A Multi-layer Perceptron based Non-linear Mixture Model to estimate class abundance from mixed pixels

Uttam Kumar[1, 2], S. Kumar Raja[5], C. Mukhopadhyay[1], and T.V. Ramachandra[2, 3, 4, *], *Senior Member, IEEE*
[1]Department of Management Studies, [2]Centre for Sustainable Technologies, [3]Centre for Ecological Sciences,
[4]Centre for *infrastructure*, Sustainable Transport and Urban Planning, Indian Institute of Science, Bangalore – 560012, India.
[5]Institut de Recherche en Informatique et Systèmes Aléatoires, 35042 Rennes cedex - France & Technicolor Research & Innovation, Cesson Sévigné, France.

***Abstract*— Sub-pixel classification is essential for the successful description of many land cover (LC) features with spatial resolution less than the size of the image pixels. A commonly used approach for sub-pixel classification is linear mixture models (LMM). Even though, LMM have shown acceptable results, pragmatically, linear mixtures do not exist. A non-linear mixture model, therefore, may better describe the resultant mixture spectra for endmember (pure pixel) distribution. In this paper, we propose a new methodology for inferring LC fractions by a process called *automatic linear-nonlinear mixture model* (AL-NLMM). AL-NLMM is a three step process where the endmembers are first derived from an automated algorithm. These endmembers are used by the LMM in the second step that provides abundance estimation in a linear fashion. Finally, the abundance values along with the training samples representing the actual proportions are fed to multi-layer perceptron (MLP) architecture as input to train the neurons which further refines the abundance estimates to account for the non-linear nature of the mixing classes of interest. AL-NLMM is validated on computer simulated hyperspectral data of 200 bands. Validation of the output showed overall RMSE of 0.0089±0.0022 with LMM and 0.0030±0.0001 with the MLP based AL-NLMM, when compared to actual class proportions indicating that individual class abundances obtained from AL-NLMM are very close to the real observations.**

Keywords*— *Hyperspectral, sub-pixel classification, multi-layer perceptron, non-linear unmixing

I. INTRODUCTION

Hyperspectral imaging spectrometers collect images in the form of image cube that represents reflected energy from the Earth's surface materials, where each pixel has the resultant mixed spectrum from reflected source radiation [1]. The mixed spectrum phenomenon causes mixed pixel problem because the intrinsic scale of spatial variation on the Earth's land cover (LC) is finer than the scale of sampling imposed by the image pixels. In other words, mixed pixels are a mixture of more than one distinct object, and exist for one of the two reasons. Firstly, if the spatial resolution of the sensor is not high enough to separate different LC types, these can jointly occupy a single pixel, and the resulting spectral measurement will be a composite of the individual spectra that reside within a pixel. Secondly, mixed pixels can also result when distinct LC types are combined into a homogeneous mixture. This occurs independent of the spatial resolution of the sensor [2].

Commonly used approach to mixed pixel classification has been linear spectral unmixing [3], supervised fuzzy-c means classification [4], artificial neural networks [5 and 6], and Gaussian mixture discriminant analysis [7], etc. which uses a linear mixture model (LMM) to estimate the abundance fractions of spectral signatures lying within a pixel. LMM assumes that the reflectance spectrum of a mixture is systematic combination of the component reflectance spectra in the mixture (called endmembers). The combination of these endmembers is linear if component of interest in a pixel appear in spatially segregated patterns. If, however, the components are in intimate association, the electromagnetic spectrum typically interacts with more than one component as it is multiple scattered, and the mixing systematics between the different components are highly non-linear. In other words, non-linear mixing occurs when radiance is modified by one material before interacting with another one under the assumption that incident solar radiation is scattered within the scene itself and that these interaction events may involve several types of ground cover materials [8] and require non-linear mixture model (NLMM) for unmixing the components of interest.

