Teen Drivers: Spatial Pattern of Crashes

2019 Mid-Continent Transportation Research Symposium
Ames, Iowa

Hemin J. Mohammed, Ph.D.

Road Safety Resource Coordinator

University of Kansas Transportation Center

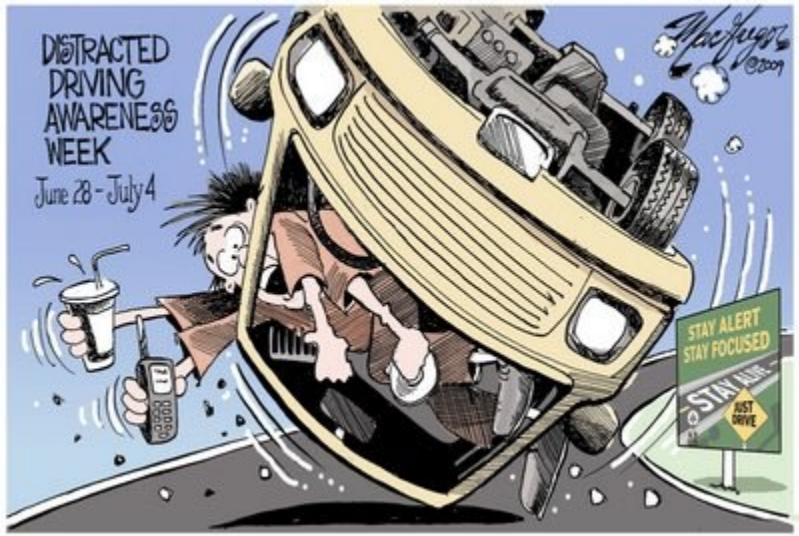
KUTC
UNIVERSITY OF KANSAS TRANSPORTATION CENTER

Thursday, Aug. 22, 2019

INTRODUCTION

Road traffic injuries are a major threat, especially for teens aged 15–19.

Traffic crashes involving teen drivers have been a worldwide concern for several decades.



Traffic crashes in the US are a leading cause of death for teens since 2001.

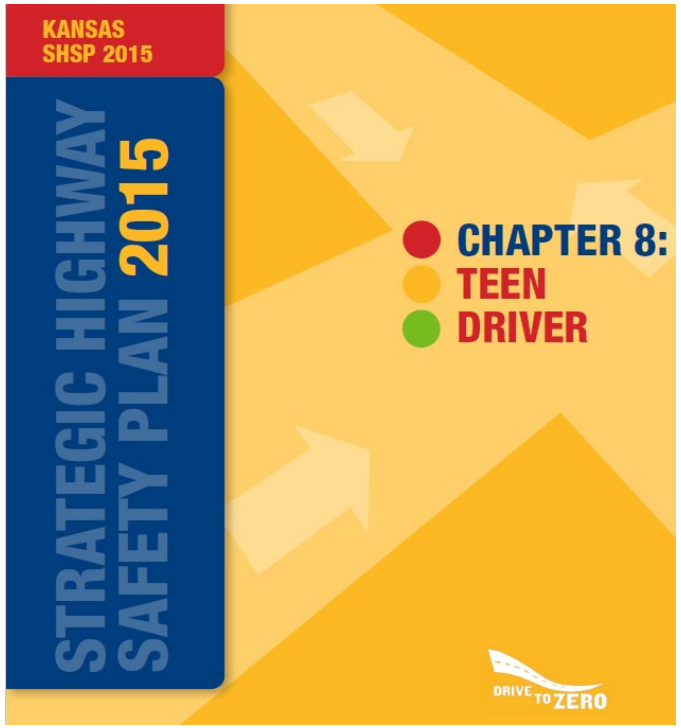
In 2015, 2,333 teens were killed (six teens every day).

In 2016 this increased by 3.6%.

Photo Source: Risk Management Group Newsletter

INTRODUCTION

In Kansas, KDOT's Strategic Highway Safety Plan (SHSP) 2015, identified teen drivers as one of its emphasis areas to reduce crashes.



The number of crashes involving teen drivers in Kansas increased by 644 (10,715 to 11,356) from 2013 to 2016.

The SHSP goals were not met.

