Highway Planning and Design in the Qinghai–Tibet Plateau of China: A Cost–Safety Balance Perspective

Chengqian Li, Lieyun Ding *, Botao Zhong

Department of Construction Management, School of Civil Engineering and Mechanics, Huazhong University of Science and Technology, Wuhan 430074, China

Article history:
Received 23 July 2018
Revised 28 December 2018
Available online 16 March 2019

Abstract

Engineering designs for mountainous highways emphasize compliance checking to ensure safety. However, relying solely on compliance checking may lead designers to minimize costs at the expense of high risk indicators, since the overall risk level of the highway design is unknown to the designers. This paper describes a method for the simultaneous consideration of traffic safety risks and the associated cost burden related to the appropriate planning and design of a mountainous highway. The method can be carried out in four steps: First, the highway design is represented by a new parametric framework to extract the key design variables that affect not only the life-cycle cost but also the operational safety. Second, the relationship between the life-cycle cost and the operational safety risk factors is established within the range of acceptable risk, in order to help decision-makers. This approach was applied and examined in the Sichuan–Tibet Highway engineering project, which is located in a complex area with a large elevation gradient and a wide range of mountains. The experimental results show that the proposed approach significantly improved both the safety and cost performance of the project in the study area.

© 2019 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The Sichuan–Tibet Highway, which is located in the Qinghai–Tibet Plateau of China, is currently facing severe problems with traffic safety hazards and budget overruns. On the one hand, the Sichuan–Tibet Highway is one of the highest and most dangerous roads in the world. The lowest and highest altitudes along the way are about 500 and 5000 m, respectively. To address this topographic height difference, long vertical slopes and a large ratio of tunnels and bridges with extraordinary length have been employed in designs to cross the enormous mountainous terrain; however, these measures result in severe traffic safety hazards. Precipitous terrain and adverse weather conditions worsen the driving environment on mountainous highways and always cause a higher frequency and severity of traffic accidents than on urban highways [1]. Furthermore, the unit price of highway construction in this area is about twice that of an urban highway in neighboring areas, and the traffic mortality rate (per 105 people) is about 2–3 times higher than the national level [2,3]. Since the construction site for the Sichuan–Tibet Highway engineering project is located in a deserted mountainous area with an average altitude of over 3000 m, the manual labor and the costs for material and machinery transportation are much more expensive than in other areas.

To improve traffic safety, traditional highway planning and design often require a more conservative solution, including road-width extension and increased length of line, as a means of dealing with the large elevation difference. However, most of these improvements are costly [4]. Of course, the life-cycle cost of a highway depends on its length, geometric parameters, and infrastructure type (i.e., bridges, tunnels, etc.), which are profound latent factors in traffic safety [5–9]. Therefore, the two most significant challenges in mountainous highway planning and design are as follows: ① Dramatic topographic changes in mountainous areas make the search algorithm more likely to be trapped in a local optimum. ② The risks of cost overrun and traffic accidents are much
more significant in mountainous highways than in urban highways; therefore, an effective optimization method is needed that can not only improve safety and reduce cost, but also strike a balance between the two.

In the existing literature, most researchers have limited the optimization objective to cost performance [10–20], since it is obvious that an optimal road design will largely reduce the total cost of a highway, including the construction, maintenance, and user costs. Some researchers have incorporated safety issues as a social cost item (i.e., accident cost) [15,17] with the aim of achieving safety at the cost of usability. Others have focused on the interactive effects of risk factors. For example, advanced research shows that downward grades interact with fatigue-driving behavior in highway tunnels with a monotonous and uniform environment. However, non-interactive effects of risk factors are important factors in driver fatigue, which prolongs the reaction time of drivers and often leads to accidents. The coupling effects of high altitude for a mountainous highway crossing a plateau and poor ventilation in a tunnel may cause a driver to become anoxic, resulting in a traffic accident.

This paper describes a cost–safety optimal balance model to identify and optimize the highway planning and design variables that affect not only life-cycle cost but also operational safety. The model can not only be used for highway alignment problems, but also be extended to multi-objective design optimization for other high-risk engineering design tasks.

This work advances research on highway alignment optimization in several ways. First, it presents a new chromosomal representation form for highway planning and design, which is the basis of the cost and safety performance assessment provided here. Second, it develops an integrated assessment model by correlating the safety risk factors in design variables to the life-cycle cost items, which bridges the relationship between cost and safety. Third, it combines a fault tree analysis (FTA) with the regression method to model traffic accident causes from the perspective of highway engineering design. A generalized regression model is introduced to formulate the mathematical function between highway planning and design and the corresponding risk level (predicted crash rate) for an alignment. Fourth, a framework for an acceptable risk assessment is proposed to determine the applicability of optimal solutions in terms of traffic safety risk. Finally, a computing algorithm considering the global trend and local features of cost-efficient variations is proposed in order to strike a balance between cost and safety.

