Journal, Advances in Mathematical & Computational Sciences
Vol. 11 No. 1, March, 2023
www.isteams.net/mathematics-computationaljournal

Error Estimation Using Possibilistic Fuzzy Geometric Regression

Onoghojobi, B. & Ibrahim, A.
Department of statistics
Federal University Lokoja
Lokoja, Kogi state, Nigeria.

ABSTRACT

In this study, an attempt is made to reduce or minimize the error of a fuzzy data using non-linear geometric regression. A model is derived and error of a non-linear geometric data set is estimated using the Tanaka possibilistic approach as it is even more compact and does not gives room for outliers when data is being fitted. The method is generalized for the linear programming problem and the process is demonstrated with a numerical example in which the data is converted to crisp before the error is then estimated and thereby making comparison with the traditional linear regression. It is concluded from the numerical illustration that the model provided a better fit to the data than the traditional linear regression model.

Keywords: Fuzzy geometry, Optimization technique, Possibilistic approach, Crisp linear regression.

Onoghojobi, B. & Ibrahim, A. (2023): Error Estimation Using Possibilistic Fuzzy Geometric Regression. Journal of Advances in Mathematical & Computational Science. Vol. 11, No. 1. Pp 75-90. Available online at www.isteams.net/mathematics-computationaljournal. dx.doi.org/10.22624/AIMS/MATHS/V11N1P6

1. INTRODUCTION

In statistics, the **non-linear regression** is a form of regression analysis in which observational data are modelled by a function which is a non-linear combina- tion of the model parameters and depends on one or more independent variables[1]. The data are fitted by method of successive approximations. in non-linear regression, a statistical model of the form $y\ f(x, \beta)$ relates a vector of inde- pendent variables x, and its associated observed dependent variables, y. The function f is non-linear in the components of the vector of parameters β, but otherwise arbitrary. Thus

$$f(x, \beta) = \frac{\beta_1 x}{\beta_2 + x} \quad [2]$$

This function is non-linear because it cannot be expressed as a linear com-bination of two β's. Error estimation as the word connotes, is the process of using statistical techniques to determine the error of a statistical model.While fuzzy regression is a fuzzy variation of classical regression[2]. It is divided into Tanaka possibilistic (LPP) approach and Celmins & Diamond least square approach also known asthe distance approach but our interest in this study is the possibilistic approach (Tanaka) for error estimation in the geometric model [3]. The possibilistic fuzzy geometric regression is a statistical method that combines the fuzzy logic and non-linear regression techniques to model data that is char- acterized by uncertainty and imprecission[4].

In this study,we wish to estimate the error of non-linear geometric data set and we may employ the knowledgeof triangular fuzzy number (TFN) as we will be using the Tanaka[5] approach. This method assumes that components of the established membership function of the triangular fuzzy are symmetric i.e the left spread equals the right spread, where the output may be written as;

$$\tilde{Y} = f(X, \tilde{A}) = \tilde{A}_0 + \tilde{A}_1 X_1 + \tilde{A}_2 X_2 + \cdots + \tilde{A}_n X_n = \tilde{A}_0 + \sum_{i=1}^{n} \tilde{A}_i X_i \cdots\cdots\cdots (1)$$

2. ERROR ESTIMATION OF THE ORDINARY REGRESSION

Ordinary Least Square regression, often called linear regression, is a common technique for estimating coefficients of linear regression equations which de- scribes the relationship between one or more independent quantitative variables and a dependent variable (simple or multiple linear regression)[6]. Least squares stand for the minimum square error (SSE). Maximum likelihood and generalized method of moments estimator are alter- native approaches to OLS. In the case of a model with p explanatory variables, the OLS regression model writes

$$Y = \beta_0 + \sum_{j=1\ldots p} \beta_j X_j + \varepsilon$$

where Y is the dependent variable, β_0, is the intercept of the model, X_j corre- sponds to the j^{th} explanatory variable of the model (j = 1 to p) and ε is the random error with expectation 0 and variance σ^2.In the case where there are n observations, the estimation of the predicted value of the dependent variable Y for the ith observation is given by:

$$y_i = \beta_0 + \sum_{j=1\ldots p} \beta_j X_i j.$$

Error Term: An error term is a residual variable produced by a statistical or mathematical model, which is created when the model does not fully represent the actual relationship between the independent variables and the dependent variables[7]. As a result of this incomplete relationship, the error term is the amount at which the equation may differ during empirical analysis. The er- ror term is also known as the residual, disturbance, or remainder term, and is variously represented in models y the letters e, ε, or u.