LITERATURE REVIEW

Curry et al. (2011)

- 822 teen drivers were involved in 795 serious crashes
- Driver error was by far the most common reason for crashes (95.6%)
- Among crashes with a driver error, a teen made the error (79.3%) of the time

Singh (2018) categorized these errors into:

- Recognition errors
- Decision errors
- Performance errors
- Non-performance errors

LITERATURE REVIEW

Shinar, 2017; OECD, 2006; Carney et al., 2016; Huestegge et al., 2010; Crundall et al., 2012

The distinguishing characteristics of teen drivers associated with high-rate crashes are:

- Inexperience
- High risk-taking behaviors

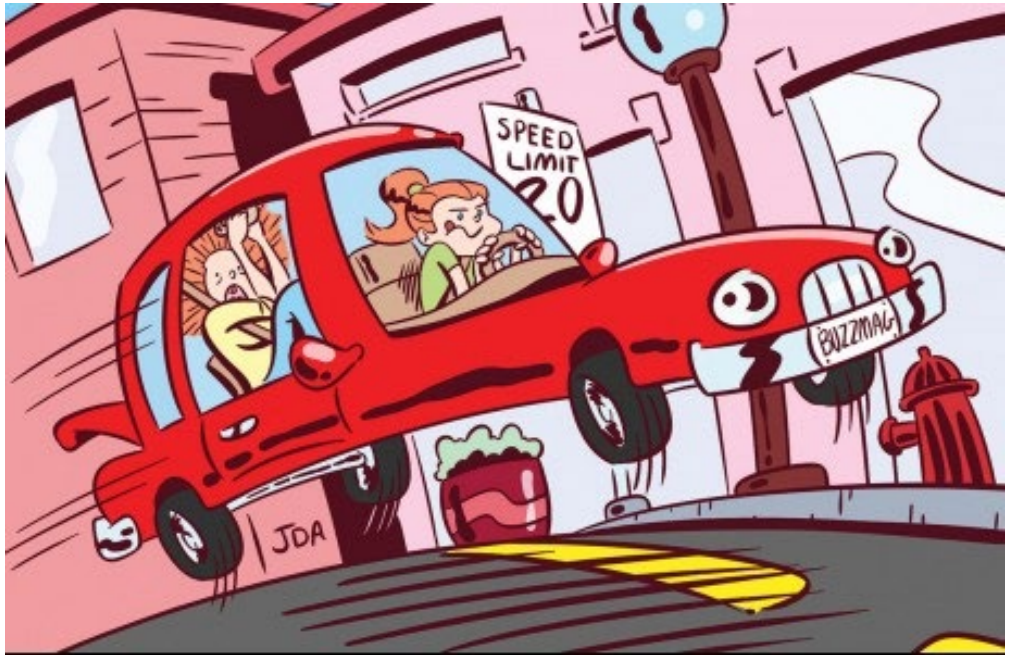
LITERATURE REVIEW

Mueller and Trick, 2012; Li et al., 2015; Scialfa et al., 2011; Borowsky et al., 2010; Borowsky et al., 2009

Teen drivers have:

- Higher hazard response times
- Greater speed and steering variability
- Higher crash rates in inclement weather,

