Analysis of the performance of the TES algorithm over urban areas

Rosa Oltra-Carrio, Manuel Cubero-Castan, Xavier Briottet, José Sobrino

Analysis of the performance of the TES algorithm over urban areas

Rosa Oltra-Carrió, Manuel Cubero-Castan, Xavier Briottet, and José A. Sobrino

Abstract—The Temperature and Emissivity Separation (TES) algorithm is used to retrieve the Land Surface Emissivity (LSE) and Land Surface Temperature (LST) values from multispectral thermal-infrared sensors. In this work, we analyze the performance of this methodology over urban areas, which are characterized by a large number of different surface materials, a variability in the lowest layer of the atmospheric profiles and a 3D structure. These specificities induce errors in the LSE and LST retrieval, which should be quantified. With this aim, the efficiency of the TES algorithm over urban materials, the atmospheric correction and the impact of the 3D architecture of urban scenes are analyzed. The method is based on the use of a 3D radiative transfer tool, TITAN, for modeling all the radiative components of the signal registered by a sensor. From the sensor radiance, an atmosphere compensation process is applied, followed by a TES methodology that considers the observed scene to be a flat surface. Finally, the retrieved LSE and LST are compared with the original parameters. Results show that: first, the TES algorithm used reproduces the LSE (LST) of urban materials within an RMSE of 0.017 (0.9 K). Second, 20% of uncertainty in the water vapor content of the total atmosphere introduces an RMSE of 0.005 (0.4 K) for the LSE (LST) product. Third, in a standard case, the 3D structure of an urban canyon leads to an RMSE of 0.005 (0.2 K) for the LSE (LST) retrieval of the asphalt at the bottom of the scene.

Index Terms—TES, LST, LSE, urban, error budget.

I. INTRODUCTION

THE growth of a city leads to local climate modifications. These changes are in most cases related to urban temperature variations, such as the urban heat island effect. Thus, the retrieval of an accurate product of Land Surface Temperature (LST) and of Land Surface Emissivity (LSE) over urban areas is needed to properly monitor all the temperature dependent phenomena. Nevertheless, the singularities of urban areas (wide diversity of man-made materials, modification of the lowest layer of the atmospheric profiles and 3D surface) introduce into the LSE and LST products an error associated with both the response of the retrieval algorithm over urban materials and structure, and with the atmospheric correction performed over the imagery. In the paper, we quantify the performance of the Temperature and Emissivity Separation (TES) algorithm [1] over a simulated 3D urban canyon with a partially shaded road composed of man-made materials. Then, as the atmosphere profiles can vary within a urban area, the error introduced when using a unique profile will be considered. These errors are evaluated from synthetic images

R. Oltra-Carrió and J.A. Sobrino are with the Image Processing Laboratory, University of València, Paterna, E-46980 Spain e-mail: rosa.oltra@uv.es.

M. Cubero-Castan and X. Briottet are with the Optronic Department, ONERA Toulouse Centre, France.

and the 3D radiative transfer code, TITAN (Thermal Infrared radiance simulaTion with AggregatioN modelling) [3].

We base our study and results on the DESIREX (Dual-use European Security IR Experiment 2008) field campaign, which took place in Madrid (Spain) during Summer 2008 [2]. This field experience has been chosen for three main reasons. Firstly, the data base of the campaign includes atmospheric and temperature data, which is used to provide realistic simulations. Secondly, Madrid is a typical European city, so the materials and urban structures studied in this work can be found in other European cities where our results might also be extended. Thirdly, the TES algorithm was used to retrieve the LST and LSE maps over the city of Madrid from the Airbone Hyperspectral Scanner (AHS) imagery, so we have the opportunity to evaluate the performance of the algorithm already used in a real scenario.

In section II, the methodology of our work is presented by describing the TITAN code. Then, the urban scenes and the sensor characteristics used to simulate the scenarios are presented in section III. In section IV, we analyze and discuss the results from the three selected sources of error. Finally, conclusions and perspectives are given in section V.

II. METHODOLOGY

Fig. 1 shows an overview of the procedure followed for estimating the sensitivity of the TES algorithm. First, a geometrical scene and an atmospheric profile are introduced as input into the radiative transfer model (see section II-A) to compute the total radiance recorded at-sensor level by each thermal band of the sensor under study, i.e. the AHS (see section III for sensor specifications). Then, we start an inversion procedure: the atmospheric effects are corrected (see section II-B, Eq. 5) and the at-surface radiance and downwelling atmospheric radiance are used as input in the TES algorithm (see section II-C for further information about the method) to retrieve the surface temperature and emissivity. Finally, the TES LST and LSE outputs are compared to the reference values of LST and LSE, respectively.

In the following section, the TITAN radiative transfer model, the atmospheric correction in the TIR region and the TES methods are described.

A. Radiative transfer model

The TITAN model [3] estimates the main radiative components contributing to the total signal received by a sensor observing a rough and heterogeneous surface in the infrared domain from 3 μm to 14 μm. In this study, only the contribution of the TIR region (from 8 μm to 14 μm) is
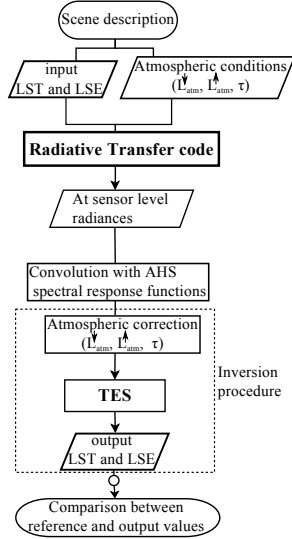
Fig. 1. Procedure overview, where $L_{atm}^{\downarrow}$ is the downwelling atmospheric radiance, $L_{atm}^{\uparrow}$ is the upwelling atmospheric radiance and τ is the upwelling atmosphere transmission.
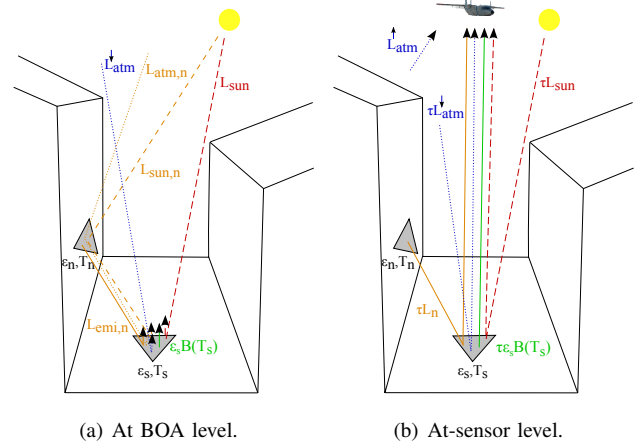

(a) At BOA level.   (b) At-sensor level.

