Interpretation and Integration of Pedological Data in Land Evaluation Systems

K. Juhos* and B. Madarász

1 Szent István University, Faculty of Horticultural Science, Department of Soil Science and Water Management, H-1118 Budapest, Hungary
2 Hungarian Academy of Sciences, Geographical Institute, Research Centre for Astronomy and Earth Sciences, H-1112 Budapest, Hungary

Abstract


Highly standardized land evaluation systems and applicable methods are necessary for an ecologically and economically sustainable land use, as the most important requirement. In general, available pedological data and their interpretation define the applicability of different land evaluation systems. Our study focuses on (1) the summary of already applied pedological land information; (2) the main methods of data interpretation and integration in the land evaluation systems; (3) the up to date quantitative land evaluation procedures, with the involvement of mathematical-statistical tools in the studies. Soil classification systems and other complex pedological qualifiers are the basis of qualitative and/or the semi-quantitative land evaluation systems. Selection, interpretation and integration of the simple soil properties and indicators are not easy to use, because those simple parameters should be evaluated in a highly multifactorial environmental context. Development of quantitatively flexible indices and mathematical and statistical methods, that are able to explore pedological and environmental factors in land productivity, beside the most accepted economic ones, seems to be highly necessary nowadays.

Key words: land capability, land suitability, soil quality index, multiple criteria decision making

Abbreviations: AHP – analytical hierarchy process; IQI – Integrated Quality Index; MCDM – multiple criteria decision making; MDS – minimum data set; NQI – Nemoro Quality Index; PCA – principal component analysis; PLS – partial squares regression; TDS – total data set; USDA-LCC – US Department of Agriculture Land Capability Classification

Introduction

Understanding the variability of soil and landscape properties and their effect on crop yield is a critical component of site-specific management systems and land use planning. Land evaluation means generally the determination of land potentials for certain purposes (Driessen and Konijn, 1992). Land potential involves the: 1) limitations and/or the capabilities of the land at/for the various environmental factors (land capability), 2) land use types, the variety of cultures to grow and the level of production (land suitability). Factors affecting the agro-ecological potential are in close relationship, however, by their relative weight in the production, which is variable both spatially and temporally (Rossiter, 1995; Herrick, 2002). The interpretation of land characteristics is taking into account the agricultural use, in which it is necessary to determine the distances among observed characteristics from the optimal, non-limiting values (Sys et al., 1991).

A data base representing the pedological properties is the most important environmental component for developing of a standardized land evaluation procedure (Davidson, 1992; Mueller et al., 2010). Soil quality has been defined by Karlen

*E-mail: Juhos.Katalin@kertk.szie.hu
et al. (1997) as “the capacity of soil to function to sustain plant and animal productivities, to maintain or enhance water and air quality and to support human health and habitation”. Biomass and grain yields depend on complex interactions among physical and chemical properties of soil (Zobel et al., 1988).

On the basis of data interpretation and integration De la Rosa and Van Diepen (2002) differentiates between: (1) qualitative, (2) semi-quantitative and (3) quantitative (modern mathematical and statistical) land and soil evaluation methods. This approach also considers a kind of chronologically development, but the differentiation cannot often be definitive. Applicability of different land evaluation systems in land-use management largely depends on the available pedological data and on the way of their interpretation.

For the last few decades there was a rapid development of informatics; the former land evaluation systems have to be considered therefore in another light. The new requirement about the available information and the possibilities of data-processing make it necessary to review them again. Through examples, mainly involving the temperate climate zone countries the present study focuses on:

- the types of pedological and other environmental information in use;
- the way of data interpretation and its integration in the land evaluation systems;
- the mathematical and statistical approach of the modern quantitative land evaluation.

An important question is whether mathematical and statistical methods or expert opinion or both combined are more useful to express the soil quality.

