ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 26, NO. 4, 2009, 736–747

A Review on Aspects of Climate Simulation Assessment

WANG Bin*[1] (王 斌), XIE Xin[1,2] (谢 歆), and LI Lijuan[1] (李立娟)

[1]*State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics,*

Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029

[2]*Graduate University of Chinese Academy of Sciences, Beijing* 100049

(Received 16 February 2009; revised 25 March 2009)

ABSTRACT

This paper reviews some aspects of evaluation of climate simulation, including the ITCZ, the surface air temperature (SAT), and the monsoon. A brief introduction of some recently proposed approaches in weather forecast verification is followed by a discussion on their possible application to evaluation of climate simulation. The authors suggest five strategies to extend the forecast verification methods to climate simulation evaluation regardless significant differences between the forecasts and climate simulations. It is argued that resolution, convection scheme, stratocumulus cloud cover, among other processes in the atmospheric general circulation model (AGCM) and the ocean-atmosphere feedback are the potential causes for the double ITCZ problem in coupled models and AGCM simulations, based on the system- and component-level evaluations as well as the downscaling strategies in some recent research. Evaluations of simulated SAT and monsoons suggest that both coupled models and AGCMs show good performance in representing the SAT evolution and its variability over the past century in terms of correlation and wavelet analysis but poor at reproducing rainfall, and in addition, the AGCM alone is not suitable for monsoon regions due to the lack of air-sea interactions.

Key words: climate simulation evaluation, forecast verification, ITCZ, surface air temperature, monsoon

Citation: Wang, B., X. Xie, and L. J. Li, 2009: A review on aspects of climate simulation assessment. *Adv. Atmos. Sci.*, **26**(4), 736–747, doi: 10.1007/s00376-009-9038-y.

1. Introduction

Climate system models (CSMs) are important scientific tools for modeling, understanding, and predicting complex behaviors and processes of the climate system. They numerically solve a series of mathematical-physical equations (including dynamic equations and parameterized schemes) to quantify the multi-sphere interactions as well as the dynamical and physical processes in the climate system. The equations are expressed in terms of codes that are run on powerful computers. CSMs have demonstrated significant and increasing skill in representing many important mean climate features, such as large-scale distributions of atmospheric temperature, precipitation, radiation, and wind, and of oceanic temperatures, currents, and sea-ice cover (Randall et al., 2007, p. 600). They have also simulated essential aspects of many of the patterns of climate variability observed across a range of time scales, e.g., the advance and retreat of the major monsoon systems, seasonal shifts of temperatures, storm tracks, and rain belts, and the hemispheric-scale seesawing of extratropical surface pressures (the Northern and Southern "Annular Modes") (Randall et al., 2007, p. 600). Some CSMs have been used to predict weather and make seasonal forecasts. Therefore, CSMs have the essential physical processes for modeling future climate change.

Discrepancies between CSM simulations and observations, or CSM simulation errors, however, remain significant. Even worse, the nature and causes of these errors are poorly understood, which must be accounted for in a systematic fashion in order to confidently use CSMs for simulation of putative global climate change. In order to improve our understanding on the nature and causes of CSM simulation errors, innovative approaches to evaluate CSM simulations are needed, as increasingly more complex models are being devel-

---

*Corresponding author: WANG Bin, wab@lasg.iap.ac.cn

oped. Up to now, simulation evaluation has been given more and more attention by the international climate modeling community and has become an essential part of CSM development. For example, the fourth assessment report (AR4) of the Intergovernmental Panel on Climate Change (IPCC) devoted a whole chapter to specially introduce the concept and method of model evaluation and to show the results from the multiple model comparison (Randall et al., 2007, 590–662). In this chapter, a summary of diverse evaluations on the results simulated by CSMs was given in terms of contemporary climate, large-scale climate variability, extreme events, climate sensitivity, and feedbacks.

Of the many methods and steps for evaluation, verification is a preliminary one, which provides further evaluation with analytical bases and methods and is an indispensable part of weather forecasting and climate modeling activities. Over the past decades, approaches for weather forecast verification have been developed from the "traditional verification" to the more-recently-developed verification methodology. The former mainly focuses on the calculation of one or more verification scores over a forecast-observation dataset, while the latter accounts for spatial structures and features characterizing weather maps or concentrates on evaluation of forecast probability distributions and further investigation into the properties of traditional verification measures for probability forecasts (Casati et al., 2008). Because of the importance of verification for NWPs, the World Weather Research Programme (WWRP) and the Working Group on Numerical Experimentation (WGNE) have built the Joint Working Group on Verification (JWGV). It has been maintaining a homepage (http://www.bom.gov.au/bmrc/wefor/staff/eee/verif/verif_web_page.html) to collect and introduce existing verification methods and to keep track of the most recently achieved research on verification in the world. This homepage provides the scientific community with a useful reference to advances in verification methods.

There are significant differences between weather forecast verification and climate simulation verification. Generally, weather forecast verification can be performed one day or several days later. However, it is impossible to verify decade predictions and centurial projections any time soon. Due to the lack of long-term observational data, climate simulation verification is also limited by a time scale of decades or shorter. No reliable observations can be obtained to verify model representation of a past climate more than a century ago. Clearly, climate simulation verification is much more difficult than weather forecast verification. Despite the many new approaches for weather forecast verification mentioned above, they can barely meet the need of climate simulation evaluation and CSM improvements. Statistical results and scores produced in the verification processes may enhance our understanding of model performance to some extent and may even provide us with directions of CSM improvements, but it is very difficult for us to design schemes for model improvements just relying on these results. Therefore, further evaluation of climate simulation is necessary so that effective approaches to CSM improvements can be finally proposed through establishing modules or parameterization schemes that can reasonably present actual mechanisms of physical phenomena based on plenty of numerical experiments and our understanding of the real climate system.

Further simulation evaluation concerns complex multi-sphere interactions of the climate system, and thus involves many analysis methods and cut-in points. It mainly falls into two classes: system- and component-level evaluations. The system-level evaluation of a CSM includes running a full model and comparing its simulation with the corresponding analysis or observation to reveal problems in the model. The component-level evaluation assesses a particular component of a model by testing it independently of the complete model to find the cause of problems hidden by the model's complexity.

At the end of the 1980s, climate simulation evaluation was recognized by the entire international climate modeling community. For example, in 1989, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) was established at the Lawrence Livermore National Laboratory (LLNL) in Livermore, California, USA, whose mission is to develop improved methods and tools for the diagnosis and intercomparison of global climate models. In the past 20 years, five projects of model intercomparison have been set up by the PCMDI, including the Atmospheric Model Intercomparison Project (AMIP), the Coupled Model Intercomparison Project (CMIP), the Seasonal Prediction Model Intercomparison Project (SMIP), the Aqua-Planet Experiment Project (APE), and the Paleoclimate Modeling Intercomparison Project (PMIP). The Climate Change Prediction Programs (CCPP)-Atmospheric Radiation Measurement (ARM) Parameterization Testbed (CAPT) was also established by the PCMDI to evaluate parameterizations of sub-grid-scale processes in global climate models.

