Fuzzy multiple regression technical and economic model of airport terminal passenger handling system

E D Guzha, T V Khvostova, V A Romanenko and M A Skorokhod*

Department of transport organization and management, Samara National Research University, Samara, 443086, Russia

*E-mail: skoro-margarita@yandex.ru

Abstract. A two-stage technique for the formation of a multiple regression model with a fuzzy dependent variable and parameters and clear independent variables is proposed. The results of the construction of a fuzzy technical and economic model of the baggage handling system, which is the technological "core" of the passenger service system at the airport terminal, are considered. The model is supposed to be used at the early stages of design and optimization of airport ground handling systems with incomplete initial information.

1. Introduction

The complex of technological and technical means of the airport passenger terminal is a complex system that is being formed and operating under the influence of various kinds of uncertainties, some of which cannot be described in terms of probability theory. Such uncertainties are caused by a lack of knowledge about the system, are characterized by the impossibility of accumulating repeated samples, and require the use of expert estimates. To describe this kind of uncertainty, the methods of the theory of fuzzy sets are used. The need to take into account uncertainties of a fuzzy nature when solving very urgent problems of designing and optimizing a passenger service system at an airport terminal requires a complex of mathematical models that allow operating with fuzzy values. The main component, the technological “core” of the airport passenger service system, determining its effectiveness, is their baggage handling system (BHS). Below we are considering a fuzzy model for estimating the costs of the acquisition, installation and commissioning of BHS depending on its technical parameters. It is supposed to use the model at the stages of designing or optimizing the passenger and baggage service system at the airport, when the composition, structure, parameters, technical and technological features of the BHS cannot be determined yet, and the specific types of devices and subsystems included in it are not selected. In this case, a significant part of the initial information for estimating BHS cost will be asked expertly, and, therefore, the introduction of fuzziness in the parameters of the technical and economic model will make it possible to take into account the uncertainty of the initial data. The technical and economic BHS model is planned to be integrated into the complex of models of various degrees of complexity and accuracy, intended for use at various stages of design and optimization of airport passenger and baggage service systems with a step-by-step refinement of their characteristics. Within the framework of this approach, it seems justified to develop a family of technical and economic BHS models, including a “normal” non-fuzzy model used for preliminary estimates at the initial stages of design, and a fuzzy one which is more informative and, accordingly, more complex,
i.e. a model for obtaining more accurate results at subsequent stages. A fuzzy model is formed on the basis of a clear one, which ensures a certain continuity of results.

2. The technique of constructing a fuzzy regression model

Fuzzy regression analysis methods developed since the first work in this area appeared in the early 80s of the 20th century are divided into two groups. The first group [1-3] is based on the method of mathematical programming, the second one [4-6] is based on the method of least squares. Below we will consider a technique that combines both approaches and involves the formation of a fuzzy regression model based on a pre-built clear statistical linear regression model that establishes a relationship:

\[
Y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_n x_{in} + \epsilon_i, \quad i = 1,2,\ldots,m,
\]

where \(Y_i\) is dependent variable (response), \(x_{i1},x_{i2},\ldots,x_{in},\ldots,x_{in}\) are independent variables (factors), \(\beta_0,\beta_1,\ldots,\beta_n\) are coefficients (parameters) of the model, \(\epsilon_i\) is random approximation error. All values in (1) are not fuzzy.

The task of the regression analysis is to determine unbiased, consistent and effective estimates \(b_0,b_1,\ldots,b_n\) of unknown coefficients \(\beta_0,\beta_1,\ldots,\beta_n\) from a given set \(\{(x_{1i},x_{2i},\ldots,x_{ni},Y_i)\}, \ i = 1,2,\ldots,m\). The solution to this problem, obtained using the least squares method, in matrix-vector form is written as [7]

\[
b = (X'X)^{-1}X'Y,
\]

where \(b = (b_0,b_1,\ldots,b_n)'\) is model coefficient vector, \(Y = (Y_1,Y_2,\ldots,Y_m)'\) is vector of dependent variables, \(X\) is matrix of independent variables:

\[
X = \begin{bmatrix}
1 & x_{11} & \cdots & x_{1n} \\
1 & x_{21} & \cdots & x_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
1 & x_{ni1} & \cdots & x_{nin}
\end{bmatrix}.
\]

