Bayesian Soft Classification for Sub-Pixel Analysis: A Critical Evaluation

J.R. Eastman and R.M. Laney

Abstract

Soft classifiers defer the decision about the class membership of a pixel in favor of an expression of the degree of membership it exhibits in each of the land-cover classes under consideration. The reasons for using a soft classifier include the examination of classification uncertainty, but are most commonly directed to the potential of uncovering the proportional constituents of mixed pixels—a process called sub-pixel classification. *In this study we examine the assumptions and procedures of a commonly cited Bayesian soft-classification procedure for sub-pixel classification, and test its ability to uncover mixture proportions. The procedure involves the use of mixed-cover training sites to estimate the underlying class signatures through the development of fuzzy mean reflectances and covariance matrices. These are then used to evaluate the Bayesian* a posteriori *probability of belonging to each land-cover class. Using an artificial data set, it was found that this Bayesian soft-classification procedure is unable to uncover constituent class proportions unless substantial overlap exists in the distributions of parent classes. It was found that the use of fuzzy training sites improves the accuracy of this procedure, but not because of any special insights it offers into the underlying distributions, but rather, because of its tendency to increase the degree of overlap between parent distributions.*

Introduction

Classification in digital image processing has traditionally been concerned with the assignment of a specific and definite land-cover class to each pixel in the image. A decision rule is thus developed such that a *hard*, or *crisp*, assignment can be made for each pixel based on its spectral characteristics. It has always been understood that uncertainty surrounds this process, with the decision rule being designed to minimize the impact of such ambiguity. Recently, interest has focused on the nature of this uncertainty, with the explicit intention of extracting additional information based on its character.

Soft classification is one approach that evaluates and utilizes this uncertainty. Soft classification defers a hard decision in preference for some intermediate statement of class (set) membership. The reasons for doing so vary, including the explicit intention to assess classification uncertainty and the possibility of incorporating additional knowledge before a final *hardening* of the decision. The primary focus, however, has been its potential for *sub-pixel classification*—the determination of constituent classes that fall below the resolution of the pixel. In this paper we provide a critical evaluation of one of the major soft-classifier approaches to sub-pixel classification with the intention of better understanding the potential and limitations of the procedure.

Sub-Pixel Classification

The basis for sub-pixel classification resides with the fact that a solid-state detector integrates the intercepted radiance within its instantaneous field of view (IFOV). Regardless of the effective resolution of a detector, it is inevitable that the IFOV will frequently intercept reflected energy from more than one land-cover class. Such cases will be uncertain, with the expectation that the pixel will exhibit spectral characteristics that are intermediate between those characteristics of each of the *end-member* (true constituent) classes. Thus, for example, a pixel equally occupied by conifers and open water should exhibit reflection characteristics that combine the characteristics of the two underlying classes in equal proportion. Given the integrating nature of the detector itself, one would expect the pixel to exhibit spectral characteristics that represent an area-weighted average of these constituent parts.

Three major approaches have been investigated for sub-pixel classification. The first considers the reflectance characteristics of pixels to be an additive composite of a small number of end-member (constituent) classes. From this additive model, linear-mixture models determine the proportion of each constituent class by solving a set of simultaneous equations expressing the relationship between pixel reflectances and the unknown fractions of end-member constituents (Detchmendy and Pace, 1972; Settle and Drake, 1993). While this approach has generally proven to be fairly successful, it suffers from the limitation that the number of end-member constituents in a single pixel generally cannot exceed the number of image bands.[1]

A second, empirical approach estimates area proportions directly, based on models that have been trained on, or calibrated to, area proportions (Lewis *et al.*, 1999; Lewis *et al.*, 2000). Both statistical models, such as nearest-neighbor algorithms, and neural networks may be designed to fit within this approach.

A third approach uses soft-classification procedures to estimate the degree of membership that each pixel has in each of the end-member classes. The magnitude of membership acts

[1]If the restriction that the proportions must add to 1.0 is removed, the number of end-members can be one more than the number of bands. Some studies have included higher order moments derived from the training statistics in the models' equations in order to identify more classes than bands (Bosdogianni *et al.*, 1997; Faraklioti and Petrou, 2000).