In as early as the 1980s, Chinese meteorologists had also recognized the significance of climate simulation evaluation. For example, Qian and Wang (1984) evaluated the simulation of diurnal variations of meteorological fields in the summer by a five-layer atmospheric general circulation model (AGCM) developed in China. Over the past 20 years, scientists at the State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics (IAP), Chinese Academy of Sciences, successively developed AGCMs (e.g., Zeng et al., 1989; Wang et al., 2004a; Zhou et al., 2005; Wu et al., 2008), ocean general circulation models (OGCMs) (e.g., Zhang and Liang, 1989), coupled ocean-atmosphere models (e.g., Zhang et al., 1992), and CSMs (e.g., Zhang et al., 2000; Yu et al., 2002; Zhou et al., 2005; Yu et al., 2008; Zhou et al., 2008a). Following the development of climate models at the LASG/IAP, climate simulation evaluation has received more attention. For example, Guo et al. (1996) evaluated the simulation of the large-scale circulation patterns of the mean climate state by a CSM. There are other recent publications on climate simulation evaluation (e.g., Zhou and Yu, 2006; Zhang et al., 2007; Zhou et al., 2007; Li et al., 2007b,c; Duan et al., 2008; Zhou et al., 2008a). Clearly, Chinese meteorologists have gradually become more active in the field of climate simulation evaluation.

In order to describe the development of model evaluation and to track its most recent progress, a brief introduction of the main methodology of forecast verification and some suggestions on how to use forecast verification methods in climate simulation verification are presented in section 2. Some recent progresses in the evaluations of several important events or meteorological fields, namely the ITCZ, surface air temperature (SAT), and monsoon, are reviewed in sections 3, 4, and 5, respectively. Finally, the summary and discussions are included in section 6.

2. Verification of climate simulation

Many approaches for weather forecast verification have been developed. These approaches may also be useful for climate simulation evaluation although they were not designed for this purpose. Here, we first briefly introduce two representative forecast verification approaches, namely spatial verification approaches and probabilistic forecasts and ensemble verification approaches, and then discuss the strategy on using these approaches in climate simulation verification.

Spatial verification approaches specifically compare forecasts with corresponding analysis or observational data over spatial domains, account for the spatial nature of forecast fields, and aim to provide feedback on the physical nature of the forecast error (Casati et al., 2008). They overcome the difficulty of standard verification methods, which are based on a point-to-point comparison between forecasts and analysis, to interpret in meaningful physical terms (i.e., mean square error, MSE). Some new approaches to spatial verification, e.g., error-decomposition (Hoffman et al., 1995; Nehrkorn et al., 2003), feature-based (e.g., Ebert and McBride, 2000; Davis et al., 2006a,b), scale-decomposition (e.g., Briggs and Levine , 1997; Casati and Wilson, 2007) and neighborhood-based or fuzzy (e.g., Tremblay et al., 1996; Ebert, 2008) approaches, have been developed in the last decade. However, most of these approaches need observations defined continuously over a spatial domain and do not allow missing values in the observations (Casati et al., 2008). This limitation needs to be addressed in future research.

Probabilistic forecasts and ensemble verification approaches are designed for verification of probability forecasts, which are different from the spatial verification methods for numerical weather forecasts. Following the advent of ensemble forecasts in the early 1990s, the existing verification methods for probability forecasts have been re-evaluated and new methods have been developed. These methods are generally classified into three types: verifying the distribution of an ensemble as the sample from a probability distribution function (PDF) (e.g., Anderson, 1996; Hersbach, 2000; Smith, 2001; Weisheimer et al., 2004), evaluating the PDF of a generic probability forecast (Good, 1952; Roulston and Smith, 2002; Wilson et al., 1999), and assessing forecasts of the probability of an event (e.g., Brier, 1950; Murphy, 1973; Brocker and Smith, 2007).

The spatial verification approaches and the probabilistic forecasts and ensemble verification approaches could be used for climate simulation verification for the following reasons. Firstly, these methods could be used to evaluate forecasts from realistically initialized CSM runs in forecast mode to determine the CSM's initial drift from the NWP analysis and/or from the available field data and thereby to gain insight on deficiencies of model parameterizations. A vivid example of this strategy is the CAPT that was established by the PCMDI to evaluate parameterizations of sub-grid-scale processes in global climate models. Secondly, these methods could be similarly applied to verify climatologically averaged simulation fields by regarding them as forecast fields. For instance, the climatological mean of the precipitation simulation in the summer or winter could be compared with the climatological mean of precipitation analysis or observation in the same season by using the methods similar to those for forecast verification. Thirdly, the forecast verification methods could be extended to verify seasonal forecasts and annual predictions by the CSMs. Fourthly, they could also be used to assess leading modes decomposed from time series of climate simulations by regarding each mode as a forecast field. Finally, the probabilistic forecasts and ensemble verification approaches could be generalized to evaluate multi-model ensemble climate simulations. Besides the use of weather forecast verification approaches, development of new methods specifically designed for climate simulation verification is also needed.

3. Evaluation on simulations of the ITCZ

The ITCZ is a narrow belt near the equator, where the low-level air converges. It plays an important role in the atmospheric general circulation, hydrological cycle, and air-sea interaction, because of its impacts on the Hadley and Walker circulations, El Niño, the Madden-Julian Oscillation (MJO), and other events. In the ITCZ, the trade winds of both the Northern and Southern Hemispheres come together, and the air rises into the upward branch of the Hadley and Walker circulations and cools, releasing the accumulated moisture in an almost perpetual series of thunderstorms and the latent heat to drive the Hadley and Walker circulations. The location and intensity of the ITCZ affect the surface wind field, and thus the ITCZ is a critical factor in air-sea interactions and a core component of El Niño. The seasonal movement of the ITCZ is driven by annual variation in solar radiation, but could be affected by the seasonal variation of the MJO intensity. Along the ITCZ, tropical cyclogenesis occurs and the tropical cyclones travel before veering toward higher latitudes. Therefore, it is very important for CSMs to correctly or reasonably represent the ITCZ, say through its precipitation.

The current CSMs, however, have great difficulty in simulating the ITCZ precipitation correctly. One of the major tropical biases, common in climate modeling of the ITCZ by CSMs, is the "double ITCZ" noted by Mechoso et al. (1995). It includes a spurious ITCZ south of the equator in the eastern and central equatorial Pacific, in addition to the observed one north of the equator. This has been a bottle-neck problem for climate forecast and simulation. For example, a large part of the difficulty in forecasting El Niño is due to CSMs' failure to simulate the ITCZ correctly. In addition, difficulties in correctly simulating the seasonal variation of the MJO intensity by most CSMs or their atmospheric components are largely caused by the models' failure in correctly representing the seasonal movement of the ITCZ.

