Long-term accuracy assessment of land surface temperatures derived from the Advanced Along-Track Scanning Radiometer

César Coll*, Enric Valor, Joan M. Galve, Maria Mira, Mar Bisquert, Vicente García-Santos, Eduardo Caselles and Vicente Caselles

Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia. 50, Dr. Moliner. 46100 Burjassot, SPAIN.

* Corresponding author. Email: cesar.coll@uv.es.

ABSTRACT

The accuracy of land surface temperatures (LSTs) derived from the Advanced Along-Track Scanning Radiometer (AATSR) was assessed in a test site in Valencia, Spain from 2002-2008. AATSR LSTs were directly compared with concurrent ground measurements over homogeneous, full-vegetated rice fields in the conventional temperature-based (T-based) method. We also applied the new radiance-based (R-based) method over bare soil and water surfaces, where ground LST measurements were not available. In the R-based method, ground LSTs are simulated from AATSR brightness temperatures in the 11 μm band and radiative transfer simulations using surface emissivity data and atmospheric water vapor and temperature profiles. The accuracy of the R-based ground LSTs depends on how well the profiles used in simulations represent the actual atmosphere at the time of AATSR observations. This can be checked with the difference $\delta(T_{11}-T_{12})$ between the actual AATSR and the profile-based simulated difference in the 11 and 12 μm brightness temperatures (T_{11} and T_{12}, respectively). We found that for -0.6 K<$\delta(T_{11}-T_{12})$<0.6 K, the R-based LSTs were accurate within ±1.0 K and can be used for LST validation. For the data analyzed here, the

AATSR operational algorithm overestimated the ground LST by 2 to 5 K, showing that the auxiliary data utilized within the retrieval scheme (biome classification and fractional vegetation cover maps at 0.5º×0.5º resolution) should be improved and provided at the same spatial resolution as the AATSR data (1 km^2). When the AATSR algorithm was optimized with biome and fractional vegetation cover selected according to the nature of each surface, LST errors showed negligible average biases and rmse=±0.5 K for full vegetation and water, and ±1.1 K for bare soil. Furthermore, we checked an alternative algorithm explicitly dependent on emissivity, which provided accurate LSTs for all the surfaces studied, with small biases, rmse from ±0.4 to ±0.6 K and most LST errors within ±1.0 K. The algorithm requires monthly emissivity maps at 1 km^2, which can be derived from classification and fractional vegetation cover estimated from optical AATSR data. The results of this paper show the high LST accuracy achievable with AATSR data in ideal conditions. While it is necessary to establish and maintain highly homogeneous T-based validation sites, the R-based method provides an alternative for the semi-operational, long-term evaluation of LST products at global scale, since it is applicable over surfaces with varied LST and atmospheric regimes where ground LST measurements are not feasible.

1. INTRODUCTION

Land Surface Temperature (LST) is a key parameter in meteorological, climatological and hydrological studies since it results from the physical interactions in the surface-atmosphere system, including the energy and water fluxes (Anderson et al., 1997; Sánchez et al., 2008). LST is sensitive to local atmospheric conditions, land cover type, soil moisture and vegetation water stress. Therefore it can be used to monitor desertification, deforestation and climate change (Allen et al., 1994; Lambin and Ehrlich, 1997). Thermal infrared (TIR) remote sensing is probably the most suitable technique to obtain LST measurements at regional and global scales. Currently, there are a number of LST products derived from TIR remote sensing observations at different spatial scales and temporal periodicities. Examples are the products generated from the Moderate Resolution Imaging Spectroradiometer (MODIS) (Wan, 2008), the Atmospheric Infrared Sounder (AIRS) (Susskind et al., 2003), the Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER) (Gillespie et al., 1998) and the Meteosat Second Generation (MSG) Spinning Enhanced Visible and Infrared Imager (SEVIRI) (PUM_LST, 2008).

The Advanced Along-Track Scanning Radiometer (AATSR) (Llewellyn-Jones et al., 2001) onboard the European Space Agency (ESA) satellite Envisat provides an operational LST product at 1 km^2 resolution, which is currently included in the Level 2 AATSR data (ATS_NR_2P). The operational LST algorithm is also applied to predecessors ATSR-1 and ATSR-2, then providing LSTs back to 1991. The AATSR has two on-board calibration targets, low-noise detectors and mechanical coolers that provide high radiometric accuracy and stability to the TIR data (better than 0.05 K for the 11 and 12 μm bands). Therefore instrumental error in AATSR LST estimation will be very small. The conical scanning mechanism of AATSR gives a dual-view of the Earth's surface, first in the forward view at zenith angles about 55º, and 150 s later in the nadir view at zenith angles from 0º to 21.7º. The nominal spatial resolution of AATSR is 1 km × 1 km in the nadir view and 1.5 km × 2 km in the forward view, with a swath width of about 500 km. Equator crossing time is 10:00 a.m. local time (descending node) and revisit time is about 3 days.

The operational AATSR LST product is based on the split-window technique using the 11 and 12 μm bands at nadir view (Prata, 2002a). Different sets of split-window coefficients are used for 14 land cover classes or biomes predefined in a static classification map, and tuned with monthly fractional vegetation cover maps based on climatology. Both classification and fractional vegetation cover maps are implemented in the operational algorithm at spatial resolution of 0.5º×0.5º in longitude and latitude. The generation of LST products involves the correction of the satellite-observed radiances for atmospheric absorption and emission (mainly due to water vapor) and non-unity of land surface emissivity. This is a challenging problem because of the water vapor and surface emissivity variability. It is therefore necessary to assess the accuracy and precision of the products to provide LST users with data quality information. Long-term validation is required to identify possible deficiencies and subsequently introduce improvements in the algorithms.

Since the launch of Envisat in 2002, the AATSR LST product has been validated with concurrent ground LST measurements performed over homogeneous surfaces such as rice fields with full vegetation cover in Valencia, Spain and inland waters in Lake Tahoe, USA (Coll et al., 2005, 2006, and 2009a). For both sites, results showed that the product overestimated the in situ LSTs by 3 to 4 K. The large LST errors were attributed to the spatial resolution of the classification and fractional vegetation cover maps (0.5º), which is too coarse to resolve the land surface heterogeneity at the AATSR 1 km^2 scale. Noyes et al. (2007a) observed both positive and negative biases, with AATSR being typically warmer (colder) in the summer (winter), which was attributed to the algorithm's sensitivity to atmospheric water vapor, temperature and LST. In order to address the above problems, Noyes et al. (2007b) proposed several modifications in the current operational algorithm.

The objective of this paper is to show the long-term accuracy assessment of AATSR derived LSTs at the Valencia test site in the period 2002-2008. We checked not only the operational LST product but also the impact of using more appropriate values for the auxiliary parameters over the test site, and the performance of an alternative split-window algorithm with explicit emissivity dependence applicable to AATSR (Galve et al., 2008). The overall aim of such validations is to stress the need for modifications in the current AATSR LST processor.

Two types of validation methods were used in the present study. First, we show conventional temperature-based (T-based) validation, where concurrent ground LST measurements were directly compared with AATSR retrievals over homogeneous rice crops with full vegetation cover, much in the same way as in Coll et al. (2005, 2006 and 2009a). Due to the dissimilarity in the spatial scale between ground instruments and satellite measurements, only highly homogeneous surfaces are suitable for T-based validation, which reduces the number of biomes and climatic conditions available for a meaningful LST assessment. Therefore, we propose the new radiance-based (R-based) method (Wan and Li, 2008) that has been used for MODIS LST validation (Wan, 2008; Coll et al 2009b) and does not require ground LST measurements. Instead, ground LSTs are calculated from satellite brightness temperatures through radiative transfer simulations using surface emissivity data, and atmospheric temperature and water vapor profiles. Since homogeneous sites in 10-13 μm emissivity are more frequent than homogeneous sites in LST, the R-based method can be potentially applied to a larger number of sites with different LST and atmospheric regimes. The R-based method is not a validation method in strict sense, because it does not rely on independently measured LSTs. However, it can be an alternative or complement to the T-based method when ground LST measurements are not feasible.

The paper follows with a brief description of the operational AATSR LST algorithm and the alternative explicit emissivity-dependent algorithm. In section 3, the T-based validation results for full vegetation cover are presented and discussed. Section 4 shows the principles and sensitivity analysis of the R-based method, and the validation results for two different biomes (bare soil and lake). Finally, the main conclusions of the paper are presented in section 5.

2. LST RETRIEVAL FROM AATSR

LST can be retrieved from AATSR brightness temperatures in the 11 and 12 μm bands, nadir view using the split-window technique. The AATSR forward view is not used due to difficulties in accounting for angular effects on temperature and emissivity, the different field of view and the non-simultaneity of the two views, which may have an important impact over heterogeneous land surfaces (Prata, 2002a; Coll et al., 2006). The application of split-window algorithms requires that the characteristics of the surface must be well known. It can be done through explicit dependence on surface emissivity data at the two bands considered (Becker and Li, 1990; Wan and Dozier, 1996; Coll and Caselles, 1997). Another possibility is to derive specific sets of algorithm coefficients for different land cover biomes weighted by fractional vegetation cover (Kerr et al., 1992). The operational AATSR LST algorithm (Prata, 2002a) adopted the latter approach. Therefore it does not depend explicitly on emissivity, but accounts implicitly for surface emissivity effects through biome-dependent coefficients and fractional vegetation cover. The operational algorithm is described next and an alternative algorithm with explicit emissivity dependence is presented in section 2.2.

2.1. Operational AATSR LST algorithm

The algorithm expresses the LST (T) as a nearly-linear combination of the brightness temperatures in the 11 and 12 μm bands, T_{11} and T_{12}, with coefficients determined by regression over simulated brightness temperatures and depending on the biome (i), the fractional vegetation cover (f), the precipitable water (W) and the satellite zenith viewing angle (θ):

$$T = a_{f,i,W} + b_{f,i}(T_{11}-T_{12})^{n} + (b_{f,i} + c_{f,i})T_{12} \quad (1)$$

$$a_{f,i,W} = 0.4[\sec(\theta)-1]W + f\, a_{v,i} + (1-f)\, a_{s,i}$$

$$b_{f,i} = f\, b_{v,i} + (1-f)\, b_{s,i}$$

$$c_{f,i} = f\, c_{v,i} + (1-f)\, c_{s,i}$$

$$n = 1/\cos(\theta/5)$$

Coefficients a, b, and c of Eq. (1) were determined for 13 land biomes defined in Dorman and Sellers (1989) plus a lake class (i=1 to 14), and are shown in Table 1 (Prata, 2002b). For each land cover biome, two separate sets of coefficients are specified for the fully vegetated surface (subscript v) and for the bare surface (subscript s), which are weighted by the fractional vegetation cover f. For some biomes (i=4, 7, 9, 10, 12, 13, and 14), coefficients are identical for the vegetated and bare surface, which makes the algorithm insensitive to f. For the lake class, different coefficients are given for day and night. In Eq. (1), all temperatures are in ºC.

Since the algorithm only uses the nadir view (θ<21.7º), the impact of W on the LST calculated from Eq. (1) is rather small, e.g., a LST difference <0.03 K for a W variation of 1 cm. For the same reason, coefficient n is close to unity and the algorithm is nearly linear on T_{11}-T_{12}. Given that $1<n<1.0029$, the difference with the LST derived using n=1 in Eq. (1) is between 0 and 0.03 K for T_{11}-T_{12}=3 K. Similarly, the impact of θ in the algorithm, both through coefficients $a_{f,i,W}$ and n, is small. Taking θ=0º instead of the correct angle yields differences in the calculated LST less than 0.06 K for W=3 cm.

The ancillary data necessary for the application of the algorithm (i, f and W) are tabulated at spatial resolution of 0.5º×0.5º in longitude/latitude. Land biomes are determined from global classification and are static. The fractional vegetation cover has monthly variability, and is obtained from Dorman and Sellers (1989) and an estimate of the "greenness" derived from global normalized difference vegetation index from the International Satellite Land-Surface Climatology Project (ISLSCP). Precipitable water data are taken from the NASA Water Vapor Project (NVAP) global climatology at monthly intervals. The target accuracy of the LST product is ±1.0 K at nighttime and ±2.5 K at daytime. For more details on the algorithm and the operational implementation see Prata (2002a).

As pointed out in previous validation studies (Coll et al., 2005, 2006, and 2009a), the 0.5º resolution of the ancillary data is too coarse to account for the natural heterogeneity of LST at the AATSR scale. Noyes et al. (2007b) and Coll et al. (2009a) also noted the high sensitivity of the LST retrievals to the fractional vegetation cover. Noyes et al. (2007b) proposed to increase the resolution of the biome and f maps to 1-3 km^2, to use the GLOBCOVER biome map derived from Envisat/Medium Resolution Imaging Spectrometer (MERIS), and to increase the accuracy of the fractional cover by introducing an observational component, e.g., estimating f by using vegetation indices derived from the AATSR optical bands.

2.2. Explicit emissivity-dependent LST algorithm

As an alternative, the split-window method can be applied with algorithms explicitly dependent on the surface emissivity at the 11 and 12 μm AATSR bands, namely the mean emissivity ($\varepsilon=(\varepsilon_{11}+\varepsilon_{12})/2$) and the band emissivity difference ($\Delta\varepsilon=\varepsilon_{11}-\varepsilon_{12}$), which are the physical magnitudes accounting for the effect of land surface biome and fractional vegetation cover on TIR measurements. This is the approach adopted by the generalized split-window algorithm (Wan and Dozier, 1996) used to generate the MODIS LST products MOD11_L2 (Terra) and MYD11_L2 (Aqua) (Wan, 2008), and by the MSG-SEVIRI LST product (PUM_LST, 2008). This approach is feasible since the emissivity values of natural surfaces at 10-13 μm show relatively small variability at global scale. According to Pinheiro et al. (2004), ε and $\Delta\varepsilon$ range from 0.968 to 0.990 and from -0.014 to 0.009, respectively, for the Advanced Very High Resolution Radiometer (AVHRR) bands 4 and 5, which are comparable to the AATSR bands. Similarly, Snyder et al. (1998) found ε varying from 0.969 to 0.990 and $\Delta\varepsilon$ from -0.006 to 0.011 for the MODIS split-window bands 31 and 32. However, the sensitivity of LST to these small emissivity changes can be large (typically ±0.8 K for uncertainty of ±0.010 in both ε_{11} and ε_{12}; Galve et al., 2008).

