Article ID: 1007-4619(2007)05-0748-08

Nonparametric and Probabilistic Classification of Agricultural Crops Using Multitemporal Images

YU Jun, Bo Ranneby
(Centre of Biostochastics Swedish University of Agricultural Sciences SE-901 83 Umea, Sweden)

Abstract: In this paper, a new approach for classification of multitemporal satellite data sets combining multispectral and change detection techniques is proposed. The algorithm is based on the nearest neighbor method and derived in order to optimize the average probability for correct classification, i e each class is equally important. The new algorithm was applied to a study area where satellite images (SPOT and Landsat TM) from different seasons were used. It showed that using five seasonal images can substantially improve the classification accuracy compared to using a single image. As a large scale application, the approach was applied to the Dal River drainage basin.
As the distributions for different classes are highly overlapping it is not possible to get satisfactory accuracy at pixel level. Instead it is necessary to introduce a new concept: pixel-wise probabilistic classifiers. The pixel-wise vectors of probabilities can be used to judge how reliable a traditional classification is and to derive measures of the uncertainty (entropy) for the individual pixels. The probabilistic classifier gives also unbiased area estimates over arbitrary areas. It has been tested on two test sites of arable land with different characteristics.

Key words: nonparametric classification; nearest neighbor method; probabilistic classifier; agricultural crops; quality assessment; multitemporal images; remote sensing; drainage basin
CLC number: TP79 Document code: A

1 INTRODUCTION

The aim with this study is to test the usefulness of remote sensing methods for classification of agricultural crops. The test area is located in Dalarna, Sweden. Characteristic for agriculture fields in that area is that they are long and narrow, which is a complication for the classification. As a consequence, several of the existing remote sensing classification methods will not give satisfactory results. For example, the traditional maximum likelihood method will not work for classification of rare classes. To overcome most of the problems with traditional methods we use a new approach for classification of multitemporal satellite data sets combining multispectral and change detection techniques. The next section is devoted to describe this new classification algorithm, and to introduce the pixel-wise probabilistic classifier, a new concept for measuring the uncertainty at the pixel level. Section 3 presents the study area and input data including masks, satellite images, and field data. Results of the classification over the study area and discussion are given in Section 4. A large scale application of the new approach is presented in Section 5.

2 METHODS

The classification algorithm developed in this paper will be based on the nearest neighbor (k-NN) meth-

Received date: 2006-08-10; Accepted date: 2007-03-15
Foundation item: Authors are grateful to MISTRA (The Foundation for Strategic Environmental Research in Sweden) and Swedish National Space Board for their financial support.

Biography: YU Jun received his Ph. D. degree in mathematical statistics from Umea University, Sweden, in 1994. In January 2002, he joined the Centre of Biostochastics, Swedish University of Agricultural Sciences. His current research interests include statistical classification of remotely sensed imagery, spatial pattern analysis, wavelet theory applied to statistical problems, spatial statistics and it applications to forest and environmental data, non-stationary biomedical signal and image processing, and time series analysis and forecasting for financial data. E-mail: jun.yu@sekon.slu.se

od introduced by Fix & Hodges[1]. The main procedures are as follows

- Define the target function. The sum of probabilities of correct classification for each class is used. Note that it differs from the overall correct classification rate which is often used in various applications
- Denoise the feature vector. The wavelet shrinkage method based on 2D-wavelet transform is used to denoise the images[2,3]. The feature vector consists of components that are pixel values from different spectral bands
- Remove outliers from the reference data. When the feature vector has high dimension and the number of classes is large, new data editing methods are needed in order to remove outliers due to poor quality of field data, and to find out the prototypes for each class in the case of nearest neighbor classifiers. This must be done so that poor quality data is removed at the same time as most of the natural variations within the different classes are kept
- Calculate the information values in the components in the feature vector. Before the feature vector can be used the components have to be rescaled. The size of the rescaling factor is a function of the dependence between the studied objects and the component. Since any natural order relation between the classes does not exist conventional dependence measures such as the correlation coefficient cannot be used. Instead measures with the origin in information theory are used[4,5].
- Determine a proper metric. We prefer to use the following metric:

$$\sqrt{\sum_{i=1}^{d} w_i^2 \left| x_i(s) - x_i(t) \right|^2} \qquad (1)$$

where $x(s) = [x_1(s), x_2(s), \cdots, x_d(s)]$ and $x(t) = [x_1(t), x_2(t), \cdots, x_d(t)]$ are two pixels at location s and t with attribute of the spectral components. d is the dimension of the feature vector. $w_i = \frac{\alpha_i^p / q_i}{\sum \alpha_i^p / q_i}$ is the rescaling factor. q_i and α_i are the interquartile range and the information value of the ith component respectively, and p is determined so the target function is maximized. For practical reasons p is restricted to the positive integers or the reciprocal of them. Usually the maximum is achieved when p equals to 1.

