Thai Journal of **Math**ematics
Special Issue: The 17^{th} IMT-GT ICMSA 2021
Pages 240–249 (2022)

http://thaijmath.in.cmu.ac.th

ISSN 1686-0209

Road Traffic Injuries in Thailand and their Associated Factors using Conway-Maxwell-Poisson Regression Model

Rattana Lerdsuwansri[1], Chonthicha Phonsrirat[1], Panicha Prawalwanna[1], Noppachai Wongsai[1,2], Sangdao Wongsai[1,2] and Teerawat Simmachan[1,2]*

[1] *Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Pathum Thani, 12120, Thailand*
e-mail : rattana@mathstat.sci.tu.ac.th; chonthicha.psr@gmail.com; panicha.prawalwanna@gmail.com

[2] *Thammasat University Research Unit in Data Learning, Faculty of Science and Technology, Thammasat University, Pathum Thani, 12120, Thailand*
e-mail : nong12@tu.ac.th; sangdao@mathstat.sci.tu.ac.th; teerawat@mathstat.sci.tu.ac.th

Abstract In road traffic accidents, the number of human injuries does not follow the traditional Poisson distribution, and often exhibits overdispersion. Based on a survey of 180 countries by the World Health Organization in 2015, Thailand is the second unsafety country for driving on roads, with 36.2 deaths per 100,000 population. In this study, a total of 20,226 road traffic accidents in Thailand in 2015 were investigated. Factors associated with the injury counts per accident were disclosed using the Conway-Maxwell-Poisson regression model. Four covariates considered were road type, road surface, road section, and festive months. The results showed that the distribution of injury count data was under-dispersed, which is rare. The number of injuries was significantly determined by road type, road section, and festive months. Driving on the Local Administration Roads was risker than on the National Highways or Rural Roads. Straight road sections were found to have a lower involvement in road accident injuries than at intersections. Our findings show that the month of the festival plays a crucial role in increasing unusual human injuries on the road. Road safety managers should pay more attention on these months, not only in Thailand but also other countries where special events occur.

MSC: 62P05; 90-10; 62J05

Keywords: road traffic accident; risk of injury; count model; underdispersion; festival effect

Submission date: 15.03.2022 / Acceptance date: 31.03.2022

1. Introduction

Based on the Global status report on road safety 2015, Thailand was the second unsafety country for driving on roads, with an estimated 36.2 deaths in road traffic accidents (RTAs) per 100,000 population [1]. RTAs have emerged as a major public health problem

*Corresponding author.

Published by The Mathematical Association of Thailand.

in the country. The total economic burden of RTAs was estimated to be over US$12,000 billion in 2015 [2]. These losses are trending to rise.

During festive and holiday seasons, RTAs in Thailand occur more frequently [3]. Two popular festivals are Thai New Year holidays, also known as Songkran festival or Water festival (the period of 11-17 April every year), and the New Year celebrations (the period of 27 December to 2 January). Families return home and friends reunite nationwide to celebrate. Such movement increases traffic volume, reduces traffic flow speed, and causes adverse consequences, such as road accident injuries (RAIs) and road accident fatalities (RAFs). Unfortunately, these festival periods are called the seven deadly days. In 2015, Thailand Road Safety Observatory [4] reported a total of 3,373 accidents with 364 fatalities and 3,559 injuries during the seven-day Songkran festival. A total of 2,997 accidents, 341 fatalities together with another 3,117 injuries were recorded during the New Year holidays. The number of RTAs has increased every year and evidently twice more during the festive periods and has continued in this direction [5].

Modelling the number of RAIs in relation to the festival periods still receives little attention in literature. Four common factors affecting injury frequencies are road, environment, human behavior, and demographics and have been reported in road safety studies elsewhere [6–9]. Poisson regression is a popular model for count data in crash analysis. It assumes an equidispersion that is the mean is equal to the variance. However, real crash data often has overdispersion and seldom underdispersion, which makes Poisson regression inappropriate. For overdispersion data, the Conway-Maxwell-Poisson (CMP) and Negative Binomial (NB) models were used by Lord, Guikema, & Geedipally [10] to analyze crash data from signalized four-legged intersections in Toronto, Ontario. Based on a study by Elvik, Sagberg & Langeland [11], the NB model was applied to analysis the number of RTAs on road bridges in Norway during 20102016. According to Lord, Geedipally & Guikema [12], the CMP, Gamma and Poisson distribution have been used for underdispersion data comprising 162 railway-highway crossings in South Korea between 1998 and 2002.

