041973

STATISTICAL SEPARABILITY OF SPECTRAL CLASSES OF BLIGHTED CORN

R. Kumar and L. Silva
Laboratory for Applications of Remote Sensing
Purdue University, 1220 Potter Drive
W. Lafayette, Indiana 47906

ABSTRACT

The purpose of this study was to determine the statistical separability of multispectral measurements from corn having varying levels of southern corn leaf blight severity. Multispectral scanner data in twelve spectral channels in the wavelength range 0.4 to 11.7 μm were analyzed for ten selected flightlines of the 1971 Corn Blight Watch Experiment. A total of 168 corn fields having 18,804 sample points were analyzed. The blight rating information for these fields were available.

Maximum average transformed divergence between spectral classes (found by LARSYS Cluster Algorithm) of <u>all possible</u> pair of blight levels, maximized over a subset of channels, was computed in each of one, two, three, and four spectral channels for each of ten flightlines. From the statistical analysis of the values of average transformed divergence, it was concluded that the greater the difference between the blight levels, the more statistically separable they are. This result is encouraging considering the fact that there are variables other than the blight severity within and between flightlines. Although the analysis was done for corn blight only, the conclusions obtained

from this analysis may well be applicable to other crop stresses because corn blight is representative of the problems of many plant stresses, especially non-systemic stresses.

The work reported in this paper was sponsored by the National Aeronautics and Space Administration (NASA) under Grant No. NGL 15-005-112.

I. Introduction

The purpose of this study was to determine statistical separability of multispectral measurements from corn having varying levels of southern corn leaf blight severity. The data were analyzed in one, two, three and four spectral channels for ten selected flightlines in the 1971 Corn Blight Watch Experiment (CBWE71). Multispectral scanner data in twelve spectral channels in the range 0.4 to 11.7 μm, collected with an optical-mechanical scanner at altitudes of 3000 to 7000 feet were analyzed by applying automatic pattern recognition techniques. The authors consulted the key persons involved in the CBWE71 and selected with their help ten flightlines for analyzing the data.

Southern corn leaf blight (SCLB) is caused by the fungus Helminthosporium maydis. The disease has been known for many years and is wide spread in corn-growing tropical areas of the world (MacDonald et. al., 1972). Symptoms of SCLB are the appearance of brown lesions on the lower leaves; the lesions grow in size and spread to upper leaves until the entire plant is prematurely killed (Figure 1). Only those corn fields were analyzed for which blight rating information (Ullstrup, 1961) from blight level 0 (healthy corn) to blight level 5 (very severe blight) based on the amount of leaf damage were available (Figure 2). A total of 168 fields having 18804 sample points taken from ten flightlines were analyzed.

II. Method of Analysis

Black and white photography and gray scale printouts of spectral channels of the flightlines were used to aid in locating the boundaries of the corn fields on the LARS Digital Display.* The LARSYS Cluster Algorithm (Swain, 1972) was then used to find the spectral distinct classes in six spectral channels. There could be more than one spectral class in one blight level but no more than one blight level was ever put in the same spectral class. A key assumption made in the cluster algorithm, statistics algorithm and feature selection algorithm is that the distributions of the classes are Gaussian. Histograms of the spectral classes defined above were used to check unimodality of the statistical distributions in individual channels. The spectral classes were redefined to eliminate distinct multiple modes. This analysis was done by various analysts at the Laboratory for Applications of Remote Sensing in 1971. The authors used the spectral classes defined by the analysts and ckecked the class unimodality.

For a pair of Gaussian distribution patterns, the divergence in n channels C_1, C_2 ... C_n, for the case of normal variables with unequal covariance matrices is given (Marill and

---

* The LARS Digital Display is a software system linked to an IBM 360/Model 67 using a cathode ray tube as the pictorial medium for gray scale multispectral imagery.

Green, 1963) by

Σ
Sigma

$$D(i,j \mid C_1, C_2 \ldots C_n) = 1/2 \,\mathrm{tr}\,[(\Sigma_i - \Sigma_j)(\Sigma_j^{-1} - \Sigma_i^{-1})]$$

$$+ 1/2 \,\mathrm{tr}\,[(\Sigma_i^{-1} + \Sigma_j^{-1})(U_i - U_j)(U_i - U_j)^T] \qquad (1)$$

where

U and Σ represent the mean vector and covariance matrix, respectively;

tr[A] (trace A) is the sum of the diagonal elements of A.

