Relationship Between Soil Moisture Content and Soil Surface Reflectance

A. L. Kaleita, L. F. Tian, M. C. Hirschi

Abstract. Depending on the topography and soil characteristics of an area, soil moisture, an important factor in crop productivity, can be quite variable over the land surface. Thus, a method for determination of soil moisture without the necessity for exhaustive manual measurements would be beneficial for characterizing soil moisture within a given region or field. In this study, soil surface reflectance data in the visible and near-infrared regions were analyzed in conjunction with surface moisture data in a field environment to determine the nature of the relationship between the two, and to identify potential methods for estimation of soil moisture from remotely sensed data in these wavelengths. Results indicate that it is feasible to estimate surface (0 to 7.6 cm) soil moisture from visible and near-infrared reflectance, although estimating moisture regimes rather than precise water content is perhaps more likely. Furthermore, an exponential model was appropriate to describe soil moisture from spectral reflectance data. In particular, the visible region of the electromagnetic spectrum works well with such a model. A partial least squares analysis with improved R² values over the single-band models indicated that multispectral data may add more useful information about soil moisture as compared to single-band data. The results also suggested that the performance of reflectance models for moisture estimation is a function of soil types; the estimation results were better for the lighter of the two soils in this study.

Keywords. Mapping, Precision farming, Remote sensing, Spectrometry.

The precision farming concept is based on the fact that crop productivity varies spatially and temporally within a field, depending on weather, crop factors, pests, soil conditions, topography, and cultural practices. One important factor in crop productivity variability is soil moisture. Numerous studies have shown that soil moisture contributes significantly to crop development at different stages (Power et al., 1961; Moore and Tyndale-Biscoe, 1999; Machado et al., 2000; Liu et al., 2001; Stewart et al., 2002). Thus, effective mapping techniques for soil moisture are important for establishing a complete site-specific management plan. With high-quality soil moisture data, precision farming decisions can be made more accurately, and map- or sensor-based equipment can be developed to carry out site-specific operations in the field (Tian, 2002; Bullock et al., 1998).

Traditionally, soil moisture mapping has been accomplished through exhaustive point measurement, which can be cost-prohibitive. Gravimetric measurements, while very reliable, are also very time- and resource-consuming. Several methods for measuring soil moisture with imbedded sensors, such as time- and frequency-domain reflectometers, have been developed. These sensors do not require quite as large an investment of time and facilities, and generate data that can be automatically logged. However, all of these methods suffer from some of the same disadvantages. In situ measurement can often be tedious, and it generally results in poor spatial resolution of soil moisture data. Depending on the topography of an area and the soil characteristics, soil moisture can be quite variable over the land surface. Thus, a method for determination of soil moisture without the necessity for exhaustive manual measurements would be beneficial in characterizing soil moisture within a given region or field. Remote sensing offers the potential for high-resolution, aggregated soil moisture mapping.

Remote sensing measurements of the soil record the amount of radiation in a given wavelength reflected off of or emitted from the surface to the sensor. There are many factors that affect the resulting spectrum from the soil. The color of the soil influences its measured reflectance in the visible wavelengths. Soil texture also affects reflectance, as incoming radiation is scattered differently by coarse particles as compared to fine particles (Thomasson et al., 2001). In general, larger aggregate soil particles will have lower measured reflectance. For the same reason and because of shadowing effects, surface roughness, including the clods and machinery tracks that are especially prevalent in agricultural fields, also affects the measured reflectance from the soil surface (Matthias et al., 2000; King and Pradhan, 2001). Furthermore, surface crusting of soil tends to increase reflectance (Cipra et al., 1980; Baumgardner et al., 1985; Ben-Dor et al., 2003). Because of the effects on soil color and texture, the mineral composition of the soil, including the soil organic matter content, also plays a role in the measured
reflectance from the soil surface (Palacios–Orueta and Ustin, 1998; Ben–Dor et al., 2002).

