A spectral soil quality index (SSQI) for characterizing soil function in areas of changed land use

Tarin Paz-Kagan a, Moshe Shachak b, Eli Zaady c, Arnon Karnieli a,⁎

a The Remote Sensing Laboratory, Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Sede Boker Campus, 84990, Israel
b Mitrani Department of Desert Ecology, Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Sede Boker Campus, 84990, Israel
c Department of Natural Resources and Agronomy, Agricultural Research Organization, Gilat Research Center, Mobile Post Negev 2, 85280, Israel

A R T I C L E   I N F O

Article history:
Received 22 May 2013
Received in revised form 13 March 2014
Accepted 6 April 2014
Available online xxxx

Keywords:
Land-use change
Semi-arid regions
Reflectance spectroscopy
Spectral soil quality index
Soil quality index

A B S T R A C T

During the last several decades, a large proportion of the planet’s terrestrial surface has transformed from natural ecosystems to human-dominated systems. These land-use dynamics affect ecosystems’ soil quality. The current study was conducted at the fringe of the northern Negev Desert, Israel, and aimed to assess and compare the soil quality in three different land-use types (afforestation, traditional grazing, and agro-pastoral) that were changed from managed to unmanaged or vice versa (e.g., shrubland was transformed to a planted forest; pastoral grazing to natural shrubland with no grazing; and agro-pastoral to abandoned agricultural). The overall aim of this research is twofold: (1) to evaluate reflectance spectroscopy changes in soil physical, biological, and chemical properties and their derived soil quality index (SQI) in the changed land uses; and (2) to develop a spectral soil quality index (SSQI) toward applying the technique of reflectance spectroscopy as a diagnostic tool of soil quality. To achieve these objectives, several mathematical/statistical procedures, consisting of a series of operations, were implemented, including a principal component analysis (PCA), a partial least squares-regression (PLS-R), and a partial least squares-discriminate analysis (PLS-DA). The PLS-R’s most suitable models successfully predicted soil properties (R2 > 0.80; ratio of performance to deviation (RPD) > 2.0), including sand–silt–clay content, NH4, NO3, and pH. Moderately well-predicted soil properties (0.50 < R2 < 0.80; RPD > 2) were residual water, soil organic matter, electric conductivity, and potassium. Poor validation (R2 < 0.50; RPD < 2) results were obtained for potential active carbon, phosphorus, and hydraulic conductivity. In addition, the PLS-R model predicted the SQI in the changed land uses. The correlations between the predicted spectral values of the calculated SQI ranged 0.65 − R2 < 0.81 with RPD > 2. The PLS-DA model was used to develop the SSQI model. The correlations between the SSQI and the SQI ranged 0.66 − R2 < 0.74 in the different land uses. This study underscores the potential application of reflectance spectroscopy as a reliable diagnostic screening tool for assessing soil quality. The classification of soils into spectral definitions provides a basis for a spatially explicit and quantitative approach for developing the SSQI. The SSQI can be used to assess hot spots of change in areas of land-use changes and to identify soil degradation.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Worldwide observations have confirmed that during the last several decades, a large proportion of the planet’s terrestrial surface has transformed from natural ecosystems to human-dominated systems (e.g., Defries et al., 2004; Foley et al., 2005). These land-use dynamics are so pervasive that they significantly affect key aspects of ecosystem structures, functions, and services (Adeel et al., 2005). Accordingly, altering ecosystem services affects the ability of biological systems to support human needs (Metzger et al., 2006; Vitousek et al., 1997) and also modifies ecosystem structure and function by changing biodiversity, productivity, and soil quality (Matson et al., 1997; Tscharntke et al., 2005). Recent assessments of the ecosystem functions of soils and their importance for global sustainability underscore the importance of the management of soil resources for different land uses for present and future societal welfare (Adeel et al., 2005; Andrews et al., 2002). The concept of soil quality is related to the capacity of soil to function in supporting important ecosystem services (Idowu et al., 2009). Soil quality involves physical, biological, and chemical attributes that are merged together to indicate soil functioning (Andrews et al., 2002; Gugino et al., 2009). However, since a practical assessment of soil quality requires the integrated consideration of key soil properties and their variations in space and time, it remains a challenging task (Doran and Parkin, 1994).
Recent studies have proposed several conceptual frameworks for monitoring soil quality (e.g., Andrews et al., 2004; Black, 1965; Stevenson, 2005; Viscarra Rossel et al., 2006). These frameworks usually share the common initial step of the choice of a minimum dataset (MDS), composed of physical, biological, and chemical properties that are essential in terms of soil functioning (Rezaei et al., 2006). The soil attributes are selected from the MDS for their suitability in assessing a particular soil function (Andrews et al., 2004), a specific soil ecosystem service (Velasquez et al., 2007), or a key threat to soils (Morvan et al., 2008). Each indicative soil property is normalized to a unitless score and, finally, integrated into a soil quality index (SQI) value (Andrews et al., 2002, 2004; Idowu et al., 2008; Karlen et al., 1997). However, because many soil analyses are involved, monitoring soil quality indices at different scales and land uses remains expensive, as well as time and labor consuming, when using the standard procedures (Cécillon et al., 2009a).

In contrast, different aspects of soil quality can be assessed by reflectance spectroscopy techniques that include visible (VIS, 400–700 nm), near-infrared (NIR, 700–1100 nm), and shortwave infrared (SWIR, 1100–2500 nm). These are rapid, non-destructive, reproducible, and cost-effective analytical methods (Ben-Dor and Banin, 1995). Reflectance and absorbance signals result from vibrations in chemical bonds and minerals that provide information about the proportion of each element in the analyzed sample (Ciurczack, 2001). Recent advances in soil analysis demonstrate that reflectance spectroscopy is a robust analytical technique suited for rapid and simultaneous analysis of the abovementioned soil attributes with various levels of prediction accuracy (Awiti et al., 2008; Cécillon et al., 2009a; Odlare et al., 2005; Velasquez et al., 2005, 2007).

Although the potential of reflectance spectroscopy as a technique for the rapid and simultaneous prediction of soil properties is rather clear, the challenge is to adapt the application of spectroscopy as a diagnostic screening tool that can aid in the development of reliable, specific spectral definitions to characterize soil quality for land management. In addition, using spectroscopy to assess the impact of changes in land use on soil properties can also help to detect environmental changes, such as soil degradation, erosion, and modifications in primary productivity. Therefore, the evaluation of soil quality, using spectroscopy as a way to generate diagnostic indices, can be used for land management. Shepherd and Walsh (2002) discussed the potential of soil spectroscopy for risk-based assessments of the effects of land use and land management on soil conditions. Vågen et al. (2006) developed a spectral fertility index and used it to investigate the effects of land use and time since forest conversion on soil conditions in Madagascar. Awiti et al. (2008) developed a spectral soil condition classification method to assess tropical forest–cropland soils in Kenya. The classification of soils that Awiti et al. (2008) developed into spectrally defined condition classes provides a basis for spatially explicit and quantitative case definitions for poor or degraded soil conditions. Such approaches need more validations, particularly to test the application of reflectance spectroscopy for detecting changes in soil condition or quality due to land management.