The double-ITCZ problem has persisted in the last several generations of CSMs (e.g., Mechoso et al., 1995; Latif et al., 2001; Davey et al., 2002; Li et al., 2004; Meehl et al., 2005), which still remains so in most of the current state-of-the-art CSMs, characterized by excessive precipitation over much of the tropics, including the Northern Hemisphere ITCZ, the South Pacific convergence zone (SPCZ), the Maritime Continent, and the equatorial Indian Ocean, and often is associated with insufficient precipitation over the equatorial Pacific (Lin, 2007). Previous studies suggest that this problem is mainly caused by the AGCM rather than by the OGCM (e.g., Schneider, 2002). For example, the 15-month ensemble hindcast by a CSM with two versions of atmospheric components (i.e., the University of California at Los Angeles-AGCM with different horizontal and vertical resolutions) indicated that the atmospheric component with a higher resolution helps the CSM to reduce the bias in its simulation of the ITCZ (Mechoso, 2006). An investigation by Zhang and Wang (2006) on the double ITCZ problem in the National Center for Atmospheric Research Community (NCAR) CSM Version 3 (CCSM3) showed that use of a modified Zhang-McFarlane convection scheme (Zhang and McFarlane, 1995; Zhang and Wang, 2006) significantly mitigates the double ITCZ problem in boreal summer, which reduces both the warm bias in the spurious ITCZ south of the equator and the cold bias in the cold tongue over the equator. However, "a synthetic view of the double-ITCZ problem is still elusive" as pointed out by Mechoso (2006). Therefore, evaluation of the ITCZ simulation is very important for our further understanding on the double ITCZ problem.

Simulation of the ITCZ is generally evaluated at a system-level or a component-level or at both through sensitivity experiments or multi-model simulations. A sensitivity-experiment-based combined system- and component-level evaluation suggests that the double ITCZ problem in CSMs results from a series of nonlocal and nonlinear adjustment processes in the coupled system, which can be traced to the uncoupled models, oceanic and atmospheric components (Li et al., 2004). In addition to the consensus that the double ITCZ problem is at least partly due to the underestimation of stratocumulus cloud cover by the AGCM that results in a warm bias of SST in the tropical southeastern Pacific, another possible reason is thought to be the overestimation of the east-west gradient of SST in the equatorial Pacific during the ocean model spin-up process (Li et al., 2004). A system-level evaluation of multi-model simulations also argues that an initial bias of surface precipitation and wind curl, i.e., the widely-known spurious eastward extension of the SPCZ associated with stronger surface southeasterly trade winds and biases of negative wind curls in stand-alone AGCMs, could amplify the double ITCZ symmetric structures in CSMs through a positive feedback mechanism (Zhang et al., 2007). This mechanism can be described through the following chain of interactions: precipitation (atmospheric latent heating), surface wind convergences, surface wind curls, Ekman pumping, South Equatorial Counter Current, and the eastward advection of ocean temperatures and SST (Zhang et al., 2007). It provides a possible method to address the longstanding double ITCZ problem in CSMs.

An analysis on the twentieth-century climate simulations of 22 IPCC AR4 CSMs together with 12 AMIP runs available shows that both CGCMs and AGCMs have an excessive tropical precipitation problem with half of CGCMs having insufficient precipitation in the equatorial Pacific (Fig. 1; Lin, 2007). It further indicates that it is an intrinsic error of the AGCMs causing the excessive tropical precipitation while the insufficient equatorial Pacific precipitation in the coupled runs of many models comes from ocean-atmosphere feedbacks (Lin, 2007). Three possible sources associated with the insufficient equatorial Pacific precipitation are revealed by Lin (2007): (1) excessive Bjerknes feedback caused by excessive sensitivity of precipitation to SST and overly strong time-mean surface wind speeds; (2) overly positive SST-latent heat flux feedbacks led by excessive sensitivity of surface air humidity to SST; and (3) insufficient SST-surface shortwave flux feedbacks as a result of insufficient sensitivity of cloud amount to precipitation. According to Lin (2007), the double-ITCZ problem in CSMs could be alleviated by reducing the excessive tropical precipitation and the aforementioned feedback-relevant errors in the AGCMs.

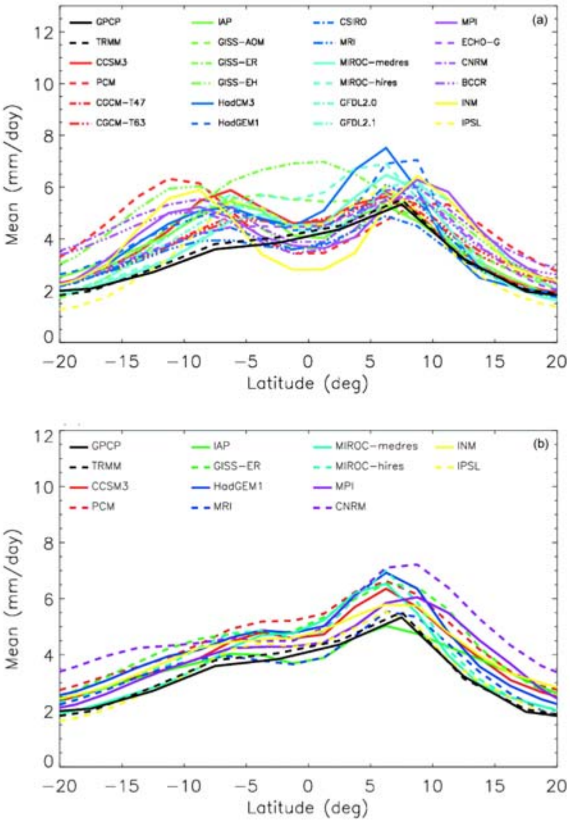

Fig. 1. (a) Meridional profiles of zonally averaged annual-mean precipitation from 22 IPCC AR4 CSMs and observations; (b) Same as (a) except for the AMIP runs of 12 models (Lin, 2007). Latitude axis ranges from −20(20°S) to 20(20°N).

The downscaling strategy could also be used in some key regions to address the double ITCZ problem. For example, based on the critical role played in the global circulation by the Maritime Continent where AGCMs tend to systematically underestimate the precipitation (Neale and Slingo, 2003), Lorenz and Jacob (2005) used a two-way nesting atmospheric model system that consists of a spectral AGCM and a grid-point regional atmospheric model in this region to investigate the influence of regional scale information on the global circulation. It is suggested that a more realistic parameterization of the convective processes is achieved in the two-way nested experiment with a 10-year integration due to the detailed representation of the complex land-sea distribution and topography within the region at the regional climate model scale (Lorenz and Jacob, 2005). It implies that the two-way nesting technique can reduce the precipitation simulation error in a stand-alone atmospheric model and thus may probably alleviate the double ITCZ problem in this model and the corresponding CSM, which needs further tests in coupled simulation experiments. Another example is the good reproduction of the northward-displaced ITCZ collocated with a zonal band of high SST by a regional ocean-atmosphere model (ROAM), which was established by coupling the version 2 of the Geophysical Fluid Dynamics Laboratory Modular Ocean Model (MOM2) with a full-physics regional atmospheric model on the domain of the eastern Pacific (Xie et al., 2007).

4. Evaluation on SAT simulation

SAT is a basic meteorological field. Its global average is one of the key variables to describe climate change, such as global warming in the past century. SAT is associated with heat exchange between the atmosphere and the ocean/land surface, and thereby plays a critical role in air-sea and air-land interactions. The importance of studying SAT is also due to its direct influence on the global economy and ecological environment.

