Foliar analysis using near infrared reflectance spectroscopy

CAROL A. WESSMAN
Department of Forestry and The Environmental Remote Sensing Center,
University of Wisconsin, Madison, WI, U.S.A. 53706

JOHN D. ABER
Department of Forestry, University of Wisconsin, Madison, WI, U.S.A. 53706

DAVID L. PETERSON
Ames Research Center, National Aeronautics and Space Administration, Moffett Field, CA, U.S.A. 94035

AND

JERRY M. MELILLO
Ecosystems Center, Marine Biological Laboratory, Woods Hole, MA, U.S.A. 02543

Received February 24, 1987
Accepted August 20, 1987


Near infrared reflectance spectroscopy was evaluated as a method for measuring nitrogen and lignin content in foliage of native forest and prairie species. Near infrared reflectance spectra (1590 to 2357 nm) were obtained for 163 samples of dried green leaves and leaf litter from 18 deciduous and 2 coniferous tree species. Forty additional spectra were obtained from grass species. Reflectance (R) spectra were recorded as log (1/R) and transformed to the first and second derivative of log (1/R). Multiple linear regressions, predicting wet chemistry values based on near infrared reflectance spectra, yielded correlation coefficients of 0.98 for Kjeldahl nitrogen and 0.78 for lignin, with standard errors of 0.11% for nitrogen and 2.9% for lignin. Results suggest that near infrared reflectance spectroscopy is very effective for rapid (approximately 2 min per sample) determination of foliar lignin and nitrogen and should be considered for use as a routine analytical method.


La spectroscopie par réflectance de l'infrarouge proche comme méthode pour mesurer le contenu en azote et en lignez dans le feuillage des espèces indigènes de la forêt et de la prairie a été évaluée. Les spectres de réflectance de l'infrarouge proche (1590 à 2357 nm) ont été obtenus pour 163 échantillons de feuilles vertes séchées et de litière de feuilles provenant de 18 espèces feuillues et de 2 conifères. Quarante spectres supplémentaires ont été obtenus d’espèces herbacées. Les spectres de réflectance (R) ont été enregistrés comme log (1/R) et transformés en première et deuxième dérivées de log (1/R). Des régressions linéaires multiples, prédisant les valeurs chimiques humides à partir des spectres de l'infrarouge proche par réflectance, ont donné des coefficients de corrélation de 0,98 pour l'azote Kjeldahl et de 0,78 pour la lignez, avec des écarts-types de 0,11% pour l'azote et de 2,9% pour la lignez. Ces résultats indiquent que la spectroscopie par réflectance de l'infrarouge est très efficace pour déterminer rapidement (environ 2 min par échantillon) la lignez et l'azote foliaire et que l'on devrait prendre en considération son emploi comme méthode analytique régulière.

[Traduit par la revue]

Introduction

Several types of research in forest ecosystems require measurements of chemical content of large numbers of samples. Foliar nitrogen content is needed for determination of fertilizer prescriptions (Leaf 1973; Brix 1981), ecosystem studies of nitrogen cycling (Bormann et al. 1977; Pastor et al. 1984) and effects of disturbance on element cycling and loss from the ecosystem (Vitousek et al. 1981). Carbon fraction chemistry (e.g., lignin content) is important in studies of leaf litter decomposition and nutrient cycling rates (Melillo et al. 1982; Berg and Ekbohm 1983). Although chemical analysis of plant tissue is fairly standardized, the expense in both time and money can severely restrict sample size. Some constituents, such as lignin, require complicated laboratory procedures and are consequently costly to measure. Moreover, important lag times may occur between collections of samples and acquisition of laboratory results.

Leaf reflectance presents a rapid, nondestructive means to measure vegetation properties and condition. A plant leaf typically has low reflectance in the visible region of the spectrum as a result of strong absorption by chlorophyll and has higher reflectance in the near infrared region, up to 1300 nm, as a result of internal leaf scattering and reduced absorption (Knipping 1970). Beyond 1300 nm, strong absorption by water reduces reflectance considerably. Field use of spectroradiometers (400 to 1100 nm) has met with some success for measurements of biomass, production, and stress detection (Press 1974; Tucker 1977). The relationship between leaf reflectance in the visible region and leaf chlorophyll and nitrogen concentrations has been demonstrated by a number of researchers (Thomas and Oerther 1972; Everitt et al. 1985). Tsay et al. (1982) estimated nitrogen content of lobolly needles to within 0.06% of Kjeldahl-determined values using leaf reflectance.

