



Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : <http://oatao.univ-toulouse.fr/>
Eprints ID : 5662

To link to this article : DOI: 10.3166/ts.27.79-108
URL : <http://dx.doi.org/10.3166/ts.27.79-108>

To cite this version :

Dobigeon, Nicolas and Moussaoui, Said and Coulon, Martial and Tourneret, Jean-Yves and Hero, Alfred O. *Algorithmes bayésiens pour le démixage supervisé, semi-supervisé et non-supervisé d'images hyperspectrales*. (2010) *Traitement du signal*, vol. 27 (n° 1). pp. 79-108. ISSN 0765-0019

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes.diff.inp-toulouse.fr

Algorithmes bayésiens pour le démixage supervisé, semi-supervisé et non-supervisé d'images hyperspectrales

Nicolas Dobigeon* — Saïd Moussaoui† — Martial Coulon* —
Jean-Yves Tourneret* — Alfred O. Hero‡

* Université de Toulouse, IRIT/INP-ENSEEIH
2, rue Charles Camichel – B.P. 7122, F-31071 Toulouse Cedex 7

† IRCCyN - CNRS UMR 6597, ECN
1, rue de la Noë – B.P. 92101, F-44321 Nantes Cedex 3

‡ University of Michigan, Department of EECS
1301 Beal Avenue – Ann Arbor, MI 48109-2122, USA
E-mail : nicolas.dobigeon@enseeiht.fr

RÉSUMÉ. Cet article présente des algorithmes totalement bayésiens pour le démixage d'images hyperspectrales. Chaque pixel de l'image est décomposé selon une combinaison de spectres de références pondérés par des coefficients d'abondances selon un modèle de mélange linéaire. Dans un cadre supervisé, nous supposons connus les spectres de références. Le problème consiste alors à estimer les coefficients du mélange sous des contraintes de positivité et d'additivité. Une loi a priori adéquate est choisie pour ces coefficients qui sont estimés à partir de leur loi a posteriori. Un algorithme de Monte Carlo par chaîne de Markov (MCMC) est développé pour approcher les estimateurs. Dans un cadre semi-supervisé, les spectres participant au mélange seront supposés inconnus. Nous faisons l'hypothèse qu'ils appartiennent à une bibliothèque spectrale. Un algorithme MCMC à sauts réversibles permet dans ce cas de résoudre le problème de sélection de modèle. Enfin, dans un dernier cadre d'étude, les algorithmes précédents sont étendus au démixage non-supervisé d'images hyperspectrales, c'est-à-dire au problème d'estimation conjointe des spectres et des coefficients de mélange. Ce problème de séparation aveugle de sources est résolu dans un sous-espace approprié.

ABSTRACT. This article describes fully Bayesian algorithms to unmix hyperspectral images. Each pixel of the hyperspectral image is decomposed as a combination of pure endmember spectra following the linear mixing model. In a supervised context, the endmembers are assumed to be known. The unmixing problem consists of estimating the mixing coefficients under positivity

and additivity constraints. An appropriate distribution is chosen as prior distribution for these coefficients, that are estimated thanks to their posterior distribution. A Markov chain Monte Carlo (MCMC) algorithm is developed to approximate the estimators. In a semi-supervised framework, the spectra involved in the mixtures are assumed to be unknown. They are supposed to belong to a known spectral library. A reversible-jump MCMC algorithm allows one to solve the resulting model selection model. Finally, in a final step, the previous algorithms are extended to handle the unsupervised unmixing problem, i.e., to estimate the endmembers and the mixing coefficients jointly. This blind source separation problem is solved in a lower-dimensional space, which effectively reduces the number of degrees of freedom of the unknown parameters.

MOTS-CLÉS : Imagerie hyperspectrale, démixage linéaire, inférence bayésienne, méthodes de Monte Carlo par chaînes de Markov

KEYWORDS: Hyperspectral imagery, linear unmixing, Bayesian inference, Markov chain Monte Carlo methods

1. Introduction

Depuis une vingtaine d'années déjà, l'imagerie hyperspectrale a démontré tout son intérêt dans un grand nombre de recherches consacrées à l'observation de la Terre. Basée sur l'acquisition d'une même scène dans un grand nombre de bandes spectrales, i.e., à différentes longueurs d'onde, elle a par exemple permis d'établir, dans le cadre d'études minéralogiques, des cartes de classification à grande échelle, évitant de longues et fastidieuses campagnes de prélèvements au sol (Jackson *et al.*, 2002; Rellier *et al.*, 2004). À l'heure où les enjeux environnementaux sont sur le devant de la scène, elle est capable de fournir des informations déterminantes relatives à des paramètres macroscopiques renseignant sur l'état des éco-systèmes, des végétaux en particuliers. Naturellement, le prix à payer pour bénéficier de la richesse d'informations contenues dans ces images est de développer de nouvelles méthodes d'analyse capables d'exploiter efficacement les données fournies par les imageurs hyperspectraux. Depuis l'acquisition des premières images hyperspectrales, le problème de démixage spectral a fait l'objet d'un nombre considérable de travaux, non seulement dans la communauté scientifique de la télédétection, mais aussi dans celle du traitement du signal et des images. Résoudre ce problème permet de répondre en effet à divers enjeux importants en imagerie hyperspectrale comme la classification de scène (Chang, 2003), la quantification de matériaux (Plaza *et al.*, 2005) ou encore la détection sub-pixellique (Manolakis *et al.*, 2001). Le démixage spectral consiste à décomposer chaque pixel de la scène observée en un ensemble de spectres de référence, appelés *pôles de mélange*, et à estimer les proportions, ou *abondances* de chacun d'eux au sein des pixels (Keshava *et al.*, 2002). Pour décrire formellement le mélange qui lie les pixels observés et les spectres de référence, le modèle le plus fréquemment rencontré est un modèle linéaire. Il constitue une bonne approximation du modèle non-linéaire proposé par Hapke (1981) dans le domaine réflectif du visible au proche infrarouge ($0.4\mu\text{m}$ à $2.5\mu\text{m}$) (Johnson *et al.*, 1983).

Ainsi qu'il l'est noté par Keshava *et al.* (2002), le démixage spectral se décompose principalement en deux étapes : l'identification des pôles de mélange et l'estimation des abondances. Dans la première étape de l'analyse, les spectres des composants purs présents dans l'image sont identifiés à l'aide d'un algorithme d'extraction de pôles de mélange (EPM). Parmi les méthodes d'EPM les plus célèbres, nous citerons l'indice de pureté de pixel (PPI) (Boardman, 1993), l'algorithme N-FINDR (Winter, 1999) et plus récemment l'algorithme VCA (Nascimento *et al.*, 2005a) qui s'attachent à retrouver les sommets du plus grand simplexe contenu dans les données représentées dans l'espace hyperspectral. Une hypothèse commune à ces algorithmes est qu'ils requièrent la présence de pixels purs dans l'image observée. Inversement, Craig (1994) et Bowles *et al.* (1995) ont proposé des transformées de volume minimal pour trouver le plus petit simplexe qui contient tous les pixels. La deuxième étape dans le démixage spectral est consacrée à l'estimation des proportions du mélange. Celles-ci doivent par définition satisfaire des contraintes de positivité et d'additivité. Les algorithmes principalement proposés dans la littérature ont recours à des tech-

niques d'optimisation sous contraintes (Heinz *et al.*, 2001; Tu *et al.*, 1998; Theys *et al.*, 2009).

Nous proposons dans cet article des algorithmes d'estimation bayésienne permettant de résoudre le problème de démixage linéaire dans des contextes supervisé, semi-supervisé et non-supervisé. Dans le premier cadre de l'étude, nous supposons que les pôles de mélange ont été préalablement identifiés à l'aide d'un algorithme EPM ou par une connaissance *a priori* de la scène observée. L'algorithme de démixage présenté consistera donc à réaliser l'étape d'inversion, i.e., à estimer les coefficients d'abondances sous les contraintes de positivité et d'additivité. Dans un deuxième temps, nous nous affranchirons partiellement de la connaissance des spectres des composants purs (i.e., pôles de mélange), en supposant que ceux-ci sont en partie inconnus. En revanche, le nombre et la nature des spectres présents au sein de chaque pixel seront supposés inconnus. Le problème d'estimation sous contrainte des abondances sera donc couplé à un problème de sélection de modèle. Enfin, nous traiterons le démixage spectral dans un contexte totalement non-supervisé, pour lequel à la fois les spectres des composants purs et leurs proportions respectives devront être estimés.

Dans chacun de ces cadres d'étude, nous verrons que l'estimation bayésienne permet de manière habile d'inclure dans le modèle les contraintes inhérentes au problème. En effet, le choix de lois *a priori* adéquates permet de prendre en compte les contraintes de positivité et d'additivité des abondances, ainsi que la positivité des spectres.

2. Modèle de mélange linéaire et position du problème

Considérons P pixels d'une image hyperspectrale acquise dans L bandes spectrales. Selon le modèle de mélange linéaire, décrit par exemple dans (Keshava *et al.*, 2002), le spectre $\mathbf{y}_p = [y_{p,1}, \dots, y_{p,L}]^T$ du $p^{\text{ème}}$ pixel ($p = 1, \dots, P$) s'exprime comme la combinaison linéaire de R signatures spectrales \mathbf{m}_r entachée d'un bruit additif \mathbf{n}_p :

$$\mathbf{y}_p = \sum_{r=1}^R \mathbf{m}_r a_{p,r} + \mathbf{n}_p, \quad [1]$$

où $\mathbf{m}_r = [m_{r,1}, \dots, m_{r,L}]^T$ est le spectre pur caractéristique du $r^{\text{ème}}$ matériau et $a_{p,r}$ est l'abondance du $r^{\text{ème}}$ matériau dans le $p^{\text{ème}}$ pixel. De plus, dans [1], $\mathbf{n}_p = [n_{p,1}, \dots, n_{p,L}]^T$ est une séquence de bruit que l'on supposera i.i.d. de loi normale centrée et de matrice de covariance¹ $\Sigma_n = \sigma^2 \mathbf{I}_L$, où \mathbf{I}_L est la matrice identité de taille $L \times L$

$$\mathbf{n}_p | \sigma^2 \sim \mathcal{N}(\mathbf{0}_L, \Sigma_n). \quad [2]$$

¹Le modèle proposé peut aisément être étendu à des modèles de bruit plus complexes, en suivant par exemple la stratégie adoptée dans (Dobigeon *et al.*, 2008b).

