9-16-2001

Validation of Land Surface Models Using Satellite-Derived Surface Temperature

Joshua Rhoads
University of Maryland - College Park

Ralph Dubayah
University of Maryland - College Park

Dennis Lettenmaier
University of Washington - Seattle Campus

Greg O'Donnell
University of Washington - Seattle Campus

Venkataraman Lakshmi
University of South Carolina - Columbia, vlakshmi@geol.sc.edu

Publication Info

Published in *Journal of Geophysical Research*, Volume 106, Issue D17, 2001, pages 20085-20099.
Rhoads, J., Dubayah, R., Lettenmaier, D., O'Donnell, G., & Lakshmi, V. (2001). Validation of land surface models using satellite-derived surface temperature. *Journal of Geophysical Research, 106* (D17), 20,085-20,099.

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 106, NO. D17, PAGES 20,085–20,099, SEPTEMBER 16, 2001

Validation of land surface models using satellite-derived surface temperature

Joshua Rhoads and Ralph Dubayah

Department of Geography, University of Maryland, College Park, Maryland

Dennis Lettenmaier and Greg O'Donnell

Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington

Venkat Lakshmi

Department of Geological Sciences, University of South Carolina, Columbia, South Carolina

Abstract. This research examines the feasibility of using remotely sensed surface temperature for validation and updating of land surface hydrologic models. Surface temperature simulated by the Variable Infiltration Capacity (VIC) hydrologic model is compared over the Arkansas-Red River basin with surface temperature retrievals from TOVS and GOES. The results show that modeled and satellite-derived surface temperatures agree well when aggregated in space or time. In particular, monthly mean temperatures agree on the pixel scale, and basin mean temperatures agree instantaneously. At the pixel scale, however, surface temperatures from both satellites were found to have higher spatial and temporal variabilities than the modeled temperatures, although the model and satellites display similar patterns of variability through space and time. The largest differences between modeled and remotely sensed surface temperature variability occur at times of maximum net radiation both diurnally and seasonally, i.e., afternoon and summer. Comparison of temporal and spatial patterns of VIC-predicted surface temperature variability with similar predictions by nine other models involved in the PILPS-2c experiment show that the VIC patterns are similar to those of the other models. Observed surface temperature and air temperature from FIFE are used to identify possible errors in satellite-retrieved surface temperatures. The FIFE comparisons show that satellite retrieved surface temperatures likely contain errors that increase variability.

1. Introduction

Evaluation of the performance of land-atmosphere models is a complex problem, yet it is necessary to determine the efficacy of hydroclimatological predictions. Traditional methods of validation, such as comparison of predicted and observed streamflow and atmospheric water vapor budgets are conducted using fluxes integrated over an entire watershed. Although useful, these approaches cannot determine whether a model correctly represents the spatial distribution of energy and moisture fluxes, e.g., overestimation in one area could compensate for underestimation in another. Point observations of surface fluxes are made at a growing number of surface flux towers, but stations are sparse, and in many parts of the world nonexistent. Furthermore, surface flux observations measure conditions at a point and may not capture the spatial variability of fluxes that depend on the heterogeneous nature of the land surface. Remote sensing offers an alternative data source for model evaluation that can mitigate some of these difficulties.

Land surface temperature (T_s) is an important climate variable that is retrievable from space and is closely linked both to the surface energy balance and to soil moisture. Land surface models predict latent, sensible, and ground heat fluxes using algorithms that either predict surface temperature (in which case observations can, in principle, be used for validation or updating) or use it as a forcing variable (in which case remote sensing could provide an alternative source of forcing data). Outgoing longwave radiation is proportional to T_s^4, sensible heat depends on the difference between surface temperature and air temperature near the surface, and surface temperature also influences the vapor pressure deficit, which in turn affects the latent heat flux. Because of these dependencies, soil moisture is effectively a buffer that under some conditions (especially in summer, when solar radiation is in ample supply) controls evapotranspiration. Therefore accurate remote sensing of surface temperature could have important implications for improving the predictability of land surface fluxes of moisture and energy.

Previous comparisons of satellite-derived surface temperatures with model predictions have shown good agreement of monthly mean temperatures [*Jin et al.*, 1997]. However, the need and potential for satellite data, e.g., for model validation and data assimilation, is at much shorter time intervals [*Dubayah et al.*, 2000]. This paper examines the variability in space and time of remotely sensed surface temperature from two sources, TOVS (TIROS Operational Vertical Sounder) and GOES (Geostationary Operational Environmental Satellite), as compared to modeled predictions, with emphasis on short timescales. The intent is to gain insight into the validity of the modeled energy balance on a spatially continuous and temporally consistent basis as well as to

Paper number 2001JD900196.
0148-0227/01/2001JD900196$09.00

evaluate the suitability of satellite observations for data assimilation.

2. Data

The study area is the combined Arkansas and Red River basin in the U.S. Southern Great Plains. These two basins have similar climatologies and are combined into a single water resource region by the U.S. Geological Survey. The combined basin covers an area of approximately 566,000 km^2 and is represented by 61 (1° × 1°) grid cells for TOVS comparisons and 272 (0.5° × 0.5°) grid cells for GOES comparisons. There is a large precipitation gradient in the basin, with conditions ranging from arid and semiarid in the west to humid in the east. The basin's hydrological processes are dominated by the humid eastern part, although large convective storms and snow occur in the west. Location, elevation, and land cover maps of the combined Red/Arkansas basin are shown in Plate 1.

2.1. TOVS

TOVS has flown on the polar orbiting satellites TIROS-N, NOAA 6-NOAA 12 and NOAA 14 from late 1978 to present. Depending upon the satellite, TOVS has equatorial, at nadir, daily overpass times of either 0230 and 1430 local time (LT) or 0730 and 1930 LT. The exact time of observation for each pixel varies with distance from the equator and ± 1 hour for off-nadir observations. TOVS is composed of HIRS2 (High Resolution Infrared Sounder), MSU (Microwave Sounding Unit) and SSU (Stratospheric Sounding Unit). These instruments were designed primarily for producing three-dimensional profiles of temperature and moisture in the atmosphere. The TOVS Pathfinder Path A data set [*Susskind et al.*, 1997], was produced as part of a NASA-NOAA joint program for the development of satellite-derived climate data sets. *Susskind et al.* [1984] developed a physically based model to estimate global land and sea surface temperatures from TOVS retrievals. Their algorithm calculates surface temperature using a 6-hour forecast of atmospheric temperature and moisture profiles produced by a 4° × 5° version of the Goddard Earth Observing System - data assimilation system (GEOS-DAS) general circulation model [*Pfaendtner et al.*, 1995]. Brightness temperatures predicted from the forecast profiles of temperature and water vapor are compared with those observed by the satellite. An iterative relaxation of atmospheric conditions is carried out on the basis of the difference between the modeled and the observed clear-sky radiances until convergence is reached or the retrieval is rejected. Retrievals can be made with up to approximately 80% cloud coverage, and for this reason, the estimated surface temperature should be considered an average of the cloud-free areas of the scene. The Pathfinder Path A data set used in this study is a gridded 1° × 1° product that includes the corresponding time of each observation.

Initial TOVS validation studies have shown encouraging agreement between TOVS T_s and ground-observed T_s. *Lakshmi and Susskind* [1999] conducted direct comparisons of TOVS surface temperature to field observations collected during the First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment (FIFE), the Boreal Ecosystem-Atmosphere Study (BOREAS), and the Hydrologic-Atmospheric Pilot Experiment (HAPEX). The standard deviations of the instantaneous T_s differences between TOVS and ground observations were found to be between 4° and 5°C and a bias of nearly zero for the length of the experiments, which range between 1 year (HAPEX), 2.5 years (FIFE), and 3 years (BOREAS).

2.2. GOES

Observations of surface temperature were derived from GOES using a split-window equation [*Czajkowski et al.*, 1998] such that

$$T_S = 0.0885 + R_4 + 0.849(R_4 - R_5) + 0.3796(R_4 - R_5)^2, \quad (1)$$

where R_4 and R_5 are channels 4 and 5 clear-sky radiances. Emissivity was assumed to be constant at 1.0 for both channels. Surface temperature was calculated hourly during daylight only for May, June, July, and August 1997 at 0.5° spatial resolution.

2.3. Model Predictions

The land surface model used in this study is the Variable Infiltration Capacity (VIC) hydrologic model [*Liang et al.*, 1994, 1996, 1999]. VIC uses a simplified soil-vegetation-atmosphere transfer (SVAT) scheme to represent the interaction between soil hydrology and evapotranspiration. VIC distinguishes itself from other SVAT models by representing the sub-grid-scale heterogeneity of soil properties, and their effect on the surface water balance, via a spatially varying probability distribution of infiltration capacity. A full surface energy balance is performed.

The VIC calculation of surface temperature is embedded within a two-layer ground heat flux formulation. Heat storage in each soil layer and the effects of vegetation on net radiation are incorporated. For each time step, and for each vegetation type, an initial surface temperature is determined from the ground heat flux. The initial surface temperature is modified iteratively to solve the energy balance, producing a surface temperature that minimizes energy balance errors. Changes in forcing variables that affect the energy balance, such as precipitation and air temperature, will therefore change surface temperature derivation.

Because the model solves the coupled energy and water balances, surface temperature estimates are dependent on various vegetation and soil parameters. Vegetation types were obtained from the 1 km EROS (Earth Resources Observation System) AVHRR vegetation data set [*Dickinson et al.*, 1986; *Eidenshink and Faundeen*, 1994]. The vegetation parameters required by VIC are based on those suggested by *Dickinson et al.* [1993] for use with BATS (Biosphere-Atmosphere Transfer Scheme). Each grid cell contains two to six vegetation types with a corresponding fractional cover for each type. For each vegetation type, parameters include leaf area index (LAI), fractional cover of vegetation, roughness length, displacement height, albedo, stomatal resistance, and architectural resistance. All parameters vary monthly but not yearly.