2. Formulations of the cost and safety functions

2.1. Identification of risk factors for safety analysis

Highway traffic safety is a complex objective that broadly depends on the drive–roadway–vehicle interaction [32]. Although the existing literature contains a great deal of analysis on the factors contributing to traffic accidents, few studies have been conducted from the perspective of engineering design. This section uses an FTA [33] to bridge the gap between engineering design details and traffic accidents.

Fig. 1 shows a fault tree that takes a traffic accident as the top event and the engineering design factors as the basic event, in order to develop an understanding of the logic leading from the engineering design variables to the undesired top event. It is notable that the fault tree in Fig. 1 is a newly proposed FTA diagram (designated herein as the NFTA) that was specially developed for the safety analysis of engineering designs. The NFTA has several differences between the NFTA and a traditional FTA diagram. First, the events in Fig. 1 are deterministic factors in a certain highway project, whereas the FTA method often uses the predetermined occurrence probabilities of basic events to deduce the probability of the top event. Second, the dependent relationships between variables are probabilistic rather than deterministic [34]. For example, even though “driver error” is an important factor leading to traffic accidents, the presence of driver error will not always result in a traffic accident. In this, the NFTA differs from the traditional FTA assumption.

The NFTA in Fig. 1 shows that traffic accidents mainly result from driver errors and vehicle-operating problems caused by poor road design.

Driver error refers to difficult situations in which poor highway design may have a negative impact on the performance of drivers. For example, super-long tunnels and a dim driving environment are important factors in driver fatigue, which prolongs the reaction time of drivers and often leads to accidents. The coupling effects of high altitude for a mountainous highway crossing a plateau and poor ventilation in a tunnel may cause a driver to become anoxic, resulting in a traffic accident.

Vehicle-operating problems refer to dangerous events in which poor highway alignment design increases the difficulty of vehicle operation and the probability of vehicle failure, even when drivers have sufficient experience and vehicles are in a good state of operation. Previous research indicates that design consistency plays a vital role in providing conformance between practical operation of a vehicle and driver expectancy [35]. A good design in terms of road consistency eliminates unexpected and abrupt changes, and thus effectively reduces the probability of an accident [27]. Design consistency can be classified into three main categories: vehicle stability, operating speed consistency on a single element, and operating speed consistency on successive elements. Vehicle stability analyzes whether the side frictions provided at a curve are enough to prevent a vehicle from skidding out.

Operating speed refers to the actual speed selected by drivers under free-flow conditions, which is always determined by the highway alignment geometrics. In addition to the above design...
consistency criteria, a large longitudinal slope and poor sight distance also bring great challenges to vehicle control and may lead to an unexpected crash.

Given that drivers’ mistakes and vehicle breakdowns are sometimes unavoidable, incorporating fault-tolerance facilities into a highway (e.g., road shoulders and emergency stopping strips) can effectively reduce the traffic accident rate.

Although the NFTA in Fig. 1 provides insight into factors in a highway-engineering project that can lead to traffic accidents, the inference approach of the traditional FTA method is inapplicable in this study, due to the major differences between the NFTA and an FTA. To solve this problem, the NFTA diagram was mapped onto a generalized linear-regression model in order to compute the traffic accident probability for a highway segment, given the engineering design information. The mapping algorithm is described as follows: First, the minimal cut sets [36], which refer to the minimal, necessary, and sufficient conditions for the occurrence of traffic accidents, are identified in the NFTA. Next, the minimal cut sets are converted to the explanatory variables in the regression model. In particular, the AND gate in NFTA will combine the information of the lower events to compute the state of upper events. Table 1 provides the risk factors deduced from the NFTA diagram, which will be used as the explanatory variables in the regression model.

The risk factors listed in Table 1 are important parameters that are associated with the planning and design of mountainous highways. It is worth noting that these risk factors not only affect road traffic safety, but also affect the life-cycle cost of highways.
Risk factors list for highway traffic accidents.

<table>
<thead>
<tr>
<th>Risk factors</th>
<th>Corresponding engineering design variables</th>
<th>Design variables relation with Construction cost</th>
<th>Maintenance cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dim tunnel environment</td>
<td>Lighting design in tunnel</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Long tunnels</td>
<td>Tunnel length</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Long tangents</td>
<td>Tangent length</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Poor ventilation in tunnel</td>
<td>Ventilation design in tunnel</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td>Poor vehicle stability</td>
<td>Pavement friction, superelevation, curve radius</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td>Poor driving consistency</td>
<td>Slope grade, curve radius, and roadside context</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td>Large steep slope</td>
<td>Slope grade and slope length</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td>Poor sight distance</td>
<td>Slope grade, curve radius, and roadside context</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td>Narrow shoulder width</td>
<td>Road shoulder width</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td>Insufficient emergency stopping strips</td>
<td>Distance between emergency stopping strips</td>
<td>√</td>
<td>×</td>
</tr>
</tbody>
</table>

2.2. Decision variables for preliminary highway design

Now that traffic safety risk factors have been identified in the section above, a representative framework for highway planning and design is developed in this section. Both the risk factors and the parameters related to life-cycle cost analysis can be extracted from this framework.

