

Patricia Melin  
Janusz Kacprzyk  
Witold Pedrycz (Eds.)

# Soft Computing for Recognition Based on Biometrics

# Automatic Dust Storm Detection Based on Supervised Classification of Multispectral Data

Pablo Rivas-Perea<sup>1</sup>, Jose G. Rosiles<sup>1</sup>,  
Mario I. Chacon Murguia<sup>2</sup>, and James J. Tilton<sup>3</sup>

<sup>1</sup> The University of Texas El Paso,  
Department of Electrical and Computer Engineering, El Paso TX 79968, USA

<sup>2</sup> Chihuahua Institute of Technology,  
Graduate Studies Department, Chihuahua Chih, Mexico

<sup>3</sup> NASA Goddard Space Flight Center,  
Computational and Information Sciences and Technology Office,  
Greenbelt MD 20771, USA

**Abstract.** This paper address the detection of dust storms based on a probabilistic analysis of multispectral images. We develop a feature set based on the analysis of spectral bands reported in the literature. These studies have focused on the visual identification of the image channels that reflect the presence of dust storms through correlation with meteorological reports. Using this feature set we develop a Maximum Likelihood classifier and a Probabilistic Neural Network (PNN) to automate the dust storm detection process. The data sets are MODIS multispectral bands from NASA Terra satellite. Findings indicate that the PNN provides improved classification performance with reference to the ML classifier. Furthermore, the proposed schemes allow real-time processing of satellite data at 1 km resolutions which is an improvement compared to the 10 km resolution currently provided by other detection methods.

## 1 Introduction

Dust aerosols are a major cause of health, environmental, and economical hazards, and can adversely impact urban areas [1]. From a scientific perspective, understanding dust storm genesis, formation, propagation and composition is important to reduce their impact or predict their effect (*e.g.*, increase of asthma cases). Multispectral instruments allow space imaging of atmospheric and earth materials based on their spectral signature. More specifically, they allow the detection of dust air-borne particles (aerosols) propagated through the atmosphere in the form of dust storms.

Current methods for dust aerosol are based on the Moderate Resolution Spectroradiometer (MODIS) Aerosol Optical Thickness (AOT) product [2, 3] which is provided by the NASA Terra satellite.

The AOT product allows tracking of pollutant aerosols by observing the aerosol optical thickness. However, AOT products require a considerable amount of processing time (*i.e.*, two days after satellite pass) before useful information on aerosol events is extracted. The use of simple band arithmetic (e.g., subtraction) has been reported as a scheme to visualize the presence of dust storms [1]. This method is highly subjective making interpretation dependent on the experience of the analyst.

Given the large amount of data produced by MODIS, it is also desirable to have automated systems that assist scientist on finding or classifying different earth phenomena with minimal human intervention. For example, Aksoy, *et al.* [4], developed a visual grammar scheme that integrates low-level features to provide a high level spatial scene description on land cover and land usage. Similar automated schemes for dust detection are highly desirable.

In this paper we present two methods for the detection of dust storms from multispectral imagery using statistical pattern classifiers. Based on reported data, we present a feature set that allows accurate and real-time detection of dust aerosol. The proposed feature set is extracted from MODIS spectral bands and evaluated with a maximum likelihood (ML) classifier and a probabilistic neural network (PNN). We will show that the PNN approach provides a better detection and representation of dust storm events.

This paper is organized as follows. Section 2 of the paper introduces the dust aerosol multispectral analysis. The ML and PNN models are explained in Section 3 and 4. Section 5 presents experimental results leading to different levels of segmentation between dust storms and other materials. Finally, conclusions are drawn in Section 6.

## 2 An Overview of MODIS Data

Remote sensing is the research area that studies how to gather and analyze information about the Earth from a distance. Uses include the mapping of fires, weather monitoring, cloud evolution, and land cover analysis. The information gathered can be used to produce images of erupting volcanoes, monitor for dust storms, view the sequential growth of a city, and track deforestation over time [6, 18].

In this paper we collected thermal information about the land, stratosphere, and atmosphere using special instruments aboard a satellite orbiting the Earth surface. This instrument is called “Moderate-Resolution Imaging Spectroradiometer” (MODIS). These remotely sensed data is collected as digital files, containing data captured at different spectral waves in the optical range (*i.e.* multispectral data). These digital files are known as “granules” and can be downloaded from the web at the NASA WIST tool.

