Building Tomorrow's Society
Bâtir la Société de Demain

Fredericton, Canada
June 13 – June 16, 2018/ *Juin 13 – Juin 16, 2018*

THROUGHPUT EVALUATION OF WORKZONES ON HIGH VOLUME HIGHWAYS OF ONTARIO, CANADA

Nahidi, Seyedata[1,4], Rahman, Sonia[2] and Tighe, Susan[3]
[1] University of Waterloo, Canada
[2] Stantec, Canada
[3] University of Waterloo, Canada
[4] snahidi@uwaterloo.ca

Abstract: Highways as the primary type of transportation infrastructure, allow for haulage of huge amounts of goods and services in North America. An aging infrastructure combined with several other parameters such as high volume of heavy truck traffic and long harsh winters could lead to faster pavement deterioration. To improve safety, traffic planning, geometry, and structural performance of the highways, transportation agencies should have a detailed program to maintain, preserve, and reconstruct these infrastructures. Workzones interrupt the regular traffic flow on the highways which could potentially lead to an increase in the number of collisions. To prevent this, agencies are encouraged to present a comprehensive strategy to minimize the queuing in the workzones. The Ministry of Transportation Ontario (MTO) have identified that accurate prediction of workzone throughput could significantly reduce the user delay costs, and increase the safety of the high volume highways. This paper involves the evaluation of the workzone's throughput performance. To that end, various strategies such as Multiple Linear Regression, Negative Binomial Regression, and Truncated Regression models were developed to identify the parameters affecting workzone throughputs. An innovative concept of random parameter modeling was also used to account for unobserved heterogeneity issue. Results show that along with traditional factors, further parameters such as number of closed lanes, length of the workzone, presence of police, time of day, and percentage of heavy trucks have a significant effect on the throughput of workzones on high volume highways.

1 Introduction

Maintenance, rehabilitation, and preservation of the major high volume highways within the province of Ontario is one the most important responsibilities of the Ministry of Transportation of Ontario (MTO). Based on the definition on the King's Highways official website, 400-series highways could be defined as follows[1],

"At the top of the classification ladder are the 400-Series King's Highways. These roads are almost exclusively controlled access freeways, and are relatively few in number (15 in total, counting the Queen Elizabeth Way & 407 Express Toll Route). These important routes form the backbone of Ontario's road infrastructure, and are being expanded constantly to meet the needs of Ontario's economy."

[1] http://www.thekingshighway.ca/intro.html

The 400-series highways inside and around the Grand Toronto Area (GTA) are responsible for transporting vast numbers of people, but also goods and sources (Statistics Canada 2016). Thus, setting up a workzone on these highways should be utterly well organized to eliminate any possible traffic flow.

High volume highways are generally confronting issues such as rapid pavement deterioration, and need of more frequent workzones in order to maintain their serviceability. Setting up workzones in high volume highways would considerably increase the user delay cost and accident frequency. These factors would significantly affect MTO's strategies regarding the policies related to the workzones.

In order to solve the issues inside and around the workzones, several advanced and traditional methods were examined by several private and public agencies all around the world. Installing portable barriers, prohibition of heavy vehicles, reducing the speed limit, and strict traffic enforcement were some of the methods that were used to control the safety and traffic of the workzones. Presence of the Police at the beginning of the workzones also aimed to reduce the violations in these vulnerable areas. Moreover, in some workzones in order to capture the influence of the trucks and heavy vehicles on possible lane blockage and secondary crashes in the workzones, entrance of these vehicle types was prohibited (Chen and Tarko 2012).

Generally, government agencies prefer to schedule the workzone activities for high volume highways at night because of lower traffic volume and less disruptions for motorists in comparison to the day time peak hours. This arrangement also avoids excessive traffic and queuing on major highways, and inconvenience for motorists (Shepard and Cottrell 1986).

This study aimed to investigate whether the workzone strategies and guidelines are appropriate for current workzone throughputs, and if not, which additional parameters should be accounted for in order to increase the performance of the workzones. To that end, statistical models was used to predict the possible range of throughput in these workzones at queuing time. Then, by analyzing the results derived from these models the most important countermeasures were identified in order to improve the performance of the workzones.