Let \( S(x_s, y_s, z_s) \) and \( E(x_E, y_E, z_E) \) be the start and end points of a highway design. The highway alignment is segmented into small homogeneous sections by inserting the points of intersection \( P_i = (x_{pi}, y_{pi}, z_{pi}, R_{pi}) \).

As shown in Fig. 2, an alignment is segmented into curves and tangents that are represented as \( \hat{E} = (C, T) \), where \( C \) and \( T \) are specified in Eqs. (2) and (4). Curves are represented by \( C \),

\[
C = [C_1, C_2, \ldots, C_n]
\]

which is comprised of a set of curved segments \( C_i \):

\[
C_i = [x_{Ci,u}, y_{Ci,u}, z_{Ci,u}, x_{Ci,v}, y_{Ci,v}, z_{Ci,v}, L_{Vi}, G_{Vi}], i = 1, 2, \ldots, n
\]

where the coordinates of transition points \((x_{CT,u}, y_{CT,u}, z_{CT,u})\) and \((x_{CT,v}, y_{CT,v}, z_{CT,v})\) can be calculated by geometric constraints [37].

![Fig. 2. Correspondence of horizontal alignment and vertical alignment.](image)

2.3. Estimation of life-cycle cost

Life-cycle cost \( f_{\text{LC}} \) is evaluated in the form of the annual average cost, as shown in Eq. (5). Let \( c_f \) be the capital recovery factor, which is usually set at 0.065 in China.

\[
f_{\text{LC}}(E) = c_f C_C(E) + C_M(E)
\]

where

\[
C_C(E) = K_{TC}L_T + K_{TC}L_L + C_R + K_{TC}(L_T + L_L)
\]

\[
C_R = C_{CB} + C_{CC} + K_{PB}L_{PB}
\]

\[
C_M(E) = C_{TM} + C_B = K_{ML}(L_T + L_L + L_R) + K_{OL_T}
\]

Given the alignment route information \( E \) and Geographic Information System (GIS) data, various highway structures (bridges, embankments, tunnels, or deep cuts) are automatically determined according to their life-cycle cost, as shown in Fig. 3.
For example, if the maximum life-cycle cost of a tunnel is lower than that of the deep cuts in a certain segment, a tunnel is selected as the ideal highway structure in this section, and vice versa. In this way, the lengths of various highway structures are determined. Let $L_T$ and $L_B$ be the length of tunnels and bridges, and let the unit prices of construction be defined as $K_{TC}$ and $K_{BC}$, respectively. The construction cost $C_C$ can be calculated as the sum of the related costs of the tunnel, bridge, road, and ancillary facilities, as shown in Eq. (6). The construction cost of road, $C_R$, equals the sum of the cost of the ground-fill $C_{GF}$ and of the cut, $C_{CC}$, and pavement construction, $K_{PA}L_{PA}$, the detailed formulation of which can be found in Ref. [37], and is not covered in this paper. In addition to the main structure construction, the total construction cost includes the cost of ancillary facility procurement and installation along the whole highway; the unit price is defined as $K_{TC}$. The costs for highway maintenance $C_M$ consist of two parts: daily maintenance, and the operating expenses of the facilities and equipment. The unit prices for these two parts are defined as $K_M$ and $K_O$, respectively. Since hardly any rescue or electrical equipment is required for bridge and road operations, the operating expenses for bridges and roads are negligible compared with those for tunnels.

Note that most unit prices, including $K_{TC}, K_{BC}, K_{PA},$ and $K_O$, are not fixed values, and depend on the state of the design variables in Table 2. Eqs. (9)–(12) show the formulations for the unit prices used in this model. $K_{X,0}$ represents the base price of $K_X$ ($K_X = [K_{TC}, K_{BC}, K_{PA}, K_O]$) when all the design variables $v$ ($v = [Li, Ve, Pa, Sh, Em]$) in Table 2 have a value of 0, while $K_{X,v}$ represents the additional cost compared with $K_{X,0}$ when the design variables $v$ have a non-zero value.

$$K_{TC} = K_{TC,0} + K'_{TC,li} + K'_{TC,ve} + K'_{TC,pa} + K'_{TC,sh} + K'_{TC,em}$$  (9)

$$K_{BC} = K_{BC,0} + K'_{BC,pa} + K'_{BC,sn} + K'_{BC,em}$$  (10)

$$K_{PA} = K_{PA,0} + K'_{PA,pa} + K'_{PA,sn} + K'_{PA,em}$$  (11)

$$K_O = K_{O,0} + K'_{O,li} + K'_{O,ve}$$  (12)

Detailed values for $K_{X,0}$ and $K_{X,v}$ will be specified in the case study.