En raison de considérations physiques décrites dans (Keshava *et al.*, 2002), les vecteurs d'abondances $\mathbf{a}_p = [a_{p,1}, \dots, a_{p,R}]^T$ dans [1] vérifient les contraintes de positivité et d'additivité suivantes

$$\begin{cases} a_{p,r} \geq 0, \forall r = 1, \dots, R, \\ \sum_{r=1}^R a_{p,r} = 1. \end{cases} \quad [3]$$

En d'autres termes, les p vecteurs d'abondances appartiennent à l'espace

$$\mathcal{A} = \{\mathbf{a} : \|\mathbf{a}\|_1 = 1 \text{ et } \mathbf{a} \succeq \mathbf{0}\}, \quad [4]$$

où $\|\cdot\|_1$ est la norme ℓ_1 telle que $\|\mathbf{x}\|_1 = \sum_i |x_i|$, et $\mathbf{a} \succeq \mathbf{0}$ résume l'ensemble des inégalités $\{a_r \geq 0\}_{r=1, \dots, R}$. De plus, les signatures spectrales $m_{r,l}$ correspondent à des mesures de réflectance et doivent donc satisfaire les contraintes de positivité

$$m_{r,l} \geq 0, \forall r = 1, \dots, R, \forall l = 1, \dots, L. \quad [5]$$

Si l'on considère tous les pixels de l'image hyperspectrale, le système d'équations [1] s'écrit sous forme matricielle

$$\mathbf{Y} = \mathbf{M}\mathbf{A} + \mathbf{N} \quad [6]$$

où \mathbf{Y} est une matrice de taille $L \times P$ contenant toutes les observations associées aux pixels de l'image, \mathbf{M} est une matrice de taille $L \times R$ contenant les signatures spectrales, \mathbf{A} est une matrice de taille $R \times P$ contenant les abondances et \mathbf{N} est une matrice de taille $L \times P$ contenant les bruits associés aux observations

$$\begin{aligned} \mathbf{Y} &= [\mathbf{y}_1, \dots, \mathbf{y}_P], & \mathbf{M} &= [\mathbf{m}_1, \dots, \mathbf{m}_R], \\ \mathbf{A} &= [\mathbf{a}_1, \dots, \mathbf{a}_P], & \mathbf{N} &= [\mathbf{n}_1, \dots, \mathbf{n}_P]. \end{aligned} \quad [7]$$

Nous proposons dans cet article une approche bayésienne pour estimer dans un premier temps les coefficients d'abondances sous les contraintes [3] lorsque les pôles de mélange sont connus. Dans les deuxième et troisième cadres d'étude, les spectres des composants purs seront supposés partiellement ou totalement inconnus.

3. Cadre supervisé : les pôles de mélange sont connus

Lorsque les pôles de mélange sont parfaitement connus, le problème de démixage linéaire se ramène à une étape d'inversion, c'est-à-dire à l'estimation sous contraintes des coefficients d'abondances. Il est assez naturel de formuler ce problème d'estimation comme un problème de régression linéaire sous contraintes dont la résolution peut être envisagée dans un cadre bayésien. En effet, les modèles bayésiens sont particulièrement adaptés dans cette situation puisque les contraintes sont alors naturellement prises en compte dans la définition des lois *a priori* des paramètres inconnus. De nombreux exemples de contraintes ont été récemment étudiés dans la littérature : monotonie (Chen *et al.*, 1996), positivité (Moussaoui *et al.*, 2006) ou parcimonie (Blumensath

et al., 2007; Févotte *et al.*, 2006). Les contraintes inhérentes à l'imagerie hyperspectrale sont la positivité et l'additivité des abondances, comme expliqué au paragraphe 2. Nous décrivons dans ce qui suit le modèle bayésien proposé pour résoudre le problème de démixage spectral supervisé. Notons que dans ce cadre supervisé, le démixage spectral sera effectué pixel par pixel. Afin d'alléger les notations, la dépendance marquée par un indice p des quantités \mathbf{y}_p , \mathbf{a}_p , \mathbf{c}_p liées au pixel p sera donc omise.

3.1. Modèle bayésien

3.1.1. Vraisemblance

Le modèle de mélange linéaire défini par [1] et les propriétés statistiques [2] du vecteur de bruit \mathbf{n} conduisent à une distribution gaussienne pour le spectre observé du $p^{\text{ème}}$ pixel :

$$\mathbf{y}|\mathbf{a}, \sigma^2 \sim \mathcal{N}(\mathbf{M}\mathbf{a}, \sigma^2\mathbf{I}_L). \quad [8]$$

Par conséquent, la vraisemblance du vecteur \mathbf{y} peut s'écrire

$$f(\mathbf{y}|\mathbf{a}, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{L}{2}} \exp\left[-\frac{\|\mathbf{y} - \mathbf{M}\mathbf{a}\|^2}{2\sigma^2}\right], \quad [9]$$

où $\|\mathbf{x}\| = (\mathbf{x}^T\mathbf{x})^{\frac{1}{2}}$ est la norme ℓ_2 du vecteur \mathbf{x} .

3.1.2. Lois a priori des paramètres

Lorsque les spectres des composants purs $\mathbf{m}_1, \dots, \mathbf{m}_R$ sont connus, le vecteur des paramètres inconnus noté $\boldsymbol{\theta}$ se compose d'un vecteur d'abondances et de la variance du bruit $\boldsymbol{\theta} = \{\mathbf{a}, \sigma^2\}$.

3.1.2.1. Coefficients d'abondances

Pour chaque pixel p , grâce aux contraintes d'additivité énoncées dans [3], le vecteur d'abondances \mathbf{a} peut être réécrit²

$$\mathbf{a} = \begin{bmatrix} \mathbf{c} \\ a_R \end{bmatrix} \quad \text{avec} \quad \mathbf{c} = \begin{bmatrix} a_1 \\ \vdots \\ a_{R-1} \end{bmatrix}, \quad [10]$$

et $a_R = 1 - \sum_{r=1}^{R-1} a_r$. Suivant le modèle proposé par Dobigeon *et al.* (2008a), la loi *a priori* choisie pour \mathbf{c} est une loi uniforme définie sur le simplexe \mathcal{S}

$$\mathcal{S} = \{\mathbf{c}; \|\mathbf{c}\|_1 \leq 1 \text{ and } \mathbf{c} \succeq \mathbf{0}\}. \quad [11]$$

²Par souci de simplification d'écriture, nous exprimerons toujours la dernière composante de \mathbf{a} en fonction des autres. En pratique, et notamment dans l'algorithme introduit dans la section suivante, la composante mise de côté peut être choisie aléatoirement à chaque itération de l'échantillonneur de Gibbs.

Choisir cette loi *a priori* pour \mathbf{c} est totalement équivalent à choisir pour le vecteur complet d'abondances \mathbf{a} une loi de Dirichlet $\mathcal{D}(1, \dots, 1)$, i.e., une loi uniforme sur l'ensemble \mathcal{A} des valeurs possibles de \mathbf{a} (ensemble défini par [4]) (Robert, 2007, Appendix A). En outre, cette reparamétrisation se montrera particulièrement adaptée à l'algorithme de Gibbs qui sera développé un peu plus loin.

3.1.2.2. Variance du bruit

Une loi conjuguée inverse-gamma est choisie comme loi *a priori* pour la variance du bruit σ^2

$$\sigma^2 | \nu, \gamma \sim \mathcal{IG}\left(\frac{\nu}{2}, \frac{\gamma}{2}\right), \quad [12]$$

où $\mathcal{IG}\left(\frac{\nu}{2}, \frac{\gamma}{2}\right)$ est une loi inverse-gamma de paramètres $\frac{\nu}{2}$ et $\frac{\gamma}{2}$. Cette loi a été utilisée avec succès dans de nombreux travaux de la littérature parmi lesquels ceux de Puns-kaya *et al.* (2002) et Dobigeon *et al.* (2007b). Par ailleurs, comme dans les références sus-citées, l'hyperparamètre ν sera fixé à $\nu = 2$.

Par ailleurs, γ est un hyperparamètre supposé inconnu, aléatoire pour lequel une loi non-informative de Jeffreys est proposée comme loi *a priori* (Jeffreys, 1961)

$$f(\gamma) \propto \frac{1}{\gamma} \mathbf{1}_{\mathbb{R}^+}(\gamma), \quad [13]$$

où \propto signifie "proportionnel à".

3.1.3. Loi *a posteriori*

La loi *a posteriori* du vecteur des paramètres inconnus $\boldsymbol{\theta} = \{\mathbf{c}, \sigma^2\}$ est calculée à partir du modèle hiérarchique suivant

$$f(\boldsymbol{\theta} | \mathbf{y}) \propto \int f(\mathbf{y} | \boldsymbol{\theta}) f(\boldsymbol{\theta} | \gamma) f(\gamma) d\gamma, \quad [14]$$

où $f(\mathbf{y} | \boldsymbol{\theta})$ et $f(\gamma)$ sont définis dans [9] et [13], respectivement. En supposant l'indépendance *a priori* entre σ^2 and \mathbf{c} , i.e. $f(\boldsymbol{\theta} | \gamma) = f(\mathbf{c}) f(\sigma^2 | \gamma)$, l'hyperparamètre γ peut être intégré dans la loi jointe $f(\boldsymbol{\theta}, \gamma | \mathbf{y})$ suivant [14], ce qui conduit à

$$f(\mathbf{c}, \sigma^2 | \mathbf{y}) \propto \frac{1}{\sigma^{L+2}} \exp\left[-\frac{\|\mathbf{y} - \mathbf{M}\mathbf{a}\|^2}{2\sigma^2}\right] \mathbf{1}_S(\mathbf{c}). \quad [15]$$

Nous introduisons dans la section suivante un échantillonneur de Gibbs qui permet de générer des échantillons distribués suivant la loi jointe $f(\mathbf{c}, \sigma^2 | \mathbf{y})$.

3.2. Échantillonneur de Gibbs

Des échantillons (notés $\cdot^{(t)}$ où t est l'indice de l'itération) peuvent être générés suivant $f(\mathbf{c}, \sigma^2 | \mathbf{y})$ à l'aide d'un échantillonneur de Gibbs décrit ci-dessous. Celui-ci génère successivement des échantillons selon les lois conditionnelles $f(\mathbf{c} | \sigma^2, \mathbf{y})$

et $f(\sigma^2|\mathbf{c}, \mathbf{y})$. Les étapes principales de cet algorithme sont détaillées ci-dessous et sont résumées par l'algorithme 1. Le lecteur intéressé pourra se reporter à (Robert *et al.*, 1999) pour plus de détails sur les méthodes MCMC.

Algorithme 1 Échantillonneur de Gibbs pour le démélange spectral **supervisé**

- 1: *% Initialisation*
 - 2: Échantillonner les paramètres $\tilde{\sigma}^{2(0)}$ et $\tilde{\mathbf{a}}^{(0)}$ à partir des lois *a priori* définies à la section 3.1.2,
 - 3: *% Itérations*
 - 4: **pour** $t = 1, 2, \dots$, **faire**
 - 5: Échantillonner $\tilde{\mathbf{c}}^{(t)}$ selon la loi [18],
 - 6: Échantillonner $\tilde{\sigma}^{2(t)}$ selon la loi [19],
 - 7: **fin pour**
-

3.2.1. Échantillonnage suivant $f(\mathbf{c}|\sigma^2, \mathbf{y})$

La loi *a posteriori* conditionnelle du vecteur partiel d'abondances est

$$f(\mathbf{c}|\sigma^2, \mathbf{y}) \propto \exp\left[-\frac{(\mathbf{c} - \mathbf{v})^T \boldsymbol{\Sigma}^{-1} (\mathbf{c} - \mathbf{v})}{2}\right] \mathbf{1}_{\mathcal{S}}(\mathbf{c}), \quad [16]$$

où

$$\begin{cases} \boldsymbol{\Sigma} = \left[(\mathbf{M}_{\cdot R} - \mathbf{m}_R \mathbf{1}_{R-1}^T)^T \boldsymbol{\Sigma}_n^{-1} (\mathbf{M}_{\cdot R} - \mathbf{m}_R \mathbf{1}_{R-1}^T) \right]^{-1}, \\ \mathbf{v} = \boldsymbol{\Sigma} \left[(\mathbf{M}_{\cdot R} - \mathbf{m}_R \mathbf{1}_{R-1}^T)^T \boldsymbol{\Sigma}_n^{-1} (\mathbf{y} - \mathbf{m}_R) \right], \end{cases} \quad [17]$$

avec $\boldsymbol{\Sigma}_n^{-1} = \frac{1}{\sigma^2} \mathbf{I}_L$, $\mathbf{1}_{R-1} = [1, \dots, 1]^T \in \mathbb{R}^{R-1}$ et où $\mathbf{M}_{\cdot R}$ est la matrice \mathbf{M} dont la $R^{\text{ème}}$ colonne a été supprimée. Par conséquent, le vecteur $\mathbf{c}|\sigma^2, \mathbf{y}$ est distribué suivant une loi normale multivariée tronquée au simplexe \mathcal{S} définie par [11]

$$\mathbf{c}|\sigma^2, \mathbf{y} \sim \mathcal{N}_{\mathcal{S}}(\mathbf{v}, \boldsymbol{\Sigma}). \quad [18]$$

Générer des échantillons distribués suivant cette loi normale tronquée peut s'effectuer en suivant la stratégie décrite dans (Dobigeon *et al.*, 2007a).

