The potential of mid- and near-infrared diffuse reflectance spectroscopy for determining major- and trace-element concentrations in soils from a geochemical survey of North America

James B. Reeves, III [a,*], David B. Smith [b]

[a] *Environmental Management and Byproduct Utilization Laboratory, ANRI, USDA, Beltsville, MD 20705, USA*
[b] *U.S. Geological Survey, Denver Federal Center, MS 973, Denver, CO 80225, USA*

* Corresponding author, *E-mail:* Jreeves@anri.barc.usda.gov (J. Reeves)

Abstract

In 2004, soils were collected at 220 sites along two transects across the USA and Canada as a pilot study for a planned soil geochemical survey of North America (North American Soil Geochemical Landscapes Project). The objective of the current study was to examine the potential of diffuse reflectance (DR) Fourier Transform (FT) mid-infrared (mid-IR) and near-infrared (NIRS) spectroscopy to reduce the need for conventional analysis for the determination of major and trace elements in such continental-scale surveys. Soil samples (n=720) were collected from two transects (east-west across the USA, and north-south from Manitoba, Canada to El Paso, Texas (USA), n=453 and 267, respectively). The samples came from 19 USA states and the province of Manitoba in Canada. They represented 31 types of land use (e.g., national forest, rangeland, etc.), and 123 different land covers (e.g., soybeans, oak forest, etc.). The samples represented a combination of depth-based sampling (0-5 cm) and horizon-based sampling (O, A and C horizons) with 123 different depths identified. The set was very diverse with few samples similar in land use, land cover, etc. All samples were analyzed by conventional means for the near-total concentration of 49 analytes (C_{total}, $C_{carbonate}$ and $C_{organic}$, and 46 major and trace elements). Spectra were obtained using dried, ground samples using a Digilab FTS-7000 FT spectrometer in the mid- (4000 to 400 cm^{-1}) and near-infrared (10,000 to 4000 cm^{-1}) at 4 cm^{-1} resolution (64 co-added scans per spectrum) using a Pike AutoDIFF DR autosampler. Partial least squares calibrations were develop using: (1) all samples as a calibration set; (2) samples evenly divided into calibration and validation sets based on spectral diversity; and (3) samples divided to have matching analyte concentrations in calibration and validation sets. In general, results supported the conclusion that neither mid-IR nor NIRS would be particularly useful in reducing the need for conventional analysis of soils from this continental-scale geochemical survey. The extreme sample diversity, likely caused by the widely varied parent material, land use at the site of collection (e.g. grazing, recreation, agriculture , etc.), and climate resulted in poor calibrations even for C_{total}, $C_{organic}$ and $C_{carbonate}$. The results indicated potential for mid-IR and NIRS to differentiate soils containing high concentrations (> 100 mg/kg) of some metals (e.g., Co, Cr, Ni) from low level samples (<50 mg/kg). However, because of the small number of high-level samples, it is possible that differentiation was based on factors other than metal concentration. Results for Mg and Sr were good, but results for other metals examined

were fair to poor, at best. In essence, it appears that the great variation in chemical and physical properties seen in soils from this continental-scale survey resulted in each sample being virtually unique. Thus, suitable spectroscopic calibrations were generally not possible.

1. Introduction

Near-infrared diffuse reflectance (DR) spectroscopy (NIRS) has over the last several decades become one of the premier methods for the rapid analysis of agricultural products and by-products such as forage, silage, grain, cotton, foodstuffs, etc. (Williams and Norris, 2001; Roberts et. al., 2004). Diffuse reflectance Infrared Fourier Transform spectroscopy (mid-IR), a mid-infrared technique, has also been examined and compared to near-infrared (NIR) for the determination of agricultural materials such as forage and grain (Reeves, 1994, 1996; Reeves et al., 1990, 1999a, b), but its use is not nearly as prevalent or advanced as that of NIRS due primarily to the long held belief that sample dilution with KBr or similar diluents was/is necessary (Coleman, 1993; Griffiths and de Haseth, 2007). Over the last two decades, both NIRS and the mid-IR have increasingly been applied to the analysis of soils particularly for soil C as driven by the interests in C sequestration (Brown et al., 2005, 2006; Janik and Skjemstad, 1995; Janik et al., 1995, 1998, 2007; Madari et al, 2005; Malley et al., 2004; McCarty and Reeves, 2001; McCarty et al., 2002; Merry and Janik, 2001; Nguyen et al., 1991; Reeves et al., 1999a, b, 2001, 2002; Reeves and McCarty, 2001; Sankey et al., 2008; Viscarra Rossel et al., 2006; Wu et al., 2005).

1.1 Calibration development

Quantitative analysis based on either NIR or mid-IR spectra requires the development of calibrations that relate the spectral information to known analyte concentrations. Presently, this is generally carried out using the entire spectra as opposed to only a few wavelengths, and utilizes multivariate calibration methods such as partial least squares (PLS), principal component regression, boosted regression trees, and neural networks (Beebe et al., 1998; Brown et al., 2006; Naes et al., 2002; PLSPlus, 1994). "Chemometrics is the field of extracting information from multivariate chemical data using tools of statistics and mathematics" (http://www.infometrix.com/chemometrics/chemometrics.html), and while the mathematical methods used are well understood and based on sound principles, there are factors beyond the mathematics which can greatly influence calibration accuracy and usefulness. For example, it has often been found using NIRS that calibrations work best when developed for a limited, defined population of samples, e.g. a different calibration for alfalfa hay, timothy hay, etc. rather than one calibration for all hays (Universal Calibration concept). Similarly, calibrations may work better for limited analyte concentration ranges (Brown et al., 2005; Madari el al., 2005).

1.2. Diffuse reflectance spectroscopy and soils

Although NIRS has been extensively used for agricultural products for decades (Williams, 1975; Norris et al., 1976; Williams and Norris, 2001; Roberts et. al., 2004), interest in soils appeared to really grow about 1999 at the International NIR Conference in Verona, Italy (Davies and Giangiacommo, 2000), although many earlier references to its use exist (see review by Malley et al., 2004). Also, most of the efforts have focused on the use of NIR to analyze the organic fraction of soils, e.g. measures of $C_{organic}$ and $N_{organic}$. These efforts have shown that NIR can accurately determine the $C_{organic}$ and $N_{organic}$ content and often forms thereof, but performs poorly, or with great variability from study to study, for mineral forms of Ag, Al, Cd, Cu, Co, Fe, K, N, P, Pb, Na, Ni, Se, Si, Zn and pH. However, it has been successful for Ca and Mg perhaps due to their "correlations with carbonate content and factors influencing cation exchange capacity, such as organic matter content" (Malley et al., 2004).

Efforts using mid-IR to analyze soils in the same manner as NIRS has generated considerably less total research. This was, and still is, largely due to the belief that samples require dilution with KBr or similar diluents for any mid-IR DR spectroscopy (Coleman, 1993, Griffiths and de Haseth, 2007; Reeves, 2003). However, research comparing mid-IR and NIRS on the same samples has invariably demonstrated that mid-IR generally outperforms NIRS in the analysis of soils, although almost all efforts have concentrated on C and N fractions (Janik et al., 1998; Madari et al., 2005; McCarty and Reeves, 2001; McCarty et al., 2002; Reeves et al., 2001, 2002). However, Janik and Skemstad (1995), Siebielec et al. (2004) and Wu et al. (2005) have demonstrated that mid-IR and NIRS also have potential for the determination of analytes such as Fe, Ca, Cd, Cr, Cu, Mg, Ni, Pb, Si, Ti and Zn.

1.3. Objectives

The objective of this study was to examine DR mid-IR and NIR spectroscopy as possible tools for the determination of major and trace elements in soils obtained from a soil geochemical survey of North America.

2. Methods

2.1. Sample collection and analysis

Soil samples were obtained from 265 sites from two continental-scale transects across the USA and Canada (Smith et al., 2005, 2009). A north-to-south transect consisting of 105 sites extended from northern Manitoba, Canada, to the USA-Mexico border near El Paso, Texas. A west-to-east transect consisting of 160 sites followed the 38th parallel from the Pacific coast of the USA just north of San Francisco, California, to the Atlantic coast in northern Virginia. Sampling protocols

at each site included one depth-based sample collected at 0-5 cm and up to 3 horizon-based samples (O, A and C horizons). Samples were collected using shovels and augers to access the deeper material with samples representative of each site visited. This study examined 720 samples representing 241 A-horizons, 252 C horizons, and 227 samples from 0-5-cm depth.

Forty-eight different analytes were determined. Aluminum, Ag, As, Ba, Be, Bi, Ca, Cd, Ce, Cs, Co, Cr, Cu, Fe, Ga, In, K, La, Li, Mg, Mn, Mo, Na, Nb, Ni, P, Pb, Rb, S, Sb, Sc, Sn, Sr, Te, Th, Ti, Tl, U, V, W, Y and Zn were determined by inductively coupled plasma-mass spectrometry and inductively coupled plasma-atomic emission spectrometry; Hg by cold vapor-atomic absorption spectrometry; Sb and Se by hydride generation-atomic absorption spectrometry (HG-AAS); carbonate-C ($C_{carbonate}$) by coulometric titration; and total C (C_{total}) and S (T-S) by combustion. Organic C ($C_{organic}$) was determined by difference. All values for Ag were below the detection limit for the samples examined in this study and, thus, no results for Ag are reported. Complete chemical results for all samples collected along the two transects can be found in Smith et al. (2005).

2.2. Spectroscopy

The 720 samples provided for the current study were scanned as dried, ground samples in the NIR and mid-IR on a Digilab (Varian, Inc., Palo Alto, CA) FTS 7000 Fourier transform spectrometer equipped with a liquid N_2 cooled InSb detector and a quartz beam splitter for the NIR range and a Peltier cooled DTGS (deuterated triglycine sulfate) detector and KBr beam splitter for the mid-IR. Spectra were obtained in diffuse reflectance mode using a Pike (Pike Technologies, Madison, WI) AutoDIFF auto-sampler with sulfur (Recommended by Karl Norris for NIR spectra) and KBr as background samples for the NIR and mid-IR, respectively. All spectra were computed as pseudo-absorbance (log [1/reflectance] as compared to absorbance which is log [1/transmittance]) due to being collected as DR spectra. The conversion to log(1/reflectance) is used as it is considered the best approximation for linearizing reflectance data with concentration. Spectral data was collected at 4 cm^{-1} resolution (64 co-added scans per spectrum) from 10,000 to 4000 cm^{-1} and 4000 to 400 cm^{-1} for the NIR and mid-IR, respectively.

2.3. Chemometrics/Calibration development/Statistical analysis

Spectra were examined qualitatively using GRAMS/AI Ver. 7.02 (Thermo Galactic, Salem, NH). Calibration development was performed using PLS regression using a SAS® (SAS, Institute, Inc., Cary, NC) program previously published (Reeves and Delwiche, 2003, 2004) and available from <www.nirpublications.com/software/index.html>. A variety of mathematical pre-treatments (1st and 2nd derivatives with gaps of 1, 2, 4, 8, 16, 32 and 64 data points, multiplicative scatter) were used to determine the best calibration method for each analyte based on leave-one-

out cross validations. In addition, all spectra were mean-centered and variance scaled. Results are reported for the final calibrations resulting in the highest R^2 and lowest root mean squared deviation (RMSD = $[\sum (\text{Actual-Predicted})^2/n]^{1/2}$). Further details on the pre-treatments may be found in Reeves and Delwiche (2003).

