OPERATIONS RESEARCH AND DECISIONS

No. 1 2015
DOI: 10.5277/ord150105

Bogdan RĘBIASZ[1]

HYBRID CORRELATED DATA IN RISK ASSESSMENT

A method for evaluating the risks in a situation has been presented where parameters in the calculation are expressed in the form of dependent fuzzy numbers and probability distributions. The procedure of risk estimation combines stochastic simulation with the execution of arithmetic operations on interactive fuzzy numbers. In order to define operations on such numbers, non-linear programming is used. Relations between the parameters presented in the form of fuzzy numbers and probability distributions are expressed by means of interval regression. The results of computations indicate that the relations between parameters have a significant impact on the ratios characterizing risk.

Keywords: *risk, interactive fuzzy numbers, random fuzzy numbers, simulation*

1. Introduction

At present, methods for risk assessment are a fundamental tool for supporting decision-making in an enterprise. Parameters affecting decision-making processes are usually burdened with uncertainty. For many years, the only tool that enabled expressing uncertainty in a mathematical language was the calculus of probability and it still remains the most commonly used tool in practice. Indeed, the probabilistic approach prevails in the literature concerning risk in business activity. In order to estimate this risk, most often a stochastic simulation is used. However, the huge workload of data preparation significantly limits its use in risk assessment. Difficulties also arise in determining the probability distributions of economic parameters. For this reason, apart from quantitative methods, qualitative methods are used in developing forecasts of the aforementioned parameters. In such a case, expert opinions and subjective probability distributions are used to model uncertainty.

[1]AGH University of Science and Technology, Faculty of Management, Gramatyka 10, 30-067 Cracow, Poland, e-mail address: brebiasz@zarz.agh.edu.pl

In numerous decision-making situations, the nature of the uncertainty regarding economic parameters does not satisfy the assumptions of the probability theory. Namely, this happens when uncertainty stems mainly from insufficient information on these parameters and is of the epistemological nature [24]. Gupta acknowledges that uncertainty regarding the forecasted values of economic parameters is often of probabilistic nature; however the available information is fuzzy [11]. In practice, we often face situations where it is not possible to estimate the probability distribution. For example, when there are no data and/or there will not be a sufficient volume of data to enable statistical tests to be performed. On the other hand, the assumption that there are "no data available at all" is also not true. In general, we do have some information available. Most commonly, estimates of unknown values made by experts.

Most real-world problems of risk analysis involve a mixture of quantitative and qualitative data. This is why the conventional probabilistic approach appears to be insufficient to model some decision-making problems (e.g., problems related to risk assessment). To cope with this, many researchers have applied alternative ways of describing uncertainty in the process of risk assessment, among which fuzzy numbers deserve particular attention.

Ward [35] was the first to utilize fuzzy numbers in financial analysis. He presented cash flows by means of trapezoidal fuzzy numbers. Buckley [2, 3] computed net present values of investment projects using fuzzy numbers. Calzi [4] presented principles for describing financial mathematics using fuzzy numbers. Choobineh and Behrens [6] presented the use of possibility distributions in economic analysis. Chiu and Park [5] calculated the efficiency of investment projects by means of fuzzy numbers. They introduced methods that allow one project to be selected from a set of mutually preclusive projects. Esogbue and Hearnes [8] utilized fuzzy numbers in problems concerning the replacement of fixed assets. They aimed to describe the economic life-cycle of fixed assets. Kahraman et al. [17] presented methods for computing the effectiveness ratios of a selected investment project assuming that certain parameters are presented in the form of fuzzy numbers. Kuchta [21] presented the use of fuzzy numbers in capital budgeting. A comparison of the results of evaluating the profitability and the risk of investment projects in the case when the uncertainty of parameters is presented both in the form of probability distributions and fuzzy numbers was presented in [26]. The factors described above caused that at present, apart from probabilistic description of the uncertainty of economic parameters, description based on possibility distributions is being more and more frequently used.

The usefulness of fuzzy and probabilistic approaches in decision-making analysis is viewed in different ways. The majority of authors voice the opinion that fuzzy and probabilistic approaches are supplementary to each other and in every single case one needs to decide which approach will be the most appropriate. The selection of the approach should be conditioned mostly by the degree of subjectivity in the available information. On the other hand, Gupta [11] and Smets [28] claim that in decision-

-making, a probabilistic description of uncertainty is more effective than a possibilistic one. They emphasize the purposefulness of transforming a possibility distribution into a probability distribution by stating that the decision-maker is not interested in "what is possible", but rather in "what is probable". According to [21], selection of the method of representing uncertainty depends mainly on the experience and habits of the decision-maker. Choobineh and Behrens [6] claim that maintaining a probabilistic approach stems more from tradition than from conscious selection.

These two methods for describing the uncertainty of economic parameters (probability distributions, fuzzy numbers) are usually used as alternatives. In practice, most commonly some parameters are specified with probability distributions, while others are given in the form of fuzzy numbers. In economic calculus, the available data are usually heterogeneous, i.e., both uncertain and imprecise, and usually come from various sources. These could be both statistical data, as well as subjective assessments of phenomena provided by experts.

To sum up, one may repeat after Baudrit et al. [1] that randomness and imprecise or missing information are two sources of uncertainty, which impact on risk in business activity. Therefore, in the process of risk assessment it is necessary to take into account both ways of describing uncertainty.

The majority of authors use probability or possibility distributions as alternatives in the process of risk assessment. There are few studies which describe the use of hybrid data, i.e., data partially described by probability distributions, and partially by possibility distributions [1, 7, 9, 10, 12]. The use of such data allows us to reflect more properly our knowledge on economic parameters. As suggested Ferson and Ginzburg [9], distinct methods are needed to adequately represent random variability (often referred to as "objective uncertainty") and imprecision (often referred to as "subjective uncertainty"). In risk assessment, no distinction is traditionally made between these two types of uncertainty, both being represented by means of a single probability distribution [1]. In the case of partial ignorance, the use of a single probability measure introduces information that is in fact not available. This may seriously bias the outcome of risk analysis in a non-conservative manner [9].

Kaufman and Gupta [18] introduce hybrid numbers which simultaneously express inaccurateness and randomness. Guyonnet et al. [12] propose a method which facilitates the estimation of risk in the case when both probability and possibility distributions are involved. This method resulted from the modification of a method proposed by Cooper et al. [7]. Methods for processing hybrid data combine stochastic simulation with the arithmetic of fuzzy numbers. The result of processing such data is defined as two cumulative distribution functions: optimistic and pessimistic [12]. Similarly, Baudrit et al. [1] use probability and possibility distributions in risk analyses. They use a procedure for data processing, which also combines stochastic simulation with the arithmetic of fuzzy numbers. As a result of processing such data, they obtain a random fuzzy variable which characterizes the examined phenomenon.

