



A new per-field classification method using mixture discriminant analysis

Nazif Çalış & Hamza Erol

To cite this article: Nazif Çalış & Hamza Erol (2012) A new per-field classification method using mixture discriminant analysis, Journal of Applied Statistics, 39:10, 2129-2140, DOI: 10.1080/02664763.2012.702263

To link to this article: <https://doi.org/10.1080/02664763.2012.702263>



Published online: 02 Jul 2012.



Submit your article to this journal [↗](#)



Article views: 219



View related articles [↗](#)



Citing articles: 1 View citing articles [↗](#)

A new per-field classification method using mixture discriminant analysis

Nazif Çalış^{a*} and Hamza Erol^b

^aDepartment of Statistics, Çukurova University, Adana, Turkey; ^bDepartment of Software Engineering, Abdullah Gül University, Kayseri, Turkey

(Received 17 December 2010; final version received 9 June 2012)

In this study, a new per-field classification method is proposed for supervised classification of remotely sensed multispectral image data of an agricultural area using Gaussian mixture discriminant analysis (MDA). For the proposed per-field classification method, multivariate Gaussian mixture models constructed for control and test fields can have fixed or different number of components and each component can have different or common covariance matrix structure. The discrimination function and the decision rule of this method are established according to the average Bhattacharyya distance and the minimum values of the average Bhattacharyya distances, respectively. The proposed per-field classification method is analyzed for different structures of a covariance matrix with fixed and different number of components. Also, we classify the remotely sensed multispectral image data using the per-pixel classification method based on Gaussian MDA.

Keywords: average Bhattacharyya distance; Gaussian mixture discriminant analysis; per-field classification; per-pixel classification; supervised classification

1. Introduction

Discriminant analysis can be used as a model-based classification method for per-field classification of remotely sensed multispectral image data of an agricultural area. In this type of classification method, a cluster of observations taken from unknown groups or test groups are assigned to one of the known pre-defined groups or training groups, indicating that it is a supervised classification method. Hastie and Tibshirani [6] proposed the mixture discriminant analysis (MDA) method to model training data set and to classify test data effectively when the structure of the training groups is heterogeneous. They used a mixture of multivariate normal distribution models with a fixed number of components for modeling training data sets in classification applications. Each component in the mixture of multivariate normal distribution models corresponds to a different sub-group in the heterogeneous data [5]. Hastie and Tibshirani [6] also used a common covariance matrix for each component in the mixture of multivariate normal distribution models of all

*Corresponding author. Email: ncalis@cu.edu.tr

training groups. The discrimination function of their per-pixel classification method was defined by the maximum posterior probabilities.

Ju *et al.* [7] estimated land-cover fractions using MDA according to the corresponding posterior probabilities of per-pixel classification. Ju *et al.* [7] made a small change to the original MDA of Hastie and Tibshirani [6] for inferring the land-cover fractions within forest stands from Landsat Thematic Mapper images. They used a different covariance matrix for all the control groups but a common one for each component of the control groups. The discrimination function of their per-pixel classification method was defined by the maximum posterior probabilities as in [6].

In addition to these studies, Win *et al.* [10] classified the texture features of images in sequences using a mixture of multivariate normal distribution models with four components. They used the average Bhattacharyya distance as the discrimination function and the minimum values of the average Bhattacharyya distances as the decision rule for their classification method.

In this paper, a new per-field classification method is proposed for supervised classification of remotely sensed multispectral image data of an agricultural area using MDA. The discrimination function of the proposed per-field supervised classification method is defined by the average Bhattacharyya distance computed between the mixtures of multivariate normal distribution models constructed for the test and control fields. The decision rule of the proposed per-field classification method is established according to the minimum value of the average Bhattacharyya distance. The mixture of multivariate normal distribution models constructed for the test and control fields can have fixed or different number of components and each component can have different or common covariance matrix in this method. So, we classify multispectral image data using the proposed per-field classification method under different structures of a covariance matrix with fixed and different number of components. Also, the multispectral image data are classified on per-pixel basis by applying the classification methods proposed by Hastie and Tibshirani [6] and Ju *et al.* [7]. For the per-pixel classification of the multispectral image data, different structures of a covariance matrix with fixed and different number of components are used as in the proposed per-field classification method.

2. Method

The proposed per-field classification method was applied to classify the remotely sensed multispectral image data of an agricultural area. A Landsat Thematic Mapper image of an agricultural area in the Seyhan plain ($\approx 37^\circ\text{N}$, 36°E) in Adana having a size of 198×200 (in total 39,600) pixels was used as the multispectral image data [4]. The data were collected on 27 March 1992 (Path 175, Row 34). Landsat Thematic Mapper bands 3, 4 and 5 were used for the proposed per-field classification method since spectral radiance data from these three bands can be used to infer the properties related to pigment absorption, leaf structure and canopy leaf water content, respectively [4]. Landsat Thematic Mapper image and the fields in the Landsat Thematic Mapper image are shown in Figure 1(a) and (b).