Calibrations were developed using a variety of sample set designs, including: (1) Calibrations based on all samples with no independent validation/test set. (2) Calibrations based on a 50/50 split of samples selected by cluster analysis using SAS FASTCLUS (SAS, 2003). (3) Calibrations using a 50/50 split based on analyte concentrations in which samples were ordered from low to high for each analyte and every other sample placed into the calibration or validation set, respectively. Calibrations were then developed using each spectral range (NIR or mid-IR). Not all possible combinations of spectral range, sample splits and analytes were tested based on initial experiments that showed some calibrations to be of little value.

Summary statistics and frequency distributions were computed using SAS® Proc Means and Proc Freq.

3. Results and discussion

3.1. Sample characterization

Table 1 shows the summary statistics for the analytes examined. Any value below the method detection limit was replaced with one-half the detection limit prior to statistical calculations.

Computation of frequency distributions showed that the 720 samples examined came from 19 different states within the USA and one province in Canada, with 453 samples from the east-west transect and 267 from the north-south transect. Due to variations in the depths of the horizons with location, samples were obtained from 193 different depths. Samples were also defined by land use (e.g. agriculture, pasture, national forest, etc.) and land cover (e.g., alfalfa, corn, cactus, etc.) with 31 different types of land use and 123 different land covers listed. Three types of land use accounted for 584 of the 720 samples: agriculture (n=329), forest (n=161) and rangeland (n=94). No single land cover accounted for more than 12% of the total samples with trees of various types (n=87), hayfields (n=40) and soybeans (n=33) being the 3 largest classifications. Cross frequency distributions of land use by land cover surprisingly only resulted in 138 different combinations with the top 3 being forest-trees (n=57), agriculture-hayfield (n=40) and agriculture-soybeans (n=33). It should be noted that these numbers are not absolute, as they are highly dependent on how land use and land cover were classified by the individual soil sampling crew. However, the frequency results alone indicate a very diverse set of samples. Diversity of the sample set with regard to soil parent material, mineralogy and climate is discussed by Eberl and Smith (2009), Garrett (2009), and Woodruff et al. (2009).

While different investigators use different metrics to state the usefulness of calibrations, the R^2 value can be a quick method if the analyte range is not arbitrarily narrow, e.g., if all the samples have values between .5 and .6 X, even a good fit may not produce a high R^2 value. Examining the data in Table 1 shows that most analytes had concentration ranges of at least 50X, which is excellent, with Ag (not presented as all values were below detection limits) In, Sb, and Te being obvious exceptions, and As, Se, Sn and Tl also being potential problems. However, it should also be noted that for some analytes (Co, Cr, Hg, Ni, P, Pb, Sr, Te and T-S) the data are highly skewed and with high kurtosis, indicating many low values and a few samples with high analyte concentrations.

The data on the structure of the sample set examined indicate that the samples in question are extremely diverse in nature, as would be expected when samples are taken with regards to a geographical plan (transect) on a continental scale. Some elements (e.g., Cr, Hg, Ni, P, Pb, Sr, Te and T-S) may not even represent a well structured data set with regard to analyte concentration distribution (i.e., even distribution of values covering a wide range of concentrations). As discussed in the introduction, previous work with other materials has shown that calibrations are often best when developed for specific classes of samples (e.g. different calibrations for each different forage or grain) rather than for many different classes together (the so called Universal Calibration). Examination of the data in Table 1 suggests this data falls more into the latter class. Finally, the highly skewed nature of some of the analyte concentrations combined with the diverse nature of the samples might allow calibrations based on factors other than the analyte concentrations themselves, e.g. if only one sample has a value which is 100X higher than any other and it also happens to be the only dark colored sample, then a calibration could separate the samples into two classes (dark and light colored) and produce a calibration which might appear acceptable if only examined from the basis of statistics such as R^2.

3.2. Mid-IR and NIR sample spectra

The mid-IR and NIR spectra of the samples with the highest (solid line) and lowest (dashed line) concentrations of C_{total} (34.2 and 0.04%, respectively) are shown in Figures 1 (mid-IR) and 2 (NIR). The results in Figure 1 show just how diverse the samples in question are. The large band(s) between 2000 and 1200 cm^{-1} in the low C_{total} soil (dashed line) are due to silica, indicating a soil high in mineral matter and low in organic matter. The small bands between 3000 and 2800 cm^{-1} indicate C-H in the high C_{total} soil (solid line). While the spectrum of the low $C_{carbonate}$ sample (not shown) appeared much like the high C_{total} spectra seen in Figure 1, the high $C_{carbonate}$ spectra looked nothing like any of the other 3, with most of the dominant bands being due to the presence of carbonates (Reeves 2003; Reeves et al., 2005).

While the mid-IR spectra appear very different, depending on the level of C_{total} or $C_{carbonate}$, the corresponding NIR spectra (spectra not shown) for the same two samples looked largely alike, except for the differences in slope and overall absorbance levels. The large differences in absorbance levels are mainly due to particle size differences, which do not cause the same baseline shifting in mid-IR spectra. The slope in the spectra between 10,000 and 6500 cm^{-1} of the high C_{total} sample (solid line, Fig. 2) is most likely due to color caused by the high C content. The large peaks centered around 7000 and 5000 cm^{-1} are due to OH, largely from residual moisture present even in dried samples, but also from OH groups in organic matter and inorganic minerals such as clays. The presence of organic matter is also shown by the smaller peaks located within the large peak at 7000 cm^{-1} and also those between 4500 and 4000 cm^{-1} seen in Figure 2 (solid line), while a small peak at ~ 4200 cm^{-1} in high $C_{carbonate}$ samples (spectra not shown) is due to carbonates. Overall, there appears to be a great deal more information in the mid-IR spectra based on visual observations, although for soils, much of this is a combination of organic and inorganic materials, while for the NIR, it is largely based on CH, OH and NH groups found in organics (Reeves et al., 2005). It was this apparent quantitative increase in information content seen with mid-IR spectra of forages that led to some of the first efforts to investigate mid-IR for quantitative analysis of ground, non-KBr diluted, forage samples (Reeves, 1994, 1996; Reeves et al., 1990), while it is the presence of specular (mirror like) reflection (data not shown) in the same samples (Reeves et al., 2005) which lead to the belief that all samples would need to be diluted with KBr to concentrations of 5% or less per sample for mid-IR to be useful (Coleman 1993; Griffiths and de Haseth, 2007).

3.3 PLS calibrations

3.3.1. Calibrations using all samples in the calibration set (n=720)

The best PLS calibrations one can hope to achieve for a specific set of samples, and thus the best indication of the potential of mid-IR or NIRS to be useful for determining composition of that specific set, are those developed using all available samples. If such a calibration is not found to be accurate, it is virtually impossible for the calibration to determine accurately the composition of new samples from their spectra alone. With the caveats previously discussed (good range of analyte values which are evenly distributed), R^2 values >0.9 generally indicate a very good calibration. R^2 values <0.7 are not very useful, with those <0.5 appearing as random scatter plots (personal observation). Comparing the root mean squared deviation (RMSD) to the mean can also be very useful, with most predicted samples falling within ± 1 RMSD of their true value. Another measure of the usefulness of a calibration is the ratio of performance to deviation (RPD) value (SD$_{(sample\ population)}$/RMSD), where SD is the standard deviation. RPD values below 5 indicate poor to fair calibrations, values between 5 and 6.5 are fair to good, values between 6.5

and 8 are considered very good and any value above 8 is excellent (Williams and Norris, 2001). It needs to be stated that these grading levels using RPD are from the originator of the measurement (Phil Williams); many researchers find calibrations to be useful with RPD values considerably lower than the proposed standards. As such, it is up to the reader to evaluate all the statistics provided and to decide if similar calibrations would be useful for their needs, screening versus quantitative analysis, etc.

The data presented in Table 2 show that the best mid-IR calibrations in terms of R^2 (>0.9) were for $C_{carbonate}$, Co, Cr, Mg and Ni, with those for Ca, Fe, Ga, Na, P, Sr, (all $R^2 > 0.8$) and possibly Al ($R^2 = 0.782$) acceptable. With only a few exceptions, calibrations based on mid-IR spectra were better (larger R^2 and smaller RMSD) than their NIR counterparts. Similarly, based on the RPD values, few of the calibrations would be judged to be useful for any application. As the same exact sample was scanned (sample cup filled once and spectrometer settings changed from mid-IR to NIR) and the same analyte values were used, this represents a true difference in the ability of the two methods (mid-IR and NIRS) to accurately determine soil composition and supports what has been found previously for many soil samples by researchers when examining dried, ground soils by DR spectroscopy; namely, mid-IR is the more accurate and robust method (Janik et al., 1998; Madari et al., 2005; McCarty and Reeves, 2001; McCarty et al., 2002; Reeves et al., 2001, 2002; Siebielec et al., 2004).

While the mid-IR calibrations may be more accurate than their NIRS counterparts, the real question is: Are they useful in the context of determining the analyte concentrations in samples from a continental scale soil geochemical survey? With so many analytes available, it would be easy to get lost in the numbers. Therefore, only a relatively few analytes will be discussed in relation to previous studies and the nature of the samples in question (diverse set from two continental-scale transects) and using only PLS calibrations based on mid-IR spectra.

Figure 3 shows the predicted results versus actual results for C_{total} using all 720 samples as a calibration set. Two things stand out. First, one high C_{total} sample was considerably under predicted. While a bad reference value is possible, it is highly unlikely that a value of >30% would be found for a sample with a true value of <5%. Second, samples with values <5% are very scattered and not generally predicted well, something not found in other soil studies (Janik et al., 1998; Madari et al., 2005; McCarty and Reeves, 2001; McCarty et al., 2002; Reeves et al., 2001, 2002). Results for $C_{carbonate}$ (Fig. 4) were similar for concentrations < 3% to those for C_{total} (Fig. 3). Examination of results for Al, where sample distribution was more even, also showed some samples to be very poorly determined with several samples with values ~2%, predicted to have ~8% Al content. These results are consistent with the concept that the samples were simply too diverse in nature for the calibration to encompass, at least using all the available samples to develop the calibration. Results for Mg (Fig. 5), however, indicated that calibrations, even in a

very diverse set of soil samples, might be quite useful and accurate, similar to the findings of Malley et al. (2004). Results for Ni are shown in Figures 6-7 (results for Co were essentially the same). In both cases, an extremely wide range of concentrations was present, the effect of which can be seen in Figure 7 where only Ni concentrations <100 mg/kg are plotted. As shown, while the overall calibration (Fig. 6) looks very good, for relatively low-level samples (Fig. 7), more typical of soils throughout the two transects, the data are highly scattered. This raises two questions: (1) the spectral basis for the calibrations, and (2) the usefulness of the calibrations. The fact that the calibration is very inaccurate at concentrations <100 mg/kg, but can differentiate high- and low-concentration samples, could be due to surrogate calibrations. There are so few samples with high levels of Ni present in the calibration that the calibration could be based on factors other than Ni content if high-concentration samples had unique properties unrelated to the Ni content, but unique, by chance, to each sample. With so few samples with high levels of Ni, Co, etc., it is impossible to determine if other samples with high levels of Ni, Co, etc., but without the unique properties of the samples used in the calibration, would be accurately determined by the calibration. Again, this is due to the nature of the transect samples where very few samples are taken from each sampling site. The supposition that the diversity of the samples in this data set is the reason for both the scatter at Ni concentrations <100 mg/kg and the ability to separate high- and low-concentration samples is supported by the work of Siebielec et al. (2004). These investigators achieved excellent results for Ni in the range of 0 to 80 mg/kg ($R^2 = 0.99$), although there were only 70 samples consisting of soils contaminated by mining activities.