In the process of data processing, the above-mentioned methods of risk assessment do not account for relations between the parameters described by means of fuzzy numbers and probability distributions. In the case of analyzing risk in business activity, consideration of the relations between parameters of economic calculus is crucial. The prices of an enterprise's products are naturally correlated with the price of the raw materials required for their production. Often, the sales volumes of particular products are also correlated.

Assessment of business activity risk requires the following actions [1]:

- acquisition and preparation of data regarding the parameters of economic calculus,
- selection of the method for representing uncertainty regarding particular parameters in the calculus,
- developing a method for processing hybrid data,
- defining a synthetic coefficient assessing risk in business activity.

In the paper, a new method for processing hybrid data has been proposed. The method takes into account the correlation between parameters expressed in the form of fuzzy numbers and probability distributions. Risk is expressed by means of optimistic and pessimistic cumulative distributions describing the examined financial ratio, according to the methodology presented in the study [1] and, alternatively, by means of fuzzy variance, according to the method proposed by Liu and Liu [23]. The results of computations obtained with and without taking into account the correlation between the parameters of calculus are compared. In the latter case, a combination of stochastic simulation and Zadeh's extension principle is used to process data, according to the procedure presented in [1].

2. Preliminaries

2.1. Arithmetic operations on fuzzy numbers

The effective processing of data expressed in the form of fuzzy numbers requires effective methods for performing arithmetic operations on such numbers. Operations can be defined on fuzzy numbers of arbitrary type, in line with Zadeh's extension principle, using operations on α-levels of these numbers [19]. Given two fuzzy numbers $\tilde{A}$ and $\tilde{B}$ and their α-levels A_α and B_α, an arithmetic operation $* \in (+, -, \times, /)$ on these numbers can be defined in accordance with the following dependence [20]:

$$\tilde{A} * \tilde{B} = \{ A_\alpha * B_\alpha \} (\alpha \in (0, 1]) \tag{1}$$

This definition implicitly assumes that all combinations of values belonging to the respective interval numbers (α-levels of the respective fuzzy numbers) are possible [19]. However, this assumption is not always true. For example, high prices for hot rolled sheets will generally result in high prices for cold rolled sheets. Combinations where one of these products has a low price and the other has a high price are unlikely to occur, as these quantities are positively correlated.

As mentioned above, variables that describe various economic phenomena are usually dependent. Often one speaks about variables that are correlated. In such a situation, the selection of schemes for performing arithmetic operations on so-called interactive fuzzy numbers often causes a problem. Undoubtedly, one should apply a scheme that enables these dependences to be taken into account. In this study, nonlinear programming methods were used to carry out arithmetic operations on fuzzy numbers. In order to take into account the relation between fuzzy numbers, interval regression was used.

2.2. Fuzzy and interval regression

Fuzzy regression involves expressing the parameters of a regression equation in the form of fuzzy numbers. Interval regression is a specific case of fuzzy regression. The parameters of a regression equation are in this case expressed in the form of bounded intervals [29–32]. Interval and fuzzy regressions are used to solve numerous practical problems [13].

Several methods have been developed to estimate the parameters of the interval regression equation. The best known method uses linear programming for this purpose [22, 31, 32]. However, this approach has many faults.

- It is often the case that some of the estimated regression parameters tend to be crisp; it even happens that the method produces only a few unexpectedly wide interval parameters, while others are crisp (a drawback called unbalancedness). This problem is generally considered to be the most serious and the most restrictive drawback that limits the usefulness of this method [14, 30, 31, 34].
- The method might produce interval regression parameters with centers that only poorly fit the data with respect to traditional goodness-of-fit measures (such as R-squared). In the literature, this problem is referred to as the non-centrality property [22, 29, 30].
- The method is highly sensitive to outliers [22, 28].

Many authors presented solutions eliminating these faults. Quadratic programming methods are often combined with the least squares method [14, 30, 31, 34]. Other methods of interval regression are based on the Minkowski distance [15] or multicriteria programming [34].

These modified methods provide more balanced intervals representing the coefficients of interval regression equations, however they require longer computation time. Moreover, the estimates of weight coefficients are made by experts, thus these methods are heuristic [13].

Another method for estimating the parameters of interval regression equations was developed by Hladik and Černy [13]. It seems to be very promising. Below, a variant of this method, which covers the case when input data (dependent and independent variables) are presented as determined numbers (crisp input-crisp output), is described. The proposed method is based on sensitivity analysis for linear systems. It consists of two stages:

- The estimation of the centers of the interval parameters using standard estimators.
- The estimation of the radii of the interval parameters.

Suppose that p observations are given:

$$\hat{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}, \quad X = \begin{bmatrix} x_{11} \dots x_{1n} \\ \vdots \quad\quad \vdots \\ x_{p1} \dots x_{pn} \end{bmatrix} \tag{2}$$

where: X – input matrix, $\hat{y}$ – output vector.

The problem is to determine interval regression parameters $\hat{a}$

$$\hat{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \tag{3}$$

comprising all the possibilities determined by the model and data. Formally, $\hat{a}$ must satisfy (see [13]):

$$y_j \in x_{j1}a_1 + x_{j2}a_2 + \dots + x_{jn}a_n, \qquad \forall j = 1, \dots, p \tag{4}$$

It is natural to try and obtain interval estimates which possess several desirable properties. They should be as tight as possible, they should be balanced and they should respect the central tendency. Robustness with respect to outliers is also an advantage.

Hladik and Černy [13] present interval regression coefficients in the form of intervals $\hat{a} = \left[\hat{b} - \hat{c},\ \hat{b} + \hat{c}\right]$. The vector $\hat{b}$ is estimated by means of the least squares method. The width of the intervals defined by the vector $\hat{c}$ is expressed in the form: $\hat{c} = \delta\hat{c}^{\Delta}$, where $\hat{c}^{\Delta}$ is a non-negative vector of sensitivity coefficients and $\delta \geq 0$ is an unknown value. Owing to the introduction of sensitivity coefficients, the width of the intervals representing regression coefficients may be controlled.