2 Background

Workzone could be defined as a section of a roadway with construction, maintenance, or utility-work activities (Tuner 1999). Previously, the MTO conducted several studies from 2007 to 2011 to evaluate the workzone throughput at queuing times in high volume 400-series highways. Since 2007 several of these workzones were visited for evaluation and data collection in order to enrich the existing dataset. The dataset was then analyzed and several multiple linear regression models were developed to predict the possible future throughput. These generic models had the benefit of providing a broader insight about the factors which can affect workzone throughput in queuing condition. These studies did not focus on the safety of the workzones (Mushtaq 2011; Hicks 2006).

Although there are significant improvements in highway design, vehicle safety, and workzone guidelines in recent decades, the number of accidents which resulted in fatalities and injuries are still unacceptably high. Canada Transport (6) reported that around 150,000 of accidents occurred in 2014 led to serious injuries for drivers, passengers, and pedestrians. Based on this report around 2000 of these accidents resulted in fatalities. Report did not focused on the workzones separately, however workzones hypothetically could be considered as high-risk locations. Factors such as low visibility, more impaired drivers, and higher speed could increase the likelihood of experiencing severe injuries and fatalities in the workzones during the night.

Traffic Management in highway workzones could be considered as one of the most challenging tasks of transportation agencies. Reducing the motorists' delay and the risk of accidents are the major concerns that should be addressed under a proper traffic management system. This problem could be predominantly captured on the segments prior to the workzones in which one or more lanes are dropped. Gipps (1986) conducted a study to capture the possible influential parameters on drivers' behavior during lane changing. He claimed that parameters such as lane selection, urgency of lane change, entry and departure of non-transit vehicles into and from transit lanes, and effect of heavy vehicles can all significantly affect the drivers' lane-changing behavior. Moreover, these sections encounter higher traffic turbulence, resulting in higher likelihood of accidents and longer delays (Tarko et al. 1998).

Generally, the likelihood of experiencing conflicts, accidents, and delays is much higher in the entry section of the workzones. Lane closures are leading to capacity reduction and subsequent aggressive lane changes. These sudden maneuvers at the merging point creates extremely hazardous environment for both motorists and workers.

Recently, several studies were conducted to investigate the factors that can significantly affect the throughput of the workzones. Some of these studies used data from high volume 400-series highways located in Ontario, Canada (Mushtaq 2011; Hicks 2006). Findings of these studies captured that the presence of police, number of closed lanes, and types of barriers are the most significant factors for throughput prediction. These studies only focused on fixed parameter multiple linear regression modelling approach which doesn't account for unobserved heterogeneity issue.

Although, statistical models are highly data dependent and coefficient magnitudes could vary from one model to another, they are still expected to capture similar behavior. For example, in most of the studies queuing will increase by closing more than two lanes.

3 Data Collection

The main idea of estimating the throughput in workzones was initiated in 2006. After the first phase of data collection, more data was collected in 2011 due to the expansion of the comprehensive database (Mushtaq 2011; Hicks 2006). Finally, in 2016 data collection continued to provide an appropriate context to develop more reliable statistical models. The data collection methodology was consistent over the three data collection periods. In 2016 data collection, it was assumed that when the traffic speed in construction zone is 10 km/hr or less, queuing occurs.

A site characteristics form, which was designed and used in early phases, was used in 2016 data collection process as well. Factors that were documented on the form included:

- Date and time of site visit
- Location
- Weather
- Speed limit (posted and temporary in km/h)
- Length of work zone
- Duration of closure
- Etc.

During the data collection, vehicle counts were split into three categories: passenger vehicles, medium heavy vehicles, and heavy vehicles. On fifteen-minute intervals, the number of vehicles passing through the workzones was recorded. The hourly throughput frequency was calculated for the sites with queuing based on the collected data. In addition to the throughput records, videos and images were recorded to capture any anomalies. These videos allowed researchers to conduct further analysis by reviewing them if necessary.

The MTO Central Region Rehabilitation Program provided the schedule and characteristics of the potential workzone projects on 400-series highways. The provided program included various information such as the location, length, and type of project. This list also presented the possible operational duration of each workzone.

Although thirty-six construction sites were identified by the Ministry as the potential sites for data collection, only 9 of the them were visited. The rest of the sites could not be visited due to:

- No queuing reported by the Construction Administrators;
- No lane closures;
- Cancellation of construction work due to weather etc.