3.2.2. Échantillonnage suivant $f(\sigma^2|\mathbf{c}, \mathbf{y})$

En étudiant la loi jointe $f(\sigma^2, \mathbf{c}|\mathbf{y})$, on peut déterminer que la loi conditionnelle de $\sigma^2|\mathbf{c}, \mathbf{y}$ est la loi inverse-gamma suivante

$$\sigma^2|\mathbf{c}, \mathbf{y} \sim \mathcal{IG}\left(\frac{L}{2}, \frac{\|\mathbf{y} - \mathbf{M}\mathbf{a}\|^2}{2}\right). \quad [19]$$

3.3. Résultats de simulation sur données synthétiques

Pour illustrer les performances de l'algorithme, un mélange synthétique de trois composants purs est tout d'abord considéré. Ces signatures spectrales sont issues de la bibliothèque distribuée avec le logiciel ENVI (RSI (Research Systems Inc.), 2003, p. 1035) et sont caractéristiques d'une scène dans un environnement urbain ou péri-urbain : béton de construction, herbe verte et terre grasse. Les proportions du mélange sont définies par $a_1 = 0.3$, $a_2 = 0.6$ et $a_3 = 0.1$. Le spectre observé a été entaché d'un bruit additif gaussien de variance $\sigma^2 = 0.025$, ce qui correspond à un rapport signal sur bruit de $RSB \approx 15\text{dB}$ où $RSB = L^{-1}\sigma^{-2} \left\| \sum_{r=1}^R \mathbf{m}_r a_r \right\|^2$. Les pôles de mélange et le spectre résultant sont représentés sur la Fig. 1.

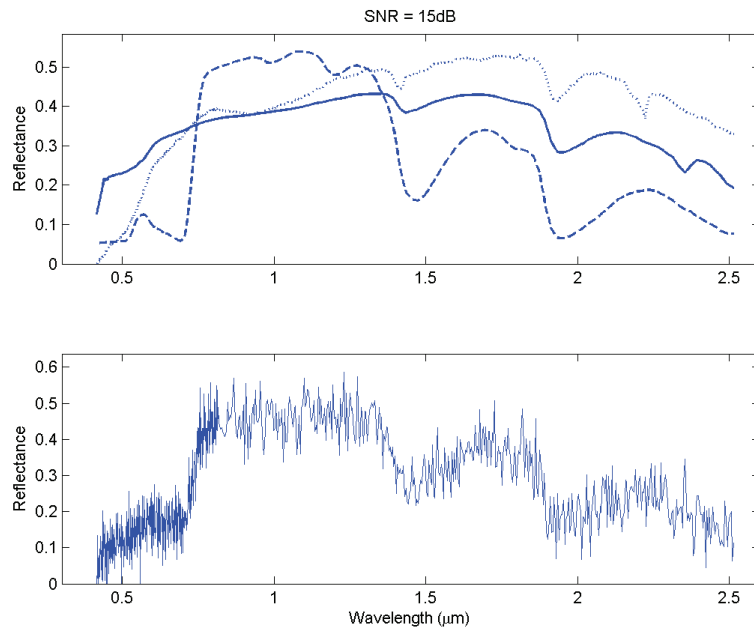


Figure 1. Haut : spectres des pôles de mélange : béton (trait plein), herbe (tirets), terre (pointillés). Bas : spectre du pixel observé pour $RSB \approx 15\text{dB}$.

La Fig. 2 montre les lois *a posteriori* des coefficients d'abondances a_r ($r = 1, 2, 3$) estimées par l'algorithme pour $N_{MC} = 20000$ itérations (dont $N_{bi} = 100$ itérations de chauffage). Ces lois sont bien en accord avec les vraies valeurs de ces coefficients $\mathbf{a} = [0.3, 0.6, 0.1]^T$. À titre de comparaison, les résultats obtenus par l'algorithme FCLS (Chang *et al.*, 2001; Heinz *et al.*, 2001) sont également représentés sur cette même figure pour N_{MC} simulations de Monte Carlo.

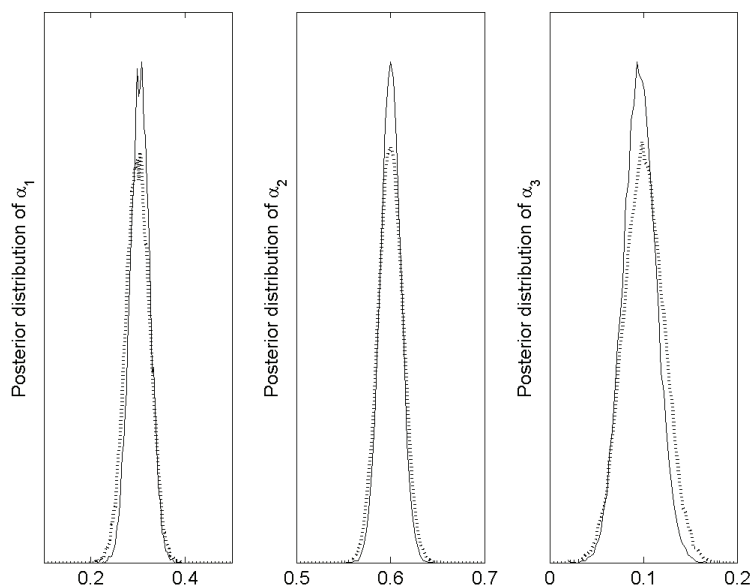


Figure 2. Lois a posteriori des coefficients d'abondances $[a_1, a_2, a_3]^T$ estimées par l'algorithme proposé (lignes continues) et histogrammes des valeurs estimées par l'algorithme FCLS (tirets).

3.4. Résultats de simulation sur données réelles

Ce paragraphe présente l'analyse d'une image hyperspectrale qui a été largement utilisée dans la communauté (Christophe *et al.*, 2005; Chen, 2005; Tang *et al.*, 2004; Akgun *et al.*, 2005). Cette image représentée sur la Fig. 3 est initialement composée de 189 bandes spectrales (après avoir enlevé les bandes d'absorption de la vapeur d'eau). Elle a été acquise en 1997 par le spectro-imageur AVIRIS au dessus de Moffett Field en Californie (voir AVIRIS Free Data (2006) pour plus de détails). Elle se compose d'un vaste point d'eau et d'une zone côtière. L'algorithme de démixage spectral bayésien a été appliqué sur une scène de taille 50×50 . La partie de l'image analysée est représentée en couleurs naturelles composites sur la Fig. 3.

3.4.1. Identification des pôles de mélange

Nous déterminons dans un premier temps les matériaux purs présents dans l'image. En absence de connaissances *a priori* concernant la scène à analyser, un algorithme EPM couplé avec une étape de réduction de dimension a été utilisé pour identifier les pôles de mélange. Une analyse en composantes principales (ACP) permet



Figure 3. Image hyperspectrale réelle de Moffett Field acquise par AVIRIS en 1997 (à gauche) et la région d'intérêt représentée en couleurs naturelles composites (à droite).

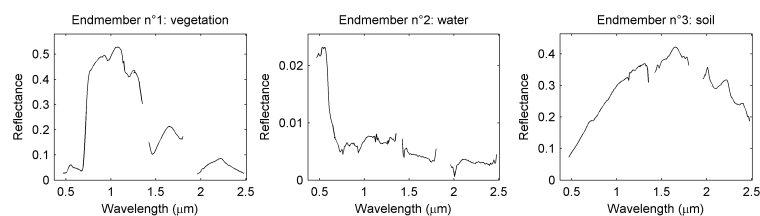


Figure 4. Les $R = 3$ pôles de mélange obtenus par l'algorithme N-FINDR.

de réduire l'espace de travail et de connaître le nombre de pôles de mélange présents dans la scène comme expliqué par Keshava *et al.* (2002). Dans ce sous-espace d'intérêt, les pôles de mélange sont identifiés comme les sommets d'un simplexe à l'aide de l'algorithme N-FINDR (Winter, 1999). Les $R = 3$ matériaux estimés correspondent à la végétation, à l'eau et au sol nu et sont représentés sur la Fig. 4

3.4.2. Estimation des abondances

L'algorithme de démixage spectral présenté aux paragraphes 3.1 et 3.2 a été appliqué sur chaque pixel de l'image hyperspectrale en utilisant les pôles de mélange identifiés au paragraphe 3.4.1. Les cartes d'abondances estimées par l'algorithme proposé pour les $R = 3$ matériaux sont représentées sur la figure 5 (en haut). Notons qu'un pixel blanc (respectivement noir) sur cette figure indique une forte (respec-

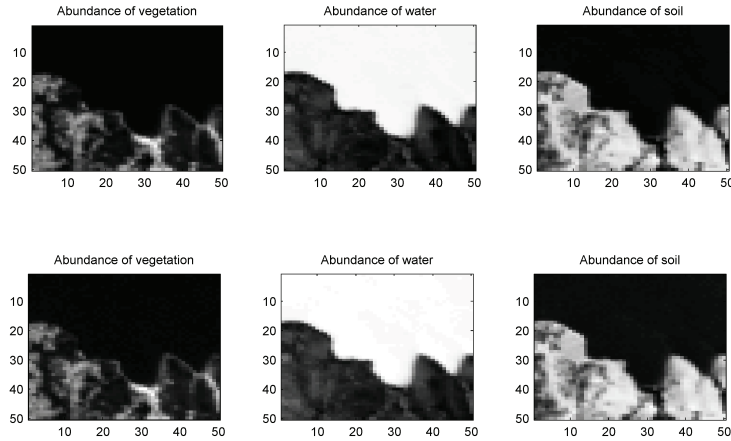


Figure 5. Haut : cartes d'abondances estimées par l'algorithme proposé. Bas : cartes d'abondances estimées par le logiciel ENVI.

tivement faible) concentration du matériau considéré. La zone de lac (représentée en blanc sur la carte d'abondance de l'eau) est clairement retrouvée. Les résultats obtenus par l'algorithme de démixage fourni avec le logiciel ENVI (RSI (Research Systems Inc.), 2003, p. 739) sont également représentés sur la Fig. 5 (en bas). Ces résultats obtenus avec un algorithme de minimisation d'un critère quadratique (sous contraintes d'additivité et de positivité) sont clairement en accord avec ceux de la figure 5 (en haut). Notons cependant que l'algorithme bayésien proposé permet également d'estimer les lois *a posteriori* des paramètres inconnus (coefficients d'abondances et variance de bruit). Ces distributions *a posteriori* peuvent être utiles pour calculer des mesures de confiance relatives aux estimations de ces paramètres.

4. Cadre semi-supervisé : les pôles de mélange sont partiellement connus

Nous généralisons dans ce paragraphe le modèle présenté précédemment au cas où le nombre de composants purs R intervenant dans le mélange est inconnu. Nous supposons à présent que les R pôles de mélange appartiennent à une bibliothèque spectrale connue $\mathcal{B} = \{s_1, \dots, s_{R_{\max}}\}$ (où s_r représente le L -spectre $[s_{r,1}, \dots, s_{r,L}]^T$ du r -ième composant de la bibliothèque). Cependant, les spectres de la bibliothèque ayant pris part au mélange sont également inconnus. Le problème de démixage spectral est alors couplé à un problème de sélection de modèle pour lequel un algorithme à sauts réversibles est proposé.

