Results of the development of the integrated evaluation of the winter barley breeding identifiers

Viktor Khlystunov[1,2*], *Nina* Kopteva[3], *Yevgeny* Filippov[1], and *Nadezhda* Udintsova[3]

[1]Agricultural Research Center "Donskoy", st. Nauchny gorodok, 3, 347740, Zernograd, Russia
[2]Don State Technical University, 1, Gagarin sq., 344003, Rostov-on-Don, Russia
[3]Azov-Chernomorsk Engineering Institute – Branch of the Federal State Budgetary Educational Institution of Higher Education "Don State Agrarian University" in Zernograd, st. named after Lenin, 21, Zernograd, Russia

Abstract. Comparative analysis of the created breeding samples is essentially a multi-criteria task, since they are evaluated by dozens of features (identifiers). So, wheat has 66, barley – 63, peas – 88, sorghum – 49, corn – 63, etc. The purpose of the study is to develop a methodology for a comprehensive assessment of features (identifiers) of winter barley breeding. The object of research is a statistical array of rank values of identifiers based on the results of their expert evaluation. A rank assessment was made of 27 winter barley identifiers, which form the basis of the variety passport and reflect morphological descriptors of traits, biological properties, disease resistance, yield, crop structure, chemical composition, technological qualities of the main product, and economic use. As a result of structuring the totality of statistical material, 14 significant and 8 insignificant features were identified, forming 69% and 25% of the identifier, respectively, in terms of rank weight. A scale of 100 points was established for the reference sample of the winter barley variety (Timofey of the ARC Donskoy breeding) for significant (C) and a set of significant and unimportant identifiers (CMC). The complex indicator was calculated as the sum of points for compared varieties Zhiguli, Vivat, Yerema, Marusya of the ANC "Donskoy" selection, taking into account the weighted share of their respective characteristics in digitized form relative to the characteristics of the reference sample. It has been established that it is sufficient to carry out a comparative assessment of various varieties of winter barley by significant identifiers, since the discrepancy between numerical indicators for C and CMC identifiers did not exceed 2.36%.

1 Introduction

Currently, one of the main tasks of the agribusiness both in Russia and around the world is to increase food production, primarily through the creation of highly productive varieties and hybrids [1–5]. This is very important for our country, since under the conditions of

* Corresponding author: viptim@gmail.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

severe foreign sanctions there was an acute shortage of breeding material coming from abroad. So, its share, in particular, for vegetable growing reaches 90%.

Today, Russia is sufficiently provided with seeds of the main cereal crops, however, the decoding of the wheat genome by transnational corporations Monsanto and Sengenta may create additional difficulties.

Therefore, a possible shortage in supplying the sowing company in our country can and should be eliminated by creating new promising domestic varieties and hybrids in the shortest possible time and, if possible, with a minimum of costs with their objective assessment. These promises are solved through innovative methods of selection, [6], as well as when using relatively low-cost and affordable methods for evaluating selection results.

In essence, the task of evaluating the created varieties and hybrids of agricultural plants is a multicriteria one.

Unlike the multicriteria task of optimizing processes and phenomena, for which a sufficient number of methods, methods and techniques have been developed [7–9], this task has neither a theoretical nor an experimental solution. By the way, in [10], an attempt was made to develop a method of multi-criteria selection at the final stage of plant breeding, based on the use of a weighted sum of normalized deviations from the optimum set by the breeder for the set of breeding traits. However, the method has two significant drawbacks. So, the number of features used for the integral assessment and their contribution to the total assessment are not substantiated in any way and are not confirmed by anything. This prevents an objective assessment of breeding efforts.

The solution of the indicated problem is possible when using the method of expert evaluation.

The purpose of the research was to develop a methodology for a comprehensive assessment of winter barley breeding identifiers.

2 Methodology

In this paper, an attempt was made to develop a methodology for comparative evaluation of the results of crop breeding based on the method of expert assessments using the example of winter barley.

The idea of developing the methodology and its implementation was (Figure 1) as follows.

