Knowledge fusion in the Agro-environmental Field: A Global Index for Soil Quality in Olive groves from Quantitative and Qualitative Variables

Victor Aranda*, Julio Calero*, Francisco Comino†, Arturo Montejo‡, Jose-Maria Serrano‡[1]

*Dept. of Geology
University of Jaen (Spain)
varanda@ujaen.es, jcalero@ujaen.es

†Dept. of Physical and Analytical Chemistry
University of Jaen (Spain)
fcr00011@estudiante.ujaen.es

‡Dept. of Computer Science
University of Jaen (Spain)
amontejo@ujaen.es, jschica@ujaen.es

***Abstract*—This work presents a synthetic, automated global soil quality index (GSQI), applied in the province of Jaen (Spain). This index has been built following a novel methodology, able to integrate quality indicators of both morphological (representing imprecise concepts) as analytical types, against other measures previously established in the bibliography. This new methodology is based on an advanced data analysis statistical technique, Categorical Principal Component Analysis (CatPCA), applied to non numerical (ordinal or nominal) morphological indicators, properly transformed by means of optimal scaling, quantified and managed along with numerical, analytical indicators. Individual (morphological and analytical) indicators integration into a global index has been locally validated with olive grove soil samples from Atanor Valley (Sierra Magina Natural Park, Spain). This validation process permits the quality comparison between conventional managed olive soils and organic managed ones, the latter clearly showing a better value. This procedure means a great advance in the establishment of a quality improvement for soil degradation due to olive grove conventional management, encouraging more sustainable, soil-protecting, agricultural practices. The discussed methodology can be applied in other highly cultivated geographical areas in order to improve sustainability in agricultural systems.**

***Keywords*—Soil quality index, Morphological soil properties, Categorical principal component analysis, Scoring functions.**

I. Introduction

Fusion of analytic and linguistic data, usually found in agrosystems, obtained from laboratory analysis, field morphological descriptions, or even surveys, is of special interest in a scientific-technical environment, its final objective being the discovery of useful knowledge for both agricultural technics and farmers. It is necessary to properly manage and analyze this information, due to its high variability and uncertainty, inherent to its spatial distribution in agricultural groves and field sampling and description.

A specially interesting issue, from the agro-evironmental point of view, is that of soil quality, very related with productivity, fertility, and soil degradation, as well as environment quality [1]. In the case of agrosystems, soil quality means to indefinitely mantain soil as a crop resource, avoiding degradation.

As mentioned in [2], a soil quality indicator can be considered as a feature related with one or more soil functions, very conditioned in turn by environmental factors. Moreover, it must be correlated with an important number of physical, chemical or biological functions, easy to measure and respond to changes in management. Indicators should explain as much possible variability of the studied system; present significantly different values in systems undergoing changes in management, have a quick response to these changes, or the agricultural system as a whole, requiring a long recovery time to the initial state, and should be easily measurable and non-redundant [3], [4]. This set of indicators is called the minimum data set, and defines the operation of a system optimally and with minimal loss of information. The process of quantifying indicators is called scoring. Different types of scoring functions for key soil indicators generate linear or curve (non-linear) funtions, which in turn represent the algorithms that best define the relationship between the indicator numerical domain and the function evaluating soil [5]. Furthermore, by means of scoring the indicator values are normalized into a range of 0.1 to 1, the latter value corresponding to the optimal quality. Each indicator has its best scoring function and in many cases it is determined by the researcher based on his expert knowledge, and therefore of a somewhat arbitrary way.

Many soil quality indicators that can be useful are of morphological type (e.g., structure, consistency, color, etc...), and being not numerically qualitative, and therefore with a high degree of uncertainty and imprecision in the measurement, have been little used, assuming a significant loss of information. By means of Nonlinear Principal Component Analysis (NLPCA), and one of its main algorithms, Categorical Principal Components Analysis (CatPCA), it is possible to manage non-numeric variables, as applied for first time in [6]. Thus, non-numeric variables, through a procedure

[1]Corresponding author

of optimal scaling, are quantified and managed along with numeric properties commonly used in classic PCA. On the other hand, maximizing the amount of explained variance, the number of indicators to be used can be reduced, selecting the most important, and hence optimizing the process of scoring nominal and ordinal indicators automatically, along with the numeric ones. Finally, the numerical assessment of soil quality is the result of integrating individual indicators into a single value (indexing), as the result of mathematical procedures that integrate the previously selected and quantified (in the case of morphological) indicator variables.