If K is the number of spectral bands in the data set, and P, the number of distinct classes of objects in the physical scene, then associated with each pixel is a K-dimensional vector $\mathbf{y}$ whose components are the gray values corresponding to the K bands. If $\mathbf{E} = [e_1, e_2, \ldots e_{p, \ldots,} e_P]$, where $\{e_p\}$ is a column vector representing the spectral signature of the p^{th} target material or category; the column vectors of the $K \times P$ matrix $\mathbf{E}$ are the endmembers. For a given pixel, the abundance fraction of the p^{th} target material present in a pixel is denoted by α_p, and these values are the components of the P-dimensional abundance vector $\boldsymbol{\alpha}$. A nonlinear mixture model is then expressed as:

$$\mathbf{y} = f(\mathbf{E}, \boldsymbol{\alpha}) + \boldsymbol{\eta} \quad (1)$$

where, f is an unknown nonlinear function that defines the interaction between $\mathbf{E}$ and $\boldsymbol{\alpha}$, and $\boldsymbol{\eta}$ is a noise vector. In this context, artificial neural network (ANN) based NLMMs outperform the traditional linear unmixing models. ANNs have been widely studied as a promising alternative to accomplish the difficult task of estimating fractional abundances of endmembers. Atkinson et al., (1997) [9] applied a MLP (milti-layer perceptron) model to decompose AVHRR imagery and

 978-1-4244-8943-5/11/$26.00 ©2011 IEEE

was superior to the linear unmixing model and a fuzzy c-means classifier. Another popular ANN model - ARTMAP was first introduced to identify the life form components of the vegetation mixture [10] using Landsat imagery which could capture non-linear effects and thus performed better than LMM [11]. ART MMAP, an extension of ARTMAP was designed specifically for mixture analysis with enhanced interpolation function and provided better prediction of mixture information than ARTMAP [12]. Regression tree has also been used as non-linear unmixing models [6]. All of these methods are stand alone and work on the data directly when endmembers are known a priori. The objective of this paper is to develop an automated procedure to unmix hyperspectral imagery for obtaining fraction that accounts for the non-linear mixture of the class types. We call this model as automatic linear-nonlinear mixture model (AL-NLMM) as shown in the block diagram in Fig. 1. This paper is structured in six sections. Methods for automatic endmember extraction, linear unmixing and MLP are discussed in section II followed by the description of AL-NLMM in section III. Data preparation is dealt in section IV with the experimental results and discussion in section V. Section VI concludes with model limitations.

II. METHODOLOGY

A. Automatic endmember extraction – The N-FINDR algorithm – N-FINDR is a fully automatic technique for endmember selection [13] as briefed below:

1. Let P denote the number of classes or endmembers to be identified.
2. Perform a Principal Component Analysis (PCA)-decomposition of the data and reduce the dimension of the data to P - 1. In what follows, we assume that the data is transformed to P - 1 dimension PCA space.
3. Pick P pixels from the set and compute the simplex volume generated by the spectra of the P pixels.
4. Replace each endmember with the spectrum of each pixel in the data set and recompute the simplex volume. If the volume increases, the spectrum of the new pixel is retained as a potential endmember.

The above steps are executed iteratively considering all the pixels, and the final set of spectra retained is taken to be the endmembers.

B. Orthogonal subspace projection (OSP) – In order to solve the linear part of the mixing problem, we used OSP [14]. OSP is based on two aspects: 1) how to best utilise the target knowledge provided a priori and 2) how to effectively make use of numerous spectral bands. Assuming the LMM, the observation vector y is related to E by

$$\mathbf{y} = \mathbf{E}\boldsymbol{\alpha} + \boldsymbol{\eta}. \quad (2)$$

We further assume that the components of the noise vector $\boldsymbol{\eta}$ are zero-mean random variables that are independent and identically distributed. Therefore, the covariance matrix of the noise vector is $\sigma^2\mathbf{I}$, where σ^2 is the variance, and $\mathbf{I}$ is K x K identity matrix.

The conventional approach [15] to extract the abundance values is to minimise $\| y - E\alpha \|$, and the estimate for the abundance is

$$\alpha = (E^T E)^{-1} E^T \mathbf{y} \quad (3)$$

which is termed as the Unconstrained Least Squares (ULS) estimate of the abundance. Imposing the sum-to-one constraint on the abundance values, gives the Constrained Least Squares (CLS) estimate of the abundance as,

$$\alpha = (E^T E)^{-1} \left(E^T \mathbf{y} - \frac{\lambda}{2} 1 \right) \quad (4)$$

