Development and analysis of fuzzy expert data for technological adjustment of a grain harvester header

Valery Dimitrov\textsuperscript{1,*}, Lyudmila Borisova\textsuperscript{1}, and Inna Nurutdinova\textsuperscript{1}

\textsuperscript{1}Don State Technical University, 1, Gagarin sq, 344003, Rostov-on-Don, Russia

Abstract. The paper considers the problems of developing and presenting fuzzy expert data on external factors and adjustable parameters of the harvester header. The object domain “Technological adjustment of the harvester header” was studied. On the basis of the data, obtained from four experts a linguistic description of the problem statements was given, linguistic variables were introduced, membership functions were developed, consistency characteristic properties were calculated. The base of fuzzy expert knowledge intended for the unit of obtaining and updating knowledge of the decision support intelligent system by an operator in the field conditions was created. In order to estimate quality of the fuzzy expert data and define the degree of its suitability for application in intelligent information system we used the algorithm which provides setting the quality criteria, availability of feedback with experts to update the data, makes it possible to choose the optimal number of terms of the membership functions. The possibility of taking into account the expert data hierarchy in the given algorithm made it possible to introduce experts ranging according to their qualification, for this purpose Fishburn numbers were used as weighting factors.

1 Introduction

Decision making on setting the technological adjustment parameters of the harvester, operating in the field conditions of harvesting conditions of the particular crops under the changing factors of the environment is one of the urgent problems for conserving and ensuring the proper quality of crops. The difficulty of choosing the technological parameters values is stipulated by a number of factors, one of them is substantial and ambiguous dependency of adjustable parameters on external factors of harvesting. Another important factor is changeability of external factors and rough estimating character of their values. As a result, even a skilled operator can have difficulties in decision making about the values of the harvester adjustable parameters, and insufficiently qualified decisions lead to direct grain losses, its mechanical damage, increase of the machine downtime and harvesting period etc, and, as a result, to financial losses. So as to help the operator in decision making, intelligent information systems (IIS) are developed, which are based on

* Corresponding author: kaf-qm@donstu.ru

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
fuzzy expert knowledge. At present, fuzzy expert systems are successfully applied in different areas of agriculture [1–8], such as processing and sorting of agricultural products, computation of yield, weeds identification, harvester positioning in the field, control of its movement etc. Along with mathematical apparatus of the theory of fuzzy sets and fuzzy logics, image analysis, genetic algorithms are used in expert systems. The questions of automatic choice of some parameters of combine operation on the basis of quality indices with the application of production rules [9, 10], fault prediction [11], assessment of the machine reliability [12] are vital today.

It was indicated previously that a harvester represents itself a multilayer hierarchal system, and to control this machine (in particular for the process of decision making on technological adjustment) it is expedient to use IIS based on the linguistic approach [13, 14].

Resulting from study of the object domain, we distinguish significant external factors and adjustable harvester parameters for description of which linguistic variables are introduced. The dependency among multiple input and output variables is described on the qualitative level in the form of statements in terms of production rules, and the procedure of forming the decision represents a fuzzy logic inference based on the system of rules of fuzzy products.

One of the important conditions of suitability and efficiency of the models obtained by fuzzy modelling is qualitative initial data, corresponding to real conditions. The present paper is devoted to the creation of the expert data block for technological adjustment of the harvester header as one of its essential parts.

2 Materials and methods

Fuzzy modelling of the problem inference is divided into three stages: fuzzification, composition, and defuzzification [13]. The problem, being set in the present paper, refers to the stage of fuzzification, which can be subdivided conventionally into two sub-problems. The first one includes linguistic description of the problem conditions, and the second one – defining membership functions (MF) of the external factors and the harvester adjustable parameters. In the process of solving the sub-problems, definition of the carrier, choice of the basic and extended term-sets, checking the requirements for the MF development, definition of the MF development method, assessment of consistency of fuzzy expert knowledge, generalized MF development are performed.

