

Neurocomputational Spatial Uncertainty Measures

Zhe Li

Graduate School of Geography
Clark University
950 Main Street
Worcester, MA 01610

Phone: 1 (508) 849-2303

Fax: 1 (508) 793-8842

Email: zheli@clarku.edu

Abstract:

Both the Self-Organizing Map (SOM) and fuzzy ARTMAP neural network are trained based upon the competitive mechanism and use the “winner-take-all” rule. Previous studies developed soft classification algorithms for the SOM. This paper introduces the idea and proposes non-parametric measures for the fuzzy ARTMAP computational neural networks to handle spatial uncertainty in remotely sensed imagery classification. These soft algorithms are neuron-triggering/committing-frequency based and are grouped into two types, i.e., *Commitment* and *Typicality*, expressing in the first case the degree of commitment a classifier has for each class for a specific pixel and in the second case, how typical that pixel’s reflectances are of the ones upon which the classifier was trained for each class. Two measures are designed for each of the two neural network models, i.e., *SOM Commitment* (SOM-C) vs. *SOM Typicality* (SOM-T) and *ART Commitment* (ART-C) vs. *ART Typicality* (ART-T). To evaluate these proposed algorithms, soft classifications of a SPOT HRV image around Westborough, Massachusetts were undertaken. Conventional soft classifiers such as Bayesian posterior probability classifier and Mahalanobis typicality classifier were used as a comparison. Correlation Analysis and Principal Components Analysis (PCA) were employed to explore the relationship between these different measures. Results indicate that great similarities exist among the ART-C, SOM-C and the Bayesian posterior probability classifier, and significant similarities exist among the ART-T, SOM-T and the Mahalanobis typicality classifier. However, the proposed models outperformed conventional ones.

1 Introduction

Recently soft classification has become an attractive means of land cover classification from remotely-sensed imagery (Bernard et al. 1997). Soft classifiers determine the class type of a pixel by using an expression of the degree of membership it exhibits in each of the land-cover classes under consideration (Eastman and Laney 2002). They allow the class membership of a pixel to be partitioned among the classes. Thus, rather than being allocated a single class as in a conventional hard classification, a pixel can have a variable degree of membership to all classes (Foody 1998). Soft classifiers are used not only for potential of uncovering the proportional constituents of mixed pixels (Eastman et al. 2002; Foody 1999), but also for the examination of classification uncertainty. Much effort has been directed in the last two decades to develop soft classification algorithms for remotely sensed data. These approaches include soft outputs based on Bayesian posterior probabilities (Foody 1992; Eastman and Laney 2002), Mahalanobis typicalities (Foody et al. 1992), Fuzzy Set membership grades (Eastman 2003), Linear Mixture fractions (Settle and Drake 1993), Dempster-Shafer beliefs (Eastman et al. 2002; Eastman et al. 2005) and Neural Networks activation levels (Atkinson et al 1997; Zhang and Foody. 2001).

Conventional soft classifiers such as those based on Bayesian and Mahalanobis constructs are parametric and limited by the assumptions about the form and distribution of input data (Foody et al. 1992). In recent years machine learning algorithms have emerged as effective non-parametric alternatives to conventional parametric algorithms when dealing with complex measurement spaces (Rogan et al, in press; Rogan and Chen, 2004). Among machine learning techniques, there has been considerable interest in the use of neural networks for the classification of remotely sensed imagery due to their numerous advantages over conventional classifiers such as no assumptions about the form and distribution of input data, non-linear decision boundaries and the capabilities of generalizing inputs and learning of complex patterns (Foody 1992, 1996, 1997, 1998; Atkinson et al. 1997; Liu et al. 2003; Tso and Mather 2001).

From the range of network types suitable for classification applications, the Multi-Layer Perceptron (MLP) network trained with feed forward/back-propagation is still the most widely used in remote sensing image classification (Foody 2004). In addition, the MLP is capable of producing both hard and soft outputs. However, the selection of an appropriate network topology and learning properties (e.g., learning rate and momentum factor) is difficult and subjective (Zhang et al. 2001; Tso et al. 2001; Foody. 2004; Qiu 2004). Moreover, for feed-forward neural networks such as MLP, it is impossible for them to incorporate new information without forgetting past learning (Carpenter 1989). Most importantly, MLP using back-propagation is potentially unstable because of its tendency to find sub-optimal solutions by getting trapped in a local minimum and over fitting (Carpenter et al. 1991; Foody 1997). In some cases, an acceptable convergence (evaluated by the RMS between predicted outputs and expected outputs) is difficult to achieve with the underlying algorithm.

Although there have been improvements in the use of MLP (Verbeke et al. 2004; Foody 2004), attention has been increasingly focused on other neural network models such as Radial Basis Function networks (RBF) (Rollet et al. 1998; Foody 2004), Kohonen's Self-Organising Map (SOM) (Ji 2000), Hopfield networks (Tatem 2002; 2003), pattern recognition techniques based on Adaptive Resonance Theory (ART/ARTMAP) (Carpenter et al. 1992; Rohwer et al. 1994; Tso et al. 2001) and Decision Tree classifiers (Breiman et al. 1984; Simard et al. 2002; Pal and Mather. 2003; Rogan et al. 2003; Yang et al. 2003). As neural approaches, Fuzzy ARTMAP (Carpenter et al. 1992) have not been explored as thoroughly as other neural network models,

although many experiments were conducted on remotely sensed imagery classification using the fuzzy ARTMAP (Mannan et al. 1998; Gopal et al. 1999; Seto and Liu 2003). Mannan et al. (1998) found that fuzzy ARTMAP were more efficient than MLP with back propagation learning in supervised classification of multi-spectral remotely sensed images. Rogan et al. (in press) conducted an experiment on mapping land cover modification over large areas by comparing different machine learning algorithms including Decision Tree (DT), Fuzzy ARTMAP and suggested that Fuzzy ARTMAP show significant promise in the classification of remotely sensed imagery.