Materials and Methods

Qualitative methods – land capability and land suitability classification

Traditional qualitative methods are using various categories for the expression of land productivity. The first step is the delineation of different complex (e.g. soil type, texture type) and simple soil qualities (e.g. pH, electrical conductivity) using (arbitrary) threshold values. This is followed by a stepwise collating, which reveals (1) the amount and nature of limitations (land capability classification), (2) suitability classes for different land tillage or plant cultivation (land and soil suitability classification).

The earliest land evaluations were applicable for certain plant species, on the bases of some general field observations (direct suitability classification). For example, in Germany soils, excellent for the wheat-, rye- and oat breeding were demonstrated as early as the 19th century (Rothkegel, 1950). In Hungary the soil rating of Géczy (1968) was created for practical purposes in 1968. It has expressed the soil quality by denoting the crop-plants, which were successfully produced on it (i.e. the cereals, the annual forage or the industrial crops, and also the papilionaceous forage plants).

The FAO system (FAO, 1976) was the basis of several national qualitative evaluations, started in 1976, and this classification is still being in use at several countries. These qualitative land evaluations are usually hierarchic.

The three level hierarchic system of Land Capability Classification (USDA-LCC) of the USA is being a classic example. It evaluates the lands in three main land use categories: (1) arable fields, (2) pastures, (3) forests (Klingebiel and Montgomery, 1961). Agricultural interpretation of soil mapping units (created by the relevant knowledge on the Soil Taxonomy) is the basis of the USDA-LCC system. The so called “units” are representing of the lowest level of the hierarchy, where management and the cultivation intensities are identical; the same kinds of plants are cultivated with similar tillage types and, soil protection ways and as the same expected yields. The subclasses are defining of the quality and most of the limiting factors (like erosion, excessive moisture, soil deficiencies in the rooting zone and the climatic hazards... etc.). Those limiting factors are definitive during the classifications; therefore this method is called as maximum-limitation system (De la Rosa and Van Diepen, 2002) or simple limitation method (Sys et al., 1991). The risk of soil damage is progressively greater from class I towards the class VIII, which can determine therefore the better land suitability.

This method was accepted and adopted without any relevant changes in several countries. The Canada Land Inventory (Agricultural and Rural Development Act, 1965), the Land Use Capability Classification of New Zealand (Lynn et al., 2009) and the Land Capability Classification for Agriculture created by the Research Institute of Macaulay Land Use in Great Britain (Bibby et al., 1991) are also based on the USDA-LCC system, with some minor modifications only.

The international Fertility Capability Classification (Sanchez et al., 1982) is the result of the qualitative agricultural interpretation of the Soil Taxonomy. This method puts significantly greater emphasize on the properties of the topsoil and of the tilled layers.

Qualitative land evaluations provide basic information on the most important production limiting factors. However, as these methods are based on some category data, its reliability depends largely on the solidity and completeness of the environmental information. The category system is best applicable in cases when only a few factors are significant in land use and all other factors play only a secondary role.
Traditional semi-qualitative soil rating systems

The traditional soil rating methods are using those various categories, which were created: 1) on the basis of some measured or established parameters, or 2) only are using some of the categories of the previous soil mapping. The combinations of those categories are placed into a numerical reference scale, in which the potential yields as fundamental in formations can provide a relatively well-applicable reference scale.

The German soil rating (Reichbodenschätzung) is one of the earliest examples for the traditional soil rating in Europe. It was registered in 1934 with the aim of developing a rather “uniform” taxation, potentially applicable for several purposes (Heide and Mückenhausen, 1980). Soils are evaluated in this method by considering the climatic effects and also the relief. In the case of arable lands, soil texture, parent rock (geological origin), degree of soil development, the climate and relief were studied. After that these data were categorized. A land evaluation table was compiled on the basis of the combinations of category data for the different land use types. For the determination of the values appearing in the tables representative soils were selected as the references. In “arable land” category for example, the optimal production site (the Chernozem soil, with 100 points) was located close to Magdeburg. Reliability of the method is supported by the estimation of soil fertility and also the expected yields. This German soil rating was the basis of the used Austrian soil ratings (Pehamberger, 1992; Wagner, 2001) and the soil rating system of the former Soviet Union (Gavrilyuk, 1974).

