Application of fuzzy set theory for integral assessment of agricultural products quality

N M Derkanosova¹, I N Ponomareva¹, G V Shurshikova², O A Vasilenko¹

¹ Voronezh State Agricultural University named after Emperor Peter I, 1, Michurina St., Voronezh, 394087, Russia
² Voronezh State University, 1, University Square, Voronezh, 394018, Russia

E-mail: kommerce05@list.ru

Abstract. The methodology of integrated assessment of quality and safety of agricultural products, approved by the example of indicators of wheat grain in relation to the provision of consumer properties of bakery products, was developed. Determination of the level of quality of the raw ingredients will allow direct using of agricultural raw materials for food production, taking into account ongoing technology, types of products, and, respectively, rational use of resource potential of the agricultural sector. The mathematical tool of the proposed method is a fuzzy set theory. The fuzzy classifier to evaluate the properties of the grain is formed. The set of six indicators normalized by the national standard is determined; values are ordered and represented by linguistic variables with a trapeziform membership function; the rules for calculation of membership functions are presented. Specific criteria values for individual indicators in shaping the quality of the finished products are considered. For one of the samples of wheat grain values of membership functions of the linguistic variable “level” for all indicators and the linguistic variable “level of quality” were calculated. It is established that the studied sample of grain obtains the 2 (average) level of quality. Accordingly, it can be recommended for the production of bakery products with higher requirements for the structural-mechanical properties bakery and puff pastry products hearth bread and flour confectionery products of the group of hard dough cookies and crackers

1. Introduction

One of the main factors of forming the quality and safety of food products is raw material characteristics. A lot of the requirements of normative legal and normative documents, as a rule, are of the restrictive character, allowing making a general decision without regard to the specific technologies, intended use of the products, and other factors. In addition, each of them plays an unequal role in the aggregate performance of the raw ingredients from the standpoint of the formation of certain characteristics of the final product such as the organoleptic, physico-chemical properties, the level of security, food, biological value, etc. [1, 2]. Discussing this problem in relation to the quality of bakery products traditionally included in the diets of various population groups, it is advisable to consider the complex assessment of the level of quality and safety of primary raw ingredient such as flour. In turn, the specification of flour is a direct consequence of the characteristics of grain. Taking into account the distribution of types of bakery products in the consumer market wheat grain is of maximum interest.

The purpose of this study was to develop and test the methods of integrated assessment of quality
and safety of agricultural products apprrobad on wheat, as a factor of ensuring the consumer properties of bakery products.

2. Materials and methods
The object of study was a wheat grain of class 3 according to GOST R 52554-2006. The study of wheat was carried out by standardized methods according to GOST 10987-76, 27676-88 GOST, GOST 10846-91, 54478-2011 GOST R, GOST R 54895-2012.

The mathematical tool of the proposed method is fuzzy set theory. The methodology provides for a consistent implementation of the four stages. Let us consider them for the solution of the quality problems. Similarly, the problem of assessing the level of security is solved.

Step 1. Determination of a plurality of indicators characterizing the properties of the grain. Each indicator is set in line to variable \( X_i \) (\( i=1,\ldots, n \)), where \( n \) is the number of indicators. The indicators are defined in such a way that the growth of each individual measure is associated with an increase or with a decrease (for example, when implying ash content) of the level of grain quality.

Step 2. The definition of the classifier to evaluate the properties of grain as varieties of a so-called "grey" scale of Pospelov [3], which is polar (the opposition) to a scale in which the transition from properties \( A^+ \) to property \( A^- \) takes place smoothly. The scales satisfy the conditions:

a) mutual compensation between the properties of \( A^+ \) and \( A^- \) (what is more evident of \( A^+ \), \( A^- \) manifests itself less, and vice versa).

b) the presence of neutral point \( A_0 \), interpreted as the point of greatest controversy, in which both properties are present equally.

Step 3. Evaluation of the indicator in terms of its influence on the quality level depending on its value using a linguistic variable. Let us introduce linguistic variable \( b_i=\text{"value } X_i \text{"} \). A universal set for variable \( b_i \) is interval \([\text{n},\text{m}]\), where \( \text{n} \) is minimal, \( \text{m} \) is the maximum value of the indicator, and the range of variable \( b_i \) is term-set \( B_i \). Let us assume that each linguistic variable has a trapeziform membership function, which can be defined by four numbers, i.e. a membership function of each term is of the form (1).

\[
\mu(x) = \begin{cases} 
0, & \text{if } x < a_1; \\
\frac{x-a_1}{a_2-a_1}, & \text{if } a_1 \leq x < a_2; \\
1, & \text{if } a_2 \leq x \leq a_3; \\
\frac{x-a_3}{a_4-a_3}, & \text{if } a_3 < x \leq a_4; \\
0, & \text{if } x > a_4. 
\end{cases}
\]  

Let us define the term set for each indicator and use the term set of three elements-values, i.e:

\( B_1 \) - "low level";

\( B_2 \) - "average index level";

\( B_3 \) - "high level".

