BSR 2941

MULTISPECTRAL DATA ANALYSIS FINAL REPORT

By:

David S. Hanson

David R. Morganstein

May 1970

Prepared Under Contract No. NAS 12-2123 by

THE BENDIX CORPORATION AEROSPACE SYSTEMS DIVISION

Ann Arbor, Michigan

Electronics Research Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

CONTENTS

ILLUSTRATIONS

ILLUSTRATIONS (CONT.)

TABLES

SECTION 1

INTRODUCTION AND SUMMARY

1.1 PURPOSE

This quarterly report describes the analyses performed by the Bendix Aerospace Systems Division on multispectral information collected over the shoal waters of Lake Michigan. Included are details of the acquisition and processing hardware. Principally, however, this report discusses the analysis approach used in studying multivariate reflectance data.

The particular problem examined is the prediction of water depth from reflectance measures. Before beginning an analysis, the physical parameters were studied and a mathematical model derived. The technique of regression analysis was chosen to evaluate the various channel measurements as predictors of water depth. To assist in the association of a depth measure with a set of reflectances, both factor analysis and time series analysis were used. The factorizing of multivariate data represents a tool whose scope reaches beyond this single requirement. The significance and use of factor analysis are hence discussed in detail.

The material reported covers the first phases of analysis. The conclusions drawn are already guiding further study.

Section 1.2 includes information on both the data acquisition and data processing hardware. The analog tapes recorded in the air are analyzed with the laboratory facilities as indicated. The possible routes of data flow are shown in the discussion on processing techniques.

1.2 DATA ACQUISITION FACILITIES

1.2.1 Bendix Multispectral Scanner

The airborne elements of the Bendix Multispectral Scanner System (Figure 1) are a multispectral scanner, an electronics subsystem, and an analog tape recorder. The scanner provides eight channels operating in the spectral region from 0.38 to 1.2 microns. The instrument is basically an imaging grating spectrometer using multiple photomultiplier detectors for data output (Figure 2).

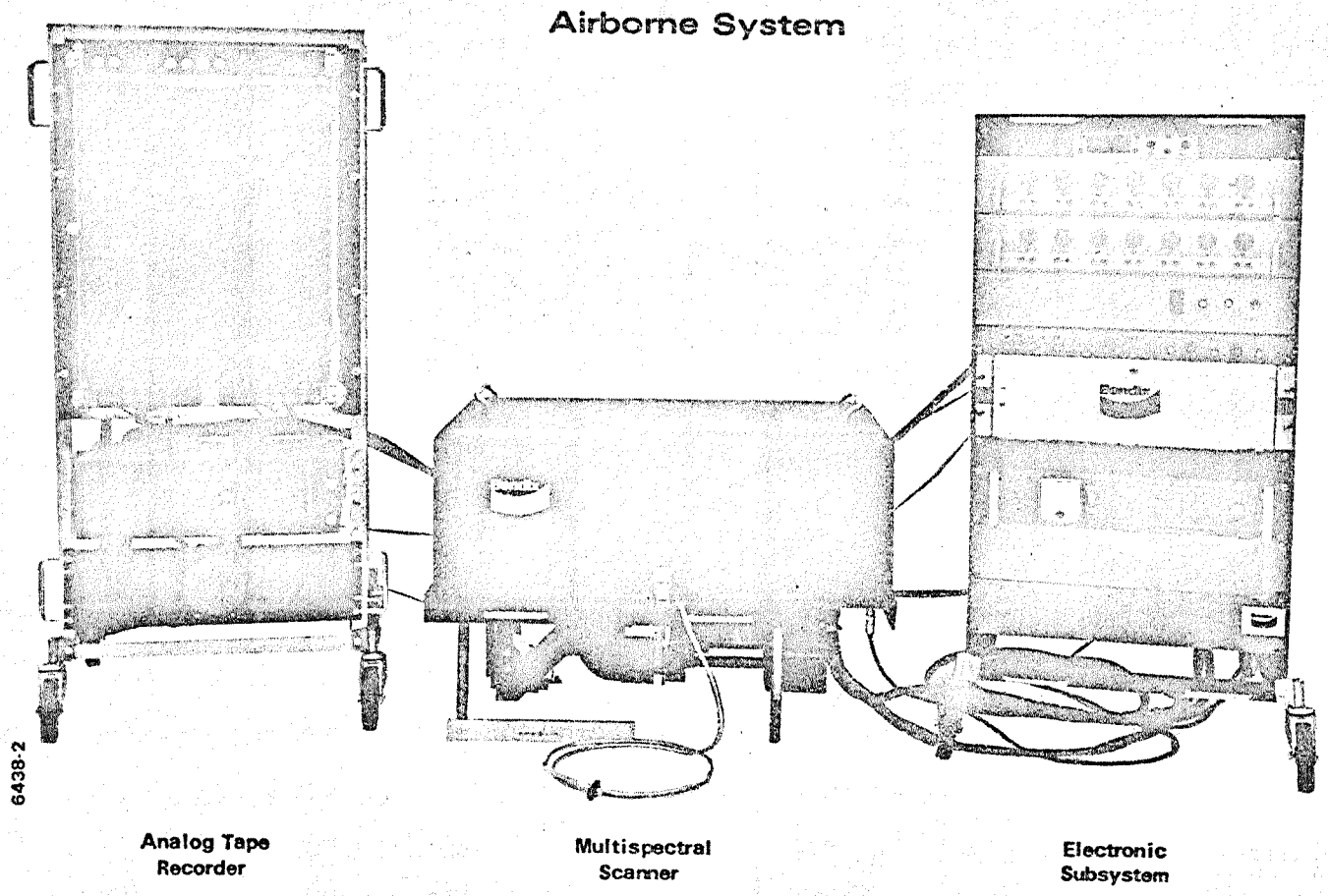
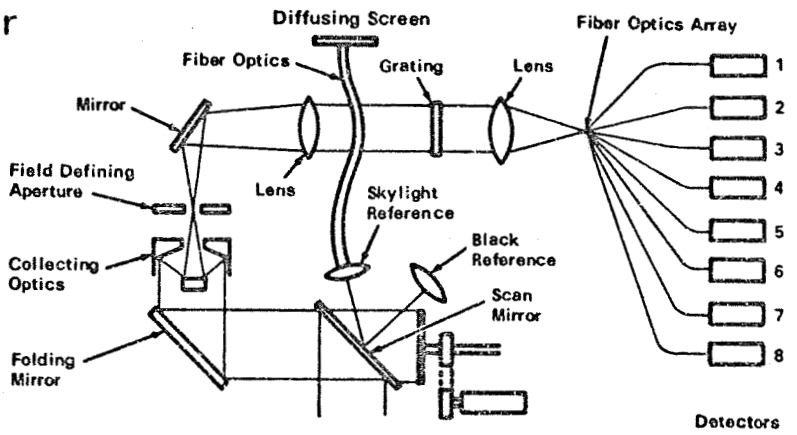


Figure 1 Multispectral Scanner Airborne System

Multispectral Scanner Pictorial Layout



Performance Characteristics

Channels (Number)	Spectral Band (Microns)	Detector (Photocathode)
1	0.38 – 0.44	S-20
2	0.44 – 0.50	S-20
3	0.50 – 0.56	S-20
4	0.56 – 0.62	S-20
5	0.62 – 0.68	S-20
6	0.68 – 0.74	S-20
7	0.74 – 0.86	S-1
8	0.86 – 1.0	S-1

Scanning rate (nominal)	100 revolutions/sec
Instantaneous Field-Of-View (Spatial Resolution)	2.5 milliradians
Lateral Scan Angle	120 degrees
Output Each Channel	Terrain Spectral Reflectance
Design V/h (aircraft velocity f.p.s/ altitude in feet)	0.25
Analog Tape Recorder	Ampex AR 1600
Tape Running Time (per reel)	30.0 min.
System Weight (Scanner, Electronics, Recorder)	400 lb.
Power Requirements (Scanner, Electronics, Recorder)	27 a @ 28V DC 7.5 a @ 115V, 400 cps Single Phase

7589-16

Figure 2 Multispectral Scanner Pictorial Layout and Performance

Of interest in the scanner system is the technique used for calibration, control, and dynamic range compression of the electronic signals. The technique used is considered a major technological advance in this type of scanner. Calibration sources are provided for dark reference and illumination level reference. Signals from these calibration sources are injected into every channel on every scan line. The illumination level reference used is diffuse solar (and sky) irradiance brought into the scanner from above the aircraft through a fiber optics bundle terminating below a diffusing surface (Figure 3).

Calibration of each channel is accomplished by a closed-loop electronic servo which automatically compensates for spectral and intensity variations in terrain illumination and variations in the optical and electronic characteristics of the scanner. The video output from each of the eight channels is calibrated to yield directly terrain spectral reflectance. System electronics permit selection of the total reflectance range sent to the tape recorder. This is a vital feature because tape recorder dynamic range is the limiting factor of the system dynamic range.

A synchronizing signal, produced by the rotation of the scan shaft, and the signal from each channel of the multispectral scanner is passed to the electronics subsystem for frequency modulation. From here it is passed to the AR-1600 analog tape recorder. The coiled element beside the scanner in Figures 1 and 2 is a fiber optics bundle which is used for automatic gain control (AGC) of the scanner electronics. This fiber optics bundle is inserted through a hole in the top skin of the aircraft and samples the illumination incident on the aircraft top. Because the flights are conducted at low altitudes, the illumination incident on the aircraft top is very nearly the same as that incident on the ground. The illumination sampled this way is conducted to a near-zero-reflectance cavity and a lens which collimates light transmitted by the fiber optic bundle. Collimated light from this fiber optics bundle is thus imaged on the aperture of the scanner when the scan mirror is directed toward the collimating lens. The signal produced when the scan mirror is looking at this reference is used to drive a closed-loop AGC system which is independent for each channel of scanner electronics. In this way, the video signal actually recorded is independent of the incident illumination and becomes a property only of the material in the field of view (FOV).

1.2.2 Data Processing Facilities

The airborne elements of the Bendix multispectral scanner system include the multispectral scanner, the electronics subsystem, and an analog tape recorder (Figure 1). The scanner has eight channels operating in the spectral range from 0.38 (ultraviolet) to 1.0 micron (near-IR). Tapes recorded in flight have eight channels of video data and one channel of synchronization signal.

7589-17

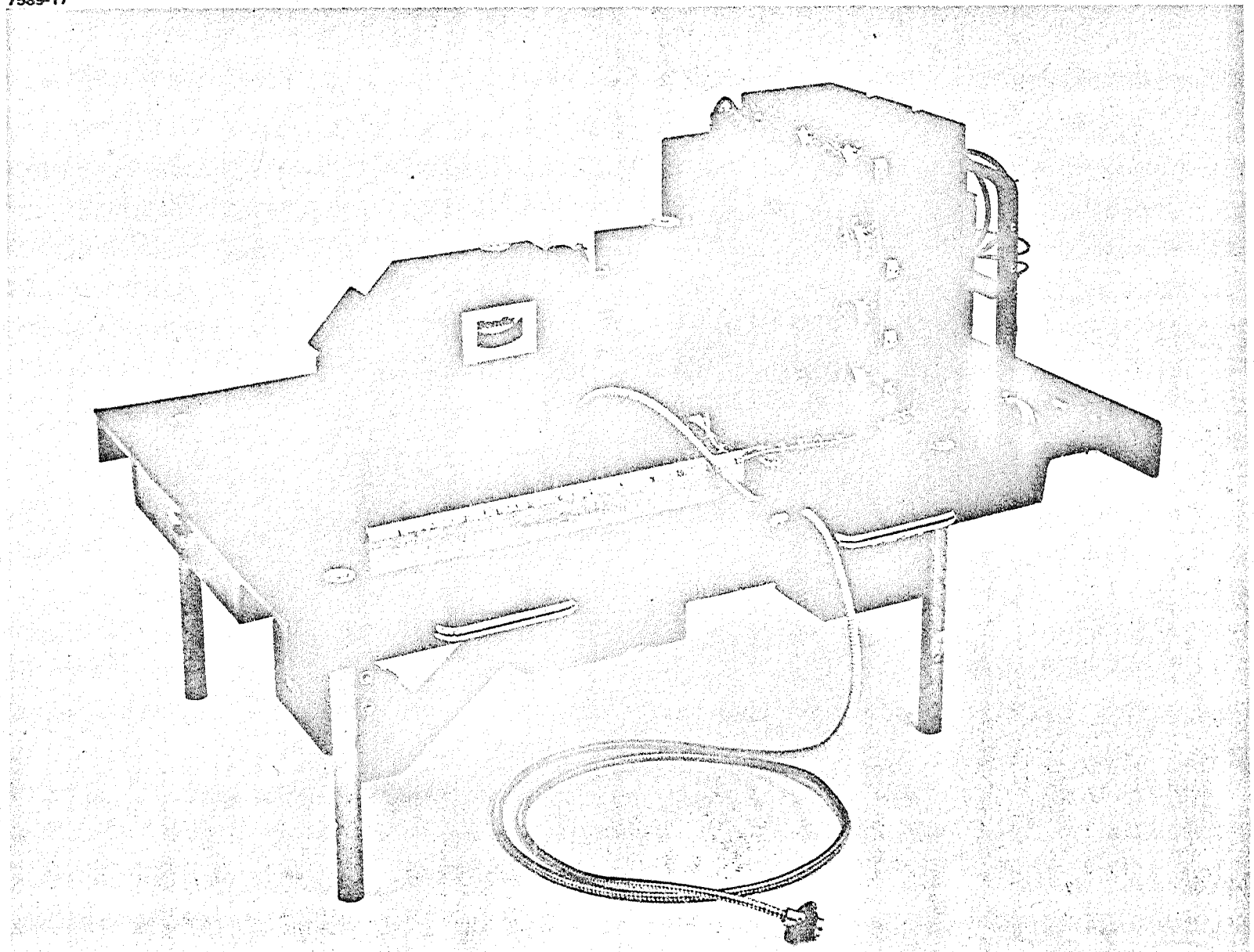


Figure 3 Multispectral Scanner, Cover Removed

The video produced by the scanner has a nominal bandwidth extending from DC to approximately 100 kHz. To preserve information over the entire frequency range, FM recording of the video signal was selected. The AR-1600 is not an FM tape recorder. Frequency modulation of the video signal is accomplished in the electronics subsystem. The input to the electronics subsystem is the video signal with the bandwidth of DC to 100 kHz. The signal obtained on the output of the electronics subsystem is an FM signal with 450-kHz carrier frequency. The modulated carrier is then passed to the AR-1600 for direct analog recording.

The video signal is displayed in the laboratory by playing the tape back through a wideband FM system. The laboratory is provided with an Ampex FR-1800H tape recorder for this purpose. The carrier frequency recorded in flight is demodulated and made available for use in the laboratory just as though it were being produced in real time by the multichannel scanner.

At this point (identified in Figure 4 as tape playback), three different operations may take place. Single-channel imagery may be produced by directing the video from the desired channel to the film recorder. Alternatively, digital samples may be taken from the analog tapes. A 5-μsec sample is taken simultaneously from each channel at the output of the recorder. The sampling point occurs at a constant, adjustable, time delay from the leading edge of the video signal. When the video signals are film recorded side by side, the sampling points appear as a straight line in the 120° FOV of the scanner running the length of the flight line. These samples are recorded on a digital tape transport and subjected to data analysis (Figure 5).

Analog data processing equipment is shown in Figure 6.

The end result of such analyses is processing constants. These coefficients are applied to an analog computer to yield enhanced signals, corresponding to the center branch of Figure 4. Each channel of the original recorded data is delivered to a separate processor and is multiplied by a constant as dictated by the computer printouts. All modified channels are then summed, forming a single new signal—a linear combination of the eight channels. At present, two such linear combinations may be generated simultaneously. Enhanced imagery may then be recorded in the same manner as a single video channel. The results of a previous analysis may be applied directly to unanalyzed data (represented by center path of Figure 4). This would occur when a particular linear combination has been previously isolated as causing a desirable enhancement.

Imagery of both continuous and binary gray levels is produced by using a fiber optics cathode ray tube (CRT). The CRT is operated at 7 kv, the beam

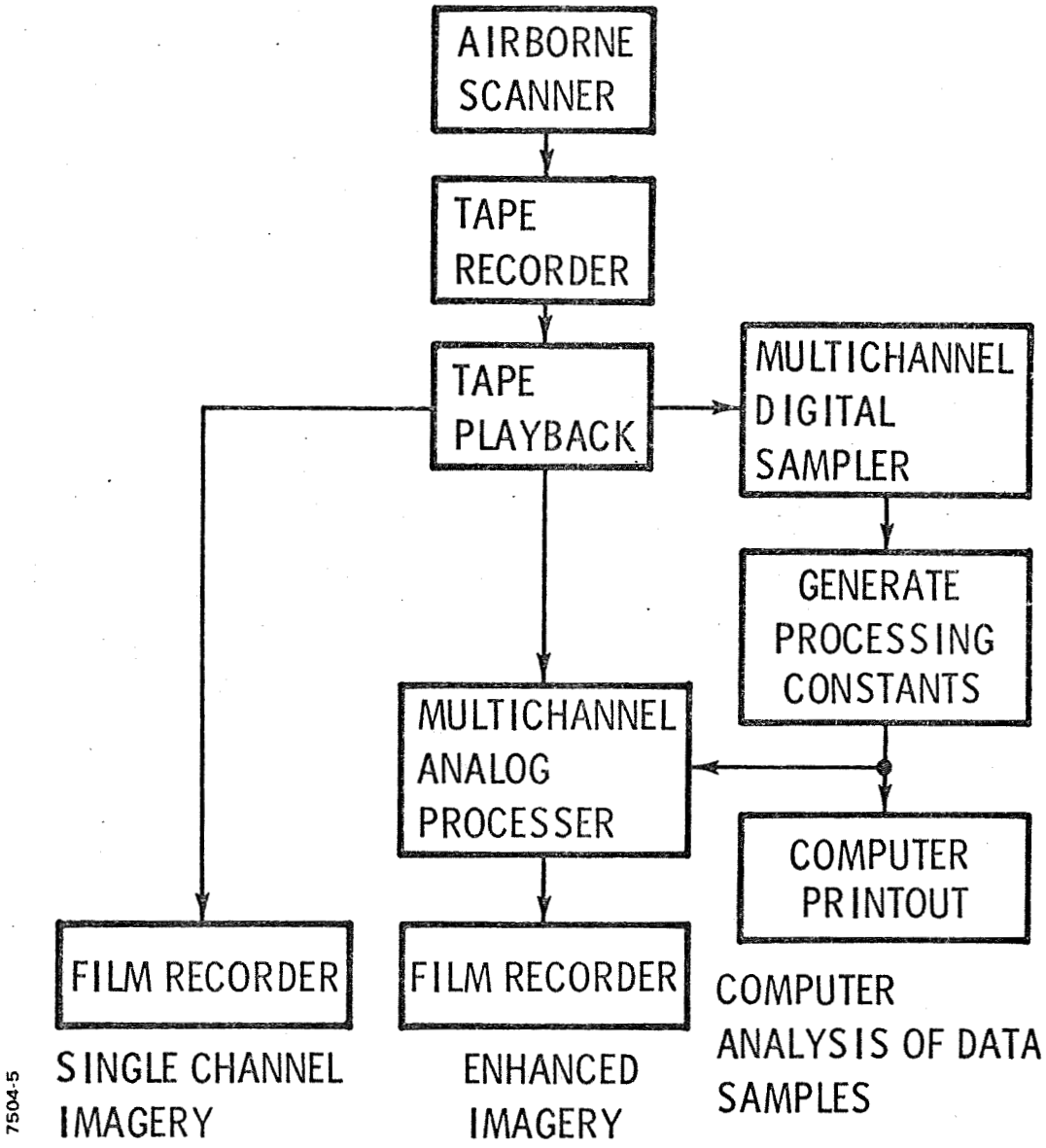


Figure 4 Laboratory Operations, Block Diagram

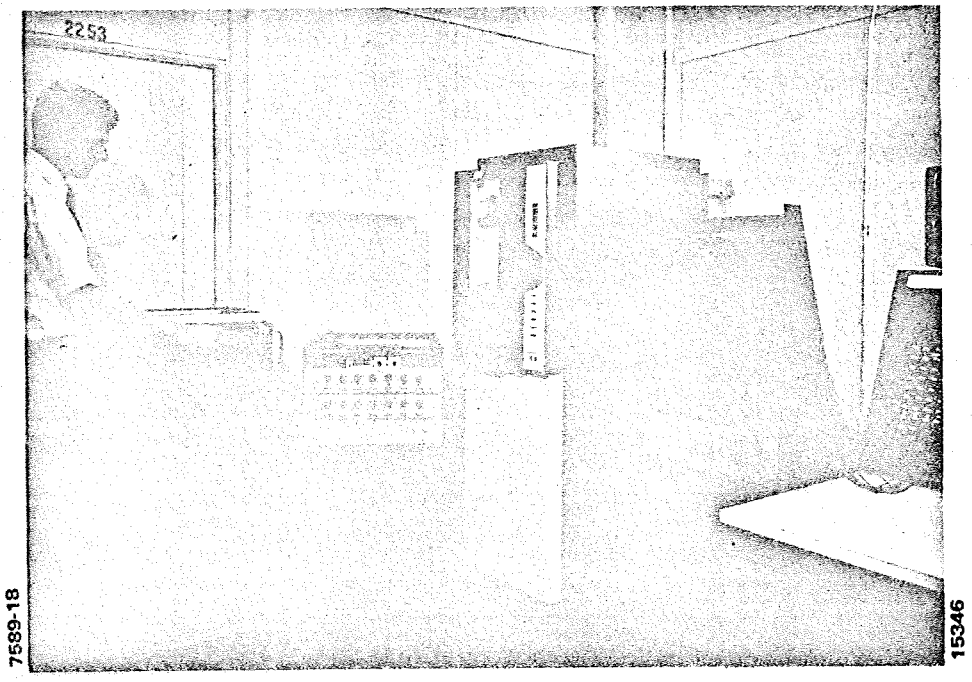


Figure 5 Digital Data Sampling

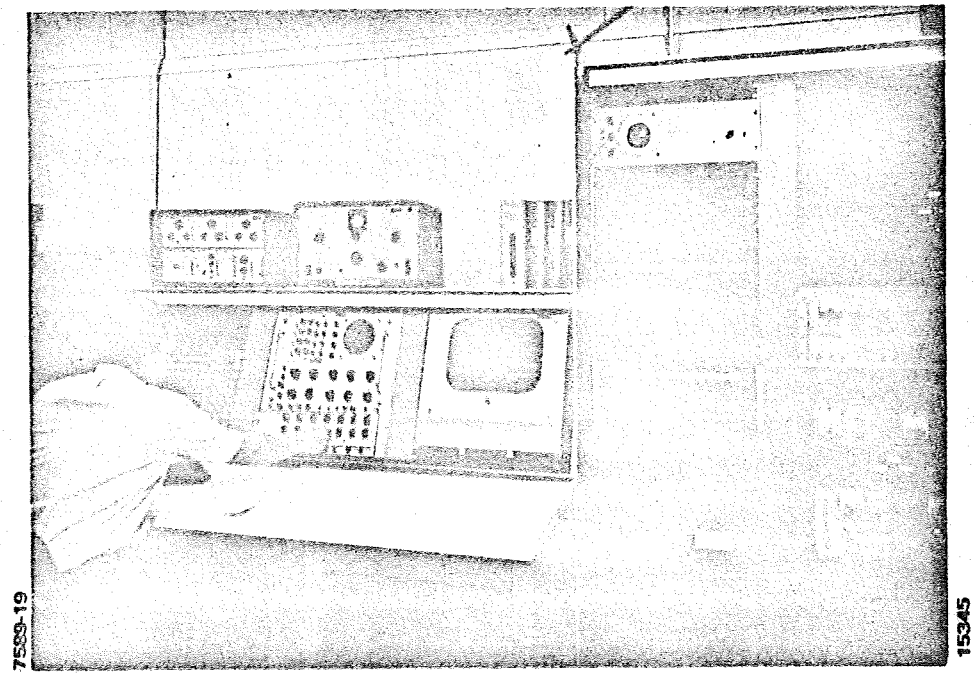


Figure 6 Analog Data Processing Equipment

being deflected in one direction only. To produce imagery, a roll of film is simply pulled across the fiber optics faceplate. The exact speed of the film is linearly proportioned to the velocity/height (V/H) ratio of the aircraft in which the data were collected. The video data become available at a rate of 100 lines per second. This is because the rotation in the scan shaft is 100 rev/sec when the data are originally collected. The duty cycle is 33.33%, corresponding to the fact that the video gate in the scanner circuitry is open only during the period of time when the scan mirror is within 60° of the vertical position (120° FOV). A video pedestal with the same duty cycle must be prepared to modulate the intensity of the CRT beam. The video signal is superimposed on the high portion of the video pedestal while beam blanking is obtained during the remaining two-thirds of the duty cycle.

Figure 7 shows the various waveforms and video signals used to modulate the intensity and deflect the CRT beam during film recording. The primary clock is the video sync signal recorded in flight. The total duration of one cycle of this signal is 0.01 sec. Its amplitude is -2 v during the time video is recorded and 0 during the unused portion of the duty cycle. The video pedestal is obtained by inverting the primary clock or video sync signal to obtain a zero voltage during the useful portion of the duty cycle and a -2 v blanking signal during the unused portion. The video signal and video pedestal are added and directed to the video amplifier which prepares the signal for modulation of the control grid voltage. This is usually an isolated video amplifier, because fiber optic CRTs must be operated with a grounded anode placing the cathode and control grid at a negative high voltage.

The CRT is magnetically deflected. The input signal for the deflection amplifier is a sawtooth signal obtained with a special circuit procured as an off-the-shelf device which is triggered by video sync signal. When the amplitude and duration of the sawtooth are adjusted to be suitable for beam deflection, the signal is directed to the deflection amplifier. This device transduces the voltage waveform into the appropriate current through the deflection coils.

These are the essential operations that must be performed on the tape recorded data to place the video signal in a form suitable for film recording. The sync signal is used for all timing operations in connection with operating the CRT. The video signal proper is used for intensity modulation. Gain and bias settings from this point are more critically related to the properties of the particular system used to produce imagery than to the properties of the recorded data used to produce the video signal.

This section has described the use of collected analog data, indicated various hardware procedures for generating imagery, as well as discussed

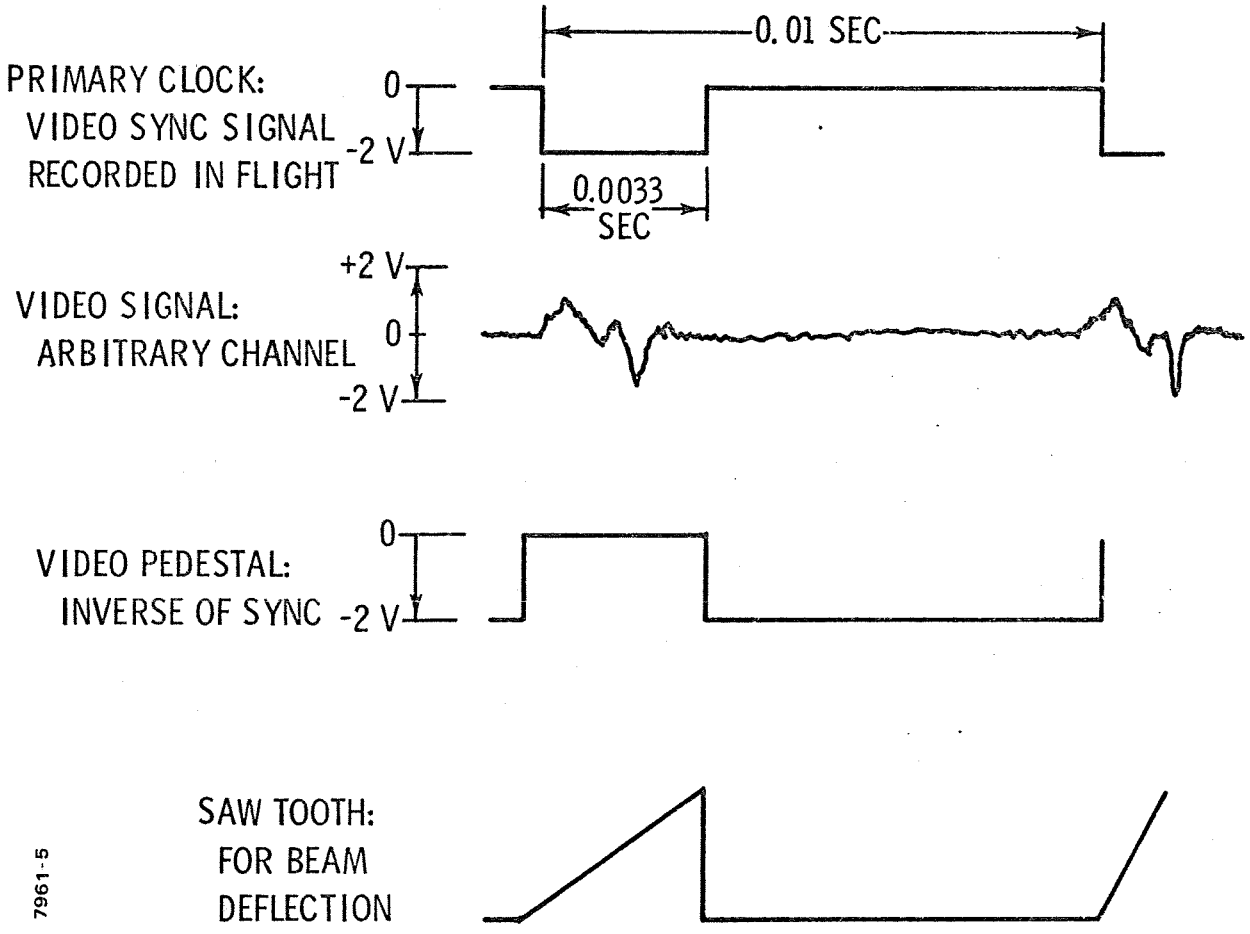


Figure 7 Waveforms and Signals Superposed to Operate the Film Recorder

some intermediate operations necessary in the production of that imagery. The data analysis techniques are detailed in following sections.

1.3 SCOPE OF ACTIVITIES

Several flights were made over the shoal water of eastern Lake Michigan (Pentwater, Michigan) to collect multispectral line scanner data and sufficiently thorough and quantitative ground truth to permit a careful data evaluation. This site was selected for two reasons: (1) the lake area is being thoroughly mapped by the US Lake Survey branch of the Army Corps of Engineers; and (2) the location is convenient to Ann Arbor, Michigan. Another flight was conducted over Hammond Bay, a portion of Lake Huron northwest of Rogers City, Michigan. Less thorough ground truth is available for the Hammond Bay data.

The data were collected for information related to water depth. By processing the multispectral video, any information related to water depth is to be isolated and quantitatively related to actual water depth measurements. Imagery from the scanning instrument may be interpreted directly as a measure of spectral reflectance. The video data, however, are not linearly related to the actual depth. A mathematical model describing the relationship between the video measurements and the actual water depth has been developed to establish the nature of the nonlinearity.

Other information in the video is of secondary interest in exploring the capabilities of multispectral equipment and developing data processing techniques to extract the information. In support of this interest, water samples of selected points and depths along the flight lines were collected and analyzed in a chemical laboratory to determine the amount and nature of the suspended material in the water. The results of the analysis are contained in Appendix A.

Hence, the information now assembled includes: (1) the multispectral video data recorded during three flights (two over the Lake Michigan site and one over Hammond Bay in Lake Huron), (2) a mathematical model parametrically describing the variations to be expected, (3) the results of water sample analysis for the two Lake Michigan flights, and (4) IR imagery and photography collected during some of the flights.

The value and interpretation of this information are the substance of the work in this report. This interpretation includes statistical analysis of the video data and its correlation with available ground truth. This process is intended to lead to a quantitative mapping of water depth contours. The study in progress is directed to the theory of interactions between water and sunlight, and a statistical study of the tape recorded data which illustrate these

variations. Remaining variations in the data will be examined to determine if they are consistent with explanations related to sun angle, weather conditions, instrument performance, or other such effects. From this examination, improvements in instrument performance and design might be suggested.

With the exception of the water sample collection and its subsequent analysis, the study activities are restricted to interpretation and analysis of the recorded video signals. While field activities directed towards providing video data suitable for analysis are not part of the study, a description of those activities is included.

This report includes a statement of each particular problem, the approach taken, the accomplishments, and further planned work. The subjects are: (1) mathematical modeling of water and light interactions, (2) factor analysis studies, (3) time series analysis of lake data, and (4) preliminary results from a regression study.

SECTION 2

APPROACH

The approach used to satisfy the primary objective of the study—the development of a methodology for extraction of information from multispectral data—has been the intensive study of a sample of data within which the information content is known in advance. The methods found to be useful for automatic extraction of this information will then be applied to similar data to evaluate the degree of generality with which the methodology developed may be applied.

The problem selected for study is the measurement of water depth in shoal water, using the Bendix multispectral scanner as a remote sensing device. Flights conducted during the week of 21 July (over Lake Michigan near Pentwater, Michigan) form the body of shoal water data selected for initial intensive study. These flights were flown over the area shown in the map of Figure 8. The best on-sight ground truth was obtained on Flight Line 4. Digital samples of the tape recorded video were taken from this flight line.

All the flight lines shown on the map were selected to coincide with lines along which the US Lake Survey Branch of the Army Corps of Engineers had plotted the shape of the bottom as a function of the distance from the water's edge. This work was completed approximately two months before the first Bendix flights. A US Lake Survey representative from the Detroit office who was in charge in the field operations noted that the bottom had not changed significantly during the three months of their study. Since the weather conditions did not include severe storms between the completion of the US Lake Survey operation and the Bendix flights, this source was used as the water depth measurement.

Bendix also conducted operations on the water during the week of the flights. Independent measurements were taken of the water depth contour. Water samples were collected and analyzed as described in Appendix A. The intended flight line was marked on the water surface with anchored floats (visible in the imagery recorded during the flights).

