LA--11694-MS

DE90 004424

Status of Image Analysis Methods to Delineate Stratigraphic Position in the Topopah Spring Member of the Paintbrush Tuff, Yucca Mountain, Nye County, Nevada

Katherine Campbell
David E. Broxton
*Joan Spaw**

DISCLAIMER

*Everest Technologies, 1800 Bering Drive, Suite 800, Houston, TX 77057.

STATUS OF IMAGE ANALYSIS METHODS TO DELINEATE STRATIGRAPHIC POSITION IN THE TOPOPAH SPRING MEMBER OF THE PAINTBRUSH TUFF, YUCCA MOUNTAIN, NYE COUNTY, NEVADA

by

Katherine Campbell, David E. Broxton, and Joan Spaw

ABSTRACT

The Topopah Spring Member of the Paintbrush Tuff is an ash-flow cooling unit that is the candidate host rock for a potential high-level nuclear waste repository at Yucca Mountain, Nevada. The repository workings will be mostly confined to the member's rhyolitic portion, which is chemically homogenous but texturally variable. This report describes the status of work to develop a useful internal stratigraphy for the rhyolitic portion of the member; our approach is to use an image analysis technique to map textural variations within the member as a function of stratigraphic height.

Fifteen petrographic thin sections of Topopah Spring rhyolitic tuff were studied in each of two drill holes (USW GU-3 and USW G-4). Digital color images were collected in transmitted light for two scenes 1 cm on a side for each thin section. Objects within a scene were classified by color, and measurements of area, elongation, and roughness were determined for each object. Summary statistics were compiled for all measurements for each color component within a scene, and each variable was statistically examined for correlations with stratigraphic position.

Our initial studies using image analysis have not yet produced a useful method for determining stratigraphic position within the Topopah Spring Member. Simplifications made in this preliminary application of image analysis may be largely responsible for these negative results. The technique deserves further investigation, and more detailed analysis of existing data is recommended.

I. INTRODUCTION

This report describes the status of investigations in which image analysis techniques are being developed to map textural variations in petrographic thin sections cut from the Topopah Spring Member of the Paintbrush Tuff. The Topopah Spring Member is an ash-flow tuff cooling unit that is the candidate host rock for a potential high-level nuclear waste repository at Yucca Mountain, Nevada.

We are attempting to develop an image analysis technique to map textural variations in the groundmass of the Topopah Spring Member as a function of stratigraphic height. The Topopah Spring Member consists of a thick layer of rhyolitic rock overlain by a relatively thin quartz latitic caprock (Lipman et al. 1966). The repository workings will be confined mostly to the rhyolitic portion of the member, which is homogenous in chemical composition (Lipman et al. 1966; Zielinski 1983) but has varied textural and phenocryst petrography (Byers 1985; Byers and Moore 1987) and mineralogy (Bish and Vaniman 1985).

Ultimately, we hope to use image analysis in conjunction with petrographic and mineralogic techniques to define recognizable stratigraphic subunits within the densely welded, devitrified portion of the rhyolite. Once defined, these subunits will be used to maintain the desired stratigraphic position within the candidate host rock during construction of drifts from the exploratory shaft and during mining of the repository. Recognition of correct stratigraphic position will be particularly important in determining amount and sense of displacement on faults crossed during mining operations. Such faults may not displace overlying tuffaceous units and thus may have no surface expression. Temperature profiles developed in Brandshaung (in review) indicate that it is desirable to maintain a buffer zone of at least 150 m between the repository and the zeolitic tuffs at the base of the Topopah Spring Member and within the tuff of Calico Hills. This buffer zone is necessary to minimize alteration of zeolite minerals that are potentially sorptive of certain radionuclides. A 150-m buffer zone is necessary to ensure that temperatures within the zeolitic tuffs do not exceed 100°C during the thermal pulse associated with the emplacement of waste. The 100°C temperature constraint is preliminary and is based on thermal stability data for the secondary minerals at Yucca Mountain (Bish 1989).

In this report we present a preliminary evaluation of the usefulness of image analysis data in defining stratigraphic subunits in the rhyolitic portion of the Topopah Spring Member. We discuss data for samples representing vertical sections sampled through the densely welded, devitrified rhyolite for two drill holes (USW GU-3 and USW G-4), which are located 3.5 km apart. Samples consisted of petrographic thin sections cut from drill core. The images were captured by digitizing nonpolarized light transmitted through the thin sections. Future studies might include the use of polarizing filters in transmitted light as well as the use of reflected light sources for both thin sections and polished rock slabs.

The image analysis data were collected under contract by Everest Technologies of Houston, Texas. The data were collected by use of the Everest Image Quantification System, which consists of hardware and software for digitizing, enhancing, mapping, analyzing, and archiving image information. The data were transferred by magnetic tape to Los Alamos for reduction and statistical analysis (see Appendix).

II. METHODS

Samples for this study were selected from three of the four stratigraphic intervals in the densely welded, devitrified rhyolite of the Topopah Spring Member, defined on the basis of petrographic observations by Byers (1985) and Byers and Moore (1987). The studied intervals consist of, in ascending order, 1) the lower nonlithophysal zone, 2) the lower lithophysal zone, and 3) the middle nonlithophysal zone (Fig. 1). A fourth zone, the upper lithophysal zone, was not included in this study because its properties are sufficiently distinct that its recognition is not expected to be a major concern during mining operations.

The thin sections used in this study were cut from drill core taken from drill holes USW GU-3 and USW G-4, located near the south and north ends of the exploration block, respectively. Four to six thin sections were analyzed from each of the three stratigraphic divisions described above. In all, 15 thin sections were studied in each of the 2 drill holes. Two nonoverlapping scenes measuring 1 cm on a side were analyzed for each thin section to examine sample heterogeneities at the thin-section scale. Individual scenes are identified by the suffixes A or B after the sample number (e.g., USW GU-3-769.1A and USW GU-3-769.1B). The number before the A or B is the depth within the drill hole in feet.

Digital images were created for the two scenes selected for each thin section. These images were captured by a solid-state high-resolution camera that measured the intensity and color of nonpolarized light transmitted through each thin section. A specially modified light table served as the light source. The camera is a linear array of 4096 silicon detectors. The images were created by exposing the row of detectors, moving the row, and exposing the detectors again; this process was continued until the entire scene was covered. In this study, the central 2048 detectors of the camera were enabled, and the collected images consist of 2048 x 2048 picture elements (4 194 304 pixels). Each image was scaled to 1 cm on a side, so each pixel represents an area measuring 5 x 5 microns. Color for each pixel is determined by sequentially scanning each image through red, green, and blue filters and combining the results to produce full-color images. Thus, each scene is represented by 3 overlapping color images (red, green, and blue); specific brightness values ranging from 0 to 256 are determined for each pixel in each of these colors. Figure 2 shows digital full-color images produced for thin sections of Topopah Spring tuff.

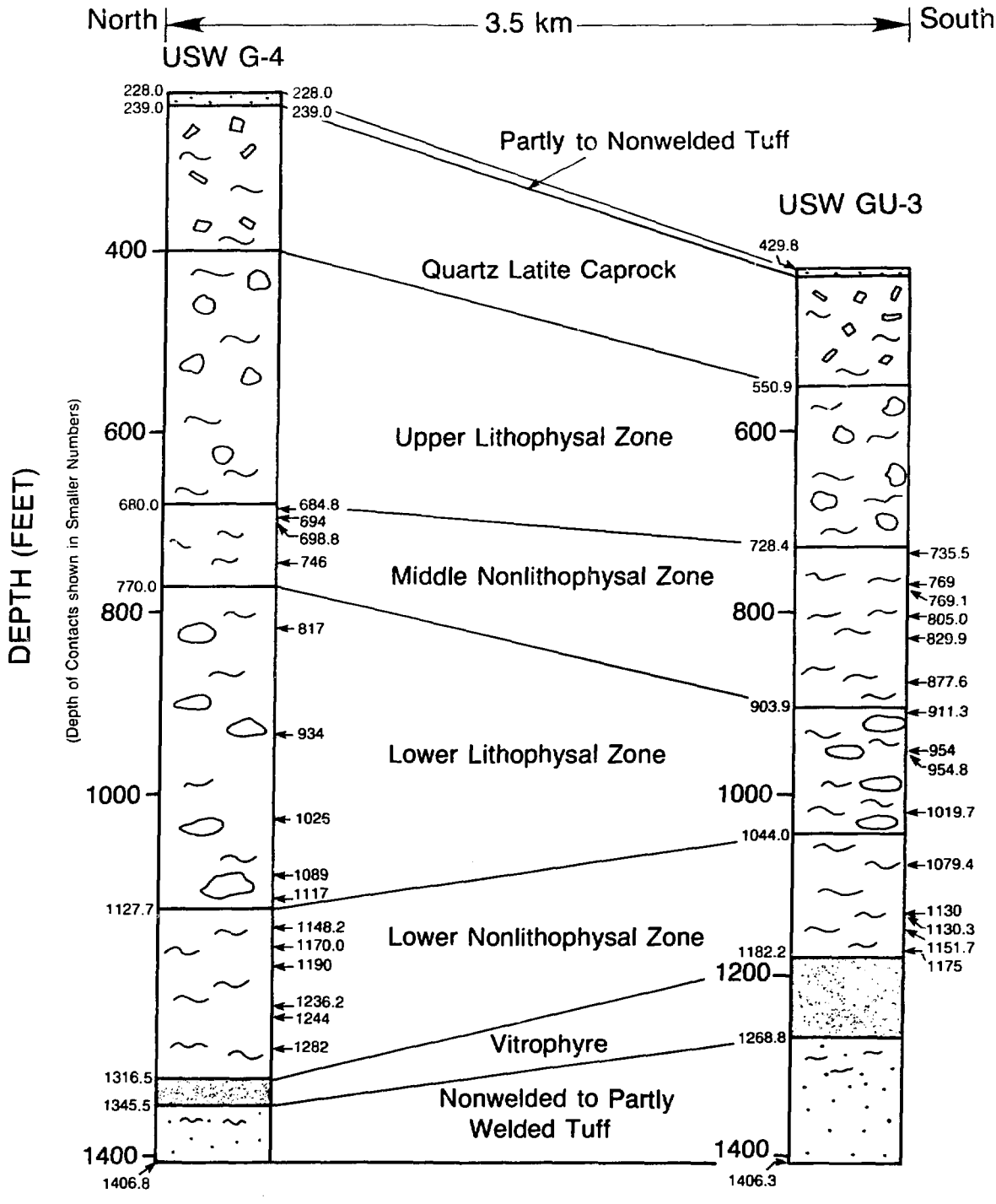

Figure 1. Stratigraphic columns of the Topopah Spring Member for drill holes USW GU-3 and USW G-4. Contacts between stratigraphic subunits taken from Byers (1985) and Byers and Moore (1987).

Figure 2A. Examples of digital full-color images of thin sections from the Topopah Spring Member. All images are 1 cm on a side and were illuminated by transmitted nonpolarized light. A) Sample USW GU-3-769.1, scene A. B) Sample USW GU-3-769.1, scene B. C) Sample USW GU-3-877.6, scene B.

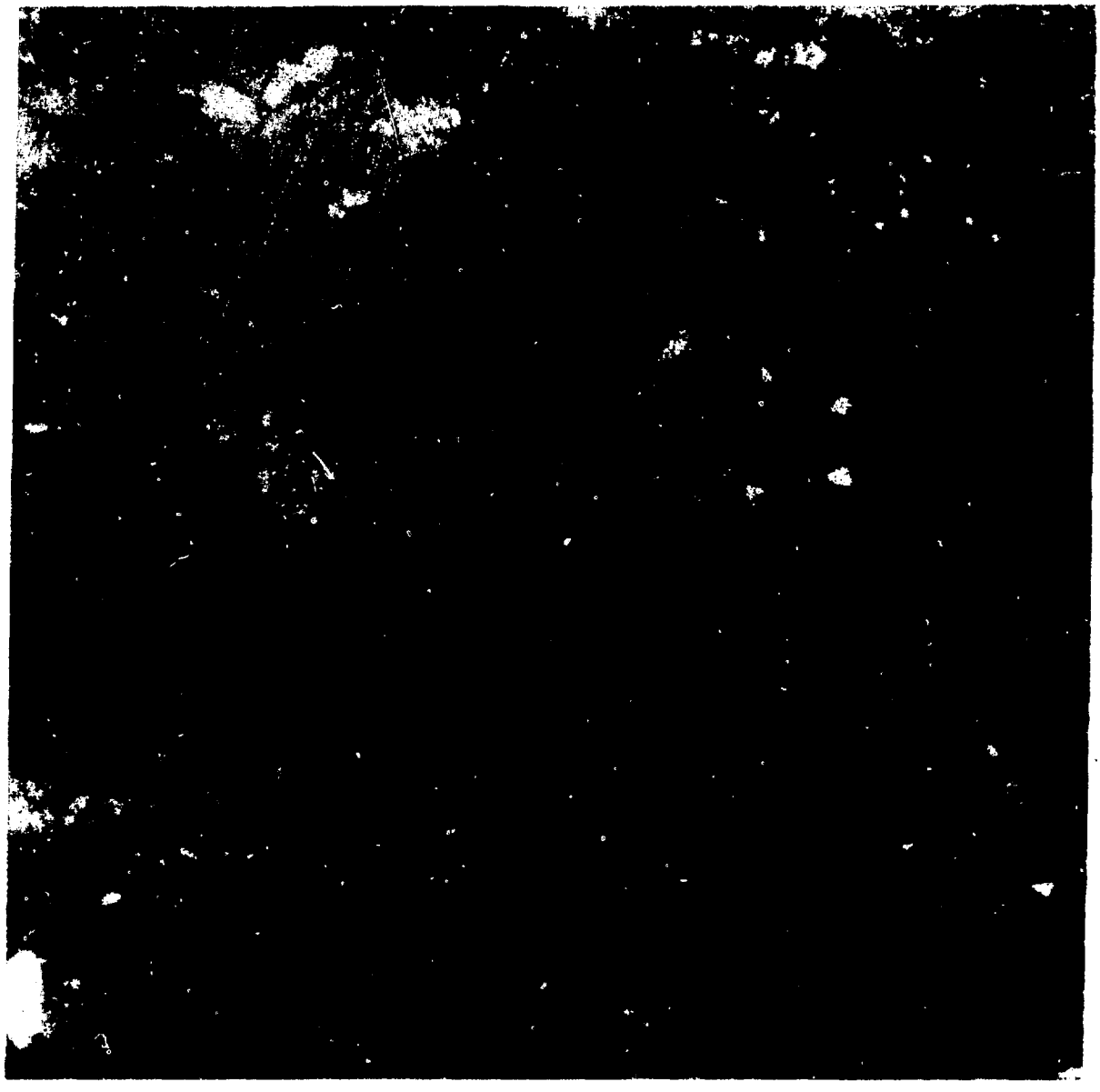

Figure 2B.

