

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/200458940>

# Near-Infrared Analysis as a Rapid Method to Simultaneously Evaluate Several Soil Properties

Article in *Soil Science Society of America Journal* · January 1995

DOI: 10.2136/sssaj1995.03615995005900020014x

CITATIONS

454

READS

1,094

2 authors, including:



[Eyal Ben-Dor](#)

Tel Aviv University

234 PUBLICATIONS 6,387 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



COST Action on Environmental Monitoring with UAS - HARMONIOUS [View project](#)



Special Issue "Selected Papers from the First "Symposia of Remote Sensing Applied to Soil Science", as part of the "21st World Congress of Soil Science, 2018"" [View project](#)

## Near-Infrared Analysis as a Rapid Method to Simultaneously Evaluate Several Soil Properties

E. Ben-Dor\* and A. Banin

### ABSTRACT

The near infrared analysis (NIRA) approach was studied to examine its capability for predicting spectral feature soil properties from the reflectance curves in the near infrared (NIR) region (1–2.5  $\mu\text{m}$ ) of arid and semiarid soils. High-resolution diffuse reflectance spectra (3113 spectral points) in the NIR region were recorded for 91 soil samples from Israel. Six soil properties (clay content, specific surface area, cation-exchange capacity, hygroscopic moisture, carbonate content, and organic matter content) were measured by routine methods employed in soil laboratories. An empirical model to predict each property from its spectral signature was developed by adapting the NIRA technique. Several data manipulations were used to obtain optimum performance. The optimum performance of all properties was found to be between 25 and 63 spectral points. Strong support for the NIRA capability was provided by its ability to examine most of the spectral assignments. A slight bias was observed for the prediction of both organic matter and hygroscopic moisture, suggesting that more attention in the prediction of these constituents is required. It was concluded that NIRA is a promising method for rapid and nonrestrictive analysis of soil materials, and further study of the synergism between NIRA and soil materials is recommended.

**N**EAR-INFRARED ANALYSIS is a laboratory approach that analyzes the diffuse reflectance radiation with regard to a material's chemistry (Stark et al., 1986). This methodology was developed 25 yr ago for rapid analysis of moisture in grains (Ben-Gera and Norris, 1968). Today, NIRA is well accepted and widely used in many other disciplines (Davies and Grant, 1987; Norris, 1988). This method is known for its rapidity, convenience, simplicity, accuracy, and ability to analyze many constituents at the same time (Stark et al., 1986). In

addition, the method does not require highly developed operator skills or a large amount of sample materials, and the analysis is nonrestrictive (Wetzel, 1983). Basically, the NIRA method assumes that a concentration of a given constituent is a linear combination of several absorption features. The method is empirical, and no physical or chemical factors or assumptions are really needed.

The spectral features of soil materials in the NIR (1–2.5  $\mu\text{m}$ ) spectral region are associated with vibration modes of functional groups. These modes are combinations and overtones of the fundamentals in the IR spectral region. Soil minerals have distinct spectral fingerprints in the NIR region because of relatively strong absorption of the overtones and the combination modes of the OH, SO<sub>4</sub>, and CO<sub>3</sub> groups (Hunt and Salisbury, 1970). Organic matter also has distinct fingerprints that relate to several functional groups (e.g., carboxyl, hydroxyl, and amine) in the organic compounds (Bowers and Hanks, 1965; Al-Abbas et al., 1972).

Both minerals and organic matter in soils were studied using the NIRA approach. Ben-Dor and Banin (1990a) pointed out that the concentration of carbonates could be predicted by NIRA in arid soils if the soils were preheated to 600°C for 8 h prior to the spectroscopic measurements. They concluded that by this pretreatment, the strong absorption features of the OH groups (both in the organic matter and in the clay minerals) are removed and the CO<sub>3</sub> groups are enhanced. Ben-Dor and Banin (1990b) used the NIRA strategy for chemical determination of clay minerals. They found that total Fe, Al, Mg, and silica can be predicted from a group

E. Ben-Dor, Dep. of Geography, Tel-Aviv Univ., P.O. Box 39040, Ramat Aviv, Tel-Aviv 69978 Israel; and A. Banin, Dep. of Soil and Water Sciences, Faculty of Agriculture, Hebrew Univ. of Jerusalem, P.O. Box 12, Rehovot 76100, Israel. Received 30 Nov. 1993. \*Corresponding author (bendor@ccsg.tau.ac.il).

**Abbreviations:** NIRA, near-infrared analysis; NIR, near infrared; SSA, specific surface area; CEC, cation-exchange capacity; HIGF, hygroscopic moisture; OM, organic matter; MRA, multiple regression analysis; SEC, standard error of calibration; SEP, standard error of performance; AE, analytical error; CPV, coefficient of prediction variation.

## Near-Infrared Analysis as a Rapid Method to Simultaneously Evaluate Several Soil Properties

E. Ben-Dor\* and A. Banin

### ABSTRACT

The near infrared analysis (NIRA) approach was studied to examine its capability for predicting spectral feature soil properties from the reflectance curves in the near infrared (NIR) region (1–2.5  $\mu\text{m}$ ) of arid and semiarid soils. High-resolution diffuse reflectance spectra (3113 spectral points) in the NIR region were recorded for 91 soil samples from Israel. Six soil properties (clay content, specific surface area, cation-exchange capacity, hygroscopic moisture, carbonate content, and organic matter content) were measured by routine methods employed in soil laboratories. An empirical model to predict each property from its spectral signature was developed by adapting the NIRA technique. Several data manipulations were used to obtain optimum performance. The optimum performance of all properties was found to be between 25 and 63 spectral points. Strong support for the NIRA capability was provided by its ability to examine most of the spectral assignments. A slight bias was observed for the prediction of both organic matter and hygroscopic moisture, suggesting that more attention in the prediction of these constituents is required. It was concluded that NIRA is a promising method for rapid and nonrestrictive analysis of soil materials, and further study of the synergism between NIRA and soil materials is recommended.

**N**EAR-INFRARED ANALYSIS is a laboratory approach that analyzes the diffuse reflectance radiation with regard to a material's chemistry (Stark et al., 1986). This methodology was developed 25 yr ago for rapid analysis of moisture in grains (Ben-Gera and Norris, 1968). Today, NIRA is well accepted and widely used in many other disciplines (Davies and Grant, 1987; Norris, 1988). This method is known for its rapidity, convenience, simplicity, accuracy, and ability to analyze many constituents at the same time (Stark et al., 1986). In

addition, the method does not require highly developed operator skills or a large amount of sample materials, and the analysis is nonrestrictive (Wetzel, 1983). Basically, the NIRA method assumes that a concentration of a given constituent is a linear combination of several absorption features. The method is empirical, and no physical or chemical factors or assumptions are really needed.

The spectral features of soil materials in the NIR (1–2.5  $\mu\text{m}$ ) spectral region are associated with vibration modes of functional groups. These modes are combinations and overtones of the fundamentals in the IR spectral region. Soil minerals have distinct spectral fingerprints in the NIR region because of relatively strong absorption of the overtones and the combination modes of the OH,  $\text{SO}_4$ , and  $\text{CO}_3$  groups (Hunt and Salisbury, 1970). Organic matter also has distinct fingerprints that relate to several functional groups (e.g., carboxyl, hydroxyl, and amine) in the organic compounds (Bowers and Hanks, 1965; Al-Abbass et al., 1972).

Both minerals and organic matter in soils were studied using the NIRA approach. Ben-Dor and Banin (1990a) pointed out that the concentration of carbonates could be predicted by NIRA in arid soils if the soils were preheated to 600°C for 8 h prior to the spectroscopic measurements. They concluded that by this pretreatment, the strong absorption features of the OH groups (both in the organic matter and in the clay minerals) are removed and the  $\text{CO}_3$  groups are enhanced. Ben-Dor and Banin (1990b) used the NIRA strategy for chemical determination of clay minerals. They found that total Fe, Al, Mg, and silica can be predicted from a group

E. Ben-Dor, Dep. of Geography, Tel-Aviv Univ., P.O. Box 39040, Ramat Aviv, Tel-Aviv 69978 Israel; and A. Banin, Dep. of Soil and Water Sciences, Faculty of Agriculture, Hebrew Univ. of Jerusalem, P.O. Box 12, Rehovot 76100, Israel. Received 30 Nov. 1993. \*Corresponding author (bendor@ccsg.tau.ac.il).

**Abbreviations:** NIRA, near-infrared analysis; NIR, near infrared; SSA, specific surface area; CEC, cation-exchange capacity; HIGF, hygroscopic moisture; OM, organic matter; MRA, multiple regression analysis; SEC, standard error of calibration; SEP, standard error of performance; AE, analytical error; CPV, coefficient of prediction variation.

### The spectral reflectance

The original spectrum consisting of 3113 spectral bands was compressed computationally to a new spectrum containing a smaller number of bands. The compressing procedure was carried out as follows. The original spectrum (of each sample) was divided into  $G$  bands ( $G = 25, 30, 63, 310$ ). Each of the new band spectrum consisted of 3113/ $G$  original bands and their paired reflectance values. The wavelength-reflectance value of each new band was obtained by using a simple average calculation across the original reflectance-wavelength bands.

The NIRA methodology requires two st

This procedure consisted of a subgroup that was selected directly from the 91 soil samples. A simple linear regression procedure between the concentration of a given constituent and the spectral response (reflectance  $[R]$ , absorption  $[a] = \log(1/R)$  or their derivatives  $[R, a', a'']$  of every band was run. Twenty-five bands (six for the  $G = 25$  compressed bands) and their corresponding wavelengths providing the highest coefficient of correlation were further selected for examination under a MRA procedure. This procedure was run for each of the spectral parameters ( $R$ ,  $a'$ , or their derivatives) and for each of the MRA stages. The result, a multiterm equation was used to determine the concentration of unknown validated samples in the validation stage.

Unknown samples that were not used in the calibration process were used to validate the calibration equation produced in Stage 1. For that purpose, the spectra of all the validation samples were mathematically manipulated exactly as were

any piecemeal...

Although the NIR method was found to be a very promising tool in many other disciplines as well, and termed a "giant" method (Davies, 1987), it has not received much attention in the soil science field. The purpose of this study was to examine the capability of the NIR technique to simultaneously measure various spectral feature properties of soil from the arid and semiarid regions of Israel. The properties were selected because of their importance in soil classification and estimating the agricultural productivity of soils. The properties selected were clay content, SSA, CEC, HIGF, OM, and  $\text{CaCO}_3$  content.