3.3.2. Calibrations based on 50/50 sample split into calibration and test sets using SAS Proc FASTCLUS

The true test of any calibration is the determination of samples not included in the samples used to develop the calibration. If one has all the samples available at the start, this is called a closed population (e.g., the present samples). In such a case, some of the samples can be used to develop a calibration, which is then used to determine the remaining samples, thus reducing the need for standard chemical analyses by some factor determined by the sample set split. Since the analyte values are not available in this scenario, one suggested procedure is to split the samples based on spectral similarities/differences. This avoids having mostly similar samples in the calibration set and missing those that are different, which can happen with a random split (Westerhaus et al, 2004). In Table 3, selected results (mid-IR calibration, $R^2 \geq 0.6$) based on splitting by spectral similarity are shown using mid-IR and NIRS. Results can be summarized as follows: (1) With the exception of Co, Cr and Ni (both calibrations still very good), calibrations based on mid-IR spectra performed better than those based on NIR spectra with respect to the

calibration set, as consistent with previously discussed results (Table 2). (2) With few exceptions (Co, Cr, Mg and Ni), validation or test set results based on mid-IR calibrations were considerably poorer than those found for the calibration set. This is especially true for the NIR calibrations (exceptions being Cr, Mg and Ni). (3) Results for validation set results for C determinations were particularly poor compared to the calibration results using either spectral range. Also, as before, the RPD values indicate few of the calibrations to be very useful. Overall, these results indicate that samples from transects covering large distances and crossing such a wide variety of parent materials, land uses, climatic zones, etc. are too diverse to split into calibration and validation/test sets at a level to be practical. For example, if 90% of the samples are needed to develop a calibration, one might as well analyze the remaining samples by the same conventional methods.

Splitting samples into calibration and validation sets based on spectral similarities and differences (Table 3) generally indicated problems with the splitting method and showed that calibrations based on only part of the sample set were not useful for predicting the remaining samples, even when using a 50/50 split. Figure 8 shows results obtained for $C_{carbonate}$. All samples with $C_{carbonate}$ levels > 2.5% were considered unique. These were placed in the calibration set by Proc FASTCLUS and, thus, do not appear in the validation results. The validation results thus appear very poor visually (Fig. 8), but are really no worse than those obtained in the calibration set for samples with levels < 2.5% if examined at the same plotting scale (data not shown). Results based on samples split by concentration ranges (Fig. 9) indicate that mid-IR calibrations might be useful for separating high- from low-level $C_{carbonate}$ samples to some degree, but not at determining, with any degree of accuracy, those with concentrations below 2.5%. This has not previously been found to be the case where problems with high levels of carbonates were found (unpublished data using soil samples from Natural Resources Conservation Service (NRCS) National Collection) and presumed to be due to specular reflection, which results in $CaCO_3$ having different spectral signatures at low and high concentrations (Reeves et al., 2005). The results indicate that sample diversity can lead both to good predictions on the one hand, when a few high analyte concentration samples are being segregated on a basis other than analyte content, and poor predictions at lower levels, where the diversity interferes with normal calibration development.

3.3.3. Calibrations based on 50/50 sample split into calibration and test sets using analyte concentrations

Upon examining the calibration plots from the results presented in Table 3, it was found that, in some cases, the poor validation R^2 might be due to the structure of the sample set (e.g., all the high concentration samples for analyte X were selected for the calibration set). Selected analytes

(C_{total}, $C_{carbonate}$ and Co) were therefore analyzed using calibration and validation sets split by analyte concentration in which samples were ordered by concentration and every other sample split into each set. Results were better for $C_{carbonate}$, but did not support concentration-based splitting as the solution for poor validation results as a whole. Also, while this method might be useful for testing calibrations, it is no help when one does not have analyte values, but rather needs to determine them.

Finally, because a calibration based on a more limited range of concentrations may sometimes perform better, calibrations for some analytes were tested using a limited range of concentrations. Results using only samples with $C_{carbonate}$ ≤2% and C_{total} ≤5% are shown in Figures 10-12. As shown, neither results for $C_{carbonate}$ nor C_{total} were particularly good with many samples being considerably under and over predicted concentrations for both analytes. As previous studies have demonstrated that both mid-IR and NIRS can accurately determine either analyte in similar concentration ranges with a high degree of accuracy (Janik and Skjemstad, 1995; Janik et al., 1998, 2007; Madari et al., 2005; McCarty and Reeves, 2001; McCarty et al., 2002; Reeves et al., 2001, 2002; Viscarra Rossel et al., 2006), the conclusion can only be that, at least for C determination, these samples just do not constitute a good calibration set. Indeed, calibrations using ~1000 samples from either the Brazilian National Collection (Madari et al., 2005) or the NRCS (USA) National Collection (unpublished data) indicated very good results over similar C ranges with R^2 >0.9. Thus, the only conclusion is that something about these samples is unique. The most obvious difference would be in the great diversity of locations and land uses identified for these soils. Soils obtained from a scenic site or a woodlot, for example, may have been affected by human activities, thus perhaps making some of these soils unique. Only further studies with many more soils obtained for a single land use throughout the continent of North America would be able to definitely answer such questions. Other possible reasons for the results include the methods used to obtain the samples and the mixing of samples from different depths in the same calibration. For some analytes, it may also be feasible to divide the present samples using cluster analysis based on their spectra. The problem with the diverse nature of the samples is similar to that found by others where the samples in the calibration set do not represent the samples being used to validate the calibration (Brown et al., 2005, 2006; Sankey et al., 2008) except in the present case the samples appear too diverse even for a good calibration to be developed.

Finally, examination of Al (Fig. 13) and Co (data not shown) calibrations indicated that calibrations for some elements (e.g., Al) may be feasible even in such a diverse set of soils over the entire concentration range found. However, calibrations for Co, at levels <35 mg/kg, indicate that calibrations would be useful only for differentiating high and low levels of Co in soils. Whether being able to differentiate levels of metals, such as Co or Ni, at some level would be

useful is difficult to determine based on these data alone. For example, Ni can be toxic to plants, but the toxicity levels depend on other factors such as the soil pH and aging effects (Dr. Rufus Chaney, personal communication). The differences in the quality of the results obtained for metals such as Al and Mg (previously discussed) and C may be due, at least in part, to the influence of human activity on the samples versus the spectral forms of the analytes being determined. However, addressing this issue is far beyond the scope of the present study.

4. Conclusions

The objective of this study was to examine the potential benefits of using DR FT mid-IR and NIR spectroscopy for the determination of major- and trace-element concentrations in soils obtained during the soil geochemical survey of North America (North American Soil Geochemical Landscapes Project). Soil samples for the current study were collected from 265 sites along two continental-scale transects across Canada and the US. In general, results supported the conclusion that neither mid-IR nor NIRS would be particularly useful in reducing the need for conventional chemical analysis of soils in the sample set from this continental-scale survey. The extreme diversity in the samples found in this continental-scale survey resulted in poor results even for measures of C, which have been shown to be easily determined in most previous studies. Considering that satisfactory results have been obtained in other studies with diverse soils obtained from national archives (Madari et al., 2005), it is highly likely that a large part of the problem is not the basic soils themselves, but differences induced by the wide variety of land use (e.g., agriculture, national forest, rangeland, woodlot, etc.), which was not present in the other cited investigations. Results indicated potential for these methods to differentiate soils containing high levels (>100 mg/kg) of some metals (Co, Cr, Ni) from those containing low levels (≤50 mg/kg). However, due to the small number of samples with metal concentrations >100 mg/kg, it is possible that the observed differentiation was based on factors other than metal concentration (surrogate calibrations). Thus, the ability to differentiate such samples would fail if soils similar to the high-metal soils, but with low-metal levels, were also included in the calibration. In essence, sampling at a continental scale, with the resulting wide diversity in soil parent material, climate and land use, can result in each sample being virtually unique and can create a sample set not suitable for spectroscopic-based calibrations.

4.1 Implications and Future Efforts

A natural question is whether calibrations could be developed for subsets of the samples in question. For example, a natural division might be by soil horizon or depth of collection. In this study, that would have resulted in 3 sets of approximately 250 samples each. One of the potential problems with spectral calibrations is over-fitting to the calibration set due to the large amount of

spectral information available (thousands of spectral data points) versus the number of samples available for calibration development (a percentage of the 250?). Personal investigations and observations indicate that at least 100 samples and preferably more are needed to develop calibrations while avoiding this problem. In addition, some independent set is needed to test the calibration. Thus, at best, one might reduce the need for traditional chemical analysis by perhaps 50% for this particular study if separate calibrations for each horizon were developed and found to be accurate. Also, the upper horizons are more likely to be directly affected by the vegetation present than deeper soil samples. So for these samples, such a division might be effective only for the C horizon, resulting in a reduction of perhaps 15% in the need for laboratory analysis (half of the C-horizon samples out of 720 total). It is also possible that better results might be obtained using a different calibration method, such as a locally weighted regression procedure (Centner and Massart, 1998; Lorber et al., 1998); boosted regression trees, which were shown to be superior to PLS by Brown et al. (2006); or local calibrations as discussed by Janik et al. (2007).

For national- and international-scale surveys involving larger data sets, this may be less of a problem due to greater numbers of samples being available for each location. In such cases, sub-setting based on spectral similarity (cluster analysis), or use of the different algorithms discussed, should allow calibrations to be developed while still significantly reducing the need for conventional chemical analysis. Finally, if soils within a specific area or of a specific type are of more concern, one could use the samples from a survey such as this to determine where future sampling efforts should concentrate, thus reducing the need to obtain more samples from each location.

Future efforts will be made to examine these possibilities. However, based on the effect of sample diversity on calibrations for other materials developed over the last several decades, it does not appear that spectroscopic calibrations for datasets of soils such as those examined here will be very useful due to the tremendous diversity in the samples due to soil types, horizons sampled, land use, etc.

