Computing the trajectories for the development of optimal routes

M. Fawad Zazai¹ · Armin R. Fügenschuh²

Received: 10 December 2019 / Revised: 19 September 2020 / Accepted: 23 September 2020 /

Published online: 15 October 2020

© The Author(s) 2020

Abstract

Planning the construction of new transport routes or power lines on terrain is usually carried out manually by engineers, with no guarantee of optimality. We introduce a new approach for the computation of an optimal trajectory for the construction of new transit routes and power lines between two locations on a submanifold $U \subset \mathbb{R}^3$ representing the topography of a terrain. U is approximatively modeled by a special weighted grid. On this grid, the shortest paths for the construction of new routes are determined, whereby we consider three optimization criteria: routes with minimum distance, routes with lowest construction costs and routes with minimum absolute altitude variations or minimum absolute gradients. Subsequently, a combination of these criteria is used to expand this problem into a multi-criteria optimization problem. A shortest path algorithm, such as the Dijkstra algorithm, is used to compute optimal compromises for the construction of new routes.

Keywords Shortest path · Graph construction · GIS · topography · Multi-criteria optimization (Pareto frontier) · Trajectory planning

1 Introduction

Worldwide, many developing countries have an undeveloped technical and logistical infrastructure. This not only causes numerous problems for the inhabitants in their daily life, but also has negative economic consequences (more on this can be found in Fay and Yepes (2003), Gaal (2017), Gurara et al. (2017), Oyedele (2016)).

✉ Armin R. Fügenschuh
fuegenschuh@b-tu.de

M. Fawad Zazai
zazaif@hsu-hh.de

¹ Helmut Schmidt University/University of the Federal Armed Forces Hamburg, Holstenhofweg 85, 22043 Hamburg, Germany

² Brandenburgische Technische Universität Cottbus-Senftenberg, Platz der Deutschen Einheit 1, 03034 Cottbus, Germany

Classical examples of a logistical infrastructure are transport networks such as roads, tracks, and waterways. On the technical side, these are supply networks such as power lines, freshwater and wastewater lines, and communication networks. An immature technical infrastructure can severely curb the energy and transport sector. These containments can not only increase the manufacturing and transport costs of a product but can also impact entire construction projects, transport networks, and economic networks. The resulting consequences can reduce the development and progress of these countries. Besides, investors lose interest in investing in under-developed countries with a high amount of natural resources since the risk of a misinvestment is too high. Infrastructure projects are usually very costly. Planning firms or consulting companies often plan the development of new routes (alignment) based on empirical values or using estimation methods. However, these methods do not guarantee optimality. Such planning projects and feasibility studies can not only cause high costs but also be very time-consuming. In the case of extensive studies in risky areas, the presence of the staff on-site can pose a threat to the safety of the personnel. Risky zones can include regions with natural disasters or conflict zones. Last but not least, manual planning involves a residual risk of corruption and manipulation, as a manual planned route allows for unnecessary detours. By planning and constructing redundant detours, companies can generate more turnover and use part of the turnover to influence the transport ministries or the relevant authorities.

The Shortest Path Problem is a well-known graph theory problem where the goal is to find a shortest path between two nodes (start and end points) within a weighted graph $G = (V, E, w)$, where V is a set of nodes, E is a set of edges, and $w : E \rightarrow \mathbb{R}_+$ is a cost function that defines the weight $w_{i,j}$ of an edge $\{i, j\} \in E$. A shortest path between two nodes is a route with a minimum sum of all used edge weights to get from the start node to the destination node. Shortest path algorithms are particularly used for path planning or finding a path. Route planning is often used for autonomous mobile robots to find an optimal route from a starting point to a destination in complex environments. There are different heuristic and non-heuristic approaches for finding a way to calculate an ideal path on a terrain. One approach is that a graph is given and a search algorithm for a path is performed on this graph. Another approach is that a certain search algorithm does not search within a given graph, but simultaneously creates a graph in the direction of the target on which a potential optimal path could lie while searching for a path. Mitchell and Sharir present in Mitchell and Sharir (2004) an efficient algorithm that calculates Euclidean shortest paths in three dimensions with obstacles, in polynomial time. An L_1 shortest path on polyhedron terrain, L_2 shortest path on parallel walls, and L_2 shortest path on a land-like, layered, and axis-aligned rectangles are analyzed and computed.

Belaidi et al. describe in Belaidi et al. (2014) an algorithm for mobile robots in a polygonal grid that plans a path to cross a given terrain with irregularities and obstacles. The objective function is, on the one hand, the distance and on the other hand the speed crossing an area. Saranya et al. modify the D^* algorithm in Saranya et al. (2016) by defining a cost function that depends on the distance and slope of the terrain. This allows an autonomous mobile robot to calculate an optimal route between two points in an unknown environment, such as an irregular terrain. In Muñoz et al. (2017), Muñoz et al. present another algorithm, the

so-called “3D Accurate Navigation Algorithm (3Dana)”, which finds the most convenient path for mobile robots on a terrain (DTM—Digital Terrain Model) by considering, for example, the maximum gradient of a terrain that can be driven on. Subsequently, they tested this algorithm on the digital topography of Mars. The geometry of a terrain influences the trajectory of a potential route. Fowler and Little show in Fowler and Little (1979) that they model a terrain with a Triangulated Irregular Network (TIN). In this work they describe a method to automatically extract a TIN model from dense raster data. A first approximation is constructed by automatically triangulating a set of feature points derived from the grid model. Basu and Snoeyink describe in Basu and Snoeyink (2007) how RTINs (rectangular, irregular networks) can be used to represent terrain as a mosaic of polynomial surface areas.

