GEOPHYSICAL RESEARCH LETTERS, VOL. 34, L15710, doi:10.1029/2007GL030730, 2007

Multimodel output statistical downscaling prediction of precipitation in the Philippines and Thailand

Hongwen Kang,[1,2] Kyong-Hee An,[1] Chung-Kyu Park,[1,3] Ana Liza S. Solis,[1,4] and Kornrawee Stitthichivapak[1,5]

Received 18 May 2007; revised 29 June 2007; accepted 10 July 2007; published 14 August 2007.

[1] Six dynamical seasonal model outputs, which are currently used in the APEC Climate Center Multimodel Ensemble (MME) prediction system, are employed for statistical downscaling prediction of station-scale precipitation in the Philippines and Thailand. Correlation analysis and Singular Value Decomposition Analysis are used to reveal atmosphere dynamic linkage based on the observed data other than model data. The observed linkage provides a robust basis for the choice of predictor and its range in predicted fields. In order to avoid spatial shift of predicted field away from observed climate, a movable window is set to select the most sensible area within the range of predictor for downscaling. The downscaled MME prediction is verified against observed station precipitation in a cross-validation manner, and the prediction skill is apparently improved compared with the simple composite of raw model predictions for most of the stations. **Citation:** Kang, H., K.-H. An, C.-K. Park, A. L. S. Solis, and K. Stitthichivapak (2007), Multimodel output statistical downscaling prediction of precipitation in the Philippines and Thailand, *Geophys. Res. Lett.*, *34*, L15710, doi:10.1029/2007GL030730.

1. Introduction

[2] Current General Circulation Models (GCMs) are able to reasonably simulate the large-scale atmospheric variables, such as sea level pressure (SLP), geopotential height at 500 hPa (Z500) [*von Storch et al.*, 1993; *Kang et al.*, 2004]. However, these GCMs show poor performance in predicting station-scale precipitation, because precipitation is governed by complicated, inherently nonlinear and extremely sensitive physical processes [*Stockdale et al.*, 1998]. One approach to improve poor prediction of precipitation is that of statistical downscaling. Statistical downscaling methods establish an empirical statistical relationship between the atmospheric circulation and precipitation, and then infer local changes by means of sensibly projecting the large scale information on the local scale [*Zorita and von Storch*, 1999]. A statistical downscaling scheme can also use GCM output as predictor data to make prediction, which is known as model output statistics (MOS) [*Wilks*, 1995]. MOS requires long and homogenous data series [*Heyen et al.*, 1996]. Every time the dynamical model undergoes a major upgrade, a long series of hindcasts must be re-computed in order to derive new MOS relations. The APEC Climate Center (APCC) aims at achieving more skillful seasonal predictions by means of operational running Multimodel Ensemble (MME) prediction system with timely contributions of GCM hindcasts and forecasts from 15 dynamical models. Therefore, these model hindcast sets are well-suited for multimodel output downscaling.

[3] Both the Philippines and Thailand are located in the Southeast Asia where the climate is under the influence of the Asian monsoon. Current GCMs lack the ability in simulating the Southeast Asian monsoon rainfall because they can not simulate correctly the relationship between local rainfall and SST anomalies over the Philippine Sea, the South China Sea and the Bay of Bengal [*Wang et al.*, 2004]. In this study, we use multimodel outputs as predictor data to make station-scale precipitation downscaling prediction for both countries, and attempt to examine the potential of precipitation downscaling using the predicted large-scale circulation by current operationally running climate models.

2. Data and Methodology

2.1. Data

[4] The observed station monthly precipitation used in this research was taken from Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAGASA) and Thailand Meteorological Department (TMD). Our target areas are the northern Philippines including 10 stations and the Bangkok region including 8 stations.

[5] The predictor data are taken from six operational seasonal prediction model outputs. The hindcast product is from Seasonal Prediction Model Intercomparison Project (SMIP) type experiment which focuses on one-month lead prediction. For CWB and MGO, the Sea Surface Temperature (SST) is the persistence of observation; For GCPS, GDAPS and JMA, SST is forecasted; NCEP is a coupled model. The basic descriptions for the individual model and data are shown in Table 1. In this study, SLP, Z500, zonal wind and temperature at 850 hPa (U850 and T850, respectively) are selected as predictor candidates. The hindcast data cover the period of 21-year from 1983 to 2003 and have the spatial resolution of $2.5° \times 2.5°$.

