

DUST STORM DETECTION THROUGH MODERATE RESOLUTION IMAGING SPECTRORADIOMETER: A MACHINE LEARNING PROBLEM

*P. Rivas-Perea, J. G. Rosiles**

The University of Texas El Paso
Dept of Electrical and Computer Eng
500 W University Ave, El Paso, TX 79968

J. C. Tilton

Comp. & Inf. Sciences & Tech. Office
NASA Goddard Space Flight Center
Greenbelt, MD 20771

ABSTRACT

Dust storms are of interest to study since they are correlated to an increase in mortality rates due to respiratory illness especially in the south-western U.S. With the aim of providing better tools to the understanding of dust storms, we present models for detection of dust storms from MODIS Terra Level 1B radiances, which can be applied in near real time with 1km resolution, in contrast to those models that are based on MODIS Aerosol Optical Thickness product that are produced several hours after reception and with a 10km resolution.

In this paper we present a multispectral multidimensional signal processing algorithm for detection of dust storms. The results showed that a Feed-forward Back-propagation Neural Network-based model perform better in classifying dust, and discriminating from other signatures such as clouds, smoke, etc.

Index Terms— Multidimensional signal processing, Image processing, Feedforward neural networks, Remote sensing.

1. INTRODUCTION

Dust storms are a major cause of several physical, environmental and economical hazards. Air pollution from dust storms is a significant health hazard for people with respiratory diseases and can adversely impact urban areas [1]. There is a direct correlation between exposure to high-levels of airborne particle concentrations (aerosols) and the increase in mortality rate from cardiovascular, respiratory illness and lung cancer. This situation is major concern for health and safety agencies as well as for the environmental [2] and geological science agencies [3].

Therefore, timely warnings of dust storms need to be fully functional in populated regions for health concerns and traffic

control [4]. However, in spite of the fact that several methods for detecting dust storms exist, there are still open questions in the detection process and in dust storms feature extraction. Furthermore, dust storms are still considered to be an open problem in analysis and modeling, as well as in the design of rapid response systems which require to minimize the processing time, and to produce results within moderate and high resolution imagery.

A number of approaches have been developed within the remote sensing community for detection and classification tasks utilizing multispectral data. However there are no specialized classification systems that use machine learning approaches to model dust storms. Detection methods based on principal components such as the one presented by Hillger *et. al.* [5] and Agarwal *et. al.* [6], improve the visualization of dust storms, however, such methods show other objects and artifacts besides the dust aerosols. Therefore, there is still a need to develop more accurate detection methods.

In this paper we will present two image processing methods for dust storm detection that are able to perform to a high level of accuracy, and are suitable for real-time applications. This models are based on feature extraction on the Moderate Resolution Imaging Spectroradiometer (MODIS) data. Such models are: Probabilistic model (PD), and Feed-forward Neural Network model (FFNN). The features extracted from multispectral MODIS data are within the near infrared reflectances. When the models are compared, the neural approach show the best numerical results compared to ground truths obtained from examples found in the literature. Furthermore, the probabilistic model show information not evident in the ground truth giving the ability to find non-trivial dust information.

In Section 2, the formation of a database of events is described. The spectral analysis of dust storms is introduced in Section 3, while in Section 4, the proposed models for dust storms are explained. In Section 5, the design of experiments are presented followed by a discussion of the results and findings. Finally, conclusions are drawn in Section 6.

*The first author performed the work while at NASA Goddard Space Flight Center (GSFC) 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 Graduate School Cotton Memorial Funding. Also the authors want to acknowledge the Texas Instruments Foundation for partially supporting this research.