The explicit emissivity-dependent split-window algorithm proposed here is based on the model of Coll and Caselles (1997) particularized for AATSR as described in Galve et al. (2008). It can be written as

$$T = T_{11} + 0.02 + 0.782(T_{11}-T_{12}) + 0.302(T_{11}-T_{12})^2 + \\ + (1-\varepsilon)[53+1.13(W/\cos\theta)-1.023(W/\cos\theta)^2] - \Delta\varepsilon[79-11.06(W/\cos\theta)] \quad (2)$$

The algorithm coefficients were obtained from radiative transfer simulations of AATSR brightness temperatures using a database of 382 continental, cloud-free radiosonde profiles with global coverage and precipitable water up to 6 cm (Galve et al., 2008). The quadratic dependence on T_{11}-T_{12} accounts for the increase of the atmospheric attenuation for large amounts of atmospheric water vapor. The algorithm depends explicitly on precipitable water in the emissivity terms only; however, the impact of W in LST is small (typically less than 0.1 K for a variation of 0.5 cm in W). Eq. (2) does not take into account directional temperature and emissivity effects, which are expected to be small in the nadir view of AATSR (θ<21.7º). In Eq. (2), temperatures can be either in K or ºC.

Since emissivity only appears explicitly in the last two terms of the right-hand side of Eq. (2), it seems that atmospheric and emissivity effects are separated in the LST retrieval. However, it should be noted that emissivity effects are also implicit in the "non-emissivity" or "atmospheric" part of Eq. (2) through the brightness temperature T_{11} and the temperature difference T_{11}-T_{12}, which are determined by both the atmosphere and the surface emissivity. For surfaces where $\Delta\varepsilon$ is positive (negative), T_{11}-T_{12} is larger (smaller) than it would be for the same atmosphere over a black-body. Therefore, when T_{11}-T_{12} is multiplied by "atmospheric" coefficients in Eq. (2), there is an over or under-correction of atmospheric effects depending on the sign of $\Delta\varepsilon$. The over or under-correction induced by emissivity is compensated by the corresponding emissivity term in Eq. (2). Similarly, emissivity effects on T_{11} are also accounted for by the emissivity terms. Note that the theoretical definitions of the "emissivity" coefficients contain the "atmospheric" coefficients (see Eq. 5 and 6 of Coll and Caselles, 1997)

The critical issue for the application of the algorithm is the need of accurate emissivity maps as ancillary data. For the present study, emissivity maps were derived using the vegetation cover method (Valor and Caselles, 1996), which is based on a physical surface model and estimations of fractional vegetation cover through spectral indices. In this method, the emissivity in band k is estimated through the relationship:

$$\varepsilon_k = \varepsilon_{kv}\, f + \varepsilon_{kg}\, (1\text{-}f) + 4\, \langle d\varepsilon_k \rangle\, f\, (1\text{-}f) \qquad (3)$$

where ε_{kv} and ε_{kg} are the vegetation and ground emissivity, respectively, $\langle d\varepsilon_k \rangle$ is the maximum cavity term, and f is the fractional vegetation cover. The cavity term accounts for the effect of radiance internal reflections between the different components of a structured and rough surface (Caselles and Sobrino, 1989).

Since ε_{kv}, ε_{kg} and $\langle d\varepsilon_k \rangle$ depend on the surface type, they were obtained by combining the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Spectral Library (Baldridge et al. 2009) and the GLOBCOVER (GLC) global classification (Bicheron et al., 2008). The ASTER Spectral Library is the most extensive published dataset of TIR reflectance spectra including both natural (soils, rocks, vegetation, minerals) and manmade (asphalt, tar, concrete, brick, tile) materials. The GLC biome map provides the global classification with best spatial resolution (300 m) to date, which is one of the key factors to improve classification accuracy (Herold et al., 2008; Heiskanen, 2008). It is generated from MERIS data (with reasonably good spectral resolution) using an unsupervised classification regional expert-tuned procedure similar to the predecessor GLC2000 classification (Bartholomé and Belward 2005), and is compatible with the standardized legend of the United Nations Food and Agriculture Organization Land Cover Classification System (LCCS). It should be noted that the approach used here for modeling surface emissivity based on land surface classes and fractional vegetation cover is similar to that used by the operational AATSR LST algorithm to define the coefficients of Eq. (1).

Surface type maps were derived from GLC data. Each GLC surface type was assigned an emissivity class depending on the soil and vegetation type contained and the surface structure (see Table 2), and then the vegetation and ground emissivities for each class were estimated from the ASTER Spectral Library data. Using this procedure, the initial 22 classes (and associated subclasses) were reduced to only 10 classes, taking into account the components included in each class and the similarity between surface structures (Table 2).

Table 3 gives the corresponding ε_{kv}, ε_{kg} and $\langle d\varepsilon_k \rangle$ values for Eq. (3), or alternatively effective emissivity values, for each class. For the case of vegetated surfaces, classes were grouped attending to structure (low grasses/crops, shrubs/trees lower than 5 m, shrubs/trees higher than 5 m), background (soil or water depending on flooding conditions), and vegetation type (green grasses, evergreen or deciduous shrubs/trees). In each case, the emissivity values for vegetation and ground (or water) were calculated using the emissivity spectra of the samples given in the ASTER Spectral Library. The spectra were first convolved with the AATSR spectral response curves for the 11 and 12 μm bands to get the band emissivity values. These values were then averaged for the selected samples. In the case of soils, all samples available in the library (52) were used, which showed low variability in these bands (standard deviation smaller than ±0.005). There are only four vegetation samples. The green grass sample was used for classes 1 and 3, the average between conifer and deciduous samples for classes 2 and 4, the conifer sample for class 6, and the deciduous sample for class 5. Rocks were excluded since they should not be usual in these surface types.

For vegetated surfaces with a significant structure (emissivity classes 2, 4, 5 and 6) the maximum cavity term was determined with a simulation procedure. According to Caselles and Sobrino (1989), the cavity term for nadir observation is given by $d\varepsilon_k=(1-\varepsilon_{kg})\varepsilon_{kv}F(1-f)$, where F is a shape factor that depends on the height and separation between the surface elements, and considers the energy transmission between them. The cavity term was simplified and parameterized in terms of fractional vegetation cover only and a maximum cavity term ($<d\varepsilon_k>$ in Eq. 3) that represents the maximum value for a given surface geometry with f ranging from 0 to 1 (Valor and Caselles, 1996). The cavity term was calculated for different geometric structures, using the vegetation and ground/water emissivities obtained above, and for fractional vegetation cover ranging from 0 to 1. The maximum value for each case was selected. Then, the average of the maximum values was calculated for each class, resulting in the $<d\varepsilon_k>$ values given in Table 3.

In the case of non-vegetated surfaces, such as bare rock, water, or snow and ice, average values were calculated from the samples provided by the ASTER Spectral Library (389 rock samples, 2 water samples and 4 snow and ice samples), and a unique effective value for each AATSR band was calculated. Certainly, rock emissivities show high standard deviations, and probably it would be necessary to distinguish them using additional rock maps. Finally, effective emissivity values were calculated for urban areas using the spectra for manmade materials (tiles, asphalt, concrete, etc.) and considering regular city structures.

Fractional vegetation cover required in Eq. (3) was calculated from normalized difference vegetation index (NDVI) and reflectance values in AATSR red (0.66 μm) and near infrared (0.87 μm) bands using the following relationship (Valor and Caselles, 1996):

$$f = \frac{\left(1 - \frac{NDVI}{NDVI_s}\right)}{\left(1 - \frac{NDVI}{NDVI_s}\right) - K\left(1 - \frac{NDVI}{NDVI_v}\right)} \tag{4}$$

where NDVI is the pixel vegetation index, $NDVI_s$ and $NDVI_v$ are the index values for bare soil and full vegetation, and factor K is

$$K = \frac{\rho_{2v} - \rho_{1v}}{\rho_{2s} - \rho_{1s}} \tag{5}$$

where ρ_{1v} and ρ_{2v} are respectively the red and near infrared reflectance values over full vegetation, and ρ_{2s} and ρ_{1s} are the corresponding reflectances over bare soil. All these coefficients can be extracted from the AATSR scene itself. As an example, Figure 1 shows mean emissivity ($\varepsilon=(\varepsilon_{11}+\varepsilon_{12})/2$) and band emissivity difference ($\Delta\varepsilon=\varepsilon_{11}-\varepsilon_{12}$) maps derived for the Eastern part of Spain using two AATSR scenes acquired on March and July, 2007. In order to minimize the impact of cloud cover, the procedure can be repeated for several dates along the year to produce monthly emissivity composites at regional or global scales.

3. T-BASED VALIDATION

The T-based method for satellite LST validation is the direct comparison with ground measurements performed at a field site concurrently with the satellite overpass. Thus, it provides an independent evaluation of the radiometric quality of the satellite instrument together with the ability of the LST retrieval algorithm to correct for atmospheric and emissivity effects. However, it is necessary that the LST observed by the ground instruments at several points over the test site is truly representative of the average LST over the instantaneous field of view of the satellite sensor, for which the site must be thermally homogeneous from the point scale to several kilometers. Since most of the Earth's surface is heterogeneous at these spatial scales, high-quality ground validation data are limited to few biomes such as lakes, silt playas, grasslands and agricultural fields collected during dedicated campaigns (Wan et al., 2002 and 2004; Coll et al., 2005). An exception is the Lake Tahoe automated validation site (Hook et al., 2007), where lake surface temperatures are continuously measured since 1999.

A homogeneous site for LST validation was established in a large (>30 km^2), flat area of rice fields close to Valencia, Spain since 2002. In July and August, rice crops are well irrigated and attain nearly full vegetation cover (see Figure 2). It makes the site highly homogeneous in terms of both surface temperature and emissivity and eases the radiometric measurement of LST. Ground data from the Valencia test site have been used in previous studies (Coll et al., 2005, 2006, 2007 and 2009a). The daytime thermal homogeneity of the site was assessed with AATSR and MODIS data at 1×1 km^2, and with ASTER TIR data at 90×90 m^2. Results showed typical variability (standard deviation) ranging from 0.2 to 0.5 K depending on the spatial resolution.

Several TIR radiometers were distributed over a 1 km^2 grid to measure the ground LST and its variability concurrent to AATSR daytime overpasses (10:20-10:40 UTC). The instruments were two CIMEL CE 312-1 radiometers with four bands 8.0-13.3 μm, 11.5-12.4 μm, 10.2-11.3 μm, and 8.3-9.3 μm) (Brogniez et al., 2003), one CIMEL CE 312-2 radiometer with six bands (8.0-13.3 μm, 8.3-8.6 μm, 8.5-8.9 μm, 9.0-9.3 μm, 10.2-11.0 μm, and 10.9-11.7 μm), and two Everest model 112.2L thermometers (8-13 μm) (www.everestinterscience.com). The instruments were calibrated against a reference blackbody before and after each field measurement day and intercompared in the field. The ground LSTs were calculated by averaging the ground temperatures measured by the available radiometers within three minutes centered on the satellite overpass time. The standard deviation of the ground temperatures was calculated as a measure of the spatial and temporal variability of LST (typically ≤0.5 K). More details on the ground LST measurements can be found in Coll et al. (2005 and 2006).

Radiometric temperatures were corrected for emissivity effects, including the reflection of the sky irradiance. Surface emissivity was measured in the field using the box method (Rubio et al., 2003) for the four channels of the CE 312-1 radiometers. The CE 312-1 channels at 10.2-11.3 and 11.5-12.4 μm are similar to the AATSR bands at 11 and 12 μm, respectively. A total of 30 emissivity measurements were taken at different spots on the test site for each channel. More details on the field emissivity measurements are given in Coll et al. (2007). We obtained uniform and high emissivity values (ε=0.985±0.005) with small spectral variation in the 8-13 μm range, as expected for full vegetation cover.

Together with the average ground LST, we estimated the total uncertainty including the radiometer calibration error, the emissivity correction error (~0.2 K for measured emissivity uncertainty of 0.005), and the LST variability. A total of 28 cloud-free, daytime concurrences of ground and AATSR data were collected in July and August, 2002-2007. Table 4 lists the ground LST (T_g) and uncertainty for each case. The center of the 1 km^2 grid was at 0º17’50’’W, 39º14’27’’N in 2002-2003; 0º17’43”W, 39º15’01”N in 2004; and 0º18’28”W, 39º15’54”N in 2005-2007. For each case, we extracted the concurrent AATSR brightness temperatures in the 11 and 12 μm bands, nadir view from L1b scenes (ATS_TOA_1P, georeferenced, top of atmosphere data). We used arrays of 3×3 pixels centered on the measurement site, for which the average value was calculated (see Table 4). The standard deviation of T_{11} and T_{12} for the 3×3 pixels was typically 0.1 K.

For each validation case, Table 4 shows the AATSR-derived LST (T_{AATSR}) as obtained from:

(i) The operational LST product in the AATSR Level 2 data. It is the result of applying Eq. (1) with the operational classification and fractional vegetation cover maps. The Valencia test site is classified as biome i=6 (broadleaf trees with ground cover) and the assigned f value is 0.40-0.47 in July-August. The data in Table 4 are the average for 3×3 pixels centered on the site.

(ii) The optimized AATSR algorithm, in which we assigned biome i=8 (broadleaf shrubs with ground cover) and f=1 (full vegetation cover) in Eq. (1) for application to the brightness temperatures in Table 4. This selection of i and f yields the best agreement of LSTs derived from Eq. (1) with the ground data (Coll et al., 2009a).

(iii) The explicit emissivity-dependent algorithm (Eq. 2) applied to the brightness temperatures using the emissivity values obtained from Eq. (3) (ε=0.986 and $\Delta\varepsilon$=-0.005 at the site, with f=0.91, emissivity class 1, and GLC class 11) and the mean precipitable water for July and August (2.5 and 2.7 cm, respectively) as obtained from the National Centers for Environmental Prediction (NCEP) global tropospheric analyses product (Kalnay et al., 1996).

Figure 3 shows the comparison between T_{AATSR} and T_g, and Figure 4 plots the LST error $\delta T = T_{AATSR} - T_g$ as a function of time for the three options of LST retrieval. The operational LST product clearly overestimated the ground LST by 2 to 5 K depending on the validation case. For the 28 cases, the average value of δT or bias was 3.6 K with standard deviation of 0.7 K. However, the optimized AATSR algorithm and Eq. (2) agreed very well with each other and with the ground LSTs, δT values ranging between -1.0 and +1.0 K for most of the cases. In Figure 3, the correlation coefficient (R) between T_{AATSR} and T_g is 0.88 for the optimized algorithm and 0.90 for Eq. (2). For the optimized algorithm the average bias was 0.2 K and the standard deviation was 0.5 K, yielding rmse=±0.5 K, while for Eq. (2) the bias was 0.4 K, and the standard deviation was 0.5 K (rmse=±0.6 K). Although the validation cases are few, covering one surface type with limited atmospheric and LST variability, the results show the good accuracy achievable with AATSR data. On the other hand, Figure 4 suggests that the positive bias in AATSR LSTs is slightly increasing in the 2006 and 2007 campaigns. This effect may be due to the small number of validation cases available for the last two years, and it will be investigated in forthcoming field campaigns.