This study aimed to identify the relationship between the number of RAIs and the festival periods as well as road conditions across the country of Thailand for year 2015 using the CMP regression model.

2. MATERIALS AND METHODS

2.1. STUDY AREA

Thailand has been one of the upper-middle income countries with a GDP of US$401 billion in 2015 and the population of 68.7 million [13]. It is divided into 77 areas including the capital Bangkok and 76 provinces. Thailand is rich in natural resources that attract international tourists from around the world. According to the 2015 Travel and Tourism Competitiveness Index by World Economic Forum report of 140 countries, Thailand ranked 35th for its destination competitiveness [14]. However, the country ranked 132nd for its safety and security. In the same year, the World Health Organization (WHO) showed that the rate of RTAs, RAFs (36.2 per 100,000 persons) [1] and RAIs (239.29 per 100,000 persons) [4] in the country was significantly high. Currently, the RTAs have a high tendency to increase. Moreover, RTAs lead to global economic losses, estimated at US$518 billion per year in RAIs costs. For developing countries, huge economic losses are a financial burden. In 2015, Thailand's rate of RAIs was 1.64 per 100 million GDP [4].

2.2. DATA

RTAs data in 2015 was collected from the Department of Disaster Prevention and Mitigation of Thailand. The data covers the number of RTAs in 76 provinces and the capital Bangkok. A total of 20,226 RTAs were reported, resulting in 15,217 non-fatal and 5,009 fatal accidents. The number of RAIs is a response variable and a set of predictors including road type, road surface, road section, and month of the year. To investigate the effect of festival events, accident months were regrouped into three categories. The Songkran festival is celebrated in April whereas the New Year holidays involve two months of December and January, representing the seven deadly days. Other months were grouped together as a baseline. It should be noted that lack of data availability prevents us from doing this study for the present period. However, data since 2019 may not reflect the number of accidents on roads due to the lockdown policy for the COVID-19 epidemic.

2.3. CONWAY-MAXWELL-POISSON REGRESSION MODEL

The CMP distribution is a two-parameter generalization of the Poisson distribution by allowing for overdispersion as well as underdispersion [15]. The CMP probability distribution function is shown as (2.1)

$$P(Y = y_i) = \frac{\lambda_i^{y_i}}{(y_i!)^{\nu} Z(\lambda_i, \nu)},\ y_i = 0, 1, 2, 3, ...,\ i = 1, 2, ..., n \tag{2.1}$$

where y_i is the human injuries at the ith accident, λ_i is the mean injury at the ith accident, $Z(\lambda_i, \nu) = \sum_{s=0}^{\infty} \frac{\lambda_i^s}{s!^{\nu}}$, $\nu \geq 0$ and ν is dispersion parameter. If dispersion parameter, $\nu = 1$ denotes equidispersion for the Poisson distribution, while $\nu < 1$ denotes overdispersion and $\nu > 1$ denotes underdispersion. Using an asymptotic approximation for $Z(\lambda_i, \nu)$, the expected value (2.2) and the variance (2.3) can be approximated as follows:

$$E(Y) = \lambda^{\frac{1}{\nu}} - \frac{\nu - 1}{2\nu} \tag{2.2}$$

$$Var(Y) = \frac{\lambda^{\frac{1}{\nu}}}{\nu} \tag{2.3}$$

Seller & Shmueli [15] further proposed the CMP regression (2.4) to model the relationship between a response variable Y and p explanatory variables, $X_1, X_2, ..., X_p$ via a function of $E(Y)$. The log link function, $\eta = logE(Y)$, is used to model the indirect relationship in a regression analysis.

$$logE(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p + \epsilon \tag{2.4}$$

where β_0, β_1,..., β_p are the regression parameters or coefficients. These coefficients are estimated by the maximum likelihood estimation (2.5) under the constraint $\nu \geq 0$, using a constrained nonlinear optimization tool. Their corresponding standard errors are derived from the generalized linear model.

