Simultaneous Land Surface Temperature–Emissivity Retrieval in the Infrared Split Window

DAVID A. FAYSASH AND ERIC A. SMITH

Department of Meteorology, The Florida State University, Tallahassee, Florida

(Manuscript received 16 April 1998, in final form 4 January 1999)

ABSTRACT

A combined land surface temperature–emissivity retrieval algorithm is developed and tested for Geostationary Operational Environmental Satellite (GOES)-Imager and National Oceanic and Atmospheric Administration Advanced Very High Resolution Radiometer (AVHRR) split-window channels. By assuming that the spectral emissivities are constant over a short time period (12–24 h), two sets of split-window radiance measurements taken at two different times are used to retrieve two spectral emissivities and two land surface temperatures (LSTs) simultaneously. The algorithm employs an optimization scheme rather than a direct solver for a system of equations because of constraint requirements. The retrieved variables minimize the rms differences between measured satellite radiances and those predicted by a spectrally detailed radiative transfer model.

A *GOES-8* version of the algorithm is validated with in situ radiometer measurements from the Department of Energy's Atmospheric Radiation Measurement Program Cloud and Radiation Testbed (ARM CART) site. In addition, an AVHRR version is validated with in situ measurements from the First ISLSCP Field Experiment (FIFE) site, the Hydrological Atmospheric Pilot Experiment–Sahel (HAPEX–Sahel) site, and an LST validation site operated near Melbourne, Australia. The biases of the retrieved LSTs for the validation sites in the Australian, FIFE, and ARM CART study areas are approximately 0.08°, 1.7°, and 1.4°C, respectively, yielding an overall bias error of better than half the current expected accuracy limit of some ±3°C. The associated bias-adjusted rmse differences are approximately 0.78°, 4.8°, and 4.5°C, respectively, mostly driven by intercomparing in situ point measurements to area-integrated satellite pixel retrievals. The bias-adjusted rmse differences for HAPEX–Sahel are larger (5° and 11°C), resulting from incomplete characterization of site heterogeneity, insufficient radiosonde launch frequency, and poor data quality of the temperature–moisture soundings, rather than intrinsic algorithm problems. Notably, the averaged retrieved emissivities for the trouble-free sites are within the expected range of emissivities for vegetated surfaces.

The *GOES-8* retrieved LSTs exhibit small amplitude, high-frequency noise, and a daily error cycle when compared to in situ measurements. The noise is attributed to random detector errors in the satellite observations for which the channel 4 noise-equivalent temperature difference is larger than that of channel 5. The systematic differences between validation measurements and retrievals are near zero during nighttime but exhibit a small semidiurnal oscillation during daytime. Notwithstanding a possible semidiurnal bias in the pyrgeometer validation measurements associated with imperfect solar dome heating corrections, plus unaccounted-for attenuation between the surface and pyrgeometer, the latter error cycle is attributed to a too-coarse sampling of the nonlinear diurnal evolution of the thermodynamic structure of the atmospheric boundary layer, particularly near the sunrise and sunset transition times. Thus, sounding frequency determines the error characteristics of the nonlinearly evolving split-window weighting functions.

1. Introduction

A major unresolved problem in applications of satellite-based land surface temperature (LST) retrieval from split-window infrared measurements is accounting for variable spectrally selective surface emissivity. There are ample data sources for thermal split-window satellite measurements, particularly from the five-channel imagers on the National Oceanic and Atmospheric Administration (NOAA)'s current fleet of Geostationary and Polar-Orbiting Operational Environmental Satellites (GOES and POES), thus motivating the development of improved methods for addressing variable surface emissivity. This study addresses the problem through a relatively new type of retrieval approach often referred to as simultaneous retrieval to designate that the dependent variables of the retrieval inversion scheme have dissimilar physical units.

There have been a variety of LST retrieval algorithms developed for use with infrared split-window measurements (Price 1984; Becker and Li 1990; Sobrino et al. 1991; Prata 1993; Coll et al. 1994). These algorithms are generally of the form

Corresponding author address: Dr. Eric A. Smith, Department of Meteorology, The Florida State University, Tallahassee, FL 32306-4520.
E-mail: esmith@metsat.met.fsu.edu

$$T_s = a_0 + a_1 T_{11} + a_2 T_{12}, \quad (1)$$

where T_s is the LST (also referred to as skin temperature); T_{11} and T_{12} are satellite equivalent blackbody temperatures (EBBTs) measured in the ''clean'' (~10.5–11.5 μm) and ''dirty'' (~11.5–12.5 μm) portions of the infrared window (denoting the relatively ''weak'' and ''strong'' water vapor continuum absorption properties within the 10.5–12.5-μm spectrum); and a_0, a_1, a_2 are empirically, semiempirically, or physically derived retrieval coefficients. These algorithms take into account the attenuating effects of the atmosphere and nearly graybody effects of the earth's surface, recognizing that a differential emissivity may exist across the clean and dirty bands. The three coefficients differ among authors according to how the two effects are treated. Price (1984) and Becker and Li (1990) used coefficients that are functions of emissivities measured in both spectral intervals. Treatment of atmospheric effects was included in their coefficients, developed from a set of model atmospheres. Coefficients of several newer algorithms (Sobrino et al. 1991; Prata 1993; Coll et al. 1994) require knowledge of the spectrally distinct surface emissivities and coincident atmospheric temperature–water vapor structure to account for atmospheric attenuation.

Therefore, use of these algorithms in their standard forms requires knowledge of the atmospheric temperature–water vapor structure over the target area and/or surface emissivities for both spectral bands. Profiles of the atmosphere can be obtained from radiosondes launched within or near the LST target area, from sounding instruments on board weather satellites, or from global data assimilation models. Atmospheric correction terms for the algorithms can then be calculated as a function of the total atmospheric water vapor path derived from the profiles (Sobrino et al. 1991; Coll et al. 1994), a calculation that may include a radiative transfer model that uses the sounding data as input (Prata 1993).

The specification of surface emissivity for both window channels can be done in several ways, noting that accurate LST retrieval depends on accurate emissivity specification. As discussed by Becker (1987), an error of 5% in emissivity can cause a 2°C error in retrieved temperature. Figure 1 provides a graphical representation of LST retrieval error as a function of split-window differential emissivity error for five different average surface emissivities. Note that, as the average emissivity decreases, the retrieval error increases along with the underlying linear error increase with respect to specified differential emissivity error. For well-vegetated surfaces, Prata (1994a) has suggested a value of 0.98 to be used for both channels, invoking a first-order graybody assumption. Caselles et al. (1997) devised a formula for an effective spectral surface emissivity (ϵ_λ) of a satellite pixel as a function of retrieved fractional vegetation cover (f_v) and known leaf–soil component spectral emissivities ($\epsilon_{v,\lambda}$ and $\epsilon_{g,\lambda}$), given by

$$\epsilon_\lambda = f_v \epsilon_{v,\lambda} + (1 - f_v)\epsilon_{g,\lambda} + d\epsilon, \quad (2)$$

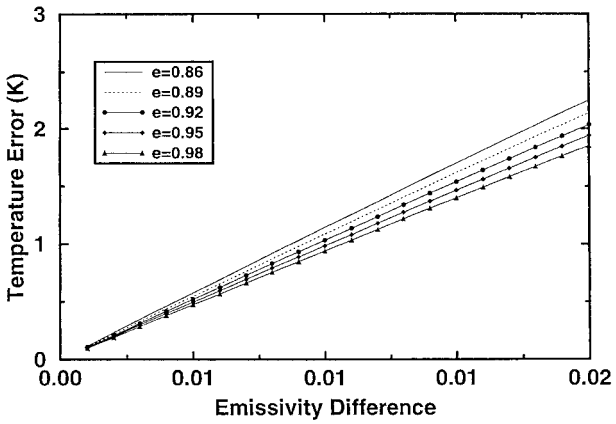

FIG. 1. Retrieved LST error as a function of differential emissivity error for five different average split-window emissivities from 0.86 to 0.98. Errors are derived from single-sample retrieval algorithm of Prata (1993).

where $d\epsilon$ is a term that deals with cavity effects. Fractional vegetation cover can be estimated from Normalized Difference Vegetation Index (NDVI) measurements while component emissivities are obtained from in situ measurements made within the target area (Caselles et al. 1997) or from published sources (Taylor 1979; Sutherland 1986; Rubio et al. 1997). It should be noted that NDVI requires measurements in both visible and near-infrared spectral bands.

In some applications, determination of the pixel-level emissivities using these approaches is not possible. For example, information on the vegetation and ground component emissivities may not be available due to the size or remote location of a target area. In the case of the NDVI approach, some environmental satellites (such as GOES) are equipped only with a visible sensor and not with a near-infrared sensor in the ~0.7–1.1-μm band, thus precluding the vegetation index approach. Of greater concern, as is evident in Fig. 1, is that the above schemes do not address the more fundamental issue that emissivity-based error in LST retrievals is driven mostly by error in the differential emissivity, which is variable within a dynamically changing biome.

To overcome the difficulties of spectral emissivity specification, several authors (e.g., Watson 1992; Xiang and Smith 1997) have examined the feasibility of simultaneous skin temperature–spectral emissivity retrieval schemes. Watson (1992) formulated a method based on using a number N channel radiances measured at M different times to retrieve the spectrally resolved surface emissivity along with a time-dependent set of temperatures. By assuming that a given spectral emissivity (ϵ_λ) is invariant over short time intervals, then $M \times N$ radiative transfer equations that govern the radiance measurements can be solved for $M \times N$ factors, that is, M skin temperatures and N spectral emissivities.

Watson (1992) developed a two-time version of this method for Thermal Infrared Multispectral Scanner

measurements and tested it on a small dataset. He did not report on the LSTs or their validation, instead focusing on a qualitative examination of retrieved emissivities. Xiang and Smith (1997) developed a modification of the method for the centimeter–millimeter spectrum (microwave), testing a two-time algorithm on several overpasses of Special Sensor Microwave Imager (SSM/I) passive microwave measurements over the African Sahel. Their results indicate an LST retrieval accuracy of perhaps 2.5°–3.0°C, but based on a small validation dataset. (They also found that three-time and greater versions were unwieldy because so few available, cloud-free pixels could be obtained beyond two successive overpasses, which is a problem in sampling with the SSM/I rather than a problem with the multitime simultaneous algorithm technique in general.)

In a related study, Li and Becker (1993) incorporated a nonsimultaneous, two-step method applied to a small dataset of channels 3, 4, and 5 infrared window measurements from the Advanced Very High Resolution Radiometer (AVHRR) flown on POES satellites, channels 4 and 5 being the split window. In this scheme, channels 3-, 4-, and 5-band emissivities are retrieved in the first step, using temperature-independent spectral indices and the assumption that measured daytime 3.7-μm reflectivity from channel 3 is the complement of 3.7-μm emissivity. In the second step, LSTs are retrieved from a two-channel split-window equation in which the coefficients of the equation are dependent on the retrieved band emissivities. No validation error analysis was provided.

