Journal of Mathematical Imaging and Vision 15: 17–37, 2001

Spectral Mixture Analysis: Linear and Semi-parametric Full and Iterated Partial Unmixing in Multi- and Hyperspectral Image Data

ALLAN AASBJERG NIELSEN
IMM, Informatics and Mathematical Modelling, Technical University of Denmark, Building 321, Richard Petersens Plads, DK-2800 Kgs. Lyngby, Denmark
aa@imm.dtu.dk, Internet www.imm.dtu.dk/∼aa

Abstract. As a supplement or an alternative to classification of hyperspectral image data linear and semi-parametric mixture models are considered in order to obtain estimates of abundance of each class or end-member in pixels with mixed membership. Full unmixing based on both ordinary least squares (OLS) and non-negative least squares (NNLS), and the partial unmixing methods orthogonal subspace projection (OSP), constrained energy minimization (CEM) and an eigenvalue formulation alternative are dealt with. The solution to the eigenvalue formulation alternative proves to be identical to the CEM solution. The matrix inversion involved in CEM can be avoided by working on (a subset of) orthogonally transformed data such as signal maximum autocorrelation factors, MAFs, or signal minimum noise fractions, MNFs. This will also cause the partial unmixing result to be independent of the noise isolated in the MAF/MNFs not included in the analysis. CEM and the eigenvalue formulation alternative enable us to perform partial unmixing when we know one desired end-member spectrum only and not the full set of end-member spectra. This is an advantage over full unmixing and OSP. The eigenvalue formulation of CEM inspires us to suggest an iterated CEM scheme. Also the target constrained interference minimized filter (TCIMF) is described. Spectral angle mapping (SAM) is briefly described. Finally, semi-parametric unmixing (SPU) based on a combined linear and additive model with a non-linear, smooth function to represent end-member spectra unaccounted for is introduced. An example with two generated bands shows that both full unmixing, the CEM, the iterated CEM and TCIMF methods perform well. A case study with a 30 bands subset of AVIRIS data shows the utility of full unmixing, SAM, CEM and iterated CEM to more realistic data. Iterated CEM seems to suppress noise better than CEM. A study with AVIRIS spectra generated from real spectra shows (1) that ordinary least squares in this case with one unknown spectrum performs better than non-negative least squares, and (2) that although not fully satisfactory the semi-parametric model gives better estimates of end-member abundances than the linear model.

Keywords: least squares regression, spectral angle mapping (SAM), orthogonal subspace projection (OSP), matched filtering, iterated constrained energy minimization (CEM), generalized eigenvalue problem, target constrained interference minimized filter (TCIMF), non-linear semi-parametric unmixing (SPU)

1. Introduction

In ordinary discriminant analysis which is often used to classify for instance multi- or hyperspectral remote sensing image data it is assumed that each observation (or pixel) is a member of one and only one of a number of pre-determined classes. Spectral mixture models allow us to estimate the abundance of each class in pixels with mixed class membership [20, 29, 33, 34, 36, 48].

This article gives a brief overview of several methods for full and partial unmixing described in the literature. Also, several new ideas are presented, namely (1) the inclusion of a constant term (α_0 below) and interactions between spectra in the linear mixture model, (2) an eigenvalue formulation of constrained energy

minimization, (3) partial unmixing by constrained energy minimization in MAF/MNF space (MAF: maximum autocorrelation factor, MNF: minimum noise fraction) to avoid matrix inversion and to exclude noise isolated in high order MAF/MNFs, (4) an iterated constrained energy minimization scheme, (5) simultaneous partial unmixing of several spectra by applying (a) non-negative least squares with special constraints or (b) a new semi-parametric unmixing (SPU) model which includes a non-linear, smooth function to represent spectra unaccounted for in the selected end-members.

Section 2 describes the linear mixture model and regression analysis. Section 3 very briefly describes full unmixing. Section 4 describes partial unmixing in some detail. Several methods are dealt with, namely orthogonal subspace projection (OSP), constrained energy minimization (CEM), an eigenvalue formulation alternative to CEM, iterated CEM, and the target constrained interference minimized filter (TCIMF). Section 5 briefly describes spectral angle mapping (SAM). Section 6 briefly introduces the reader to additive models and describes the semi-parametric unmixing (SPU) model. Section 7 briefly describes some useful computer programs to carry out this type of analysis. Section 8 gives two cases, one is based on simple, generated data with two spectral bands, the other case uses AVIRIS data. Section 9 concludes.

2. Linear Mixing

We assume that the signal measured at each pixel consists of a linear combination of p so-called end-members. End-members are pure pre-determined classes with 100% abundance of one element and with no mixtures. We think of our l-dimensional signal for end-member i as a vector $\boldsymbol{m}_i = [m_{i1} \ \ldots \ m_{il}]^T$, $i = 1, \ldots, p$ and represent the end-members by a matrix

$$\boldsymbol{M} = [\boldsymbol{m}_1 \quad \ldots \quad \boldsymbol{m}_p] = \begin{bmatrix} m_{11} & \cdots & m_{p1} \\ \vdots & \ddots & \vdots \\ m_{1l} & \cdots & m_{pl} \end{bmatrix} \tag{1}$$

with one column for each end-member. l varies over wavelength λ and p varies over end-members. We write each observation $\boldsymbol{r}(x, y) = [r_1(x, y) \ \ldots \ r_l(x, y)]^T$ as a linear combination of the end-members $\boldsymbol{M}$; the abundances $\boldsymbol{\alpha}(x, y) = [\alpha_1(x, y) \ \ldots \ \alpha_p(x, y)]^T$ are the coefficients we wish to estimate

$$\boldsymbol{r}(x, y) = \boldsymbol{M}\boldsymbol{\alpha}(x, y) + \boldsymbol{n}(x, y) \tag{2}$$

where $\boldsymbol{n}(x, y) = [n_1(x, y) \ \ldots \ n_l(x, y)]^T$ is the residual or the noise, i.e., the variation in $\boldsymbol{r}(x, y)$ not explained by the model. The noise is independent, identically distributed (iid) Gaussian with expectation value $\mathrm{E}\{\boldsymbol{n}\} = \boldsymbol{0}$. This is the linear mixture model. The term linear means linear in the coefficients. Interactions between the end-member spectra, if for instance the influence of one spectrum depends on the level of another spectrum, can be allowed for by including products between the spectra. In linear models a constant term α_0 is often introduced (from now on we omit (x, y) from the notation). Here, α_0 represents effects not explained by the chosen end-members as far as these can be represented by a constant. If we introduce α_0 we get

$$\boldsymbol{M} = \begin{bmatrix} 1 & m_{11} & \cdots & m_{p1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & m_{1l} & \cdots & m_{pl} \end{bmatrix} \tag{3}$$

$$\boldsymbol{\alpha} = [\alpha_0 \ \alpha_1 \ \ldots \ \alpha_p]^T. \tag{4}$$

Sometimes the column of ones is replaced by a column of zeros. This represents the end-member "total shade."

To solve the system of equations involved we minimize the sum of squared residuals $\boldsymbol{n}^T\boldsymbol{n}$ or more generally $\boldsymbol{n}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{n}$ where $\boldsymbol{\Sigma}_n$ is the dispersion or covariance matrix of the residuals. This is done by setting the partial derivative $\partial(\boldsymbol{n}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{n})/\partial\boldsymbol{\alpha} = \boldsymbol{0}$. The result is

$$\hat{\boldsymbol{\alpha}} = \left(\boldsymbol{M}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{M}\right)^{-1}\boldsymbol{M}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{r}. \tag{5}$$

The estimator $\hat{\boldsymbol{\alpha}}$ is central ($\mathrm{E}\{\hat{\boldsymbol{\alpha}}\} = \boldsymbol{\alpha}$) with dispersion $(\boldsymbol{M}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{M})^{-1}$. When $\boldsymbol{\Sigma}_n = \sigma^2\boldsymbol{I}$ where $\boldsymbol{I}$ is the $l \times l$ unit matrix and σ^2 is the variance of all residuals, $\mathrm{V}\{n_i\} = \sigma^2$ (this is the ordinary least squares (OLS) case)

$$\hat{\boldsymbol{\alpha}} = (\boldsymbol{M}^T\boldsymbol{M})^{-1}\boldsymbol{M}^T\boldsymbol{r} \tag{6}$$

with dispersion $\sigma^2(\boldsymbol{M}^T\boldsymbol{M})^{-1}$.

To evaluate the goodness of the model we use $R^2 = (\tilde{\boldsymbol{r}}^T\tilde{\boldsymbol{r}} - \boldsymbol{n}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{n})/\tilde{\boldsymbol{r}}^T\tilde{\boldsymbol{r}}$, the coefficient of determination (if $\boldsymbol{M}$ contains the column of ones $\tilde{\boldsymbol{r}}$ is $\boldsymbol{r}$ centered, if not $\tilde{\boldsymbol{r}} = \boldsymbol{r}$), and the estimate of residual variance, $s^2 = (\boldsymbol{n}^T\boldsymbol{\Sigma}_n^{-1}\boldsymbol{n})/(l - p - 1)$. s is the root mean square error, RMSE. $l - p - 1$, the number of degrees of freedom, must be positive. If an extra column is added to $\boldsymbol{M}$ p is replaced by $p + 1$.

In some cases l is not large enough for the number of degrees of freedom to be positive. [30] suggests the use of SAM (Section 5) as a preprocessor to identify a so-called optimum end-member subset by choosing from the total set only the few end-members with the largest projections for a given observation.

3. Full Unmixing

To perform full unmixing one needs to know the spectra for all end-members present in the scene. In this case we demand that the the non-negative abundances add to 100%, i.e.,

$$\sum_{i=1}^{p} \alpha_i = \mathbf{1}^T \boldsymbol{\alpha} = 1 \text{ and } \alpha_i \geq 0, \tag{7}$$

where $\mathbf{1}$ is a $p \times 1$ vector of ones. The first constraint can be dealt with by introducing a Lagrange multiplier -2λ and minimizing $F = \boldsymbol{n}^T \boldsymbol{\Sigma}_n^{-1} \boldsymbol{n} + 2\lambda(\mathbf{1}^T \boldsymbol{\alpha} - 1)$ without constraints. The solution obtained by setting the partial derivatives $\partial F/\partial \boldsymbol{\alpha} = \mathbf{0}$ and $\partial F/\partial \lambda = 0$ is

$$\begin{bmatrix} \boldsymbol{M}^T \boldsymbol{\Sigma}_n^{-1} \boldsymbol{M} & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ \lambda \end{bmatrix} = \begin{bmatrix} \boldsymbol{M}^T \boldsymbol{\Sigma}_n^{-1} \boldsymbol{r} \\ 1 \end{bmatrix}. \tag{8}$$

The latter constraint can be dealt with by means of methods from convex quadratic programming, see Section 7 on computer programs.

In [53] an iterative scheme for full unmixing based on the RMSE image is suggested. This image can be used to select additional end-members or replace an existing end-member with one from a region of maximum RMSE.

Often, knowledge of all end-member spectra is not available. Therefore partial unmixing methods where we estimate the presence of one or a few desired, known spectra only are important. α_0 with a column of ones in $\boldsymbol{M}$ above will to some extent reflect the presence of end-members not accounted for in $\boldsymbol{M}$ and so will R^2 and RMSE.

4. Partial Unmixing

Partial unmixing builds on the usual linear mixture model in Eq. (2). We split the $\boldsymbol{M}\boldsymbol{\alpha}$ term into two terms, one which is the desired, known end-member $\boldsymbol{d}$ with a corresponding abundance α_p (without loss of generality we place $\boldsymbol{d}$ in the last column of $\boldsymbol{M}$), and one which consists of the undesired (and often unknown) end-members $\boldsymbol{U}$ with a corresponding $(p-1) \times 1$ vector, $\boldsymbol{\gamma}$, of abundances. $\boldsymbol{U}$ contains the first $p-1$ columns of $\boldsymbol{M}$ and $\boldsymbol{\gamma}$ contains the first $p-1$ elements of $\boldsymbol{\alpha}$. Hence

$$\begin{aligned} \boldsymbol{r} &= \boldsymbol{M}\boldsymbol{\alpha} + \boldsymbol{n} \\ &= \boldsymbol{d}\alpha_p + \boldsymbol{U}\boldsymbol{\gamma} + \boldsymbol{n}. \end{aligned} \tag{9}$$

$\boldsymbol{U}\boldsymbol{\gamma}$ is often termed the interference. In partial unmixing we want to develop methods to eliminate or minimize the effect of $\boldsymbol{U}$ and $\boldsymbol{\gamma}$. Often the term matched filtering is applied to such methods.