## Soil Sampling

Clay content was obtained by the hydrometer method (Gee and Bauder, 1986), SSA by ethylene glycol monoethyl ether adsorption (Rathner-Zohar et al., 1983), CEC by quantitative saturation and subsequent displacement of Na cations onto the soil surface according to Thomas (1986), HGF by gravimetric measurements before and after heating to 105°C for 24 h according to Gardner (1986), OM by a loss-on-ignition method according to Ben-Dor and Banin (1989), and CaCO<sub>3</sub> by the gasometric method according to Nelson (1986). All determinations (except of CLAY) consisted of three or more replications to assure accuracy.

carried out using the fine soil samples ( $<0.355$  mm).

The MRA procedure selects a best multiple correlation equation between a constituent concentration  $C_p$  in a given sample and the selected spectral parameters as follows:

The MRA procedure selects a best

$$SEP = \sqrt{\frac{\sum_{d=1}^N (C_d^m - C_d^p)^2}{N_d - 1}} \quad [3]$$

Reflectance spectra of 11 soil samples that represent most of the soil groups within the examined population are given in Fig. 1. It is apparent that all of the spectra consist of spectral information that can be easily attributed to some of the soil compounds. The most likely-

most of the soil groups within the examined

SSA, and HIGF content (e.g., high: A-4, H-2, P-2; low: C-2, W-1, and S-5) as well as between the absorption intensity at  $\approx 2.3$   $\mu\text{m}$  and the  $\text{CaCO}_3$  content (e.g., high: C-2, W-1, O-2, S-5; low: E-1, A-4, W-1). In some soils, some absorption features that may be assigned to Fe oxides can be observed at  $\approx 1.05$   $\mu\text{m}$  (e.g., B-1, K-2, P-2, and S-5; Hunt and Salisbury, 1970). The above discussion is limited only to the 11 representative soil samples. For a broader point of view and for more adequate conclusions, the whole soil population must be systematically examined (chemically and spectrally). Judging from the chemical range of the whole population.

Table 2 presents a matrix of the linear coefficient correlation between all the examined soil properties. Three essential groups can be observed: Group I, which consists of clay content, SSA, CEC, and HIGF (all closely correlated among themselves with  $r > 0.643$ ) Group II, which consists of OM only (moderately correlated with Group I;  $0.412 < r > 0.665$ ); and Group III, which consists of  $\text{CaCO}_3$  content (does not correlate with any of the other properties). Properties in Group I are all related to the clay content and its SSA. Bantin and Amiel (1970) pointed out that in Israeli soils, SSA is for the most part contributed by smectitic clay minerals and soil clay minerals (2.2  $\mu\text{m}$ ).

studied here.

of smectite minerals by using the NIRA strategy without any pretreatment.

Organic matter and its compounds (C and N) have also been predicted by using the NIRA procedure in soils. Dalal and Henry (1986) pointed out that organic C, total N, and soil moisture can be predicted for Australian soils using the NIRA technique. They concluded that the NIRA technique can be used as a routine testing method if further consideration is given to the instrumentation and the data handling. Morra et al. (1991) studied the NIRA capability to predict organic C and N in a relatively homogenous soil group (12 soils with their chemically treated subsets). They concluded that the method is very promising for the prediction of these constituents and feel, also, that the method is very promising for soil applications. Recently, Ben-Dor and Banin (1994, unpublished data) were able to show that the NIRA methodology can be used to predict some spectral featureless constituents in soils via an intercorrelation with soil properties that are spectrally active in the NIR region.

Although the NIRA method was found to be a very promising tool in many other disciplines as well, and termed a "giant" method (Davies, 1987), it has not received much attention in the soil science field. The purpose of this study was to examine the capability of the NIRA technique to simultaneously measure various spectral feature properties of soil from the arid and semiarid regions of Israel. The properties were selected because of their importance in soil classification and estimating the agricultural productivity of soils. The properties selected were clay content, SSA, CEC, HIGF, OM, and CaCO<sub>3</sub> content.

## MATERIALS AND METHODS

### Soil Sampling

Ninety-one soil samples that represent 12 groups of Israeli soils were selected to cover the arid and semiarid climate zone of Israel. The samples ( $\approx 5$ –10 kg) were collected mostly from the A<sub>0</sub> horizon of the soils (upper 5 cm) within an area of  $\approx 1$  m<sup>2</sup>. The soils were air dried and gently crushed to pass through a <2-mm sieve. Subsamples were further ground (quantitatively) by agate mortar to pass through a <0.355-mm sieve. All chemical and spectroscopic determinations were carried out using the fine soil samples (<0.355 mm).

### Chemical Analyses

Clay content was obtained by the hydrometer method (Gee and Bauder, 1986), SSA by ethylene glycol monoethyl ether adsorption (Rathner-Zohar et al., 1983), CEC by quantitative saturation and subsequent displacement of Na cations onto the soil surface according to Thomas (1986), HIGF by gravimetric measurements before and after heating to 105°C for 24 h according to Gardner (1986), OM by a loss-on-ignition method according to Ben-Dor and Banin (1989), and CaCO<sub>3</sub> by the gasometric method according to Nelson (1986). All determinations (except of CLAY) consisted of three or more replications to assure accuracy.

## Spectral Measurements and Compression

### Spectral Measurements

The spectral reflectances of the soil powder were recorded by an Alpha Centauri FTIR spectrophotometer (Mattson Co., Madison, WI) optimized to the NIR (1–2.5  $\mu$ m) region. It has a high optical resolution of 1 cm<sup>-1</sup> throughout the spectral region and has the capability of producing up to 6000 spectral bands. For the current measurements, we used a constant spectral resolution value of 8 cm<sup>-1</sup>, which produced 3113 spectral points across the 1- to 2.5- $\mu$ m spectral region. In the wavelength mode, this resulted in a variation of the spectral resolution between 0.8 nm for the lower wavelength (1  $\mu$ m) and 1.8 nm for the higher wavelength (2.5  $\mu$ m). Diffuse reflectance was measured by a praying mantis bi-directional diffuse reflectance attachment (DEA-dimensional model, Harrick Co., Ossining, NY). The soil powders (<0.355 mm) were packed into a stainless steel sample holder, using a standard sample preparation procedure. The soil reflectance was reported relative to a BaSO<sub>4</sub> standard spectrum, which was recorded under similar spectroscopic and geometric conditions.

### Spectral Compression

The original spectrum consisting of 3113 spectral bands was compressed computationally to a new spectrum containing a smaller number of bands. The compressing procedure was carried out as follows. The original spectrum (of each sample) was divided into  $G$  bands ( $G = 25, 30, 63, 310$ ). Each of the new band spectrum consisted of 3113/ $G$  original bands and their paired reflectance values. The wavelength-reflectance value of each new band was obtained by using a simple average calculation across the original reflectance-wavelength bands.

### Spectral Analysis

The NIRA methodology requires two stages: (i) the calibration stage where the prediction equation is developed and (ii) the validation stage, where the previous stage is validated. Because NIRA is basically an empirical method, the validation stage is extremely critical. The complete NIRA procedure requires both calibration and validation stages to analyze unknown samples.

### Calibration Stage

This procedure consisted of a subgroup that was selected directly from the 91 soil samples. A simple linear regression procedure between the concentration of a given constituent and the spectral response (reflectance  $[R]$ , absorption  $[A = \log(1/R)]$  or their derivatives  $[R', A', A'']$  of every band was run. Twenty-five bands (six for the  $G = 25$  compressed bands) and their corresponding wavelengths providing the highest coefficient of correlation were further selected for examination under a MRA procedure. This procedure was run for each of the spectral parameters ( $R$ ,  $A$ , or their derivatives) and for each of the MRA stages. The result, a multiterm equation, was used to determine the concentration of unknown validated samples in the validation stage.

### Validation Stage

Unknown samples that were not used in the calibration process were used to validate the calibration equation produced in Stage 1. For that purpose, the spectra of all the validation samples were mathematically manipulated exactly as were

the spectra that provided the prediction equation. Then the concentration of each constituent was calculated using the corresponding prediction equation produced in Stage 1. The calculated results (predicted) were compared with the chemical results (measured). This stage was run for each constituent until a relatively high prediction performance was obtained. At that point, the corresponding calibration equation was declared the preferred equation for predicting samples from the calibration and validation stages.

### Statistical Analysis

The MRA procedure selects a best multiple correlation equation between a constituent concentration  $C_p$  in a given sample and the selected spectral parameters as follows:

$$C_p = b_0 + b_1L_1 + b_2L_2 + b_3L_3 + \dots + b_nL_n \quad [1]$$

where  $b_0$  is an intercept,  $b_1, b_2, \dots, b_n$  are weighting factors for the spectra readings in various selected wavelengths 1 to  $n$ ,  $L_1, L_2, \dots, L_n$  are the values of the spectral parameters at wavelengths 1, 2,  $\dots, \lambda$  and could be either reflectance data ( $R$ ), absorption data ( $A$ ), or their derivatives (first  $R'$ ,  $A'$ , or second  $A''$ ). Predicted values of the constituent concentrations were calculated using Eq. [1] and the SEC was then calculated according to Davies and Grant (1987) as follows:

$$SEC = \sqrt{\frac{\sum_{i=1}^{N_c} (C_m - C_p)^2}{N_c - n - 1}} \quad [2]$$

where  $C_m$  is the value measured by the chemical analysis,  $C_p$  is the predicted value on the basis of the spectral analysis,  $N_c$  is the number of the samples in the calibration stage, and  $n$  is the number of terms in the prediction equation.

The calibration Eq. [1] was used to predict the constituent concentration in another set of soil samples not involved in the calibration stage (validation stage). The SEP was then calculated as follows:

$$SEP = \sqrt{\frac{\sum_{i=1}^{N_p} (C_m - C_p)^2}{N_p - 1}} \quad [3]$$

where  $N_p$  is the total number of tested samples in this stage.

A bias test, adopted from Miller and Miller (1988), was used to examine the regression line ( $C_m = aC_p + b$ ) of the validation results. It uses three null hypotheses to reject or accept the regression line as a 1:1 line. The assumptions of the null hypotheses are (i) the slope of the regression line ( $a$ ) is unity, (ii) the intercept of the regression line ( $b$ ) is zero, and (iii) the regression coefficient ( $r^2$ ) is unity.

