Automatic Detection of Fire Smoke Using Artificial Neural Networks and Threshold Approaches Applied to AVHRR Imagery

Zhanqing Li, Alexandre Khananian, Robert H. Fraser, and Josef Cihlar

Abstract—In this study, satellite-based remote sensing techniques were developed for identifying smoke from forest fires. Both artificial neural networks (NN) and multithreshold techniques were explored for application with imagery from the Advanced Very High Resolution Radiometer (AVHRR) aboard NOAA satellites. The NN was designed such that it does not only classify a scene into smoke, cloud, or clear background, but also generates continuous outputs representing the mixture portions of these objects. While the NN approach offers many advantages, it is time consuming for application over large areas. A multithreshold algorithm was thus developed as well. The two approaches may be employed separately or in combination depending on the size of an image and smoke conditions. The methods were evaluated in terms of Euclidean distance between the outputs of the NN classification, using error matrices, visual inspection, and comparisons of classified smoke images with fire hot spots. They were applied to process daily AVHRR images acquired across Canada. The results obtained in the 1998 fire season were analyzed and compared with fire hot spots and TOMS-based aerosol index data. Reasonable correspondence was found, but the signals of smoke detected by TOMS and AVHRR are quite different but complementary to each other. In general, AVHRR is most sensitive to low, dense smoke plumes located near fires, whereas smoke detected by TOMS is dispersed, thin, elevated, and further away from fires.

Index Terms—AVHRR, classification, fire, neural networks, smoke.

I. INTRODUCTION

BIOMASS burning emits a large amount of greenhouse gases and aerosols into the atmosphere. Approximate estimation showed that the annual amount of CO_2 released into the atmosphere due to biomass burning is about 114 Tg in the tropics [28] and 62.3 Tg in boreal zone [37]. Trace gases and aerosol particles produced by fires play important roles in atmospheric chemistry, cloud microphysics, temperature, and radiation balance in the lower atmosphere. Fire can thus impinge significantly on local weather and climate [7], [19]. Fire impact on weather is mainly due to the attenuation of sunlight by smoke particles, which is usually short lived. Robock [35] attempted to relate temperature forecast errors to large fire events occurring in the world's major boreal forests in Canada, China, and Siberia. It was found that, without considering the direct influence of the fires, temperature prediction in a nearby region tends to be overestimated by 1.5–7 °C due to the cooling effect of smoke. To a lesser extent, smoke can have an impact that extends far beyond the region of fire activity. Smoke plumes may travel over hundreds, or even thousands of kilometers horizontally and reach up to the stratosphere under certain atmospheric circulation conditions [9]. A major fire episode in northwestern Canada was found to influence significantly air quality in the southeastern U.S. and eastern seaboard [39].

The climatic impact of smoke is twofold: cooling due to smoke particles and warming due to greenhouse gases. Smoke particles scatter and absorb incoming solar radiation, thereby having a cooling effect at the surface, but warming effect on the atmosphere [23]. Since the magnitude of the scattering effect outweighs that of absorption, smoke has a net cooling effect at the top of the atmosphere-surface system [14]. Smoke can also modify the short wave reflective properties of clouds by acting as cloud condensation nuclei [31]. Under a limited supply of water vapor, an increased number of nuclei result in smaller cloud droplets that have higher reflectivity than larger cloud droplets [18]. The cooling due both to the direct and indirect effects of smoke could potentially offset the warming effect of increasing CO_2 content [28], but they act on different temporal and spatial scales. The latter has a much longer lifetime and covers larger areas. Understanding such numerous and complex effects of smoke on weather and climate requires a good knowledge of the spatial and temporal variation of smoke and its optical properties, which is only feasible by means of satellite observation. Discrimination of smoke on satellite imagery is a prerequisite to study and retrieve physical, chemical, and optical properties of smoke.

Identification of smoke is by no means a trivial task using spaceborne data. As is demonstrated later, there is a large overlap in the spectral signature of satellite measurements between smoke and other scene types such as clouds and background surfaces. So far, very few investigations have focused on the identification of smoke except for some studies that used somewhat *ad hoc* approaches to identify smoke for pursuing other research themes. The most commonly used method of identifying smoke is to assign different colors to different channels or channel combinations [5], [32]. The resulting false-color images can provide visual separation of smoke from other objects. For example, Kaufman *et al.* [17] assigned AVHRR channel 1 to red, channel 2 to green, and inverse channel 4 to blue, generating a composite image

showing smoke plumes. Such an approach can hardly be used for automatic processing masses of satellite imageries. Another popular approach is thresholding. Christopher *et al.* [4] examined various AVHRR channels and their combinations for distinguishing smoke. He then applied a texture analysis to these channels and their combinations.

This study developed new remote sensing methods for detecting smoke. Unlike many previous studies dealing mainly with tropical fires [17], [30], [4], the methods proposed here address smoke from boreal forest fires, although the principles of the methods are applicable to other types of biomass burning as well. Due to relatively poor knowledge and limited investigations on boreal fires, more attention needs to be paid to this biome. We have developed a suite of remote sensing techniques for systematically monitoring and studying boreal forest fires, including the detection of hot spots [22], [25], mapping of burned areas [26], [8], and identification of smoke plume (this study), retrieval of smoke optical properties [38], and studying the radiative impact of smoke on Earth's radiation budget [23], [24]. The algorithms are designed for routine operational application to daily satellite data from the advanced very high resolution radiometer (AVHRR) aboard the National Oceanic and Atmospheric Administration (NOAA), Washington, DC, series of satellites.

Data used in the study are introduced in the following section. Section III describes the algorithms, which employ both neural network and threshold techniques. The performance of the algorithms is evaluated by various means that are also addressed in this section. Section IV presents some routine products generated by applying the algorithms to AVHRR data obtained in 1998. The smoke product is compared with fire hot spots and an aerosol index data set derived from the total ozone mapping spectrometer (TOMS) aboard on a different platform [12], [13].

II. Data

This study employs AVHRR images from NOAA-14 acquired in 1998, while the algorithms have also been used to process AVHRR data covering Canada in other years. NOAA-14 has a daytime overpass around 2–3 PM in Canada with a viewing plane of $45°$ relative to the principal plane. After the data were received at the Prince Albert station in Saskatchewan, they were radiometrically calibrated and geometrically referenced using the geocoding and compositing (GEOCOMP) AVHRR data processing system [36]. The calibration for visible (ch. 1) and near-IR (ch. 2) measurements was based on the method of Rao and Chen [33], with their coefficients updated from time to time. The thermal channels (3–5) were calibrated using onboard blackbody references. Pixel locations were first computed from an orbit model that takes into account spacecraft orbit, velocity, attitude and altitude, earth rotation and curvature. Daily AVHRR images composited across Canada (5700×4800 pixels) were used. The data contain top of atmosphere reflectance in channels 1 and 2 ($R1$, $R2$) and brightness temperatures in channels 3 to 5 ($BT3$, $BT4$, and $BT5$).

These channels exhibit some distinction in the characteristics of smoke, clouds, and underlying surface, which is the basis of smoke identification. On the other hand, there exists a considerable overlap in the magnitude of observations between smoke
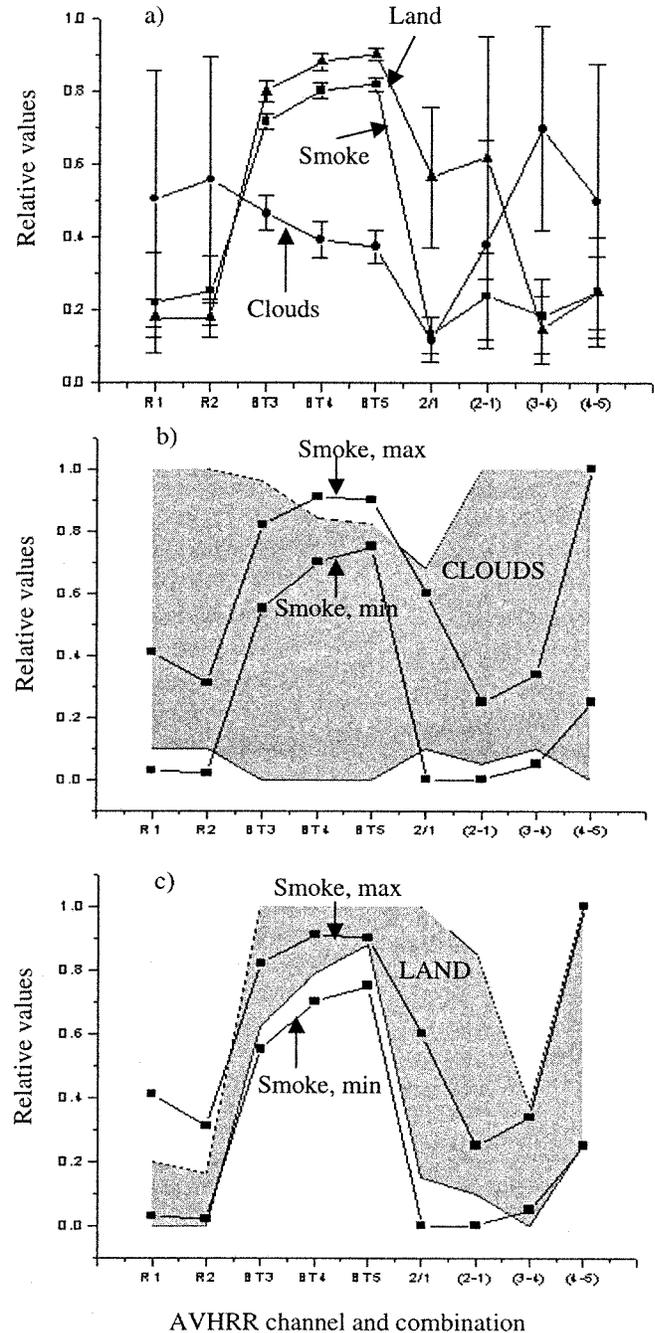

