ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 22, NO. 4, 2005, 479–495

Evaluation of East Asian Climatology as Simulated by Seven Coupled Models

JIANG Dabang (姜大膀)*[1,2], WANG Huijun (王会军)[1], and LANG Xianmei (郎咸梅)[1]

[1] *Nansen-Zhu International Research Center, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing* 100029

[2] *Key Laboratory of Regional Climate-Environment Research for Temperate East Asia, Chinese Academy of Sciences, Beijing* 100029

(Received 10 October 2004; revised 11 April 2005)

ABSTRACT

Using observation and reanalysis data throughout 1961–1990, the East Asian surface air temperature, precipitation and sea level pressure climatology as simulated by seven fully coupled atmosphere-ocean models, namely CCSR/NIES, CGCM2, CSIRO-Mk2, ECHAM4/OPYC3, GFDL-R30, HadCM3, and NCAR-PCM, are systematically evaluated in this study. It is indicated that the above models can successfully reproduce the annual and seasonal surface air temperature and precipitation climatology in East Asia, with relatively good performance for boreal autumn and annual mean. The models' ability to simulate surface air temperature is more reliable than precipitation. In addition, the models can dependably capture the geographical distribution pattern of annual, boreal winter, spring and autumn sea level pressure in East Asia. In contrast, relatively large simulation errors are displayed when simulated boreal summer sea level pressure is compared with reanalysis data in East Asia. It is revealed that the simulation errors for surface air temperature, precipitation and sea level pressure are generally large over and around the Tibetan Plateau. No individual model is best in every aspect. As a whole, the ECHAM4/OPYC3 and HadCM3 performances are much better, whereas the CGCM2 is relatively poorer in East Asia. Additionally, the seven-model ensemble mean usually shows a relatively high reliability.

Key words: coupled model, East Asian climatology, evaluation

1. Introduction

The fully coupled global atmosphere-ocean general circulation model (CGCM) has been one of the most important tools to explore climate and climate changes in the past, present and future Earth system, especially in the studies on the present day human-induced climate change and near-future climate scenario due to continually increased atmospheric greenhouse gases and aerosols (e.g., Hu et al., 2003). As is well known, any numerical model is only ideally similar to a real climate system (Zeng et al., 1989). Therefore, it is important to grasp such a model's climate simulation ability through effective examination in a given region. After that, we will very likely find problems and consequently improve climate models step by step. On the other hand, model evaluation can give us knowledge of what kind of scientific work is suitably performed by CGCMs, and of what kind of results derived from climate models can be further taken into account.

Studies on anthropogenic climate changes and near-future climate prediction have been carried out for over a decade in China. As for East Asian climate change under different atmospheric greenhouse gas and aerosol emission scenarios over the 21st century, many results have been derived from climate model integrations (e.g., Hulme et al., 1994; Hu et al., 2000; Bueh et al., 2003; Xu et al., 2003a; Xu et al., 2003b; Zhao et al., 2003; Jiang et al., 2004a, b). It is usually recognized that the horizontal resolution of global climate models is generally coarse, and they consequently cannot reliably describe regional topography and surface cover conditions or capture mesoscale and small-scale climate evolution processes, and this finally leads to relatively large simulation errors when global models are utilized on a regional scale

*E-mail: jiangdb@mail.iap.ac.cn

(e.g., Fu et al., 1998; Zhao and Luo, 1998). Therefore, it is quite necessary to investigate regional reliability of global CGCMs when one wants to use such model integrations to analyze climatic issues in East Asia dominated by a well-known monsoon climate system.

Previously, Zhao et al. (1995) found that large simulation uncertainties existed when they assessed several global CGCMs' performances in East Asia. They further revealed that the CGCMs' simulation abilities for surface air temperature were better relative to precipitation, and the CGCMs' reliabilities were much higher in boreal winter compared with boreal summer. CGCMs can qualitatively capture the geographical distribution pattern of East Asian surface air temperature and precipitation, and quantitative differences are, however, also noticeable. As for surface air temperature and precipitation, multi-model ensemble products are usually better than individual models. Later, Xu et al. (2002) indicated that CGCMs' performances for East Asian climatology have been improved compared to previous climate models earlier than 2002 when they briefly examined the signal of human-induced climate changes in East Asia on the basis of ECHAM4, HadCM2, GFDL-R15, CGCM1 and CSIRO-Mk2 integration outputs.

In the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC, 2001), based upon the full range of 35 emission scenarios (Nakićenović et al., 2000), a number of climate models have been used to make projections of climate change over the 21st century, and surface air temperature change trends were given out on global and regional scales. Therefore, it is of interest to examine such CGCMs' climatology performances under the present day conditions on a regional scale, such as the East Asian monsoon region, in consideration of its close relationship to regional prediction confidence. In this study, seven CGCMs, having been used to predict anthropogenic climate change over the 21st century, are systematically evaluated in East Asia, with emphasis on the performance of individual and multi-model ensemble mean climatology throughout 1961 to 1990 because it bears partly on whether such climate models can provide useful prediction products for the near future.

2. CGCM output and observation-based data

Annual and seasonal surface air temperature, precipitation and sea level pressure climatology are evaluated in the following because they are the most suitable climatic indicators for climate change, at the same time taking into account the fact that they are the only climatic variables provided by every CGCM through public websites (see next paragraph). Additionally, we follow the seasons according to the Northern Hemisphere throughout this study, such as summer corresponding to June, July and August.

The seven climate models involved in this study have previously made predictions of climate change over the 21st century under several atmospheric greenhouse gas and aerosol emission scenarios, such as SRES A2 and B2 scenarios. All model outputs are provided by the IPCC Data Distribution Centre (http://pluto.dkrz.de/IPCC_DDC) and the Hadley Centre for Climate Prediction and Research. For model information in more detail, see Table 1. During 1961 to 1990, the focus period in this analysis, each model is forced by the present day boundary conditions and observational annual atmospheric CO_2 levels (including effects of other greenhouse gases) and sulphate aerosol concentrations. The climatic data used to assess model performance include observation-based sea level pressure reanalysis data from the U.S. National Centers for Environmental Prediction and National Center for Atmospheric Research (NCEP/NCAR) (Kalnay et al., 1996) and observational monthly terrestrial surface air temperature and precipitation climatology throughout 1961 to 1990 from the Climatic Research Unit of the University of East Anglia in the United Kingdom (New et al., 1999). In consideration of the differences of horizontal resolution among the models, linear interpolation or extrapolation methods are applied to different model climatologies in order to generate a uniform dataset with a consistent horizontal grid resolution of 3.75° by 3.75°.

3. Evaluation method

Climate models are generally evaluated by investigating the validity of the physical processes and parameterizations, climate reproducibility under the present conditions, and sensitivities of perturbation experiments (Zhao et al., 1995). In the present study, the models' ability to reproduce climatology during 1961–1990 is examined by the following two assessments. Firstly, the simulated climatologies are directly compared with observation in order to get an overview of the models' capability in East Asia. Then, the following six statistical variables are used to quantitatively assess the simulated East Asian climatologies within the region of 15°–60°N and 70°–140°E, namely regional average (RA, V_{RA}), regional average error (RAE, V_{RAE}), spatial correlation coefficient (SCC, V_{SCC}), ratio of standard deviation (RSD, V_{RSD}), root

Table 1. Information on the seven climate models used in this study.

Model name	Country, Affiliation acronym	Atmospheric model resolution (longitude by latitude)	Oceanic model resolution (longitude by latitude)	Reference
CCSR/NIES	Japan, CCSR/NIES	T21(5.625° × 5.6°)/L20	2.8° × 2.8°/L17	Emori et al. (1999)
CGCM2	Canada, CCCma	T32(3.75° × 3.7°)/L10	1.8° × 1.8°/L29	Flato and Boer (2001)
CSIRO-Mk2	Australia, CSIRO	R21(5.625° × 3.2°)/L9	3.2° × 5.6°/L21	Gordon and O′Farrell (1997)
ECHAM4/OPYC3	German, DKRZ	T42(2.8125° × 2.8°)/L19	2.8° × 2.8°/L11	Roeckner et al. (1996)
GFDL-R30	America, GFDL	R30(3.75° × 2.25°)/L14	1.875° × 2.25°/L18	Knutson et al. (1999)
HadCM3	England, UKMO	3.75° × 2.5°/L19	1.25° × 1.25°/L20	Gordon et al. (2000)
NCAR-PCM	America, NCAR	T42(2.8125° × 2.8°)/L18	0.67° × 0.67°/L32	Washington et al. (2000)

mean square error (RMSE, V_{RMSE}), and RMSE excluding systematic model errors (RMSE2, $\mathrm{V}_{\mathrm{RMSE2}}$). Each statistical variable is listed in the following.

$$V_{\mathrm{RA}}=\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i\ ,$$

$$V_{\mathrm{RAE}}=\bar{x}-\bar{y}=\frac{1}{n}\sum_{i=1}^{n}x_i-\frac{1}{n}\sum_{i=1}^{n}y_i\ ,$$

$$V_{\mathrm{SCC}}=\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2}}\ ,$$

$$V_{\mathrm{RSD}}=\frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2}}\ ,$$

$$V_{\mathrm{RMSE}}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-y_i)^2}\ ,$$

$$V_{\mathrm{RMSE2}}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}[x_i-y_i-(\bar{x}-\bar{y})]^2}\ .$$

Here, x_i (y_i) denotes the simulation (observation) value at the i-th spatial gridpoint of the climatology, and n is the total number of gridpoints within the domain. According to statistical theory, RAE can measure regional differences of CGCM simulation against observation, SCC denotes the similarity of geographical distribution pattern between simulation and observation, RSD describes the contrast of spatial variational spread of simulated climatology relative to observation, RMSE quantifies absolute model-observation differences, and RMSE2 compared with RMSE indicates the magnitude of systematic model errors.