2. Motivation for simultaneous LST–emissivity retrieval

One of the standard variables predicted by surface–atmosphere exchange models (also called biosphere or land surface process models) is surface upwelling longwave radiative flux (e.g., see Smith et al. 1993 or Colello et al. 1998). With an appropriate broadband emissivity, this variable can be transformed into a skin temperature, thus enabling the use of an observationally based LST data assimilation scheme in the biosphere model. The motivation for this research is to develop an algorithm capable of retrieving representative LSTs at a time interval that is commensurate with biosphere modeling (30–60 min) for assimilation into the Florida State University (FSU) Biosphere–Atmosphere Transfer Scheme Experimental Version (Ex-BATS) biosphere model (see Smith et al. 1993 and Crosson et al. 1993 for a description of Ex-BATS). A horizontally heterogeneous version of Ex-BATS produces output over a spatial grid at a half-hour time step. In this framework, LSTs retrieved from high-resolution GOES satellite imagery represent suitable data for assimilating into the biosphere model for improving model performance over a diurnal cycle, particularly for processes related to soil moisture.

The Department of Energy's Atmospheric Radiation Measurement Program Cloud and Radiation Testbed (ARM CART) site has been selected as the main study area for the retrieval and assimilation experiments, since this region is well instrumented. The ARM CART domain, encompassing an area of roughly 760 km × 320 km within Kansas and Oklahoma (Peppler et al. 1996), contains a network of surface meteorological and flux sites and a radiosonde station that is capable of high-resolution profiling for atmospheric correction; see Fig. 2 for a schematic representation and Table 1a for locations of the 16 ARM CART sites that were selected for validation. Note that each of these sites in the network provides measurements of pressure, temperature, mixing ratio, wind speed and direction, rainfall, downwelling solar and infrared fluxes, and soil moisture, which are the standard nine variables used for initializing and forcing a biosphere model.

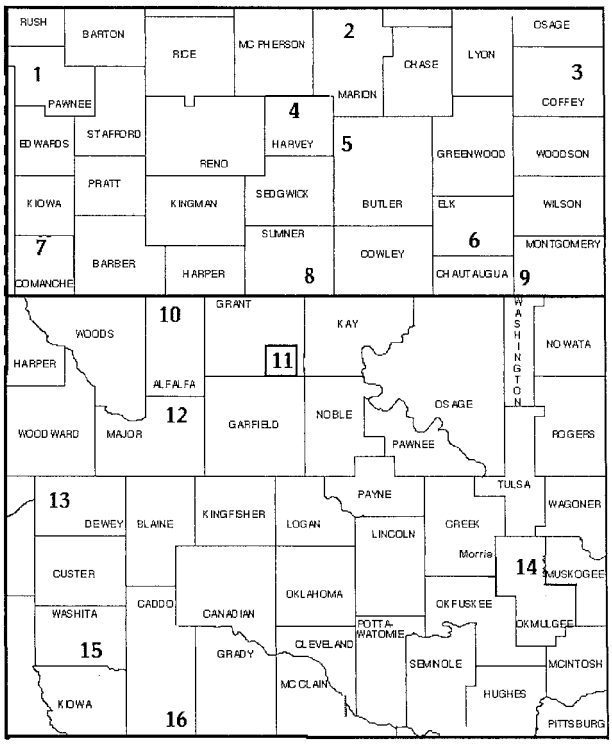

FIG. 2. Map of ARM CART domain (northern Oklahoma–southern Kansas) indicating locations of 16 validation sites. The central facility is collocated with site 11 enclosed by a box. Kansas–Oklahoma border is indicated by a solid line running horizontally through the map.

Land use maps show that the surface cover over ARM CART is a mixture of croplands, rangelands, pasture, and forest. The fractional vegetation covers of these biomes change with respect to their annual growth cycles, a process that is complementary with soil exposure. The effective surface emissivities also change, mainly because of the tendency of increased soil exposure, to lower emissivities within their typical range of 0.95–0.99 for partially to fully vegetated surfaces (van de Griend et al. 1991; Prata 1994a; Rubio et al. 1997), but

TABLE 1a. ARM CART pyrgeometer site locations.

Site	ARM CART site ID	Latitude (north)	Longitude (west)
1	EF-01	38.20	99.31
2	EF-02	38.30	97.30
3	EF-03	38.20	95.59
4	EF-05	38.14	97.51
5	EF-06	37.84	97.02
6	EF-07	37.38	96.18
7	EF-08	37.33	96.18
8	EF-09	37.13	97.26
9	EF-10	37.06	95.78
10	EF-11	36.08	98.28
11	EF-13	36.60	97.48
12	EF-15	36.43	98.28
13	EF-16	36.06	98.95
14	EF-18	35.68	97.85
15	EF-22	35.35	98.97
16	EF-24	34.88	98.20

TABLE 1b. FIFE IR thermometer site locations.

Site	FIFE site ID	Latitude (north)	Longitude (west)
1	0847-SAM	39.11	96.52
2	1463-PAM	39.10	96.48
3	2123-SAM	39.09	96.57
4	2139-PAM	39.09	96.53
5	2428-PAM	39.08	95.56
6	3221-PAM	39.07	96.58
7	4168-SAM	39.05	96.47
8	4439-PAM	39.05	96.54
9	4609-PAM	39.05	96.61
10	6469-PAM	39.01	96.47
11	6912-PAM	39.00	96.60
12	8639-SAM	38.97	96.54

also because of green–dead leaf emissivity variations. Since *GOES-8* does not measure radiation in a near-infrared channel (0.7–1.1 μm), the spectral emissivities cannot be estimated from fractional vegetation retrievals. Therefore it is necessary to retrieve the spectral emissivities simultaneously with the LSTs.

A simultaneous LST–emissivity retrieval algorithm, based on the multispectral/multitemporal methods of Watson (1992) and Xiang and Smith (1997), has been developed for large dataset applications with *GOES-8* split-window measurements. Section 3 describes the theory and methodology of the retrieval algorithm. Validation of the algorithm using *GOES-8* measurements as input has been conducted based on in situ ground measurements made over the ARM CART site. In addition, an AVHRR version of the algorithm has been validated using in situ ground measurements obtained from the First International Satellite Land Surface Climatology Project Field Experiment site (FIFE) (Sellers et al. 1992), the Hydrological Atmospheric Pilot Experiment–Sahel site (HAPEX–Sahel) (Prince et al. 1995; Goutorbe et al. 1997) (Tables 1b,c), and the LST validation site operated at Walpeup near Melbourne, Australia, by Prata (1994a). Retrieved skin temperatures are compared with the in situ measurements to determine accuracy and precision factors. Satellite, radiosonde, and in situ skin temperature data used from each experiment are described in section 4, while section 5 presents retrieval results and intercomparisons with the in situ measurements. Conclusions are presented in section 6.

3. Simultaneous retrieval theory and methodology

Although mainly intended for the *GOES-8* imager, the LST retrieval methodology developed here is applicable to any set of split-window measurements. Assuming that the earth's surface is a Lambertian emitter–reflector and that the atmosphere is nonscattering, then the radiation measured by a satellite in channel i at zenith angle Θ is a combination of the surface emission attenuated by the atmosphere, the upwelling atmospheric emitted radiation, and the downwelling atmospheric emitted radiation that is reflected by the earth's surface back toward the satellite. In radiative transfer form,

$$I_i(\Theta) = \epsilon_i \tau_i B_i(T_s) + I_i(\Theta)\uparrow + (1 - \epsilon_i)\tau_i F_i\downarrow/\pi, \quad (3)$$

ϵ_i is surface emissivity, τ_i is atmospheric transmittance, $B_i(T_s)$ is the Planck function evaluated at skin temperature T_s, $I_i(\Theta)\uparrow$ is radiation emitted by the atmosphere towards the satellite, and $F_i\downarrow/\pi$ is radiation emitted by the atmosphere that is reflected by the earth toward space. The three terms τ_i, $I_i(\Theta)\uparrow$, and $F_i\downarrow/\pi$ account for atmospheric correction and can be calculated using a radiative transfer model.

For one set of split-window measurements, such as channels 4 and 5 from *GOES-8*, Eq. (3) becomes an underdetermined system of two equations and three unknowns (T_s, ϵ_4, and ϵ_5). By assuming that the surface emissivity is time invariant over a short time interval (6–24 h) and introducing a second set of split-window measurements for a different time, Eq. (3) can be expanded to a closed system of four equations and four unknowns. An optimization technique called the truncated Newton minimization package (TNPACK) (Schlick and Folgelson 1992) is then used to find the four retrieval variables (two emissivities and two LSTs at two times) that minimize a cost function J:

$$J = \sum_{j=1,2} \sum_{i=4,5} (I_{\text{sat},i,j} - I_{i,j})^2, \quad (4)$$

where $I_{\text{sat},i,j}$ is the satellite radiance measured in channel i (GOES uses imager channels 4 and 5 for the split window) at time j, and $I_{i,j}$ is a radiative transfer equation (RTE) model-generated satellite radiance following Eq. (3). By linearizing the Planck function, it is possible to solve algebraically the system of equations represented by Eqs. (3) and (4). However, an optimization approach is preferred since physically realistic constraints cannot be placed on an algebraic solver, and noise in the channel radiances will produce pathological solutions in an algebraic framework. Based on field and laboratory

TABLE 1c. HAPEX–Sahel IR thermometer site locations and surface covers. Values in parentheses indicate number of measurements over specific type of surface cover.

Site	HAPEX–Sahel site ID	Surface cover	Latitude (north)	Longitude (east)
Southern supersite				
1	2424	Tiger bush (3), bare soil (2)	13.2508	2.2513
2	2924	millet (3), sand–weed (2)	13.2466	2.3094
3	2319	tiger bush (3)	13.2081	2.2436
Western supersite				
4	5655	Fallow (1)	13.5375	2.5681
5	5153	Millet (1)	13.5361	2.5147
Eastern supersite				
6	6855	Grass (1)	13.5635	2.6925
7	6754	Millet (1)	13.5488	2.6805

measurements, the lower and upper bounds placed on the emissivity solutions are 0.90 and 0.999, respectively. The upper range limit for emissivity is set just below 1.0 to prevent emissivity solutions that are greater than or equal to the physically valid upper limit. [Note that Platt and Prata (1993) have found satellite ''effective'' emissivities that were slightly greater than 1.0 for AVHRR pixels completely covered by vegetation, but that such retrievals do not represent true emissivity.] The surface temperatures are constrained to be within ± 10°C of the clean window EBBTs.

TNPACK uses a truncated Newton method to find solutions that minimize multivariate functions such as that expressed in Eq. (4). To use TNPACK, the exact form of the cost function and its gradient (first derivatives) is specified along with an initial guess to the solution. To achieve computational efficiency when doing retrievals for ARM CART *GOES-8* infrared image sectors (a typical sector is 180 $\times$ 80 pixels), the maximum allowable iteration count is set to 100, with the convergence criterion assigned to 0.0001 (defined as the absolute normalized difference between the cost function at two successive iterations). Initial guesses of the four retrieval variables are obtained by using a brute-force, low-resolution LST–emissivity scan of Eq. (4) across the allowable solution ranges, using the smallest J to select the initial guess variable set. Based on these settings and constraints, solutions typically are found in less than 10 iterations.