Estimating the interval regression parameters reduces to the problem of finding the minimum value $\delta \geq 0$, such that for the vector $\hat{a} = \left[\hat{b} - \delta\hat{c}^{\Delta},\ \hat{b} + \delta\hat{c}^{\Delta}\right]$, the following is met:

$$\forall j \in \{1,\ \ldots,\ p\}\ \exists \hat{a}' \in \left[\hat{b} - \delta\hat{c}^{\Delta},\ \hat{b} + \delta\hat{c}^{\Delta}\right] : y_j = X_{j*}\hat{a}' \tag{5}$$

Hladik and Černy [13] present a simple formula to calculate δ. Namely, when there exists $j \in \{1, \ldots, p\}$ such that $|X|_{j*}\hat{c}^{\Delta} = 0$ and simultaneously $y_j \neq X_{j*}\hat{b}$, then there does not exist any δ, which would satisfy condition (4). In any other case, δ may be calculated based on the formula:

$$\delta^* = \max_{j:|X|_{j*}>0} \frac{\left|y_j - X_{j*}\hat{b}\right|}{|X|_{j*}\hat{c}^{\Delta}} \tag{6}$$

The value δ^* thus defined is the minimum value of δ. This relation constitutes an effective method for solving the problem of interval regression. It eliminates the issue of significant imbalance between intervals representing particular regression coefficients and the problem of non-centric location of the estimated intervals.

The sensitivity coefficients are most often assumed to satisfy $c_i^{\Delta} = 1$ or $cc_i^{\Delta} = |b_i|$ for $i = 1, \ldots, n$. The former case is one of the most natural choices, since it minimizes the sum of the radii of the interval estimators. The latter case is called the relative tolerance approach. In this case, the minimum value of δ (δ^*) gives the following information to a user: it is sufficient to perturb the regression parameters by no more than $100\delta^*$% in order to cover all the observations. Hence, it is an alternative measure of goodness-of-fit.

2.3. Fuzzy random variables

Liu and Liu [23] define fuzzy random variables in the following manner. Let us assume that Z is a set of fuzzy variables. Each element $\boldsymbol{z}$ of the set Z is characterized by a membership function $\mu_{\boldsymbol{z}}$. Let us assume that (Ω, Σ, P) is a probability space.

A fuzzy random variable is then defined as a map $\xi : \Omega \to Z$ such that for each closed subset C of the space $\Re$,

$$\xi^{*}(C)(\omega) = \mathrm{Pos}\{\ \xi(\omega) \in C\} = \sup_{x \in C} \mu_{\xi(\omega)}(x) \tag{7}$$

is a measurable function ω, where $\mu_{\xi(\omega)}$ is the possibility distribution of the fuzzy variable $\xi(\omega)$.

In order to effectively use fuzzy random variables in risk assessment, it is necessary to define the expected value and the variance of such variables. The aforesaid authors define the expected value of such variables in various ways. Most frequently, it is defined in the form of a fuzzy set [25]. However, in decision-making problems, it is desirable that the expected value is expressed in the form of a scalar [23]. This facilitates the interpretation of results. Methods using such values are readily accepted by practitioners.

Liu and Liu [23] proposed a new method for calculating the expected value and the variance of a fuzzy random variable. They express these values in the form of a scalar. The expected value $E(\xi)$ of a normal fuzzy variable ξ defined in the probability space (Ω, Σ, P) is given by the following formula:

$$E(\xi) = \int_{\Omega} \left[\int_{0}^{\infty} Cr\{\xi(\omega) \geq x\} dx - \int_{0}^{\infty} Cr\{\xi(\omega) \leq x\} dx \right] P(d\omega) \tag{8}$$

where $Cr\{\xi(\omega)\}$ is the credibility distribution of ξ.

Further, the above authors define the variance of a fuzzy random variable according to the formula:

$$Dev(\xi) = E\left[(\xi - E[\xi])^{2} \right] \tag{9}$$

Based on a fuzzy random variable, upper and lower distribution functions may be estimated. These functions characterize uncertainty as to future realizations of the analyzed variable. It is known that each fuzzy variable $\tilde{Z}$ with possibility distribution π induces a random set [27]. Let the α-levels of this variable be denoted by π_{α}. The focal elements of a random set generated by the fuzzy variable $\tilde{Z}$ are α-levels $(\pi_{\alpha})_{j=1,\cdots,q}$, where $\alpha_0 = \alpha_1 = 1 > \alpha_2 > \dots \alpha_q > \alpha_{q+1} = 0$. The values $\left(v_j = \alpha_j - \alpha_{j+1}\right)_{j=1,\dots,q}$ constitute the probability mass of the random set generated. If there are n fuzzy variables $\left(\tilde{X}_i\right)_{i=1,2,\ldots,n}$ with corresponding probabilities $(p_i)_{i=1,2,\ldots,n}$, then $\left(\pi_\alpha\right)_{j=1,\ldots,q;\ i=1,2,\ldots,n}$ are focal elements of a random set, while the values p_{ij} ($p_{ij} = p_i \times (\alpha_{ji} - \alpha_{j+1;i})$) constitute the probability mass of a random set. Based on such a random set, one may determine the upper $\overline{F}(x)$ and lower $\underline{F}(x)$ distribution function according to the formulas:

$$\overline{F}(x) = \mathrm{Pl}\left(X \in [-\infty, x]\right) \tag{10}$$

$$\underline{F}(x) = \mathrm{Bel}\left(X \in [-\infty, x]\right) \tag{11}$$

where Pl and Bel are the Belief and Plausibility functions, respectively.

3. The hybrid propagation method

In order to solve the stated problem, one needs to determine the value of $f(\hat{\mathbf{X}})$, where $\hat{\mathbf{X}} = \left[X_1, X_2, \ldots, X_m\right]$ is a vector of variables burdened with uncertainty. It is assumed that k variables ($k < m$) are random variables $[X_1, X_2, \ldots, X_k]$ and $m - k$ variables are fuzzy variables $[X_{k+1}, X_{k+2}, \ldots, X_m]$. Additionally, it is assumed that there may be defined subsets X^K of correlated variables X_i; $X^K = \left\{X_i,\ i \in K\right\}$, $K \in K_s$. In such a case, K is a subset of the indices of the correlated variables, and K_s is the set of the indices of the selected subsets of correlated variables.

The proposed procedure for determining the value of $f(\hat{\mathbf{X}})$ involves two stages. It combines a stochastic simulation procedure with the execution of arithmetic operations on dependent fuzzy numbers. In order to execute such arithmetic operations, non-linear programming is used. In this case, the computational procedure is as follows: The realizations $[x_1, x_2, \ldots, x_k]$ of the random variables are drawn using a procedure which accounts for the correlation of variables. These realizations and the fuzzy variables $[X_{k+1}, X_{k+2}, \ldots, X_m]$ allow us to determine $f(x_1, x_2, \ldots, x_k, X_{k+1}, X_{k+2}, \ldots, X_m)$ as a fuzzy number. This can be achieved using the concept of α-levels. The upper bound (sup) and lower bound (inf) of an α-level of a fuzzy number $f(x_1, x_2, \ldots, x_k, X_{k+1}, X_{k+2}, \ldots, X_m)$ may be determined by solving the following non-linear programming tasks.