SAT simulation evaluation mainly focuses on simulations of its evolution, variability, and regional features. The approaches to perform the evaluation mainly include correlation analysis, wavelet analysis, sensitivity analysis, among others. A correlation analysis on the SAT simulations over China and the globe in the twentieth century by 19 IPCC AR4 CSMs driven by historical natural and anthropogenic forcing shows that most models perform well in simulating both the global and the Northern Hemispheric mean SAT evolutions of the twentieth century, and acceptably in representing the SAT averaged over China, although the correlation between the observation and simulation over China (0.52) is lower than those over the globe (0.87) and over the Northern Hemisphere (0.82) (Zhou and Yu, 2006). The same analysis on the land SAT simulation over the globe in the twentieth century by the Grid-point Atmospheric Model of IAP LASG (GAMIL) Version 1.1.0 (simply GAMIL 1.1.0 hereinafter) (Li et al., 2007a) driven by historical natural and anthropogenic forcing reveals a better performance from the stand-alone AGCM, because of the higher correlation between the observations and its simulation (0.884) (Li et al., 2007b).

Based on a Morlet wavelet analysis, Zhou and Yu (2006) found the powers concentrating within the bands of inter-annual, decadal, or even longer inter-decadal time scales in the observed global mean SAT, while the powers in the ensemble mean simulation of nine IPCC AR4 CSMs are mainly concentrated within inter-decadal time scales, indicating poor performance of these CSMs at shorter time scales, e.g., inter-annual time scales. The same approach was used to evaluate land SAT simulation by GAMIL 1.1.0 following the standard coordinated experiment design of the Climate Variability and Predictability (CLIVAR) International Climate of the Twentieth Century Project (C20C), Phase II (Li et al., 2007b). The results showed that GAMIL 1.1.0 not only performed well in simulating decadal and inter-decadal variations of the land mean SAT but also successfully represented its inter-annual variability (IAV) (Li et al., 2007b). The evaluation on the SAT simulations from the coupled CSMs and the stand-alone AGCMs suggests that inclusion of natural and anthropogenic forcing can improve simulation of the SAT variations in inter-annual (for AGCM), decadal, and inter-decadal (for both AGCMs and CSMs) time scales (Zhou and Yu, 2006; Li et al., 2007b).

There are significant regional features in SAT; for example, under the background of global warming, a surface cooling trend from spring to summer has been observed in Central Eastern China (27°–36°N) in the past half century (e.g., Li et al., 1995). Can a CSM or an AGCM represent this cooling trend? If not, what are the possible factors that lead to their failures in simulating the trend? To answer these questions, careful investigations in the cooling trend and its simulations need to be carried out. Zhou and Yu (2006) examined the SAT simulations by 19 IPCC AR4 CSMs and found that very few of them could reproduce the summertime cooling (Fig. 2). The models that failed to represent this cooling trend include the Flexible Global Ocean-Atmosphere-Land System Model (FGOALS) with GAMIL 1.0 (the initial version of GAMIL) (Wang et al., 2004a) as the atmospheric component (FGOALS-g1.0) (Yu et al., 2008), which presented a warming trend in the same location. In order to find the reason of this bias in the SAT simulation by FGOALS-g1.0, we examined the same SAT simulation from the C20C experiment using its stand-alone atmospheric component, GAMIL 1.0. The results show that GAMIL1.0 presents a warming trend as FGOALS-g1.0 does, instead of a cooling trend (the corresponding figure is not shown here). It indicates that the bias in simulating the SAT over Central Eastern China by FGOALS-g1.0 is most likely due to the failure of GAMIL 1.0. In further experiments, however, GAMIL 1.1.0 successfully represented the surface cooling in the summer, although there is a northward shift in the simulated cooling (Fig. 3; Li et al., 2007c). The only difference between the two versions of GAMIL is their cumulus convective schemes: the Zhang-McFarlane scheme in GAMIL 1.0 and the improved Tiedtke scheme (Li et al., 2007a) in GAMIL 1.1.0. It implies the important role of the cumulus convective scheme in simulating the summertime cooling over Central Eastern China by the GAMIL. Further tests are required to investigate whether the FGOALS could reduce the bias in simulating the cooling trend through updating the cumulus convective scheme in its atmospheric component.

5. Evaluation on monsoon simulation

Monsoon is a prevailing surface wind, which changes direction after several months and alternates between a wet summer and a dry winter. It brings persistent and heavy rainfall to a region (i.e., monsoon region) during a particular season (e.g., summer), and thus defines one of the essential features of the Earth's climate. Because the monsoon plays an important role in the global atmospheric circulation and affects atmospheric variability of the surrounding areas, studies on its various aspects, e.g., its annual cycle, onset time, and intensity variability are of great scientific significance (Zhang and Li, 2007). Following the development of AGCMs and CSMs, numerical simulation has become more and more important for monsoon study. However, most of the current CSMs show large errors in the summer-monsoon simulation, although the

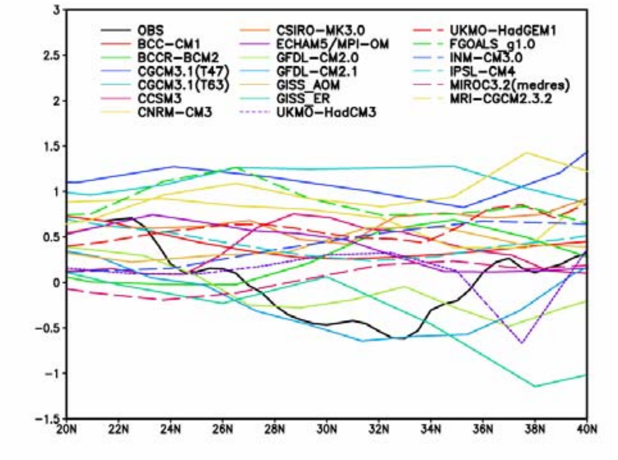

Fig. 2. Zonal mean of the JJA SAT linear trends during 1951–1999 over East China (102.5°–122.5°E) for IPCC AR4 runs by 19 different models. The observation is shown as the thick black line [Units: °C (50 yr)$^{-1}$] (Zhou and Yu, 2006).

annual and seasonal SAT, precipitation, and sea-level pressure climatology in the East Asian monsoon (EAM) region are successfully reproduced (Jiang et al., 2005). Therefore, the monsoon simulation evaluation is necessary to investigate the causes of models' deficiencies and related mechanisms.

Monsoon simulation evaluation is mainly concentrated on the simulations of monsoon variability, onset and retreat, and spatiotemporal evolution. Evaluation approaches include traditional ones, such as point error statistics and correlation analysis (e.g., Wang et al., 2004b), and most recently developed ones, e.g., the analyses based on the season-reliant empirical orthogonal function (S-EOF) (Wang et al., 2008a), directed angle (Zhang and Li, 2007), and wind onset and withdrawal (Li and Zhang, 2008). For example, a pattern correlation analysis on the ensemble simulations of the Asian-Australian monsoon (A-AM) anomalies by eleven AMIP AGCMs indicates that the models' simulations of anomalous Asian summer rainfall patterns in the A-AM region (30°S–30°N, 40°–160°E) are considerably poorer than those in the El Niño region (30°S–30°N, 160°E–80°W), mainly due to a lack of skill over Southeast Asia and the western North Pacific (5°–30°N, 80°–150°E), which is a striking characteristic of all the models (Wang et al., 2004b). Failing to represent correctly the negative correlations between the local summer rainfall and the SST anomalies over the Philippine Sea, the South China Sea, and the Bay of Bengal is the main reason of the model deficiencies (Wang et al., 2004b). The profound mechanisms of these deficiencies are related to the uncertainties in the models' physical parameterizations and the unreasonable experiment design of AMIP itself in which the atmosphere is forced by the prescribed SSTs without any air-sea interactions (ASIs) (Wang et al., 2004b). The importance of ASI in simulating the Asian summer monsoon (ASM) is also exhibited in a comparison between two AGCM runs with and without interactive coupled processes in the great warm pool (GWP) using the Spectral Atmospheric Model of IAP LASG (SAMIL) and an ocean mixed layer model (Duan et al., 2008). Their results show that inclusion of local ASI in the GWP leads to a substantial improvement in both the precipitation distribution and monsoon onset timing in the coupled simulation. The improvement benefits from a mechanism that ASIs modulate SST via wind-evaporation and cloud-radiation processes and