at horizontal curves, and at intersections



RESEARCH OBJECTIVE

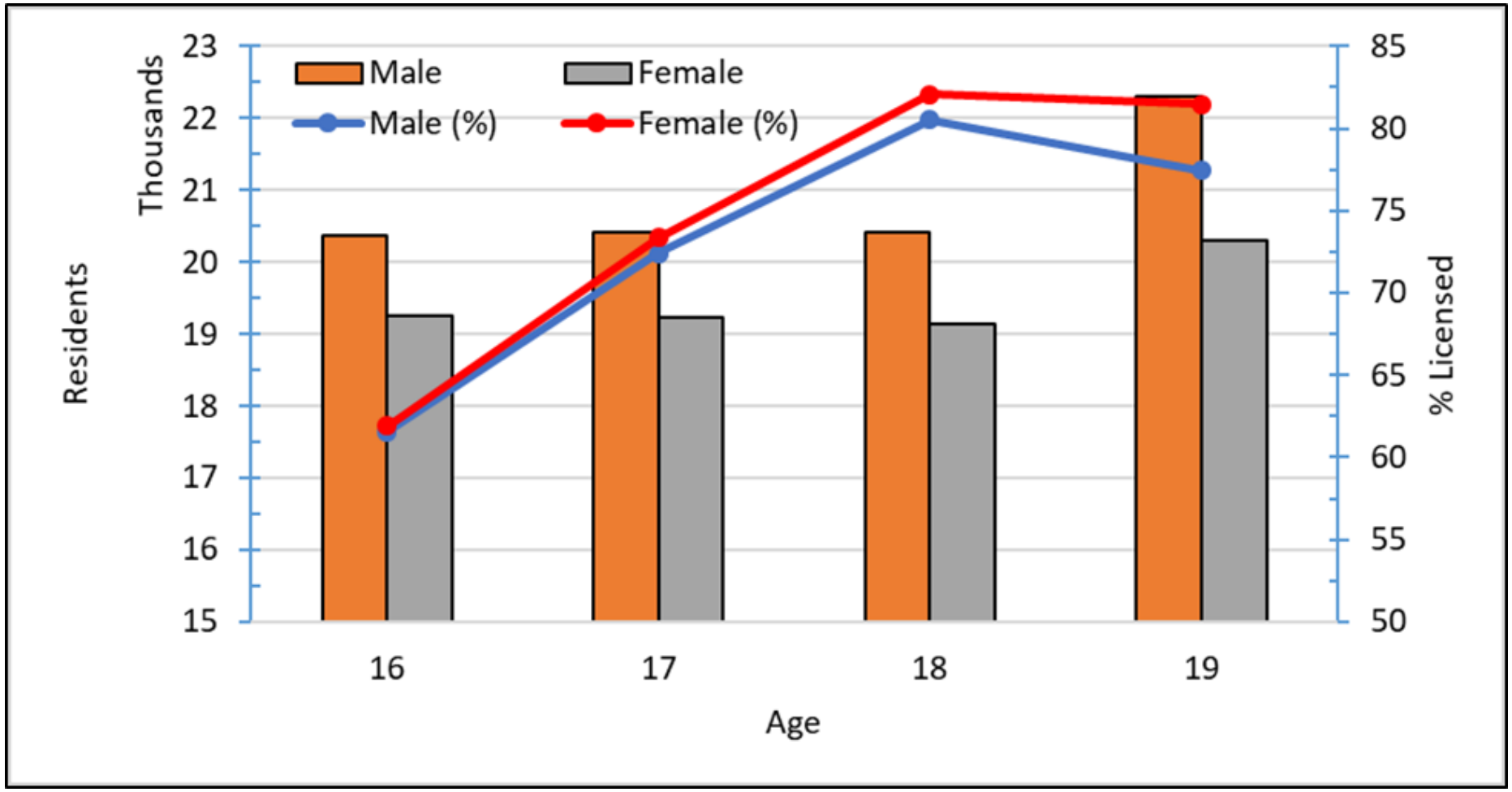
- Modeling spatial characteristics at:
  - County level
  - State level

DATA

- The Fatality Analysis Reporting System (FARS) database
- Kansas Department of Transportation (KDOT)
- U.S. Census Bureau Database
- Kansas State Department of Education