J.R. Eastman is with the Graduate School of Geography, Clark University, 950 Main Street, Worcester, MA 01602 (REastman@clarku.edu).

R.M. Laney is with the Department of Geography, Sonoma State University, 1801 East Cotati Avenue, Rohnert Park, CA 94928-3609 (laney@sonoma.edu).

Photogrammetric Engineering & Remote Sensing
Vol. 68, No. 11, November 2002, pp. 1149–1154.

0099-1112/02/6811–1149$3.00/0

as a surrogate for proportions. Soft-classification procedures are sometimes called *Fuzzy Classifiers* because of their evaluation of the membership grade that each pixel possesses with respect to the constituent end-member classes. However, the term is somewhat of a misnomer because many different measures of set membership have been employed, including Bayesian *probabilities* (Wang, 1990a; Foody *et al.*, 1992), neural network *activation levels* (Foody and Aurora, 1996; Zhang and Foody, 2001), Dempster-Shafer *beliefs* (Eastman, 1997; Eastman, 1999), Fuzzy Set *possibilities* (Eastman, 1997; Eastman, 1999), and Fuzzy c-mean *partitioning* (Foody, 2000).[2]

A calibration procedure may be required to translate membership values to proportions (Foody, 2000), although some techniques assume that statements of set membership are directly proportional to the relative area occupied by each end-member class within the pixel (Fisher and Pathirana, 1990). Those that estimate *a posteriori* probabilities, such as Bayesian probabilities, typically assume a direct linear relationship with area proportions (e.g., Wang, 1990a).

Within the soft-classification approach, class membership may or may not be subject to the constraint that a pixel's total membership to all possible classes adds to one. If the constraint holds (e.g., fuzzy-c-means or Bayesian probabilities), classification accuracy is significantly compromised by the potential presence of unknown classes (Foody, 2000). If the constraint is relaxed (e.g., possibilistic c-means or Dempster-Shafer beliefs), unknown classes will have less of an impact on classification accuracy (Foody, 2000).

In this study we focus on the use of Bayesian soft classification, as set out by Wang (1990a; 1990b) and implemented in the IDRISI software system.

Bayesian Soft Classification and Fuzzy Signature Development

In general, the soft-classification process involves two stages: a training stage in which a statistical characterization is developed of the expected reflectance values for each class over all bands, and a classification stage where the grade of membership that each pixel possesses in each class is assessed.

In theory, the training stage should lead to a statistical characterization of the pure end-member classes. In practice, however, this is difficult to achieve because the training site pixels will also likely contain mixtures. Wang (1990a) therefore proposed the use of a procedure that directly accommodates the impurity of training site pixels through a consideration of the probability of fuzzy events (Zadeh, 1968). In effect, he develops a weighted mean and a weighted variance/covariance matrix, using the estimated fuzzy membership grades of training sites as the weights. Thus, the fuzzy mean is calculated as

$$\mu_c = \frac{\sum_{i=1}^{n} f_c(x_i)x_i}{\sum_{i=1}^{n} f_c(x_i)} \tag{1}$$

where n is the total number of pixels in the training site, x_i is a sample pixel measurement vector, and f_c is the membership function of class c (the proportion of cover class c in the training site sample).

Similarly, the fuzzy covariance matrix is calculated as (Wang, 1990a)

$$\Sigma_c^* = \frac{\sum_{i=1}^{n} f_c(x_i)(x_i - \mu_c)(x_i - \mu_c)^{\mathrm{T}}}{\sum_{i=1}^{n} f_c(x_i)}. \tag{2}$$

Once the training statistics have been developed, the classification stage can be undertaken. In Wang's (1990a) procedure, class membership is defined on the basis of the underlying logic of maximum likelihood classification, but using the fuzzy means and fuzzy variance/covariance matrix as determined above: i.e.,

$$f_c(x) = \frac{p_c(x)}{\sum_{i=1}^{m} p_i(x)} \tag{3}$$

where f_c is the membership function of class c (the proportion of cover class c in the training site sample) and $p_i(x)$ is the normal probability density for class i,
and

$$p_i(x) = \frac{1}{(2\pi)^{N/2}|\Sigma_i^*|^{1/2}} e^{-[(x-\mu_i)^{\mathrm{T}} \Sigma_i^{*-1}(x-\mu_i)]/2} \tag{4}$$

where Σ_i^* is the fuzzy variance/covariance matrix and μ_i is the fuzzy mean.

