Evaluations of Land–Ocean Skin Temperatures of the ISCCP Satellite Retrievals and the NCEP and ERA Reanalyses

BEN-JEI TSUANG
Department of Environmental Engineering, National Chung Hsing University, Taichung, Taiwan

MING-DAH CHOU
Department of Atmospheric Sciences, National Taiwan University, Taipei, Taiwan

YUANCHONG ZHANG
Department of Applied Physics and Applied Mathematics, NASA Goddard Institute for Space Studies, Columbia University, New York, New York

ANDREAS ROESCH
Institute for Atmospheric and Climate Science, Zürich, Switzerland

KUN YANG*
Department of Civil Engineering, University of Tokyo, Tokyo, Japan

(Manuscript received 19 May 2006, in final form 25 January 2007)

ABSTRACT

This study evaluates the skin temperature (ST) datasets of the International Satellite Cloud Climatology Project (ISCCP) D satellite product, the ISCCP FD satellite product, the 40-yr ECMWF Re-Analysis (ERA-40), the NCEP–NCAR Reanalysis, and the NCEP–Department of Energy (DOE) Atmospheric Model Intercomparison Project (AMIP)-II Reanalysis. The monthly anomalies of all the datasets are correlated to each other and to most of the ground-truth stations with correlation coefficients >0.50. To evaluate their qualities, the 5 ST datasets are used to calculate clear-sky (CS) outgoing longwave radiation (OLR) and upward surface longwave radiation (USLR); the results are compared with the Earth Radiation Budget Experiment (ERBE) satellite observation and 14 surface stations. The satellite-derived STs and ERA-40 ST tend to bias high on hot deserts (e.g., Sahara Desert), and the reanalyzed STs tend to bias low in mountain areas (e.g., Tibet). In Northern Hemisphere high-latitude regions (tundra, wetlands, deciduous needle-leaf forests, and sea ice), the CS OLR anomalies calculated using the satellite-derived STs have higher correlations and lower root-mean-squared errors with the ERBE satellite observation than those derived from using the reanalyzed STs. ERA-40 underestimates the amplitude of the seasonal ST over glaciers. All the reanalysis products (ERA-40, NCEP–NCAR, and NCEP–DOE AMIP-II) overestimate the ST during partial sea ice–covered periods in the middle-high-latitude oceans. Nonetheless, suspected spurious noises with an amplitude of 2 K in the satellite-derived STs produce a physically unviable anomaly over earth's surface where the amplitude of the anomaly is weak (such as open-water bodies, croplands, rain forest, grasslands, hot deserts, and cold deserts). Better land–ocean–ice schemes for a reanalysis should be developed for desert regions, high plateaus, fractional sea ice–covered oceans, and seasonally snow-covered lands, where the largest ST errors are identified.

1. Introduction

Sea surface temperature (SST) anomalies, such as the El Niño, have been found to have profound impact on climate (e.g., Gates et al. 1999). On the other hand, the characteristics and features of land skin temperature (LST) anomalies have not been well explored (Prigent et al. 2003). The LST is an important parameter for determining the energy exchange between the atmosphere and the land (e.g., Brutsaert 1982). The accuracy in the LST of a climate model represents how well the land surface processes are parameterized in the model (e.g., Jin et al. 1997; Tsuang 2003).

Conventionally, LST data are only available from surface stations at a limited number of research sites (e.g., Betts and Ball 1998; Baldocchi et al. 2001). Nowa-

* Current affiliation: Insitute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China.

Corresponding author address: Ben-Jei Tsuang, Department of Environmental Engineering, National Chung Hsing University, Taichung 402, Taiwan.
E-mail: tsuang@nchu.edu.tw

DOI: 10.1175/2007JCLI1502.1

© 2008 American Meteorological Society

days, because of rapid progress in the development of remote sensing and reanalysis techniques, LST data with global coverage have been derived from satellite observations (Jin 2004; Zhang et al. 2004) or calculated using data assimilation systems (Simmons and Gibson 2000; Kanamitsu et al. 2002a). Without extensive evaluation of the data however, it is not clear how reliable these products are. This study attempts to conduct such an evaluation.

In the present study, we compare skin temperature (ST; including both LST and SST) datasets, especially their interannual variability of the International Satellite Cloud Climatology Project (ISCCP) D2 satellite product (ID; Rossow and Schiffer 1999), the ISCCP FD product (IFD; Zhang et al. 2004), the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) Reanalysis (NCEP1; Kalnay et al. 1996; Kistler et al. 2001), the NCEP–Department of Energy (DOE) Atmospheric Model Intercomparison Project (AMIP)-II Reanalysis (NCEP2; Kanamitsu et al. 2002a), and the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40; Simmons and Gibson 2000). ERA-40 is a newer version of the ECMWF Re-Analysis, which succeeds the prior ERA-15 (Gibson et al. 1997).

In the absence of routine in situ ST measurements, Prigent et al. (2003) evaluated STs by comparison with surface air temperature (SAT). The SAT data compiled by the Climate Research Unit (CRU, version CRU TS2.1; Mitchell and Jones 2005) is also used in this study for the comparison with ST data.

Since there is no global ground-truth dataset for ST, we evaluate the STs of the 5 products by using them to calculate clear-sky (CS) outgoing longwave radiation (OLR) at the top of the atmosphere and the upward surface longwave radiation at the land surface (USLR). The results are compared with observations from satellites and a few ground-truth stations. It is well understood that the contribution to CS OLR anomalies may be separated into two factors: surface properties and atmospheric properties (Zhang et al. 2006, 2007). The surface properties include ST and emissivity; the atmospheric properties include both temperature and humidity (T/q) profiles and near-SAT. These two factors are equally important in the calculation of CS OLR (See the appendix for details). To separate the effect of ST from the others for determining CS OLR, all the CS OLR calculations are calculated using the same emissivity dataset, the same T/q profile data and the same SAT data as inputs.

These comparisons provide information for understanding the strength and weakness of the satellite retrieval and the reanalysis skill on the global scale. In addition, since the reanalyzed ST of each land type is usually calculated with its own parameters (Simmons and Gibson 2000), the results of the comparisons are grouped for each land type, based on a modified International Geosphere-Biosphere Programme (IGBP) land classification (Fig. 1; after Loveland et al. 2001; Masson et al. 2003), of which the criteria are listed in Table 1. The practical relevance may lie in that the paper gives a better understanding of the ST error sources/uncertainties for both the satellite-derived and reanalyzed products, and therefore indicates possible future improvements.

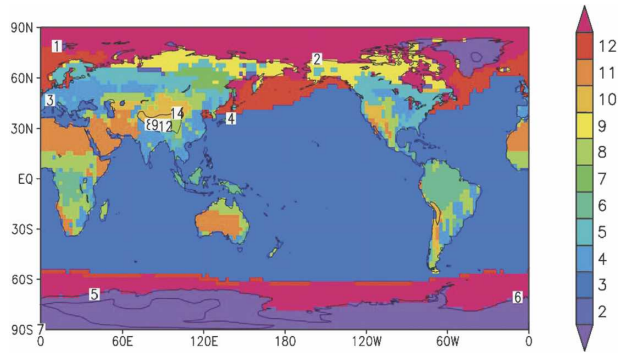

FIG. 1. Map of modified IGBP land types, where 1 = glacier, 2 = wetlands, 3 = open water, 4 = croplands, 5 = other forests, 6 = rain forest, 7 = deciduous needle-leaf forests, 8 = grasslands, 9 = tundra, 10 = cold desert, 11 = hot desert, 12 = ocean with ice >0% but ≤10%, and 13 = ocean with sea ice >10%. The numbers in the map denote the locations for the USLR stations as listed in Table 2. In addition, the elevation at 3000 m is shown in contours.

2. Data descriptions

The ID product (Brest et al. 1997; Rossow and Schiffer 1999) ran from July 1983 until recently. It is the monthly average of D1 3-hourly data on a global equal-area grid. It is calibrated as follows: 1) Prelaunch calibration of the radiometer to the temperature of a reference blackbody was conducted. 2) On the spacecraft, calibration of the infrared channels was carried out actively, once per scan, by having the radiometer view space and a standard blackbody with a known temperature. 3) Further calibration was then incorporated to eliminate the changes between different reference polar orbiters. As a result, the relative calibrations of the infrared radiances (IR) used by ISCCP are on average uncertain by no more than 1–2 K absolute, ±0.3%–1.0% relative, and the absolute calibration uncertainty is estimated to be about 2% for IR (Brest et al. 1997).

TABLE 1. Criteria and descriptions of the modified IGBP land classification.

Number	Land type	Description and original IGBP land type
1	Glacier	Glacier, snow, and ice.
2	Wetlands	Permanent wetlands.
3	Open water	Iced-free water bodies.
4	Croplands	Croplands, urban and built up, or cropland/natural vegetation mosaics.
5	Other forests	Evergreen needle-leaf forests, deciduous broadleaf forests, or mixed forests.
6	Rainforest	Evergreen broadleaf forests.
7	Deciduous needle-leaf forests	Deciduous needle-leaf forests.
8	Grasslands	Closed shrublands, woody savannahs, savannahs, or grasslands with latitude <52.5°N/52.5°S.
9	Tundra	Barren, open shrublands, closed shrublands, woody savannahs, savannahs, or grasslands with latitude ≥52.5°N/52.5°S.
10	Cold desert	Open shrublands, or barren with latitude <52.5°N/52.5°S and annual surface air temperature <288 K.
11	Hot desert	Open shrublands, or barren with latitude <52.5°N/52.5°S and annual surface air temperature ≥288 K.
12	Ocean with ice >0% but ≤10%	Water bodies with ice fraction >0% but ≤10%.
13	Ocean with ice >10%	Water bodies with ice fraction >10%.
14	Tibet	Tibet with elevation >3000 m asl.
15	Andes	Andes Mountains with elevation >3000 m asl.

ST is retrieved from the clear radiance values of narrowband IR (about 11 microns) by assuming a surface emissivity of 1. For the narrowband IR, atmospheric effects are small and depend on ozone and water abundances, the temperature profile, and aerosol optical thickness. Complete information on ozone, water, and temperature are obtained from correlative data. Since little information is available concerning aerosol properties, their effects are neglected.

The IFD ST (Zhang et al. 2004) was taken from the clear-sky composite values of the ID dataset with nonunit emissivity and diurnal cycle adjustments. As discussed by Rossow and Garder (1993), these values accurately represent the surface temperature under clear-sky conditions, although they are biased in different ways regionally and seasonally when used to represent LST for cloudy conditions (Prigent et al. 2003; Aires et al. 2004). Surface emissivity values have been derived for 33 k intervals based on combining field and laboratory measurements of the emissivities of various rock, soil, and vegetation types for land and the Fresnel reflection formula for open water (Zhang et al. 2007). The IFD calculations incorporate a diurnal adjustment scheme for ST over land areas only, but retain the original SST diurnal cycle amplitude from ID.

All the SSTs of the reanalyses (ERA-40, NCEP2, and NCEP1) are taken from Reynolds' SST (Reynolds and Smith 1994) with slight modification. On land, ERA-40 calculates the LST using a four-level self-contained soil parameterization scheme developed by Viterbo and Beljaars (1995) and Viterbo et al. (1999), in which a zero heat flux condition is set at the bottom as a boundary condition. Viterbo and Beljaars' (1995) scheme has been tested in a stand-alone mode with the help of several long observational time series from field experiments in the United States, at Cabauw in the Netherlands, and in the Amazonian rain forest in Brazil.

NCEP1 and NCEP2 calculate the LST according to a forced-restore rate equation (Pan and Mahrt 1987), in which the soil temperature of the deepest numerical layer is prescribed (Deardorff 1978; Kanamitsu 1989). In NCEP2, simple rainfall assimilation over land surfaces for improving soil wetness is adopted (Kanamitsu et al. 2002a). The assimilation system uses an observed 5-day mean "pentad" precipitation, based on a newly available NCEP–Climate Prediction Center (CPC) global precipitation analysis that merges satellite and gauge measurement on a 2.5° × 2.5° latitude–longitude grid. This retrospective global precipitation analysis is described further in Gruber et al. (2000). Using the assimilated pentad precipitation analysis prevents long-term climate drift of soil wetness in NCEP2.

We use the STs of various products to calculate CS OLRs. The STs are then evaluated by comparing the calculated CS OLRs with the Earth Radiation Budget Experiment (ERBE) observations (Barkstrom 1984; Green et al. 1990). The ERBE was designed to measure the earth's radiation fields. The ERBE radiometric packages were placed into orbit aboard three satellite platforms, including the Earth Radiation Budget satellite (ERBS), the *NOAA-9* satellite and the *NOAA-10* satellite. We use the mean value of the three satellite observations of CS OLR measured by the scanning radiometers for comparisons. ERBS covered the period

from November 1984 to February 1990, *NOAA-9* covered the period from February 1985 to January 1987, and *NOAA-10* covered the period from November 1986 to May 1989. Note that during the overlapping period of the three satellites (November 1986–January 1987), the CS OLR measured by different satellite can be different. On a monthly average, the uncertainties of the ERBE CS OLR are about 5 W m^{-2} (Barkstrom et al. 1989) with a higher uncertainty in the Antarctic.

In addition, the calculated USLRs are compared with observations at Baseline Surface Radiation Network (BSRN) stations (Ohmura et al. 1998) available from 1992 onward, Global Energy and Water Cycle Experiment (GEWEX) Asian Monsoon Experiment (GAME)-Tibet stations (Yang et al. 2006) in 1998, and Heihe river basin Field Experiment (HEIFE)-Gobi stations (Wang and Mitsuta 1991; Tamagawa 1996) in 1991 and 1992 (Table 2 and Fig. 1). The BSRN global network measures surface radiative fluxes at the highest possible accuracy with well-calibrated state-of-the-art instrumentation at selected sites in major climate zones. The accuracy of USLR of BSRN is 5% or 10 W m^{-2}.