- Determine prototypes for the classes. Here the information about the occurrence of different classes in the reference data can be used
- Run a nonparametric classification. Here the nearest neighbor classifier is used[1,6].
- Declare the quality of classification result by using probability matrices. The probability matrix is based on the confusion matrix and defined as $P = [P_{ij}]$ where P_{ij} is the probability that class i is classified as class j

The classification algorithm is derived in order to optimize the average probability for correct classification, i.e. each class is equally important. If we want to have good classification performance also for small classes they have to be overestimated (unless the classification is 100% perfect). Traditional methods as Maximum likelihood classification or Linear Discriminant Analysis (LDA) with or without prior information fail completely[7].

However, as the distributions for different classes are highly overlapping it is not possible to get satisfactory accuracy at pixel level. Instead it is necessary to introduce a new concept: pixel-wise probabilistic classifiers. This concept is closely related to fuzzy classification: see Bezdek et al (1999)[8] for different definitions. Instead of classifying each pixel to a specific class, each pixel is given a probability distribution describing how likely the different classes are. The probabilistic classifier derived in this project will be based on the k-NN method. Articles on fuzzy classification and k-NN are available, see e.g. Bereau & Dubuisson (1991)[9] and Kissiov & Hadjitodorov (1992)[10], but they have a different scope. To be useful for area estimation the probabilities must be extracted from proper class distributions. When the classifier uses distances in feature space to derive the probabilities it is intuitively obvious that the probabilities will be inversely proportional to the distances raised to some power. However, there is only one value that is correct and that value is given by the number of components in the feature vector, that is the dimension of the feature space. Actually, by using the nonparametric density estimation[11], the density for class c at $x(s)$ will be inversely proportional to the volume of the nearest neighbor ball with center $x(s)$. Note that the volume is proportional to the radius (based on the distance defined in (1)) raised to the dimension of the feature space.

Specifically, let the class labels be 1, 2, …, C. For each pixel with location s and feature vector $x(s) = [x_1(s), x_2(s), \cdots, x_d(s)]$, the probability that this pixel belongs to class c is defined as

$$p_c(s) = \frac{\mathrm{dist}(x(s), c)^{-d}}{\sum_{c=1}^{C} \mathrm{dist}(x(s), c)^{-d}} \qquad (2)$$

where dist$(x(s), c)$ = minimum distance in the feature space (using the metric defined in (1)) from $x(s)$ to the prototypes of class c. Thus, the probability vector at pixel s is obtained as $p(s) = [p_1(s), p_2(s), \cdots, p_C(s)]$.

The pixel-wise vectors of probabilities can be used to judge how reliable a traditional classification is and to derive measures of the uncertainty (entropy) for the individual pixels. Based on the probability vector, the entropy at pixel s is defined as $-\sum_{c=1}^{C} p_c(s) \log p_c(s)$.

It is extremely important that proper probability distributions allowing frequency interpretation are derived; otherwise misleading results are obtained. The probabilistic classifier can also be used to obtain unbiased area estimates over arbitrary regions.

3 STUDY AREA AND INPUT DATA

The study area is located in the eastern part of Dalarna in Sweden. Input data includes various masks, satellite data (Landsat TM and SPOT images), difference images, field data ("block database").

- Map masks. Masks from the topographic maps will be used for stratification before classification. Agricultural area was defined by the agricultural mask from the 1∶100000 scale "blue map".
- Cloud cover masks. All scenes have been manually interpreted for cloud cover and other image data errors. Areas covered by clouds, haze and shadows from clouds were digitized on the screen. Other pixels having erroneous measurements for different other reasons were also marked and included in the "cloud mask", which in reality is a mask defining pixels not to be used for classification. Examples of this are data dropouts, resampling effects at the scene edges and "no data" pixels outside the imaged area.
- Satellite data. The images to be used are either Landsat TM images or SPOT images. The satellite data were geometric corrected with ortho-correction methods. The data are resampled to 25m pixel size. The following table presents the satellite images used in the study area.