Wang's fuzzy membership is thus equivalent to the *a posteriori* probability of class membership for the case of equal *a priori* probabilities.

In the IDRISI system, the FUZSIG module can be used to create signature statistics incorporating the fuzzy means and fuzzy variance/covariance matrices suggested by Wang (1990a). The BAYCLASS module can subsequently be used to create images of the *a posteriori* probability of class membership for each class. The procedure does allow the incorporation of *a priori* probabilities. However, for the purposes of this study, *a priori* probabilities were kept equal for all classes.

This concept is illustrated further in Figure 1. The figure shows the conditional probability density functions of two classes (A and B) along a single band of data. The likelihood that the image pixel at x belongs to each class is determined by calculating the conditional probability density at x for each class (i.e., the height of each curve at x), multiplied by its *a priori* probability, normalized by the sum of such results over all classes (Foody *et al.*, 1992).

Considerable uncertainty exists in the direct relationship between probabilities and proportions. Uncertainty may arise from the statistical inseparability of signatures, or from the presence of unknown cover types (which Bayesian probability theory does not recognize and thus must apportion to known

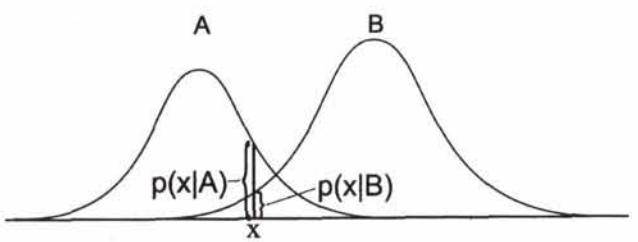

Figure 1. Conditional probability functions of classes A and B at x.

[2]Each of these expressions of intermediate set membership belongs to a larger family of *Fuzzy Measures* (Dubois and Prade, 1982). Thus, in this sense, soft classifiers can justifiably be called fuzzy.

classes). Uncertainty may also be introduced by the Point Spread Function of the sensor, which commonly leads to higher sensitivity (and thus enhanced representation) near the center of the IFOV (Cracknell, 1998; Manslow and Nixon, 2000).[3]

Note that, because the conditional probability density functions are assumed to be normal, they actually extend to infinity, allowing the *a posteriori* probability to be evaluated no matter how far it is from the central value of any end-member class. In *hard* maximum-likelihood classification, it is common practice to set a threshold chi-square value that, when exceeded, indicates the pixel may belong to an unspecified, or unknown, class. In soft Bayesian classification, however, end-members may lie far from each other in spectral space, in which case pixels containing their mix will necessarily lie at the far tails of the end-member probability curves. Thresholds typical to hard classification procedures no longer apply.

Assumptions of the Bayesian Classification Procedure

It is important at this stage to review a series of important assumptions about the use of the fuzzy soft-classification procedures described above for sub-pixel classification.

- Unlike the methodology of linear spectral unmixing, fuzzy classifiers based on a decomposition of the Bayesian evaluation of evidence assume (and even require) that there be within-class variability in reflectance. As a consequence, the constituents of a mixture should probably not be called end-members. We will therefore call these *parent* classes.
- It is assumed that the fuzzy means vector and fuzzy variance/covariance matrix for each parent class describe the underlying *pure* parent class distributions (Wang, 1990a: p. 197; Foody, 1999: p. 446). Pure in this instance implies homogeneity of class contribution rather than an absence of variability.
- By the logic of the Bayesian maximum-likelihood rule, it is assumed that all possible parent classes are known, and have been characterized in the training stage.
- It is assumed that the membership grades are linearly related to the *a posteriori* probabilities of class membership.

Potential Problems

Although the Bayesian classification and fuzzy signature development approach to sub-pixel classification described above has been widely described (e.g., Jensen, 1996; Foody, 1999), some immediate questions arise which suggest that a closer examination is needed.

