JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 111, D13106, doi:10.1029/2005JD006873, 2006

Comparison of different global information sources used in surface radiative flux calculation: Radiative properties of the near-surface atmosphere

Yuanchong Zhang,[1] William B. Rossow,[2] and Paul W. Stackhouse Jr.[3]

Received 9 November 2005; revised 16 February 2006; accepted 23 March 2006; published 11 July 2006.

[1] Direct estimates of surface radiative fluxes that resolve regional and weather-scale variability over the whole globe with reasonable accuracy have only become possible with the advent of extensive global, mostly satellite, data sets within the past couple of decades. The accuracy of these fluxes, estimated to be about 10–15 W/m^2, is largely limited by the uncertainties of the input data sets. This study presents a fuller, more quantitative evaluation of these uncertainties, mainly for the near-surface air temperature and humidity, by comparing the main available global data sets from the European Centre for Medium-Range Weather Forecasts, NASA, the National Centers for Environmental Prediction, the International Satellite Cloud Climatology Project (ISCCP) and the Laboratoire de Météorologie Dynamique that are treated as ensemble realizations of actual climate such that their differences represent an estimate of the uncertainty in their measurements (because we do not know the absolute truth). The results are globally representative and may be taken as a generalization of our previous ISCCP-based uncertainty estimates for the input data sets. Near-surface atmospheric properties have the primary role in determining the surface downward longwave (LW) flux. From this study, the most important quantity, the surface air temperature, has a uncertainty of about 2–4 K (3 K on average), which would easily induce about 15 W/m^2 uncertainty for surface downward LW flux. The humidity profile comparison suggests an uncertainty of 20–25% for the atmospheric column precipitable water below the 300 hPa level, which would cause $\lesssim$10 W/m^2 uncertainty for surface downward LW flux, making it the second largest source of uncertainty. The comparison for the difference between surface skin and air temperature shows its uncertainty is about 2–3 K, which translates into 10–15 W/m^2 uncertainty for surface net LW flux. The used atmospheric data set from ISCCP represents the diurnal variations better than the other available sources (as it was designed to do) and the synoptic variations only slightly better than the other sources, but it still has notable clear-cloudy sky biases and interannual variations that are dominated by spurious changes introduced by methodology changes in the original TOVS product. In a companion paper, the work is extended to evaluate the uncertainties of surface radiative properties.

Citation: Zhang, Y., W. B. Rossow, 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.

1. Introduction

[2] The era of truly global observations of Earth by satellites is now more than 40 years long. A number of significant scientific advances have been facilitated by the study of these comprehensive observations. Over the past couple of decades research effort has focused on determining from satellite observations the temperature and humidity of the atmosphere, the properties of clouds and the distribution of precipitation to characterize their variability and their effects on the planetary and surface energy budgets. In particular, such studies have resulted in the production of several extensive and detailed surface radiative flux products [*Rossow and Lacis*, 1990; *Darnell and Staylor*, 1992; *Whitlock et al.*, 1995; *Rossow and Zhang*, 1995; *Stackhouse et al.*, 2004; *Zhang et al.*, 2004] that are produced using radiative transfer models and measurements of the physical properties of clouds, the atmosphere and surface obtained from a suite of global data sets.

[3] However, the usefulness of these flux products for monitoring long-term variations of global radiation budgets

[1]Applied Physics and Applied Mathematics, Columbia University, at NASA Goddard Institute for Space Studies, New York, New York, USA.
[2]NASA Goddard Institute for Space Studies, New York, New York, USA.
[3]NASA Langley Research Center, Hampton, Virginia, USA.

0148-0227/06/2005JD006873$09.00

and climate changes, as well as in other applications such as biological and oceanic modeling, is dependent on two separate issues: (1) the current uncertainty (or reliability) that can be estimated by evaluation of the calculated radiative fluxes against more direct measurements and (2) how much and by what means may the flux products be improved. Addressing issue 1, *Zhang et al.* [1995, 2004] have compared their flux products against surface (and the top of atmosphere) flux observations to estimate their uncertainty. Their 2004 estimate is about 10–15 W/m^2 uncertainty for surface fluxes. For issue 2, they have conducted all the important sensitivity studies by varying the input data sets (and the radiative transfer model parameters) to provide quantitative uncertainty estimates of the impacts of realistically assumed uncertainties of the input data sets. The ''realistically assumed'' uncertainties for the input data sets were based on earlier estimates, mainly from the International Satellite Cloud Climatology Project (ISCCP) D1 [*Rossow and Schiffer*, 1991]. Some of their conclusions are summarized here. (1) Although clouds have, for a long time, been highlighted as the major source of uncertainty in both planetary and surface radiation budgets, the advent of extensive cloud data sets has reduced this source of uncertainty. (2) As a result, other uncertainty sources have now become relatively more important and comparable to the uncertainty from cloud data sets. (3) For surface radiative flux estimates in particular, the accuracy is now largely limited by two aspects of Earth observations that have long been assumed to be adequately observed, especially during the satellite era, namely, the near-surface atmospheric radiative properties (temperature and humidity) that are usually obtained from atmospheric profiling instruments and the surface radiative properties (surface skin temperature, solar albedo and infrared emissivity).

[4] In order to improve the accuracy of estimating surface radiative fluxes further, the pressing problem is evidently to improve these inputs. As a first step we must find out how accurate are the most representative global climatological data sets used as inputs by most radiation models so that we know where and on which parameters we should concentrate our efforts to reduce the uncertainties.

[5] This issue is investigated in the current work mainly for the near-surface atmospheric properties (but also for some closely related atmospheric profile and surface properties). We try to quantify their uncertainties by comparing the differences of the main sources of global, long-term information concerning the near-surface atmospheric properties (section 2). In the final section, we summarize the uncertainty estimates and the uncertainties they induce into the surface radiation budget and discuss some implications for understanding and modeling the surface-atmosphere interactions. In a companion paper we use the similar approach but for surface properties (Y.-C. Zhang et al., Comparison of different global information sources used in surface radiative flux calculation: Radiative properties of the surface, submitted to *Journal of Geophysical Research*, 2005, hereinafter referred to as Zhang et al., submitted manuscript, 2005).

2. Near-Surface Atmospheric Properties

[6] It is difficult to incorporate the latest versions of all the available atmospheric data sets for this study since they have been updated relatively frequently. When beginning to prepare this paper several years ago, we chose representative data sets that provided complete (or nearly complete) global atmospheric information with a preference for those that also provide both a long time record and high time resolution, at least daily or preferably diurnal resolution. We start with our comparisons for atmospheric temperature/humidity profiles and then focus on the near-surface properties that are derived from them.

2.1. Atmospheric Profiles

[7] The traditional direct measurements of atmospheric temperature and humidity profiles come from the radiosonde (or upper air) network. In their analysis of long-term climate trends, *Oort and Liu* [1993] processed radiosonde data for a 380-month period (May 1958 through December 1989), with once-a-day, early on, to twice-a-day instrumented balloon launchings at more than 500 stations worldwide in support of weather forecasts. However, these routine observations do not resolve the diurnal cycle, may not be too representative of the daily mean for atmospheric properties and have rather sparse spatial coverage, particularly over oceans and the Southern Hemisphere. Hence these data are usually ''assimilated'' (i.e., combined with a global weather forecast model) to produce a globally complete description of the atmosphere. Later analyses supplement the radiosondes with satellite infrared and microwave radiances sensitive to temperature and humidity. Three systematic products of such analyses (called the reanalyses), produced by using fixed versions of the models and assimilation procedures, are used in this study (while maintaining their fundamental nature, some of these have been revised since we started this work): (1) the first version of the the European Centre for Medium-Range Weather Forecasts reanalysis for a 15-year period (1979–1993, called ERA15 or ER in our figures/tables) [*Gibson et al.*, 1999], (2) the first version of the NASA Data Assimilation Office Goddard Earth Observing System reanalysis for a 15-year period (March 1980 to February 1995, called GEOS-1 or GE in our figures/tables) [*Takacs et al.*, 1994], and (3) the second version of the National Centers for Environmental Prediction reanalysis for a 51-year period (1948–1998, called NCEP or NC in our figures/tables) [*Kistler et al.*, 2001]. All three data sets provide global atmospheric temperature and humidity profiles covering many years with spatial resolutions around 280 km. The ERA15 and GEOS-1 data sets have a temporal sampling interval of 6 hours, while the 6-hour version of NCEP was not available when we began this study (but it is now, see http://www.cdc.noaa.gov/).

[8] The NOAA polar weather satellites have carried infrared spectroradiometers that have been used to determine profiles of atmospheric temperature and humidity since 1973; this task was augmented by the advent of microwave spectroradiometers in 1979. However, these data have not been generally used or assimilated in weather analyses and forecasts until recently. The NOAA operational analysis of these measurements is called the operational Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) product and provides profiles for all locations, except those that are very cloudy, at least once-daily at spatial intervals of about 280 km. Since

Table 1. Five Major Data Sets for Temperature/Humidity Profiles and Their Associated Surface Properties Plus the Sixth (NVAP) for Humidity Profile Only[a]

Original Property	ERA15 ("ER")	GEOS-1 ("GE")	NCEP ("NC")	FDTV[b] ("TV")	3I ("3I")	NVAP ("NT")[c]
Covered time period	1979 to 1993	Mar 1980 to Feb 1995	1948 to 1998	Jul 1983 to 2003	1987 to 1995	1988 to 1997
Temporal resolution	daily mean, monthly and monthly hourly	daily mean, monthly and monthly hourly	daily mean and monthly	daily, monthly and monthly hourly	AM/PM daily, 5-day and monthly	daily and monthly
Horizontal spatial resolution	2.5° × 2.5° (144,73)	2.5° × 2.0° (144,91)	2.5° × 2.5° (144,73)	equal-area 2.5° (6596)	1° × 1° (360,180)	1° × 1° (360,180)
Vertical spatial resolution for T	17 levels (1000 → 10 hPa)	18 levels (1000 → 20 hPa)	8 levels (925 → 100 hPa)	≤9 layers (SRF → 10 hPa)	9 layers (SRF → 15 hPa)	N/A
Q specification	specific humidity at 17 levels	specific humidity at 18 levels	specific humidity at 5 levels	PW for 5 layers (SRF → 310 hPa)	PW for 5 layers (SRF → TOA)	PW for 3 layers (SRF → 300 hPa)
Is there surface air T?	yes (2 m T)	no	yes (2 m T)	yes (profile-extrapolated)	no	N/A
Is there surface air Q?	yes, specific humidity	yes, specific humidity	yes, specific humidity	no	No	no
Surface skin T	available with the same horizontal resolution as above	available with the same horizontal resolution as above	available with the same horizontal resolution as above	available with the same horizontal resolution as above	available with the same horizontal resolution as above	N/A
Is there surface albedo	yes (but we do not use them, see text)	yes (but we do not use them, see text)	yes (but we do not use them, see text)	yes (from ISCCP-FD, see text)	no	N/A

[a] T and Q are for temperature and humidity, respectively, and defined at either on levels or at the centers of layers. Accordingly, the humidity profiles are either in (level) specific humidity or (layer) precipitable water, but they usually do not have data in the stratosphere/uppermost troposphere. For a unified comparison, all the surface air temperature used in this comparison work are extrapolated from their unified temperature profiles (see text). PW is for precipitable water.