Table 1 SPOT and Landsat images used in the study area

	Scene	Date	Spectral bands used
SPOT-2	054-226/8	1998-10-24	XS 1, 2, 3
Landsat-5	194-018	1999-05-07	TM 2, 3, 4, 5
Landsat-5	194-018	1999-07-10	TM 2, 3, 4, 5
SPOT-4	054-226/227	1999-07-30	XS 1, 2, 3, 4
Landsat-7	195-018	1999-09-11	TM 2, 3, 4, 5

All images used were registered by the Landsat satellites with the TM sensor (Thematic Mapper) as the principal instrument. The TM sensor has 7 spectral bands in visible, near infrared, mid infrared and thermal infrared with 30 meter ground resolution. Most of the satellite data were geometric corrected with ortho-correction methods, using the digital elevation model to remove image parallaxes. The geometric accuracy requirement for multitemporal data classification is to aim for less than one pixel RMS error. This quality was not reached for all scenes used, leading to local misalignments between scenes in some cases. The main reason for this has been the deteriorated geometric quality of the ageing Landsat-5 TM sensor, which was launched already in 1984 and could not be replaced completely until 2000 by Landsat-7. The Thematic Mapper data are resampled to 25 meter pixel size for the continued processing. The image data is stored as 8 bit integer data in each band, corresponding to digital numbers from 0–255.

- Field data. The "block database", i.e. a GIS database with polygons grouping the parcels of agricultural units ("blocks") that are registered in the IAKS database (the administrative database for agricultural aid in Sweden). The units can contain either one crop (one field unit) or several crops (multi-field unit).

4 RESULTS AND DISCUSSION

The crops and agricultural land were predefined into 25 classes, including cereals (autumn-sown and spring-sown), oil seed crops, potatoes, grassland on arable land, energy forest (salix), and so on. Tables 2 and 3 give the definition of crop classes and different thematic levels. Because the spectral signature differs between the center and the boundary of the arable units, and from one field units to multi-field units, the pixels along the edges and transition zones were grouped into their own classes.

Table 2 Class definition and different thematic levels for the study area

Crops	Level 3	Level 2	Level 1
Winter barley	1	1	1
Spring barley	2	2	1
Oats	3	2	1
Winter wheat	4	1	1
Spring wheat	5	2	1
Mixed grain	6	2	1
Rye	7	1	1
Spring turnip rape	8	3	1
Peas	9	7	1
Linseed	10	3	1
Potatoes	11	4	1
Pasture or grass for hay or silage on arable land	12	5	1
Seed ley	13	5	1
Grazing land	14	7	2
Meadow	15	5	2
Woodland pasture	16	7	2
Fallow (one year)	17	7	1
Fallow (>one year)	18	7	1
Energy forest (salix)	19	6	1
Reed canary grass (1999)	20	5	1
Strawberries	21	7	1
Other cultivation of berries	22	7	1
Forest plantation on arable land	23	7	1
Other land use	24	7	2
Boundary pixels	25	8	3

Table 3 Definition of the thematic levels

Code	Level 2	Level 1
1	Autumn-sown cereals	Arable land
2	Spring-sown cereals	Grazing land
3	Spring-sown oil seed crops	Boundary pixels
4	Potatoes	
5	Grass land on arable land for hay or silage	
6	Energy forest (salix)	
7	Other crops	
8	Boundary pixels	

In field data, only few were used as reference data for classification, i.e. build-up of prototypes for each class. The major part was used for validation.

For the purpose of comparison, classification was done using a single image and five images from different seasons, respectively. The probability matrices at level 1 are shown in Table 4, and the corresponding results at level 2 are presented in Tables 5 and 6. It is evident that the use of multitemporal images over seasons can substantially improve the classification result compared to a single image. With a single image, however, spring-sown cereals and spring-sown oil seed crops (C2 and C3 in Table 6) have quite high accuracy (53% and 70% respectively). At the thematic level 1 we have very good classification accuracy for arable land (84%), but much lower for grazing land (50%). This is mainly depending on confusion with pasture or grass for hay or silage on arable land.