3. GEOMETRIC REGRESSION

A regression model provides a function that descries the relationship between one or more independent variables and a response, dependent, or target vari- able. For example the relationship between height and weight may be described by a linear regression model[2]. In the geometric regression, the mean of the y is determined by the exposure time t and a set of k regressor variables (the x's). The expression relating these quantities is;

$$\mu_i = exp(In(t_i) + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki})$$

Often $x_1 \equiv 1$, In which case β_1 is called the intercept. The regression coefficients $\beta_1, \beta_2, \ldots, \beta_k$ are unknown parameters that are estimated from a set of data[8]. Their estimates are symbolized as $b_1, b_2, \ldots, b_k$ using this notation, the fundamental geometric regression model for an observation is written as

$$pr(Y = y_i|\mu_i) = \frac{\Gamma(y_i + 1)}{\Gamma(y_i + 1)} \left(\frac{1}{1 + \mu_i}\right)^1 \left(\frac{\mu_i}{1 + \mu_i}\right)^{y_i}$$

3.1 Possibilistic Fuzzy Geometric Regression

The possibilistic fuzzy geometric regression is a statistical method that combines the fuzzy logic and non-linear regression techniques to model that is character- ized by uncertainty and imprecision[3]. In this paper, we attempt to estimate the error of a possibilistic fuzzy geometric regression with a symmetric triangular fuzzy number(STFN) coefficients using the possibilistic approach where the data is first transformed into a fuzzy set membership fuction(Using the basic knowledge of triangular fuzzy number).

4. THE LPP FORMULATION

The LPP technique was first introduced in 1930 by Russian mathematician Leonid Kantorovich[9] in the field of manufacturing schedules and by American economist Wassily Leontief[9] in the field of economics. These techniques were heavily adopted to solve problems related to transportation, scheduling, allo- cation of resources, etc. The elements of a basic linear programming problem includes:

- Decision Variables: They are unknown quantities that are expected to be estimated as an output of the LPP solution.
- Objective Function: All linear programming problems aim to either maximize or minimize some numerical value.
- Objective Function Coefficient: Known as the amount y which the objective function value would change when one unit of a decision variable is altered.

Suppose we wish to establish the membership function which must lie between the interval [0,1]before proceeding to minimize the spreads subjected to the constraints obtained from the membership functions.

