The Unsupervised Gravitational Mass Weighted Probability PCA for Pixel-Wise and Sub-Pixel Wise Classification

Nandhini K, Porkodi R

Abstract: Recent advancements of the hyperspectral remote sensing data extend high spectral information along with the huge dimensional features. Feature extraction helps to project the high dimensional features to low dimensional features without missing any spectral information of the hyperspectral data. This paper highlights the two issues of the hyperspectral data pixel-wise and sub-pixel-wise classification paradigms by proposing the novel weighted feature extraction to improve the classification accuracy. The first issue is the curse of dimensionality of hyperspectral data and the second issue is the former method in estimating model requires the proper weighted method due to lack in estimation of the covariance matrix. This paper proposed gravitational mass weighted method in probability PCA (GMWProbPCA) which is inspired from the gravitational search algorithm to project the high dimensional features to the low dimensional features. In pixel-wise-classification paradigm, the proposed algorithms was compared with the state-of-art techniques supervised LDA and unsupervised PCA and Probability PCA using SVM classification accuracy. The proposed method yields high classification average accuracy of 93.75% and overall accuracy of 89.74% respectively. In sub-pixel-wise classification, the fully constrained least square method and khype are employed and justified as with and without proposed feature extraction GMWProbPCA. The less average root mean squared error is achieved for both FCLS and Khype with proposed GMWProbPCA are 0.044 % and 0.037%.

Index Terms: Feature extraction, Hyperspectral Data, Gravitational Mass Weighted, Gravitational Mass Weighted Probability PCA, Probability PCA

INTRODUCTION

The customary way of hyperspectral image classification over the decades is the pixel-wise where it assumes that pure pixels are presented in it. The following are the three different learning models often used; 1) Supervised 2) Unsupervised 3) Semi-supervised methods. The use of spectral informative in each pixel is also called as the spectral signature. The different types of classification will be adapted based on homogeneity or heterogeneity spatial area. The homogeneity spatial area considers that each pixel contain only pure pixel, then the classification of the class labels for the particular scene will be classified and it is referred as a pixel-wise classification. But most of the real time hyperspectral data are not falls under the homogeneity spatial area where the classification of different fraction of pure substances are mixed in a single pixel is also referred as sub-pixel-wise classification or spectral unmixing. For several decades, the feature extraction techniques try to solve the curse of dimensionality in any domain without loss of any information especially in hyperspectral data. Hyperspectral feature extraction techniques helps to improve the accuracy for any classifiers. The detailed study of the literatures helps to derive the vast knowledge in either way by enhancing or developing the appropriate framework or algorithms for hyperspectral feature extraction to improve the classification accuracy by pixel-wise and sub-pixel-wise classification.

- *Nandhini K, Department of Computer Science, Bharathiar University, Coimbatore-641046, India. E-mail: nandhinik2@gmail.com*
- *Dr. R. Porkodi, Associate Professor, Department of Computer Science, Bharathiar University, Coimbatore-641046, India. E-mail: porkodi_r76@buc.edu.in*