4. Evaluation results

4.1 *Surface air temperature*

It is interesting to examine whether the climate models can successfully capture the annual and seasonal surface air temperature climatology in East Asia because it is directly connected with our confidence to use such models to forecast anthropogenic climate changes in the near future. As shown in Fig. 1, observational annual surface air temperature is distributed in a band parallel with latitude which gradually enlarges northward. A large extent of cooling is present over the Tibetan Plateau, which can be ascribed to in situ high topography. In contrast to the cooling, surface air temperature warms over most parts of Xinjiang Province situated northwest of the plateau, with a maximum of about 9°C, where the topography is much lower than the plateau. It can be found that the seven-model ensemble mean coincides well with the above spatial pattern in East Asia. However, discrepancies are also exhibited. In particular, the seven-model ensemble mean is cooled by 0° to 4.5°C compared with the observation in East Asia, with two maximum cooling bias centers occurring over southeastern Xinjiang Province and the middle reaches of the Yangtze River Valley, respectively. In addition, the simulated surface air temperature contours are ob-

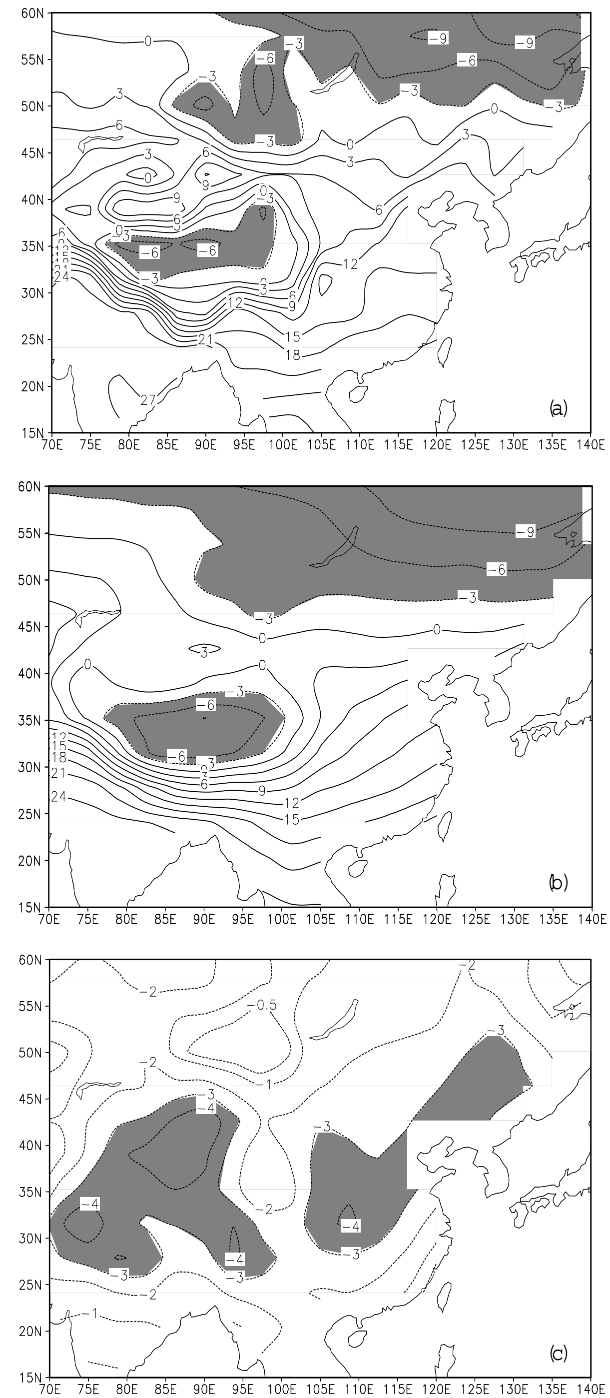

Fig. 1. Annual terrestrial surface air temperature climatology throughout 1961 to 1990 (units: °C). (a) CRU data, (b) the seven-model ensemble mean, and (c) the nine-point running average differences of the ensemble mean against the CRU data. The regional minimum and/or maximum centers, with variational value level, are shaded.

viously smoother than the observation and therefore cannot capture observational regional climate signals, and especially lacking is the regional warming center in Xinjiang Province. It is plausible that the CGCMs' ability to simulate surface air temperature is poor nearby steep topography, such as around the Tibetan Plateau.

In general, the observational spatial distribution pattern of surface air temperature in East Asia is reliably reproduced by the models on the seasonal timescale (Figs. 2–5). However, the difference field of seasonal surface air temperature between the seven-model ensemble mean and observation is also pronounced. It is noted that seasonal surface air temperature is simulated as uniformly lower relative to the observation over mainland China, with a relatively larger cooling bias in winter (Fig. 2) and spring (Fig. 3).

To quantitatively assess the models' performance, the values of the six statistical variables described in section 3 are given for each of the models and the seven-model ensemble mean in Table 2. As for annual surface air temperature, the RAE of the seven-model ensemble mean reaches −2.07°C, denoting the models' cool simulation bias. The RMSE and RMSE2 are 3.40°C and 2.70°C, respectively, suggesting both absolute and systematic simulation errors are embedded in the models, and simulation errors are mainly derived from the models' intrinsic physical processes and parameterizations. However, the SCC and RSD are as high as 0.97 and 0.98, respectively, indicating that the spatial distribution pattern of annual surface air temperature is quite well depicted by the seven-model ensemble mean in East Asia. In addition, the HadCM3 is found to be the best model to describe East Asian annual surface air temperature, and the ECHAM4/OPYC3 is the second best model. Moreover, both of the above two models show a better performance than the seven-model ensemble mean. It is also noted that the CGCM2 and CCSR/NIES exhibit a relatively poor simulation, with the RMSE reaching 7.10°C and 5.17°C, respectively, which is obviously greater than the ensemble mean.

For the seasonal climatology, cold simulation errors are generally present in each of the four seasons. In winter, the RAE and RMSE of the seven-model ensemble mean are respectively −2.64°C and 4.46°C. However, the SCC and RSD still keep their high level as that of annual mean, which means that the ensemble mean can capture the geographical distribution features of the East Asian winter surface air temperature. As a whole, both the ECHAM4/OPYC3 and HadCM3 performances are better than the ensemble mean, and the simulation ability of the CGCM2, CCSR/NIES, and CSIRO-Mk2 is relatively poor, especially for the

Table 2. The values of six statistical variables for surface air temperature in East Asia.

	RA (°C)	RAE (°C)	SCC	RSD	RMSE (°C)	RMSE2 (°C)
Annual mean (observational RA equals 6.49°C)						
CCSR/NIES	3.94	−2.55	0.92	1.06	5.17	4.50
CGCM2	1.34	−5.15	0.90	0.97	7.10	4.88
CSIRO-Mk2	3.84	−2.65	0.95	0.97	4.46	3.59
ECHAM4/OPYC3	6.67	0.18	0.97	1.04	2.86	2.85
GFDL-R30	6.04	−0.45	0.96	1.00	3.24	3.21
HadCM3	5.03	−1.46	0.98	1.03	2.61	2.16
NCAR-PCM	4.05	−2.44	0.97	0.94	3.76	2.87
Ensemble mean	4.42	−2.07	0.97	0.98	3.40	2.70
Winter (observational RA equals −7.45°C)						
CCSR/NIES	−9.61	−2.16	0.95	1.03	5.68	5.26
CGCM2	−15.07	−7.62	0.89	0.95	10.56	7.31
CSIRO-Mk2	−10.22	−2.77	0.94	0.87	6.08	5.41
ECHAM4/OPYC3	−6.83	0.62	0.98	1.01	3.34	3.28
GFDL-R30	−10.83	−3.38	0.97	1.02	5.37	4.17
HadCM3	−9.20	−1.75	0.98	0.99	3.84	3.42
NCAR-PCM	−8.90	−1.45	0.96	0.88	4.67	4.44
Ensemble mean	−10.09	−2.64	0.98	0.95	4.46	3.59
Spring (observational RA equals 7.49°C)						
CCSR/NIES	2.66	−4.83	0.90	1.32	8.31	6.76
CGCM2	−2.03	−9.52	0.87	1.12	11.43	6.33
CSIRO-Mk2	0.82	−6.67	0.95	1.20	8.09	4.59
ECHAM4/OPYC3	6.88	−0.61	0.97	1.12	3.41	3.35
GFDL-R30	6.16	−1.33	0.95	1.16	4.54	4.34
HadCM3	6.31	−1.18	0.98	1.07	2.66	2.38
NCAR-PCM	3.77	−3.72	0.96	0.99	4.93	3.24
Ensemble mean	3.51	−3.98	0.96	1.11	5.28	3.47
Summer (observational RA equals 18.86°C)						
CCSR/NIES	17.38	−1.48	0.77	0.81	4.64	4.40
CGCM2	16.45	−2.41	0.86	1.05	4.48	3.78
CSIRO-Mk2	18.93	0.07	0.85	0.96	3.62	3.62
ECHAM4/OPYC3	19.29	0.43	0.91	1.07	3.15	3.12
GFDL-R30	22.69	3.83	0.81	0.88	5.55	4.01
HadCM3	18.18	−0.68	0.96	1.09	2.22	2.11
NCAR-PCM	15.79	−3.07	0.91	0.97	4.22	2.90
Ensemble mean	18.39	−0.47	0.92	0.92	2.65	2.60
Autumn (observational RA equals 7.07°C)						
CCSR/NIES	5.33	−1.74	0.92	0.94	4.54	4.20
CGCM2	6.03	−1.04	0.90	0.80	4.89	4.77
CSIRO-Mk2	5.83	−1.24	0.94	0.94	3.90	3.69
ECHAM4/OPYC3	7.42	0.35	0.96	0.99	3.03	3.01
GFDL-R30	6.17	−0.90	0.96	0.98	3.22	3.09
HadCM3	4.88	−2.19	0.98	1.03	3.09	2.17
NCAR-PCM	5.64	−1.43	0.97	0.91	3.25	2.92
Ensemble mean	5.90	−1.17	0.97	0.92	2.98	2.74

former whose RAE and RMSE reach −7.62°C and 10.56°C, respectively. The models' simulation capacity in spring is similar to that in winter, except for larger cold simulation errors (−3.98°C) and systematic model errors. In summer, the RAE of the seven-model ensemble mean is down to −0.47°C, meaning relatively small systematic model errors. However, the model performance in this season declines obviously because