There are totally 269 fields in the agricultural area used for the proposed per-field classification method based on MDA. These fields are coded with field codes from F001 to F269 on the Government Irrigation Department (Devlet Su İşleri – DSI) parcelization map. There are five land cover types and totally 24 sub-species of these cover types in the agricultural area. These are wheat, potato, vegetable garden, citrus and bare soil. So, there are totally 24 control fields in this area. These control fields, acquired through a land cover survey, are coded with control codes from CF001 to CF024 on the Government Irrigation Department parcelization map [4]. The remaining 245 test parcels or fields are coded with test codes from TF001 to TF245 on the Government Irrigation Department parcelization map.

When we applied the proposed per-field classification method, first, multivariate normal or mixture of normal distribution models were constructed for all fields in the remotely sensed

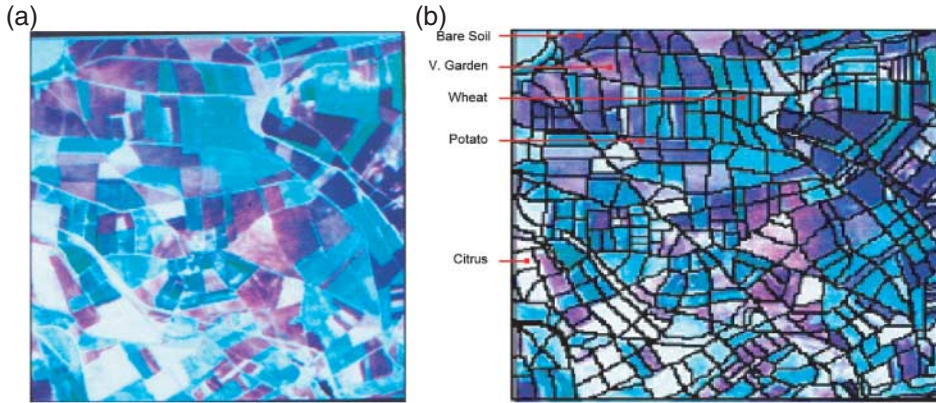


Figure 1. The remotely sensed multispectral image data of an agricultural area for the proposed per-field classification method using MDA: (a) Landsat Thematic Mapper image and (b) the parcels or fields in the Landsat Thematic Mapper image.

multispectral image data according to the fields having a homogeneous or heterogeneous data structure. Then, the discrimination function values, and so the average Bhattacharyya distances, were calculated between these multivariate normal or mixture of normal distribution models constructed for each of the 245 test fields and 24 control fields. After that, minimum values of the average Bhattacharyya distances were used as the decision rule to assign each of the 245 test fields to one of the 24 control fields. An error matrix was constructed at the end of the proposed per-field classification method to compute the classification accuracies.

2.1 Model description

In the proposed per-field classification method based on MDA, a multivariate Gaussian mixture model assumption is made for p -dimensional data with n observations and G groups. Thus, the probability density function $f(x; \theta)$ for the multispectral image data has the following form:

$$f(x_i; \theta) = \sum_{j=1}^G \pi_j f_j(x_i; \Psi_j), \quad i = 1, 2, \dots, n, \tag{1}$$

where π_j is the mixture proportion of the j th component for $j = 1, 2, \dots, G$ such that $0 < \pi_j < 1$ and $\sum_{j=1}^G \pi_j = 1$. In Equation (1), $f_j(x_i; \Psi_j)$ is the component probability density function which has a multivariate normal distribution with mean μ_j and covariance matrix Σ_j for the j th component and $f_j(x_i; \Psi_j)$ is given as

$$f_j(x_i; \Psi_j) = \frac{1}{(2\pi)^{p/2} |\Sigma_j|^{1/2}} \exp \left\{ -\frac{1}{2} (x_i - \mu_j)^T \Sigma_j^{-1} (x_i - \mu_j) \right\}, \tag{2}$$

where $\Psi_j = (\mu_j, \Sigma_j)$ denotes the parameter vectors in component densities and $\theta = ((\pi_1, \mu_1, \Sigma_1), \dots, (\pi_G, \mu_G, \Sigma_G))$ denotes all the parameters of the mixture of multivariate normal distribution models in the parameter space Ω . In Equations (1) and (2), $x_i \in \mathbf{x} : n \times p$ is a random variable which denotes the spectral brightness band (pixel) values for 3, 4 and 5 bands.