- The second assumption, that the fuzzy mean vector and fuzzy variance/covariance matrix for each class describe the underlying *pure* parent classes, is generally hard to achieve. Despite the widely held view that the fuzzy signature procedure can produce *pure* parent class statistics without need of pure training sites (Foody, 1999; Jensen, 1996; Wang, 1990a; Wang, 1990b), consideration of the procedure would suggest otherwise. The procedure described by Wang (1990a; 1990b) for estimating the fuzzy mean vector and fuzzy variance/covariance matrix cannot lead to unbiased estimates of parent class signatures. Equation 1 is a weighted average. As a result of the averaging process, impure training sites will cause the means to migrate towards their impure constituents. Impurities thus bias the mean vectors—the larger the impurity (i.e., mixture), the larger the degree of bias. Similarly, the fuzzy variance/covariance matrix is calculated through a set membership weighting procedure that should also serve to degrade, rather than enhance, the purity of the training site statistics.
- The presence of any additional classes intermediate to two mixing classes in spectral space will cause the mixture to be misidentified. For example, assume that a pixel truly contains a mixture of only two parent classes (A and B). The means of all mixed pixels will occur along a line joining A and B in spectral space. If no other class mean occurs between A and B, the procedure should be able to uncover the proper sub-pixel composition. If, however, a third class (C) were located along that line, the mixes would erroneously be interpreted as mixes between A and C or B and C, but never A and B as they truly are. While non-fuzzy signature development and *hard* maximum-likelihood classification procedures are susceptible to this problem as well, it becomes especially severe in soft signature development and Bayesian classification procedures. These techniques elongate the probability curves between parent classes that lie far from each other in spectral space, thereby increasing the likelihood that these curves will intercept a third intermediate class.
- The normal distribution of the conditional probability function has tails that extend to infinity, allowing the potential evaluation of mixtures even for widely separated parent classes. For example, in Figure 1, the distributions of A and B substantially overlap. However, when spectral distances between classes are great, the probabilities associated with their mix will be very small, and based on the elongated tails of the probability curves more than on actual training information. Moreover, the small probabilities will reach beyond the precision of the digital representations of floating point real numbers in the computer. Thus, the probability of the mix will be considered zero.

These observations cast significant doubt on the ability of the soft-classification procedure to correctly identify mixtures in all cases. As a result, the following empirical evaluation was undertaken to test the quality of the estimate derived from this interesting procedure.

An Empirical Evaluation

Methodology

In order to evaluate the soft-classifier approach to sub-pixel evaluation, an image data set of completely known character was required. We chose to create an artificial four-band set of images representing a landscape composed of only two cover types, deciduous (mixed hardwood) forest and white pine forest. These two cover types were chosen because they have spectral signatures that are quite different, and because they commonly occur in both pure and mixed stands in actual landscapes. The images were then aggregated to a coarser resolution by integrating the reflectances of the higher resolution pixels. This yielded a new four-band set with perfectly known sub-pixel mixtures of the two cover types.

The data set was then classified using Wang's (1990a) fuzzy classification procedure by means of the FUZSIG and BAYCLASS procedures in IDRISI. The output, indicating the degree of membership of each pixel in the two cover classes, was then compared to the actual proportional content. In order to test the sensitivity of the technique to the training stage, the classification procedure was repeated four times using different proportional mixes in the fuzzy signatures.

Creation of the Artificial Data Set

The spectral signatures of the two parent cover types (deciduous and white pine forest) were obtained from large, homogeneous training sites within a Landsat TM image of central Massachusetts which were verified through site visits. These training statistics (mean vector and variance/covariance matrix for each cover type) were then used to construct a four-band set of images (typifying Landsat bands 2, 3, 4, and 5). In order to replicate the means and variances/covariances of the two cover types, random draws from corresponding multivariate normal distributions were made for each band at each pixel location. The software used for this purpose was *Crystal Ball* (Decisioneering, Inc., 2000). This stage yielded two sets of four bands

[3]IDRISI's BAYCLASS outputs a classification uncertainty image, which, in the context of this paper, would be interpreted to express the degree of mixing present, but in a real world example could incorporate any of these sources of uncertainty.