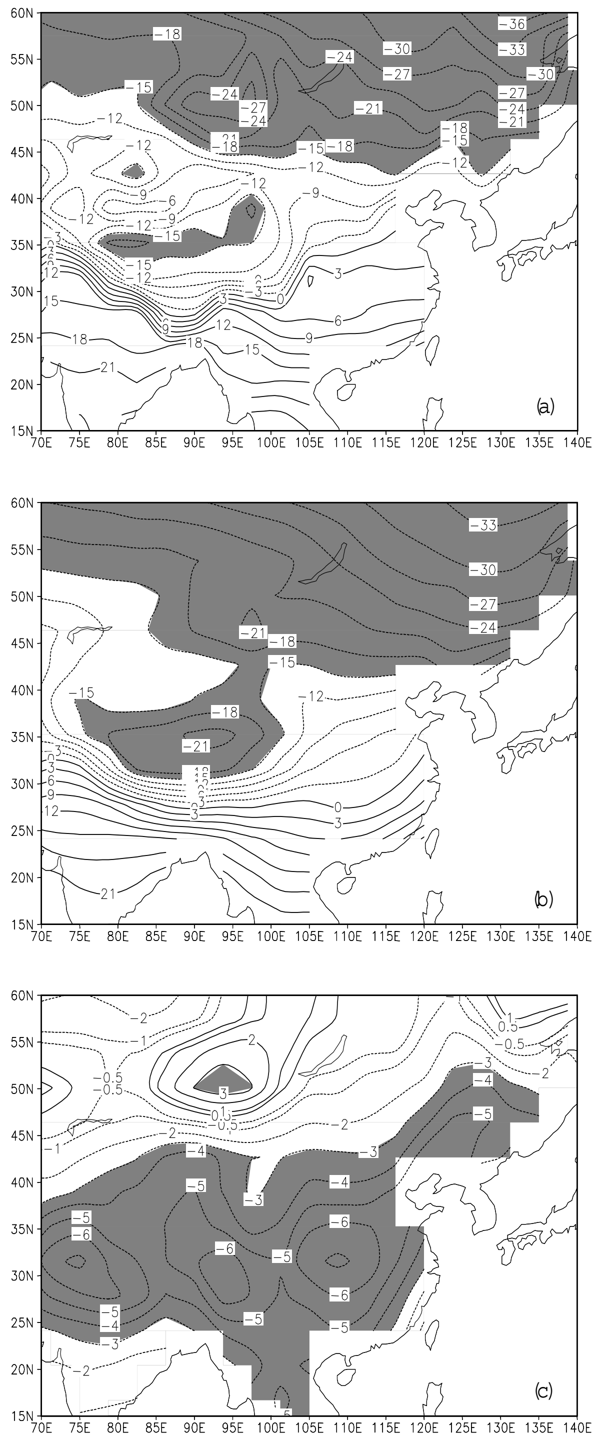

Fig. 2. Same as Fig. 1, but for winter.

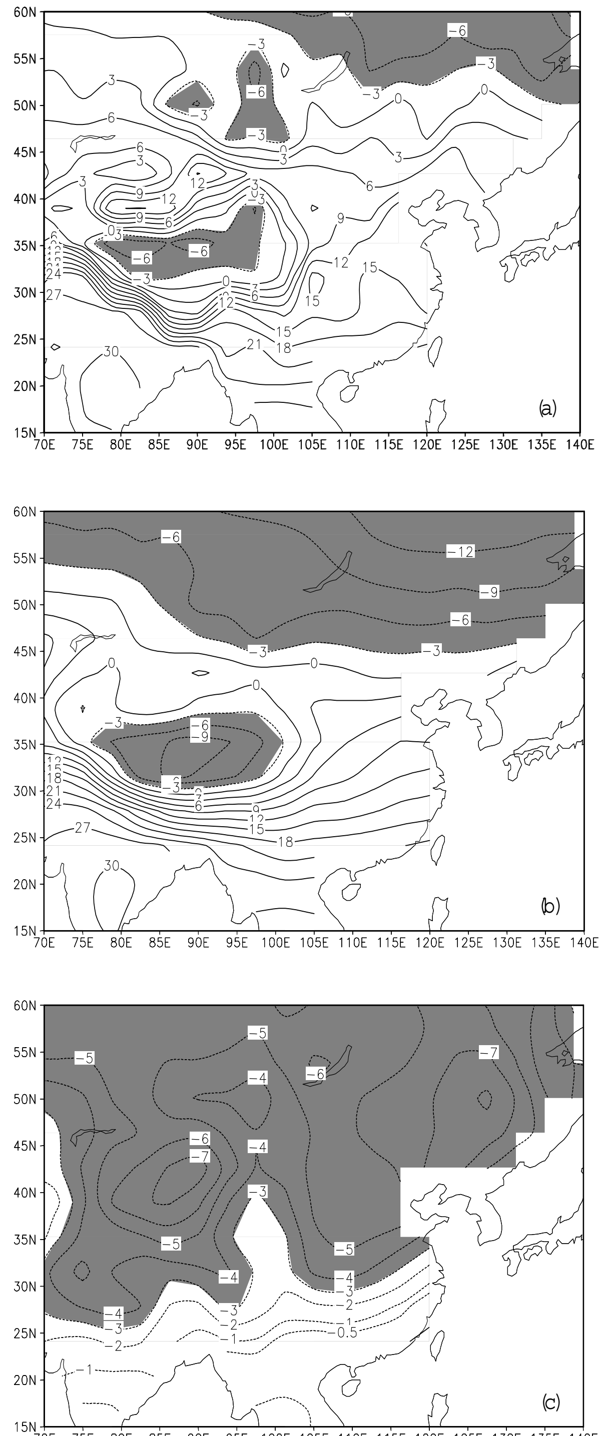

Fig. 3. Same as Fig. 1, but for spring.

both the SCC and RSD are only 0.92, being much smaller than the other seasons. In addition, it is found that the HadCM3 still exhibits a better score than the ensemble mean, and the CGCM2, CCSR/NIES, CSIRO-Mk2 and GFDL-R30 are relatively poor. In general, the models' simulation capacity is best in autumn, with relatively smaller regional average simulation errors (−1.17°C) than winter and spring and a larger spatial correlation coefficient (0.97) than summer for the seven-model ensemble mean.

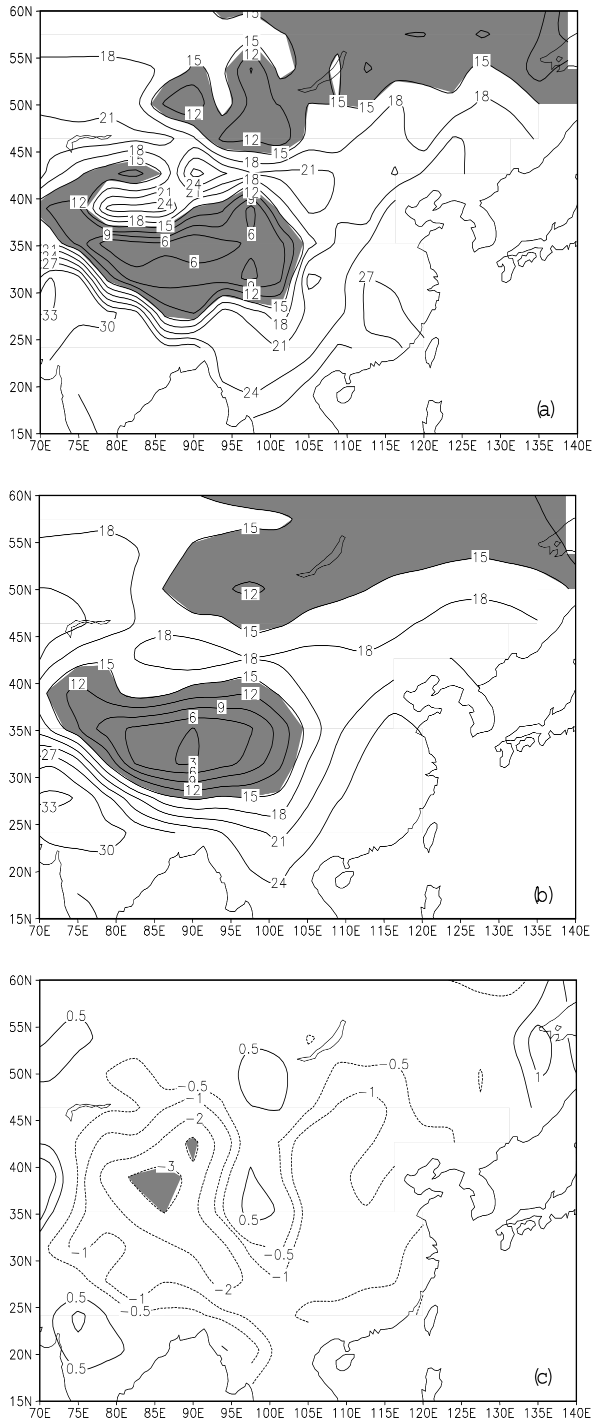

Fig. 4. Same as Fig. 1, but for summer.

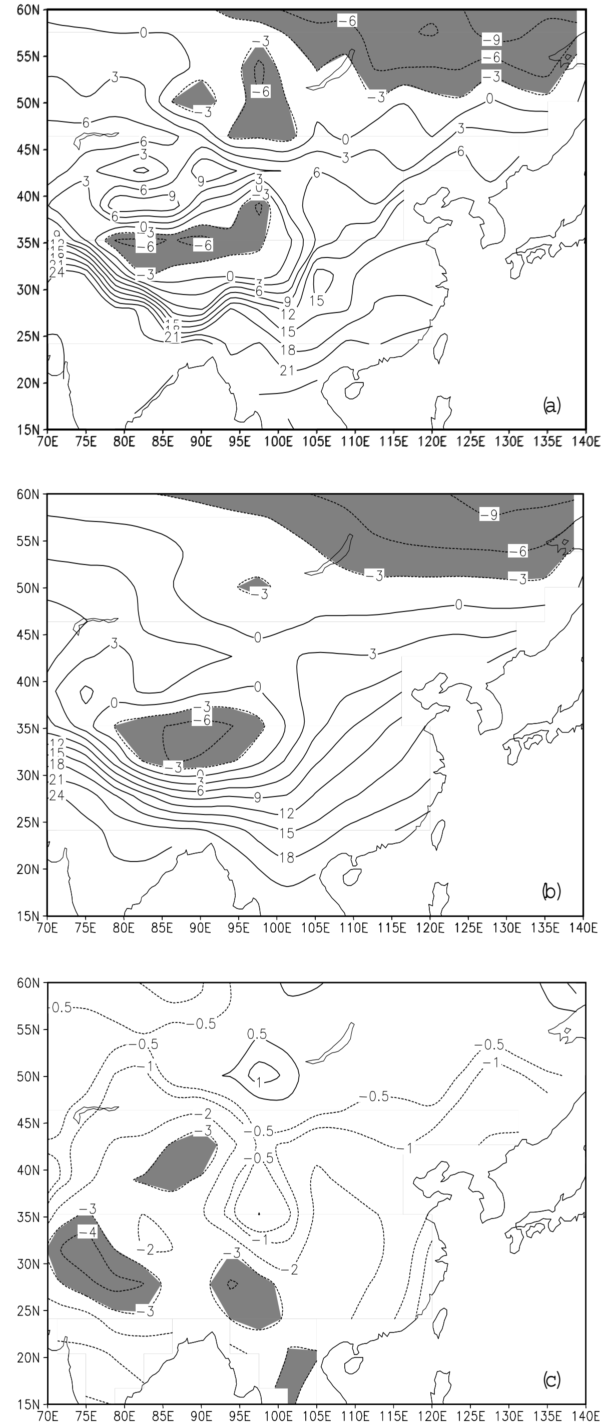

Fig. 5. Same as Fig. 1, but for autumn.