4. Description of datasets

The retrieval algorithm has been applied to *GOES-8* split-window measurements collected over the ARM CART study area and to AVHRR split-window measurements from *NOAA-9* and *NOAA-11* POES platforms collected over the FIFE study area in central Kansas (grassland prairie), HAPEX–Sahel study area in southern Niger (grassland Sahel), and Walpeup study area in southern Australia (wheat field).

Radiosonde profiles made for each of these experiments allow for calculation of atmospheric correction terms needed for the retrievals. The radiosonde data quality for ARM CART, FIFE, and Walpeup is very high, as these data were collected for research purposes at high vertical resolution under rigorous scientific management. The radiosonde data quality for HAPEX–Sahel was poor, with part of the problem stemming from nonquality-controlled measurements provided by the operational station at Niamey airport, and the rest of the problem stemming from equipment troubles at the HAPEX–Sahel experimental high-resolution sounding site. Expanded discussion on the radiosonde data and their processing for this study can be found in Faysash (1998).

In situ IR thermometers (IRTs), contact thermometers, or downfacing broadband flux pyrgeometers situated at various sites within the individual study areas provide the validation measurements. Tables 1a–c give site coordinates and identify the instruments used for ARM CART (pyrgeometers), FIFE (IRTs), and HAPEX–Sahel (IRTs). At Walpeup (35.12°S, 142.02°E), which was designed as an LST validation site for AVHRR, averages of nine contact thermometer measurements made over a 1 km $\times$ 1 km area are used for validation skin temperatures at the AVHRR pixel scale (Prata 1994b).

The FSU direct-readout GOES ground station was used to provide navigated *GOES-8* images for 28–29 September 1996 over ARM CART at full spatial resolution and a temporal resolution of 15 min during periods of the day when 15-min continental United States (CONUS) scanning was invoked. The calibration procedure for the infrared channels is discussed in Weinreb et al. (1997). Gu et al. (1998, manuscript submitted to *J. Geophys. Res.*) discuss residual uncertainties in the calibrated split-window measurements. A total of 130 time-matched images were acquired for the two days, corresponding to the standard complement of ~65–70 CONUS images generated per day. These images were cloud free with the exception of a small cluster of cumulus humilis that drifted over the southeast portion of the study area on the afternoon of 28 September.

Radiosonde and pyrgeometer measurements for ARM CART were obtained from the ARM CART Web site (http://www.arm.gov). Full-resolution AVHRR mea-

TABLE 2a. Summary information on satellite and radiosonde data used for FIFE.

Yearday	Date	Satellite overpass time (UTC)	Zenith angle	Radiosonde launch time (UTC)
155	4 Jun 1987	1031	31.48	1155
155	4 Jun 1987	2157	55.04	2132
156	5 Jun 1987	2147	46.29	21328
157	6 Jun 1987	1010	4.57	1149
157	6 Jun 1987	2136	34.89	2100
176	25 Jul 1987	2132	28.30	2146
177	26 Jul 1987	0955	25.40	1158
178	27 Jul 1987	0944	39.13	1149
179	28 Jul 1987	2100	20.85	2003
221	9 Aug 1987	2150	44.40	2131
222	10 Aug 1987	1013	5.12	1153
222	10 Aug 1987	2139	32.19	2256
223	11 Aug 1987	2128	16.91	2204
227	15 Aug 1987	2045	44.30	2005
228	16 Aug 1987	2034	53.42	2134
229	17 Aug 1987	1038	31.23	1148
229	17 Aug 1987	2204	55.86	2138
232	20 Aug 1987	1005	17.90	1159
233	21 Aug 1987	2120	3.77	2022
279	6 Oct 1987	2127	6.60	2106
280	7 Oct 1987	2116	10.42	2108
284	11 Oct 1987	2209	57.98	2211
285	12 Oct 1987	2203	49.98	2136

TABLE 2b. Summary information on satellite and radiosonde data used for HAPEX–Sahel.

Yearday	Date	Satellite overpass time (UTC)	Zenith angle	Radiosonde launch time (UTC)*	Launch site
247	3 Sep 1992	1509	28.57	1352	Hamdallay
249	5 Sep 1992	1445	18.62	1032	Niamey
250	6 Sep 1992	1433	38.53	1528	Niamey
254	10 Sep 1992	1526	49.61	1455	Hamdallay
255	11 Sep 1992	1514	34.09	1032	Niamey
261	17 Sep 1992	1543	63.03	1454	Hamdallay
262	18 Sep 1992	1531	53.18	1459	Hamdallay
266	22 Sep 1992	1442	28.11	1435	Hamdallay
267	23 Sep 1992	1430	45.24	1502	Hamdallay
268	24 Sep 1992	1559	72.07	1029	Niamey
276	2 Oct 1992	1422	54.34	1031	Niamey
277	3 Oct 1992	0216	31.31	2230	Niamey
278	4 Oct 1992	0203	47.75	2227	Niamey
279	5 Oct 1992	1527	47.62	1030	Niamey
280	6 Oct 1992	1515	31.05	1045	Niamey
281	7 Oct 1992	0309	54.22	2228	Niamey

* Radiosondes close to 2200 UTC launched day before day indicated in first column.

surements for FIFE and HAPEX–Sahel, along with necessary ancillary measurements (radiosonde profiles, IRT temperatures, navigational identifiers), were extracted from the respective CD-ROM datasets created for each experiment's final data archives; see Strebel et al. (1992, 1994) and Kerr et al. (1993a,b). Full-resolution, cloud-free AVHRR, and ancillary measurements at Walpeup were provided by Dr. Fred Prata. Cloud screening for FIFE and HAPEX–Sahel AVHRR imagery was done using the method of Saunders and Kriebel (1988). Summaries of the AVHRR imagery are given in Tables 2a–c.

For three of the study areas (HAPEX–Sahel being the exception), radiosonde profiles were available at or near the time of the satellite overpasses (AVHRR) or sampling times (GOES). The Moderate-Resolution Transmittance Model and Code (MODTRAN) 3.5 radiative transfer model (Berk et al. 1989) was used with these soundings to calculate atmospheric correction terms. The spectral resolution of MODTRAN 3.5 is 1 cm^{-1}; the spectral radiances are integrated with a particular satellite's channel filter response functions to determine the correction terms. At ARM CART, radiosonde profiles were launched every 3 h beginning at 0200 UTC each day. Two-point linear interpolation of the correction terms calculated at the launch times is used to estimate correction terms at the satellite sampling times. As evident in the comparison between calculated and independently measured downwelling radiances for 28 September (as shown in Fig. 3), the interpolation procedure produces reliable estimates of correction terms between launches at ARM CART. The independent comparison measurements were obtained from the Atmospheric Emitted Radiance Interferometer (AERI), a precision and spectrally detailed upward-pointing instrument situated at the ARM-CART Central Facility site located near Enid, Oklahoma (see http://www.arm.gov/docs/instruments/static/aeri.html for a discussion of the AERI instrument). Summaries of radiosonde launch times relative to AVHRR overpass times are in Tables 2a–c.

Excepting Walpeup where surface contact thermometers were used, downfacing IRTs and pyrgeometers were used to estimate surface skin temperatures at the other three study areas. A radiometric temperature (T_r) measured by an IRT at FIFE or HAPEX–Sahel is converted to a skin temperature (T_s) representing LST according to

$$T_s = \{[\sigma T_r^4 - (1 - \epsilon_{BB})L\downarrow]/\epsilon_{BB}\sigma\}^{1/4}, \qquad (5)$$

where $L\downarrow$ is the measured broadband downwelling longwave flux (see Norman and Becker 1995), σ is the Stefan–Boltzmann constant, and ϵ_{BB} is the broadband emissivity. These measurements were obtained for FIFE from the averaged dataset of Betts and Ball (1998) based on the pyrgeometer measurements of Smith et al. (1992), and for HAPEX–Sahel from the CD-ROM archives based on pyrgeometer measurements made at the south-

TABLE 2c. Summary information on satellite and radiosonde data used for Walpeup.

Yearday	Date	Satellite overpass time (UTC)	Zenith angle	Radiosonde launch time (UTC)
228	16 Aug 1990	1610	4.0	1610
229	17 Aug 1990	1555	19.7	1555
230	18 Aug 1990	1550	35.3	1550

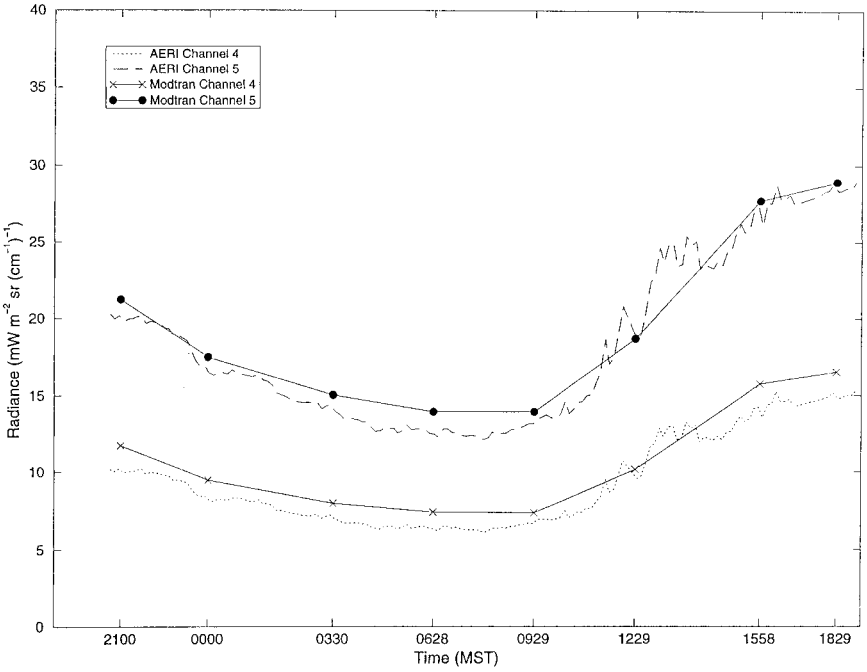

FIG. 3. Time series of downwelling radiances associated with *GOES-8* channels 4 and 5, calculated using MODTRAN 3.5 RTE model based on radiosonde profiles taken from ARM CART Central Facility for 27–28 Sep 1996, along with analogous band radiances measured by AERI instrument. Times of radiosonde launches are indicated with solid dots and crosses. MST = mean solar time.

ern and central-west supersites. Note that Cooper and Asrar (1989), Sugita and Brutsaert (1993), Kalluri and Dubayah (1995), and Schmugge and Schmidt (1998) all conducted intercomparisons of AVHRR-retrieved LSTs (based on one or more nonsimultaneous single-sample algorithms) with IRT validation measurements made during FIFE. Their results indicate retrieval uncertainties of ±3.0°, 3.5°, 2.8°–4.7°, and 5.7°C, respectively.

For ARM CART, the σT_r^4 term is replaced by the pyrgeometer-measured broadband upwelling flux $L\uparrow$, where pyrgeometer-measured $L\downarrow$ measurements also are provided. A graybody ϵ_{BB} value of 0.98 is used, as suggested by B. Blad of the University of Nebraska (1998, personal communication) for the grassland-type biomes of FIFE and ARM CART (equivalent to the graybody spectral value used by Prata 1994a). The broadband emissivities used for HAPEX–Sahel are given in Table 3. Since Walpeup used contact thermometers to obtain estimates of T_s, no assumptions on emissivity at that site are required.

