مقارنة بين نماذج الانحدار المكاني والتقديرات الضبابية باستخدام المحاكاة

امنه حسين شوق علي

أ .م. د أسماعيل هادي جلوب

الكلية التقنية الادارية بغداد

dac0001@mtu.edu.iq

Comparison Between Spatial Regression models and Fuzzy Estimations using simulation

Amna Hussein Shawq

a. Dr. Ismail Hadi Globe

[1] postgraduate student / Technical Administrative College of Baghdad

الخلاصة

نظرًا لانتشار سرطان الثدي في العالم ككل وأعلى معدل إصابة بين النساء ومعدلات الإصابة المختلفة في مناطق مختلفة من العالم ، يمكننا دراسة هذه الظاهرة وفقًا لتأثير المكان وليس الزمان حيث تتأثر البيانات المكانية بشكل مباشر في العديد من العوامل المؤثرة ، بما في ذلك: (العمر ، والموقع الجغرافي ، والحالة الاجتماعية والاقتصادية والإنجابية ، وتناول الهرمونات ، وعوامل الخطر المتعلقة بنمط الحياة (التدخين ، والنظام الغذائي ، والسمنة ، والنشاط البدني) والتاريخ العائلي الذي يساهم في المرض.في بحثنا ، تم استخدام نماذج الانحدار المكاني لتحليل البيانات المكانية لسرطان الثدي ومقارنة التقدير بين نماذج الانحدار المكاني للبيانات غيرمضببة والمضببة للعثور على أفضل مقدر بين نماذج الانحدار المكاني. تم تطبيق نموذج الانحدار الذاتي المكاني (SAR) ، ونموذج الخطأ المكاني (SEM) ، ونموذج الانحدار التلقائي المكاني الضبابي (FSAR). كانت أفضل طريقة احتمالية لنموذج الانحدار التلقائي المكاني هي الأفضل وفقًا لطريقة المقارنة المستخدمة في MSE **الكلمات المفتاحية : نماذج الانحدار المكاني ، مصفوفة الاوزان المكانية , طريقة المربعات الصغرى ، طريقة الامكان الاعظم , متوسط مربعات الخطا, منطق الضبابي**

Abstract

Given the prevalence of breast cancer in the world as a whole and the highest infection rate among women and the different incidence rates in different regions of the world, we can study this phenomenon according to the effect of place rather than time as spatial data is affected directly in many influencing factors, including: (age, geographical location, and socioeconomic and reproductive status, hormone intake, lifestyle risk factors (smoking, diet, obesity, physical activity) and family history that contribute to the disease.In our research, spatial regression models were used to analyze the spatial data of breast cancer and compare the estimation between the spatial regression models for unfuzzy and fuzzy data to find the best estimator among the spatial regression models. The spatial autoregressive (SAR) model, the spatial error model (SEM), and the fuzzy spatial autoregressive model (FSAR) were applied. The maximum likelihood method of the spatial automatic regression model was the best according to the comparison method used MSE. **Keywords: Spatial regression models, spatial weights matrix, least squares method, maximum likelihood estimated , mean squares error, Fuzzy Logic**

الرموز	المصطلحات
(spatial autoregressive model(SAR))	١- الانحدار الذاتي المكاني
(spatial error model(SEM))	٢- الانحدار الخطا المكاني
(Fuzzy spatial autoregressive model(FSAR)	٣- الانحدار الذاتي المكاني المضبب
(Ordinary Least Squares (OLS))	٤- طريقة مربعات الصغرى
(maximum likelihood estimated (MLE))	٥- طريقة الامكان الاعظم
(Mean Square Error (MSE))	٦- متوسط مربعات الخطا

1- Introduction

Spatial regression is one of the modern statistical methods for analyzing the relationship between regression variables in the presence of a spatial correlation (spatial dependence), as a relationship between regression variables and the spatial lag itself. It is one of the directions of spatial econometrics that deals with spatial phenomena, where spatial series are distributed for each variable on the basis of space, not time [2]. Spatially characterized by the spatial dependence between observations of the sample data at different points, which means that the observations are close to each other; It has a greater degree of spatial dependence than those farther from the center, namely, the strength of spatial dependence between observations decreases with the distance between them, and ignoring spatial dependence may lead to weaknesses in statistical methods for spatial data analysis. The theory of aggregates is fuzzy, which is concerned with phenomena whose variables are measured in points Measured in terms of periods, or what are described as cases with fuzzy data Because of its characteristics that make it unclear, such as variables that belong to a certain degree to its groups and do not have a complete affiliation, as well as the case for linguistic variables that cannot be measured numerically, and studying how to formulate spatial regression models for fuzzy data based on fuzzy logic information [1], In this research, the issue of the spread of breast cancer in Iraq to a number of regions was addressed. The causes of cancer are still not well understood. But there are many causes and factors that influence the development of cancer (breast cancer). These risk factors include age, geographic location (country), socioeconomic and reproductive status, hormone intake, lifestyle risk factors (smoking, diet, obesity, physical activity), and history. Familial breast cancer contributes to a better understanding of the risk of breast cancer. Therefore, this study came to address this dangerous phenomenon in its various dimensions and by using simulation to generate non-fuzzy data and fuzzy data, then apply spatial regression models to obtain estimations and apply estimation methods: the usual least squares method and the method of greatest possibility, then compare between spatial regression models with a measure of average squares error to obtain On the best estimate.