To allow better comparison with the SEP values, the AE of the chemical analysis was calculated as follows:

$$AE = \sqrt{\frac{\sum_{i=1}^{91} \sum_{j=1}^{N_i} (X_{ij} - C_{mi})^2}{N_i - 1}} \quad [4]$$

Where  $X_{ij}$  is the direct chemical measurement of Sample  $i$  in Replication  $j$ ,  $C_{mi}$  is the mean of Sample  $i$ , and  $N_i$  is the total number of the entire replications for a given property.

## RESULTS AND DISCUSSION

Reflectance spectra of 11 soil samples that represent most of the soil groups within the examined population are given in Fig. 1. It is apparent that all of the spectra consist of spectral information that can be easily attributed to some of the soil compounds. The most likely

components in the NIR region to have strong absorption are the OH groups in both the adsorbed water (at  $\approx 1.4$  and 1.9  $\mu$ m) and the crystal lattice water (at  $\approx 1.45$  and 2.2  $\mu$ m; Hunt and Salisbury, 1970). Also, the CO<sub>3</sub> in calcite mineral (and hence in calcite-enriched soils) is very active in the NIR region (major peak at  $\approx 2.33$   $\mu$ m; Gaffey, 1986). Table 1 provides some of the analytical data from the 11 soils for each of the six properties. A rough correlation occurs between the absorption intensity at  $\approx 1.4, 1.9$ , and 2.2  $\mu$ m and the Clay content, SSA, and HIGF content (e.g., high: A-4, H-2, P-2; low: C-2, W-1, and S-5) as well as between the absorption intensity at  $\approx 2.3$   $\mu$ m and the CaCO<sub>3</sub> content (e.g., high: C-2, W-1, O-2, S-5; low: E-1, A-4, W-1). In some soils, some absorption features that may be assigned to Fe oxides can be observed at  $\approx 1.05$   $\mu$ m (e.g., B-1, K-2, P-2, and S-5; Hunt and Salisbury, 1970). The above discussion is limited only to the 11 representative soil samples and does not necessarily apply to other soil samples. For a broader point of view and for more adequate conclusions, the whole soil population must be systematically examined (chemically and spectrally). Judging from the chemical range of the whole population (Table 1), it can be concluded that the 11 soil samples are a good representative of the entire population. However, we further selected an additional soil sample (B-7) that accurately represents the entire population (Table 1) on which to present the spectral compression process.

Figure 2 presents the compressed spectra of Sample B-7. It is apparent that some spectral features are weakened by the compression process, decreasing the number of spectral bands from 3113 to 25. Nonetheless, three essential absorption bands (which appear in the 11 representative soil spectra as well) are evident throughout all the compressed spectra (at  $\approx 1.4, 1.9$ , and 2.2  $\mu$ m). These features are attributed to the OH groups of soil moisture (1.4 and 1.9  $\mu$ m) and to the crystal lattice in soil clay minerals (2.2  $\mu$ m).

Table 2 presents a matrix of the linear coefficient correlation between all the examined soil properties. Three essential groups can be observed: Group I, which consists of clay content, SSA, CEC, and HIGF (all closely correlated among themselves with  $r > 0.643$ ); Group II, which consists of OM only (moderately correlates with Group I;  $0.412 < r < 0.665$ ); and Group III, which consists of CaCO<sub>3</sub> content (does not correlate with any of the other properties). Properties in Group I are all related to the clay content and its SSA. Banin and Amiel (1970) pointed out that in Israeli soils, SSA is for the most part contributed by smectite clay minerals, which constitute some 70% of the clay portion of the soil and which affect the adsorption properties (CEC and HIGF). Organic matter (Group II) is also moderately related to the adsorption properties of soil (SSA, CEC, and HIGF), either directly or indirectly throughout the correlation of clay content. Carbonates (Group III) are not correlated with any of the other five soil properties studied here.

An important prerequisite for successful application of the NIRA approach for analysis of a given constituent is that it has spectral absorption features in the NIR region

Table 1. The analytical values (measured) of 11 representative soils within the whole population examined. Also given are the chemical values of a representative soil sample (B-7) that represents these 11 soils and the chemical range of each property within the whole population.

Soil sample	Clay content	SSA	CEC	HIGF	OM	CaCO <sub>3</sub> content
E-1	13.9	48.5	8.4	1.12	1.27	1.8
C-2	10.1	80.0	12.89	2.70	1.02	40.3
B-1	39.4	196.8	35.61	4.68	4.91	41.9
A-4	61.2	338.2	60.3	11.63	5.65	8.4
H-2	61.2	425.3	51.4	10.14	5.34	18.5
J-1	50.6	189.4	31.4	11.12	2.42	38.9
W-1	12.7	81.0	10.9	1.34	0.87	16.1
O-2	24.6	176.3	23.8	3.69	2.30	74.3
K-2	38.9	116.2	14.8	3.19	2.87	58.9
P-2	59.9	228.7	27.5	4.49	1.19	55.2
S-5	7.6	38.5	7.2	1.17	1.05	18.2
B-7	41.1-65.3	10.9-425.3	1.4-62.1	0.20-11.63	0.09-13.23	0-74.3
Range†	45.91	306.2	51.43	6.0	3.46	12.78

† SSA = specific surface area; CEC = cation-exchange capacity; HIGF = hygroscopic moisture; OM = organic matter.  
† Chemical range of the whole soil population (n = 91).

† SSA = specific surface area; CBC = cation-exchange capacity; HIGF = hygroscopic moisture; OM = organic matter.  
‡ Chemical range of the whole soil population ( $n = 91$ ).

Absorption bands in the NIR region for phyllosilicates are assigned to OH in free water at 1.4- and 1.9- $\mu\text{m}$  and 2.1- to 2.35- $\mu\text{m}$  bands (Hunt and Salisbury, 1970) bands and to OH in crystal lattice at 1.3- to 1.45- $\mu\text{m}$

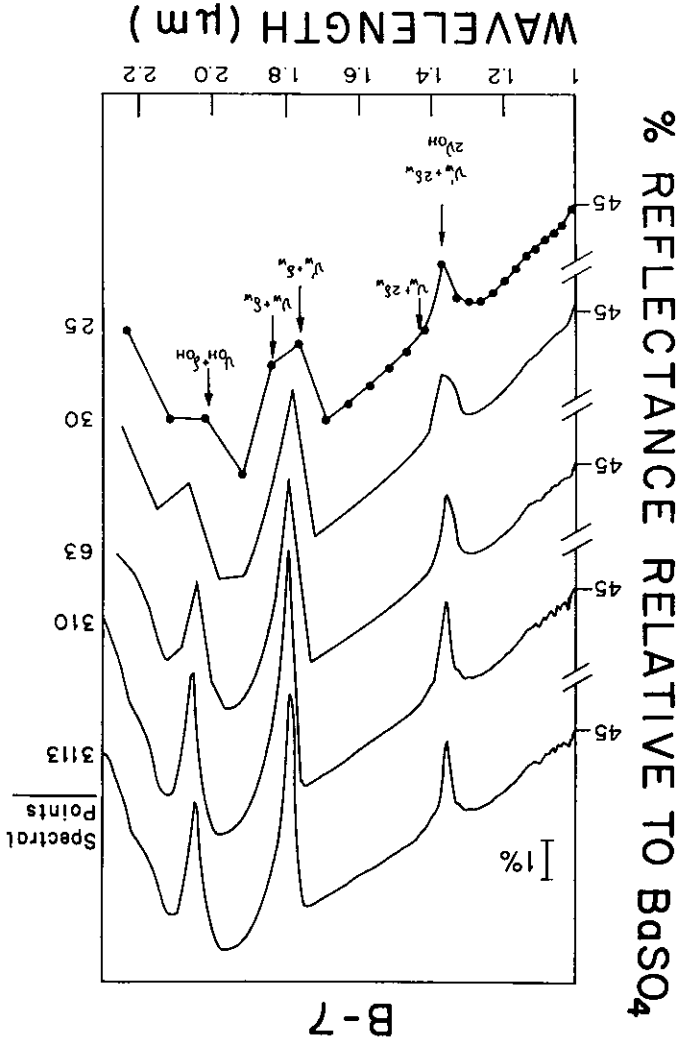


Fig. 2. The effect of compressing the original reflectance spectrum of a representative soil sample (B-7). The number of spectral bands,  $G$ , was varied by mathematical manipulations of the originally measured spectrum ( $G = 3113$ ).

Table 2. Linear coefficient of correlation ( $r$ ) between the properties of the study soils. The correlation was obtained using the chemical values ( $n = 91$ ).†

reflectance spectrum. In Group III, carbonates have significant absorbance features in the NIR, corresponding to CO<sub>3</sub> radicals at

throughout the entire NIR region. However, it is difficult to specifically assign an absorption mode to a particular wavelength because many combinations and overtone modes can occur. This situation occurs because of the chemical complexity of organic matter in soils. They contain many and varied functional groups (Bowers and Hanks, 1965). Furthermore, the organic matter content of soils below 2% does not permit simple spectral detection of the absorption features related to the functional groups of organic matter (Baumgardner et al., 1970). This is because of a strong mask provided by other soil components (e.g., hygroscopic and crystal lattice water and carbonates) and because of low absorption rates of the OM functional groups in the soil mixture. Nevertheless, by using the NIRA strategy, Dalal and Henry (1986) were able to predict the OM content despite a low chemical range (0–2.6% of C). This suggests that a sophisticated analytical method such as NIRA can be used for extracting hidden information from the diffuse soil re-

bands and to OH in crystal lattice at 1.3- to 1.45- $\mu\text{m}$  and 2.1- to 2.35- $\mu\text{m}$  bands (Hunt and Salisbury, 1970). These bands have been found to be quantitatively correlated with the chemical composition of smectite minerals using the NIRA strategy (Ben-Dor and Banin, 1990b). In Group II, organic matter is spectroscopically active