Using the obtained estimates \(b_0,b_1,\ldots,b_n\), a consistent, unbiased, and effective estimate \(\hat{y}_i\) of the mathematical expectation \(M[Y_i]\) of the dependent variable \(Y_i\) is determined according to the linear regression dependence

\[
\hat{y}_i = b_0 + b_1 x_{i1} + \ldots + b_n x_{in}, \quad i = 1,2,\ldots,m.
\]

If the relationship between the values of the dependent and independent variables is not well defined, then the linear regression dependence should be converted to a fuzzy form:

\[
\tilde{y}_i = \tilde{b}_0 + \tilde{b}_1 x_{i1} + \ldots + \tilde{b}_n x_{in}, \quad i = 1,2,\ldots,m.
\]  

(2)

Here and below, fuzzy values are indicated by the tilde sign (\(~\) ). Let the parameters of model (2) \(\tilde{b}_0,\tilde{b}_1,\ldots,\tilde{b}_n\) be triangular fuzzy numbers (TNFs), which are denoted by triples of parameters \(\tilde{b}_j = (b_j^l, b_j^m, b_j^r)\), where the indices “L” and “R” are assigned, respectively, to the left and right boundaries of the medium, and “\(M\)” is the modal value of the corresponding TNF \(\tilde{b}_1\), \(j = 0,1,\ldots,n\). In this case, the dependent variable is also described by the VLF, inside the carrier of which the initial value \(Y_i\) should be located. In a formalized form, this condition is defined by two inequalities
\[ Y_i \geq y_i^t = b_i^l + b_i^r x_{i1} + \ldots + b_i^r x_{im}, \quad Y_i \leq y_i^r = b_i^l + b_i^r x_{i1} + \ldots + b_i^r x_{im}, \quad i = 1,2,\ldots,m. \] (3)

Let us note that a fuzzy number can be converted to a clear quantitative value using one of the defuzzification methods. For example, the defuzzification of TNF \( \tilde{B}_j = (b_j^l, b_j^m, b_j^r) \) by the center of gravity method is performed according to the formula

\[ b_j^o = \frac{b_j^l + b_j^m + b_j^r}{3}. \] (4)

To ensure the continuity of the results in the transition from a clear model to a fuzzy, considering the second development of the first, we require that the values \( \bar{y}_i \) and \( \tilde{y}_j \) defuzzified by the center of gravity method and coincide with clear ones \( y_i \) and \( b_j \), accordingly:

\[ y_i^m = y_i, \quad i = 1,2,\ldots,m. \] (5)

\[ b_j^o = b_j, \quad j = 0,1,\ldots,n. \] (6)

Condition (5), taking into account (4), allows us to determine the value \( y_i^m \), if \( y_i^t \), \( y_i^r \) are known:

\[ y_i^m = 3y_i - y_i^t - y_i^r, \quad i = 1,2,\ldots,m. \]

Similarly, from condition (6) with known \( b_j^o \), \( b_j^r \), \( j = 1,2,\ldots,n \) the value \( b_j^m \) is found:

\[ b_j^m = 3b_j - b_j^o - b_j^r, \quad j = 0,1,\ldots,n. \] (7)

In this case, for the coordinates of the reference points of the membership functions of the TNP \( \tilde{B}_b, \tilde{B}_1,\ldots,\tilde{B}_n \), the condition following from the definition of the TNP must be satisfied:

\[ b_j^o \leq b_j^m \leq b_j^r, \quad j = 0,1,\ldots,n. \] (8)