TABLE 1. NUMBER OF PIXELS REPRESENTING EACH PROPORTION OF THE DECIDUOUS COVER CATEGORY

Proportion	Sample Size
0%	985
11%	710
22%	576
33%	525
44%	579
56%	599
67%	710
78%	600
89%	489
100%	627

TABLE 2. PROPORTIONAL MIXES USED TO DEVELOP THE FOUR SETS OF FUZZY SIGNATURES

Signature	Percentage of cover in mix	
Set A	Deciduous	White pine
Fuzzy deciduous	100	0
Fuzzy white pine	0	100
Set B		
Fuzzy deciduous	89	11
Fuzzy white pine	11	89
Set C		
Fuzzy deciduous	78	22
Fuzzy white pine	22	78
Set D		
Fuzzy deciduous	54	45
Fuzzy white pine	45	55

(one set for each cover type) expressing the expected pixel values if the entire image were covered by only one cover type.

Next, the two sets of images were combined into a single set, where the new set contained a spatial mix (but not a spectral mix at this stage) of the deciduous and white pine covers. To create this new set, a Boolean image was made in a patchy and mottled spatial pattern. A second reverse Boolean image was also created. Each band in the deciduous set was multiplied by the first Boolean image while the white pine bands were multiplied by the reverse Boolean image. Corresponding bands were then added, creating a single set of images.

Finally, each band was aggregated by a factor of three such that each pixel in the new set represented an average of nine pixels in the original set.[4] This new set thereby contained a weighted average of the two cover types, according to the proportion of each cover type represented in the nine pixels. The final set of images embodied ten proportional mixes of the two covers, with each mix represented by at least 489 pixels (Table 1). The Boolean images were also aggregated, producing *verification images* indicating the actual proportion of each cover type in each pixel.

Signature Development Stage

Fuzzy signatures were developed using the FUZSIG module of the IDRISI software system. Four sets of fuzzy signatures were produced, with each set based on a different proportional mix of parent covers within the training pixels.

Two training sites were used to develop each set of signatures, with the two sites always containing complementary proportions of the two cover types. For example, to create a set of fuzzy signatures based on a 22%/78% proportional mix, the first training site (called white pine) included pixels with a 22 percent deciduous and 78 percent white pine mix, and the second training site (called deciduous) included pixels with a 78 percent deciduous and 22 percent white pine mix. Four proportional mixes were used to create the four fuzzy signature sets: 0%/100%, 11%/89%, 22%/78%, and 44%/55% (Table 2). The first of these is clearly non-fuzzy, and was used as a control.[5]

Classification Stage

After creating the signature statistics, the artificial data set was classified four times using the BAYCLASS soft classifier in IDRISI. Each run produced a separate image for each cover class, where pixel values represent the probability of the pixel belonging to the cover class (and thus, by interpretation, the relative areas occupied by each cover class within each pixel). Output is presented here for the deciduous cover class only because the output for white pine is its complement.

Results are illustrated in cross-tabular form—a close corollary to a classification error matrix (Tables 3 through 6). Columns indicate the actual proportion of deciduous cover in the pixel. Rows represent the probability of the pixel belonging to the cover class, as determined by BAYCLASS. Matrix cells report the percentage of pixels within each *actual* proportion of cover (i.e., columns sum to 100). In the ideal case, 100 percent of pixels in each image mixture class would be classified as possessing the same mixture percentage: i.e., the principal diagonal would contain values of 100 while all other cells would contain 0.

Results

Table 3 shows the results of the soft classification based on signatures with a 0%/100% training site mix (i.e., the non-fuzzy control). As can readily be seen, the procedure performed very poorly. In fact, it was not able to uncover any of the true mixtures that existed in the data. The fact that many true mixture combinations were left unclassified by the BAYCLASS procedure suggests that the two cover types were so different that their conditional probability distributions did not overlap enough to yield calculated probabilities that exceeded the precision of the data.