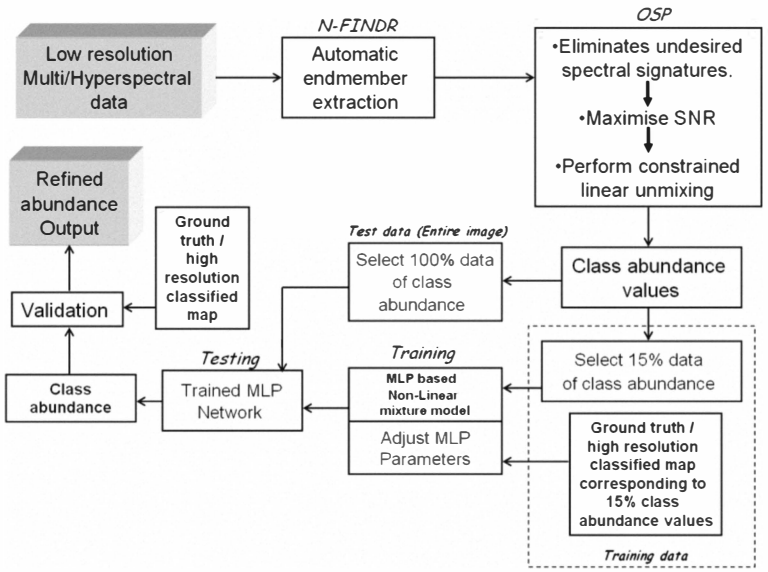

Fig. 1. Block diagram of the proposed AL-NLMM method for spectral mixture analysis.

where,

$$\lambda = \frac{2(1^T(E^TE)^{-1}E^T\mathbf{y}-1)}{1^T(E^TE)^{-1}1}. \quad (5)$$

In order to find the abundance of the p^{th} target material (α_p), let the corresponding spectral signature of the desired target material be denoted as **d**. The term $\mathbf{E}\boldsymbol{\alpha}$ in (2) can be rewritten to separate the desired spectral signature **d** from the rest as:

$$E\alpha = \mathbf{d}\alpha_p + \mathbf{Rr} \quad (6)$$

where **r** contains the abundance of the rest of the end-members, and **R** is a (K x P– 1) matrix containing the columns of **E** except for the column vector **d**. We rewrite (2) as

$$\mathbf{y} = \mathbf{d}\alpha_p + \mathbf{Rr} + \boldsymbol{\eta}. \quad (7)$$

The interfering signatures present in **R** can be removed from (7) by the operator,

$$P = (\mathbf{I} - \mathbf{R}(\mathbf{R}^T\mathbf{R})^{-1}\mathbf{R}^T) \quad (8)$$

which is used to project the vector **y** into a space orthogonal to the space spanned by the interfering spectral signatures. Therefore, operating on **y** with **P**, and noting that **PR** = 0,

$$Py = P\mathbf{d}\alpha_p + P\boldsymbol{\eta}. \quad (9)$$

The next step is to find an operator w^T which maximises the SNR given by

$$SNR = \frac{(w^TP\mathbf{d}\alpha_p)^2}{\varepsilon\{(w^TP\boldsymbol{\eta})^2\}} = \frac{w^TP\mathbf{d}\alpha_p^2d^TP^Tw)^2}{w^TP\varepsilon\{(\boldsymbol{\eta}\boldsymbol{\eta}^T)P^Tw} \quad (10)$$

$$= \left(\frac{\alpha_p^2}{\sigma}\right)^2 \frac{w^TP\mathbf{d}\mathbf{d}^TP^Tw}{w^TPP^Tw}$$

Maximising the SNR leads to the generalised eigenvalue problem: $P\mathbf{d}\mathbf{d}^TP^Tw = \lambda PP^Tw$. The eigenvector corresponding to the maximum eigenvalue is the vector '**w**'. It can be shown that the **w** which maximises the SNR is given by

$$\mathbf{w} = k\mathbf{d}. \quad (11)$$

Therefore, an optimal estimate of α_p is given by

$$\hat{\alpha}_p = \frac{y^TP^TPy}{\mathbf{d}^TP^TP\mathbf{d}} \quad (12)$$

In the absence of noise, the estimate matches with the exact value as in (7). The value of α is the abundance of the p^{th} class (in an abundance map) ranging from 0 to 1 in any given pixel. 0 indicates absence of a particular class and 1 indicates full presence of that class in that particular pixel.

C. ANN based multi-layer perceptron (MLP) – The advent of ANN approaches is mainly due to their power in pattern recognition, interpolation, prediction, forecasting, classification and process modeling [16]. A MLP network comprises a number of identical units organised in layers, with those on one layer connected to those on the next layer so that the output of one layer are used as input to the next layer.