3.1 Descriptive Statistics

Based on the available dataset, it was captured that police officials were not present in 91% of the closures. 83% of the dataset was collected during the weekdays. Around 88% of the closures occurred at night time which was expected. In 94 % of the workzones barrel were used for closure, and in only 6% of the construction areas concrete barrier were installed. Right and left lanes were closed almost equally in the visited workzones. In addition, in 73% of the construction sites 2 or more lanes were closed to conduct the required maintenance, rehabilitation, and reconstruction activities. Among all visited construction sites 401 and Queen Elizabeth Way (QEW) were visited the most.

4 Methodology

4.1 Multiple Linear Regression

Multiple linear regression could be defined as a model which includes at least two independent variables. These models are common for continuous dependent variables (Washington et al. 2011). Most regression applications generally try to detect a list of explanatory variables that are significantly influence the dependent variable.

In general, explanatory models are based on data achieved from well-controlled experimentations in the laboratory, analytical models are based on data obtained from observational studies such as data collected from workzones. It has been investigated that any changes on dependent variable due to changes of explanatory variables is depend on numerous factors other than independent variables. To observe the effect of the independent variable, some assumptions should be considered (Washington et al. 2011). When any of the requirements is not satisfied, corrective actions should be considered, and in some cases, alternative modeling approaches should be tried.

The main goal of the regression is to state the relationship between dependent variable Y with at least one independent variable. Generally, these regression models aim to present some realistic information and properties of the population by only reviewing properties of the sample.

The matrix expression of the linear regression model can be presented as follows,

[1] $\mathrm{Y} = \mathrm{X}\beta + \varepsilon$

$$\varepsilon = \mathrm{Y} - \hat{\mathrm{Y}} = \mathrm{Y} - \mathrm{X}\hat{\beta}$$

where X is properties matrix, β is the factor's coefficient matrix, Y is dependent variable, $\hat{\mathrm{Y}}$ is the prediction matrix, and ε is the difference between actual and predicted values.

One of the most common ways to solve the above problem is Least Squares Method. This method is also known as "Ordinary least squares" or "OLS". In this method, the goal is to estimate the model parameters by using the sample data. To that end, the following formula could be used to estimate these parameters,

[2] $\beta = (\mathrm{X}^{\mathrm{T}}\mathrm{X})^{-1}\mathrm{X}^{\mathrm{T}}\mathrm{Y}$

Equation 2 presents the estimation of model parameters.

4.2 Truncated Regression

The truncated regression model is common for the no limit observations in the tobit formulation. This model is not a linear regression model and could be considered as non-linear regression (LIMDEP 2012). For general explanation of this model in which the range of the dependent variable is truncated in both tails, the conditional mean function is,

[3] $E[Y_i|x_i, L_i \leq Y_i \leq U_i] = \acute{\beta} x_i + \sigma \frac{\phi_L - \phi_U}{\Phi_U - \Phi_L}$

Where $x \sim N[\mu, \sigma^2]$ and σ is a constant. $\phi(.)$ and $\Phi(.)$ are standard normal pdf and cdf, respectively. L represents lower limit; however, U is upper limits representative.

The specification of the both lower and upper limits are same as the tobit model,

[4] $\alpha_{limit} = [limit - \acute{\beta} x]/\sigma\ ], limit = upper\ or\ lower$

$\Phi_{limit} = \Phi(\alpha_{limit})$

$\phi_{limit} = \phi(\alpha_{limit})$

The least squares analysis is inconsistent and inaccurate in solving this type of problems. To that end, maximum likelihood method was chosen to estimate the coefficients.

The log likelihood function of the truncated regression could be maximized by using Olsen's transformation parameters. After transformation, this function could be written as,

[5] $\sum_i logL_i = \sum_i log\theta - \frac{1}{2}\log 2\pi - \frac{1}{2}\varepsilon_i^2 - \log[\Phi(\theta U_i - \acute{\gamma} x_i) - \Phi(\theta L_i - \acute{\gamma} x_i)]$

where $\theta = 1/\sigma$ and $\gamma = \beta/\sigma$.

Since the truncated regression is a nonlinear regression, coefficients are not the marginal effects. Differentiation of conditional mean function with respect to the x_i produces the vector of slopes,

[6] $\frac{\partial E[Y_i|x_i, L_i \leq Y_i \leq U_i]}{\partial x_i} = \beta \left[1 - \frac{\alpha_L \phi_L - \alpha_U \phi_U}{\Phi_U - \Phi_L} - \left(\frac{\phi_L - \phi_U}{\Phi_U - \Phi_L}\right)^2\right]$

The term in the bracket is the scale factor of the marginal effect.