Alternatively, we can set up simultaneous partial unmixing constraints in convex quadratic programming by demanding that the non-negative abundances add to a quantity less than or equal to 100%, i.e.,

$$\sum_{i=1}^{p} \alpha_i = \mathbf{1}^T \boldsymbol{\alpha} \leq 1 \quad \text{and} \quad \alpha_i \geq 0. \tag{10}$$

4.1. *Orthogonal Subspace Projection (OSP)*

The idea in OSP [16, 33] is to project $\boldsymbol{r}$ onto a subspace orthogonal to $\boldsymbol{U}$. If we inspired by Eq. (6) apply the $l \times l$ matrix $\boldsymbol{P} = \boldsymbol{I} - \boldsymbol{U}(\boldsymbol{U}^T\boldsymbol{U})^{-1}\boldsymbol{U}^T$ we obtain

$$\begin{aligned} \boldsymbol{P}\boldsymbol{r} &= \boldsymbol{P}\boldsymbol{d}\alpha_p + \boldsymbol{U}\boldsymbol{\gamma} - \boldsymbol{U}(\boldsymbol{U}^T\boldsymbol{U})^{-1}\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{\gamma} + \boldsymbol{P}\boldsymbol{n} \\ &= \boldsymbol{P}\boldsymbol{d}\alpha_p + \boldsymbol{P}\boldsymbol{n}. \end{aligned} \tag{11}$$

$\boldsymbol{\gamma}$ and $\boldsymbol{U}$ are removed from the linear mixture model but as with full unmixing we need $\boldsymbol{U}$, i.e., we need all the end-member spectra, both desired (known) and undesired (typically unknown). However, OSP can be used to remove the effect of known spectra collected in $\boldsymbol{U}$. Another way of removing undesired signal based on band ratios is hinted in [8].

In [47] it is shown that full linear unmixing and OSP as used above and described in [16] are identical (except that OSP is computationally slightly more expensive).

For a noise subspace projection method, see [52].

4.2. *Constrained Energy Minimization (CEM)*

Constrained energy minimization, CEM [18, 33, 45, 49], builds on the linear mixture model in Eq. (9). In CEM we project $\boldsymbol{r}$ onto $\boldsymbol{w}$ with the intent to highlight presence of the desired end-member, and to suppress

the presence of the undesired end-members and noise. We do this by requesting the following

1. we want the output (the projected value) to be one when we project the desired spectrum, $\boldsymbol{d}$, i.e., we want $\boldsymbol{w}^T\boldsymbol{d} = 1$;
2. in general, we want the output to be close to zero, i.e., we want its expected value to be 0 or $\mathrm{E}\{\boldsymbol{w}^T\boldsymbol{r}\} = 0$;
3. also, we want to minimize the expected value of the squared difference between the output, $\boldsymbol{w}^T\boldsymbol{r}$, and the desired output, 0, i.e., we want to minimize $\mathrm{E}\{(\boldsymbol{w}^T\boldsymbol{r} - 0)^2\}$.

Since $\mathrm{E}\{\boldsymbol{w}^T\boldsymbol{r}\} = 0$ we get $\mathrm{E}\{(\boldsymbol{w}^T\boldsymbol{r} - 0)^2\} = \mathrm{V}\{\boldsymbol{w}^T\boldsymbol{r}\}$. Hence the job is to minimize $\mathrm{V}\{\boldsymbol{w}^T\boldsymbol{r}\} = \boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}$ with the constraint $\boldsymbol{w}^T\boldsymbol{d} = 1$. $\boldsymbol{\Sigma}$ is the dispersion matrix of $\boldsymbol{r}$. To do this we introduce a Lagrange multiplier -2λ and minimize $F = \boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w} + 2\lambda(\boldsymbol{w}^T\boldsymbol{d} - 1)$ without constraints. This is done by setting the partial derivatives $\partial F/\partial \boldsymbol{w} = \boldsymbol{0}$ and $\partial F/\partial\lambda = 0$. This leads to

$$\begin{bmatrix} \boldsymbol{\Sigma} & \boldsymbol{d} \\ \boldsymbol{d}^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{w} \\ \lambda \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ 1 \end{bmatrix} \tag{12}$$

or

$$\boldsymbol{w} = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{d}}{\boldsymbol{d}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{d}} \tag{13}$$

with $\lambda = -1/\boldsymbol{d}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{d}$. $(\boldsymbol{w}^T\boldsymbol{r})^2$ the expectation of which is minimized is often termed "energy", hence the name CEM.

Note that nothing in the above ensures the ideal: $0 \le \boldsymbol{w}^T\boldsymbol{r} \le 1$. On the contrary, we have requested that $\mathrm{E}\{\boldsymbol{w}^T\boldsymbol{r}\} = 0$ which means that some projections must necessarily be negative.

As opposed to OSP and full linear unmixing, CEM does not require knowledge of all end-member spectra. Only the desired spectrum is needed.

4.3. An Eigenvalue Formulation Alternative to CEM

As an alternative approach to CEM consider Eq. (9) and the projection $\boldsymbol{w}^T\boldsymbol{r}$ again

$$\boldsymbol{w}^T\boldsymbol{r} = \boldsymbol{w}^T\boldsymbol{d}\alpha_p + \boldsymbol{w}^T\boldsymbol{U}\boldsymbol{\gamma} + \boldsymbol{w}^T\boldsymbol{n}. \tag{14}$$

Consider the variance of $\boldsymbol{w}^T\boldsymbol{r}$

$$\begin{aligned}\mathrm{V}\{\boldsymbol{w}^T\boldsymbol{r}\} = {} & \mathrm{V}\{\boldsymbol{w}^T\boldsymbol{d}\alpha_p\} + \mathrm{V}\{\boldsymbol{w}^T\boldsymbol{U}\boldsymbol{\gamma}\} + \mathrm{V}\{\boldsymbol{w}^T\boldsymbol{n}\} \\ & + 2\,\mathrm{Cov}\{\boldsymbol{w}^T\boldsymbol{d}\alpha_p, \boldsymbol{w}^T\boldsymbol{U}\boldsymbol{\gamma}\}\end{aligned} \tag{15}$$

where we assume no covariation between the abundance of the desired spectrum and noise, and the abundances of the undesired spectra and noise. $\mathrm{Cov}\{\cdot\}$ denotes covariance. This can be written as

$$\begin{aligned}\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w} = {} & \mathrm{V}\{\alpha_p\}\boldsymbol{w}^T\boldsymbol{d}\boldsymbol{d}^T\boldsymbol{w} + \boldsymbol{w}^T\boldsymbol{U}\,\mathrm{D}\{\boldsymbol{\gamma}\}\boldsymbol{U}^T\boldsymbol{w} \\ & + \boldsymbol{w}^T\boldsymbol{\Sigma}_n\boldsymbol{w} + 2\boldsymbol{w}^T\boldsymbol{d}\,\mathrm{Cov}\{\alpha_p, \boldsymbol{\gamma}\}\boldsymbol{U}^T\boldsymbol{w} \\ = {} & \mathrm{V}\{\alpha_p\}\boldsymbol{w}^T\boldsymbol{d}\boldsymbol{d}^T\boldsymbol{w} + \boldsymbol{w}^T\boldsymbol{E}\boldsymbol{w}\end{aligned} \tag{16}$$

where $\boldsymbol{E}$ represents all undesired effects, namely dispersions of interference and noise, and covariance between abundances of desired and undesired spectra. $\boldsymbol{E}$ is unknown. $\mathrm{D}\{\cdot\}$ denotes dispersion. From this we get

$$1 = \mathrm{V}\{\alpha_p\}\frac{\boldsymbol{w}^T\boldsymbol{d}\boldsymbol{d}^T\boldsymbol{w}}{\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}} + \frac{\boldsymbol{w}^T\boldsymbol{E}\boldsymbol{w}}{\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}}. \tag{17}$$

To minimize the variance of all the undesired effects we must minimize the last term on the right-hand-side of Eq. (17) and therefore since the sum is constant and $\mathrm{V}\{\alpha_p\} \ge 0$ we must maximize the Rayleigh coefficient in the first term. Since $\boldsymbol{d}\boldsymbol{d}^T$ is rank 1 we get one solution only namely the $\boldsymbol{w}$ that satisfies the generalized eigenvalue problem

$$\boldsymbol{d}\boldsymbol{d}^T\boldsymbol{w} = \rho\boldsymbol{\Sigma}\boldsymbol{w}. \tag{18}$$

The solution $\boldsymbol{w}$ is proportional to $\boldsymbol{\Sigma}^{-1}\boldsymbol{d}$ which gives $\rho = \boldsymbol{d}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{d}$. If we want $\boldsymbol{w}^T\boldsymbol{d} = 1$ we get the same solution as in Eq. (13) with $\rho = -1/\lambda$.

For hyperspectral data these operations can be performed on (a subset of) orthogonally transformed data such as signal maximum autocorrelation factors, MAFs, or signal minimum noise fractions, MNFs [6, 7, 12, 15, 25, 28, 32, 35, 39, 42, 43, 50, 51]. In this case matrix inversion is not needed since $\boldsymbol{\Sigma}_M = \boldsymbol{I}$. Hence $\boldsymbol{w}_M = \boldsymbol{d}_M/\boldsymbol{d}_M^T\boldsymbol{d}_M$ where the subscript M denotes dispersion, projection vector and desired spectrum after the MAF or MNF transformation. If the data are not sampled on a regular grid MAF/MNF analysis as described in [32, 37, 38, 40] can be applied.

4.4. Iterated CEM

From Eq. (17) we see that for the first term on the right-hand-side which represents the desired spectrum to dominate over the second term which represents undesired effects, we could estimate $\boldsymbol{\Sigma}$ from the pixels

where $\boldsymbol{d}$ is present rather than over the entire image. Thus $\boldsymbol{\Sigma}$ in a second iteration could be estimated by weighting with $\boldsymbol{w}^T\boldsymbol{r}$ from the first iteration or we could threshold $\boldsymbol{w}^T\boldsymbol{r}$ and estimate $\boldsymbol{\Sigma}$ only where $\boldsymbol{w}^T\boldsymbol{r}$ is large. Of course, more iterations could be performed (and stopped when relative changes become small).

4.5. *Target Constrained Interference Minimized Filter (TCIMF)*

In case we have several desired spectra, say t, as well as several undesired spectra, say q, we project $\boldsymbol{r}$ onto $\boldsymbol{w}$ to highlight all the desired spectra written as columns in the $l \times t$ matrix $\boldsymbol{D}$, to null all the undesired spectra written as columns in the $l \times q$ matrix $\boldsymbol{U}$ and to minimize noise. This is done by replacing the CEM constraint $\boldsymbol{w}^T\boldsymbol{d} = 1$ with $\boldsymbol{w}^T[\boldsymbol{DU}] = [\mathbf{1}^T\ \mathbf{0}^T]$ where $[\boldsymbol{DU}]$ is an $l \times (t+q)$ matrix of desired and undesired spectra as columns, $\mathbf{1}$ is a $t \times 1$ vector of ones and $\mathbf{0}$ is a $q \times 1$ vector of zeros. This leads to

$$\boldsymbol{w} = \boldsymbol{\Sigma}^{-1}[\boldsymbol{DU}]([\boldsymbol{DU}]^T\boldsymbol{\Sigma}^{-1}[\boldsymbol{DU}])^{-1}\begin{bmatrix}\mathbf{1}\\ \mathbf{0}\end{bmatrix}. \quad (19)$$

In [44] this is called the target constrained interference minimized filter or TCIMF. Obviously this filter can be applied in a MAF/MNF sub-space also. Also iteration applies in this case.

5. Spectral Angle Mapping (SAM)

As a supposedly more physically oriented way of establishing a measure of closeness to a desired spectrum, spectral angle mapping, SAM, has been suggested [21]. In SAM the angle between the desired spectrum, $\boldsymbol{d}$, and the spectrum in each pixel, $\boldsymbol{r}$, is measured. This angle is the inverse cosine of the normalized inner product $\boldsymbol{d}^T\boldsymbol{r}/(\|\boldsymbol{d}\|\ \|\boldsymbol{r}\|)$. Apart from the scaling this inner product corresponds to CEM with $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}$, i.e., CEM without allowing for the covariance between the spectral bands. The result from SAM is ideally insensitive to illumination effects.