Fig. 2. Radiances simulated by TITAN at BOA (a) and at-sensor level (b). Each parameter depends on the wavelength, it is not indicated to simplify the notation. Two facets are highlighted: the observed facet with temperature T_s and emissivity ε_s and the neighboring facet with temperature T_n and emissivity ε_n. Note that facets are homogeneous and isotropic and all computations are done at the center of each facet.

evaluated. The inputs required by TITAN are the 3D scene and the atmosphere. The synthetic scene was modeled by triangular facets. Each facet was considered as homogeneous and isotropic in terms of optical properties and temperature. TITAN uses the MODTRAN (Moderate Resolution Atmospheric Transmission) code [6] to estimate the atmosphere contribution. Therefore, when a flat surface is introduced to the model, only the MODTRAN code is required.

A schematic view of all the radiative terms simulated by TITAN is shown in Fig. 2. The model computes every component of incident irradiance as well as at bottom of atmosphere (BOA) and at-sensor level radiances for each facet and for all wavelengths, considering just one reflection inside the urban canyon. This last constraint does not imply a limitation of our work: around 2% of the TIR radiance outgoing from an asphalt facet placed in the middle of the urban canyon comes from the neighboring influence. If we consider a typical urban material with a reflectance of 5%, a second reflection will contribute just for 0.1% to the total radiance, which will not have any impact on the retrieval of the LSE and LST products.

In the following section, we describe each radiative component in the model; a full description is given in [3].

At ground level, TITAN simulates the incident irradiance components, which are then followed by the reflection components (BOA radiance): the downwelling atmospheric irradiance ($E_{atm,\lambda}^{\downarrow}$); the direct Sun irradiation at ground level ($E_{sun,\lambda}$); the irradiance incident to each facet from the neighboring facets ($E_{n,\lambda}$); the downwelling atmospheric BOA radiance $L_{atm,\lambda}^{\downarrow}$ (that is, after being reflected by the observed facet); the direct Sun radiance after the reflection by each facet ($L_{sun,\lambda}$); the radiance emitted by each facet at BOA level ($\varepsilon_{s,\lambda}B_\lambda(T_s)$, where $B_\lambda(T_s)$ is the Planck radiance at surface temperature T_s); the radiance arriving from the atmosphere that is reflected by the neighboring facets and then reflected by the observed facet ($L_{atm,n,\lambda}$); the direct Sun radiance reflected by the neighboring facets and then

reflected by the observed facet ($L_{sun,n,\lambda}$); the radiance emitted by the neighboring facets and reflected by the observed facet ($L_{emi,n,\lambda} = \varepsilon_{n,\lambda}B_\lambda(T_n)(1 - \varepsilon_{s,\lambda})$); the total radiance that arrives from the neighboring facets and is reflected by the observed facet ($L_{n,\lambda} = L_{atm,n,\lambda} + L_{sun,n,\lambda} + L_{emi,n,\lambda}$); and the total radiance at the BOA, after reflection and emission, that is the land-leaving radiance ($L_{g,\lambda}$), given by Eq. 1.

$$L_{g,\lambda} = \varepsilon_{s,\lambda}B_\lambda(T_s) + L_{atm,\lambda}^{\downarrow} + L_{sun,n,\lambda} + L_{n,\lambda}. \quad (1)$$

At sensor level, TITAN simulates the radiance observed by a sensor (L_λ) and its components: the atmospheric radiance that arrives at the sensor after being reflected by the facet ($\tau_\lambda L_{atm,\lambda}^{\downarrow}$); the Sun radiance that arrives at the sensor after being reflected by the facet ($\tau_\lambda L_{sun,\lambda}$); the radiance emitted by the facet that arrives at the sensor ($\tau_\lambda \varepsilon_{s,\lambda} B_\lambda(T_s)$); the radiance that comes from the neighboring facets, then it is reflected by the observed facet, and arrives to the sensor ($\tau_\lambda L_{n,\lambda}$). The τ_λ parameter is the atmospheric transmittance from the surface to the sensor and it is computed by MODTRAN, as well as the atmospheric transmittance from the top of the atmosphere to the surface ($\tau_\lambda^{\downarrow}$) and the upwelling atmosphere radiance ($L_{atm,\lambda}^{\uparrow}$). Finally, the at sensor radiance obtained from TITAN is given by Eq. 2.

$$L_\lambda = \tau_\lambda(\varepsilon_{s,\lambda}B_\lambda(T_s) + L_{atm,\lambda}^{\downarrow} + L_{sun,\lambda} + L_{n,\lambda}) + L_{atm,\lambda}^{\uparrow}. \quad (2)$$

B. Atmospheric correction

Thermal bands on remote sensing instruments work in an atmospheric window region of the electromagnetic spectrum, approximately between 8 and 14 μm. Nevertheless, an atmospheric window does not mean that the atmospheric effect is negligible and the thermal signals are sensitive to the water vapor content and temperature profiles. The radiant energy detected by thermal sensors is a composite of two sources:

the energy emitted and reflected by the land surface, which is
then transmitted through the atmosphere, and the energy that
is emitted by the atmosphere. Thus, supposing the surface as
flat, the radiative transfer equation for the TIR region is given
by Eq. 3.

$$L_i = \tau_i L_{g,i} + L_{atm,i}^{\uparrow}, \qquad (3)$$

where the index i indicates that each parameter is weighted
averaged with the spectral response of each i channel of the
sensor (f_i), as seen in Eq.4, for the radiance case.

$$L_i = \frac{\int f_i L_\lambda \, d\lambda}{\int f_i d\lambda}. \qquad (4)$$

The atmospheric correction leads to the estimation of the
land-leaving radiance:

$$L_{g,i} = \frac{L_i - L_{atm,i}^{\uparrow}}{\tau_i}. \qquad (5)$$

C. The TES algorithm

The TES algorithm [1] aims to estimate the LST and
LSE from Eq. 3 using the estimated land-leaving radiance
in Eq. 5 in every channel. The method is organized into
three modules. The NEM (Normalized Emissivity Method)
obtains preliminary values of LST and LSE. The second
module, RATIO, estimates the emissivity normalized spectrum
(β spectrum):

$$\beta_i = \frac{\varepsilon_{s,i}}{\frac{1}{N} \sum_{i=1}^{N} \varepsilon_{s,i}}, \qquad (6)$$

where N is the number of thermal channels considered.