Compared to the above papers, the main focus of our research is not the development of an algorithm for finding a shortest path, but the development of algorithms for a discrete modeling of a Riemannian submanifold $U \subset \mathbb{R}^3$ through a special grid $\Gamma = (V, E, W)$ of nodes V , edges E and weights W , where U is a topographic terrain with obstacles. On Γ a shortest path for the expansion of transport infrastructure on U is searched.

A regular grid Γ is constructed on a geographical area A by mapping Γ to the topography U of A and representing U approximatively. A family of grids $(\Gamma_a, \Gamma_b, \Gamma_c)$ with different weights (W_a, W_b, W_c) characterized by certain criteria is calculated. Each edge $\{i, j\}$ represents a segment of a route, and the non-negative weight $w_{i,j}$ of that edge indicates (a) the length, (b) the construction cost, or (c) the absolute elevation change between nodes i and j . For the calculation and evaluation of the edge weights (W_a, W_b, W_c) different algorithms were developed. For these algorithms different data are used, e.g., the soil condition, the elevation data of a terrain (topography) and the construction costs per unit length. The resulting routes can be either transport routes or transmission lines connecting two locations or running from one starting point through several locations to a destination. The aim is to construct the shortest route in terms of distance, the most cost-effective route, and the most convenient route with minimum vertical fluctuation. Compared to the above approaches, we will not use heuristics to search for an optimal route but solve the problem with a search algorithm, such as the Dijkstra (1959), which guarantees a global optimum. Through the Pareto method, these different grids are combined into multi-criteria grids $\Gamma_{\text{Pareto}} = (\Gamma_1^{\text{Pareto}}, \dots, \Gamma_n^{\text{Pareto}})$ so that a family of n grids is generated based on compromises between these criteria (Ehrgott 2005). On each multi-criteria grid Γ_k^{Pareto} of Γ_{Pareto} an optimal route is determined. These solutions represent a set of optimal trade-offs to design ideal transport and energy infrastructures. The presented work is based on Zazai (2020) and is a largely extended version of Zazai and Fügenschuh (2018).

The advantages and differences of our method compared to the above approaches in the papers are as follows:

1. We present a new method by planning infrastructure development with discrete optimization and graph theory, considering different construction criteria. This

approach is different from the approach of autonomous mobile robots, where a vehicle moves from its starting point to its destination in an environment. The environment remains invariant while an autonomous mobile robot crosses it, but when a route is laid out, the environment may change as a result of construction work. This means that certain topographical outliers can be tolerated because they are removed when a route is developed.

2. Although the TIN (Triangulation) method is a good approximation for modeling a terrain, it does not guarantee efficiency for calculating routes on that terrain. Especially when the geometry of a terrain is only a secondary consideration, such as when designing the most cost-effective route, where the ground conditions must be taken into account.
3. Unlike the path-finding of autonomous mobile robots, the edges of a grid Γ can consist of different infrastructure segments, such as roads, tracks, pipelines or power lines. For the selection of a specific infrastructure segment, the input data is modified and adjusted, such as the average cost or the structure of the ground.
4. On Γ no locally, but globally optimal routes are computed.
5. Once a grid Γ of a topographic submanifold U is computed, it is possible to calculate between any start and destination point on U an optimal path with a running time of $\mathcal{O}(|V|\log(|V|) + |E|)$ Turau (2004). For each changed start and destination point on U , the cost for the computation of the edge weights W is being saved.
6. All planning results, such as trajectory coordinates, construction costs, distances, and total absolute altitude changes, are given separately for each optimal route.
7. The Pareto method is used to define multi-criteria grids by using the different grids Γ_a , Γ_b , and Γ_c . The results can model an environment more realistically. These grids are used to determine various optimal trajectories with the corresponding results, which represent optimal compromises between the mentioned criteria.

2 Topographical grid generation

In this section, we construct a special grid Γ on a spheroid S . For this we take an area U' from S , which is curved. This curved area has to be mapped on a Euclidean plane in $\mathbb{R}^3$ by a homeomorphism. Thus the construction of Γ in the Euclidean space will be simplified. A spheroid S can be described by an implicit representation

$$\frac{x^2 + y^2}{a^2} + \frac{z^2}{c^2} = 1, \quad (1)$$

or a parameter representation

$$\begin{pmatrix} a \cos(\alpha) \cos(\beta) \\ a \cos(\alpha) \sin(\beta) \\ c \sin(\alpha) \end{pmatrix} \quad (2)$$

where the numbers a and c are the half-axes of the rotating half-ellipse, $-\frac{\pi}{2} \leq \alpha \leq \frac{\pi}{2}$, $0 \leq \beta \leq 2\pi$ and $a, c > 0$.