6. Semi-parametric Unmixing (SPU)

As an extension to the ordinary linear model

$$Y = (\,\alpha_0 + \,)\sum_{i=1}^{p}\alpha_i X_i + \epsilon \quad (20)$$

consider an additive model

$$Y = \sum_{i=1}^{p} f_i(X_i) + \epsilon, \quad (21)$$

where f_i which can subsume α_0 if present are smooth functions, typically (cubic smoothing) splines but (locally weighted) running-line or kernel smoothers could be used also. Based on the conditional expectations $\mathrm{E}\{Y - \sum_{k\neq i} f_k(X_k) \mid X_i\},\ i = 1, \ldots, p$ where f_k can be initialized by ordinary least squares regression, we see that the additive model can be estimated by so-called backfitting: fit the smooth functions f_i one at a time by smoothing the residuals $Y - \sum_{k\neq i} f_k(X_k)$ against X_i using a 1-D smoother. Known in numerical analysis as the Gauss-Seidel algorithm the process is repeated until convergence is obtained. Backfitting is also known as iterative smoothing of partial residuals.

Here, consider the simpler combined linear and additive model with only one non-linear term, the so-called semi-parametric model

$$\boldsymbol{r}(x, y) = \boldsymbol{M}\boldsymbol{\alpha}(x, y) + \boldsymbol{f}(x, y) + \boldsymbol{n}(x, y), \quad (22)$$

where $\boldsymbol{f}(x, y)$ which can subsume $\alpha_0(x, y)$ if present consists of combined unknown spectra described by a smooth function (of wavelength λ), again typically a spline. We note that although written as in the full unmixing case this model specifically allows for unknown spectra through the covariate $\boldsymbol{f}$. We do not consider the estimation of $\boldsymbol{f}$ itself interesting but in the presence of unknown spectra we need to estimate $\boldsymbol{f}$ ideally to avoid bias in $\boldsymbol{\alpha}$. As in the fully linear case, n_i is white noise, i.e., n_i is independent, identically distributed (iid) Gaussian with zero mean and variance σ^2, $\boldsymbol{n} \in N(\mathbf{0}, \sigma^2\boldsymbol{I})$. The semi-parametric model can be estimated directly without iteration [17].

For further description of (generalized) additive models see [5, 17, 46, 56]. [46] also introduces mean weighted least squares as an alternative to backfitting.

Other non-linear techniques for solving the mixing problem not described here include the expectation maximization algorithm [9] and artificial neural networks [4].

7. Computer Programs

Eight computer programs written at IMM are useful in this type of analysis, `maf`, `osp`, `project`, `tcimf`, `seed`, `eqndist`, `spam` and `unmix`.

`maf` finds principal components, (rotated) principal factors, maximum autocorrelation factors, minimum noise fractions, canonical discriminant functions, (multiset) canonical variates and linear combinations that give maximal multivariate differences of two sets of variables [39]. All covariance matrices are found by the method of provisional means [10]. The eigenvalue problems associated with the analysis are solved by means of LAPACK routines [1]. For a fuller description, see [32].

`osp` performs orthogonal subspace projection of one or more spectra, Eq. (11). Typically this is done to reduce the influence of undesired spectra.

`project` projects data in feature space onto a unit vector representing a desired end-member spectrum.

`tcimf` performs target constrained interference minimized filtering, Eq. (19), including constrained energy minimization, Eq. (13).

`seed` grows a training area from one or a few pixels by requesting spatial as well as spectral closeness. Spatial closeness is ensured by requesting 8-neighbor connectivity. Spectral closeness is ensured by requesting low Euclidean or Mahalanobis distance in feature space. For a fuller description, see [13, 26, 27, 41].

`eqndist` performs a Wishart distribution based test to check if pairs of data classes simultaneously have equal mean and dispersion, see [2]. Here this is done to check that no two regions grown by `seed` actually represent the same class.

`spam` performs spectral angle mapping, Section 5.

`unmix` performs full unmixing either without constraints or with the natural constraints that the non-negative abundances sum to one, and marginal (one spectral band at a time) or simultaneous (more bands at a time) partial unmixing with the natural constraints that the non-negative abundances sum to a quantity not greater than one. The unconstrained OLS problem (Eq. (6)) is solved by LINPACK routines [11]. The constrained NNLS problems (Eq. (7) and (10)) are solved by a linearly constrained least squares algorithm, LSSOL [14], which solves the problem: minimize $\frac{1}{2}\|\boldsymbol{r}-\boldsymbol{M}\boldsymbol{\alpha}\|^2$ over $\boldsymbol{\alpha}$ in this case with $\alpha_i \geq 0$ and $\mathbf{1}^T\boldsymbol{\alpha}=1$ respectively $\mathbf{1}^T\boldsymbol{\alpha}\leq 1$. By working in MAF/MNF space we allow for inter-variable covariances in the spectral distance measure rather than working with simply the Euclidean distance. The simultaneous partial unmixing can be performed iteratively as described for the marginal partial unmixing, Section 4.4.

The above computer programs are written to comply with the HIPS standard [22–24].

Estimation in the semi-parametric model is done by means of the S/S-PLUS `gam` function [5].

8. Case Studies

Two sets of data are used in the case studies. One is based one simple generated data, another on hyperspectral imaging spectrometer data. The generated data are well suited for simple tests of both full and partial unmixing methods because the truth is known. On the other hand the data generated are so simple that the results do not reflect a realistic behavior of the methods. Therefore also real data are applied.

8.1. Simple Generated Data

The data used consist of two bands, one with a centered horizontal bar and one with a centered vertical bar. The images are 130×130 pixels with 10 pixels wide bars. Both bands have Gaussian noise with standard deviation equal to half the difference between foreground (which equals 1.0) and background (which equals 0.0) added. Figure 1 shows the two bands without noise in the first column, the two bands with noise in the second column, the CEM results (Eq. (13)) stretched linearly from minimum to maximum in the third column, and the CEM results stretched linearly from 0 to 1 in the fourth column. In spite of the high noise level CEM nicely picks up the two desired spectra.

Iterated CEM images (second iteration, Section 4.4) are shown in columns two and four in Fig. 2. $\boldsymbol{\Sigma}$ in the second iteration is estimated by weighting with $\boldsymbol{w}^T\boldsymbol{r}$ from Eq. (13) stretched linearly from 0 to 1. To facilitate comparison the original CEM images are shown

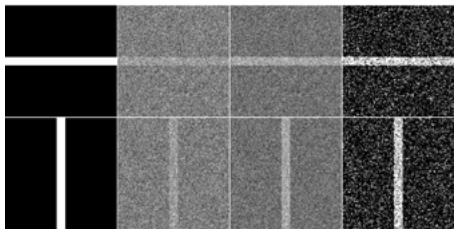

Figure 1. Two generated bands (without noise in column one, with noise in column two) and the CEM results, Eq. (13) (columns three and four).

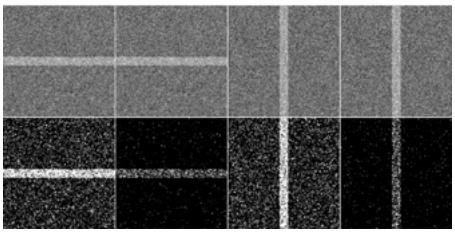

Figure 2. Two generated bands, original CEM results, Eq. (13) (columns one and three) and iterated CEM results, Section 4.4 (columns two and four).

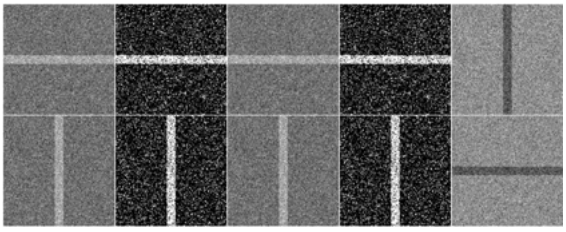

Figure 3. Two generated bands, original CEM results, Eq. (13) (columns one and two), TCIMF results, Eq. (19) (columns three and four) and differences (column five).

again in columns one and three. In the top row all quantities are stretched linearly between minimum and maximum, in the bottom row all quantities are stretched linearly between 0 and 1. Iterated CEM seems to suppress noise better than ordinary CEM at the cost of missing true positives.

TCIMF images (Eq. (19)) with one desired and one undesired spectrum are shown in columns three and four in Fig. 3. Columns one and two repeat the CEM results. Columns two and four are stretched linearly between 0 and 1. As the results are very similar column five shows the CEM result subtracted from the TCIMF result. Columns one, three and five are stretched linearly between minimum and maximum. In this example we see that since dark gray levels represent low values TCIMF tends to suppress what is not desired rather than put extra highlight on what is desired.

Figure 4 shows results from a full unmixing without constraints based on ordinary least squares, Eq. (6). Column one is abundances of band 1, column two is abundances of band 2, column three is R^2 and column four is RMSE. In the top row all quantities are stretched linearly between minimum and maximum, in the bottom row all quantities are stretched linearly between 0 and 1 (which makes no sense for RMSE). Figure 5

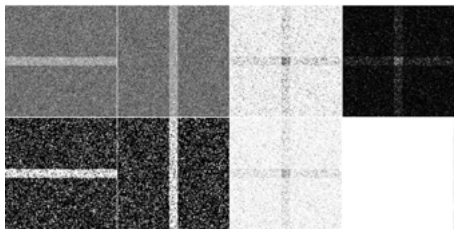

Figure 4. Two generated bands, full unmixing without constraints based on OLS, Eq (6).

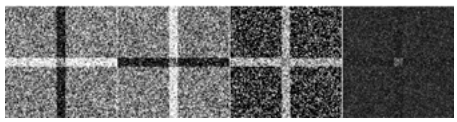

Figure 5. Two generated bands, full unmixing with constraints based on NNLS, Eq. (7).

shows the same results for the full, constrained unmixing based on non-negative least squares, Eq. (7).

Both full unmixings are carried out with three variables, namely the two generated bands and their product. This causes the cross-shaped appearances of the R^2 and RMSE images. Compared to unconstrained unmixing we see that the constrained unmixing in this case better suppresses the undesired spectrum at the cost of more false positives.

8.2. AVIRIS Data

The data used here is the 30 bands subset of AVIRIS data, [3, 54, 55], over a small part of the Mojave Desert, California, USA, that come with the LinkWinds software [19]. These bands cover the spectral range 0.52–2.33 μm. The images have 180 rows and 360 columns. Figure 6 shows every other of the 30 bands (row-wise). AVIRIS (the Airborne Visible/Infra-Red Imaging Spectrometer) from NASA/JPL features 224 approximately 10 nm wide channels measured by 4 detectors covering the spectral range from 0.370–2.500 μm. Pixels are 20 m $\times$ 20 m.

The data are used in four different ways. First (Table 3), we generate new spectra from measured spectra by adding known abundances of spectra for end-members 1, 3 and 5. These spectra are used to test unmixing by OLS (Eq. (6)) and NNLS (Eq. (10)) in situations with known abundances and with different levels of independent, identically distributed (iid) Gaussian noise added. Second, (Table 4), similarly

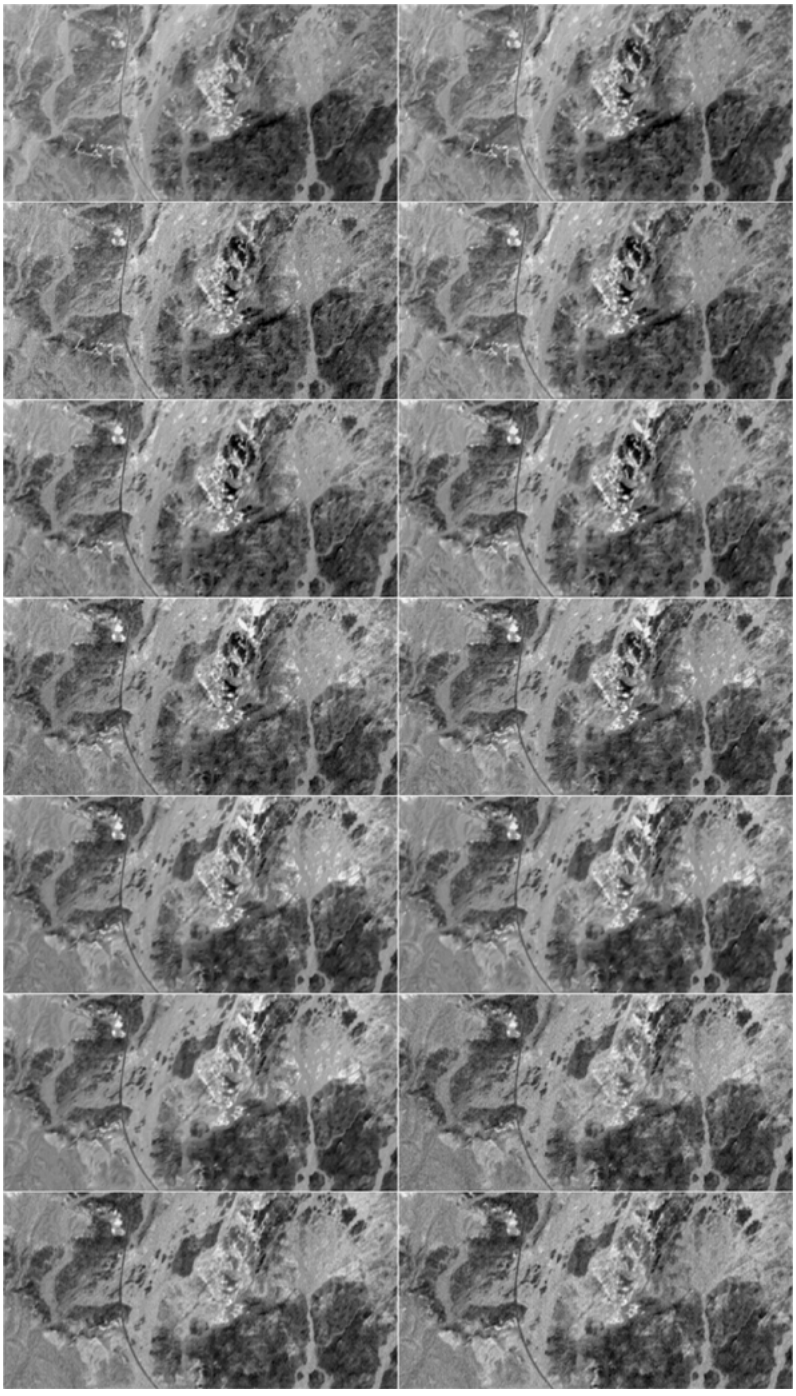

Figure 6. Every other of the 30 bands subset of AVIRIS data over the Mojave Dessert, California, USA.