The bottom contour along Flight Line 4, as reported by the US Lake Survey, is illustrated in Figure 9. Using multiple regression analysis, quantitative measures of a directly observable parameter such as water depth can

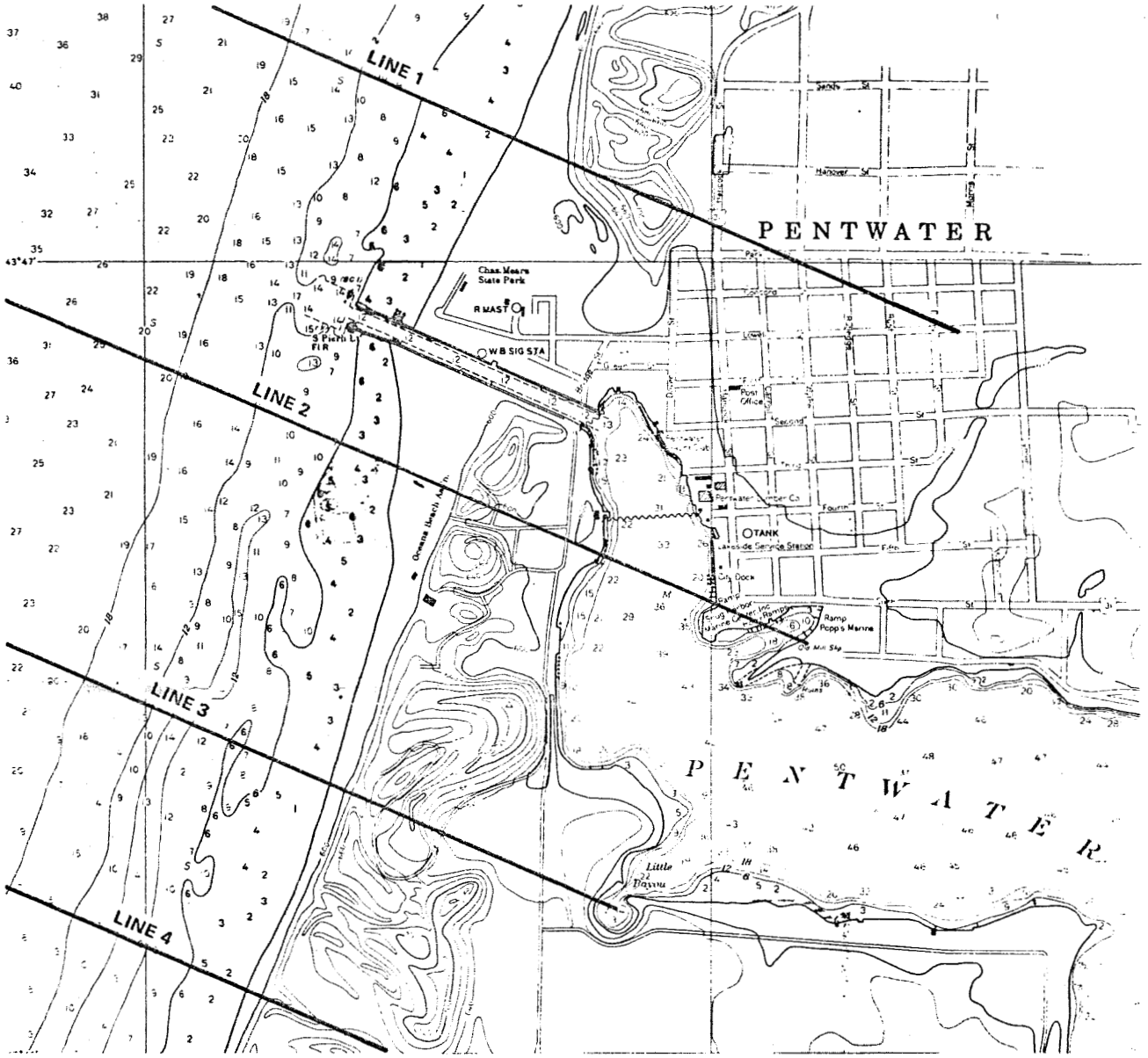


Figure 8 Map of Flight Lines Over Pentwater, Michigan

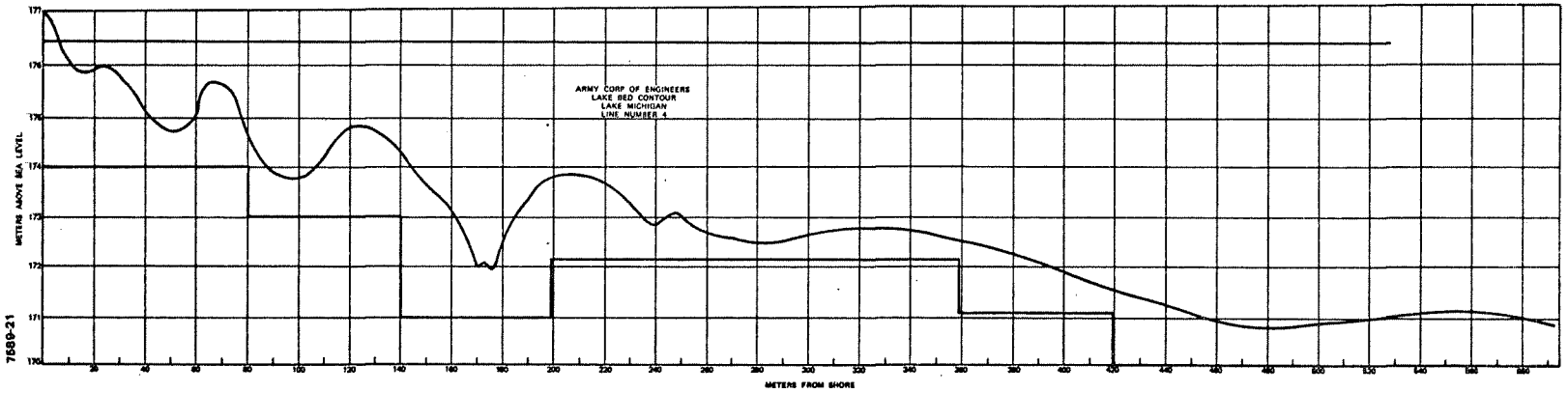


Figure 9 Bottom Contour of Flight Line 4

frequently be obtained from remotely observable parameters such as the video voltage levels of several channels of a multispectral scanner. The regression requires that each observation of the independent variables (the digital sampling of eight channels of multispectral video) be accompanied by the corresponding value of a dependent variable (the water depth).

The following description of the approach used in depth prediction begins with a discussion of a mathematical model. Simulations using the results of the model are revealed. The hardware technique used in providing digital samples of reflectance measures for computer analysis is then described. A method of associating, with each sample, a value of the dependent variable, water depth, utilizes time series and factor analysis. These two sections indicate how variation in the multispectral information, as isolated with factor analysis, is correlated with the lake bottom contour. Software improvements in both cross-correlation and factor analysis programs are mentioned. Completion of this method will make available a series of independent and dependent variables which can be subjected to regression analysis, as described in Section .

2.1 MATHEMATICAL MODEL

Color changes in shoal waters are usually associated with the bottom contour and composition. Since these changes are more pronounced in the spectral bands at which water transmits well, and since the effect is large, color change phenomena were one of the first selected for study with the Bendix 9-channel scanner. Because the color change is obviously not linearly proportional to depth, a model was developed to predict the relationship between depth and color change.

The intuitive requirements placed on the model were:

1. The bottom reflectance must influence the returned energy for shallow depths but not for large depths

2. The water alone determines the coloration of returned energy for large depths.

The problem is solved for the effective reflectance of the water and the bottom is defined as the ratio of upwelling to downwelling light intensity, as observed just below the water surface. This eliminates the effect of reflections from the air-water boundary. While this is an important effect in terms of what is observed with a line scanner, the effect is DC in nature and is not affected in the same way by depth variations. The model achieved is specialized to the case of diffuse illumination. The restriction is imposed

because the upwelling light intensity in the water is diffuse; under this condition, scattering is equally probable in the upward and downward directions. If the downwelling incident light intensity is not also diffuse, then downward scattering of this component is more probable than upward scattering. The effect of removing the restriction is to extend to greater depths the utility of the model for water depth studies; the incident light is then carried to greater depths by downward scattering of downwelling light intensity.

The ratio of upwelling to downwelling radiation intensity just below the air-water boundary surface of a body of water, depth equal to Z, is described by the function:

$$\frac{I_u(Z)}{I_d(Z)} = \frac{(\alpha_A r_- R_b + \alpha_s A_b)\, e^{\alpha_A r Z} + (\alpha_A r_+ R_b - \alpha_s A_b) e^{-\alpha_A r Z}}{(\alpha_A r_+ + \alpha_s A_b) e^{\alpha_A r Z} + (\alpha_A r_- - \alpha_s A_b) e^{-\alpha_A r Z}}. \qquad (2\text{-}1)$$

The following definitions apply:

α_s = scattering coefficient

α_A = absorption coefficient

R_b = reflectance of the bottom

$r = \sqrt{1 + 2\,\alpha_s/\alpha_A}$

$r_+ = r + 1$

$r_- = r - 1$

$A_b = 1 - R_b$ = absorption of the bottom.

The resulting ratio,

$$\frac{I_u}{I_d} = R_s(Z), \qquad (2\text{-}2)$$

the apparent reflectance of the water body, neglecting surface effects, approaches a limit which depends only on α_S and α_A for large Z. For the case of no scattering, $\alpha = 0$, the expression reduces to:

$$R_s(Z) = e^{-2mZ}. \tag{2-3}$$

The ratio

$$R_s(Z) = \frac{I_u}{I_d} \tag{2-4}$$

has been evaluated as a parametric function of the bottom reflectance and the ratio,

$$A = \frac{\alpha_s}{\alpha_A}, \tag{2-5}$$

of scattering coefficient to absorption coefficient. If numerator and denominator are divided by α_A, these two parameters completely specify the result as a function of the product depth measure, $\alpha_A Z$. This is simply the water depth measured in units of $1/\alpha_A$. Figures 10 to 13 are plotted against Z in these units.

The asymptotic value of $R_s(Z)$ is a function only of $A = \alpha_s/\alpha_A$. This function is also plotted. Since A can be determined as a function of wavelength, Figure 14 is titled "Limiting Value of Water Color."

2.1.1 Solution

The depth measurement, Z, is taken with zero at the bottom of the body of water. The downwelling and upwelling components of light intensity in the water are $I_d(Z)$ and $I_u(Z)$. The effective reflectance just below the surface of water having a depth Z is:

$$R_s(Z) = \frac{I_u(Z)}{I_d(Z)}. \tag{2-6}$$

The same ratio, taken at the bottom, Z = 0, is:

$$R_B(O) = \frac{I_u(O)}{I_d(O)} = R_B = \text{Bottom Reflectance.} \tag{2-7}$$

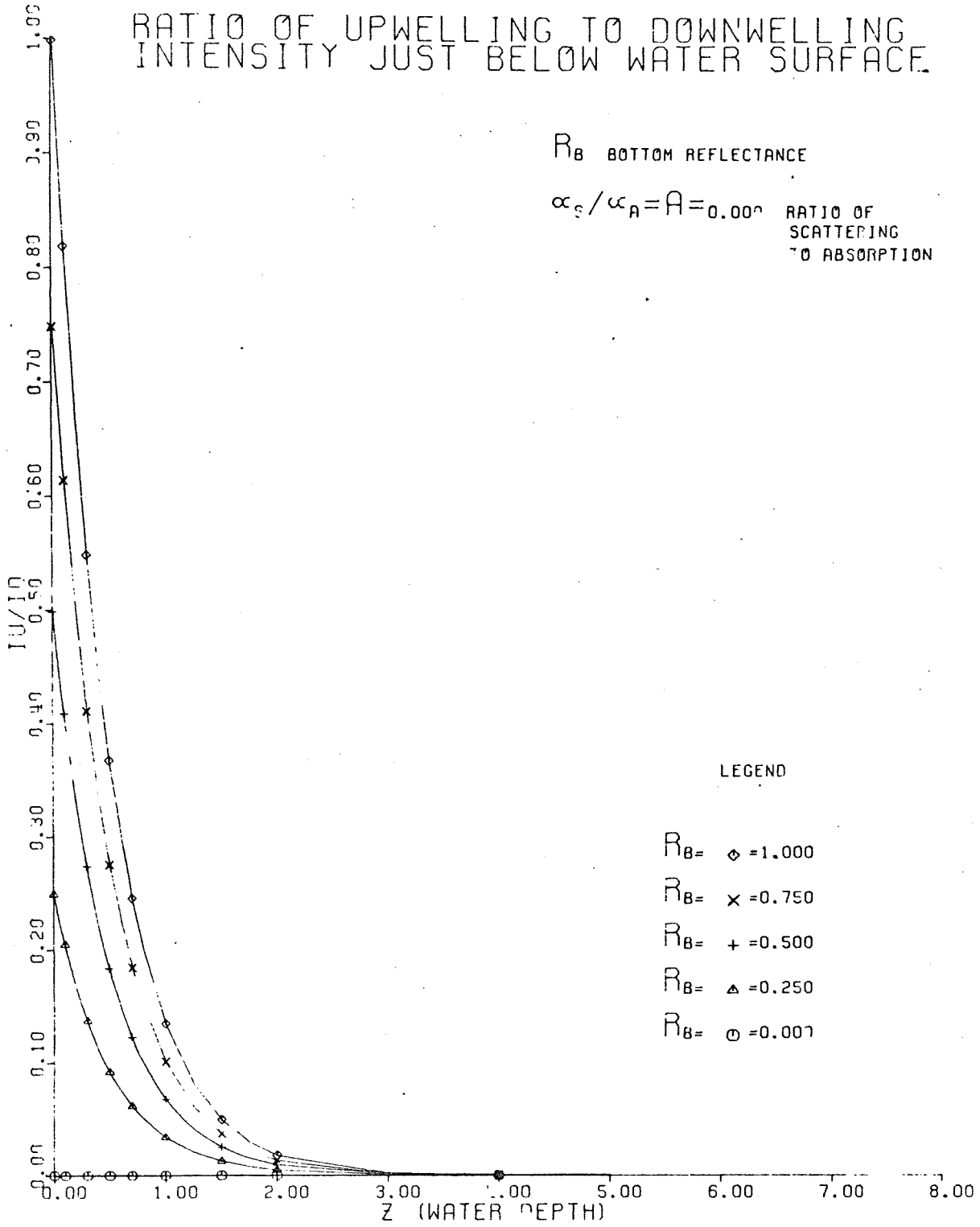


Figure 10 Ratio of Upwelling to Downwelling Intensity Just Below Water Surface (A = 0.0)

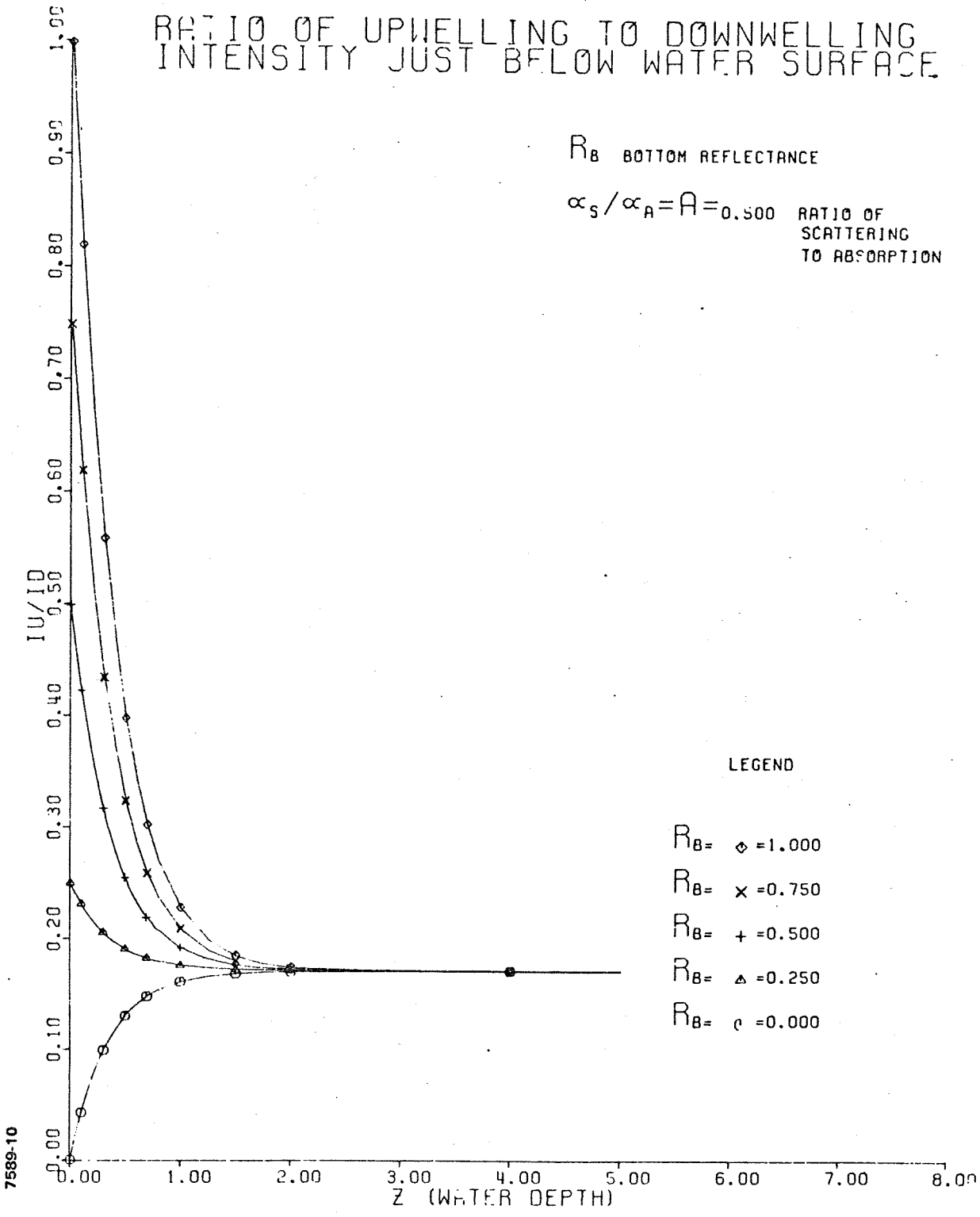


Figure 11 Ratio of Upwelling to Downwelling Intensity Just Below Water Surface (A = 0.50)

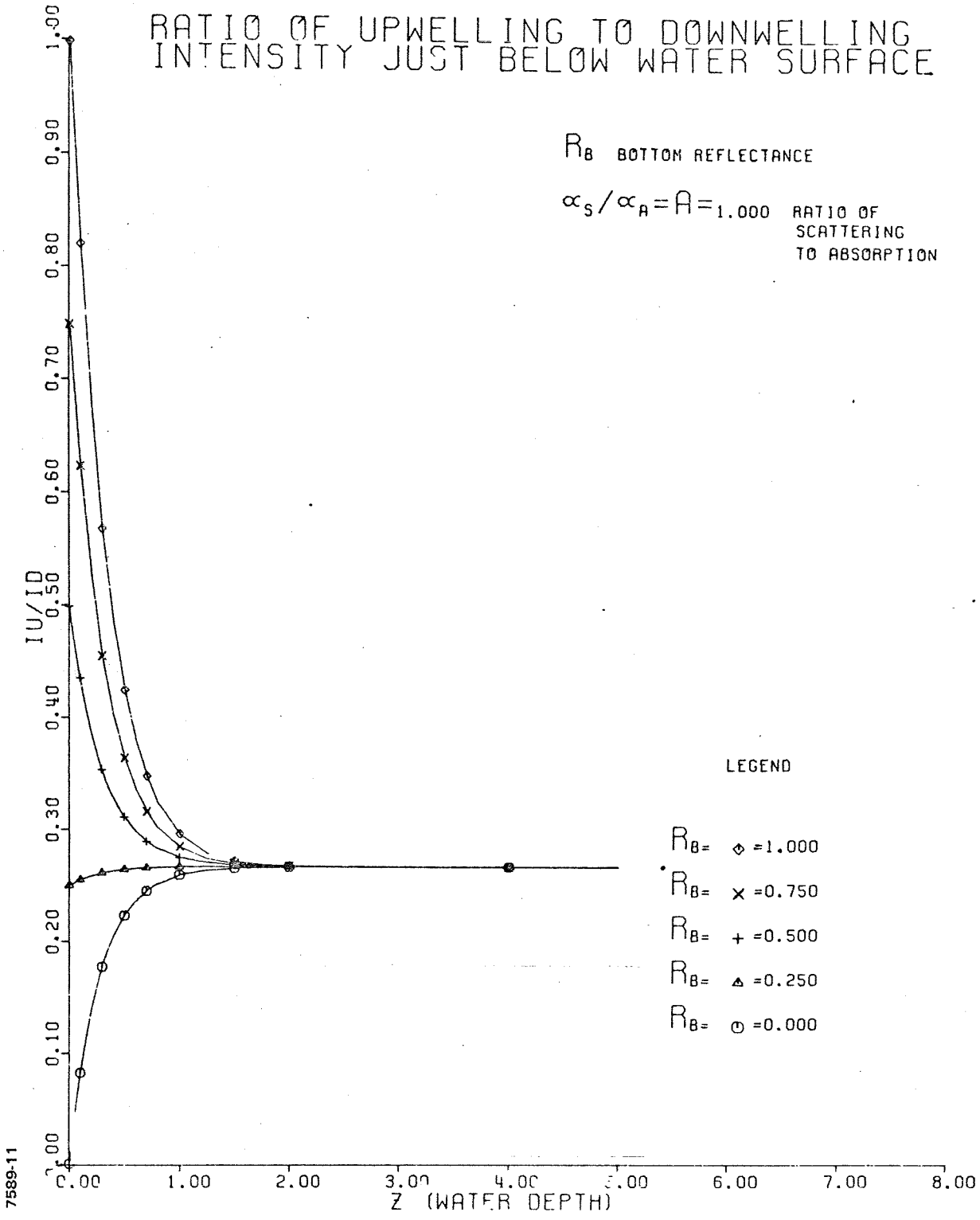


Figure 12 Ratio of Upwelling to Downwelling Intensity Just Below Water Surface (A = 1.0)

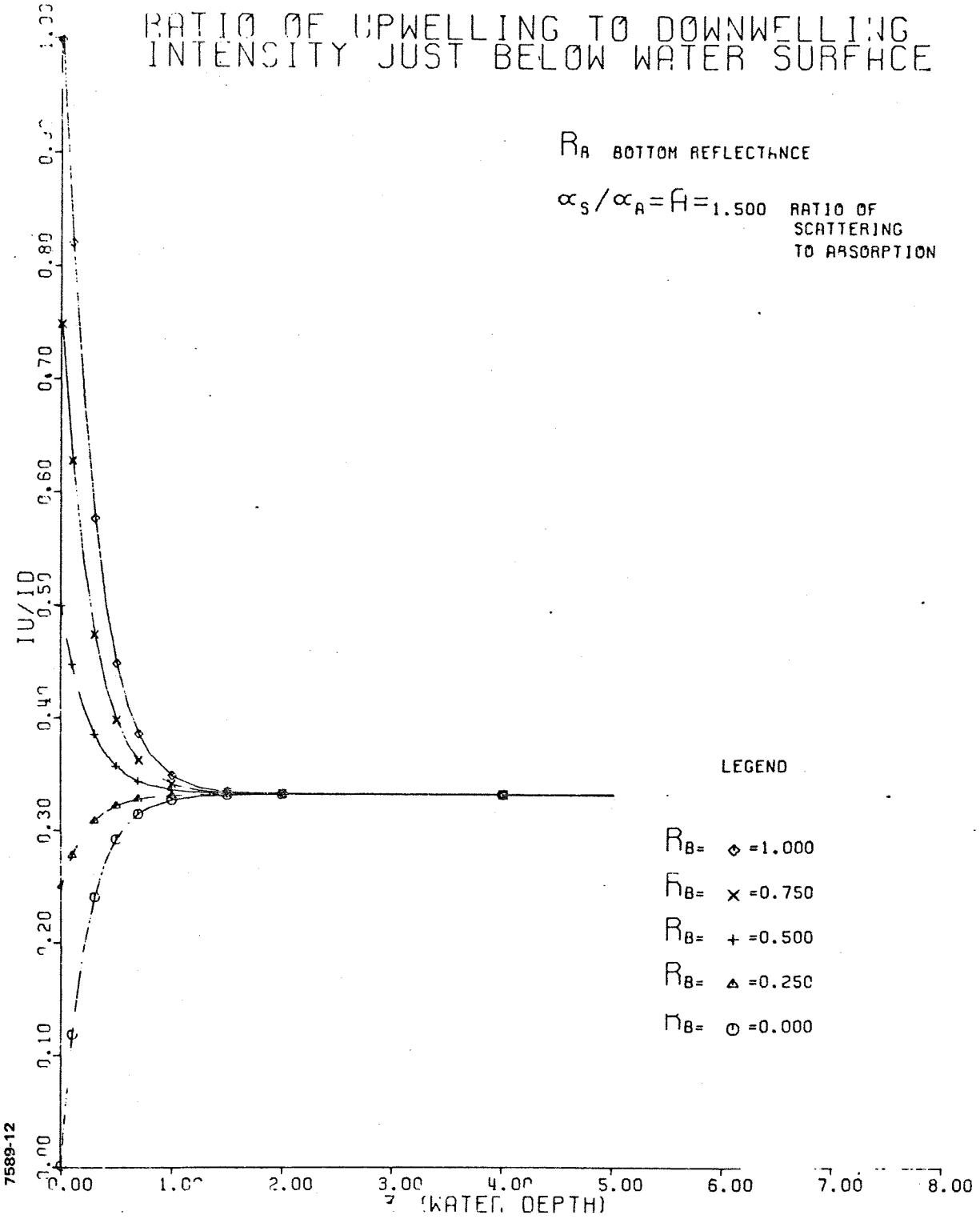


Figure 13 Ratio of Upwelling to Downwelling Intensity Just Below Water Surface (A = 1.5)

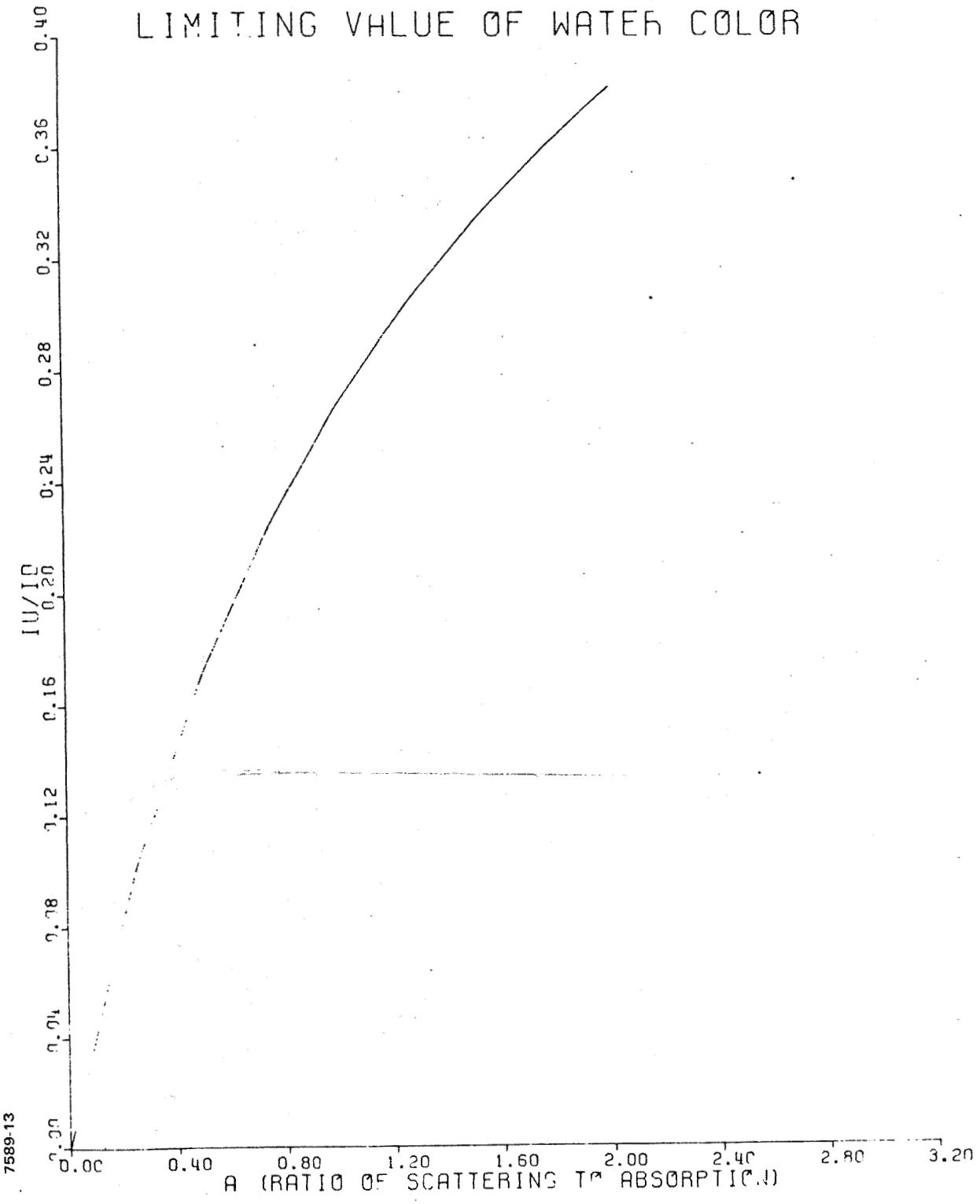


Figure 14 Limiting Value of Water Color

2.1.2 The Equations

The variation of light intensity with depth is due to absorption and scattering. Because the case considered is a medium uniformly illuminated at a plane, components of scattering interactions in directions other than up or down can be disregarded. Scattering in these directions provides coupling between the two equations required to describe separately the downwelling and the upwelling intensity. To obtain these equations, consider the variations in each direction over a thin horizontal layer. With the aid of Figure 16 and remembering that the sign of the variations must be selected considering the origin of the Z axis to be below the horizontal layer, the equations to be solved can be written by inspection.

Consider the contributions to the change in I_d across this horizontal layer. Absorption and scattering decrease the value of I_d moving downward over a negative ΔZ; this provides a net positive contribution to dI_d. Scattering of upwelling radiation into the downward direction increases the value of I_d moving downward over a negative ΔZ; this provides a net negative contribution to dI_d.

Thus:

$$I_d + dI_d = I_d (1 + \alpha_A dZ)(1 + \alpha_s dZ) - I_u \alpha_s dZ. \qquad (2\text{-}8)$$

Conversely, absorption and scattering decrease the value of I_u moving upward over a positive ΔZ; this provides a net negative contribution to dI_u. Scattering of downwelling radiation into the upward direction increases the value of I_u moving upward over a positive ΔZ; this provides a net positive contribution to dI_u.

Thus:

$$I_u + dI_u = I_u (1 - \alpha_A dZ)(1 - \alpha_s dZ) + I_d \alpha_s dZ . \qquad (2\text{-}9)$$

These equations reduce to

$$dI_d = + I_d(\alpha_s + \alpha_A)dZ - I_u \alpha_s dZ \qquad (2\text{-}10)$$

$$dI_u = -I_u(\alpha_s + \alpha_A)dZ + I_d \alpha_s dZ \qquad (2\text{-}11)$$

7589-14

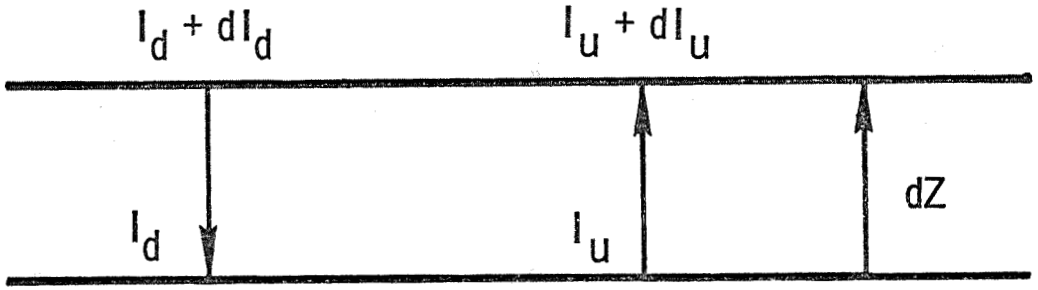


Figure 15 Light Scattering Example

when second-order terms are dropped. The scattering coefficient, α_S, and the absorption coefficient, α_A, when multiplied by an infinitesimal length, dZ, represent the relative likelihood of a particular photon being either scattered into the opposite direction or absorbed. Since the model is restricted to diffuse illumination, α_S is equal for all depths and all scattering directions; a photon is either scattered or not, and if scattered, its direction changes. On a microscale, variations of I_u and I_d in a horizontal plane are zero; the problem can therefore be solved in a single dimension.

The solution to these coupled first-order equations is obtained by standard methods involving conversion to a single second-order equation. The result is:

$$I_u(Z) = C_1 e^{mZ} + C_2 e^{-mZ} \tag{2-12}$$

$$I_d(Z) = C_3 e^{mZ} + C_4 e^{-mZ}. \tag{2-13}$$

In terms of the scattering and absorption coefficients,

$$m = \pm \sqrt{\alpha_A^2 + 2\alpha_A\alpha_S}\ . \tag{2-14}$$

Only two of the four constants are independent. It is convenient, for the solution desired, to assign $C_1 = 1$ temporarily, then solve for those remaining by using three boundary conditions of which only two are independent. The solutions are evaluated at Z = 0, requiring that:

$$\frac{I_u(O)}{I_d(O)} = R_b$$

$$dI_d = I_d(\alpha_s + \alpha_A)dZ - I_u\alpha_s dZ \tag{2-15}$$

$$dI_u = -I_u(\alpha_s + \alpha_A)dZ + I_d\alpha_s dZ.$$

Solutions of the form

$$C_2 = A/D$$
$$C_3 = B/D \qquad (2\text{-}16)$$
$$C_4 = C/D$$

are obtained, suggesting that a better value for C_1 is $C_1 = D$. Then the constants become:

$$C_1 = \alpha_A\left[\left(\sqrt{1 + 2\,\alpha_s/\alpha_A}\right) + 1\right] + \alpha_s(1 - R_b)$$
$$C_2 = \alpha_A\left[\left(\sqrt{1 + 2\,\alpha_s/\alpha_A}\right) - 1\right] - \alpha_s(1 - R_b)$$
$$C_3 = \alpha_A R_b\left[\left(\sqrt{1 + 2\,\alpha_s/\alpha_A}\right) - 1\right] + \alpha_s(1 - R_b) \qquad (2\text{-}17)$$
$$C_4 = \alpha_A R_b\left[\left(\sqrt{1 + 2\,\alpha_s/\alpha_A}\right) + 1\right] - \alpha_s(1 - R_b)\,.$$

The function described in the result section makes use of this solution.

2.2 DIGITAL SAMPLING OF MULTISPECTRAL DATA

Figure 16 is an image of the area known as Flight Line 4. This image corresponds roughly to a photograph with extended-range panchromatic film; it is obtained from the multispectral video collected over this flight line by adding together all eight channels of video. The spectral range extends from 0.38 to 1.0 micron, with spectral resolution as indicated in Figure 2. An optical schematic of the scanner used to collect these data is also shown in Figure 2.