[6] NCEP/NCAR reanalysis data of the 4 circulation variables (SLP, Z500, U850 and T850) are also used in this study. The reanalysis data have the same spatial resolution as the model data.

2.2. Methodology

[7] The strategy for multimodel outputs downscaling prediction of precipitation is described as the following three steps:

[1]APEC Climate Center (APCC), Busan, South Korea.
[2]Also at State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing, China.
[3]Also at Korea Meteorological Administration, Seoul, South Korea.
[4]Now at Philippine Atmospheric, Geophysical and Astronomical Services Administration, Quezon City, Philippines.
[5]Now at Thailand Meteorological Department, Bangkok, Thailand.

0094-8276/07/2007GL030730$05.00

Table 1. Description of the GCMs Used in This Study

Model	Institution (Member Economy)	Data Type	SST
CWB	Central Weather Bureau (Chinese Taipei)	SMIP/HFP	Observe Persistence
GCPS	Seoul National University (Korea)	SMIP/HFP	Forecast
GDAPS	Korea Meteorological Administration (Korea)	SMIP/HFP	Forecast
JMA	Japan Meteorological Agency (Japan)	SMIP/HFP	Forecast
MGO	Main Geophysical Observatory (Russia)	SMIP/HFP	Observe Persistence
NCEP	Climate Prediction Center/NCEP (United States)	SMIP/HFP	Forecast

2.2.1. Choice of Predictor and the Range of the Predictor for Downscaling

[8] In our research, the predictor and the range of the large scale circulation for downscaling are selected based on correlation analysis and Singular Value Decomposition Analysis (SVDA). The analyses use the observed station precipitation and the observed atmospheric circulation variables (SLP, Z500, T850 and U850). The variable with the highest correlation coefficient with precipitation will be selected as predictor. SVDA is carried out to reveal atmospheric dynamical coupled relationship between local precipitation and the predictor. The observed mapping pattern provide a robust basis for the choice of the predictor and its range. However, the choice of the predictor range will also need to check whether the dynamical predictions can reproduce major mode of observed atmospheric variability realistically well.

2.2.2. Searching for Optimal Window

[9] As mentioned above, the range for downscaling is chosen based on the observed data. However, downscaling prediction should use predicted large-scale circulation information. Current GCMs suffer from some spatial drifts away from the observed climate [*Palmer et al.*, 2004]. In order to avoid the model bias, a movable window is set to scan over the predictor range. The optimal window is the most sensible area with the maximum area average of correlation coefficient with precipitation. The precipitation is finally specified by the large-scale information at the optimal window.

2.2.3. Downscaling and MME

[10] Suppose the predictand and predictor are $Y(t)$ and $X(i, j, t)$, respectively. $Y(t)$ is observed station precipitation and $X(i, j, t)$ is model predicted large-scale variable.

$$Y(t) = \alpha X_p(t) + \beta,$$

Where $X_p(t)$ is the projection of the predictor in the optimal window

$$X_p(t) = \sum_{i,j} COR(i,j)^* X(i,j,t).$$

The correlation coefficient is obtained as

$$COR(i,j) = \frac{\frac{1}{N}\sum (Y(t) - Y_m)^*(X(i,j,t) - X_m(i,j))}{\sigma_x(i,j)^*\sigma_y},$$

Where N is the training year, the subscript m means the average of the variable during the training period; σ denotes the variance [*Kang and Shukla*, 2006; *Kug et al.*, 2007].

[11] The downscaling is processed at each station for each model in a cross-validation manner. Then two MME predictions are made: one is average of downscaled precipitation from the 6 models; another is that of raw model predicted precipitation. Because the raw model outputs are gridded data, raw MME prediction is interpolated into the station for comparison.

3. Results

3.1. Philippines

[12] Our target area is the northern Philippines with 10 observation stations including the capital Manila. From the correlation analysis between the observed station precipitation in the northern Philippines and observed SLP, Z500, U850 and T850, it is found that Z500 in the tropical western Pacific has stronger correlation than other variables. Thus Z500 is selected as the predictor in this study.

