A comparison of MODIS land surface temperature with in situ observations

Michael G. Bosilovich[1]

Received 10 July 2006; revised 13 September 2006; accepted 20 September 2006; published 31 October 2006.

[1] MODerate resolution Imaging Spectroradiometer (MODIS) land surface temperatures (LSTs) are compared to in situ observations during the Coordinated Enhanced Observing Period (CEOP). The purpose is to test the utility of global enhanced station data to provide additional information on the consistency of large volumes of remotely sensed data. While comparisons are limited by unresolved spatial and temporal representativeness, many of the comparisons are quite favorable, especially in mid-latitude regions. We note the extent of cloud contamination in the data product, and also some biases that may vary seasonally. Upscaling to 25km, as would be needed for global model comparisons or some mesoscale models, did not overly change the comparison results. The veracity of remotely sensed observations is important to identify and understand as these data begin to be applied to research questions. **Citation:** Bosilovich, M. G. (2006), A comparison of MODIS land surface temperature with in situ observations, *Geophys. Res. Lett.*, *33*, L20112, doi:10.1029/2006GL027519.

1. Introduction

[2] Global land surface temperature observations are a crucial component in the study of the surface energy and water budgets [*Park et al.*, 2005; M. G. Bosilovich et al., Skin temperature analysis and bias correction in a coupled land-atmosphere data Assimilation System, submitted to *Journal of the Meteorological Society of Japan*, 2006], global and regional climates [*Wan et al.*, 2004a; *Jin et al.*, 2005] and epidemiology [*Rogers et al.*, 2002; *Tatem et al.*, 2004], to name a few examples. The MODerate resolution Imaging Spectroradiometer (MODIS) instrument (onboard Aqua and Terra satellites) level 1 and level 2 data are produced at extremely fine spatial scales, from a global point of view (250 m to 1 km). MODIS calibration/validation stations provide ground based observations for validation data. These sites are very well controlled, and have general spatial homogeneity and instantaneous correlation with instrument overflights. The validation exercises generally consider this pixel level data for relatively small numbers of locations and observations (e.g., *Wan et al.* [2002, 2004b] for land surface temperature).

[3] More extensive well-controlled experiments are expensive and presently not available. Yet data is required for many different conditions through the progression of seasons and over a variety of climate regimes. Here, we will compare the MODIS land surface temperature (LST) level 3 product with a multitude of in-situ station observations over a three month period during the Coordinated Enhanced Observing Period (CEOP) EOP1 (July through September 2001). While some stations are listed in the EOS validation suite [*Morisette et al.*, 2002], in-situ LST is not presently available there. The present comparison is not controlled as tightly as was done in the MODIS calibration and validation experiments. However, we consider that this is more of a real world exercise that may be more generally applicable to the typical data user, for example, someone who might use this data to validate a global land parameterization.

2. MODIS Level 3 (5 km) LST

2.1. Observation Data

[4] The MODIS Level 3 (5 km) LST data are gridded uniformly across the globe [*Hall et al.*, 2002; *Brubaker et al.*, 2005]. A daytime and nighttime pass are provided once per day at each grid box for cloud free pixels, and following a certain amount of quality control. A time stamp is provided for each 5 km grid box. The data used here is based on the MODIS collection 4 [*Wan et al.*, 2004b].

[5] The in situ data has been collected at a number of stations, all contributing to the Coordinated Enhanced Observing Period (CEOP). Table 1 shows the station names and geographical location. The data was collected in conjunction with the first testing phase of CEOP from 1 July through 30 September, 2001 [*Bosilovich and Lawford*, 2002]. The stations measure surface temperature with radiometers. The stations are managed by different institutions and have different instruments. The MODIS LST is defined by the radiation emitted from the surface [*Wan et al.*, 2004b], which should be generally similar to the station observations. However, the comparison must consider several sources of representativeness differences (discussed below).