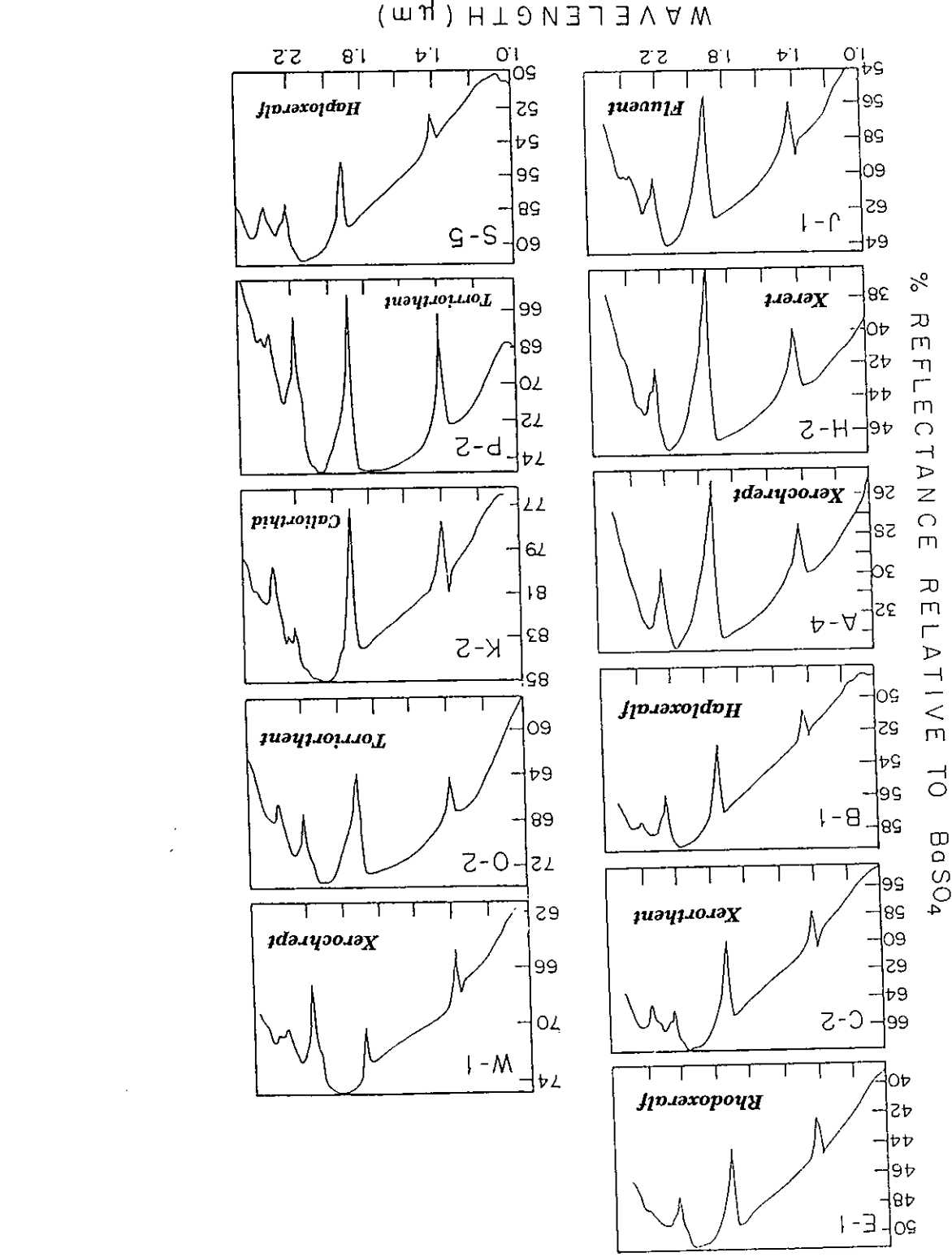


Fig. 1. Reflectance spectra of 11 representative soil samples. Each of the soils corresponds to a different soil group and is classified according to the USDA classification system.

(Murray, 1988). All three groups satisfy this condition, as discussed below.

of both the crystal lattice and water adsorbed on the mineral surface. The amount of OH is rather specific to certain phyllosilicates. The OH groups are spectrally active throughout the NIR region because of several combinations and overtones of the OH fundamentals.





Table 3. Categorizing of the soil properties into three groups. Also given is the number of samples in the calibration and validation sets and their corresponding chemical range.

Property†	Group	N‡	Calibration range	Validation range
Clay, %	I	35	4.1-61.2	6.3-65.3
SSA, m <sup>2</sup> g <sup>-1</sup>	I	35	10.9-425.3	10.8-369.0
CEC, cmol. kg <sup>-1</sup>	I	35	1.4-60.3	1.3-62.1
HIGF, %	I	35	0.20-11.63	0.39-9.21
OM, %	II	39	0.09-13.23	0.09-10.70
CaCO <sub>3</sub> , %	III	22	0-74.3	0-68.3

† SSA = specific surface area; CEC = cation-exchange capacity; HIGF = hygroscopic moisture; OM = organic matter.  
‡ N<sub>c</sub> = Number of samples in the calibration stage.  
§ N<sub>v</sub> = Number of samples in the validation stage.

1.6, 1.7, 2.2, 2.3, and 2.5  $\mu\text{m}$  (Gaffey, 1986). Ben-Dor and Banin (1990a) pointed out that if masking OHs (from lattice and hygroscopic mineral sources) are diminished by preheating the soil to 600°C, the CO<sub>2</sub> groups of calcite minerals are enhanced and the NIRA capability of prediction of the CaCO<sub>3</sub> content of soils is improved.

The NIRA technique requires the selection of a representative calibration set that is chemically matched with the validation set and with a large number of samples (Norris and Williams, 1984). Because we identified three different groups of properties, three independent testing groups were employed for the NIRA examination. Each testing group contained different samples (both types and numbers) for both the calibration and the validation sets. Table 3 gives the chemical range of both the calibration and the validation sets and the number of samples in each set. From Table 3, it can be seen that a positive consistency occurs between the two sets in all of the examined properties.

Basically, the NIRA approach is empirical, allows (actually requires) many combinations of samples and data manipulations to obtain optimal prediction. No a priori information is available regarding the best manipulation (Murray, 1988). To keep the number of combinations to a manageable set, we kept constant the number and identity of samples in both calibration and validation sets for each soil property group and changed only the data manipulation procedures (R, R', A, A', and A'') in both the expanded and the compressed spectra. The criteria of lowest SEP and highest  $r^2$  of the validation set were selected to indicate the optimal data manipulation for obtaining the best analytical performance. Also, the bias of the validation results were inspected to better estimate the prediction power of the selected routine (manipulation and wavelength assignments).

Statistical parameters from the calibration, validation, and analytical stages of the optimal combinations are given in Table 4. It was found that for all soil properties, the high-resolution spectra (3113 spectral bands) do not necessarily yield the best validation results. Moreover, with one property (HIGF), optimal performance was obtained within the low spectral resolution (25 spectral bands) using A'' manipulation. In the other properties of Group I (clay content, CEC, and SSA), optimal performance was obtained using 63 spectral bands and A' manipulation. The OM property shows the best performance using 30 spectral bands and R' manipulation, whereas for the CaCO<sub>3</sub> property, no significant difference was obtained between any given resolution. The highest performance of the CaCO<sub>3</sub> was obtained by using the R' manipulation and 25 to 3113 spectral bands. These findings suggest that spectral resolution is not necessarily a limitation for predicting these soil properties. Furthermore, in some cases, low spectral resolution is strongly recommended for better utilization of the soil reflectance spectra. Averaging several raw (high-resolution) spectral bands into a new (low-resolution) spectral band may better enhance spectral signals that relate to the examined constituent.

Table 5 shows the best prediction equations determined for each of the properties. For clay content, SSA, and CEC, a similar mathematical manipulation (63 bands and A' spectra) and selected wavelengths resulted in the optimal performance. The wavelengths are associated with OH in both crystal lattice and surface water. These observations support the group selection based on the chemical correlation discussed above. The best prediction for HIGF (part of Group I) occurred with 25 spectral bands and A' manipulation. Moreover, the wavelengths selected by NIRA for HIGF are quite different from the wavelengths in the equations for clay content, SSA, and CEC. One possible explanation for this is that the high extinction coefficient of free water molecules (HIGF) leads to oversaturation by absorption of water. The shape of a highly compressed spectrum loses its natural peaks and hence apparently provides peak undersaturation, which is closely related to water content.

In Group II (OM), the spectral assignments might correspond to both humic substances (and their compo-

Table 4. Statistical parameters of the calibration, validation, and analytical stages.†

Property‡	Spectra			Calibration			Validation			Analytical		
	N <sub>c</sub>	R <sup>2</sup>	SEP	N <sub>v</sub>	r <sup>2</sup>	SEP	CPV	N <sub>a</sub>	AE	CAV		
Clay	35	0.76	8.6	35	0.56	10.3	0.174	91	4.8			0.065
SSA	35	0.84	37.7	35	0.70	50.2	0.140	91	4.0			0.010
CEC	35	0.82	6.72	35	0.64	8.46	0.139	91	1.2			0.020
HIGF	25	0.58	1.9	35	0.62	1.55	0.175	91	0.22			0.019
OM	30	0.69	1.63	52	0.55	1.34	0.126	91	0.30			0.023
CaCO <sub>3</sub>	22	0.70	12.9	47	0.69	11.6	0.169	91	1.3			0.017

† N<sub>c</sub> = Number of samples in the calibration stage; R<sup>2</sup> = Multiple coefficient of correlation; SEP = Standard error of calibration; N<sub>v</sub> = Number of samples in the validation stage; r<sup>2</sup> = Linear coefficient of correlation; SEP = Standard error of performance; N<sub>a</sub> = Number of samples in the analytical stage; CPV = Coefficient of prediction variation (SEP/Max-Min); AE = Analytical error; CAV = Coefficient of analytical variation (AE/Max-Min).  
‡ SSA = specific surface area; CEC = cation-exchange capacity; HIGF = hygroscopic moisture; OM = organic matter.  
§ Estimated value.

nents) and surface properties (as obtained from the chemical correlation in Table 2).

In Group III (CaCO<sub>3</sub>), the wavelengths entered into the calibration are not those characteristic of carbonate bands. The two selected wavelengths (1.6473 and 1.4825  $\mu\text{m}$ ) are typically located in nonabsorption regions of the spectrum. Furthermore, the optimal prediction was obtained using the reflectance (R), and analyses of all the compressed spectra resulted in practically the same SEP. This leads us to the conclusion that the prediction of CaCO<sub>3</sub> in this case is based largely on the position within the spectrum continuum (albedo) rather than in any specific absorption band. Large amounts of CaCO<sub>3</sub> generally cause soils to be more reflective in the NIR (as in the visible spectrum).