Fig. 1. Relative variability of the spectral signals of smoke, clouds, and clear land surfaces for five AVHRR channels and some of their combinations. (a) Means and standard deviations for all three subjects, (b) relative maximum and minimum values for smoke (curves) and clouds (shaded area), and (c) same as (b) but the shaded area represents clear land.

scenes and for other scene types, most notably between cloud and land. Fig. 1 illustrates the overlapping of spectral signatures in all the channels and some channel combinations among three distinct scene types [Fig. 1(a)], namely, smoke and clouds [Fig. 1(b)], and smoke and land [Fig. 1(c)]. For different channels to be comparable, relative values (R) are used that are computed from the absolute values (V) according to

$$R = (V - V_{\min})/(V_{\max} - V_{\min}) \qquad (1)$$

where $V_{\max}$ and $V_{\min}$ denote the maximum and minimum observation values. R thus varies between 0 and 1. Fig. 1(a) shows the mean and standard deviations, while Fig. 1(b) and (c) are the maximum (top curves) and minimum (bottom curves) values.

The figure reveals the potential and limitations in separating smoke from clouds and land using AVHRR data. In general, the reflectance and brightness temperature of dense smoke have intermediate values between those of clouds and land. The reflectance of smoke is usually less than that of clouds, but higher than that of the underlying surface, while the converse is true for brightness temperature. From Fig. 1(a), it appears possible to differentiate smoke from most clouds and land based on the relatively large difference in brightness temperatures in channels 3–5 ($BT3$, $BT4$, $BT5$). The ratio of the reflectance at channels 2 and 1 ($R2/R1$) is useful to identify smoke over land, and the difference between $BT3$ and $BT4$ is useful to separate smoke from clouds. Overall, the three thermal channels are superior to the two shortwave channels. Although reflectance of smoke is generally less than that of clouds, the latter has so large a range of variation that it is difficult to use it to discriminate smoke pixels from cloudy pixels.

In fact, it is a general problem facing any classification using AVHRR that the large ranges of variation in all AVHRR channels cause the overlap between different scene types. This is illustrated more clearly in Fig. 1(b) and Fig. 1(c), which show the entire ranges of variation in terms of relative maximum and minimum values, with the curves denoting smoke and shaded areas for clouds [Fig. 1(b)] and for land [Fig. 1(c)]. It is observed that the reflectance and brightness temperature for smoke, clouds, and land overlap considerably. Although the number of overlapping pixels is small relative to the total number of land or cloudy pixels, it is comparable to, or even greater than, the number of smoke pixels. The spectral overlap is due partially to turbulent diffusion processes associated with smoke and clouds, which produces large variability in the parameters and leads to fuzzy boundaries between different scene objects.

The results shown in Fig. 1 were obtained by analyzing AVHRR data acquired across Canada. For regional studies, the overlap range is smaller depending on smoke amount and cloud thickness, meteorological conditions, as well as the spatial and temporal distributions of fires and smoke. In some special circumstances, smoke, clouds, and land are readily separated by reflectance and brightness temperatures using even single channel measurements, but in general this is very difficult.

III. ALGORITHM

In order to cope with a variety of smoke conditions, the detection algorithms proposed in this study are based on both artificial neural networks and multithreshold approaches. Each consists of two major steps: identifying potential areas covered by smoke using the neural networks or threshold classifier, then removing false-classified pixels by applying additional tests, texture analysis, and spatial filtration. The threshold and texture parameters were chosen and optimized following thorough investigations and analyses of the spectral signature and texture of smoke, clouds, and land with allowance for their spatial and temporal variability.

A. Neural Network Method

The neural networks (NN) approach has the capability to learn patterns whose complexity makes them difficult to analyze using other conventional approaches [3], [10], [20], [21]. The NN is useful for smoke identification due to its ability to find and learn complex linear and nonlinear relationships in the radiometric data between smoke, clouds, and land.

In the present study, a commercial NN package, named NeuroSolutions Professional from NeuroDimension, Inc., is used. The multilayer perceptron (MLP) neural network of NeuroSolutions package used for the image analysis is a two-layer forward feed network (FFN) with five inputs from the five channels of AVHRR, one hidden layer with ten processing elements, and one output layer. The output layer included three neurons. The number of neurons in the output layer is equal to the number of desired parameters of the output vector, which are "smoke," "clouds," and "land" in this study. Individual computational elements of an FFN are referred to as neurons or processing elements (PE). Each neuron consists of a vector of modifiable weights or connection strength. The task of a neuron is to map a given input vector into a single output that is transmitted to other neurons. Each element of an input vector is multiplied by a corresponding weight and added together to produce a net input. The neuron uses an activation function to transform the net input into a single output. In our NN, we used two kinds of activation functions. The hyperbolic tangent activation function is used for the hidden layer, and an additional softmax activation function is used for output layer [29]. The softmax function is used to interpret the output of the NN classification in terms of posterior probabilities whose outputs for all classes sum to one. Neurons are arranged in successive layers with connections between the neurons of two layers but with no connections between neurons within the same layer. In this layer arrangement, data flow is unidirectional starting from the input layer. Weights are commonly computed by minimizing the difference between network outputs, once a set of input data vectors or patterns have been propagated through the network. The network was trained to distinguish smoke from clouds and the underlying surfaces, including both land and water, with the standard backpropagation method.

The training data were selected from AVHRR images containing active forest fires. Input parameters to the NN include reflectance from channels 1 and 2, and brightness temperatures from channels 3, 4, and 5 without considering any of their combinations. Training pixels were obtained from representative polygons containing smoke, clouds, land cover, and water. Three outputs were generated by the NN corresponding to the three types of classified objects (smoke, cloud, and land). Each output is encoded to denote one classified object [27], [2]. To this end, in the training data, an input vector of a class was assigned a desirable output. To encode the outputs, a softmax output activation function was used. According to the softmax activation function, the output vector for smoke, clouds, and land categories was represented using binary encoding as shown in the matrix in Table I.

To train the NN and test its performance, we employed AVHRR images containing forest fires in northern Quebec in

TABLE I
ENCODING MATRIX USED IN THE NEURAL NETWORKS CLASSIFICATION

	Smoke code	Cloud code	Land code
Smoke	1	0	0
Cloud	0	1	0
Land	0	0	1

TABLE II
MEANS AND STANDARD DEVIATIONS OF THE NN OUTPUTS FOR PRESELECTED
SCENE TYPES OF SMOKE, CLOUD, AND LAND

	"Smoke"	"Cloud"	"Land"
Output NN "Smoke"	**0.798 (0.045)**	0.040 (0.009)	0.068 (0.005)
Output NN "Cloud"	0.046 (0.007)	**0.958 (0.009)**	0.003 (0.0002)
Output NN "Land"	0.156 (0.051)	0.002 (0.0007)	**0.929 (0.005)**

July 1998 and in northern Saskatchewan and Manitoba in the middle of August 1998. The training data set included dense and thin smoke, different types of clouds, and various land cover types typical of the boreal forest zone. The total number of pixels used for training and testing the NN was more than 200 000. 30% of the pixels were randomly selected from each class and used for training the NN, while the remaining pixels served as test samples. The averaged NN output values are presented in Table II. In accordance with the softmax function of the NN output, the values in the diagonal describe the probabilities of correct classification or the resemblance to a "pure" scene. The off-diagonal values denote the probabilities of misclassification or deviation from a pure scene. The diagonal values in the Table are close to unity, as the data include rather pure scenes: dense smoke, thick clouds, and clear land. In the case of optical thin smoke or cloud, the output values are more dispersed due to class mixing.