A detailed introduction on MLP can be found in literatures [16-19]. The main aspects here are: (i) the order of presentation of training samples should be randomised from epoch to epoch; and (ii) the momentum and learning rate parameters are typically adjusted (and usually decreased) as the number of training iterations increases.

III. AUTOMATIC LINEAR-NONLINEAR MIXTURE MODEL

Despite many attempts of using ANN for unmixing models, ANN-based non-linear unmixing techniques remain largely unexplored for general purpose applications [20]. Only [1], [6], [8] and [21] have been some of the pioneering work in NLMM to be considered as a general model for ANN-based non-linear unmixing independent of physical properties of the observed classes. Some of these applications are however, difficult and complex in their implementation that makes LMM easy to implement, generalise and reconstruct. Therefore, in our approach, we make use of the LMM output as the input to NLMM to refine the fraction estimates. The MLP architecture can be extended to produce a continuous-values output for sub-pixel classification problem. The entries to the MLP model are the abundance **(a)** output obtained from LMM, which is denoted by a^i_{LMM} where i = 1, …, E, and the neuron count at the input layer equals the number of endmember classes (estimated by a fully constrained LMM) as shown in Fig. 2. The training process is based on error back-propagation algorithm [21], where the respective weights in the output and hidden nodes (W and V in Fig. 3) are modified depending on the error (δe), the input data and the learning parameter alpha (α). The activation rule used here for the hidden and output layer nodes is defined by the logistic function

$$f(x) = \frac{1}{1+e^{-x}}. \quad (13)$$

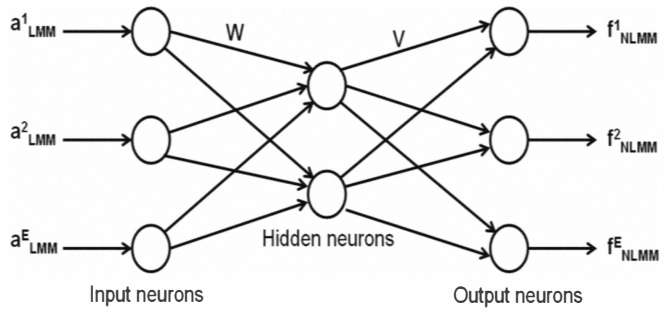

Fig. 2. Architecture of the MLP model.

δe of the output layer is calculated as the difference between the fraction (f) estimation outputs f^i_{NLMM}, i = 1, … E, provided by the network architecture and a set of desired output given by actual fractional abundances available for the training samples. The resulting error is back-propagated until the convergence is reached. One of the earlier works by [20] attempted a similar NLMM methodology, which made use of a modified MLP neural network (NN), whose entries were determined by a linear activation function provided by a Hopfield NN (HNN).

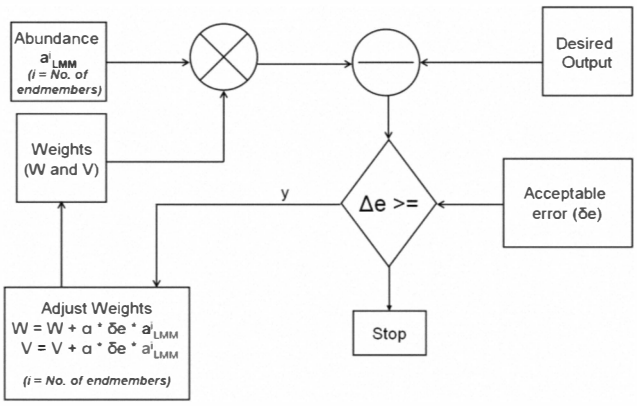

Fig. 3. MLP structural diagram.

The combined HNN/MLP method used the LMM to provide an initial abundance estimation and then refined the estimation using a non-linear model. As per [20], this was the first and the only approach in the literature that integrated linear and NLMM.

IV. Computer Simulations

One of the major problems involved in analysing the quality of fractional estimation methods is the fact that ground truth information about the real abundances of materials at sub-pixel levels is difficult to obtain in real scenarios [21]. In order to avoid this shortcoming, simulation of hyperspectral imagery was carried out to examine the algorithm's performance in a controlled manner.