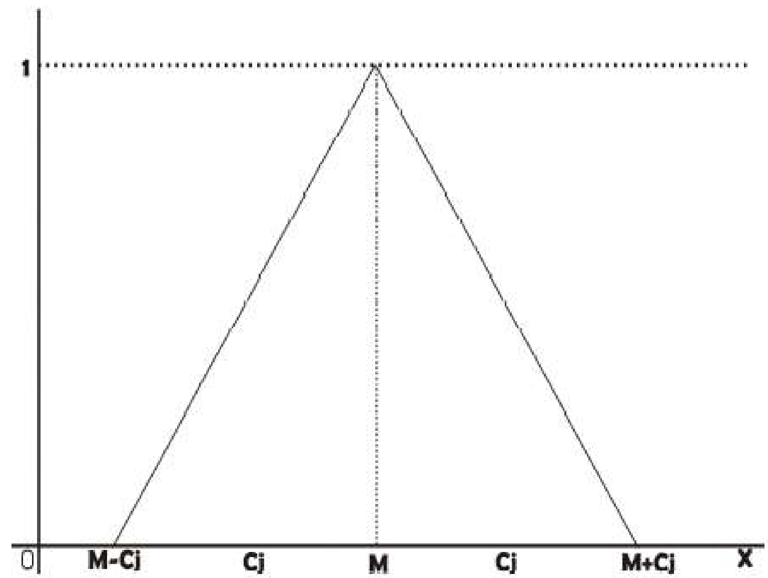

Fig.1 Triangular Fuzzy Coefficients

The diagram above is called the triangular fuzzy number. As indicated, the features of the TFN are its mode or center(M), its left and right spreads(M-cj, M+cj) and its supports (cj's). when the two spreads are equal,the triangular fuzzy number is known as a symmetric triangular fuzzy number (STFN). The basic idea of Tanaka approach often referred to as possibilistic regression is to minimize the fuzziness of the model by minimizing the total spread of the fuzzy coefficients subjected to the constraints[10], which in our own case will be formed from the data set linearized using geometric transformation.
The purpose of this article is to derive a model ad estimate the error of a non- linear data set using Tanaka possibilistic approach.

where:
m-cj= left spread
m+cj= right
spread m= center
cj= supports

from the basic knowledge of triangular fuzzy, we can form a membership function $\mu^{(x)}$ such that the spread is subjected to the constraints obtained from $\_{(x)A}$.
From _g(1),

$$\mu_A^{(x)} = \begin{cases} 0, x < m - cj \\ \frac{x-(m-cj)}{m-(m-cj)},\ m - cj \leq x \leq m \\ \frac{(m+cj)-x}{(m-cj)-m},\ m \leq x \leq m + cj \\ \\ 0, x > m + cj \end{cases}$$

Simplifying the denominators, we have:

$$\mu_A^{(x)} = \begin{cases} 0, x < m - cj \\ \frac{x-(m-cj)}{cj},\ m - cj \leq x \leq m \\ \frac{(m+cj)-x}{cj},\ m \leq x \leq m + cj \\ \\ 0, x > m + cj \end{cases}$$

Recall from (1)

we can also transform the center and spreads of the TF even more compactly such that:

$$m = m_0 + \sum_{j=1}^{n} m_j \text{ and } cj = c_0 + \sum_{j=1}^{n} cj|xj| \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (2)$$

we substitute equation (2) into the membership function $\mu_A^{(x)}$ above.

$$\mu_A^{(y)} = \begin{cases} 0, y < [m_0 + \sum_{j=1}^{n} m_j] - [c_0 + \sum_{j=1}^{n} c_j|x_j|] \\ \frac{y-[(m_0+\sum_{j=1}^{n} m_j|x_j|)-(c_0+\sum_{j=1}^{n} c_j|x_j|)]}{c_0+\sum_{j=1}^{n} c_j|x_j|}, \ [m_0 + \sum_{j=1}^{n} m_j|x_j|] - [c_0 + \sum_{j=1}^{n} c_j|x_j|] \leq y \leq m_0 + \sum_{j=1}^{n} m_j|x_j| \\ \frac{[(m_0+\sum_{j=1}^{n} m_j|x_j|)+(c_0+\sum_{j=1}^{n} c_j|x_j|)]-y}{c_0+\sum_{j=1}^{n} c_j|x_j|}, \ [m_0 + \sum_{j=1}^{n} m_j|x_j|] \leq y \leq [m_0 + \sum_{j=1}^{n} m_j|x_j|] + [c_0 + \sum_{j=1}^{n} c_j|x_j|] \\ 0, y > [m_0 + \sum_{j=1}^{n} m_j|x_j|] + [c_0 + \sum_{j=1}^{n} c_j|x_j|] \end{cases}$$

we introduce the h-certain factor into the membership function.

$$\frac{y-[(m_0+\sum_{j=1}^{n} m_j|x_j|)-(c_0+\sum_{j=1}^{n} c_j|x_j|)]}{c_0+\sum_{j=1}^{n} c_j|x_j|} \geq h \cdots\cdots (3)$$

$$\frac{[(m_0+\sum_{j=1}^{n} m_j|x_j|)+(c_0+\sum_{j=1}^{n} c_j|x_j|)]-y}{c_0+\sum_{j=1}^{n} c_j|x_j|} \geq h \cdots\cdots (4)$$

from equation (3),

cross multiply and make y the subject of formula

$$y - [(m_0 + \sum_{j=1}^{n} m_j|x_j|) - (c_0 + \sum_{j=1}^{n} c_j|x_j|)] \geq h(c_0 + \sum_{j=1}^{n} c_j|x_j|)$$