2- Spatial regression : [9][17]

Spatial regression is a statistical method used to determine the relationship between the repressor variable and the regression variables taking into account the correlation between regions. it must use a spatial model spatial the data, because the regression variables that influence the regression variable can be different at each location. the fundamental concepts of spatial dependence and spatial autocorrelation, which is a property of data that arises whenever there is a spatial pattern in the values, as opposed to a random pattern that indicates no spatial autocorrelation., The general model of spatial is:[13]

$Y = \rho WY + Z\beta + \lambda W_u + \varepsilon$ ... (1)

Where

$\varepsilon \sim N(0, \sigma^2 In)$

Y: is a vector (n × 1) for the observation regression variable.

W: is the spatial weights matrix (n × n).

ρ: parameter of spatial autoregressive .

λ: parameter of spatial autoregressive for error.

Z: matrix (n × (k+1)) for the observation regression variables. β: is a vector ((K+1) × 1) Parameter to be estimated.

ε: is the vector (n × 1) for the error term.

u: is the vector (n × 1) for the error term which is spatial correlated.

There are several models for spatial regression as follows:

2- **Spatial Autoregressive Model (SAR):[4][5]**this model (SAR) is a spatial regression model whose regression variables are spatially correlated meaning that this model has a dependence on one observation in a region with observations in its neighboring region. The model yields better classification and prediction accuracy for many spatial data sets exhibiting strong spatial autocorrelation. It is the most straightforward way of incorporating the notion of spatial dependence in a linear regression framework. The general forms for SAR can be written as:-$Y = \rho W y + Z\beta + \varepsilon$ ... (2)Where$\varepsilon \sim N(0, \sigma^2 I_n)$

Y: is a vector (n × 1) for the observation regression variable.

W: is the spatial weights matrix (n × n).

ρ :parameter of spatial autoregressive model.

Z: matrix (n × (k+1)) for the observation regression variables.

β: is a vector ((K+1) × 1) Parameter to be estimated.

ε: is the vector (n × 1) for error term

4- Spatial Error Model (SEM): [14][15]One of the most important violations that plague regression model is the independence of the error term, so it will be studied with this model. It is assumed that the error or (model errors are linked spatially) reversed the presumption of independence of errors, one of the aims of THE model spatial error model (SEM) to spatial error correction$Y = Z\beta + e$ ... (3)

$e = \lambda W_2 + \varepsilon$

$e = (I - \lambda W)^{-1} \varepsilon$

$Y = Z\beta + [(I\_N - \lambda W\_2)]^{-1} \varepsilon$

$\varepsilon \sim N(0, \sigma^2 I_n)$

Where

Y: is a vector (n ×1) for the views depend variable. W: is the spatial weights (n×n) matrix

Z: matrix (n× (k+1)) the observation of explanatory variables

β: vector ((K+1) ×1) Parameter required estimation In: identity (n×n) matrix

λ: is the spatial parameter In: is the identity (n×n) matrix

u: is a vector of (n×1) error term which spatially correlated e: is a vector of (n×1) random error term

5 - Fuzzy Logic Subtract d. Lutfi Zadeh in 1965, the concept of the fuzzy group, Set Fuzzy, which differs from the classical group in that it allows each element to have partial affiliation, where each element has a degree of belonging to the group with values ranging between zero and one (one affiliation is full membership and zero is lack of membership,[8] and the values between them indicate degrees of partial membership)Fuzzy logic is based on the concept of a fuzzy number:

5-1 Fuzzy number[7]: describes uncertain cases or unclear data as variables that belong in proportion to their totals and have no complete affiliation, or meaningful variables that cannot be measured numerically, but in the form of periods. Its affiliation is from one to zero,

5-2 Membership function (affiliation[7]): It is used to determine how any element x belongs to the fuzzy group A within an inclusive group X. A value is determined for it in the range between [1,0] and it is symbolized by μA (x)Function: [7] Membership 5-2-1 TriangularThis function has three parameters a,b,c, and it can be represented by the following formula:

$$\mu_{A(x)} = \begin{cases} \dfrac{x-a}{b-a} & if\ a \le x \le b \\ \dfrac{c-x}{c-b} & if\ b \le x \le c \\ 0 & otherwise \end{cases}$$

5-2-2 Trapezoidal Membership Function:[7]This function has three parameters a,b,c,d and it can be represented by the following formula:

$$\mu_A(x) = \begin{cases} \dfrac{x-a}{b-a}, & \text{if } a \le x \le b \\ 1, & \text{if } b \le x \le c \\ \dfrac{d-x}{d-c} & if\ c \le x \le d \\ 0 & otherwise \end{cases}$$

5-2-3 Bell-shaped Membership Function:[7]

It is also called the Gaussian function and is represented as in the following formula:

$$\mu_{A(x)} = e^{c} - \frac{(x-a)^2}{b}, \qquad if \quad -\infty \le x \le \infty$$

