Linear Mixture Modelling Solution Methods for Satellite Remote Sensing

College of Aeronautics Report 9603

ISBN 1 871564 83 2

Stephen Hobbs
College of Aeronautics, Cranfield University
Cranfield, Bedford MK43 0AL, UK
s.e.hobbs@cranfield.ac.uk

27 March 1996, revised 4 Nov 2003

Abstract

This report documents the algorithms used in the program *MIXMOD* to analyse mixed pixel data (assuming linear mixing). The report describes the mathematical algorithms rather than acting as a user's manual for *MIXMOD*. The algorithms described obtain the desired solutions, quantify the quality of the solution, and estimate error bounds, using a variety of methods. A novel feature is the ability to handle uncertainty in the assumed end-member spectra, which in practice may be the dominant source of error. The report includes a brief literature review to place the work in its broader context (with references to the mathematics of linear systems and other applications of mixture modelling in remote sensing).

Contents

Chapter 1

Introduction

The principal aim of this report is to document the algorithms used in the program *MIXMOD* for the analysis of "mixed" pixel data. Subsidiary aims are to summarise the principles on which linear mixture modelling are based and to provide a brief overview of the literature relevant to the subject.

Linear mixture modelling refers to cases in remote sensing in which the signal measured is, for practical purposes, a linear mixture of several different "fundamental" signals (usually called "end-members" in the literature). The goal is often to estimate the relative contributions of the different "end-member" components given only measurements of the composite (or "mixed") signals for a scene.

The report lists references to applications and discussions of the methods specifically in the area of satellite remote sensing. References are also given to the much larger body of literature concerned with general linear systems and to some related applications of the techniques.

Program *MIXMOD* which is used at Cranfield University for the analysis of mixed pixel data uses several different solution algorithms; it also evaluates parameters to quantify the goodness-of-fit of the solution, and calculates appropriate error bounds. All the methods currently used for these functions are documented in this report, which therefore serves as both a reference for *MIXMOD* and also as a (partial) review of the mathematical methods appropriate to linear mixture modelling.

1.1 Linear mixture modelling

The general case being considered is that in which the outputs of some linear combination process are measurable, but in which the parameters of interest are actually the inputs. In satellite remote sensing, this occurs when the elemental area of the Earth's surface being imaged (a pixel) contains several different fundamental cover components (e.g. water, bare soil, grass, forest) and so the measured spectrum is a composite of the spectra from the different individual components. If the sensor's resolution limit is greater than the pixel size then neighbouring pixels are not independent measurements of surface properties. In many cases of interest, the composite spectrum is a linear combination of the spectra of the individual components, and it is the areas covered by the

individual components which are required.

The relationship between the measured signal in band i (y_i), the value of the reference signal in *band* i for a pixel wholly filled by *component* j ($[A]_{ij} = a_{ij}$), and the proportion of the pixel actually occupied by cover component j (x_j), is given by:

$$\mathbf{y} = A\mathbf{x} \tag{1.1}$$

In this case, the "inputs" are the class proportions and the "outputs" are the signals in each band. In general, the number of bands does not equal the number of cover classes, and the matrix equation cannot be inverted directly since A is not square. There are usually more bands than classes, and a conventional least squares matrix solution method is used (which may or may not account for uncertainty in the measurements). An important feature of the various least-squares methods is their ability to cope with a degree of noise in the system. Conventionally, this noise is assumed to lie solely in the signal measurements made in each spectral band. Methods reported in the literature do not yet include the ability to deal with noise in the end-member reference spectra which is likely to be the dominant source of noise in most applications. Appropriate solution methods are presented in sections 5 and 6 of this report, and are also discussed in [7].

Practical methods of estimating the pixel proportions take many forms, and often make use of additional information to try to improve the solution estimate. The sophistication of solution methods also varies widely, from relatively straightforward inversion methods for a fixed number of bands (e.g. [25]) to much more general algorithms able to incorporate a wide range of additional information (e.g. [24]).

1.2 Related analysis techniques and applications

Many fields of study use linear models to relate inputs and outputs, and the mathematics of such systems is a highly developed subject. [15] discuss the mathematics of solution methods for linear systems in good detail, and refer to related topics such as fitting models to data, and regularised solutions (see below).

In remote sensing, a mature field in terms of the algorithms used is atmospheric profiling. Rodgers [19, 9, 20] authoritatively reviews the inversion methods used in this field, covering linear methods as well as more sophisticated techniques.

1.3 Structure of the report

Following this Introduction, Section 2 gives an overview of literature relevant to linear mixture modelling. This is not intended to be a comprehensive review, but does provide pointers to a wide range of work, covering not only applications of "conventional" mixture modelling, but also a range of related fields, and some relevant mathematical texts discussing the algorithms used. This overview places mixture modelling in its broader context.