Based on the results of an expert assessment of the traits taken into account in the selection of winter barley in this case, they are ranked and the significant (most significant), insignificant and insignificant traits are distinguished from them. Considering separately significant (option I) and a set of significant and unimportant features (option II), determine the contribution (share) of each feature relative to the sums of the rank assessment for each of the above options. Taking into account this contribution and the ratio of the digitized value of each feature of the compared sample and the selected reference sample, the value of the complex indicator in conditional scores of the compared sample is determined as the sum in points for all features. The calculation results are implemented for several varieties.

Before the start of the research, 27 factors (characteristics) were selected, which, as a rule, are used when compiling a variety passport and in accordance with the International classifier of the genus HORDEUM. Next, we prepared a questionnaire and randomized the sequence of recording features in it using a table of random numbers [11] in order to exclude a hint of a programmed decision-making by an expert.

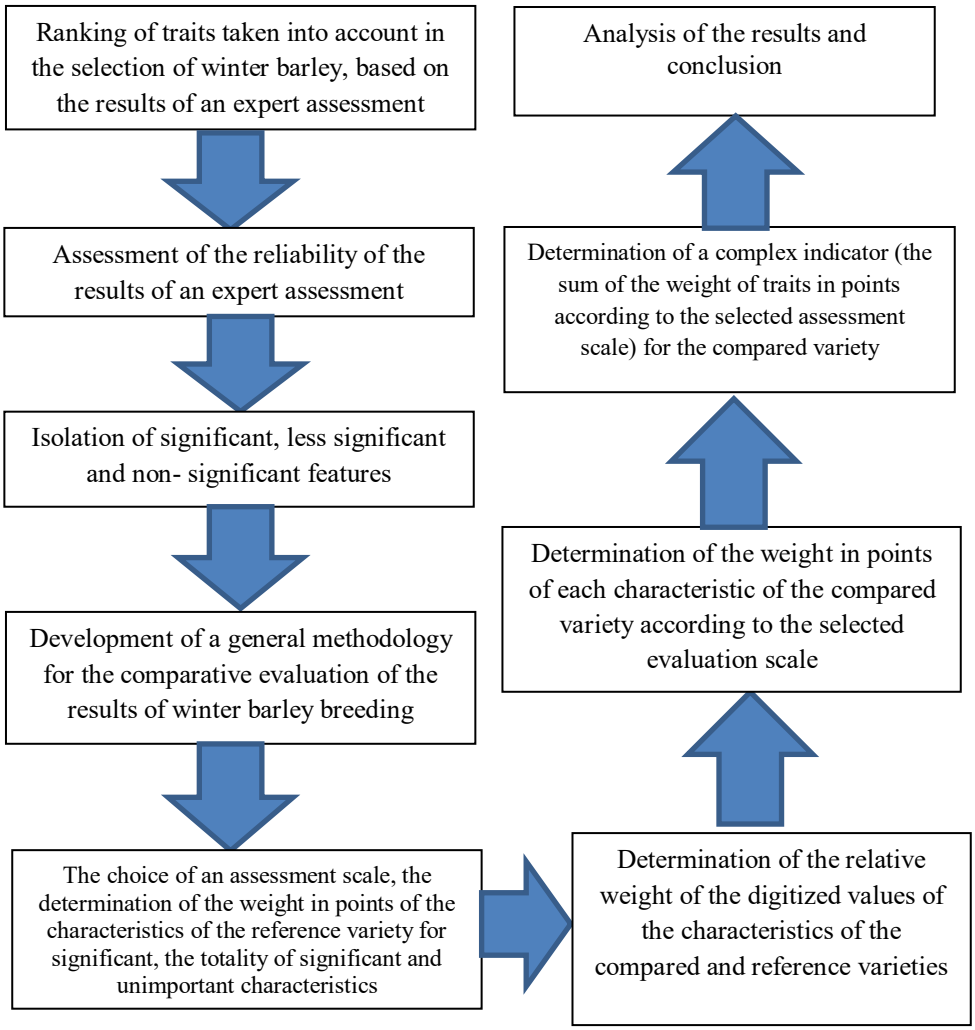


Fig. 1. The structure of the idea and its implementation in the development of a method for comparative evaluation of the results of winter barley breeding