[6] For this period, only MODIS Terra (Collection 4) is available from 4 July 2001 onward. The station data is quality controlled by the station managers, and also during a reformatting process for CEOP [*Lawford et al.*, 2006]. The instruments at the stations are infrared thermometers. In order to compare the station data with the MODIS level 3 LST, we determine the time of each available MODIS observation at the closest 5 km grid point to the station location. Using that time, we find the closest (in time) in situ observation from each station. The time series of both MODIS and station temperatures used to calculate the statistics and create the figures can be found in the auxiliary material.[1]

[1]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

Published in 2006 by the American Geophysical Union.

[1]Auxiliary materials are available at ftp://ftp.agu.org/apend/gl/2006gl027519.

Table 1. Station Locations for the in Situ Data From the CEOP Data Archives[a]

Station	Location	Latitude, deg	Longitude, deg
GAPP	ARM SGP	36.61	−97.49
GAPP	Bondville	40.01	−88.29
GAPP	Fort Peck	48.31	−105.10
LBA	Rondonia	−10.08	−61.93
CAMP	Mongolia(130)	45.74	106.26
CAMP	Mongolia(330)	46.13	106.37
CAMP	Mongolia(230)	46.21	106.71
CAMP	Mongolia(430)	46.78	107.14
ARM	NSA(Atquasuk)	70.47	−157.41
ARM	NSA (Barrow)	71.32	−156.61
ARM	TWP (Manus)	−2.06	147.43
ARM	TWP (Darwin)	−0.52	166.92

[a]The acronyms are GEWEX Americas Prediction Project (GAPP), LBA, CAMP, Atmospheric Radiation Measurement (ARM), Southern Great Plains (SGP) North Slope of Alaska (NSA) and Tropical Western Pacific (TWP).

[7] There are two fundamental differences in the spatial and temporal representativeness between the station data and 5 km MODIS data. First, the station data are hourly averaged and the time stamp is the center of the mean, while MODIS LST are instantaneous relative to the over flight of the instrument. Second, the MODIS level 3 data are aggregated to the 5 km grid center, while the observations are point locations and may not represent the grid box area. Even in regions of apparent surface homogeneity, spatial variability of the surface temperature can add uncertainty to the comparisons. This spatial and temporal variability can increase the uncertainty of the comparison. In addition, while the station data is documented, the stations may not be stringently calibrated as those in the validation suite. For these reasons, this study cannot be considered to be validation of the MODIS instrument or LST algorithm, but rather an evaluation or comparison.

[8] Despite the limitations discussed previously, we want to make the comparison to determine if the station data can be used to make some qualitative assessment of the MODIS data. These representativeness issues will be present in any study that utilizes the level 3 data. By using stations at many locations, we can identify systematic variations between the remotely sensed and station observations. Therefore, we compare the instrument data with observations from around the globe, and over several months.

2.2. The 5 km Comparison

[9] Table 1 shows the stations available during CEOP (including their acronyms and affiliations) that provide in-situ measurements of surface temperature. The sites represent several climate regimes, including plains, forest and high latitudes. Table 2 shows the mean of the available MODIS level 3 LST 5 km observations, and the statistics (mean difference, standard deviation of the differences and correlation) when compared to the matching in-situ measurements. There are some very good comparisons. However, over the TWP sites there are substantial differences, possibly related to tropical water vapor or cloudiness near the island station. With such clear discrepancy between the different observations, we cannot learn much from that comparison and will not pursue in depth in this study. It does identify a region of interest for further diagnosis in the algorithm development or data reprocessing. Over the continental plains, (CAMP and GAPP) the comparisons seem quite reasonable, though not all statistical values are always favorable. It may be important to note that in almost all the cases the mean bias shows that the remotely sensed observation is lower than the in situ measurements.

[10] Figure 1 shows a scatter diagrams for four of the stations. The ARM Southern Great Plains site night time temperatures are quite comparable to the MODIS retrievals over the period. However, there is a distinct difference between the observations at high daytime temperatures (this will be evaluated later in this section). The Bondville site also shows generally reasonable comparison during the night, with the exception of two outliers, where MODIS temperatures are much lower than the in situ observation. In general, there is more scatter in the daytime retrievals than the nighttime (Figure 1 and Table 2). At the Fort Peck site, the station temperatures are regularly higher than MODIS in both day and night (Figure 1). The daytime high station bias is greatest when the temperatures are largest, and the bias is uniformly high at night. A similar pattern is apparent for the CAMP Mongolia station.