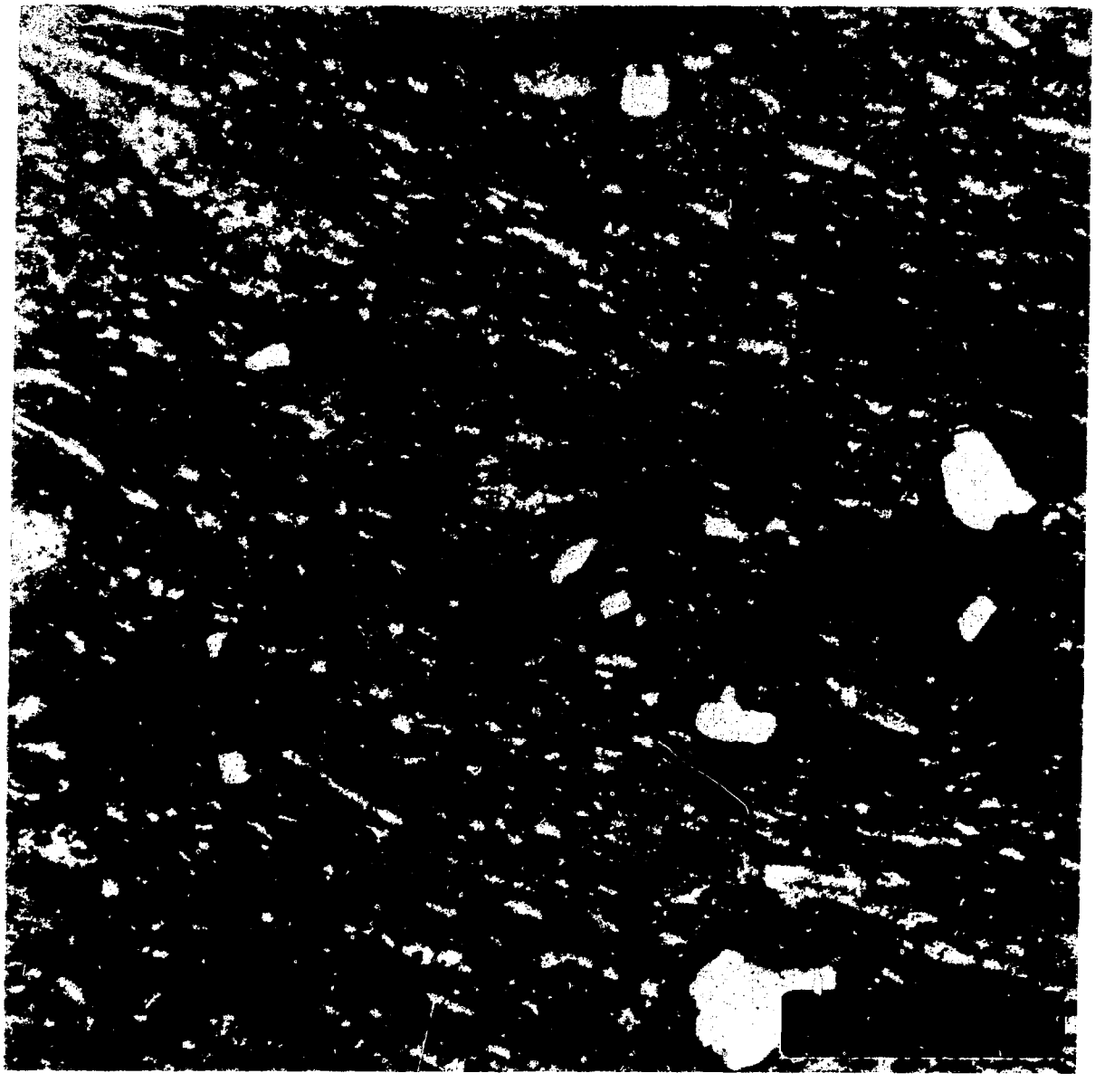

Figure 2C.

All pixels in a scene were plotted in a three-dimensional color cube in which each axis of the cube represents brightness in the red, green, and blue images. During data analysis, the color cube was reduced to a series of two-dimensional plots in which the covariance of color pairs (red vs blue, blue vs green, red vs green) was compared (Fig. 3). Each plot contains 4 194 304 data points, each representing color brightness values for individual pixels. Pixels associated with low brightness values near the origin of the plot and pixels with high brightness are represented by points in the upper right portion of the plots. The distribution and shape of data points in color space depends upon the overall color and brightness of the scenes.

We identified 11 visually distinct color components based upon observations with an optical microscope; these 11 components roughly correspond to textural components of the tuff identified by Byers (1985) and Byers and Moore (1987) (Table I). A "pick" file demarcating the coordinates for each of these color components was created by drawing polygons around the appropriate pixels in the two-color X-Y plots (Fig. 4). These picks were defined for the 11 chosen components by using sample G4-1236.2A as a training set, were archived to a file, and were subsequently recalled to identify the color components in the rest of the scenes. In most cases the original picks obtained from the training set were applied to the other scenes unmodified. In a few cases, however, small lobes or polygons were added to or deleted from the original picks to map adequately a particular component in a scene. These slight modifications, documented in a separate file, did not change the abundance of a color component by more than 1% of the value determined by the unmodified pick.

During this initial investigation, we attempted to match color components to textural features described by Byers (1985) and Byers and Moore (1987). No attempt was made to map all color components within a scene; instead, our objective was to examine variations in the 11 chosen color components. The percentage of image area covered by these color components ranges from 33 to 92%. Future work will investigate the variations in all color components regardless of whether they can be associated with identifiable textural features.

Objects within a scene were classified by color, and the boundaries of all objects of interest were converted into a series of vectors and their positions were mapped. The position and boundary data yield measurements of object area, object elongation, and object roughness. Areas of individual objects were measured and compiled in ASCII files for each of the 11 color components. Total area of a component was determined by summing the areas of all objects within a color component.

Elongation and roughness were also measured for each object within a scene in hopes that these variables would provide information about systematic changes in welding characteristics of individual pyroclasts within this thick, multiflow tuff cooling unit. Elongation is defined as the maximum length of an

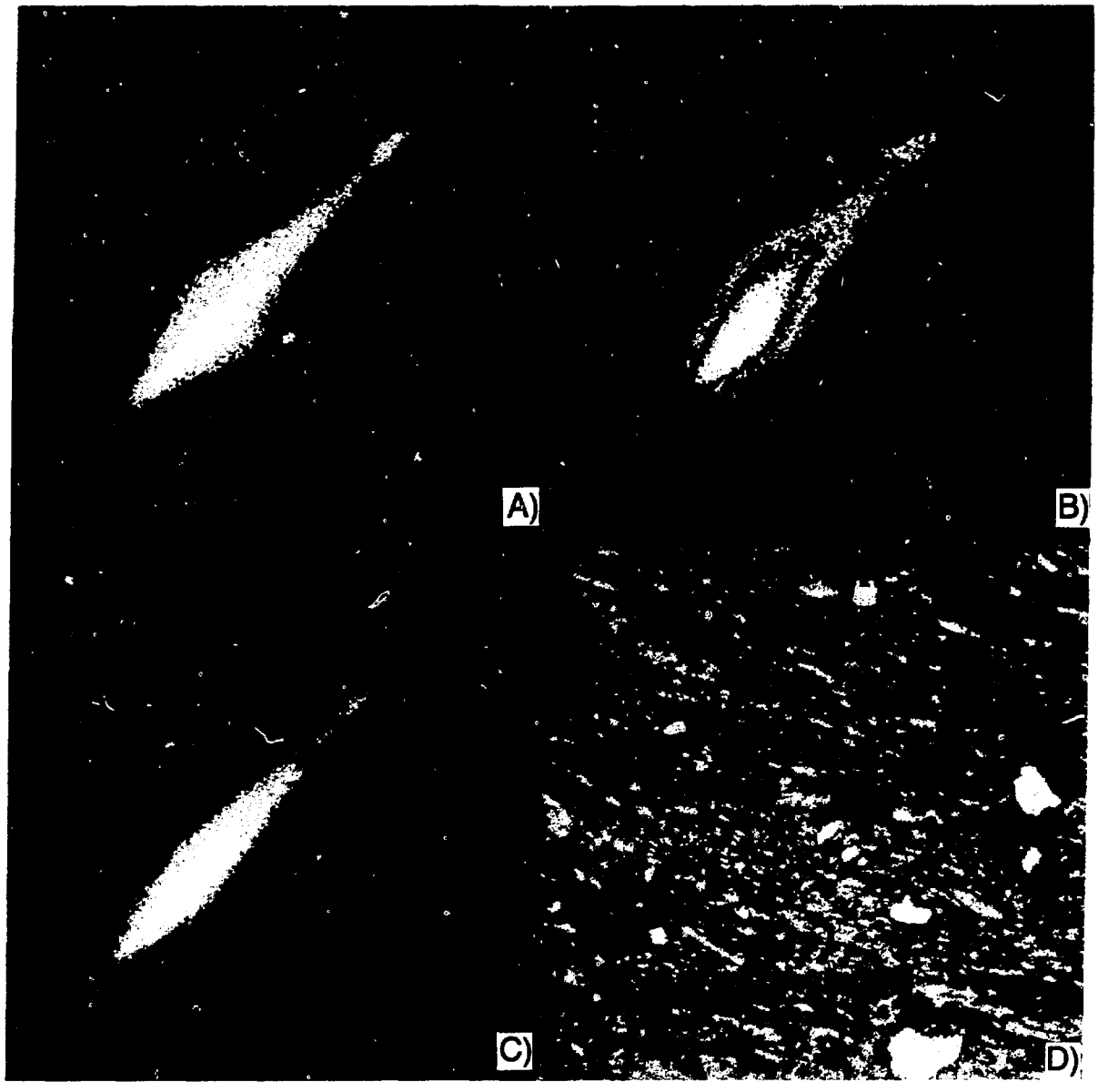

Figure 3. Two-dimensional representations of red, green, and blue color clouds for the digital image of sample USW GU-3-877.6. These X-Y plots represent faces of the three-dimensional color cube described in text. Colors plotted are shown on X and Y axes of plots. Color intensity increases to the right and upward for each plot: A) green vs blue color cloud; B) red vs blue color cloud; C) red vs green color cloud; D) full-color digital image represented by the color clouds.