A modified form of the divergence D_T, referred to as the "transformed divergence," has a behavior more like probability of correct classification than D (Figure 3).

$$D_T = 2[1 - \exp(-D/8)] \qquad (2)$$

Transformed divergence has been used throughout this study.

Let i and j denote the blight levels i and j. Let $W_{i_k j_\ell}$ denote the weight between k^{th} spectral class of blight level i and ℓth spectral class of blight level j, for computing the weighted average of transformed divergence over the pairs of spectral classes. Throughout this analysis, $W_{i_k j_\ell}$ was taken either = 0 or = 1. The weights which were taken = 0 and the weights which were taken = 1 shall be specified in each individual case.

Let $$D^n_{TAVG} = \frac{1}{N} \sum_{i=0}^{s} \sum_{j=i}^{s} \sum_{k=1}^{n_i} \sum_{\ell=1}^{n_j} W_{i_k j_\ell} D^n_T(i_k, j_\ell \mid C_1, C_2 \ldots C_n) \qquad (3)$$

where

n_i = no. of spectral classes in blight level i

n_j = no. of spectral classes in blight level j

N = total no. of spectral class pairs whose weights $W_{i_k j_\ell}$ were taken = 1.

$D^n_T (i_k, j_\ell \mid C_1, C_2 \ldots C_n)$ = transformed divergence between k^{th} spectral class of blight level i and ℓ^{th} spectral class of blight level j in n spectral channels - C_1, $C_2 \ldots C_n$.

Let D^n_{TMAVG} denote maximum of D^n_{TAVG}, maximized over a set of n spectral channels

Throughout this analysis, $W_{i_k j_\ell} = 0$ when i=j and k=ℓ.

III. Maximum Average Statistical Separability Between the Spectral Class Pairs of Mild, Intermediate and Severe Blight

As pointed out in the Introduction, a total of 168 corn fields having 18804 sample points taken from ten flightlines were analyzed. The corn blight severity was divided into three groups - mild blight (blight levels 0 and 1), intermediate blight (blight levels 2 and 3) and severe blight (blight levels 4 and 5).

The subset of each of one, two, three and four spectral channels which maximized the average transformed divergence given by Eq. (3) was selected and the maximum average transformed divergence (maximized over a subset of channels) was tabulated. Let $\{D^n_{TMAVG}\}$ denote a vector whose p^{th} component (where p = 1, 2 ... 10) represents D^{n}_{TMAVG} in p^{th} flightline. Each of $\{D^{1}_{TMAVG}\}$, $\{D^{2}_{TMAVG}\}$, $\{D^{3}_{TMAVG}\}$ and $\{D^{4}_{TMAVG}\}$ was computed using Eq. (3) for the Sections III (a) to III (d). The values of $W_{i_k j_\ell}$ given below for Sections III (a) to III (d) hold true for all values of k and ℓ.

(a) Maximum Average Statistical Separability Between all Spectral Class Pairs

Spectral classes were found with the help of the LARSYS Cluster Processor (Sec. II). Between all spectral class pairs, weights are set equal to one, i. e., $W_{i_k j_\ell} = 1$ for all i, j, k and ℓ. In all Sections III (b) to III (d) given below, $W_{i_k j_\ell} = 0$ for i=j, and the weights whose values are not specified are taken = 1.

(b) Maximum Average Statistical Separability Between All Spectral Class Pairs of Mild Blight (Blight Levels 0 and 1), Intermediate Blight (Blight Levels 2 and 3) and Very Severe Blight (Blight Levels 4 and 5).

Between all spectral class pairs of blight levels 0 and 1, blight levels 2 and 3, blight levels 4 and 5, weights are set equal to zero, i. e., $W_{0_k 1_\ell} = 0$, $W_{2_k 3_\ell} = 0$, $W_{4_k 5_\ell} = 0$.

(c) Maximum Average Statistical Separability Between the Spectral Class Pairs of Moderate (Blight Levels 0, 1 and 2) and Severe (Blight Levels 3, 4 and 5) Blight.

The corn blight severity was divided into two groups -- moderate blight (blight levels 0, 1 and 2) and severe blight (blight levels 3, 4 and 5). Between all spectral class pairs of blight levels 0, 1 and 2, weights are set equal to zero. Similarly, between all spectral class pairs of blight levels 3, 4 and 5, weights are set equal to zero, i.e., $W_{0_k 1_\ell} = 0$, $W_{0_k 2_\ell} = 0$, $W_{1_k 2_\ell} = 0$, $W_{3_k 4_\ell} = 0$, $W_{3_k 5_\ell} = 0$, $W_{4_k 5_\ell} = 0$.