TABLE 3. Emissivities of surface covers for 8–14-μm band.

Surface cover	Emissivity
Bare soil (fallow field)	0.914*
Sand–weed mixture	0.949*
Grass	0.976**
Tiger bush	0.984**

* van de Griend et al. (1991).
** Rubio et al. (1997).

5. Applications and validation of retrieval algorithm

The retrieval algorithm requires two sets of split-window images taken close in time so that the time-invariant emissivity assumption holds. The 65 *GOES-8* images taken on 28 September (yearday 272) 1996 each were paired with the same images made 24 h later on 29 September (yearday 273). The AVHRR images were paired if precipitation did not occur between overpasses. In order to maximize the number of retrievals, some of the FIFE and HAPEX–Sahel AVHRR images were used more than once (e.g., the FIFE 1031 UTC image on yearday 155 was paired with all other images taken between yeardays 155 and 157). There were 31 pairs of images for FIFE, 10 for HAPEX–Sahel, and 3 for Walpeup. The retrieval algorithm was not applied to any of the cloud-flagged pixels within the ARM CART, FIFE, and HAPEX–Sahel images. The Walpeup AVHRR images were cloud free and did not require cloud screening or precipitation checks.

Skin temperature and emissivity retrievals were made for the pixels located closest to the validation sites in the *GOES-8* and AVHRR images. In addition, retrievals for the pixels for the four FIFE sites 1563, 2123, 2139,

TABLE 4. Total number of temperature retrievals for the five study areas.

Study area	Number of retrievals
ARM CART	1994
FIFE	630
Konza	62
HAPEX–Sahel	98
Walpeup	6

and 3221 located within the Konza Prairie Reserve (a homogeneous grassland) were averaged to form a 2 km $\times$ 2 km superpixel for comparison to the four-site validation average (hereinafter referred to as the Konza site), analogous to the multiple in situ measurement scheme used at Walpeup. The total retrieval counts for each of the five single sites or site groups are given in Table 4.

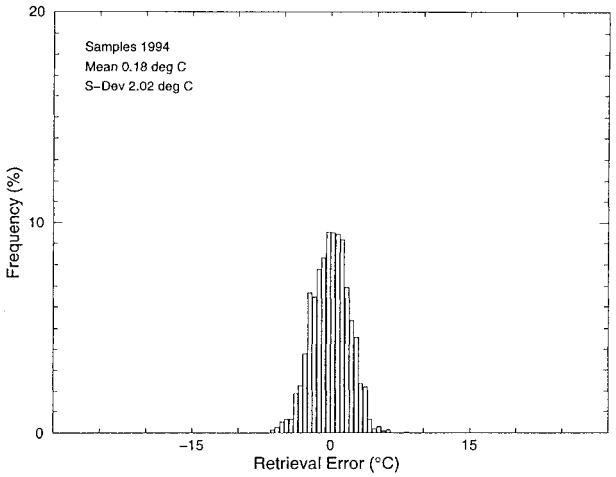

FIG. 4. Histogram of skin temperature retrieval errors (validation minus retrieval) at ARM CART site during 28–29 Sep 1996. Bin size of histogram is 0.5°C.

a. GOES-8 *retrievals*

A histogram of the ARM CART temperature retrieval errors (validation minus retrieval) is given in Fig. 4. The average retrieval error is 0.18°C with a bias-adjusted root mean square error (rmse) of 2.02°C. LST bias and rmse for the 16 individual validation sites are given in Table 5. Taken altogether, absolute values of the bias and rmse factors all are less than 2.0° and 4.5°C, respectively. Note that 14 of the 16 sites exhibit biases of less than 1.2°C, with 10 of 16 rmse's being less than 2.5°C.

A time series of the retrieved skin temperatures and emissivities for site EF-10 is given in Fig. 5. Note that this diagram effectively illustrates features that are common to all of the individual site validations. First, note that the retrieval algorithm captures the strong diurnal cycle in skin temperature. There is also high-frequency and seemingly random noise in the retrieved temperatures that is on the order of 1°–2°C. Inspection of the time series shows that there is some correlation between the temperature noise and variations in the retrieved emissivities, which is not surprising in simultaneous retrieval calculations. In addition, the retrieval algorithm produces generally negative errors (retrieval less than validation) during nighttime and early morning hours, with generally positive errors (retrieval greater than validation) during daytime.

The source of the retrieval noise is linked to spurious variations in the satellite EBBTs. To analyze the problem, a split-window retrieval algorithm developed by Prata (1993) of the form

$$T_s = a_0 + a_1 T_4 + a_2 T_5 \qquad (6)$$

is applied to the *GOES-8* EBBTs at each site. The three coefficients for this algorithm, defined in the appendix, are functions of the surface emissivities, atmospheric transmittances, and downwelling radiances. The emissivities are set to the value of 0.98 that is recommended by Prata (1994a) for a vegetated surface. Since the Prata algorithm uses a single sample set of EBBTs, its solution directly pinpoints the dependence of the retrieved temperatures on the satellite radiances. Figure 6 illustrates the comparison between the Prata algorithm LST retrievals and the channels 4 and 5 EBBTs for ARM CART site EF-10. The Prata LSTs exhibit a similar noise pattern to that produced by the simultaneous algorithm, although the magnitude of the noise is slightly smaller.

Inspection of the values shown in Fig. 6 shows that the retrieval noise is associated with variations in satellite EBBTs. One example can be seen at 1432 MST (mean solar time) on yearday 272 (marked ''1'' in Fig. 6). The retrieved temperatures are nearly steady state between 1315 and 1515 MST except for a 1.6°C drop in the retrieved LST at 1432 MST. At that same time

TABLE 5. Temperature biases and bias-adjusted rmse's for ARM CART validation sites.

Site	Bias (°C)	Bias-adjusted rmse (°C)
EF-01	0.8047	2.0684
EF-02	0.1730	1.7356
EF-03	−0.8273	2.6678
EF-05	−1.6276	3.5736
EF-06	0.4248	1.7526
EF-07	0.1798	1.8345
EF-08	−0.1100	2.3391
EF-09	0.4785	3.1133
EF-10	−0.4369	1.7760
EF-11	0.2845	2.2781
EF-13	0.9661	2.4877
EF-15	−0.3734	2.0959
EF-16	1.1630	2.9115
EF-18	−1.0140	2.8330
EF-22	−0.1665	1.7092
EF-24	1.9707	4.4949

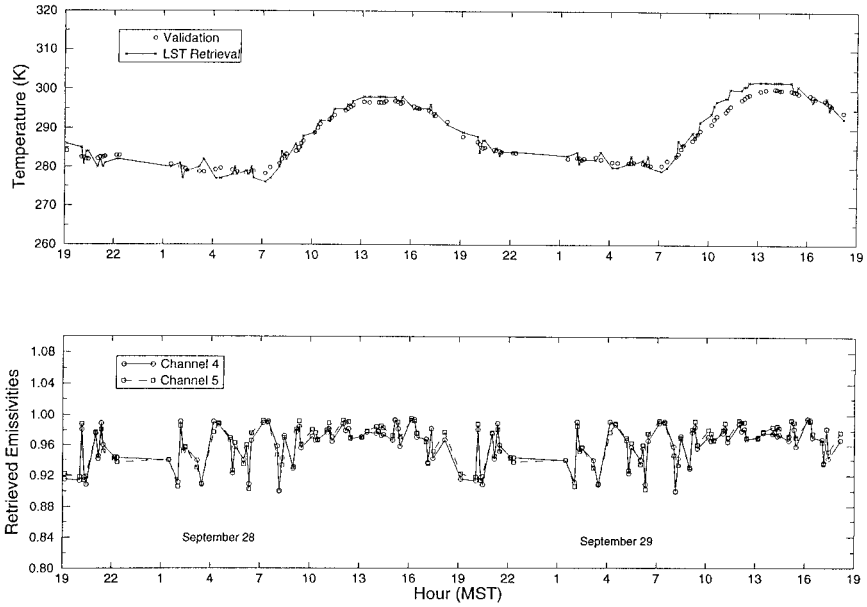

FIG. 5. Time series of retrieved and validation temperatures (top panel) and retrieved channel 4 and 5 emissivities (bottom panel) for ARM CART site EF-10 during 28–29 Sep 1996.

there is a drop in the channel 4 EBBT without a corresponding drop in the channel 5 EBBT. Five similar events can be seen between 0145 and 0700 MST on yearday 273. Last, two rise events occurring between 1400 and 1600 MST are associated with increases in channel 4 without concomitant increases in channel 5. As evident in Fig. 5, there are stepwise image-to-image variations in the retrieved LSTs based on the simultaneous algorithm, while the in situ measurements exhibit smoother behavior. During the two days, channel 4 and 5 EBBT differences exhibit oscillations in both sign and magnitude. This behavior suggests that the noise-equivalent temperature difference (NEΔT) of *GOES-8* channel 4 is larger than for channel 5.

The Prata single-sample algorithm produces less noise than the simultaneous multisample algorithm, because the error is compounded when pairing two sets of observations, each of which contains noise, in the cost function minimization.

Another feature common to the individual site retrievals is a semidiurnal oscillation in the retrieval errors. The average bias error at all 16 sites for each retrieval time for the two days is given in Fig. 7. During the nighttime period, the average errors are near zero. A short time after sunrise (marked "×"), the bias error exhibits a sinusoidal-type oscillation with the largest negative errors occurring during midmorning and the largest positive errors occurring during midafternoon. The errors again approach zero just after sunset (marked "+").

Although the source of this semidiurnal error cycle cannot be confirmed, the evidence suggests that it is caused by unaccounted for nonlinear changes in the channel weighting functions that occur between radiosonde launches. Since linear interpolation is used in conjunction with the bracketing soundings to generate atmospheric correction factors that represent conditions at precisely the time at which the satellite image was

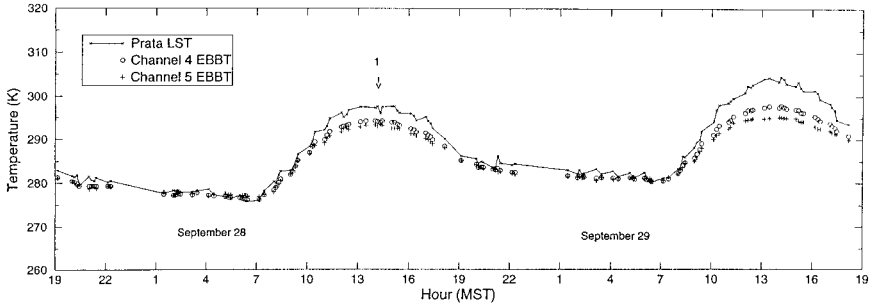

FIG. 6. Time series of skin temperature retrievals based on Prata (1993) algorithm and satellite EBBTs for ARM CART site EF-10 during 28–29 Sep 1996.

