Scientiae Mathematicae Japonicae

CONSTRUCTION OF A POSSIBILISTIC REGRESSION MODEL BASED ON POSSIBILITY GRADES WITH VAGUENESS AND RELATIONSHIP WITH PARAMETERS

Yoshiyuki Yabuuchi

Abstract. A possibilistic regression model is an interval-type model. An interval-type model intuitively helps us to understand the possibilities of the target system. The data distribution defines the possibility interval of the system, which may hinder our understanding of the analysis results. Improved models have reported using outlier problem approaches. We propose models to deal with the vagueness included in a possibility grade derived from a possibilistic regression model and samples. Unfortunately, the results obtained by the proposed models were not as expected. Then, the improved model was proposed to handle the vagueness included in possibility grades. The numerical example confirmed that the proposed model could eliminate the influence of unusual samples and describe the possibilities of a focal system. The paper reports the improved model and the results by using a numerical example.

1 Introduction The interval-type possibilistic regression model proposed by Tanaka and Watada [16], as used in this paper, includes all samples. An interval output illustrates the possibility distribution of a focal system. This interval type is rewritten in linear programming (LP), and can be obtained easily. Furthermore, there are various models [1, 4, 7, 12] using possibilistic regression in addition to the least-squares model proposed by Diamond [2, 3]. Fuzzy least squares based on a fuzzy random variable [9, 10] provides a lot of information. However, we use an interval type from the viewpoint of soft computing, because an interval model helps us to understand the analysis object intuitively.

An interval type illustrates the possibilities of an analyzed system by including all samples. The shape of a model is defined by that of the data distribution. For this reason, an interval type is susceptible to the shape of the data distribution. Therefore, processing of outliers for an interval type [14, 15], in which a model coincides with a focal system [5, 8, 11, 18, 19, 20, 21, 22, 23], a linguistic regression model [17], and so forth, are reported. We have proposed a model to deal with the vagueness included in a possibility grade derived from a possibilistic regression model and samples [24, 25]. The objectives of the proposed model are to remove the influence of unusual samples and describe the possibilities of a focal system so that it can be understood subjectively. Unfortunately, the results obtained by the proposed method were not as expected. That model is sometimes unable to remove the influence of unusual samples and distortion of the model. Therefore, a model dealing with the vagueness included in the possibility grade has been built [24, 25]. The proposed model made it possible to eliminate the influence of unusual samples and describe the possibilities of a focal system [26].

This paper is organized as follows. Section 2 briefly explains the interval type of the possibility regression model dealt with in this paper. Section 3 explains the proposed model to process vagueness included in possibility grades. In Section 4, we confirm the usefulness of the proposed model using a simple numerical example. Section 5 concludes this paper.

Key words and phrases. Fuzzy Regression Model, Fuzzy Number, Possibility Grade, Vagueness, Error.

2 Possibilistic Regression Model Consider a possibilistic regression equation using triangular fuzzy regression coefficients:

$$\boldsymbol{Y}_i = (a_0, c_0) + (a_1, c_1)x_{i1} + \cdots + (a_p, c_p)x_{ip} = (\boldsymbol{a}\boldsymbol{x}_i, \boldsymbol{c}|\boldsymbol{x}_i|). \tag{1}$$

The independent and dependent variables are $\boldsymbol{x}_i = (1, x_{i1}, \ldots, x_{ip})$ and y_i in samples $(\boldsymbol{x}_i, y_i)(i = 1, 2, \ldots, n)$. The center and width of the coefficient shown in equation (1) are $\boldsymbol{a} = (a_1, a_2, \ldots, a_p)$ and $\boldsymbol{c} = (c_1, c_2, \ldots, c_p)$, respectively. An output of equation (1) contains this dependent variable. In addition, the vagueness of this model, that is the widths, should be small. Therefore, a possibilistic regression model can be rewritten in the following LP:

$$\begin{aligned} &\text{min.} && F \\ &\text{s.t.} && \boldsymbol{a}\boldsymbol{x}_i - \boldsymbol{c}|\boldsymbol{x}_i| \leq y_i \leq \boldsymbol{a}\boldsymbol{x}_i + \boldsymbol{c}|\boldsymbol{x}_i|, i = 1, 2, \ldots, n. \end{aligned} \tag{2}$$

In equation (2), F employs various functions such as widths of coefficients, $F = \sum_j^p c_j$, and widths of forecasted values, $F = \sum_i^n \boldsymbol{c}|\boldsymbol{x}_i|$.