(d) Maximum Average Statistical Separability Between Spectral Class Pairs of Mild Blight (Blight Levels 0 and 1) and Very Severe Blight (Blight Levels 4 and 5).

Between all spectral class pairs of blight levels 0 and 1, blight levels 4 and 5, weights are set equal to zero, i.e., $W_{0_k 1_\ell} = 0$, $W_{4_k 5_\ell} = 0$. Spectral classes of blight levels 2 and 3 were not included in the analysis.

Let $\{D^n_{TMAVG}\}_a$, $\{D^n_{TMAVG}\}_b$, $\{D^n_{TMAVG}\}_c$, and $\{D^n_{TMAVG}\}_d$ denote the values of $\{D^n_{TMAVG}\}$ in Sections III(a), III(b), III(c) and III(d) (mentioned above), respectively. Bartlett's Test (Ostle, 1969) was used to test for the homogeneity of variances of $\{D^1_{TMAVG}\}_a$, $\{D^2_{TMAVG}\}_b$, $\{D^1_{TMAVG}\}_c$ and $\{D^1_{TMAVG}\}_d$. No evidence was found to reject the hypothesis of homogeneous variances for α level of 0.001; hence, means of $\{D^1_{TMAVG}\}_a$, $\{D^1_{TMAVG}\}_b$, $\{D^1_{TMAVG}\}_c$ and $\{D^1_{TMAVG}\}_d$ could be compared on the same basis. A low value of α was taken to guard against rejecting the hypothesis when it is actually true. The same analysis was done for each of n (no. of spectral channels) = 2, 3 and 4, individually and the same results were obtained as for n = 1. Let $\{\overline{D}^n_{TMAVG}\}$ denote the mean of the components of $\{D^n_{TMAVG}\}$. It was found that each of $\overline{D}^1_{TMAVG}$, $\overline{D}^2_{TMAVG}$, $\overline{D}^3_{TMAVG}$ and $\overline{D}^4_{TMAVG}$ individually, increased consistently as the authors went from Section III(a) to Section III(b) .... to Section III(d) (Figure 4). Also, this trend was generally found in most of the flightlines. Since the average transformed divergence between spectral class pairs is the measure of average separability between them, this trend shows that the greater the difference between the severity of spectral blight classes, the more separable they are.

IV. Maximum Average Statistical Separability Between the Spectral Class Pairs of All Possible Pairs Of Blight Levels.

A more detailed analysis of the statistical separability of spectral classes of blighted corn is done in this Section than in Section III using the same data as used in Section III. A subset of each of one, two, three and four spectral channels which maximized the average transformed divergence given by Eq. (4) was selected and the maximum average transformed divergence (maximized over a subset of channels) was tabulated.

Let $$D^n_{TAVG\ ij} = [\frac{1}{N} \sum_{k=1}^{n_i} \sum_{\ell=1}^{n_j} W_{i_k j_\ell} D^n_T (i_k, j_\ell \mid C_1, C_2 \dots C_n)] \qquad (4)$$

where

$D^n_{TAVG\ ij}$ denotes average transformed divergence between spectral class pairs of blight level i and blight level j in n spectral channels $C_1, C_2 \dots C_n$.

The rest of the notation is the same as in Eq. (3).

Let $D^n_{TMAVG\ ij}$ denote maximum of $D^n_{TAVG\ ij}$, maximized over a set of n spectral channels. Note that unlike Eq. (3), summation over indices i and j (which denote the blight levels) is not done in Eq. (4).

The maximum average transformed divergence between spectral class pairs of all possible pairs of blight levels, in n spectral channels of p^{th} flightline can be conveniently represented by a triangular matrix, as follows.

Let $[D^{n}_{TMAVG\ ij}]_{p}$, i = 0, 1 ... 4 and j = i+1, ... 5, denote a 5 × 5 triangular matrix with elements $D^{n}_{TMAVG\ ij}$ in p^{th} flightline. $[D^{n}_{TMAVG\ ij}]_{p}$ was computed for n=1, 2, 3, 4 using Eq. (4) with $W_{i_k j_\ell}=1$ for all values of i, j, k and ℓ, in each of the ten flightlines individually (p = 1, 2 ... 10), i.e., maximum average transformed divergence between spectral class pairs of all possible pair of blight levels in one, two, three and four spectral channels was calculated in each of ten flightlines individually.

