Identification of soil classification units from VIS-NIR spectral signatures

Ádám Csorba, Kornél D. Szalay and László Fenyvesi, Hungarian Institute of Agricultural Engineering, H-2100, Gödöllő, Tessédik S. u. 4.
Vince Láng, Erika Michéli, Szent István University, H-2100, Gödöllő, Páter K. u. 1.

Abstract

In this paper we present the first steps of a methodological development, which aims the classification of soil profiles based on their spectral characteristics. In the frame of this study the properties of 12 soil profiles were investigated in Hungary. To being able to characterize the spectral characteristics of the soils we took samples by 5 cm increments, to 1.5 m depth. The reflectance spectra of the samples were acquired by using the Analytical Spectral Devices (ASD) FieldSpec 3 MAX portable spectroradiometer. To reduce the high dimensionality of the spectral data, and to compress the information into a few variables, principal component analysis was performed on the spectral dataset. The profile-wise spectral distribution of the principal component score values were used for the vertical characterization of the soils. To define the similarities and dissimilarities between the studied soils Euclidean distance calculation was performed based on the profile-wise distribution of factor score values. A second principal component analysis was performed based on the similarity matrix, thus the similarity and dissimilarity of soils can be visualized in a two dimensional data space. The study showed that reflectance spectroscopy can efficiently support the soil classification procedure.

Keywords: soil, reflectance, spectroscopy, classification, World Reference Base

1. Introduction

The modern soil classification systems are definition, standard measurement and quantitative data based, so called diagnostic systems. If the diagnostic approach is being followed, the classification of soils is based on the identification of diagnostic horizons, properties and materials.

Diagnostic soil horizons and properties are characterized by a combination of attributes that reflect widespread, common result of soil formation processes, or indicate specific conditions of soil development. Some of their features are perceptible to the naked eye, or can be measured by using simple methods in the field. However, a wide range of diagnostic attributes are to be measured in the laboratory. The application of most of these methods is time-consuming; the measurements require expensive instruments, furthermore environmentally harmful chemicals.

A large number of publications deal with the possibilities to use spectroscopic methods in prediction such soil parameters whose determination is necessary to define a diagnostic horizon or property (VISCARRA ROSSEL et al., 2009). Most of the efforts is successful, so it is
widely known that spectroscopy is suitable to enhance, or in some cases to substitute the conventional laboratory methods.

The objective of our work is to develop a method to classify the soils based on their differences in spectral characteristics through their profile.

Most absorption features in the VIS-NIR-SWIR region are overtone or combination bands in the infrared region of the electromagnetic spectrum which are due to vibrational and rotational transition processes (CLARK, 1999). Specifically, the VIS range provides a measure of soil color; broad and shallow absorption bands near 500-700 nm are due to iron oxides, oxihydroxides, hydroxides (BEN-DOR et al., 2006); narrow, well defined absorptions near 1400 and 1900 nm can be related to hydroxyl and water molecules (HAUBROCK et al., 2008); absorption features beyond 2000 nm are due to clay minerals (CHABRILLAT et al., 2002), organic constituents (STEVENS et al, 2008), carbonates (LAGACHERIE et al., 2008) and a wide range of salt minerals (METTERNICH and ZINCK, 2003). Various other absorptions due to soil constituents also occur throughout the 350-2500 nm range, however these are often difficult to identify because they may represent much weaker large-order overtones of the soil constituents that can also overlap.

Thus, the reflectance measurements in the visible and near-infrared region of the electromagnetic spectrum (350-2500 nm) provide an alternate or supplementary method to the conventional physical and chemical laboratory soil analysis for estimation of a wide range of key soil properties. The method is rapid, cheap, non-destructive, furthermore a single spectrum allows for simultaneous determination of various soil properties.

2. Materials and methods

In the frame of this study the profile-wise spectral properties of 12 soil profiles were investigated in Hungary. During the field work mollic, voronic, calcic, cambic, argic and vertic horizons were identified. During the field work, each soil profile was classified by using the World Reference Base for Soil Resources (WRB) guideline (IUSS WORKING GROUP WRB, 2006). To being able to characterize the spectral characteristics of the soils we took samples by 5 cm increments, to 1.5 m depth.

The reflectance spectra of the samples were acquired by using the Analytical Spectral Devices (ASD) FieldSpec 3 MAX portable spectroradiometer.