Fig. 2(a) shows an output image from the NN classification of a large smoke plume (400×100 km^2) observed on August 30, 1998 in northern Saskatchewan. The image is a three-band false color composite based on the three NN outputs, with "smoke" in red, "clouds" in green, and "land" in blue. Also presented in Fig. 2(b)–(d) are the frequency histograms of each output value. Here the y-axis shows the percentage of the number of pixels, and the x-axis shows the corresponding outputs of the NN (encoding values). They demonstrate sufficient separation in the NN outputs between smoke and the other two scene types. The majority of smoke pixels have an NN output larger than 0.5, while the outputs for land and cloud are infrequently larger than 0.5. The red color in the image corresponds to relatively thick smoke that dominates the image. In the yellow-green portion of the image lie pixels that are attributed more to clouds. Note that these clouds were probably formed inside a smoke plume. The violet part of image corresponds to optically thin smoke ($R1 < 0.2$), which has a spectral signal influenced by the under-
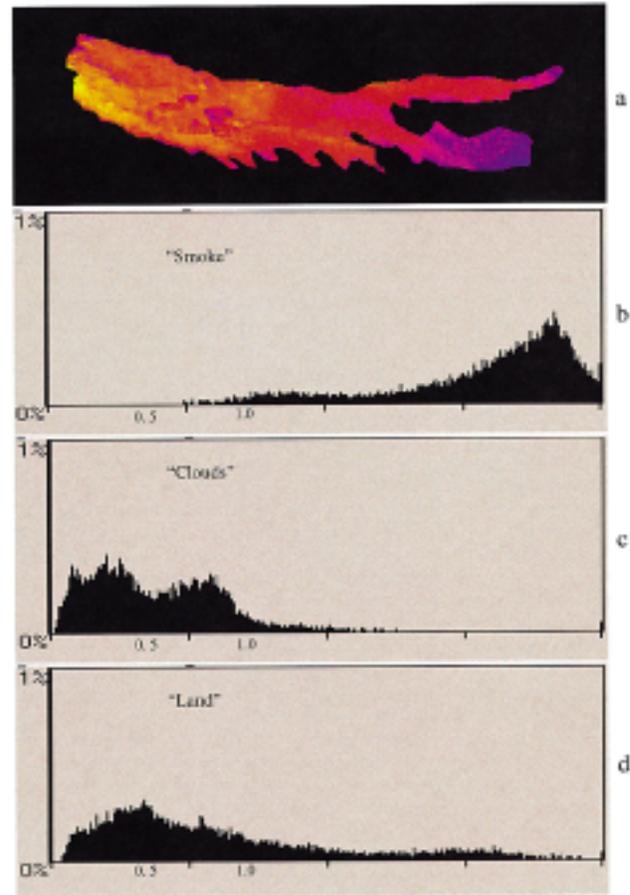

Fig. 2. (Top) False color composite image of smoke, clouds, and land generated by assigning the NN outputs of smoke, clouds, and land to red, green, and blue colors, respectively. (Second panel) Frequency histograms corresponding to each NN output for smoke, (third panel) clouds, and (bottom panel) land.

lying land surface. Reduced smoke concentration due to turbulent diffusion of the smoke plume leads to gradually decreasing values of the NN output for "smoke" and increasing values of the NN output for "land."

Fig. 3 shows changes in the NN output along a transact across a smoke plume starting at the core of the plume and moving toward its edge. It illustrates gradual changes in NN outputs for smoke and land during a transition from smoke to land. The vertical lines show the range of variation. Reflectance in AVHRR channel 1, which is proportional to smoke concentration and the NN output for "smoke," shows a strong correlation. The spatial variations of the NN output for smoke across a plume follows the Gaussian distribution due to turbulent diffusion [1]. The visible radius of the smoke plume in the study area is approximately 60 km. At this distance, smoke optical depth is reduced to approximately 0.1, leading to reflectance close to that of the land background ($R1 \approx 0.07$–0.08). The NN output for "smoke" also decreases to its background value around 0.1 near the boundary of the smoke. Fig. 3 also shows that the NN outputs for "smoke" and for "land" are negatively correlated with a combined value near unity. Therefore, the NN approach does not only identify smoke, but also provides information on its loading.
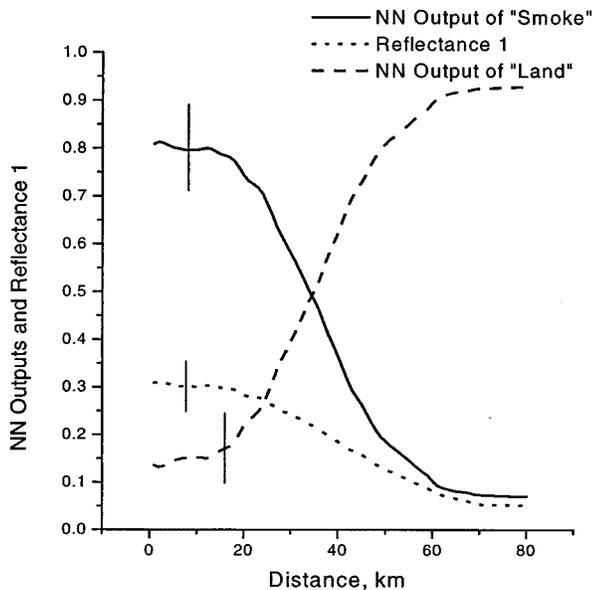
Fig. 3. Comparison of the NN outputs and channel 1 reflectance along a transact between smoke and land. The bars show the range of variability.

B. Noise Reduction

Although the NN is powerful enough to make optimal use of all signals pertaining smoke, its performance is inherently limited by the input satellite data as demonstrated in Fig. 1. To reduce inevitable false classifications, additional tests are necessary to eliminate noisy pixels from the "smoke" NN output. The majority of misclassified smoke pixels are located near the boundaries of relatively warm clouds over land covered by ice and snow such as the Rocky mountains in British Columbia and Alberta. A number of discrimination tests are employed to reduce the noise levels in the "smoke" outputs. The first test uses a median filter (5 × 5 or 9 × 9 pixels). It computes the median value of the output of smoke index within a rectangular filter window surrounding each pixel. The median filter smooths image data, preserves the edges of a smoke plume, and removes small clusters of noisy pixels produced by small clouds and randomly distributed sources of underlying background. The second test eliminates "smoke" pixels of output lower than 0.1. The third test eliminates false smoke pixels caused by cloud boundaries using difference in spatial variance between smoke and noise. Spatial variance is computed from the standard deviation of "smoke" output values within a 5 × 5 pixel neighborhood. Since smoke is normally more homogeneous than cloud boundaries, this test rejects false alarm pixels of the "smoke" channel with variance larger then 1.1. In addition to the above threshold tests, one special screening is applied. It is designed to eliminate false smoke pixels occurring over areas covered by permanent ice and snow, reflective barren land (bare soil and rock), and some individual pixels that have very low contrast between smoke and other scene types in all AVHRR channels. This test resorts to the use of a land cover classification derived from AVHRR data [6].

Fig. 4 shows a false-color composite of the image over Canada with fires and smoke observed in Northwest Territories on July 16, 1998. It illustrates the final results of the NN classification separating smoke, clouds, and land. Smoke is shown in yellow, clouds in pink and white, and land in green. It was created by combining two NN outputs ("smoke" and "clouds") in red, reflectance from channel 2 AVHRR in green, and output NN of "cloud" alone in blue. Fire pixels, detected using the algorithm of Li *et al.* [25], [26], are shown in red. The color image and smoke mask illustrate the potential of the NN to identify dense and thin fresh smoke located close to the fires. In some cases, it is difficult to confirm the presence of smoke due to the lack of independent ground-truth data. For example, to the right of fresh smoke plumes there is a long stretch feature also selected as smoke. Since there are no adjacent hotspots, it is difficult to confirm these "smoke" pixels, which have the same or similar spectral and textural properties as smoke. They may be dispersed downwind smoke, clouds mixed with smoke, or clouds having the same parameters as smoke. The problem stems partly from the use of a large training data set that covers the whole country. Of consequence, the radiometric signatures of the three typical scenes overlaps, rendering uncertainties in the classification. To solve the problem, we also used multithreshold approach to fine tuning the classification on a regional basis, which can also be run alone for fast operational detection of smoke from AVHRR composite image over Canada.

