Received: 2 March 2023 | Accepted: 12 July 2023

DOI: 10.1111/mice.13073

RESEARCH ARTICLE

COMPUTER-AIDED CIVIL AND INFRASTRUCTURE ENGINEERING WILEY

# Mountain railway alignment optimization based on landform recognition and presetting of dominating structures

Xinjie Wan[1,2] | Hao Pu[1,2] | Paul Schonfeld[3] | Taoran Song[1,2] | Wei Li[1,2] | Lihui Peng[4] | Jianping Hu[5] | Ming Zhang[6]

[1]School of Civil Engineering, Central South University, Changsha, China

[2]National Engineering Research Center of High-Speed Railway Construction Technology, Changsha, China

[3]Department of Civil and Environmental Engineering, University of Maryland, College Park, Maryland, USA

[4]China Railway Siyuan Survey and Design Group Co. Ltd., Wuhan, China

[5]China Railway Eryuan Engineering Group Co. Ltd., Chengdu, China

[6]China Railway First Survey and Design Institute Group Co. Ltd., Xi'an, China

**Correspondence**

Hao Pu, School of Civil Engineering, Central South University, Changsha 410075, China.
Email: haopu@csu.edu.cn

Paul Schonfeld, Department of Civil and Environmental Engineering, University of Maryland, College Park, MD 20742, USA.
Email: pschon@umd.edu

**Funding information**

National Key R&D Program of China, Grant/Award Number: 2021YFB2600403; National Science Foundation of China, Grant/Award Number: 52078497; Science and Technology Research and Development Program Project of China railway group limited, Grant/Award Number: 2022-Major-20

## Abstract

Mountain railway alignment optimization has always been a challenge for designers and researchers in this field. It is extremely difficult for existing methods that optimize alignments before major structures to generate a better alignment than the best one provided by human designers when the terrain is drastically undulating between the start and endpoints. To fill this gap, a "structures before alignments" design process is proposed in this paper. Primarily, a landform recognition method is devised for recognizing dominating landforms. Then, a bi-level alignment optimization model is proposed, with the upper level dedicated to characterizing dominating structures and the lower level focusing on optimizing the entire alignments. To solve this bi-level model, a three-stage optimization method is designed. At the first stage, a scanning process and screening operators are devised for generating all the possible locations of dominating structures. At the second stage, a hierarchical multi-criteria decision-making procedure is applied for selecting the optimized dominating structure layouts. At the third stage, alignments are optimized based on the determined structure layouts using a bi-objective optimization method, which minimizes construction cost and geo-hazard risk simultaneously. The proposed model and solution method are applied to two real-world cases whose results verify their capabilities in producing alignment alternatives with better combinations of construction cost and geo-hazard risk than manually designed alternatives.

# 1 | INTRODUCTION

## 1.1 | Problem statement

Railway alignment development is a complex problem, which significantly influences the construction cost and operation safety of a railway project (Zhang et al., 2022). The main tasks of alignment determination are to find the optimized railway locations (Kang et al., 2009), geometric configurations (Gao et al., 2022), and structural components (Kim et al., 2007) between two given endpoints in a landform (Jha & Schonfeld, 2004). This work is complicated because there are theoretically infinite number of possible alternatives (Song , Pu, Schonfeld, & Hu, 2022) that should consider multiple constraints (Li et al., 2016) and hard-to-evaluate objectives (Karlson et al., 2016). Particularly, it is quite challenging to find satisfactory structure layouts for alignments traversing drastically undulating terrain (Song, Pu, Schonfeld, Liang, et al., 2022). Circuitous alignments as well as high bridges and long tunnels are needed when there are deep valleys and high mountains between the start and endpoints. These high bridges and long tunnels are designated as dominating structures, while the deep valleys and high mountains are designated as dominating landforms in this study. It is important to find proper layouts for these structures because

1. excessively high bridges and long tunnels are impractical due to limitations of the existing construction technology and
2. inappropriate structure layouts could pose potential risks (including geo-hazard risk and structural instability) to railway construction and operation. The impact of these risks is significantly magnified due to the enormous maintenance complexity of these structures.

In this regard, the determination of dominating structures has a decisive role in the entire alignment design process, and hence should be prioritized. Thus, human designers generally determine the locations of these dominating structures first and then develop alignments that connect the dominating landforms to the start and end points. It usually takes considerable time and effort to compare numerous alignment solutions and find the best one among them. However, with limited resources and time, many potential alternatives are still overlooked during such a time-consuming and laborious process. In the 21st century, with the thriving development of computer technology and geographic information systems (GIS), many researchers have proposed numerous computer-aided approaches for automatically generating optimized alignment alternatives between two given points (Jong et al., 2000; Kang et al., 2012; Song et al., 2023).

## 1.2 | Literature review

In general, computer-aided alignment optimization approaches can be characterized by two components, namely their model formulation and solution methods.

1. An optimization model is used for quantitatively describing a practical problem by formulating objective functions and constraints based on specific design variables. For alignment design, an alignment is typically specified by its points of intersection (PIs) and corresponding curve configurations. Thus, the design variables for alignment optimization model are the coordinates of PIs and parameters of curves. Objective functions are formulated for evaluating the fitness of alignments, which may include multiple factors. The construction cost (Kazuhiro, 2005; Mondal et al., 2015) is a vitally important factor that should be considered in alignment optimization models. In addition, land availability and cost, geologic and ecologic factors, operating costs, and demand are also quantified by various researchers for considering their impacts on alignment fitness. The geologic and ecologic factors can be converted into costs for forming a single-objective optimization model (Jha & Schonfeld, 2004) that also includes the construction cost. They can also be considered separately from the construction cost for building a bi-objective (Hirpa et al., 2016) or a multi-objective (Sadek et al., 1999) optimization model. Regarding the constraints, geometric (Monnet et al., 2022), locational (Lee & Cheng, 2001), and structural (Zhang et al., 2019) constraints are considered in increasingly complex alignment optimization models.
2. The solution methods for alignment optimization can be categorized into three kinds according to their search processes, namely, mathematical programming (Vázquez-Méndez et al., 2018), evolutionary (Jong et al., 2000), and growth-based (Li et al., 2016).

Various mathematical programming methods, such as linear programming (LP) (Easa, 1988; Moreb, 1996), mixed integer linear programming (MILP) (Momo et al., 2023; Monnet et al., 2020; Vázquez-Méndez et al., 2021), and sequential quadratic programming (SQP) (Casal et al., 2017) have been successfully applied for alignment optimization. Some two-stage optimization methods that integrate MILP and SQP (Vázquez-Méndez et al., 2018), or MILP and derivative-free optimization solvers (Mondal et al., 2015) have also been designed for improving the performance of a single optimization method.

For evolutionary optimization methods, initial solutions are generated first. Then, the alignment alternatives are adjusted iteratively based on their fitness, which are determined by their surrounding environment (e.g., topography and geology). In this field, Jong and Schonfeld (2003) first developed a genetic algorithm (GA) for highway alignment optimization. Afterward, a GIS was designed to store basic geographic data, such as environmentally sensitive regions and ground elevations (Jha et al., 2007). The GIS was integrated with a GA for providing environmental information during the alignment search process in order to solve realistic alignment optimization problems. In addition to GAs, other evolutionary optimization methods, such as a particle swarm optimization (PSO; Babapour et al., 2018; Shafahi & Bagherian, 2013) and ant colony algorithms (ACO; Samanta & Jha, 2012; Sushma et al., 2022), have also been customized for optimizing highway and railway alignments.

For growth-based optimization approaches, local optimized paths are produced gradually. Then, all the local optimized paths are accumulated to form a global optimized alignment. A distance transform (DT) is a typical growth-based method and was first introduced to alignment optimization by de Smith (2006). Subsequently, Li et al. (2016) improved it with a bidirectional scanning strategy intended to solve railway alignment optimization problems in constrained mountainous regions. Moreover, a comprehensive geographic information model (CGIM) was devised to consider more detailed geographic data than in a basic GIS. In this model, besides basic topographic data, geologic conditions and forbidden areas are also stored. Thus, necessary environmental information can be conveniently obtained during the alignment search process. Afterward, Pu et al. (2019) further extended the search spaces and developed a three-dimensional DT (3D-DT). In addition to DT-based methods, Dijkstra's algorithm (Pradhan & Mahinthakumar, 2013; Song, Pu, Schonfeld, Liang, et al., 2022), a motion planning algorithm (Sushma & Maji, 2020; Yang et al., 2020) and a k-shortest path algorithm (Pushak et al., 2016), are also growth-based methods, which have been widely used in alignment optimization.

Furthermore, an alignment optimization method utilizing a deep reinforcement learning technique (Gao et al., 2022) has also been proposed for solving alignment optimization problem.

Alignment design is a very complex problem, which can be separated into two subproblems (i.e., determining alignment profiles and structure layouts). These two subproblems are highly interrelated and their relative importance varies with environmental conditions. In a study area with relatively simple topographic and geologic conditions, these two subproblems may be regarded as equally important or the alignment profile determination is prioritized over structure layouts. However, these two subproblems present a strong primary–secondary relation when bridges and tunnels account for a large fraction of an alignment. The primary subproblem is the layout of dominating structures because it basically determines the construction cost and duration for an entire alignment. Therefore, in a manual alignment design process, the primary subproblem is addressed first and other alignment sections are optimized while constrained by the results of the primary subproblem.

These two subproblems are optimized concurrently in all known alignment optimization methods that are automated (such as in Pu et al., 2021, 2023; Song, Pu, Schonfeld, & Hu, 2022). The alignment design in both evolutionary and growth-based approaches uses an "alignments before structures" process. Specifically, an alignment spatial profile is determined first and then the structures are configured according to the elevation differences between the vertical alignment and the ground profile as well as the predefined fill-bridge height and cut-tunnel depth. Afterward, the structural constraints are checked and the alignment fitness is computed. This "alignments before structures" design process is based on a trial-and-error principle. An alignment is regarded as infeasible when its structure layouts cannot satisfy the structural constraints. In a study area with dominating landforms, alignment-feasible regions may be extremely limited and the generated structures would mostly violate the structural constraints. In this condition, it is extremely difficult for this "alignments before structures" design process to find a better dominating structure layout compared with a good manually designed one. Among mathematical programming optimization methods, a branch-and-bound approach (Momo et al., 2023) is widely applied for configuring structures before computing the alignments, but the structures and alignments are still optimized concurrently. In this situation, the influence of nondominating alignment sections on alignment fitness may be exaggerated and the dominating structures may not be properly located in a study area with dominating landforms.

### 1.3 | Focus of this study

To fill this gap, a "structures before alignments" alignment design process is proposed for presetting dominating structure layouts before alignment optimization in study areas with dominating landforms, as shown in Figure 1. The main flow of the proposed solution method is demonstrated in Figure 2 and the main contributions of this work are listed below:

1. A landform recognition method combining alignment spatial reachability analysis and elevation tendency is devised for segmenting alignment-unreachable