Section 3 provides a brief introduction to the different solution methods available for linear mixture modelling and shows how they relate to each other. The core of the material presented in the report describes the algorithms implemented in the program *MIXMOD* (Sections 4 to 6). This gives a sound background understanding for users of *MIXMOD*. The descriptions are however quite general and allow the reader to develop his or her own applications as necessary as well as illustrating the methods discussed earlier in the report. Section 7 discusses briefly some extensions of the standard methods.

Chapter 2

Literature Overview

The references listed with this report are a sample of the literature relevant to the subject of linear mixture modelling. The articles generally deal with either the mathematical principles and relevant algorithms, specific applications, or with comparative studies for several different approaches to the same problem. The following notes are included to guide the reader to appropriate articles for specific topics.

2.1 Mathematical Methods

[15] give an excellent guide to the use of mathematical methods for data analysis. Topics covered include linear systems (Chapter 2), fitting models to data (Chapter 15), quantifying solution quality and errors (sections 15.1 and 15.6 of Chapter 15), and regularising solutions (sections 18.4ff of Chapter 18). Papers covering the application of these methods to remote sensing include [24] and the work of Rodgers [19, 9, 20].

2.2 Applications

One of the most active groups using mixture modelling techniques is that based at the University of Washington (Seattle, USA) [26, 1, 18, 27, 21, 25, 22, 28]. Their work, principally for land surfaces includes hyperspectral geological applications, and has spawned projects at related institutions [8, 14]. Monte Carlo methods have been used to try to analyse the ability of measurements to discriminate surface types and to evaluate error bounds (e.g. [22]), but the approach is largely empirical, without always justifying underlying assumptions (e.g. the use of shade as an end-member, or the application of the method to water colour due to sediment loading).

Perhaps the most comprehensive work, mathematically, (although it does not deal with end-member uncertainty) is that of [24]. Topics such as the regularisation of solutions are discussed. Non-mathematical readers may find the work daunting.

Various case studies have been reported [3, 5, 16, 17, 2, 12, 30, 29]. The work of [31] makes an interesting link between mixture modelling and data fusion.

Outside the area of surface remote sensing, Rodgers [19, 9, 20] gives authoritative reviews of atmospheric profiling, including the use of linear mixing models.

2.3 Comparative Studies

Several studies have been carried out comparing mixture modelling with alternative techniques such as artificial neural networks and fuzzy logic [4, 6, 2]. Neural networks, when properly trained, perform well, but do not provide clear measures of solution quality or likely error bounds.

Chapter 3

Solution Methods Overview

3.1 The assumption of linearity

The *linear* mixing model is assumed in most work related to the problem of estimating the proportions of different cover components occupying a single pixel from the spectral signature of that pixel. This model is equivalent to assuming that radiation scattered from the pixel is scattered by one component only, i.e. photons only interact with one cover type. In general this is often a good assumption and leads to a number of applications.

Note however, that several developments of this method need careful thought:

1. The signal must be *linearly* related to the component proportions. This usually requires it to be linearly related to the power received by the detector. Quantities such as radiance, DN (usually) and reflectance are suitable parameters, but in general, quantities such as the radar backscatter intensity (σ^0) *in dB* and the brightness temperature (T_B) are not. For small dynamic ranges quantities which are strictly non-linear may be close enough to linear to be dealt with by linear mixture modelling.

2. Signals produced by *coherent scattering* (e.g. radar, lidar) do not in general satisfy the required linear dependence on component proportion.

3. *Shade*: shade is inherently a product of multiple scattering and is therefore unlikely in general to be measurable by this technique.

4. *Variable strength*: the linear method can solve for either variable areas of fixed spectra (the usual land-cover application), or variable strengths of given spectra for fixed areas. The two effects cannot however be separated.

3.2 Constraining the solution

Information in addition to the measured signals in each band is often used to improve the estimated solution. This information may concern only individual pixels or it may be applied to groups of pixels so that neighbouring pixels help determine the values estimated for any individual pixel. Constraints are usually applied to force the solution to be more "realistic" in some sense.

It should be noted that constraints can be applied in different ways. Constraints can be applied so that they are satisfied exactly. For example, the equations can be rewritten to ensure that the sum of components is exactly equal to 1 (Sections 4, 5; the number of classes to solve for is then reduced by 1). Alternatively, the constraint can be applied on the same basis as the agreement between the data and modelled signals is required, i.e. small departures from exactly satisifying the constraint are allowed (an example of this is referred to below as the Lagrange multiplier method).

3.2.1 Constraints applied to pixels individually

There are two principal constraints which are applied to individual pixels:

1. No proportion value should be less than 0.
2. The sum of proportions should be equal to 1.

The first of these is usually applied iteratively (as is an option in *MIXMOD*) while the second can be applied either exactly (by modifying the basic linear mixing model equations) or approximately (treating the constraint on the same basis as the data). A third constraint, that no proportion value should be greater than 1, is implicit if both the above constraints are applied, and is rarely applied explicitly.