TABLE I. DESCRIPTION OF COLOR COMPONENTS MAPPED IN THIS STUDY

Color Component	Color Symbol Used in Text, Tables, and Figures	Corresponding Textural Component
Dark red	1	Varieties of cryptocrystalline groundmass
Intermediate red	2	
Light red	2B	
Dark green	3	
Intermediate green	3B	
Dark gray	4	Varieties of microlitic/spherulitic groundmass
Intermediate gray	5	
Light gray	6	
White	7	Granophyric groundmass and felsic phenocrysts
Black	8	Opaque material including oxide microphenocrysts and dark groundmass material
Blue	9	Pore space highlighted by epoxy dyed blue

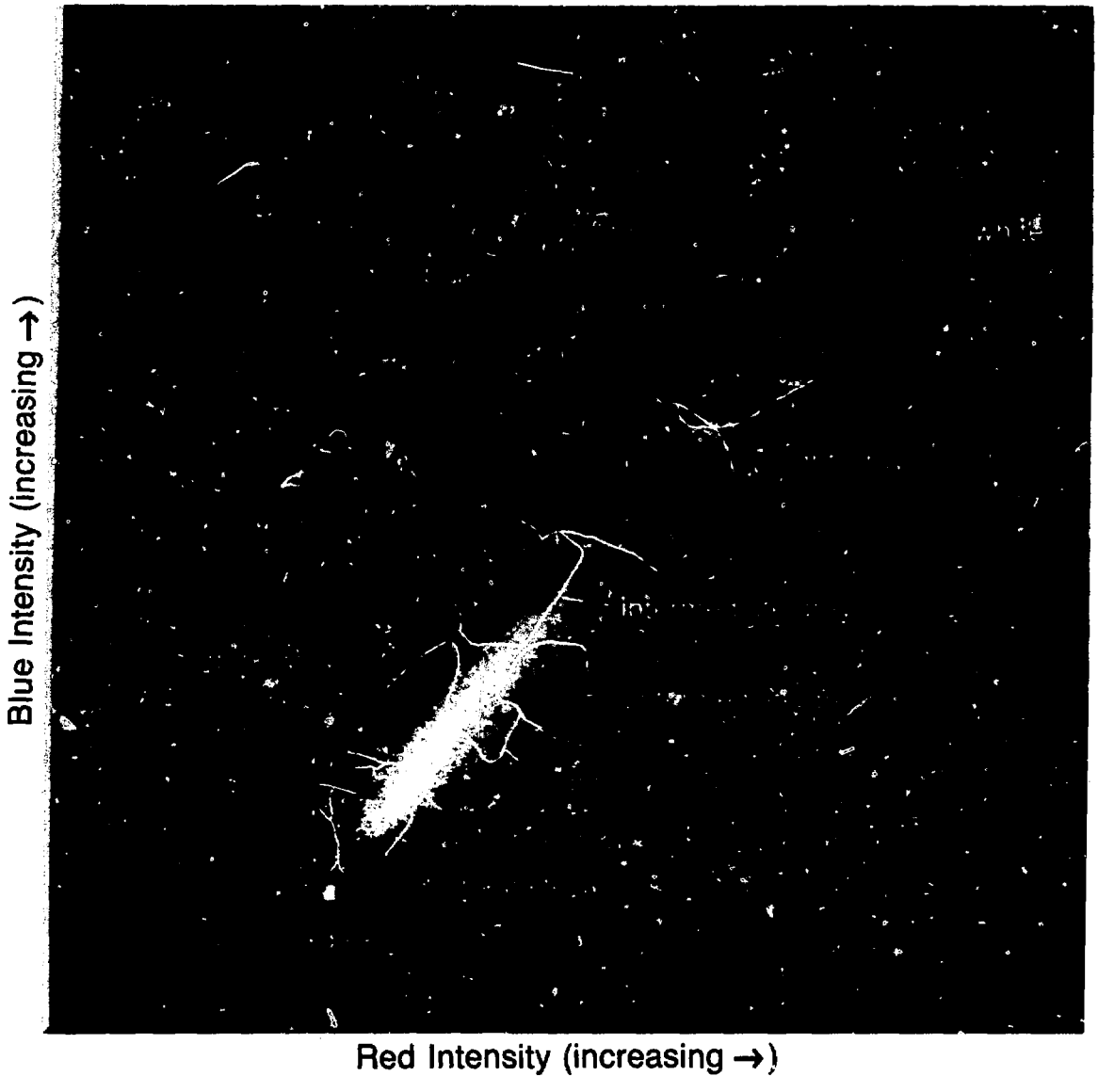

Figure 4. Example of red vs blue color cloud for sample GU3-769.1A (see Fig. 2A), showing the "picks" used to define color components mapped in this study.

object (in the plane of the thin section) divided by its throat; the throat of an object is the length of the shortest chord that connects points that are on opposite sides of the object. Elongation is measured from the external boundary of the object only. Internal objects or holes within the object of interest are ignored during the calculation.

Roughness is a measurement of how jagged an object boundary is. It is defined as the average absolute value second-order difference of the orientation angles of boundary segments of the object. The calculation for roughness is made by first listing the angles that each boundary segment makes with the horizontal. Differences between adjacent angles are calculated and these difference angles are subtracted from adjacent difference angles. The absolute value of angular differences determined from the last subtraction are summed and divided by the number of boundary segments making up the object to arrive at the average value of roughness. The roughness of a circle is zero because the difference angles will be the same and second differences will be zero. High values of roughness indicate many changes in the curvature of an object boundary. The more abrupt the changes in curvature are, the higher the value of roughness.

Other object descriptors have also been defined by Everest Technologies, but these three–area, elongation, and roughness–seem to be most closely related to traditional petrographic observations.

III. RESULTS

A. Summary of Available Data

A total of 60 scenes (2 per thin section for 15 thin sections from USW GU-3 and 15 from USW G-4) were analyzed. The area, elongation, and roughness of each object mapped in each of the 11 color components were reported. The number of objects in a given color component in a scene varies from 0 to over 10 000 and averages in the hundreds (see Table II); this size resulted in an enormous amount of data. For this preliminary study, we decided to work only with summary statistics for each scene and color component, such as

1. the number of objects;
2. their total, mean, median, minimum, and maximum areas;
3. their mean or median elongation and roughness; and
4. second-order statistics such as the standard deviation of object area, elongation, or roughness within a scene.

This use of summary statistics is a crude and nonoptimal way of working with these data, but even this much information generated a data base associating more than 150 "variables" with each of the 60 scenes. Therefore, this report must be considered preliminary. Future work will explore other ways of summarizing

TABLE II. NUMBERS OF OBJECTS IN COLOR COMPONENTS

	COLOR COMPONENT									
	CRYPTOCRYSTALLINE MATRIX					MICROLITIC/ SPHERULITIC			GRANOPHYRE/ PHENOCRYSTS	OPAQUES
SCENE	1	2	2B	3	3B	4	5	6	7	8
G3-735.5A	1096	3315	2524	4321	2389	1787	2171	673	92	7
G3-735.5B	105	158	6086	240	8979	4223	2243	2174	279	7
G3-769.0A	2615	8208	790	11221	1375	4585	1603	773	150	128
G3-769.0B	2344	6920	760	5499	447	4903	1341	510	84	47
G3-769.1A	133	2393	10858	5819	5505	7259	6417	910	211	12
G3-769.1B	333	3745	6985	7132	3927	6645	3487	732	147	9
G3-805.0A	38	.	418	.	.	523	4652	917	1426	3
G3-805.0B	.	.	3417	.	.	1832	2046	3073	279	.
G3-829.9A	981	2226	2903	2749	3331	1613	2975	879	164	17
G3-829.9B	5490	6783	890	10263	1569	5050	1048	269	29	73
G3-877.6A	88	154	586	212	2425	2336	5991	2925	378	14
G3-877.6B	97	66	.	.	.	5160	7312	3628	559	31
G3-911.3A	1356	2613	529	3078	653	5282	3968	1141	249	16
G3-911.3B	2426	3914	491	1902	548	7230	3379	588	142	4
G3-954.0A	648	952	3216	1150	1239	3322	1538	1538	307	17
G3-954.0B	137	293	1666	222	3923	2965	2504	1841	295	5
G3-954.8A	.	318	6226	2150	3918	2338	2781	1457	.	21
G3-954.8B	.	1441	1726	338	3113	516	7904	2375	709	5
G3-1019.7A	1848	4563	372	9268	479	3260	2203	565	53	53
G3-1019.7B	2134	3810	3328	3935	2257	4113	4231	1107	246	23
G3-1079.4A	2198	3744	.	.	.	3481	992	231	7	12
G3-1079.4B	.	3588	1991	6625	1788	3457	3464	1049	140	27
G3-1130.0A	2916	1973	1633	6097	699	3666	846	297	65	121
G3-1130.0B	4323	4497	3014	9615	1937	6664	1752	681	164	55
G3-1130.3A	370	7935	.	4310	4495	1531	1112	1736	74	10
G3-1130.3B	431	2298	8677	7842	5306	4290	2245	976	260	26
G3-1151.7A	1324	3690	7801	6904	4446	9369	2534	931	120	76
G3-1151.7B	670	2358	6424	4262	2882	5998	1992	342	43	14
G3-1175.0A	2177	1308	3890	11842	2124	7357	2075	609	89	72
G3-1175.0B	1012	3583	6601	8589	.	5349	1840	704	108	32
G4-684.8A	1389	7630	.	6187	.	3627	1363	547	57	31
G4-684.8B	1516	4268	4901	.	.	2866	2530	979	165	.
G4-694.0A	3376	3075	634	6187	1159	3883	1358	536	97	81
G4-694.0B	2746	2279	1045	5125	1003	2866	864	266	19	36
G4-698.8A	2137	3721	4400	5166	3028	3190	2524	1053	374	21
G4-698.8B	619	2146	10878	6222	6913	3597	3944	1030	173	5
G4-746.0A	3461	2330	37	81	88	1490	637	216	19	455
G4-746.0B	4930	1938	.	173	.	1094	354	82	6	9
G4-817.0A	171	284	6745	247	6904	3679	3581	3132	667	2
G4-817.0B	89	107	3034	49	3507	1641	5869	2332	622	8
G4-934.0A	2759	6458	159	5741	523	3852	1482	593	98	12
G4-934.0B	1865	3918	174	2903	439	2955	1085	483	65	40
G4-1026.0A	1877	3426	2705	7340	1790	5207	1742	915	254	25
G4-1026.0B	1801	6792	929	4967	983	3734	1528	708	78	41
G4-1089.0A	3182	4307	565	6230	722	4776	1895	835	118	40
G4-1089.0B	2809	3734	692	2806	978	3947	2682	659	121	5
G4-1117.0A	3325	7855	284	.	.	5162	1685	652	112	51
G4-1117.0B	2805	8254	129	2432	441	5172	1407	436	40	.
G4-1148.2A	.	.	2749	.	4705	2325	2877	3358	666	11
G4-1148.2B	127	416	11167	1068	4688	4631	4813	2298	603	14
G4-1170.0A	67	73	2675	26	3028	583	4164	2370	1051	19
G4-1170.0B	1190	1374	7468	3582	4511	4486	2746	1304	227	14
G4-1190.0A	3086	6750	312	4160	900	4264	2042	998	232	98
G4-1190.0B	3699	3981	1827	5923	658	4522	2735	1110	260	22
G4-1236.2A	1451	1590	.	4424	.	2137	1030	378	76	12
G4-1236.2B	528	3257	.	3414	.	2497	946	430	106	4
G4-1244.0A	1976	3730	299	3569	481	3040	1108	451	110	42
G4-1244.0B	1930	5816	597	6385	513	3566	1194	438	71	61
G4-1282.0A	2663	3777	113	1386	353	1832	974	355	80	969
G4-1282.0B	4649	6172	392	3510	646	2777	1361	648	140	225

the data; in particular, the use of only the larger objects in some color components might correspond more closely to the traditional measurements made by a petrographer.

Of the 11 color components measured, 1, a blue component corresponding to pore space, produced data (i.e., had at least 1 object) in only a couple of scenes and thus was not useful. Occasional scenes had no objects in some of the other color components, as shown in Table II. Most often these gaps fall in the color components that correspond to elements of the cryptocrystalline matrix in the principal training set G4-1236.2A. (The training set itself had no objects in components 2B or 3B.) For these cases the only available summary statistics are the number of objects and their total area (both zero), a fact with considerable impact later in this study when multivariate techniques that discard "incomplete" observations are used.

B. Area Measurements

The variables that appear to correspond most closely to the percent areas of textural types, traditionally measured by modal petrographic analysis, are the total areas of the objects in each color component (Table III). Apparent trends in some modal count observations with stratigraphic position are reported by Byers (1985). For example, Byers found that "[t]he percentage of cryptocrystalline groundmass is inversely related to stratigraphic position above the chilled base" (Byers 1985, p. 21), and "[i]ncreasing granophyre [upwards]...suggest[s] stratigraphic position within the upper part of the lower lithophysal zone" (Byers 1985, p. 22). These and other favorable observations have suggested that textural variations might provide a useful index to stratigraphic position.

The corresponding image analysis variables, however, appear to have little stratigraphic discriminatory power, either between or within zones. For example, the sum of color components 1, 2, 2B, 3, and 3B, which correspond to elements of the cryptocystalline groundmass, is plotted in Fig. 5. This plot shows considerable scatter and no detectable trends with depth. In some cases, the difference between two scenes on the same slide is large (for example, the two scenes from slide G3-877.6 near the bottom of the middle nonlithophysal zone). For a second example, the total area in color component 7, corresponding to granophyre and phenocrysts, is plotted in Fig. 6 (on a logarithmic scale, to minimize the visual impact of a rather large observation for G4-911.3 at the top of the lower lithophysal zone), and here a trend with depth in the lower lithophysal zone is certainly suggested, but the variability is considerable as well.

