A study on atmospheric pressure in krg using spatial regression (SAR and SEM)

Sham Azad Rahim

Statistics and Computer department, College of Commerce, University of Sulaimani, Sulaimani, Kurdistan Region, Iraq

sham.rahim@univsul.edu.iq

Dr. Mohammad Mahmood Faqe Hussein

Statistics and informatics department, College of Administration & Economic, university of Sulaimani, Sulaimani, Kurdistan Region, Iraq

Mohammad.Faqe@univsul.edu.iq

ARTICLE INFO

ABSTRACT

The development of using location analysis in all aspects of life encouraged us to use it in this paper for analyzing weather in some Kurdistan regions which are (Sulaimaniyah, Erbil and Duhok) based on the data recorded by the devices placed in 27 different areas which had been recording the climate for all months during 2015. Moran test is used to check spatial dependency in a data and Spatial autoregressive model (SAR), spatial error model (SEM) is created also Lagrange test used to select the significant and alternative spatial model between (SAR and SEM). Different criteria or measures like (R2adj, RMSE, MAPE, AICc) are used for finding the best fit model also for all of them used (rook, bishop and queen) matrixes to specify the spatial correlation. The important result in the practical part shows that SAR model for queen and rook matrixes is significant while SEM model is not

significant for all three matrixes and SAR model for queen matrix is best appropriate model.

1. INTRODUCTION[4]

Spatial data refers to all types of data objects or elements that are presented in a geographical space or horizon. The spatial regression model worked as time series but lag refers to any place or far between places and not to time as time series, so in time series the trend only for one sides but it is different here because the trends have four sides in north ,east ,west and south while the relation between places studied by weight matrices and we can use the spatial regression models if we have spatial dependency between places and in spatial error term the errors assume to be correlated spatially that is the two condition for spatial regression models how ever in time series the time is active variable for explain the result but in spatial analysis the places and correlation between places is active and strong variable that depended for discuses the result .The Paper divided in two sections witch first section include the theoretical part for the spatial analysis and tests with rules while in second section is practical part with discuses results also include conclusion and recommendations.

2.Literature Review

1.In 2004 Baris M. Kazar, Shashi Shekhar, David J. Lilja, Ranga Raju Vatsavai and R. Kelley Pace studies the Comparing Exact and Approximate Spatial Auto-regression Model Solutions for Spatial Data Analysis.GIScience.

Applications that use spatial auto-regression (SAR) for data mining are working with ever increasing sizes of geo-spatial databases. The explosive growth in databases coupled with the demand for exact solutions for estimating SAR parameters are both computationally expensive and memory intensive. This paper presents two candidate approximate-semi-sparse solutions of the SAR model based on the Taylor series expansion and Chebyshev polynomials. When accuracy of these new approximation algorithms and an exact algorithm were compared, both provided accurate results. However, the approximation algorithms outperformed the exact algorithm in both terms of computation and memory usage. It was also noted that the exact algorithm was unable to solve any problem with over 10K observation points. They performed experiments on satellite imagery. Authors suggested exploring better model based on this approach to get better prediction

2.In 2006 Prasanna Man Shrestha study the Comparison of Ordinary Least Square Regression, Spatial Autoregression, and Geographically Weighted Regression for Modeling Forest Structural Attributes Using a Geographical Information System (GIS)/Remote Sensing (RS) Approach

The performances of three modeling techniques: (i) ordinary least square (OLS) regression, (ii) spatial autoregression (SAR) and (iii) geographically weighed regression (GWR) were compared for the task of predicting a key forest structural parameter crown closure – across a study area in west-central Alberta using a series of spectral and topographic variables.

3.In 2011 J. Paul Elhorst study the Spatial panel models

This paper provides a survey of the existing literature on spatial panel data models. Both static and dynamic models will be considered. The paper also demonstrates that spatial econometric models that include lags of the dependent variable and of the independent variables in both space and time provide a useful tool to quantify the magnitude of direct and indirect effects, both in the short term and in long term

4.In 2014 Omar Abdul Mwhsen Ali with student Sawsan Qasm Hadi studies the Spatial Regression Models Estimation for the poverty Rates In the districts of Iraq

This paper shows spatial regression model and model possessory error in an attempt to provide a general guide Shows the importance of spatial loading, with particular on the importance of using spatial regression models,which Each of which includes spatial reliability testing and that is whether or not find tests the Moran, and ignore this Reliability may lead to the loss of information important

for empowerment reflected in end up on the strength of estimate Statistical index extracted, these models are the link between the usual regression models with change models

5.In 2016 Hamid Saed Nwr with master student Swhad Ali Shahid for master project studies estimate spatial dynamic panel data model (SDPD) with fixed effects-stable state using the direct approach

Although spatial dynamic panel data model sparked a lot of attention in the last decade, however, the econometric analysis of spatial models and dynamic panel data rare so far, with capabilities there are no available take into account the study the lagged of dynamic spatial model of panel data for the presence of one or more of the endogenous variables (dependent) as explanatory variables, with lagged in time along or both. The presence of the dependent variable and endogenous lagged variables in a spatial lag model invalidate the use of known estimation methods such as(OLS) and (ML)

3.Spatial Regression Models:

3.1Spatial Autoregressive Model (SAR):[7]

