Fuzzy linguistic forecasting of the results of the implementation of regional socio-economic projects

Alexander Lomazov[1*], *Vadim* Lomazov[2,3], *Olga* Akupian[3], *Alina* Dobrunova[3], and *David* Petrosov[1]

[1]Financial University under the Government of the Russian Federation, 125993 Moscow, Russia
[2]Belgorod State National Research University, 308015 Belgorod, Russia
[3] Belgorod State Agricultural University named after V. Gorin, 308503 Belgorod, Russia

Abstract. The article is devoted to the problem of improving the intellectual digital support for managing the implementation of regional socio-economic projects, which are an effective mechanism for the development of regions. Given the high level of uncertainty in the description of the subject area, it is proposed to use expert technologies to assess the degree of achievement of the intermediate and target indicators of the project. L. Zadeh's linguistic modeling apparatus was used as the basis for intellectual tools for representing and processing expert judgments. A linguistic indicative-temporal-territorial model of the project has been built in the form of a set of linguistic variables and a system of fuzzy production rules linking these variables. At the same time, temporal and territorial linguistic variables reflect the time and place of measurement of project indicators, the incomplete correspondence of which determines the degree of reliability of the measured values of indicators. The use of the constructed system of indicative-temporal-territorial fuzzy production rules makes it possible to link the intermediate values of the indicators with the target values of the project indicators and thereby predict the possible outcomes of its implementation. For this, a fuzzy logic inference procedure based on the use of the Mamdani algorithm has been developed. Preliminary results of using a research prototype of an intelligent expert system for analyzing the intermediate results of the project stages and predicting its results may indicate the effectiveness of the proposed approach.

1 Introduction

One of the effective mechanisms for the development of regions is state support in the form of regional socio-economic (health, educational, environmental, etc.) private-state projects (in particular, in relation to projects of sustainable development of rural areas [1-3]). In modern conditions, which are characterized by a high level of instability and unpredictability of project implementation conditions, the ability to predict results when external conditions change [4], which is especially important for long-term multi-stage projects of an innovative

* Corresponding author : vlomazov@yandex.ru

nature, is of greater importance [5]. Forecasting the final results allows you to make timely adjustments to the project activities, implementing adaptive management [6,7].

The purpose of the article is to develop a theoretical apparatus for intelligent support for predicting the final results of a multi-stage socio-economic project based on fuzzy linguistic modeling of the subject area of the project and the external conditions for its implementation. The use of models and methods of fuzzy linguistic analysis by L. Zadeh [8] (unlike traditional forecasting methods [9-11]) makes it possible to take into account the significant blurring of concepts and patterns of the subject area, attracting experts to evaluate project performance indicators and external factors. environment.

2 Materials and methods

To describe socio-economic projects, we will use the approach of L. Zade, in which project indicators are presented in the form of linguistic variables, which are an extension of fuzzy numbers and are constructions of the form [12]:

$$< Ind, DInd, TBInd, SintRInd, SemRInd> \tag{1}$$

where *Ind* is the name of the indicative linguistic variable; *DInd* is the range of numerical values of the indicator under consideration (taking into account the normalization, we can assume that *DInd* =[0, 1]); *TBInd* – basic set of indicative terms (we will assume *TBInd* = {low, medium, high}); *SintRInd* is a syntactic rule that allows you to generate the names of indicative terms from the names of the elements *TBInd* and thereby expand the set of the base set of terms *TBInd* to the set of terms *TInd*; *SemRInd* is a semantic rule establishing correspondence between temporal terms and fuzzy subsets of *DInd*. The semantics of terms is set by experts based on their knowledge of the subject area of the project. In the future, it is assumed that the membership functions of fuzzy subsets are trapezoidal, as shown in Fig. 1.