3. Calculation of CS OLR

For evaluating the ST datasets, the IR radiation scheme of Chou et al. (2001) is used to calculate the OLR. The scheme includes the absorption due to major gaseous absorption (water vapor, CO_2, O_3) and most of the minor trace gases (N_2O, CH_4, CFCs), as well as clouds and aerosols. The thermal infrared spectrum is divided into nine bands. To achieve a high degree of accuracy and speed, various approaches of computing the transmission function are applied to different spectral bands and gases. The gaseous transmission function is computed by using the table lookup method. The scheme can accurately compute fluxes to within 1% of the high-spectral-resolution line-by-line calculations. The cooling rate can be accurately computed in the region extending from the surface to the 0.01-hPa level. This longwave radiation scheme has been implemented in the Goddard general circulation models (Sud and Mocko 1999; Bacmeister and Suarez 2002), a cloud ensemble model (Tao et al. 2002), and a version of the NCEP model (NCEP2000; Kanamitsu et al. 2002b).

Above the atmospheric boundary layer (ABL), the monthly T/q profile data of the ERA-40 are used. Near the surface, the 2-m T/q data of CRU are used. To have a consistent vertical profile, the T/q profiles within the ABL are linearly interpolated between the T/q of the ERA-40 at the top of the ABL and the 2-m T/q of CRU. The global surface emissivity is taken from Zhang et al. (2004). Monthly mean thermal infrared

TABLE 2. Characteristics of USLR stations used in this study.

ID	Site	Lat	Lon	Location	Elevation (m)	IGBP land type of the grid	Simple biosphere model land type of the grid	Data period	Source
1	Nya	78.93°N	11.95°E	Ny Alesund, Spitsbergen	11	Water body	Water body	1992–2001	BSRN
2	Bar	71.32°N	156.40°W	Barrow, Alaska	8	Water body	Water body	1993–2001	BSRN
3	Pay	46.82°N	6.93°E	Payerne, Switzerland	491	Croplands	Agriculture	1992–2001	BSRN
4	Tat	36.05°N	140.13°E	Tateno, Japan	25	Water body	Water body	1996–2001	BSRN
5	Syo	69.00°S	39.58°E	Syowa, Antarctic	18	Snow and ice	Glacier	1998–2001	BSRN
6	Gvn	70.65°S	8.25°W	Georg von Neumayer, Antarctic	42	Water body	Water body	1992–2001	BSRN
7	Spo	89.82°S	0°	South Pole, Antarctic	2800	Snow and ice	Glacier	1994–2001	BSRN
8	Sqh	32.5°N	80°E	Shiquanhe, Tibet	4282	Open shrublands	Broadleaf shrubs with bare soil	May–Sep 1998	GAME-Tibet
9	Gerze	32.3°N	84.5°E	Gerze, Tibet	4420	Open shrublands	Broadleaf shrubs with bare soil	May–Sep 1998	GAME-Tibet
10	Amdo	32.241°N	91.635°E	Amdo, Tibet	4700	Grasslands	Agriculture or C3 grassland	Jun–Aug 1998	GAME-Tibet
11	MS3478	31.85°N	91.65°E	MS3478, Tibet	4700	Grasslands	Agriculture or C3 grassland	May–Sep 1998	GAME-Tibet
12	Naqu	31.36°N	91.54°E	Naqu, Tibet	4496	Grasslands	Agriculture or C3 grassland	Jun–Aug 1998	GAME-Tibet
13	Desert	39.38°N	100.17°E	Station 65, Gansu, China	1391	Barren or sparsely vegetated	Bare soil	1991–1992	HEIFE
14	Gobi	39°N	100°E	Station 95, Gansu, China	1680	Barren or sparsely vegetated	Bare soil	1991	HEIFE

TABLE 3. Comparisons of the global monthly means and anomalies of the calculated CS OLRs with the ERBE observation from 1985–89, where Std, MAE, and Corr denote the standard deviation, the mean-absolute error, and the correlation coefficient of the monthly means, respectively. AStd, AMAE, and ACorr denote the standard deviation, the mean-absolute error, and the correlation coefficient of the monthly anomalies, respectively. The names "ID," "IFD," "ERA-40," "NCEP2," and "NCEP1" denote the ST of the dataset used as an input for the modeling system for determining CS OLR by this study. The name "org. ERA-40" denotes that the comparison is with the subset of the original CS OLR of the ERA-40 (Simmons and Gibson 2000). The name "I0" denotes that the run is simulated using the ST of invariant seasonal cycle of ID as input, which is determined to be the monthly composite of the ST of ID from 1985 to 2000.

Name	Number of grid month	Mean (W m^{-2})	Std (W m^{-2})	AStd (W m^{-2})	Bias (W m^{-2})	MAE (W m^{-2})	Corr	AMAE (W m^{-2})	ACorr
ERBE	518273	256	41.74	3.88					
ID	518273	259	42.34	3.88	2.82	9.01	0.98	3.42	0.48
IFD	518273	259	41.40	3.98	3.47	9.46	0.97	3.43	0.47
ERA-40	518273	259	40.66	3.30	3.60	9.44	0.97	3.47	0.45
NCEP2	518273	259	40.43	3.56	3.00	8.49	0.98	3.46	0.45
NCEP1	518273	259	41.30	3.49	2.81	9.11	0.98	3.48	0.44
org. ERA-40	518273	255	39.96	3.40	−0.58	8.71	0.98	3.44	0.46
I0	518273	258	41.85	2.76	2.31	8.76	0.98	3.55	0.40

radiative fluxes are calculated using the five ST datasets for the period September 1983–December 2000 globally with a 2.5 × 2.5 latitude–longitude spatial resolution. All the runs are calculated using the same spatial and temporal resolutions. All the derived radiative fluxes are in monthly intervals from September 1983 to December 2000 globally at 2.5° resolution. Vertical resolutions of the calculated fluxes are the same as ERA-40. They have 24 levels: at 1000, 925, 850, 775, 700, 600, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, 10, 7, 5, 3, 2, and 1 hPa, and at surface.

The anomaly of CS OLR is due to 1) the anomaly of ST and 2) the anomalies of the T/q profiles of the atmosphere. To separate the effects of the simulated CS OLR anomaly caused by the anomaly of ST from those of T/q profiles of the atmosphere, an additional run is conducted. The run is simulated using ST with invariant seasonal cycle as an input, which is determined to be the monthly composite of the ST of ID from 1985 to 2000 (denoted as I0).

Nonetheless, the modeling system used in this study may contain errors. The errors can be caused by the discrepancy in surface emissivity, in T/q profiles, in SAT, in the radiative code, or in the time resolution of inputs. Note that we are not able to generate a product with time scales shorter than a month since one of the input datasets, CRU dataset, is in monthly intervals. In addition, CS values between ERBE and modeling are different. The CS values of the modeling refer to values that are computed whether or not clear sky is actually detected. The CS values of ERBE are derived from measurements that occurred when clear sky was detected.

Table 3 compares the CS OLR calculated in this study using the ST of ERA-40 as an input (denoted as ERA-40) with the original CS OLR of ERA-40 (denoted as org. ERA-40; Simmons and Gibson 2000). It can be seen that the global mean CS OLR calculated using the modeling system of this study is 259 W m^{-2}, while that of the original ERA-40 dataset is 255 W m^{-2}. That is, the modeling system of this study produces higher CS OLR than the original ERA-40 dataset by 4 W m^{-2}. Nonetheless, their correlation coefficients of the monthly means and monthly anomalies are as high as 0.99 and 0.90 (Table 4), respectively. This implies that the modeling system used here may bias high, but it should be able to reproduce the seasonal and interannual variabilities of CS OLR.

Furthermore, calibration of the modeling system is conducted over open-water oceans. The Reynolds' SST over open-water oceans is based on both satellite estimates of SST and on direct observations of SST. It should therefore be more reliable than the STs observed elsewhere although a more accurate, diurnally resolved SST dataset is a current focus of research (Webster et al. 1996; Curry et al. 2004; Tu and Tsuang 2005). Figure 2 shows the bias of the simulated CS OLR from the ERBE observations for open-water ocean grids during 1986–89. The CS OLR is simulated using the ERA-40 SST as input. Note again that the SST of ERA-40 is taken from Reynolds' SST. It shows that the bias of the simulated CS OLR is 3.74 W m^{-2} with a standard deviation of 3.46 W m^{-2} over the open-water oceans.

4. Mean states

The spatial distributions of the mean STs of ID, IFD, ERA-40, NCEP2, and NCEP1 and the mean of the SAT of CRU are close to each other. Figure 3 shows the differences of the STs of ERA-40 and NCEP2, and

TABLE 4. Correlation of simulated CS OLR from 1985 to 1989 globally.

a. Monthly mean correlation

	ERBE	ID	IFD	ERA-40	NCEP2	NCEP1	org. ERA-40	I0
ERBE	1.00	0.98	0.97	0.97	0.98	0.98	0.98	0.98
ID		1.00	1.00	0.99	0.99	0.99	0.99	1.00
IFD			1.00	0.99	0.99	0.99	0.99	1.00
ERA-40				1.00	0.99	0.99	0.99	0.99
NCEP2					1.00	1.00	1.00	0.99
NCEP1						1.00	1.00	0.99
org. ERA-40							1.00	0.99
I0								1.00

b. Monthly anomaly correlation

	ERBE	ID	IFD	ERA-40	NCEP2	NCEP1	org. ERA-40	I0
ERBE	1.00	0.48	0.47	0.45	0.45	0.44	0.46	0.40
ID		1.00	0.97	0.86	0.85	0.85	0.83	0.82
IFD			1.00	0.84	0.84	0.84	0.81	0.81
ERA-40				1.00	0.92	0.93	0.90	0.87
NCEP2					1.00	0.97	0.91	0.82
NCEP1						1.00	0.92	0.83
org. ERA-40							1.00	0.78
I0								1.00

the SAT of CRU compared with the ST of ID from 1985 to 2000.

Table 5 shows the differences for the areal mean of each of the modified IGBP land types (Fig. 1) for the period 1985–89, where the grids with the standard error of ERBE CS OLR >5 W m^{-2} are discarded. Both Fig. 3 and Table 5 show that ID has the highest temperature among those datasets over desert regions and Tibet, the lowest temperature over tropical rain forest, and the lowest SST on middle-high-latitude oceans (40°–70°S and 48°–56°N). The largest differences between the ID ST and the reanalysis products (ERA-40, NCEP2, and NCEP1) are found over cold deserts, hot deserts, Tibet, and the Andes Mountains. Their differences are about −7 K. In contrast, the magnitudes of differences are <3 K for the other land types. Table 6 shows that the 5 products are all close to each other with monthly correlation >0.97. The two satellite products are closer to each other than to the reanalyses. The three reanalyses are closer to each other than to the satellite products. In addition, higher correlations of monthly STs are found in the extratropics (>0.8) than in the tropics among all the ST products for both land and ocean (not shown).

Table 3 compares the calculated CS OLRs using the STs of ID, IFD, ERA-40, NCEP2, NCEP1, and I0 as inputs, with the ERBE observation. It shows that, using the NCEP2 ST as an input, the calculated monthly

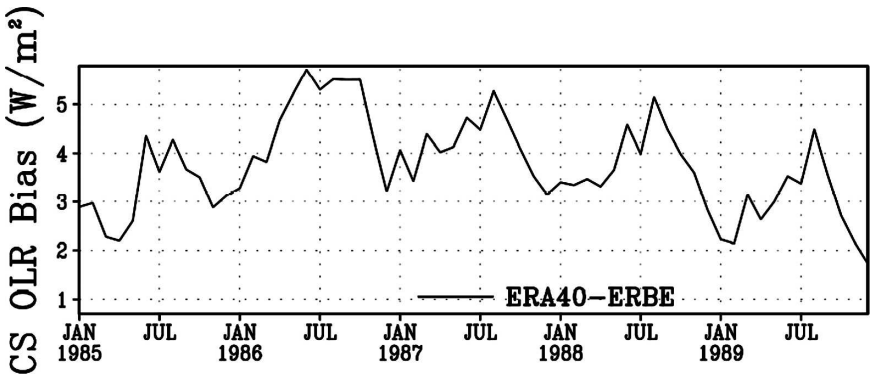

FIG. 2. Bias of the simulated CS OLR using ERA-40 SST comparing with the ERBE observations for open ocean during 1986–89.