[13] Figure 1 shows the first SVD mode between the observed station precipitation and observed Z500, and the time series of expansion coefficients for the leading SVD mode. This SVD mode accounts for 92% of total covariance, and the correlation coefficient between the expansion coefficients is 0.75. It is found a weakening of Z500 over the region from the Bay of Bengal eastward to the western North Pacific with a center over the Luzon Strait is accompanied by an enhanced precipitation over the northern Philippines. The leading covariant mode represents a reasonable dynamical link: for example, a weakening of Z500 centered over the Luzon Strait favors the southwesterly winds to bring more precipitation over the northern Philippines from the South China Sea. But for precipitation pattern, there are two stations (No.8 and 10 in Figure 2a) with very weak or negative value. It is found that both stations are located in hilly regions with complicated terrain, which may influence local precipitation. In order to clarify the dynamic link between observed precipitation and predicted Z500, the SVD modes for these models are also checked. Basically, these models can reproduce the leading mode of observation well. We take GCPS and NCEP as example and present the SVD modes in Figure 1. For GCPS, there is a little shift to the west; for NCEP, there is a little twist with anti-clockwise. However, both models can reproduce the leading mode of observation well. It is very interesting to note that the two stations (No. 8 and 10) have also very week covariance with predicted Z500.

[14] Based on the correlation analysis and the SVD analysis between the observed precipitation and Z500, we think the domain of 80–180E, 20S–40N contains the large-scale circulation information for downscaling over the northern Philippines. But for different model, the sensible area for downscaling should be different because of model spatial bias. In order to avoid the model bias, we set a movable window with the size of 60 longitude × 15 latitude. The window scans over the chosen domain for downscaling. The optimal window is that with the highest correlation coefficient between the observed station precipitation and predicted Z500 in the training period.

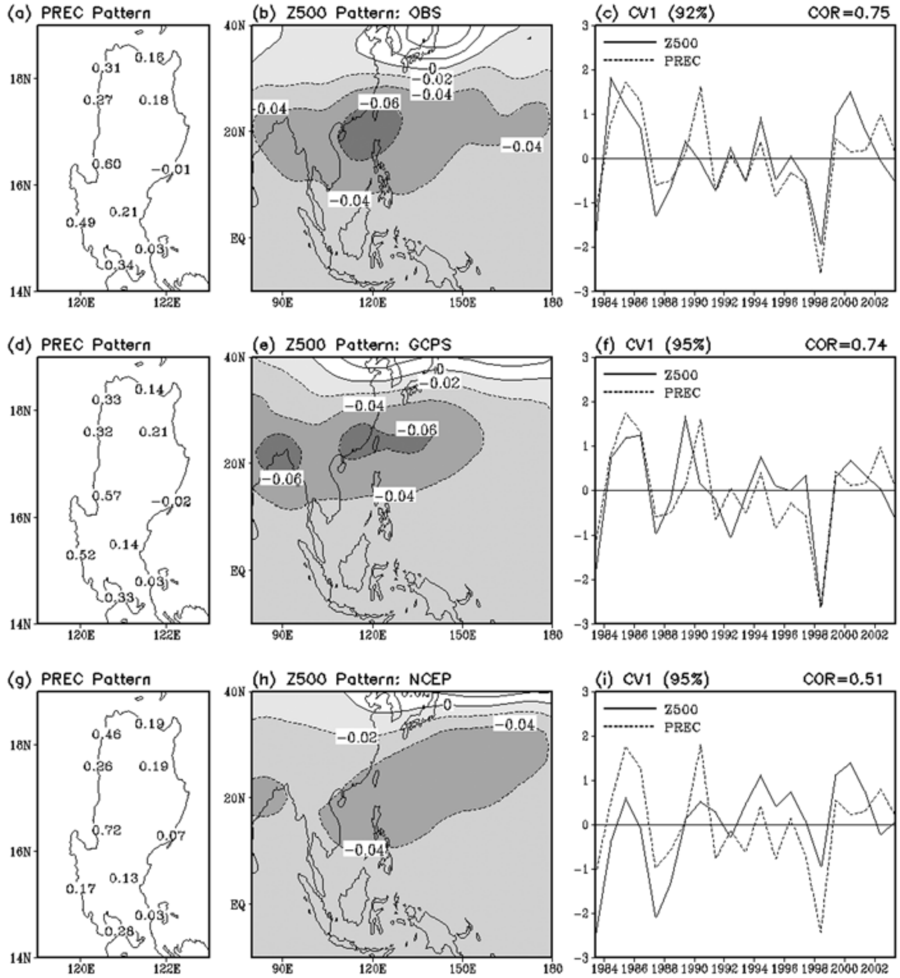

Figure 1. (a)–(c) The first SVD mode between the observed station precipitation in northern Philippines and observed Z500, and the time series of the expansion coefficient for the leading mode. (d)–(f) The first SVD mode and the time series of the expansion coefficient for GCPS. (g)–(i) The first SVD mode and the time series of the expansion coefficient for NCEP.