2.4. Estimation of traffic accident rate

To estimate the crash frequency for a highway segment, traffic accident information from the same region as the construction project must be collected from the traffic administration ministry. With the highway geometrical data, traffic data, and crash data extracted from the traffic accidents report, a multivariate regression analysis can be performed to model the mathematical function between the crash rate and the related risk factors [38,39]. This paper employs a widely used metric called the safety performance function (SPF), which was first proposed in the Highway Safety Manual [40] by the American Association of State Highways and Transportation Officials (ASSHTO) to evaluate the traffic risk level of a proposed alignment. The SPF is a surrogate function [41,42] that uses known information for a road segment to predict its traffic safety in terms of accident frequency; it has been widely used in several highway design tools, such as the Interactive Highway Safety Design Model [43], SafetyAnalyst [44], and Crash Modification Factors Clearinghouse [45]. However, all of these tools are designed for safety assessment in rural/urban highways, and thus assume a driving environment that is significantly different from the environment of a mountainous highway with a large proportion of tunnel segments. Therefore, the SPF $f_{sa}$ is defined by combining all the information from segments located in a mountainous driving environment, as in Eq. (13).

$$f_{sa}(X) = \frac{\sum_i \lambda_i \tilde{e}_i}{\sum_i \lambda_i}$$  (13)

where $L_i$ is the length of segment $i$ (a curve segment or tangent segment); and $\tilde{e}_i$ is the SPF of the $i$th segment, which uses a generalized linear-regression model to establish the mathematical relationship between the explanatory variables and the annual crash frequency [5]:

$$\tilde{e}_i = E(Y_i) = e^{\beta_0 + \sum_{i=1}^p \beta_i X_i}$$  (14)

$C. Li et al. / Engineering 5 (2019) 337–349$
where \( Y_i \) represents the predicted annual crash frequency of segment \( i \), \( X_i = (X_{i1}, X_{i2}, \ldots, X_{in}) \) denotes the explanatory variables of segment \( i \) extracted from the NFTA diagram (see Fig. 1); and \( \tilde{a} = (a_0, a_1, \ldots, a_n) \) denotes the corresponding coefficients, which are described in Table 3.

In general, it is assumed that \( Y_i \) follows a negative binomial distribution [46] with a variance of \( \text{Var}(Y_i) = \tilde{e}_i + \tilde{e}_i^2 \). The fundamental negative binomial regression model for a segment \( i \) is written as follows:

\[
\Pr(Y_i = y_i|\tilde{e}_i, \tilde{a}) = \frac{\Gamma(y_i + \tilde{e}_i)}{\Gamma(y_i)\Gamma(\tilde{e}_i)} \left( \frac{\tilde{e}_i}{\tilde{e}_i + \tilde{a}} \right)^{y_i} \left( \frac{\tilde{a}}{\tilde{e}_i + \tilde{a}} \right)^{(y_i + \tilde{e}_i)}
\]

where \( \Gamma \) is the gamma function and \( \tilde{e}_i \) is the dispersion parameter of the negative binomial distribution.

The regression coefficients \( \tilde{a} \) are estimated using the method of maximum likelihood by the GENMOD procedure in Statistical Analysis System (SAS) software [47]. Note that various statistical tests, including goodness-of-fit test, significance test, and residual analysis, should be adopted before the employment of the SFF model in practical applications [48].

3. Cost–safety optimization and balance for highway planning and design

3.1. Acceptable level of risk examination

Let \( R^n \) be the decision space defined on an \( n \)-dimension Euclidean vector space. Since each candidate highway planning and design set can be represented by \( \tilde{E} = (C, T) \), it can be converted into a chromosome \( \tilde{E} = (P, L, G, V) \). \( P = (P_1, \ldots, P_k) \) contains the 3D coordinates and curve radius information for \( k \) intersection points. \( L = (L_{V1}, \ldots, L_{Vk+1}) \) contains the local variables information for the \( k \) curve segments and the \( k+1 \) tangent segments. \( GV \) includes the global variables information for all the segments.

A bi-objective optimization problem for the highway planning and design in terms of cost and safety can be formulated as follows:

\[
\min_{E \in R^n} f_o(E), f_s(E)
\]

To find the best candidates for the decision-maker, the Non-Dominated Sorting Genetic Algorithm II (NSGA-II) is employed [49]. Every individual in the population requires two entities in order to be computed: non-domination rank and crowding distance. Based on a comparison of these two entities between different solutions, the algorithm selects excellent individuals to form the new parent population, and generates a new offspring population through the basic operation (crossover operator, mutation operator, insert operator, and straight operator [13,14,24]) of the genetic algorithm. The Pareto front is constantly optimized by successive generations until the convergence condition is satisfied, as shown in Fig. 4.

All the solutions on the final Pareto front can be regarded as optimal solutions, considering that there are no other solutions with a better simultaneous safety and cost performance. However, unlike other objectives such as environmental impact [50] and travel time [17], driver and passenger safety is a complex but vital problem that cannot be evaluated based solely on the cost–benefit analysis. All these solutions must therefore be further examined by a risk assessment framework to determine whether they fall within the range of acceptable risk.