Photo Source: community.qlik.com

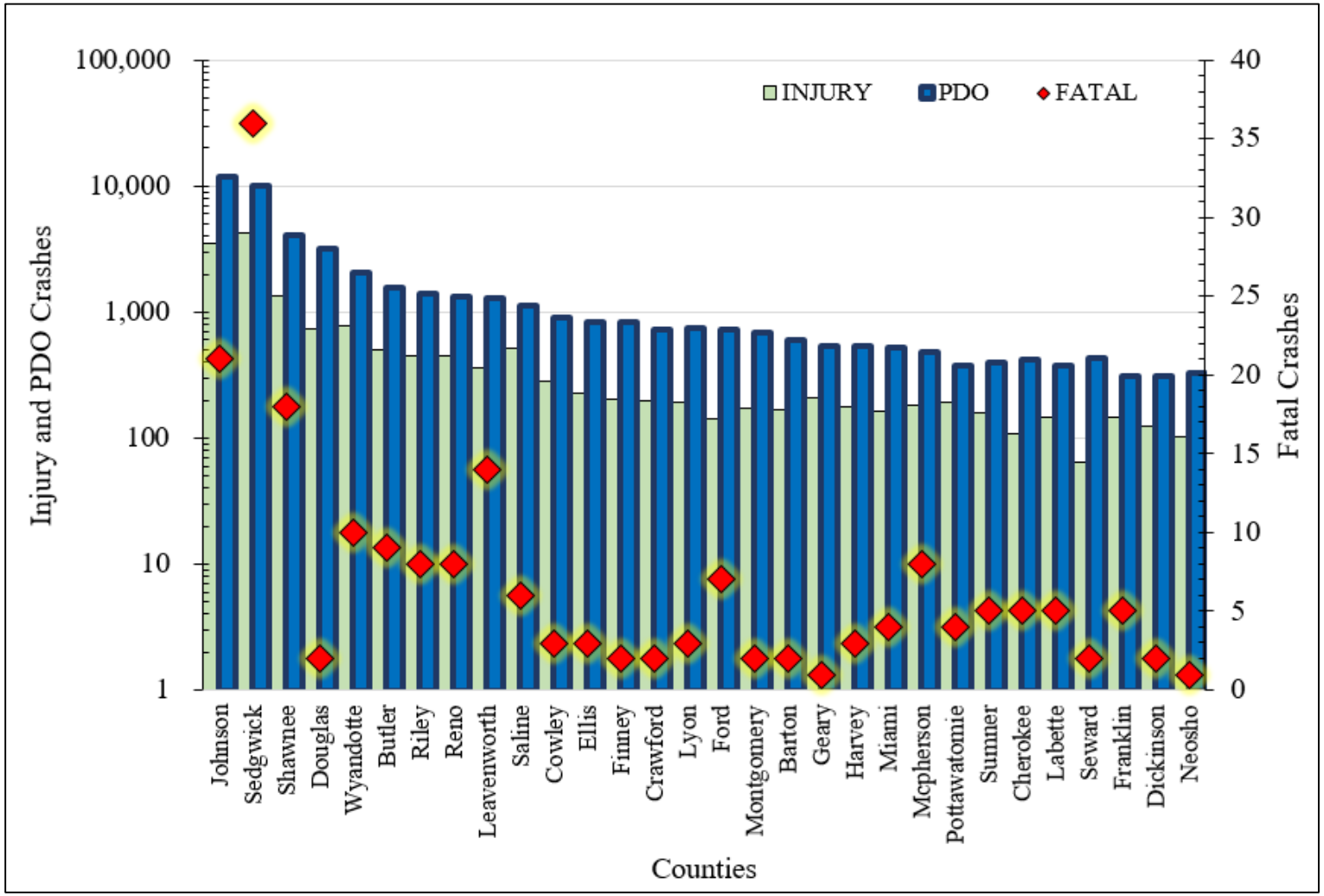
Population and Licensed Teenagers by Age and Gender



Crashes by Year

Year	Total Crashes	Crashes Involving Teen Drivers	Crashes Involving Teen Drivers (%)
2010	60,562	11,765	19.43
2011	60,270	10,978	18.21
2012	58,373	10,587	18.14
2013	59,265	10,472	17.67
2014	59,938	10,508	17.53
2015	61,440	10,709	17.43
2016	62,390	11,172	17.91
Total	422,238	76,191	**18.04**

Counties & Crashes



Modeling Spatial Relationships

- Ordinary Least Squares (OLS)
- Geographically Weighted Regression (GWR)

Picture Source: Solutions illustrate Inc.

POTENTIAL EXPLORATORY VARIABLES

Potential Exploratory Variables

- DVMT on all types of roads
- DVMT on rural non-state roads
- Household income
- Miles of all types of roads
- Miles of rural non-state roads
- Number of families (income is below poverty level)
- No. of High Schools
- No. of postsecondary schools
- No. of workers >15 years & commute to work

POTENTIAL EXPLORATORY VARIABLES

- No. of non-commercial trucks
- No. of passenger cars
- Population of >15 years in the labor force
- Population of 18-24 yrs old < high school degrees
- Population of counties
- Population of Females >15 years in the labor force
- Population of Males >15 years in the labor force
- Population of teens
- Precipitation

EXPLORATORY REGRESSION

Exploratory Variables

OLS

Test all variable combinations for:

- Redundancy
- Significance
- Bias
- Performance

Outputs

EXPLORATORY REGRESSION

Every variable should be statistically significant to be included in the model

Six variables were statistically significant

PC	The average number of passenger cars
DVMT_ALL	Average DVMT on all types of roads
RD	Miles of rural non-state roads in counties
HIGH.SCHOOL	Number of High Schools
POP.TEEN	The population of teens
POP	The population of counties

EXPLORATORY REGRESSION

- Variance Inflation Factor (VIF) Value
- Adjusted R Squared
- p-value
- Koenker (BP) Statistic p-value
- Joint F-Statistic p-value
- Joint Wald Statistic p-value
- Jarque-Bera p-value