Figure 3 shows plots of C<sub>p</sub> (predicted) vs. C<sub>m</sub> (measured) at the calibration and validation stages for each of the six examined properties. For all properties, most of the points fell in the vicinity of the 1:1 line with a few outliers. Murray (1988) noted that removing outliers from the calibration-validation set may increase prediction accuracy. However, because the prediction accuracy for all six properties was within reasonable limits of detection for rapid laboratory measurements (CPV = 0.126-0.175, Table 4), no recalculations were done to compensate for reducing the sample sets by omitting outliers. An examination to evaluate possible bias within the validation sets was performed by using the Miller and Miller (1988) test. This test was originally designed to examine a new analytical method as it relates to an old one. It statistically examines the difference between

the regression and ideal 1:1 lines. Table 6 presents the regression line (C<sub>p</sub> = aC<sub>m</sub> + b) and the parameters a, b, and r<sup>2</sup> (taken from the validation stage) along with their statistical significance. The r<sup>2</sup> values for all properties are not significantly different from 1 at the 0.05 significance level. However, based on the test of a and b parameters only clay content, CEC, and CaCO<sub>3</sub> were found to significantly match the 1:1 line at the 0.05 significance level (clay content, CEC, and CaCO<sub>3</sub> even matched at the 0.01 significance level) and therefore held no bias. The regression line parameter a for OM and HIGF were significantly different from the ideal 1:1 line and, therefore, did hold some bias. These results may suggest that further research is needed to improve the prediction of OM and HIGF properties. The NIRA prediction of organic C and N contents of organic matter (0-10% C) was found to be possible only if a homogeneous soil group was used (Morra et al., 1991). Morra et al. (1991) pointed out that the necessity of using a well-defined closed sample population for the NIRA procedure for any constituent is an inherent limitation of the procedure. We believe that the soil population in our study represents two essential groups with regard to the OM decomposition stages: (i) a group of soils that consists of 0 to 4% OM, representing greatly decomposed OM soils, and (ii) a group of soils that consists of 4 to 14% OM/OM, representing less decomposed OM soils. In the first group, the functional properties of OM are more likely similar to humic substances, whereas in the second group the functional properties are more likely similar to fresh litter and vegetation.

Table 5. The optimal equations for predicting each of the soil constituents. Also given are the suggested wavelength assignments.

Property†	Mathematics	Constant (b <sub>0</sub> )	Weighting coefficient (b <sub>1</sub> )	Wavelength (λ)	Suggested assignment
Clay	A'	63	59.7	121.9	U <sub>mg-OH</sub> + 2δ <sub>Al-OH</sub>
				83.9	U <sub>Al-OH</sub> + 2δ <sub>Al-OH</sub>
				152.7	U <sub>Al-OH</sub> + 2δ <sub>Al-OH</sub>
				35.8	U <sub>u</sub> + δ <sub>u</sub>
				9.3	U <sub>u</sub> + δ <sub>u</sub>
				1.8703	U <sub>u</sub> + δ <sub>u</sub>
				1.4720	U <sub>u</sub> + 2δ <sub>u</sub>
SSA	A'	63	268.2	457.0	U <sub>mg-OH</sub> + 2δ <sub>mg-OH</sub>
				2.3333	U <sub>mg-OH</sub> + 2δ <sub>mg-OH</sub>
				2.0973	U <sub>mg-OH</sub> + 2δ <sub>mg-OH</sub>
				1.9423	U <sub>u</sub> + δ <sub>u</sub>
				434.0	U <sub>u</sub> + δ <sub>u</sub>
				1.8700	U <sub>u</sub> + δ <sub>u</sub>
				1.4033	U <sub>u</sub> + 2δ <sub>u</sub>
				1.3560	U <sub>u</sub> + 2δ <sub>u</sub>
CEC	A'	63	73.4	208.0	U <sub>mg-OH</sub> + 2δ <sub>mg-OH</sub>
				2.3333	U <sub>mg-OH</sub> + 2δ <sub>mg-OH</sub>
				2.0973	U <sub>mg-OH</sub> + 2δ <sub>mg-OH</sub>
				87.4	U <sub>u</sub> + δ <sub>u</sub>
				1.4313	U <sub>u</sub> + δ <sub>u</sub>
				1.4119	U <sub>u</sub> + δ <sub>u</sub>
				94.8	U <sub>u</sub> + δ <sub>u</sub>
				1.3219	U <sub>u</sub> + δ <sub>u</sub>
				1.0641	U <sub>u</sub> + δ <sub>u</sub>
HIGF	A'	25	5.1	24.3	U <sub>u</sub>
				2.3620	U <sub>u</sub>
				2.1202	U <sub>u</sub> + 2δ <sub>Al-OH</sub>
				2.3883	U <sub>u</sub> + 2δ <sub>Al-OH</sub>
				2.0165	U <sub>u</sub> + 2δ <sub>Al-OH</sub>
				1.9410	U <sub>u</sub> + 2δ <sub>Al-OH</sub>
				1.5849	U <sub>u</sub> + 2δ <sub>Al-OH</sub>
				1.4122	U <sub>u</sub> + 2δ <sub>Al-OH</sub>
				0.20	U <sub>u</sub> + 2δ <sub>Al-OH</sub>
				0.07	U <sub>u</sub> + 2δ <sub>Al-OH</sub>
CaCO <sub>3</sub>	R	25-3113	-74.4	12.0	No assignments involved
				1.6473	No assignments involved
				1.4825	No assignments involved

† SSA = specific surface area; CEC = cation-exchange capacity; HIGF = hygroscopic moisture; OM = organic matter.



Table 3. Categorizing of the soil properties into three groups. Also given is the number of samples in the calibration and validation sets and their corresponding chemical range.

Property†	Group	N <sub>c</sub> ‡	N <sub>v</sub> §	Calibration range	Validation range
Clay, %	I	35	56	4.1–61.2	6.3–65.3
SSA, m <sup>2</sup> g <sup>-1</sup>	I	35	56	10.9–425.3	10.8–369.0
CEC, cmol <sub>c</sub> kg <sup>-1</sup>	I	35	56	1.4–60.3	1.3–62.1
HIGF, %	I	35	56	0.20–11.63	0.39–9.21
OM, %	II	39	52	0.09–13.23	0.09–10.70
CaCO <sub>3</sub> , %	III	22	47	0–74.3	0–68.3

† SSA = specific surface area; CEC = cation-exchange capacity; HIGF = hygroscopic moisture; OM = organic matter.

‡ N<sub>c</sub> = Number of samples in the calibration stage.

§ N<sub>v</sub> = Number of samples in the validation stage.

1.6, 1.7, 2.2, 2.3, and 2.5  $\mu\text{m}$  (Gaffey, 1986). Ben-Dor and Banin (1990a) pointed out that if masking OHs (from lattice and hygroscopic mineral sources) are diminished by preheating the soil to 600°C, the CO<sub>3</sub> groups of calcite minerals are enhanced and the NIRA capability of prediction of the CaCO<sub>3</sub> content of soils is improved.

The NIRA technique requires the selection of a representative calibration set that is chemically matched with the validation set and with a large number of samples (Norris and Williams, 1984). Because we identified three different groups of properties, three independent testing groups were employed for the NIRA examination. Each testing group contained different samples (both types and numbers) for both the calibration and the validation sets. Table 3 gives the chemical range of both the calibration and the validation sets and the number of samples in each set. From Table 3, it can be seen that a positive consistency occurs between the two sets in all of the examined properties.

Basically, the NIRA approach is empirical, allows (actually requires) many combinations of samples and data manipulations to obtain optimal prediction. No a priori information is available regarding the best manipulation (Murray, 1988). To keep the number of combinations to a manageable set, we kept constant the number and identity of samples in both calibration and validation sets for each soil property group and changed only the data manipulation procedures (*R*, *R'*, *A*, *A'*, and *A''*) in both the expanded and the compressed spectra. The criteria of lowest SEP and highest *r*<sup>2</sup> of the validation set were selected to indicate the optimal data manipulation for obtaining the best analytical performance. Also, the bias of the validation results were inspected to better

estimate the prediction power of the selected routine (manipulation and wavelength assignments).

Statistical parameters from the calibration, validation, and analytical stages of the optimal combinations are given in Table 4. It was found that for all soil properties, the high-resolution spectra (3113 spectral bands) do not necessarily yield the best validation results. Moreover, with one property (HIGF), optimal performance was obtained within the low spectral resolution (25 spectral bands) using *A''* manipulation. In the other properties of Group I (clay content, CEC, and SSA), optimal performance was obtained using 63 spectral bands and *A'* manipulation. The OM property shows the best performance using 30 spectral bands and *R'* manipulation, whereas for the CaCO<sub>3</sub> property, no significant difference was obtained between any given resolution. The highest performance of the CaCO<sub>3</sub> was obtained by using the *R* manipulation and 25 to 3113 spectral bands. These findings suggest that spectral resolution is not necessarily a limitation for predicting these soil properties. Furthermore, in some cases, low spectral resolution is strongly recommended for better utilization of the soil reflectance spectra. Averaging several raw (high-resolution) spectral bands into a new (low-resolution) spectral band may better enhance spectral signals that relate to the examined constituent.

Table 5 shows the best prediction equations determined for each of the properties. For clay content, SSA, and CEC, a similar mathematical manipulation (63 bands and *A'* spectra) and selected wavelengths resulted in the optimal performance. The wavelengths are associated with OH in both crystal lattice and surface water. These observations support the group selection based on the chemical correlation discussed above. The best prediction for HIGF (part of Group I) occurred with 25 spectral bands and *A'* manipulation. Moreover, the wavelengths selected by NIRA for HIGF are quite different from the wavelengths in the equations for clay content, SSA, and CEC. One possible explanation for this is that the high extinction coefficient of free water molecules (HIGF) leads to oversaturation by absorption of water. The shape of a highly compressed spectrum loses its natural peaks and hence apparently provides peak undersaturation, which is closely related to water content.

In Group II (OM), the spectral assignments might correspond to both humic substances (and their compo-

Table 5. The optimal equations for predicting each of the soil constituents. Also given are the suggested wavelength assignments.