In [7] and [8] it is shown that ecological crop management affects positivelty soil quality, and this induces an olive tree health improvement, an important aspect regarding the system agroecological sustainability. Nevertheless, these works use sophisticated techniques (e.g., infrared spectroscopy, X-ray diffraction, pyrolisis, etc.), completely out of range for farmers. The main purpose of this paper is the development and validation of a new global soil quality index (GSQI), based on the use of CatPCA over both morphological (categorical) and analytical (numerical) indicators, obtained from a regional database. This index allows us to evaluate soil quality in a easy, low cost, way, with the same reliability as other sophisticated measures in the field of advanced scientific research, according to farmers available measures (usual analysis, field linguistic descriptions). It aims to establish a comprehensive and automatic procedure for computing a soil quality index able to be integrated in a user-friendly Decision Support System. Validation will be performed on olive grove agrosystem in Jaen (Spain), where it is present a huge negative impact on soil quality, mainly soil erosion, caused by conventional intensive agriculture carried out for decades.

II. Material and methods

A. Data sources

A soil database with 131 profiles was made from the combination of 10 existing soil surveys [9]–[19]. On this database, a total of 41 soil quality indicators (SQI) was present, from which 18 indicators were field soil morphological indicators and 23 were analytical soil quality indicators. The soil quality index was generated only from superficial horizons (*Ah* and *Ap* types) with mean thickness of 18 cm. Regarding Land Use Types (LUTs), 8 types were considered, characterized by the classification scheme proposed by FAO [20] including modifiers related to crop type, vegetation classification and human influence. These types are: (1) Mediterranean Xeromorphic Woodland (*MXW*), (2) Mediterranean Xeromorphic Scrub (*MXS*), (3) Little-disturbed Forest (*LDF*), (4) Pine Plantation Forestry (*PPF*), (5) Olive groves including Conventional Olive Groves (*COG*), (6) Alpha Grass communities (*AG*), (7) Pastures and degraded grassland (*PDG*), and (8) frequent but disperse Herbaceous Annual Cultures (*HC*).

B. Global Soil Quality Index (GSQI) calculation

To obtain and validate the final soil quality index proposed in this work, a procedure divided into 6 different steps is applied:

Soil indicator	**Type**	**Scale**
$Sand(\%)$	Analytical	Number
$W33(\%)$	Analytical	Number
$O.C.(\%)$	Analytical	Number
$W1500(\%)$	Analytical	Number
$N_{total}(\%)$	Analytical	Number
$CaCO_{3eq}(\%)$	Analytical	Number
$CEC\ (cmol+kg^{-1})$	Analytical	Number
$Dry\,value$	Morphological	Ordinal
$Dry\,hue$	Morphological	Nominal
$Root\,abundance$	Morphological	Ordinal
$Texture$	Morphological	Nominal
$Clay(\%)$	Analytical	Number
$Bulk\,density(gcm^{-3})$	Analytical	Number
$Stickiness$	Morphological	Ordinal
$Moist\,value$	Morphological	Ordinal

TABLE I. Selected morphological and analytical indicators in the building of GSQI. Abbreviations: $W33$: Moisture content at field capacity; OC: Organic carbon; $W1500$: Moisture content at permanent wilting point; N_{total}: Total nitrogen; $CaCO_{3eq}$: Carbonates; CEC: Cation exchange capacity

1) Indicator selection: Minimum data set: According to the correlation matrix obtained on an exploratory CatPCA (correlations between indicators higher than $r = \pm 0.5$), most relevant soil indicators were selected from the 41 available. Table I lists these 15 relevant indicators.

2) Indicator scaling: Using CatPCA [21] nominal and ordinal morphological indicators were transformed into numerical ones. Also, optimal scalings y_j, Variances Accounted For per variable (VAF_j) and per component (VAF_s) were obtained in the process (being s the number of dimensions in CatPCA). Its fitness was measured by Cronbach's α, with a value near to 1, ensuring a valid percentage of variance kept by the model [22].

3) Interpretation of PC in terms of soil quality: With the generated model, it is possible to interpret the resulting components in terms of soil quality [23]. With the lowest P-value in a Kruskal-Wallis test, it was found that the soil quality component holds those scores that best discriminate between land use types.

4) Normalized scores: We need to obtain a score s_{ij} for each horizon i and indicator j, ranging from 0.1 to 1, so horizon quantifications q_{ij} had to be normalized. The loadings of the model for each q_{ij} indicate whether the component increases or decreses soil quality. In the former case, a *more is better* function is applied:

$$s_{ij} = s_{more\_is\_better}(q_{ij}) = 0.1 + \left(\frac{q_{ij} - y_{jmin}}{y_{jmax} - y_{jmin}}\right) \times 0.9 \quad (1)$$

In the second case (when the component decreses soil quality), a *less is better* function is used:

$$s_{ij} = 1.1 - s_{more\_is\_better}(q_{ij}) \quad (2)$$

where y_{jmin} and y_{jmax} are the minimum and maximum category quantification values of y_j (y_{jmin} and y_{jmax} should be considered as the reference values from [24]).