It has been suggested that if the color bands defining the color components are too narrow, then possibly more consistent results could be obtained by exploiting the correlations among them. The correlation matrix of the 10 areas is shown in Table IV, where the upper number of each pair is the Pearson

TABLE III. AREA COVERED BY COLOR COMPONENTS

SCENE	1	2	2B	3	3B	4	5	6	7	8	TOTAL
G3-735.5A	6.32	9.72	1.38	3.53	1.94	46.73	11.00	2.62	0.87	0.06	84.15
G3-735.5B	0.12	0.12	4.23	0.37	7.70	15.70	47.58	9.38	0.70	0.05	85.96
G3-769.0A	0.50	14.57	0.29	10.32	0.83	21.17	4.92	1.75	0.89	0.25	85.50
G3-769.0B	3.07	16.88	0.30	3.77	0.17	16.43	2.57	1.25	0.51	0.24	75.17
G3-769.1A	0.17	5.74	10.88	8.36	16.72	16.10	3.93	2.77	1.21	0.03	65.92
G3-769.1B	0.37	21.55	7.26	18.00	7.21	22.75	3.05	2.42	0.79	0.07	83.48
G3-805.0A	0.12	0.0	0.62	0.0	0.0	0.43	25.05	61.42	3.39	0.04	91.07
G3-805.0B	0.0	0.0	2.36	0.0	0.0	3.40	51.11	20.56	2.52	0.0	79.94
G3-829.9A	1.50	4.77	1.27	7.57	2.12	48.43	14.80	3.93	1.04	0.07	85.50
G3-829.9B	6.25	22.30	0.39	11.23	0.73	23.45	2.06	0.55	0.59	0.14	77.69
G3-877.6A	0.08	0.13	0.25	0.10	1.69	36.16	25.46	7.15	1.91	0.09	73.01
G3-877.6B	0.09	0.07	0.0	0.0	0.0	25.68	25.22	9.48	3.50	0.05	64.08
G3-911.3A	2.78	3.10	0.18	2.48	0.27	39.12	11.19	2.36	1.14	0.54	63.14
G3-911.3B	2.15	9.54	0.17	1.32	0.24	27.22	4.76	1.25	0.47	0.01	57.13
G3-954.0A	1.38	1.01	2.29	0.77	0.59	34.82	7.73	7.73	1.92	0.55	58.80
G3-954.0B	0.22	0.53	0.98	0.27	3.78	28.56	40.45	7.99	2.57	0.08	85.43
G3-954.8A	0.0	0.52	20.02	6.61	27.76	1.28	2.05	3.47	0.0	0.17	61.89
G3-954.8B	0.0	2.84	1.76	0.47	17.67	0.43	5.61	3.07	2.60	0.06	34.51
G3-1019.7A	2.09	10.74	0.14	10.85	0.20	38.43	7.09	1.67	1.35	0.36	82.93
G3-1019.7B	5.79	7.39	1.80	3.56	1.18	35.97	10.58	2.76	0.87	0.22	70.11
G3-1079.4A	8.25	8.61	0.0	0.0	0.0	42.49	1.52	0.30	0.13	0.08	61.38
G3-1079.4B	0.0	5.34	1.11	5.92	1.11	40.68	7.84	1.78	2.35	0.27	66.40
G3-1130.0A	0.60	39.44	1.36	10.64	0.51	11.14	1.83	1.40	1.08	0.74	88.74
G3-1130.0B	0.70	32.11	2.10	10.25	1.73	11.50	2.62	2.02	0.82	0.07	73.93
G3-1130.3A	0.51	16.38	0.0	7.53	18.02	0.83	0.81	6.15	2.18	0.02	52.43
G3-1130.3B	0.65	3.99	15.75	8.67	16.69	2.05	1.59	4.95	2.80	0.24	57.39
G3-1151.7A	1.81	7.83	10.47	16.37	7.51	18.85	3.17	1.81	0.70	0.39	68.90
G3-1151.7B	1.62	14.85	12.02	18.58	7.06	20.28	6.33	1.19	0.75	0.15	82.81
G3-1175.0A	4.74	15.88	4.06	15.27	1.99	10.10	1.90	2.40	2.54	0.36	59.22
G3-1175.0B	1.58	10.58	7.30	9.56	0.0	5.14	2.15	2.78	0.97	0.14	40.18
G4-684.8A	1.79	8.32	0.0	2.98	0.0	18.00	4.01	1.50	0.78	0.04	77.43
G4-684.8B	0.04	12.74	2.78	0.0	0.0	30.78	12.29	3.92	1.42	0.0	73.96
G4-694.0A	2.61	24.21	0.21	7.42	0.79	13.51	2.98	1.62	0.88	0.16	64.35
G4-694.0B	5.67	36.36	0.55	12.37	0.70	10.43	1.66	0.73	0.38	0.09	78.93
G4-698.8A	9.70	13.86	4.06	6.91	2.73	26.95	13.19	5.99	1.44	0.26	85.09
G4-698.8B	0.76	4.24	10.54	9.21	6.46	34.27	12.13	3.79	1.07	0.03	82.51
G4-746.0A	0.88	1.76	0.01	0.02	0.03	3.49	1.30	0.47	0.51	0.42	38.88
G4-746.0B	3.73	1.28	0.0	0.06	0.0	1.19	0.45	0.23	0.32	0.02	27.29
G4-817.0A	0.18	0.37	10.07	0.15	8.80	21.38	37.66	11.10	4.53	0.07	94.32
G4-817.0B	0.24	0.08	9.82	0.02	4.97	4.25	27.55	16.28	10.81	0.36	74.39
G4-934.0A	4.50	22.82	0.05	8.33	0.20	9.04	3.22	1.44	0.39	0.21	80.19
G4-934.0B	6.36	25.56	0.06	2.21	0.19	5.80	1.37	1.06	2.44	0.09	85.13
G4-1026.0A	8.46	19.21	1.68	12.30	1.12	20.04	7.58	3.25	1.58	0.25	75.48
G4-1026.0B	0.80	27.14	0.45	6.08	0.53	14.25	5.38	1.91	0.92	0.25	77.70
G4-1089.0A	2.10	26.20	0.19	12.18	0.31	17.53	4.04	1.66	0.58	0.48	85.26
G4-1089.0B	0.68	11.26	0.29	2.86	0.50	34.21	4.42	1.28	0.51	0.02	66.04
G4-1117.0A	7.16	8.42	0.10	0.0	0.0	28.84	2.41	1.11	0.65	0.10	58.79
G4-1117.0B	1.60	12.89	0.04	1.05	0.17	17.05	1.78	0.73	1.09	0.0	66.40
G4-1148.2A	0.0	0.0	7.34	0.0	5.52	4.85	79.24	14.33	3.14	0.04	114.45
G4-1148.2B	0.12	0.41	10.05	0.92	4.03	28.12	27.24	5.98	3.56	0.14	80.56
G4-1170.0A	0.13	0.04	5.25	0.01	13.14	0.60	32.32	19.79	7.02	0.18	78.47
G4-1170.0B	0.54	2.97	21.85	7.18	6.13	27.63	9.63	3.34	2.03	0.15	81.47
G4-1190.0A	1.49	22.60	0.11	5.36	0.36	10.00	2.19	2.02	0.99	0.20	75.32
G4-1190.0B	5.41	25.77	0.95	8.45	0.24	18.91	3.92	2.45	1.59	0.25	77.95
G4-1236.2A	0.16	37.15	0.0	13.21	0.0	21.54	4.66	2.15	1.58	0.09	90.55
G4-1236.2B	1.27	21.22	0.0	25.30	0.0	29.74	5.71	3.27	3.55	0.02	90.09
G4-1244.0A	9.28	28.10	0.15	5.60	0.19	8.68	2.62	1.95	1.04	0.81	78.40
G4-1244.0B	1.14	21.75	0.32	9.23	0.23	8.78	1.87	0.92	0.59	0.49	85.33
G4-1282.0A	8.45	15.07	0.06	1.48	0.13	3.88	1.61	1.22	1.81	2.34	66.06
G4-1282.0B	7.89	13.97	0.22	3.39	0.34	4.81	2.76	1.56	1.70	0.63	57.26

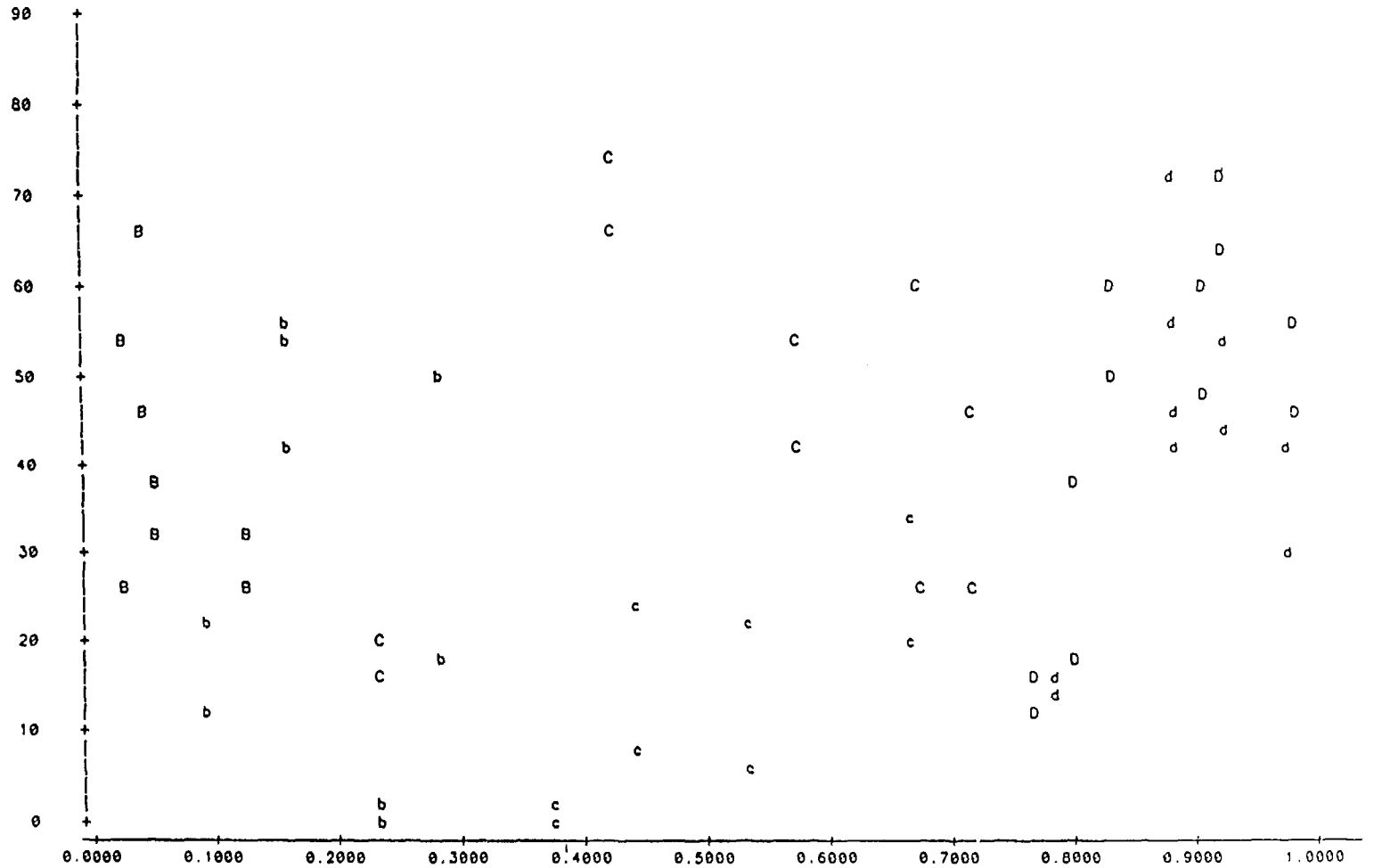

Figure 5. Plot of AREA 1 + AREA 2 + AREA 2B + AREA 3 + AREA 3B as a function of relative depth. Lower-case letters represent observations from USW G-3, upper-case from USW G-4. Observations labeled b or B come from the middle nonlithophysal zone; those labeled c or C come from the lower lithophysal zone; and those labeled d or D come from the lower nonlithophysal zone.

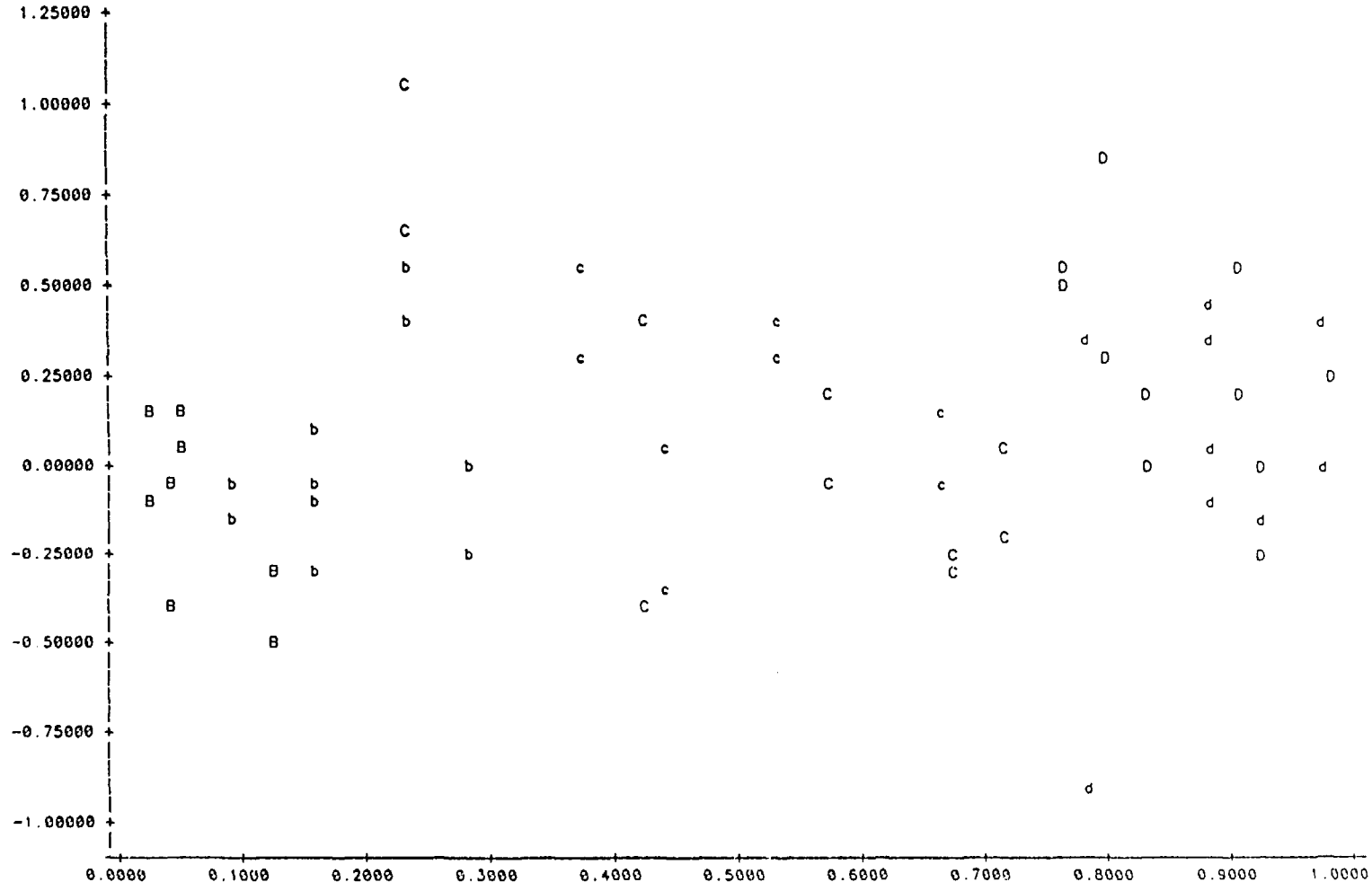

Figure 6. Plot of log10 (AREA 7) as a function of relative depth. Lower-case letters represent observations from USW G-3, upper-case from USW G-4. Observations labeled b or B come from the middle nonlithophysal zone; those labeled c or C come from the lower lithophysal zone; and those labeled d or D come from the lower nonlithophysal zone.