The reliability of the various evaluation systems are depending mostly on the different kind of experiences, which has been the basis of the used reference systems (scale). During the interpretation of pedological data it is important to consider the pedological requirements of certain land use types or that of certain plants, the expected yields and the effect of other most relevant key-important environmental factors. The Hungarian land-evaluation was developed in 1970, by creating specific “soil-rating numbers” between the values of 1 and 100. In this system the yield elements was involved, as calculated data of the main national genetic soil-map categories (Stefanovits et al., 1972). The potential yield could not be truly predicted, especially in the context of the great variability of the various management practices (of using different amount of organic and inorganic fertilizers for instance). Solving this problem was the objective of the new D-e-Meter land evaluation system in Hungary (Tóth, 2011).

Semi-quantitative parametric land evaluations – soil quality indexes

Traditional parametric land evaluation assesses directly the simple soil qualities (single factor valuation). This evaluation is followed by the integration of those interpreted data into an index via additive, multiplicative or combined mathematical methods (Driessen and Konijn, 1992). The interpretation of soil characteristics is taking into account the agricultural use, in which it is necessary to determine of the distances among observed characteristics from the optimal, non-limiting values (Sys et al., 1991).

Requirements of certain land use types are taken into account during the ranking of the resulting indexes. The Storie Index developed in California in the 1930’s (Storie, 1978) is considered as the first mathematical soil quality assessment method. The Storie Index had a role in the taxation system and has been modified several times since its creation. Four groups of factors are apprised by the evaluation system based on certain indicators. Optimal value of these factors are 100 %, thus unfavourable conditions are expressed in percentages, which are used to calculate multiplicative non-weighted index:

\[ \text{Storie Index} = A \times B \times C \times X, \]

where: \(A\): character of physical profile (indicators: depth and quality of parent rock, present sedimentation), \(B\): surface texture, \(C\): shape and angle of the slope, \(X\): other factors influencing the fertility (i.e. the product of drainage, alkalinity, acidity, nutrient level, erosion and micro relief).

This index was adopted and reworked in several countries, e.g. in Brazil (Bacic et al., 2003) and in Poland (Koreleski, 1988).

Soil productivity index of Neill (1979) is also multiplicative. This index integrates the change of soil properties with depth. Factors (sufficiency of available water capacity, aeration, bulk density, pH, electrical conductivity) are weighted at each layer in consideration of the rooting depth of the produced plant.

It is also possible to insert parameters in land evaluation in an additive way, by simple or weighted sum or by subtraction from a maximum value. Limiting effect of the soil properties is expressed by point subtraction in the Canadian Land Suitability Rating System for Agricultural Crops (Agronomic Interpretations Working Group, 1995). The most productive soils, where limiting factors are small or negligible have 100 points. This system evaluates the production-limiting effect of pedological factors in consideration of climatic and hydrologic conditions. In Europe the French method is a similar, point subtracting system (Mori et al., 1984), which gives a weight to each factor.
A combination of additive and multiplicative methods makes the evaluation even more complex. A good example is the Muencheberg Soil Quality Rating, which evaluates soil productivity on cropping and grazing (Mueller et al., 2007). Basic soil indicators (e.g. substrate, A horizon depth, rooting depth, available water retention capacity, slope) are scored by using scoring tables. Single scores are on a scale ranging from best conditions to worst (2 to 0 values) and are summarized by weighting (1 to 3 values). Weight of the known values is assessed. The so called hazard indicators relevant in the study area (e.g. contamination, salinization, acidification, low total nutrient status) are evaluated on a 0 to 2.94 scale. Then the basic soil value index is multiplied by these hazard values. The resulting index expresses soil productivity on a 100 point scale.

Major problem with the additive systems is that production-limiting effect of certain extremely unfavourable parameters is not expressed and/or emphasized enough. Additive methods are improved significantly by weighting. On the other hand, weighting is a source of errors, especially when it is done arbitrarily. Advantage of the multiplicative methods is that the effect of each factor is well expressed in the resulting index. Factors may be weighted in accordance with the objectives of different land use.