Figure 17 is a series of eight images showing Flight Line 4 in each of the eight spectral channels of the scanner. The western shore of Lake Michigan has a series of parallel sand bars which characterize the structure of the bottom from Chicago to Traverse City. These sand bars are most easily visible in the channels within which visible light most easily penetrates the water and is reflected from the bottom. The mathematical relationship of the water, bottom, and light interactions is discussed in Section . The direction of flight over this line was to the west (from right to left in the illustration). The first step necessary in accomplishing

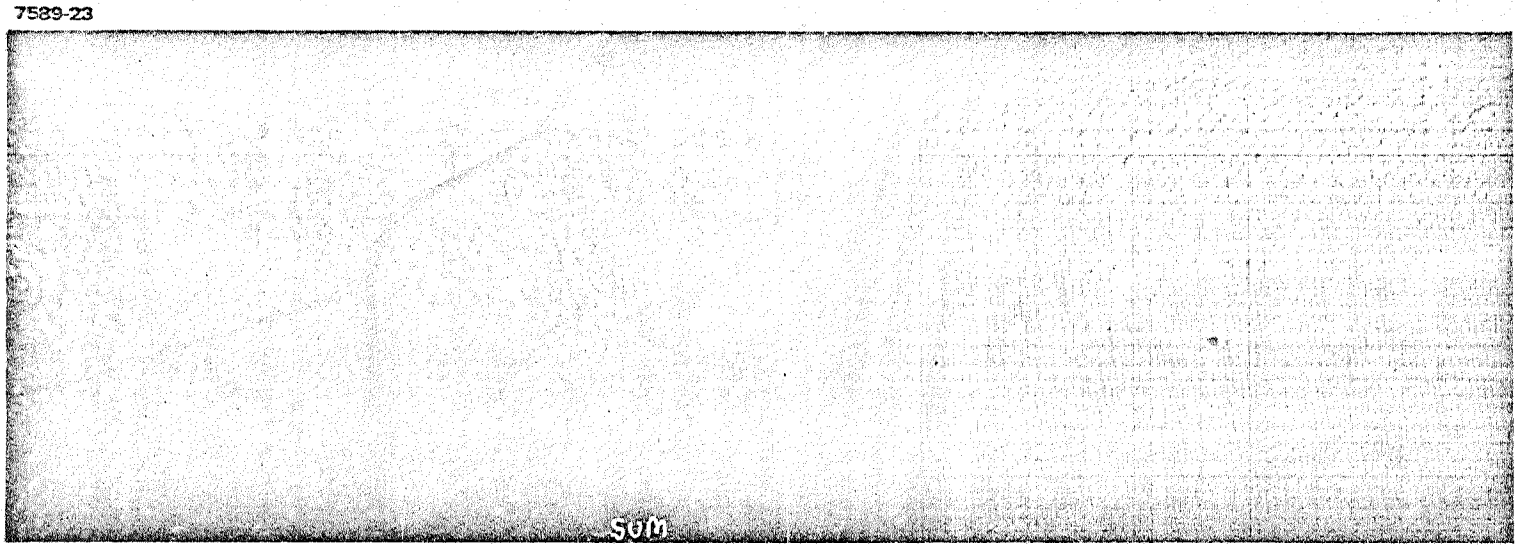


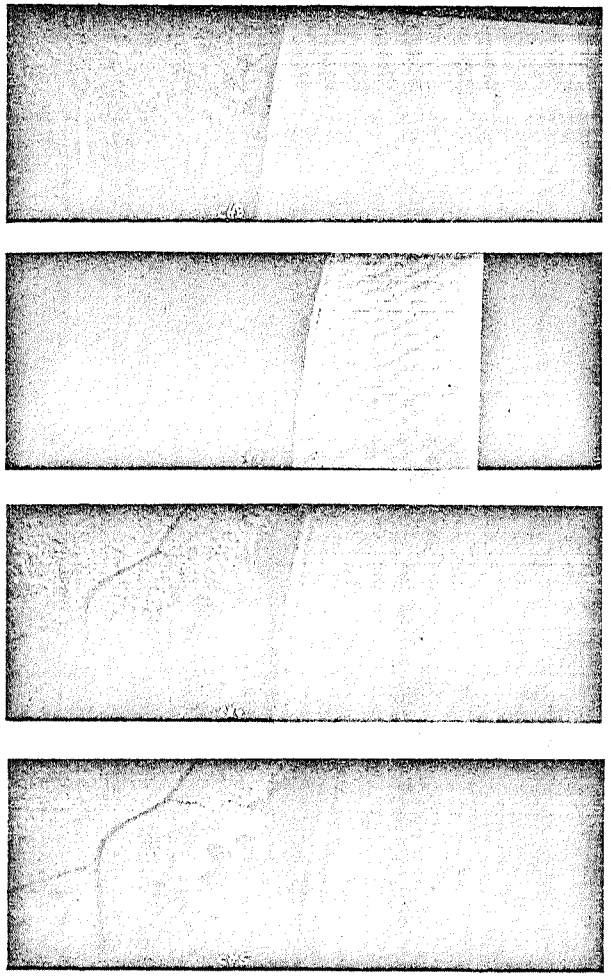
Figure 16 Imagery From Flight Line 4 (Sum)

the data processing is to extract digital data from the analog tapes. The playback and film recording equipment used to produce imagery from the tapes includes a multichannel A/D converter and a digital tape deck. The A/D converter is used to select, sample, and record spectral signatures of the interesting areas of the scene. The A/D converter is used with the film recorder in operation; the imagery produced at this time is electronically identified in such a way that each signature in the computer listing can be reliably identified with the image feature which produced it. The digital tape thus produced is the link that permits communication of the flight recorded analog data with the computer. The procedure is summarized in Figure 18.

Figure 19 is a sample of imagery produced on the film recorder during the digitizing operations. Note the dark black cursor line at either end of the frame. This is produced by a video blanking signal superimposed on the imagery during the 5-μ sec sampling period used to obtain the digital signature; when the user is satisfied with the vertical position of the cursor line, the sampling is initiated by a pushbutton which also removes the blanking signal. Digital samples are recorded on the tape transport at the vertical position indicated by the cursor line, once for each scan line.

The horizontal density of samples and scan lines is estimated from the aircraft speed and rotational speed of the scan shaft on the scanner. During the flights of interest, the aircraft speed of 150 mph (statute) and the scan shaft speed of about 6000 rpm give an estimated horizontal density of 2.2 ft per sample. The terms sample and case are used interchangeably, and the statistical results are often plotted as a function of case number.

Notice in Figure 17, that the sand bar structure of the Lake Michigan bottom is most easily visible in Channels 3, 4, and 5, is barely perceptible in Channels 2 and 6, and has no effect at all in Channels 1, 7, and 8. The individual channel reflectances are plotted in Figures 20 through 27. Prior to automated electronic systems, a reflectance graph could be made by utilizing densitometric techniques. A severe dependence on the quality of the collected photographs limited the extent of data extraction. In this study the situation is reversed. The imagery is generated from the reflectance information and depends primarily on controllable laboratory parameters. The reflectance graphs of the sampled information show little relation to the lake bottom contour in the ultraviolet and near-IR data.