Application of the methodology of linguistic approach for studying complex technical systems, the harvester belongs to, provides development of the models of the environmental factors (linguistic variables - \(X_i\)) and the harvester adjustable parameters (linguistic variables - \(Y_j\)) in the form of semantic spaces and the corresponding them MF (\(\mu_R\)) [13, 15]:

\[
\{X_i, T(X_i), U, G, M\}, \quad \mu_R(x_1, x_2, \ldots, x_n) \in (0; 1),
\]

\[
\{Y_j, T(Y_j), U, G, M\}, \quad \mu_R(y_1, y_2, \ldots, y_j) \in (0; 1),
\]

where \(T\) is a set of values of linguistic variables (LV), or terms representing the names of fuzzy variables defined by the set \(U\), \(G\) is a syntactic procedure defining the process of generating new values of LV from the set \(T\), \(M\) is a semantic procedure allowing to represent new values generated by the procedure \(G\) into fuzzy variables.

The model of the object domain “preliminary technological adjustment of the harvester” has the form:

\[
R = X \rightarrow Y,
\]
where $R$ is a fuzzy relation “environmental factors – technological adjustment parameters” which represents a fuzzy set on a direct product of $X \times Y$ of the full space of the suppositions $X$ and the full space of conclusions $Y$.

For all LVs, corresponding to the sets of external factors and adjustable parameters, it is necessary to determine basic term-sets, which are laid down on the basis of expert judgements and have the form [13]:

$$T_i = \{T_{i1}, T_{i2}, \ldots, T_{im}\}, \quad i \in K = \{1, 2, \ldots, l\}.$$ 

Here $\{T_i, X, \tilde{C}_i\}$ is a fuzzy variable, corresponding to the term $T_i \in T$; $\tilde{C}_i = \{\mu_{C_i}(x) / x\}, \quad x \in X$; $C_i$ is a carrier of the fuzzy set $\tilde{C}_i$, for which subset of LV values is used; $\mu_{C_i}(x)$ is MF.

Let us mention briefly the questions of the MF development. In the fuzzy models of decision making different procedures for the MF development are used [14]. In order to solve the given problem, it is preferable to present the MF with the help of standard functions being assigned parametrically. On the basis of such special feature as uncertainty, assumptions on the MF properties, functions form is assigned, as a rule, typical forms are used (triangle, trapezoidal etc.), parameters of which are estimated by experts.

To choose the optimal model as a criterion it is expedient to use indices of pair consistency, indices of fuzziness, indices of overall and pair consistency [14, 16].

General consistency of the set of models of expert estimation of an attribute is determined by additive $k_a$ and multiplicative $k_m$ indices [13]:

$$k_a = \frac{1}{m} \sum_{i=1}^{m} \frac{\int_{0}^{1} \min_{x \in X} \mu_{ij}(x) dx}{\int_{0}^{1} \max_{x \in X} \mu_{ij}(x) dx}, \quad k_m = \prod_{i=1}^{m} \frac{\int_{0}^{1} \min_{x \in X} \mu_{ij}(x) dx}{\int_{0}^{1} \max_{x \in X} \mu_{ij}(x) dx},$$

where $l=1, 2, \ldots, m$ - term number, $i=1, 2, \ldots, k$ - expert number, $\mu_{ij}(x)$ - MF, which was assigned by $i$-th expert for $l$-th term.

Fuzziness index $d_{ij}^l$ between the models of two experts, $i$-th and $j$-th, in terms of $l$-th term is determined as a linear distance (Hemming) between fuzzy sets with MF $\mu_{ij}(x)$ and $\mu_{ji}(x)$ [13], and the index of pair consistency between the models – as a magnitude $k_{ij}^l$ [14, 16]:

$$d_{ij}^l = \int_{0}^{1} |\mu_{ij}(x) - \mu_{ji}(x)| dx, \quad k_{ij}^l = \frac{\int_{0}^{1} \min[\mu_{ij}(x), \mu_{ji}(x)] dx}{\int_{0}^{1} \max[\mu_{ij}(x), \mu_{ji}(x)] dx}.$$ 

The magnitude $d_{ij}^l$ form a matrix of fuzziness indices $D^l$, and the magnitudes $k_{ij}^l$ - a matrix of pair consistency $K'$ of the models of the $i$-th and $j$-th experts for the $l$-th term.

