Linear-quadratic mixing model for reflectances in urban environments

Inès Meganem[1,2], Philippe Déliot[1], Xavier Briottet[1], Yannick Deville[2] and Shahram Hosseini[2]

***Abstract*—In the field of remote sensing, the unmixing of hyperspectral images is usually based on the use of a mixing model. Most existing spectral unmixing methods, used in the reflective range [0.4-2.5 μm], rely on a linear model of endmember reflectances. Nevertheless, such a model supposes the pixels at ground level to be uniformly irradiated and the scene to be flat. When considering a 3D landscape, such a model is no longer valid as irradiated and shadowed areas are present as well as radiative interactions between facing surfaces. This paper introduces a new mixing model, adapted to urban environments, and which aims to overcome these limitations. This model is derived from physical equations based on radiative transfer theory and its analytic expression is linear-quadratic. Similar models have already been used in the literature for unmixing purpose but without being justified by physical analysis. Our proposed model is validated using a synthetic but realistic European 3D urban scene. Then, simplifications are introduced, based on a study of the different radiative components contributing to the signal in a way to make the model easy to use for spectral unmixing. This paper also shows that the quadratic term cannot be neglected in many cases in urban environments, since it can e.g. range from 15% to 20% of reflectances in canyons.**

***Index Terms*—urban images, linear-quadratic mixing model, reflectances, physical modelling, spectral unmixing.**

I. Introduction

BECAUSE of low spatial resolution of hyperspectral image sensors compared to panchromatic ones, a pixel rarely represents a homogeneous surface and the signal incident to a sensor into its IFOV (instantaneous field of view) often results from contributions of several materials. For example, in [1] authors found that in HyMap images acquired over Dresden, Germany, 52% of pixels comprise spectral mixtures. In case of mixed pixels it is possible to retrieve the constituent components of a given pixel using spectral mixture analysis techniques owing to the spectral richness of the data. This process of sub-pixel retrieval is called spectral unmixing.

Our work is focused on the reflective domain (wavelengths from 0.4 to 2.5 μm). Most available unmixing methods in this domain rely on linear mixing models for reflectances (see e.g. [2]–[9] and references therein), where the mixing coefficients are the abundances. Such an assumption is valid when the scene is flat with a homogeneous incident irradiance.

In this paper, we aim at deriving a mixing model of reflectances adapted to urban images. This involves a more complicated model, because towns are characterized by a high spatial variability at a meter scale which implies different illumination levels (sunny and shadowed areas). Besides, the 3D structure of such environments induces multiple scattering of light between surfaces (see Figure 1). As a consequence, the reflectance of one pixel can contain the contribution of a material reflectance from its neighbourhood, even if the surface represented by the pixel does not contain this material. Then, the linear model is no longer valid, since this model assumes that only the content of the pixel contributes to its reflectance.

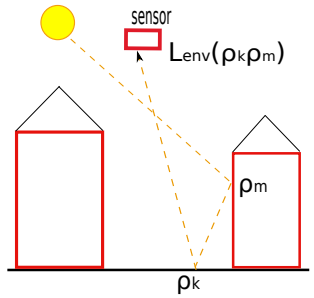


Fig. 1. Contribution of the reflections due to the neighbourhood, denoting ρ_i the reflectance of surface i (where i can be k or m) and L_{env} the radiance due to these reflections.

Taking into account multiple reflections of light between surfaces yields a non-linear model. Nonlinearity in unmixing has been considered in previous works (e.g. [10]–[13]) where the proposed methods essentially make use of neural networks. However, these methods are presented without an explicit mixing model. These papers show, though, that considering possible non-linearities in the model improves unmixing performance. In [14], experimental results obtained with an image acquired in a laboratory (from a synthetic scene) showed that a second-order model better describes the mixtures than a linear one, in the case of 3D structures like in urban environments, since it gives more accurate unmixing results.

Starting from physical equations based on radiative transfer theory, we derive a new physical based mixing model that deals with the non-linearities faced in urban environments. The derived linear-quadratic model only takes into account second-order reflections, neglecting higher-order interactions, which is very realistic given the range of values for reflectances (values between 0 and 1). Similar models were used in some unmixing approaches in the literature, for different types of scenes, but without any physical justification: orchards [15], [16], forest [17] and scenes composed of vegetation, water and soil materials [18]. The assumptions made on the mixing coefficients are not either physically justified in those works. In this paper we propose a justification and validation of

[1] ONERA, The French Aerospace Lab, 31055 Toulouse, France.
[2] IRAP, CNRS, Université de Toulouse, UPS-OMP, 31400 Toulouse, France.

the linear-quadratic model in presence of 3D structures. A substantiation of the possible assumptions concerning the mixing coefficients is also given. This work thus aims at giving a realistic model that could be used for unmixing methods.

In Section II, the method used to derive our physical mixing model for urban scenes is explained. This model is then validated in Section III and the relative impact of radiative terms is evaluated, based on synthetic images. In Section IV, simplifications of this model, using reasonable assumptions, are presented. This yields an invariant instantaneous linear-quadratic mixing model. Eventually, in Section V, we finalize this model to make it more convenient for spectral unmixing. We also discuss the possible assumptions on the mixing coefficients, compared to what exists in the literature.

II. Physical mixing model

In this section, we show how our physical mixing model is derived, starting from physical equations based on radiative transfer theory. The development of this model followed the same procedure as in [19], developed in the thermal infrared domain. The approach, originally proposed by [20], consists in comparing the sensor radiance at a coarse resolution with all the radiance coming from the different elementary surfaces inside the coarse pixel.

A. Presentation of the method leading to the model

At a coarse resolution, a pixel often corresponds to a heterogeneous surface composed of different elementary homogeneous surfaces. Thus, for each pixel, the sensor receives contributions resulting from reflectances of all these elementary surfaces. This is illustrated in Figure 2(a). However, the signal received by the sensor from this heterogeneous surface can be considered as the resulting signal coming from an equivalent horizontal homogeneous flat surface. We thus assume that, at a coarse resolution, each pixel of the image can be associated with such a flat equivalent surface at the height of the urban canopy. Therefore, we attribute an equivalent reflectance $\langle\rho\rangle$ (corresponding to the flat equivalent surface) to this pixel (see Figure 2(b)). Then, to derive the equation expressing the equivalent reflectance of a pixel as a function of contributions of elementary surface reflectances, the following method is applied:

- Express the total radiance of a given pixel as the result of the contributions of the different materials composing it at the ground level (Section B).
- Express the radiance of a pixel considering its associated equivalent surface (Section C).
- Perform a term by term radiative identification of the two preceding expressions to derive the equivalent reflectance expression for the pixel (Section D).

B. Radiance from the ground level at a fine scale

Assuming each pixel is composed of N homogeneous elementary surfaces indexed by k, the total radiance L_{mix} received by the sensor, in the field of view delimited by the solid angle Ω corresponding to a given pixel, can be expressed as

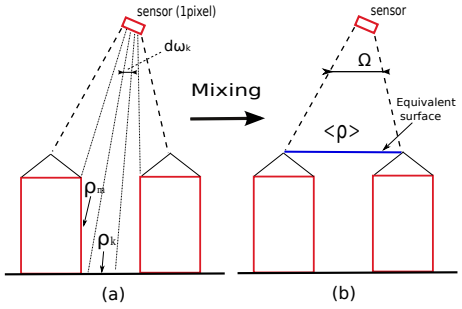


Fig. 2. Mixing principle - the dashed lines define the coarse IFOV Ω and the continuous lines the fine IFOV $d\omega_k$.

$$\Omega L_{mix} = \sum_k L_{D,k} d\omega_k + \sum_k L^{\downarrow}_{atm,k} d\omega_k + \sum_k L_{env,k} d\omega_k + \sum_k L_{coupling,k} d\omega_k + \sum_k L^{\uparrow}_{atm,k} d\omega_k + \sum_k L^{\uparrow}_{diff,k} d\omega_k \quad (1)$$

with
- $d\omega_k$: solid angle corresponding to an elementary surface k.
- $L_{D,k}$: solar radiance due to the solar irradiance on the surface k directly reflected toward the sensor.
- $L^{\downarrow}_{atm,k}$: atmospheric downwelling radiance incident on the surface k and directly transmitted toward the sensor.
- $L_{env,k}$: radiance due to the neighbourhood of the considered surface k.
- $L_{coupling,k}$: radiance due to multiple scatterings between the atmosphere and the ground.
- $L^{\uparrow}_{diff,k}$: radiance due to the total downwelling irradiance, incident on the neighbourhood of the surface k, reflected toward the sensor direction.
- $L^{\uparrow}_{atm,k}$: upwelling atmospheric radiance, reaching the sensor in the solid angle $d\omega_k$.

By replacing the radiances of interest by their expressions given in [21], this yields

$$L_{mix} = \sum_k S_k t^{\uparrow k-c} \frac{\rho^{dd}_k(\overrightarrow{U_s}, \overrightarrow{U_{c,k}})}{\pi} E_{D,k} + \sum_k S_k L^{\downarrow}_{atm,k} + \sum_k S_k t^{\uparrow k-c} \iint_{m\in V_k} \frac{\rho^{dd}_k(\overrightarrow{U_{km}}, \overrightarrow{U_{c,k}})}{\pi} g_{m,k} \left[\frac{\rho^{dd}_m(\overrightarrow{U_s}, \overrightarrow{U_{mk}})}{\pi} E_{D,m} + \frac{\rho^{hd}_m(\overrightarrow{U_{mk}})}{\pi} E_{diff,m} \right] dS_m + \sum_k S_k L_{coupling,k} + \sum_k S_k L^{\uparrow}_{atm,k} + \sum_k S_k L^{\uparrow}_{diff,k} \quad (2)$$

where
- $S_k = \frac{d\omega_k}{\Omega}$: normalized elementary solid angle over which the sensor sees an elementary surface k (with $\sum_k S_k = 1$).
- $t^{\uparrow k-c}$: atmospheric transmission from surface k up to the sensor.
- $\overrightarrow{U_s}$: normalized vector of sun's direction at surface k.
- $\overrightarrow{U_{c,k}}$: normalized vector of sensor's direction with respect to surface k.
- $\overrightarrow{U_{km}}$: normalized vector defining the direction between

surface k and surface m (vector linking the centres of these surfaces).
• ρ_k^{dd} : bidirectional reflectance of elementary surface k, the directions being defined by $\overrightarrow{U_s}$ and $\overrightarrow{U_{c,k}}$ (unitless).
• ρ_k^{hd} : hemispheric directional reflectance of surface k, depending only on the observation direction (unitless).
• $E_{D,k}$: direct solar irradiance on surface k.
• $E_{diff,m}$: atmospheric irradiance on surface m.
• V_k : neighbourhood of surface k.
• $g_{m,k} = \langle \overrightarrow{n_k}, \overrightarrow{U_{km}} \rangle . \frac{\langle \overrightarrow{n_m}, \overrightarrow{U_{mk}} \rangle}{r^2}$: geometric factor (named so because it only depends on the geometry of the scene), where r is the distance between surfaces m and k, and $\overrightarrow{n_k}$ is the normal to surface k.

