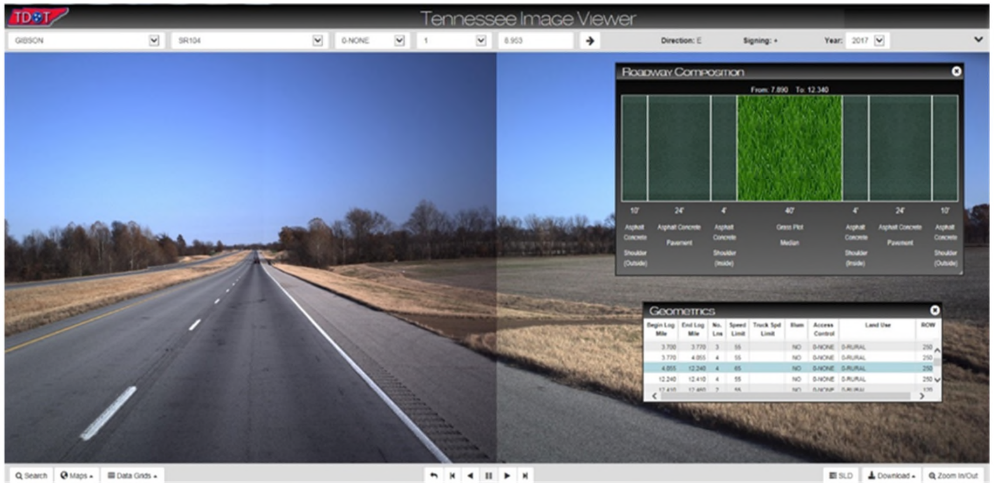
Application of Geographically and Temporally Weighted Regression Models for Estimating Safety Performance Functions of Multilane Rural Highways in Tennessee

Introduction

- Studies have used crash predictive models to estimate Safety Performance Functions (SPFs) defined in the Highway safety manual (HSM).
- These studies mostly use "one-size fits all" approach in estimating "global" models.
- However, the effect of a specific variable on crash frequency may be different in different locations and over different periods of time.
- This study explores the effects of key factors on crash frequency while simultaneously adresses spatial and temporal heterogeneity.
- This study used Geographically and Temporally Weighted Regression (GTWR) models to account for spatio-temporal heterogeneity in crash data.