The large overestimation given by the operational LST product is due to incorrect classification and fractional vegetation cover assigned to the test site. Using the same biome (i=6) but setting f=1 in Eq. (1), the AATSR algorithm would yield better agreement with the ground data (δT between 0.4 and 2.4 K, average bias of 1.4 K). As an alternative to the AATSR classification, Noyes et al. (2007b) considered high-resolution (1 km) global biome classification maps such as those from the University of Maryland (UMD) (Hansen et al., 2000) and the International Geosphere-Biosphere Project (IGBP) (Loveland et al., 1998). The Valencia site is classified as cropland in both cases, which corresponds to AATSR biome 12 (broadleaf-deciduous trees with winter wheat). Using i=12 and f=1 in Eq. (1), the AATSR-derived LSTs overestimated the ground LSTs by a range from 1.2 to 3.3 K for the 28 validation cases, with average bias of 2.3 K.

Excluding the optimized algorithm mentioned above (i=8, f=1), the best agreement with the ground LST was obtained for vegetation biomes 2 (broadleaf deciduous trees), 3 (broadleaf and needleleaf trees) and 5 (needleleaf-deciduous trees) using f=1, resulting in biases of -0.3, -0.4 and 0.5 K, respectively. Other vegetation biomes (e.g., 1, 4, 9, 10 and 12 in Table 1) yielded large positive biases (2.3-2.5 K) even with f=1. In terms of surface emissivity, all the above fully-vegetated biomes should be quite similar, with ε and Δε values close to those used for the Valencia site and thus the retrieved LSTs should be similar. However, they produce rather different LST estimates showing the high sensitivity of the AATSR LST algorithm to the biome assignation. Noyes et al. (2007b) also noted a high sensitivity of the algorithm to the fractional vegetation cover. For comparison, an uncertainty of ±0.005 in both ε_{11} and ε_{12} results in an LST uncertainty of ±0.5 K in the explicit emissivity-dependent algorithm (Eq. 2).

4. R-BASED VALIDATION

The R-based method (Wan and Li, 2008) does not require ground LST measurements and thus provides an alternative for satellite LST validation over a wide range of biomes and LST and atmospheric regimes. The ground temperatures are calculated from radiative transfer simulations using at-sensor brightness temperatures in the 11 μm band, near concurrent atmospheric temperature and water vapor profiles and surface emissivity measurements. As mentioned above, emissivity in AATSR bands at 11 and 12 μm is high and exhibits small variations for most land cover biomes, so the R-based method could be applied, with care, globally. In this study, radiative transfer simulations were performed using the MODTRAN 4 code (Berk et al., 1999). In the next section, the R-based method is described. In section 4.2, the feasibility of the method is assessed through a comparison with ground LST measurements and a sensitivity analysis is performed. R-based validation results for two new biomes (bare soil and lake) are presented in section 4.3.

4.1. Description of the R-based method

The R-based method is physically based on the radiative transfer equation. For a surface at temperature T and with spectral emissivity ε_λ, the at sensor spectral radiance measured in band k at zenith viewing angle θ (L_k^{sen}) can be written as

$$L_k^{sen} = \int_0^\infty f_k(\lambda)\,\{[\varepsilon_\lambda B_\lambda(T) + (1-\varepsilon_\lambda)\frac{F_\lambda^{\downarrow}}{\pi}]\tau_\lambda(\theta) + L_\lambda^{\uparrow}(\theta)\}d\lambda \qquad (6)$$

where $f_k(\lambda)$ is the normalized spectral response function of band k ($\int_0^\infty f_k(\lambda)d\lambda = 1$), B_λ is the Planck function for blackbody spectral radiance, τ_λ is the atmospheric transmittance, $L_\lambda^{\uparrow}$ is the atmospheric radiance emitted towards the sensor, and $F_\lambda^{\downarrow}$ is the downwelling sky irradiance (Lambertian reflection assumed). The brightness temperature, T_k, corresponding to the at-sensor radiance is defined as $L_k^{sen}=B_k(T_k)$, B_k being the band-averaged Planck function,

$$B_k(T^*) = \int_0^\infty f_k(\lambda) B_\lambda(T^*) d\lambda \qquad (7)$$

where T* is a generic temperature. Using Eq. (7), look-up tables for T*-B_k(T*) conversion can be generated for a given band at small temperature steps in order to calculate T_k from L_k^{sen}. Equation (6) is the basis of the so-called forward simulation, in which the at-sensor radiances or brightness temperatures are simulated from surface temperature and emissivity data, and atmospheric parameters (τ_λ, $L_\lambda^{\uparrow}$, and $F_\lambda^{\downarrow}$) calculated with MODTRAN 4 from temperature and water vapor profiles. Similarly, Eq. (6) can also be used for inverse simulation, that is, the calculation of surface temperature (T) from satellite T_k or L_k^{sen}. In this case, the right-hand side of Eq. (6) can be calculated iteratively for different T values until it agrees with the prescribed L_k^{sen} value.

The R-based method requires accurate temperature and water vapor profiles representing the actual atmospheric conditions at the time of the satellite observation. The accuracy of the atmospheric profiles can be assessed with the test suggested by Wan and Li (2008), which involves the calculation of $\delta(T_{11}-T_{12})=(T_{11}-T_{12})_{obs}-(T_{11}-T_{12})_{sim}$, that is, the difference between the $T_{11}-T_{12}$ value observed by AATSR and the $T_{11}-T_{12}$ value simulated from the atmospheric profiles and the surface emissivity data. The test relies on the fact that the atmospheric effect is larger at 12 μm, owing to the water vapor continuum absorption. Then $T_{11}-T_{12}$ is usually positive and increases with the atmospheric water vapor. When the atmospheric profile used for the R-based LST calculation is over (under) correcting the atmospheric effect, then $\delta(T_{11}-T_{12})<0$ (>0) since the calculated, profile based $T_{11}-T_{12}$ value is larger (smaller) than the actual AATSR value. Therefore, $\delta(T_{11}-T_{12})$ should be close to zero when the atmospheric profiles used in simulations represent the real atmospheric conditions and the effect of the surface emissivity uncertainties is small. The R-based method can be summarized as follows:

a) Calculation of the R-based in situ LST (T_{R-b}) from the 11 μm band at-sensor radiance using atmospheric profiles and surface emissivity data (inverse simulation). The 11 μm band is used because it is less affected by variations in atmospheric water vapor and temperature. The LST error (δT) is the difference between the product LST and T_{R-b}.

b) Calculation of brightness temperatures in bands at 11 and 12 μm using T_{R-b} as ground LST, atmospheric profiles and surface emissivity data (direct simulation). The difference $\delta(T_{11}-T_{12})$ between the actual and the simulated $T_{11}-T_{12}$ value is obtained. Note that the brightness temperature T_{11} simulated here is equal to the measured T_{11} used in step a). Thus, $\delta(T_{11}-T_{12})$ is equal to the difference between the simulated and the measured brightness temperature T_{12}. Since the 12 μm band is not used in the T_{R-b} calculation, it provides an independent assessment of the atmospheric profiles.

A good knowledge of the spectral emissivity of the site is necessary for the application of the method. It should be applied over long time periods at each site to analyze the relationship between δT and δ(T_{11}-T_{12}) and select the cases with δ(T_{11}-T_{12}) values around zero for which the error introduced by the atmospheric profiles is small.

4.2. Feasibility and sensitivity analysis of the R-based method

In order to show the feasibility of the R-based method, we compared the ground LSTs derived from radiative transfer simulations with concurrent, ground-measured LSTs. Two ground LST datasets were used. (1) The Valencia data shown in Table 4 corresponding to full vegetation cover, with LSTs ranging from 25 to 31 ºC and W ranging from 2 to 4 cm. (2) A set of 99 ground measurements of lake surface temperature from Lake Tahoe, USA (Hook et al., 2007) concurrent to AATSR overpasses in 2002 and 2003, both daytime and nighttime, which have been used previously in Coll et al. (2009a). Cloud-free conditions were ensured by means of a statistical cloud masking algorithm (any scenes with a mean <4.4 ºC and/or a standard deviation >0.44 K were excluded). For the Lake Tahoe data, surface temperature ranges from 7 to 22 ºC, and W from 0.2 to 2 cm. Therefore, the two datasets cover a wide range of LST and atmospheric conditions. Moreover, the two sites are homogeneous in terms of emissivity with well-known emissivity values based on ground measurements (ε_{11}=0.985 and ε_{12}=0.980 for the rice fields) or spectral libraries (ε_{11}=0.991 and ε_{12}=0.985 for water; Baldridge et al., 2009), which facilitates the application of the R-based method.

For the Valencia site, we used three types of atmospheric profiles. Radiosonde balloons were launched at the test site near-concurrently with AATSR overpasses for 5 cases (cases 16, 19, 21, 23, and 27 in Table 4). Since local radiosondes measurements are seldom and limited to dedicated campaigns, we also used the NCEP global tropospheric analyses product (Kalnay et al., 1996), which provides global atmospheric data at 1º×1º grids at 00:00, 06:00, 12:00 and 18:00 UTC. For each case in Table 4, the NCEP profiles for the four grids closest to the site and the two times closest to the overpass were linearly interpolated to obtain the NCEP profile at the site and the overpass time. Finally, we also used atmospheric profiles derived from the Aqua/AIRS instrument (Susskind et al., 2003) with spatial resolution of 45 km and overpass time at the site between 12:30 and 13:30 UTC. AIRS profiles were available for 16 cases in Table 4 (10-12, 14, 16-23, and 25-28). For the Lake Tahoe dataset, we only used spatially and temporally interpolated NCEP atmospheric profiles for the 99 validation cases.

The profiles were entered in MODTRAN 4 to calculate the atmospheric transmittance and emitted upwelling and downwelling radiances. Using Eq. (6), we calculated the ground LST ($T_{R\text{-}b}$) from the corresponding brightness temperature T_{11}, and the difference $\delta(T_{11}-T_{12})$ for each case. Figure 5 plots the difference between the ground-measured LSTs and the R-based LSTs ($\delta T=T_g-T_{R\text{-}b}$) against $\delta(T_{11}-T_{12})$ for all the sites and atmospheric profiles considered (49 cases in Valencia and 99 in Lake Tahoe). For all datasets, a close relationship (R=0.91) is observed between δT and $\delta(T_{11}-T_{12})$, with large positive (negative) values of δT corresponding to large positive (negative) values of $\delta(T_{11}-T_{12})$, and small LST errors associated with a narrow range of $\delta(T_{11}-T_{12})$ values in the vicinity of zero. The correlation is slightly higher for the Lake Tahoe dataset alone because the uncertainties in the ground LST measurements and the water emissivity are smaller than in Valencia, as well as the precipitable water.

For most of the cases with local radiosondes, the δT and $\delta(T_{11}-T_{12})$ values were relatively small, showing that the atmospheric profiles were accurate. The NCEP profiles provided large, negative values of δT and $\delta(T_{11}-T_{12})$ in some cases, meaning that the simulations overestimated the atmospheric effect. For other cases, however, the NCEP profiles yielded accurate results with δT and $\delta(T_{11}-T_{12})$ close to zero. Most of the cases with AIRS profiles resulted in large, positive δT and $\delta(T_{11}-T_{12})$ values. The bad performance of AIRS profiles may be due to the temporal gap between Envisat and Aqua (2-3 h.), and the coarse spatial resolution of AIRS data (45 km). The data in Figure 5 correspond to a wide set of LST and atmospheric regimes and cover wide ranges of δT and $\delta(T_{11}-T_{12})$. The linear regression for all data yields $\delta T=1.78\times\delta(T_{11}-T_{12})+0.02$. From it, we can obtain $-0.6\ K<\delta(T_{11}-T_{12})<0.6\ K$ as the condition for which the error in the calculated R-based LST is within ±1.0 K.

The data in Figure 5 are affected by uncertainties in both the ground LSTs and the radiative transfer calculations involved in the R-based method. A sensitivity analysis was performed to assess the accuracy of the simulated T_{R-b} and $T_{11}-T_{12}$ values using the radiosonde profile of case 16 as a representative case (W=2.5 cm). The effect of atmospheric uncertainties was simulated in two ways. First, the water vapor mixing ratio was increased by 10% at each profile level. Second, the air temperature was increased by 1 K at each level. These changes represent small temporal and spatial atmospheric variations or typical radiosonde errors. Table 5 shows the effect of the atmospheric variations on T_{R-b} (inverse simulation), T_{11}, T_{12} (forward simulation) and the difference $T_{11}-T_{12}$. Atmospheric effects on T_{11} and T_{12} have the same sign, so they cancel in part in $T_{11}-T_{12}$. The uncertainty in the MODTRAN 4 code was assumed as the root square sum (rss) of the errors due to the water vapor and air temperature changes.

We assumed an uncertainty of 0.005 in ε_{11} and ε_{12}, which is justified for vegetated surfaces and is a typical uncertainty in field emissivity measurements (Rubio et al., 2003). However, larger uncertainties may be expected for bare surfaces. For the present case, the resulting uncertainty in the simulated temperatures is shown in Table 5. Emissivity uncertainties are regarded as random and may have opposite signs in the two bands. For this reason, the emissivity uncertainty in the simulated T_{11}-T_{12} is obtained as the rss of the errors in T_{11} and T_{12}. Finally, the total uncertainty in the simulated temperatures is estimated as the rss of the above atmospheric, MODTRAN 4 and emissivity errors, resulting in ±0.7 K for $T_{R\text{-}b}$ and ±0.4 K for the simulated T_{11}-T_{12}.