Note that ρ_k^{hd}, ρ_k^{dd} and all reflectances considered in this work are defined as in [22], they are thus unitless.

C. Radiance from the equivalent surface level

Let $\langle L \rangle$ be the equivalent radiance corresponding to the equivalent surface, defined at the height of the urban canopy, and associated with our pixel. This equivalent radiance $\langle L \rangle$ is the radiance measured by a sensor viewing a flat surface:

$$\begin{aligned} \langle L \rangle &= \langle L_{es} \rangle + L_{atm}^{\downarrow} + L_{coupling} + L_{atm}^{\uparrow es-c} + L_{diff}^{\uparrow} \\ &= \frac{\langle \rho \rangle}{\pi} E_D t^{\uparrow es-c} + L_{atm}^{\downarrow} + L_{coupling} + L_{atm}^{\uparrow es-c} + L_{diff}^{\uparrow} \end{aligned} \tag{3}$$

where $\langle L_{es} \rangle$ is the radiance at the equivalent surface level, resulting from the reflection of solar irradiance on the equivalent surface. $L_{atm}^{\uparrow es-c}$ and $t^{\uparrow es-c}$ are respectively the atmospheric upwelling radiance and transmission from the equivalent surface toward the sensor. All other terms are the same as in Equation (2), except they are here defined at the equivalent surface level (and not for an elementary surface on the ground).

As we can see, the terms due to the environment are not present here (compared to Equation (2)) because the surface is flat.

The atmospheric upwelling transmission and the atmospheric upwelling radiance of Equation (2) can be expressed as functions of those of Equation (3):

$$t^{\uparrow k-c} = t^{\uparrow k-es} \times t^{\uparrow es-c} \tag{4}$$

$$\begin{aligned} \sum_k S_k L_{atm,k}^{\uparrow} &= \sum_k S_k L_{atm,k}^{\uparrow k-c} = \sum_k S_k L_{atm,k}^{\uparrow k-es} \times t^{\uparrow es-c} \\ &\quad + L_{atm}^{\uparrow es-c} \end{aligned} \tag{5}$$

where
- $t^{\uparrow k-es}$: atmospheric upwelling transmission from a surface k to the equivalent surface.
- $L_{atm}^{\uparrow k-es}$: atmospheric upwelling radiance from surface k toward the equivalent surface.

D. Identification leading to the mixing model

The total radiance expressions for a pixel defined in Equations (2) and (3) are equal, i.e. $L_{mix} = \langle L \rangle$. Thus, considering the physical conservation of radiant flux at every scale, a term by term identification is performed. By identifying the radiative terms of the same nature in these two expressions, we can easily simplify by the following terms: $L_{atm}^{\downarrow}$, $L_{coupling}$, $L_{diff}^{\uparrow}$. Using (5) the term $L_{atm}^{\uparrow es-c}$ can also be eliminated. This leads to

$$\begin{aligned} &\frac{\langle \rho \rangle}{\pi} E_D t^{\uparrow es-c} = \sum_k S_k t^{\uparrow k-c} \frac{\rho_k^{dd}(\overrightarrow{U_s}, \overrightarrow{U_{c,k}})}{\pi} E_{D,k} \\ &+ \sum_k S_k t^{\uparrow k-c} \iint_{m \in V_k} \frac{\rho_k^{dd}(\overrightarrow{U_{km}}, \overrightarrow{U_{c,k}})}{\pi} g_{m,k} \left[\frac{\rho_m^{dd}(\overrightarrow{U_s}, \overrightarrow{U_{mk}})}{\pi} E_{D,m} \right. \\ &\left. + \frac{\rho_m^{hd}(\overrightarrow{U_{mk}})}{\pi} E_{diff,m} \right] dS_m + \sum_k S_k L_{atm,k}^{\uparrow k-es} \times t^{\uparrow es-c} \end{aligned} \tag{6}$$

We then simplify by $t^{\uparrow es-c}$, in all terms, using Expression (4). The last terms $L_{atm,k}^{\uparrow k-es}$ can be removed, considering that the atmosphere between the ground and the urban canopy can be neglected. For the same reason, atmospheric transmission under the canopy level is $t^{\uparrow k-es} \simeq 1$ (assuming we are outside a strong absorption band). We can finally express the equivalent reflectance of a pixel as a function of the elementary reflectances:

$$\begin{aligned} \langle \rho \rangle &= \sum_{k=1}^{N} S_k \left[\rho_k^{dd}(\overrightarrow{U_s}, \overrightarrow{U_{c,k}}) \frac{E_{D,k}}{E_D} \right. \\ &+ \iint_{m \in V_k} \rho_k^{dd}(\overrightarrow{U_{km}}, \overrightarrow{U_{c,k}}) g_{m,k} \left(\frac{\rho_m^{dd}(\overrightarrow{U_s}, \overrightarrow{U_{mk}})}{\pi} \frac{E_{D,m}}{E_D} \right. \\ &\left. \left. + \frac{\rho_m^{hd}(\overrightarrow{U_{mk}})}{\pi} \frac{E_{diff,m}}{E_D} \right) dS_m \right] \end{aligned} \tag{7}$$

Note that this equation is valid for a pixel totally or partially in the sun ($E_D \neq 0$).

We now assume, without loss of generality, that reflectances have a Lambertian behaviour in the second additive term (ρ_k and ρ_m do not depend on the sun and sensor directions: $\rho^{hd} = \rho^{dd} = \rho$). This assumption is possible because multiple reflections tend to make the material behaviour Lambertian and because this term is small compared to the first one. The reflectances in the latter are therefore denoted ρ_k and ρ_m. We also approximate the integrals by sums, assuming that the neighbourhood of an elementary surface k is composed of elementary surfaces m of area ΔS and each one made up of only one material. This yields

$$\langle \rho \rangle = \langle \rho \rangle_D + \langle \rho \rangle_{env} \tag{8}$$

with

$$\langle \rho \rangle_D = \sum_{k=1}^{N} S_k \frac{E_{D,k}}{E_D} \rho_k^{dd}(\overrightarrow{U_s}, \overrightarrow{U_{c,k}}) \tag{9}$$

$$\langle \rho \rangle_{env} = \langle \rho \rangle_{env,D} + \langle \rho \rangle_{env,diff} \tag{10}$$

with

$$\begin{cases} \langle \rho \rangle_{env,D} = \sum_{k=1}^{N} \sum_{m \in V_k} S_k \frac{g_{m,k}}{\pi} \Delta S \frac{E_{D,m}}{E_D} \rho_k \, \rho_m \\ \langle \rho \rangle_{env,diff} = \sum_{k=1}^{N} \sum_{m \in V_k} S_k \frac{g_{m,k}}{\pi} \Delta S \frac{E_{diff,m}}{E_D} \rho_k \, \rho_m \end{cases} \tag{11}$$

This is our physical mixing model in a general case. The first term of the right-hand side of Eq. (8) gives the linear mixing of the pixel's own components ρ_k (see Eq. (9)). The second term, which is quadratic with respect to the reflectances (see Eq. (10)-(11)), is due to the reflections from surfaces in the environment of each elementary surface. If the landscape is flat and the incident irradiance is homogeneous, the second term vanishes and the well-known linear mixing model is obtained.

Equation (8) is valid for one wavelength and one pixel. All the irradiances depend on both wavelength and pixel location. Reflectance only depend on wavelength if it is considered Lambertian. The other terms are due to geometrical layout and only depend on pixel location.

Note that $\sum_{k=1}^{N} S_k = 1$, but we also have another important property. The incident flux conservation at a pixel permits us to write: $\Omega E_D = \sum_{k=1}^{N} d\omega_k E_{D,k}$. This leads to the following property:

$$\sum_{k=1}^{N} S_k \frac{E_{D,k}}{E_D} = 1 \tag{12}$$

This flux conservation property will be verified with simulations in the next section.

For the sake of simplicity, reflectances in (9) are also assumed to be Lambertian hereafter. Equation (9) then becomes

$$\langle \rho \rangle_D = \sum_{k=1}^{N} S_k \frac{E_{D,k}}{E_D} \rho_k \tag{13}$$

III. Validation of the physical model

In this section, a "validation" of the above model is presented, using several synthetic multispectral images. These images were simulated with the 3D radiative transfer code, AMARTIS V2 [21], which is described hereafter.

A. AMARTIS description

AMARTIS is a radiative transfer code specifically dedicated to urban areas. It allows to simulate airborne and space-borne multi-angular acquisitions, in the [0.4; 2.5μm] domain, over scenes defined by their 3D geometry, the optical properties of every material composing the scene, the atmosphere conditions and the viewing geometry of the sensor.

The synthetic 3D scene is composed of uniform triangular shape facets considered as homogeneous in terms of reflectance which be either Lambertian either bidirectional. The atmospheric radiative properties are modelled thanks to the radiative transfer code 6S [23]. The aerosols can be modelled by the standard models of 6S, by their physical properties (with notably Junge or multimodal distributions) or directly by their optical properties (spectral optical thickness, single scattering albedo and phase function). The gaseous atmosphere is modelled by the standard models of 6S. AMARTIS allows the simulation of airborne or satellite sensors. The sensor is defined by the following parameters: its zenith and azimuth viewing angles defined by the optical axis orientation pointed at the centre of the scene, its pixels matrix (number of pixels by rows and columns and pixel size), its spatial resolution, the wavelengths of observation, and the focal length of the instrument. The altitude of the sensor is deduced from the previous geometrical parameters.