7589-22

Figure 17 Eight Channel Images of Line 4

7589-1

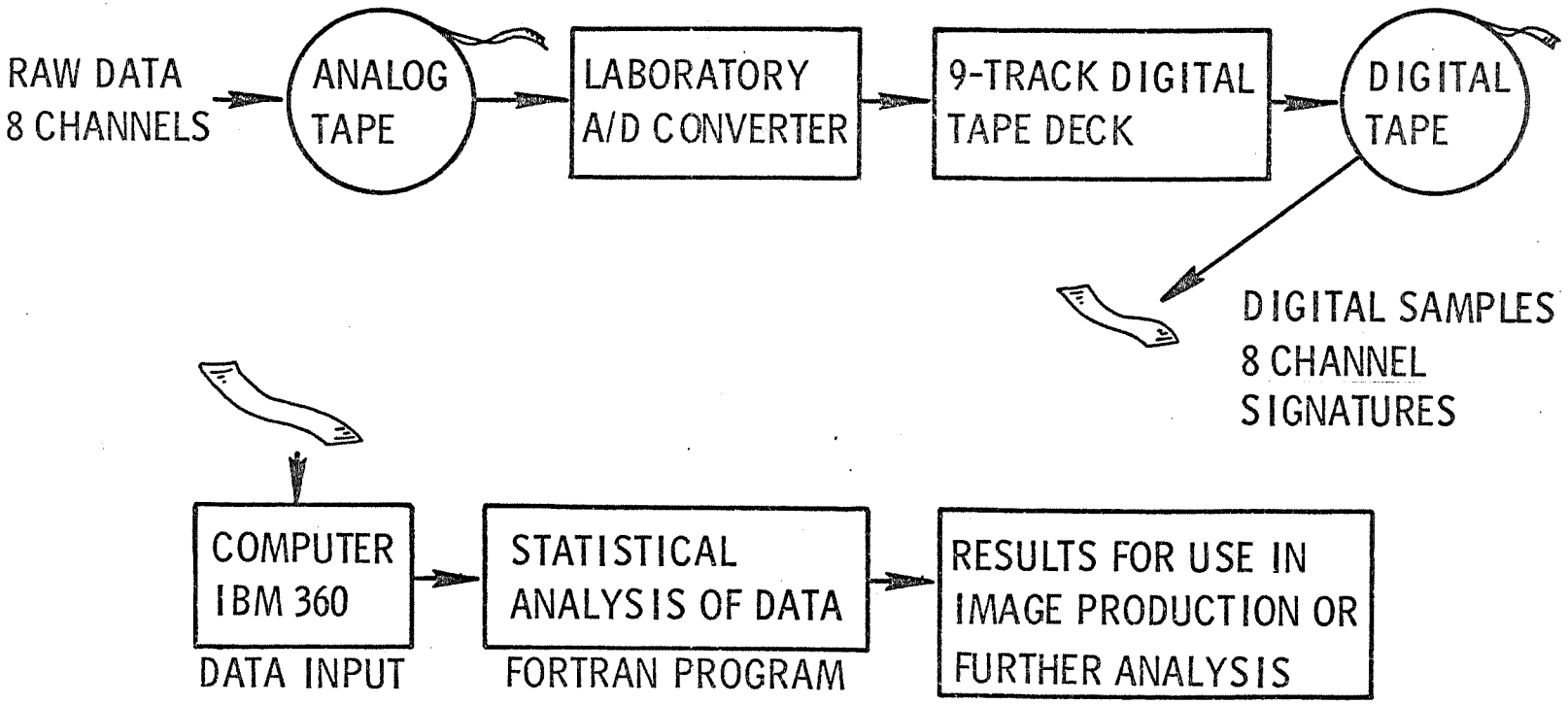


Figure 18 Data Analysis Procedures

7589-26

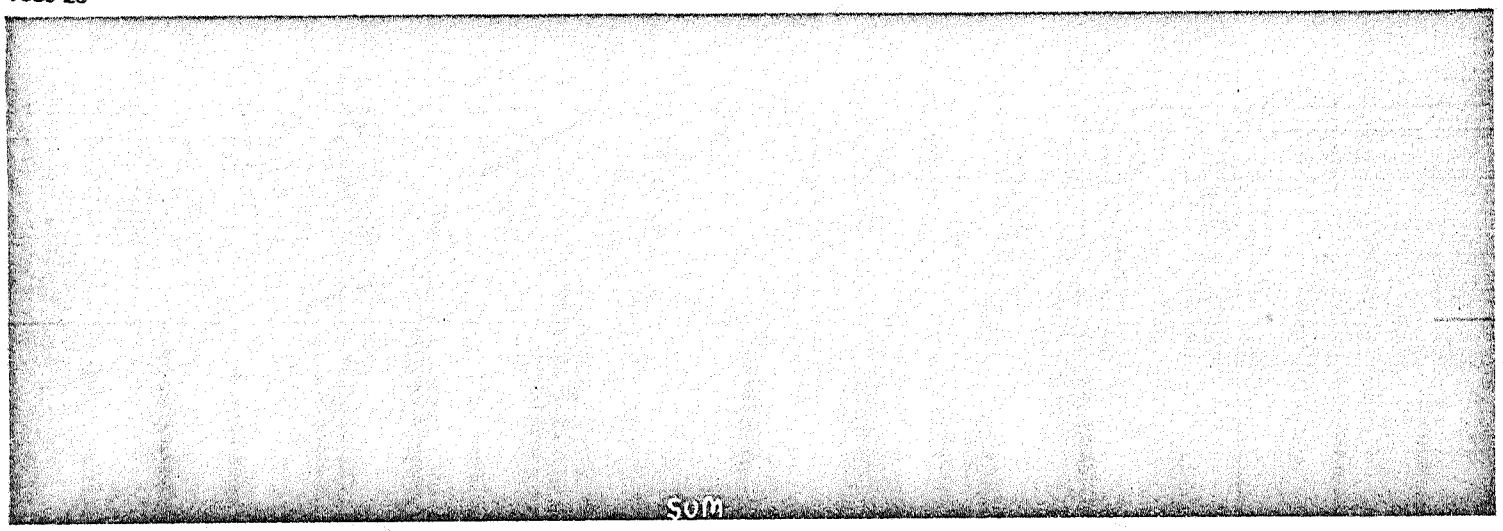


Figure 19 Digitizing Imaging With Cursor

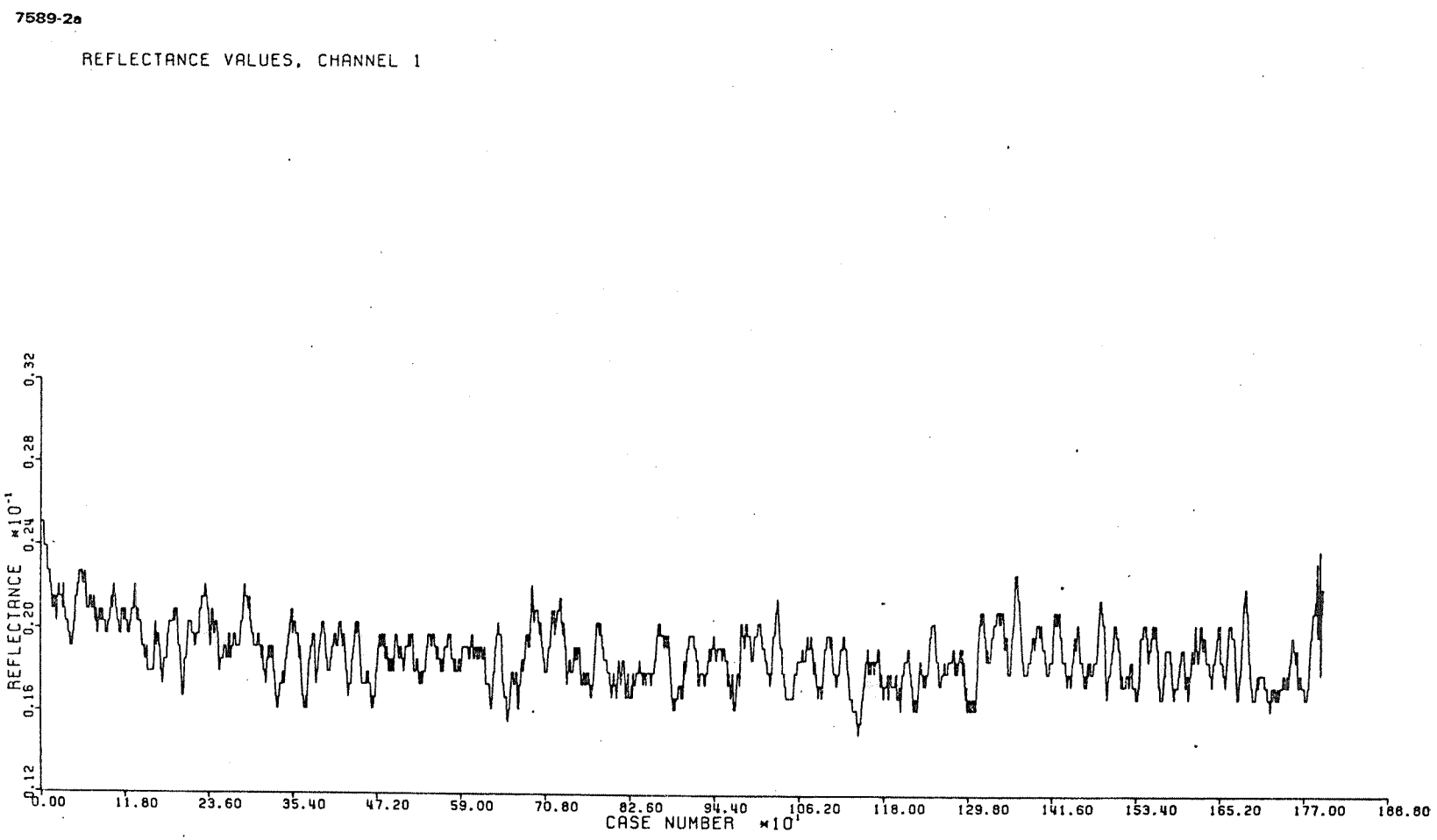


Figure 20 Individual Channel Reflectances

7589-2b

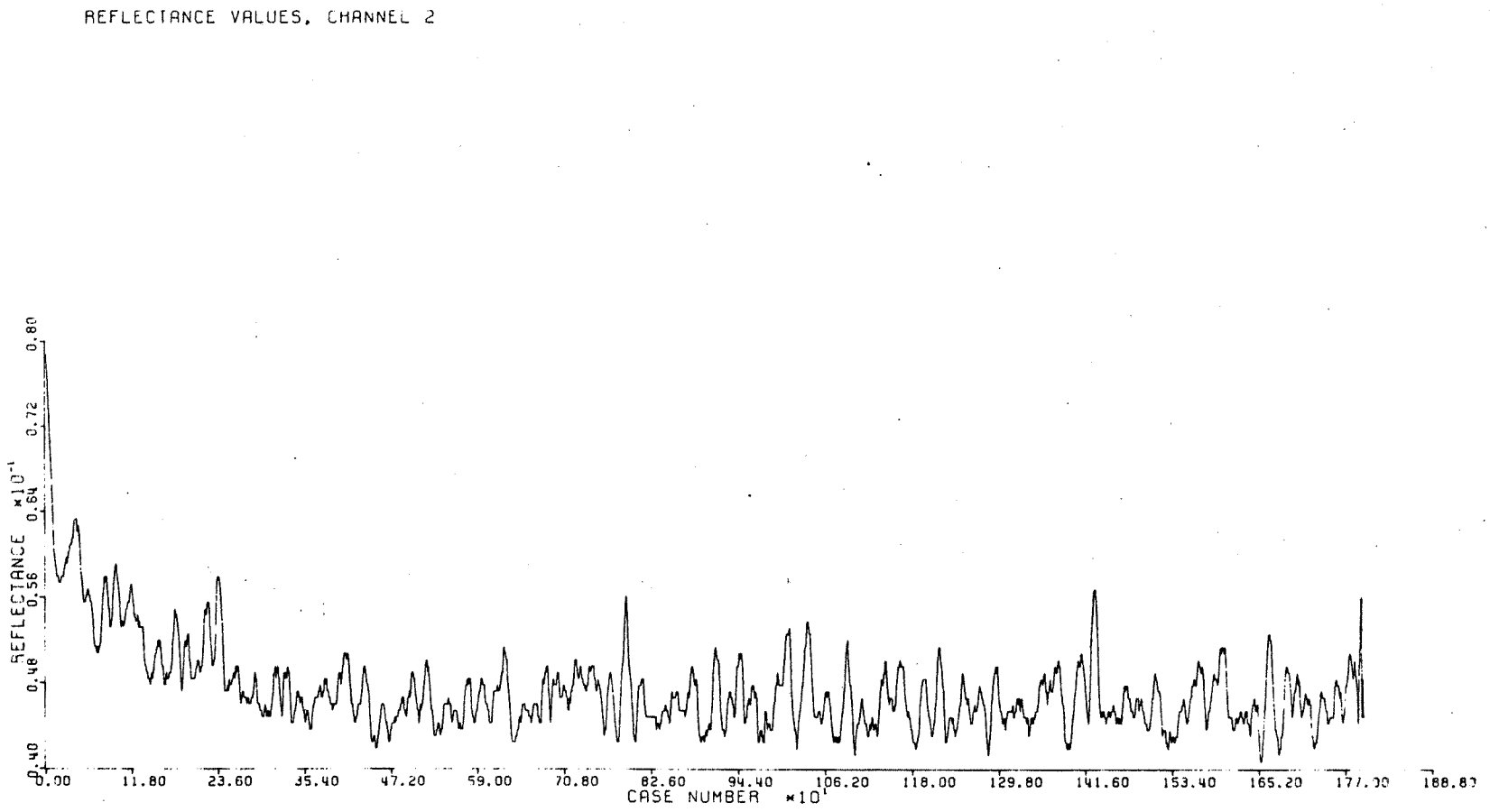


Figure 21 Individual Channel Reflectances

7589-2c

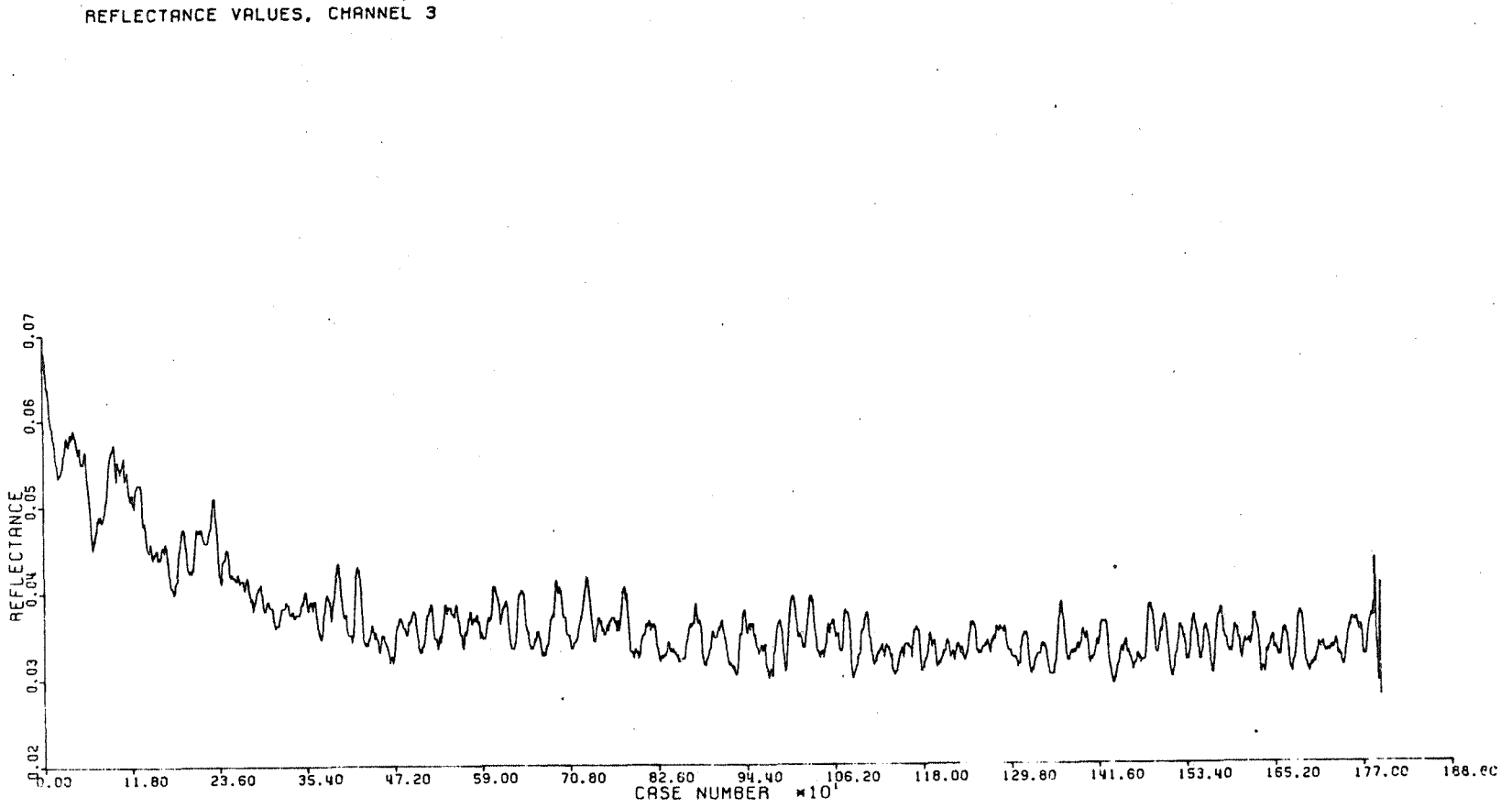


Figure 22 Individual Channel Reflectances

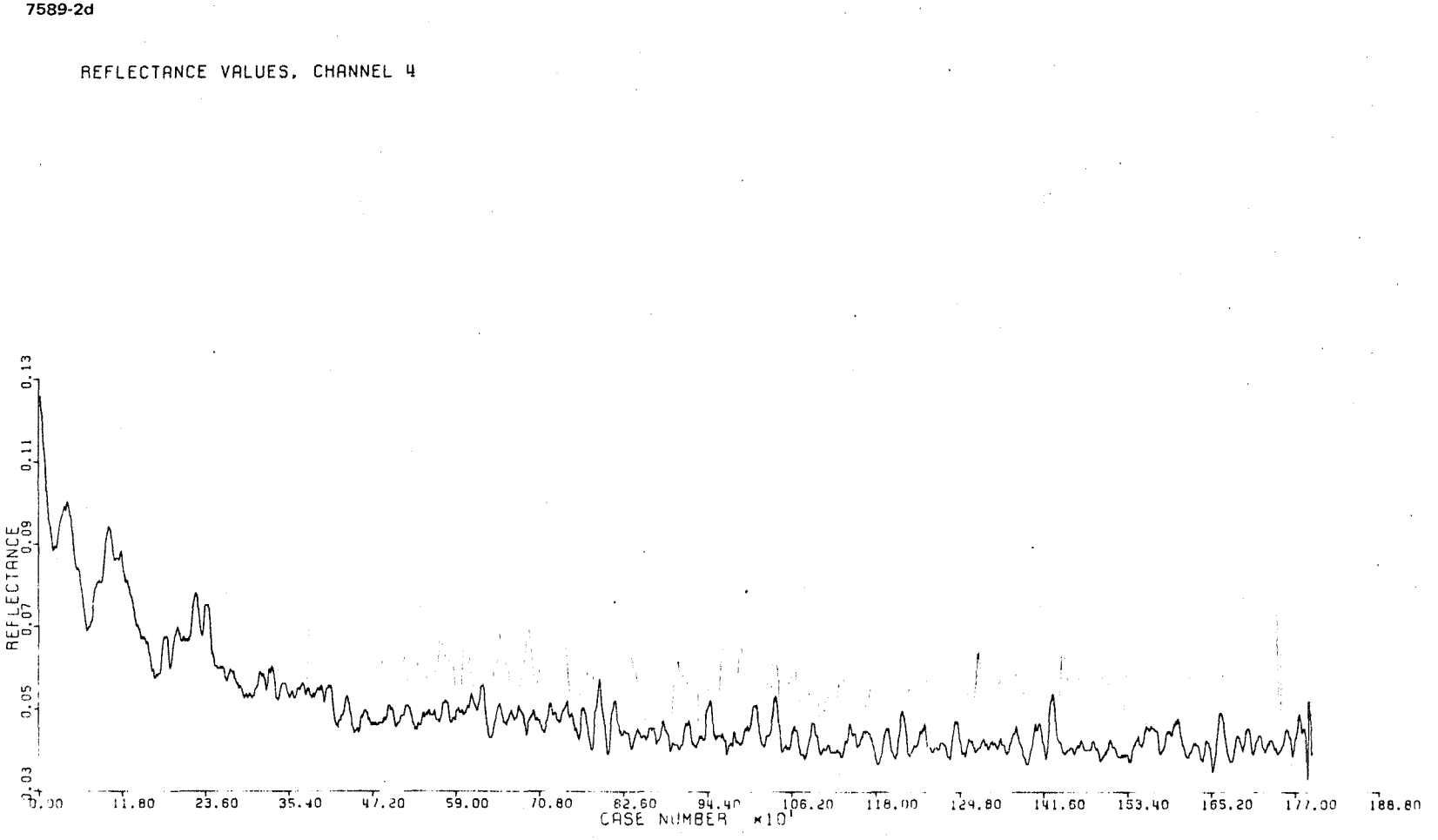


Figure 23 Individual Channel Reflectances

7589-2e

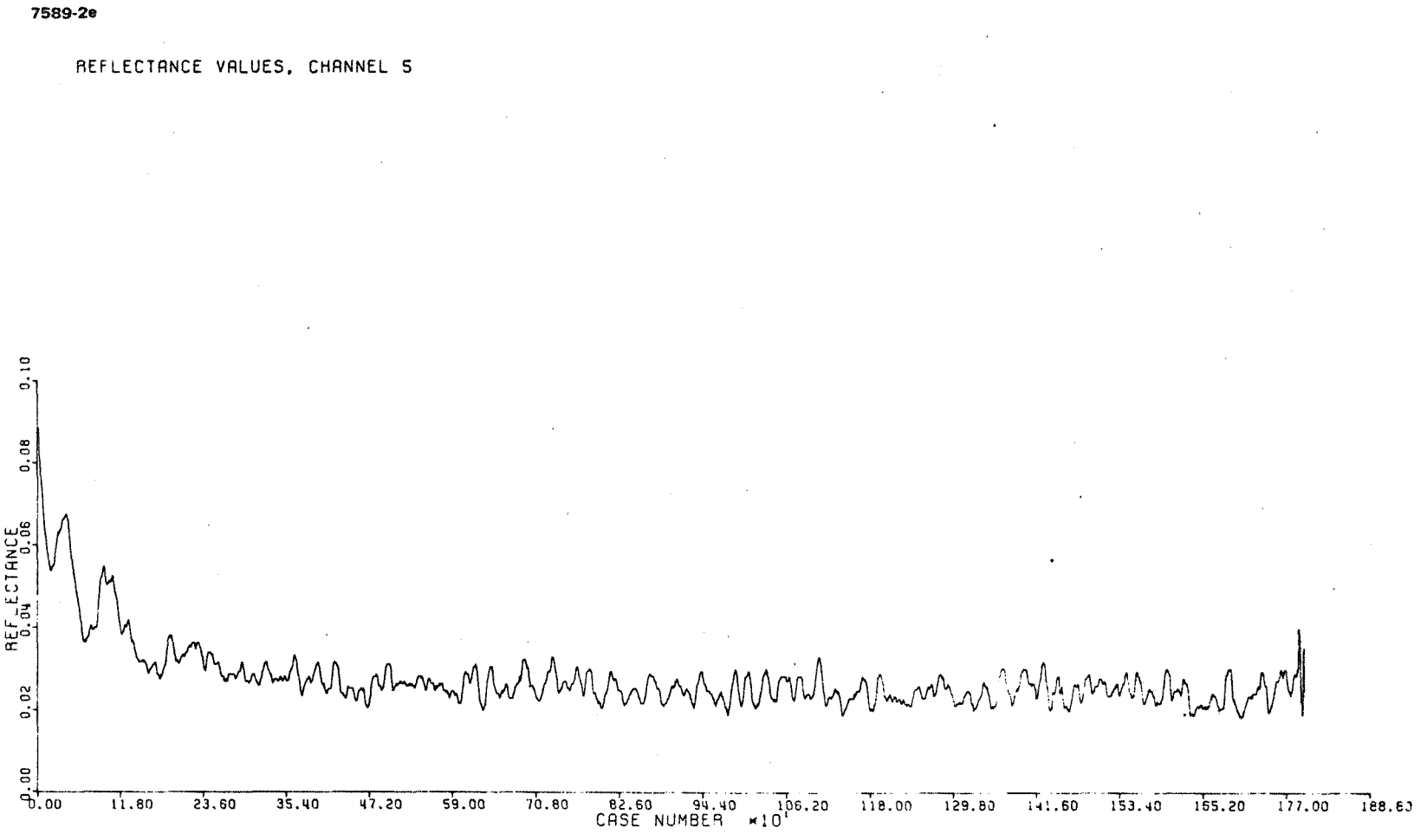


Figure 24 Individual Channel Reflectances

7589-2f

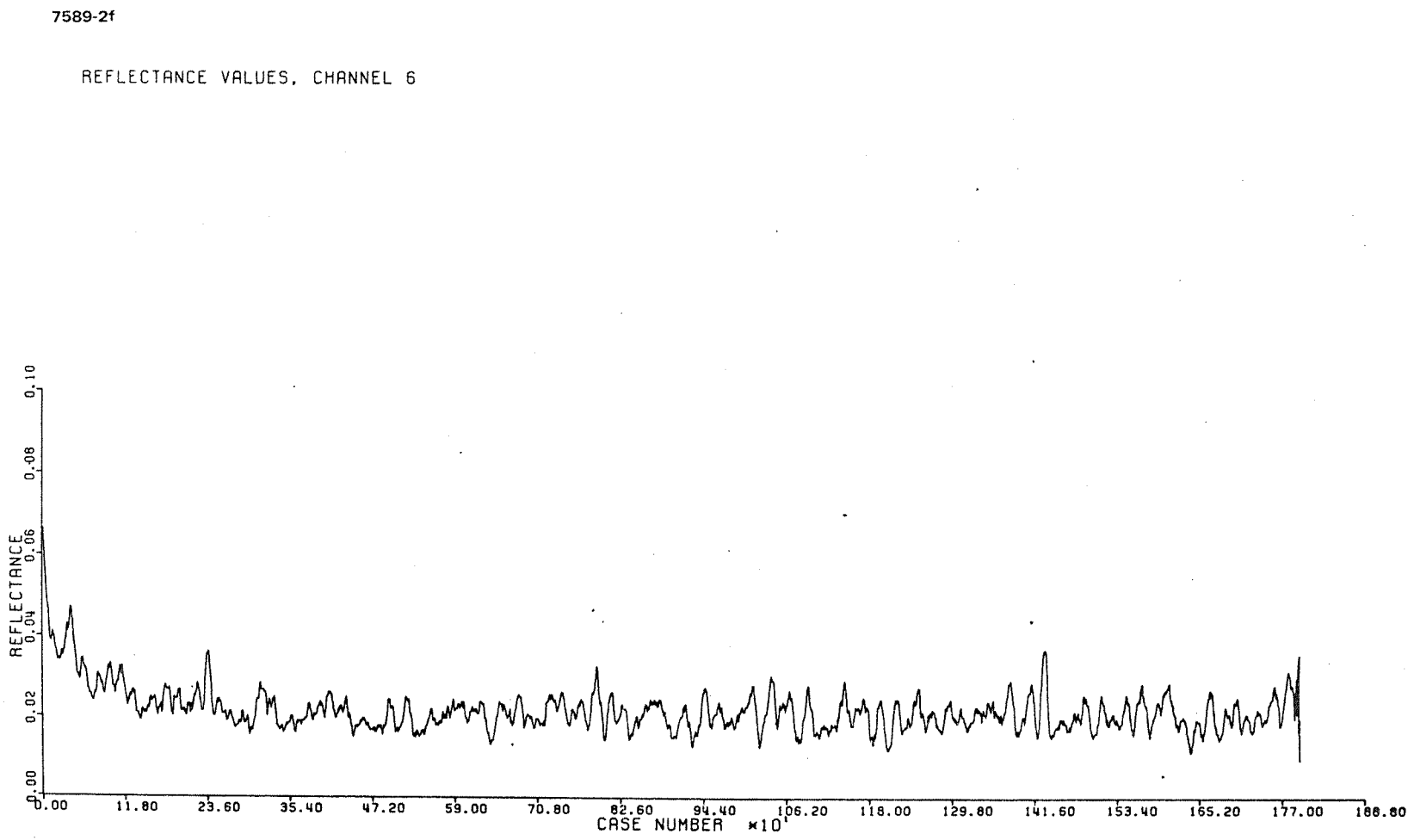


Figure 25 Individual Channel Reflectances

7589-2g

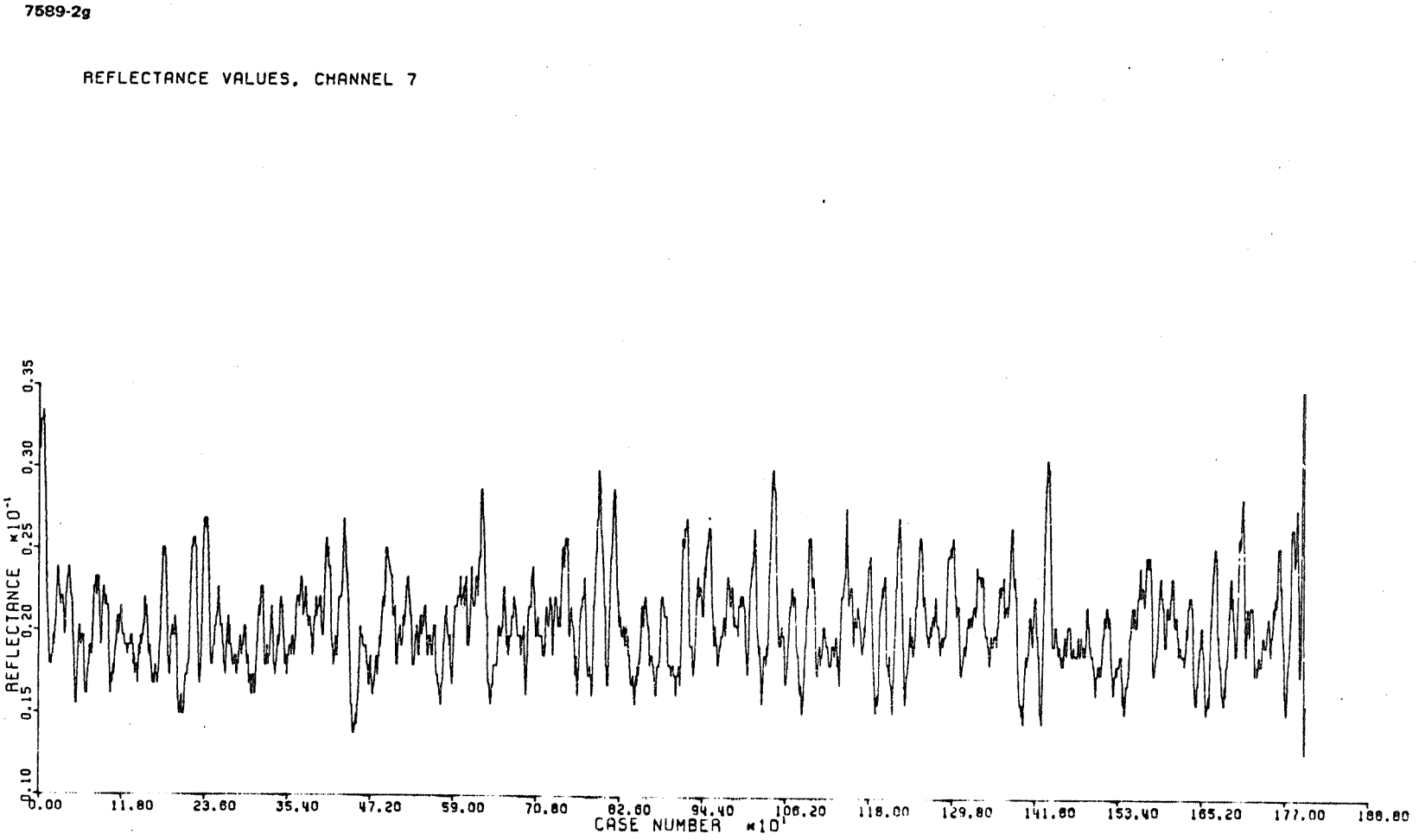


Figure 26 Individual Channel Reflectances

7589-2h

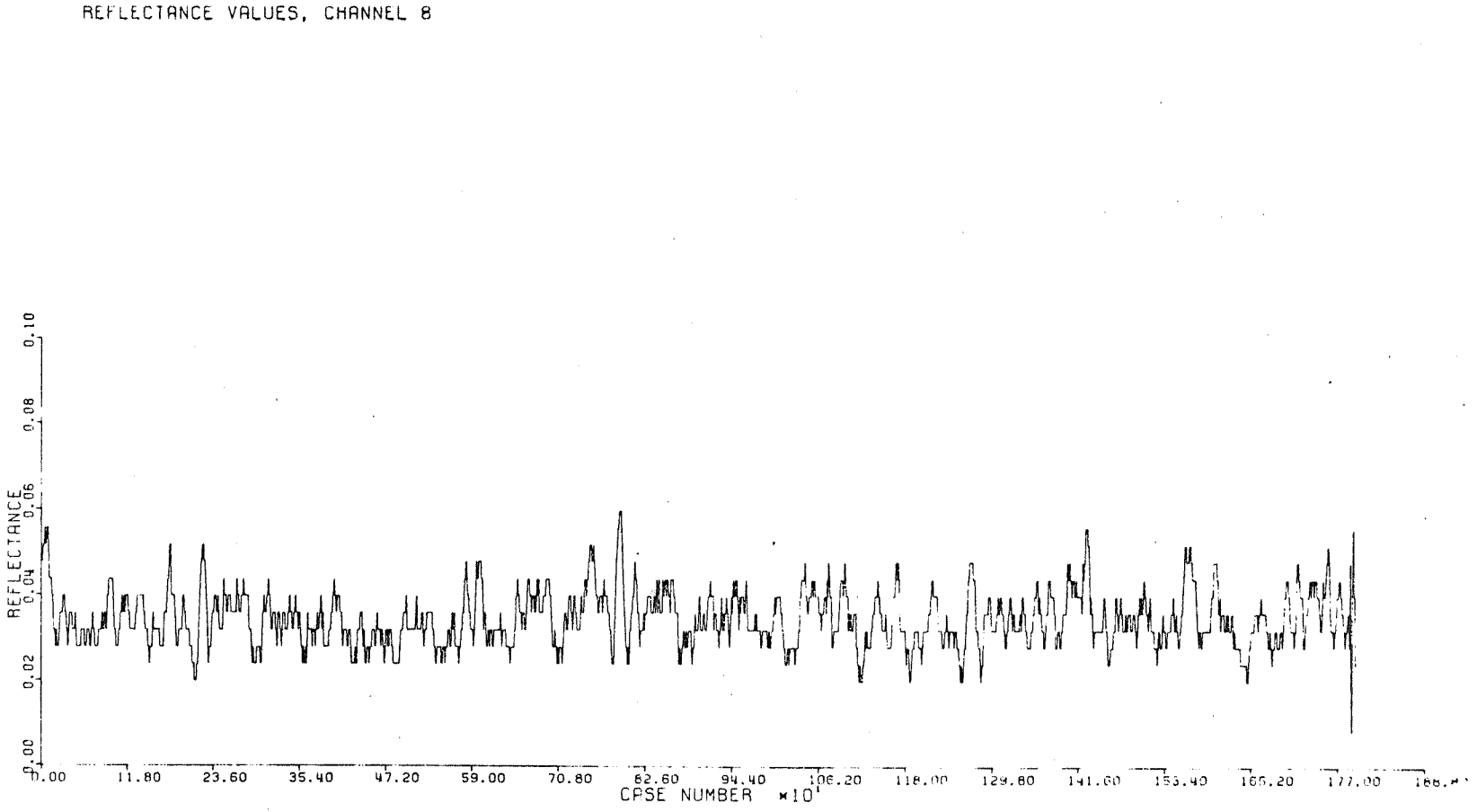


Figure 27 Individual Channel Reflectances

2.3 FACTOR ANALYSIS

Digital samples collected, as previously indicated, are stored in the IBM-360. Selection of particular cases for study is readily accomplished. The following paragraphs describe the use of a factor analysis program in extracting information from the stored reflectance measures. In the present problem, an exact variable is sought from the collected data; factor analysis is evaluated to isolate water depth as that variable.

The primary analysis technique used in this study of multispectral data is factor analysis.* As one multivariate approach, principal component analysis** is used to locate underlying trends in collected measures. In the present study, the measures are of spectral reflectance in the ultraviolet, visible, and near-IR regions. The new variables (factors) sought may represent various physical phenomena. Factor analysis is used to emphasize differences in the 8-channel spectral data sampled from the collected analog tapes. When these differences are properly emphasized with factor analysis, desired characteristics are automatically enhanced.

There are several excellent reasons for deriving new variables from digitized scanner data. The several channels from the scanner are not independent. Generally, they possess high first-order correlation. The ranges of variation observed in different channels are unequal even though the terrain observed is the same for all channels. The scanner data also contain noise which confuses their interpretation.

An examination of the bandpass of each of the eight channels (see Figure 28) reveals a certain degree of overlap. This affects the eight-dimension reflectance measure by adding to any existing correlation between channels. The effect is greatest in adjacent channels, but may be evident over a wider range of wavelengths. A two-dimensional plot of digitizer values for two channels reveals a linear trend. Factor analysis derives variables that are uncorrelated. These effects are illustrated by the scatter diagrams of heterogenous data in Figures 29 and 30. Notice the obviously high correlation in the adjacent channel pairs scatter diagrammed in Figure 29. The nonadjacent pairs do not possess such high correlation. In contrast, there is a lack of correlation in the factor pairs of Figure 30. Furthermore, in Figure 30 the clustering of points is most obvious in the Factor 1 and 2 pair, and nearly disappears for higher numbered pairs. In the plots of channel pairs, less distinct groups may be found in many of the scatter diagrams.

---

*, **Appendix B describes the formal mathematical properties of factor analysis, both the principal component analysis and the operations known as rotations.

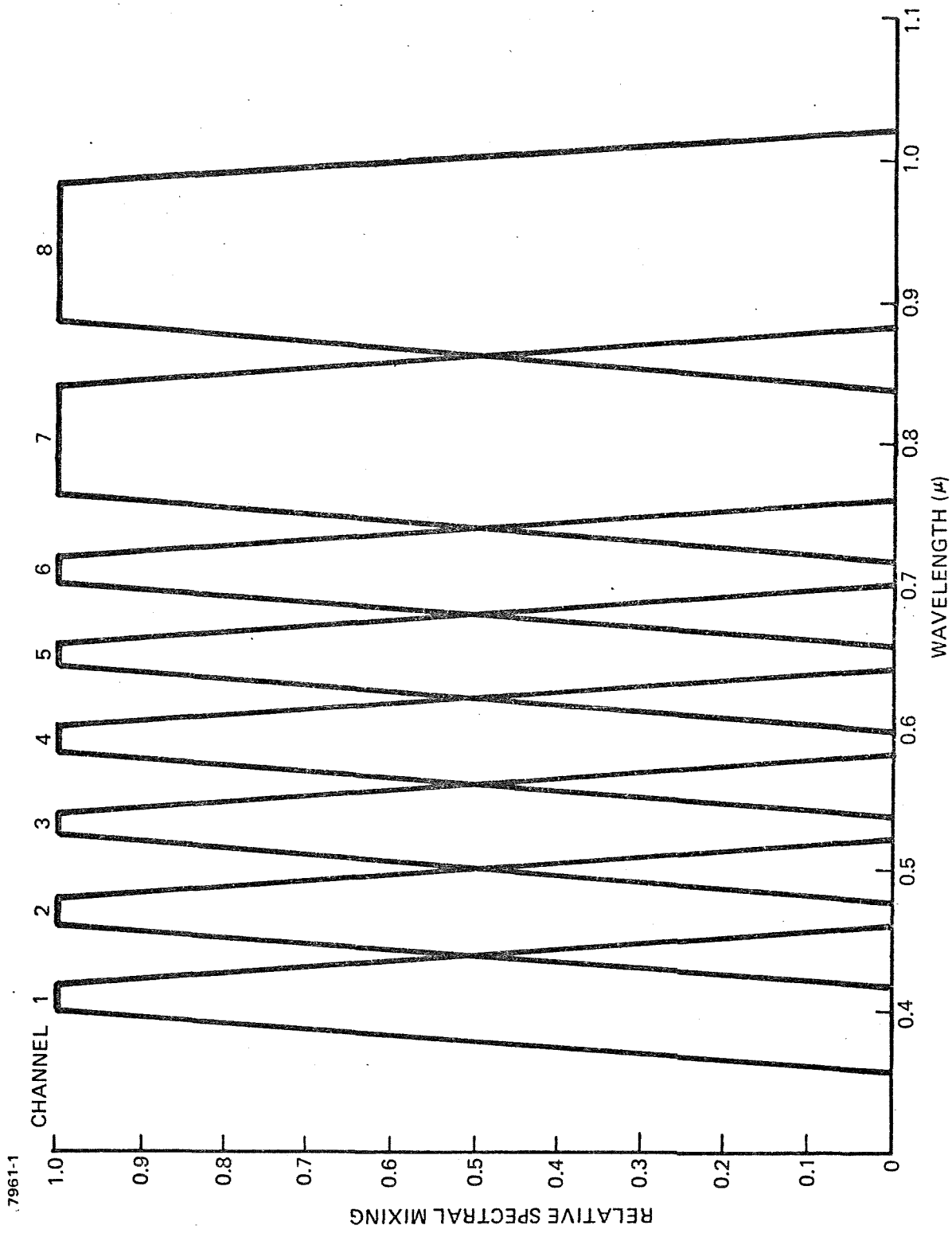


Figure 28 Channel Profiles

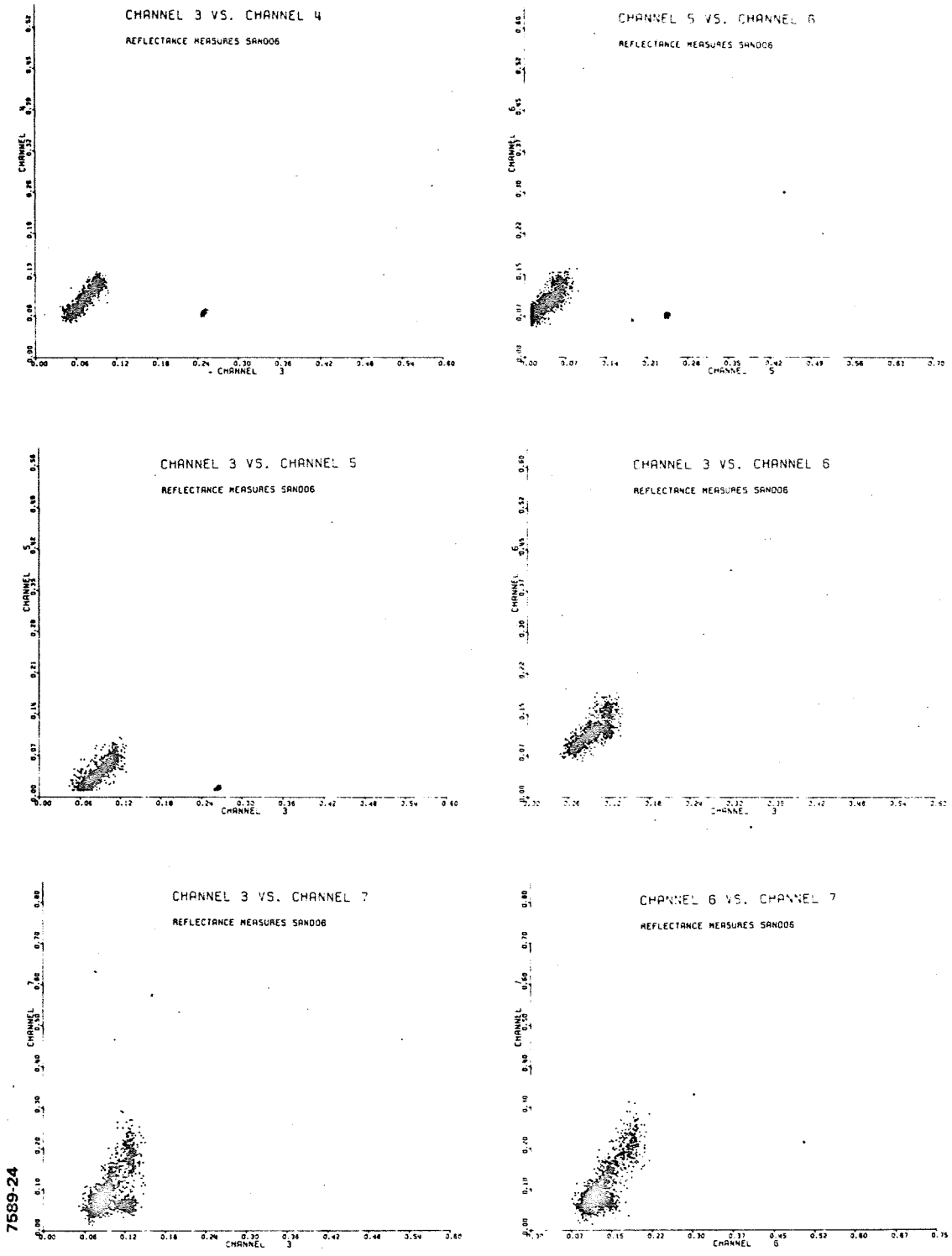


Figure 29 Channel Pairs Scatter Diagram

7589-25

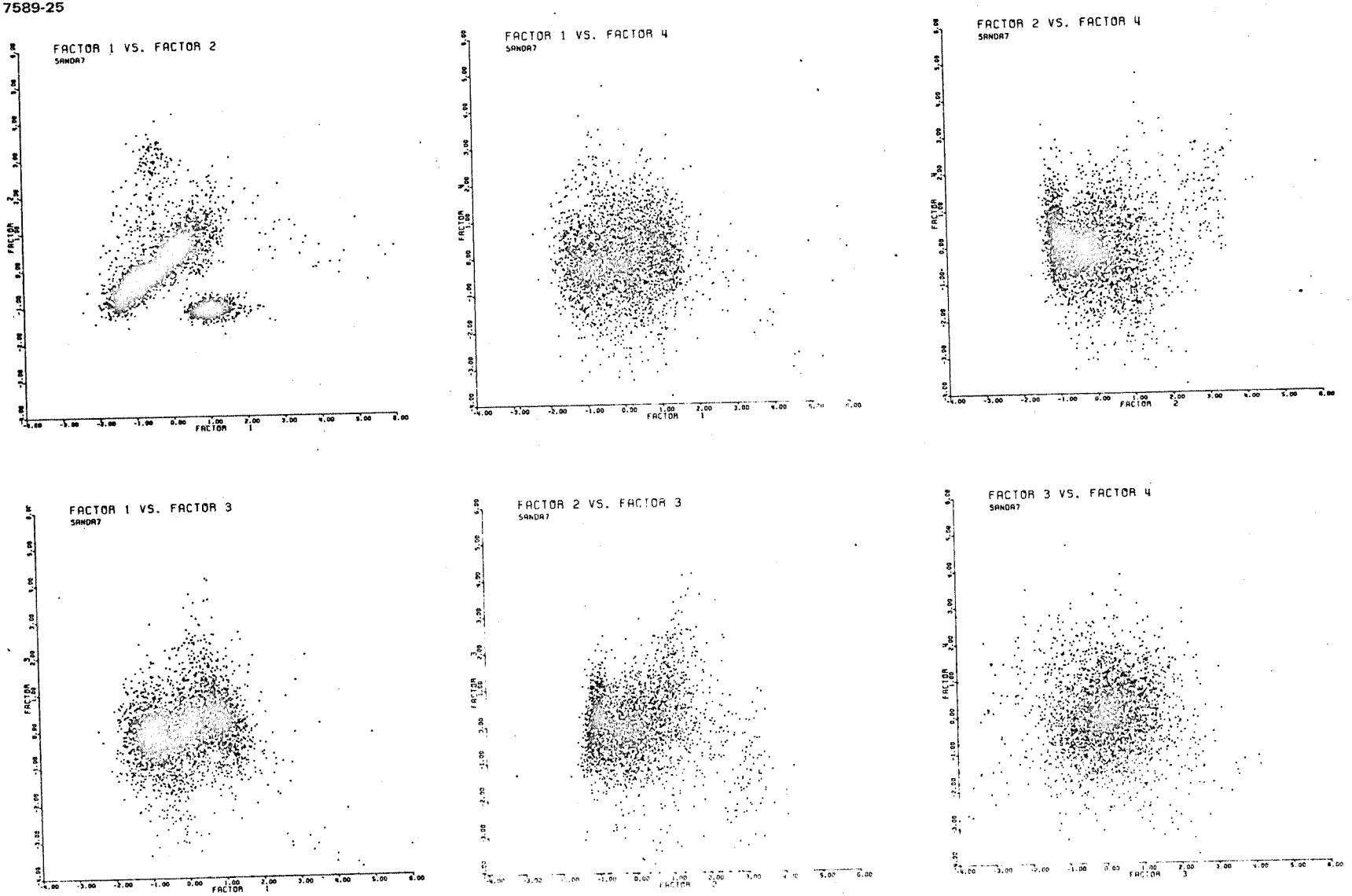


Figure 30 Factor Scores Scatter Diagram

This illustrates the utility of factor analysis; data interpretation is simplified by bringing the important correlatable differences into a smaller number of variables.

The reflectance ranges used to calibrate the scanner during the Lake Michigan flights are listed in Table I. These were derived from previous water flights. Optimally, the calibration scheme would use the full dynamic range of all channels. Typically, the ratio of variation between channels for some digital samples may be as great as 4:1. The factors resulting from the analysis all possess unit variance.

As in most electronic devices, there is noise (introduced by both the measuring device and the real world) in each channel. The presence of noise increases the difficulty in drawing conclusions based on digitized sample [illegible]. A technique that minimizes the effects of noise by taking into [illegible] rchannel correlations would prove very useful. Factor analysis [illegible] technique.

In addition to alleviating some of the above difficulties, factor analysis reveals several relevant characterististics of the observed measures. The variables resulting from a factor analysis, known as factor scores, possess zero means and unit variance. The principal component factor analysis is most useful for this application because it expresses the variance of the measured data in terms of new directions. The analysis associates the direction of maximum data variation with the first factor. The direction of second greatest variance is isolated as the second factor, and so forth, until all the variance has been redistributed among a new set of orthogonal, or linearly independent, variables. Several other useful phenonena may then be discovered. Usually, most of the information will appear in a fewer number of factors than the number or original measures or one factor may describe or isolate the noise in the data. Briefly, the use of factor analysis may improve the signal-to-noise (S/N) ratio as well as diminish the number of variables needed to characterize particular targets.

The development of the factor analysis program used in this study stemmed from work contained in the IBM Scientific Subroutine Package (SSP). A double-precision version of the SSP program was found to be more accurate when compared with several other available routines. The test for accuracy was made by generating data containing known factors and submitting this synthetic data for analysis.

The use of the class of transformations called Hadamard matrices permits a priori knowledge of the principal component solution. Such a matrix when multiplied by the appropriate constant defines orthogonal eigenvectors. The product of this matrix with a diagonal matrix of eigen-

TABLE I

CALIBRATION FOR LAKE MICHIGAN FLIGHT

Channel	Reflectance
1	0 - 15%
2	0 - 15%
3	0 - 15%
4	0 - 15%
5	0 - 15%
6	0 - 15%
7	0 - 15%
8	0 - 100%

values ($1/\sqrt{\lambda}$; along the diagonal) yields a factor matrix. The factor matrix transforms normalized random numbers into channel readings. By beginning with the directions (eigenvectors) and amounts (eigenvalues) of data variation, synthetic data are generated.

The double-precision version of the SSP program was expanded for additional versatility. The developed program is capable of numerous rotation schemes, some of which are discussed later. The input/output format was generalized to permit complete identification of data flow. The program printouts were detailed to provide simpler interpretation of the various steps printed. After testing, the program was used, as described, in the present study.

A number of experiments were performed using the factor analysis program. Initially, a principal component analysis was performed on sets of reflectance measures. This involved the diagonalizing of a correlation matrix formed from the raw data. Alternatively, it is possible to perform this operation on the covariance matrix. Since the variance in each channel is related to the calibration scale used for the flight, standardizing the channel readings removes some arbitrariness. The process of diagonalizing a correlation matrix results in a number of factors equal to or less than the number of input measures. The eight factors resulting from a principal component analysis of multispectral data, about the means and origin, are shown in Figures 31 and 32.

The definition of directions of successively smaller amounts of data variance by principal component analysis was discussed earlier.

In such an analysis of sampled water reflectances, Factor 1 represented 46% of the total variance in the collected data. Factor 2 accounted for 20%; each of the third and fourth factors accounted for another 10% of the variance. The remaining variables contributed the remainder of the variation (< 15%).

A plot of the individual channel readings (Figure 31) when compared with a graph of Factor 1 (Figure 32) demonstrates the redistribution of data variation. No single reflectance band reveals a similarity to the lake bottom as does the resulting new variable. The principal component analysis represents, however, only a cursory use of factor analysis as a multivariate analysis tool.

In an effort to relate factors to physical phenomena, a simplification of their form is undertaken. A factor may be viewed as a direction in measure space. Ideally, this direction may be explained by some characteristic of the observations. Frequently, the coefficients derived in the analysis reveal a complex combination of the reflectance channels. To simplify the structure

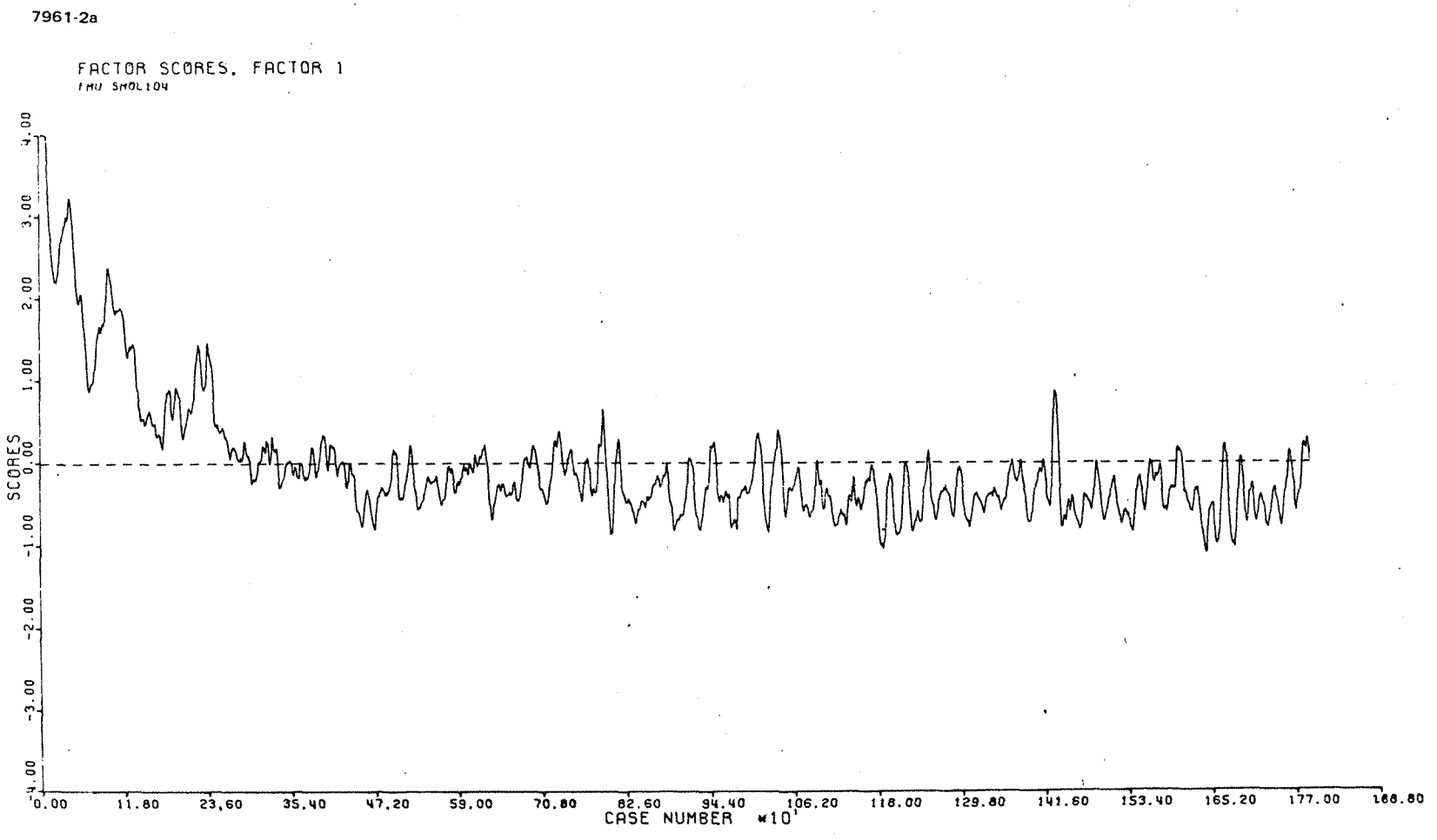


Figure 31 Eight Factors Resulting From Principal Components (Mean) (Page 1 of 8)

7961-2b

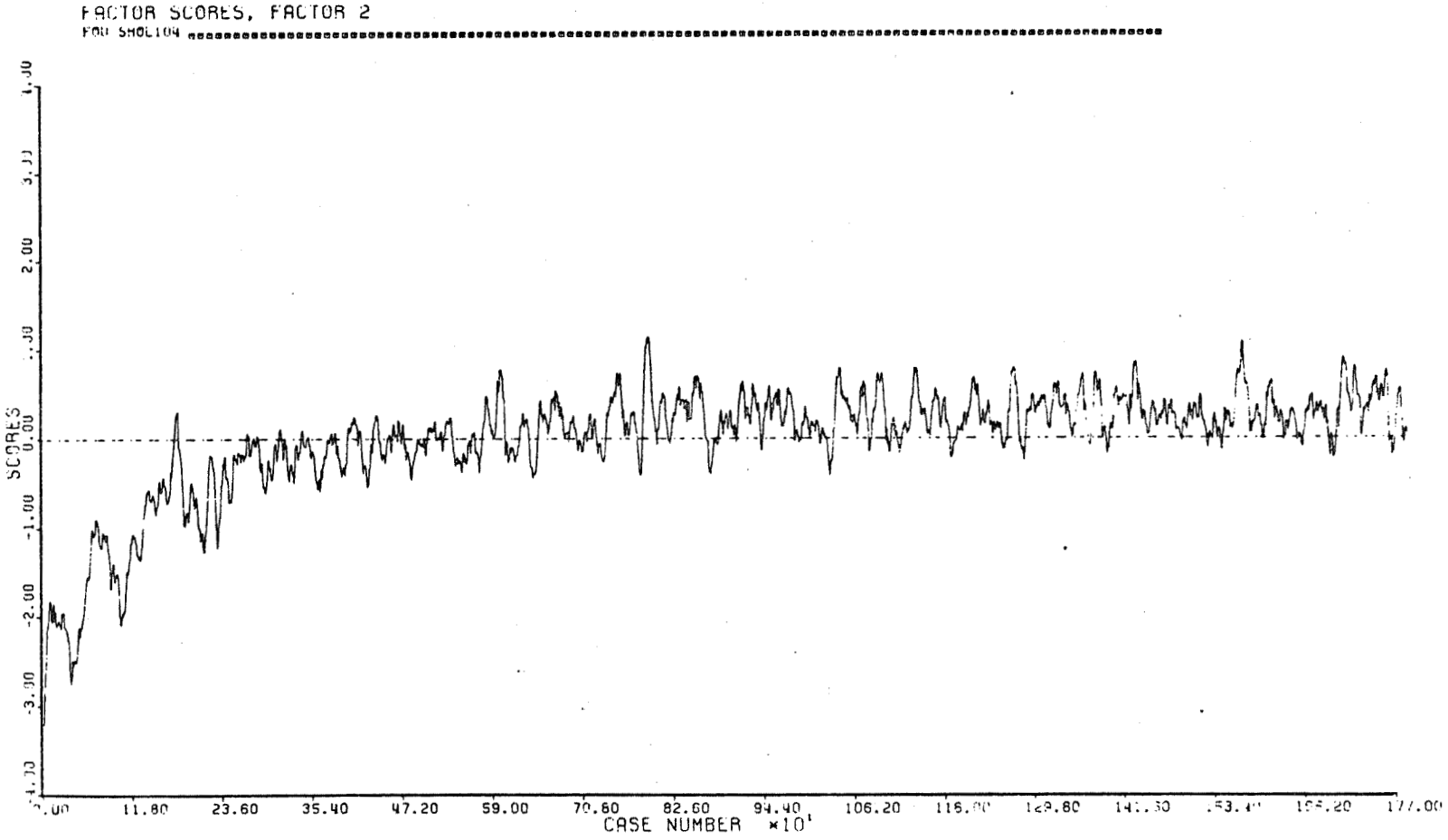


Figure 31 Eight Factors Resulting From Principal Components (Mean) (Page 2 of 8)

7961-2c

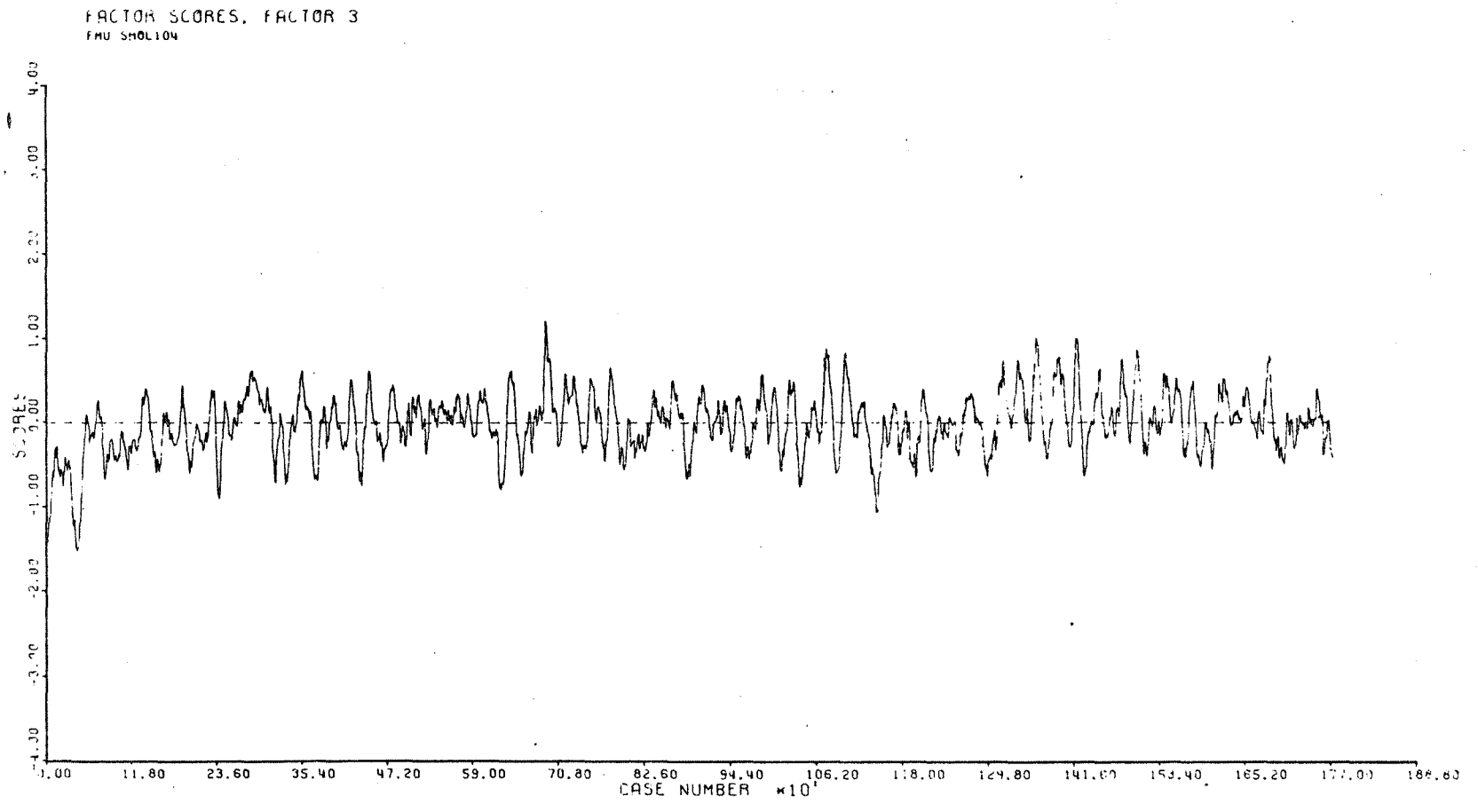


Figure 31 Eight Factors Resulting From Principal Components (Mean) (Page 3 of 8)