For the MDA, suppose we have n_j training observations from population j for $j = 1, \dots, G$ and we divide each class j into R_j artificial subclasses denoted by c_{jr} for $r = 1, \dots, R_j$. Denote R by $\sum_j R_j$ and $n = \sum_j n_j$. The observations in the r th subgroup of the j th group have a multivariate

normal distribution with $\boldsymbol{\mu}_{jr}$ and $\boldsymbol{\Sigma}_{jr}$ parameters in the model-based cluster analysis or mixture cluster analysis [2]. Then, the mixture density function $f_j(x)$ for class j is

$$f_j(x) = \sum_{r=1}^{R_j} \pi_{jr} \frac{1}{(2\pi)^{\frac{p}{2}} |\boldsymbol{\Sigma}_{jr}|^{\frac{1}{2}}} \exp \left[-\frac{D^2(x, \boldsymbol{\mu}_{jr})}{2} \right], \quad j = 1, 2, \dots, G, \quad (3)$$

where π_{jr} is the unknown mixture proportion of the r th subgroup of the j th group such that $\sum_r \pi_{jr} = \pi_j, j = 1, 2, \dots, G$, and $D(\mathbf{x}, \boldsymbol{\mu}_{jr}) = \sqrt{(\mathbf{x} - \boldsymbol{\mu}_{jr})^T \boldsymbol{\Sigma}_{jr}^{-1} (\mathbf{x} - \boldsymbol{\mu}_{jr})}$ is the Mahalanobis distance between \mathbf{x} and $\boldsymbol{\mu}_{jr}$.

An observation is classified into class j , which has the highest posterior probability, based on Bayes rule, denoted by

$$P(G = j | X = \mathbf{x}) \propto \pi_j P(\mathbf{x} | j) = \pi_j \sum_{r=1}^{R_j} \pi_{jr} |\boldsymbol{\Sigma}_{jr}|^{-1/2} \exp \left[-\frac{D^2(\mathbf{x}, \boldsymbol{\mu}_{jr})}{2} \right], \quad j = 1, 2, \dots, G, \quad (4)$$

where π_j is the prior probability for class j and can be easily estimated using the data. In this classification, the discrimination rules depend on the unknown parameters π_{jr} , $\boldsymbol{\mu}_{jr}$ and $\boldsymbol{\Sigma}_{jr}$, which are to be estimated from the training data using the expectation–maximization (EM) algorithm [3,8]. Let $z_i^{(jr)}$ denote the component label vectors for the data set $x_i^{(jr)} \in \{x_1, x_2, \dots, x_{n_j}\}$. Then, $z_i^{(jr)}$ is given as follows:

$$z_i^{(jr)} = \begin{cases} 1, & x_i \in c_{jr}, \\ 0, & x_i \notin c_{jr}, \end{cases} \quad (5)$$

and they are used in the estimation of the unknown parameters in the mixture of multivariate normal distribution models. The unknown parameters are estimated using the EM algorithm from the training data set. The EM equations for the unknown parameters are given in Equations (6)–(8):

$$\hat{\boldsymbol{\mu}}_{jr} = \frac{1}{n_{jr}} \sum_{i=1}^{n_{jr}} x_i^{(jr)} = \frac{\sum_{i=1}^{n_j} z_i^{(jr)} x_i}{\sum_{i=1}^{n_j} z_i^{(jr)}}, \quad (6)$$

$$\hat{\boldsymbol{\Sigma}}_{jr} = \frac{1}{n_{jr}} \sum_{i=1}^{n_{jr}} (x_i^{(jr)} - \boldsymbol{\mu}_{jr})^T (x_i^{(jr)} - \boldsymbol{\mu}_{jr}) = \frac{\sum_{i=1}^{n_j} z_i^{(jr)} (x_i - \boldsymbol{\mu}_{jr})^T (x_i - \boldsymbol{\mu}_{jr})}{\sum_{i=1}^{n_j} z_i^{(jr)}}, \quad (7)$$

$$\hat{\pi}_{jr} = \frac{n_{jr}}{n_j} = \frac{\sum_{i=1}^{n_j} z_i^{(jr)}}{n_j}, \quad (8)$$

where the estimation of parameters π_{jr} , $\boldsymbol{\mu}_{jr}$ and $\boldsymbol{\Sigma}_{jr}$ in Equations (6)–(8) is used in the EM steps as follows:

E step: $(\boldsymbol{\mu}_{jr}^{(t)}, \boldsymbol{\Sigma}_{jr}^{(t)})$ is the estimation of the $(\boldsymbol{\mu}_{jr}, \boldsymbol{\Sigma}_{jr})$ parameters in step t . The expected value $E[z_i^{(jr)} | \boldsymbol{\mu}_{jr}^{(t)}, \boldsymbol{\Sigma}_{jr}^{(t)}]$ of the label vector $z_i^{(jr)}$ is computed when $(\boldsymbol{\mu}_{jr}^{(t)}, \boldsymbol{\Sigma}_{jr}^{(t)})$ is given.

M step: If the expected value $E[z_i^{(jr)} | \boldsymbol{\mu}_{jr}^{(t)}, \boldsymbol{\Sigma}_{jr}^{(t)}]$ is taken for $z_i^{(jr)}$, $(\pi_{jr}^{(t+1)}, \boldsymbol{\mu}_{jr}^{(t+1)}, \boldsymbol{\Sigma}_{jr}^{(t+1)})$ new maximum-likelihood estimation of $(\pi_{jr}^{(t)}, \boldsymbol{\mu}_{jr}^{(t)}, \boldsymbol{\Sigma}_{jr}^{(t)})$ is calculated in step $t + 1$.