$$logL = \sum_{i=0}^{n} y_i log\lambda_i - \nu \sum_{i=1}^{n} logy_i! - \sum_{i=1}^{n} logZ(\lambda_i, \nu) \tag{2.5}$$

The analysis in this study was carried out using the R programming language [16]. The R package, COMPoissonReg (available on CRAN) by Sellers, Borle & Shmueli [17] contains procedures for estimating the CMP regression coefficients and standard errors under the constant dispersion assumption, dispersion testing, and diagnostics. The latest version by Sellers, Lotze & Raim [18] has been available since 2019. The Pearson Chi-squared statistic is used to assess the adequacy of models. This statistic asymptotically follows a Chi-squared distribution with degrees of freedom equal to the difference in the number of parameters for a proposed model and a model with a constant only.

3. RESULTS

Table 1 shows the distribution of the injury counts for a total of 20,226 RTAs in 2015 Thailand. Over the year of the study, 66.19% of accidents (13,387 injuries) resulted in one person being injured and 9.05% (4,338 injuries) injured at least two people. 24.77% of accidents caused no human injuries, marked with a value of zero injury count. Table 2 shows the number of RTAs according to the explanatory covariates studied. RTAs occurred more often on the Local Administration Roads (43.29%) and the National Highway (43.19%) and less on the Rural Roads (13.53%). Dry surfaces have the most RTAs (92.70%). The largest number of RTAs were observed on curve sections (67.97%), following by straight sections (19.78%), and intersections (12.25%). RTAs were reported in ascending order during Songkran festival (18.28%), New Year Celebration Festival (23.23%) and other periods (58.49%).

TABLE 1. Distribution of injury counts due to RTAs in Thailand for 2015

Observed injury count	RTAs (percentage)	Human injury (percentage)
0	5,009 (24.77)	0 (0.00)
1	13,387 (66.19)	13,387 (75.53)
2	1,408 (6.96)	2,816 (15.89)
3	286 (1.41)	858 (4.84)
4	81 (0.40)	324 (1.83)
5	23 (0.11)	115 (0.65)
6	13 (0.06)	78 (0.44)
7	10 (0.05)	70 (0.39)
8	4 (0.02)	32 (0.18)
9	5 (0.02)	45 (0.25)
Total	20,226 (100.00)	26,099 (100.00)

TABLE 2. Variable definitions and descriptive statistics of RTAs in Thailand for 2015

Variable	Description	RTAs (percentage)
Road type	1: National Highway	8,735 (43.19)
	2: Rural Road	2,736 (13.53)
	3: Local Administration	8,755 (43.29)
Road surface	1: Dry	18,750 (92.70)
	2: Wet	1,476 (7.30)
Road section	1: Curve	13,748 (67.97)
	2: Intersection	2,478 (12.25)
	3: Straight	4,000 (19.78)
Festive period	1: New Year Festival	4,699 (23.23)
	2: Songkran Festival	3,697 (18.28)
	3: Other	11,830 (58.49)

Table 3 shows how well the Poisson and CMP distributions fit to the injury data. It was surprising that our data was under-dispersed. Consequently, the CMP distribution provided a better fit to the data than the Poisson distribution. Moreover, the presence of excessive zeros can be assessed. Since the number of observed accidents with zero injury were not greater than expected counts in both theoretical models, there was no signal of zero-inflation. Table 4 shows the factors affecting human injuries using the CMP regression. The dispersion parameter was more than one ($e^{1.221} = 3.390$) indicating underdispersion in our count data. Based on the Pearson Chi-squared statistic (Chi-squared value = 9,365.160, df = 20,217), the model provided an adequate fit to the data. Three variables were found to be statistically significant with p-value less than 5%. In terms of driving on different road types, we found that both estimated coefficients were negative. It suggested that the injury counts on the National Highways and the Rural Roads compared to the Local Administration Roads were lower. In more interpretable term, mean number of injuries decreased by 26% (since 1-0.740) on the Highways and by 27.3% (since 1-0.727) on the Rural Roads, when holding other covariates constant. We compared the RAIs based on road section. The number of injury counts were in general 1.363 times higher at intersections than on straight road sections. There was no evident difference in the number of injuries between curve sections and straight sections. The months of the festival caused more accident injuries than other non-festive months. The Songkran festival had a larger coefficient, suggesting that higher injuries (1.316 times) were found during the New Year celebration, even more (1.460 times) in the Songkran festival. Road surface was not statistically significant.