Step 4. Determination of compliance of a set of values of the indicators characterizing the level of quality \( X_i \) (\( i=1,\ldots, n \)), where \( n \) is the number of indices of the statements on the level of quality.

For the formation of the transition rules from the indices values to the linguistic variables, it is advisable to determine the weight (importance) of the indicator according to the degree of contribution to the level of quality, i.e. to compare each indicator \( X \), its weight \( r \) that determines the contribution of the indicator to the level of quality. If the weights of the indicators are ordered, i.e., there is information on that, and there is no more information about these values, the weight is determined by the Fishburn rule [4]:

\[
\text{Fishburn rule: } r_i = \frac{X_i}{\sum X_j}. 
\]
\[ r_i = \frac{2(n - i + 1)}{(n-1)n} \]  

(2)

If the indicators are equally preferable or there is no system of preferences, let us assume that they have equal weight:

\[ r_i = \frac{1}{n} \]  

(3)

With the selected weighting system of the indicators, the rule of transition from the values of quality to the scales of terms of linguistic variable \( g \) is of the form [4]:

\[ p_k = \sum_{i=1}^{n} r_i \mu_{ki}, \quad k = 1, 2, 3. \]  

(4)

Calculating the observed weight of each term of linguistic variable \( H_i \), let us produce the value of variable \( h \) according to the formula:

\[ h = \sum_{k=1}^{3} p_k \bar{h}_k, \]  

(5)

where \( \bar{h}_k \) is the middle of the interval, which is a support of term \( H_k \in (a_{k1} , a_{k4}) \).

3. Results and discussion

Let us approbate the methods for the assessment of the level of quality by the example of wheat of the third class. The studies of the wheat grain parties by the norms of GOST R 52554-2006 were conducted in the Voronezh branch of the Federal centre of safety and quality of grain and products of its processing assessment [5].

Let us approbate the methods for the assessment of the level of quality by the example of wheat of the third class. The studies of the wheat grain parties by the norms of GOST R 52554-2006 were conducted in the Voronezh branch of the Federal centre of safety and quality of grain and products of its processing assessment [5].

In accordance with the above mentioned stages:

1. from the indicators specified by the national standard defining consumer properties of bakery products, let us choose: \( X_1 \) being a mass fraction of protein, %; \( X_2 \) being a mass fraction of wet gluten, %; \( X_3 \) being quality of wet gluten, values of device IDG; \( X_4 \) number of falling c; \( X_5 \) being vitreousness, %; \( X_6 \) being nature, g/l.

2. Let us form the classifier as a fuzzy linguistic description on interval \([0,1]\) [4]. Let us introduce linguistic variable \( h = " \) - the value of the quality level of the wheat grain of the 3\textsuperscript{rd} class." A universal set for variable \( g \) is interval \([0,1]\), and a set of values of variable \( h \) being term set \( H = \{ H_1 , H_2 , H_3 \} \), where \( H_i \) is determined as follows: \( H_1 = " \) quality level 1", the highest from the point of view of the influence on the quality of bakery products; \( H_2 = " \) quality level 2"; \( H_3 = " \) quality level 3" being the lowest.

Each term from set \( H \) is the name of a fuzzy subset of interval \([0,1]\). Let us consider these subsets as fuzzy trapeziform numbers (formula 1).

Let us set up a table of membership functions of each term (tab. 1), using the formula of the membership function of trapeziform fuzzy number \( x = (a_1 , a_2 , a_3, a_4) \) (1).

| Term \( H_k \) | Functions of fuzzy set membership \( H \) |
|----------------|----------------------------------------|
| \( H_1 = " \) quality level 1" \( H_1 \in [0; 0.4] \) | \( \mu_{i} = \begin{cases} 
1, & \text{if } 0 \leq h \leq 0.2; \\
5(0.4 - h), & \text{if } 0.2 < h \leq 0.4 
\end{cases} \) |
H₂⇒«quality level 2»
\[ H₂ ∈ (0.2; 0.8) \]
\[ \mu₂ = \begin{cases} 1 - 5(0.4 - h), & \text{if } 0.2 < h \leq 0.4; \\ 1, & \text{if } 0.4 < h \leq 0.6 \\ 5(0.8 - h), & \text{if } 0.6 < h \leq 0.8 \end{cases} \]

H₃⇒«quality level 3»
\[ H₃ ∈ (0.6; 1) \]
\[ \mu₃ = \begin{cases} 1 - 5(0.8 - h), & \text{if } 0.6 < h \leq 0.8; \\ 1, & \text{if } 0.8 < h \leq 1 \end{cases} \]

In the formulas of membership functions, discarded intervals on which the function take the zero value.