2.2 Discrimination function and decision rule for the proposed per-field classification method

In this study, the average Bhattacharyya distance was considered as the discrimination function of the proposed per-field supervised classification method [10]. Before describing the average

Bhattacharyya distance, the Bhattacharyya distance between two multivariate normal distribution models, denoted by BD, is defined as

$$BD = \frac{1}{8}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^T \left[\frac{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}{2} \right]^{-1} (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) + \frac{1}{2} \ln \frac{|\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2/2|}{\sqrt{|\boldsymbol{\Sigma}_1|}\sqrt{|\boldsymbol{\Sigma}_2|}}. \tag{9}$$

In Equation (9), $\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1$ and $\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2$ are the mean vector and covariance matrix of the first and second multivariate normal distribution models, respectively. To describe the average Bhattacharyya distance, suppose that we have two fields denoted as T (test) and C (control) with g and k components of the mixture of multivariate normal distribution models. The Bhattacharyya distance between the components i and j of the mixture of multivariate normal distribution models for the test field T and the control field C, denoted by $d_{ij}^{T,C}$, is computed as

$$d_{ij}^{T,C} = \frac{1}{8}(\boldsymbol{\mu}_{j,C} - \boldsymbol{\mu}_{i,T})^T \left[\frac{\boldsymbol{\Sigma}_{i,T} + \boldsymbol{\Sigma}_{j,C}}{2} \right]^{-1} (\boldsymbol{\mu}_{j,C} - \boldsymbol{\mu}_{i,T}) + \frac{1}{2} \ln \frac{\left| \frac{\boldsymbol{\Sigma}_{i,T} + \boldsymbol{\Sigma}_{j,C}}{2} \right|}{\sqrt{|\boldsymbol{\Sigma}_{i,T}|}\sqrt{|\boldsymbol{\Sigma}_{j,C}|}} \tag{10}$$

for $i = 1, 2, \dots, g$ and $j = 1, 2, \dots, k$. Then, multiplying mixing weights with $d_{ij}^{T,C}$ Bhattacharyya distances in Equation (10), we obtain the weighted Bhattacharyya distance, denoted by d_{ij}^W , in terms of components as

$$d_{ij}^W = \pi_i \pi_j d_{ij}^{T,C}. \tag{11}$$

After that, the average Bhattacharyya distance in terms of components is computed by averaging the weighted Bhattacharyya distance denoted by $d_{avg}^{T,C}$:

$$d_{avg}^{T,C} = \frac{1}{kg} \sum_{i=1}^g \sum_{j=1}^k d_{ij}^W. \tag{12}$$

The average Bhattacharyya distance $d_{avg}^{T,C}$ in Equation (12) is considered for the discrimination function of the proposed per-field classification method.

The decision rule is established according to the minimum of the average Bhattacharyya distances. For this purpose, the average Bhattacharyya distances were computed between the 245 test fields and the 24 control fields. Then, a test field was classified to one of the control fields such that the average Bhattacharyya distances between them were minimum. While the lowest value of the average Bhattacharyya distances refers to a strong similarity between the test and control fields, the highest value of the average Bhattacharyya distances refers to a strong difference for the proposed per-field classification method.

3. Per-field classification of the agricultural area using MDA

In this section, remotely sensed multispectral image data of the agricultural area are classified with the proposed supervised per-field classification method based on MDA. All fields, 245 test fields and 24 control fields, in the multispectral image data were determined applying an edge detection algorithm and using the parcelization map of this agricultural area [4, 11].

We analyzed the proposed method for different covariance matrix structures ($\boldsymbol{\Sigma}_{jr}, \boldsymbol{\Sigma}_j$ and $\boldsymbol{\Sigma}$) with fixed (3) and different number of components for each field in the agricultural area. So, we computed the classification accuracy using the proposed method for six different types of distribution models. Also, we applied the per-pixel classification method based on Bayes classification rule for these different models. Thus, we can see how the classification accuracy changes

Table 1. The control field codes, field codes, field types, number of pixels used in the training data set, number of total pixels in the fields and number of components in the mixture of multivariate normal distribution models constructed for the control field.

Control field code	Field code	Field type	Number of pixels used in the training data	Number of total pixels in the fields	Number of components in the mixture of distribution models for the control field
CF001	F214	Wheat	144	429	2
CF002	F078	Wheat	204	412	2
CF003	F204	Wheat	180	346	2
CF004	F152	Wheat	351	432	1
CF005	F257	Wheat	200	204	3
CF006	F233	Wheat	160	190	2
CF007	F237	Potato	135	275	3
CF008	F223	Potato	72	262	3
CF009	F234	Potato	142	162	3
CF010	F165	Potato	144	624	3
CF011	F219	Vegetable garden	143	262	2
CF012	F173	Vegetable garden	108	427	3
CF013	F104	Vegetable garden	255	343	3
CF014	F077	Vegetable garden	42	98	3
CF015	F101	Vegetable garden	75	127	2
CF016	F011	Vegetable garden	88	339	2
CF017	F012	Citrus	154	436	2
CF018	F027	Citrus	72	169	3
CF019	F120	Citrus	96	176	3
CF020	F148	Bare soil	224	581	2
CF021	F176	Bare Soil	336	609	2
CF022	F068	Bare soil	289	600	2
CF023	F099	Bare soil	104	222	3
CF024	F150	Bare soil	190	257	3

for the proposed per-field and per-pixel classification methods under different covariance matrix structures with fixed (3) and different number of components.