Property†	Mathematics spectral points	Constant (b <sub>0</sub> )	Weighting coefficient (b <sub>1</sub> )	Wavelength ( $\lambda$ ) $\mu\text{m}$	Suggested assignment
Clay	<i>A'</i> 63	59.7	121.9	2.3333	$U_{\text{Mg-OH}} + 2\delta_{\text{Mg-OH}}$
			83.9	2.1857	$U_{\text{Al-OH}} + 2\delta_{\text{Al-OH}}$
			152.7	2.0973	$2U_{\text{Al-OH}} + 2\delta_{\text{Al-OH}}$
			35.8	1.9046	$U_w + \delta_w$
			9.3	1.8703	$U_w + \delta_w$
SSA	<i>A'</i> 63	268.2	226.6	1.4720	$U_w + 2\delta_w$
			457.0	2.3333	$U_{\text{Mg-OH}} + 2\delta_{\text{Mg-OH}}$
			-201.2	2.0973	$2U_{\text{Al-OH}} + 2\delta_{\text{Al-OH}}$
			-143.1	1.9423	$U_w + \delta_w$
			434.0	1.8700	$U_w + \delta_w$
CEC	<i>A'</i> 63	73.4	-1538.3	1.4033	$U_w + 2\delta_w$
			-491.6	1.3560	$2U_{\text{Al-OH}}$
			208.0	2.3333	$U_{\text{Mg-OH}} + 2\delta_{\text{Mg-OH}}$
			-185.4	2.0973	$2U_{\text{Al-OH}} + 2\delta_{\text{Al-OH}}$
			87.4	2.0157	$2U_{\text{Al-OH}} + 2\delta_{\text{Al-OH}}$
HIGF	<i>A'</i> 25	5.1	133.5	1.4313	$U_w + \delta_w$
			-65.7	1.4119	$2U_{\text{Fe-OH}}$
			94.8	1.3219	$2U_{\text{Al-OH}}$
			-34.4	1.0641	
			24.3	2.3620	$U_w$
OM	<i>R'</i> 30	10.1	-27.8	2.1202	$2U_{\text{Al-OH}} + 2\delta_{\text{Al-OH}}$
			-0.19	2.3883	$U_{\text{C-N}} \text{ amines}$
			-0.34	2.0165	$U_{\text{C-N}} \text{ phenol}$
			-0.15	1.9410	$U_w + 2\delta_w \text{ water}$
			0.50	1.5849	$U_{\text{COOH}} \text{ amino acids}$
CaCO <sub>3</sub>	<i>R</i> 25–3113	-74.4	-0.20	1.4122	$2U_w \text{ water}$
			0.07	1.0429	
			12.0	1.6473	No assignments
			-10.6	1.4825	involved

† SSA = specific surface area; CEC = cation-exchange capacity; HIGF = hygroscopic moisture; OM = organic matter.

nents) and surface properties (as obtained from the chemical correlation in Table 2).

In Group III (CaCO<sub>3</sub>), the wavelengths entered into the calibration are not those characteristic of carbonate bands. The two selected wavelengths (1.6473 and 1.4825  $\mu\text{m}$ ) are typically located in nonabsorption regions of the spectrum. Furthermore, the optimal prediction was obtained using the reflectance (*R*), and analyses of all the compressed spectra resulted in practically the same SEP. This leads us to the conclusion that the prediction of CaCO<sub>3</sub> in this case is based largely on the position within the spectrum continuum (albedo) rather than in any specific absorption band. Large amounts of CaCO<sub>3</sub> generally cause soils to be more reflective in the NIR (as in the visible spectrum).

Figure 3 shows plots of *C<sub>p</sub>* (predicted) vs. *C<sub>m</sub>* (measured) at the calibration and validation stages for each of the six examined properties. For all properties, most of the points fell in the vicinity of the 1:1 line with a few outliers. Murray (1988) noted that removing outliers from the calibration-validation set may increase prediction accuracy. However, because the prediction accuracy for all six properties was within reasonable limits of detection for rapid laboratory measurements (CPV = 0.126–0.175, Table 4), no recalculations were done to compensate for reducing the sample sets by omitting outliers. An examination to evaluate possible bias within the validation sets was performed by using the Miller and Miller (1988) test. This test was originally designed to examine a new analytical method as it relates to an old one. It statistically examines the difference between

the regression and ideal 1:1 lines. Table 6 presents the regression line ( $C_p = aC_m + b$ ) and the parameters *a*, *b*, and *r*<sup>2</sup> (taken from the validation stage) along with their statistical significance. The *r*<sup>2</sup> values for all properties do not significantly differ from 1 at the 0.05 significance level. However, based on the test of *a* and *b* parameters, only clay content, SSA, CEC, and CaCO<sub>3</sub> were found to significantly match the 1:1 line at the 0.05 significance level (clay content, CEC, and CaCO<sub>3</sub> even matched at the 0.01 significance level) and therefore held no bias. The regression line parameter *a* for OM and HIGF were significantly different from the ideal 1:1 line and, therefore, did hold some bias. These results may suggest that further research is needed to improve the prediction of OM and HIGF properties. The NIRA prediction of organic C and N contents of organic matter (0–10% of C) was found to be possible only if a homogenous soil group was used (Morra et al., 1991). Morra et al. (1991) pointed out that the necessity of using a well-defined or closed sample population for the NIRA procedure for any constituent is an inherent limitation of the procedure. We believe that the soil population in our study represents two essential groups with regard to the OM decomposition stages: (i) a group of soils that consists of 0 to 4% OM, representing greatly decomposed OM soils, and (ii) a group of soils that consists of 4 to 14% OM, representing less decomposed OM soils. In the first group, the functional properties of OM are more likely similar to humic substances, whereas in the second group, the functional properties are more likely similar to fresh litter and vegetation.

Table 4. Statistical parameters of the calibration, validation, and analytical stages.†

Property‡	Spectra		Calibration			Validation				Analytical		
	Manipulation	Bands	N <sub>c</sub>	R <sup>2</sup>	SEC	N <sub>v</sub>	r <sup>2</sup>	SEP	CPV	N <sub>a</sub>	AE	CAV
Clay	<i>A'</i>	63	35	0.76	8.6	56	0.56	10.3	0.174	91	4§	0.065
SSA	<i>A'</i>	63	35	0.84	37.7	56	0.70	50.2	0.140	91	4.0	0.010
CEC	<i>A'</i>	63	35	0.82	6.72	56	0.64	8.46	0.139	91	1.2	0.020
HIGF	<i>A''</i>	25	35	0.58	1.9	56	0.62	1.55	0.175	91	0.22	0.019
OM	<i>R'</i>	30	39	0.69	1.63	52	0.55	1.34	0.126	91	0.30	0.023
CaCO <sub>3</sub>	<i>R</i>	25–3113	22	0.70	12.9	47	0.69	11.6	0.169	91	1.3	0.017

† N<sub>c</sub> = Number of samples in the calibration stage; R<sup>2</sup> = Multiple coefficient of correlation; SEC = Standard error of calibration; N<sub>v</sub> = Number of samples in the validation stage; r<sup>2</sup> = Linear coefficient of correlation; SEP = Standard error of performance; N<sub>a</sub> = Number of samples in the analytical stage; CPV = Coefficient of prediction variation (SEP/[Max-Min]); AE = Analytical error; CAV = Coefficient of analytical variation (AE/[Max-Min]).

‡ SSA = specific surface area; CEC = cation-exchange capacity; HIGF = hygroscopic moisture; OM = organic matter.

§ Estimated value.

Table 6. The regression parameters of the validation sets ( $r^2$  = coefficient of correlation,  $a$  = slope, and  $b$  = the intercept) of six soil properties in the validation stage. Also significant is that  $r^2$  does not differ from unity,  $a$  does not differ from 1, and  $b$  does not differ from 0 (see text).

Property†	$r^2$	$a$	$b$
Clay	0.56*	0.76**	5.96**
SSA	0.69*	0.81**	23.6*
CEC	0.64*	0.68**	8.50**
HIGF	0.62*	0.56	1.99**
OM	0.51*	0.66	0.82**
CaCO <sub>3</sub>	0.67*	0.78**	7.42**

\*, \*\* Significant at the 0.05 and 0.01 probability levels, respectively. † SSA = specific surface area; CEC = cation-exchange capacity; HIGF = hygroscopic moisture; OM = organic matter.

evaluation of soil properties in the laboratory and in the field, soil survey, and remote sensing).

## SUMMARY AND CONCLUSIONS

Simultaneous and rapid, nonrestrictive determination of clay content, SSA, CEC, HIGF, OM, and CaCO<sub>3</sub> in soils is possible by using the NIRa technique. The original spectral data have to be converted into new spectral data by using several mathematical manipulations such as spectral compression, conversion to absorbance, and derivation. Many manipulations should be examined to select the optimal prediction equation. A low SEP and a high coefficient of correlation ( $r^2$ ) in the validation stage enable selection of the best manipulation. Compressing high-resolution spectra into low-resolution spectra yields better validation results. Most of the wavelengths entered into the calibration equations are spectroscopically explained, providing strong support for the NIRa prediction. A slight bias occurs in the validation of OM and HIGF, suggesting that great care is needed in the preparation and handling of these properties for NIRa measurement and analysis. Careful utilization of the method was required for each of the properties examined. Results can be improved by selecting appropriate calibration samples for a given soil constituent and population. It should be remembered that the NIRa approach will never replace the traditional analytical methods. However, for rapid, nonrestrictive measurements of a large number of samples within a relatively heterogeneous population, the NIRa achievements presented in this study are likely to be good. Further study and use of the NIRa approach in the soil science field is strongly recommended.