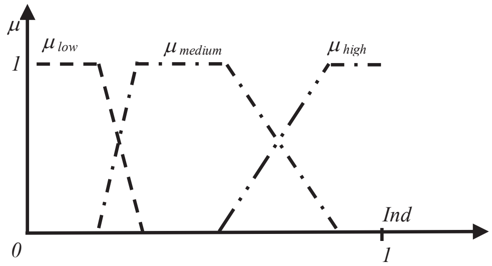

Fig. 1. Graphs of membership functions μ_{low}, μ_{medium}, μ_{high} , defining the semantics of the terms *low*, *medium*, *high*.

Formal linguistic representation (1) can be used to describe target indicators TargInd, intermediate indicators InterInd, as well as external factors ExInd, reflecting the conditions for the implementation of projects.

Based on the methodological apparatus of knowledge representation [13], the dependence of the final values of the target indicators of the project *TargIndj(N); j=1,2,…,J* at the end of the final stage numbered *n=N*, from the values of intermediate indicators (InterIndi(n); *i=1,2,…,I; n=n*,n*+1,…,n***) at the end of several (starting from the stage numbered n* and up to the stage numbered *n***) already past stages of the project, we will represent them in the form of Rule-Based Fuzzy production rules:

$$IF\ (Flow(InterInd_i(n);\ i=1,2,\ldots,I;\ n=n^*,n^*+1,\ldots,n^{**}))\quad (2)$$
$$THEN\ (TargInd_j(N)-low)$$
$$IF\ (Fmedium(InterInd_i(n);\ i=1,2,\ldots,I;\ n=n^*,n^*+1,\ldots,n^{**}))$$
$$THEN\ (TargInd_j(N)-medium)$$
$$IF\ (Fhigh(InterInd_i(n);\ i=1,2,\ldots,I;\ n=n^*,n^*+1,\ldots,n^{**}))$$
$$THEN\ (TargInd_j(N)-high)$$

where *Flow, Fmedium, Fhigh* included in the left parts of the productions are fuzzy logical formulas regarding the linguistic values of intermediate indicators, built using the operations of negation, conjunction and disjunction, defined in the usual way (Zadeh operations) [14].

Finding the predicted values of target indicators can be performed using the fuzzy inference procedure [15], for example, using the Mamdani algorithm, which has a high level of interpretability [16]. However, model (1), (2) (which is a standard application of the linguistic modeling apparatus to solving the problem of forecasting) does not reflect all the specifics of socio-economic projects and it is advisable to expand it, taking into account the fuzzy territorial and temporal binding of indicator values characteristic of such projects. An extended linguistic model description of regional socio-economic projects in combination with a constructed system of fuzzy production rules and a fuzzy inference procedure is the main result of this work.

3 Results and discussion

3.1 Temporal and territorial model description of projects

Let's supplement the system of indicative linguistic variables (1) with the temporal linguistic variable Stage, whose values are the names of the project stages. The formal linguistic description of the Stage temporal variable in accordance with [17] has the form of a tuple:

$$< Stage,\ D_{Stage},\ TB_{Stage},\ SintR_{Stage},\ SemR_{Stage}> \quad (3)$$

where *Stage* is the name of the temporal linguistic variable; *DStage* – range of numerical values of the considered indicator *DStage=[tstart, tfinish]; TBStage* – basic set of temporal terms *TBStage = {Stage1, Stage2 , ..., StageN}*; *SintRStage* is a syntactic rule that allows you to generate names of temporal terms from the names of *TBStage* elements (for example, the beginning of the Stagen or the union of Stagen and Stagem), forming a set *TStage*; SemRStage is a semantic rule that establishes a correspondence between temporal terms and fuzzy subintervals *DStage*. The fuzziness of the temporal linguistic variable Stage is due to the fact that the preparation of the work that makes up the content of the next stage, as a rule, begins earlier than the normative start of the stage, and the full completion of these works may occur later than the normative end of the stage, i.e. some moments of time can be attributed to the stage under consideration to some extent, corresponding to the membership function of the fuzzy subinterval, denoted by μStage(t), and taking values from the interval [0. 1]. In this case 1 means complete compliance, and 0 complete non-compliance of time t with stage Stagen (n=1,2,…,N).