To determine and limit the risk level, the Dutch Ministry of Housing, Spatial Planning, and Environment (VROM) proposed a framework for risk evaluation [51,52], which was further improved by later scholars [53–56]. Acceptable risk has two implications, according to the perspective of the decision-makers. The first implication is the personally acceptable level of risk, which is defined as “the frequency at which an individual may be expected to sustain a given level of harm from the realization of a specific hazard” [55]; this can be calculated in the following way [54]:

\[
p_f = \frac{\tilde{a}_i}{10^5p_{df}}
\]

where \( p_f \) denotes the personally acceptable level of risk and \( p_{df} \) is the death rate in a traffic accident, which is typically around 4.4–10.0% [57–59]; here, we assume it to be a conservative 10%. \( \tilde{a}_i \) is a policy factor that reflects the public attitude toward the acceptable risk level for activity \( i \), and varies from 0.01 (in the case of the most prudent attitude and the strongest tendency for risk-aversion strategies, such as in liquefied petroleum gas (LPG)-station construction) to 100 (in a case with complete freedom of choice, such as in mountaineering); \( \tilde{a}_i \) should be 1.0 in the general case of road safety [55]. From a personal perspective, the probability of a severe traffic accident involving fatalities should meet the following requirement:

\[
p_f < p_f
\]

### Table 3

**Description of significant explanatory variables [3].**

<table>
<thead>
<tr>
<th>Categories</th>
<th>Variable ( X_i )</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic situation</td>
<td>AADT (veh d(^{-1}))</td>
<td>Annual average daily traffic</td>
</tr>
<tr>
<td></td>
<td>( Li )</td>
<td>Lighting design in tunnel</td>
</tr>
<tr>
<td></td>
<td>( L_{tu} ) (km)</td>
<td>Tunnel length</td>
</tr>
<tr>
<td></td>
<td>( L_{tm} ) (km)</td>
<td>Tangent length</td>
</tr>
<tr>
<td></td>
<td>( V_0 )</td>
<td>Ventilation design in tunnel</td>
</tr>
<tr>
<td>Vehicle-operating problems</td>
<td>( \Delta V_{ij} ) (km h(^{-1}))</td>
<td>Operating speed consistency on single element, ( \Delta V_{ij} = V_{ij} - V_{ai} ) where ( V_{ij} ) is the operating speed and ( V_{ai} ) is the design speed in the observed segment ( i )</td>
</tr>
<tr>
<td></td>
<td>( \Delta V_{k} ) (km h(^{-1}))</td>
<td>Operating speed consistency on successive elements, ( \Delta V_{k} = V_{k} - V_{ai-1} )</td>
</tr>
<tr>
<td></td>
<td>( \Delta f_t )</td>
<td>Vehicle stability, the difference between the skid resistance assumed ( f_{ai} ) and demanded ( f_{ai} ) where ( C = 1 ) if pavement surface material is concrete, ( C = 0.5 ) if pavement surface material is asphalt</td>
</tr>
<tr>
<td></td>
<td>( L_i ) (km)</td>
<td>Large slope (slope grade ( \geq 2% )), ( L_c = L_c ), interaction between slope grade ( G % ), and length of slope segment ( l ) (km)</td>
</tr>
<tr>
<td></td>
<td>( S_a ) (m)</td>
<td>Stopping sight distance, ( S_a = 0.278V_0\tau + 0.039V_0^2C/a ), where ( \tau ) is the brake reaction time that the AASHO assumes to be 2.5 s in open roads and 3.0 s in tunnels, and ( C ) is the deceleration rate (3.4 m s(^{-2}))</td>
</tr>
<tr>
<td>Weak fault tolerance</td>
<td>( Sh )</td>
<td>Road shoulder width</td>
</tr>
<tr>
<td></td>
<td>( Em )</td>
<td>Distance between emergency stopping strips</td>
</tr>
</tbody>
</table>

\( V_0 = 135.49 \times 10^{-3} \times L_{tu} = 1.29C_1 \times 14.4277\pi - 0.08383C_1 \), where \( Tu(Br) \) is the binary variable, which is equal to 1 if the segment is on a tunnel (bridge); \( f_{ai} = 0.22 \times 1.79 \times 10^{-5}V_0 + 0.06 \times 10^{-5}V_0^2 \), \( f_{ai} = V_{ai}/(1278C_1) - C \), where \( e \) is the supererelevation.
where $P_i$ is the probability of an individual suffering an accident on the highway, and its calculation is based on the predicted accident frequency $f_{sa}$ (acc km$^{-1}$ a$^{-1}$) for this section of the highway:

$$P_i = \frac{f_{eq} \times V_d}{365 \times \text{AADT} \times L}$$

(19)

where $V_d$ is the design speed, $L$ is the total length of the road, and AT is the annual traffic. Given that drivers (i.e., users) only spend a portion of each day ($L/24V_d$) on this section of the highway, the equivalent annual accident frequency $f_{eq}$ should be increasing the $f_{sa}$ by this portion.