3.2.2 Constraints applied to groups of pixels

Constraints which apply to groups of pixels, usually groups of contiguous pixels, are described by the general term "regularisation", and are discussed by [24] and [15]. Regularisation methods allow *a priori* information concerning spatial properties of the solution (e.g. the size of inter-pixel variations, or the (presumably low) likelihood of large departures from an area mean) to be incorporated into the solution.

3.3 Testing solution quality

Any serious attempt to solve for the pixel proportions must address the issue of checking the quality of the solution obtained. This is usually done by comparing the discrepancies between the modelled and measured signals in each band with the level of discrepancy which can be explained using known uncertainties in the measurement system. Several methods of making the comparisons are available. Residuals (the actual differences between the measured and modelled signals) are used by some authors (e.g. [22]), while others use quasi or fully quantitative statistical measures based on the conventional χ^2 goodness-of-fit parameter (see Section 6 of this report). However this task is done, it is a crucial step.

3.4 Solution error bounds

If a solution is of acceptable quality, the next step to make the solution estimate usable is to estimate the level of uncertainty of the solution. If there is no knowledge of the uncertainty of the solution then the solution is useless in

Constraint(s) applied		No end-member uncertainty	End-member uncertainty treatment method:	
			Approximate	Exact
None		OK	OK	OK
$\sum x_j = 1$	(a) $x_{nc} = 1 - \sum_{j=1}^{nc-1} x_j$	OK	OK	OK
	(b) Lagrange multipliers	ok	ok	ok
$x_j \geq 0$		OK	OK	OK
$\sum x_j = 1^{(a)}$ & $x_j \geq 0$		OK	OK	OK

Table 3.1: Different solution methods for linear mixture modelling. ('OK' indicates that this case has been implemented fully in program *MIXMOD* (current version = 0.12), and 'ok' that a partial implementation is available, i.e. excluding dedicated error analysis.) The various solution methods referred to are explained in the text.

practice. The main quantitative approaches to estimating error bounds are (1) a Monte Carlo method, and (2) an analytical method. The Monte Carlo method simulates several different sets of measurements with different sets of "random" noise added, and then looks at the spread of solutions obtained. The analytical method uses the curvature of the cost function minimum at the solution to estimate error bounds and covariances. Both methods assume that the noise distributions are known. The Monte Carlo method can deal more easily with general noise models while the analytical method is more effective at revealing the factors which make the measurements susceptible to noise.

3.5 The *MIXMOD* program

Table 3.1 desribes the solution methods implemented in program *MIXMOD* (nc = number of classes). The current version of the program runs under the MS-DOS operating system, and is written in C to be relatively portable.

MIXMOD has been applied to a range of linear mixing problems ranging from the standard mixed pixel analysis [29] to atmospheric composition profiling and the analysis of fluorescence spectra, illustrating the common ground between these different application areas.

Chapter 4

Solution Methods Ignoring End-Member Noise

Most of the methods reported so far in the literature completely ignore the contribution of end-member noise. Section 5 and 6 of this report include methods of handling end-member uncertainty. This section describes a variety of methods for dealing with the simpler case which ignores end-member uncertainty. A range of standard methods is available to solve for the pixel proportions.

All the methods reported here are based on a goodness-of-fit measure G. G is a sum of the squared differences between the measured (y_i) and modelled ($\hat{y}_i$) signals in each band (label i), normalised by a quantity σ_i.

$$G = \sum_{i=1}^{nb} \frac{(y_i - \hat{y}_i)^2}{\sigma_i^2} \tag{4.1}$$

$$\hat{y}_i = \sum_{j=1}^{nc} a_{ij} x_j \tag{4.2}$$

where a_{ij} is the array of end-member reference spectra, and x_j is the proportion of component j contributing to the (modelled) signal.

Depending on the definition of σ_i, different types of solution (including those used in most published work) are obtained.

1. σ_i is constant, independent of signal band (i.e. all bands are weighted equally): the method repeats the standard least squares solution.

2. σ_i varies with signal band, usually according to the level of noise present with the signal, but only the *relative* weights are controlled: gives a weighted least squares solution but the goodness-of-fit parameter is not quantitative.

3. σ_i is set equal to the expected noise standard deviation in each band, and the noise is close to Gaussian: a weighted least-squares solution is obtained together with a quantitative measure of the goodness-of-fit.

Note that in certain cases, e.g. noise level set correctly and the same in all bands, the different cases may be equivalent.

4.1 No constraint on the sum of components

It is not necessary to apply any constraint on the sum of components, although this is not usually the approach taken in published work.