Soil moisture also affects the reflectance of the soil, although the manner in which it does so varies across the electromagnetic spectrum. Many different regions of the spectrum have been used to estimate soil moisture, including gamma radiation (Carroll, 1981), thermal infrared (Price, 1982), and passive and active microwave (Jackson et al., 1996). Microwave radiation has the advantage of being able to penetrate into the soil to a wavelength–dependent depth and is, in some wavelengths, able to penetrate through vegetation cover. These are significant advantages of microwave remote sensing and are why there has been a vast body of research into determination of soil moisture with microwave sensors (Jackson and Schmugge, 1995; Jackson et al., 1999). However, the resolution that is practically achievable with microwave sensors is not as fine as that with comparable optical sensors because footprint size increases as wavelength increases. Even ground–based microwave sensors mounted on a boom or truck have spatial resolutions on the order of several meters. Furthermore, due to the limited applications of microwave imagery for agricultural studies and the currently relatively high cost of acquiring microwave data, microwave sensors are not as practical for agricultural applications from an economic standpoint as sensors in visible and near–infrared (VIS–NIR) wavelengths.

Several researchers have found bands in the shortwave–infrared region (SWIR, 1.5 to 2.5 μm) to be suitable for soil moisture prediction (Whiting et al., 2004; Liu et al., 2002). However, SWIR sensors are currently an order of magnitude greater in price than similar VIS–NIR sensors.

Remote sensing in VIS–NIR, while more affordable, is less effective for soil moisture measurement due to the influence of confounding factors, as discussed above (Engman, 2000). Nonetheless, in a precision farming environment, many of these factors can be and are quantified on an appropriate scale. Despite the stated difficulties, there have been successful efforts to use visible and infrared wavelengths in estimation of soil moisture from remote sensing data. Moran et al. (1994) and Gillies et al. (1997) explain a method for determining crop water deficit (which translates to soil moisture) using remotely sensed vegetation indices and surface temperatures from visible and infrared data. There are several drawbacks to the temperature–vegetation index method. One is that in order to have enough points in a remote sensing image to use in determination of the boundaries of extreme conditions, which is necessary for interpolation of the most common conditions, a sufficiently large number of pixels needs to be sampled. In addition, all four of the extreme conditions need to be present in a given image: totally wet vegetated pixels, totally wet bare pixels, totally dry vegetated pixels, and totally dry bare pixels. For climate studies, which are the types of studies most likely to utilize these methods, this is not a severe limitation because the scale of imagery used in these types of studies is generally sufficient. This limitation, however, is more of a handicap when dealing with smaller–scale imagery on the order of the size of a typical farm field.

There is also research to suggest that significant information on soil moisture can be gleaned directly from VIS–NIR remote sensing bands. This is based on the commonly observed concept that a soil is usually darker when wet than when dry. Bach and Mauser (1994) note that “the spectral resolution of multispectral sensors ... is not sufficient to determine soil water content from spectral reflectance. Instead, hyperspectral data ... is needed to allow the identification of specific water absorption features.” They also note that in the spectral region of 500 to 800 nm, darkening of the soil is due to internal reflection of incoming radiation in the soil water surrounding the soil particles. Absorption of radiation by water in the soil is the primary contributor to wet soil darkening at wavelengths over about 800 nm, including a minor water absorption band around 950 nm (Ben–Dor et al., 2002; Bach and Mauser, 1994).

Muller and Décaps (2001) used simulated SPOT reflectance data derived from an aerial platform to model soil moisture over French agricultural fields under bare–soil conditions. They compared the reflectance data to moisture in the top several millimeters of soil and found that an exponential model of the following form could be used to quantify this relationship:

\[
\rho_s(\lambda) = \rho_{so}(\lambda) \exp(a_s(\lambda)M)
\]

where \(\rho_s(\lambda)\) is the reflectance of the wet soil at a given band \(\lambda\), \(a_s(\lambda)\) is the reflectance attenuation factor for the soil \(s\) in the spectral band \(\lambda\) due to the soil moisture \(M\), and \(\rho_{so}(\lambda)\) is the theoretical reflectance of the soil in the spectral band \(\lambda\) with a soil water content at air dryness. Furthermore, their study, which included several different soil types, suggested that \(\rho_{so}(\lambda)\) is representative of the hue of the soil, which is intrinsically stable and reasonably constant for soils in the same region with similar geologic origin.