Model parameters related to soil properties were derived using soil information from the U.S. Soil Conservation Service State Soil Geographic (STATSGO) [*STATSGO*, 1994] database. Total soil depth data are directly available while saturated hydraulic conductivity, porosity, soil moisture at field capacity, and wilting point were derived from STATSGO soil texture (i.e., percent sand and clay) using methods in the Handbook of Hydrology [*Rawls et al.*, 1993]. The STATSGO 1 km data were averaged up to the 1° × 1° resolution.

The VIC forcing data include precipitation, air temperature, wind speed, humidity, and downwelling shortwave and longwave radiation. Precipitation and air temperature observations were obtained from Summary of the Day (NOAA, NCDC, Earthinfo Inc., CD-ROM, 1997) and the wind observation data from Surface Airways (NOAA, NCDC, Earthinfo Inc., CD-ROM, 1997). Shortwave radiation, longwave radiation, and vapor

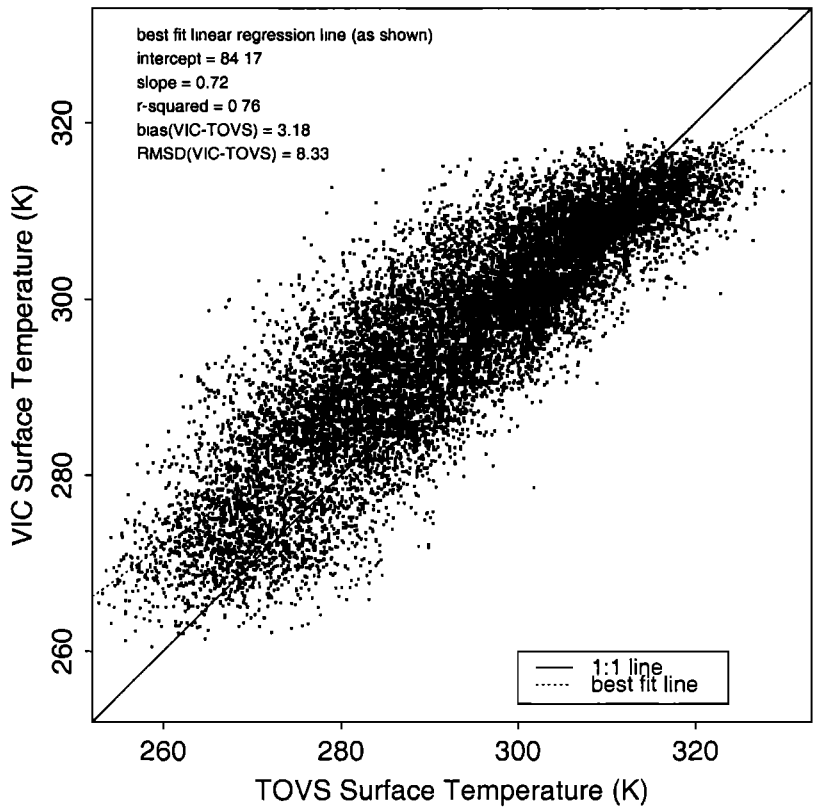

Figure 1. Scatterplot of VIC versus TOVS surface temperature. Comparison is for 1985 at the 1430 LT overpass.

pressure were parameterized from air temperature using methods described by *Thornton and Running* [1999], *Bras et al.* [1990], and *Kimball et al.* [1997], respectively. Daily precipitation and temperature data from the observing stations were interpolated to the grid cells using the Symap algorithm [*Shepard*, 1984], which incorporates an inverse distance and directional weighting scheme. The precipitation data were rescaled using monthly PRISM (precipitation-elevation regressions on independent slopes mode) fields, in which slope orientation and orography are incorporated [*Daley et al.*, 1984]. Temperatures were lapsed to a reference height, and the final interpolated grid values were lapsed to the mean grid cell elevations at a rate of 6.5 K/1000 m. Gridded wind speed values were derived using an inverse distance squared weighting scheme. The gridded daily air temperature observations were disaggregated to the model time step (i.e. 1 hour), by first assuming that the minimum occurred at dawn and the maximum at two-thirds the interval between dawn and dusk.

3. TOVS Comparisons

Surface temperature was compared at the time of satellite observation for each grid cell. The TOVS sensor is onboard the NOAA polar orbiting series of satellites so the time of observation is unique to each pixel (up to ±1 hour around the overpass time depending on latitude and distance from nadir) and varies slightly from day to day. The exact time of each observation is included in the Path A data set. Therefore for each grid cell, the VIC output was matched to the exact time of TOVS observation by isolating the two VIC temperatures at the hourly time steps that surround the time of satellite observation, and linearly interpolating. If TOVS data were not available due to clouds or gaps in the scan swaths, no comparison was made with the VIC predictions. For ease of reference, times of comparison will be referred to as the local equatorial, at nadir overpass time (0230, 1430, 0730, or 1930 LT, depending on the year).

The comparisons of TOVS versus VIC are made using 1985 1430 LT data unless otherwise stated (i.e., one observation per day, per grid cell at approximately 1430 LT). The 1430 LT overpass was chosen because, of the times available, it represents the most active time of energy exchange between the surface and the atmosphere. A scatterplot of all 1985 1430 LT data is presented in Figure 1. The bias (VIC − TOVS) is 3.18 K, root-mean-square (RMS) difference is 8.33 K and the R^2 value is 0.76.

The 1430 LT monthly mean was calculated for each pixel over the days within each month. Monthly means display good agreement, both geographically and seasonally (Plate 2a). Differences between the VIC and the TOVS means fall within the range of −5.8 to 11.3 K over the 61 grid cells, with 80% of the differences falling between −2 and 6 K. The model and the satellite exhibit similar spatial patterns of warming and cooling. For example, both the VIC and the TOVS mean temperatures are generally homogeneous throughout the basin in March and October. In the winter months (November through February) the VIC and the TOVS north-south temperature gradients are very similar. Elevation effects are evident in both data sets with temperatures decreasing in the northwest portion of the basin that contains the foothills of the Rocky Mountains. The arid south central sector of the basin warms up in May and stays warm through September, while eastern areas, typically wetter and more densely vegetated, remain relatively cooler. Thus both the model and the satellite capture the seasonal variations of climate within the region.

The monthly temporal standard deviation (σ_t) was calculated for each pixel over the days in the month, i.e., up to a possible 31 observations for each pixel. A threshold was applied to insure

that at least 10 days of the month must have had data before σ_t was computed. The images in Plate 2b show that TOVS has a higher variability through time than VIC. TOVS monthly σ_t ranges from 3.4 to 16.3 K over the region, while VIC σ_t ranges from 0.5 to 11.8 K. VIC σ_t decreases in the summer months with the lowest values occurring in the southern and eastern portions of the basin. There is a marked increase of variability for both the model and the satellite in the region within 3° longitude of the 100th meridian during the month of September and a return to lower values in October. Beginning in November, VIC σ_t in the far northwest portion of the basin becomes depressed relative to the rest of the basin and remains low through March. Examination of these northwestern grid cells revealed that snow was on the ground, and this caused a dampening of surface temperature amplitude in the model. Excluding the northwest, TOVS and VIC show good agreement of σ_t throughout the basin for the months of November, December, and February. Both the model and the satellite have similar spatial trends of σ_t with lower variability in the east, increasing for longitudes between 98°W and 103°W, then decreasing towards the west end of the basin.

Spatial (basin) mean temperatures were compared as anomalies of seasonality. Anomalies were derived as

$$T_{s(i,j)\text{anomaly}} = T_{s(i,j)} - T'_{s(i,j)} , \quad (2)$$

where $T_{s(i,j)}$ is an instantaneous observation and $T'_{s(i,j)}$ is the 2-year monthly mean at the time of overpass for each pixel of latitude i and longitude j. Anomalies were derived independently for VIC and TOVS. Data from 1985 and 1986 were used to compute the monthly mean since TOVS retrievals for these years (NOAA 9) are at 1430 LT. By averaging these values through space then plotting them in time (Figure 2), we can examine the spatial average of daily departures from the 2-year monthly climatology. The time series plot shows good agreement between VIC and TOVS, and even when values differ, daily changes are similar to each other in direction (positive or negative) and magnitude. Both the satellite and the model capture the daily variation of basin-wide surface temperature independent of the effects of seasonality.

Spatial standard deviation (σ_s) was calculated as the standard deviation of the surface temperatures of the 61 grid cells at the 1430 LT overpass each day. Spatial standard deviation was only calculated on days for which at least 21 of the 61 possible observations (grid cells) had TOVS estimates of T_s. Exclusion of days with a small number of TOVS-observed temperatures creates the gaps seen in the time series. The time series plot (Figure 3) shows the greater spatial variability of TOVS, which ranges between 4 and 12 K (average of 7.2 K) for the year, as compared with a range for VIC from 2 to 8 K with a yearly average of 4.0 K. During June and July, VIC has depressed spatial variability, averaging 4.3 K less than TOVS. November and December comparisons show better agreement of σ_s, with both model and satellite in the same range, and having similar trends through time. This improved aggreement might possibly be attributed to the onset of senescence and therefore the effect of vegetation properties on surface temperature could be reduced. Also, cooler temperatures and reduced net radiation may reduce errors in satellite retrievals and model predictions of surface temperature. In all months, both VIC and TOVS generally rise and fall together, though occasionally out of phase, representing similar daily changes of σ_s.

