2. Methdology

2.1. Models

Three models- Ordinary least square model (OLS), spatial lag model (SLM) and spatial error model (SEM) are applied in this study with the spatial data to compare their performance and accuracy. The ordinary least square model can be written as

$$Y_i = \beta_i X_i + \xi, \tag{1}$$

where Y_i is the number of RTAs as the dependent variable for $i = 1,2,\ldots,n$, β_i is a vector of parameters, X_i is a matrix of independent variables and, ξ is a vector of unobserved error terms that assumed to be distributed normally $N(0,\sigma^2)$. The estimates of parameters β, in matrix notation, [10] can be given as

$$\hat{\beta}_i = (X^t X)^{-1} X^t Y, \tag{2}$$

where X^t is the transpose of data matrix, X, and Y is the observed vector of the dependent variable.

In this study, we evaluate two spatial models SLM and SEM, where the neighbouring (spatial) effect for a region i is considered to be affected by the other neighbouring regions $j, j = 1,2,\ldots,n$ [11]. Let us define the spatial dependence as

$$Y_i = f(Y_j), \quad i,j = 1,2,\ldots,n; i \neq j, \tag{3}$$

where Y_i is the natural logarithm of the number of RTAs in i spatial units. However, this study applied two spatial models (SLM and SEM) compared with the OLS model as given in Eq (1).

The spatial lag model (SLM) models can be written as

$$Y_i = \rho W_y + \beta_i X_i + \xi, \tag{4}$$

where Y_i is the vector of natural logarithm of RTAs for ith spatial unit (region), ρ is the spatial lag coefficient, W_y is the spatial weight matrix, X_i is the matrix of independent variables, β_i is the vector of parameters and ξ is the unobserved error terms vector.

The spatial error model (SEM) can be written as

$$Y_i = \beta_i X_i + \mu, \quad \text{with} \quad \mu = \lambda W_\mu + \xi, \tag{5}$$

where μ is the function of unobserved error terms, λ is a spatial error coefficient and W_μ is the spatial error weight matrix. The spatial weights are distinct elements in any cross-sectional analysis of spatial dependence. They are necessary components in the setting up of spatial autocorrelation statistics and provide calibration between units. For both SLM and SEM, the spatial weight matrix W can be defined as

$$W = \begin{bmatrix} W_{11} & W_{12} & \ldots & W_{1n} \\ W_{21} & W_{22} & \ldots & W_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ W_{n1} & W_{n2} & \ldots & W_{nn} \end{bmatrix}. \tag{6}$$

2.2. Data

Data for this study have been gathered from multiple sources. RTA data have been collected from the published reports by Royal Omani Police [12]. Other secondary data on different factors such as population, area, density, etc. from the publication of the National Centre for Statistics & Information (NCSI) in Oman [13, 14, 15, 16]. Unemployment statistics were provided by the Public Authority of Manpower Register in Oman.

In this study, the number of road traffic accidents in Oman in 2017 is considered as the dependent variable and the independent variables are density, population, number of vehicles, number of job seekers (unemployment), number of accidents caused by the speed in 2017 and number of accidents occurring during seasonal months (May, June, July 2017). There are eleven Governorates in Oman which represent the spatial units in this study.

3. Results and Discussion

In this study, three models have been applied which are ordinary least square (OLS), spatial lag model (SLM) and spatial error model (SEM) to the RTA data of Oman for the year 2017. In the regression models, the number of road traffic accidents in different Governorates (regions) in Oman is considered as the dependent variable. The independent variables in this study are density, population, number of vehicles, number of job seekers (unemployment), number of accidents caused by speed in 2017 and number of accidents occurring during seasonal months (May, June, July 2017) in the Sultanate of Oman. There are eleven Governorates in Oman are representing the spatial units in this study. The results are showed in Table 1 that are obtained from the fitting of the three models– OLS, SLM and SEM.

The results showing in the Table 1 gives the comparison of the estimated parameters and their significance levels ($* : P < 0.05$, $** : P < 0.001$, $*** : P < 0.0001$). The effects of two variables (number of speed accidents and seasonal accidents) obtained are similar in all three models while there is a clear difference in density, population and number of vehicles. Spatial error model gives very close coefficient to ordinary least square to the number of job seekers variable.