Table 2. MODIS Level 3 LST Data Compared to CEOP in Situ Station Measurements[a]

		Day				Night			
5 km Station	Location	Mean	Diff	SD	Corr	Mean	Diff	SD	Corr
GAPP	ARM SGP	312.0	−4.1	4.19	0.84	295.9	−0.5	1.35	0.98
GAPP	Bondville	301.3	0.8	3.52	0.50	289.1	−1.0	2.18	0.90
GAPP	Fort Peck	307.1	−6.8	4.32	0.80	286.3	−2.3	1.27	0.97
LBA	Rondonia	302.3	−0.7	2.01	0.52	295.6	0.6	1.19	0.76
CAMP	Mongolia(130)	309.2	−3.3	3.91	0.90	285.1	−2.2	2.76	0.86
CAMP	Mongolia(330)	309.6	−0.9	3.88	0.88	284.1	−1.9	3.33	0.84
CAMP	Mongolia(230)	308.8	0.3	3.98	0.84	285.6	−2.1	2.44	0.89
CAMP	Mongolia(430)	308.2	−1.1	3.84	0.86	283.4	−3.6	2.43	0.89
ARM	NSA (Atquasuk)	279.1	−3.7	3.23	0.90	270.5	−5.1	5.14	0.59
ARM	NSA (Barrow)	274.6	−5.7	6.25	0.50	269.4	−5.3	5.24	0.26
ARM	TWP (Manus)	299.2	−14.0	3.27	0.54	298.5	−0.5	3.38	−0.20
ARM	TWP (Darwin)	305.9	−1.9	6.09	0.04	294.4	−4.0	3.19	−0.12

[a]Mean indicates the MODIS mean temperature averaged for each measurement in the time series at that location for the period of study. Similarly, Diff indicates the mean difference from the reference site data, SD indicates the standard deviation of the differences, and Corr is the Correlation between the two observed time series. Day passes are separated from night passes. Units are Kelvin, except for correlation (nondimensional).

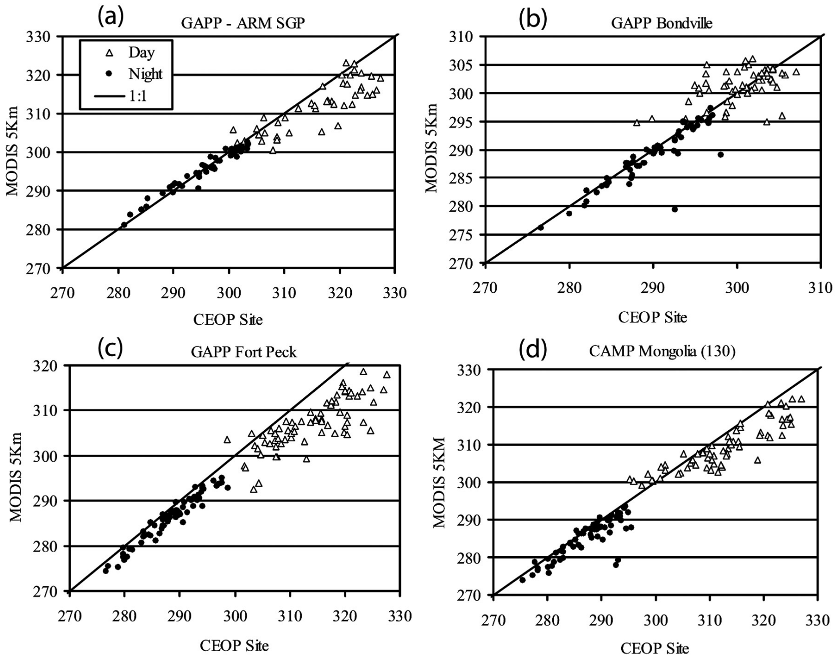

Figure 1. Scatter diagram for surface temperature from the MODIS 5 km LST product compared to station observations at each of 4 sites collected by CEOP, including (a) ARM SGP, (b) Bondville (c) Fort Pect and (d) CAMP Mongolia (130). Open triangles are daytime MODIS overflights, and closed circles are night time overflights. The solid line indicates 1:1. Units are Kelvin. Acronyms are defined in Table 1.