On the basis of matrices $D^l$ and $K'$ there are index matrices of fuzziness $D$ and consistency $K$ of the models on all terms. Their elements are determined by the formulae:
\[ d_{ij} = \frac{1}{m} \sum_{l=1}^{m} d_{ij}^l, \quad k_{ij} = \frac{1}{m} \sum_{l=1}^{m} k_{ij}^l. \] (3)

The quality of expert data and the adequacy degree of formal description of a real situation are mainly determined by the consistency indices, therefore, expert data consistency estimation occupies a central position in the algorithm of expert data quality estimation, suggested in [17]. The given algorithm provides assignment of the expert data quality criteria, availability of feedback with experts in order to correct data in case of failure of the assigned criteria, makes it possible to choose the optimal number of terms of the membership functions, gives the opportunity to take into account expert data hierarchy. However, in this algorithm there is no detailed description of the step where the generalized MF development takes place, and it is of vital importance, since these same MF are used in the system of fuzzy logic inference. Estimates of the MF parameters, obtained from different experts, are averaged. An essential point is the choice of averaging mode. Arithmetic mean or weighted average of the parameters estimates is used, and in the function of weighting coefficients Fishburn numbers are applied. The fundamental principle for ranking may be the consistency degree of the data, obtained from each expert, with the others [16]. Another criterion may be the experts’ qualification, which, in a number of cases, is not insignificant, since such an approach makes it possible to give the priority to well-trained experts and, in addition, to take into account the input of the others. The algorithm of the expert data quality estimation, reflecting the problems of choice of the MF terms number, is presented in fig.1.

![Diagram](image-url)

**Fig. 1.** The diagram of the improved algorithm for estimating quality of the expert data.
Results of investigations

The study of the object domain “preliminary technological adjustment of the harvester header” has made it possible to identify external factors, exerting a significant impact on the header adjustable parameters. Among these are crop yield, grain dockage, humidity, height, lodged grain, density, stand of grain disarrangement. In accordance with the mentioned factors, LVs were introduced, linguistic scales were determined. It should be noted that these scales depend on the crops being harvested and may be different. In addition, it may become necessary to differentiate such factors as crop yield, for example, in case of harvesting wheat it is expedient to consider the yield for different values, in particular, yield is about 50 q/ha, 40 q/ha etc. Output LVs are adjustable parameters of the harvester header: vertical position of a reel relative to the cutting tool, a reel’s horizontal position (reel offset), cutting height, a clearance between auger flights and the header bottom, a clearance between the auger pins and the header bottom, a clearance between a spacer beater and the bottom. Let us present a linguistic description of some external factors and the adjustable parameters for one crop – wheat (at the yield of about 40 q/ha). The stated below number of terms is optimal and chosen on the account of the result of the algorithm application for estimating the quality of the expert data.

LV tuples “Crop yield - 40”, “Stand of grain dockage”, “Stand of grain height”, “Stand of grain lodging”, “Stand of grain density”, “Reel vertical position”, “Reel horizontal position (reel offset)” correspondingly have the form:

\[
\text{CY} = \{\text{CYL40, CYA40, CYM40}\};
\]

\[
\text{SGD} = \{\text{LSGD, HSGD, %}\};
\]

\[
\text{SGH} = \{\text{LSGH, BMSG, MSG, HSG, cm}\};
\]

\[
\text{SGL} = \{\text{USGL, LSGL, MSGL, HSGL}\};
\]

\[
\text{SGDe} = \{\text{SD, BMD, MD, TD}\};
\]

\[
\text{RVP} = \{\text{VL, L, M, H cm}\};
\]

\[
\text{RHP} = \{\text{S, BM, M, F cm}\}.
\]

Four experts were involved to estimate the LP parameters. As the LVs under study were determined on numerical intervals, parametric fuzzy numbers with typical (trapezoidal and triangle) MFs were used as LV terms [15]. The developed MFs of the LV terms ought to meet necessary conditions of sequence, normality, completeness, consistency, limitation [15].

Initial conditions to be established for applying the algorithm (Fig.1) of expert data quality estimation are as follows: expert ranks will be defined according to experts’ qualification, weight factors for averaging parameters will be calculated as Fishburn numbers, coefficients critical values (1) - (3) will be assumed as follows $k_{cr} = 0.75; k_{ij}^{cr} = 0.75; d_{ij}^{cr} = 0.15$.