NCEP and AIRS profiles may be less accurate than assumed in the previous uncertainty analysis, thus the errors in the calculated $T_{R\text{-}b}$ and T_{11}-T_{12} may be larger. To assess the accuracy of the NCEP and AIRS profiles, we compared the five cases of the Valencia dataset with local radiosondes with the same cases with NCEP and AIRS profiles. If we assume that the local radiosondes represent accurately the atmospheric state, then the difference between $T_{R\text{-}b}$ calculated from radiosonde and NCEP/AIRS profiles is the LST error due to the incorrect profile. Similarly, the difference between the radiosonde T_{11}-T_{12} and the corresponding NCEP/AIRS value is the $\delta(T_{11}$-$T_{12})$ difference due to the incorrect profile. Figures 6a and 6b show the difference δT between radiosonde and NCEP/AIRS $T_{R\text{-}b}$ values against the precipitable water difference W_{rad}-$W_{NCEP/AIRS}$ and $\delta(T_{11}$-$T_{12})$, respectively. δT shows a good correlation (R=0.82) with the water vapor differences, with NCEP/AIRS overestimating (underestimating) radiosonde $T_{R\text{-}b}$ when NCEP/AIRS water vapor overestimated (underestimated) radiosonde water vapor. Furthermore, the correlation between δT and $\delta(T_{11}$-$T_{12})$ is excellent (R=0.94) and shows a similar behavior as in Figure 5. The correlation in Figure 6b is higher than in Figure 6a because the difference in the R-based LSTs depends not only on the precipitable water difference but also on the difference in temperature profiles and viewing angle, and all these effects are also largely reflected in $\delta(T_{11}$-$T_{12})$.

The results of this section show that large errors in R-based LSTs are due to inappropriate atmospheric profiles, and that the $\delta(T_{11}-T_{12})$ test may be used to select the profiles that reasonably represent the actual atmosphere for the AATSR observation. Cases meeting the condition -0.6 K<$\delta(T_{11}-T_{12})$<0.6 K correspond to errors within ±1.0 K in the calculated in situ LST, which is appropriate for LST validation. This condition was derived using ground measured LSTs covering a wide range of conditions (Figure 5), and is compatible with the preceding uncertainty analysis and the δT-$\delta(T_{11}-T_{12})$ relationship shown in Figure 6b.

4.3. Application of the R-based method

The R-based method was applied to two different biomes in the Valencia test site for a total of 47 cloud-free, daytime and nighttime AATSR scenes from March to early May in 2003-2008. Cloud-free conditions were assessed by means of visual inspection and threshold tests using LST-T_{11}, T_{11}-T_{12}, and the standard deviation of T_{11}. On these dates, rice fields are fallow and the soil is left dry before the start of the growing season in mid May. Therefore the site appears as a wide area of uniform, dry bare soil. As for the summer full-cover conditions, we assessed the uniformity of the site with satellite data at different resolutions (AATSR and MODIS at 1 km, ASTER and Landsat at ~100 m) and field surveys. We tried to use the bare soil for T-based validation, but preliminary tests showed that the ground-scale thermal heterogeneity of the site was higher than for the full-vegetation case, and thus the ground LSTs may not be accurate enough for AATSR LST validation. For this reason, we applied the R-based method to the bare soil site using concurrent NCEP atmospheric profiles (local radiosondes were not available), and laboratory emissivity measurements. Different samples were taken from the site for mineralogy analysis, resulting in homogeneous composition of 14% sand, 35% clay, 4.5% organic matter and 44% carbonate (Mira et al., 2007). According to the measurements (30 per channel), the soil emissivity was quite uniform with values

ε_{11}=0.957 and ε_{12}=0.954 (uncertainty of ±0.005) for dry conditions (Mira et al., 2007). Additionally, we applied the R-based method over the nearby Albufera Lake (about 7 km in diameter, a few kilometers north of the rice fields site, see Fig. 2b) using the same AATSR scenes and NCEP profiles, and taking ε_{11}=0.991 and ε_{12}=0.985 for water.

The AATSR viewing angle (θ) and brightness temperatures (T_{11} and T_{12}) for the bare soil and lake cases are shown in Tables 6 and 7, respectively. These data are the average values for 3×3 pixels centered on the validation sites. In both Tables, we also show the AATSR-derived LSTs (T_{AATSR}) using the three algorithms as in section 3: (i) the operational algorithm, which assigns biome i=6 and f=0.35-0.40 in March-May for both sites; (ii) the optimized AATSR algorithm selecting biome i=11 (bare soil) for the bare soil cases and i=14-d/n (lake day/night) for the lake cases; and (iii) the explicit emissivity-dependent algorithm (Eq. 2) using the emissivity values derived from Eq. (3) (ε=0.974 and Δε=-0.007 with f=0.06 for bare soil, ε=0.988 and Δε=0.006 for water) and the NCEP mean precipitable water for March-May (1.2-1.9 cm). T_{AATSR} values from options (i) and (ii) are quite similar for the bare soil site (Table 6), with differences of a few tenths of K. For the lake site (Table 7), options (ii) and (iii) agree very well with each other and differ by 1-4 K from option (i).

Table 8 shows the calculated R-based LST ($T_{R\text{-}b}$) and the difference $\delta(T_{11}\text{-}T_{12})$ for each bare soil and lake validation case from Tables 6 and 7. The precipitable water obtained from the NCEP profiles (W_{NCEP}) is also shown. When comparing each case, the differences between $\delta(T_{11}\text{-}T_{12})$ values for bare soil and lake are relatively small (root mean square difference of 0.2 K), being compatible with the uncertainty due to small emissivity errors (most likely affecting the soil). Nearly all the validation cases met the condition -0.6 K<$\delta(T_{11}\text{-}T_{12})$<0.6 K derived in section 4.2. The outliers are marked with ***** in Table 8. This means that, for most cases, the NCEP profiles provided a sufficiently accurate description of the atmosphere over the site at the time of AATSR observations, and the R-based LSTs can be used for AATSR LST validation. Figure 7 shows the AATSR LST error ($\delta T=T_{AATSR}-T_{R\text{-}b}$) against $\delta(T_{11}-T_{12})$ for the three algorithm options and the two sites. For all algorithms and sites, Figure 7 points out a high correlation between δT and $\delta(T_{11}-T_{12})$, especially for algorithm (iii) in both sites, and algorithm (ii) in the lake site. The relationship between δT and $\delta(T_{11}-T_{12})$ is similar as in Figure 5, which gives confidence in the R-based results.

Considering only the δT values for the cases meeting the above $\delta(T_{11}-T_{12})$ condition, we computed the average bias, standard deviation and rmse for each algorithm and site, which are given in Table 9 together with the range of δT. Algorithm (iii) yielded the smallest LST errors for both sites (rmse=±0.4 K), with δT values within ±1.0 K for all cases. The optimized algorithm (ii) resulted in small errors for the lake site (rmse=±0.5 K), in good agreement with the T-based validation in Lake Tahoe shown in Coll et al. (2009a). For the bare soil site, LST errors were still acceptable (rmse=±1.1 K) and generally met the target accuracy of the AATSR algorithm (±1.0 K at nighttime and ±2.5 K at daytime). Again, the largest LST errors were provided by the operational algorithm (i) for the lake cases because of the misclassification of the site. However, results were better for bare soil (rmse=±1.3 K), with similar LST errors as the optimized algorithm.

We did not find any significant long-term temporal variability in the R-based LST errors for the data used in this section (2003-2008). Comparing the daytime and nighttime results, the operational algorithm showed lower biases at nighttime, with rmse close to ±1 K for both bare soil and water. For the optimized algorithm over bare soil, the bias changed from 0.9 K at daytime to -0.9 K at nighttime. For Eq. (2), no significant differences were found between daytime and nighttime at both sites.

We observed a high dependence of the LST errors of the operational and the optimized algorithm on the actual AATSR brightness temperature difference T_{11}-T_{12}, a magnitude that plays a key role in the split-window algorithms (see Eqs. 1 and 2). T_{11}-T_{12} is a function of atmospheric precipitable water, but also depends on the temperature profile, LST, surface emissivity and observation angle. Figure 8 plots the LST errors against T_{11}-T_{12} for the three algorithms over bare soil. It appears that AATSR LSTs based on Eq. (1) overestimate (underestimate) the R-based LSTs for large (small) values of T_{11}-T_{12}, which suggests that coefficients $b_{f,i}$ in Eq. (1) may be overestimated. For the algorithm of Eq. (2), the LST error dependence on T_{11}-T_{12} was much smaller. A few validation cases (32, 34, 38, 40 and 41, all corresponding to nighttime in March) showed small values of T_{11}-T_{12} (-0.05 to 0.13 K over water, see Table 7) and T_{11} values close to the ground LST, which suggests that they may be affected by near surface atmospheric temperature inversions (Platt and Prata, 1993). According to the NCEP profiles, air temperatures at 100-135 m were 9 to 13 K warmer than T_{11} in these cases. In these possible inversion conditions, Eq. (2) resulted still in reasonable LST errors (biases of -0.5 K over water and -0.7 K over bare soil).

5. CONCLUSIONS

AATSR-derived LSTs were validated over the Valencia test site during the period 2002-2008. Validation included the conventional T-based method over full-vegetated rice fields, and the new R-based method (Wan and Li, 2008) over bare soil and lake. T-based validation provides a direct assessment of the AATSR and algorithm performance since it relies on independently measured ground LSTs. However, validation sites must be large enough, highly homogeneous in surface temperature from ground to satellite scales, and equipped with accurate ground radiometers for multiple spatial and temporal sampling in an area of at least 1 km^2. While it is necessary to establish and maintain high-quality validation sites, the T-based method is not suitable for the global validation of satellite LST products because of the large heterogeneity of land surfaces.

The R-based method provides an alternative for the semi-operational, long-term validation and diagnostics of the AATSR LST product at global scale. It could be applied over surfaces with varied LST and atmospheric regimes where ground LST measurements are not feasible (forests, partially vegetated surfaces, semi-arid areas, deserts, remote regions, etc.). Accurate measurements of surface emissivities are necessary. The R-based method requires atmospheric profiles representing adequately the atmospheric state at the time of the AATSR observation. This can be checked with the $\delta(T_{11}-T_{12})$ test, which must be applied for each atmospheric profile at each site in long periods of time to select a range of $\delta(T_{11}-T_{12})$ values around zero for which the R-based ground LSTs are sufficiently accurate. In this paper, we used spatially and temporally interpolated NCEP profiles for R-based validation over the Valencia and Lake Tahoe sites. Local radiosonde measurements and AIRS profiles were also used at the Valencia site. Another possibility for R-based validation is to select suitable validation areas around permanent radio sounding stations with launching times close to AATSR overpasses.

Results of this paper stress the need for modifications on the operational AATSR LST algorithm. The most obvious conclusion is that the ancillary data (classification and fractional vegetation cover maps) must be provided at the same spatial resolution as the AATSR data (1 km^2). Otherwise, LST errors from 2 to 5 K can be obtained. When the AATSR algorithm (Eq.

1) was optimized with i and f values selected for each validation site according to their nature, LST errors were within the target accuracy of the LST product, with negligible biases and rmse=±0.5 K for full-vegetated and lake surfaces, and ±1.1 K for bare soil. For the vegetated surface, we found that the algorithm is very sensitive to the assigned biome since the various vegetation biomes of Table 1 (with f=1) resulted in rather different biases (AATSR minus ground) ranging from -0.4 to 2.5 K. Furthermore, LST errors from Eq. (1) showed certain dependence on the brightness temperature difference T_{11}-T_{12}.

The explicit emissivity-dependent algorithm (Eq. 2) provided accurate LSTs for all the biomes studied, with small biases, rmse from ±0.4 to 0.6 K and LST errors within ±1.0 K for most validation cases. This shows the high accuracy achievable with AATSR data in ideal conditions. The ancillary data required for Eq. (2) (ε and $\Delta\varepsilon$ maps at 1 km^2 resolution) can be derived from classification maps and monthly fractional vegetation cover estimated from visible and near infrared AATSR data. It is the approach adopted by other satellite LST products such as MODIS and SEVIRI (Wan and Dozier, 1996; PUM_LST, 2008). The land cover biomes studied here (vegetation, water and bare soil) likely span the emissivity range of natural surfaces. However, these results only apply to the geographical area used in the study and the accuracy of the AATSR LST retrievals may be different in other regions and under different conditions. More validation datasets are necessary especially over deserts where emissivity may have a larger variability in response to differences in soil composition, texture and moisture content.

ACKNOWLEDGMENTS

This work was funded by the Spanish *Ministerio de Educación y Ciencia* (CGL2007-64666/CLI, and CGL2007-29819-E, and research grants of J.M. Galve, M. Mira, and V.

García-Santos), and the *Conselleria d'Educació* of the *Generalitat Valenciana* (PROMETEO/2009/086). We thank Dr. Z. Wan (UCSB) for fruitful discussion on the R-based method, Dr. S. Hook (JPL) for providing the Lake Tahoe ground LSTs, and the Centro de Estudios Ambientales del Mediterraneo (CEAM) for the radiosonde data. AATSR data were provided by ESA under Cat-1 project 3466. We also thank ESA and the ESA GLOBCOVER Project, led by MEDIAS-France, for the GLOBCOVER classification data. NCEP data were provided by the Data Support Section of the Computational and Information Systems Laboratory at the National Center for Atmospheric Research (NCAR). NCAR is supported by grants from the National Science Foundation. AIRS data were obtained thanks to the AIRS Project and Science Team through NASA's Goddard Earth Sciences Data Information and Services Center. The comments and suggestions of the anonymous reviewers are also acknowledged.

REFERENCES

Allen, M. R., Mutlow, C. T., Blumberg, G. M. C., Christy, J. R., McNider, R. T., and Llewellyn-Jones, D. T. (1994). Global change detection. *Nature*, *370*(6484), 24-25.

Anderson, M. C., Norman, J. M., Diak, G. R., Kustas, W. P., and Mecikalski, J. R. (1997). A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. *Remote Sensing of Environment*, *60*, 195-216.

Baldridge, A. M., Hook, S. J., Grove, C. I., and Rivera, G. (2009). The ASTER spectral library version 2.0. *Remote Sensing of Environment*, *113*, 711–715.

Bartholomé, E., and Belward, A. S. (2005). GLC2000: A new approach to global land cover mapping from earth observation data. *International Journal of Remote Sensing*, *26* (9), 1959-1977.

Becker, F., and Li, Z.-L. (1990). Towards a local split-window method over land surfaces. *International Journal of Remote Sensing*, *11*, 369-394.

Berk, A., Anderson, G. P., Acharya, P. K., Chetwynd, J. H., Bernstein, L. S., Shettle, E. P. , Matthew, M. W., and Adler-Golden, S. M. (1999). MODTRAN 4 user’s manual. Air Force Research Laboratory, Space Vehicles Directorate, Air Force Materiel Command, Hanscom AFB, MA, 95 pp.