2. EVENTS DATABASE, LABELING AND SELECTION OF SPECTRAL BANDS

We have collected 31 different dust storm events using the alerts record from the National Weather Service in Santa Teresa, New Mexico [7], as well as many other events reported in the literature. The data is labeled according to the standard used in weather forecast services. This labels are: DS, BLDU, SM, and C0. The events labeled as DS (10 events) are considered to be dust storms, and also the BLDU (15 events) correspond to blowing dust. The DS and BLDU both are dust events. However, the SM (smoke, 2 cases) and C0 (4 cases) are considered to be non-dust (background) information. The C0 correspond to land, oceans, clouds, etc. We force this distinction since smoke and clouds have an aerosol optical thickness very similar to the dust. The total number of events was then separated in design/test and validation sets, with 23 and 8 cases respectively.

In MODIS level 1B are available the bands needed for analysis and modeling. Hao *et. al.*[8] have demonstrated that bands $B20, B29, B31,$ and $B32$ (corresponding to $3.75\mu\text{m}, 8.55\mu\text{m}, 11.03\mu\text{m}$ and $12.02\mu\text{m}$ respectively) can be effectively utilized to enhance the visual perception of dust storms.

In this research we also used the band-math approach introduced by Ackerman *et. al.*[9] which proposes that band $B32$ and $B31$ should be subtracted to provide a better visual contrast on the images containing dust storms. Based on the previously cited research work we have selected four bands for our analysis and modeling. These four bands compose a five element feature vector:

$$F = [B20, B29, B31, B32, B32 - B31]. \quad (1)$$

3. DUST STORM MODELING

The proposed classification methodologies for remotely sensed data involve the usage of a uni-variate probabilistic method based on the estimation of the probability density function of the observed data. The reason for this is their reliability, and also the fact that the expected output result can be determined intuitively. On the other hand, machine learning methods such as supervised learning artificial neural networks, also provide with an ability to approximate the true distribution of observed data up to a given level of error. Thus, in this section we describe the two models for classification. First, we consider a simple probabilistic method based on the individual probabilities as a function of two random variables (the band-math approach). Second, a method based on a four-layered feed-forward back-propagation neural network.

3.1. Simplistic Probabilistic Modeling as a Function of Two Random Variables

Let X be a discrete random variable associated to the universe $\Gamma \in \mathfrak{R}$ of values for multispectral remote sensing data. Let $X^{(m)}$ be the a random variable associated with the values of the m -th spectral band of MODIS. Let $X_{\mathbf{n}}^{(m)}$ be the random variable associated with the n -th pixel of the m -th spectral band of MODIS. Let $f_{X_{\mathbf{n}}^{(m)}}(X_{\mathbf{n}}^{(m)} = x)$ denote the probability density function of the n -th pixel element of the matrix \mathbf{n} of the m -th spectral band of MODIS to have a value equal to x .

In this classification method we are interested on displaying the probability of the presence of a dust storm based on MODIS spectral band subtraction $B32 - B31$. Thus we are interested in the modeling of $f_{X_{\mathbf{n}}^{32-31}}(X_{\mathbf{n}}^{32-31} = x)$, which could be modeled assuming a Gaussian distribution on the form

$$f_{X_{\mathbf{n}}^{(m)}}(X_{\mathbf{n}}^{(m)} = x) = \frac{1}{\sqrt{2\pi\sigma_{X_{\mathbf{n}}^{(m)}}^2}} e^{-\frac{1}{2}\left(\frac{x - \mu_{X_{\mathbf{n}}^{(m)}}}{\sigma_{X_{\mathbf{n}}^{(m)}}}\right)^2} \quad (2)$$

where $\mu_{X_{\mathbf{n}}^{(m)}}$ is the expected value of the random variable x and $\sigma_{X_{\mathbf{n}}^{(m)}}^2$ is the variance associated with the random variable. MODIS band subtraction $B32 - B31$ is assumed to produce the random variable X .