The SAR regression model follows an autoregressive process, which is indicated by the presence of dependence relationship among a set of observations or spatial units. While the variable (λ) describes the power of neighboring from a place to another place and the depend variable(Y) convert to spatial depend variable (WY)

$$Y = \lambda WY + X\beta + e \dots\dots(1)$$

WhereW: is the spatial weight $(n \times n)$ matrix, λ: is spatial auto regressive regression Parameter The value of λ located between $-1 < \lambda < 1$ we have a positive spatial correlation if $\lambda > 0$ that is indicated the places is resemble but if $\lambda < 0$ we have negative spatial correlation and the places is not look like other also if $\lambda = 0$ the model back to classic regression model

3.1.1 Maximum Likelihood Estimation for (SAR) Model [7][2][11]

Maximum likelihood estimation of the SAR model described here and in Anselin (1988) involves maximizing the log likelihood function (concentrated with respect to β and $\sigma 2$, the noise variance associated with ε) with respect to the spatial parameter λ one of the most important ways for estimation is maximum likelihood estimation because it gives the accurate value for a parameter from among several possible estimates in SAR model, in a time series the depend variable is time while the time not be problem for OLS estimation but for spatial regression model the OLS is not suitable because the *e* and WY is not independent from other

$$Y = X\beta + \lambda W\underline{Y} + e \dots\dots(2)$$

$$Y = (I - \lambda W)^{-1} X \beta + (I - \lambda W)^{-1} e$$

$$e \sim N(0, \sigma^2)$$

$$e = (I - \lambda W)\underline{Y} - X\beta \dots\dots(3)$$

Thus, the log likelihood function for y of the spatial lag model is obtained by adding the term ln |I-λW| to the log likelihood function of the standard regression model

$$LnL\left(\beta, \lambda, \frac{\sigma2}{y}, x\right) = \frac{-n}{2}Ln2\prod -\frac{n}{2}Ln\,\sigma2 + \,Ln\,|I - \lambda W| - \frac{1}{2\sigma2}(e'e) \quad Where$$

$$e = \,Y - \lambda WY - X\,\beta$$

$$LnL\left(\beta, \lambda, \frac{\sigma2}{y}, x\right)$$

$$= \frac{-n}{2}Ln2\prod -\frac{n}{2}Ln\,\sigma2 + \,Ln\,|I - \lambda W| - \frac{1}{2\sigma2}(Y - \lambda WY - X\,\beta)' + (Y$$
$$- \lambda WY - X\,\beta\,)$$

On account of this correction the MLE estimates will differ from the OLS estimates. They coincide for λ=0 where the spatial lag model approaches the standard regression model.

In this step we get the derivative for β, σ^2 in log of likelihood and equal to zero we get

$$b_{mle} = (X'X)^{-1} + X'A\underline{Y}.....(4)$$

Let;

$$A = (I - \lambda W)$$
$$= (X'X)^{-1} + X'(I - \lambda W\,)\underline{Y}$$
$$b_o = (X'X)^{-1} + X'\underline{Y}.....(5)$$
$$b_L = (X'X)^{-1} + X'W\underline{Y}......(6)$$

$$b_{mle} = b_o - \lambda \ b_L \dots (7)$$

$$e_{mle} = Y - \lambda \ W\underline{Y} - X(b_o - \lambda \ b_L) \dots (8)$$

$$e_o = \underline{Y} - X \ b_o \dots (9)$$

$$e_L = W\underline{Y} - X \ b_L \dots (10)$$

$$e = e_o - \lambda \ e_L \dots (11)$$

According the first order condition the error variance can be estimation by

$$\sigma^2_{MLE} = \frac{(e_o - \lambda e_L)^{'} \ (eo - \lambda \ e_L)}{n} \dots (12)$$

b_o:Vector of the regression parameter Y for X

b_L:Vector of regression parameter   WY for X

e_o:Vector of another regression model

e_L: Vector of another regression model WY for X

by putting the value of ($\sigma^{^2}{}_{MLE}$, $\beta^{^2}{}_{MLE}$) in a likelihood function we get

$$LC = \frac{-n}{2} Ln[\frac{1}{n}(e_o - \lambda e_L)^{'} (e_o - \lambda e_L)] + \ln |I - \lambda W| \dots (13)$$

3.2 Spatial Error Model (SEM) [1][5]:

One of the most important violations that necessary in regression model is independence of the error term , so it will be studied in this model .along time ago and in classical models we learned that error or model errors must be independence

but in SEM it is reverse the errors must be correlated with other and aims of this correlation in spatial error model (SEM) to correction spatiality in error .

$$Y = X\beta + \underline{u} \dots (14)$$

$$\text{or} \quad \underline{u} = (I - \theta W)^{-1} \underline{e} \dots (15)$$

Where   W: is the spatial weights (n*n) matrix

θ :is the spatial parameter

if $\theta \neq 0$ we have a spatial dependency between the errors for neighbor observation

if $\theta = 0$ we go to the classical regression model and not have any spatial dependency between errors for neighbor observation

3.2.1Maximum Likelihood Estimation for (SEM) Model: [6][8]

The spatial error model attention to θ where show the correlation between the residuals :

Then the Maximum likelihood estimation for (SEM) model is :

$$Y = X\beta + \underline{u}$$

$$\underline{u} = \theta W \underline{u} + \underline{e}$$

$$\text{or} \quad \underline{Y} = X\beta + \theta W \underline{u} + \underline{e} \dots (16)$$