3. Let us define the term set of three elements, for each indicator \( X_i \), i.e. \( B_i = \{ B_{i1}, B_{i2}, B_{i3} \} \). In table 2, values of indicators are ordered by terms. The term is represented by four numbers \( (a_1, a_2, a_3, a_4) \), which correspond to trapeziform membership functions. The numbers are determined either on the basis of expert polls, or formally. So, in connection with the specific quality index wet gluten (\( X_3 \)), the low and medium level is defined by pairs of numbers \( x = (a_1, a_2, a_3, a_4) \).

For formal definitions of four numbers, one can use a number of factors. Thus, if the index value is proportional to the value of the quality level, the number of coefficients \( k_j \) is represented as sequence \((0; 0.2; 0.4; 0.6; 0.8; 1)\). Thus, to obtain the ranges for the linguistic variable according to the formula, for example, (by the example of the mass fraction of protein of the class 3 wheat grain):

\[ a_1 = X_{1\text{min}} + 0 \times (X_{1\text{max}} - X_{1\text{min}}) = 12.0; \]
\[ a_2 = X_{1\text{min}} + 0 \times (X_{1\text{max}} - X_{1\text{min}}) = 12.3; \]
\[ a_3 = X_{1\text{min}} + 0.2 \times (X_{1\text{max}} - X_{1\text{min}}) = 12.6; \]
\[ a_4 = X_{1\text{min}} + 0.4 \times (X_{1\text{max}} - X_{1\text{min}}) = 12.9 \]

This rule was used to determine the values of terms, as shown in table. 2.

**Table 2.** Evaluation of acceptable values of indicators in the form of linguistic variables for wheat of the 3rd class

| Index \( X_i \) | Value of index | \( B_{i1} \) | \( B_{i2} \) | \( B_{i3} \) |
|-----------------|----------------|-------------|-------------|-------------|
| \( X_1 \) | From 12.0 to 13.5 % of SV | 12.0; 12.0; 12.3; 12.6; 12.9; 13.2; 13.5; 13.8 | 12.0; 12.0; 12.3; 12.6; 12.9; 13.2; 13.5; 13.8 | 12.0; 12.0; 12.3; 12.6; 12.9; 13.2; 13.5; 13.8 |
| \( X_2 \) | From 23 to 28 % | 23.0; 23.0; 24.0; 25.0; 26.0; 27.0; 28.0; 29.0 | 23.0; 23.0; 24.0; 25.0; 26.0; 27.0; 28.0; 29.0 | 23.0; 23.0; 24.0; 25.0; 26.0; 27.0; 28.0; 29.0 |
| \( X_3 \) | From 20 to 100 units of device IDG | 20.0; 20.0; 25.0; 30.0; 40.0; 45.0; 50.0; 60.0 | 20.0; 20.0; 25.0; 30.0; 40.0; 45.0; 50.0; 60.0 | 20.0; 20.0; 25.0; 30.0; 40.0; 45.0; 50.0; 60.0 |
| \( X_4 \) | From 150 to 200 c | 150; 150; 160; 170; 180; 190; 200; 200 | 150; 150; 160; 170; 180; 190; 200; 200 | 150; 150; 160; 170; 180; 190; 200; 200 |
| \( X_5 \) | From 40 to 60 % | 40; 40; 44; 48; 52; 56; 60; 60 | 40; 40; 44; 48; 52; 56; 60; 60 | 40; 40; 44; 48; 52; 56; 60; 60 |
| \( X_6 \) | From 730 to 750 g/l | 730; 730; 734; 738 | 730; 730; 734; 738 | 730; 730; 734; 738 |