Spectral libraries of four minerals - alunite, buddingtonite, kaolinite and calcite (available at http://speclib.jpl.nasa.gov/) were used to generate synthetic data. Plaza et al. [21], used the signatures of soil (e_1) and vegetation (e_2) to create a simulated image with non-linear mixtures using a simple logarithmic function. The abundance of e_1 and e_2 were assigned according to (14)

$$\mathbf{y}(x,y) = \sum_{P=1}^{2} \mathbf{e}_p \cdot s_p(x,y) \tag{14}$$

where, y denotes a vector containing the simulated discrete spectrum of the pixel at spatial coordinates (x,y) of the simulated image, $s_p(x,y) = \log \alpha_p(x,y)$ is the contribution of endmember e_p and $\alpha_p(x,y)$ is the fractional abundance of e_p at (x,y). A limitation here is that even though all the pixels are mixed in different proportions, there are no instances of pure pixels. If α is 1, we expect the observed hyperspectral signature to be solely from one material, and therefore, ideally it should be identical to the endmember itself. Here, as the abundance increases towards 1, log (α) approaches 0, thereby suppressing the contribution of that particular endmember. On the other hand, if α is 0, log (α) approaches $-\infty$, and therefore, starts dominating in the observed spectral signature. Of course, it will appear as negative numbers. This is against our physical understanding as to how a material which is almost not present in the pixel, contribute to the observation in a dominant way? That is, the model is not able to highlight the endmember of the correct material when its contribution is 1 and gives a wrong endmember when its contribution is 0. To overcome this limitation, we modify the model in (14) by (15):

$$\mathbf{y}(x,y) = \sum_{P=1}^{4} \mathbf{sig}_p \cdot s_p(x,y) \tag{15}$$

where, sig_p is the signature corresponding to p^{th} mineral, $s_p(x,y) = \log(1+\alpha_p(x,y))$ is the contribution of endmember e_p and $\alpha_p(x,y)$ is the fractional abundance of e_p in the pixel at (x,y).

Simulated synthetic non-linear mixture hyperspectral data of 200 bands (250 x 250) using four minerals were classified using Maximum Likelihood Classifier (MLC) with signatures from the spectral libraries. This constitutes high-resolution (HR) images. These images were used to generate synthetic mixed pixels of 25 x 25 (referred as low-resolution - LR images). Four endmembers were extracted from LR images, and subsequently, abundance images were estimated corresponding to each endmember. Percentage abundance for a group of 10 x 10 pixels were computed for this entire HR classified image (250 x 250) obtained from MLC. This new image of size 25 x 25 was used as reference for validating the LR abundance output. However, the HR MLC based classified output (250 x 250) was not validated as the same spectral library which was used for generating the individual class signatures for classification of the HR image was also used to create the synthetic images. Abundance values from 15% of the pixels obtained from linear unmixing along with the corresponding proportions obtained from the 250 x 250 classified image obtained by MLC were used for training the neurons in MLP. For example, each input sample to the MLP has the abundance values obtained from OSP for each of the four classes (0.2, 0.3, 0.1, 0.4 = 1 or 100 % of a pixel) and the proportion of each class as derived from HR MLC based classified map (0.18, 0.27, 0.2, 0.35 = 1 or 100 % of a pixel) by considering 25 x 25 classified pixels and finding the percentage of each class separately which is equivalent to 1 x 1 LR pixel spatially. Testing was done on the entire output abundance images (100% pixels).

V. Results and Discuussions

Three images from the 200-bands are shown in Fig. 4 and the classified output of the 250 x 250 hyperspectral 200 bands data is shown in Fig. 5. The proportions of each of the 4 minerals were computed based on 10 x 10 groups of pixels for 625 groups [(250 x 250) divided by (10 x 10)]. N-FINDR was used to extract the endmembers from the synthetic mixed pixels, which are shown in Fig. 6. The endmembers identified by the algorithm (drawn in red) have a good match with the actual ones (green in color). Abundances of each of the minerals from the artificial mixed pixels obtained from LMM are as shown in Fig. 7 (b)–(e). Fig. 7 (a) is the 10 times down-sampled image of the original mineral classified image (250 x 250) shown in Fig. 5 to compare hard classification with the abundance map visually. A three-layer MLP architecture was made with 4 input, 1 hidden and 4 output layers.