When searching for sup, find:

$$f(x_1, x_2, \ldots, x_k, X_{k+1}, X_{k+2}, \ldots, X_m) \rightarrow \max \tag{12}$$

When searching for inf, find:

$$f(x_1,\ x_2,\ ...,\ x_k,\ X_{k+1},\ X_{k+2},\dots,\ X_m)\to \min \tag{13}$$

subject to the following constraints:

$$\inf\left(X_i\right)_\alpha \le x_i \le \sup(X_i)_\alpha \qquad \text{for} \qquad i=1,2,\dots,m \tag{14}$$

$$x_i \ge \inf\left(a_1^{iz}\right)\times x_z + \inf\left(a_2^{iz}\right) \quad \text{for } i\in K, z\in K;\ i\ne z;\ i\le k,\ z\le k,\ K\in K_s \tag{15}$$

$$x_i \le \sup\left(a_1^{iz}\right)\times x_z + \sup\left(a_z^{iz}\right) \quad \text{for } i\in K, z\in K;\ i\ne z;\ i\le k,\ z\le k,\ K\in K_s \tag{16}$$

$$x_i \ge \inf\left(a_1^{iz}\right)\times x_z + \inf\left(a_2^{iz}\right) \quad \text{for } i\in K, z\in K;\ i>k, z\le k;\ K\in K_s \tag{17}$$

$$x_i \le \sup\left(a_1^{iz}\right)\times x_z + \sup\left(a_2^{iz}\right) \quad \text{for } i\in K, z\in K;\ i>k, z\le k;\ K\in K_s \tag{18}$$

The values a_1^{iz}, a_2^{iz} are the coefficients of the interval regression equations determining the relation between the variables X_i and X_z. These coefficients are calculated based on historical data. Historical data concerning the appropriate variables (e.g. the prices of products and raw materials, sales volumes) are available and expressed in the form of determined volumes. In order to obtain the values of the coefficients, one may use the method developed by Hladik and Černy described above [13].

Drawing values $[x_1, x_2, ..., x_k]$ and determining $f(x_1,\ x_2,\ ...,\ x_k,\ X_{k+1},\ X_{k+2},...,\ X_m)$ is repeated $\ddot{n}$ times. As a result $\ddot{n}$ fuzzy sets characterized by membership functions $(\mu_1^f, ..., \mu_{\ddot{n}}^f)$ are obtained. In this case, the value $f(\hat{X})$ is represented by a random fuzzy variable. The above hybrid procedure may be described by the following algorithm.

START

Step1. Define α_0, $\wp, n=1$

Step 2. Randomly generate a vector $[x_1, x_2, ..., x_k]$ taking into account the correlation of variables.

Step 3. $\alpha = \alpha_0$

Step 4. Define α-levels $(X_i)_\alpha$ for $i=k+1, k+2,\dots,m$

Step 5. Define (sup) and (inf) for α-levels of the fuzzy numbers defining $f(\hat{X})$

Find $f(x_1, x_1, ..., x_k, x_{k+1}, x_{k+2}, ..., x_m)\to \max$

and $f(x_1, x_1, \ldots, x_k, x_{k+1}, x_{k+2}, \ldots, x_m) \to \min$ with the problem constraints specified by inequalities (14)–(18)
Step 6. $\alpha = \alpha + \wp$
Step 7. If $\alpha \le 1$ go to step 4
Step 8. $n = n + 1$
Step 9. If $n \le \ddot{n}$ go to step 2
Step10. Define the set of fuzzy numbers ($\mu_1^f, \ldots, \mu_{\ddot{n}}^f$)
STOP

4. Numerical examples

In order to illustrate the effectiveness of the proposed method, the results of calculations made for a simple model are presented below. The method was tested based on the example of calculating the operating profit for a metallurgical industrial enterprise.

The calculation was performed for the production system presented in Fig. 1.

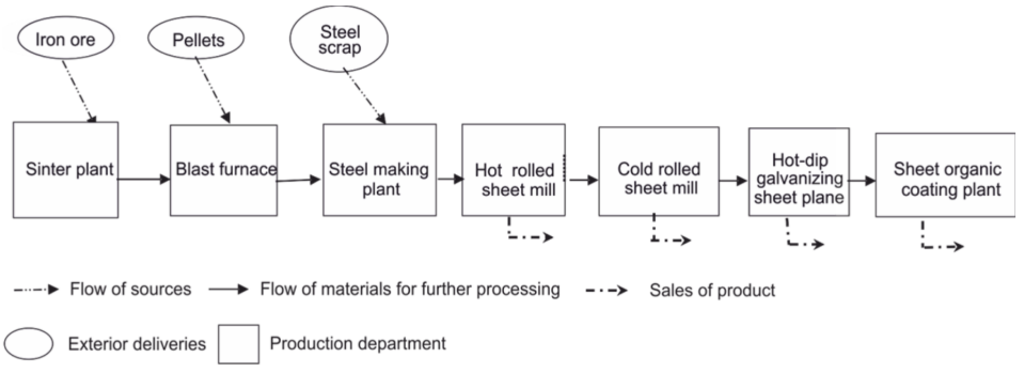

Fig. 1. Flow diagram of the production system analyzed

The operating profit can be expressed by the formulas (19)–(33):

$$ZO = c_{bg}G_{bg} + c_{bz}G_{bz} + c_{bo}G_{bo} + c_{bp}G_{bp} - k_{sp}Pr_{sp} - k_{su}Pr_{su} - k_{st}Pr_{st} - k_{bg}Pr_{bg} - k_{bz}Pr_{bz} - k_{bo}Pr_{bo} - k_{bp}Pr_{bp} - Zu_{zl}c_{zł} - Zu_{gr}c_{gr} - Zu_{ru}c_{ru} - kf \tag{19}$$

$$Pr_{sp} = Pr_{su}m_{sp} \tag{20}$$

$$Pr_{su} = Pr_{st} m_{su} \tag{21}$$

$$Pr_{st} = Pr_{bg} m_{pł} \tag{22}$$

$$Pr_{bg} = Pr_{bz} m_{bg} + G_{bg} \tag{23}$$

$$Pr_{bz} = Pr_{bo} m_{bz} + G_{bz} \tag{24}$$

$$Pr_{bo} = Pr_{bp} m_{bo} + G_{bo} \tag{25}$$

$$Pr_{bp} = G_{bp} \tag{26}$$

$$G_{bg} = JZ_{bg} u_{bg} \tag{27}$$

$$G_{bz} = JZ_{bz} Zu_{bz} \tag{28}$$

$$G_{bo} = JZ_{bo} Zu_{bo} \tag{29}$$

$$G_{bp} = JZ_{bp} Zu_{bp} \tag{30}$$

$$Zu_{zł} = Pr_{st} m_{zł} \tag{31}$$

$$Zu_{gr} = Pr_{su} m_{gr} \tag{32}$$

$$Zu_{ru} = Pr_{sp} m_{ru} \tag{33}$$

where:

G_{bg}, G_{bz}, G_{bo}, G_{bp} are the sales of hot rolled strip, cold rolled sheets, hot dip galvanized strip and sheets, organic coated sheets, respectively.