4. Let us define the values of the membership function, where \( i \) is the index (\( i = 1..6 \)), \( j \) is the index term (\( j = 1..3 \)) (formula 1). The results for one of the selected samples of wheat grain are presented in table 4. In the last line of table 3, there is the calculated weight of terms \( p_k, k = 1, 2, 3 \) by formula (4),
provided that all the indicators are still preferred: \( r_i = 1/n = 1/6 \), for \( i=1..6 \). In this case, the weight of a term is the arithmetic mean of the values of membership functions \( \mu_{ik} \) of term \( B_{ik} \).

\[
\text{Table 3. Primary processing of quality indicators for a sample of wheat}
\]

| Index \( X_i \) | Actual value | \( \mu_{i1} \) | \( \mu_{i2} \) | \( \mu_{i3} \) |
|-----------------|--------------|---------------|---------------|---------------|
| \( X_1 \)      | 12.5         | 0.33          | 0.67          | 0             |
| \( X_2 \)      | 25.0         | 0             | 1             | 0             |
| \( X_3 \)      | 81.0         | 0             | 1.0           | 0             |
| \( X_4 \)      | 193.0        | 0             | 0             | 1.0           |
| \( X_5 \)      | 55.0         | 0             | 0.25          | 0.75          |
| \( X_6 \)      | 745          | 0             | 0.25          | 0.75          |

Weight term \( p_k \) by linguistic variable \( h \)

\[
p_k = \frac{1}{6} \sum_{i=1}^{6} \mu_{ik}, \quad k = 1, 2, 3.
\]

Next, let us compute the value of the membership function of linguistic variable \( h=\text{"quality level"} \) for a sample in accordance with formulas (4) and (5) (table. 4). The transition from weight term \( p_k \) to the linguistic variable value of \( h \) due to the fact that the highest, the first quality level \( H_1=\text{"quality level } 1\)\) is achieved by two values \( B_{ij} \text{ "high level } X_i \)\), so \( p_j = p_k, \quad j = 3 - k + 1, \quad k = 1...3 \).

\[
\text{Table 4. Calculation of the values of linguistic variable } h=\text{"quality level"} \text{ for a sample of wheat}
\]

| Weight term \( p_j \) by linguistic variable \( h \) | Set transmitter \( j\)-th term by linguistic variable \( h \) | Middle interval \( h_j = p_j \tilde{h}_j \) |
|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|
| weight \( p_1 = 0.417 \)                      | \( H_1 \in [0; 0.4] \)                        | 0.2                                           | 0.0834                                        |
| weight \( p_2 = 0.528 \)                      | \( H_2 \in (0.2; 0.8] \)                      | 0.5                                           | 0.2640                                        |
| weight \( p_3 = 0.055 \)                      | \( H_3 \in (0.6; 1] \)                        | 0.8                                           | 0.0440                                        |
| \( h = \sum_{j=1}^{3} p_j \tilde{h}_j = \)   |                                               |                                               | 0.3914                                        |

*See the last line of table 3.

Using the standardized national standard, let us find values of the membership functions if \( h=0.3914 \):

for \( H_1=\text{"1, the highest level of quality"} \)

\[
\mu_1(0.3914) = 5 \cdot (0.4 - 0.3914) = 0.043
\]

for \( H_2=\text{"2, quality level 2"} \)

\[
\mu_2(0.3914) = 1 - 5 \cdot (0.4 - 0.3914) = 0.957
\]

for \( H_3=\text{"3, the lowest quality level"} \)

\[
\mu_3(0.3914) = 0.
\]

The results show that the studied sample of grain is the 2\text{nd} (average) level of quality with a sufficiently high value of the membership function. Accordingly, it can be recommended for
processing in technology of bakery products with high demands on structural-mechanical properties – bakery and puff pastry products of hearth bread and flour confectionery products of the group lingering and dry biscuits

4. Conclusion
The proposed method allows gradation of raw ingredients from the position of formation of consumer properties of finished products and, consequently, implementation of a mechanism of rational use of potential raw material.

References
[1] Chernykh V Y 2014 Management of the baking properties of wheat flour Bakery and confectionery production 10 12-14
[2] Chernykh V Y 2015 Methodology of control of the rheological properties of semi-finished and finished products for bakery production Bakery Russia 1 14-17
[3] Pospelov D S 1994 "Grey" and/or "black and white" [scale] Applied ergonomics. Special issue “Reflexive processes” 1 29-33
[4] Konysheva L K, Nazarov D M 2011 Fundamentals of the theory of fuzzy sets (St. Petersburg: Piter)
[5] Derkanosova N M, Gorozhanina M D, Dovlatov L S, Popova L P 2014 Statistical evaluation of the baking potential of wheat Bakery products 2 58-61