5-3 fuzzy spatial autoregressive model[16] : So, the general formula for the fuzzy spatial autoregressive model, after converting the fuzzy number into a normal number with the trigonometric affiliation function, is the formula:

$$y_c = \lambda w y_c + e_c \quad \dots(4)$$

$$\breve{y}_c = (I - \lambda w)^{-1} e_c \qquad when\ \lambda = 0$$

$$y_c = \rho w y_c + e_c \qquad |\rho| < 1$$

$$\breve{y}_c = (I - \rho w)^{-1} e_c$$

y_c= vector (n*1) representing the fuzzy numbers of the dependent variable

I = unit matrix with dimension (N*N)

ρ= spatial regression parameter

W= spatial weight matrix

=is a vector (n*1) that represents the random errors of the fuzzy modele_c

6- modalities: Estimation

6-1 method of estimating the Maximum likelihood:[17]

The method of Maximum likelihood (MLE), is one of the most important methods because it gives the best estimate of the parameter among several possible estimations, and it is possible to estimate in this way for the two models (SAR), (SEM),(FSAR), which models were defined for the first time by Ord (1975)[17] [3] Assuming it is a natural state of error terms. Then follows the joint probability from the multivariate normal distribution to Y in contrast to what applies to the classical model

6-1-1 Estimation Maximum likelihood for (SAR) model : [2][10]

The first comprehensive treatment of maximum likelihood estimation of regression models that incorporate spatial autocorrelation in the form of a spatial lag by Ord (1975). An important aspect of this likelihood function is the Jacobian of the transformation, which takes the form $|I-\rho W|$ in respectively the spatial lag models:

$Y = X\beta+\rho WY + \varepsilon$ …(7)

Mximumliklehood for model is:

$$\text{Ln } L(\beta, \rho, \sigma^2) = -\frac{n}{2}\text{Ln } 2\prod - \frac{n}{2}\text{Ln } \sigma^2 + \text{Ln}|(I-\rho W)| -(1/2\sigma^2)[(Y - \rho Wy - X\beta)' (Y-\rho Wy - X\beta)] \quad \dots \quad (8)$$

By differentiating with respect to β and σ2 and setting them equal to zero, we get

$b_{(MLE)} = [\ (X'X)^{-1}\ +X'(I-\rho W)Y\ ]$ (9)

$e = Y-X[\ (X'X)^{-1}\ +X'Y\ ]-\rho[\ (X'X)^{-1}\ +X'WY\ ]$ (10)

Using the iterative method for the greatest possibility function steps, the variance estimate is obtained as follows:

$$\tilde{\sigma}^2 = \frac{(y - \rho Wy)'(y - \rho W y)}{n} \qquad (11)$$

6-1-2 Estimating the greatest potential of the SEM model:[2][17]

The interest in this model is (θ), which explains the correlation between the residuals [6].

The maximum possibility function for this model is:

$Y = X\beta + u$ …(12)

$$L(\beta, \theta, \sigma^2) = \frac{-n}{2}\text{Ln } 2\prod - \frac{n}{2}\text{Ln } \sigma2 + \text{Ln } |(I-\theta W)| - (1/2\sigma^2)[(Y-X\beta)'(I-\theta W)'(I-\theta W)(Y-X\beta)]) \dots (13)$$

By differentiating with respect to β and σ2 and setting them equal to zero, we get

$b_{(MLE)} = [\ (X'(I-\theta W)'(I-\theta W)X\ ]^{-1}\ X(I-\theta W)'(I-\theta W)Y$ … (14)

$e = [\ Y - X\,b_{(MLE)}\ ]$

$\sigma^2_{(MLE)} = e'e\ /\ n$ … (15)

6-1-3 Estimating the maximam liklood of the FSAR model:[16]

$Y_C - \rho WY_C = Z_C\beta + e_C$ **…(17)**

$$L(\beta،\rho،\sigma^2/Y_C\ ،Z_C) = \frac{1}{(2\pi\sigma^2)^{n/2}}|I-\rho W|\exp\left[-\frac{1}{2\sigma^2}e'e\right] \dots (18)$$

We can get:

$$Ln(L) = -\frac{n}{2}Ln(2\pi) - \frac{n}{2}Ln\sigma^2 + |I-\rho W| - \frac{1}{2\sigma^2}(Y_C - \rho WY_C - Z_C\beta)'(Y_C - \rho WY_C - Z_C\beta)\ (19)$$

$$\sigma^2_{MLE} = \frac{(e_{0C}-\rho e_{1C})'(e_{0C}-\rho e_{1C})}{n} \quad \dots (20)$$

6-2 method of estimating of the Ordinary Least Squares (OLS):[13][17]

The classical general regression model : is the most well-known of all regression techniques, and aims to estimate the regression coefficient vector β by the OLS method such that the total squared difference between the observed and predicted values for the response variable and the explanatory variables is minimized, This type of regression is known as "universal" because of the spatial invariance of its modulus estimates, which means that one model can be applied equally to different areas of interest. The mathematical formula is:

$y = X\beta + \varepsilon$ …(21)

6-2-1 Ordinary least squares estimation of the SAR model: [11][2]

The explanatory variables are independent and follow a normal distribution. Under these assumptions, the OLS estimation is unbiased and normal and can be statistically effective for spatial regression models. As shown below, the SAR model and its formula:[43]

$\underline{\varepsilon} = Y\ \rho WY - Z\underline{\beta}$ …(22)

$E(\varepsilon'\varepsilon) = Y'Y - \rho Y'WY - \rho Y'W'Y + \rho^2 Y'W'WY - 2\beta'Z'Y + 2\rho\,\beta'Z'WY + \beta'Z'Z\beta$ …(23)

After the derivation and substitution operations the estimation formula for the spatial autoregressive parameter (ρ) is obtained.