JZ_{bg}, JZ_{bz}, JZ_{bo}, JZ_{bp} are the apparent consumption of hot rolled strip, cold rolled sheets, hot dip galvanized strip and sheets, organic coated sheets, respectively.

u_{bg}, u_{bz}, u_{bo}, u_{bp} are the market shares of hot rolled strip, cold rolled sheets, hot dip galvanized strip and sheets, organic coated sheets, respectively.

c_{ru}, c_{gr}, $c_{zł}$, c_{bg}, c_{bz}, c_{bo}, c_{bp} are the prices of iron ore, pellets, steel scrap, hot rolled strip, cold rolled sheets, hot dip galvanized strip and sheets, organic coated sheets, respectively.

m_{sp}, m_{su}, $m_{pł}$, m_{bg}, m_{bz}, m_{bo} are sinter consumption per tonne of pig iron, pig iron consumption per tonne of continuous casting stands, continuous casting stand consumption per tonne of hot rolled sheets, hot rolled sheet consumption per tonne of cold rolled sheets, cold rolled sheet consumption per tonne of galvanized sheets, galvanized sheet consumption per tonne of organic coated sheets, respectively.

$m_{zł}$, m_{gr}, m_{ru} are scrap consumption per tonne of steel, pellet consumption per tonne of pig-iron, iron ore consumption per tonne of sinter, respectively.

Pr_{sp}, Pr_{su}, Pr_{st}, Pr_{bg}, Pr_{bz}, Pr_{bo}, Pr_{bp} are the production levels of sinter, pig iron, continuous casting stands, hot rolled strip, cold rolled sheets, hot dip galvanized strip and sheets, organic coated sheets, respectively.

k_{sp}, k_{su}, k_{st}, k_{bg}, k_{bz}, k_{bo}, k_{bp} are the appropriately adjusted unit variable costs of sinter, pig iron, continuous casting stands, hot rolled strip, cold rolled sheets, hot dip galvanized strip and sheets, organic coated sheets, respectively. This cost does not account for the values of steel products manufactured in previous stages of the cycle or the value of raw materials used; corrections are made in order to avoid multiple calculation of the same cost components during the calculation of profit, according to formula (19).

k_f are the company's fixed costs.

$Zu_{zł}$, Zu_{gr}, Zu_{ru} are scrap, pellet and iron ore consumption, respectively.

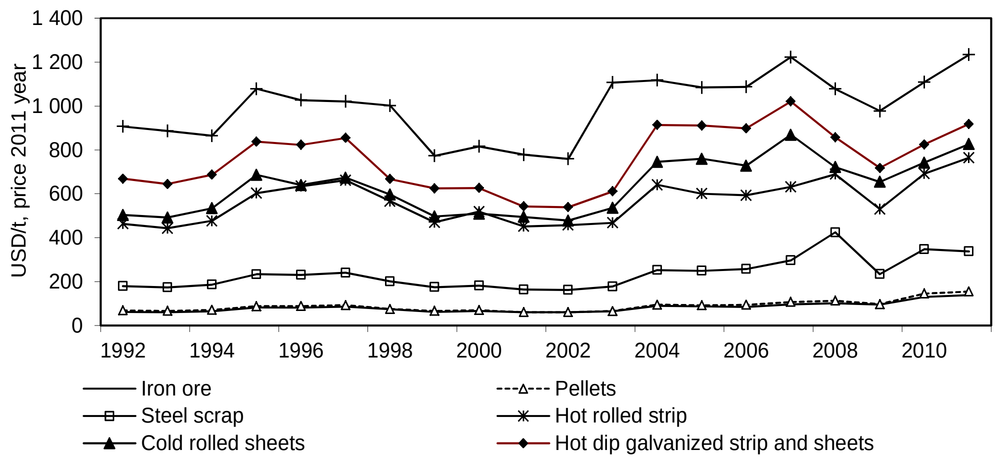

Fig. 2. Prices of metallurgical products manufactured by the company analyzed and the prices of iron ore, pellets and steel scrap in 1992–2011

Figure 2 shows the prices of metallurgical products manufactured by the company analyzed and the prices of iron ore, pellets and steel scrap in 1992–2011. Figure 3 shows the apparent consumption of metallurgical products manufactured by the company analyzed.

Table 1 describes the trapezoidal fuzzy numbers and parameters of the normal probability density function (average value – m and standard deviation – σ) specifying the forecast of parameters for calculating the operating profit for the company analyzed.

Table 1. Trapezoidal fuzzy numbers representing forecasts of the prices of products and raw materials, material consumption rates and parameters of the normal probability density function characterizing the apparent consumption of metallurgical products

Price	Trapezoidal fuzzy numbers [USD/t]
Iron ore	(111.7; 120.0; 133.3; 141.7)
Pellets	(125.0; 133.3; 146.7; 156.7)
Steel strap	(313.3; 320.0; 336.7; 345.0)
Hot rolled strip	(666.7; 680.0; 711.7; 728.3)
Cold rolled sheets	(715.0; 730.0; 763.3; 781.7)
Hot dip galvanized strip and sheets	(805.0; 821.7; 860.0; 880.0)
Organic coated sheets	(1080.0; 1101.7; 1153.3; 1175.0)
Material consumption rates	Trapezoidal fuzzy numbers [t/t]
Iron ore – sinter	(0.918; 0.920; 0.920; 0.922)
Sinter – pig iron	(1.352; 1.354; 1.359; 1.362)
Pellets – pig iron	(0.338; 0.339; 0.340; 0.341)
Scrap – continuous casting stands	(0.269; 0.276; 0.279; 0.288)
Pig iron – continuous casting stands	(0.855; 0.860; 0.870; 0.875)
Continuous casting stands – hot rolled strip	(1.058; 1.064; 1.075; 1.078)
Hot rolled strip – cold rolled sheets	(1.105; 1.111; 1.124; 1.130)
Cold rolled sheets – hot dip galvanized strip and sheets	(1.010; 1.020; 1.026; 1.031)
Hot dip galvanized strip and sheets – organic coated sheets	(0.998; 0.999; 1.000; 1.001)
Apparent consumption	(Average; standard deviation) [10^3 t]
Hot rolled strip	(2704.0; 117.5)
Cold rolled sheets	(1162.3; 51.4)
Hot dip galvanized strip and sheets	(1147.9; 52.4)
Organic coated sheets	(708.4; 30.8)

The relations between the prices of the ranges of steel products analyzed, prices of iron ore, pellets and scrap, as well as the apparent consumption of particular product ranges were expressed by means of an interval regression model. The coefficients of the regression equation were estimated using the method described in Section 2. Table 2 presents the coefficients of the interval regression equations characterizing the relations between the prices of the products manufactured by the producer analyzed and the prices of iron ore, pellets and scrap. Table 3 presents the correlation matrix for the apparent consumption of particular product ranges manufactured by the producer.