For checking the adequacy of models, the residual plots can be used. The residuals of the three models were also evaluated as displayed in Fig 1. This suggests that the SLM for RTAs data has higher accuracy than others. To select the most suitable model, Akaike information criteria (AIC) and Bayesian information criteria (DIC) are the two major criteria widely used including spatial data analysis [11].

Table 1: Results of parameter estimates in different spatial models for RTA data.

Dependent variable	OLS	SLM	SEM
Spatial lag coefficient,ρ	-	0.061498	-
Spatial error coefficient,λ	-	-	-1.5208
Constant	-1.96E+01**	-3.76E+01***	-2.06E+01***
Density	4.90E-01*	2.39E-01***	5.11E-01***
Population	2.85E-05	-3.76E-05***	4.40E-05***
Number of vehicles	-3.99E-04**	-2.20E-04***	-4.04E-04***
Number of job seekers	-1.10E-03	-1.07E-04	-1.20E-03*
Number of speed accidents	6.64E-01**	6.00E-01***	6.53E-01***
Number od seasonal accidents	2.56E+00***	2.76E+00***	2.53E+00***

Level of signficance: $*: P < 0.05$, $**: P < 0.001$, $***: P < 0.0001$

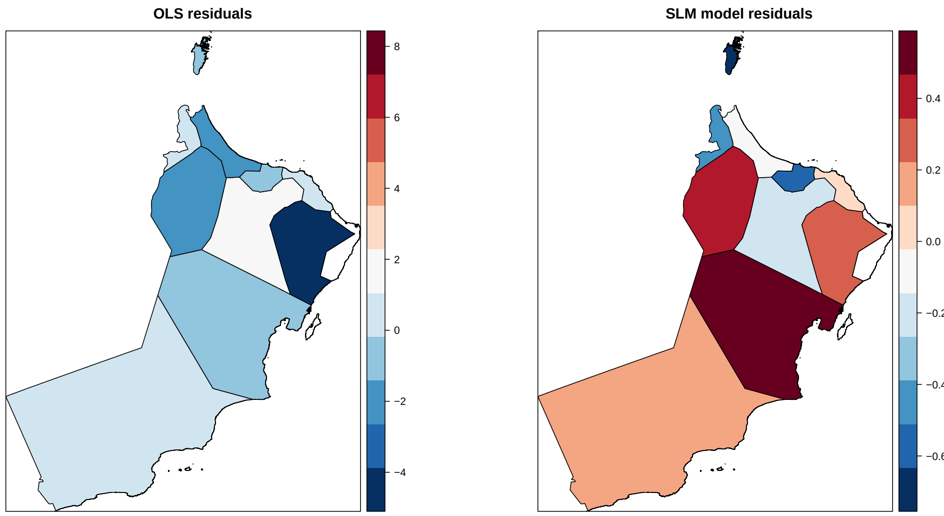

Fig. 1: Models residuals through Oman's Governorates.

As showing in Table 2, the values of the log-likelihood, AIC and BIC of OLS, SLM and SEM models were compared in order to determine the best model. The most interesting finding is that the SLM outperforms the SEM due to best values of log-likelihood $= -6.34$, AIC $= 30.69$ and BIC $= 34.27$ while the corresponding values for SEM are -25.78, 69.55 and 73.13, respectively. Therefore, the SLM model found to be the best to identify associated factors for the road traffic accidents in Oman.

Table 2: Evaluation of different spatial models for RTA data.