Given their close temporal proximity to daily T_s maxima and minima, the NOAA 9 overpass times of 0230 LT (night) and 1430 LT (day) are well suited for exploring diurnal ranges of temperatures. Figure 4 shows histograms of day minus night T_s differences. Because the minimum observed air temperature (T_a)

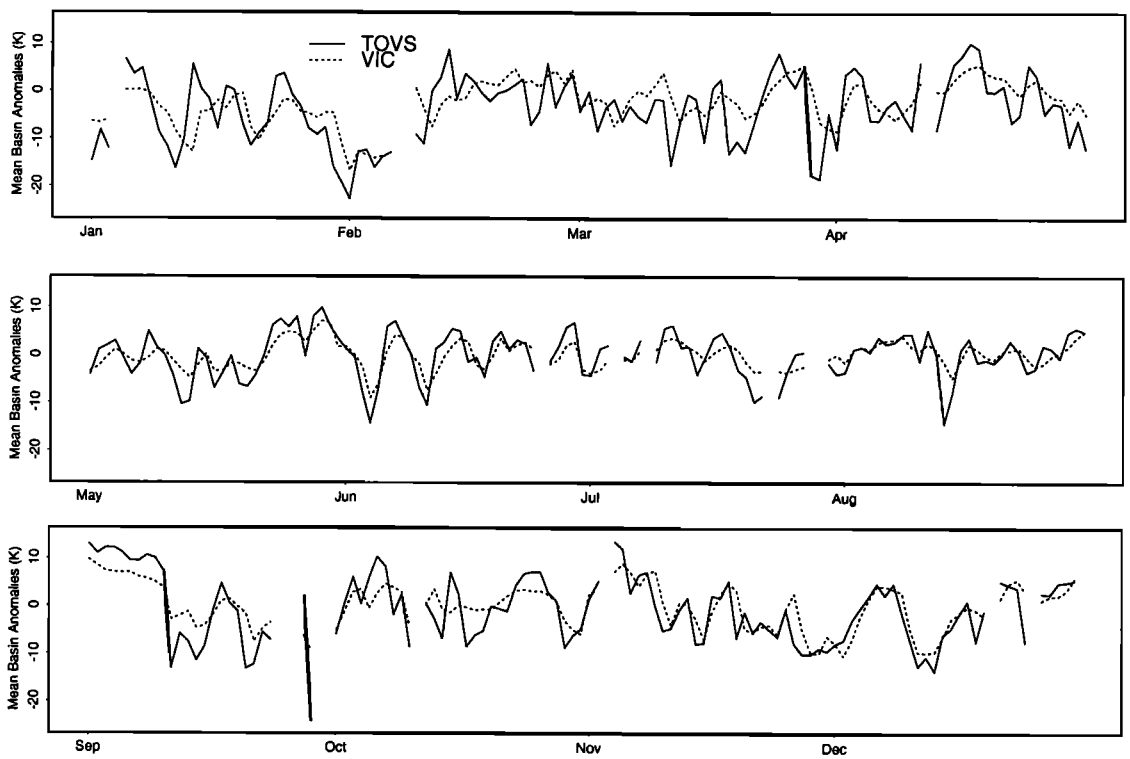

Figure 2. Time series of VIC and TOVS mean basin surface temperature anomalies. The 2-year monthly average temperature for each pixel is calculated for and subtracted from each observation prior to finding the basin (spatial) mean. Calculated daily for 1985 at the 1430 LT overpass.

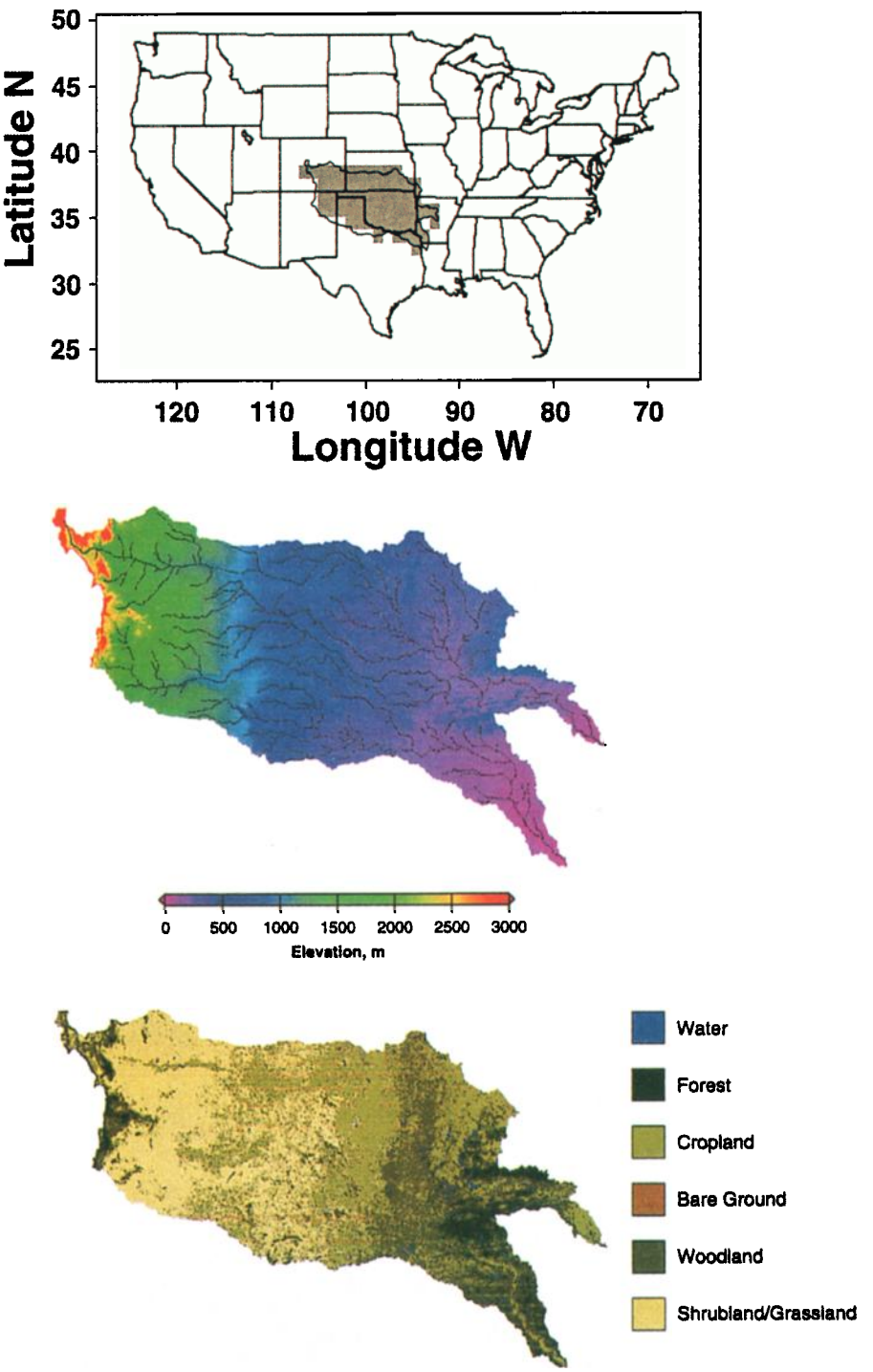

Plate 1. Maps of the Red/Arkansas River basin. (a) Location of the Red/Arkansas within the continental United States (shaded in gray are 1° grid cells). (b) Elevation and streams. (c) Land cover classifications.

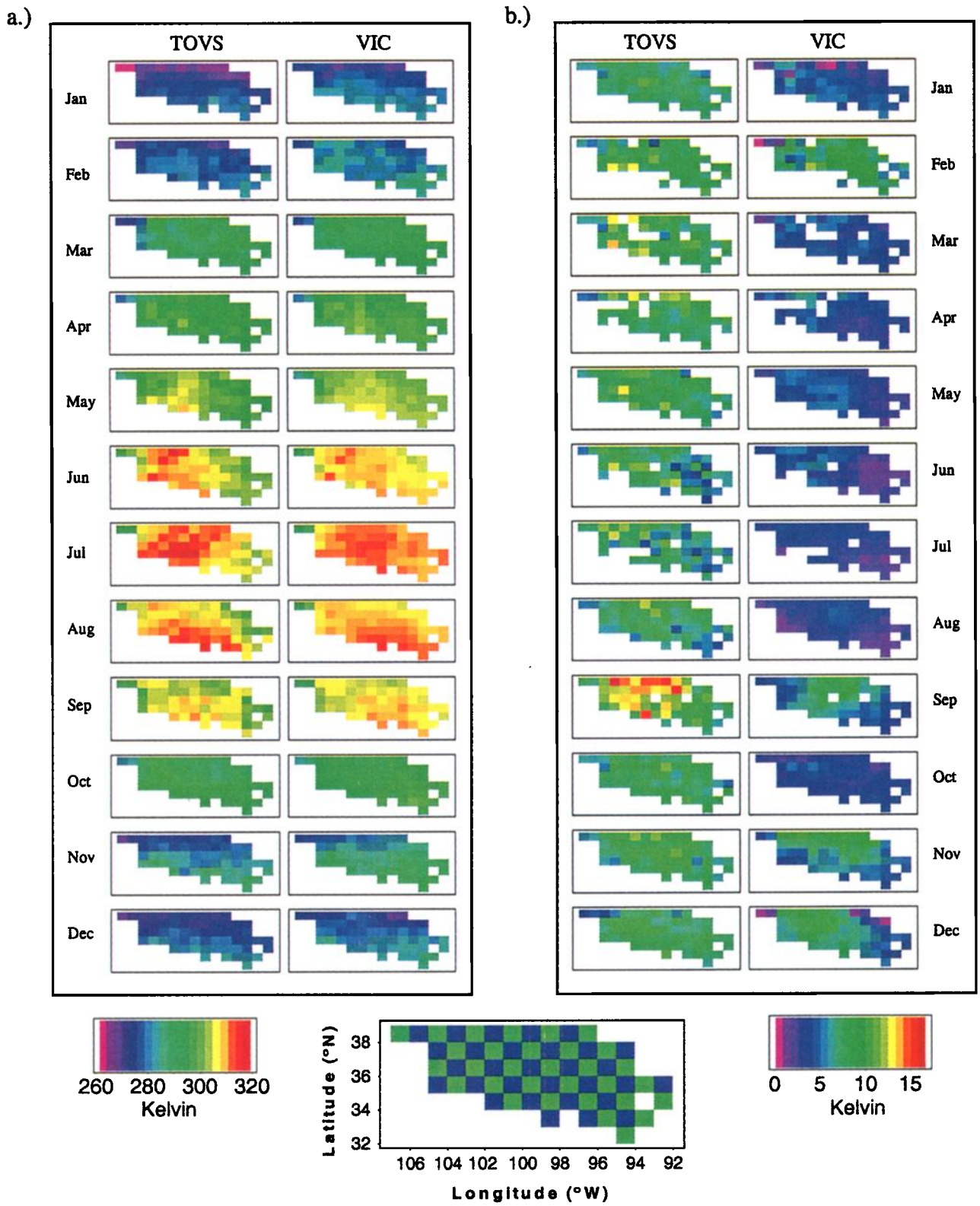

Plate 2. (a, b) Left column is TOVS, right column is VIC; x and y axes are longitude and latitude; comparisons are for 1985 at the 1430 LT overpass time. (a) Images of monthly mean surface temperatures. (b) Images of monthly temporal standard deviation.