Finally, the third module, the MMD (Maximum Minimum
Difference), transforms the β spectrum into the absolute LSE
and recalculates the LST by using an empirical relation be-
tween the minimum emissivity (ε_{min}) and the spectral contrast
(the MMD, defined by Eq. 7), which is the hypothesis that
solves the coupling problem.

$$MMD = max(\beta_i) - min(\beta_i). \qquad (7)$$

This function (Eq. 7) depends on the spectral band con-
ficutration of a sensor. In this work, the AHS sensor in the
DESIREX 2008 experimental field campaign was chosen.
As this sensor was the same used in the SEN2FLEX field
campaign [4] (see table I), the same empirical relationship
was considered. It is given by Eq. 8 and it was obtained
using 108 emissivity spectra from the ASTER spectral library
[5] (including soils, vegetation, water, ice and whole rock
chips. Note that no man-made materials were used to calibrate
this law). The statistical fit of this relation, with the spectra
considered, gave a correlation coefficient of 0.997 and a
standard deviation of 0.005. As mentioned previously, the
same equation has been applied over the urban area of Madrid
and the validation of TES LST results with in-situ data gave
a root mean square error (RMSE) of 1.4 K for night imagery,
that is, when no shadows are present [2], and of 0.039 for the
LSE product [7].

$$\varepsilon_{min} = 0.999 - 0.777 MMD^{0.815}. \qquad (8)$$

III. DATA SET DESCRIPTION

The input scenes of TITAN varied depending on the ana-
lyzed source of error; an accurate description of each is pro-
vided in section IV. The scenarios were selected to reproduce
the conditions of the DESIREX experimental campaign: the
surface materials were representative of the urban surfaces and
specifically of the city of Madrid, the atmospheric conditions
and LST values were chosen in accordance with the values
recorded during the campaign; and all the wavelength depen-
dent data were processed with the spectral configuration of
the AHS sensor. The AHS spectrometer, with a field of view
(FOV) of 90° and an instantaneous FOV of 2.5 $mrad$, contains
four types of detectors organized into five optical ports. Port 1
covers the VNIR (Visible and Near infrared) range. Its bands
are relatively broad (28-30 nm wide). The spectral coverage
is continuous from 443 nm up to 1025 nm. In the SWIR
(Short-Wave infrared) range, port 2A has an isolated band
centered at 1.6 μm and 90 nm wide. Next, port 2, has a set
of continuous, fairly narrow bands (17-18 nm wide) between
1907 nm and 2558 nm. Port 3, in the MWIR (Medium-Wave
Infrared) region, has 7 bands operating from 3.1 to 5.5 μm
with a mean bandwidth of 350 nm. Finally, the LWIR (Long-
Wave Infrared) photons, from 8.1 to 13.4 μm, are collected by
optical port 4, with a 10-element Mercury Cadmium Telluride
MCT detector array. Its average bandwidth is 480 nm and the
spectral arrangement of the AHS thermal bands, used in this
work, is given in Table I.

TABLE I
AHS THERMAL BANDS (PORT 4) USED IN THE DESIREX 2008
EXPERIMENTAL CAMPAIGN.

Band	Effective wavelength (μm)
71	8.18
72	8.66
73	9.15
74	9.60
75	10.07
76	10.59
77	11.18
78	11.78
79	12.35
80	12.93

IV. ERROR BUDGETS

In the case of the DESIREX campaign, the TES algorithm
was applied to the AHS imagery with satisfactory results
[2], [7]. Nevertheless, our instrument and methods are not
ideal, and we should do a more accurate sensitivity study by
evaluating the error sources that must be taken into account
in order to quantify the accuracy of the results. The first
source of error lies in the algorithm itself. We analyzed the
error committed when applying the TES algorithm with a
ε_{min}-MMD relationship that was not generated from man-
made materials. The next step was to verify atmospheric com-
pensation. Usually, remote sensing data are atmospherically
corrected using atmospheric profiles. Although soundings are
launched in synchronization with the sensor overpass, profile
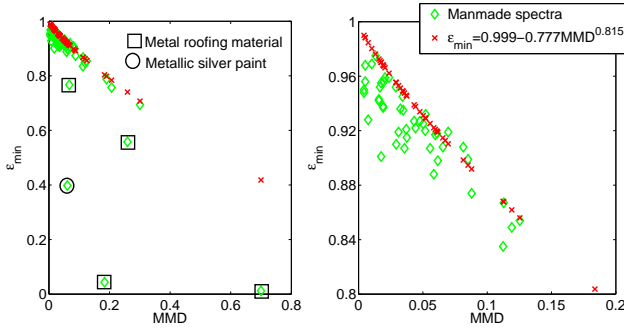measurements may not represent the atmosphere of the entire

Fig. 3. MMD relationship used in the DESIREX campaign which was assesed on 54 man-made materials from the ASTER library. Right plot is a detail of left plot.

image. For example, in the specific case of the DESIREX campaign, the atmospheric sounding was launched not inside the urban area but next to it. So we studied how small variations in water vapor content may affect the final LST and LSE maps. To conclude, we showed the uncertainty that was introduced due to the 3D structure of the city.

A. Impact of the TES algorithm: ε_{min} and MMD relationship

In this section, we estimate the performance of our MMD relationship (Eq. 8) over urban surfaces, mostly composed of constructed materials. From the ASTER spectral library, we selected 54 man-made spectra. This library contains directional-hemispherical reflectance spectra so emissivity spectra were deducted from Kirchhoff's law. Moreover, all the spectra were integrated into the spectral response of the AHS.

The ε_{min} and the MMD (see Eq. 7) were obtained for each spectrum. Then, from each MMD a new ε'_{min} value was retrieved using Eq. 8. Comparison of both ε_{min} values is given in Fig. 3. Apart from four materials for which ε_{min} was slightly underestimated by our MMD relationship (in ascending order of underestimation we had reddish asphalt, shingle roofing , terra cotta tile and brick), the ε_{min} was overestimated for all the materials. It can be observed that metal surfaces (identified by circles and squares) were badly modeled by our empirical relationship. This result has already been mentioned by [8] and [9].

Excluding metal materials and comparing both ε_{min} values plotted in Fig. 3, we determined that the man-made emissivity spectra can be recovered within a Root Mean Square Error (RMSE) of 0.025. Similar values are achieved in the literature, e.g. in [8] the TES accuracy was investigated with numerical simulations using ASTER spectral library (including 521 spectra of soils, minerals, vegetation, water and man-made materials excluding metals) and MODTRAN code. A multiband study was carried out in which all the spectra considered were interpolated to match the spectral response of the CIMEL 312-2 multiband radiometer, whose bands are very similar to ASTER bands, plus one broad band. The study showed that the original MMD relationship proposed by [1] ($\varepsilon_{min} = 0.994 - 0.687 MMD^{0.737}$) allows ε_{min} retrieval within an RMSE of 0.024, which is similar to our RMSE value.

In light of these results, for most urban materials, the use of our MMD relationship will lead to an overestimation of the output LSE. We have confirmed this hypothesis using the following approach.

1) Description of the scene: To generate a dataset of thermal infrared radiances, nine man-made materials representative of the urban surfaces were selected from the ASTER spectral library: brick, glass, tile, 2 types of asphalt, concrete, asphalt shingle, marble and cement. Four different LST values were assigned to each material, varying from 295 K to 310 K in steps of 5 K. The 36 targets (nine different materials at four LST) were established over a flat surface. Notice that a flat surface was selected to leave out the errors committed when a 3D structure is viewed by a sensor.

The atmospheric conditions of the scene introduced in the TITAN model were described by an atmospheric sounding acquired during the DESIREX campaign (total column of atmospheric water vapor content of $w = 1.59 g/cm^2$ and air temperature at the bottom of the atmosphere of 300.8 K). This profile provided information about pressure, temperature and humidity but did not include information about other atmospheric absorbers, such as CO_2 and O_3. Nevertheless, the MODTRAN code takes them into account and they were set to the values of the standard Mid-Latitude Summer Atmosphere. The Sun position was similar to its position during the AHS image acquisition: zenith angle $\theta_s = 20.11\,°$ and azimuth angle $\varphi_s = 144.5\,°$. The sensor was set up at the same altitude as the AHS, 1866 m over the surface, with a nadir viewing angle.