![Figure 1.: The instrumental set-up for the spectral measurements](image-url)
Considering the spectral properties of the organic carbon, clays and carbonates (that play key role in the classification of soils) spectral sub-regions were defined. (Table 1. and Figure 2.).

Table 1.: The selected spectral sub regions for the analysis

<table>
<thead>
<tr>
<th>Soil property</th>
<th>Spectral sub region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soil organic carbon content</td>
<td>350 – 1100 nm</td>
</tr>
<tr>
<td>Calcium Carbonate content</td>
<td>2280 – 2390 nm</td>
</tr>
<tr>
<td>Clay content</td>
<td>2170 – 2290 nm</td>
</tr>
</tbody>
</table>

Figure 2.: The selected spectral sub regions

To reduce the high dimensionality of the spectral data, and to compress the information into a few variables, principal component analysis was performed on these spectral sub-regions. To define the similarities and dissimilarities between the studied soils Euclidean distance calculation was performed that is a widely used method to define taxonomic relationships between soils. The calculation was based on the profile-wise distribution of the principal component score values.

The results of the calculation was visualized in a similarity matrix that shows numerically the relatedness of soil pairs, namely how close the soils are each other in a multidimensional data space based on their spectral properties. Soils having similar spectral properties along their profile are close to each other and show small distance values.

A second principal component analysis was performed based on the similarity matrix, thus the similarity and dissimilarity of soils can be visualized in a two dimensional coordinate system.
3. Results and discussion

The similarity matrix (Table 2.) shows the similarity and dissimilarity of the soils.

Table 2.: The similarity matrix showing the similarity of soils

<table>
<thead>
<tr>
<th></th>
<th>Endocalcic cutanic Luvisol (AGREA)</th>
<th>Endocalcic Luvisol (AGREF)</th>
<th>Calcic Chernozem (JMF)</th>
<th>Calcic Chernozem (JMACs)</th>
<th>Luvic Calciisol (JMA)</th>
<th>Gleyic Chernozem (VRA)</th>
<th>Haplic Chernozem (VRF)</th>
<th>Luvic Chernozem (SZPB)</th>
<th>Bathygleyic Chernozem (MH1)</th>
<th>Gleyic Chernozem (MH2)</th>
<th>Endostagnic albic Alisol (G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endocalcic cutanic Luvisol (AGREA)</td>
<td>1.28</td>
<td>2.45</td>
<td>2.41</td>
<td>2.23</td>
<td>2.43</td>
<td>2.53</td>
<td>1.34</td>
<td>1.48</td>
<td>2.38</td>
<td>2.71</td>
<td>2.76</td>
</tr>
<tr>
<td>Endocalcic Luvisol (AGREF)</td>
<td>2.10</td>
<td>1.23</td>
<td>2.33</td>
<td>2.27</td>
<td>2.33</td>
<td>2.41</td>
<td>1.46</td>
<td>1.66</td>
<td>2.05</td>
<td>2.49</td>
<td>2.65</td>
</tr>
<tr>
<td>Calcic Chernozem (JMF)</td>
<td>2.45</td>
<td>1.82</td>
<td>3.13</td>
<td>1.56</td>
<td>1.55</td>
<td>1.62</td>
<td>1.66</td>
<td>1.24</td>
<td>2.03</td>
<td>2.57</td>
<td>2.05</td>
</tr>
<tr>
<td>Calcic Chernozem (JMACs)</td>
<td>2.41</td>
<td>2.25</td>
<td>1.31</td>
<td>1.03</td>
<td>1.84</td>
<td>1.93</td>
<td>1.66</td>
<td>1.89</td>
<td>2.89</td>
<td>2.89</td>
<td>2.89</td>
</tr>
<tr>
<td>Luvic Calciisol (JMA)</td>
<td>2.23</td>
<td>2.72</td>
<td>3.13</td>
<td>2.25</td>
<td>2.58</td>
<td>2.50</td>
<td>2.72</td>
<td>2.74</td>
<td>2.55</td>
<td>3.20</td>
<td>3.20</td>
</tr>
<tr>
<td>Gleyic Chernozem (VRA)</td>
<td>2.43</td>
<td>1.16</td>
<td>1.31</td>
<td>2.72</td>
<td>1.16</td>
<td>1.85</td>
<td>1.84</td>
<td>1.61</td>
<td>1.87</td>
<td>3.34</td>
<td>3.34</td>
</tr>
<tr>
<td>Haplic Chernozem (VRF)</td>
<td>2.53</td>
<td>1.16</td>
<td>1.55</td>
<td>1.03</td>
<td>2.58</td>
<td>1.16</td>
<td>1.82</td>
<td>1.82</td>
<td>1.40</td>
<td>1.57</td>
<td>1.57</td>
</tr>
<tr>
<td>Luvic Chernozem (SZPB)</td>
<td>1.34</td>
<td>1.84</td>
<td>1.62</td>
<td>1.84</td>
<td>2.50</td>
<td>1.85</td>
<td>1.82</td>
<td>0.56</td>
<td>1.78</td>
<td>2.30</td>
<td>2.48</td>
</tr>
<tr>
<td>Bathygleyic Chernozem (MH1)</td>
<td>2.38</td>
<td>2.72</td>
<td>1.62</td>
<td>1.93</td>
<td>2.72</td>
<td>1.84</td>
<td>1.82</td>
<td>0.56</td>
<td>1.95</td>
<td>1.23</td>
<td>2.81</td>
</tr>
<tr>
<td>Gleyic Chernozem (MH2)</td>
<td>2.71</td>
<td>1.84</td>
<td>2.03</td>
<td>1.89</td>
<td>2.55</td>
<td>1.87</td>
<td>1.57</td>
<td>2.30</td>
<td>2.44</td>
<td>1.23</td>
<td>3.54</td>
</tr>
<tr>
<td>Endostagnic albic Alisol (G)</td>
<td>2.76</td>
<td>2.48</td>
<td>2.89</td>
<td>3.20</td>
<td>3.34</td>
<td>3.16</td>
<td>2.48</td>
<td>2.65</td>
<td>2.81</td>
<td>3.54</td>
<td>3.54</td>
</tr>
</tbody>
</table>