These methods may be considered as semi-quantitative also, if the interpretation of those parameters is based on category tables and it is partly subjective. The proper interpretation of those “simple” parameters needs specific and highly qualified expertise, due to the fact, that soil qualities can only be evaluated in context of the other qualities and the affecting natural factors. Beside this, their importance in the produced yields can be also different.

These semi-quantitative soil quality indices can be well-applicable if (1) they are simple and consist of the least number of parameters; (2) the parameters show significant differences in the productivity; (3) the parameters are independent. The benefit of such parametric method is that the showing significant variability in the study area and is definitive in productivity.

**Mathematical-statistical tools in modern land evaluation and soil quality assessment systems**

Soil quality indices still are the most commonly used today, because they are quantitatively flexible (Andrews et al., 2002; Qi et al., 2009; Lima et al., 2013; D’Hose et al., 2014; Rahmanipour et al., 2014). If the mathematical-statistical tools are used for the selection, weighting and interpretation of factors and indicators and perhaps for their interpretation, the quantitative parametric evaluation methods are being developed. The development of a soil quality index should follow three steps: (1) selection of indicators, (2) scoring the selected indicators and (3) the integration of indicators in an index (Karlen et al., 1997).

Among the indicators selection methods, Total Data Set (TDS) and Minimum Data Set (MDS) have been widely used (Rahmanipour et al., 2014). The TDS is a collection of indicators selected according to specific characteristics of soil analysed. The MDS is a selected collection of indicators chosen according to expert opinion and soil functions (Andrews et al., 2002; Lima et al., 2013) or using several multivariate statistical methods: principal component analysis and correlation among indicators (Andrews et al., 2002; Qi et al., 2009; D’Hose et al., 2014; Rahmanipour et al., 2014), these combination with discriminant analysis among different land use/management categories (Xu et al., 2006; Yao et al., 2013) and multiple regression analysis with dependent variables crop yield (Andrews et al., 2002; Rezaei et al., 2006).

The second step in development a soil quality index is normalizing the MDS indicators by different numerical scales using linear (Liebig et al., 2001) and nonlinear scoring (Andrews et al., 2002; Qi et al., 2009; Lima et al., 2013; D’Hose et al., 2014; Rahmanipour et al., 2014) functions. Mathematical basis of this is the Fuzzy set theory, which works with possibilities and membership functions: the membership is expressed with numerical values between 0 and 1 (where 0 means that the data is outside of the cluster and 1 means the full inclusion in a cluster) (Pedrycz and Gomide, 1998). This method is suitable for the indicators show the similarity between the attribute data (soil properties, local conditions) and the demands of the applying groups (e.g. farmers).

The integration of dimensionless normalized indicators into quality indices is possible through many procedures based on additive, multiplicative or weighed mean techniqes. Integrated Quality Index (IQI) (Karlen et al., 2007; Andrews et al., 2002; Lima et al., 2013; D’Hose et al., 2014) and Nemoro Quality Index (NQI) (Qi et al., 2009; Rahmanipour et al., 2014) are examples of this calculation.

Integrated Quality Index (IQI) were calculated in the following way (Doran and Parkin, 1996):

\[ IQI = \sum_{i=1}^{n} W_i N_i, \]

where \( W_i \) is the weight of each indicator and \( N_i \) is the indicator score.