4.3 Poisson and Negative Binomial Regression Models

One of the goals of this study is to determine the number of vehicles passing the work zone during the queuing time. Equation 7 presents the Poisson Regression model, where the probability of work zone i having y_i throughput per hour.

[7] $P(y_i) = \frac{EXP(-\lambda_i)\lambda_i^{y_i}}{y_i!}$

Where $P(y_i)$ is the probability of workzone i having y_i throughput per hour and λ_i is the Poisson parameter of workzone i, which is equal to workzone i's expected number of throughput per hour, $E[y_i]$.

The relationship between independent (explanatory) variables with Poisson parameter can be presented as Equation 8.

[8] $\lambda_i = EXP(\beta X_i)$ or, equivalently $Ln\ N\ (\lambda_i) = \beta X_i$

X_is are the vector of the statistically significant explanatory variables directly derived from the collected dataset; βs are vector of the estimated parameters.

Sometimes the data do not meet the requirement of equality of mean and variance to permit the use of Poisson Regression. Negative Binomial Regression is one of the most popular models that can account for this inequality (Washington et al. 2011).

The Negative Binomial model has a similar formulation to Poisson model and it only counts for error term (ε) (Washington et al. 2011):

[9] $\lambda_i = \text{EXP}(\beta X_i + \varepsilon_i)$ or, equivalently $\text{Ln N}(\lambda_i) = \beta X_i + \varepsilon_i$

Where EXP(ε_i) is Gamma-distributed disturbance term with the mean 1 and variance α.

One of the methods to estimate the accuracy of the Poisson and negative binomial models is McFadden ρ^2. This value is equivalent to R^2 in linear regression and can be calculated as Equation 10.

[10] $\rho^2 = 1 - \frac{LL(\beta)}{LL(0)}$

where LL(β) is the log-likelihood at convergence with parameter vector β, and LL(0) is the log-likelihood at zero. A perfect model has ρ^2 equal to 1, however, the model with ρ^2 equal to zero is considered flawed (Washington et al. 2011).

4.4 Random parameters models

The modeling approaches that were discussed earlier consider the coefficients constant across the observations. However, this fixed-parameter assumption might not be corrected in all cases. This fluctuating behavior of some parameters is known as unobserved heterogeneity. Random parameter models account for this issue and led the coefficients change across the observations. This is a significant consideration, as if unobserved heterogeneity is not addressed, the parameter estimates can be incorrect, the inferences unreliable, and the predictions inaccurate (Washington et al. 2011; Anastasopoulos 2009; Nahidi et al. 2017).

In random parameters modeling, the coefficients become:

[11] $\beta_n = \beta + \omega_n$

where β_n is a vector of estimable parameters, and ω_n is a randomly distributed term (for example, a normally distributed term with mean zero and variance σ^2). Based on these assumptions, β_n is the vector of the coefficients for an individual parameter across the observations. The variation of this factor could be described by a density function - mixing distribution - q(β|ω), where ω presents the parameters of the density distribution (LIMDEP 2012). A simulation-based maximum likelihood estimation method was used to simplify the additional complexity due to presence of random parameters in the calculation of outcome probabilities. To achieve simulated probabilities numerous iterations of the same procedure should be conducted (Washington et al. 2011).

One of the major concerns in random parameter models is the sampling technique of β values from the selected distribution. In order to solve this issue, the technique of Halton draws (Halton 1960) is chosen. This method can provide an appropriate context to efficiently approximate probabilities by limiting the number of draws to the smallest possible extent. Moreover, the non-random selection of draws provides more robust estimates, as compared to the use of merely random draws (Bhat 2003; Train 2003) for further details on the simulation approach. Based on the literature, selecting 200 Halton draws can possibly yield to have stable and accurate estimation of parameters (Bhat 2003; Milton et al. 2008). For the functional form of the parameter density function several distributions, such as normal, Weibull, lognormal, uniform, and triangular distributions, will be considered and results will be compared in order to choose the best candidate in terms of statistical fit. A parameter will be considered as random when the mean and standard deviation of the parameter density function is statistically significant with 95% confidence level.

