UDC 51-76 **doi: 10.20998/2522-9052.2020.1.02**

S. A. Hasanova[1], A. B. Pashayev[1], E. N. Sabziev[2]

[1] Institute of Control Systems of ANAS, Baku, Azerbaijan
[2] Kiber Ltd Company, Baku, Azerbaijan

FUZZY DIAGNOSTICS OF SOILS ACCORDING TO THE WORLD REFERENCE BASE FOR SOIL RESOURCES

Abstract. Soil classification remains one of the most controversial topics in the world soil science because of differences in the principles underlying it. As of today, many countries have developed and use their own national classifications. Since scientific classification facilitates successful investigation and proper use of soils, the issue arises of representation of one classification in another. In the paper the possibility of applying fuzzy technologies to soil diagnostics based on the World Reference Base for Soil Resources is considered. A trapezoidal form of membership functions of the parameter set of diagnostic horizons is proposed. The algorithm of fuzzy diagnostics is presented on the basis of the proposed functions. **Conclusion.** The paper proposes a fuzzi fication mechanism for the set of indicators of diagnostic parameters and describes the algorithm of the system of fuzzy inference about the identification of the values of the corresponding parameters in accordance with the compositional inference rule. Thus, the effectiveness of the use of fuzzy technologies in soil diagnostics by the World Reference Base for Soil Resources is demonstrated.

Keywords: information model; soil; fuzzy technologies; soil classification; diagnostic horizon; diagnostic parameters.

Introduction

Such classifications include, for instance, the soil classification of Australia (Australian Soil Classification, 2002) [1], Azerbaijan (M.P. Babayev's classification, 2006), Belarus (Belarusian soil classification, 1997) [2], Brazil (Sistema Brasileiro de Classificacaode Solos, 1999) [3, p.1-3], Germany (Harmonization of Soil Survey Classification-Blending East with West, 1990), Canada (Canadian System of Soil Classification, 1995) [4], Russia (V.V. Dokuchaev's classification, 1977) [5], USA (Keys to Soil Taxonomy, 2003) [6], France (Referentael Pedologique, 2008) [7, p.166] etc. The foundation of the work on the creation of an international soil classification was laid back in the 1970s by Professor R. Dudal [8]. Afterwards, a consolidated program was developed, which was later named "World Reference Base for Soil Resources" (WRB) [9]. The main objective of WRB was the introduction of the latest achievements in soil science related to the study of global soil resources and their interrelationships.

Since scientific classification facilitates successful investigation and proper use of soils, the issue arises of representation of one classification in another. The process of determining whether a soil belongs to a particular class of soils in the classification in question is called soil diagnostics. Each of these classifications has its own diagnostic algorithm. There is a certain number of soil parameters, by which a soil is placed in one class or another. However, the measurement range of these parameters is often approximate, fuzzy. In view of the above, this paper is devoted to the study of the possibility of applying fuzzy technologies to soil diagnostics based on the World Reference Base for Soil Resources.

On the priority of the defining classification factor

It should be noted that, depending on the structural principle, different classifications use different parameters and use them in different sequences. In soil classification, different soil science schools prioritize different criteria based on the objective at hand. Here is an example showing the importance of the sequence of classification criteria applied.

We shall consider three notional tomato cultivars. Let each cultivar be characterized by three parameters: yield, disease resistance and water requirement. Assume that the first cultivar has a high yield, but is not resistant to diseases and requires abundant watering (Table 1). The second cultivar is resistant to diseases and has a high water requirement but an average yield. And the third cultivar is not disease-resistant, has an average yield but a low water requirement.

Table 1 – **Values of the indicators for different tomato cultivars in conventional units**

Indicator	Cultivar		
	I	II	III
Yield	1.0	0.8	0.6
Disease resistance	0.8	0.1	1.0
Water requirement	0.7	0.5	1.0

Assume that, for the purpose of cultivation, three farmers need to classify these tomato cultivars into three categories, "Preferred", "Allowed", "Undesirable". It is known that the first farmer grows tomatoes in a greenhouse, where all the conditions required for the optimal development of plants can be created; the second farmer works in an arid zone with limited irrigation water, and the third one – in the ecological area, where the use of chemicals is not allowed.

It is obvious that the first farmer prioritizes the criterion of yield, the second farmer proceeds from cultivation in the conditions of limited water resources, and the third farmer prioritizes the resistance of the cultivar to diseases. Consequently, the preferences of these farmers will be as specified in Table 2. This example demonstrates that changing the sequence of the applied classification criteria can lead to completely different results.