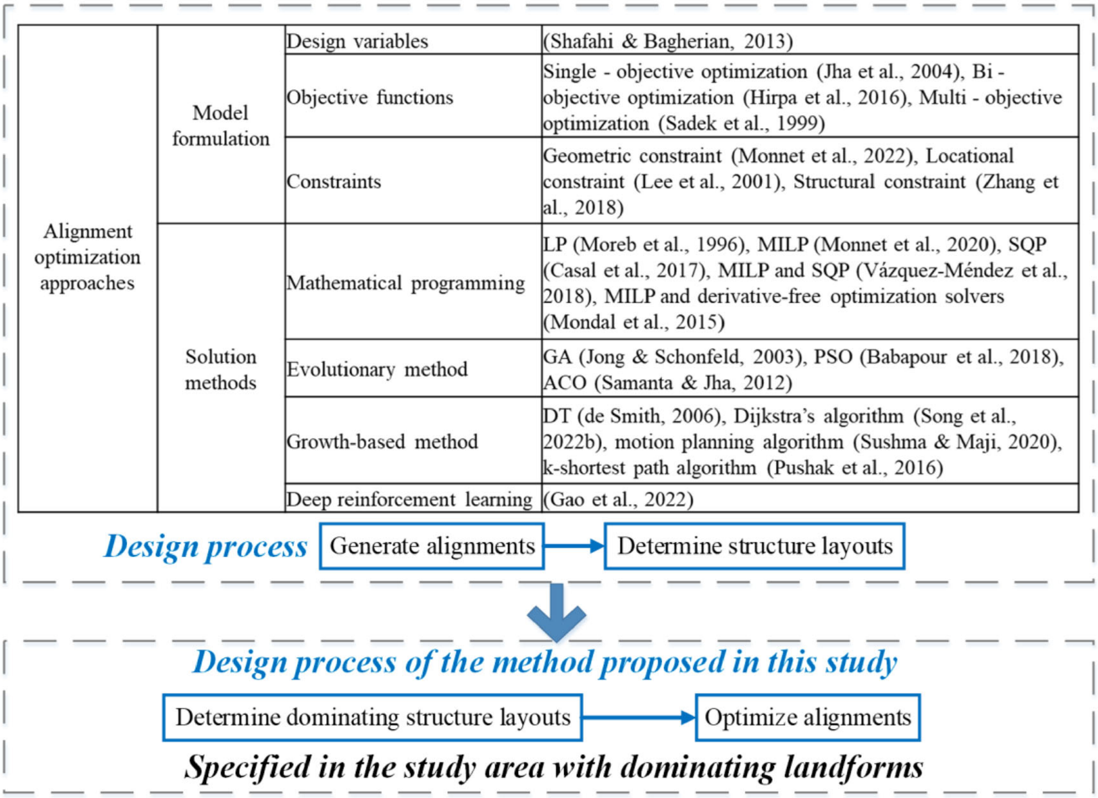



**FIGURE 1** The alignment optimization process of existing methods vs. the proposed method.

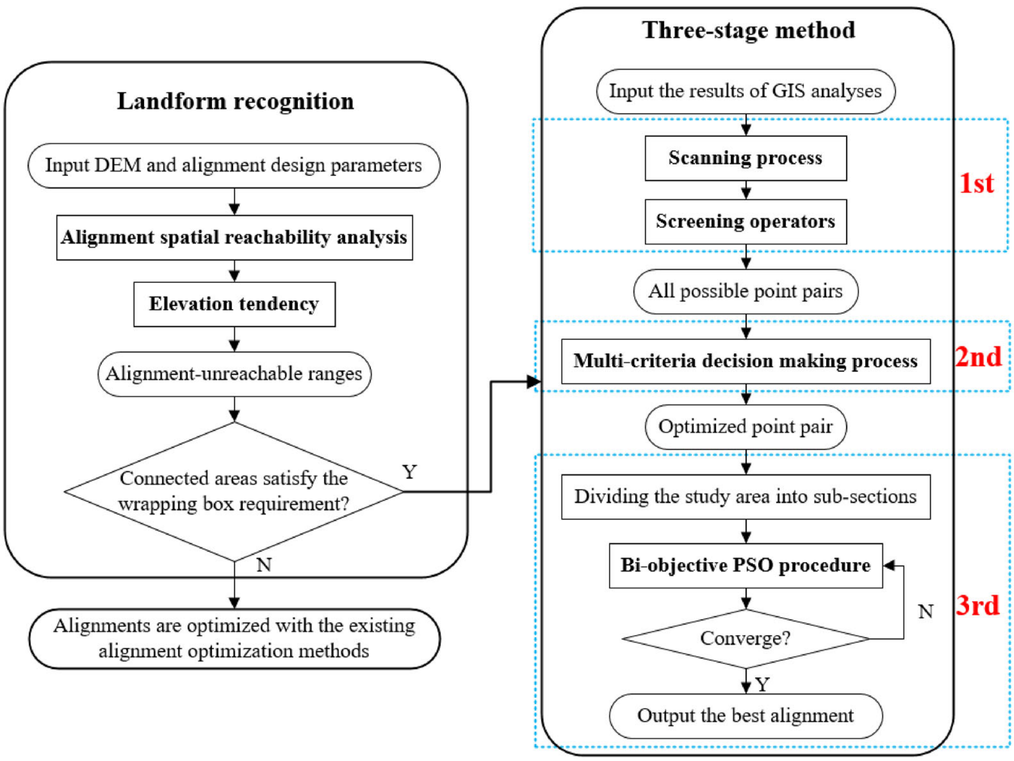



**FIGURE 2** Main flow of the solution method.

regions. A two-pass scanning method is then adopted for identifying the dominating landforms.

2. A dominating structure optimization model is developed for presetting dominating structures. In this model, the structure fitness is evaluated according to seven factors (regarding topology, geology, and structure–alignment relations) and structure-related constraints are imposed. The output of this model is used as the input to the subsequent alignment optimization process for constructing a bi-level optimization model.
3. A three-stage method is proposed for solving this bi-level optimization model. At the first stage, a devised scanning process and specific screening operators are used for generating all the possible endpoints of dominating structures. At the second stage, a hierarchical multi-criteria decision-making (MCDM) process is employed for selecting the optimized dominating structure layouts. At the third stage, a bi-objective PSO, which concurrently minimizes construction cost and geo-hazard risk, is adopted for alignment optimization based on the optimized structure layouts.
4. The devised model and solution method are applied to two realistic cases with drastically undulating terrain. Their effectiveness in finding optimized layouts for dominating structures and generating optimized alignments is verified by the results of case studies. Specifically, they can find an alignment with lower construction cost and geo-hazard risk than the best available manually designed one.

The rest of this paper is organized as follows: Section 2 shows the landform recognition process. Section 3 presents the developed bi-level optimization model. Section 4 demonstrates the procedures of the solution method. Section 5 provides two real-word case studies and Section 6 summarizes this research.

## 2 | LANDFORM RECOGNITION

In practice, advanced terrain analysis is an indispensable step in alignment design. Through this procedure, dominating landforms (which refer to the valleys or mountains that alignments cannot avoid traversing) can be recognized in an alignment optimization process. There are various studies devoted to landform classification in the geomorphology field (Jasiewicz & Stepinski, 2013). Different entities (e.g., peak, ridge, slope, valley, and flat) are categorized based on the texture similarity of terrains (Matsuura & Aniya, 2012). For alignment optimization problems, alignment characteristics should be integrated into landform classification through which only valleys and mountains unreachable by alignments need to be recognized. Thus, two GIS procedures, namely alignment spatial reachability analysis and elevation tendency analysis, are implemented to segment the alignment-unreachable ranges. Afterward, a two-pass scan method is employed to identify the dominating landforms. To record the landforms' characteristics, the study area is discretized as a series of grids based on the resolution of digital elevation model (DEM).

### 2.1 | Segmentation of alignment-unreachable valleys and mountains

#### 2.1.1 | Alignment spatial reachability analysis

To connect two given points in topographic space, an alignment should be circuitous to overcome without excessive gradients the elevation difference between two endpoints and to avoid obstacles. The maximum circuity threshold $\gamma_{\max}$ for a railway is determined by the natural gradient ($G_T$) of the study area and the maximum allowable gradient ($G_{\max}$). $\gamma_{\max}$ is formulated as follows (Song et al., 2021):

$$\gamma_{\max} = \max\left\{\frac{G_T}{G_{\max}}, 1\right\} + \sigma \tag{1}$$

where $\sigma$ = an empirical parameter used to measure the additional alignment circuity for gradient losses in curve or tunnel sections and obstacle avoidance.

Based on $G_{\max}$ and $\gamma_{\max}$, the horizontally ($\Omega$) and vertically ($\Phi$) reachable ranges for a railway can be determined (Pu et al., 2023). Specifically, all the grids that satisfy the $G_{\max}$ and $\gamma_{\max}$ constraints constitute the alignment-reachable ranges:

$$\begin{aligned}
&\Omega = \frac{\sqrt{(r-r_S)^2+(c-c_S)^2}+\sqrt{(r-r_E)^2+(c-c_E)^2}}{\sqrt{(r_S-r_E)^2+(c_S-c_E)^2}} < \gamma_{\max} \\
&\Phi = (H_S - G_{\max}\cdot L_S\cdot\gamma_{\max}, H_S + G_{\max}\cdot L_S\cdot\gamma_{\max}) \\
&\cap (H_E - G_{\max}\cdot L_E\cdot\gamma_{\max}, H_E + G_{\max}\cdot L_E\cdot\gamma_{\max})
\end{aligned} \tag{2}$$

where $r$, $c$, $r_S$, $c_S$, $r_E$, $c_E$ = the row and column numbers of a grid in an alignment-reachable range as well as the start and end points, respectively; $L_S$, $L_E$ = the distance between this grid to the start and end points, respectively. $H_S$, $H_E$ = the elevations of the start and endpoints, respectively.

Geometrically, the horizontally reachable range for a railway alignment is an elliptical region with the start and end points as the focal points and $1/\gamma_{\max}$ as the eccentricity. The vertically reachable range is a rhombic region with $G_{\max}$ as the gradient. The regions within the horizontally reachable ranges whose elevations are below

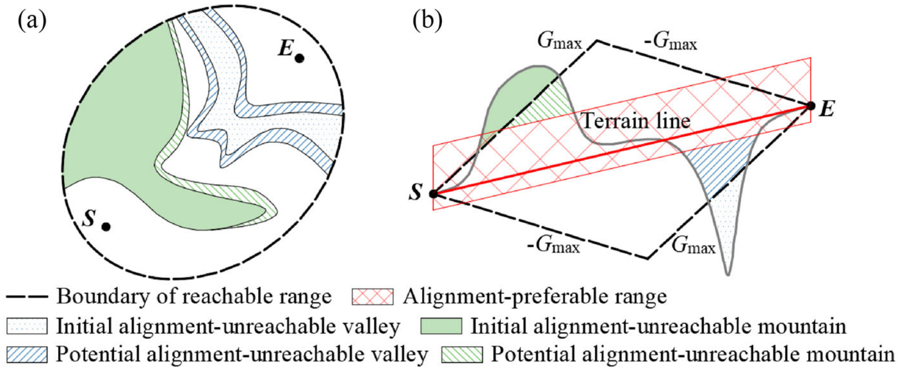



**FIGURE 3** (a) Horizontal and (b) vertical views of alignment-unreachable valleys and mountains.

or above the vertically reachable range are deemed as the initial alignment-unreachable valleys or mountains. These regions can only be traversed by an alignment through bridges or tunnels.

### 2.1.2 | Elevation tendency

In complicated mountainous regions, the structure layouts vary with different vertical alignments. Although their geometric shapes are dissimilar, their common aim is to overcome the elevation difference between two endpoints. Thus, the vertical alignment is expected to be closer to the straight line between the two endpoints. In this regard, the sections with elevations ranging between 100 m below or 500 m above this vertical straight line are regarded as the preferable range for an alignment, while other sections in the vertical reachable region are defined as potential alignment-unreachable valleys or mountains (Code for Design of Railway Line, 2017). Alignments tend to traverse these potential alignment-unreachable valleys or mountains through bridges or tunnels. Therefore, they are combined with initial alignment-unreachable ranges to jointly constitute the alignment-unreachable valleys and mountains, with details provided in Figure 3.

## 2.2 | Recognition of dominating landforms

Alignment-unreachable valleys and mountains are generally scattered in complicated mountainous regions. Among them, only the regions that entirely separate two endpoints are defined as dominating landforms in this paper. To recognize these dominating landforms, a two-pass scan method (He et al., 2008) used for classifying different connected areas is adopted. Different connected areas can be distinguished after two scanning processes are conducted with an 8-connectedness mask (Wu et al., 2009). The steps customized for identifying dominating landforms are listed below:

Step 1. Initializing: The grids in alignment-unreachable regions are defined as foreground grids and others are designated as background grids. A provisional label is specified as 1.

Step 2: First scan: scans from the top left to the right bottom:

1. The provisional label is assigned to the current grid if the grids in its 8-connectedness mask (which is shown in Figure 4) are all background grids. The provisional label's value is increased by 1.
2. When there are foreground grids in the mask of the current grid, the minimum provisional label of these foreground grids is assigned to the current grid.
3. All provisional labels assigned to the same connected regions are defined as equivalent labels. Each group of equivalent labels is recorded in an equivalence sequence, as shown in Figure 4a.

Step 3: Determining representative labels: The minimum label in an equivalence sequence is designated as the representative label in this connected region.

Step 4: Second scan: proceeds in the same order as the first scan and relabels each foreground grid by the representative label, as shown in Figure 4b.

After different connected areas are labeled, a rectangular wrapping box for each area can be generated as illustrated in Figure 5. It can be noted that when at least one pair of diagonal points of the wrapping box is outside the railway's horizontally reachable range (which is defined as the wrapping box requirement in this paper), this connected area is unavoidable for traversing alignment, and hence, is designated as a dominating landform.