Based on the above, it can be said that the models addressed in the present research can reliably reproduce the annual and seasonal surface air temperature climatology in East Asia. Comparatively, the spatial distribution pattern of surface air temperature is poorly simulated in summer relative to the other three seasons and annual mean, the regional average of simulation errors is larger for winter, spring and annual mean, and the intrinsic and systematic model errors are pronounced in winter and spring. As a whole, the model performance in autumn is best, and annual mean is second best. In addition, it is revealed that

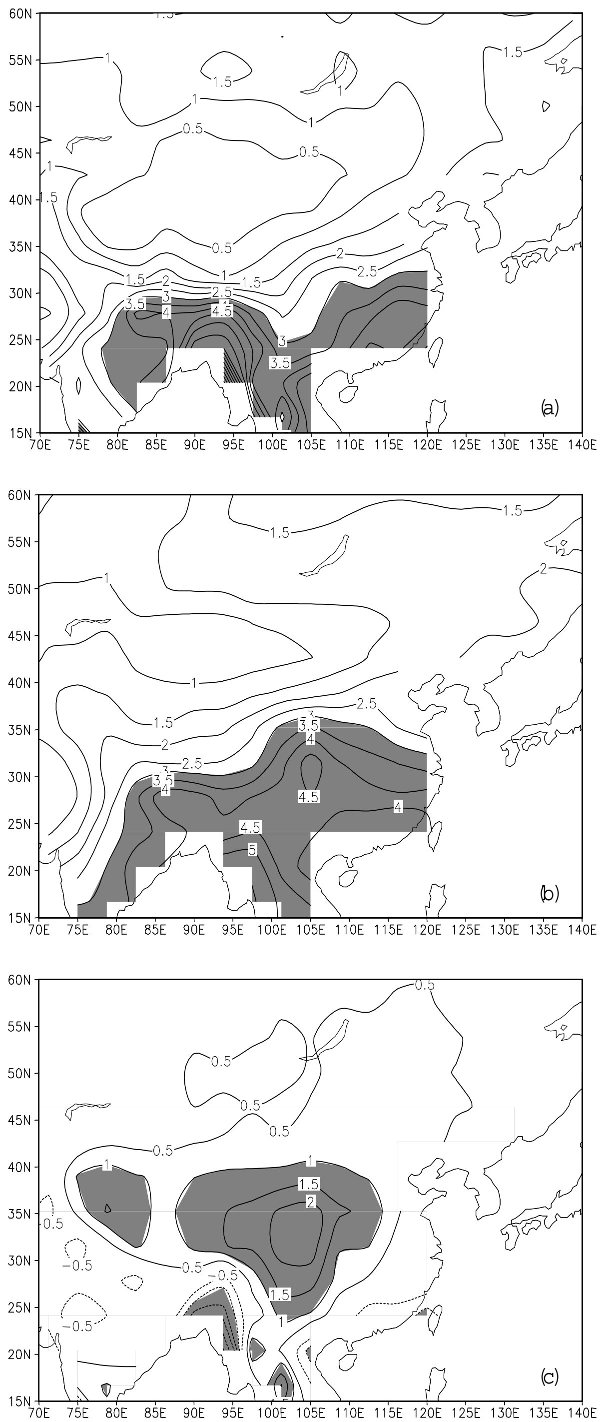

Fig. 6. Annual terrestrial precipitation climatology throughout 1961 to 1990 (units: mm d^{-1}). (a) CRU data, (b) the seven-model ensemble mean, and (c) the nine-point running average differences of the ensemble mean against the CRU data. The regional minmum and/or maximum centers, with variational value level, are shaded.

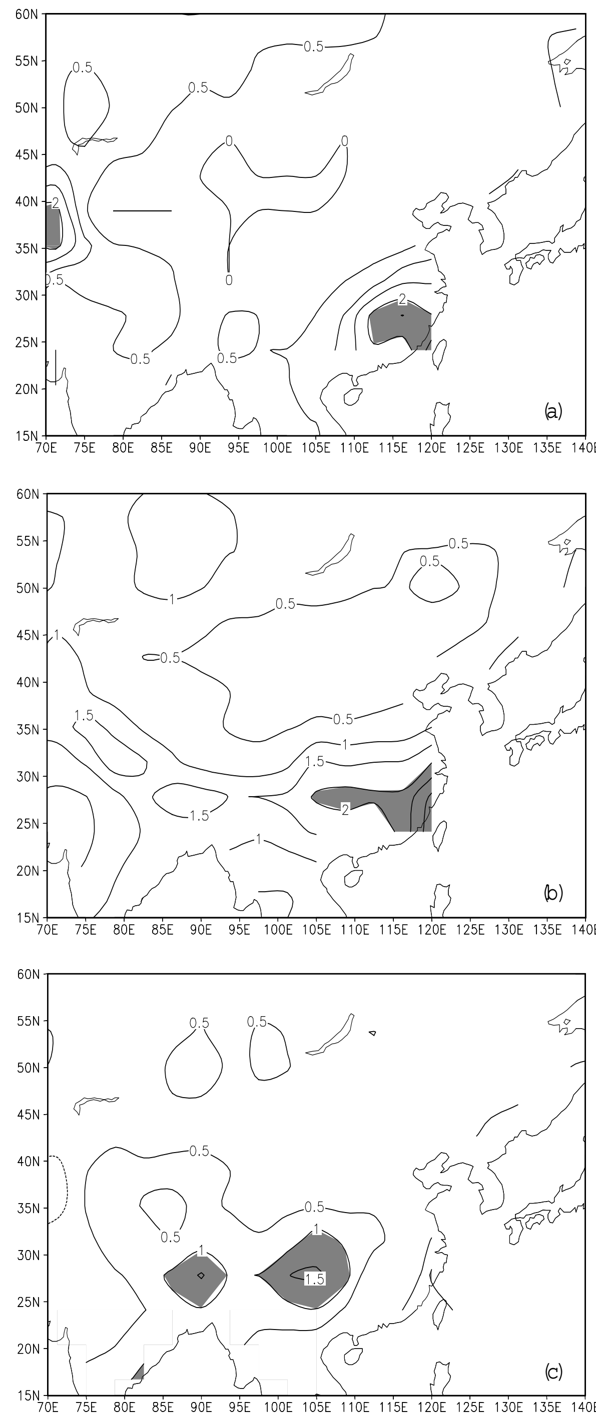

Fig. 7. Same as Fig. 6, but for winter.

the HadCM3 and ECHAM4/OPYC3 simulation abilities to reproduce annual and seasonal surface air temperature in East Asia are best among the models and generally better than the seven-model ensemble mean. In contrast, the CGCM2 gives relatively larger simulation errors.

4.2 *Precipitation*

Precipitation is one of the climatic variables to

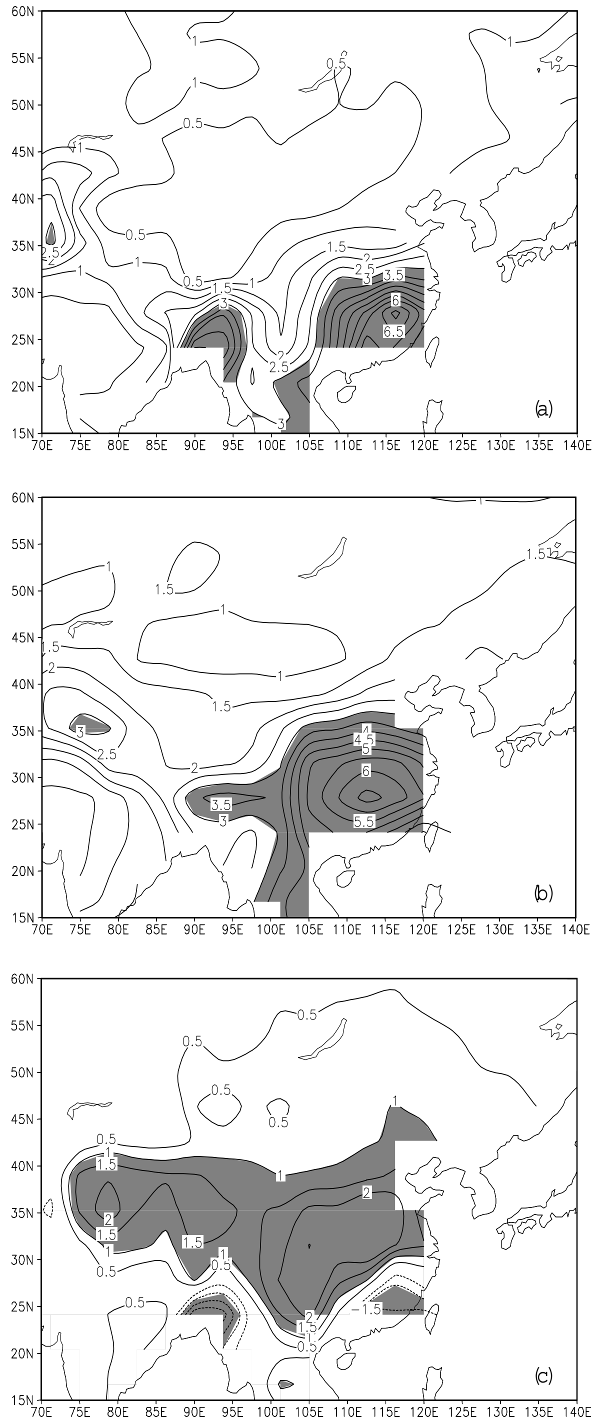

Fig. 8. Same as Fig. 6, but for spring.

