Initial CEOP-based Review of the Prediction Skill of Operational General Circulation Models and Land Surface Models

Kun YANG, Mohamed RASMY, Surendra RAUNIYAR, Toshio KOIKE, Kenji TANIGUCHI, Katsunori TAMAGAWA, Petra KOUDELOVA

Department of Civil Engineering, The University of Tokyo, Tokyo, Japan

Masaru KITSUREGAWA, Toshihiro NEMOTO, Masaki YASUKAWA

Institute of Industrial Science, The University of Tokyo, Tokyo, Japan

Eiji IKOMA

Center for Spatial Information Science, The University of Tokyo, Tokyo, Japan

Michael G. BOSILOVICH

NASA Data Assimilation Office, Greenbelt, Maryland, USA

and

Steve WILLIAMS

UCAR / Joint Office for Science Support, NCAR, Colorado, USA

(Manuscript received 24 February 2006, in final form 15 November 2006)

Abstract

Using data archived in the Coordinated Enhanced Observing Period (CEOP) project, this study presents an initial evaluation of the prediction skill of five General Circulation Models (GCMs) and three Land Surface Models (LSMs). Comparisons between observations and the GCMs show that all the models are able to produce an afternoon peak in precipitation, but other major features are not well produced, including the total amount of precipitation, onset time of the afternoon peak, the early-evening low (around 1800 LST), and the partition between convective and stratiform rainfall. The ratios of evaporation to precipitation differ among the GCMs. Evaporation in some of the GCMs is even greater than precipitation, perhaps due to the model spin-up effect. In terms of the surface radiation budget, the GCMs generally over-predict downward shortwave radiation and under-predict downward longwave radiation; further investigations of the causes of these trends require cloudiness observations. In terms of the surface energy budget, the GCMs generally over-predict nighttime downward sensible heat fluxes

Corresponding author: Kun Yang, Dr., River Lab, Department of Civil Engineering, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8656, Japan.

E-mail: yangk@hydra.t.u-tokyo.ac.jp

© 2007, Meteorological Society of Japan

and under-predict diurnal ranges of surface–air temperature difference, as heat transfer resistances are under-predicted. Finally, three offline LSMs driven by identical forcing are evaluated, and we note that the reproduction of surface temperature is not a sufficient condition for a LSM to reproduce surface energy partition.

1. Introduction

Diagnosing model errors obtained by comparing model output with observations are the basis for improving representations of key physical processes in any General Circulation Model (GCM) (Phillips 2004). Many such studies have been conducted via international cooperation projects such as PILPS (Project for Inter-comparison of Land-Surface Schemes) (Pitman et al. 1999; Henderson-Sellers et al. 2003) and AMIP (Atmospheric Model Inter-comparison Project) (Gates et al. 1999). PILPS has designed several phases used to investigate model discrepancies, model errors, and, jointly with AMIP, investigate land–atmosphere interactions. These studies have found large differences in surface energy partitioning among PILPS models and reveal that simple biosphere Land Surface Models (LSMs) usually perform better than bucket models; however, as only limited observational data are available with which to evaluate PILPS models, it is difficult to identify common model deficiencies. AMIP provides a community-based infrastructure in support of climate-model diagnosis, validation, and inter-comparisons. These models are driven by observed sea-surface temperature and sea-ice distributions over the period 1979–1988. AMIP validation is mainly based on reanalysis data and satellite data. The AMIP models' ensemble mean has produced reasonable average large-scale seasonal distributions of pressure, temperature, and circulation (Gates et al. 1999), but details related to land–atmosphere interactions need to be further addressed using in situ data.

As a coordinated international activity of the World Climatic Research Programme (WCRP) initiated by the Global Energy and Water Cycle Experiment (GEWEX), the Coordinated Enhanced Observing Period (CEOP) is seeking to archive data of common measurements from in situ and satellite remote sensing measurements, as well as matching model output with four-dimensional data assimilations. CEOP

not only integrates these types of data at different levels (reference site, monsoon region, and global-scale), but also develops corresponding facilities that enable data analysis (Koike 2002, 2004). The details of the project and reports on progress can be found in relevant CEOP newsletters. There are four Enhanced Observing Periods (EOP1–EOP4) in CEOP phase 1 (from 2002 July to the end of 2004), and data collection will continue in the forthcoming CEOP Phase 2. CEOP-archived data have been used to calibrate/evaluate models, improve model skill at global and regional scales, and develop land-data assimilation systems (Koike et al. 2003; Hirai and Matsumura 2004; Lu and Mitchell 2004; Yang et al. and several other papers in the current issue).

In particular, CEOP data provide a golden opportunity for model assessment. First, most of the GCMs that contribute to CEOP are weather-forecasting models. They usually have higher spatial resolutions than those of AMIP GCMs and are reinitialized six-hourly or daily for operational weather forecasting. It is therefore expected that they will have superior prediction skill to AMIP GCMs. Second, CEOP reference sites have been selected to ensure that the sites cover a variety of climate regimes. The observations include high-accuracy measurements of major variables of relevance to hydro-meteorological studies (surface, subsurface, and atmospheric variables), making it possible to identify the effectiveness and deficiencies of parameterization schemes of physical processes. Third, CEOP has archived and will continue to archive field observation data with high temporal resolution (hourly or half-hourly) and specialized model output (hourly or three-hourly); this provides the opportunity to study diurnal variations in the variables of interest. These data have been used for inter-comparisons over short periods (Yang et al. 2005), for individual sites, regions (e.g., Beyrich and Adam 2004; Ruane and Roads this issue; Rikus this issue), and single parameters (Tamagawa et al. 2003). These studies are important in de-

Table 1. Major characteristics of the six GCMs used in this study and their MOLTS data (sourced from the CEOP webpage: <http://monsoon.t.u-tokyo.ac.jp/ceop/index.html>).

Model Name	Model Horizontal Resolution	Number of Vertical Levels	Land Surface Scheme	CEOP Products	Output Interval
BMRC Operational Global Medium Range Prediction Model	T239L29 ($\Delta x \sim 80$ km)	29 sigma	Bucket hydrology	36 hr forecast/analysis (once per day)	1 hr
ECPC SFM (Seasonal Forecast Model)	T62L28 ($\Delta x \sim 280$ km)	28 sigma	Oregon state	6 hr forecast (4 times per day);	3 hrs
ECPC RII (Reanalysis-II)		17 pressure	University LSM ver. 2	36 hr forecast (once per day)	
JMA GSM (Global Spectral Model)	T213L40 ($\Delta x \sim 60$ km)	23 pressure	Simple Biosphere	6 hr forecast (4 times per day)	1 hr
NCEP GFS (Global Forecast System since 29 Oct 2002)	T254L64 ($\Delta x \sim 50$ km)	64 pressure	Simple Biosphere	84 hr forecast (once per day)	3 hrs
UKMO Global Unified Model	0.83×0.56 deg.	38 model levels	MOSES ¹	36 hr forecast/analysis (once per day)	3 hrs

¹MOSES: Meteorological Office Surface Exchange Scheme.

testing model deficiencies and thus improving the representations of key processes, as well as assisting users to identify useful data from within rapidly expanding databases.

