International Journal of
Science and Engineering Investigations vol. 8, issue 95, December 2019
Received on December 2, 2019 ISSN: 2251-8843

Modeling Spatial Relationships for Crashes Involving Teen Drivers

Hemin J. Mohammed[1], Steven D. Schrock[2]
[1]Ph.D., University of Kansas Transportation Center (KUTC), University of Kansas, Lawrence, Kansas, 66045
[2]Ph.D., P.E., ACTAR, F.ITE, Department of Civil, Environmental & Architectural Engineering, University of Kansas, Lawrence, Kansas, 66045
([1]hemin@ku.edu, [2]schrock@ku.edu)

Abstract- In the United States, the risk of vehicle crashes is higher among teens than among any other age group. Most previous studies investigated the effects of demographic differences and nonspatial factors associated with crashes such as gender, age, driving under the influence of drugs or alcohol, the presence of passengers, and distractions. This research was conducted to model spatial relationships between teen-related crashes and factors that significantly influence the number of these crashes using an ordinary least squares (OLS) model and geographically weighted regression (GWR). The data of seven years (2010-2016) of crashes involving drivers aged 15-19 that occurred in Kansas were investigated using ArcGIS Pro Software. From 18 candidate exploratory variables, two statistically significant exploratory variables were used to build a predictive model using OLS and GWR. The two exploratory variables were the miles of rural non-state roads and the number of passenger cars in counties. The predictive model showed that the number of crashes involving teen drivers was expected to be lower by more than three percent by 2026. Additional independent variables could have been examined to reveal their association with the number of crashes but the unavailability of related data prevented further examination.

Keywords*-Teen Driver, Spatial Analysis, Traffic Safety, Ordinary Least Squares (OLS), Geographically Weighted Regression (GWR)*

I. INTRODUCTION

Globally, around 1.25 million people died in 2013, as a result of road traffic crashes, and up to 50 million people sustained non-fatal injuries. Unfortunately, the number of road traffic deaths has increased by 13 percent from 2000 to 2013. Therefore, road traffic injuries represent a major threat to the world population, especially for teens, where it is the main cause of death among people aged 15–19 in the world [1, 2]. The root of this problem is not new. In the first annual international symposium of youth enhancement service in 1995, Simpson [3] stated that the traffic crashes involving teen drivers aged 16 to 19 “have been a worldwide road safety and public health concern for several decades.”

In the United States, the risk of vehicle crashes is higher among teens than among any other age group [4]. The motor vehicle traffic crashes in the US are a leading cause of death for young people aged 16-20 years since 2001 [5-10]. In 2015, 2,333 teens were killed meaning six teens died every day from road traffic injuries, and this number increased by 3.6 percent in 2016 [11]. Per mile driven, teen drivers are nearly three times more likely to be in a fatal crash than drivers age 20 and older [12]. Hersman and Rosekind listed the factors that put young drivers at the highest risk: driving under the influence (DUI) of alcohol and drugs, speeding, and seatbelt usage. These factors have a higher prevalence of male than female drivers [13].

In Kansas, the Kansas Department of Transportation (KDOT) [14] has determined teen drivers aged 14 to 19 as one of the foci of KDOT’s Strategic Highway Safety Plan (SHSP). After several years of improving metrics, it appears that overall teen crashes have begun to increase in the past few years. The number of crashes involving teen drivers in Kansas increased by six percent from 2013 to 2016. This concerning trend shows a need to identify the associated factors for this group in order to better target safety improvements. The objective of this study included modeling spatial factors that contribute to changes in the number of all type crashes involving teen drivers aged 14 to 19 from 2010 to 2016 using ArcGIS Pro software. The crash data were extracted from: KDOT traffic crash database; the Fatality Analysis Reporting System (FARS) database; and the U.S. Census Bureau database while the spatial layers were downloaded from Esri, USGS, and KDOT websites. The number of crashes involving this age group during the study period extracted from the KDOT crash database that used for this study consisted of 72,656 crashes.