$$y \geq h(c_0+) + [(m_0 + \sum_{j=1}^{n} m_j|x_j|) - (c_0 + \sum_{j=1}^{n} c_j|x_j|)]$$

$$y \geq hc_0 + h\sum_{j=1}^{n} c_j|x_j| + m_0 + \sum_{j=1}^{n} m_j|x_j| - c_0 + \sum_{j=1}^{n} c_j x_j$$

$$y \geq hc_0 - c_0 + h\sum_{j=1}^{n} c_j|x_j| - \sum_{j=1}^{n} c_j|x_j| + m_0 + \sum_{j=1}^{n} m_j|x_j|$$

$$y \geq c_0(h-1) + \sum_{j=1}^{n} c_j|x_j|(h-1) + m_0 + \sum_{j=1}^{n} m_j|x_j|$$

$$y \geq (h-1)[c_0 + \sum_{j=1}^{n} c_j|x_j|] + m_0 + \sum_{j=1}^{n} m_j|x_j|$$

$$\Rightarrow (h-1)[c_0 + \sum_{j=1}^{n} c_j|x_j|] + m_0 + \sum_{j=1}^{n} m_j|x_j| \leq y \cdots\cdots (5)$$

similarly from equation (4),

cross multiply and make y the subject of formula also

$$[(m_0 + \sum_{j=1}^{n} m_j|x_j|) + (c_0 + \sum_{j=1}^{n} c_j|x_j|)] - y \geq h(c_0 + \sum_{j=1}^{n} c_j|x_j|)$$

$$-y \geq h(c_0 + \sum_{j=1}^{n} c_j|x_j|) - [(m_0 + \sum_{j=1}^{n} m_j|x_j|) + (c_0 + \sum_{j=1}^{n} c_j|x_j|)]$$

$$y \leq -h(c_0 + \sum_{j=1}^{n} c_j|x_j|) + [(m_0 + \sum_{j=1}^{n} m_j|x_j|) + (c_0 + \sum_{j=1}^{n} c_j|x_j|)]$$

$$y \leq -hc_0 - h\sum_{j=1}^{n} c_j|x_j| + m_0 + \sum_{j=1}^{n} m_j|x_j| + c_0 + \sum_{j=1}^{n} c_j|x_j|$$

$$y \leq c_0 - hc_0 + \sum_{j=1}^{n} c_j|x_j| - h\sum_{j=1}^{n} c_j|x_j| + m_0 + \sum_{j=1}^{n} m_j|x_j|$$

$$y \leq c_0(1-h) + \sum_{j=1}^{n} c_j|x_j|(1-h) + m_0 + \sum_{j=1}^{n} m_j|x_j|$$

$$y \leq (1-h)[c_0 + \sum_{j=1}^{n} c_j x_j] + m_0 + \sum_{j=1}^{n} m_j x_j$$

$$\Rightarrow (1-h)[c_0 + \sum_{j=1}^{n} c_j|x_j|] + m_0 + \sum_{j=1}^{n} m_j|x_j| \cdots\cdots\cdots\cdots (6)$$

Hence,our newly generated constraints are equation (5) & (6)

$$(1-h)[c_0 + \sum_{j=1}^{n} c_j|x_j|] + m_0 + \sum_{j=1}^{n} m_j|x_j| \leq y$$

$$(1-h)[c_0 + \sum_{j=1}^{n} c_j|x_j|] + m_0 + \sum_{j=1}^{n} m_j|x_j| \geq y$$

By minimizing the spreads,

$$Z = min[c_0 + \sum_{j=1}^{n} c_j|x_j|] \cdots\cdots\cdots\cdots (7)$$

Now, let us consider a non-linear geometric data set by first linearizing and then substituting into the spread and constraints above.