Within the framework of a full annual cycle and multiple sites, parameters, and models, this paper presents an initial review of the prediction skill of five GCMs and three off-line LSMs in a land data assimilation system (LDAS). We achieve this by comparing observed and modeled diurnal cycles of precipitation, surface water budgets, radiation budgets, and energy budgets.

2. Data and methodology

CEOP has defined 35 reference sites worldwide, where field observations from research institutes and national services, as well as model output location time series (MOLTS) from nine numerical weather prediction (NWP) centers and two data assimilation centers, have been or are to be archived. Currently, the data for EOP3 (from 1 October 2002 to 30 September 2003) are the most widely archived, including in situ data for 27 reference sites, 3 LSMs' MOLTS of the Global Land Data Assimilation System (GLDAS) of the National Aeronautics and Space Administration (NASA), and 6

GCMs' MOLTS from the Bureau of Meteorology Research Centre, Australia, (BMRC), the Experimental Climate Prediction Center (ECPC) at Scripps Institution of Oceanography, the Japan Meteorological Agency (JMA), National Centers for Environmental Prediction (NCEP), USA, and the Met Office in the UK (UKMO). Note there are two GCMs from ECPC: SFM (Seasonal Forecast Model) and RII (Reanalysis-II). GLDAS MOLTS are three-hourly output of three offline LSMs that are driven by identical forcing data with a horizontal resolution of $1^\circ \times 1^\circ$. The major characteristics of the 6 GCMs and their MOLTS are briefly summarized in Table 1. Additional details of the GCMs are described in several accompanying papers in this issue. For convenience, each model is herein referred to using the name of its corresponding center.

This study is based on EOP 3 data. The focus of the paper is a comparison of observational data with BMRC concatenated 12–36 hr data of 36-hr forecasts (the available data), JMA concatenated 6-hr forecasts (the available data), ECPC SFM concatenated 6-hr forecasts, the first 24-hr data of NCEP 84-hr forecasts, and the first 24-hr data of UKMO 36-hr forecasts. Additional comparisons with other forecasts

and analysis products are also conducted as required for specific analyses. ECPC RII data are not used because SFM is the more updated model in ECPC (Alexander Ruane, pers. com. 2006). Table 2 lists the location and data availability at the reference sites, among which there are 24 air temperature sites, 21 air humidity sites, 19 precipitation sites, 14 surface temperature sites, 21 surface radiation sites, and 8 turbulent fluxes sites. Table 3 shows the in situ and GCM land-use types at these sites.

In this study, we evaluate the skill of the GCMs from the viewpoints of both water and energy. Because the spatial variability of a variable may contribute to differences between patch-scale in situ observations and grid-based model output (e.g., Bosilovich 2002; Beyrich and Adam 2004; Yang et al. 2005), we generally compare multi-site composite monthly mean, annual mean, and annual mean diurnal cycles rather than hourly, three-hourly, or daily values at individual sites. The monthly and annual mean data are calculated from data of monthly mean diurnal variations for each site and each month; however, when more than half of the observed data are not available for a month at a site (e.g., Sodankylä) or when the monthly mean temperature difference between observations and the simulation is more than 20 K (e.g., East Siberia–Taiga), the relevant data are excluded from the comparisons. The latter case occurred occasionally, but this is not considered to be a model error. In addition, sites with contrasting in situ and model land-use types are excluded from the comparisons of surface temperature and fluxes, as the values recorded at these sites are potentially sensitive to surface conditions. For example, observations of surface variables at small island sites cannot be used to represent the values for a GCM grid that encompasses the surrounding sea surface.

3. Diurnal cycle of precipitation in GCMs

An accurate representation of the diurnal cycle of precipitation over land and ocean provides a key test of many aspects of the physical parameterizations in a climate model, from radiative transfer and surface exchanges through to boundary layer, convection, and cloud processes (Randall et al. 1991; Lin et al. 2000). Yang et al. (2001) showed that the UKMO

model has considerable difficulty in capturing the observed phase of the diurnal cycle in convection. Many studies have focused on the diurnal cycle, using surface observations, satellite data, and GCM output. Most of these studies found that the precipitation maximum tends to occur during the early morning over open oceans and in the late afternoon/early evening over land; however, most of these studies were confined to particular regions (e.g., Betts and Ball 1995; Nesbitt et al. 2003; Lau et al. 2004; Hirose and Nakamura 2005).

This section evaluates the diurnal cycle of precipitation using data for the summer season. This season is defined as JJA in the Northern Hemisphere, DJF in the Southern Hemisphere, and MAM in equatorial regions. In total, there are 16 sites (bold sites in the “Code” column of Table 2) with available data for the defined summer.

First, the diurnal cycle derived from in situ data was compared with those in the literature, revealing that the two sets of data are in general agreement. The NSC site (full names are listed in Table 2) is located on a subtropical island and is subjected to heavy rainfall. EOP3 observations reveal that the maximum rainfall hour at this site is 1700–1900 LST, which is similar to the findings (1600–1700 LST) of Chen et al. (1999) based on more than 300 stations and the findings (1700 LST) of Kishtawal and Krishnamurti (2001) based on Tropical Rainfall Measuring Mission (TRMM) data. At the HIM site, EOP3 data reveal that the maximum precipitation hour is 2300 LST; the same result was also reported by Tartari et al. (1998) and Ueno et al. (2001) for nearby stations and Hirose and Nakamura (2005) for a station located south of Mt. Everest (>5500 m ASL). We also identify a daytime peak at 1400–1600 LST at the East Tibet reference site, which is consistent with the findings of Hirose and Nakamura (2005). At the DAR site, EOP3 observations reveal two precipitation peaks at 1500 LST and 0200 LST, while Soman et al. (1995) identified two peaks at 1500 LST and 2300 LST based on a short-term (19 days) dataset. The 1700 LST maximum in precipitation recorded at the ARM SGP site is similar to the findings (1600–1800 LST) of Dai (2001), who used three-hourly weather reports of precipitation and thunderstorm events for the period 1975–97. Lin et al.

Table 2. Availability of EOP3 geographic information and data from CEOP reference sites (blank: no data; ● : data used; ○ : data used but q_a unavailable; ○ : data questionable or vegetation-type mismatched. In the Code column, the bold sites are those used to analyze diurnal cycles of precipitation; grayed sites were used for the study of local minimum precipitation described in Section 3.4).