MODELING SPATIAL RELATIONSHIPS

Ordinary Least Squares Regression Model

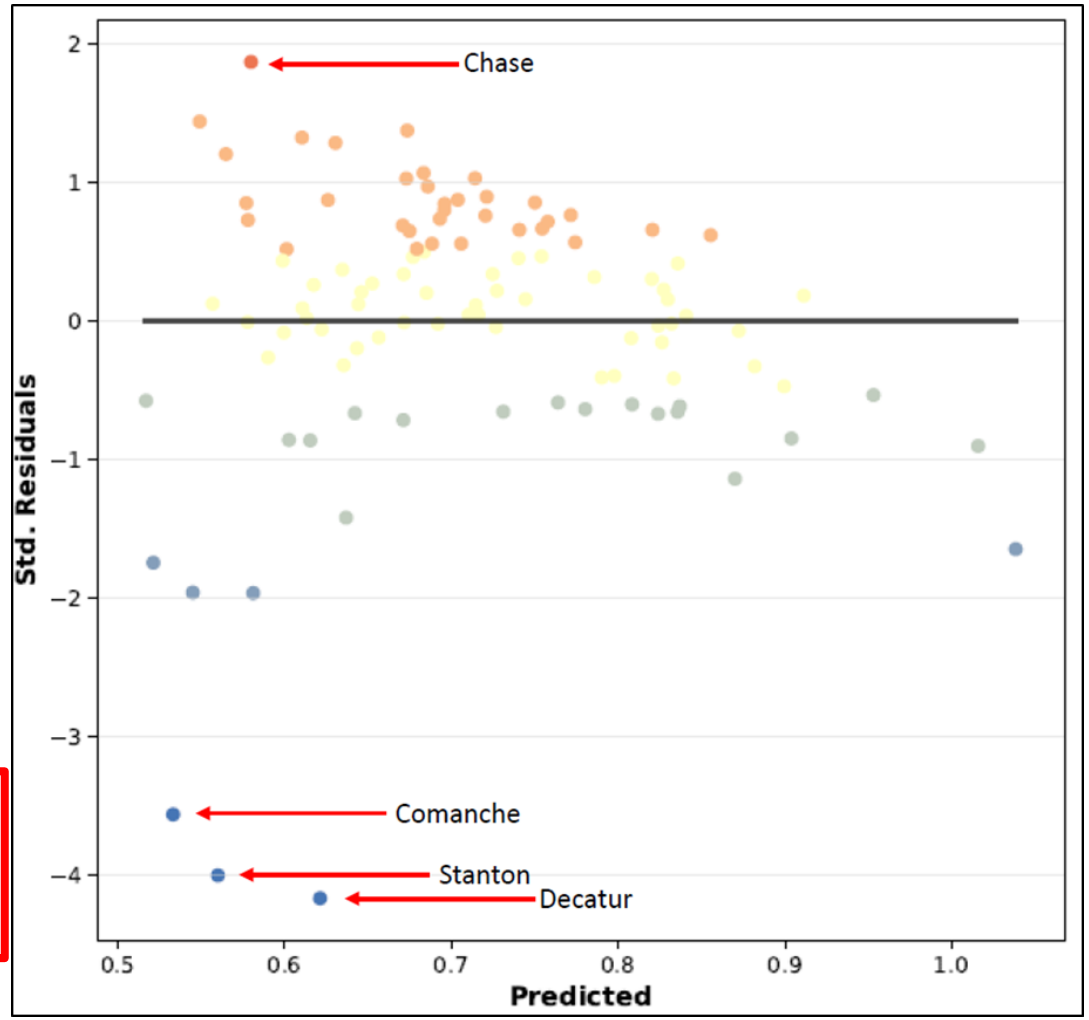
Multiple R-Squared	0.9101
Adjusted R-Squared	0.9084
Joint F-Statistic	<0.0001
Joint Wald Statistic	<0.0001
Koenker (BP) Statistic	0.0133
Jarque-Bera Statistic	<0.0001

Variable	Intercept	RD	PC
Coefficient	-1.0655	0.0192	1.8058
Std. Error	0.05693	0.00713	0.06028
Probability	<0.0001*	0.00826*	<0.0001*
Robust SE	0.07802	0.00717	0.06540
Robust Pr.	<0.0001*	0.00868*	<0.0001*
VIF	-------	1.08688	1.08688