[b] As defined in the text, ISCCP-FD-TOVS, simplified as FDTV, incorporates the temperature/humidity profile from ISCCP-TOVS, surface air and skin temperature and total column PW from ISCCP-FD radiative flux data set, in which the ISCCP-TOVS surface air temperature for land is diurnal adjusted and then logarithmically extrapolated, the surface skin temperature is emissivity corrected with also diurnal adjustment (for land only), and PW is also adjusted (see text and *Zhang et al.* [2004]).

[c] NVAP's PW is relayed using FDTV surface pressure since NVAP does not supply it.

1979, the analysis method used has undergone many changes [*Reale*, 2001], four of which caused notable changes in the climatological averages (as we illustrate later): (1) In 1988, the original statistical retrieval method [*Smith et al.*, 1979] was replaced by a new physical-statistical method [*Smith et al.*, 1984]; (2) in 1992–1994, the first guess was changed to include cloudy situations, the use of the High-resolution Infrared Radiation Sounder (HIRS)water vapor channels was expanded and an improved cloud detection was introduced; (3) in late 1997 (but in mid-1998 in our data set), the physical-statistical retrieval was replaced by the RTOVS system which includes AVHRR-based cloud detection and improved treatment of satellite zenith angle dependence [*Reale et al.*, 1995]; and (4) in early 1999 (but in late 2001 in our data set), the RTOVS method was replaced by the ATOVS method to exploit the replacement of the Microwave Sounding Unit (MSU) by the AMSU system [*Reale et al.*, 1999], which added microwave water vapor profiles. In this study, we use a version of TOVS from ISCCP that is produced by relayering, regridding and filling with climatological data. We call this data set ISCCP-TOVS, which is also used in ISCCP retrieval of cloud parameters (and surface skin temperature) [*Rossow et al.*, 1996]. For the ISCCP-FD radiative flux calculations, we further modified the daily TOVS temperature profiles to introduce a diurnal variation of near-surface air temperatures for land [*Zhang et al.*, 2004]. The ISCCP-TOVS temperature/humidity profiles combined with the modified surface air temperature (and other parameters, see footnote b of Table 1) that are from ISCCP-FD are referred to as ISCCP-FD-TOVS, shortened as FDTV (called TV in most figures/tables), which is the fourth major data set we use for this work.

[9] Using the Improved Initialization Inversion (called 3I) retrieval algorithm, the Laboratoire de Météorologie Dynamique group has developed a new analysis that produced temperature and humidity profiles data set from TOVS (Path B) (consisting of HIRS, MSU and the Stratospheric Sounding Unit [see *Scott et al.*, 1999]). The 3I product provides atmospheric temperature/humidity profiles about twice-daily (for less cloudy locations) per satellite at spatial intervals of 100 km and is currently available for the time period 1987–1995 from one satellite.

[10] These five major data products also supply other parameters such as surface skin temperature and surface albedo, which are described in a companion paper (Zhang et al., submitted manuscript, 2005). Table 1 shows the original basic information.

[11] Table 1 also lists an additional humidity profile data set from the NASA Water Vapor Project (NVAP, called NT in figures/tables). The NVAP data set was produced by merging the radiosonde and TOVS profiles over land and TOVS with an analysis of the Special Sensor Microwave Imager (SSM/I) total precipitable water amounts over oceans covering the time period 1988–1997 [*Randel et al.*, 1996]. This product has a spatial interval of 1° × 1° with 3 layers (from surface to 300 hPa) and time interval of 1 day.

[12] To simplify, unless otherwise indicated, we restrict the comparisons to monthly means and, if available, monthly-hourly means for every sixth hour (UTC = 0000, 0600, 1200 and 1800), for January, April, July and October of

1992. Except for some long-term variations in the character of these data sets that we note below, this year is taken to be typical of these products.

2.2. Homogenization of Spatial Resolution and Derivation of Near-Surface Properties

[13] ERA15, GEOS-1, NCEP, FDTV, 3I and NVAP all have different horizontal and vertical spatial resolutions as well as other differences (Table 1), notably their rather different treatment of surface pressures (P_s). In order to compare all the temperature/humidity profiles and the near-surface properties that are derived (by extra/interpolation) from them, all data sets must be projected into one common 3-D coordinate system (while retaining fidelity with their original results). The common system we chose is the ISCCP-D1 system [see *Rossow and Schiffer*, 1991] with one slight modification of its vertical projection: instead of a varying tropopause pressure level, we fix all the pressure levels to a maximum of ten levels (nine layers) defined by: Surface ($\leq$1000 hPa), 800, 680, 560, 440, 310, 180, 70, 30 and 0 hPa, where surface is defined by the normalized P_s. The horizontal 2-D grid system is the ISCCP-D1 280-km equal-area map. The procedure for the homogenization of spatial resolution and derivation of near-surface properties is described in Appendix A.

2.3. Near-Surface Properties From Surface Observations

[14] Surface atmospheric temperature and humidity are also reported by the surface weather station network at more than 6000 stations, mostly over land, but also including ship reports. *Hahn and Warren* [1999] compiled and published this collection of surface synoptic weather reports in the Extended Edited Synoptic Cloud Reports Archive (here called surface observations, SOBS) that covers 1971–1996 for land and 1951–1997 for ocean [*Hahn and Warren*, 1999]. There are other, larger, collections available. These measurements are usually collected at 3-hour intervals and, for most continental areas (except desert, mountainous and permanent frozen areas), provide nearly complete spatial coverage for the land areas of the Northern Hemisphere with an average (minimum) station separation interval of <100 km (based on the four months from SOBS with about 6400 stations). The oceanic observations do not provide true diurnal resolution because the ships move (but different ships may cover the diurnal cycle at a given map grid cell if it is taken to be sufficiently large). This extensive data set was not used in the earlier versions of the weather analyses or reanalyses and does not seem to have been used much by researchers except for monitoring the long-term changes of global temperature. Since these surface station reports constitute direct in situ measurements of surface air temperature and humidity with good diurnal resolution, we use them as the reference standard for comparison of all the other temperature-humidity data sets.

2.4. Comparison of Temperature and Humidity Profiles

[15] To quantify the temperature profile differences among the five data sets, we have averaged the profiles over eight spatial domains: five latitude zones (90–65°S, 65–25°S, 25°S–25°N, 25–65°N and 65–90°N), both hemispheres and the globe, separately for land and ocean (including near-coastal areas), for each of the four months and then differenced these profiles with respect to FDTV. The mean profiles from all five data sets agree reasonably well: differences range from 2 K in the lower troposphere to 2 K in the mid to upper troposphere. Figures 1a and 1b show the averaged differences of the temperature profiles over land in wintertime and ocean in summertime in the Northern Hemisphere. Although the hemispheric mean differences over land are small at all levels, where all the products use or are tuned to radiosondes (Figure 1a), the largest discrepancies appear over wintertime land areas at high latitudes where significant surface temperature inversions occur (not shown). The worst disagreements appear over wintertime Antarctica, where the near-surface-layer temperature difference are as large as 10 K for 3I minus FDTV, >13 K for 3I minus NCEP and >6 K for GEOS-1/NCEP versus ERA15. The explanation is that the 3I profile

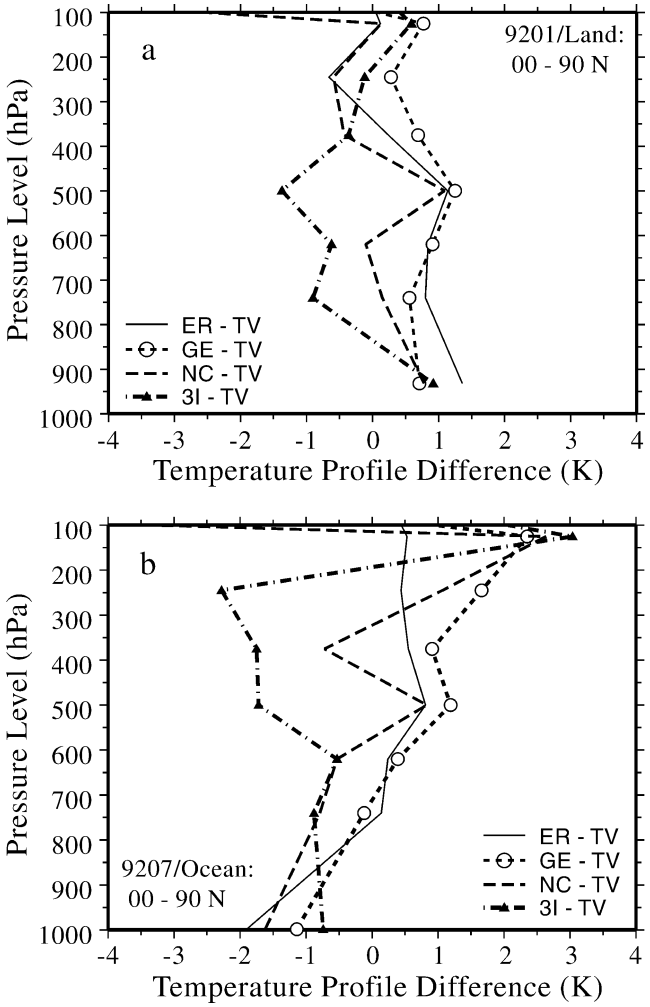

Figure 1. (a) Northern Hemispheric averages of monthly mean temperature profile differences (in K) for ERA15 (ER), GEOS-1 (GE), NCEP (NC) and 3I minus FDTV (TV), respectively, for land in wintertime (January 1992) in the Northern Hemisphere. (b) Same as Figure 1a but for ocean in summertime (July 1992).