During the survey (psychological experiment), the expert (specialist) was asked to indicate the place of each factor from 1 (the most significant) to 27 (the least significant of all). The rank assessment was carried out using the principle of "inverted scale", that is, the most significant feature of the j-th expert was assigned a rank score of 27, decreasing to 1 for the most insignificant. The rank assessment of each i-th factor was the sum of the assessments for all experts.

Breeding scientists from leading specialized scientific and educational institutions of Moscow, Voronezh, Krasnoyarsk, Krasnodar, Stavropol, Simferopol acted as experts. Their scientific work experience ranges from 15 to 40 years, the number of varieties and hybrids of barley bred is from 10 to 40.

3 Results and discussion

The results obtained (Table 1) were checked for reliability according to two criteria: the consistency of the opinions of experts and the non-randomness of their consistency.

To establish the consistency of the opinions of experts, we calculated [11] the Kendall criterion (coefficient of concordance)

$$W_{calc.} = \frac{S}{\frac{1}{12}m^2(K^3-K)-m\sum_{j=1}^{m}\frac{1}{12}\sum_{j=1}^{m}\left(t_j^3-t_j\right)}, \quad (1)$$

where m – number of experts, 7 чел.;

K –number of factors (features) in the matrix, 27;

t_j – количество одинаковых рангов у j- ого эксперта, 2.

S – the sum of squared deviations of the rank assessment of all factors for all experts is defined as

$$S = \sum_{i=1}^{K}\left(\sum_{j=1}^{m} X_{ij} - L\right)^2, \quad (2)$$

where y_{ij} – the sum of the rank assessment of the i-th factor for all experts;

L – the arithmetic mean value of the sum of the rank score for all factors for all experts, defined as

$$L = \frac{\sum_{i=1}^{K}\sum_{j=1}^{m} X_{ij}}{K}. \quad (3)$$

As a result, the calculated value of the concordance coefficient was

W_p=0,6257,

which indicates a fairly high degree of agreement among experts (upon request $0 \leqslant W_p \leqslant 1,0$).

Next, we checked the condition of non-randomness of the consistency of the opinions of experts. It is necessary [11] that, according to the Pearson criterion ($\chi 2$), the condition

$$\chi^2_{calc.} > \chi^2_{obs.} \quad (4)$$

The calculated value of the Pearson criterion was determined by the formula

$$\chi^2_{расч.} = \frac{S}{\frac{1}{12}mK(K+1)-\frac{1}{K-1}\sum_{j=1}^{m}\frac{1}{12}\sum_{j=1}^{m}\left(t_j^3-t_j\right)}, \quad (5)$$

which amounted to 113,8800, and table value [11] at 0,01 significance $\chi^2_{tab.}$=45,6420.

Thus, the consistency of expert opinions is more than accidental.

Taking into account the reliability check, the diagram of the rank evaluation of features was obtained in the following form (Figure 2).

To solve the issue of identifying significant, unimportant and insignificant features, in accordance with [13 –15], their totality was considered in the form of a multidimensional variable – an identifier. Having considered the sum of the rank assessment in this capacity, we obtained the integral empirical distribution function *yi = F(xi)* of the accumulated frequencies of the share of the value of each feature in their total set (Figure 3). The form of this function characterizes the increase in the probability of considering a feature with an increase in their number.

Since the rank estimation was carried out in accordance with the frequency of accumulation, the sharp increase in *F(x)* in the initial period turned into a relatively smooth flow, asymptotically approaching 1 at the final stage.