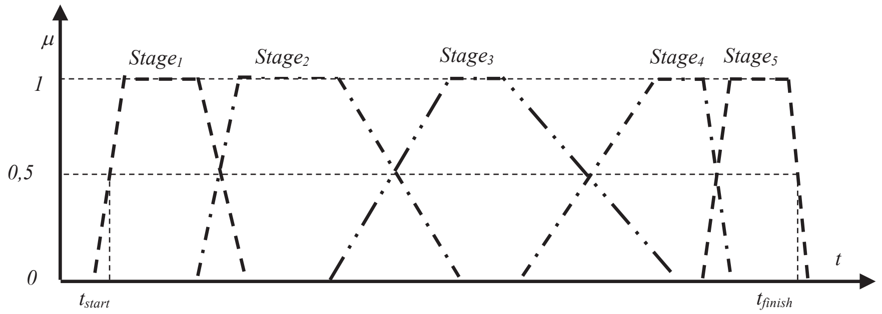

Fig. 2. Membership functions of elements of the base vector of temporal terms from TB_{Stage}.

Since the implementation of a regional socio-economic project (and, therefore, each planned or actual value of an intermediate and target indicator) is associated with a certain territory (a set of settlements in the region), it is advisable to supplement system (1) with territorial linguistic variables. We restrict ourselves to consideration of one territorial linguistic variable Ter, which corresponds to the selection of small, medium and large settlements in the region and is formally defined by a tuple

$$< Ter, D_{Ter}, TB_{Ter}, SintR_{Ter}, SemR_{Ter}> \quad (4)$$

where Ter is the name of the territorial linguistic variable; DTer is the range of numerical values of the considered indicator DTer (for example, up to 1 million people); TBTer = {small settlements, medium settlements, large settlements} – basic set of temporal terms; SintRTer is a syntactic rule that allows you to generate the names of territorial terms from the names of TBTer elements (in particular, using strengthening / weakening modifiers (for example, very small settlements) or logical connectives (for example, small and medium settlements)), forming a set TTer; SemRTer is a semantic rule that establishes a correspondence between territorial terms and fuzzy subsets of DTer.

It should be noted that, depending on the used classification feature of settlements (according to the specifics of a particular socio-economic project), other territorial linguistic variables can be constructed, and thus the territorial component of the linguistic description of the project may contain several linguistic static variables. The considered linguistic variable is universal for all socio-economic projects.

3.2 Indicative-temporal-territorial production dependencies

Let's expand the system of fuzzy production rules (2) by indicating the time stage and territory that correspond to the value of the intermediate indicator. As a result, we get

$$IF\ F_{p,k}\ (InterInd_i - InterIndTerm_s, i=1,2,..I,\ s=1,2,\ldots,S) \quad (5)$$

$$WHEN\ Time - TimeTerm^*$$

$$WHERE\ Ter - TerTerm^*$$

$$WITH\ ExInd - ExIndTerm^*$$

$$THEN\ TargInd_p - TargIndTerm_k$$

$$p=1,2,..,P,\ k=1,2,\ldots,K$$

where *Fp,k(InterIndi – InterIndTerms, i=1,2,..I, s=1,2,…,S)* – fuzzy propositional formulas for fuzzy propositions InterIndi – InterIndTerms, meaning that the intermediate indicator InterIndi takes (with some degree of truth) a linguistic meaning (term) InterIndTerms; InterIndTerms∈TInd, i=1,2,..n, s=1,2,…m.