The shape of the earth is spheroidal which defines an equivalent manifold to S . In the following, S represents the surface of the earth. Since the earth is represented by a flattened spheroid, we set $a > c$. A submanifold U' of S is mapped to a submanifold U of $\mathbb{R}^3$ by a homeomorphism $\varphi_{U'}$ so that U is a Euclidean representation of the curved surface U' in $\mathbb{R}^3$. It is assumed that U describes the topography of a geographical area $A \subset \{(x, y, 0) \in \mathbb{R}^3\}$ on Earth. So the surface of U can have irregular altitudes and depths. The topography of an area is collected from discrete elevation data which is given by a matrix H . A rectangular weighted grid Γ of nodes and edges is constructed on a geographical area A , so that $\Gamma(A)$ defines a geographical grid of A . $\Gamma(A)$ consists of rectangular grid cells that have vertices and borderlines in the plane in $\mathbb{R}^3$. The boundary of each cell consists of four borderlines. The upper and lower borderlines consist of $m \in \mathbb{N}$ equidistant edges and the left and right borderlines consist of $k \in \mathbb{N}$ equidistant edges (see Fig. 1). A cell of Γ has $2m + 2k$ associated nodes on the borderlines. All node pairs of these cells (except those on the same borderline) are connected by edges. These edges are added to the set E_A . A grid Γ consists of $X \cdot Y$ cells, where Y cells in Γ are stacked vertically and in every i -th row X cells are placed horizontally next to each other.

A grid with rectangular cells has the advantage that it is a structured grid. The cells are regularly sorted (parqueted) so that they lie next to each other in a grid. This allows the nodes of the cells to be easily indexed and counted one after the other. All node pairs of a rectangular cell of Γ with $2k$ vertical and $2m$ horizontal segments are connected by $m^2 + 4km + k^2$ edges. On the surface of U , a grid $\Gamma(U) = (V_U, E_U)$ is constructed from nodes V_U and edges E_U , which approximately models the topography of A . Next, the topographic grid is to be extended with graphs from altitude contours that lie on local areas of submanifold U , such as the contour line of a mountain or the altitude edges of a canyon. These contours improve the modeling quality of terrain. Since the elevation data H are discrete, only discrete contours can be generated from these data.

Calculating these contours for each altitude $h_i \in h = \{h_1, \dots, h_k\}$ of U the Marching Squares algorithm (see Ho et al. 2005; Hanisch 2004) is used, where k is the number of contours for different altitudes from H . Consequently, the result is a set of edge lines in different altitudes of h . Each contour consists of nodes and

Fig. 1 On the right a grid Γ consists of $X \cdot Y = 4 \cdot 3 = 12$ cells and on the left a cell of Γ with $m = 2$ and $k = 2$ decomposition consisting of $2m + 2k = 8$ nodes and $m^2 + 4km + k^2 = 24$ edges

edges. This algorithm generates the sets of contours $\gamma(h_1), \dots, \gamma(h_k)$ where $\gamma(h_i) = \{\gamma_1(h_i), \dots, \gamma_{s_i}(h_i)\}$ is a set of discrete contours on U at the level h_i for $i \in \{1, \dots, k\}$. Each contour defines a path of edges on U . The contours are required later for the computation of a mountain road. These contours which are adjacent and lie on top of each other are interlinked by additional edges, so that one has a transition between the contours in different levels. As an illustration, two-layered contours with level h_i and h_{i+1} are considered which are connected as follows: Let $\gamma(h_i) = (V_i, E_i)$ and $\gamma(h_{i+1}) = (V_{i+1}, E_{i+1})$ be two adjacent contours lying one above the other that generate a graph through nodes V_i, V_{i+1} and edges E_i, E_{i+1} in $\mathbb{R}^3$. For each node $v_j \in V_i$ the next neighboring node is determined from V_{i+1} . Let p_j^i and p_m^{i+1} be points in $\mathbb{R}^3$ representing the nodes v_j and $v_m \in V_{i+1}$. The nearest neighboring nodes of V_i are calculated as nodes $V_u := \{u_1, \dots, u_{|V_i|}\} \subset V_{i+1}$ with minimal distances:

$$\forall v_j \in V_i : q_j^{i+1} = p_m^{i+1}, \text{ if } \text{dist}(p_j^i, p_m^{i+1}) \leq \text{dist}(p_j^i, p_s^{i+1}), \quad (3)$$

for all $p_s^{i+1} \in V_{i+1}$, where q_j^{i+1} is a point representing a node $u_j \in V_u$ and dist is a distance function between any two points in $\mathbb{R}^3$. Each node v_j is connected to its nearest neighboring node u_j through an additional edge $e_j = \{v_j, u_j\}$. Repeat this step for all contours except the second last. This generates a mesh between neighboring contours that defines an unstructured grid. Figure 2 shows two neighboring contours $\gamma(h_i)$ and $\gamma(h_{i+1})$ where the red edges connect all nodes of V_i to the nodes of V_{i+1} . Some nodes of V_{i+1} located far away have no direct connections to the nodes of V_i . The crosshatched red edges make a connection to the next nodes of the neighboring contours.

In the last step, these contours are connected to the grid $\Gamma(U) = (V_U, E_U)$ so that a transition between them is possible. To connect $\Gamma(U)$ with the set of contours $\gamma(h) = \gamma(h_1) \cup \dots \cup \gamma(h_k)$ with $\gamma(h_i) = \{\gamma_1(h_i), \dots, \gamma_{s_i}(h_i)\}$ for $i \in \{1, \dots, k\}$, edges with minimal distance are searched, which connect these two sets. As far as possible, all nodes of the contours should be connected with V_U . Seven cases are considered for this type of meshing to keep the number of connections to a minimum except for the most necessary: Let $\gamma(h_{i-1})$, $\gamma(h_i)$ and $\gamma(h_{i+1})$ be three contours lying on top of each other.