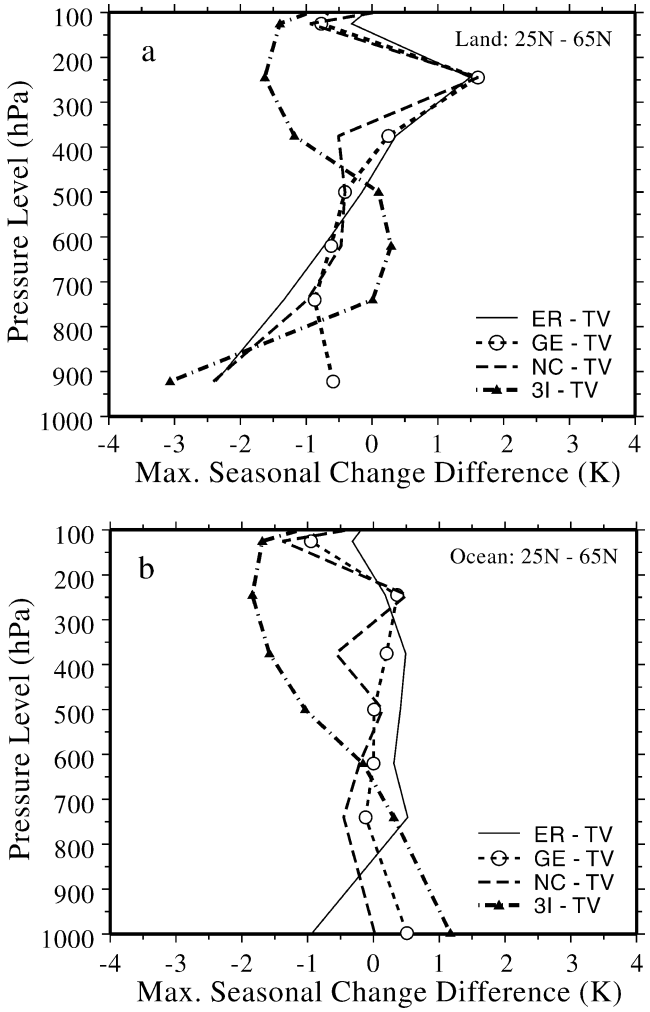

Figure 2. (a) Average differences of the maximum seasonal change of monthly mean temperature profile for ERA15, GEOS-1, NCEP and 3I minus FDTV, respectively, averaged over land of 25–65°N. (b) Same as Figure 2a but for ocean.

shows no inversions and the ERA15 profile shows only a very weak inversion (not shown). The second-largest differences, as large as >5 K (GEOS-1/NCEP/ERA15 versus 3I), appear in the tropical (25°S–25°N) upper troposphere over oceans. Over oceans, where much fewer radiosonde data are available, the reanalyses exhibit systematically shallower temperature lapse rates (Figure 1b) relative to the two TOVS-based results (e.g., the lapse rate for 1000–800 hPa is 5.6 K per 10^4 Pa for TOVS/3I while 4.8, 4.8 and 5.3 K per 10^4 Pa for NCEP, ERA15 and GEOS-1, respectively), which could be associated with the concentration of TOVS and 3I results in less cloudy locations. The differences for all other zones and seasons are smaller than these extreme values, except for summertime polar regions.

[16] The global mean differences of the seasonal cycle amplitudes are generally 2 and 1 K for land and ocean, respectively, except in the polar regions where the seasonal amplitude differences can be >8 K (NCEP/3I versus FDTV). Figures 2a and 2b show the mean differences over land and ocean, respectively, of the seasonal range of temperature at 25–65°N, where the seasonal variation is largest. The largest differences appear in the near-surface and upper tropospheric layers over land: >2–3 K. The three reanalyses are more similar to each other than they are to either TOVS or 3I, whereas TOVS and 3I are more similar to each other than to any reanalysis.

[17] We have conducted similar studies of the precipitable water (PW) profiles from the six data sets (but only for the five lower layers since NVAP and FDTV do not report values above the 300 hPa level). For the global mean PW profiles, the mid to low tropospheric differences range from ~2–3 mm up to 3.5 mm for land and ocean, respectively. Since the tropical areas are most important for PW profiles, Figures 3a and 3b show the mean profile differences (all minus FDTV) for the tropics (25°S–25°N) for January 1992 for land and ocean, respectively, where the largest differences appear in the midtroposphere over land (2 mm, >30% of the mean layer amount) and in the near-surface layers over ocean (as large as ≈5 mm, ≈25% of the layer amount), respectively. Note that the agreement among these PW data

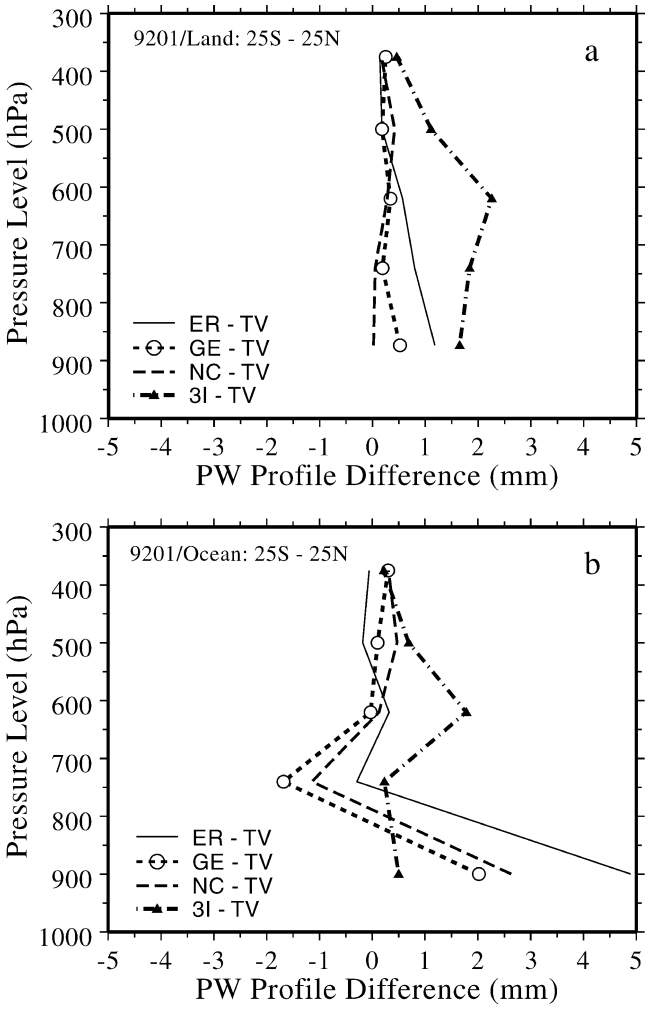

Figure 3. (a) Averages of monthly mean precipitable water profile differences (in mm) for ERA15, GEOS-1, NCEP and 3I minus FDTV, respectively, averaged over land of 25°S–25°N for January 1992. (b) Same as Figure 3a but for ocean.

Table 2. Comparison of Monthly Mean Surface Air Temperature From ERA15, GEOS-1, NCEP, FDTV and 3I for the Four Seasonal Months of 1992[a]

	Jan 1992		Apr 1992		Jul 1992		Oct 1992	
X/Y	Ocean	Land	Ocean	Land	Ocean	Land	Ocean	Land
ER/GE	−0.31 (1.5)	0.73 (2.6)	−0.01 (1.6)	1.21 (2.8)	−0.09 (1.9)	0.69 (2.9)	−0.71 (1.8)	0.96 (2.7)
ER/NC	−0.18 (1.1)	0.45 (1.8)	0.03 (1.2)	1.69 (3.3)	−0.13 (1.4)	1.51 (3.7)	−0.27 (1.1)	1.02 (2.4)
ER/TV	−1.52 (1.9)	0.33 (3.6)	−1.03 (1.9)	1.32 (4.3)	−1.09 (2.1)	2.45 (3.7)	−1.56 (2.0)	1.29 (3.1)
ER/3I	−1.26 (2.3)	0.89 (4.0)	−0.90 (2.0)	1.50 (3.5)	−1.42 (2.8)	0.68 (3.5)	−1.53 (2.6)	0.49 (3.2)
GE/NC	0.13 (1.4)	−0.28 (2.0)	0.05 (1.1)	0.48 (2.5)	−0.04 (1.4)	0.82 (3.0)	0.43 (1.6)	0.07 (2.3)
GE/TV	−1.21 (1.8)	−0.39 (4.0)	−1.02 (2.0)	0.10 (5.4)	−1.00 (2.0)	1.76 (4.0)	−0.86 (1.8)	0.33 (3.4)
GE/3I	−0.95 (1.8)	0.16 (4.5)	−0.89 (1.8)	0.29 (4.5)	−1.33 (2.4)	−0.01 (4.7)	−0.82 (2.2)	−0.47 (4.3)
NC/TV	−1.34 (2.0)	−0.11 (3.7)	−1.07 (2.1)	−0.37 (6.1)	−0.97 (2.0)	0.94 (4.0)	−1.29 (2.0)	0.27 (2.7)
NC/3I	−1.08 (2.2)	0.44 (4.1)	−0.94 (1.9)	−0.19 (5.2)	−1.29 (2.5)	−0.83 (6.0)	−1.26 (2.5)	−0.53 (4.0)
TV/3I	0.26 (1.5)	0.56 (3.6)	0.13 (1.8)	0.19 (4.3)	−0.33 (1.8)	−1.77 (4.9)	0.03 (1.6)	−0.80 (3.5)
abs mean	0.82 (1.8)	0.43 (3.4)	0.61 (1.7)	0.74 (4.2)	0.77 (2.0)	1.15 (4.0)	0.88 (1.9)	0.62 (3.2)

[a]The shown numbers are global land/oceanic means and standard deviations of regional differences (X − Y) on 280-km equal-area grid map in Kelvin. Ocean areas include coast cells (defined as area > 33% of water; otherwise, land). "ER," "GE," "NC," "TV" and "3I" are for ERA15, GEOS-1, NCEP, FDTV and 3I, respectively. The ten X/Y pairs are all their possible combinations. The last row ("abs mean") is the modulus (i.e., absolute value) average of the mean differences (standard deviation) of the ten pairs, and it is not affected by the X/Y assignment. When averaging over all the four months of the last row, we have the pseudoannual/all-pair mean-modulus difference (standard deviation) = 0.77 (1.9) and 0.73(3.7) K, for ocean and land, respectively. For the global (ocean and land) mean, it becomes 0.53 (2.6) K.

sets is much better over land, where they all use the radiosondes or have been tuned to the radiosondes (FDTV); the 3I profile, which has not been tuned to radiosondes, shows a moister tropical middle troposphere [cf. *Escoffier et al.*, 2001]. The hemispheric mean differences for the seasonal range are up to ≈2 mm (roughly 15% of the seasonal range) or so for near-surface-layer PW for both land and ocean (not shown).