Table 2 – **Farmers' preferences for cultivars**

Class	Farmer		
	1	2	3
Preferred	I	II	III
Allowed	II	I	I
Undesirable	III	III	II

Fuzzy nature of diagnostic parameters

The defining factors in the soil diagnostics are genetic horizons. There are 37 of them in the WRB system. Each of these horizons is estimated on the basis of the values of specific sets of parameters. Table 3 gives examples of genetic horizons [9].

Table 3 – **Some genetic horizon according to the WRB classification**

Genetic horizon	Parameters
Protovertic horizon	• Clay; • Wedge-shaped soil aggregates; • Shrink-swell cracks; • Coefficient of linear extensibility; • Thickness.
Salic horizon	• pH water; • Soluble salts; • Thickness; • Consist of organic or mineral material.
Petrogypsic horizon	• Secondary gypsum (CaSO4·2H2O); • Gypsum (by mass); • Thickness.
Fragic horizon	• Consists of mineral material; • Soil organic carbon; • Penetration resistance at field capacity; • Effervescence.

At present, the WRB classification includes 32 reference soil groups. However, when assigning a soil to a reference soil group, not all parameters are used.

Depending on the reference soil group, the parameters under consideration also change. Let us illustrate this with the following example. Take two reference soil groups, Fluvisols and Histosols. When assigning a soil to Histosols, the thickness of the horizon of organic materials is considered first, and in Fluvisols, it is the presence of alluvial material in the soil profile that is primarily taken into account.

The total number of parameters that characterize a soil is 28, and most of them have a fuzzy range of measurement.

We give some of these parameters as an example below in Table 4.

It should also be noted that the values of these parameters are measured with a certain error; therefore, their mean values are used in soil diagnostics.

Thus, sets of indicators can be determined as corresponding elements of fuzzy sets [10].

In the following paragraphs, we shall perform the fuzzification of a set of indicators of diagnostic parameters, applying the fuzzy technology to the description of the data, after which we shall use fuzzy inference rules to carry out clustering of soils according to the WRB system.

Table 4 – **Some diagnostic parameters of soils**

Parameter, unit of measurement	Parameter value	Lower limit value	Upper limit value
Granulometric composition, mm	Clay	0	0.002
	Ooze	0.002	0.05
	Very fine grained sand	0.05	0.10
	Fine grained sand	0.10	0.25
	Medium sand	0.25	0.50
	Coarse sand	0.50	1
	Very coarse sand	1	2
Boundary between horizon, sm	Sharp	0	2
	Clear	2	5
	Gradual	5	15
	Diffuse	15	
Gypsum content, %	There is no gypsum		0
	Weak-gypsum	0	5
	Medium-gyp-sum bearing	5	15
	Strongly hygroscopic	15	60
	Very strongly hyposonic	60	
Acid of water in soil (pH), conventional unit from 0 to 14	Ultra acidic	0	3.5
	Extremely acidic	3.5	4.4
	Very strongly acidic	4.5	5.0
	Strongly acidic	5.1	5.5
	Moderately acidic	5.6	6.0
	Slightly acidic	6.1	6.5
	Neutral	6.6	7.3
	Slightly alkaline	7.4	7.8
	Moderately alkaline	7.9	8.4
	Strongly alkaline	8.5	9.0
	Very alkali	9.0	14

Description of the parameters as fuzzy numbers

As a rule, when choosing the membership function type, preference is given to the forms that are easier to compute. To describe the membership function of a set of diagnostic parameters of soil, we shall use a trapezoidal function. This choice is based on the following reasoning: the numerical values of the parameters under consideration vary within certain intervals; accuracy of measurement of point values has errors; the boundaries of the intervals of values are points.

Using the parameter of soil pH (acidity) as an example (Table 4), we analytically build a trapezoidal membership function.

To this end, we number the values of the soil pH parameter $i = 1,...,11$. Let a_i be the sequence of the upper limit values of these parameters, and $a_0 = 0$ – the

lower limit value "Ultra acidic". We call the interval $[a_i, a_{i+1}]$ the original carrier of the i-th value of the parameter in question. For the value with the serial numbers $i = 2,3,...,10$, we shall propose the membership function of the following form:

$$\mu(x,i) = \begin{cases} 0, & x \le \dfrac{a_{i-1}+19a_i}{20}, \\ \dfrac{20x-(19a_i+a_{i-1})}{a_{i+1}-a_{i-1}}, & \dfrac{a_{i-1}+19a_i}{20} \le x \le \dfrac{19a_i+a_{i+1}}{20}, \\ 1, & \dfrac{19a_i+a_{i+1}}{20} \le x \le \dfrac{a_i+19a_{i+1}}{20}, \\ \dfrac{20x-(19a_{i+1}+a_i)}{a_{i+2}-a_i}, & \dfrac{a_i+19a_{i+1}}{20} \le x \le \dfrac{19a_{i+1}+a_{i+2}}{20}, \\ 0, & \dfrac{19a_{i+1}+a_{i+2}}{20} \le x. \end{cases} \quad (1)$$

