NASA CR-
141877

(NASA-CR-141877) OPTIMAL SELECTION OF PASSES (Texas A&M Univ.) 19 p HC $3.25 CSCL 05B

N75-26480

Unclas
G3/43 27284

DEPARTMENT OF MATHEMATICS

TEXAS A&M UNIVERSITY
COLLEGE STATION TEXAS

OPTIMAL SELECTION OF PASSES

by

L. F. Guseman, Jr. and Bruce P. Marion

Department of Mathematics
Texas A&M University

Prepared For

Earth Observations Division
NASA/Johnson Space Center
Houston, Texas

Contract NAS-9-13894-1S

January, 1975

Report #2

OPTIMAL SELECTION OF PASSES

L. F. Guseman, Jr. and Bruce P. Marion

1. Introduction

This report presents preliminary numerical results obtained from the application of a linear feature selection technique to the determination of combinations of passes which best discriminate between a given set of crops in a given area of interest. The results obtained are not purported to hold in a general situation, but only for the given set of crops and the given (but unknown) levels of several factors (such as soil type, fertilizer practice, etc.) holding in the area of interest. Nevertheless, some insight into optimal choices of passes can be gained through the application of the technique to data sets with different levels of those factors which are known to affect the spectral signatures. The technique could also be useful in determining those factors which significantly affect the choice of passes used for discrimination between a given set of crops.

For a given set of r registered ERTS passes, the input data consists of a 4r-dimensional mean vector and a $4r \times 4r$ covariance matrix for each of the m classes (assumed to be multivariate normal). For each of the combinations of r passes taken k at a time ($1 \leq k \leq r$) the portions of the mean vectors and covariance matrices pertaining to the k passes $i_1, i_2, \ldots, i_k$ are used to compute a 4k-dimensional vector $B(i_1, \ldots, i_k)$ of norm one whose linear combination of the channels in the passes $i_1, i_2, \ldots, i_k$ minimizes the probability of misclassification, $g(B)$, for the transformed (one-dimensional) densities of the m classes. The feature selection technique and associated computer program are described in [1].

In Section 2 we present numerical results obtained when using 6 registered passes of ERTS data from Hill County North.

2. Numerical Results--Hill County North

Input data for the runs described in this section consisted of 24-dimensional mean vectors and 24 x 24 covariance matrices computed from training data using six registered passes of ERTS data from Hill County North. The six registered ERTS passes (identified below) were, for purposes of this note, assigned the following pass numbers and channel numbers. The associated biological phases for Winter Wheat are also given.

Pass Number	Channel Numbers	Date of Pass	Scene Number	Biological Phase (Winter Wheat)
1	1-4	Aug. 20, 1973	1393-17392	Harvest
2	5-8	Aug. 3, 1973	1376-17452	Yellow/Harvest
3	9-12	July 16, 1973	1358-17453	Heading/Yellow
4	13-16	June 27, 1973	1339-17400	Heading
5	17-20	June 11, 1973	1332-17460	Heading
6	21-24	May 23, 1973	1304-17461	Jointing (Greening)

(Note that the passes are numbered in reverse chronological order)

The crops under consideration and their designated class numbers were:

Class	Crop
1	Barley
2	Stubble
3	Grass
4	Winter Wheat
5	Spring Wheat
6	Fallow

For each of the sixty-three combinations of 6 passes taken k at a time ($1 \leq k \leq 6$), the appropriate signatures were used to compute a vector B of norm one (and appropriate dimension) whose linear combination of the channels minimizes the one-dimensional probability of misclassification. The resulting values, g(B), of the one-dimensional probability of misclassification determined by the optimal B vectors for the various combinations of passes appear in Tables 1-5 below. The value for g(B) using all channels within <u>all six passes</u> was .2699. The graphs of the six transformed density functions determined by the optimal B vector for the optimal one, two, three, four, five and six passes appear in Figures 1-6, respectively.

Additional runs were made using all six passes where Spring Wheat and Winter Wheat were alternately deleted and the remaining five crops (renumbered in same order) considered. The results from these runs appear in Figures 7 and 8, respectively.

The confusion matrices associated with Figures 1-8 appear in Tables 6 and 7.

Pass	1	2	3	4	5	6
g(B)	.5226	.4811	.4326	.6600	.4488	.4832

Table 1. Values of g(B) for $\binom{6}{1}$

Passes	1,2	1,3	1,4	1,5	1,6
g(B)	.4700	.4045	.4873	.4276	.4703

Passes	2,3	2,4	2,5	2,6	3,4
g(B)	.4342	.4730	.3873	.3807	.4256

Passes	3,5	3.6	4,5	4,6	5,6
g(B)	.3551	.4332	.4453	.4756	.3962

Table 2. Values of g(B) for $\binom{6}{2}$

Passes	1,2,3	1,2,4	1,2,5	1,2,6	1,3,4
g(B)	.3980	.4349	.3762	.3305	.3964

Passes	1,3,5	1,3,6	1,4,5	1,4,6	1,5,6
g(B)	.3453	.3618	.4239	.4432	.2910

Passes	2,3,4	2,3,5	2,3,6	2,4,5	2,4,6
g(B)	.4164	.3266	.3789	.3844	.3762

Passes	2,5,6	3,4,5	3,4,6	3,5,6	4,5,6
g(B)	.2951	.3485	.4259	.3048	.3104

Table 3. Values of g(B) for $\binom{6}{3}$

Passes	1,2,3,4	1,2,3,5	1,2,3,6	1,2,4,5	1,2,4,6
g(B)	.3917	.3215	.3226	.3732	.3247