In relations (5), the values of the considered intermediate indicator InterIndi are assumed to be measured during TimeTermon the territory of TerTerm* under external conditions ExIndTerm*,, which corresponds to the fulfillment (with some degree of truth) of the statements: Time – TimeTerm*, Ter – TerTerm*, ExInd – ExIndTerm* (TimeTerm*∈TStage, TerTerm*∈TTer, ExIndTerm*∈TInd). If the left part of the fuzzy production rule (5) is satisfied (with some degree of truth), then its right part is satisfied (with some degree of truth), which is a fuzzy statement TargIndp – TargIndTermk regarding the target indicator TargIndp, which consists in the fact that this indicator accepts a linguistic value (term) TargIndTermk.

Relations (5) (in contrast to relations (2) having the structure IF-TIEN-) having the structure IF-WHEN-WHERE-WHITTH-TIEN- will be called W-rules. We will assume that W-rules (as well as fuzzy rules (2)) are set by experts based on their knowledge of the subject area of the project.

3.3 W-rule transformation

The right parts of fuzzy production rules (5) contain, along with the values of intermediate indicators, also the conditions for their measurement (time, territory and external conditions), which may not fully comply with the regulatory conditions. As part of the task of constructing a knowledge model about the subject area (project implementation) in the form of a system of fuzzy production rules, since the circumstance of time, the circumstance of place and an additional circumstance determine the conditions for measuring the main circumstance (intermediate indicator), it is advisable to introduce a linguistic variable *Rel* (Reliability), reflecting the degree of compliance with these conditions and formally represented as:

$$< Rel, DRel, ITBRel, SintRRel, SemRRel> \quad (6)$$

where *Rel* is the name of the linguistic variable; *DRel* – range of numeric values, *DRel* =[0, 1]); *TBRel* = {low, medium, high}.– basic set of indicative terms; *SintRRel* – syntactic rule that allows you to generate term names from the names of *TBRel* elements; *SemRRel* – is a semantic rule that establishes a correspondence between temporal terms and fuzzy subsets of *DRel*. In what follows, it is assumed that the membership functions of fuzzy subsets (as in the case of indicative variables) are trapezoidal.

Let us first estimate the actual deviations of the measurement conditions from the standard conditions. Let's introduce the linguistic variables *DevTime, DevTer, DevExInd*, reflecting the deviations (deviations) Time, Ter, ExInd, from the normative values in the terms {low, medium, high}. Then the system of fuzzy production rules relating the deviations of the measurement conditions with the reliability of the result will take the form (7)-(9):

$$IF\ Flow\ (DevTime – low,\ DevTime – medium,\ DevTime – high,\ DevTer – low,\ DevTer – medium,\ DevTer – high,\ DevExInd – low,\ DevExInd – medium,\ Dev\ ExInd – high)\ \text{THEN}\ (Rel – low) \quad (7)$$

$$\text{IF}\ F_{medium}\ (DevTime – low,\ DevTime – medium,\ DevTime – high,\ DevTer – low,\ DevTer – medium,\ DevTer – high,\ DevExInd – low,\ DevExInd – medium,\ Dev\ ExInd – high)\ \text{THEN}\ (Rel – medium) \quad (8)$$

$$\text{IF } F_{high}\ (DevTime - low,\ DevTime - medium,\ DevTime - high,\ DevTer - low,\ DevTer - medium,\ DevTer - high,\ DevExInd - low,\ DevExInd - medium,\ DevExInd - high) \text{ THEN } (Rel - high) \quad (9)$$

where F_{low}, F_{medium}, F_{high} – fuzzy propositional formulas corresponding to the levels *low, medium, high* of value *Rel*.

Integral accounting for the possible deviation of the measurement conditions of the InterInd indicator *InterInd* from the standard values *NormTime*, *NormTer*, *NormExInd* in the form of a reliability indicator *Rel* allows you to go from *InterInd(Time*, *Ter*, *ExInd)* to a Z-linguistic variable (Z-number [18] with linguistic values) represented as:

$$< InterInd(NormTime, NormTer, NormExInd), Rel > \quad (10)$$

representing a pair of linguistic variables, including the linguistic variable *InterInd(NormTime*, *NormTer*, *NormExInd)*, corresponding to the indicator measured under standard conditions, and the linguistic variable *Rel*, reflecting the degree of reliability of the value of the first linguistic variable.