[Saurabh Prasad and Lori Mann Bruce, 2008] proposed the subspace LDA for hyperspectral target detection. Basically for supervised feature extraction most of the time Fishers linear discriminant analysis (LDA) model is adopted and for unsupervised learning, PCA model is adapted for hyperspectral data. The limitation of the LDA is refers to the ill-posed problems where the solution retrieves from LDA lead to another issue. To overcome the ill-posed problem in LDA the PCA technique is employed. The subspace LDA method takes the advantages of the two algorithms and hyperspectral targets are recognized effectively [1]. [Michael J. Mendenhall and Erzsébet Merényi, 2008] proposed generalized relevance learning vector quantization (GRLVQ) and improved version of GRLVQ (GRLVQI) which derived from learning vector quantization and generalized learning vector quantization. Based on the learning vector quantization the winner selection method is used in GRLVQ. There are three major issues found in the GRLVQ as 1) address and update in the rule 2) speed while converging and 3) classification accuracy. The above issues are the addressed in the GRLVQI algorithm. [Barat Mojaradi et.al. 2009] introduced the prototype space spectral feature extraction (PSFE) method for reducing the hyperspectral dimensionality. The author drawn major contributions to this study are; 1) the search algorithm is handled by the SFE and it replaces with the spectral channels to investigate the potential spectral regions by considering the measure of OA (Overall accuracy) and training dataset. 2) The disjoint spectral features are identified and combined with the extracted features. 3) The spectral library is acquired 4) the estimation of scatter matrices are avoided in the huge dimensional feature space due to first order statistics. The PSFE algorithm can be applied in the scenario of unsupervised feature extraction. Three different hyperspectral datasets are introduced in this study are Botswana, Kennedy space centre and AVIRIS Indian pines datasets are used [2] [3]. [Jinn-Min Yang et.al. 2010] introduced the new nonparametric cosine based nonparametric feature extraction (CNFE) method to reduce the dimensionality, The CNFE method uses the cosine based distance metric in with-in-class and between-class matrices in which weights are generated. The CNFE incorporated with the KNN classifier (CKNN) which uses the nearest neighbor metric for the effective feature extraction in hyperspectral real time data. [Hsiao-Yun Huang and Bor-Chen Kuo, 2010] constructed the method which is based on the criterion the scatter matrices are acquired and it boosts the classifier such as LDA and NDA to have maximum bound separability between classes in case of overlap between classes. The criterion based new double weighted proportion feature extraction technique is proposed for hyperspectral data to reduce the dimensionality and the cause of overlaps between class distributions [4] [5]. [Inmaculada Dopido et.al. 2012] proposed the new feature extraction algorithm to reduce the hyperspectral dimensionality. The different aspects that needs to considered in feature extraction are the following the spatial information cannot be included in the unmixing based chains but the spatial information plays the major role in investigating the correlated spectral features. The next is needed the apriori spectral information combined with the spatial information. The optimal features need to be extracted before classification of the hyperspectral data. The proposed research constructed based on considering the aforementioned aspects. It gains the sub-pixel information using clustering method which combines both spectral and spatial features. Analysis of the feature extraction strategies before SVM classification is highly considered in this research work. The unsupervised feature extraction techniques are PCA, Minimum noise fraction, ICA, mixture tuned matched filtering (MTMF), Average MTMF are used for the comparative study. the proposed work shows the significant improvements in all the hyperspectral dataset [6]. [Xudong Kang et.al. 2014] proposed the feature extraction technique for hyperspectral data by using image fusion and recursive filtering method. The steps involved in this study are the following; initially, the hyperspectral images are divided into several subsets of hyperspectral bands where the image subsets are fused using average method. [Zexuan Zhu et.al. 2015] introduced the three dimensional Gabor Wavelet feature extraction using memetic framework (M3DGFE). The conditional mutual information with the concept of redundancy free relevance is used as a fitness evaluation function in M3DGFE to extract the potential features of the hyperspectral data. The pruning local search is adapted to eliminate the irrelevant features and redundancy features in hyperspectral data [7] [8]. [Renbo Luo.et.al. 2016] proposed the semi-supervised graph learning method for extracting the features of hyperspectral data. The graph is constructed by considering the aspects such as class maximization and acquires the local neighborhood information by fusion of unlabelled and labeled samples. Based on the neighborhood information the unlabeled samples and labeled information are connected. Estimating the mean distances between labeled and unlabeled samples for all the classes by considering nearest neighborhood information. The four different datasets are considered for this research study is Kennedy space centre, university of Pavia, Botswana and Pavia centre. LPP, NWFE, PCA, SDA, Semi-supervised local fisher discriminant analysis, and improved SLED are employed for this research work. The classifiers are employed to assess the feature extraction methods are randomforest, INN and SVM. [Heming Liang and Qi Li, 2016] proposed the deep Convolutional neural network with sparse representation to extract the deep learning features to improve the classification accuracy. There are three major folds in this methodology framework 1) By using CNN the spatial features are extracted 2) Extracted spatial features from CNN are represented through sparse to reduce the dimensionality 3) from the generated sparse coding the dictionaries are obtained and with this information; classification and predictions results are estimated. The AVIRIS Indian pines and Pavia University are used for this experimental study. The proposed algorithm is compared with the spectral spatial deep Convolutional neural network (SSDCNN) and SVM based classification such as deep features, extended morphological attribute profiles –EMAP. [Shaohui Mei et.al. 2017] proposed the Convolutional neural network with feature learning by exploiting spectral and spatial features. The main contributions of this research work are the following: 1) the five layer CNN is constructed to combine both spatial and spectral features 2) the new learning concept introduced the sensor specific spectral and spatial features and tune the other sensor images [9] [10] 11]. [Renlong Hang et.al. 2017] proposed the multi-scale robust matrix discriminant analysis (MRMDA) for hyperspectral feature extraction. In huge dimensional hyperspectral data, there might be more chance of corrupted data which influence the factors such as interruptions in sensor, noisy, error in calibration etc are some of the issues. The above issues with hyperspectral data referred to as corrupted data which reduces the performance of the MDA. But by adapting the prior information about neighborhood pixels (spatial features) the corrupted data will be eliminated. The RMDA algorithm addresses the above issues and experimental study conducted on Pavia university dataset, AVIRIS Indian pines and Kennedy space centre hyperspectral dataset are used. [Leyuan Fang et.al. 2018] proposed the feature extraction method based on the new representation of covariance matrix based on spectral and spatial information. The spectral correlations between various bands are evaluated using cosine based distance metric. The spatial and spectral features are extracted through the proposed local covariance matrix (LCMR) by deriving the concept of CMR. Extracted features are the input of SVM classifier with log Euclidean based kernel. [Junjun Jiang et.al. 2018] proposed an unsupervised superPCA feature extraction technique for reducing the huge hyperspectral dimensionality. The fundamental properties of the superPCA are 1) SuperPCA address the issue of homogenous regions in which different regions different projection will be carried out 2) Based on the superpixel segmentation the spatial features are acquired 3) the potential low dimensional features can also be extracted in noise 4) the unsupervised method is compared with supervised method. The entropy is measured in superpixel segmentation. [Bing Liu et.al. 2018] introduced the deep feature extraction Siemese Convolutional neural network (S-CNN) for improving the hyperspectral classification accuracy. The following are the major contributions; 1) construct the 2D Convolutional network to extract the spectral and spatial features of the given hyperspectral dataset by considering three fully connected layers with maxpooling function. The proposed method contains two CNN for classes' discrimination with shared weights and passed it to the loss function which uses the Euclidean distance for embedding feature vectors [12] [13] [14] [15]. From the decades of the literature study derived that Feature extraction techniques helps to enable the useful information from the huge dimensionality of hyperspectral data which also eliminates the redundancy features and noise. The computational efforts are reduced due to the low dimensional features. The supervised, semi-supervised and unsupervised feature extraction techniques are adapted as per the sample labels present in the hyperspectral data. LDA, SDA, NDA, QDC, etc. are some of the often used supervised learning feature extraction models, The PCA, ICA, and other sparse based algorithms are used to derive the spectral and spatial features. PCA is often used as the stat-of-art method which always shows the significant improvement with some limitations in estimating the covariance matrix. Also CNN is often experimented to extract the deep features of hyperspectral data. The SVM classifier is often adopted for assessing the performance of the different feature selection algorithms. Thus from the derived knowledge, the Novel PCA technique need to be developed for extracting the hyperspectral data by adding some weight function to enhance the classification performance in both pixel-wise and sub-pixel wise classification. The AVIRIS Indian Pines dataset will be used for experiment study. The major contributions of this paper involved two steps: i) The Novel Gravitiation Mass weighted method is introduced in the probability PCA (GMWProbPCA) to extract the low dimensional features from high dimensionality ii) the proposed algorithm is compared with LDA, PCA, and ProbPCA ans justified with the SVM classification results as accuracy. The remaining of this paper organized as follows section 2 and 3 discusses about the existing probability PCA and proposed methodology framework, section 4 discusses about the proposed GMWProbPCA algorithm , section 5 discusses about the experimental results and discussion and section 6 concludes the research work.

2 PROBABILITY PCA (PROBPCA)

In the traditional PCA model, the isotropic error arises due to the residual variances where the criterion required being equal and also the evaluation likelihood subspace is not fit into the observed principal subspaceω. Assume if there is a linear relationship from $\Re^h$ where h is the dimensional vector T and latent η dimensional vector with unobserved values f, is in the linear relationship in a factor analysis is represented as shown in the below equation (1).

$$T = A_f + \Delta + \varepsilon \text{ ....(1)}$$

h x η matrix A contains the two different set of variables where Δ referred as non zero mean. In addition ε is to specify the noise or Gaussian error. Also represented as $\forall\,(0, \partial^2 I)$ for ε

$$T|f \sim \forall\,(A_f + \Delta, \partial^2 I) \quad \text{....(2)}$$

The Gaussian latent variables also marginal distribution clearly and traditionally defined as f~ $\forall\,(0, I)$. The integration of observed values and Gaussian latent variables are represented as;

$$T \sim \forall(\partial, \mathbb{Z}) \text{ ......(3)}$$

The covariance matrix is specified as $\mathbb{Z}$=AA^T $+\partial^2 I$; the likelihood

$$\delta = \frac{1}{N}\sum_{n=1}^{N}(T_n - \Delta)(T_n - \Delta)^t \quad \text{......(4)}$$

Finally based on the Bayes rule the conditional probability is estimated in Gaussian distribution are the following

$$f|T \sim \forall\,(\exists^{-1}A^t(T - \Delta), \partial^2\exists^{-1} \text{ ...(5)}$$

The probabilistic PCA model is constructed based on the above shown equations. Basically the maximum likelihood estimators are used for factor analysis [182].