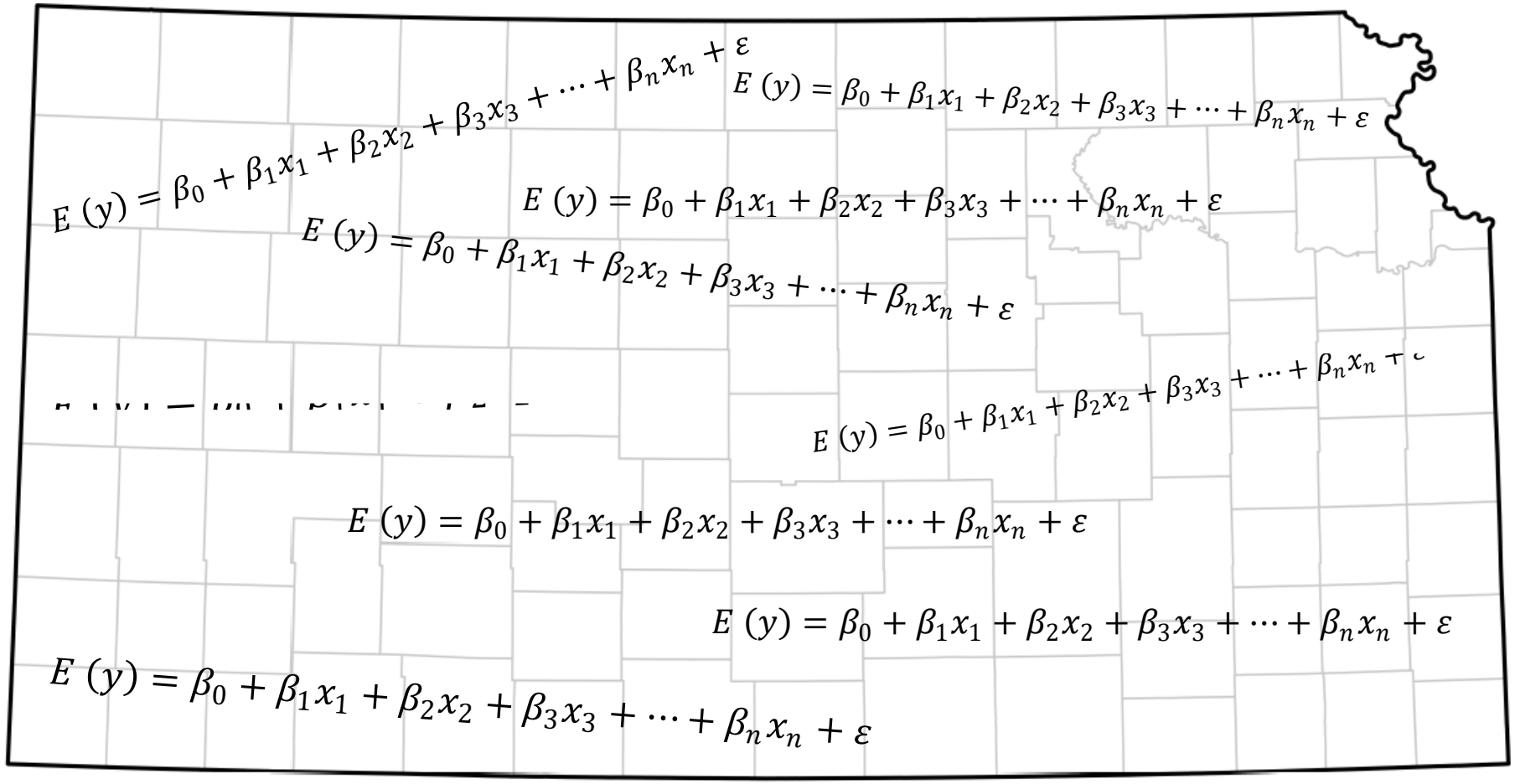
ORDINARY LEAST SQUARES REGRESSION (OLS) MODEL