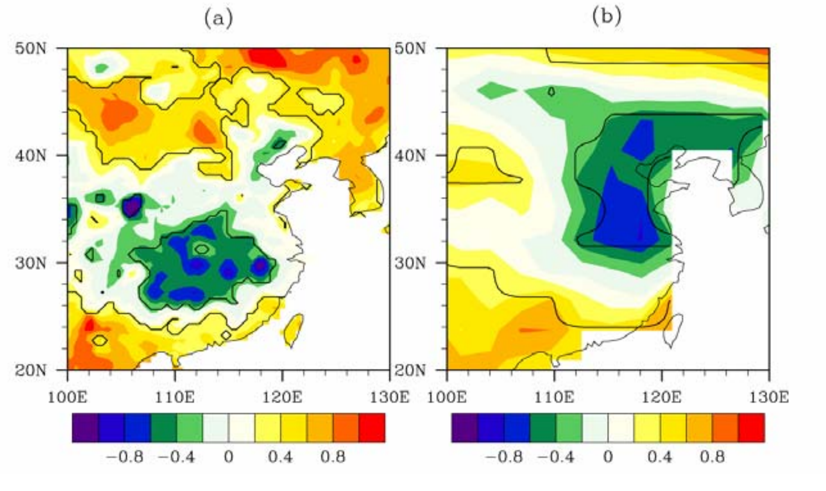

Fig. 3. The linear trends of summer (June–August) SAT from (a) CRUTS2.1 and (b) GAMIL1.1.0. The statistical time covers 1951–2000. The interval in the shaded regions is 0.2°C per 50 yrs. The solid lines are statistically significant at the 5% level using a student *t*-test (Li et al., 2007c).

in turn influence atmospheric circulations and precipitation patterns (Duan et al., 2008). ASIs also play a crucial role in representing the ENSO-monsoon interaction, the most dominant phenomenon in the climate system, according to the dramatic improvement of the simulation of the ENSO-monsoon relationship in the experiment with the interactive coupled ensemble (ICE) comparing with the results in the experiment without the ICE in which the impact of monsoon on the ENSO is absent (Kirtman and Shukla, 2002). The role of ASI is also implied in an evaluation on the simulation of the ENSO-Asian monsoon interaction by the HadCM3, a version of the Hadley Centre coupled ocean-atmosphere general circulation model, which well captures the main features of two anomalous anticyclones closely related to the developing and decaying phases of ENSO and with a crucial role in linking the Asian monsoon to ENSO (Li et al., 2007d), due to the full inclusion of the ASI processes in the HadCM3. Their results also indicated that variability in the ASM seems to be responsible for the appearance of the western North Pacific anticyclonic anomaly, which is crucial for the transition from an El Niño to a La Niña.

Another pattern correlation analysis on the priority of different components in the A-AM circulations in terms of reproducibility of a multi-model set of atmospheric simulations produced by 13 AGCMs, of which 12 models are from the C20C project participants, indicates that among the subsystems of the wide A-AM, both the South Asian monsoon and the Australian monsoon circulations are better reproduced according to Zhou et al. (2008a). The same study also showed that the western North Pacific monsoon circulation is reasonably represented with a slightly lower reproducibility because of its delayed response to the eastern tropical Pacific forcing. However, the EAM is poorly simulated by these AGCMs with the lowest reproducibility mainly due to the failure of specifying historical SST in capturing the zonal land-sea thermal contrast change across East Asia. Their evaluation also shows that the multi-model ensemble (MME) has the best performance in simulating all the A-AM circulation indices except for the Indian monsoon and the EAM circulation indices.

In order to obtain an integral view of the year-to-year variability across the entire A-AM system, Wang et al. (2008a) used the S-EOF to identify two major modes of variability in the observations for the period of 1956–2004, one with a prominent biennial tendency corresponding to the turnabout of ENSO, and the other leading ENSO by one year. These two major modes could be applied to evaluate the performances of both CSMs and AGCMs in simulating monsoon variability. For example, a comparison exhibits that the two leading modes from the coupled MME predictions, in terms of both spatial pattern and temporal evolution, are in general better than or at least comparable to those from the European Centre for Medium-Range Weather Forecasts (ECMWF) 40 year Re-analysis (ERA-40) and the National Centers for Environmental Prediction (NCEP)/Department of Energy Global Reanalysis 2 (NCEP-2) (Fig. 4), which again suggested that treating the atmosphere as a slave may be inherently unable to simulate summer monsoon rainfall variations in the heavily precipitating regions (Wang et al., 2008b). A similar conclusion on stand-alone AGCMs made by Zhou et al. (2008b) indicates that the SST-forced simulation of the seasonal rainfall anomalies from the MME of 11 AMIP AGCMs shows a skill comparable to NCEP-2 in reproducing the first two leading modes of variability, in terms of the spatial patterns and the corresponding temporal variations as well as their relationships with ENSO evolution. As for an individual AGCM simulation, Wu and Li (2008) evaluated the performance of GAMIL on the retrospective prediction of the A-AM IAV and its capability in capturing the two major observed modes of A-AM rainfall IAV for the period of 1979–2003 (Wu and Li, 2008). Their results show that the one-month lead prediction of the seasonal precipitation anomalies by GAMIL can primarily capture the major features of the two observed leading modes of the IAV, with the first mode better predicted than the second. The model is able to well describe the relationship between the first mode and ENSO, but has deficiencies in representing the relationship between the second mode and ENSO (Wu and Li, 2008). For global monsoon studies, Wang and Ding (2008) used a three-parameter metrics: the annual mean and two major modes of annual variation, namely, a solstitial mode and an equinoctial asymmetric mode, which together account for 84% of the annual variance, to represent the primary climatological features of the tropical precipitation and low-level circulation. The observed annual modes are used to verify the simulation of the global monsoon by FGOALS-s1.1 whose atmospheric component is SAMIL (Zhang et al., 2008a). Their verification also reveals that FGOALS-s1.1 reasonably captures the major features of the annual modes of tropical precipitation, e.g., the maximum centers of annual-mean rainfall well match the observations and an equatorial anti-symmetric structure of the monsoon mode close to the observations, although the model overestimates the rainfall intensity over the equatorial Pacific and the tropical South Pacific, underestimates the rainfall intensity over the northern equatorial Pacific, and fails to simulate the spring-fall asymmetric mode.