Passes	1,2,5,6	1,3,4,5	1,3,4,6	1,3,5,6	1,4,5,6
g(B)	.2764	.3377	.3521	.2873	.2873

Passes	2,3,4,5	2,3,4,6	2,3,5,6	2,4,5,6	3,4,5,6
g(B)	.3229	.3738	.2925	.2909	.3005

Table 4. Values of g(B) for $\binom{6}{4}$

Passes	1,2,3,4,5	1,2,3,4,6	1,2,3,5,6	1,2,4,5,6	1,3,4,5,6	2,3,4,5,6
g(B)	.3171	.3140	.2733	.2718	.2842	.2886

Table 5. Values of g(B) for $\binom{6}{5}$

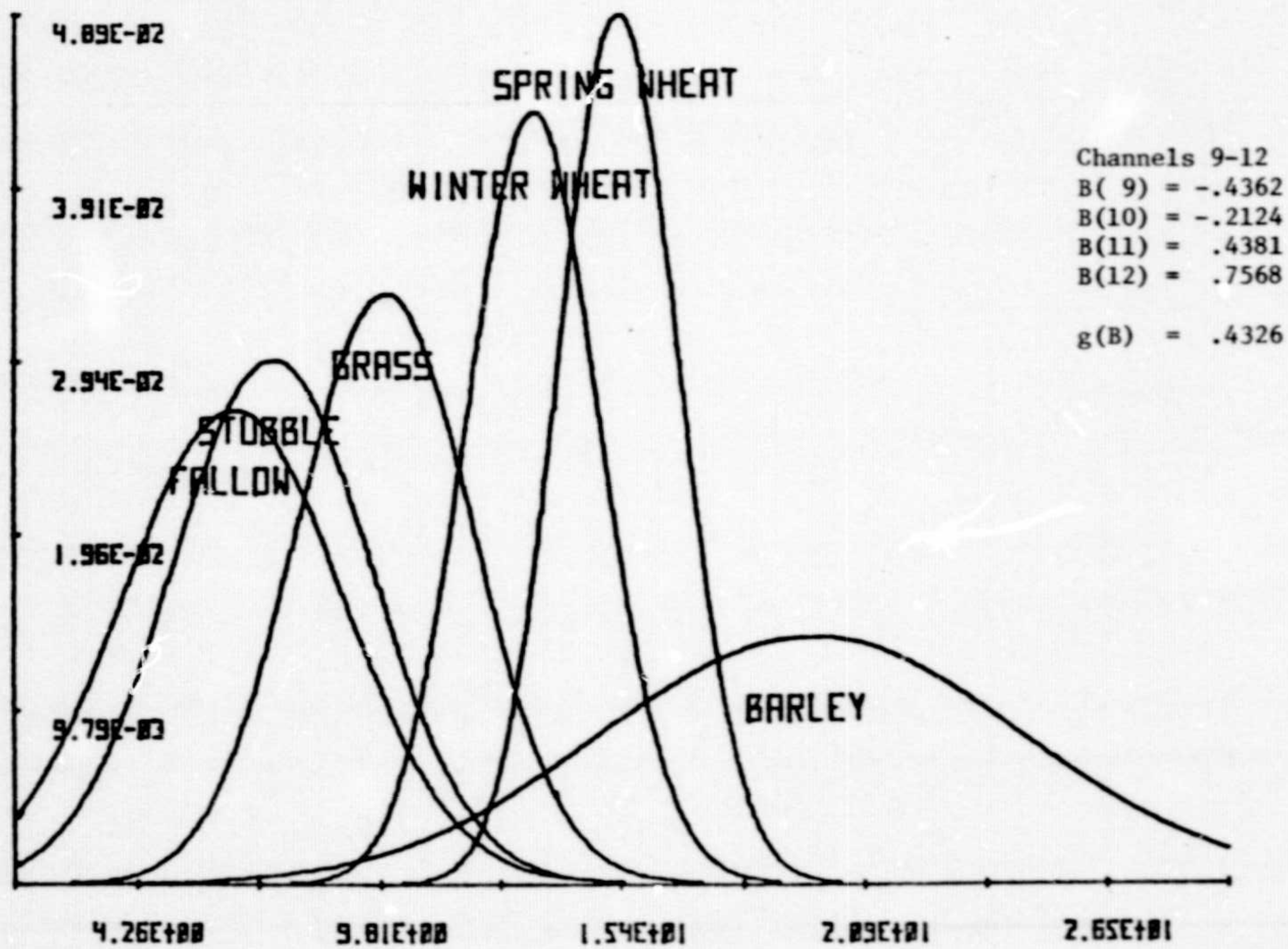
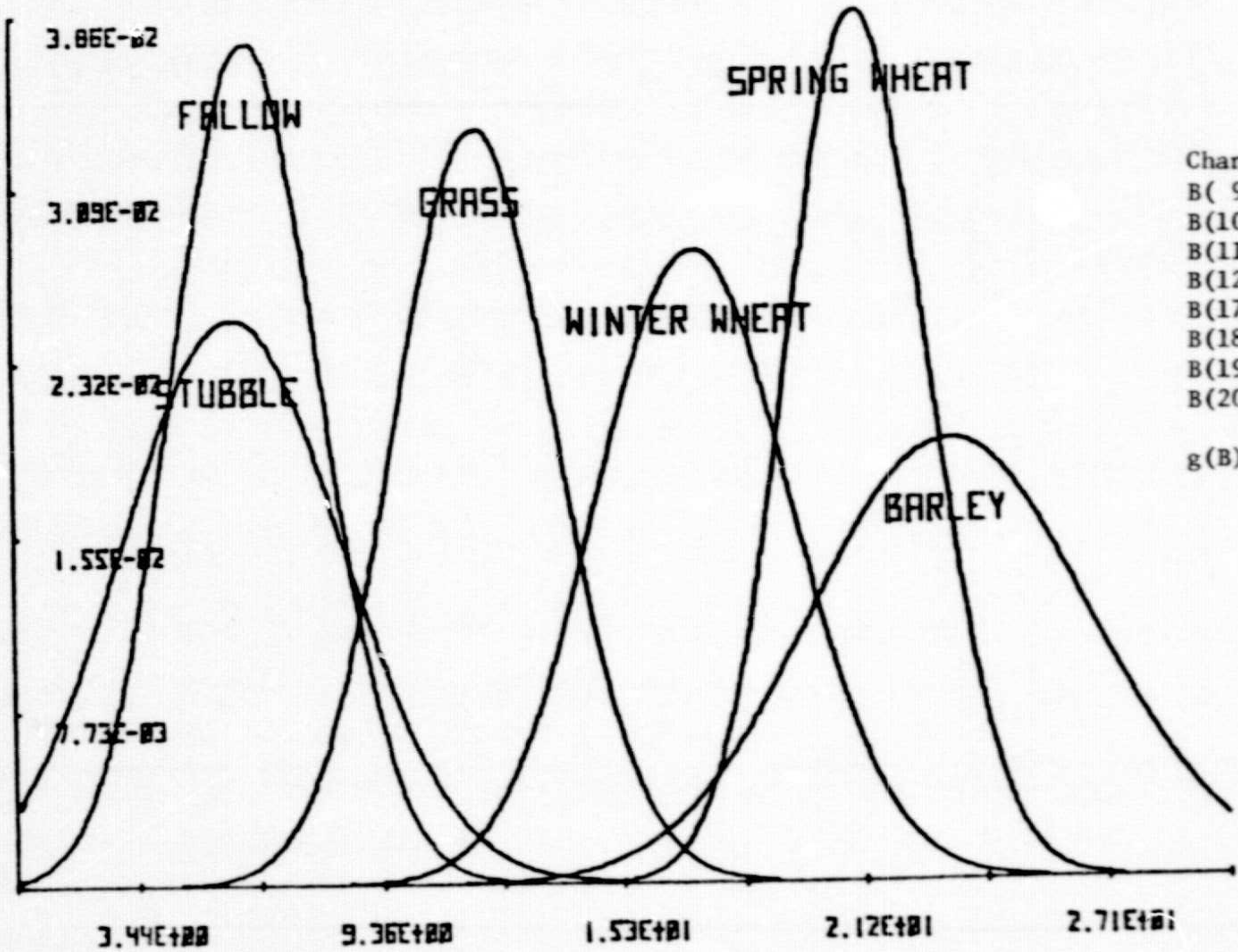


