Computer Identification of Multispectral Satellite Cloud Imagery

Li Jun (李 俊) and *Zhou Fengxian* (周凤仙)
Institute of Atmospheric Physics, Academia Sinica, Beijing 100011

ABSTRACT

A dynamic clustering method based on multispectral satellite imagery to identify the different features is described. The channel combinations selected are for the different purposes in classification. Several cases are presented using the polar-orbiting satellite imageries.

I. INTRODUCTION

Mapping of cloud and weather system is a fundamental task in the operational weather forecasting. The satellite data are the important source in monitoring and predicting the weather system. Since the satellite image contains huge amount of data, the processing must be done by computer to extract the interesting information for the forecasters.

The computer classification of cloud imagery consists of two steps. The first step is cloud detection which determines the percentage of cloud cover. The second step is analysis which determines the properties of clouds found in the first step. These methods can be separated into two categories:the threshold method in which the cloud cover detection is performed on the individual pixels using VIS, IR, NIR radiances and physical thresholds, and in the statistical method or clustering method(Seze et al., 1987). The cloud detection and analysis are performed globally on image segments using both spectral and spatial properties of the VIS, IR and NIR radiances. Liljas(1982; 1986) uses the threshold method. But the threshold values vary with the time and season, while the major problem of the clustering method is the choice of the initial clusters which has an influence on final classification. But the clustering method has more potential than the threshold one because the later uses only spectral information of the data, whereas the clustering method uses both spectral and spatial properties of the data. In order to solve the problem which clustering method encounters, a combination of threshold method and clustering method is presented in this paper. The initial clusters are selected by the threshold method, then the clustering technique is performed on the image classification. On the other hand, the threshold method also can be used after the imageries are classified by the clustering method.

II. MAIN FEATURES OF AVHRR

The spectral features of AVHRR are summarized on Table 1 together with the objectives assigned to each channel. The AVHRR instrument provides images with very high spatial resolution (1.1 km on foot track) and the radiance value is represented by 10-bit digits. For simplicity, all the analyses here are done directly on AVHRR counts(here 8-bit is used) instead of on brightness temperatures and albedos. Before applying any processing to those

data, we assume that the solar zenith angle should not be changed too much through the analysis area.

Table 1. AVHRR Channel Characteristics (NOAA-9)

Channel	Wavelength (μm)	Objectives
1	0.58–0.68	Recognition of clouds
2	0.725–1.10	Delineation of land, water, melting & non-melting snow
3	3.55–3.93	SST and atmospheric correction in partly cloudy areas
4	10.3–11.3	SST and atmospheric correction
5	11.5–12.5	Same as channel 4

III. CHOICE OF CHARACTERISTICS

1. *Variance Image*

A variance image was constructed for each image. In the derived image, the value attributed to each pixel is the Local Standard Deviation (LSD) from 3 × 3 of the pixels.

If $f(i,j)$ is the original image, for i = 1,2.....N, j = 1,2.....N, then its variance image or LSD image is

$$g(i,j)=\frac{1}{9}\sum_{k=i-1}^{i+1}\sum_{l=j-1}^{j+1}[f(k,l)-\bar{f}(i,j)]^2\}^{1/2}$$

where

$$\bar{f}(i,j)=\frac{1}{9}\sum_{k=i-1}^{i+1}\sum_{l=j-1}^{j+1}f(k,l)$$

The variance image (or LSD image) plays an important role in cloud identification. LSD image also shows the characteristics of the original image.

2. *Choice of Characteristics*

Many clouds are easily identified on the IR image because their temperature is lower than the sea surface temperature and the ground temperature. The light cirrus cloud, which is almost invisible in VIS and NIR, is easy to be seen in IR channel. Conversely, some low cloud which temperature closes to the sea surface temperature (SST) is best identified by VIS and NIR channels. So VIS or NIR and IR characteristics can be selected in classification. But some pixels may not be attributed to a specific category only from its VIS / IR spectral signature, because their spatial or textural characteristics in the corresponding VIS and / or IR image are different. For instance, the adjacent pixels covered by low or middle clouds may have very different VIS values whereas their IR values are similar. The behavior of thin cirrus is reversed. These textural properties can be quantified by LSD images. The general interpretation of VIS and IR LSD features can be summarized as follow:

a) Low VIS variance, low IR variance: Surface(sea, ground), Homogeneous thick clouds

b) Relatively low VIS variance, high IR variance: Cirrus

c) Relatively high VIS variance, low IR variance: Middle–low clouds

d) High VIS variance and high IR variance: Thick high and middle clouds

According to this interpretation, the objects which could not be separated on VIS and IR images can be well identified by using VIS and IR variance images. So channel 1, channel 4 and their LSD images are selected to classify cloud types and surfaces. On the other hand, NIR can be chosen for water–land separation.