References

Beebe, K.R., Pell, R.J., Seasholtz, M.B., 1998. Chemometrics: A Practical Guide. John Wiley & Sons, New York.

Brown, D.J., Bricklemyer, R.S. Miller, P. R. 2005. Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. Geoderma: 129, 251-267.

Brown, D.J., Shepherd, K.D., Walsh, M.G., Mays, M.D., Reinsch, T.G., 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma 132, 273-290.

Centner, V., Massart, D. L., 1998. Optimization in locally weighted regression. Anal. Chem. 70, 4206-4211.

Coleman, P.B., 1993. Practical Sampling Techniques for Infrared Analysis. CRC Press, Inc., Boca Raton, FL, 39-44.

Davies, A.M.C., Giangiacomo, R., 2000. Near Infrared Spectroscopy. Proc. 9th Internat. Conf.. NIR Publications, West Sussex, England.

Eberl, D.D., Smith, D.B., 2009. Mineralogy of soils from two continental-scale transects across the United States and Canada and its relation to soil geochemistry and climate. Appl. Geochem., this issue.

Garrett, R.G., 2009. Relative spatial soil geochemical variability along two transects across the United States and Canada. Appl. Geochem., this issue.

Griffiths, P.R., de Haseth, J.A., 2007. Diffuse Reflectance. In: Fourier Transform Infrared Spectroscopy, 2nd ed.. John Wiley & Sons, Inc., Hoboken, NJ, 349-362 (Chapter 16).

Janik, L.J., Skjemstad, J.O., 1995. Characterization and analysis of soils using mid-infrared partial least-squares. 2. Correlations with some laboratory data. Aust. J. Soil Res. 33, 637-650.

Janik, L.J., Merry, R.H., Skjemstad, J.O., 1998. Can mid-infrared diffuse reflectance analysis replace soil extractions? Aust. J. Exp. Agric. 38, 681-696.

Janik, L.J., Skjemstad, J.O., Raven, M.D., 1995. Characterization and analysis of soils using mid-infrared partial least-squares. 1. Correlations with XRF-determined major-element composition. Aust. J. Soil Res. 33, 621-636.

Janik, L.J., Skjemstad, J.O., Shepherd, K.D., Spouncer, L.R., 2007. The prediction of soil carbon fractions using mid-infrared-partial least-square analysis. Aust. J. Soil Res. 45, 73-81.

Lorber , A., Faber, K., Lowalski, B. R., 1998. Local centering in multivariate calibrations. J. Chemom. 10, 215-220.

Madari, B.E., Reeves, III, J.B., Coelho, M.R., Machado, P.L.O.A., De-Polli, H., Coelho, R.J., Benites, V.M., Souza, L.F., McCarty, G.W., 2005. Mid- and near-infrared spectroscopic determination of total and organic carbon in a diverse set of soils from the Brazilian national soil collection. Spectrosc. Lett. 38, 721-740.

Malley, D.F., Martin, P.D., Den-Dor, E., 2004. Application in Analysis of Soils. In: Roberts, C.A., Workman, Jr., J., Reeves, III, J.B. (Eds), Near-Infrared Spectroscopy in

Agriculture. Am. Soc. Agron., Madison, WI, 729-784 (Chapter 26).

McCarty, G.W., Reeves, III, J.B., 2001. Development of rapid instrumental methods for measuring soil organic carbon. In: Lal. R., Kimble, J.M., Follett, R.F., Stewart, B.A. (Eds), Assessment Methods for Soil Carbon. Lewis Publishers, Boca Raton, FL, 371-380.

McCarty, G.W., Reeves, III, J.B., Reeves, V.B., Follett, R.F., Kimble, J., 2002. Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. Soil Sci. Soc. Am. J. 66, 640-646.

Merry, R.H., Janik, L.J., 2001. Mid infrared spectroscopy for rapid and cheap analysis of soils. Proc. 10th Australian Agronomy Conference 2001. Available from: <http://www.regional.org.au/au/asa/2001/3/c/merry.htm>.

Naes, T., Isaksson, T., Fearn, T., Davies, T. 2002. A User Friendly Guide to Multivariate Calibration and Classification. NIR Publications, Chichester, West Sussex, UK.

Nguyen, T.T., Janik, L.J., Raupach, M., 1991. Diffuse reflectance infrared Fourier transform (DRIFT) spectroscopy in soil studies. Aust. J. Soil Res. 29, 49-67.

Norris, K.H., Barnes, R.F., Moore, J.E., Shenk, J.S., 1976. Predicting forage quality by infrared reflectance spectroscopy. J. Animal Sci. 43, 889-897.

PLSplus. 1994. PLSplus, V2.1G, Galactic Industries Corporation, Salem, NH.

Reeves, III, J.B., 1994. Near- versus mid-infrared diffuse reflectance spectroscopy for the quantitative determination of the composition of forages and by-products. J. Near Infrared Spectrosc. 2, 49-57.

Reeves, III, J.B., 1996. Improvement in Fourier near- and mid-infrared diffuse reflectance spectroscopic calibrations through the use of a sample transport device. Appl. Spectrosc. 50, 965-969.

Reeves, III, J.B., 2003. Mid-infrared diffuse reflectance spectroscopy: Is sample dilution with KBr neccesary, and if so, when? Am. Lab. 35, 24-28.

Reeves, III, J.B., Delwiche, S.R., 2003. SAS partial least squares regression for analysis of spectroscopic data. J. Near Infrared Spectrosc. 11, 415-431.

Reeves, III, J.B., Delwiche, S.R., 2004. Using SAS for PLS calibrations of spectroscopic data. NIRS News 15,10-15.

Reeves, III, J.B., McCarty, G.W., 2001 Quantitative analysis of agricultural soils using near infrared reflectance spectroscopy and a fiber-optic probe. J. Near Infrared Spectrosc. 9, 25-43.

Reeves, III, J.B., Francis, B.A., Hamilton, S.K., 2005. Specular reflection and diffuse reflectance spectroscopy of soils. Appl. Spectrosc. 59, 39-46.

Reeves, III, J.B., Galletti, G.C., Buta, G., Zadrazil, F., 1990. Near and mid-infrared spectroscopies and wet chemistry as tools for the study of treated feedstuffs. In: Coughlan, M.P., Amaral Collaço, M.T. (Eds), Advances in Biological Treatment of Lignocellulosic Materials. Elsevier Applied Science, Essex, 89-106.

Reeves, III, J.B., McCarty, G.W., Meisenger, J.J., 1999a. Near infrared reflectance spectroscopy for the analysis of agricultural soils. J. Near Infrared Spectrosc. 7, 179-193.

Reeves, III, J.B., McCarty, G.W., Mimmo, T.V., Reeves, V.B., Follett, R.F., Kimble, J.M., Galletti, G. C., 2002. Spectroscopic calibrations for the determination of carbon in soils. 17th World Congress of Soil Science, Thailand, Paper no. 398, 1-9.

Reeves, III, J.B., McCarty, G.W., Reeves, V.B., 2001. Mid-infrared diffuse reflectance spectroscopy for the quantitative analysis of agricultural soils. J. Agric. Food Chem. 49, 766-772.

Reeves, III, J.B., Zapf, C.M., Simovik, I., Delwiche, S.R., 1999b. Mid-infrared versus near-infrared diffuse reflectance spectroscopy for quantitative and qualitative analysis of agricultural products. Recent Res. Devel. Agric. Food Chem. 3, 201-222.

Roberts, C., Workman, J.J., Reeves, III, J.B. (Eds), 2004. NIR in agriculture. Am. Soc. Agronomy-Crop Sci. Soc. Am.-Soil Sci. Soc. Am., Madision, WI.

Sankey, J.B., Brown, D.J., Bernard, M.L., Lawrence, R.L. 2008. Comparing local vs. global visible and near-infrared (VisNIR) diffuse reflectance spectroscopy (DRS) calibrations for the prediction of soil clay, organic C and inorganic C. Geoderma 148, 149-158.

SAS, 2003. SAS/STAT User's Guide, Release 9.1. SAS Institute, Cary, NC.

Siebielec, G.W., McCarty, G.W., Stuczynski, T.I., Reeves, III, J.B., 2004. Near- and mid-infrared diffuse reflectance spectroscopy for measuring soil metal content. J. Environ. Qual. 33, 2056-2069.

Smith, D.B., Cannon, W.F., Woodruff, L.G., Garrett, R.G., Klassen, R., Kilburn, J.E., Horton, J.D., King, H.D., Goldhaber, M.B., Morrison, J.M. 2005. Major- and trace-element concentrations in soils from two continental-scale transects of the United States and Canada. U.S. Geol. Surv. Open-File Rep. 2005-1253. Available from: <http://pubs.usgs.gov/of/2005/1253/>.

Smith, D.B. Woodruff, L.G., O'Leary, R.M., Cannon, W.F., Garrett, R.G., Kilburn, J.E., Goldhaber, M.B., 2009. Pilot studies for the North American Soil Geochemical Landscapes Project – site selection, sampling protocols, analytical methods, and quality control protocols. Appl. Geochem., this issue.

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O., 2006. Visible, near infrared, mid-infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 131, 59-75.

Westerhaus, M., Workman, J., Reeves, III, J.B., Mark, H., 2004. Quantitative analysis. In: Roberts, C.A., Workman, Jr., J., Reeves, III, J.B. (Eds), Near-Infrared Spectroscopy in Agriculture. Am. Soc. Agron., Madison, WI, 133-174 (Chapter 7).

Williams, P.C., 1975. Application of near infrared reflectance spectroscopy to analysis of cereal grains and oilseeds. Cereal Chem. 52, 561-576.

Williams. P., Norris, K. (Eds), 2001. Near-infrared technology in the agricultural and food industries, 2nd ed. Am. Assoc. Cereal Chemists, Inc. St. Paul, MN.

Woodruff, L. G., Cannon, W.F., Eberl, D.D., Smith, D.B., Kilburn, J.E., Horton, J.D., Garrett, R.G., Klassen, R., 2009. Continental-scale patterns in soil geochemistry and mineralogy: Results from two transects across the United States and Canada. Appl. Geochem., this issue.

Wu, Y., Chen, J., Wu, X., Tian, Q., Junfeng, J., Qin, Z., 2005. Possibilities of reflectance spectroscopy for the assessment of contaminant elements in suburban soils. Appl. Geochem. 20, 1051-1059.

Table 1. Summary statistics for 720 soils studied.