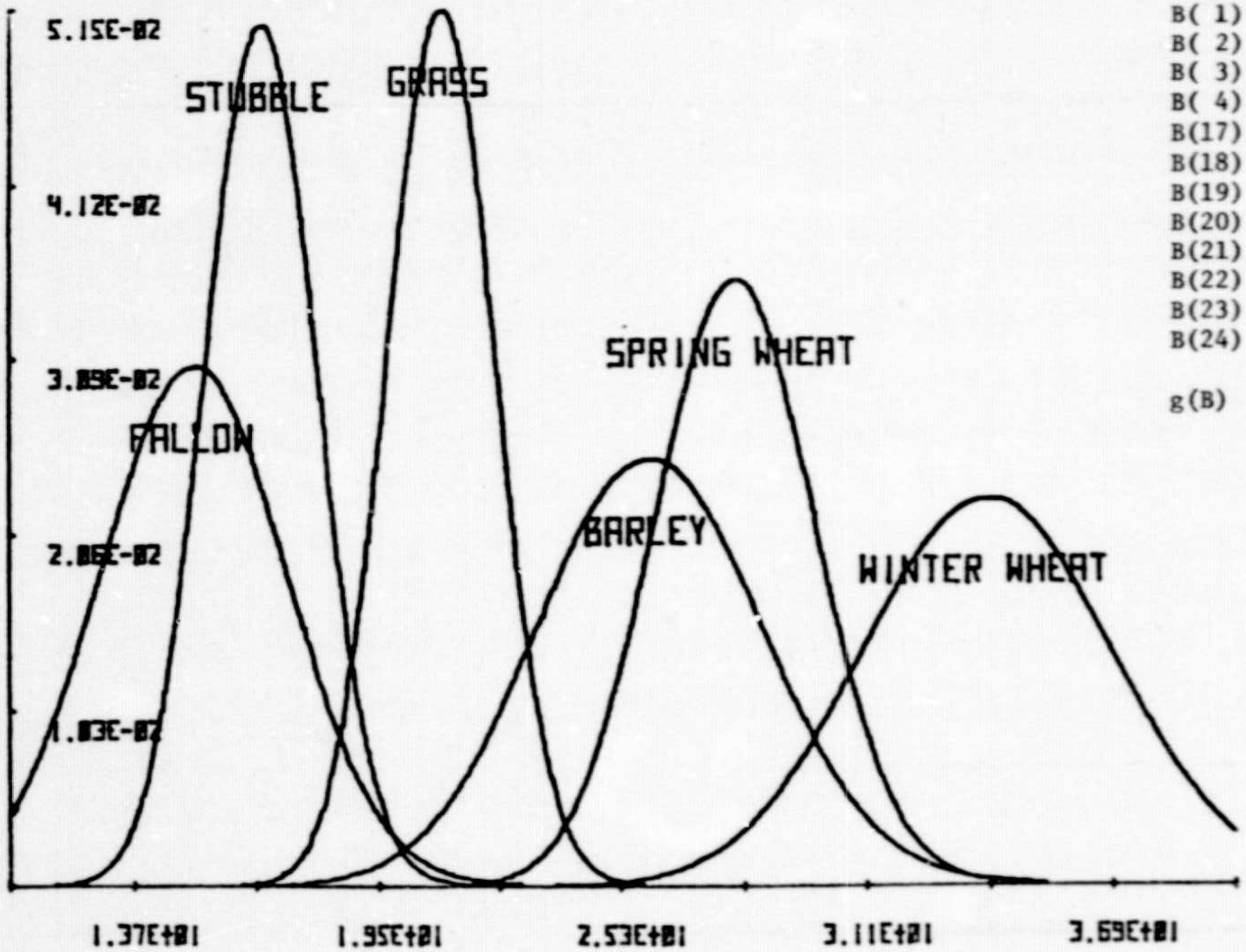
Fig. 1. Optimal B, g(B), Transformed densities: Best one pass-3.