A solution is obtained by requiring the derivative of the goodness-of-fit parameter G to be zero with respect to each of the components x_j.

$$0 = \sum_{i=1}^{nb} \frac{(y_i - \hat{y}_i)}{\sigma_i^2} \frac{\partial \hat{y}_i}{x_j} \tag{4.3}$$

$$= \sum_{i=1}^{nb} \frac{(y_i - \sum_{k=1}^{nc} a_{ik} x_k)}{\sigma_i^2} a_{ij} \tag{4.4}$$

$$\sum_{i=1}^{nb} \frac{1}{\sigma_i^2} y_i a_{ij} = \sum_{k=1}^{nc} \sum_{i=1}^{nb} \frac{1}{\sigma_i^2} a_{ij} a_{ik} x_k \tag{4.5}$$

$$B_j = \sum_{k=1}^{nc} C_{jk} x_k \tag{4.6}$$

where

$$C_{jk} = \sum_{i=1}^{nb} \frac{1}{\sigma_i^2} a_{ij} a_{ik} \tag{4.7}$$

$$B_j = \sum_{i=1}^{nb} \frac{1}{\sigma_i^2} y_i a_{ij} \tag{4.8}$$

The solution is obtained by matrix inversion:

$$\mathbf{x} = C^{-1}\mathbf{B} \tag{4.9}$$

Equations written in matrix form

Using conventional matrix notation, the equations may be written as

$$G = (\mathbf{y} - \hat{\mathbf{y}})^T W (\mathbf{y} - \hat{\mathbf{y}}) \tag{4.10}$$

$$\hat{\mathbf{y}} = A\mathbf{x} \tag{4.11}$$

W is the matrix of weightings (a diagonal matrix if all the bands are independent, with $w_{ii} = \frac{1}{\sigma_i^2}, w_{ij} = 0\ (i \neq j)$), and A is the matrix of end-member reference spectra. The solution is then

$$\mathbf{x} = (A^T W A)^{-1} A^T W \mathbf{y} \tag{4.12}$$

4.2 Sum of components equal to 1

This is a natural constraint to apply since in practice the sum of components should be exactly one (if the end-members are correctly chosen) and the problem can be reduced in scope by one dimension using this constraint.

4.2.1 Forced exact compliance: $x_{nc} = 1 - \sum_{j=1}^{nc-1} x_j$

The simplest way to apply this constraint is to express one component, typically the last component, in terms of all the other components. A disadvantage is that the method loses the explicit symmetry between all the components.

A solution may be obtained by requiring the first derivative of the goodness-of-fit parameter to be zero. This gives a matrix equation which can be solved by matrix inversion methods (e.g. Singular Value Decomposition (SVD) or Gauss-Jordan [15] - SVD methods are generally preferred because they can be made more robust and also make the sensitivity of the solution to noise more explicit).

$$\sum_{k=1}^{nc-1} \sum_{i=1}^{nb} \frac{1}{\sigma_i^2}(a_{ij} - a_{inc})(a_{ik} - a_{inc})x_k = \sum_{i=1}^{nb} \frac{1}{\sigma_i^2}(a_{ij} - a_{inc})(y_i - a_{inc}) \quad (4.13)$$

$$\sum_{k=1}^{nc-1} C_{jk}x_k = B_j \quad (4.14)$$

where

$$C_{jk} = \sum_{i=1}^{nb} \frac{1}{\sigma_i^2}(a_{ij} - a_{inc})(a_{ik} - a_{inc}) \quad (4.15)$$

$$B_j = \sum_{i=1}^{nb} \frac{1}{\sigma_i^2}(a_{ij} - a_{inc})(y_i - a_{inc}) \quad (4.16)$$

These equations may be solved for the $nc - 1$ component fractions x_j ($j = 1, 2, \ldots nc - 1$). Note that the value of the final component then has to be obtained from $x_{nc} = 1 - \sum_{j=1}^{nc-1} x_j$.

$$\mathbf{x} = C^{-1}\mathbf{B} \quad (4.17)$$

4.2.2 Lagrange Multiplier Method

Lagrange multiplier methods can incorporate the contraint on the sum of the components on the same basis as requiring that the signal in a particular band is well-modelled. An extra "band" row is provided for the end-member data to generate the sum of the component proportions. The corresponding "data" value is 1, added to the signal input. The band measurement noise chosen acts as the Lagrange multiplier and sets the scale of permitted deviation of the sum of the components from 1.

Lagrange multiplier methods appear to be appropriate to be based on methods with no other constraints applied to the sum of components (or else double constraints are being applied for the same feature of the problem).

4.3 All components positive

The current method is to check the first solution for any components which are negative. Any that are found are set to zero, and the problem is re-run using

a restricted set of components. The process continues until no components are negative.

A development of this constraint for the case without the sum of components required to equal 1 is to test the solution for component values set above 1.

4.4 Multiple constraints

The constraints on the sum of components and their permitted range can be applied independently.

The constraint on components not exceeding 1 could be applied independently, although once one component is set to 1 all the others must be zero (*if the constraint on the sum of components is also active*).