2.5. Comparison of Surface Air Temperatures

[18] The five (monthly mean) surface air temperature (T_a) data sets produce ten different pairs. The spatial correlation coefficients from the ten pairs of monthly mean maps are all >0.96 with an average value = 0.98 (ocean is slightly better than land). However, such large correlations may be misleading because a substantial part of them are caused by equator-to-polar tendency (compare daily time series correlation based on calculation from each grid cell). Table 2 shows the global average (rms = RMS difference with bias removed throughout the paper) regional differences for all ten pairs for land and ocean, respectively, for four monthly means. For a more compact summary, we determine the average and RMS for the moduli of the global monthly mean differences (i.e., absolute values of the mean differences, not affected by subtraction order [see, e.g., *Reynolds*, 1988]). The overall modulus mean (rms) difference for T_a for all ten combinations is 0.53 (2.6) K, 0.73 (3.7) and 0.77 (1.9) for land and ocean, separately (Table 2 footnotes), indicating that the land and ocean differences partially cancel. The differences over land have larger scatter than over ocean. The largest regional differences occur on or near Antarctica in April: >40 and 30 K for land (FDTV versus 3I) and ocean (FDTV versus NCEP and GEOS-1), respectively.

[19] Figures 4a and 4b show the zonal mean differences (all minus FDTV) for the monthly mean T_a for April, 1992, for land and ocean, respectively. The largest differences appear in the polar regions: up to >15 K (near Antarctica in Figure 4a). For the other zones, the differences are generally 5 K and 3 K for land and ocean, respectively. Comparable differences also appear in the other three months (not

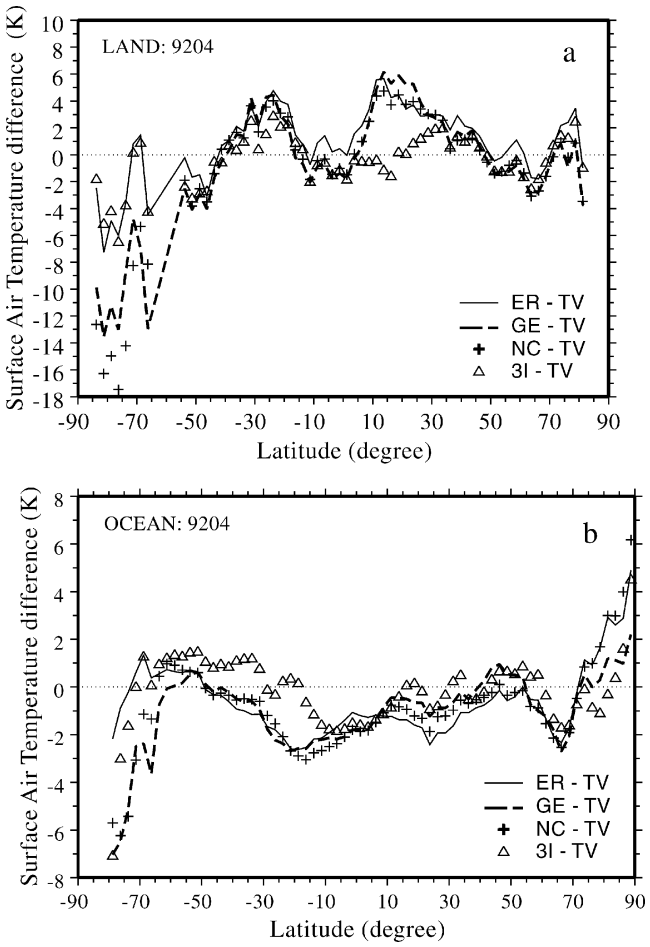

Figure 4. (a) Differences of zonal monthly mean surface air temperature (in K) for ERA15, GEOS-1, NCEP and 3I minus FDTV, respectively, for land and April 1992. (b) Same as Figure 4a but for ocean.

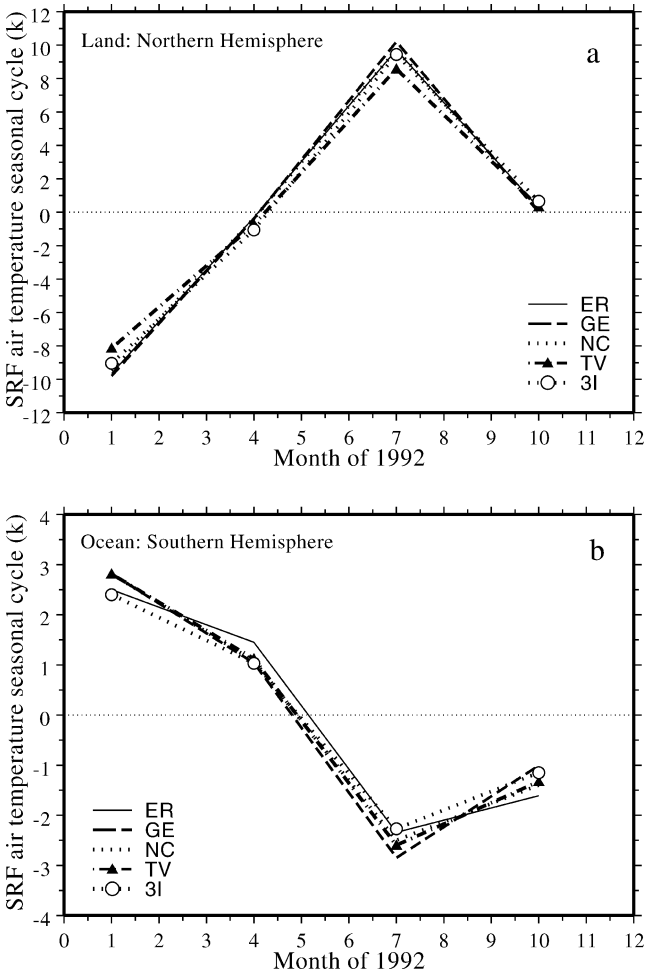

Figure 5. (a) Average seasonal cycle of surface air temperatures (in K) for the Northern Hemispheric land for ERA15, GEOS-1, NCEP, 3I and FDTV, shown as deviations of individual monthly means from their averages. (b) Same as Figure 5a but for the Southern Hemispheric ocean.

shown). Note that, especially over land, the three reanalyses are generally more similar to each other than they are to either the FDTV or 3I which are derived more directly from satellite sounder observations.

[20] Using the four monthly mean maps to represent the seasonal cycle of T_a shows that all the data sets exhibit very similar seasonal variations (Figure 5), with hemispheric mean differences in amplitude being only about 2 K and 0.5 K for land and ocean, respectively. It is notable that the differences among these data sets are slightly larger in the southern than in the Northern Hemisphere. Zonal mean seasonal amplitudes differ somewhat more, up to 10 K (between NCEP and 3I for January) and 5 K (between GEOS-1 and FDTV for July) for the southern and northern polar land regions; over the ocean, the largest zonal mean differences in seasonal amplitude are about 2.5 K (between NCEP and ERA15) and 1 K (between GEOS-1 and FDTV for January), respectively.

[21] At the time this study began, there were only three data sets with 6-hourly values of T_a available: ERA15, GEOS-1 and FDTV (the original TOVS has approximately daily time resolution). Figure 6 illustrates the seasonal variation of the hemispherical mean diurnal cycles (with respect to their own monthly-daily mean) over Northern Hemisphere land (January and July). FDTV (which is based on the SOBS climatology, see Table 1, footnote b, also *Zhang et al.* [2004]) exhibits a peak temperature at 1200 local time (LT) while the other two data sets peak at 1800 LT. Since we cannot accurately determine the peak times because of the 6-hour interval, these shapes should be interpreted to mean that the peak time for FDTV is closer to 1200 LT, while that of the two reanalyses is closer to 1800 LT. FDTV also has larger diurnal amplitudes, up to ~3 K, than the other two, <1.5 K (compare section 2.8).

2.6. Comparison of Near-Surface-Layer Precipitable Water Amounts

[22] With NVAP, we have six precipitable water (PW) data sets. Figure 7 illustrates the zonal mean differences (all minus FDTV) for the near-surface-layer PW amount (surface to 800 hPa, PW-1) from the six data sets over land and

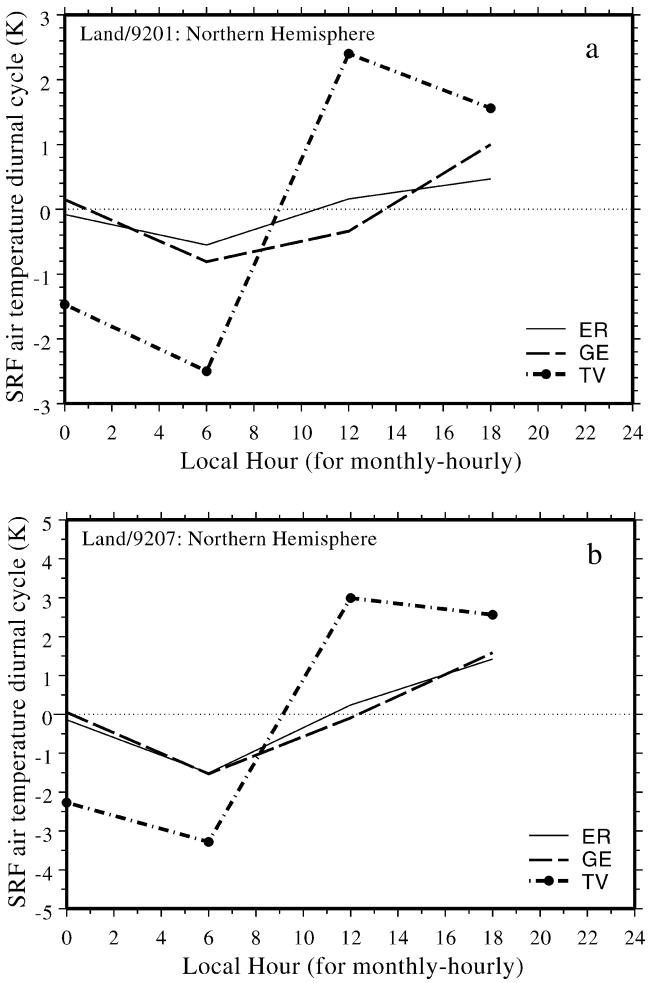

Figure 6. (a) Average diurnal cycle for monthly mean-hourly surface air temperature (in K) for ERA15, GEOS-1 and FDTV for the Northern Hemisphere land for January 1992, shown as deviations from their individual monthly averages over time of day. (b) Same as Figure 6a but for July 1992.