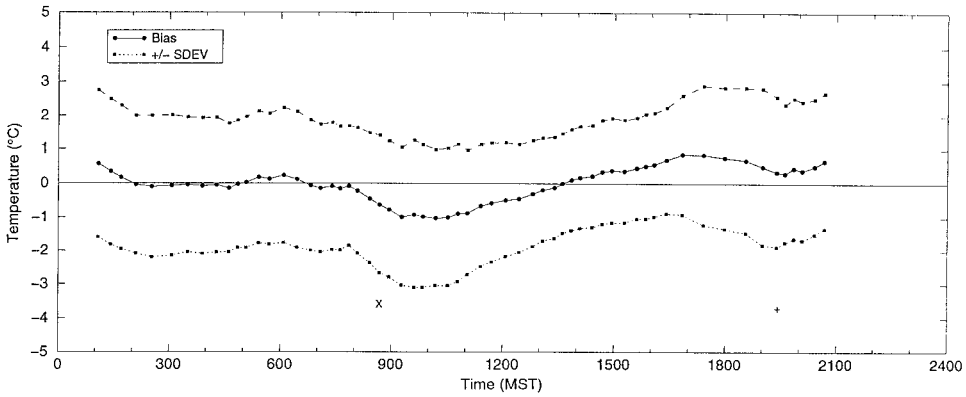

FIG. 7. Time series of average temperature retrieval errors (validation minus retrieval) and standard deviations for all ARM CART sites for each retrieval time during 28–29 Sep 1996. Mean solar time is used; ''×'' denotes time of sunrise, while ''+'' denotes time of sunset.

made, any nonlinear behavior in the temperature–moisture profile structures related to nonlinear development of the boundary layer cascades into error in representing the weighting functions. [Note that the weighting functions are defined in the standard manner, that is, as vertical derivatives of the atmospheric cumulative transmittance functions for each of the split-window spectral bands.]

Channel 4 weighting functions for the 16 ARM-CART soundings have been calculated using the MODTRAN 3.5 RTE model and are illustrated in Fig. 8. The weighting functions follow a diurnal cycle that is linked to the evolution of the boundary layer. Note in the left-hand panel of Fig. 8 that at 0630 MST the weighting function decreases rapidly from the surface to ~0.7 km within the surface inversion, with the inversion top indicated by the advent of the ''pause'' behavior extending from ~0.7 to ~0.95 km. This profile is indicative of the nighttime collapsed stable boundary layer. Three hours later (0930 MST) there has been little change other than erosion of the inversion strength. At 1230 MST there is a large increase in the weighting function values between ~0.45 and 0.8 km, a process continuing up through 1830 MST. This increase is related to increased moisture in the steadily deepening boundary layer and dissipation of the capping inversion. At 2130 MST (right-hand panel of Fig. 8) as the surface inversion reappears, the nighttime collapse is taking place, continuing on to 0330 MST when the weighting function profile is nearly equivalent to the prior day's 0630 MST profile.

The atmospheric correction terms assigned to the images taken between radiosonde launches are calculated using a linear interpolation of the correction terms found at the times of the radiosonde profiles. However, the atmospheric boundary layer evolves in a complex, nonlinear fashion, particularly during the daytime period when convective turbulence is strong. The use of a two-point interpolation scheme to generate representative correction factors in the presence of a nonlinearly evolving boundary layer can produce wavelike behavior in the daytime error cycle, mostly due to generation of error in establishing the boundary layer depth and vertically distributed moisture up to the boundary layer top.

The problem with linear interpolation is most evident for the retrievals done shortly after sunrise as the boundary layer makes a transition from a stable nocturnal stratification (capped by a near-surface inversion) to a convectively active and deepening layer controlled by surface fluxes, advection, and boundary layer top entrainment. A slightly smaller amplitude pattern is seen also near sunset as the boundary layer begins to collapse and switches from unstable to stable conditions. Thus, it is during the boundary layer transition times where one expects to find the greatest uncertainty in the retrievals, unless there are almost exactly matched temperature–moisture profiles to go with the satellite overpass times.

It also is noted that part of the daytime error may be related to problems with the validation measurements, which are based on downfacing Eppley pyrgeometers. There are always unresolved errors in pyrgeometer measurements under solar illumination because of solar-induced radiative heating of the pyrgeometer dome and a concomitant source of contaminating thermal radiation from the inner dome to the pyrgeometer's thermopile. Although a standard procedure has been followed for the ARM-CART pyrgeometer measurements (Dr. M. Wesely, Argonne National Laboratory, 1998, personal communication), which is designed to correct for dome heating according to a simple Fickian diffusion formula and is based on dome–sink thermistor measurements within the radiometer (see Smith et al. 1992), there is always a potential for residual error in the correction scheme that is diurnal in nature and is related directly to the diurnal modulation of downwelling solar flux (see Hodges and Smith 1997).

There is a final problem concerning uncertainty in broadband pyrgeometer validation measurements that is unwieldy: the moisture-induced absorption attenuation

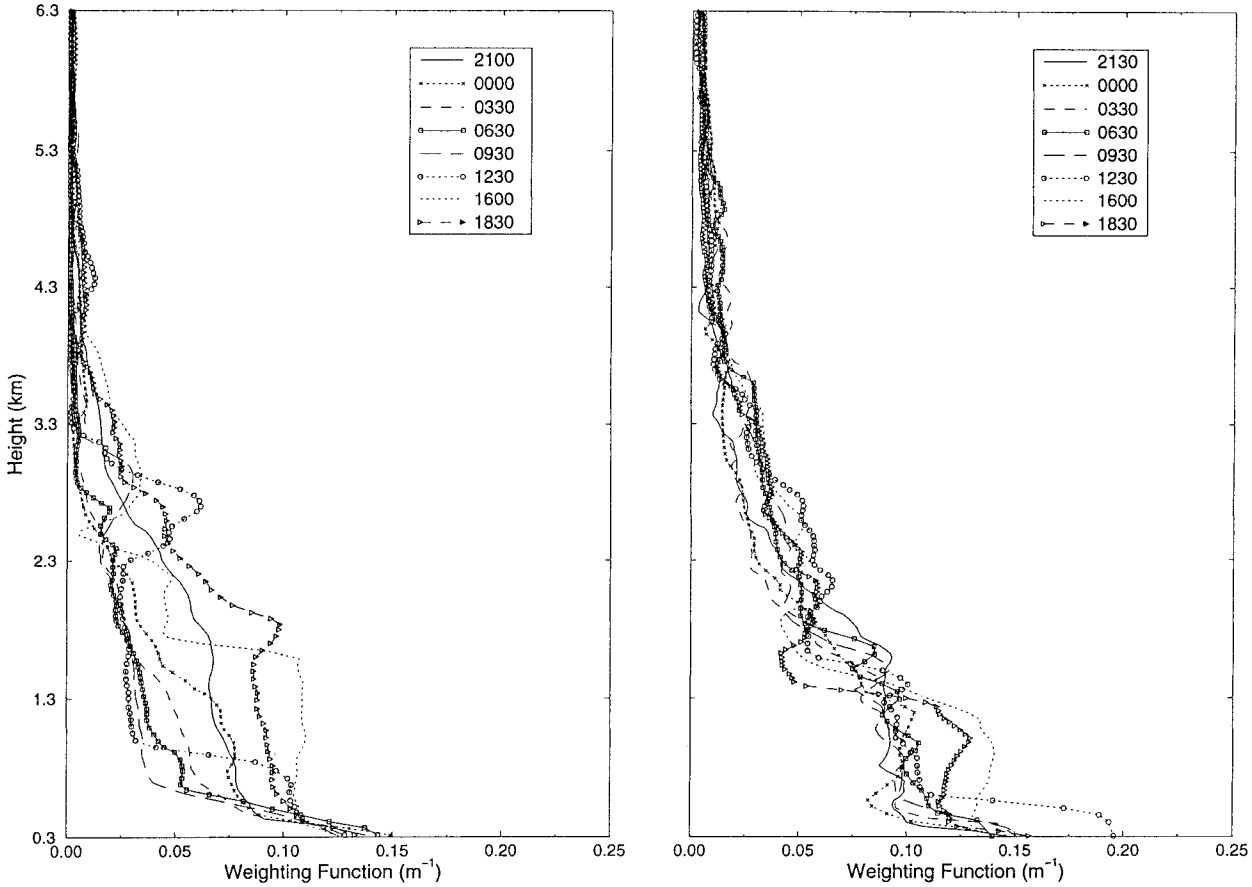

FIG. 8. *GOES-8* channel 4 weighting functions calculated from MODTRAN 3.5 RTE model based on radiosonde profiles obtained from ARM CART Central Facility during 27–28 (left panel) and 28–29 (right panel) Sep 1996. Radiosonde launch times given in key are in MST.

across the layer between the surface and the plane extending through the height of the radiometer. If there is no temperature jump across the surface–air interface, this term is negligible, on the order of -0.05°C for typical midlatitude summer conditions. However, if there is a skin temperature jump (typically positive in the daytime and negative at night), the attenuation increases significantly. For example, for typical summertime midlatitude conditions, a 6°C positive (negative) skin temperature jump leads to in situ attenuation errors of approximately $+1.3$°C (-1.3°C) for a pyrgeometer height of 10 m, and approximately $+1.9$°C (-1.9°C) for a pyrgeometer height of 50 m (ARM CART pyrgeometers are mounted at 10 m). Since daytime skin temperature jumps generally are more positive than nighttime jumps are negative, this attenuation effect will lead to greater positive biases in the validation intercomparisons if the validation sampling is equal over day and night, and serves to overestimate the ARM CART biases by approximately 0.5°–1.0°C.

b. AVHRR results

The six individual skin temperature and channel 4–5 emissivity retrievals for the Walpeup site are given in Table 6. Histograms of the retrieval errors for the FIFE, Konza, and HAPEX–Sahel sites are given in Fig. 9. Averaged LST biases and rmse's for all 21 of the in-

TABLE 6. Individual skin temperature and emissivity retrievals for Walpeup site. Date 1 and date 2 columns indicate dates and times of two overpasses used for retrievals. The T_{v1} and T_{r1} columns indicate in situ and retrieved temperatures for date 1; T_{v2} and T_{r2} columns indicate similar quantities for date 2. Also, ϵ_4 and ϵ_5 indicate associated channel 4 and 5 emissivities.