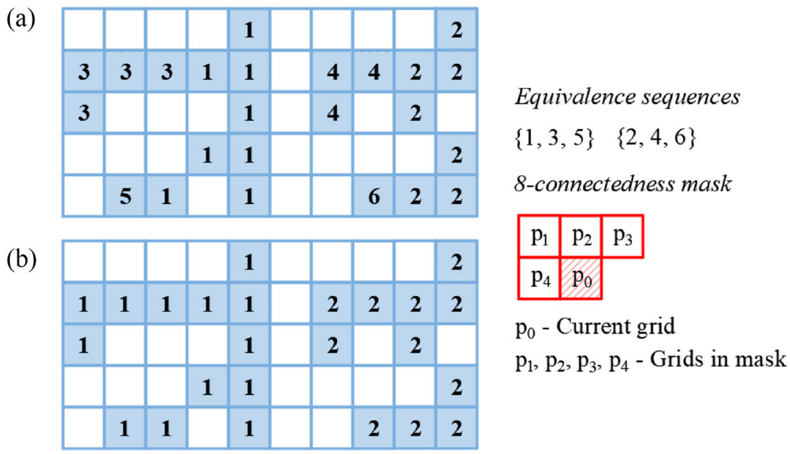



**FIGURE 4** The outcomes of (a) first scan and (b) second scan.

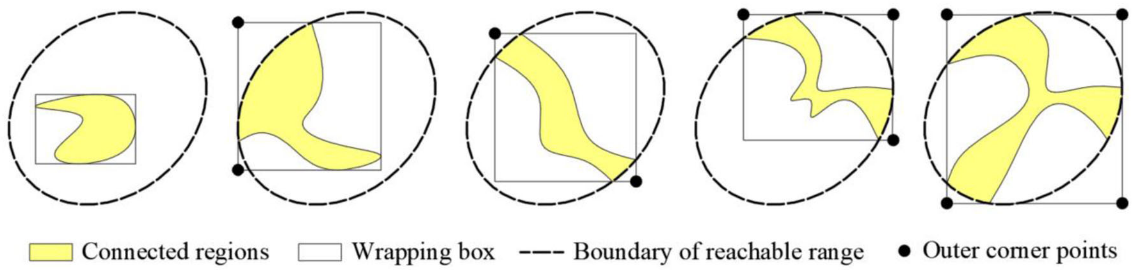



**FIGURE 5** Distribution of wrapping boxes and diagonal points for different connected areas.

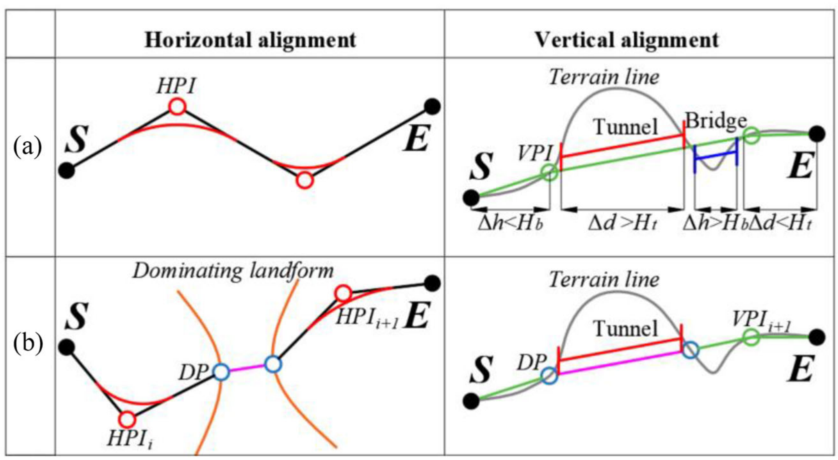



**FIGURE 6** (a) "Alignments before structures" design process in the existing methods, where $H_b$ is the threshold height of fill and bridge, and $H_t$ is the threshold depth of cut and tunnel and (b) "structures before alignments" design process in this paper.

## 3 | BI-LEVEL OPTIMIZATION MODEL

In previous studies, the layouts of alignment structures are determined by the geometric shape of an alignment (which is determined by horizontal points of intersection [HPIs], vertical points of intersection [VPIs]), and the elevation differences ($\Delta h$ for fill height and $\Delta d$ for cut deep) between design elevations and corresponding ground profile (Figure 6a). However, this "alignments before structures" design process is not applicable in complex mountainous regions with dominating landforms. In these areas, the layouts of dominating structures should be prioritized and used to determine the locations of an entire alignment. To reduce the construction cost and difficulty, these dominating structures should be as short and straight as possible. In this regard, they are represented by their two endpoints, which are specified as dominating points (DPs) in this study, as shown in Figure 6b.

After the dominating structures are preconfigured, the study area is divided into several subsections (i.e., the sections between and beyond the structures' DPs). Thus, a bi-level optimization model is developed in this study for

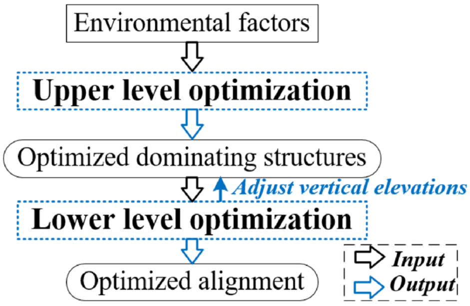



**FIGURE 7** Flow chart of the bi-level alignment optimization model.

presetting dominating structures and generating a complete alignment accordingly. At the upper level, the layouts of dominating structures are optimized while considering various topographic and geologic factors as well as alignment–structure relations (which represent the coupling degrees between the alignment and structures). At the lower level, the complete alignments are optimized by using the construction cost and geo-hazard risk as objective functions. The flow chart of this bi-level alignment optimization model is provided in Figure 7.

## 3.1 | Upper level—Dominating structures optimization

### 3.1.1 | Design variables

For a dominating structure, DPs can be described by the spatial coordinates of its abutments ($X_{A\text{-}S}$, $Y_{A\text{-}S}$, $Z_{A\text{-}S}$), ($X_{A\text{-}E}$, $Y_{A\text{-}E}$, $Z_{A\text{-}E}$) or portals ($X_{P\text{-}S}$, $Y_{P\text{-}S}$, $Z_{P\text{-}S}$), ($X_{P\text{-}E}$, $Y_{P\text{-}E}$, $Z_{P\text{-}E}$). A pair of DPs is defined as a point pair in this work and all structures' point pairs are taken as their design variables ($D_S$). Supposing the number of dominating valleys and mountains are $m$ and $n$, respectively, for a study area, $D_S$ for this case are provided below:

$$\begin{aligned} &\text{design variables}: D_S = (X_S, Y_S, Z_S) \\ &\text{where } X_S = [X_{A-S1}, X_{A-E1}, X_{P-S1}, X_{P-E1}, \\ &\ldots, X_{A-Sm}, X_{A-Em}, X_{P-Sn}, X_{P-En}]^T \\ &Y_S = [Y_{A-S1}, Y_{A-E1}, Y_{P-S1}, Y_{P-E1}, \\ &\ldots, Y_{A-Sm}, Y_{A-Em}, Y_{P-Sn}, Y_{P-En}]^T \\ &Z_S = [Z_{A-S1}, Z_{A-E1}, Z_{P-S1}, Z_{P-E1}, \\ &\ldots, Z_{A-Sm}, Z_{A-Em}, Z_{P-Sn}, Z_{P-En}]^T \end{aligned} \tag{3}$$

### 3.1.2 | Determining influential factors

Multiple factors should be considered for evaluating the fitness of dominating structures ($F_S$). Three types of factors (i.e., topographic factors, geologic factors, and factors related to alignment–structure relations) are considered in this paper, as shown in Figure 8. Their detailed quantitative criteria are presented below.

*Topographic factors*

The location of a dominating structure and its corresponding topographic conditions jointly determine its construction cost and stability. For construction cost, the decisive factors are the length ($L_P$) of a dominating structure and the maximum elevation difference ($H_M$) between this structure and its corresponding ground profile. The construction cost for a bridge or a tunnel equals $L_P$ multiplied by its unit cost ($U_C$). $U_C$ for a bridge or tunnel varies with that bridge's or tunnel's $L_P$ and $H_M$. Their formulations are as follows:

$$U_C = \begin{cases} U_0 & H_M < H_1 \\ U_1 & H_M \geq H_1, L_P < L_1 \\ U_2 & H_M \geq H_1, L_1 < L_P < L_2 \\ \ldots \end{cases} \tag{4}$$

where $H_1$, $L_1$, $L_2$ = the elevation difference and structure length thresholds for different unit costs; $U_0$, $U_1$, $U_2$ = the unit costs for a structure with different $H_M$ and $L_P$ ($U_0 < U_1 < U_2$).

It can be found that the construction cost increases with the increase of $L_P$ and $H_M$. Hence, these two criteria should be reduced to decrease the construction cost of bridges or tunnels.

Regarding the structure stability, the gradient ($G_S$) of an abutment or a portal should decrease to enhance the stability of corresponding component, and thus, to further ensure the stability of the entire bridge or tunnel. Moreover, to reduce the pressure bias and guarantee the stability of an entire tunnel, the average intersection angle ($I_A$) between the tunnel's horizontal-projected directions ($I_S$) and tangent directions ($I_T$) of the contour lines it traverses should be close to 90°. This factor is evaluated by the following criterion $I_A$':

$$\begin{aligned} I'_A &= \left| I_S - \frac{\sum_{i=1}^{k} I_{Ti}}{k} - 90 \right| \\ &= \left| \arctan \frac{Y_{P-S} - Y_{P-E}}{X_{P-S} - X_{P-E}} - \frac{\sum_{i=1}^{k} I_{Ti}}{k} - 90 \right| \end{aligned} \tag{5}$$

where $k$ = the total number of contour lines the tunnel traverses and $I_{Ti}$ = the tangent direction of $i$th contour line; $X_{P\text{-}S}$, $Y_{P\text{-}S}$, $X_{P\text{-}E}$, $Y_{P\text{-}E}$ = the horizontal coordinates of the entrance and exit of a tunnel.

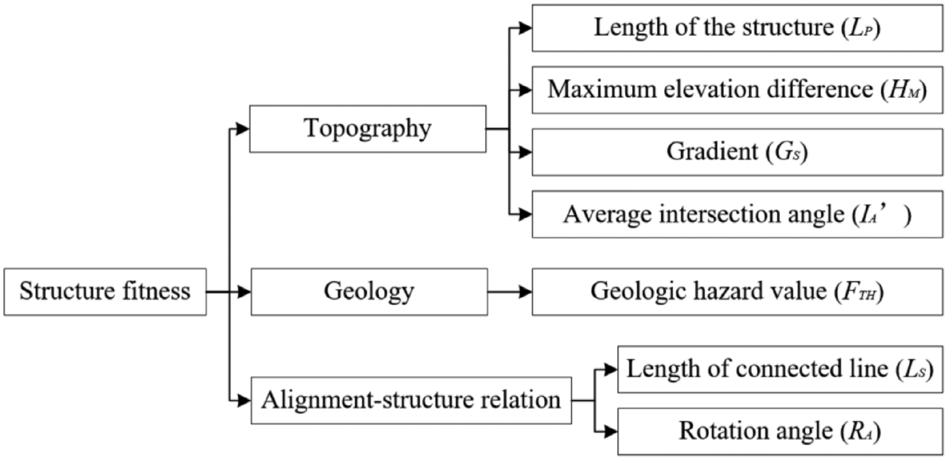



**FIGURE 8** Hierarchical structure of the multi-criteria decision-making (MCDM) problem in this study.

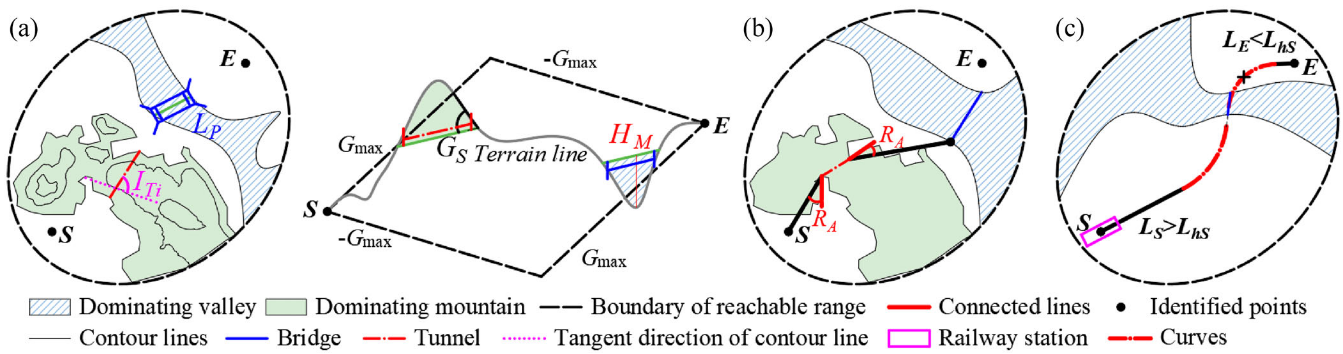



**FIGURE 9** (a) The topographic factors; (b) $R_A$ of a dominating structure's dominating points (DPs); and (c) geometric constraints for dominating structures.

The pressure bias is reduced as the value of $I_A$' decreases, and the tunnel's stability is improved. It is noted that since the bridge has no direct contact with the rock mass, this factor is not considered in evaluating the bridge's topographic fitness. These factors are illustrated in Figure 9a.

*Geologic factors*

Geo-hazard risk for an alignment is difficult to evaluate due to numerous contributing factors. These factors are hard to quantify and aggregate. At present, there are only few specialized studies considering geo-hazard risk in the alignment optimization realm. Therefore, a geo-hazard risk value $F_H$ proposed in Pu et al. (2021) is employed here as the geologic influencing factor. $F_H$ is assigned to each grid in railway-reachable ranges and its value in every grid cell traversed by a dominating structure can be accumulated to form the total geo-hazard risk value ($F_{TH}$) for this structure:

$$F_{TH} = \sum_{i=1}^{p} F_{Hi}(x_i, y_i, z_i) \tag{6}$$

where $p$ = the number of grids traversed by the structure, $F_{Hi}$ = the geo-hazard risk value of $i$th grid, and $x_i, y_i, z_i$ are its spatial coordinates.