Table 1. Input data for calculating the Kendall and Pearson criteria

Epert (j)	**Rank assessment by i=th factor**													
	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	Y_{11}	Y_{12}	Y_{13}	Y_{14}
1	20	25	14	17	13	9	16	15	3	12	11	27	7	19
2	16	26	15	25	13	22	12	21	8	15	18	27	4	11
3	4	20	11	13	23	2	17	19	22	10	7	27	3	12
4	23	24	7	9	19	5	14	13	16	8	20	27	3	10
5	19	27	9	22	22	7	12	11	16	2	20	26	3	8
6	7	19	16	21	9	15	24	23	8	17	11	21	2	14
7	11	15	17	19	14	7	20	18	13	4	10	27	3	24
$\sum_{j=1}^{m} y\,ij$	123	156	89	126	113	67	115	120	86	68	97	182	25	98
$\sum_{j=1}^{m} y\,ij - L$	22.78	55.78	-11.22	25.78	12.78	-33.22	14.78	19.78	-14.22	-32.22	-3.22	81.78	-75.22	-2.22
Δ_i^2	518.93	3111.41	125.89	664.61	163.33	1103.57	218.45	391.25	202.21	1038.13	10.37	6687.97	5658.05	4.93

Epert (j)	**Rank assessment by i=th factor**												
	Y_{15}	Y_{16}	Y_{17}	Y_{18}	Y_{19}	Y_{20}	Y_{21}	Y_{22}	Y_{23}	Y_{24}	Y_{25}	Y_{26}	Y_{27}
1	24	10	26	26	21	23	3	6	18	5	4	22	21
2	19	9	10	24	23	17	14	3	6	1	7	20	2
3	16	1	21	26	24	8	9	5	27	1	15	18	14
4	22	4	26	25	12	21	6	2	18	1	15	17	11
5	23	10	25	24	5	18	6	4	14	1	13	17	15
6	6	4	18	27	13	10	12	3	25	1	5	26	20
7	9	5	24	23	26	8	6	2	22	1	12	16	15
$\sum_{j=1}^{m} y\,ij$	119	43	150	175	126	103	56	25	128	11	75	132	98
$\sum_{j=1}^{m} y\,ij - L$	18.78	-57.12	49.78	74.78	25.7	2.78	-44.22	-75.22	27.78	-89.22	-25.22	31.78	-2.78
Δ_i^2	352.69	3274.13	2478.05	5542.05	664.61	7.73	1955.41	5658.05	771.73	7960.21	363.05	1009.97	7.73

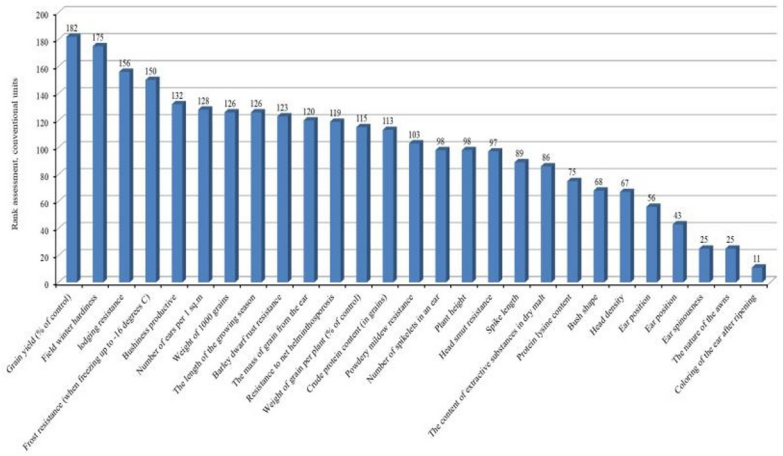
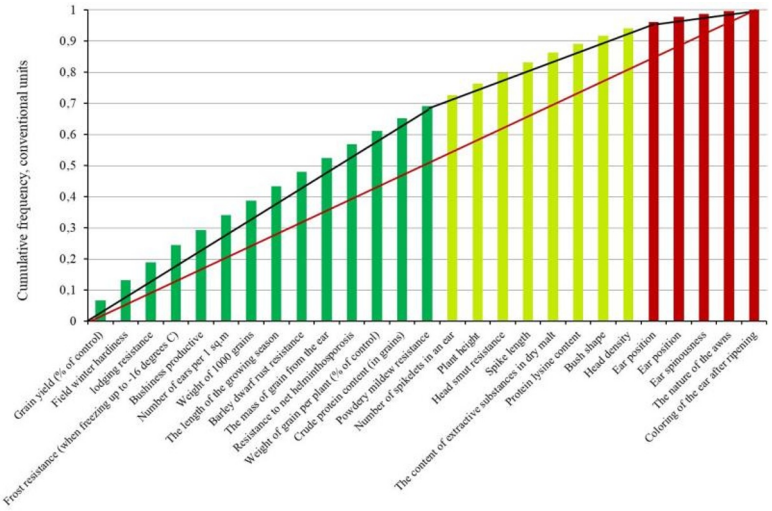