Our study strived to assess and compare soil quality in changed land uses using laboratory analyses of soil physical, biological, and chemical properties, as well as through spectral measurements. The overall aim of this research is twofold: (1) to evaluate the ability of reflectance spectroscopy to detect changes in soil quality across changed land uses; and (2) to develop a spectral soil quality index (SSQI) toward applying the reflectance spectroscopy technique as a diagnostic tool of land management. To achieve these objectives, several mathematical/statistical procedures, consisting of a series of operations, were used, including a principal component analysis (PCA) for establishing the SQI based on the soil physical, biological, and chemical properties (Fig. 1, step 1); a partial least squares-regression (PLS-R) to relate spectral reflectance for measuring soil properties with (SQI) scoring (Fig. 1, steps 2 and

Fig. 1. The scheme for assessing the spectral soil quality index (SSQI) and the soil quality index (SQI) by laboratory and spectroscopy measurements in different land-use categories.
and a partial least squares-discriminate analysis (PLS-DA) to relate spectral reflectance classification with different land-use categories and to develop the SSQI using a scoring function (Fig. 1, steps 4–5). The last operation involved a correlation between the SQI and the developed SSQI (Fig. 1, steps 6).

2. Materials and methods

2.1. Site descriptions

Measuring the response of soil condition to management over long time scales at large spatial scales is feasible but highly demanding and requires long-term experiments. Therefore, the selected study areas are located in long-term ecological research (LTER) sites in the northern Negev Desert of Israel, across the transition between the arid and semi-arid zones (Fig. 2A). The area is characterized by a mean annual rainfall of 200–300 mm that is concentrated during the rainy season between November and April. Average daily temperature ranges from 10 °C in the winter to 30 °C in the summer. The changed land uses include several treatments (managements): (1) a natural shrubland area was changed to a planted forest; (2) a traditional grazing area was changed to a shrubland area where grazing has been excluded; (3) an agro-pastoral area was changed to an abandoned field where grazing has been excluded; and (4) an abandoned field where grazing had been excluded was changed to an abandoned field with grazing.

2.1.1. Afforestation system

The study was conducted at the Yatir Forest (Fig. 2B), which is a pine forest of approximately 2800 ha comprising predominantly *Pinus halepensis* Mill, planted mainly between 1965 and 1969 (35°03′ E, 31°20′N 650 m a.m.s.l.). The silty-loam soil is of aeolian origin deposited on a chalk and limestone substrate. There is no distinct organic horizon other than an occasional thin, 1–3 cm, litter layer (Grunzweig et al., 2003). The Yatir Forest is the largest planted forest in the semi-arid northern Negev Desert (Bonneh, 2000; Volcani et al., 2005). The surrounding native vegetation is a two-phase mosaic dominated by the shrub *Sarcopoterium spinosum* (L.), embedded in a soil biological crust matrix. During the rainy season, herbaceous annuals and perennials (vegetation height: 30–50 cm) are prominent in both patches (Sprintsin et al., 2009). Samples were taken from the area according to the landscape patchiness in the afforestation system and the adjacent shrubland system, under the canopy (understory) and in the open patches of the forest, and under the shrub and in the open patches in the adjacent shrubland (Table 1).

2.1.2. Traditional grazing ecosystem

The traditional grazing study was conducted at the Lehavim experimental farm (34°49′ E, 31°21′N, 350–500 m a.m.s.l.) (Fig. 2C). The bedrock lithology is chalk of the Eocene, and the soil is loamy (Stevi et al., 2008). The vegetation physiognomy is a shrubland in which

![Fig. 2. Study sites of the different land uses: (A) location map of the sites in the northern Negev Desert in Israel overlaid on the rainfall isohyets; (B) afforestation — Yatir Forest; (C) traditional grazing — Lehavim experimental farm; and (D) agro-pastoral grazing — Migda farm.](image-url)
Sarcopoterium spinosum, Coridothymus capitatus and Thymelaea hirsuta are the dominant shrub species, embedded in soil biological crust matrices. Annual species represent 56% of the regional flora (Danin and Orshan, 1990). The herbaceous vegetation appears in the mid-winter and persists for 2–5 months, depending on the amount and distribution of rainfall (Karnieli, 2003; Karnieli et al., 2002).

The 800-ha experimental farm was established in 1980 under the auspices of the Israeli Ministry of Agriculture and Rural Development and has since been moderately grazed by flocks of sheep and goats (Stevi et al., 2008). Four permanent enclosure plots (10 × 10 m) to prevent grazing were established on each slope in 1993. The area was grazed every year by a flock grazing were established on each slope in 1993. The area was grazed every year by a flock grazing under different grazing regimes. The main crop is spring barley (Hordeum vulgare), harvested from October to April. The bedrock lithology is chalk of the Eocene, and the overlaying soil is a sandy-loam.

The farm has a second-order transformation of the polynomial Savitzky–Golay smoothing (Luo et al., 2005; Savitzky and Golay, 1964) was performed to minimize variance between samples caused by grinding and the

Table 2
Soil quality indicators for the afforestation land use with the following treatments: understory forest canopy, open patches, shrubland under the shrub, and shrubland soil biogenic crust in the open patches. Statistics include: average value, standard deviation, and significant differences between treatments.

<table>
<thead>
<tr>
<th>Treatments</th>
<th>Sand (%)</th>
<th>Silt (%)</th>
<th>Clay (%)</th>
<th>AGG (1–6)</th>
<th>RW (m/m)</th>
<th>SH (psi)</th>
<th>HC (mm/h)</th>
<th>SOM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest understory</td>
<td>35.82 ± 4.15&lt;sup&gt;a&lt;/sup&gt;</td>
<td>43.78 ± 1.92&lt;sup&gt;a&lt;/sup&gt;</td>
<td>204.6 ± 3.58&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.92 ± 0.1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.65 ± 0.22&lt;sup&gt;a&lt;/sup&gt;</td>
<td>300.8 ± 4.17&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.136 ± 0.07&lt;sup&gt;ab&lt;/sup&gt;</td>
<td>11.18 ± 1.81&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Forest clearing</td>
<td>41.02 ± 2.22&lt;sup&gt;a&lt;/sup&gt;</td>
<td>41.58 ± 2.92&lt;sup&gt;a&lt;/sup&gt;</td>
<td>17.4 ± 1.95&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.2 ± 0.12&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.99 ± 0.13&lt;sup&gt;b&lt;/sup&gt;</td>
<td>286.13 ± 8.79&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.14 ± 0.06&lt;sup&gt;a&lt;/sup&gt;</td>
<td>8.38 ± 1.16&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Shrubland under the shrub</td>
<td>33.42 ± 1.224&lt;sup&gt;b&lt;/sup&gt;</td>
<td>49.38 ± 1.09&lt;sup&gt;b&lt;/sup&gt;</td>
<td>17.2 ± 1.923&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.43 ± 0.39&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.171 ± 0.115&lt;sup&gt;b&lt;/sup&gt;</td>
<td>316.15 ± 5.83&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.107 ± 0.049&lt;sup&gt;ab&lt;/sup&gt;</td>
<td>6.23 ± 0.72&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Shrubland open patches</td>
<td>35.42 ± 2.65&lt;sup&gt;b&lt;/sup&gt;</td>
<td>49.58 ± 1.23&lt;sup&gt;b&lt;/sup&gt;</td>
<td>15 ± 2.6&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.18 ± 0.06&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.91 ± 0.18&lt;sup&gt;b&lt;/sup&gt;</td>
<td>336.15 ± 6.73&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.094 ± 0.04&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4.24 ± 0.4&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Abbreviations: AGG — aggregation; RW — residual water; SH — surface hardness (penetration); HC — hydraulic conductivity (infiltration); SOM — soil organic matter; PAC — potential active carbon; RH — root health; EC — electric conductivity; NH<sub>4</sub> — ammonium; NO<sub>3</sub> — nitrate; P — phosphorus; K — potassium; NS — no significant differences. Small letters indicate significant differences between treatments. Values in each vertical column followed by the same letter do not differ significantly at P < 0.05 using ANOVA Turkey test.