$\tilde{\rho} = [\ Y'W'WY - \underline{b}_L'Z'WY\ ]^{-1}\ [\ Y'WY - \underline{b}_o'Z'WY\ ]$ …(24)

6-2-2 Ordinary least squares estimation of the SEM model:[11] [2]

To estimate the spatial error model using the OLS method, whereby it is assumed that the errors are independent and follow a normal distribution under these assumptions, as shown below: whereas

$Y = Z\beta + (I-\rho w)^{-1} u$

$E(\varepsilon\varepsilon') = \sigma^2 [(I-\rho W)'(I-\rho w)]^{-1}$ …(25)

$= [W(I-\rho W)^{-1}(Z\beta - \rho WZ\beta) E(u^T) + W(I-\rho W)^{-1} E[uu^T]$

$\sigma_\varepsilon^2 W (I-\rho W)^{-1}$ …(26)

6-2-3 Ordinary least squares estimation of the FSAR model:[16]

The estimate for the model after converting the fuzzy number into a regular number with the affiliation function is:

$$e_C = Y_C - \rho W Y_C - Z_C\beta \quad \dots (27)$$

After a number of derivations and substitutions we get

$$\rho[Y'_C W' W Y_C - b'_1 Z'_C W Y_C] = Y'_C W Y_C - b'_0 Z'_C W Y_C$$

$$\sigma^2 = \frac{(Y_C - \hat{\rho} W Y_C - Z_C\hat{\beta})'(Y_C - \hat{\rho} W Y_C - Z_C\hat{\beta})}{n-k} \quad \dots (28)$$

7- Weight Matrix:

It is a square matrix in which elements have positive values and are denoted by W and not necessary to be symmetric and create this matrix based on neighbouring and relation neighbour from the location for another location in the same row in the rows of the matrix and value for the diagonal usually equal to zero and chose weight matrix is very important for Determine the spatial effects so we must create an Appropriate weight matrix and there for some way to create this matrix [2][13]

To build the spatial weights matrix there are two types, in this thesis we use Binary Contiguity Weights Matrix.

7-1 Binary Contiguity Weights Matrix:[12][13]

The matrix W is positive and square (n×n), if i,j are contiguous wij=1 and if not contiguous wij =0

$$W_{ij} = \begin{cases} 1 & \text{if } i \text{ neighbour } j \\ 0 & \text{otherwis} \end{cases}$$

The other type is the distance weight

$$W_{ij} = \begin{cases} 1/dij^2 & \text{for } i=j \\ 0 & \text{for } i \neq j \end{cases}$$

d_i: is the distance between the geographical centres of the two regions.

To build the contiguity weights matrix dependence on common borders, we review some important methods from:

7-1-1 Rook Contiguity:[12][13]

This matrix the value of the element equal to one if the two areas neighbor by limited and have relation between the two areas in any side, and it is equal to zero if it is not contiguous. This matrix used in applications more than the others.

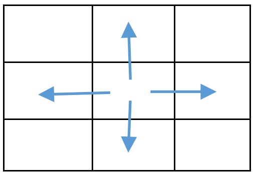

Figure (2) for the Rook Weight Matrix

7-1-2 Bishop Contiguity:[12][13]

This matrix arises if the two areas connected a point and this point is the connected limited between the two location are smallest connected limited and be the elements value is equal to one and another element is equal to zero .

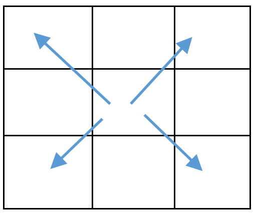

Figure (3) for the Bishop Weight Matrix

7-1 - 3 Queen Contiguity:[12][13]

This matrix gets its elements from the sum of (rook) and (bishop) matrix elements and neighbor in this matrix is based on connect point or connect limited.

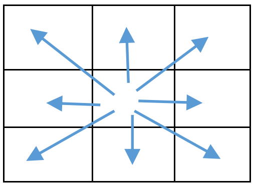

Figure (2-4) shows the Queen Weight Matrix

8- (Comparison Criteria)

8-1 Mean squared error (MSE)[17]: It is a preference comparison criterion that is widely used to predict accuracy, and it is the average squared difference between the actual observations and the expected observations. (Z1,Z2…Zn) A random sample drawn from the population of the distribution function formula (Z;)F is known, but depends on the unknown parameter λ and estimated for $\hat{\lambda}$, the value is:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{\lambda})\,(\lambda)^2$$

n = the number of times the experiment was repeated

$\hat{\lambda}$ = estimated value

λ = true value

. Experimental side:

In this research, the simulation method was used to generate the data required in the research. Simulation can be defined as a numerical technique used to carry out numerical computer tests. The logical and mathematical relationships interact with each other to describe a behavior or phenomenon in the real world. Simulation is distinguished in that it reduces the high costs, time and effort that It is required to work in obtaining samples from the practical reality. The Monte Carlo method is among the most important, best and most widely used simulation methods. To generate the necessary data for the purpose of comparison between the methods with different sample sizes and different values, the simulation method was implemented through the use of the MATLAB statistical program and the use of the comparison criterion between the estimation methods (mean square error (MSE). By means of simulations, spatial regression models are generated and applied

Spatial autoregressive model (SAR) : $y = \rho Wy + \beta X + \varepsilon$

Spatial error model (SEM) : $y = \beta X + e$

Fuzzy Spatial autoregressive model (FSAR);

$$y_c = \lambda w y_c + e_c$$

To generate spatial data in a Monte Carlo way, where the simulation includes several proxy steps:

The first step: Determining two sample sizes (n = 150) concerned with the study and assuming two sets of values of explanatory variables p: (P = 5, P = 10) and one response variable (Yi).

The second step: In this step, an independent random variable (X) and an error-limit random variable (ε) are generated.

The third step: In this step, matrices are found, W_{ij} (QueenContiguity) , The criterion is that adjacency is when two cells share a common side as well as a common vertex

$$W_{ij} = \begin{cases} 1 & if\ i\ and\ j\ are\ contiguous \\ 0 & if\ i\ and\ j\ are\ contiguous \end{cases}$$

Fourth step: estimation methods: Two types of estimation methods were used for the spatial regression model parameters ((FSAR), (SAR), (SEM)), each estimation method for each regression model, namely: The method of ordinary least squares (OLS) and its formula is as follows:

- Estimation using the least squares method of the SAR spatial error model, according to the mathematical formula

$b_{(OLS)}$= (X' X)-1 X' Y – ρ(X' X) -1 X' WY …(28)

- Estimation using the least squares method of the SEM spatial error model, according to the mathematical formula -

$$b_{ols} = (X'^{W^{-1}}X)^{-1}X'^{W^{-1}}Y \quad \ldots (29)$$

- Estimation using the least squares method of the SEM spatial error model, according to the mathematical formula

$$\widehat{b_{ols}} = (x'_C x_C)^{-1} x'_C Y_C - \rho (x'_C x_C)^{-1} x'_C W Y_C \quad ...(30)$$

The method for estimating the maximum likelihood (MLE) and its formula is as follows:

$$L = \prod_{i=1}^{n} f(x_i, \lambda)$$

- Estimation using the maximum likelihood of the SAR model, the spatial auto regression, according to the mathematical formulas:

$$b_{(MLE)} = [\ (X'\ X)^{-1}\ +X'\ (I - \rho W)Y\ ] \quad ...(31)$$

- Estimation using the maximum likelihood of the SEM model, the spatial auto regression, according to the mathematical formulas:

$$b(MLE) = [\ (X'\ (I - \theta W)'\ (I - \theta W)\ X\ ]\text{-}1\ \ X\ (I - \theta W)'\ (I - \theta W)\ Y \quad ...(32)$$

- Estimation using the maximum likelihood of the FSAR model, the Fuzzy spatial auto regression, according to the mathematical formulas:

$$\hat{b}_{MLE} = (x'_C x_C)^{-1} x'_C (I - \rho W) Y_C \quad ...(33)$$

The sixth step: to compare and compare between statistical estimators, the most important statistical measure will be used, which is the mean squared error (MSE), as it is considered the most common, as it measures how close and far the estimator is from the real values.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (\hat{\lambda})\ (\lambda)^2$$

7- The results of the simulation experiment:

The first experiment: includes the sample size (n = 150) and for the group of explanatory variables P = 5, and by applying the estimation methods (OLS, MLE) for each model (SAR, SEM, FSAR) as in Table No. (2) to estimate the parameters of the models, as for finding the efficiency The relative methods, a comparison table was made using the comparison scale (MSE), and we had the results as shown in Tables (1). **Table (1) which includes comparison results using the mean square error measure (MSE) for both estimation methods (OLS, MLE) applied to each spatial regression model (SAR, SEM, FSAR).**

Model	Method	MSE
SAR	OLS	0.041856
	MLE	0.042533
SEM	OLS	0.040916
	MLE	0.022283
FSAR	OLS	0.021892
	MLE	0.047265
Best model	FSAR -OLS	0.021892
Worst model	FSAR -MLE	0.047265

Table (2) includes an estimate of the parameters of the probabilistic regression models (SAR, SEM, FSAR) and by applying the estimation methods (OLS, MLE) for each model

Estimation Method	SAR		SEM		FSAR	
	OLS	*MLE*	*OLS*	*MLE*	*OLS*	*MLE*
$\hat{\beta}_0$	0.198735	0.199719	0.197342	0.155429	0.153588	0.206227
$\hat{\beta}_1$	0.764251	0.765236	0.762859	0.720945	0.719104	0.771744