And put the value of error in a likelihood function

$$L\left(\beta, \theta, \frac{\sigma^2}{Y}, X\right) = \frac{-n}{2}\text{Ln }2\prod -\frac{n}{2}\text{Ln }\sigma^2 + \text{ Ln }|I - \theta W| - \frac{1}{2\sigma^2}(Y - X\beta)'(I - \theta W)'(Y - X\beta)(I - \theta W)$$

In this step we get the derivative for β, σ^2 in log of likelihood and equal to zero we get

$$Y^* = (I - \theta W)\underline{Y} \quad .....(17)$$

and $X^* = (I - \theta W)(X)......(18)$

$$b_{mle} = (X^{*'}X^*)^{-1}X^{*'}Y^*....(19)$$

put the value for X* an d Y* we get

$$b_{mle} = [X'(I - \theta W)'(I - \theta W)X]^{-1}X^{\backslash}(I - \theta W)'(I - \theta W)\underline{Y}......(20)$$

$$e = Y - Xb_{mle}......(21)$$

$$\sigma^{\wedge 2}{}_{mle} = \frac{e'e}{n}......(22)$$

After putting the value of ($\sigma^{\wedge 2}{}_{MLE}, b_{MLE}$) in likelihood function we get

$Lc = -\frac{n}{2}\ln[\frac{1}{n}e'(I - \theta W)'(I - \theta W)e] + \ln |I - \theta W|.......(23)$

4. Weight Matrix:[13].

It is square matrix which is elements positive value and denoted by W not necessary to be symmetric while create this matrix based on neighboring, relation neighboring from location for another location in same row in the rows of 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 a Appropriate weight matrix and there for some way to create this matrix

4-1 Methods to Create Weight Matrix

1.Rook Contiguity: the elements are equal to 1 if the two areas neighbor by limited and have relation between the two areas in any side so other is zero. This matrix is used so much than other

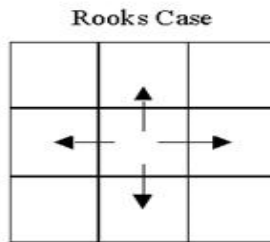**Figure (1) Show the rook weight matrix**

2.Bishop Contiguity: Neighbor would occur if the two areas connected by point and this point is the is smallest area neighbor then the value will be equal to 1 and other value is zero in this type of matrix
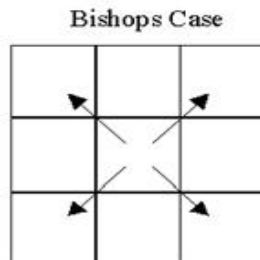
Figure (2) Show the bishop weight matrix

3.Queen Contiguity: This matrix gets its elements from the sum of (rook) and (bishop) matrixes elements and neighbor in this type is based on connect by point or connect by limited. [13]
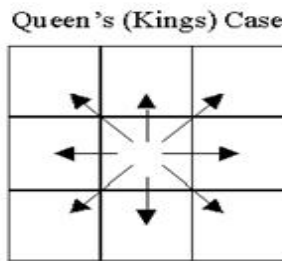
Figure (3) Show the queen weight matrix

4-2 Binary Contiguity Weights Matrix: [1]

The matrix is positive and square $(n \times n)$ matrix ,if i,j contiguous $w_{ij=}$ 1and if not contiguous w_{ij} =0

4-3 Row - Standardized Weights Matrix [13]:

In this matrix sum of row equal to 1 and this matrix depend on (Binary Contiguity Weights Matrix)

$$W_{ij}^{std} = \left\{ \begin{array}{ll} \dfrac{w_{ij}}{\sum_i w_{ij}} & \text{i neighbor j} \qquad 0 < W_{ij}^{std} \le 1 \\ 0 & \text{otherwise} \end{array} \right\}$$

5 Moran's Test [10][3][15][5]:

There are several tests to detect spatial dependence and Moran test is one of them this test is measure to show that we have spatial dependency in data or not before creating models it is a general measure depends on the (GLM) model which indicates that the near things more of a relationship than far things any phenomenon related to each other phenomenon .

$$I_M = \frac{n(e'We)}{S_0(e'e)} \dots (24)$$

$S_0 : \sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij}$: sum of every element in W matrix

where we using row – standardized so sum of row equal to 1 in this case ($n = S_o$) that is work to simplify the above formula for the follows

$$I_M = \frac{e'We}{e'e}$$

Hypotheses test for Moran's

Null hypotheses $\quad\quad$ H_0: $\lambda = 0$, $\theta = 0$ no have spatial dependency

Alternative hypotheses $\quad$ H_1: at least one of $\lambda \neq 0$ or $\theta \neq 0$ spatial dependency is exist

$$Z = \frac{I - E(I)}{\sqrt{V(I)}} \dots (25)$$

$$E(I) = \frac{n(tr(MW))}{S0(n - k)}$$

929

$$V(I) = \frac{tr(MWMW`) + tr(MW)^2 + (tr(\,MW))^2}{(n-k)\,(n-k+1)} - E(I)^2 \ldots\ldots(26)$$