Channels 9-12, 17-20
B(9) = -.6167
B(10) = -.2063
B(11) = .3985
B(12) = .0084
B(17) = .1407
B(18) = -.2038
B(19) = .5936
B(20) = .6725

g(B) = .3551

Fig. 2. Optimal B, g(B), Transformed densities: Best two passes-3,5.

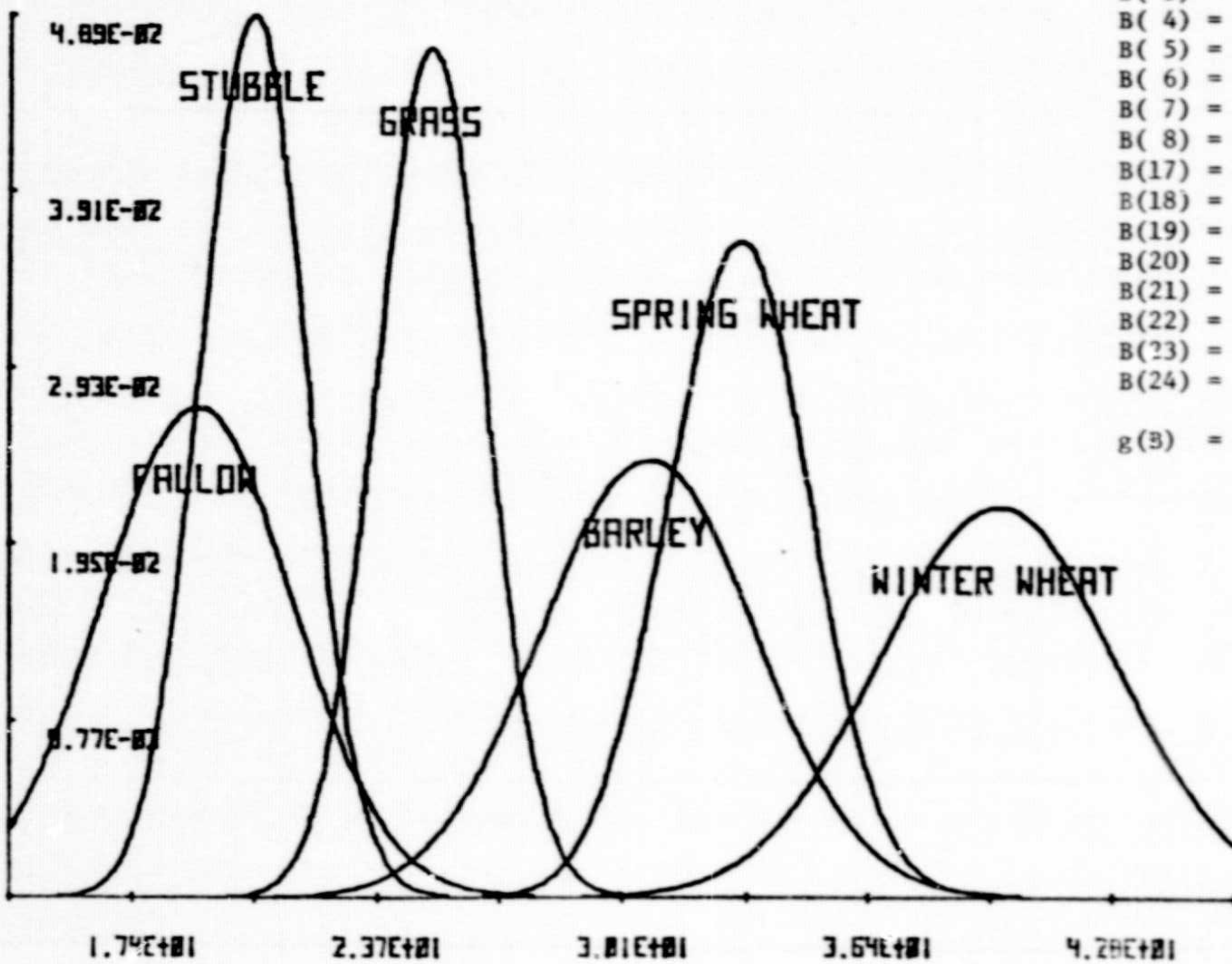


Channels 1-4, 17-24

B(1)	=	.0016
B(2)	=	.1567
B(3)	=	.0805
B(4)	=	.0767
B(17)	=	-.0779
B(18)	=	-.2536
B(19)	=	.0968
B(20)	=	.3538
B(21)	=	.0655
B(22)	=	-.2743
B(23)	=	-.0438
B(24)	=	.8227

g(B) = .2910

Fig. 3. Optimal B, g(B), Transformed densities: Best three passes-1,5,6.