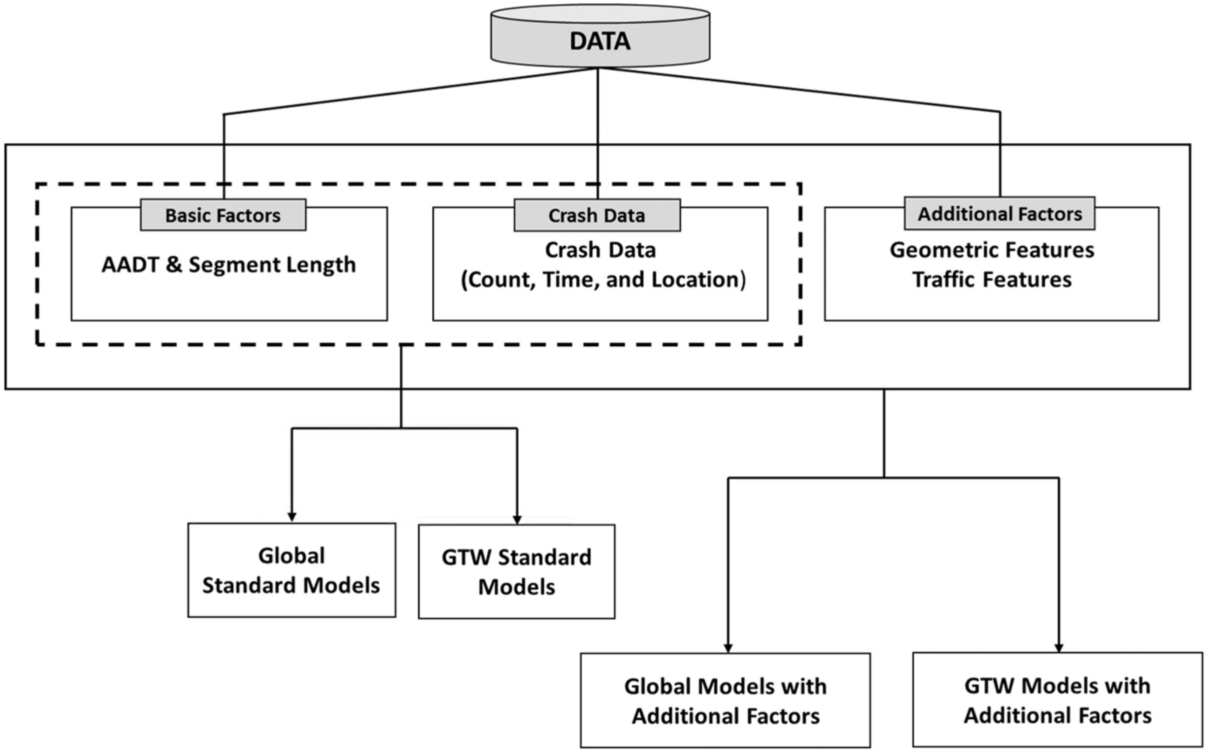
Data Collection

- Initial N = 558 segments were identified from the Enhanced Tennessee Roadway Information Management System (E-TRIMS).
- Then, crash data of the segments between 2013-2017 was extracted.
- Next, geometric data were manually collected from TDOT's image viewer.
- Finally, after data cleaning, 2444 crash observations belonging to 490 segments remained as the final dataset.



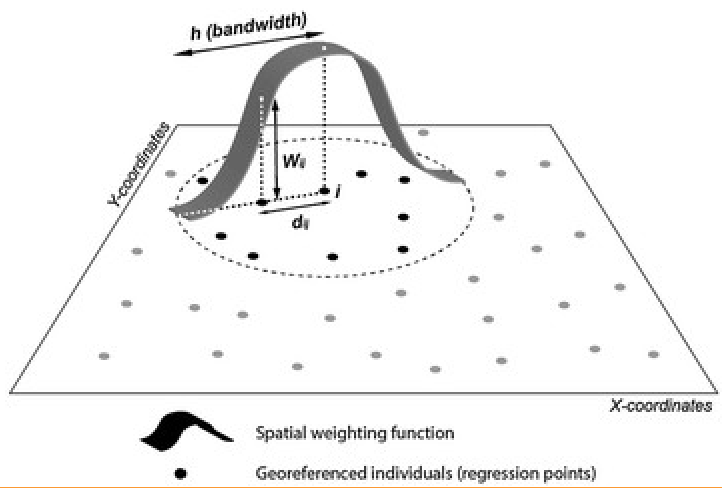
TDOT's Image Viewer

Methodology



GTWR models

- In GTWR models, every observation is considered as a centroid
- Then, a sample of observations that geographically and temporally have the most proximity to the target observation are used for model estimation.
- The neighboring observations of each sample are weighted in accordance with their spatiotemporal distance from the central observation in a way that the closer an observation is to the centroid, the more weight it receives.



Spaciotemporal SPFs

Step 1: Determining Weights Using Adaptive Bi-Square Kernel Function:

Geographical Weights → $$w_j^s(u_i, v_i, t_i) = \left[1 - \left(\frac{d_{ij}}{b}\right)^2\right]^2$$

Temporal Weights → $$w_j^t(t_i) = \left[1 - \left(\frac{t_{ij}}{b}\right)^2\right]^2$$

Spatio-temporal Weights → $$w_j^{st}(u_i, v_i, t_i) = w_j^s(u_i, v_i, t_i) \times w_j^t(t_i)$$

Step 2: Sample Selection for Neighboring Observations:

Optimum Sample Size → $$AIC = -2L + 2k + \frac{2k(k+1)}{n-k-1}$$

Step 3: Performing Weighted Regression Models:

$$N_{SPF} = \exp(a(u_i, v_i, t_i) + b(u_i, v_i, t_i) \times \ln(AADT) + c(u_i, v_i, t_i) \times \ln(L) + \sum_k \beta_k(u_i, v_i, t_i) \times X_{ik})$$

Results

Model estimation and Non-stationary test results of Geographically and Temporally Weighted Negative Binomial Regression (GTWNBR)