In addition to personally acceptable risk, a socially (nationally) acceptable level of risk is an important consideration that reflects the risk evaluation from a national perspective. To meet the requirements of this aspect, the following conditions should be implemented [55]:

$$E(N_{di}) + R_a(N_{di}) < 7 \times 10^{-6} \hat{a}_i P_N$$

(20)

$$E(N_{di}) = N_{di} f_{sa} P_{df}$$

(21)

$$R_a(N_{di}) = 3P_{df} \sqrt{N_{di} f_{sa} (1 - f_{sa})}$$

(22)

where $E(N_{di})$ represents the expected number of fatalities in traffic accidents in the country, $R_a(N_{di})$ is the impact of risk aversion by society toward rare accidents with large fatalities, $P_N$ is the national population size, and $N_{di}$ is the total length of highway in the country (which is about $1.31 \times 10^5$ km in China).

If both Eqs. (18) and (20) are satisfied for a proposed solution's safety performance, that solution is perceived as falling within the range of acceptable risk. In practice, it is only necessary to compare the predicted risk level of the proposed solution with the lower values of Eqs. (18) and (20). All non-dominated solutions should be examined by the risk assessment program in order to rule out solutions with unacceptable risk in terms of traffic safety (see Fig. 5).

### 3.2. Marginal efficiency analysis of safety improvement

Given the residual solutions, decision-makers are recommended to adopt one according to their preference. This paper provides three commonly used decision-making perspectives. The first scenario is to choose a solution with the highest traffic safety performance as the optimal solution; this decision-making method is...

![Fig. 4. Final Pareto front generation process.](image)

![Fig. 5. Risk assessment for decision-making in highway alignments.](image)
Marginal efficiency analysis is provided in the following section. Given the residual solutions within the budget, an algorithm for selecting a cost-efficient solution based on a marginal efficiency analysis is provided in the following section. Marginal efficiency refers to the increase in safety performance that results from extra investment. Considering the law of diminishing marginal utility [60], which is depicted in Fig. 6, an efficient point exists in the Pareto front that satisfies the requirement of minimal marginal efficiency.

Suppose a fitting curve for the Pareto front is represented by \( y = h(x) \). The approximate marginal efficiency \( M_e \) of each point on the Pareto front can be calculated as shown in Eq. (23). Since the Pareto front can be regarded as a decision for safety improvement at the expense of extra investment, a decision criterion to select the cost-efficient solution from the Pareto set is defined as follows:

\[
M_e = \frac{dy}{dx} = |h'(x)|
\]

(23)

It can be seen from Fig. 6 that the marginal efficiency of the safety improvement is remarkable at the lower cost, but decreases as the cost investment increases, and gradually grows to 0. More specifically, even a substantial increase in investment will only minimally improve safety when the cost performance exceeds a certain threshold \( M_e \geq \tau \). Point \( P^* \), which completes \( M_e = \tau \), is seen as a turning point for the cost–benefit balance, and is also regarded as the cost-efficient solution for safety improvement.

Since the fitting curve only reflects the overall shape, rather than the local distribution of the Pareto set, a two-stage filtering algorithm is applied for the marginal efficiency analysis.

**Step 1:** A curve-fitting method is employed to exhibit the Pareto front through a Matlab tool, and then the fitting curve is transformed into a marginal efficiency curve through derivatives, as in Eq. (23). A predefined threshold \( \tau \) (which is set at 0.5 in this paper) is used to find the local region in which the best solution lies. The local region is defined as the five points \( P_{i,j} \) that are closest to \( P^* \) among the non-dominated solutions, where \( P_{i,j} = (f_{coj}, f_{sai}) \) is subject to \( M_e(f_{coi}) = \tau \).

**Step 2:** The five points \( P_{i,j} \) are used as a sequence to calculate their marginal efficiency \( M_{e2i,j} = (f_{sai,j} - f_{sai}) / (f_{coj} - f_{coi}) \) \((i = 1, 2, 3, 4, 5)\).

Every improvement \( P_{i,j} \rightarrow P_{i,j+1} \) that satisfies \( M_{e2i,j} \geq \tau \) is considered worthwhile. Thus, the cost investment is increased step by step until \( M_{e2i,j} < \tau \). Finally, the last improved point \( P_{i,j} \) is chosen as the cost-efficient solution for decision-making.

### 4. Case study

The proposed method was applied to the highway alignment design across Damala Mountain, one of the most difficult sections of the northern route of the China Sichuan–Tibet Highway. The start and end points of this section are Qamdo (97°10’18” E, 31°08’34” N, 3283.000) and Tuoba Township (97°31’33” E, 31°17’07” N, 4000.000). As shown in Fig. 7, an existing country road (G317) connected these two locations through the valley. In order to reduce the travel time and enlarge the traffic capacity, a new highway alignment is being designed for this area. Four planning and design sets have been proposed by the designers, which are displayed in Fig. 7. Corridor A is excluded because it is basically the same as the existing road and does not satisfy the constraint of maximum length (≤ 80 km). The remaining three corridors (B, C, and D) proposed by designers represent the initial population and input into the NSGA-II algorithm.

#### 4.1. Data preparation

Three types of data were required in this study: terrain elevation data, crash data from the study area, and the cost and other model-related parameters.

First, a digital elevation model (DEM) of the study area was developed using Global Mapper to transform the terrain data of...
the study area into point-cloud data, which stores the elevation data of every planar point. Next, the point-cloud data were input to the Matlab software in order to reconstruct a DEM (Fig. 8).