4.1. Derivation and numerical illustration of Geometric fuzzy regression.

We wish to transform a fuzzy data which is characterized by a high level of uncertainty. In this data set, the explanatory variables (x) cannot be used to predict the response variable (y). However, the transformation is illustrated below.

log(y) = log(a) + βlog(x)

let,

$log(y) = Y^*$,

$log(a) = \alpha$

$log(x) = X^*$, so that,

$Y^* = \alpha + \beta X^*$

$$\Rightarrow Y^* = \alpha + \sum_{i=1}^{n} \beta_i X_i$$

Table 1: Geometric transformation of fuzzy values

X	Y	log(X)	log(Y)
1.37	5285.55	0.3148	8.5727
2.75	19.02	1.0116	2.9455
3.82	1.77	1.3403	0.5709
2.78	6.44	1.0225	1.8625
1.78	5.41	0.5481	1.6882
3.46	23.44	1.2413	3.1455
3.36	562.34	1.2119	6.3321

Recall:

Our aim is to obtain an objective function and a constraint which we would resolve into a linear programming problem. For the constraints, we shall sum all X values characterized by uncertainty in table(1) above.

$$Z = min[c_0 + \sum_{j=1}^{n} c_j |x_j|]$$

$$\Sigma xj = 1.37 + 2.75 + 3.82 + 2.78 + 1.78 + 3.46 + 3.36$$

$$\Rightarrow \Sigma xj = 19.32$$

Hence,

$Z = min[c_0 + 19.32cj]$, where j=1

For the constraints, we introduce the h-certain factor which balances the spreads of the triangular fuzzy number. Now, let the h-certain factor (h=0.1). So that,

$$(h-1)[c_0 + \sum_{j=1}^{n} c_j|x_j|] + m_0 + \sum_{j=1}^{n} m_j|x_j| \leq y$$

$$(1-h)[c_0 + \sum_{j=1}^{n} c_j|x_j|] + m_0 + \sum_{j=1}^{n} m_j|x_j| \geq y$$

Substituting the value of h into the constraint

$$(0.1-1)[c_0 + c_1x_1] + m_0 + m_1x_1 \leq y$$
$$(1-0.1)[c_0 + c_1x_1] + m_0 + m_1x_1 \geq y$$

$$-0.9(c_0 + c_1x_1) + m_0 + m_1x_1 \leq y$$
$$0.9(c_0 + c_1x_1) + m_0 + m_1x_1 \geq y$$

Expanding the brackets,

$$m_0 + m_1x_1 - 0.9c_0 - 0.9c_1x_1 \leq y \cdots\cdots\cdots (8)$$

$$m_1x_1 + 0.9c_0 + 0.9c_1x_1 + m_0 \geq y \cdots\cdots\cdots (9)$$

Then we substitute the values for the transformed (x's) and (y's) from table(1) above into the constraints i.e equation 8 & 9.

$$m_0 + m_1(0.3148) - 0.9c_0 - 0.9(0.3148)c_1 \leq 3.1544$$
$$m_0 + m_1(0.3148) + 0.9c_0 + 0.9(0.3148)c_1 \geq 3.1544$$

$$m_0 + m_1(1.0116) - 0.9c_0 - 0.9(1.0116)c_1 \leq 6.3321$$
$$m_0 + m_1(1.0116) + 0.9c_0 + 0.9(1.0116)c_1 \geq 6.3321$$

$$m_0 + m_1(1.3403) - 0.9c_0 - 0.9(1.3403)c_1 \leq 8.5727$$
$$m_0 + m_1(1.3403) + 0.9c_0 + 0.9(1.3403)c_1 \geq 8.5727$$

$$m_0 + m_1(1.0225) - 0.9c_0 - 0.9(1.0225)c_1 \leq 1.8625$$
$$m_0 + m_1(1.0225) + 0.9c_0 + 0.9(1.0225)c_1 \geq 1.8625$$

$$m_0 + m_1(0.5481) - 0.9c_0 - 0.9(0.5481)c_1 \leq 1.6882$$
$$m_0 + m_1(0.5481) + 0.9c_0 + 0.9(0.5481)c_1 \geq 1.6882$$

$$m_0 + m_1(1.2413) - 0.9c_0 - 0.9(1.2413)c_1 \leq 0.5709$$
$$m_0 + m_1(1.2413) + 0.9c_0 + 0.9(1.2413)c_1 \geq 0.5709$$

$$m_0 + m_1(1.2119) - 0.9c_0 - 0.9(1.2119)c_1 \leq 2.9454$$
$$m_0 + m_1(1.2119) + 0.9c_0 + 0.9(1.2119)c_1 \geq 2.9454$$