2) Results:

a) LSE analysis: Table II shows the statistical values from the comparison between the input LSE (LSE_{input}) and the output LSE (LSE_w hereafter), with $Bias = LSE_w - LSE_{input}$, $RMSE = \sqrt{bias^2 + \sigma^2}$ and σ the standard deviation. The spectral shapes are very similar (r^2 of 0.99) over all the targets except for the marble, whose r^2 decreases to 0.84. When all the targets were considered, we obtained an RMSE of 0.017, which indeed indicates the accuracy of the TES algorithm. This value is within the interval of error of the algorithm, considered to be 0.015 in terms of LSE retrieval [1]. The TES algorithm performed well for surfaces such as concrete and asphalt. Values for brick and tile surfaces were slightly underestimated (negative bias). For the rest of the materials, LSE_w values were overestimated. These results agreed with the conclusions extracted from the analysis of the ε_{min}-MMD relationship.

TABLE II
COMPARISON BETWEEN THE TITAN LSE_{input} AND THE LSE_w
RETRIEVED FROM TES

Material	r^2	Bias	σ	RMSE
Brick	0.99	-0.008	0.002	0.008
Asphalt 1	0.99	0.001	0.001	0.001
Glass	0.99	0.032	0.004	0.032
Tile	0.99	-0.004	0.001	0.005
Asphalt 2	0.99	0.001	0.001	0.001
Concrete	0.99	0.001	0.001	0.001
Asphalt Shingle	0.99	0.015	0.002	0.016
Marble	0.84	0.030	0.005	0.031
Cement	0.99	0.015	0.002	0.015
All targets	0.94	0.009	0.014	0.017

b) LST analysis: The statistical values from the comparison between the input LST (LST_{input}) and the output LST_w are shown in Table III ($Bias = LST_w - LST_{input}$). The RMSE obtained when considering all targets was lower than the accuracy of the algorithm (i. e. 1.5 K, [1]) and the LST was underestimated. Four material types (cement, asphalt shingle, marble and glass) had differences of around 0.8 K to 1.8 K between the input LST and the LST_w. For the other materials (brick, both asphalts, tile and concrete) both LST values were almost similar. As expected, LST_w was underestimated for the materials whose LSE_w was overestimated in the TES approach.

TABLE III
STATISTICAL VALUES FROM THE COMPARISON BETWEEN THE TITAN
LST_{input} AND THE LST_w RETRIEVED FROM TES.

Material	Bias (K)	σ (K)	RMSE (K)
Brick	0.3	0.0	0.3
Asphalt 1	-0.1	0.0	0.1
Glass	-1.8	0.1	1.8
Tile	0.1	0.0	0.1
Asphalt 2	-0.1	0.0	0.1
Concrete	-0.1	0.0	0.1
Asphalt Shingle	-1.0	0.1	1.0
Marble	-1.5	0.1	1.5
Cement	-0.8	0.1	0.8
All targets	-0.5	0.7	0.9

B. Impact of the uncertainty over the water vapor content

1) Description of the scene: The scene and dataset of TIR radiances used in section IV-A were also exploited in this case. In section II-B, we said that water vapor and temperature profiles are the primary atmospheric factors contributing to the thermal signal. Two different water vapor profiles were considered in the atmosphere compensation process to simulate the error in the estimation of the water vapor content. One with 20% of the total water content less than the original profile ($w = 1.27 g/cm^2$), and the second one with a 20% increase in total water content from the original profile ($w = 1.91 g/cm^2$). In the case of the temperature, different profiles were used to correct the atmosphere effects but we had not recorded any differences between the output LSE and LST products and the reference ones.

2) Results:

a) LSE analysis: Fig. 4 plots the output LSE values against the input LSE values. Results are shown for three different cases, first, for the ideal case, when the atmospheric conditions did not change during the entire procedure (LSE_w). Second, when the atmospheric water content increased 20% in comparison to the original atmospheric profile ($LSE_{+20\%w}$ hereafter). And third, when this value decreased by 20 % ($LSE_{-20\%w}$ hereafter). Each material is represented by a different symbol in order to differentiate each kind of behavior. We obtain a wide range of bias values, from -0.012 to 0.026 for the $LSE_{+20\%w}$ points, and from -0.005 to 0.038 for the $LSE_{-20\%w}$ values. Glass and marble have the highest bias values, while brick has the lowest. The materials whose bias were closest to 0 were: tile in the driest atmosphere and the two types of asphalts in the wettest. It should also be noted that
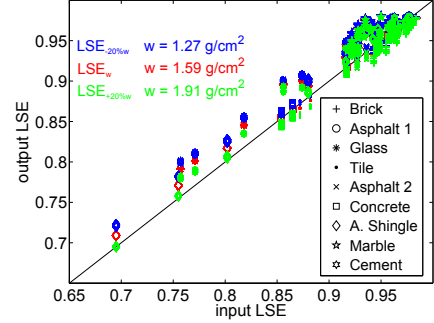

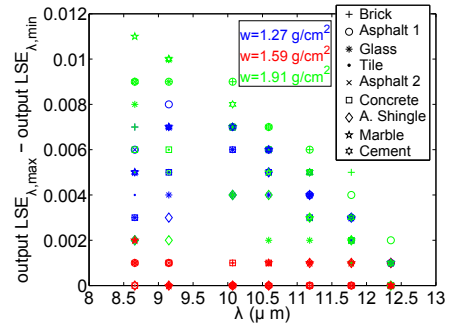Fig. 4. Comparison between the output LSE for different values of atmospheric water vapor content and the LSE_{input}.

Fig. 5. Difference between the output LSE maximum and the output LSE minimum for each effective wavelength of the AHS.

in the highest w case, the materials whose $LSE_{+20\%w}$ were underestimated were not just brick and tile, but also concrete and the two asphalts.

Notice from Fig. 4 that the same value of LSE_w was retrieved for all targets of the same material, even when the input LST changed. Nevertheless, this was not the general result when w was modified. For the same input of LSE and for some wavelengths, we obtained different values of $LSE_{+20\%w}$ which differed up to 0.01 for marble and cement, and up to 0.009 for brick and both asphalt surfaces. When the presence of the atmosphere was reduced, the variation of the $LSE_{-20\%w}$ retrieved for the same material was lower and went up to 0.007 for brick, both asphalts, marble and cement. Fig. 5 shows the difference between the maximum output LSE value retrieved and the minimum output LSE value retrieved (both for the same input LSE) versus the wavelength. It can be observed that higher differences, and therefore higher fluctuations, were obtained when water vapor increased, that is, when the atmospheric effects are over-corrected. Moreover, the uncertainty depended on the wavelength, being higher at lower λ.