*Alignment–structure relations*

In addition to the structure itself, the evaluation of a structure should also be integrated with the complete alignment. Therefore, the alignment–structure relations should be taken into account in the optimization of structure layouts. Since a structure can be represented by its point pair, the alignment–structure relations can be simplified as the circuity conditions of the point pair. To form an entire alignment, two DPs of a point pair should be connected with two identified points (either one DP of another dominating structure or an endpoint). The connected line should preferably be straight and short to reduce the construction cost as well as to facilitate operation condition. However, due to the distribution of dominating landforms, the connected line should be circuitous in order to entirely bypass the dominating landform in the entrance and exit of the point pair. In this regard, the circuity condition of one DP is quantified by two criteria.

First, the length of line ($L_S$) that connects a DP and another identified point is expected to be short and its value is computed as follows:

$$L_S = \max\left\{\frac{|H_{IP} - H_{EP}|}{G_{\max}} \cdot L_{IE}, L_{IE}\right\} \tag{7}$$

where $H_{IP}, H_{EP}$ = the elevations of the identified point and the DP; $L_{IE}$ = the linear distance between these two points.

Besides, if the connected line in one step-size range (whose value is taken from a PSO method proposed in Song et al., 2020) from the DP is still in the dominating landform, this part is rotated until the dominating landform is entirely bypassed, as shown in Figure 9b. The circuity condition for this DP increases with the reduction of the corresponding rotation angle ($R_A$).

### 3.1.3 | Constraints

There are specific constraints for dominating structures that should be satisfied. These constraints can be categorized into three types:

*Geometric constraints*
As a part of an alignment, the geometric profile of the dominating structures should be consistent with the alignment. Specifically, when the DP is close to the start or endpoint (i.e., there is only one intersection angle between one end of a dominating structure and the start or endpoint, as shown in Figure 9c), the minimum horizontal circular curve radius ($R_{\min}$) is used at this intersection. Afterward, the tangent length between one end of this curve and the start or endpoint ($L_S, L_E$) is computed. Then, the following requirements must be satisfied:

1. $L_S$ or $L_E$ should exceed half the length of a railway station ($L_{hS}$).
2. This curve ($CV$) should not overlap the dominating structure ($SS$).

$$L_{S(E)}\left(R_{\min}, X_S, Y_S\right) > \frac{L_{hS}}{2} \tag{8}$$

$$CV\left(R_{\min}\right) \cap SS\left(D_S\right) = \emptyset \tag{9}$$

If any one of the above conditions is not satisfied, the location of this dominating structure violates the geometric constraints. These geometric constraints are denoted as $C_G(D_S) \leq 0$.

*Locational constraints*
For a railway alignment, forbidden regions ($R_F$), such as mining areas, military bases, and nature reserves, should be avoided for dominating structures:

$$R_F \cap SS\left(D_S\right) = \emptyset \tag{10}$$

In addition to these forbidden regions, geologic hazard regions ($R_H$, including rock-falls, debris flows, and landslides) should also be considered. Generally, these geologic hazards do not affect tunnels and have a slight influence on bridges, except for abutments and portals.

Abutments ($B_A$) and portals ($T_P$) are the most important and fragile components of bridges and tunnels. As these components are in direct contact with the ground surface, they are greatly affected by geologic hazards and crucially influence the stability of the whole structures. Therefore, they should entirely bypass geologic hazard regions:

$$\begin{aligned} R_H \cap B_A &= \emptyset \\ R_H \cap T_P &= \emptyset \end{aligned} \tag{11}$$

Moreover, the clearances ($H_C$) must be satisfied when a dominating structure (whose design elevation is designated as $H_D$) traverses rivers, existing roads, or railways (with the elevation $H_E$):

$$\left|H_D\left(D_S\right) - H_E\left(D_S\right)\right| < H_C \tag{12}$$

These locational constraints are denoted as $C_L(D_S) \leq 0$.

*Construction constraints*
Due to the limitations of construction technique and construction period, exceedingly high bridges and long tunnels are not acceptable for railway alignments. Thus, the length of tunnels and the maximum height of bridges should be within the required ranges, which is denoted as $C_C(D_S) \leq 0$.

Thus, the dominating structure optimization model is formulated as:

$$\begin{aligned} \text{Maximize } &: F_S\left(D_S\right) \\ \text{Subjectto } &: C_G\left(D_S\right) \leq 0 \\ & C_L\left(D_S\right) \leq 0 \\ & C_C\left(D_S\right) \leq 0 \end{aligned} \tag{13}$$

## 3.2 | Lower level—Alignment optimization

### 3.2.1 | Design variables

As mentioned above, a 3D alignment is generally represented by its PIs and the corresponding tangent and curve configurations. According to Li et al. (2016), the design variables for this problem include the coordinates ($X_H$, $Y_H$) and circular curve radius ($R_H$) of HPIs as well as the mileage ($K_V$) and design elevation ($H_V$) of VPIs. In this study, the layouts of dominating structures are prioritized and the study area is divided into several subsections. After the study area is separated into subsections, each subsection possesses its own PIs set. The geometric parameters of these PIs are denoted as alignment subsections' design variables ($D_A$). Moreover, the DPs' design elevations can be adjusted during alignment optimization process. Thus, $Z_S$ for dominating structures are treated as additional design variables for alignment optimization.

Assuming the number of alignment subsegments is $i$, the design variables for the study area are provided below:

$$
\begin{aligned}
&\text{design variables } : D_T = [Z_S, D_A] \\
&= [Z_S, (X_H, Y_H, R_H, K_V, H_V)] \\
&\text{where } X_H = [X_{H1}, \dots, X_{Hi}]^T \\
&\quad Y_H = [Y_{H1}, \dots, Y_{Hi}]^T \\
&\quad R_H = [R_{H1}, \dots, R_{Hi}]^T \\
&\quad K_V = [K_{V1}, \dots, K_{Vi}]^T \\
&\quad H_V = [H_{V1}, \dots, H_{Vi}]^T
\end{aligned}
\tag{14}
$$

### 3.2.2 | Objective function

In alignment optimization problems, the construction cost is usually treated as the objective function. According to Li et al. (2016), the construction cost of a railway alignment includes costs for earthwork ($C_E$), bridges ($C_B$), tunnels ($C_T$), length-dependent component ($C_L$), and right-of-way areas ($C_R$). Typically, the division of the alignment structures ($C_S$) is determined by the elevation difference ($\Delta h$ for fill and $\Delta d$ for cut) between the design elevation and its corresponding ground profile. The construction cost for an alignment ($C_A$) is expressed as follows:

$$
\begin{aligned}
&C_A = C_S + C_L + C_R \\
&\text{where } C_S = \begin{cases} C_E & \Delta h < H_b \text{or} \Delta d < H_t \\ C_B & \Delta h > H_b \\ C_T & \Delta d > H_t \end{cases}
\end{aligned}
\tag{15}
$$

In this paper, the layouts of dominating structures are pregenerated and, for simplicity, their construction costs are determined by the spatial coordinates of their point pairs. Relatively, the construction cost of alignment in subsection is still computed with Equation (15). Thus, the total construction cost ($C_{TC}$) in this study consists of two terms and is formulated as follows:

$$
C_{TC} = \sum_{i=1}^{n} C_{DSi}(D_{Si}) + \sum_{j=1}^{m} C_{Aj}\left(D_{Aj}\right) \tag{16}
$$

where $n$, $m$ = the numbers of dominating structures and alignment subsections, respectively; $C_{DSi}$, $D_{Si}$ = the construction cost and design variables of $i$th dominating structure; $C_{Aj}$, $D_{Aj}$ = the construction cost and design variables of $j$th subsection.

Besides, the geologic conditions should also be considered in mountainous regions with widely distributed geologic hazards. According to Pu et al. (2021), a geologic hazard value $F_H$ is assigned to every grid in railway-reachable ranges. Thus, the geologic hazard value ($C_{HA}$) of an alignment can be accumulated by the $F_H$ of each grid cell traversed by the alignment. Supposing the number of traversed grids is $I$, $C_{HA}$ is formulated as:

$$
C_{HA} = \sum_{i=1}^{I} F_{Hi}(D_T) \tag{17}
$$

### 3.2.3 | Constraints

Multiple constraints must be satisfied to find an appropriate railway alignment in complex mountainous regions. These alignment constraints include the general alignment constraints (such as on maximum allowable bridge height and tunnel length) $C_{AG}(Z_S,\ D_A) \leq 0$ proposed in Li et al. (2017) and the geo-hazard constraints (such as on forbidden geo-hazard regions and tunnel portal locations) $C_{AH}(Z_S,\ D_A) \leq 0$ developed in Song et al. (2020).

In summary, the alignment optimization model is formulated as:

$$
\begin{aligned}
&\text{Objective1} - \text{Minimize } : C_{TC}(Z_S, D_A) \\
&\text{Objective2} - \text{Minimize } : C_{HA}(Z_S, D_A) \\
&\text{Overallobjective } : \\
&\text{Minimize } : C_T(Z_S, D_A) = [C_{TC}(Z_S, D_A), C_{HA}(Z_S, D_A)] \\
&\text{Subjectto } : \begin{cases} C_{AG}(Z_S, D_A) \leq 0 \\ C_{AH}(Z_S, D_A) \leq 0 \end{cases}
\end{aligned}
\tag{18}
$$

## 4 | SOLUTION METHOD

A three-stage optimization method is proposed for solving the bi-level optimization model. At the first stage, all the endpoints of dominating structures are generated by a scanning process and the infeasible structure alternatives are precluded by specific screening operators. At the second stage, an MCDM procedure is adopted for evaluating the fitness of dominating structure layouts. Afterward, the optimized dominating structure layouts are selected and the study area is divided into several subsections. At the third optimization stage, the endpoints of determined dominating structures are assigned as required traversed points (RTPs) and a bi-objective PSO (Song et al., 2020) is used for optimizing alignments incorporating RTPs.

## 4.1 | Structure locations determination

### 4.1.1 | Scanning process

In this process, the whole railway horizontally reachable area is scanned with line segments in various directions to obtain all point pairs crossing the dominating landforms. The parametric formulas of these line segments $\boldsymbol{L}$ $(\alpha, i, j)$ are specified as follows:

$$\begin{cases} x_L = \dfrac{x_S + x_E}{2} + \sin\left(f_{SE} + \alpha\right) \cdot w \cdot i + \cos\left(f_{SE} + \dfrac{\pi}{2} + \alpha\right) \cdot w \cdot j \\ y_L = \dfrac{y_S + y_E}{2} + \cos\left(f_{SE} + \alpha\right) \cdot w \cdot i + \sin\left(f_{SE} + \dfrac{\pi}{2} + \alpha\right) \cdot w \cdot j \end{cases}$$
$$i = \left(-\frac{\gamma_{\max} \cdot l_{SE}}{2 \cdot w}, \ldots, 1, 2, \ldots, \frac{\gamma_{\max} \cdot l_{SE}}{2 \cdot w}\right)$$
$$j = \left(-\frac{\gamma_{\max} \cdot l_{SE}}{2 \cdot w}, \ldots, 1, 2, \ldots, \frac{\gamma_{\max} \cdot l_{SE}}{2 \cdot w}\right)$$
$$\alpha = \left(-\theta, \ldots, 0, \frac{\pi}{36}, \ldots, \theta\right) \tag{19}$$

where $x_L$, $y_L$ = the coordinates of the points in these line segments; $x_S$, $y_S$, $x_E$, $y_E$ = the coordinates of start and endpoints; $f_{SE}$ = azimuth of the start and endpoints; $w$ = width of DEM's grid; $i, j, \alpha$ = the $[i + (\gamma_{\max} \bullet l_{SE})/2/w]$th point in $[j + (\gamma_{\max} \bullet l_{SE})/2/w]$th line segment and the angle between this line segment and the straight line between two endpoints (Figure 10).

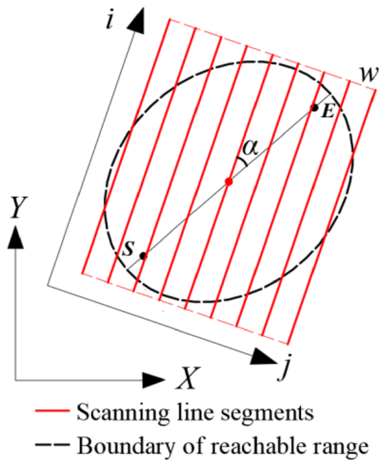



**FIGURE 10** Scanning line segments.

### 4.1.2 | Screening operators

There might be several point pairs obtained from the above scanning process whose connecting segments do not entirely traverse the dominating landform. These point pairs must be eliminated from the feasible solutions integrated with landform characteristics. Generally, the geometric characteristics of valleys and mountains are different. Valleys are typically caused by surface erosion, and thus they usually have V-shaped or U-shaped profiles in longitudinal and quasi-dendritic forms in horizontal (Williams & Phillips, 2001), while mountains are formed by rock uplift or orogeny and usually spread widely in a belt-like configuration (Jin et al., 2009; Prasicek et al., 2015). Therefore, the point pairs of dominating bridges and tunnels are screened with different operators, which consider specific landform features.