we further simplify,

$$m_0 + 0.3148m_1 - 0.9c_0 - 0.28332c_1 \leq 3.1544$$
$$m_0 + 0.3148m_1 + 0.9c_0 + 0.28332c_1 \geq 3.1544$$

$$m_0 + 1.0116m_1 - 0.9c_0 - 0.91044c_1 \leq 6.3321$$
$$m_0 + 1.0116m_1 + 0.9c_0 + 0.91044c_1 \geq 6.3321$$

$$m_0 + 1.3403m_1 - 0.9c_0 - 1.20627c_1 \leq 8.5727$$
$$m_0 + 1.3403m_1 + 0.9c_0 + 1.20627c_1 \geq 8.5727$$

$$m_0 + 1.0225m_1 - 0.9c_0 - 0.92025c_1 \leq 1.8625$$
$$m_0 + 1.0225m_1 + 0.9c_0 + 0.92025c_1 \geq 1.8625$$

$$m_0 + 0.5481m_1 - 0.9c_0 - 0.49329c_1 \leq 1.6882$$
$$m_0 + 0.5481m_1 + 0.9c_0 + 0.49329c_1 \geq 1.6882$$

$$m_0 + 1.2413m_1 - 0.9c_0 - 1.11717c_1 \leq 0.5709$$
$$m_0 + 1.2413m_1 + 0.9c_0 + 1.11717c_1 \geq 0.5709$$

$$m_0 + 1.2119m_1 - 0.9c_0 - 1.09071c_1 \leq 2.9454$$
$$m_0 + 1.2119m_1 + 0.9c_0 + 1.09071c_1 \geq 2.9454$$

We thereby standardize the constraints by adding slacks for the less than constraints and reducing surplus for the greater than constraints(with additional variable) as seen below:

$$m_0 + 0.3148m_1 - 0.9c_0 - 0.28332c_1 + S_1 = 3.1544$$
$$m_0 + 0.3148m_1 + 0.9c_0 + 0.28332c_1 - S_2 + P_1 = 3.1544$$

$$m_0 + 1.0116m_1 - 0.9c_0 - 0.91044c_1 + S_3 = 6.3321$$
$$m_0 + 1.0116m_1 0.9c_0 + 0.91044c_1 - S_4 + P_2 = 6.3321$$

$$m_0 + 1.3403m_1 - 0.9c_0 - 1.20627c_1 + S_5 = 8.5727$$
$$m_0 + 1.3403m_1 + 0.9c_0 + 1.20627c_1 - S_6 + P_3 = 8.5727$$

$$m_0 + 1.0225m_1 - 0.9c_0 - 0.92025c_1 + S_7 = 1.8625$$
$$m_0 + 1.0225m_1 + 0.9c_0 + 0.92025c_1 - S_8 + P_4 = 1.8625$$

$$m_0 + 0.5481m_1 - 0.9c_0 - 0.49329c_1 + S_9 = 1.6882$$
$$m_0 + 0.5481m_1 + 0.9c_0 + 0.49329c_1 - S_{10} + P_5 = 1.6882$$

$$m_0 + 1.2413m_1 - 0.9c_0 - 1.11717c_1 + S_{11} = 0.5709$$
$$m_0 + 1.2413m_1 + 0.9c_0 + 1.11717c_1 - S_{12} + P_6 = 0.5709$$

$$m_0 + 1.2119m_1 - 0.9c_0 - 1.09071c_1 + S_{13} = 2.9454$$
$$m_0 + 1.2119m_1 + 0.9c_0 + 1.09071c_1 - S_{14} + P_7 = 2.9454$$

Haven standardize the above constraints, the next step is to proceed y combining the objective function in (7) with the standardized constraints so that we may have a linear programming problem with a single objective function and 14 constraints (because we are considering 7 observations only). Hence, our LPP is of the form:

$Z = \min[c_0 + 19{:}32c_j]$; $j = 1$

subjected to:

$$m_0 + 0.3148m_1 - 0.9c_0 - 0.28332c_1 + S_1 = 3.1544$$
$$m_0 + 0.3148m_1 + 0.9c_0 + 0.28332c_1 - S_2 + P_1 = 3.1544$$

$$m_0 + 1.0116m_1 - 0.9c_0 - 0.91044c_1 + S_3 = 6.3321$$
$$m_0 + 1.0116m_1 + 0.9c_0 + 0.91044c_1 - S_4 + P_2 = 6.3321$$

$$m_0 + 1.3403m_1 - 0.9c_0 - 1.20627c_1 + S_5 = 8.5727$$
$$m_0 + 1.3403m_1 + 0.9c_0 + 1.20627c_1 - S_6 + P_3 = 8.5727$$

$$m_0 + 1.0225m_1 - 0.9c_0 - 0.92025c_1 + S_7 = 1.8625$$
$$m_0 + 1.0225m_1 + 0.9c_0 + 0.92025c_1 - S_8 + P_4 = 1.8625$$

$$m_0 + 0.5481m_1 - 0.9c_0 - 0.49329c_1 + S_9 = 1.6882$$
$$m_0 + 0.5481m_1 + 0.9c_0 + 0.49329c_1 - S_{10} + P_5 = 1.6882$$

$$m_0 + 1.2413m_1 - 0.9c_0 - 1.11717c_1 + S_{11} = 0.5709$$
$$m_0 + 1.2413m_1 + 0.9c_0 + 1.11717c_1 - S_{12} + P_6 = 0.5709$$

$$m_0 + 1.2119m_1 - 0.9c_0 - 1.09071c_1 + S_{13} = 2.9454$$
$$m_0 + 1.2119m_1 + 0.9c_0 + 1.09071c_1 - S_{14} + P_7 = 2.9454$$

We resolve using the Tora software. hence, the values of the center and spreadsof the membership function ($m_0, m_1, c_0, c_1,$) are given by:

$m_0 = 0$, $m_1 = 3.5418$, $c_0 = 4.2506$, $c_1 = 0$

where the minimized value of the objective function on the spread is:

$Z = 4.2506$

Hence, the TFN coefficients are as follows:
$\tilde{A}_0 = [0; 4.2506, 4.2506]$
$\tilde{A}_1 = [3.5418; 0, 0]$
If the Fuzzy regression model is given by:

$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 X_1$ ____(10)
We substitute the transformed values of X's from table(1) into equation

(10) toobtain our predicted $\tilde{Y}$ which will be converted into crisps fuzzy values.

$\tilde{Y} = [0; 4.2506, 4.2506] + [3.5418; 0, 0](0.3418)$

$\tilde{Y}_1 = [1.1149586; 4.2506, 4.2506]$

In summary:

$Y_2 = [3.58288; 4.2506, 4.2506]$

$Y_3 = [4.74707; 4.2506, 4.2506]$

$Y_4 = [3.62149; 4.2506, 4.2506]$

$Y_5 = [1.941261; 4.2506, 4.2506]$

$Y_6 = [4.39644; 4.2506, 4.2506]$

$Y_7 = [4.29231; 4.2506, 4.2506]$

Then the value of $\tilde{Y}$ low is converted to fuzzy crisp values using the formula be- $Y = a + \frac{1}{7}(m)$, so that:

$Y_1 = 1.1149586 + \frac{1}{7}(4.2506)$

$Y_1 = 1.722187$

Similarly, the predicted values are presented below:

$Y_2 = 4.19011$

$Y_3 = 5.354299$

$Y_4 = 4.228719$

$Y_5 = 2.5484896$

$Y_6 = 5.00367$

$Y_7 = 4.89954$

We proceed to estimate the error by taking the difference of the predicted $\tilde{Y}_j$ and the observed y_j. in summary;

$e_j = (Y_j - y_j), j = 1, 2, \cdots, n$

Table 2: Error of GFR model with h=0.1

sx_j	y^*_j	Estimated Y_j	e_j	e^2_j
1.37	5285.55	1.722187	-5283.83	279188859.47
2.75	19.02	4.19011	-14.82989	219.92563
3.82	1.77	5.3543	3.58429	12.84719
2.78	6.44	4.22872	-2.21128	4.88976
1.78	5.41	2.54849	-2.86151	8.18824
3.46	23.44	5.00367	-18.43633	339.89826
3.36	562.34	4.89954	-557.44046	310739.8664
				279500185.1

In this paper we have estimated the error of a geometric data set, we may wish to estimate the error of the data set using the ordinary regression modelso as to draw conclusion from former.