In the proposed classification method, multivariate normal or mixture of multivariate normal distribution models were constructed for each field in the agricultural area. So without any restrictions, the 245 test fields and 24 control fields were modeled with a different covariance matrix (Σ_{jr}) and a different number of components. When creating the distribution models, the MIXMOD software and EMMIX software were used parallelly for determining the number of components and constructing the mixture of multivariate normal distribution models [2,9]. The control field codes, field codes, field types, number of pixels used in the training data set, number of total pixels in the fields and number of components in the mixture of multivariate normal distribution models constructed for the control fields are given in Table 1.

When choosing the number of components, the Akaike information criteria were used in the MIXMOD and EMMIX software [1,2,9]. The number of components in the mixture of multivariate normal distribution models constructed for 269 fields and the corresponding percentages were as follows:

- 23 fields of totally 269 fields, 9% of total fields, had a multivariate normal distribution,
- 123 fields of totally 269 fields, 46% of total fields, had a mixture of multivariate normal distributions with two components,
- 117 fields of totally 269 fields, 43% of total fields, had a mixture of multivariate normal distributions with three components and

- 6 fields of totally 269 fields, 2% of total fields, had a mixture of multivariate normal distributions with four components.

This means that 91% of the fields in the agricultural area had a mixture of multivariate normal distributions. The parameters in the mixture of multivariate normal distribution models were estimated based on the determined number of components using the MIXMOD and EMMIX software [2,9]. The examples of the graphs for the multivariate normal and mixture of normal distribution models with one to four components are shown in Figure 2.

After determining the parameters of the distribution models, the average Bhattacharyya distance in Equation (12) was computed between the distribution models of the test and control fields for the situation of a different covariance matrix (Σ_{jr}) and a different number of components. The minimum values of the average Bhattacharyya distances were used as a decision rule to classify the test fields to the control fields. As a result of this classification, a constructed error matrix given in Table 2 was obtained.

In the same way, we modeled the test and control fields for the remaining five situations, two of them were Σ_j and Σ with a different number of components and three of them were Σ_{jr} , Σ_j and Σ with a fixed (3) number of components. Then, we computed the average Bhattacharyya distances between the test and control fields. The classification results of these situations resulted in constructed error matrixes, which are given in Tables 3–7, respectively.

We also classified the test fields to the control fields using the per-pixel classification method based on MDA proposed by Hastie and Tibshirani [6] and by Ju *et al.* [7] to determine the performances of their methods for different structures of a covariance matrix with fixed and different number of components. For this purpose: first, the mixture of multivariate normal distribution

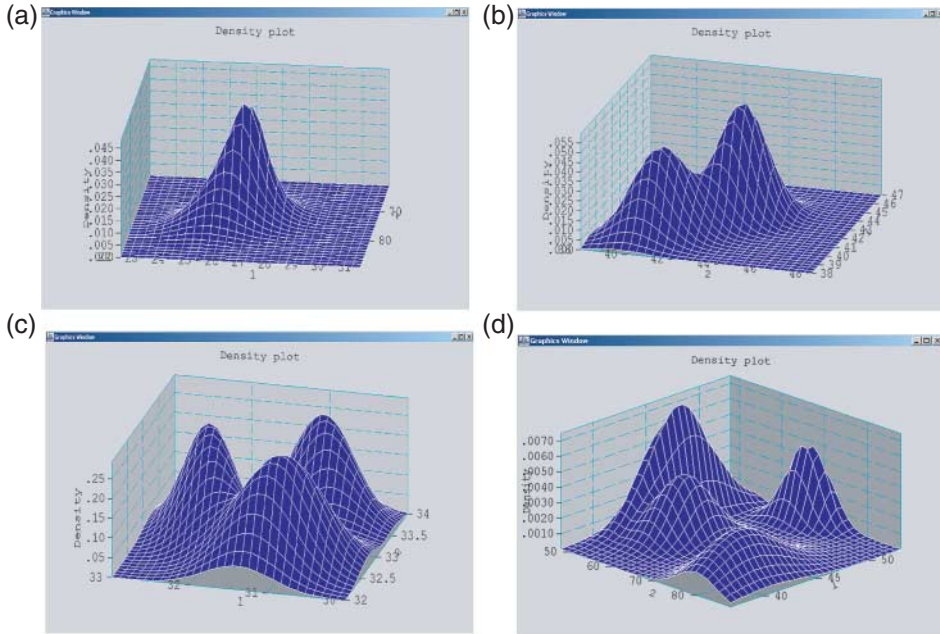


Figure 2. (a) The graph of a multivariate normal distribution model for control field CF004 containing wheat. (b) The graph of a mixture of multivariate normal distribution models with two components for control field CF020 containing bare soil. (c) The graph of a mixture of multivariate normal distribution models with three components for control field CF009 containing potato. (d) The graph of a mixture of multivariate normal distribution models with four components for test field TF152 containing citrus.