The MODIS instrument is built in NASA Terra and Aqua satellites. MODIS multispectral data is currently used in the analysis of different

phenomena like sea temperature and surface reflectivity. MODIS provides information in 36 spectral bands between wavelengths 405nm and 14.385 $\mu$ m.

MODIS multispectral data is available in different levels. These levels depend on the level of data processing. Level 0 is raw telemetry data (*i.e.* satellite unorganized data). Level 1A is raw data organized by spectral bands. Level 1B consists of corrected multispectral data (*i.e.* bad sensor information is pointed out). Subsequent levels are processed for particular analysis that include aerosol, water vapor, and cloud. In this paper we use the multispectral bands available in MODIS Level 1B.

### 3 Selection and Analysis of Spectral Bands for Feature Extraction

In this section we described the proposed feature extraction process based on the analysis of spectral bands reported in the literature. These studies have focused on the visual identification of the image channels that reflect the presence of dust storms through correlation with meteorological reports. Visual assessment of dust storms can be achieved using MODIS bands  $B1$ ,  $B3$ , and  $B4$  which are within human visual range [5]. An RGB-like composite image can be produced by the mapping red to  $B1$ , green to  $B4$ , and blue to  $B3$ . Hao *et al.*[6] demonstrated that bands  $B20, B29, B31$  and  $B32$  can also be utilized for dust aerosol visualization. Ackerman *et al.*[7] demonstrated that band subtraction  $B32 - B31$  improves dust storm visualization contrast. Based on these findings, we will form feature vectors using pixels values from the recovered bands  $B20, B29, B31$ , and  $B32$ .

A "recovered" radiance is a 16 bit MODIS band mapped to its original units ( $W/m^2/\mu m/sr$ ). The recovery process is given by

$$L = \kappa(\iota - \eta), \quad (1)$$

where  $L$  denotes the recovered radiance,  $\kappa$  is the radiance scale,  $\eta$  denotes the radiance offset, and  $\iota$  is the scaled intensity (raw data). For each pixel location  $(n, m)$ , a feature vector  $F \in \mathbb{R}^4$  is formed by

$$F_{nm} = [L_{nm}^{B20}, L_{nm}^{B29}, L_{nm}^{B31}, L_{nm}^{B32}]^T. \quad (2)$$

corresponding to the recovered radiances of the dust sensitive wavelengths.

### 4 Dust Storm Detection Using the Maximum Likelihood Classifier

The Maximum Likelihood Classifier (ML) has been extensively studied in remotely sensed data classification and analysis [9, 4]. Here we present a straightforward adaptation of the ML classifier to dust storm detection

using the feature set described in the previous section. Let  $f_{X|k}(x) = (X = x|C = k)$  be the conditional probability density function of feature vector  $X$  having a value  $x$ , given the probability that the  $k$ -th class occurs. This might be referred as the “data likelihood” function. Assuming normally distributed features (*i.e.*, pixel values), we can define a discriminant function

$$\psi_k(x) = -\det(\Sigma_k) - (x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k) \quad (3)$$

for each class  $k$ , where  $\Sigma_k$  the covariance matrix,  $\mu_k$  denotes the mean feature vector, and  $\det(\cdot)$  is the determinant function. Then, the decision rule can be simply stated as

$$x \in C = j \quad \text{if} \quad \psi_j(x) > \psi_i(x) \quad \forall j \neq i. \quad (4)$$

The parameters  $\Sigma_k$  and  $\mu_k$  were obtained from the training data described in the previous section using the maximum likelihood estimators (*e.g.*, sample mean and sample covariance matrix).

## 5 Neuro-Probabilistic Modeling: The Probabilistic Neural Network

Specht’s Probabilistic Neural Network (PNN) is a semi-supervised neural network [10]. It is widely used in pattern recognition applications [11]. The PNN is inspired in Bayesian classification and does not require training. It estimate the PDF of each feature assuming they are normally distributed. The PNN has a four-layered architecture as shown in Figure 1. The first layer is an input layer receiving the feature vectors  $F_{nm}$ . The second layer consists of a set of neurons which are fully connected to the input nodes. The output of this layer is given by

$$\varphi_{jk}(F) = \frac{1}{(2\pi)^{\frac{d}{2}} \sigma^d} e^{-\frac{1}{2\sigma^2} (F - \nu_{jk}^F)^T (F - \nu_{jk}^F)}. \quad (5)$$