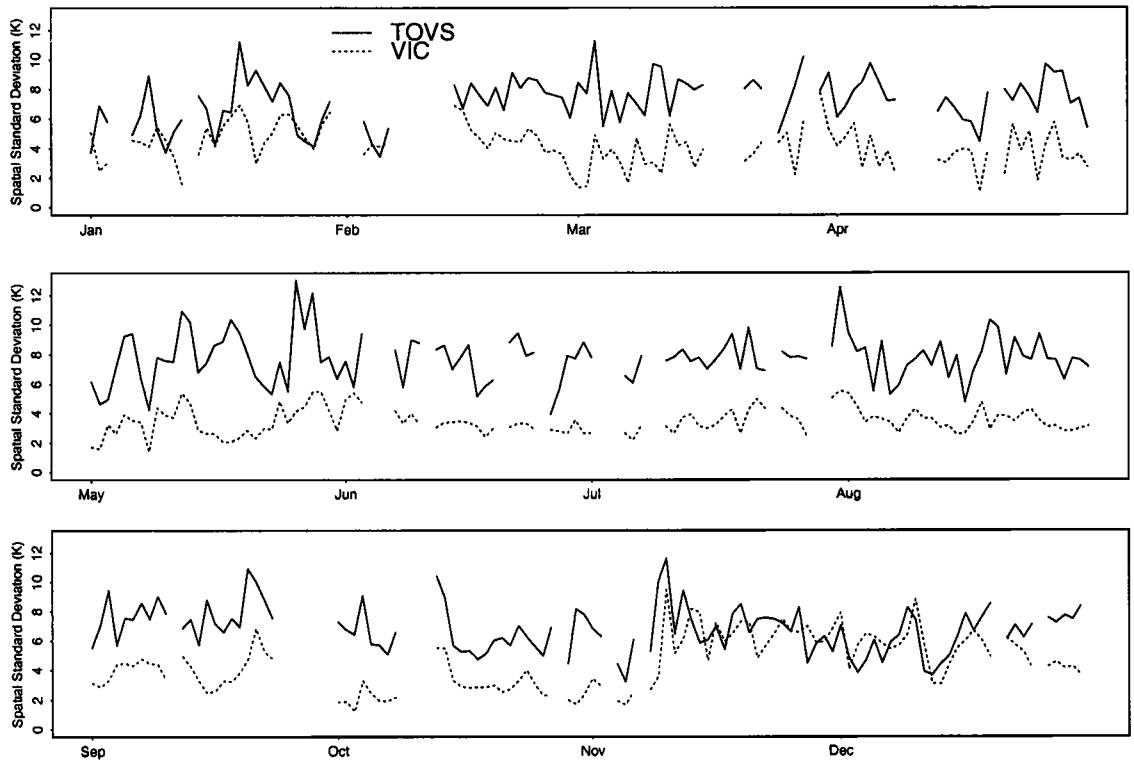

Figure 3. Time series of TOVS and VIC spatial standard deviation of surface temperature. Calculated daily for 1985 at the 1430 LT overpass. Gaps in the time series are caused by an insufficient number of observations.

is restricted from exceeding the maximum, and because the VIC modeled T_s is dependent on T_a, all values for VIC differences are positive. The frequency of VIC differences has a mode of 17 K and has a range of 0 to 33 K. There is a noticeable increase in frequency of VIC differences between 0 and 2 K which corresponds to diurnal range damping caused by snow. TOVS differences have a much wider range (−26 to 47 K), and there are instances where night T_s is higher than day T_s. The high frequency of negative differences is suspect based on the conservative assumption that it is rare that surface temperature

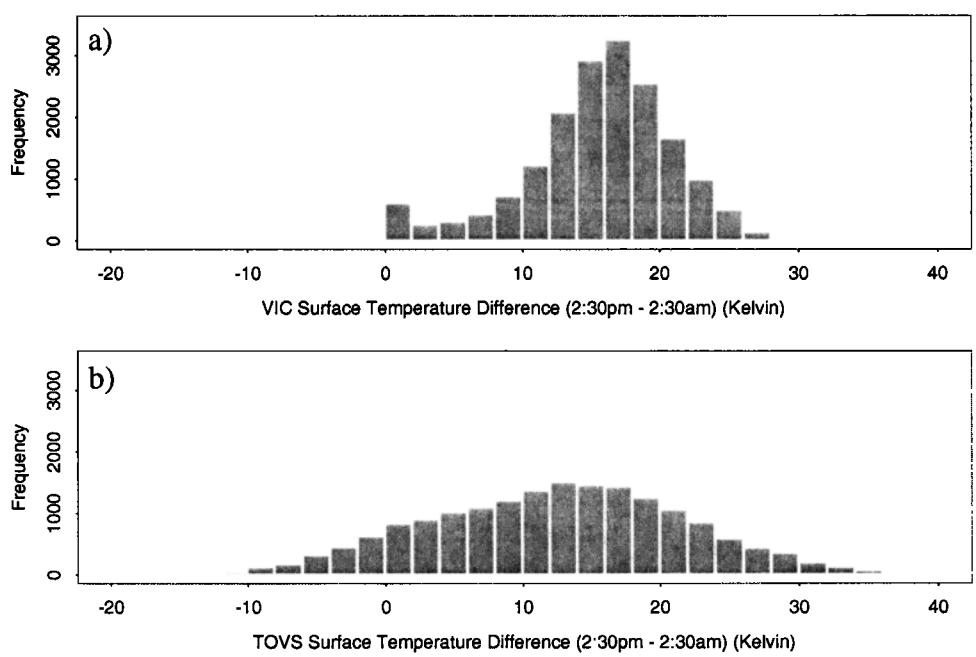

Figure 4. Histograms of temperature difference between 0230 LT and 1430 LT overpasses. Computed for each pixel as $T_s(1430) - T_s(0230)$. (a) VIC, (b) TOVS.

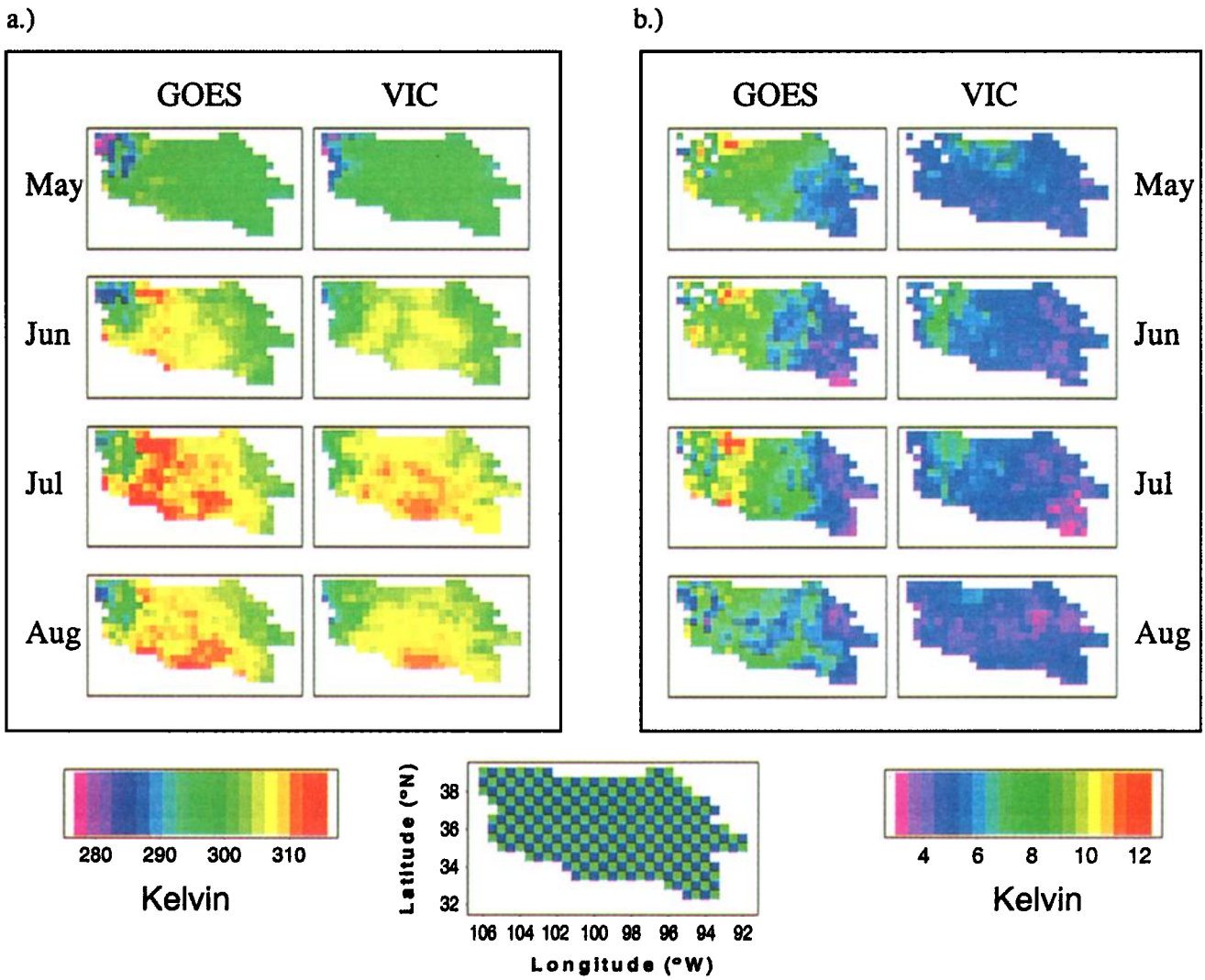

Plate 3. (a, b) Left column is GOES, right column is VIC; x and y axes are longitude and latitude; comparisons are for 1997 daylight hours only. (a) Images of monthly mean surface temperatures, (b) images of monthly temporal standard deviation.