4.1. Modèle bayésien

Dans le cas où le nombre et la nature des pôles de mélange sont inconnus, le vecteur des paramètres inconnus est $\theta = \{\mathbf{c}, \mathbf{M}, R, \sigma^2\}$. La loi *a posteriori* de θ peut alors s'écrire

$$f(\mathbf{a}, \mathbf{M}, R, \sigma^2 | \mathbf{y}) \propto f(\mathbf{y} | \mathbf{a}, \mathbf{M}, \sigma^2, R) f(\mathbf{a} | R) f(\mathbf{M} | R) f(\sigma^2) f(R), \quad [20]$$

où

$$f(\mathbf{y} | \mathbf{a}, \mathbf{M}, \sigma^2, R) = \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{L}{2}} \exp \left[-\frac{\|\mathbf{y} - \mathbf{M}(R)\mathbf{a}(R)\|^2}{2\sigma^2} \right]. \quad [21]$$

Notons que les dimensions de $\mathbf{M}(R)$ et $\mathbf{a}(R)$ dépendent du paramètre inconnu R . Les lois *a priori* $f(\mathbf{a} | R)$ et $f(\sigma^2)$ ont été précédemment définies au paragraphe 3.1.2. Une loi uniforme discrète sur $[2, \dots, R_{\max}]$ est choisie comme loi *a priori* du nombre de composants de mélange R :

$$f(R) = \frac{1}{R_{\max} - 1}, \quad R = 2, \dots, R_{\max}. \quad [22]$$

Par ailleurs, conditionnellement à R , toutes les combinaisons de R spectres appartenant à la bibliothèque \mathcal{B} sont supposées équiprobables :

$$f(\mathbf{M} | R) = \frac{1}{\binom{R_{\max}}{R}} = \frac{\Gamma(R+1)\Gamma(R_{\max}-R+1)}{\Gamma(R_{\max}+1)}. \quad [23]$$

4.2. Algorithme de Gibbs hybride

Nous proposons dans ce paragraphe un algorithme de Gibbs capable de générer des échantillons distribués suivant la loi *a posteriori* $f(\mathbf{a}, \mathbf{M}, \sigma^2, R | \mathbf{y})$. Le vecteur de paramètres inconnus appartient à un espace dont la dimension dépend de R , imposant le recours à une stratégie permettant de passer d'un espace donné à un espace de dimension différente comme dans (Richardson *et al.*, 1997; Richardson *et al.*, 1998). Plus précisément, l'algorithme noté "Algorithme 2" comporte trois mouvements principaux :

- 1) mise à jour de la variance du bruit σ^2 .
- 2) mise à jour des vecteurs d'abondances \mathbf{a} ,
- 3) mise à jour des spectres des pôles de mélange \mathbf{M} ,

Ces trois mouvements possibles sont réalisés systématiquement comme dans (Richardson *et al.*, 1997). Les deux premières étapes sont identiques à l'algorithme proposé dans le cadre du démelange spectral supervisé au paragraphe précédent (voir l'algorithme 1).

Algorithme 2 Échantillonneur de Gibbs hybride pour le démélange **semi-supervisé**

% Initialisation

Échantillonner le paramètre $\tilde{R}^{(0)}$,

Choisir $\tilde{R}^{(0)}$ spectres dans la bibliothèque \mathcal{B} pour construire $\tilde{\mathbf{M}}^{(0)}$,

Échantillonner les paramètres $\tilde{\sigma}^{2(0)}$ et $\tilde{\mathbf{a}}^{(0)}$,

% Itérations

pour $t = 1, 2, \dots$, **faire**

% Mettre à jour la matrice des spectres $\tilde{\mathbf{M}}^{(t)}$

tirer $u_1 \sim \mathcal{U}_{[0,1]}$,

si $w \leq b_{\tilde{R}^{(t-1)}}$ **alors**

proposer un mouvement de *naissance*,

sinon si $w \leq b_{\tilde{R}^{(t-1)}} + d_{\tilde{R}^{(t-1)}}$ **alors**

proposer un mouvement de *mort*,

sinon

proposer un mouvement d'*échange*,

fin si

tirer $u_2 \sim \mathcal{U}_{[0,1]}$,

si $u_2 < \rho$ (voir [24] ou [25]) **alors**

poser $(\tilde{\mathbf{a}}^{(t)}, \tilde{\mathbf{M}}^{(t)}, \tilde{R}^{(t)}) = (\mathbf{a}^*, \mathbf{M}^*, R^*)$,

sinon

poser $(\tilde{\mathbf{a}}^{(t)}, \tilde{\mathbf{M}}^{(t)}, \tilde{R}^{(t)}) = (\tilde{\mathbf{a}}^{(t-1)}, \tilde{\mathbf{M}}^{(t-1)}, \tilde{R}^{(t-1)})$,

fin si

fin pour

Échantillonner $\tilde{\mathbf{a}}^{(t)}$ d'après la loi [18],

Échantillonner $\tilde{\sigma}^{2(t)}$ d'après la loi [19],

Les spectres des composants purs prenant part au mélange sont mis à jour à l'aide de trois sortes de mouvement, appelés mouvements de “*naissance*”, “*mort*” et “*échange*” comme dans (Denison *et al.*, 2002, p. 53). Les deux premiers mouvements consistent respectivement à augmenter ou à diminuer le nombre de composants purs R d'un élément. C'est pourquoi, ils nécessitent l'utilisation d'une méthode MCMC à sauts réversibles (Green, 1995). En revanche, le troisième mouvement proposé ne modifie pas le nombre de composants du mélange R . Si on suppose qu'à l'itération t , le modèle courant est défini par $(\tilde{\mathbf{a}}^{(t)}, \tilde{\mathbf{M}}^{(t)}, \tilde{\sigma}^{2(t)}, \tilde{R}^{(t)})$, les mouvements de *naissance*, *mort* et *échange* sont définis comme suit :

– *Naissance* : un mouvement de *naissance* $R^* = \tilde{R}^{(t)} + 1$ est proposé avec la probabilité $b_{\tilde{R}^{(t)}}$. Un nouveau spectre \mathbf{s}^* est aléatoirement choisi parmi les spectres des composants purs disponibles dans la bibliothèque \mathcal{B} pour construire $\mathbf{M}^* = [\tilde{\mathbf{M}}^{(t)}, \mathbf{s}^*]$. Un nouveau vecteur de coefficients d'abondances est proposé selon une règle inspirée de (Richardson *et al.*, 1997) :

- tirer un nouveau coefficient d'abondance w^* d'après la loi beta $\mathcal{B}e\left(1, \tilde{R}^{(t)}\right)$,
- re-normaliser les poids existants tels que tous les poids somment à 1,
- construire $\mathbf{a}^* = \left[a_1^*, \dots, a_{\tilde{R}^{(t)}}^*, w^*\right]^T$,

- *Mort* : un mouvement de *mort* $R^* = \tilde{R}^{(t)} - 1$ est proposé avec la probabilité $d_{\tilde{R}^{(t)}}$. Un des spectres de $\tilde{\mathbf{M}}^{(t)}$ est enlevé, ainsi que le coefficient d'abondance correspondant. Les coefficients restants sont re-normalisés pour sommer à 1,

- *Mouvement d'échange* : un mouvement d'*échange* est proposé avec la probabilité $u_{\tilde{R}^{(t)}}$. Un spectre aléatoirement choisi dans $\tilde{\mathbf{M}}^{(t)}$ est remplacé par un spectre aléatoirement choisi dans la bibliothèque \mathcal{B} .

À chaque itération, un des mouvements de *naissance*, *mort* ou d'*échange* est aléatoirement choisi avec les probabilités $b_{\tilde{R}^{(t)}}$, $d_{\tilde{R}^{(t)}}$ et $u_{\tilde{R}^{(t)}}$ respectivement (avec $b_{\tilde{R}^{(t)}} + d_{\tilde{R}^{(t)}} + u_{\tilde{R}^{(t)}} = 1$). Bien sûr, un mouvement de mort n'est pas autorisé pour $R = 2$ et le mouvement de naissance est impossible si $R = R_{\max}$ (i.e. $d_2 = b_{R_{\max}} = 0$). Par conséquent, les probabilités de mouvements suivantes ont été choisies : $b_R = d_R = u_R = \frac{1}{3}$ pour $R \in \{3, R_{\max} - 1\}$ et $b_2 = d_{R_{\max}} = u_2 = u_{R_{\max}} = \frac{1}{2}$. Quand le vecteur d'abondances \mathbf{a} est distribué *a priori* suivant une loi de Dirichlet $\mathcal{D}_R(\delta, \dots, \delta)$, les probabilités d'acceptation pour les mouvements de *naissance* et *mort* sont $\rho = \min\{1, A_b\}$ et $\rho = \min\{1, A_b^{-1}\}$ avec :

$$\begin{aligned}
A_b = \exp & \left[-\frac{\|\mathbf{y} - \mathbf{M}^* \mathbf{a}^*\|^2 - \|\mathbf{y} - \tilde{\mathbf{M}}^{(t)} \tilde{\mathbf{a}}^{(t)}\|^2}{2} \right] \\
& \times \frac{d_{\tilde{R}^{(t)}+1}}{b_{\tilde{R}^{(t)}}} \frac{1}{g_{1, \tilde{R}^{(t)}}(w^*)} (1 - w^*)^{\tilde{R}^{(t)}-1} \\
& \times \frac{\Gamma(\delta \tilde{R}^{(t)} + \delta)}{\Gamma(\delta \tilde{R}^{(t)}) \Gamma(\delta)} w^{\delta-1} (1 - w^*)^{(\delta-1)\tilde{R}^{(t)}},
\end{aligned} \tag{24}$$

où $g_{a,b}(\cdot)$ représente la densité de probabilité d'une loi Beta $\mathcal{B}e(a, b)$.

La probabilité d'acceptation pour le mouvement d'*échange* est le rapport classique de Metropolis-Hastings $\rho = \min\{1, A_s\}$ avec :

$$A_s = \exp \left[-\frac{\|\mathbf{y} - \mathbf{M}^* \mathbf{a}^*\|^2 - \|\mathbf{y} - \tilde{\mathbf{M}}^{(t)} \tilde{\mathbf{a}}^{(t)}\|^2}{2} \right]. \tag{25}$$

Le détail des calculs conduisant à ces probabilités d'acceptation est disponible dans (Dobigeon, 2007).

4.3. Résultats de simulation sur données synthétiques

Nous évaluons les performances de l'algorithme à sauts réversibles proposés en étudiant le spectre du pixel synthétique utilisé au paragraphe 3.3. Nous rappelons que ce pixel est le résultat du mélange de trois pôles de mélange (béton de construction, herbe verte et terre grasse) avec le vecteur d'abondances $\mathbf{a} = [0.3, 0.6, 0.1]^T$. Les observations sont entachées d'un bruit additif gaussien de rapport signal à bruit RSB = 15dB. Les résultats sont obtenus pour $N_{MC} = 20000$ itérations, dont $N_{bi} = 200$ itérations de chauffage. Cette simulation utilise une bibliothèque extraite du logiciel ENVI contenant six spectres, dont ceux réellement utilisés dans le mélange : béton de construction, herbe verte, terre grasse, peinture verte, brique, acier galvanisé. Les spectres de ces composants purs sont représentés sur la figure 6.