Near infrared reflectance spectroscopy (NIRS) makes use of the diffuse reflectance of dried, ground samples. Each constituent (e.g., cellulose, protein) of a complex organic

---

1Author to whom correspondence should be sent at the following address: Cooperative Institute for Research in Environmental Sciences (CIRES), Campus Box 449, University of Colorado, Boulder, CO, U.S.A. 80309.
mixture has unique absorption properties in the near infrared (NIR) region (700 to 2500 nm) of the spectrum. These properties are the result of stretching and bending vibrations of strong molecular bonds between light atoms; e.g., hydrogen bonds associated with carbon, oxygen, and nitrogen. The fundamental absorption band is located at the energy level (or wavelength) that allows the molecule to rise to higher vibrational states. These bands occur in the midinfrared region and their overtones (½, ½, ¼, ..., of the wavelength of the fundamental band) occur in the near infrared region. Overtones are specific to each constituent and are, with combination bands (sums and differences between overtones of different constituents), more sensitive to changes in the chemical environment of the absorbing molecules than the fundamental of the same vibration (Wetzel 1983). Slight disturbances in the bonding scheme will produce only small changes in the fundamental while the near infrared region will experience large frequency shifts and amplitude changes. Consequently, these properties of the NIR spectra can be correlated to changes in the chemical composition of a sample measured by wet chemical means.

Measurement of a sample’s diffuse reflectance, the sum of NIR absorption properties and the scattering characteristics of the sample, is rapid and requires little or no sample preparation. The sample’s weight is not required, no chemical reagents are necessary, and analytical time is short. Shenk et al. (1979) state that four criteria must be met for meaningful NIR predictions: (i) selection of calibration samples representative of the population to be predicted, (ii) accurate laboratory analysis of the calibration samples, (iii) choice of the correct mathematical treatment of the NIR data for optimum information extraction, and (iv) choice of wavelengths relevant to the total population of samples. In the simplest scenario, one measurement at an absorption maximum and another to serve as a reference of the overall reflectivity are required for analysis. Other wavelengths may be used to function as reference terms or to give positive or negative correlations to the constituent. Simultaneous estimation of two or more constituents requires multiple linear regression equations.

Over the past 20 years, extensive research has been performed on NIRS applications to quantitative agricultural work. NIRS techniques have been developed for predicting oil, protein, and moisture content of grains and oilseeds (Norris and Hart 1965; Ben-Gara and Norris 1968; Hymowitz et al. 1974), and for determination of protein, lignin, fiber fractions, and in vitro or in vivo digestible dry matter of forage (Norris et al. 1976; Shenk et al. 1981). In general, the analytical precision of the NIRS predictions approaches the repeatability range of standard wet chemical analyses. Winch and Major (1981) applied NIRS to prediction of Kjeldahl nitrogen in legumes and grasses, with standard errors of the NIR calibration (SEC) only slightly higher than those of the laboratory (0.08 – 0.11 vs. 0.08%N). Standard errors of prediction (SEP) ranged from 0.16 – 0.20%N. Shenk et al. (1981) reported SEP of 0.96% for crude protein, 1.13% for lignin, 1.27% for cellulose, 0.16% for calcium, 0.04% for phosphorus and 0.37% for potassium. Redshaw (1985) reported that calcium and phosphorus analyses of forages were fairly erratic, although NIRS analyses were successfully used to detect trends as a result of experimental treatments.

In conjunction with an ongoing study of nitrogen cycling in temperate forests, we examined the potential of NIRS for determining lignin and nitrogen content of foliage from native forest and prairie species. This report summarizes the results obtained using a commercially available NIR instrument to determine nitrogen and lignin values in tree leaves across several species, canopy heights, and site conditions, as well as nitrogen concentration in prairie grass species.

Materials and methods

Samples were taken from established field sites on Blackhawk Island, Wisconsin, located in the Wisconsin River north of Wisconsin Dells (43°40' N, 89°45' W) (Pastor et al. 1984) and in the University of Wisconsin – Madison Arboretum (43°3' N, 89°25' W) (Nadelhoffer et al. 1983). Green leaves were collected in August 1984 and July 1985 from four canopy positions for each major species at a site: emergent, sunlit codominant, shaded codominant, and suppressed. Collections of leaf litter fall were made weekly from October to November 1984 using five 0.25 m² traps at each site, sorted by species and pooled for all dates.