Date 1	Date 2	T_{v1}	T_{r1}	T_{v2}	T_{r2}	ϵ_4	ϵ_5
228 × 210	229 × 155	277.2	277.9	276.4	275.9	0.9714	0.9744
229 × 155	230 × 145	276.4	274.9	272.4	273.0	0.9880	0.9901
229 × 210	230 × 145	277.2	276.9	272.4	272.9	0.9883	0.9923

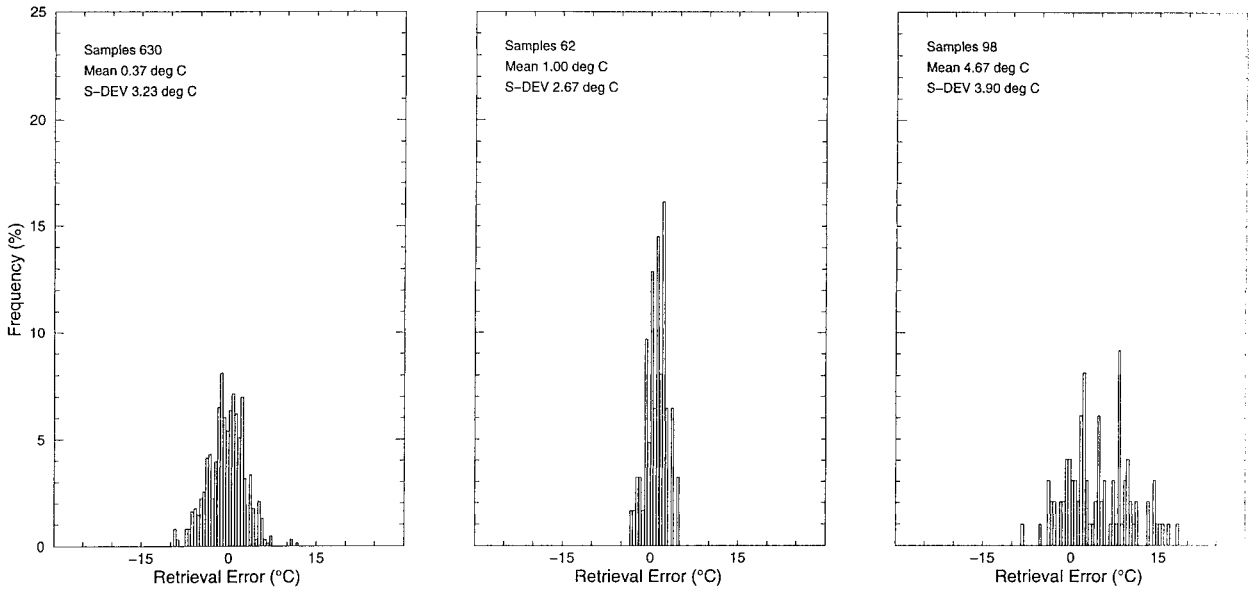

FIG. 9. Histograms of skin temperature retrieval errors (validation minus retrieval) at FIFE (left), Konza (center), and HAPEX–Sahel (right) sites. Bin size of histograms is 0.5°C. Also given are sample sizes, means, and standard deviations.

dividual AVHRR validation sites are given in Table 7. A graphical summary of all 36 individual site biases and rmse's is given in Fig. 10 (excepting Konza but including those for ARM CART). Walpeup had the lowest bias and rmse of the 21 AVHRR sites. The individual retrieval errors for Walpeup range between −0.7° and 1.5°C. These errors are close in magnitude to those reported by Prata (1994a) for several retrieval algorithms applied to the Walpeup AVHRR measurements. The Walpeup, FIFE, and Konza sites (with the exception of

TABLE 7. Biases and bias-adjusted rmse's for 21 individual sites within four AVHRR study areas.

Study area	Site	Bias (°C)	Bias-adjusted rmse (°C)
FIFE	0847	0.127	3.204
FIFE	1563	1.683	4.174
FIFE	2123	−1.153	3.488
FIFE	2139	−0.152	2.251
FIFE	2428	1.089	3.084
FIFE	3221	0.609	2.069
FIFE	4168	3.517	7.997
FIFE	4439	−1.315	3.858
FIFE	4609	−1.518	4.860
FIFE	6469	−3.115	−3.115
FIFE	6912	−1.706	4.437
FIFE	8639	0.428	2.482
Konza		1.001	2.670
HAPEX–Sahel	2319	4.912	10.541
HAPEX–Sahel	2424	5.532	12.678
HAPEX–Sahel	2924	−0.501	3.539
HAPEX–Sahel	5153	3.951	9.102
HAPEX–Sahel	5655	3.983	9.828
HAPEX–Sahel	6754	6.733	14.374
HAPEX–Sahel	6855	6.814	14.432
Walpeup		0.083	0.779

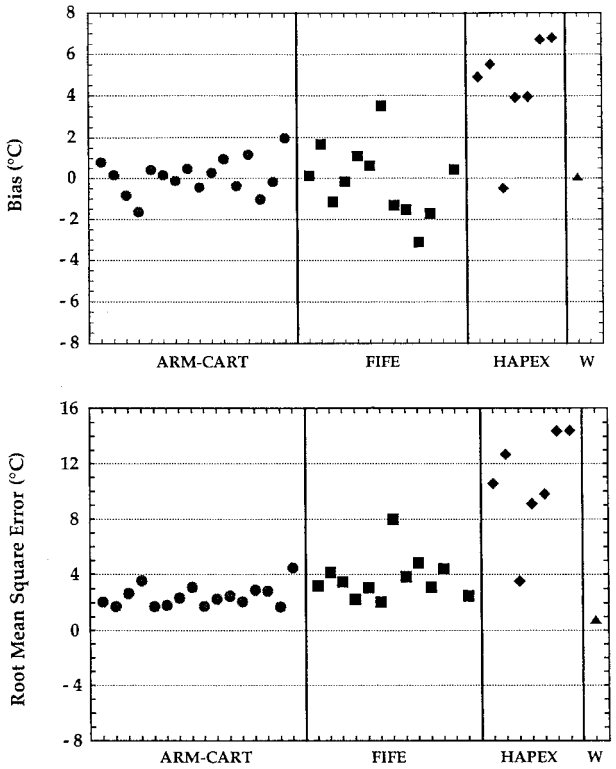

FIG. 10. Surface temperature retrieval bias errors (top panel) and bias-adjusted rmse's (bottom panel) at all 36 individual ARM CART, FIFE, HAPEX–Sahel, and Walpeup (W) validation sites.

FIFE sites 4168 and 6912) have absolute value biases and rmse's less than 1.7° and 5.0°C, respectively. These factors are similar in magnitude to those found for the site-level ARM CART validations.

On the other hand, the HAPEX–Sahel validation analysis produced biases and rmses that are much larger than those from the other AVHRR validation sites. There are a number of possible causes for the larger validation differences associated with HAPEX–Sahel, in addition to the basic problem that the data quality of the operational soundings is poor. To begin with, the surface covers at the scale of an AVHRR pixel are mixtures of vegetation and exposed soil (mostly sandy). A single IRT at four of the sites measured the temperature of only one component of the surface cover (typically vegetation). Andersen (1997) showed that the spatial variability of skin temperature at one of the IRT sites (0179, not used in this study) was ~20°C over 2 m. The variability is caused primarily by differences in heat budget processes between the canopy and soil components. Therefore, validation temperatures themselves are highly vulnerable to bias in representing sites when only one IRT is used.

The three HAPEX–Sahel sites at the U.K. southern supersite (2319, 2424, 2924) had multiple IRTs that measured the temperatures of the vegetation and soil components. Validation temperatures for these sites were made by averaging the component temperatures. However, the correct weighting factors based on the fractional amount of vegetation covers are not known, which means that representative skin temperatures at the scale of the satellite pixels cannot be obtained. This problem represents another source of bias in the ''truth'' data in seeking to validate the retrievals.

Atmospheric correction terms for 7 of 16 overpasses for HAPEX–Sahel are calculated from radiosonde profiles taken within an hour of overpass. The remaining terms are calculated using radiosonde profiles that are displaced in time by 4–5 h. As discussed previously, since the boundary layer undergoes significant evolution over that length of time, especially for daytime conditions, the correction terms calculated by linear interpolation cannot be expected to give accurate renditions of atmospheric conditions at the times of the overpasses. This failure leads to errors in the retrieved skin temperatures, since the cost function is contaminated by biased atmospheric correction terms, and thus is another source of increased uncertainty in the validation analysis.

To gauge the magnitude of this effect, a Monte Carlo experiment is devised based on the high frequency ARM CART soundings. A random selection of 100 pixels is used for a first set of retrievals based on evaluating correction terms at five *GOES-8* image times from 2115 MST 27 September to 1545 MST 28 September 1996 (images 2115, 0345, 0932, 1232, and 1545 MST), using interpolation to the image times based on the high-frequency profiles. A second set of retrievals then is made in which the interpolation of correction terms is made based only on the 1900, 0700, 1900 MST launches over the two days (i.e., at conventional synoptic times of 0000, 1200, 0000 UTC). Here the interpolations are based on 12-h intervals. Averaged differences (100 samples) at the five analysis times are made between the ''good'' interpolation case and the ''poor'' interpolation case. The bias/bias-adjusted rmse factors for the five image times are 3.38/1.90, 3.55/1.94, 1.14/2.53, −0.75/2.15, and 0.55/2.03°C. The largest bias plus rmse total error is approximately 5.5°C. Such an uncertainty is considerably larger than that associated with inaccurate specification of the differential emissivity, and points to the two main pitfalls in interpreting most previous studies on LST retrieval. These pitfalls are the characterization of error in atmospheric correction that stems from the timing of the sounding profiles relative to the satellite sampling time, and the vulnerability of a validation analysis conducted in only one locale with only one set of instrumentation.

c. Comparison of emissivity retrievals

Since direct measurements of surface emissivities are unavailable for any of the sites under study, only qualitative evaluations of the retrieved infrared emissivities can be carried out. Laboratory and field measurements of the emissivities of surface covers common to the retrieval sites in the 8–12.5- (Chen and Zhang 1989), 8–14- (van de Griend et al. 1991; Rubio et al. 1997), and 2–18-μm (Sutherland 1986) bands show a range of 0.90 to 0.99. The emissivities of bare soils tend to be near 0.91 (Sutherland 1986; van de Griend et al. 1991), while soils with some organic matter mixed in have higher emissivities (Rubio et al. 1997). Green vegetation with little or no exposed soil typically has an emissivity between 0.97 and 0.99 (Rubio et al. 1997), while the range for vegetation that is dry or has some exposed soil is 0.94 to 0.96 (van de Griend et al. 1991; Rubio et al. 1997). Rubio et al. (1997) found that differential emissivities (channel 4 minus channel 5) of most surface covers range between −0.001 and −0.014.

Table 6 shows that the retrieved emissivities for Walpeup ranged between 0.971 and 0.992, with means of 0.983 and 0.986 for channels 4 and 5, respectively. Platt and Prata (1993) found a mean emissivity of 0.984 for the Walpeup site (the mean was calculated using channel 4 and 5 emissivities) using an equation similar to Eq. (3). This calculation was made in conjunction with measured AVHRR radiances and surface temperatures, along with RTE-modeled atmospheric radiances and transmittances based on radiosondes close to the satellite overpasses. The average retrieved emissivities from this study along with those estimated by Platt and Prata (1993) lie within a range of values (0.981–0.986) reported by Chen and Zhang (1989) in the 8.0–12.5-μm band for a wheat crop in various stages of growth.