The regression coefficients are a symmetrical triangular fuzzy number, and the model describes the possibility distribution of the target system. The predicted value $\boldsymbol{Y}_{=}(Y_i^C, Y_i^W)$ in the independent variable x_i is the interval value with the center $Y_i^C = \boldsymbol{a}\boldsymbol{x}_i$ and the width $Y^W = \boldsymbol{c}|\boldsymbol{x}_i|$. The possibility grade $\mu(y_i, \boldsymbol{x}_i)$ is written as follows:

$$\mu(y_i, \boldsymbol{x}_i) = \max\left(0, 1 - \frac{|y_i - Y_i^C|}{Y_i^W}\right). \tag{3}$$

As shown by equation (3), the range of possibility grades is $[0, 1]$. When the regression coefficients are symmetric triangular fuzzy regression coefficients, their outputs are also symmetric triangles. The possibility grade is the maximum value 1 at the center of the distribution, and becomes the minimum value 0 when leaving the center. The conventional possibility regression model does not consider the possibility grade because it is a model with the least vagueness. On the other hand, the models we propose maximize the possibility grade. The model proposed in this paper deals with vagueness included in the possibility grade. For this reason, the proposed model can eliminate the influence of unusual samples and illustrate the possibility of the focal system. The next section describes the proposed model.

3 Possibilistic Regression Model with Vagueness in Possibility Grades Observed variables include various errors. Errors included in sample attribute values are discussed in statistics and probability, and many research results have been reported. For a possibility grade [6], research results dealing with grade fluctuations are reported using Type-2 fuzzy sets. However, the method using Type-2 fuzzy sets is more complicated than handling using Type-1 fuzzy sets. Therefore, we do not use Type-2 fuzzy sets in this work, and consider a method to easily handle the vagueness included in possibility grades.

Here, because attribute values contain an error, it is natural to think that possibility grades obtained from attribute values also contain an error. Therefore, although possibility grades can be obtained depending on a relationship between membership functions and samples, we assume that a grade has flexibility [24, 25].

In this paper, the proposed regression model handles samples with vagueness in the possibility grade to illustrate the possibility of the focal system. For that purpose, this section explains handling with samples and LP problems to obtain the proposed model.

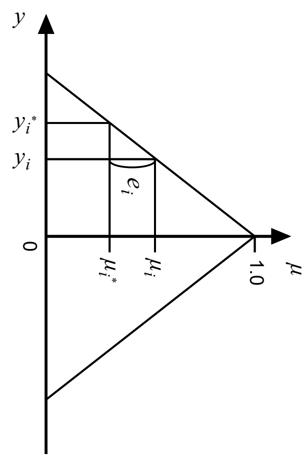


Figure 1: Vagueness included in a possibility grade

3.1 Dealing with Vagueness Including Possibility Grades The possibility grade of attribute value y_i is assumed as μ_i. That is, let us consider that possibility grades, μ_i, contain an error, e_i. At this time, as shown in Fig. 1, let the true possibility grade be μ_i^*. Then, the attribute value corresponding to the true possibility grade μ_i^* will be the value corresponding to y_i^* in Fig. 1. Let Y^C be the center of the membership function and Y^W be the width, then we can obtain the following:

$$y_i^* = y_i + e_i Y^W. \tag{4}$$

Then we replace y_i^* and y_i to find a possibility regression model.

A possibilistic regression model as shown by equation (2) explains the proposed method. A possibility grade μ_i of the ith sample contains an error e_i, and the following relationship holds with the true possibility grade μ_i^* that contains none of error:

$$\mu_i = \mu_i^* + e_i. \tag{5}$$

Here, because a possibility grade takes values of $[0, 1]$, e_i also takes values of $[-1, 1]$.