5 Discussion of results

Tables 1 and 3 show the statistically significant parameters of the various modeling strategies used for this study in order to estimate the best throughput estimation for the workzones in high volume highways in Ontario, Canada. A concise argument about each significant factor in the developed models will be presented in the following section.

5.1 Model results

Table 1 presents the estimated fixed-parameter models. The constant in these models is a representation of the average capacity of the highways from where the data was collected. Queuing usually occurs when departure rate at the termination area is lower than the arrival rate at the transition area. Developed models show that throughput decreases when two or more lanes with a length of more than 3km are closed. In addition, multiple linear regression and truncated regression models, captured that when only one lane is closed (in comparison with 2 or more closed lanes) more vehicles are passing through the workzones, which was expected. 401 highway is responsible for transportation of vast amount of people and trade goods, and is one of the busiest highways of the region. Accounting for this fact guarantees the justification of the significancy factor attributed to this highway in the models. Highway 401 was identified as a highway which has more queuing than the other highways due to the amount of traffic. On the other hand, 400, 409, and 410 highways behaved in an opposite way and models show that less queuing was experienced in the workzones located in these highways. The developed models also show that an increase in the percentage of trucks would lead to a significant increase in queuing, which is in-line with previous studies (Mushtaq 2011; Hicks 2006). In order to capture a more detailed effect of the heavy vehicles, this factor was divided in two major categories:1- more than 20% trucks, 2- less than 6% trucks. when the truck percentage is higher than 20%, throughput decreases, and queuing increases which was expected. Moreover, when this factor is less than 6%, throughput increases. This factor can emphasize that low percentage of trucks balances the speed of the vehicles in highways.

Surprisingly, it was captured that the presence of Police in workzone areas could significantly reduce the number of vehicles passing through the workzones. It highlights the fact that drivers act more cautiously by observing police cars prior to the workzones and aim to reduce their speed more rapidly. In addition, if two or more lanes are closed throughput frequency decreases. The RMSE for each of these models were calculated and this measure is presented for each model in Table 2. Lower RMSE shows better fit. Among all these models, truncated regression model performs better than the other two. Also, this model predicts more correct data points within 200 vcl/h/l, which is around 87%.

Table 3 shows the model results from random parameter modeling process. One the major differences of the random parameter models is that the constant values in these models are much closer to the design capacity, and are therefore expected to be more accurate. Similar to the fixed parameter multiple linear regression model, presence of police shows a significant reduction on throughput in workzones at the queuing time. Most of the factors in random parameter models are similar to the fixed parameter models. One of the parameters which was significant in the truncated random parameter model is left lane closure. This model captured that when the left lane is closed more vehicles are passing the workzones. This could be because of higher speed in left lanes and hence a quicker merging process. Based on the standard deviation and mean value of the night factor in the negative binomial model, when lane closure occurs during the night time, throughput increases which means less queuing. This was expected but note that the magnitude of this factor changes across the observations.

Another interesting finding of this study was that installation of the barrels instead of concrete barriers could reduce the throughput. Type of barriers could possibly affect the behavior of drivers and motorists while they are passing the workzones. Because of this fact, this factor was significant as random parameter, which means the effect of this factor changes across the observations. Finally, based on the Truncated regression model, having barrels as barriers in the workzones increase the queuing in 97% of the occasions. However, in 3% of the workzones concrete barriers increased the queuing. This fact slightly changes under the Negative Binomial model. In this model, in 93% of the closures with barrels queuing increases.

Table 4 shows the percentage of the correct predictions of random parameter models within 25, 50, 75, 100, 200, and 300 vcl/h/l. This table also presents the RMSE value for three different random parameter models for sake of comparison. Based on prediction performance and RMSE values Truncated Regression model performs better than the other two models.

6 Summary and Conclusion

This article aimed to investigate a practical analysis of workzone characteristics, configurations, and conditions on the throughput of workzones on high volume highways in Ontario, Canada. The proposed statistical modeling approach provided an appropriate context in order to identify the parameters that affect the throughput of the workzones. Also, this study accounted for unobserved heterogeneity by benefiting from random parameter concept in the modeling procedure. Modelling strategies such as Multiple Linear regression, Negative Binomial regression and Truncated regression were used in this paper. The results from statistical analysis presented a broad perspective about factors which can play a significant role in predicting the throughput of the workzones. Lengthier workzones, and high percentage of trucks were the two major factors that increased queuing in these regions. Therefore, by setting up shorter workzones and also limiting trucks' access to the highways within active construction zones could help reduce queuing. In addition, based on the estimated models, it is recommended to have only one lane closure in high volume highways instead of two or more lanes closures.