TABLE IV. CORRELATION MATRIX FOR COLOR COMPONENT AREAS

			AREA1	AREA2	AREA2B	AREA3	AREA3B	AREA4	AREA5	AREA6	AREA7	AREA8
RED:	DARK	AREA1	1.00000	0.48750	-0.48161	-0.05863	-0.45344	-0.27117	-0.44522	-0.34758	-0.34896	0.37233
			0.0000	0.0001	0.0001	0.6563	0.0003	0.0361	0.0004	0.0065	0.0063	0.0034
	INTERMEDIATE	AREA2	0.48750	1.00000	-0.36649	0.56571	-0.33220	-0.13331	-0.52070	-0.38382	-0.34860	0.17905
			0.0001	0.0000	0.0040	0.0001	0.0095	0.3099	0.0001	0.0025	0.0063	0.1710
	LIGHT	AREA2B	-0.48161	-0.36649	1.00000	0.15006	0.66023	-0.11473	0.15965	0.06424	0.21929	-0.11548
			0.0001	0.0040	0.0000	0.2525	0.0001	0.3827	0.2230	0.6258	0.0923	0.3796
GREEN:	DARK	AREA3	-0.05863	0.56571	0.15006	1.00000	0.03851	0.04998	-0.42113	-0.30726	-0.23247	-0.00420
			0.6563	0.0001	0.2525	0.0000	0.7702	0.7045	0.0008	0.0169	0.0739	0.9746
	INTERMEDIATE	AREA3B	-0.45344	-0.33220	0.66023	0.03851	1.00000	-0.30981	0.08640	0.06590	0.18207	-0.16134
			0.0003	0.0095	0.0001	0.7702	0.0000	0.0160	0.5116	0.6169	0.1638	0.2181
GRAY:	DARK	AREA4	-0.27117	-0.13331	-0.11473	0.04998	-0.30981	1.00000	-0.01844	-0.24921	-0.21328	-0.18199
			0.0361	0.3099	0.3827	0.7045	0.0160	0.0000	0.8888	0.0548	0.1018	0.1640
	INTERMEDIATE	AREA5	-0.44522	-0.52070	0.15965	-0.42113	0.08640	-0.01844	1.00000	0.52541	0.47041	-0.21255
			0.0004	0.0001	0.2230	0.0008	0.5116	0.8888	0.0000	0.0001	0.0001	0.1030
	LIGHT	AREA6	-0.34758	-0.38382	0.06424	-0.30726	0.06590	-0.24921	0.52541	1.00000	0.48259	-0.14236
			0.0065	0.0025	0.6258	0.0169	0.6169	0.0548	0.0001	0.0000	0.0001	0.2779
WHITE		AREA7	-0.34896	-0.34860	0.21929	-0.23247	0.18207	-0.21328	0.47041	0.48259	1.00000	-0.00079
			0.0063	0.0063	0.0923	0.0739	0.1638	0.1018	0.0001	0.0001	0.0000	0.9952
BLACK		AREA8	0.37233	0.17905	-0.11548	-0.00420	-0.16134	-0.18199	-0.21255	-0.14236	-0.00079	1.00000
			0.0034	0.1710	0.3796	0.9746	0.2181	0.1640	0.1030	0.2779	0.9952	0.0000

moment correlation estimate and the lower number is the probability that this statistic would exceed this value (in magnitude) under the null hypothesis that the true correlation is zero. Clearly, many of the color components are significantly correlated. Factor analysis suggests that not more than 4 or 5 independent factors underlie these 10 variables and tends to group together color components 5, 6, and 7 (2 of the 3 components for microlitic/spherulitic elements and the component corresponding to granophyre and phenocrysts); color components 2B and 3B (2 of the 5 components for the cryptocrystalline matrix); and color components 2 and 3 (also corresponding to the cryptocrystalline matrix). These are not particularly intuitive groupings, and plots of these factors suggest that they are no less variable than the original areas.

These and other similar preliminary examinations of the total area variables fail to bear out the promising results obtained by Byers (1985). One probable cause of this disappointing result is the use of a single scene as a "training set" to select the color components. Apparently there is more variability within a single petrographic texture than could be captured with this limited training set. Second, the variability between scenes on the same slide, although generally significantly smaller than between-slide variability, suggests that perhaps areas of this size (1 cm on a side) are too small to be representative; the areas scanned in modal count petrography are generally five or six times as large. (Additional support for this observation will be found below, where better classification results were obtained by working with the pairs observations than with the single scenes.) Finally, it is possible that only the larger objects correspond to what a traditional petrographer observes, and working with the statistics of objects from only the upper tail of the size distribution might come closer to matching the results of traditional analyses.

C. Other Measurements

Area measurements. The largest object in a given color component can account for 80 or 90% of the area covered by that color component, even when there are more than 1000 objects in the scene. The ratio of the mean object area to the median object area is typically between two and four. The estimated skewness is always large. These are all indicators that the distributions of object areas within a scene are far from normal. For such asymmetric distributions, moment-based estimators are highly unreliable, and the distribution of such a statistic–for example, the mean area of objects in color component 4–is likewise far from normally distributed. It is preferable, in the statistical analyses carried out below, to use variables that are more nearly normal.

For area measurements, more normal statistics can be derived by assuming that the areas of the objects in a given scene are lognormally distributed. Thus, if X_j is the area of the j^{th} object in a scene, and

$Y_j = \log(X_j)$, then we assume that the Y_j come from a normal distribution with mean μ_A^* and variance σ_A^{*2}, which are related to the mean μ_A and variance σ_A^2 of the distribution of the X_j by

$$\mu_A = \exp\left(\mu_A^* + \frac{\sigma_A^{*2}}{2}\right)$$

$$\sigma_A^2 = \mu_A^2[\exp(\sigma_A^{*2}) - 1] \ .$$

Thus the parameters μ_A^* and σ_A^{*2} are computable as functions of the variables in the data base. Some validation of the assumption that the X_j are in fact lognormally distributed is provided by the observation that $\mu_A^*/\log(m_A)$, where m_A is the median X_j, is close to one for all color components and scenes, varying both below and above one.

Therefore, we consider the following variables related to area:

Ni	the number of objects in color component i of the scene;
AREAi	the total area of color component i;
LGMEANi	the logarithmic mean of object areas in color component i;
LGMEDi	the logarithmic median of object areas in color component i;
RATIOi	the ratio of logarithmic mean to logarithmic median in color component i; and
LGSDi	the logarithmic standard deviation of object areas (i.e., σ_A^*) in color component i.

Elongation measurements. The elongation of an object is a number greater than or equal to 1 and ranges up to about 50 in these data. In general, the statistics for elongation are well behaved; the distributions do not appear to be contaminated by a few outliers. However, the elongations of objects within a scene are almost certainly not normally distributed. They might better be modeled by a shifted (so that the minimum value is zero) gamma distribution. The parameters α and β of the shifted gamma distribution are related to the mean μ_E and variance σ_E^2 of the elongation distribution by

$$\beta = \frac{\sigma_E^2}{\mu_E - 1}$$

$$\alpha = \frac{\mu_E - 1}{\beta} \ .$$

Thus again these parameters are computable as functions of the variables in the data base. α is a "shape" parameter: $\alpha = 1$ corresponds to an exponential distribution, whereas $\alpha < 1$ for a distribution that is more concentrated at the lower end (i.e., near 1 in this case) than an exponential, and $\alpha > 1$ for a distribution with mode > 1. β is a scale parameter. Estimates of α for these distributions are generally less than 1, but for some color components they become quite large (up to 8 or 9). The scale parameter β is generally between 1 and 10 but occasionally as large as 20.

We consider the following variables related to elongation:

EMEANi	the mean elongation of objects in a scene in color component i;
EMEDi	the median elongation of objects;
ESDi	the standard deviation of elongation of objects;
ALPHAi	the estimated shape parameter α;
BETAi	the estimated scale parameter β; and
ECVi	the shifted coefficient of variation $\frac{\text{ESDi}}{\text{EMEANi - 1}}$.

Roughness measurements. Roughness ranges generally between 70 and 450, with an average value between 150 and 200. In general, the statistics for roughness are well behaved (i.e., look reasonably normal), and we consider

RMEANi	the mean roughness of objects in scene from color component i;
RMEDi	the median roughness of objects;
RSDi	the standard deviation of roughness of objects; and
RCVi	the coefficient of variation $\frac{\text{RSDi}}{\text{RMEANi}}$.

Thus we have defined 16 variables to be associated with each color component of each scene, for a total of 160 variables per observation. The goal of the remaining work described in this report was to see whether multivariate approaches using subsets of these 160 variables might be more successful than the simple univariate analogies to ordinary modal count variables described in the preceding section.

D. Classification by Zone

Univariate analyses of the variance of each of the 160 variables listed for consideration above were undertaken with the aim of discovering which of these variables have potential for distinguishing among the middle nonlithophysal, lower lithophysal, and lower nonlithophysal zones. The model included a hole effect to determine whether there were significant differences between the two holes as well as a zone

effect and a term for interaction between hole and zone. The significance of these effects was tested 1) against the error term for the model containing only these effects and 2) against the between-slides sum of squares in a model with an additional term for the effect of slides nested within hole and within zone. The second significance test was generally more stringent. With the first model, it was also possible to test the significance of the contrasts between the lithophysal and nonlithophysal zones and between the middle and lower zones.

Table V summarizes the results of the variance tests, showing only those variables for which zone was a significant main effect in the first model. An "x" in a column denotes significance (10% level) of the effect named at the top of the column: MODEL for the first model overall, Z for zone, H for hole and Z*H for their interaction in each of the two models, plus the two contrasts (LITH for lithophysal vs nonlithophysal zones, LEVEL for middle vs lower zones).

The 51 variables of Table V were offered to a stepwise discriminant procedure, which selects subsets of these variables on which to base a discriminant function for classifying observations into 1 of the 3 zones. The selections made by the procedure were cross-validated by repeating the experiment 24 times, each time with a different random subset of about 80% of the data (the number of observations available for any 1 run varied between 26 and 37 of the 43 that are complete with respect to all of the variables in Table V). Seven of the 51 variables were selected in 8 or more trials, and 40 were selected by 2 or fewer (out of a possible 24).

The most promising variables for discrimination appear to be

in color component 2B:	RMEAN and RMED, the average and median roughness;
in color component 4:	LGMEAN and LGMED, the logarithmic mean and median area;
in color component 5:	N, the number of objects,
	LGMEAN and LGMED, the logarithmic mean and median area,
	EMEAN, the average elongation;
in color component 6:	EMEAN and EMED, the average and mean elongation; and
in color component 8:	LGMEAN, the logarithmic mean area.

(There is competition between paired means and medians for a position in the model, a fact that has been taken into consideration in evaluating the results of the stepwise discriminant procedure.) Thus of the five color components associated with the cryptocrystalline matrix (1, 2, 2B, 3, and 3B), only one produces any significant discriminant, and this discriminant is associated with the roughness of its objects, perhaps a reasonable result for a "groundmass" component. Color components 4, 5, and 6 are associated with the microlitic / spherulitic elements, and both the shape (elongation) and the size of the individual objects

TABLE V. SIGNIFICANT EFFECTS IN ANALYSES OF VARIANCE

		MODEL 1			MODEL 2			CONTRASTS	
VARIABLE	MODEL	Z	H	Z*H	Z	H	Z*H	LITH	LEVEL
N3	x	x		x				x	
N5	x	x	x	x		x		x	
N6	x	x		x				x	
AREA3	x	x						x	
AREA8		x							x
LGMEAN3		x	x		x			x	
LGMEAN4	x	x			x			x	x
LGMEAN5	x	x			x			x	x
LGMEAN7	x	x		x	(x)		(x)	x	
LGMEAN8	x	x			x			x	x
LGMED2B		x							
LGMED3	x	x	x					x	
LGMED4	x	x	x			x			x
LGMED5	x	x	x		x				x
LGMED8		x			x			x	x
RATIO4		x						x	x
LGSD2B		x							
LGSD3		x							x
EMEAN2B		x							
EMEAN3	x	x			x			x	
EMEAN5	x	x			x				x
EMEAN6	x	x	x		x	x			x
EMEAN7	x	x		x	x		x		x
EMED1		x			x			x	
EMED3	x	x			x			x	
EMED5		x							x
EMED6	x	x	x		x	x			x
EMED7	x	x			x		x		x
ESD2B		x							
ESD3		x			x			x	
ESD7	x	x			x				x
ECV2B		x							x
ECV7		x			x				
ALPHA2B		x			x				x
BETA2B		x							
BETA3		x			x			x	
BETA7	x	x			(x)				
RMEAN2B	x	x			x			x	
RMEAN3	x	x			x			x	
(RMEAN4)	x	x							x
RMED2		x						x	
RMED2B	x	x			x			x	
RMED3	x	x	(x)		x			x	
RMED4	x	x			(x)				x
RSD5	x	x			x				x
RSD6		x							
RCV2B	x	x			x			x	
RCV3	x	x						x	
RCV3B	x	x		x	x			x	
RCV4		x							x
RCV6		x			x				

provide some usable discriminants as well as the number of objects in color component 5. Color component 8 is associated with opaques, and the size of opaque objects is the most consistently selected variable.

