A PROBABILISTIC MODEL FOR PREDICTING TRAFFIC ACCIDENTS BASED ON THE PARAMETERS OF THE THREE-DIMENSIONAL GEOMETRY OF THE ROAD PROFILE

Baghdasaryan I.
Junior member of the Institute of Applied Problems of Physics, National Academy of Sciences of Armenia, Yerevan, Armenia
Barseghyan A.
Department of Economics and Law, National University of Architecture and Construction of Armenia, Yerevan, Armenia
https://doi.org/10.5281/zenodo.19292681

Abstract

The article proposes a probabilistic model for assessing the risk of road accidents on sections of highways with three-dimensional pedestrian markings. The aim of the study is to develop an approach to quantifying the risk of an accident based on the parameters of the spatial geometry of the road profile and traffic conditions. The model is based on the integration of methods of three-dimensional geometric modeling of road infrastructure, analysis of curve radii, longitudinal and transverse slopes, visibility zones and statistical forecasting of accidents using logistic regression. The model also takes into account external factors, including the speed of vehicles, weather conditions, and time of day. To form a vector of features, key parameters of the road infrastructure were identified, the impact of which on the probability of accidents was assessed using weighting coefficients. The results show that small curve radii, significant longitudinal slopes and limited visibility range have the greatest impact on the risk of an accident. The practical significance of the study lies in the possibility of using the developed model in analyzing the safety of road infrastructure and identifying potentially dangerous sections of highways. The use of three-dimensional pedestrian markings can help reduce the speed of movement and increase the attention of drivers, which is confirmed by the results of the analysis.

Keywords: 3D marking, road geometry, logistic regression, binary classification, accident risk, spatial modeling, road safety, integral risk

1. Introduction

Road traffic accidents remain one of the main causes of injuries and deaths on the roads in many countries of the world. Improving road safety is one of the most important tasks of transport engineering and road infrastructure management [11]. One of the key factors in accidents is failure to observe safe driving speeds on sections of roads with complex geometries, including sharp turns, longitudinal slopes and limited visibility. In recent years, special attention has been paid to the development of new methods to improve road safety. One of the modern solutions is the use of three-dimensional (3D) pedestrian markings that create a visual effect of a three-dimensional obstacle on the surface of the roadway [1]. Such solutions are aimed at attracting the attention of drivers and reducing the speed of vehicles on potentially dangerous sections of the road.

The issues of forecasting road accidents are actively being investigated in transportation engineering. Statistical methods, including logistic regression models, are widely used to analyze the probability of accidents [2,7,8]. These models make it possible to quantify the impact of various factors, such as road geometry, speed, weather conditions and visibility conditions, on the likelihood of an accident [4].

Modern research also actively uses machine learning methods to analyze accidents and identify dangerous sections of the road network [5,10,14,15,16]. The use of such methods makes it possible to take into account the complex interrelationships between the parameters of the road infrastructure and the characteristics of the traffic flow [3, 12]. In addition, the use of three-dimensional models of road infrastructure makes it possible to more accurately take into account the spatial geometry of the road, including slopes, curve radii and visibility zones, which helps to more accurately identify potentially dangerous sections of the road network. In recent years, accident forecasting models based on machine learning and big data analysis have been actively developing [15,16].

With the increasing intensity of traffic flows and the increasing complexity of the road infrastructure, the development of methods for predicting road accidents based on the analysis of geometric road parameters and behavioral factors of drivers is becoming particularly relevant. Modern modeling methods allow us to take into account the complex relationships between the characteristics of the road environment, traffic conditions and the likelihood of accidents. In this regard, the use of three-dimensional models of road infrastructure and statistical analysis methods opens up new opportunities for a more accurate assessment of the risk of accidents and the development of effective measures to improve road safety.

«Recent advances in integrating 3D road modeling with statistical and machine learning approaches enable more precise risk prediction. These methods consider not only geometric parameters but also dynamic traffic and environmental factors, improving safety planning and infrastructure design.».

The purpose of this study is to develop a probabilistic model for quantifying the risk of road accidents on road sections with three-dimensional pedestrian markings based on the parameters of the spatial geometry of the road profile.