2.1.3. Agro-pastoral ecosystem

The agro-pastoral study was conducted at the Migda experimental farm in the northern Negev Desert of Israel, located northwest of Beer-Sheva (34°25′E, 31°22′N, 100–120 m a.m.s.l.) (Fig. 2D). The farm was established in 1960 by the Agricultural Research Organization with an area of 400 ha supporting extensive agriculture that includes grazing under different grazing regimes. The main crop is spring wheat (Pimstein et al., 2009), growing annually during the rainy season from October to April. The bedrock lithology is chalk of the Eocene, and the overlying soil is a sandy-loam.

The farm has been grazed every year by a flock of about 800 Awassi sheep and 600 goats, starting in late February, when the field is fully covered, continuing until May (green pasture), and again from August to December (dry pasture). The grazing in the farm is managed by controlling the intensity, stocking density, flock size, and timing of herd introduction in the field. Three fields were examined (Table 1): (1) an abandoned agricultural field with natural vegetation in an area of 5 ha of mainly annual plants. In this field, no cultivation, irrigation, fertiliation, or grazing has been performed; (2) an abandoned field in an area of 9.6 ha, mainly with annual plants, with grazing, but with no cultivation, fertiliation, or irrigation; and (3) a monocultural agro-pastoral field of wheat in an area of 9.5 ha, with moderate grazing and cultivation but with no fertilization or irrigation.

2.2. Soil sampling and laboratory analysis

Soil samples were collected in late August of 2011, at the peak of the dry season, at a depth of 0–0.15 m; thus, the soil water content was negligible. The sampling was conducted following a stratified random survey methodology. For each treatment, the sampling included five quadrates of 1 m<sup>2</sup>, randomly placed (n = 5). In each quadrate, four soil samplings of about 700 g of soil were collected for laboratory measurement (N = 55). In total, 220 soil samples were collected (repeated measurements of four soil samples in each quadrate). The soil measurements that were conducted in the field included surface hardness (SH) and hydraulic conductivity (HC). All soil samples were transferred to the laboratory and were stored unopened at room temperature until analysis.

The Cornell Soil Health Test (CSHT) protocols were adopted for analyzing 14 physical, biological, and chemical soil properties (Gugino et al., 2009; Idowu et al., 2008; Schindelbeck et al., 2008). The physical properties included soil texture (fractions of clay, silt, and sand), net aggregate stability (AGG), residual water (RW), surface hardness (SH), and hydraulic conductivity (HC). The biological properties included soil organic matter (SOM), potential active carbon (PAC) and root health (RH). The chemical properties included pH, electrical conductivity (EC), extractable phosphorus (P), extractable potassium (K), extractable ammonium (NH<sub>4</sub>), and extractable nitrate (NO<sub>3</sub>). All laboratory measurements were performed with CSHT’s standards; however, minor modifications were introduced due to specific management practices, climatic regions, and available tools. These included: (1) wet aggregate stability that was measured by an aggregate stability kit (Herrick, 2000); (2) residual water (RW) that was measured by the soil moisture sensor (Black, 1965); (3) NH<sub>4</sub> and NO<sub>3</sub> that was measured by potassium chlorate extracts (Stevenson, 2005); and finally (4) hydraulic conductivity property that was measured by a mini-disk infiltrometer in the field (Ankeny et al., 1991).

2.3. Spectral measurement and processing

In each of the abovementioned quadrates, four soil samples were collected for spectral measurements. A total of 440 soil samples were separated into two different datasets. The dataset of 2011 (220 samples) was used for model calibration while the remaining 220 samples from 2010 were used for prediction. Prior to the spectral measurements, the soil samples were ground and passed through a 2-mm sieve. Soil samples were measured with the portable Analytical Spectral Devices (ASD) Field Spec® Pro spectrometer with a spectral range of 350–2500 nm and a 25° field of view. Once every few minutes, the spectrometer was calibrated to spectral reflectance using a standard white reference panel (Spectralon Labsphere Inc., North Sutton, NH, USA). Reflectance data were measured under stable illumination from two directions while the spectrometer’s fiber aperture was fixed at a constant height of 20 cm above the sample platform. The bidirectional illumination reduced the effects of micro-topography shadowing. For the same reason, each sample was measured four times, while rotating it 90° between each reading. These four readings were later averaged to represent the sample. The output spectral resolution of the data is 1 nm along the whole spectrum. Finally, the data were resampled homogeneously to 5 nm to minimize noise effects caused by the optical setup while maintaining a high spectral resolution."
optical setup. This transformation was found to be an optimal spectral pre-treatment in similar studies (e.g. Chang et al., 2001; Shepherd and Walsh, 2002; Vågen et al., 2006); (2) an autoscale transformation was performed; this is an exceptionally common pre-processing method for the prediction of known samples from the dataset that was collected (Efron and Gong, 1983) by the toolbox of Eigen-

The prediction process included randomly selecting spectral samples from the 2010 dataset (25%) to calibrate the model for the prediction of known samples from the dataset that was collected in 2011.

2.4. Soil quality index (SQI)

Evaluation of the soil quality was carried out using the general approach of the soil quality indices, involving scoring functions for each of the abovementioned soil properties (Andrews et al., 2004). The scoring functions were defined in a simple nonlinear polynomial framework. Each soil property was transformed through a scoring algorithm into a unitless score (0 to 1) representing the associated level of function in that system so that the scores could be combined to form a single value (e.g. Karlen et al., 2001, 2003). The interpretation of the scoring function was then integrated into an index calculated by a PCA (Efron and Gong, 1983; Masto et al., 2007). The index values ranged from 0 to 1; low values indicated poor soils, while high values indicated healthy soils (Gugino et al., 2009).

Thresholds for each property measurement were set based on the range of values measured in natural ecosystems and on the critical values in the literature. After finalizing the thresholds, the soil property values were recorded by the different algorithms (scoring functions) to transform them to unitless scores (Si) for each soil property, using the following equation (Kinoshita et al., 2012; Masto et al., 2007, 2008):

\[
Si = \left(1 + e^{-\beta(x-a)}\right)^{-1}
\]

where \(x\) is the normally distributed soil property value, \(a\) is the baseline value of the soil property where the score equals 0.5 (infection point) or the population mean of the natural ecosystems, and \(b\) is the slope tangent of the baseline curve. Three types of scoring functions were considered: (1) more is better — an upper asymptotic sigmoid curve (negative slope) that characterizes aggregation, residual water, soil organic matter, potential active carbon, ammonium, nitrate, and potassium; (2) less is better — lower asymptote (positive slope) that characterizes root health and surface hardness; and (3) an optimum midpoint-Gaussian function that characterizes pH, EC, phosphorus, and hydraulic conductivity. The shape curves were determined by the literature (e.g., Masto et al., 2007, 2008). All the soil measurement scores were integrated from the previous stage into a single additive index value termed a soil quality index (SQI) (Eq. (2)). This value is considered to be an overall assessment of soil quality, reflecting management practice effects on soil function. To evaluate the index, the PCA statistical method, a common tool in chemometrics for data compression and information extraction, was used. A PCA finds combinations of variables that describe major trends in the data:

\[
SQI = \sum_{i=1}^{n} PWi \times Si
\]

where \(PWi\) is the PCA weighing factor. Standardized PCAs of all (untransformed) data that differed significantly between treatments in the different land uses were performed using the MATLAB package (Wise et al., 2006). The equation was normalized to obtain a maximum SQI with a score of one. Principal components (PCs) with eigenvalues higher than 1 that explained at least 5% of the variations of the data were examined (Andrews et al., 2002; Masto et al., 2008). Under a particular PC, only variables with high factor loading were retained for soil quality indicating. High factor loading was defined as having an absolute value within 20% of the highest factor loading. When more than one variable was retained under a single PC, a multivariate correlation was employed to determine whether the variables could be considered redundant and, therefore, eliminated from the SQI. If the highly loaded factors were not correlated, then each was considered important and, thus, retained in the SQI. Among well-correlated variables, the variable with the highest factor loading (absolute value) was chosen for the SQI. Each PC explained a certain amount of variation (percent) in the total dataset, and this percentage provided the weight for the variables chosen under a given PC.