Table 2. Coefficients of the interval regression equations depicting the interrelations between the prices of particular product ranges produced by the manufacturer in question and prices of raw materials

Independent variable	*a*	Dependent variable						
		Iron ore	Pellets	Steel scrap	Hot rolled strip	Cold rolled sheets	Hot dip galvanized strip and sheets	Organic coated sheets
Iron ore	1		[0.80; 0.89]	[0.20; 0.34]	[0.17; 0.26]	[0.10; 0.20]	[0.06; 0.17]	[0.07; 0.16]
	2		[6.87; 7.69]	[13.00; 22.39]	[–30.09; –46.54]	[–9.18; –18.17]	[–2.80; –7.70]	[–21.79; –48.12]
Pellets	1	[1.11; 1.25]		[0.24; 0.42]	[0.19; 0.30]	[0.11; 0.24]	[0.07; 0.21]	[0.08; 0.19]
	2	[–7.37; –8.30]		[8.46; 14.91]	[–39.89; –63.98]	[–15.30; –32.08]	[–7.89; –23.86]	[–28.57; –67.82]
Steel strap	1	[1.46; 4.07]	[1.32; 3.42]		[0.40; 0.87]	[0.21; 0.71]	[0.32; 0.60]	[0.13; 0.56]
	2	[5.16; 14.37]	[14.65; 38.10]		[–80.04; –172.16]	[–26.03; –86.59]	[–38.69; –73.93]	[–42.87; –178.41]
Hot rolled strip	1	[3.48; 4.51]	[2.90; 3.80]	[1.01; 1.36]		[0.60; 0.79]	[0.48; 0.66]	[0.39; 0.66]
	2	[211.97; 275.29]	[237.92; 311.40]	[247.72; 334.56]		[112.16; 148.39]	[112.61; 155.43]	[31.68; 53.32]
Cold rolled sheets	1	[3.55; 5.80]	[3.04; 4.83]	[1.12; 1.70]	[0.87; 1.41]		[0.75; 0.92]	[0.57; 0.96]
	2	[191.11; 312.26]	[221.35; 351.44]	[239.84; 364.71]	[–10.59; –17.12]		[–2.18; –2.65]	[–95.30; –159.73]
Hot dip galvanized strip and sheets	1	[3.59; 5.76]	[3.14; 4.89]	[1.19; 1.78]	[0.96; 1.50]	[0.98; 1.22]		[0.60; 1.07]
	2	[292.01; 468.72]	[318.79; 497.23]	[331.01; 492.93]	[45.67; 71.29]	[57.88; 72.08]		[–48.89; –87.59]
Organic coated sheets	1	[3.96; 6.14]	[3.28; 5.20]	[1.16; 1.79]	[0.35; 0.87]	[0.81; 1.31]	[0.63; 1.12]	
	2	[458.75; 712.03]	[482.39; 765.86]	[511.55; 790.07]	[385.57; 941.97]	[248.18; 403.48]	[236.97; 420.89]	

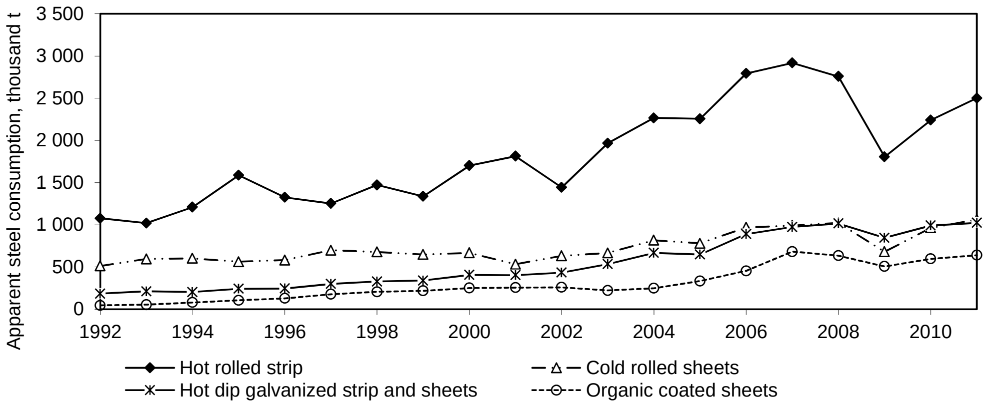

Fig. 3. Apparent consumption of metallurgical products manufactured by the company analyzed

Table 3. Correlation matrix for the apparent consumption of metallurgical products manufactured by company analyzed

Material	Hot rolled strip	Cold rolled sheets	Hot dip galvanized strip and sheets	Organic coated sheets
Hot rolled strip	1.000	0.878	0.911	0.863
Cold rolled sheets	0.878	1.000	0.915	0.888
Hot dip galvanized strip and sheets	0.911	0.915	1.000	0.966
Organic coated sheets	0.863	0.888	0.966	1.000

The value of the fixed costs was assumed to be USD 315 090 thousand/year. The adjusted unit variable processing costs for particular product ranges were also adopted at the levels given in Table 4. The market shares of particular product ranges adopted in the calculations are given in Table 5.