Table 4 Probability matrices at level 1

	5 seasonal scenes			1 single scene		
	C1	C2	C3	C1	C2	C3
C1	0.90	0.09	0.01	0.84	0.14	0.02
C2	0.36	0.62	0.02	0.49	0.49	0.03

Table 5 Probability matrix using five seasonal scenes at level 2

	C1	C2	C3	C4	C5	C6	C7	C8
C1	0.49	0.35	0.00	0.00	0.03	0.00	0.14	0.00
C2	0.04	0.78	0.01	0.00	0.03	0.00	0.14	0.01
C3	0.00	0.07	0.72	0.00	0.01	0.00	0.20	0.00
C4	0.01	0.09	0.01	0.65	0.04	0.00	0.20	0.00
C5	0.02	0.02	0.02	0.00	0.63	0.01	0.27	0.02
C6	0.00	0.04	0.00	0.00	0.11	0.56	0.27	0.01
C7	0.01	0.04	0.00	0.00	0.20	0.01	0.71	0.02

Table 6 Probability matrix using one single scene at level 2

	C1	C2	C3	C4	C5	C6	C7	C8
C1	0.19	0.30	0.01	0.03	0.11	0.00	0.34	0.01
C2	0.04	0.53	0.01	0.01	0.10	0.01	0.27	0.02
C3	0.01	0.07	0.70	0.00	0.05	0.00	0.17	0.00
C4	0.02	0.02	0.01	0.17	0.37	0.00	0.39	0.01
C5	0.02	0.04	0.00	0.05	0.47	0.02	0.38	0.02
C6	0.01	0.14	0.00	0.00	0.12	0.28	0.42	0.03
C7	0.02	0.06	0.01	0.02	0.28	0.04	0.54	0.03

Looking further into the classification accuracy at level 3 in Table 7, it can be found that even with images from five occasions it is difficult to identify winter barley (C1) and winter wheat (C4). One possible reason is that the image from 7th of May is too early. This is confirmed by the information value for spectral band 5. The value is only half of that from the image taken in July. The probabilities of correct classification for these two classes are slightly below 40%. Both crops occupy, however, less than 1% each of the cultivated area in the study area. The other classes of grains have considerably higher probabilities of correct classification (55% − 93%). For other crops and agricultural landscapes the probability of correct classification is above 50%, except for meadow (C15), fallow (>1 year) (C18) and other land use (C24), which as expected has low accuracy (Table 7).

Even though the classification accuracy at pixel level

Table 7 Probability matrix using five seasonal scenes at level 3

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25
C1	0.38	0.35	0.14	0.04	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C2	0.02	0.65	0.11	0.01	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.03	0.01	0.03	0.00	0.01	0.08	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01
C3	0.02	0.22	0.55	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.02	0.01	0.03	0.00	0.01	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01
C4	0.08	0.22	0.03	0.36	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.04	0.00	0.03	0.00	0.00	0.15	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C5	0.00	0.00	0.00	0.00	0.93	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C6	0.00	0.03	0.00	0.02	0.00	0.90	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C7	0.03	0.15	0.00	0.07	0.00	0.00	0.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C8	0.00	0.04	0.01	0.00	0.00	0.00	0.00	0.77	0.12	0.00	0.01	0.00	0.01	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C9	0.00	0.07	0.04	0.00	0.00	0.00	0.00	0.00	0.82	0.00	0.00	0.01	0.00	0.03	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C10	0.01	0.11	0.02	0.00	0.00	0.00	0.01	0.00	0.00	0.46	0.00	0.02	0.00	0.00	0.00	0.00	0.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
C11	0.00	0.06	0.03	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.65	0.01	0.02	0.03	0.01	0.01	0.14	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00
C12	0.00	0.01	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.02	0.00	0.62	0.00	0.10	0.00	0.02	0.13	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.02
C13	0.02	0.13	0.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.49	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C14	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.48	0.02	0.11	0.04	0.02	0.01	0.00	0.00	0.01	0.00	0.01	0.03
C15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.12	0.00	0.68	0.13	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C16	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.04	0.00	0.23	0.00	0.65	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01
C17	0.00	0.05	0.01	0.01	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.07	0.00	0.08	0.00	0.02	0.63	0.02	0.01	0.02	0.00	0.00	0.00	0.02	0.00
C18	0.00	0.02	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.02	0.00	0.25	0.00	0.17	0.00	0.03	0.16	0.21	0.01	0.02	0.00	0.00	0.00	0.04	0.03
C19	0.00	0.02	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.12	0.00	0.03	0.05	0.05	0.56	0.00	0.00	0.00	0.01	0.01	0.01
C20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.02	0.00	0.10	0.15	0.00	0.00	0.71	0.00	0.00	0.00	0.00	0.00
C21	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.03	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.72	0.00	0.00	0.00	0.00
C22	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.02	0.42	0.00	0.02	0.07	0.00	0.00	0.00	0.00	0.35	0.00	0.00	0.02
C23	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.19	0.00	0.10	0.00	0.02	0.10	0.00	0.00	0.00	0.59	0.02	0.00
C24	0.00	0.05	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.13	0.00	0.32	0.00	0.04	0.17	0.06	0.03	0.00	0.00	0.00	0.00	0.16	0.01