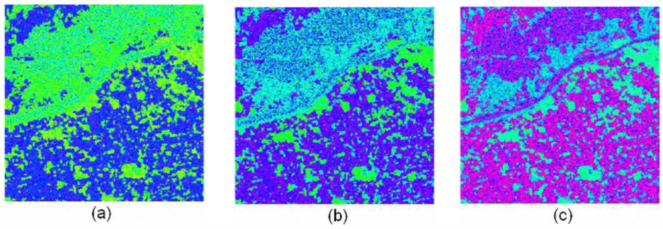

Fig. 4. A 200 band hyperspectral images generated from spectral libraries of four different minerals - (a) band 1 (b) band 100 (c) band 200.

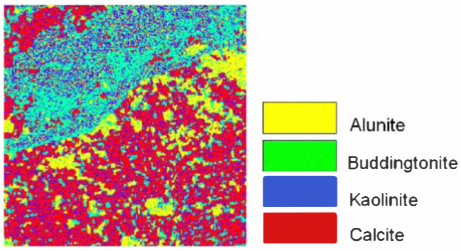

Fig. 5. Mineral classified map.

The number of hidden nodes in the hidden layer, learning rate, momentum and epoch were varied in steps to estimate the best abundance values that could account for the non-linearity in the mineral mixtures; until the performance saturated, as shown in Fig. 8. Table I lists the values of the training parameters along with the training time and the overall RMSE of the MLP network for every 500 epochs. Three measures of performance were used to evaluate the output from artificial dataset – RMSE, correlation, Bivariate Distribution Functions (BDFs). BDF is helpful to visualise the accuracy of prediction by mixture models. BDFs were plotted against the real proportions as shown in Fig. 9. Pearson's product-moment correlation at 95% confidence interval and RMSE between the actual and estimated proportion from LMM and NLMM are given in table II. The average RMSE of the LMM was 0.0089±0.0022 while the average RMSE of the NLMM was 0.0030±0.0001 demonstrating the superiority of the NLMM over the LMM. The MLP network can successfully approximate virtually any function when trained correctly. This implies that training is the most important step in MLP-based classification. 15% of the LMM based abundances along with the actual abundances were used for training and the MLP model could interpolate to produce many more combinations of class proportions to match the testing samples. The interpolation function minimises the difference between estimated and real class proportions. Since the approach in this work is a hybrid of linear and non-linear estimator, it can be easily adapted to either the linear on non-linear mixture model. Initialisation and training are the two issues in a MLP network that has to be dealt carefully. However, an effective learning algorithm should not depend on initial conditions, which can only affect the convergence rate but should not alter the final results [21]. Often, this is not the case of learning algorithms used for NNs. In order for a mixture model to be effective, initial values must be

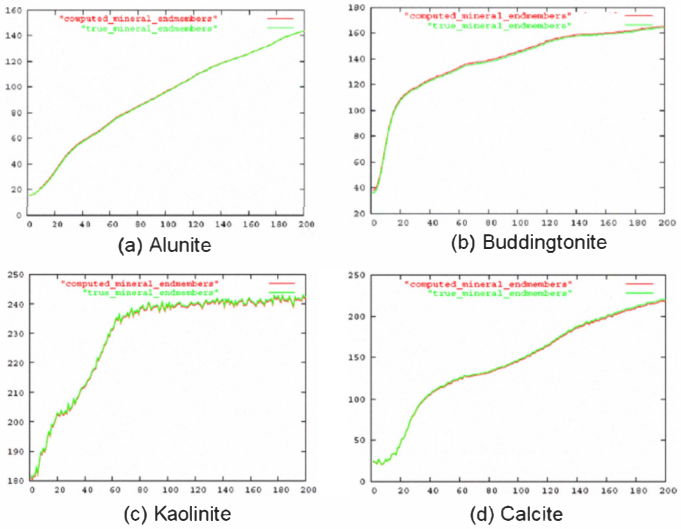

Fig. 6. Comparison between the true endmembers and endmembers computed from the N-FINDR algorithm. **[X-axis: Band number, Y-axis: Reflectance value.]**

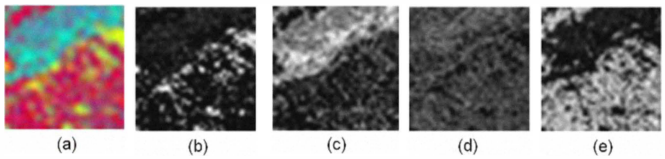

Fig. 7. (a) 10 times down-sampled mineral classified image; (b) – (d) show abundances of the four mineral obtained from the LMM.