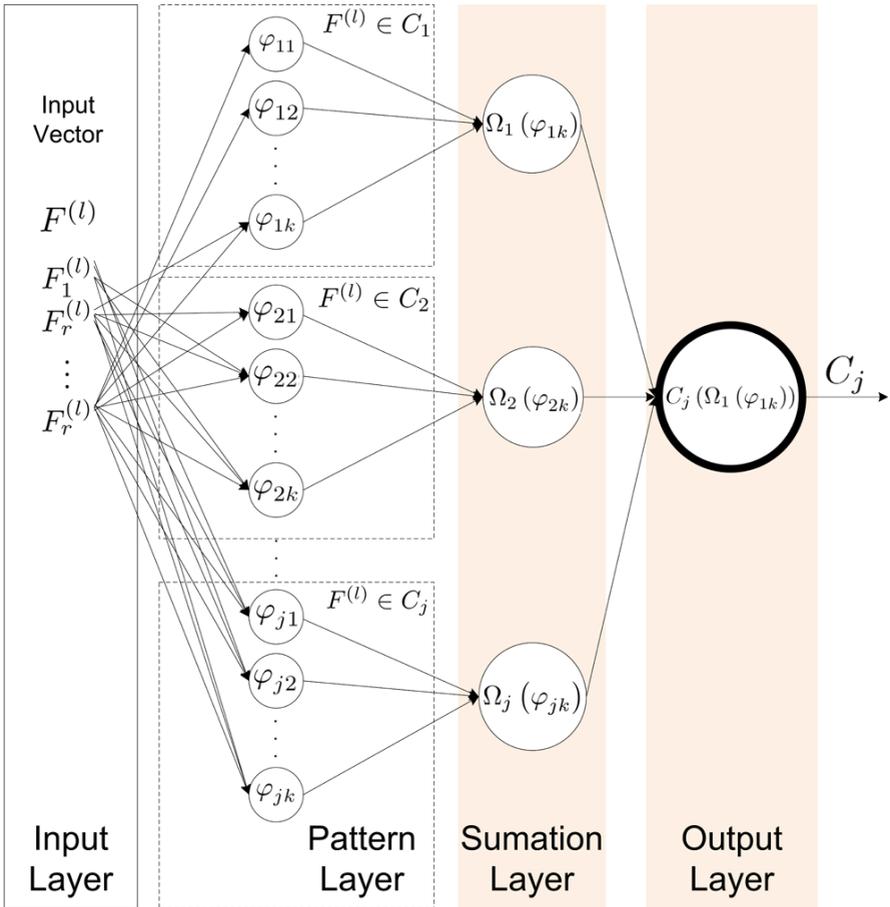
where  $j$  is an index labeling each design vector and  $k$  is its the corresponding class. The pattern units  $\nu_{jk}^F$  correspond to the mean feature vector for each class. The parameter  $\sigma$  is estimated with the method developed by Srinivasan *et al.* [12].

The third layer contains summation units to complete the probability estimation. There are as many summation units as classes. The  $j$ -th summation unit denoted as  $\Omega_j(\cdot)$ , receives input only from those pattern units belonging to the  $j$ -th class. This layer computes the likelihood of  $F$  being classified as  $C$ , averaging and summarizing the output of neurons belonging to the same class. This can be expressed as

$$\Omega_j(\varphi_{jk}(F)) = \frac{1}{(2\pi)^{\frac{d}{2}} \sigma^d N_j} \times \dots \sum_{i=1}^{N_j} e^{-\frac{1}{2\sigma^2}(\varphi_{ik}(F) - \varpi_i)^T(\varphi_{ik}(F) - \varpi_i)}. \quad (6)$$

The last layer classifies feature input vector  $F_{nm}$  according to the Bayesian decision rule given by

$$F \in C_j \text{ if, } \dots C_j(\Omega_j(\varphi_{jk}(F))) = \arg \max_{1 \leq i \leq j} \Omega_i(\varphi_{ik}(F)). \quad (7)$$



**Fig. 1.** The hybrid architecture of the Probabilistic Neural Network. Note the probabilistic nature embedded in a neural architecture.