II. METHODOLOGY

The spatial statistics in ArcGIS are a set of techniques in different toolsets used for describing and modeling spatial distributions, spatial patterns, processes, and relationships [15]. The modeling spatial relationships toolset provides modeling, examining, and exploring spatial relationships among features using regression analysis to better understand the contributing factors behind observed spatial patterns or to predict spatial outcomes [16]. Regression analysis attempts to answer most of the *why* and/or *what* questions such as: why are the expected traffic crashes involving teen drivers exceptionally high in particular locations in Kansas? Or what are the potential factors that make some areas have more than the expected traffic crashes involving teen drivers? The salient tools were considered in this toolset were ordinary least squares (OLS), and geographically weighted regression (GWR). The output of the OLS is a single equation that best describes the data relationships between a response variable and each one of the explanatory variables in the study area. However, the GWR is a local model that creates an equation for every feature in the dataset and the coefficients in the model rather than being global estimates specific to a targeted location [16, 17]. The GWR is treated in this research as a spatial disaggregation of the OLS. This is–the OLS was used to produce a single equation for the whole study area (Kansas) while the GWR was used to provide different equations for each county (as a unit of analysis) in the study area but with the same independent variables of the OLS.

Since the OLS and GWR are both linear regression methods, the relationship between all of the explanatory variables and the dependent variable needs to be linear; otherwise, the resultant model will perform poorly. The dependent variable (CRASH) in this study is the number of crashes involving teen drivers during the study period. A scatter plot matrix graph was used to clarify the relationships among the proposed variables. The variables that had nonlinear relationships or curvilinear relationships were treated by transforming their values using square roots or logarithmic transformations such as the Common Logarithm (log: a logarithm with base 10) and/or Natural Logarithm (ln: a logarithm with base e). For instance, the dependent variable (CRASH) and the exploratory variable (PC) were transformed by applying ln and log to their values while the exploratory variable (RD) was transformed using the square root and ln.

III. ORDINARY LEAST SQUARES (OLS) RESULTS

Shults et al. [18] define the OLS as a global model that creates a single equation that best describes the data relationships between a response variable and each one of explanatory variables in the study area. The output of the OLS is a single equation that best describes the data relationships between a response variable and each one of the explanatory variables in the study area. Several related variables were prepared for modeling. The selected variables depended on their relativity to the study topic, the availability, and accessibility to the targeted variables. The scope of this topic made obtaining the desired variables a challenging task.

This global model was used to create a single equation that describes the relationship between the dependent variable (the number of traffic crashes involving teen drivers from 2010 to 2016) and each of the explanatory variables. There were 18 exploratory variables that were examined by the exploratory regression in order to select appropriate variables for the OLS model. Table 1 shows the first outcome of the exploratory regression, which includes the threshold criteria and also the number of trials and number and percentage of time that the trials passed the criterion cutoff. These models were listed based on the number of exploratory variables and then the models that had the highest adjusted R-squared results. However, not all the listed models were satisfied with all the threshold criteria. Therefore, investigating the significance of each exploratory variable was the next step to select proper variables for more in-depth investigations.

TABLE I. PERCENTAGE OF SEARCH CRITERIA PASSED

Search Criterion Cutoff	Trials	Passed	Passed (%)
Min Adjusted R-Squared > 0.50	11,706	11,684	99.81
Max Coefficient p-value < 0.05	11,706	323	2.76
Max VIF Value < 7.50	11,706	1,865	15.93
Min Jarque-Bera p-value > 0.10	11,706	64	0.55
Min Spatial Autocorrelation p-value > 0.10	28	24	85.71

The significance of the exploratory variables (Table 2) defines how statistically significant each variable was during analyzing every possible combination in the Significant (%) column and how stable variable relationships were by examining the Negative (%) and Positive (%) columns. The strong candidate variables were those variables that were significant over 50 percent of the time [19]. Accordingly, the first six variables (Table 2) were selected, and they are listed below:

- The average number of passenger cars (PC);
- Miles of rural non-state roads in a county;
- The population of teens (TN);
- The population of counties (P);
- Number of high schools (HS); and
- Average DVMT on all types of roads (DVMT).

However, the only model that includes these variables and satisfies the VIF, Jarque-Bera p-value, and adjusted R-squared threshold criteria were the model that contained:

- Miles of rural non-state roads in a county (RD); and
- The average number of passenger cars (PC).

Therefore, these two explanatory variables are the only variables that qualified to be in the reduced OLS and GWR models.

OLS was applied using the two exploratory variables that passed most of the significant threshold criteria of the exploratory regression. The number of traffic crashes involving teen drivers was related to the average number of passenger cars (PC) and the number of miles of rural non-state in a county (RD). The rural non-state roads comprise all routes that are not Interstate, US and Kansas routes located outside city limits with a population under 5,000 people. Passenger cars are road motor vehicles, excluding motorcycles, intended for the carriage of passengers and designed to seat no more than eight persons plus the driver [20].