C. Multithreshold Tests

The multithreshold approach is based on differences in the reflectance of AVHRR channel 1 and channel 2 and in the brightness temperature of channel 4. Like the NN technique, it involves two major steps: marking potential smoke pixels and removing false pixels. Both steps are accomplished using threshold tests. The first employs thresholds of the ratio of AVHRR channel 1 and channel 2 reflectance ($R1$ and $R2$) and a brightness temperature of AVHRR channel 4 ($BT4$):

$$0.9 \leq R2/R1 \leq 1.5 \quad \text{and} \quad BT4 \leq 298 \text{ K}. \quad (2)$$

Pixels passing this test are considered to be either smoke or cloud. Otherwise, they are deemed land pixels. The second test employs channel 4 to further separate smoke and cloudy pixels

$$BT4 \leq 280 \text{ K}. \quad (3)$$

Pixels passing this test are regarded as clouds as they are usually higher and colder than smoke. The third test is introduced to eliminate warm nonbright clouds from "smoke" pixels

$$BT4 \leq 284 \text{ K} \quad \text{and} \quad R1 \geq 0.35. \quad (4)$$

Fig. 5 shows a flowchart of the multithreshold tests. The threshold values were established using the training database from AVHRR observation of forest fires over Canada in 1998 [25], [26]. If the threshold approach is run standalone, i.e., the NN approach is not run ahead, the various screening processes described earlier ought to be applied including median spatial filtering, variance testing, and background checking using a land cover mask.

Fig. 6 (left panel) shows a smoke image produced using the threshold algorithm applied to the AVHRR image across
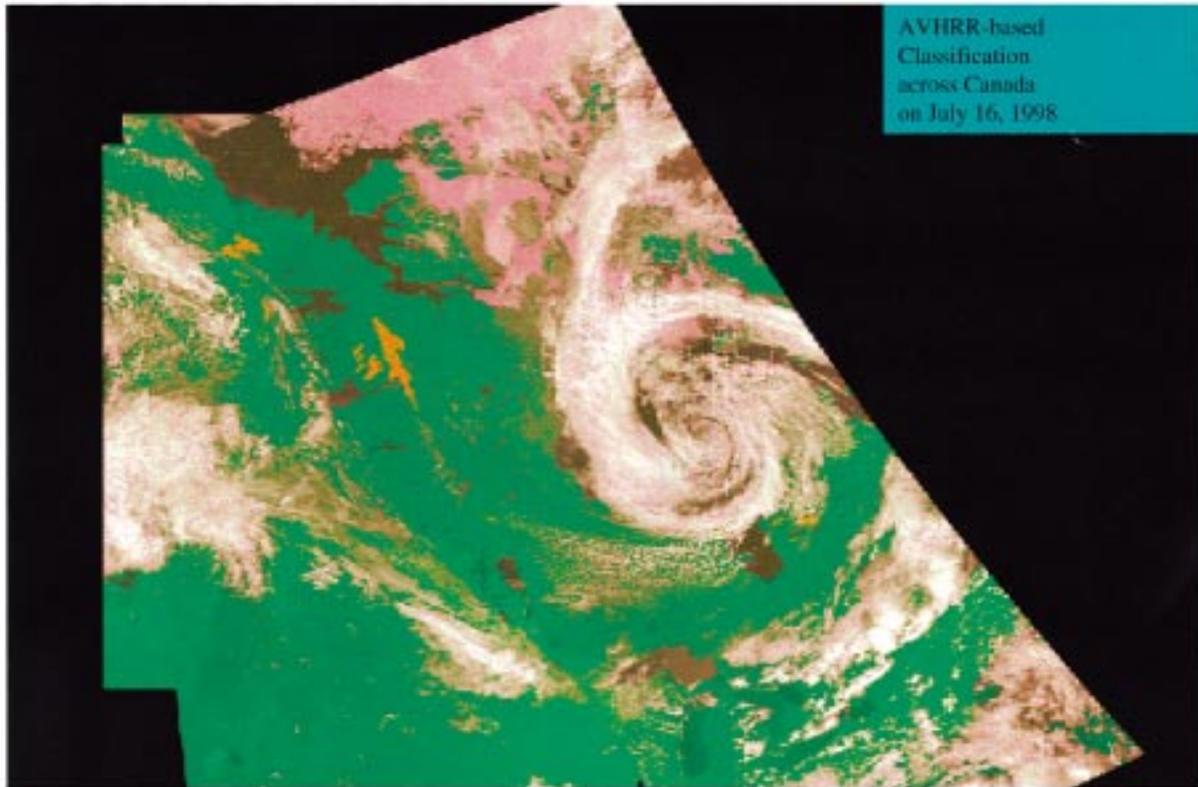
Fig. 4.　Forest fire and smoke across Canada on July 16, 1998, identified by the neural networks. Smoke plumes, clouds, and clear land are shown in orange, white/gray, and green, respectively.

Canada on August 11, 1998. By assigning the combined "smoke" and "clouds" masks in red, reflectance from channel 2 AVHRR in green, and cloud mask in blue, this creates a composite image where smoke appears orange, clouds white and grey, and land green. Shown on the right panel is an image over a much smaller region encompassing most of the fires occurring in Saskatchewan and Alberta on the same day, identified by the NN. There were more than a dozen of large fires over an area of 1000 km^2. The image shows not only the thick smoke plumes, but also the widespread, persistent smoke haze. Some thin smoke is not clearly seen over land, but it is readily discernible over dark water bodies. The overlapped clouds and smoke are also successfully separated.

D. Performance Evaluation

Relative to the use of individual AVHRR channels and their simple combinations, both the NN and multithreshold approaches described above have improved capability to identify smoke plume and to produce images providing visual contrast between smoke, clouds, and the underlying background. To demonstrate this quantitatively, the degree of separation is measured by the three-dimensional (3-D) Euclidean distance between each pair of scene objects. Of course, our main interest here is the separation between smoke and cloud, and between smoke and land, although distinctions between cloud and other two subjects are useful for other studies. Each of the objects is designated by three variables (X, Y, Z) that are assigned three basic colors (red, blue and green) to generate false color images. These variables include the outputs of the NN

classification and the combinations of AVHRR channels used in the threshold algorithms. For comparison, some selected individual channels are also tested, which were employed in some previous studies. Since the images are linearly enhanced in this study, the Euclidean distance defined in the following quantifies the capability of separation between two subjects (A and B):

$$D_{AB} = \sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2 + (Z_A - Z_B)^2}. \quad (5)$$

Note that all the variables (X, Y, Z) are normalized to the same scale (0–1). The distances between three pairs, namely, smoke-clouds, smoke-land, and clouds-land are computed and shown in Fig. 7. The figure indicates that the combination of three NN output channels produces the best visual contrast between smoke, clouds, and land. Other methods may render good separation between one pair of objects but often fail for other pairs.

In addition, the classification accuracy of the algorithms is assessed using an error matrix [15] which describes probabilities of each scene type being correctly identified (diagonal elements in the matrix), and misidentified into different categories (off-diagonal elements). From this matrix, the overall accuracy and commission and omission errors can be computed. The overall accuracy was computed as the ratio of the sum of numbers in the diagonal divided by the total number of all scenes. The commission error is the ratio of the number of cases misclassified as one scene divided by the total number of this scene, while the omission error is the ratio of the number of scenes misclassified
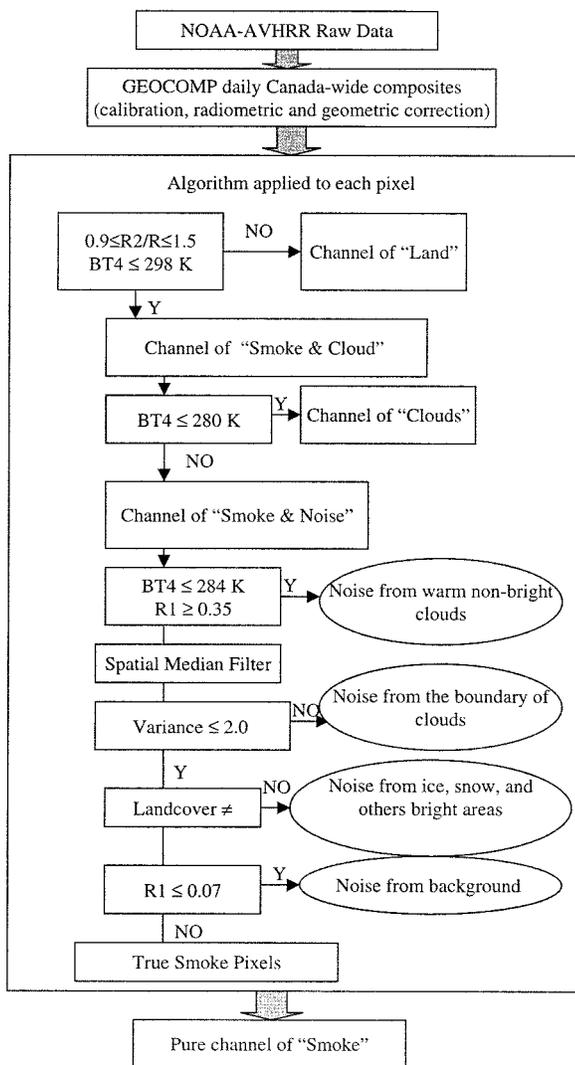
Fig. 5. Flowchart of the multithreshold algorithm.