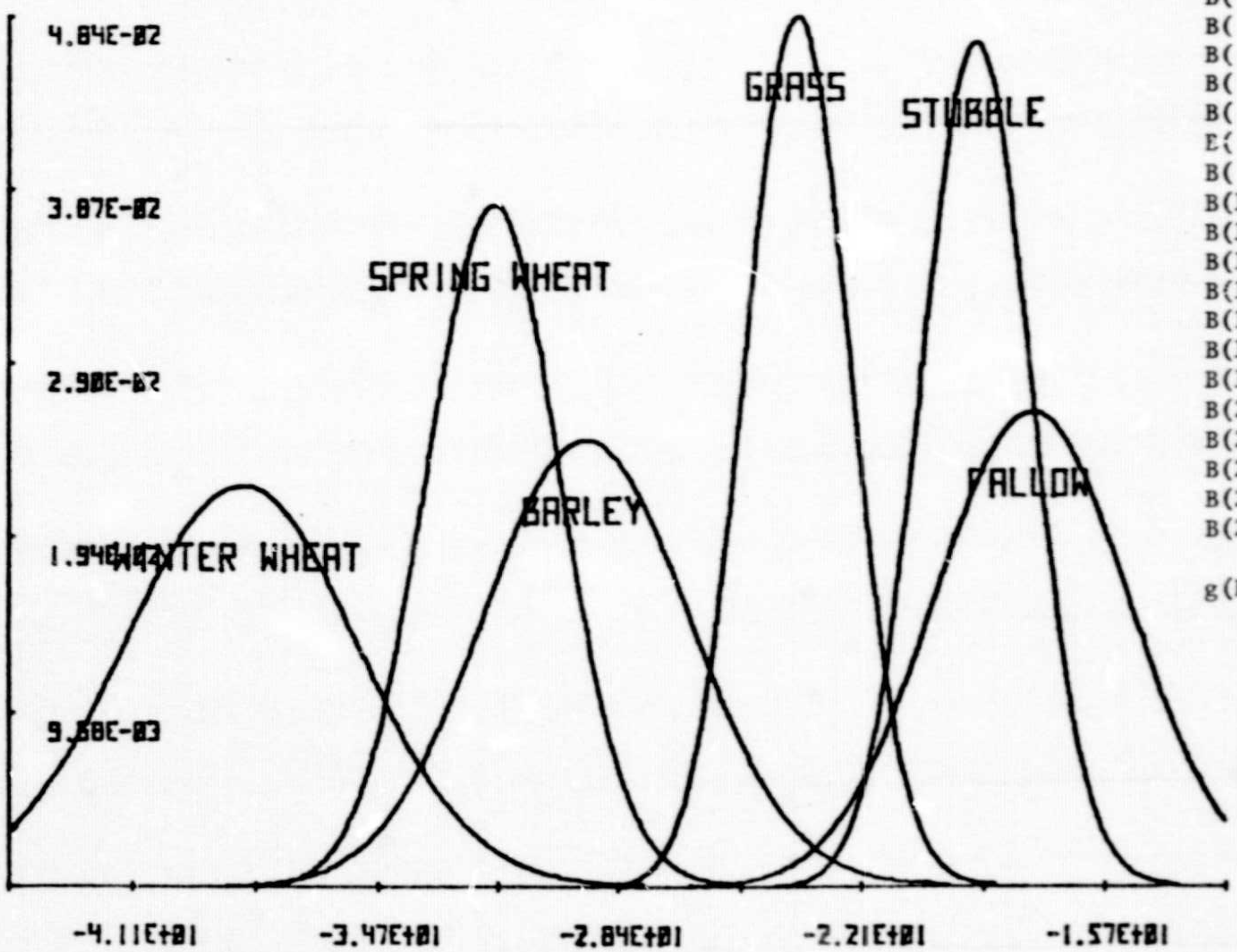


Channels 1-8, 17-24

B(1) = .0038
B(2) = .0908
B(3) = .0875
B(4) = .1604
B(5) = .0092
B(6) = .0578
B(7) = .0508
B(8) = .0034
B(17) = -.0789
B(18) = -.2124
B(19) = .0946
B(20) = .3456
B(21) = .0695
B(22) = -.3161
B(23) = -.0417
B(24) = .8162

g(B) = .2764

Fig. 4. Optimal B, g(B), Transformed densities: Best four passes-1,2,5,6.