Formally assuming e.g. $a_{-1} = -1$, $a_{12} = 15$, we can also extend formula (1) to the values $i = 1$ and $i = 11$.

As an example, let us give the value "Very strongly acidic" for the soil pH parameter:

$$\mu_{pH}(x,3) = \begin{cases} 0, & x \le 4.355, \\ (20x-89)/1.5, & 4.355 \le x \le 4.43, \\ 1, & 4.43 \le x \le 4.97, \\ (20x-954.4)/1.2, & 4.97 \le x \le 5.03, \\ 0, & 5.03 \le x. \end{cases}$$

For clarity, the graph of the function $\mu_{pH}(x,3)$ is given in Fig. 1.

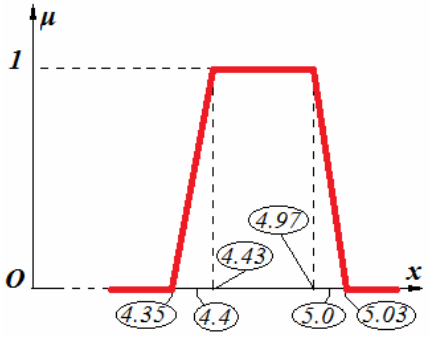

Fig. 1. Graph of the membership function for the value "Very strongly acidic" of the soil pH parameter

It should be noted that the functions $\mu(x,i)$ are not symmetric functions, i.e. their graphs are not isosceles trapezoids. It is not too difficult to see that the function $\mu(x,i) = 1$ is within 90% of the original carrier of the values of the parameter in question and covers 5% of the original carriers of adjacent values of the same parameter (Fig. 2).

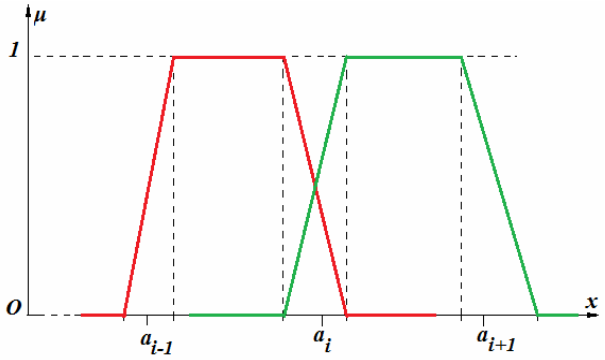

Fig. 2. Overlapping of the supports of the membership function in adjacent intervals

The overlapping of the intervals of the supports of the constructed functions $\mu(x,i)$ provides an explanation to the disputes related to soil clustering in various soil science schools.

In the following paragraphs, we demonstrate that in the case of coincidence of the values of the membership function with the adjacent values, the soil classification algorithm chooses a value with a smaller serial number, thereby introducing a certain clarity in the process of clustering.

Description of the soil classification algorithm

During soil diagnostics, soil is studied first: the soil is sectioned, the soil parameters are measured (some parameters are determined in the field and some in the laboratory).

To identify the diagnostic horizons, the soil parameters are compared with the parameters of the horizons, using the compositional inference rule. Thus, it is established which horizons exist in the given soil. This operation is additive and does not depend on the sequence of application of the values of some or other parameter.

Further, proceeding from the available diagnostic horizons in the soil under investigation, it is possible to identify to which of the 32 reference soil groups it belongs. For this purpose, the main diagnostic indicators of the soil are successively checked in accordance with Table 5.