For a dominating valley, to preclude the point pairs on the dendrite branches, each point pair is disturbed within a certain range. If the line segment formed by this point pair could entirely bypass the valley after disturbance, the corresponding point pair should be eliminated from the possible solutions, as shown in Figure 11a. With the implementation of this bridge-related screening operator, all the remaining point pairs that satisfy the aforementioned constraints are stored in an external alternative set.

A dominating mountainous area can be divided into several irregular belt-like components. Although a segment connecting a point pair may traverse a fraction of the mountain, the circuitous alignments could also be obstructed by other parts of these mountains. Thus, the tunnel-related screening operator is described as follows. Primarily, the line segment formed by each point pair is disturbed in a certain range. If the disturbed line segment entirely bypasses the mountainous region, the corresponding point pair is eliminated, as shown in Figure 11b. Otherwise, the point pairs are connected with the start or end point or the DP of another dominating structure. Afterward, the line segments ($L_{P1}$, $L_{P2}$) formed by the point pair and their connected points are checked. If $L_{P1}$ or $L_{P2}$ traverses the mountain regions, the traversed part is disturbed within the horizontally reachable range (Figure 11c). If it can entirely bypass the mountain after this disturbance (Figure 11d) and the aforementioned constraints are sat-

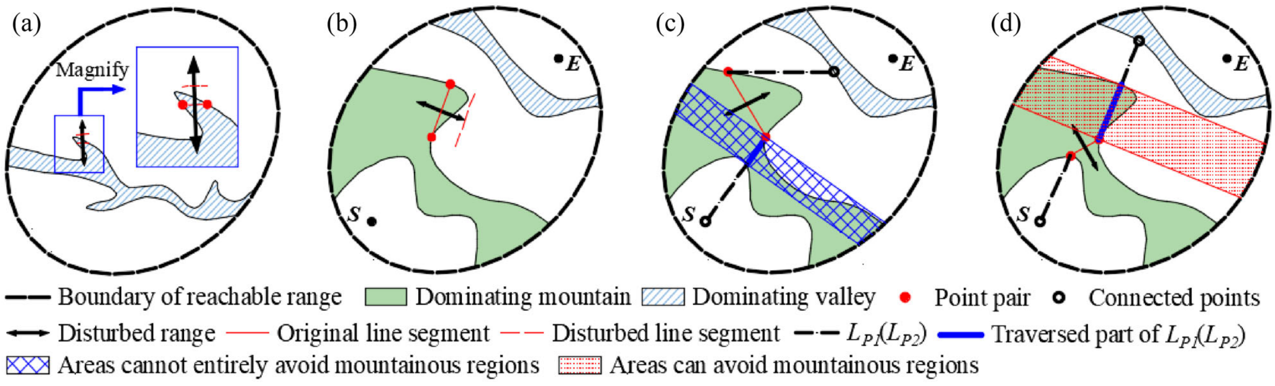



**FIGURE 11** (a) Bridge-related screening operator; (b), (c), and (d) tunnel-related screening operator.

isfied, the corresponding point pair is recorded in the external alternative set.

## 4.2 | MCDM analysis

Through the analyses in Section 3.1.2, it is found that there are three categories, which are further subdivided into seven kinds of factors (as shown in Figure 8), that significantly influence the layouts of structures. These factors should be aggregated to evaluate the comprehensive fitness of structure layouts. The relations among these factors are very complex and mutually constrained. The trade-offs between the structures themselves and their relations with other alignment sections should be properly considered. This is intrinsically a hierarchical MCDM (Karlson et al., 2016) problem and thus, it can be solved with an analytical hierarchy process (AHP; Saaty, 2008) combined with a criteria importance through inter-criteria correlation (CRITIC; Diakoulaki et al., 1995) method for obtaining the criterion weight of each factor. Specifically, the AHP is used for determining the criterion weight of each category and the CRITIC is adopted for computing the weights of subfactors in the categories.

After considering the advice of experts from the China Railway Siyuan Survey and Design Group Co. Ltd, the China Railway Eryuan Engineering Group Co. Ltd, and the China Railway First Survey and Design Institute Group Co. Ltd, the comparison matrix for all categories is obtained and shown in Table 1.

Then, the weights ($\omega$) of subfactors in each category are computed based on the specified data's fluctuation ($\varepsilon$) and correlation ($\rho$) (Diakoulaki et al., 1995):

$$\omega = \varepsilon \cdot (1 - \rho) \tag{20}$$

where $\varepsilon$ reflects the deviation degree of a criterion and the criterion with a larger $\varepsilon$ value can provide more information and take greater weight; $\rho$ represents the differentiation from other criteria, which should preferably be smaller for providing more information.

**TABLE 1** The comparison matrix for all the categories.

| Structure fitness | Topography | Geology | Alignment–structure relations | Criterion weight |
|---|---|---|---|---|
| Topography | 1 | 1 | 2 | 0.4 |
| Geology | 1 | 1 | 2 | 0.4 |
| Alignment–structure relations | 1/2 | ½ | 1 | 0.2 |

Supposing the number of dominating structures is $N$ in a study area, then, the structure combination fitness ($F_S$) can be expressed as follows:

$$\begin{aligned} F_S &= \sum_{q=1}^{N}\sum_{i=1}^{2}\left[W_{Ci}\cdot\sum_{j=1}^{Si}\left(W_{Sij}\cdot x_{qij}\right)\right] + W_{C3}\cdot\sum_{j=1}^{S3}\left(W_{S3j}\cdot x_{Sj}\right) \\ &\quad + W_{C3}\cdot\sum_{j=1}^{S3}\left(W_{S3j}\cdot x_{Ej}\right) + \sum_{q=1}^{N-1} W_{C3}\cdot\sum_{j=1}^{S3}\left(W_{S3j}\cdot x_{qj}\right) \\ &= F_{S1} + F_{S2} + F_{S3} + F_{S4} \end{aligned} \tag{21}$$

where $W_{Ci}$ = criterion weight of $i$th category; $Si$ = number of the subfactors in $i$th category; $W_{Sij}$ = subcriterion weight of $j$th subfactor in $i$th category; $x_{qij}$ = normalized value of $j$th subfactor in $i$th category for $q$th structure; $x_{Sj}$, $x_{Ej}$ = normalized value of $j$th subfactor that influences the alignment–structure relations between the structure and the start or end points, respectively; $x_{qj}$ = normalized value of $j$th subfactor that influences the alignment–structure relations between two structures. The detailed illustrations are shown in Figure 12a.

The point pairs in the external alternative sets can be combined, as shown in Figure 12b, and their $F_S$ are computed with Equation (21). Then, the set of dominating

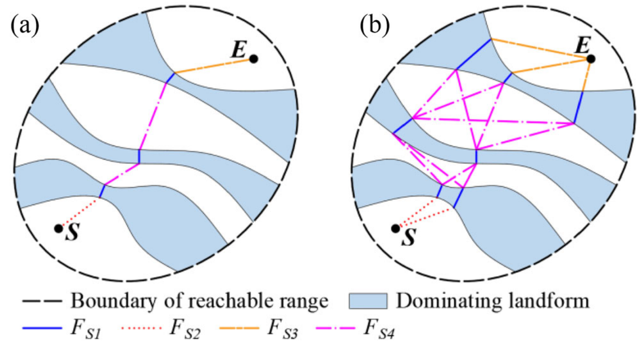



**FIGURE 12** Illustrations of (a) structure combination fitness and (b) structure combined forms.

**TABLE 2** Unit costs.

| Item | Value | Item | Value |
|---|---|---|---|
| Rail track (USD/m) | 646.71 | Right-of-way (USD/m$^2$) | 12.22 |
| Fill earthwork (USD/m$^3$) | 3.59 | Cut earthwork (USD/m$^3$) | 4.02 |
| ($L$ < 500 m) Bridge ($10^4$ USD/m) | 0.88 | ($L$ > 500 m) Bridge ($10^4$ USD/m) | 1.09 |
| ($L$ < 10 km) Tunnel ($10^4$ USD/m) | 0.96 | ($L$ > 10 km) Tunnel ($10^4$ USD/m) | 1.37 |
| ($L$ > 20 km) Tunnel ($10^4$ USD/m) | 1.77 | One bridge abutment/tunnel portal ($10^4$ USD) | 2.87 |

structures on an alignment alternative with the maximum $F_S$ value is designated as the optimized combination form of dominating structures. The corresponding point pairs are treated as the RTPs for the subsequently alignment optimization process.

### 4.3 | Alignment generation

After the locations of dominating structures are determined, the study area is separated into several subsections based on the assigned point pairs. Then, the alignment generation is a bi-objective alignment optimization problem incorporating RTPs. Hence it can be solved by a customized bi-objective PSO proposed by Song et al. (2020).

If the length of the line segment formed by the point pair is less than a step size assigned in the PSO method, this line segment is extended to both sides from its center point until its length equals the step size. In this regard, the endpoints of the newly formed line segment are considered as RTPs. RTPs are assigned as new start or endpoints of each subsection, and the PSO method is applied for generating optimized alignments. RTPs' design elevation can be adjusted for decreasing construction cost and geo-hazard risk concurrently when the PSO is employed to each subsection. The detailed information is provided in Figure 13.

## 5 | APPLICATIONS

Two realistic cases in complex mountainous regions are used to demonstrate the effectiveness of the proposed model and method in presetting dominating structures and optimizing alignments accordingly. This program is run on a computer with Intel Core i7-7700 CPU @ 3.60 GHz. The optimized alignments ($A_S$ and $A_{S1}$ for cases 1 and 2, respectively) are generated after 936 and 1,560 s. For these two cases, both a 3D-DT (Pu et al., 2019) and a PSO (Song et al., 2020), which are customized for alignment optimization in constrained mountainous regions, failed to generate a better alignment alternative (regarding construction cost and geo-hazard risk) than manually designed alignments ($A_M$ and $A_{M2}$ for cases 1 and 2, respectively).

Specifically, the best alignments generated by the 3D-DT ($A_{3D\text{-}DT}$, $A_{3D\text{-}DT2}$) and PSO ($A_{PSO}$, $A_{PSO2}$) for cases 1 and 2, respectively, are compared with the manual solutions. Their horizontal, vertical alignments and structure data are shown in Figures 14 and 15. For case 1, the dominating bridge lengths of the $A_{3D\text{-}DT}$ and $A_{PSO}$ are both longer than that of $A_M$ (i.e., 3,927 and 536 m for $A_{3D\text{-}DT}$ and $A_{PSO}$, respectively). For case 2, $A_{3D\text{-}DT2}$ presents a dominating bridge of 2,016 m compared to $A_{M2}$ and $A_{PSO2}$. Although the total railway length of $A_{PSO2}$ is shorter (i.e., 10,007 m) than that of $A_{M2}$, $A_{PSO2}$' dominating tunnel length increases by 12,686 m compared with that of $A_{M2}$.

### 5.1 | Case 1

#### 5.1.1 | Case profile

In this case, a railway section covered by an extremely deep valley is chosen as an actual example. The linear distance between the start and end points is 53,033 m, while the elevation difference between them is 975 m. Thus, the natural gradient ($G_T$) of the study area is 975/53,033 = 18.4‰. The topographic and geologic maps of these region are provided by China Railway Eryuan Engineering Group Co. Ltd. and shown in Figure 16. The unit costs are provided in Table 2.