5) GSQI calculation: Now all the needed elements are under appropriate constraints to compute the horizon quality index $GSQI_i$ as a linear combination of scores and weights. The weighting factors are selected according to certain strategies, detailed by [5], [25], and applying the method explained in [26]. The w_j weight factor is the VAF_{js}, i.e. the vector coordinate of indicator j in the soil quality component. Thus, GSQI is calculated for the horizons in the soil data as specified in the following equation, with a subsequent re-scale to unity:

$$GSQI_i = \sum_{j=1}^{m} s_{ij} w_j \tag{3}$$

6) GSQI validation: It only remains to check the validity of the obtained index. To this end, olive grove data not included in the original data set has been used. These measurements comes from Atanor valley [8], providing data about 10 A type horizons from two different olive soil treatments: *COG* and Organic Olive Groves (*OOG*), with N = 20, the latter being not included in the original data set due to its lack of extension, making it not statistically relevant in previous studies. The data was sampled and the soil properties collected from estimations. The obtained categories for each horizon were transformed according to optimal scalings y_j from CatPCA, so normalized scores were produced. Using the same weight factors and these scores, final GSQI values were calculated. Student's-t test ($P < 0.05$) was applied to obtain statistical differences between the means of GSQI for both treatments.

III. Index Development

A model of 3 principal components with eigenvalues higher than one was selected. It gives to a good fit, with a Cronbach's α of 0.971 [21]. Table II shows the component loadings and variances explained by components and individual variables. The 3 components combined explained as much as 71% of system variance. This improves the total percentage of system variance with respect to previous studies by [27], [28].

Principal Component	PC1	PC2	PC3	VAF_j	$PVAF_j$
Analytical					
$Sand(\%)$	-.096	.851	-.121	0.748	4.986
$W33(\%)$	-.624	-.663	-.122	0.843	5.622
$O.C.(\%)$	**-.868**	-.033	-.108	0.767	5.111
$W1500(\%)$	-.585	-.657	-.170	0.804	5.357
$N_{total}(\%)$	-.163	.017	**.946**	0.923	6.150
$CaCO_{3eq}(\%)$	.740	-.160	-.116	0.587	3.914
$CEC\ (cmol+kg^{-1})$	-.719	-.418	.119	0.706	4.707
$Clay(\%)$	.110	-.837	.116	.726	4.838
$Bulk\ density(gcm^{-3})$	-.039	.051	**.976**	.957	6.377
Morphological					
$Dry\ value$	**.915**	-.132	-.006	0.855	5.700
$Dry\ hue$	-.498	.215	.049	0.297	1.977
$Root\ abundance$	-.682	.296	-.095	0.562	3.748
$Texture$	-.125	**.884**	.061	0.800	5.335
$Stickiness$	.115	-.403	.151	0.198	1.322
$Moist\ value$	**.902**	-.269	.029	0.886	5.909
VAF_s (eigenvalue)#	4.993***	3.666*	1.998^{ns}	**10.658**	
$PVAF_s$ (percent)	33.3	24.4	13.3	**71.053**	

TABLE II. Categorical Principal Component Analysis (CatPCA) of field morphological soil indicators: component loadings and variance explained by components and variables. #Differences between Land Use Types: Kruskall-Wallis test. *Significant at $P < 0.05$; ***significant at $P < 0.001$; ns = not significant

In general, analytical variables show a higher explanatory power ($PVAF_j$ of 47% opposite 24% for morphological).

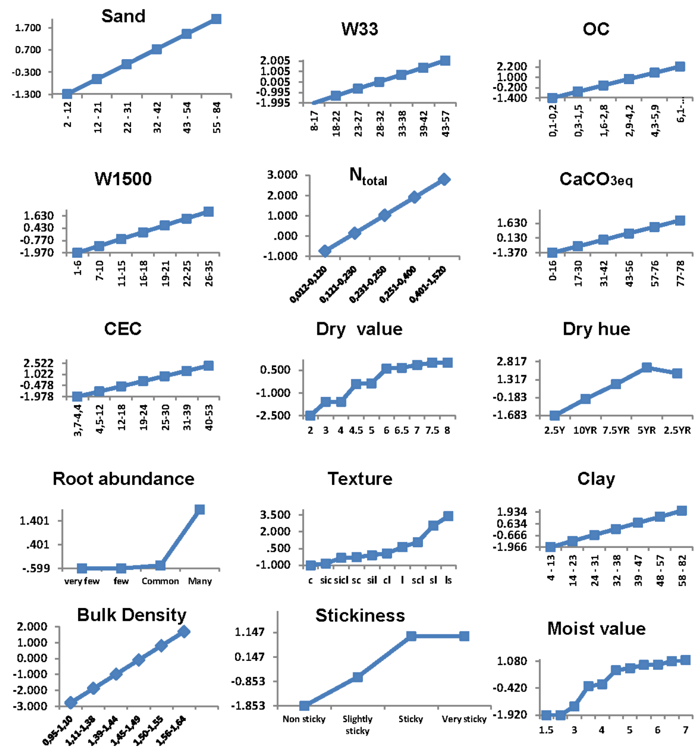

Fig. 1. Optimal scalings of soil indicators. Legend for 'Texture' attribute: c: clay; sc: sandy clay; l: loam; scl: sandy clay loam; sicl: silty clay loam; sic: silty clay; cl: clay loam; sil: silty loam; sl: sandy loam; ls: loamy sand.