2.5. Correlation between soil and spectroscopy analyses

A partial least-squares-regression (PLS-R) cross-validation procedure was used to correlate the spectral data with the laboratory soil measurements. PLS-R is a predictive module technique used in spectroscopy, and it is commonly used for quantitative spectral analysis (Viscarra Rossel et al., 2006; Wold et al., 2001). It can construct predictive models when there are many predictor variables that are highly collinear. The technique is closely related to principal component regression (PCR). However, unlike PCR, the PLS-R algorithm selects orthogonal factors that maximize the covariance between predictor (X spectra) and response variables (Y soil laboratory data or scores). The PLS-R analysis was applied with the full cross-validation of the Ve-netian Blinds method (Efron and Gong, 1983) by the toolbox of Eigenvector© software. In this study, the correlations were between 14 soil
properties and the calibrated dataset. The prediction was evaluated by the Root Mean Square Error of Calibration and Cross Validation (RMSEC and RMSECV), as well as by the coefficient of determination ($R^2$) values of the relation between the predicted and observed soil properties, and by the ratio of performance-to-deviation (RPD), calculated as $\text{RPD} = \text{SD} / \text{RMSECV}$.

In order to evaluate the relative importance of each waveband in each of the PLS-R models, the variable importance in projection (VIP) was computed to reveal the score for each wavelength (Cécillon et al., 2008). VIP scores are a measure of the importance of each explanatory variable (i.e., wavelength). Since the average of squared VIP scores equals 1 (Wold et al., 2001), only influential wavelengths with a VIP score greater than 1 are identified as important wavelengths. The VIP can be used in cases of multi-collinearity among the predictors. The VIP analysis was applied in the MATLAB numerical computing environment using the PLS toolbox of Eigenvector®. A PLS-R cross-validation was also used to correlate the spectral data (indirect spectral measurement) with the SQI. Despite the common scientific use of the PLS-R in spectroscopy and the extensive research into soil quality, to the best of the authors’ knowledge, no such correlations have been carried out previously. The same validating procedures were used as mentioned above.

2.6. Soil spectra classification and spectral soil quality index (SSQI)

A partial least squares-discriminant analysis (PLS-DA) was performed to quantify the changes in soil quality in the changed land uses (afforestation, traditional grazing, and agro-pastoral grazing) and in the different treatments. The PLS-DA is a variant of PLS modeling and aims to find the variables and directions in multivariate space that determine the known classes in a calibration set. The predictor (X spectra) and the response variables are the class membership for each sample (land-use category or treatment). The PLS-DA provides an understandable graphical means of identifying the spectral regions of difference between the classes and also allows a statistical evaluation as to whether the differences between classes are significant. The strength of the model is defined by the $\kappa$ coefficient and the total accuracy calculated from the model’s confusion matrix (Wise et al., 2006). The discriminant analysis has been used in soil analysis to classify soil attributes. Carroll et al. (2006), for example, used a discriminant analysis for the classification of various soil types, based on their respective physical, biological, and chemical characteristics, and to identify relative changes in each soil type after an extended period of application of effluent. In addition, Awiti et al. (2008) used a discriminant analysis for classifying spectral case definitions to define poor or degraded soil classes throughout a tropical forest–cropland. There are two main purposes for which a discriminant analysis is commonly used: (1) to analyze the differences between two or more groups of multivariate data using one or more discriminant functions in order to maximally separate the identified groups; and (2) to obtain linear mathematical functions that can be used to classify the original data or new unclassified data, into the respective groups (Awiti et al., 2008).

The PLS-DA output was used to develop a scoring function in an attempt to evaluate the soil quality only by spectral differences. A proportional odds logistic model was used for evaluating the scores of the spectral SQI from the PLS-DA output. The proportional odds model is based on the cumulative probabilities of the coefficient of variation (CV) and the latent variable (LV). Consequently, the proposed spectral SSQI is a function of the cumulative probability scoring class, ranging from 0 to 1 as in the SQI:

$$\text{SSQI} = \left( e^{\beta_0 + \beta_1 \cdot (CV_{LV})} \right)^{-1}$$

where $T^2$ refers to the Hotelling’s T-squared distribution value (Hotelling, 1931; Wise et al., 2006) that represents a measure of the variation in each sample within the model. The $T^2$ indicates how far each sample is from the center of the model (score = 0) and represents the score distance (SD) within the PLS-DA. $\beta$ is the slope of the function and is calculated from a range of the minimum and the maximum value of the function. Under a particular CV, only the variables with high factor loading were retained for the SSQI, and under a particular LV, only the variables with high factor loading were retained. High factor loading was defined as having an absolute value within 20% of the highest factor loading (Andrews et al., 2002). The cumulative variance of the model is the scores for the individual samples, and the coefficients of the LV are the weighing factors obtained from the PLS-DS model. Each LV explains a certain amount of variation in the total dataset; this percentage provides the weight for variables chosen under a given LV.

2.6.1. Spectral soil quality index (SSQI)

The SSQI is a function of the ratio of the square of the coefficient of variation (CV) and the latent variable (LV). Consequently, the proposed spectral SSQI is a function of the cumulative probability scoring class, ranging from 0 to 1 as in the SQI:

$$\text{SSQI} = \left( e^{\beta_0 + \beta_1 \cdot (CV_{LV})} \right)^{-1}$$

where $T^2$ refers to the Hotelling’s T-squared distribution value (Hotelling, 1931; Wise et al., 2006) that represents a measure of the variation in each sample within the model. The $T^2$ indicates how far each sample is from the center of the model (score = 0) and represents the score distance (SD) within the PLS-DA. $\beta$ is the slope of the function and is calculated from a range of the minimum and the maximum value of the function. Under a particular CV, only the variables with high factor loading were retained for the SSQI, and under a particular LV, only the variables with high factor loading were retained. High factor loading was defined as having an absolute value within 20% of the highest factor loading (Andrews et al., 2002). The cumulative variance of the model is the scores for the individual samples, and the coefficients of the LV are the weighing factors obtained from the PLS-DS model. Each LV explains a certain amount of variation in the total dataset; this percentage provides the weight for variables chosen under a given LV.

Table 3

Soil quality properties for the traditional grazing land use with the following treatments: northern and southern slope with grazing and with no grazing. Statistics include: average value, standard deviation, and significant differences between treatments.