Fuzzy signatures based on the 11%/89% mix in the training sites (Table 4) produced more pixels with partial membership to deciduous, but again there were high rates of inconsistent and inaccurate membership assignment. Pixels with 67 percent and 78 percent actual deciduous were very

TABLE 3. RESULTS BASED ON 0%/100% (NON-FUZZY) PROPORTIONAL MIX, WHERE UNCL = UNCLASSIFIED PIXELS

Crosstabulation of actual proportion of deciduous (columns) against predicted proportion (rows)										
	0%	11%	22%	33%	44%	56%	67%	78%	89%	100%
0%	**100**	100	17							
11%			17							
22%										
33%										
44%										
56%										
67%										
78%										
89%										
100%							1	54	100	**100**
uncl			66	100	100	100	99	46		

[4]Aggregation by a factor of three provided a range of proportional mixes great enough to analyze the sensitivity of the procedure to those mixes, and yet restricted enough to simplify the interpretation of the results.

[5]The result of using the non-fuzzy signatures with the FUZSIG module was identical to that achieved using the conventional procedure of developing signature statistics.

TABLE 4. CLASSIFICATION RESULTS BASED ON FUZZY SIGNATURES WITH 11%/89% PROPORTIONAL MIX

Crosstabulation of actual proportion of deciduous (columns) against predicted proportion (rows)

	0%	11%	22%	33%	44%	56%	67%	78%	89%	100%
0%	**100**	94	49	15	5	1				
11%		**6**	44	31	5	1	1			
22%			**6**	32	15	1				
33%			1	**15**	16	1				
44%				5	**21**	7				
56%				2	23	**12**	1			
67%					12	22	**3**			
78%					2	27	6	**1**		
89%					1	26	27	1		
100%						2	62	98	100	**100**

TABLE 6. RESULTS BASED ON FUZZY SIGNATURES BASED ON 45%/55% PROPORTIONAL MIX

Crosstabulation of actual proportion of deciduous (columns) against predicted proportion (rows)

	0%	11%	22%	33%	44%	56%	67%	78%	89%	100%
0%	**96**	7							6	88
11%										
22%	4	93	**45**							
33%			55	**84**	1					
44%				16	**94**	7				
56%					5	**93**	19			
67%							**81**	65		
78%								**35**	91	6
89%										
100%									3	**6**

poorly classified, with less than 10 percent of these pixels assigned membership probabilities within 22 percent of the actual proportion. For all other mixes, however, at least 40 percent of all pixels were assigned within 22 percent of the actual proportion.

Significant improvement resulted with the fuzzy signatures based on the 22%/78% training mix (Table 5). At least 84 percent of all pixels were assigned membership probabilities within 22 percent of the actual proportion. Even higher association between actual cover proportion and predicted membership probability was realized with the fuzzy signatures based on a 45%/55% training mix (Table 6), with one notable exception. Eighty-eight percent of pixels of 100 percent deciduous were assigned a membership probability of 100 percent white pine.

Discussion

What is immediately clear from Tables 3 through 6 is that the degree of fuzziness in the training site data makes a substantial difference in the success with which the procedure can uncover sub-pixel composition. However, the reason for this has nothing to do with its success in uncovering the true nature of the parent classes, but rather, with the degree to which the calculation of the fuzzy means and variance/covariance matrices causes the conditional probability distributions of the parent classes to overlap. Table 7 tabulates the calculated fuzzy means for each of the training site mix combinations of the deciduous category, while Table 8 presents the corresponding variance/covariance matrices. It can readily be seen that increased fuzziness in the training site data causes the means to migrate towards those of the contaminating class (white pine in this case), thereby drawing the conditional probability distributions together. In addition, it can be noted from Table 8 that variances/covariances are increased by training site fuzziness, thereby enhancing the degree to which the conditional probability distributions overlap.

Overlap in the conditional probability distributions of the parent classes would thus seem to be essential if sub-pixel composition is to be uncovered. When they did not (0%/100% mix; Table 3), no sub-pixel information could be recovered. In addition, the accuracy improved enormously as the degree of overlap became substantial (22%/78% mix, Table 5; and 45%/55% mix, Table 6). However, when the two distributions approach coincidence of their means (45%/55% mix, Table 6), the procedure becomes unstable in its estimation of pure pixels.