3.2 Formulation of Model Handling Vagueness Included in Possibility Grades From the above, the inclusion relation between y_i and a model output $\boldsymbol{Y}_i = (\boldsymbol{a}\boldsymbol{x}_i, \boldsymbol{c}|\boldsymbol{x}_i|)$ are as follows:

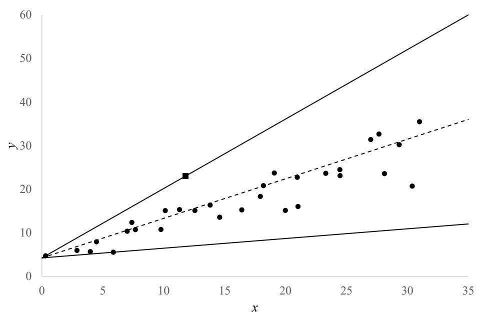
$$\boldsymbol{a}\boldsymbol{x}_i - \boldsymbol{c}|\boldsymbol{x}_i| \leq y_i + e_i\boldsymbol{c}|\boldsymbol{x}_i| \leq \boldsymbol{a}\boldsymbol{x}_i + \boldsymbol{c}|\boldsymbol{x}_i|, i = 1, 2, \ldots, n. \tag{6}$$

As a result, equation (2) can be rewritten as follows:

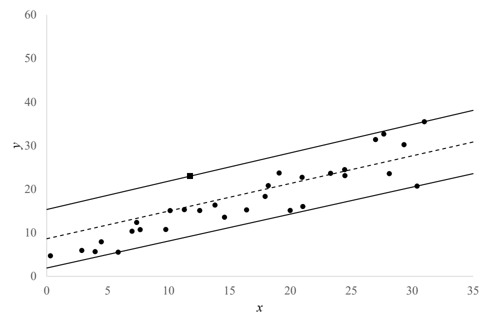
$$\begin{array}{ll} \text{min.} & F \\ \text{s.t.} & \boldsymbol{a}\boldsymbol{x}_i - \boldsymbol{c}|\boldsymbol{x}_i| \leq y_i + e_i\boldsymbol{c}|\boldsymbol{x}_i| \leq \boldsymbol{a}\boldsymbol{x}_i + \boldsymbol{c}|\boldsymbol{x}_i|, \\ & |e_i| \leq \varepsilon, i = 1, 2, \ldots, n. \end{array} \tag{7}$$

Here, ε is a parameter that specifies the range of vagueness included in the possibility grade. As possibility grades are real numbers, ε is also a real number. Furthermore, the objective function F uses an appropriate function according to the data, similar to the conventional possibilistic regression model.

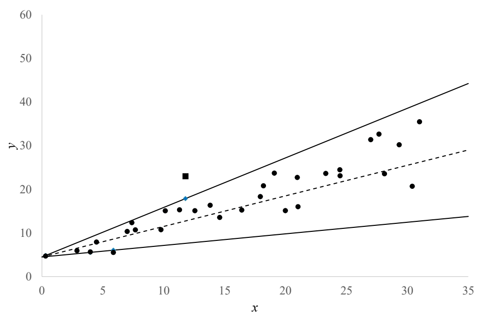
Using only this, the influence of unusual samples can be removed. We confirm this concretely using a numerical example.



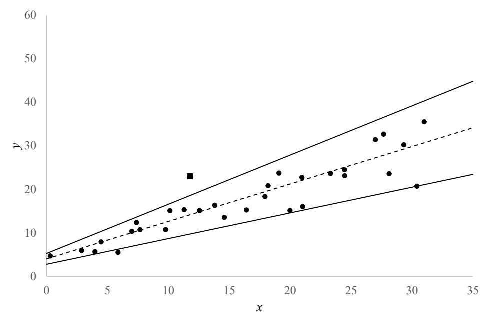
(a) Model 1 denoted by Eq.(9)



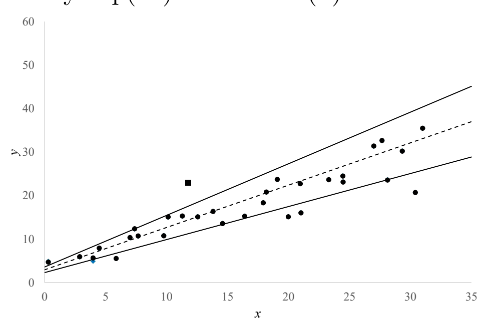
(b) Model 2 denoted by Eq.(10)



(c) Model 3 denoted by Eq.(11)



(d) Model 4 denoted by Eq.(12)



(e) Model 5 denoted by Eq.(13)