by using one single image is low, the results can be useful in applications concerning the source apportionment. Here it is, however, important to develop sensible area estimates for different subcatchments. One problem in area estimation based on the classified image is that small classes will be overestimated whereas large classes often underestimated. Therefore, the observed area has to be corrected. By using the Monte Carlo method and information from the block database, unbiased estimate of the confusion matrix can be derived. Then it is possible to calculate an "area correction matrix", which can be used to obtain unbiased area estimates for the whole satellite scene.

The area estimation in the study area before and after correction is summarized in Table 8 for all types of units. Although the observed proportions based on one image and five images respectively, may differ significantly, the differences after corrections become small. For the single image case, one can see that the observed (classified) proportion for spring barley (C2), pasture on arable land (C12) and meadow (C15) was 13.2%, 19.5% and 2.16%, respectively. After correction the proportions became 29%, 34% and 0.27% respectively.

Table **8** Area estimation for all types of units

Class	5 seasonal scenes		one single scene	
	Observed/%	Corrected/%	Observed/%	Corrected/%
C1	1.41	0.89	0.45	0.81
C2	25.89	30.96	13.20	29.05
C3	11.06	11.60	13.05	10.64
C4	0.99	0.98	1.93	0.92
C5	0.17	0.06	0.70	0.04
C6	0.42	0.06	0.02	0.08
C7	0.95	0.09	0.33	0.08
C8	0.84	0.95	1.07	0.94
C9	0.37	0.24	2.32	0.31
C10	1.39	0.22	0.55	0.21
C11	1.47	1.30	2.60	1.05
C12	25.82	32.50	19.51	34.00
C13	0.61	0.32	4.43	0.32
C14	6.70	8.82	7.22	10.21
C15	0.33	0.21	2.16	0.27
C16	1.60	0.29	3.44	0.34
C17	14.86	7.35	13.98	7.35
C18	1.05	1.21	0.76	1.33
C19	1.25	1.24	1.66	1.28
C20	0.58	0.05	2.48	0.06
C21	0.49	0.07	3.66	0.05
C22	0.11	0.05	0.15	0.07
C23	0.12	0.05	1.45	0.07
C24	0.59	0.48	0.52	0.52
C25	0.93		2.38	

It is worth noticing that the above proportions are only valid at the scene level. In practice, often unbiased area estimates at local level are required, as well as the quality of classification at pixel level. Hence, the probabilistic classifier is derived and run on two small test sites. These two sites were selected with different characteristics of spatial distributions (homogeneous versus heterogeneous), each of size 16×16 pixels.

The probability for each class at pixel level and the entropy for each pixel were calculated. Fig. 1 and Fig. 2 present the map of entropy for test site 1 and test site 2, respectively. Note that they are based on five scenes and at thematic level 2. Higher values of entropy indicate uncertain classification at that pixel. By setting some threshold, we are able to identify areas where the classification accuracy is not satisfactory. In our test site 1, most of the entropies are very low, except a few pixels with medium size entropies. This indicates that the quality of classification at this