TABLE 3. Empirical, Poisson, and CMP distributions of injury counts

Observed injury count	Empirical	Poisson	CMP
0	5,009	8,420	5,481
1	13,387	7,379	11,928
2	1,408	3,233	2,657
3	286	944	156
4	81	207	4
5	23	36	0
6	13	5	0
7	10	1	0
8	4	0	0
9	5	0	0
Total	20,226	20,226	20,226
Mean	0.876	0.876	0.919
Variance	0.472	0.876	0.385
Poisson parameter	-	0.876	2.176
Dispersion parameter	-	1	3.288
$Z(\lambda_i, \nu)$	-	-	3.690
Chi-squared	-	10,897.21	9,586.27
d.f.	-	20,225	20,224

TABLE 4. Results from the CMP regression at significant level of 5% (indicated by *)

	Coefficient	Estimate	Standard error
Intercept	0.845	2.327	0.046
Road type			
(Baseline: Local Administration Road)			
1: National Highway	-0.301	0.740	0.024*
2: Rural Road	-0.319	0.727	0.035*
Road surface			
(Baseline: Wet)			
1: Dry	-0.071	0.932	0.042
Road section			
(Baseline: Straight)			
1: Curve	0.029	1.029	0.034
2: Intersection	0.310	1.363	0.028*
Festival period			
(Baseline: Other)			
1: New Year Festival	0.275	1.316	0.027*
2: Songkran Festival	0.379	1.460	0.030*
Dispersion parameter	1.221	3.390	0.011
Log-likelihood	-19,366.410		
Chi-squared	9,365.160		
d.f.	20,217		

4. DISCUSSION

The National Highway in Thailand is a primary highway that connect between provinces and regions and the Rural Road is a secondary road network connecting districts within provinces and from one city to another. The two networks cover 14% (100,759 kilometers) of road networks in the country. The Local Administration Roads connecting cities and towns cover 86% (597,667 kilometers). It is thus not surprising that there is a greater chance of RTAs occurring on the Local Administration Roads than other roads. Recently in 2020, 82% of the total number of RTAs in Thailand occur more frequently on the Local Administration Roads [19]. An effect of different road types on road safety was reported by Yahaya et al. [20]. They found that a higher number of non-fatal accidents were found on Urban roads while many fatal accidents were found on Rural Highways. Malin, Norros & Innamaa [21] carried out a comparison study of accident risk by road types. They summarized that relative accident risks were higher on motorways than on two-lane and multiple-lane roads.

In terms of road sections, RTAs occurred significantly at intersections. A similar result was reported in the United Kingdom for the year of 2017 [22]. For the severity analysis in China from 2000 to 2010, RTAs occurred more often at intersections, according to Wang, Liu, Ma, Zhang & Cong [23]. Similarly, Eboli, Forciniti & Mazzulla [24] summarized that intersection road accidents resulted in non-fatalities on the data of road accidents occurred in Italy during 2016.

In our study, road surface was not statistically significant. Thailand has a typically tropical climate consisting of two seasons: dry (summer and winter 9 months) and wet (raining 3 months). An adverse weather due to heavy rainfall may lead to a decrease in injured accidents since drivers will be more cautious. This explanation is supported by a study of crashes under bad weathers (such as snow, ice, or wet) on rural highways in North Carolina from 2009 to 2013 [25]. Other studies reported different results. Sangare et al. [22] and Yu et al. [26] concluded that the road surface variable (dry, wet, snow, ice and other) has a minor impact on accidents. Malin, Norros & Innamaa [21] studied 43 Finnish main roads during 2014 to 2016 for different levels of slippery road. The results showed that the relative accident risk was over four times higher for slushy road conditions and over two times higher for slippery and very slippery road conditions.