Table 5 – **Sequence of genetic horizons for soil identification according to the WRB classification**

#	Genetic horizon	Parameters
1	**2**	**3**
1	Anthraquic horizon	• Munsell colours; • Platy structure; • Bulk density; • Thickness of soil.
2	Argic horizon	• Illuvial accumulation of clay; • Predominant pedogenetic formation of clay in the subsoil; • Selective surface erosion of clay; • Biological activity; • Coefficient of linear extensibility; • Has a texture class of loamy sand or finer;

1	2	3
		• Thickness; • Has clay content or clay in the fine earth fraction; • Soil organic carbon.
3	Calcic horizon	• Calcium carbonate; • The degree of effervescence • Secondary carbonates; • Thickness of soil; • Porosity; • Munsell colours; • pH water; • Consist of gypsum;
4	Cambic horizon	• Texture class; • Munsell colours • Structure; • Clay content; • Carbonates; • Consist of gypsum.
…	…	…
35	Thionic horizon	• pH water; • Mottles or coatings with accumulations of iron or aluminium sulfate or hydroxysulfate minerals; • Direct superposition on sulfidic material; • Thickness of soil.
36	Umbric horizon	• Structure soil; • Soil organic carbon; • Munsell colour; • Teksture soil; • Thickness of soil.
37	Vertic horizon	• Clay content; • Wedge-shaped soil aggregates; • Slickensides; • Shrink-swell cracks; • Thickness of soil; • Coefficient of linear extensibility.

Further, comparing the diagnostic indicators of soils with the obtained diagnostic horizons (in the strict sequence of the given algorithm), the first reference soil group on the list is found.

Example

Let us give an example by considering sierozem soils. The monograph by M. P. Babayev [11] defines the indicators of these soils as follows:

- soil with anthropogenically transformed horizon thicker than 50 cm, root penetration depth 20-30 cm;
- no horizon over 40 cm, by two thirds consisting of weakly decomposed organic residues;
- no technogenic inclusions;
- no cryic horizon;
- thin section in sierozem soil over 30 cm;
- 42% clay in the 25-30 cm thick horizon in sierozem soil.

The membership function for such indicators as "soil thickness", "organic residues", "root penetration depth", "technogenic inclusions", "thickness of the cryic horizon", "thickness of thin section", "thickness of the argic horizon".

The results of the fuzzy inference about the identification of the parameter values are given below in Table 6.

Table 6 – **Description of the sequence of operations in the soil identification on the basis of WRB**

Sequence of reference soil groups	Diagnostic indicators of soils	Result
Histosols	• Thickness of the horizon; • Organic residues	No
Anthrosols	• Depth of penetration of the root mass in the anthropogenically transformed horizon	No
Technosols	• The presence of technogenic inclusions in the upper layer	No
Cryosols	• There is a frozen horizon	No
Leptosols	• Thickness of thin profile	No
Vertisols	• Presence of clay in the Vertic horizon with a thickness of clay horizon.	Yes

Thus, the soil under investigation, according to the granulometric composition data for sierozem soils based on M. Babayev's classification of Azerbaijan soils, can be assigned to the soil group "Vertisols".

Conclusion

The paper proposes a fuzzi fication mechanism for the set of indicators of diagnostic parameters and describes the algorithm of the system of fuzzy inference about the identification of the values of the corresponding parameters in accordance with the compositional inference rule. Thus, the effectiveness of the use of fuzzy technologies in soil diagnostics by the World Reference Base for Soil Resources is demonstrated.

References

1. Isbell, R.F. (2002), "The Australian Soil Classification", *CSIRO Publishing*, Revised Edition, Melbourne, pp.1-4.
2. Tsyrybka, V. and Ustsinava, H. (2018), *Soil classification in Belarus: history and current problems*, Physical Geography Series, 14 p.
3. (1999), *Sistema Brasileiro de Classificacaode Solos*, DF: Embrapa, Brasília, 20 p., available at: https://livimagens.sct.embrapa.br/amostras/00053080.pdf
4. McKeague, J.A. (1995), *Soil suruey and genesis and classification research in Canada*, 7 p., available at: https://www.nrcresearchpress.com/doi/pdf/10.4141/cjss95-002
5. (1997), *Classification of soils of Russia*, Ins.of Soil named V.V. Dokuchayev, Moscow, 235 p.
6. (2003), *Keys to Soil Taxonomy*, United States Department of Agriculture, Natural Resources Conservation Service, 9 Edition, 372 p., available at: https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcs142p2_051544.pdf
7. Samofalova, I.A. (2012), *Modern problems of classification of soils*, Tutorial, Perm, 175 p.