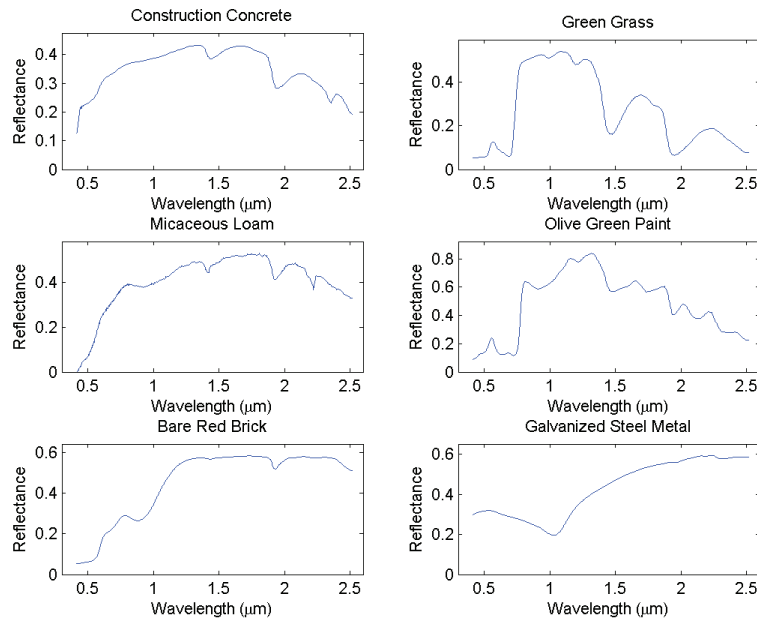


Figure 6. Spectres des pôles de mélange de la librairie.

Tout d'abord, l'ordre du modèle, i.e., le nombre de composants ayant pris part au mélange R , est estimé. La loi *a posteriori* de R représentée sur la Fig. 7 est clairement en accord avec la valeur réelle de R puisque le maximum est obtenu pour $R = 3$. Ensuite les probabilités *a posteriori* estimées des combinaisons de $\hat{R}_{MAP} = 3$ spectres permet d'identifier la configuration la plus probable. Pour cette simulation, seulement deux combinaisons ont été générées après la période de chauffage : $[s_1, s_2, s_3]$ et $[s_1, s_2, s_5]$ avec les probabilités respectives $P_{1,2,3} = 0.87$ et $P_{1,2,5} = 0.13$. La probabilité maximale correspond à la combinaison de spectres réellement employés pour

réaliser le mélange. Les lois *a posteriori* des coefficients d'abondances correspondants sont finalement estimées conditionnellement à $\hat{\mathbf{M}}_{\text{MAP}} = [\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3]$ et représentées sur la figure 8. Ces distributions sont en accord avec les valeurs réelles des abondances $\mathbf{a} = [0.3, 0.6, 0.1]^T$.

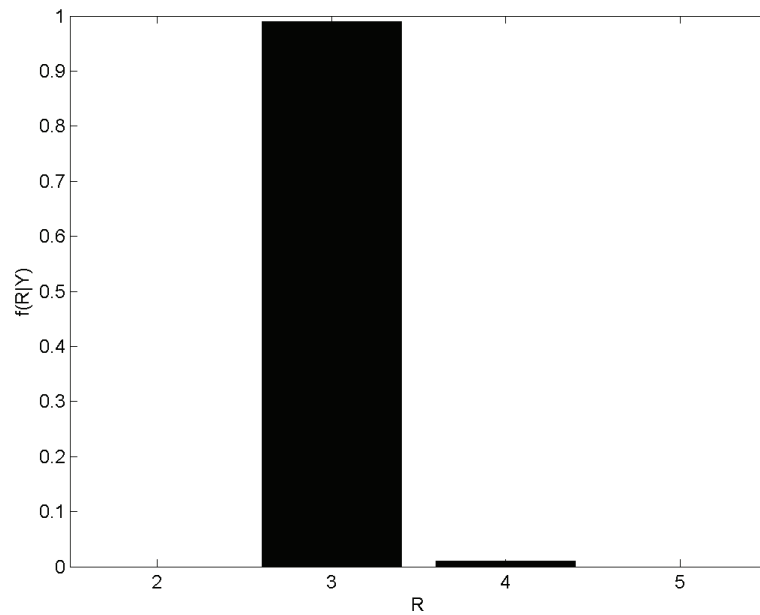


Figure 7. Loi a posteriori de l'ordre du modèle estimé R .

5. Cadre non-supervisé : les pôles de mélanges sont inconnus

Comme expliqué dans (Keshava *et al.*, 2002), le démixage spectral a très souvent été effectué en deux étapes distinctes : l'identification des signatures spectrales à l'aide d'un algorithme EPM puis l'estimation des abondances par inversion. Cependant, résoudre le problème de démixage spectral en deux étapes distinctes et successives fournit parfois de mauvais résultats. En particulier, lorsqu'il n'y a pas de pixels purs dans l'image, les algorithmes VCA, N-FINDR ou encore PPI fournissent des estimations des pôles de mélange qui sont parfois de mauvaise qualité. Pour pallier ces défauts, nous proposons dans cette partie de résoudre le problème de démixage spectral dans un cadre totalement non-supervisé, en estimant conjointement les spectres des matériaux purs et leurs abondances respectives. Cette approche conjointe fait du problème considéré un problème de séparation aveugle de sources. Le problème de séparation aveugle de sources a reçu un grand intérêt dans la littérature. Pourtant,

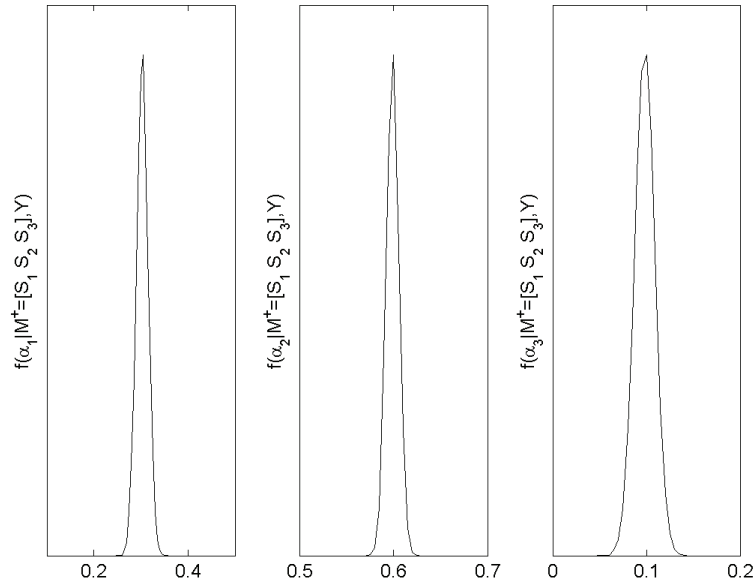


Figure 8. Loïs a posteriori des abondances estimées $\mathbf{a} = [a_1, a_2, a_3]^T$ sachant $\mathbf{M} = [S_1, S_2, S_3]$.

comme relevé par Nascimento *et al.* (2005b) ou (Dobigeon *et al.*, 2005), les algorithmes basés sur l'analyse en composantes indépendantes se sont montrés inefficaces en imagerie hyperspectrale, à cause notamment de la trop grande dépendance entre les signaux sources. D'autres stratégies exploitant des techniques de factorisation de matrices non-négatives (Paatero *et al.*, 1994) peuvent être utilisées pour l'estimation conjointe des pôles de mélange et des abondances. Cependant, celles-ci ne permettent pas de prendre en compte explicitement toutes les contraintes physiques liées au problème de démixage linéaire rencontrées en imagerie hyperspectrale : la positivité des signatures spectrales et des abondances, ainsi que l'additivité des abondances. Comme suggéré plus haut, la modélisation bayésienne adoptée ici est un moyen efficace de prendre en compte toutes ces contraintes, notamment en choisissant des lois *a priori* adéquates pour les paramètres inconnus.

Par ailleurs, une analyse simple de la géométrie du problème nous a permis de montrer qu'il suffit d'estimer les signatures spectrales dans un sous-espace approprié, de dimension bien inférieure à celle de l'espace d'observation. Cette estimation dans un sous-espace permet de réduire considérablement les degrés de liberté des paramètres inconnus, et ainsi de s'affranchir du problème d'identifiabilité tout en respectant les contraintes physiques sus-citées.

5.1. Modèle bayésien

Le problème de démixage spectral étant formulé comme un problème de séparation aveugle de sources, l'estimation conjointe des pôles de mélange et des abondances n'est possible que lorsque plusieurs pixels d'une image sont analysés conjointement. D'un démixage pixel par pixel présenté dans les paragraphes 3.1 et 4.1, nous passons donc maintenant à une analyse complète d'une image hyperspectrale, pour laquelle nous supposons connu le nombre de pôles de mélange.

En particulier, nous étendons le modèle bayésien hiérarchique proposé au paragraphe 3.1 en définissant un modèle *a priori* pour les pôles de mélange et en calculant la nouvelle loi *a posteriori* associée à l'ensemble des paramètres inconnus. La fonction de vraisemblance n'est pas rappelée ci-dessous mais s'exprime comme le produit des fonctions de vraisemblance marginales [3.1.1] attachées à chaque pixel. Par ailleurs, pour chaque sous-vecteur d'abondances \mathbf{c}_p ($p = 1, \dots, P$), une loi uniforme définie sur le simplexe \mathcal{S} défini par [11] est choisie comme loi *a priori*. Outre le respect des contraintes inhérentes [3] au modèle, le choix de cette loi *a priori* permet d'imposer une contrainte forte sur la taille du simplexe formé dans l'espace hyperspectral par les pôles de mélange estimés par l'algorithme. En effet, comme montré par (Dobigeon *et al.*, 2009), parmi deux solutions au problème de démixage aveugle *a priori* équiprobables, la loi *a priori* uniforme permet de favoriser la solution qui correspond au polytope de volume minimal. Cette propriété a également été exploitée par Craig (1994) et Bowles *et al.* (1995).

5.1.1. Modèle a priori des pôles de mélange

5.1.1.1. Réduction de dimensionalité

Remarquons tout d'abord que l'ensemble

$$\mathcal{S}_M = \left\{ \mathbf{x} \in \mathbb{R}^L; \mathbf{x} = \sum_{r=1}^R \lambda_r \mathbf{m}_r, \sum_{r=1}^R \lambda_r = 1, \lambda_r \geq 0 \right\}$$

est un polytope convexe de \mathbb{R}^L dont les sommets $\mathbf{m}_1, \dots, \mathbf{m}_R$ sont les $R \ll L$ signatures à estimer. Par conséquent, les données cachées $\mathbf{X} = \mathbf{M}\mathbf{A} = \mathbf{Y} - \mathbf{N}$ peuvent être représentées dans un sous-espace \mathcal{V}_K de \mathbb{R}^K avec $R - 1 \leq K \ll L$ sans perte d'information. Dans ce sous-espace, les données non bruitées \mathbf{X} forment un $(R - 1)$ -simplexe dont les sommets sont les projections des signatures spectrales à estimer. Comme énoncé dans (Keshava *et al.*, 2002), la réduction de dimensionalité est une étape classique dans le démixage spectral, requise par de nombreuses techniques d'EPM, comme les algorithmes N-FINDR (Winter, 1999) ou PPI (Boardman, 1993). Dans ce papier, nous proposons d'estimer dans un cadre bayésien les projections \mathbf{t}_r ($r = 1, \dots, R$) des signatures spectrales \mathbf{m}_r sur le sous-espace \mathcal{V}_K . Cette approche permet de réduire considérablement les degrés de liberté des paramètres à estimer. Nous supposons que ce sous-espace a préalablement été estimé par une technique de réduction de dimensionalité (ACP par exemple).

