# NON-LINEAR APPROACH IN MULTISPECTRAL DATA CLASSIFICATION

*Hristo Nikolov*

*Solar-Terrestrial Influences Lab.-BAS*

## Introduction

During the past 20 years, statistical classification methods, such as the minimum distance and the maximum likelihood classifiers, have been widely used. However, these methods have their restrictions, related particularly to the distribution assumptions and limitations in the input data types. In the past decade, the non-linear approaches, theoretically a more sophisticated and robust methods of image classification has been introduced and employed in remote sensing applications. Although these methods have been used in a wide range of scientific disciplines for a variety of applications since the early 1980s, their use in remote sensing area is relatively new, dating only from the early 1990s. Studies have shown that non-linear methods are more robust than conventional statistical methods in terms of producing classification results with higher accuracies and requiring fewer training samples. One of the most important of their characteristics is perhaps the non-parametric nature of the model, assuming no a priori knowledge, particularly of the frequency distribution of the data. Because of their adaptability and their ability to produce high-quality results, the use of non-linear methods has increased the in the remote sensing field research. Often comparison is made with model which applies maximum likelihood classification and this will be the approach stated hereafter.

### *Method and Data*

In this study only classification methods based on supervised learning will be considered. The non-linear methods used to achieve the results are:

- ***NN method*** - the most common neural network model is the multilayer perceptron (MLP) - type networks that work in a feed-forward direction where information progresses from an input layer to an output layer in the learning phase. Such networks contain an extra layer or layers termed the hidden layer(s) to overcome the problems of the perceptron. Due to the involvement of one or more extra layers and the use of nonlinear rather than linear transfer functions (in this research sigmoidal function), the MLP can approximate and map any kind of problem. Bostock (1994) emphasizes that the major reason for the popularity of MLP models is that whilst some problems are more efficiently modeled by other more specialized networks, such as radial basis function networks or binary tree structures, the multilayer perceptron trained by backpropagation learning algorithm, is a good general learning tool for a wide range of applications. A typical neural network consists of one input layer, one or two hidden layers and one output layer. Training a feed-forward neural network using the backpropagation algorithm involves setting several initial parameters including network structure, learning rate, momentum term and activation function. According to Hand (1997), 'a network with two hidden layers allows convex regions to be combined, producing no convex, even disconnected regions i.e. two hidden layers are enough for any task while retaining good generalization.

- ***SVM method –*** Support Vector Machines (SVM) have been recently introduced in the statistical learning theory domain for regression and classification problems, and applied to the classification of multispectral images. The technique consists in finding the optimal separation surface between classes thanks to the identification of the most representative training samples of the side of the class called support vectors. If the training data set is not linearly separable, a kernel method (linear, polynomial kernels) is used to simulate a non-linear projection of the data in a higher dimension space, where the classes are linearly separable. Actually, the projection can be simulated using a kernel method. Besides, unless statistical estimations, a small number of training samples is enough to find the support vectors. This classifier proposes a very interesting property for multispectral image processing- it does not suffer from the Hughes phenomenon and it may perform class separation even with means very closed to each other with a small number of training samples (Gualtieri'99). In this investigation a multiple class separation is performed by so called one against one approach - ***M(M-1)/2*** classifiers are applied on each on each pair of classes, the most often computed label is kept for each vector.

- ***Bayes classifier*** (BC)– this method rely on the local distribution functions $p(x_i/pa_i;\ \Theta_i;\ Sh)$ are essentially classification/regression models. Therefore, if we are doing supervised learning where the explanatory (input) variables cause the outcome (target) variable and data is complete, then the Bayesian-network and classification approaches are identical. A Bayesian network is a graphical model for probabilistic relationships among a set of variables. One of its most prominent properties is the fact that Bayesian networks can readily handle incomplete data sets (Chickering'96). Bayesian methods in conjunction with Bayesian networks and other types of models offers an efficient and principled approach for avoiding the over fitting of data. When dealing with incomplete data sets Monte-Carlo methods yield accurate results, but they are often intractable in large sample size. Another approximation that is more efficient than Monte-Carlo methods and often accurate for relatively large samples is the Gaussian approximation, which we used in this study.

The effectiveness of the above methods was tested with data from laboratory experiments (contracts with MES МУНЗ 1201/03 and Б1306/03) and satellite data from ETM instrument of Landsat7. Both data sets were divided in training and validation sets with total number of samples 130 and 2200 respectively.

**Results and Discussion**

The results for training and classification are shown in tables below. Maximum likelihood (ML) statistic method is used for reference for the ones introduced above.

*Training data for classification (number of multispectral pixels- samples)*

| *Type of land cover* | *Laboratory* | *Satellite* |
|---|---|---|
| Vegetation | 45 | 275 |
| Rocks | 35 | 600 |
| Water | 50 | 580 |
| Urban | - | 745 |

*Accuracy over the validation set after training for both data sets (percent)*

| *Type of land cover* | ML | | NN | | SVM | | BC | |
|---|---|---|---|---|---|---|---|---|
| | *Lab* | *Sat* | *Lab* | *Sat* | *Lab* | *Sat* | *Lab* | *Sat* |
| Vegetation | 58.7 | 56.3 | 63.2 | 65.8 | 68.4 | 70.6 | 71.2 | 73.6 |
| Rocks | 60.8 | 58.6 | 70.3 | 75.8 | 82.3 | 79.4 | 84.3 | 82.6 |
| Water | 75.8 | 78.4 | 76.9 | 80.1 | 82.5 | 81.5 | 85.6 | 90.3 |
| Urban | - | 82.3 | - | 86.4 | - | 92.6 | - | 94.2 |

**Conclusions**

As it can be seen the Bayes classifier offers better performance for all the data and especially for incomplete training data since it is closer to the optimal classification. On the other hand the other methods discussed should not be underestimated and considered as a competitive equivalent. The non-linear methods for classification stated here, are seen as an introduction of larger feature space (i.e. 10 or more features) that could be used by conventional statistical methods.

**Acknowledgments:** This paper was supported by the NSFB under contracts MU-NZ-1201/02, B-1306/03 and NZ-1410/04.

**References:**

1. Bateson, A. and Curtiss, B. A method for manual endmember selection and spectral unmixing. Remote Sensing of Environment, 55, 229-243., 1996;
2. Gualtieri J. A. and R. F. Cromp, “Support vector machines for hyperspectral remote sensing classification” in Proceedings of the SPIE, vol. 3584, 1999, pp. 221–232.;
3. Hand D. J., Construction and Assessment of Classification Rules. New York: John Wiley & Sons, 1997.
4. Heckerman David, A Tutorial on Learning with Bayesian Networks, Microsoft Research Advanced Technology Division, 1996;
5. Chickering, D. and Heckerman, D., Efficient approximations for the marginal likelihood of incomplete data given a Bayesian network, Technical Report MSR-TR-96-08, Microsoft Research, Redmond, WA