Table 2. The error matrix for the proposed per-field classification method using MDA with Σ_{jr} and different number of components.

	Wheat	Potato	Vegetable garden	Citrus	Bare soil	Row total	User's accuracy
Wheat	10,660	0	0	857	0	11,517	0.92
Potato	225	1596	230	1077	190	3318	0.48
Vegetable garden	369	219	6626	274	2498	9986	0.66
Citrus	0	0	0	568	218	786	0.72
Bare soil	0	55	543	130	5558	6286	0.88
Column total	11,254	1870	7399	2906	8464	31,893	
Producer's accuracy	0.94	0.85	0.89	0.19	0.65		0.7841

Table 3. The error matrix for the proposed per-field classification method using MDA with Σ_j and different number of components.

	Wheat	Potato	Vegetable garden	Citrus	Bare soil	Row total	User's accuracy
Wheat	11,280	217	20	0	0	11,517	0.98
Potato	973	2097	0	138	110	3318	0.63
Vegetable garden	432	648	7837	141	928	9986	0.78
Citrus	0	0	238	326	222	786	0.41
Bare soil	0	1159	1293	826	3008	6286	0.48
Column total	12,685	4121	9388	1431	4268	31,893	
Producer's accuracy	0.89	0.51	0.83	0.23	0.70		0.7697

Table 4. The error matrix for the proposed per-field classification method using MDA with Σ and different number of components.

	Wheat	Potato	Vegetable garden	Citrus	Bare soil	Row total	User's accuracy
Wheat	7031	3174	882	430	0	11,517	0.61
Potato	1391	834	1016	0	77	3318	0.25
Vegetable garden	2890	145	6384	54	513	9986	0.64
Citrus	57	0	0	273	456	786	0.35
Bare soil	371	899	2594	608	1814	6286	0.29
Column total	11,740	5052	10,876	1365	2860	31,893	
Producer's accuracy	0.60	0.17	0.59	0.20	0.63		0.5122

Table 5. The error matrix for the proposed per-field classification method using MDA with Σ_{jr} and fixed (3) number of components.

	Wheat	Potato	Vegetable garden	Citrus	Bare soil	Row total	User's accuracy
Wheat	11,364	0	0	153	0	11,517	0.99
Potato	1922	1225	171	0	0	3318	0.37
Vegetable garden	441	0	6913	183	2449	9986	0.69
Citrus	0	0	234	552	0	786	0.70
Bare soil	137	0	487	131	5531	6286	0.88
Column total	13,864	1225	7805	1019	7980	31,893	
Producer's accuracy	0.82	1.00	0.89	0.54	0.69		0.8022

Table 6. The error matrix for the proposed per-field classification method using MDA with Σ_j and fixed (3) number of components.

	Wheat	Potato	Vegetable garden	Citrus	Bare soil	Row total	User's accuracy
Wheat	11,409	0	0	108	0	11,517	0.99
Potato	1267	1875	66	0	110	3318	0.57
Vegetable garden	581	88	7933	141	1243	9986	0.79
Citrus	0	0	238	330	218	786	0.42
Bare soil	0	761	1578	88	3859	6286	0.61
Column total	13,257	2724	9815	667	5430	31,893	
Producer's accuracy	0.86	0.69	0.81	0.49	0.71		0.7966

Table 7. The error matrix for the proposed per-field classification method using MDA with Σ and fixed (3) number of components.

	Wheat	Potato	Vegetable garden	Citrus	Bare soil	Row total	User's accuracy
Wheat	5565	2312	3444	87	109	11,517	0.48
Potato	27	659	2517	0	115	3318	0.20
Vegetable garden	269	117	6886	0	2714	9986	0.69
Citrus	0	0	568	218	0	786	0.28
Bare soil	0	934	2467	608	2277	6286	0.36
Column total	5861	4022	15,882	913	5215	31,893	
Producer's accuracy	0.95	0.16	0.43	0.24	0.44		0.4893

Table 8. The error matrix for the per-pixel classification method using MDA with Σ_{jr} and different number of components.

	Wheat	Potato	Vegetable garden	Citrus	Bare soil	Row total	User's accuracy
Wheat	9238	1263	520	276	220	11,517	0.80
Potato	1026	1486	429	80	297	3318	0.45
Vegetable garden	592	799	5237	357	3001	9986	0.52
Citrus	7	20	210	362	187	786	0.46
Bare soil	206	952	1618	371	3139	6286	0.50
Column total	11,069	4520	8014	1446	6844	31,893	
Producer's accuracy	0.83	0.33	0.65	0.25	0.46		0.6102

models for different structures of a covariance matrix (Σ_{jr} , Σ_j and Σ) with fixed (3) and different number of components was constructed for each of the 24 control fields using the MIXMOD and EMMIX software [2,9]. Second, the discrimination function values of the per-pixel classification method based on MDA were computed as the maximum posterior probabilities between the data points, which were 31,893 pixel values for 3, 4 and 5 bands in the remaining 245 test fields and the mixture of multivariate normal distribution models. Finally, error matrices were constructed at the end of this method for different structures of a covariance matrix (Σ_{jr} , Σ_j and Σ) with fixed (3) and different number of components. The error matrices for situations of Σ_{jr} , Σ_j and Σ with a different number of components are given in Tables 8–10, respectively.