2. Research methodology

2.1 Mathematical model of accident probability

Statistical analysis of accidents using 3D technologies makes it possible to more accurately reproduce the spatial characteristics of the road infrastructure and simulate the trajectory of a vehicle. Unlike traditional two-dimensional analysis schemes, three-dimensional models allow you to take into account the slopes of the road surface, the radii of curves and visibility zones, which significantly increases the accuracy of assessing the risk of accidents.

The mathematical model is based on the description of vehicle motion as a complex motion of a solid body in three-dimensional space. The model is based on statistical analysis and machine learning methods, which are widely used to predict accidents on highways [3, 12].

1.2. Calculation of stopping and stopping distance

The braking distance of a car on a section of road with a longitudinal slope is determined by the expression:

$$S = \frac{v^2}{(2g(f \pm Sin\ \alpha)} \quad (1)$$

Where **S** is the stopping distance, m;

v is the speed of the car, m\s;

g = 9.81 m\s^2 – acceleration of gravity

f is the coefficient of wheel adhesion to the road (dimensionless);

α - the longitudinal slope of the road (rad or hail);

The « + » sign is used for lifting;

« - » the sign is used for descent.

The longitudinal slope of the road is automatically calculated for each point of the vehicle's trajectory, displayed from the position of the camera (the driver's eye or virtual camera), the coordinates of the marking point on the road, as well as from the direction of movement. Therefore, α it is the angle between the line of sight and the road plane, or the angle between the direction of travel and the vector to the marking point.

The probability of a traffic accident occurring on the road section under consideration is determined using a logistic regression model, which is one of the most common binary classification methods in accident analysis tasks [2,7,8].

The probability of an accident on a section of road P is determined using logistic regression:

$$P = \sigma \cdot (W \cdot G + b) = \frac{1}{1+e^{-(W \cdot G + b)}} \quad (2)$$

where P is the probability of an accident on a road section (0-1);

σ - sigmoid function

$\mathbf{G = [R, \alpha, \beta, S_{vis}, V, W_{weather}, W^{time}]^T}$ - is the vector of site features;

$\mathbf{W = [w_R, w_\alpha, w_\beta, w_S, w_V, w_{weather}, w^{time}]}$ - vector of weights reflecting the influence of each feature;

b is the bias (the baseline probability level under neutral conditions).

The stopping distance (SD) is considered a critical, as well as the main aspect of road safety, and, as a result, ignorance of the principle of safe distance provokes the cause of one in three accidents in large cities [5]. When assessing the probability of an accident risk, special attention is paid to the stopping distance of the vehicle. It includes the driver's reaction distance, the delay time for the start of braking, and the braking distance itself until it comes to a complete stop. Ignorance or underestimation of a safe stopping route is one of the key causes of accidents on the roads, especially in areas with limited visibility, steep slopes or a small radius of curves.

A correct assessment of the stopping distance allows you to take into account how much time and distance the driver needs to safely respond to a potential danger. Combined with factors such as speed, curve radii, and road slopes, it is possible to predict the risk of an accident on specific road sections much more accurately.

Thus, taking into account the braking distance in the model enhances its practical value and makes it possible to more effectively identify potentially dangerous areas on the highway where the use of safety measures, such as 3D markings, would be most justified.

$$S_{ocm} = S_p + S_c + S_t \quad (3)$$

The SD is the length of the route from the point of danger detection to the final cessation of movement, as well as the distance to the start of braking t_p, the delay time of braking, and the braking distance St (formula 3).

2. Risk factors and mathematical model.

3.1. The influence of bias

The following indicators are used to assess the influence of the geometric parameters of the road: radius of the curve of the road; longitudinal slope; transverse slope; range of visibility; speed of movement; weather conditions; time of day.

The use of logistic models makes it possible to quantify the impact of various factors of road infrastructure and traffic conditions on the likelihood of emergency situations [4.10].

The linear combination of features used in the sigmoid function, i.e. the total weighted risk indicator will take the following form:

$$Z = W \cdot G + b = \sum_{i=1}^{m} wigi + b \quad (4)$$

where **m** is the number of features in the vector G;

Z - indicates the overall threat level at the site; as Z increases, the probability of an accident increases.