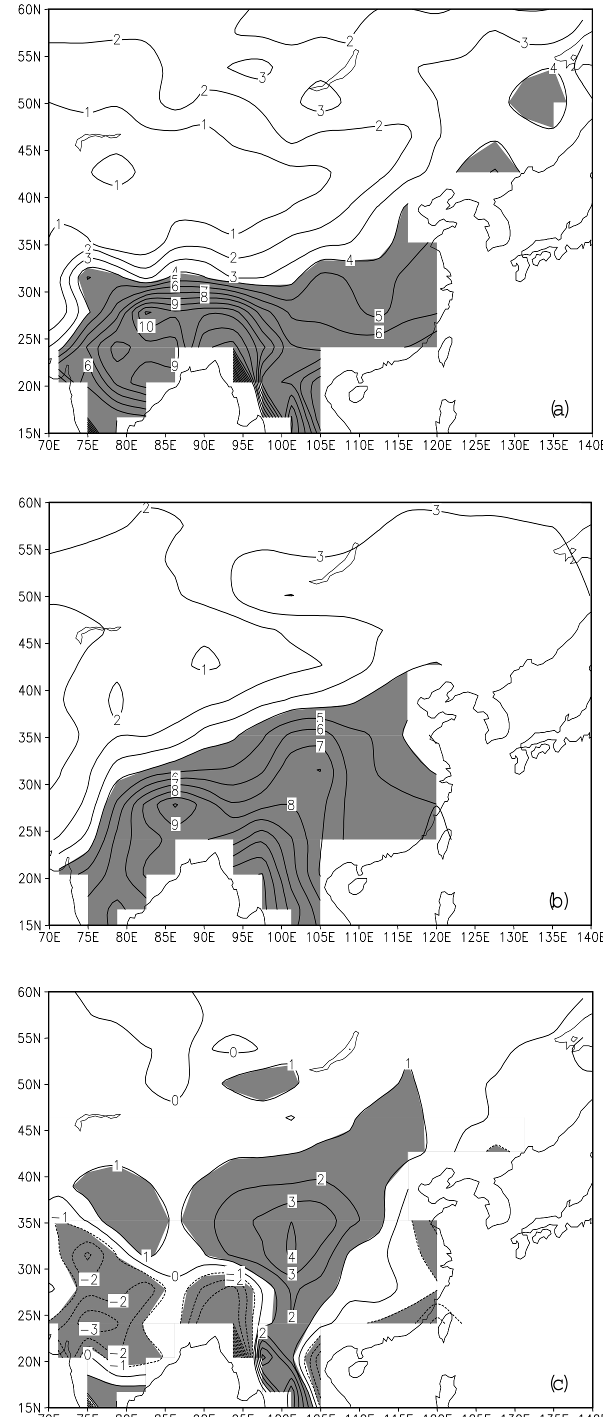

Fig. 9. Same as Fig. 6, but for summer.

have notably important influences on the Earth environment and on socio-economic development. Therefore, it is necessary to examine to what extent the models can reproduce the East Asian annual and seasonal precipitation. As displayed in Fig. 6, the annual precipitation as described by the seven-model ensemble mean is in general agreement with the observation, for example much more precipitation in southern China, larger longitudinal precipitation gradient south of 35°N, gradual decrease toward high latitudes, and so on. However, the ensemble mean is generally greater than the observation, especially in the regions south

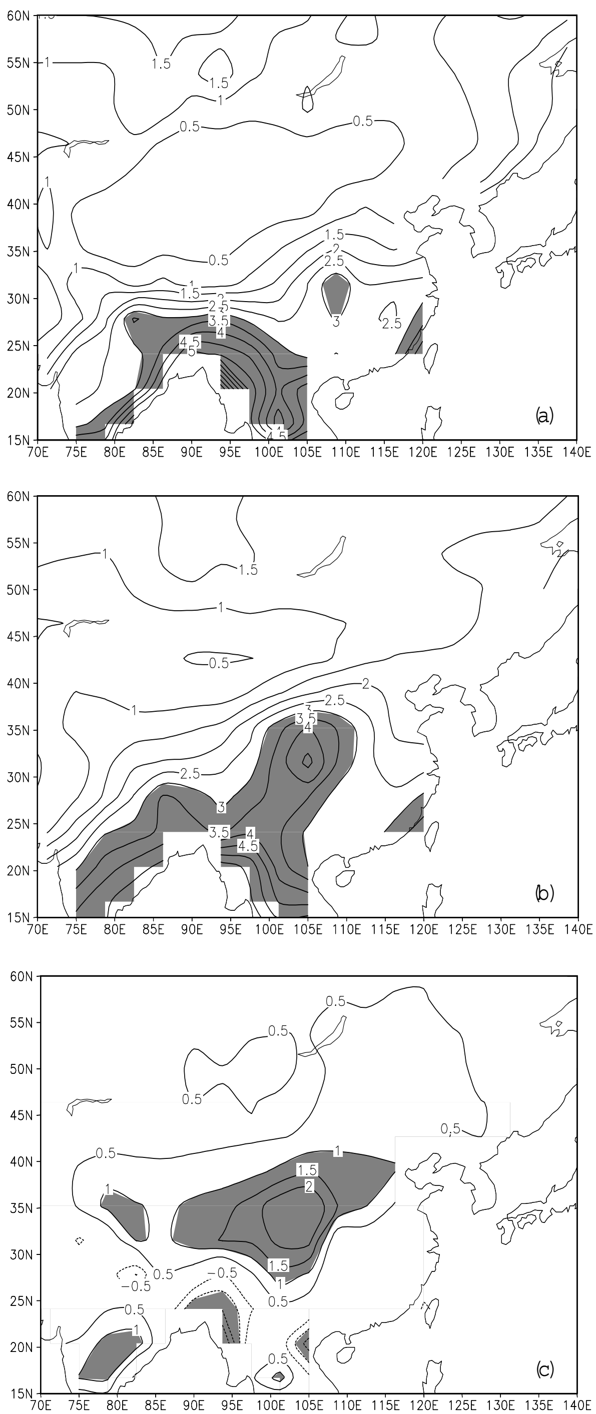

Fig. 10. Same as Fig. 6, but for autumn.

of 40°N where annual precipitation is generally overestimated by greater than 0.5 mm d^{-1}. It is noted that the simulated annual precipitation is 1.0–2.0 mm d^{-1} greater than the observation over the Tibetan Plateau and about 2.0 mm d^{-1} above the observation east of the plateau.

Although seasonal precipitation is usually overestimated by the models over mainland China, its spatial distribution pattern is, to a certain extent, reliably reproduced in East Asia (Figs. 7–10). Overall, the seven-model ensemble mean precipitation agrees with the observation in summer (Fig. 9), which is consistent with the ten atmospheric general circulation models' ensemble results as examined by Kang et al. (2002). However, it is found that the simulated summer precipitation is 0–1 mm d^{-1} weaker than the observation in mid-eastern Northeast China, southern China, and eastern China along the coastline. It is very likely that the simulated East Asian summer monsoon in the central part of eastern China (along the middle and lower reaches of the Yangtze River Valley) is weaker than the observation. It is unfortunate that we have no access to low-tropospheric wind data and consequently cannot perform a further examination. On the other hand, summer precipitation is overestimated in the rest of China. The simulated precipitation is 1–4 mm d^{-1} higher relative to the observation over the western, mid- to eastern, and east of the Tibetan Plateau. In the other three seasons (Figs. 7–8, 10), the models can reproduce the geographical distribution pattern of East Asian precipitation, such as much more precipitation in southern China. However, the simulated magnitude is generally greater than the observation over mainland China, especially in the regions south of 40°N.

As listed in Table 3, the RAE of ensemble mean annual precipitation is only 0.36 mm d^{-1}, the SCC and RSD are 0.86 and 0.78, and the RMSE and RMSE2 respectively reach 0.89 mm d^{-1} and 0.82 mm d^{-1}, all of which denote that the models reliably reproduce the spatial distribution pattern of East Asian annual precipitation. In addition, it is found that the ensemble mean generally exhibits much better performance than any individual model. Comparatively, the ECHAM4/OPYC3 and HadCM3 performances are slightly worse than the ensemble mean, yet the above two models are, however, still most reliable for annual precipitation among the models. Except for the GFDL-R30, all of the six other models underestimate the spatial variational magnitude of annual precipitation in East Asia.

The models' ability to simulate East Asian seasonal precipitation climatology varies with season. Intercomparison indicates that the models' scores are highest in autumn, when the SCC of the seven-model ensemble mean reaches 0.88, and the RAE, RMSE and RMSE2 are relatively smaller than the other three seasons. It is found that the spread of inter-model SCC is larger in winter relative to the other seasons, with

Table 3. The values of six statistical variables for precipitation in East Asia.

	RA (mm d^{-1})	RAE (mm d^{-1})	SCC	RSD	RMSE (mm d^{-1})	RMSE2 (mm d^{-1})
Annual mean (observational RA equals 1.89 mm d^{-1})						
CCSR/NIES	2.67	0.78	0.77	0.69	1.29	1.02
CGCM2	2.62	0.73	0.65	0.91	1.46	1.26
CSIRO-Mk2	2.01	0.12	0.78	0.77	1.00	1.00
ECHAM4/OPYC3	1.93	0.04	0.83	0.71	0.91	0.91
GFDL-R30	2.81	0.92	0.79	1.44	1.68	1.41
HadCM3	2.12	0.23	0.81	0.95	0.98	0.95
NCAR-PCM	1.57	−0.32	0.66	0.70	1.23	1.19
Ensemble mean	2.25	0.36	0.86	0.78	0.89	0.82
Winter (observational RA equals 0.52 mm d^{-1})						
CCSR/NIES	1.20	0.68	0.52	1.17	0.96	0.67
CGCM2	1.46	0.94	0.41	1.46	1.29	0.88
CSIRO-Mk2	0.91	0.39	0.76	1.62	0.78	0.68
ECHAM4/OPYC3	0.79	0.27	0.83	1.28	0.54	0.46
GFDL-R30	1.21	0.69	0.64	2.25	1.32	1.21
HadCM3	0.77	0.25	0.84	1.35	0.54	0.47
NCAR-PCM	0.48	−0.04	0.76	0.89	0.42	0.42
Ensemble mean	0.98	0.46	0.86	1.12	0.58	0.36
Spring (observational RA equals 1.40 mm d^{-1})						
CCSR/NIES	2.28	0.88	0.69	0.88	1.41	1.10
CGCM2	2.03	0.63	0.57	1.02	1.53	1.39
CSIRO-Mk2	1.73	0.33	0.78	0.98	1.04	0.99
ECHAM4/OPYC3	1.89	0.49	0.76	0.89	1.10	0.99
GFDL-R30	2.54	1.14	0.72	1.53	1.94	1.59
HadCM3	2.05	0.65	0.80	1.19	1.25	1.07
NCAR-PCM	1.51	0.11	0.61	0.72	1.19	1.19
Ensemble mean	2.01	0.61	0.82	0.90	1.06	0.87
Summer (observational RA equals 3.95 mm d^{-1})						
CCSR/NIES	4.76	0.81	0.70	0.78	2.65	2.52
CGCM2	4.54	0.59	0.61	0.91	3.02	2.96
CSIRO-Mk2	3.42	−0.53	0.65	0.57	2.74	2.68
ECHAM4/OPYC3	3.24	−0.71	0.82	0.64	2.22	2.10
GFDL-R30	5.12	1.17	0.75	1.47	3.62	3.43
HadCM3	3.94	−0.01	0.79	0.80	2.13	2.13
NCAR-PCM	3.03	−0.92	0.59	0.76	3.04	2.89
Ensemble mean	4.01	0.06	0.82	0.73	2.04	2.03
Autumn (observational RA equals 1.68 mm d^{-1})						
CCSR/NIES	2.43	0.75	0.80	0.90	1.23	0.98
CGCM2	2.41	0.73	0.65	0.98	1.51	1.32
CSIRO-Mk2	1.96	0.28	0.68	0.84	1.22	1.18
ECHAM4/OPYC3	1.79	0.11	0.80	0.71	0.97	0.96
GFDL-R30	2.38	0.70	0.82	1.36	1.43	1.25
HadCM3	1.73	0.05	0.81	0.92	0.95	0.95
NCAR-PCM	1.25	−0.43	0.64	0.78	1.32	1.24
Ensemble mean	1.99	0.31	0.88	0.79	0.82	0.76

the maximum of 0.86 for the ensemble mean. That means the multi-model ensemble mean is preferable compared with any individual model in the season if one aims for much more reliable information. Larger RMSE and RMSE2 in summer reflect that the intrinsic model errors are robust in this season. Additionally, it is revealed that the annual and seasonal precipitation are reproduced best by the seven-model ensemble