[15] The downscaling procedure is carried out at each station for each model. Then two MME predictions are made for comparison. Figure 2 shows the correlation coefficients between the two MME predictions and observation for each station and area average. The location of the stations is shown in Figure 2a. For downscaled MME, the skill of the area average of precipitation over the northern Philippines reaches 0.75; while for raw MME, the skill is −0.47. For most of the stations, the skills have been substantially improved. The No.9 station is the capital

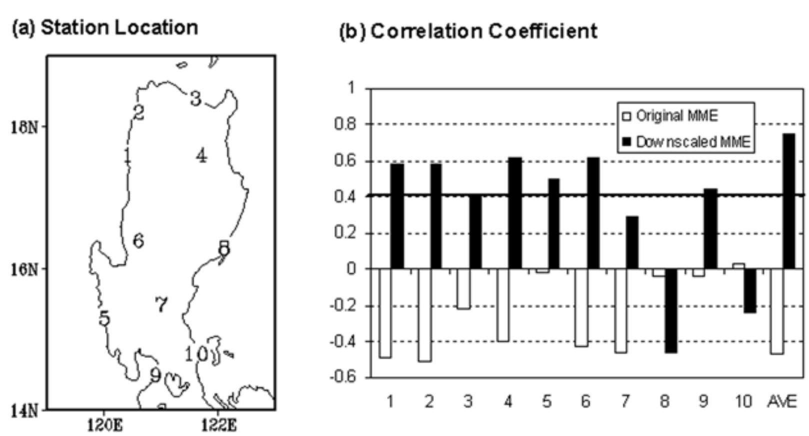

Figure 2. (a) Station location in the northern Philippines and (b) correlation coefficients between the observed station precipitation and two MME predictions: one is downscaled MME; another is raw model output MME. The solid line in Figure 2b indicates the critical value of correlation coefficient at 5% significance level.