From Tables 7 and 8, it can also be noted that the fuzzy mean and variance/covariance calculation does not recover the underlying *true* parent class distributions (the 0%/100% mix). In fact, in the case of the 45%/55% mix, the parent class statistics were inaccurate enough to allow pure deciduous pixels to appear as pure white pine (Table 6). Yet, with moderate training mixes, the fuzzy means and fuzzy variance/covariance calculations do lead to a remarkably accurate determination of sub-pixel composition. It would appear, then, that the utility of the fuzzy mean and fuzzy variance/covariance procedure is in enhancing the degree of overlap in the conditional probability distributions, and not in uncovering the true underlying parent class characteristics. In addition, it is clear that the use of mixed training sites is vital to the success of this procedure.

Finally, it should be noted that this research does not consider the case in which another unique class may occupy the overlap region between the true parent class distributions, described earlier as the second of the Potential Problems. Nevertheless, the implications of this problem are significant. Any procedure that evaluates a mixed pixel's position in spectral space must take into account the possibility of a third class occupying a similar place.

TABLE 5. CLASSIFICATION RESULTS BASED ON FUZZY SIGNATURES WITH 22%/78% PROPORTIONAL MIX

Crosstabulation of actual proportion of deciduous (columns) against predicted proportion (rows)

	0%	11%	22%	33%	44%	56%	67%	78%	89%	100%
0%	**100**	64	13	5	1					
11%		**36**	66	11	3					
22%			**21**	50	8	1				
33%				**32**	29	3	1			
44%				2	**49**	17	1			
56%					10	**40**	4			
67%						37	**25**	1		
78%						2	59	**20**		
89%							9	76	**31**	1
100%								3	69	**99**

Conclusions

From this very controlled experiment, it can be concluded that the fuzzy signature/Bayesian soft-classifier approach to sub-pixel decomposition can indeed work, but with important limitations. It is critical that the conditional probability distributions (from the training site data) overlap over the complete

TABLE 7. MEANS OF FOUR DECIDUOUS SIGNATURES, BASED ON DIFFERENT PROPORTIONAL MIXES (DECIDUOUS/WHITE PINE), AND MEAN OF WHITE PINE SIGNATURE

Proportional Mix	Band 1	Band 2	Band 3	Band 4
100% deciduous	26.0	21.0	114.1	71.0
89%/11%	25.8	21.1	101.5	64.2
78%/22%	25.7	21.1	95.4	61.1
56%/44%	25.5	21.2	87.1	56.6
100% white pine	25.0	21.3	59.4	55.9

TABLE 8. COVARIANCE MATRICES FOR EACH PROPORTIONAL MIX (DECIDUOUS/WHITE PINE)

100%/0% proportional mix				
	Band 1	Band 2	Band 3	Band 4
Band 1	0.064	0.0032	−0.0079	0.034
Band 2		0.089	0.00071	0.039
Band 3			3.62	1.28
Band 4				1.64
89%/11% proportional mix				
	Band 1	Band 2	Band 3	Band 4
Band 1	0.23	0.045	4.55	2.64
Band 2		0.32	−0.73	−0.056
Band 3			238.96	128.42
Band 4				70.97
78%/22% proportional mix				
	Band 1	Band 2	Band 3	Band 4
Band 1	0.24	0.061	3.14	1.95
Band 2		0.39	−0.38	0.32
Band 3			158.34	86.07
Band 4				48.96
56%/44% proportional mix				
	Band 1	Band 2	Band 3	Band 4
Band 1	0.29	0.16	0.48	0.74
Band 2		0.53	0.43	1.12
Band 3			12.88	7.81
Band 4				7.51

range of mixtures. This will commonly not be the case with real world data, particularly with reasonable training sites. Thus, some intervention is needed to produce the required overlap. Using heterogeneous training sites and fuzzy signature development can help achieve this, but not in the manner that has previously been assumed. The weighting procedure used to create fuzzy means and fuzzy variance/covariance matrices does not uncover the *pure* parent distribution parameters, but rather, serves only to cause further overlap in the distributions. Finally, cases where one or more additional classes exist in an intermediate spectral position between the two parent classes of a true mixture will cause the procedure to fail to uncover the proper mixture. Thus, we can expect that, under real world conditions, the success of this procedure for sub-pixel classification will be highly variable and not easily controlled.