Figure 2: Obtained models in the numerical example

4 Numerical Example In this section, the same numerical example as in [25] is used. The numerical example adds errors with probability to the two variables, x and y, in the relationship of $y = x$. In addition, samples contain one unusual sample, and the model parameter constraint is set to $|e_i| \leq \varepsilon = 1$. In the numerical example, the following possibilistic regression equation will be found:

$$\boldsymbol{Y} = (a_0, c_0) + (a_1, c_1)x. \tag{8}$$

We obtain model 1 with $F = \sum_j^p c_j$ as the objective function of the interval-type possibilistic regression model shown by equation (2), and model 2 with the objective function $F = \sum_i^n \boldsymbol{c}|\boldsymbol{x}_i|$. In addition to models 3 and 4, which add the vagueness of grades to models 1 and 2, we also obtain model 5 that considers the vagueness of possibility grades to the model proposed by Yabuuchi [24].

The outputs of models 1 to 5 are denoted as $\boldsymbol{Y}_1$ to $\boldsymbol{Y}_5$, respectively. The five models

Table 1: Features of obtained models in the numerical examples

	Model 1	Model 2	Model 3	Model 4	Model 5
Sum of widths of regression coefficients	0.685	6.733	0.435	1.513	0.857
Sum of widths of forecasted values	344.521	432.172	437.480	348.199	254.751
Sum of possibility grades derived from the model and samples	21.473	15.927	16.420	17.003	14.018
Sum of possibility grades for widths of forecasted values	2.542	2.287	3.077	3.569	4.244
Outside samples of intervals			3	3	7

are as follows:

$$\boldsymbol{Y}_1 = (4.232, 0) + (0.908, 0.685)x, \tag{9}$$
$$\boldsymbol{Y}_2 = (8.627, 6.717) + (0.634, 0.016)x, \tag{10}$$
$$\boldsymbol{Y}_3 = (4.495, 0) + (0.701, 0.435)x, \tag{11}$$
$$\boldsymbol{Y}_4 = (4.039, 1.244) + (0.859, 0.270)x, \tag{12}$$
$$\boldsymbol{Y}_5 = (2.944, 0.643) + (0.973, 0.214)x. \tag{13}$$

The least squares is as follows:

$$\boldsymbol{Y}_S = 4.316 + 0.827x. \tag{14}$$

In Fig. 2, the original sample is rounded, and the values converted by equation (4) are indicated by a rhombus. Fig. 2 shows that the models handling vagueness included in possibility grades are not distorted. However, the value of the constant term seems to be large owing to the influence of a specific sample. For this reason, the center of model 3 has a small inclination. The center of model 3 is similar to model 1, the constant term is slightly larger, and the inclination seems to be smaller. On the other hand, in model 5, the centers of the model and the data distribution almost coincide, the width of the forecasted value becomes small, and the possibility of the system can be understood intuitively.

The information obtained from these models is listed in Table 1. The possibility grade is large when the sample is close to the center, so the model with the small width of the interval has the small sum of possibility grade. For this reason, the sum of the possibility grades of model (9) has the maximum value, and that of model (13) has the minimum value. However, in Table 1, the sum of the possibility grade for the width of the forecasted value is opposite to the sum of the possibility grades. This is, the sum of the possibility grade for the width of the forecasted value of the model (9) has the second smallest value, and that of the model (13) is the maximum value.

From the above, we can summarize the features of the proposed model that consider the vagueness included in the possibility grade. First, it was subjectively perceptible that the model describes the data distribution. Second, the influence of the outlier was eliminated, and a mode without distortion in shape was obtained.

In addition, its effect was improved by using the model in conjunction with the model proposed by Yabuuchi [24] that maximizes the sum of the possibility grade for the width of the forecasted value.

In the above, the parameter ε of models 3–5 has been set to 1 because the range of possibility grades is $[0, 1]$. On the other hand, because models 1 and 2 are conventional models, this parameter was not used. Here, the models are obtained by using 0.5, 1.0,

Table 2: The coefficients of the three models using $\varepsilon = \{0.5, 1.0, 1.5, 2.0\}$