The procedure for the formation of a regression model (2) with clear input and fuzzy output data, consisting in determining the coefficients \( b_j^l, b_j^r, \), \( b_j^m \), \( b_j^m, b_j^m, b_j^m, b_j^m, b_j^m, b_j^m, b_j^m, b_j^m, b_j^m, b_j^m \), will be divided into two stages. At the first stage, a clear regression model is constructed using the least squares method, that is, coefficients \( b_0, b_1,\ldots,b_n \) are determined that coincide, according to (6), with \( b_0^m, b_1^m,\ldots,b_n^m \) respectively. The found coefficients are used in the second stage, which consists in solving the linear programming problem in order to determine those \( b_0^l, b_1^l,\ldots,b_n^l \) and \( b_0^r, b_1^r,\ldots,b_n^r \) that, satisfying the constraints (3), (8), minimize the width of the fuzzy corridor including the actual values of the dependent variable by the sum of all observations:

\[ \sum_{i=1}^{m} (y_i^r - y_i^t) \rightarrow \min. \]

The value \( b_j^m \) required at checking condition (8) at each step of solving the linear programming problem is determined by formula (7) for those accepted \( b_j^o, b_j^r \) and found at the first stage \( b_j \), \( j = 0,1,\ldots,n \).

3. The object of modeling

Characteristics of BHS vary widely, which is confirmed by the data of manufacturers of a number of systems given in table 1. Costs \( (Y) \) for BHS are presented in US dollars, having a purchasing power
corresponding to the end of 2017. Structural and functional analysis of BHS, carried out in order to select the most significant factors of the technical and economic model, allows us to distinguish the following main subsystems: baggage starting passengers (S₀), entering baggage of transfer passengers (S₉) into the BHS, multi-level screening (S₀), sorting (S₅), picking baggage for departing passengers on flights (S₆), storing baggage checked in advance (S₆), baggage claim (S₆), control (S₀).

The main elements of the subsystems S₀, S₉ are belt conveyors designed to transport luggage. The length of the conveyor lines of the subsystem S₀ and the complexity of their configuration depends on the number of places of registration served by these conveyor lines. The main type of equipment of the subsystem S₀ is inspection devices, the number and productivity of which are one of the most important factors determining the throughput of BHS.

### Table 1. Main BHS characteristics.

| i  | Technology | x₁ | x₂ | x₃ | x₄ | x₅ | x₆ | Y, mln.USD |
|----|------------|----|----|----|----|----|----|------------|
| 1  | M          | 2  | 1  | 1  | 0  | 0  | 0  | 1.0        |
| 2  | C, M       | 20 | 1  | 0  | 0  | 0  | 0  | 1.9        |
| 3  | C, M       | 14 | 2  | 5  | 0  | 0  | 0  | 3.0        |
| 4  | C          | 30 | 4  | 3  | 0  | 0  | 0  | 14.0       |
| 5  | C          | 48 | 3  | 5  | 0  | 0  | 1  | 3.3        |
| 6  | C, M       | 10 | 4  | 5  | 2  | 0  | 0  | 9.0        |
| 7  | C          | 30 | 9  | 0  | 0  | 0  | 0  | 9.3        |
| 8  | C          | 30 | 6  | 3  | 0  | 0  | 0  | 10.0       |
| 9  | C          | 36 | 6  | 4  | 1  | 0  | 2  | 27.8       |
| 10 | C, M       | 96 | 10 | 8  | 6  | 0  | 0  | 15.6       |

The most important factor affecting the performance and cost of BHS is the baggage sorting technology implemented by the system. The most important quantitative factors determining the technical appearance and cost of individual subsystems and BHS in general should include the number of check-in desks (x₁), baggage carousels (x₂), inspection vehicles (x₃), baggage packaging carousels (x₄), pre-checked baggage storage capacity (x₅), number of transfer baggage handling lines (x₆).

### 4. The results of a feasibility model formation

For the formation of a fuzzy technical and economic model, the data of table 1 were used. The BHS is divided into two groups of equal size, the first of which (“group I”) includes the BHS i = 1,...,10 with a conveyor belt, including “manual”, baggage sorting scheme, and the second (“Group II”) includes BHS i=11,...,20, equipped with lines of deflecting pallets or autonomous luggage trolleys. Since,
according to the data in Table 1, there are no subsystems $S_e$ in the BHS of group I, and subsystems $S_T$ are used to a limited extent, then in order to simplify the regression model of the BHS of this group, the variables $x_5$ and $x_6$ are excluded from its factors. Table 2 presents the results of the formation of regression models for the BHS of both groups, including clear estimates of the coefficients $b_j$ and three parameters $(b_j^T, b_j^H, b_j^E)$ of fuzzy estimates of the coefficients $\tilde{b}_j$, moreover $j = 0, ..., 6$ - in the BHS model of group I and $j = 0, ..., 6$ - in the BHS model of group II.