References

Bosdogianni, P., M. Petrou, and J. Kettler, 1997. Mixture models with higher order moments, *IEEE Transactions on Geoscience and Remote Sensing*, 35(2):341–353.

Cracknell, A.P., 1998. Synergy in remote sensing—What's in a pixel?, *International Journal of Remote Sensing*, 19(11):2025–2047.

Decisioneering, Inc., 2000. *Crystal Ball 4.0*, Decisioneering, Inc., Denver, Colorado (software).

Detchmendy, D.M., and W.H. Pace, 1972. A model for spectral signature variability for mixtures, *Proceedings of the Conference on Earth Resources Observations and Information Analysis*, 13–14 March, Tullahoma, Tennessee, pp. 596–620.

Dubois D., and H. Prade, 1982. A class of fuzzy measures based on triangular norms: A general framework for the combination of uncertain information, *International Journal of General Systems*, 8(1):43–61.

Eastman, J.R., 1997. *IDRISI for Windows. Version 2.0*, Clark University, Worcester, Massachusetts, 306 p.

———, 1999. *Idrisi32*, Volume 2, Clark University, Worcester, Massachusetts, 170 p.

Faraklioti, M., and M. Petrou, 2000. Recovering more classes than available bands for sets of mixed pixels in satellite images, *Image and Vision Computing*, 18:705–713.

Fisher, P.F., and S. Pathirana, 1990. The evaluation of fuzzy membership of land cover classes in the suburban zone, *Remote Sensing of Environment*, 34:121–132.

Foody, G.M., 1999. The continuum of classification fuzziness in thematic mapping, *Photogrammetric Engineering & Remote Sensing*, 65(4):443–451.

———, 2000. Estimation of sub-pixel land cover composition in the presence of untrained classes, *Computers & Geosciences*, 26:469–478.

Foody, G.M., and M.K. Arora, 1996. Incorporating mixed pixels in the training, allocation and testing of supervised classifications, *Pattern Recognition Letters*, 17:1389–1398.

Foody, G.M., N.A. Campbell, N.M. Trodd, and T.F. Wood, 1992. Derivation and application of probabilistic measures of class membership from maximum-likelihood classification, *Photogrammetric Engineering & Remote Sensing*, 58(9):1335–1341.

Jensen, J., 1996. *Introductory Digital Image Processing: A Remote Sensing Perspective*, Prentice-Hall, Inc., Upper Saddle River, New Jersey, 316 p.

Lewis, H.G., M.S. Nixon, A.R.L. Tatnall, and M. Brown, 1999. Appropriate strategies for mapping land cover from satellite imagery, *Proceedings of the 25th Annual Conference of The Remote Sensing Society*, 08–10 September, Cardiff, United Kingdom, pp. 717–724.

Lewis, H.G., M. Brown, and A.R.L. Tatnall, 2000. Incorporating uncertainty in land cover classification from remote sensing imagery, *Advanced Space Research*, 26(7):1123–1126.

Manslow, J., and M. Nixon, 2000. On the representation of fuzzy land cover classifications, *Proceedings of The 26th Annual Conference of the Remote Sensing Society*, 12–14 September, Remote Sensing Society, Leicester, United Kingdom (unpaginated CD-ROM).

Settle, J.J., and N.A. Drake, 1993. Linear mixing and the estimation of ground cover proportions, *International Journal of Remote Sensing*, 14:1159–1177.

Wang, F., 1990a. Fuzzy supervised classification of remote sensing images, *IEEE Transactions on Geoscience and Remote Sensing*, 28(2):94–201.

———, 1990b. Improving remote sensing image analysis through fuzzy information representation, *Photogrammetric Engineering & Remote Sensing*, 56(8):1163–1169.

Zadeh, L.A., 1968. Probability measures of fuzzy events, *Journal of Mathematical Analysis and Application*, 10:421–427.

Zang, J., and G.M. Foody, 2001. Fully-fuzzy supervised classification of sub-urban land cover from remotely sensed imagery: Statistical and artificial neural network approaches, *International Journal of Remote Sensing*, 22(4):615–628.

(Received 17 July 2001; accepted 21 March 2002; revised 10 April 2002)