Analyte	Mean	Min.	Max	SD[a]	Skew	Kurtosis
Al	5.0	0.13	11.2	1.8	0.21	0.07
As	5.7	0.5	23	3.2	1.12	2.63
Ba	540	20	3670	240	3.36	39.16
Be	1.4	0.05	4.3	0.6	0.94	1.51
Bi	0.2	0.02	0.9	0.1	2.89	15.91
Ca	2.4	0.02	23.7	3.6	2.89	9.86
$C_{carbonate}$	0.6	<0.003	10.4	1.3	3.39	13.91
$C_{organic}$[b]	1.8	0.04	34.2	2.6	5.56	53.55
C_{total}	2.4	0.04	34.2	2.7	4.64	38.92
Cd	0.3	0.05	8.2	0.6	9.45	105.00
Ce	50	1	204	24	1.20	4.57
Co	9.5	0.7	190	15	8.92	92.03
Cr	73	0.5	6030	440	11.19	127.86
Cs	3.3	0.1	15	1.8	1.37	4.74
Cu	15	0.2	103	10	3.47	19.63
Fe	2.2	0.06	11.7	1.4	2.50	9.63
Ga	11	0.3	35	4.3	0.69	1.44
Hg	0.03	0.01	0.7	0.03	11.19	205.91
In	0.04	0.01	0.1	0.02	0.53	0.67
K	1.7	0.06	4.4	0.7	0.45	0.63
La	27	0.6	79	12	0.84	2.33
Li	23	4	260	16	5.73	65.11
Mg	0.9	0.04	18.4	1.8	6.64	52.54
Mn	590	22	4190	480	2.70	11.27
Mo	1.1	0.08	21	1.6	7.97	81.40
Na	0.8	0.01	3.2	0.6	0.99	0.91
Nb	7	0.3	42	4	2.33	13.03
Ni	41	1.6	3450	240	11.18	128.21
P	630	69	22640	1120	13.58	234.77
Pb	20	0.2	320	19	9.39	117.74
Rb	70	2.4	190	28	0.91	2.34
S	0.10	0.01	13	0.7	12.77	176.40
Sb[c]	0.6	0.05	2.4	0.4	1.75	4.29
Sb[d]	0.4	0.3	1.5	0.2	2.84	7.64
Sc	7	0.2	48	4	3.83	28.56
Se	0.3	0.1	3.7	0.4	3.53	20.00
Sn	1.4	0.2	8.6	0.7	2.65	19.01
Sr	200	11	5630	260	13.34	255.71
Te	0.05	0.05	0.7	0.03	16.62	341.92
Th	8.8	0.2	31.4	4.4	0.94	1.93
Ti	0.2	0.01	1.0	0.1	2.69	13.53
Tl	0.5	0.05	1.8	0.2	1.29	6.54
T-S	0.09	0.03	11.3	0.6	13.02	182.96
U	2.2	0.4	9.7	1.1	1.14	4.25
V	64	6	430	41	3.30	20.12
W	0.8	0.05	10	0.8	6.92	67.83
Y	14	0.4	91	7	2.72	22.73
Zn	58	3	433	34	3.71	33.32

All concentrations in mg/kg except for Al, $C_{carbonate}$, Ca, Fe, K, Mg, Na, S (ICP-MS), Ti, C_{total}
    and T-S (combustion) in %.
[a] SD, standard deviation; [b] $C_{organic}$ = C_{total} − $C_{carbonate}$, [c] ICP-MS, [d] HG-AAS

Table 2. Cross validation results using all 720 samples.

Analyte	Mid-Infrared				Near-Infrared		
	R^2	RMSD[a]	RPD[b]		R^2	RMSD	RPD
Al	0.782	0.828	2.1		0.515	1.235	1.4
As	0.587	2.076	1.6		0.346	2.613	1.2
Ba	0.438	181.086	1.3		0.427	182.930	1.3
Be	0.668	0.373	1.7		0.472	0.470	1.4
Bi	0.421	0.076	1.3		0.165	0.091	1.1
Ca	0.860	1.333	2.7		0.807	1.566	2.3
$C_{carbonate}$	0.933	0.326	3.9		0.832	0.514	2.5
$C_{organic}$[c]	0.581	1.656	1.5		0.534	1.745	1.5
C_{total}	0.601	1.730	1.6		0.632	1.662	1.6
Cd	0.470	0.426	1.4		0.500	0.414	1.4
Ce	0.654	14.025	1.7		0.519	16.538	1.4
Co	0.956	3.058	4.8		0.910	4.387	3.3
Cr	0.972	72.294	6.0		0.997	24.540	17.8
Cs	0.712	0.953	1.9		0.465	1.300	1.4
Cu	0.559	6.892	1.5		0.358	8.311	1.2
Fe	0.816	0.581	2.3		0.594	0.862	1.6
Ga	0.805	1.918	2.3		0.528	2.982	1.5
Hg	0.109	0.033	0.9		0.141	0.032	0.9
In	0.705	0.010	2.0		0.415	0.014	1.4
K	0.738	0.352	2.0		0.624	0.422	1.6
La	0.679	6.697	1.8		0.429	8.428	1.4
Li	0.558	10.909	1.5		0.320	13.594	1.2
Mg	0.978	0.263	6.7		0.937	0.444	4.0
Mn	0.570	315.737	1.5		0.370	382.445	1.3
Mo	0.424	1.195	1.3		0.199	1.409	1.1
Na	0.822	0.246	2.4		0.733	0.302	1.9
Nb	0.651	2.391	1.7		0.533	2.764	1.5
Ni	0.996	15.701	15.5		0.994	18.112	13.4
P	0.850	432.120	2.6		0.099	1059.09	1.1
Pb	0.094	18.271	1.1		0.124	17.966	1.1
Rb	0.720	14.870	1.9		0.494	19.985	1.4
S	0.597	0.469	1.6		0.646	0.439	1.7
Sb[d]	0.552	0.262	1.5		0.515	0.273	1.4
Sb[e]	0.240	0.178	1.1		0.262	0.176	1.1
Sc	0.648	2.555	1.7		0.431	3.248	1.3
Se	0.467	0.258	1.4		0.293	0.297	1.2
Sn	0.465	0.517	1.4		0.187	0.637	1.1
Sr	0.886	89.003	3.0		0.898	84.124	3.1
Te	0.021	0.029	1.0		0.074	0.028	1.1
Th	0.705	2.372	1.8		0.348	3.528	1.2
Ti	0.736	0.059	2.0		0.489	0.085	1.4
Tl	0.565	0.130	1.5		0.393	0.154	1.3
T-S	0.635	0.393	1.7		0.683	0.366	1.8
U	0.611	0.674	1.6		0.270	0.923	1.2
V	0.502	29.064	1.4		0.283	34.860	1.2

W	0.219	0.689	1.1		0.122	0.730	1.1
Y	0.730	3.756	1.9		0.382	5.680	1.3
Zn	0.292	28.308	1.2		0.155	30.930	1.1

[a]RMSD = Root mean square deviation = $[\Sigma \text{ (Actual-Predicted)}^2/n]^{1/2}$; all concentrations in mg/kg except for Al, $C_{carbonate}$, Ca, Fe, K, Mg, Na, S (ICP-MS), Ti, C_{total} and T-S (combustion) in %.

[b]RPD = SD/RMSD, [c] $C_{organic} = C_{total} - C_{carbonate}$, [d]ICP-MS, [e]HG-AAS.

Table 3. Calibration results using an independent test set based on cluster analysis.

| Analyte | Mid-Infrared Calibration Results | | | | | Near-Infrared Calibration Results | | | | |
| | Calibration Set | | Validation Set | | | Calibration Set | | Validation Set | | |
	R^2	RMSD	R^2	RMSD[a]	RPD[b]	R^2	RMSD	R^2	RMSD	RPD
Al	0.822	0.855	0.649	0.865	1.7	0.575	1.320	0.304	1.272	1.1
Be	0.683	0.425	0.568	0.342	1.5	0.521	0.522	0.289	0.462	1.1
$C_{carbonate}$	0.906	0.500	0.590	0.257	1.4	0.805	0.721	0.255	0.487	0.7
$C_{organic}$[c]	0.693	1.519	0.353	1.904	1.2	0.578	1.783	0.335	1.943	1.2
C_{total}	0.790	1.393	0.302	1.964	1.2	0.509	2.131	0.354	1.913	1.2
Ce	0.578	18.037	0.543	12.905	1.5	0.237	24.269	0.215	16.898	1.1
Co	0.902	4.786	0.905	5.846	2.4	0.907	4.661	0.842	6.165	2.3
Cr	0.989	57.208	0.981	100.530	3.0	0.996	36.097	0.974	84.275	3.5
Cs	0.690	1.171	0.512	1.015	1.4	0.487	1.505	0.211	1.309	1.1
Fe	0.881	0.535	0.665	0.643	1.7	0.591	0.993	0.379	0.871	1.3
Ga	0.824	2.108	0.663	1.997	1.7	0.532	3.438	0.301	2.999	1.1
In	0.795	0.009	0.550	0.010	1.5	0.388	0.016	0.083	0.015	1.0
K	0.758	0.390	0.640	0.341	1.7	0.555	0.528	0.504	0.433	1.3
La	0.642	8.300	0.526	6.456	1.4	0.312	11.503	0.237	8.173	1.1
Mg	0.978	0.310	0.962	0.253	5.1	0.883	0.718	0.924	0.375	3.4
Na	0.853	0.261	0.677	0.255	1.7	0.785	0.316	0.498	0.356	1.2
Ni	0.982	34.230	0.981	60.543	3.9	0.982	33.624	0.968	73.496	3.2
Th	0.666	2.958	0.573	2.262	1.5	0.394	3.986	0.229	3.124	1.1

[a]RMSD = Root mean square deviation = $[\Sigma \, (Actual-Predicted)^2/n]^{1/2}$; all concentrations in mg/kg except for Al, $C_{carbonate}$, Fe, K, Mg, Na, and C_{total} in %.
[b]RPD = SD/RMSD, [c] $C_{organic} = C_{total} - C_{carbonate}$

Figure captions

Figure 1. Mid-infrared spectra of samples with highest (solid line) and lowest (dashed) concentrations of C_{total}.

Figure 2. Near-infrared spectra of samples with highest (solid line) and lowest (dashed) concentrations of C_{total}.

Figure 3. Mid-infrared calibration for C_{total} concentrations based on all 720 samples (Table 2).

Figure 4. Mid-infrared calibration for $C_{carbonate}$ concentrations based on all 720 samples (Table 2).

Figure 5. Mid-infrared calibration for Mg concentrations based on all 720 samples (Table 2).

Figure 6. Mid-infrared calibration for Ni concentrations based on all 720 samples (Table 2).

Figure 7. Mid-infrared calibration for Ni concentrations based on all 720 samples (Table 2), but showing only results for samples with Ni <100 mg/kg (ppm).

Figure 8. Mid-infrared validation results for $C_{carbonate}$ (shown on the graph as Total Carbonate; n=360) using calibration based on spectral differences (Table 3).

Figure 9. Mid-infrared validation results for $C_{carbonate}$ (n=360, samples divided based on even distribution of $C_{carbonate}$ concentration in calibration and validation sets).

Figure 10. Mid-infrared validation results for $C_{carbonate}$ (n=327, based on $C_{carbonate}$ concentration and limited to concentrations ≤2%).

Figure 11. Mid-infrared calibration results for C_{total} (n=328, based on C_{total} concentration and limited to concentrations ≤5%).

Figure 12. Mid-infrared validation results for C_{total} (n=328, based on C_{total} concentration and limited to concentrations ≤5%).