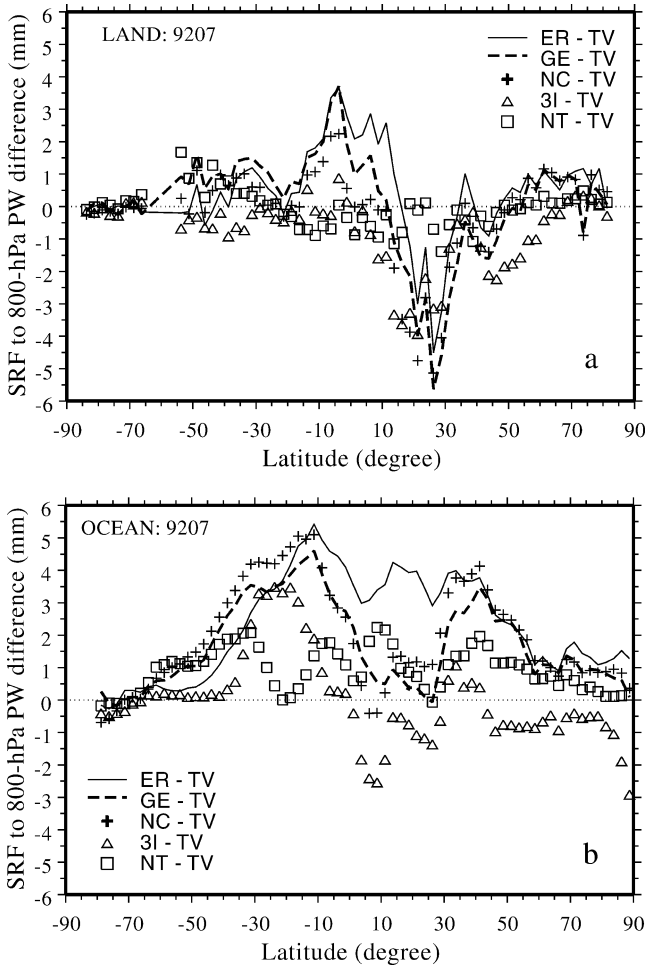

Figure 7. (a) Difference of zonal monthly mean surface-to-800-hPa precipitable water amounts (in mm) for ERA15, GEOS-1, NCEP, 3I and NVAP (NT) minus FDTV, respectively, for land July 1992. (b) Same as Figure 7a but for ocean.

ocean for July. There are substantial differences in the tropical and subtropical regions: up to ~6 mm (>30% of the total). For most tropical zones, ERA15 has the largest values (wettest); over the Sahara the three reanalyses are much drier than the sounder-based products; the three reanalyses are generally wetter than the sounder products over subtropical oceans. These grouped differences (reanalysis versus sounder) may be indicative of the land-ocean difference of the mix of radiosonde and satellite information. The overall modulus mean (rms) difference for PW-1 from all 15 different combinations of the six data sets is 0.10 (0.23) cm, or 7 (16)% of their mean PW-1 amount, 0.04 (0.25) cm and 0.13 (0.21) cm [or 4 (26)% and 8 (13)% of mean PW-1 amount] for land and ocean, separately.

[23] More detail from the comparisons is given in Table 3 that shows the surface-310 hPa PW column amounts (≈total column PW) for all fifteen possible combinations of the six data sets. Comparison of zonal mean total column PW differences (versus FDTV as above) indicates that the relative difference magnitudes are comparable to the differences of the near-surface-layer PW. However, ERA15 no longer has the largest difference with FDTV everywhere; instead, 3I has the largest difference for tropical land areas (not shown).

2.7. Synoptic Variations

[24] To compare how similarly synoptic variations are represented in these data sets, we collected daily averages (ERA15, GEOS-1 and NCEP) or just daily values (3I, FDTV and NVAP) for all 123 days in January, April, July and October, 1992, for all data sets. The magnitude of the synoptic variations at each location is summarized by the RMS differences between daily and monthly mean values for each month. The global averages of the synoptic variations of T_a are between 2.6 K (FDTV) and 2.9 K (ERA15 and 3I) with an overall mean = 2.8 K. If land and ocean are separated, the overall mean values are 3.7 (range of 3.5–4.0 K) and 2.4 K (range of 2.1–2.7 K), respectively. Therefore the magnitudes of the synoptic variations with

Table 3. Comparison of Monthly Mean Surface-to-310 hPa (≈Total Column) Precipitable Water (PW) From ERA15, GEOS-1, NCEP, FDTV, 3I and NVAP for the Four Seasonal Months of 1992[a]

	Jan 1992		Apr 1992		Jul 1992		Oct 1992	
X/Y	Ocean	Land	Ocean	Land	Ocean	Land	Ocean	Land
ER/GE	0.22 (0.3)	0.04 (0.3)	0.24 (0.4)	0.09 (0.3)	0.12 (0.3)	0.10 (0.2)	0.11 (0.3)	0.04 (0.2)
ER/NC	0.06 (0.3)	0.04 (0.2)	0.14 (0.5)	0.08 (0.3)	−0.01 (0.4)	0.08 (0.3)	−0.04 (0.3)	0.03 (0.2)
ER/TV	0.20 (0.3)	0.03 (0.4)	0.14 (0.4)	−0.04 (0.5)	0.07 (0.3)	−0.04 (0.4)	−0.08 (0.2)	−0.07 (0.4)
ER/3I	0.10 (0.3)	−0.21 (0.4)	0.26 (0.4)	−0.16 (0.4)	0.15 (0.3)	0.01 (0.5)	0.07 (0.3)	−0.09 (0.3)
ER/NT	0.13 (0.3)	0.02 (0.4)	0.15 (0.3)	−0.02 (0.4)	0.01 (0.2)	−0.03 (0.4)	−0.02 (0.2)	−0.04 (0.3)
GE/NC	−0.16 (0.3)	0.00 (0.2)	−0.10 (0.3)	−0.01 (0.2)	−0.13 (0.3)	−0.03 (0.3)	−0.15 (0.3)	−0.01 (0.2)
GE/TV	−0.03 (0.3)	−0.01 (0.3)	−0.11 (0.4)	−0.13 (0.5)	−0.05 (0.3)	−0.15 (0.5)	−0.19 (0.4)	−0.11 (0.3)
GE/3I	−0.12 (0.4)	−0.25 (0.4)	0.02 (0.4)	−0.25 (0.5)	0.04 (0.4)	−0.10 (0.5)	−0.04 (0.4)	−0.13 (0.4)
GE/NT	−0.10 (0.3)	−0.02 (0.3)	−0.09 (0.3)	−0.12 (0.4)	−0.11 (0.3)	−0.14 (0.4)	−0.13 (0.3)	−0.09 (0.3)
NC/TV	0.13 (0.3)	−0.01 (0.3)	0.00 (0.4)	−0.12 (0.5)	0.08 (0.3)	−0.12 (0.5)	−0.04 (0.3)	−0.10 (0.3)
NC/3I	0.04 (0.4)	−0.25 (0.5)	0.12 (0.4)	−0.24 (0.5)	0.17 (0.4)	−0.07 (0.5)	0.11 (0.4)	−0.12 (0.4)
NC/NT	0.06 (0.3)	−0.02 (0.3)	0.01 (0.4)	−0.10 (0.4)	0.02 (0.4)	−0.11 (0.4)	0.02 (0.3)	−0.08 (0.3)
TV/3I	−0.10 (0.3)	−0.24 (0.4)	0.13 (0.3)	−0.12 (0.5)	0.09 (0.2)	0.05 (0.5)	0.15 (0.3)	−0.02 (0.4)
TV/NT	−0.07 (0.3)	−0.01 (0.2)	0.01 (0.3)	0.02 (0.3)	−0.06 (0.2)	0.01 (0.3)	0.06 (0.3)	0.03 (0.2)
3I/NT	0.03 (0.4)	0.23 (0.4)	−0.11 (0.3)	0.13 (0.5)	−0.15 (0.3)	−0.04 (0.4)	−0.09 (0.3)	0.05 (0.4)
ab mn	0.10 (0.3)	0.09 (0.3)	0.11 (0.4)	0.11 (0.4)	0.08 (0.3)	0.07 (0.4)	0.08 (0.3)	0.07 (0.3)

[a]Same conventions as explained for Table 2 except for PW (cm). The pseudoannual/all-pair mean-moduli (rms) differences are 0.10 (0.3) and 0.09 (0.4), for ocean and land, respectively. For global (all land and ocean), it is 0.08 (0.3), where "NT" is for NVAP using FDTV surface pressure in relayering PW.

Table 4a. Definition of Six Regions

Region	Latitude Range	Longitude Range
Central USA	32.5°N → 45.0°N	105.0°W → 90.0°W
Central Asia	30.0°N → 50.0°N	55.0°E → 105.0°E
Brazil	20.0°S → 0.0°	65.0°W → 45.0°W
North Pacific	30.0°N → 60.0°N	150.0°E → 135.0°W
South Pacific	60.0°S → 30.0°S	180.0°W → 90.0°W
West Tropical Pacific	15.0°S → 15.0°N	120.0°E → 180.0°E

respect to the monthly averages are comparable in all the data sets. The global average of the time record mean (rms) differences for each map grid cell (i.e., 123 days for each grid cell) between each data set and FDTV ranges from 0.1 (3.6) K (3I - FDTV) to −0.8 K (3.4) (NCEP - FDTV). The overall average difference is 0.6 (5.1) and −0.8 (2.9) K for land and ocean, respectively. These values indicate that the differences in the daily variations with respect to FDTV (and each other) are somewhat larger than the daily variations themselves. However, the fact that the global mean correlation coefficients of the daily time series from 10 possible pairs range from 0.54 (FDTV and 3I) to 0.86 (ERA15 and NCEP) with overall average = 0.66, and 0.74 and 0.64 for land and ocean, respectively, means that all the data sets have some skill in representing the synoptic variations. That the mean difference for land is less than for ocean seems counterintuitive, but it may be explained by the very large difference in the number of radiosondes used between land and ocean since, except for 3I, all the temperature profiles are somehow calibrated to radiosondes (primarily operated on land, and T_a is extrapolated from the profile). It is also reflected in the monthly mean comparison (section 2.5, where the mean differences are 0.73 and 0.77 for land and ocean, respectively, but Table 2 shows the opposite for most 3I-related pairs) as well as higher correlation value for land as just stated above. In other words, land T_a values are more correlated (and surface-observation based) than ocean for four of five data sets. However, that the ocean has much less variation (rms difference) than land is realistic. Thus, in terms of total uncertainty (mean and RMS difference), land T_a values still are more uncertain than ocean values.