CSE	Reference Site Name	Code	Lat. (N)	Lon. (E)	Ele. (m)	Data availability					References/PI
						Ta&q _a	Tsfc	Radiation	Flux	Rain	
BALTEX	Lindenbergl	LIN	52.2	14.1	112	●	●	●	●	●	Beyrich & Adam (2004)
	Cabauw	CAB	52.0	4.9	−1	●	●	●	●	●	Isemer (2002)
	Sodankylä	SOD	67.4	26.7	179	●		●		○	Isemer (2002)
GAPP	ARM Southern Great Plains	SGP	36.6	−97.5	313	●		●	●	●	Raymond McCord
	Bondville	BON	40.0	−88.3	300	●	●	●	●	●	Tilden P. Meyers
	Fort Peck	FPE	48.3	−105.1	800	●	●	●	●	●	Tilden P. Meyers
	Oak Ridge	ORI	36.0	−84.3	275	●	●	●	●	●	Tilden P. Meyers
	Eastern Siberian Tundra	ES1	71.6	128.8	38	●	●	●			Ohata et al. (1999)
CAMP	Eastern Siberian Tiaga	ES2	62.3	129.6	220	●	●	●			Ohta et al. (2001)
	Mongolia	MON	45.7	106.3	1393	●	●			●	Kaihotsu et al. (2003)
	Tongyu	TON	44.4	122.9	184	●		●	●	●	Liu et al. (2006)
	Tibet	TIB	31.4	91.9	4580	●	●	●		●	Ishikawa et al. (2001)
	West Tibet	GAI	32.5	84.1	4416	●	●	●			Ishikawa et al. (2001)
	Himalayas	HIM	28.0	86.8	5050	●		●		●	Bollasina et al. (2002)
	North South China Sea	NSC	25.0	121.2	8	○		○		●	Chen et al. (2004)
	Korean Haenam	KHA	34.6	126.6	14	⊙		●			Kim et al. (2002)
	Korean Peninsula	KPE	37.4	127.9	330	●		●		●	Kim et al. (2002)
	Chao-Phraya River—Lampang	CPL	18.4	99.5	241	⊙		●		●	Masatoshi AOKI
	North-East Thailand	NET	14.5	102.4	311	⊙		●		●	Masatoshi AOKI
LBA	Western Pacific Ocean	WPO	7.1	134.3	40	○		○		●	Kubota et al. (2002)
	Equatorial Island	EIS	−0.2	100.3	699	●				●	Mori et al. (2004)
	Manaus	MAN	−2.6	−60.2	130	●	●	○	○	●	Marengo et al. (2003)
	Santarem	SAN	−3.0	−55.0	N/A	●		●		●	Marengo et al. (2003)
	Pantanal	PAN	−19.6	−57.0	N/A	●	●	●		●	Marengo et al. (2003)
ARM	ARM NSA-Barrow	NSA	71.3	−156.6	8	●	●	●		●	Tilden P. Meyers
	ARM TWP-Manus	MNS	−2.1	147.4	4	○		○		●	Raymond McCord
	ARM TWP-Darwin	DAR	−12.4	130.9	29.9	●	●	●		●	Raymond McCord

Table 3. In situ and GCM land-use types at CEOP sites. Italics indicates turbulent flux sites. NET: Needleleaf evergreen trees; NDT: Needleleaf deciduous trees; BDT: broadleaf deciduous trees; GMF: grass and mixed farming. The UKMO column shows the land-use type with the maximum area fraction. No explicit vegetation is considered in the BMRC model.

CSE	Reference Site Name	Land use					
		In-situ	JMA	NCEP	ECPC	UKMO	GLDAS
BAITEX	<i>Lindenberg</i>	<i>grass</i>	<i>savanna</i>	<i>cultivations</i>	<i>GMF</i>	<i>grass</i>	<i>cropland</i>
	<i>Cabauw</i>	<i>short grass</i>	<i>savanna</i>	<i>groundcover</i>	<i>grass</i>	<i>grass</i>	<i>cropland</i>
	Sodankylä	NET	NET	NET	needleleaf forest	needleleaf trees	bareground
GAPP	ARM SGP	<i>grass</i>	<i>mixed forest</i>	<i>cultivations</i>	<i>grass</i>	<i>grass</i>	<i>cropland</i>
	<i>Bondville</i>	<i>cropland</i>	<i>savanna</i>	<i>cultivations</i>	<i>GMF</i>	<i>grass</i>	<i>cropland</i>
	<i>Fort Peck</i>	<i>grass</i>	<i>grass</i>	<i>groundcover</i>	<i>grass</i>	<i>grass</i>	<i>grass</i>
	<i>Oak Ridge</i>	<i>mixed forest</i>	<i>mixed forest</i>	<i>mixed forest</i>	<i>deciduous forest</i>	<i>broadleaf trees</i>	<i>wooded grass</i>
	Eastern Siberian tundra	open shrubland	sea	tundra	tundra	bare Soil	open shrubland
CAMP	Eastern Siberian Tiaga	NDT	NDT	NDT	deciduous forest	Needleleaf trees	wooded grass
	Mongolia	grass	grasses	bare soil	grass	bare Soil	grass
	<i>Tongyu</i>	<i>cropland</i>	<i>grasses</i>	<i>cultivations</i>	<i>GMF</i>	<i>grass</i>	<i>cropland</i>
	Tibet	grass	grasses	bare soil	ice	grass	grass
	West Tibet	grass	bare soil	bare soil	ice	shrubs	bareground
	Himalayas	grass	grasses	mixed forest	grass	grass	bareground
	North South China Sea	N/A	sea	groundcover	sea	grass	shrubs
	Korean Haenam	groundcover	savanna	groundcover	GMF	grass	
	Korean Peninsula	BDT	BDT	cultivations	grass	grass	cropland
	Lampang	deciduous forest	savanna	cultivations	tropical forest	broadleaf trees	N/A
	North-East Thailand	BDT	savanna	BDT	GMF	broadleaf trees	cropland
	Western Pacific sea	N/A	sea	sea	sea	sea	sea
	Equatorial Island	N/A	tropical forest	cultivations	tropical forest	broadleaf trees	grass

of precipitation intensity occurs during 1400–1600 LST, with rapid development and decay. The observed diurnal cycle of precipitation frequency is similar to that of precipitation intensity. All the models generally produce an afternoon peak in precipitation intensity and frequency. The starting time of the afternoon peak predicted by JMA and NCEP is similar to the observed time (1300 LST), while the other GCMs predict start times that are earlier than the observed time. For the afternoon peak, there is no remarkable difference between the forecast and analysis diurnal cycles for either BMRC or UKMO. This implies that the diurnal cycle is mainly determined by the nature of the model. The patterns of the diurnal cycle of ECPC 0–6 hr and 12–36 hr forecasts are different (Fig. 1), indicating that the spin-up effect is significant for ECPC SFM; however, this effect is small for NCEP and UKMO (not shown).