Thus, the system of fuzzy W-rules (5) can be reduced to fuzzy Z-rules of the form

$$IF\ F_{p,k}\ (InterInd_i - InterIndTerm_s, i=1,2,..I,\ s=1,2,\dots,S),\ Rel - RelTerm^* \quad (11)$$
$$\text{THEN } TargInd_p - TargIndTerm_k$$
$$p=1,2,..,P,\ k=1,2,\dots,K$$

To further transform the system of W-rules (5), we use the approach proposed in [19] based on the concept of a weighted Z-number obtained by multiplying the membership function μ(x) of the fuzzy set corresponding to the first component of this number by the value λ (0≤λ≤1) in accordance with formula (12):

$$Z\lambda=\{(x,\ \mu\lambda(x)) \setminus \mu\lambda(x) = \lambda\mu(x)\ \} \quad (12)$$

In this case, the value of λ is defined as the result of defuzzification of the second component of the Z-number (reflecting a measure of confidence in the value of the first component of the Z-number), i.e. is an ordinary crisp number.

The transition from a Z-number to a weighted Z-number can be seen as a partial defuzzification accompanied by some loss of information. However, the use of a weighted Z-number is inconvenient, since, generally speaking, the normalization condition is not satisfied for the corresponding membership function, i.e. max { μλ(x)}= λ≠ 1.

Therefore, in accordance with [20], a normalized fuzzy number Zλ* is constructed, which has the same fuzzy expectation as the weighted Z-number Zλ. By definition, the fuzzy expectation of a fuzzy number has the form (13)

$$E_A(x) = \int_X x\mu_A(x)dx \quad (13)$$

Based on this, the fuzzy number Zλ* is characterized by the membership function μλ*(x), determined by formula (14):

$$\mu^{\lambda^*}(x)= \mu\ (x/(\lambda)^{0.5}) \quad (14)$$

which corresponds to the change of variables in the original membership function of the first Z-number component.

Since Z-numbers correspond to the terms of linguistic variables, the Z-rules (11) constructed in the previous section can be reduced to a system of ordinary fuzzy rules regarding modified (taking into account the confidence measure λ) terms $InterIndTerm_s^{\lambda *}$, that have the form (15):

$$IF\ F_{p,k}\ (InterInd_i - InterIndTerm_s^{\lambda *}, i=1,2,..I,\ s=1,2,\dots,S) \quad (15)$$
$$THEN\ TargInd_p - TargIndTerm_k$$
$$p=1,2,..,P, \quad k=1,2,\dots,K$$

4 Procedure for determining the predicted values of project targets

The general procedure for fuzzy inference of predicted values of target indicators of a regional program for the development of rural areas contains the following steps:

1. Determination (measurement) of the numerical values of the intermediate indicators of the program and the values of the characteristics of the measurement conditions (time and territory on which the measurements were taken, as well as (in a more general case) acting external factors).
2. Fuzzification of the measured numerical values, within which the degree of correspondence of these values to the basic terms of the corresponding fuzzy linguistic variables is established
3 Fuzzy inference based on auxiliary fuzzy production rules (7) - (9), aggregation of subconditions, and then activation of conclusions, as a result of which numerical values of the degree of confidence are determined for the values of the measured intermediate indicators.
4. Recalculation of the parameters of the semantics of indicative terms (corresponding to intermediate indicators) taking into account the measure of confidence in accordance with formula (14).
5. Aggregation of subconclusions of the main fuzzy production rules that connect intermediate indicators (left parts of the rules) with target indicators (right parts of the rules).
6.Activation of subconclusions regarding target indicators.
7. Accumulation of sub-conclusions regarding target indicators.
8. Defuzzifications for target indicators, i.e. determination of their numerical values.