The statistical report (Table 3) shows both the Multiple R-Squared and Adjusted R-Squared values were higher than 90 percent, which showed a strong correlation in model performance. The Adjusted R-Squared value of 0.91 indicates that the model explains approximately 91 percent of the variation in the dependent variable. The resultant model is shown in (Table 4). The most critical parameters in the table are Coefficient, Probability (p-value), and VIF. Both of the coefficients have a positive relationship with the dependent variable, which is the number of crashes involving teen drivers. Thus, the more rural non-state road miles and the higher number of passenger cars, the more crashes involving teen drivers are expected. The p-value shows that the exploratory variables are statistically significant for the model.

TABLE II. SUMMARY OF VARIABLE SIGNIFICANCE FROM THE EXPLORATORY REGRESSION

No.	Explanatory Variable	Significant (%)	Negative (%)	Positive (%)
1	Average number of passenger cars (PC)	100.00	0.00	100.00
2	Miles of rural non-state roads in a county (RD)	94.26	0.49	99.51
3	Population of teens (TN)	91.45	0.00	100.00
4	Population of counties (P)	74.49	23.95	76.05
5	Number of high schools (HS)	52.04	100	0.00
6	Average Daily Vehicle Miles Traveled on all types of roads (DVMT)	51.55	0.00	100.00
7	Average number of non-commercial trucks (TK)	38.89	24.67	75.33
8	Average DVMT on rural non-state roads (DVMT_RD)	30.96	22.19	77.81
9	Population of over 15 in the labor force (L15)	30.24	43.8	56.20
10	Number of workers over 15 commuting to work (CW)	29.78	37.83	62.17
11	Population of males over 15 in the labor force (M15)	25.86	66.50	33.50
12	Population of 18-24 under high school degrees (P_HS)	24.24	77.91	22.09
13	Miles of all types of roads (A_RD)	23.26	26.62	73.38
14	Population of females over 15 in the labor force (F15)	22.63	40.16	59.84
15	Average precipitation in inches (PCT)	15.14	3.23	96.77
16	Average household income (INCOM)	11.19	96.44	3.56
17	Number of postsecondary schools (POST_S)	9.72	59.54	40.46
18	Number of families below the poverty level (POV)	1.31	44.11	55.89

TABLE III. THE STATISTICAL REPORT OF THE OLS REGRESSION

Multiple R-Squared	0.9101	Adjusted R-Squared	0.9084
Joint F-Statistic	516.4001	Prob.(>F), (2,102) dof	<0.0001*
Joint Wald Statistic	774.8932	Prob.(>chi-squared), (2) dof	<0.0001*
Koenker (BP) Statistic	8.6422	Prob.(>chi-squared), (2) dof	0.0133*
Jarque-Bera Statistic	178.9214	Prob.(>chi-squared), (2) dof	<0.0001*

An asterisk next to a number indicates a statistically significant p-value ($\alpha = 0.05$)*

TABLE IV. THE RESULTANT MODEL FROM THE OLS REGRESSION

Variable	Intercept	RD	PC
Coefficient	-1.065522	0.019208	1.805782
Std. Error	0.056927	0.007131	0.060281
Probability	<0.00001*	0.008261*	<0.00001*
Robust SE	0.078023	0.007178	0.065401
Robust Pr.	<0.00001*	0.008683*	<0.00001*
VIF	--------	1.086884	1.086884

An asterisk next to a number indicates a statistically significant p-value ($\alpha = 0.05$)*

The resultant model from the OLS regression takes the form shown in Equation 1:

$$CRASH = e^{10^{\left(-1.0655+0.0192\left(\frac{\sqrt{RD}}{\ln RD}\right)+1.8058\,(\log(\ln PC))\right)}} \quad (1)$$

In Table 3, the Koenker (BP) Statistic is statistically significant at $\alpha = 0.05$, and similarly, the robust probability (Table 4) is statistically significant. Therefore, the null hypothesis is rejected and there is a nonstationary condition in the model, which is expected as mentioned before. That is, the relationships between the number of crashes involving teen drivers and exploratory variables change across the study area. One or both of the exploratory variables might be a significant predictor of the number of crashes involving teen drivers in some counties, but perhaps a weak predictor in other counties.