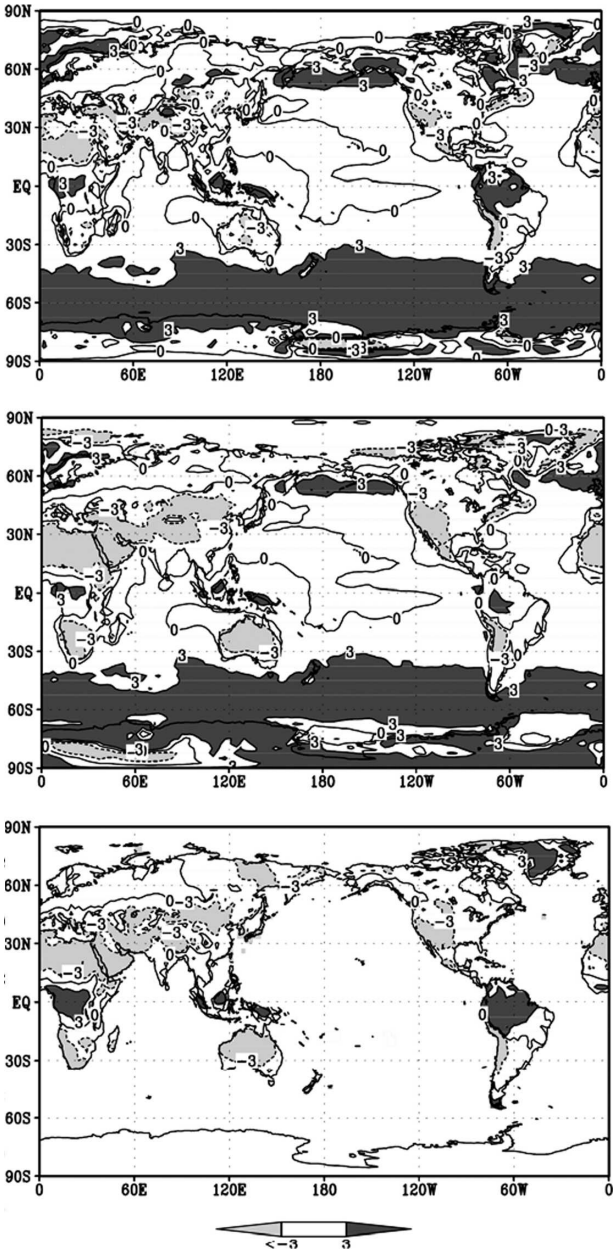

FIG. 3. Differences (K) of the skin temperatures of (top) ERA-40 and (middle) NCEP2, and (bottom) the SAT of CRU compared with the skin temperature of ID during 1985–2000.

value of CS OLR has the lowest mean-absolute error (MAE) and the highest correlation coefficient (Corr) among the calculated CS OLRs. Table 4 shows that all the simulated CS OLRs of ID, IFD, ERA-40, NCEP2, NCEP1, and I0 are close to each other with monthly mean correlation >0.99. The two satellite products are closer to each other than to the reanalyses. The three reanalyses are closer to each other than to the satellite products. However, the I0 run shows that atmospheric properties alone can obtain a monthly mean correlation of 0.98 for CS OLR simulation when compared with the ERBE satellite observation.

Figure 4 shows the spatial distribution of the biases of the calculated CS OLRs of ID, ERA-40, and NCEP2 compared with the ERBE observation, where the regions with bias larger than three standard deviation are shaded (i.e., >13 W m^{-2} or <−5 W m^{-2}). Note that from the previous assessment from the open ocean, we find that the modeling system systematically biases high by 4 W m^{-2} with the standard deviation at 3 W m^{-2}. Table 7 lists the biases for the areal mean of each of the modified IGBP land types. Note that the spatial pattern of IFD is close to ID, and that of NCEP1 is close to NCEP2. The largest bias compared with ERBE is found for the satellite products (ID, IFD) and ERA-40 over hot deserts. It can be seen that using the STs of ID, IFD, and ERA-40 as inputs, CS OLR is overestimated by >13 W m^{-2}, which is much larger than the uncertainty of the ERBE observation at 5 W m^{-2} and the uncertainty of the CS OLR simulation by this study (4 ± 3 W m^{-2}). In contrast, the reanalyses (ERA-40, NCEP2, NCEP1) bias is low in the Tibetan Plateau and the Andes Mountains. Using the STs of all the reanalyses (ERA-40, NCEP2, and NCEP1) underestimates CS OLR over Tibet and the Andes Mountains by >10 W m^{-2}. The same underestimation over Tibet in NCEP1 has been documented in Yang et al. (1999). It is likely due to a low bias in the STs of the reanalyses. If the modeling system used in this study systematically biases high by only 4 W m^{-2} in CS OLR calculation, it can be inferred that the STs of the reanalyses are too cold in Tibet and the Andes Mountains, and that the STs of all the satellite products and ERA-40 are too hot in hot deserts, including Saharan and Australian Deserts. Nonetheless, the above statements are not conclusive since there are potential errors in the modeling system. In the following, further data are used to clarify the above statements.

a. Middle-high-latitude ocean

Figure 5 shows that in the middle-high-latitude (48°–56°N) Northern Hemisphere ocean, the simulated CS OLRs of ERA-40 and NCEP2 bias high over the sea ice grids. This infers that in iced periods the SSTs of ERA-40 and NCEP2 over this ocean area are too high. Cavalieri et al. (1996) found, by using satellite observation, that over this region the ocean was occasionally iced. The zonal mean of the ice fraction ranged from 1% to 5%. However, the reanalyses products (ERA-40, NCEP2) indicate the SST being above 3°C even during the iced period, which contradicts the nonzero ice fraction. On the other hand, the satellite-derived products

TABLE 5. Differences (K) of the skin temperatures of IFD, ERA-40, NCEP2, and NCEP1 and the surface air temperature of CRU (CRU_SAT) compared with the skin temperature of ID for the areal mean of each of the modified IGBP land types for the period 1985–89, where the grids with the standard error of ERBE CS OLR > 5 W m^{-2} are discarded.

Land type	Grid number	Mean (K)	Bias (K)				
		ID	IFD	ERA-40	NCEP2	NCEP1	CRU_SAT
Glacier	109	253	2	0	1	0	2
Wetlands	18	267	2	0	−1	−1	−1
Open water	4467	294	0	0	0	0	—
Croplands	398	289	0	−1	−1	−2	−1
Other forests	371	277	0	0	−1	−1	−1
Rainforest	203	294	0	3	2	2	3
Deciduous needle-leaf forests	44	267	1	1	0	−1	−2
Grasslands	344	295	0	−2	−3	−3	−2
Tundra	294	265	2	0	−1	−2	−2
Cold desert	98	285	0	−3	−5	−6	−4
Hot desert	345	301	−1	−3	−7	−6	−5
Ocean with ice >0% but ≤10%	348	278	1	2	1	1	—
Ocean with ice >10%	1209	262	1	1	−1	−2	—
Tibet	9	277	−1	−4	−7	−9	−4
Andes Mountains	2	290	1	−7	−6	−7	−7

(ID, IFD) show that the ST during the iced period was about 0°C, which is reasonable. The same higher-than-0°C error in the reanalyses has been found in the northern part of the Caspian Sea at 46°N during its ice period (Tsuang et al. 2001). As a consequence, the simulated CS OLRs of ID and IFD are closer to ERBE than those of ERA-40 and NCEP2. Note again that both SSTs of ERA-40 and NCEP2 are taken from Reynolds' SST (Reynolds and Smith 1994). Reynolds' SST is derived by combining ship observations and satellite observations. The error in the reanalyses is likely due to ships avoiding cruising the iced fraction of an ocean grid, therefore the temperature at 0°C was not measured. Errors likely occur also in the middle-high-latitude (40°–70°S) Southern Hemisphere ocean, where the simulated CS OLRs of ERA-40 and NCEP2 bias high as well.

b. *Glacier*

Over glacier regions, the seasonal range of the ST of ERA-40 appears to be too weak compared with those of the STs of ID, IFD, and NCEP2 (Fig. 6). As a consequence, the simulated range of CS OLR of ERA-40 is smaller than that observed from ERBE. Therefore, the seasonal range of the ST of ERA-40 is likely to be

TABLE 6. Correlation of skin temperature from 1985 to 2000, where the correlations with CRU_SAT are only for land grids (Antarctica excluded) and the correlations among the others are global.

a. Monthly mean correlation

	ID	IFD	ERA-40	NCEP2	NCEP1	CRU_SAT
ID	1.00	0.99	0.97	0.97	0.98	0.98
IFD		1.00	0.97	0.97	0.97	0.97
ERA-40			1.00	0.98	0.98	0.98
NCEP2				1.00	1.00	0.98
NCEP1					1.00	0.98
CRU_SAT						1.00

b. Monthly anomaly correlation

	ID	IFD	ERA-40	NCEP2	NCEP1	CRU_SAT
ID	1.00	0.93	0.54	0.59	0.56	0.62
IFD		1.00	0.50	0.55	0.53	0.58
ERA-40			1.00	0.69	0.67	0.72
NCEP2				1.00	0.91	0.78
NCEP1					1.00	0.76
CRU_SAT						1.00

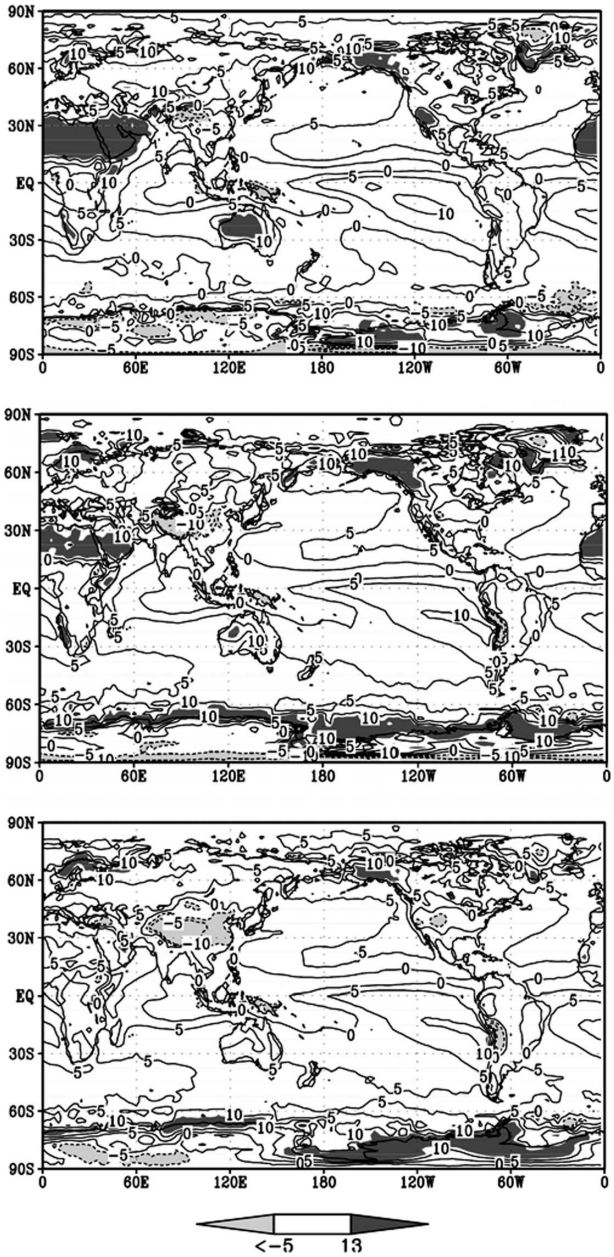

FIG. 4. Bias (W m^{-2}) of the simulated CS OLRs compared with ERBE during 1985–89, where regions with bias >3 std dev are shaded (i.e., >13 or <−5 W m^{-2}). Note that the CS OLR modeling system of this study systematically biases high by 4 W m^{-2} with the std dev 3 W m^{-2}: (top) ID − ERBE, (middle) ERA40 − ERBE, and (bottom) NCEP2 − ERBE.

underestimated. The same argument applies to the simulated USLR at the South Pole station as shown in Fig. 7. The figure shows that when using the ERA-40 ST as an input, the amplitude of the simulated seasonal variation of the USLR is too weak. In addition, according to the CS OLR simulation (Fig. 6) and the USLR at the South Pole station (Fig. 7), it can be seen that the STs of all the reanalyses are overestimated during the winter season of the corresponding hemisphere.

c. *Coastal stations*

For the coastal BSRN stations (Ny Alesund, Barrow, Tateno, Syowa, and Georg von Neumayer), Fig. 7 shows that the seasonal ranges of the simulated USLRs are generally smaller than the observations. This is expected since the spatial resolutions of the simulations of this study are as large as 2.5°, and the grid of a coastal station usually contains ocean fraction. Note that the seasonal range of SST is usually smaller than that of LST at the same grid. Therefore, the simulated seasonal ranges of USLRs are usually smaller than the coastal BSRN observations on land.

d. *Tibet and Gobi Desert*

Table 7 and Fig. 4 show that in Tibet the simulated CS OLRs of the reanalyses (ERA-40, NCEP2, NCEP1) bias low especially in summer season. This infers that the LSTs of the reanalyses over the Tibet area are too cold in summer. In addition, Fig. 8 shows that, from May to September 1998, at 5 GAME-Tibet stations (SQH, Gerze, Naqu, Amdo, MS3478) the simulated USLRs using the STs of ID, IFD, ERA-40, NCEP2, and NCEP1 as inputs are much lower than observations. This further confirms that the LSTs of the reanalyses over the Tibet area are too cold in summer. Besides, Yang et al. (2006) shows that the ST of IFD biases low in summer as well. Note that the ST of IFD is higher than those of the reanalyses.

Figure 4 shows that, in the Gobi Desert, the simulated CS OLRs of ID and ERA-40 bias high. This infers that the LSTs of ERA-40 and ID over the Gobi area are too hot. In addition, Fig. 8 shows that at 2 HEIFE Gobi stations (Desert, Gobi), the simulated USLRs using the ST of ERA-40 as an input is much higher than observations. This further confirms that the LST of ERA-40 over the Gobi Desert is too hot.