Fig. 2. Diagram of rank evaluation of traits taken into account when analyzing the results of winter barley breeding



Fi. 3. Diagram of the integral function of the accumulated frequencies of the share of the value of each feature in the total population

Here (Figure 3) it was possible to identify three characteristic areas of change in the function *F(x)*:

– the first area is the area where the growth rate of the volume of features F(x) is greater than the growth rate of the features that form this volume

$$\frac{\Delta F(x)}{\Delta x} > 1, \tag{6}$$

– the second area is the area where the growth rate of F(x) is approximately proportional to the growth rate of features

$$\frac{\Delta F(x)}{\Delta x} \leq 1, \quad (7)$$

– the third area is the area where the rate of increase F(x) is less than the rate of increase in features

$$\frac{\Delta F(x)}{\Delta x} < 1. \quad (8)$$

According to these areas, the features of the effectiveness of winter barley breeding can be classified as significant, insignificant and non-significant. Using formulas (6–8), one can single out the ranks of those signs that satisfy these conditions. Thus, the fuzzy set described by the function *F(x)* can be structured.

Applying this method to the studied array, we obtained a graph of the accumulated frequencies δ_{af} (Figure 3).

In this figure, a straight line characterizes the average rate of formation of the array, assuming that each value makes the same contribution to the total population.

The broken line highlights three areas: K_S – significant, K_{LS} – less significant, K_{NS} – non-significant features for evaluating the effectiveness of winter barley breeding.

As a result of the implementation of the algorithm according to this method, it was found that the hypothetical rate of formation of the array (Δ1) =0,037. The condition $\delta_i > \Delta 1$ is satisfied by the first 14 elements of the array in question. That is K_S =14.

Thus, the first K_S = 14 features of assessing the effectiveness of winter barley breeding form 69% of the rank weight of the total flow of requirements (51,85% from all 27 traits by number).

These include: 12, 18, 2,17, 26, 23, 4,19, 1, 8, 15, 7, 5, 20 features. These are significant features.

According to the algorithm, Δ2=0,0769. Of the remaining 13 elements of the array, the condition $\delta_k > \Delta 2$ satisfy the following 8 criteria for evaluating the effectiveness of winter barley breeding. That is K_{LS} =8.

Thus, the following K_{LS} = 8 features form 25% of the requirements (but 29,63% of the 27 features by number).

These include: 14, 27, 11, 3, 9, 25, 10, 6. These are less significant features.

For the rest K_{NS} = 5 traits account for 6% (but 18.52% if counted from 27 traits). These are non-significant features. These include: 21, 16, 13, 22, 24 features.

Next, we determined the scale of points for significant and the totality of significant and unimportant features on the assumption that the weight of each feature in percent was:

– for significant features

$$B_{iC} = \frac{R_{iC} * 100\%}{\sum_{i=1}^{K_C} R_{iC}} \quad (9)$$

where R_{iS} – rank assessment of the i-th significant feature;

K_S – the number of significant features (values-identifiers), numerically equal to 14.