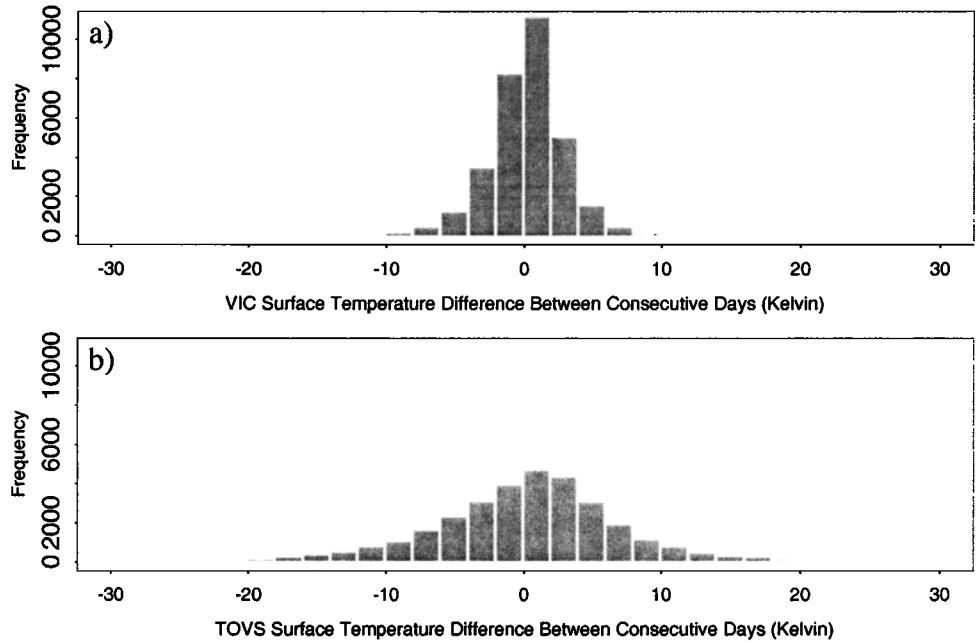

Figure 5. Histogram of temperature differences between consecutive days (1985, 1430 LT overpass). Computed for each pixel as $T_{s(i+1)} - T_{s(i)}$, where i is Julian day. (a) VIC, (b) TOVS.

will be higher at 0230 LT than at 1430 LT. We attribute this to cloud contamination.

Daily surface temperature variability was compared by calculating T_s changes from day to day (i.e., successive days) at 1430 LT for each pixel. TOVS exhibits a higher daily variability with temperature changes ranging from −25 to 25 K (Figure 5). VIC temperatures fall within a much smaller range of −10 to 10 K.

4. GOES Comparisons

GOES-derived surface temperatures were compared to VIC hourly values during daylight only for May, June, July, and August 1997 at 0.5° × 0.5° spatial resolution. Using all data, the bias (VIC − GOES) is −1.12 K, with a RMS difference of 5.21 K and the R^2 value is 0.68 (Figure 6a). The eastern portion of the basin appeared to have better agreement than the west. To test

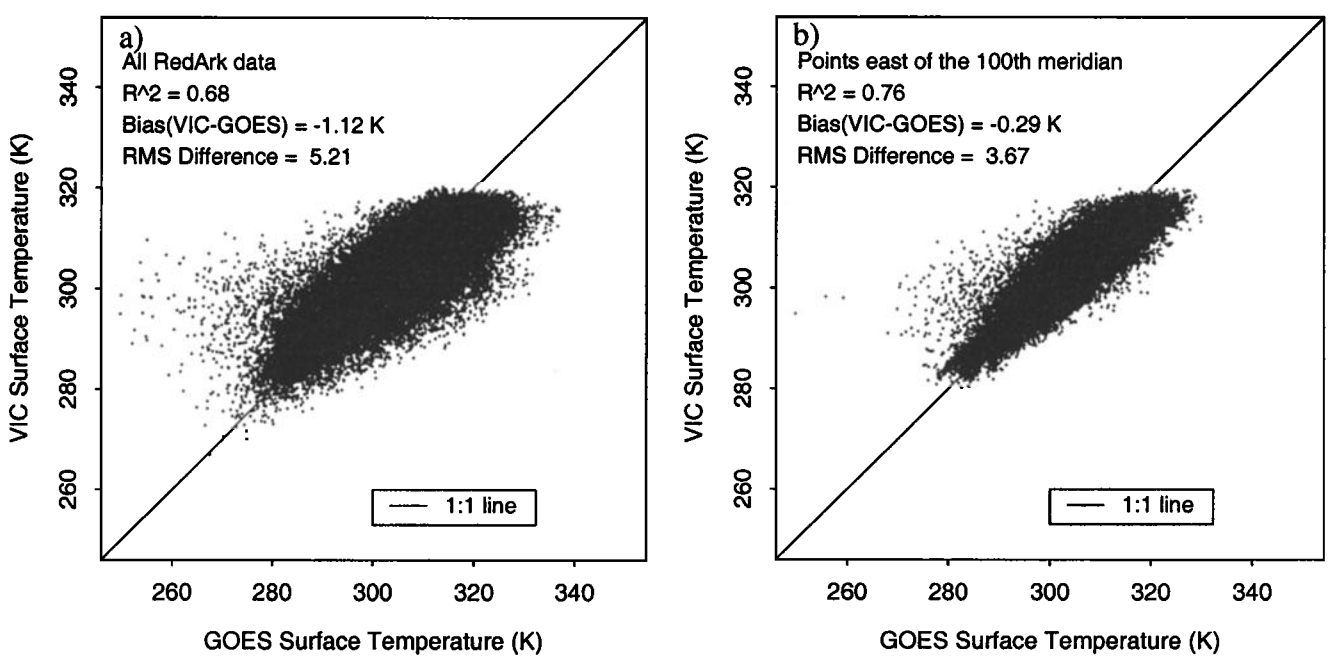

Figure 6. Scatterplots of GOES versus VIC T_s for daylight hours in May, June, July, and August 1997. (a) all data, (b) only pixels east of the 100th meridian.

whether this was true, a comparison was made using only data from the eastern part of the basin. By eliminating all data west of the 100th meridian, the bias and RMS difference were decreased to −0.29 K and 3.67 K respectively, and the R^2 value increased to 0.76 (Figure 6b). This east to west pattern is visible in the images of average monthly T_s (Plate 3a). The eastern portion of the basin shows good agreement through space and time, i.e., from north to south and from month to month. Additionally, the mountainous region in the upper northwest displays similar patterns for both the model and the satellite. The greatest differences occur in the longitudes between 100°W and 104°W during the months of June, July, and August, where it is evident that GOES temperatures have a warm bias relative to VIC.

Images of temporal standard deviation are shown in Plate 3b. GOES σ_t ranges from 3.4 to 12.2 K, whereas VIC σ_t ranges from 3.2 to 7.6 K. Temporal standard deviation varies through space, primarily on an increasing east to west gradient, for both GOES and VIC. While the gradient is steeper for GOES, both data sets display similar patterns of changing σ_t through time (month to month) and space. For example, the southeast corner of the basin encompasses the least variable grid cells that reach their minimum in July, maximum variability occurs in the region surrounding the mountainous northeast corner, and basin-wide temporal variability is at its lowest in August.

The hourly resolution of the GOES satellite platform allows for examination of the diurnal variability of surface temperature. Although we restrict our analysis to daylight hours, we can nonetheless examine the daytime rise and fall of both the spatial mean and the spatial standard deviation of modeled versus remotely sensed surface temperature. The monthly mean diurnal curve was calculated for each month (Figure 7) by first finding the average basin temperature at each time step and then averaging through the month for each time step. These curves display good agreement, reinforcing earlier findings (i.e., TOVS comparisons) that averaging in space improves the agreement between the satellite observations and the model predictions. The bias of GOES over VIC is evident, while further revealing that this bias is at its maximum at or near the time of maximum net radiation (~3:00 p.m.).

Figure 8 shows the daytime variability of σ_s. As with the mean basin temperatures, the standard deviation for the basin is found for each daylight hour, and then the monthly average of σ_s is computed for each time step. This figure shows a pronounced daytime variation of GOES σ_s, which reaches a maximum that coincides approximately with maximum net radiation. In contrast, VIC displays a brief increase in σ_s during the morning and a subsequent gradual decline throughout the rest of the day. A possible explanation is that spatial variability of GOES T_s is more closely related to net radiation than is the spatial variability of model predictions.