It can be observed that the minimum and maximum elevations of this study area are 2,802 m and 5878 m, respectively. The terrain is drastically undulating with high

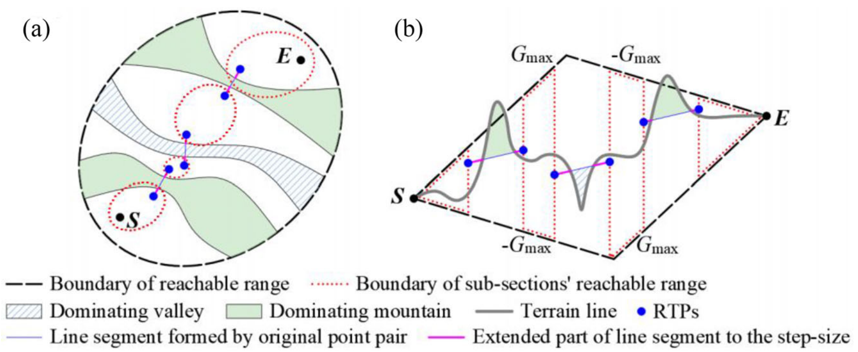
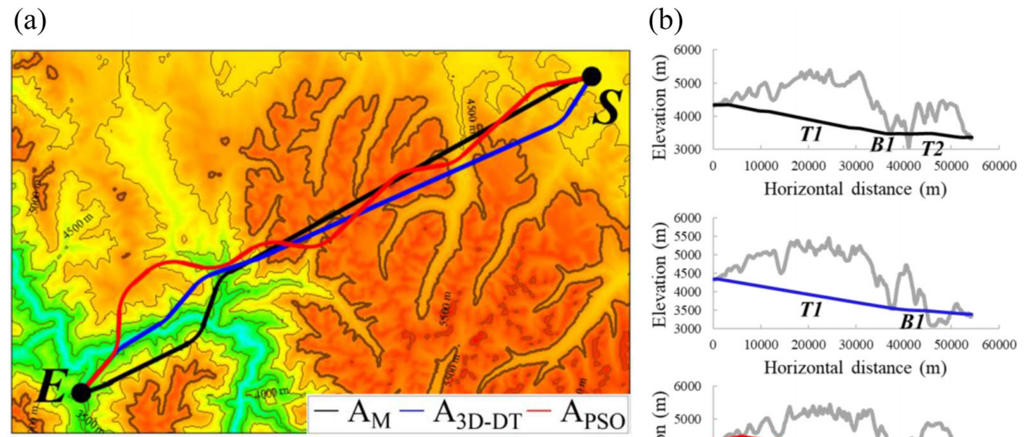
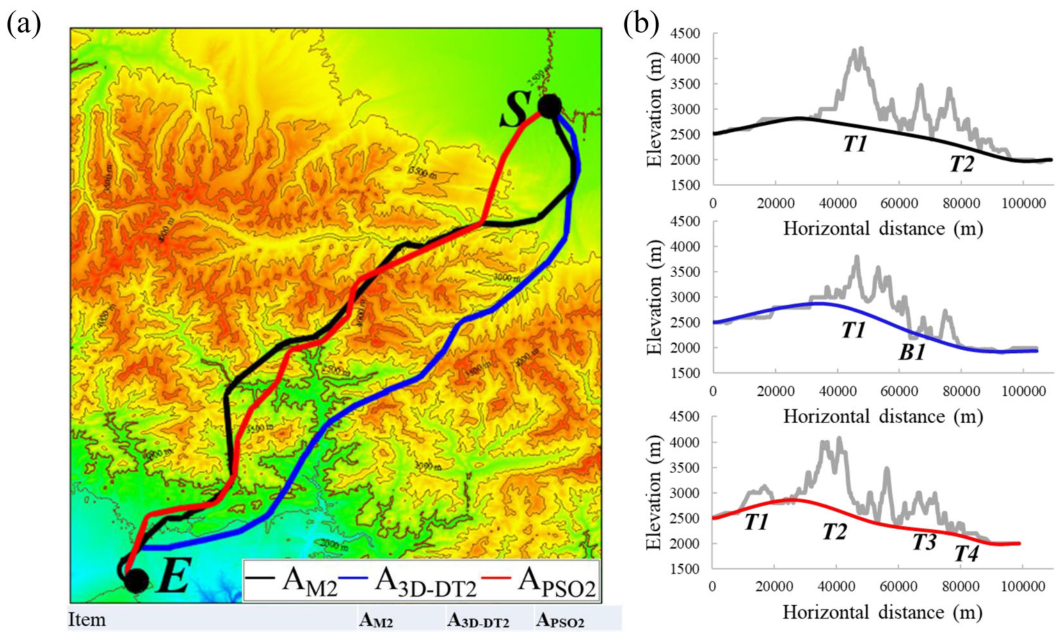



**FIGURE 13** (a) Horizontal and (b) vertical demonstrations of required traversed points (RTPs) and corresponding subsections.




| Item | $A_M$ | $A_{3D\text{-}DT}$ | $A_{PSO}$ |
|---|---|---|---|
| Railway length (m) | 54,300 | 54,259 | 58,007 |
| Tunnel number-length (L>10km) | 2-48,633 | 1-36,109 | 2-44,110 |
| Bridge number-length (H>100m) | 1-861 | 1-4,788 | 1-1,397 |

**FIGURE 14** (a) Horizontal alignments and (b) vertical alignments of $A_M$, $A_{3D\text{-}DT}$ and $A_{PSO}$, for case 1.




| Item | $A_{M2}$ | $A_{3D\text{-}DT2}$ | $A_{PSO2}$ |
|---|---|---|---|
| Railway length (m) | 108,804 | 104,138 | 98,797 |
| Tunnel number-length (L>10km) | 2-62,833 | 1-27,678 | 4-75,519 |
| Bridge number-length (H>100m) | N/A | 1-2,016 | N/A |

**FIGURE 15** (a) Horizontal alignments and (b) vertical alignments of $A_{M2}$, $A_{3D\text{-}DT2}$ and $A_{PSO2}$, for case 2.

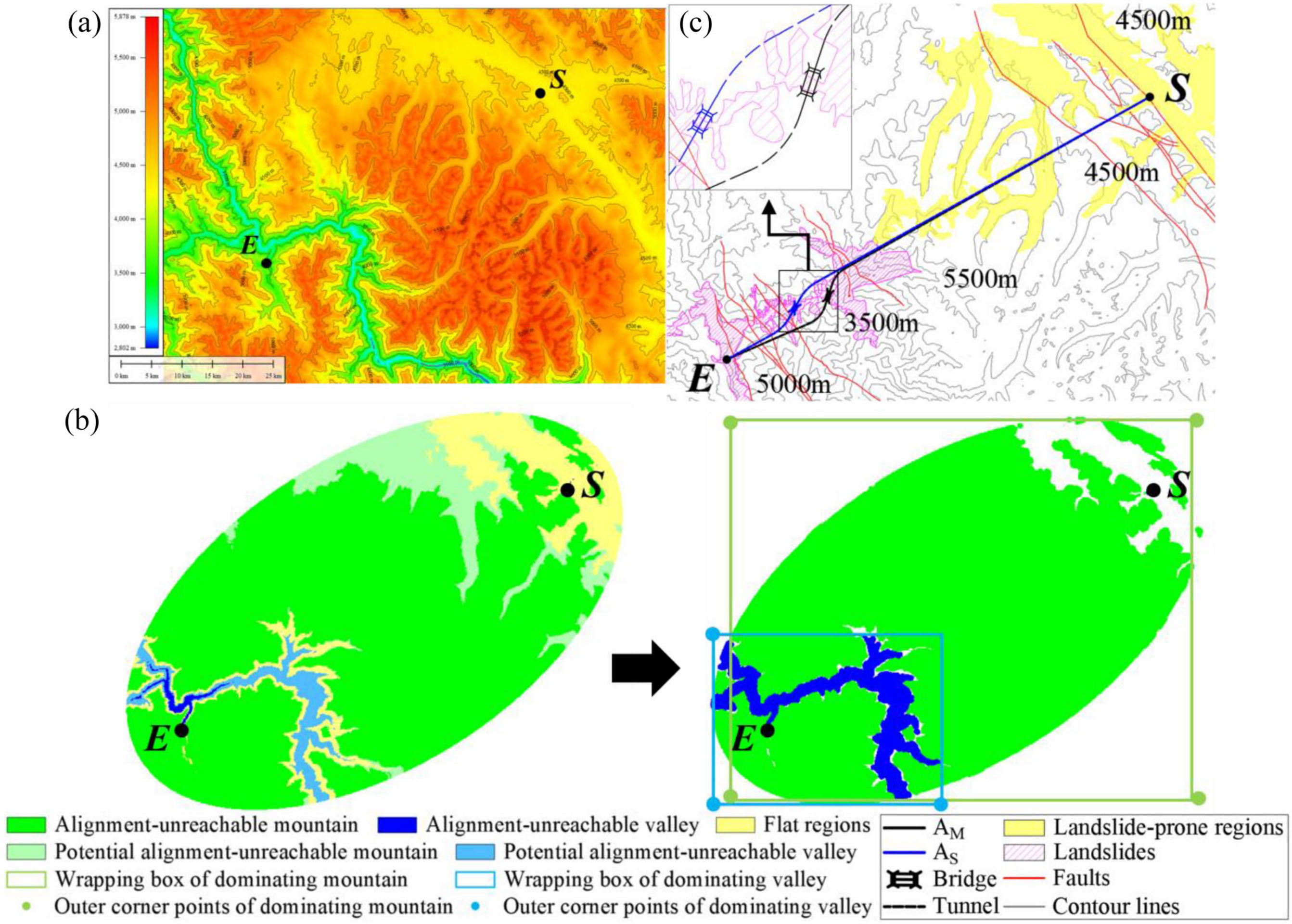



**FIGURE 16** (a) Terrain map of the study area; (b) dominating landform determination; and (c) the horizontal alignments of $A_M$ and $A_S$.

mountains and deep valleys between the start and endpoints. Moreover, the shape characteristics of mountains and valleys are belt-like and quasi-dendritic, respectively. Regarding the geologic conditions, there are four landslides, three landslide-prone areas, and eight faults spread widely over the study area. It is noted that the landslides are mainly concentrated at the endpoint and the landslide-prone areas are clustered at the start point. The faults are densely distributed around the start and endpoints. The sensitive locations such as tunnel portals and bridge abutments must entirely bypass these geologic hazards regions.

### 5.1.2 | Determining dominating landforms

First, the railway spatial reachability analysis is conducted in this section. According to the design standard, the maximum gradient ($G_{\max}$) of this railway is 30‰. Compared with $G_T$, the elevation difference between two endpoints can be overcome without additional circuity. However, in such drastically undulating terrain, the alignment should usually be somewhat circuitous to achieve proper structure layouts. Thus, the maximum circuity coefficient $\gamma_{\max}$ is set at 1.2. Afterward, the alignment-unreachable valleys and mountains can be obtained based on $G_{\max}$ and $\gamma_{\max}$, as shown in Figure 16b.

It is observed that this study area is almost entirely covered by high mountains and deep valleys. In particular, 83.5% and 3.8% of the area, respectively, consist of alignment-unreachable mountains and valleys. Moreover, there are two and one pairs of diagonal points outside the railway horizontally reachable range, respectively, for the dominating mountain and valley. Thus, the alignments must traverse these dominating mountains and valleys with tunnels and bridges. Specifically, the alignments must first traverse the dominating mountain with tunnels from the start point, and then a bridge is necessary for traversing the quasi-dendritic valleys. Then, tunnels are required until the end point is connected.

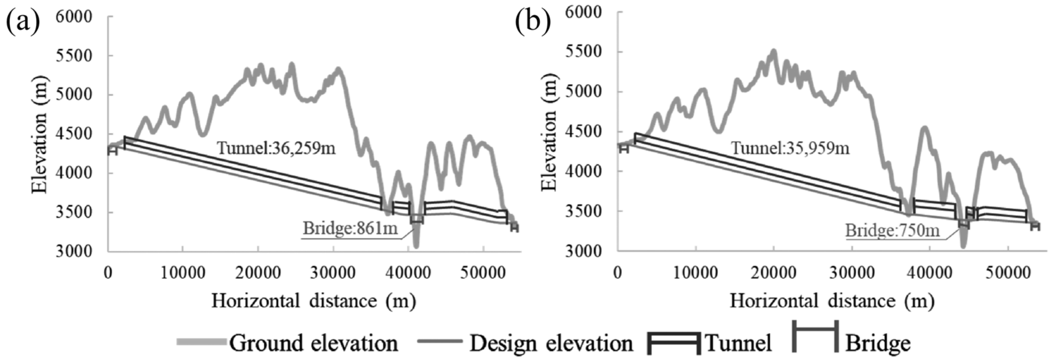



FIGURE 17 The vertical alignments of (a) $A_M$ and (b) $A_S$.

TABLE 3 Detailed comparisons of $A_M$ and $A_S$.

| Item | $A_M$ | $A_S$ |
|---|---|---|
| Railway length (m) | 54,300 | 53,570 |
| Right-of-way area ($m^2$) | 255,372 | 251,496 |
| Embankment / excavation volume ($m^3$) | 31,975 / 31,681 | 28,910 / 30,305 |
| Bridge number-length (m) | 3-1379 | 3-1410 |
| Tunnel number-length (m) | 3-51,838 | 4-50,794 |
| Total construction cost (million USD) | 893.75 | 830.81 |
| **Improvement** | N/A | 7.0% |
| Total geo-hazard risk | 1,622 | 1,629 |
| **Improvement** | 0.4% | N/A |

TABLE 4 Unit costs.

| Item | Value | Item | Value |
|---|---|---|---|
| Rail track (USD/m) | 646.71 | Right-of-way (USD/$m^2$) | 13.65 |
| Fill earthwork (USD/$m^3$) | 3.59 | Cut earthwork (USD/$m^3$) | 4.02 |
| ($H < 50$ m, $L > 3$ km) Bridge ($10^4$ USD/m) | 1.06 | ($L < 10$ km) Tunnel ($10^4$ USD/m) | 0.96 |
| ($L > 20$ km) Tunnel ($10^4$ USD/m) | 1.37 | One bridge abutment/tunnel portal ($10^4$ USD) | 2.87 |

### 5.1.3 | Alignment solutions

Afterward, the optimized alignment ($A_S$) generated by the three-stage solution method (which is proposed in Section 4) is compared with the best alignment provided by experienced designers in China Railway Eryuan Engineering Group Co. Ltd. ($A_M$). Their horizontal and vertical alignments are illustrated in Figures 16c and 17 and their numerical comparisons are provided in Table 3.