In remote sensing, a flat ground assumption is usually made to model the signal at ground and sensor levels. However, in cities, at very high spatial resolution, this hypothesis is no longer valid because of the complexity introduced by the relief which induced specific radiative effects. Thus, AMARTIS has been developed to overcome these limitations. It is able to model independently all the radiative contributors at ground (irradiance unit) and at sensor (radiance unit) levels (Figure 3). The irradiance at ground level (E_{tot}) is the sum of four components (Figure 3(a)): the direct irradiance (E_D), the scattered and diffused irradiance (E_{diff}), the Earth-atmosphere coupling irradiance ($E_{coupling}$) and the downward reflected irradiance (E_{env}). The radiance at sensor level (L_{tot}) is the sum of the different components defined in Section II by Equations (1)-(2) (see Figure 3(b)). The analytic expression of each radiative component is detailed in [21]. Due to the strong heterogeneity of the scene, $E_{coupling}$, E_{env} and L_{env} are solved using ray tracing and Monte Carlo methods.

The outputs of AMARTIS are radiance images at sensor level (total, direct and due to the neighbourhood and all other terms present in our equations) and irradiance images at surface level (total, direct and diffused).

AMARTIS results are reliable and have already been validated, however the code has two little disadvantages that can induce some artefacts in our results:

- AMARTIS uses a Monte Carlo process to estimate the radiative terms due to reflections, which induces statistical fluctuations, essentially in shadowed areas. As this method needs a lot of photons to converge well, its convergence is less reliable in partly shadowed regions.
- Before AMARTIS calculation, a triangulation of the 3D scene is made. A hyperfine meshing is needed to achieve a good accuracy. Unfortunately we are limited by computational cost and memory size, and the used mesh size often implies triangles bigger than our small elementary surfaces (three or four times bigger). This can

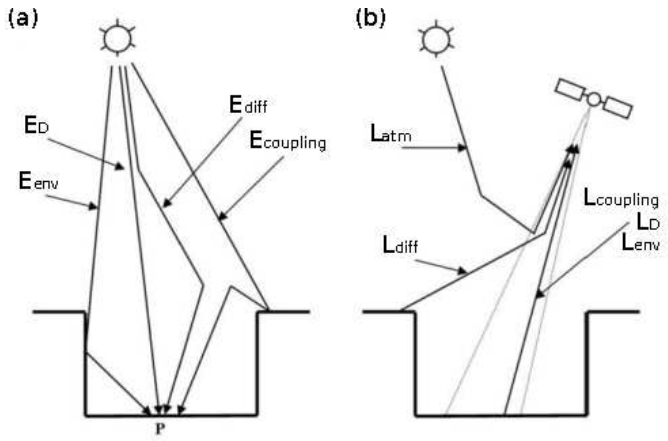


Fig. 3. Description of the different modelled radiative contributors (a) the irradiances at ground level (b) the radiance at sensor level

induce some artefacts on the computed radiances, mainly around the shadow/light transitions.

B. Methodology

For each studied image, at a coarse resolution X, the same image at a high resolution x is used, whose pixels would be used as the elementary surfaces k and m in our mixing model.

To validate our model we first proceed as follows (Section III-D1):

1) Compute an image of reflectances $\langle\rho\rangle$ for pixels at a coarse resolution X (by applying our mixing model in Equations ((8), (13), (10) and (11)), using values provided at a fine resolution x for the elementary surfaces.
2) Compute the total radiance image $\langle L\rangle$ (Equation (3)) at resolution X, using the obtained image of reflectances $\langle\rho\rangle$.
3) Compare our computed image $\langle L\rangle$ with the total radiance image L given by AMARTIS at resolution X.
   We then focus on $\langle L_{es}\rangle$ (in Eq. (3)). Due to (8), it can be defined as: $\langle L_{es}\rangle = \langle L\rangle_D + \langle L\rangle_{env}$, with:
   - $\langle L\rangle_D = \frac{\langle\rho\rangle_D}{\pi} E_D t^{\uparrow es-c}$ (with $\langle\rho\rangle_D$ defined in (13))
   - $\langle L\rangle_{env} = \frac{\langle\rho\rangle_{env}}{\pi} E_D t^{\uparrow es-c}$ (with $\langle\rho\rangle_{env}$ defined in (10)). (this corresponds to Equation (2)).
4) Compare our computed image $\langle L\rangle_D$ with the radiance image L_D given by AMARTIS at resolution X. Both correspond to the solar radiance due to the solar irradiance on the surface directly reflected toward the sensor.
5) Compare L_{env} given by AMARTIS for the resolution X, with our computed $\langle L\rangle_{env}$. These terms correspond to the radiance due to the neighbourhood of the considered surface (induced by reflections).

To show the discrepancy we get when comparing our results with AMARTIS' ones, we define what we will call "errors" hereafter:

- a global absolute error in radiance units:

$$Err_L(Z) = \sqrt{mean_i(\langle Z\rangle_i - Z_i)^2} \tag{14}$$

- a global relative error in percentage:

$$Err_{\%}(Z) = \sqrt{\frac{mean_i(\langle Z\rangle_i - Z_i)^2}{mean_i(Z_i^2)}} \times 100 \tag{15}$$

- an error in radiance units computed at a pixel i (difference between the two images):

$$err_{L_i}(Z) = |\langle Z\rangle_i - Z_i| \tag{16}$$

In all these definitions Z stands for L, L_{env} or L_D, and i corresponds to a pixel in the area of interest.

In a second step, a further analysis is proposed to show the usefulness of taking into account the environment radiative component (Section III-D2):

- Study the proportion of the terms due to environment in radiances but also in reflectances (non-linear part in the mixed reflectance).
- Compare the importance of the term due to direct irradiance, $\langle\rho\rangle_{env,D}$, to that of the one due to diffused irradiance, $\langle\rho\rangle_{env,diff}$, in reflectances (see definitions in (11)).

We finally validate, in Section III-D3, the flux conservation property earlier discussed at the end of Section II-D.

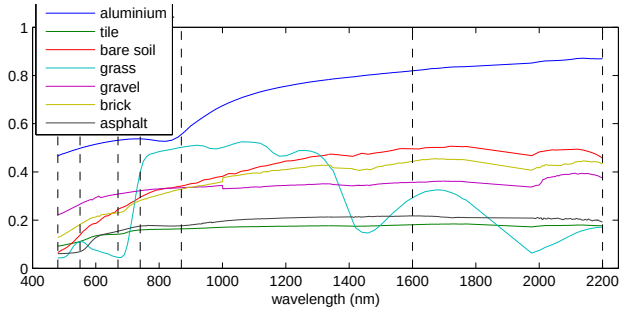


Fig. 4. Reflectance spectra of materials present in our scenes

C. Data description

We start our analysis with simple geometric scenes, then a synthetic urban scene is used. Images used in this study are noise-free. For each studied image at a coarse resolution X, the same image at a high resolution $x = 0.2m$ is used, whose pixels would be used as the elementary surfaces k and m in our mixing model.

All used images have been generated with the following atmospheric conditions: urban aerosol model with a visibility of 23 km, mid-latitude summer atmosphere.

This study concerns the reflective range [0.4 - 2.5 μm]. Selected wavelengths are: 480, 550, 670, 740, 870, 1600 and 2200 nm.

Figure 4 shows the reflectance spectra (coming from MEMOIRES data base [24]) involved in our images. The dashed lines correspond to the 7 studied wavelengths.

1) Simple geometric scenes:

Figure 5 represents a scene with a canyon (between two buildings), which is the most common geometrical shape in urban environments. From this scene four images are generated. In these images, the roofs are made up of tiles and the bottom of the canyon is covered, for its left half, by gravel and the right half by asphalt. Walls are made up of bricks, except for one image where aluminium is used instead.

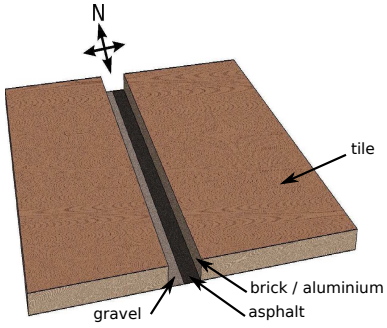


Fig. 5. The simple geometric scenes (canyon shape)

TABLE I
DESCRIPTION OF THE 4 GEOMETRIC SCENES (GENERATED FROM THE SIMPLE SCENE) : SOLAR ANGLES, PERCENTAGE OF SHADOW IN THE CANYON, AND MATERIAL COVERING THE WALLS.

	sun (zenith,azimuth)	shadow in the canyon	walls
C1	(16°,180°)	0%	brick
C2	(27°,239°)	20%	brick
C3	(45°,264°)	40%	brick
C4	(45°,264°)	40%	aluminium

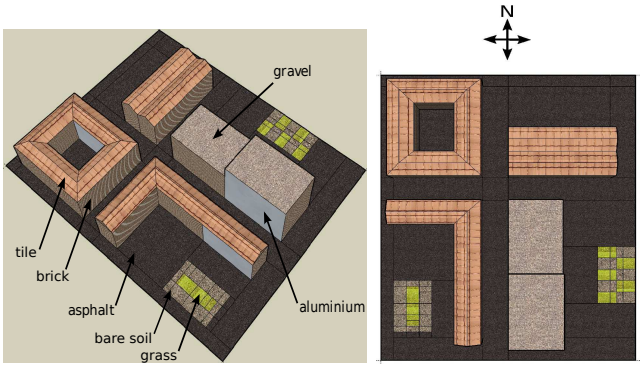
In order to perform a general investigation, also to see how the proportion of the non-linear term (the contribution of the reflections on the environment of a pixel) can vary depending on the sun direction, several cases of irradiation are studied. We also aim at see the effect of having walls covered with aluminium instead of brick (modern buildings are often covered with very reflective surfaces: glass, aluminium, etc.), with unchanged solar conditions. Our four images are presented in Table I where are given the solar angles, the percentage of induced shadow in the canyon and the material covering the walls.

For all these cases, the studied images have a coarse resolution of $X = 5$ m. In the coarse-resolution images, the canyon width corresponds to two 5 m pixels.

2) Urban scene:

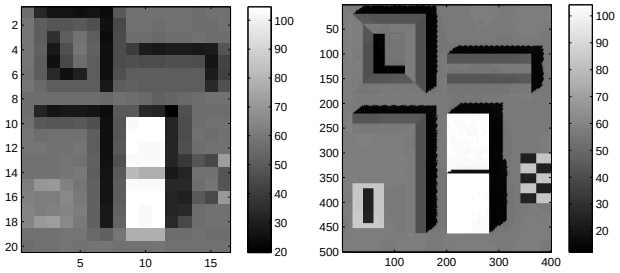
A second more complex and realistic urban scene is also simulated (Figure 6) which contains different kinds of building configurations that are common to European towns. Road is covered with asphalt, two small gardens include vegetation and bare soil, walls are mostly brick with some comprising aluminium cover (grey walls in Fig. 6), and finally sloped roofs are covered with tiles with flat ones covered with gravel. As the effect of solar angles is studied with the simple geometric scenes, we here consider only one configuration: solar zenith angle = 26°, solar azimuth angle = 227°. These angles where chosen because they permit to have no pixels entirely shadowed, as our model in not valid for entirely shadowed pixels.