5. PILPS-2c comparisons

Although there are substantial differences between both the temporal and the spatial variability of surface temperature derived from the model and satellite sensors, there is no independent means of determining whether the model or remotely sensed data are correct. We were, however, able to determine whether VIC is an outlier among other SVAT models, through examination of the results from PILPS-2c (Project for the Intercomparison of Land Surface Parameterization Schemes) [*Wood et al.*, 1998]. The PILPS-2c experiment was conducted over the Red River/Arkansas River basin for the years 1979-1988 and therefore allowed comparison with 1985 TOVS data. Surface temperature from VIC and nine other models involved in PILPS-2c were compared to TOVS 1985 1430 LT overpass. The 10 models used are a subset of the total of 16 PILPS-2c participants for which results were output on an hourly time step. The PILPS-2c VIC results are slightly different than those used in section 3 because of model upgrades and modifications to forcing data subsequent to the PILPS-2c experiment, but they are generally comparable. Scatterplots of all data for all models for the 1985 1430 LT overpasses of TOVS are presented in Figure 9. The only model that has a cool bias (relative to TOVS) is

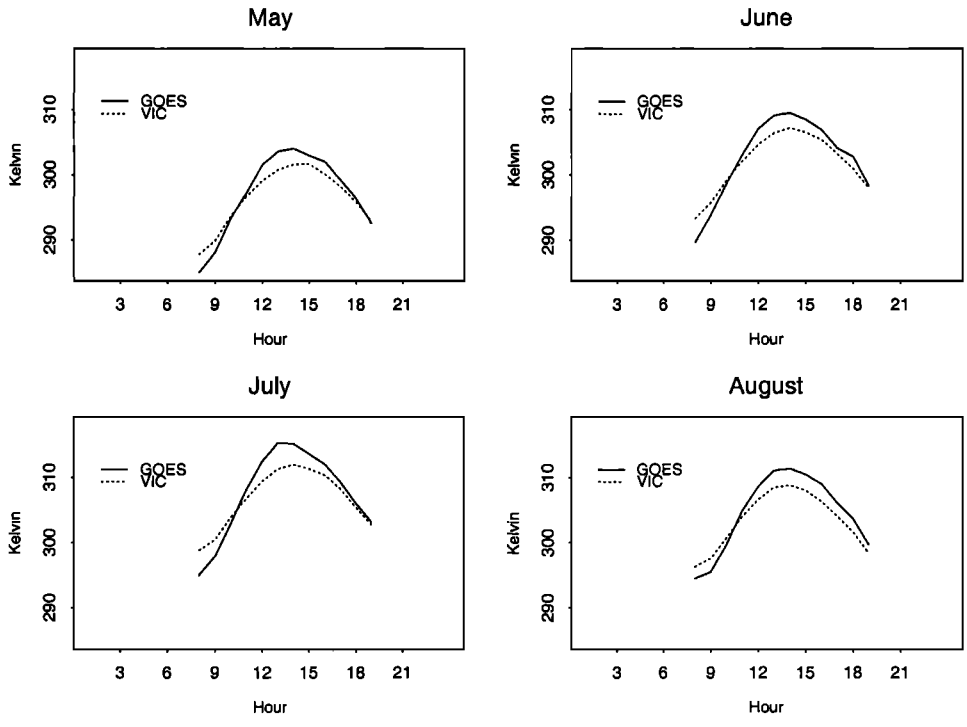

Figure 7. 1997 GOES and VIC average monthly diurnal curves of surface temperature.

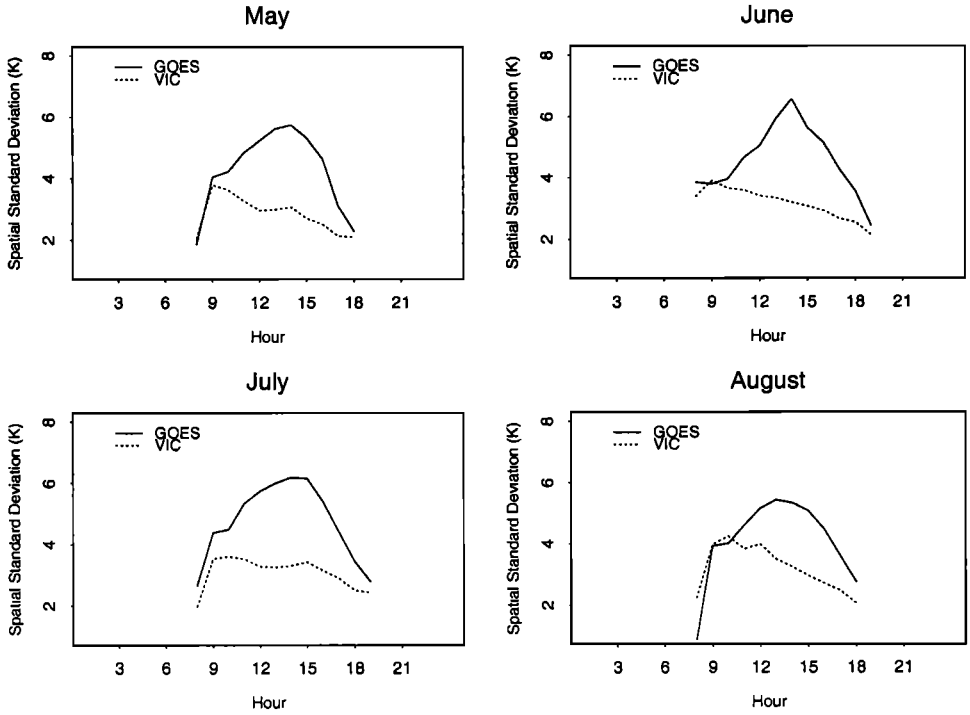

Figure 8. 1997 GOES and VIC average monthly diurnal curves of spatial standard deviation.

MOSAIC at −1.24 K. All other models are positively biased over TOVS, ranging from 0.34 K for SSiB to 2.74 K for NCEP. RMS differences are all similar ranging from 7.18 K to 8.02 K, and R^2 values all fall between 0.77 and 0.80. Table 1 summarizes the bias, RMS, and R^2 comparisons of PILPS-2c models to TOVS.

Spatial variability trends are similar for all PILPS-2c models (Figure 10) As seen in the VIC analysis (section 3), all models show depressed variability in summer, better agreement with TOVS during November, and a pronounced departure in mid-December. The spatial variabilities of all models display similar behavior with no particular model standing apart from the rest. Monthly temporal standard deviation was also calculated for all models and was found to show strong intermodel agreement, while all models are less variable through time than TOVS.

6. Data Selection Based on FIFE Observations

The concerns raised in section 3, regarding possible TOVS errors, are addressed in this section through use of a screening system based on surface observations. When models are compared to satellite retrievals, it is difficult to interpret the results unless the satellite data are first validated. Ground-based validation of TOVS and GOES is difficult because of the lack of in situ surface temperature observations. Even when validated, surface temperature observations for any specific application may be in error because of improper choice of model parameters, emissivity, and especially by subpixel cloud contamination or poor cloud screening. The last of these can lead to unrealistically cold temperatures.

The relationship between surface temperature and near-surface air temperature was used to screen TOVS data for possible errors. The data used to create boundaries for the screening process are from the First ISLSCP Field Experiment (FIFE), during which observations were taken at 20 sites within a 15 km^2 site near Manhattan, Kansas. Although the FIFE area is not representative of the entire basin, it is typical of grassland regions in the central part of the Arkansas-Red River basin and is the only source of well-validated surface temperature observations. Site-averaged

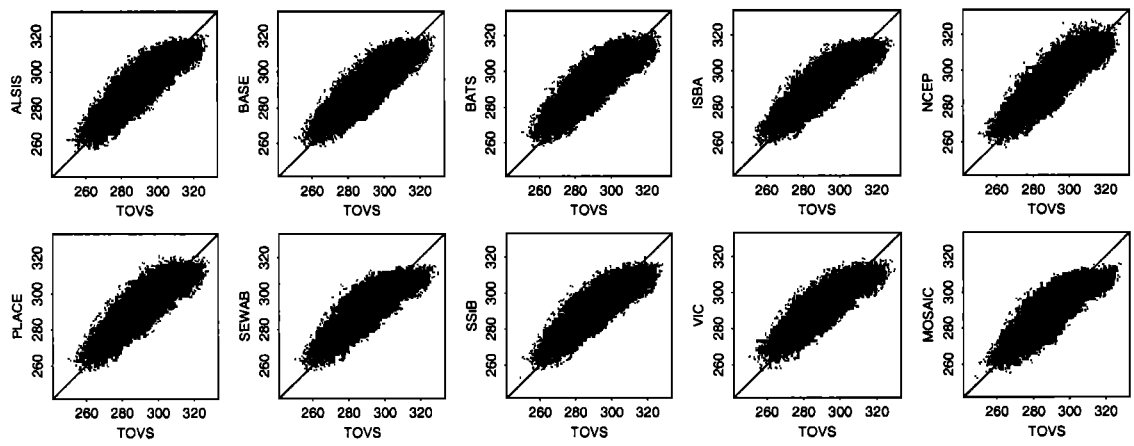

Figure 9. Scatterplots of PILPS-2c models versus TOVS surface temperature. Comparisons are for 1985 at 1430 LT overpass.