Chapter 5

Solution Methods Allowing for End-Member Noise

Uncertainty in the end-member measurements is probably the most important source of error in the solution estimates. Such uncertainty is not easily handled by the conventional methods reported in Section 4, and so two different methods of handling end-member uncertainty have been developed and are presented below. Both methods gave identical results in numerical simulations carried out to test them. This more general problem is not yet dealt with in the literature.

5.1 Approximate Methods

An iterative method based on the case with no end-member noise is possible, and is identical except for the evaluation of the expected noise standard deviation in each signal band. Iterations repeat until the solution values for successive iterations differ by less than some preset threshold. See section 2.5 of [15] for more discussion of iterative methods for improving solutions to sets of equations.

The sections above describe how the different constraints may be applied.

5.2 Exact Methods

The general problem requiring solution is to minimise G (the goodness of fit parameter) which is the sum of the squared normalised errors over the number of bands (nb). This is a non-linear minimisation problem, discussed in Chapter 15 of [15].

$$G = \sum_{i=1}^{nb} \frac{(y_i - \hat{y}_i)^2}{\sigma_i^2} \tag{5.1}$$

where the modelled signal ($\hat{y}_i$) and variance (σ_i^2) in band i are given by

$$\hat{y}_i = \sum_{j=1}^{nc} a_{ij} x_j \tag{5.2}$$

$$\sigma_i^2 = \sum_{k=1}^{nc}\sum_{l=1}^{nc} x_k x_l \overline{\delta a_{ik}\delta a_{il}} + 2\sum_{k=1}^{nc} x_k \overline{n_i \delta a_{ik}} + \overline{n_i n_i} \tag{5.3}$$

In general the cross-product $\overline{n_i \delta a_{ik}}$ may be ignored.

If the constraint on the sum of the components being equal to 1 is applied by writing the last class component as the complement of the other components, then the required solution (x_j) is a vector in $nc-1$ dimensions, and the modelled signal and variance may be written

$$\hat{y}_i = a_{inc} + \sum_{j=1}^{nc-1} x_j (a_{ij} - a_{inc}) \tag{5.4}$$

$$\begin{aligned}\sigma_i^2 = {} & \overline{n_i^2} + 2\sum_{k=1}^{nc-1} x_k \left[e_{ik,inc} - e_{inc,inc}\right] + e_{inc,inc} \\ & + \sum_{k=1}^{nc-1}\sum_{l=1}^{nc-1} x_k x_l \left[e_{ik,il} - e_{inc,il} - e_{ik,inc} + e_{inc,inc}\right]\end{aligned} \tag{5.5}$$

where the end-member array statistics are written as

$$e_{ik,jl} = \overline{\delta a_{ik}\delta a_{jl}} \tag{5.6}$$

Note that $e_{ik,jl} = e_{jl,ik}$.

The standard method of solution in multiple dimensions requires the ability to evaluate the function and its n-dimensional derivative.

$$\frac{\partial G}{\partial x_j} = -2\sum_{i=1}^{nb} \frac{(y_i - \hat{y}_i)}{\sigma_i^2}\frac{\partial \hat{y}_i}{\partial x_j} - \sum_{i=1}^{nb} \frac{(y_i - \hat{y}_i)^2}{\sigma_i^4}\frac{\partial \sigma_i^2}{\partial x_j} \tag{5.7}$$

To evaluate the effect of measurement errors the second derivative may also be required:

$$\begin{aligned}\frac{\partial^2 G}{\partial x_k \partial x_j} = {} & 2\sum_{i=1}^{nb} \frac{1}{\sigma_i^2}\frac{\partial \hat{y}_i}{\partial x_k}\frac{\partial \hat{y}_i}{\partial x_j} - \sum_{i=1}^{nb} \frac{(y_i - \hat{y}_i)^2}{\sigma_i^4}\left[\frac{\partial^2 \sigma_i^2}{\partial x_k \partial x_j} - \frac{2}{\sigma_i^2}\frac{\partial \sigma_i^2}{\partial x_j}\frac{\partial \sigma_i^2}{\partial x_k}\right] \\ & -2\sum_{i=1}^{nb} \frac{(y_i - \hat{y}_i)}{\sigma_i^2}\left(\frac{\partial^2 \hat{y}_i}{\partial x_k \partial x_j} - \frac{1}{\sigma_i^2}\frac{\partial \hat{y}_i}{\partial x_j}\frac{\partial \sigma_i^2}{\partial x_k} - \frac{1}{\sigma_i^2}\frac{\partial \sigma_i^2}{\partial x_j}\frac{\partial \hat{y}_i}{\partial x_k}\right)\end{aligned} \tag{5.8}$$

$$E\left(\frac{\partial^2 G}{\partial x_k \partial x_j}\right) \simeq 2\sum_{i=1}^{nb} \frac{1}{\sigma_i^2}\frac{\partial \hat{y}_i}{\partial x_k}\frac{\partial \hat{y}_i}{\partial x_j} - \sum_{i=1}^{nb} \frac{(y_i - \hat{y}_i)^2}{\sigma_i^4}\left[\frac{\partial^2 \sigma_i^2}{\partial x_k \partial x_j} - \frac{2}{\sigma_i^2}\frac{\partial \sigma_i^2}{\partial x_j}\frac{\partial \sigma_i^2}{\partial x_k}\right] \tag{5.9}$$

This expression is *approximately* correct in the sense that terms in odd powers of the difference between the measured and modelled signals have been ignored. (If the errors are random, which they should be for a good model, then this is a good approximation for the purposes of evaluating typical error bounds.)