Fig. 2 Two neighboring contours $\gamma(h_i)$ and $\gamma(h_{i+1})$ in black for $i \geq 1$ interconnected by red edges

- (a) If nodes of V_U exist between these three contours, then all nodes of $\gamma(h_i)$ are connected to the nearest nodes of V_U that lie between $\gamma(h_{i-1})$ and $\gamma(h_{i+1})$. This means that each node of $\gamma(h_i)$ is connected to two nodes of V_U with one node being above $\gamma(h_i)$ and the other node being below $\gamma(h_i)$.
- (b) If no nodes of V_U exist between the two contours $\gamma(h_{i-1})$ and $\gamma(h_{i+1})$, then the middle contour $\gamma(h_i)$ has no direct connection to V_U .
- (c) If there are nodes of V_U between $\gamma(h_{i-1})$ and $\gamma(h_i)$, then there are direct connections from $\gamma(h_i)$ to V_U , so that each node of $\gamma(h_i)$ is connected to the nearest node of V_U .
- (d) If there are nodes between $\gamma(h_i)$ and $\gamma(h_{i+1})$, then the nodes of $\gamma(h_i)$ are connected to the nearest nodes of V_U above $\gamma(h_i)$.
- (e) If $\gamma(h_{i-1}) = \emptyset$, then $\gamma(h_i)$ is locally the lowest contour. Each node of $\gamma(h_i)$ is connected to the nearest nodes of V_U below $\gamma(h_i)$ and between $\gamma(h_i)$ and $\gamma(h_{i+1})$ (if such a node exists).
- (f) If $\gamma(h_{i+1}) = \emptyset$, then $\gamma(h_i)$ is locally the highest contour. Each node of $\gamma(h_i)$ is connected to the nearest nodes of V_U above $\gamma(h_i)$ and between $\gamma(h_i)$ and $\gamma(h_{i+1})$ (if such a node exists).
- (g) If $\gamma(h_{i-1}) = \gamma(h_{i+1}) = \emptyset$, then each node of $\gamma(h_i)$ is connected to the nearest nodes of V_U below and above $\gamma(h_i)$.

Figure 3 shows as an example three discrete height contours of U , which on the one hand are connected by dashed red edges and on the other hand these contours are connected to the (red) nodes of the grid $\Gamma(U) = (V_U, E_U)$ by orange edges. The dashed blue connections are the edges that connect the nodes of V_U .

The minimum distance from $v_j \in V_i$ to V_U can be determined for the above seven cases as follows: Let p_j^i be an associated point of $v_j \in V_i$,

$$V_i^+ := \begin{cases} \{v_u \in V_U \mid v_u \text{ is between } \gamma(h_i) \text{ and } \gamma(h_{i+1})\}, & \text{if } \gamma(h_{i+1}) \neq \emptyset, \\ \{v_u \in V_U \mid v_u \text{ is above } \gamma(h_i)\}, & \text{if } \gamma(h_{i+1}) = \emptyset, \end{cases} \quad (4)$$

and

Fig. 3 The discrete elevation contours $\gamma(h_{i-1})$, $\gamma(h_i)$ and $\gamma(h_{i+1})$ of U are connected by (red) edges to the (red) nodes of the grid $\Gamma(U)$

$$V_i^- := \begin{cases} \{v_u \in V_U \mid v_u \text{ is between } \gamma(h_i) \text{ and } \gamma(h_{i-1})\}, & \text{if } \gamma(h_{i-1}) \neq \emptyset, \\ \{v_u \in V_U \mid v_u \text{ is below } \gamma(h_i)\}, & \text{if } \gamma(h_{i-1}) = \emptyset, \end{cases} \quad (5)$$

two subsets of V_U and $q_1, \dots, q_{|V_i^+ \cup V_i^-|}$ points in $\mathbb{R}^3$ representing nodes of these subsets. To minimize the distance of each point p_j^i of the node $v_j \in V_i$ the following applies

$$\sum_{k=1}^{|V_i^+ \cup V_i^-|} \mu_k \text{dist}(p_j^i, q_k) \longrightarrow \min!, \quad \forall \mu_k \in \{0, 1\}, \quad (6)$$

with the auxiliary condition

$$\sum_{k=1}^{|V_i^+ \cup V_i^-|} \mu_k = \begin{cases} 1, & \text{if } V_i^+ \neq V_i^- = \emptyset \text{ or } V_i^- \neq V_i^+ = \emptyset, \\ 2, & \text{if } V_i^+ \neq \emptyset \text{ and } V_i^- \neq \emptyset, \\ 0, & \text{if } V_i^+ = V_i^- = \emptyset. \end{cases}$$

For each $v_j \in V_i$ with $\mu_m = 1$ and $m \in \{1, \dots, |V_i^+ \cup V_i^-|\}$ an edge is generated with minimal distance from $v_j \in V_i$ to a node $q_m \in V_U$. Repeat this algorithm for each node of the contours of U until they have meshed with the grid $\Gamma(U)$. This mesh, together with the mesh of all neighboring and superimposed contours of the set $\gamma(h)$, generates an unstructured mesh $\Gamma(\gamma) = (V_\gamma, E_\gamma)$ consisting of nodes V_γ and edges E_γ . Since $\Gamma(U)$ is structured, together with $\Gamma(\gamma)$ it forms a hybrid grid

$$\Gamma_{\mathbb{R}^3}(U) = \Gamma(U) \cup \Gamma(\gamma) = (V_U \cup V_\gamma, E_U \cup E_\gamma). \quad (7)$$

This grid is used in the next section to calculate optimal routes or shortest paths to the surface of U .