In general, the median is to be preferred in the mean/median pairs. There is evidence that the logarithmic mean area, in particular, is distorted by some low outliers in the distribution (with very large negative logarithmic areas). Mean elongation is consistently larger than median elongation, as expected for a gamma-shaped distribution with small shape parameter, but not significantly worse behaved. Thus we will use the median in each of the four paired cases above.

Seven variables–RMED2B, LGMED4, N5, LGMED5, EMEAN5, EMED6, and LGMEAN8–were used to form a discriminant function. Again the procedure was cross-validated by using about 80% of the data to construct the discriminant function and then using the result to classify the remaining samples. Observations were reserved for the test sample in pairs when there are two observations (i.e., scenes) per slide. (Recall that the multivariate discriminant technique discards "incomplete" observations, which here means those scenes with no objects in color component 2B, those scenes with no objects in color component 8, or both scenes.) This was done because the variation between scenes within slides is often small, and the presence of one scene in a calibration set would probably increase the likelihood of the other scene's being correctly classified by the resulting discriminant function, whereas in practice a whole slide would have to be classified on the basis of one or more scenes. This procedure was repeated 16 times with randomly selected calibration and test sets.

There are 47 observations (out of 60) that are complete with respect to the 7 variables used in this part of the exercise. The randomly selected calibration subsets contained between 33 and 42 of these observations, and misclassification rates within the calibration subsets averaged 15% and did not exceed 20%. In the test sets of size 5 to 14, on the other hand, misclassification rates for scenes ranged from 11% all the way up to 71% and averaged 40%. However, if slides for which both scenes are complete (20 of the 30 slides) are classified more realistically by using the classification of the scene with the highest posterior probability, there appear to be about four troublemakers: slides G3-769.1, G3-954.8, G4-1148.2, and G4-1190.0 are misclassified every time they are placed in the test set. (The two G-4 slides are also frequently misclassified even when they are included in the calibration set.) Thus the base misclassification rate for slides appears to be about 20%.

In addition, three unpaired scenes, G3-877.6A, G3-1079.4B, and G4-1117.0A, have serious problems. It is interesting that all three of these scenes, as well as G4-1148.2, come from within a few tens of feet of a boundary between zones. G3-877.6A, in particular, is persistently classified as middle

nonlithophysal, where it would be assigned by Scott and Castellanos (1984), although Byers and Moore's (1987) assignment to the lower lithophysal was used here.

E. Depth Discrimination within the Lower Nonlithophysal Zone

The variables were screened for this part of the exercise by computing the sample correlation of each variable with depth within the lower nonlithophysal zone. Variables with the most significant correlations are shown in Table VI. Of these, the most consistently selected by stepwise regression performed on subsets of the lower nonlithophysal observations are LGSD8, AREA4, RATIO6, and RATIO7. These are plotted vs relative depth in Figs. 7 through 11. A regression model constructed with the 21 observations that are complete for these 4 variables explained about 75% of the variance with depth. The residuals, plotted in Fig. 10, suggest that depth cannot be predicted to within better than about 25% within the unit using this model. Scene G4-1148.2A is an influential outlier in this model and probably should be eliminated.

IV. CONCLUSIONS

In general, the results of attempts to relate the computed variables to stratigraphic position failed to produce a useful stratigraphic indicator, despite the promise of the results Byers (1985) obtained using standard (and possibly more subjective; see Moore et al. 1989) textural analysis measurements. However, the simplifications made in this preliminary application of image analysis, especially the use of a very limited training set for determining the interesting color components and the restriction of the choice of variables to statistics summarizing the object distributions in each component, may be largely responsible for these preliminary results. The technique deserves further investigation.

TABLE VI. VARIABLES CORRELATED WITH DEPTH IN LOWER NONLITHOPHYSAL ZONE

VARIABLE	CORRELATION	P-VALUE
N3	.49	.0221
N5	-.61	.0025
N6	-.63	.0018
N7	-.56	.0067
AREA3	.48	.0254
AREA4	-.45	.0336
AREA5	-.54	.0088
AREA6	-.45	.0346
AREA8	.46	.0333
LGMEAN6	.62	.0024
LGMEAN7	.47	.0290
LGMEAN8	-.50	.0211
LGSD8	.64	.0019
RATIO6	-.58	.0043
RATIO7	-.57	.0059
RATIO8	.53	.0132
EMED3	.52	.0192
EMEAN5	-.45	.0340
RSD2B	.53	.0248
RSD3	.51	.0218
RSD3B	.52	.0283
RSD5	.49	.0212
RSD6	.63	.0016
RCV6	.66	.0007

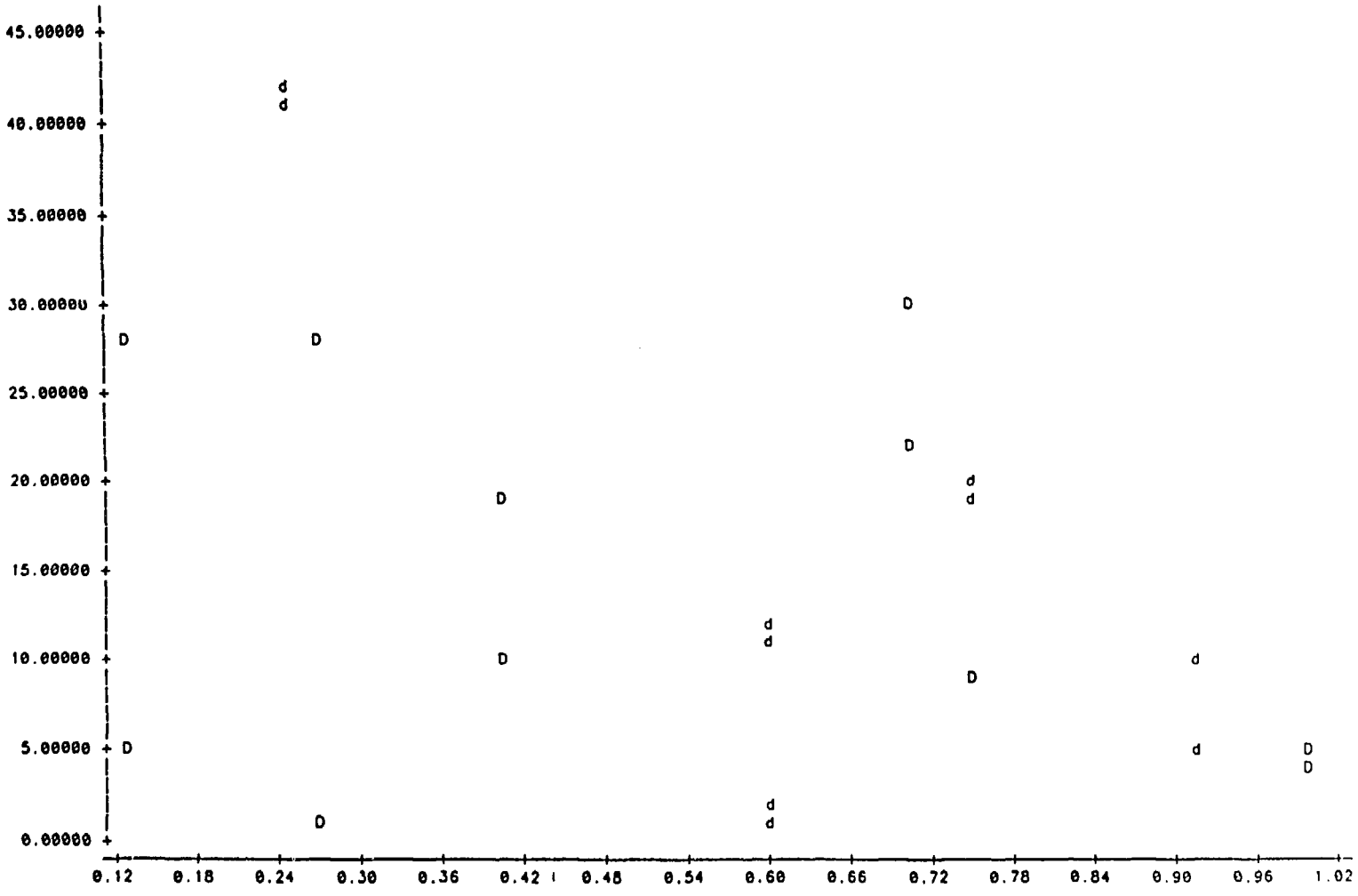

Figure 7. Plot of AREA4 as a function of relative depth in the lower nonlithophysal zone. Points plotted using lower-case letters come from USW G-3; those plotted with upper-case letters come from USW G-4.

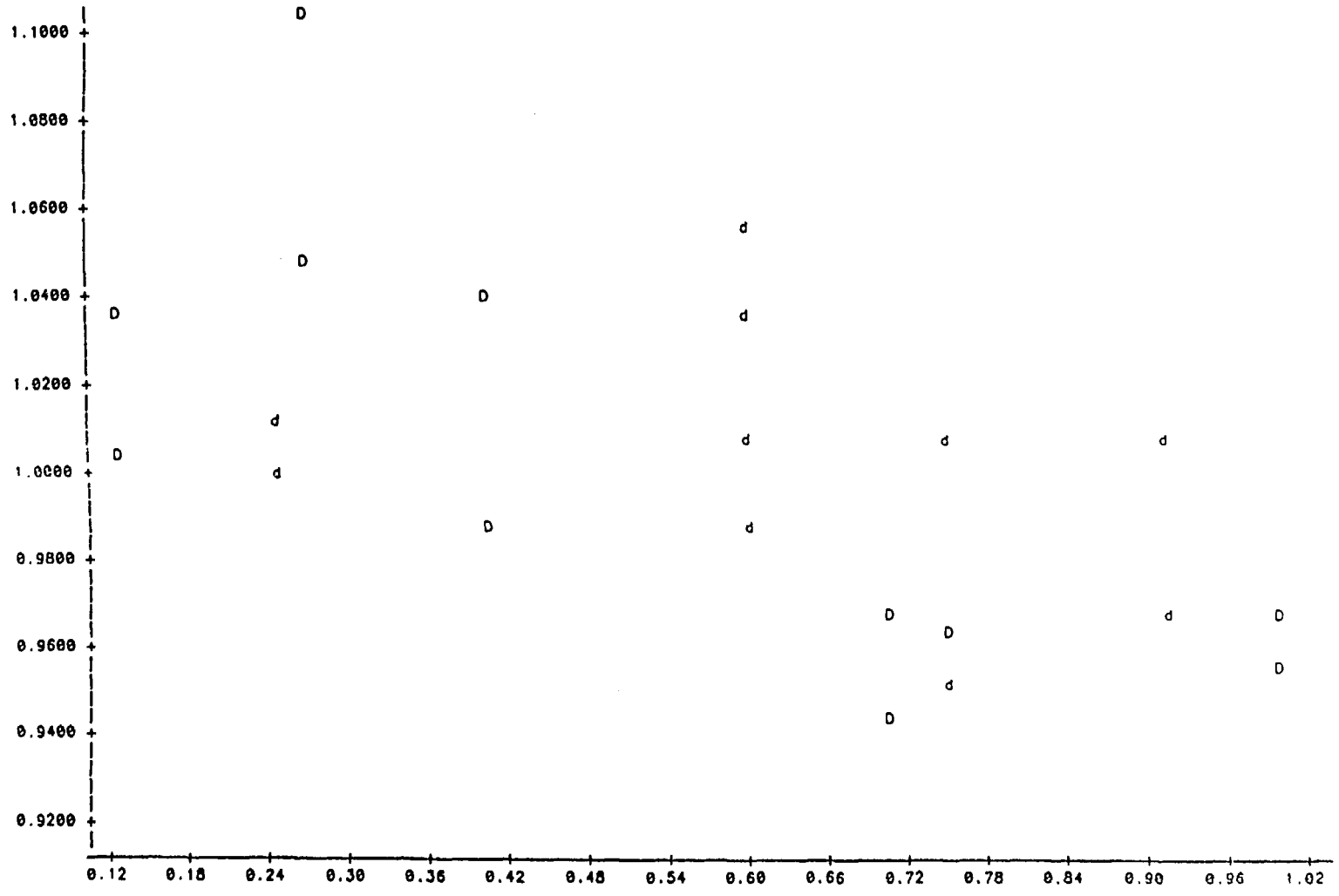

Figure 8. Plot of RATIO6 as a function of relative depth in the lower nonlithophysal zone. Points plotted using lower-case letters come from USW G-3; those plotted with upper-case letters come from USW G-4.