$\hat{\beta}_2$	0.520029	0.521014	0.518637	0.476724	0.474883	0.527522
$\hat{\beta}_3$	0.405618	0.406603	0.404226	0.362313	0.360472	0.413111
$\hat{\beta}_4$	0.340959	0.341944	0.339566	0.297653	0.295812	0.348452
$\hat{\beta}_5$	0.294927	0.295912	0.293534	0.251621	0.249781	0.30242

The comparative results are shown as in Tables No. (9) for the second experiment, where the sample size is: n = 150, and for the group of explanatory variables P = 5 and one response variable Yi, and by applying the estimation methods (OLS, MLE) to estimate the parameters of each spatial regression model (SEM, SAR, FSAR).) and using the comparison scale MSE Mean squares error, the comparison ratios are almost close, but the MLE estimation method for the SAR model showed the lowest mean squares error compared to the rest of the estimates for the spatial regression models, reaching a percentage of (0.022537), which is the best model, and the OLS method for the SEM model showed the largest mean squares The error is clear from the rest of the estimates, as it amounted to (0.0428175), and thus it is considered the worst model

The second experiment: where the sample size is 150, but for the group of explanatory variables P = 10 and for one response variable (Y), and by applying the methods (OLS, MLE) for each model (SAR, SEM, FSAR) as in Table No. (4) to estimate the parameters of the spatial regression models, As for finding the relative efficiency of the methods, a comparison table was made using the comparison scale (MSE).**Table (3), which includes comparison results using the mean square error measure (MSE) for both estimation methods (OLS, MLE) applied to each spatial regression model (SAR, SEM, FSAR)**

Model	Method	MSE
SAR	OLS	0.059208231
	MLE	0.095740539
SEM	OLS	0.071436977
	MLE	0.127238258
FSAR	OLS	0.109673131
	MLE	0.028079074
Best model	FSAR -MLE	0.028079074
Worst model	SEM- MLE	0.127238258

Table 4) includes an estimate of the parameters of the probabilistic regression models (SAR, SEM, FSAR) and by applying the estimation methods (OLS, MLE) for each model.

Estimation Method	SAR		SEM		FSAR	
	OLS	***MLE***	***OLS***	***MLE***	**OLS**	**MLE**
$\hat{\beta}_0$	**0.04843**	**0.070464**	**0.056604**	**0.085525**	**0.077441**	**0.018952**
$\hat{\beta}_1$	**0.729687**	**0.751722**	**0.737862**	**0.766782**	**0.758698**	**0.70021**
$\hat{\beta}_2$	**0.505889**	**0.527923**	**0.514063**	**0.542984**	**0.5349**	**0.476411**
$\hat{\beta}_3$	**0.405904**	**0.427938**	**0.414078**	**0.442999**	**0.434915**	**0.376426**
$\hat{\beta}_4$	**0.323928**	**0.345962**	**0.332102**	**0.361023**	**0.352939**	**0.29445**
$\hat{\beta}_5$	**0.267238**	**0.289272**	**0.275412**	**0.304333**	**0.296249**	**0.23776**
$\hat{\beta}_6$	**0.23464**	**0.256675**	**0.242815**	**0.271735**	**0.263652**	**0.205163**
$\hat{\beta}_7$	**0.216936**	**0.23897**	**0.22511**	**0.254031**	**0.245947**	**0.187458**
$\hat{\beta}_8$	**0.185984**	**0.208019**	**0.194158**	**0.223079**	**0.214996**	**0.156507**
$\hat{\beta}_9$	**0.181571**	**0.203605**	**0.189745**	**0.218666**	**0.210582**	**0.152093**
$\hat{\beta}_{10}$	**0.185072**	**0.207107**	**0.193246**	**0.222167**	**0.214084**	**0.155594**

- The comparative results are shown as in Tables No. (11) for the second experiment, where the sample size is n = 150, and for the group of explanatory variables P = 10 and one response variable Yi, and by applying the estimation methods (OLS, MLE) to estimate the parameters of each spatial regression model (SEM, SAR, FSAR) and using By comparison scale MSE Mean squares error The MLE estimation method for the FSAR model showed the lowest mean square error compared to the rest of the estimates for the spatial regression models, reaching a percentage of (0.028079074), which is the best model. (0.127238258), thus it is considered the worst model**The following steps show how to convert data into fuzzy data by using the trigonometric membership function: The first step**: Determine the highest value and the lowest value of each column for the independent or dependent variables in all tables, the results of the research experiments.**The second step**: we extract the rang for each value from the value of a column for all the tables of the experiments.**The third step**: the membership number is found and 7 is selected to divide the affiliation periods for each group of numbers to complete the fuzzing process by role. **The fourth step**: obtaining the fuzzy numbers and applying the previous research methods with the experiments that were conducted

- **The third experiment**: includes fuzzy data with a sample size of (n = 150) and for the group of explanatory variables (P = 5), and by applying the estimation methods (OLS, MLE) for each model (SAR, SEM, FSAR) as in Table No. (5) to estimate the parameters of the models As for finding the relative efficiency of the methods, a comparison table was made using the comparison scale (MSE), and we had the results as shown in Tables (6). **Table (6) which includes the results of a comparison using the mean square error measure (MSE) for both estimation methods (OLS, MLE) applied to each spatial regression model (SAR, SEM, FSAR).**

