provided in the upper left. The 'xycorr' represents the temporal correlations of each model with observation. Figure reproduced from Lee *et al*. (2011).

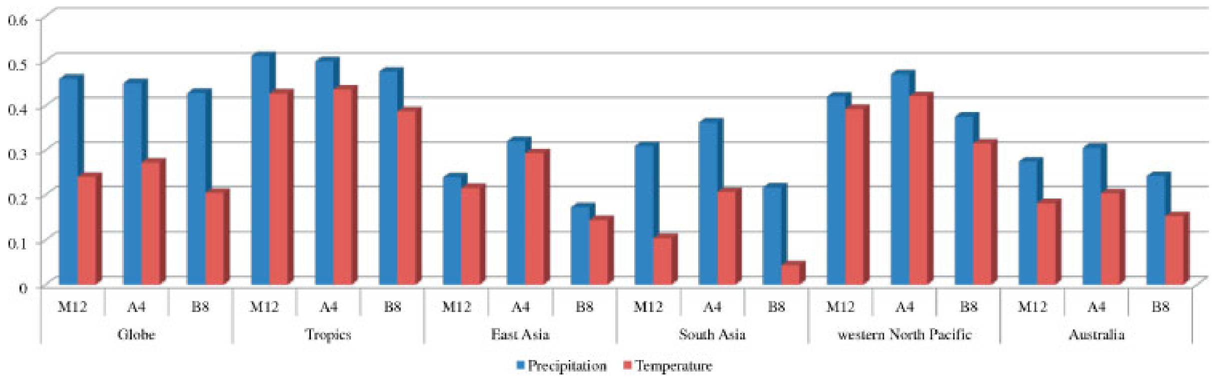

Figure 2. Time average of spatial pattern correlations between the observed and simulated precipitations (blue bars) and those for the temperature at 850 hPa (red bars) from M12, A4 and B8 over the six regions of the Global region (0°–360°E, 90°S–90°N), Tropics (0°–360°E, 20°S–20°N), East Asia (90–150°E, 20–50°N), South Asia (60–120°E, 10–40°N), western North Pacific (120–160°E, 10–40°N) and Australia (110–180°E, 50–10°S). M12, A4 and B8 are the multi-model ensemble predictions based on a simple composite method using the total of 12 models, the four more skillful models and the eight less skillful models, respectively.

