A new look into development of accident prediction model for mountainous highways in Thailand

Assistant Professor Dr. Rattaphol Pueboobpaphan

Suranaree University of Technology

Summary

Road accidents have caused substantial loss in economic and human resources worldwide. There have been several attempts to develop accident prediction models in order to understand the factors that cause the accidents and provide an insight about effective measures to alleviate the accidents. So far, general findings from past studies are that the amount and the severity of accidents in each particular road segment are affected by, among others, the geometric characteristics of that same particular segment. However, in case of mountainous highways, the geometric characteristics of a few hundred meters before and after the subject segment may also be relevant. Based on this premise, this research explicitly considers the geometric characteristics of the neighbor road segments when developing the accident prediction model for mountainous highways using Poisson and Negative Binomial regression. Results show that the gradient of the neighbor segments has strong contribution to the total number of accidents and severities.

Aim of Research

The final goal of this research is to have a comprehensive understanding on what are the factors that are responsible for the accidents on mountainous highways so that the proper accident remedial measures can be developed. With this research goal in mind, the specific objectives of this research are as follows:

- To explore the relationship between the geometric characteristics of the neighbor road segment (or sub-segments) and the number and severity of the accidents on the subject segment. In particular, the gradient of the neighbor (sub-)segment is the main geometric characteristic that will be the focus of this research,
- To develop the accident prediction model for mountainous highways that considers explicitly the characteristics of the neighbor (sub-)segments.

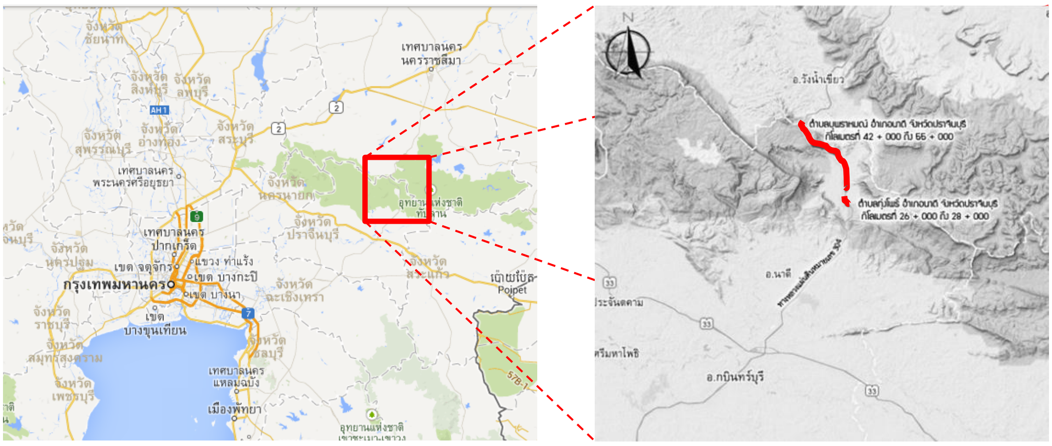
Method of Research & Progression

To accomplish the above objectives, the following research processes were adopted.

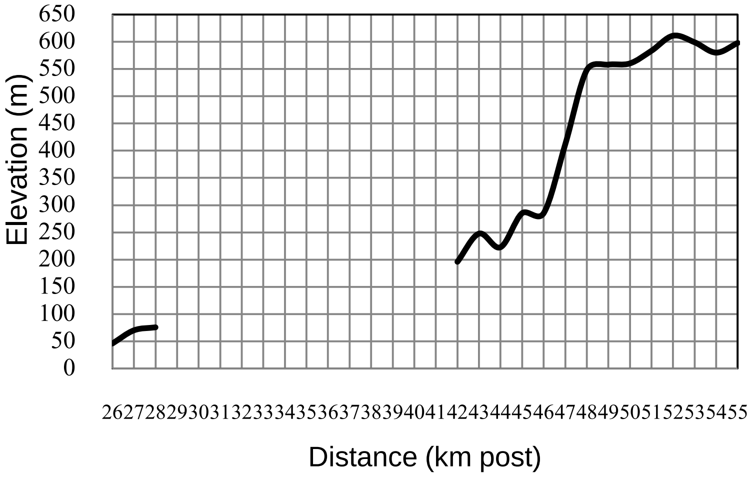
Literature review
To review existing literature in order to gain more insight about what factors that affect the accidents on highways both in the flat terrain and the mountainous areas, as well as the modeling approaches available in the literature. We found that Poisson and Negative Binomial Regressions are the two most employed modeling framework.