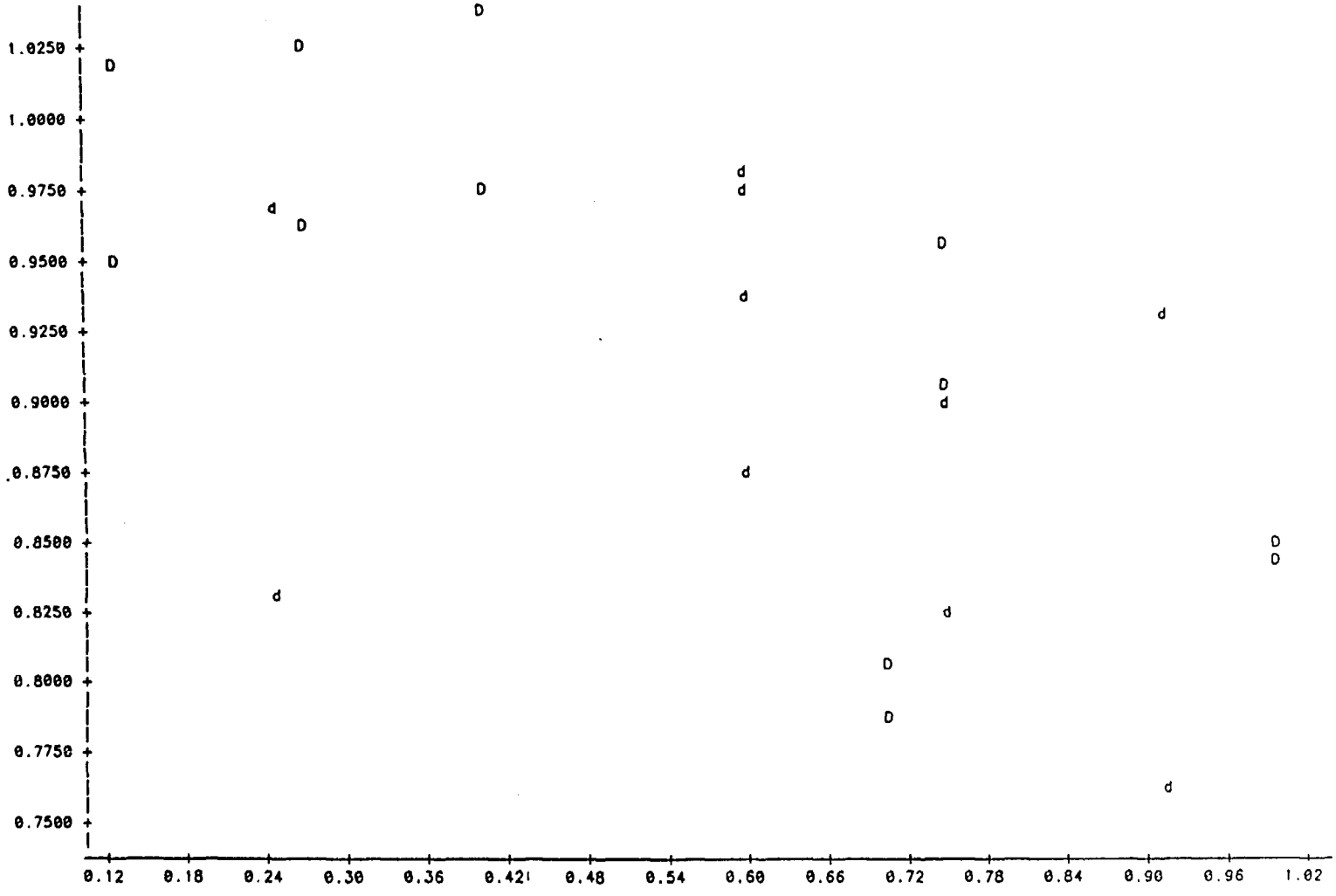

Figure 9. Plot of RATIO7 as a function of relative depth in the lower nonlithophysal zone. Points plotted using lower-case letters come from USW G-3; those plotted with upper-case letters come from USW G-4.

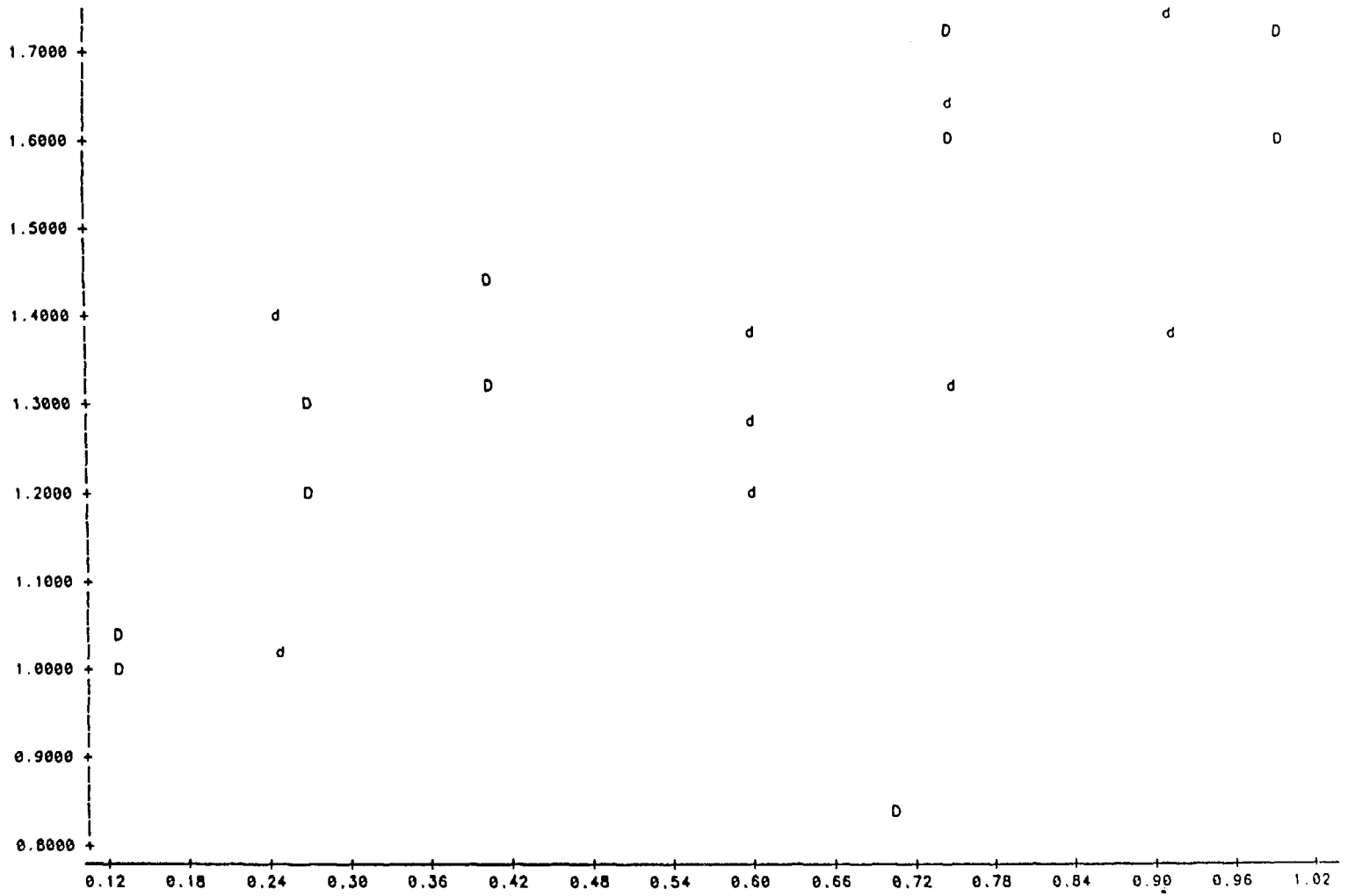

Figure 10. Plot of LGSD8 as a function of relative depth in the lower nonlithophysal zone. Points plotted using lower-case letters come from USW G-3; those plotted with upper-case letters come from USW G-4.

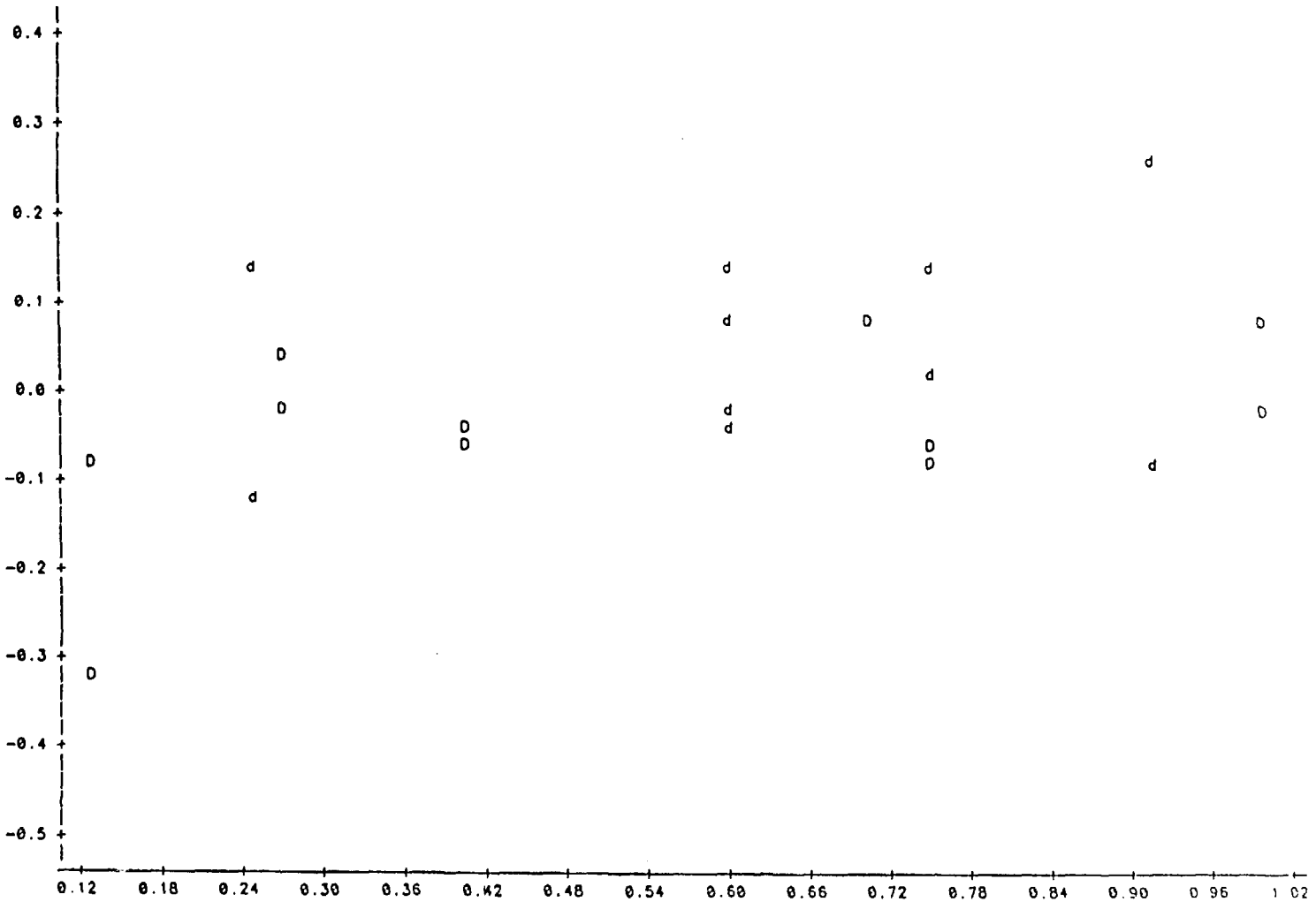

Figure 11. Plot or residual from regression of relative depth in the lower nonlithophysal zone on AREA4, RATIO6, RATIO7, and LGSD8, as a function of relative depth in the lower nonlithophysal zone. Points plotted using lower-case letters come from USW G-3; those plotted with upper-case letters come from USW G-4.

REFERENCES

Bish, D. L., "Evaluation of Past and Future Alterations in Tuff at Yucca Mountain, Nevada, Based on the Clay Mineralogy of Drill Cores USW G-1, G-2, and G-3," Los Alamos National Laboratory report LA-10667-MS (March 1989). NNA.890126.0207

Bish, D. L., and D. T. Vaniman, "Mineralogic Summary of Yucca Mountain, Nevada," Los Alamos National Laboratory report LA-10543-MS (October 1985). NNA.870407.0330

Brandshaung, T., "A Sensitivity Study of the Thermomechanical Far-Field Model of Yucca Mountain," Sandia National Laboratories report (in preparation).

Byers, F. M., Jr., "Petrochemical Variation of Topopah Spring Tuff Matrix with Depth (Stratigraphic Level), Drill Hole USW G-4, Yucca Mountain, Nevada," Los Alamos National Laboratory report LA-10561-MS (December 1985). HQS.880517.1103

Byers, F. M., Jr., and L. M. Moore, "Petrographic Variation of the Topopah Spring Tuff Matrix Within and Between Cored Drill Holes, Yucca Mountain, Nevada," Los Alamos National Laboratory report LA-10901-MS (December 1987). HQS.880517.2630

Lipman, P. W., R. L. Christiansen, and J. T. O'Connor, "A Compositionally Zoned Ash-Flow Sheet in Southern Nevada," US Geological Survey Professional Paper 524-F, pp. F1-F47 (1966). HQS.880517.1319

Moore, L. M., F. M. Byers, Jr., and D. E. Broxton, "Statistical Test of Reproducibility and Operator Variance in Thin-Section Analysis of Textures and Phenocrysts in the Topopah Spring Member, Drill Hole USW VH-2, Crater Flat, Nye County, Nevada," Los Alamos National Laboratory report LA-11452-MS (June 1989). NNA.890502.0089

Scott, R. B., and M. Castellanos, "Stratigraphic and Structural Relations of Volcanic Rocks in Drill Holes USW GU-3 and USW G-3, Yucca Mountain, Nye County, Nevada," US Geological Survey Open-File Report 84-491 (1984). HQS.880517.1444

Zielinski, R. A., "Evaluation of Ash-Flow Tuffs as Hosts for Radioactive Waste: Criteria Based on Selective Leaching of Manganese Oxides," US Geological Survey Open-File Report 83-480 (1983). HQS.880517.1582

APPENDIX
NOTEBOOKS AND DATA FILES

Data in this report were under contract (TWS-ESS-1-10/87-2) by Everest Technologies, Houston, Texas. Notebook entries pertinent to this report can be found in notebook TWS-ESS-1-4/84-36 on pages 111-113, 134-135, 137-138, and 149.