Bicheron, P., Defourny, P., Brockmann, C., Schouten, L., Vancutsem, C., Huc, M., Bontemps, S., Leroy, M., Achard, F., Herold, M., Ranera, F., Arino, O. (2008). GLOBCOVER Products Description and Validation Report, MEDIAS-France, Toulouse, 47 pp.

Brogniez, G., Pietras, C., Legrand, M., Dubuisson, P. and Haeffelin, M. (2003). A high-accuracy multiwavelength radiometer for in situ measurements in the thermal infrared.

Part II: behavior in field experiments. *Journal of Atmospheric and Oceanic Technology*, *20*(7), 1023-1033.

Caselles, V. and Sobrino, J. A. (1989). Determination of frosts in orange groves from NOAA-9 AVHRR data. *Remote Sensing of Environment*, *29*, 135-146.

Coll, C., and Caselles, V. (1997). A global split-window algorithm for land surface temperature from AVHRR data: Validation and algorithm comparison. *Journal of Geophysical Research*, 102D, 16697-16713.

Coll, C., Caselles, V., Galve, J. M., Valor, E., Niclòs, R., Sánchez, J. M. and Rivas R. (2005). Ground measurements for the validation of land surface temperatures derived from AATSR and MODIS data. *Remote Sensing of Environment*, *97,* 288-300.

Coll, C., Caselles, V., Galve, J. M., Valor, E., Niclòs, R., and Sánchez, J. M. (2006). Evaluation of split-window and dual-angle correction methods for land surface temperature retrieval from Envisat/AATSR data. *Journal of Geophysical Research*, *111*, 12105 doi 10.1029/2005J D006830.

Coll, C., Caselles, V., Valor , E., Niclòs, R., Sánchez, J. M., Galve, J. M. and Mira M. (2007). Temperature and emissivity separation from ASTER data for low spectral contrast surfaces. *Remote Sensing of Environment*, *110*, 162-175.

Coll, C., Hook, S. J., and Galve, J. M. (2009a). Land surface temperature from the Advanced Along-Track Scanning Radiometer: Validation over inland waters and vegetated surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(1), 350-360.

Coll, C., Wan, Z., and Galve, J. M. (2009b). Temperature-based and radiance-based validations of the V5 MODIS land surface temperature product, *Journal of Geophysical Research*, *114*, D20102, doi:10.1029/2009JD012038.

Dorman, J. L., and Sellers, P. J. (1989). A global climatology for albedo, roughness length and stomatal resistance for atmospheric general circulation models as represented by the simple biosphere model (SiB). *Journal of Applied Meteorology*, *28*, 833-855.

Galve J. M., Coll, C., Caselles V. and Valor E. (2008). An Atmospheric Radiosounding database for generating Land Surface Temperature Algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, *46* (5),1547-1557.

Gillespie, A. R., Matsunaga, T., Rokugawa, S., and Hook, S. J. (1998). Temperature and emissivity separation from Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images. *IEEE Transactions on Geoscience and Remote Sensing*, *36*, 1113–1125.

Hansen, M. C., DeFries, R. S., Townshend, J. R. G., and Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *Remote Sensing of Environment*, *21*(6 and 7), 1331–1364.

Heiskanen, J. (2008). Evaluation of global land cover data sets over the tundra-taiga transition zone in northernmost Finland. *International Journal of Remote Sensing*, *29*(19), 3727-3751.

Herold, M., Mayaux, P., Woodcock, C. E., Baccini, A., Schmullius, C. (2008). Some challenges in global land cover mapping: an assessment of agreement and accuracy in existing 1 km datasets. *Remote Sensing of Environment*, *11*2, 2538-3556.

Hook, S. J., Vaughan, R. G., Tonooka, H. and Schladow, S. G. (2007). Absolute radiometric in-flight validation of mid infrared and thermal infrared data from ASTER and MODIS on the Terra spacecraft using the Lake Tahoe, CA/NV, USA, Automated Validation Site. *IEEE Transactions Geoscience and Remote Sensing*, *45*, 1798-1807.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebissuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D. (1996). The NCEP/NCAR 40 Year Reanalysis Project. *Bulletin of the American Meteorological Society*, *77*, 437-471.

Kerr, Y. H., Lagouarde, J.-P., and Imbernon, J. (1992). Accurate land surface temperature retrieval from AVHRR data with the use of an improved split-window algorithm. *Remote Sensing of Environment*, *41*, 197-209.

Lambin E. F., and Ehrlich, D. (1997). Land-cover changes in sub-Saharan Africa (1982-1991). Application of a change index based on remotely sensed surface temperature and vegetation indices at a continental scale. *Remote Sensing of Environment*, *61*, 181-200.

Llewellyn-Jones, D., Edwards, M. C., Mutlow, C. T., Birks, A. R., Barton, I. J., and Tait, H. (2001). AATSR: Global-change and Surface-Temperature measurements from ENVISAT. *ESA Bulletin*, *February 2001*, 11-21.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., and Merchant, J. (1998). Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1-km AVHRR Data. *International Journal of Remote Sensing*, *21*(6-7), 1303-1330.

Mira, M., Valor, E., Boluda, R., Caselles, V., and Coll, C. (2007). Influence of soil water content on the thermal infrared emissivity of bare soils: Implication for land surface temperature determination. *Journal of Geophysical Research*, *112*(F4), F04003, doi: 10.1029/2007JF000749.

Noyes, E., Corlett, G., Remedios, J., Kong, X., and Llewellyn-Jones, D. (2007a). An Accuracy Assessment of AATSR LST Data Using Empirical and Theoretical Methods. *Proceedings of the Envisat Symposium 2007*, Montreux, Switzerland, ESA SP-636 (July 2007).

Noyes, E., Corlett, G., Remedios, J., Kong, X., and Llewellyn-Jones, D. (2007b). Future development of the operational AATSR LST product. *Proceedings of the Envisat Symposium 2007*, Montreux, Switzerland, ESA SP-636 (July 2007).

Pinheiro, A. C., Privette, J., Mahoney, R., and Tucker, C. J. (2004). Directional effects in a daily AVHRR land surface temperature dataset over Africa. *IEEE Transactions on Geoscience and Remote Sensing*, *42*(9), 1941–1954.

Platt, C. M. R., and Prata, A. J. (1993). Nocturnal effects in the retrieval of land surface temperatures from satellite measurements, *Remote Sensing of Environment*, *45*, 127-136.

Prata, A. J (2002a). Land surface temperature measurement from space: AATSR algorithm theoretical basis document. Contract Report to ESA, CSIRO Atmospheric Research, Aspendale, Victoria, Australia.

Prata, A. J. (2002b). Land surface temperature measurement from space: Global surface temperature simulations for the AATSR. Contract Report to ESA, CSIRO Atmospheric Research, Aspendale, Victoria, Australia.

PUM_LST (2008). Product User Manual. Land Surface Temperature. Land SAF Project Team SAF/LAND/IM/ 621, PUM_LST/2.2. Available at http://landsaf.meteo.pt.

Rubio, E., Caselles, V., Coll, C., Valor, E. and Sospedra, F. (2003). Thermal-infrared emissivities of natural surfaces: Improvements on the experimental set-up and new measurements, *International Journal of Remote Sensing*, *24* (24), 5379-5390.

Sánchez J. M., Kustas W. P., Caselles V., Anderson M. C., (2008). Modeling surface energy fluxes over maize using a two-source patch model and radiometric soil and canopy temperature observations. *Remote Sensing of Environment*, *112*, 1130-1143.

Snyder, W., Wan, Z., Zhang, Y., and Feng, Y.-Z. (1998). Classification-based emissivity for land surface temperature measurement from space. *International Journal of Remote Sensing*, *19*, 2753-2774.

Susskind, J., Barnet, C. D., and Blaisdell, J. M. (2003). Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(2), 390-409.

Valor, E., and Caselles, V. (1996). Mapping land surface emissivity from NDVI. Application to European, African and South American areas. *Remote Sensing of Environment, 57,* 167-184.

Wan, Z. (2008). New refinements and validation of the MODIS Land-Surface Temperature/Emissivity products. *Remote Sensing of Environment*, *112*, 59−74.

Wan, Z., and Dozier, J. (1996). A generalized split-window algorithm for retrieving land surface temperature from space. *IEEE Transactions on Geoscience and Remote Sensing*, *34* (4), 892-905.

Wan, Z., and Li, Z.-L. (2008). Radiance-based validation of the V5 MODIS land-surface temperature product. *International Journal of Remote Sensing*, *29*(17-18), 5373-5395.

Wan, Z., Zhang, Y., Zhang, Q. and Li, Z.-L. (2002). Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data. *Remote Sensing of Environment*, *83*, 163-180.

Wan, Z., Zhang, Y., Zhang, Q. and Li, Z.-L. (2004). Quality assessment and validation of the MODIS global land surface temperature. *International Journal of Remote Sensing*, *25* (1), 261-274.

TABLES

Table 1. Coefficients for the operational AATSR LST algorithm (Eq. 1) for the different biomes (from Prata, 2002b).

Biome	***i***	a_{vi}	a_{si}	b_{vi}	b_{si}	c_{vi}	c_{si}
Broadleaf evergreen trees	1	0.6907	6.0951	3.8129	4.5637	-2.8456	-3.3617
Broadleaf deciduous trees	2	-0.5393	4.6301	3.6472	4.3652	-2.7218	-3.2155
Broadleaf and needleleaf trees	3	-0.6885	4.8786	3.6472	4.3652	-2.7218	-3.2155
Needleleaf-evergreen trees	4	1.0801	1.0801	3.2972	3.2972	-2.2909	-2.2909
Needleleaf-deciduous trees	5	0.7804	1.491	3.2721	3.8117	-2.3374	-2.7233
Broadleaf trees with groundcover	6	0.9089	0.0348	3.3511	3.9038	-2.389	-2.7891
Groundcover	7	0.7994	0.7994	3.5088	3.5088	-2.5065	-2.5065
Broadleaf shrubs with groundcover	8	1.5662	0.7833	3.1384	3.656	-2.2419	-2.6121
Broadleaf shrubs with bare soil	9	0.8965	0.8965	3.4867	3.4867	-2.4908	-2.4908
Dwarf trees, shrubs with groundcover	10	1.0817	1.0817	3.3039	3.3039	-2.2955	-2.2955
Bare soil	11	0.7075	0.7041	3.7832	3.7832	-2.7868	-2.7868
Broadleaf-deciduous trees with winter wheat	12	0.881	0.881	3.4106	3.4106	-2.4133	-2.4133
Perennial land ice	13	1.0801	1.0801	3.2972	3.2972	-2.2909	-2.2909
Lake-day	14-d	-0.0005	-0.0005	2.4225	2.4225	-1.4344	-1.4344
Lake-night	14-n	-0.3658	-0.3658	2.3823	2.3823	-1.3556	-1.3556

Table 2. Emissivity classes by surface type and their correspondence with the biomes defined by the GLOBCOVER (GLC) dataset.

Emissivity class	GLC Class	Description
1. Flooded vegetation, crops and grasslands	11	Post-flooding or irrigated croplands (or aquatic)
	13	Post-flooding or irrigated herbaceous crops
	180	Closed to open (>15%) grassland or woody vegetation on regularly flooded or waterlogged soil - Fresh, brackish or saline water
	185	Closed to open (>15%) grassland on regularly flooded or waterlogged soil - Fresh or brackish water
2. Flooded forest and shrublands	170	Closed (>40%) broadleaved forest or shrubland permanently flooded - Saline or brackish water
3. Croplands and grasslands	14	Rainfed croplands
	15	Rainfed herbaceous crops
	20	Mosaic cropland (50-70%) / vegetation (grassland/shrubland/forest) (20-50%)
	21	Mosaic cropland (50-70%) / grassland or shrubland (20-50%)
	120	Mosaic grassland (50-70%) / forest or shrubland (20-50%)
	140	Closed to open (>15%) herbaceous vegetation (grassland, savannas or lichens/mosses)
	141	Closed (>40%) grassland
	150	Sparse (<15%) vegetation
	151	Sparse (<15%) grassland
4. Shrublands	16	Rainfed shrub or tree crops (cash crops, vineyards, olive tree, orchards…)
	30	Mosaic vegetation (grassland/shrubland/forest) (50-70%) / cropland (20-50%)
	130	Closed to open (>15%) (broadleaved or needleleaved, evergreen or deciduous) shrubland (<5m)
	131	Closed to open (>15%) broadleaved or needleleaved evergreen shrubland (<5m)
	134	Closed to open (>15%) broadleaved deciduous shrubland (<5m)
	152	Sparse (<15%) shrubland
5. Broadleaved/ needleleaved deciduous forest	40	Closed to open (>15%) broadleaved evergreen or semi-deciduous forest (>5m)
	50	Closed (>40%) broadleaved deciduous forest (>5m)
	60	Open (15-40%) broadleaved deciduous forest/woodland (>5m)
	90	Open (15-40%) needleleaved deciduous or evergreen forest (>5m)
	91	Open (15-40%) needleleaved deciduous forest (>5m)
6. Broadleaved/ needleleaved evergreen forest	32	Mosaic forest (50-70%) / cropland (20-50%)
	70	Closed (>40%) needleleaved evergreen forest (>5m)
	92	Open (15-40%) needleleaved evergreen forest (>5m)
	100	Closed to open (>15%) mixed broadleaved and needleleaved forest (>5m)
	101	Closed (>40%) mixed broadleaved and needleleaved forest (>5m)
	110	Mosaic forest or shrubland (50-70%) / grassland (20-50%)
7. Urban area	190	Artificial surfaces and associated areas (Urban areas >50%)
8. Bare rock	200	Bare areas
	201	Consolidated bare areas (hardpans, gravels, bare rock, stones, boulders)
	202	Non-consolidated bare areas (sandy desert)
	203	Salt hardpans
9. Water	210	Water bodies
10. Snow and ice	220	Permanent snow and ice

Table 3. Coefficients for the vegetation cover method emissivity (Eq. 3) for AATSR bands at 11 and 12 μm based on the classes shown in Table 2.