3 PROPOSED METHODOLOGY FRAMEWORK

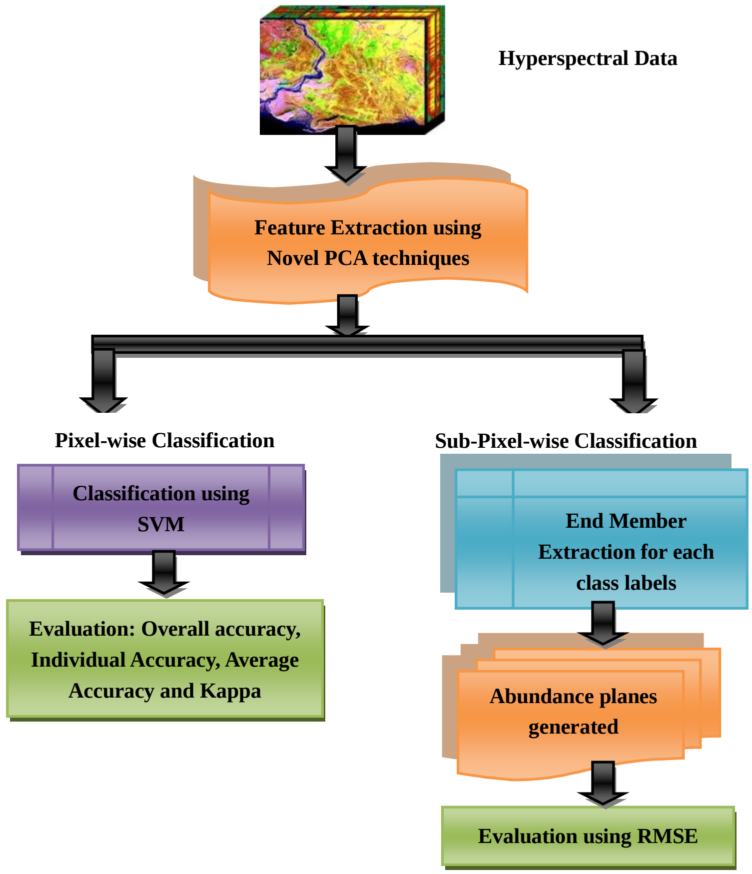


***Figure 1**. Proposed Methodology Framework for Hyperspectral Data Feature Extraction in Pixel-wise and Sub-Pixel-wise classification*

The proposed methodology framework constructed for the both pixel-wise and sub-pixel-wise classification shown in figure 1. The novel PCA feature extraction method is constructed by using the concept of gravitational force in estimating the mass is used to compute the weights by applying objective functions [18]. Once the features are extracted from the original hyperspectral data then pixel-wise and sub-pixel-wise classification will be carried to assess the performance of the proposed feature extraction method. In pixel-wise classification SVM technique is used to assess the proposed method by using evaluation metric such OA, AA, Individual accuracy and Kappa statistics. In the sub-pixel-wise classification FCLS and K-hype method are adopted to assess the performance of the proposed method. The evaluation metric is used for Sub-pixel wise classification is the RMSE for individual and the average RMSE.

3.1 PROPOSED NOVEL GRAVITATIONAL MASS WEIGHTED PROBPCA (GMWPROBCA)

The proposed algorithm derives the novel proposed Gravitational Mass Weighted Probability PCA (GMWProbPCA) technique is used to extract the features by involving the following steps. Since the hyperspectral data are in huge dimensional, first population is initialized for G (hyperspectral data). The properties for initializing the population are d dimension, upper value, lower value and number of bands presented in the hyperspectral data. The fitness function must be evaluated. The worst and best fits are to be estimated using minimizing and maximizing function. Estimate and update mass. Once the GMW weighted matrix is generated then pass through the probability PCA transformation function to acquire the low dimensional potential spectral feature component by estimating the covariance matrix and Eigen vectors. Based on the Eigen vectors the features are extracted and it is projected into the principal components. Here the ProbPCA adopt the isotropic Gaussian model with the conditional probability distribution is applied. Further the section will discuss about the other feature extraction techniques for the comparative study and analysis.

Algorithm: Proposed GMWProbPCA

Input:
HS data (G)
Output:
Extracted features
Projected Matrix
Steps:
Initialize the population G : d, u, l and N (where d= dimension, u=upper value, l= lower value, N= number of bands)
Estimate the fitness value by applying

$$f_i(j) = \sum_{i=1}^{N} -G_i \sin\left(\sqrt{|G_i|}\right)$$

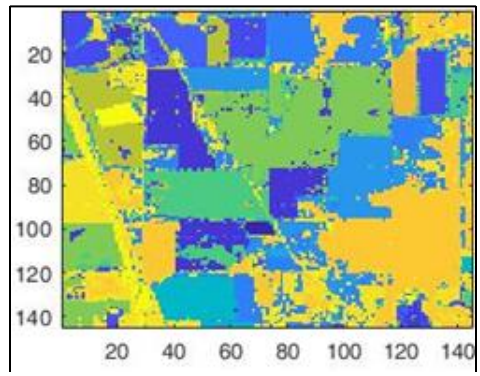
Calculate the minimization problem for both worst and best fit by applying $W(j) = \max_{i \in 1,2,\ldots..N} f_i(j)$ and $B(j) = \min_{i \in 1,2,\ldots..N} f_i(j)$.
Calculate the maximization problem for both worst and best fit by applying equations $W(j) = \min_{i \in 1,2,\ldots..N} f_i(j)$ and $B(j) = \max_{i \in 1,2,\ldots..N} f_i(j)$
Update the mass by applying $m_i(j) = \frac{f_i(j) - W(j)}{f_i(j) - B(j)}$ and $M_i(j) = \frac{m_i(j)}{\sum_k^N m_k(j)}$
The average of the mass Mi(j) is generated as a weight matrix W
The weighted matrix W is the input and with the Gaussian prior on the latent representation
Based on Gaussian latent representation of the matrix the conditional probability is evaluated
Calculate the covariance matrix
Calculate the eigenvectors of the covariance matrix
Extract the features
Project the data into the matrix

4 EXPERIMENTAL RESULTS AND DISCUSSION

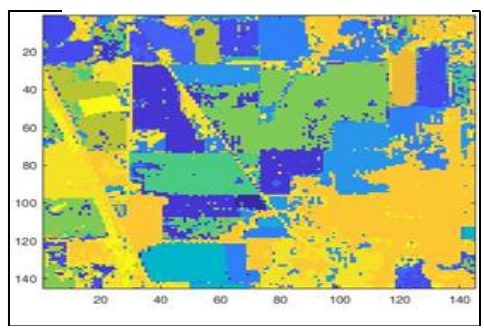
The feature extraction techniques are played crucial role in both pixel-wise and sub-pixel wise classification. The AVIRIS Indian pines dataset are used throughout this research work. In this phase 200 hundred original bands are considered for feature extraction process. The existing feature extraction algorithms supervised LDA, unsupervised PCA and Probability PCA are used for comparison with the performance of the proposed novel algorithm GMWProbPCA. The nonlinear unmixing models such as FCLS and Khype are employed for the sub-pixel-wise classification. The experimental dataset used for this research work is AVIRIS (Airborne Visible/ Infrared Imaging Spectrometer) Indian Pines hyperspectral dataset downloaded from Purdue University. This site is the agriculture area of Western Tippecanoe County, Indiana. The spectral reflectance wavelengths range of bands is 0.4- 2.5 X 10^{-6} m [19]. The structure of the AVIRIS Indian Pine site is in the form of 145X145X224. The water absorption bands are 104-108, 150-163, and 220 are removed. There are total 16 classes present in the scene with the ground truth reference.