The directed angle, a new concept introduced by

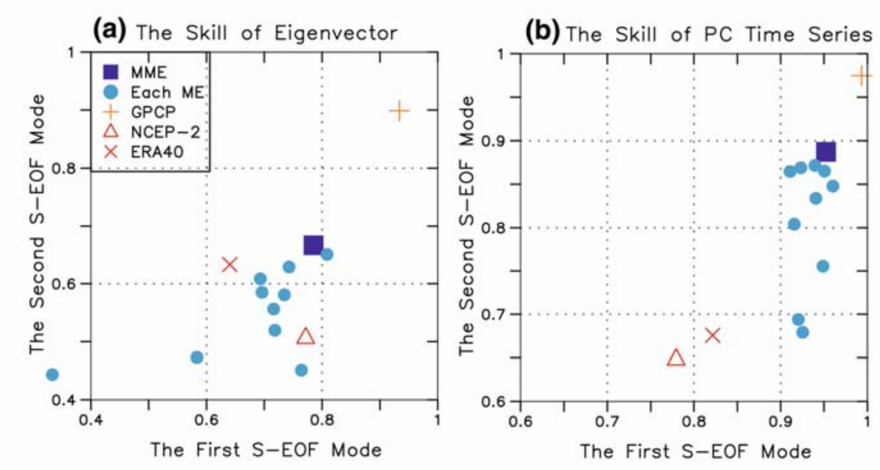

Fig. 4. Comparison of the performances of the MME and individual predictions, GPCP estimate, and two (NCEP-2 and ERA-40) reanalyses against the two observed (CMAP) dominant S-EOF modes of seasonal mean precipitation anomalies. The abscissa and ordinate represent, respectively, correlation coefficients between the observed and predicted (reanalyzed) anomalies for the first and second modes. The left panel is for the spatial correlation skill of the eigenvector, and the right panel is for the temporal correlation skill of the principal component (Wang et al., 2008b).

Zhang and Li (2007), with six categories of wind vector rotation in a seasonal cycle provides an effective measure for evaluating model performance in simulating the spatiotemporal evolution of the monsoon, because of different rotation styles for wind vectors in different monsoon subsystems in an annual cycle. This approach was employed to assess the monsoon simulations by eight AMIP AGCMs in the IPCC AR4, which reveals good reproducibility of the global wind-vector rotation regimes by most models, very little skill in simulating monsoon rotation styles by some models, especially in the South China Sea and West Africa, and poor model performance during the transitional season (Zhang and Li, 2007).

The wind onset and retreat are new concepts introduced by Li and Zhang (2008), which provides unique variables as well as convenience for evaluating simulations of monsoon onset and retreat. These concepts have been used in assessing monsoon onset and retreat simulations by seven AMIP AGCMs from the IPCC AR4 by Li and Zhang (2008), which indicates that (1) the multi-model ensemble mean simulations are generally better than any individual model results; (2) the wind retreat is better reproduced than the wind onset; (3) three of the AGCMs represent the wind onset well, namely GAMIL 1.0 (the atmospheric component of FGOALS-g1.0), MIROC3.2 (medres), and ECHAM5; and (4) most models can capture the behaviors of the wind retreat in the tropics. The study also pointed out the model discrepancies, including the failure in representing the dates of sudden change in the monsoon wind direction by a few models, and the poor reproducibility of the onset and retreat of both rainfall and OLR in most models.

A sensitivity experiment is also applied to evaluate the monsoon simulation. For example, Kitoh and Kusunoki (2008) revealed the impact of AGCMs' horizontal resolution on simulating IAV of the summer mean precipitation and seasonal march of the monsoon rain band, by comparing the outputs from a global 20-km mesh AGCM with those from a lower resolution (180-km mesh) model, both forced by the global SST during the period of 1979–1998. Their evaluation reveals that the higher-resolution model cannot only capture correctly orographic rainfall but also represent small-scale features such as Meiyu/Baiu rain bands and tropical cyclones than the lower-resolution model. Zhang et al. (2008b) suggest that the increase in model resolution can improve the simulation of the structure, seasonal evolution and IAV of the EASWJ that are closely related to the EAM, based on a comparison between the results from two versions of CSMs with different resolutions. The performance of a regional climate model (RCM) in simulating the EAM and precipitation over China during 1998 to 2002 is also assessed, and the results show that the RCM well reproduces the seasonal patterns of the mean circulation as well as the intensity and seasonal march of the EAM, although it has difficulty in simulating the persistent abnormal precipitation pattern over China (Li and Ding, 2005).

6. Summary and discussions

A review of climate simulation evaluation is presented with an emphasis on the discussion of methodologies with potential application in the assessment of climate simulation and the mechanical analysis on the ITCZ, SAT, and monsoon simulated by models. The use of two recently developed weather forecast verification measures, namely the spatial verification approaches and the probabilistic forecasts and ensemble verification approaches, in climate simulation verification are also discussed due to some similarities between these two verifications. The probable causes of the double ITCZ by coupled models, such as resolution, convection scheme, stratocumulus cloud cover among other processes in AGCMs and ocean-atmosphere feedbacks, are examined, using the system- and component-level evaluation as well as downscaling strategy. It is pointed out that both coupled and atmospheric models show acceptable skills in simulating the SAT evolution and variability of the last century through the correlation and wavelet analysis, but different performances in representing the regional features. Some traditional and newly proposed techniques (including the S-EOF and the directed angle) are used for evaluating the simulated monsoon variability, onset and retreat, and spatiotemporal evolution. The results indicate the inadequate use of AGCMs in monsoon regions, the poor reproducibility of rainfall, and the capability of simulating the wind retreat in the tropics.

We should keep in mind that CSM development and climate simulation evaluation supplement each other. Simulation evaluation improves our understanding of model deficiencies and their causes and mechanisms, which is the fundamental premise for improving the model's description of crucial processes in the climate system and even proposing new schemes for models. For this reason, more in-depth assessments on simulations of many other important events (e.g., diurnal cycle, MJO, ENSO, etc.), as well as examinations of simulated mechanisms, forcings, and feedbacks by individual CSM components should be conducted. On the other hand, more in-depth simulation evaluations not only depend on the development of new and better measures but also desire more accurate CSMs and more reliable simulations. Currently, they are limited by the imperfectness of CSMs, such as the conceptual problems in conventional model physics: artificial separation of processes and artificial separation of scales, in which different physical processes interacted mainly through the model's prognostic variables, missing most of the small-scale interactions between the processes (Arakawa, 2004).

Acknowledgements. The authors very much appreciate the valuable suggestions from Prof. Xuehong Zhang on the double ITCZ problem and the useful revision by Prof. Bin Wang from University of Hawaii. This paper is jointly supported by the National Basic Research Program of China (973 Program, Grant No. 2005CB321703) and the National Natural Science Foundation of China (NSFC) Innovation Research Group Fund (Grant No. 40821092).

REFERENCES

Arakawa, A., 2004: The cumulus parameterization problem: Past, present and future. *J. Climate*, **17**, 2493–2524.

Anderson, J. L., 1996: A method for producing and evaluating probabilistic forecasts from ensemble model integrations. *J. Climate*, **9**, 1518–1530.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1–3.