The objectives of this study were to: (1) study the relationship between high–resolution remote sensing data for and near–surface soil moisture, and (2) determine if the exponential model is applicable to near–surface layers.

**STUDY DESCRIPTION**

A University of Illinois research farm field in southeast Urbana, Illinois, was used in this study, which was carried out during the summer 2002 growing season. This field, named Grein, is an 8.2–acre cornfield with moderate topographic variation. Grein has two soil types, Dana and Drummser. Dana (fine–silty, mixed, superactive, mesic Oxyaquic Argiudoll) is a grayish–brown, moderately well drained silt loam typically found on till plains and moraines, while Drummer (fine–silty, mixed, superactive, mesic Typic Endoaquoll) is a blackish, poorly drained, moderately permeable silty clay loam (Mount, 1982). No drainage tile lines are known to be located within Grein. The field was tilled after harvest the previous fall, and cultivated in the spring just prior to planting. During the 2002 season, the field was planted so that there were alternating strips of corn (16 rows at 76.2 cm spacing) and bare soil (6.1 m wide). Because of the limitations of optical remote sensing of soil characteristics, specifically the interference of vegetation cover in these wavelengths, the best data for investigating a relationship between reflectance and soil moisture are those collected under bare–soil conditions, prior to significant canopy development. Consequently, this work focused on collecting spectral and moisture data under minimal to no vegetation cover. Sampling points were therefore located within the bare–soil swaths in the field. Markers were placed in the field at each of the sampling locations so that they could be easily
identified during repeat sampling, and GPS data were collected with a Leica GPS system.

Ground–based spectral reflectance data were collected in the field on each sampling occasion with an HR2000 high–resolution miniature fiber optic spectrometer (Ocean Optics, Inc., Dunedin, Fla.). The HR2000 is a small–footprint, modular spectrometer with a spectral resolution of 0.065 nm, measuring in wavelengths from 331.54 nm to 1068.97 nm. The view angle is approximately 25°. Upwelling radiation is recorded in digital counts.

In order to normalize the digital count data recorded from the spectrometer, it was necessary to take a reference scan immediately prior to every soil surface scan. The reference scan was taken from a piece of standard white office paper held about 40 cm underneath the sensor. Multiple comparisons with a white diffuse reflectance standard (Spectralon SRS–99–020, LabSphere, Inc., North Sutton, N.H.) under both sunny and shaded conditions were used to confirm that the white paper had a sufficiently uniform reflectance across the spectral range of interest. White paper was used so that it could be easily replaced when smudged or dirtied in the field, a significant concern in often dusty field conditions. The soil surface scan was then taken by holding the spectrometer 1.5 m above the soil surface, being sure to avoid shadows from the operator or equipment. From this height, the spectrometer measured a circular area on the ground with a diameter of about 66 cm. The difference in distance from the sensor to the reference and to the soil surface may have introduced a small amount of error.

In order to maximize the strength of the signal detected by the spectrometer, the integration time for the spectrometer data acquisition was adjusted so that for each location, the signal from the white paper reference was near saturation. As long as the solar conditions remained relatively constant, the integration time did not need significant adjustment during the data collection window.