[25] In contrast to T_a, the synoptic variations for PW-1 show more diversity and disagreement, in which 3I stands out. The overall average of the global mean RMS synoptic variations is 25 mm, 40 mm for 3I but <25 mm for the other four. The unusually large variations in 3I appears mostly over oceans. However, the mean difference between 3I and FDTV is smaller than between FDTV and the three reanalyses [cf. *Escoffier et al.*, 2001]. The global mean temporal correlation coefficients are all >0.51 except for all the pairs with 3I present (<0.46). The lowest correlations (<0.40) are with 3I over oceans; all other data set pairs have correlations >0.49 over oceans. 3I also exhibits the largest variations in the total-column PW (rms = 74 mm versus 38–46 mm for the other data sets).

2.8. Comparison of FDTV and SOBS

[26] In previous sections we have compared all the other data sets to FDTV. Here we compare FDTV to the best "truth" available, the surface measurements (SOBS) of temperature and humidity made at surface weather stations and by ships at more than 6000 locations. To make an appropriate comparison, we first spatially and temporally average and grid the SOBS station/ship-observed surface air temperature and specific humidity to our standard 280-km equal-area map and 3-hour time window for eight GMTs (0000, 0300, . . ., 2100). The overall monthly mean values are determined from the four monthly hourly means for each month (January, April, July and October, 1992). Since FDTV only gives layer PW, we convert the near-surface-layer (surface to 800 hPa, about 130 hPa thick in average for land) PW to an estimate of the mean surface layer specific humidity, which is associated with but usually smaller than surface specific humidity. Because SOBS data are globally (and diurnally) incomplete, the comparisons T_a as well as surface specific humidity/mean surface layer specific humidity are conducted for averages on the basis of matched grid cells/GMTs for FDTV and SOBS.

[27] Table 4b compares the collected (monthly hourly mean based) monthly means of T_a for the global map and the six regions defined in Table 4a. The global mean (rms) difference (FDTV minus SOBS) is −0.2 (3.4) K with normal deviation (with respect to the best fit straight line) = 2.4 K, comparable to the overall difference = 0.53 (2.6) K among the five data sets (section 2.5). The correlation coefficient is 0.95. Given differences between the space-time sampling of SOBS and FDTV, as well as other errors (see below), their agreement is within their uncertainties. Figure 8 shows the scatterplot of the monthly mean T_a values from these two data sets. The points in the upper left corner with SOBS > 295 K and FDTV $\lesssim$ 273 K are located over the latitudinal zones from 88°S to 54°S, essentially sea ice and coastal snow, very possibly caused by erroneous ship reports contained in SOBS. On the other hand, the points in the lower left corner with FDTV < 260 K and SOBS < 250 K are caused by FDTV underestimates of the strength of wintertime temperature inversions at the surface that leads to interpolated T_a values higher than SOBS by up to 10 K or so in snow covered regions.

[28] Figures 9a and 9b show the (global mean) diurnal cycles of monthly hourly mean T_a (against local hours) for

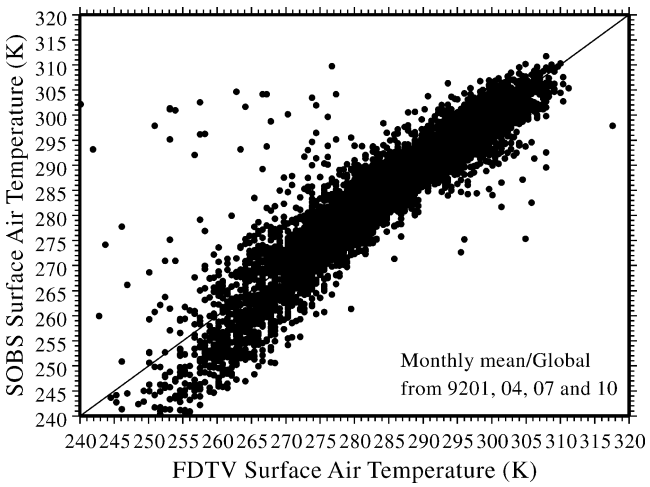

Figure 8. Scatterplot of monthly mean surface air temperatures (in K) from FDTV and SOBS accumulated for January, April, July and October 1992.

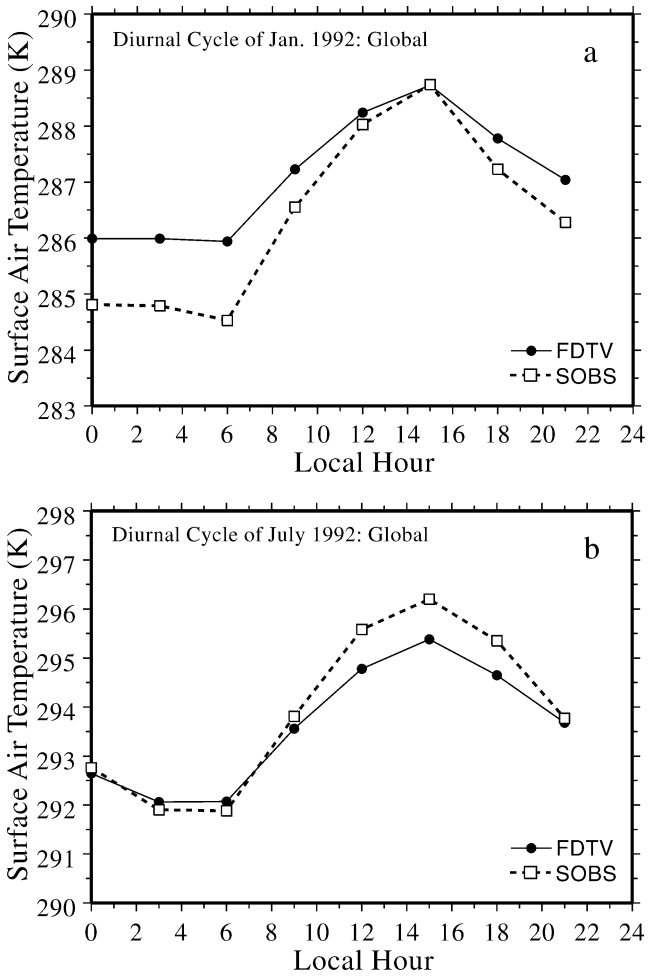

Figure 9. (a) Comparison of the global mean diurnal cycle of surface air temperature (in K) from FDTV and SOBS for January 1992. (b) Same as Figure 9a but for July 1992.

January and July 1992, respectively. The largest discrepancy is about 1 K or so, indicating good agreement for diurnal variations (by design since FDTV is based on a SOBS climatology of the diurnal variations [see *Zhang et al.*, 2004]). Note, however, that FDTV tends to overestimate nighttime T_a in winter and underestimate daytime T_a in summer, likely because of systematic differences between the lapse rates of the free troposphere and atmospheric boundary layer and a possible overestimate of cloud effects on the diurnal temperature extremes.

[29] Table 4b also shows the comparison of T_a for the six regions. The West Tropical Pacific ocean and Brazil show the smallest mean differences (0.2 and 0.1 K, respectively) and also the lowest correlation coefficients (0.4 and 0.3) because both of them are tropical regions that exhibit smaller variations: standard deviations over Brazil are much smaller for FDTV (2.6 K) than for SOBS (5.9 K). The RMS differences for both regions are about the same as the SOBS variations.

[30] Similarly, Table 4c shows the comparison for mean surface layer specific humidity of FDTV and surface specific humidity of SOBS. Again, except for the two tropical regions, all the correlation coefficients are >0.74. The normal deviation about the best fit straight line for the global comparison is 1.4 g/kg, about 14% of the mean values of the FDTV and SOBS, comparable to the previous estimate of RMS differences (16%) for the near-surface-layer PW-1 (section 2.6). Figure 10 shows the global scatterplot, in which, there are a few points with values >20 g/kg for SOBS and <4 g/kg for FDTV that are located in high latitudinal zones >60°S or 70°N, primarily in sea ice and coastal snow covered areas, possibly caused by mislocated SOBS reports.

[31] Using SOBS as the reference standard, the comparison for all the five data sets for the four months show overall quantitative agreement of global monthly mean T_a values to within a few K: FDTV and 3I agree slightly better with SOBS in the global mean than the three reanalyses (all differences $\lesssim$1 K, not shown). All five T_a data sets generally agree well with SOBS (to within a few degrees rms) in the range of temperatures from 265 to 305 K; FDTV and 3I are systematically warmer than the reanalyses over oceans and somewhat colder over land except in wintertime. All three reanalyses underestimate T_a in the warmest regions but overestimate it in the coldest regions in the Northern Hemisphere. Neither FDTV nor 3I do any better in the coldest locations but FDTV provides the best estimate of T_a in the warmest locations where 3I has slightly larger departures from SOBS than FDTV. The SOBS values of T_a suggest that the FDTV and 3I values are slightly better over oceans, but still a slight underestimate particularly at higher latitudes. Over land the FDTV and 3I values are usually biased low with respect to SOBS and the three reanalyses are usually biased high. Regional differences are larger. All the data sets tend to overestimate wintertime

Table 4b. Comparison of Monthly Mean Surface Air Temperature Between FDTV and SOBS for the Four Seasonal Months of 1992[a]

Region	FDTV	SOBS	M Diff	Stdv	Correlation Coefficient	Slope	Intercept	Nm Dev	Cells
Global	291.1 (10.7)	291.3 (11.3)	−0.19	3.38	0.954	0.997	1.04	2.40	19,557
Central USA	286.1 (6.8)	285.5 (8.4)	0.61	3.47	0.917	1.135	−39.33	2.21	68
Central Asia	280.2 (11.8)	281.5 (12.1)	−1.32	5.17	0.907	0.938	18.83	3.73	404
Brazil	298.9 (2.6)	298.8 (5.9)	0.12	5.71	0.301	0.694	91.25	4.65	157
North Pacific	283.3 (7.5)	284.6 (7.0)	−1.26	2.08	0.961	0.901	29.45	1.45	960
South Pacific	285.3 (4.8)	286.7 (4.7)	−1.46	2.20	0.892	0.884	34.60	1.60	474
West Tropical Pacific	301.0 (1.5)	300.8 (1.3)	0.23	1.49	0.435	0.239	228.74	1.11	977

[a]The monthly mean values are based on monthly hourly mean from matched 280-km equal-area cells and GMTs for FDTV and SOBS (see text). The statistical values are based on the coordinates of points in a scatterplot with SOBS values on the ordinate and FDTV values on the abscissa. Regression statistics are from a linear least squares fit to the scatter of points. All values are in K, except the correlation coefficients and slopes, which are unitless. The first two value columns are mean (spatial standard deviation) for FDTV and SOBS, respectively. "M Diff" ("Stdv") is mean (rms) difference between FDTV and SOBS, and "Nm Dev" is the RMS distance of all the points from the regression line.