From this scene, of size 100 x 80 m^2, two images with different coarse resolutions have been generated to be studied separately: one with a resolution $X = 4$ m (size 25 x 20 pixels) and one with a resolution $X = 5$ m (size 20 x 16 pixels).



(a) Perspective view (b) Top view

Fig. 6. The urban scene



(a) 5 m resolution (b) 0.2 m resolution

Fig. 7. Urban scene - Total radiance images at resolutions of 5 m (left) and 0.2 m (right) - wavelength 670 nm

Figure 7 represents the total radiance images for this urban scene at resolutions of 0.2 and 5 m.

D. Results ans discussions

1) Global analysis:

a) Simple geometric scenes: Results correspond to values of radiances obtained only for the pixels inside the canyon (two columns of 5 m pixels), since neighbouring reflections can occur only in the bottom of the canyon, contributing then to the non-linear term of the equivalent reflectance. In Table II and III, results are presented for the seven studied wavelengths. Scene C1 is not considered in these tables. As the solar irradiation direction is almost parallel to the walls, they yield no reflections in this scene. The term due to the environment is negligible, and the well-known linear model remains valid. The results show a good agreement between L and $\langle L \rangle$.

In these tables, the mean value of the radiances of interest is presented, to have an idea of the global level of values. We also show the discrepancy we get when comparing our results with AMARTIS' ones.

Table II shows that, globally, errors are very low compared to the mean values of $\langle L \rangle$ (the corresponding $Err_{\%}(L)$, shown in Table II, varies globally from 0 to 3 %).

To understand the source of errors, it is interesting to compare error values Err_L in the two tables (Table II for $\langle L \rangle$ and Table III for $\langle L \rangle_{env}$): they are almost the same. This means that errors seen for $\langle L \rangle$ are essentially due to errors obtained with $\langle L \rangle_{env}$. Moreover, errors for $\langle L \rangle_D$ are

TABLE II
SIMPLE SCENES - TOTAL RADIANCES $\langle L \rangle$ (IN $W/m^2/sr/\mu m$) AND ERRORS

wavelength (nm)		480	550	670	740	870	1600	2200
$\langle L \rangle$	C2	64.1	59.5	65.6	56.2	48.5	15.1	3.8
	C3	40.5	33.5	36.2	30.8	26.2	8.3	1.9
	C4	42.6	35.7	39.4	33.2	28.2	9.5	2.3
Err_L (L)	C2	0.2	0.3	0.3	0.3	0.3	0.1	0.0
	C3	0.1	0.2	0.4	0.4	0.4	0.2	0.0
	C4	0.5	0.6	0.8	0.7	0.7	0.3	0.1
$Err_{\%}$ (L)	C2	0.3	0.4	0.8	0.5	0.6	0.9	0.6
	C3	0.3	0.7	1.0	1.3	1.6	2.1	2.1
	C4	1.0	1.6	2.0	2.0	2.3	3.2	3.5

TABLE III
SIMPLE SCENES - $\langle L \rangle_{env}$ (IN $W/m^2/sr/\mu m$) AND ERRORS

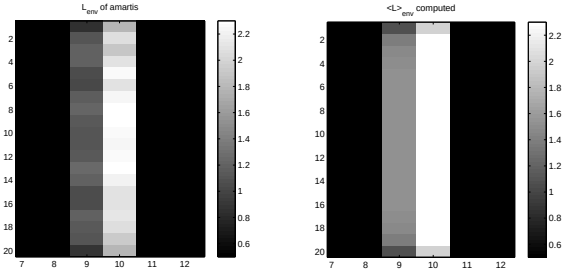
wavelength (nm)		480	550	670	740	870	1600	2200
$\langle L \rangle_{env}$	C2	0.4	0.7	1.2	1.3	1.3	0.6	0.1
	C3	0.6	1.0	1.9	2.1	2.2	1.0	0.2
	C4	2.0	2.6	4.3	3.9	3.7	1.9	0.5
Err_L (L_{env})	C2	0.1	0.2	0.2	0.2	0.2	0.1	0.0
	C3	0.1	0.2	0.3	0.4	0.4	0.2	0.0
	C4	0.4	0.5	0.7	0.7	0.6	0.3	0.1
$Err_{\%}$ (L_{env})	C2	28	27	20	19	19	17	18
	C3	27	27	20	19	19	18	19
	C4	26	26	19	19	19	18	19

negligible. Note that we only compute $\langle L \rangle_{env}$ and $\langle L \rangle_D$ with our equations (other radiance terms are provided by AMARTIS and do not play any role in this study).

We now focus on Table III. When comparing the three scenes, it can be noticed that values of Err_L for C4 are higher than for the other two scenes. With C3, errors are a little higher than for C2 (essentially when comparing with the radiance values) but values remain close to each other. These differences in errors between the three scenes are linked to the differences for the values of $\langle L \rangle_{env}$: Err_L is higher for higher values of $\langle L \rangle_{env}$, whereas percentages ($Err_{\%}$) are globally stable from one scene to another.

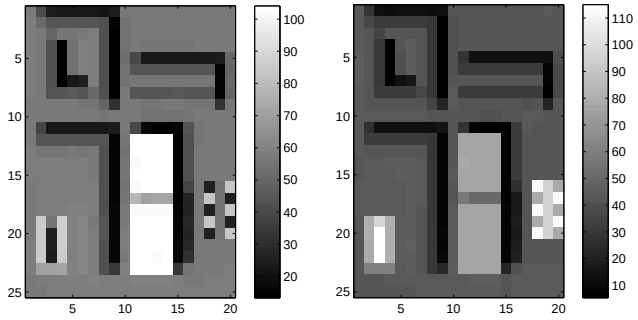
In spite of these errors that will be explained hereafter, results show that our model correctly describes the reflections phenomena in this simple canyon case. Indeed, the error is in the worst case equal to 0.7 and in most cases below 0.3 $W/m^2/sr/\mu m$. This means our errors are noticeably in the order of the instrumental noise level, which is about 0.2 $W/m^2/sr/\mu m$ for airborne sensors (see e.g. [25]). Therefore, most errors are reasonably weak even if errors in percentages can look high. Furthermore, most errors are due to artefacts induced by the simulation process, as explained below.

Figure 8 is an example of an image of L_{env} from scene C3, at 670 nm. The image of L_{env} (left one) can be compared with the image of $\langle L \rangle_{env}$ (right one). It can be seen that there is more discrepancy for column 9 than for column 10. Obtained results, for the three scenes, showed that errors are essentially localized in column 9 (about 33% of error), while errors are much lower in column 10 (max 13%). Note that column 9 contains a certain percentage of shadow (80% of the column area for scene C3), while, in column 10, pixels



(a) L_{env} (b) $\langle L \rangle_{env}$

Fig. 8. Scene C3 (zoom on the canyon) - Comparison of L_{env} and $\langle L \rangle_{env}$ (in $W/m^2/sr/\mu m$) at 670 nm



(a) wavelength 670 nm (b) wavelength 870 nm

Fig. 9. Urban scene at 4 m resolution - $\langle L \rangle$ (in $W/m^2/sr/\mu m$)

are totally irradiated by the sun. These errors are artefacts due to AMARTIS code already explained in Section III-A:

- The result of Monte Carlo in partly shadowed regions. The induced effect can be seen in the AMARTIS image (Figure 8.a): values along column 9 are not homogeneous as they should be, with fluctuations around 16% of the radiance values.
- The scene triangulation. This induces some artefacts on the computed radiances, mainly around the shadow/light transitions. These errors can hardly be quantified. However, a simulation made on a smallest scene showed that, by increasing the number of triangles, the global error Err_L decreases by 28%.

For these reasons, we consider that the errors mainly result from the simulation process itself rather than from our modelling. Furthermore, errors are globally low and, if we consider that a great part of these errors is due to simulation, the residual errors become very close to the instrumental noise level. We thus consider the results as satisfactory.

b) Urban scene at 4 m resolution: The ground pixels[1] are all pure at this resolution except for the two gardens, since buildings dimensions have been chosen as multiples of 4 m. This will permit us to isolate the effect of reflections, to better see their behaviour and easily localize causes of errors. Presented results, for the urban scenes, correspond only to the pixels where $\langle L \rangle_{env}$ is higher than 0.2 $W/m^2/sr/\mu m$ (instrumental noise level), so only for the pixels where the environment plays a role.

Using the errors defined in (14) and (16), results obtained pixel by pixel (over all the studied wavelengths) can be summarized as follows, for the pixels of interest (47% of the image pixels). For L_D the mean error, $Err_L(L_D)$, is equal to 0.2 $W/m^2/sr/\mu m$ and for the total radiance we have 0.3 $W/m^2/sr/\mu m$. Concerning L_{env}, $Err_L(L_{env}) = 0.2$ $W/m^2/sr/\mu m$. For this term, Figure 11(a) shows the histogram of all obtained errors $err_{L_i}(L_{env})$ (over all wavelengths). The maximum value is of 1.3 but 95% of values are below 0.4 $W/m^2/sr/\mu m$. Errors are globally low compared with instrumental noise level ($\sim$ 0.2 $W/m^2/sr/\mu m$).

Figure 9 shows two images of total radiance $\langle L \rangle$ at 670

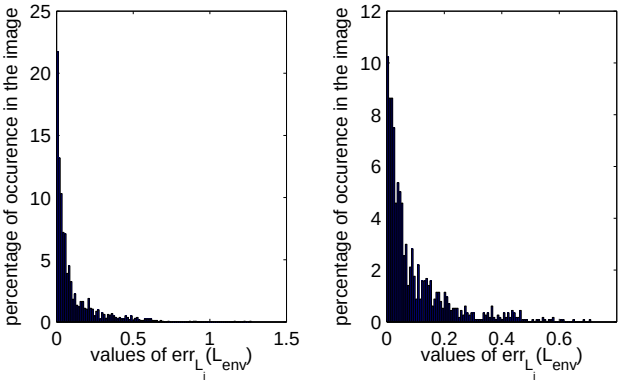
[1]Surface on the ground corresponding to the pixel projection.

nm and 870 nm. By comparing the above values of errors ($Err_L \simeq 0.2$ for all terms) with the radiance values in the images presented in this figure, it can be seen that errors are generally very low.