4.1 Pixel-Wise Classification

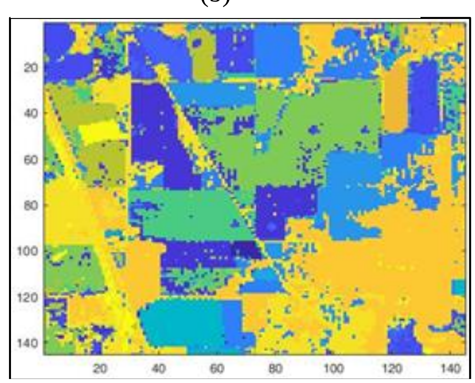
The SVM classifier technique is employed for assessing the performance of the proposed algorithm and existing models such as LDA, PCA and Probability PCA.



(a)



(b)



(c)

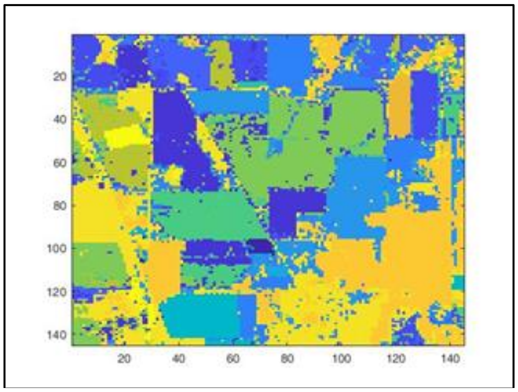


Figure 2. *The OA results for a) LDA b) PCA c) ProbPCA* d) GMWProbPCA

The proposed novel gravitational methods are compared with existing state-of-art- feature extraction techniques LDA, PCA, and ProbPCA using AA, OA and Kappa evaluation measures. The figure 2 shows the overall accuracy obtained from the SVM classifiers with different feature extraction models.

Table 1 *SVM Classification Results for the Existing and Proposed Feature Extraction Algorithms*

Classes	*Supervised*	*Unsupervised*		*Proposed Technique*
	LDA	*PCA*	*ProbPCA*	*GMWProb-PCA*
1	*84.09%*	*93.18%*	*79.55%*	*93.18%*
2	*82.01%*	*84.68%*	*85.98%*	*89.52%*
3	*91.33%*	*89.03%*	*88.14%*	*87.86%*
4	*94.02%*	*88.59%*	*93.48%*	*94.25%*
5	*89.93%*	*95.97%*	*93.06%*	*94.97%*
6	*97.70%*	*96.56%*	*97.56%*	*97.38%*
7	*81.25%*	*87.50%*	*87.50%*	*93.75%*
8	*99.32%*	*98.63%*	*98.63%*	*99.30%*
9	*100.00%*	*100.00%*	*100.00%*	*100.00%*
10	*88.24%*	*84.53%*	*83.66%*	*89.54%*
11	*83.87%*	*82.05%*	*88.88%*	*82.18%*
12	*90.60%*	*89.18%*	*90.96%*	*90.07%*
13	*98.77%*	*97.53%*	*99.38%*	*97.37%*
14	*94.61%*	*96.70%*	*96.30%*	*92.14%*
15	*93.03%*	*96.36%*	*96.67%*	*98.44%*
16	*100.00%*	*95.56%*	*100.00%*	*100.00%*
Average	*91.80%*	*92.25%*	*92.48%*	*93.75%*
OA	*89.21%*	*88.97%*	*90.79%*	*89.74%*
Kappa	*87.70%*	*87.43%*	*89.46%*	*88.30%*

The table 1 shows the SVM classification results for the existing and proposed feat[illegible]action algorithms. The less training samples are avail[illegible] (d) [illegible]he few classes are Alfalfa, Grass-pasture-mowed, Oats, and stone steel towers. For the other classes more than hundred training samples are available. From the above experimental study, even with the less training samples all the classes of GMWProbPCA achieved AA of 93.75%, OA is 89.74% and Kappa statistics value is 88.30%. The figure 3 shows the comparative analysis of the proposed, supervised and unsupervised models. The OA, Kappa and AA are shown as the performance measures and the novel proposed method GMWProbPCA outperforms well when compared with the unsupervised and supervised existing models such as LDA, PCA and ProbPCA. From the above context identified that OA of proposed algorithm shows the significant improvement in the pixel-wise classification accuracy for AVIRIS Indian Pine hyperspectral data.

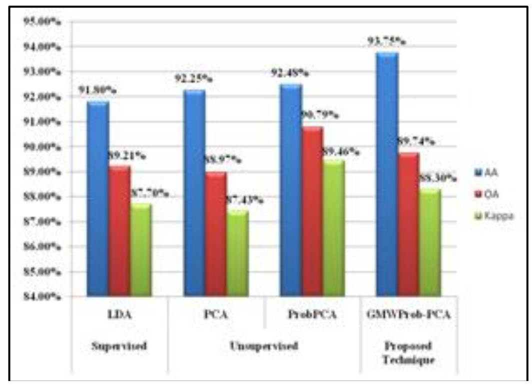


Figure 3 *Comparative Analysis of Proposed and Existing Feature Extraction Algorithms*

4.2 Sub-Pixel-Wise Classification

This research study considers the original features of 200 bands without feature extraction for assessing the classification performance analysis. GMWProbPCA employed FCLS and Khype. The nonlinear unmixing models used for this study are K-Hype and FCLS (Fully Constraint Linear Squared) methods to assess the performance of the hyperspectral feature extraction. The figure 3 and 4 shows the clear visualization of abundance planes generated for individual classes present in the AVIRIS Indian pine site scene. From the above context identified that GMWProbPCA shows the significant results in both nonlinear unmixing models FCLS and Khype respectively. The RMSE is considered as the evaluation measure for sub-pixel-wise classification. The table 2 shows the RMSE values for individual classes and the average RMSE rate is less in GMWProbPCA method when compared with the other models. But the both proposed algorithm outperforms upon existing supervised LDA and unsupervised PCA and ProbPCA. The RMSE for the GMWProbCA feature extraction technique shows the less error rate in both FCLS and Khype are 0.044 and 0.03 respectively.