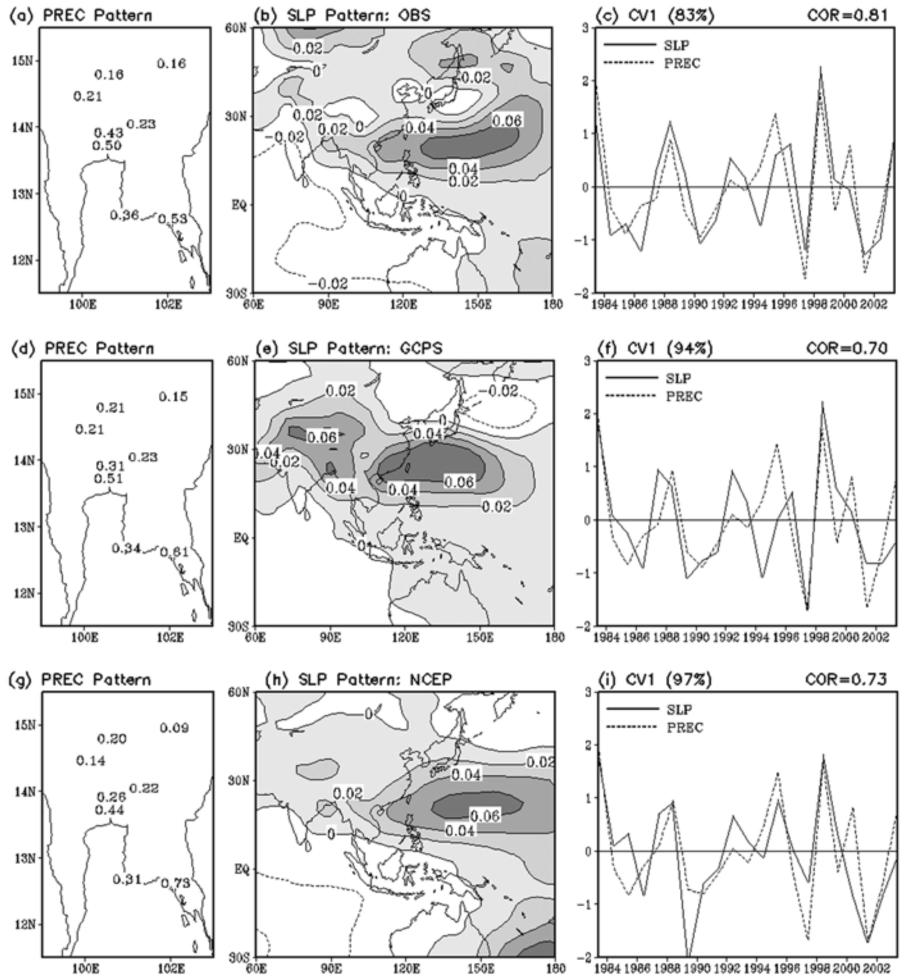

Figure 3. (a)–(c) The first SVD mode between the observed station precipitation in the Bangkok region and observed SLP, and the time series of the expansion coefficient for the leading mode. (d)–(f) The first SVD mode and the time series of the expansion coefficient for GCPS. (g)–(i) The first SVD mode and the time series of the expansion coefficient for NCEP.

Manila, the skill for downscaled MME at Manila is 0.44, while for raw MME the skill is only −0.04. However, two stations have negative correlation coefficients. From the leading covariant mode (see Figure 1), it is found that precipitation for the two stations show very weak covariant mode with the corresponding Z500 pattern, and the leading mode accounts for most of total covariance. Thus, the downscaling from Z500 can not predict variability of precipitation for the two stations.

3.2. Thailand

[16] The target area in Thailand is the Bangkok region which contains 8 stations. From the correlation analysis between the observed station precipitation and observed SLP, Z500, U850 and T850, it is found that SLP in the western North Pacific has stronger correlation than other variables. Thus SLP is selected as the predictor for downscaling.

[17] Figure 3 shows that the first SVD mode between the observed station precipitation and observed SLP explains 83% of total covariance, and the correlation coefficient of the expansion coefficients for the leading SVD mode reaches 0.81. It is found a strengthening of SLP over the South China Sea, the Philippine Sea and the western North Pacific is accompanied by an enhanced precipitation in the Bangkok region. This is dynamically reasonable: for example, a strengthening of SLP centered over the Philippine Sea favors the southeasterly passing over the Bangkok region along the South China Sea and the Gulf of Thailand, bringing more rainfall in the region. It is noted that the eigenvector values of precipitation at the 8 stations show same sign, which suggests that all the precipitation is controlled mainly by the same large-scale circulation process. This is probably because all the 8 stations are located in the same plain which is open to the Gulf of Thailand (see Figure 3a). We also take GCPS and NCEP as examples and show the leading SVD mode between observed precipitation and predicted SLP in Figure 3. Basically, both models can reproduce the leading mode of observation well.

[18] Based on the correlation analysis and the SVD analysis between observed precipitation and SLP, we think the domain of 60–180E, 30S–60N contains the large-scale circulation information for downscaling over the Bangkok region. In the same way as in the Philippines case, we set a movable window to search for the most sensible area, and then make downscaling and MME predictions. Figure 4 shows that the skill of the area average of precipitation over

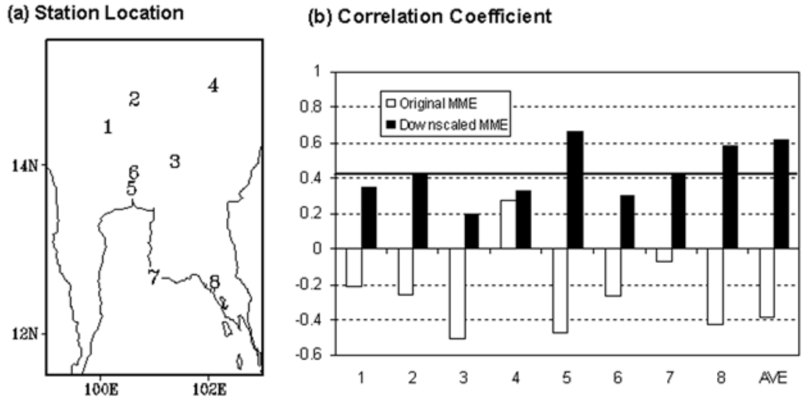

Figure 4. (a) Station location in the Bangkok region and (b) correlation coefficients between the observed station precipitation and two MME predictions: one is downscaled MME; another is raw model output MME. The solid line in Figure 4b indicates the critical value of correlation coefficient at 5% significance level.

the Bangkok region for downscaled MME reaches 0.62; while for raw MME, the skill is −0.39. For all the stations, the skills have been substantially improved. This is probably because precipitation for all the stations shows similar covariant mode with the SLP pattern. In other words, the downscaling from the coupled SLP pattern can predict variability of precipitation for all the stations if these models have good performance in predicting SLP. It is interesting to note that the correlation coefficient in the capital Bangkok (No. 5 station) reaches 0.66.