3 Optimal routes

The construction of an optimal route on a topographical area U is based on the computation of a shortest path. To solve this problem, the routes are discretized to a grid Γ . The edge weights of Γ are non-negative values, which are determined by a cost function $w : E \rightarrow \mathbb{R}_{\geq 0}, e_{i,j} \mapsto w_{i,j}$. Since these values are always non-negative, algorithms such as the Dijkstra (1959) can be used to compute a global optimal route that represents the shortest, most cost-effective, or most convenient route (minimum absolute elevation changes along the route). The higher the number of cells is chosen at a constant size of Γ , the more accurately the routes are determined with simultaneously higher computing time.

Obstacles on U can be modeled by finitely many discrete polygons $O_1, \dots, O_m$. If a trajectory is not allowed to cross an obstacle, then it must first be checked whether an edge $e_{i,j}$ intersects one of the polygons $O_1, \dots, O_m$. If $e_{i,j}$ has at least one intersection with these polygons, then the weight of this edge is set as $w_{i,j} = +\infty$. If $e_{i,j}$ touches these polygons at the boundary or has no intersection point, then the weight of this edge

is given by the equation (12). If a set of obstacles are $O := \{O_u, \dots, Q_k\}$ for $u \geq 1$ and $k \leq m$ lies between a starting node and a destination node then a shortest route is computed around $O_u, \dots, Q_k$.

The geographical coordinates on earth are usually given in latitude and longitude coordinates. The latitude coordinate is the geographical width measured from the equator to the north and south. The latitude values range from -90° at the south pole over 0° at the equator to $+90^\circ$ at the north pole. The longitude coordinate is the geographical length going from the zero meridian 0° to 180° eastwards and from 0° to -180° westwards.

3.1 Shortest routes

Let S be a spheroid and U_S a submanifold of S , where S has the size of the earth and the subset U_S has a regular non-topographic surface. A submanifold U is defined as a connected topographical area on earth which represents the submanifold U_S in $\mathbb{R}^3$ as a topographical surface. The shortest routes with respect to a distance are routes within a weighted topographic grid $\Gamma_{\mathbb{R}^3}(U) = (V, E, W_D)$, where the weights W_D indicate the distances of the edges of E_U . First W_D is computed and then the weights of W_D are assigned to the corresponding edges of the grid $\Gamma_{\mathbb{R}^3}(U)$. For U_S and U the latitude and longitude coordinate systems are used. The distances W_D of the edges of $\Gamma_{\mathbb{R}^3}(U)$ are computed with the formulas of Vincenty for spheroids (Vincenty 1975) since U is an approximation of U_S except for the topography. Let

$$\begin{aligned} \text{dist}: U \times U &\longrightarrow \mathbb{R}_{\geq 0}, \\ (p_i, p_j) &\longmapsto d_{ij} \end{aligned} \quad (8)$$

be a function that gives a distance between two points p_i and p_j of U according to Vincenty formulas for distances on U_S . Due to the topological conditions, the edges of a topographic grid $\Gamma_{\mathbb{R}^3}(U)$ have different slopes. Usually, traffic routes may not exceed a certain slope otherwise, the crossing of this route is not possible or the consumption of a vehicle increases significantly. Therefore, when building a route, a maximum gradient angle is always taken into account. An edge $e_{ij} = \{v_i, v_j\}$ is decomposed in $u_{ij} \in \mathbb{N}$ segments $e_{ij}(k) = \{v_{ij}(k-1), v_{ij}(k)\}$ for $k = 1, \dots, u_{ij}$ where $v_{ij}(k)$ for $k = 1 \dots u_{ij} - 1$ are elements of a set of auxiliary nodes V_{aux} lying on the edge e_{ij} (see Fig. 4). For the first and last auxiliary node of an edge e_{ij} is

$$v_{ij}(k) = \begin{cases} v_i \in V, & \text{if } k = 0, \\ v_j \in V, & \text{if } k = u_{ij}. \end{cases} \quad (9)$$

Fig. 4 An edge $e_{ij} = \{v_i, v_j\}$ is divided into the segments $e_{ij}(1), \dots, e_{ij}(u_{ij})$ with the auxiliary nodes $v_{ij}(1), \dots, v_{ij}(u_{ij}-1)$