Values of the consistency parameters were determined with the help of the original software package. For the purpose of illustration table 1 presents matrices $K$ and $D$ of pair consistency and fuzziness indices (2) - (3), and table 2 presents indices $k_a$ and $k_m$ of general consistency (1) for some LVs, which are most essential for technological adjustment of the header.
Table 1. Matrices $K$ and $D$ of pair consistency and fuzziness indices.

| LV | Matrix $K$ | Matrix $D$ |
|----|------------|------------|
|    |            |            |
| «Stand of grain height» | 1 | 0.945 | 0.901 | 0.812 | 0 | 0.037 | 0.071 | 0.137 |
|    | 0.945 | 1 | 0.951 | 0.862 | 0.037 | 0 | 0.034 | 0.1 |
|    | 0.901 | 0.951 | 1 | 0.911 | 0.071 | 0.034 | 0 | 0.066 |
|    | 0.812 | 0.862 | 0.911 | 1 | 0.137 | 0.1 | 0.066 | 0 |
| «Stand of grain lodging» | 1 | 0.868 | 0.856 | 0.929 | 0 | 0.109 | 0.114 | 0.056 |
|    | 0.868 | 1 | 0.912 | 0.839 | 0.109 | 0 | 0.067 | 0.124 |
|    | 0.856 | 0.912 | 1 | 0.927 | 0.114 | 0.067 | 0 | 0.057 |
|    | 0.929 | 0.839 | 0.927 | 1 | 0.056 | 0.124 | 0.057 | 0 |
| «Stand of grain density» | 1 | 0.958 | 0.896 | 0.872 | 0 | 0.3 | 0.66 | 0.084 |
|    | 0.958 | 1 | 0.857 | 0.832 | 0.3 | 0 | 0.096 | 0.114 |
|    | 0.896 | 0.857 | 1 | 0.867 | 0.66 | 0.096 | 0 | 0.085 |
|    | 0.872 | 0.832 | 0.867 | 1 | 0.084 | 0.114 | 0.085 | 0 |
| «Reel vertical position» | 1 | 0.868 | 0.868 | 0.874 | 0 | 0.026 | 0.051 | 0.052 |
|    | 0.868 | 1 | 1 | 0.747 | 0.026 | 0 | 0.025 | 0.026 |
|    | 0.868 | 1 | 1 | 0.747 | 0.051 | 0.025 | 0 | 0.023 |
|    | 0.874 | 0.747 | 0.747 | 1 | 0.052 | 0.026 | 0.023 | 0 |
| «Reel horizontal position (reel offset)» | 1 | 0.86 | 0.86 | 0.892 | 0 | 0.056 | 0.056 | 0.056 |
|    | 0.86 | 1 | 0.96 | 0.81 | 0.056 | 0 | 0 | 0 |
|    | 0.86 | 0.96 | 1 | 0.92 | 0.056 | 0 | 0 | 0 |
|    | 0.892 | 0.81 | 0.92 | 1 | 0.056 | 0 | 0 | 0 |

Table 2. Indices $k_a$ and $k_m$ of overall consistency.

| LV | $k_a$ | $k_m$ |
|----|------|------|
| «Stand of grain height» | 0.812 | 0.812 |
| «Stand of grain lodging» | 0.82 | 0.812 |
| «Stand of grain density» | 0.787 | 0.779 |
| «Reel vertical position» | 0.908 | 0.907 |
| «Reel horizontal position (reel offset)» | 0.902 | 0.895 |

At step 11 of the algorithm (Fig. 1) of expert data quality estimation, the MF parameters averaging and the generalized MF development are performed. The average values of the parameters are determined as weighted average according to the formula:

$$p = \frac{1}{n} \sum_{i=1}^{n} \omega_i p_i,$$

where $p$ is an average value of the parameter $p$, $\omega_i$ are weighting factors, $p_i$ are the parameter $p$ values, estimated by the experts, $n$ is a number of experts. Weighting factors $\omega_i$ will be calculated as Fishburn numbers according to the chosen system of expert ranking in accordance with their qualification. The 2-nd and 3-rd experts possess the highest qualification, then the 1-st expert goes, and finally he is followed by the 4-th one. For the mixed system of preferences, when alongside with preferences, the system includes indifference ratios, Fishburn weighting factors have the form:

$$\omega_i = \frac{a_i}{b}, \quad a_{i-1} = \begin{cases} a_i, & \text{if } r_{i-1} \approx r_i \\ a_{i+1}, & \text{if } r_{i-1} > r_i \end{cases}, \quad r_N = 1, \quad i = N, ..., 2, \quad b = \sum_{i=1}^{N} a_i.$$
According to the formulae (4) we obtain weighting factors \( \omega_1 = 2/9, \omega_2 = \omega_3 = 1/3, \omega_4 = 1/9 \), applying them we find average values of the MF parameters. Fig. 2 presents the obtained MF graphs of several LVs.