Mean and variance of each of the elements of $[D^{1}_{TMAVG\ ij}]_{p}$ over all the flightlines (p = 1, 2 ... 10) were calculated. Bartlett's Test (Ostle, 1969) was used to test for the homogeneity of the variances of the elements of $[D^{1}_{MAVG\ ij}]$ found above. No evidence was found to reject the hypothesis of homogeneous variances for α level of 0.001; hence, the means of elements of $[D^{1}_{TMAVG\ ij}]_{p}$, p = 1, 2 ... 10, could be compared on the same basis. The same analysis was done on the elements of $[D^{n}_{TMAVG\ ij}]_{p}$ for each of n = 2, 3 and 4, individually, and the same results were found as for n = 1.

Let the elements of $[\overline{D}^{n}_{TMAVG\ ij}]$ denote the average of the corresponding elements of $[D^{n}_{TMAVG\ ij}]_p$ over all flightlines. It was found that $\bar{D}^{n}_{TMAVG\ ij+1}$ was greater than (in most cases) or almost equal to (in some cases) $\bar{D}^{n}_{TMAVG\ ij}$, for $i = 0, 1 \ldots 4$ and $j = i+1, \ldots 5$ for each of $n = 1, 2, 3$ and 4 spectral channels *individually*, i.e., it means that the greater the difference between the blight levels, the greater the maximum average transformed divergence between their spectral class pairs, and, hence, more separable the blight levels are (Figure 5). This trend was also generally found in most of the flightlines.

V. Conclusion

The conclusions based on the analysis of limited amount of data of ten flightlines are:

The greater the difference between the blight levels, the more statistically separable they usually are. This result is very encouraging considering the fact that there are other variables within and between the flightlines -- soil variables, plant variables (percent ground cover is likely to be very important) and environmental variables -- which can cause significant differences in reflection and emission from the plant canopy (whereas blight levels 1 and 2 cause brown lesions on the *lower* leaves which is likely to cause only relatively subtle changes in the reflection and emission from the healthy plant canopy). In addition, a number of human decisions and errors are involved in this analysis. For example, errors in rating the blight level in ground observations, scanner errors (geometric distortion, detector and system noise, calibration uncertainty, etc.). The analysis presented here has much practical application for it gives the maximum average transformed divergence between the spectral class pairs of blight levels, from which hopefully in the near future, classification accuracy will be reasonably predicted.

The authors believe that no previous authors have investigated the statistical separability of spectral classes of blighted corn in so much detail, data quantity (168 fields having 18804 sample points in ten flightlines) and depth as the authors have. Although the analysis was done for corn blight only, the conclusions obtained from this analysis may well be applicable to other crop stresses, because corn blight is representative of the problems of many plant stresses, especially non-systemic stresses.

Acknowledgements

We gratefully acknowledge the assistance, continued encouragement and guidance of Mrs. G. F. Santner, Dr. P. H. Swain and Dr. V. L. Anderson of Purdue University. We also wish to thank Dr. D. A. Landgrebe, Dr. M. E. Bauer, Dr. J. C. Lindenlaub, Dr. J. E. Cipra, and Mr J. G. Stockton of Purdue University for their suggestions and advice.

References

MacDonald, R. B., M. E. Bauer, R. D. Allen, J. W. Clifton, J. D. Erickson, and D. A. Landgrebe (1972), Proc. Eighth Symp. Remote Sens. Environ., Ann Arbor, Michigan.

Marill, T. and D. M. Green (1963), I. E. E. E. Trans. Information Theory IT-9, 11-17.

Ostle, B. (1969), Statistics in Research, The Iowa State University Press.

Swain, P. H. (1972), LARS Information Note 111572, LARS, Purdue University, W. Lafayette, Indiana.

Ullstrup, A. J. (1961), USDA Handbook No. 199.

Figure Captions

Fig. 1. Southern corn leaf blight. The disease is characterized by small, brown lesions which increase rapidly in size and number.

Fig. 2. Scale for estimating southern corn blight severity.

Fig. 3. Relationship of separability and
(a) probability of correct classification,
(b) divergence, (c) transformed divergence

(Taken from Swain (1972))

Fig. 4. Statistical separability between the spectral class pairs of mild, intermediate and severe blight.