The PDF $f_{X_{\mathbf{n}}^{(m)}}(X_{\mathbf{n}}^{(m)} = x)$, indeed, is theoretically defined as a function of two random variables $g(X_{\mathbf{n}}^{(m)})$, more specifically, the difference of two random variables

$$X_{\mathbf{n}}^{(32-31)} = g(X_{\mathbf{n}}^{(m)}) = X_{\mathbf{n}}^{(32)} - X_{\mathbf{n}}^{(31)} \quad (3)$$

with mean $\mu_g(X_{\mathbf{n}}^{(32-31)})$ and variance $\sigma_g^2(X_{\mathbf{n}}^{(32-31)})$. In spite of the fact that these parameters are unknown, they can be estimated by observation of the data. The only important variable to observe at this point is the function $g(\cdot)$ from (3) used in (2).

The estimations are computed over the number of events selected for modeling and design (see Section 2). In the approximation process we observe several samples (pixels), in the order of millions leading to a more accurate parameter estimation. We have computed the histogram for all the events and from the total frequency observed we have estimated the sample mean $\hat{\mu}_g(X_{\mathbf{n}}^{(32-31)})$, and the standard deviation $\hat{\sigma}_g^2(X_{\mathbf{n}}^{(32-31)})$.

3.2. Modeling Based on Multilayered Feed-forward Back-propagation Neural Networks

Multilayered feed-forward Neural Networks (FFNN) are of particular interest in pattern recognition and classifica-

tion applications because they can approximate any square-integrable function to any desired degree of accuracy, and can exactly implement any arbitrary finite training set. There exist many remote sensing data classification problems that have been successfully solved using neural networks besides dust storms. Therefore we have designed a FFNN to model a dust storm by approximating the probability density function $f_{C_j|X_n^{(m)}}(C_j = c|X_n^{(m)} = x)$. A simple FFNN contains an input layer and an output layer, separated by l layers (known as the hidden layer) or neuron units. Given an input sample clamped to the input layer, the other units of the network compute their values according to the activity of the units that they are connected to in the previous layers. In this model we consider the particular topology where the input layer is fully connected to the first hidden layer, which is fully connected to the second layer and so on up to the output layer.

Given an input $x \in X_n^{(m)}$, the value of the j -th unit in the i -th layer is denoted $h_j^i(x)$, with $i = 0$ referring to the input layer, $i = l + 1$ referring to the output layer. We refer to the size of a layer as $|\mathbf{h}^i(x)|$. The default activation level is determined by the internal bias b_j^i of that unit. The set of weights W_{jk}^i between $h_k^{i-1}(x)$ to in layer $i - 1$ and unit $h_j^{i-1}(x)$ in layer i determines the activation of unit $h_j^i(x)$ as follows:

$$h_j^i(x) = \Phi(a_j^i(x)), \quad (4)$$

where $a_j^i(x) = \sum_k W_{jk}^i h_k^{i-1}(x) + b_j^i$, $\forall i \in \{1, \dots, l\}$, with $h^0(x) = x$, and where $\Phi = \text{sigm}(\cdot)$ is the sigmoid activation function defined as $\text{sigm}(a) = \frac{1}{1+e^{-a}}$, which could be replaced by any desired activation function. Given the last hidden layer, the output layer is computed similarly by

$$\begin{aligned} \mathbf{o}(x) &= \mathbf{h}^{l+1}(x), \\ \mathbf{h}^{l+1}(x) &= \mathbf{\Phi}(\mathbf{a}^{l+1}(x)), \end{aligned} \quad (5)$$

where $\mathbf{a}^{l+1}(x) = \mathbf{W}^{l+1}\mathbf{h}^l(x) + \mathbf{b}^{l+1}$, and the activation function $\mathbf{\Phi}$ is to be defined later in Section 4. Thus, when an input sample x is presented to the network, the application of (4) at each layer will generate a pattern of activity in the different layers of the neural network and produce an output at the end.

3.2.1. Features And Events Selection

To model the dust storms with FFNN we used data from the same events utilized in the previous classification method. And, we'll use all the features described in Section 2. That is

$$F = [B20, B29, B31, B32, B32 - B31]. \quad (6)$$

4. EXPERIMENTS AND DISCUSSION

As established in the models for classification of dust storms, the millions of data points (elements of a feature vector) were used to estimate the parameters of the probabilistic model.