Emissivity class	**11 μm**			**12 μm**		
	ε_v	ε_g	<dε>	ε_v	ε_g	<dε>
1. Flooded vegetation, crops and grasslands	0.983±0.005	0.970±0.005 (soil)	0	0.989±0.005	0.977±0.004 (soil)	0
		0.991±0.001 (water)	0		0.985±0.001 (water)	0
2. Flooded forest and shrublands	0.981±0.008	0.970±0.005 (soil)	0.014±0.004 (soil)	0.982±0.009	0.977±0.004 (soil)	0.010±0.003 (soil)
		0.991±0.001 (water)	0.004±0.001 (water)		0.985±0.001 (water)	0.007±0.002 (water)
3. Croplands and grasslands	0.983±0.005	0.970±0.005	0	0.989±0.005	0.977±0.004	0
4. Shrublands	0.981±0.008	0.970±0.005	0.014±0.004	0.982±0.009	0.977±0.004	0.010±0.003
5. Broadleaved/ needleleaved deciduous forest	0.973±0.005	0.970±0.005	0.019±0.006	0.973±0.005	0.977±0.004	0.015±0.004
6. Broadleaved/ needleleaved evergreen forest	0.989±0.005	0.970±0.005	0.019±0.005	0.991±0.005	0.977±0.004	0.015±0.004
	Effective emissivity, 11 μm			**Effective emissivity, 12 μm**		
7. Urban area	0.969±0.006			0.976±0.004		
8. Bare rock	0.93±0.05			0.95±0.05		
9. Water	0.991±0.001			0.985±0.001		
10. Snow and ice	0.990±0.004			0.971±0.014		

Table 4. Ground-measured LSTs (T_g) and uncertainties (σ) over full-vegetated rice fields in the Valencia test site, and concurrent AATSR brightness temperatures in the 11 and 12 μm bands, nadir view. θ is the satellite zenith viewing angle. The AATSR-derived LSTs (T_{AATSR}) using different options (see text) are shown in the last three columns.

case	**date (d/m/y)**	**$T_g \pm \sigma$ (ºC)**	**θ (º)**	**T_{11} (ºC)**	**T_{12} (ºC)**	**operational T_{AATSR} (ºC)**	**optimized T_{AATSR} (ºC)**	**Eq. (2) T_{AATSR} (ºC)**
1	10/07/02	28.6±0.6	3.7	25.04	22.99	32.1	28.6	28.8
2	13/07/02	27.6±0.9	13.8	22.28	19.26	31.8	28.3	28.3
3	26/07/02	27.9±0.6	1.11	23.39	20.68	32.0	28.6	28.6
4	08/08/02	26.5±0.7	16.2	20.29	17.31	29.4	26.5	26.2
5	14/08/02	28.5±0.5	3.9	23.69	21.52	30.8	27.7	27.7
6	17/08/02	29.1±0.6	13.91	22.81	19.84	32.0	28.7	28.7
7	05/09/02	28.0±0.8	19.06	24.10	22.03	31.2	27.9	27.9
8	08/07/03	28.3±0.7	11.13	25.30	23.03	33.1	29.4	29.5
9	11/07/03	29.1±0.7	1.20	27.03	25.50	33.0	29.2	29.9
10	14/07/03	28.6±0.6	8.66	24.73	22.39	32.7	29.0	29.1
11	24/07/03	28.8±0.6	16.25	24.68	22.36	32.5	28.9	29.0
12	30/07/03	28.9±0.6	3.74	23.44	20.63	32.5	28.9	28.9
13	12/08/03	31.3±0.6	11.13	28.10	26.51	33.9	30.3	31.0
14	28/06/04	29.2±0.6	8.66	26.41	24.36	34.0	29.8	30.2
15	08/07/04	25.7±0.6	16.33	23.15	21.60	28.8	25.8	26.0
16	14/07/04	27.2±0.7	3.74	22.45	19.78	31.1	27.7	27.6
17	27/07/04	27.7±0.4	11.05	25.02	23.35	31.2	27.8	28.1
18	30/07/04	27.8±0.4	1.19	23.37	20.58	32.1	28.8	28.8
19	12/08/04	28.4±0.6	16.25	25.50	23.98	31.0	27.9	28.3
20	12//07/05	27.0±0.6	11.13	24.64	23.04	30.6	27.1	27.6
21	21/07/05	28.5±0.6	18.97	25.40	23.63	31.8	28.4	28.6
22	28/07/05	28.8±0.5	16.33	24.75	22.67	32.0	28.5	28.6
23	06/08/05	28.0±0.5	13.66	25.35	23.68	31.0	28.1	28.4
24	03/07/06	29.5±0.6	8.7	27.50	25.90	33.6	29.8	30.4
25	22/07/06	29.5±0.5	13.7	26.39	24.44	33.3	29.6	30.0
26	04/07/07	27.8±0.9	3.7	23.40	20.63	32.3	28.8	28.8
27	20/07/07	27.8±0.4	1.3	24.03	21.65	32.0	28.4	28.5
28	26/07/07	27.5±0.4	19.1	24.70	22.90	31.2	27.8	28.0

Table 5. Uncertainty (in K) of the simulated temperatures in the R-based method ($T_{R\text{-}b}$, T_{11}, T_{12}, and T_{11}-T_{12}) for different sources of uncertainty (see text).

	T_{R_b}	T_{11}	T_{12}	T_{11}-T_{12}
W (10%)	0.33	0.26	0.39	0.13
T_{air} (1 K)	0.28	0.21	0.33	0.12
MODTRAN 4	0.43	0.33	0.51	0.18
ε (0.005)	0.35	0.28	0.22	0.36
Total (rss)	0.71	0.55	0.76	0.43

Table 6. R-based validation cases over bare soil in the Valencia test site. The AATSR-derived LSTs (T_{AATSR}) using different options (see text) are shown in the last three columns.

case	Date (d/m/y)	Time (UTC)	θ (º)	T_{11} (ºC)	T_{12} (ºC)	operational T_{AATSR} (ºC)	optimized T_{AATSR} (ºC)	Eq. (2) T_{AATSR} (ºC)
1	09/03/2003	10:25	7.0	11.49	10.46	15.2	15.0	14.5
2	12/03/2003	10:31	4.1	12.76	12.06	15.7	15.4	15.3
3	25/03/2003	10:22	12.5	11.45	9.98	16.4	16.2	15.1
4	10/04/2003	10:19	18.0	13.18	11.91	17.7	17.4	16.5
5	02/05/2003	10:28	1.3	15.25	13.71	20.6	20.2	18.9
6	21/05/2003	10:31	4.1	17.41	15.61	23.5	23.1	21.6
7	12/03/2004	10:28	1.4	11.18	10.05	15.2	15.0	14.3
8	03/04/2004	10:37	15.2	12.33	11.04	16.8	16.6	15.7
9	13/04/2004	10:22	12.4	12.63	11.39	17.0	16.8	15.9
10	06/05/2005	10:28	1.4	16.99	15.91	21.2	20.6	19.9
11	25/05/2005	10:31	4.1	19.16	17.72	24.4	23.8	22.7
12	14/03/2006	10:22	12.5	11.89	10.93	15.5	15.2	14.8
13	30/03/2006	10:19	18.2	13.45	12.13	18.1	17.8	16.9
14	18/04/2006	10:22	12.5	14.56	12.85	20.3	20.0	18.6
15	21/04/2006	10:28	1.5	14.64	12.95	20.3	20.0	18.6
16	26/05/2006	10:28	1.5	17.36	15.61	23.3	22.9	21.4
17	29/05/2006	10:34	9.6	19.01	17.51	24.4	23.8	22.6
18	05/03/2007	10:34	9.6	12.68	11.71	16.3	16.1	15.6
19	15/03/2007	10:19	18.0	12.16	10.98	16.4	16.1	15.4
20	18/03/2007	10:25	6.9	13.08	12.02	17.0	16.7	16.1
21	03/04/2007	10:22	12.6	11.97	10.44	17.1	16.9	15.7
22	08/05/2007	10:22	12.5	16.88	15.58	21.7	21.2	20.2
23	14/05/2007	10:34	9.6	17.14	15.81	22.0	21.5	20.5
24	30/05/2007	10:31	4.1	17.85	16.29	23.4	22.8	21.6
25	02/03/2008	10:25	7.0	13.28	12.28	17.0	16.7	16.2
26	05/03/2008	10:31	4.1	12.34	11.42	15.8	15.6	15.2
27	18/03/2008	10:22	12.5	12.74	11.66	16.7	16.4	15.8
28	21/03/2008	10:28	1.4	11.55	9.84	17.1	17.0	15.6
29	03/04/2008	10:19	18.0	13.92	12.80	18.1	17.7	17.0
30	25/04/2008	10:28	1.5	12.65	11.30	17.3	17.1	16.1
31	09/03/2003	21:40	16.7	8.94	8.70	10.5	10.3	11.0
32	19/03/2003	21:26	10.9	4.77	4.73	5.5	5.6	6.7
33	25/03/2003	21:37	11.1	1.65	0.82	4.3	4.7	4.4
34	03/03/2004	21:26	11.0	4.84	4.68	5.9	6.0	6.8
35	22/03/2004	21:29	5.4	6.67	6.19	8.7	8.7	9.0
36	13/04/2004	21:38	11.1	7.45	7.18	9.0	8.9	9.5
37	15/05/2004	21:32	0.1	12.54	11.49	16.4	16.1	15.5
38	07/03/2005	21:29	5.5	0.71	0.70	1.1	1.4	2.6
39	26/03/2005	21:32	0.1	11.90	11.44	14.2	13.8	14.2
40	12/03/2007	21:29	5.5	5.26	5.02	6.6	6.6	7.3
41	15/03/2007	21:35	5.5	6.53	6.42	7.6	7.5	8.5
42	16/04/2007	21:29	5.4	10.12	9.33	13.2	13.0	12.7
43	19/04/2007	21:35	5.6	11.42	10.64	14.5	14.3	14.0
44	08/05/2007	21:37	11.1	15.63	14.92	18.8	18.3	18.1
45	02/03/2008	21:40	16.6	11.46	11.37	12.7	12.4	13.4
46	18/03/2008	21:37	11.1	10.91	10.17	13.9	13.6	13.5
47	31/03/2008	21:29	5.5	11.09	10.65	13.3	13.0	13.3

Table 7. R-based validation cases over Albufera Lake in the Valencia test site. The AATSR-derived LSTs (T_{AATSR}) using different options (see text) are shown in the last three columns.

case	Date (d/m/y)	Time (UTC)	θ (º)	T_{11} (ºC)	T_{12} (ºC)	operational T_{AATSR} (ºC)	optimized T_{AATSR} (ºC)	Eq. (2) T_{AATSR} (ºC)
1	09/03/2003	10:25	7.5	11.91	10.91	15.6	13.2	13.3
2	12/03/2003	10:31	3.5	13.01	12.32	15.9	13.8	14.0
3	25/03/2003	10:22	13.0	12.04	10.62	16.9	13.9	14.0
4	10/04/2003	10:19	18.5	13.75	12.48	18.3	15.4	15.5
5	02/05/2003	10:28	1.9	16.05	14.55	21.3	18.0	18.2
6	21/05/2003	10:31	3.5	17.85	15.95	24.3	20.4	20.7
7	12/03/2004	10:28	1.9	9.45	7.67	15.1	11.9	12.1
8	03/04/2004	10:37	14.6	11.95	10.69	16.3	13.6	13.7
9	13/04/2004	10:22	12.9	12.50	11.22	17.0	14.2	14.3
10	06/05/2005	10:28	1.9	17.45	16.35	21.7	18.8	19.0
11	25/05/2005	10:31	3.5	18.82	17.44	23.9	20.6	20.8
12	14/03/2006	10:22	13.0	11.91	10.96	15.5	13.1	13.2
13	30/03/2006	10:19	18.7	13.64	12.29	18.4	15.5	15.5
14	18/04/2006	10:22	13.0	14.81	13.16	20.4	17.0	17.2
15	21/04/2006	10:28	2.0	14.83	13.02	20.8	17.2	17.5
16	26/05/2006	10:28	2.0	17.15	15.36	23.2	19.5	19.8
17	29/05/2006	10:34	9.0	19.09	17.52	24.7	21.1	21.4
18	05/03/2007	10:34	9.1	12.91	11.94	16.6	14.2	14.2
19	15/03/2007	10:19	18.6	12.33	11.14	16.6	13.9	14.0
20	18/03/2007	10:25	7.5	12.92	11.91	16.7	14.2	14.3
21	03/04/2007	10:22	13.1	12.04	10.51	17.1	14.1	14.2
22	08/05/2007	10:22	13.0	16.76	15.49	21.5	18.4	18.5
23	14/05/2007	10:34	9.0	16.54	15.36	21.0	18.0	18.2
24	30/05/2007	10:31	3.5	17.99	16.33	23.8	20.2	20.4
25	02/03/2008	10:25	7.5	13.49	12.53	17.2	14.7	14.8
26	05/03/2008	10:31	3.5	12.19	11.21	15.8	13.5	13.5
27	18/03/2008	10:22	13.1	12.89	11.82	16.8	14.3	14.4
28	21/03/2008	10:28	2.0	11.05	9.28	16.7	13.5	13.7
29	03/04/2008	10:19	18.6	14.32	13.23	18.4	15.7	15.8
30	25/04/2008	10:28	2.0	11.32	9.91	16.1	13.2	13.3
31	09/03/2003	21:40	16.5	8.11	7.94	9.4	8.2	8.5
32	19/03/2003	21:26	11.2	4.36	4.41	4.8	4.0	4.6
33	25/03/2003	21:37	10.9	6.44	5.81	8.8	7.1	7.3
34	03/03/2004	21:26	11.2	3.68	3.65	4.3	3.5	4.0
35	22/03/2004	21:29	5.7	5.78	5.37	7.6	6.1	6.4
36	13/04/2004	21:38	10.9	7.56	7.26	9.2	7.8	8.1
37	15/05/2004	21:32	0.1	11.33	10.53	14.5	12.4	12.5
38	07/03/2005	21:29	5.7	0.04	0.04	0.4	-0.3	0.3
39	26/03/2005	21:32	0.2	11.59	11.12	13.9	12.2	12.3
40	12/03/2007	21:29	5.7	4.56	4.43	5.5	4.5	5.0
41	15/03/2007	21:35	5.3	5.73	5.68	6.6	5.6	6.1
42	16/04/2007	21:29	5.7	9.57	8.87	12.3	10.4	10.6
43	19/04/2007	21:35	5.4	11.59	10.81	14.7	12.6	12.7
44	08/05/2007	21:37	10.9	14.39	13.84	17.0	15.2	15.2
45	02/03/2008	21:40	16.4	11.03	10.87	12.5	11.2	11.4
46	18/03/2008	21:37	10.9	10.27	9.57	13.1	11.1	11.3
47	31/03/2008	21:29	5.7	9.96	9.63	11.8	10.3	10.5

Table 8. Calculated R-based LST ($T_{R\text{-}b}$) and difference $\delta(T_{11}\text{-}T_{12})$ for each validation case in Table 6 (bare soil) and Table 7 (lake). W_{NCEP} is the precipitable water from NCEP profiles. Cases not meeting -0.6 K<$\delta(T_{11}\text{-}T_{12})$<0.6 K are marked with *****.