In analyzing the trend of the retrieved LSE versus the input temperature, we observed that for all the materials, the $LSE_{+20\%w}$ values decreased when the LST_{input} increased, while for the $LSE_{-20\%w}$ retrievals, the opposite was true, higher values of LST_{input} resulted in higher values of $LSE_{-20\%w}$. In accordance with the data trend, we proposed a linear fit: $LSE_{\pm20\%w} = a \cdot LST_{input} + b$. Linear parameters and

the correlation coefficient for the asphalt material are shown in Table IV for each wavelength. We obtained high correlation coefficients, though in all cases, the slope (a) was small. Therefore, it must be concluded that, when the atmospheric compensation is performed with an uncertainty on the water vapor content, the error introduced in the LSE retrieval is related to the LST_{input} of the surface, but this dependence is not very strong.

Fig. 6. Comparison between the output LST for different values of atmospheric water vapor content and the LST_{input}.

Fig. 7. Comparison between the LST values retrieved with the original atmosphere (LST_w) and the LST retrieved with the other atmospheres considered ($LST_{\pm20\%w}$).

TABLE IV

LINEAR FIT PARAMETERS BETWEEN THE OUTPUT $LSE_{\pm20\%w}$ AND THE LST_{input} FOR THE ASPHALT 2 MATERIAL: $LSE_{\pm20\%w}=a\cdot LST_{input}+b$; r^2 INDICATES THE CORRELATION COEFFICIENT.

λ (μm)	$LSE_{-20\%w}$ ($w = 1.27g/cm^2$) r^2	a	b	$LSE_{+20\%w}$ ($w = 1.91g/cm^2$) r^2	a	b
8.66	0.95	$4\cdot10^{-4}$	0.819	0.98	$-6\cdot10^{-4}$	1.107
9.15	0.99	$5\cdot10^{-4}$	0.787	1.0	$-6\cdot10^{-4}$	1.102
10.07	1.00	$4\cdot10^{-4}$	0.833	0.99	$-5\cdot10^{-4}$	1.082
10.59	1.00	$4\cdot10^{-4}$	0.843	1.00	$-4\cdot10^{-4}$	1.076
11.18	0.98	$3\cdot10^{-4}$	0.883	0.98	$-3\cdot10^{-4}$	1.063
11.78	0.90	$1\cdot10^{-4}$	0.937	1.00	$-2\cdot10^{-4}$	1.028
12.35	0.8	$8\cdot10^{-5}$	0.953	0.8	$-8\cdot10^{-5}$	1.001

To isolate the error introduced by an inaccurate atmospheric compensation from the error of the TES algorithm, we compared the results to the LSE_w. Results are shown in Table V, in which the example of the asphalt material is also shown. An uncertainty of -20% on water vapor content led to an overestimation of the LSE with an RMSE of 0.005, while an uncertainty of +20% led to an underestimation with an RMSE of 0.006. It can be concluded that, when the at-sensor radiances are atmospherically corrected using an atmosphere with an overestimation or underestimation of 20% of the total water content, the output $LSE_{\pm20\%w}$ obtained differs around 0.5% from the output LSE_w retrieved with the correct atmospheric conditions.

TABLE V

COMPARISON BETWEEN THE LSE VALUES RETRIEVED WITH THE ORIGINAL ATMOSPHERE (LSE_w) AND THE LSE RETRIEVED WITH THE OTHER ATMOSPHERES CONSIDERED ($LSE_{\pm20\%w}$).

$w(g/cm^2)$	Material	Bias (K)	σ (K)	RMSE (K)
1.27	Asphalt 1 and 2	0.003	0.003	0.004
1.27	All targets	0.004	0.003	0.005
1.91	Asphalt 1 and 2	-0.004	0.003	0.005
1.91	All targets	-0.005	0.003	0.006

b) LST analysis: In Fig. 6 the LST_{input} used as input for each target on the flat surface and the LST obtained from the TES algorithm are compared. On the one hand, when the water content of the atmospheric profile was increased to 20 % ($w = 1.91\ g/cm^2$), the bias ranged from -1.6 K for the glass surface to 0.7 K for the brick. The output LST ($LST_{+20\%w}$ hereafter) values were overestimated for brick, asphalt, tile and concrete surfaces. On the other hand, when the water content was decreased a 20 % ($w = 1.27g/cm^2$) during the inversion procedure, all the output surface temperatures ($LST_{-20\%w}$ hereafter) were underestimated, except in the case of the brick material, which was coincident with the LST_{input} (bias=0 K). The highest bias value was once again achieved for the glass material (-2 K).

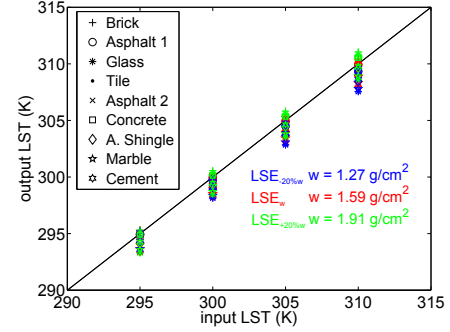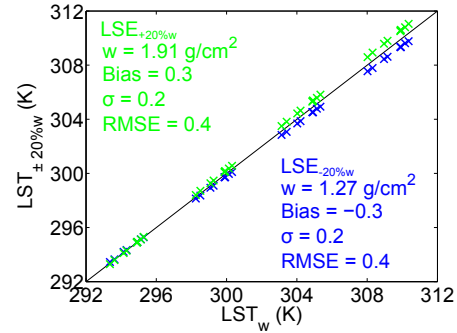Fig. 7 shows the comparison between $LST_{\pm20\%w}$ and LST_w

when all targets were taken into account, with bias defined by $LST_{\pm20\%w}$-LST_w. The plotted points were distributed into four groups, corresponding to the 4 different LST_{input}. The $LST_{+20\%w}$ values were slightly overestimated while the $LST_{-20\%w}$ were underestimated. In both cases, for the lowest LST_{input}, temperatures were almost coincident with LST_w values , while the difference between them grew to 0.6 K for the highest LST_{input}. The same bias (although with a different sign: $\pm$ 0.3 K), σ (0.2 K) and RMSE (0.4 K) were obtained in both cases. Moreover, these statistical values were the same as those achieved for both asphalt materials. In short, if we validate the LST retrieved when an uncertainty of 20% in the water vapor content is introduced, we obtain an RMSE of 0.4 K, which is in fact the error introduced by the atmospheric correction. Nevertheless, we have noticed that the error will depend on the LST_{input}.

C. Impact of the 3D geometry

The TES algorithm is usually applied when the observed image is considered to be a flat surface. However, a urban scene contains heterogeneous 3D structures mostly arranged following the structure of a urban canyon. This section analyzes the results of the influence of the urban structure on the retrieval of the LSE and the LST at the bottom of the urban structure.

1) Description of the scenes: A urban canyon is a basic urban surface that consists of walls and ground between two

adjacent buildings [10]. Four urban canyons (named as Canyon A, B, C and D) were modeled, and their structure can be seen in Fig. 8. The dimensions of each simulated canyon were 10 m wide (i. e. street width) and 40 m long and they ran north to south. The height of each building (i. e. the canyon height) varied depending on each canyon, see Table VI for each canyon geometry.