However, this is impractical to do for the FFNN model. Instead, we decided to reduce the number of data points for training following the criteria in [10] that establishes that the number of samples required for training the networks must be at least 3 times the number of bands used as features. Therefore, in the case of the FFNN, we utilized 500 times the size of F , thus, exceeding the minimum size criteria. Also it is important to say that the FFNN has a linear transfer function (*purelin*) at the output. The back-propagation method used to update the weights and biases was the Levenberg-Marquardt optimization method (*trainlm*). Also as a learning function we used the gradient descent with momentum weight and bias learning function (*learnngdm*). The stop conditions for the FFNN are either: 1) 100 epochs, 2) Performance=0, 3) Validation failures=5, and 4) Minimum performance gradient= 1×10^{-10} . The performance metric is the mean squared error (MSE). An internal set of samples for training, testing and validation was randomly selected to evaluate the generalization ability of the network. For testing purposes we select all the features extracted from all events for both methods, and tested to obtain the probability for that input to be dust storm. As a performance metric, we used the traditional standards: precision, accuracy, and one of the ultimate estimation of general performance, the Area Under the Receiver Operating Characteristics (ROC) curve (AUC).

4.1. Quantitative and Visual Results

The numerical results are shown in Table 1, where it is clear that the FFNN methods have the best performance metrics. The results of our algorithms are displayed for visual assessment of the outputs in Figure 1. Here we present two different kinds of figures, the first is a true color image re-projected using the traditional Mercator approach. Follows the second graphic that shows the probability of the presence of a dust storm.

Table 1. Performance in detection and in time response of the models. The processing time is in milliseconds per pixel vector (ms/pix).

	Precision	Accuracy	AUC	P. Time
PD	0.3938	0.4964	0.4993	0.0141
FFNN	0.4554	0.5426	0.7402	0.0472

5. CONCLUSION

The problem of dust storm detection has been addressed in this paper. First, we constructed a database of events from satellite observations of MODIS Terra satellite. From the samples in this database, we modeled a simple probabilistic

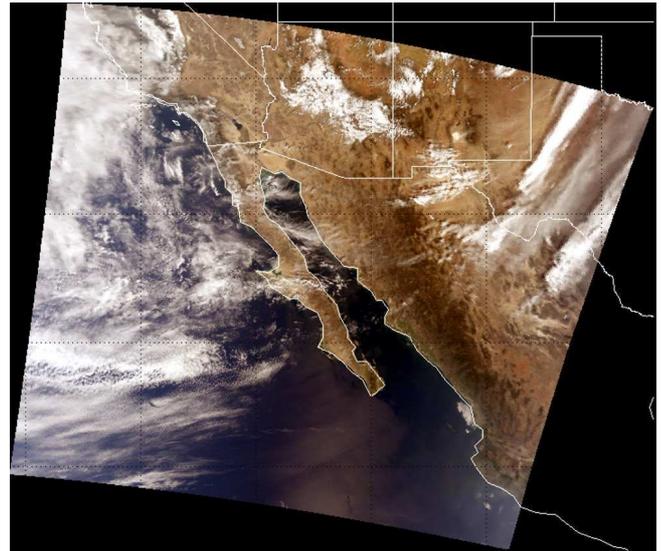
method specialized on measuring the probability of the presence of dust storms given MODIS Level 1B data. Then we designed a neural network architecture for dust storms detection. When we compared the probabilistic model against the Feed-forward Back-propagation Neural Network, FFNN, the latter reported a strong ability to inferring the relationship between spectral bands to classify dust and discriminate from other signatures, such as clouds, smoke, etc.

Moreover, the proposed probabilistic models are suitable for near real-time applications, such as direct broadcast, rapid response analysis, emergency alerts, etc. The probabilistic models are suitable for fast prototyping due to their simplicity, besides, the theory behind is easy to understand.