The original spectrometer data has 2047 bands. This amount of data is difficult to work with because of the time necessary for data processing, especially when performing statistical analyses. Furthermore, these data are somewhat noisy. To reduce the effect of both of these problems, the data were smoothed with a k–smoothing algorithm having a five–band window, and then pared down by eliminating every other waveband. This process resulted in a less noisy data set of a much more manageable size (1023 bands per spectrum). Subsequently, because the signal–to–noise ratio at the edges of the spectral range of the spectrometer is fairly low, the data in these bands were eliminated. The 807 bands in the range 408.22 nm to 999.55 nm were retained.

Near–surface soil moisture measurements were made via gravimetric sampling at each location on each date. Immediately following the spectral data collection at a given location, two 3.8 cm diameter, 7.6 cm deep soil cores were removed from within the spectrometer footprint and placed in a labeled zip–seal plastic bag. After sampling of the entire field was completed, the samples were taken back to the laboratory, where they were removed from the bags and weighed. The samples were then oven–dried at approximately 100°C for 23 h, removed from the oven, and re–weighed. Several initial tests of an additional 6 h at this temperature indicated sufficiently minimal additional drying of these samples after 23 h. The dry–based gravimetric soil moisture was then calculated for each sample.

Table 1 lists the dates for this study. In order to minimize both the effect of time of day on the spectral data and the effect of drying time on the moisture data, every effort was made to perform the data collection within a consistent and minimal time frame. Data collection was limited to approximately a 2 h window. Scheduling conflicts caused some inconsistency in the times of data collection, but sampling generally was either late morning or early to mid–afternoon. Furthermore, not all locations were sampled on every date, primarily due to battery failure or other technical problems.

In addition to the ground data that were collected specifically for this research, ancillary data were also collected at Grein. Distributed soil samples were collected throughout the field and sent out for analysis for organic matter content and nutrient information. A second reflectance dataset was collected on four dates in May and June to assess the impact of time of day on the soil reflectance measurements in this field. One location in the field was monitored with a Theta Probe (Delta–T Devices, Cambridge U.K., marketed in the U.S. by Dynamax, Inc., Houston, Texas) to determine the moisture content at sub–minute intervals, and reflectance spectra were collected using the same method described above every 20 min over the course of a day.

**RESULTS AND DISCUSSION**

To ensure that the differences in time of day of data collection were not introducing significant effect on the relationship between moisture and reflectance, the second independent data set was analyzed by comparing the time of day (converted to military decimal) to the reflectance. The time of day compared was limited to approximately 10:00 a.m. to 3:30 p.m. Results indicated that time of day accounted for between 0% and 8.7% of the variance in the data, depending on wavelength, with the higher numbers in the 800–900 nm region. Average reflectance variance explained by time of day across all wavelengths was 5.5%. These low numbers indicated that the reflectance–moisture dataset was not likely to include prohibitive differences due solely to time of day.

At the Grein field, the two soil types present are evident to the naked eye because of organic matter differences. Histograms of measured organic matter content, shown in figure 1, illustrate the two soil groups.

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>No. of Sampled Locations</th>
</tr>
</thead>
<tbody>
<tr>
<td>17 June</td>
<td>13:35 – 15:25</td>
<td>43</td>
</tr>
<tr>
<td>18 June</td>
<td>13:49 – 15:18</td>
<td>42</td>
</tr>
<tr>
<td>19 June</td>
<td>13:49 – 15:02</td>
<td>43</td>
</tr>
<tr>
<td>19 July</td>
<td>11:59 – 13:45</td>
<td>40</td>
</tr>
<tr>
<td>20 July</td>
<td>10:16 – 12:15</td>
<td>38</td>
</tr>
<tr>
<td>21 July</td>
<td>9:59 – 11:01</td>
<td>43</td>
</tr>
<tr>
<td>15 Aug.</td>
<td>11:02 – 12:55</td>
<td>43</td>
</tr>
<tr>
<td>20 Aug.</td>
<td>10:31 – 12:33</td>
<td>40</td>
</tr>
<tr>
<td>12 Sept.</td>
<td>13:35 – 15:24</td>
<td>40</td>
</tr>
</tbody>
</table>
The data suggest that the two soil groups split at an organic matter content of about 3.0%. While it is not likely to result in perfect categorization because of a certain amount of overlap, values of 3.0% or less were classified into one group (the light soil group), and values of 3.1% or more were classified into another (the dark soil group). For the purpose of delineating soil groups for their potential differences in spectral reflectance, organic matter differences are likely to be the most influential. Thus, for the spectral analysis, the sampling locations for each field were assigned to the light or dark soil group based on organic matter content.