Fig. 8. 3D model of the urban canyons simulated. Measures are indicated in Table VI.

<div style="text-align:center">

TABLE VI
BUILDINGS HEIGHT FOR EACH CANYON.

</div>

	Canyon A	Canyon B	Canyon C	Canyon D
Left building height (H_l) (m)	15	22	30	50
Right building height (H_r) (m)	12	16	24	44

The materials used were the same for all scenes: asphalt for the street, red brick for the walls and red terracotta tiles for the roofs. The material spectra were extracted from the ASTER spectral library and the input LSTs are shown in Table VII. For the street, the values have been chosen according to the in situ measurements performed during the DESIREX campaign. For the walls LST (LST_{wall}) six different cases (Test 1 to Test 6) were selected in order to perform a sensitivity analysis. Test 1 was chosen to have the sunlit brick facets at the same temperature ($LST_{sunlit\ wall}$) as the air temperature (AT). Then, their temperature was increased 5 K for each test. For Test 2, the temperature of the wall-shaded facets ($LST_{shaded\ wall}$) and the $LST_{sunlit\ wall}$ were lower than the $LST_{shaded\ street}$ and the $LST_{sunlit\ street}$, respectively. $LST_{shaded\ wall}$ was equal to the AT and to the $LST_{shaded\ street}$ at Test 3. For Test 4, the $LST_{sunlit\ wall}$ and the $LST_{sunlit\ street}$ values were coincident. For Test 5, $LST_{shaded\ wall}$ and $LST_{sunlit\ wall}$ were higher than the $LST_{shaded\ street}$ and $LST_{sunlit\ street}$, respectively. Finally, for Test 6, both sunlit and shaded brick temperatures were higher than the $LST_{sunlit\ street}$.

Atmosphere, illumination and sensor conditions were set similar to those in previous sections. With these configurations, we studied the influence of the temperature of the wall facets over the thermal properties of the bottom of the canyon (asphalt), when the height of the buildings changed. In other words, when the sky view factor from the bottom of the canyon changed.

<div style="text-align:center">

TABLE VII
LST FOR EACH TEST.

</div>

		Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Wall (Brick)	Sunlit LST (K)	301	306	311	316	321	326
	Shaded LST (K)	291	296	301	306	311	316
Street (Asphalt)	Sunlit LST (K)	316					
	Shaded LST (K)	301					
Roof (Tile)	Sunlit LST (K)	311					

2) Results: In the following section, we describe the plots of RMSE and the bias as a function of the position of the facets inside each urban scene. The errors introduced due to the 3D geometry, without the influence of the error that the TES itself introduces to the LST and LSE for the asphalt material, are to be evaluated. To this end, RMSE and bias were obtained from the comparison between the LSE (LST) retrieved with the TES over the asphalt material on a flat surface (LSE_w and LST_w retrieved for the asphalt in section IV-A) and the output LSE (LST) when the 3D scene was analyzed. The output LSE and
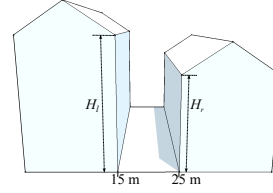
LST were retrieved from the average value of the at-sensor radiances which came from all the facets at the same distance from the walls. Moreover, we have focused the analysis on the facets placed in the middle of the canyon, assuming an infinite size in order to avoid borders.

a) LSE analysis: For all scenes, the main contribution to the RMSE was the bias from the sunlit facets, while for the shaded area, σ dominates at some points. Fig. 9 plots the RMSE and bias values for each canyon.

For Canyon A, due to the triangular mesh distribution, we did not have information on the behavior of positions closer than 1.5 m to the sunlit wall. *Far* from the wall (in the middle of the canyon), the RMSE went from approximately $1.8 \cdot 10^{-3}$ for Test 1 to $4.7 \cdot 10^{-3}$ for Test 6. When going from the sunlit facets to the shaded ones, the RMSE increased for Tests 1 and 2, while it decreased for Tests 3, 4, 5 and 6, and then it increased when approaching the shaded wall. Finally, next to the shaded wall, RMSE decreased with the increase of $LST_{shaded\ wall}$. Regarding the bias, for every test at the sunny area, we obtained an overestimation of the output LSE. Moreover, for Tests 1, 2 and 3 all the shaded facets were underestimated.

Regarding Canyon B, for all positions, the RMSE was the same for Tests 5 and 6 and for Tests 1 and 2. The RMSE decreased near the sunlit wall. In the middle of the canyon, the RMSE for Test 6 was around 50% greater than that for the Test 1. The shady values were lower than those in the sunlit area for tests 3, 4, 5 and 6. Again, near the shaded wall the RMSE decreased with the increase of LST_{wall}. Regarding the bias, it must be noted that the values were overestimated, except for the shaded points of Tests 1 and 2. The bias decreased near the walls and in the shady area.

The RMSE values for Canyon C were slightly higher than those for Canyon B although the observed behavior was similar. The error rose when going away from the directly irradiated wall and almost stabilized; the RMSE for Test 6 was almost 3 times higher than that for Test 1. Finally, it decreased at the shaded zone, except for Test 1. Once again, bias was lower in the shaded area than in the sunlit one, and it decreased next to both walls. All the values retrieved in the shaded area for Test 2 were underestimated.

Finally, we analyzed the values obtained for Canyon D. Results for Tests 4 and 5 are similar. The minimum values of RMSE were found next to both walls (apart from Test 1, when it increased next to the shady wall). For this canyon, only one asphalt facet was directly irradiated, and the RMSE obtained for the other ones were lower than the maximum RMSE of the Canyon C. However, if we compare both shaded areas, the

(a) RMSE Canyon A.

(b) Bias Canyon A.

(c) RMSE Canyon B.

(d) Bias Canyon B.

(e) RMSE Canyon C.

(f) Bias Canyon C.

(g) RMSE Canyon D.
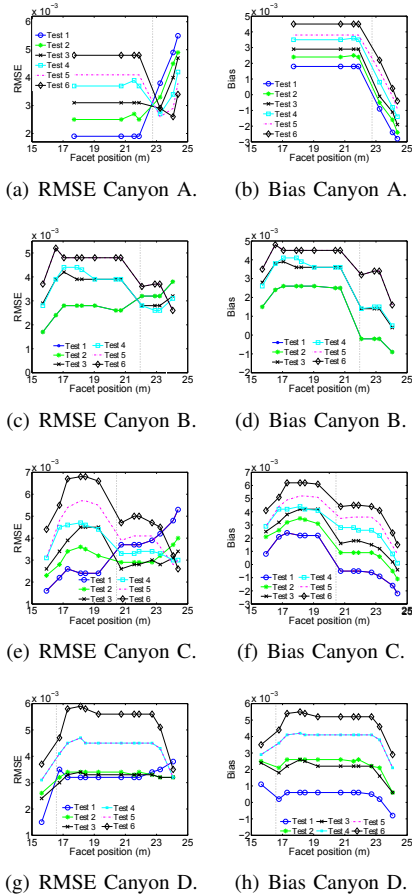
(h) Bias Canyon D.