Thus, variables N_{total} and $Bulk\,density$ are the ones explaining more variance (percentage over 6%). Nevertheless, morphological variables relating soil color (i.e. moist and dry values) also show a relatively high explanatory power.

Through a Kruskal-Wallis test, differences between distributions of PC1, PC2 and PC3 component scores values for each LUT were trialled. It was assumed that the most significant component in this test was the one most related to soil quality. This approximation is also followed by [23] when considering the soil quality component. According to this, PC1, with $P < 0.001$ (Table II) was selected.

The optimal scalings given by the model are shown in Figure 1. Numerical indicators show a linear behavior, while morphological ones vary between smooth nonlinear tendencies (i.e. dry value) to markedly nonlinear ones (i.e. root abundance). The ordering of nominal variables given by the model show a logical tendency regarding their physical mean. In the case of texture, categories are ordered from clay (clayey texture, c) to sand (loamy sand texture, ls). In the case of dry hue, colors are ordered from yellow (2.5Y) to red (5YR), even when intense red (2.5YR) quantification is lightly below red. This effect should be explained by human eye saturation when perceiving intense red by daylight, as supposed by [29].

According to [30], nonlinear transformations should be interpreted as differences in information content for each category, as they represent the category centroid projections over the variable vector (given by the coordinates in table II). This way, centroids related to highly overlapping clusters show very close optimal scaling values (Figure 1). Inversely, objects in highly detached categories (for the same variable), are characterized by sudden jumps between quantifications. An example is root abundance, with a quantification value of 1.886 for 'many', while the rest of the categories had the same value (-0.599). This allows the evaluation of the amount of information carried out by each category, not just the amount of total information explained by the whole indicator, given by its VAF_j. As the soil quality component PC1 accounted for most of the system variance (33%, Table II), it is possible to conclude that categories with relatively high scalings with respect to the others (in absolute values), show a greater response to soil quality, being it negative or positive terms. This capability to assess the 'weight' of each category respect to the soil quality has not yet been tried by other indexing methodologies.

For the case of analytical indicators, the category points represent intervals in the variable numerical domain. Intervals obtained in the discretization process are consistent with interpretation intervals given by the experts. E.g., sand shows a increment of 10% (Table II), agreeing with texture categories proposed by [31]. In this sense, authors like [22] use CatPCA as a discretization method for numerical domains.

x	y	Equation	r
Moist value	Organic carbon	y = -0.8856x + 0.0504	-0.754
	$CaCO_{3eq}$	y = 0.7464x - 0.0269	0.700
Root abundance	Organic carbon	y = 0.8914x - 0.081	0.601
Texture	W1500	y = -0.4691x + 0.5324	- 0.475
CEC	Dry value	y = -0.493x + 0.3539	-0.573

TABLE III. SOME SIGNIFICANT ($P < 0.001$) CORRELATIONS BETWEEN TRANSFORMED FIELD SOIL MORPHOLOGICAL INDICATORS AND ANALYTICAL PROPERTIES IN THE JAEN DATABASE

Table II also permits the interpretation of the principal components. The first component was the soil quality component. It positively correlated with the moist and dry value (component loadings of 0.902 and 0.915, respectively) and carbonates (loading of 0.740), and negatively with the organic carbon (loading of -0.868) and, in lesser extent, CEC and root abundance (component loadings of -0.719 and -0.682, respectively). Thus, when the value increases and the soil become clearer (e.g., value 8 is white color), a loss in its fertility (drop in organic matter, less roots and CEC decreasing) can be revealed. Also, the negative correlation between carbonates and CEC could be highlighted; when calcium from carbonates saturates the exchange complex, a loss in soil fertility occurs. As stated above, this component could be defined as the soil quality component where soil quality decreases with increasing component scores (fewer roots, lighter, less fertility, etc.).

A remarkable point is that by means of morphological indicators quantification, rigorous relations can be established, from the mathematical point of view, between these and analytical indicators. As an example, Table III shows some simple linear regression models between both parameter types. Interesting correlation were found, for instance, between texture and W1500 ($r = -0.475, P < 0.001$). When texture increased (towards lighter textures) soil aptitude to retain water in its micropores ($< 0.2\mu m$ diameter) decreases. Bearing in mind the nominal nature of texture, this numerical relationship would be very difficult or impossible to obtain without optimal scalings.