These nodes generate between v_i and v_j the segments $e_{ij}(1), \dots, e_{ij}(u_{ij})$ on the surface of U . Let $\varphi_{\max} \in [0, 90]$ be a maximum allowed angle of a segment and $\varphi_{ij}(k) \in [0, 90]$ the angle of a segment $e_{ij}(k)$ with respect to a horizontal plane. For each node $v_{ij} \in V \cup V_{\text{aux}}$ there is a map $h : V \cup V_{\text{aux}} \rightarrow \mathbb{R}$ so that $h(v_{ij})$ gives the altitude of v_{ij} on an area $A \subset \{(x, y, 0) \in \mathbb{R}^3\}$ where A has the altitude 0. The absolute elevation change between two neighboring auxiliary nodes $v_{ij}(k-1)$ and $v_{ij}(k)$ of $V \cup V_{\text{aux}}$ is defined as

$$\Delta h_{ij}(k) := |h(v_{ij}(k)) - h(v_{ij}(k-1))|, \quad (10)$$

for $k = 1, \dots, u_{ij}$. A segment $e_{ij}(k)$ is called **critical** if this segment has at least one intersection with O so that this segment is either entirely or partially in O . If $e_{ij}(k)$ is a critical segment then the edge e_{ij} is also critical. All critical segments on $\Gamma_{\mathbb{R}^3}(U)$ with respect to O define a set E_{crit} . Each segment $e_{ij}(k)$ is assigned a distance $d'_{ij}(k)$: Let $\pi : U \rightarrow A$ be an orthogonal projection of the typographic area U on the area A . With Vincenty's formulas, first the length

$$d'_{ij}(k) := \text{dist}(\pi(p_{ij}(k)), \pi(p_{ij}(k-1))) \quad (11)$$

of the segment $e_{ij}(k)$ is calculated where the points $p_{ij}(k-1)$ and $p_{ij}(k)$ represent the nodes $v_{ij}(k-1)$ and $v_{ij}(k)$ for $k = 1, \dots, u_{ij}$ in $\mathbb{R}^3$. Finally, the length of a segment is computed by considering a absolute slope angle or maximum slope angle that can cause an additional lengthening:

$$d_{ij}(k) = \begin{cases} \frac{d'_{ij}(k)}{\cos(\varphi_{ij}(k))}, & \text{if } \varphi_{ij}(k) \leq \varphi_{\max} \text{ and } e_{ij}(k) \notin E_{\text{crit}}, \\ \frac{\Delta h_{ij}(k)}{\sin(\varphi_{\max})}, & \text{if } \varphi_{ij}(k) > \varphi_{\max} \text{ and } e_{ij}(k) \notin E_{\text{crit}}, \\ N, & \text{if } e_{ij}(k) \in E_{\text{crit}}, \end{cases} \quad (12)$$

for all $k \in \{1, \dots, u_{ij}\}$, $\varphi_{\max} > 0$ and $N \in \mathbb{R}_+$. Serpentine are mostly built in mountain slopes to be able to go up or down steep areas. The second entry in equation 12 describes the segment $e_{ij}(k)$ as a serpentine with an maximal slope angle $\varphi_{\max}$. That is, if this segment exceeds the angle $\varphi_{\max}$, then this segment is assigned the length of a serpentine which has a maximum slope angle $\varphi_{\max}$. A serpentine is due to its shape longer than a straight segment. If a segment $e_{ij}(k)$ is within one of the polygons of O or it intersects with O then the value N in the third entry of equation 12 is the weight of this segment. The polygons of O are usually unfavorable areas in which the construction of a route is unsuitable. If these areas are forbidden then set $N = +\infty$. If these areas are not prohibited but unsuitable for construction or crossing, then $N < \infty$ is a sufficiently large positive number. For the computation of a shortest route regarding the construction of a start node v and destination node t , the following objective function has to be minimized:

$$\sum_{e_{ij} \in E} x_{e_{ij}} \sum_{k \in I_{ij}} d_{ij}(k) \longrightarrow \min!, \quad (13)$$

for $I_{ij} = \{1, \dots, u_{ij}\}$ such that for all $v \in V$:

$$\sum_{e_{ij} \in \delta^-(v)} x_{e_{ij}} - \sum_{e_{ij} \in \delta^+(v)} x_{e_{ij}} = \begin{cases} -1, & \text{if } v = s, \\ 1, & \text{if } v = t, \\ 0, & \text{else,} \end{cases}$$

$$\forall e_{ij} \in E : x_{e_{ij}} \geq 0.$$

3.2 Cheapest routes

The construction of a route can depend on several factors, such as the construction price, the ground conditions on which the route is to be built, and personnel costs. In this paper, the weights of a grid $\Gamma_{\mathbb{R}^3}(U) = (V, E, W_C)$ are the average construction costs W_C of the edges E_U . The cost of building a route per length unit is made up of the average construction cost per length unit and the ground condition. The objective is to compute a route between two nodes on $\Gamma_{\mathbb{R}^3}(U)$ that can be built most cost-effectively. The soil condition can be described by a set of discrete polygons $P = \{P_1, \dots, P_n\}$ on U where each $P_i = (V_i, E_i)$ consists of nodes K_i and edges E_i . Each polygon is assigned a color that describes the structure of the ground. For example, if the color of a polygon is j then this polygon represents a ground structure s_j . All polygons of P with a color j create a polygon class $G_j \subset P$. Each polygon of P is assigned to exactly one polygon class. All polygons of P are classified in $k \leq n$ different subsets so that $P = G_1 \cup \dots \cup G_k$ consists of k disjunctive polygon classes. Each polygon class G_j is assigned a surface structure value $\alpha_j > 0$. If a route passes through G_j , this value is the multiple of the construction costs of this route. This means if a route is built over an area P_u of G_j the construction costs on this area increase by the factor α_j . Let $e_{i,l}$ be an edge of s segments such that the edge $e_{i,l} = (e_{i,l}(1), \dots, e_{i,l}(s_{i,l}))$ can be decomposed into the segments $e_{i,l}(1), \dots, e_{i,l}(s_{i,l})$. The cost $c_{i,l}$ of an edge $e_{i,l}$ can be expressed as the sum of all incurred costs of the segments