This paper is an extension of a previous study on the commitment and typicality measures for the Self-Organizing Map (Li and Eastman 2006) and is concerned with the development of soft classification procedures for the Fuzzy ARTMAP neural network model. Although a few studies have explored mixed pixel analysis using Fuzzy ARTMAP (Liu et al. 2004; Carpenter et al. 1999; Liu et al. 2004; Ju 2003), soft classification for uncertainty analysis has not been undertaken. The objective of this study is to develop soft classification algorithms for the Fuzzy ARTMAP neural network model, i.e., *ART Commitment* and *ART Typicality* to handle uncertainty analysis in remotely sensed image classification.

2 An Overview of the Fuzzy ARTMAP Procedure

Adaptive Resonance Theory (ART) based neural networks have evolved from the biological theory of cognitive information processing and have been developed by Grossberg and Carpenter (Carpenter et al. 1991). ART networks are designed, in particular, to resolve the stability-plasticity dilemma: they are stable enough to preserve significant past learning, but nevertheless remain adaptable enough to incorporate new information whenever it might appear (Carpenter 1989). A comprehensive description of ART models is detailed in many literatures (Carpenter et al. 1991, 1992; Mannan et al. 1998; Gamba and Dell'Aqua 2003). Fuzzy ART is a clustering algorithm that operates on vectors with analog-values elements.

Fuzzy ART, which is for unsupervised classification, has two layers, i.e. F_1 (input layer) and F_2 (category layer). These two layers make up of ARTa model (Figure 1). F_1 layer represents the input feature vector and thus has neurons for each measurement dimension. Fuzzy ART contains $2n$ input for the complement coded input feature vector to preserve amplitude information (Carpenter et al. 1991; Mannan et al. 1998)

The number of F_2 layer neurons of fuzzy ART is automatically determined, i.e., F_2 layer begins with a single neuron and dynamically grows in number during the procedure of learning.

Fuzzy ARTMAP, designed for supervised classification, has two additional layers, i.e. map field layer and out put layer. The latter makes up of ARTb model (Figure 1). The map field layer connects ARTa and ARTb models. The output and map field layers consist of m neurons each, where m is the output class dimension. There exists a one-to-one connection between these two layers. Figure 1 illustrates the basic architecture of a Fuzzy ARTMAP model.

During the supervised training stage of the fuzzy ARTMAP, every input pattern will compare with existing neurons to determine if it is similar enough to one of the existing patterns. If yes, the neuron associated with this input pattern will be selected as the winner, often known as “to become committed”, which is based on the degree of similarities between the input vector and the F_2 neurons, i.e., the winner is the neuron that has the maximum similarity with the current input feature vector, which is determined using a fuzzy intersection operation, i.e.

$$Winner = \arg \max_j \left(\frac{|x_j \wedge w_j|}{\alpha + |w_j|} \right) = \arg \max_j \frac{\sum_{i=1}^n \min(x_j, w_j)}{\alpha + \sum_{i=1}^n w_j} \quad (1)$$

where the term in the parenthesis is the choice function, and α is the choice parameter. Once the winner is determined (committed), it is subject to a vigilance test according to:

$$\frac{\sum_{i=1}^n \min(x_j, w_j)}{\sum_{i=1}^n x_j} \geq \rho \quad (2)$$

where ρ is the vigilance parameter, a predefined positive number. If the test in the above equation fails, this winner is considered invalid and ruled out (reset) and the search is repeated until a new satisfactory winner is found (resonance). If no winner is selected, a new neuron will be generated to accommodate this input pattern. Carpenter (1989) gave a vivid analogy that the category choice can be conceived as a hypothesis, whilst vigilance test be a statistical significant test.

When a winner is selected (a resonance is found), weights connect it with the F1 neurons are updated as

$$w_j^{t+1} = \beta(I \wedge w_j^t) + (1 - \beta)w_j^t \quad (3)$$

where I is output vector, β is learning rate between F1 layer and F2 layer.

This algorithm is interesting in that it can accommodate novel patterns without forgetting old ones, given the different values of the learning rate β used. A larger β makes the model keep more new memory and forgetting the old.

The supervised training for ART models is achieved by adding MAP and is obtained through "match tracking". Namely, when the match ratio at mapfield is equal or greater than a predefined vigilance ρ_2 , weight vector w_{j2} between a chosen neuron of F2 layer to mapfield layer will be updated as

$$w_{j2}^{t+1} = \beta_2(O \wedge w_{j2}^t) + (O - \beta_2)w_{j2}^t \quad (4)$$

where O is output vector, β_2 is learning rate between F2 layer and mapfield layer.