		$\varepsilon = 0.5$	$\varepsilon = 1.0$	$\varepsilon = 1.5$	$\varepsilon = 2.0$
Model 3	$\boldsymbol{A}_0$	(4.232, 0)	(4.495, 0)	(4.232, 0)	(4.363, 0)
	$\boldsymbol{A}_1$	(0.908, 0.457)	(0.701, 0.435)	(0.908, 0.274)	(0.891, 0.230)
Model 4	$\boldsymbol{A}_0$	(3.697, 1.137)	(4.039, 1.244)	(3.990, 2.169)	(4.132, 1.495)
	$\boldsymbol{A}_1$	(0.956, 0.359)	(0.859, 0.270)	(0.808, 0.138)	(0.786, 0.191)
Model 5	$\boldsymbol{A}_0$	(2.866, 0.709)	(2.944, 0.643)	(2.944, 0.643)	(2.944, 0.643)
	$\boldsymbol{A}_1$	(1.056, 0.475)	(0.973, 0.214)	(0.973, 0.214)	(0.973, 0.214)

Table 3: Features of the three models using $\varepsilon = \{0.5, 1.0, 1.5, 2.0\}$

		$\varepsilon = 0.5$	$\varepsilon = 1.0$	$\varepsilon = 1.5$	$\varepsilon = 2.0$
Model 3	Index 1	459.361	437.480	275.617	231.563
	Index 2	18.403	16.420	13.721	12.253
	Index 3	3.151	3.077	3.731	3.912
Model 4	Index 1	431.065	348.199	273.368	284.759
	Index 2	18.044	17.003	13.359	13.483
	Index 3	3.175	3.569	3.326	3.393
Model 5	Index 1	521.995	254.751	254.751	254.751
	Index 2	18.663	14.018	14.018	14.018
	Index 3	2.928	4.244	4.244	4.244

Index 1: Sum of widths of forecasted values

Index 2: Sum of possibility grades derived from the model and samples

Index 3: Sum of possibility grades to widths of forecasted values

1.5, and 2.0 as the parameter ε, and the characteristics are confirmed. Table 2 lists the coefficients obtained by the models. Even if the parameter is changed, the center of the models does not change significantly. In addition, the width of the model decreased by increasing the value of the parameter. Furthermore, in model 5, the same model was obtained when $\varepsilon \geq 1.0$.

Table 3 lists the features of the model obtained by changing the parameter ε. When ε was changed from 1.5 to 2.0, the possibility grade of models 3 and 4 did not change significantly. In particular, when ε was increased, the width of the predicted value and the value of the possibility grade became smaller. However, index 3, which divided the possibility grade by the width of the predicted value, increased. In general, if the width of the predicted value is small, the sum of the possibility grade is also small. Then, the relationship between indices 1 and 2 can understand. Index 3 has a large value when the samples are near the center of the possibility interval. Therefore, increasing the value of the parameter ε gathers samples near the center of the possibility interval.

As described above, the width of the solution search space is increased by increasing the value of the parameter, and an unexpected solution is obtained from LP. Although the upper limit of the number of samples processed with fuzziness possibility grade was limited, index 1 of model 4 is larger for $\varepsilon = 2.0$ than for $\varepsilon e = 1.5$. In addition, model 5 uses possibility grades for the objective function. For this reason, model 5 might not be influenced by the parameter ε more than necessary.

To confirm these results, the models are shown in Figs. 3–5. In Figs. 3–5, the boundaries of the model when ε is changed to 0.5, 1.0, 1.5, and 2.0 are shown by a dashed-dotted line, dashed-two dotted line, dashed line, and dotted line, respectively. The features listed in Tables 2 and 3 are confirmed by the results in Figs. 3–5.