<table>
<thead>
<tr>
<th>Treatments</th>
<th>Sand (%)</th>
<th>Silt (%)</th>
<th>Clay (%)</th>
<th>AGG (1–6)</th>
<th>RW (m/m)</th>
<th>SH (psi)</th>
<th>HC (mm/h)</th>
<th>SOM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No grazing northern slope</td>
<td>35.87 ± 2.22a</td>
<td>48.38 ± 1.41a</td>
<td>15.75 ± 0.96a</td>
<td>4.21 ± 0.33a</td>
<td>1.31 ± 0.13a</td>
<td>295.78 ± 10.41a</td>
<td>0.14 ± 0.04a</td>
<td>6.43 ± 0.51ab</td>
</tr>
<tr>
<td>Grazing northern slope</td>
<td>34.74 ± 1.59a</td>
<td>47.68 ± 0.99a</td>
<td>17.57 ± 2.58ab</td>
<td>4.63 ± 0.34a</td>
<td>1.23 ± 0.12a</td>
<td>279.06 ± 16.28a</td>
<td>0.34 ± 0.04a</td>
<td>5.81 ± 0.83b</td>
</tr>
<tr>
<td>No grazing southern slope</td>
<td>31.34 ± 0.95b</td>
<td>49.46 ± 2.53b</td>
<td>19.20 ± 2.16b</td>
<td>4.76 ± 0.22b</td>
<td>1.36 ± 0.16b</td>
<td>288.89 ± 13.54b</td>
<td>0.08 ± 0.016b</td>
<td>6.59 ± 0.62b</td>
</tr>
<tr>
<td>Grazing southern slope</td>
<td>33.62 ± 2.24ab</td>
<td>46.98 ± 2.30b</td>
<td>19.40 ± 0.89b</td>
<td>4.75 ± 0.26b</td>
<td>1.05 ± 0.19b</td>
<td>244.32 ± 29.27b</td>
<td>0.21 ± 0.03b</td>
<td>6.01 ± 0.97b</td>
</tr>
<tr>
<td>P &lt; α</td>
<td>&lt;0.01</td>
<td>NS</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

Table 4

Soil quality indicators for the agro-pastoral grazing land use with the following treatments: abandoned field with no grazing; monocultural agro-pastoral grazing on wheat field; and abandoned field with grazing. Statistics include: average value, standard deviation, and significant differences between treatments.

<table>
<thead>
<tr>
<th>Treatments</th>
<th>Sand (%)</th>
<th>Silt (%)</th>
<th>Clay (%)</th>
<th>AGG (1–6)</th>
<th>RW (m/m)</th>
<th>SH (psi)</th>
<th>HC (mm/h)</th>
<th>SOM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abandoned field no grazing</td>
<td>49.12 ± 1.34a</td>
<td>37.08 ± 1.09a</td>
<td>13.8 ± 0.44a</td>
<td>2.87 ± 0.19a</td>
<td>2.29 ± 0.16a</td>
<td>311.9 ± 4.16a</td>
<td>0.14 ± 0.05a</td>
<td>3.96 ± 1.05ab</td>
</tr>
<tr>
<td>Agro-pastoral grazing</td>
<td>44.32 ± 1.14b</td>
<td>38.68 ± 1.95b</td>
<td>17.0 ± 1.01b</td>
<td>2.33 ± 0.21b</td>
<td>2.62 ± 0.06b</td>
<td>265.89 ± 12.6b</td>
<td>0.58 ± 0.07b</td>
<td>3.21 ± 0.14b</td>
</tr>
<tr>
<td>Abandoned field with grazing</td>
<td>39.52 ± 4.15a</td>
<td>38.08 ± 3.11a</td>
<td>22.4 ± 2.88a</td>
<td>2.77 ± 0.26a</td>
<td>3.18 ± 0.19a</td>
<td>306.18 ± 8.43a</td>
<td>0.22 ± 0.07b</td>
<td>7.38 ± 1.21a</td>
</tr>
<tr>
<td>P &lt; α</td>
<td>&lt;0.01</td>
<td>NS</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

*Abbreviations: AGG – aggregation; RW – residual water; SH – surface hardness (penetration); HC – hydraulic conductivity (infiltration); SOM – soil organic matter; PAC – potential active carbon; RH – root health; EC – electric conductivity; NH4 – ammonium; NO3 – nitrate; P – phosphorus; K – potassium; NS – no significant differences. Small letters indicate significant differences between treatments. Values in each vertical column followed by the same letter do not differ significantly at P < α using ANOVA Tukey test.
Analyses of variances for all parameters were tested using: (1) a General Linear Model (GLM) analysis of random effect (nested-ANOVA); and (2) a one-way ANOVA for the average sample for each quadrat in a treatment \((n = 5)\). The separation of means was subjected to a Tukey’s test for a significant difference test. A correlation matrix analysis was conducted to identify the relationships between the measured parameters. The statistical analysis was performed with STATISTICA Version 10, 2011 software. The soil quality indices (PCA, PLS-R, and PLS-DA) included the RMSEC and RMSECV, and the confusion matrices were performed in MATLAB Version 7, 2011 software with a PLS toolbox (EIGENVECTOR research) and using Microsoft Excel packages. Soil quality properties and soil quality indices were tested for their level of significance at \(P = 0.05\) between changed land uses and treatments and by the result of the F-statistic test.

### 3. Results and discussion

#### 3.1. Soil quality indices (SQIs)

The mean values of all the soil properties across the three land-use types are presented in Tables 2–4, along with their standard deviations and significance values. The afforestation soil properties are shown in Table 2. The results of the afforestation land use were significantly higher than the shrubland system in the AGG, SOM, and PAC properties. In addition, significant differences in soil texture were observed from the silt-loam soil in the natural shrubland to the loam soil in the forest. The soil with a higher clay content in the forest has a higher ability to retain nutrients (higher cation exchange capacity) and can bind more organic matter (Idowu et al., 2008). The results in the forest understory and under the shrubs in the shrubland, with respect to the open patches, showed significant increases in AGG, RW, SOM, EC, NO₃⁻, and P. A high negative correlation was found between SH and HC (\(R^2 = 0.72\); \(P < 0.01\)) and between SOM and AGG (\(R^2 = 0.87\); \(P < 0.01\)); these relations were found and explained in earlier studies (e.g., Carter, 2002; Dexter, 2004). To avoid properties that could be considered as redundant, the multivariate correlation was tested for \(R \geq 0.8\). Then, the soil properties with a high factor loading were eliminated from the SQI (Masto et al., 2007, 2008).

The SQI results from the afforestation land-use type are shown in Fig. 3. The SQI score in the forest understory is 0.62, while in the forest open patches, it is 0.67. The natural adjacent shrubland exhibits opposite trends in which the high SQI values are in the natural shrub area with a significant change in the soil texture was found (due to a higher clay component). The abandoned field with grazing, compared to the abandoned field with no grazing, showed significant increases in the soil properties of SOM, EC, NO₃⁻, P, and K. On the other hand, significant reductions in RW, NH₄⁺, and pH were found.

#### Table 2

<table>
<thead>
<tr>
<th>PAC (ppm)</th>
<th>RH (1–9)</th>
<th>pH</th>
<th>EC (dS/m)</th>
<th>N (NH₄⁺) (mg/kg)</th>
<th>N (NO₃⁻) (ml/kg)</th>
<th>P (mg/kg)</th>
<th>K (mg/kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>567.9 ± 57.7a</td>
<td>1.3 ± 0.4a</td>
<td>7.08 ± 0.11a</td>
<td>0.85 ± 0.12a</td>
<td>71.7 ± 11.06a</td>
<td>8.465 ± 3.05a</td>
<td>14.18 ± 6.53a</td>
<td>10.96 ± 2.65a</td>
</tr>
<tr>
<td>460.87 ± 60.1b</td>
<td>2.25 ± 0.58b</td>
<td>6.95 ± 0.09b</td>
<td>0.82 ± 0.09b</td>
<td>79.61 ± 11.36b</td>
<td>10.36 ± 4.59b</td>
<td>8.25 ± 3.5b</td>
<td>9.06 ± 2.31b</td>
</tr>
<tr>
<td>588.09 ± 64.1a</td>
<td>1.7 ± 0.4a</td>
<td>7.02 ± 0.05a</td>
<td>0.66 ± 0.04b</td>
<td>57.51 ± 11.03b</td>
<td>8.14 ± 1.88b</td>
<td>11.01 ± 3.16b</td>
<td>8.71 ± 1.03b</td>
</tr>
<tr>
<td>466.65 ± 82.01b</td>
<td>2.75 ± 0.55a</td>
<td>6.97 ± 0.05b</td>
<td>0.68 ± 0.05b</td>
<td>51.39 ± 9.28a</td>
<td>8.26 ± 1.54a</td>
<td>8.69 ± 3.1b</td>
<td>7.96 ± 1.41b</td>
</tr>
<tr>
<td>&lt;0.01</td>
<td>NS</td>
<td>&lt;0.01</td>
<td>&lt;0.05</td>
<td>&lt;0.01</td>
<td>NS</td>
<td>&lt;0.05</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