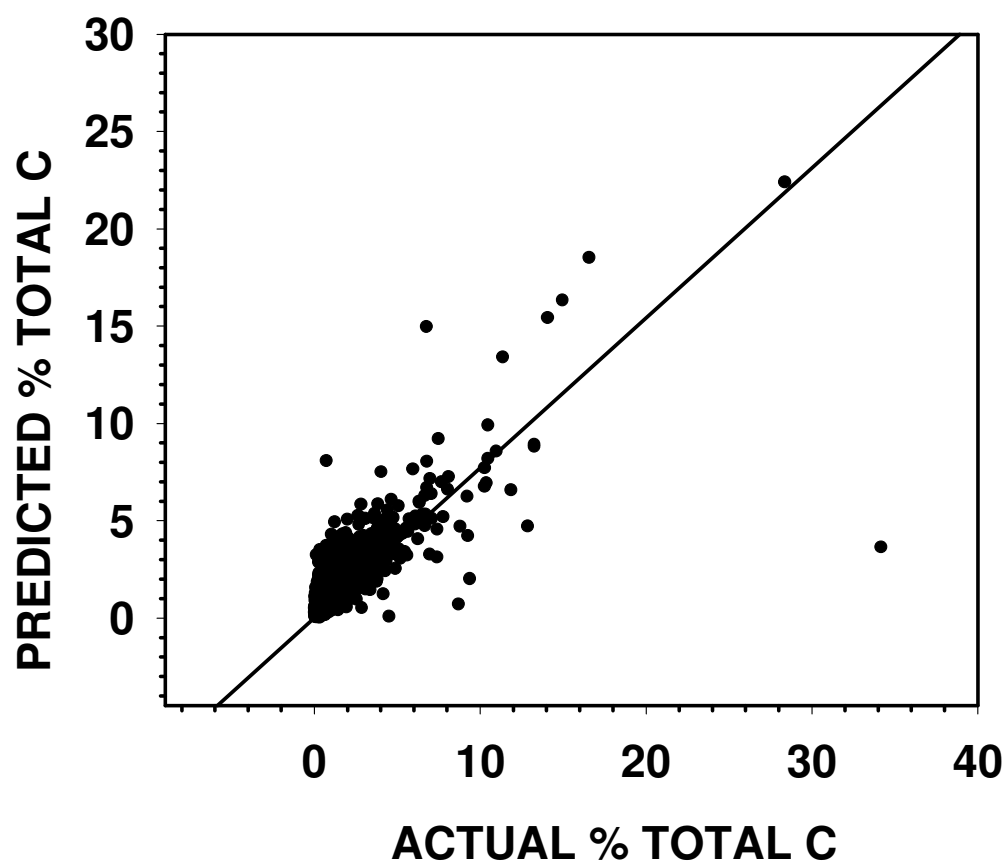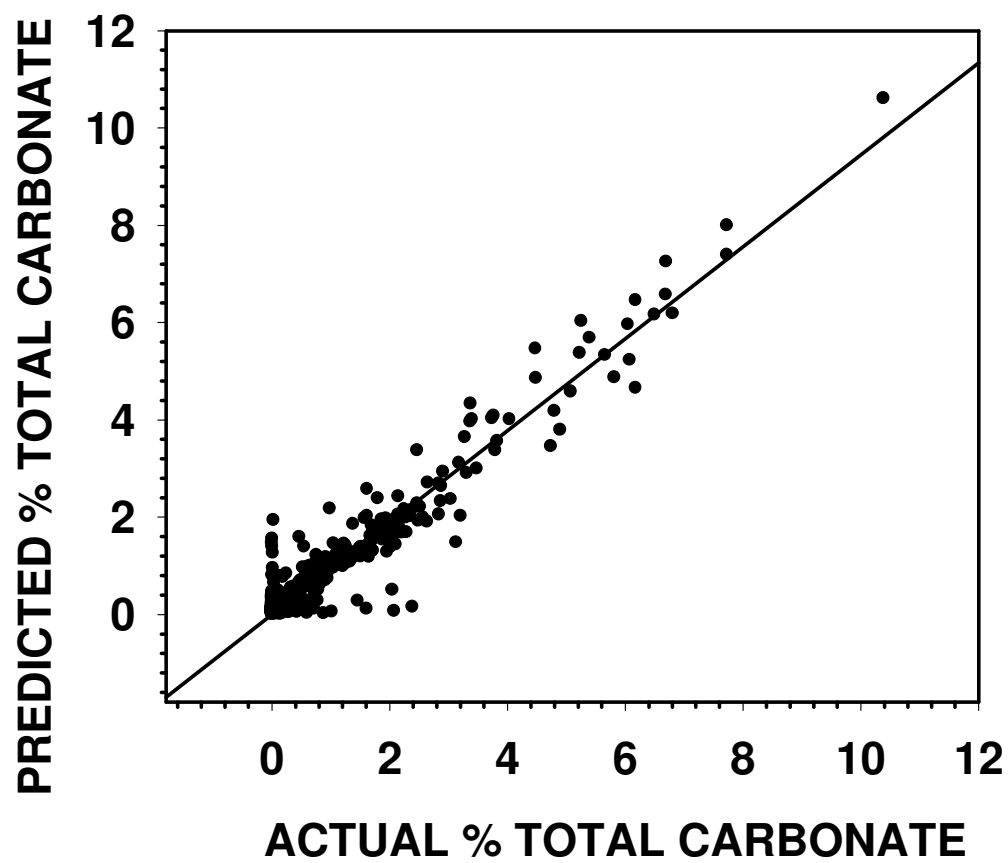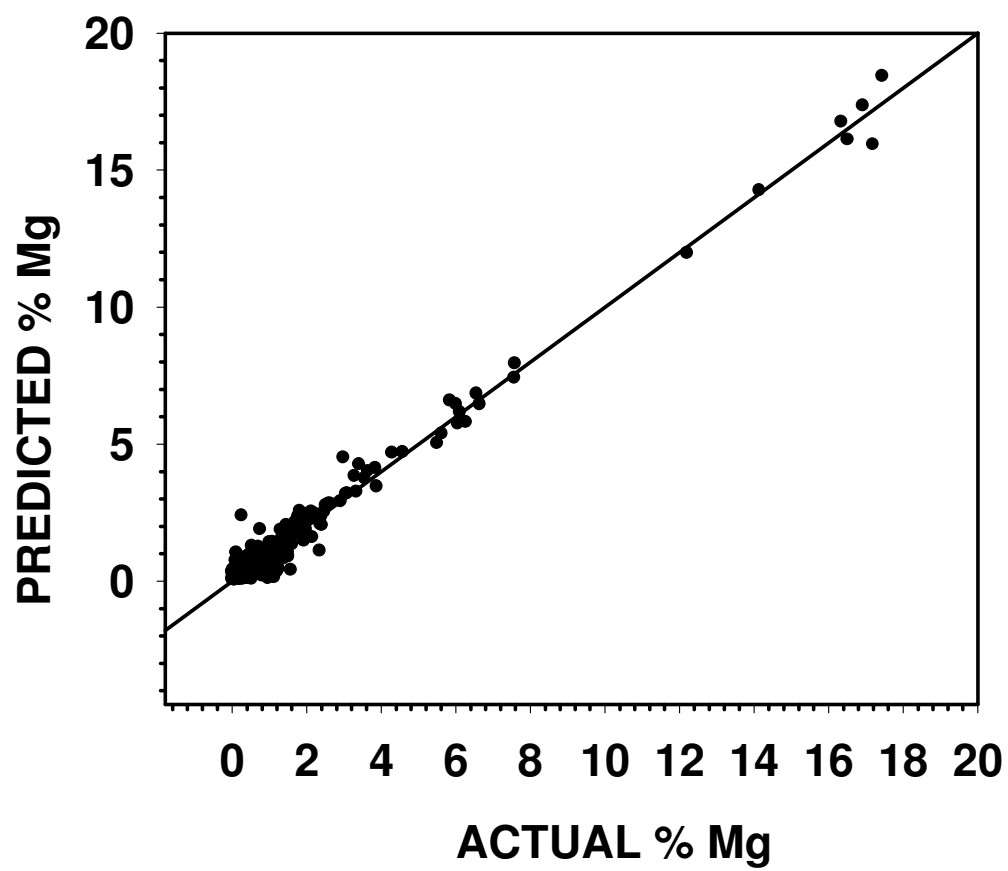**Figure 13**. Mid-infrared validation results for Al (n=360) using