The statistical model emphasizes samples away from the center of gravity of the

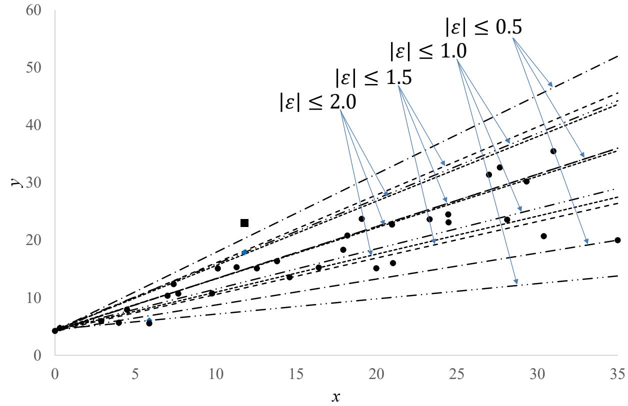


Figure 3: Model 3 using $\varepsilon = \{0.5, 1.0, 1.5, 2.0\}$

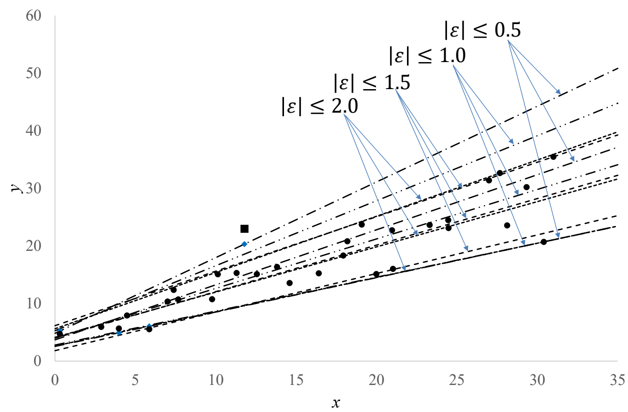


Figure 4: Model 4 using $\varepsilon = \{0.5, 1.0, 1.5, 2.0\}$

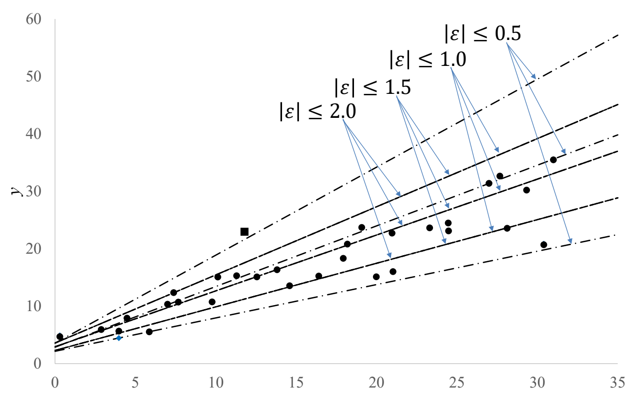


Figure 5: Model 5 using $\varepsilon = \{0.5, 1.0, 1.5, 2.0\}$

data distribution. On the other hand, samples away from the center of the interval model distort the model. We also found that the parameter of the proposed model adjusts the influence of samples away from the center of this model.

5 Conclusion In this paper, we have proposed a possibility regression model considering the vagueness included in possibility grades. Then, the usefulness of the proposed model was confirmed by using the numerical example with outliers. The proposed technique improved the forecast accuracy of models and eliminated the influence of unusual samples. In addition, by adjusting the parameter ε, it is possible to adjust the influence of samples away from the center of the model.

Furthermore, it has been improved by using it in conjunction with the model proposed by Yabuuchi to maximize the sum of the possibility grade for the width of the forecasted value. Finally, the proposed model only arranges the constraints as shown in equation (6), and sufficient results have been obtained.

References

[1] R. Coppi, P. D'Urso, P. Giordani, A. Santoro: Least squares estimation of a linear regression model with LR Fuzzy response. *Comput. Statist. Data Anal.* 51(1), 267–286 (2006)

[2] P. Diamond: Fuzzy Least Squares. *Inform. Sci.*, 46(3), 141–157 (1988)

[3] P. Diamond: Least squares and maximum likelihood regression for fuzzy linear models. In J. Kacprzyk and M. Fedrizzi (Eds.), *Fuzzy Regression Analysis*, Omnitech Press, 137–151 (1992)

[4] P. D'Urso and T. Gastaldi: A least-squares approach to fuzzy linear regression analysis. *Comput. Statist. Data Anal.*, 34(4), 427–440 (2000)

[5] P. Guo and H. Tanaka: Fuzzy DEA: a perceptual evaluation method. *Fuzzy Sets and Systems*, 119(1), 149–160 (2001)

[6] A. Honda and Y. Okazaki: Identification of Fuzzy Measures with Distorted Probability Measures. *J. Advanced Computational Intelligence and Intell. Informatics*, 9(5), 467–76 (2005)

[7] D.H. Hong, C. Hwang, C. Ahn: Ridge estimation for regression models with crisp inputs and Gaussian fuzzy output. *Fuzzy Sets and Systems*, 142(2), 307–319 (2004)

[8] M. Inuiguchi and T. Tanino: Interval Linear Regression Methods Based on Minkowski Difference: A Bridge between Traditional and Interval Linear Regression Models. *Kybernetika*, 42(4), 423–440 (2006)