5. Interannual variability

The standard deviations (SD) (or the amplitudes) of the ST monthly anomalies and the SAT monthly anomalies from 1985 to 2000 are shown in Fig. 9. The datasets shown in the figure include the ST anomalies of ID, IFD, ERA-40, NCEP2, and NCEP1, and the SAT anomalies of CRU. Note that the SD of IFD is similar to ID, and the SD of NCEP1 is similar to NCEP2. Monthly anomaly is the value departing from its long-term mean of the month in multiple years, 1985–2000 in this case. The "standard deviation of anomalies" is defined as the root-mean-squared value of the deviation from the mean of anomalies. Note that

TABLE 7. Same as Table 5, but for the biases of the simulated CS OLRs (W m^{-2}) of ID, IFD, ERA-40, NCEP2, and NCEP1 runs, and the original ERA-40 dataset (Simmons and Gibson 2000) compared with ERBE.

Land type	Grid number	ID	IFD	ERA-40	NCEP2	NCEP1	org. ERA-40
Glacier	59	6.8	9.9	8.3	8.7	7.1	3.8
Wetlands	16	7.5	10.9	8.1	6.5	6.4	2.8
Open water	4461	3.0	3.3	3.7	3.6	3.7	−0.6
Croplands	381	3.5	3.6	2.2	1.2	0.8	−3.2
Other forests	333	5.8	6.8	5.7	4.5	3.7	0.7
Rainforest	198	0.3	0.4	1.9	1.5	1.1	−1.6
Deciduous needle-leaf forests	39	4.4	6.4	5.5	3.6	2.6	−0.2
Grasslands	336	5.3	5.8	2.6	0.8	0.6	−2.6
Tundra	210	8.5	12.2	9.2	6.9	6.3	3.7
Cold desert	92	8.0	8.1	2.8	−1.3	−2.3	−3.8
Hot desert	345	17.4	15.9	11.4	5.8	7.5	3.8
Ocean with ice >0% but ≤10%	310	3.1	4.0	5.9	5.3	5.4	1.2
Ocean with ice >10%	637	6.3	7.5	8.0	4.6	4.1	3.5
Tibet	8	2.4	0.4	−10.2	−13.9	−17.8	−16.1
Andes Mountains	2	−0.1	2.8	−13.6	−11.3	−13.1	−13.0

the mean of anomalies is 0. All the datasets show that the largest SD with magnitude >2 K occurs in the high-latitude regions. Over tropical and subtropical oceans, the SDs of ID and IFD are higher than the other datasets. In these areas, the SDs of ID and IFD are about 1–2 K, but those of the other datasets are <0.5 K. Since the accuracies of ID and IFD are about 1–2 K (Brest et al. 1997), it is suspected that the anomalies of ST shown in ID and IFD on tropical and subtropical oceans are partly produced because of the inaccuracy of ID and IFD. However, locations with the SD >2 K retrieved in ID and IFD may contain true signals. On glaciers (Greenland and Antarctica), the SDs (or the amplitudes) of the ST monthly anomalies shown in ERA-40 are much weaker than those shown in ID and NCEP2. On glaciers, the SDs in ID and NCEP2 are about 2.5 K, but that in ERA-40 is about 1.5 K.

Over ocean, the time series of the zonal means of the anomalies of the SSTs of ID and ERA-40 from 1985 to 2000 are shown in Fig. 10. Note that the zonal mean of

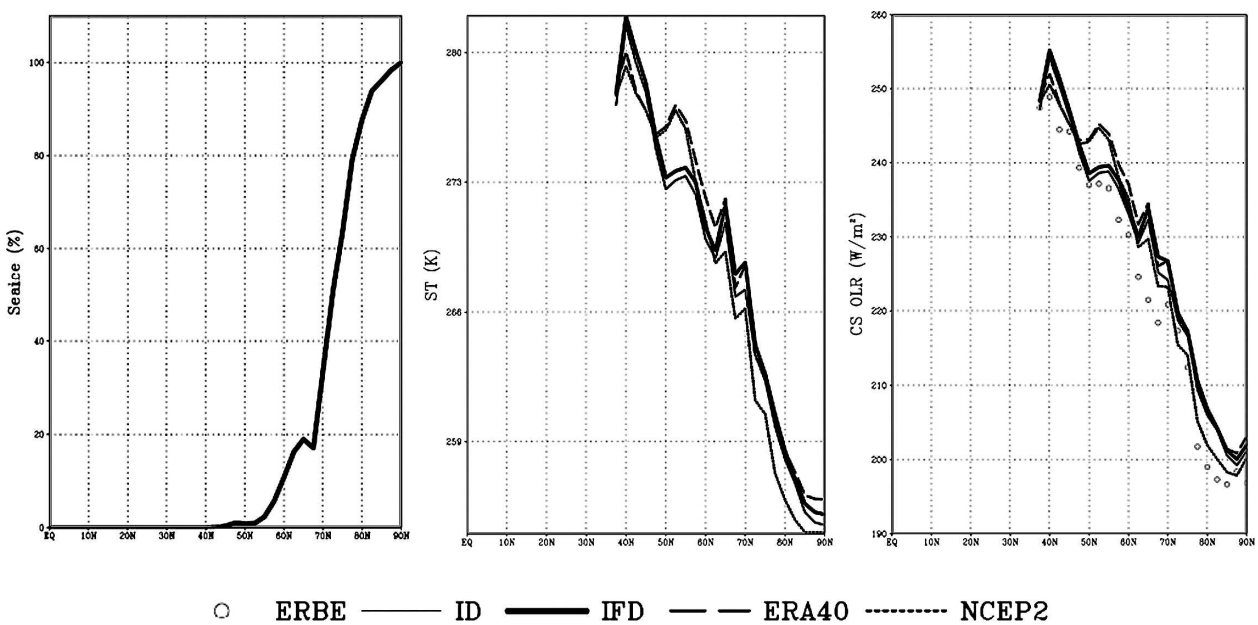

FIG. 5. Zonal mean of (left) sea ice fraction of ocean grids, and the corresponding (middle) SST and (right) CS OLR of the iced grids during 1986–89.

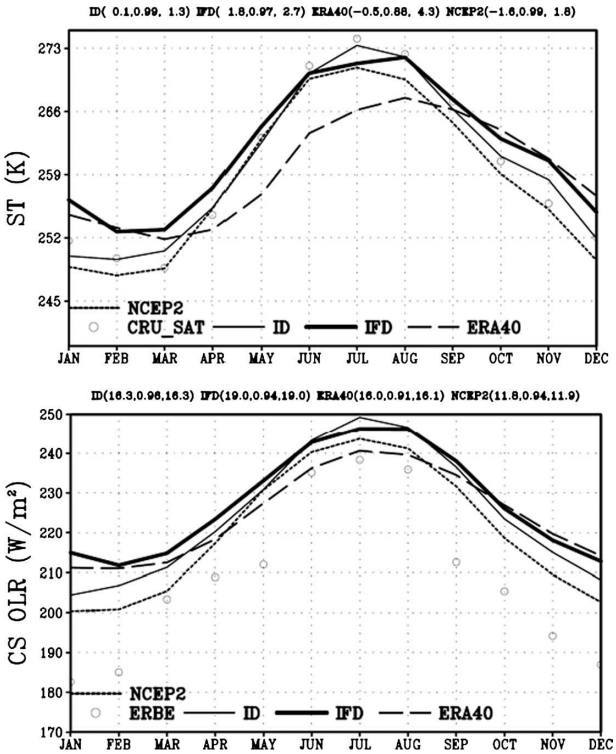

FIG. 6. (top) LST and (bottom) CS OLR for the Northern Hemisphere glaciers during 1985–89. Numbers in the parentheses denote bias (K), correlation coefficient, and MAE (K) compared with the ERBE observations.

IFD is close to ID, and the zonal means of NCEP2 and NCEP1 are close to ERA-40. The El Niño events in 1987, 1993, 1994, and 1997 can be clearly detected in the ERA-40 dataset. In addition, the figure shows that the SST anomalies of ID are much stronger than those of ERA-40, especially for latitude <50° and for dates after May 1998.

Over land, the time series of the zonal means of the anomalies of the LSTs of ID, ERA-40 and NCEP2 and the SAT of CRU from 1985 to 2000 are shown in Fig. 11. Note that the zonal mean of IFD is close to ID. It can be seen that the anomalies of CRU, ERA-40, and NCEP2 are similar to each other. Simmons et al. (2004) also found that ERA-40 and CRU have similar trends and low-frequency variability of SAT from 1979 onward, especially for the Northern Hemisphere. From the figure, it can be seen that the pattern between NCEP2 and CRU is even more similar. The anomaly correlations of the STs of ID, ERA-40, NCEP2, and NCEP1 with the SAT of CRU are 0.62, 0.72, 0.78, and 0.76, respectively (Table 6b). That is, the anomaly of CRU SAT is closest to NCEP2. But CRU, ERA-40, NCEP2, and NCEP1 are slightly different from the anomaly of ID LST. The ID satellite observation has much stronger anomalies than the other 3 datasets, especially in the subtropics and the tropics. According to ERA-40, NCEP2, NCEP1, and CRU datasets, in the tropics, there was a hot event, that is, of amplitude >0.5 K, occurring in 1998; and in the high-latitude regions hot events occur almost every other year.

Table 6b shows that the monthly anomalies of the STs of all the products (ID, IFD, ERA-40, NCEP2, NCEP1) are close to each other, with correlations >0.50 (so-called anomaly correlation). The two satellite products are closer to each other than to the reanalyses. The three reanalyses are closer to each other than to the satellite products. Between the satellite-retrieved STs and the reanalyzed STs, the anomaly correlations over land and eastern tropical oceans are mostly positive. Nonetheless, at the equator in the western Pacific Ocean (the warm pool) and the western Atlantic Ocean, their anomaly correlations are very low and become negative at a few grids, where the inaccuracy of ID and IFD may cause the low anomaly correlations.

Figure 12 shows the SDs of the monthly anomalies of CS OLRs as simulated using the STs of ID, IFD, ERA-40, NCEP2, NCEP1, and I0, and that observed from ERBE. It can be seen that the SD of I0 is low at middle-high-latitude regions (latitude >40°N/40°S). Nonetheless, a strong SD over the high-latitude Northern Hemisphere is observed by ERBE. This implies that the CS OLR anomaly over the high-latitude region is mainly due to the anomaly of ST. This point is confirmed according to the CS anomaly as simulated using the STs of ID, IFD, ERA-40, NCEP2, and NCEP1 as inputs. In contrast, ERBE shows very little variation in clear-sky OLR over the tropical and subtropical oceans, much less than the others. Furthermore, the figure shows that the SD of ERA-40 over Greenland and Antarctica is much weaker than ERBE. It implies that the ST anomaly of ERA-40 is too weak over Greenland and the Antarctic. In addition, the figure shows a high variability adjacent to the ice edge in the Antarctic (ERBE). Moreover, all the datasets show that the largest MAE of anomalies (>4 W m^{-2}) occurs over high-latitude coastal areas, such as the Antarctic (60°–70°S; not shown). This high variability is probably due to different positions of the ice margin in different years. The same partly applies to the ice edge in the Northern Hemisphere and for seasonally snow-covered regions. In contrast, a low SD in ID0 is found in these regions. Nonetheless, it is found that the anomaly correlations (0.44–0.48) between ERBE and the derived CS OLR are very small compared to the anomaly correlations (0.84–0.97) among the derived CS OLRs (Table 4b). It is of interest that all the derived CS OLRs are much

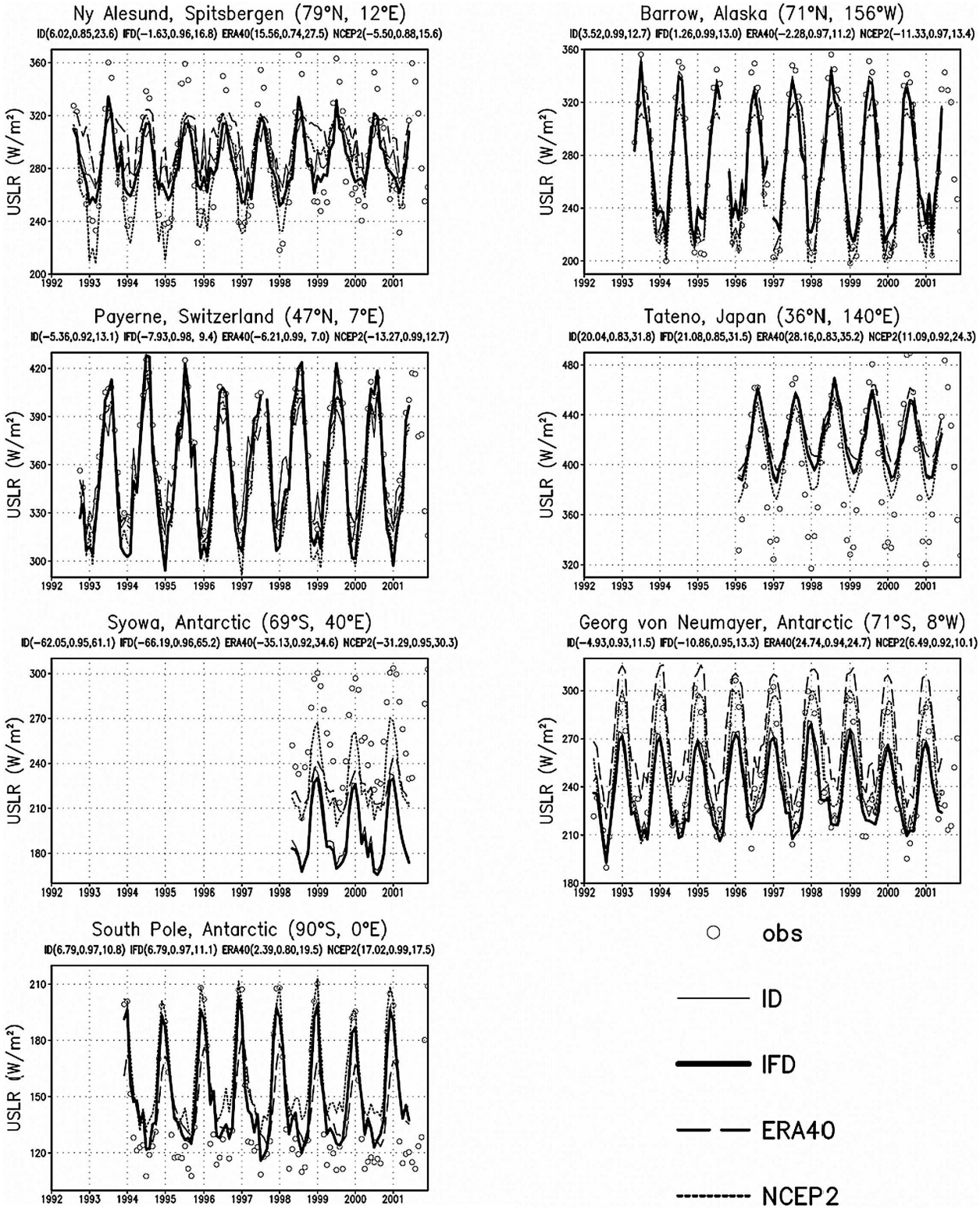

FIG. 7. Time series of USLR at BSRN stations, as simulated using the skin temperatures of ID, IFD, ERA-40, and NCEP2, where the numbers in the parentheses denote the biases (W m^{-2}), correlation coefficients, and MAEs (W m^{-2}) compared with the BSRN observation, respectively.