generated spectra are used to test unmixing by OLS (Eq. (6)), NNLS (Eq. (10)) and SPU (Eq. (22)) in situations with known abundances and with different levels of iid Gaussian noise added, this time with one spectrum considered as unknown. Third, (Table 5), to avoid problems that may arise from results based on just one realization of the iid Gaussian noise added, generated spectra based on 100 noise realizations are used to test

Table 1. Dispersion and *correlation* matrices of the six end-member spectra.

	EM1	EM2	EM3	EM4	EM5	EM6
EM1	1645.927	2125.770	1608.495	2780.741	2641.368	2809.189
EM2	*0.9366*	3129.980	2154.799	3696.108	3121.738	3758.973
EM3	*0.9530*	*0.9257*	1730.953	2829.628	2609.729	2753.942
EM4	*0.9859*	*0.9503*	*0.9783*	4833.341	4505.905	4823.767
EM5	*0.9568*	*0.8201*	*0.9219*	*0.9525*	4629.797	4467.913
EM6	*0.9890*	*0.9596*	*0.9454*	*0.9910*	*0.9378*	4902.158

Table 2. Eigenvalues and amount of variance explained by principal components of dispersion and correlation matrices of the six end-member spectra.

	PC1	PC2	PC3	PC4	PC5	PC6
Dispersion matrix	$1.999 \cdot 10^4$	$7.021 \cdot 10^2$	$1.479 \cdot 10^2$	$2.609 \cdot 10^1$	$8.977 \cdot 10^0$	$5.961 \cdot 10^{-1}$
	95.8%	3.36%	0.709%	0.125%	0.0430%	0.00286%
Correlation matrix	5.737	0.1826	0.06645	0.01169	0.002441	0.0001492
	95.6%	3.04%	1.11%	0.195%	0.0407%	0.00249%

unmixing by OLS (Eq. (6)) and SPU (Eq. (22)) in situations with known abundances and with different levels of iid Gaussian noise added, again with one spectrum considered as unknown. Finally, the AVIRIS data are used directly to illustrate results of different partial unmixing methods.

To establish which regions in the image contain extreme values and therefore are potential end-members we look for the minimum and maximum values in the MAFs. The first 14 MAFs are shown in Fig. 7 (rowwise).

Since we don't want our partial unmixing results to be based on possible noise spectra we use training areas grown from the pixels with extreme values as seeds to calculate average spectra instead of using the spectra from the extreme pixels themselves directly [13, 26, 27, 41]. To ensure that training areas do not represent the same class a Wishart distribution based test to check that no pair of classes simultaneously have equal mean and dispersion is performed [2].

This is a "true remote sensing situation", we don't know what is on the ground. Our aim here is to illustrate the unmixing methods and not to classify or identify material on the ground. We arbitrarily choose six potential end-members corresponding to the extreme values of MAFs 1–3. Figure 8 shows the six training areas grown from these extremes by `seed`, see Section 7.

Table 1 shows dispersion and correlation matrices for the six end-member spectra. Numbers on and above the diagonal are from the dispersion matrix, numbers below the diagonal and in *italics* are from the correlation matrix. The high correlations in Table 1 are confirmed by the two eigenanalyses shown in Table 2. In both cases more than 95% of the total variation is explained by principal component one. Figure 9 shows all pairwise scatterplots of seven variables, namely wavelength and the six end-member spectra. Again the high correlations between the end-member spectra are confirmed but although correlations are high there are large differences in the levels between the six spectra especially in the $0.5\ \mu\text{m} < \lambda < 1.7\ \mu\text{m}$ region.

Table 3 shows results of unmixings of generated spectra by means of ordinary least squares regression (OLS, Eq. (6)) and non-negative least squares regression with weights adding to ≤ 1 (NNLS, Eq. (10)). The spectra are constructed as weighted sums of end-members 1, 3 and 5 with different amounts of iid Gaussian noise. The true $\alpha_0 = 0$ in this case and the true abundances are listed along with their estimates in Table 3. Also the true $\alpha_5 = 0$ here. The noise content is characterised by the signal-to-noise ratio (SNR) defined as the variance of the signal relative to the variance of the noise. Coefficients are listed whether they are significant or not. We see that apart from the intercept estimates in this situation with all end-members known OLS and NNLS perform fairly similarly for reasonable SNR levels (1000 and 100).

Table 4 shows results of simultaneous partial unmixings of generated spectra by means of ordinary least squares regression (OLS, Eq. (6)), non-negative

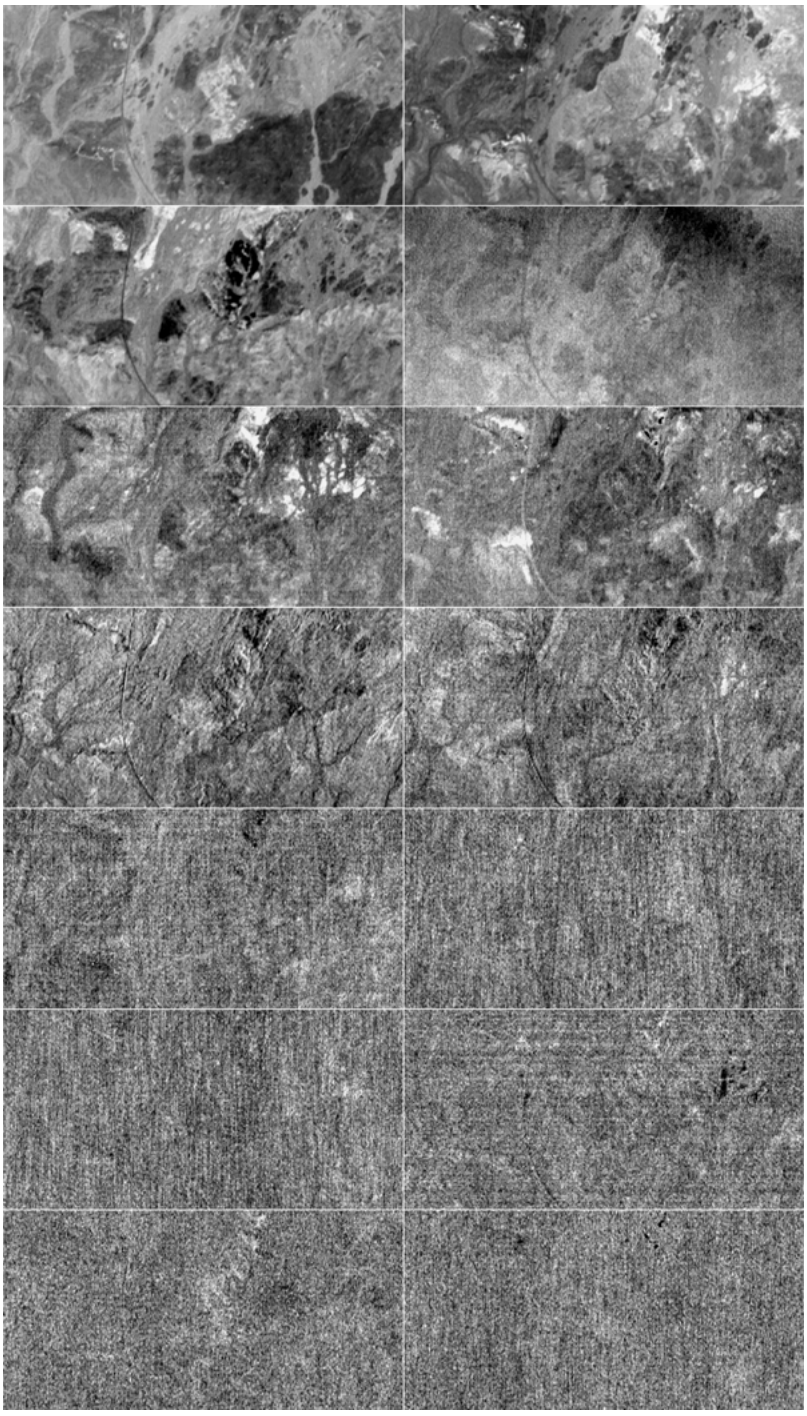

Figure 7. The first 14 MAFs of the AVIRIS data.

least squares regression with weights adding to ≤ 1 (NNLS, Eq. (10)) and semi-parametric unmixing (SPU, Eq. (22)). The spectra are constructed as weighted sums of end-members 1, 3 and 5 with different amounts of iid Gaussian noise (here we have a situation with no noise added also). End-member 5 is considered as unknown, i.e., only abundances of end-member spectra 1 and 3 are estimated (but in the presence of end-member

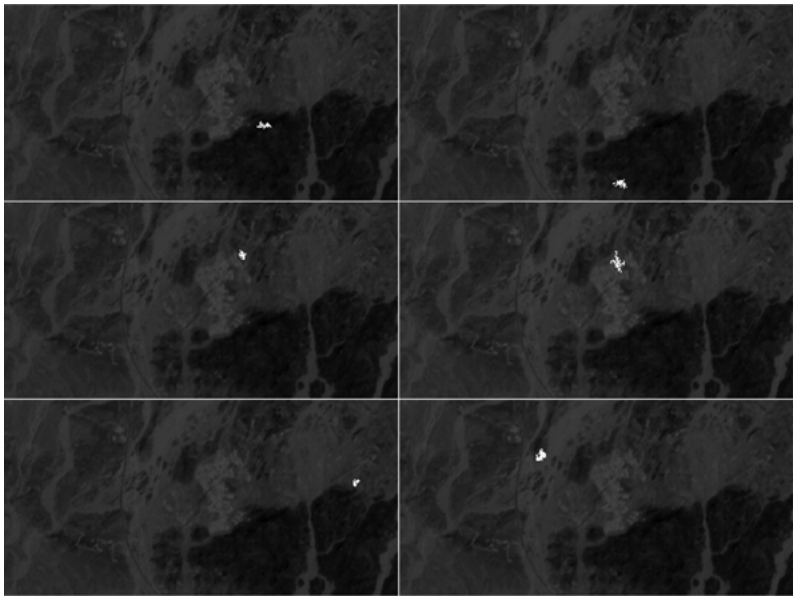

Figure 8. Seed grown training areas under which average spectra are calculated shown in white on top of MAF1. The plotting order is (row-wise) min MAF1, max MAF1, min MAF2, max MAF2, min MAF3, max MAF3.