M = I − X (X′X)$^{-1}$ X′: it is Idempotent Matrix that is square and symmetric

tr: Sum of diagonal element

k: Number of explanatory variables

comparison between calculate Z value with Z table value for(α, two sided test) where Moran test is significant that is mean relation between geographic location exist and using spatial regression model allowed so we have spatial dependency

6 Lagrange Multiplier test or Lag Test (LM)

6 -1 Lagrange Multiplier (LM) Lag Test for (SAR) Model [10][3]:

Test of Lagrange multiplier test is more use than Moran test because Moran is use only to test spatial dependency does exist or not and we cannot find by Moran test what is the alternative model for (GLM) model but Lagrange is tell we which is the model is alternative (SAR) or (SEM) and by this test occur the model you created is significant or not assumption for this test is

Hypotheses test for Lagrange (SAR) model:

H$_0$: $\lambda = 0$ spatial dependence exist

H$_1$: $\lambda \neq 0$ at least one of the λ does not equal to 0 spatial dependence not exist

Where reject null hypotheses is mean spatial dependence not exist and accept alternative hypnoses where alternative hypnoses is mean exist spatial dependence and model alternative is (SAR) model while the rule show in down and sometime denoted by (LM-SAR)

$$LM_\lambda = \frac{\left(\frac{e`WY}{S^2}\right)^2}{D} \text{.....(27)}$$

$$D = \frac{(WXb)`M(WXb)}{S^2} + tr(W'W + WW) \text{......(28)}$$

$$S^2 = \frac{e`e}{n} \text{.....(29)}$$

S^2: Variance of error for general linear model regression

We compare the calculate value with table value for $\chi^2(1,\alpha)$ after that we decide to the hypotheses

6-2 Lagrange Multiplier (LM) error Test for (SEM) Model [10][9]:

Hypotheses test for Lagrange (SEM) model

H_0: θ = 0   spatial dependence exist in error

H_1: θ ≠ 0   at least one of the θ does not equal to zero 0spatial dependence not exist in error

Where reject null hypotheses and accept alternative hypotheses that is mean spatial dependency does exist in error and alternative model is spatial error model (SEM) and the rule is show in dawn where some time denoted by (LM-SEM)

$$LM_\theta = \frac{\left(\frac{e`We}{S^2}\right)^2}{T} \dots.(30)$$

$$T = tr[(W + W')W \dots\dots(31)$$

S^2: variance for general linear model regression

Comparison between (LM_θ, LM_λ) with χ^2 value table byonce degree freedom and once level significant where Lagrange test for (SAR) or (SEM) significant or exist spatial dependency in each of them we must go to robust test and robust rule for (SAR) and (SEM) model is:

$$Robust - LM_\lambda = \frac{\left[\left[\frac{e`WY}{S^2}\right] - \left[\frac{e`We}{S^2}\right]\right]^2}{D - T} \dots\dots (32)$$

$$Robust - LM_\theta = \frac{\left[\left[\frac{e`WY}{S^2}\right] - \frac{T}{D}\left[\frac{e`We}{S^2}\right]\right]^2}{T - \frac{T^2}{D}} \dots\dots(33)$$

$$D = \frac{(Wxb)`M(WXb)}{S^2} + tr\left(W'W + WW\right) \dots\dots(34)$$

$$T = tr(W'W + WW) \dots\dots(35)$$

Comparison each (robust-LM $_\lambda$, robust-LM $_\theta$)with χ^2table value foronce degree freedom and once level significant.

7. Comparison Criteria for Choosing the Best Model

The choice of a specific model from a range of models style is an important part of the analysis data as it leads us to choose the best model , where the use of certain statistical criteria which are as described below

7-1 Root Mean Squares Error[3]**:**It is a square root of the sum of squares errors divided by the (n-k-1) is calculated for each models, the model which is the root mean square error smaller it is the best model. The formula for calculating this standard is

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{n-k-1}}\ldots\ldots\ldots(36)$$

7-2Mean Absolute Percentage Error (MAPE) [3]**:**It is calculated by dividing the sum of the absolute value of the error divided by the number of observation (n) While, the smaller value is the best modeland the formula for calculating is

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}}{y_i}\right|\ldots\ldots\ldots(37)$$

7-3 Akaike Information Criterion(AIC) [14]**:**This standard introduced for the first time in 1973 , this standard criterion is equal to twice number of parameters minus twice log maximum likelihood function of the observations and is expressed mathematically as the following

$$\text{AIC} = -2\log(L) + 2K\ldots\ldots(38)$$

L: Great value for the value of the logarithm (M.L.E) K: Number for the parameters of the model

the corrected Akaike criterion (AIC) , which takes the formula described in the below if the sample size was small ((n / k)<40) it is better to use the standard corrected Akaike

$$\text{AIC}_c = \text{AIC} + \frac{2K(K+1)}{(n-k-1)} \ldots\ldots(39)$$

7-4 Adjusted Determinations of Coefficient [2]: The coefficient of determination R^2 is not only a good indicator to show the quality of explanatory variables because the value of the coefficient of determination increases with each additional variable, so using Adjusted Determinations of Coefficient which consider for the number of explanatory variables is better than R^2so ,calculated by the following formula

$$R^2_{adj} = 1 - \left[\frac{\frac{\sum_{i=1}^{n}(y_i-\hat{y})^2}{n-k}}{\frac{\sum_{i=1}^{n}(y_i-\hat{y})^2}{n-1}}\right]\ldots\ldots(40)$$