However, some pixels yield high errors for $\langle L \rangle_{env}$. In Figure 10, are given the images of $\langle L \rangle_{env}$ (left image) and the difference between the image of L_{env} and $\langle L \rangle_{env}$ (right image), at 670 nm. It can be seen that most errors are low but higher errors exist for some pixels (the brightest ones). Most of these errors are due to the same reasons as detailed above (Section III-D1a). However, another type of errors occurs in this scene, that concerns some pixels at the canyon crossings. This is due to our aggregation algorithm (when computing Equation (8)): for simplification, we assumed that a surface on the ground can only receive reflections from walls of the same canyon. This induces an underestimation of $\langle L \rangle_{env}$ for some pixels on the corners. The maximum error obtained in such pixels is 0.9 $W/m^2/sr/\mu m$, for pixel (row=24,col=8), because we ignored there contributions due to reflections on aluminium which are not negligible. Considering pixels around the two canyons crossing, the maximum error is 0.6 $W/m^2/sr/\mu m$ for pixel (row=10,col=8). These errors are thus well identified and we just have to keep in mind that they are not due to our model.

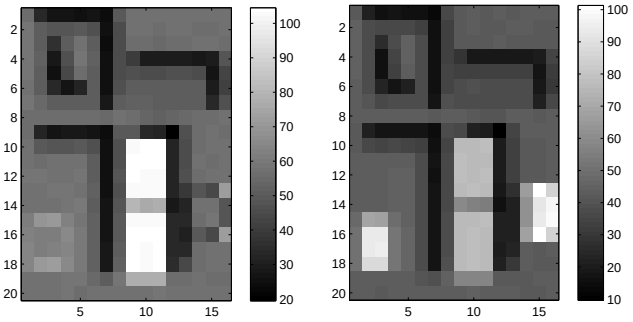
c) Urban scene at 5 m resolution: We here face a more realistic case: a ground pixel can be a mixture of two or three materials.

The global results over all the wavelengths, for the pixels of interest (here 51% of the image pixels) gave, for L, a mean error $Err_L(L)$ of 0.2 $W/m^2/sr/\mu m$. For L_D we have $Err_L(L_D) = 0.1$ $W/m^2/sr/\mu m$. And finally $Err_L(L_{env}) = 0.1$ $W/m^2/sr/\mu m$. Figure 11(b) shows the histogram of all errors $Err_{L_i}(L_{env})$. The maximum value is here 0.7 and for 95% of the pixels values are below 0.4 $W/m^2/sr/\mu m$. Most of error values are low comparing with the level of sensor noise. Furthermore, comparing the level of the errors with the values of total radiance faced in our images (see Figure 12), we can note that they are globally low. Figure 13 shows our computed image of $\langle L \rangle_{env}$ (left) and the difference between L_{env} and $\langle L \rangle_{env}$ (right), at 670 nm. The pixels with the highest errors are due to the same reasons as previously noted (in III-D1a and III-D1b).



(a) 4 m resolution image (b) 5 m resolution image

Fig. 11. Urban scenes - Histograms for error $err_{L_i}(L_{env})$ (in $W/m^2/sr/\mu m$) over all wavelengths



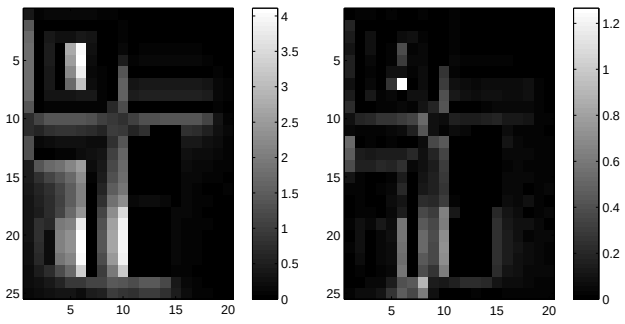
(a) wavelength 670 nm (b) wavelength 870 nm

Fig. 12. Urban scene at 5 m resolution - $\langle L \rangle$ (in $W/m^2/sr/\mu m$)

Through all these simulations, it can be concluded that the main errors are due to the generation of the synthetic scenes or to our aggregation algorithm. Moreover, these errors are comparable in level to the radiometric sensitivities of airborne sensors. We can thus consider that our model defined by Equations (8), (13), (10) and (11) is validated.

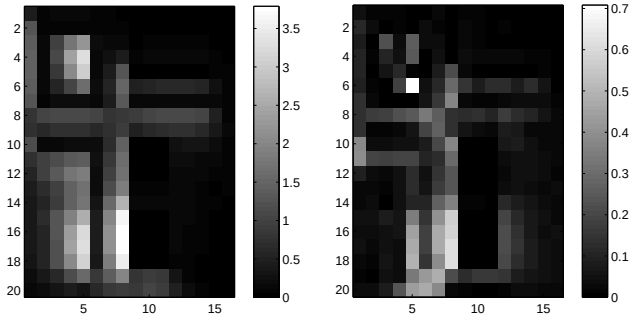
2) Analysis of the different terms:

In this section we evaluate the relative importance of the terms related to the reflections, using the same simulated



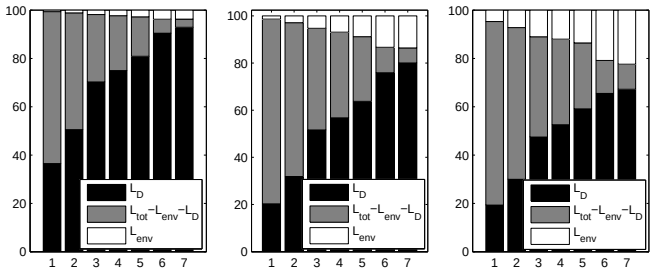
(a) $\langle L \rangle_{env}$ (b) error: $|\langle L \rangle_{env} - L_{env}|$

Fig. 10. Urban scene at 4 m resolution - $\langle L \rangle_{env}$ and error (in $W/m^2/sr/\mu m$) at 670 nm



(a) $\langle L \rangle_{env}$ (b) error: $|\langle L \rangle_{env} - L_{env}|$

Fig. 13. Urban scene at 5 m resolution - $\langle L \rangle_{env}$ and error (in $W/m^2/sr/\mu m$) at 670 nm



(a) C2 (b) C3 (c) C4

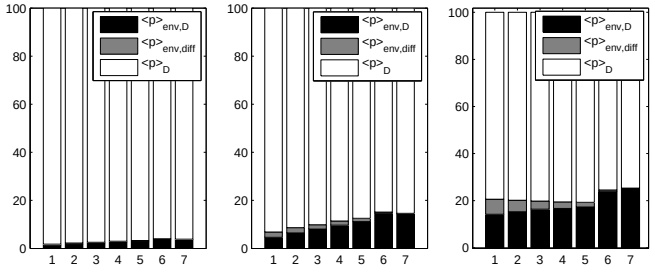
Fig. 14. Simple scenes - Percentages of $\langle L \rangle_D$ and $\langle L \rangle_{env}$ compared with total radiance, for the 7 studied wavelengths (nm): 1- 480, 2- 550, 3- 670, 4- 740, 5- 870, 6- 1600, 7- 2200

data. For the urban scene only the 5m image is considered here, since it is the more realistic example that includes pixel mixtures.

a) Simple geometric scenes: Figure 14 shows that the proportion of $\langle L \rangle_{env}$ depends on the sun direction (see differences between scenes C3 and C2), which induces different levels of irradiances on the walls (higher values of irradiance when the sun direction is perpendicular to the surface). Furthermore, when the proportion of shadow increases, the amount of direct radiance is lower, so radiance due to the environment takes a more important proportion. Proportions of L_{env} depend also on the wavelength (L_{env} has lower importance in the blue domain). Finally, it is obvious that the materials present in the scene play an important role too (more reflections with the aluminium, for scene $C4$, compared with scene $C3$ which is characterized by the same sun direction).

Similar comments can be made concerning the proportion of $\langle \rho \rangle_{env}$ in total reflectance, presented in Figure 15. Percentages are higher than 5 % in most cases (and higher than 10% when looking only at scenes C3 and C4). Such proportions cannot be reasonably neglected in a mixing model for reflectances. We also notice that reflections induced by direct radiance $\langle \rho \rangle_{env,D}$ are more important than those induced by diffused radiances $\langle \rho \rangle_{env,diff}$ and that the latter term is almost equal to zero for wavelengths higher than 1600 nm.

b) Urban scene at 5 m resolution: In this section we focus only on the reflectance terms, which are the terms of main concern in this paper, since values of radiance essentially



(a) C2 (b) C3 (c) C4

Fig. 15. Simple scenes - Percentages of $\langle \rho \rangle_D$, $\langle \rho \rangle_{env,D}$ and $\langle \rho \rangle_{env,diff}$ in total reflectance $\langle \rho \rangle$, for the 7 studied wavelengths (nm): 1- 480, 2- 550, 3- 670, 4- 740, 5- 870, 6- 1600, 7- 2200

TABLE IV
URBAN SCENE 5M - PERCENTAGE OF $\langle \rho \rangle_{env,D}$ AND $\langle \rho \rangle_{env}$ IN $\langle \rho \rangle$, FOR ONE ZONE IN THE IMAGE (DESCRIBED IN THE TEXT)

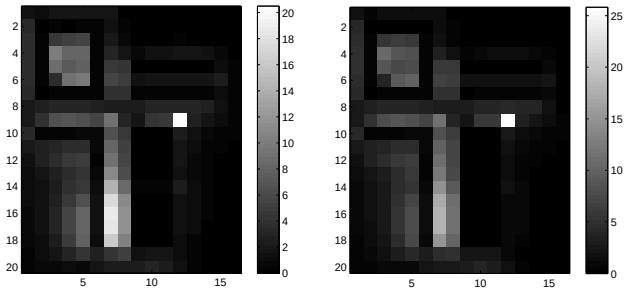
wavelength (nm)	480	550	670	740	870	1600	2200
$\langle \rho \rangle_{env}$ (%)	8	8	10	11	11	15	14
$\langle \rho \rangle_{env,D}$ (%)	6	6	9	9	10	14	14

helped us to validate the model with AMARTIS. Figure 16 contains two images that represent $\langle \rho \rangle_{env}$ percentages in $\langle \rho \rangle$ for wavelengths 670 nm and 870 nm. We thus can see, as previously (in III-D2a), that the importance of the term due to reflections depends on the materials present in the scene, on the presence of shadow and on the area in the scene (the geometry and orientation of buildings towards the sun, etc.). It seems e.g. obvious that $\langle \rho \rangle_{env}$ is higher for pixels situated near the walls irradiated by the sun and percentages of $\langle \rho \rangle_{env}$ are higher where walls are covered with aluminium (see Figure 6 to localize the concerned zones). Note that the white pixel in the images corresponds to a surface with a high proportion of shadow, which explains this high percentage of $\langle \rho \rangle_{env}$.