## REFERENCES

- Al-Abbas, H. H., P. H. Swain, and M. F. Baumgardner. 1972. Relating organic matter and clay content to the multispectral radiance of soils. *Soil Sci. Soc. Am.* 36:477-485.
- Banin, A., and A. Arniel. 1970. A correlation study of the chemical and physical properties of a group of natural soils of Israel. *Geoderma* 3:185-198.
- Baumgardner, M. F., S. J. Kristof, C. J. Johansen, and A. L. Zachary. 1970. Effect of organic matter on the multispectral properties of soils. *Proc. Indian Acad. Sci.* 79:413-422.
- Ben-Dor, E., and A. Banin. 1989. Determination of organic matter content in arid-zone soils using simple "loss-on-ignition" method. *Commun. Soil Sci. Plant Anal.* 20:1675-1695.
- Ben-Dor, E., and A. Banin. 1990a. Near infrared reflectance analysis of carbonate concentration in soils. *Appl. Spectrosc.* 44:1064-1069.
- Ben-Dor, E., and A. Banin. 1990b. Diffuse reflectance spectra of smectite minerals in the near infrared and their relation to chemical composition. *Sci. Geol. Bull.* 43(2-4):117-128.
- Ben-Gera, I., and K. H. Norris. 1968. Determination of moisture content in soybeans by direct spectrophotometry. *Isr. J. Agric. Res.* 18:124-132.
- Bowers, S. A., and R. J. Hanks. 1965. Reflection of radiant energy from soils. *Soil Sci. Soc. Am.* 100:130-138.
- Datal, R. C., and R. J. Henry. 1986. Simultaneous determination of moisture, organic carbon and total nitrogen by near infrared reflectance spectroscopy. *Soil Sci. Soc. Am.* 50:120-123.
- Davies, A. M. 1987. Near infrared spectroscopy: Time for the giant to wake up. *Eur. Spectrosc. News* 73:10-16.
- Davies, A. M., and A. Grant. 1987. Review: Near infrared analysis of food. *Int. J. Food Sci. Technol.* 22:191-207.
- Garrey, W. 1986. Spectral reflectance of carbonate minerals in the visible and near infrared (0.35-2.55  $\mu$ m): Calcite, aragonite and dolomite. *Am. Mineral.* 71:151-162.
- Gardner, W. H. 1986. Water content. p. 493-544. *In* A. Klute (ed.), *Methods of soil analysis*. Part 1. 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.
- Ge, G. W., and J. W. Bauder. 1986. Particle-size analysis. p. 383-411. *In* A. Klute (ed.), *Methods of soil analysis*. Part 1. 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.
- Hunt, G. R., and J. W. Salisbury. 1970. Visible and near infrared spectra of minerals and rocks. I: Silicate minerals. *Mod. Geol.* 1:283-300.
- Miller, J. C., and J. N. Miller. 1988. Statistics for analytical chemistry. John Wiley & Sons, New York.
- Morra, M. J., M. H. Hall, and L. L. Freeborn. 1991. Carbon and nitrogen analysis of soil fractions using near-infrared reflectance spectroscopy. *Soil Sci. Soc. Am.* 55:288-291.
- Murray, I. 1988. Aspect of interpretation of NIR spectra. p. 9-21. *In* C. S. Creaser and A. M. C. Davies (ed.), *Analytical application of spectroscopy*. R. Soc. Chem., London.
- Nelson, R. E. 1986. Carbonate and gypsum. p. 181-197. *In* A. L. Page et al. (ed.), *Methods of soil analysis*. Part 2. 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.
- Norris, K. H. 1988. History, present state and future prospects for near-infrared spectroscopy. p. 3-9. *In* C. S. Creaser and A. M. C. Davies (ed.), *Analytical application of spectroscopy*. R. Soc. Chem., London.
- Norris, K. H., and P. C. Williams. 1984. Optimization of mathematical pretreatment for surface-area determination of soils and clays. *Soil Sci. Soc. Am.* 47:1056-1058.
- Stark, E., K. Luchter, and M. Margoshes. 1986. Near-infrared analysis (NIRa): A technology for quantitative and qualitative analysis. *Appl. Spectrosc. Rev.* 22:335-399.
- Thomas, G. W. 1986. Exchangeable cations. p. 159-165. *In* A. L. Page et al. (ed.), *Methods of soil analysis*. Part 2. 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.
- Wezel, D. L. 1983. Near-infrared reflectance analysis: Sleeper among spectroscopic techniques. *Anal. Chem.* 55:1165A-1176A.

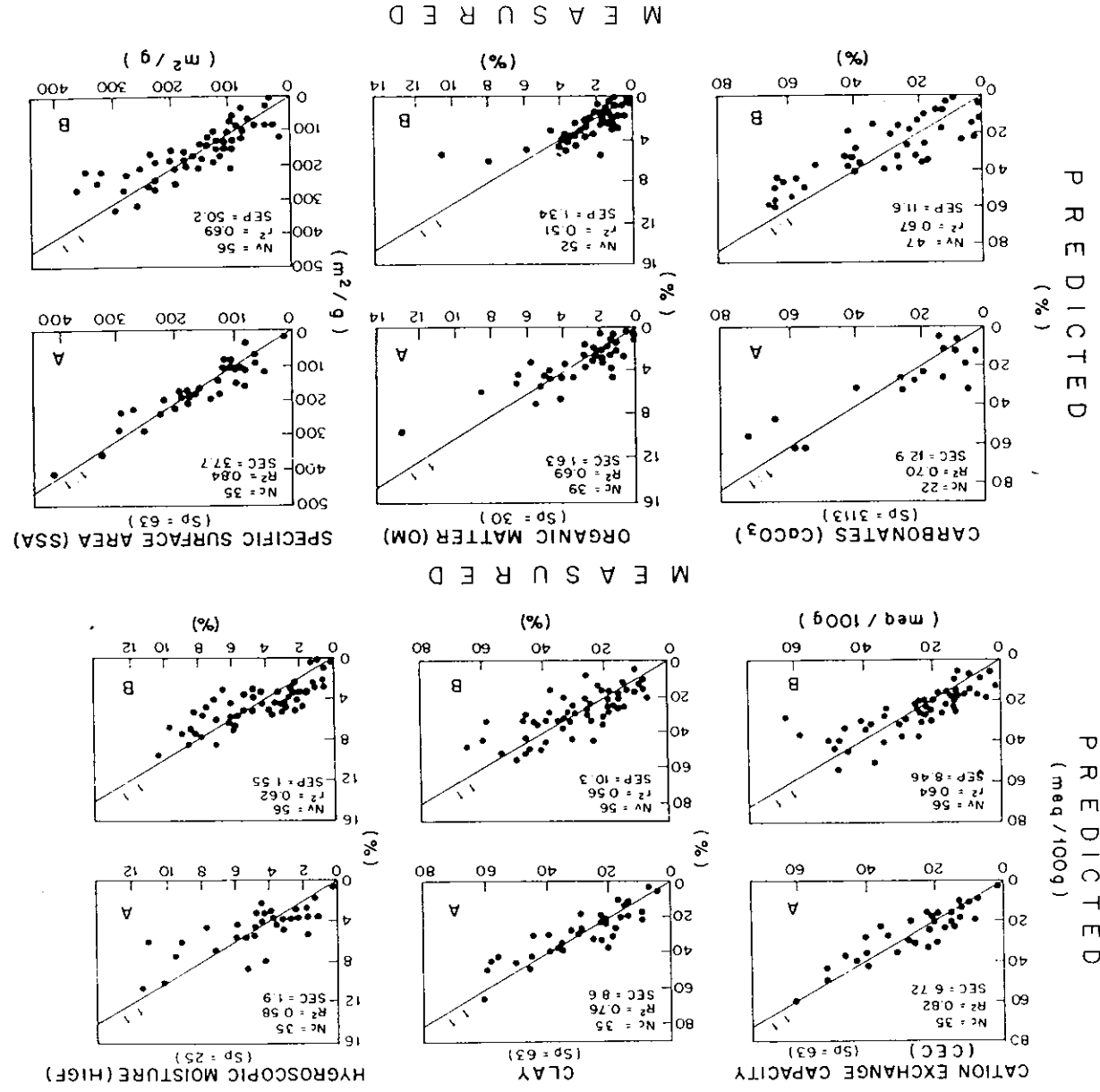


Fig. 3. Plots of the predicted ( $C_p$ ) vs. measured ( $C_m$ ) values at the (A) calibration and (B) validation stages for each property, using the optimal prediction equation given in Table 5 ( $N_c$  = number of samples in the calibration stage;  $N_v$  = number of samples in the validation stage; SEP = standard error of performance; SEC = standard error of calibration).

As a result, the prediction of the OM across a wide chemical range (0.09-13.23%) is not consistent within the two populations and, hence, significant bias occurs. As mentioned before, Datal and Henry (1986) were able to predict soil organic C and N in soil using the NIRa approach within the low-OM range (0-2.6% of C). Their prediction was possible probably because the decomposition stage in most of the soil samples was similar (the C/N ratio, which may shed some light on this issue, was not reported). Careful examination of Datal and Henry's (1986) results shows, apparently, some bias in both low and high OM content. Because they performed no bias analysis, it is speculated that careful consideration of the OM decomposition stage in soil (and hence the correct OM composition) could yield more accurate results.

The significant bias of the HIGF property may be a result of change in atmosphere (such as relative humidity) that occurred between the time that the analysis of the HIGF content was performed and the spectral measurements were acquired.

A comparison of the SEPs obtained by NIRa with the experimental errors of the independent analyses (Table 4) indicates better detection and higher accuracy for the analytical measurements (compare the coefficients of analytical variation for the independent analytical analyses to the CPVs used in the validation stage, 0.010-0.065 and 0.126-0.175, respectively). However, for rapid field and laboratory measurements, the accuracy obtained by the NIRa method is quite sufficient. This may rank the NIRa approach as an applied and promising technique for many applications in the soil science field (e.g., rapid