The results of the soil quality indices in the traditional grazing land use are presented in Fig. 3. The SQI scores in the no-grazing treatments in the northern- and southern-facing slopes are 0.70 and 0.67, respectively, and are significantly higher than those in the grazing treatments (\(F_{1, 79} = 21.55; P < 0.01\)). The scores of the grazing treatments on the northern- and southern-facing slopes are 0.61 and 0.59, respectively. The changes in soil quality caused a reduction in the functionality of the soil after long-term livestock grazing in the traditional grazing land use in semi-arid areas. More results and detailed interpretations can be obtained at Paz-Kagan et al. (2014).

#### Table 4

<table>
<thead>
<tr>
<th>PAC (ppm)</th>
<th>RH (1–9)</th>
<th>pH</th>
<th>EC (dS/m)</th>
<th>N (NH₄⁺) (mg/kg)</th>
<th>N (NO₃⁻) (ml/kg)</th>
<th>P (mg/kg)</th>
<th>K (mg/kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>547.52 ± 60.2a</td>
<td>2.1 ± 0.31a</td>
<td>7.14 ± 0.08a</td>
<td>0.73 ± 0.12b</td>
<td>264 ± 10.79b</td>
<td>8.68 ± 6.03b</td>
<td>8.72 ± 3.35b</td>
<td>16.93 ± 3.01b</td>
</tr>
<tr>
<td>564.55 ± 64.2a</td>
<td>2.3 ± 0.47b</td>
<td>7.06 ± 0.09b</td>
<td>0.89 ± 0.12b</td>
<td>43.188 ± 10.1b</td>
<td>31.43 ± 11.7b</td>
<td>8.14 ± 1.47b</td>
<td>10.63 ± 6.52b</td>
</tr>
<tr>
<td>456.52 ± 53.2b</td>
<td>2.7 ± 0.47b</td>
<td>7.01 ± 0.1b</td>
<td>1.04 ± 0.114a</td>
<td>17.73 ± 9.68b</td>
<td>24.63 ± 7.93b</td>
<td>14.2 ± 4.52a</td>
<td>34.72 ± 11.26b</td>
</tr>
<tr>
<td>&lt;0.01</td>
<td>NS</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>NS</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

The results of the SQI in the monocultural agro-pastoral system (wheat field) is 0.67; in the abandoned field with no grazing, the score is 0.66, and in the abandoned field with grazing, the score is 0.54 and is significantly lower than the agro-pastoral and abandoned field with no grazing scores \(F_{1, 77} = 68.21; P < 0.01\) (Fig. 3). The monocultural agro-pastoral treatment showed higher sustainability to grazing than the monocultural agro-pastoral system (wheat field).
the abandoned field with grazing. The SQI model showed that long-term management of a land-use type changes the soil quality. More results and detailed interpretations can be obtained at Paz-Kagan et al. (2014).

In summary, the SQI equation is valid for establishing the degree of the degradation of soils as a function of soil properties. In addition, it enables the identification of changes in the soil properties and the monitoring of them. However, it is essential that these indices be validated under various other land management systems before further use.

3.2. Correlation between laboratory and spectroscopy analyses

Reflectance spectroscopy is directly influenced by the combinations and overtones of the fundamental vibrations of mineral and organic functional groups, as well as electronic transitions (Awiti et al., 2008). Table 5 presents the results of the PLS-R analysis in terms of wavelengths that are indicative for each of the soil properties. Statistics include the coefficient of determination, the RMSEC, the RMSECV, and the RPD. The RPD was calculated since RMSEC and RMSECV alone cannot provide sufficient information on model predictability due to the variable standard deviation of the soil properties, while the RPD can be compared across the soil properties measured in different units. No critical value exists for the RPD in soil science, but an RPD > 2 is denoted as satisfactory (Chang et al., 2001). Three categories of predictability were suggested by Chang et al. (2001); these are category A (R² > 0.80; RPD > 2.00), category B (0.50 < R² < 0.80; 1.40 < RPD < 2.00), and category C (R² < 0.50; RPD < 1.40). In addition, Table 5 summarizes the examples from the literature review of indicative spectral regions for each soil property, along with the best reported correlation of determination. One may note that many of the PLS-R-derived wavelengths are in agreement with those that were previously found in other studies.

Fig. 4 shows scatterplots of correlations between soil spectroscopy and laboratory-measured soil values of several soil properties for the calibration dataset with a coefficient of determination range between 0.72 and 0.91 (Fig. 4 represents the results by RMSC, RMSCV, R², RPD and number of LV for each soil property). The highest results of the PLS-R model prediction were associated with the RPD category A (R² > 0.80; RPD > 2.00), including sand–silt–clay content, NH₄, NO₃, and pH. Satisfactory results of category B (0.50 < R² < 0.80; RPD > 2.00) were obtained for RW, SOM, EC, and potassium. Poor validation results of category C (R² < 0.50; RPD < 2.00) were obtained for PAC, phosphorus, and HC (Table 5, Fig. 4).

Among the soil physical properties, residual water was found to be sensitive at 1200, 1450, 1962, and 2200 nm. The soil particle size that influences light scattering, and soil texture (sand, silt, and clay) is sensitive at 1900 and 2000–2200 nm (Islam et al., 2003). The VIS–NIR–SWIR, characterized by two strong absorption features at 1400 and 1900 nm, can usually be visible based on the specific surface area (SSA) content of the solid soil phase that eventually controls the hygroscopic moisture content (residual water, RW). These wavelengths are related to the O–H bands (Banin and Amiel, 1970; Bowers and Hanks, 1965; Chang et al., 2001; Confalonieri et al., 2001; Dalal and Henry, 1986; Demattê et al., 2006; Lagacherie et al., 2008; Moron and Cozzolino, 2003). The water absorption wavelength shows moderate or weak absorption features in the SWIR region due to clay minerals (2200 nm), carbonate (2330 nm), salt or primary minerals, and organic compounds in the soil (Ben-Dor and Banin, 1995; Ben-Dor et al., 2008; Chang et al., 2001; Dalal and Henry, 1986; Malley et al., 2004). NIR is directly influenced by the organic matter and bound water (Ben-Dor et al., 2008).