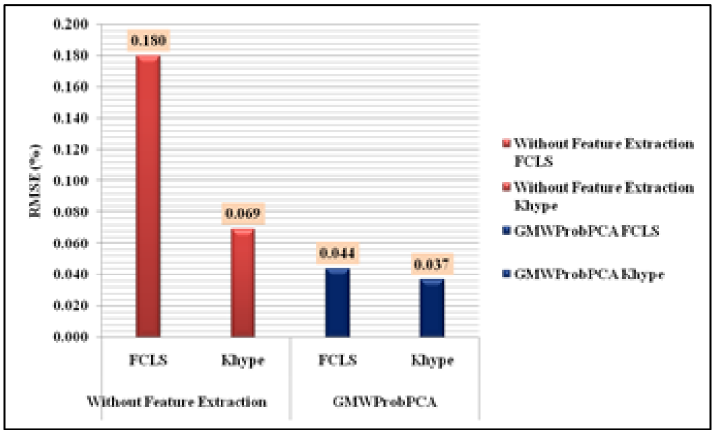
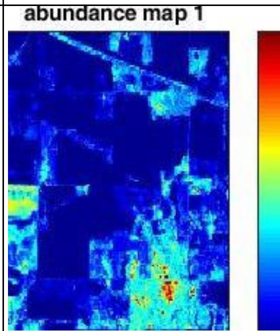
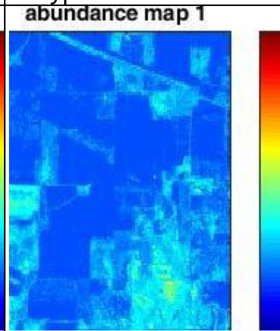
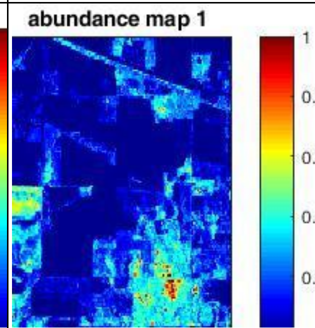
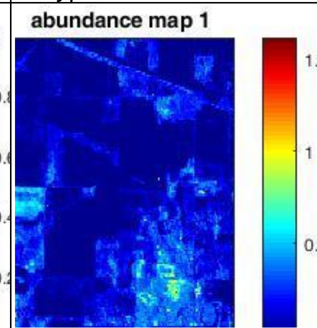
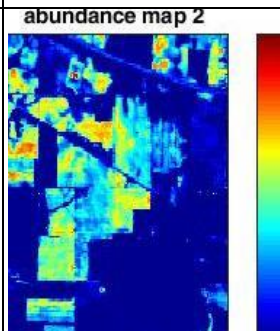
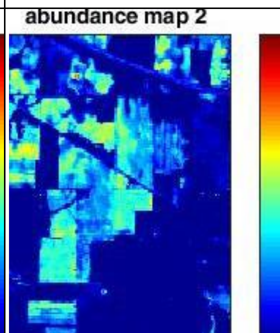
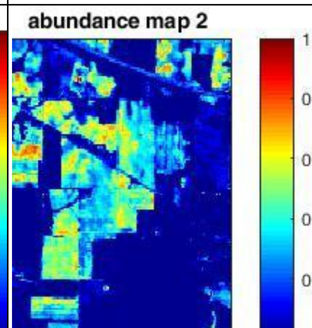
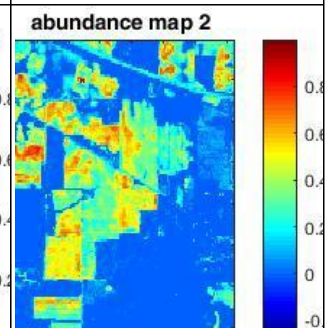
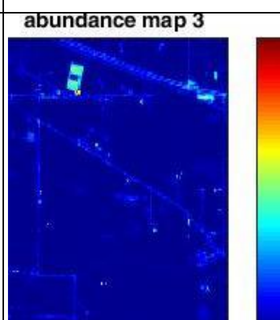
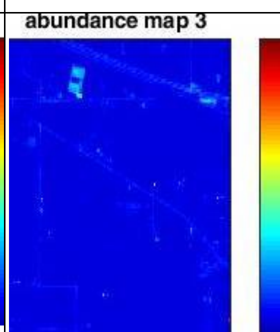
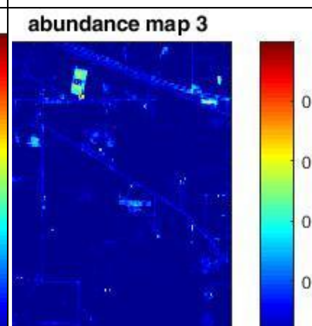
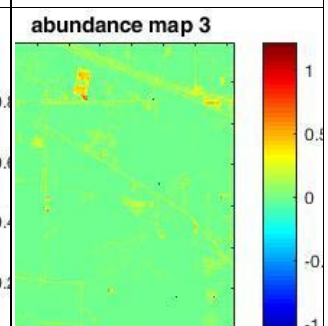
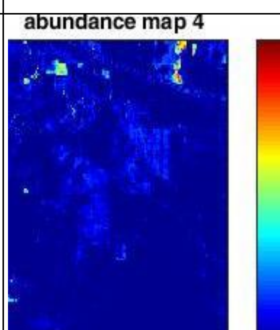