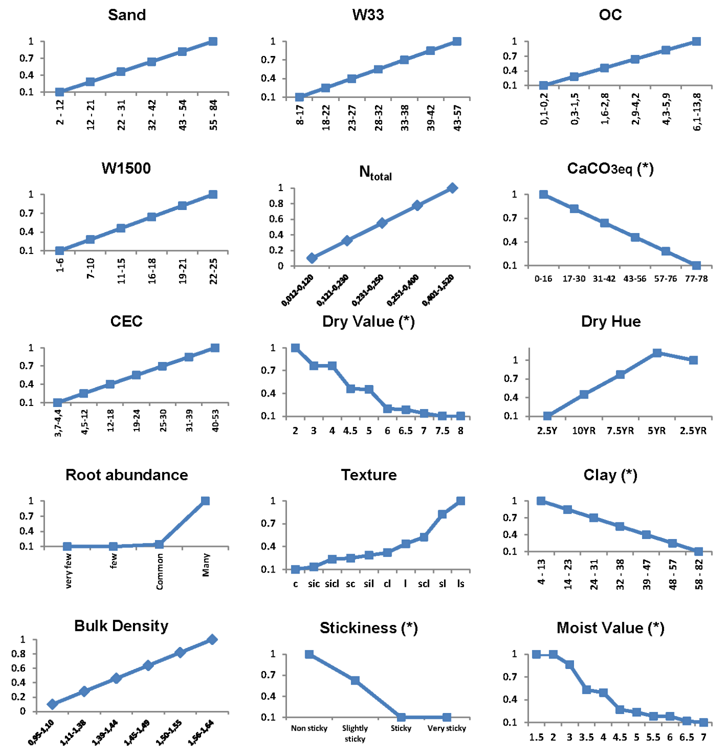

Fig. 2. Scores of soil indicators. See figure 1 for 'Texture' attribute legend. (*) marks indicators where 'less is better' (eq. 2) was applied

Next step in index building was the optimal scalings scoring. The numerical values of optimal scalings were normalized between 0.1 (poorest soil quality) and 1 (best soil quality), depending on the relative signs of the loadings of each indicator in the soil quality component (PC1). Hence, we applied less is better (eq. 2) when the indicator showed a positive loading on PC1, which was interpreted above as a loss of soil quality, changing from a monotonically growing (Fig. 1) to a monotonically decreasing function (Fig. 2), but maintaining its form. When more is better (eq. 1) is applied (negative loadings of the indicator on PC1), the functions in Fig. 2 maintain this order.

The normalization of categories quantifications permitted to normalize as well the q_{ij} (soil horizons) quantifications. Starting from these and the variables eigenvectors in the soil quality component (VAF_{js}, where $s = PC1$, Table II), the soil quality index $GSQI_i$ (eq. 3) was obtained for each horizon. The mean values for each LUT index are shown in Table IV. LUTs were ordered according to their soil quality, high to low: LDF (mean GSQI of 0.940) $> AG$ (0.899) $> PPF$ (0.732) $> MXW$ (0.673) $> MXS$ (0.609) $> PDG$ (0.528) $> HG$ (0.278) $> COG$ (0.263). This ordering correctly reflects the ecosystem disturbing degree, according to extensive studies of the area ecology, as [32]–[35].

IV. INDEX VALIDATION

We used Atanor Valley soil data as validation set, in order to find any significant difference between soil quality in two olive grove types: conventional (COG) and organic

	N	Mean	SD
LDF	5	0.940	0.094
AG	5	0.899	0.091
PPF	6	0.732	0.108
MXW	16	0.673	0.123
MXS	5	0.609	0.119
PDG	9	0.528	0.085
HC	14	0.278	0.106
COG	71	0.263	0.098

TABLE IV. GSQI VALUES (MEANS AND STANDARD DEVIATIONS SD) FOR LAND USE TYPES

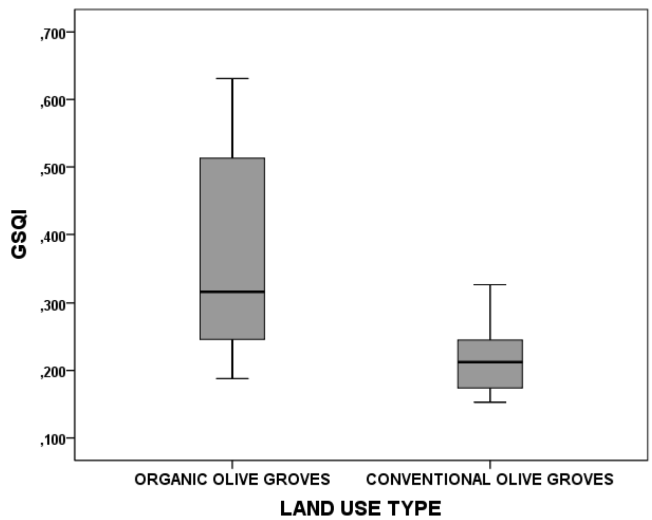

Fig. 3. Box-and-whiskers plot of Global Soil Quality Index (GSQI), applied to Organic olive groves (OOG) and Conventional olive groves (COG) in the validation set (Atanor valley)