calibration based on spectral differences (Table 3).
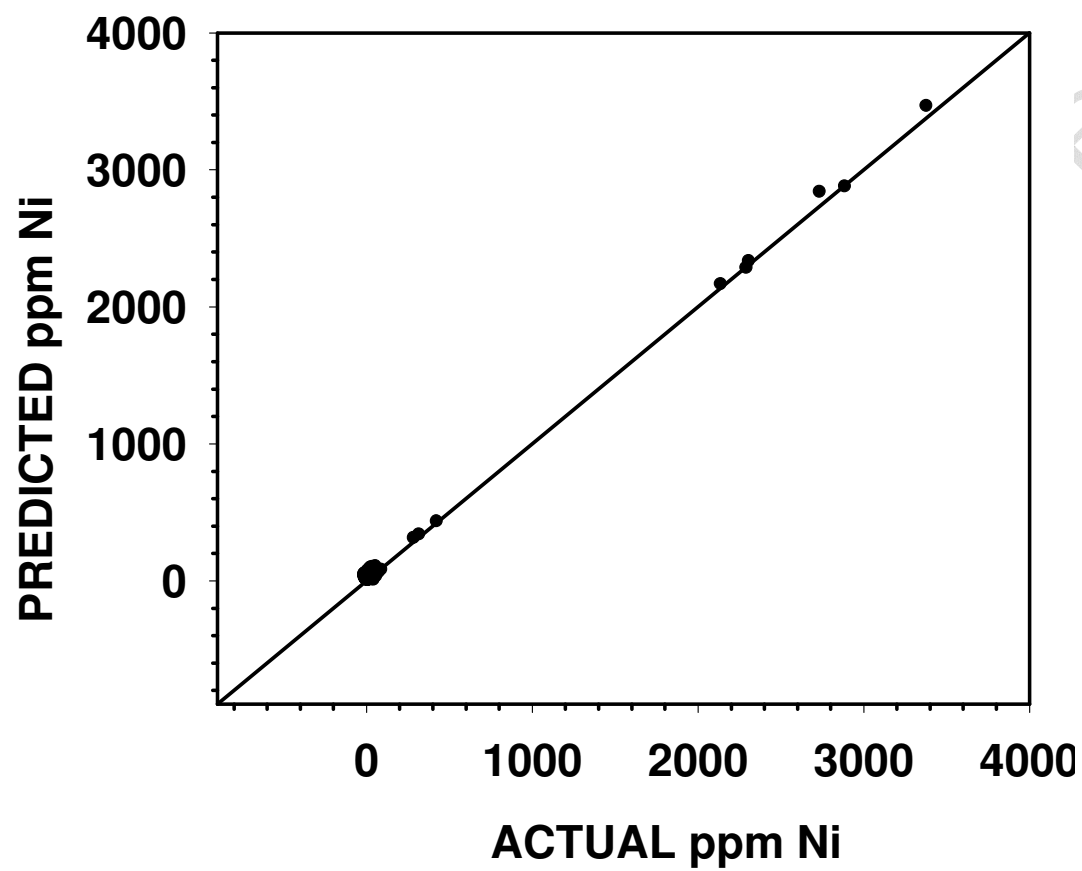
Fig. 1
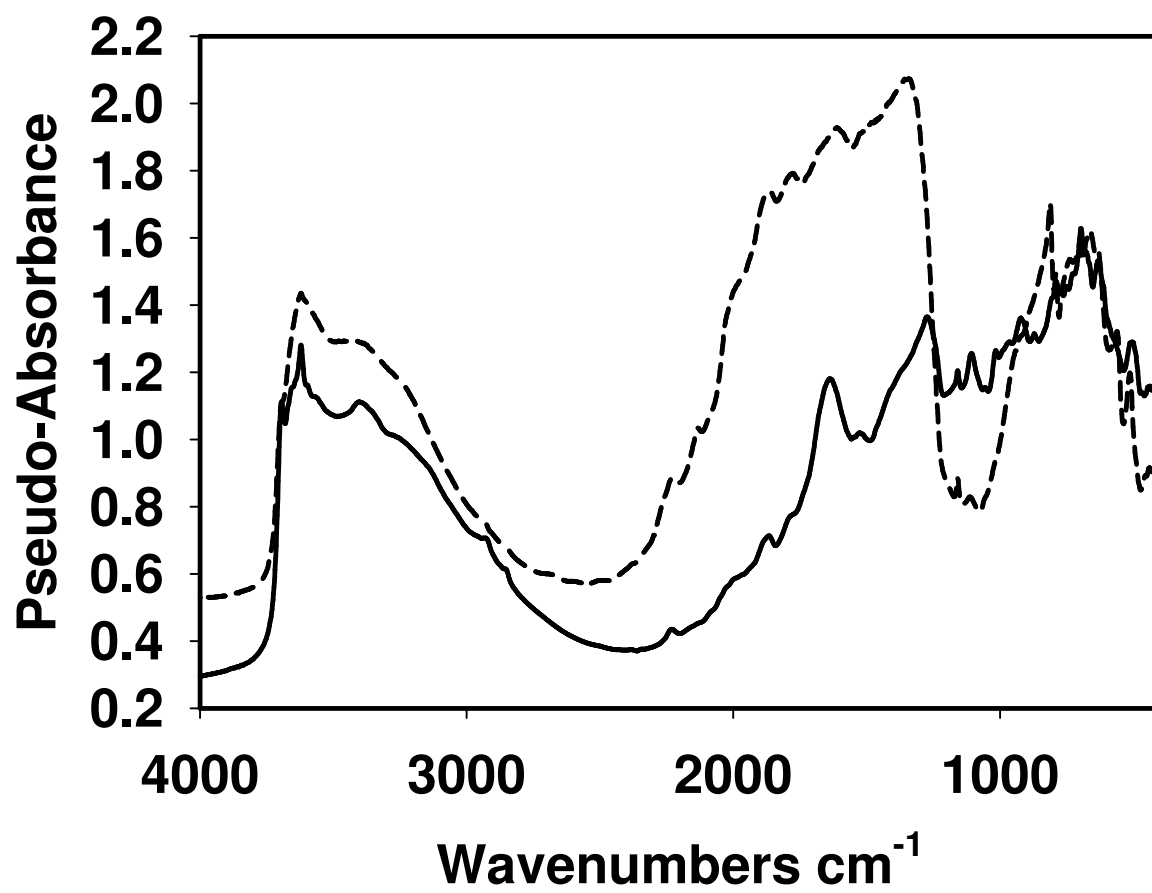
Fig. 2
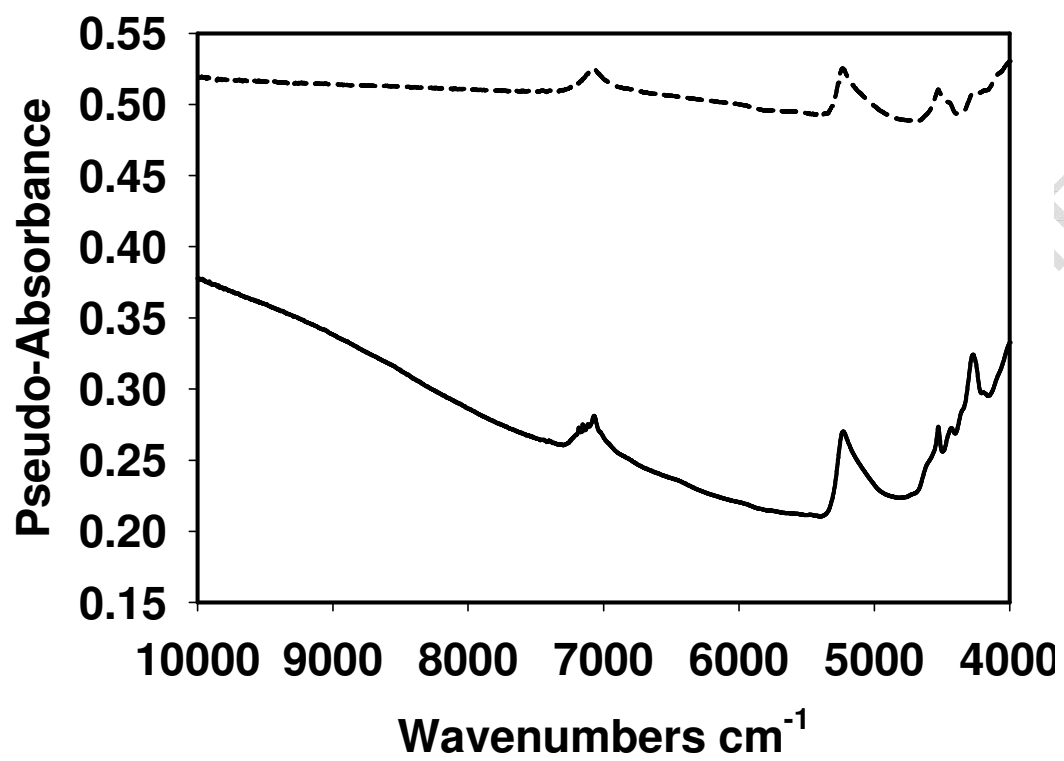
Fig. 3.