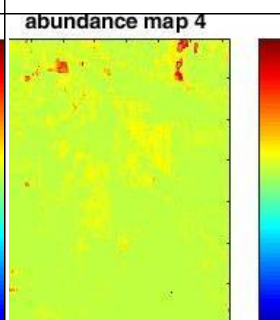
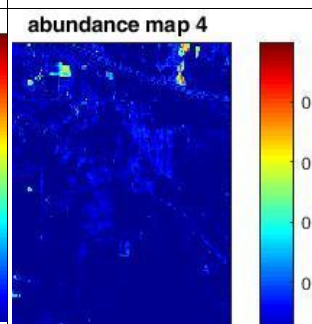
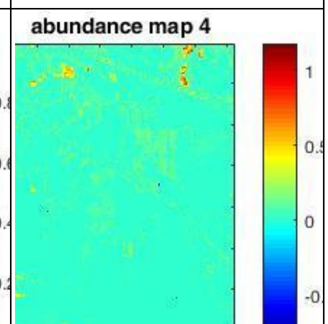
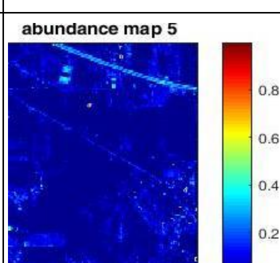
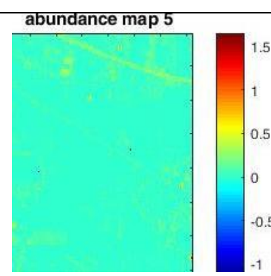
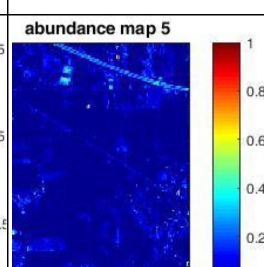
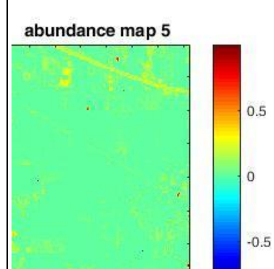
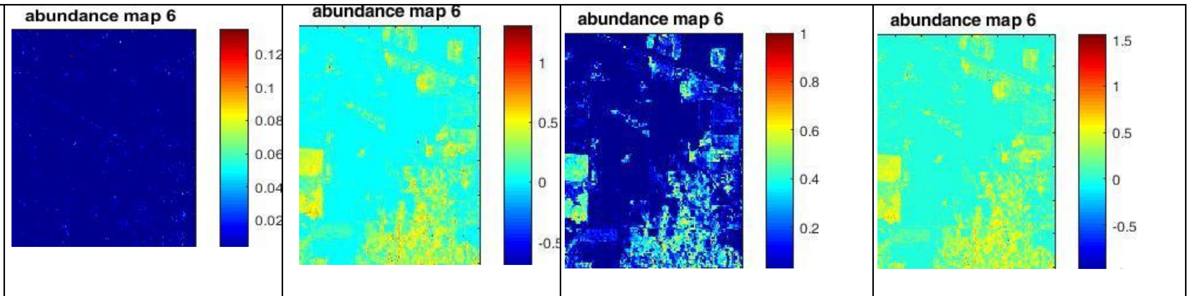
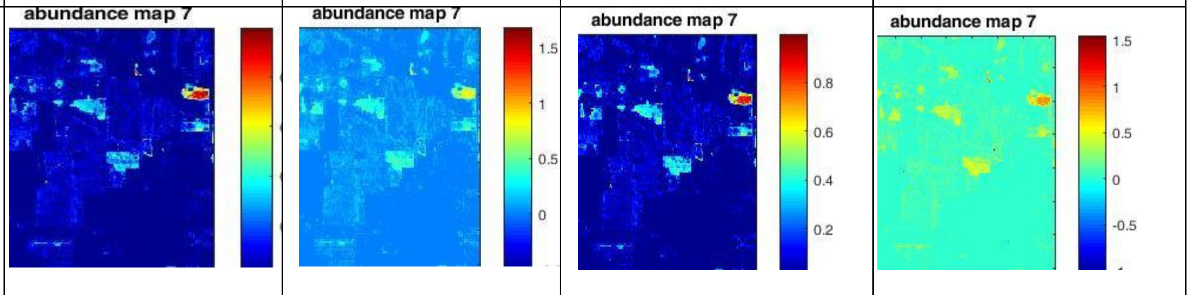
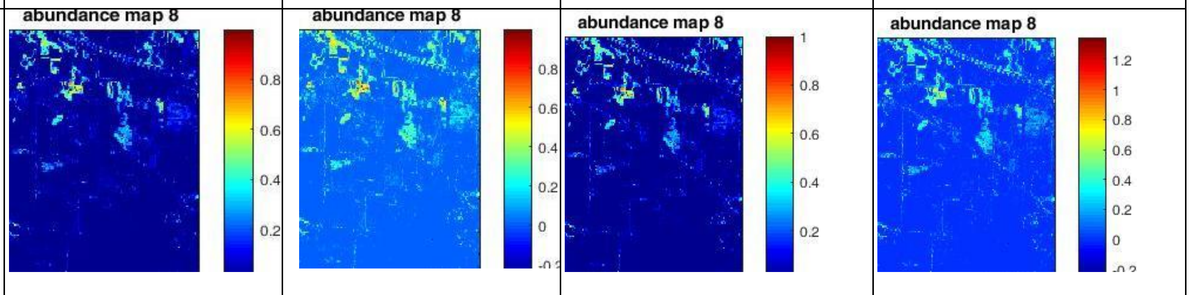
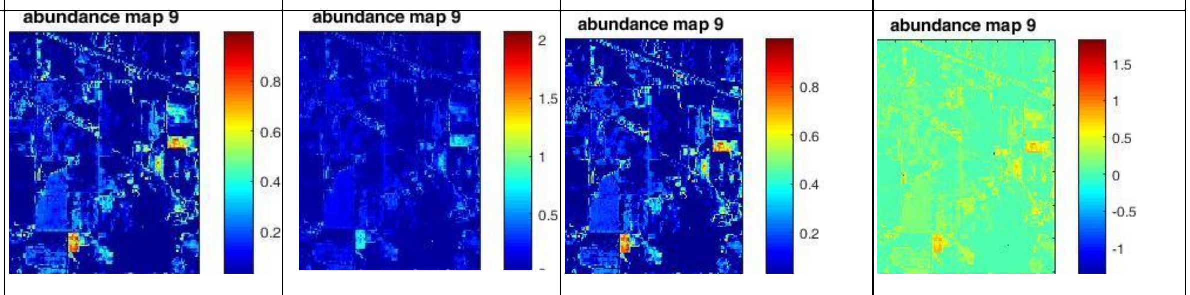
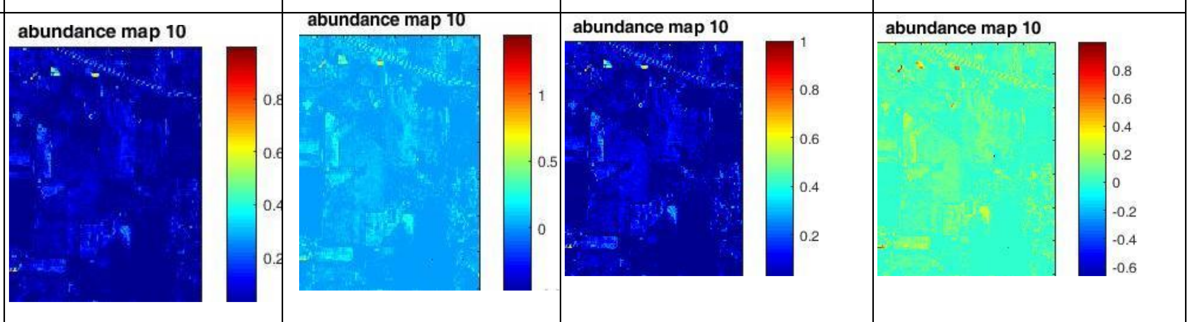
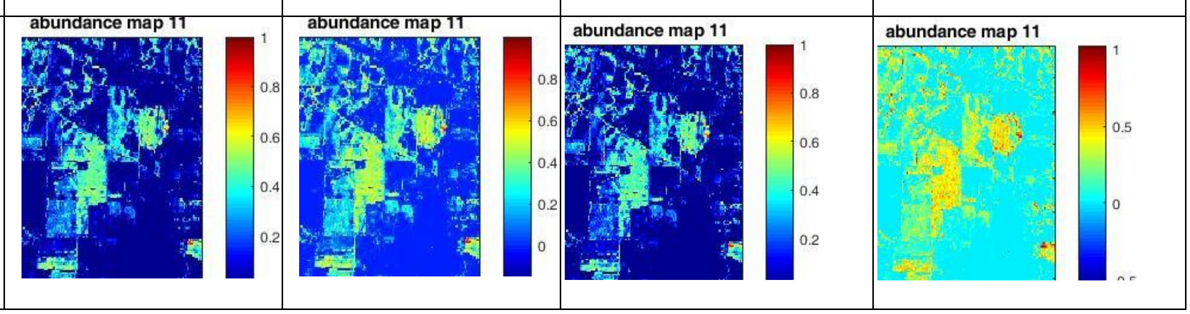