**Fig. 2.** Membership functions of LV terms a) «Stand of grain height»; b) «Stand of grain lodging»; c) «Stand of grain density»; d) «Reel vertical position».
4 Conclusions

The base of expert knowledge, including membership functions of linguistic variables of external factors and the adjustable parameters of the harvester header, has been created. The knowledge base is intended for IIS of the harvester technological adjustment in the field. Four experts were involved for the purpose of preparing the data. Application of the algorithm of the expert data quality estimation has made it possible to take into account the hierarchy of expert knowledge according to their qualification and define the sufficient degree of the data adequacy to real conditions of the harvester operation, which is one of important conditions of efficiency and gains in performance in decision making with the help of IIS.

References

1. L. Jain, H. Kumar, R.K. Singla, Special issues on IP Multimedia Communications 1, 55-60 (2011) https://research.ijcaonline.org/ipmc/number1/ipmc013.pdf
2. M. Sujaritha, S. Annadura, J. Satheeshkuma, S. Kowshik Sharan, L. Mahesh, Computers and Electronics in Agriculture 134, 160–171 (2017) doi:10.1016/j.compag.2017.01.008
3. H. Navarro-Hellin, J. Martinez-del-Rincon, R. Domingo-Miguel, F. Soto-Valles, R. Torres-Sánchez, Computers and Electronics in Agriculture 124, 121–131 (2016) doi:10.1016/j.compag.2016.04.003
4. H. Zareiforoush, S. Minaei, M. Reza Alizadeh, A. Banakar, B. Hosseinzadeh Samani, Computers and Electronics in Agriculture 124, 14–22 (2016) doi:10.1016/j.compag.2016.01.024
5. L. Sulistiyo Budi, R. Mustika Wardhani, International Journal on Advanced Science Engineering Information technology 7(3), 1089-1097 (2017) doi:10.18517/ijaselt.7.3/2194
6. J.T. Walls, P. Caciagli, J. Frazier Tooker, J.M. Russo, E. George Rajotte, C. Rosa, Computers and Electronics in Agriculture 127, 775–786 (2016) doi: 10.1016/j.compag.2016.08.005
7. R.F, Chevalier, G. Hoogenboom, R.W. McClendon, J.O. Paz, Environmental Modelling & Software 35, 84–91 (2012) doi : 10.1016/j.envsoft.2012.02.010
8. W. Goodridge, M. Bernard, R. Jordan, R. Rampersad, Computers and Electronics in Agriculture 133, 80–87 (2017) doi : 10.1016/j.compag.2016.12.003
9. M. Omid, M. Lashgar, H. Mobli, R. Alimardani, S. Saeid Mohtasebi, R. Hesamifard, Expert Systems with Applications 37, 7080-7085 (2010) doi:10.1016/j.eswa.2010.03.010
10. G. Craessaerts, J. De Baerdemaeker, B.M.A. Missotten, W. Saeyes, Biosystems Engineering 106, 103–111 (2010) doi:10.1016/jbiosystemseng.2009.12.012
11. E. Ebrahimi, A. Zarei, N. Astan, Agricultural Communications 3(3), 26 -32 (2015)
12. P. Rybacki, Z. Grześ, Journal of Research and Applications in Agricultural Engineering 63(1), 80-91 (2018)
13. L. Borisova, V. Dimitrov, I. Nurutdinova, Advances in Intelligent Systems and Computing 680, 95-105 (2018) doi:10.1007/978-3-319-68324-9_11
14. V. Dimitrov, L. Borisova, I. Nurutdinova, Advances in Intelligent Systems and Computing 875, 236-245 (2019) doi:10.1007/978-3-030-01821-4_25
15. L.V. Borisova, I.N. Nurutdinova, V.P. Dimitrov, Don State Technical University Bulletin 15(2), 100-107 (2015) doi:10.12737/11611

16. V. Dimitrov, L. Borisova, I. Nurutdinova, MATEC Web of conference 132, 04009 (2017) doi: 10.1051/matecconf/201713204009

17. L. Borisova, V. Dimitrov, I. Nurutdinova, Proceedings of IEEE East-West Design and Test Symposium (EWDTS’2017), 319-322 (2017) doi: 10.1109/EWDTS.2017.8110107