Selection of study area
A section of undivided highway on the highway number 304 between Kabinburi and Wang Nam Keaw was chosen as a study area. This corresponds to the section between the 26th –

28th kilometer post and the section between 42nd – 55th kilometer post. Figure 1 shows the location of the study area and the elevation of the chosen section.



(a) Location of the study area (b) Chosen study section



(c) Elevation of the study section

Figure 1 Location of the study section and elevation

Data collection

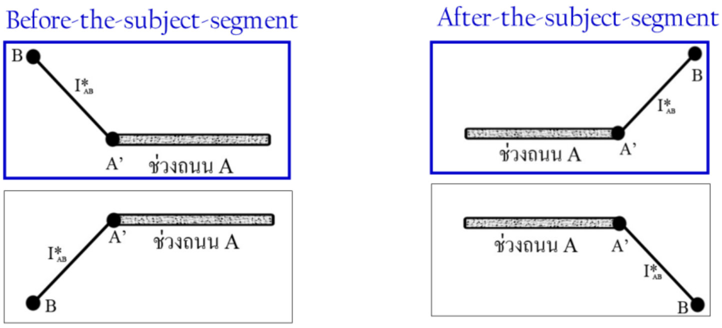
The following data were collected and used in this study.

- Accident statistics: this research used the database of accident statistics from the Bureau of Highway Safety, Department of Highway, Prachinburi Highway District, and also from the police report. The statistics include the number of accidents, number of injuries, and number of fatalities between 2006 and 2011 (6 years) of the study highway section. These aforementioned statistics were then manipulated to obtain the corresponding statistics for each sub-section per year.
- Traffic data: the data of traffic volume and proportion of heavy vehicles were obtained from the Bureau of Highway Safety, Department of Highway.
- Highway geometric: the highway geometric characteristics of each sub-section were extracted from the detailed construction plan and profile obtained from the Prachinburi Highway District. The data include number of lane, minimum horizontal curvature, number of horizontal curve, number of vertical curve, maximum design speed, gradient etc.

Data manipulation

Data collected from the related agencies and from the field were extracted into a digital format. The whole length of the study section was divided into smaller segments (every one kilometer) and several important geometric characteristics and accident statistics were manipulated accordingly. In particular, highway gradient was considered separately for the local and the neighbor. The local gradient is defined as the gradient of the segment under consideration (called as the “subject segment”). The neighbor gradients are defined as the gradient of the segment just right before and after the subject segment.

- Localized gradient: three measures were used to represent the localized gradient. These include the maximum gradient (%), average gradient (%), and the sum of vertical displacement (m).
- Neighbor gradient: this research considered the gradient of the small segments right before and after the subject segment. The gradient was averaged over the distance of 100, 200, 300, 400, and 500 meters before and after the subject segment. In addition, this research considers the effect of the neighbor segment only when it is a descending segment toward the subject segment. The concept of this consideration is illustrated in figure 2.



$$I_{BA'} = \begin{cases} \left|I^*_{BA'}\right|, if\ I^*_{BA'} < 0 \\ 0, otherwise \end{cases} \qquad I_{A'B} = \begin{cases} I^*_{A'B}, if\ I^*_{A'B} > 0 \\ 0, otherwise \end{cases}$$

Figure 2 Manipulation of the neighbor’s gradient

Model and variables

Poisson regression and Negative Binomial regression were used to model the relationship between the accident statistics and the explanatory variables in this research which can be described in the following form.

$$y = Exp(\beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n)$$

Where y is the accident statistics (no. of accidents, no. of injuries, and no. of fatalities), x_i denotes the i-th explanatory variable, β_i denotes the coefficient of the variable i. The variables used during the model development process are summarized in Table 1. Several model structures were tested and compared in order to search for the best possible model.