Channels 1,2,4 and their LSD images reveal the main cloud types and surface objects during summer with a sun elevation angle over twenty degrees(Liljas, 1982; Liljas, 1984). But these characteristics are sometimes not able to classify imagery at low sun elevation angles in winter. They are failed during the night for obvious reasons. Other limitations are sun–glints in morning and evening passages during spring and fall. The sun–glints are classified as stratus, fog or mist. Cirrus clouds over lower clouds are sometimes classified as nimbostratus, which is highly dangerous in automatic precipitation analysis. Snow has signatures which are close to that of different clouds. The imagery of channel 3 (3.7 μm) is complex during the day since it includes both reflected solar radiation and emitted terrestrial radiation.

The radiative properties of water droplet and ice clouds are related by equation (Hunt, 1973)

$$\varepsilon + re + tr = 1 \tag{1}$$

where ε,re and tr are the emissivity, reflectivity and transmissivity of the cloud layer, respectively. The following facts are to be considered:

a) An increase of the particle sizes has the effect of increasing the transmissivity and emissivity and decreasing the reflectivity of the cloud layer;

b) At infrared windows water droplet and ice clouds do not behave like blackbody radiators;

c) The emissivities of the ice clouds are less than one of the water droplet clouds as a result of their lower water content;

d) At 8.5–13 m window there is a large contribution to the emergent field of radiation reflected by the cloud top at the shorter wavelength that is absent for the 8.5–13 m radiance;

e) In the 8.5–13 m window the emergent radiance is dominated by the emitted radiation from the cloud layer.

For optically thick clouds the transmissivity vanishes and equation (1) is simplified to

$$re + \varepsilon = 1 \tag{2}$$

Land, water and snow surfaces have reflectivities close to 0 in the window around 3.7 μm.

To simplify the interpretation of the 3.7 μm imagery Liljas (1986) subtracts radiance of long–wave window channel from short–wave window channel. It means that the new image is constructed by subtracting channel 4 from channel 3. This new image is defined as channel 3 / 4 image. In the new image the dominant features result from reflected solar radiance at 3.7 μm.

Figure 1 shows a result of the described groupings of different reflected surfaces in the two–dimensional intensity space of channel 3 / 4 and channel 1. The bands are normalized by dividing by sin (α), where α is the sun elevation angle.

So channel 1, channel 3 / 4 images can be chosen for distinguishing sun glints, snow, and

cloud types in winter. Simulation results show that channel 1 variance image plays an important role in winter classification when it is regarded as one of the characteristics.

Detection of clouds at night using satellite images is possible mainly because clouds are relatively cold and appear as light grey or white color. The main problems at night are that the thermal contrast between the fog or stratus top and the surface is usually very small and that cirrus and cirrostratus often have the same appearance as deep clouds. This can be solved by using channel 3 / 4 and its LSD images, because surface appears smooth texture characteristic. So channel 3 / 4, channel 4 and their LSD images can be selected for nighttime cloud classification.

The different characteristics can be chosen according to different purposes in classification. Generally to identify the land, water, and clouds, we should choose channel 1, channel 2 and their LSD images. If we combine channel 1, channel 4 and their LSD images, then all cloud types will be well identified. Use of the radiance difference between channel 3 and channel 4 provides additional information. The combination of channel 3 / 4, channel 1 and channel 1 variance images can effectively identify the sun glints, snow, surface and cloud types in winter. At night time channel 3 / 4, channel 4 and their LSD images are selected for identifying stratus or fog, stratocumulus, altocumulus, cumulonimbus, cirrus etc.