The even powers of the difference between the measured and modelled signals are related to the expected signal variance.

$$E\left(\sum_{i=1}^{nb}\frac{(y_i-\hat{y}_i)^2}{\sigma_i^2}\right) = Number\ of\ degrees\ of\ freedom \quad (5.10)$$

$$= nb+n0-nc \quad (5.11)$$

(Writing $n0$ for the number of constraints applied.) Contributions from different bands are independent, therefore

$$E\left(\frac{(y_i-\hat{y}_i)^2}{\sigma_i^2}\right) = \frac{nb+n0-nc}{nb} \quad (5.12)$$

The expression for the expected value of the second derivative of G is therefore

$$E\left(\frac{\partial^2 G}{\partial x_k \partial x_j}\right) = 2\sum_{i=1}^{nb}\frac{1}{\sigma_i^2}\frac{\partial\hat{y}_i}{\partial x_k}\frac{\partial\hat{y}_i}{\partial x_j} - \frac{2(nb+n0-nc)}{nb}\sum_{i=1}^{nb}\frac{1}{\sigma_i^2}\left[\frac{1}{2}\frac{\partial^2\sigma_i^2}{\partial x_k\partial x_j}-\frac{1}{\sigma_i^2}\frac{\partial\sigma_i^2}{\partial x_j}\frac{\partial\sigma_i^2}{\partial x_k}\right] \quad (5.13)$$

These expressions apply whether or not constraints are applied to the sum of the components. The constraints affect the expressions used to evaluate the derivatives of $\hat{y}_i$ and σ_i^2.

5.2.1 No constraint on sum of components

In this case, the derivatives are given by:

$$\frac{\partial\hat{y}_i}{\partial x_j} = a_{ij} \quad (5.14)$$

$$\frac{\partial\sigma_i^2}{\partial x_j} = 2\sum_{k=1}^{nc} x_k e_{ij,ik} \quad (5.15)$$

$$\frac{\partial^2\sigma_i^2}{\partial x_k\partial x_j} = 2e_{ik,ij} \quad (5.16)$$

5.2.2 Sum of components equal to 1

The required derivatives are more complicated if written out in full:

$$\frac{\partial\hat{y}_i}{\partial x_j} = a_{ij}-a_{inc} \quad (5.17)$$

$$\frac{\partial\sigma_i^2}{\partial x_j} = 2\left[e_{ij,inc}-e_{inc,inc}\right] + 2\sum_{k=1}^{nc-1} x_k\left[e_{ik,ij}-e_{inc,ij}-e_{ik,inc}+e_{inc,inc}\right] \quad (5.18)$$

$$\frac{\partial^2\sigma_i^2}{\partial x_k\partial x_j} = 2\left[e_{ik,ij}-e_{inc,ij}-e_{ik,inc}+e_{inc,inc}\right] \quad (5.19)$$

An alternative may be to use the expressions for the derivatives in section 5.2.1 and then use a second stage of the partial derivatives to express the dependence of the final component on the remaining $nc - 1$ components.

5.2.3 All components positive

The constraints on the permitted range of the components is implemented the same way as for the other solution methods.

Chapter 6

Estimating Error Bounds

(This discussion is based closely on Chapter 15 of [15], to which reference should be made for full details.)

The variation of χ^2 with the parameter values can be used to estimate error bounds on the parameters (if assumptions about the model fit etc. are satisfied).

$$\Delta\chi^2 = \delta\mathbf{x}\,[\alpha]\,\delta\mathbf{x} \tag{6.1}$$

where, identifying G with χ^2,

$$\begin{aligned} \alpha &= \frac{1}{2}\frac{\partial^2\chi^2}{\partial x_j \partial x_k} \\ &= \frac{1}{2}\frac{\partial^2 G}{\partial x_j \partial x_k} \end{aligned} \tag{6.2}$$

The matrix $[\alpha]$ thus needs to be evaluated at the solution point to estimate the expected error bounds. Different definitions apply according to the constraints applied.

$[\alpha]$ needs to be inverted to obtain the covariance matrix, which is then projected onto the sub-space of interest for the error bounds before being reinverted to obtain the coefficients describing the error bound ellipsoid.

The chosen confidence level and number of degrees of freedom for the solution determine the actual value used for $\Delta\chi^2$ (the permitted increase in χ^2 relative to its value at the minimum corresponding to the solution point) to evaluate error bounds in a particular case.