The VIF values were less than 7.5, which means the variables were inconsistent in predicting the number of crashes.

The Joint F-statistic and Joint Wald Statistic p-values (Table 3) supported that the model was statistically significant. The OLS residuals (Figure 1) indicated the over predictions in blue and under predictions in red. Since the Jarque-Bera Statistic's p-value was statistically significant, the null hypothesis was rejected. This means there was heteroscedasticity because of influential outliers in the data, as shown in the map and residual plot (Figure 1 and Figure 2), respectively. The red-colored county on the map is the red dot on the scatterplot, which represents Chase County. This indicates that the model underpredicted the number of crashes involving teen drivers in Chase County and the actual number was larger than the model predicted. However, the blue-colored counties represent the counties where the model overpredicted the number of crashes, which means in these counties the actual numbers were smaller than the model predicted.

OLS regression models the relationships between dependent and independent variables precisely when they were consistent across the study area, but when these relationships were heterogeneous and nonstationary across the study area, the regression equation created an average of the mixed relationships present. The dominant method that deals with the regional variation and that eliminates their impact is the GWR regression model.

IV. GEOGRAPHICALLY WEIGHTED REGRESSION (GWR) RESULTS

The GWR is a local model that creates an equation for every county in the state. In other words, the OLS used every single county in Kansas to calibrate the resultant equation, but the GWR models the nonstationary relationships over the study area so that each county gets a separate OLS equation calibrated based on the neighboring counties while using the same explanatory variables applied in the OLS model. Therefore, the coefficients of the exploratory variables were different for each county in the study area.

The GWR was applied similarly to the OLS. The number of traffic crashes involving teen drivers was entered as the dependent variable and both miles of rural non-state roadways in a county (RD), and the average number of passenger cars (PC) were used as explanatory variables. The GWR tool produces an attribute table that contains coefficients, local R-Squared, residuals, and some other parameters. Each of these parameters could be mapped to visualize their impact on the study area. The coefficient of the average number of passenger cars is shown in (Figure 3). The dark areas show where the coefficient values were large and these were the locations having the strongest relationship between the number of passenger cars variable and the number of crashes involving teen drivers. The resultant map of the other coefficient (miles of rural non-state roads) is shown in (Figure 4). The dark areas show where the coefficient values were large and they represent a strong indicator for the number of crashes