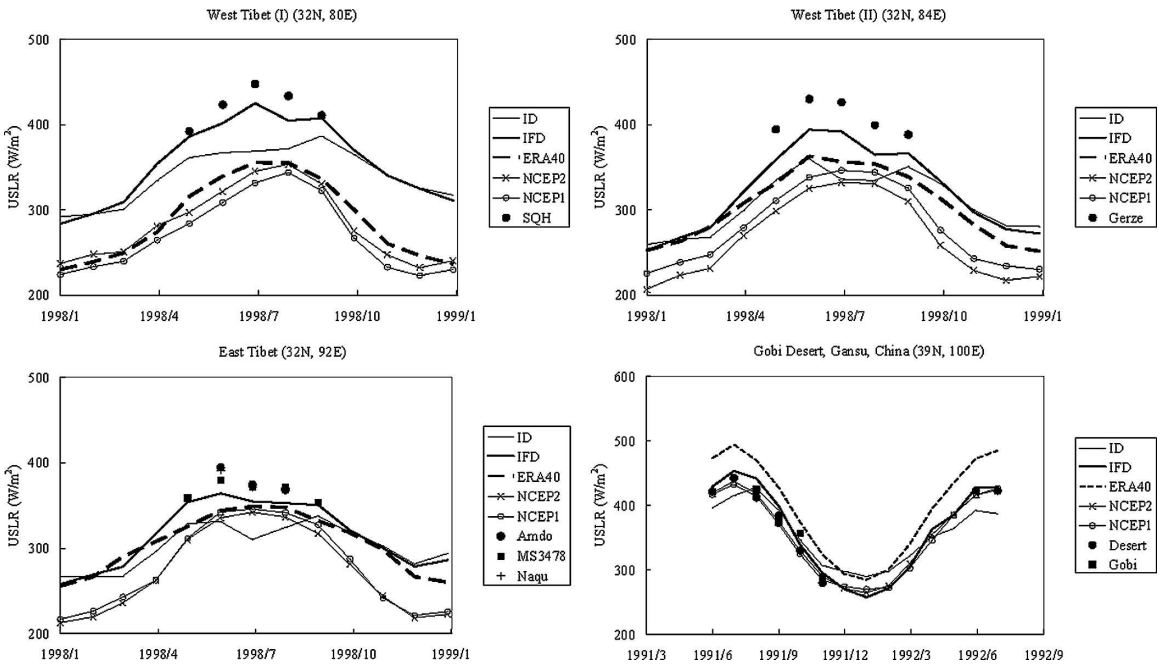

FIG. 8. Time series of USLR in Tibet and the Gobi Desert, as simulated using the skin temperatures of ID, IFD, ERA-40, NCEP2, and NCEP1.

more "similar" in this respect. They therefore all lack a similar deficiency, probably erroneous/poor T/q profiles over the tropical and subtropical regions, or oversimplified algorithms.

Table 3 shows that using the ID ST as an input, the calculated monthly anomaly of CS OLR has the lowest mean-absolute error (AMAE) at 3.42 W m^{-2} and the highest anomaly correlation (ACorr) at 0.48 among the calculated CS OLRs. Table 4b shows that all the calculated monthly anomalies of the CS OLRs of ID, IFD, ERA-40, NCEP2, NCEP1, and I0 are close to each other with anomaly correlation >0.81. The two satellite products are closer to each other than to the reanalyses. The three reanalyses are closer to each other than to the satellite products. However, the I0 run shows that atmospheric properties alone can obtain an anomaly correlation of 0.4 for CS OLR simulation when compared with the ERBE satellite observation. In contrast, the anomaly correlations of ID, IFD, ERA-40, NCEP2, and NCEP1 range from 0.44 to 0.48 compared with ERBE. That is, the anomaly correlations of ID, IFD, ERA-40, NCEP2, and NCEP1 are higher than I0. Since I0 does not contain the interannual variability of ST, the higher anomaly correlations of ID, IFD, ERA-40, NCEP2, and NCEP1 are due to their interannual variabilities of ST. In addition, it can be seen that the anomaly correlation of NCEP2 is slighter higher than NCEP1. It probably benefits from adopting the rainfall assimilation system over land in NCEP2 although the anomaly correlation of the simulated CS OLR of NCEP2 is only 1% better than that of NCEP1 (Table 4b).

Figure 13 shows the explained variance of the anomaly of CS OLR due to the anomalies of the STs of ID, IFD, ERA-40, NCEP2, and NCEP1. They are determined by subtracting the squared anomaly correlation of I0 run from those of ID, IFD, ERA-40, NCEP2, and NCEP1 runs for the period 1985–89. From the figure, it can be seen that over high-latitude Northern Hemisphere (latitude >50°N) using the ST of ID as an input generally improves the simulation of the variance. In contrast, no significant improvement is found when using the STs of NCEP2 and ERA-40 over the high-latitude Northern Hemisphere region. Over high-latitude Southern Hemisphere (latitude >50°S) all the ST datasets (ID, IFD, NCEP2, and ERA-40) are unable to explain the variance as observed in ERBE. It is suspected that there is an error in the ERBE observation. Over low- and middle-latitude oceans (10°–40°), the variance is slightly better explained by the STs of ERA-40 and NCEP2 rather than that of ID. Over tropical land, the variance is slightly better explained by the ST of ID rather than those of ERA-40 and NCEP2. Over eastern tropical Pacific Ocean in 120°–150°W, where the El Niño occurs, the variance of CS OLR can be explained by all the ST datasets (not shown).

Table 8 shows the explained variances for each of the modified IGBP land types. It is found that for all of the

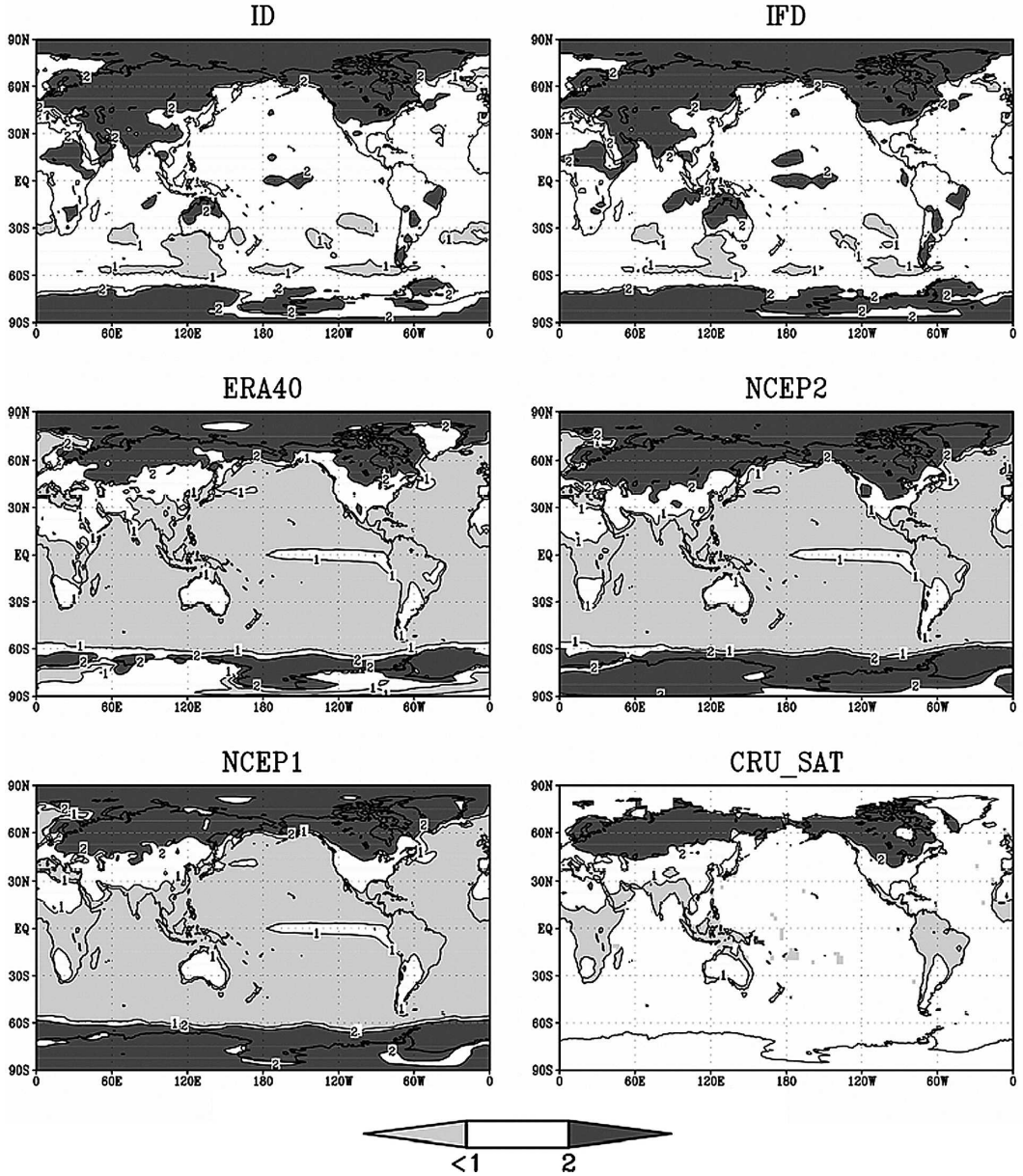

FIG. 9. Std dev (K) of the monthly skin temperature anomalies of ID, IFD, ERA-40, NCEP2, and NCEP1, and the near-SAT anomalies of CRU derived from the period 1985–2000. The anomalies are calculated by removing the climatic annual cycle of the monthly composites from the period 1985–2000.

land types, the interannual variabilities of the STs of the reanalyses (ERA-40, NCEP2, NCEP1) tend to be correlated with independent observations, calculated CS OLR in this study. Nevertheless, generally, in high-latitude regions (tundra, wetlands, deciduous needleleaf forests, and sea ice), the simulated CS OLRs derived from using the satellite STs have higher anomaly correlations than those derived using the reanalyses' STs. But suspected spurious noises with amplitude at 2 K in the satellite-derived STs produce physically unlikely anomalies to land types of which the amplitude of the anomaly is weak (such as open-water bodies, grasslands, rain forests, croplands, hot deserts, and cold deserts). It can be concluded that, although ID ST describes the largest variance of the natural interannual variability compared with the other ST datasets, its uncertainty of 1–2 K also introduces fake variance. Note that the errors in the reanalysis are usually systematic, while the errors in the observation are usually random. This means that the errors in the reanalysis induce bias