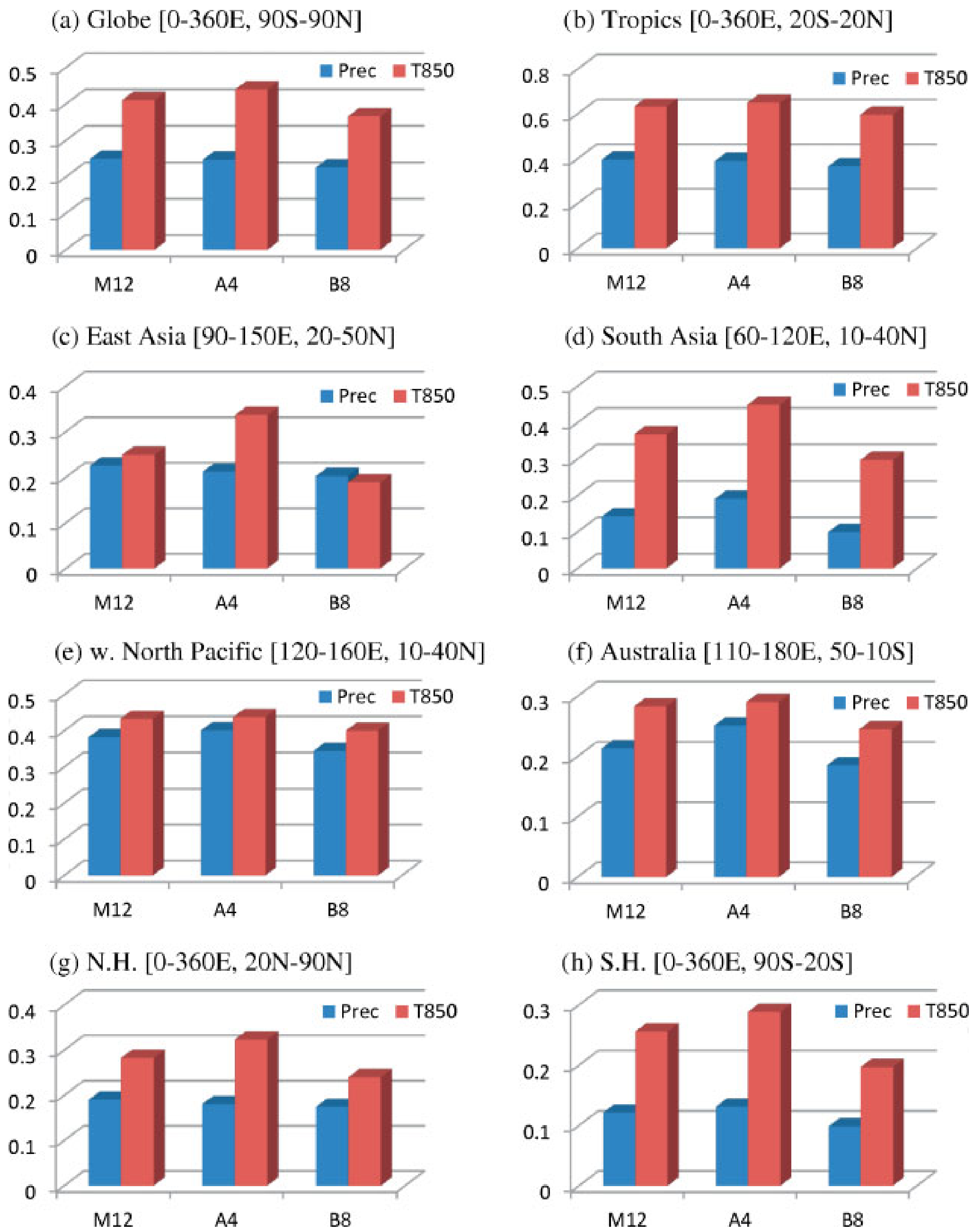

Figure 3. Area average of the spatial distribution of temporal correlation coefficients for precipitation (blue bars) and temperature at 850 hPa (red bars) between the observation and hindcasts from M12, A4 and B8 over the eight regions of the (a) Globe, (b) Tropics, (c) East Asia, (d) South Asia, (e) western North Pacific, (f) Australia, (g) Northern Hemisphere and (h) Southern Hemisphere during the period of 1983–2005.

prediction skills (APCC, ENS, M12 and A4) are compared to each of the individual model's skill and to the averaged skill of all individual models by using the scatter diagram between the spatial pattern correlations and normalized root mean square error (by the corresponding observed standard deviation) for precipitation and temperature at 850 hPa over the four mid-latitude regions. From these results, it can be seen that the skills of MME prediction are generally better than those of individual models, and it is also better than the averaged skill of all constituent models over all regions. The values of pattern correlation of MME prediction for precipitation are considerably higher than those for temperature over the same area. Differences between the normalized errors of MME predictions are relatively smaller than those between the spatial correlation skills. We can also find that the MME prediction skills by ENS are generally better than those for APCC MME. However, the APCC MME exhibits significantly better hindcast skills for Australia, which experiences its summer monsoon, and also in predicting precipitation in South Asia. While the performance of the M12 grand MME prediction is relatively inconsistent and subject to the location, it is apparent that the A4 filtered grand MME gives more