### Table 2. Estimates of the model coefficients, mln. USD.

| $j$ | BHS group I | BHS group II |
|-----|-------------|-------------|
|     | $b_j$ | $b_j^T$ | $b_j^H$ | $b_j^E$ | $b_j$ | $b_j^T$ | $b_j^H$ | $b_j^E$ |
| 0   | 0.050 | 0.010 | 0.065 | 0.075 | 0.500 | 0.100 | 0.500 | 0.900 |
| 1   | 0.040 | 0.010 | 0.050 | 0.060 | 0.040 | 0.020 | 0.045 | 0.055 |
| 2   | 0.300 | 0.290 | 0.300 | 0.310 | 0.370 | 0.227 | 0.301 | 0.581 |
| 3   | 0.770 | 0.428 | 0.871 | 1.011 | 1.001 | 0.914 | 1.038 | 1.048 |
| 4   | 0.318 | 0.218 | 0.318 | 0.418 | 0.400 | 0.390 | 0.400 | 0.410 |
| 5   | -     | -     | -     | -     | 0.010 | 0.005 | 0.010 | 0.015 |
| 6   | -     | -     | -     | -     | 0.999 | 0.724 | 1.088 | 1.188 |

The resulting estimates of costs per unit of equipment should be interpreted as the costs incurred by the structural unit of the BHS, including, in addition to the device in question, a number of other devices and / or subsystems, such as, for example, conveyors of various types and purposes. The costs of equipment of similar types for BHS of group II are on average higher, which is explained by the need to use more productive and technically advanced equipment in the systems of this group compared to the BHS of group I, which is consistent with practice.

Examples of the results of using a fuzzy model to determine the cost of existing BHS are shown in Figure 1, which shows clear ($y$) and fuzzy ($\tilde{y}$) model estimates, as well as, for comparison, the actual values ($Y$) of the cost of baggage systems described in Table 1. Under the influence of the uncertainties, the difference in estimates of the costs of installation and operation of two different BHS, even if their main design characteristics coincide, can be significant. In general, the results of comparing the actual values and their model estimates indicate the possibility of practical use of the formed technical and economic model.

**Figure 1.** Estimation of the BHS cost of groups I (a) and II (b).
5. Conclusion

Using the proposed two-stage methodology allowed the formation of a multidimensional regression model with fuzzy parameters and clear independent variables, designed to assess the cost of airport BHS. The first stage of the methodology involves the development of a statistical regression technical and economic model of independent value, which then, at the second stage, is used to obtain a more informative fuzzy model for a similar purpose. The inclusion of the developed models in the complex of models of different levels of complexity will allow us to move on to solving urgent problems of designing and optimizing airport systems and complexes with phased refinement of the results.

References

[1] Tanaka H, Vegima S and Asai K 1982 Linear regression analysis with fuzzy modelTransactions on Systems, Man and Cybernetics 12(6) 903-7
[2] Aliev R A, Cerkovnyj A Je and Mamedova G A 1991 Production management in the fuzzy initial information (Moscow: Jenergoatomizdat) p 240
[3] Tanaka H, Ishibuchi H and Yoshikawa S 1995 Exponential possibility regression analysis Fuzzy Sets and Systems 69(3) 305-18
[4] Diamond P 1988 Fuzzy least squares Information Science 46 141-57
[5] Albrecht M 1992 Approximation of functional relationships to fuzzy observations Fuzzy Sets and Systems 49(3) 301-5
[6] Yang M S and Ko C H On cluster-wise fuzzy regression analysis 1997 Transactions on Systems, Man and Cybernetics 27(1) 1-13
[7] Förster E and Rönz B 1992 Regressions- und Korrelationsanalyse: Grundlagen, Methoden, Beispiele (Wiesbaden: Gabler) p 370