[9] H. Kwakernaak: Fuzzy random variables-I. definitions and theorems, *Inform. Sci.*, 15(1), 1–29 (1978)

[10] H. Kwakernaak: Fuzzy random variables-II. Algorithms and examples for the discrete case, *Inform. Sci.*, 17(3), 253–278 (1979)

[11] H. Lee and H. Tanaka: Upper and lower approximation models in interval regression using regression quantile techniques. *Eur. J. Oper. Res.*, 116(3), 653–666 (1999)

[12] M. Modarres, E. Nasrabadi and M.M. Nasrabadi: Fuzzy linear regression model with least square errors. *Appl. Mathematics and Computation*, 163(2), 977–989 (2005)

[13] M. L. Puri and D. A. Ralescu: The Concept of Normality for Fuzzy Random Variables, *Ann. Probab.*, 13(4), 1373–1379 (1985)

[14] A.A. Ramli, J. Watada and W. Pedrycz: Real-time fuzzy regression analysis: A convex hull approach. *Eur. J. Oper. Res.*, 210(3), 606–617 (2011)

[15] A.A. Ramli, J. Watada and W. Pedrycz: A Combination of Genetic Algorithm-based Fuzzy C-Means with a Convex Hull-based Regression for Real-Time Fuzzy Switching Regression Analysis: Application to Industrial Intelligent Data Analysis. *IEEJ Trans. Elect. Electron. Eng.*, 9(1), 71–82 (2014)

[16] H. Tanaka and J. Watada: Possibilistic Linear Systems and Their Application to The Linear Regression Model. *Fuzzy Sets and Systems*, 27(3), 275–289 (1988)

[17] J. Watada and W. Pedrycz: A Fuzzy Regression Approach to Acquisition of Linguistic Rules. In W. Pedrycz, A. Skowron and V. Kreinovich (Eds.), *Handbook of Granular Computing*, Wiley, 719–740 (2008)

[18] Y. Yabuuchi and J. Watada: Model Building Based on Central Position for a Fuzzy Regression Model. *Proc. Czech-Japan Seminar 2006*, 114–119 (2006)

[19] Y. Yabuuchi and J. Watada: Fuzzy Regression Model Building through Possibility Maximization and Its Application. *Innovative Computing, Inform. Control Express Lett.*, 4(2), 505–510 (2010)

[20] Y. Yabuuchi and J. Watada: Fuzzy Robust Regression Model by Possibility Maximization. *J. Advanced Computational Intelligence and Intell. Informatics*, 15(4), 479–484 (2011)

[21] Y. Yabuuchi: Japanese Economic Analysis by a Fuzzy Regression Model building through Possibility Maximization. *Proc. 6th Int. Conf. Soft Computing and Intelligent Syst., and 13th Int. Symp. Advanced Intelligent Syst.*, 1772–1777 (2012)

[22] Y. Yabuuchi and J. Watada: Fuzzy Robust Regression Model building through Possibility Maximization and Analysis of Japanese Major Rivers. *Innovative Computing, Inform. Control Express Lett.*, 9(4), 1033–1041 (2015)

[23] Y. Yabuuchi: Centroid-Based Fuzzy Robust Regression Model. *Innovative Computing, Inform. Control Express Lett.*, 9(12), 3299–3306 (2015)

[24] Y. Yabuuchi: The Difference between the Formulations of Possibilistic Robust Regression Model. *Proc. Joint 17th World Congr. Int. Fuzzy Syst. Assoc. and 9th Int. Conf. Soft Computing Intelligent Syst.*, N.P. (6 pages) (2017)

[25] Y. Yabuuchi: Possibility Grades with Vagueness in Fuzzy Regression Models. *Proc. KES 2017*, 1470–1478 (2017)

[26] Y. Yabuuchi: Construction of a Possibilistic Regression Model based on Possibility Grades with Vagueness. *Proc. Joint Conf. 26th Forum of Interdisciplinary Mathematics and 14th Int. Symp. Manage. Eng.*, 32–41 (2017)

Shimonoseki City Univerisity

2-1-1 Daigaku-cho, Shimonoseki, Yamaguchi 751-8510, Japan

Email: yabuuchi@shimonoseki-cu.ac.jp