7961-2d

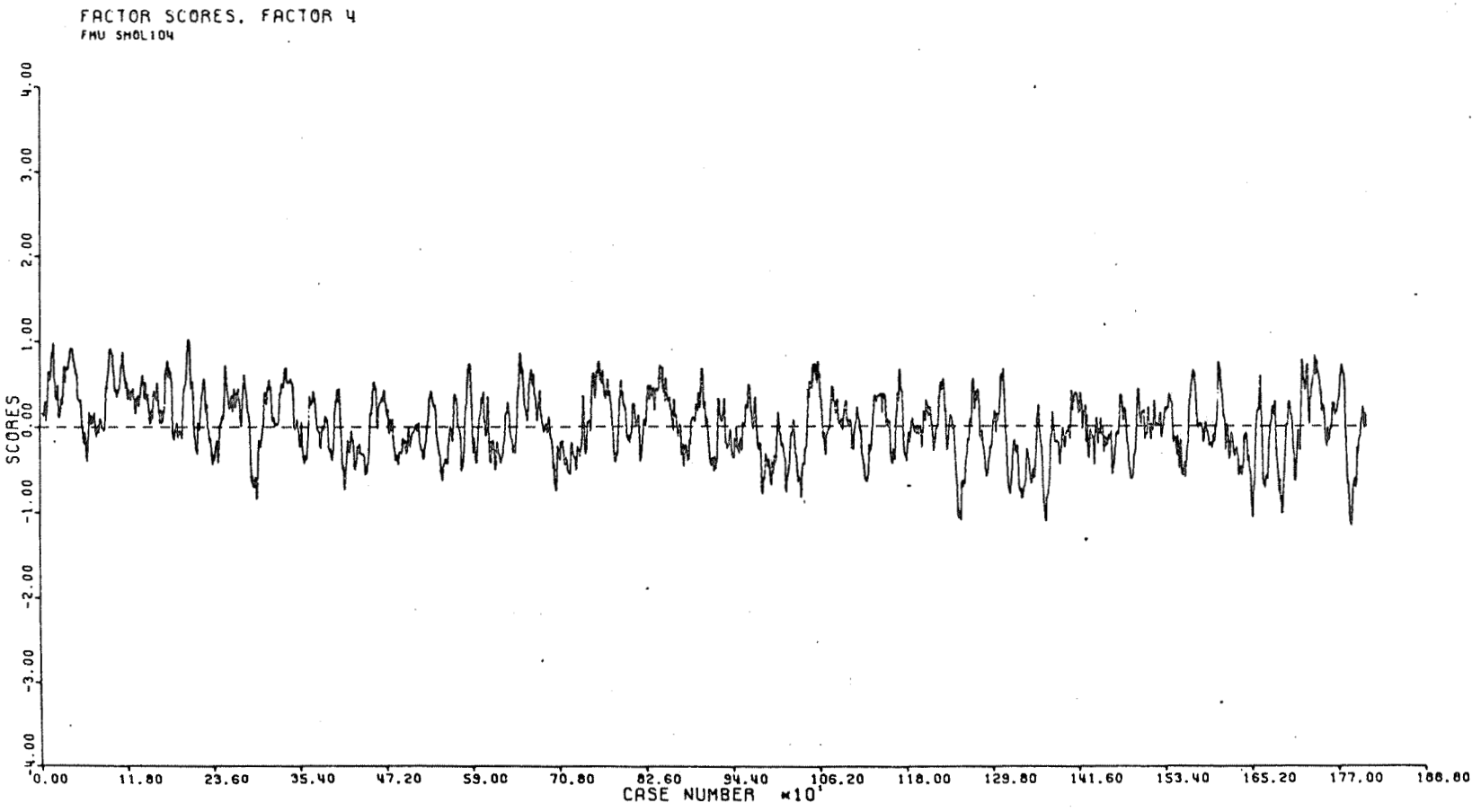


Figure 31 Eight Factors Resulting From Principal Components (Mean) (Page 4 of 8)

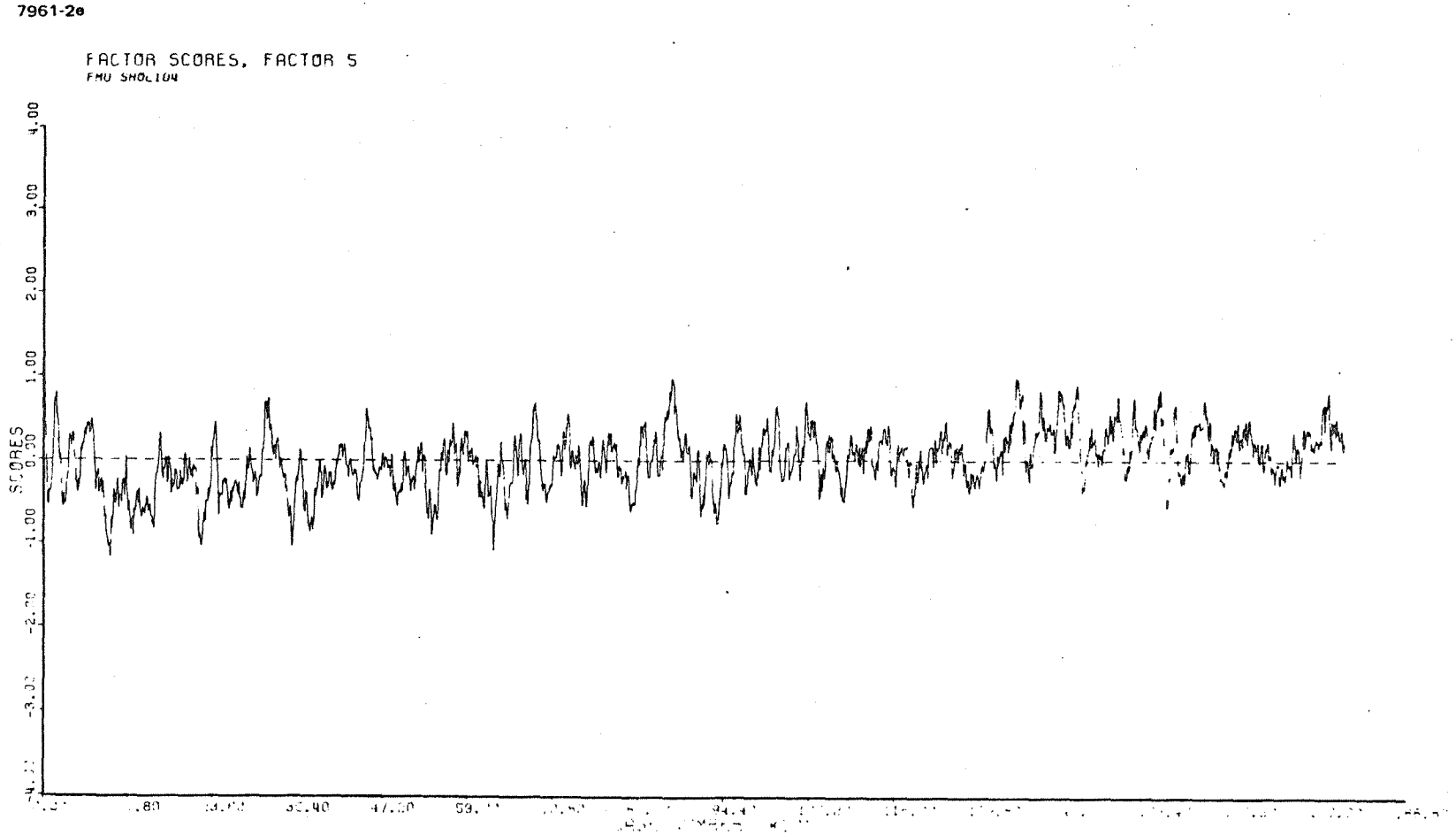


Figure 31 Eight Factors Resulting From Principal Components (Mean) (Page 5 of 8)

7961-2f

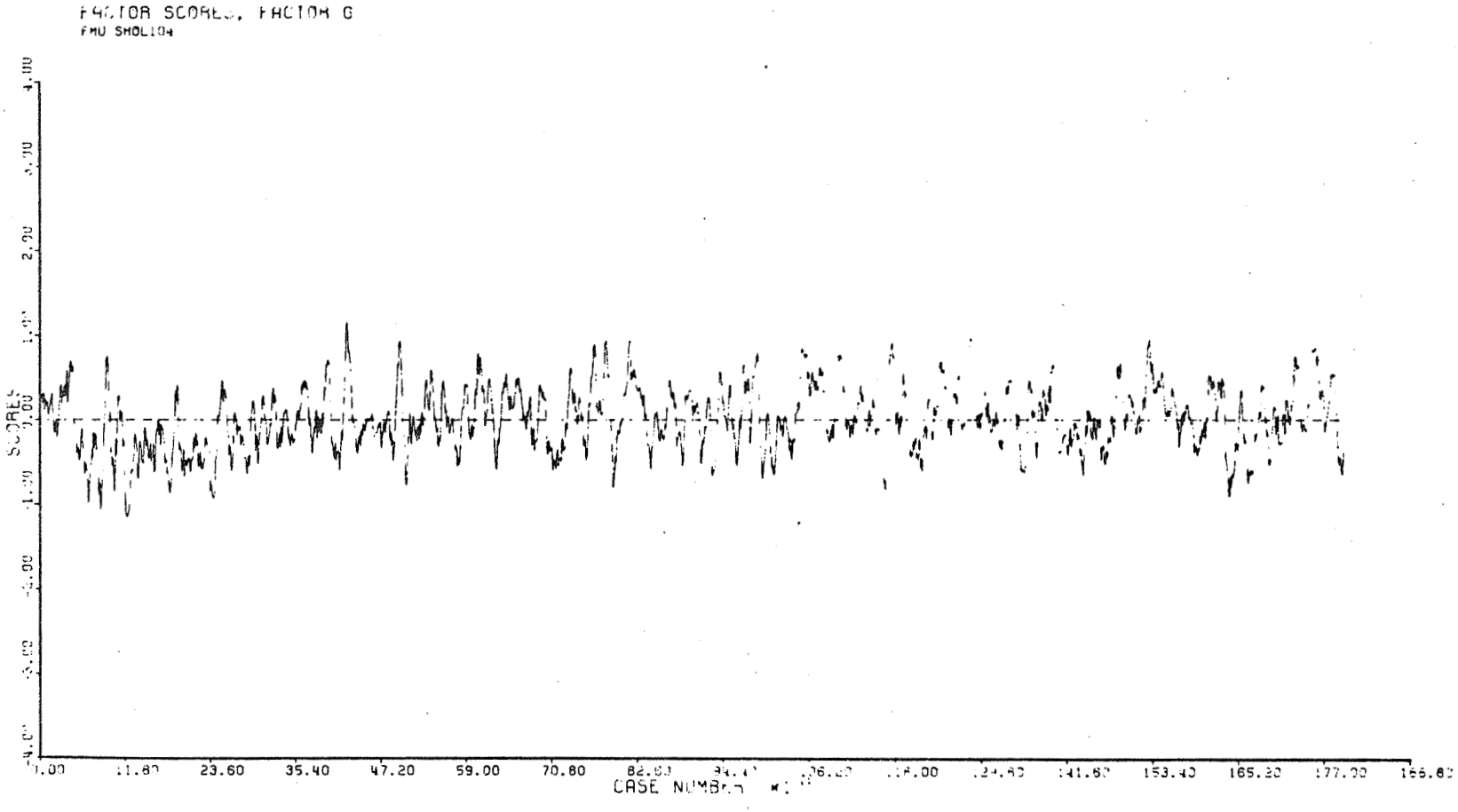


Figure 31 Eight Factors Resulting From Principal Components (Mean) (Page 6 of 8)

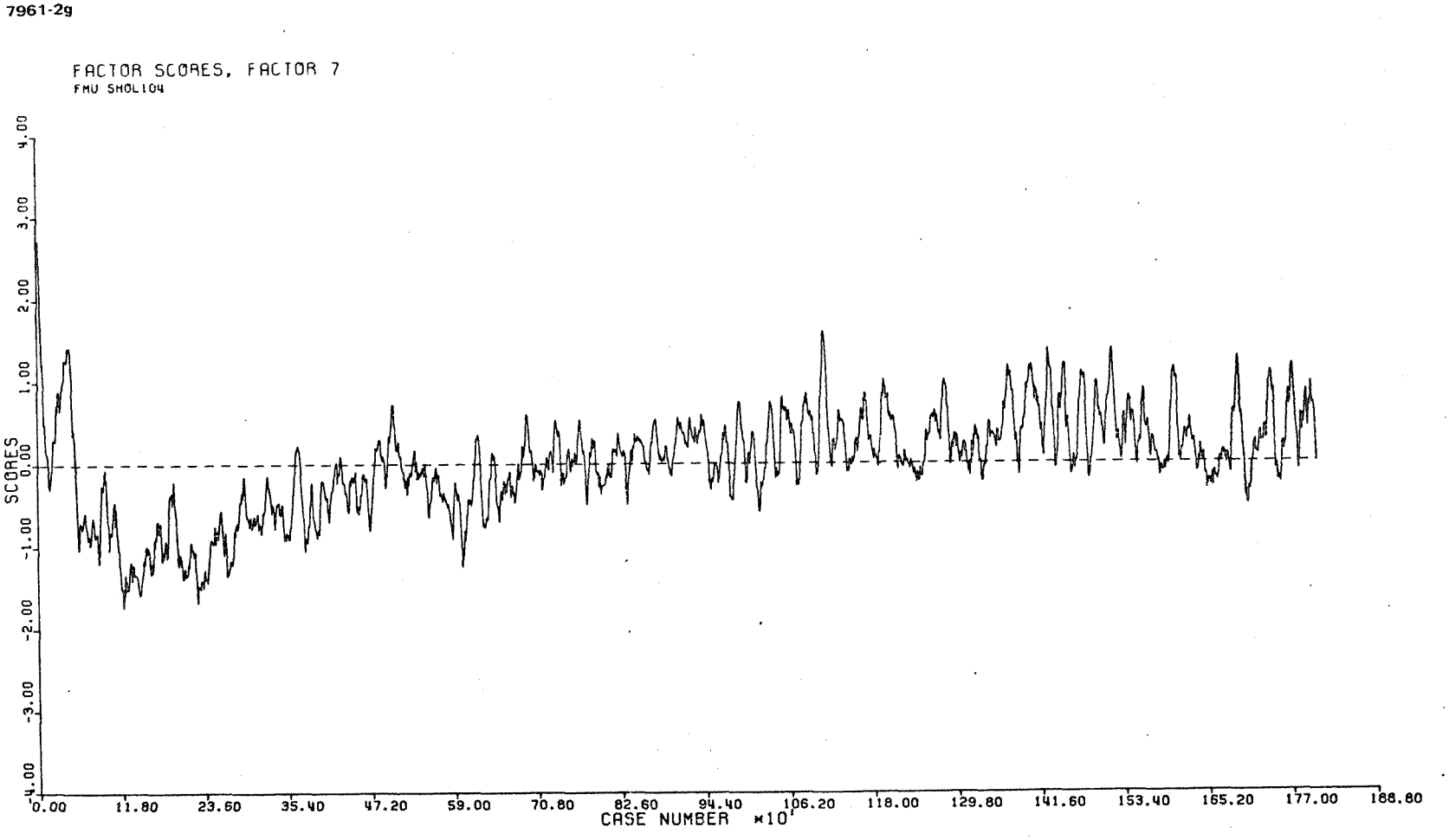


Figure 31 Eight Factors Resulting From Principal Components (Mean) (Page 7 of 8)

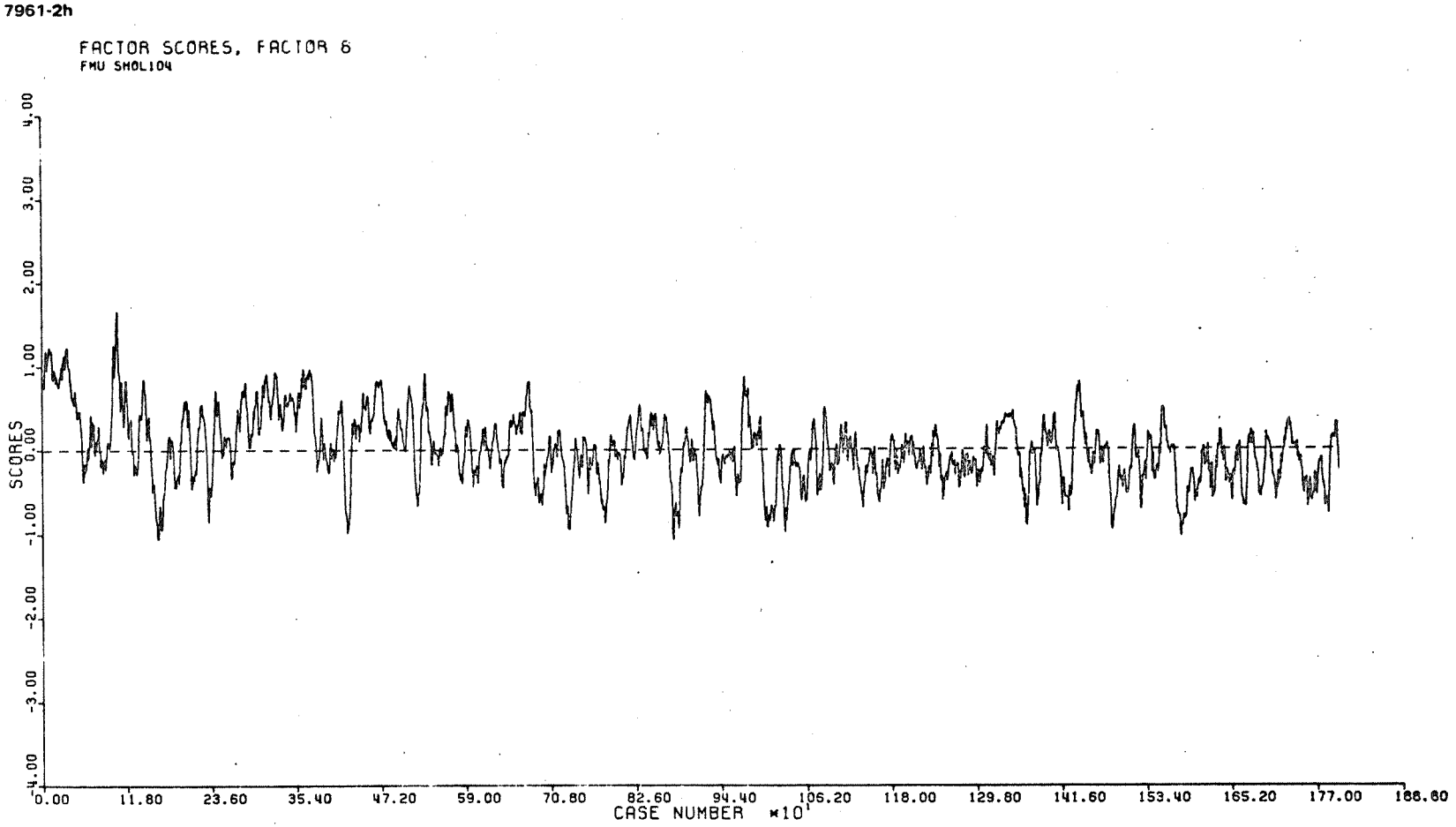


Figure 31 Eight Factors Resulting From Principal Components (Mean) (Page 8 of 8)

7961-3a

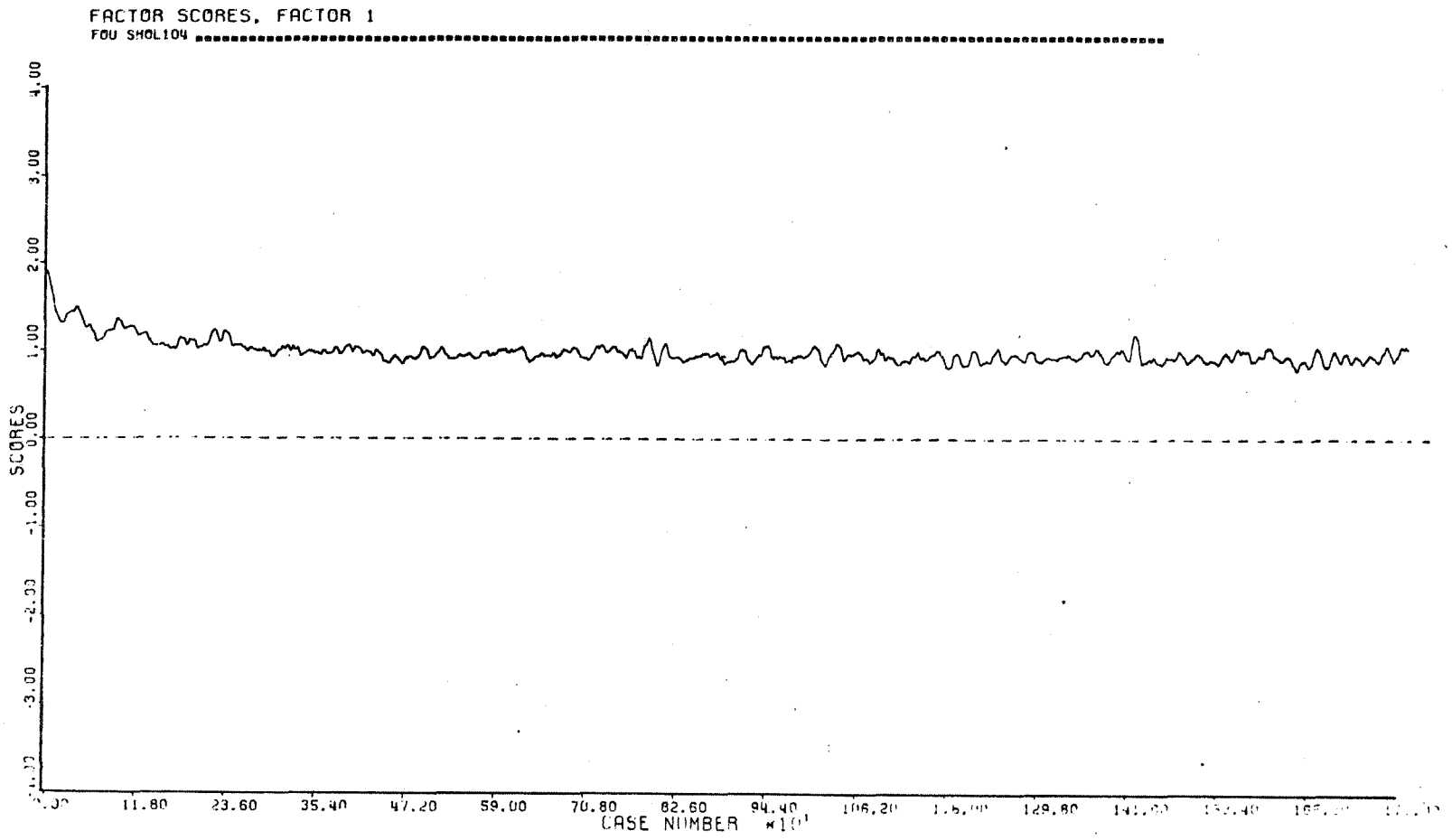


Figure 32 Eight Factors Resulting From Principal Components (Origin) (Page 1 of 8)

7961-3b

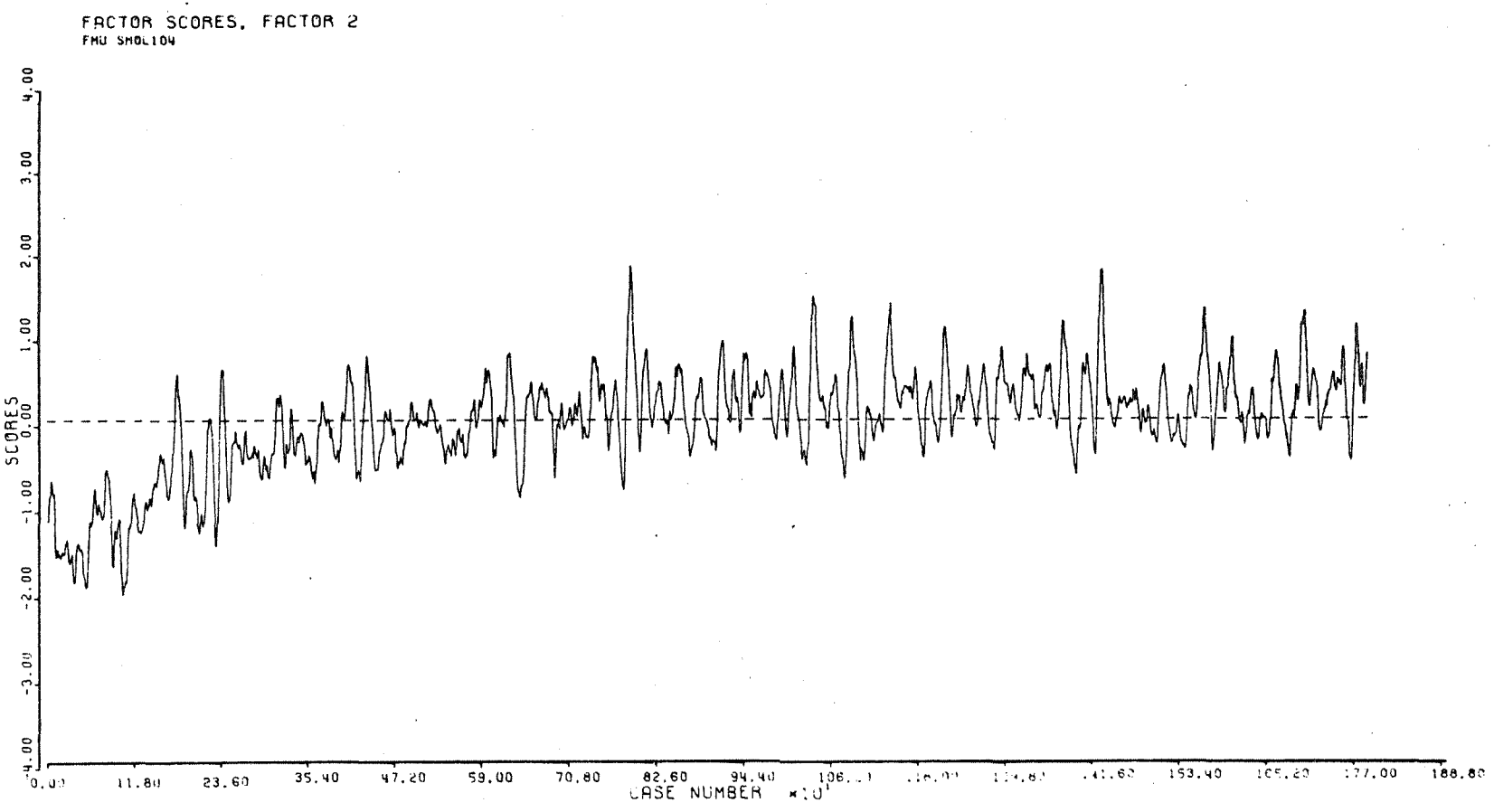


Figure 32 Eight Factors Resulting From Principal Components (Origin) (Page 2 of 8)

7961-3c

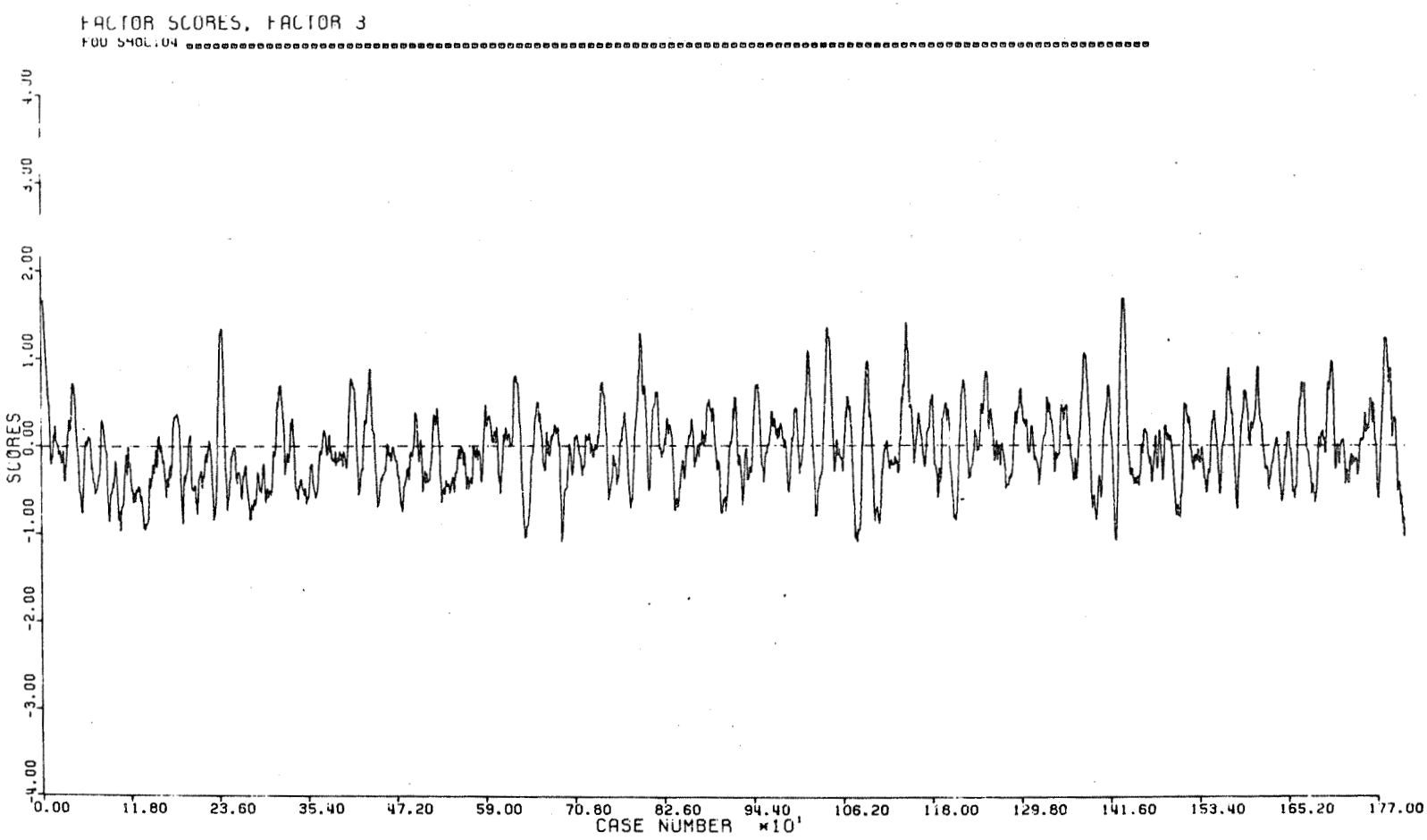


Figure 32 Eight Factors Resulting From Principal Components (Origin) (Page 3 of 8)

7961-3d

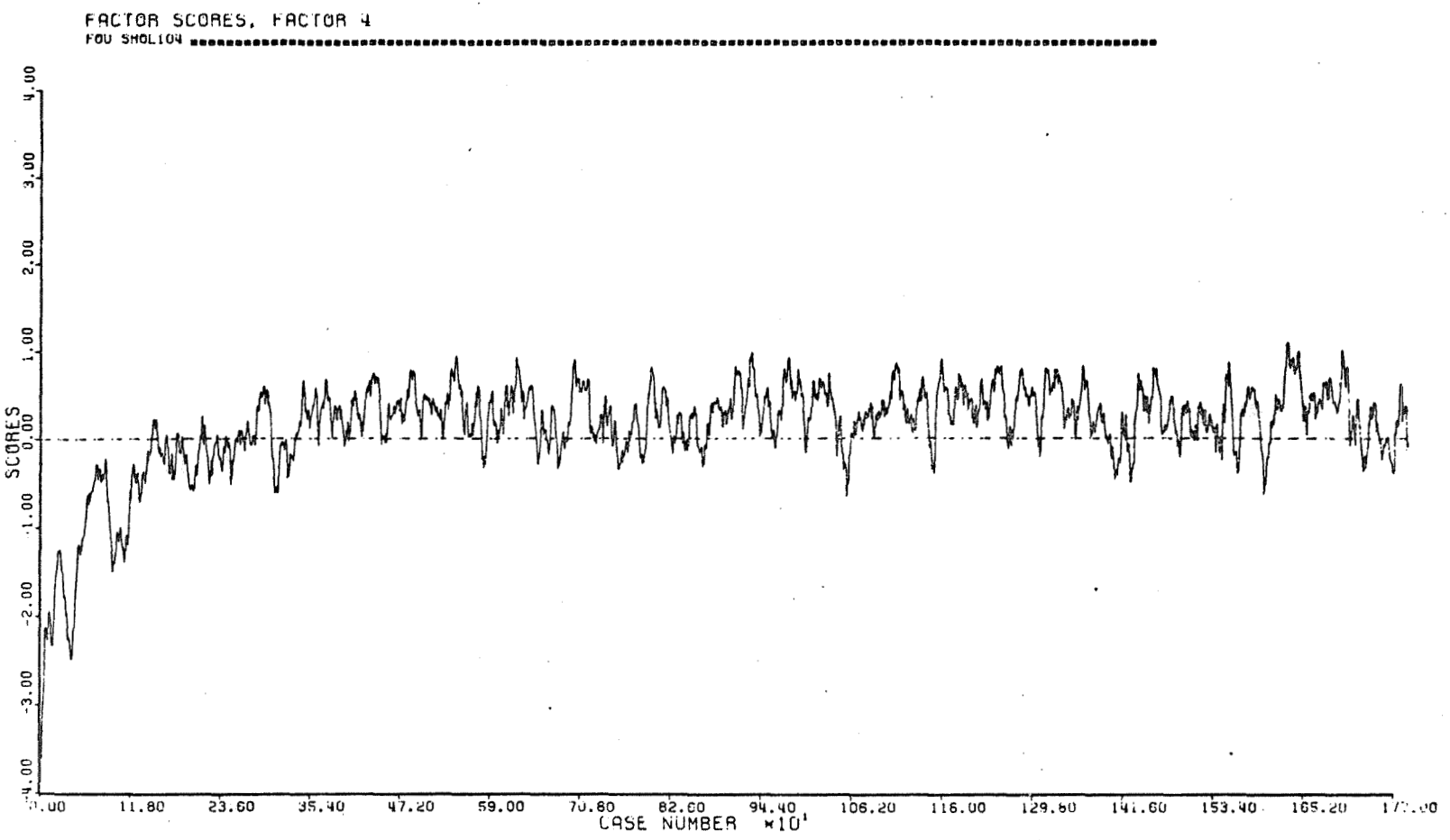


Figure 32 Eight Factors Resulting From Principal Components (Origin) (Page 4 of 8)