### 5.1 The PNN Large Sample Size Problem

To avoid the overwhelming processing of millions training samples, we limited the training samples number. We based our reduction method on Kanellopoulos criteria [13] which establishes that the number of training samples must be at least three times the number of feature bands. Therefore, in our PNN design we used six times the feature vector size (*e.g.*, four) requiring 24 training samples per class. In order to select the testing vectors (24 per class), principal component analysis (PCA) was applied to a training set consisting of millions of feature vectors. Then the test feature vectors associated to the 24 largest eigenvalues were selected as the PNN training set.

## 6 Results and Discussion

In our experiments we selected 31 different events corresponding to the south-western US, and north-western Mexico area. The 31 events are known dust storm cases reported in [8]. From these events, 23 were selected to train and test the classifiers. Each event contains multispectral images of size  $2030 \times 1053$  pixels. We manually segmented the images using the MODIS visual range into four classes  $C = \{dust\ storm, blowing\ dust, smoke, background\}$ . The selection of modeling (training) and testing feature vectors was performed by PCA as explained in the last section. The complete data set provides approximately 75 million feature vectors from which 97.5% correspond to the background class. The feature vectors are sliced into 0.005% for training and the remaining are for testing.

In order to evaluate the performance of the classifiers, we need to select a figure of merit. Typically accuracy, received operating characteristic (ROC) or area under the ROC curve (AUC) have been used individually. However, as reported in [16] these measures can only be used interchangeably when the positive and negative test sets are large and balanced. Hence, it is now recognized that using more than one figure of merit is necessary to have a good assessment of a classifier. We evaluate our results using accuracy defined as

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN}, \quad (8)$$

where  $TP$  is the number of true positives,  $FP$  is the number of false positives,  $TN$  is the number of true negatives and  $FN$  is the number of false negatives. Hence accuracy corresponds to the correct classification rate over all classes. A related measure is precision or positive predictive value (PPV) given by

$$\text{Precision} = \frac{TP}{TP + FP}. \quad (9)$$

In this case, precision represents the fraction of true positives from all the vectors classified as a positive. Finally we use AUC which is has been recognized in the machine learning community to provide a better metric than

**Table 1.** Classifiers Performance.

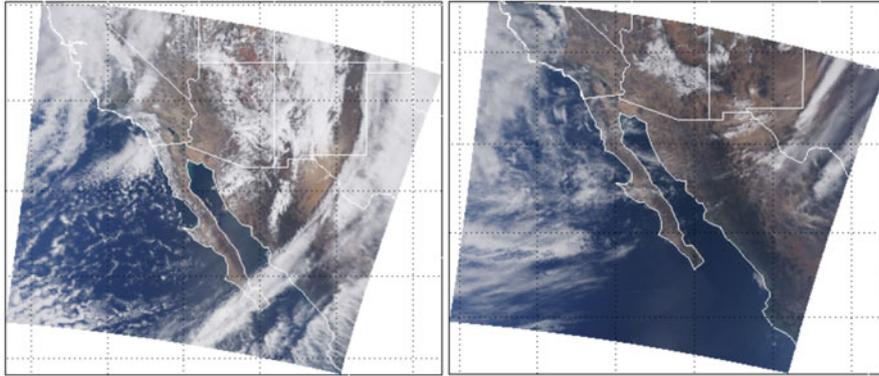
	Precision	Std. Dev.	Accuracy	Std. Dev.	AUC	Std. Dev.	P. Time	Std. Dev.
ML	0.5255	0.2610	0.6779	0.1282	0.4884	0.0036	0.1484	0.0021
PNN	0.7664	0.1616	0.8412	0.1612	0.6293	0.0654	2.5198	0.0018

accuracy [14]. In summary, higher precision and accuracy reflect that a system produces more true positives results while reducing the number of false negatives. Similarly, a higher AUC reflects how a classifier is able to correctly classify feature vectors and at the same time minimize the misclassification errors.