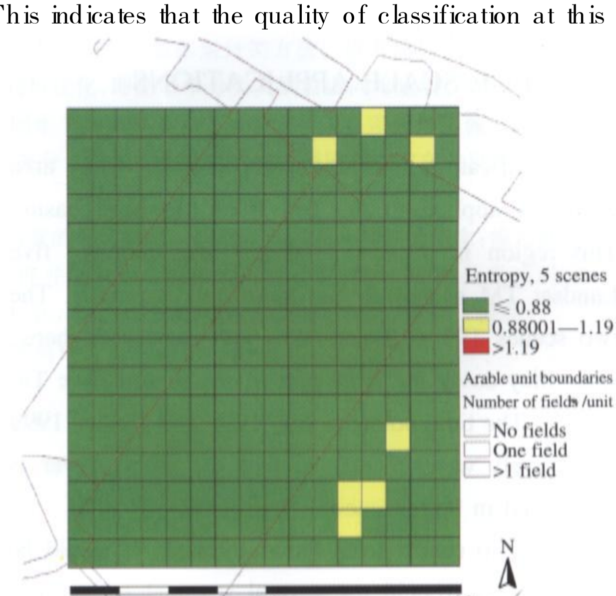

Fig. 1 Entropy map of test site 1

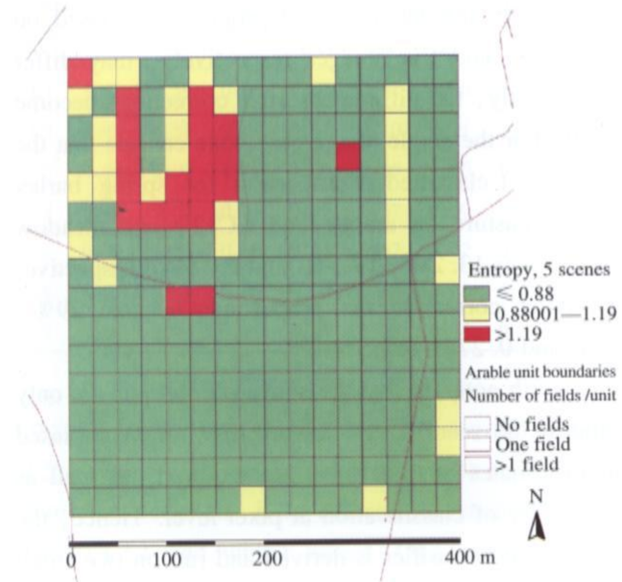

Fig. 2 Entropy map of test site 2

site is high. As also expected, the accuracy of classification at site 2 (heterogeneous area) became lower, so that one obtain a number of pixels with high entropies (red colored).

In areas having an unsatisfactory classification it is possible to take additional field plots and perform an improved classification. Here it is important that the classification method can handle different types of reference data. The final product will be two maps, one showing the pixel-wise classification and the other giving the classification accuracy at pixel level.

5 LARGE SCALE APPLICATIONS

The classification method developed in the study area was later applied to the Dal River drainage basin. This region is of size 29000km^2 and requires five Landsat TM scenes to be completely covered. The two scenes 194-018 and 196-017 were used here, since they cover the large part of the region (see Table 9). The time points at July 1999 and August 1999 respectively were chosen so that the cloud cover is minimized in both scenes.

According to results from the study area, it would be desirable if we could use several seasonal scenes also for the whole catchment. It was, however, impossible to get Landsat TM images to fulfill this requirement.

Table 9 Landsat images used in classification for the Dal River drainage basin

	Scene	Date	Spectral bands used
Landsat-5	194-018	1999-07-10	TM 2, 3, 4, 5
Landsat-5	196-017	1999-08-01	TM 3, 4, 5

Alternatively, one can consider using either SPOT images (10 meter resolution) or MERIS images (200 meter resolution) to handle the agricultural fields. The former demands much more work in preprocessing to make a useful mosaic spatially over the whole drainage basin as well as temporally over different seasons, whereas the latter has too low spatial resolution to distinguish different agricultural units.

The crops and agricultural land were predefined in the same way as in the study area. Note that the number of classes differs from those in the study area. This is simply because of the lack of field data in different scenes. Summarized results of our classification are shown in Table 10 for the thematic level 1 and Tables 11 and 12 for the thematic level 2.

Table 10 Probability matrices at level 1

	scene 194-018		scene 196-017	
	C1	C2	C1	C2
C1	0.82	0.18	0.81	0.19
C2	0.35	0.65	0.25	0.75

Table 11 Probability matrix for scene 194-018 at level 2

	C1	C2	C3	C4	C5	C6	C7	C8
C1	0.20	0.26	0.07	0.05	0.05	0.01	0.34	0.03
C2	0.08	0.48	0.05	0.02	0.03	0.02	0.26	0.04
C3	0.02	0.08	0.61	0.01	0.03	0.01	0.19	0.04
C4	0.03	0.02	0.24	0.21	0.07	0.00	0.38	0.04
C5	0.02	0.05	0.12	0.05	0.24	0.02	0.44	0.05
C6	0.02	0.19	0.01	0.00	0.06	0.36	0.32	0.04
C7	0.02	0.06	0.07	0.02	0.14	0.04	0.58	0.05