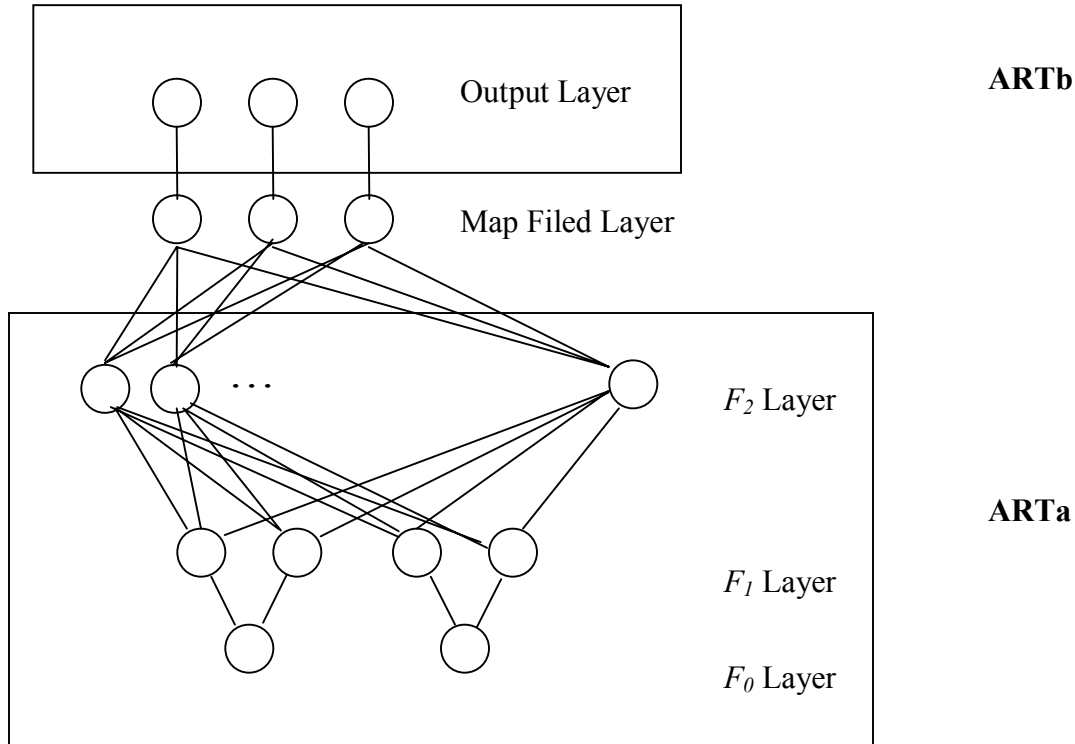


Figure 1. Example of the architecture of a Fuzzy ARTMAP neural network with an F1 layer (made up of $2n$ input for the complement coded input), an F2 layer (dynamically grows during the procedure of learning), a map field layer and output layer.

3 Proposed Soft Classification Algorithms

The fuzzy ARTMAP procedure adopted in this study was closely modeled on that discussed by Mannan et al. (1998). Since both the SOM and fuzzy ARTMAP are trained based upon competitive mechanism and use the “winner-take-all” rule, the development of soft algorithms for the fuzzy ARTMAP can adopt the similar concepts of the *SOM Commitment* and *SOM Typicality* (Li and Eastman 2006). Here they are named *ART Commitment* and *ART Typicality*.

3.1 ART Commitment (ART-C)

Unlike the SOM, every F2 layer neuron has opportunities to be triggered (committed) in that F2 layer neurons dynamically grow in number. Thus there are no redundant neurons in the ART model, which are analogous to the dead units or disconnected units in the SOM (Li and Eastman 2006). Therefore, in ambiguous cases each F2 neuron must represent at least one cluster, and clusters with higher variability will associate with more neurons.

Similar to the SOM, the degree of commitment that the fuzzy ARTMAP can make to an input pattern belonging to a particular class can thus be determined by calculating the committing proportion of the class on the committed neuron, i.e.,

$$C_i = \frac{P_i(j)}{\sum_{i=1}^m P_i(j)} \quad (5)$$

where $P_i(j)$ is the proportion of training site of class i ($i = 1, 2, \dots, m$) committing neuron j ($j = 1, 2, \dots, n$), and $P_i(j)$ can be calculated as:

$$P_i(j) = \frac{f_i(j)}{N_i} \quad (6)$$

where $f_i(j)$ is the frequency of neuron j committed by pixels labeled as class i , and N_i is the total number of samples of class i in the training sites.

With the SOM, in unambiguous cases neurons are either associated with only one class or are left unlabelled, given the training data are good, there are adequate neurons to capture class variability and the bands are such that they can yield completely distinctive signatures. In cases of overlapping class distributions, it would be expected that neurons will be triggered by more than one class leading to the possibility that the neuron could be one of several classes. Thus frequently encountered mixtures would be expected to have organized their own neurons which are unlabelled and thus yield a commitment of 0 to any of the classes upon which it was trained. For the ART model, however, in unambiguous cases neurons are associated with at least one of the training classes. In cases of overlapping class distributions, F2 layer neurons are also expected to be committed by more than one class. The expression in Equation 5 thus allows one to determine the degree of commitment that the classifier can make to each of these possible class labels.

It is not difficult to see that this algorithm yields a measure that is analogous to the concept of a posterior probability under the assumption of equal priors. Posterior probabilities express the probability of a pixel belonging to a class given the conditional probability of encountering such a vector if the pixel were to belong to that class. The $P_i(j)$ described by Equation 6 can be understood as an empirical estimation of the conditional probability of each neuron's weight structure. Further, $P_i(j)$ divided by its summation over each class is a procedure of normalization which is similar to the use of Bayes' Theorem to derive the posterior probability from the conditional and prior probabilities associated with a class. Thus *ART Commitment* is consistent with the concept of posterior probabilities using equal prior probabilities. This measure is not called a probability, however, because of its indirect use in image classification. The measure is really a statement about the neuron's weight structure and not the pixel's reflectance vector per se. Soft classification is achieved by finding the neuron whose weight structure most closely matches that of each pixel and then assigning it the value of that neuron's commitment to each class. Given this indirect link, it is called a measure of *commitment* rather than probability. In a later section of this paper, this reasoning will be justified through a comparison with a non-neural parametric soft classifier that outputs true posterior probabilities. As a consequence, *ART-C* is not considered as a measure of classification ambiguity and not sub-pixel class fractions.

3.2 ART Typicality (ART-T)

The *ART-C* algorithm assigns class commitments based on the relative frequency of class associations with a specific neuron. Thus similar to the calculation of posterior probability in the Maximum Likelihood procedure, it may assign a high commitment to a class even when the frequency with which training pixels fired that neuron was quite small. This is caused by the normalization (Eq. 1) step that effectively treats the lack of evidence that it belongs to other classes as support for the classes that it is associated with. To derive a measure of support that is free of this assumption, an alternative algorithm can be used, i.e.,

$$T_i = \frac{f_i(j)}{\max_j \{f_i(j)\}} \quad (7)$$

where $f_i(j)$ is the frequency of neuron j committed by pixels labeled as class i .