Another outstanding problem is the fact that no model is able to produce an early-evening (around 1800 LST) local minimum in rainfall intensity, which is clear in the observed composite diurnal cycle and also at many sites (see the nine gray-colored cells in Table 2). This minimum intensity is seldom reported and has not been addressed in the literature. As this minimum has not been reproduced in the model, we take the alternative approach of investigating the mechanism using hourly GMS (Geostationary Meteorological Satellite) and GOES (Geostationary Operational Environmental Satellite) data. A Convection and Stratiform Technique (CST; Goldenburg et al. 1990) is used to separate observed rainfall into convective and stratiform rainfall at the nine sites (Fig. 2). It is clear that the early-evening minimum in rainfall intensity is caused by the sudden cessation of convective rainfall in the early evening followed by a start of stratiform rainfall. This suggests that the rainfall minimum may be related to the lifetime of convective activity in the late afternoon. The derived ratio of convective to total rainfall is approximately 36%, which is much lower than the predictions of some GCMs (Table 4).

4. Surface water and energy budget in GCMs

Figure 3 shows the Mean Bias Errors (MBEs) and Root Mean Square Errors (RMSEs) of

Fig. 2. Composite diurnal cycle of convective precipitation and stratiform precipitation derived using a convective and stratiform technique (CST) from in situ and satellite data recorded at 9 sites (grayed sites in Table 2) during their summer seasons.

Table 4. Partitioning of convective and stratiform rainfall [%] at the nine sites that record an early-evening minimum in precipitation intensity (grayed sites in Table 2). N/A for the BMRC and JMA models.

Methods/Models	Convective	Stratiform
CST Method	36	64
UKMO-F	84	16
NCEP	62	38
ECPC_SFM (0~6 hr)	80	20
ECPC_SFM (12~36 hr)	81	19

GCMs, as calculated using monthly mean data at all of the sites of interest. The values in parentheses along the horizontal axes denote the measured annual mean values. We emphasize the following points. (1) When the model elevation is different from the in situ elevation, modeled air temperature is corrected using a lapse rate of 0.0065 K m^{-1} and LWD is corrected with the addition of 2.8 W m^{-2} per 100 m, following Wild et al. (2001). (2) ECPC represents SFM 0–6 hr forecasts throughout this section unless stated otherwise. (3) A surface energy imbalance has been widely observed but has yet to be solved. The following comparisons do not take this energy imbalance issue into account, and this may affect our comparisons to some degree.

Fig. 3. Mean Bias Error (*MBE*) and Root Mean Square Error (*RMSE*) of GCMs calculated using monthly mean values at all the sites of interest during the period October 2002 to September 2003. Values in parentheses along the horizontal axes represent the measured mean values. *MBE* and *RMSE* have the same unit as the mean value. *SWD*: Downward Shortwave Radiation; *LWD*: Downward Longwave Radiation; *Rn*: Net Radiation; *H*: Sensible heat flux; *IE*: Latent heat flux; *P*: Precipitation; *E*: Evaporation.

4.1 Surface water budget

Figure 3a1 shows that NCEP and JMA over-predict precipitation ($MBE > 0$), while BMRC and UKMO both under-predict ($MBE < 0$). The precipitation difference between NCEP and BMRC is comparable to the magnitude of the observed precipitation ($\sim 2.9 \text{ mm d}^{-1}$). Figure 3a2 shows that the *RMSE* values ($\sim 3 \text{ mm d}^{-1}$) are also comparable to this magnitude, indicating that the predicted precipitation contains large errors. All of the models show significant over-estimates of precipitation for some sites (Himalayas) and under-estimates for other sites (Darwin, Manus, NSA-Barrow) for the summer season. These results highlight the difficulties involved in predicting precipitation using a GCM, as has been reported in many previous studies.

An important issue is the relationship between precipitation and evaporation at the land surface. There are a total of eight turbulent flux sites. Table 5 shows the annual mean evapora-

Table 5. Composite annual mean evaporation (*E*) and precipitation (*P*) and their ratio (*E/P*) at the eight turbulent flux sites.

Data	<i>E</i>	<i>P</i>	<i>E/P</i>
Observed	1.76	2.09	0.84
BMRC	2.29	1.53	1.50
ECPC	2.00	1.24	1.61
JMA	1.99	2.65	0.75
NCEP	3.16	3.32	0.95
UKMO	2.16	2.82	0.77

tion (E), precipitation (P), and evaporation–precipitation ratio (E/P) for the eight sites. It is clear that evaporation is over-predicted by all of the GCMs, particularly by NCEP. JMA and UKMO under-predict the ratio of evaporation to precipitation and NCEP over-predicts the ratio. ECPC and BMRC predict that evaporation exceeds precipitation, which is contrary to the observed fact that evaporation is less than precipitation over land. This problem might be associated with model spin-up. Using ECPC 36-hr forecast, we found that the ratio is sensitive to the hours of interest, e.g., the ratio is 0.9 when the first 24-hr data are used and 0.72 if concatenated 12–36 hr data are used. In contrast, the E/P ratios derived from NCEP and UKMO are insensitive to model spin-up; for BMRC and JMA, additional forecast data serials are required to assess the spin-up effect.

We also examined the correlation between evaporation error (δE) and precipitation error (δP) in each GCM. The correlation in NCEP (Fig. 4a) is stronger than that in JMA (Fig. 4b) and UKMO (not shown). Mitchell et al. (2005) showed that precipitation biases decreased after replacing the LSM in the NCEP model with a new LSM that was able to reduce warm-season surface evaporation. It therefore appears that the correlation between δE and δP represents the intensity of land–atmosphere coupling in the GCMs.

4.2 Surface radiation budget

Figures 3b1–b2 show the errors in the annual mean surface radiation budget. It is clear that downward shortwave radiation (SWD) is over-predicted by all of the models ($MBE > 0$), and downward longwave radiation (LWD) is under-predicted by all the models except for BMRC, which yields a slight over-prediction. Large biases for SWD ($\sim 30 \text{ W m}^{-2}$) are produced by all of the models except UKMO, and large biases for LWD (-30 W m^{-2}) are produced by JMA and ECPC. UKMO records relatively small biases for both SWD (8 W m^{-2}) and LWD (-4 W m^{-2}); BMRC also records a small LWD bias (2 W m^{-2}). The SWD and LWD errors counteract each other; this generally results in small biases for the total downward radiation. Accordingly, ECPC, JMA, and UKMO are able to well simulate annual mean total downward radiation and net radiation; how-

Fig. 4. Errors in GCM monthly mean precipitation (δP) and evaporation (δE) for eight flux sites over the period October 2002 to September 2003. (a) NCEP data. (b) JMA data.

ever, seasonal variations in SWD and LWD and their total amounts are better simulated in UKMO than ECPC or JMA (Fig. 5). Figure 3b2 also shows that UKMO has a smaller $RMSE$ value in predicting radiation.