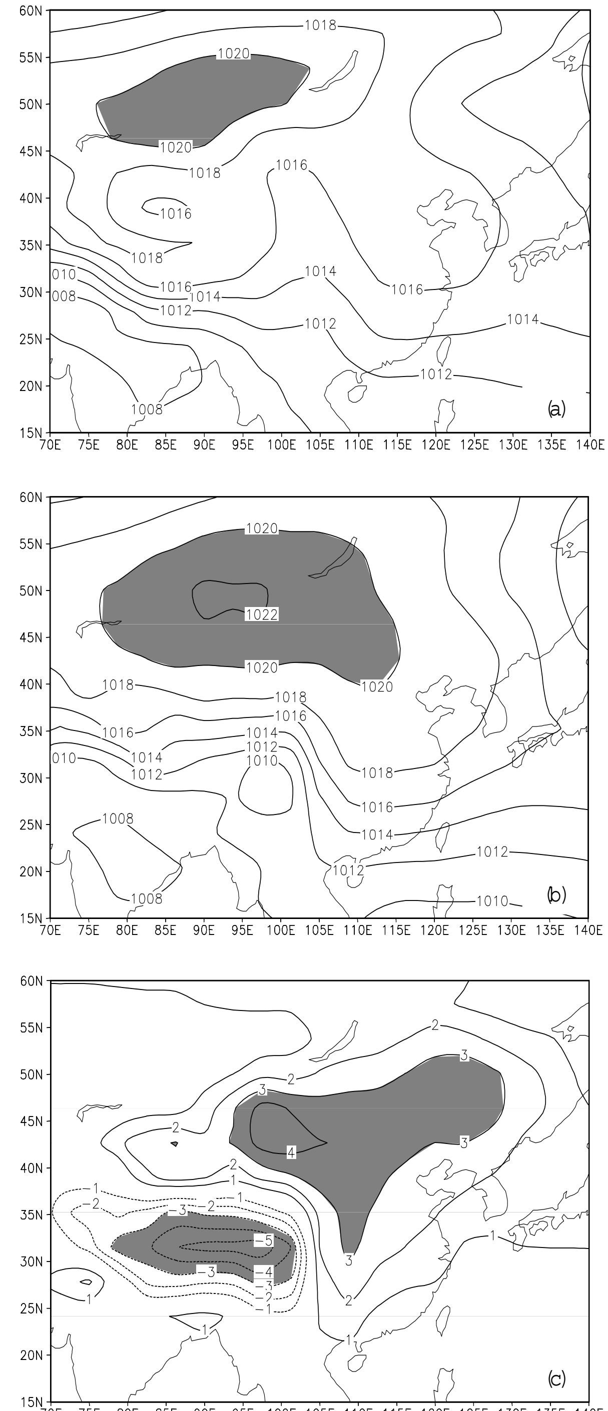

Fig. 11. Annual sea level pressure climatology throughout 1961 to 1990 (units: hPa). (a) NCEP/NCAR reanalysis data, (b) the six-model ensemble mean except for the GFDL-R30, and (c) the differences of the ensemble mean against the NCEP/NCAR reanalysis data. The regional minimum and/or maximum centers, with variational value level, are shaded.

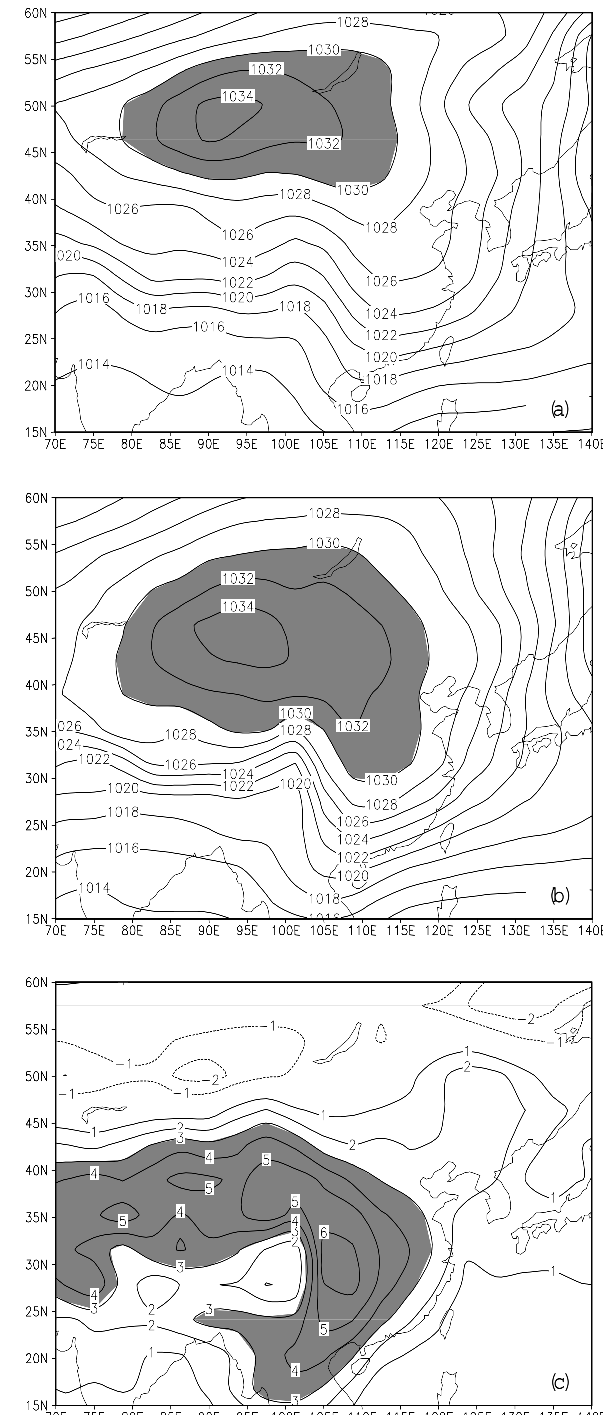

Fig. 12. Same as Fig. 11, but for winter.

mean than any one individual model. No single model is best in every aspect. Comparatively, the ECHAM4/OPYC3 and HadCM3 are most excellent, and the CGCM2 and NCAR-PCM are poorest among the models. Overall, all of the models overestimate the annual and seasonal precipitation amount in East Asia. The models' ability to reproduce precipitation climatology is obviously weaker relative to surface air

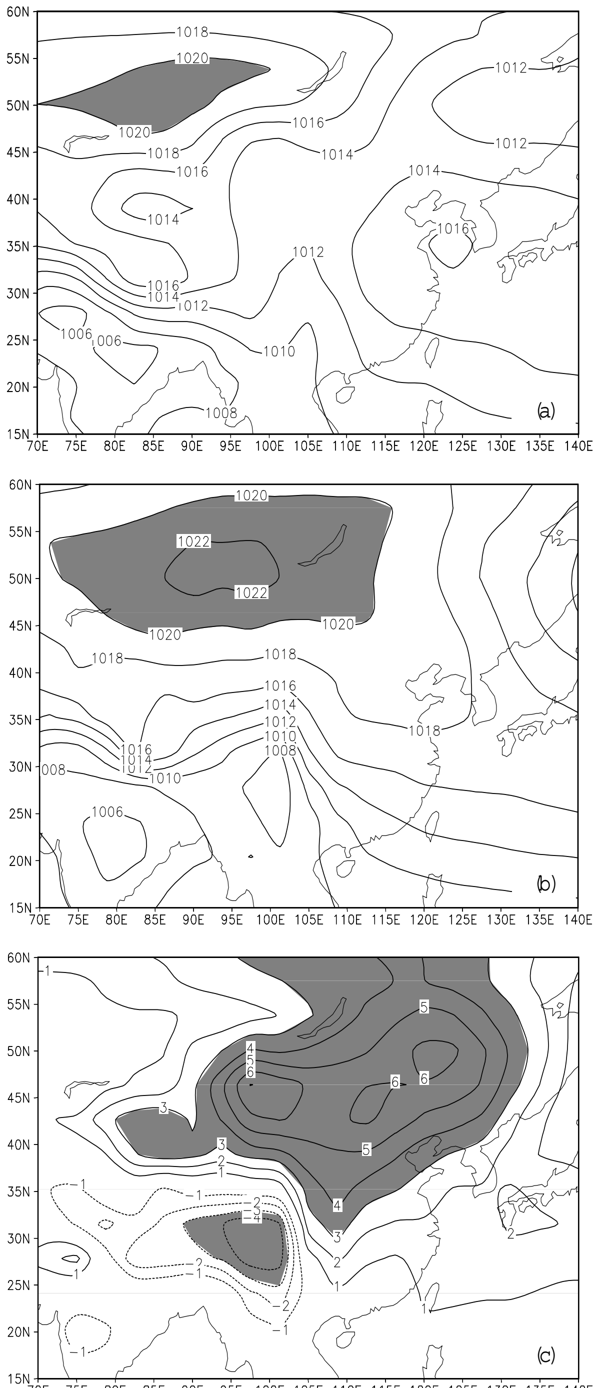

Fig. 13. Same as Fig. 11, but for spring.

temperature.