For comparison models the best model is the models which have bigger R^2_{adj}value

8. Practical Part

This part covers the practical aspects of this paper, applying the data practically to the statistical Models SAR and SEM for estimating regression parameters and comparing the results between different models. The data used in this paper are

taken from sulaimani meteorological organization and seismology. The sample consists of (27) places (observations) (every observations are a yearly average) about the (Atmospheric Pressure, Air Temperature, Relative Humidity, wind Speed) of Kurdistan Region (Sulaimaniyah, Erbil, and Dhouk) stations .the geoda,minitab,matlab and excel are used to analyze the data

$$\underline{Y} = X\underline{\beta} + \underline{e}$$

Where $\underline{Y} = (Y_1, Y_2, ..., Y_{27})'$ is an (27 ×1) observation vector, $\underline{\beta} = (\beta_0, ..., \beta_3)'$ is an (4 ×1) vector of unknown parameters, X is an (27 × 4) matrix of full column rank

A.P (Atmospheric Pressure) =β₀+β₁W.S(Wind Speed) +β₂ A.T(Air Temperature (%))+β₃ (Relative Humidity (%))R.H +e_i.....(41)

Table (1) Observed parameter and standard error in classic regression

Model	Unstandardized Coefficients		Standardized Coefficients	T	Sig.P value
	B	Std. Error	Beta		
Constant	6.083	0.781		7.791	0.000
W.S	-0.196	0.101	-0.281	1.948	0.064
A.T	0.121	0.027	0.712	4.501	0.000
R.H	2.847	0.754	0.612	3.774	0.001

Table (1) Show that the p-value of wind speed (W.S) is greater than level of significant (α = 0.05) this mean that the explanatory variable (W.S) is not significant and the p-value of other explanatory variables (A.T and R.H) are less than (α = 0.05) that means

that the explanatory variables (A.T and R.H) are statistically significant, so they remain in the model.

The fitted regression model:

$$A.P = 6.083 + 0.121 \, A.T + 2.847 \, R.H$$

9 Test for Problem Econometric and Assumption of Regression Model

9-1 Test of Hetroscedasticity

Now consider to test the heteroscedasiticity brush-pagan&Konker-basset

Table (2) Diagnostics for heteroskadasticity random coefficients test

Test	D.F	Value	P-value
Breusch-pagan	3	1.1553	0.7637
Konker-Basset	3	1.7544	0.6249

Table (2) Show that the P-value of the test are greater than ($\alpha = 0.05$),this indicate that the null hypothesis of homogeny is accept or the data will have not problem of hetroskadasticity

9-2 Test of Autocorrelation

consider to testing the autocorrelation using the Durbin-watson test

Table (3) Show the interval accept for Durbin test

D.W	D.U	4-D.U
1.698	1.65	2.35

The value of D.W statistical test is 1.698 then value is between 1.65≤1.698≤4-1.65=2.35 then the data has not the Auto correlation problem. The level of (D.W) is greater than (du=1.65) and less than(4-du) .This indicates that there is not any serious autocorrelation among the variables, and so the null hypothesis is accepted.

9-3 Test of Multicollinearity

For test the multicollinearity using the variance inflation factor

Table(4) Show the test of Multicollinearity

Factors	Tolerance	Variance inflation factor
Wind Speed	0.919	1.089
Air Temperature	0.762	1.313
Relative Humidity	0.725	1.379

The output of Table (4) show that the (VIF) of the variables (W.S) and (A.T) and (R.H) are near to one, which indicate that there are no multicollinearity problem among the variables.

9-4 Test of Normality

Consider for testing the standard residuals using jarque-Bera test i.e

H_0: The distribution of residual standard is very close to normal standard

H_1: The standard residual are significantly different from the normal standard

The null hypothesis H_0 is not rejected if the P-value for jarque-Bera test is greater than 0.05 , otherwise rejects H_0 and fails to reject the alternative hypothesis H_a.

Table (5) Show test of normality

Test	D.F	Value	P-value
Jarque-Bera	2	1.0108	0.6032

Table (5) Shows that the P-value for residuals using (jarque-Bera)test is 0.6032 this value is greater than($\alpha = 0.05$), then we accept the null hypothesis. This indicates that residuals are normally distribution

10.Moran Test:

Before building a model we must test the data for find spatial dependency of the data is exist or not because if the spatial dependency not exist we return to general model and we can't use spatial models.

Null hypotheses $\qquad$ H$_0$: $\lambda = 0$, $\theta=0$ having no spatial dependency

Alternative hypotheses $\qquad$ H$_1$: at least one of $\lambda \neq 0$ or $\theta \neq 0$ spatial dependency is exist

Table (6) Show the Moran test of spatial dependency for Rook, Bishop and Queen Matrix

Weight Matrix (W)		
Rook	Bishop	Queen
2.3119**	0.1332	2.3250**

Based on Table (6) the value of Moran's test compared with $Z_{0.025}$ =1.96 so the spatial dependency is significant in rook and queen matrix.

11.Building models according matrixes

11-1 Building SAR model according matrixes

11-1-1 SAR Model by using Rook Matrix : According the table below positive spatial parameter (λ) indicated that the neighbor places have similarity according weather and Atmospheric Pressure, while in SAR model where using the rook weight matrix the variable wind speed is not significant.