Second, crash data (including road and traffic-related factors) of the study area were collected from the local highway administration ministry. A total of 1375 crashes on this highway section were recorded from 2014 to 2017. The regression coefficients were estimated by means of the maximum likelihood method, and the results are shown in Table 4.

Third, the cost and other model-related parameters were collected, and are specified in Table 5.

### 4.2. Analysis of the results

The combination of DEM in the study area and the initial population constitute the input data, which is the basic required information to start the entire program. As the optimization procedure runs for hundreds of generations, the NSGA-II explores a variety of possibilities for the values of the design variables. Excellent individuals are preserved and selected as the parent alignments to generate new populations through genetic algorithm operators, while poor individuals are eliminated. Fig. 9 shows the searched positions of the intersection points by the NSGA-II. The surrounding space along the local optimal corridors has been carefully exploited in both the horizontal and vertical alignments.

The safety performance and cost performance are gradually improving, which demonstrates the great potential of the NSGA-II for performance optimization. Fig. 10 shows the process of safety improvement. Safety-related indicators, such as $ΔV_d$, $ΔV_o$, $Δf$, and $S_o$, have been gradually improved with the iteration process, resulting in a remarkable decrease in expected traffic accidents. The global variables, such as $Li$, $Ve$, $Sh$, and $Em$, converge earlier than other local variables, which proves that these safety-improvement measures are necessary. Fig. 11 presents the improvement in the expected annual costs of the generated alignments, as well as the improvement details of their components.

#### Table 5

<table>
<thead>
<tr>
<th>Cost type</th>
<th>Cost items</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction cost</td>
<td>Ground cut (RMB m⁻³)</td>
<td>$K_{GC} = 54$</td>
</tr>
<tr>
<td></td>
<td>Ground-fill (RMB m⁻³)</td>
<td>$K_{GF} = 34$</td>
</tr>
<tr>
<td></td>
<td>Base price of pavement (RMB m⁻¹)</td>
<td>$K_{PA,o} = 8750$</td>
</tr>
<tr>
<td></td>
<td>Base price of bridge (RMB m⁻¹)</td>
<td>$K_{BC,o}(H ≤ 20 m) = 1.6 \times 10^5$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$K_{BC,o}(20 m &lt; H ≤ 40 m) = 2.0 \times 10^5$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$K_{BC,o}(40 m &lt; H ≤ 60 m) = 2.5 \times 10^5$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$K_{BC,o}(60 m &lt; H ≤ 80 m) = 3.0 \times 10^5$</td>
</tr>
<tr>
<td></td>
<td>Base price of tunnel (RMB m⁻¹)</td>
<td>$K_{TC,c} = 5.6 \times 10^3$</td>
</tr>
<tr>
<td>Highway apportionment (RMB m⁻¹)</td>
<td>$K_{HR} = 7300$</td>
<td></td>
</tr>
<tr>
<td>Additional cost compared to $K_{X,0}$ (RMB m⁻¹)</td>
<td>$K_{IC,Δi,j} = 900$</td>
<td></td>
</tr>
<tr>
<td>Maintenance cost</td>
<td>Maintenance cost (RMB m⁻³a⁻¹)</td>
<td>$K_{0,0} = 2000$</td>
</tr>
<tr>
<td></td>
<td>Base cost of tunnel operation (RMB m⁻³a⁻¹)</td>
<td>$K_{0,0} = 2000$</td>
</tr>
<tr>
<td></td>
<td>Additional cost compared to $K_{0,0}$ (RMB m⁻³a⁻¹)</td>
<td>$K_{G,Δi,j} = 200$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$K_{G,Δi,j} = 3.02 \times 10^3$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$K_{G,Ve,j} = 108$</td>
</tr>
</tbody>
</table>

$H$: bridge pier's height.

Fig. 8. Digital elevation model of the study area.
Fig. 9. Searching space for highway alignment optimization.

Fig. 10. The transition of the safety performance during the search. (a) Safety performance; Explanatory variables including (b) $\Delta V_o$, (c) $\Delta V_d$, (d) $\Delta V_v$, (e) $S_{d}$, (f) $L_{s}$, (g) $L_{u}$, (h) $L_{c}$, (i) $V_{e}$, (j) $S_{h}$, (k) $S_{l}$, (l) $E_{m}$.

Fig. 11. The transition of the cost performance during the search. (a) Cost performance; (b) construction cost of tunnel; (c) construction cost of bridge; (d) maintenance cost; (e) construction cost of road.
including construction costs for various highway structures and maintenance costs. Note that Figs. 10 and 11 only show the average value of part of the indicators from a certain generation.