(OOG). Mean differences were significant in a Student's-test ($P = 0.023$). The box-and-whiskers diagram in figure 3 shows the results. GSQI mean value for COG is of 0.218, a bit lower than the one found in the original data set, from which the index was calculated. On the other hand, OOG shows a GSQI higher value, 0.362, indicating that ecological management of olive grove soils gradually improves in quality, clearly stepping aside of conventional management. These results have already been reported by other authors [27], [28], [36], [37], even though these studies were exclusively based on analytical variables. Moreover, as noted before, [7] and [8], where advanced techniques were applied in the same location, led to similar results. The advantage is that GSQI measure is simpler and more accessible than these techniques.

From the methodological point of view, optimal scaling functions permit to automatically relate soil quality with indicator domain, being the latter either numerical or categorical. This is accomplished by optimizing the variance explained by the system. As a result, it is an improvement on scoring functions applications over soil quality indexing, a process highly dependant on expert knowledge [5], [38]. Thus, this automatic process simplifies the calculation of soil quality indices, as it is not required intensive participation of soil experts. Moreover, as easy to estimate, cheaper and suitable, parameters as morphological ones are included, the farmer can be more involved in the evaluation process of soil quality. Advances have been made in this sense. E.g., [39] developed score cards for untrained farmers evaluation of their olive grove soil degradation. Nevertheless, as pedo-morphological parameters were not taken into account, this type of data would be more difficult to interpret and merge with scientific knowledge than pedological data [40], as has been proven in this paper.

V. CONCLUSIONS

As stated in [41], one of the main features of a soil quality index must be its applicability in a real scenario, at farmers level, in this case. An index to integrate both analytical (quantitative) and morphological (qualitative) indicators into a global measure of soil quality has been proposed. The process is automatically driven, requiring little expert knowledge about soil, and extending its applicability to very interesting real scenarios as olive cultivation in the South of Spain. In this sense, the index can be applied by farmers themselves, and the indicators being soil survey standard parameters (i.e., soil maps), fusion with scientific knowledge is immediate. Moreover, it can be seen as an alternative in those scenarios where no other expensive and complex measures could be applied. This allows, as well, to apply the farmer (user) knowledge about his property at a higher scale, in regional or national land planning. This type of numerical and quantitative approach to soil quality could be highly useful for modeling and prediction in future physical, agronomical or socio-economic scenarios.

ACKNOWLEDGEMENTS

Work partially supported by research project PYR-2014-24, as part of the GENIL-PYR-2014 'Start-up Projects for Young Researchers' Program.

REFERENCES

[1] USDA. (2006, February) Soil quality institute. natural resources conservation service. [Online]. Available: http://soils.usda.gov/sqi/

[2] E. Gregorich, C. Drury, and J. Baldock, "Changes in soil carbon under long-term maize in monoculture and legume-based rotation," *Canadian Journal of Soil Science*, vol. 81, pp. 21–23, 2001.

[3] D. L. Karlen, M. J. Mausbach, J. W. Doran, R. G. Cline, R. F. Harris, and G. E. Schuman, "Soil quality: A concept, definition and framework for evaluation," *Soil Science Society of America Journal*, vol. 61, p. 410, 1997.

[4] M. Yemefack, V. Jetten, and D. Rossiter, "Developing a minimum data set for characterizing soil dynamics in shifting cultivation systems," *Soil and Tillage Research*, vol. 86, pp. 84–98, 2006.

[5] S. S. Andrews, D. Karlen, and C. Cambardella, "The soil management assessment framework: A quantitative soil quality evaluation method," *Soil Sci. Soc. Am. J.*, vol. 68, p. 19451962, 2004.

[6] J. Calero, R. Delgado, G. Delgado, and J. Martín-García, "Transformation of categorical field soil morphological properties into numerical properties for the study of chronosequences," *Geoderma*, vol. 145, p. 278287, 2008.

[7] V. Aranda, M. Ayora-Cañada, A. Domínguez-Vidal, J. Martín-García, J. Calero, R. Delgado, T. Verdejo, and F. González-Vila, "Effect of soil type and management (organic vs. conventional) on soil organic matter quality in olive groves in a semi-arid environment in sierra mgina natural park (s spain)," *Geoderma*, vol. 164, pp. 54–63, 2011.