Table 1. Summary of PILPS-2c Models versus TOVS.

model	R^2	RMSD	Bias
ALSIS	0.80	7.18	1.37
BASE	0.79	7.19	0.53
BATS	0.78	7.74	2.32
ISBA	0.79	7.19	1.02
NCEP	0.77	8.02	2.74
MOSAIC	0.77	7.61	−1.24
PLACE	0.79	7.39	1.40
SEWAB	0.79	7.32	0.62
SSiB	0.79	7.32	0.34
VIC	0.79	7.42	1.17

Bias is calculated as (model − TOVS).

air temperature and surface temperature from *Betts et al.* [1998] were processed to match TOVS overpass times [*Lakshmi and Susskind*, 1999]. Differences between FIFE observed T_s and T_a were produced and are shown in Figure 11. Although this study is concerned with the 1430 LT time step, the FIFE observations were for 4 times of day (0230, 1430, 0730, and 1930 LT), all of which were used to include as wide a range of temperature differences as possible. Using 99% of the FIFE differences, i.e., removing 0.5% of the data from each extreme of the distribution, the acceptable range of $T_s - T_a$ was determined to be −7.8 to 14.2 K. Therefore when TOVS is cooler than the air temperature used to force VIC at the time of overpass by more than 7.8 K or warmer than T_a by more than 14.2 K, the data were removed.

After implementing this screening process, the warm bias of VIC relative to TOVS was reduced from 3.18 to 0.43 K, the R^2 value increased from 0.76 to 0.87, and the RMS difference decreased from 8.33 to 5.25 K (Figure 12). The screening process also affected spatial standard deviation. Table 2 contains monthly averaged spatial standard deviation values for before (σ_{sm}) and after (σ_{sms}) the screening process. TOVS spatial variability was reduced for all months, ranging from a decrease of 1.00 K in November to a decrease of 1.99 K in May. VIC values changed relatively little, ranging from a slight increase of 0.09 K in September to a decrease of 0.95 K in November.

There are a possible 22,265 observations in the basin using one overpass per day (61 pixels × 365 days); 10,004 observations were not available due to clouds or gaps in the TOVS scan swaths. The FIFE-based screening process reduced the remaining 12,261 by 3023 or 24.6%. Of the 3023 screened values 2813 were removed due to low T_s, while only 210 were removed due to high T_s. This is a desirable effect because T_s less than T_a in excess of 7.8 K at 1430 LT are more suspect than high temperatures. Additionally, excessively cold T_s could be the result of low cloud interference that was not detectable during the initial screening of the Pathfinder Path A data set.

7. Discussion and Conclusions

Modeled and satellite-derived surface temperatures show good agreement when aggregated in space or time, for example, monthly (temporal) mean temperatures agree at the pixel scale and basin (spatial) mean temperatures agree instantaneously. However, space-time variability is consistently higher for remotely sensed surface temperatures compared to model predictions. The true variability of surface temperature is probably between modeled and satellite-derived.

The largest differences between remotely sensed and modeled surface temperature variability occur primarily at times of maximum net radiation both diurnally and seasonally, i.e., afternoon and summer. The spatial variability of GOES T_s reaches a maximum that coincides approximately with maximum net radiation, while VIC σ_s does not. During these periods we

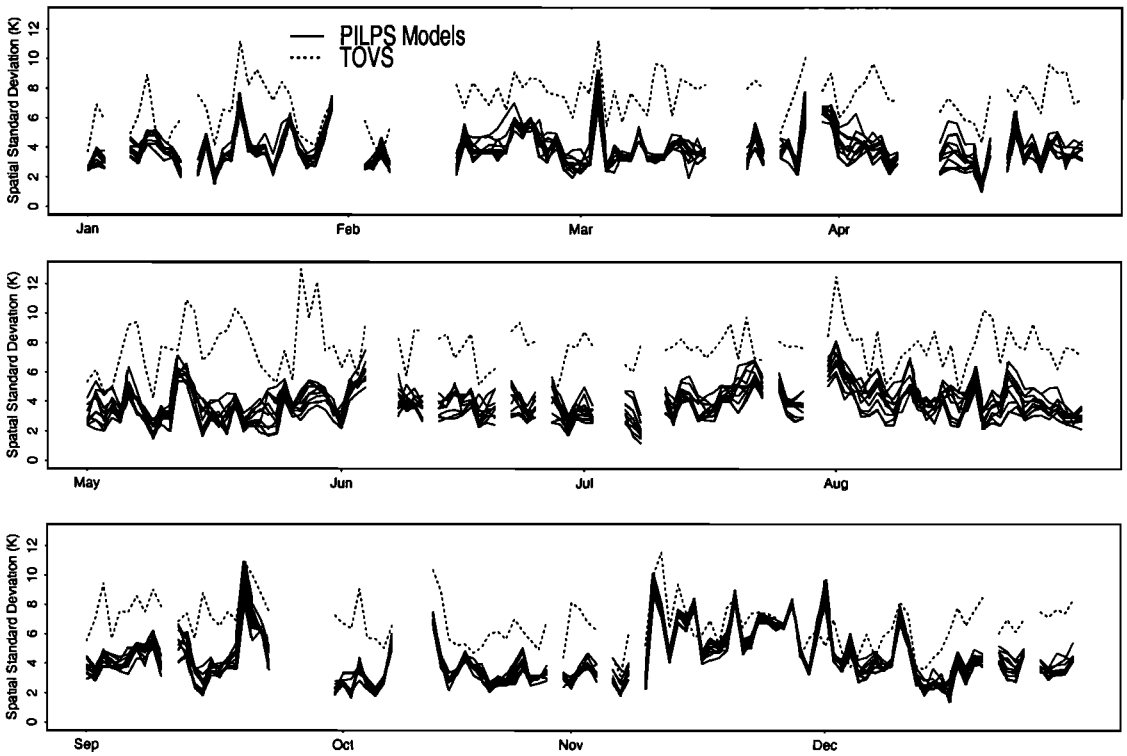

Figure 10. Time series of spatial standard deviation. PILPS-2c models and TOVS 1985 calculated daily at the 1430 LT overpass.

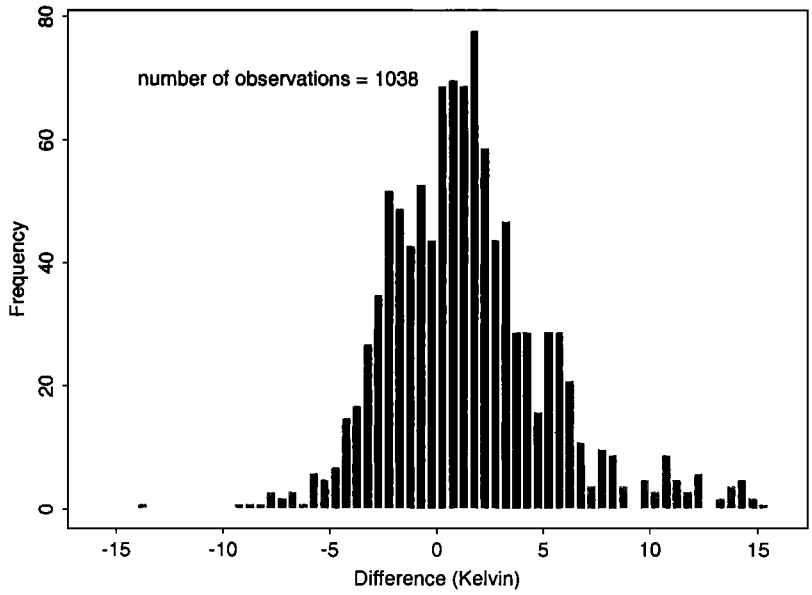

Figure 11. Histogram of $T_s - T_a$ differences from FIFE. FIFE observations were at 0230, 1430, 0730, and 1930 LT.

expect increased spatial variability since the control of heterogeneous land surface properties on surface temperature would be accentuated at times of maximum energy exchange. The depression of space-time variability of model predictions in summer and afternoon is therefore suspect.

The higher variability of the TOVS data relative to the model predictions may be caused in part by TOVS retrieval errors. Analysis of diurnal ranges and daily changes of surface temperature shows that the satellite data probably contain errors, e.g., excessively low temperatures, possibly as a result of cloud contamination, that artificially increase the range of remotely sensed surface temperatures.

TOVS and GOES both retrieved much higher surface temperatures than the model predicted. One possible reason is that modeled surface temperature is linked to air temperature. An initial surface temperature that represents the thermal state of the uppermost layer in the ground heat flux algorithm is derived from net radiation, and the parameterization of net radiation is partially

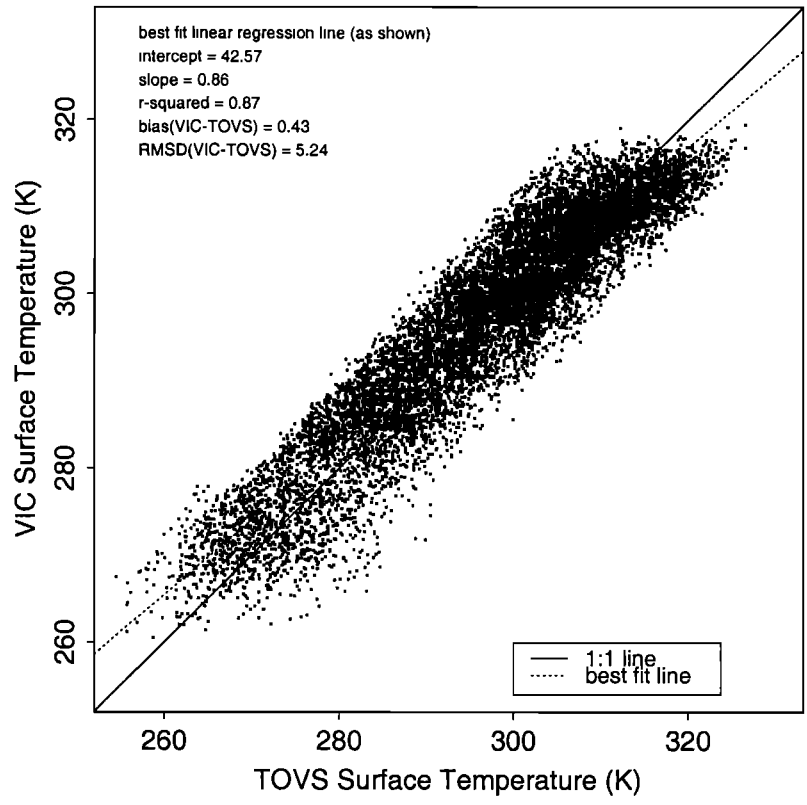

Figure 12. Scatterplot of TOVS versus VIC. 1985, 1430 LT T_s after screening based on FIFE observations.