Model	Method	MSE
SAR	OLS	0.053316
	MLE	0.068888
SEM	OLS	0.042817
	MLE	0.020035
FSAR	OLS	0.058281
	MLE	0.066322
Best model	SEM -MLE	0.020035
Worst model	SAR -MLE	0.068888

Table (5) includes an estimate of the parameters of the probabilistic regression models (SAR, SEM, FSAR) and by applying the estimation methods (OLS, MLE) for each model

Estimation Method	**SAR**		**SEM**		**FSAR**	
	OLS	*MLE*	*OLS*	*MLE*	OLS	MLE
$\hat{\beta}_0$	0.213775	0.230503	0.200128	0.138125	0.219467	0.227953
$\hat{\beta}_1$	0.779291	0.79602	0.765645	0.703641	0.784983	0.79347
$\hat{\beta}_2$	0.535069	0.551798	0.521423	0.45942	0.540761	0.549248
$\hat{\beta}_3$	0.420658	0.437387	0.407012	0.345009	0.426351	0.434837
$\hat{\beta}_4$	0.355999	0.372728	0.342353	0.280349	0.361691	0.370178
$\hat{\beta}_5$	0.309967	0.326696	0.296321	0.234318	0.315659	0.324146

The comparative results are shown in Tables No. (6) for the second experiment, where the sample size is: n = 150, and for the set of explanatory variables, P = 5, and one response variable, Yi, and by applying the estimation methods (OLS, MLE) to estimate the parameters of each spatial regression model (SEM, SAR, FSAR). And by using the comparison scale MSE Mean squares error, the comparison ratios are almost close, but the MLE estimation method for the SEM model showed less average squares error than the rest of the estimates for the spatial regression models, reaching a ratio of (0.020035), which is the best model, and the MLE method for the SAR model showed the largest mean squares error than The rest of the estimates are clear and amounted to (0.068888), and thus it is considered the worst model.

- **The fourth experiment:** for fuzzy data, where the sample size is 150, and for a group of illustrative variables, P = 10, and for one response variable (Y), and by applying the methods (OLS, MLE) for each model (SAR, SEM, FSAR) as in Table No. (8) to estimate the parameters of the spatial regression models As for finding the relative efficiency of the methods, a comparison table was made using the comparison scale (MSE), and we had the results as in Tables (7). **Table (7) which includes the results of a comparison using the mean square error measure (MSE) for both estimation methods (OLS, MLE) applied to each spatial regression model (SAR, SEM, FSAR).**

Model	Method	MSE
SAR	OLS	0.085723
	MLE	0.091137
SEM	OLS	0.021327
	MLE	0.045275
FSAR	OLS	0.040447
	MLE	0.066058
Best model	SEM -OLS	0.021327
Worst model	SAR- MLE	0.091137

Table (8) includes an estimate of the parameters of the probabilistic regression models (SAR, SEM, FSAR) and the application of estimation methods (OLS, MLE) for each model

Estimation Method	**SAR**		**SEM**		**FSAR**	
	OLS	*MLE*	*OLS*	*MLE*	OLS	MLE
$\hat{\beta}_0$	0.065056	0.068024	0.007192	0.037481	0.033063	0.053136
$\hat{\beta}_1$	0.746313	0.749282	0.688449	0.718738	0.71432	0.734393
$\hat{\beta}_2$	0.522515	0.525484	0.464651	0.49494	0.490522	0.510595
$\hat{\beta}_3$	0.42253	0.425498	0.364666	0.394955	0.390537	0.41061
$\hat{\beta}_4$	0.340554	0.343522	0.28269	0.312979	0.308561	0.328634
$\hat{\beta}_5$	0.283864	0.286832	0.226	0.256289	0.251871	0.271944
$\hat{\beta}_6$	0.251267	0.254235	0.193403	0.223691	0.219274	0.239347
$\hat{\beta}_7$	0.233562	0.236531	0.175698	0.205987	0.201569	0.221642
$\hat{\beta}_8$	0.20261	0.205579	0.144746	0.175035	0.170617	0.19069
$\hat{\beta}_9$	0.198197	0.201166	0.140333	0.170622	0.166204	0.186277
$\hat{\beta}_{10}$	0.201698	0.204667	0.143834	0.174123	0.169705	0.189778

- The comparative results are shown as in Tables No. 7)) for the second experiment, where the sample size is n = 150, and for the set of explanatory variables, P = 10, and one response variable, Yi, and by applying the estimation methods (OLS, MLE) to estimate the parameters of each spatial regression model (SEM, SAR, FSAR), and with the help of By comparison scale MSE Mean squares error The OLS estimation method of the SEM model showed the lowest average square error of the rest of the estimates for the spatial regression models, reaching 0.021327, which is the best model, and the MLE method of the SEM model showed the largest mean square error of the rest of the estimates clearly.