The over-prediction of SWD and under-prediction of LWD have been reported in previous studies (e.g., Cess et al. 1995; Wild et al. 1995; Garratt and Prata 1996; Wild et al. 2001). For SWD , under-estimation of clear-sky absorption and cloud absorption in GCMs has been suggested by Ramanathan et al. (1995) and Wild et al. (2006), respectively. For LWD , Wild et al. (2001) suggested that the LWD schemes themselves play a major role in the under-estimation. The error in surface radiation might also reflect an under-prediction of

Fig. 5. Composite seasonal variations in downward shortwave and longwave radiation and their total values derived from in situ data and GCMs at 21 radiation sites during the period October 2002 to September 2003.

cloudiness, as suggested by Wild et al. (1995) for a GCM, Siebesma et al. (2004) for several GCMs in stratocumulus regions, Milton and Earnshaw (this issue) for the UKMO model, and Hirai and Sakamoto (pers. com. 2006) for the JMA model. As CEOP Phase 1 did not collect cloudiness data, it is not readily possible to evaluate whether the observed discrepancies between the models and observations are based

on cloud errors or shortcomings in the employed radiation scheme.

In addition, the diurnal variation of LWD (data not shown) shows that all the GCMs are not able to predict LWD peak timing. The observed LWD peak occurs in the early afternoon, but these GCMs produce a peak in the middle or late afternoon.

4.3 Surface energy budget

For the eight turbulent flux sites, Fig. 3b2 shows that the *RMSEs* for the sensible heat fluxes of all of the models ($>20 \text{ W m}^{-2}$) are comparable to the observed mean value of sensible heat fluxes (23 W m^{-2}), indicating that the surface energy budget is poorly modeled. JMA and UKMO produce relatively small *RMSE* values for both sensible and latent heat fluxes, ECPC predicts relatively accurate latent heat fluxes but estimates sensible heat fluxes that are nearly double observed values, and NCEP over-predicts latent heat fluxes and under-predicts sensible heat fluxes. The total turbulent fluxes are best simulated by JMA and UKMO; the other models over-predict the total by 25 W m^{-2} .

Figure 6 shows the multi-site composite seasonal variations in surface energy fluxes. It is clear that the ranges of seasonal variations of all the energy components are over-predicted by all of the models. The over-prediction of net radiation is attributed to over-prediction of the total downward radiation and under-prediction of outgoing longwave radiation (or surface temperature; see Fig. 3a1). Sensible heat fluxes are predicted more accurately by JMA and UKMO but are substantially over-predicted by BMRC and ECPC for the boreal summer. It is noticeable that all of the models over-predict downward sensible heat fluxes in the boreal winter. Latent heat fluxes are more accurately predicted by ECPC and JMA than the other models. BMRC uses a bucket hydrological model without any explicit representation of vegetation and therefore yields poor predictions of the seasonal trend in latent heat fluxes (over-predicted in the MAM season and under-predicted in the JJA season).

Figure 7 shows the composite annual-mean diurnal variations in surface energy fluxes. All the models predict reasonable phases of net radiation and sensible heat fluxes, but the

Fig. 6. Composite seasonal variations in the surface energy budget derived from in situ data and GCMs during the period October 2002 to September 2003.

predicted peak time of latent heat fluxes is 1–2 hours later than the observed time. All the models yield significant over-predictions of nighttime downward heat fluxes. To explain this phenomenon, we plotted the composite diurnal variations in surface–air temperature differences in Fig. 8a; the figure shows that the diurnal range is under-predicted by all of the models. This under-prediction is more significant for arid and semi-arid sites than for forested sites (not shown). We then calculated the

Fig. 7. Composite diurnal variations in the surface energy budget derived from in situ data and GCMs at eight flux sites during the period October 2002 to September 2003.

heat transfer resistance according to the following formula:

$$r_h = \frac{\rho c_p (T_{sfc} - T_{air})}{H} \quad (1)$$

where r_h [$s\ m^{-1}$] is the heat transfer resistance, H [$W\ m^{-2}$] is the sensible heat flux, ρ [$kg\ m^{-3}$] is the air density, T_{sfc} [K] is the surface skin temperature, T_{air} [K] is the surface air temper-

Fig. 8. (a) Composite diurnal variations in surface-air temperature difference derived from in situ data and GCMs at 14 surface temperature sites. (b) Annual mean diurnal variations in heat transfer resistance for a semi-arid site (Tongyu) during the period October 2002 to September 2003. Not available for BMRC.

ature, and c_p ($= 1004 \text{ J kg}^{-1} \text{ K}^{-1}$) is the specific heat of air at a constant pressure.

Figure 8b shows the heat transfer resistance for a semi-arid site (Tongyu). Oscillations at times close to sunset and sunrise are caused by near-zero heat fluxes and temperature differences. It is clear that all of the GCMs under-predict the resistances for this site; this in turn leads to the small surface-air temperature differences displayed in Fig. 8a. The resistances are strongly dependent on the setting of momentum roughness lengths and thermal roughness lengths. The empirical relationships between the two lengths are developed from field experiments conducted in small and uniform surfaces where momentum fluxes are unaffected by meso-scale topographic undulations. Such undulations are common in GCM grids and may strongly enhance momentum loss while contributing little to heat transfer (Yang et al. 2002). Accordingly, a large ratio of the length of momentum roughness to that of thermal roughness is required to produce realistic momentum and heat fluxes; such a large ratio can differ markedly from that derived from patch-scale experiments. This upscaling issue for GCM simulations requires further investigation.

The accuracy of modeled air temperature and humidity relies on the accuracy of the surface energy partition. Figure 9 shows an example for the BALTEX/Cabauw site. Clearly, BMRC and ECPC over-predict air temperature (Fig.

9a) and under-predict humidity (Fig. 9b) during the period June–August. These large biases are caused by the excessive transfer of heat and insufficient transfer of moisture into surface boundary layers in the GCMs (Fig. 9c–9d). Similar error-correlations are also found for other sites (Manaus, Pantanal, and Santarem). The other models show similar phenomena for different sites. Accordingly, a highly accurate LSM plays a key role in obtaining accurate forecasts of these near-surface variables.