Chernozems are soils that show thick, black, well-structured surface horizon with a high base saturation and moderate to high content of organic matter. Chernozems that show subsurface secondary calcium-carbonate accumulation (Calcic horizon) are close to each other (JMF, JMACs). The ones whose deeper parts are saturated with groundwater during most of the year (Gleyic color pattern) show high similarity values (MH1, MH2, VRA). Chernozems showing the evidence of subsurface clay enrichment (Argic horizon), are close to each other (SZPB, SZPJ).

Luvisols are soils that have higher clay content in the subsoil than in the topsoil as a result of pedogenetic processes (especially clay migration), leading to an Argic subsurface horizon. Luvisols have high base saturation at certain depths and high activity clays throughout the argic horizon. It is evident that the two Luvisols are close to each other (AGREA, AGREF). They are similar to the Chernozems that show the subsurface clay enrichment, and are transition soils between Luvisols and Chernozems (SZPB, SZPJ).

Calcisols are soils in which there is substantial secondary accumulation of lime. The colluvial JMA soil is far from the other ones.

Alisols are soils that have higher clay content in the subsoil than in the topsoil as a result of pedogenetic process, leading to an argic horizon. Alisols have low base saturation at certain
depths and high activity clays throughout the Argic horizon. The soil named G is far from every other one, showing the different physical-chemical properties of this soil.

The Figure 3. shows the result of the PCA performed on the similarity matrix.

Figure 3.: The distribution of soil profiles based on the PC analysis performed on the similarity matrix

The scatterplot shows the distribution of the soil profiles on the first two principal components. It describes not just the similarities and dissimilarities but the transitions as well. For example the Luvic Chernozems are transitional soils between Chernozems and the Luvisols. It shows the evidence the Mollic that is diagnostic to Chernozems. They have an Argic horizon as well that diagnostic to the Luvisols. Based on the Euclidean distances this „transitional” aspect do not shows up.

4. Conclusions

The work demonstrates the significance of VIS-NIR spectroscopy in soil characterization and classification. It can supply integrative measurements of soils, and improve the efficiency of data collection, facilitate the collection of large amount of spatial data and provide more information than conventional surveying, where only few – accurate, but expensive and complicated – measurements are used.

5. Acknowledgements

The research was supported by the HU-KTIA-AIK-12-12012-0012 Research Grant.
6. References