[8] J. Calero, M. Cordovilla, V. Aranda, R. Borjas, and C. Aparicio, "Effect of organic agriculture and soil forming factors on soil quality and physiology of olive trees," *Agroecology and Sustainable Food Systems*, vol. 37, pp. 193–214, 2013.

[9] D. Aguilar, J. Fernández, J. Sánchez, T. de Haro, Rodríguez, and E. Fernández, *Memoria y Mapa de suelos escala 1:100,000, Iznalloz (991). Proyecto LUCDEME*. Ministerio de Agricultura, Alimentación y Pesca- ICONA, Madrid, Spain, 1993.

[10] D. Aguilar, J. Fernández, S. de Haro, A. Martínez, T. Rodríguez, and E. Fernández, *Memoria y Mapa de suelos escala 1:100,000, Alcaudete (968). Proyecto LUCDEME*. Ministerio de Agricultura, Alimentación y Pesca- ICONA, Madrid, Spain, 1995.

[11] D. Aguilar, M. Simón, A. Marañés, I. García, C. Asensio, and A. Iriarte, *Memoria y Mapa de suelos escala 1:100,000, Úbeda (906). Proyecto LUCDEME*. Ministerio de Agricultura, Alimentación y Pesca- ICONA, Madrid, Spain, 1997.

[12] R. Delgado, E. Ortega, B. de Quiros, C. Sierra, G. Delgado, C. Aguilar, and M. Simón, *Memoria y Mapa de suelos escala 1:100,000, Andújar (904). Proyecto LUCDEME*. Ministerio de Agricultura, Alimentación y Pesca- ICONA, Madrid, Spain, 1997.

[13] G. Delgado, M. Sánchez-Marañón, J. Párraga, R. Delgado, E. Gámiz, J. Martín-García, M. Soriano, and R. Temsamani, *Memoria y Mapa de suelos escala 1:100,000, Mengíbar (926). Proyecto LUCDEME*. Ministerio de Agricultura, Alimentación y Pesca- ICONA, Madrid, Spain, 1997.

[14] G. Delgado, M. Sánchez-Marañón, J. Párraga, R. Delgado, E. Gámiz, J. Martín-García, P. García-Corral, M. Soriano, and R. Temsamani, *Memoria y Mapa de suelos escala 1:100,000, Baeza (927). Proyecto LUCDEME*. Ministerio de Agricultura, Alimentación y Pesca- ICONA, Madrid, Spain, 1997.

[15] C. Sierra, F. Martínez, A. Roca, I. Saura, M. Sierra-Aragón, A. Cirre, F. Moral, and A. Pérez, *Proyecto LUCDEME, Memoria y Mapa de suelos escala 1:100,000, Alcalá la Real (990)*. Ministerio de Agricultura, Alimentación y Pesca- ICONA y Universidad de Granada. Spain, 2003.

[16] S. de Haro, "Génesis y cartografía de suelos en la comarca olivarera de martos. características agronómicas (in spanish)," Master's thesis, Univ. of Granada (Spain), 1992.

[17] V. Aranda, "Caracterización y análisis de la fracción orgánica de los horizontes superficiales en suelos de ecosistemas mediterráneos (in spanish)," Master's thesis, Univ. of Granada, Spain, 1998.

[18] J. Martín-García, R. Jiménez, G. Delgado, R. Delgado, V. Aranda, J. Calero, M. Sánchez-Marañón, G. Liébanas, F. Ariza, A. Alcalá, C. Pinilla, and C. Enríquez, *Mapa Geomorfoedáfico del Parque Natural de Sierra Mágina (Jaén-España). Memoria y Mapa. Informe científico-técnico de la Consejería de Medio Ambiente*. Junta de Andalucía, Sevilla, 2000.

[19] J. Calero, "Génesis de la fracción mineral y de la ultramicrofábrica en una cronosecuencia de suelos sobre terrazas del río guadalquivir (in spanish)," Master's thesis, Univ. of Granada, Spain, 2005.

[20] IUSS WORKING GROUP WRB, *World reference base for soil resources 2006 - a framework for international classification, correlation and communication*, 1st ed. Roma: Food and Agriculture Organization of the United Nations, 2006, world Soil Resources Reports No. 103.

[21] J. Meulman and W. Heiser, *SPSS Categories 10.0*. SPSS Inc, Chicago, 1999.

[22] J. Meulman, A. van der Kooij, and W. Heisser, *Handbook of Quantitative Methodology for the Social Sciences*. Sage Publications, Inc., Thousand Oaks, CA, 2004, ch. Principal components analysis with nonlinear optimal scaling transformations for ordinal and nominal data, p. 4970.