TABLE I
DETAILS OF TRAINING FOR UNMIXING OF SIMULATED DATASET

No. of epochs	Learning rate	Momentum term	Training time (sec)	Unmixing time (sec)	Overall RMSE
500	0.90	0.5	4	8	0.0160
1000	0.85	0.4	5	8	0.0117
1500	**0.80**	**0.3**	**7**	**7**	**0.0030**
2000	0.70	0.2	7	6	0.0071
2500	0.60	0.1	8	5	0.0115

TABLE II
CORRELATION AND RMSE BETWEEN ACTUAL AND PREDICTED PROPORTIONS

Classes	Correlation (r) ($p < 2.2e^{-16}$)		RMSE	
	LMM	NLMM	LMM	NLMM
Alunite	0.67	0.97	0.0120	0.0032
Buddingtonite	0.71	0.98	0.0073	0.0029
Kaolinite	0.73	0.98	0.0088	0.0031
Calcite	0.75	0.99	0.0076	0.0029

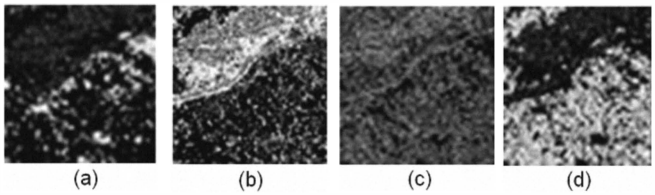

Fig. 8: Abundances maps of the four mineral obtained from the NLMM.

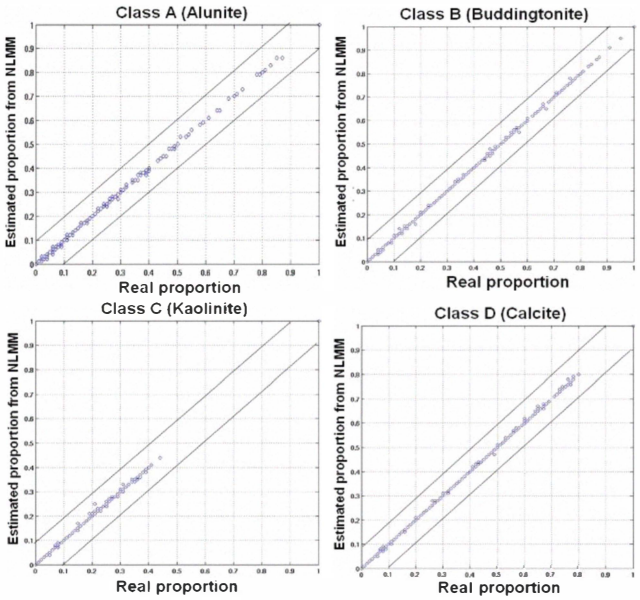

Fig. 9. BDFs of simulated test data for the four minerals obtained LMM.

representative and cannot be arbitrary. Therefore, an important issue for successful application of this approach is the selection of endmembers. One of the potential draw back of the N-FINDR algorithm is that it requires at least one pixel in the image to be pure. This may not be the situation always. In such cases, alternative methods of endmember extraction have to be studied and integrated in the algorithm. Also, the simulated dataset in our work is noise-free (has no noise component), so, further research is required to analyse the impact of the interference of noise.

VI. CONCLUSIONS

The proposed AL-NLMM algorithm integrates the concept of both linear and non-linear mixing. The method is based on extracting the endmembers from the image, followed by unmixing in the second step and then interpolating the fractions for the whole image based on training data that has both the estimated abundance obtained from second step and real proportion. The results on simulated hyperspectral data shows that the method gives acceptable results for unmixing pixels with an overall RMSE of 0.0089±0.00215 with LMM and 0.0030±0.0001 with the MLP based NLMM when compared to real class proportions. It may be concluded that influence due to multiple reflections among ground cover targets has to be considered for the abundance estimation. While a linear detection method might work adequately for many scenarios, a non-linear model might perform better. The future work will involve developing methods to obtain pure pixels when there are no endmembers in the scene.

REFERENCES

[1] K. J. Guilfoyle, M. L. Althouse and C.-I. Chang, "A quantitative and comparative analysis of linear and nonlinear spectral mixture models using radial basis neural networks," IEEE Trans., Geosci. Remote Sens., vol. 39, no. 10, pp. 2314-2318, 2001.