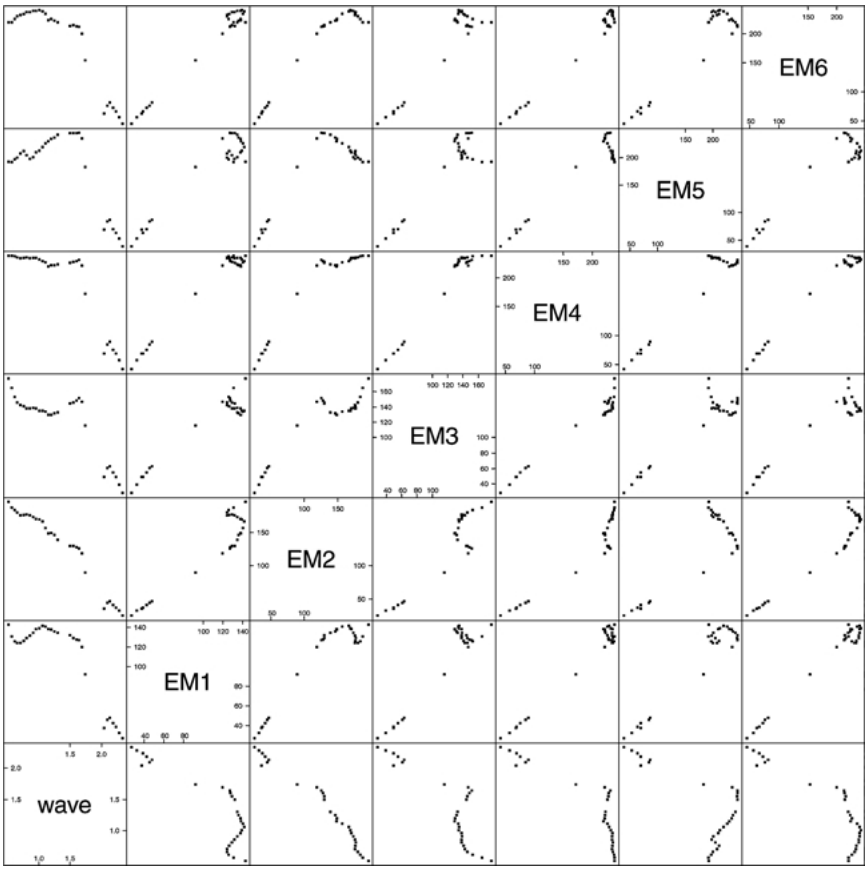

Figure 9. All pairwise scatterplots of wavelength and the six end-member spectra. The first column shows the actual spectra.

Table 3. Abundances and RMSEs estimated from generated spectra by means of ordinary least squares (OLS, Eq. 6) and non-negative least squares with weights adding to ≤ 1 (NNLS, Eq. 10). The spectra are constructed as weighted sums of end-members 1, 3 and 5 with different amounts of iid Gaussian noise.

				OLS					NNLS				
α_1	α_3	α_5	SNR	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_3$	$\hat{\alpha}_5$	RMSE	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_3$	$\hat{\alpha}_5$	RMSE
0.00	1.00	0.00	1000	0.05	−0.03	0.98	0.03	1.50	0.01	0.00	0.97	0.02	1.52
0.10	0.90	0.00	1000	0.05	0.07	0.91	0.01	1.63	0.01	0.07	0.91	0.01	1.63
0.20	0.80	0.00	1000	−0.37	0.22	0.80	−0.01	1.27	0.00	0.66	0.31	0.02	1.54
0.30	0.70	0.00	1000	1.17	0.30	0.70	−0.01	1.08	0.00	0.29	0.71	0.00	1.15
0.40	0.60	0.00	1000	0.03	0.44	0.60	−0.02	1.51	0.00	0.39	0.61	0.00	1.59
0.50	0.50	0.00	1000	−1.48	0.49	0.51	0.01	1.04	0.00	0.50	0.49	0.01	1.16
0.60	0.40	0.00	1000	0.56	0.57	0.42	0.00	1.06	0.00	0.56	0.43	0.01	1.07
0.70	0.30	0.00	1000	1.13	0.67	0.30	0.02	1.50	0.01	0.66	0.31	0.02	1.54
0.80	0.20	0.00	1000	−0.53	0.79	0.22	−0.00	1.26	0.00	0.78	0.22	0.00	1.29
0.90	0.10	0.00	1000	0.65	0.87	0.10	0.02	1.14	0.01	0.86	0.10	0.02	1.16
1.00	0.00	0.00	1000	0.43	1.01	−0.03	0.01	1.26	0.01	0.98	0.00	0.01	1.33
0.00	1.00	0.00	100	−0.50	−0.10	1.05	0.03	4.00	0.00	0.00	1.00	0.00	4.12
0.10	0.90	0.00	100	−3.49	0.04	0.92	0.03	5.36	0.00	0.06	0.90	0.03	5.49
0.20	0.80	0.00	100	−0.76	0.28	0.85	−0.08	3.78	0.00	0.17	0.83	0.00	4.17
0.30	0.70	0.00	100	1.87	0.37	0.66	−0.03	4.18	0.00	0.32	0.68	0.00	4.27
0.40	0.60	0.00	100	−0.72	0.39	0.66	−0.03	4.64	0.00	0.35	0.65	0.00	4.69
0.50	0.50	0.00	100	−3.54	0.52	0.50	0.01	2.85	0.00	0.53	0.47	0.01	3.10
0.60	0.40	0.00	100	−0.10	0.65	0.31	0.04	3.58	0.00	0.65	0.31	0.04	3.58
0.70	0.30	0.00	100	−2.13	0.68	0.27	0.05	5.16	0.00	0.69	0.26	0.04	5.21
0.80	0.20	0.00	100	3.41	0.92	0.15	−0.06	4.75	0.00	0.82	0.18	0.00	5.02
0.90	0.10	0.00	100	−4.59	1.08	0.05	−0.06	4.02	0.00	1.00	0.00	0.00	4.55
1.00	0.00	0.00	100	−0.39	0.95	−0.04	0.06	3.84	0.00	0.91	0.00	0.05	3.89
0.00	1.00	0.00	10	−15.17	0.09	1.18	−0.09	11.84	0.00	0.00	0.99	0.01	13.22
0.10	0.90	0.00	10	−2.65	0.11	0.97	−0.06	15.28	0.00	0.04	0.93	0.00	15.35
0.20	0.80	0.00	10	−0.31	−0.01	0.92	0.05	13.52	0.00	0.00	0.92	0.05	13.52
0.30	0.70	0.00	10	7.43	0.04	0.81	0.04	14.26	0.07	0.06	0.87	0.05	14.49
0.40	0.60	0.00	10	3.55	0.69	0.54	−0.14	11.74	0.00	0.39	0.58	0.03	12.32
0.50	0.50	0.00	10	5.13	0.45	0.39	0.08	14.22	0.05	0.43	0.44	0.08	14.33
0.60	0.40	0.00	10	−0.81	0.62	0.53	−0.09	13.07	0.00	0.49	0.50	0.00	13.21
0.70	0.30	0.00	10	−4.05	0.50	0.38	0.10	11.94	0.00	0.52	0.35	0.09	12.02
0.80	0.20	0.00	10	0.94	0.99	0.29	−0.16	13.38	0.00	0.68	0.31	0.01	13.92
0.90	0.10	0.00	10	−6.55	0.72	0.10	0.13	13.04	0.00	0.75	0.05	0.12	13.23
1.00	0.00	0.00	10	−5.16	1.13	−0.20	0.08	13.28	0.00	0.92	0.00	0.04	13.82
0.00	1.00	0.00	1	−0.86	0.47	0.42	0.14	38.84	0.00	0.44	0.41	0.15	38.84
0.10	0.90	0.00	1	−47.68	0.41	1.53	−0.42	48.21	0.00	0.00	0.91	0.00	51.93
0.20	0.80	0.00	1	−3.46	0.04	0.53	0.35	41.16	0.00	0.05	0.50	0.34	41.18
0.30	0.70	0.00	1	−16.81	−1.46	1.72	0.53	55.07	0.00	0.00	0.87	0.13	57.04
0.40	0.60	0.00	1	−1.26	0.37	0.74	0.01	39.79	0.00	0.10	0.74	0.16	39.94
0.50	0.50	0.00	1	−44.75	0.83	0.18	0.24	43.46	0.00	0.80	0.00	0.15	46.21

(*Continued on next page*.)

Table 3. (*Continued*).

				OLS					NNLS				
α_1	α_3	α_5	SNR	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_3$	$\hat{\alpha}_5$	RMSE	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_3$	$\hat{\alpha}_5$	RMSE
0.60	0.40	0.00	1	18.44	0.47	0.29	0.05	30.28	0.08	0.40	0.45	0.07	30.93
0.70	0.30	0.00	1	−19.06	0.70	0.91	−0.35	32.46	0.00	0.23	0.66	0.00	34.04
0.80	0.20	0.00	1	21.48	0.87	0.09	−0.12	31.93	0.09	0.64	0.26	0.00	32.82
0.90	0.10	0.00	1	−9.35	0.53	0.38	0.11	32.32	0.00	0.56	0.30	0.09	32.48
1.00	0.00	0.00	1	11.09	−0.20	0.59	0.33	41.40	0.15	0.00	0.58	0.27	41.64

spectrum 5). We see that OLS in this situation with one unknown spectrum performs better than NNLS. Also, the RMSEs for OLS are lower than for NNLS. SPU tends to perform slightly better than OLS for reasonable SNR levels (1000 and 100) with high abundance (30%) of the unknown component. In all cases with reasonable SNR (1000 and 100) the RMSEs for SPU are lower than for OLS.

Table 4. Abundances and RMSEs estimated from generated spectra by means of ordinary least squares (OLS, Eq. (6)), non-negative least squares with weights adding to ≤ 1 (NNLS, Eq. (10)) and semi-parametric unmixing (SPU, Eq. (22)). The spectra are constructed as weighted sums of end-members 1, 3 and 5 with different amounts of iid Gaussian noise. End-member 5 is considered as unknown (and is therefore not estimated).

				OLS				NNLS				SPU			
α_1	α_3	α_5	SNR	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_3$	RMSE	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_3$	RMSE	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_3$	RMSE
0.00	0.90	0.10		0.50	0.14	0.92	2.04	0.00	0.00	1.00	7.46	−0.51	0.09	0.97	0.82
0.45	0.45	0.10		0.50	0.59	0.47	2.04	0.00	0.25	0.75	6.63	−0.51	0.54	0.52	0.82
0.90	0.00	0.10		0.50	1.04	0.02	2.04	0.00	0.70	0.30	6.63	−0.51	0.99	0.07	0.82
0.00	0.90	0.10	1000	1.15	0.15	0.91	2.39	0.00	0.00	1.00	7.37	4.05	0.05	0.98	1.62
0.45	0.45	0.10	1000	1.14	0.58	0.48	2.54	0.00	0.23	0.77	7.05	10.78	0.46	0.53	1.78
0.90	0.00	0.10	1000	−0.45	1.04	0.03	2.07	0.00	0.72	0.28	6.31	−2.17	1.01	0.06	1.67
0.00	0.90	0.10	100	0.90	0.13	0.92	4.49	0.00	0.00	1.00	7.82	−11.78	0.10	1.00	4.23
0.45	0.45	0.10	100	2.33	0.67	0.39	4.54	0.00	0.29	0.71	8.42	−3.12	0.67	0.40	4.37
0.90	0.00	0.10	100	1.78	1.02	0.02	4.73	0.00	0.71	0.29	7.53	38.18	0.76	0.08	3.31
0.00	0.90	0.10	10	7.44	0.12	0.89	10.62	0.00	0.00	1.00	13.14	−6.25	0.15	0.92	11.01
0.45	0.45	0.10	10	−4.32	0.63	0.49	13.97	0.00	0.21	0.79	16.12	39.37	0.28	0.61	14.62
0.90	0.00	0.10	10	4.961	1.19	−0.19	14.29	0.00	0.99	0.01	14.84	−42.01	1.27	−0.05	14.65
0.00	0.90	0.10	1	17.51	0.43	0.53	46.09	0.00	−0.00	1.00	47.34	357.59	−1.79	1.00	43.32
0.45	0.45	0.10	1	−10.80	−0.07	1.21	40.84	0.00	0.00	1.00	41.96	191.96	−1.40	1.51	40.17
0.90	0.00	0.10	1	−1.12	0.96	0.20	38.03	0.00	0.14	0.86	41.06	−8.33	1.19	0.03	40.25
0.00	0.80	0.20		1.00	0.29	0.84	4.07	0.00	−0.00	1.00	14.92	−1.02	0.17	0.93	1.64
0.40	0.40	0.20		1.00	0.69	0.44	4.07	0.00	0.01	0.99	13.27	−1.02	0.57	0.53	1.64
0.80	0.00	0.20		1.00	1.09	0.04	4.07	0.00	0.41	0.59	13.27	−1.02	0.97	0.13	1.64
0.00	0.80	0.20	1000	0.66	0.29	0.84	4.19	0.00	0.00	1.00	14.93	5.56	0.14	0.93	1.97
0.40	0.40	0.20	1000	0.64	0.69	0.43	3.93	0.00	0.04	0.96	12.85	−5.25	0.61	0.52	1.96
0.80	0.00	0.20	1000	1.97	1.08	0.03	4.41	0.00	0.41	0.59	13.30	5.34	0.93	0.14	2.07

(*Continued on next page*.)