5. Surface energy budget in GLDAS

GLDAS embeds three land models: Mosaic (Koster and Suarez 1996), Noah LSM (Chen et al. 1996; Koren et al. 1999), and the Common Land Model (CLM; Dai et al. 2003). GLDAS users can select forcing data from an Atmospheric Data Assimilation System (ADAS) or a reanalysis product. Observation-derived fields, including precipitation and radiation, are optionally specified to replace the corresponding ADAS forcing fields when and where possible (Rodell et al. 2004). Compared with the GCMs, GLDAS land surface forcings show relatively good or moderate accuracy (not shown).

For EOP3, GLDAS generally produces better surface temperature than the GCMs, but the improvement for arid regions (e.g., West Tibet) is limited (not shown). Figures 10a1,b1 show two examples of monthly mean diurnal variations in surface skin temperature at Cabauw and Bondville. CLM and Noah are able to re-

Fig. 9. Observed and GCM-modeled monthly mean diurnal variations in air temperature and specific humidity (a and b, respectively) and monthly mean sensible and latent heat fluxes (c and d, respectively) for the Cabauw site during the period October 2002 to September 2003.

Fig. 10. Monthly mean diurnal variations in surface skin temperature and monthly mean turbulent heat fluxes derived from in situ data and three GLDAS products (Mosaic, CLM, and Noah) for two sites during the period October 2002 to September 2003. (a) Cabauw data. (b) Bondville data.

produce the observed surface temperatures for the two sites, but Mosaic produces cold biases at nighttime. This is attributed to a deficiency in the model, as the three LSMs are driven by identical forcing data.

Relative to GCMs, these offline LSMs do not show significant improvements in the simulation of surface energy partition. The surface energy budget simulated by the offline LSMs (Fig. 10) is no better than that simulated by JMA and UKMO (Fig. 9) for the Cabauw site. Noah and CLM produce different energy fluxes, although their surface skin temperatures are similar. For the boreal summer, not only do the modeled energy fluxes deviate far from the observed values, but the differences among the models are comparable to the differences between observations and models. This demonstrates large uncertainties among the LSMs; this point is also addressed by Kato et al. (this issue).

6. Concluding remarks and recommendations

Based on CEOP/EOP3 in situ observations and model output, we presented an initial evaluation of the prediction skill of five weather-forecast GCMs in terms of surface water and energy at diurnal, seasonal, and annual scales. The surface energy budget in three offline GLDAS LSMs was also evaluated.

Comparisons of the diurnal cycle of precipitation between observations and GCMs reveal that all of the models are able to reproduce an afternoon peak in precipitation, but none of them are able to reproduce other major features of precipitation, including onset time of the afternoon peak, the early-evening low intensity, and the partition between convective and stratiform precipitation. Because of the complexity of rainfall mechanisms, obtaining an understanding of the model deficiencies is only possible via international collaborative projects such as the EUROCS (EUROpean Cloud Systems study) scientific community, which focuses on process studies of diurnal cycles of clouds and precipitation (Siebesma et al. 2004; Bechtold et al. 2004).

The water budget in the different GCMs shows significant variation. ECPC and BMRC produce evaporation that exceeds precipitation over land-based sites, perhaps due to the model

spin-up effect. JMA and UKMO yield lower ratios of evaporation to precipitation than observed values, while NCEP produces excessively high ratios. The over-prediction of precipitation in NCEP is related to its over-prediction of evaporation. This strong coupling may imply that seasonal weather forecasting is potentially sensitive to the initialization of soil moisture, and that soil moisture may have a longer-term memory in the model than in reality. Because of this problem, the LSM in NCEP has been replaced with an LSM (Mitchell et al. 2005).

In terms of the surface radiation budget, *SWD* is generally over-predicted and *LWD* is generally under-predicted, particularly by ECPC and JMA. The diurnal peak of *LWD* is also delayed in the models. UKMO provides the best prediction of surface downward radiation. BMRC and NCEP provide good simulations of *LWD*, but yield poor results for *SWD*. The errors in the two downward components counteract each other, resulting in smaller errors in the total downward radiation than in the individual downward components. Errors in predicted radiation might reflect shortcomings in the radiation scheme and/or errors in cloudiness. Cloudiness data are required to confirm the cause of the radiation discrepancies, but these data were not collected during CEOP Phase 1. This lack of cloudiness data represents a major gap in the CEOP data, and the next CEOP phase must consider the current gaps in in situ observational data.

The surface energy budget is poorly predicted by all of the models. JMA and UKMO show the best skill in modeling the surface energy budget, while NCEP under-predicts sensible heat fluxes and over-predicts latent heat fluxes, and ECPC over-predicts the sensible heat flux but predicts latent heat fluxes relatively well. BMRC uses a simple bucket hydrological model without explicit vegetation; consequently, it yields unrealistic surface energy budgets for certain sites during certain seasons. We identified a number of biases that are common to all of the GCMs: over-prediction of the ranges in seasonal variations of all of the flux components, late daytime peaking time of latent heat fluxes, over-prediction of downward sensible heat fluxes during the winter season and at nighttime, and under-prediction of heat trans-

fer resistance. The latter bias results in under-prediction of the diurnal range of surface–air temperature difference. It is therefore important to strengthen theoretical and experimental studies of heat transfer in stable boundary layers, particularly the setting of momentum roughness lengths and thermal roughness lengths in GCM grids.

Inter-comparisons among three offline GLDAS LSMs revealed large uncertainties in surface energy budgets among the models. Therefore, it remains crucial to strengthen studies of parameterizations of land processes and calibrations of model parameters. In addition, Yang et al. (this issue) demonstrated that LSM errors could be effectively reduced by assimilating soil moisture-relevant satellite data.

In summary, this study is our first step in evaluating the prediction skill of CEOP-participating models; as such, it provides clues to the identification of additional model deficiencies. The next step will focus on inter-comparisons and improvements of physical schemes, for which data analysis groups must collaborate with NWP centers. It is crucial to develop an inter-comparison platform, establish the priorities of target processes, provide offline source code for each scheme and auxiliary input data, collect benchmark data sets, and implement comparisons. As satellite data are also being archived in CEOP, these data can be used to upscale in situ observations to GCM grids, leading to a strong potential for use in future inter-comparison studies.

Acknowledgements

This study was funded by the Japan Science and Technology Agency and the Special Coordination Funds for Promoting Science and Technology. The authors thank Dr. H.-J. Isemer for his constructive comments. The authors are also grateful to Dr. L. Rikus (BMRC); Dr. M. Hirai, Dr. H. Kitagawa, and M. Sakamoto (JMA); Dr. K. Mitchell and S. Katz (NCEP); Dr. P. Earnshaw and Dr. S. Milton (UKMO); Dr. J. Roads and Dr. A. Ruane (ECPC); Dr. M. Rodell and Dr. H. Kato (GFSC/NASA); Dr. H. Luthardt and Dr. H. Thiemann (MPI); principal investigators and researchers involved in CEOP in situ observations (see Table 2), and data archiving centers for their data collections, the provision and archiving of data, dis-

cussions, and information related to the updated status of relevant models.