The error matrices for the situations of Σ_{jr} , Σ_j and Σ with a fixed (3) number of components are given in Tables 11–13, respectively.

Table 9. The error matrix for the per-pixel classification method using MDA with Σ_j and different number of components.

	Wheat	Potato	Vegetable garden	Citrus	Bare soil	Row total	User's accuracy
Wheat	9158	1335	502	294	228	11,517	0.80
Potato	1001	1505	426	78	308	3318	0.45
Vegetable garden	564	813	5295	302	3012	9986	0.53
Citrus	7	20	244	306	209	786	0.39
Bare soil	189	931	1636	311	3219	6286	0.51
Column total	10,919	4604	8103	1291	6976	31,893	
Producer's accuracy	0.84	0.33	0.65	0.24	0.46		0.6043

Table 10. The error matrix for the per-pixel classification method using MDA with Σ and different number of components.

	Wheat	Potato	Vegetable garden	Citrus	Bare soil	Row total	User's accuracy
Wheat	7508	2648	652	565	144	11,517	0.65
Potato	1212	851	779	141	335	3318	0.26
Vegetable garden	891	356	5085	851	2803	9986	0.51
Citrus	138	2	99	403	144	786	0.51
Bare soil	337	545	1667	646	3091	6286	0.49
Column total	10,086	4402	8282	2606	6517	31,893	
Producer's accuracy	0.74	0.19	0.61	0.15	0.47		0.5311

Table 11. The error matrix for the per-pixel classification method using MDA with Σ_{jr} and fixed (3) number of components.

	Wheat	Potato	Vegetable garden	Citrus	Bare soil	Row total	User's accuracy
Wheat	9298	1174	556	303	186	11,517	0.81
Potato	710	1759	508	74	267	3318	0.53
Vegetable garden	585	767	5543	279	2812	9986	0.56
Citrus	12	20	218	331	205	786	0.42
Bare soil	234	770	1806	346	3130	6286	0.50
Column total	10,839	4490	8631	1333	6600	31,893	
Producer's accuracy	0.86	0.39	0.64	0.25	0.47		0.6290

Table 12. The error matrix for the per-pixel classification method using MDA with Σ_j and fixed (3) number of components.

	Wheat	Potato	Vegetable garden	Citrus	Bare soil	Row total	User's accuracy
Wheat	9188	1215	520	291	303	11,517	0.80
Potato	638	1761	532	70	317	3318	0.53
Vegetable garden	444	509	5455	298	3280	9986	0.55
Citrus	3	10	219	272	282	786	0.35
Bare soil	227	713	1718	358	3270	6286	0.52
Column total	10,500	4208	8444	1289	7452	31,893	
Producer's accuracy	0.88	0.42	0.65	0.21	0.44		0.6254

Table 13. The error matrix for the per-pixel classification method using MDA with Σ and fixed (3) number of components.

	Wheat	Potato	Vegetable garden	Citrus	Bare soil	Row total	User's accuracy
Wheat	8708	121	1870	405	413	11,517	0.76
Potato	842	182	1536	186	572	3318	0.05
Vegetable garden	509	1547	2079	700	5151	9986	0.21
Citrus	7	4	161	334	280	786	0.42
Bare soil	269	432	1030	403	4152	6286	0.66
Column total	10,335	2286	6676	2028	10,568	31,893	
Producer's accuracy	0.84	0.08	0.31	0.16	0.39		0.4756

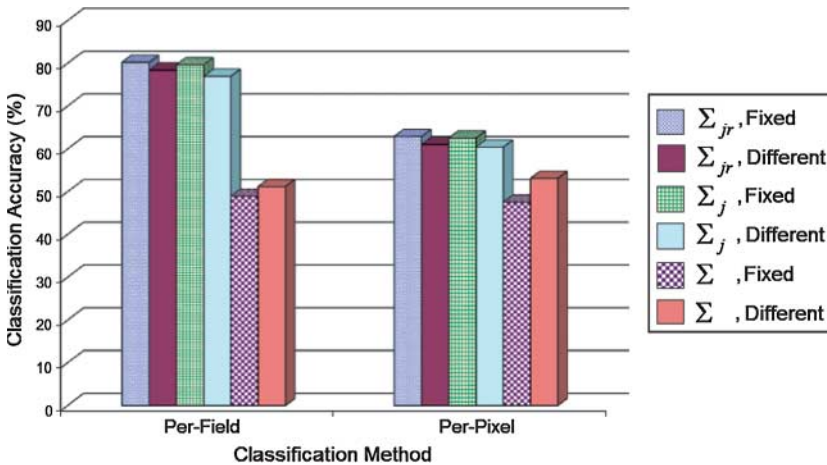


Figure 3. The classification results of the proposed per-field classification method and the per-pixel classification method for different structures of a covariance matrix with fixed (3) and different number of components.