Table 4. (*Continued*).

				OLS				NNLS				SPU			
α_1	α_3	α_5	SNR	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_3$	RMSE	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_3$	RMSE	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_3$	RMSE
0.00	0.80	0.20	100	1.45	0.31	0.80	5.19	0.00	0.00	1.00	14.18	10.14	0.09	0.94	3.42
0.40	0.40	0.20	100	1.01	0.71	0.42	5.93	0.00	0.00	1.00	14.97	−1.43	0.65	0.48	5.40
0.80	0.00	0.20	100	6.33	1.07	0.01	4.46	0.00	0.38	0.62	13.73	−3.21	0.97	0.14	2.93
0.00	0.80	0.20	10	16.73	0.20	0.80	11.29	0.00	0.00	1.00	18.69	−23.25	0.18	0.97	9.17
0.40	0.40	0.20	10	10.54	0.57	0.44	16.10	0.00	0.09	0.91	18.62	83.35	0.00	0.58	14.65
0.80	0.00	0.20	10	10.86	1.06	0.00	13.57	0.00	0.31	0.69	19.65	50.40	0.80	0.04	14.11
0.00	0.80	0.20	1	−46.01	0.43	1.00	45.91	0.00	0.04	0.96	49.27	41.91	−0.07	1.09	46.65
0.40	0.40	0.20	1	28.90	0.00	0.80	49.89	0.00	0.00	1.00	51.02	72.15	−1.13	1.57	48.79
0.80	0.00	0.20	1	6.96	1.27	−0.19	29.50	0.00	0.54	0.46	32.49	84.62	0.90	−0.20	30.87
0.00	0.90	0.30		1.51	0.43	0.75	6.11	0.00	−0.00	1.00	22.38	−1.53	0.26	0.90	2.46
0.35	0.35	0.30		1.51	0.78	0.40	6.11	0.00	0.00	1.00	20.32	−1.53	0.61	0.55	2.46
0.70	0.00	0.30		1.51	1.13	0.05	6.11	0.00	0.11	0.89	19.90	−1.53	0.96	0.20	2.46
0.00	0.90	0.30	1000	1.57	0.40	0.78	6.52	0.00	0.00	1.00	23.02	2.93	0.21	0.92	3.38
0.35	0.35	0.30	1000	1.74	0.83	0.36	6.01	0.00	0.00	1.00	20.57	2.23	0.63	0.51	2.82
0.70	0.00	0.30	1000	0.46	1.15	0.09	6.14	0.00	0.12	0.88	20.24	−5.73	0.99	0.19	2.38
0.00	0.90	0.30	100	1.03	0.42	0.76	6.98	0.00	0.00	1.00	21.94	10.41	0.12	0.97	3.83
0.35	0.35	0.30	100	−1.98	0.81	0.42	7.17	0.00	0.00	1.00	22.41	15.89	0.57	0.68	4.81
0.70	0.00	0.30	100	2.75	1.11	0.07	7.44	0.00	0.06	0.94	20.88	−6.80	0.96	0.22	3.94
0.00	0.90	0.30	10	2.58	0.66	0.55	12.12	0.00	0.00	1.00	27.16	−54.74	0.86	0.62	11.32
0.35	0.35	0.30	10	9.78	0.62	0.51	14.47	0.00	0.00	1.00	25.76	5.88	0.47	0.63	14.23
0.70	0.00	0.30	10	−2.32	1.17	0.04	15.72	0.00	0.17	0.83	24.41	34.59	0.75	0.21	13.81
0.00	0.90	0.30	1	−8.35	0.71	0.63	37.26	0.00	0.00	1.00	48.71	−278.16	2.14	0.58	36.20
0.35	0.35	0.30	1	−12.94	0.83	0.55	46.40	0.00	0.00	1.00	54.83	−214.52	2.03	0.45	47.05
0.70	0.00	0.30	1	−4.26	0.38	0.83	43.21	0.00	0.00	1.00	47.55	−84.99	1.34	0.29	44.31

Table 5 shows results of simultaneous partial unmixings of 100 realizations of generated spectra (see Fig. 10) by means of ordinary least squares regression (OLS, Eq. (6)) and semi-parametric unmixing (SPU, Eq. (22)). The spectra are constructed as immediately above, this time with 100 realizations of the iid Gaussian noise. Again, end-member 5 is considered as unknown. We see that SPU correctly tends to give more equal but not satisfactory estimates of α_1 and α_3 than OLS. Also, we see that the RMSEs for SPU are generally lower than for OLS especially for high SNR.

To give an impression of the appearance of some smooth functions in the semi-parametric model (Eq. (22)) Fig. 11 shows residuals from OLS with examples of differently smoothed versions of $\boldsymbol{f}$, see Section 6.

Figure 12 shows results from a non-negative least squares (NNLS) simultaneous partial unmixing (i.e., the non-negative abundances add to a quantity not greater than one, Eq. (10)) of the six chosen end-member spectra. Abundances and R^2 are stretched linearly from 0 to 1. No intercept term is estimated in this case. Where RMSE is high R^2 is low and vice versa. High RMSE regions potentially hold additional end-members.

Figure 13 shows the resulting abundance images as estimated from the first nine MAFs by CEM (Eq. (13)) stretched linearly from 0 to 1.

Figure 14 shows second iteration abundance images as estimated from the first nine MAFs by CEM (Section 4.4) stretched linearly from 0 to 1. $\boldsymbol{\Sigma}$ in the second iteration is estimated by weighting with $\boldsymbol{w}^T\boldsymbol{r}$ stretched linearly from 0 to 1.

Table 5. Abundances and RMSEs estimated from 100 realizations of generated spectra by means of ordinary least squares (OLS, Eq. (6)) and semi-parametric unmixing (SPU), Eq. (22). First rows give mean values, second rows give standard deviations. The spectra are constructed from end-members 1, 3 and 5 with different amounts of iid Gaussian noise. End-member 5 is considered as unknown (and is therefore not estimated).

				OLS				SPU			
α_1	α_3	α_5	SNR	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_3$	RMSE	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_3$	RMSE
0.45	0.45	0.10	1000	0.36	0.59	0.47	2.48	−0.82	0.53	0.52	1.59
				0.74	0.02	0.02	0.26	4.91	0.04	0.03	0.24
0.45	0.45	0.10	100	0.57	0.59	0.47	4.82	−0.07	0.53	0.52	4.40
				2.43	0.07	0.06	0.64	16.24	0.11	0.08	0.64
0.45	0.45	0.10	10	0.71	0.60	0.46	13.73	0.41	0.54	0.51	13.65
				7.86	0.19	0.19	1.88	42.84	0.32	0.26	1.95
0.45	0.45	0.10	1	−2.79	0.57	0.51	43.66	−27.91	0.65	0.56	43.55
				24.06	0.79	0.77	5.40	174.92	1.15	0.95	5.58
0.25	0.25	0.50	1000	2.49	0.97	0.34	10.29	−2.34	0.68	0.58	4.47
				1.01	0.03	0.03	0.36	5.88	0.04	0.03	0.28
0.25	0.25	0.50	100	2.46	0.98	0.33	11.30	−2.27	0.70	0.56	6.57
				2.91	0.08	0.08	0.97	17.69	0.14	0.10	0.78
0.25	0.25	0.50	10	2.37	0.94	0.36	19.56	−2.15	0.64	0.62	17.17
				8.06	0.28	0.28	2.88	60.01	0.41	0.30	2.54
0.25	0.25	0.50	1	3.35	1.05	0.26	54.34	13.23	0.82	0.38	53.71
				33.46	0.82	0.80	7.96	186.46	1.33	0.96	7.75

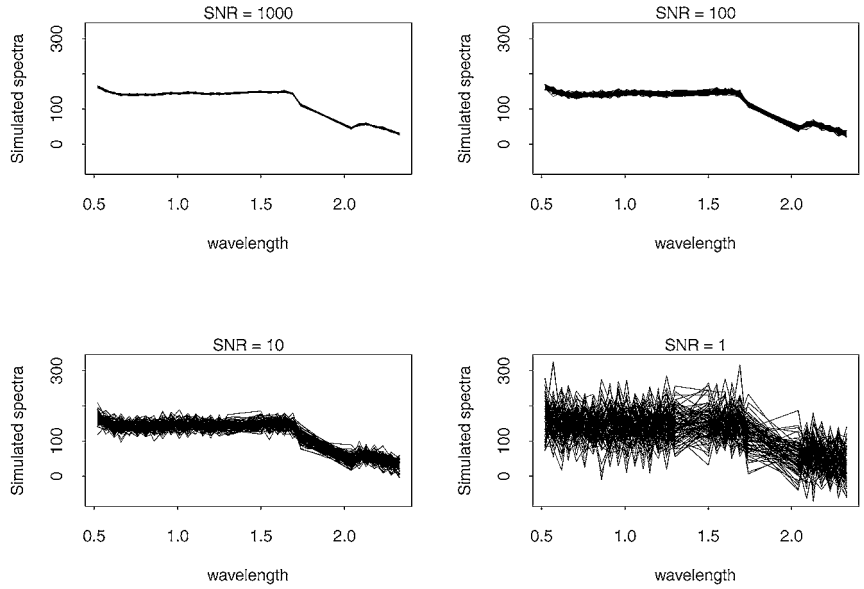

Figure 10. 100 realizations of generated spectra with different amounts of iid Gaussian noise. $\alpha_1 = \alpha_3 = 0.45$ and $\alpha_5 = 0.10$.

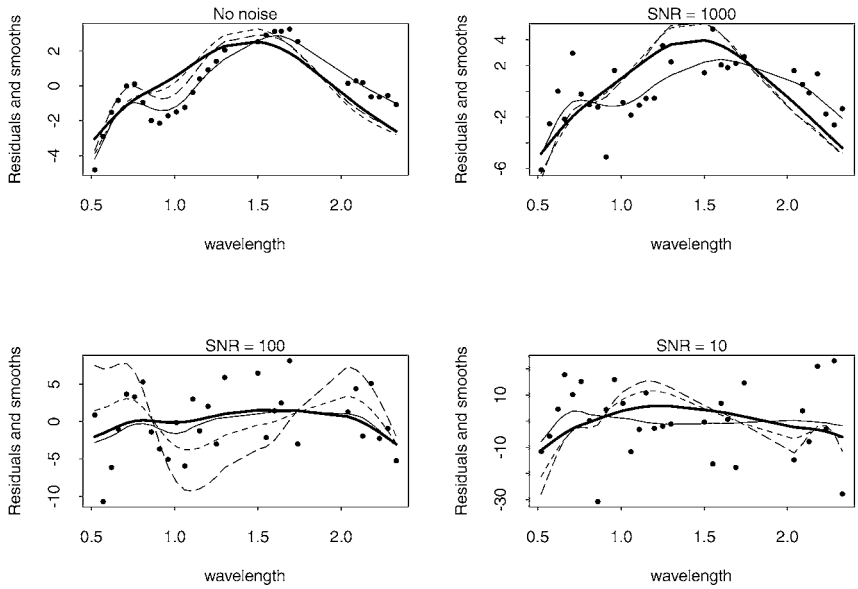

Figure 11. Examples of residuals from the ordinary least squares regression model. The thin line is a local linear regression of these residuals, dashed lines are different smooths that can be chosen by the back-fitting algorithm with the actually used smooth as the fat line.

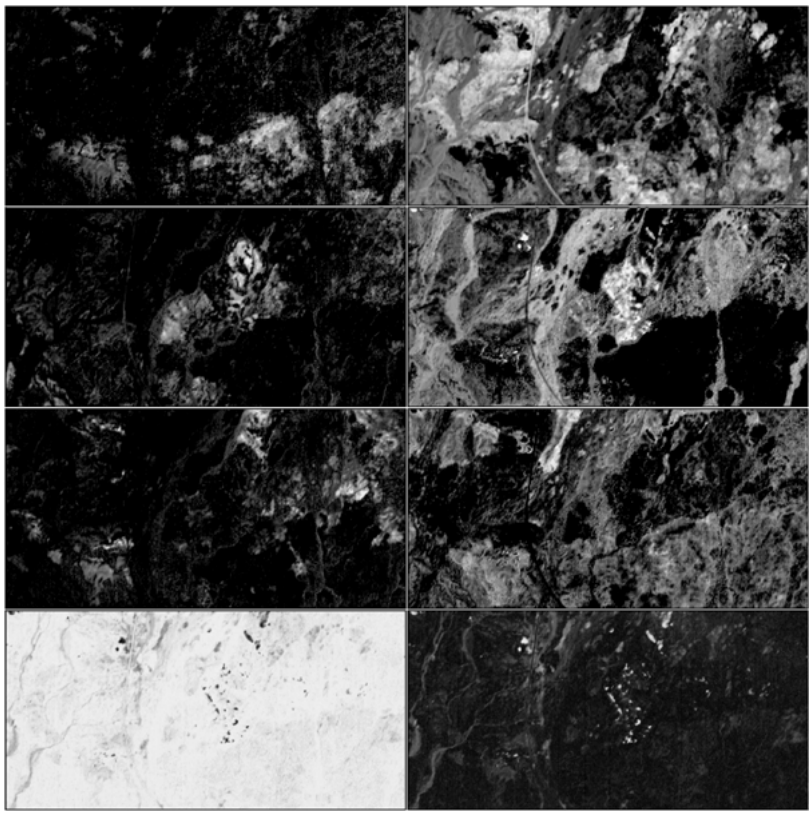

Figure 12. Results from a non-negative least squares simultaneous partial unmixing (i.e., the non-negative abundances add to a quantity not greater than one, Eq. (10)) of the six chosen end-member spectra. Row-wise: abundances of the six end-member spectra, R^2, and RMSE.