References

- Bechtold, P., J.-P. Chaboureaud, A. Beljaars, A.K. Betts, M. Miller, M. Köhler, and J.-L. Redelsperger, 2004: The simulation of the diurnal cycle of convective precipitation over land in a global model, *Q. J. R. Meteorol. Soc.*, **130**, 3119–3137.
- Betts, A.K. and J.H. Ball, 1995: The FIFE surface diurnal cycle climate, *J. Geophys. Res.*, **100**(D12), 25679–25693.
- Beyrich, F. and W. Adam, 2004: A note on the use of CEOP reference site data for comparison with the output of global models: The Lindenberg example, *CEOP Newsletter*, **6**, 6–7.
- Bollasina, M., L. Bertolani, and G. Tartari, 2002: Meteorological observations in the Khumbu Valley, Nepal Himalayas, 1994–1999, *Bull. Glac. Res.*, **19**, 1–11.
- Bosilovich, M.G., 2002: On the use and validation of mosaic heterogeneity in atmospheric numerical models, *Geophys. Res. Lett.*, **29**, 10.1029/2001GL013925.
- Cess, R.D., M.H. Zhang, P. Minnis, and co-authors, 1995: Absorption of solar radiation by clouds: observations versus models, *Science*, **267**(5197), 496–499.
- Chen, F., K. Mitchell, J. Schaake, Y. Xue, H.-L. Pan, V. Koren, Q.-Y. Duan, M. Ek, and A. Betts, 1996: Modeling of land-surface evaporation by four schemes and comparison with FIFE observations, *J. Geophys. Res.*, **101**(D3), 7251–7268.
- Chen, T.-C., M.-C. Yen, J.-C. Hsieh, and R.W. Arritt, 1999: Diurnal and seasonal variations of the rainfall measured by the Automatic Rainfall and Meteorological Telemetry System in Taiwan, *Bull. Amer. Meteor. Soc.*, **80**, 2299–2312.
- Dai, A., 2001: Global precipitation and thunderstorm frequencies. Part II: Diurnal variations. *J. Climate*, **14**, 1112–1128.
- Dai, Y., X. Zeng, R.E. Dickinson, I. Baker, G.B. Bonan, M.G. Bosilovich, A.S. Denning, P.A. Dirmeyer, P.R. Houser, G.-Y. Niu, K.W. Oleson, C.A. Schlosser, and Z.-L. Yang, 2003: The Common Land Model (CLM), *Bull. Amer. Meteor. Soc.*, **84**, 1013–1023.
- Edwards, J.M. and A. Slingo, 1996: Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model, *Q. J. R. Meteorol. Soc.*, **122**, 689–719.
- Garratt, J.R. and A.J. Prata, 1996: Downwelling longwave fluxes at continental surfaces—A comparison of observations with GCM simulations and implications for the global land-

- surface radiation budget, *J. Climate*, **9**, 646–655.
- Gates, W.L. and co-authors, 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP), *Bull. Amer. Meteor. Soc.*, **80**, 29–55.
- Henderson-Sellers, A., P. Irannejad, K. McGuffie, and A.J. Pitman, 2003: Predicting land surface climates-better skill or moving targets? *Geophys. Res. Lett.*, **30**(14), 1777, doi:10.1029/2003GL017387.
- Hirai, M. and T. Matsumura, 2004: A validation study of a new land surface scheme—with the CEOP EOP-1 reference site data set, *CEOP Newsletter*, **6**, 3–5.
- Hirose, M. and K. Nakamura, 2005: Spatial and diurnal variation of precipitation systems over Asia observed by the TRMM Precipitation Radar, *J. Geo. Res.*, **110**, D05106.
- Isemer, H.J., 2002: Report on European CEOP Reference Sites, *CEOP Newsletter*, **2**, 6.
- Ishikawa, H. and GAME-Tibet Boundary Layer Group, 2001: What has been known and what has not in GAME/Tibet BL observation, *Proceedings of the Fifth International Study Conference on GEWEX in Asia and GAME*, 691.
- Kaihotsu, I., T. Yamanaka, D. Oyunbaatar, T. Ganbold, G. Davaa, and T. Koike, 2003: Long-term monitoring by water cycle stations in the central part of the Mongolian Plateau, *Proceed. 1st Inter. Conf. APHW*, **1**, 117–120.
- Kato, H., M. Rodell, F. Beyrich, H. Cleugh, E. van Gorsel, H. Liu, and T.P. Meyers, 2007: Sensitivity of land surface simulations to model physics, land characteristics, and forcings, at four CEOP sites, *J. Meteor. Soc. Jpn*, this issue.
- Kim, J., W. Kim, C. Cho, B. Choi, H. Chung, B. Lee, K. Kim, M. Kim, B. Lee, D. Lee, G. Lee, J. Lee, J. Lim, J. Oh, E. Park, J. Shim, J. Yun, and C. Rho, 2002: KOFLUX: A new tool to study the biosphere-atmosphere interactions in Asia, in *Ecology of Korea* edited by D. Lee, 215–229.
- Kishtawal, C.M. and T.N. Krishnamurti, 2001: Diurnal variation of summer rainfall over Taiwan and its detection using TRMM observations, *J. Appl. Meteor.*, **40**, 331–344.
- Koike, T., 2002: CEOP Starts—A step for predictability improvement of the water cycle and water resources, *CEOP Newsletter*, **1**, 1.
- Koike, T., 2004: The Coordinated Enhanced Observing Period—an initial step for integrated global water cycle observation, *WMO Bulletin*, **53**(2), 1–8.
- Koike, T., Y. Nakamura, I. Kaihotsu, G. Davaa, and N. Matsuura, 2003: AMSR-E soil moisture product validated at the CEOP Mongolia reference site, *CEOP Newsletter*, **7**, 5.
- Koren, V., J. Schaake, K. Mitchell, Q.Y. Duan, F. Chen, and J.M. Baker, 1999: A parameterization of snowpack and frozen ground intended for NCEP weather and climate models, *J. Geophys. Res.*, **104**, 19,569–19,585.
- Koster, R.D. and M.J. Suarez, 1996: Energy and water balance calculations in the Mosaic LSM, *NASA Tech. Memo. 104606*, Vol. **9**, 76 pp.
- Kubota, H., R. Shiroyaka, T. Ushiyama, J. Chen, T. Chuda, S. Iwasaki, and K. Takeuchi, 2002: Seasonal variability observed at Palau over western Pacific and water vapor variation observed prior to the large-scale convective activity, *Western Pacific Geophysical Meeting Abstracts, AGU, WP8-WP9*.
- Lau, W.K.M., J. Matsumoto, M. Bollasina, and H. Berbery, 2004: Diurnal variability in the monsoon region: Preliminary results from the CEOP Inter-Monsoon Studies (CIMS), *CEOP Newsletter*, **5**, 2–4.
- Lin, X., D.A. Randall, and L.D. Fowler, 2000: Diurnal Variability of the Hydrologic Cycle and Radiative Fluxes: Comparisons between Observations and a GCM, *J. Climate*, **13**, 4159–4179.
- Liu, H., G. Tu, W. Dong, C. Fu, and L. Shi, 2006: The diurnal and seasonal variations of water vapor and CO₂ fluxes at a land surface in a semi-arid region, *Chinese J. Atmos. Sci.*, **30**(1), 108–118.
- Lu, C.-H. and K. Mitchell, 2004: Land surface processes simulated from the Noah LSM in the NCEP global model: A comparative study using the CEOP EOP-1 reference site observations, *CEOP Newsletter*, **5**, 5–6.
- Marengo, J.A. and L. Horta, 2003: CEOP reference sites in the LBA Continental Scale Experiment, *CEOP Newsletter*, **4**, 6.
- Milton, S. and P. Earnshaw, 2007: Evaluation of surface water and energy cycles in the Met Office Global NWP Model using CEOP data, *J. Meteor. Soc. Jpn*, this issue.
- Mitchell, K., H. Wei, S. Lu, G. Gayno, and J. Meng, 2005: NCEP implements major upgrade to its medium-range global forecast system, including land-surface component, *GEWEX Newsletter*, Nov. 2005, 8–9.
- Nesbitt, S.W. and E.J. Zisper, 2003: The diurnal cycle of rainfall and convective intensity according to three year TRMM measurements, *J. Climate*, **16**, 1456–1475.
- Ohata, T. and Y. Fukushima, 1999: Progress of GAME-Siberia 1997–98. *GAME Publication No. 14: Activity Report of GAME-Siberia*, 1998, 1–5.
- Ohta, T., T. Hiyama, H. Tanaka, T. Kuwada, T.C. Maximov, T. Ohata, and Y. Fukushima, 2001:

- Seasonal variation in the energy and water exchanges above and below a larch forest in Eastern Siberia, *Hydro. Process.*, **15**, 1459–1476.
- Phillips, T.J. and co-authors, 2004: Evaluating parameterizations in general circulation models, climate simulation meets weather prediction, *Bull. Amer. Meteor. Soc.*, **85**, DOI: 10.1175/BAMS-85-12-1903.
- Pitman, A.J. and PILPS team co-authors, 1999: Key results and implications from phase 1(c) of the Project for Intercomparison of Land-surface Parameterization Schemes, *Clim. Dynamics*, **15**, 673–684.
- Ramanathan, V., B. Subasilar, G.J. Zhang, W. Conant, R.D. Cess, J.T. Kiehl, H. Grassl, and L. Shi, 1995: Warm pool heat budget and short-wave cloud forcing: A missing physics? *Science*, **267**(5197), 499–503.
- Randall, D.A., Harshvardhan, and D.A. Dazlich, 1991: Diurnal variability of hydrological cycle in a general circulation model, *J. Atmos. Sci.*, **48**, 40–62.
- Rikus, L., 2007: A comparison of some surface variables in the BMRC MOLTS with CEOP in-situ data for EOP3, *J. Meteor. Soc. Jpn*, this issue.
- Rodell, M. and co-authors, 2004: The Global Land Data Assimilation System, *Bull. Amer. Meteor. Soc.*, **85**, 381–394.
- Ruane, A.C. and J.O. Roads, 2007: The diurnal cycle of water and energy over the Continental United States from three reanalyses, *J. Meteor. Soc. Jpn*, this issue.
- Siebesma, A.P., C. Jakob, G. Lenderink, R.A.J. Neggers, J. Teixeira, E. Van Meijgaard, J. Calvo, A. Chlond, H. Grenier, C. Jones, M. Köhler, H. Kitagawa, P. Marquet, A.P. Lock, F. Müller, D. Olmeda, and C. Severijns, 2004: Cloud representation in general-circulation models over the northern Pacific Ocean: A EUROCS inter-comparison study, *Q. J. R. Meteorol. Soc.*, **130**, 3245–3267.
- Soman, V.V., J.B. Valdes, and G.R. North, 1995: Satellite sampling and the diurnal cycle statistics of Darwin rainfall data, *J. Appl. Meteor.*, **34**, 2481–2490.
- Tamagawa, K., T. Koike, and S. Williams, 2003: First CEOP EOP-1 data comparison, *CEOP Newsletter*, **7**, 4.
- Tartari, G., G. Verza, and L. Bertolami, 1998: Meteorological data at the pyramid observatory laboratory. *Mem. Ist. Ital. Idrobiol.*, **57**, 23–40.
- Ueno, K. and co-authors, 2001: Meteorological observations during 1994–2000 at the Automatic Weather Station (GEN-AWS) in khumbu region, Nepal Himalayas, *Bull. Glacio. Res.*, **18**, 23–30.
- Wild, M., C.N. Long, and A. Ohmura, 2006: Evaluation of clear-sky solar fluxes in GCMs participating in AMIP and IPCC-AR4 from a surface perspective, *J. Geophys. Res.*, **111**(D01104), doi:10.1029/2005JD006118.
- Wild, M., A. Ohmura, H. Gilgen, J.-J. Morcrette, and A. Slingo, 2001: Evaluation of downward long-wave radiation in general circulation models, *J. Climate*, **15**, 3227–3239.
- Yang, G.-Y. and J. Slingo, 2001: The diurnal cycles in the Tropics, *Mon. Wea. Rev.*, **129**, 784–801.
- Yang, K., T. Koike, H. Fujii, K. Tamagawa, and N. Hirose, 2002: Improvement of surface flux with a turbulence-related length, *Q. J. R. Meteor. Soc.*, **128**, 2073–2087.
- Yang, K., T. Koike, K. Tamagawa, and P. Koudelova, 2005: Model uncertainties correlated with spatial scale of prognostic variables, *The 85th AMS Annual Meeting (CD-ROM)*, San Diego, USA, 9–13.
- Yang, K., T. Watanabe, T. Koike, X. Li, H. Fujii, K. Tamagawa, Y. Ma, and H. Ishikawa, 2007: Auto-calibration system developed to assimilate AMSR-E data into a land surface model for estimating soil moisture and the surface energy budget, *J. Meteor. Soc. Jpn*, this issue.