Table (7) Estimation the parameter of SAR model by using rook matrix

Model	Weight Matrix- Rook(WMR)			
	Coefficient	Std. Error	t-calculate	t- table
(Constant)	6.3189	0.7626	8.2852**	
b_1	-0.1834	0.09841	1.8635	
b_2	0.1113	0.0262	4.2335**	2.052
b_3	2.5050	0.7369	3.39906**	
(λ)	0.0141			

The SAR model is estimated as:$Y^{\wedge}$ =6.3189+0.1113b_2+ 2.5050b_3+0.0141 λ

11-1-2 SAR Model by using Bishop Matrix : According table below the positive spatial parameter (λ) indicated that the neighbor places have similarity according weather for Atmospheric Pressure,so in SAR model where using the bishop weight matrix only the variable wind speed is not significant

Table (8) Estimation the parameter of SAR model by using bishop matrix

Model	Weight Matrix- Bishop(WMB)			
	Coefficient	Std. Error	t- calculate	t-table
(Constant)	6.0688	0.7792	7.7877**	2.052
b_1	-0.1995	0.1005	1.9840	
b_2	0.1221	0.0267	4.5587**	
b_3	2.8149	0.7530	3.7381**	
(λ)	0.0047			

The SAR model is estimated as: $\hat{Y}$ = 6.0688 + 0.1221 b_2+2.8149 b_3+0.0047 λ

11-1-3 SAR Model by Queen Matrix: The table below show that a positive spatial parameter (λ)is positive that is indicated the neighbor places have the same weather only the wind speed isn't significant too

Table (9) Estimation the parameter of SAR model by using queen matrix

Model	Weight Matrix Queen(WMQ)			
	Coefficient	Std. Error	t-calculate	t-table
(Constant)	6.4880	0.7384	8.7855**	
b_1	-0.1944	0.0952	2.0401	
b_2	0.1096	0.0253	4.3180**	2.052
b_3	2.0739	0.7136	2.9062**	
(λ)	0.0240			

The SAR model is estimated as: $\hat{Y}$ = 6.4880 + 0.1096 b_2+2.0739 b_3+0.0240 λ

11-2 Building SEM model according matrixes

11-2-1 SEM Model by using Rook Matrix: The table below shows positive spatial parameter (θ) so that is mean the condition correlation between the errors exists in SEM model by this types of matrix , the variable wind speed is not significant too

Table (10) Estimation the parameter of SEM model by using rook matrix

Model	Weight Matrix-Rook(WMR)			
	Coefficient	Std. Error	t-calculate	t- table
(Constant)	6.3462	0.7875	8.0585**	
b1	-0.1599	0.1016	1.5736	
b2	0.1097	0.02706	4.0529**	2.052
b3	2.5834	0.7609	3.3948**	
(θ)	0.2763			

The SEM model is estimated as: $Y^{\wedge}$ = 6.3462 +0.1097b_2+2.5834 b_3+0.2763 θ

11-2- 2SEM Model by using Bishop Matrix : The table below show the negative spatial parameter (θ) so that is mean the condition correlation between the errors not exists in SEM model by bishop matrix and this types of matrix not use because not satisfy the condition , while the variable wind speed is not significant

Table (11) Estimation the parameter of SEM model by using bishop matrix

Model	Weight Matrix-Bishop(WMB)			
	Coefficient	Std. Error	t-calculate	t- table
(Constant)	6.0867	0.7806	7.7965**	
b1	-0.1967	0.1007	1.9526	
b2	0.1206	0.0268	4.4945**	2.052
b3	2.8453	0.7543	3.7716**	
(θ)	-0.0096			

The SEM model is estimated as: $Y^{\wedge}$ = 6.0867 + 0.1206 b_2+2.8453 b_3-0.0096 θ

11-2-3 SEM Model by using Queen Matrix: The table below show the positive spatial parameter (θ) so that is mean satisfy correlation between errors, and the variable wind speed is not significant.

Table (12) Estimation the parameter of SEM model by using queen matrix

Model	Weight Matrix-Queen (WMQ)			
	Coefficient	Std. Error	t- calculate	t-table
(Constant)	6.2555	0.7890	7.9275**	
b_1	-0.1479	0.1018	1.4526	
b_2	0.1123	0.02712	4.1407**	2.052
b_3	2.6142	0.7624	3.4284**	
(θ)	0.2630			

The SEM model is estimated as: $\hat{Y} = 6.2555 + 0.1123\, b_2 + 2.6142\, b_3 + 0.2630\, θ$

12-1 Lagrange Multiplier Test for SAR (LM λ):

This test is used for finding alternative and significant model in spatial autoregressive model (SAR) and spatial error model (SEM)

Table (13) Lagrange multiplier test for SAR Model

Test	Weight Matrix (W)		
	Rook	Bishop	Queen
LM λ	7.4654**	0.086381	9.65248**
Robust LM λ	6.7468**	0.087028	8.7016**