7961-3e

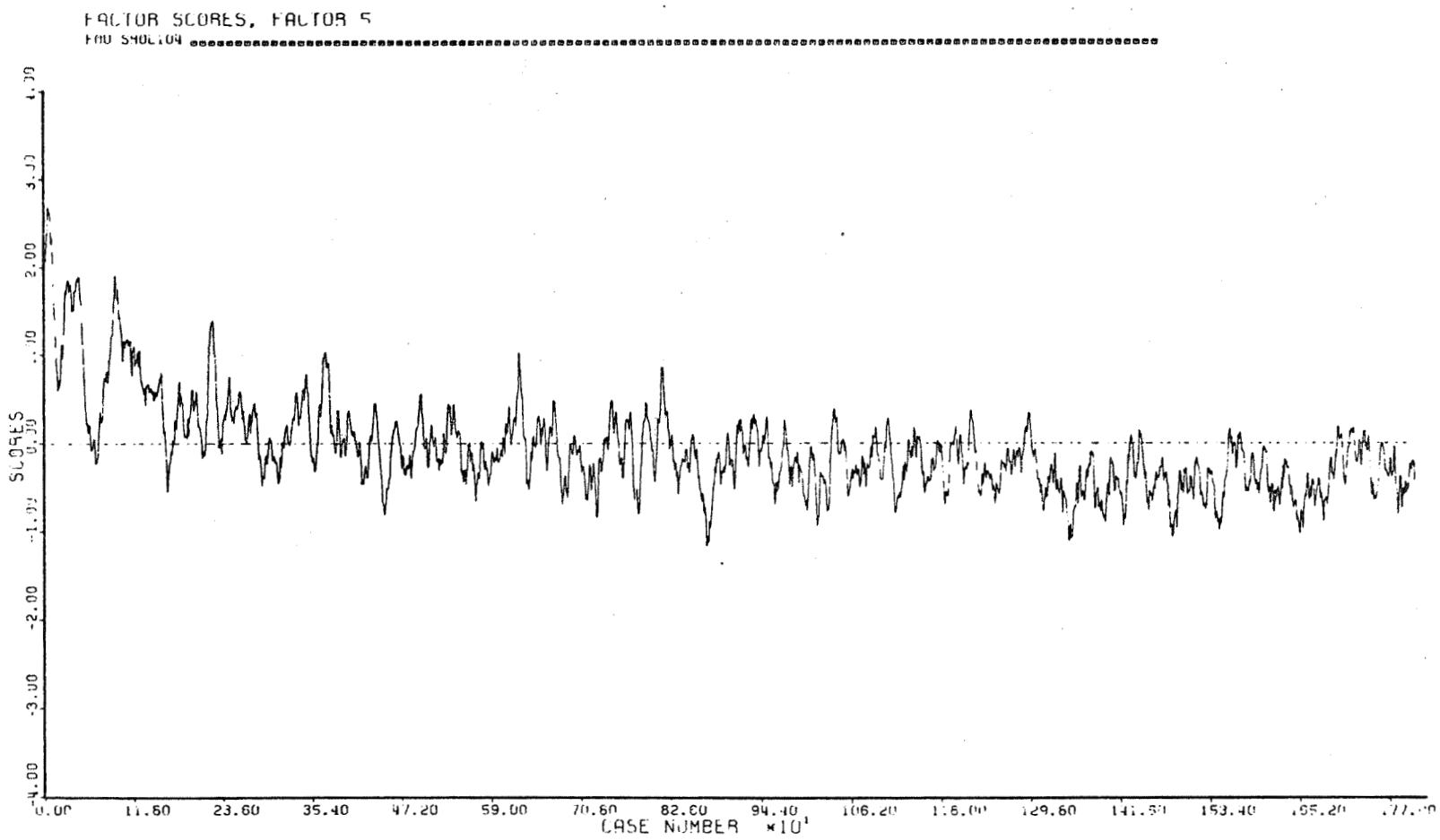


Figure 32 Eight Factors Resulting From Principal Components (Origin) (Page 5 of 8)

7961-3f

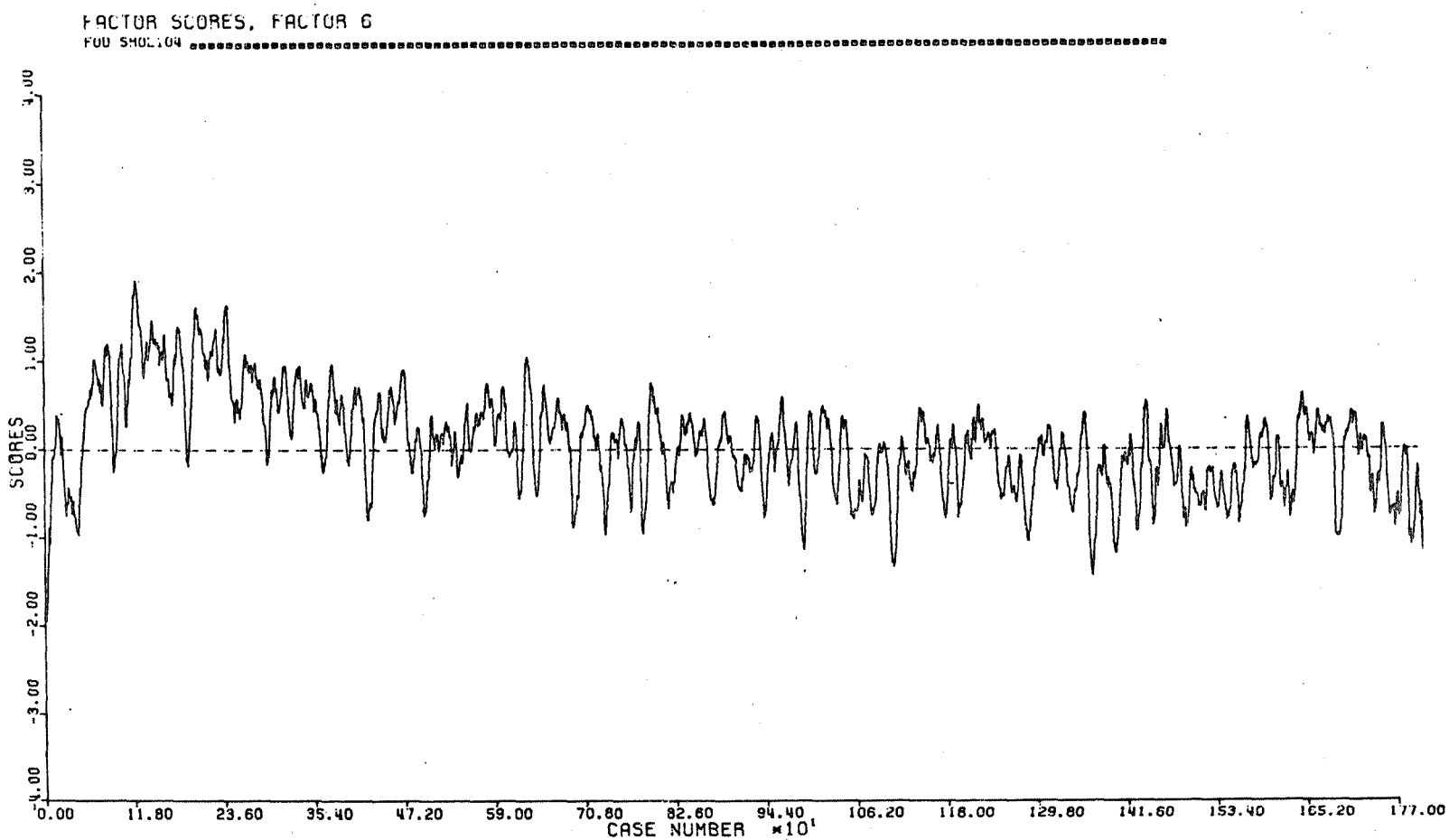


Figure 32 Eight Factors Resulting From Principal Components (Origin) (Page 6 of 8)

7961-3g

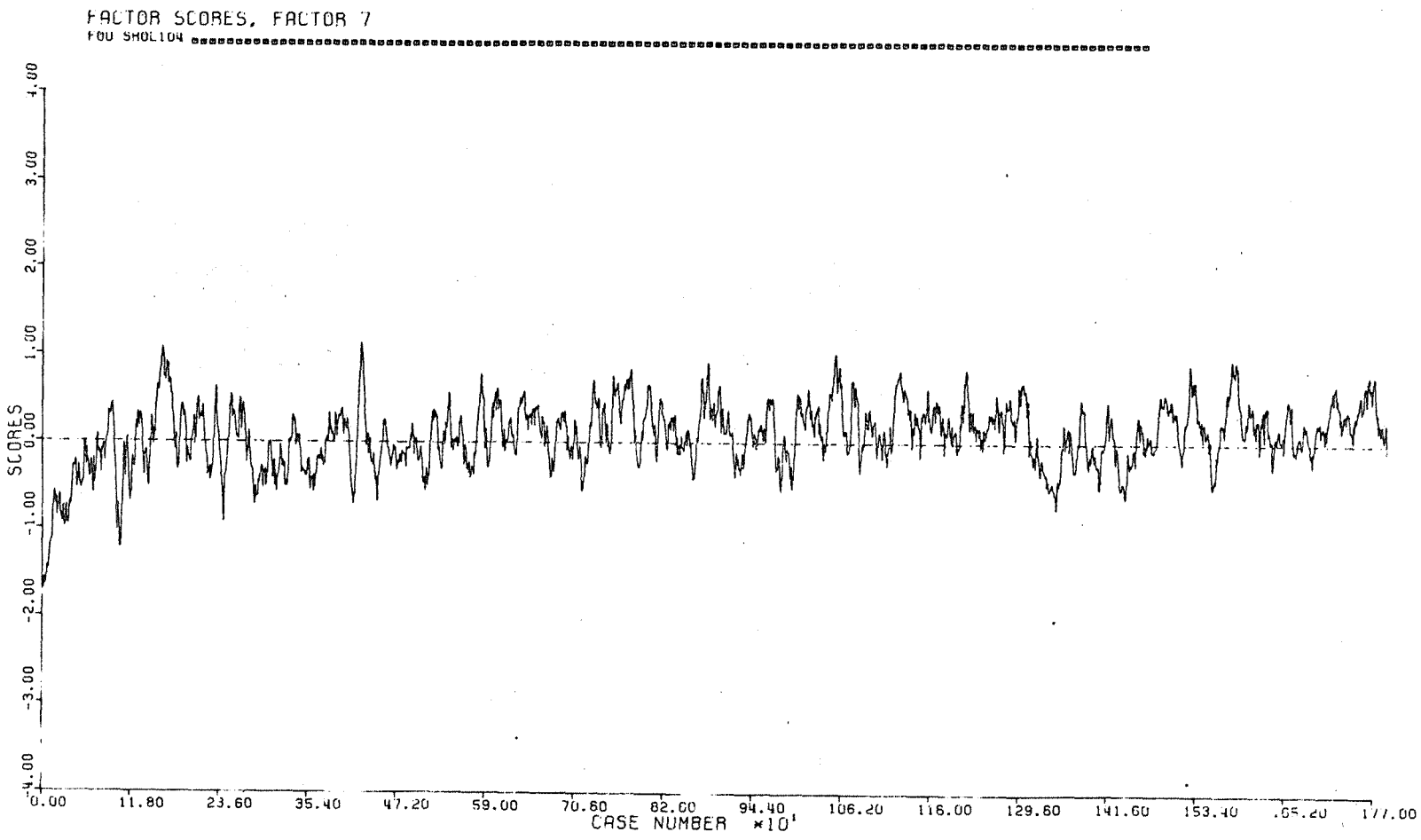


Figure 32 Eight Factors Resulting From Principal Components (Origin) (Page 7 of 8)

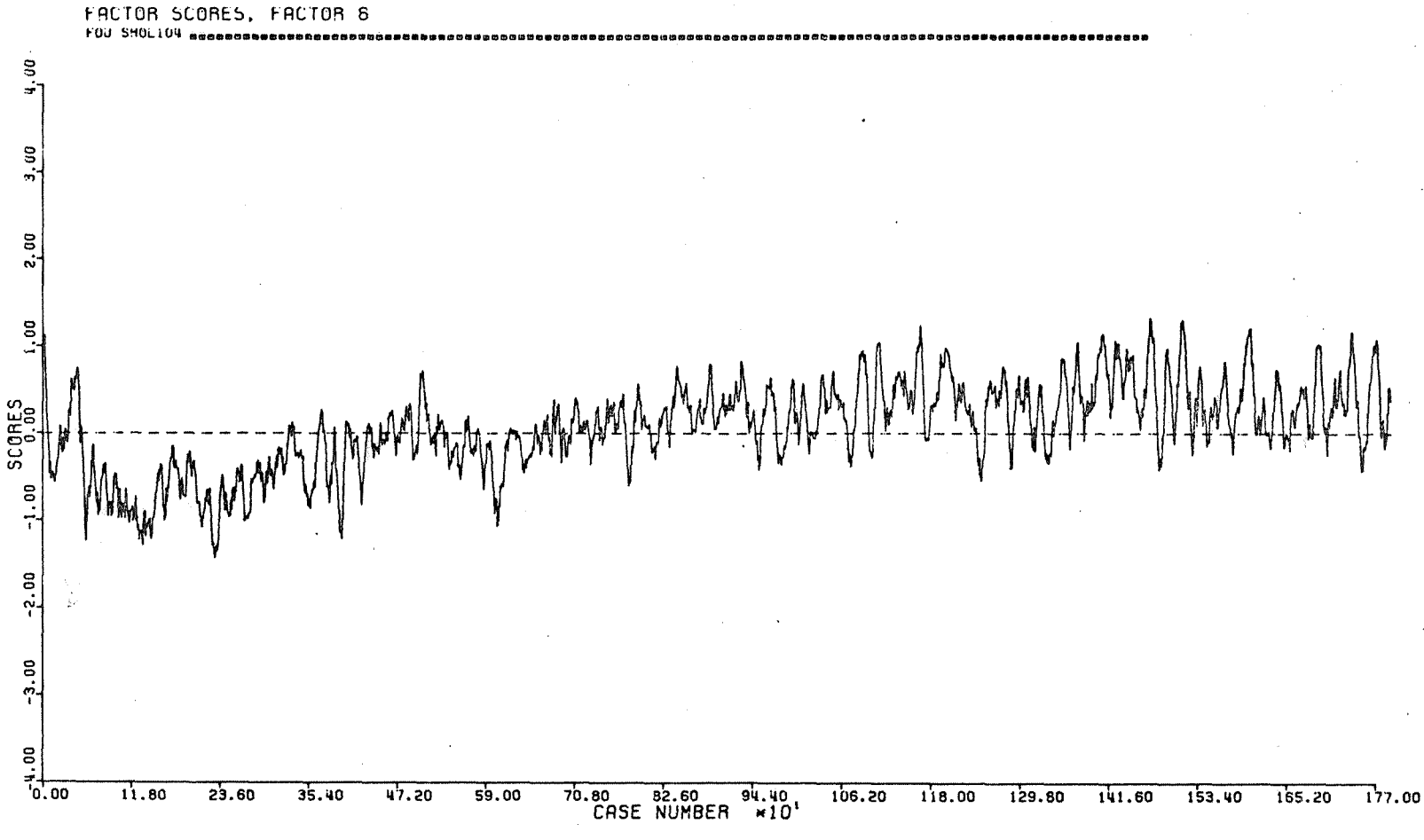


Figure 32 Eight Factors Resulting From Principal Components (Origin) (Page 8 of 8)

of extracted factors, rotations of the factors is undertaken. A transformation of the coefficient matrix can force various elements to approach zero or unity. Either the factors or channels (rows or columns of the matrix) can be treated in this manner. Analytic expressions for maximizing or minimizing the matrix elements are available for producing orthogonal or oblique rotations of the factors. The effect of applying these expressions is to expand the role of variables found to be important in the particular factor and minimize the contribution of less relevant variables.

There are no established criteria for determining the number of factors to be rotated or the type of rotation best suited for a given problem. There are, however, general goals. Traditionally, the term "parsimony" was used to describe the objective of rotating the results of a principal factor analysis: to describe actual underlying phenomena in the fewest number of measured variables. To provide some insight into these conditions, various rotation schemes were tried, including varying numbers of factors in the scheme. In the present study, it is preferred to simplify the linear combination of each factor rather than simplify the distribution of each variable. The two major classes of rotation schemes for accomplishing this are Varimax (orthogonal rotation) and Oblimin (oblique rotation). The latter form requires a constant (between zero and one) to provide a unique solution. When the constant is zero, the Oblimin factors are greatly simplified but highly correlated. At the other extreme, though less correlated, the factors retain considerable complexity. Thus, a constant of 0.5 was used for the oblique rotations (referred to as a Bi-Quartimin solution).

Aside from the transformation matrix (channel coefficients) and transformed input data (factor scores), the program provides additional information on the nature of the derived factors. The eigenvalues reveal the amount of variance accounted for by each factor. In an eight-variable case, the total variance in the data would be eight units. The sum of the individual factor eigenvalues of the same data must, therefore, also equal eight. Any factor with greater than unit variance might be viewed as above average. The factor matrix displays the distribution of variance from each channel over the various factors. (See Table II for standard case of normal data.) The square of the element in the ith row, jth column, is the percent of variation from channel i contained in factor j. A deviation matrix (shown in Table III also for normal data), which could reproduce digitizer readings from the factor scores, demonstrates the amount of change in channel readouts which would be produced when a factor score deviates by one unit. The two matrices shown were derived during the principal component analysis.

Using the descriptions of factor scores and their relations to the multispectral data from which they were derived, it is often possible to determine some physical notion of the meaning of a given factor. To view

TABLE II

7589-A

FACTOR MATRIX : U [D] F = S^T x (8 FACTORS) STANDARD DATA
UNROTATED

VARIABLE								
1	0.41103227	-0.40730221	0.52038291	-0.61701725	0.09766091	-0.05354960	-0.00743080	0.03447166
2	0.82650689	0.31455903	-0.10508086	-0.03035109	0.00804349	-0.41389082	0.12955730	-0.13349311
3	0.79527545	-0.46491553	0.02708363	0.14250891	-0.16973150	0.14947771	-0.15539117	-0.23462610
4	0.91085476	-0.06687463	-0.13017040	0.14456658	-0.06926199	-0.14648433	-0.22161616	0.22888305
5	0.82845461	-0.37981085	-0.03146161	0.17416691	-0.08729722	0.17389074	0.29904861	0.10389348
6	0.71396072	0.37362578	-0.16716513	-0.06047460	0.52101721	0.21420794	-0.03335365	-0.02461857
7	0.38728626	0.71103401	-0.10967470	-0.39353278	-0.37457035	0.19260630	-0.00364396	0.01140149
8	0.21694464	0.51823385	0.72770929	0.39172624	-0.01528942	0.03251434	-0.00511897	0.00630201

TABLE III.

7589-B

FACTOR SCORE COEFFICIENT MATRIX : F = [D]^-1 U S^-1/2 x STANDARD DATA
BEFORE ROTATION

FACTOR								
1	0.11067066	0.22293959	0.21451531	0.24569135	0.22346497	0.19258171	0.10446548	0.05851802
2	-0.26344447	0.20376678	-0.30116555	-0.04461599	-0.24603597	0.24202937	0.46059753	0.33570438
3	0.59811265	-0.12077683	0.03112912	-0.14961400	-0.03616104	-0.19213465	-0.12605685	0.83640748
4	-0.80640326	-0.03966684	0.18624961	0.18893885	0.22762449	-0.07903626	-0.51432101	0.51196260
5	0.21101052	0.01737914	-0.36672549	-0.14965056	-0.18861828	1.12573311	-0.80931346	-0.03303501
6	-0.16117174	-1.24571425	0.44989282	-0.44088347	0.52337033	0.64471563	0.57969976	0.09786054
7	-0.04113627	0.71704408	-0.86002351	-1.22655042	1.65510585	-0.18459813	-0.02016777	-0.02833132
8	0.24976714	-0.96723478	-1.70000198	1.65839012	0.75276839	-0.17837582	0.08261041	0.04566174

the linear combinations correctly, the factor scores for normalized data are examined. This matrix would represent channel coefficients if each channel had the same variance. It shows, therefore, the mixture of frequencies comprising each factor, unencumbered by channel variation.

By varying the minimum eigenvalue (variation attributed to a factor) necessary to be included in a rotation, analyses of 2, 3, 4, 5, and 8 factors were executed using Varimax and Oblimin criteria. Plots of the various first factors (labeled Factor 1) are shown in Figures 33 through 36. (The eight-factor solution for Oblimin rotation produced physically uninterpretable results. The minimizing conditions are under examination to explain the effect. It appears that the inclusion of factors which explain very small amounts of variation in this rotation produces the observed failure.) In all of the analyses, at least one factor demonstrated a similarity to the known bottom contour for a distance of more than 400 ft from the shoreline. Just as in the principal component analysis, the water depth factors possess improved S/N ratio when compared with signals from any individual channel. In all rotation schemes, the first factor was most similar in appearance to the lake bottom contour.

To compare the spectral combinations suggested by the analyses, the factor matrix and factor score coefficients for the first factor 1 from each rotation is plotted. The factor matrix reveals the amount of change produced in each channel resulting from a unit change in any factor. The coefficient matrix indicates the particular linear combination of channels necessary to recreate the various factors. Since the analyses are performed using the correlation matrix (standard deviations removed), both the coefficient and factor matrices may be examined in both normal and nonstandard form (with and without the individual channel standard devications). Tables IV and V show the Factor Matrix and Factor Score coefficient matrix multiplied by the respective standard deviations. The equations at the top of each table more clearly reveal the relation between the channel readings, X, and the factors scores F. The underline portion describes the matrix below. The plots in Figures 33 through 36 are of the nonstandard type. Also, the magnitudes have been scaled for direct comparison. Of primary interest is the relative changes occurring for the various solutions. In the case of the orthogonal rotations, several patterns are seen in the coefficient schemes (Figure 33). The two- and three-factor solutions appear very similar as do the five- and eight-factor solutions. The solution containing four factors contains some similarities to both pairs. All of the coefficient rotations demonstrate sign differences for various channels, which is not true for the principal component solution for the first factor. The two-, three-, and five-factor solutions for the oblique rotation (Figure 35) indicate almost identical combinations as the orthogonal rotations. The four-factor solution gives a Factor 1 with opposite direction to the orthogonal case (the coefficients are

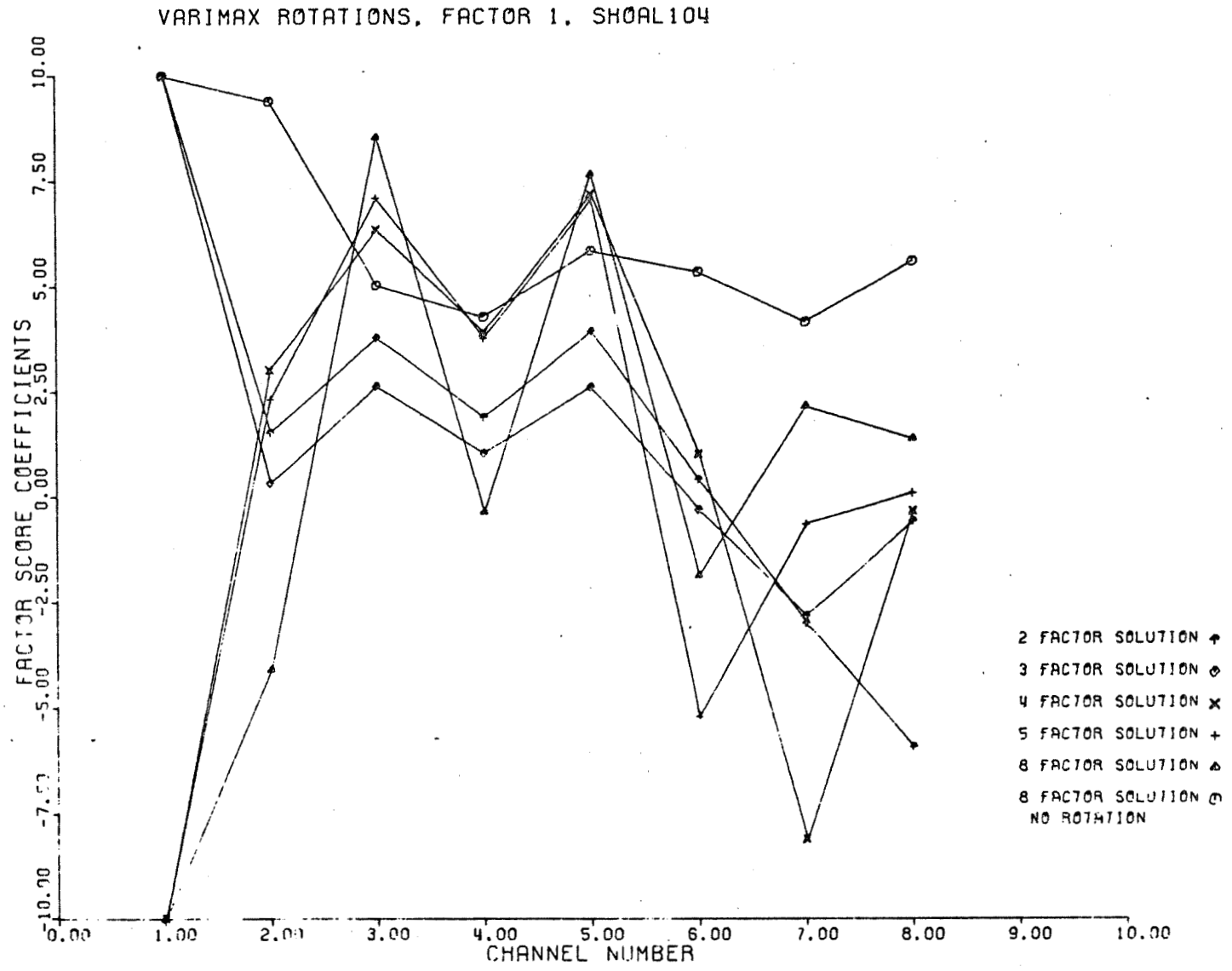


Figure 33 Coefficient Matrix Orthogonal

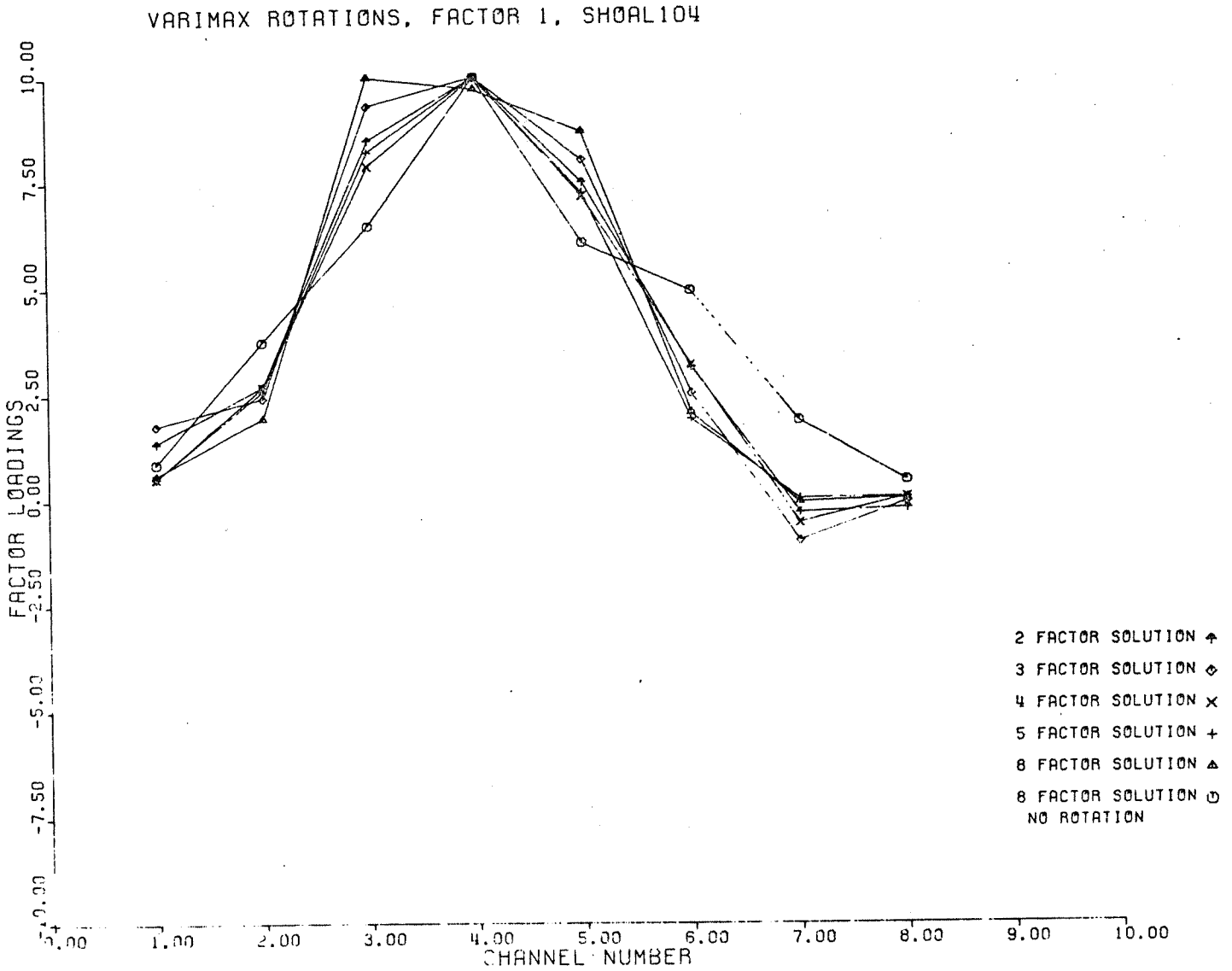


Figure 34 Derivation Matrix Orthogonal

7589-7

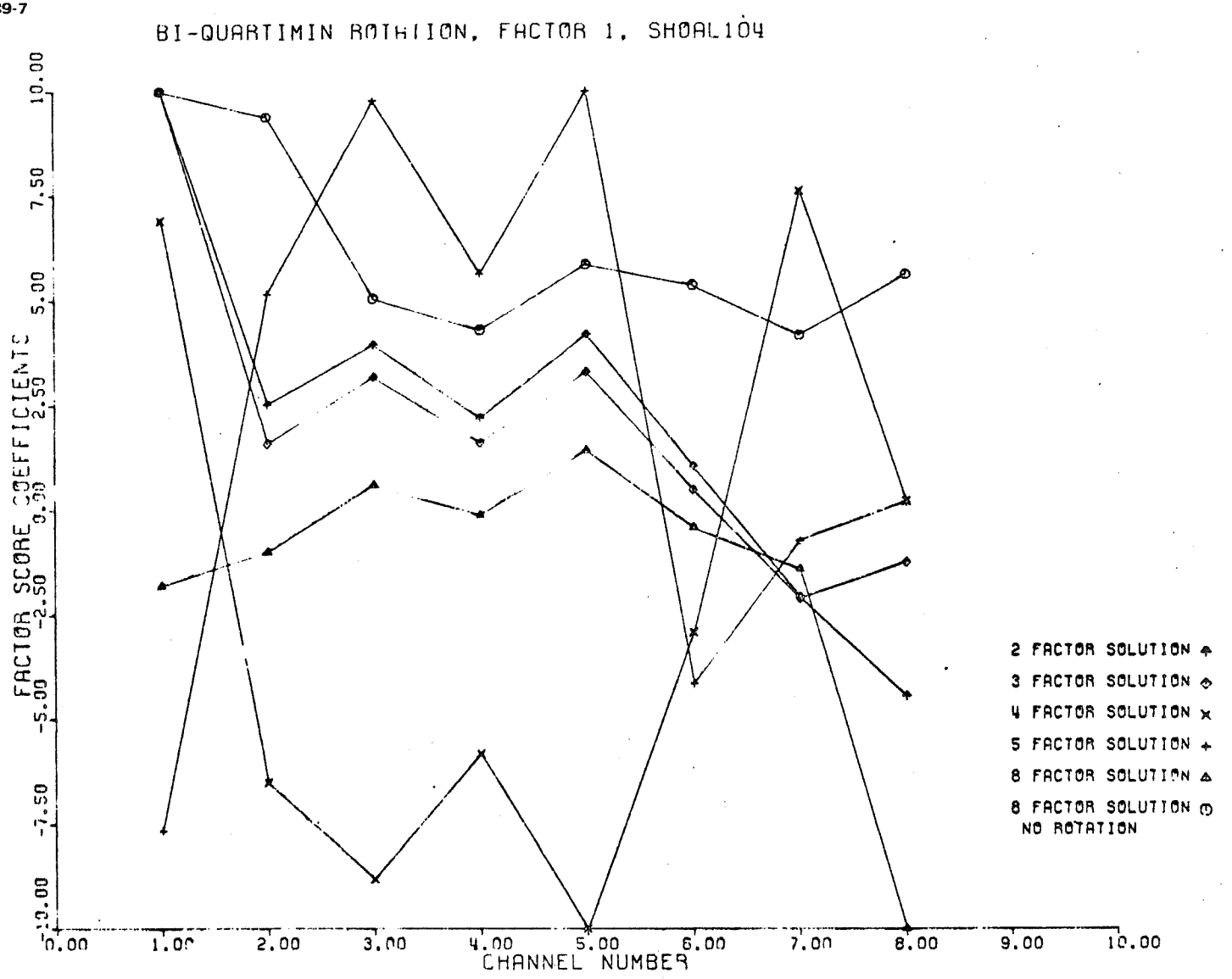


Figure 35 Coefficient Matrix Bi Quartimin

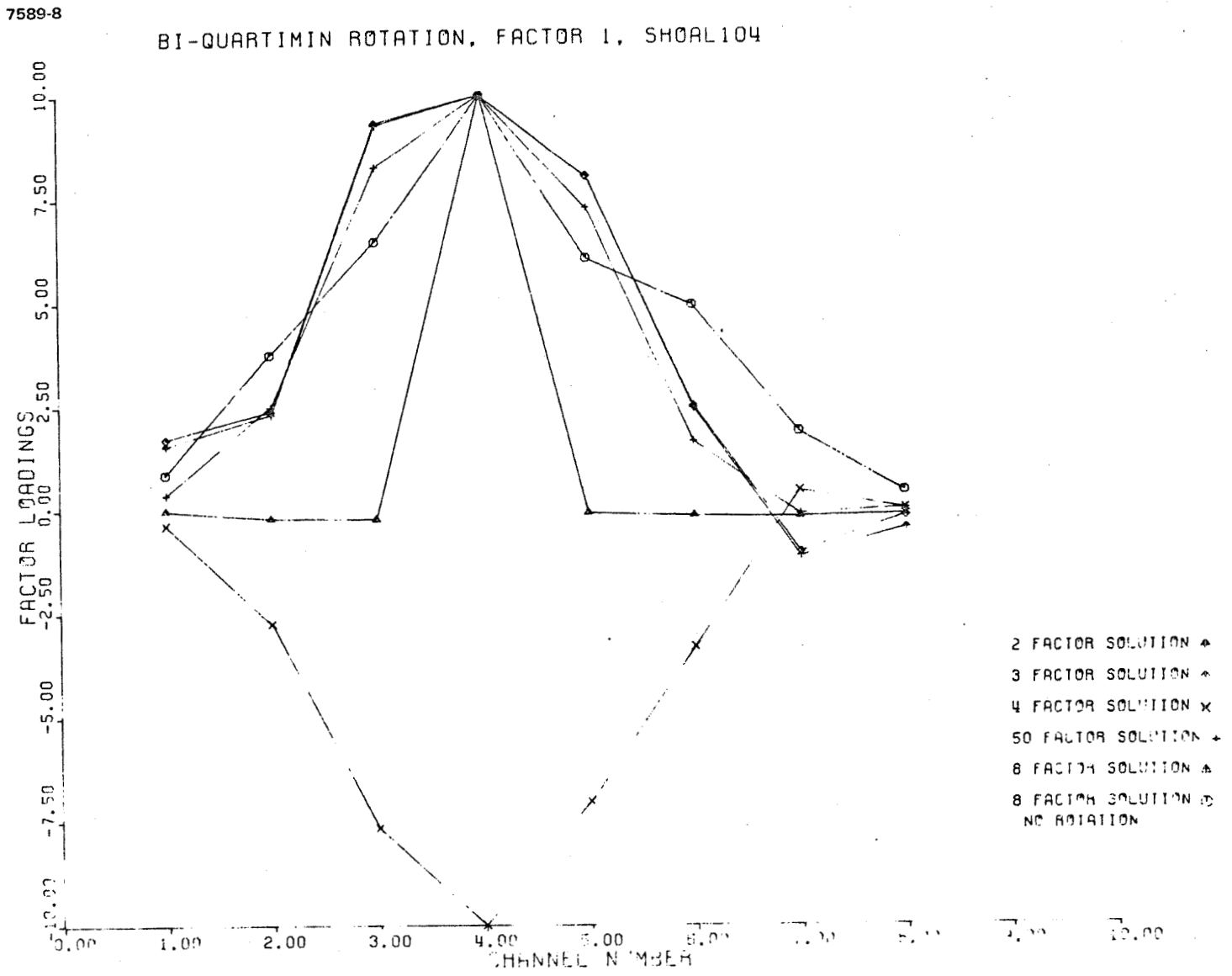


Figure 36 Derivation Matrix Bi Quartimin

TABLE IV

7589-C

FACTOR MATRIX : $S_U^{1/2}[0]^T$ F = X (8 FACTORS) NON-STANDARD DATA
UNROTATED

VARIABLE								
1	2.00842380	-1.99019767	2.54274300	-3.01492662	0.47719973	-0.76165901	-0.03630904	0.16843862
2	8.63075725	3.28476712	-1.09730170	-0.31693976	0.08399380	-4.32203440	1.35289565	-1.39399515
3	14.91054244	-8.71665635	0.50778835	2.67188581	-3.18227956	2.80254306	-2.91341407	-4.39898213
4	23.00610424	-1.73961524	-3.28780613	3.65142052	-1.74939921	-3.69985859	-5.59751645	5.78106137
5	13.98382136	-6.41098137	-0.53105325	2.93983362	-1.47352518	2.93517236	5.04776282	1.75366022
6	11.34792252	5.93852898	-2.65697668	-0.96120291	8.28121602	3.40469025	-0.53013363	-0.39129559
7	4.30125332	7.89683928	-1.21806194	-4.37062794	-4.16002861	2.13911143	-0.04047034	0.12662651
8	1.00347694	2.39708949	3.36601765	1.81193807	-0.07072119	0.15039501	-0.02367779	0.02914995

TABLE V

7589-D

FACTOR SCORE COEFFICIENT MATRIX : $F = [0]^{-1} U S^{-1/2}$ X NON-STANDARD DATA
BEFORE ROTATION

FACTOR								
1	0.02269014	0.02134936	0.01144149	0.00972738	0.01323891	0.01211638	0.00940611	0.01265118
2	-0.05396587	0.01951331	-0.01606310	-0.00176643	-0.01457610	0.01522741	0.04147235	0.07257692
3	0.12240624	-0.01156595	0.00166032	-0.00592350	-0.00214232	-0.01208826	-0.01135020	0.18082540
4	-0.16503316	-0.00379861	0.00993389	0.00748044	0.01348534	-0.00497261	-0.04630963	0.11068271
5	0.04318416	0.00166428	-0.01956005	-0.00592495	-0.01117446	0.07082611	-0.07287085	-0.00714194
6	-0.03298447	-0.11929329	0.02399569	-0.01745540	0.03100644	0.04956263	0.05219636	0.02115676
7	-0.00841666	0.06866627	-0.04587060	-0.04856143	0.09805475	-0.01161409	-0.00181591	-0.00612503
8	0.05111588	-0.09262527	-0.09067207	0.06565877	0.04459685	-0.01122261	0.00743827	0.00987175

multiplied by minus one). The factor loadings (Figures 34 and 36) demonstrate the same agreement between orthogonal and oblique solutions. The deviation of channels is very similar for all derived solutions, being greatest in the visible region of Channels 3 to 5. Since light penetration is most effective in this frequency range, a natural characteristic of a water depth variable is large expected change in those channels.