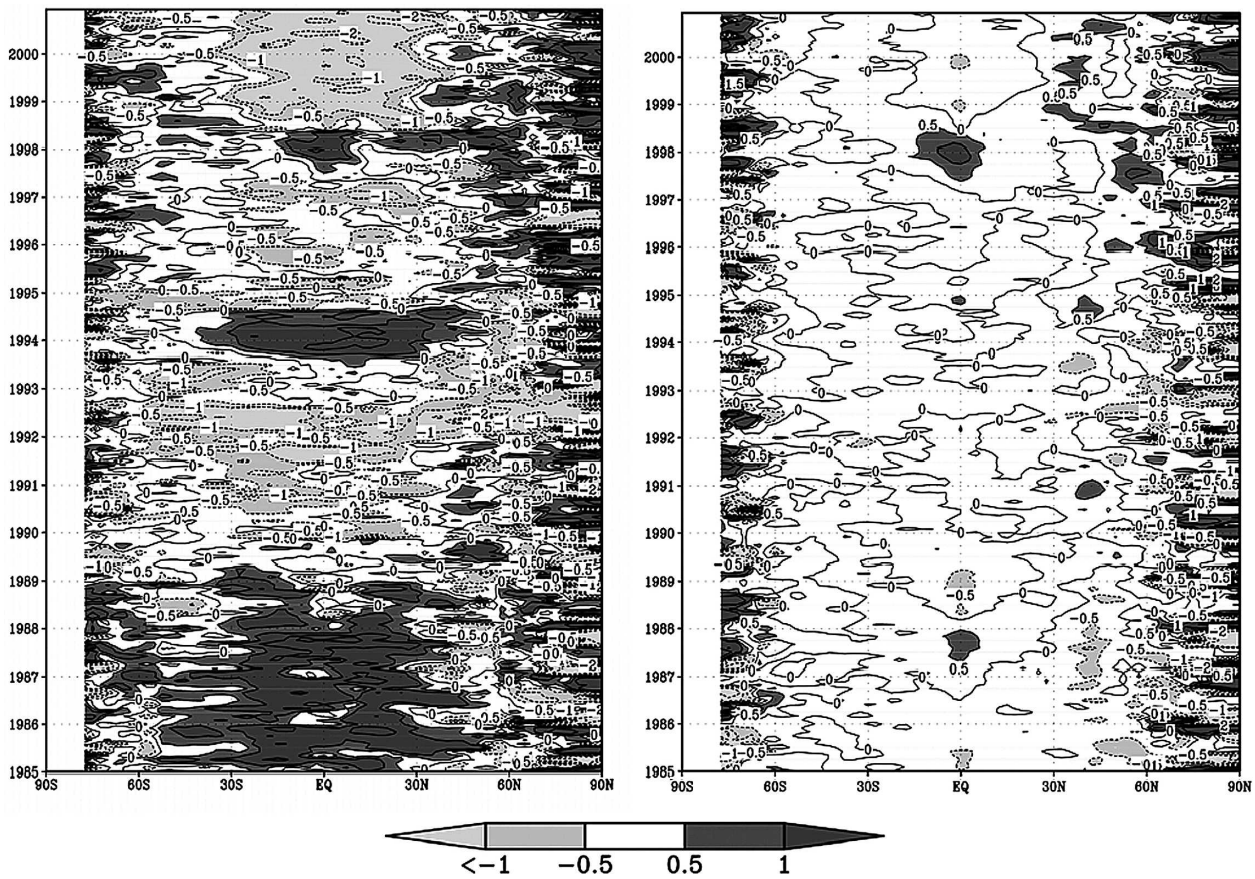

FIG. 10. Zonal means (K) of the SST anomalies of (left) ID and (right) ERA-40 from the period 1985–2000.

but little variance in the results (Roads 2003; Brotzge 2004; Werth and Avissar 2004); but the errors in the satellite data cause both bias and variance.

a. Subtropical ocean

The anomaly SST of the tropical and subtropical oceans in ID is much stronger than that in ERA-40 (Figs. 9 and 10). Nonetheless, the simulated CS OLR anomalies of these regions, especially in the subtropical oceans (10°–40°S, 20°–30°N), using the monthly STs of ID and IFD as inputs are worse than those using I0 ST as an input (Fig. 13). These imply that over the subtropical and tropical oceans, the SSTs of ID and IFD do not contain trustworthy signals of interannual variability. Robertson and Lu (2004) also made the same conclusion over these parts of ocean, and indicated that ID SST did not agree with any currently accepted measures of SST variability. They believe that the problem is related to the use of operational TOVS moisture soundings to correct for water vapor absorption–emission effects in the 11 micron data for retrieving the satellite ST. The TOVS problem has also been found by the ISCCP scientists (Zhang et al. 2006).

b. Glacier and coastal stations

The anomaly correlations of the simulated USLRs at the Antarctic BSRN stations (Syowa, Georg von Neumayer, and South Pole) using the STs of ID, IFD, ERA-40, and NCEP2 as inputs are >0.41 (Fig. 14). This implies that the STs of ID, IFD, ERA-40, and NCEP2 do contain their natural interannual variability in the Antarctic, which was not shown while comparing with the ERBE CS OLR (Fig. 13). Similarly, the STs of ID, IFD, ERA-40, and NCEP2 also contain their natural interannual variability in the coastal regions. The anomaly correlations of the simulated USLRs at most of the coastal BSRN stations (Ny Alesund, Barrow, Tateno, Syowa, and Georg von Neumayer), except for Tateno, using the STs of ID, IFD, ERA-40, and NCEP2 as inputs are >0.48 (Fig. 14).

c. Cropland station

Tables 5 and 7 show that over croplands, no significant biases of the simulated CS OLRs are found when using the STs of all the 5 products as inputs. Nonetheless, both of the satellite products (ID, IFD) seem un-

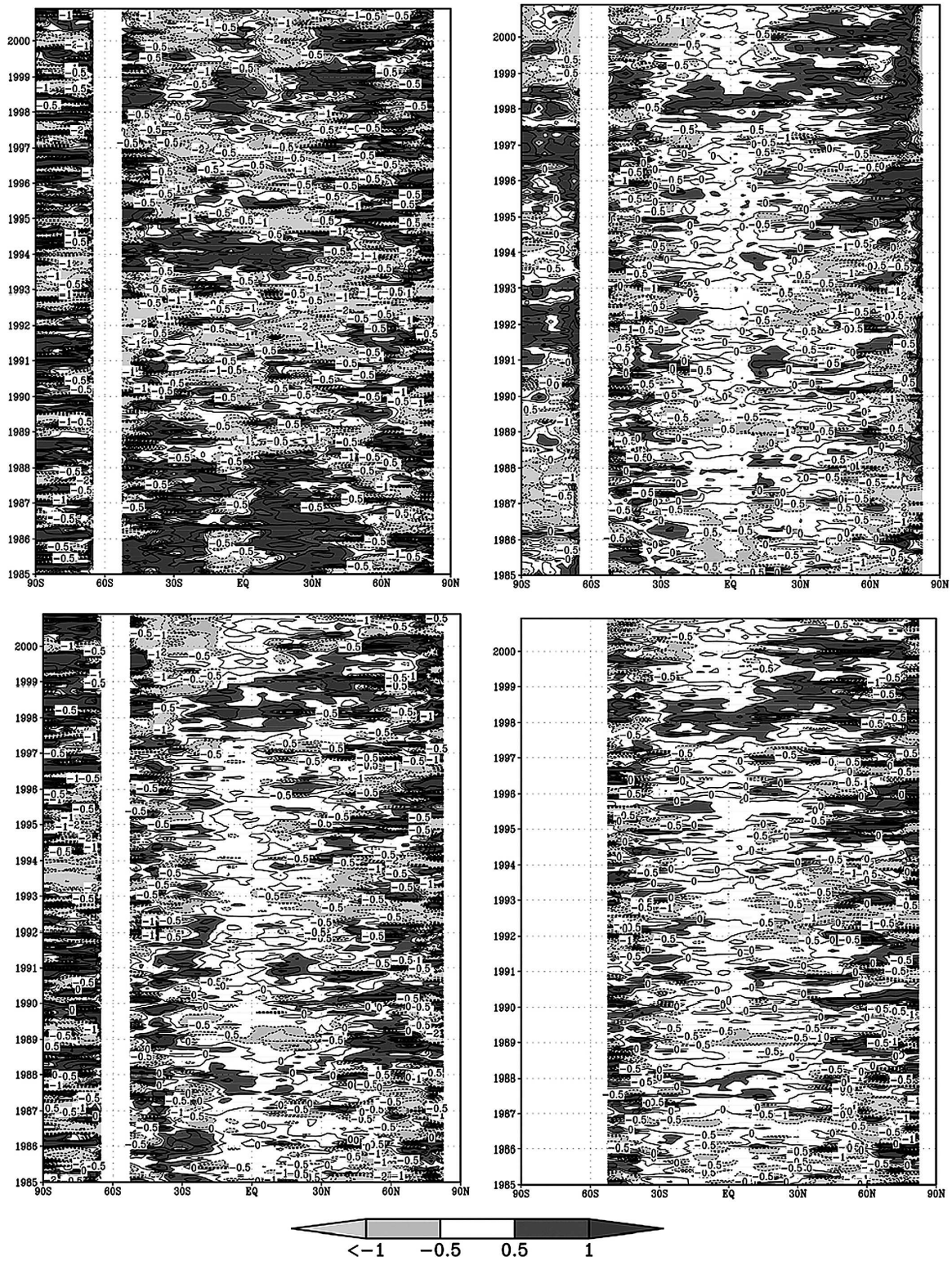

FIG. 11. Same as in Fig. 10, but for LST anomalies (K) of (top) (left) ID and (right) ERA-40; and (bottom) (left) NCEP2 and (right) the near-SAT anomaly of CRU from the period 1985–2000.

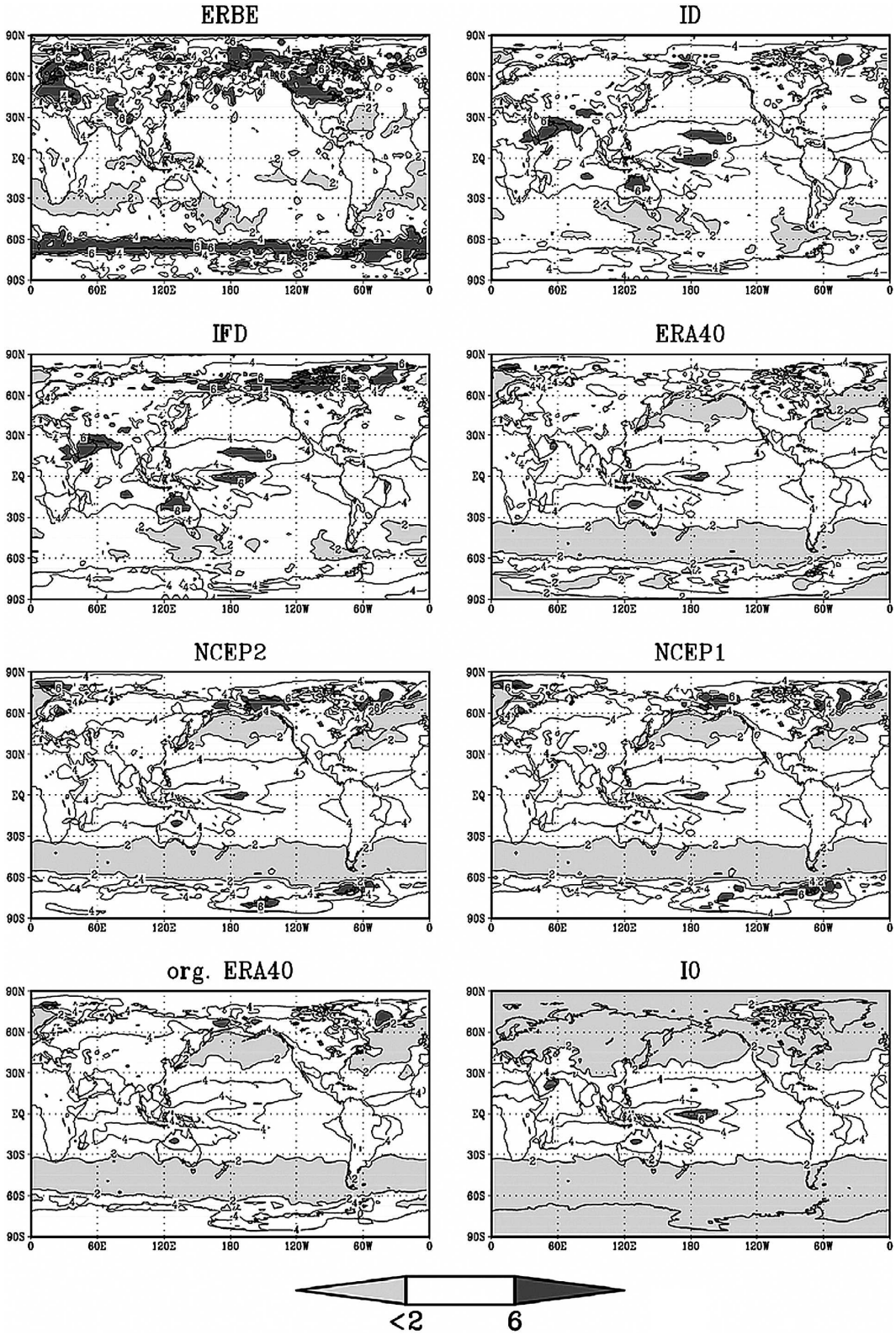

FIG. 12. Standard deviation (W m^{-2}) of the monthly anomalies of CS OLR from the period 1985–89, where ID, IFD, ERA-40, NCEP2, and NCEP1 are simulated using their corresponding monthly skin temperatures, I0 is simulated using the monthly composite of the skin temperature of ID, and ERBE is observed from ERBE.