[11] Figure 2 shows the time series of MODIS daytime retrievals at ARM SGP and the coincident station observation for the three month period (July–September 2001). A nine point running average of each time series is also included for convenience. The bias generally decreases in time from summer into autumn. While the daytime temperatures are decreasing, following their seasonal cycle, the atmospheric water vapor and land cover would also be changing, following their seasonal cycles. A longer time series, and multiple years would be useful to determine better the characteristics of this bias. However, the time series of Fort Peck and Mongolia show similar temporal variations in the bias as Figure 2 (not shown). The offset in time between the MODIS overflight and the center of the hourly average station measurement shows some sensitivity during the daytime; however, this is small in comparison to the time series of the bias and does not explain the seasonal variations.

[12] Several night time outlying points at the Bondville site have been evaluated further, and one case is presented here. On 19 September 2001, the MODIS report was 281.9K while the concurrent in situ measurement was 292.6K. Figure 3 shows the map of the MODIS 5km level 3 LST data for this day. The black box encompasses the Bondville site (at the center), outlining the area of 5x5 MODIS 5 km boxes. At this time there was substantial cloudiness. The white areas of Figure 3 do not show any temperature values, generally because of the presence of clouds. In the vicinity of the station, there are broken areas where some retrievals were made, but some areas where they were not. The range of the level 3 observations is from 274K to 294K. Given the predominance of clouds near the station, it seems apparent that the lowest values have some cloud contamination. The 1km level 2 data from the core validation archive [*Morisette et al.*, 2002] shows that most of the 200km^2 validation area is cloudy. This data set also shows a few points, as low as 230K, which are not apparent in the 5 km gridded data.

3. Spatial Representativeness

[13] Most general circulation models and global data assimilation systems have grid spacing around ½ - 2 degrees, though some global numerical predictions are being run at resolutions of 20 km. In order to better

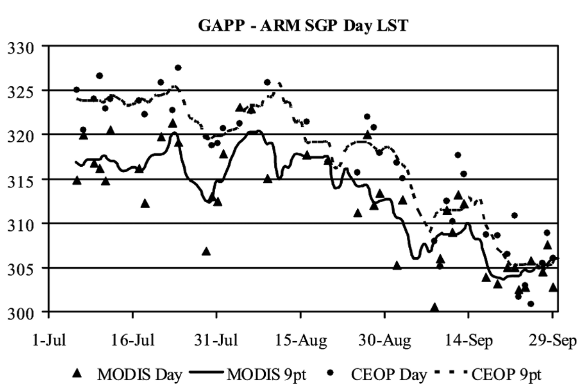

Figure 2. Time series of daytime MODIS 5 km LST (solid triangles) and the in situ observations from the ARM SGP (solid circles). To emphasize the time variation of the mean bias, 9 day centered running averages of the time series' are also plotted. Only in situ data are used when a corresponding MODIS observation is available. When data drops out, the time average is only for days that are available, so the average may be based on less than 9 values. Units are Kelvin.