IV. APPLICATION

Smoke from forest fires is the primary disturbance to the relatively low loading of background aerosols across the world's major boreal forests in Canada and Russia. These forests are subject to widespread periodic burning induced primarily by lightning. On average, tens of thousands of fires occur each year across the boreal zone [16]. Because of the high biomass content, boreal forest fires tend to be more intense and last much longer than tropical fires. As a result, smoke from these fires can usually travel a long distance and extend to a high altitude. For example, Hsu *et al.* [12] found a close linkage between fires occurring in western Canada and the high loading of absorbing aerosol over Greenland using aerosol index (AI) data derived from the TOMS. The AI is a measure of the wavelength-dependent reduction of Rayleigh scattered radiance by aerosol absorption relative to pure Rayleigh atmosphere [11], [13]. The AI was defined such that positive values generally correspond to UV-absorbing aerosols situated 1.5 km above the Earth's surface. TOMS is not sensitive to boundary layer absorbing aerosols [11]. Using several spaceborne data sets such as the Polar Ozone and Aerosol Measurement (POAM) III and Stratospheric Aerosol and Gas Experiment (SAGE) II, Fromm *et al.* [9] found that aerosol from large fires in Canadian and Russian forests can travel over half the globe and extend into the stratosphere (15 km).

While both studies referred to AVHRR images, no direct comparisons were made against either fire hot spots or smoke plume. Since the observation principles and characteristics of these instruments (TOMS, POAM/SAGE, AVHRR) are very different, comparisons of different fire smoke products provide a means of validation and/or consistency check and more complete information concerning the smoke. Note that the smoke identified by AVHRR is usually fresh and located near fire locations at relatively low altitudes, whereas that identified by TOMS or POAM/SAGE is older and located at higher altitude away from the fire origin. If correctly identified, smoke detected with these sensors should correspond to each other with lags in time and space dictated by atmospheric circulation conditions.

We implemented the smoke detection algorithm as described above to process daily AVHRR imagery across Canada during the entire fire season (May–October) for a few years. In this paper, only the results obtained for 1998 are analyzed. As is shown in Fig. 8, there were two major fire episodes in 1998 that occurred in July and August. The one in August is rather intensive and widespread, which was studied by Hsu *et al.* [12], [13] using TOMS AI data. For the sake of comparison, the same episode is investigated here. Fig. 9 presents side-by-side comparisons of AVHRR-based hot spots (left panels), smoke-cloud-land composite images (middle panels), and TOMS-based aerosol index images (downloaded from the NASA/TOMS home page: http://toms.gsfc.nasa.gov). It is seen that on August 3, there was widespread fire activity across western and southern Northwest Territories and scattered fires in Yukon Territories, Manitoba, and Ontario. Almost all of the fires detected by satellite are confirmed by Canadian forest fire agencies [26]. Smoke associated with these fires was detected

into other scenes divided by the total number of this scene. To this end, 13 nationwide AVHRR images of 4400 × 4400 pixels with significant fire activities were selected. Each of the selected images contain at least 300 fire pixels with more than 25 000 identifiable smoke pixels and 50% cloud cover. The smoke detection algorithms were applied to these images. The resulting masks of smoke, clouds, and land were compared against results obtained by a supervised classification. Table III is the error matrix including the statistics of the NN classification. Reference data given in the columns represent the real number of pixels belonging to each category identified by supervised classification, i.e., the sums of the elements appearing in the same column are deemed as true. The numbers in the rows are classification results obtained by the NN. Also included in the table are the overall classification accuracy, omission, and commission errors. Both the commission and omission errors for smoke scenes are in the neighborhood of 27%, which are considerably larger than for clouds and land. This is because the areas of cloud and clear land scenes are much larger than that of smoke. It should be noted that this accuracy assessment contains uncertainties due to absence of real ground truth information.
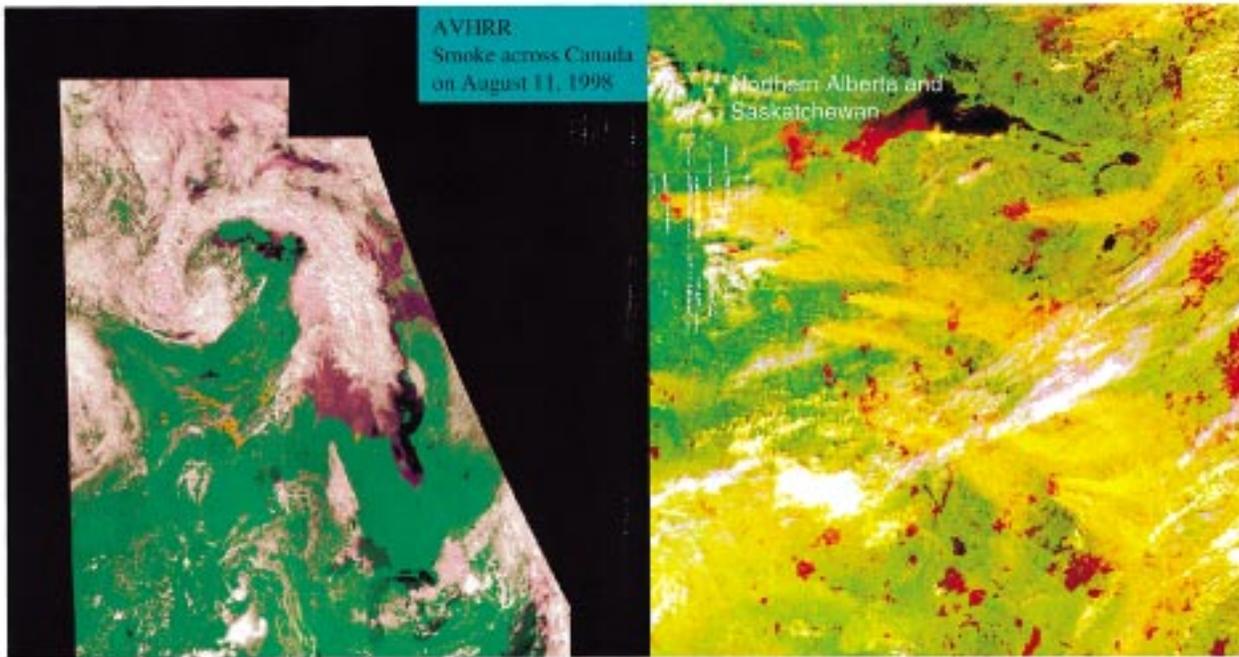
Fig. 6. (Left) False-color composite image of forest fires (red), smoke (orange), clouds (white and gray), and land (green) identified by the multithreshold approach across Canada on August 11, 1998. (Right) Regional false-color image classified by the NN in northern Alberta and Saskatchewan on the same day.
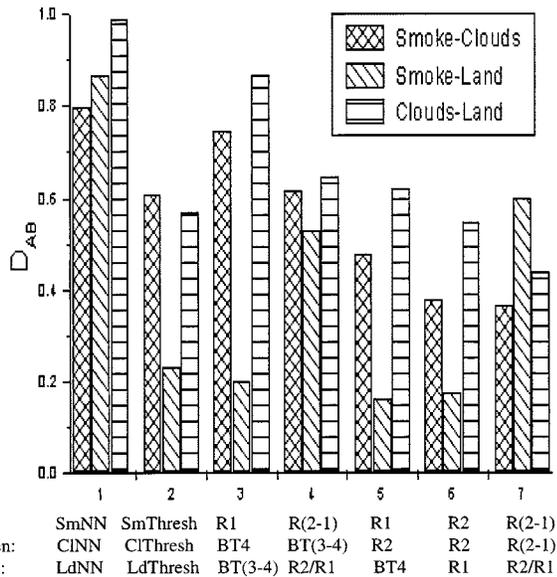

Fig. 7. Euclidean distance between smoke and clouds, smoke and land, and clouds and land, determined by three variables (the outputs of the NN and multithreshold algorithms, AVHRR channels and their combinations) that characterize the three subjects.