Table 4c. Comparison of Monthly Mean Surface Layer Mean Specific Humidity From FDTV and Surface Specific Humidity From SOBS for the Four Seasonal Months of 1992[a]

Region	FDTV	SOBS	M Diff	Stdv	Correlation Coefficient	Slope	Intercept	Nm Dev	Cells
Global	8.4(4.1)	11.8(5.8)	−3.33	2.57	0.918	1.287	0.91	1.40	18,474
Central USA	6.2(3.1)	7.3(3.9)	−1.06	1.18	0.972	1.237	−0.41	0.58	67
Central Asia	5.8(3.6)	5.6(3.9)	0.15	2.69	0.740	0.800	1.00	2.03	331
Brazil	12.4(1.7)	16.8(3.5)	−4.36	2.81	0.612	1.258	1.15	1.73	140
North Pacific	5.1(2.3)	7.6(3.4)	−2.52	1.43	0.943	1.390	0.52	0.65	941
South Pacific	5.0(1.5)	7.9(2.6)	−2.91	1.80	0.749	1.302	1.39	1.06	436
West Tropical Pacific	13.3(1.6)	18.6(1.6)	−5.25	1.54	0.548	0.561	11.10	1.19	939

[a]Same as Table 4b except now it is for specific humidity (in g/kg). Note that FDTV (from ISCCP-FD full-sky PW profile, slightly different from the original TOVS [see *Zhang et al.*, 2004]) presents mean specific humidity of the near-surface layer (from surface to 800 hPa, usually around 130 hPa thick for land and up to 200 hPa thick for coean) while SOBS surface (air) specific humidity.

temperatures by failing to properly represent the near-surface inversions. Seasonal and synoptic variations are very similar in all the data sets but FDTV synoptic variability agrees with SOBS slightly better than the reanalyses. The reanalyses systematically underestimate the diurnal temperature amplitude and overestimate the lag with the sun, whereas FDTV used in ISCCP-FD captures the diurnal cycle well (as it was designed to do).

2.9. Comparison of Difference Between Surface Skin and Air Temperature

[32] Since it is the difference of surface skin temperature T_s (see Zhang et al., submitted manuscript, 2005), and surface air temperature T_a that controls the net longwave (LW) at the surface, we also compare the differences, $T_s - T_a$, from FDTV with those from 3I and the three reanalyses. Figure 11 shows zonal mean values ($T_s - T_a$) for land and ocean, respectively, for April (we chose the month with more scatter in both the polar regions). The discrepancies in ($T_s - T_a$) are very large in the polar regions, up to >15 K (ERA15 versus 3I for land and GEOS-1 versus 3I for ocean). In other zones, the values of ($T_s - T_a$) are mostly <2–3 K except for FDTV, especially in the subtropical deserts. In general, comparable magnitudes also appear in other months (not shown). A curious feature in Figure 11a is that all the reanalyses show negative values of ($T_s - T_a$), instead of the positive values that are expected for sensible heat transfer from the surface to the air over most nonpolar regions. This suggests that the T_a values that we have inferred for these data sets by extrapolation from the temperature profiles are biased slightly high (or possibly that the model-based skin temperatures are too low); if this is the case, then the comparisons of T_a values in Table 2 would be changed to indicate a low bias relative to FDTV over both land and ocean. This possibility is supported by Zhang et al. (submitted manuscript, 2005, Figure 7), which suggests that the FDTV values of T_s are anomalously low during 1992, which would result from the ISCCP retrieval with TOVS values of T_a being too high. Such a situation would also explain why the FDTV values of ($T_s - T_a$) over ocean are nearly zero instead of being slightly positive. All of these effects are associated with biases in the 1992 FDTV temperatures of only 1–2°.

3. Summary and Discussion

[33] Estimating the surface radiation budget has long been pursued [*Simpson*, 1929] because, together with latent and sensible heat fluxes, it composes one fundamental coupling of the atmosphere with the ocean and land surfaces. Direct estimates that resolve regional and weather-scale variability with reasonable accuracy have only become possible with the advent of complete global, mostly satellite, data sets within the past couple of decades. *Zhang et al.* [2004] now estimate that surface radiative fluxes can be determined to within about 10–15 W/m^2 (regional monthly mean flux values), improved by about 5 W/m^2 over *Rossow and Zhang* [1995]. The main limitation is still the accuracy of the input data sets. As stated in Introduction (section 1), sensitivity studies using ISCCP-based estimates ("realistically assumed") of the input uncertainties show that the leading uncertainties in the surface fluxes are no longer predominately associated with clouds but are now more associated with uncertainties in the near-surface atmospheric and the surface properties.

[34] This study presents a fuller, more quantitative evaluation of these uncertainties for the near-surface air temperature and humidity by comparing the main available global data sets that are treated as ensemble realizations of actual climate such that their differences represent an estimate of the uncertainty in their values. The results are globally representative and may be taken as a generalization

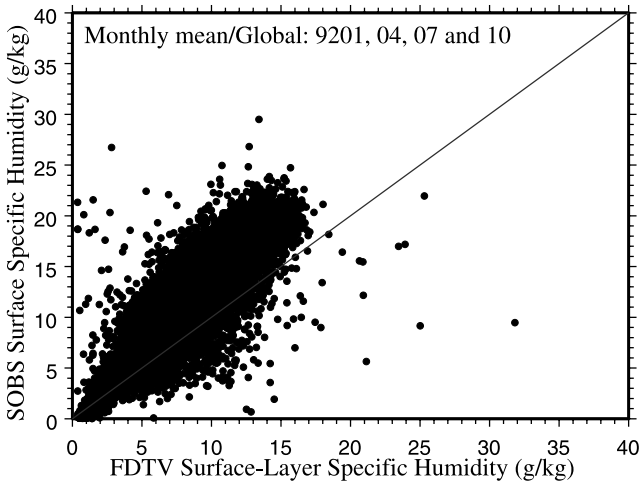

Figure 10. Scatterplot of monthly mean surface-to-800-hPa mean specific humidity from FDTV and surface specific humidity from SOBS (both in g/kg) accumulated for January, April, July and October 1992.

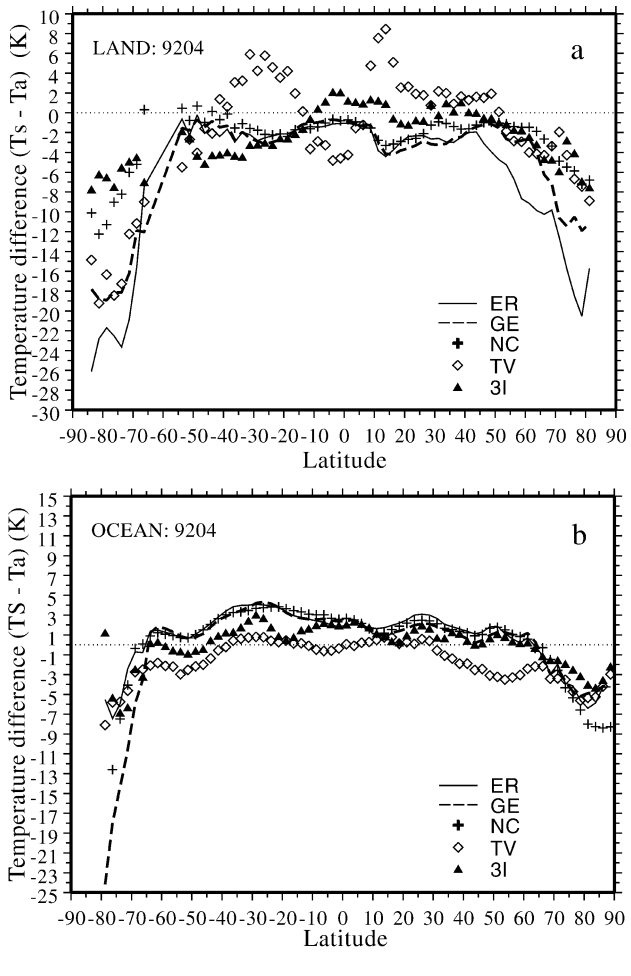

Figure 11. (a) Zonal monthly mean differences between surface skin and air temperatures (in K) for ERA15, GEOS-1, NCEP, FDTV and 3I, respectively, for land for April 1992. (b) Same as Figure 11a but for ocean.

of our previous ISCCP-based estimates in the works by *Zhang et al.* [1995, 2004].

[35] In summary of the differences, we concentrate on the first two largest factors that affect downward LW at surface [see *Zhang et al.*, 1995]. (1) The uncertainty in mean (rms) difference of the surface air temperature is about 0.7 (3.7) and 0.8 (1.9) K for land and ocean, respectively, which may be interpreted as a uncertainty of ~3 K from bias-included RMS (= square root for the sum of square of mean and RMS difference, weighted by separated land and ocean, where RMS is bias-removed as we presented in the previous sections), 1 K larger than assumed in the previous sensitivity study [*Zhang et al.*, 1995]. (2) For the total column precipitable water (most contributions come from lowest three layers 560 hPa), the uncertainty is 20–25%, close to what was assumed in the sensitivity study [*Zhang et al.*, 1995]. These two leading factors for surface downward LW would induce the uncertainty of $\gtrsim$10 W/m^2 and $\lesssim$10 W/m^2, respectively, for downward LW flux values based on our past and current sensitivity studies. This study confirms that the uncertainty in the surface air temperatures is the leading factor causing uncertainty in the surface downward LW flux calculations. Even though the disagreements among the various data sets are generally only 2–4 K, such differences can easily cause 10–15 W/m^2 uncertainty in surface (downwelling) LW fluxes. In order to meet the accuracy requirements for climatological studies, say within 10 W/m^2, we still need substantial improvements of those input parameters that contribute major errors as summarized above and by Zhang et al. (submitted manuscript, 2005).

[36] The comparison for the difference between surface skin and air temperature shows its uncertainty is about 2–3 K, which translates into 10–15 W/m^2 for surface net LW flux uncertainty.