2.4 TIME SERIES ANALYSIS

The prediction of one variable from a set of independent variables can only be accomplished if the data are in phase. Since the independent and dependent variables in this study were not measured simultaneously, some technique must be established for labeling the individual reflectance measures with the appropriate depth of water. Each sample is an average reflectance over a rectangular shape with the cell centers being approximately 2.2 ft apart. Strict adherence to this estimate, however, could cause serious error at large distances from the shore. Hence, a time series analysis of the water depth data and a factor isolated in the previous section is performed.

The lake bottom contour is designated with the graph constructed by the Corps of Engineers. To provide a time series, an interpolation program is used to create any point density requested from a set of values read from the graph. The time series analysis is performed to establish an exact distance between resolution cell centers as well as to locate the shoreline boundary in the digital samples.

To begin the analysis, m values of water depth at various unevenly spaced distances from shore are read from the water depth plot, Figure 9. From these m points, a series of n depths are interpolated at a designated interval. A third-order-polynomial function is used in the interpolation process. The scheme reduces to simple matrix computations and thus easily generates the water depth series at various densities.

Analytically, the time series problem is expressed as follows:

W_n = distance from shore of nth depth measurement

X_n = distance from shore of nth reflectance measurement

$\Delta W = W_n - W_{n-1}$ (a constant distance between interpolated points)

$\Delta X = X_n - X_{n-1}$ (the distance between sample centers)

$W_1 = 0$ m

$W_m = 200$ m.

Find X_1 and ΔW_n to obtain the maximum correlation between the two series. This should occur when $\Delta W = \Delta X$ and $X_1 = 0$, thus revealing the correct ratio of distance per sample and the positioning of the water-shore boundary.

The correlation is computed by assigning a one-to-one correspondence between the two series:

$$\begin{array}{ccccccccc} W_1, & W_2, & \ldots., & W_n, & \ldots., & W_m \\ \updownarrow & \updownarrow & & \updownarrow & & \updownarrow \\ X_1, & X_2, & \ldots., & X_n, & \ldots., & X_m \end{array}$$

The correlation coefficient is calculated using the cross-product sum and the standard duration of each series. One series is then shifted to the right (or left) and the calculation repeated. The coefficients when compared indicate the best spatial registration of the two series.

The results of the analysis isolate the position of the sample indicating the lake edge and suggest the ratio value of 2.3 ft per sample. In Figure 37, densities from 2.0 to 2.4 are shown with lead and lag of 125 cases. A maximum correlation appears on the 2.3 ft/case curve when a lag of four cases is used. Possessing independent and dependent variables, the particular problem of predicting water depth from reflectance measures can be approached. For this purpose, regression analysis is attempted.

2.5 STEPWISE REGRESSION ANALYSIS

The program used in the analysis of water depth is a modification of a regression analysis from the BIMED package of the University of California. A sequence of multiple linear regression equations is computed. In a stepwise manner, variables are added to the equation. The variable chosen is the one which best reduces the error sum of squares. Consequently, it is the variable which, if added, would have the highest F value. Partialling on the variables already added, the new variable has the highest partial correlation with the dependent variable.

Several useful outputs are available from the program. Prior to the regression, the means, standard deviations, covariance, and correlation matrices may be displayed.

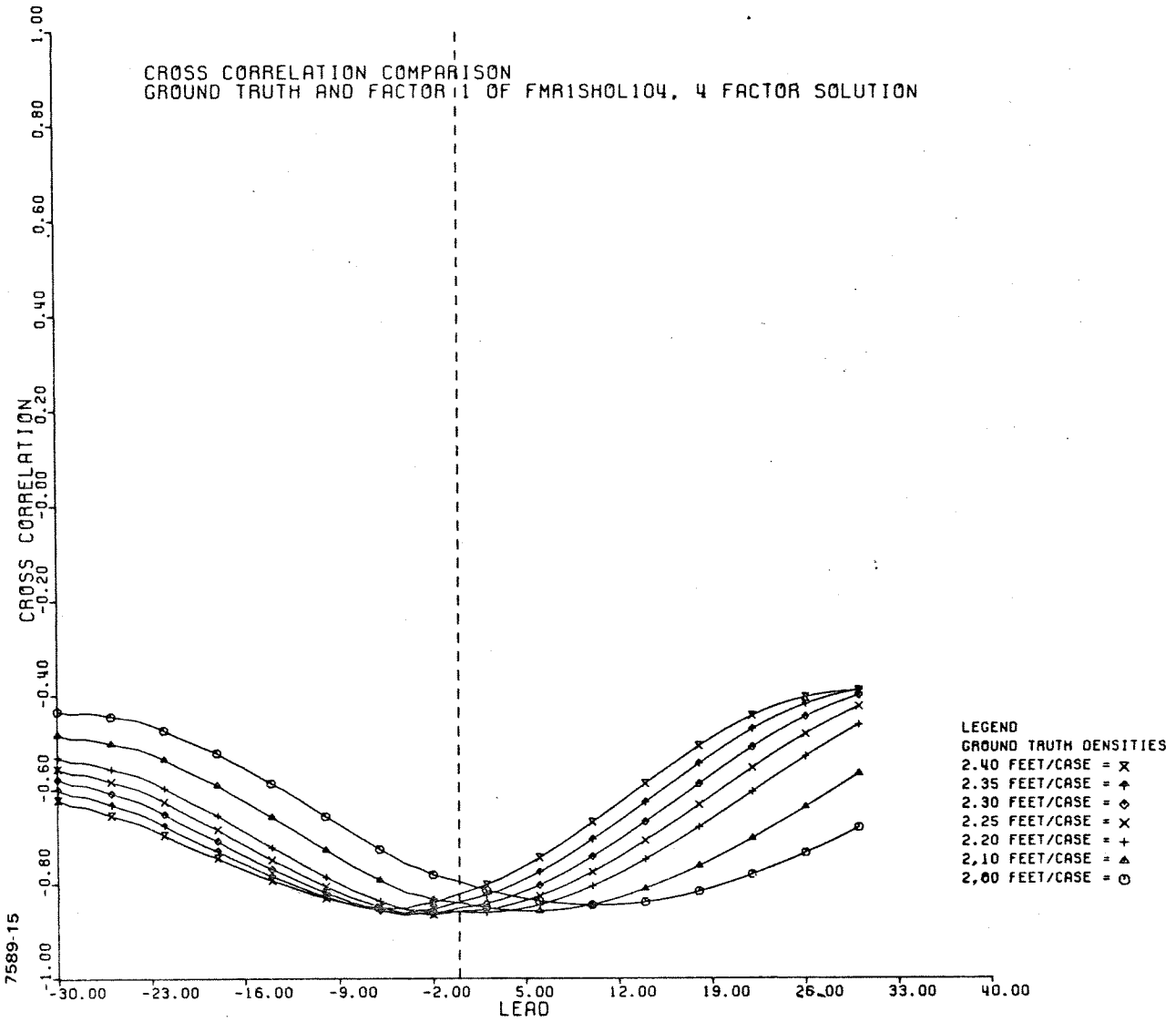


Figure 37 Results of Cross-Correlation (Lead & Lag)

At each step the standard error of estimate, multiple R, and analysis of variance table are provided. Also, for each variable in the equation, its regression coefficient, standard error, and F to remove are given. For variables not in the equation, their tolerance, partial correlation coefficient, and F to enter are printed. Finally, the list of residuals may also be provided.

A useful program option is the transgeneration of input variables. Great flexibility is given for a choice of functional form. For the problem of water depth, the model previously described required only three functional forms.

A linear regression of water depth (in meters) on reflectance measures (in digitizer units) was performed. The initial data used were evaluated at 2.2 ft/case. The following analyses were performed prior to obtaining the final results from the time series, principally to test the program. Also, the inclusion of some information about the method in this report was desirable. Two-hundred sets of nine values were inputted in spatial order. The correlation matrix was printed and revealed physically expected results. Correlations of greater than 0.75 were found between the third, fourth, and fifth channels and the dependent variable. Since light penetration is greatest in this frequency range, these variables would be expected to follow a water depth contour. The correlation of both Channel 2 and Channel 6 was somewhat less, 0.61 and 0.67. Little correlation (less than 0.3) was found between the ultraviolet or near-IR and the depth of water.

All eight variables were added to the regression equation in the following order: Channels 4, 5, 7, 2, 3, 1, 8, and 6. The multiple R began at 0.885 and continually increased to 0.896. At each step, the F test indicated significance at the 0.001 level. The standard error of estimate was 0.421 for the first step and decreased to 0.375. (The standard deviation of the dependent variable was 0.828). A disturbing result, however, appeared in the list of residuals. Here a grouping of positive and negative signs was apparent. Changes of sign occurred four times.

The functional form of first and second powers added the first several variables as follows: Channel 4, 5, Channel 5 squared, Channel 6 squared, Channel 4 squared, Channel 1, 6, 2, and so on. The multiple R for all 16 independent variables was 0.934, and the standard error of estimate was 0.307. The F-ratio was, as in the previous case, significant at the 0.001 level at all steps of the regression. Examination of the residuals revealed a similar pattern of sign changes to the one found in the first-order functional form.

The same data were regressed in a third transgenerated form. The water depth was predicted from the natural log of the reflectance measures. For this analysis, the variables were added as follows: Channels 5, 4, 7, 1, 8, 6, 3, and 2. The multiple R was somewhat improved over the previous first-order results. The first variable yielded 0.881. With the eight variables in the equation, this statistic had risen to 0.920. As with both previous functional forms, all steps were significant at the 0.001 level. The standard error of estimate approached 0.329. Examination of the residuals showed a similar pattern as in the previous models.

In an earlier section, overlap in the spatial resolution cell was mentioned. The pattern in the residual list suggests an autocorrelated series. Use of the d-statistic (added to the regression program by the author) and the "significance points of the Von Neumann ratio of least-squares estimated regression disturbances" (Theil and Nager) imply the rejection of the hypothesis of no autocorrelation at the 1% level for the regressions performed. This effect is known to produce needlessly large sampling variances for the estimates of the regressor coefficients, through they are unbiased. Also, the F tests indicated by the program are not valid. Since the overlap of the resolution cell causes the mth reflectance sample of the nth channel, X_m, to be related to the (m-1)th sample as follows:

$$X_m = X'_m + r(X_{m-1}) + E_m, \quad 0 < r < 1$$

$$X'_m = \text{true reflectance}$$

$$E_m = \text{error term}. \qquad (2\text{-}20)$$

then choosing an r and replacing X_m with $X_m - r(X_{m-1})$ may remove the autocorrelation (Johnston, p. 187).* The dependent variable is treated in the same manner.

When r was chosen as 0.25 and both the linear and logarithmic forms were used, the d-statistic rejected the null hypothesis of no serial correlation. The standard error and multiple R were lower than the previous cases (r = 0). An r of 0.5 and 1.0 was also tried, using only the direct linear combination (natural logs of occasionally negative differences being difficult to come by). The residual sign pattern in all three cases was not removed or appreciably altered; the d-statistic remained too low. As with r = 0.25, the standard error was smaller than that obtained using raw data and the

*See Bibliography.

percentage of variance accounted for was also smaller. Using the results of the F test, all equations extracted would be judged significant at the 0.001 level.

The best estimate of case density was used in a straightforward linear model. The water depth curve was evaluated at 2.3 ft/case to associate a depth value with a set of reflectance measures. The first 200 cases (representing a distance of 460 ft from the shore) were submitted for analysis. The channels were added to the predictor equation in the following order: 5, 4, 7, 1, 6, 8, 3, and finally 2. The multiple R for the single Channel 5 prediction was 0.851. With all eight variables included, the value had risen slightly to 0.896. The standard error of estimate was 0.369 with the inclusion of all channels. The F-test indicated significance at the 0.001 level, as in the previous regressions. However, this estimate of case density did not remove the apparent autocorrelation in the residuals. The value of the d-statistic confirmed a visual inspection in suggesting a serial dependency.

The method of ordering by dependent variable and grouping was next applied to attempt a more satisfying result. The first 300 samples were ordered by water depth. Thirty sets of measures were arrived at by averaging over 10 successive samples. The multiple regression of the natural log of the samples added the channels in the following order: 4, 7, 1, 3, 6, 2, 8, and 5. The multiple R with all eight variables in the equation was 0.977, the standard error of estimate being 0.275. Each step of the analysis was significant at the 0.001 level. The residuals, however, showed no patterns but appeared in random order. The d-statistic was comfortably above the minimum for acceptance of the hypothesis of no autocorrelation. The coefficients of the regression were largest for Channels 3 and 4 and smallest for Channels 5, 6, and 8.

In all the stepwise regression runs, the fourth channel was entered as the first or second variable to be included. The resulting predictor equations suggested this reflectance range receives a coefficient of magnitude greater than most of the other variables. The third or fifth channels appeared as important variables in the prediction of water depth, one or the other being entered at an early step and receiving a large coefficient. The persistent pattern in the residual signs was never removed. The regressions performed indicate a better fit from the logarithmic form, as suggested by the model in Section 2.1. The standard error of estimate was approximately 20-25% of the mean water depth (which ranged up to 2.2 m at a distance of 460 ft from shore). An increase in S/N ratio could decrease this error. (Reference is made to this statement in the Program Summary.)

SECTION 3

PROGRAM RESULTS

3.1 SUMMARY OF COMPLETED STUDIES

The results obtained during this report are described in the following paragraphs.

3.1.1 Theoretical Model

A mathematical model defining the relationship between the apparent reflectance of a body of water and its depth and bottom reflectance was derived. The expression is given as a parametric function of scattering and absorption interactions of light intensity with water. Section 2.1 gives the derivation. The expression was evaluated for various ratios of scattering to absorption and bottom reflectances.

3.1.2 Factor Analysis Development

Several factor analysis programs were comparatively evaluated by comparing performance on specially contrived data sets. The best of these was selected for development to include the most desirable features of all routines considered. The numerical accuracy of the resulting program is a considerable improvement upon that provided by the original program. The development and use of this technique are illustrated in Section 2.2. One primary application of this analysis dealt with the assignment of ground truth (water depth) to sample spectra.

3.1.3 Sample Scanner Data

Portions of multispectral video from a test site selected over Lake Michigan were sampled and recorded in digital form for computer analysis. Portions of the sampled spectra were factor analyzed, correlated with known water depths, and treated as independent variables in multiple regression analyses. Film imagery of these portions was produced to illustrate the information content of single channels, factor-enhanced imagery, and regression imagery. The scanner, digitizing technique, and the film recorder are described in Section 1.

3.1.4 Ground Truth Collection

Several field exercises were conducted to collect information about lake conditions and water depth at the flight times in the area over which the multispectral scanner was flown. Water depth measurements were made on-site during the time of data acquisition. Mappings of the area made just previous to the flight by the Army Corps of Engineers were obtained. A detailed sampling to measure various chemical parameters of the lake was prepared by Professor F. Bevis of Grand Valley State College and is presented in Appendix A.

3.1.5 Association of Data and Ground Truth

Section 2.3 describes the use of factor and time series analyses and cross-correlation in the assignment of a value of water depth to sample spectra. A factor isolated by the analysis showed resemblance to a variable measuring depth which was sufficiently adequate to permit estimates of the spatial coverage of each spectral sample. Using a polynomial estimating procedure, lake bottom contours were converted to digital values which were subsequently compared to the water depth factor. Time series analysis and calculation of cross-correlation permitted the locating of the water-beach boundary in the data, as well as the exact spatial density of the digital samples.

3.1.6 Regression Analysis

Several functional forms were used in the generation of a regression predictor equation. All regression analyses were found to be significant at a 0.1% confidence level. A pattern of alternating sign changes appeared in the residuals in every computation. Such a pattern suggests the existence of a serial dependency in the regression parameters. Several plausible reasons explain the existence of serial dependency. Spatial overlap of scan lines would produce this effect. A portion of the regression problem required the estimation of the exact distance between scan lines. The assumption of a constant value may not be accurate; if so, no single density estimate would be adequate. For evaluative purposes, however, it is evident that information can be used in multispectral data for water depth prediction. Two methods used to display this information are described in the following section.

3.1.7 Water Depth Prediction

The water depth information (as related to measured reflectance) in multispectral data is shown in Figures 38 and 39. In Figure 38, depth curves as measured by the Army Corps of Engineers are compared with the depth as predicted by the regression analysis results. Figure 39 shows a contour mapping of equal depth lines (as measured by the depth prediction equation). The mapping was made by level thresholding the output of the analog computer while the computer was generating the linear regression equation:

$$y \text{ depth} = \sum_{i=1}^{8} a_i \chi_i + C \quad (\chi_i = \text{reflectance in ith channel})$$

where

c = constant

The thresholding of the depth voltage was accomplished with comparators and logic circuitry.

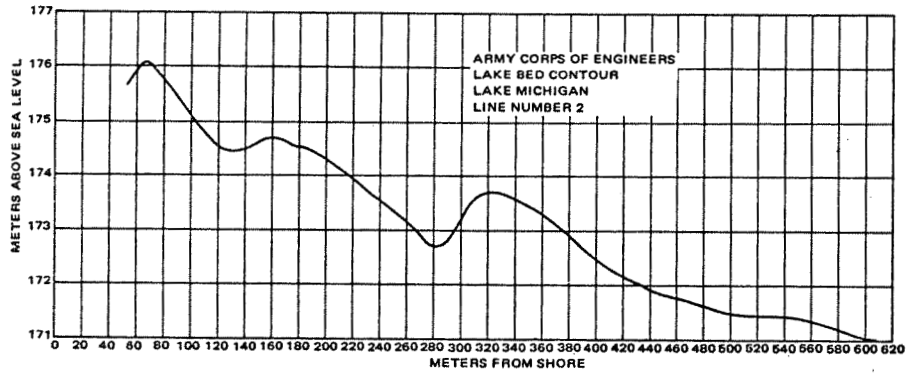
3.2 OBSERVATIONS AND CONCLUSIONS

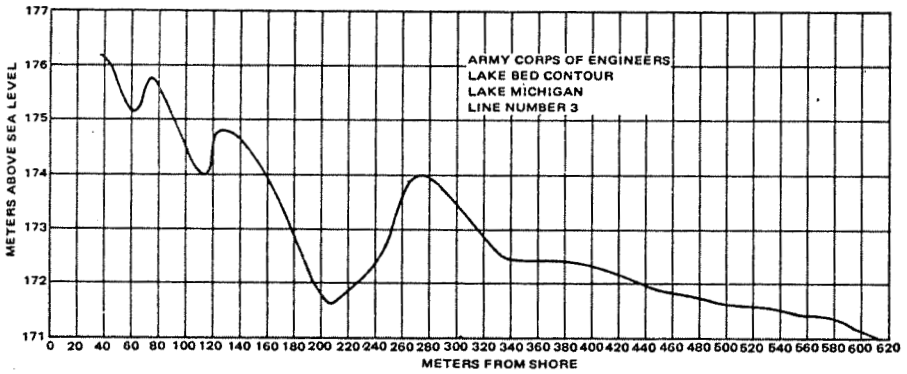
In this section, the limitations posed on the research study are discussed. Some comments are made in retrospect which may guide future multispectral data analysis. From the results, certain generalities are inferred for consideration.

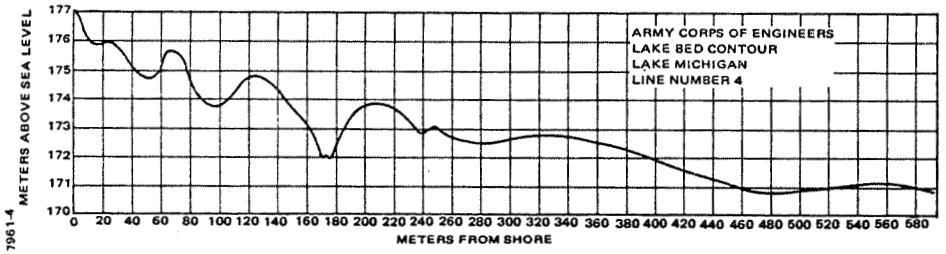
The principal conclusion drawn is the existence of water depth information in multispectral data, and further, that this information is recoverable. Both the contour maps and prediction lines demonstrate this. The most useful tool applied to the information extraction problem was the numerical technique of linear regression. This is due to the nature of the quantitative problem, i.e., the prediction of a dependent variable. The physical model suggested an exponential relation between reflectance and depth, thus posing initial restrictions on linear processing techniques. Measurable reflectance changes were shown to be asymptotic to a common value dependent only on scattering and absorption properties of the water. Multispectral data would not contain depth information beyond that value. Generalizing the results would require consideration of the "light" properties of the particular body of water being surveyed. Changes in mineral (or pollutant) content or salinity in altering the scattering and absorption properties would yield different quantitative answers. However, over water regions which may be assumed homogeneous, numerical results may be consistent.

In summary, the study provided important insight into the quantitative use of multispectral data. The water depth mapping is apparently feasible within the limitations discussed. A degree of operational accuracy sufficient to distinguish navigable from unnavigable depths would not be a difficult extension of the work completed.

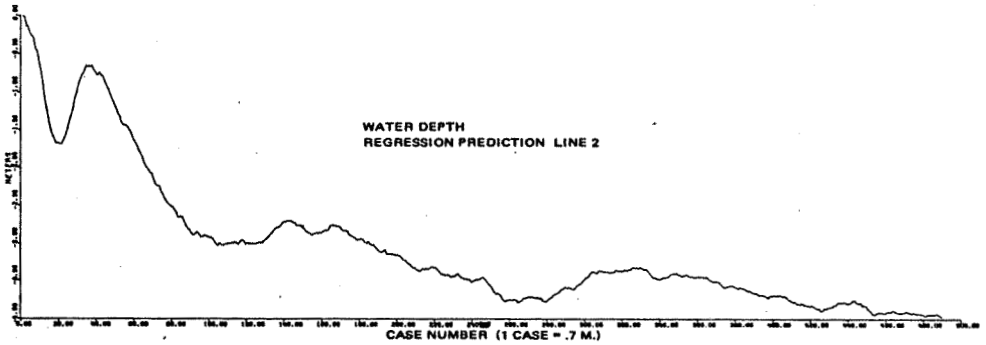
A hybrid processing approach employing three levels of processing would be appropriate for this purpose. The information contained in multispectral data is of the categorical type rather than the measurement type for which regression analysis is best suited. In recognizing this, a three-step processing scheme would first apply tests to determine the nature of the water observed. This categorical decision would specify the type of suspended and dissolved material affecting the optical properties of the water. The second step would determine the amount of such material suspended; in both cases, the conclusions are based on observation of water deep enough to have the asymptotic reflectance. With these conclusions, a "best" set of regression coefficients would be selected for the actual depth measurements. Such a scheme would make better use of the available information and could draw upon a base of research oriented towards solution of the more usual categorical decision problems put to multispectral data.

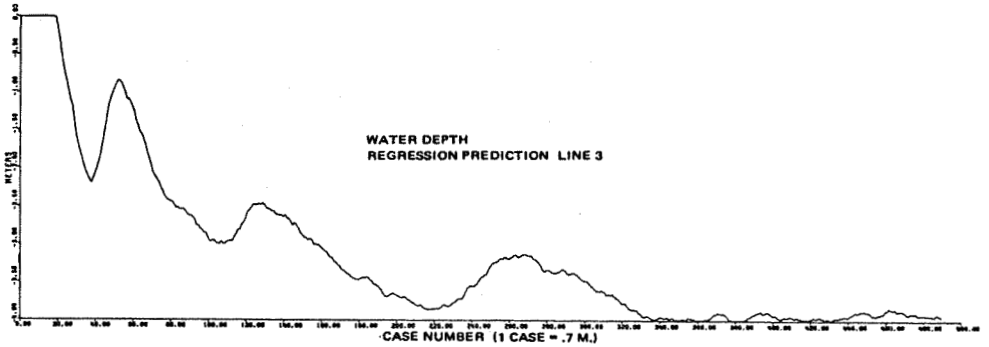






7961-4





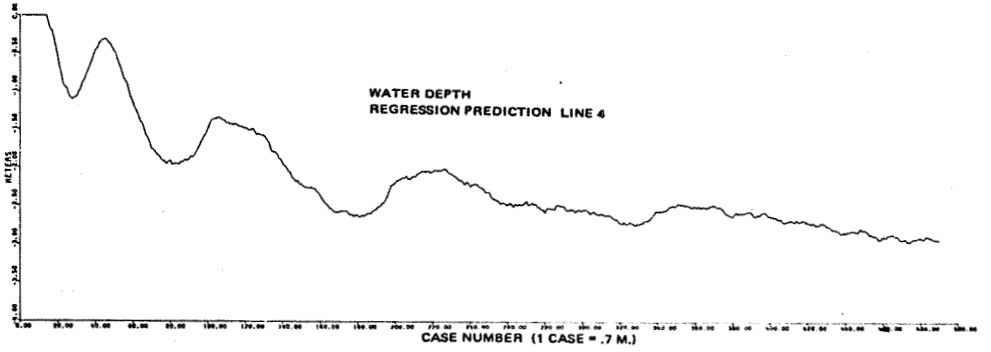


Figure 38 Actual and Predicted Water Depth Curves

7961-6

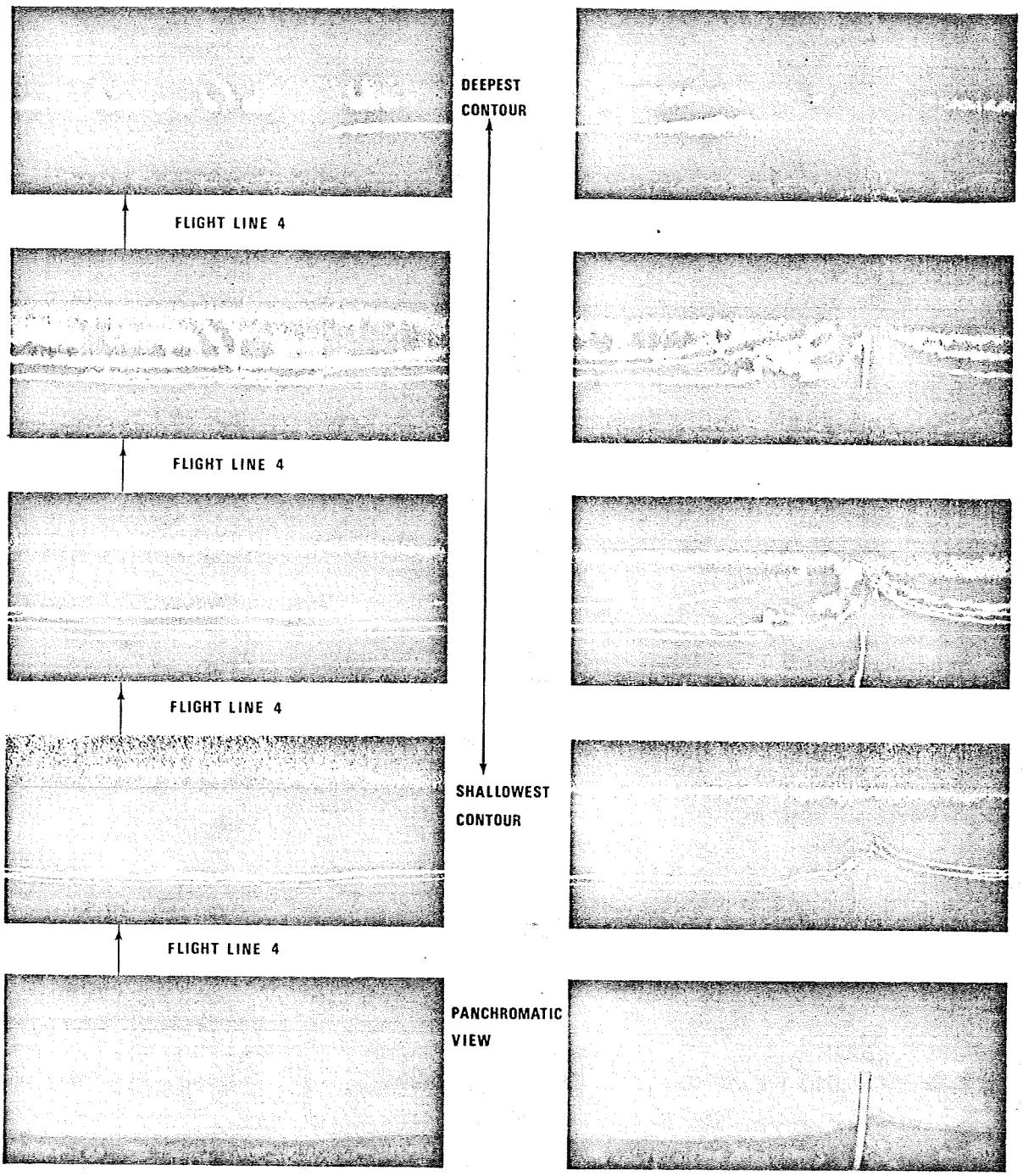


Figure 39 Water Depth Contours

APPENDIX A

GROUND TRUTH

The following report was prepared by Frederick Bevis for the Bendix Aerospace Systems Division. It describes the physical and chemical parameters of the Lake Michigan, Pentwater nearshore area, during the months of August and September 1969.

Physical and Chemical Parameters of Lake Michigan, Pentwater Nearshore Area.

A
report
for

Bendix Corporation
Aerospace Systems Division
Ann Arbor, Michigan 48107

by

Frederick B. Bevis
Department of Biology
Grand Valley State College
Allendale, Michigan 49401

August 1969

This study was undertaken at the request of Bendix Corporation, Aerospace Systems Division by Frederick B. Bevis, Grand Valley State College under the provisions of Bendix purchase order No. K1734, as amended 7/28/69. Data presented and the conclusions drawn in this report are those of the principal investigator, Frederick B. Bevis, and not those of Grand Valley State College. Release or publication of this report shall be at the discretion of Bendix Corporation, Aerospace Systems Division.

Contents

Comments on Observations

Table 1. All oxygen values are 100% saturated or greater, for all depths in the water column. Surface water temperature and bottom water temperature differences did not exceed 3.5°C on July 21, and 1.0°C on July 24, as shown in the table below.

Temperature Range in °C

Station	Date 7/21	7/24
2-1 (N, offshore)	2.7	0.9
2-2 (N, inshore)	3.5	0.1
4-1 (S, offshore)	2.6	1.0
4-2 (S, inshore)	0.7	1.0

Cool offshore Lake Michigan water from the NW mixed with the nearshore water on July 21 resulted in wider temperature differences. This condition did not obtain on July 24.