TABLE III
STATISTICS OF THE CLASSIFICATION RESULTS

Scene Types (Output)	Reference Data (Input)				
	Smoke	Clouds	Land	Relative Errors	
				Omission	Commission
Smoke	70 000	20 000	8 000	26.4 %	28.6 %
Clouds	10 000	11 295 000	3 000	0.3 %	0.2 %
Land	15 000	9 000	7 930 000	0.2 %	0.4 %
Overall Accuracy				99.6 %	
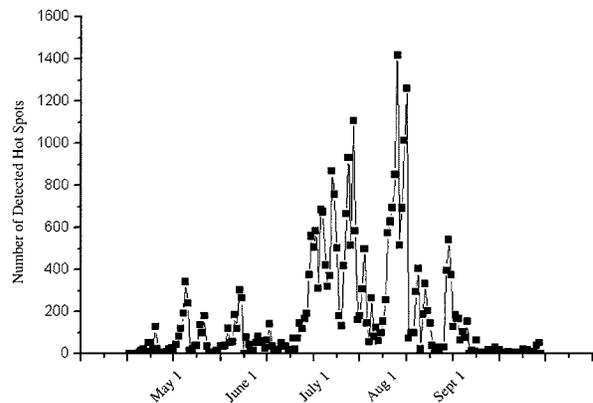
Fig. 8. Number of fire hot spots detected daily across Canada in 1998.

by both AVHRR and TOMS in northern Northwest Territories. TOMS detected a massive smoke plume with AI more than 2.7 over a large area (length $\approx$ 800 km, width $\approx$ 250 km). Presumably, the fires detected by AVHRR in the south are fresh and the associated smoke is light and close to the ground, leaving a very weak signal to be detected by AVHRR. By contrast, the smoke detected in the north by both AVHRR and TOMS is from older fires and has been elevated to higher altitudes. This is in agreement with the atmospheric circulation condition as shown

in Hsu *et al.* [12], [13]. On this day and several days following, a high-pressure ridge resided in western Canada and the U.S., accompanied by prevailing wind blowing northeast. As a result,
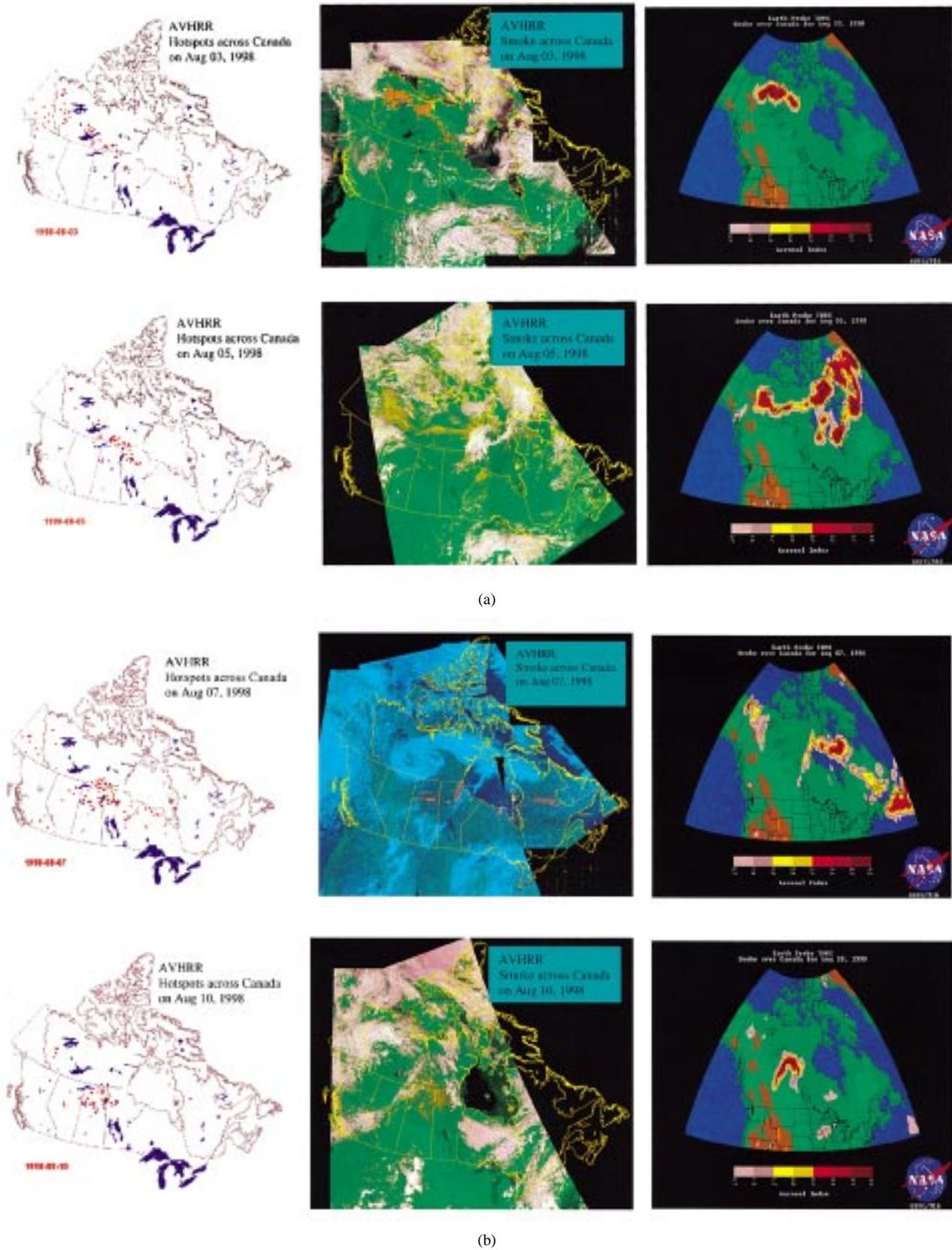
(a)

(b)

Fig. 9. Comparisons of satellite detected fire hot spots (left) from AVHRR, (middle) smoke AVHRR, and (right) aerosol index from TOMS on a series of days in August 1998.

two days later (August 5), smoke extended to eastern Canada according to AVHRR. TOMS detected smoke around Hudson Bay and across the northern Atlantic with higher values of AI. This may be attributable to increasing altitude of smoke as it
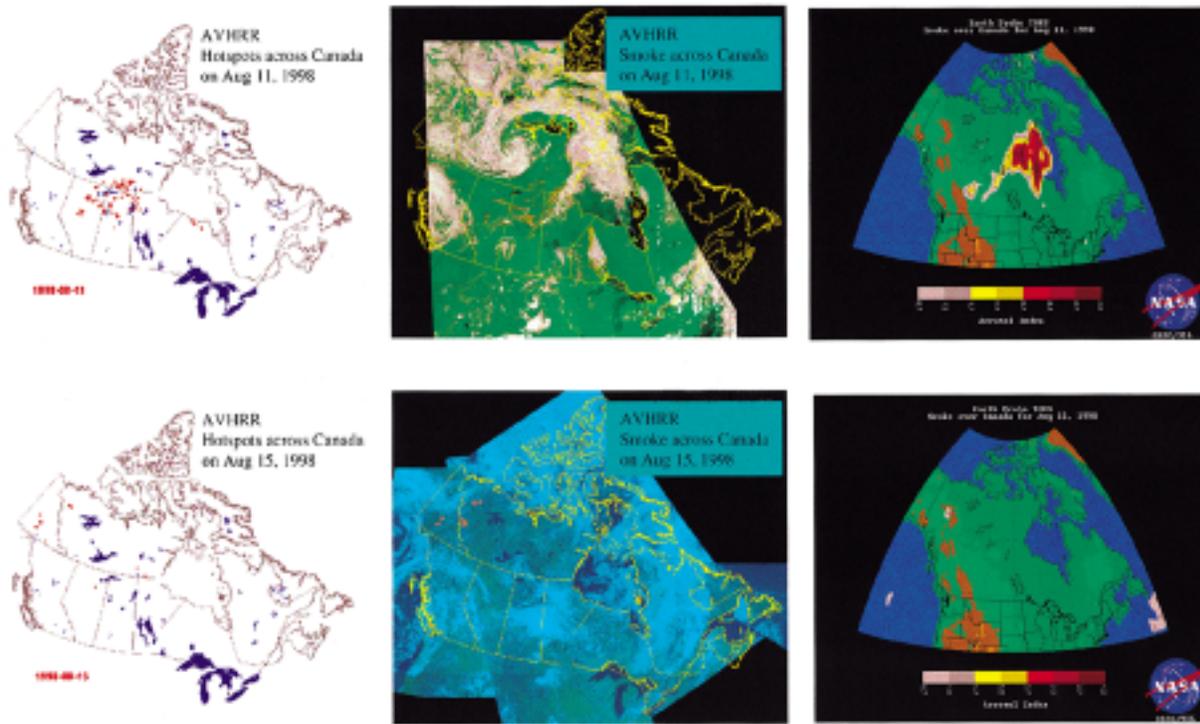
(c)