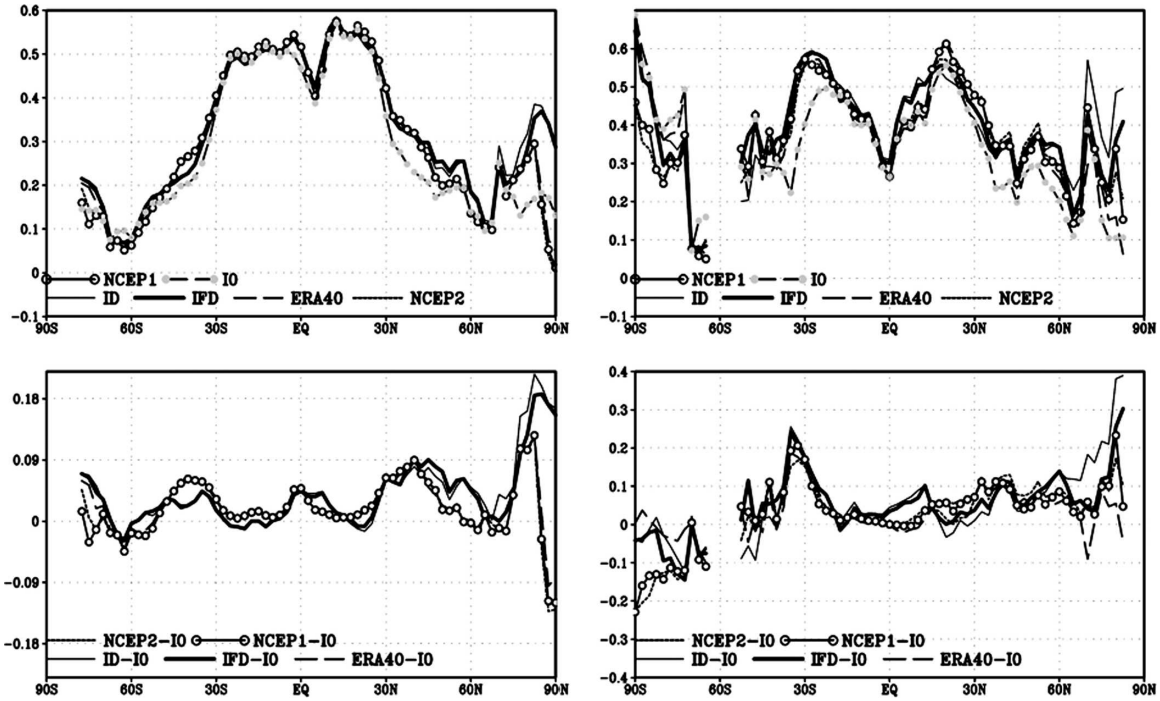

FIG. 13. Over (left) ocean and (right) land. (top) Total explained variance (Total EV) for the anomalies of CS OLRs (W m^{-2}) and (bottom) skin temperature (K) explained variance (ST EV) for the anomalies of CS OLRs due to the anomalies of the STs of ID, IFD, ERA-40, NCEP2, and NCEP1 (denoted as ID-I0, IFD-I0, ERA-40-I0, NCEP2-I0, and NCEP1-I0, respectively). The ST EVs are determined by subtracting the square of the anomaly correlation of I0 run from those of ID, IFD, ERA-40, NCEP2, and NCEP1 runs for the period 1985–89.

able to reproduce the ST-induced interannual variability of CS OLR (Table 8). In contrast, all the reanalyses (ERA-40, NCEP2, NCEP1) are able to reproduce the interannual variability. Nonetheless, Fig. 14 shows that at the only cropland BSRN station (Payerne), the anomaly correlations of the simulated USLRs are >0.62 using the STs of ID and IFD as inputs. This implies that the STs of ID and IFD do capture the

TABLE 8. Explained variances of the anomalies of CS OLRs 1) due to the anomalies of atmospheric properties (I0 run) and 2) due to the anomalies of skin temperatures at each land type. The skin temperature explained variances are determined by subtracting the squared anomaly correlation of I0 run from those of ID, IFD, ERA-40, NCEP2, and NCEP1 runs for the period 1985–89. AStd denotes the std dev of the monthly anomalies of CS OLR of the land type.

Land type	AStd (W m^{-2})	Atmospheric property (T/q profile) (I0) (%)	Surface property (skin temperature) (%)				
			ID	IFD	ERA-40	NCEP2	NCEP1
Glacier	2.18	17	16	12	1	20	23
Wetlands	2.58	8	20	13	2	8	7
Open water	0.45	1	0	−1	1	2	2
Croplands	1.26	53	−6	−3	15	20	17
Other forests	1.56	28	11	16	15	28	22
Rainforest	0.95	54	−7	−5	4	5	7
Deciduous needle-leaf forests	2.87	19	32	34	28	29	29
Grasslands	0.72	33	−10	−9	13	18	17
Tundra	2.07	18	13	7	4	9	4
Cold desert	1.42	26	−12	−11	32	20	31
Hot desert	1.53	57	−14	−6	13	6	13
Ocean with ice >0% but ≤10%	1.23	17	9	13	10	8	7
Ocean with ice >10%	1.42	25	10	8	8	7	7
Tibet	4.88	0	2	1	1	12	4
Andes Mountains	2.27	61	−25	−19	1	−1	−4

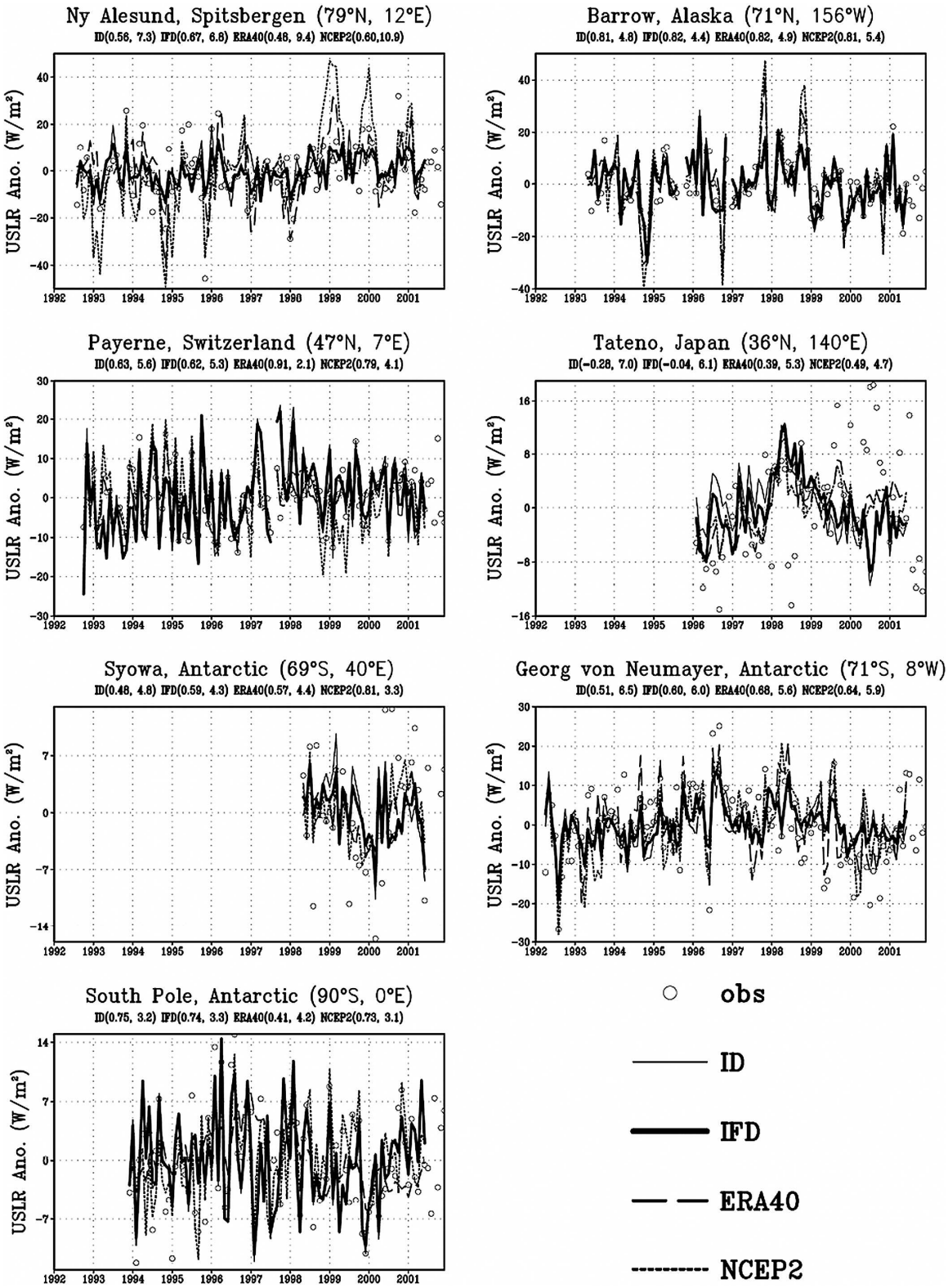

FIG. 14. Same as in Fig. 7, but for the anomalies of USLRs, where the numbers in the parentheses denote the correlation coefficients and MAEs (W m^{-2}) of the anomalies compared with the BSRN observation, respectively.

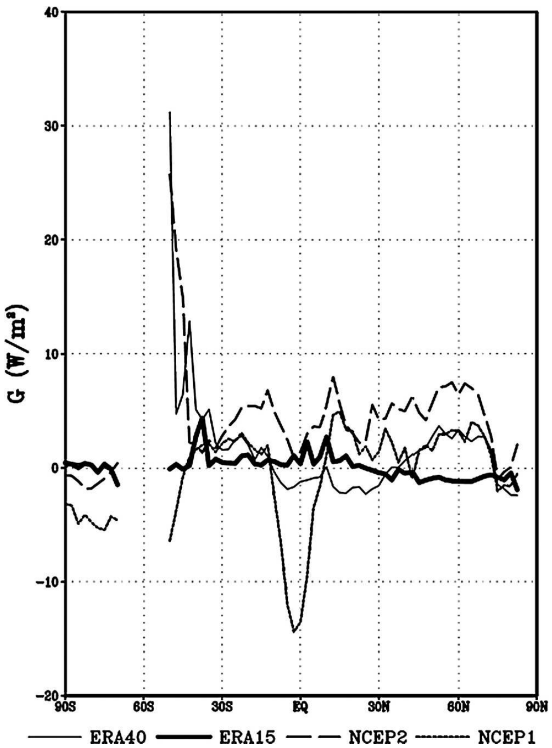

FIG. 15. Zonal mean of imbalanced ground heat flux over land for the period 1985–2000.

natural interannual variability at the cropland station although this statement may not be true at other cropland sites.

6. Discussion

That the results (e.g., Table 3) show the ST product of NCEP2 superior to ERA-40 is not surprising, since the force-restore rate equation is used in the NCEP reanalyses. This rate equation adjusts LST by forcing it to restore to climatic deep-soil temperature. If the deep-soil temperature is properly assigned, LST can be well reproduced (Deardorff 1978; Tsuang and Yuan 1994). Nonetheless, the force-restore rate equation may produce artificial ground heat fluxes as noted for NCEP1 in Tsuang (2005). As a result, the imbalanced heat fluxes of the land column energy budget components of the NCEP reanalyses are larger than those of the ERA reanalyses (Fig. 15). Alternatively, without using the force restore rate equation, better simulations of the skin temperature can be achieved by using improved parameterizations of the surface albedo, aerodynamic resistance, canopy resistance, area heat capacity, and snow and ice morphology of each land types (Tsuang and Tu 2002). Note that the land scheme of ERA-40 has been tested at field experiments in the United States, at Cabauw in the Netherlands, and in the Amazonian rain forest in Brazil. Over the tested sites, this study does not find significant errors for ERA-40. The errors are mostly found at sites where the land scheme is not well tested such as in Tibet, desert regions, and glacier regions.

7. Conclusions

This study evaluates the global skin temperature (ST) datasets of the ISCCP D satellite product, the ISCCP FD satellite product, ERA-40, NCEP2, and NCEP1. All the 5 ST datasets (ID, IFD, ERA-40, NCEP2, NCEP1) agree well with mean differences <3 K for most land types, except in desert and mountain areas. Nonetheless, all the reanalyses overestimate ST during the partial sea ice–covered period in the 48°–56°N zone. ERA-40 underestimates the seasonal variation of ST over glaciers. All the reanalyses (ERA-40, NCEP2, NCEP1) underestimate CS OLR over Tibet and the Andes Mountains by >10 W m^{-2}. All the satellite-derived STs and the ERA-40 ST overestimate CS OLR over hot deserts by >10 W m^{-2}. If the modeling system used in this study systematically biases high by only 4 W m^{-2} in CS OLR calculation, it can be inferred that the STs of the reanalyses are too cold on Tibet and the Andes Mountains, and that the STs of all the satellite products and ERA-40 are too high for hot deserts.

It is found that all the interannual variabilities of the CS OLRs calculated by the reanalyzed STs (ERA-40, NCEP2, and NCEP1) are positively correlated with the ERBE observation for all the land types. However, the magnitude of the largest CS OLR anomalies observed over fractional sea ice–covered oceans and seasonally snow-covered lands by the ERBE satellites were underestimated using the reanalyzed STs. On the other hand, those calculated using the satellite-retrieved STs (ID, IFD) generally agree better in high-latitude regions; however, they have spurious signals in subtropical regions.

For developing a better reanalysis product, more thorough tests are suggested over desert regions (e.g., Sahara Desert, Middle East deserts, Australia desert), mountain areas (e.g., Tibet, the Andes Mountains), and ice sheets (e.g., Greenland, the Antarctic), where more field data should be collected. In addition, better schemes should be developed for fractional sea ice–covered oceans and seasonally snow-covered lands.

Acknowledgments. This work is supported by National Science Council/Taiwan under Contracts NSC-93-2621-Z-005-001, 94-2621-Z-005-005, 95-2621-Z-005-004, and 95-2111-M-005-001. ERA-40 data are taken

from European Centre for Medium-Range Weather Forecasts (more information is available online at http://www.ecmwf.int/), NCEP reanalysis data are provided by the National Weather Service–Climate Prediction Center (more information is available online at http://nomad3.ncep.noaa.gov/ncep_data/index.html), BSRN data are taken from the Baseline Surface Radiation Network (more information is available online at http://bsrn.ethz.ch/), CRU data are taken from Climatic Research Unit, University of East Anglia (more information is available online at http://www.cru.uea.ac.uk/), and HEIFE data are taken from Heihe river basin Field Experiment (more information is available online at http://ssrs.dpri.kyoto-u.ac.jp/~heife/). The GAME-Tibet project was supported by the MEXT, FRSGC, NASDA of Japan, the Chinese Academy of Science, and the Asian Pacific Network. Thanks to Noel Dallow for proofreading and to two anonymous reviewers for enriching the manuscript significantly.