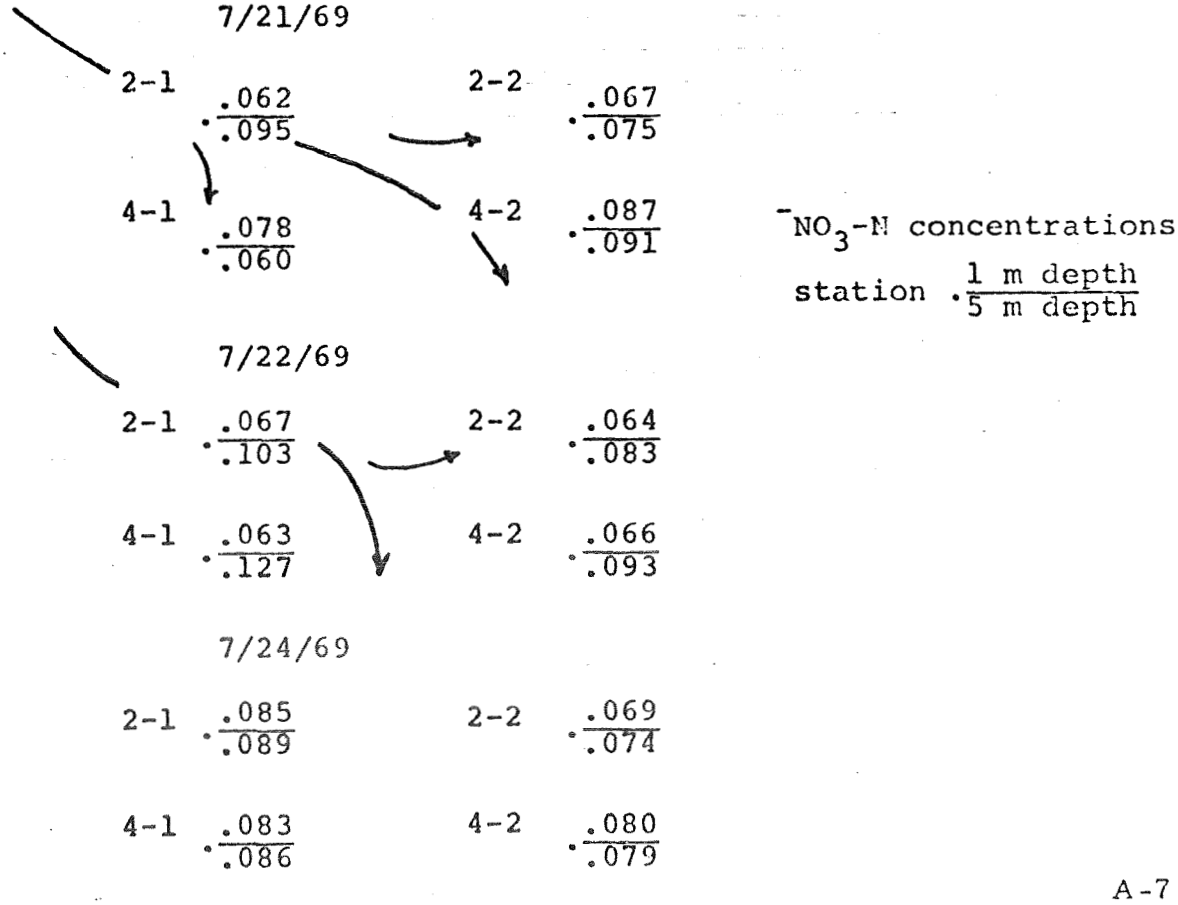
Table 2. No significant differences in pH, percent transmittance, and turbidity were found. Secchi disc readings indicate turbulence in the water column on July 21 and 22, with calm water (little mixing) on July 24.

Table 3. Evaluation of the pigment assemblage indicates a phytoplankton dominated by Bacillariophyceae (diatoms), with some Chlorophyceae (green algae) and Cyanophyceae (blue-green algae). Counts of the phytoplankton, of which diatom frustules/ml. are summarized in Table 5, enhance this conclusion. Diatoms generally exceeded all other algae by 100 - 1000 x. The concentrations of chlorophyll -a, -b, -c, total chlorophyll (C_t), and total carotenoids (TC) are as expected for 'unpolluted' Lake Michigan water. Higher concentrations of chlorophyll -c (and total chlorophyll -C_t) support the pulse (bloom) in diatoms reaching a peak on July 22 as shown for stations 2-2 (N, inshore) and 4-1 (S, offshore) in Table 5.

Table 4. Seston levels indicate mixed water on July 21, with considerable detritus (non-living organic matter) in the water column. Settling of suspended particulate matter occurred on July 22, particularly along the south line (stations 4-1 and 4-2). Low seston levels obtain in the calm water of July 24, where essentially all seston recorded was living phytoplankton.

Table 5. A diatom 'pulse' is indicated for all stations reaching a peak on July 22. Maximum frustules/ml. were counted for stations 2-2 (N, inshore) and 4-1 (S, offshore). This 'pulse' subsided by July 24, apparently associated with Nitrate Nitrogen ($\bar{N}O_3$-N) depletion as shown in Table 6.

Table 6. Nitrate nitrogen ($\bar{N}O_3$-N) concentrations increased for all stations, except 4-2 (S, inshore), from July 21 to July 22. This increase may be related to prevailing NW winds and southern drift of the nearshore water carrying nitrates from the source at the outlet of the Pentwater channel into the study area as shown in the following diagrams.



Nitrate nitrogen is the 'critical' nutrient in the regulation of phytoplankton blooms. The decrease in nitrate from July 22 to July 24 is apparently related to utilization by phytoplankton during the diatom 'pulse.'

Table 7. Ferric iron (Fe^{+++}) concentrations show the same distribution pattern as nitrate nitrogen in the above diagrams. Pronounced ferric iron stratification on July 24 is most likely related to precipitation as Fe_3PO_4 or the settling of particulate organic matter with 'bond' ferric ion. Ferrous iron (Fe^{++}) was not detectable (test sensitivity 0.001 PPM).

Table 8. Ortho-phosphate ($\bar{P}O_4$-P) and total phosphate (T-$\bar{P}O_4$) levels show no significant differences from station to station, depth to depth, and day to day. Values as expected for Lake Michigan nearshore water.

Table 9. CO_2, $\bar{S}O_4$, $\bar{N}O_2$, and TDS levels show no differences from station to station, depth to depth, and day to day. Values as expected for Lake Michigan nearshore water.

Table 10. Calcium hardness as ppm $CaCO_3$ shows no differences from station to station, depth to depth, and day to day. Values as expected for Lake Michigan nearshore water.

Table 11. ODI and COD levels show little difference from station to station, depth to depth, and day to day, with the possible exception of July 24 where the slight decrease in ODI and COD for all stations may be related to the decrease in amount of organic matter in the water column (see Table 4, Seston in mg/m^3).

Table 12. All parameters measured for the Pentwater Channel surface water were as expected for a 'polluted' lake discharge into Lake Michigan. Relatively low nitrate nitrogen concentrations are probably directly related to the 'bloom' of blue-green and green algae occurring on both sample days, July 23 and July 24. Seston levels support this argument, as does the high concentrations of chlorophyll -a and other pigments.

Procedures

1. Temperature (°C) and dissolved oxygen (ppm) measured direct with YSI temperature compensated Oxygen Analyzer.

2. pH measured on Beckman Expandomatic pH Meter. Percent transmittance and turbidity (JTU) measure direct on Hach DR Colorimeter.

3. Chlorophyll -a, -b, -c, total chlorophyll (C_t) and total carotenoid (TC) calculated from the Strickland - Parsons equations (A Practical Handbook of Seawater Analysis, Bull. 167, Fisheries Res. Board, Canada, 1968) as given below.

   $C\text{-}a = 11.6E_{6650} - 1.31\ E_{6450} - 0.14\ E_{6300}$.

   $C\text{-}b = 20.7\ E_{6450} - 4.34\ E_{6650} - 4.42\ E_{6300}$.

   $C\text{-}c = 55\ E_{6300} - 4.64\ E_{6650} - 16,3\ E_{6450}$.

   $TC = 4.0\ E_{4800}$.

4. Seston determinations follow procedure of Strickland and Parsons (3 above).

5. Diatoms counted according to method of Holland, R. 1969, "Seasonal FLuctuations of Lake Michigan Diatoms," Limno. and Oceanogr. 14(2): 423-436.

6. Nitrogen, Nitrate - Cadmium reduction method with 1-Naphthylamine-Sulfanilic Acid.

7. Iron, Ferric - 1, 10-Phenanthroline and 2,4,6-Tripyridyl-s-triazine methods used.

8. Phosphate determinations by ammonium Molybdate-Stannous reduction method.

9. CO_2 - titration with sodium hydroxide with phenolphthalein indicator.

   $^{-}SO_4$- Barium sulfate turbidimetric method.

   $^{-}NO_2$- similar to $^{-}NO_3$-N method with cadmium reductor omitted.

   TDS - direct in ppm NaCl on Hach Conductivity Meter.

10. Calcium Hardness - titration with Disodium Dihydrogen -1, 2-Cyclohexanediaminetetraacetate (CDTA).

11. ODI - dichromate-sulfuric acid oxidation.

    COD - sulfuric acid - silver sulfate oxidation with potassium dichromate.

Table 1. Temperature in °C[1] and Dissolved Oxygen (DO)[2] in PPM, Lake Michigan, Pentwater Nearshore Area, July 1969.

Station	Depth (m)	T(°C)	DO(PPM)
	- July 21, 1969 -		
2-1 (N, offshore)	1	17.4	10.3
	5	17.5	11.3
2-2 (N, inshore)	1	20.2	11.3
	5 (4 1/2)	16.8	10.9
4-1 (S, offshore)	1	17.3	10.3
	5	16.7	11.2
4-2 (S, inshore)	1	18.6	10.8
	5 (4 1/2)	18.1	11.2
	- July 24, 1969 -		
2-1	1	19.8	9.9
	5	19.4	11.0
2-2	1	19.0	9.9
	5 (4 1/2)	18.9	11.1
4-1	1	20.0	9.7
	5	19.4	10.6
4-2	1	20.1	9.9
	5 (4 1/2)	19.2	10.7

1) and 2) $\bar{X}$ of 3 observations immediately above and including the reading at the depth given.

Table 2. pH, % Transmittance, Turbidity (JTU)[1], and Secchi Disc Readings, Lake Michigan, Pentwater Nearshore Area, July 1969.

Station	Depth (m)	pH	% T	JTU[2]	Secchi Disc[3]
	- July 21, 1969 -				
2-1 (N, offshore)	1	8.79	99.8	7	15
	5	8.81	99.7	8	
2-2 (N, inshore)	1	8.79	99.7	7	10
	5 (4 1/2)	8.79	100.0	8	
4-1 (S, offshore)	1	8.80	99.8	5	16
	5	8.77	99.6	7	
4-2 (S, inshore)	1	8.78	99.5	7	11
	5 (4 1/2)	8.78	99.7	4	
	- July 22, 1969 -				
2-1	1	8.77	99.7	9	15.5
	5	8.82	99.8	10	(13)
2-2	1	8.80	99.7	5	10
	5 (4 1/2)	8.77	99.7	6	(10)
4-1	1	8.82	99.8	7	15
	5	8.81	99.7	6	(14)
4-2	1	8.79	99.7	6	12
	5 (4 1/2)	8.81	99.8	5	(12.5)
	- July 24, 1969 -				
2-1	1	8.80	100.0	7	16
	5	8.82	98.6	12	
2-2	1	8.79	99.4	10	13
	5 (4 1/2)	8.80	99.6	8	
4-1	1	8.82	99.7	8	16
	5	8.81	99.8	7	
4-2	1	8.83	99.6	7	14
	5 (4 1/2)	8.80	99.8	6	

1) and 2) JTU - Jackson Turbidity Units. Test sensitive to 1 JTU.

3) Secchi disc readings in Ft. by Robert Breault, Bendix Corporation, Aerospace Systems Division. July 21st reading at 3:00 p.m., July 22nd at 10:45 a.m., except for ( ) at 8:30 a.m., and July 24th at 8:00 a.m.

Table 3. Chlorophyll -a, -b, -c, Total Chlorophyll (C_t) and Total Carotenoid (TC) Levels[1] in mg pigment/m^3, Lake Michigan, Pentwater Nearshore Area, July 1969.

Station	Depth (m)	Level -a	-b	-c	C_t	TC
	- July 21, 1969 -					
2-1 (N, offshore)	1	2.87	3.51	5.02	11.40	3.44
	5	3.02	3.04	1.11	7.17	4.60
2-2 (N, inshore)	1	2.86	2.12	3.62	8.60	4.48
	5 (4 1/2)	7.43	2.82	3.45	13.72	5.12
4-1 (S, offshore)	1	2.85	3.52	3.67	10.04	3.20
	5	2.29	2.06	2.96	7.31	3.52
4-2 (S, inshore)	1	1.40	1.40	3.33	6.13	2.68
	5 (4 1/2)	2.03	2.00	2.26	6.29	3.20
	- July 22, 1969 -					
2-1	1	1.51	2.26	9.82	13.59	3.64
	5	2.49	2.10	8.00	12.59	3.96
2-2	1	1.36	1.82	9.50	12.68	4.04
	5 (4 1/2)	13.85	9.31	30.46	53.62	6.64
4-1	1	2.69	3.53	12.56	18.78	5.20
	5	5.26	3.86	10.70	19.82	4.32
4-2	1	1.77	3.40	0.25	5.42	2.76
	5 (4 1/2)	2.55	2.81	8.95	14.31	5.24
	- July 24, 1969 -					
2-1	1	1.57	2.09	10.46	14.12	3.24
	5	1.75	2.59	11.45	15.79	3.76
2-2	1	2.12	2.77	11.56	16.45	3.52
	5 (4 1/2)	2.45	3.79	11.74	17.98	3.72
4-1	1	2.01	2.42	8.84	13.27	3.48
	5	1.80	1.59	7.39	10.78	2.36
4-2	1	1.15	1.51	7.98	10.64	2.48
	5 (4 1/2)	1.56	1.67	8.77	12.00	2.36

1) 20 hour extraction @ 4°C in 90% acetone. Extinction values @ 7500 Å, 6650 Å, 6450 Å, 6300 Å, and 4800 Å.

Table 4. Seston[1] Levels in mg/m^3, Lake Michigan, Pentwater Nearshore Area, July 1969.

Station	Depth (m)	Level
	- July 21, 1969 -	
2-1 (N, offshore)	1	95
	5	110
2-2 (N, inshore)	1	70
	5 (4 1/2)	95
4-1 (S, offshore)	1	70
	5	70
4-2 (S, inshore)	1	95
	5 (4 1/2)	45
	- July 22, 1969 -	
2-1	1	45
	5	45
2-2	1	95
	5 (4 1/2)	95
4-1	1	45
	5	110
4-2	1	45
	5 (4 1/2)	110
	- July 24, 1969 -	
2-1	1	24
	5	21
2-2	1	15
	5 (4 1/2)	19
4-1	1	24
	5	19
4-2	1	19
	5 (4 1/2)	14

1) Seston includes both living organisms, phytoplankton and zooplankton, and the non-living particles of organic matter floating or held in suspension in the water.

Table 5. Diatoms (Frustules)/ml.[1], Lake Michigan, Pentwater Nearshore Area, July 1969.

Station	Depth (m)	Count
	- July 21, 1969 -	
2-1 (N, offshore)	1	142
	5	110
2-2 (N, inshore)	1	219
	5 (4 1/2)	733
4-1 (S, offshore)	1	187
	5	605
4-2 (S, inshore)	1	195
	5 (4 1/2)	107
	- July 22, 1969 -	
2-1	1	225
	5	430
2-2	1	318
	5 (4 1/2)	1736
4-1	1	262
	5	1038
4-2	1	373
	5 (4 1/2)	429
	- July 24, 1969 -	
2-1	1	182
	5	461
2-2	1	513
	5 (4 1/2)	1105
4-1	1	357
	5	296
4-2	1	304
	5 (4 1/2)	361

1) $\bar{X}$ of 3 - 1 ml. counts/depth/station.

Table 6. Nitrate Nitrogen (NO_3-N) Levels in PPM[1], Lake Michigan, Pentwater Nearshore Area, July 1969.

Station	Depth (m)	Level
	- July 21, 1969 -	
2-1 (N, offshore)	1	0.062
	5	0.095
2-2 (N, inshore)	1	0.067
	5 (4 1/2)	0.075
4-1 (S, offshore)	1	0.078
	5	0.060
4-2 (S, inshore)	1	0.087
	5 (4 1/2)	0.091
	- July 22, 1969 -	
2-1	1	0.067
	5	0.103
2-2	1	0.064
	5 (4 1/2)	0.083
4-1	1	0.063
	5	0.127
4-2	1	0.066
	5 (4 1/2)	0.093
	- July 24, 1969 -	
2-1	1	0.085
	5	0.089
2-2	1	0.069
	5 (4 1/2)	0.074
4-1	1	0.083
	5	0.086
4-2	1	0.080
	5 (4 1/2)	0.079

1) $\bar{X}$ of 3 tests/depth/station

Table 7. Ferric Iron (Fe^{+++}) Levels in PPM[1], Lake Michigan, Pentwater Nearshore Area, July 1969.

Station	Depth (m)	Level
	- July 21, 1969 -	
2-1 (N, offshore)	1	0.045
	5	0.021
2-2 (N, inshore)	1	0.006
	5 (4 1/2)	0.041
4-1 (S, offshore)	1	0.025
	5	0.040
4-2 (S, inshore)	1	0.036
	5 (4 1/2)	0.035
	- July 22, 1969 -	
2-1	1	0.039
	5	0.042
2-2	1	0.013
	5 (4 1/2)	0.014
4-1	1	0.028
	5	0.038
4-2	1	0.026
	5 (4 1/2)	0.014
	- July 24, 1969 -	
2-1	1	0.005
	5	0.034
2-2	1	0.006
	5 (4 1/2)	0.019
4-1	1	0.005
	5	0.025
4-2	1	0.006
	5 (4 1/2)	0.033

1) $\bar{X}$ of 3 tests/depth/station

Table 8. Ortho-Phosphate[1] ($^{-}PO_4$-P) and Total Phosphate (T-$^{-}PO_4$) Levels[1] in PPM, Lake Michigan, Pentwater Nearshore Area, July 1969.

Station	Depth (m)	Level	
		$^{-}PO_4$-P	T-$^{-}PO_4$
	- July 21, 1969 -		
2-1 (N, offshore)	1	0.001	0.005
	5	0.005	0.005
2-2 (N, inshore)	1	0.006	0.006
	5 (4 1/2)	0.003	0.004
4-1 (S, offshore)	1	0.005	0.005
	5	0.005	0.005
4-2 (S, inshore)	1	0.002	0.004
	5 (4 1/2)	0.004	0.004
	- July 22, 1969 -		
2-1	1	0.003	0.005
	5	0.003	0.004
2-2	1	0.003	0.003
	5 (4 1/2)	0.004	0.004
4-1	1	0.002	0.003
	5	0.002	0.002
4-2	1	0.001	0.003
	5 (4 1/2)	0.004	0.005
	- July 24, 1969 -		
2-1	1	0.004	0.005
	5	0.004	0.005
2-2	1	0.005	0.007
	5 (4 1/2)	0.003	0.006
4-1	1	0.003	0.005
	5	0.004	0.005
4-2	1	0.004	0.006
	5 (4 1/2)	0.002	0.004

1) $\bar{X}$ of 3 tests/depth/station. Test sensitive to 0.001 PPM.

Table 9. CO_2, $^{-}SO_4$, $^{-}NO_2$, and TDS[1] Levels[2] in PPM, Lake Michigan, Pentwater Nearshore Area, July 1969.

Station	Depth (m)	Level CO_2	$^{-}SO_4$	$^{-}NO_2$	TDS
	- July 21, 1969 -				
2-1 (N, offshore)	1	7.0	24	0.001	130
	5	6.9	23	0.002	130
2-2 (N, inshore)	1	7.0	23	0.001	129
	5 (4 1/2)	7.0	24	0.001	130
4-1 (S, offshore)	1	7.0	22	0.001	130
	5	7.0	24	0.002	129
4-2 (S, inshore)	1	7.0	23	0.001	130
	5 (4 1/2)	7.0	22	0.002	130
	- July 22, 1969 -				
2-1	1	7.0	24	0.001	129
	5	7.0	24	0.002	130
2-2	1	7.0	23	0.001	131
	5 (4 1/2)	7.0	24	0.002	131
4-1	1	7.0	23	0.001	129
	5	7.0	24	0.002	129
4-2	1	7.0	23	0.001	132
	5 (4 1/2)	6.9	23	0.002	132
	- July 24, 1969 -				
2-1	1	7.0	24	0.001	134
	5	7.0	24	0.001	134
2-2	1	6.9	23	0.001	134
	5 (4 1/2)	7.0	23	0.002	135
4-1	1	7.0	23	0.001	134
	5	7.0	24	0.002	132
4-2	1	7.0	23	0.001	135
	5 (4 1/2)	7.0	24	0.002	137

1) CO_2-Carbon Dioxide; $^{-}SO_4$-Sulfate; $^{-}NO_2$ Nitrite Nitrogen; and TDS - Total Dissolved Solids as PPM NaCl.

2) $\bar{X}$ of 3 tests/depth/station. Test sensitivity: CO_2- 0.2 PPM; $^{-}SO_4$- 1 PPM; $-NO_2$- 0.001 PPM; and TDS - 5 PPM.

Table 10. Calcium Hardness Levels[1] in PPM $CaCO_3$, Lake Michigan, Pentwater Nearshore Area, July 1969.

Station	Depth (m)	Level
	- July 21, 1969 -	
2-1 (N, offshore)	1	96
	5	95
2-2 (N, inshore)	1	92
	5 (4 1/2)	89
4-1 (S, offshore)	1	89
	5	94
4-2 (S, inshore)	1	93
	5 (4 1/2)	89
	- July 22, 1969 -	
2-1	1	92
	5	97
2-2	1	94
	5 (4 1/2)	92
4-1	1	95
	5	91
4-2	1	93
	5 (4 1/2)	95
	- July 24, 1969 -	
2-1	1	98
	5	92
2-2	1	95
	5 (4 1/2)	95
4-1	1	92
	5	93
4-2	1	93
	5 (4 1/2)	90

1) $\bar{X}$ of 6 tests/depth/station. Test sensitive to 5 PPM $CaCO_3$.

Table 11. Oxygen Demand Index[1] (ODI) and Chemical Oxygen Demand (COD) Levels[1] in PPM, Lake Michigan, Pentwater Nearshore Area, July 1969.

Station	Depth (m)	ODI[2]	COD
	- July 21, 1969 -		
2-1 (N, offshore)	1	0.9	10
	5	1.0	11
2-2 (N, inshore)	1	1.0	14
	5 (4 1/2)	1.0	12
4-1 (S, offshore)	1	0.9	13
	5	1.0	12
4-2 (S, inshore)	1	1.0	10
	5 (4 1/2)	0.9	11
	- July 22, 1969 -		
2-1	1	1.0	14
	5	1.0	11
2-2	1	1.0	9
	5 (4 1/2)	0.9	14
4-1	1	0.9	15
	5	0.9	13
4-2	1	1.0	15
	5 (4 1/2)	1.1	10
	- July 24, 1969 -		
2-1	1	0.8	8
	5	0.9	10
2-2	1	0.8	9
	5 (4 1/2)	0.8	10
4-1	1	1.0	11
	5	0.9	10
4-2	1	0.9	9
	5 (4 1/2)	0.9	8

1) Test sensitivity: ODI - 0.1 PPM; COD - 5 PPM.

2) ODI, an approximation of normal BOD (Biochemical Oxygen Demand).

Table 12. Some Physical and Chemical Parameters of the Pentwater Channel Surface Water, July 1969.

	Temperature (°C)	Turbidity (JTU)	TDS[1]	pH
7/23	21.3	25	125	8.81
7/24	--	32	148	8.79

PPM[2]

	DO	CO_2	$-NO_3-N$	$-NO_2$	$-PO_4-P$	$T-PO_4$	ODI	COD
7/23	10.5	5.5	0.043	0.000	0.007	0.007	1.8	21
7/24	9.8	6.5	0.042	0.001	0.007	0.008	1.9	26

PPM[2]

	Calcium Hardness[3]	Fe^{+++}-Iron	Fe^{++}-Iron	$-SO_4$
7/23	114	0.081	0.002	19
7/24	105	0.088	0.002	20

mg/m^3

	Seston	Chlorophyll -a	-b	-c	C_t[4]	TC[4]
7/23	139	16.80	10.22	14.36	41.38	18.88
7/24	195	11.95	12.69	6.53	31.17	16.12

1) TDS - Total Dissolved Solids as PPM NaCl.

2) $\bar{X}$ of 3 tests/sample.

3) as PPM $CaCO_3$.

4) C_t - Total Chlorophyll; TC - Total Carotenoids.

"Pentwater 2" Physical and Chemical Parameters of Lake Michigan, Pentwater Nearshore Area.

A
report
for

Bendix Corporation
Aerospace Systems Division
Ann Arbor, Michigan 48107

by

Frederick B. Bevis
Department of Biology
Grand Valley State College
Allendale, Michigan 49401

September 1969

This study was undertaken at the request of Bendix Corporation, Aerospace Systems Division, by Frederick B. Bevis, Grand Valley State College, under the provisions of Bendix purchase order No.K-3879. Data and conclusions presented in this report are those of the principal investigator, Frederick B. Bevis, and not those of Grand Valley State College. Release or publication of this report shall be at the discretion of Bendix Corporation, Aerospace Systems Division.

Contents

Comments on Observations

Table 1. All oxygen values are 100% saturated or greater, for all depths in the water column. This situation obtained in July (see 'Pentwater 1' report, Table 1, August 1969). Surface water temperature and bottom water temperature differences did not exceed 0.5°C, inshore or offshore, indicating well 'mixed' conditions, as shown in the table below.

Temperature Range in °C

Station	Temp. Range
4-1 (S, offshore)	0.4
4-2 (S, inshore)	0.5

Table 2. Evaluation of the pigment assemblage indicates a phytoplankton dominated by Bacillariophyceae (diatoms) with a greater abundance of chlorophyceae (green algae), particularly the filamentous algae - Cladophora glomerata, then previously noted in the July survey ('Pentwater 1', August 1969 report). This green algae forms 'windrows' of decaying filaments along the Lake Michigan shoreline particularly in the vicinity of the outflow of the larger rivers bordering Lake Michigan.

In general, the chlorophyll pigment concentrations for September are lower than those measured in July. Total carotenoid levels (TC) appear similar to those measured in July.

Table 3. Seston levels apparently indicate mixed water with considerable detritus (non-living organic matter) in the water column. The maximum level occurring at 1m. for both stations 4-1 and 4-2 correlate well with the diatom counts (Table 4) for this depth at both stations. No evidence of a diatom 'pulse' is indicated for any depth at the sampling stations. The maximum of 465 frustules/ml. is approximately one-quarter of the high count of 1736 for 2-2 Station at 5 meters, obtained July 22, 1969 (Table 5, 'Pentwater 1', August 1969 report).

Table 4. These are relatively low counts of frustules/ml. for September in the nearshore waters of Lake Michigan.

Table 5. Iron (Fe+++) depletion is evident for both station 4-1 and 4-2, at the 1m. depth, whereas the $\bar{N}O_3$-N levels are greater at this depth. Ortho-phosphate ($\bar{P}O_4$-P) levels are approximately 100 X those of July (Table 8., 'Pentwater 1; August 1969). The July phosphate levels reflect the normal depletion associated with plankton utilization, while the September levels indicate the typical phosphate 'replenishment' associated with the onset of fall and decay of detritus in the water column.

The pH and TDS levels of the stations and depths sampled are summarized in the table below.

pH and TDS[1)]

Station	Depth (m).	pH	TDS
4-1 (S, offshore)	0	8.87	115
	1	8.91	114
	5	8.93	114
4-2 (S, inshore)	0	8.91	115
	1	8.89	112

1) TDS-Total dissolved solids in PPM NaCl.

Table 6. Nitrate nitrogen ($\bar{N}O_3$-N) and ortho-phosphate ($\bar{P}O_4$-P) levels are considerably higher than reported for July (Table 12, 'Pentwater 1', August 1969). This condition is typical of the fall period of nutrient recycling ('replenishment') in fresh water lakes.

Table 1. Temperature in °C[1] and Dissolved Oxygen (DO)[2] in PPM, Lake Michigan, Pentwater Nearshore Area, September 4, 1969.

Station	Depth (m)	T(°C)	DO(PPM)
4-1 (S, offshore)	0	21.8	9.85
	1	21.9	9.25
	5	21.8	8.73
4-2 (S, inshore)	0	22.2	10.30
	1	22.1	9.95

1) and 2) $\bar{X}$ of 3 observations immediately above and including the reading at the depth given, with the exception of surface water temperature and DO (0) which is as recorded in the field.

Table 2. Chlorophyll -a, -b, -c, Total Chlorophyll (C_t) and Total Carotenoid (TC) Levels[1] in mg pigment/m^3, Lake Michigan, Pentwater Nearshore Area, September 4, 1969.

Station	Depth (m)	Level				
		-a	-b	-c	C_t	TC
4-1 (S, offshore)	0	1.80	1.93	4.28	8.01	4.20
	1	2.22	2.92	4.63	9.77	4.40
	5	2.60	2.57	3.94	9.11	4.60
4-2 (S, inshore)	0	1.17	1.67	3.28	6.66	3.00
	1	1.74	2.29	2.44	6.47	4.16

1) 20 hour extraction @ 4°C in 90% acetone. Extinction values @ 7500Å, 6650Å, 6450Å, 6300Å, and 4800Å.

Table 3. Seston [1] Levels in mg/m^3 Lake Michigan, Pentwater Nearshore Area, September 4, 1969.

Station	Depth (m)	Level
4-1 (S, offshore)	0	18
	1	72
	5	40
4-2 (S, inshore)	0	34
	1	74

1) Seston includes both living organisms, phytoplankton and zooplankton, and the non-living particles of organic matter floating or held in suspension in the water.

Table 4. Diatoms (Frustules)/ml.[1], Lake Michigan, Pentwater Nearshore Area, September 4, 1969.

Station	Depth (m)	Count
4-1 (S, offshore)	0	207
	1	416
	5	311
4-2 (S, inshore)	0	176
	1	465

1) $\bar{X}$ of 3 -1 ml. counts/depth/station.

Table 5. Nitrate Nitrogen ($\bar{N}O_3$-N), Ortho-Phosphate ($\bar{P}O_4$-P), and Ferric Iron (Fe+++) Levels in PPM[1], Lake Michigan, Pentwater Nearshore Area, September 4, 1969.

Station	Depth (m)	$\bar{N}O_3$-N	$\bar{P}O_4$-P	Fe+++
4-1 (S, offshore)	0	0.053	0.07	0.042
	1	0.078	0.05	0.005
	5	0.059	0.06	0.019
4-2 (S, inshore)	0	0.067	0.06	0.020
	1	0.094	0.05	0.004

1) $\bar{X}$ of 3 tests/depth/station.

Table 6. Some Physical and Chemical Parameters of the Pentwater Channel Surface Water, September 4, 1969.

Temperature		TDS[1]			pH
20.7		143			8.97
PPM[2]					
DO	$\bar{N}O_3$-N		$\bar{P}O_4$-P		Fe+++
9.65	0.056		0.07		0.041
mg/m^3					
Seston	Chlorophyll -a	-b	-c	C_t[3]	TC[3]
74	8.79	6.22	7.56	22.57	9.56

1) TDS - Total Dissolved Solids as PPM NaCl.

2) $\bar{X}$ of 3 tests/sample.

3) C_t - Total Chlorophyll; TC- Total Carotenoids.

APPENDIX B

FACTOR ANALYSIS

The matrix demonstration which follows shows the basis for factor analysis and that the unrotated factors first obtained by diagonalization of the correlation matrix are statistically independent.

For a data matrix, ${}_mX_n$, consisting of n column vectors, each column representing a complete spectral observation, and m rows, each row representing a single channel of the observation, the m by m correlation matrix, R, is

$$ {}_mR_m = \frac{1}{n} = \frac{1}{n} S^{-1} (X - \bar{X}) (X - \bar{X})^T S^{-1} \qquad (B-1) $$

where the following definitions have been used, m being the number of channels and n the number of observation:

(1) $\bar{X}$ = the m by n matrix of mean values

$$ = \begin{bmatrix} \bar{X}_1 & \bar{X}_1 & \bar{X}_1 \\ \bar{X}_2 & \bar{X}_2 & \\ \bar{X}_m & & \bar{X}_m \end{bmatrix} $$

where

$$ \bar{X} = \frac{\sum_{y=1}^{n} X_{ij}}{n} \quad , i = 1, \ldots, m $$

(2) S = the diagonal matrix of standard deviations

$$= \begin{bmatrix} S_1 & & & O \\ & S_2 & & \\ & & \ddots & \\ O & & & S_m \end{bmatrix}$$

where

$$S_i = \left[\frac{\sum_{j=1} \left(X_{ij} - \bar{X}_i \right)^2}{n} \right]^{1/2}$$

and

$$S^{-1} = \begin{bmatrix} 1/S_1 & & & O \\ & 1/S_2 & & \\ & & \ddots & \\ O & & & 1/S_m \end{bmatrix}$$

These operations transform the data to standard form. Subsequent expressions using the matrix X are presented in a simpler form by assuming that the data are in standard form.

R is the real symmetric correlation matrix which can be diagonalized by a unitary transformation to obtain:

$$D(\lambda_1, \lambda_2, \ldots \lambda_m) = U^T R U \qquad \text{(B-2)}$$

Define

$$\sqrt{D} \stackrel{d}{=} D\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots \sqrt{\lambda_m}\right) . \qquad \text{(B-3)}$$

It follows that

$$U^T R U = \sqrt{D}\sqrt{D} \qquad \text{(B-4)}$$

and

$$\sqrt{D}^{-1} U^T R U \sqrt{D}^{-1} = I \tag{B-5}$$

which can also be written as

$$\sqrt{D}^{-1} U^T X X^T U \sqrt{D}^{-1} = nI . \tag{B-6}$$

A factor score matrix, F, can now be identified and defined as

$$\begin{aligned} {}_m F_n &= \sqrt{D}^{-1} U^T X \\ &\triangleq AX . \end{aligned} \tag{B-7}$$

The correlation matrix of factor scores follows from the definition and from Equation B-6. Notice that

$$\frac{1}{n} {}_m F_n {}_n F^T_m = {}_m I_m \tag{B-8}$$

is the identity matrix. Since all off diagonal elements of the correlation matrix are zero, the factor scores are uncorrelated.

The eigenvalues, $\lambda_1, \lambda_2, \ldots \lambda_m$, of a real symmetric matrix such as R, are positive definite. Each eigenvalue represents the amount of variation found in the direction of its corresponding eigenvector.

The communality of a variable is the fraction of the variance of the variable which is explained by the common factors, the remainder being attributable to a specific factor for the variable and the error variance. The reduced correlation matrix is obtained by substituting estimates of the communalities for the diagonal elements of the correlation matrix. Estimates for the communalities are acceptable only if the reduced matrix is positive semi-definite. The foregoing derivation also holds for the reduced matrix. Since the matrix is real and symmetric, an orthogonal transformation U exists for obtaining the diagonal matrix D, and the eigenvalues are all positive or zero, because of the positive definite character of the reduced correlation matrix. The principal component analysis diagonalizes the correlation matrix (ones along the diagonal); whereas, the principal-factor analysis operates on the reduced correlation matrix (communalities on the diagonal).

Rotated analyses are derived from the initial factor matrix. A straightforward rotation of factor axes following an exact analytic criterion results in a new set of factors. The rotation is accomplished by applying a minimizing or maximizing process. The VARIMAX procedure attempts to simplify the

factors by extracting components that tend toward unity or zero while maintaining the zero correlation of the factors. OBLIMIN rotations involve minimizing similar expressions without the requirement of maintaining zero correlation. An arbitrary constant ranging from 0 to 1 provides an infinite number of solutions. When the constant is close to unity, the resulting factors are generally uncorrelated. A near-zero constant yields the most simple factor components (many zero weights) but the factors are highly correlated. Typically, a constant of 0.5 is used to control correlation while providing some simplification.