Formulas are used to assess the influence of individual road geometry factors:

a) the effect of the longitudinal slope on the braking distance

$$K_{slope} = 1 + a \cdot |\alpha|^b \quad (5)$$

where **a** and **b** are empirical constants.

b) the effect of the radius of the curve:

$$K_{curve} = \{c/R\}, \text{ при } R<Rmin \quad (6)$$

where **c** is the coefficient of influence of the curve; R_{min} is the minimum safe radius.

2.2. The impact of visibility

The effect of visibility using the inverse dependence of the sigmoid:

$$K_{vis} = \frac{1}{1+\exp(k \cdot (Sfact - Sstop)} \quad (7)$$

where S_{fact} is the actual visibility, m;

S_{stop} - stopping distance, m;

k is the steepness coefficient (the greater the k, the sudden jump in the risk of RTA).

The essence of the formula is as follows

- with $S_{fact} \ll S_{stop}$, then $K_{vis} \approx 1$ is a high risk;

- with $S_{fact} \gg S_{stop}$, then $K_{vis} \approx 0$ is a low risk.

The technique consistently reproduces the justified risk of an accident.

2.3. The influence of the curve radius

When exposed to the radius of the curve, an exponential dependence (pattern of change) is applied, which smoothly increases the risk when the radius of the curve decreases as follows:

$$\mathbf{K_{curve} = exp\left(-\frac{R}{R_0}\right)} \quad (8)$$

Where **R** – is the radius of the curve, m;

$\mathbf{R_0}$ **–** is the characteristic radius (empirical constant) at which the risk is the considered moderate,

$\mathbf{K_{curve}}$ - increases at R_1, for very small curve radii.

The exponential formula is simple, easy to understand, and effective for creating models.

2.4. Integral risk model

The total integral risk, taking into account geometric and external factors, has the form:

$$\mathbf{R_{total} = P \cdot K_{slope} \cdot K_{curve} \cdot K_{vis}} \quad (9)$$

This is what the final stage of the accident probability looks like, taking into account the synergistic effect.

Thus, the model takes into account the influence of road geometry, visibility, speed, and external factors, allowing you to quantify the probability of an accident on each road section.

3D zebra pedestrian markings are a type of road marking that creates the illusion of a three-dimensional obstacle due to graphic techniques that form the effect of a bulge on a flat surface of the roadway. This visual effect is based on the mechanisms of optical perception and is aimed at increasing the attention of drivers and reducing the speed of vehicles. The driver's consciousness captures the differences between "chiaroscuro", three-dimensionality and modeling as a space. When depicting zebra stripes with "chiaroscuro" and displacement, the driver fixes them as an obstacle on the road. This representation is based on the mechanisms of binocular and monocular depth perception, as well as the processing of light and shadow by the visual cortex. If the car driver notices reality in 3D, i.e. perceives depth, it should be taken into account that binocular vision provides a fusion reflex (merging two images into one) [3]. In this sense, the driver's mental abilities are understood to combine images into a single whole [7]. Binocular vision, unlike monocular vision, gives a person the opportunity to navigate normally in space. With monocular vision, complete loss of functioning of one eye, severe strabismus, amblyopia ("lazy eye") or the consequences of injuries (operations) [4]. Whereas binocular vision creates a three-dimensional image, monocular vision forms a plane, i.e. a 2D picture. If the mind captures a possible danger, then the orientation reflex is activated, attention is increased, then the driver automatically reduces speed. The whole process is associated with the work of emotional reactions and an instant threat monitoring system.

If there are strange objects on the roadway on motorways, then they concentrate more intensely here, they are easier to fix, and they occasionally improve concentration. Such a study is called the novelty effect, and it is scientifically based in the field of behavioral psychology and transportation safety.

The review analysis allows you to mathematically substantiate the distance at which the driver noticed the obstacle from afar. The software simulates viewing areas from the point of view of the car driver to the elements of road construction. In 3D markup, the integrated risk management system (ERM) on highways is considered a comprehensive digital plan with geometric data, as well as with characteristics and probabilistic safety indicators.

To assess the impact of individual road infrastructure factors on the probability of an accident, feature vectors G and corresponding weighting coefficients W were determined (Table 1). To quantify the risk of road accidents on the road sections under consideration, the key parameters forming the feature vector G and the corresponding weighting coefficients W were determined.