Table 4. Adjusted unit variable processing costs for particular product ranges

Product	Sinter	Pig iron	Continuous casting stands	Hot rolled strip	Cold rolled sheets	Hot dip galvanized strip and sheets	Organic coated sheets
Cost, USD/t	16.6	153.8	25.4	28.4	28.0	116.7	175.3

Table 5. Market shares of the particular product ranges

Hot rolled strip	Cold rolled sheets	Hot dip galvanized strip and sheets	Organic coated sheets
42.5%	40.0%	46.0%	45.0%

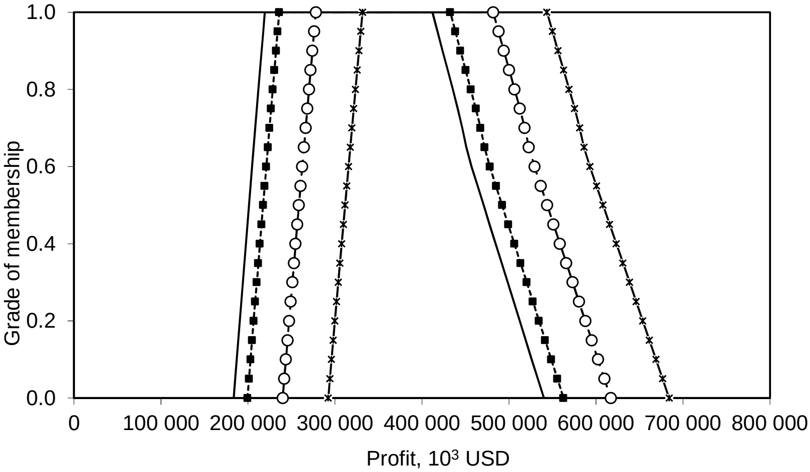

Fig. 4. Example of fuzzy numbers depicting the operating profit calculated in selected iterations of the computer simulation

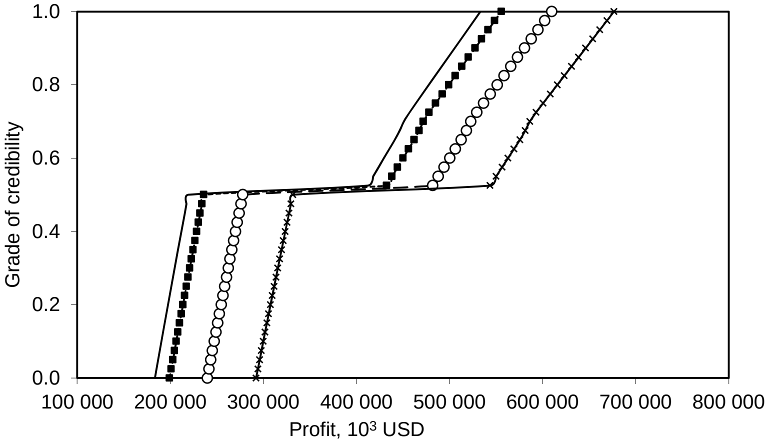

Fig. 5. Credibility distribution functions depicting the operating profit calculated in selected iterations of the computer simulation

Figures 4 and 5 depict illustrative results of the calculations of the operating profit for the manufacturer analyzed. Figure 4 presents four fuzzy numbers characterizing the operating profit calculated in selected iterations of the computer simulation and Fig. 5 presents the corresponding credibility distribution functions. The average value of the operating profit was USD 361 333.6 thousand. and the standard deviation USD 102 303.6 thousand. Figure 6 depicts pessimistic and optimistic cumulative distribution functions resulting from these computations.

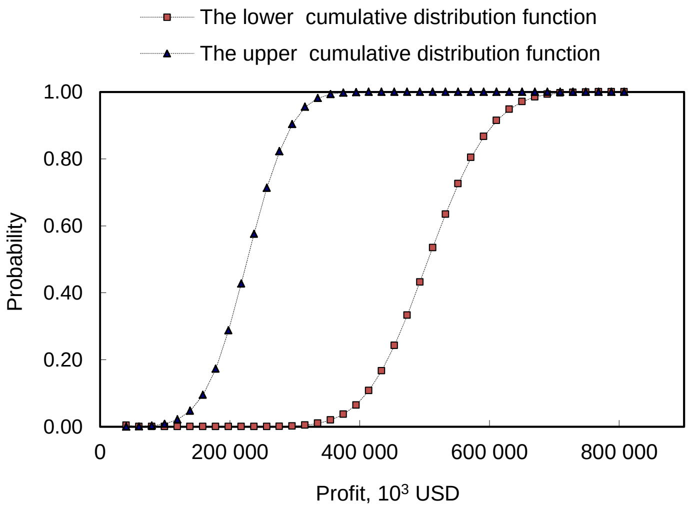

Fig. 6. Upper and lower distribution functions depicting the operating profit

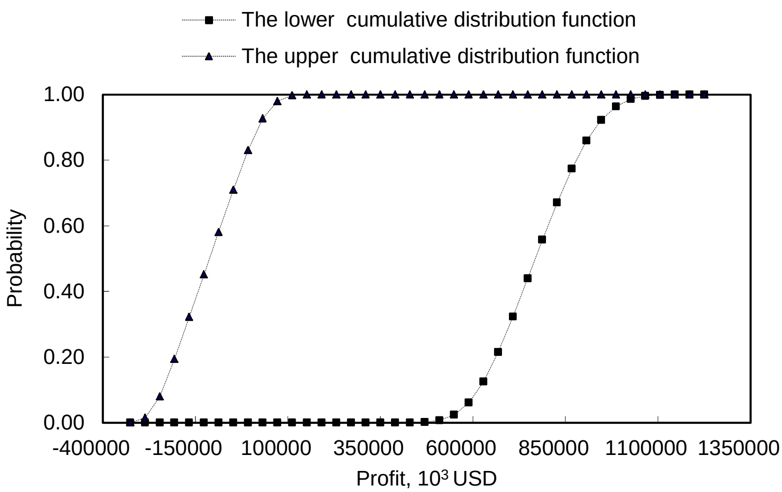

Fig. 7. Upper and lower distribution functions depicting the operating profit obtained by ignoring the correlation of parameters in the economic calculation

Additionally, calculations were carried out without considering the correlation of economic parameters. In this case, the average operating profit obtained was 315 488.9 and the standard deviation was 308 475.5. Figure 7 presents the resulting optimistic and pessimistic cumulative distribution functions. These figures indicate that accounting for the correlation of the variables has a considerable impact on estimating the operating profit.

5. Conclusion

In the literature devoted to the issues of risk, approaches based on probability theory are dominant. Nevertheless, numerous publications indicate the possibility of using other methods to describe economic parameters. Fuzzy numbers are used most frequently. Recent publications present methods which allow us to assess risk in situations when hybrid data are involved, i.e., when some parameters are described by probability distributions and some by fuzzy numbers.

This study considers a proposed method of risk assessment which facilitates the processing of hybrid data, taking into account the correlation between economic parameters. The procedure for processing such data combines stochastic simulation with Zadeh's extension principle (a method for the execution of arithmetic operations on fuzzy numbers). Non-linear programming was used in the execution of arithmetic operations on interactive fuzzy numbers. The relations between economic parameters were expressed in the form of interval regression. The proposed method may be used in the case of fuzzy numbers with arbitrary membership functions, because arithmetic operations are executed on α-levels. This method is a universal and flexible tool for processing hybrid data, which enables existing correlations to be taken into account.