Rather than dividing by the sum of frequencies from classes as shown in Equation 5, the maximum triggering (committing) frequency within the underlying class of interest is used. This algorithm is interesting in that it is analogous to the concept of typicality, which can be derived from the Mahalanobis distance measurements. The Mahalanobis distance between a pixel and the centroid of a multivariate normally distributed class is known to vary with a χ^2 distribution and can thus be used to derive a measure of how typical the pixel is of that class (Foody et al. 1992). These typicality probabilities indicate whether it is reasonable to assume that a case actually belongs to a class – a statement that is derived only from the information available about that class. The ART-T algorithm (Eq. 7) is similar in that it requires no information about other classes to derive its measure and it similarly conveys a sense of how typical a pixel is based on the distribution of training site pixels. A T_i value of 1.0 thus indicates that the pixel in question has reflectances similar to those that were most commonly encountered in the training data for that class. Thus unlike Equation 5, this algorithm considers the variability within the class of interest itself, i.e., intra-class relationship.

4 A Case Study Evaluation

4.1 Data description

To evaluate the proposed two algorithms, we undertook soft classifications of a SPOT HRV image (with 3-bands: G, R, and IR) from 1991. A sub-image of 565×452 pixels covering the region (11.3 km×9.0 km) around Westborough Massachusetts was extracted as the study site. Both training and testing sites for twelve land use/cover classes were digitized and extracted from the imagery. A total of 4597 and 3085 samples were selected for training and validating the fuzzy ARTMAP model respectively (Table 1).

Table 1. Information classes and training/testing sites

Class ID	Land use/cover	Training site pixels	Testing site pixels
1	High density residential area	193	72
2	Low density residential area	220	132
3	Industrial and commercial area	189	158
4	Roads/Transportation	92	76
5	Deep water	808	379
6	Cropland	152	166
7	Deciduous forest	1623	905
8	Wetland	420	488
9	Grass	178	126
10	Conifer forest	85	95
11	Shallow water	584	469
12	Reeds	53	19
	Total	4597	3085

4.2 Methods

Soft classifications using the two proposed algorithms were undertaken after training the Fuzzy ARTMAP neural network. In order to examine the effect of the learning rate and the vigilance parameter on its soft output images, i.e., both the Commitment and Typicality images, a series of tests were conducted using learning rates ranging from 0.7 to 1.0 and vigilance parameters ranging from 0.97 to 0.999. All parameters used in this experiment are listed in Table 2. To illustrate the similarity between the SOM and fuzzy ARTMAP, *SOM Commitment* and *SOM Typicality* measures were also used. Additionally, Bayesian posterior probability soft classifier (Bayes) and a Mahalanobis typicality soft classifier (Mahal) were also used¹ to demonstrate the relationship between these different approaches. Both Correlation Analysis and Principal Components Analysis (PCA) were used to explore the relationship between the measures. The software used for this research was a Fuzzy ARTMAP neural network developed by the author as part of a suite of neural net classifiers for the IDRISI GIS and Image Processing software system (Eastman 2006).

A set of soft images from a Fuzzy ARTMAP model (learning rate = 0.9 and vigilance parameter = 0.999) were selected for analysis because this data set yielded a highest average correlation coefficient (0.6576) with the Mahalanobis and a relatively high overall Kappa (0.8642) in terms of hard classification (Table 3). With this Kappa, the Fuzzy ARTMAP model outperformed the hardening results from Bayesian and Mahalanobis classifier, which were 0.8571 and 0.8073 respectively. Another reason to use this set of results is that, among the twelve classes, those with less intra spectral variabilities such as the shallow water (class 11) and deep water class (class 5), and with higher variabilities such as deciduous forest (class 7) created

¹ The Bayesian posterior probability and Mahalanobis typicality soft classifiers used were the BAYCLASS and MAHALCLASS procedures in IDRISI Kilimanjaro, respectively.

by ART-T show extremely high similarity to those by Mahalanobis, with correlation coefficients of 1.0, 0.98 and 0.94, respectively.

Soft classification maps for each of the twelve classes (Table 1) from all the six classifiers were created. Due to limited space, only maps of the deciduous forest class have been chosen as shown in Figure 2. Note the similarity in the outputs of the Bayes, SOM-C and ART-C classifiers and between Mahalanobis, SOM-T and ART-T. Similarly note the differences between these two groups. Visually it would appear that the SOM-C and ART-C algorithm do produce a measure similar to a posterior probability while SOM-T and ART-T produce a form of typicality.

Table 2. Parameters used in the experiment

Parameter	Value
F1 layer neuron number	6
F2 layer neuron number	Dynamically growing
Output layer neuron number	12
Mapfield layer neuron number	12
Choice parameter	0.01
ARTa Learning rate	0.7 – 1.0
ARTa vigilance parameter	0.97 – 0.999
ARTa Learning rate	1.0
ARTa vigilance parameter	1.0

Table 3. Correlation coefficient between ART-T and Mahalanobis model

C 1	C 2	C 3	C 4	C 5	C 6	C 7	C 8	C 9	C 10	C 11	C 12	AVE	Overall Kappa	C: Classes
0.50	0.59	0.65	0.61	0.98	0.66	0.94	0.66	0.53	0.31	1.00	0.49	0.66	0.86	

Table 4. Correlation matrix between different models

	Bayes	mahalanobis	SOM-C	SOM-T	ART-C	ART-T
Bayes	1.00	0.64	0.83	0.65	0.72	0.64
mahalanobis	0.64	1.00	0.66	0.89	0.45	0.94
SOM-C	0.83	0.66	1.00	0.68	0.61	0.62
SOM-T	0.65	0.89	0.68	1.00	0.45	0.87
ART-C	0.72	0.45	0.61	0.45	1.00	0.51
ART-T	0.64	0.94	0.62	0.87	0.51	1.00

5 Results and Discussion

Correlation analysis was undertaken between all these different models as an initial exploration of their relationships. Table 4 lists a correlation matrix for the deciduous forest class (Class 7). Results show that higher correlations exist between Bayes, SOM-C and ART-C (0.61-0.83), and between Mahal, SOM-T and ART-T (0.87-0.94), while lower correlation (0.45-0.65) exist between the

probability models (i.e., Bayes, SOM-C and ART-C) and typicality models (i.e., Mahal, SOM-T and ART-T).