- Chase County underpredicted
- Comanche, Stanton, Decatur overpredicted

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



GEOGRAPHICALLY WEIGHTED REGRESSION (GWR) MODEL



$E\ (y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_n x_n + \varepsilon$

$E\ (y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_n x_n + \varepsilon$

$E\ (y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_n x_n + \varepsilon$

$E\ (y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_n x_n + \varepsilon$

$E\ (y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_n x_n +$

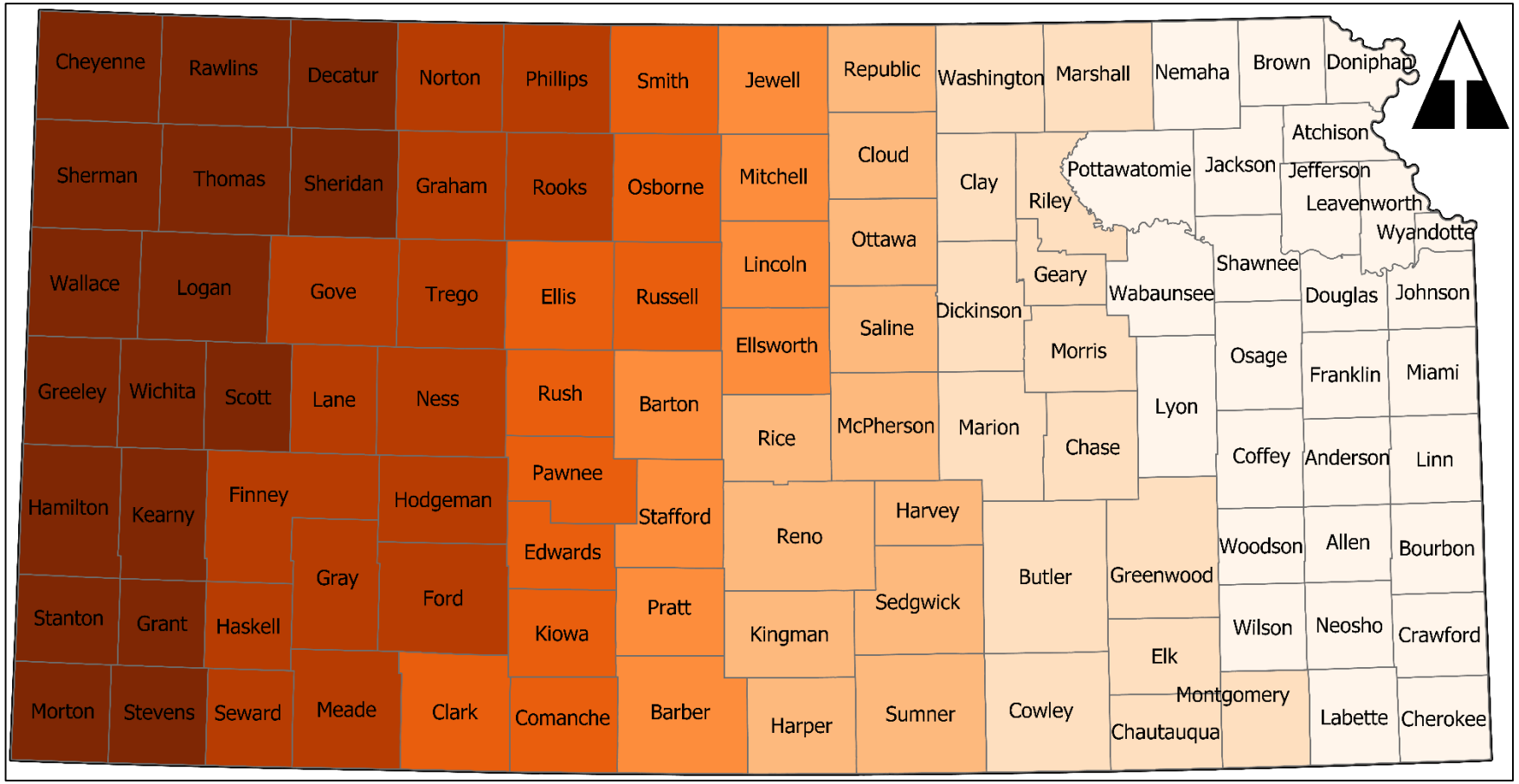
$E\ (y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_n x_n + \varepsilon$