The injury counts per accident evidently increases during festive periods. This situation is worsening as a recent report in 2020 by the Transport Statistics Sub-Division, Planning Division, Department of Land Transport [27] and an official website http://roadsafety.-disaster.go.th operating by the Road Safety Center, the Department of Disaster Prevention and Mitigation of Thailand, Ministry of Interior [28]. The number of road deaths during the Songkran festival has risen from 364 deaths in 2015 to 418 in 2018. The number of injuries ranged from 3,559 to 3,897 persons. Over the period of three years, the number of road deaths during the New Year holidays has risen from 341 deaths in 2015 to 423 in 2018. The number of injuries ranged from 3,117 to 4,005 persons. Drink-driving is a major cause of injuries and deaths during the festive seasons, followed by speeding. Based on a report by WHO in 2020 [29], Thailand has achieved in announcing a comprehensive approach for controlling the COVID-19 epidemic in April 2020. It not only prevents the spread of infectious diseases, but also saves the lives of road users in just one month, with a 50% reduction in road deaths over the past five years. During the seven deadly days, a total of 1,260 injuries and 167 deaths in Songkran festival and 3,499 injuries and 373 deaths for the New Year holidays were reported [30], which was clearly lower than in previous years. Such evidence indicates that reducing the vehicle volume on the road and access to alcohol can greatly reduce road traffic deaths and injuries. To solve the problem of road injury and death in Thailand, sustainable strategies for road safety management are urgently needed.

5. CONCLUSION

This research studied the number of injuries due to road accidents in Thailand in 2015. The count regression model used in the study was the CMP regression. It is not surprising that the equidispersion criterion is not satisfied with the traditional Poisson distribution. Our data was under-dispersed since the mean was greater than the variance. This is rare in road safety analysis where overdispersion is common. Therefore, the CMP regression model was used to account for underdispersion in the data. The results show that the expected injuries per accident was 0.919, or about 92% of accident victims injured. Three independent variables: road type, road section, and festival period were significant factors affecting RTAs in which people were injured. Our findings highlight that festival month plays a critical role in raising unusual human injuries on the road. Road safety managers should pay more attention on these months, not only in Thailand but also other countries where special events occur. Other variables such as human factors (e.g., gender, age, driver's license, alcohol use, drug use) and environmental factors (e.g., road conditions, accident location, climate, daytime, nighttime) can be analyzed in further studies to improve a model accuracy. Other count regression models, such as the generalized Poisson regression, may be considered.

Acknowledgements

The authors acknowledge with thanks the availability of data coming from the Department of Disaster Prevention and Mitigation. The authors also thank the editor along with anonymous reviewers for their useful comments. This work was jointly supported by Thammasat University Research Unit in Data Learning and Thammasat University Research Fund [grant number TUFT 043/2563].

References

[1] World Health Organization (WHO), Global Status Report on Road safety 2015, Geneva, Switzerland, (2015).

[2] P. Taneerananon and P. Klungboonkrong, Thailand Road Safety Crisis: Time for Urgent Actions, presented at the 20th National Convention on Civil Engineering, Thailand. 14 (2015) 110.

[3] Y. Tanaboriboon, T. Satiennam, Traffic accidents in Thailand, IATSS Research. 29 (2005) 88-100.

[4] Thailand Road Safety Observatory, Festival Road Accidents. Accessed on: Oct. 27, 2021. [Online]. Available: http://trso.thairoads.org

[5] Transport and Traffic Policy Plan Office. Road Accident Situation Analysis Report of the Ministry of Transport 2019, (2020).

[6] M.M.A. Manan,A. Vrhelyi, Motorcycle fatalities in Malaysia,IATSS Research. 36 (2012) 30-39.

[7] M. Yu, C. Zheng, C. Ma, J. Shen,The temporal stability of factors affecting driver injury severity in run-off-road crashes: A random parameters ordered probit model with heterogeneity in the means approach, Acci. Anal. Prev. 144 (2020) 105677.

[8] F. M.A Hassouna, and I. Pringle, Analysis and prediction of crash fatalities in Australia, The Open Transportation J. 13 (2019) 134-140.

[9] P. Wu, L. Song, X. Meng, Influence of built environment and roadway characteristics on the frequency of vehicle crashes caused by driver inattention: A comparison between rural roads and urban roads, J. of Safe. Research, 2021 (in press).

[10] D. Lord, S.D. Guikema, S.R. Geedipally, Application of the Conway-Maxwell-Poisson generalized linear model for analyzing motor vehicle crashes, Acci. Anal. Prev. 40 (2008) 1123-1134.