For an analytical confirmation of these observations, a PCA was used to produce six components from the results for each class. The results of these 12 analyses were very similar. Thus they have been chosen to illustrate through an examination of the results for the deciduous forest class. Figure 3 shows these component images for the deciduous forest class while Table 5 shows the component loading and Table 6 indicates the variance explained by each component.

The results show that Component 1 explains 75.8% of the variance of the six models, and has high loadings on the Bayesian soft classifier (0.94), SOM-C (0.90) and ART-C (0.86), and moderate loadings on Mahal (0.69), SOM-T (0.72) and ART-T (0.72). These results are interesting in that Component 1 shows that probability models are the key elements in common. Bayes, SOM-C and ART-C are similarly associated with this component while ART-T, SOM-T and Mahal show lower but almost equal correlations with this component. It is logical that the typicality measures would be moderately correlated with posterior probability since the pixels which are most typical of a class would also be highly probable to belong to that class.

Independent of Component 1, the second component explains 13.4% of the variance. This component is positively associated with ART-C, while negatively with Bayes, Mahal, SOM-C, SOM-T and ART-T. This interesting result indicates that ART-C model is different from all the others. Component 3 also accounts for a significant variance (5.89%) and is similarly associated with Mahal, SOM-T and ART-T with loadings of 0.54, 0.49 and 0.58 respectively. In contrast, Bayes, SOM-C and ART-C show loadings of -0.12 and -0.20 and 0.06 respectively. Thus it would appear that Component 3 expresses what the typicality measures provide which is independent from the probability measures. The fact that the component looks similar to the typicality measures and that their loadings are very similar gives credence to our position that the SOM-T and ART-T measure do express typicality.

Table 5. PCA loadings for the deciduous forest class

Loading	Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6
Bayes	0.94	-0.10	-0.12	-0.30	0.00	0.00
mahalanobis	0.71	-0.39	0.54	0.01	-0.17	-0.13
SOM-C	0.90	-0.32	-0.20	0.23	-0.01	0.00
SOM-T	0.72	-0.39	0.50	0.03	0.28	0.00
ART-C	0.86	0.51	0.06	0.07	0.00	0.00
ART-T	0.72	-0.27	0.58	-0.01	-0.17	0.19

Table 6. Variance explained by the components for the deciduous forest class

Component	Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6
% var.	75.81	13.43	5.89	3.90	0.75	0.23

Component 4 shows a positive association with SOM-C (0.23) and a negative association with Bayes (-0.30). Component 5 and 6 account for small variance of input variables. Component 5 again shows a positive loading with SOM-T (0.28), but negative and equal loadings with Mahalanobis (-0.17) and ART-T (-0.17), thus indicates again ART-T measure is

similar to Mahalanobis' typicality. Component 6 accounts for only 0.23% of the total variance and has significant and opposite loadings on ART-T (0.19) and Mahalanobis (-0.13), which indicates Component 6 highlights a difference between these two algorithms. It is believed that Component 6 may have picked up a difference between the assumption of normality for Mahalanobis and the non-parametric character of ART-T. However a definitive explanation of character of Components 5 and 6 is beyond the scope of this study. They are not large components, but were consistently found in all 12 covers. Thus they are expected to become topics of future work.

6. Conclusions

In this paper soft supervised classification algorithms i.e., *ART-Commitment* and *ART-Typicality*, for fuzzy ARTMAP neural network were developed and the relationship between the conventional parametric measures and the neurocomputational non-parametric measures were examined and verified. These algorithms are explicit, meaningful and easily implemented. There are three important implications of this development. First, great similarities exist among the SOM-C, ART-C and a parametric Bayesian posterior probability classifier, and among the SOM-T, ART-T and a Mahalanobis typicality classifier. Second, they are non-parametric and free from assumptions about the form and distribution of input data, and outperformed the conventional methods. Thus they act as appropriate substitutes for posterior probability and typicality soft classifiers in the context of non-normal data. Finally, similar to the SOM-T, ART-T procedure has the additional attraction of requiring only intra-class information. Thus it is suitable for analysis of a single class. This can be of value in remote sensing where evidence is sought of the presence of a specific cover type or to fields such as ecology that need to model species distributions on the basis of presence data alone.