4. Summary

[19] Multimodel outputs statistical downscaling prediction of precipitation for the stations over the northern Philippines and the Bangkok region are carried out in this research. In order to choose the predictor and its range for downscaling, we used observation data other than model data to make correlation analysis and SVD analysis. The analyses based on the observed data can reveal real atmospheric dynamic linkage between precipitation and the large-scale circulation, and provide robust physical basis for the choice of predictor and its range. For the Philippines and Thailand, Z500 and SLP are selected as predictors, respectively. However, downscaling prediction of precipitation should be specified by predicted large-scale circulation information. Current GCM generally suffers from the systematic bias due to some shift of spatial pattern in prediction. In order to avoid the model bias, a movable window is set to scan over the range to select the most sensible large-scale circulation area for downscaling. The downscaling is carried out at each station for each model in a cross-validation manner. Then two MME predictions are made: one is average of downscaled precipitation from the 6 models; another is that of raw model predicted precipitation. The downscaled MME and raw model MME predictions are verified against observed station precipitation, separately. Downscaled MME achieve apparently better prediction skill than the original one. The results suggest that the predicted large-scale circulation by current GCMs have the potential in predicting station-scale precipitation by means of statistical downscaling method. In this way, we may make operational forecasts with timely collected GCM outputs for some regions, especially for some important cities such as Manila and Bangkok where a large amount of people live. However, statistical downscaling in this study has also limitation for some stations, where precipitation is governed mainly by local complicated terrain other than large-scale process. Further research is needed to use multi-predictor to make downscaling prediction for these stations.

[20] **Acknowledgments.** The authors thank the Climatology and Agrometeorology Branch of the Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAGASA) and the Thailand Meteorological Department for providing the station data of precipitation. The authors also appreciate those institutes participating in the APCC multimodel ensemble operational system for providing the hindcast experiment data.

References

Heyen, H., E. Zorita, and U. Cubasch (1996), Statistical downscaling of monthly mean North Atlantic air-pressure to sea level anomalies in the Baltic Sea, *Tellus, Ser. A*, *48*, 312–313.

Kang, I.-S., and J. Shukla (2006), Dynamical seasonal prediction and predictability of the monsoon, in *The Asian Monsoon*, edited by B. Wang, pp. 585–612, Praxis, Chichester, U.K.

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. Clim.*, *17*, 834–844.

Kug, J.-S., J.-Y. Lee, and I.-S. Kang (2007), Global sea surface temperature prediction using a multimodel ensemble, *Mon. Weather Rev.*, in press.

Palmer, T. N., et al. (2004), Development of a European multimodel ensemble system for seasonal to inter-annual prediction (DEMETER), *Bull. Am. Meteorol. Soc.*, *85*, 853–872.

Stockdale, T. N., D. L. T. Anderson, J. O. S. Alves, and M. A. Balmaseda (1998), Global seasonal rainfall forecasts using a coupled ocean-atmosphere model, *Nature*, *392*, 370–373.

von Storch, H., E. Zorita, and U. Cubasch (1993), Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime, *J. Clim.*, *6*, 1161–1171.

Wang, B., I.-S. Kang, and J.-Y. Lee (2004), Ensemble simulation of Asian-Australian monsoon variability by 11 AGCMs, *J. Clim.*, *17*, 803–818.

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

Zorita, E., and H. von Storch (1999), The analogue method as a simple statistical downscaling technique: Comparison with more complicated methods, *J. Clim.*, *12*, 2474–2489.

K.-H. An, H. Kang, and C.-K. Park, APEC Climate Center (APCC), National Pension Corporation Busan Bldg., Busan 611-705, South Korea. (hkang@apcc21.net)

A. L. S. Solis, Philippine Atmospheric, Geophysical and Astronomical Services Administration, Science Garden Complex, Agham Road, Quezon City, Philippines 1100.

K. Stitthichivapak, Thailand Meteorological Department, 4353 Sukhumvit Road, Bangkok, Thailand 10260.