$E\ (y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_n x_n + \varepsilon$

$E\ (y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_n x_n + \varepsilon$

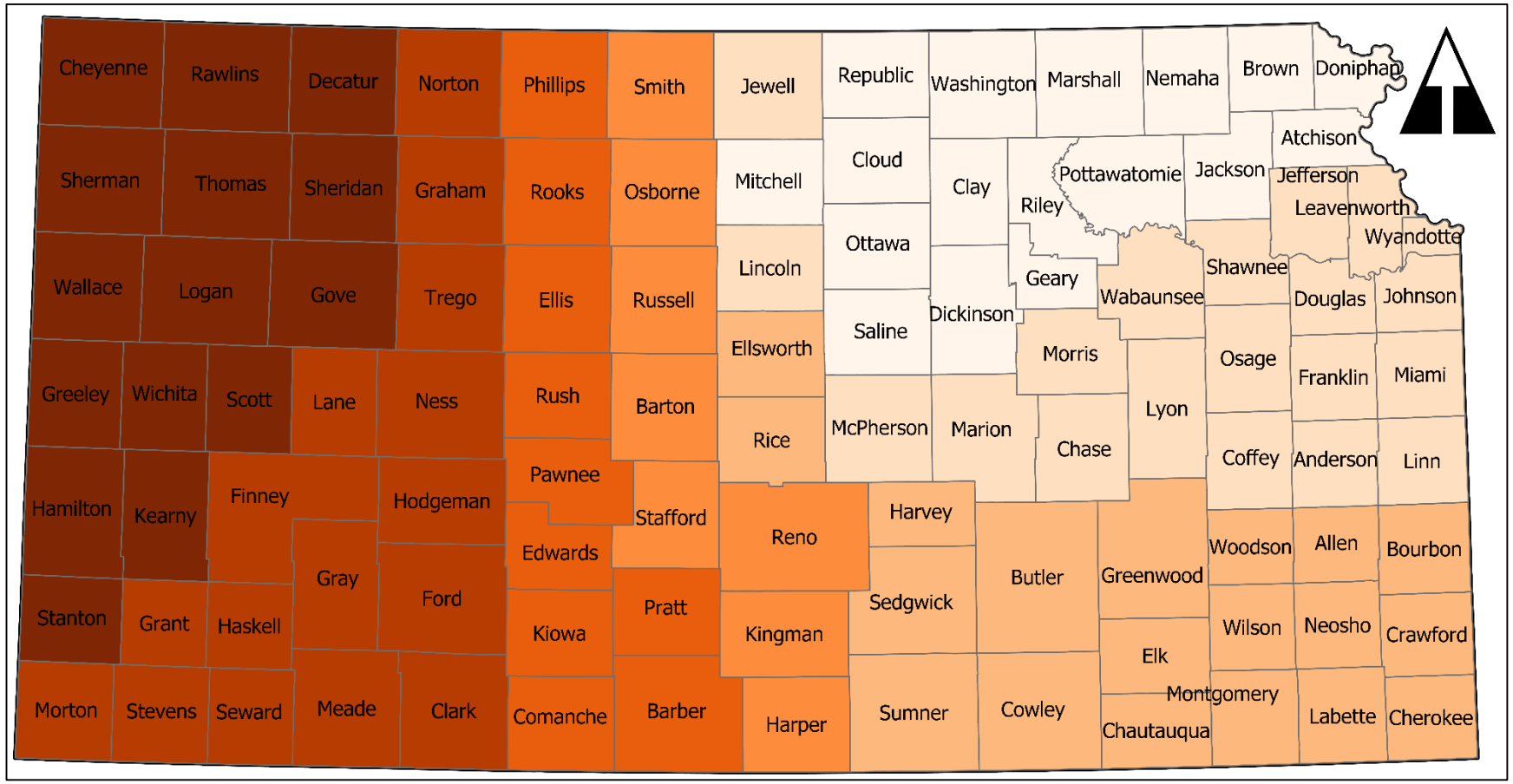
GEOGRAPHICALLY WEIGHTED REGRESSION (GWR) MODEL

Coefficient of average number of passenger cars



GEOGRAPHICALLY WEIGHTED REGRESSION (GWR) MODEL

Coefficient of miles of rural non-state roads



VALIDATION OF MODELS

- Validate by number of crashes in 2017
- Use number of registered passenger cars in each county in 2017
- Use miles of non-state roads in each county in 2017
- OLS underestimated overall number of crashes was by 3.66 %
- GWR underestimated overall number of crashes was by 2.94 %

VALIDATION OF MODELS

- Rawlins County (5 crashes in 2017)
- Predicted number (7.65 crashes)
- Overestimated by 53% (2.65 crash)

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

Among 30 counties that had the highest number of crashes, only Jefferson and Wyandotte had the predicted number of crashes off by >25%

PREDICTION BY MODELS

- Predict number of crashes in 2026
- Use growth factor
- Assume no other changes
- OLS predicts 3.11% reduction
- GWR predicts 3.37% reduction

Photo source: medium.com

FINDINGS AND RECOMMENDATIONS

- Developed statistically significant models (OLS and GWR) to predict future teen-related crashes at the county and state levels
- Among 18 related exploratory variables, the number of miles of non-state roads and the number of passenger cars were found to be statistically significant
- OLS and GWR 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

FUTURE STUDIES

- This research methodology can be used to analyze other subsets such as alcohol-related crashes, and older driver crashes
- Conducting temporal analysis using ArcGIS to compare crash rates between teen drivers and other age groups
- Comparing the findings of the current research with the network screening results

Thanks

Mid-America Transportation Center (MATC)

KDOT

Hemin J. Mohammed
University of Kansas Transportation Center (KUTC)
Hemin@ku.edu