[2] N. Keshava and J. F. Mustard, "Spectral unmixing," IEEE Signal Processing Magazine, vol. 19, pp. 44-57, 2002.

[3] U. Kumar, N. Kerle, and T.V. Ramachandra, "Constrained linear spectral unmixing technique for regional land cover mapping using MODIS data," In: Innovations and advanced techniques in systems, computing sciences and software engineering/ed by Khaled Elleithy, Berlin: Springer, pp. 87-95, 2008.

[4] G. M. Foody and D. P. Cox, "Subpixel land cover composition estimation using a linear mixture model and fuzzy membership functions," *Int. J. Remote Sensing*, vol. 15, no. 3, pp. 619-631, 1994.

[5] I. Kanellopoulos, A. Varfis, G. G. Wilkinson, and J. Megier, "Land-cover discrimination in SPOT imagery by artificial neural network-a twenty class experiment," *Int. J. Remote Sensing*, vol. 13, no. 5, pp. 917–924, 1992.

[6] W. Liu and E. Y. Wu, "Comparison of non-linear mixture models: sub-pixel classification," *Remote Sens. Environ.*, vol. 94, pp. 145-154, 2005.

[7] J. Ju, E. D. Kolaczyk and S. Gopal, "Gaussian mixture discriminant analysis and sub-pixel and sub-pixel land cover characterization in remote sensing," *Remote Sens., Environ.*, vol. 84, pp. 550 – 560, 2003.

[8] K. Arai, "Non-linear mixture model of mixed pixels in remote sensing satellite images based in Monte Carlo simulation," *Advances in Space Research*, vol. 42, pp. 1715-1723, 2008.

[9] P. M. Atkinson, M. E. J. Cutler and H. Lewis, "Mapping sub-pixel proportional land cover with AVHRR imagery," *Int. J. Remote Sensing*, vol. 18, no. 4, pp. 917-935, 1997.

[10] G. A. Carpenter, S. Gopal, S. Macomber, S. Martens and C. E. Woodcock, "A neural network method for mixture estimation for vegetation mapping," *Remote Sens. Environ.*, vol. 70, pp. 138- 152, 1999.

[11] W. Liu, S. Gopal and C. Woodcock, "ARTMAP Multisensor/resolution framework for land cover characterization," *in Proc. 4th Annu. Conf. information fusion*, Montreal, Canada, 2001, pp. 11-16.

[12] W. Liu, K. Seto, E. Wu, S. Gopal and C. Woodcock, "ART-MMAP: A neural network approach to subpixel classification," *IEEE Trans., Geosci. Remote Sens.*, vol. 42, no. 9, pp. 1976-1983, 2004.

[13] M. E. Winter, "N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data," *in Proc. of the SPIE: Imaging Spectrometry*, 1999, vol. 3753, pp. 266-275.

[14] C.-I. Chang, "Orthogonal Subspace Projection (OSP) Revisited: A Comprehensive Study and Analysis," I*EEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 502-518, 2005.

[15] A. A. Nielsen, "Spectral Mixture Analysis; Linear and Semi-parametric Full and Iterated Partial Unmixing in Multi- and Hyperspectral Image Data," *International Journal of Computer Vision,* vol. 42, no. (1/2), pp. 17-37, 2001.

[16] S. Haykin, Neural Networks: a comprehensive foundation, NY: Macmillan, New York, 1998.

[17] R. O. Duda, P. E. Hart and D. G. Stork, Pattern classification, A Wiley-Interscience Publication, Indianapolis, IN., 2000, pp. 517-598.

[18] T. Kavzoglu and P. M. Mather, "The use of backpropagating artificial neural networks in land cover classification," *Int. J. Remote Sensing*, vol. 24, no. 23, pp. 4907-4938, 2003.

[19] J. F. Mas, "Mapping land use/cover in a tropical coastal area using satellite sensor data, GIS and artificial neural networks," *Estuarine, Coastal and Shelf Science*, vol. 59, pp. 219-230, 2003.

[20] J. Plaza, P. Martinez, R. Pérez and A. Plaza, "Nonlinear neural network mixture models for fractional abundance estimation in AVIRIS Hyperspectral Images," *Proceedings of the NASA Jet Propulsion Laboratory AVIRIS Airborne Earth Science Workshop, Pasadena, California*, March 31-April 2, 2004, pp. 1-12.

[21] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–535, 1986.