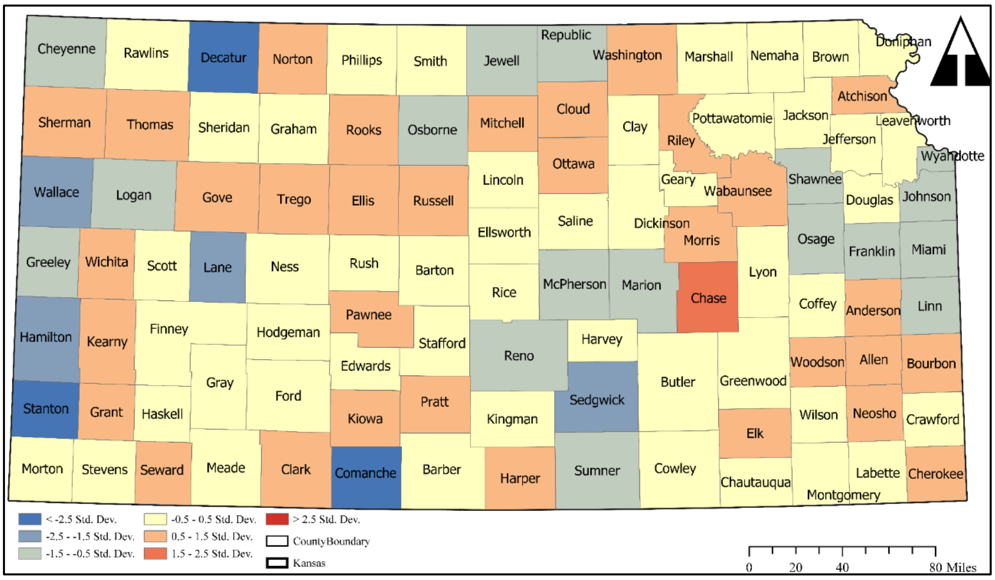

Figure 1. The OLS Mapped Residuals

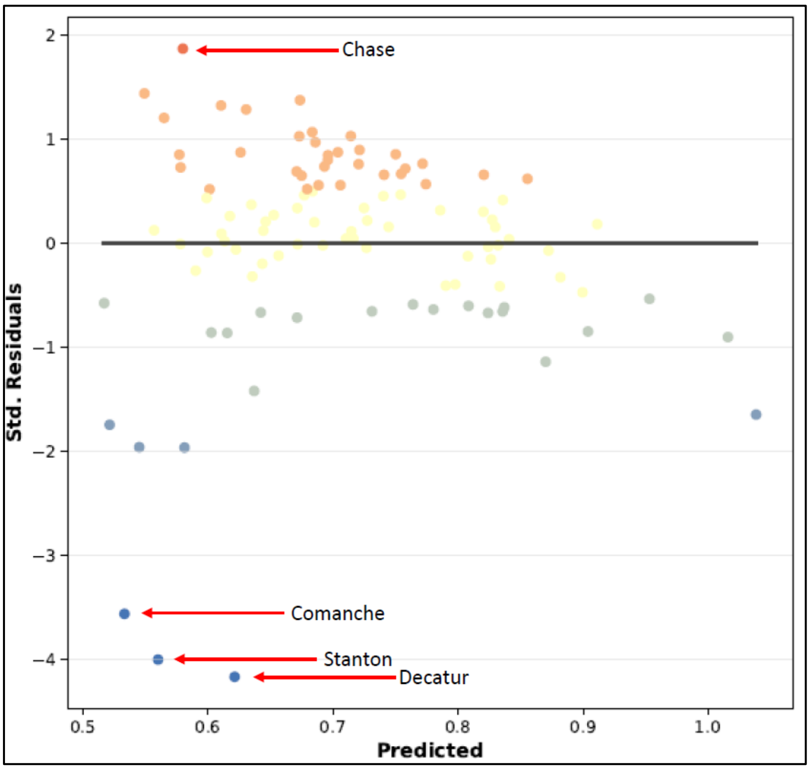

Figure 2. The OLS Residual vs. the Predicted Dependent Variable

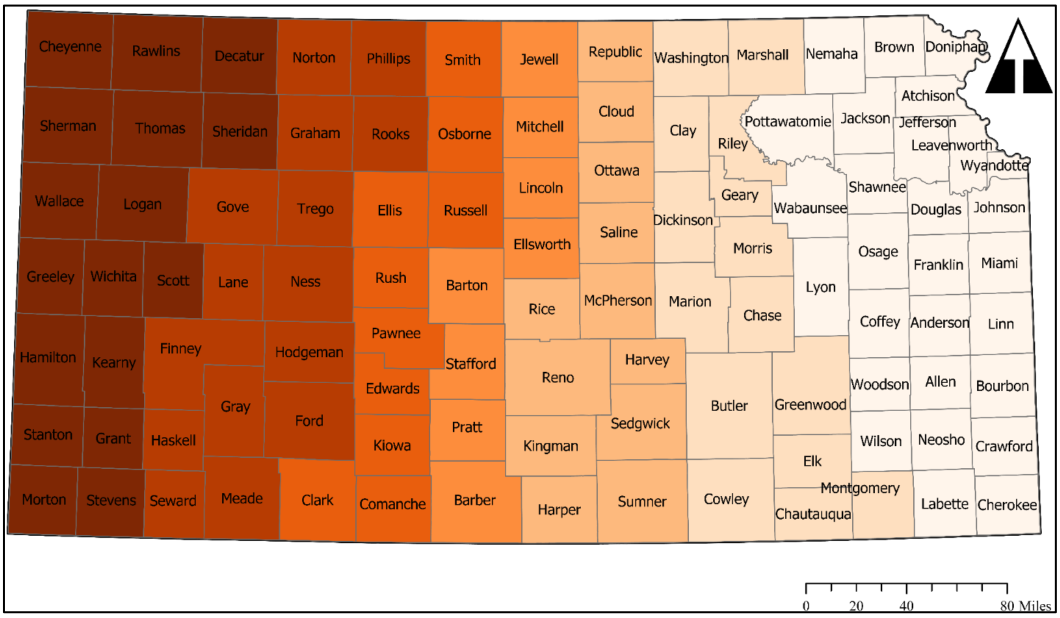

Figure 3. The Coefficient of Passenger Cars

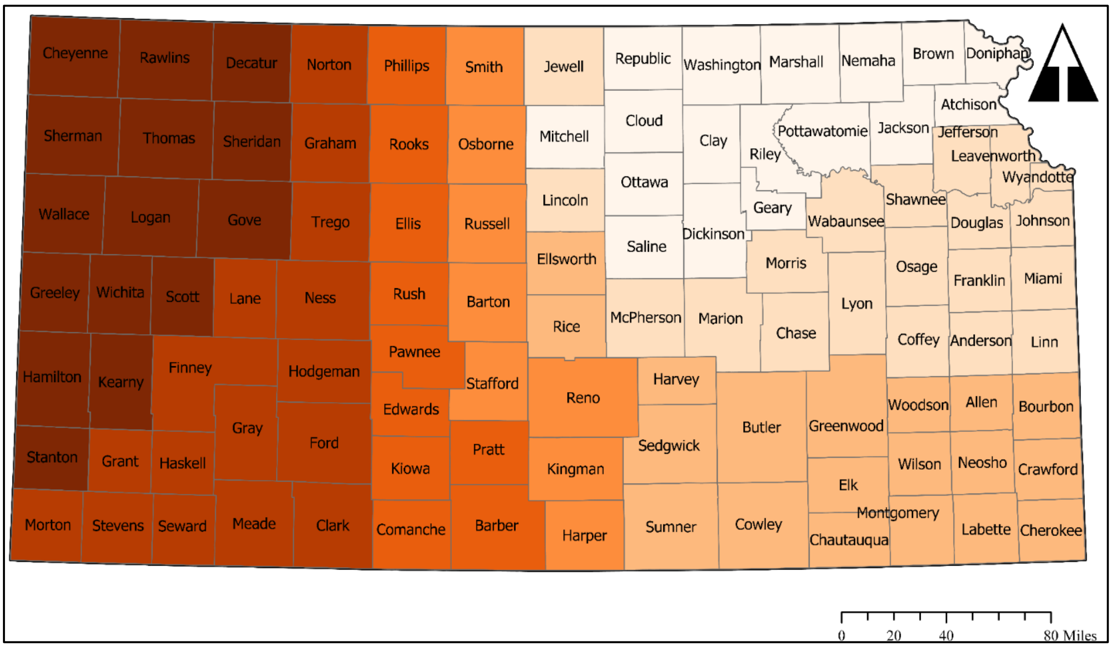

Figure 4. The Coefficient of Rural Non-State Roads

V. PREDICTION BY MODELS

The OLS and GWR models were used to predict the number of crashes involving teen drivers in 2026. Based on the available number of passenger cars and miles of non-state roads data from 2007 to 2016, the growth rates of these variables for 2026 were calculated for each county. The statewide average growth rates for the number of passenger cars and miles of non-state roads were 0.002 percent and 0.047 percent, respectively. The resultant growth rates were used to predict the number of passenger cars and miles of non-state roads in 2026. Consequently, the predicted values were entered into the OLS and GWR models in order to predict the number of crashes involving teen drivers in 2026 for each county.

Based on the best available data, assuming nothing else changes in the state generally and counties specifically, the number of expected crashes based on the OLS model was predicted to be 10,824 statewide. However, the GWR model predicted the crash number in 2026 to be 10,795 crashes in the state. Given that the number of crashes involving teen drivers in 2016 was 11,172 the OLS and GWR models predict a 3.11 percent and 3.37 percent crash reduction, respectively. This predicted reduction is due to slight downward trends of growth rates in the two exploratory variables in the most populated counties (which also have the highest number of crashes) such as Douglas, Johnson, Leavenworth, Shawnee, and Wyandotte. At the county level, the models predict the future trend of each county. This provides a useful indicator for related parties to identify where they need to target their resources. For instance, the number of crashes that occurred in Shawnee County in 2016 was 860 crashes and the GWR model predicted that this number will increase to 983 (14.3%) by 2026.

VI. VALIDATION OF MODELS

Validation is a process to test the performance of model prediction when applied to an independent dataset that was not used in the modeling. The independent dataset used was the number of crashes involving teen drivers, the number of registered passenger cars, and miles of non-state roads in each county of Kansas in 2017. The number of registered passenger cars and miles of non-state roads dataset was entered into the models as an exploratory variable, to predict the number of crashes involving teen drivers in 2017 in each county. Accordingly, the predicted numbers were compared to the real number of crashes. The validation step was performed for both OLS and GWR models.