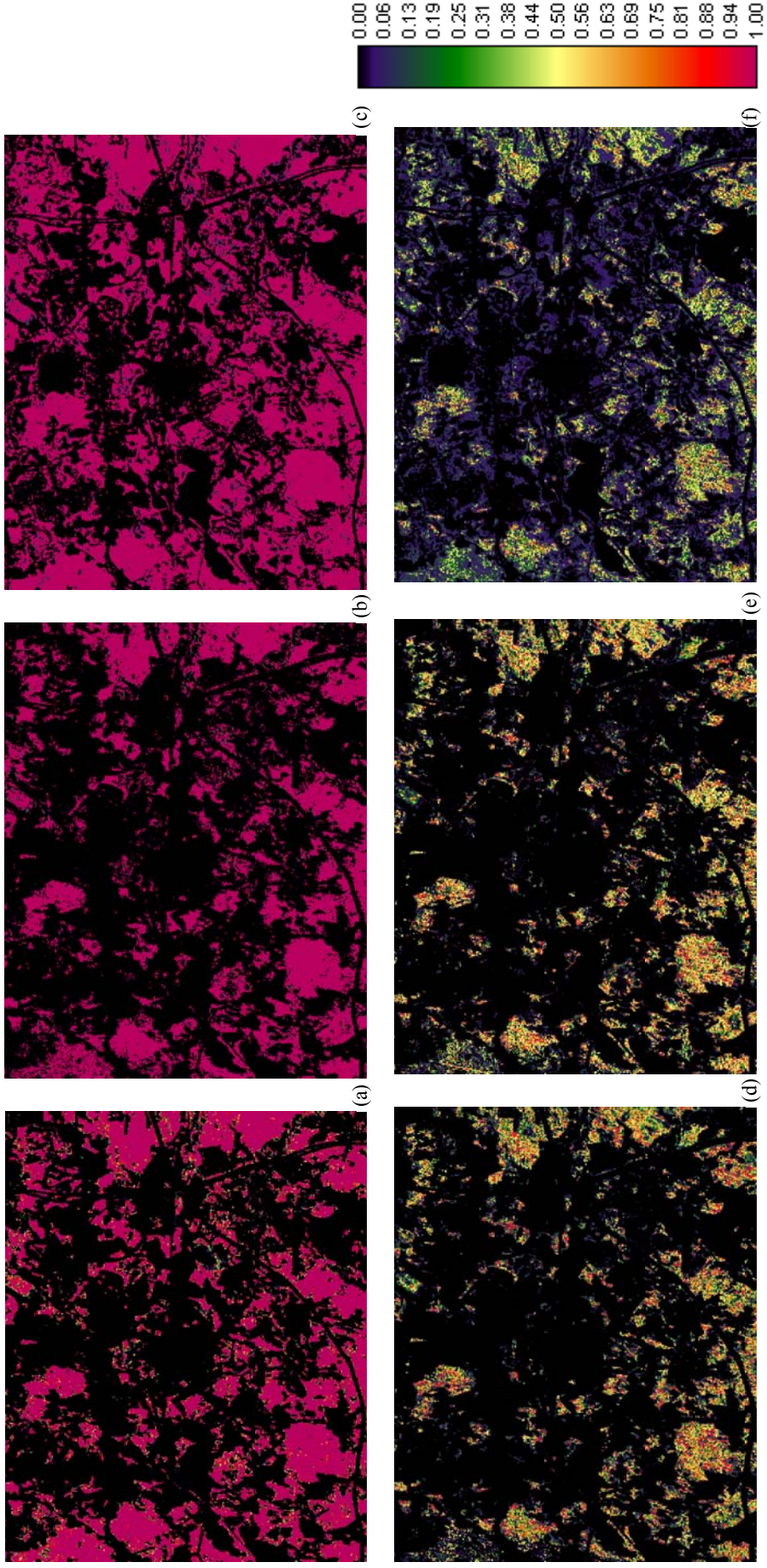


Figure. 2. Soft classification maps for the deciduous forest class created from (a) Bayesian soft classifier; (b) SOM Commitment; (c) ART Commitment; (d) Mahalanobis typicality classifier; (e) SOM Typicality; and (f) ART Typicality.