[11] R. Elvik, F. Sagberg, P.A. Langeland, An analysis of factors influencing accidents on road bridges in Norway, Acci. Anal. Prev. 129 (2019) 1-6.

[12] D. Lord, S.R. Geedipally, S.D. Guikema Extension of the application of Conway-MaxwellPoisson models: analyzing traffic crash data exhibiting underdispersion, Risk Anal. 30 (2010) 12681276.

[13] World bank, People. Accessed on: Nov. 20, 2021. [Online].
Available: https://datatopics.worldbank.org/world-development-indicators /themes/people.html #population

[14] ThailandBusinessNews, Thailand ranks 35th in Travel and Tourism Competitiveness Report 2015. Accessed on: Nov. 20, 2021. [Online]. Available: https://www.thailand-business-news.com/travel/50749-thailand-ranks-35th-in-travel-and-tourism-competitiveness-report-2015.html

[15] K. F. Sellers, G. Shmueli, A flexible regression model for count data, Ann. Appl. Stat. 4 (2010) 943-961.

[16] R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Australia; 2021. http://www.Rproject.org (version 4.1.0).

[17] K. F. Sellers, S. Borle, G. Shmueli, The COM-Poisson model for count data: a survey of methods and applications, Appl. Stochastic Models Bus. Ind. 28 (2012) 104-116.

[18] K. F. Sellers, T. Lotze, and A. Raim, Conway-Maxwell Poisson (COM-Poisson) Regression, R Package Version 0.7.0, 2019. Accessed on: Jul. 1, 2021. [Online]. Available: https://CRAN.R-project.org/package=COMPoissonReg

[19] The Secretariat of the House of Representatives, Road safety and transportation by Transport Commission, 2020. Accessed on: Oct. 1, 2021. [Online]. Available: https://dl.parliament.go.th/handle/lirt/564558

[20] M. Yahaya, R. Guo, W. Fan, K. Bashir, Y. Fan, S. Xu, X. Jiang, Bayesian networks for imbalance data to investigate the contributing factors to fatal injury crashes on the Ghanaian highways, Acci. Anal. Prev. 19 (2021) 105936.

[21] F. Malin, I. Norros, S. Innamaa, Accident risk of road and weather conditions on different road types, Acci. Anal. Prev. 122 (2019) 181-188.

[22] M. Sangare, S. Gupta, S. Bouzefrane, S. Banerjee, P. Muhlethaler, Exploring the forecasting approach for road accidents: Analytical measures with hybrid machine learning, Expert Syst. Appl. 167 (2021) 113855.

[23] D. Wang, Q. Liu, L. Ma, Y. Zhang, H. Cong. (2019). Road traffic accident severity analysis: A census-based study in China. J. of Safe. Research, 70, 135-147.

[24] L. Eboli, C. Forciniti, G. Mazzulla. (2020). Factors influencing accident severity: an analysis by road accident type. Transportation Research Procedia, 47, 449-456.

[25] L. Gong, W. (David) Fan, Modeling single-vehicle run-off-road crash severity in rural areas: Accounting for unobserved heterogeneity and age difference, Acci. Anal. Prev. 101 (2017) 124-134.

[26] M. Yu, C. Zheng, C. Ma, J. Shen, The temporal stability of factors affecting driver injury severity in run-off-road crashes: A random parameters ordered probit model with heterogeneity in the means approach, Acci. Anal. Prev. 144 (2020) 105677.

[27] Transport Statistics Sub-Division, Planning Division of Department of Land Transport, 2020, Analysis of accident statistics during Songkran festivals 2017-2019 (nationwide). Bangkok, Thailand.

[28] Road Safety Center, the Department of Disaster Prevention and Mitigation of Thailand, Ministry of Interior. Statistics for road accidents. Accessed on: Oct. 20, 2021. [Online]. Available: http://roadsafety.disaster.go.th

[29] World Health Organization (WHO). Thailands status against 12 global road safety performance targets, (2020).

[30] Department of Disaster Prevention and Mitigation Ministry of Interior Thailand, Road traffic accident statistics in Thailand. Accessed on: Jan. 28, 2021. [Online]. Available: http://roadsafety.disaster.go.th/in.roadsafety-1.196/