It noteworthy that the maximum bridge lengths of $A_M$ and $A_S$ are 861 and 750 m, respectively. Moreover, compared with $A_M$, the railway length of $A_S$ decreases by (54,300 − 53,570) = 730 m and the maximum tunnel length of $A_S$ also decreases by (36,259 − 35,959) = 300 m. By shortening the most expensive structures (i.e., high bridge and long tunnel), the construction cost of $A_S$ is reduced by 7.0% compared to that of $A_M$.

Regarding the geo-hazard risk, the alignment locations are nearly the same for $A_S$ and $A_M$. Thus, these two alternatives basically traverse the similar geo-hazard regions. The major geometric dissimilarity between $A_S$ and $A_M$ is in the section between the bridge exit and the end point. $A_M$ traverses one fault, while $A_S$ traverses two faults and a longer landslide region than $A_M$. However, because both of these alignments traverse these geo-hazard regions through tunnels, their geo-hazard risk values are only slightly different (i.e., $A_M$ has a 0.4% improvement over $A_S$).

The above analyses confirm that the proposed method can find more appropriate bridge abutment locations as well as a lower cost complete alignment in a study area with a dominating valley, compared with the manually designed one.

## 5.2 | Case 2

### 5.2.1 | Case profile

A mountain-covered railway section is used as another case. The linear distance and elevation difference between the start and end points are 88,808 and 632 m, respectively. The topographic and geologic maps of this study area are provided by China Railway First Survey and Design Institute Group Co., Ltd. and shown in Figure 18. The unit costs of this railway are provided in Table 4.

It can be found that although the natural gradient of this region is only 632/88,808 = 7.1‰, there are mountains with elevations ranging from 3,000 m to 4,725 m that

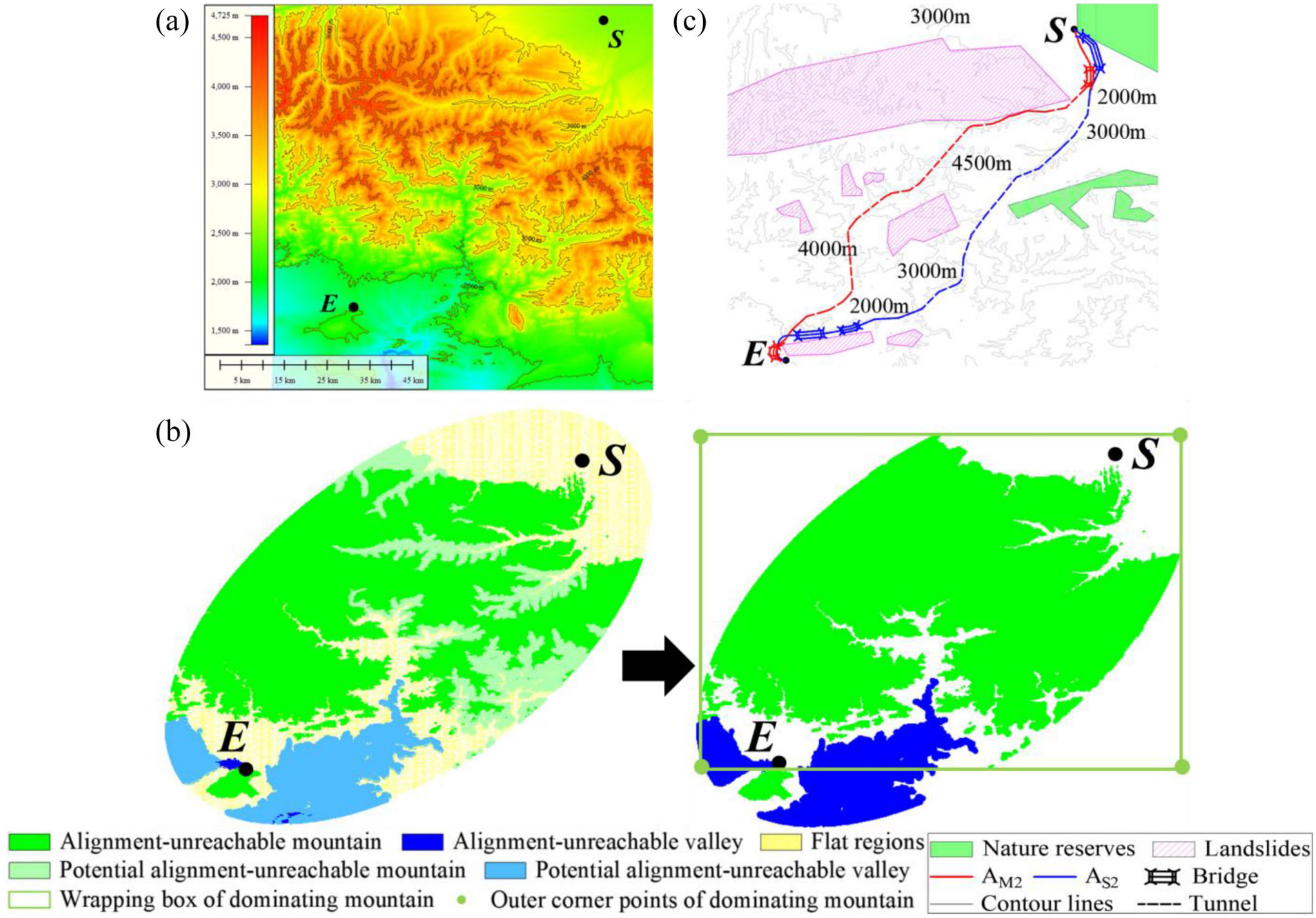



**FIGURE 18** (a) Terrain map of the study area; (b) dominating landform determination; and (c) the horizontal alignments of $A_{M2}$ and $A_{S2}$.

would be inevitably traversed by the alignments. The geologic conditions include seven landslides and three natural conservation regions which are scattered and occupy most of the study area. Therefore, alignment diversity is severely constrained in these regions.

### 5.2.2 | Determining dominating landforms

To bypass the alignment unfavorable regions as well as to find an appropriate location for crossing the mountains, the maximum circuity coefficient ($\gamma_{max}$) and the maximum gradient ($G_{max}$) of this railway are set at 1.2 and 20‰, respectively. Then, the alignment-unreachable mountains can be found according to these two parameters, as illustrated in Figure 18b.

The landform segmentation results further verify that the alignments must traverse large-scale mountainous ranges and their geometric shapes are strictly restricted by the valleys surrounding the end point. Specifically, 54.5% and 7.5% of the horizontally reachable regions are covered by alignment-unreachable mountains and valleys, respectively.

### 5.2.3 | Alignment solutions

After the implementation of the three-stage optimization method, an optimized alignment ($A_{S2}$) is produced and compared with the best alignment provided by experienced designers in China Railway First Survey and Design Institute Group Co., Ltd. ($A_{M2}$). Their horizontal and vertical alignments are shown in Figures 18c and 19 and their detailed comparisons are shown in Table 5.

It is noted that the total lengths of tunnels are 71,955 m and 50,506 m for $A_{M2}$ and $A_{S2}$, respectively. From the start point, $A_{S2}$ traverses a relatively narrow section of the dominating mountainous ranges with a tunnel whose length is 42,403 m. Then, the end point is connected with bridges, cuts and fills. Regarding $A_{M2}$, although its tunnel length in dominating mountainous regions is (42,403 − 37,599) = 4,804 m shorter than that of $A_{S2}$, an additional

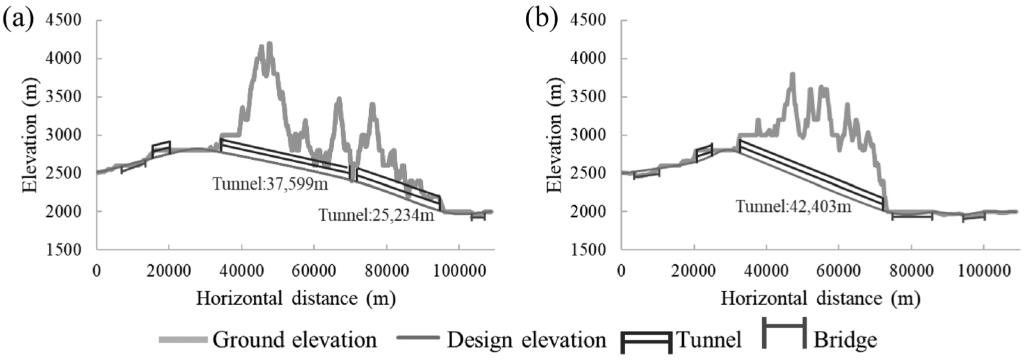



**FIGURE 19** The vertical alignments of (a) $A_{M2}$ and (b) $A_{S2}$.

**TABLE 5** Detailed comparisons of $A_{M2}$ and $A_{S2}$.

| Item | $A_{M2}$ | $A_{S2}$ |
|---|---|---|
| Railway length (m) | 108,804 | 108,824 |
| Right-of-way area ($m^2$) | 210,559 | 226,793 |
| Embankment / Excavation volume ($m^3$) | 71,604 / 42,662 | 40,193 / 39,847 |
| Bridge number-length (m) | 2-7,669 | 3-16,723 |
| Tunnel number-length (m) | 3-71,955 | 2-50,506 |
| Total construction cost (million USD) | 1,101.27 | 908.84 |
| **Improvement** | N/A | 17.5% |
| Total geo-hazard risk | 2,855 | 2,572 |
| **Improvement** | N/A | 10.0% |

25,234 m tunnel is imposed due to a landslide region near the tunnel portal. Thus, there are two adjacent long tunnels for $A_{M2}$. For the complete alignment alternatives, although the railway length of $A_{S2}$ (i.e., 20 m) slightly exceeds that of $A_{M2}$ and the total bridge length of $A_{S2}$ increases by (16,723 − 7,669) = 9,054 m compared with that of $A_{M2}$, the decrease in the total length of long tunnels results in a 17.5% lower construction cost for $A_{S2}$ than $A_{M2}$.

Regarding the geo-hazard risk, the geometric shapes of the alignments are restricted by the widely distributed landslides. Specifically, one earthwork section of $A_{M2}$ is quite close to a massive landslide region in order to reduce the length of first tunnel from the start point, but also increases the geo-hazard risk. Similarly, there are two successive bridges surrounding the end point of $A_{S2}$. The remainder sections of $A_{M2}$ and $A_{S2}$ are in the areas with relatively low geo-hazard risk. Since the geo-hazard risk is lower when the alignment traverses through bridges rather than earthworks, the total geo-hazard value of $A_{S2}$ is improved by 10.0% compared that of $A_{M2}$.

The above results indicate that the proposed method is capable of finding a better alignment solution (in terms of both construction cost and geo-hazard risk) than the alternative produced by human designers in the study area covered by dominating mountains.

## 6 | CONCLUSION

Alignment design in complex mountainous regions has always been a challenging problem. Particularly when there are high mountains and deep valleys between the start and endpoints, dominating structures are necessary for connecting two endpoints. These dominating structures have priority over other alignment sections because they basically determine the construction cost and duration for an entire alignment. To preset the dominating structures before the alignment optimization process, a bi-level optimization model and a three-stage solution method are proposed. The major contributions of this work are as follows:

1. A landform recognition method combining GIS analyses (including alignment spatial reachability analysis and elevation tendency) is designed to find the alignment-unreachable valleys and mountains. Then, a two-pass scan method is adopted to determine whether an alignment-unreachable area is a dominating landform.
2. A bi-level optimization model for determining dominating structures before alignment optimization is developed. At the upper level, seven factors related to structure fitness are quantified in terms of topography, geology and alignment–structure relations. Then, the structure-related constraints are formulated. At the lower level, a complete alignment is optimized based on locations of dominating structures and with construction cost and geo-hazard risk as objective functions.
3. A three-stage solution method is proposed to solve the bi-level optimization model. Initially, a scanning process and specific screening operators are devised to find all possible point pairs of the dominating structures.

Then, a hierarchical MCDM analysis is conducted to evaluate the combined fitness of structures as well as to find the optimized structure combined form. Afterward, a bi-objective PSO is developed with selected point pairs as RTPs for yielding a complete optimized alignment.
4. It is found that the "structures before alignments" design process performs better than the "alignments before structures" design process in study areas with dominating landforms. More specifically, the existing optimization methods used for complex mountain railway design have difficulties in generating a better alignment than the best one provided by human designers. In contrast, the proposed model and method can find an optimized alignment with lower construction cost (i.e., 7% and 17.5% for cases 1 and 2, respectively) and geo-hazard risk (i.e., 10% for case 2) as well as shorter dominating bridges and tunnels compared to the best manually designed ones.

The limitations of this work are summarized below:

1. The alignments are optimized with a PSO method at the third optimization stage and the effectiveness of the proposed three-stage method is only compared with solutions from experienced human designers. In further studies, other computer-aided optimization methods such as a neural dynamic model (Adeli & Park, 1995), a harmony search algorithm (Siddique & Adeli, 2015), a simulated annealing (Siddique & Adeli, 2016), and a spider monkey optimization method (Akhand et al., 2020) may also be applied in the third optimization stage and compared with the human designers for further verifying the effectiveness of the proposed model and method.
2. The proposed model and method are customized for optimizing alignments in study areas with dominating landforms. In further works, they should be adapted to more cases with various characteristics for continuously validating their effectiveness in the alignment optimization field.
3. The summarized and extended human experiences derived from machine learning techniques may be integrated into the alignment optimization process. Moreover, in addition to construction cost and geology, other factors that influence alignments such as ecologic conditions and social impacts should also be considered.