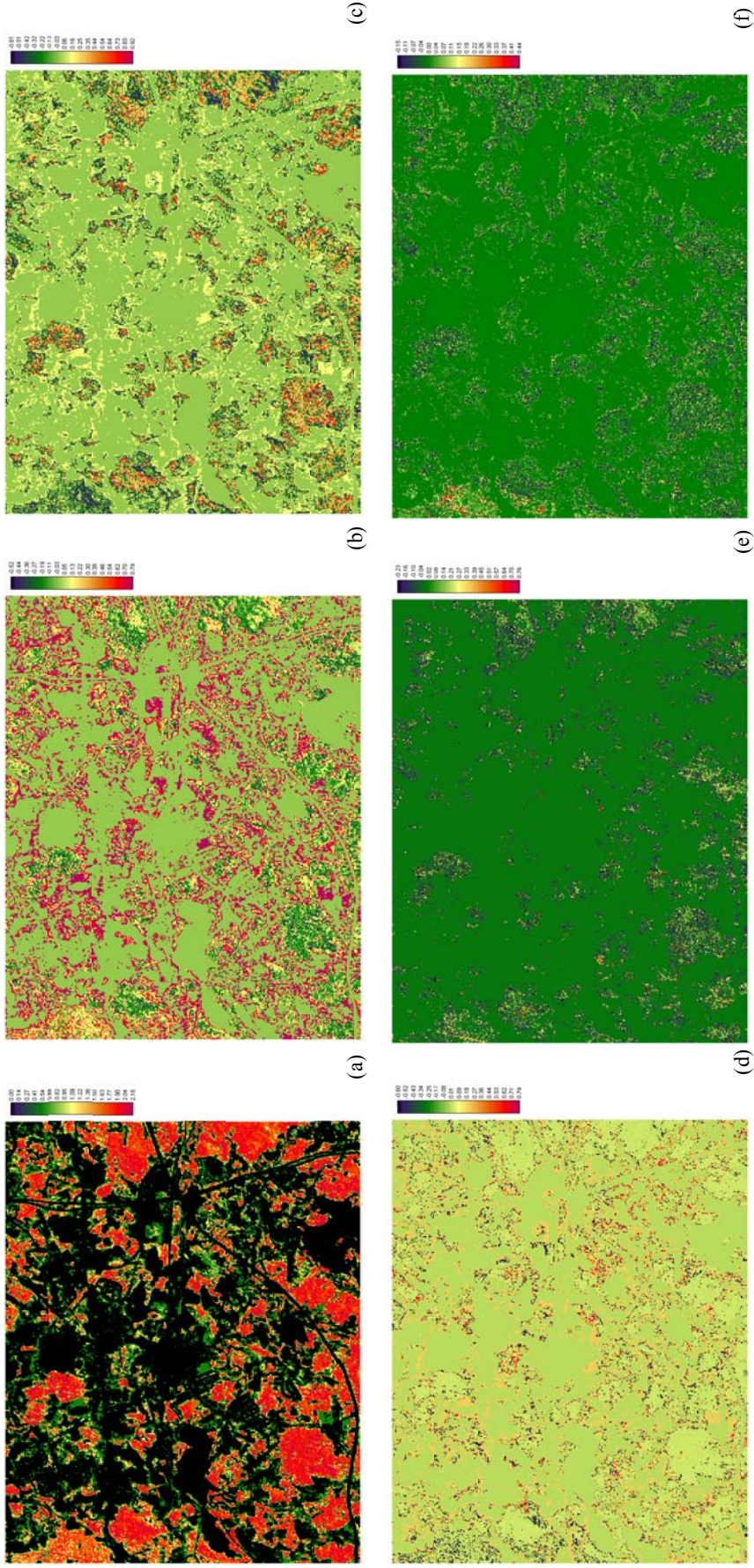


Figure 3. PCA for the deciduous forest class: (a) Component 1; (b) Component 2; (c) Component 3; (d) Component 4; (e) Component 5; and (f) Component 6.

Acknowledgments

The author is grateful to Dr. J. Ronald Eastman for his direction and Dr. David Weiguo Liu of the University of Toledo for his insightful discussion and suggestion. The procedures covered in this paper have been implemented using the IDRISI GIS and Image Processing System.

References

- Atkinson PM, Cutler MEJ, Lewis H (1997) Mapping sub-pixel proportional land cover with AVHRR imagery. *International Journal of Remote Sensing* 18(4): 917-935
- Bernard AC, Wilkinson GG, Kanellopoulos I (1997) Training strategies for neural network soft classification of remotely-sensed imagery. *International Journal of Remote Sensing* 18(8): 1851-1856
- Carpenter G, Gillison AN, Winter J (1993) DOMAIN: a flexible modeling procedure for mapping potential distributions of plants, animals. *Biodiversity Conservation* 2: 667-680
- Carpenter GA (1989) Neural Network Models for Pattern Recognition and Associative Memory. *Neural Networks* 2: 243-257
- Carpenter GA, Crossberg S, Markuzon N, Reynolds JH, Rosen DB (1992) Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supervised Learning of Analog Multidimensional Maps. *IEEE Transactions on Neural Networks* 3(5): 698-713
- Carpenter GA, Crossberg S, Reynolds JH (1991) ARTMAP: Supervised Real-Time Learning and Classification of Nonstationary Data by a Self-Organizing Neural Network. *Neural Network* 4: 565-588
- Carpenter GA, Gajja MN, Gopal S, Woodcock, CE (1997) ART neural networks for remote sensing: Vegetation classification from Landsat TM and terrain data. *IEEE Transactions on Geoscience and Remote Sensing*, 35(2): 308-325
- Carpenter GA, Gopal S, Macomber S, Martens S, Woodcock CE (1999) A neural network method for mixture estimation for vegetation mapping. *Remote Sensing of Environment*, 70: 138-152
- Eastman JR (2003) *IDRISI : The Kilimanjaro Edition*. (Clark Labs, Clark University, Worcester, Mass)
- Eastman JR, Laney RM (2002) Bayesian soft classification for sub-pixel analysis: A critical evaluation. *Photogrammetric Engineering & Remote Sensing* 68(11): 1149-1154
- Eastman JR, Toledano J, Crema S, Zhu H, and Jiang H (2005) In-Process Classification Assessment of Remotely Sensed Imagery. *GeoCarto International* 20: 33-44
- Foody GM (1996) Relating the land-cover composition of mixed pixels to artificial neural network classification output. *Photogrammetric Engineering & Remote Sensing* 62(5): 491-499
- Foody GM (1997) Fully fuzzy supervised classification of land cover from remotely sensed imagery with an artificial neural network. *Neural Computing & Application* 5: 238-247
- Foody GM (1998) Sharpening fuzzy classification output to refine the representation of sub-pixel land cover distribution. *International Journal of Remote Sensing* 19(13): 2593-2599
- Foody GM (1999) The continuum of classification fuzziness in thematic mapping. *Photogrammetric Engineering & Remote Sensing* 65(4): 443-451