5. Numerical illustration

Table 3: Observed Non-Linear Data

X	Y
1.37	5285.55
2.75	19.02
3.82	1.77
2.78	6.44
1.78	5.41
3.46	23.44
3.36	562.34

.The model is given by:

$$Y = \alpha + \beta X$$

By computation, we obtain the slope and intercept:

$\alpha = 4495.835605$, $\beta = -3805.197654$

Hence the model:
$Y = 4495.835605 - 3805.197654X$

In summary, the error estimates are gotten by sqyaring the difference of each estimated $\tilde{Y}$ and the observed y values i.e e^2 as demonstrated in the table below;

5.1 Error Estimates

Table 3: Error of Ordinary Regression Model.

x_ja height1.37	y_j^* 5285.55	Estimated Y_j -717.285181	e_j 6002.83518	e^{2j} 36034030.21
2.78	19.02	-5968.45794	5987.47794	35849892.08
3.82	1.77	-10040.01943	10041.78943	100837535
2.78	6.44	-6082.61387	6089.05387	37076577.03
1.78	5.41	-2277.41622	2282.82622	5211295.551
3.46	23.44	-8670.14828	8693.58828	75578477.18
3.36	562.34	-8289.62851	8851.96851	78357346.5
				368945153.6

6. CONCLUSION

The geometric fuzzy regression model derived is used to estimate the error ofa non-linear data set and from observation it is more compact and performed better than the traditional regression as presented above. Using the fuzzy geometric model indicates the error accumulated is (279500185.1). While using the ordinary regression model, the error accumulated is (368945153.6) which is much larger than the former. Hence it is safe to say that the fuzzy geometric regression minimize error of a non-linear data set and performs better than the traditional linear regression.

REFERENCE

1. Arnold F. Shapiro(2005): Fuzzy regression models.
2. Y. J. Lai(2015): A new approach to fuzzy regression analysis.
3. H. Tanaka, Fuzzy data analysis by possibilistic linear model (2016), FuzzySets and Systems 24, pp.363-375.
4. D. Redden, W. Woodall(2019), Further examination of fuzzy linear regres- sion, Fuzzy Sets and Systems 79, pp. 203-211.
5. H. Tanaka, H. Ishibuchi (2016), Identification of possibilistic linear systemsby quadratic membership functions of fuzzy parameters, Fuzzy Sets and System 41, pp. 145-160.
6. Nonlinear Least Squares Regression. (n.d). Engieering Statistics Hand-book(2015). Retrieved from http://www.itl.nist.gov/div898/hanbook/pmd/section1/p.
7. Seber, G.A.F(George Arthur Frederick)(2018). Linear regression analysis,Hoboken, N.J.: Wiley-Interscience.
8. D. Redden, W. Woodall (2005), Properties of certain fuzzy regression meth-ods, Fuzzy Sets and Systems 64, pp.361-375.
9. Pedrycz W.(1994), Why triangular membership function?. Fuzzy Sets and Systems 64(1), pp. 21-30.
10. Hideo Tanaka, Haekwan Lee(2014), Fuzzy linear regression combining cen- tral tedency and possibilistic properties,IEEE Trans. on Fuzzy Systems, pp. 63-68.
11. Xie H.(2012): Fuzzy theory and application. prentice hall.

12. Pavel Skrabanek and Natalia Martinkova(2022): Getting started with fit-ting Fuzzy linear regression models.
13. Draper, N.R., Smith, H.(2019): Applied Regression Analysis. Wiley Hobo-ken.
14. D. Dubois, H. Prade (1988),Possibility Theory, Plenum Press, New York,.
15. H. Tanaka, J. Watada(1988), Possibilistic linear systems and their applica-tions to the linear regression model, Fuzzy Sets and Systems 27, pp. 275-289.
16. H. Tanaka, S. Uejima, Kazak(2018): Fuzzy linear regression analysis with fuzzy model, IEEE Systems, Trans. Systems Man, Cybernet. SMC-2, pp. 903-907,.
17. Diamond, P.(2018): Fuzzy least square. inf.sci.46.