

The authors sincerely appreciate the reviewers' time and constructive comments to help us improve this manuscript. We have addressed the review comments as follows:

Response to Reviewer #1

1、As there are many machine learning models, the reasons for selecting the MLP should be specified.

Response: The authors would like to thank the reviewer for the comment. First of all, the MLP model used in this paper belongs to the set counting model, whose dependent variable accident frequency is the set counting result of accidents on a certain road length at a certain time, so the dependent variable is a non-negative positive value. Secondly, although the current machine learning model can achieve a better fit after the completion of the tuning, some of the results would still have negative values, which fundamentally conflicts with the modeling theory of the set counting model. In this regard, the neural network model has better operability, and the ReLU function (Rectified Linear Unit) was chosen as the activation function of the final output layer of the neural network model, as shown in eq. 1. Therefore, the MLP model is finally selected as the correlation model in this paper.

$$\text{ReLU}(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (1)$$

$$y_i = f \left(\sum_{h=1}^q w_{hj} \cdot b_h + \zeta_j \right) \quad (2)$$

And the introduction has been modified as follows:

This study aims to analyze the relationship between highway geometric characteristics and accident risk. First, the geometric alignment characteristics of the high-way plane and its longitudinal section is extracted from satellite maps by using the In-verse Micro Element clustering algorithms, the reliability of the method was verified on Nanfu highway in

Chongqing, China. Then a multi-layer perceptron neural network (MLP) approach is developed to model the relationship between the geometric characteristics and accident risk. The influence of road geometric linear variables on accident risk is further analyzed using the SHAP model. With the corresponding 36439 traffic accidents from 2010 to 2016. The negative binomial model is used as the benchmark for comparison. ◦

2、 It would be better to add a separate data section to describe the accident data as well as the basic freeway data.

Response: The authors would like to thank the reviewer for the comment. This study mainly focuses on the correlation between highway geometry and accident risk, so the accident data and highway data mentioned by the experts are the basic research data of the paper and should be reflected. The authors provide a brief description of both in Table 1 (215 lines, accident data) and Table 4 (258 lines, highway alignment data).

3、 Usually when developing the crash prediction model, the AADT should be considered as an exposure, which is not the case in this study. Please explain the reason.

Reponse: Thank you so much for the comment. For these four highway roads, there is no AADT data in the database from the operation companies. The authors will consider these parameters in further study.

4、 Page 12, line 347-351: the sentence needs to be rephrased. The Nanfu highway seems to be only used for verifying the geometric alignment extraction method.

Response: The sentence has been rewritten as follows.

This study aims to collect road design data in large-scale research and analyzes the accident risk of highway geometric alignment. Accordingly, a method based on satellite maps and clustering algorithms is proposed to calculate the geometric alignment of the highway plane and its longitudinal section and the reliability of the method was verified on Nanfu highway in Chongqing, China. With the corresponding 36439 traffic accidents which occurred from 2010 to 2016 being used as the research objects. The accident risk of the highway geometry was analyzed based on the SHAP and MLP theories. The conclusions are as follows:

5、 Please add the literatures that using MLP for road safety analysis.

Response: Thank you very much for the suggestions put forward by the reviewers. According to the expert opinions, relevant literature has been added in the introduction to explain

6、 Something seems to be missing for Figure 5(b)

Response: New figure 5 (b) is added in the paper. Figure 5(b) shows the global plot of the SHAP dependence plot of flat curve radius versus roadway unit accident frequency, including straight line roadway units and flat curve roadway units (shown in Figure 1 below). The aggregation of curve section units in Figure 5(b) is highlighted.

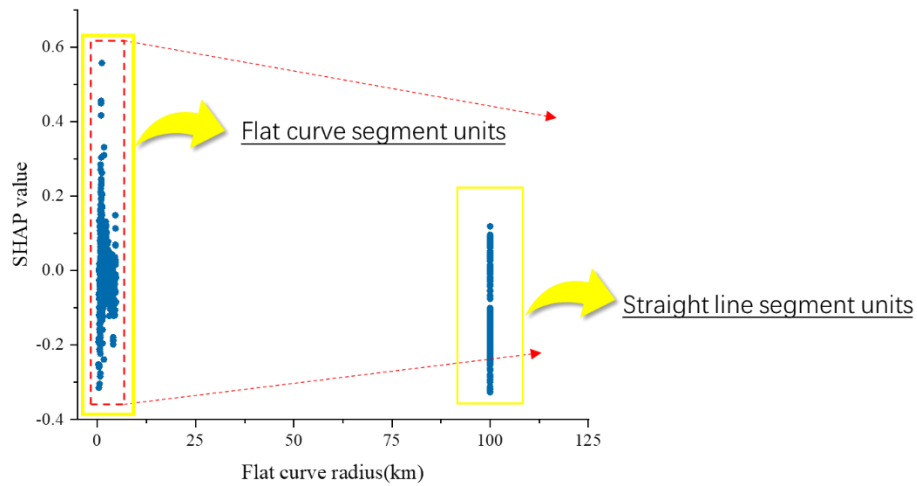


Figure 5(b)