Fig. 9. (*Continued.*) Comparisons of satellite detected fire hot spots (left) from AVHRR, (middle) smoke AVHRR, and (right) aerosol index from TOMS on a series of days in August 1998.

travels eastward, as implied by the atmospheric circulation condition [12], [13]. However, such a signal is not detected using AVHRR due both to reduced smoke density and to an extensive large cloud cover. On August 7, AVHRR detected several low and dense smoke plumes, whereas TOMS missed almost all of them. On August 10, the smoke is widespread in northern Saskatchewan as revealed by AVHRR-based hot spots and smoke, which is also captured by TOMS. On August 11, the smoke intensified and moved to the northeast. It is interesting to note that TOMS detects smoke with stronger signals downstream, whereas AVHRR is easier to identify smoke near the fire origin.

The above analyses demonstrate the feasibility of smoke detection using both AVHRR and TOMS. More importantly, the two sensors and methods have very different sensitivities and response to smoke located in different layers. It is fortunate that the two methods complement each other. The linkage of smoke detected by the two sensors may be better established using a chemical transport model coupled with an atmospheric circulation model that can trace the movement of smoke species.

V. Summary

Smoke from wildfires is an important source of atmospheric aerosols and chemicals, especially in the boreal forest environment. Smoke aerosol has a significant impact on atmospheric chemistry, weather and climate. To better understand these impacts requires a good knowledge of smoke distribution and temporal variation. On regional or global scales, this may only be achieved using satellite remote sensing techniques. So far,

there have been few studies dedicated to the identification of smoke using satellite imagery data with automatic procedures.

This study developed satellite-based classification algorithms and used them to automatically process large volumes of daily AVHRR imagery data acquired across Canada during several fire seasons. The spectral characteristics of smoke relative to other major scene types (clouds and land background) were first investigated. While both the AVHRR visible and infrared channels convey certain information pertaining to smoke, they are not distinct enough from nonsmoke scenes due to large overlaps of signals, although a combination of the channels may work under certain conditions over a small region. To alleviate these difficulties, both neural networks (NN) and multithreshold approaches were explored. The NN approach has the capability of learning from training data sets and handling complex relationships between various channels in linear or nonlinear forms. Moreover, it provides quantitative and continuous indices of smoke as well as other objects. The smoke index provides a measure of both the concentration of smoke and the mixing with other scene types (i.e., smoke/cloud, smoke/land, and smoke/land/cloud). The main disadvantage of the NN approach is that it is time consuming to process large images like those covering entire Canada (it should be noted that training is most time consuming. Once trained, the network can classify scenes relatively quickly). In addition, when employed over such a large area, misclassification is inevitable. A multithreshold approach was thus also introduced that has certain advantages (quick processing) and disadvantages (categorized output) relative to the NN. Both approaches suffer from limitations inherent in the characteristics of the input

satellite data. They also have greater difficulty in identifying thin dispersed smoke compared to fresh dense smoke. For handling large data volume, an effective approach is to apply the threshold method to identify dense smoke and then to apply the NN to deal with thin smoke. The performance of the algorithms was evaluated in terms of Euclidean distance between the output channels using error matrices and visual inspection and comparisons of classified smoke images with fire hot spots detected using independent algorithms.

The algorithm has been applied to process daily, Canada-wide AVHRR data. The output smoke images for 1998 were assessed with reference to AVHRR-based hot spots, and TOMS-based aerosol index images. Overall, the three types of images show a reasonable correspondence. Both AVHRR and TOMS can detect smoke downwind of fires but have rather different response and capability. AVHRR-based detection is more sensitive to the density of smoke, while TOMS is also affected by the altitude of smoke. AVHRR can detect most smoke near the origin of a fire, which is often missed by TOMS unless the smoke is dense and widespread. However, AVHRR is not effective for monitoring older, highly diffuse smoke distant from a fire source.

REFERENCES

[1] A. K. Blackadar, *Turbulence and Diffusion in the Atmosphere*. Berlin, Germany: Springer-Verlag, 1996, p. 180.

[2] J. A. Benediktsson, P. H. Swain, and O. K. Ersoy, "Neural network approaches versus statistical methods in classification of multisource remote sensing data," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 540–551, May 1990.

[3] C. Clark and A. Canas, "Spectral identification by artificial neural network and genetic algorithm," *Int. J. Remote Sensing*, vol. 16, pp. 2255–2275, 1995.

[4] S. A. Christopher, D. V. Kliche, J. Chou, and R. M. Welch, "First estimates of the radiative forcing of aerosols generated from biomass burning using satellite data," *J. Geophys. Res.*, vol. 101, pp. 21 265–21 273, 1996.

[5] Y. S. Chung and H. V. Le, "Detection of forest-fire smoke plumes by satellite imagery," *Atmos. Environ.*, vol. 18, pp. 2143–2151, 1984.

[6] J. Cihlar, H. Ly, and Q. Xiao, "Land cover classification with AVHRR multichannel composites in northern environments," *Remote Sensing Environ.*, vol. 58, pp. 36–51, 1996.

[7] P. J. Crutzen and M. O. Andreae, "Biomass burning in the tropics: Impact on atmospheric chemistry and biogeochemical cycles," *Science*, vol. 250, pp. 1669–1678, 1990.

[8] R. Fraser, Z. Li, and J. Cihlar, "Hotspot and NDVI differencing synergy (HANDS): A new technique for burned area mapping over boreal forest," *Rem. Sens. Environ.*, vol. 74, pp. 362–375, 2000.

[9] M. Fromm, J. Alfred, K. Hoppel, J. Hornstein, R. Bevilacqua, E. Shettle, R. Servranckx, Z. Li, and B. Stocks, "Observations of boreal forest fire smoke in the stratosphere by POAM III, SAGE II, and lidar in 1998," *Geophys. Res. Lett.*, vol. 27, pp. 1407–1410, 2000.

[10] W. W. Hsieh and B. Tang, "Applying neural network models to prediction and data analysis in meteorology and oceanography," *Bull. Amer. Meteorol. Soc.*, vol. 79, pp. 1855–1870, 1998.

[11] J. R. Herman, P. K. Bhartia, O. Torres, C. Hsu, and C. Seftor, "Global distribution of UV-absorbing aerosols from Nimbus 7/TOMS data," *J. Geophys. Res.*, vol. 102, pp. 16 911–16 922, 1997.

[12] N. C. Hsu, J. R. Herman, J. F. Gleason, O. Torres, and C. J. Seftor, "Satellite detection of smoke aerosols over a snow/ice surface by TOMS," *Geophys. Res. Lett.*, vol. 26, pp. 1165–1168, 1999.

[13] N. C. Hsu, J. R. Herman, O. Torres, B. N. Hoblen, D. Tanre, T. F. Eck, A. Smirnov, B. Chatenet, and F. Lavenu, "Comparison of the TOMS aerosol index with sun-photometer aerosol optical thickness: Results and applications," *J. Geophys. Res.*, vol. 104, no. D6, pp. 6269–6279, 1999.

[14] P. V. Hobbs, J. S. Reid, R. A. Kotchenruther, R. J. Ferek, and R. Weiss, "Direct radiative forcing by smoke from biomass burning," *Science*, vol. 275, pp. 1776–1778, 1997.

[15] J. R. Jensen, *Introductory Digital Image Processing. A Remote Sensing Perspective*. Upper Saddle River, NJ: Prentice-Hall, 1996, p. 315.

[16] E. S. Kasischke, N. L. Christensen, and B. J. Stocks, "Fire, global warming, and the carbon balance of boreal forests," *Ecol. Appl.*, vol. 5, pp. 437–451, 1995.