Figure 3 *Comparative Analysis of FCLS and Khype Nonlinear Unmixing Models without Feature Extraction and with Proposed GMWProbPCA*

End Members	*Without Feature Extraction*		*With Proposed GMWProbPCA Feature Extraction*	
	FCLS	*Khype*	*FCLS*	*Khype*
Alfalfa	*0.000*	*0.000*	*0.000*	*0.000*
Corn-no till	*0.000*	*0.000*	*0.000*	*0.000*
Corn-min till	*0.000*	*0.000*	*0.000*	*0.000*
Corn	*0.000*	*0.000*	*0.000*	*0.009*
Grass Pasture	*0.000*	*0.000*	*0.000*	*0.001*
Grass-Trees	*0.000*	*0.000*	*0.001*	*0.002*
Grass-Pasture-Mowed	*0.000*	*0.126*	*0.001*	*0.034*
Hay-Windrowed	*0.000*	*0.141*	*0.009*	*0.059*
Oats	*0.140*	*0.163*	*0.009*	*0.007*
Soybean-no till	*0.166*	*0.164*	*0.012*	*0.085*
Soybean-min till	*0.271*	*0.044*	*0.045*	*0.010*
Soybean-clean	*0.282*	*0.058*	*0.086*	*0.014*
Wheat	*0.392*	*0.073*	*0.095*	*0.203*
Woods	*0.452*	*0.087*	*0.096*	*0.027*
Buildings-Grass-Trees-Drives	*0.535*	*0.116*	*0.164*	*0.054*
Stone steel -Towers	*0.637*	*0.131*	*0.183*	*0.081*
Average RMSE	*0.180*	*0.069*	*0.044*	*0.037*

Table 2. The Individual Endmembers RMSE values for with proposed and without feature extraction

Endmember	Without Feature Extraction		With proposed GMWProbPCA Feature Extraction	
	FCLS	Khype	FCLS	Khype
Alfalfa				
Corn-no till				
Corn-min till				
Corn				
Grass Pasture				

Grass-Trees	
Grass-Pasture-Mowed	
Hay-Windrowed	
Oats	
Soybean-no till	
Soybean-min till	

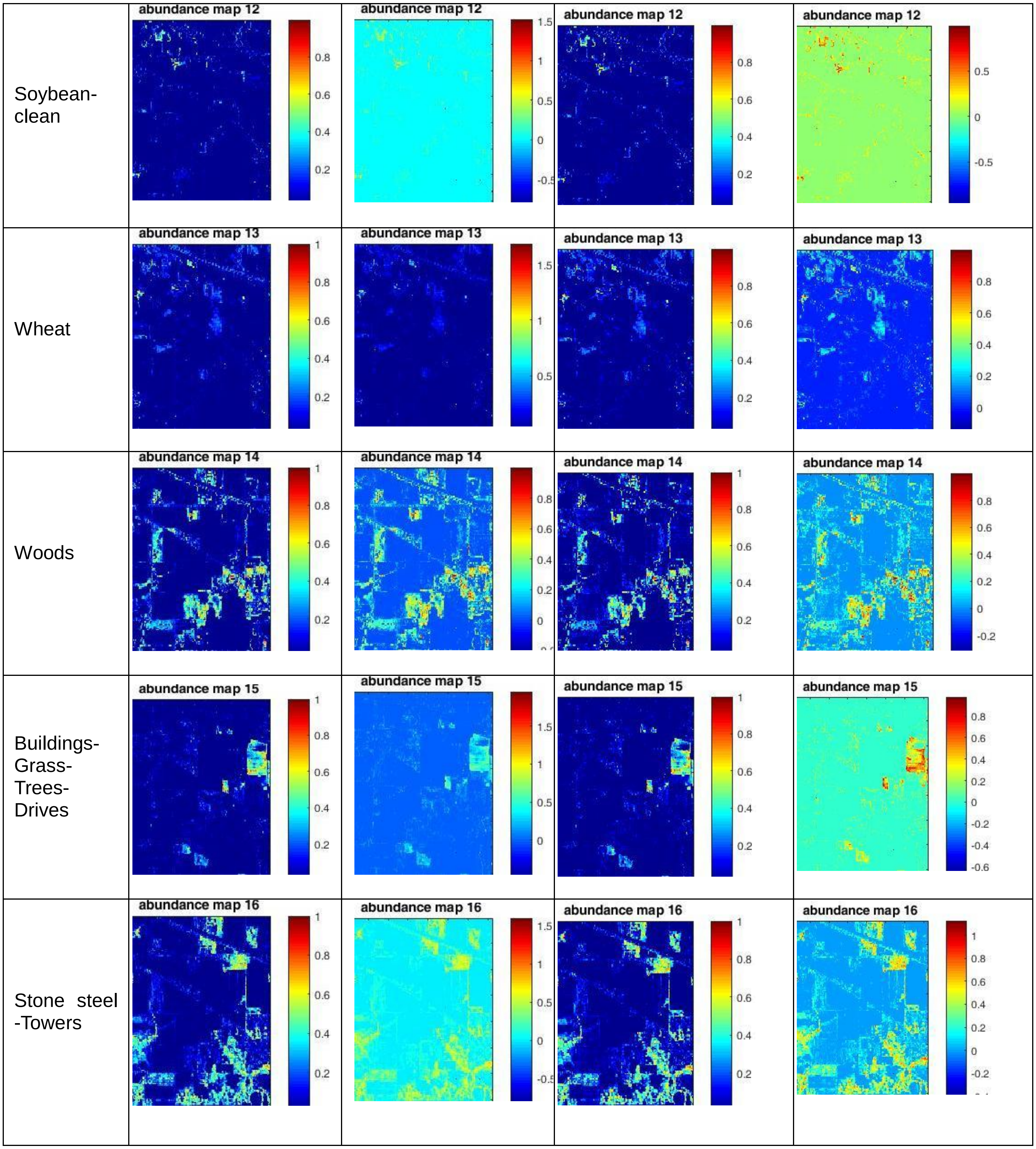


***Figure 4**. Abundance planes for Nonlinear unmixing models (FCLS and Khype) performance for the novel proposed feature extraction algorithm GMWProbPCA*