Each factor characterizes a specific aspect of the road infrastructure or traffic conditions and affects the likelihood of an accident. Thus, *the radius of the curve (R)* reflects the curvature of the road section - reducing the radius increases the risk of accidents. *The longitudinal slope (α)* shows the effect of ascents and descents on braking distance and driving dynamics, while *the transverse slope (β)* has a more moderate effect. *The range of visibility (S_{vis})* is critically important for a timely driver response: reducing it increases the risk of a collision. *The speed of movement (V)* directly increases the probability of an accident, and *weather conditions ($W_{weather}$)* and *time of day (W_{time})* modify the risk of an accident, taking into account the influence of the external environment and visibility.

Table 1.

The main indicators of the factors are vectors of features and weighting coefficients

Names	**Symbols**	**Measures of quantities**	**Weight, w_i**	**Notes**
Radius of the curve	R	m	0.45	Reducing the radius leads to an increase in danger
Longitudinal slope	A	rad	0.60	Steep descent increases the risk
Lateral slope	B	rad	0.25	Light impact
Range of visibility	S_{vis}	m	- 0.50	Keeping a distance reduces the risk of an accident
The speed of movement	V	m\s	0.55	As the speed increases, the threat increases
Weather conditions	$W_{weather}$	Immeasurable	0.35	Precipitation increases the risk
Time of day	W_{time}	immeasurable	0.20	At night, the probability of an accident is slightly higher

All indicators and their numerical values are presented in Table 1. Analysis of this table makes it possible to understand which factor has the greatest impact on the risk of an accident and how their combined impact forms an integral indicator of danger on a road section.

The proposed model reflects the cumulative effect of location, i.e. the synergetic effect of territorial circumstances (combining a steep curve with a steeply sloping descent), which cannot be said about traditional 2D schemes.

3. Features of visual perception of 3D markup

The effectiveness of three-dimensional pedestrian markings is related to the peculiarities of human visual perception. The illusion of a three-dimensional obstacle is formed through the use of chiaroscuro effects, perspective distortions, and lane shifts that create the visual effect of a bulge on a flat road surface [1,9,13].

The driver's binocular vision makes it possible to estimate the depth of space and the distance to objects, which increases the accuracy of orientation on the road [6, 13]. When perceiving 3D markings, the orientation reflex is activated, attention is increased, and the driver automatically reduces the speed of movement. In addition, psychological effects such as the novelty effect enhance the driver's concentration: an unusual visual element on the road instantly attracts attention and improves concentration. This is especially important in areas with limited visibility or complex road geometry.

Combined with an analysis of the geometric parameters of the road and an assessment of external factors (speed, weather conditions, time of day), taking into account the driver's visual perception makes the accident risk assessment model more accurate and practical. Thus, the integration of 3D markings into the risk forecasting system allows not only to identify potentially dangerous sections of the road, but also to take into account the behavioral aspect of the driver when assessing the likelihood of accidents.

4. Conclusion

As a result of the conducted research, a probabilistic model for assessing the risk of road accidents on sections of highways with three-dimensional pedestrian markings has been developed. The proposed approach is based on the integration of methods of three-dimensional geometric modeling of road infrastructure and statistical analysis using logistic regression**.**

The analysis showed that the combined effect of the geometric characteristics of the road and external factors forms a synergistic effect that significantly affects the likelihood of road accidents. In particular, the combination of small radii of curvature, significant longitudinal slopes, and limited visibility leads to a marked increase in the integral risk level.

The use of three-dimensional modeling makes it possible to more accurately take into account the spatial structure of the road profile and identify potentially dangerous sections of the road network. This significantly expands the analytical capabilities compared to traditional two-dimensional analysis methods.

The results obtained confirm the prospects of integrating statistical analysis, machine learning, and three-dimensional modeling methods to develop intelligent traffic safety systems.

The proposed model can be used in analyzing the safety of individual sections of highways, as well as in intelligent road infrastructure management systems. The results obtained can serve as a basis for developing practical recommendations on the placement of road infrastructure elements, including the use of three-dimensional pedestrian markings on potentially dangerous road sections. Further research may be aimed at expanding the set of model factors and using machine learning methods to improve the accuracy of predicting traffic accidents.