4.3 *Sea level pressure*

Sea level pressure is a key climate element to describe the geographical distribution of atmospheric mass and to partly reflect the atmospheric general circulation behavior. As displayed in Fig. 11, the simulated spatial distribution of annual mean sea level pressure is in accord with the reanalysis data on the whole, although sea level pressure is usually underestimated by 1–5 hPa over the Tibetan Plateau, and overestimated by 1–4 hPa in the rest of the East Asian continent. Furthermore, it is noted that the geographical distribution of spring and autumn sea level pressure is generally reproduced by the six-model ensemble mean (Figs. 13 and 15), with a similar spatial difference pattern to the annual mean being present.

As a widely recognized monsoon region, the difference of thermal capacity between the ocean and continent leads to a huge longitudinal and latitudinal heat forcing contrast in East Asia both in winter and summer (Tao and Chen, 1987), which can be clearly reflected by the seasonal sea level pressure pattern. Therefore, reliable simulation of winter and summer sea level pressure bears directly on whether climate models can accurately reproduce East Asian monsoon general circulation. As illustrated in Figs. 12 and 14, the six-model ensemble mean is in good accordance with the reanalysis data both in winter and summer on a large scale. For instance, both the systematic continental cold high pressure centered over Mongolia in winter and warm low pressure situated over the Tibetan Plateau in summer are successfully reproduced by the ensemble mean. However, simulation biases are also pronounced. In winter (Fig. 12), the simulated center of cold high pressure over Mongolia is displayed southward by about 4° relative to the reanalysis data. In addition, Fig. 12c shows that there exists a boundary of 46°N, south (north) of which the simulated sea level pressure is 1–6 (0–2) hPa above (below) the reanalysis data in East Asia. Except for northern Xinjiang Province, the winter sea level pressure is usually overestimated by all six of the models in East Asia. In summer (Fig. 14), the ensemble mean fails to capture the observational low pressure system over Mongolia, with the minimum at 1002 hPa. At the same time, the observational intensity of the low pressure system over the Tibetan Plateau is overestimated by 1–14 hPa, and its position is simulated to shift northward by an average of 4° or so. In addition, the subtropical western Pacific High is simulated to be weaker than the reanalysis data, at least implying a slightly decreased East Asian summer monsoon.

The values included in Table 4 indicate that the models can successfully capture the geographical distribution pattern of the East Asian annual and seasonal sea level pressure climatology, except for summer. In winter, the RAE of the six-model ensemble mean is only 1.62 hPa, and the SCC and RSD are respectively as high as 0.95 and 1.01, at the same time showing a small value for RMSE and RMSE2. All of the above denotes that the ensemble mean can reliably reproduce

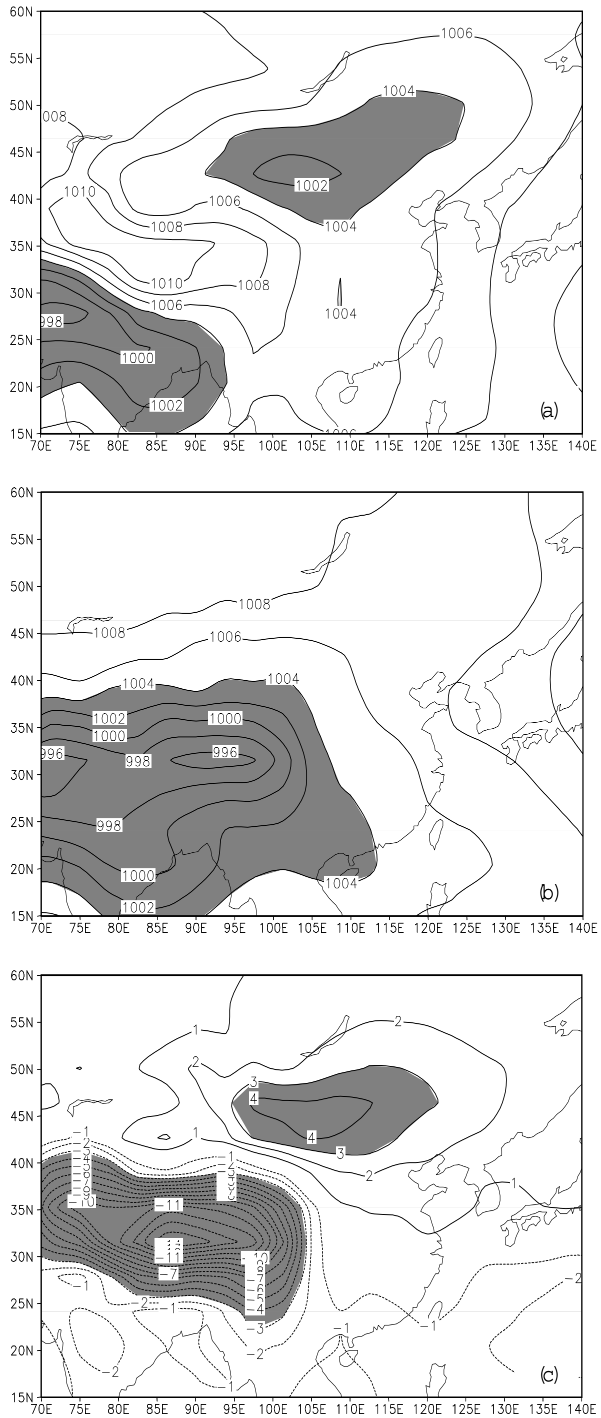

Fig. 14. Same as Fig. 11, but for summer.

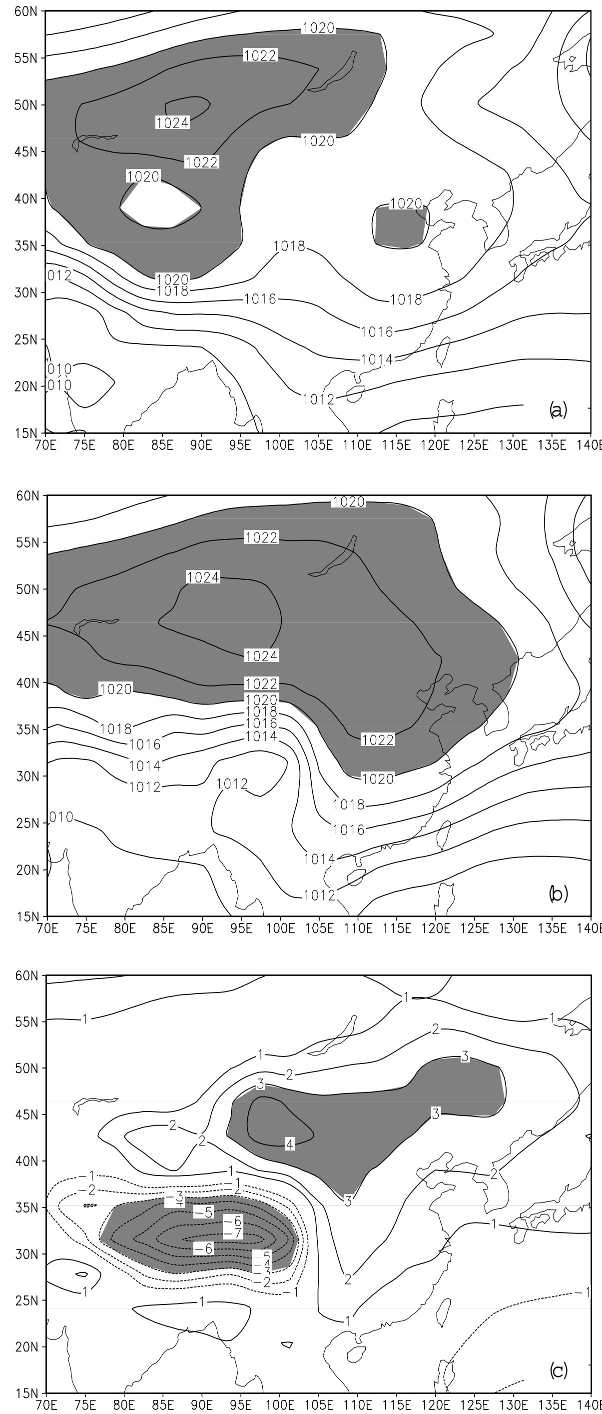

Fig. 15. Same as Fig. 11, but for autumn.

the geographical distribution pattern and spatial variational magnitude of East Asian winter sea level pressure. The ECHAM4/OPYC3 presents a much better performance than the ensemble mean, and the CCSR/NIES and CGCM2 are relatively worst among the models.

In summer, the model reproducibility of sea level pressure is significantly weaker than in winter. Although the RAE is relatively smaller compared with winter, the decrease of SCC is quite notable. The models' SCC varies from 0.15 (NCAR-PCM) to 0.45 (ECHAM4/OPYC3 and CSIRO-Mk2), being only 0.39

Table 4. The values of six statistical variables for sea level pressure in East Asia.

	RA (hPa)	RAE (hPa)	SCC	RSD	RMSE (hPa)	RMSE2 (hPa)
Annual mean (observational RA equals 1014.81 hPa)						
CCSR/NIES	1015.65	0.84	0.87	1.10	2.12	1.94
CGCM2	1016.33	1.52	0.80	1.40	3.40	3.04
CSIRO-Mk2	1017.42	2.61	0.89	1.28	3.38	2.15
ECHAM4/OPYC3	1016.03	1.22	0.92	1.26	2.21	1.84
HadCM3	1012.30	−2.51	0.80	1.35	3.82	2.88
NCAR-PCM	1015.78	0.97	0.78	1.33	3.11	2.96
Ensemble mean	1015.58	0.77	0.91	1.19	1.98	1.82
Winter (observational RA equals 1022.75 hPa)						
CCSR/NIES	1026.20	3.45	0.68	1.15	6.38	5.36
CGCM2	1025.85	3.10	0.77	1.24	5.73	4.82
CSIRO-Mk2	1025.73	2.98	0.93	1.19	4.09	2.80
ECHAM4/OPYC3	1023.50	0.75	0.97	1.18	2.13	1.99
HadCM3	1020.31	−2.44	0.90	0.93	3.64	2.71
NCAR-PCM	1024.64	1.89	0.89	1.01	3.50	2.94
Ensemble mean	1024.37	1.62	0.95	1.01	2.60	2.03
Spring (observational RA equals 1013.72 hPa)						
CCSR/NIES	1015.08	1.36	0.75	1.34	3.44	3.16
CGCM2	1018.30	4.58	0.74	1.58	5.98	3.85
CSIRO-Mk2	1017.70	3.98	0.80	1.40	4.99	3.02
ECHAM4/OPYC3	1015.29	1.57	0.88	1.34	2.82	2.34
HadCM3	1011.57	−2.15	0.84	1.53	3.80	3.13
NCAR-PCM	1014.06	0.34	0.74	1.37	3.32	3.30
Ensemble mean	1015.33	1.61	0.87	1.29	2.80	2.29
Summer (observational RA equals 1006.11 hPa)						
CCSR/NIES	1004.19	−1.92	0.40	1.48	4.18	3.72
CGCM2	1004.48	−1.63	0.35	1.46	4.14	3.81
CSIRO-Mk2	1007.26	1.15	0.45	1.24	3.34	3.13
ECHAM4/OPYC3	1007.63	1.52	0.45	1.12	3.31	2.94
HadCM3	1002.80	−3.31	0.44	1.83	5.47	4.36
NCAR-PCM	1005.86	−0.25	0.15	1.87	5.23	5.22
Ensemble mean	1005.37	−0.74	0.39	1.38	3.63	3.55
Autumn (observational RA equals 1016.67 hPa)						
CCSR/NIES	1017.13	0.46	0.91	1.15	2.10	2.05
CGCM2	1016.68	0.01	0.79	1.36	3.52	3.52
CSIRO-Mk2	1019.03	2.36	0.91	1.21	3.17	2.11
ECHAM4/OPYC3	1017.65	0.98	0.96	1.07	1.64	1.31
HadCM3	1014.52	−2.15	0.83	1.25	3.64	2.94
NCAR-PCM	1018.55	1.88	0.81	1.33	3.78	3.28
Ensemble mean	1017.26	0.59	0.91	1.17	2.13	2.04