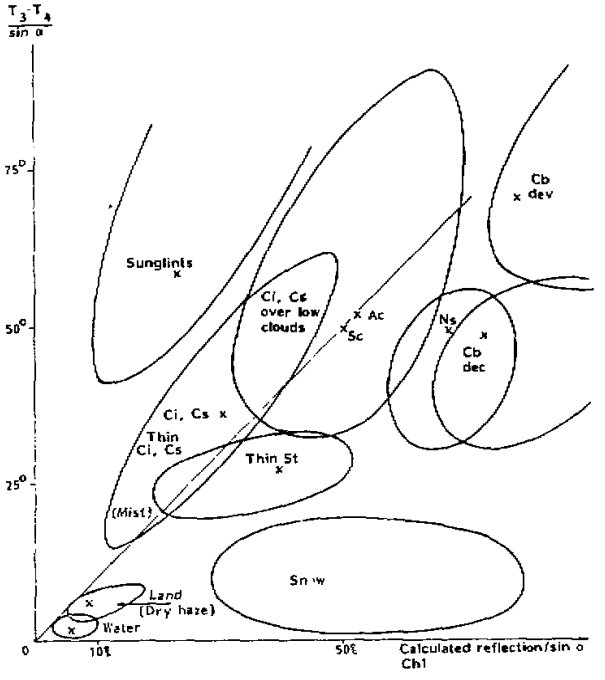

Fig.1. Multispectral cloud classification using polar orbiting NOAA-Series satellite imagery (Liljas, 1984). Reflection of channel 3 – channel 4 is displayed in ordinate and reflection from channel 1 in abscissa;the values are normalized dividingby sin (α) for overhead sun.

IV. DYNAMIC CLUSTERING TECHNIQUE

We assume that the multispectral imagery (including their LSD images) data are normally distributed. The explicit expression of $p(x)$ is

$$p(x) = (2\pi)^{\frac{-k}{2}} \left|\sum_x\right|^{\frac{1}{2}} \exp[-\frac{1}{2}(x-M)^T \sum_x^{-1}(x-M)] = N(x,M,\sum_x)$$

Where $M = E(x)$ is the mean vector and $\sum_x$ is the covariance matrix of x. Its k–dimensional multispectral images (including their LSD images) with N × N pixels are represented by the matrix X

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1N} & x_{1N+1} & \cdots & x_{1N^2} \\ x_{21} & x_{22} & \cdots & x_{2N} & x_{2N+1} & \cdots & x_{2N^2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{k1} & x_{k2} & \cdots & x_{k_N} & x_{k\,N+1} & \cdots & x_{k\,N^2} \end{bmatrix}$$

$k \leqslant 4$ is used in this paper.

Clustering is commonly used in classification. It does not use any knowledge of the feature vector statistics. It utilizes only the "inherent" features of the multispectral image data. If the data set consists of m–subsets, each of them corresponds to a class of object. The characteristics of the feature vectors within subset are more similar than those of inter–subsets. Clustering partitions the feature vectors into subsets or clusters according to the measurement of similarities. The distance between the feature vector and the cluster mean vector is the major measurement for the partition of the feature vectors. The clustering can be summarized as the flow chart shown in Fig 2.

The clustering algorithm can be described with the following steps:π

(1) Specify the number of clusters needed.

(2) Use of box–classification to select m initial cluster mean vectors or m initial clusters M1, M2,...,Mm.

(3) Calculate the "distance" of a feature vector to each cluster mean vector and assign the feature vector to the nearest cluster.

(4) Update the cluster mean vectors after all feature vectors have been assigned to the corresponding clusters.

(5) Calculate the distance between the updated cluster mean vectors M'_j, for j = 1,2,...,m. If the average distance between previous and updated mean vector exceeds some given threshold, go back to step (3), otherwise go to step (6) or if the maximum number of iterations is reached, also go to step (6).

(6) Analyze the separability of the clusters produced, reject some feature vectors in clusters, and merge or split some clusters if necessary.

1. *Selection of Initial Cluster Mean Vectors*

The clustering algorithm is related to the choice of initial cluster mean vectors. More iterations may be required for convergence if the initial cluster centers are too far from the final cluster mean vectors, and the choice of the initial cluster has an influence on the final partition. In order to solve this problem, we use a multispectral box–classification (Liljas 1982; 1985) to do the selection of initial clusters, then calculate the cluster mean vectors according

to m clusters produced.