For the OLS model the intercept, coefficient of miles of non-state roads, and coefficient of the number of registered passenger cars were fixed for all counties at -1.065522, 0.019208, and 1.805782, respectively, but residuals were different based on counties. The results were overestimated for some counties and underestimated for others. The overall predicted number of crashes was underestimated by 3.66 percent (411 crashes). That is_ the total number of predicted crashes was 10,801 crashes, whereas the number of crashes involving teen drivers that occurred in Kansas in 2017 was 11,212 crashes.

However, the GWR's prediction number of crashes was overall better than the OLS's prediction number. The total predicted number by the GWR model was 10,883 crashes, which means it underestimated the crashes by 2.94 percent (329 crashes). Since each county in the GWR model had its own equation, intercept, coefficients, and residuals they were used separately to predict the number of crashes in each county.

At the county level, the model estimation for the number of crashes involving teen drivers is shown in (Table 5). The table shows that the prediction of the OLS and GWR models were off by less than one percent for six counties, underestimated for four counties and overestimated for two counties. However, for 13 or 14 counties (depending on the model used), the estimated number of crashes was off by more than 50 percent. The reason of these differences between the predicted and actual number of crashes is not clear, and it could be caused by different factors, such as unusual weather or traffic patterns in those counties in 2017 compared to 2010-2016. An unusually high level of roadway construction or some other one-time event could also have been a factor.

TABLE V. THE NUMBER OF COUNTIES UNDERESTIMATED OR OVERESTIMATED FOR MODELS

Percentage	OLS		GWR	
	Underestimated	Overestimated	Underestimated	Overestimated
< 1%	4	2	4	2
(1-4.9)%	3	6	3	6
(5-9.9)%	16	9	16	9
(10-24.9)%	18	21	17	21
(25-49.9)%	6	6	8	6
> 50%	4	10	3	10
Total	51	54	51	54

However, the counties that had been overestimated or underestimated by more than 25 percent were generally counties that had a low number of crashes. When the predicted number was off by a few crashes, the percentage of variance increased dramatically. For instance, the number of crashes in Rawlins County was five crashes in 2017 while the predicted number of crashes was 7.65 crashes, which means it was overestimated by 53 percent, but the numerical difference between the actual number and the predicted number was only 2.65.

Furthermore, among the 30 counties that had the highest number of crashes, only two counties (Jefferson and Wyandotte) had the predicted number of crashes off by more than 25 percent. It was not clear why the predicted number of crashes in Jefferson County was off by 26 percent. Further analysis on Wyandotte County revealed that only 5.89 miles of non-state roads were reported in the list or county roadway miles provided by KDOT, but a brief review of the county's map [21] revealed that there are many more miles, which clearly shows that there is an error in the dataset for the non-state miles. If the correct number were available, it is believed that the predicted number of crashes would be much closer to the actual number.

VII. CONCLUSIONS

The OLS and GWR tools were used to determine the statistical association factors behind observed spatial patterns of teen-related crashes and to predict the number of crashes involving teen drivers in each county in Kansas. The 18 related exploratory variables that were prepared for modeling and projected to provide a better understanding of associated factors to the number of crashes involving teen drivers, only two were found to be statistically significant and were used to build the predictive OLS and GWR models. The two exploratory variables were the number of miles of rural non-state roads in a county and the number of passenger cars in a county. With OLS a single model was built to represent the entire state, while with GWR a separate model was created for each county in the state. The OLS and GWR models were used to predict the number of crashes involving teen drivers in the future for each county based on the growth rates of the exploratory variables. Assuming that no other global changes happened which could influence the number of teen-related crashes, the models predicted a three percent reduction in the number of crashes, statewide by 2026.

This research provides a useful indicator for related parties to identify where they can target their resources in order to improve teen driver safety. For instance, the number of crashes that occurred in Shawnee County in 2016 was 860 crashes and the GWR model predicted that this number will increase to 983 (an increase of 14.3 percent) by 2026 if nothing else was changed. Unavailability of teen-related data, such as vehicle-miles traveled by teen drivers, the number of passenger cars was driven by teens, and the number of licensed teen drivers in each county was evident limitations that if addressed could improve the utility of future research of this type.

ACKNOWLEDGMENTS

The authors would like to thanks the Kansas Department of Transportation for providing access to the motor vehicle crash database and offering assistance whenever necessary.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: H. Mohammed, S.D. Schrock; data collection: H. Mohammed; analysis and interpretation of results: H. Mohammed, S.D. Schrock; draft manuscript preparation: H. Mohammed, S.D. Schrock. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

[1] WHO, World Health Organization, *Youth and Road Safety*. Geneva, Switzerland: World Health Organization, 2007, p. 47.