– for a set of significant and non- significant features

$$B_{iCMC} = \frac{R_{iCMC} * 100\%}{\sum_{i=1}^{K_{CMC}} R_{iCMC}} \quad (10)$$

where R_{iSLS} – rank assessment of the i-th by the totality of significant and unimportant features;

K_{SLS} – the number of a set of significant and non-significant features, numerically equal to 22.

For convenience, a 100-point rating scale was adopted, corresponding to the percentage distribution of the weight of features. Such a scale of 100 points corresponds to the sum of the points of the reference variety.

The calculation results are summarized in Table 2.

The calculation of a complex indicator for the compared variety (hybrid) is carried out according to the formulas, respectively, based on the score for significant and the totality of significant and unimportant features

$$B_{\sum C} = \sum_{i=1}^{14} B_{iC} * \delta_i \quad (11)$$

$$B_{\sum CMC} = \sum_{i=1}^{22} B_{iCMC} * \delta_i \quad (12)$$

where $B_{\sum C}$, $B_{\sum CMC}$ – the value, respectively, of the complex indicator of the compared variety (hybrid) for significant and the totality of significant and unimportant traits of winter barley breeding;

δ_i – the weight of the compared trait relative to the corresponding trait of the reference variety (hybrid), calculated as

$$\delta_i = \frac{B_{iCV}}{B_{iE}}, \quad (13)$$

B_{iCV}, B_{iE} – respectively, the value in absolute digitized expression of the value of the trait of the compared variety (hybrid) and the corresponding trait of the reference variety (hybrid).

Table 2. The results of determining the scoring scale for the characteristics of the reference variety of winter barley

№	Name of features	R_{iS}, conv. units	$Б_{iS}$, points	$R_{i(S+LS)}$, conv. units	$Б_{i(S+LS)}$, points
1.	Grain harvest	182	9,74	182	7,15
2.	Field winter hardiness	175	9,37	175	6,87
3.	Lodging resistance	156	8,35	156	6,13
4.	Frost resistance (at –16 ˚C)	150	8,03	150	5,83
5.	Bushiness productive	132	7,07	132	5,18
6.	Number of ears per 1 m^2	128	6,85	128	5,03
7.	Weight of 1000 grains	126	6,75	126	4,95
8.	The length of the growing season	126	6,75	126	4,95
9.	Dwarf rust resistance	123	6,58	123	4,83
10.	Mass of grain from the ear	120	6,42	120	4,71
11.	Resistance to net helminthosporiasis	119	6,37	119	4,67
12.	Weight of grain per plant (% of control)	115	6,16	115	4,52
13.	Crude protein content (in grain)	113	6,05	113	4,44
14.	Powdery mildew resistance	103	5,51	103	4,05
15.	Number of spikelets in an ear			98	3,85
16.	Plant height			98	3,85
17.	Dust head resistant			97	3,81
18.	Spike length			89	3,50

19.	The content of extractive substances			86	3,38
20.	Protein lysine content			75	2,95
21.	Bush shape			68	2,67
22.	Head density			67	2,62
Sum		1868	100	2546	100

For example, this value was calculated (Table 3) for the varieties of winter barley Zhiguli, Timofey, Vivat, Yerema and Marusya of the ARC "Donskoy" breeding.

As a result of calculating the complex indicator, a diagram (Figure 4) was obtained for assessing winter barley varieties for the winter barley varieties under consideration.

Table 3. The results of the calculation of the weight of the characteristics of the compared varieties of winter barley