We present a graphical representation in Figure 3 for a better view of all the classification accuracy results of the proposed per-field classification and the per-pixel classification methods for different structures of a covariance matrix with fixed (3) and different number of components obtained at the end of all the classification applications.

4. Conclusions and discussions

For fields in agricultural areas, most of the digitalized multispectral image data can have a mixture of multivariate normal distribution models with a different number of components as shown in Figure 2(b)–(d). Therefore, a new per-field classification method based on MDA has been proposed for supervised classification of remotely sensed multispectral image data of an agricultural area. The discrimination function of the proposed method is defined by the average Bhattacharyya distance and the decision rule is established according to the minimum values of the average Bhattacharyya distances.

For the proposed method, multivariate normal or mixture of normal distributions can be used since the average Bhattacharyya distance can be computed between different binaries of these distribution models. The mixture of multivariate normal distribution models constructed for the

test and control fields can have a different number of components. Each component of these mixture models can have different covariance matrices in the proposed method.

In Section 3, we applied the proposed method for different covariance matrix structures (Σ_{jr} , Σ_j and Σ) with fixed (3) and different number of components for each field in the agricultural area. So, we computed the classification accuracy for six different types of distribution models. The best classification accuracy was obtained, 80.22%, for the Σ_{jr} covariance matrix with a fixed (3) number of components given in Table 5. Also, the second best classification accuracy, 79.66%, was obtained for the Σ_j covariance matrix with a fixed (3) number of components given in Table 6. In addition to the proposed per-field classification method, we applied the per-pixel classification method based on Bayes classification rule under the same situations. In parallel to the proposed per-field classification method, the first and second best classification accuracies, 62.90% and 62.54%, were obtained for the per-pixel classification method for the Σ_{jr} and Σ_j covariance matrices with a fixed (3) number of components given in Tables 11 and 12, respectively.

Despite the fact that different methods have been used, the classification results of the per-field and per-pixel methods were parallel according to the covariance matrices and number of components. When the number of components was fixed (3) and the covariance matrices were different for sub-groups (Σ_{jr}) or groups (Σ_j), the best classification results were obtained for both the methods. Also, the worst classification results of these methods were obtained for the Σ covariance matrix with a fixed (3) number of components. Most of the classification results of the proposed per-field classification method were better than those of the per-pixel classification method. But instead of comparing the classification methods, in this study, it has been shown as to how the classification accuracies of the per-field and per-pixel classification methods change for different covariance matrix structures and a different number of components.

Acknowledgements

This work was supported by the TUBITAK BAYG (No.2211) and Çukurova University Scientific Research Project Unit (No. FEF 2008D16 LTP). The authors thank the editor and especially two of the anonymous referees for carefully reading the manuscript and making some valuable comments which had greatly improved the earlier draft of the manuscript.

References

- [1] H. Akaike, *A new look at the statistical model identification*, IEEE Trans. Autom. Control 19 (1974), pp. 716–723.
- [2] C. Biernacki, G. Celeux, G. Govaert, and F. Langrognet, *Model-based cluster analysis and discriminant analysis with the MIXMOD software*, Comput. Statist. Data Anal. 51 (2006), pp. 587–600.
- [3] A.P. Dempster, N. Laird, and D. Rubin, *Maximum likelihood from incomplete data via the EM algorithm*, J. R. Stat. Soc. 39 (1977), pp. 1–38.
- [4] H. Erol and F. Akdeniz, *A per-field classification method based on mixture distribution models and an application to Landsat Thematic Mapper data*, Int. J. Remote Sens. 26 (2005), pp. 1229–1244.
- [5] C. Fraley and A.E. Raftery, *Model-based clustering, discriminant analysis, and density estimation*, J. Amer. Statist. Assoc. 97 (2002), pp. 611–631.
- [6] T. Hastie and R. Tibshirani, *Discriminant analysis by Gaussian mixtures*, J. R. Stat. Soc. 58 (1996), pp. 155–176.
- [7] J. Ju, E.D. Kolaczyk, and S. Gopal, *Gaussian mixture discriminant analysis and sub-pixel land cover characterization in remote sensing*, Remote Sens. Environ. 84 (2003), pp. 550–560.
- [8] G.J. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*, Wiley and Sons, New York, 1992.
- [9] G.J. McLachlan, D. Peel, K.E. Basford, and P. Adams, *The EMMIX software for the fitting of mixtures of normal and t-components*, J. Statist. Softw. 4 (1999), pp. 1–4.
- [10] K. Win, S. Baik, R. Baik, S. Ahn, S. Kim, and Y. Jo, *Texture feature extraction and selection for classification of images in a sequence*, IWCI, LNCS 3322, 2004, pp. 750–757.
- [11] D. Ziou and S. Tabbone, *Edge detection techniques an overview*, Tech. Rep. No.195, Dept Math & Informatique, Universit de Sherbrooke, 1997.