		bare soil		**lake**	
case	**W_{NCEP} (cm)**	**$T_{R\text{-}b}$ (ºC)**	**$\delta(T_{11}\text{-}T_{12})$ (K)**	**$T_{R\text{-}b}$ (ºC)**	**$\delta(T_{11}\text{-}T_{12})$ (K)**
1	0.98	14.4	0.47	13.1	0.13
2	0.84	15.5	0.29	13.9	0.09
3	1.23	14.7	0.54	13.6	0.24
4	1.46	16.8	0.13	15.7	-0.19
5	2.53	18.4	0.29	18.1	-0.13
6	1.83	21.1	0.37	19.9	0.26
7	1.68	13.8	0.40	10.1	1.05*****
8	1.15	15.6	0.29	13.3	0.16
9	0.81	15.9	0.32	13.9	0.22
10	1.16	20.2	0.18	18.9	-0.05
11	1.54	22.8	0.11	20.6	-0.14
12	1.09	14.8	0.26	13.2	0.10
13	2.71	17.5	-0.44	16.3	-0.67*****
14	1.77	18.5	0.29	17.1	-0.09
15	2.11	18.2	0.31	17.0	0.17
16	2.20	21.4	0.15	19.5	-0.02
17	3.34	23.3	-0.49	22.2	-0.67*****
18	1.30	15.9	0.12	14.5	-0.13
19	1.13	15.4	0.16	13.8	0.03
20	1.28	16.4	0.10	14.4	-0.13
21	1.33	15.5	0.32	13.9	0.17
22	1.57	20.1	0.30	18.2	0.06
23	1.37	20.6	0.23	18.1	-0.09
24	2.22	21.3	0.39	20.0	0.22
25	1.62	16.0	0.40	14.8	0.04
26	0.55	15.7	-0.15	13.5	-0.19
27	1.02	15.8	0.21	14.2	0.01
28	1.45	15.1	0.48	12.7	0.38
29	1.48	16.8	0.36	15.6	0.12
30	1.69	15.0	0.91*****	11.9	0.93*****
31	0.85	11.6	-0.20	8.9	-0.39
32	0.81	7.3	-0.32	5.1	-0.54
33	1.53	3.2	0.91*****	7.3	-0.07
34	0.53	7.5	-0.29	4.4	-0.47
35	0.89	9.6	-0.19	6.8	-0.36
36	0.93	10.2	-0.24	8.5	-0.41
37	1.60	15.9	0.01	12.8	-0.32
38	0.38	3.3	-0.40	0.8	-0.48
39	1.68	15.3	-0.63*****	13.3	-0.81*****
40	1.24	7.9	-0.26	5.4	-0.49
41	0.87	9.1	-0.31	6.5	-0.48
42	1.80	13.1	-0.05	10.9	-0.38
43	2.42	14.3	-0.29	13.1	-0.55
44	1.80	18.2	0.06	15.2	-0.17
45	1.59	13.5	-0.08	11.5	-0.18
46	1.15	13.9	-0.08	11.4	-0.26
47	1.06	14.3	-0.53	11.2	-0.73*****

Table 9. Statistics of the R-based validation ($\delta T=T_{AATSR}-T_{R\text{-}b}$) for the AATSR-derived LSTs using the three options (see text) over the bare soil and lake sites.

LST algorithm	operational		optimized		Eq. (2)	
site	bare soil	lake	bare soil	lake	bare soil	lake
bias (K)	0.6	2.2	0.3	-0.2	-0.2	0.0
std. dev. (K)	1.2	1.3	1.1	0.4	0.4	0.4
rmse (K)	1.3	2.5	1.1	0.5	0.4	0.4
min. δT (K)	-2.2	-0.4	-1.9	-1.1	-0.8	-0.5
max. δT (K)	2.4	4.4	2.0	0.7	0.6	0.9

FIGURE CAPTIONS

Figure 1. Mean emissivity (ε=(ε_{11}+ε_{12})/2) and emissivity difference ($\Delta\varepsilon$=ε_{11}-ε_{12}) in AATSR bands at 11 and 12 μm derived from Eq. (3) over East Spain on March and July, 2007. Clouds and water are masked in white.

Figure 2. (a) Photograph showing part of the Valencia test site in July. (b) RGB composite of ASTER bands 2 (0.66 μm), 3 (0.81 μm) and 1 (0.56 μm) showing the Valencia rice field area and environs on August 3, 2004. (c) RGB composite of AATSR bands at 0.87, 0.66 and 0.55 μm over Valencia on July 11, 2003.

Figure 3. AATSR-derived LST (T_{AATSR}) against ground-measured LST (T_g) for the three LST retrieval options. The dashed line is the 1:1 line.

Figure 4. Difference between AATSR-derived and ground-measured LST (T_{AATSR}-T_g) as a function of time.

Figure 5. Difference between ground-measured and R-based LSTs against δ(T_{11}-T_{12}) for the datasets and atmospheric profiles indicated. The linear regression for all data is shown.

Figure 6. Difference between R-based LSTs obtained from local radiosonde and NCEP/AIRS profiles against a) precipitable water difference, and b) difference in T_{11}-T_{12} simulated from local radiosonde and NCEP/AIRS profiles.

Figure 7. AATSR-derived LST minus R-based LST calculated with NCEP profiles against δ(T_{11}-T_{12}) in the bare soil and lake sites for a) the operational algorithm, b) the optimized algorithm, and c) the explicit emissivity-dependent algorithm.

Figure 8. AATSR-derived LST minus R-based LST calculated with NCEP profiles against the actual AATSR brightness temperature difference T_{11}-T_{12} in the bare soil site for the operational, the optimized and the explicit emissivity-dependent algorithm (Eq. 2). The solid lines are the linear regressions for each algorithm.

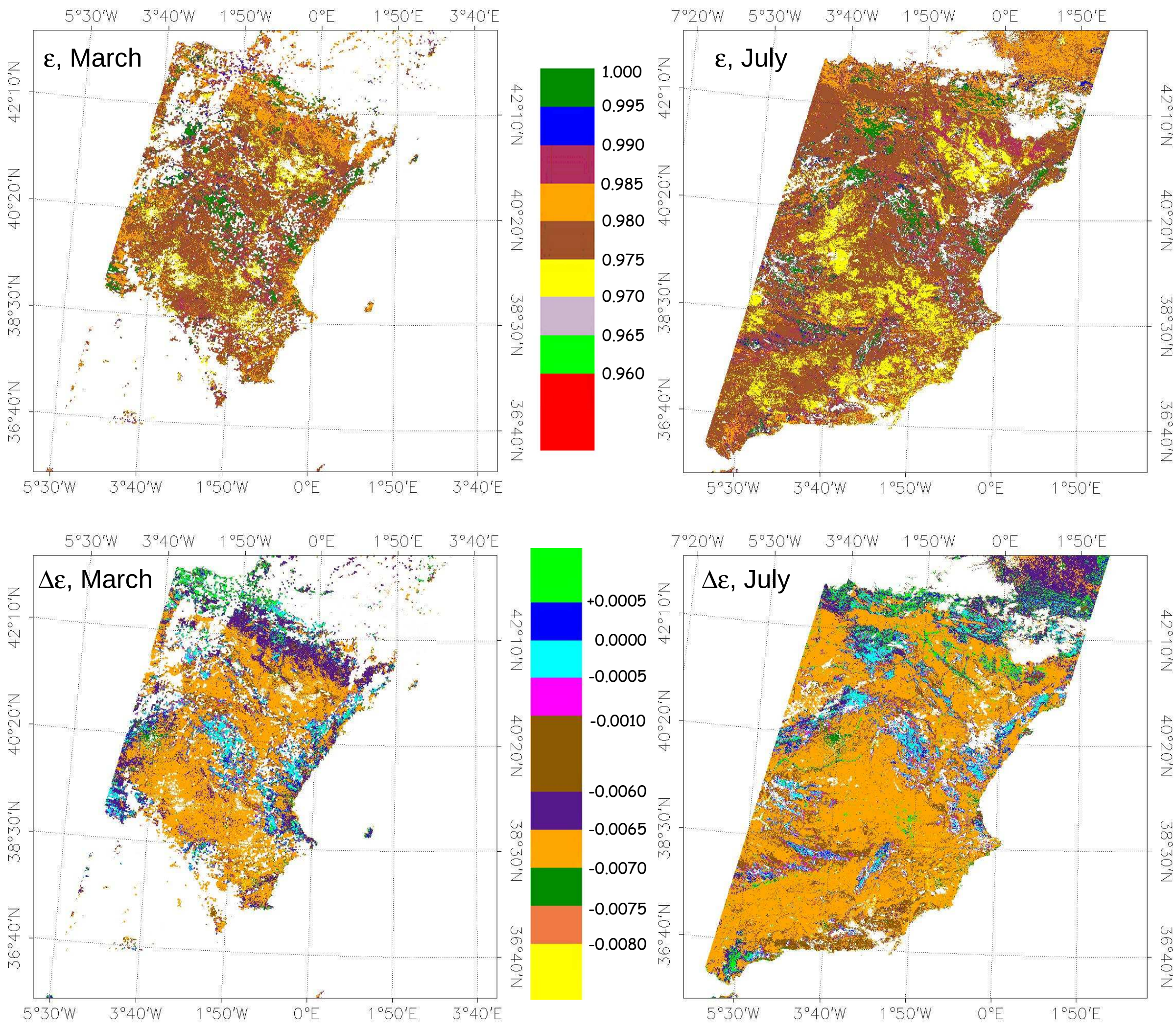
ε, March
ε, July
1.000
0.995
0.990
0.985
0.980
0.975
0.970
0.965
0.960
Δε, March
Δε, July
+0.0005
0.0000
-0.0005
-0.0010
-0.0060
-0.0065
-0.0070
-0.0075
-0.0080

Figure 1

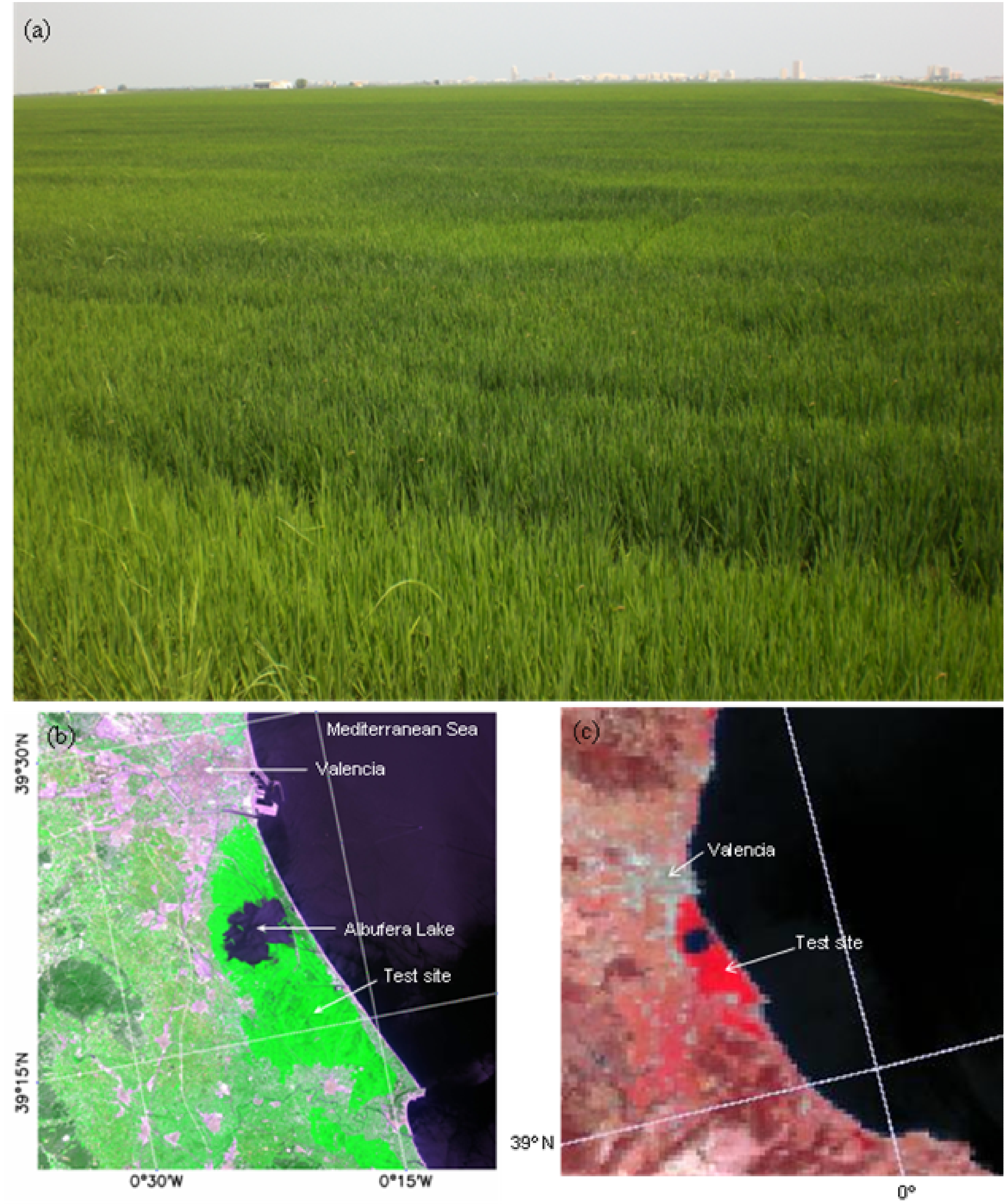
(a)
(b)
Mediterranean Sea
Valencia
Albufera Lake
Test site
39°30'N
39°15'N
0°30'W
0°15'W
(c)
Valencia
Test site
39° N
0°