[17] Y. J. Kaufman, C. J. Tucker, and I. Fung, "Remote sensing of biomass burning in the tropics," *J. Geophys. Res.*, vol. 95, pp. 9927–9939, 1990.

[18] Y. J. Kaufman and T. Nakajima, "Effect of Amazon smoke on cloud microphysics and albedo—Analysis from satellite imagery," *J. Appl. Meteorol.*, vol. 32, pp. 729–744, 1993.

[19] Y. J. Kaufman and R. Fraser, "The effect of smoke particles on clouds and climate forcing," *Science*, vol. 277, pp. 1636–1639, 1997.

[20] D. S. Kimes, R. F. Nelson, M. T. Manry, and A. K. Fung, "Attributes of neural networks for extracting continuous vegetation variables from optical and radar measurements," *Int. J. Remote Sensing*, vol. 19, pp. 2639–2663, 1998.

[21] J. Lawrence, *Introduction to Neural Networks. Design, Theory, and Applications*: California Sci. Software, 1994, p. 347.

[22] Z. Li, J. Cihlar, L. Moreau, F. Huang, and B. Lee, "Monitoring fire activities in the boreal ecosystem," *J. Geophys. Res.*, vol. 102, pp. 29 611–29 624, 1997.

[23] Z. Li, "Influence of absorbing aerosols on the inference of solar surface radiation budget and cloud absorption," *J. Climate*, vol. 11, pp. 5–17, 1998.

[24] Z. Li and L. Kou, "Atmospheric direct radiative forcing by smoke aerosols determined from satellite and surface measurements," *Tellus B*, vol. 50, pp. 543–554, 1998.

[25] Z. Li, S. Nadon, and J. Cihlar, "Satellite detection of Canadian boreal forest fires: Development and application of an algorithm," *Int. J. Remote Sensing*, vol. 21, pp. 3057–3069, 2000.

[26] Z. Li, S. Nadon, J. Cihlar, and B. J. Stocks, "Satellite mapping of Canadian boreal forest fires: Evaluation and comparison," *Int. J. Remote Sensing*, vol. 21, pp. 3071–3082, 2000.

[27] J. D. Paola and R. A. Schowengerdt, "A review and analysis of back-propagation neural networks for classification of remotely sensed multi-spectral imagery," *Int. J. Remote Sensing*, vol. 16, pp. 3033–3058, 1995.

[28] J. E. Penner, R. E. Dickinson, and C. A. O'Neill, "Effects from aerosol from biomass burning on the global radiation budget," *Science*, vol. 256, pp. 1432–1433, 1992.

[29] J. C. Principe, N. R. Euliano, and W. C. Lefebvre, *Neural and Adaptive Systems*. New York: Wiley, 1999, p. 656.

[30] E. M. Prins and W. P. Menzel, "Trends in South American biomass burning detected with the GOES visible infrared spin scan radiometer atmospheric sounder from 1983 to 1991," *J. Geophys. Res.*, pp. 16 719–16 735, 1994.

[31] L. F. Radke, "Airborne observation of cloud microphysics modified by anthropogenic forcing," in *Proc. Symp. Atmospheric Chemistry and Global Climate*, Anaheim, CA, Jan. 29–Feb. 3 1989.

[32] T. Randriambelo, S. Baldy, M. Bessafi, M. Petit, and M. Despinoy, "An improved detection and characterization of active fires and smoke plumes in south-eastern Africa and Madagascar," *Int. J. Remote Sensing*, vol. 19, pp. 2623–2638, 1998.

[33] C. R. N. Rao and J. Chen, "Post-launch calibration of the visible and near-infrared channels of the advanced very high resolution radiometer on NOAA-14 spacecraft," *Int. J. Remote Sensing*, vol. 17, pp. 2743–2747, 1996.

[34] J. A. Richards, *Remote Sensing Digital Image Analysis*. Berlin, Germany: Springer-Verlag, 1986, p. 281.

[35] A. Robock, "Surface cooling effect due to forest fire smoke," *J. Geophy. Res.*, vol. 96, pp. 20 869–20 878, 1991.

[36] B. Robertson, A. Erickson, J. Friedel, B. Guindon, T. Fisher, R. Brown, P. Teillet, M. D'Iorio, J. Cihlar, and A. Sancz, "GEOCOMP, a NOAA AVHRR geocoding and compositing system," in *Proc. Int. Soc. Photogrammetry and Remote Sensing Conf.*, Washington, DC, 1992, pp. 223–228.

[37] B. J. Stocks, D. R. Cahoon, J. S. Levine, W. R. Cofer, III, and T. J. Lynham, *Fire in Ecosystems of Boreal Eurasia*, J. G. Goldammer and V. V. Furyaev, Eds. Norwood, MA: Kluwer, 1996, pp. 139–150.

[38] J. Wong and Z. Li, "Retrieval of optical depth for heavy smoke aerosol plumes: Uncertainties and sensitivities to the optical properties," J. Atmos. Sci., 2000, submitted for publication.

[39] G. Wotawa and M. Trainer, "The influence of Canadian forest fires on pollutant concentrations in the United States," *Science*, vol. 288, pp. 324–328, 2000.

Zhanqing Li received the B.Sc. and M.Sc. degrees from Nanjing Institute of Meteorology, Nanjing, China, in 1983 and 1985, respectively, and the Ph.D. degree from McGill University, Montreal, QC, Canada, in 1991.

Currently, he is a Professor at the University of Maryland, College Park, after eight years with the Canada Centre for Remote Sensing (CCRS), Ottawa, ON, as a Research Scientist and Group Leader. He has engaged in a variety of studies concerning cloud, radiation budget, aerosol, UV and ozone, photosynthetically active radiation, BRDF, terrestrial environment, and forest fires. He serves as a Principal Investigator and Science Team Member in numerous Canadian, U.S., and international programs such as the FIRE/M3, ARM, GACP, GOFC, EOS, CloudSat, and LCLUC. He has authored over 70 peer-reviewed journal articles.

Dr. Li has received seven major awards, including four national ones.

Alexandre Khananian received the M.S. degree in nuclear physics from the University of Tashkent, Uzbekistan, USSR, in 1969, and the Ph.D. degree in geophysics from the Institute of Experimental Meteorology, Obninsk, in 1974.

He was a Senior Research Scientist at the Institute of Experimental Meteorology from 1976 to 1994, and Professor of the Obninsk University of Nuclear and Power Engineering, Obninsk, from 1994 to 1997. From 1998 to the present, he has been at the Canada Centre for Remote Sensing (CCRS), Ottawa, ON, as a Research Associate. His research interests include many aspects of atmospheric physics, environmental monitoring, and remote sensing data analysis.

Robert H. Fraser received the B.S. in biology from the University of Ottawa, Ottawa, ON, Canada, in 1993, and the M.S. and Ph.D. degrees from the School of Forestry and Environmental Studies, Yale University, New Haven, CT, both in 1995.

He is an applications Research Scientist with the Environmental Monitoring Section, Canada Centre for Remote Sensing (CCRS), Ottawa, ON, Canada. Since joining CCRS in 1998, his research has focused on satellite remote sensing of boreal forest fires. He is a member of the Natural Resources Canada Team that developed the Fire Monitoring, Mapping, and Modeling (Fire M3) System, which came online in May 1999. Fire M3 uses NOAA/AVHRR imagery and Internet mapping software to provide daily interactive maps of forest fires across Canada.

Dr. Fraser received a NASA Graduate Fellowship in Earth systems science.

Josef Cihlar received the B.Sc. degree from the University of Agriculture, Prague, Czechoslovakia, in 1967, the M.Sc. degree from the University of Guelph, Guelph, ON, Canada, in 1971, and the Ph.D. degree from the University of Kansas, Lawrence, in 1975.

He is currently a Senior Research Scientist and Head of the Environment Monitoring Section, Canada Centre for Remote Sensing (CCRS), Ottawa, ON, Canada. His research interests include biosphere dynamics, especially boreal ecosystems, and the use of satellite-based methods to provide measurements for mechanistic models; corrections of satellite data to allow derivation of biophysical variables such as leaf area index, active fires and burned areas, land cover, photosynthetically active radiation; the application of satellite data and associated process models to the biosphere of Canada and circumpolar boreal forest; and the design and establishment of global observing systems to detect, monitor, and assess the impact of climate on the terrestrial biosphere, including associated feedbacks. He has authored over 100 papers on these subjects.

Dr. Cihlar is a Fellow of the Royal Society of Canada.