$$c_{i,l} = C \sum_{k \in K_{i,l}, j \in J_{i,l}} d_k \cdot \alpha_j, \quad (14)$$

where $K_{i,l} = \{1, \dots, s_{i,l}\}$ is the index set of the segments, $J_{i,l}$ is the index set of the occurring surface structures on an edge $e_{i,l}$, C is the average costs for the construction and maintenance of a route per length unit, d_k is the length of the k -th segment of an edge $e_{i,l}$. To identify a segment $e_{i,l}(k)$ on G_j , an auxiliary node m_k is generated in the middle of this segment. This point is checked by the Point-in-Polygon algorithm (Sunday 2001) to see if it lies in a closed polygon P_u of P . If not, this procedure is continued for the next polygons until the result is positive. If m_k is in P_u with

the color j , the segment $e_{i,l}(k)$ is assigned the value α_j . Calculating the length d_k of a segment $e_{i,l}(k)$ lead to one of the following cases:

1. If there are no intersections and the auxiliary node m_k is not in P_u , then

$$d_k = 0.$$

2. If there are no intersections and the auxiliary node m_k lies in P_u , then the following is applied with $k = 1$

$$d_k = \text{dist}(v_i, v_l).$$

3. If there is exactly one intersection s_1 with $e_{i,l}(k) = \{v_i, s_1\}$, then

$$d_k = \text{dist}(v_i, s_1).$$

4. If there is exactly one intersection s_1 with $e_{i,l}(k) = \{s_1, v_l\}$, then

$$d_k = \text{dist}(s_1, v_l).$$

5. If there are two intersections s_1 and s_2 with $e_{i,l}(k) = \{s_1, s_2\}$, then

$$d_k = \text{dist}(s_1, s_2).$$

If a segment $e_{i,l}(k)$ lies in a polygon P_u with the color j , then the segment length d_k is multiplied by the corresponding surface structure value α_j . Note that the distances $\{d_k\}$ have to be calculated analogously to the equation (12). Figure 5 shows an edge $e_{i,l}$ lying on the three polygons P_1 in red, P_2 in green and P_3 in blue. It has two intersections s_1 and s_2 . The first intersection s_1 is at the boundary between P_1 and P_2 and the second intersection s_2 is between P_2 and P_3 . The edge $e_{i,l}$ is divided into three segments $e_{i,l}(1)$, $e_{i,l}(2)$ and $e_{i,l}(3)$. In the middle of these segments are the midpoints m_1 , m_2 and m_3 which determine the surface structure values α_1 , α_2 , and α_3

Fig. 5 The decomposition of an edge $e_{i,l}$ in the three segments $e_{i,l}(1)$, $e_{i,l}(2)$, and $e_{i,l}(3)$ on the three polygons P_1 , P_2 , and P_3 representing different soil conditions

of these segments using the Point-in-Polygon test. The yellow dotted line represents the length d_1 , the red dotted line the length d_2 , and the white line the length d_3 .

The following objective function has to be minimized to compute the cheapest route in terms of construction:

$$\sum_{e_{i,l} \in E, K_{i,l} \subset K, J_{i,l} \subset J} x_{e_{i,l}} \left(\sum_{k \in K_{i,l}, j \in J_{i,l}} d_k \cdot \alpha_j \right) \rightarrow \min! \quad (15)$$

where K is the index set of all existing segments and J is the index set of the occurring surface structures on the edges of E . The complexity of the Point-in-Polygon algorithm is $\mathcal{O}(N_{\text{Poly}})$ and of the calculation of the intersections of an edge with the polygons also $\mathcal{O}(N_{\text{Poly}})$ where N_{Poly} is the number of all nodes of the polygons P . Consequently, the time complexity of calculating the weight $c_{i,l}$ of any edge $e_{i,l}$ is also $\mathcal{O}(N_{\text{Poly}})$.

3.3 Most convenient routes

In this part two variants for the calculation of the trajectory of a most convenient or passable route on U are presented. In the first variant a method is presented in which a route between two nodes is computed with a minimum sum of absolute elevation changes. This means that a route with the least vertical variations is searched for. Similarly, in the second variant, a route with a minimum sum of absolute gradients from the starting point to the destination is calculated.