Chapter 7

Developments of the Standard Methods

It is possible to obtain greater accuracy or a wider range of parameters if the signals from several pixels are used. The penalty for using these methods is that there is some loss of spatial resolution.

7.1 Reducing uncertainty on estimated proportions

If N independent measurements of pixel proportions are made then the error standard deviation on the mean is $\frac{1}{\sqrt{N}}$ of the error on a single pixel. Since the underlying model is linear, it is possible to average the signals and then carry out the inversion rather than invert the signals for many pixels and then average the resulting proportions.

To estimate the uncertainty on the proportions estimated for the collection of pixels it is necessary to use variances appropriate to the number of pixels involved, i.e. the measurement uncertainty variance is $\frac{1}{N}$ that of a single pixel. The end-member uncertainty values must also be scaled appropriately. It may be more convenient to simply scale the variance for a single pixel by the factor $\frac{1}{N}$.

7.2 Estimating spectral signatures using known proportions

If the proportions are known then it is possible to use m pixels to estimate the spectral signature in n bands, averaged over the m pixels, as long as $m > n$. It is assumed that the m pixels are independent in the sense that they provide independent items of information, i.e. the signal for any one pixel is not a linear function of the signals of a (sub-)set of the other $m - 1$ pixels.

The result applies to the area covered by the pixels, i.e. there is a loss of spatial resolution.

Chapter 8

Conclusions

This report describes the solution algorithms used by the program *MIXMOD* to analyse mixed pixel data, assuming the linear mixing model. A unique feature of the program is its ability to handle uncertainty in the assumed end-member spectra. *MIXMOD* includes routines to perform the three principal analysis steps: (1) estimate a solution, (2) evaluate the quality of the solution, and (3) provide a measure of the error for the solution.

The literature reviewed gives an overview of current applications of linear mixture modelling, as well as some pointers to the wider literature concerned with linear systems. Several studies comparing mixture modelling with alternative methods of analysis are also identified. The principal strength of mixture modelling relative to the alternatives is that because it is based on an explicit model, its limitations can be clearly identified, and within those limitations it is possible to quantify the goodness-of-fit and to evaluate appropriate error bounds.

Although a large number of mixture modelling studies have been presented, relatively few deal with issues of quantifying the solution quality or of estimating proper error bounds. This report presents methods of performing these tasks, taking account of *all* the principal sources of uncertainty. As the field of mixture modelling matures, these issues must be addressed.

Bibliography

[1] Adams, J.B., Smith, M.O., and Gillespie, A.R. (1989), Simple models for complex natural surfaces: a strategy for the hyperspectral era of remote sensing. *Proceedings of the IGARSS' 89, Vancouver, Canada: IGARSS '89, Vol. I*, pp. 16-21.

[2] Atkinson, P.M., Cutler, M.E.J., and Lewis, H. (1995), Mapping sub-pixel variation in land cover in the UK from AVHRR imagery. *Proc. 21st Annual Conference of the Remote Sensing Society, 11-14 September 1995, Southampton, UK*. Published by the Remote Sensing Society, pp. 58-65, ISBN 0 946226 20 0.

[3] Campbell, J.G. (1981) The Use of Landsat MSS Data for Ecological Mapping, *Proc. 9th Ann. Conf. Remote Sensing Society, Univ. London, Dec. 1981.* pp 143-161.

[4] Campbell, J.G., and A.A. Hashim (1992), Fuzzy Sets, Pattern Recognition, and Neural Networks ..., *Proc. 18th Ann. Conf. Remote Sensing Socy., Univ. Dundee, Sept. 1992* (AP. Cracknell and RA. Vaughan, eds.) pp. 508 -517.

[5] Cross, A.M. (1990). AVHRR as a data source for a GIS: deforestation in Amazonia. *Proceedings of the International Geoscience and Remote Sensing Symposium, 1990. (IGARSS '90)*, Washington, D.C. (New York: IEEE), pp.223-226.

[6] Foody, G.M., and Cox, D.P. (1994), Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. *Int. J. Remote Sensing*, **15**(3), 619-631.

[7] Hobbs, S.E., and Thomas, G. (1996), A goodness-of-fit method applied to spectral mixture modelling. *Proc. Image Processing: Mathematical methods and applications. Conference of The Institute of Mathematics and its Applications at Cranfield University, September 1994.* Published by Oxford University Press (in press).

[8] Holben, B.N., and Shimabukuro, Y.E. (1993), Linear mixing model applied to coarse spatial resolution data from multispectral satellite sensors. *Int. J. Remote Sensing*, **14**(11), 2231-2240.

[9] Houghton, J.T., Taylor, F.W., and Rodgers, C.D. (1984), *Remote sounding of atmospheres*, Cambridge University Press, Cambridge (ISBN 0 521 31065 2).