5.1.1.2. Réduction de dimensionalité par ACP

La matrice de covariance empirique Υ des données \mathbf{Y} est donnée par

$$\Upsilon = \frac{1}{P} \sum_{p=1}^P (\mathbf{y}_p - \bar{\mathbf{y}}) (\mathbf{y}_p - \bar{\mathbf{y}})^T \quad [26]$$

où $\bar{\mathbf{y}}$ est la moyenne empirique des observations :

$$\bar{\mathbf{y}} = \frac{1}{P} \sum_{p=1}^P \mathbf{y}_p. \quad [27]$$

Nous notons respectivement

$$\begin{cases} \mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_K), \\ \mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_K]^T \end{cases} \quad [28]$$

la matrice diagonale des K plus grandes valeurs propres et la matrice correspondantes des vecteurs propres de Υ . Le vecteur projeté $\mathbf{t}_r \in \mathbb{R}^K$ du pôle de mélange $\mathbf{m}_r \in \mathbb{R}^L$ est alors obtenu par la transformation

$$\mathbf{t}_r = \mathbf{P} (\mathbf{m}_r - \bar{\mathbf{y}}) \quad [29]$$

où $\mathbf{P} = \mathbf{D}^{-\frac{1}{2}} \mathbf{V}$. De manière équivalente, on a

$$\mathbf{m}_r = \mathbf{U} \mathbf{t}_r + \bar{\mathbf{y}} \quad [30]$$

où $\mathbf{U} = \mathbf{V}^T \mathbf{D}^{\frac{1}{2}}$. Notons que dans le sous-espace \mathcal{V}_{R-1} obtenu pour $K = R - 1$, les vecteurs $\{\mathbf{t}_r\}_{r=1, \dots, R}$ forment un simplexe que tentent d'estimer les algorithmes EPM classiques tels que N-FINDR (Winter, 1999), MVT (Craig, 1994) ou ICE (Berman *et al.*, 2004). Nous proposons d'estimer les sommets projetés \mathbf{t}_r ($r = 1, \dots, R$) de ce simplexe en utilisant un approche bayésienne. Les lois *a priori* des projections \mathbf{t}_r ($r = 1, \dots, R$) sont présentées dans le paragraphe suivant.

5.1.1.3. Lois *a priori* des spectres projetés

La signature spectrale $\mathbf{m}_r \in \mathbb{R}^L$ et sa projection $\mathbf{t}_r \in \mathbb{R}^K$ sur \mathcal{V}_K sont liées par les relations $\mathbf{t}_r = \mathbf{P} (\mathbf{m}_r - \bar{\mathbf{y}})$ et $\mathbf{m}_r = \mathbf{U} \mathbf{t}_r + \bar{\mathbf{y}}$, où \mathbf{P} est la matrice de projection, \mathbf{U} est la pseudo-inverse de \mathbf{P} et $\bar{\mathbf{y}}$ est la moyenne empirique des observations. Les lois *a priori* choisies pour les spectres projetés doivent permettre de respecter les contraintes de positivité introduites dans [5]. De simples calculs permettent d'identifier l'ensemble $\mathcal{T}_r \subset \mathcal{V}_K$ tel que

$$\{m_{l,r} \geq 0, \forall l = 1, \dots, L\} \Leftrightarrow \{\mathbf{t}_r \in \mathcal{T}_r\} \quad [31]$$

à l'aide des L inégalités suivantes

$$\mathcal{T}_r = \left\{ \mathbf{t}_r; \bar{y}_l + \sum_{k=1}^K u_{l,k} t_{k,r} \geq 0, l = 1, \dots, L \right\}, \quad [32]$$

L'originalité de la méthode de séparation de sources considérée dans cet article consiste à définir des lois *a priori* pour les projections des pôles de mélange \mathbf{m}_r sur le sous espace \mathcal{V}_K notées \mathbf{t}_r (et non pas des lois *a priori* pour les spectres des divers composants du mélange \mathbf{m}_r , ce qui pourrait paraître naturel). Plus précisément, nous choisissons une loi normale multivariée

$$\mathbf{t}_r \sim \mathcal{N}_{\mathcal{T}_r}(\mathbf{e}_r, s_r^2 \mathbf{I}_K) \quad [33]$$

tronquée à l'ensemble \mathcal{T}_r comme loi *a priori* pour chaque vecteur \mathbf{t}_r ($r = 1, \dots, R$). Les vecteurs moyennes \mathbf{e}_r de ces lois *a priori* sont fixés comme les solutions fournies par un algorithme rapide d'EPM comme N-FINDR ou VCA. En absence d'informations *a priori* supplémentaires, les variances s_r^2 ($r = 1, \dots, R$) sont fixées à de grandes valeurs $s_1^2 = \dots = s_R^2 = 50$, ce qui permet de modéliser l'écart entre les véritables pôles de mélange \mathbf{t}_r et les estimations grossières \mathbf{e}_r issues des algorithmes N-FINDR ou VCA.

5.1.2. Loi a posteriori

Les lois *a priori* des spectres projetés définies précédemment permettent d'obtenir la loi *a posteriori* jointe suivante

$$\begin{aligned} f(\mathbf{C}, \mathbf{T}, \sigma^2 | \mathbf{Y}) &\propto \prod_{r=1}^R \exp \left[-\frac{\|\mathbf{t}_r - \mathbf{e}_r\|^2}{2s_r^2} \right] \mathbf{1}_{\mathcal{T}_r}(\mathbf{t}_r) \\ &\times \prod_{p=1}^P \left[\left(\frac{1}{\sigma^2} \right)^{\frac{L}{2}+1} \exp \left(-\frac{\|\mathbf{y}_p - (\mathbf{U}\mathbf{T} + \bar{\mathbf{y}}\mathbf{1}_{R-1})\mathbf{a}_p\|^2}{2\sigma^2} \right) \right] \\ &\times \prod_{p=1}^P \mathbf{1}_{\mathcal{S}}(\mathbf{c}_p), \end{aligned} \quad [34]$$

où $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_P]^T$ est une matrice issue de la reparamétrisation [10] des vecteurs d'abondances. Un algorithme de Gibbs, détaillé au paragraphe suivant, est proposé pour générer des échantillons $\{\mathbf{C}^{(t)}, \mathbf{T}^{(t)}, \sigma^{2(t)}\}$ distribués suivant cette loi, où t est l'indice de l'itération. Ces échantillons sont ensuite utilisés pour approcher les estimateurs des paramètres d'intérêt.

5.2. Échantillonneur de Gibbs

Nous détaillons ci-dessous un échantillonneur de Gibbs qui permet de générer des échantillons distribués asymptotiquement suivant la loi *a posteriori* [34]. Cet algorithme se résume à l'algorithme 1 détaillé au paragraphe 3.2 auquel s'ajoute une étape qui permet d'échantillonner suivant la loi conditionnelle $f(\mathbf{T} | \mathbf{C}, \sigma^2, \mathbf{Y})$ (voir aussi l'algorithme 3).

Algorithme 3 Échantillonneur de Gibbs pour le démélange spectral **non-supervisé**

```
1: % Pré-traitement
2: Calculer la moyenne empirique  $\bar{\mathbf{y}}$  selon [27],
3: Calculer les matrices  $\mathbf{D}$  et  $\mathbf{V}$  selon [28] à l'aide d'une ACP,
4: Poser  $\mathbf{U} = (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \mathbf{D}^{\frac{1}{2}}$ ,
5: Choisir les vecteurs approchés  $\mathbf{e}_r \in \mathcal{V}_K$  requis dans [33],
6: % Initialisation
7: pour  $r = 1, \dots, R$  faire
8:   Échantillonner  $\mathbf{t}_r^{(0)}$  selon la loi [33],
9:   Poser  $\mathbf{m}_r^{(0)} = \mathbf{U} \mathbf{t}_r^{(0)} + \bar{\mathbf{y}}$ ,
10: fin pour
11: Échantillonner  $\sigma^{2(0)}$  selon la loi [12],
12: % Itérations
13: pour  $t = 1, 2, \dots$ , faire
14:   pour  $p = 1, \dots, P$  faire
15:     Échantillonner  $\mathbf{c}_p^{(t)}$  selon la loi [36],
16:   fin pour
17:   pour  $r = 1, \dots, R$  faire
18:     pour  $k = 1, \dots, K$  faire
19:       Échantillonner  $t_{k,r}^{(t)}$  selon la loi [40],
20:     fin pour
21:     Poser  $\mathbf{m}_r^{(t)} = \mathbf{U} \mathbf{t}_r^{(0)} + \bar{\mathbf{y}}$ ,
22:   fin pour
23:   Échantillonner  $\sigma^{2(t)}$  selon la loi [42].
24: fin pour
```

5.2.1. Échantillonnage suivant $f(\mathbf{C}|\mathbf{T}, \sigma^2, \mathbf{Y})$

Pour chaque pixel p , comme au paragraphe 3.2.1, on a

$$f(\mathbf{c}_p | \mathbf{T}, \sigma^2, \mathbf{y}_p) \exp \left[-\frac{(\mathbf{c}_p - \mathbf{v}_p)^T \boldsymbol{\Sigma}_p^{-1} (\mathbf{c}_p - \mathbf{v}_p)}{2} \right] \mathbf{1}_{\mathcal{S}}(\mathbf{c}_p), \quad [35]$$

où $\boldsymbol{\Sigma}_p$ et \mathbf{v}_p ont été définis pour un pixel donné au paragraphe 3.2.1. Le vecteur $\mathbf{c}_p | \mathbf{T}, \sigma^2, \mathbf{y}_p$ est donc distribué suivant une loi normale multivariée tronquée au simplexe \mathcal{S} défini par [11]

$$\mathbf{c}_p | \mathbf{T}, \sigma^2, \mathbf{y}_p \sim \mathcal{N}_{\mathcal{S}}(\mathbf{v}_p, \boldsymbol{\Sigma}_p). \quad [36]$$

5.2.2. Échantillonnage suivant $f(\mathbf{T}|\mathbf{C}, \sigma^2, \mathbf{Y})$

Si \mathbf{T}_{-r} est la matrice \mathbf{T} dont la r ème colonne a été supprimée, alors la loi conditionnelle *a posteriori* de \mathbf{t}_r ($r = 1, \dots, R$) est

$$f(\mathbf{t}_r | \mathbf{T}_{-r}, \mathbf{c}_r, \sigma^2, \mathbf{Y}) \propto \exp \left[-\frac{1}{2} (\mathbf{t}_r - \boldsymbol{\tau}_r)^T \boldsymbol{\Lambda}_r^{-1} (\mathbf{t}_r - \boldsymbol{\tau}_r) \right] \mathbf{1}_{\mathcal{T}_r}(\mathbf{t}_r), \quad [37]$$

avec

$$\begin{cases} \boldsymbol{\Lambda}_r = \left[\sum_{p=1}^P a_{p,r}^2 \mathbf{U}^T \boldsymbol{\Sigma}_n^{-1} \mathbf{U} + \frac{1}{s_r^2} \mathbf{I}_K \right]^{-1}, \\ \boldsymbol{\tau}_r = \boldsymbol{\Lambda}_r \left[\sum_{p=1}^P a_{p,r} \mathbf{U}^T \boldsymbol{\Sigma}_n^{-1} \boldsymbol{\epsilon}_{p,r} + \frac{1}{s_r^2} \mathbf{e}_r \right], \end{cases} \quad [38]$$

et

$$\boldsymbol{\epsilon}_{p,r} = \mathbf{y}_p - a_{p,r} \bar{\mathbf{y}} - \sum_{j \neq r} a_{p,j} \mathbf{m}_j. \quad [39]$$

Générer des vecteurs distribués suivant cette loi est délicat, principalement à cause la troncature à l'espace \mathcal{T}_r . L'alternative consiste à générer chaque composante $t_{k,r}$ de \mathbf{t}_r conditionnellement aux autres $\mathbf{t}_{-k,r} = \{t_{j,r}\}_{j \neq k}$. En notant $\mathcal{U}_k^+ = \{l; u_{l,k} > 0\}$, $\mathcal{U}_k^- = \{l; u_{l,k} < 0\}$ et $\varepsilon_{l,k,r} = \bar{y}_l + \sum_{j \neq k} u_{l,j} t_{j,r}$, on obtient

$$t_{k,r} | \mathbf{t}_{-k,r}, \mathbf{T}_{-r}, \mathbf{c}_r, \sigma^2, \mathbf{Y} \sim \mathcal{N}_{[t_{k,r}^-, t_{k,r}^+]}(w_{k,r}, z_{k,r}^2), \quad [40]$$

avec

$$\begin{cases} t_{k,r}^- = \max_{l \in \mathcal{U}_k^+} -\frac{\varepsilon_{l,k,r}}{u_{l,k}}, \\ t_{k,r}^+ = \min_{l \in \mathcal{U}_k^-} -\frac{\varepsilon_{l,k,r}}{u_{l,k}}, \end{cases} \quad [41]$$

où $w_{k,r}$ et $z_{k,r}^2$ sont les moyenne et variance conditionnelles calculées suivant (Kay, 1993, p. 324) (voir aussi les calculs similaires de Dobigeon *et al.* (2007a)). Générer des échantillons distribués suivant la loi normale doublement tronquée [40] peut se faire en utilisant la méthode décrite dans (Robert, 1995).