With respect to soil biological properties, the soil organic matter is sensitive at 1350–1450 and 2200 nm. These wavelengths are related to the C–H absorption in the spectral wavelength (e.g. Ben-Dor and Banin, 1994; Ben-Dor et al., 2008; Confalonieri et al., 2001; Haberern, 1992; Malley et al., 2002; Rinnan and Rinnan, 2007). In relation to soil chemical properties, the mineral N (NH₄ and NO₃) was found to be sensitive at 1700–1800 and 2180–2270 nm (Morra et al., 1991; Rinnan and Rinnan, 2007), although it is characterized by poor predictability (Malley et al., 2002). Theoretical second- and third-order bands exist in the NIR region for N bonds of this nature. Possibly, the absorbance is interfered by water, organic matter, and/or iron oxides. The pH was found to be sensitive at 1477, 1932, and 2200 nm; measurement of pH using NIR spectroscopy has generally been well-predicted in the literature (Table 5). Hydrogen ions are not primary absorbers in the NIR region, but the basis of prediction may be through O–H groups or other basic ions, such as carbonate (Malley et al., 2004; Mouazen et al., 2007; Reeves and McCarty, 2001). The EC does not have spectral absorption bands; however, it can be evaluated from indirect spectral absorption bands of gypsum and salts in the soil at 1760–1780 and 1830–1970 nm (Ben-Dor et al., 2008, 2009; Farifteh et al., 2008; Malley et al., 2004). There are several tentative wavelengths in the NIR region for compounds involving P. The majorities of these wavelengths are...
third order, with some being second or first order. The prediction of K was relatively high ($R^2 = 0.76$) and included sensitivity bands 1850, 1940, 2180, and 2290 nm. K is generally not amenable to analysis by NIR spectroscopy (Malley et al., 2002), although it has been reported to have a high spectral correlation (Daniel et al., 2003).

### 3.2.1. Correlation between SQI and spectroscopy

Scatterplots of the calculated SQI versus spectral measurements in the changed land uses, along with their RMSEC, RMSECV, and RPD, are presented in Fig. 5. Coefficients of determination range between 0.65 and 0.8. The highest validation of the PLS-R model prediction value was obtained for the afforestation land use ($R^2 = 0.81$; RPD = 2.51), lower for the traditional grazing land use ($R^2 = 0.75$; RPD = 2.05), and the lowest for the agro-pastoral land use ($R^2 = 0.65$; RPD = 2.01). Our results show that reflectance spectroscopy in the VIS, NIR, and SWIR spectral ranges can be used for assessing various soil properties and for calculating the SQI for land management. Furthermore, after finalizing the SQI, the use of the PLS-R technique is rapid, making it possible to analyze a large number of samples in a practical and timely manner as a prediction model. These merits make spectroscopic analysis, combined with PLS−R, an attractive method for environmental monitoring, especially for modeling soil quality in land−use changes. The disadvantage with this procedure is that it requires verification of the SQI with direct soil measurements. As mentioned previously, laboratory analyses are complex, time and labor consuming, and expensive.

### 3.3. Soil spectra classification in land-use types

The proportional odds in the PLS-DA classification of the spectral samples, in each treatment within each land use, during 2010–2011, are presented in Fig. 6A−C, and in all land uses in both years in Fig. 6D. Table 6 shows the Kappa coefficient, the total accuracy of the model, and the prediction model of unknown samples (prediction of samples from the 2010 dataset is based on the 2011 dataset). Fig. 7 shows the PLS-DA classification of the land use and the prediction model of unknown samples. The PLS-DA provides an explicitly quantitative approach to predict the cumulative probability of soil spectral samples that belong to different soil conditions.

### Determination of the optimal number of PLS LVs

The optimal number of PLS LVs was performed by the cross-validation procedure. Three LVs were selected since they have eigenvalues higher than 1. These eigenvalues explain at least 5% of the variation of the data. The total accuracy of the model was calculated to be of a Kappa coefficient >0.92 and a total accuracy >0.9 and, in the prediction of unknown samples, of a Kappa coefficient >0.75 and a total accuracy >0.7 (Table 6). The PLS-DA model describes the maximum possible separation of predefined soil conditions. This technique can

### Table 5

Results of partial least squares-regression (PLS-R) analysis in terms of spectral regions that are indicative for the soil properties. Statistics include the correlation of determination ($R^2$), the root mean square error of calibration (RMSEC), the root mean square error of cross-validation (RMSECV), and the ratio of performance-to-deviation (RPD). Examples of literature findings for indicative spectral regions are provided along with the best reported correlations of determination by different models. Bold numbers refer to agreements between the present spectral regions and those that are known from the literature.

<table>
<thead>
<tr>
<th>Soil indicators</th>
<th>$R^2$</th>
<th>RMSEC</th>
<th>RMSECV</th>
<th>RPD</th>
<th>Sensitivity bands identified</th>
<th>$R^2$ from literature</th>
<th>Sensitivity bands from literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sand (%)</td>
<td>0.91</td>
<td>1.3</td>
<td>1.78</td>
<td>3.21</td>
<td>1802, 1872, 1870, 2067, 2159, 2249, 2362</td>
<td>0.82 (Chang et al., 2001)</td>
<td>2067, 2159, 2249, 2362</td>
</tr>
<tr>
<td>Clay (%)</td>
<td>0.80</td>
<td>0.52</td>
<td>1.02</td>
<td>2.03</td>
<td>2097, 2139, 2232, 2278, 2329</td>
<td>0.74 (Ben-Dor and Banin, 1995)</td>
<td>2067, 2159, 2249, 2362</td>
</tr>
<tr>
<td>RW</td>
<td>0.77</td>
<td>0.72</td>
<td>0.78</td>
<td>2.19</td>
<td>1450, 1750, 1850, 1962, 2000</td>
<td>0.98 (Denatté et al., 2006)</td>
<td>1850, 1962, 2000</td>
</tr>
<tr>
<td>PAC</td>
<td>0.42</td>
<td>110.84</td>
<td>119.5</td>
<td>1.32</td>
<td>592, 872, 1887, 2217</td>
<td>0.86 (Kimoshita et al., 2012)</td>
<td>1887, 2217</td>
</tr>
<tr>
<td>NH$\text{}_4$</td>
<td>0.83</td>
<td>9.81</td>
<td>11.03</td>
<td>2.43</td>
<td>592, 872, 1887, 2217</td>
<td>0.74 (Reeves and McCarty, 2001)</td>
<td>1887, 2217</td>
</tr>
<tr>
<td>NO$\text{}_3$</td>
<td>0.81</td>
<td>3.32</td>
<td>3.78</td>
<td>2.68</td>
<td>562, 1282, 1422, 1752, 1962, 2217</td>
<td>0.74 (Reeves and McCarty, 1999)</td>
<td>1887, 2217</td>
</tr>
<tr>
<td>pH</td>
<td>0.80</td>
<td>0.107</td>
<td>0.132</td>
<td>2.29</td>
<td>517, 657, 747, 1477, 1492, 1932, 2062, 2227</td>
<td>0.7 (Shepherd and Walsh, 2002)</td>
<td>1887, 2217</td>
</tr>
<tr>
<td>EC</td>
<td>0.75</td>
<td>0.086</td>
<td>0.101</td>
<td>2.94</td>
<td>572, 847, 992, 1202, 1537, 1622, 1767, 1942, 2222, 2352, 2397</td>
<td>0.72 (Moran and Cozzolino, 2003)</td>
<td>1887, 2217</td>
</tr>
<tr>
<td>Potassium</td>
<td>0.76</td>
<td>3.45</td>
<td>4.40</td>
<td>2.09</td>
<td>537, 1542, 1862, 1947, 2187, 2297</td>
<td>0.72 (Moran and Cozzolino, 2003)</td>
<td>1887, 2217</td>
</tr>
<tr>
<td>Phosphorous</td>
<td>0.36</td>
<td>1.76</td>
<td>1.95</td>
<td>1.98</td>
<td>537, 1542, 1862, 1947, 2187, 2297</td>
<td>0.81 (Daniel et al., 2003)</td>
<td>1887, 2217</td>
</tr>
<tr>
<td>Hydraulic conductivity</td>
<td>0.47</td>
<td>0.11</td>
<td>0.12</td>
<td>1.51</td>
<td>2187, 2297</td>
<td>0.80 (Confalonieri et al., 2001)</td>
<td>1887, 2217</td>
</tr>
</tbody>
</table>

### References

Fig. 4. Scatterplots of cross-validation (CV) predicted values versus measured values for several soil properties for the calibration dataset for all land uses. Calibration models were developed with partial least-squares regression. RMSEC = root mean square error of calibration; RMSECV = root mean square error of cross-validation. RW — residual water; SOM — soil organic matter; NH₄⁺ — ammonium; EC — electric conductivity; NH₃ — nitrate.