Briggs, W. M., and R. A. Levine , 1997: Wavelets and field forecast verification. *Mon. Wea. Rev.*, **125**, 1329–1341.

Brocker, J., and L. A. Smith, 2007: Increasing the reliability of reliability diagrams. *Wea. Forecasting*, **22**(3), 651–661.

Casati, B., and L. J. Wilson, 2007: A new spatial scale decomposition of the Brier score: Application to the verification of lightning probability forecasts. *Mon. Wea. Rev.*, **135**, 3052–3069.

Casati, B., and Coauthors, 2008: Review forecast verification: Current status and future directions. *Meteoroogical Application*, **15**, 3–18

Davey, M. K. and Coauthors, 2002: STOIC: A study of coupled model climatology and variability in tropical ocean regions. *Climate Dyn.*, **18**, 403–420.

Davis, C., B. Brown, and R. Bullock, 2006a: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134**, 1772–1784.

Davis, C., B. Brown, and R. Bullock, 2006b: Object-based verification of precipitation forecasts. Part II: application to convective rain systems. *Mon. Wea. Rev.*, **134**, 1785–1795.

Duan, A., C. Sui, and G. Wu, 2008: Simulation of local air-sea interaction in the great warm pool and its influence on Asian monsoon. *J. Geophys. Res.*, **113**, D22105, doi: 10.1029/2008JD010520.

Ebert, E. E., 2008: Fuzzy verification of high resolution gridded forecasts: A review and proposed framework. *Meteorological Applications*, **15**, 53–66.

Ebert, E. E., and J. L. McBride, 2000: Verification of precipitation in weather systems: Determination of systematic errors. *J. Hydrol.*, **239**, 179–202.

Good, I. J., 1952: Rational decisions. *Journal of the Royal Statistical Society Series B-Methodological*, **14**, 107–114.

Grams, J. S., W. A. Gallus, S. E. Koch, L. S. Wharton, A. Loughe, and E. E. Ebert, 2006: The use of a modified Ebert-McBride technique to evaluate mesoscale model QPF as a function of convective system morphology during IHOP 2002. *Wea. Forecasting*, **21**, 288–306.

Guo, Y., Y. Yu, K. Chen, X. Jin, and X. Zhang, 1996: Mean climate state simulated by a coupled ocean-atmosphere general circulation model. *Theor. Appl. Climatol.*, **55**, 99–112.

Hersbach, H., 2000: Decomposition of the continuous rank probability score for ensemble prediction systems. *Wea. Forecasting*, **15**, 559–570.

Hoffman, R. N., Z. Liu, J. F. Louis, and C. Grassotti, 1995: Distortion representation of forecast errors. *Mon. Wea. Rev.*, **123**, 2758–2770.

Jiang, D., H. Wang, and X. Lang, 2005: Evaluation of East Asian climatology as simulated by seven coupled models. *Adv. Atmos. Sci.*, **22**, 479–495.

Kirtman, B. P., and J. Shukla, 2002: Interactive coupled ensemble: A new coupling strategy for CGCMs. *Geophys. Res. Lett.*, **29**, 1367, doi: 10.1029/2002GL014834.

Kitoh, A., and S. Kusunoki, 2008: East Asian summer monsoon simulation by a 20-km mesh AGCM. *Climate Dyn.*, **31**, 389–401.

Latif, M., and Coauthors, 2001: ENSIP: The El Niño simulation intercomparison project. *Climate Dyn.*, **18**, 255–276.

Li, J. L., X. H. Zhang, Y. Q. Yu, and F. S. Dai, 2004: Primary reasoning behind the double ITCZ phenomenon in a coupled ocean-atmosphere general circulation model. *Adv. Atmos. Sci.*, **21**(6), 857–867.

Li, J.-P., and L. Zhang, 2008: Wind onset and withdrawal of Asian summer monsoon and their simulated performance in AMIP models. *Climate Dyn.*, doi: 10.1007/s00382-008-0465-8.

Li, L. J., B. Wang, Y. Q. Wang, and H. Wan, 2007a: Improvements in climate simulation with modifications to the Tiedtke convective parameterization in the Grid-point Atmospheric Model of IAP LASG (GAMIL). *Adv. Atmos. Sci.*, **24**(2), 323–335, doi: 10.1007/s00376-007-0323-3.

Li, L. J., B. Wang, and T. J. Zhou, 2007b: Impacts of external forcing on the 20th century global warming. *Chinese Science Bulletin*, **52**(22), 3148–3154.

Li, L. J., B. Wang, and T. J. Zhou, 2007c: Contributions of natural and anthropogenic forcings to the summer cooling over eastern China: An AGCM study. *Geophys. Res. Lett.*, **34**, L18807, doi: 10.1029/2007GL030541.

Li, Q. P., and Y. H. Ding, 2005: Multi-year simulation of the east Asian monsoon and precipitation in China using a regional climate model and evaluation. *Acta Meteorologica Sinica*, **19**(3), 302–316.

Li, X., X. Zhou, W. Li, and L. Chen, 1995: The cooling of Sichuan province in recent 40 years and its probable mechanisms. *Acta Meteorologica Sinica*, **9**, 57–68.

Li, Y., R. Lu, and B. Dong, 2007d: The ENSO-Asian monsoon interaction in a coupled ocean-atmosphere GCM. *J. Climate*, **20**, 5164–5177.

Lin, J. L., 2007: The double-ITCZ problem in IPCC AR4 coupled GCMs: Ocean-atmosphere feedback analysis. *J. Climate*, **20**(18), 4497–4525.

Lorenz, P., and D. Jacob, 2005: Influence of regional scale information on the global circulation: A two-way nesting climate simulation. *Geophys. Res. Lett.*, **32**, L18706, doi: 10.1029/2005GL023351.

Mechoso, C. R., 2006: Modeling the south eastern Pacific climate: Progress and challenges. NCEP EMC seminar. [Available online from http://www.emc.ncep.noaa.gov/seminars/presentations/2006/Mechoso.NCEP.Jan.06.ppt.]

Mechoso, C. R., and Coauthors, 1995: The seasonal cycle over the tropical Pacific in general circulation model. *Mon. Wea. Rev.*, **123**, 2825–2838.

Meehl, G. A., C. Covey, B. McAvaney, M. Latif, and R. J. Stouffer, 2005: Overview of the Coupled Model Intercomparison Project. *Bull. Amer. Meteor. Soc.*, **86**, 89–93.

Murphy, A. H., 1973: A new vector partition of the probability score. *J. Appl. Meteor.*, **12**, 595–600.

Neale, R., and J. Slingo, 2003: The Maritime Continent and its role in the global climate: A GCM study. *J. Climate*, **16**, 834–848.

Nehrkorn, T., R. N. Hoffman, C. Grassotti, and JF Louis, 2003: Feature calibration and alignment to represent model forecast errors: Empirical regularization. *Quart. J. Roy. Meteor. Soc.*, **129**(587), 195–218.

Qian, Y. F., and Q. Wang, 1984: A numerical simulation of diurnal variations of meteorological fields in summer. *Adv. Atmos. Sci.*, **1**, 40–52.

Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 600.

Roulston, M. S., and L. A. Smith, 2002: Evaluating probabilistic forecasts using information theory. *Mon. Wea. Rev.*, **130**, 1653–1660.