As a result of processing such data, a solution in the form of a random fuzzy set was obtained. Each element of this set results from one replication of a stochastic simulation. Assessment of the level of risk is obtained by computing the standard deviation or by estimating the upper and lower cumulative distribution functions of the operating profit.

The results of the computations presented indicate that the correlation of economic parameters has a considerable impact on the estimates of the operating profit. Risk assessment without consideration of these interrelations is subject to a considerable, systematic error.

References

[1] BAUDRIT C., DUBOIS D., GUYONET D., *Joint propagation and exploitation of probabilistic and possibilistic information in risk assessment*, IEEE Transaction on Fuzzy Systems, 2006, 14 (5), 593–607.

[2] BUCKLEY J.J., *The fuzzy mathematics of finance*, Fuzzy Sets and Systems, 1987, 21 (4), 257–273.

[3] BUCKLEY J.J., *Solving fuzzy equations in economics and finance*, Fuzzy Sets and Systems, 1992, 48 (4), 289–296.

[4] CALZI M.L., *Toward a general setting for the fuzzy mathematics of finance*, Fuzzy Sets and Systems, 1990, 35 (4), 265–280.

[5] CHIU C.Y., PARK S.C., *Fuzzy cash flow analysis using present worth criterion*, England Economic, 1994, 39 (2), 113–138.

[6] CHOOBINEH F., BEHRENS A., *Use of intervals and possibility distribution in economic analysis*, Journal of the Operational Research Society, 1992, 43 (9), 907–918.

[7] COOPER A., FERSON S., GINZBURG L., *Hybrid processing of stochastic and subjective uncertainty data*, Risk Analysis, 1996, 16 (6), 785–791.
[8] ESOGBUE A.O., HEARNES W.E., *On replacement models via a fuzzy set theoretic framework*, IEEE Transactions on Systems Manufacturing and Cybernetics. Part C. Applications and Reviews, 1998, 28 (4), 549–558.
[9] FERSON S., GINZBURG L.R., *Different methods are needed to propagate ignorance and variability*, Reliability Engineering System Safety, 1996, 54, 133−144.
[10] FERSON S., *What Monte Carlo Methods cannot do*, Human and Ecological Risk Assessment, 1996 (2), 990–1007.
[11] GUPTA C.P., *A note on transformation of possibilistic information into probabilistic information for investment decisions*, Fuzzy Sets and Systems, 1993, 56 (2), 175–182.
[12] GUYONNET D., BOURGINE B., DUBOIS D., FARGIER H., CME B., CHILS P.J., *Hybrid approach for addressing uncertainty in risk assessment*, Journal of Environmental Engineering, 2003, 126, 68–76.
[13] HLADIK M., ČERNY M., *Interval regression by tolerance analysis approach*, Fuzzy Sets and Systems, 2012, 193 (1), 85–107.
[14] HUANG C.H., KAY H.-Y., *Interval regression analysis with soft-margin reduced support vector machine*, Lecture Notes in Computer Science, 2009, NLAI 5579, 826–835.
[15] INUIGUCHI M., FUJITA H., TANINO T., *Robust interval regression analysis based on Minkowski difference*, SICE 2002, Proc. 41st SICE Annual Conference, Osaka, Japan, 2002, 4, 2346–2351.
[16] JENG T.-J., CHUANG C.-C., SU F.-S., *Support vector interval regression networks for interval regression analysis*, Fuzzy Sets and Systems, 2002, 138 (2), 283–300.
[17] KAHRAMAN C., RUAN D., TOLGA E., *Capital budgeting techniques using discounted fuzzy versus probabilistic cash flows*, Information Science, 2002, 42 (1), 57–76.
[18] KAUFMANN A., GUPTA M.M., *Introduction to fuzzy arithmetic. Theory and application*, Van Nostrand Reinhold, New York 1985.
[19] KLIR G.J., *A principle of uncertainty and information invariance*, International Journal of General Systems, 1990, 17 (2), 249–275.
[20] KLIR G.J., YUAN B., *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
[21] KUCHTA D., *Fuzzy capital budgeting*, Fuzzy Sets and Systems, 2000, 111 (4), 367–385.
[22] LEE H., TANAKA H., *Lower and upper approximation models in interval regression using regression quantile techniques*, European Journal of Operational Research, 1999, 116 (3), 653–666.
[23] LIU Y.-K., LIU B., *Fuzzy random variables. A scalar expected value operator*, Fuzzy Optimization and Decision Making, 2003, 2 (2), 143–160.
[24] MOHAMED S., MCCOWAN A.K., *Modelling project investment decisions under uncertainty using possibility theory*, International Journal of Project Management, 2001, 19 (4), 231–241.
[25] RALESCU D.A., PURI M.L., *The concept of normality for fuzzy random variables*, Annals of Probability, 1985, 13, 1371–1379.
[26] REBIASZ B., *Fuzziness and randomness in investment project risk appraisal*, Computers and Operations Research, 2007, 34 (1), 199–210.
[27] SMETS P., *Constructing the pignistic probability function in a context of uncertainty*, [in:] M. Henrion, E.D. Schachter, L.N. Kanal, J.F. Lemmer (Eds.), *Uncertainty in Artificial Intelligence*, Vol. 5, North Holland, 1990, 29–39.
[28] SCHAFER G., *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, 1976.
[29] TANAKA H., LEE H., *Fuzzy linear regression combining central tendency and possibilistic properties*, Proc. 6-th IEEE International Conference on Fuzzy Sets, 1997, 1, Barcelona, Spain, 63–68.
[30] TANAKA H., LEE H., *Interval regression analysis by quadratic programming approach*, IEEE Transactions on Fuzzy Systems, 1998, 6 (4), 473–481.

[31] TANAKA H., WATADA J., *Possibilistic linear systems and their application to the linear regression model*, Fuzzy Sets and Systems, 1988, 27 (3), 275–289.
[32] TANAKA H., *Fuzzy data analysis by possibilistic linear models*, Fuzzy Sets and Systems, 1987, 24 (3), 363–375.
[33] TANAKA K., UEJIMA S., ASAI K., *Linear regression analysis with fuzzy model*, IEEE Transactions on Systems Man and Cybernetics, SMC-12, 1982 (6), 903–907.
[34] TRAN L., DUCKSTEIN L., *Multiobjective fuzzy regression with central tendency and possibilistic properties*, Fuzzy Sets and Systems, 2002, 130 (1), 21–31.
[35] WARD T.L., *Discounted fuzzy cash flows analysis*, Proc. Fall Industrial Engineering Conference, Institute of Industrial Engineers, Stanford, USA, 1985, 476–481.

Received 9 May 2014
Accepted 21 January 2015