Table 12 Probability matrix for scene 196-017 at level 2

	C1	C2	C3	C4	C5	C6	C7	C8
C1	0.09	0.38	—	0.09	0.14	—	0.28	0.00
C2	0.02	0.49	—	0.09	0.09	—	0.29	0.02
C3	—	—	—	—	—	—	—	—
C4	0.01	0.26	—	0.34	0.08	—	0.27	0.04
C5	0.02	0.14	—	0.11	0.33	—	0.35	0.05
C6	—	—	—	—	—	—	—	—
C7	0.01	0.09	—	0.01	0.12	—	0.75	0.02

From Table 10, one can see that classification accuracies at thematic level 1 are similar and satisfactory for both scenes. Looking at Tables 11 and 12 for thematic level 2, two things are apparent and common in both scenes, that is C1 (autumn-sown cereals) is confused with C2 (spring-sown cereals), and C7 (other crops) is confused with all other classes. One of the reasons is that C2 and C7 are both dominating classes in arable land and grazing land respectively.

For the sake of space limitation, details of probability matrices at thematic level 3 for these two scenes are omitted here but they are available upon request.

REFERENCES

[1] Fix E, Hodges J L. Discriminatory Analysis-Nonparametric Discrimination: Consistency Properties. Report no. 4, US Air Force School of Aviation Medicine, Random Field, Texas, 1951.

[2] Yu J, Ekstrom M, Nilsson M. Image Classification Using Wavelets with Application to Forestry[A]. Proceedings of the 4th International Conference on Methodological Issues in Official Statistics[C]. Stockholm, 2000.

[3] Yu J, Ekstrom M. Multispectral Image Classification Using Wavelets: A Simulation Study[J]. Pattern Recognition, 2002, **36**(4): 889-898.

[4] Rajski C. A Metric Space of Discrete Probability Distributions [J]. Information and Control, 1961, **4**: 373-377.

[5] Rajski C. On the Normed Information Rate of Discrete Random Variables[A]. Trans. Third Prague Conference[C]. 1964.

[6] Ripley B D. Pattern Recognition and Neural Networks. Cambridge University Press, 1996.

[7] Ranneby B, Yu J. Remote Sensing Classification of Environmental Indicators [A]. Proceedings of the 30th International Symposium on Remote Sensing of Environment-Information for Risk Management and Sustainable Development[C]. Hawaii, 2003.

[8] Bezdek J C, Keller J M, Krishnapuram R, Pal N R. Fuzzy Models and Algorithms for Pattern Recognition and Image Processing. Kluwer Academic Publishers, 1999.

[9] Bereau M, Dubuisson B. A Fuzzy Extended k-nearest Neighbor Rule[J]. Fuzzy Sets and Systems, 1991, **44**: 17-32.

[10] Kissiov V T, Hadjitodorov S T. A Fuzzy Version of the k-nn Method[J]. Fuzzy Sets and Systems, 1992, **49**: 323-329.

[11] Loftsgaarden D O, Quesenberry C P. A Nonparametric Estimate of a Multivariate Density Function[J]. Annals of Mathematical Statistics, 1965, **36**: 1049-1051.

基于多时相影像的农业作物非参数与概率分类

俞　军，Bo Ranneby

(瑞典农业大学生物推测中心，SE-901 83于默奥，瑞典)

摘　要： 本文提出了一种新的结合多光谱和变化检测技术的多时相卫星数据集分类方法。该方法以数理统计中的最近邻法为基础，其目标函数是使得正确分类的平均概率得到最优化，即把每个分类类别看成同等重要。该新算法被应用于一个农业作物分类的研究区域，并利用覆盖该区的不同季节的 SPOT和 LANDSAT TM 多时像影像。结果表明，与单时像影像相比，使用五个不同季节的多时像影像可以充分地提高分类精度。为了说明该方法在大尺度范围内的效果，本文选取瑞典道拉河流域作为研究区。

由于不同地物的分布高度重叠，不可能得到像元水平上满意的分类精度。这就需要引进一种新的概念：像元概率分类法。基于像元的概率向量可用于判别传统分类法的可靠性并测量单个像元的不确定性(熵)。概率分类法同时提供了不同地物的面积的无偏估计，无论所感兴趣的区域的大小。这已经在不同特性的耕地试验点进行了检验。

关键词： 非参数分类；近邻方法；概率分类法；农业作物；质量评价；多时相影像；遥感；流域