As shown in Fig. 11 and tabulated in Table 8, the

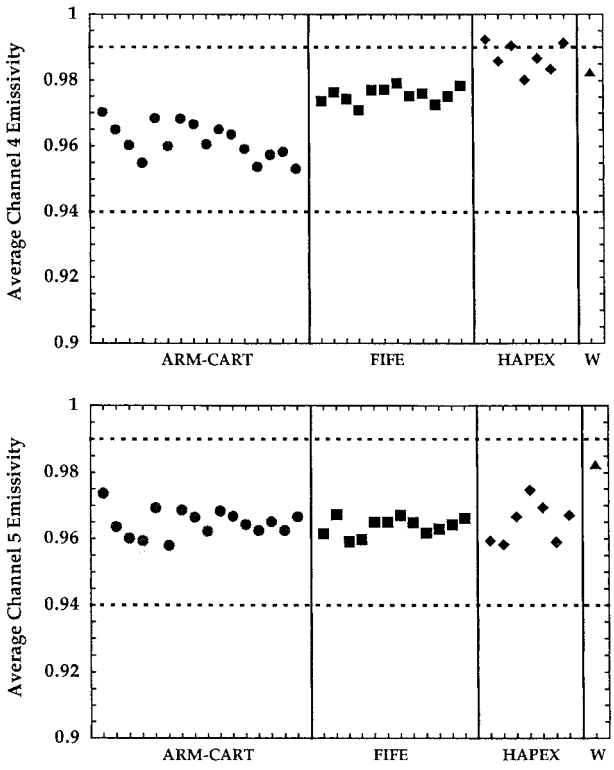

FIG. 11. Averaged retrieved channel 4 (top panel) and channel 5 (bottom panel) emissivities at all 36 individual ARM CART, FIFE, HAPEX–Sahel, and Walpeup (W) validation sites.

average retrieved channel emissivities for the ARM CART, FIFE, Konza, HAPEX–Sahel, and Walpeup sites range between 0.95 and 0.99. The ARM CART sites have slightly lower emissivities (0.95–0.97) than do FIFE and Konza (0.96–0.98). The differential emissivities are on the order of 0.01. Given the grassland nature of these sites, the retrieved emissivities are in the range expected for vegetated surface covers. Since the HAPEX–Sahel sites had exposed soil in the satellite's field of view, it might be expected that the channel 4 emissivities would be lower than 0.99 and closer to the channel 5 values of 0.95. However, results for HAPEX–Sahel are contaminated by sounding data quality problems—problems that affect the reliability of both the retrieved LSTs and the spectral emissivities.

The spreads of the retrieved emissivities in Table 8 are worth noting. A portion of the emissivity variance for the FIFE, Konza, and HAPEX–Sahel sites can be attributed to variations in surface composition since the retrievals are done over several months, while part of the variance for the ARM CART sites stems from the geographical distribution of the site network. In addition, changes in the satellite view angle will create differences in the projected surface, and will cause a change in the magnitude of projected exposed soil for low values of leaf area index. Both these factors change the effective emissivities between overpasses. Therefore, in obtaining representative emissivities based on instantaneous retrievals it is essential to merge statistically the individual realizations according to the expected mode of variability.

6. Conclusions

The LSTs generated by the simultaneous retrieval algorithm for ARM CART, FIFE, Konza, and Walpeup exhibit bias errors of 1.9°, 1.7°, 1.0°, and 0.08°C, respectively, with bias-adjusted rmse's of 4.5°, 4.8°, 2.7° and 0.78°C, respectively. The ARM CART bias estimate could be reduced to ~1.4°C by assuming a conservative estimate of bias error (~0.5°C) in the ARM CART validation measurements caused by unaccounted for moisture attenuation between surface and pyrgeometer. Most important, the average retrieved emissivities for these three sites are in the expected range of emissivities for vegetated surfaces (0.95–0.98).

The LST bias uncertainties are about half or better than the expected ±2°–3°C accuracy of LST algorithms based on current validation measurements (as suggested in a review by Pozo Vazquez et al. 1997), or the accuracy limits of 2.8°–5.7°C suggested in the FIFE studies of Cooper and Asrar (1989), Sugita and Brutsaert (1993), Kalluri and Dubayah (1995), and Schmugge and Schmidt (1998). The results are more promising when considering the small errors at the Walpeup site, which was instrumented specifically for satellite-based LST validation purposes. Notably, Prata et al. (1995) suggest that, with properly instrumented validation sites, biases of ±0.5°C are achievable. In any case, the retrieval biases are small enough to warrant exploring the use of GOES-retrieved LSTs as assimilation data in a biosphere model to improve surface flux predictability, particularly with regard to the troublesome soil moisture factors.

The retrieved skin temperatures for HAPEX–Sahel exhibit larger bias and rmse (5° and 11°C). The increases are attributed to heterogeneities around the measuring sites, the concomitant nonrepresentativeness of point measurements within a heterogeneous ecotone, highly noncoincident sampling of radiosonde launches with respect to satellite overpass times, and poor-quality operational sounding data. This intercomparison draws attention to the importance of validating algorithms at different locales using different instruments because of the susceptibility of any given validation site with its specific instrumentation to generate nonrepresentative difference errors—errors that are not attributable to the algorithm but to the validation data themselves.

There are additional issues concerning retrieval uncertainty that should be considered for any given application. The *GOES*-8 retrieved temperatures follow the expected diurnal behavior but contain a high-frequency noise component. The retrieved skin temperature noise and associated emissivity noise are correlated with noise in the satellite radiances in which the NEΔT

TABLE 8. Averages and standard deviations of retrieved channel (Ch) 4 and channel 5 emissivities for *GOES-8* and AVHRR validation sites.

Study area	Site	Avg emissivity (Ch 4)	Std dev emissivity (Ch 4)	Avg emissivity (Ch 5)	Std dev emissivity (Ch 5)	Avg difference (Ch 4 − Ch 5)	Std dev difference (Ch 4 − Ch 5)
ARM CART	EF-01	0.9703	0.0233	0.9738	0.0232	−0.0229	0.0063
ARM CART	EF-02	0.9650	0.0243	0.9636	0.0261	0.0014	0.0076
ARM CART	EF-03	0.9603	0.0240	0.9602	0.0247	0.0001	0.0075
ARM CART	EF-05	0.9555	0.0294	0.9595	0.0292	−0.0040	0.0054
ARM CART	EF-06	0.9685	0.0244	0.9695	0.0264	−0.0010	0.0091
ARM CART	EF-07	0.9600	0.0271	0.9580	0.0292	0.0020	0.0081
ARM CART	EF-08	0.9683	0.0189	0.9688	0.0212	−0.006	0.0080
ARM CART	EF-09	0.9666	0.0264	0.9666	0.0279	0.000	0.0073
ARM CART	EF-10	0.9606	0.0253	0.9624	0.0266	−0.0018	0.0062
ARM CART	EF-11	0.9651	0.0232	0.9685	0.0247	−0.033	0.0087
ARM CART	EF-13	0.9636	0.0263	0.9669	0.0289	−0.0033	0.0072
ARM CART	EF-15	0.9592	0.0270	0.9644	0.0286	−0.0051	0.0080
ARM CART	EF-16	0.9538	0.0250	0.9625	0.0262	−0.0086	0.0066
ARM CART	EF-18	0.9574	0.0303	0.9653	0.0301	−0.0080	0.0066
ARM CART	EF-22	0.9583	0.0279	0.9625	0.0283	−0.0043	0.0070
ARM CART	EF-24	0.9532	0.0274	0.9667	0.0304	−0.0134	0.0066
FIFE	0847	0.9737	0.0222	0.9615	0.0192	0.0122	0.0199
FIFE	1563	0.9764	0.0205	0.9673	0.0182	0.0182	0.0172
FIFE	2123	0.9743	0.0163	0.9592	0.0180	0.0150	0.0220
FIFE	2139	0.9709	0.0204	0.9598	0.0187	0.0110	0.0118
FIFE	2428	0.9771	0.0197	0.9651	0.0191	0.0121	0.0197
FIFE	3221	0.9772	0.0178	0.9651	0.0172	0.0121	0.0198
FIFE	4168	0.9791	0.0169	0.9671	0.0151	0.0119	0.0196
FIFE	4439	0.9753	0.0218	0.9650	0.0208	0.0103	0.0182
FIFE	4609	0.9761	0.0180	0.9617	0.0181	0.0143	0.0215
FIFE	6469	0.9727	0.0208	0.9629	0.0219	0.0098	0.0178
FIFE	6912	0.9753	0.0218	0.9643	0.0207	0.0110	0.0188
FIFE	8639	0.9785	0.0178	0.9662	0.01726	0.0122	0.0199
Konza		0.9762	0.0155	0.9638	0.1490	0.0122	0.0051
HAPEX–Sahel	2319	0.9923	0.0119	0.9594	0.0412	0.0328	0.0330
HAPEX–Sahel	2424	0.9859	0.0170	0.9583	0.0439	0.0276	0.0343
HAPEX–Sahel	2924	0.9905	0.0115	0.9667	0.0389	0.0239	0.0275
HAPEX–Sahel	5153	0.9803	0.0294	0.9747	0.0387	0.0056	0.0207
HAPEX–Sahel	5655	0.9867	0.0240	0.9694	0.0391	0.0173	0.0247
HAPEX–Sahel	6754	0.9834	0.0285	0.9590	0.0407	0.0244	0.0306
HAPEX–Sahel	6855	0.9913	0.0122	0.9671	0.0346	0.0243	0.0315
Walpeup		0.9825	0.0079	0.9856	0.0081	−0.0030	0.0076

of GOES channel 4 exceeds that of channel 5, confirmed by the GOES calibration engineer at the National Environmental Satellite, Data and Information Service (NESDIS) (Dr. M. Weinreb, NOAA/NESDIS, 1998, personal communication). Satellite noise will introduce into any LST retrieval algorithm error that can become compounded in a multitime/simultaneous algorithm framework. However, such noise can be suppressed with an appropriate filter once the noise properties are defined adequately.

The timing of the temperature–moisture profiles relative to the satellite sampling time is important for retrieval accuracy, as is evident from a comparison of the uncertainties between the ARM CART and HAPEX–Sahel retrievals, which lie on the two ends of the data quality/data sampling spectrum. As noted, the split-window channel weighting functions change nonlinearly throughout the day in response to the growth and decay of the boundary layer. A two-point interpolation scheme for obtaining atmospheric correction terms coincident with a satellite overpass time cannot always account reliably for such nonlinear evolution over a lengthy time interval. Thus the errors in atmospheric correction increase as the discrepancies between the launch times and the satellite overpass times increase, in addition to any problems associated only with sounding data quality. As shown in a Monte Carlo simulation experiment using the high quality ARM CART soundings, the use of radiosonde profiles for calculating atmospheric correction terms that are launched 4–5 h before or after a satellite overpass can contribute a degree of uncertainty in the retrievals that exceeds 5°C.

The impact of this type of error is most evident shortly after sunrise when the boundary layer goes from a nocturnal stable stratification to being influenced by unstable convective conditions. During these transitions, the estimated >5°C errors exceed that of the generally accepted source of major retrieval error in LST algorithms: incorrect specification of the spectrally distinct split-window channel emissivities. By going to a simulta-

neous skin temperature–spectral emissivity retrieval approach, the differential emissivity problem is mitigated. However, in the process, a perhaps more serious problem is revealed that likely has influenced most methods of LST retrieval for a long time, namely, the strong sensitivity to the space–time coincidence of the atmospheric temperature–moisture profile measurements used for atmospheric correction.