APPENDIX

Relevant Sensitivity Study Results Based on NASA GISS Radiative Transfer Model and ISCCP Datasets

Zhang et al. (1995, 2004, 2006, 2007) have conducted various (>60) sensitivity tests based on the radiative transfer model of the National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies and ISCCP cloud climatological datasets (of C- and D-series; Rossow and Schiffer 1991). Their sensitivity studies show how much uncertainty appears in the calculated fluxes if uncertainty only in a single input parameter (of atmospheric, surface, or cloud properties) is introduced. Each test calculation with altered model or input parameter values covers the whole globe for one particular day (average of 8 times of day for 15 July 1986) to provide a realistic range in flux uncertainty estimate. The magnitudes of the input parameter changes for the calculation are based on the realistic uncertainty estimate of the input parameters. Daily mean fluxes are compared at each grid cell of 280-km equal-area map, either with the original base values or with another calculation that has a parameter change of the same magnitude but opposite sign. For the latter, and based on the previous and current calculations by one of the coauthors (Zhang), we here report the global mean of the flux changes in response to the changes of the concerned inputs for this work. The most important clear-sky OLR flux uncertainties (in order of flux uncertainty magnitude) for each input variable change are as follows: 1) A 2-K increase of the whole atmospheric profile temperature causes an increase of 9.7 W m^{-2} with the SD of 2.1 W m^{-2}, 2) a 25% increase of column precipitable water vapor causes a change of -8.5 W m^{-2} with the SD of 4.2 W m^{-2}, 3) a 50% increase of 400–100-mb precipitable water vapor causes a change of -5.4 W m^{-2} with the SD of 2.0 W m^{-2}, 4) a 2-K increase of surface skin temperatures causes an increase of 4.3 W m^{-2} with the SD of 1.1 W m^{-2}, 5) a 3-K increase of 400–100-mb temperatures causes an increase of 3.8 W m^{-2} with the SD of 0.9 W m^{-2}, 6) a 2-K increase of the near-surface layer [1000–800 mb from mean sea level, but generally varies with topography; see Rossow and Schiffer (1991)] temperature causes an increase of 2.5 W m^{-2} with the SD of 0.9 W m^{-2}, 7) a 2-K increase of the surface air temperature causes an increase of 1.4 W m^{-2} with the SD of 0.5 W m^{-2}, and 8) a 25% increase of the near-surface layer precipitable water vapor causes a change of -1.2 W m^{-2} with the SD of 1.5 W m^{-2}.

REFERENCES

Aires, F., C. Prigent, and W. B. Rossow, 2004: Temporal interpolation of global surface skin temperature diurnal cycle over land under clear and cloudy conditions. *J. Geophys. Res.,* **109,** D04313, doi:10.1029/2003JD003527.

Bacmeister, J. T., and M. J. Suarez, 2002: Wind stress simulations and equatorial momentum budget in an AGCM. *J. Atmos. Sci.,* **59,** 3051–3073.

Baldocchi, D., and Coauthors, 2001: FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bull. Amer. Meteor. Soc.,* **82,** 2415–2434.

Barkstrom, B. R., 1984: The Earth Radiation Budget Experiment (ERBE). *Bull. Amer. Meteor. Soc.,* **65,** 1170–1185.

——, E. Harrison, G. Smith, R. Green, J. Kibler, and R. Cess, 1989: Earth Radiation Budget Experiment (ERBE) archival and April 1985 results. *Bull. Amer. Meteor. Soc.,* **70,** 1254–1262.

Betts, A. K., and J. H. Ball, 1998: FIFE surface climate and site-average dataset 1987–89. *J. Atmos. Sci.,* **55,** 1091–1108.

Brest, C. L., W. B. Rossow, and M. D. Roiter, 1997: Update of radiance calibrations for ISCCP. *J. Atmos. Oceanic Technol.,* **14,** 1091–1109.

Brotzge, J. A., 2004: A two-year comparison of the surface water and energy budgets between two OASIS sites and NCEP–NCAR reanalysis data. *J. Hydrometeor.,* **5,** 311–326.

Brutsaert, W., 1982: *Evaporation into the Atmosphere*. Reidel, 299 pp.

Cavalieri, D., C. Parkinson, P. Gloerson, and H. J. Zwally, cited 1996: Sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I passive microwave data. National Snow and Ice Data Center, Boulder, CO, digital media. [Available online at http://nsidc.org/data/docs/daac/nsidc0051_gsfc_seaice.gd.html.]

Chou, M.-D., M. J. Suarez, X.-Z. Liang, and M. M.-H. Yan, 2001: A thermal infrared radiation parameterization for atmospheric studies. NASA Goddard Space Flight Center Tech. Rep. TM-2001-104606, Vol. 19, 55 pp.

Curry, J. A., and Coauthors, 2004: SEAFLUX. *Bull. Amer. Meteor. Soc.,* **85,** 409–424.

Deardorff, J. W., 1978: Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation. *J. Geophys. Res.,* **83,** 1889–1903.

Gates, W. L., and Coauthors, 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Amer. Meteor. Soc.,* **80,** 29–55.

Gibson, J. K., P. Kållberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano, 1997: The ERA description. ECMWF Re-Analysis Project Report Series 1, 86 pp.

Green, R. N., F. B. House, P. W. Stackhouse, X. Wu, S. A. Ackermann, W. L. Smith, and M. J. Johnson, 1990: Intercomparison of scanner and nonscanner measurements for the Earth Radiation Budget Experiment (ERBE). *J. Geophys. Res.,* **95,** 11 785–11 798.

Gruber, A., D. Su, M. Kanamitsu, and J. Schemm, 2000: The comparison of two merged rain gauge–satellite precipitation datasets. *Bull. Amer. Meteor. Soc.,* **81,** 2631–2644.

Jin, M., 2004: Analysis of land skin temperature using AVHRR observations. *Bull. Amer. Meteor. Soc.,* **85,** 587–600.

——, R. E. Dickinson, and A. M. Vogelmann, 1997: A comparison of CCM2–BATS skin temperature and surface-air temperature with satellite and surface observations. *J. Climate,* **10,** 1505–1524.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kanamitsu, M., 1989: Description of the NMC Global Data Assimilation and Forecast System. *Wea. Forecasting,* **4,** 335–342.

——, W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002a: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.,* **83,** 1631–1643.

——, and Coauthors, 2002b: NCEP dynamical seasonal forecast system 2000. *Bull. Amer. Meteor. Soc.,* **83,** 1019–1037.

Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.,* **82,** 247–267.

Loveland, T. R., B. C. Reed, J. F. Brown, D. O. Ohlen, J. Zhu, L. Yang, and J. W. Merchant, 2001: Development of a global land cover characteristics database and IGBP DISCover from 1-km AVHRR data. *Int. J. Remote Sens.,* **21,** 1303–1330.

Masson, V., J.-L. Champeaux, F. Chauvin, C. Meriguet, and R. Lacaze, 2003: A global database of land surface parameters at 1-km resolution in meteorological and climate models. *J. Climate,* **16,** 1261–1282.

Mitchell, T. D., and P. D. Jones, 2005: An improved method of constructing a database of monthly climate observations and associated high resolution grids. *Int. J. Climatol.,* **25,** 693–712.

Ohmura, A., and Coauthors, 1998: Baseline Surface Radiation Network (BSRN/WRMC): New precision radiometry for climate research. *Bull. Amer. Meteor. Soc.,* **79,** 2115–2136.

Pan, H.-L., and L. Mahrt, 1987: Interaction between soil hydrology and boundary-layer development. *Bound.-Layer Meteor.,* **38,** 185–220.

Prigent, C., F. Aires, and W. B. Rossow, 2003: Land surface skin temperatures from a combined analysis of microwave and infrared satellite observations for an all-weather evaluation of the differences between air and skin temperatures. *J. Geophys. Res.,* **108,** 4310, doi:10.1029/2002JD002301.

Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate,* **7,** 929–948.

Roads, J., 2003: The NCEP–NCAR, NCEP–DOE, and TRMM tropical atmosphere hydrologic cycles. *J. Hydrometeor.,* **4,** 826–840.

Robertson, F. R., and H.-I. Lu, 2004: How well are recent climate variability signals resolved by satellite radiative flux estimates? Preprints, *13th Conf. on Satellite Meteorology and Oceanography,* Norfolk, VA, Amer. Meteor. Soc., CD-ROM, 5.4.

Rossow, W. B., and R. A. Schiffer, 1991: ISCCP cloud data products. *Bull. Amer. Meteor. Soc.,* **72,** 1–20.

——, and L. C. Garder, 1993: Validation of ISCCP cloud detections. *J. Climate,* **6,** 2370–2393.

——, and R. A. Schiffer, 1999: Advances in understanding clouds from ISCCP. *Bull. Amer. Meteor. Soc.,* **80,** 2261–2288.

Simmons, A. J., and J. K. Gibson, Eds., 2000: The ERA-40 project plan. ECMWF Re-Analyses Project Report Series 1, 62 pp.

——, and Coauthors, 2004: Comparison of trends and low-frequency variability in CRU, ERA-40, and NCEP/NCAR analyses of surface air temperature. *J. Geophys. Res.,* **109,** D24115, doi:10.1029/2004JD005306.

Sud, Y., and D. M. Mocko, 1999: New snow-physics to complement SSiB. Part I: Design and evaluation with ISLSCP initiative I datasets. *J. Meteor. Soc. Japan,* **77,** 335–348.

Tamagawa, I., 1996: Turbulent characteristics and bulk transfer coefficients over the desert in the HEIFE area. *Bound.-Layer Meteor.,* **77,** 1–20.

Tao, W.-K., and Coauthors, 2002: Microphysics, radiation and surface processes in the Goddard Cumulus Ensemble (GCM) model. *Meteor. Atmos. Phys.,* **82,** 97–137.

Tsuang, B.-J., 2003: Analytical asymptotic solutions to determine interactions between the planetary boundary layer and the Earth's surface. *J. Geophys. Res.,* **108,** 8608, doi:10.1029/2002JD002557.

——, 2005: Ground heat flux determination according to land skin temperature observations from in situ stations and satellites. *J. Hydrometeor.,* **6,** 371–390.

——, and H. C. Yuan, 1994: The ideal numerical surface thickness to determine ground surface temperature and schemes comparison (in Chinese). *Atmos. Sci.,* **22,** 189–218.

——, and C.-Y. Tu, 2002: Model structure and land parameter identification: An inverse problem approach. *J. Geophys. Res.,* **107,** 4096, doi:10.1029/2001JD000711.

——, ——, and K. Arpe, 2001: Lake parameterization for climate models. Max Planck Institute for Meteorology Rep. 316, 72 pp.

Tu, C.-Y., and B.-J. Tsuang, 2005: Cool-skin simulation by a one-column ocean model. *Geophys. Res. Lett.,* **32,** L22602, doi:10.1029/2005GL024252.

Viterbo, P., and A. C. M. Beljaars, 1995: An improved land surface parameterization scheme in the ECMWF model and its validation. *J. Climate,* **8,** 2716–2748.

——, ——, J. F. Mahfouf, and J. Teixeira, 1999: The representation of soil moisture freezing and its impact on the stable boundary layer. *Quart. J. Roy. Meteor. Soc.,* **125,** 2401–2426.

Wang, J., and Y. Mitsuta, 1991: Turbulence structure and transfer characteristics in the surface layer of the HEIFE Gobi area. *J. Meteor. Soc. Japan,* **69,** 587–593.

Webster, P. J., C. A. Clayson, and J. A. Curry, 1996: Clouds, radiation, and the diurnal cycle of sea surface temperature in the tropical western Pacific. *J. Climate,* **9,** 1717–1730.

Werth, D., and R. Avissar, 2004: The regional evapotranspiration of the Amazon. *J. Hydrometeor.,* **5,** 100–109.

Yang, K., T. Koike, P. Stackhouse, C. Mikovitz, and S. J. Cox, 2006: An assessment of satellite surface radiation products

for highlands with Tibet instrumental data. *Geophys. Res. Lett.,* **33,** L22403, doi:10.1029/2006GL027640.

Yang, S.-K., Y.-T. Hou, A. J. Miller, and K. A. Campana, 1999: Evaluation of the Earth Radiation Budget in the NCEP–NCAR Reanalysis with ERBE. *J. Climate,* **12,** 477–493.

Zhang, Y., W. B. Rossow, and A. A. Lacis, 1995: Calculation of surface and top of atmosphere radiative fluxes from physical quantities based on ISCCP data sets, 1: Method and sensitivity to input data uncertainties. *J. Geophys. Res.,* **100,** 1149–1166.

——, ——, ——, V. Oinas, and M. I. Mishchenko, 2004: Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data. *J. Geophys. Res.,* **109,** D19105, doi:10.1029/2003JD004457.

——, ——, and P. W. Stackhouse Jr., 2006: Comparison of different global information sources used in surface radiative flux calculation: Radiative properties of the near-surface atmosphere. *J. Geophys. Res.,* **111,** D13106, doi:10.1029/2005JD006873.

——, ——, and ——, 2007: Comparison of different global information sources used in surface radiative flux calculation: Radiative properties of the surface. *J. Geophys. Res.,* **112,** D01102, doi:10.1029/2005JD007008.