Table IV shows the mean value of percentages of $\langle \rho \rangle_{env}$ as a function of the wavelength for one zone in the image: the canyon situated in the bottom of the image (rows 10 to 18 - columns 7 and 8). As previously (for the simple scenes in III-D2a), the importance of the term due to the environment cannot be neglected (all values higher than 5 %). It can be noticed again that $\langle \rho \rangle_{env}$ is less important in the blue domain. It is also clear, as for the simple scenes, that the reflections induced by direct radiance, $\langle \rho \rangle_{env,D}$, have the highest percentage (between 75% and 100% of $\langle \rho \rangle_{env}$) , and this will be useful for the next section.

3) Verification of the flux conservation property:

To entirely validate our model, it is necessary to verify the property of the incident flux conservation, at a pixel level, cited in Section II, by validating Equation (12). A new image at 5 m resolution is simulated, with only a part of our urban scene (rows 1 to 10 and columns 1 to 7). Taking a smaller scene permits a finer triangulation, and thus better accuracy in results. Note that all pixels in this image do not correspond to flat homogeneous surfaces, so the assumptions of the classical linear model are not met here. The sum-to-one property is obtained: the left-hand term of (12) has a mean value of



(a) wavelength 670 nm (b) wavelength 870 nm

Fig. 16. Urban scene at 5 m resolution - Images of percentage of $\langle \rho \rangle_{env}$ in $\langle \rho \rangle$

0.999 over the whole image and a standard deviation of 0.01. Property (12) is then true. This result is very important for the unmixing community as we will see in Section V.

IV. Derivation of a simplified mixing model

As explained previously (in Section I), the aim of this work is to derive a mixing model for reflectances, valid in urban environments, and that could be used for spectral unmixing. However, the mixing model defined by Equations (8), (13), (10) and (11) is not convenient yet for unmixing methods. It is a bit complicated because of the dependency of the irradiance terms on the wavelength.

Rewriting the model by highlighting the above-defined dependencies of the different terms vs. the 2D pixel position y and the wavelength λ, we have:

$$\begin{aligned}\langle\rho\rangle(y,\lambda) &= \sum_{k=1}^{N} S_k(y)\frac{E_{D,k}}{E_D}(y,\lambda)\rho_k(\lambda) \\ &+ \sum_{k=1}^{N}\sum_{m\in V_k} S_k(y)\frac{g_{m,k}(y)}{\pi}\Delta S\left(\frac{E_{D,m}}{E_D}(y,\lambda)\right. \\ &+ \left.\frac{E_{diff,m}}{E_D}(y,\lambda)\right)\rho_k(\lambda)\rho_m(\lambda)\end{aligned} \tag{17}$$

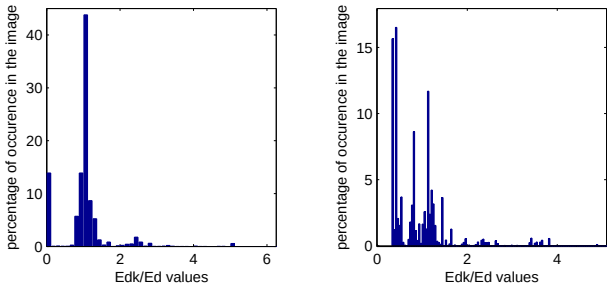
Using the source separation terminology [26]–[29] and considering wavelength-dependent sources, this model is an instantaneous (or memoryless) one, since each mixture at one wavelength only depends on the sources (reflectances here) and the mixing coefficients at the same wavelength. It is a linear-quadratic model with respect to the reflectances (our sources) but the mixing coefficients include terms with a double dependency on the position and on the wavelength: the irradiance ratios (i.e. $\frac{E_{D,k}}{E_D}$, ...). This makes the model difficult to use for unmixing. Therefore, a study of the spectral behaviour of these terms is necessary to see if some simplifications are possible. To do this, we use the urban scene at 5 m resolution described in Section III.

A. Results

1) Terms $\frac{E_{D,k}}{E_D}$: Although E_D and $E_{D,k}$ depend on the wavelength, the ratios $\frac{E_{D,k}}{E_D}$ only depend on the scene geometry. As shown in [21], these terms can be reduced to ratios of cosines so they do not depend on the wavelength.

These terms are equal to one when the surface is flat, i.e. when the linear mixing model is valid. This is not the case here, because of the 3D structures, so it is interesting to see which values can be encountered. As an example, Figure 17 shows the histogram of the values of these ratios in our urban image, and here is a summary of the results:

- $\frac{E_{D,k}}{E_D}$ in the linear part of our model: almost all values are between 0 and 5, 54% of values are almost equal to one (with a precision of 0.01). These terms are often different from one, which is not the case when the surface is flat.
- $\frac{E_{D,m}}{E_D}$ in the quadratic part of our model (where m here correspond to elementary surfaces irradiated by the sun): almost all values are between 0.3 and 5, 94% of values are lower than 2. Note that values are high when E_D corresponds to partially shadowed pixels.



(a) in the linear part (b) in the quadratic part

Fig. 17. Histogram of values of $\frac{E_{D,k}}{E_D}$ in the linear part and quadratic part of our model

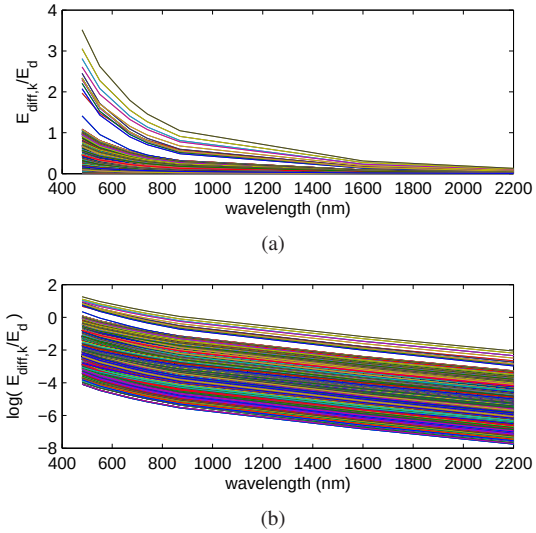
2) Terms $\frac{E_{diff,m}}{E_D}$: Figure 18 graphs $\frac{E_{diff,m}}{E_D}$ vs. the wavelength, for different elementary surfaces m and different pixel positions. Figure 18(a) shows that these terms vary very slowly in the infra-red domain but they significantly depend on the wavelength in the visible domain. Looking at these curves in a semi-log scale (Fig. 18(b)), we can notice that all curves look like translated versions of one curve. This means that, in a linear scale, these curves are proportional. This property can be verified by computing the cross-correlation coefficients between all couples of curves, which are always higher than 0.98. We thus can approximate all our curves by one curve multiplied by a scalar factor: $u_k(\lambda) \simeq a_k v(\lambda)$, with:

- $v(\lambda)$ the curve chosen as a reference.
- $a_k = \dfrac{mean_\lambda(u_k(\lambda)v(\lambda))}{mean_\lambda(v(\lambda)^2)}$.

To check the proportionality of our curves, we also compute the following normalized root mean square error:

$$error = \sqrt{\frac{mean_\lambda((u_k(\lambda) - a_k v(\lambda))^2)}{mean_\lambda(u_k^2(\lambda))}}.$$

We obtain the following results. The maximum encountered error is of 8 % and we have a mean error of 4%, for all curves. The errors are lower than 6% for more than 90% of the curves and lower than 7% for 98% of the curves. The errors obtained



(a)

(b)

Fig. 18. (a) $E_{diff,m}/E_D$ vs. wavelength (nm), (b) $log(E_{diff,m}/E_D)$ vs. wavelength (nm)

with this approximation are thus very low, so we can consider that these curves are almost proportional. This result is very interesting, considering the fact that whatever the materials and the geometric conditions are, the terms $\frac{E_{diff,m}}{E_D}$ have the same spectral behaviour.

B. Resulting simplifications and the obtained model

The above study yields very interesting results: we have separability between the dependency of the irradiance ratios vs. the wavelength and vs. the spatial position, and $E_{d,k}/E_D$ does not depend on λ. We can thus write

$$\frac{E_{D,k}}{E_D}(y,\lambda) = \alpha_k(y) \quad (18)$$

$$\frac{E_{diff,m}}{E_D}(y,\lambda) = \beta_m(y) f(\lambda) \quad (19)$$

This allows us to derive a simplified mixing model from Equation (17)

$$\begin{aligned}\langle\rho\rangle(y,\lambda) &= \sum_{k=1}^{N} S_k(y)\alpha_k(y) \times \rho_k(\lambda) \\ &+ \sum_{k=1}^{N}\sum_{m\in V_k} S_k(y)\frac{g_{m,k}(y)}{\pi}\Delta S \\ &\quad \times(\alpha_m(y)+\beta_m(y)f(\lambda)) \times \rho_k(\lambda)\rho_m(\lambda)\end{aligned} \quad (20)$$

By grouping all coefficients depending only on the spatial position together, we finally obtain

$$\begin{aligned}\langle\rho\rangle(y,\lambda) &= \sum_{k=1}^{N} b_k(y) \times \rho_k(\lambda) \\ &+ \sum_{k=1}^{N}\sum_{m\in V_k} (c_{k,m}(y)+d_{k,m}(y)f(\lambda)) \times \rho_m(\lambda)\rho_k(\lambda)\end{aligned} \quad (21)$$

with $b_k(y) = S_k(y)\alpha_k(y)$,
$c_{k,m}(y) = S_k(y)\frac{g_{m,k}(y)}{\pi}\Delta S\alpha_m(y)$
and $d_{k,m}(y) = S_k(y)\frac{g_{m,k}(y)}{\pi}\Delta S\beta_m(y)$.

This is our general mixing model for urban environments, without notable approximation. It is a *spectrally-variant* linear-quadratic model, because of the term $f(\lambda)$, in the last mixing coefficients, which still depends on the wavelength. Quantifying the error induced on $\langle\rho\rangle$ by this approximation is not easy. However, we can be sure that it is globally lower than 3%, which corresponds to the maximum expected proportion of $\langle\rho\rangle_{env,diff}$ in $\langle\rho\rangle$ (see results in Figure 15 and Table IV).