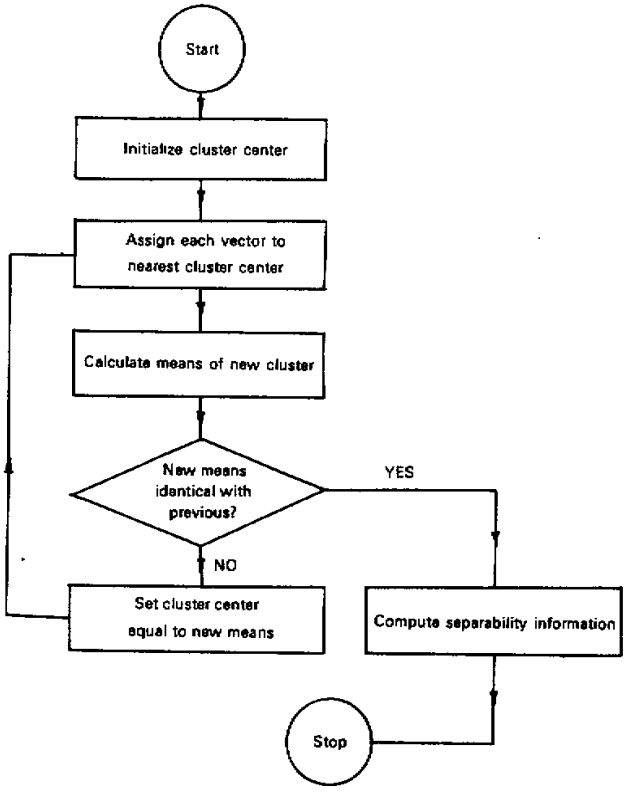

Fig. 2. The clustering algorithm

According to box–classification method, different cloud type, land and water have different–spectral signatures, these signatures define parallelepipeds in a three–dimensional intensity space, these boxes for different clouds do not overlap. In most cases all pixels in the image are classified. VIS and IR separate different cloud types; NIR is used for land–sea separation;channel 1 and channel 3 / 4 separate sunglints, snow, and cloud types;channel 3 / 4 and channel 4 are used for night classification. Fig. 3 is an example of using box–classification in winter. (Figs. 3, 4, 5 and 6 are displayed as false colour images.)

2. *The Distance Measurement*

In clustering, the distance between the feature vector and the cluster mean vector is the most important measure. The feature vector will be assigned to the cluster that is the nearest. There are several different definitions of distance (as defined in multi–dimensional space) existed. The most commonly used distances are:

(1) *Euclidean distance*

$$D_e(x,M_i)=[(x-M_i)^T(x-M_i)]^{1/2}=\left[\sum_{j=1}^{k}(x_j-m_{ji})^2\right]^{1/2}$$

Where x_j and m_{ji} are the j–th components of the feature vectors x and the cluster mean vectors M_i for j = 1,2,...,k

(2) *Weighted euclidean distance*

$$D_{we}(x,M_i)=\left[\sum_{j=1}^{k}w_j(x_j-m_{ji})^2\right]^{1/2}$$

where w_j is the weighting factor of the j–th component, the contribution of each component to the distance is weighted.

(3) *Mahalanobis distance*

$$D_m(x,M_i)=[(x-M_i)^T\sum_x^{-1}(x-M_i)]^{1/2}$$

where $\sum_x$ is the covariance matrix of x. Usually we choose Euclidean distance or Mahalanobis distance in clustering.

3. *Analyzing the Clusters*

After the m clusters are finally produced, they are then analyzed to determine whether the clusters are actually separated or lie too close together in the feature space that they should be merged. Some clusters are so large that they should be split.

The cluster should be split into two clusters if one or a few components of the standard deviation exceeds a certain threshold. A more useful strategy is adopted. In the first step of the clustering procedure, the number of clusters is specified. This number should be larger than required. After termination of the clustering iteration, superfluous clusters are merged. This may be done by computing the intercluster distances and merging those clusters which are separated by less than a prespecified threshold. The threshold for merging must be set small enough so that the resulting composite clusters will not have multi–modal probability functions.

V. SOME RESULTS OF CLUSTERING FOR NOAA–9 AVHRR IMAGE DATA

Fig.4 shows the image produced from channel 1, channel 4 and their LSD images combination.

Four main classes have been found, whose centers of gravity parameter values are given in Table 2.