Several authors weighted the normalized indicators according to expert opinion (Andrews et al., 2002; Lima et al., 2013), principal component analysis (Andrews et al., 2002; D’Hose et al., 2014; Rahmanipour et al., 2014) or using analytical hierarchy process (AHP) (Li et al., 2002; Ananda and
AHP is a useful mathematical method to express how the factors of the lowest hierarchy level (soil indicators) effect on the highest factor (e.g. crop yield), i.e. what is their intensity or priority. First step in this process is the grouping of decision-making factors by their natural qualities (Saaty, 1980). This is followed by the creation of a new hierarchy level by the re-classification of the groups until the uppermost element; the objective of the decision is reached. Strength of the connection between nodes of the hierarchic structure (graph) is expressed by a number in the [0;1] interval. A positive entire number is assigned to one of the elements and the reciprocal of this number is assigned to the number directly above the previous one. The so called Saaty-matrix can be scored up in each level with the above described importance values. Weight of the hierarchy connections is defined by the elements of the eigen-vector belonging to the highest eigen-value of the matrix.

The Nemoro Quality Index (NQI) is not multiplicative but it takes into account the average of the scores and the least suitable pedological parameter (Qi et al., 2009):

$$NQI = \sqrt{\frac{P_{ave}^2 + P_{min}^2}{2}},$$

where $P_{ave}$ is the average and $P_{min}$ is the minimum of the scores of the selected indicators.

These indices have been used to evaluate impacts of agricultural practices. But land use decision support systems (De la Rosa et al., 1992; Kelgenbaeva, 2002; Shabbazi et al., 2008; De la Rosa et al., 2009; Lamelas et al., 2009; Lugo-Morin and Rey, 2008), the so-called Multiple Criteria Decision Making (MCDM) (Kurtener and Badenko, 2000; Malczewski, 2002; Baja et al., 2007; Busscher et al., 2007; Zhang et al., 2004) and dynamic biomass simulation models (Krupkin and Toptygin, 1999; Kersebaum, 2007; Xiong et al., 2008; Reidsma et al., 2009) demand usually quantitatively flexible and complex pedological data and mathematically interpreted input indicators.

Several authors found that soil quality has a significant relationship with crop yield using simple linear correlation/regression or multivariate linear regression analysis (Brubaker et al., 1994; Andrews and Carrol, 2001; Andrews et al., 2002; Rezaei et al., 2006; De Araujo et al., 2009). Simple linear functions, however, are usually inappropriate to describe the effect of soil indicators on the productivity. Furthermore, intercorrelation among soil properties can result in multicollinearity problems regarding relationships between soil properties and crop yield. To solve multicollinearity problems between independent variables Ping et al. (2004) carried partial least squares regression (PLS). Loadings from linear combinations of variables in PLS allowed identifying soil properties that have the greatest influence on yields. Another option is combining variables using principal component analysis (PCA) and then having the derived factors conducted a linear multivariate regression processes (Mallinarino et al., 1999; Cox et al., 2003; Shukla et al., 2004; Masto et al., 2008; Juhos et al., 2016). However these PCA factors do not explain the total variance of the entire set of variables and they are not easy interpretable.

The relationship between crop yield and soil quality is very complex and it depends on complex interactions among physical and chemical properties of soil and other external natural factors. Therefore a multivariate approach is necessary.

Conclusions

Currently, there are no available simple, standard methods for the land evaluations in the World, which is being reflected from the study. The great variability in land use types might require diverse land evaluation methods; exact methodologies of which the local differences in land use categories and also in the ecological conditions might be considered.

Today, there is a trend of explore the relationship among main pedological and environmental factors in land productivity. This can be helpful also in creating the complex decision support systems, in which mainly soil interpretation methods might be involved on the bases of some mathematical-statistical analysis of measurable pedological parameters. Tools of mathematical-statistical methods can be helpful in interpretations and also in the integrations of pedological data, becoming to be more exact and less subjective.

The modern possibilities of information technologies and also the traditional methods are complementary to each other. Mathematics and statistics might be highly complementary tools for the transformation of a certain land evaluation method to the specific area or culture. However, no any quantitative method is suitable without the appropriate expertise of the stakeholders, establishing of the right values of soils.

References


Interpretation and Integration of Pedological Data in Land Evaluation Systems


Storie, R. E., 1978. Storie index for soil rating revised. Special Publications of Division of Agriculture and Natural Resources 3203, University of California, Davis.


Received March, 12, 2015; accepted for printing February, 19, 2016