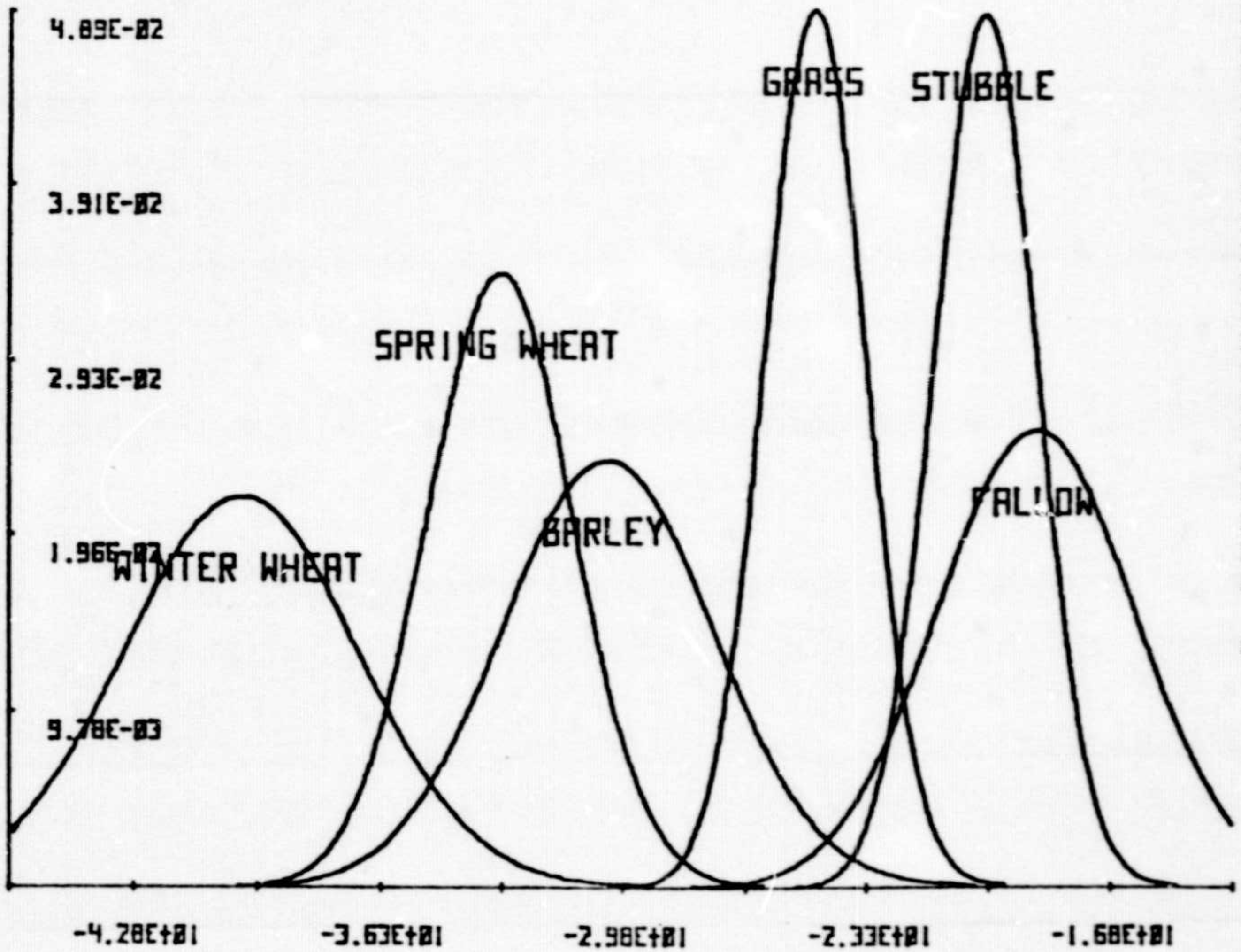


Channels 1-8, 13-24

B(1) = -.0244
B(2) = -.0678
B(3) = -.0980
B(4) = -.1546
B(5) = -.0045
B(6) = -.0468
E(7) = -.0589
B(8) = -.0181
B(13) = .2297
B(14) = -.1478
B(15) = -.0488
B(16) = .0296
B(17) = .0830
B(18) = .2098
B(19) = -.1269
B(20) = -.2987
B(21) = -.0444
B(22) = .3041
B(23) = .0175
B(24) = -.7914

g(B) = .2718

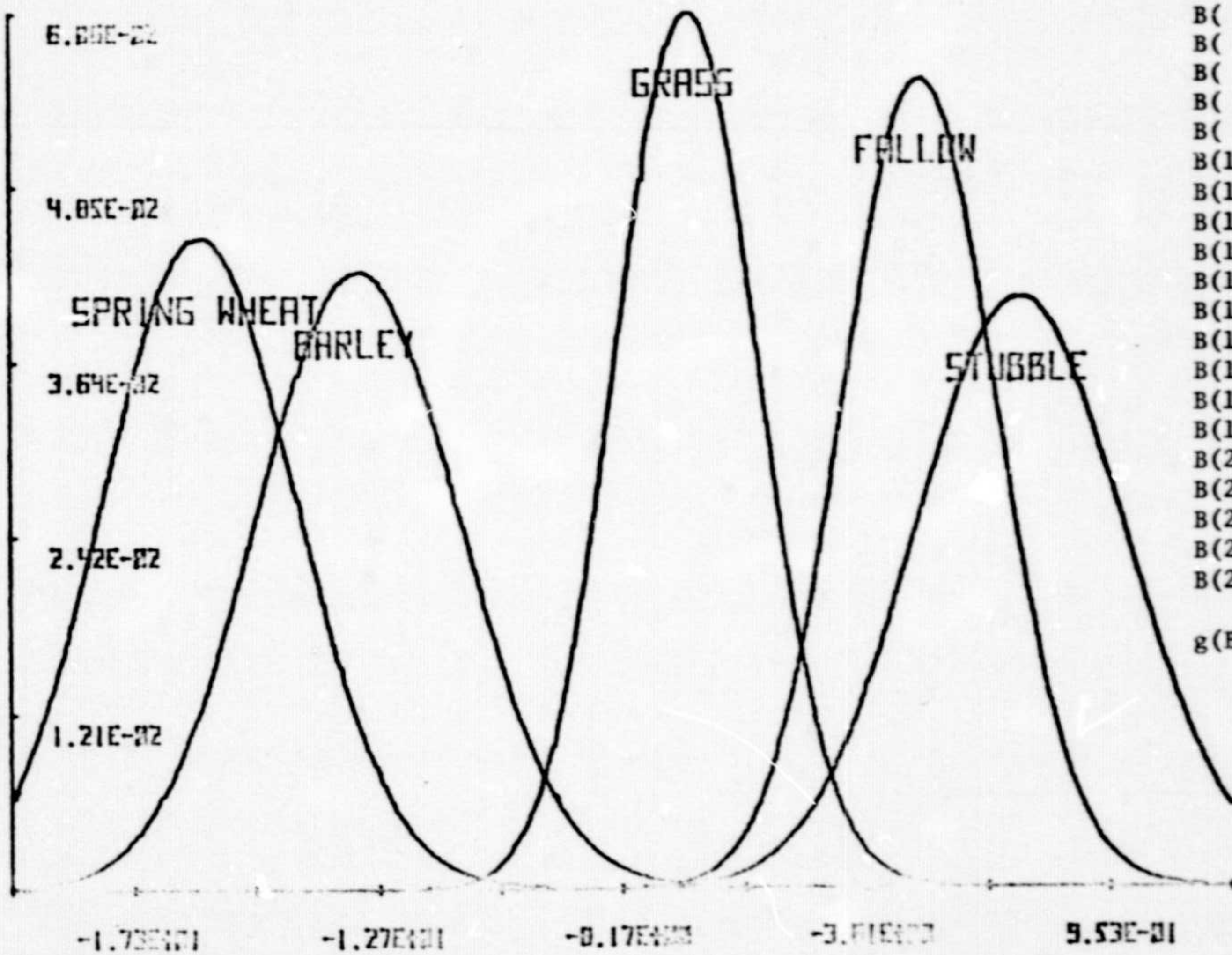
Fig. 5. Optimal B, g(B), Transformed densities: Best five passes-1,2,4,5,6.