Equation 1 was rearranged to:

\[
M = \log \left( \frac{\rho_s(\lambda)}{\rho_{so}(\lambda)} \right) \frac{a_s(\lambda)}{a_o(\lambda)}
\]

and models of this logarithmic form were fit to all of the data by wavelength. For comparison, linear models were also fit. Models and their corresponding R² values were also computed for the data segmented by soil type. R² values for these models are presented by wavelength in figure 3.

Figure 3 illustrates that the visible region, from about 500 nm to 700 nm, has the strongest relation to soil moisture. The exponential model is a better fit to the data than the linear model. Although not shown in the figure, this was true for the individual soil groups as well as both groups together. Furthermore, the models for the light soil have higher R² values than those for the dark soil. This may be because the dark soil reflects so little already that incremental differences in darkening due to moisture are harder to detect.

In order to visualize the nature of the relationship between moisture and relative reflectance, the data and model for the 535 nm waveband are plotted in figure 4. This waveband was selected because it was in the range of the maximum R² for both light and dark soils as well as the aggregated data. Other wavebands within 500–700 nm have similar behavior. Both the light and dark soil groups when visualized separately have the same overall behavior.
Figure 3. $R^2$ values for a linear model and for logarithmic models of the form of equation 2, fitted by wavelength.

Figure 4. Soil surface reflectance at 535 nm versus gravimetric moisture content for both soil groups. Data from 19 July 2002 are plotted as “+”, while the rest of the data are plotted as “o”.

Figure 4 illustrates that the reflectance decreases with increasing soil moisture, as expected, but that the rate of decrease in relative reflectance becomes more moderate with increasing soil moisture. This is likely because at very high soil moisture contents, the soil is already quite dark, and further moisture added to the soil has less and less of an effect on the reflectance. This implies that a logarithmic model such as equation 2 is indeed suitable.

Also notable in figure 4 is that there appears to be a cluster of points that does not follow the same trend as the rest of the data. Further investigation reveals that most of it is from 19 July 2002. These data are plotted with a different symbol in figure 4. Although lighting condition and antecedent weather data from this date were compared to those from the other dates, there is no obvious explanation for the deviation of the 19 July 2002 data from the pattern followed by most of the rest of the Grein data. Muller and Décamps (2001) found that drier soils exhibited more random spectral reflectance. It could be that because the 19 July 2002 moisture contents were low, the spectral data from this date exhibit a more random behavior. Other dry data, however, are not as anomalous.
The fitted models do not perform well for a single day’s data. The range of moisture values present in the field is small when compared to the range for a series of days, and the models do not have the precision that would be necessary to differentiate between these comparatively minor differences in moisture content.

In order to investigate the potential for using combinations of bands rather than a single band of reflectance data to estimate surface soil moisture, a partial least squares analysis (PLS) was performed on the 807-band data set. This type of analysis identifies independent signals within a given data set that are the most related to a reference variable. The new signals, or factors, are linear combinations of the original data and are successively combined to provide improving prediction of the reference variable. More details on PLS for remote sensing data analysis can be found in Wold (1982). PLS was performed on the data by soil group. Table 2 lists R\(^2\) values for the first four PLS factors related to moisture. There is a substantial improvement in the PLS model from one factor to two factors, with only minor improvement thereafter. Furthermore, use of three and four factors results in weightings for the original bands that are harder to determine useful information from, because the weightings begin to have a noisier appearance, rather than varying more smoothly across the spectrum. Figure 5 thus shows the band weightings, or loadings, for the PLS model with two factors.