Fig. 9. Spatial distribution across the urban canyon of the LSE RMSE and bias. The vertical dotted line indicates the separation between the sunlit zone (left) and the shaded zone (right).

(a) Canyon A.

(b) Canyon B.
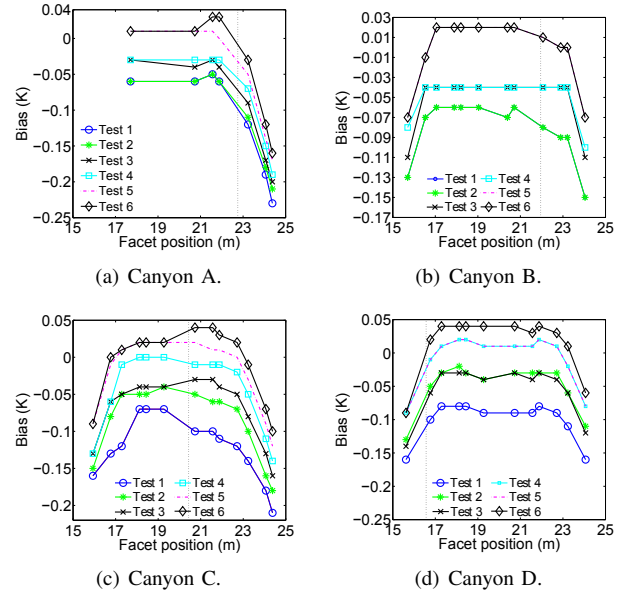
(c) Canyon C.

(d) Canyon D.

Fig. 10. Spatial distribution across the urban canyon of the LST bias. The vertical dotted line indicates the separation between the sunlit zone (left) and the shaded zone (right).

Canyon D RMSE was higher than the Canyon C error.

From the analysis of the retrieval of the LSE, seven main conclusions can be drawn: (1) the RMSE increased when the wall temperature was increased, except for the points closest to the shaded wall; (2) except for Canyon D, where only one facet is in the sunlit area, when going from the sunlit to the shaded facets, the RMSE increased for Test 1 (and also for Test 2 in Canyons A and B) while it decreased for Tests 3, 4, 5 and 6; (3) in the shaded area, the influence of the wall LST was lower than it was in the sunlit zone; (4) at the center of the canyon, we observed a tiny increase in the RMSE with canyon height; (5) even though the maximum RMSE, considering all tests, was always lower than 1%, these maximum values measured at the middle of the canyons were around 2 and 6 times (for Tests 1 and 6, respectively) the RMSE obtained for the asphalt when a flat surface is observed; (6) the mean bias between the output and the LSE_w values decreased next to both walls; (7) while our procedure overestimated the asphalt emissivity (see section IV-A), when a 3D scene is observed, the LSE of the shaded facets was underestimated for the tests with lower wall LST (Test 1 and in some cases, Test 2).

b) LST analysis: For the LST only the bias is presented, since from the averaged radiance only one value of LST is retrieved for each position (so the σ computation makes no sense).

The LST results for all canyons are presented in Fig. 10. Similar behavior was detected in all four cases. Values of LST were underestimated except for the central points (those points *far* from the walls) of Tests 5 and 6 (also Test 4 in the case of Canyon D), for which the bias was near zero and positive. The lowest values (highest in absolute value) of bias were registered near both walls. Moreover, differences between Test 6 and Test 1 were always around 0.1 K.

Finally, we can extract some conclusions from our LST analysis (statements are written in terms of the absolute value of the bias): (1) the bias rose when the wall LST decreases; (2) the bias plot across the street canyon was more or less symmetrical, with a maximum next to the walls and a minimum around the central facets; (3) the maximum value was around twice the error introduced by the algorithm for the asphalt; (4) the bias for Tests 5 and 6, that is when both wall temperatures are higher than the street ones, was always the lowest one and around zero; (5) in general terms, the retrieved LST at the bottom of a urban canyon was underestimated.

V. CONCLUSIONS

This paper aims to evaluate three types of errors usually encountered through an entire processing line of remote sensing images over a urban environment: TES algorithm error over urban materials, water vapor uncertainty and finally, error due to the three-dimensional structure of the urban environment. To carry out our study, different structures and situations were simulated with the 2D and 3D radiative transfer models (MODTRAN and TITAN, respectively). The atmospheric conditions, sun position, sensor altitude and spectral configuration were extracted from the DESIREX campaign, where the AHS sensor flew over the city of Madrid. The synthesis of the error budgets when different urban materials are considered is presented in table VIII.

TABLE VIII
SYNTHESIS OF THE ERROR BUDGETS.

Error source	LSE		LST (K)	
	RMSE	Bias	RMSE	Bias
TES Algorithm	0.017	0.009	0.9	-0.5
Atmospheric correction	0.005	±0.005	0.4	± 0.3

We have seen that the TES algorithm used for the data of the DESIREX campaign tends to overestimate the value of urban LSE materials, retrieving it within an RMSE of 1.7%. For LST, the error obtained is almost 1 K, with a tendency to underestimate the values. The uncertainty of the water vapor content during the atmospheric correction process introduces an error to the LSE and LST retrieval that depends on the ground truth temperature of the target and, in the case of the LSE, on the wavelength. An uncertainty of 20% of the water vapor content may introduce a general RMSE of 0.5% on LSE retrieval and of 0.4 K on LST. LSE values are over-estimated (underestimated) when the water vapor content is underestimated (overestimated). We find the opposite effect for the LST: an overestimation (underestimation) of the w leads to an overestimation (underestimation) of the LST product. The Root Sum Square (RSS) is defined as the square root of the sum of the squares and can be used to calculate the average accuracy of a measurement when the accuracies of all the sources of error are known and independent. In our case, the RSS of the TES algorithm and the atmospheric compensation was $\sqrt{0.017^2 + 0.005^2} \simeq 0.018$ for LSE and $\sqrt{0.9^2 + 0.4^2} \simeq 1$ K for LST. We see that the maximum contributor to the final error is the algorithm itself and its behavior over urban materials.

Finally, we have analyzed the problem of the retrieval of LST and LSE at the bottom of a urban canyon. We have seen that the error depends on the canyon geometry, on the position inside the canyon and also on the thermal properties of the vertical surfaces. The at-bottom material considered was asphalt and Table IX shows the estimation of the different errors obtained for this material. For the 3D structure, the error shown was approximately the highest one produced when Test 3 was analyzed, that is, when the most realistic temperatures were assessed; shaded walls were at the same LST as the AT and the sunlit wall LST was 10 K higher. Therefore, the RSS of the LSE retrieval at the bottom of a urban canyon is $\sqrt{0.001^2 + 0.005^2 + 0.005^2} \simeq 0.007$, and that of the LST retrieval is $\sqrt{0.1^2 + 0.4^2 + 0.2^2} \simeq 0.5$ K. For Test 3, the highest error for the LST is introduced by the wrong atmospheric correction, while the influence of the 3D structure and the atmospheric compensation is the same for the LSE case.