Acknowledgments. The authors thank Dr. Harry Cooper and Mr. James Merritt for their scientific and technical assistance and Dr. Fred Prata for providing his Walpeup data. ARM CART data were obtained from the Atmospheric Radiation Measurement Program sponsored by the U.S. Department of Energy, Office of Health and Environmental Research, Environmental Sciences Division. This research was supported by a graduate fellowship grant from the American Meteorological Society, NASA Global Change Fellowship Grant NGT5-30116, and NASA Research Grant NAG5-6258.

APPENDIX

Coefficients for Split-Window Retrieval Algorithm

The coefficient formulations for the Prata (1993) algorithm, following the form of Eq. (6), are given by

$$a_0 = -\Delta I\downarrow[\partial B_4(T_4)/\partial T]^{-1}(1 - \delta)/\delta + [1 - (a_1 + a_2)][T_4 - L_4],$$

$$a_1 = (1 + \gamma)/\delta, \quad \text{and} \quad a_2 = -\gamma/\delta,$$

where the factors on the right-hand sides are defined by

$$\gamma = (1 - \tau_4)/(\tau_4 - \tau_5),$$

$$\delta = \epsilon_4 + \gamma\tau_5(\epsilon_4 - \epsilon_5),$$

$$L_4 = B_4(T_4)/[\partial B_4(T_4)/\partial T], \quad \text{and}$$

$$\Delta I\downarrow = I_5\downarrow - I_4\downarrow,$$

and the following definitions are used:

$B_i(T) \equiv$ Planck function integrated over the satellite's filter response function for channel i,
$T_i \equiv$ satellite EBBT for channel i,
$\tau_i \equiv$ atmospheric transmittance integrated over the satellite's filter response function for channel i,
$\epsilon_i \equiv$ surface emissivity for channel i, and
$I_i\downarrow \equiv$ atmospheric emitted radiance integrated over the satellite's filter response function for channel i.

REFERENCES

Becker, F., 1987: The impact of spectral emissivity on the measurement of land surface temperature from a satellite. *Int. J. Remote Sens.,* **8,** 1509–1522.

——, and Z.-L. Li, 1990: Towards a local split-window method over land surfaces. *Int. J. Remote Sens.,* **11,** 369–393.

Berk, A., L. S. Bernstein, and D.C. Robertson, 1989: MODTRAN: A moderate resolution model for LOWTRAN 7. Tech. Rep. GL-TR-89-0122. [Available from Spectral Sciences Inc., 99 South Bedford Street #7, Burlington, MA 01803.]

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

Caselles, V., C. Coll, and E. Valor, 1997: Land surface emissivity and temperature determination in the whole HAPEX–Sahel area from AVHRR data. *Int. J. Remote Sens.,* **18,** 1009–1027.

Chen, J. M., and R. H. Zhang, 1989: Studies on the measurements of crop emissivity and sky temperature. *Agric. For. Meteor.,* **49,** 23–34.

Colello, G. D., C. Grivet, P. J. Sellers, and J. A. Berry, 1998: Modeling of energy, water, and CO_2 flux in a temperate grassland ecosystem with SiB2: May–October 1987. *J. Atmos. Sci.,* **55,** 1141–1169.

Coll, C., V. Caselles, J. A. Sobrino, and E. Valor, 1994: On the atmospheric dependence of the split-window equation for land surface temperature. *Int. J. Remote Sens.,* **15,** 105–122.

Cooper, D. I., and G. Asrar, 1989: Evaluating atmospheric correction models for retrieving surface tmperatures from the AVHRR over a tall grass prairie. *Remote Sens. Environ.,* **27,** 93–102.

Crosson, W. L., H. J. Cooper, and E. A. Smith, 1993: Estimation of surface heat and moisture fluxes over a prairie grassland. Part 4: Impact of satellite remote sensing of slow canopy variables on accuracy of a hybrid biosphere model. *J. Geophys. Res.,* **98,** 4979–4999.

Faysash, D. A., 1998: Simultaneous land surface temperature and spectral emissivity retrieval in the infrared split window. M.S. thesis, Dept. of Meteorology, The Florida State University, 87 pp.

Goutorbe, J. P., and Coauthors, 1997. An overview of HAPEX–Sahel: A study in climate and desertification. *J. Hydrol.,* **188–189,** 4–17.

Hodges, G. B., and E. A. Smith, 1997: Intercalibration, objective analysis, intercomparison, and synthesis of BOREAS surface net radiation measurements. *J. Geophys. Res.,* **102,** 28 885–28 900.

Kalluri, S. N. V., and R. O. Dubayah, 1995: Comparison of atmospheric correction models for thermal bands of the Advanced Very High Resolution Radiometer over FIFE. *J. Geophys. Res.,* **100,** 25 411–25 418.

Kerr, Y., T. Valero, and S. Wagner, 1993a: HAPEX–Sahel Information System, CD-ROM 1 and 2: AVHRR 92. CESBIO, BPI 2801. [Available from 18 Avenue E. Belin, 31055 Toulouse Cedex, France.]

——, ——, and ——, 1993b: HAPEX–Sahel Information System, CD ROM 3: Ground data Vol. 1 (12/94). CESBIO, BPI 2801. [Available from 18 Avenue E. Belin, 31055 Toulouse Cedex, France.]

Li, Z.-L., and F. Becker, 1993: Feasibility of land surface temperature and emissivity determination from AVHRR data. *Remote Sens. Environ.,* **43,** 67–85.

Norman, J. M. and F. Becker, 1995: Terminology in thermal infrared remote sensing of natural surfaces. *Remote Sens. Rev.,* **12,** 159–173.

Peppler, R. A., P. J. Lamb, and D. L. Sisterson, 1996: Site scientific mission plan for the southern Great Plains CART site: January–June 1996. National Technical Information Services Report, 86 pp. [Available from U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22161.]

Platt, C. M. R., and A. J. Prata, 1993: Nocturnal effects in the retrieval of land surface temperatures from satellite measurements. *Remote Sens. Environ.,* **45,** 127–136.

Pozo Vazquez, D., F. J. Olmo Reyes, and L. Alados Arboledas, 1993: A comparative study of algorithms for estimating land surface temperature from AVHRR data. *Remote Sens. Environ.,* **62,** 215–222.

Prata, A. J., 1993: Land surface temperatures derived from the Ad-

vanced Very High Resolution Radiometer and the Along-Track Scanning Radiometer: 1. Theory. *J. Geophys. Res.,* **98,** 16 689–16 702.

——, 1994a: Land surface temperatures derived from the Advanced Very High Resolution Radiometer and the Along-Track Scanning Radiometer: 2. Experimental results and validation of AVHRR algorithms. *J. Geophys. Res.,* **99,** 13 025–13 058.

——, 1994b: Validation data for land surface temperature determination from satellites. CSIRO, Division of Atmospheric Research Tech. Paper 33, 36 pp.

——, J. Caselles, J. Coll, C. Ottle, and J. Sobrino, 1995: Thermal remote sensing of the land surface from satellite: Current status and future prospects. *Remote Sens. Rev.,* **12,** 175–224.

Price, J. C., 1984: Land surface temperature measurements from the split window channels of the *NOAA 7* Advanced Very High Resolution Radiometer. *J. Geophys. Res.,* **89,** 7231–7237.

Prince, S. D., and Coauthors, 1995: Geographical, biological, and remote sensing aspects of the Hydrologic Atmospheric Pilot Experiment in the Sahel (HAPEX–Sahel). *Remote Sens. Environ.,* **51,** 215–234.

Rubio, E., V. Caselles, and C. Badenas, 1997: Emissivity measurements of several soils and vegetation types in the 8–14 μm wave band: Analysis of two field methods. *Remote Sens. Environ.,* **59,** 490–521.

Saunders, R. W., and K. T. Kriebel, 1988: An improved method for detecting clear sky and cloudy radiances from AVHRR data. *Int. J. Remote Sens.,* **9,** 123–150.

Schlick, T., and A. Fogelson, 1992: TNPACK—A truncated Newton minimization package for large-scale problems. I: Algorithm and usage. *ACM Trans. Math. Software,* **18,** 46–70.

Schmugge, T. J., and G. M. Schmidt, 1998: Surface temperature observations from AVHRR in FIFE. *J. Atmos. Sci.,* **55,** 1239–1246.

Sellers, P. J., F. G. Hall, G. Asrar, D. E. Strebel, and R. E. Murphy, 1992: An overview of the First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment (FIFE). *J. Geophys. Res.,* **97,** 18 345–18 371.

Smith, E. A., W. L. Crosson, and B. D. Tanner, 1992: Estimation of surface heat and moisture fluxes over a prairie grassland. Part 1: In situ energy budget measurements incorporating a cooled mirror dew point hygrometer. *J. Geophys. Res.,* **97,** 18 557–18 582.

——, ——, H. J. Cooper, and H. Y. Weng, 1993: Estimation of surface heat and moisture fluxes over a prairie grassland. Part 3: Design of a hybrid physical/remote sensing biosphere model. *J. Geophys. Res.,* **98,** 4951–4978.

Sobrino, J. A., C. Coll, and V. Caselles, 1991: Atmospheric correction for land surface temperatures using *NOAA-11* AVHRR channels 4 and 5. *Remote Sens. Environ.,* **38,** 19–34.

Strebel, D. E., D. R. Landis, J. A. Newcomer, D. van Elburg-Obler, B. W. Meeson, and P. A. Agbu, 1992: Satellite imagery, 1987–1989. *Collected Data of the First ISLSCP Field Experiment.* Vol. 2, NASA, CD-ROM.

——, ——, K. F. Huemmrich, and B. W. Meeson, 1994: Surface observations and non-image data sets. *Collected Data of the First ISLSCP Field Experiment.* Vol. 1, NASA, CD-ROM.

Sugita, M., and W. Brutsaert, 1993: Comparison of land surface temperatures derived from satellite observations with ground truth during FIFE. *Int. J. Remote Sens.,* **14,** 1659–1676.

Sutherland, R. A., 1986: Broadband and spectral emissivities (2–18 μm) of some natural soils and vegetation. *J. Atmos. Oceanic Technol.,* **3,** 199–202.

Taylor, S. E., 1979: Measured emissivity of soils in the southeast United States. *Remote Sens. Environ.,* **8,** 359–364.

van de Griend, A. A., M. Owe, M. Groen, and M. P. Stoll, 1991: Measurement and spatial variation of thermal infrared surface emissivity in a savanna environment. *Water Resour. Res.,* **3,** 371–379.

Watson, K., 1992: Two-temperature method for measuring emissivity. *Remote Sens. Environ.,* **42,** 117–121.

Weinreb, M. P., M. Jamieson, N. Fulton, Y. Chen, J. X. Johnson, J. Bremer, C. Smith, and J. Baucom, 1997: Operational calibration of *Geostationary Operational Satellite-8* and *-9* imagers and sounders. *Appl. Optics,* **36,** 6895–6904.

Xiang, X., and E. A. Smith, 1997: Feasibility of simultaneous surface temperature–emissivity retrieval using SSM/I measurements from HAPEX–SAHEL. *J. Hydrol.,* **188–189**, 330–360.