The two–factor model loadings are very similar for the light and dark soil groups and both groups together. These loadings are also similar in shape to the relationship between individual reflectance bands and moisture, shown in figure 4, particularly in the area between 500 and 800 nm. However, the PLS analysis also points to contribution from wavelengths around 900 and 1000 nm. There is some correspondence here to the minor water absorption band around 950 nm, although the 970 nm band is more influential for this data. Because these wavelengths are not well correlated individually with moisture, these results imply that wavelengths in this region are most useful in conjunction with wavelengths in the visible region of the spectrum. Furthermore, the improved R\(^2\) values with the two–factor PLS models compared to the single–band exponential models indicates some advantage to using multispectral data over single–band data.

It is unclear from this analysis how much spectral resolution is needed to achieve these results. In order to compare the hyperspectral data to lower–resolution multispectral data, the spectra were resampled to create bands that correspond to VIS–NIR channels for two common multispectral remote sensors: Landsat 7 ETM (500–590 nm, 610–680 nm, and 790–890 nm), and SPOT 5 HRG (450–520 nm, 520–600 nm, 630–690 nm, and 760–900 nm). A PLS analysis was run on these spectra; results for the exponential and two–factor PLS models for both soil groups are shown in table 3.

These results are similar to the results using the full hyperspectral data. For all three sensor designs, the maximum R\(^2\) value using an exponential channel is 0.54 and occurs between 500 and 600 nm. However, results using a

![Figure 5. Band weightings for the two–factor PLS models by soil group.](image-url)
Table 3. Results of exponential and two–factor PLS models of reflectance for simulated multispectral data.

<table>
<thead>
<tr>
<th>Exponential Model</th>
<th>Landsat 7 ETM</th>
<th>SPOT 5 HRG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channel 1</td>
<td>R² = 0.57</td>
<td>R² = 0.47</td>
</tr>
<tr>
<td>Channel 2</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>Channel 3</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Channel 1</td>
<td>R² = 0.47</td>
<td></td>
</tr>
<tr>
<td>Channel 2</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>Channel 3</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>Channel 4</td>
<td>0.26</td>
<td></td>
</tr>
</tbody>
</table>

Two–factor PLS

<table>
<thead>
<tr>
<th>SPOT 5 HRG</th>
<th>Landsat 7 ETM</th>
</tr>
</thead>
<tbody>
<tr>
<td>R² = 0.57</td>
<td>0.57</td>
</tr>
</tbody>
</table>

PLS analysis to combine data from multiple bands are slightly worse for the Landsat and SPOT data (two–factor R² of 0.57 in both cases) than for the full hyperspectral data (two–factor R² of 0.59).

CONCLUSIONS

The results of this research indicate that it is feasible to estimate surface (0 to 7.6 cm) soil moisture from visible and near–infrared reflectance, although estimating moisture regimes rather than precise water content is perhaps more likely. Furthermore, this study confirms that the relationship between soil moisture and spectral reflectance proposed by Muller and Décamps (2001) is appropriate for the depths and soils present in this study. In particular, the visible region of the electromagnetic spectrum was useful with such a model. However, the results suggested that the usefulness of reflectance models for moisture estimation depends on soil type; in this study, results were better for the lighter of the two soils.

The results of the PLS analysis indicate that the added data provided by multispectral data may add more useful information about soil moisture, specifically as compared to single–band data. Furthermore, the relationship between the PLS results and soil moisture for narrow–band hyperspectral data was slightly stronger than for wide–band multispectral data; however, obtaining and processing data at high spectral resolution is more expensive.

ACKNOWLEDGEMENTS

This material is based upon work supported by the Illinois Council for Food and Agricultural Research under Award No. IDA CF 01–DS–3–1 AE, and the USDA National Needs Graduate Student Fellowship. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the Iowa State University nor the University of Illinois.

REFERENCES