5.2.3. Échantillonnage suivant $f(\sigma^2 | \mathbf{C}, \mathbf{T}, \mathbf{Y})$

La loi conditionnelle *a posteriori* de $\sigma^2 | \mathbf{C}, \mathbf{T}, \mathbf{Y}$ est la loi inverse-gamma

$$\sigma^2 | \mathbf{C}, \mathbf{T}, \mathbf{Y} \sim \mathcal{IG} \left(\frac{PL}{2}, \frac{1}{2} \sum_{p=1}^P \|\mathbf{y}_p - \mathbf{M}\mathbf{a}_p\|^2 \right). \quad [42]$$

5.3. Résultats de simulation sur données synthétiques

Pour illustrer l'intérêt de la méthode proposée, l'algorithme a été appliqué sur une image synthétique de 100×100 pixels, dans laquelle apparaissent trois signatures spectrales représentatives d'un environnement semi-urbain : béton de construction, herbe verte et brique rouge. Ces signatures, issues de la bibliothèque spectrale fournie avec le logiciel ENVI, sont mesurées dans $L = 413$ bandes spectrales et représentées sur la figure 9 (haut, noir). Ces composants macroscopiques sont mélangés au sein des pixels de l'image dans des proportions aléatoires avec un bruit i.i.d. correspondant à un niveau de bruit RSB = 15dB.

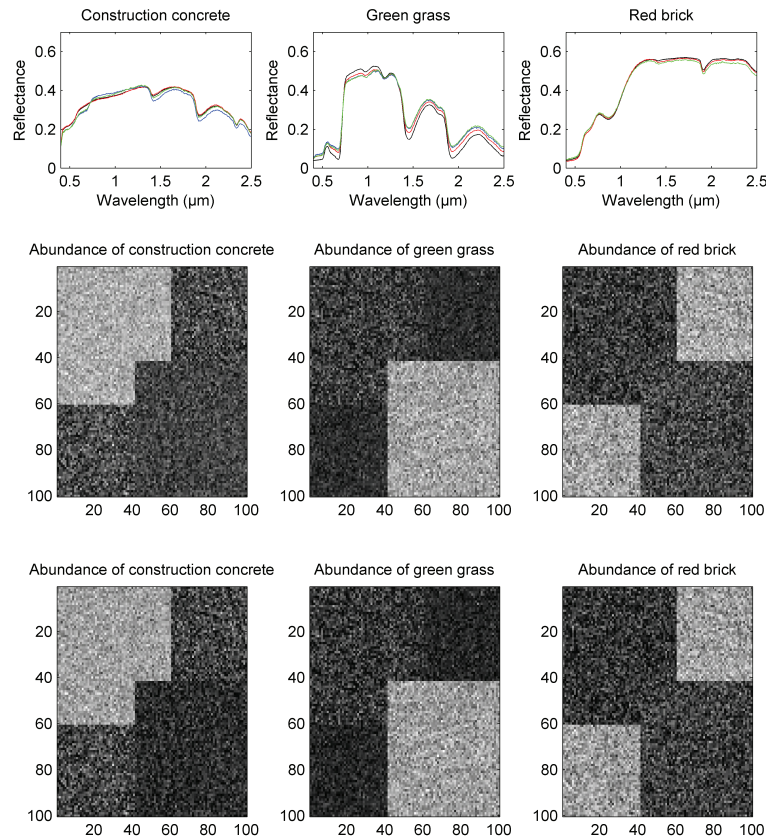


Figure 9. Haut : spectres réels (noir), estimés par N-FINDR (bleu), estimés par VCA (vert) et estimés par l'approche proposée (rouge). Milieu et bas : cartes d'abondances réelles et estimées.

Les résultats d'estimation des signatures spectrales obtenus par notre algorithme, représentées sur la figure 9 (haut, rouge) ont été comparés à ceux fournis par les algorithmes VCA et N-FINDR (voir Tableau 1, haut). Les résultats concernant l'estimation des 10^4 vecteurs d'abondances (voir figure 9, bas), rapportés dans le Tableau 1 (bas) illustrent également la précision de la méthode proposée. Comme on peut l'observer, les résultats d'estimation obtenus avec l'algorithme bayésien sont meilleurs que ceux obtenus avec les autres algorithmes. Ceci est dû au fait que l'algorithme i) fournit une estimation conjointe de \mathbf{M} et \mathbf{A} , ii) respecte les contraintes de positivité et d'additivité liées au modèle de mélange.

Spectres	Modèle bayésien	VCA	N-FINDR
Matériau #1	0.10	1.29	0.54
Matériau #2	2.68	15.59	5.19
Matériau #3	0.16	4.35	0.57
Abondances	Modèle bayésien	VCA	N-FINDR
Matériau #1	25.68	57.43	30.66
Matériau #2	29.97	74.48	46.45
Matériau #3	3.19	83.02	11.22

Tableau 1. Comparaison de performances d'estimation entre les algorithmes VCA, N-FINDR et l'approche proposée : erreurs quadratiques moyennes entre les $R = 3$ spectres estimés et les spectres réels (haut), erreurs quadratiques moyennes globales entre les abondances réelles et estimées (bas).

5.4. Résultats de simulation sur données réelles

Pour terminer, nous proposons d'appliquer l'algorithme de démixage spectral aveugle proposé ci-dessus à l'image réelle de Moffett Field présentée au paragraphe 3.4. Les $R = 3$ pôles de mélange identifiés par l'algorithme sont représentés sur la Fig. 10 (haut). Les cartes d'abondances correspondantes sont également représentées sur la Fig. 10 (bas). Ces résultats sont en accord avec ceux des figures 4 et 5 obtenus dans le cas d'une approche supervisée.

6. Conclusion et discussion

Dans cet article, nous avons présenté des algorithmes totalement bayésiens pour résoudre le problème de démixage linéaire d'images hyperspectrales dans des contextes supervisé, semi-supervisé et non-supervisé. Dans chacun de ces cadres d'étude, des lois *a priori* adéquates sont choisies pour les paramètres inconnus. Notamment un soin particulier a été apporté pour le choix de ces lois afin de satisfaire les contraintes inhérentes au modèle de mélange : positivité et additivité des coefficients d'abondances et positivité des spectres. Des algorithmes MCMC ont été développés

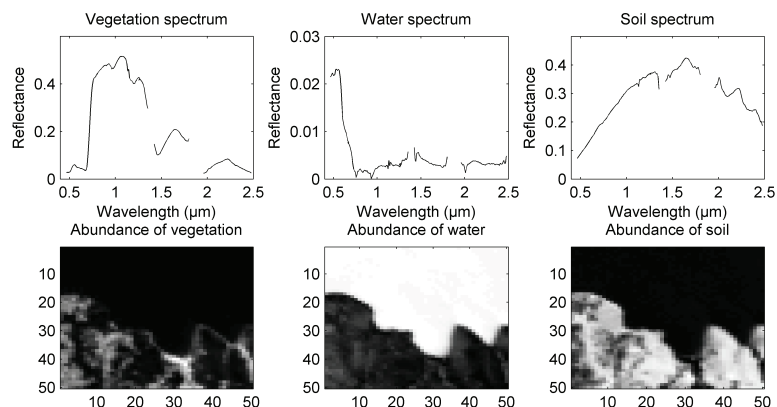


Figure 10. Haut : les $R = 3$ pôles de mélange estimés par l’algorithme non-supervisé dans la scène Moffett Field. Bas : les cartes d’abondances correspondantes.

pour générer des échantillons distribués suivant les lois *a posteriori* des paramètres inconnus. Des résultats de simulations, obtenus sur des images synthétiques et réelles, ont permis de mettre en évidence l’intérêt des méthodes proposées. Une piste de travail futur concerne l’estimation du nombre de spectres purs impliqués dans le mélange.

Par souci de concision, nous avons présenté dans cet article des résultats obtenus avec un seul niveau de bruit $RSB \approx 15\text{dB}$. Les comportements des algorithmes proposés face à des niveaux de bruits plus élevés ont été étudiés dans (Dobigeon *et al.*, 2008a) et (Dobigeon *et al.*, 2009). Ces résultats montrent naturellement que les variances des lois *a posteriori* des paramètres d’intérêts augmentent lorsque le niveau de bruit augmente. La qualité des estimations fournies par les algorithmes reste toutefois supérieure à celle de la plupart des algorithmes de démixage rencontrés dans la littérature. Notons par ailleurs que les capteurs hyperspectraux les plus couramment utilisés fournissent des images dont les niveaux de RSB sont rarement inférieurs à 15dB (Pearlman *et al.*, 2003).

Enfin, nous noterons que les algorithmes de démixage supervisé, semi-supervisé et non supervisé ont été développés et testés dans le cas de scénarii où le nombre de composants purs est limité. En particulier, dans un contexte semi-supervisé, lorsque ce nombre de composants devient important, il est clair que la stratégie à sauts réversibles peut se montrer coûteuse en temps de calculs. L’algorithme est en effet contraint d’explorer des espaces dont la dimension peut devenir prohibitive. Considérer le cas de démixages supervisé ou semi-supervisé avec une grande bibliothèque spectrale nécessite alors de mettre en oeuvre des techniques appropriées. Les avancées théoriques et algorithmiques récentes dans le domaine des représentations parcimonieuses ouvrent la voie pour des travaux de recherches originaux. Sous l’hypothèse plausible

qu'au sein d'un pixel, un petit nombre de composants purs participe au mélange, l'estimation des coefficients d'abondances peut être considérée comme la résolution d'un problème inverse sous contraintes de parcimonie. Des résultats préliminaires ont récemment été présentés dans (Iordache *et al.*, 2009; Iordache *et al.*, 2010). Il apparaît l'utilisation d'algorithmes gloutons tels que "matching pursuit" (MP) ou sa version orthogonale (OMP) représente une alternative intéressante dans le cas où les algorithmes bayésiens proposés dans cet article seraient mis en défaut.

7. Remerciements

Une partie de ce travail a été effectué en collaboration avec C.-I Chang de l'Université du Maryland. Quelques résultats ont également été obtenus lors d'un projet "Jeunes Chercheurs" financé par le GdR-ISIS. Les auteurs tiennent à remercier Jérôme Idier et Eric le Carpentier pour les discussions intéressantes concernant ce travail.