Data in this report can be found in the following computer files:

1) MASS storage files residing at the Los Alamos Central Computing Facility (CCF).
Directory Pathway = 086049/images and 086049/images2.
Data File:

```
trunk path  /086049
this root was structured on 89/04/05 at 09:57:56 on machine 2
images-dir
  310b1.stat    310b2.stat    310b2b.stat   310b3.stat    310b3b.stat   310b3c.stat
  310b4.stat    310b5.stat    310c6.stat    310b7.stat    310b8.stat    311.stat
  3111.stat     3112.stat     3113a.stat    3114.stat     3115.stat     3116.stat
  3117.stat     3118.stat     3119.stat     312.stat      312b.stat     313.stat
  3131.stat     3132.stat     3133.stat     3133b.stat    3133c.stat    3134.stat
  3135.stat     3136.stat     3137.stat     3138.stat     313b.stat     313b1.stat
  313b2.stat    313b2b.stat   313b3.stat    313b3b.stat   313b3c.stat   313b4.stat
  313b5.stat    313b6.stat    313b7.stat    313b8.stat    314.stat      3141.stat
  3142.stat     3142b.stat    3143.stat     3143b.stat    3144.stat     3145.stat
  3146.stat     3147.stat     3148.stat     315.stat      316.stat      317.stat
  318.stat      31b1.stat     31b2.stat     31b2b.stat    31b3.stat     31b3b.stat
  31b4.stat     31b5.stat     31b6.stat     31b7.stat     31b8.stat     321.stat
  322.stat      322b.stat     323.stat      323b.stat     324.stat      325.stat
  326.stat      327.stat      328.stat      32b1.stat     32b2.stat     32b3.stat
  32b3b.stat    32b4.stat     32b5.stat     32b6.stat     32b7.stat     32b8.stat
  331.stat      332.stat      332b.stat     333.stat      333b.stat     334.stat
  335.stat      336.stat      337.stat      338.stat      33b1.stat     33b2.stat
  33b2b.stat    33b3.stat     33b3b.stat    33b4.stat     33b5.stat     33b6.stat
  33b7.stat     33b8.stat     341.stat      342b.stat     344.stat      345.stat
  346.stat      347.stat      348.stat      34b4.stat     34b5.stat     34b6.stat
  34b7.stat     351.stat      352.stat      352b.stat     353.stat      353b.stat
  354.stat      355.stat      356.stat      357.stat      358.stat      35b1.stat
  35b2.stat     35b2b.stat    35b3.stat     35b3b.stat    35b4.stat     35b5.stat
  35b6.stat     35b7.stat     35b8.stat     361.stat      362.stat      362b.stat
  367.stat      367b.stat     364.stat      365.stat      366.stat      367.stat
  368.stat      36b1.stat     36b2.stat     36b4.stat     36b5.stat     36b6.stat
  36b7.stat     36b8.stat     36b9.stat     371.stat      372.stat      372b.stat
  373.stat      373b.stat     374.stat      375.stat      376.stat      377.stat
  378.stat      37b1.stat     37b2.stat     37b2b.stat    37b3.stat     37b3b.stat
  37b4.stat     37b5.stat     37b6.stat     37b7.stat     37b8.stat     38b1.stat
  38b2.stat     38b2b.stat    38b3.stat     38b3b.stat    38b4.stat     38b5.stat
  38b6.stat     38b7.stat     38b8.stat     392.stat      392b.stat     393.stat
  393b.stat     394.stat      395.stat      396.stat      397a.stat     398.stat
  39b2.stat     39b2b.stat    39b3.stat     39b3b.stat    39b4.stat     39b5.stat
  39b6.stat     39b7.stat     39b8.stat     411.stat      412.stat      413.stat
  414.stat      415.stat      416.stat      417.stat      418.stat      441.stat
  442.stat      442b.stat     443.stat      443b.stat     444.stat      445.stat
  446.stat      447.stat      448.stat      451.stat      451b.stat     452.stat
  452b.stat     453.stat      453b.stat     454.stat      455.stat      456.stat
  457.stat      458.stat      461.stat      462.stat      462b.stat     463.stat
  463b.stat     464.stat      465.stat      466.stat      467.stat      468.stat
  471.stat      472.stat      472b.stat     473.stat      473b.stat     474.stat
  475.stat      476.stat      477.stat      478.stat      test
images2-dir
  3101.stat     3102.stat     3102b.stat    3103.stat     3103b.stat    3104.stat
  3105.stat     3106.stat     3107.stat     3108.stat     311b2.stat    311b2b.stat
  311b3.stat    311b3b.stat   311b4.stat    311b5.stat    311b6.stat    311b7.stat
  311b8.stat    311b9.stat    3121.stat     3122.stat     3122b.stat    3123.stat
  3123b.stat    3124.stat     3125.stat     3126.stat     3127.stat     3128.stat
  3128b1.stat   312b2.stat    312b2b.stat   312b3.stat    312b3b.stat   312b4.stat
  312b5.stat    312b6.stat    312b7.stat    312b8.stat    314b1.stat    314b2.stat
  314b2b.stat   314b3.stat    314b3b.stat   314b4.stat    314b5.stat    314b6.stat
  314b7.stat    314b8.stat    314b9.stat    3151.stat     3152.stat     3152b.stat
  3153.stat     3153b.stat    3154.stat     3155.stat     3156.stat     3157.stat
  3158.stat     315b1.stat    315b2.stat    315b2b.stat   315b3.stat    315b4.stat
  315b5.stat    315b6.stat    315b7.stat    315b8.stat    32b2b.stat    34b2b.stat
  381.stat      382.stat      382b.stat     383.stat      383b.stat     384.stat
  385.stat      386.stat      387.stat      388.stat      4101.stat     4101z.stat
  4102.stat     4102b.stat    4104.stat     4105.stat     4106.stat     4107.stat
  4108.stat     410b1.stat    410b2.stat    410b2b.stat   410b3.stat    410b3b.stat
  410b4.stat    410b5.stat    410b6.stat    410b7.stat    4112b.stat    4113b.stat
  4114.stat     4115.stat     4116.stat     4117.stat     4118.stat     411b1.stat
  411b2.stat    411b2b.stat   411b3.stat    411b3b.stat   411b4.stat    411b5.stat
  411b6.stat    411b7.stat    411b8.stat    4121.stat     4122.stat     4122b.stat
  4123.stat     4123b.stat    4124.stat     4125.stat     4126.stat     4127.stat
  4128.stat     412b1.stat    412b2.stat    412b2b.stat   412b3.stat    412b3b.stat
  412b4.stat    412b5.stat    412b6.stat    412b7.stat    412b8.stat    4131.stat
  4132.stat     4132b.stat    4133.stat     4133b.stat    4134.stat     4135.stat
  4136.stat     4137.stat     4138.stat     413b1.stat    413b2.stat    413b2b.stat
  413b3.stat    413b3b.stat   413b4.stat    413b5.stat    413b6.stat    413b7.stat
  413b8.stat    4141.stat     4142.stat     4143.stat     4144.stat     4145.stat
  4146.stat     4147.stat     4148.stat     414b1.stat    414b2.stat    414b3.stat
  414b4.stat    414b5.stat    414b6.stat    414b7.stat    414b8.stat    414b9.stat
  4151.stat     4152.stat     4152b.stat    4153.stat     4153b.stat    4154.stat
  4155.stat     4156.stat     4157.stat     4158.stat     415b1.stat    415b2.stat
  415b2b.stat   415b2bz.stat  415b3.stat    415b3b.stat   415b4.stat    415b5.stat
  415b6.stat    415b7.stat    415b8.stat    4161.stat     4162.stat     4162b.stat
  4163.stat     4163b.stat    4164.stat     4165.stat     4166.stat     4167.stat
  4168.stat     416b1.stat    416b2.stat    416b2b.stat   416b3.stat    416b3b.stat
  416b4.stat    416b5.stat    416b6.stat    416b7.stat    416b8.stat    41b1.stat
  41b2.stat     41b2b.stat    41b3.stat     41b3b.stat    41b4.stat     41b5.stat
  41b6.stat     41b7.stat     421.stat      422.stat      422b.stat     423.stat
  423b.stat     424.stat      425.stat      426.stat      427.stat      428.stat
  42b1.stat     42b2.stat     42b2b.stat    42b3.stat     42b3b.stat    42b4.stat
  42b5.stat     42b6.stat     42b7.stat     42b8.stat     44b1.stat     44b2.stat
  44b2b.stat    44b3.stat     44b3b.stat    44b4.stat     44b5.stat     44b6.stat
  44b7.stat     44b8.stat     451.stat      452.stat      452b.stat     453.stat
  453b.stat     454.stat      455.stat      456.stat      457.stat      458.stat
  45b1.stat     45b2.stat     45b3.stat     45b4.stat     45b5.stat     45b6.stat
  45b7.stat     45b8.stat     46b1.stat     46b2.stat     46b2b.stat    46b3.stat
  46b3b.stat    46b4.stat     46b5.stat     46b6.stat     46b7.stat     46b8.stat
  471.stat      472.stat      472b.stat     473.stat      473b.stat     474.stat
  475.stat      476.stat      477.stat      478.stat      47b1.stat     47b2.stat
  47b2b.stat    47b3.stat     47b3b.stat    47b4.stat     47b5.stat     47b6.stat
  47b7.stat     47b8.stat     481.stat      482.stat      482b.stat     483.stat
  483b.stat     484.stat      485.stat      486.stat      487.stat      488.stat
  48b1.stat     48b2.stat     48b2b.stat    48b3.stat     48b3b.stat    48b4.stat
  48b5.stat     48b6.stat     48b7.stat     48b8.stat     491.stat      492.stat
  492b.stat     493.stat      493b.stat     494.stat      495.stat      496.stat
  497.stat      498.stat      49b1.stat     49b2.stat     49b2b.stat    49b3.stat
  49b3b.stat    49b4.stat     49b5.stat     49b6.stat     49b7.stat     49b8.stat
```

2) Data Tables in RS-1 Data Base Manager
Residing on ESSXRF MicroVax II.
Directory Pathway = ESSXRF::DUB0:[086049]
Within RS-1, Directory Pathway = @Broxton@image
Data Tables: ALLAREA
ALLROUGHNESS
ALLELONGATION
IMAGE_LIST
IMAGE2_LIST
IMAGE3_LIST

3) The following data files stored on the Los Alamos Central Computing Facility's Common File System (directory /v81583/everest):

area1.dat area2.dat area3.dat elong1.dat elong2.dat elong3.dat ident1.dat ident2.dat rough1.dat rough2.dat rough3.dat.

Those with "1" in their name have 256 records each and are slightly edited (to remove headers and insert place holders so all lines are the same length) versions of files containing statistics for area, elongation, and roughness, with the list of file names in ident1.dat. The "2" files have 327 records each, and the "3" files have 23 records, representing some data received in the latter part of the study.

Three SAS command files stored in the same cfs directory (makeone.sas, maketwo.sas and makethree.sas) read these files and create SAS-formatted data sets in which each observation corresponds to a bitmap and the variables given for each observation are the statistics for objects in that bitmap, some derived statistics (assuming other distributions for area--namely, log normal, and elongation--namely gamma, and a few other ratios) and some identifying information such as hole (USW GU-3 or G-4), scene (A or B), slide number and associated depth, and lithology as supplied by ESS-1 (mnl, ll or lnl). The SAS data files were constructed on the S1VAX as [081583.B321.EVEREST]ONE.SSD, TWO.SSD and THREE.SSD. These files were not saved, but they could be reconstructed at any time, given the files stored in the cfs.

Finally, however, what is really wanted is a data base organized so that each observation corresponds to one "scene" and has all the statistics for all of the bitmaps for that scene. This data base is constructed by the SAS command file SPLMERGE.SAS (all these SAS command files are stored in /v81583/everest on cfs), which begins with ONE, TWO and THREE above. All analyses in this report are based on the resulting SAS data file, which was constructed as [081583.B321.EVEREST]SPECIMEN.SSD.

The research that this report documents is assigned a Quality Assurance Level of III.

The following number is for Office of Civilian Radioactive Waste Management purposes only and should not be used when ordering this document:

Accession Number: NNA.890502.0090

☆ U.S. GOVERNMENT PRINTING OFFICE: 1989—773-205/00026

Page Range	NTIS Price Code	Page Range	NTIS Price Code	Page Range	NTIS Price Code	Page Range	NTIS Price Code
001–025	A02	151–175	A08	301–325	A14	451–475	A20
026–050	A03	176–200	A09	326–350	A15	476–500	A21
051–075	A04	201–225	A10	351–375	A16	501–525	A22
076–100	A05	226–250	A11	376–400	A17	526–550	A23
101–125	A06	251–275	A12	401–425	A18	551–575	A24
126–150	A07	276–300	A13	426–450	A19	576–600	A25
						601-up*	A99

*Contact NTIS for a price quote.