Fig. 5. Scatterplot correlation of soil quality indices (SQIs) and reflectance spectroscopy values for the changed land uses. Calibration models were developed with a partial least-squares regression.
also be useful to determine that a sample does not belong to any of the predefined classes. Furthermore, these results demonstrate the sensitivity of reflectance spectroscopy to management changes in converted land-use soil conditions.

3.4. Spectral soil quality index (SSQI)

The soil condition scores in the different land uses were well separated, and each sample represents SSQI scores. The comparison of the SSQIs to the SQIs is shown in Figs. 8–9. The calibration sets between the SSQI and the SQI were, in the afforestation land use, $R^2 = 0.66; P < 0.01$, in the traditional grazing land use, $R^2 = 0.67; P < 0.01$, and in the agro-pastoral land use, $R^2 = 0.74; P < 0.01$. The SSQI indicates higher values than the SQI; the model of the SSQI is a proportional model that is not based on the individual probability of each class but on the cumulative probabilities. Therefore, the proportions between classes that explain the changes caused by management are more essential than the actual values.

Although there have been several attempts to predict soil physical, biological, and chemical properties using reflectance spectroscopy (Cécillon et al., 2009a), the current study pioneers the use of PCA, PLS-R, and PLS-DA of the VIS–NIR–SWIR spectra for assessing the soil

![image](image.png)

**Table 6.** Accuracy assessment of the soil spectral classification in each treatment and in all land uses.

<table>
<thead>
<tr>
<th>Land use</th>
<th>Treatments</th>
<th>Kappa coefficient model</th>
<th>Total accuracy model</th>
<th>Kappa coefficient prediction</th>
<th>Total accuracy prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afforestation</td>
<td>Treatment 1–4 Forest canopy; forest open; natural shrub; natural open</td>
<td>0.94</td>
<td>0.91</td>
<td>0.75</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>Traditional grazing Treatment 1–4</td>
<td>0.93</td>
<td>0.90</td>
<td>0.88</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>Agro-pastoral system Treatment 1–3</td>
<td>1</td>
<td>1</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>All land uses Treatment 1–3 Traditional grazing, agro-pastoral and afforestation</td>
<td>1</td>
<td>1</td>
<td>0.97</td>
<td>0.95</td>
</tr>
</tbody>
</table>
quality. The ability of the model to predict soil quality or functionality by spectroscopy is an essential tool for soil monitoring (Awiti et al., 2008; Cécillon et al., 2008, 2009a,b). In addition, the model is capable of evaluating the SSQI between the different land uses, a comparison that is not possible in the SQI due to the need to specify a land-use threshold (Fig. 10). The results show that the afforestation, traditional grazing, and agro-pastoral grazing land uses have SSQI scores of 0.78, 0.87, and 0.57, respectively. The differences between the SSQIs of the land uses are significant ($F_{(2,216)} = 128.49; P < 0.01$).

Assessment of soil conditions with reflectance spectroscopy, as presented above, enables the rapid tracing of the states of soil quality or of its changes after long-term management (land-use change). In addition, successful classifications of sites regarding the land use and soil conditions were performed. This SSQI can be used for assessing hot spots of land-use changes and for identifying soil degradation. However, in order to examine the causes of the change in soil quality, much more detailed research should be done on direct soil measurements, in terms of laboratory analyses as inputs for models such as the SQI.

Soil quality policies usually address specific management goals, such as productivity, waste recycling, and environmental protection (Andrews et al., 2004). Thus, methodologies for soil quality assessment should be able to measure the specific soil functions and the soil ecosystem services associated with these management goals. Assessing SSQI by VIS–NIR–SWIR spectroscopy as a preliminary tool for detecting hot spots of soil degradation can provide an alternative to laboratory analyses. The benefits of this technique include a reduction of the sampling processing time and an increase in the number of samples that can be analyzed within time and budget constraints, and hence, an improvement of the detection of changes in soil quality in a given area. It is a rapid, non-destructive, reproducible, and cost-effective analytical method, and therefore, it is a promising tool for soil quality assessment. However, the challenge is to move from the field and the laboratory scale to a larger scale for the effective monitoring of soil quality with airborne or spaceborne spectroscopy. In addition, assessments of soil conditions at the regional scale across various soil types are needed. This model is based on laboratory measurements under controlled conditions that avoid the disruptive factors that characterize field measurements, such as soil moisture content, soil roughness, and vegetation cover (Stevens et al., 2008). There is a need for additional studies to identify the potential of the SSQI in a spatial resolution for large ranges of spatial domains, from laboratory point measurements, through ground-level and airborne, and up to spaceborne images. In addition, there is a need to examine the model’s capability in different soil types and climatic regions.

4. Conclusions

This study underscores the potential application of reflectance spectroscopy as a reliable diagnostic screening tool for assessing soil quality. Classification of soils into spectrally defined entities provides a basis for spatially explicit and quantitative definitions for developing SSQI. This paper proposes a framework for the use of VIS–NIR–SWIR spectroscopy as a tool for assessing soil quality. The utilized linear parametric models, PCA, PLS-R, and PLS-DA, improved the ability to predict soil properties and SQI, and they enabled us to overcome problems of multivariate...
and high co-linearity data as hyperspectral spectroscopy. The ability and motivation to assess soil quality varies among the different indices:

(1) The SQI was found to be a good tool for diagnosing the degradation or amplification of different soil properties and to identify the critical changes in soil functionality. The SQI is an integrative approach that recognizes the physical, biological, and chemical processes in soils. In addition, it requires many soil analyses; monitoring such soil quality indices at different scales and land uses is expensive and time consuming. This problem can be partially solved by combining field and laboratory measurements with reflectance spectroscopy methods and models, such as PLS-R.

(2) The SSQI is a diagnostic tool for assessing soil quality in changed land uses and treatments. The ability of reflectance spectroscopy was proved to be a reliable diagnostic tool for identifying and separating various soil properties, for identifying SQI in areas of land-use change, and for classifying soil conditions. The SSQI can be used to assess hot spots of change in different land-use areas and to identify soil degradation. However, examining the cause of the soil degradation requires an extensive assessment of soil quality by soil measurements and the use of models such as SQI.

Our findings could have strong implications regarding the monitoring of soil quality for changed land-use assessments. Future studies should test the success of this framework for larger and more variable datasets, including different soil types and climate regions. In addition, the challenge is to move from the datasets, including different soil types and climate regions. In addition, one should test the success of this framework for larger and more variable soils. Future studies can provide new information that cannot be extracted by field work that uses traditional soil sampling or point spectrometry measurements. The advantages of using reflectance spectroscopy to detect soil quality and degradation processes in land-use changes are highly important. Thus, methods may expand to image spectroscopy and to the fields of soil surveying and soil mapping in the future.

Acknowledgments

The research was funded by the FP7 Experimentation in Ecosystem Research (ExpeER) project. The research leading to these results has received funding from the ExpeER Programme of the European Union’s Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no. 262060.

References


Ben-Dor, E., et al., 2009. Using imaging spectroscopy to study soil properties. Remote Sens. Environ. 113 (Supplement 1(0)), S38–S55.


Gugino, B.K. et al., 2009. Corn Soil Health Assessment Training Manual. New York Agricultural Experiment Station Cornell University, New-York, US.