5 Conclusion

Adaptive management of a regional socio-economic project (allowing you to respond flexibly to possible changes in external factors, taking into account the results obtained earlier) requires tools for predicting its results. The approach proposed in the article is based on the use of the apparatus of fuzzy linguistic modeling, as the one that best corresponds to the high level of uncertainty of the concepts of the subject area and allows the use of expert technologies.

The main result of the research is the construction of a procedure for determining the predicted values of the target indicators of a regional socio-economic project, which is based on:

- a system of fuzzy indicative, temporal and territorial linguistic variables describing the progress of the project ((1),(3),(4));

- a system of fuzzy production W-rules (5) linking the values of intermediate indicators, as well as circumstances of time, place and additional circumstances of measuring these indicators with predicted values of target indicators;
- a set of procedures that implement a sequential transition from W-rules to Z-rules, and then to ordinary fuzzy rules;
- fuzzy inference procedure (Mamdani's algorithm).

Preliminary results of using a research prototype of an intelligent expert system for predicting project results may indicate the effectiveness of the proposed approach. One of the directions for further development and use of the approach can be its application in solving problems of synthesis of multi-scenario projects (in combination with evolutionary methods for selecting parameters [21] and neural network tuning [22]).

References

1. V.L. Anichin, A.V. Turyanskiy, A.I. Dobrunova, International Journal of Economic Research **14(7),** 203-210 (2017)
2. V. Anichin, et.al, AIP Conference Proceedings **2661,** 020020 (2022)
3. V.A. Lomazov, et.al, Proc. SPIE **12637,** 126370P (2023)
4. A. Shtub, M. Rosenwein, Project management: processes, methodologies, and economics (Pearson, Hoboken, NJ, USA, 2017)
5. V.A. Lomazov, et.al, IOP Conference Series: Earth and Environmental Science, **839(2),** 022068 (2021)
6. R. Pellerin, N. Perrier, Int. J. Prod. Res. **57,** 2160–2178 (2019)
7. J.R. San Cristóbal Mateo, Management science, operations research and project management: modelling, evaluation, scheduling, monitoring (Routledge, Abingdon 2019)
8. L.A. Zadeh, Computational Statistics & Data Analysis **51(1),** 15-46 (2006)
9. J. Scott Armstrong, Principles of forecasting: a handbook for researchers and practitioners (Norwell, Massachusetts: Kluwer Academic Publishers, 2001)
10. O.S. Glinskaya, et.al, Lecture Notes in Networks and Systems **57** (2019)
11. V.Y. Trofimets, et.al, Journal of Advanced Research in Dynamical and Control Systems **11(8 Special Issue),** 935-947 (2019)
12. H.H. Lieb, Linguistic Variables: Towards a unified theory of linguistic variation. Amsterdam (Philadelphia, John Benjamins Publishing, 1993)
13. S.J. Russell, P. Norvig, Artificial intelligence artificial intelligence: a modern approach (New Jersey, Prentice Hall, 2020)
14. A. Piegat, Fuzzy modeling and control (Heidelberg, Physica-Verlag, 2001)
15. E.S. Khorasani, et.al, Journal of Ambient Intelligence and Humanized Computing **4(4),** 409-410 (2013)
16. E.H. Mamdani, Proceedings of the Institution of Electrical Engineers **121(12),** 1585–1588 (1974)
17. A.V. Lomazov, et.al, Journal of Physics: Conference Series **1661,** 012058 (2020)
18. L.A. Zadeh, Information Sciences **181(14),** 2923–2932 (2011)
19. B. Kang, et.al, Journal of Information & Computational Science **8(7),** 2807-2814 (2012)
20. B. Kang, et.al, Journal of Information &Computational Science **9(3),** 703-709 (2012)

21. D.A. Petrosov, et.al, Journal of Advanced Research in Dynamical and Control Systems **10(10 Special Issue),** 1840-1846 (2018)
22. D.A. Petrosov, et.al, Applied Sciences **11(9),** 3899 (2021)