The proposed probabilistic model can be integrated into intelligent transportation systems to support real-time risk assessment and guide strategic placement of 3D pedestrian markings. Future work may focus on expanding the set of environmental and behavioral factors to enhance predictive accuracy.

Discussion of the results

An analysis of the results shows that the geometric parameters of the road and the speed of vehicles have the greatest impact on the likelihood of road accidents. At the same time, the combined effect of several factors forms a synergistic effect, in which the risk level increases significantly more than with the isolated influence of each individual parameter.

The synergistic effect of multiple risk factors indicates that interventions targeting combined geometric and environmental risks are more effective than isolated measures. Furthermore, incorporating behavioral aspects of drivers into predictive models enhances the accuracy of traffic safety assessments.

Scientific novelty and practical significance of the research

The scientific novelty of the research lies in the development of a probabilistic model for predicting road accidents, taking into account the parameters of the three-dimensional geometry of the road profile, visibility factors and external traffic conditions. The proposed approach makes it possible to integrate several critical risk factors into a single model, which provides a more accurate assessment of the probability of an accident compared to traditional methods.

The practical significance of the work is shown in the possibility of using the model in analyzing the safety of road infrastructure, identifying potentially dangerous areas and developing recommendations for the use of 3D pedestrian markings to increase driver attention and reduce speed.

Contribution of the authors

Irina Baghdasaryan – research concept, methodology development, mathematical modeling, writing the initial text of the article.

Angela Barseghyan – data analysis, visualization of results, editing and scientific review of the text.

Conflict of interests.

The authors declare that there is no conflict of interest.

List of abbreviations

3D - three- dimensional;

Traffic accident - a traffic accident;

ERM - risk management system;

Svis - range of visibility;

R is the radius of the curve;

V is the speed of movement.

References:

1. Bagdasaryan I.R., Barsegyan A.A. Current Application of Unusual Three-Dimensional Pedestrian Crossing for Road Safety. Scientific Works of the National University of Architecture and Construction of Armenia. Vol. 3(87), 2023, pp. 3–9.

2. Chen M., Chen M. Modeling Road Accident Severity with Logistic Regression, Decision Tree and Random Forest. Information, 2020.

3. Diakonov A.G. Machine Learning and Data Analysis. Moscow, 2024, p. 36.

4. Islam M., Reza I., Rahman M. Predicting Road Crash Severity Using Classifier Models and Crash Hotspots. Applied Sciences, 2022.

5. Kotsyubynska Y., et al. Machine Learning and Deep Learning for Predicting Traffic Crash Injury Severity: A Systematic Review (2014–2025). Journal of Road Safety, 2026.

6. Nikiforov A.S. Neurology: Complete Explanatory Dictionary. RoyalLib.com Electronic Library.

7. Prediction and Analysis of Road Accident Severity Using Logit Model and Machine Learning. Discover Applied Sciences, 2021.

8. Prediction of Crash Severity: Logit Models and Machine Learning Models. Transportation Engineering, 2025.

9. Rychkova S.I., Likhvantseva V.G. Monocular Depth Perception: Literature Review. 24(1), 2022, pp. 43–54. https://doi.org/10.33791/2222-4408-2022-1-43-54

10. Shiran G., Imaninasab R., Khayamim R. Crash Severity Analysis of Highways Based on Multinomial Logistic Regression, Decision Tree and Neural Network. Sustainability, 2021.

11. Timovsky A.A., Nesterenko V.B. Fundamentals of Vehicle Control and Road Safety. Kyiv: Arij Publishing, 2009, p. 146.

12. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Williams Publishing, 2020, p. 768.

13. Ochkov.net. Advantages of Binocular Vision. Available from: https://www.ochkov.net/wiki/preimushestva-binokulyarnogo-zreniya.htm (accessed 12 February 2026).

14. Yumak A., et al. Machine Learning Approach to Identify High-Risk Road Segments. Applied Sciences, 2025.

15. Wang Y., Li Z. Machine Learning Models for Road Accident Risk Prediction. Transportation Research Part C, 2024. doi:10.1016/j.trc.2024.104215

16. Zhang H. Road Safety Analysis Using Logistic Regression Models. Safety Science, 2023. doi:10.1016/j.ssci.2023.10596a