Channels 1-24
B(1) = -.0069
B(2) = -.0759
B(3) = -.1168
B(4) = -.1916
B(5) = .0542
B(6) = -.0424
B(7) = -.0729
B(8) = -.0960
B(9) = -.0960
B(10) = -.0217
B(11) = .0359
B(12) = .0660
B(13) = .2060
B(14) = -.1399
B(15) = -.0319
B(16) = .0081
B(17) = .1042
B(18) = .2067
B(19) = -.1302
B(20) = -.3086
B(21) = -.0255
B(22) = .3128
B(23) = .0148
B(24) = -.7626

g(B) = .2699

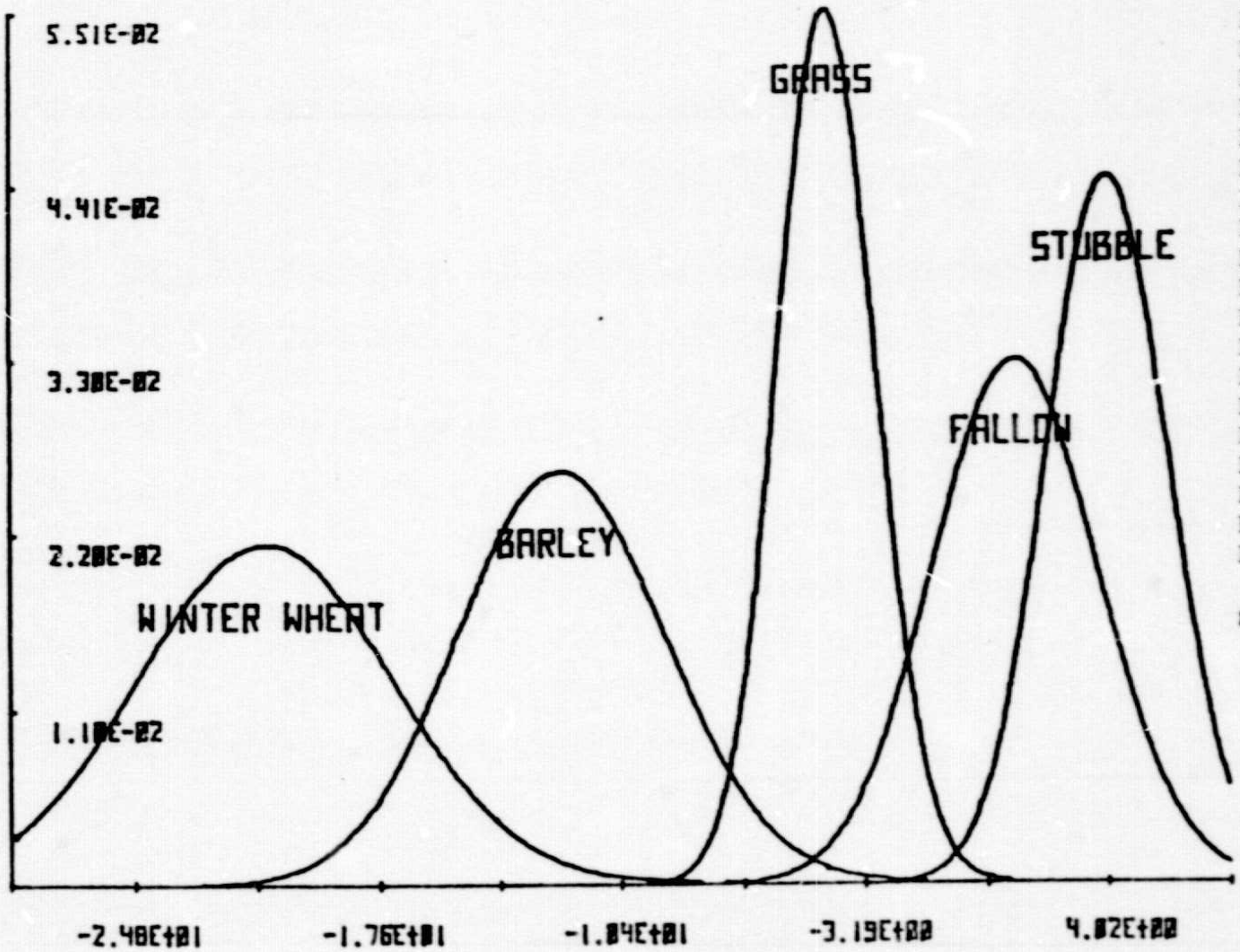
Fig. 6. Optimal B, g(B), Transformed densities: All passes-1,2,3,4,5,6.