№	Feature name	Value, dimension of the reference (normative) trait (variety Timofey) V_E	Name of winter barley variety							
			Yerema		Vivat	Marusya		Zhiguli		
			Sign value							
			B_{icom}	δ_i	B_{icom}	δ_i	B_{icom}	δ_i	B_{icom}	δ_i
1	2	3	4	5	6	7	8	9	10	11
1.	Barley dwarf rust resistance	90%	80%	0,889	90%	1	90%	1	70%	0,778
2.	Lodging resistance	90%	100%	1,111	100%	1,111	100%	1,111	100%	0,778
3.	Spike length	6 cm	6 cm	1,000	7 cm	1,167	7,2 cm	1,200	6 cm	1,000
4.	Weight of 1000 grains	40 г	38 g	0,950	42 g	1,05	48 g	1,200	38 g	0,95
5.	Crude protein content (in grain)	12%	12%	1,000	12,0%	1,000	12,2%	1,017	12,0%	1000
6.	Spike density (number of segments per 4 cm of spike rod length)	17 pcs.	18 pcs.	1,059	17 pcs.	1,000	17,3 pcs.	1,018	16,1 pcs.	0,947
7.	Weight of grain per plant (% of control)	100%	110%	1,100	110%	1,100	140%	1,406	80%	0,800
8.	Mass of grain from the ear	2 g	1,8 g	0,900	2,1 g	1,050	2,6 g	1,300	1,3 g	0,900
9.	The content of extractive substances in dry malt (in malting barleys)	76%	75%	0,987	75%	0,987	75%	0,987	75%	0,987
10.	Bush shape (angle of inclination to the horizontal)	75˚	80^0	1,067	65^0	0,867	77^0	1,028	65^0	0867

Table 3 continued

1	2	3	4	5	6	7	8	9	10	11
11.	Dust head resistant	80%	80%	1,000	90%	1,125	95%	1,188	80%	1,000
12.	Grain yield (% of control)	100%	110%	1,100	115%	1,150	130%	1,300	90%	0,900
13.	Ear awn (awn length)	12 cm	11 cm	0,917	12,3 cm	1,042	10,0 cm	0,833	10,0 cm	0,833
14.	Number of spikelets in an ear	50 pcs.	46 pcs.	0,920	52 pcs.	1,040	56 pcs.	1,120	46 pcs.	0,920
15.	Resistance to net helminthosporiasis	70%	80%	1,143	75%	1,071	80%	1,143	75%	1,071
16.	Ear position (angle of inclination to the horizontal)	75%	90%	1,000	70%	0,933	80%	1,067	65%	0,867
17.	Frost resistance (with artificial freezing up to –16 ˚C	50%	70%	1,000	60%	1,200	70%	1,400	70%	1,400
18.	Field winter hardiness	75%	90%	1,000	75%	0,833	80%	1,143	80%	1,143
19.	Duration of the growing season (compared to control)	270 days	273 days	1,050	268дн.	1,019	268 days	1,019	275 days	0,993
20.	Powdery mildew resistance	100%	100%	0,800	90%	0,900	90%	0,900	80%	0,800
21.	Caryopsis shapes (ratio of diameters in elliptical interpretation)	3,0	25	0,900	3,0	1,000	2,5	0,833	3	1,000
22.	The nature of the awns, pcs. (including smooth)	0	0	0	0	0	0	0	0	0

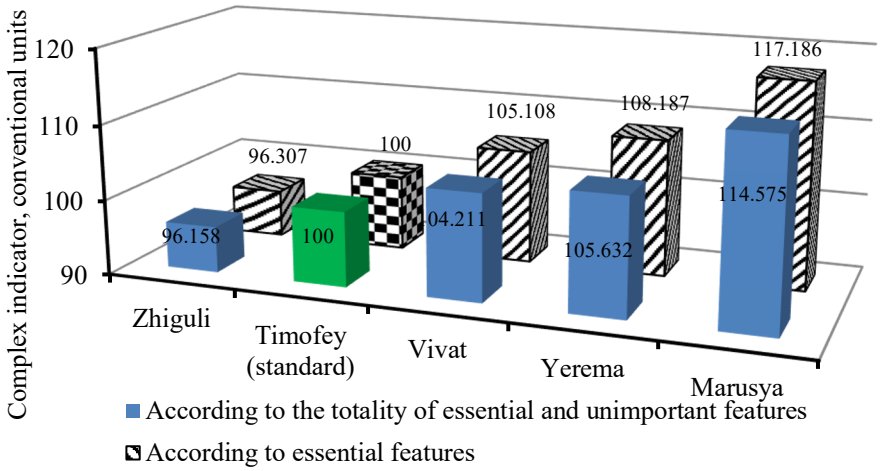


Fig. 4. Diagram of assessment of varieties of winter barley by a complex indicator