Table 1: Fixed Parameter Models' Summary

Fixed Parameter Models	MLR*		TR**		NBNR***	
Factors	Coef.	t-value	Coef.	t-value	Coef.	t-value
Constant	1397.450	28.170	1049.090	13.610	7.138	193.160
Presence of Police (1 if present, 0 otherwise)	-229.571	-3.940				
WZ**** Configuration (1 if only one lane closed and it is day time, 0 otherwise)	499.296	7.330	765.145	6.190		
WZ Configuration (1 if two or more lanes closed and length of the workzone is greater than 3 km)	-158.116	-3.780	-212.311	-5.100	-0.144	-3.600
WZ Configuration (1 if it's weekend night and only one lane closed, 0 otherwise)					0.284	3.590
WZ Condition (1 if 2 or more lanes closed, 0 otherwise)					-0.323	-4.210
WZ Location (1 if 401 Highway, 0 otherwise)	-76.099	-2.100			-0.053	-1.670
WZ Location (1 if 400, 409, and 410 Highways, 0 otherwise)			87.965	2.000		
WZ Condition (1 if speed limit is 80 km/h and it' is night time, 0 otherwise)			-273.785	-4.600	-0.226	-3.780
WZ Condition (1 if it's night, 0 otherwise)			164.958	2.150		
Natural Logarithm of Truck Percentage	-70.050	-3.400				

Truck Percentage (1 if it's higher than 20%, 0 otherwise)			-112.874	-2.440	-0.077	-1.870
Truck Percentage (1 if it's less than 6%, 0 otherwise)			79.702	1.990	0.088	2.070
Sigma			143.440	11.530		
Dispersion Parameter					0.014	4.890
R-Squared or Rho-Square	0.61		0.69		0.68	
Adjusted R-Squared or Rho-Square	0.58		0.66		0.66	

* Multiple Linear Regression ** Truncated Regression *** Negative Binomial Regression **** Workzone

Table 2: Fixed Parameter Models' RMSE and Prediction Performance

Models	Within 100 vcl/h/l	Within 200 vcl/h/l	Within 300 vcl/h/l	RMSE
TR	57%	87%	98%	134.369
NBNR	55%	82%	98%	141.795
MLR	59%	79%	98%	146.946

Table 3: Random Parameter Models' Summary

Random Parameter Model	MLR		TR		NBNR	
Factors	Coef.	t-value	Coef.	t-value	Coef.	t-value
Constant	1853.730	49.580	1670.230	59.320	7.012	271.98
Presence of Police (1 if present, 0 otherwise)	-195.473	-7.130				
WZ Configuration (1 if two or more lanes closed and length of the workzone is greater than 3 km)	-114.172	-5.680			-0.152	-9.410
WZ Configuration (1 if two or more lanes closed, 0 otherwise)	-118.377	-6.220	-102.688	-6.720	-0.461	-14.190
WZ Configuration (1 if left lane closure, 0 otherwise)			33.516	2.480		
WZ Configuration (1 if it's weekend night and only one lane closed, 0 otherwise)					0.424	12.410
WZ Condition (1 if night, 0 otherwise)	-120.459	-4.230	-111.753	-4.780	0.104	4.820
Standard deviation of parameter distribution					*0.131*	*20.080*
WZ Condition (1 if speed limit is 80 km/h and it's night, 0 otherwise)			-242.796	-10.900	-0.209	-9.650
WZ Location (1 if 401 Highway, 0 otherwise)	-64.968	-3.680				
WZ Location (1 if 401 or QEW Highways, 0 otherwise)			-67.483	-3.950		

Natural Logarithm of Truck Percentage	-66.029	-7.000				
Truck Percentage (1 if it's higher than 20%, 0 otherwise)	-0.088				-0.088	-5.120
Truck Percentage (1 if it's higher than 18%, 0 otherwise)			-103.014	-6.200		
Truck Percentage (1 if it's less than 6%, 0 otherwise)					0.060	4.130
Type of Barrier (1: Barrel, 0: concrete)	-279.830	-7.890	-215.003	-6.460		
Standard deviation of parameter distribution	*152.883*	*18.570*	*150.927*	*23.010*		
Sigma			56.428	12.330		
Dispersion Parameter					524.384	4.320