[10] Jolliffe, I.T. (1990), Principal components analysis: a beginner's guide - I. Introduction and application. *Weather*, **45**, 375-382.

[11] Jolliffe, I.T. (1993), Principal components analysis: a beginner's guide - II. Pitfalls, myths and extensions. *Weather*, **48**(8), 246-253.

[12] Kerdiles, H., and Grondona, M.O. (1995), NOAA-AVHRR NDVI decomposition and subpixel classification using linear mixing in the Argentinean Pampa. *International Journal of Remote Sensing*, **16**(7), 1303-1325.

[13] Lillesand, T.M., and RW. Kiefer, (1994) *Remote Sensing and Image Interpretation, 3rd. ed.*, John Wiley.

[14] Novo, E.M., and Shimabukuro, Y.E. (1994), Spectral mixture analysis of inland tropical waters. *Int. J. Remote Sensing*, **15**(6), 1351-1356.

[15] Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P., (1992), *Numerical Recipes in C; Second Edition*, Cambridge University Press: Cambridge (ISBN 0 521 43108 5).

[16] Quarmby, N.A., (1992), Towards continental scale crop area estimation, *International Journal of Remote Sensing*, 13: 981-989.

[17] Quarmby, N.A., Townshend, J.R.G., Settle, J.J., White, K.H., Milnes, M., Hindle, T.L. and Silleos, N., (1992). Linear mixture modelling applied to AVHRR data for crop estimation. *International Journal of Remote Sensing*, **13**:415-425.

[18] Roberts, D.A., Adams, J.B., and Smith, M.O. (1990), Transmission and scattering of light by leaves: effects of spectral mixtures. *Proc. IGARSS '90, Remote Sensing for the Nineties, Vol. II*, 967-970.

[19] Rodgers, C.D. (1976), Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. *Reviews of Geophysics and Space Physics*, **14** (4), November 1976, pp 609-624.

[20] Rodgers, C.D. (1990), Characterization and error analysis of profiles retrieved from remote sounding measurements. *Journal of Geophysical Research*, **95**(D5), pp 5587-5595, 20 April 1990.

[21] Sabol, D.E., Adams, J.B., and Smith, M.O. (1990), Predicting the spectral detectability of surface materials using spectral mixture analysis. *Proc. IGARSS '90, Remote Sensing for the Nineties, Vol. II*, 967-970.

[22] Sabol, D.E., Adams, J.B., and Smith, M.O. (1992), Quantitative subpixel spectral detection of targets in multispectral images. *J. Geophysical Research*, **97** (E2), 2659-2672, 25 February 1992.

[23] Schanzer, D.L. (1993), Comments on"The least squares mixing models to generate fraction images derived for remote sensing multispectral data". *IEEE Transactions on Geoscience and Remote Sensing*, **31**(3), May 1993.

[24] Settle, J.J., and Drake, N.A. (1993), Linear mixing and estimation of ground cover proportions. *Int. J. Remote Sensing*, **14**(6), 1159-1177.

[25] Shimabukuro, Y.E., and Smith, J.A. (1991), The least-squares mixing models to generate fraction images derived from remote sensing multispectral data. *IEEE Transactions on Geoscience and Remote Sensing*, **29**(1), 16-20, January 1991.

[26] Smith, M.O., Johnson, P.E. and Adams, J.B. (1985). Quantitative determination of mineral types and abundances from reflectance spectra using principal component analysis. *Journal of Geophysical Research*, **90**:792-804.

[27] Smith, M.O., Adams, J.B., and Gillespie, A.R. (1990), Reference endmembers for spectral mixture analysis. *Proc. 5th Australian Remote Sensing Conference, 8-12 October 1990, Perth, Australia, Vol. I*, 331-340.

[28] Smith, M.O., Adams, J.B., and Sabol, D.E. (1994), Spectral mixture analysis - new strategies for the analysis of multispectral data. *Imaging Spectrometry - a Tool for Environmental Observations*, Ed. J. Hill and J. Megier, Kluwer Academic Publishers, Dordrecht, pp 125-143 (ISBN 0-7923-2965-1).

[29] Thomas, G, Hobbs, S E and Dufour, M J D (1996). Woodland area estimation by spectral mixing: applying a goodness-of-fit solution method. *International Journal of Remote Sensing*, **17**, pp 291-301.

[30] van Kootwijk, E.J., van der Voet, H., and Berdowski, J.J.M., (1995), Estimation of ground cover composition per pixel after matching image and ground data with subpixel accuracy. *International Journal of Remote Sensing*, **16**(1), 97-111.

[31] Zhukov, B., Berger, M., Lanzl, F., and Kaufmann, H., (1995), A new technique for merging multispectral and panchromatic images revealing subpixel spectral variation. *Proceedings International Geoscience and Remote Sensing Symposium 1995, IGARSS '95, Firenze, Italy, 10-14 July 1995, Vol. III*, IEEE.