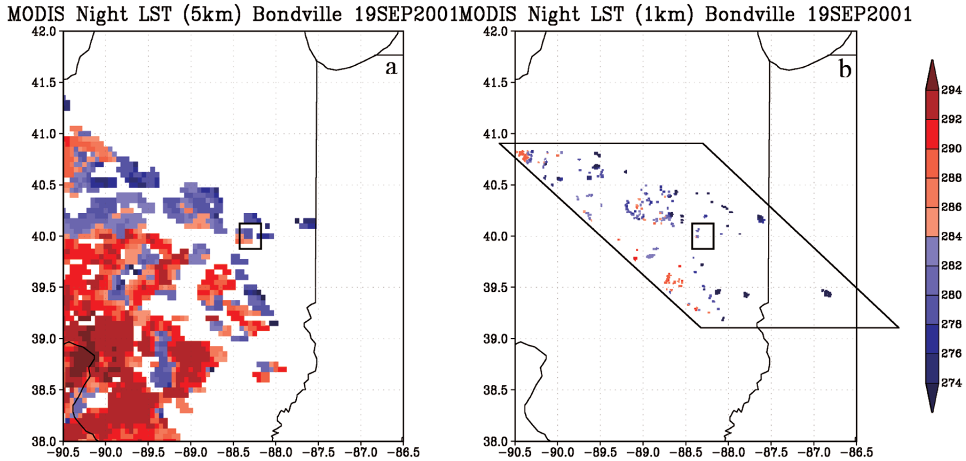

Figure 3. MODIS (a) 5 km LST Climate Model Grid data and (b) 1 km Core Validation site data on 19 Sep 2001 for in the vicinity of the Bondville site (located near the center of the black box). The size of the black box is 5 × 5, MODIS 5 km spaces. The polygon in Figure 3b denotes the 201 × 201 1 km points provided by the Core Validation archive. While the colors are plotted with identical intervals for comparison purposes, a cluster of points in Figure 3b to the east of the Illinois-Indiana border register 230K. Units are Kelvin.

understand the spatial representativeness of the comparison, we upscaled the MODIS data to a 5x5 area around the site locations (nominally referred to as the 25 km data). Table 3 reproduces Table 2, using instead the 25 km MODIS values. Using a larger area allows more days to provide an observation for the comparison. In general, the statistics are quite similar. Scatter diagrams (as in Figure 1, but not shown) identified more outlying points than the 5 km comparison. For these additional outliers, maps of surface temperature show that the box average was represented by one or only a few points, and these appeared too low compared to surrounding area (similar to Figure 3), and presumably clouds were interfering with the retrieval. Even with increasing disparity due to spatial representativeness in the comparison, it seems that we can utilize surface station temperature records, without stringent calibration and validation control, to better evaluate the character of the remotely sensed data products. Extending the spatial range of the MODIS observations to 25 km allows for increasing number of data points in the comparisons.

4. Summary and Future Directions

[14] Here, we compare remotely sensed observations with openly available in-situ surface temperature measurements. The exercise demonstrates that we can gain some understanding of the remotely sensed data. While the in-situ measurements are likely reliable, it cannot be ruled out that some drift in the instruments, or very fine scale spatial variability is affecting the comparisons. Also, the MODIS team recognizes the continued presence of cloud contamination and provides some documentation to effect further quality control that cannot be done during the operational processing. The cloud contamination was not more aggressively screened here for illustrative purposes, to identify the extent that contaminated points can be seen in this type of comparison.

[15] Each of the data sets used in this study evolved during 2006. MODIS collection 5 is currently being processed. The revised algorithm should have improved cloud clearing. The CEOP main data collection period, October 2002 through December 2004, has completed and the new observations are being compiled and quality controlled in the same way as the test period. By lengthening the time series, we will be able to further study the seasonality in the records, as well as evaluate the updated cloud clearing algorithm. While CEOP includes global distributions, some sites, such as ARM SGP, have records as long as the MODIS Aqua and Terra. These could be used to evaluate the interannual variability of the MODIS product. In addi-

Table 3. As in Table 2, Except for Upscaled (25 km) Averages of MODIS

		Day				Night			
25 km Station	Location	Mean	Diff	SD	Corr	Mean	Diff	SD	Corr
GAPP	ARM SGP	311.7	−4.2	4.42	0.82	295.8	−0.7	1.42	0.97
GAPP	Bondville	300.8	0.6	3.68	0.52	288.8	−1.6	3.46	0.78
GAPP	Fort Peck	306.0	−7.1	4.43	0.80	285.9	−2.7	1.94	0.93
LBA	Rondonia	302.1	−0.6	2.22	0.54	295.1	0.0	1.32	0.68
CAMP	Mongolia(130)	307.6	−3.5	4.46	0.88	284.8	−2.8	2.30	0.90
CAMP	Mongolia(330)	308.8	−1.3	3.73	0.89	283.9	−2.7	3.60	0.81
CAMP	Mongolia(230)	308.9	0.3	3.75	0.85	284.7	−3.2	2.47	0.89
CAMP	Mongolia(430)	307.5	−1.4	3.44	0.91	283.1	−4.0	2.72	0.86
ARM	NSA (Atquasuk)	278.1	−3.9	4.55	0.78	269.8	−6.6	6.44	0.33
ARM	NSA (Barrow)	275.0	−4.9	5.09	0.59	268.4	−6.0	5.11	0.15
ARM	TWP (Manus)	297.6	−14.7	3.83	0.44	296.3	−2.7	4.37	0.00
ARM	TWP (Darwin)	302.4	−5.4	5.01	0.08	295.5	−2.9	2.80	−0.19