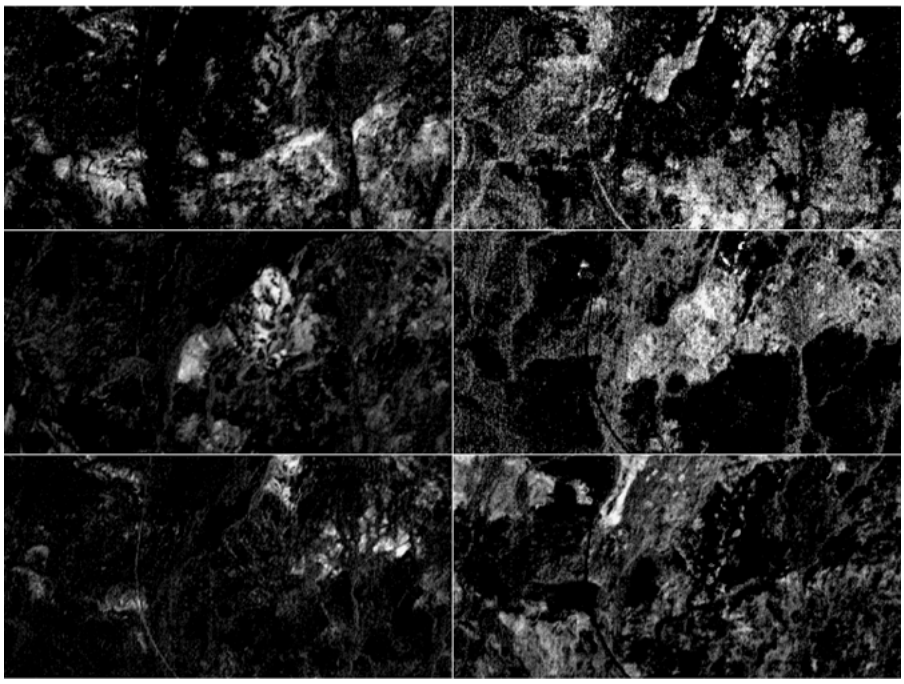

Figure 13. CEM abundance estimates stretched linearly from 0 to 1, Eq. (13).

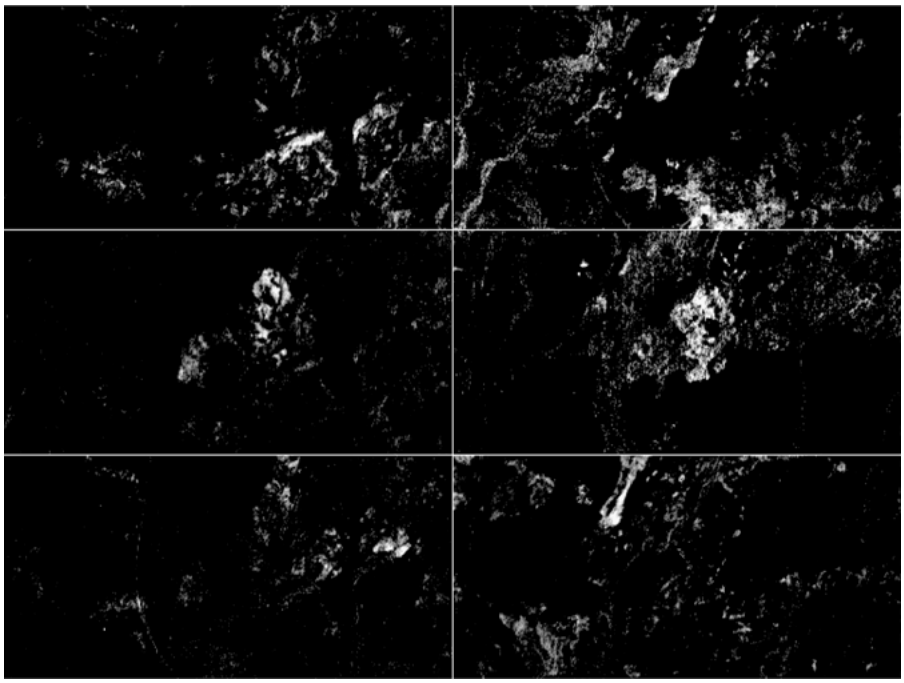

Figure 14. Iterated CEM abundance estimates stretched linearly from 0 to 1, Section 4.4.

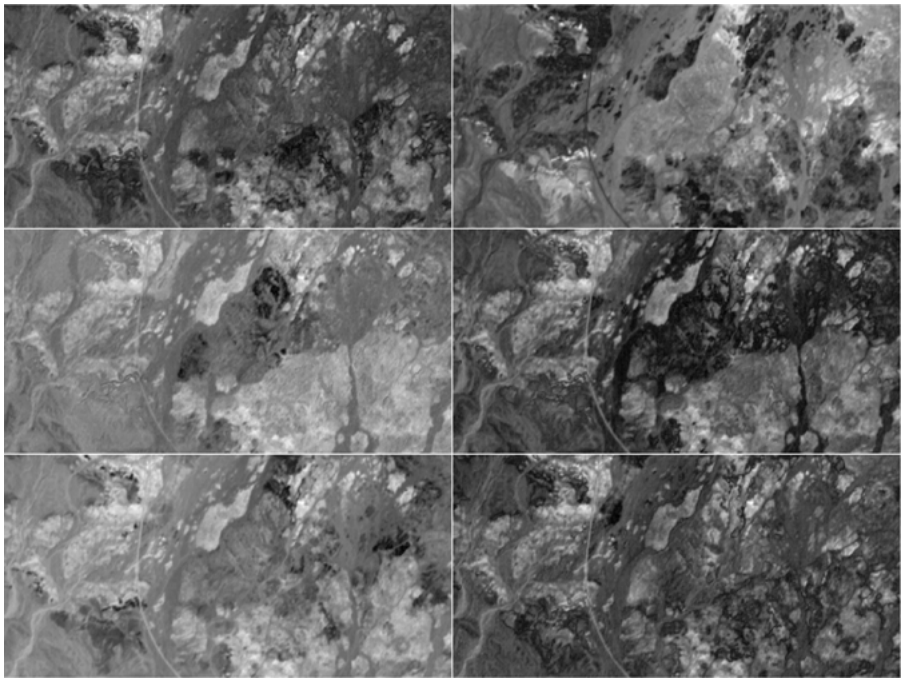

Figure 15. Spectral angles, Section 5.

Figure 15 shows the spectral angles from a SAM analysis (Section 5). Obviously, the spectral angle is small where the abundance is high.

With stretching from 0 to 1 the CEM abundance estimates give a more distinct and visually pleasing impression of the spatial distribution of spectrally similar material than does spectral angles. SAM seems to give information on what is very different from the desired spectrum whereas CEM seems to give information on what is very similar to the desired spectrum. Unlike CEM results there is no obvious stretching to be applied to SAM results.

Again, this is a "true remote sensing situation". With no knowledge of "the truth" it is difficult to make hard statements on the above results on the AVIRIS data. However, the partial simultaneous constrained unmixing based on non-linear least squares (Eq. (10) and Fig. 12), and the partial marginal unmixing based on (iterated) constrained energy minimization (Eq. (13), Section 4.4, and Figs. 13 and 14) all agree on the locations of the highest and to some extent also the lowest abundances. There are large differences between the three results for the intermediate abundances. Especially the iterated CEM method gives smaller target areas for the six end-member spectra used. This complies with the results from the example given in Section 8.1 on simple generated data where this characteristic is seen to be coupled with the cost of missing true positives.

9. Conclusions

CEM and a new eigenvalue formulation alternative enable us to perform partial unmixing when we know the desired end-member spectrum only and not the full set of end-member spectra which is a very realistic situation. This is an advantage over full unmixing and OSP. Based on the eigenvalue formulation a new iterative scheme for calculating CEM is suggested. Results from the first iteration shows promising results for both the case with generated data and for the case with AVIRIS data where it seems to suppress noise better than the ordinary CEM method at the cost of missing true positives. When applying the CEM method or the eigenvalue formulation alternative to MAF or MNF transformed data, matrix inversion is not needed and also the noise isolated in the MAF/MNFs excluded from the analysis does not influence the matched filtering performed. In a very simple 2-spectrum example TCIMF tends to suppress the undesired spectrum rather than put extra highlight on the desired spectrum.

Compared to unconstrained full unmixing by OLS the constrained full unmixing by NNLS in the same simple example better suppresses the undesired spectrum at the cost of more false positives. A study with generated AVIRIS spectra constructed as weighted sums of real spectra and different amounts of iid Gaussian noise shows (1) that ordinary least squares in this case with one unknown spectrum performs better than non-negative least squares, and (2) that unmixing based on the semi-parametric model with backfitting gives better but still not satisfactory estimates of endmember abundances than linear unmixing. Hopefully, future work on mean weighted least squares estimation in the semi-parametric model will improve on this.

Acknowledgments

Professor Knut Conradsen, Dr. Bjarne K. Ersbøll, Dr. Jens Michael Carstensen and Dr. Rasmus Larsen, all IMM, provide a strong environment for discussing and working with (multivariate) statistics and image analysis. Rasmus Larsen wrote the `maf` and `seed` programs in close cooperation with the author. Johan Doré, formerly IMM, wrote part of the `seed` program. Thanks to Dr. Henrik Aalborg Nielsen, IMM, for good discussions on additive and semi-parametric models. Thanks also to two anonymous reviewers whose comments helped improve the readability of the manuscript.