Table 1 Description of the dependent and independent variables

Dependent variables			
Accidents	No. of accidents (times/year/km)		
Injury	No. of injuries (persons/year/km)		
Fatality	No. of fatalities (persons/year/km)		
Independent variables			
AADT	Annual average daily traffic (1000 veh/day)	**BFVG1**	100 m Before-the-subject-segment gradient (%)

HV	Proportion of heavy vehicles (%)	**BFVG2**	200 m Before-the-subject-segment gradient (%)
DS	Max design speed (km/hr)	**BFVG3**	300 m Before-the-subject-segment gradient (%)
LANE	No. of lane (lanes)	**BFVG4**	400 m Before-the-subject-segment gradient (%)
HC	No. of horizontal curve per km (km^{-1})	**BFVG5**	500 m Before-the-subject-segment gradient (%)
VC	No. of vertical curve per km (km^{-1})	**AFVG1**	100 m After-the-subject-segment gradient (%)
RAHC	Min. radius of horizontal curve (m)	**AFVG2**	200 m After-the-subject-segment gradient (%)
VG	Average gradient (%)	**AFVG3**	300 m After-the-subject-segment gradient (%)
MAXVG	Max gradient (%)	**AFVG4**	400 m After-the-subject-segment gradient (%)
DIFVG	Sum of the vertical difference (m)	**AFVG5**	500 m After-the-subject-segment gradient (%)

Results of Research

It was found that the Negative Binomial regression is more suitable in modeling the accident statistics in this research than the Poisson regression. The results of the final models are shown in Table 2.

Table 2 Results of the final model

Variables	Coefficient of Negative Binomial Regression Models		
	No. Accidents	No. Injuries	No. Fatalities
Intercept	-1.145 (0.100)	-4.055 (0.000)	-7.357 (0.000)
RAHC	-0.003 (0.026)	-	-
MAXVG	-	0.519 (0.003)	0.796 (0.002)
DIFVG	0.014 (0.002)	-	-
AFVG4	-	-	0.414 (0.000)
AFVG5	0.316 (0.000)	0.494 (0.000)	-

Remarks: the value shown in parenthesis is the p-value, which reflects the level of significance of the coefficient

From the results, it can be seen that the minimum radius of the horizontal curve is one of the significant factors that contribute to the number of accidents. The negative value indicates that the larger the radius of the horizontal curve, the less the number of accidents. In addition, the variables that characterize about the gradient of the subject segment (MAXVG and DIFVG) are also responsible for the number of accidents and their severity. Moreover, the results also witness a significant effect of the gradient of the neighbor around 400-500 meters right after the subject segment on the number of accidents and their severity. All gradient variables have a positive value, indicating that it is riskier and more dangerous if the subject segment as well as its neighbor are steeper. In this study, only the neighbor right after the subject segment, not the neighbor before the subject segment, that has a significant effect. This is intuitively reasonable as the elevation of the neighbor after the subject segment is higher than that of the subject segment.

Future Areas to Take Note of, and Going Forward

This research was conducted using a small set of data from highway 304 in Thailand. It is worth to explore whether or not the findings from this research can be generalized to other mountainous highway sections.

Means of Official Announcement of Research Results

The results from this research are about to be submitted to one of the leading journals in safety and accident analysis such as Accident Analysis and Prevention, Safety Science, Journal of Safety Research, IATSS Research.