Fig. 5. Statistical separability between the spectral class pairs of all possible pairs of blight levels.

Fig. 1. Southern corn leaf blight. The disease is characterized by small, brown lesions which increase rapidly in size and number.

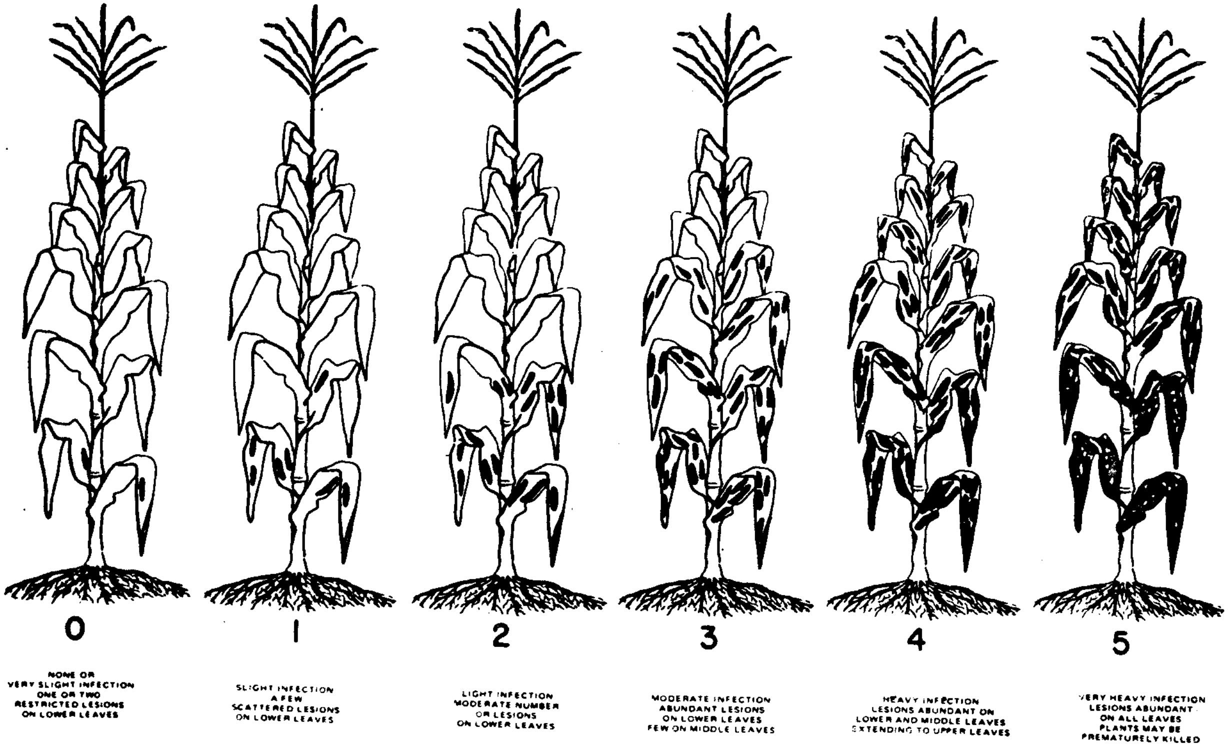
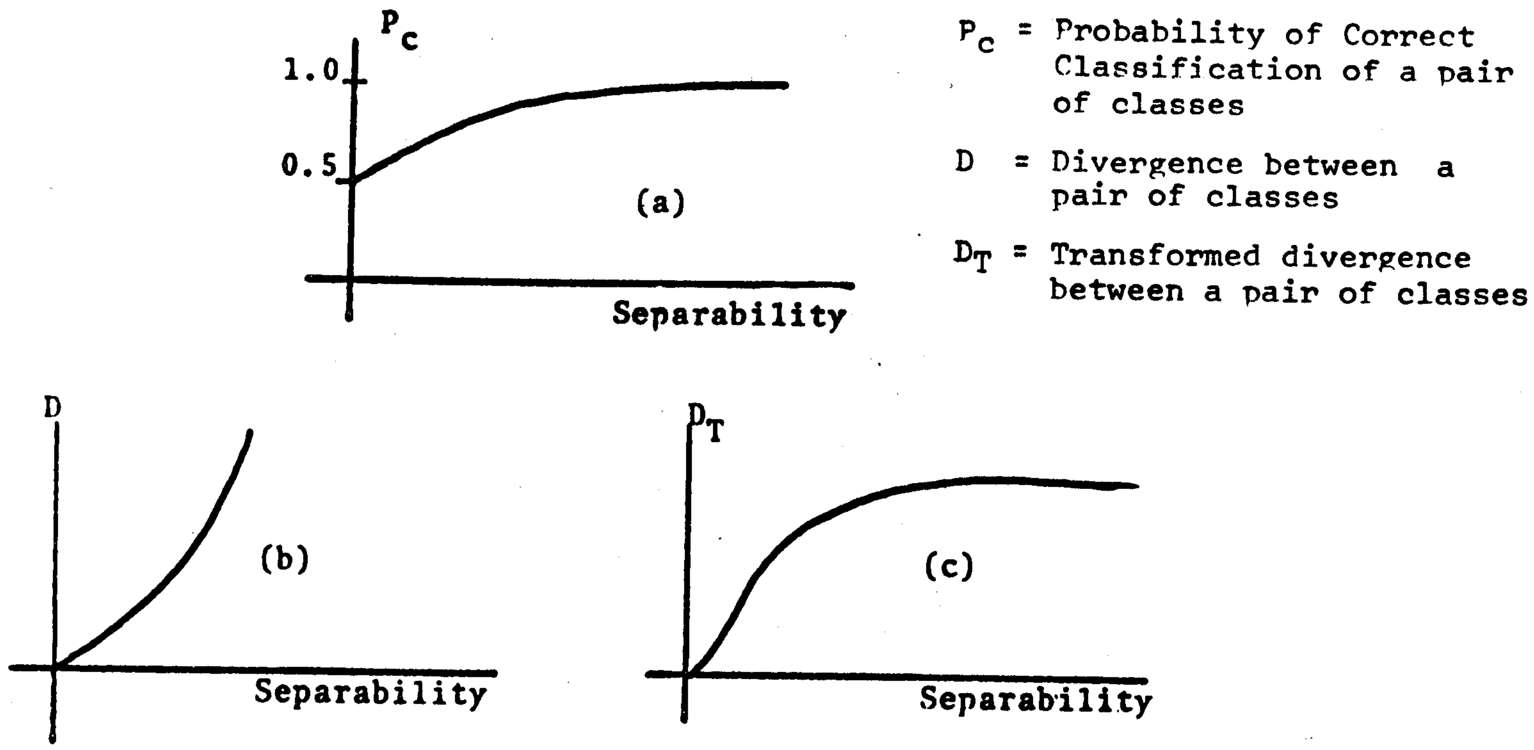