From a total of 203 samples, 163 represented 20 forest tree species from 19 sites (both deciduous and conifer). Of those 163 samples, 83 were dried green leaves and 80 samples were dried brown litter. Forty additional green samples represented prairie grasses from four sites in the Arboretum Curtis Prairie.

All samples were dried at 70°C and ground in a Wiley mill through a 1-mm mesh. Percent nitrogen was determined for all samples by the micro-Kjeldahl method. Percent lignin was measured as the fraction insoluble in sulfuric acid following extraction in polar and nonpolar solution (Effland et al. 1977; McClaugherty et al. 1985). Only 132 (all forest species) of the 203 samples were analyzed for lignin.

Consistency is crucial to the successful use of NIRS and all samples should be prepared using the same technique. After drying, samples are left to equilibrate to atmospheric temperature and humidity (Abrams 1985). Samples should be stored in sealed containers to prevent moisture changes if there is a long period between drying and NIRS analysis.

Reflectance measurements were made with a Pacific Scientific Neotec 51A Scanning Filter Instrument owned by the University of Wisconsin Soil and Plant Analysis Laboratory. The instrument houses six interference filters, each of which transmits radiation over a limited wavelength region as it is tilted away from the perpendicular of the incident light source. Spectral data are collected at 1.0 nm wavelength intervals over 1590 to 2557 nm with an effective band pass of 12 nm. An internal ceramic reference is scanned for each sample scanned. The average of 40 scans taken for each sample to improve the signal-to-noise ratio. Approximately 2 g of leaf material are required for NIRS analysis using the Neotec 51A. Reflectance (R) data in NIRs are typically recorded as log (1/R), a form analogous to absorbance and, assuming Beer’s Law, approximately linear with concentration. A variety of data treatments have been developed by others with the common objective of reducing system and sample noise and emphasizing diagnostic differences between samples. Least-square curve fitting has been investigated as a means to express a sample spectrum as a linear combination of spectra of known constituents (Hruscha and Norris 1982). Derivative spectroscopy (multiple order differentiation of the log (1/R) data) has been found to be effective in reducing errors from baseline shifts in the data and interfering absorptions (Williams et al. 1983; Dixit and Ram 1985). Figure 1 shows three spectra from a sample that was stirred prior to each scan. This change in particle arrangement caused by stirring resulted in a change in the scattering characteristic of the sample, shown as a shift in the baseline reflectance. The derivative transformation compensates for that shift and the transformed spectra are essentially equivalent.

A modified stepwise linear regression technique was used to determine optimal wavelengths for predicting chemical composition (Infrasoft International; Westerhaus, 1985). Regressions were approached in two ways. (1) by allowing the stepwise procedures
TABLE 1. Equations used in NIRS regression analysis

\[ Y_i = \beta_0 + \beta_1 \log (1/R_1) + \ldots \]

First difference
\[ Y_i = \beta_0 + \beta_1[\log (1/R_2) - \log (1/R_1)] + \beta_2[\log (1/R_3) - \log (1/R_2)] + \ldots \]

Second difference
\[ Y_i = \beta_0 + \beta_1[2 \log (1/R_2) - \log (1/R_1) - \log (1/R_3)] + \beta_2[2 \log (1/R_3) - \log (1/R_2) - \log (1/R_4)] + \ldots \]

where
- \( Y_i \) = % constituent
- \( R_{ij} \) = reflectance at wavelength \( j \)
- \( \beta_k \) = regression coefficient for term \( k \)

---

**Figure 1.** Three spectra of a milled wood sample stirred prior to each scan: (a) \( \log(1/R) \) and (b) first derivative spectra. Breaks in spectra are due to gaps between filters.

To statistically select the “best” wavelengths and (ii) by initializing the selection process with wavelengths chosen on the basis of absorption features in spectra of pure organic compounds. The objective of this second approach is to create an analytical model with a theoretical basis that will serve as a stable calibration for future samples.

Approximately 30% of the samples were reserved as a validation set for testing all possible equations within each mathematical transformation. Fifty samples are recommended as the minimum for calibration, although a higher percentage should yield more accurate predicted values (Shenk et al. 1978). For the 203 nitrogen values, 138 were randomly selected for calibration and 65 were retained as validation samples. For lignin, 91 of the total 132 values were used for calibration and the remaining 41 for prediction. Derivative transformations of the \( \log(1/R) \) data were approximated by the software using first and second difference calculations (Table 1). The best regression equation was selected as that with the highest correlation coefficient, the lowest SEC, and an \( F \)-ratio greater than 10.0 for each selected wavelength. Validation samples were combined with the calibration set to derive the final regression equations.