Channels 1-24
B(1) = .1185
B(2) = .0259
B(3) = -.0827
B(4) = -.3607
B(5) = -.0676
B(6) = .1548
B(7) = -.1239
B(8) = -.3585
B(9) = .3096
B(10) = -.1510
B(11) = .2829
B(12) = .2331
B(13) = .1013
B(14) = -.1326
B(15) = -.0466
B(16) = -.0326
B(17) = -.1502
B(18) = .2032
B(19) = -.3334
B(20) = -.3820
B(21) = .0890
B(22) = .0627
B(23) = -.0558
B(24) = .2361

g(B) = .2298

Fig. 7. Five classes (no Winter Wheat): Optimal B, g(B), Transformed densities: All passes.



Channels 1-24
B(1) = .1783
B(2) = -.1129
B(3) = -.1789
B(4) = -.4214
B(5) = -.1422
B(6) = .0319
B(7) = -.3098
B(8) = -.1218
B(9) = .2948
B(10) = -.0689
B(11) = .1912
B(12) = .2480
B(13) = -.0114
B(14) = .0239
B(15) = -.0611
B(16) = -.0349
B(17) = -.1142
B(18) = .2285
B(19) = -.1552
B(20) = -.3886
B(21) = .1882
B(22) = .3325
B(23) = -.0322
B(24) = .1762

g(B) = .1842

Fig. 8. Five classes (no Spring Wheat): Optimal B, g(B), Transformed densities: All passes.

.668	.000	.000	.003	.050	.000
.007	.373	.203	.001	.000	.316
.037	.274	.562	.131	.003	.187
.084	.028	.183	.593	.216	.019
.201	.001	.015	.272	.731	.001
.003	.324	.036	.000	.000	.478

Best one pass - 3

.525	.000	.000	.005	.121	.000
.000	.307	.019	.000	.000	.165
.004	.116	.825	.122	.000	.059
.102	.001	.096	.680	.122	.000
.369	.000	.000	.192	.757	.000
.000	.576	.060	.000	.000	.776

Best two passes - 3, 5

.423	.000	.053	.005	.193	.000
.004	.794	.049	.000	.000	.410
.137	.050	.897	.000	.005	.043
.037	.000	.000	.864	.074	.000
.399	.000	.000	.131	.729	.000
.000	.156	.000	.000	.000	.547

Best three passes - 1, 5, 6

.475	.000	.049	.003	.175	.000
.002	.813	.050	.000	.000	.441
.115	.048	.901	.000	.001	.065
.028	.000	.000	.890	.055	.000
.380	.000	.000	.106	.769	.000
.000	.139	.000	.000	.000	.494

Best four passes - 1, 2, 5, 6

Table 6. Confusion matrices.

.493	.000	.047	.003	.168	.000
.002	.810	.047	.000	.000	.446
.112	.049	.906	.000	.001	.066
.025	.000	.000	.894	.052	.000
.368	.000	.000	.103	.779	.000
.000	.141	.000	.000	.000	.487

Best five passes - 1, 2, 4, 5, 6

.528	.000	.047	.002	.176	.001
.002	.825	.048	.000	.000	.451
.117	.048	.905	.000	.001	.083
.017	.000	.000	.894	.059	.000
.335	.000	.000	.104	.764	.000
.000	.127	.000	.000	.000	.465

All passes - 1, 2, 3, 4, 5, 6

.749	.000	.021	.199	.000
.000	.629	.000	.000	.195
.032	.018	.926	.000	.059
.219	.000	.000	.801	.000
.000	.353	.053	.000	.746

Five classes (no Winter Wheat)

.863	.000	.021	.111	.000
.000	.817	.000	.000	.322
.052	.001	.929	.000	.096
.085	.000	.000	.889	.000
.000	.183	.049	.000	.581

Five classes (no Spring Wheat)

Table 7. Confusion matrices

3. Remarks

The preliminary numerical results in the previous section indicate that fewer optimally chosen passes could be used while maintaining most of the discrimatory power inherently available in the original data. Again, we note that the results give insight only for the crops considered and levels of factors reflected by the data used. However, by identifying the various factors (an levels thereof) affecting the spectral signatures, and by formulating a regression model one could use the feature selection technique to determine the regression coefficients for predicting optimal passes for a given set of crops.

Another use of the feature selection technique as applied to multiple-pass registered data is the generation of enhanced grey scale displays by using a single linear combination of all channels of all designated passes as opposed to a single channel within a single pass.

References

1. L. F. Guseman, Jr. and Bruce P. Marion, LFSPMC: Linear Feature Selection Program Using the Probability of Misclassification, Report #3, NAS-9-13894-1S, Texas A&M University, Department of Mathematics, January, 1975.