8. Bibliographie

- Akgun T., Altunbasak Y., Mersereau R. M., « Super-Resolution Reconstruction of Hyperspectral Images », *IEEE Trans. Image Processing*, vol. 14, n° 11, p. 1860-1875, Nov., 2005.
- AVIRIS Free Data, Jet Propulsion Lab. (JPL), California Inst. Technol., Pasadena, CA, 2006. <http://aviris.jpl.nasa.gov/html/aviris.freedata.html>.
- Berman M., Kiiveri H., Lagerstrom R., Ernst A., Dunne R., Huntington J. F., « ICE : A Statistical Approach to Identifying Endmembers in Hyperspectral Images », *IEEE Trans. Geosci. and Remote Sensing*, vol. 42, n° 10, p. 2085-2095, Oct., 2004.
- Blumensath T., Davies M. E., « Monte-Carlo methods for adaptive sparse approximations of time-series », *IEEE Trans. Signal Processing*, vol. 55, n° 9, p. 4474-4486, Sept., 2007.
- Boardman J., « Automating spectral unmixing of AVIRIS data using convex geometry concepts », *Summaries 4th Annu. JPL Airborne Geoscience Workshop*, vol. 1, JPL Pub., Washington, D.C., p. 11-14, 1993.
- Bowles J. H., Palmadesso P. J., Antoniadis J. A., Baumbach M. M., Rickard L. J., « Use of filter vectors and fast convex set methods in hyperspectral analysis », in M. Strojnik, B. F. Andresen (eds), *Infrared Spaceborne Remote Sensing III*, vol. 2553, SPIE, p. 148-157, Sept., 1995.
- Chang C.-I., *Hyperspectral Imaging : Techniques for Spectral detection and classification*, Kluwer, New York, 2003.
- Chang C.-I., Ji B., « Weighted abundance-constrained linear spectral mixture analysis », *IEEE Trans. Geosci. and Remote Sensing*, vol. 44, n° 2, p. 378-388, Feb., 2001.
- Chen F. W., « Archiving and Distribution of 2-D Geophysical Data Using Image Formats With Lossless Compression », *IEEE Geosci. and Remote Sensing Lett.*, vol. 2, n° 1, p. 64-68, Jan., 2005.
- Chen M.-H., Deely J. J., « Bayesian Analysis for a Constrained Linear Multiple Regression Problem for Predicting the New Crop of Apples », *J. of Agricultural, Biological and Environmental Stat.*, vol. 1, p. 467-489, 1996.

- Christophe E., Léger D., Mailhes C., « Quality Criteria Benchmark for Hyperspectral Imagery », *IEEE Trans. Geosci. and Remote Sensing*, vol. 43, n° 9, p. 2103-2114, Sept., 2005.
- Craig M. D., « Minimum Volume Transforms for Remotely Sensed Data », *IEEE Trans. Geosci. and Remote Sensing*, vol. 14, n° 3, p. 542-552, May, 1994.
- Denison D. G. T., Holmes C. C., Mallick B. K., Smith A. F. M., *Bayesian methods for nonlinear classification and regression*, Wiley, Chichester, England, 2002.
- Dobigeon N., Modèles bayésiens hiérarchiques pour le traitement multi-capteurs, PhD thesis, Institut National Polytechnique (INP) de Toulouse, Toulouse, France, 2007.
- Dobigeon N., Achard V., « Performance comparison of geometric and statistical methods for endmembers extraction in hyperspectral imagery », in L. Bruzzone (ed.), *Image and Signal Processing for Remote Sensing XI*, vol. 5982, SPIE, p. 335-344, Oct., 2005.
- Dobigeon N., Moussaoui S., Coulon M., Tournet J.-Y., Hero A. O., « Joint Bayesian end-member extraction and linear unmixing for hyperspectral imagery », *IEEE Trans. Signal Processing*, vol. 57, n° 11, p. 4355-4368, Nov., 2009.
- Dobigeon N., Tournet J.-Y., Efficient sampling according to a multivariate Gaussian distribution truncated on a simplex, Technical report, IRIT/ENSEEIH/TéSA, March, 2007a.
- Dobigeon N., Tournet J.-Y., Chang C.-I., « Semi-Supervised Linear Spectral Unmixing Using a Hierarchical Bayesian Model for Hyperspectral Imagery », *IEEE Trans. Signal Processing*, vol. 56, n° 7, p. 2684-2695, July, 2008a.
- Dobigeon N., Tournet J.-Y., Davy M., « Joint segmentation of piecewise constant autoregressive processes by using a hierarchical model and a Bayesian sampling approach », *IEEE Trans. Signal Processing*, vol. 55, n° 4, p. 1251-1263, April, 2007b.
- Dobigeon N., Tournet J.-Y., Hero A. O., « Bayesian linear unmixing of hyperspectral images corrupted by colored Gaussian noise with unknown covariance matrix », *IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Las Vegas, USA, p. 3433-3436, March, 2008b.
- Févotte C., Godsill S. J., « A Bayesian Approach for Blind Separation of Sparse Sources », *IEEE Trans. Audio, Speech, Language Processing*, vol. 14, n° 6, p. 2174-2188, Nov., 2006.
- Green P. J., « Reversible jump MCMC computation and Bayesian model determination », *Biometrika*, vol. 82, n° 4, p. 711-732, Dec., 1995.
- Hapke B. W., « Bidirectional reflectance spectroscopy. I. Theory », *J. Geophys. Res.*, vol. 86, p. 3039-3054, 1981.
- Heinz D. C., Chang C.-I., « Fully constrained least-squares linear spectral mixture analysis method for material quantification in hyperspectral imagery », *IEEE Trans. Geosci. and Remote Sensing*, vol. 29, n° 3, p. 529-545, March, 2001.
- Iordache M.-D., Bioucas-Dias J., Plaza A., « Unmixing of Sparse Hyperspectral Mixtures », *Proc. IEEE Int. Conf. Geosci. and Remote Sensing (IGARSS)*, Cape Town, South Africa, March, 2009.
- Iordache M.-D., Bioucas-Dias J., Plaza A., « On the Use of Spectral Libraries to Perform Sparse Unmixing of Hyperspectral Data », *Proc. IEEE GRSS Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing (WHISPERS)*, Reykjavík, Iceland, June, 2010.

- Jackson Q., Landgrebe D. A., « An Adaptive Method for Combined Covariance Estimation and Classification », *IEEE Trans. Geosci. and Remote Sensing*, vol. 40, n° 5, p. 1082-1087, May, 2002.
- Jeffreys H., *Theory of Probability*, 3 edn, Oxford University Press, London, 1961.
- Johnson P. E., Smith M. O., Taylor-George S., Adams J. B., « A semiempirical method for analysis of the reflectance spectra of binary mineral mixtures », *J. Geophys. Res.*, vol. 88, p. 3557-3561, 1983.
- Kay S. M., *Fundamentals of Statistical Signal Processing : Estimation theory*, Prentice Hall, Englewood Cliffs NJ, 1993.
- Keshava N., Mustard J. F., « Spectral Unmixing », *IEEE Signal Processing Magazine*, vol. , p. 44-57, Jan., 2002.
- Manolakis D., Siracusa C., Shaw G., « Hyperspectral Subpixel Target Detection Using the Linear Mixing Model », *IEEE Trans. Geosci. and Remote Sensing*, vol. 39, n° 7, p. 1392-1409, July, 2001.
- Moussaoui S., Brie D., Mohammad-Djafari A., Carteret C., « Separation of Non-Negative Mixture of Non-Negative Sources Using a Bayesian Approach and MCMC Sampling », *IEEE Trans. Signal Processing*, vol. 54, n° 11, p. 4133-4145, Nov., 2006.
- Nascimento J. M., Bioucas-Dias J. M., « Vertex Component Analysis : A Fast Algorithm to Unmix Hyperspectral Data », *IEEE Trans. Geosci. and Remote Sensing*, vol. 43, n° 4, p. 898-910, April, 2005a.
- Nascimento J. M. P., Bioucas-Dias J. M., « Does independent component analysis play a role in unmixing hyperspectral data ? », *IEEE Trans. Geosci. and Remote Sensing*, vol. 43, n° 1, p. 175-187, Jan., 2005b.
- Paatero P., Tapper U., « Positive Matrix Factorization : a non-negative factor model with optimal utilization of error estimates of data values », *Environmetrics*, vol. 5, p. 111-126, 1994.
- Pearlman J. S., Barry P. S., Segal C. C., Shepanski J., Beiso D., Carman S. L., « Hyperion, a Space-Based Imaging Spectrometer », *IEEE Trans. Geosci. and Remote Sensing*, vol. 41, n° 6, p. 1160-1173, June, 2003.
- Plaza J., Pérez R., Plaza A., Martínez P., Valencia D., « Mapping oil spills on sea water using spectral mixture analysis of hyperspectral image data », in J. O. Jensen, J.-M. Thériault (eds), *Chemical and Biological Standoff Detection III*, vol. 5995, SPIE, p. 79-86, 2005.
- Punskaya E., Andrieu C., Doucet A., Fitzgerald W., « Bayesian curve fitting using MCMC with applications to signal segmentation », *IEEE Trans. Signal Processing*, vol. 50, n° 3, p. 747-758, March, 2002.
- Rellier G., Descombes X., Falzon F., Zerubia J., « Texture Feature Analysis Using a Gauss Markov Model in Hyperspectral Image Classification », *IEEE Trans. Geosci. and Remote Sensing*, vol. 42, n° 7, p. 1543-1551, July, 2004.
- Richardson S., Green P. J., « On Bayesian analysis of mixtures with unknown number of components », *J. Roy. Stat. Soc. B*, vol. 59, n° 4, p. 731-792, 1997.
- Richardson S., Green P. J., « Corrigendum : On Bayesian analysis of mixtures with unknown number of components », *J. Roy. Stat. Soc. B*, vol. 60, n° 3, p. 661, 1998.
- Robert C. P., « Simulation of truncated normal variables », *Statistics and Computing*, vol. 5, p. 121-125, 1995.

- Robert C. P., *The Bayesian Choice : from Decision-Theoretic Motivations to Computational Implementation*, Springer Texts in Statistics, 2 edn, Springer-Verlag, New York, 2007.
- Robert C. P., Casella G., *Monte Carlo Statistical Methods*, Springer-Verlag, New York, 1999.
- RSI (Research Systems Inc.), *ENVI User's guide Version 4.0*, Boulder, CO 80301 USA. Sept., 2003.
- Tang X., Pearlman W. A., « Lossy-to-lossless block-based compression of hyperspectral volumetric data », *Proc. IEEE Int. Conf. Image Processing (ICIP)*, vol. 5, p. 3283-3286, Oct., 2004.
- Theys C., Dobigeon N., Tourneret J.-Y., Lantéri H., « Linear unmixing of hyperspectral images using a scaled gradient method », *Proc. IEEE-SP Workshop Stat. and Signal Processing*, Cardiff, UK, p. 729-732, Aug., 2009.
- Tu T. M., Chen C. H., Chang C.-I., « A Noise Subspace Projection Approach to Target Signature Detection and Extraction in an Unknown Background for Hyperspectral Images », *IEEE Trans. Geosci. and Remote Sensing*, vol. 36, n° 1, p. 171-181, Jan., 1998.
- Winter M., « Fast Autonomous Spectral End-member Determination in Hyperspectral Data », *Proc. 13th Int. Conf. on Applied Geologic Remote Sensing*, vol. 2, Vancouver, p. 337-344, April, 1999.