References

1. E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J.D. Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorenson, *LAPACK Users' Guide*, Society for Industrial and Applied Mathematics: Philadelphia, 2nd ed., 1995.
2. T.W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley: New York, 2nd ed., 1984.
3. AVIRIS, Jet Propulsion Laboratory, National Aeronautics and Space Administration, Pasadena, California, USA. Also available at http://makalu.jpl.nasa.gov/aviris.html.
4. A. Bernard, I. Kanellopoulos, and G.G. Wilkinson, "Training strategies for neural network soft classification of remotely sensed imagery," *International Journal of Remote Sensing*, Vol. 18, No. 8, pp. 1851–1856, 1997.
5. J.M. Chambers and T.J. Hastie (Eds.), *Statistical Models in S,* Wadsworth and Brooks/Cole, 1992, Pacific Grove, CA.
6. K. Conradsen, B.K. Ersbøll, and A.A. Nielsen, "Noise removal in multichannel image data by a parametric maximum noise fractions estimator," in *Proceedings of the 24th International Symposium on Remote Sensing of Environment*, Rio de Janeiro, Brazil, 1999, ERIM (Ed.), pp. 403–416.
7. K. Conradsen, B.K. Nielsen, and T. Thyrsted, "A comparison of min/max autocorrelation factor analysis and ordinary factor analysis," in *Proceedings from Symposium in Applied Statistics*, Lyngby, Denmark, 1985, pp. 47–56.
8. R.E. Crippen and R.G. Bloom, "Unveiling the lithology of vegetated terrains in remotely sensed imagery via forced decorrelation," in *Proceedings of the Thirteenth International Conference on Applied Geologic Remote Sensing*, Vancouver, British Columbia, Canada, ERIM (Ed.), 1999, Vol. 1, p. 150.
9. A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, Series B*, Vol. 39, pp. 1–38, 1977.
10. W.J. Dixon (Ed.), *BMDP Statistical Software*, University of California Press, 1985, p. 734.
11. J.J. Dongarra, J.R.Bunch, C.B. Moler, and G.W. Stewart, *LINPACK Users' Guide*, Society for Industrial and Applied Mathematics: Philadelphia, 1979.
12. B.K. Ersbøll, "Transformations and classifications of remotely sensed data: Theory and geological cases," Ph.D Thesis, Institute of Mathematical Statistics and Operations Research, Technical University of Denmark, Lyngby, 1989, P. 297.
13. H. Flesche, A.A. Nielsen, and R. Larsen, "Supervised mineral classification with semi-automatic training and validation set generation in scanning electron microscope energy dispersive spectroscopy images of thin sections," *Mathematical Geology*, Vol. 32, No. 3, pp. 337–366, 2000.
14. P.E. Gill, S.J. Hammarling, W. Murray, M.A. Saunders, and M.H. Wright, "User's guide for LSSOL (Version 1.0): A FORTRAN package for constrained linear least-squares and convex quadratic programming," Department of Operations Research, Stanford University, California, USA, Technical Report 86-1, 1986.
15. A.A. Green, M. Berman, P. Switzer, and M.D. Craig, "Transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 26, No. 1, pp. 65–74, 1994.
16. J.C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 32, No. 4, pp. 779–785, 1994.
17. T.J. Hastie and R.J. Tibshirani, *Generalized Additive Models*, Chapman and Hall, London, 1990.
18. A. Jacobsen, K.B. Heidebrecht, and A.A. Nielsen, "Monitoring grasslands using convex geometry and partial unmixing—A case study," in *Proceedings of the First EARSeL Workshop on Imaging Spectroscopy*, Zürich, Switzerland, 1998, M. Schaepman, D. Schläpfer, and K. Itten (Eds.). pp. 309–316.
19. A.S. Jacobson, A.L. Berkin, and M.N. Orton, "Linkwinds: Interactive scientific data analysis and visualization," *Communications of the ACM*, Vol. 37, No. 4, pp. 42–52, 1994. http://linkwinds.jpl.nasa.gov/.
20. J.T. Kent and K.V. Mardia, "Spatial classification using fuzzy membership models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 5, pp. 659–671, 1988.
21. F.A. Kruse, A.B. Lefkoff, J.B. Boardman, K.B. Heidebrecht, A.T. Shapiro, P.J. Barloon, and A.F.H. Goetz, "The spectral image processing system (SIPS)—Interactive visualization and analysis of imaging spectrometer data," *Remote Sensing of Environment*, Vol. 44, pp. 145–163, 1993.

22. M.S. Landy, "HIPS-2 software for image processing: Goals and directions," in *Proceedings of the SPIE 1964 Applications of Artificial Intelligence 1993: Machine Vision and Robotics*, K.L. Boyer and L. Stark (Eds.), pp. 382–391. http://www.cns.nyu.edu/ home/msl/hipsdescr.cgi/.
23. M.S. Landy, Y. Cohen, and G. Sperling, "HIPS: A UNIX based image processing system," *Computer Vision, Graphics and Image Processing*, Vol. 25, No. 3, pp. 331–347, 1984. http://www.cns.nyu.edu/ home/msl/hipsdescr.cgi/.
24. M.S. Landy, Y. Cohen, and G. Sperling, "HIPS: Image processing under UNIX. Software and applications," *Behavior Research Methods, Instrumentation, and Computers*, Vol. 16, No. 2, pp. 199–216, 1984. http://www.cns.nyu.edu/ home/msl/hipsdescr.cgi/.
25. R. Larsen, A.A. Nielsen, and K. Conradsen, "Restoration of hyperspectral push-broom scanner data," in *Proceedings of the 17th EARSeL Symposium on Future Trends in Remote Sensing*, Lyngby, Denmark, 1997, P. Gudmandsen (Ed.), pp. 157–162.
26. R. Larsen, A.A. Nielsen, and H. Flesche, "Sensitivity study of a semi-automatic supervised classifier applied to minerals from X-ray mapping images," in *Proceedings of the Scandinavian Image Analysis Conference (SCIA'99)*, Kangerlussuaq, Greenland, 1999, B.E. Ersbøll and P. Johansen (Eds.), Vol. 2, pp. 785–792.
27. R. Larsen, A.A. Nielsen, and H. Flesche, "Sensitivity study of a semi-automatic training set generator," *Pattern Recognition Letters*, Vol. 21, No. 13–14, pp. 1175–1182, 2000.
28. J.B. Lee, A.S. Woodyatt, and M. Berman, "Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 28, No. 3, pp. 295–304, 1990.
29. S.E. Marsh, P. Switzer, W.S. Kovalik, and R.J.P. Lyon, "Resolving the percentage of component terrains within single resolution elements," *Photogrammetric Engineering and Remote Sensing*, Vol. 46, pp. 1079–1086, 1980.
30. F. Maselli, "Multiclass spectral decomposition of remotely sensed scenes by selective pixel unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 36, No. 5, pp. 1809–1820, 1998.
31. J.W.V. Miller, J.B. Farison, and Y. Shin, "Spatially invariant image sequences," *IEEE Transactions on Image Processing*, Vol. 1, No. 2, pp. 148–161, 1992.
32. A.A. Nielsen, "Analysis of regularly and irregularly sampled spatial, multivariate, and multi-temporal data," Ph.D. Thesis, Department of Mathematical Modelling, Technical University of Denmark, Lyngby, 1994. http://www.imm.dtu.dk/∼aa/phd/.
33. A.A. Nielsen, "Linear mixture models and partial unmixing in multi- and hyperspectral image data," in *Proceedings from the 1st EARSeL Workshop on Imaging Spectroscopy*, Zürich, Switzerland, 1998, M. Schaepman, D. Schläpfer, and K. Itten (Eds.), pp. 165–172.
34. A.A. Nielsen, "Linear mixture models, full and partial unmixing in multi- and hyperspectral image data," in *Proceedings of the Scandinavian Image Analysis Conference (SCIA'99),* Kangerlussuaq, Greenland, 1999, B.E. Ersbøll and P. Johansen (Eds.), Vol. 2, pp. 898–902.
35. A.A. Nielsen, "Multi-channel remote sensing data and orthogonal transformations for change detection," in *Machine Vision and Advanced Image Processing in Remote Sensing*, I. Kanellopoulos, G.G. Wilkinson, and T. Moons (Eds.), Springer, 1999.
36. A.A. Nielsen, "Partial unmixing in hyperspectral image data," in *Proceedings from the Fourth International Airborne Remote Sensing Conference and Exhibition,* Ottawa, Ontario, Canada, 1999, ERIM (Ed.), Vol. II, pp. 535–542.
37. A.A. Nielsen, K. Conradsen, J.L. Pedersen, and A. Steenfelt, "Spatial factor analysis of stream sediment geochemistry data from South Greenland," in *Proceedings of the Third Annual Conference of the International Association for Mathematical Geology (IAMG'97)*, Barcelona, Spain, 1997, V. Pawlowsky-Glahn (Eds.), pp. 955–960.
38. A.A. Nielsen, K. Conradsen, J.L. Pedersen, and A. Steenfelt, "Maximum autocorrelation factorial kriging," in *Proceedings of the Sixth International Geostatistics Congress (Geostats 2000)*, Cape Town, South Africa, 2000, W.J. Kleingeld and D.K. Krige (Eds.), paper no. E13 on CD Rom.
39. A.A. Nielsen, K. Conradsen, and J.J. Simpson,"Multivariate alteration detection (MAD) and MAF post-processing in multispectral, bi-temporal image data: New approaches to change detection studies," *Remote Sensing of Environment*, Vol. 64, pp. 1–19, 1998.
40. A.A. Nielsen, B. Ersbøll, W. Pälchen, and G. Rank, "Spatial analysis of multivariate, (Ir-) regularly sampled data: Geochemistry from the eastern Erzgebirge," in *Geostatistics Wollongong '96*, Wollongong, Australia, 1997, E.Y. Baafi and N.A. Schofield (Eds), pp. 1173–1184, 1996.
41. A.A. Nielsen, H. Flesche, R. Larsen, J.M. Rykkje, and M. Ramm, "Semi-automatic supervised classification of minerals from X-ray mapping images," in *Proceedings of the Fourth Annual Conference of the International Association of Mathematical Geology (IAMG'98)*, Ischia, Italy, 1998, A. Buccianti, G. Nardi, and R. Potenza (Eds.), pp. 473–478.
42. A.A. Nielsen and R. Larsen, "Restoration of GERIS data using the maximum noise fractions transform," in *Proceedings from the First International Airborne Remote Sensing Conference and Exhibition,* Strasbourg, France, 1994, ERIM (Ed.), Vol. II, pp. 557–568.
43. N. Pendock and A.A. Nielsen, "Multispectral image enhancement neural networks and the maximum noise fraction transform," in *Proceedings of the Fourth South African Workshop on Pattern Recognition*, Simon's Town, South Africa, P. Cilliers (Ed.), pp. 2–13.
44. H. Ren and C.-I. Chang, "A target-constrained interference-minimized filter for subpixel detection in hyperspectral imagery," in *Proceedings of the IEEE 2000 IGARSS*, Honolulu, Hawaii, USA, 2000.
45. R.G. Resmini, M.E. Kappus, W.S. Aldrich, J.C. Harsanyi, and M. Anderson, "Mineral mapping with hyperspectral digital imagery collection experiment (HYDICE) sensor data at cuprite, Nevada, U.S.A.," *International Journal of Remote Sensing*, Vol. 18, No. 7, pp. 1553–1570, 1997.
46. P. Sadegh, H.A. Nielsen, and H. Madsen, "A semi-parametric approach for decomposition of absorption spectra in the presence of unknown components," Department of Mathematical Modelling, Technical University of Denmark, Technical Report 1999-17, http://www.imm.dtu.dk/documents/ftp/tr99/tr17_99.abstract.html.
47. J.J. Settle, "On the relationship between spectral unmixing and subspace projection," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 34, No. 4, pp. 1045–1046, 1996.

48. J.J. Settle and N.A. Drake, "Linear mixing and the estimation of ground cover proportions," *International Journal of Remote Sensing*, Vol. 14, No. 6, pp. 1159–1177, 1993.
49. S.S. Stan, "Mineral identification and mapping by imaging spectroscopy data analysis," Ph.D. Thesis, Computational and Applied Mathematics Department, University of the Witwatersrand, Johannesburg, South Africa, 1993.
50. P. Strobl, R. Richter, A. Müller, F. Lehmann, D. Oertel, S. Tischler, and A.A. Nielsen, "DAIS system performance, first results from the 1995 evaluation campaigns," in *Proceedings from the Second International Airborne Remote Sensing Conference and Exhibition,* San Francisco, California, USA, 1996, ERIM (Ed.), Vol. II, pp. 325–334.
51. P. Switzer and A.A. Green, "Min/max autocorrelation factors for multivariate spatial imagery," Department of Statistics, Stanford University, Technical Report 6, 1984.
52. T.-M. Tu, C.-H. Chen, and C.-I. Chang, "A noise subspace projection approach to target signature detection and extraction in an unknown background for hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 36, No. 1, pp. 171–181, 1998.
53. F. van der Meer, "Iterative spectral unmixing," *International Journal of Remote Sensing*, Vol. 20, No. 17, pp. 3431–3436, 1999.
54. G. Vane and A.F.H. Goetz, "Terrestrial imaging spectroscopy," *Remote Sensing of Environment*, Vol. 24, pp. 1–29, 1988.
55. G. Vane, R.O. Green, T.G. Chrien, H.T. Enmark, E.G. Hansen, and W.M. Porter, "The airborne/infrared imaging spectrometer (AVIRIS)," *Remote Sensing of Environment*, Vol. 44, pp. 127–143, 1993.
56. W.N. Venables and B.D. Ripley, *Modern Applied Statistics with S-PLUS*, Springer, New York, 3rd ed., 1999.

Allan Aasbjerg Nielsen is an Associate Professor with the Section for Image Analysis, Informatics and Mathematical Modelling, Technical University of Denmark. M.Sc. in 1978 from the Department of Electrophysics, Technical University of Denmark. Ph.D. in 1994 from Informatics and Mathematical Modelling, Technical University of Denmark. With the Danish Defense Research Establishment 1977-1978. Work on energy conservation in housing with the Thermal Insulation Laboratory, Technical University of Denmark, from 1978 to 1985. Since 1985 with the Section for Image Analysis, Informatics and Mathematical Modelling, Technical University of Denmark. Since then work on several national and international projects on application of statistical methods and remote sensing in mineral exploration, mapping, geology, environment, geodesy and oceanography agriculture, funded by industry, the European Union, Danida (the Danish International Development Agency), and the Danish Research Councils. National point of contact (NPOC) for concerted actions on neurocomputing and advanced methods in remote sensing funded by the European Union. Member of host, program and scientific committees for international conferences. Referee for international journals and conferences. Homepage www.imm.dtu.dk/∼aa.