[2] WHO, World Health Organization, "World Health Statistics 2017," in *Monitoring Health for the SDGs, Sustainable Development Goals* vol. CC BY-NC-SA 3.0 IGO, ed. France, 2017, p. 116.

[3] H. Simpson, *New to the road: Reducing the risks for young motorists*. Youth Enhancement Service, Brain Information Service, UCLA School of Medicine, 1996.

[4] L. Lonero and D. Mayhew, "Large-scale evaluation of driver education: Review of the literature on driver education evaluation 2010 update," *Washington, DC: AAA Foundation for Traffic Safety*, 2010.

[5] U.S. Department of Transportation. (2006). *HS-809 956, Race and ethnicity in fatal motor vehicle traffic crashes 1999-2004*. Available: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/809956

[6] U.S. Department of Transportation. (2015). *DOT HS 812 203, Motor vehicle traffic crashes as a leading cause of death in the United States, 2010 and 2011*.

[7] U.S. Department of Transportation. (2012). *DOT HS-811 620, Motor vehicle traffic crashes as a leading cause of death in the United States, 2008 and 2009*.

[8] U.S. Department of Transportation. (2016). *DOT HS 812 297, Motor Vehicle Traffic Crashes as a Leading Cause of Death in the United States, 2012–2014*.

[9] NHTSA, National Highway Traffic Safety Administration, "Motor Vehicle Traffic Crashes as a Leading Cause of Death in the United States, 2015," vol. DOT HS 812 499, USDOT, Ed., ed: NHTSA, 2018.

[10] R. Subramanian, "Motor vehicle traffic crashes as a leading cause of death in the United States, 2001," *Young*, vol. 1, p. 3, 2005.

[11] NHTSA, National Highway Traffic Safety Administration, "2016 Fatal Motor Vehicle Crashes: Overview," in *Traffic safety facts research note* vol. 2017, ed: NHTSA, 2017, pp. 1-9

[12] IIHS, Insurance Institute for Highway Safety. (2016, November 12, 2017). *Fatality facts: teenagers 2015*. Available: http://www.iihs.org/iihs/topics/t/teenagers/fatalityfacts/teenagers

[13] D. A. Hersman and M. R. Rosekind, "The road to zero deaths from motor vehicle crashes," *JAMA internal medicine*, vol. 177, no. 12, pp. 1717-1718, 2017.

[14] KDOT. (2015). *Strategic Highway Safety Plan 2015*. Available: http://www.ksdot.org/Assets/wwwksdotorg/bureaus/burTrafficSaf/reports/reportspdf/SHSP.pdf

[15] Esri Events. (2017, May 20, 2017). *Using Spatial Statistics to do More: Simple Approaches* [Presentation]. Available: https://www.esri.com/videos/watch?videoid=3d_8nQpSCgE

[16] M. M. Fischer and A. Getis, *Handbook of applied spatial analysis: software tools, methods and applications*. Springer Science & Business Media, 2009.

[17] 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.

[18] R. A. Shults, E. Olsen, and A. F. Williams, "Driving among high school students-United States, 2013," *MMWR. Morbidity and mortality weekly report*, vol. 64, no. 12, pp. 313-317, 2015.

[19] Esri Events. (2018, April 20, 2019). *Beyond Where: Modeling Spatial Relationships and Making Predictions* [Presentation]. Available: https://www.youtube.com/watch?v=5_tbFFeYXWM

[20] O. The Organisation for Economic Co-operation and Development, ,. (2013, November 30, 2019). *Passenger Car*. Available: https://stats.oecd.org/glossary/detail.asp?ID=3524

[21] KDOT, Kansas Department of Transportation. (2015, May 14, 2019). *General Highway Map of Wyandotte County*. Available: http://wfs.ksdot.org/arcgis_web_adaptor/rest/directories/arcgisoutput/County/halfInch/WyandotteCounty.pdf

How to Cite this Article:

Mohammed, H. J. & Schrock, S. D. (2019) Modeling Spatial Relationships for Crashes Involving Teen Drivers. International Journal of Science and Engineering Investigations (IJSEI), 8(95), 1-8. http://www.ijsei.com/papers/ijsei-89519-01.pdf