Table 4: Random Parameter Models' RMSE and Prediction Performance

Models	Within 25 vcl/h/l	Within 50 vcl/h/l	Within 75 vcl/h/l	Within 100 vcl/h/l	Within 200 vcl/h/l	Within 300 vcl/h/l	RMSE
TR	74%	94%	98%	100%	100%	100%	24.668
NGBN	67%	91%	99%	99%	100%	100%	30.886
MLR	59%	88%	98%	100%	100%	100%	33.271

References

Anastasopoulos, P., Mannering, F. 2009. A Note on Modeling Vehicle Accident Frequencies with Random-parameter Count Models. *Journal of Accident Analysis and Prevention*, 41**(1):** 153-159.

Bhat, C. 2003. Simulation Estimation of Mixed Discrete Choice Models Using Randomized and Scrambled Halton Sequences. *Journal of Transportation Research Part B: Methodological*, 2003**(9)**: 837-855. DOI: 10.1016/S0191-2615(02)00090-5

Chen, E., and Tarko, A. P. 2012. Best Practices for INDOT-Funded Work Zone Police Patrols. Publication FHWA/IN/JTRP-2012/36. *Joint Transportation Research Program*, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana.

Gipps, P. A. 1986. Model for the Structure of Lane-changing Decisions. *Transportation Research Part B: Methodological*, 20: 403-414.

Halton, J. H. 1960. On the Efficiency of Certain Quasi-Random Sequences of Points in Evaluating Multi-Dimensional Integrals. *Journal of Numerische Mathematik*, 2**(1)**: 84-90. DOI: 10.1007/BF01386213

Hicks, C. 2006. Measuring Highway Work Zone Throughput." MA thesis. U of Toronto.

LIMDEP Econometric Modeling Guide. 2012. LIMDEP, Plainview, NY. 1162–1167.

Milton, J., Shankar, V. and Mannering, F. 2008. Highway Accident Severities and the Mixed Logit Model: An Exploratory Empirical Analysis. *Journal of Accident Analysis and Prevention*, 40: 260-266.

Ministry of Transportation of Ontario, Book 7: Ontario Traffic Manual for Temporary Conditions. Queen's Printer, Ontario, 2014. ISBN: 978-1-4606-3505-6

Mushtaq, A. 2011. Measuring Work Zone Throughput and User Delays. MA thesis. U of Waterloo.

Nahidi, S., Fountas, G., Sarvani SP, Sarwar, M. T., Anastasopoulos, P.C. 2017. Project Discrepancies in Roadway Construction and Preservation: A Statistical Analysis of Public–Private Partnership Contract Types in the US. *Journal of Frontier Built Environment*, 3. DOI=10.3389/fbuil.2017.00015

Shepard, F. D., Cottrell, B.H. Jr., 1986. Benefits and Safety Impacts of Night Work-Zone Activities. Transportation Research Record: *Journal of the Transportation Research Board*, TRR, 1083: 31-36.

Statistics Canada. Population by Year, by Province and Territory. 2016. Government of Canada, CANSIM (database). http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/demo02a-eng.htm.
Tarko, A., Kanipakapatnam, S., Wasson, J. 1998. Modeling and Optimization of the Indiana Lane Merge Control System on Approaches to Freeway Workzones, Part I. Publication FHWA/IN/JTRP-97/12. *Joint Transportation Research Program*, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana.
Train, K. 2003. Discrete Choice Methods with Simulation, Cambridge University Press, Cambridge, UK .
Transports Canada. Transportation in Canada. Government of Canada. https://www.tc.gc.ca/media/documents/policy/2014_TC_Annual_Report_Overview-EN.pdf. ISSN: 1920-0846. Accessed 2015.
Turner, J.N. 1999. What's a Work Zone? Journal of Public Road, 62**(6)**.
Washington, S., Karlaftis, M., Mannering, F. 2011. Statistical and Econometric Methods for Transportation Data Analysis (2nd ed.), Chapman and Hall/CRC, Boca Raton, FL.