**Results**

Nitrogen concentration was best predicted using a first derivative transformation of the \( \log(1/R) \) data. Four samples were considered outliers based on large \( t \)-statistics, which indicated that the laboratory values did not represent the samples at the time of the scan. These four samples were omitted from the analysis. The full statistical procedure and the regression forcing four wavelengths known to be predictive of protein content produced six-term and five-term linear equations, respectively, for the best correlation with Kjeldahl values. Summary statistics were identical between the two procedures, although different wavelengths were selected. The variance explained in all sets (calibration, validation, and final) was equal to 0.98, with each set SEC = 0.11 (Table 2; Fig. 2).

Only one known (to us) absorption feature of lignin is found in the spectral range of the instrument and it was included in the regression equation chosen by statistical criteria. Therefore, only one “best” equation was produced and this was arrived at with a second derivative transformation (Table 2; Fig. 3). Two outliers were omitted based on large \( t \)-statistics. The six terms of the equation explained 78% of the variance, SEC = 3.02. Validation samples were predicted with an \( R^2 \) of 0.71 and an SEP = 3.14. The final regression equation, with the total 130 samples, explained 78% of the variance, with an SEC = 2.90.

**Discussion**

The regression equations were developed from a wide range of species and different tissue conditions (green and brown). Species differences did not appear to affect the calibration nor did differences in tissue condition cause any outliers; i.e., there were no obvious trends in the small
Table 2. Summary of multiple-term regression analysis relating NIR data to chemical analyses

<table>
<thead>
<tr>
<th></th>
<th>Calibration</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>$R^2$</td>
<td>SEC</td>
<td>$n$</td>
<td>$R^2$</td>
<td>SEP</td>
<td>$n$</td>
<td>$R^2$</td>
</tr>
<tr>
<td>Nitrogen</td>
<td>135</td>
<td>0.98</td>
<td>0.11</td>
<td>68</td>
<td>0.98</td>
<td>0.11</td>
<td>199</td>
<td>0.98</td>
</tr>
<tr>
<td>Lignin</td>
<td>91</td>
<td>0.78</td>
<td>3.02</td>
<td>41</td>
<td>0.71</td>
<td>3.14</td>
<td>130</td>
<td>0.78</td>
</tr>
</tbody>
</table>

The number of outliers. Calibrations on subpopulations of only green or only litter samples were less successful as a result of smaller sample size.

Near infrared reflectance predictions compare favorably with the standard laboratory methods (Table 3). While the NIR results cannot be better than the data used for calibration, the standard error of the NIR prediction for nitrogen shows precision similar to the laboratory technique. In a broader field sampling context, the total error in both sampling and analysis combined can be greatly reduced using NIRs because much larger numbers of samples can be collected and analyzed. Approximately 2 min are required per sample.

The complexity of the procedures for measuring lignin increases the sources of error in the chemical method. The NIR SEP for lignin is approximately 27% higher than the laboratory error (Table 3). The lower standard errors reported by investigators in agriculture (Norris et al. 1976; Shenk et al. 1981; Marten et al. 1983) are likely a reflection of the much lower lignin content of forages. Standard errors may also be influenced by the different procedure used in forage analysis for lignin determination (commonly, acid detergent lignin using permanganate (Van Soest and Wine 1968)) or the extended spectral range of the NIR instruments. Of those investigations in which regression terms were reported, all included at least one term in the shorter wavelengths of the NIR, out of the range of the Neotec 51A used here.

In agreement with the results presented here, Redshaw (1985) reported best performance with a first derivative transformation for crude protein prediction. Others have used both log ($1/R$) and second derivative values (Norris et al. 1976; Shenk et al. 1981; Marten et al. 1983). The second derivative transformation has been found by these same investigators to be most useful for lignin prediction.

The exact wavelengths of pure compound absorption peaks are not expected to appear in the equations since the presence of other compounds in the mixture will tend to broaden and confound their effect. However, the regression equations may include terms that appear close to the constituent’s absorbing wavelengths and others that occur to compensate for background constituents.