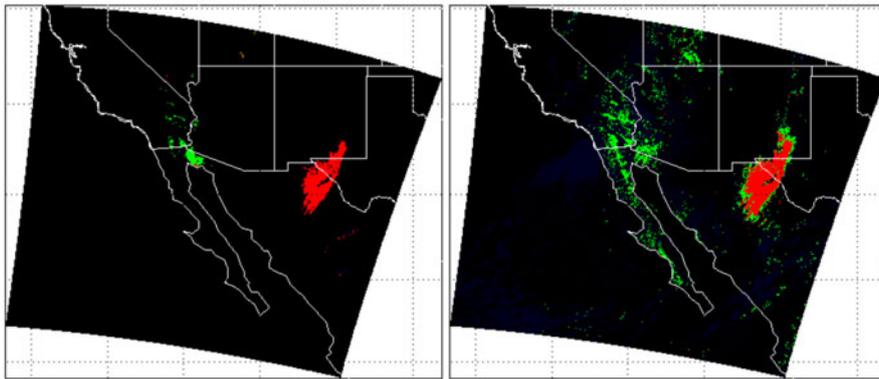
Since in our case we have four classes, generalizing precision and accuracy is obtained by considering a  $4 \times 4$  confusion matrix where the main diagonal entries represent the true positives for each class. We can drop the idea of a negative set and use  $TP_i$  to identify the true positives for class class  $i$ . The idea of false negatives is now represented by the off-diagonal elements of the confusion matrix. For instance, the false negatives for class *dust storm* consists of those vectors misclassified as *blowing dust*, *smoke* or *background*. Similarly, false positives consists of all those vectors classified as *dust storm* that belong to any of the other three classes. Based on these considerations, expressions for precision and accuracy are straightforward to derive. The case of multi-class ROCs and therefore AUCs is an open problem. Some multi-class AUCs are described in [17]. In this paper we resorted to a simpler method where we create a binary classifier by grouping both types of dust as a single (i.e., positive) class, and lumping *smoke* and *background* as the negative.

We present metric results on Table 1. These results were obtained from the whole set of 26 events by averaging each event results. Overall the PNN approach provides better classification than ML. In particular, the AUC indicates that the ML classifier should not be used in the dust storm detection problem. On the other hand, the other metrics show a modest level of performance. Hence, using multiple metrics provides a better understanding on the capabilities of each classifier.

Beyond classifier performance, it is important to integrate the results of the classification with actual images. Ultimately, the output of the classifiers should be used as a tool to help scientists develop insights about dust storms. We present two typical dust storm events in Figure 2. These color images were obtained by mapping three MODIS bands to red, blue and green respectively. The classification results can be visualized as the segmentations shown in Figure 3 for the ML and PNN classifiers. Pixels classified as *dust storm* are labeled red, *blowing dust* to green, *smoke* to blue, and *background* to black. Both classifiers detect the presence of the storms, albeit the PNN detects larger regions. This can be directly explained by the higher PNN metric values on Table 1. From a detection perspective, both classifiers are successful. The ML classifier would be attractive as a detector given its lower computational



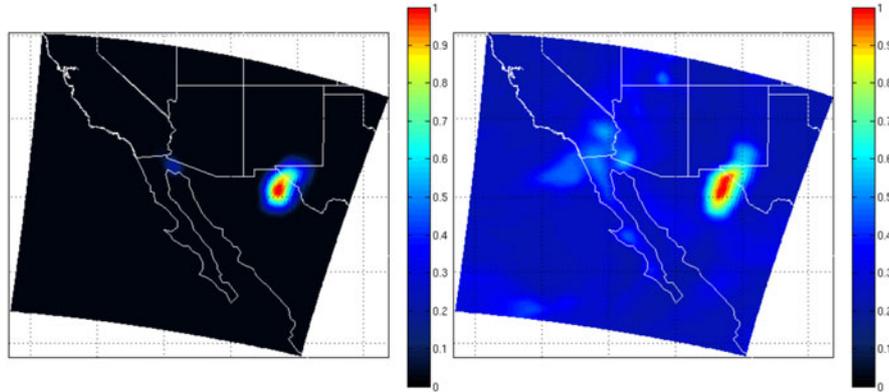
**Fig. 2.** Left, dust storm event on April 6th 2001. True color image R=B1, G=B4, and B=B3. Right, dust storm event on December 15, 2003. True color image R=B1, G=B4, and B=B3.



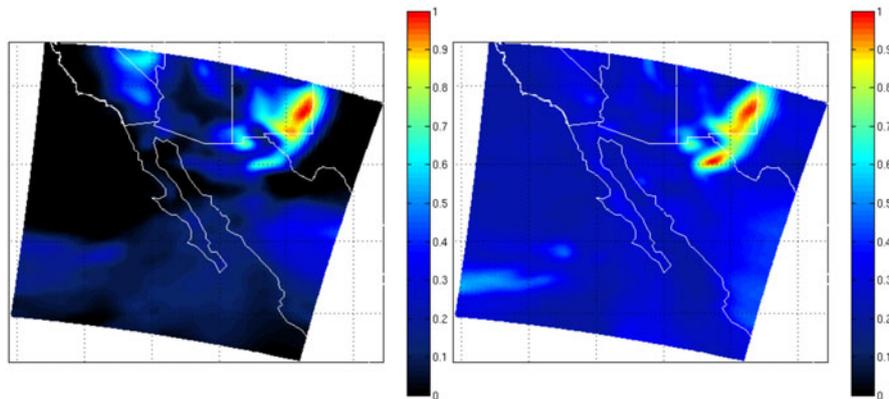
**Fig. 3.** Dust storm event on April 6th 2001. Left, segmentation using ML. Right segmentation using PNN.