In the search, it was gradually discovered that corridor B has the greatest potential for optimization of the three preliminary corridors (B, C, and D) since it has the minimum tunnel length, which results in a slump in cost. Fig. 12 shows the performance results of the optimization generations of the search program. The traffic safety and construction cost are optimized simultaneously in the iteration. The program terminated at gen#729, when the convergent condition was reached. The personally acceptable safety performance \( f_{psa} = 52.731 \text{ acc.} \cdot \text{km}^{-1} \cdot \text{a}^{-1} \) and the nationally acceptable safety performance \( f_{nsa} = 0.744 \text{ acc.} \cdot \text{km}^{-1} \cdot \text{a}^{-1} \) were calculated according to Eqs. (19) and (20). It was found that all the non-dominated solutions at gen#729 fell within the personally and nationally acceptable risk levels, so no solutions were eliminated because they exceeded the safety threshold.

There are three optional optimal solutions, as shown in Fig. 12: Optimal solution 1 has the best safety performance without considering a budget limitation, optimal solution 2 has the best safety performance considering a budget limitation of 700 million RMB, and optimal solution 3 is the cost-efficient solution within the budget limitation. Information on these three solutions is provided in Fig. 13; this information includes feasible and commonly used recommendations for decision-makers. Decision-makers are encouraged to adopt one of these solutions according to the practical situation and their preference.

5. Discussion and conclusion

The optimization results show that the proposed model is capable of generating high-quality solutions, and can achieve low costs and low accident rates. However, it was found that the nationally acceptable safety performance was far lower than the

---

**Fig. 12.** Performance results of optimization generations.

**Fig. 13.** Computer-generated optimal solution with respect to cost and safety.
personally acceptable safety performance. This is because, for individuals, the possibility of being killed on this highway is relatively low—even though many traffic accidents happen, most are property damage only (PDO) or crashes that cause injuries [31]; thus, the estimated number of fatalities is not particularly high when compared with aircraft accidents [55]. However, for society (i.e., the nation), the highway death rate (fatalities/population), which does not take the total highway length of a country into consideration, is an important index for evaluating traffic safety [61,62]. Considering that the total length of all Chinese highways is very large, the permissible accident rate per kilometer must be seriously limited in order to lower the total death rate on all highways at the national level. It should be noted that the nationally acceptable risk levels are different for each country, and the personally acceptable risk levels of different countries (i.e., different death rates in cases of traffic accidents) are also different. Therefore, both acceptable risk levels must be computed, since it is not certain which one is lower (dominating).

Throughout the experiment, the optimization process and the scope of the searching space were displayed. In the NSGA-II, 143,314 offspring were generated during the search. The mean annual cost value of the final generation was 610.08 million RMB, while the mean cost value of the first generation was 821.17 million RMB, which is 1.35 times higher. Similarly, the mean safety performance value of the final generation was 5.84, while the mean value of the first generation was 9.74, which is 1.67 times higher. Fig. 10 shows that the safety optimization of the planning and design was mainly achieved by reducing the length of tunnels and steep slopes and improving the design consistency. The quick convergence of the global variables may indicate that the improvements in the lighting, ventilation, road width, and emergency stopping strips design are valuable, since they significantly improve operation safety at a low cost. Fig. 11 shows that cost optimization was mainly achieved by replacing the tunnel structures with road structures.

This paper presents a novel framework that strikes a balance between cost and safety for mountainous highway alignment design. The contribution of this paper to the highway alignment problem is as follows: ① An FTA was combined with regression analysis methods to model the relations between traffic accidents and highway design variables. An NFTA diagram was presented in order to identify the risk factors affecting both operation safety and life-cycle cost among the design variables. Next, a generalized linear-regression model was introduced in order to evaluate the safety performance of a highway alignment. ② The highway planning and design variables were arrayed in a new form of chromosome genes, which formed the basis of the cost and safety analysis. ③ The relations between the life-cycle cost of the highway planning and design and the states of the risk factors were developed in the cost-assessment functions. The impact of risk factors on the life-cycle cost was modeled by Eqs. (9)–(12). ④ The theory of acceptable risk was introduced to discuss individual and social attitudes toward highway safety in the proposed alignments. ⑤ A marginal efficiency-based computing method was proposed in order to solve for the cost-efficient solution among the candidate optimal solutions.

Despite all the effort described above, there is still room for future improvement. Since the SPF model was trained using crash data collected from the same region as the engineering project, it is suggested that other scholars not apply this model directly to other regional projects. The neglect of spatial heterogeneity might lead to biased parameters and model misspecification. Spatially varying coefficients models can be investigated in order to account for unstructured and spatially structured heterogeneity in future research.

In conclusion, this paper identifies the relations between life-cycle cost and traffic safety in mountainous highway construction projects, and strikes an optimal balance between them within acceptable risk levels. The proposed method can be applied to a reasonable investment analysis and multi-objective optimization in other high-safety-risk engineering projects.

Acknowledgements

The authors would like to acknowledge Qi Fang, Jian Guo, Xiaomin Wu for their help. We are also grateful for the support of the National Natural Science Foundation of China (71732001, 51875311, and 51678265) and the Research Project of the Chinese Academy of Engineering (2017-XZ-12).

Compliance with ethics guidelines

Chengqian Li, Liefyun Ding, and Botao Zhong declare that they have no conflict of interest or financial conflicts to disclose.

References