Diagnostic Criteria	OLS	SLM	SEM
log-likelihood	-28.18	-6.34	-25.78
AIC	72.37	30.69	69.55
BIC	75.54	34.27	73.13

4. Conclusion

The main goal of the current study is to compare the spatial lag model (SLM) and the spatial error model (SEM) with road traffic accidents (RTA) data. The study evaluated and diagnosed three models, which are OLS, SLM and, SEM. The study showed that the similar impact of three models OLS, SLM and SEM in two variables; the number of speed accidents and seasonal accidents. In contrast, it was shown a clear difference on the effect of the variables such as density, population and number of vehicles in all three models. One of the significant findings to emerge from this study is that the SLM outperformed the SEM due to the best values in log-likelihood, AIC and, BIC. However, study [5] found the SEM has higher accurancy than SLM model for RTA with different spatial dataset while considered smaller spatial units. Whilst this study did not conform to the finding of the study by [5], this study yet offers some model selection techniques for spatial analysis.

Acknowledgments

The authors would like to thank the Academic Skill Tutors at the Staffordshire University Library for their support. We also thank the Royal Omani Police (ROP), Authority of Manpower Register and the National Centre for Statistics & Information NCSI in the Sultanate of Oman for providing support to gather data.

References

[1] H. Al Reesi, J. Freeman, J. Davey, S. Al Adawi, and A. Al Maniri, “Measuring risky driving behaviours among young drivers: Development of a scale for the Oman setting,” *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 55, pp. 78–89, 2018.

[2] F. Mannering and C. Bhat, “Analytic methods in accident research: methodological frontier and future directions,” *Analytic Methods in Accident Research*, vol. 1, pp. 1–22, 2014.

[3] Z. Li, W. Wang, P. Liu, J. M. Bigham, and D. R. Ragland, “Using geographically weighted poisson regression for county-level crash modeling in california,” *Safety science*, vol. 58, pp. 89–97, 2013.

[4] S. Barua, K. El-Basyouny, and M. T. Islam, “Effects of spatial correlation in random parameters collision count-data models,” *Analytic Methods in Accident Research*, vol. 5, pp. 28–42, 2015.

[5] K.-A. Rhee, J.-K. Kim, Y.-I. Lee, and G. F. Ulfarsson, “Spatial regression analysis of traffic crashes in seoul,” *Accident Analysis & Prevention*, vol. 91, pp. 190–199, 2016.

[6] M. A. Quddus, “Modelling area-wide count outcomes with spatial correlation and heterogeneity: An analysis of london crash data,” *Accident Analysis & Prevention*, vol. 40, no. 4, pp. 1486–1497, 2008.

[7] J. Aguero-Valverde and P. Jovanis, “Analysis of road crash frequency with spatial models,” *Transportation Research Record: Journal of the Transportation Research Board*, no. 2061, pp. 55–63, 2008.

[8] B. Lu, P. Harris, M. Charlton, and C. Brunsdon, “The gwmodel r package: further topics for exploring spatial heterogeneity using geographically weighted models,” *Geo-spatial Information Science*, vol. 17, no. 2, pp. 85–101, 2014.

[9] A. A. N. Al-Maniri, H. Al-Reesi, I. Al-Zakwani, and M. Nasrullah, “Road traffic fatalities in Oman from 1995 to 2009: evidence from police reports,” *International Journal of Preventive Medicine*, vol. 4, no. 6, p. 656, 2013.

[10] C. Brunsdon, A. S. Fotheringham, and M. E. Charlton, “Geographically weighted regression: a method for exploring spatial nonstationarity,” *Geographical analysis*, vol. 28, no. 4, pp. 281–298, 1996.

[11] L. Anselin, “Model validation in spatial econometrics: a review and evaluation of alternative approaches,” *International Regional Science Review*, vol. 11, no. 3, pp. 279–316, 1988.

[12] R. O. Police, *Facts and Figures*. Director General of Traffic, 2017.

[13] NCSI, *Monthly Statistical Bulletin April 2017*. National Centre for Statistics & Information, Sultanate of Oman, 2017.

[14] ——, *Monthly Statistical Bulletin July 2017*. National Centre for Statistics & Information, Sultanate of Oman, 2017.

[15] ——, *Monthly Statistical Bulletin September 2017*. National Centre for Statistics & Information, Sultanate of Oman, 2017.

[16] ——, *Monthly Statistical Bulletin January 2018*. National Centre for Statistics & Information, Sultanate of Oman, 2018.