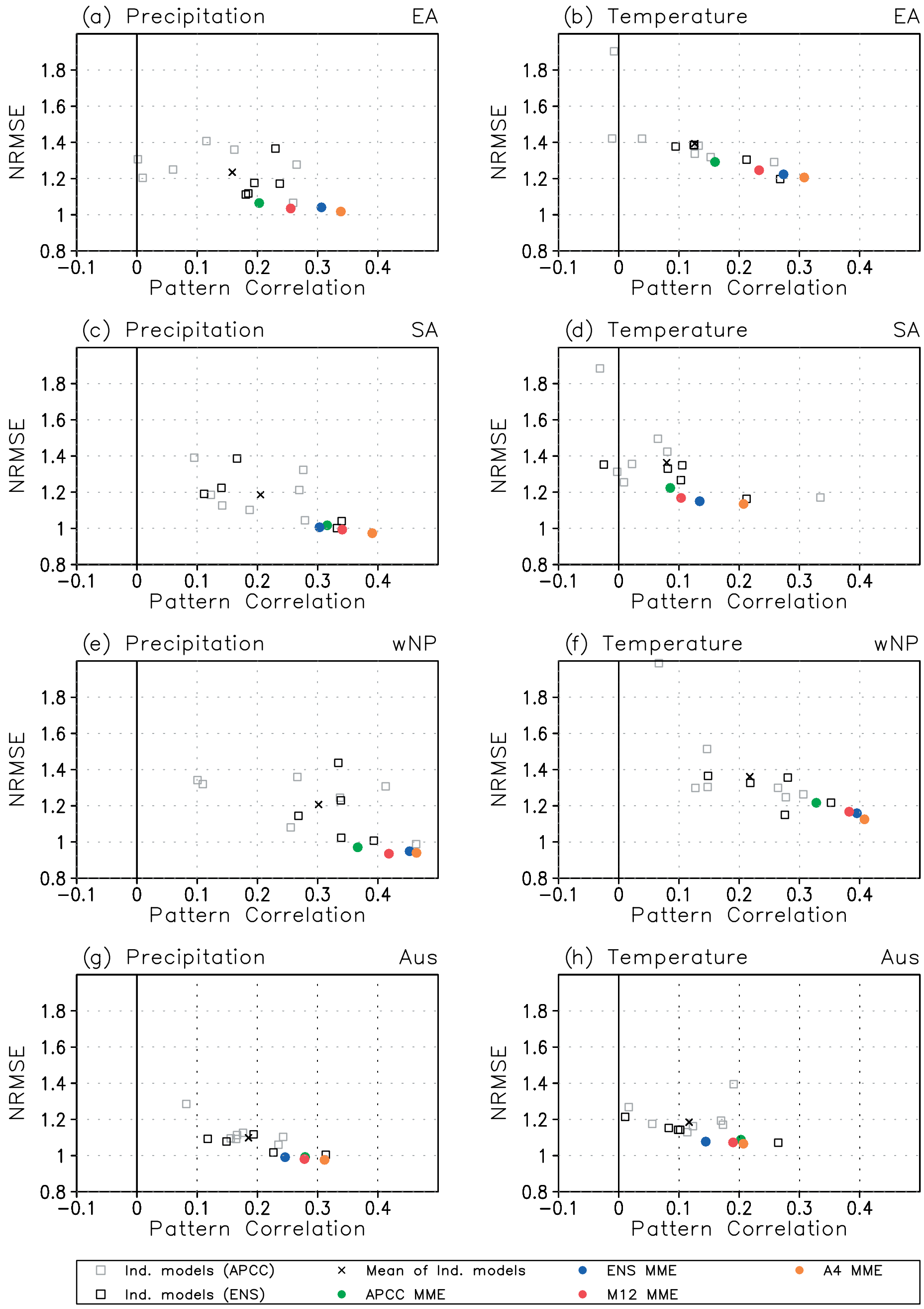

Figure 4. Scatter diagram between the spatial pattern correlation skills and normalized root mean square errors (NRMSEs) for precipitation (left panels) and temperature at 850 hPa (right panels) over the four regions of the (a, b) East Asia, (c, d) South Asia, (e, f) western North Pacific and (g, h) Australia derived from the four MMEs (APCC, ENS, M12 and A4) and individual model predictions.

skillful predictions as compared with other MMEs in the all four midlatitude regions.

4. Summary and conclusion

We apply a climate filter, proposed by Lee *et al*. (2011), to rank the individual model hindcast performances of two MME systems through evaluation of the relative capabilities of each model in reproducing the observed relationship between the tropical Pacific rainfall and the local ENSO-associated Walker circulation for boreal winter season (DJF) during the period 1983–2005. Twelve coupled model hindcast simulations of the APCC operational MME prediction system and ENSEMBLES are used in this study. We explore the usefulness of this climate filter method to filter models with better fidelity, and finally introduce an optimized MME suite with enhanced seasonal prediction skills. In agreement with Lee *et al*. (2011), we find that the MME prediction skills from four better performing models are indeed significantly higher as compared with those from the rest of the non-performing models, and those from the all-inclusive 12 model grand MME. Particularly, it is noteworthy that the difference of MME prediction skills between the performing and nonperforming models considerably widens over the extratropics; most of the models generally exhibit good skills in predicting the tropical climate. Incidentally, the performance of the grand MME including all the available models, on the other hand, shows the limitations that vary in different localities in the extratropics as compared to MME suites of the APCC and ENS. Importantly, the revised grand MME, constituting the four significant skillful models that are cleared by the climate filter, provides significantly better performance. The results, built on the earlier work by Lee *et al*. (2011), confirm that selection of models that can reproduce realistic climate associations provides a better prediction skill. From a specific and practical point, these also emphasize the necessity of the climate filter for enhancement of MME seasonal prediction of the boreal winter climate.

Acknowledgements

The authors appreciate the participating institutes of the APCC coupled MME prediction system for providing the hindcast experiment data. We are also grateful to the modeling centers for the generous support of ENSEMBLES data.

References

Alessandri A, Borrelli A, Navarra A, Aribas A, Deque M, Rogel P, Weisheimer A. 2011. Evaluation of probabilistic quality and value of the ENSEMBLES multi-model seasonal forecasts: comparison with DEMETER. *Monthly Weather Review* **139**: 581–607.

Barnston AG, Mason S, Goddard L, Dewitt DG, Zebiak SE. 2003. Multimodel ensembling in seasonal climate forecasting at IRI. *Bulletin of the American Meteorological Society* **84**: 1783–1796.

Hagedorn R, Doblas-Reyes FJ, Palmer TN. 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting – I. Basic concept. *Tellus A* **57**: 219–233.