[23] M. Shukla, R. Lal, and M. Ebinger, "Determining soil quality indicators by factor analysis," *Soil & Tillage Research*, vol. 87, p. 194204, 2004.

[24] C. Seybold, M. Mausbach, D. Karlen, and H. Rogers, *Soil Processes and the Carbon Cycle. Advances in Soil Science 2*. Boca Ratón, Florida, USA, 1998, ch. Quantification of Soil Quality.

[25] K. Sharma, U. Mandal, K. Srinivas, K. Vittal, B. Mandal, J. Kusuma-Grace, and V. Ramesh, "Long-term soil management effects on crop yields and soil quality in a dryland alfisol," *Soil & Tillage Research*, vol. 83, p. 246259, 2005.

[26] E. Velasquez, P. Lavalle, and M. Andrade, "Gisq, a multifunctional indicator of soil quality," *Soil Biology and Biochemistry*, vol. 39, pp. 3066–3080, 2007.

[27] J. Gómez, S. Álvarez, and M. Soriano, "Development of a soil degradation assessment tool for organic olive groves in southern spain," *Catena*, vol. 79, pp. 9–17, 2009.

[28] S. Álvarez, M. A. Soriano, B. B. Landa, and J. A. Gómez, "Soil properties in organic olive groves compared with that in natural areas in a mountainous landscape in southern spain," *Soil Use and Management*, vol. 23, pp. 404 – 416, 2007.

[29] M. Sánchez-Marañón, P. García, R. Huertas, J. Hernández-Andrés, and M. Melgosa, "Influence of natural daylight on soil color description: Assessment using a color-appearance model," *Soil Sci. Soc. Am. J.*, vol. 75, p. 984993, 2011.

[30] M. Linting and A. Van der Kooij, "Nonlinear principal components analysis with catpca: A tutorial," *Journal of Personality Assessment*, vol. 94, no. 1, p. 1225, 2012.

[31] P. Schoeneberger, D. A. Wysocki, E. C. Benham, and W. D. Broderson, *Field Book for Describing and Sampling Soils*. USDA-NRCS National Soil Survey Center, Lincoln, Nebraska, 1998.

[32] F. González-Vila, G. Almendros, and F. Martín, "An evaluation of the differences in the composition of humic acids in soils under oak and pine forest by gc-ms after mild degradation," *Plant Soil*, vol. 103, pp. 83–88, 1987.

[33] V. Aranda and C. Oyonarte, "Effect of vegetation with different evolution degree on soil organic matter in a semi-arid environment (cabo de gata-níjar natural park, se spain)," *Journal of Arid Environments*, vol. 62, pp. 631–647, 2005.

[34] ——, "Characteristics of organic matter in soil surface horizons derived from calcareous and metamorphic rocks and different vegetation types from the mediterranean high-mountains in se spain," *European Journal of Soil Biology*, vol. 42, pp. 247–258, 2006.

[35] G. Montserrat-Martí, J. Camarero, S. Palacio, C. Pérez-Rontomé, R. Milla, J. Albuixech, and M. Maestro, "Summer-drought constrains the phenology and growth of two coexisting mediterranean oaks with contrasting leaf habit: implications for their persistence and reproduction," *Trees*, vol. 23, p. 787799, 2009.

[36] R. García-Ruiz, M. Ochoa, M. Hinojosa, and B. Gómez-Muñoz, "Improved soil quality after 16 years of olive mill pomace application in olive oil groves," *Agron. Sustain. Dev*, vol. 32, no. 3, pp. 803–810, 2012.

[37] R. Marzaioli, R. DAscoli, R. De Pascale, and F. Rutigliano, "Soil quality in a mediterranean area of southern italy as related to different land use types," *Applied Soil Ecology*, vol. 44, p. 205212, 2010.

[38] S. Andrews, D. Karlen, and M. J.P., "A comparison of soil quality indexing methods for vegetable production systems in northern california," *Agriculture, Ecosystems and Environment*, vol. 90, p. 2545, 2002.

[39] J. Calero, J. Serrano, V. Aranda, D. Sánchez, M. Vila, and G. Delgado, "Analysis and characterization of olive tree cultivation system in granada province (south of spain) with optimal scaling and multivariate techniques," *Agrochimica*, vol. 49, pp. 118–127, 2005.

[40] G. Delgado, V. Aranda, J. Calero, M. Sánchez-Marañón, J. Serrano, D. Sánchez, and M. Vila, "Using fuzzy data mining to evaluate survey data from olive grove cultivation," *Computers and Electronics in Agriculture*, vol. 65, pp. 99–113, 2009.

[41] F. Bastida, A. Zsolnay, T. Hernández, and C. García, "Past, present and future of soil quality indices: A biological perspective," *Geoderma*, vol. 147, p. 159171, 2008.