requirements. However, if a better understanding on the spatial distribution of the storm is needed, then the PNN should be the selected classifier.

Processing time is an important measure when modeling real-time processing systems. In the case of the MODIS instrument, image swaths of  $10 \times 1053 \times 36$  pixels known as “scans” are produced every 2.96 seconds (*i.e.* 20.3 scans per minute). Thus, a real-time system must perform a classification in less than or equal to this time. The fourth column on Table 1 shows the processing time per scan in seconds. The time shown is computed by taking the time average over all scans for all the events. The times were measured with a MATLAB implementation running on a 2 GHz PC. The time was measured using the *tic()*, *toc()* functions that give the true CPU



**Fig. 4.** Dust storm event on April 6th 2001. Left, dust likelihood probability ML. Right, dust likelihood probability PNN.



**Fig. 5.** Dust storm event on December 15, 2003. Left, dust likelihood probability ML. Right, dust likelihood probability PNN.

processing time. The ML approach takes less than one second to classify the complete scan, and the PNN approach takes about 2.5 seconds to produce the classification result. In conclusion, both can be considered suitable for real time detections at 1km resolution. In contrast, the MODIS AOT product takes two days to be produced and released at a 10km resolution [15].

Finally, as a byproduct of the classification stage, it becomes possible to extract more information about a dust storm by visualizing the dust likelihood over the whole image. Both classifiers produce a parametrization of the likelihood probability density function of dust  $f(\mathbf{x}|\text{dust storm})$  under a multivariate Gaussian assumption. Namely, a new image is formed by assigning a value of  $f(\mathbf{x}|\text{dust storm})$  for each feature vector  $F_{nm}$ . This visualization

over an image provides unique information about the spatial distribution of dust at the moment the image was acquired. This can be utilized to track dust aerosols with a particular degree of confidence. The degree of confidence is proportional to the probability of a pixel being classified as dust storm. With this kind of visualization we can show only those pixels classified as dust storm with a high degree of confidence (*e.g.* above 90% of confidence), that resemble a conservative detection with a high degree of exigence. On the other hand, we can use a low confidence interval (*e.g.* above 5% of confidence) to study how the dust storm spreads across land. This analysis is known as “dust transport,” and is relevant on establishing the origin and extensions of a dust storm. Since the dust aerosol concentration is reduced as the storm advances, dust transport can be studied by analyzing the pixels classified as dust storm but with lower probability.

The dust likelihood visualization of the April 6th, 2001 event is shown in Figure 4. One particularly interesting case is shown in Figure 5, where the visual composite of the satellite image (Figure 2) shows one dust cloud; however, when we observe the dust likelihood visualization we can notice that there were two different dust storm outbreaks at different sources. This information is difficult to see using only the visual composite of MODIS, neither is possible using the AOT product because of the lack of spatial resolution.

## 7 Conclusion

The dust aerosol detection problem has been addressed in this paper. We have modeled probabilistic approaches for dust storm detection and classification. These models are specialized on measuring the dust aerosol probability given MODIS Level 1B data. Machine learning techniques were utilized to model a dust aerosol detection neural architecture. To the best of the authors knowledge, the presented work is first in its kind. We compared the Maximum Likelihood classification (ML) model, and the Probabilistic Neural Network (PNN). The PNN showed a strong ability classifying dust, and discriminating other classes, such as clouds, smoke, and background. Moreover, the proposed probabilistic models are suitable for near real-time applications, such as direct broadcast, rapid response analysis, emergency alerts, etc. The reported work has relevancy in dust aerosol analysis, since the algorithms can show the dust presence to a resolution of 1km. This represents an improvement over Aerosol Optical Thickness index (AOT) methods which lack resolution and have a two day generation delay.