[37] The product that we used, FDTV, represents the diurnal variations better than the other available sources (by design) and the synoptic variations slightly better than these other sources, but still has notable clear-cloudy sky biases (although they are smaller than in the original ISCCP-TOVS data set) and interannual variations that are dominated by spurious changes introduced by methodology changes in the original TOVS (section 2.1).

[38] With all the important surface properties included, a more comprehensive summary of the uncertainties of the surface radiative flux calculation is presented in the work of Zhang et al. (submitted manuscript, 2005).

Appendix A: Procedure for Homogenization of Spatial Resolution and Derivation of Near-Surface Properties

[39] As mentioned in the main text, the common coordinate system we chose is the ISCCP-D1 system [*Rossow and Schiffer*, 1991] with one slight modification of its vertical projection: instead of a varying tropopause pressure level, we fix all the pressure levels to a maximum of ten levels (nine layers) defined by: Surface ($\leq$1000 hPa), 800, 680, 560, 440, 310, 180, 70, 30 and 0 hPa, where surface is defined by P_s (normalized as described below), from which the first (near-surface) atmospheric layer begins (since topography/P_s vary spatially, the total layer number and the thickness of the first layer also vary, e.g., if P_s = 780 hPa, the first layer becomes 780–680 hPa and the total layer number is reduced to eight from nine). The horizontal 2-D grid system is the standard ISCCP 280-km equal-area map.

[40] The profiles from the five products report temperature values either at fixed standard pressure levels (ERA15, GEOS-1 and NCEP) or at the centers of standard atmospheric pressure layers (3I and FDTV), beginning from mean sea level (regardless of topography) or a "true" surface pressure (P_s) to near the top of the atmosphere. Similarly, the six humidity profile data sets report either specific humidity at standard pressure levels or give the precipitable water (PW) amount for pressure layers. In our reprojection of all the levels/layers we pay special attention to the treatment of surface air temperature and near-surface-layer temperature and humidity as well as surface pressure.

[41] We treat the different temperature and humidity profiles as follows. (1) In the new pressure projection system with respect to the new grid, but not the final P_s, the projection conserves the lapse rate of the temperature in the two nearest-surface layers and changes surface specific humidity (PW of the near-surface layer) to be proportional to the ratio of the new P_s to old P_s (new to old near-surface-

layer thickness, conserving mass). (2) P_s is normalized by setting values >1000 hPa to 1000 hPa to produce the final P_s map. (3) Surface air temperature (T_a) and near-surface humidity are normalized to the final P_s such that T_a (linearly interpolated from the original two nearest-surface layer temperatures) and PW of the near-surface layer are both preserved unchanged (but the near-surface-layer temperature has to be changed for consistency with the T_a-interpolation). (4) If necessary, the PW profiles are rescaled to preserve the column total PW (after layer mean specific humidity is converted to PW amount) using the method described by *Rossow et al.* [1996], i.e., the final humidity profile is a layer-PW profile.

[42] The homogenized and normalized P_s maps from the five data sets (NVAP does not supply a P_s map so we just used the values from FDTV) are still a little different: their global mean (rms) differences (for regional 280-km equal-area cells) are as large as 3.7 hPa (17 hPa) but the maximum regional differences are as large as ~300 hPa in mountainous regions, suggesting that the underlying topography maps being used are not the same.

[43] All of the surface air temperatures, T_a, are then obtained by linear extrapolation from the homogenized temperature profiles (but see footnote b in Table 1 for FDTV).

[44] NCEP and ERA15 also report 2-m "surface air temperature" (T_{ao}) values (Table 1) that come from a first guess for NCEP (model forecast) and either analysis or forecast for ERA15. We find that the differences between T_a and T_{ao} are so large that they cannot be explained by the effects of the reprojection and the extrapolation: over all four monthly means (9201, 04, 07 and 10), the global mean differences (on the unified equal-area map) are up to 0.9 K and 0.12 K for ERA15 and NCEP, respectively, with a standard deviation of around 3 K. Although these differences are affected to a smaller degree by the linear extrapolation of T_a with a vertical resolution coarser than the original, the values of T_{ao} are apparently somewhat independent of the corresponding temperature profiles (perhaps a different lapse rate obtains in the atmospheric boundary layer). Even when we extrapolate surface air temperature (T_{aoi}) using the original higher-resolution temperature profiles, however, T_{aoi} still differs from T_a by about the same amount. To preserve consistency with the temperature profiles, we use only T_a in this work regardless T_{ao} or T_{aoi}.

[45] **Acknowledgments.** We thank Calleen Mikovitz and Steve Cox for supplying the ERA15 and GEOS-1 and SRB data sets. The work has been funded by the NASA Radiation Sciences Program (D. Anderson and H. Maring). We appreciate all the beneficial comments and suggestions from three anonymous reviewers that help improve this publication.

References

Darnell, W. L., and W. F. Staylor (1992), Seasonal variation of surface radiation budget derived from International Satellite Cloud Climatology Project C1 data, *J. Geophys. Res.*, *97*, 15,741–15,760.

Escoffier, C., J. Bates, A. Chedin, W. B. Rossow, and J. Schmetz (2001), Comparisons of upper tropospheric humidity retrievals from TOVS and METEOSAT, *J. Geophys. Res.*, *106*, 5227–5238.

Gibson, J. K., P. Kållberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano (1999), ERA-15 description (version 2 - January 1999), *Reanal. Proj. Rep. Ser.*, *1*, Eur. Cent. for Med.-Range Weather Forecasts, Reading, U. K.

Hahn, C. J., and S. G. Warren (1999), Extended edited synoptic cloud reports from ships and land stations over the globe, 1952–1996, *ORNL/CDIAC-123 NDP-026C*, 71 pp., Carbon Dioxide Inf. Anal. Cent., Oak Ridge Natl. Lab., Oak Ridge, Tenn.

Kistler, R., et al. (2001), The NCEP-NCAR 50 year reanalysis: Monthly means CD-ROM and documentation, *Bull. Am. Meteorol. Soc.*, *82*(2), 247–268.

Oort, A. H., and H. Liu (1993), Upper-air temperature trend over the globe, 1958–1989, *J. Clim.*, *6*, 292–307.

Randel, D. L., T. H. Vonder Haar, M. A. Ringgerud, G. L. Stephens, T. J. Greenwald, and C. L. Combs (1996), A new global water vapor dataset, *Bull. Am. Meteorol. Soc.*, *77*, 1233–1246.

Reale, A. L. (2001), NOAA operational sounding products from advanced-TOVS polar orbiting environmental satellites, *NOAA Tech. Rep. NESDIS 102*, 61 pp.

Reale, A. L., H. J. Bloom, and T. J. Gardner (1995), Revised TOVS (RTOVS) sounding system status, paper presented at 8th International TOVS Study Conference, Int. TOVS Working Group, Queenstown, New Zealand.

Reale, A. L., M. W. Chalfant, and L. Wilson (1999), Scientific status of NOAA advanced TOVS sounding products, paper presented at 10th International TOVS Study Conference, Int. TOVS Working Group, Boulder, Colo., 26 Jan. to 2 Feb.

Reynolds, R. W. (1988), A real-time global sea surface temperature analysis, *J. Clim.*, *1*, 75–86.

Rossow, W. B., and A. A. Lacis (1990), seasonal cloud variations from satellite radiance measurements. Part II: Cloud properties and radiative effects, *J. Clim.*, *3*, 1204–1253.

Rossow, W. B., and R. A. Schiffer (1991), ISCCP cloud data products, *Bull. Am. Meteorol. Soc.*, *72*, 1–20.

Rossow, W. B., and Y.-C. Zhang (1995), Calculation of surface and top-of-atmosphere radiative fluxes from physical quantities based on ISCCP datasets, part II: Validation and first results, *J. Geophys. Res.*, *100*, 1167–1197.

Rossow, W. B., A. W. Walker, D. E. Beuschel, and M. D. Roiter (1996), International Satellite Cloud Climatology Project (ISCCP) documentation of new cloud datasets, *WMO/TD 737*, 115 pp., World Clim. Res. Programme, Geneva, Switzerland.

Scott, N. A., A. Chédin, R. Armante, J. Francis, C. Stubenrauch, J.-P. Chaboureau, F. Chevallier, C. Claud, and F. Cheruy (1999), Characteristics of the TOVS pathfinder Path-B dataset, *Bull. Am. Meterorol. Soc.*, *80*, 2679–2701.

Simpson, G. C. (1929), The distribution of terrestrial radiation, *Mem. R. Meteorol. Soc.*, *3*, 53–78.

Smith, W. L., H. M. Woolf, C. M. Hayden, D. Q. Wark, and L. M. McMillin (1979), The TIROS-N operational vertical sounder, *Bull. Am. Meteorol. Soc.*, *60*, 1177–1187.

Smith, W. L., H. M. Woolf, C. M. Hayden, D. Q. Wark, A. J. Schreiner, and J. F. LeMarshall (1984), The physical retrieval TOVS export package, paper presented at 1st International TOVS Study Conference, Int. TOVS Working Group, Madison, Wis.

Stackhouse, P. W., Jr., S. K. Gupta, S. J. Cox, J. C. Mikovitz, T. Zhang, and M. Chiacchio (2004), 12-year surface radiation budget data set, *GEWEX News*, 10 –12 Nov.

Takacs, L. L., A. Molod, and T. Wang (1994), Documentation of the Goddard Earth Observing System (GEOS) general circulation model - version 1, *NASA Tech. Memo. 104606*, *1*.

Whitlock, C. H., et al. (1995), First global WCRP shortwave surface radiation budget dataset, *Bull. Am. Meterorol. Soc.*, *76*, 905–922.

Zhang, Y.-C., W. B. Rossow, and A. A. Lacis (1995), Calculation of surface and top-of-atmosphere radiative fluxes from physical quantities based on ISCCP datasets, part I: Method and sensitivity to input data uncertainties, *J. Geophys. Res.*, *100*, 1149–1165.

Zhang, Y.-C., W. B. Rossow, A. A. Lacis, V. Oinas, and M. I. Mishchenko (2004), Calculation of radiative fluxes from the surface to top-of-atmosphere based on ISCCP and other global datasets: Refinements of the radiative transfer model and the input data, *J. Geophys. Res.*, *109*, D19105, doi:10.1029/2003JD004457.

Y. Zhang, Applied Physics and Applied Mathematics, Columbia University, at NASA Goddard Institute for Space Studies, New York, NY 10025, USA. (yzhang@giss.nasa.gov)

W. B. Rossow, NASA Goddard Institute for Space Studies, New York, NY 10025, USA.

P. W. Stackhouse Jr., NASA Langley Research Center, Hampton, VA 23681, USA.