tion, CEOP is collecting global operational analyses for numerical weather prediction centers, and some field experiments, like ARM SGP and CAMP, have multiple stations in and are of a couple hundred square kilometers. These sites, along with MODIS observations could be used to investigate the spatial representativeness and uncertainty in remotely sensed observations and heterogeneity in land parameterizations.

[16] **Acknowledgments.** The contributions by the reference station managers to the CEOP data base is greatly appreciated. Steve Williams quality assurance on the reference data greatly added to the accessibility of the station data. Zhengming Wan, Ana Pinhiero, Jeff Morisette and Christa Peters-Lidard provided many discussions on the MODIS LST data set. Two anonymous reviewers' comments helped improve the quality and clarity of the final manuscript.

References

Bosilovich, M. G., and R. Lawford (2002), Coordinated Enhanced Observing Period (CEOP) International Workshop, *Bull. Am. Meteorol. Soc.*, *83*, 1495–1499.

Brubaker, K. L., R. T. Pinker, and E. Deviatova (2005), Evaluation and comparison of MODIS and IMS snow-cover estimates for the continental United States using station data, *J. Hydrometeorol.*, *6*(6), 1002–1017.

Hall, D. K., G. A. Riggs, V. V. Salomonson, N. E. DiGirolamo, and K. J. Bayr (2002), MODIS snow-cover products, *Remote Sens. Environ.*, *83*(1–2), 181–194.

Jin, M. L., R. E. Dickinson, and D. L. Zhang (2005), The footprint of urban areas on global climate as characterized by MODIS, *J. Clim.*, *18*(10), 1551–1565.

Lawford, R., et al. (2006), U.S. contributions to the CEOP, *Bull. Am. Meteorol. Soc.*, *87*, 927–939.

Morisette, J. T., J. L. Privette, and C. O. Justice (2002), A framework for the validation of MODIS Land products, *Remote Sens. Environ.*, *83*(1–2), 77–96.

Park, S., J. J. Feddema, and S. L. Egbert (2005), MODIS land surface temperature composite data and their relationships with climatic water budget factors in the central Great Plains, *Int. J. Remote Sens.*, *26*(6), 1127–1144.

Rogers, D. J., S. E. Randolph, R. W. Snow, and S. I. Hay (2002), Satellite imagery in the study and forecast of malaria, *Nature*, *415*(6872), 710–715.

Tatem, A. J., S. J. Goetz, and S. I. Hay (2004), Terra and Aqua: New data for epidemiology and public health, *Int. J. Appl. Earth Obs.*, *6*, 33–46.

Wan, Z. M., Y. L. Zhang, Q. C. Zhang, and Z. L. Li (2002), Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data, *Remote Sens. Environ.*, *83*(1–2), 163–180.

Wan, Z., P. Wang, and X. Li (2004a), Using MODIS land surface temperature and normalized difference vegetation index products for monitoring drought in the southern Great Plains, USA, *Int. J. Remote Sens.*, *25*(1), 61–72.

Wan, Z., Y. Zhang, Q. Zhang, and Z. L. Li (2004b), Quality assessment and validation of the MODIS global land surface temperature, *Int. J. Remote Sens.*, *25*(1), 261–274.

M. G. Bosilovich, Global Modeling and Assimilation Office, Earth Sciences Directorate, NASA GSFC Mail Code 610.1, Greenbelt, MD 20771, USA. (michael.g.bosilovich@nasa.gov)