Variable Estimate		GTWNBR				Non-stationary Test			
		Min β	Max β	Mean β	std.dev	IQR	S.E.	Z value	Test
Intercept		-2.289	6.065	0.033	2.094	2.517	0.977	0.772	NO
Ln (AADT)		-0.172	0.700	0.498	0.228	0.229	0.052	107.920	YES
Ln (Segment Length)		0.222	0.705	0.520	0.158	0.317	0.022	163.322	YES
Speed limit		-0.017	0.017	-0.004	0.010	0.016	0.004	-22.799	YES
Terrain (Base: Flat)	Rolling	0.168	1.467	0.488	0.358	0.270	0.134	67.427	YES
	Mountainous	-0.149	1.988	0.628	0.544	0.451	0.238	57.048	NO
Land Use Base: Rural	Commercial	-0.036	0.980	0.611	0.282	0.497	0.151	107.193	YES
	Mix of Residential and Commercial)	0.085	0.905	0.380	0.182	0.265	0.110	103.474	YES
	Industrial	-0.753	1.010	0.003	0.459	0.799	0.371	0.295	NO
	Residential	0.160	0.907	0.454	0.130	0.176	0.202	172.802	NO
Access control (Base: No Access control)	No access control	-1.987	1.347	-0.444	1.036	1.359	0.337	-21.176	YES
	Partial	-1.073	-0.296	-0.533	0.165	0.148	0.110	-159.437	NO
Median Width		-0.038	0.003	-0.015	-0.038	0.019	0.003	-66.569	YES
Median type (Base: Grass)	Barrier	-2.194	-0.233	-1.016	-2.194	1.142	0.180	-82.537	YES
	Mountable Barrier	-0.596	1.182	0.393	0.559	1.030	0.210	34.737	YES
Lane width		-0.483	-0.160	-0.299	-0.483	0.141	0.063	-178.746	YES
Presence of acceleration or deceleration lane		-0.058	1.284	0.474	-0.058	0.862	0.144	53.393	YES
Number of observations		2444							
Dispersion parameter		0.872							
Log likelihood		-4183.528							
Pseudo – R^2		0.820							
AIC		8401.057							