The identifications given to the classes are based on the following measurements:

1) Values of the different parameters for the centers of gravity.

2) Projection of each center on Liljas's "BOX".

3) Comparison of the classified images to the corresponding original images.

So we have obtained:

class 1 has the very high LSDVIS and is more homogeneous in IR, and also relatively dark, is the class of low level clouds.

* class 2 are characterized by high LSDVIS and high LSDVIS and very bright, which denotes the presence of middle–low clouds.

* class 3 :the most homogeneous in VIS and IR, and also the darkest, is characteristic of the surface.

* class 4:relatively homogeneous in VIS channel, but has very high LSDIR, corresponds to cirrus.

Table 2. Values of the Centers of Gravities in Counts of the Four Classes Obtained with the Four Parameters (VIS, IR, LSDVIS, LSDIR) Classification

C	VIS	IR	LSDVIS	LSDIR	Percentage	Object
1	23	81	126	55	.215	Low level clouds
2	32	88	204	139	.090	Middle level clouds
3	17	78	29	16	.632	Surface(*)
4	36	91	59	171	.063	Cirrus

* To use channel 1 and channel 2 by threshold method, surface is divided into two clusters:water (sea, lake) and land.

Figure 5 covers Southeast of Japan and its vicinity. The image is produced from channel 1, channel 2 and their LSD images.

Five classes have been obtained, whose centers of gravity parameter values are given in Table 3.

Table 3. As in Table 2 but for Four Parameters (VIS, NIR, LSDVIS, LSDNIR)

C	VIS	NIR	LSDVIS	LSDNIR	Percentage	Object
1	6	4	0	1	.532	Sea
2	11	23	2	7	.453	Land
3	18	25	12	133	.003	Cirrus
4	32	39	186	178	.004	Low level clouds
5	22	29	74	78	.008	Middle or low clouds

Fig.6 shows the result of clustering classification use three parameters: channel 1, channel 3 / 4 and LSD channel 1. It is the same region and the same time of Fig 2. In this step, snow is well identified.

Four main classes are obtained, whose centers of gravity parameter values are given in Table 4.

Table 4. As in Table 2

C	VIS	CHANNEL 3 / 4	LSDVIS	Percentage	Object
1	25	22	11	.414	Land
2	40	13	6	.486	Snow
3	40	77	19	.075	Developing cumulonimbus
4	33	24	78	.024	Middle low clouds

VI. CONCLUSION

The results of classification fit the appearance of true color–composite image. The application of the clustering technique considering both radiances and local variance confirms the

usefulness of these parameters for cloud, snow / surface separation, as well as for the separation of the clouds with different types and levels. The results obtained are preliminary. More experiments are being done. Verification with ground truth and real weather observation will be carefully checked out. The data used in the paper are just for testing the algorithm. The meaningful cases are being prepared. Since the precipitable cloud, for instance the deep–convective cloud, can be well identified by the clustering method, the application of the technique in nowcasting keeps promising.

The authors would like to particularly acknowledge the AVHRR data provided by Remote Sensing Division, National Research Center for Marine Environmental Forecasts. We are also grateful to Professor Zhang Xuding at the Information Laboratory, Department of Mathematics, Peking University for offering FORTRAN program of Dynamic Clustering.