Table 2. Monthly Mean Spatial Standard Deviation

month	TOVS σ_{sm}	TOVS σ_{sms}	difference	VIC σ_{sm}	VIC σ_{sms}	difference
Jan.	6.48	5.43	1.05	4.72	4.48	0.24
Feb.	7.00	5.35	1.65	4.52	4.11	0.41
March	7.87	6.58	1.29	3.78	3.67	0.11
April	7.44	5.82	1.62	3.68	3.49	0.19
May	7.80	5.81	1.99	3.34	3.32	0.02
June	7.61	5.81	1.80	3.43	3.47	-0.04
July	7.84	6.14	1.70	3.42	3.28	0.14
Aug	7.79	5.98	1.81	3.55	3.50	0.05
Sept	7.59	6.24	1.35	3.96	4.05	-0.09
Oct.	6.43	4.96	1.47	2.94	2.77	0.17
Nov.	6.80	5.80	1.00	5.63	4.68	0.95
Dec.	6.33	4.95	1.38	5.35	4.87	0.48

σ_{sm} = mean spatial σ before screening, σ_{sms} = mean spatial σ after screening, difference = $\sigma_{sm} - \sigma_{sms}$.

dependent upon air temperature. The initial VIC T_s is then iteratively modified to minimize energy balance errors. In contrast, remotely sensed radiometric surface temperature represents the thermal state of a very thin layer, on the order of millimeters, which can be interpreted as having near-zero thermal inertia. Methods that reconcile the physical differences between remotely sensed and modeled surface temperatures must be developed before the full potential of combining thermal remote sensing and land surface modeling can be realized.

The geolocation of land surface properties within a grid cell may also contribute to differences in variability. The VIC model calculates energy and water budgets, and thus surface temperature, for each grid cell N times, where N is the number of vegetation types for each cell and $2 \leq N \leq 6$. The final VIC output for each cell is an average of N computational model runs, weighted by the fractional coverage of each vegetation type. This mosaic scheme does not place vegetation types at any specific location within the cell. Conversely, satellites sense radiances at specific cloud-free locations. In the case of TOVS, surface temperature for a pixel could be based on cloud-free areas that represent as little as 20% of the scene. Additionally, satellite-derived T_s can be an average of multiple land cover types with satellite view angle, in combination with vegetation height and density, determining the percentages of vegetated versus bare surface that will be sensed by the satellite. These differences between remote sensing and modeling of surface temperature are accentuated at the $1° \times 1°$ scale because land surface heterogeneity occurs at much smaller scales.

Variability in land surface emissivity has been ignored in this study. The VIC model always uses an emissivity of 1.00, TOVS emissivity is fixed at 0.85 over land, and the GOES split window equation assumes an emissivity of 1.00. The TOVS methodology [*Susskind et al.*, 1984] attempts to minimize the effects of uncertainties in emissivity. Incorporation of land surface variability in emissivity into model predictions and satellite retrievals could have important implications for understanding the spatial and seasonal variability of land surface temperature.

The VIC model was shown to be representative of land surface models for surface temperature comparisons with satellite data; that is, VIC is not an outlier among SVAT models. The comparisons with land surface models involved in the PILPS-2c experiment show that when driven by identical data, all of the models tested had reduced space-time variability of surface temperature relative to satellite data. The question that remains is, at the scale of this inquiry, which source of surface temperature provides a truer representation of temporal and spatial variability? The remote sensing retrievals, whether by GOES or TOVS, are not direct measurements but rather the result of a modeling process. Thus one possible explanation, given the assessment of the PILPS-2c models, is that the satellite retrievals have inflated variability due to errors. The error-screening process performed with FIFE data did, in fact, reduce some of the variability but not enough to bring TOVS and PILPS-2c models into agreement. However, there are some possible sources for the reduced variability seen in the models. The forcing data used to drive these models may not be realistically variable because of aggregation, interpolation, or parameterization methods. Additionally, model inputs of land cover type may not adequately represent land surface heterogeneity.

Given the complexity of the computational structure of land surface models and the limits of extrapolating continuous fields of model inputs from point data sources, the use of satellite data as a source of unmodeled surface temperature variability may prove to be an effective means of incorporating this variability. However, if satellite data are to be used in this fashion, much more validation of satellite retrievals of land surface temperature will be required.

References

Betts, A. K., and J. H. Ball, FIFE surface climate and site-average data set 1987-89, J. Atmos. Sci., *55*(7) 1091-1108, 1998.

Bras, R., Hydrology, 643 pp., Addison-Wesley-Longman, Reading, Mass., 1990.

Czajkowski, K. P., S. N. Goward, and H. Ouaidrari, Impact of AVHRR filter functions on surface temperature estimation from the split window approach, Int. J. Remote Sens., *19*(10) 2007-2012,1998.

Daley, C., R. Neilson, and D. Phillips, A Statistical-topographic model for mapping climatological precipitation over mountainous terrain, J. Appl. Meteorol., *33*, 140-158, 1984.

Dickinson, R. E., A. Henderson-Sellers, P. J. Kennedy, and M.F. Wilson, Biosphere-Atmosphere Transfer Scheme (BATS) for the NCAR community climate model, NCAR Tech. Note, *NCAR/TN275+STR*, 69 pp., Natl. Cent. for Atmos. Res., Boulder, Colo., 1986.

Dickinson, R. E., A. Henderson-Sellers, and P. J. Kennedy, Biosphere-Atmosphere Transfer Scheme (BATS) version 1e as coupled to the NCAR Community Climate Model, NCAR Tech. Note, *TN-387+STR*, 80 pp., Natl. Cent. for Atmos. Res., Boulder, Colo., 1993.

Dubayah, R. O., E. F. Wood, E. T. Engman, K. P. Czajkowski, M. Zion, and J. D. Rhoads, Remote sensing in hydrological modeling, in *Remote Sensing in Hydrology and Water Management*, edited by G.A. Schultz and E.T. Engman, Springer-Verlag, New York, 2000.

Eidenshink, J. C. and J. L. Faundeen, The 1 km AVHRR global land data set – First stages in implementation, Int. J. Remote Sens., *15*(17), 3443-3462, 1994.

Jin, M., R. E. Dickinson, and A. M. Vogelmann, A Comparison of CCM2-BATS skin temperature and surface-air temperature with satellite and surface observations, J. of Clim., *10*, 1505-1524, 1997.

Kimball, J. S., S. W. Running, and R. Nemani, An improved method for estimating surface humidity from daily minimum temperature, Agric. For. Meteorol., *85*, 87-98, 1997.

Lakshmi, V., and J. Susskind, Comparison of TOVS-derived land surface variables with ground observations, J. Geophys. Res., *105*, 2179-2190, 1999.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges, A simple hydrologically based model of land surface water and energy fluxes for GCMs, J. Geophys. Res., *99*, 14,415-14,428, 1994.

Liang, X., D. P. Lettenmaier, and E. F. Wood, One-dimensional statistical dynamic representation of subgrid spatial variability of precipitation in the two-layer variable infiltration capacity model, J. Geophys. Res., *101*, 21,403-21,422, 1996.

Liang, X., E. F. Wood, and D. P. Lettenmaier, Modeling ground heat

flux in land surface parameterization schemes, J. Geophys. Res., *104*, 9581-9600, 1999.

Pfaendtner, J., S. Bloom, D. Lamich, M. Seablom, M. Sienkiewicz, J. Stobie, and A. da Silva, Documentation of the Goddard Earth Observing System data assimilation system - Version 1, *NASA Tech. Memo.* 104606, vol. 4, 1995.

Rawls, W. J., L. R. Ahuja, D. L. Brakensiek, and A. Shirmohammadi, Infiltration and soil water movement, in *Handbook of Hydrology*, edited by D. Maidment, McGraw-Hill, New York, 1993.

Shepard, D. S., Computer mapping: The SYMAP interpolation algorithm, in Spatial Statistics and Models, edited by G. L. Gaile and C. J. Wilmott, pp. 133-145, D. Reidel, Norwell, Mass., 1984.

STATSGO, U.S. State Soil Geographic (STATSGO) database: Data use information, *Tech. Rep. 1492*, U.S. Dep. of Agric., Washington, D.C., 1994.

Susskind, J., J. Rosenfield, D. Reuter, and M. T. Chahine, Remote sensing of weather and climate parameters from HIRS2/MSU on TIROS-N, J. Geophys. Res., *89*, 4677-4697, 1984.

Susskind, J., P. Piraino, L. Rokke, L. Iredell, and A. Mehta, Characteristics of the TOVS Pathfinder Path A data set, Bull. Am. Meteorol. Soc., *78*(7), 1449-1472, 1997.

Thornton, P. E., and S. W. Running, An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation, Agric. For. Meteorol., *93*, 211-228, 1999.

Wood, E. F., et al., The Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) phase 2(c) Red-Arkansas River basin experiment, 1, Experiment description and summary intercomparisons, *Global and Planet. Change*, *19*(1-4), 115-135, 1998.

R. Dubayah and J. Rhoads, Department of Geography, University of Maryland, College Park, MD 20742-8225.

G. O'Donnell and D. Lettenmaier, Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195.

V. Lakshmi, Department of Geological Sciences, University of South Carolina, Columbia, SC 29208.

(Received September 21, 2000; revised February 26, 2001; accepted February 28, 2001.)