Spatial Variation of Parameters

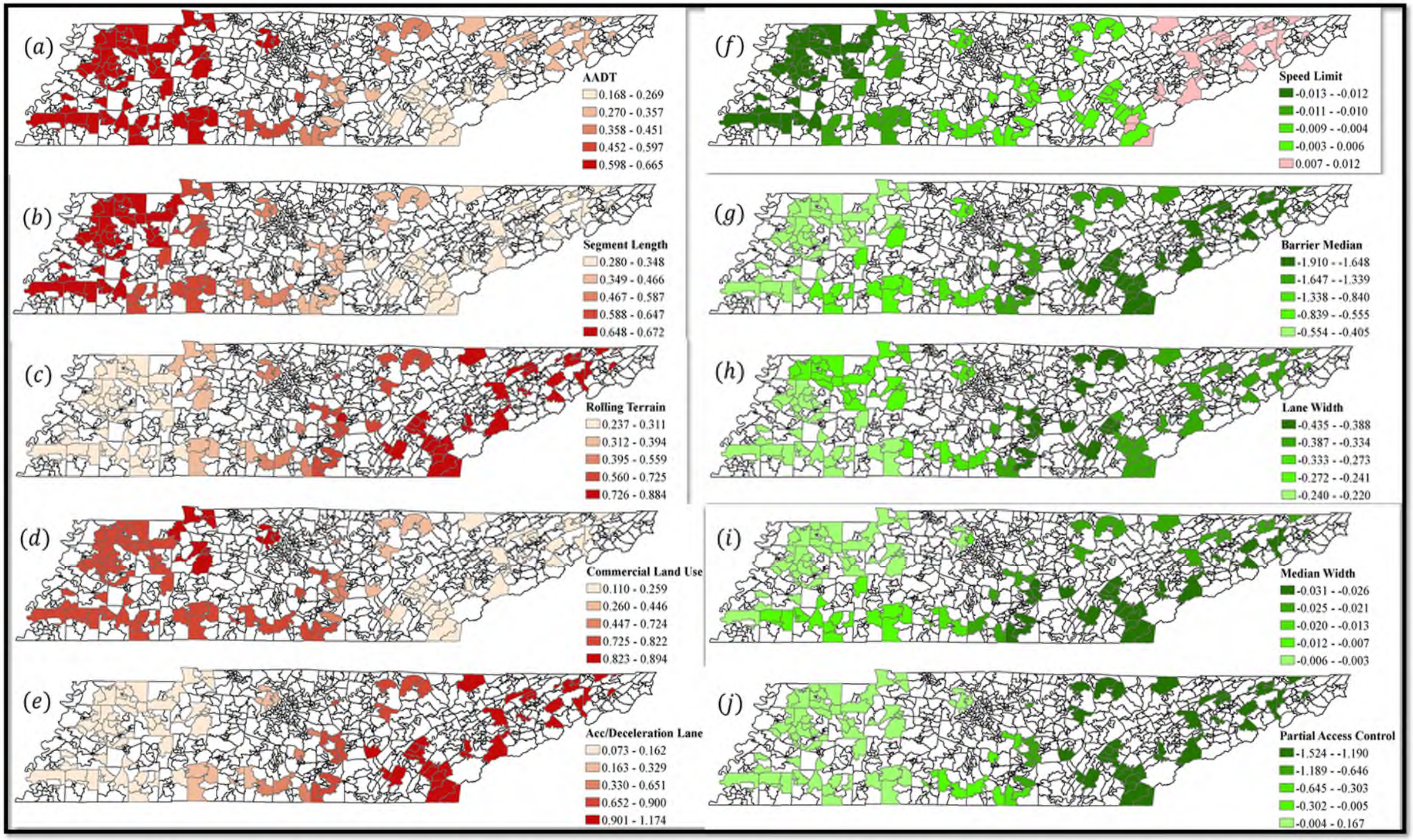
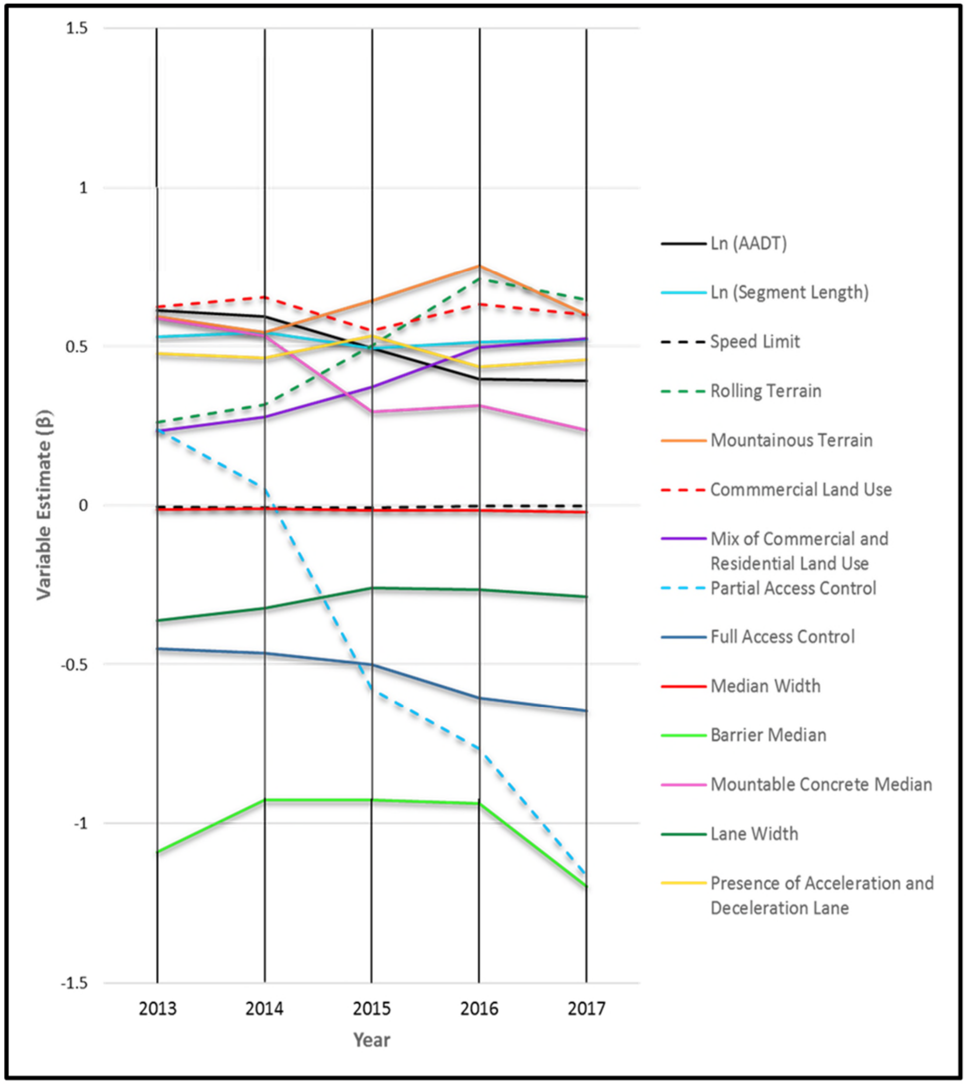


Illustration of the spatial variation of local parameter estimates in Tennessee in a county level

Temporal Variation of Parameters



Conclusion

- The majority of contributing factors varied significantly across space and over time.
- Variable estimates are stationary within specific regions, but vary from one side to the other side of the state of Tennessee.
- Also, some variables were stable over time (e.g. Median Width, and Speed Limit), but some variables vary considerably from one year to another (e.g. Access Control).
- Overall, the spatio-temporal variation of contributing factors in this study implies that considering variable estimates to be constant for an entire jurisdiction is a biased decision


Thank you
HASLAM
JOURNEY