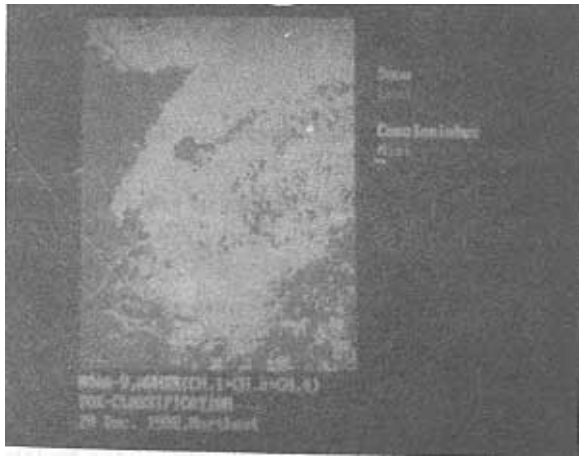

Fig 3. Box–classification over Northeast China

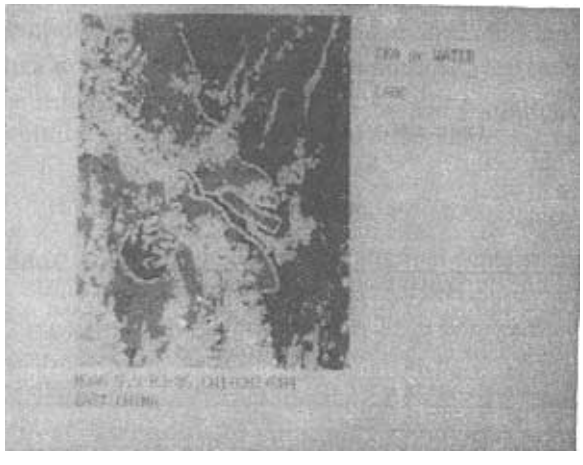

Fig 4. Clustering classification over East China using four parameters.

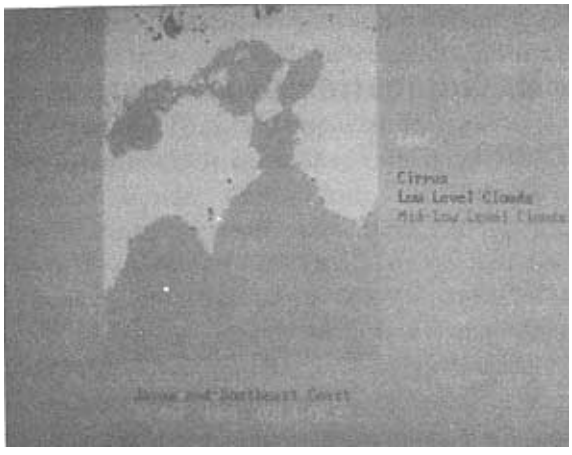

Fig 5. Clustering classification over Southeast of Japan using four parameters.

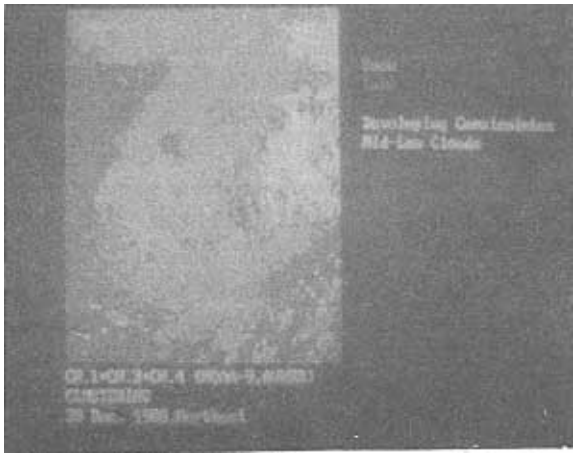

Fig 6. AS in figure 4 but using clustering method.

REFERENCES

Coakley, J. A. and D. G. Baldwin (1984), Towards the Objective Analysis of Clouds from Satellite Imagery Data *J. Climate Appl. Meteor* **23:** 1065–1099.

Hunt, G. E. (1973), Radiative properties of terrestrial clouds at visible and infrared thermal window wavelengths, *Q. J. R. Meteorol. Soc.*, **99:** 346–369.

Kitting, R. L. and D. A. Langrebe, Classification of Multispectral Image Data by Extraction and Classification of Homogeneous Objects, Trans.GE–14, 10.

Liljas, E. (1982), Automated Techniques for the Analysis of Satellite Cloud Imagery. *Nowcasting*, K. A. Browning, Academic Press, 177–190.

Liljas, E. (1986), Use of the AVHRR 3.7 Micrometer Channel in Multispectral Cloud Classification, SMHI.

Phulpin, T., M. Derrien and A. Brard (1983), A Two–Dimensional Histogram Procedure to Analyze Cloud Cover from NOAA Satellite High–Resolution Imagery, *J. Climate Appl. Meteor.*, **22:** 1332–1345.

Seze, G. and M. Desbois (1987), Cloud Cover Analysis from Satellite Imagery Using Spatial and Temporal Characteristic of the Data, *J. Climate & Appl. Meteor.*, **26:** 287–303.

Tsonis, A. A. (1984), On the separability of various classes from GOES visible and infrared data, *J. Climate Appl. Meteor.*, **23:** 1393–1410.