C. Obtaining the invariant linear-quadratic model

Now, if we want to derive an *invariant* linear-quadratic model in which the sources are the reflectances, we need to neglect $f(\lambda)$ variations. As noticed in Section III, the contribution of $\langle\rho\rangle_{env,diff}$ (so the terms $\frac{E_{diff,m}}{E_D}$) in $\langle\rho\rangle_{env}$ is weak compared to the contribution of $\langle\rho\rangle_{env,D}$ (the terms $\frac{E_{D,m}}{E_D}$). Therefore, the approximation that $f(\lambda)$ is constant is acceptable and can be made without important consequences. We then include $f(\lambda)$ in the definition of the term $\beta_m(y)$.

By grouping together all coefficients depending only on the spatial position, Eq. (21) becomes

$$\langle\rho\rangle(y,\lambda) = \sum_{k=1}^{N} b_k(y)\times\rho_k(\lambda) + \sum_{k=1}^{N}\sum_{m\in V_k} c_{k,m}(y)\times\rho_k(\lambda)\rho_m(\lambda) \quad (22)$$

with here $b_k(y) = S_k(y)\alpha_k(y)$
and $c_{k,m}(y) = S_k(y)\frac{g_{m,k}(y)}{\pi}\Delta S(\alpha_m(y)+\beta_m(y))$.
It must be clear that this simplification can be made without demonstrating the separability of the terms $\frac{E_{diff,m}}{E_D}$ vs. wavelength and position (Section IV-A). The error induced by the simplification made here will be evaluated in the following section, at the pixel level.

V. Towards spectral unmixing

A. Final mixing model adapted to unmixing

To perform spectral unmixing or source separation, we need to have the same sources in all the image, so sources have to be independent of the pixels. In our case, the sources are the reflectances, so the model assumes that each material corresponds to only one reflectance spectrum. In Equation (22), a reflectance is associated with each elementary surface k or m, so we regroup elementary surfaces corresponding to each material together by associating a reflectance ρ_j with each material. Our equation, for one pixel i, thus becomes a sum over all M materials:

$$\langle\rho\rangle_i(\lambda) = \sum_{j=1}^{M} a_j(i)\times\rho_j(\lambda) + \sum_{j=1}^{M}\sum_{\ell=j}^{M} a_{j,\ell}(i)\times\rho_j(\lambda)\rho_\ell(\lambda) \quad (23)$$

with, for example, $a_j(i) = \sum_{k\in D_j} b_k(y)$, with D_j = elementary surfaces composed of material j. We thus have an invariant instantaneous linear-quadratic mixing model, with M sources that are the reflectances of the materials present in the image. This model, already used in the literature for the unmixing (see e.g. [15], [16], [17] and [18]), is now justified by our physical equations.

B. Possible assumptions about the mixing coefficients?

The above linear-quadratic model has been used, in the literature, for unmixing, with different assumptions concerning the mixing coefficients $a_j(i)$ and $a_{j,\ell}(i)$.

In [15], it is considered that the sum of all coefficients ($a_j(i)$ and $a_{j,\ell}(i)$) is equal to one, which does not have a really physical meaning. In [16] different possibilities are proposed for the $a_{j,\ell}(i)$, not justified either. [17] assumes that $a_{j,\ell}(i) = a_j(i)a_\ell(i)$, which means that a surface not seen by the sensor for the considered pixel, can not contribute by reflections on its reflectance, and this is often not true (see Figure 20 and discussed examples below).

Concerning the linear coefficients $a_j(i)$, most works keep the assumption used for the linear model [17] [18], considering that the $a_j(i)$ are the abundances (proportion of each material in a pixel) and then:
$\sum_{j=1}^{M} a_j(i) = 1$ and $0 \le a_j(i) \le 1$.
In Section II, we showed that this assumption (Eq. (12)) is

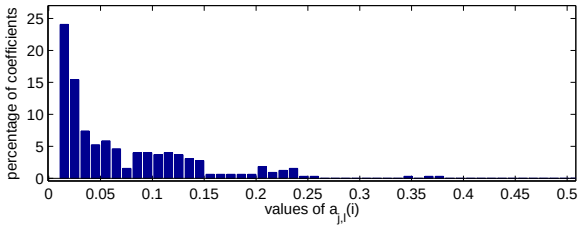


Fig. 19. Histogram of values of $a_{j,\ell}(i)$ - urban image at 5 m resolution

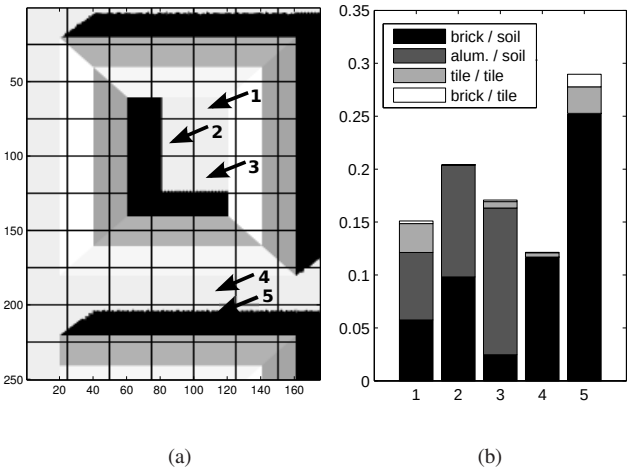
true for our general model, i.e. even if surfaces are not flat and homogeneous. This was confirmed by simulation results in Section III-D3. In fact, we can write for a pixel i:

$$\sum_{j=1}^{M} a_j(i) = \sum_{k=1}^{N} S_k \frac{E_{D,k}}{E_D} = 1$$

For the classical linear model, which is valid when the surface is flat and homogeneous, the terms $\frac{E_{D,k}}{E_D}$ are equal to one, and we obtain the classical sum-to-one assumption for the abundances.

Concerning the $a_{j,\ell}(i)$, related to the quadratic part of the mixing model, Figure 19 shows a histogram of the non-zero coefficients obtained with our urban image at 5 m resolution. It may appear that most values are low (72% below 0.1), but this is normal since our urban image is, in some areas, not very representative of towns: there are no buildings on the edges of the image, so no possible multiple reflections there. Focusing on the previously chosen canyon (rows 10 to 18 - columns 7 and 8), values of $a_{j,\ell}(i)$ have another distribution, which is more representative of urban environments:
- $a_{j,\ell}(i) \in [0, 0.1]$ for 49% of the values,
- $0.1 < a_{j,\ell}(i) < 0.2$ for 16% ,
- $0.2 < a_{j,\ell}(i) < 0.25$ for 35% of them.

Next, we analyse the values of $a_{j,\ell}(i)$ for some chosen pixels in the urban image. To this end, we use the same image as in Section III-D3 (a 5 m image obtained by simulating one zone of our urban scene with a finer triangulation than for the entire scene). Figure 20(a) shows a 0.2 m resolution image of the used scene (image of direct irradiance E_D), with a grid showing the 5 m pixels. This figure highlights the surface orientations towards the sun and the shadowed areas (in black). The five chosen pixels are indicated. Figure 20(b) shows the values of non-zero $a_{j,\ell}(i)$ for these pixels. As shown in the figure, these pixels have been chosen as they correspond to different configurations including: shadow proportion, distance from a wall, surface orientation (roof or soil). We can also see to which materials each coefficient corresponds (to know by which reflectances it is multiplied in the model). The aim again is to see which values can be encountered in an urban image. It is also interesting to see how the position on the scene and the proportion of the shadow can play a role on the values of $a_{j,\ell}(i)$:



(a) (b)

Fig. 20. Values of $a_{j,\ell}(i)$ for 5 chosen pixels in the used urban image (top left corner of the image in Fig. 7(a)): (a) the E_D image, with the pixel locations, (b) values of $a_{j,\ell}(i)$ for the 5 pixels

• Pixel 1: there are contributions due to reflections between the roofs (tiles), to reflections from the wall of brick and from the wall of aluminium. However the total coefficient is not very high since the reflections between roofs are less important than reflections due to walls.
• Pixel 2: it corresponds to a surface more distant from the walls than the surface of pixel 1, and though the coefficient is higher than for pixel 1. This shows again that reflections due to walls give higher contributions than reflections between roofs (pixel 1). This is due essentially to the orientation of the involved surfaces.
• Pixel 3 and 4: pixel 3 has a high coefficient corresponding to the aluminium (wall on the right of the pixel), which is e.g. higher than the coefficient due to the brick for pixel 4, this can be due to the surface orientation towards the sun.
• Pixel 5: the coefficient corresponding to the brick is high compared with other pixels, due to the presence of shadow, although the pixel is not close to the reflecting wall.

The values of $a_{j,\ell}(i)$ are certainly lower than for the terms $a_j(i)$, but many of them cannot be neglected if we intend to perform unmixing with a good precision for urban images.

There is one last point that must be treated now: the possible error on the model induced by the approximation of Section IV-C. It has been assumed that the $a_{j,\ell}(i)$ coefficients do not depend on the wavelength. Note that the above results concerning these coefficients were obtained at the wavelength 670 nm. Figure 21 shows the actual variations of the terms $a_{j,\ell}(i)$ vs. the wavelength. As it can be seen, the coefficients vary at the beginning of the wavelength interval but are almost stable beyond 800 nm. To see the effect of this approximation, we need to remind that assuming that the $a_{j,\ell}(i)$ are constant (vs. the wavelength) corresponds to an approximation of only the term $\langle\rho\rangle_{env,diff}$. This represents globally 3% in the total reflectance (see Figure 15 and Table IV), and corresponds to 6% at most (image C4, for 480 nm). This term is thus globally low compared to the total reflectance, so the considered approximation is sufficient. Therefore, neglecting the variations of terms $a_{j,\ell}(i)$ vs. the wavelength is a reasonable assumption and the proposed model (23) is quite accurate.