TABLE IX
SYNTHESIS OF THE ERROR BUDGETS FOR THE ASPHALT SURFACE.

Error source	LSE		LST (K)	
	RMSE	Bias	RMSE	Bias
TES Algorithm	0.001	0.001	0.1	-0.1
Atmospheric correction	0.005	±0.004	0.4	± 0.3
3D structure (Test 3, retaining one reflection)	0.005	0.005	0.2	-0.2

Thus, we can conclude that the TES algorithm performs well over urban areas (excluding some materials such as metallic roofs), though we have to point out the applicability and limitations of our study. Although simulations have been focused on the city of Madrid, the materials and 3D structures used have been chosen because they are representative of the majority of European cities. Nevertheless, further work has to be done to extrapolate our results to urban agglomerations with different building materials (such as wood or glass) and density; two extreme examples would be financial centers with skyscrapers and residential streets with green areas. The error introduced by the uncertainty of the water vapor content has to be taken as a reference value, but higher variations are expected if we study wetter atmosphere environments. The errors detected because of the effects of the neighboring facets in a 3D scene are linked to an increment of the down-welling irradiation at the bottom of the urban canyon, and this parameter would not change at different spatial resolutions, so similar results are expected for satellite thermal imagery at lower resolution. Lastly, we encourage TES users to apply the algorithm with a ε_{min}-MMD relationship calibrated with a data base of spectra representative of the surfaces seen by the sensor, which will improve the retrieval of LST and LSE parameters.

REFERENCES

[1] A. Gillespie, S. Rokugawa, T. Matsunaga, J. S. Cothern, S. Hook, and A. B. Kahle. A temperature and emissivity separation algorithm for Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images. *IEEE Transactions on Geoscience and Remote Sensing*, 36(4):1113–1126, 1998.

[2] J.A. Sobrino, R. Oltra-Carrió, G. Sòria, J.C. Jiménez-Muñoz, B. Franch, V. Hidalgo, C. Mattar, Y. Julien, J. Cuenca, M. Romaguera, J.A. Gomez, E. De Miguel, R. Bianchi, and M. Paganini. Evaluation of the surface urban heat island effect in the city of Madrid by thermal remote sensing. *International Journal of Remote Sensing*, 34:3177–3192, 2013.

[3] G. Fontanilles, X. Briottet, S. Fabre, and T. Tremas. Thermal infrared radiance simulation with aggregation modeling (TITAN): an infrared radiative transfer model for heterogeneous three-dimensional surface-application over urban areas. *Applied Optics*, 47(31):5799–5810, 2008.

[4] J.A. Sobrino, J.C. Jiménez-Muñoz, G. Sòria, Gómez M., A. Barella-Ortiz, M. Romaguera, M. Zaragoza, Y. Julien, J. Cuenca, M. Atitar, V. Hidalgo, B. Franch, C. Mattar, A. Ruescas, L. Morales, A. Gillespie, L. Balick, Z. Su, F. Nerry, L. Peres, and R. Lionati. Thermal remote sensing in the framework of the SEN2FLEX project: field measurements, airborne data and applications. *International Journal of Remote Sensing*, 29(17-18):4961–4991, 2008.

[5] A. M. Baldridge, S. J. Hook, C. I. Grove, and G. Rivera. The ASTER spectral library version 2.0. *Remote Sensing of Environment*, 113(4): 711–715, 2009.

[6] A. Berk, G. P. Anderson, L. S. Bernstein, P. K. Acharya, H. Dothe, M. W. Matthew, S. M. Adler-Golden, J. H. Chetwynd, S. C. Richtsmeier, B. Pukall, C. L. Allred, L. S. Jeong, and M. L. Hoke. MODTRAN4 radiative transfer modeling for atmospheric correction. *Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research III*, 3756:348–353, 1999.

[7] R. Oltra-Carrió, J. A. Sobrino, B. Franch, and F. Nerry. Land surface emissivity retrieval from airborne sensor over urban areas. *Remote Sensing of Environment*, 123(0):298–305, 2012.

[8] V. Payan and A. Royer. Analysis of Temperature Emissivity Separation (TES) algorithm applicability and sensitivity. *International Journal of Remote Sensing*, 25(1):15–37, 2004.

[9] J.A. Sobrino, R. Oltra-Carrió, J.C. Jiménez-Muñoz, Y. Julien, G. Sòria, B. Franch, and C. Mattar. Emissivity mapping over urban areas using a classification-based approach: Application to the Dual-use European Security IR Experiment (DESIREX). *International Journal of Applied Earth Observation and Geoinformation*, 18:141–147, 2012.

[10] M. Nunez and T. R. Oke. The energy balance of an urban canyon. *Journal of Applied Meteorology*, 16(1):11–19, 1977.

José A. Sobrino is Professor of Physics and Remote Sensing, President of the Spanish Association of Remote Sensing and head of the Global Change Unit at the University of Valencia, Spain. Author of more than 150 papers. Coordinator of the European projects WATERMED and EAGLE. He has been a member since November of 2003 of ESAC (Earth Science Advisory Commitee) of ESA (European Space Agency). Chairperson of the series of International Symposiums RAQRS (Recent Advances in Quantitative Remote Sensing). His research interest include atmospheric correction in visible and infrared domains, the retrieval of emissivity and surface temperature from satellite images, and the development of remote sensing methods for land cover dynamic monitoring.

Rosa Oltra-Carrió received her Degree in Physics from the Universitat de València in 2007 and the PhD degree in Remote Sensing from the Universitat de València in 2013. During her PhD she was interested on the study of the urban environments from thermal remote sensing data. Her research activities currently concern hyperspectral imagery and soil properties characterization.
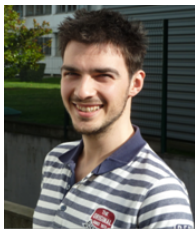
Manuel Cubero-Castan received the M.Sc. degree in electrical and signal processing engineering from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 2011. In collaboration between ONERA, the French Aerospace Lab in Toulouse, GIPSA-Lab in Grenoble and RMA in Brussels, he is currently pursuing a Ph.D. degree from Grenoble Institute of Technology on unmixing models applied on thermal images. His current research interests also include temperature/emissivity separation and source separation methods in thermal remote sensing.

Xavier Briottet received his PhD degree in Electonics from the Ecole Nationale Supérieure de l'Aéronautique et de l'Espace (Toulouse, France) in 1986 and his "Habilitation à Diriger les Recherches" in 1999 from Université Paul Sabatier. In 1987, he became an assistant professor in signal and image processing at the Ecole Supérieure d'Electricité. Since 1988, he is working in remote sensing at ONERA in the optic department (Toulouse, France). He received his Habilitation à Diriger les Recherches from Paul Sabatier University (Toulouse, France) in 1999. He is now Director of Research and the advisor of the scientific politic at DOTA. His current research interests are on 3D radiative transfer in urban area, shadow detection, atmosphere, unmixing in the reflective and thermal domain, using high spatial hyperspectral imagery and lidar. He is one of the 4 PI of HYPXIM hyperspectral space mission.