While both approaches to the regression procedure produced suitable results for nitrogen prediction, all terms of the equation produced by initiation with selected wavelengths corresponded closely to known absorption peaks of leaf constituents (Table 4). Terms in the other equation, where all were selected by the stepwise regression, were less strongly associated with known absorption features and were generally more difficult to interpret. Of the two equations, the former is closer to being a true analytical model in the sense that it has a theoretical basis and should serve for prediction of future samples. The terms are wavelengths

![Fig. 2. Relationship between Kjeldahl nitrogen values and near infrared reflectance values predicted from multiple regression reflectance data at five wavelengths ($n = 199$).](image1)

![Fig. 3. Relationship between wet chemistry lignin values and near infrared reflectance values predicted from multiple regression reflectance data at six wavelengths ($n = 130$).](image2)

Table 3. Summary of NIR prediction vs. laboratory values

<table>
<thead>
<tr>
<th></th>
<th>$n$</th>
<th>Range of values (% dry matter)</th>
<th>Mean</th>
<th>SE</th>
<th>Lab</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nitrogen</td>
<td>199</td>
<td>0.43–3.06</td>
<td>1.48</td>
<td>0.11</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>Lignin</td>
<td>130</td>
<td>3.32–30.51</td>
<td>15.82</td>
<td>2.90</td>
<td>2.28</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Wavelengths (in nm) used in NIRS regression equations

<table>
<thead>
<tr>
<th>Term</th>
<th>Wavelength</th>
<th>Coefficient</th>
<th>Wavelength</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>2063 (P)</td>
<td>358.608</td>
<td>2135 (S)</td>
<td>-2.832</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>2181 (P)</td>
<td>31.075</td>
<td>1963 (L)</td>
<td>6855.375</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>2293 (F)</td>
<td>22.381</td>
<td>2076 (P)</td>
<td>4577.982</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>2129 (S?)</td>
<td>-42.112</td>
<td>2270 (C)</td>
<td>-2862.437</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>2053 (P)</td>
<td>-260.933</td>
<td>1654 (L?)</td>
<td>5173.482</td>
</tr>
<tr>
<td>$\beta_6$</td>
<td>2327 (S)</td>
<td>3170.335</td>
<td>2230 (P)</td>
<td>2330 (S)</td>
</tr>
</tbody>
</table>

Note: L, lignin; P, protein; C, cellulose; S, starch; ?, wavelength > 10 nm from the absorption peak. Chemical assignments of NIR bands are as follows: 1685 (L); 2055 (P); 2060 (P); 2100 (S); 2180 (P); 2270 (C); 2300 (P); 2330 (S).

of known absorption peaks of nitrogen and other leaf constituents and are not only specific to the population of samples in hand.

Despite the paucity of known lignin absorption features in the NIR part of the spectrum, the terms of the regression equation can all be shown to have some physical significance (Table 4). Again, this is useful if the equation is to be considered for general prediction purposes. Log (1/R) and first and second derivative spectra for a green Acer saccharum sample are shown in Fig. 4 with the location of wavelengths used in the final regression equations for each component.

NIRS exhibits several advantages over conventional wet chemistry methods for foliar analysis. It is rapid (approximately 2 min per sample), requires little sample preparation, and a number of constituents can be measured simultaneously. Recalibration is only necessary when the sample population or instrument changes. The precision reported here for nitrogen determination meets the performance of current laboratory methods. If we are to improve the precision and accuracy of lignin determinations, the precision of the chemical analysis method needs to be improved or the spectral range of the measurements be increased to include the shorter wavelengths of the NIR. We suggest NIRS be applied as a routine analytical method for nitrogen and lignin determination. Further research is recommended to examine the applicability of NIRS for determination of other foliar constituents that are at concentrations above 0.1% the approximate detection limit of the instrument.

Acknowledgements

This research was funded by NASA under Joint Research Interchange No. NCA-2-28 with Ames Research Center, by the National Science Foundation (grant No. BSR-8317531) and by the McIntire-Stennis program through the College of Agriculture and Life Science, University of Wisconsin–Madison. The assistance of the staff of the University of Wisconsin Soils and Plant Analysis Laboratory, in particular Dr. Emmett Schulte and Jan Jarman, is gratefully acknowledged. We thank John Obst of the U.S. Department of Agriculture Forest Products Laboratory for milled wood enzyme lignin samples. Steve McNulty, Sue Kaat, and Robin Harrington collected the foliage samples and prepared them for analysis. Lignin analyses were performed by Andrea Turner. Technical support was provided by the University of Wisconsin Institute for Environmental Studies and the Environmental Remote Sensing Center.
WESSMAN ET AL.