- Foody GM (2004) Supervised image classification by MLP and RBF neural networks with and without an exhaustively defined set of classes. *International Journal of Remote Sensing* 25 (15): 3091-3104
- Foody GM, Campbell NA, Trodd NM, Wood TF (1992) Derivation and Applications of Probabilistic Measures of Class Membership from the Maximum Likelihood Classification. *Photogrammetric Engineering and Remote Sensing* 58: 1335-1341
- Foody GM, Campbell NA, Trodd NM, Wood TF (1992) Derivation and Applications of Probabilistic Measures of Class Membership from the Maximum Likelihood Classification. *Photogrammetric Engineering and Remote Sensing*, 58(9): 1335-1341
- Gamba P, Dell'Aqua D (2003) Increased accuracy multiband urban classification using a neuro-fuzzy classifier. *International Journal of Remote Sensing* 24(4): 827-834
- Guisan A, Zimmerman NE (2000) Predictive habitat distribution models in ecology. *Ecological Modeling* 135: 147-186
- Ito Y, Omatu S (1997a) Category classification method using a self-organizing neural network. *International Journal of Remote Sensing* 18: 829-845
- Ito Y, Omatu S (1998) Polarimetric SAR data classification using competitive neural networks. *International Journal of Remote Sensing* 19: 2665-2684
- Ji CY (2000) Land-use classification of remotely sensed data using Kohonen self-organizing feature map neural network. *Photogrammetric Engineering & Remote Sensing* 66: 1451-1460
- Ju J, Kolaczyk E, Gopal S (2003) Gaussian mixture discriminant analysis and sub-pixel land cover characterization in remote sensing. *Remote Sensing of Environment* 84: 550-560
- Kangas JA, Kohonen TK, Laaksonen JT (1990) Variants of Self-Organizing Maps. *IEEE Transactions on Neural Networks* 1: 93-99
- Kohonen T (1989) *Self-Organization and Associative Memory*. 3rd edition, (Springer, Berlin)
- Kohonen T (1990) The Self-Organizing Map. *Proceedings of the IEEE* 78: 1464-1480
- Lehmann A, Overton JM, Leathwick JR (2002) GRASP: generalized regression analysis and spatial prediction. *Ecological Modeling* 160: 165-183
- Li Z, Eastman JR (2006) Commitment and Typicality Measurements for the Self-Organizing Map. *UCGIS 2006 Summer Assembly*, Vancouver, Washington, June 28 – July 1, 2006. http://www.ucgis.org/summer2006/studentpapers/li_zhe.pdf
- Li Z, Eastman JR (2006) The nature of and classification of unlabelled neurons in the use of Kohonen's Self-Organizing Map (SOM) for supervised classification. *Transactions in GIS* 10(4): 599-613
- Liu W, Gopal S, Woodcock CE (2004) Uncertainty and confidence in land cover classification using a hybrid classifier approach. *Photogrammetric Engineering & Remote Sensing* 70(8): 963-971
- Liu W, Seto KC, Wu EY, Gopal S, Woodcock, CE (2004) ART-MMAP: a neural network approach to subpixel classification. *IEEE Transactions on Geoscience and Remote Sensing* 42(9): 1976-1983
- Mannan B, Ray AK (2003) Crisp and fuzzy competitive learning networks for supervised classification of multispectral IRS scenes. *International Journal of Remote Sensing* 24: 3491-3502
- Mannan B, Roy J (1998) Fuzzy ARTMAP supervised classification of multi-spectral remotely-sensed images. *International Journal of Remote Sensing* 19: 767-774

- Mather PM (1999) *Computer Processing of Remotely Sensed Images: An introduction*. Second Edition. (Wiley, New York)
- Mertens KC, Verbeke LPC, Ducheyne EI, De Wulf RR (2003) Using genetic algorithms in sub-pixel mapping. *International Journal of Remote Sensing* 24(21): 4241–4247
- Pal M, Mather PM (2003) An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment* 86 (4): 554-565
- Pal NR, Laha A, Das J (2005) Designing fuzzy rule based classifier using self-organizing feature map for analysis of multispectral satellite images. *International Journal of Remote Sensing* 24: 2219-2240
- Phillips SJ, Anderson RP, Schapire, RE (2005) Maximum entropy modeling of species geographic distributions. *Ecological Modeling* 190: 231-259
- Qiu F, Jensen JR (2004) Opening the Black Box of Neural Networks for Remote Sensing Image Classification. *International Journal of Remote Sensing* 25(9): 1749-1768
- Richards JA, Jia X (1999) *Remote Sensing Digital Image Analysis*. (Springer, New York) 223-238
- Rogan J, Chen D (2004) Remote sensing technology for mapping and monitoring land-cover and land-use change. *Progress in Planning* 61: 301–325
- Rogan J, Franklin J, Stow D, Miller J, Roberts DA, Woodcock C (in press) Mapping land cover modifications over large areas: a comparison of machine learning techniques. *Remote Sensing of Environment*
- Rohwer R, Wynne-Jones M, Wysotzki F (1994) *Neural Network*. In Michie D, Spiegelhalter DJ (eds) *Machine Learning, Neural & Statistical Classification* (Prentice Hall) 84-105
- Rollet R, Benie GB, Li W, Wang S (1998) Image classification algorithm based on the RBF neural network and K-means. *International Journal of Remote Sensing* 19(15): 3003-3009
- Settle JJ, Drake NA (1993) Linear mixing and the estimation of ground proportions. *International Journal of Remote Sensing* 14: 1159-1177
- Simard M, Grand GD, Saatch S, Mayaux P (2002) Mapping tropical coastal vegetation using JERS-1 and ERS-1 radar data with a decision tree classifier. *International Journal of Remote Sensing* 23: 1461-1474
- Stockwell D, Peters D (1999) The GARP modeling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science* 13: 143–158
- Tatem AJ, Lewis HG, Atkinson PM, Nixon MS (2002) Super-resolution land cover pattern prediction using a Hopfield neural network. *Remote Sensing of Environment* 79(1): 1-14
- Tatem AJ, Lewis HG, Atkinson PM, Nixon MS (2003) Increasing the spatial resolution of agricultural land cover maps using a Hopfield neural network. *International Journal of Remote Sensing* 17(7): 647-672
- Tso B, Mather PM (2001) *Classification Methods for Remotely Sensed Data*. (Taylor and Francis, New York)
- Ultsch A, Siemon H (1989) *Technical Report 329*. (University of Dortmund, Dortmund, Germany)
- Villmann T, Merenyi E., Hammer B (2003) Neural maps in remote sensing analysis. *Neural Networks* 16: 389-403
- Waldemark J (1997) An automated procedure for cluster analysis of multivariate satellite data. *International Journal of Neural Systems* 8(1): 3-15

- Wiley EO, McNyset KM, Peterson AT, Robins CR, Stewart AM (2003) Niche modeling and geographic range predictions in the marine environment using a machine-learning algorithm. *Oceanography* 16(3): 120–127
- Yang CC, Prasher SO, Enright P, Madramootoo C, Burgess M, Goel PK, Callum I (2003) Application of decision tree technology for image classification using remote sensing data. *Agricultural System* 76: 1101-1117
- Zaniewski AE, Lehmann A, Overton JM (2002) Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. *Ecological Modeling* 157: 261–280
- Zhang J, Foody GM (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: 615-628