for the ensemble mean. Therefore, large disagreements exist between the simulation and reanalysis data. The RSD varies from 1.12 (ECHAM4/OPYC3) to 1.87 (NCAR-PCM), denoting that the spatial variational magnitude of sea level pressure is uniformly overestimated by all six of the models. In general, the ECHAM4/OPYC3 and CSIRO-Mk2 exhibit good performance, better than the ensemble mean, whereas the NCAR-PCM is highly unreliable. Based on the above analysis with respect to sea level pressure, one must be careful when trying to use the above model outputs to address and even forecast the East Asian summer atmospheric circulation change related to human activities in the domain addressed in the present study.

5. Conclusion

East Asian annual and seasonal surface air temperature, precipitation and sea level pressure climatology as simulated by seven fully coupled atmosphere-ocean models are systematically evaluated on the basis of observation and reanalysis data. The primary conclusions are as follows.

(1) The models can reliably reproduce the annual and seasonal surface air temperature climatology in East Asia. In general, relatively better (poorer) performance is exhibited for autumn and annual mean (summer and spring). Comparatively, the HadCM3 and ECHAM4/OPYC3 abilities to simulate surface air temperature are best among the models and generally much more reliable than the seven-model ensemble mean, whereas the CGCM2 produces relatively larger simulation errors.

(2) The models can reliably reproduce the geographical distribution pattern of the East Asian annual and seasonal precipitation climatology, although all of the models overestimate the annual and seasonal precipitation amount in East Asia. The models' ability to simulate the East Asian seasonal precipitation climatology varies with season, with the best performance in autumn. Additionally, it is revealed that the annual and seasonal precipitation is reproduced best by the seven-model ensemble mean than any single model. Comparatively, the ECHAM4/OPYC3 and HadCM3 are most excellent, whereas the CGCM2 and NCAR-PCM are poorest among the models. The models' ability to reproduce precipitation is obviously weaker than surface air temperature.

(3) The models can capture the main spatial distribution pattern of annual, winter, spring and autumn sea level pressure. In contrast, large simulation errors are present for the summer sea level pressure.

(4) No single model is best in every aspect. The multi-model ensemble mean always exhibits a high reliability.

In Xu et al. (2002), it is found that SCC is 0.68 for annual precipitation based on five CGCMs' outputs, in response to 0.86 in this study. In addition, comparing the results presented here with those of Zhao et al. (1995, 1998) also shows that the CGCMs' simulation abilities to reproduce East Asian climatology has been improved largely in the past several years, and its reproducibility has been generally reliable up to now, although there still exist uncertainties to a certain degree, such as the relatively large simulation errors for summer sea level pressure.

The complex physically-based climate model is an important tool to explore the climate and climatic changes. Simulation errors are inevitable when models are used to simulate the complex real climate system. It is important to learn that we need to have a basic grasp on their ability and the errors in a given region, such as East Asia in this study. Taking into account the results mentioned above, the East Asian climate changes as simulated by the above seven models under different atmospheric greenhouse gas and aerosol emission scenarios are worth considering in more detail.

Acknowledgments. The authors sincerely thank the IPCC Data Distribution Centre and the Hadley Centre for Climate Prediction and Research for providing model outputs. Thanks are also due to the anonymous reviewers for their valuable comments. This research was jointly supported by the Chinese Academy of Sciences (CAS) under Grant No. KZCX3-SW-221, by the National Natural Science Foundation of China under Grant No. 40405015, by the Chinese Ministry of Science and Technology under Grant No. 2001BA611B (part 1), and by the CAS "Hundred Talent Project" funding awarded to Gao Yongqi.

REFERENCES

Bueh Cholaw, U. Cubasch, Lin Yonghui, and Ji Liren, 2003: The change of North China climate in transient simulations using the IPCC SRES A2 and B2 scenarios with a coupled atmosphere-ocean general circulation model. *Adv. Atmos. Sci.*, **20**, 755–766.

Emori, S., T. Nozawa, A. Abe-Ouchi, A. Numaguti, M. Kimoto, and T. Nakajima, 1999: Coupled ocean atmosphere model experiments of future climate change with an explicit representation of sulphate aerosol scattering. *J. Meteor. Soc. Japan*, **77**, 1299–1307.

Flato, G. M., and G. J. Boer, 2001: Warming asymmetry in climate change simulations. *Geophys. Res. Lett.*, **28**, 195–198.

Fu Congbin, Wei Helin, Chen Ming, Su Bingkai, Zhao Ming, and Zheng Weizhong, 1998: Simulation of the evolution of summer monsoon rainbelts over eastern China from regional climate model. *Chinese J. Atmos. Sci.*, **22**, 522-534. (in Chinese)

Gordon, C., C. Cooper, C. A. Senior, H. T. Banks, J. M. Gregory, T. C. Johns, J. F. B. Mitchell, and R. A. Wood, 2000: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dyn.*, **16**, 147–168.

Gordon, H. B., and S. P. O′Farrell, 1997: Transient climate change in the CSIRO coupled model with dynamic sea ice. *Mon. Wea. Rev.*, **125**, 875–907.

Hu, Z. Z., S. Yang, and R. Wu, 2003: Long-term climate variations in China and global warming signals. *J. Geophys. Res.*, **108** (D19), 4614, doi: 10.1029/2003JD003651.

Hu, Z. Z., L. Bengtsson, and K. Arpe, 2000: Impact of global warming on the Asian winter monsoon in a coupled GCM. *J. Geophys. Res.*, **105** (D4), 4607–4624.

Hulme, M., Zhao Zhongci, and T. Jiang, 1994: Recent and future climate change in East Asia. *International Journal of Climatology*, **14**, 637–658.

IPCC, 2001: *Climate Change 2001: The Scientific Basis*. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, Houghton et al., eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 881pp.

Jiang Dabang, Wang Huijun, and Lang Xianmei, 2004a: East Asian climate change trend under global warming background. *Chinese Journal of Geophysics*, **47**, 675–681.

Jiang Dabang, Wang Huijun, and Lang Xianmei, 2004b: Multimodel ensemble prediction for climate change trend of China under SRES A2 scenario. *Chinese Journal of Geophysics*, **47**, 878–886.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–472.

Kang, I. -S., and Coauthors, 2002: Intercomparison of the climatological variation of Asian summer monsoon precipitation simulated by 10 GCMs. *Climate Dyn.*, **19**, 383–395.

Knutson, T. R., T. L. Delworh, K. W. Dixon, and R. J. Stouffer, 1999: Model assessment of regional surface temperature trends (1949–1997). *J. Geophys. Res.*, **104**, 30981–30996.

Nakićenović, N., and Coauthors, 2000: *IPCC Special Report on Emissions Scenarios*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 599pp.

New, M., M. Hulme, and P. D. Jones, 1999: Representing twentieth century space-time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology. *J. Climate*, **12**, 829–856.

Roeckner, E., J. M. Oberhuber, A. Bacher, M. Christoph, and I. Kirchner, 1996: ENSO variability and atmospheric response in a global coupled atmosphere-ocean GCM. *Climate Dyn.*, **12**, 737–754.

Tao Shiyan and Chen Longxun, 1987: A review of recent research on the East Asian monsoon in China. *Monsoon Meteorology*. Oxford University Press, 60–92.

Washington, W. M., and Coauthors, 2000: Parallel climate model (PCM) control and transient simulations. *Climate Dyn.*, **16**, 755–774.

Xu Ying, Ding Yihui, and Zhao Zongci, 2002: Detection and evaluation of effect of human activities on climatic change in East Asia in recent 30 years. *Journal of Applied Meteorological Science*, **13**, 513–525. (in Chinese)

Xu Ying, Ding Yihui, Zhao Zongci, and Zhang Jin, 2003a: A scenario of seasonal climate change of the 21st century in Northwest China. *Climatic and Environmental Research*, **8**, 19–25. (in Chinese)

Xu Yinlong, Xue Feng, and Lin Yihua, 2003b: Changes of surface air temperature and precipitation in China during the 21st century simulated by HadCM2 under different greenhouse gas emission scenarios. *Climatic and Environmental Research*, **8**, 209–217. (in Chinese)

Zeng Qingcun, Zhang Xuehong, and Yuan Chongguang, 1989: Conception, methodology, and actualities of climate model. *Advances in Earth Science*, **4**, 1–26. (in Chinese)

Zhao Zongci, and Luo Yong, 1998: Advance on investigations of regional climate modelling since 1990. *Acta Meteorologica Sinica*, **56**, 225–246. (in Chinese)

Zhao Zongci, Ding Yihui, Li Xiaodong, and Wang Shaowu, 1995: Evaluation of CGCM climate simulation in East Asian region. *Journal of Applied Meteorological Science*, **6**(suppl.), 9–18. (in Chinese)

Zhao Zongci, Ding Yihui, Xu Ying, and Zhang Jin, 2003: Detection and prediction of climate change for the 20th and 21st century due to human activity in Northwest China. *Climatic and Environmental Research*, **8**, 26–34. (in Chinese)