8. Conclusions: After conducting the description and implementation of simulation experiments on spatial regression models (SAR), (SEM)) (FSAR) and applying the two estimation methods (OLS), (MLE) and the results presented to obtain the best method, the researcher concluded the following:

1- We note by using the measure of the average sum of squares error of un fuzzy data for comparison of the estimation methods (OLS), MEL) in estimating spatial regression models (FSAR), (SEM), (SAR) according to the spatial weight matrix that the best estimation method is (OLS) for the model Fuzzy spatial autoregressive (FSAR) at a sample size of 150 and explanatory variables P = 5

2- We note when using the estimation methods (OLS), (MLE) on spatial regression models. The comparison results showed the comparison scale Mean Squares Error (MSE) in the second experiment, where the sample size is 150 items and explanatory variables P = 10 that the method of greatest possibility (MLE)) is better than the method of estimation (OLS) of the (FSAR) model, as it achieved the lowest value of the comparison standard (MSE) by increasing the number of explanatory variables with the same sample size for un fuzzy data.

3- As we notice at the same sample size 150 using estimation methods (OLS), (MLE) on spatial regression models and using the comparison scale Mean Squares Error (MSE)) but for fuzzy data and the explanatory variables were P = 5, the results showed that the method of greatest possibility (MLE) of the (SEM) model is the best because it achieved the highest value of the comparison standard (MSE).

4- Results for the same sample size of 150 for fuzzy data and at the number of explanatory variables P = 10 showed that the method (OLS) of the (SEM) model is the best method of estimation, since the comparison standard (MSE) was the lowest.

9- Recommendations: Based on the conclusions reached through the experimental results, the most important recommendations can be included as follows:

1- Using the method (OLS) and the method of greatest probability (MLE) to estimate fuzzy spatial autoregressive models on non-fuzzy data, due to its effectiveness. The method (OLS) and the method of greatest probability (MLE) can also be used to estimate spatial error models to estimate fuzzy data.

2- The use of a proposed spatial weight matrix (W) for spatial models.

3- Use the presented estimation methods and apply them to real data.

4- Using other methods to find the distances between the values of the observations as a standard for comparison.

10- Reference .

[1]& هبةعليطه الصباغ ، ٢٠٠٩ , "دراسة في المتغيرات المضببة الانحدار المتعدد المضبب " مجلة تكريت للعلوم الإدارية , محمد طه احمد
(المجلد والاقتصادية 5) العدد (14.)

[2] Anselin, L.(1988) “Spatial Econometrics: Methods and Models” Kluwer Academic, Dordrecht Anselin, L.: Under the hood Issues in the specification and interpretation of spatial regression models. Agricul. Econ.MLA

[3] Anselin, L. (1999); **Spatial Econometrics**; Bruton Center School of Social Sciences University of Texas at Dallas Richardson, TX 75083 – 0688

[4] Anselin, L. Under the hood: Issues in the specification and interpretation of spatial regression models. Agric. Econ. 2002, 27, 247–267

[5] Anselin, Luc, and Anil K. Bera. "Spatial dependence in linear regression models with an introduction to spatial econometrics." *Statistics textbooks and monographs* 155 (1998): 237-290

[6] L. Anselin and H. H. Kelejian,“*Testing for Spatial Error Autocorrelation in the Presence of Endogenous Repressors*.” International Regional Science Review, vol .20, no.1-2.pp. 153–182,1997.

[7]Bolotin A., “Replacing indicator variables by fuzzy membership func- tions in statistical regression models: Examples of epidemiological stud- ies” In: Lecture Notes in Computer Science: Biological and Medical Data Analysis, Springer, 2004, pp. 251-258.

[8] Corral, Norberto, María Ángeles Gil, and Pedro Gil. "Interval and fuzzy-valued approaches to the statistical management of imprecise data." *Modern mathematical tools and techniques in capturing complexity*. Springer, Berlin, Heidelberg, 2011. 453-468

[9] Elhorst, Paul, and Solmaria Halleck Vega. "On spatial econometric models, spillover effects, and W." (2013).

[10] Florax, Raymond, and Henk Folmer. “Specification and Estimation of Spatial Linear Regression Models.” Regional Science and Urban Economics 22.3 (1992): 405–432. Web

[11] González-Rodríguez, Gil, et al. "Least squares estimation of linear regression models for convex compact random sets." *Advances in Data Analysis and Classification* 1.1 (2007): 67-81

[12] Holloway, Garth, D. Lacombe, and James P. LeSage. “Spatial Econometric Issues for Bio-Economic and Land-Use Modelling.” Journal of Agricultural Economics 58.3 (2007): 549–588. Web

[13]Lesage, James P. (1999); **The Theory and Practice of Spatial Econometrics**; Department of Economics University of Toledo, pp(1 – 148).

[14] LeSage J. P., (1998). Spatial Econometrics , Department of Economics University of Toledo; pp(30-43).

[15] LeSage, J. P. (2008). An introduction to spatial econometrics. Revue d'économie industrielle, (123), 19-

[16] Mohammed, Jaufar Mousa, Ammer Fadel Tawfeeq, and Maysoon M. Aziz. "Maximum likelihood for fuzzy pure spatial autoregressive model." *International Journal of Nonlinear Analysis and Applications* 13.1 (2022): 2267-2277

[17] Ord, J.K.(1975); Estimation Methods for Models of Spatial Interaction; Journal of the American Statistical Association, VOL. 70, NO. 349, pp (120-126(