Based on the method of expert assessments, the ranking of signs (identifiers) of winter barley was carried out. Using a mathematical algorithm for structuring a stylistic array, significant, insignificant and insignificant estimated identifiers of the selection of the agricultural crop in question were identified. Together with the established ranking list of each trait, this served as the basis for the development of a methodology for comparative evaluation of the results of winter barley breeding according to a complex (integral) indicator, which was implemented in the analysis of varieties of breeding ARC "Donskoy"

At the same time, it was found that the nature of the change in the indicator of a comprehensive assessment of the results of winter barley breeding when using an evaluation scale that takes into account both essential and a complex of essential and unimportant features are similar and equidistant. Therefore, for practical use, it is quite sufficient to use data on a significant part of winter barley traits, whose contribution, according to experts, is up to 69% by rank estimate.

The reliability of the result obtained is confirmed by a sufficient degree of agreement between the opinions of experts [the concordance coefficient is 0.6257 and the non-randomness of this agreement (the calculated value of the Pearson criterion exceeds its tabular value at a significance of 0.01 by 2.5 times)].

The signs that significantly determine (up to 69%) the results of the evaluation of the selection work of winter barley in the amount of 14 have been established.

The methodology can be diversified in relation to other selectable agricultural crops.

References

1. S.S. Gosal, S.Kh. Vanya, Accelerated Plant Breeding Volume 1. Cereals. Springer International. (2020). https://doi.org/10.1007/978-3-030-41866-3
2. B. Lenarts, B.K. Collard, M. Demont, Improving global food security through accelerated plant breeding. plant science **287**, 110207 (2019). https://doi.org/10.1016/j.plantsci.2019.110207
3. H.F. Williamson, S. Leonelia, Studies in the history and philosophy of science **95**, 167-176 (2022). https://doi.org/10.1016/j.shpsa.2022.08.006

4. A. Riaz, et.al., Agronomy **11(5),** 894 (2021). https://doi.org/10.3390/agronomy11050894
5. L.V. Tsatsenko, Innovative technologies in agronomy: selection and seed production: textbook (KubdAU. Krasnodar. 2020)
6. O.V. Podinovskaya, V.V. Podinovsky, Control problems **6,** 2-8 (2014)
7. A.G. Protchenko, E.A. Kulagina, V.N. Smoryakova, Introduction to multi-criteria optimization: teaching aid. Nizhny Novgorod. Nizhny Novgorod University (2017)
8. A.S. Terteryan, A.V. Brovko, Modeling of the system and processes **15**, 1, 107-114 (2022)
9. S.P. Martynov, Agricultural biology **6**, 122-124 (1987)
10. A.D. Nasledov, Mathematical methods of psychological research. Analysis and interpretation of data: textbook. St. Petersburg. Saint Petersburg State University of Information Technologies, Mechanics and Optics (2020)
11. A.V. Goremin, O.V. Sherstyukov, Modern science-intensive technologies **5(2),** 179-183 (2022)
12. N.A. Kopteva, N.M. Udintsova, V.V. Seregina, L.N. Shapovalova, International scientific journal **6**, 49-55 (2014)
13. N.A. Kopteva, V.F. Khlystunov, M.V. Khlystunov, *Mathematical apparatus for identifying industry (product) clusters.* Socio-economic and technological development of the agro-industrial complex: state, trends, forecast: Proceedings of the international scientific-practical conference June 18-19, 2008, part 2. Rostov-on-Don. VNIIEiN, "RINH". 208. P. 77-82 (2008)
14. N.A. Kopteva, *Method for determining a sufficient number of elements of a representative array*, Improvement of processes and technical means in the agro-industrial complex, p. 124-128 (Zernograd, 1999)