Figure 2

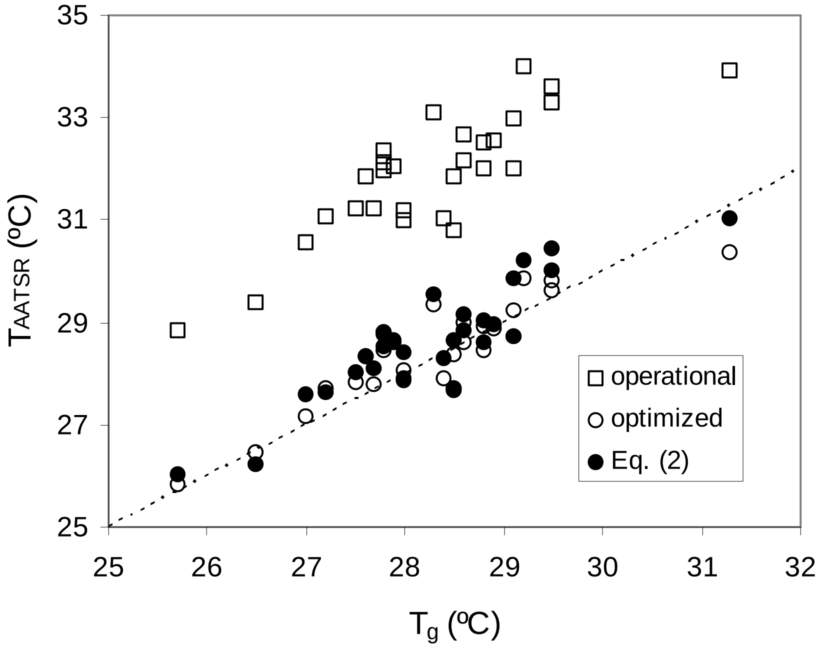
35
33
31
29
27
25
25
26
27
28
29
30
31
32
TAATSR (ºC)
Tg (ºC)
operational
optimized
Eq. (2)

Figure 3

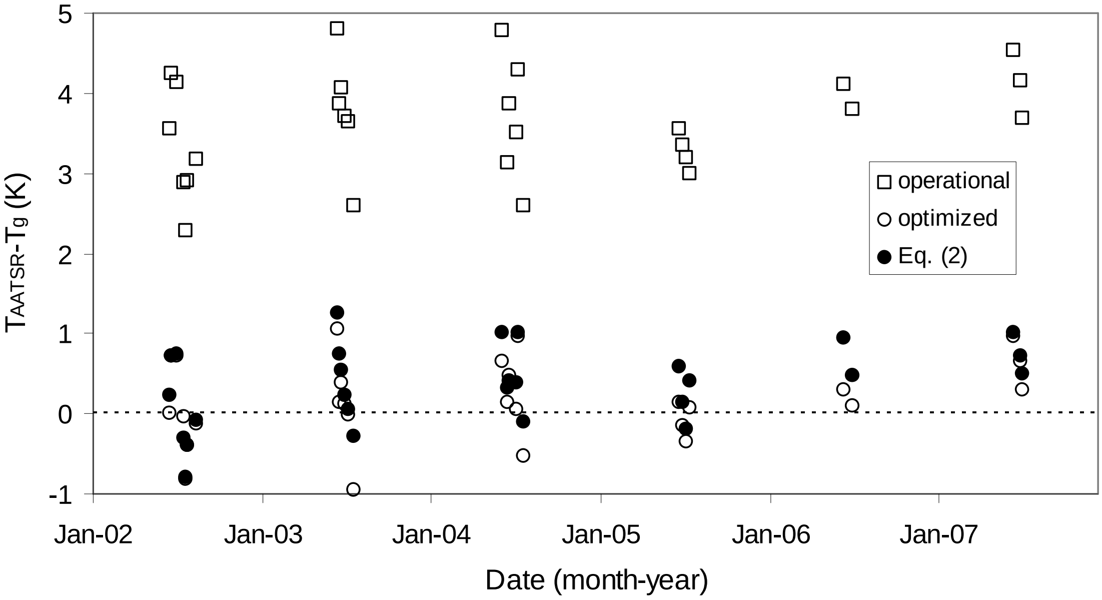
5
4
3
2
1
0
-1
TAATSR-Tg (K)
Jan-02
Jan-03
Jan-04
Jan-05
Jan-06
Jan-07
Date (month-year)
operational
optimized
Eq. (2)

Figure 4

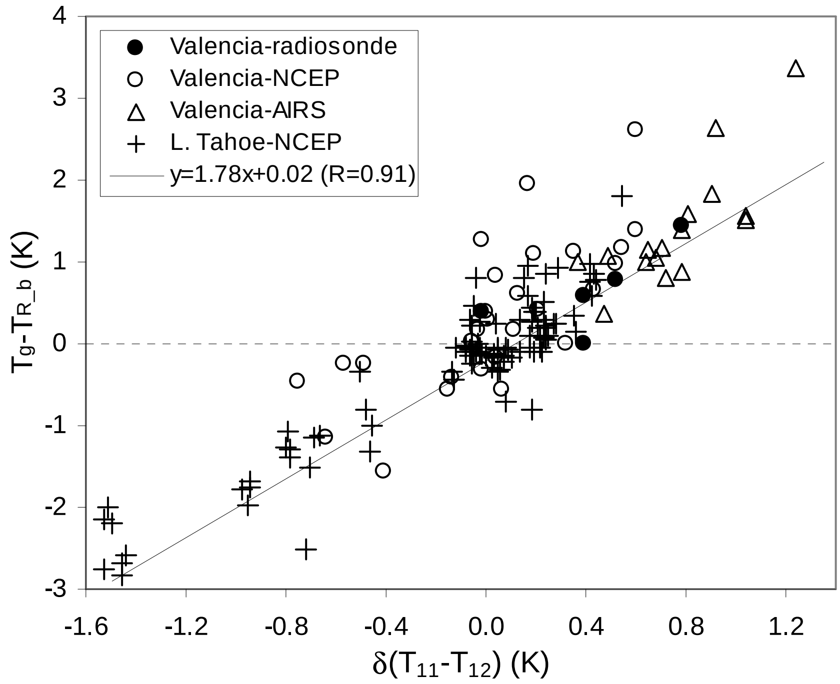
Valencia-radiosonde
Valencia-NCEP
Valencia-AIRS
L. Tahoe-NCEP
y=1.78x+0.02 (R=0.91)
Tg-TR_b (K)
δ(T11-T12) (K)
4
3
2
1
0
-1
-2
-3
-1.6
-1.2
-0.8
-0.4
0.0
0.4
0.8
1.2

Figure 5

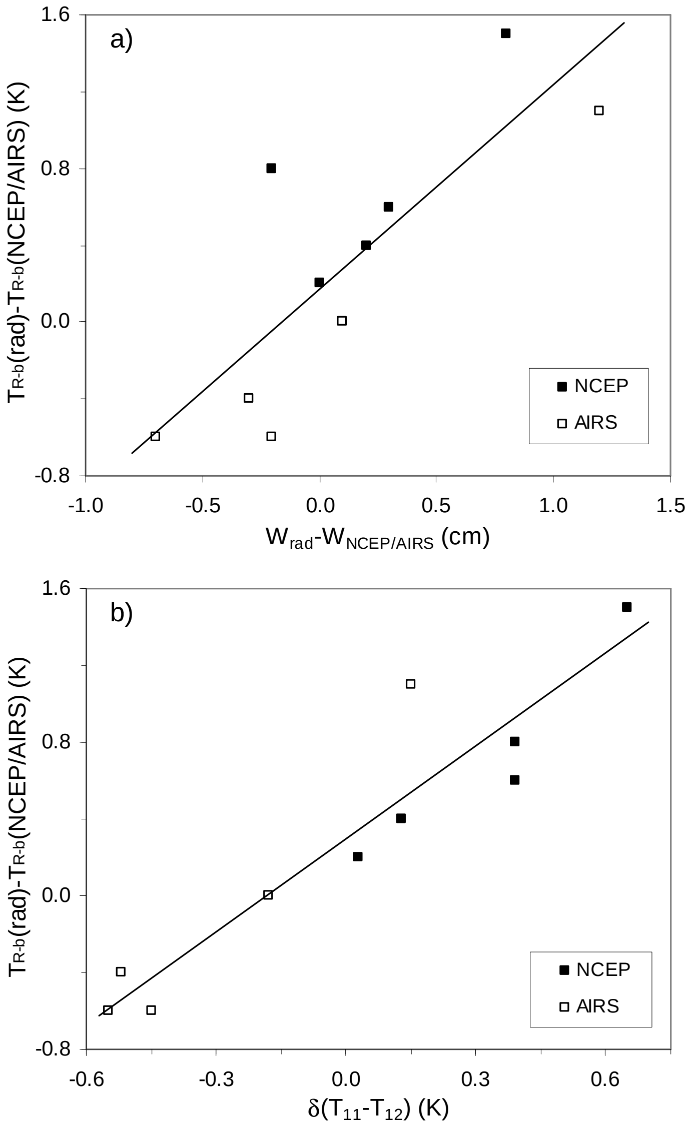
a)
1.6
0.8
0.0
-0.8
-1.0
-0.5
0.0
0.5
1.0
1.5
TR-b(rad)-TR-b(NCEP/AIRS) (K)
Wrad-WNCEP/AIRS (cm)
NCEP
AIRS
b)
1.6
0.8
0.0
-0.8
-0.6
-0.3
0.0
0.3
0.6
TR-b(rad)-TR-b(NCEP/AIRS) (K)
δ(T11-T12) (K)
NCEP
AIRS

Figure 6

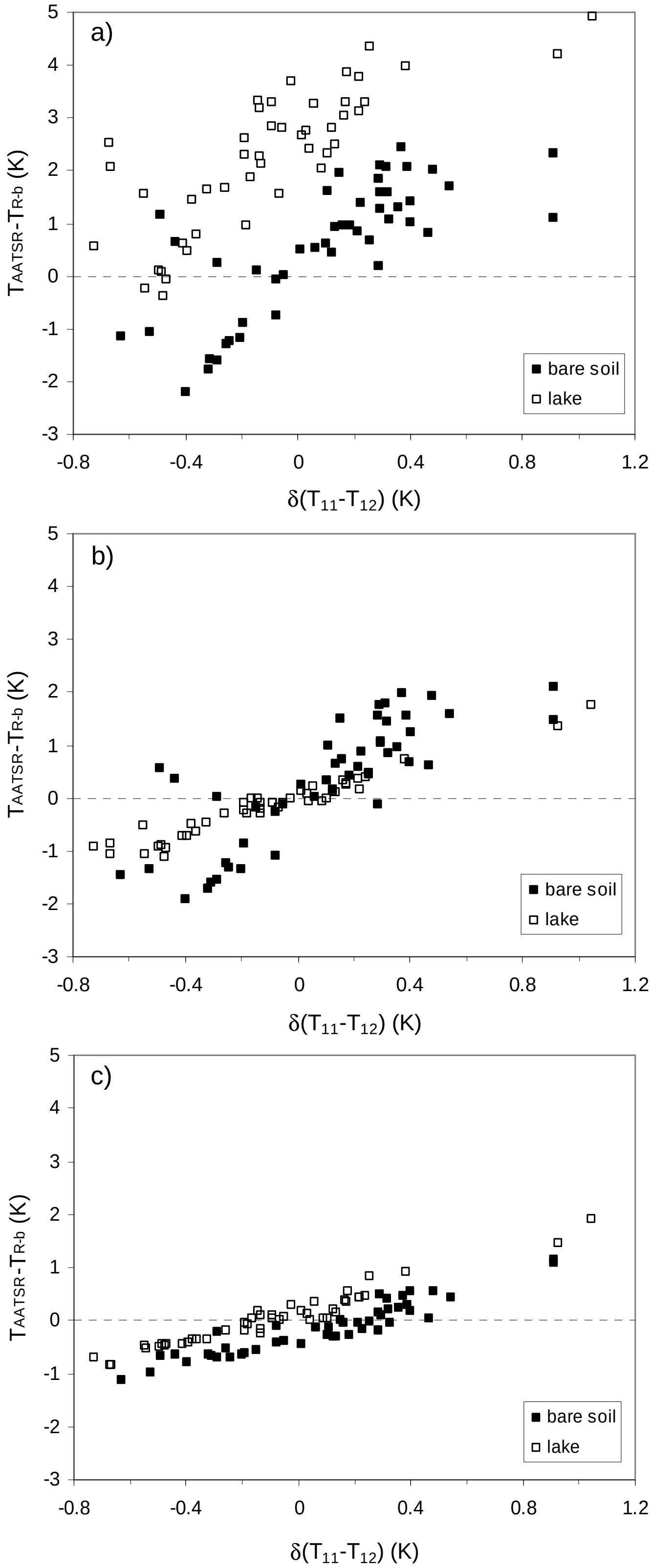
a)
b)
c)
TAATSR-TR-b (K)
δ(T11-T12) (K)
bare soil
lake
5
4
3
2
1
0
-1
-2
-3
-0.8
-0.4
0
0.4
0.8
1.2

Figure 7

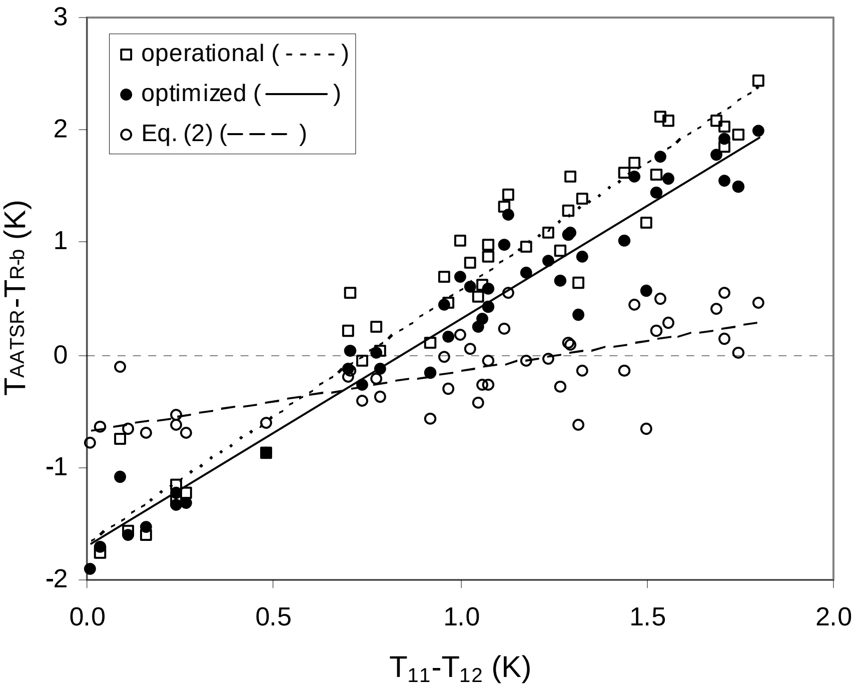
operational (- - - -)
optimized (———)
Eq. (2) (— — —)
T_{AATSR}-$T_{R\text{-}b}$ (K)
T_{11}-T_{12} (K)
3
2
1
0
-1
-2
0.0
0.5
1.0
1.5
2.0

Figure 8