## Acknowledgments

The author P.R.P performed the work while at NASA Goddard Space Flight Center under the Graduate Student Summer Program (GSSP). This work

was partially supported by the National Council for Science and Technology (CONACyT), Mexico, under grant 193324/303732, as well as by the University of Texas El Paso Graduate School Cotton Memorial Funding. The author M.I.C.M. wants to thank DGEST for the support in the dust storm project.

## References

1. Rivera Rivera, N.I., Gebhart, K.A., Gill, T.E., Hand, J.L., Novlan, D.J., Fitzgerald, R.M.: Analysis of air transport patterns bringing dust storms to El Paso, Texas. In: Symposium on Urban High Impact Weather, AMS Annual Meeting, Phoenix, January 2009, vol. JP2.6 (2009)
2. Khazenie, N., Lee, T.F.: Identification Of Aerosol Features Such As Smoke And Dust, In NOAA-AVHRR Data Using Spatial Textures. Geosc. and Rem. Sens. Symp. (May 1992)
3. Levy, R.C., Remer, L.A., Kaufman, Y.J., Tanr, D., Mattoo, S., Vermote, E., Dubovik, O.: Revised Algorithm Theoretical Basis Document: MODIS Aerosol Products MOD/MYD04 (2006)
4. Aksoy, S., Koperski, K., Tusk, C., Marchisio, G., Tilton, J.C.: Learning bayesian classifiers for scene classification with a visual grammar. *IEEE Trans. on Geosc. and Rem. Sens.* 43(3), 581–589 (2005)
5. Gumley, L., Descloitres, J., Schmaltz, J.: Creating Reprojected True Color MODIS Images: A Tutorial. Technical Report. Version: 1.0.2, January 14, (2010)
6. Hao, X., Qu, J.J.: Saharan dust storm detection using MODIS thermal infrared bands. *Journal of Applied Remote Sensing* 1, 013510 (2007)
7. Ackerman, S.A., Strabala, K.I., Menzel, W.P., Frey, R., Moeller, C., Gumley, L., Baum, B.A., Seeman, S.W., Zhang, H.: Discriminating clear-sky from cloud with MODIS algorithm theoretical basis document (MOD35). ATBD-MOD-06, version 4.0, NASA Goddard Space Flight Center
8. Novlan, D.J., Hardiman, M., Gill, T.E.: A synoptic climatology of blowing dust events in El Paso, Texas. In: 16th Conference on Applied Climatology, AMS, vol. J3.12, p. 13
9. Richards, J.A., Jia, X.: *Remote Sensing Digital Image Analysis*. Springer, Heidelberg (2006)
10. Specht, D.F.: Probabilistic neural networks for classification, mapping, or associative memory. In: *IEEE international conference on neural networks*, vol. 1(7), pp. 525–532 (1988)
11. Duda, R.O., Hart, P.E., Stork, D.G.: *Pattern Classification*, pp. 1–654. John Wiley and Sons, New York (2001)
12. Ramakrishnan, S., Selvan, S.: Image texture classification using wavelet based curve fitting and probabilistic neural network. *Int. J. Imaging Syst. Technol.*, 266–275 (2007)
13. Kanellopoulos, G., Wilkinson, G., Austin, J.: *Neurocomputation in Remote Sensing Data Analysis*. Springer, Heidelberg (1997)
14. Hanley, J.A., McNeil, B.J.: The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143(1), 29 (1982)

15. Crosson, W., Rickman, D., Limaye, A.: Optimal temporal scale for the correlation of AOD and ground measurements of PM<sub>2.5</sub> in a real-time air quality estimation system. *Jnl. Atmospheric Environment* 43(28), 4303–4310 (2009)
16. Mazurowski, M.A., Tourassi, G.D.: Evaluating classifiers: Relation between area under the receiver operator characteristic curve and overall accuracy. In: IEEE - INNS - ENNS International Joint Conference on Neural Networks, 2009 International Joint Conference on Neural Networks, pp. 2045–2049 (2009)
17. Fawcett, T.: ROC graphs: Notes and practical considerations for researchers. *Machine Learning* 31 (2004)
18. Tao, H., Yaohui, L., Hui, H., Yongzhong, Z., Yujie, W.: Automatic Detection of Dust Storm in the Northwest of China Using Decision Tree Classifier Based on MODIS Visible Bands Data. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium IGARSS 2005, Korea (July 2005)