8. (2006), *World reference base for soil resources 2006: A framework for international classification, correlation and communication*, Food and agriculture organization of the united nations Rome, 145 p., available at: http://www.fao.org/3/a-a0510e.pdf
9. (2015), *IUSS Working Group WRB*, "World Reference Base for Soil Resources 2014", update 2015, "International soil classification system for naming soils and creating legends for soil maps". World Soil Resources Reports No. 106. Rome: FAO, 198 p.
10. Zadeh, L.A. (1975), "The Concept of a Linguistic Variable and its Application to Approximate Reasoning-I", *Information Sciences*, vol. 8(4), pp. 301-357.
11. Babaev, M.P., Hasanov, V.Kh., Dzhafarova, Ch.M. and Huseynova, S.M. (2011), *Morphogenetic diagnostics, nomenclature and classification of Azerbaijan soils*, Science, Baku, 201 p.

Received (Надійшла) 20.12.2019
Accepted for publication (Прийнята до друку) 29.01.2020

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Хасанова Самира Афрасияб – науковий співробітник, Інститут систем управління Національної академії наук Азербайджану, Баку, Азербайджан;
Samira Afrasiyab Hasanova – Scientific researcher, Control Systems Institute of Azerbaijan National Academy of Sciences, Baku, Azerbaijan;
e-mail: samirahasanova75@gmail.com, ORCID ID: http://orcid.org/0000-0002-1595-6107.

Пашаєв Адалат Бахтіяр – завідувач лабораторією, Інститут систем управління Національної академії наук Азербайджану, Баку, Азербайджан;
Adalat Bakhtiyar Pashayev – Head of laboratory, Control Systems Institute of Azerbaijan National Academy of Sciences; Baku, Azerbaijan;
e-mail: adalat.pashayev@gmail.com, ORCID ID: http://orcid.org/0000-0001- 9208-8430.

Сабзієв Ельхан Наріман– завідувач відділу, компанія Kiber Ltd, Баку, Азербайджан;
Elhan Nariman Sabziev – Head of department, Kiber Ltd Company, Baku Azerbaijan;
e-mail: elkhan.sabziev@gmail.com, ORCID ID: http://orcid.org/0000-0001-8150-9439.

Нечітка діагностика ґрунтів відповідно до Світової реферативної бази ґрунтових ресурсів

С. А. Хасанова, А. Б. Пашаєв, Е. Н. Сабзієв

Анотація. Класифікація ґрунтів залишається однією з найбільш суперечливих тем в світовій науці про ґрунті через відмінності в основоположних принципах. На сьогоднішній день багато країн розробили і використовують свої власні національні класифікації. Оскільки наукова класифікація полегшує успішне дослідження і правильне використання ґрунтів, виникає проблема представлення однієї класифікації в інший. У статті розглядається можливість застосування нечітких технологій для діагностики ґрунтів на основі Всесвітньої Довідкової Бази ґрунтових ресурсів. Запропоновано трапецеїдальних форма функцій приналежності набору параметрів діагностичних горизонтів. Алгоритм нечіткої діагностики представлений на основі запропонованих функцій. Висновок. У статті запропоновано механізм фазифікації для набору індикаторів діагностичних параметрів і описаний алгоритм системи нечіткого виведення про ідентифікацію значень відповідних параметрів відповідно до правила композиційного виведення. Таким чином, доведена ефективність використання нечітких технологій в діагностиці ґрунту Світовий реферативної базою ґрунтових ресурсів.

Ключові слова: інформаційна модель; ґрунт; нечіткі технології; класифікація ґрунтів; діагностичний горизонт; діагностичні параметри.

Нечеткая диагностика почв в соответствии с Мировой реферативной базой почвенных ресурсов

С. А. Хасанова, А. Б. Пашаев, Э. Н. Сабзиев

Аннотация. Классификация почв остается одной из самых противоречивых тем в мировой науке о почве из-за различий в основополагающих принципах. На сегодняшний день многие страны разработали и используют свои собственные национальные классификации. Поскольку научная классификация облегчает успешное исследование и правильное использование почв, возникает проблема представления одной классификации в другой. В статье рассматривается возможность применения нечетких технологий для диагностики почв на основе Всемирной Справочной Базы почвенных ресурсов. Предложена трапецеидальная форма функций принадлежности набора параметров диагностических горизонтов. Алгоритм нечеткой диагностики представлен на основе предложенных функций. **Вывод.** В статье предложен механизм фаззификации для набора индикаторов диагностических параметров и описан алгоритм системы нечеткого вывода об идентификации значений соответствующих параметров в соответствии с правилом композиционного вывода. Таким образом, доказана эффективность использования нечетких технологий в диагностике почвы Мировой реферативной базой почвенных ресурсов.

Ключевые слова: информационная модель; почва; нечеткие технологии; классификация почв; диагностический горизонт; диагностические параметры.