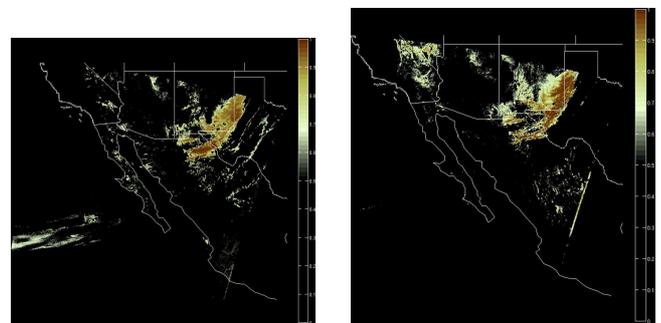
The work reported in this document is suitable for the study of the dust storm problem since the algorithms can output the dust presence to a resolution of 1km, which is an improvement over the methods based on the Aerosol Optical Thickness index (AOT) which has a 10km resolution. Besides, the MODIS AOT product is generated after several hours of the satellite pass, increasing the response time in the analysis and study of the dust storms.

6. REFERENCES

- [1] N.I. Rivera Rivera, K.A. Gebhart, T.E. Gill, J.L. Hand, D.J. Novlan, and R.M. Fitzgerald, "Analysis of air transport patterns bringing dust storms to el paso, texas," in *Symposium on Urban High Impact Weather*. American Meteorological Society, 2009, vol. J, p. 2.
- [2] S.A. Loomer, "Academic research opportunities at the national geospatial-intelligence agency (nga)," in *Proc. of SPIE*, 2006, vol. 6233, p. 623324.
- [3] W.B. Gail, "Remote sensing in the coming decade: the vision and the reality," *Journal of Applied Remote Sensing*, vol. 1, pp. 012505, January 2007.
- [4] H. El-Askary, M. Kafatos, X. Liu, and T. El-Ghazawi, "Introducing new approaches for dust storms detection using remote sensing technology," in *Proc. of IEEE International Geoscience and Remote Sensing Symposium*, 2003, vol. 4, pp. 2439–2441.
- [5] D.W. Hillger and G.P. Ellrod, "Detection of important atmospheric and surface features by employing principal component image transformation of goes imagery," *Journal of Applied Meteorology*, vol. 42, pp. 611–629, January 2003.
- [6] A. Agarwal, H.M. El-Askary, T. El-Ghazawi, M. Kafatos, and J. Le-Moigne, "Hierarchical pca techniques for fusing spatial and spectral observations with application to misr and monitoring dust storms," *IEEE Geoscience and Remote Sensing Letters*, vol. 4, pp. 678–682, October 2007.
- [7] D.J. Novlan, M. Hardiman, and T.E. Gill, "A synoptic climatology of blowing dust events in el paso, texas from 1932–2005," in *Proc. 16th Conference on Applied Climatology*. American Meteorological Society, 2007, vol. J3.13, p. 12.
- [8] X. Hao and JJ. Qu, "Saharan dust storm detection using modis thermal infrared bands," *Journal of Applied Remote Sensing*, vol. 1, pp. 013510, April 2007.
- [9] S. Ackerman, K. Strabala, P. Menzel, R. Frey, C. Moeller, L. Gumley, B. Baum, C. Schaaf, and G. Riggs, "Discriminating clear-sky from cloud with modis algorithm theoretical basis document (mod35)," *NASA Algorithm Tech. Background Doc. Publ. Prod. ID MOD35 (NASA Goddard Space Flight Center)*, vol. 1, pp. 195, November 1997.
- [10] I. Kanellopoulos, G. G. Wilkinson, F. Roli, and J. Austin, "Neurocomputation in remote sensing data analysis," Springer-Verlag, 1997.



(a) True color image $R= B1, G= B4, BB3$



(b) Dust probability FFNN

(c) Dust probability PD

Fig. 1. Dust storm event of February 19th 2004.