In rook matrix: The values of the two tests (LM λ=7.4654, Robust LM λ=6.7468) are significant when we compared the value of the tests with the value of chi-square (1,α) degree of freedom χ2 (1,0.05)=3.841 in bishop matrix the value of the two test (LM λ=0.086381, Robust LM λ=0.087028)are non-significant when we compared the value of tests with the value of chi-square (1,α)degree freedom i.e χ $^2_{(1,0.05)}$ =3.841 are not significant and in queen matrix The values of the two test (LM λ=9.65248, Robust LM λ=8.7016) are significant when we compared the value of tests with the value of chi-square (1,α) degree freedom i.e χ $^2_{(1,0.05)}$ =3.841

12-2 Lagrange Multiplier Test for SEM (LM θ):

This test is used for finding spatially dependency in spatial error model with finding alternative model

Table (14) Lagrange multiplier test for SEM Model

Test	Weight Matrix (W)		
	Rook	Bishop	Queen
LM θ	1.203816	0.00027	1.035462
Robust LM θ	0.486513	0.000956	1.986976

In table (14), the values of the two test LM θ and robust LM θ according to the weight matrix rook, bishop and queen are not significant because the values of the two tests are less than the value of Chi-Square with degree of freedom χ ^{2}(1,0.05)=3.841 we can concluded that the SAR Model is better than the SEM model so SAR is alternative model.

13 Comparison Criteria according matrixes:

After create a model by using three types of weight matrix and test it For find alternative model we must use same criteria for finding the best model according matrixes

Table (15): Comparison between the models according matrices

DData	Models	Weight Matrix	R^2_{adj}	AIC	RMSE	MAPE
Data	SAR	Rook	0.5468	3.9415	0.2903	0.000803
		Bishop	0.5266	3.956	0.2967	0.000837
		Queen	0.5747	3.9206	0.2811	0.000748

Result and discussion:

According to the criteria the best and suitable model is SAR for QUEEN matrix because for R^2_{adj} it have bigger value according queen than other types of matrixes ,so for (AIC,RMSE,MAPE) the value is smaller according queen too. That is indicate the best model is SAR for QUEEN matrixes

14 conclusion and recommendation

14-1 Conclusions

After using various methods of estimating parameters of the regression model and using some Criteria to produce the best fitted model for A.P, the following conclusions are drawn

1. The results of data analysis show that the rook and queen matrixes is significant in test for spatial dependency Moran test and by Lagrange test SAR model for rook and queen is significant too while the bishop is not significant in both Moran and Lagrange test

2.The results in practice part shows the positive of spatial parameter (λ=0.0141,0.0047,0.0240) according matrixes in SAR model appear the correlation between places for Atmospheric Pressure

3. According Lagrange test (LM θ =1.203816,0,000027,1.035462) for matrixes SEM model is not significant which means the comparison between two models SAR and SEM not allowed, however in SEM model for bishop matrix the spatial parameter (θ) is negative which indicates the errors is not correlate with other and the condition not satisfy

4. SAR model is the most appropriate model for queen weighted matrix according criteria's

$$\text{Atmospheric pressure}(\mathbf{y_i}) = \mathbf{6.4880} + \mathbf{0.1096}\,(\text{Air}$$

$$\text{Temperature}) + \mathbf{2.0739}\,(\text{Relative Humidity}) + \mathbf{0.0240}\,\lambda$$

5.The explanatory variable wind speed is not significant in classic regression and spatial regression but the variables air temperate and relative humidity are significant

14-2 Recommendations

 The authors recommend the following points:

1.Use those models for any aspects of life such as the cancer disease according to places or study the rank of university according to the place and decide which university is most efficient than the others.

2.Use another type of weight matrix instead of rook bishop and queen such as linear weight matrix for finding the neighbor places.

3.Calculate the parameters by using the panel model whichis correlation between time and place. when using the panel data, the researcher can work in time and place simultaneously.

Appendix A : Matlab Code for Finding Spatial Parameter and Concentrated Likelihood Function for SAR Model

```
n = 27;
eO =[ input eO ];
eL = [ input eL];
eOt = eO';
eLt = eL';
I27 = eye(27);

W =[input weight matrix];
x=-1:0.0001:1;
LC=x;
rho=-0.9999;
```

```matlab
for j=1:20000

    if j<10000

.

.

        LCmax=LC(1);

    end

    if LC(j)>LCmax

        rho=d;

        LCmax=LC(j);

    end

end

    plot(x,LC)

    grid on;

xlabel(' \rho');

ylabel('ln L( \rho)');

'rho = '; rho;'ln Lc(rho) = '; Lcma
```

REFERENCES:

1. philip,viton (2010). "Notes on spatial economtric model" Retrieved from https://www.yumpu.com/en/document/read/3858779/notes-on-spatial-econometric-models-the-ohio-state-universitdownloaded in5/12/2019

2. Samprit.Bertam (1997). "*Regression analysis by example".* new york: john wily and sons.

3. Gan, j (2009). "*Spatial combenation interpolation model based on panal data and its empirical study".* china: fujian institute.

4. Danel H.p A griffith " Non-standard Spatial Statistics and Spatial Econometrics " Retrievedfromhttps://books.google.iq/books?id=wk439dKDxDIC&lpg=PR6&ots=OIN1zdYBSD&dq=Some%20Notes%20on%20Spatial%20Statistics%20And%20Spatial%20Econometrics%E2%80%99%E2%80%99&lr&pg=PR2#v=onepage&q=Some%20Notes%20on%20Spatial%20Statistics%20And%20Spatial%20Econometrics downloaded in24/4/2016

5. J. P.LeSage (1998). "Spatial Econometrics Retrieved from"https://www.spatial econometrics.com/html/wbook.pdf downloaded in7/5/2016

6.J. P.LeSage (2004). "*Maximum likelihood estemation of spatial regression model" .* lucas U.S.: uneversity of toledo.