5 CONCLUSION AND SUMMARY

The feature extraction techniques are basically partitioned into three types: unsupervised, supervised and semi-supervised. An unsupervised feature extraction method doesn't require known labels and also the class discrimination information. Principal component analysis (PCA), Independent Component Analysis (ICA) and Maximum Noise Fraction (MNF) are often used in HS data as unsupervised feature extraction techniques. Over the decades, supervised and unsupervised techniques are widely used in extraction of the features of HS data. Linear Discriminant Analysis (LDA), General Discriminant analysis (GDA), Median-mean line based discriminant analysis (MMLDA), nonparametric weighted feature extraction (NWFE), etc. are some of the supervised feature extraction techniques often been in comparison with other techniques and also widely used methods. This research study employed LDA, PCA and Probabilistic PCA (ProbPCA) for further experimental work. The weight matrix often scores the high performance based on the extraction of heavier mass. Hence, this proposed study considered the Gravitational Mass Weighted (GMW) which is inspired from the Gravitational Search Algorithm. Fundamentally, the objects with Gravitational Force (GF) contain high mutual attraction. The novel proposed method GMWProbPCA are experimented in both pixel-wise and sub-pixel-wise classification. In pixel-wise classification GMWProbPCA achieves high accuracy of 93.75% when compared with the other techniques. In sub-pixel-wise classification the GMWPCA technique achieves slightest RMSE error of 0.044% and 0.037% for FCLS and Khype respectively. Overall the proposed novel methods achieve high performance in accuracy when compared with the existing methods.

6 ACKNOWLEDGMENTS

This research study was supported under the University Research Fellowship (URF NO: C2/2016/2170) from the Bharathiar University, Coimbatore, India.

7 REFERENCES

[1] Prasad, S., & Bruce, L. M. (2008). Limitations of Principal Components Analysis for Hyperspectral Target Recognition. IEEE Geoscience and Remote Sensing Letters, 5(4), 625–629. doi:10.1109/lgrs.2008.2001282

[2] Mendenhall, M. J., & Merenyi, E. (2008). Relevance-Based Feature Extraction for Hyperspectral Images. IEEE Transactions on Neural Networks, 19(4), 658–672. doi:10.1109/tnn.2007.914156

[3] Mojaradi, B., Abrishami-Moghaddam, H., Zoej, M. J. V., & Duin, R. P. W. (2009). Dimensionality Reduction of Hyperspectral Data via Spectral Feature Extraction. IEEE Transactions on Geoscience and Remote Sensing, 47(7), 2091–2105. doi:10.1109/tgrs.2008.2010346

[4] Jinn-Min Yang, Pao-Ta Yu, & Bor-Chen Kuo. (2010). A Nonparametric Feature Extraction and Its Application to Nearest Neighbor Classification for Hyperspectral Image Data. IEEE Transactions on Geoscience and Remote Sensing, 48(3), 1279–1293. doi:10.1109/tgrs.2009.2031812

[5] Huang, H.-Y., & Kuo, B.-C. (2010). Double Nearest Proportion Feature Extraction for Hyperspectral-Image Classification. IEEE Transactions on Geoscience and Remote Sensing. doi:10.1109/tgrs.2010.2058580

[6] Dopido, I., Villa, A., Plaza, A., & Gamba, P. (2012). A Quantitative and Comparative Assessment of Unmixing-Based Feature Extraction Techniques for Hyperspectral Image Classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5(2), 421–435. doi:10.1109/jstars.2011.2176721

[7] Kang, X., Li, S., & Benediktsson, J. A. (2014). Feature Extraction of Hyperspectral Images With Image Fusion and Recursive Filtering. IEEE Transactions on Geoscience and Remote Sensing, 52(6), 3742–3752. doi:10.1109/tgrs.2013.2275613

[8] Zhu, Z., Jia, S., He, S., Sun, Y., Ji, Z., & Shen, L. (2015). Three-dimensional Gabor feature extraction for hyperspectral imagery classification using a memetic framework. Information Sciences, 298, 274–287. doi:10.1016/j.ins.2014.11.045

[9] Luo, R., Liao, W., Huang, X., Pi, Y., & Philips, W. (2016). Feature Extraction of Hyperspectral Images With Semisupervised Graph Learning. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9(9), 4389–4399. doi:10.1109/jstars.2016.2522564

[10] Liang, Heming, and Qi Li. "Hyperspectral imagery classification using sparse representations of convolutional neural network features." Remote Sensing 8.2 (2016): 99.

[11] Mei, S., Ji, J., Hou, J., Li, X., & Du, Q. (2017). Learning Sensor-Specific Spatial-Spectral Features of Hyperspectral Images via Convolutional Neural Networks. IEEE Transactions on Geoscience and Remote Sensing, 55(8), 4520–4533. doi:10.1109/tgrs.2017.2693346

[12] Hang, R., Liu, Q., Sun, Y., Yuan, X., Pei, H., Plaza, J., & Plaza, A. (2017). Robust Matrix Discriminative Analysis for Feature Extraction from Hyperspectral Images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10(5), 2002–2011. doi:10.1109/jstars.2017.2658948

[13] Fang, L., He, N., Li, S., Plaza, A. J., & Plaza, J. (2018). A New Spatial–Spectral Feature Extraction Method for Hyperspectral Images Using Local Covariance Matrix Representation. IEEE Transactions on Geoscience and Remote Sensing, 56(6), 3534–3546. doi:10.1109/tgrs.2018.2801387

[14] Jiang, J., Ma, J., Chen, C., Wang, Z., Cai, Z., & Wang, L. (2018). SuperPCA: A Superpixelwise PCA Approach for Unsupervised Feature Extraction of Hyperspectral Imagery. IEEE Transactions on Geoscience and Remote Sensing, 56(8), 4581–4593. doi:10.1109/tgrs.2018.2828029

[15] Liu, B., Yu, X., Zhang, P., Yu, A., Fu, Q., & Wei, X. (2018). Supervised Deep Feature Extraction for Hyperspectral Image Classification. IEEE Transactions on Geoscience and Remote Sensing, 56(4), 1909–1921. doi:10.1109/tgrs.2017.2769673

[16] Kang, X., Xiang, X., Li, S., & Benediktsson, J. A. (2017). PCA-Based Edge-Preserving Features for Hyperspectral Image Classification. IEEE Transactions on Geoscience and Remote Sensing, 55(12), 7140–7151. doi:10.1109/tgrs.2017.2743102

[17] Xing, C., Ma, L., & Yang, X. (2016). Stacked Denoise

Autoencoder Based Feature Extraction and Classification for Hyperspectral Images. Journal of Sensors, 2016, 1–10. doi:10.1155/2016/3632943

[18] Rashedi, Esmat, Hossein Nezamabadi-Pour, and Saeid Saryazdi. "GSA: a gravitational search algorithm." Information sciences 179.13 (2009): 2232-2248.

Baumgardner, M. F., Biehl, L. L., Landgrebe, D. A. (2015). 220 Band AVIRIS Hyperspectral Image Data Set: June 12, 1992 Indian Pine Test Site 3. Purdue University Research Repository. doi:10.4231/R7RX991C