Fig. 2. Scale for estimating southern corn leaf blight severity.



P_c = Probability of Correct Classification of a pair of classes

D = Divergence between a pair of classes

D_T = Transformed divergence between a pair of classes

Fig. 3. Relationship of separability and
(a) probability of correct classification,
(b) divergence, (c) transformed divergence

(Taken from Swain (1972))

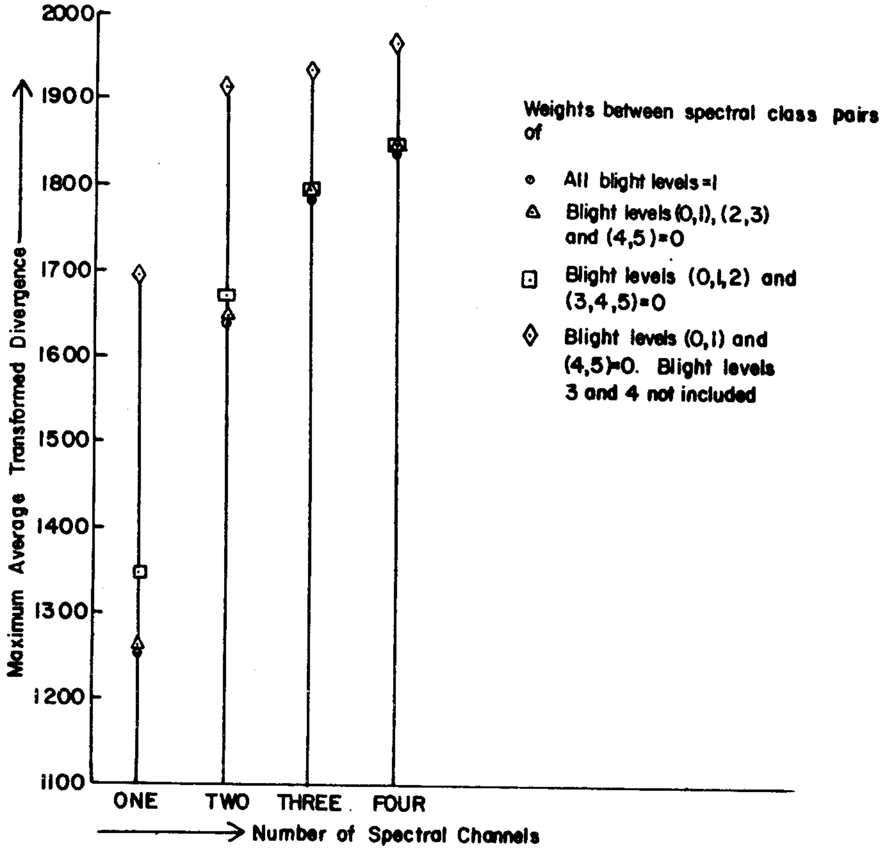


Fig. 4. Statistical separability between the spectral class pairs of mild, intermediate and severe blight.

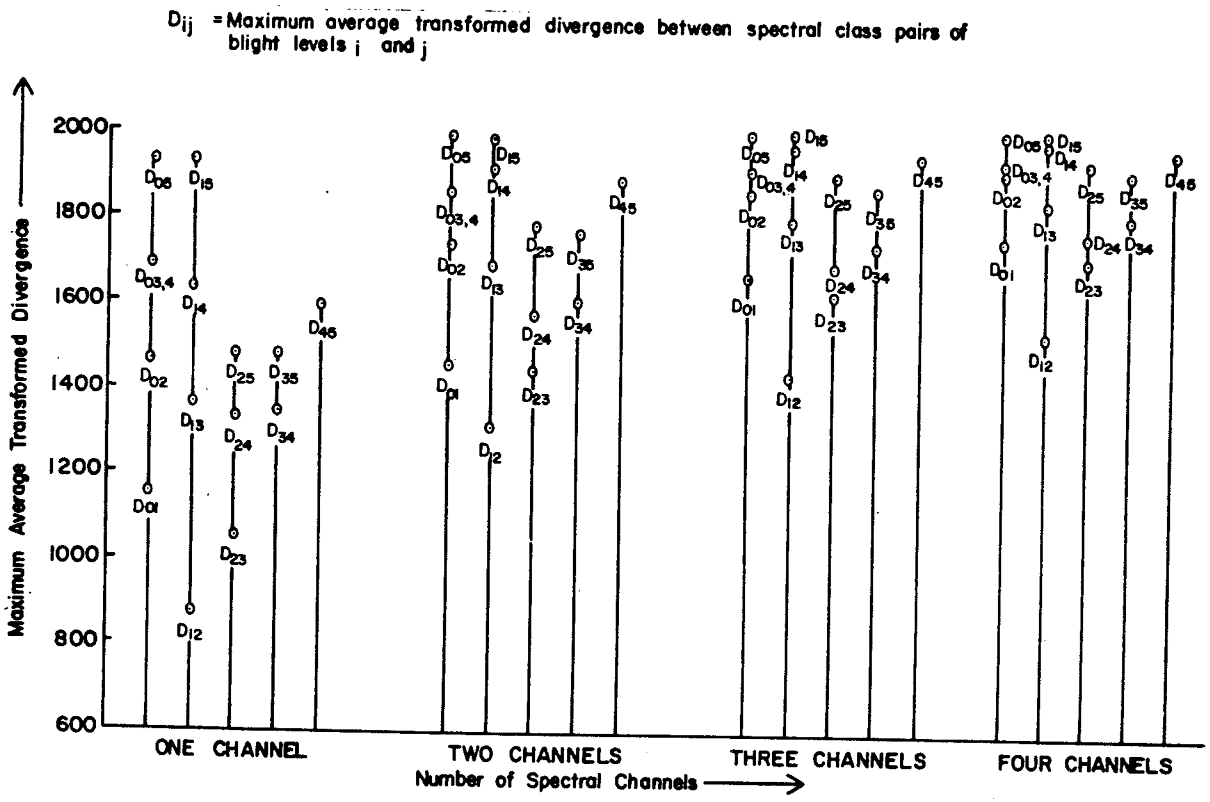


Fig. 5. Statistical separability between the spectral class pairs of all possible pairs of blight levels.