## ACKNOWLEDGMENTS

This work is partially funded by the National Key R&D Program of China with award number 2021YFB2600403; National Science Foundation of China: 52078497; Science and Technology Research and Development Program Project of China railway group limited (2022-Major-20).

## REFERENCES

Adeli, H., & Park, H. S. (1995). Optimization of space structures by neural dynamics. *Neural Networks*, *8*(5), 769–781.

Akhand, M. A. H., Ayon, S. I., Shahriyar, S. A., Siddique, N., & Adeli, H. (2020). Discrete spider monkey optimization for travelling salesman problem. *Applied Soft Computing*, *86*(C), 105887.

Babapour, R., Naghdi, R., Ghajar, I., & Mortazavi, Z. (2018). Forest road profile optimization using meta-heuristic techniques. *Applied Soft Computing*, *64*, 126–137.

Casal, G., Santamarina, D., & Vázquez-Méndez, M. E. (2017). Optimization of horizontal alignment geometry in road design and reconstruction. *Transportation Research: Part C*, *74*, 261–274.

China Ministry of Railways. (2017). *Code for design of railway line. TB 10098–2017*.

de Smith, M. J. (2006). Determination of gradient and curvature constrained optimal paths. *Computer-Aided Civil and Infrastructure Engineering*, *21*(1), 24–28.

Diakoulaki, D., Mavrotas, G., & Papayannakis, L. (1995). Determining objective weights in multiple criteria problems: The critic method. *Computers & Operations Research*, *22*(7), 763–770.

Easa, S. M. (1988). Selection of roadway grades that minimize earthwork cost using linear programming. *Transportation Research Part A: General*, *22A*, 121–136.

Gao, T., Li, Z., Gao, Y., Schonfeld, P., Feng, X., Wang, Q., & He, Q. (2022). A deep reinforcement learning approach to mountain railway alignment optimization. *Computer-Aided Civil and Infrastructure Engineering*, *37*(1), 73–92.

He, L., Chao, Y., & Suzuki, K. (2008). A run-based two-scan labeling algorithm. *IEEE Transactions on Image Processing*, *17*(5), 749–756.

Hirpa, D., Hare, W., Lucet, Y., Pushak, Y., & Tesfamariam, S. (2016). A bi-objective optimization framework for three-dimensional road alignment design. *Transportation Research Part C: Emerging Technologies*, *65*, 61–78.

Jasiewicz, J., & Stepinski, T. F. (2013). Geomorphons—A pattern recognition approach to classification and mapping of landforms. *Geomorphology*, *182*, 147–156.

Jha, M. K., & Schonfeld, P. (2004). A highway alignment optimization model using geographic information systems. *Transportation Research Part A: Policy and Practice*, *38*(6), 455–481.

Jha, M. K., Schonfeld, P., & Samanta, S. (2007). Optimizing rail transit routes with genetic algorithms and geographic information system. *Journal of Urban Planning and Development*, *133*(3), 161–171.

Jin, W., Tang, L., Yang, K., Wan, G., Lü, Z., & Yu, Y. (2009). Tectonic evolution of the middle frontal area of the Longmen Mountain thrust belt, western Sichuan basin, China. *Acta Geologica Sinica*, *83*(1), 92–102.

Jong, J. C., & Schonfeld, P. (2003). An evolutionary model for simultaneously optimizing three-dimensional highway alignments. *Transportation Research Part B: Methodological*, *37*(2), 107–128.

Jong, J. C., Jha, M. K., & Schonfeld, P. (2000). Preliminary highway design with genetic algorithms and geographic information systems. *Computer-Aided Civil and Infrastructure Engineering*, *15*(4), 261–271.

Kang, M., Jha, M., & Schonfeld, P. (2012). Applicability of highway alignment optimization models. *Transportation Research Part C: Emerging Technologies*, *21*(2C), 257–286.

Kang, M., Schonfeld, P., & Yang, N. (2009). Prescreening and repairing in a genetic algorithm for highway alignment optimization. *Computer-Aided Civil and Infrastructure Engineering*, *24*(2), 109–119.

Karlson, M., Karlsson, C. S. J., MÍortberg, U., Olofsson, B., & Balfors, B. (2016). Design and evaluation of railway corridors based on spatial ecological and geological criteria. *Transportation Research Part D: Transport and Environment*, *46*, 207–228.

Kazuhiro, A. (2005). Tabu search optimization of horizontal and vertical alignments of forest roads. *Journal of Forest Research*, *10*(4), 275–284.

Kim, H. S., Schonfeld, P., Kim, E., & Jha, M. K. (2007). Highway alignment optimization incorporating bridges and tunnels. *Journal of Transportation Engineering*, *133*(2), 71–81.

Lee, Y., & Cheng, J. F. (2001). A model for calculating optimal vertical alignments of interchanges. *Transportation Research Part B: Methodological*, *35*, 1–23.

Li, W., Pu, H., Schonfeld, P., Zhang, H., & Zheng, X. (2016). Methodology for optimizing constrained 3-dimensional railway alignments in mountainous terrain. *Transportation Research Part C: Emerging Technologies*, *68*, 549–565.

Li, W., Pu, H., Schonfeld, P., Yang, J., Zhang, H., Wang, L., & Xiong, J. (2017). Mountain railway alignment optimization with bidirectional distance transform and genetic algorithm. *Computer-Aided Civil and Infrastructure Engineering*, *32*(8), 691–709.

Matsuura, T., & Aniya, M. (2012). Automated segmentation of hillslope profiles across ridges and valleys using a digital elevation model. *Geomorphology*, *177*, 167–177.

Momo, N. S., Hare, W., & Lucet, Y. (2023). Modeling side slopes in vertical alignment resource road construction using convex optimization, *Computer-Aided Civil and Infrastructure Engineering*, *38*(2), 211–224.

Mondal, S., Lucet, Y., & Hare, W. L. (2015). Optimizing horizontal alignment of roads in a specified corridor, *Computers and Operations Research*, *64*, 130–138.

Monnet, D., Hare, W., & Lucet, Y. (2020). Fast feasibility check of the multi-material vertical alignment problem in road design. *Computational Optimization & Applications*, *75*(2), 515–536.

Monnet, D., Hare, W., & Lucet, Y. (2022). A successive relaxation algorithm to solve a MILP involving piecewise linear functions with application to road design. *Computational Optimization and Applications*, *81*(3), 741–767.

Moreb, A. A. (1996). Linear programming model for finding optimal roadway grades that minimize earthwork cost, *European Journal of Operational Research*, *93*(1), 148–154.

Pradhan, A., & Mahinthakumar, G. (2013). Finding all-pairs shortest path for a large-scale transportation network using parallel Floyd-Warshall and parallel Dijkstra algorithms. *Journal of Computing in Civil Engineering*, *27*(3), 263–273.

Prasicek, G., Larsen, I. J., & Montgomery, D. R. (2015). Tectonic control on the persistence of glacially sculpted topography. *Nature Communications*, *6*, 8028.

Pu, H., Song, T., Schonfeld, P., Li, W., Zhang, H., Wang, J., Hu, J., & Peng, X. (2019). A three-dimensional distance transform for optimizing constrained mountain railway alignments. *Computer-Aided Civil and Infrastructure Engineering*, *34*(11), 972–990.

Pu, H., Wan, X., Song, T., Schonfeld, P., Li, W., & Hu, J. (2023). A geographic information model for 3-D environmental suitability analysis in railway alignment optimization. *Integrated Computer-Aided Engineering*, *30*(1), 67–88.

Pu, H., Xie, J., Schonfeld, P., Song, T., Li, W., Wang, J., & Hu, J. (2021). Railway alignment optimization in mountainous regions considering spatial geological hazards: A sustainable safety perspective. *Sustainability*, *13*(4), 1661.

Pushak, Y., Hare, W., & Lucet, Y. (2016). Multiple-path selection for new highway alignments using discrete algorithms. *European Journal of Operational Research*, *248*(2), 415–427.

Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, *1*, 83–98.

Sadek, S., Bedran, M., & Kaysi, I. (1999). GIS platform for multicriteria evaluation of route alignments. *Journal of Transportation Engineering*, *125*(2), 144–151.

Samanta, S., & Jha, M. K. (2012). Applicability of genetic and ant algorithms in highway alignment and rail transit station location optimization. *International Journal of Operations Research and Information Systems (IJORIS)*, *3*(1), 13–36.

Shafahi, Y., & Bagherian, M. (2013). A customized particle swarm method to solve highway alignment optimization problem. *Computer-Aided Civil and Infrastructure Engineering*, *28*(1), 52–67.

Siddique, N., & Adeli, H. (2015). Harmony search algorithm and its variants, *International Journal of Pattern Recognition & Artificial Intelligence*, *29*(8), 1.

Siddique, N., & Adeli, H. (2016). Simulated annealing, its variants and engineering applications. *International Journal On Artificial Intelligence Tools*, *25*(6), 1.

Song, T., Pu, H., Schonfeld, P., & Hu, J. (2022). Railway alignment optimization under uncertainty with a minimax robust method. *IEEE Intelligent Transportation Systems Magazine*, *15*(1), 2–15.

Song, T., Pu, H., Schonfeld, P., Liang, Z., Zhang, M., Hu, J., Zhou, Y., & Xu, Z. (2022). Mountain railway alignment optimization integrating layouts of large-scale auxiliary construction projects. *Computer-aided Civil and Infrastructure Engineering*, *38*(4), 433–453. https://doi.org/10.1111/mice.12839

Song, T., Pu, H., Schonfeld, P., Zhang, H., Li, W., Hu, J., & Wang, J. (2020). Mountain railway alignment optimization considering geological impacts: A cost-hazard bi-objective model. *Computer-Aided Civil and Infrastructure Engineering*, *35*(12), 1365–1386.

Song, T., Pu, H., Schonfeld, P., Zhang, H., Li, W., Peng, X., Hu, J., & Liu, W. (2021). GIS-based multi-criteria railway design with spatial environmental considerations. *Applied Geography*, *131*, 102449.

Song, T., Schonfeld, P., & Pu, H. (2023). A review of recent alignment optimization research for roads, railways and rail transit lines. *IEEE Transactions on Intelligent Transportation Systems*, *24*(5), 4738–4757. https://doi.org/10.1109/TITS.2023.3235685

Sushma, M. B., & Maji, A. (2020). A modified motion planning algorithm for horizontal highway alignment development. *Computer Aided Civil and Infrastructure Engineering*, *35*(8), 818–831.

Sushma, M. B., Roy, S., & Maji, A. (2022). Exploring and exploiting ant colony optimization algorithm for vertical highway alignment development. *Computer-Aided Civil and Infrastructure Engineering*, *37*(12), 1582–1601.

Vázquez-Méndez, M. E., Casal, G., Castro, A., & Santamarina, D. (2021). An algorithm for random generation of admissible horizontal alignments for optimum layout design. *Computer-Aided Civil and Infrastructure Engineering*, *36*(8), 1056–1072.

Vázquez-Méndez, M. E., Casal, G., Santamarina, D., & Castro, A. (2018). A 3D model for optimizing infrastructure costs in road design. *Computer-Aided Civil and Infrastructure Engineering*, *33*(5), 423–439.

Williams, R. M. E., & Phillips, R. J. (2001). Morphometric measurements of Martian valley networks from Mars Orbiter Laser Altimeter (MOLA) data. *Journal of Geophysical Research: Planets*, *106*(E10), 23737–23751.

Wu, K., Otoo, E., & Suzuki, K. (2009). Optimizing two-pass connected-component labeling algorithms. *Pattern Analysis and Applications*, *12*(2), 117–135.

Yang, D., He, Q., & Yi, S. (2020). Underground metro interstation horizontal-alignment optimization with an augmented rapidly exploring random-tree connect algorithm. *Journal of Transportation Engineering Part A Systems*, *146*(11), 04020129.

Zhang, H., Pu, H., Li, W., Song, T., Wang, J., Hu, J., & Peng, X. (2019). Multi-objective optimization of railway alignment with distance transform and particle swarm optimization. *International Conference on Computers & Industrial Engineering*.

Zhang, H., Pu, H., Schonfeld, P., Song, T., Li, W., & Hu, J. (2021). Railway alignment optimization considering lifecycle costs. *IEEE Intelligent Transportation Systems Magazine*, *14*(5), 22–40. https://doi.org/10.1109/MITS.2021.3071032

**How to cite this article:** Wan, X., Pu, H., Schonfeld, P., Song, T., Li, W., Peng, L., Hu, J., & Zhang, M. (2024). Mountain railway alignment optimization based on landform recognition and presetting of dominating structures. *Computer-Aided Civil and Infrastructure Engineering*, *39,* 242–263. https://doi.org/10.1111/mice.13073