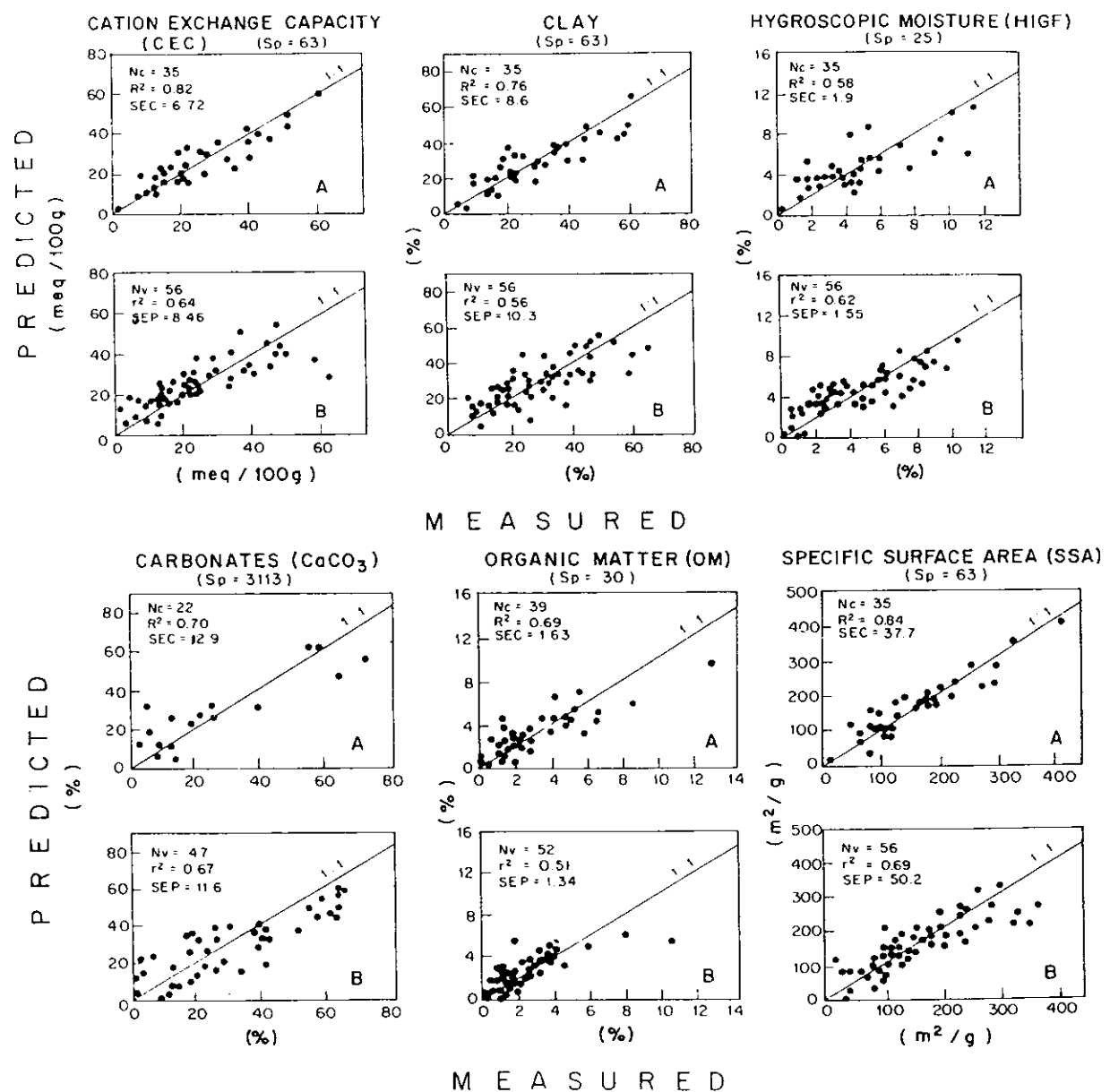


Fig. 3. Plots of the predicted ( $C_p$ ) vs. measured ( $C_m$ ) values at the (A) calibration and (B) validation stages for each property, using the optimal prediction equation given in Table 5 ( $N_c$  = number of samples in the calibration stage;  $N_v$  = number of samples in the validation stage; SEP = standard error of performance; SEC = standard error of calibration).

As a result, the prediction of the OM across a wide chemical range (0.09–13.23%) is not consistent within the two populations and, hence, significant bias occurs. As mentioned before, Dalal and Henry (1986) were able to predict soil organic C and N in soil using the NIRA approach within the low-OM range (0–2.6% of C). Their prediction was possible probably because the decomposition stage in most of the soil samples was similar (the C/N ratio, which may shed some light on this issue, was not reported). Careful examination of Dalal and Henry's (1986) results shows, apparently, some bias in both low and high OM content. Because they performed no bias analysis, it is speculated that careful consideration of the OM decomposition stage in soil (and hence the correct OM composition) could yield more accurate results.

The significant bias of the HIGF property may be a result of change in atmosphere (such as relative humidity) that occurred between the time that the analysis of the HIGF content was performed and the spectral measurements were acquired.

A comparison of the SEPs obtained by NIRA with the experimental errors of the independent analyses (Table 4) indicates better detection and higher accuracy for the analytical measurements (compare the coefficients of analytical variation for the independent analytical analyses to the CPVs used in the validation stage, 0.010–0.065 and 0.126–0.175, respectively). However, for rapid field and laboratory measurements, the accuracy obtained by the NIRA method is quite sufficient. This may rank the NIRA approach as an applied and promising technique for many applications in the soil science field (e.g., rapid

Table 6. The regression parameters of the validation sets ( $r^2$  = coefficient of correlation,  $a$  = slope, and  $b$  = the intercept) of six soil properties in the validation stage. Also significant is that  $r^2$  does not differ from unity,  $a$  does not differ from 1, and  $b$  does not differ from 0 (see text).

Property†	$r^2$	$a$	$b$
Clay	0.56*	0.76**	5.96**
SSA	0.69*	0.81**	23.6*
CEC	0.64*	0.68**	8.50**
HIGF	0.62*	0.56	1.99**
OM	0.51*	0.66	0.82**
CaCO <sub>3</sub>	0.67*	0.78**	7.42**

\*, \*\* Significant at the 0.05 and 0.01 probability levels, respectively.

† SSA = specific surface area; CEC = cation-exchange capacity; HIGF = hygroscopic moisture; OM = organic matter.

evaluation of soil properties in the laboratory and in the field, soil survey, and remote sensing).

## SUMMARY AND CONCLUSIONS

Simultaneous and rapid, nonrestrictive determination of clay content, SSA, CEC, HIGF, OM, and CaCO<sub>3</sub> in soils is possible by using the NIRA technique. The original spectral data have to be converted into new spectral data by using several mathematical manipulations such as spectral compression, conversion to absorbance, and derivation. Many manipulations should be examined to select the optimal prediction equation. A low SEP and a high coefficient of correlation ( $r^2$ ) in the validation stage enable selection of the best manipulation. Compressing high-resolution spectra into low-resolution spectra yields better validation results. Most of the wavelengths entered into the calibration equations are spectroscopically explained, providing strong support for the NIRA prediction. A slight bias occurs in the validation of OM and HIGF, suggesting that great care is needed in the preparation and handling of these properties for NIRA measurement and analysis. Careful utilization of the method was required for each of the properties examined. Results can be improved by selecting appropriate calibration samples for a given soil constituent and population. It should be remembered that the NIRA approach will never replace the traditional analytical methods. However, for rapid, nonrestrictive measurements of a large number of samples within a relatively heterogeneous population, the NIRA achievements presented in this study are likely to be good. Further study and use of the NIRA approach in the soil science field is strongly recommended.

## REFERENCES

- Al-Abbas, H.H., P.H. Swain, and M.F. Baumgardner. 1972. Relating organic matter and clay content to the multispectral radiance of soils. *Soil Sci.* 114:477–485.
- Banin, A., and A. Amiel. 1970. A correlation study of the chemical and

- physical properties of a group of natural soils of Israel. *Geoderma* 3: 185–198.
- Baumgardner, M.F., S.J. Kristof, C.J. Johansen, and A.L. Zachary. 1970. Effect of organic matter on the multispectral properties of soils. *Proc. Indian Acad. Sci.* 79:413–422.
- Ben-Dor, E., and A. Banin. 1989. Determination of organic matter content in arid-zones soils using simple "loss-on-ignition" method. *Commun. Soil Sci. Plant Anal.* 20:1675–1695.
- Ben-Dor, E., and A. Banin. 1990a. Near infrared reflectance analysis of carbonate concentration in soils. *Appl. Spectrosc.* 44:1064–1069.
- Ben-Dor, E., and A. Banin. 1990b. Diffuse reflectance spectra of smectite minerals in the near infrared and their relation to chemical composition. *Sci. Geol.* 43(2–4):117–128.
- Ben-Gera, I., and K.H. Norris. 1968. Determination of moisture content in soybeans by direct spectrophotometry. *Isr. J. Agric. Res.* 18:124–132.
- Bowers, S.A., and R.J. Hanks. 1965. Reflection of radiant energy from soils. *Soil Sci.* 100:130–138.
- Dalal, R.C., and R.J. Henry. 1986. Simultaneous determination of moisture, organic carbon and total nitrogen by near infrared reflectance spectroscopy. *Soil Sci. Soc. Am. J.* 50:120–123.
- Davies, A.M. 1987. Near infrared spectroscopy: Time for the giant to wake up. *Eur. Spectrosc. News* 73:10–16.
- Davies, A.M., and A. Grant. 1987. Review: Near infrared analysis of food. *Int. J. Food Sci. Technol.* 22:191–207.
- Gaffey, W. 1986. Spectral reflectance of carbonate minerals in the visible and near infrared (0.35–2.55  $\mu$ m): Calcite, aragonite and dolomite. *Am. Mineral.* 71:151–162.
- Gardner, W.H. 1986. Water content. p. 493–544. *In* A. Klute (ed.) *Methods of soil analysis. Part 1.* 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.
- Gee, G.W., and J.W. Bauder. 1986. Particle-size analysis. p. 383–411. *In* A. Klute (ed.) *Methods of soil analysis. Part 1.* 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.
- Hunt, G.R., and J.W. Salisbury. 1970. Visible and near infrared spectra of minerals and rocks. I: Silicate minerals. *Mod. Geol.* 1:283–300.
- Miller, J.C., and J.N. Miller. 1988. *Statistics for analytical chemistry.* John Wiley & Sons, New York.
- Morra, M.J., M.H. Hall, and L.L. Freeborn. 1991. Carbon and nitrogen analysis of soil fractions using near-infrared reflectance spectroscopy. *Soil Sci. Soc. Am. J.* 55:288–291.
- Murray, I. 1988. Aspect of interpretation of NIR spectra. p. 9–21. *In* C.S. Creaser and A.M.C. Davies (ed.) *Analytical application of spectroscopy.* R. Soc. Chem., London.
- Nelson, R.E. 1986. Carbonate and gypsum. p. 181–197. *In* A.L. Page et al. (ed.) *Methods of soil analysis. Part 2.* 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.
- Norris, K.H. 1988. History, present state and future prospects for near-infrared spectroscopy. p. 3–9. *In* C.S. Creaser and A.M.C. Davies (ed.) *Analytical application of spectroscopy.* R. Soc. Chem., London.
- Norris, K.H., and P.C. Williams. 1984. Optimization of mathematical treatment of raw near-infrared signal in the measurements of protein in hard red spring wheat. I: influence of particle size. *Cereal Chem.* 61(2):158–165.
- Rathner-Zohar, Y., A. Banin, and Y. Chen. 1983. Oven drying as pretreatment for surface-area determination of soils and clays. *Soil Sci. Soc. Am. J.* 47:1056–1058.
- Stark, E., K. Luchter, and M. Margoshes. 1986. Near-infrared analysis (NIRA): A technology for quantitative and qualitative analysis. *Appl. Spectrosc. Rev.* 22:335–399.
- Thomas, G.W. 1986. Exchangeable cations. p. 159–165. *In* A.L. Page et al. (ed.) *Methods of soil analysis. Part 2.* 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.
- Wetzel, D.L. 1983. Near-infrared reflectance analysis: Sleeper among spectroscopic techniques. *Anal. Chem.* 55:1165A–1176A.