Jeong H-I, Lee DY, Ashok K, Ahn J-B, Lee J-Y, Luo J-J, Schemm J, Hendon HH, Braganza K, Ham Y-G. 2012. Assessment of the APCC coupled MME suite in predicting the distinctive climate impacts of two flavors of El Niño during boreal winter. *Climate Dynamics* **39**: 475–493.

Jin EK, Kinter JL III, Wang B, Park C-K, Kang I-S, Kirtman BP, Kug J-S, Kumar A, Luo J-J, Schemm J, Shukla J, Yamagata T. 2008. Current status of ENSO prediction skill in coupled ocean–atmosphere model. *Climate Dynamics* **31**: 647–664.

Jolliffe IT, Stephenson DB. 2003. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. John Wiley & Sons: Chichester.

Kanamitsu M, Ebisuzaki W, Woollen J, Yang S-K, Hnilo JJ, Fiorino M, Potter GL. 2002. NCEP-DOE AMIP-II Reanalysis (R-2). *Bulletin of the American Meteorological Society* **83**: 1631–1643.

Krishnamurti TN, Kanamitsu M, Koss WJ, Lee JD. 1973. Tropical east–west circulations during the Northern winter. *Journal of the Atmospheric Sciences.* **30**: 780–787.

Krishnamurti TN, Kishtawal CM, LaRow TE, Bachiochi DR, Zhang Z, Williford CE, Gadgil S, Surendran S. 1999. Improved weather and seasonal climate forecasts from multimodel superensemble. *Science* **285**: 1548–1550.

Lee W-J, Saji NH, Ashok K, Rafisura K, Lee DY, Ahn K-H, Shon S-J, Min Y-M, Jeong H-I, Kang S-C, Heo J-Y, Jeong G-K, Song B-G, Yi H, Kim S-C, Yeo W-I, Kim H-R, Kang J-W, Ha K-J, Joo S-Y. 2009. *APEC Climate Center for Climate Information Services*, APCC 2009 Final Report. http://www.apcc21.org/eng/research/pub/repo/japcc040602_lst.jsp

Lee DY, Ashok K, Ahn J-B. 2011. Toward enhancement of prediction skills of multimodel ensemble seasonal prediction: A climate filter concept. *Journal of Geophysical Research* **116**: D06116, DOI: 10.1029/2010JD014610.

Palmer TN, Alessandri A, Andersen U, Cantelaube P, Davey M, Délécluse P, Déqué M, Diez E, Doblas-Reyes FJ, Feddersen H, Graham R, Gualdi S, Guérémy JF, Hagedorn R, Hoshen M, Keenlyside N, Latif M, Lazar A, Maisonnave E, Marletto V, Morse AP, Orfila B, Rogel P, Terres JM, Thomson MC. 2004. Development of a European multi-model ensemble system for seasonal to inter-annual prediction (DEMETER). *Bulletin of the American Meteorological Society* **85**: 853–872.

Peng P, Kumar A, van den Dool H, Barnston AG. 2002. An analysis of multimodel ensemble predictions for seasonal climate anomalies. *Journal of Geophysical Research* **107**: 4710, DOI: 10.1029/2002JD002712.

Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W. 2002. An improved in situ and satellite SST analysis for climate. *Journal of Climate* **15**: 1609–1625.

Ropelewski CF, Halpert MS. 1989. Precipitation patterns associated with the high index phase of the Southern Oscillation. *Journal of Climate* **2**: 268–284.

Saha S, Nadiga S, Thiaw C, Wang J, Wang W, Zhang Q, van den Dool HM, Pan H-L, Moorthi S, Behringer D, Stokes D, Pena M, Lord S, White G, Ebisuzaki W, Peng P, Xie P. 2006. The NCEP climate forecast system. *Journal of Climate* **19**: 3483–3517.

Wang X, Shen SS. 1999. Estimation of spatial degrees of freedom of a climate field. *Journal of Climate* **12**: 1280–1291.

Weisheimer A, Doblas-Reyes FJ, Palmer TN, Alessandri A, Arribas A, Déqué M, Keenlyside N, MacVean M, Navarra A, Rogel P. 2009. ENSEMBLES: a new multi-model ensemble for seasonal-to-annual predictions-skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. *Geophysical Research Letter* **36**: L21711, DOI: 10.1029/2009GL040896.

Wilks DS. 1995. *Statistical Methods in the Atmospheric Sciences*. Academic Press: New York, NY.

Xie P, Arkin PA. 1997. Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bulletin of the American Meteorological Society* **78**: 2539–2558.

Yun WT, Stefanova L, Mitra AK, VijayaKumar TSV, Dewar W, Krishnamurti TN. 2005. A multi-model superensemble algorithm for seasonal climate prediction using DEMETER forecasts. *Tellus A*, **57**: 280–289.