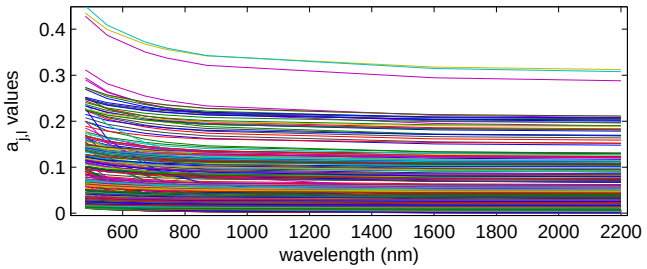


Fig. 21. Values of $a_{j,\ell}(i)$ vs. wavelength - urban image at 5 m resolution

VI. Conclusion

In this paper, a mixing model for reflectances was derived from physical equations based on the radiative transfer theory. A linear-quadratic model was thus obtained, which was validated on simulated urban images. It was demonstrated that the quadratic term cannot be neglected in urban scenes, essentially when there are many buildings and canyons (it can e.g. range from 15% to 20% of the total equivalent reflectances). We also found that the sum-to-one property used in linear models for the coefficients is still true for the coefficients of the linear part of our model. Finally, an invariant linear-quadratic model adapted to unmixing methods was derived, with some additional information concerning the mixing coefficients. All our simulations were done without any noise. Future work will consider synthetic images with noise to analyse its impact on the importance of the terms due to environment (quadratic terms). It would also be interesting to compare this model with the linear one while unmixing simulated urban hyperspectral images. Furthermore, we intend to validate the model with real hyperspectral images. An airborne experiment is planned in the near future over two French cities (Toulouse and Amiens) with the Onera hyperspectral cameras (HYSPEX). Finally, note that this model is true for pixels partially or totally in the sun, and it would be interesting to also derive a model for the completely shadowed pixels.

Acknowledgment

The authors thank HPc-SA/ RayCREATIS, provider of Raybooster software, used in the AMARTIS simulations.

References

[1] U. Heiden, S. Roessner, K. Segl, and H. Kaufmann, "Analysis of spectral signatures of urban surfaces for their identification using hyperspectral hymap data," in *Remote Sensing and Data Fusion over Urban Areas, IEEE/ISPRS Joint Workshop 2001*, 2001, pp. 173 –177.

[2] N. Keshava and J. F. Mustard, "Spectral unmixing," *Signal Processing Magazine, IEEE*, vol. 19, no. 1, pp. 44 – 57, 2002.

[3] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, pp. 650–663, 2004.

[4] M. Parente and A. Plaza, "Survey of geometric and statistical unmixing algorithms for hyperspectral images," in *IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2010, pp. 1 –4.

[5] J. M. P. Nascimento and J. M. Bioucas-Dias, "Vertex Component Analysis: A fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898–910, 2005.

[6] S. Jia and Y. Qian, "Spectral and spatial complexity-based hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 12, pp. 3867–3879, 2001.

[7] R. Heylen, D. Burazerovic, and P. Scheunders, "Fully constrained least squares spectral unmixing by simplex projection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4112 –4122, nov. 2011.

[8] J. Silván-Cárdenas and L. Wang, "Fully constrained linear spectral unmixing: Analytic solution using fuzzy sets," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 11, pp. 3992 –4002, nov. 2010.

[9] X. Liu, W. Xia, B. Wang, and L. Zhang, "An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 2, pp. 757 –772, feb. 2011.

[10] J. Plaza, A. Plaza, R. Perez, and P. Martinez, "Joint linear/nonlinear spectral unmixing of hyperspectral image data," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2007.

[11] W. Liu and E. Y. Wu, "Comparison of non-linear mixture models: Subpixel classification," *Remote Sensing of environment*, vol. 94, no. 2, pp. 145–154, 2005.

[12] U. Kumar, S. K. Raja, C. Mukhopadhyay, and T. V. Ramachandra, "A multi-layer perceptron based non-linear mixture model to estimate class abundance from mixed pixels," in *IEEE Students' Technology Symposium (TechSym)*, 2011, pp. 148 –153.

[13] R. Pu, P. Gong, R. Michishita, and T. Sasagawa, "Spectral mixture analysis for mapping abundance of urban surface components from the terra/aster data," *Remote Sensing of Environment*, vol. 112, no. 3, pp. 939 – 954, 2008.

[14] P. Huard and R. Marion, "Study of non-linear mixing in hyperspectral imagery - a first attempt in the laboratory," in *IEEE 3rd Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing*, 2011.

[15] J. M. P. Nascimento and J. M. Bioucas-Dias, "Nonlinear mixture model for hyperspectral unmixing," in *SPIE Conference on Image and Signal Processing for Remote Sensing XV*, 2009.

[16] B. Somers, K. Cools, S. Delalieux, J. Stuckens, D. V. der Zande, W. W. Verstraeten, and P. Coppin, "Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards," *Remote Sensing of Environment*, vol. 113, no. 6, pp. 1183 – 1193, 2009.

[17] W. Fan, B. Hu, J. Miller, and M. Li, "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated forest hyperspectral data," *International Journal of Remote Sensing*, vol. 30, no. 11, pp. 2951–2962, 2009.

[18] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4153 –4162, 2011.

[19] G. Fontanilles, X. Briottet, S. Fabre, S. Lefebvre, and P.-F. Vandenhaute, "Aggregation process of optical properties and temperature over heterogeneous surfaces in infrared domain," *Applied Optics*, vol. 49, no. 24, pp. 4655–4669, 2010.

[20] F. Becker and Z. Li, "Surface temperature and emissivity at various scales: Definition, measurement and related problems," *Remote Sensing Reviews*, vol. 12, no. 3-4, pp. 225–253, 1995.

[21] C. Miesch, X. Briottet, Y. H. Kerr, and F. Cabot, "Radiative transfer solution for rugged and heterogeneous scene observations," *Applied Optics*, vol. 39, pp. 6830–6846, 2000.

[22] M. von Schönermark, B. Geiger, and H. P. Röser, *Reflection properties of vegetation and soil: with a BRDF data base*. Berlin, Germany: Wissenschaft und Technik Verlag, 2004, ch. 2.

[23] E. F. Vermote, D. Tanre, J. L. Deuze, M. Herman, and J.-J. Morcette, "Second simulation of the satellite signal in the solar spectrum, 6S: an overview," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 3, pp. 675–686, may 1997.

[24] "Memoires data base," http://www.onera.fr/dota/memoires/index.php.

[25] J. A. Gómez, E. de Miguel, O. G. de la Cámara, and A. Fernández-Renau, "Status of the INTA AHS sensor," in *5th EARSeL Workshop on Imaging Spectroscopy*, 2007.

[26] P. Comon and C. Jutten, *Handbook of Blind Source Separation, Independent Component Analysis and Applications*. Oxford UK, Burlington USA: Academic Press, 2010.

[27] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis*. John Wiley & Sons Inc, a volumee in Series on Adaptive and Learning Systems for Signal Processing, Communications and Control, 2001.

[28] P. D. O'Grady, B. A. Pearlmutter, and S. T. Rickard, "Survey of sparse and non-sparse methods in source separation," *International Journal of Imaging Systems and Technology*, vol. 15, no. 1, pp. 18–33, 2005.
[29] C. Jutten and J. Karhunen, "Advances in blind source separation (BSS) and independent component analysis (ICA) for nonlinear mixtures." *International Journal of Neural Systems*, vol. 14, no. 5, pp. 267–292, 2004.

Inès Meganem received the Engineering degree in electronics and signal processing from ENSEEIHT, Toulouse, France, and the M.Sc. degree in signal processing from the National Polytechnical Institute of Toulouse, both in 2009 and, in December 2012, the Ph.D. degree in signal and image processing from the University of Toulouse. She is currently a teaching and research assistant in signal and image processing at the University of Toulouse. Her main research interests concern the study of Blind Source Separation methods for the analysis of hyperspectral images, with applications to astrophysics and spectral unmixing in remote sensing.

Philippe Déliot obtained a physical engineer grade in 1987. He has been working at ONERA (French Aerospace Lab) since 1989 in the applied and theoretical optics department. He was firstly in charge of infrared sensor conception and realization. Since 2000, he has been working on subjects in relation with image quality in remote sensing and particularly with calibration of airborne system (ONERA-IGN-CNES PELICAN) and development of defocusing correction method for sensor after launch (SPOT, PLEIADES). He is currently involved in hyperspectral acquisition and calibration in lab and airborne system implementation (SETHI). Its main interests are in relation with hyperspectral image processing and its application to security and environment.

Xavier Briottet received his Engineer Doctorate degree in Electonics from the "Ecole Nationale Supérieure de l'Aéronautique et de l'Espace" (Toulouse, France) in 1986. In 1987, he became assistant professor in signal and image processing at the "Ecole Nationale Supérieure d'Electricité". Since 1988, he is working in remote sensing at ONERA in the Theoretical and Applied Optics Department (DOTA in Toulouse, France). He received his "Habilitation à Diriger les Recherches" from Paul Sabatier University (Toulouse, France) in 1999. He is now director of research in charge of the scientific politics in the Theoretical and Applied Optics Department. His main research interests are radiative transfer modelling and calibration of airborne and satellite image (SPOT 1-5, Polder, Vegetation, Meris) sensors in the whole optical domain using on-board devices or terrestrial targets. In radiative transfer, his studies aim to model the aggregation processes in the reflective and thermal domain over heterogeneous and rough surfaces taking into account its directional effects. He is contributing to 2 future space missions: MISTIGRI (multispectral thermal infrared sensor), HYPXIM (hyperspectral sensor in the reflective domain)

Yannick Deville was born in Lyon, France, in 1964. He graduated from the Ecole Nationale Supérieure des Télécommunications de Bretagne (Brest, France) in 1986. He received the D.E.A and Ph.D. degrees, both in Microelectronics, from the University of Grenoble (France), in 1986 and 1989 respectively. From 1986 to 1997, he was a Research Scientist at Philips Research Labs (Limeil, France). His investigations during this period concerned various fields, including GaAs integrated microwave RC active filters, VLSI cache memory architectures and replacement algorithms, neural network algorithms and applications, and nonlinear systems. Since 1997, he has been a Professor at the University of Toulouse (France). From 1997 to 2004, he was with the Acoustics lab of that University. Since 2004, he has been with the Astrophysics lab in Toulouse, which is part of the University and of the French National Center for Scientific Research (CNRS). His current major research interests include signal and image processing, higher-order statistics, time-frequency analysis, neural networks, and especially Blind Source Separation methods (including Independent or Sparse Component Analysis) and their applications to remote sensing, astrophysics, quantum information processing, acoustics and communication/electromagnetic signals.

Shahram Hosseini was born in Shiraz, Iran, in 1968. He received the B.Sc. and M.Sc. degrees in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 1991 and 1993, respectively, and the Ph.D. degree in signal processing from the Institut National Polytechnique, Grenoble, France, in 2000. He is currently an Associate Professor at the Université Paul Sabatier Toulouse 3, Toulouse, France. His research interests include BSS, artificial neural networks, and adaptive signal processing.