Smith, L. A. 2001: Disentangling uncertainty and error: On the predictability of nonlinear systems. *Nonlinear Dynamics and Statistics*, A. I. Mees, Ed., Birkhauser, Boston, MA, 31–64.

Schneider, E. K., 2002: Understanding differences between the equatorial Pacific as simulated by two coupled GCMs. *J. Climate*, **15**, 449–469.

Tremblay, A., S. G. Cober, A. Glazer, G. Isaac, and J. Mailhot, 1996: An intercomparison of mesoscale forecasts of aircraft icing using SSM/I retrievals. *Wea. Forecasting*, **11**, 66–77.

Wang, B., and Q. Lee, 2008: Global monsoon: Dominant mode of annual variation in the tropics. *Dyn. Atmos. Oceans*, doi: 10.1016/j.dynatmoce.2007.05.002.

Wang, B., H. Wan, Z. Z. Ji, X. Zhang, R. C. Yu, Y. Q. Yu, and H. B. Liu, 2004a: Design of a new dynamical core for global atmospheric models based on some efficient numerical methods. *Science in China (A)*, **47**(Suppl), 4–21.

Wang, B., I. S. Kang, and J. Y. Lee, 2004b: Ensemble simulations of Asian-Australian monsoon variability by 11 AGCMs. *J. Climate*, **17**, 803–818.

Wang, B., J. Yang, and T. J. Zhou, 2008a: Strengthening relationship with ENSO since the late 1970s. *J. Climate*, **21**, 1771–1789.

Wang, B. and Coauthors, 2008b: How accurately do coupled climate models predict the leading modes of Asian-Australian monsoon interannual variability? *Climate Dyn.*, **30**, doi: 10.1007/s00382-007-0310-5.

Weisheimer, A., L. A. Smith, and K. Judd, 2004: A new view of forecast skill: Bounding boxes from the DEMETER ensemble seasonal forecasts. *Tellus*, **57**(3), 265–279.

Wilson, L. J, W. R. Burrows, and A. Lanzinger, 1999: A strategy for verification of weather element forecasts from an ensemble prediction system. *Mon. Wea. Rev.*, **127**, 956–970.

Wu, T. W., and Coauthors, 2008: The Beijing Climate Center atmospheric general circulation model: Description and its performance for the present-day climate. *Climate Dyn.*, doi: 10.1007/s00382-008-0487-2.

Wu, Z. W., and J. P. Li, 2008: Prediction of the Asian-Australian monsoon interannual variations with the Grid-Point Atmospheric Model of IAP LASG (GAMIL). *Adv. Atmos. Sci.*, **25**(3), 387–394, doi: 10.1007/s00376-008-0387-8.

Xie, S. P., and Coauthors, 2007: A regional ocean-atmosphere model for eastern Pacific climate: Toward reducing tropical biases. *J. Climate*, **20**(8), 1504–1522.

Yu, Y. Q., R. C. Yu, X. Zhang, H. L. Liu, C. Yuan, and S. Chen, 2002: A flexible global coupled climate model. *Adv. Atmos. Sci.*, **19**, 169–190.

Yu, Y., and Coauthors, 2008: Coupled model simulations of climate changes in the past century and future climate. *Adv. Atmos. Sci.*, **25**(4), 641–654, doi: 10.1007/s00376-008-0641-0.

Zeng, Q. C., X. Zhang, X. Liang, C. Yuan, and S. Chen, 1989: Documentation of IAP two-level atmospheric general circulation model. TR044: DOE/ER/60314-H1.

Zhang, G. J., and N. A. McFarlane, 1995: Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Climate Centre general circulation model. *Atmos.-Ocean*, **33**, 407–446.

Zhang, G. J., and H. Wang, 2006: Toward mitigating the double ITCZ problem in NCAR CCSM3. *Geophys. Res. Lett.*, **33**, L06709, doi: 10.1029/2005GL025229.

Zhang, L., and J. Li, 2007: Seasonal rotation features of wWind vectors and application to evaluate monsoon simulations in AMIP models. *Climate Dyn.*, **31**, doi: 10.1007/s00382-007-0327-9.

Zhang, X., and X. Liang, 1989: A numerical world ocean general circulation model. *Adv. Atmos. Sci.*, **6**, 43–61.

Zhang, X., N. Bao, R. C. Yu, and W. Wang, 1992: Coupling scheme experiments based on an atmospheric and an oceanic GCM. *Chinese J. Atmos. Sci.*, **16**(2), 129–144. (in Chinese)

Zhang, X. H., G. Y. Shi, H. Liu, and Y. Q. Yu, 2000: *IAP Global Ocean—Atmosphere-Land System Model*, Science Press, Beijing, New York, 252pp.

Zhang, X., G. Shi, H. Liu, and Y. Yu, Eds., 2000: *IAP Global Ocean-Atmosphere-Land System Model*, Science Press, Beijing, 252pp.

Zhang, X., W. Lin, and M. Zhang, 2007: Toward understanding the double Intertropical Convergence Zone pathology in coupled ocean-atmosphere general circulation models. *J. Geophys. Res.*, **112**, D12102, doi: 10.1029/2006JD007878.

Zhang, L.-X., T. -J. Zhou, B. Wu, and Q. Bao, 2008a: The annual modes of tropical precipitation simulated by LASG/IAP ocean-atmosphere coupled model FGOALS_s1.1. *Acta Meteorologica Sinica*, **66**(6), 968–981. (in Chinese)

Zhang, Y. -C., M. Takahashi, and L.-L. Guo, 2008b: Analysis of the East Asian Subtropical Westerly Jet Simulated by CCSR/NIES/FRCGC Coupled Climate System Model. *Journal of the Meteorological Society of Japan*, **86**, 257–278.

Zhou, T. J., and Coauthors, 2005: The climate system model FGOALS-s using LASG/IAP Spectral AGCM SAMIL as its atmospheric component. *Acta Meteorologica Sinica*, **63**(5), 702–715. (in Chinese)

Zhou, T. J., and R. C. Yu, 2006: Twentieth century surface air temperature over China and the globe simulated by coupled climate models. *J. Climate*, **19**(22), 5843–5858.

Zhou TJ (Zhou Tianjun)Yu, Y. Q., H. L. Liu, W. Li, X. B. You, and G. Q. Zhou, 2007: Progress in the development and application of climate ocean models and ocean-atmosphere coupled models in china. *Adv. Atmos. Sci.*, **24**(6), 1109–1120, doi: 10.1007/s00376-007-1109-3.

Zhou, T. J., and Coauthors, 2008a: The CLIVAR C20C Project: Which components of the Asian-Australian monsoon variability are forced and reproducible? *Climate Dyn.*, doi: 10.1007/s00382-008-0501-8.

Zhou, T. J., B. Wu, and B. Wang, 2008b: How well do atmospheric general circulation models capture the leading modes of the interannual variability of Asian-Australian monsoon? *J. Climate*, **21**, doi: 10.1175/2008JCLI2245.1.

Zhou, T. J., B. Wu, X. Wen, L. Li, and B. Wang, 2008c: A fast version of LASG/IAP climate system model and its 1000-year control integration. *Adv. Atmos. Sci.*, **25**(4), 655–672, doi: 10.1007/s00376-008-0655-7.