Fig. 4

Fig. 5

Fig. 6

Fig. 7

Fig. 8

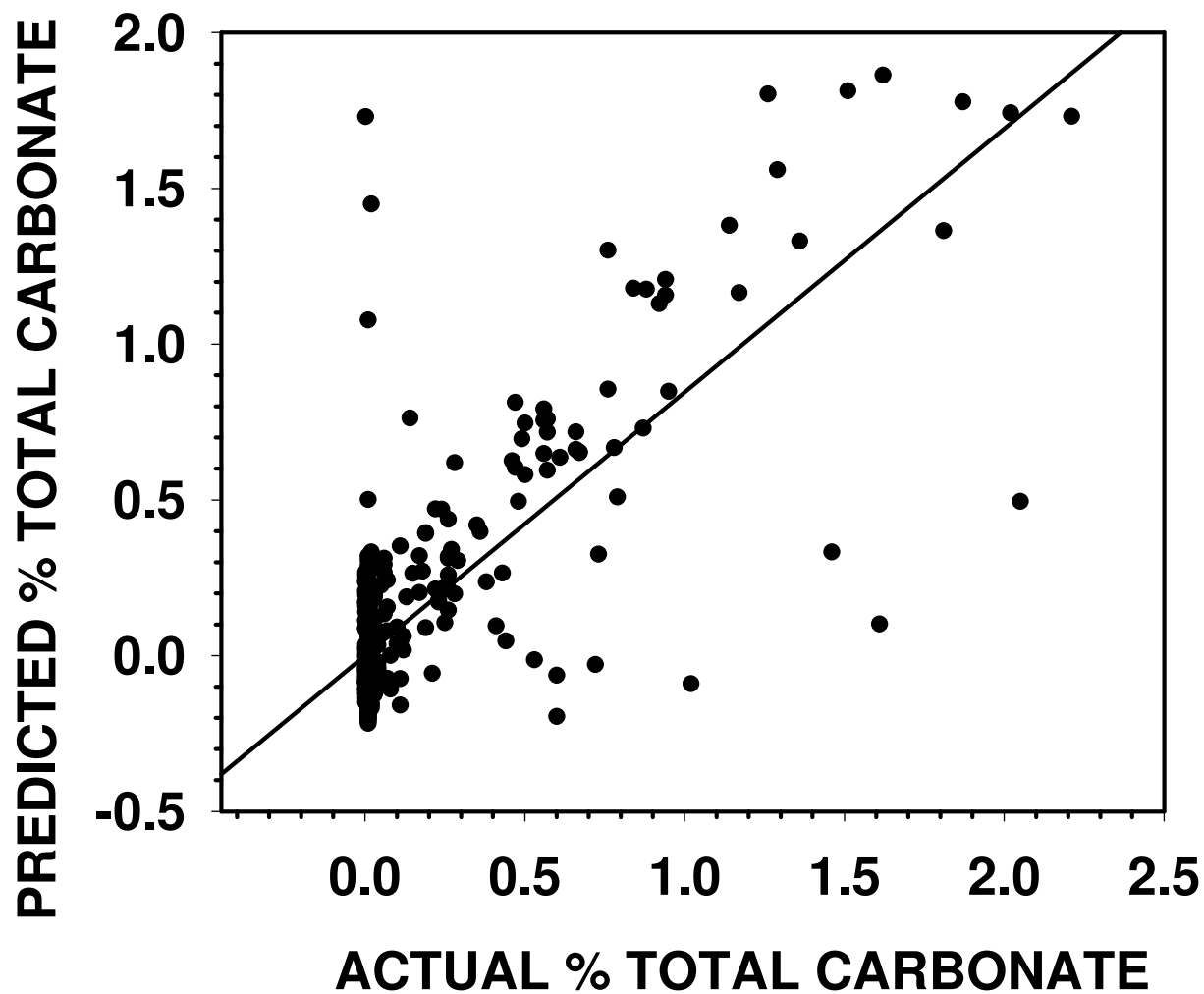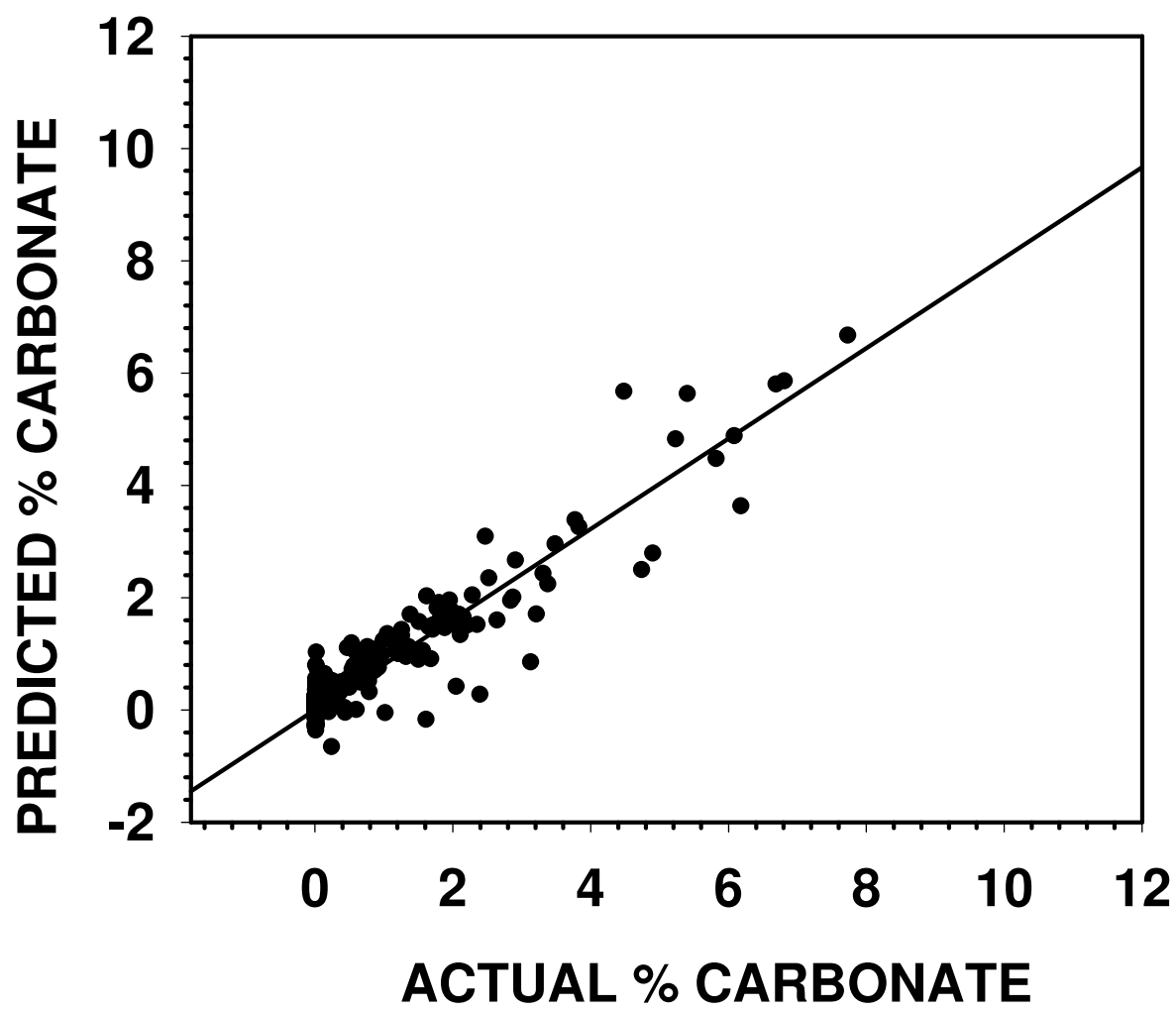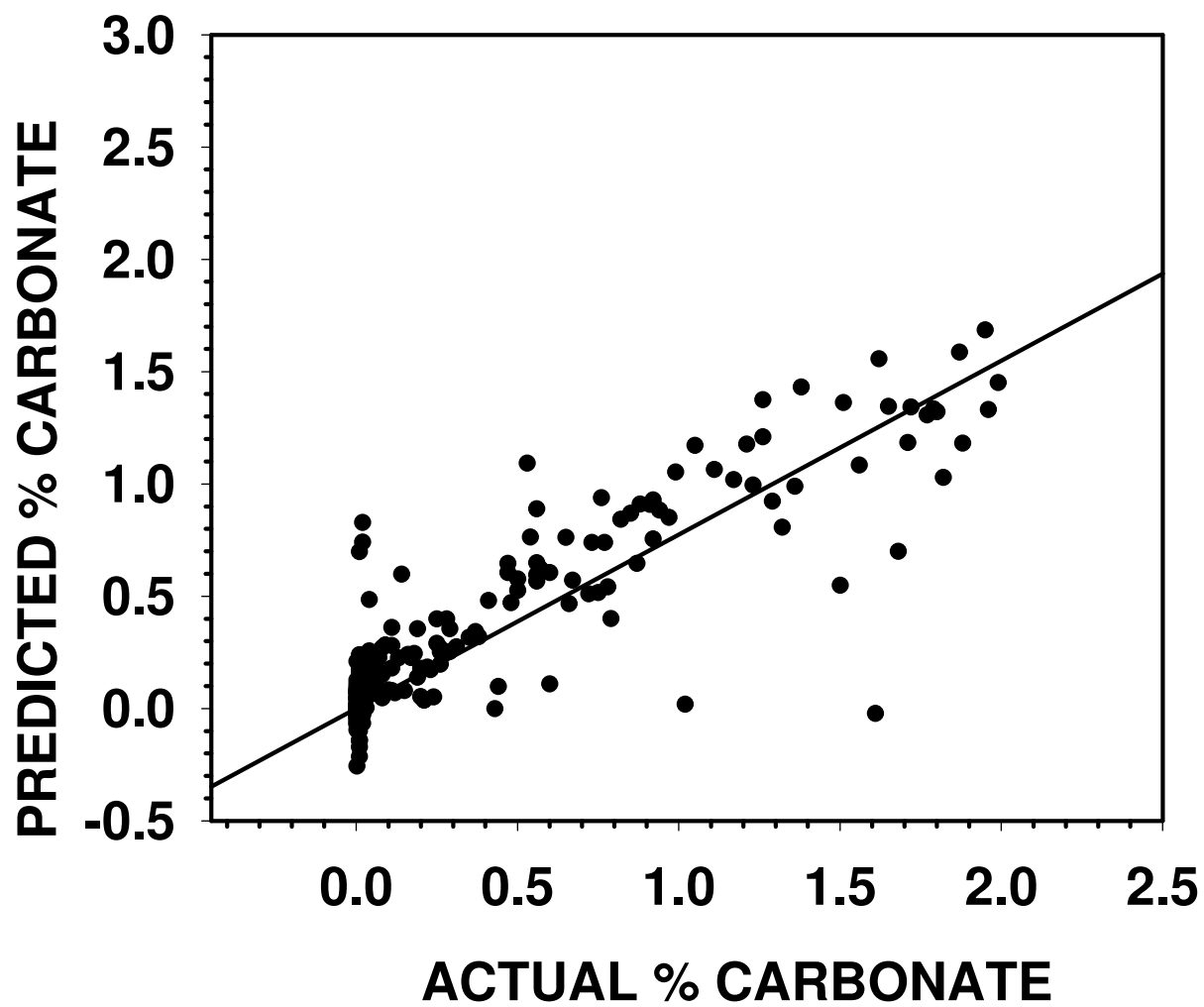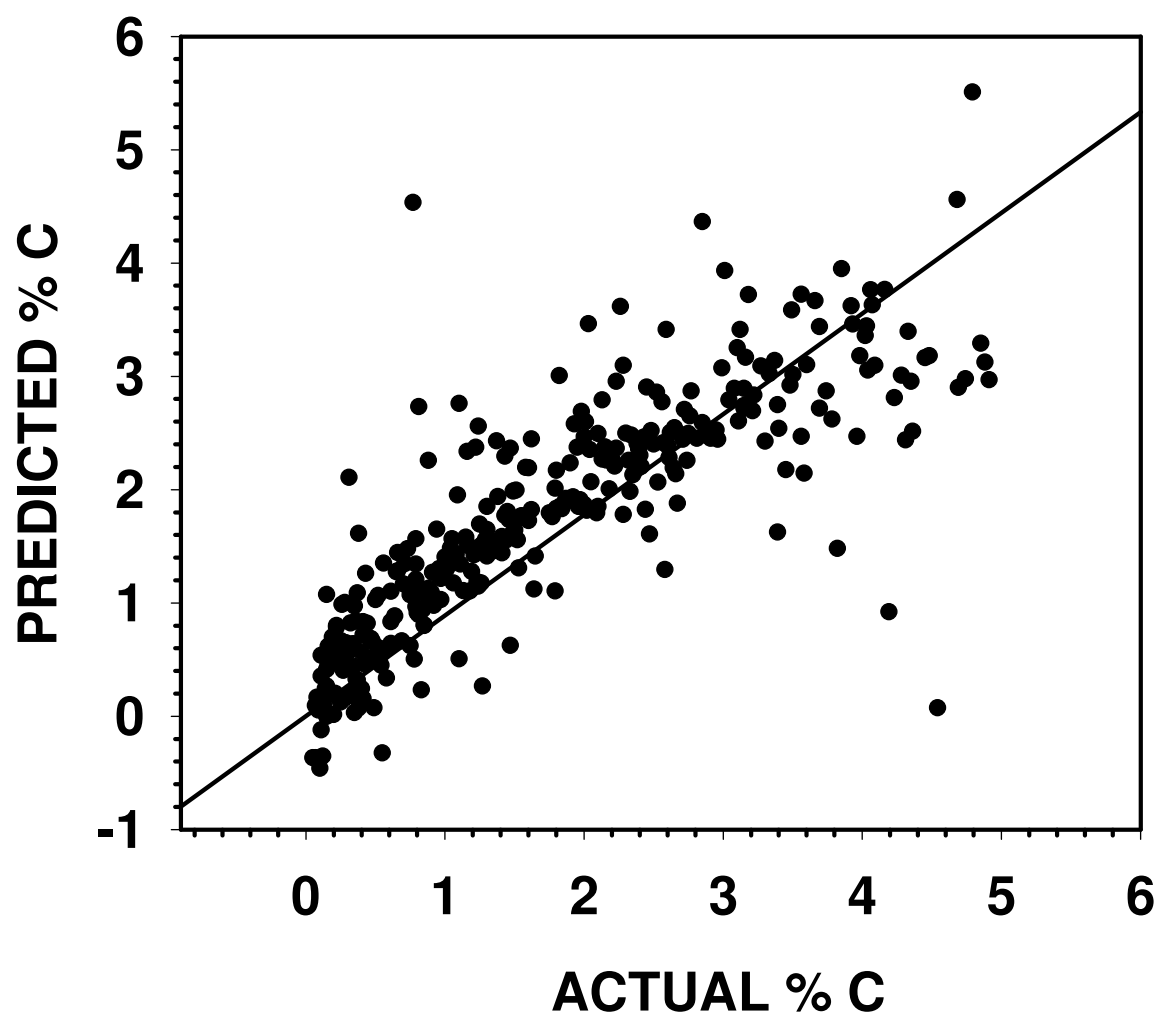Fig. 9

Fig. 10

Fig. 11

Fig. 12

Fig. 13