7.J. P.LeSage (1997)."*Regression analysis of spatial data ".* university of toledo.

8.zongji , Y (2007)."*A spatial econometric analysis on the relation shiop btween power consumptionand regional economic devlopment".* china: chongqing unversity.

9.G.S ,Maddala (2003)."*Introduction to econometric" .* new york: john wiley.

10. L , Nselin (2001). "*Spatial econometrics".*London : blackwell.

11. j.K Ord (1975). "*Estimation methods for models of spatial interaction".* china: chongqing.

12.Timothy J Ross (2016). Retrieved from https://www.wiley.com/en-iq/Fuzzy+Logic+with+Engineering+Applications%2C+4th+Edition-p-9781119235866 downloaded in12/10/2015

13.KarstenRusche (2010). " Quality of Life in the Regions: An Exploratory Spatial Data Analysis for West German Labor Markets" Retrieved from https://mpra.ub.uni-muenchen.de/19759/ downloded in 1/8/2015

14.K. P Urham (2002). *model selection and multimodel infrence .* newyork: springer-verlag.

15. سوسنقاسم , (2014) ."أقضيةالعراقللعام تقديرنماذجالأنحدارالحيزيلنسبالفقرفيمجلة العلوم الاقتصادية 2012 عدد 79

پوخته:

بـەکارهێنـانـە گرنگەکـانی چـەمـاوەی شـوێنی لـەهـەموو بوارەکانی ژیـانـدا وایکردووەلـەم تۆژینەوەیەداەبابەتی کەشـوەهەوادابەکاربهێنرێت .داتا و زانیارییەکانکۆکراوەتەوەلە 27 شـوێن یانوێسـتگەلەهەرێمیکوردسـتان (سـلێمانی ,هەولێر,دهۆك)کە ئەم شـوێنانە ئامێری پێوانە کردنی رەگەزەکانی کەش و هەوایان تێدا بووە وەچەماوەکانی شـوێن بەکاردێت بۆدۆزینەوەی کاریگەری ئەوشـوێنـانـەی کـەدراوسـێن.چـەمـاوەیخۆیشـوێن SAR وەچـەمـاوەیهەلـەشـوێن SEM کەهەردووکیانپارامیتەرشـوێنیانتێدایە. بەکارهێنانی هەندێکپێوەری بەراوردی جیاواز وەك (R2adj ,RMSE , MAPE, AICc) بۆدیاریکردنی باشترین مۆدێل پاشان پشکنینی مۆرانبەکارهاتووە بۆئەوەی بزانین داتاکەکاریگەری شـوێنی لەسـەرە .هەروەهاپشکنینی لاگرانج بەکارهاتووە بۆ دۆزینەوەی رێژەی قبووڵ بوونی مۆدێـل لـەنێـوان مۆدێلـەکانی(SAR,SEM) لـەگـەڵ سـێ جۆر لـە مـاتریکسـی (nپپrook,bishop,qu). لە گرنگترین ئەو ئەنجامانەی لەبەشـی پراکتیکدا دەرچووە چەماوەکانی کاتی خۆیی (SAR) بۆماتریکسـەکانی (nپپrook,qu) قبوڵدەکرێت وچەماوەی هەلـەیشـوێن(SEM) قبووڵنـاکرێت بۆهـەر سـێ جۆرەکـەی مـاتریکس وەبـاشـترین مۆدێـل بریتیـەلـە (SAR) بەبەکارهێنانمـاتریکسی هاوسەنگی (queen)

دراسة الضغط الجوي في اقليم كردستان باستخدام تقنية الانحدار الحيزي (SAR and SEM)

الملخـص:

ان التطور الحاصل في التحليل الحيزي في كل مجالات الحياة ادى الى استخدامه في هذه البحث لتحليل الطقس في اقليم كوردستان (السليمانية , أربيل ,دهوك) وبناءا على البيانات المسجلة من 27 موقع والذي يحتوي على

الاجهزة الخاصة لتسجيل البيانات لكل شهور عام2015 .استخدام اختبار موران لتحليل اعتماد مكاني للبيانات وثم استخدام نماذحج الانحدار الذاتي الحيزي (SAR) ونموذج الخطأ الحيزي (SEM) التي تضم كل منها عامل الحيزية واستخدام معايير المقارنة (R^2_{adj} , RMSE , MAPE, AICc) لاختيار افضل النماذج .وتم استخدام اختبار لاكرانج لايجاد مستوى المعنوية وايجاد نموذج بديل (SEM, SAR) مع ثلاث انواع من المصفوفات (rook,queen,bishop) .

ومن اهم الاستنتاجات بعد اجراء التحليل للجانب التطبيقي كانت نتائج تحليل نماذج الانحدار الحيزي الذاتي باستخدام مصفوفتي (rook,queen) معنوي ونموذج الخطأ الحيزي (SEM) غير معنوي لكل انوع المصفوفات الثلاثة .أفضل نموذج هو الانحدار الذاتي الحيزي(SAR) باستخدام مصفوفة الموزون (queen)