3.3.1 Absolute elevation change

Both for ordinary traffic and for the transport of heavily loaded vehicles, routes are preferred which are easy and comfortable to drive on. Due to the topography of U , the surface of U has both ups and downs corresponding to the grid $\Gamma_{\mathbb{R}^3}(U)$. An edge $e_{i,j} = (e_{i,j}(1), \dots, e_{i,j}(u_{i,j}))$ of V is divided into $u_{i,j}$ segments. Let V_{aux} be a set of auxiliary nodes on the grid $\Gamma_{\mathbb{R}^3}(U)$ and $A \subset \{(x, y, 0) \in \mathbb{R}^3\}$ an area with an elevation of 0. Analogue to Sect. 3.1, the weight of a segment $e_{i,j}(k)$ is defined as an absolute elevation change between two neighboring nodes $v_{i,j}(k)$ and $v_{i,j}(k-1)$ of $V \cup V_{\text{aux}}$ as follows

$$\Delta h_{i,j}(k) := |h(v_{i,j}(k)) - h(v_{i,j}(k-1))|, \quad (16)$$

for $k = 1, \dots, u_{i,j}$. A route between the starting node s and the destination node t is called the **most convenient route regarding the absolute elevation change**, if the sum of the absolute elevation changes on $\Gamma_{\mathbb{R}^3}(U)$ is minimal. Figure 6 shows that there are two nodes v_i and v_j on U and between these nodes there are horizontal-equidistant auxiliary nodes $v_{i,j}(k)$ of V_{aux} for $k = 1, \dots, u_{i,j} - 1$. These nodes generate between v_i and v_j the segments $e_{i,j}(1), \dots, e_{i,j}(u_{i,j})$ on the surface of U . The surface of U is depicted between v_i and v_j by a red connection curve. An edge line consisting of the segments $e_{i,j}(1), \dots, e_{i,j}(u_{i,j})$ approximates this connecting curve in order to discretely model the surface of U . Discrete modeling of U is sufficient since the

Fig. 6 A depiction of the decomposition of an edge $e_{ij} = \{v_i, v_j\}$ in smaller segments with weights of absolute elevation changes. The red curve describes the topology U of a geographical area A , the black nodes $\{v_{ij}(1), \dots, v_{ij}(u-1)\}$ are the auxiliary nodes between v_i and v_j on U and the blue connections are the segments of an edge e_{ij}

elevation data of terrain are also discrete. The red arrows illustrate the elevation changes between all neighboring auxiliary nodes along an edge e_{ij} .

If an absolute elevation change $\Delta h_{ij}(k)$ of a segment of a route exceeds a maximum elevation change $\Delta h_{\max} > 0$ then this route may not be passable. A segment $e_{ij}(k)$ is called critical with respect to the absolute elevation change if $\Delta h_{ij}(k) > \Delta h_{\max}$. An alternative most convenient route with $\Delta h_{ij}(k) \leq \Delta h_{\max}$ for all $k \in \{1, \dots, u_{ij}\}$ and $e_{ij} \in E$ is searched where u_{ij} is the maximum number of segments of e_{ij} . All weights of the segments within $\Gamma_{\mathbb{R}^3}(U)$ will be modified when considering a maximum elevation change. Let $h_{\max} > 0$ be a maximal allowed absolute elevation change of a segment $e_{ij}(k)$. Then the weight of the segment $e_{ij}(k)$ is defined as

$$w_{ij}(k) = \begin{cases} \Delta h_{ij}(k), & \text{if } \Delta h_{ij}(k) \leq \Delta h_{\max} \text{ and } e_{ij}(k) \notin E_{\text{crit}}, \\ N_1, & \text{if } \Delta h_{ij}(k) > \Delta h_{\max} \text{ and } e_{ij}(k) \notin E_{\text{crit}}, \\ N_2, & \text{if } e_{ij}(k) \in E_{\text{crit}}. \end{cases} \quad (17)$$

where $N_1, N_2 \in \mathbb{R}_+$ are two sufficiently large positive numbers. The weight of an edge e_{ij} is given by

$$w_{ij} = \sum_{k=1}^{u_{ij}} w_{ij}(k). \quad (18)$$

If we assume that the absolute altitude change of $e_{ij}(k)$ may not be greater than $\Delta h_{\max}$ then $N_1 = +\infty$. If for $e_{ij}(k)$ the areas from O are prohibited then $N_2 = +\infty$ when this segment is in O or intersects this set. A route is searched between s and t of V with minimal absolute elevation changes:

$$\sum_{e_{ij} \in E} x_{e_{ij}} w_{ij} \longrightarrow \min! \quad (19)$$

Note that modifying the weights from equation (17) for $N_1 = +\infty$ and/or $N_2 = +\infty$ may result in not finding a solution in the worst-case.

3.3.2 Absolute gradients

Another variant of the most convenient route within a grid $\Gamma_{\mathbb{R}^3}(U)$ is a route with segment weights of absolute gradients

$$s_{ij}(k) = \frac{\Delta h_{ij}(k)}{d'_{ij}(k)}, \quad (20)$$

where $s_{ij}(k)$ is the weight of a segment $e_{ij}(k)$ and $d'_{ij}(k)$ is the distance of $e_{ij}(k)$ using the formulas of Vincenty. In this subsection, we consider the absolute gradients of the segments of $\Gamma_{\mathbb{R}^3}(U)$. A **most convenient route in terms of the absolute gradient** on $\Gamma_{\mathbb{R}^3}(U)$ is a route between a starting node s and a destination node t that has a minimum sum of absolute gradients in $\Gamma_{\mathbb{R}^3}(U)$. Such a route is a route with minimal fluctuations from the starting point to the destination. Let $s_{\max} \in \mathbb{R}$ be a maximum gradient a vehicle can drive. A segment $e_{ij}(k)$ is critical with respect to the absolute gradient if $s_{ij}(k) > s_{\max}$. This critical segment is part of the route that is unsuitable for vehicles due to the high gradient of this segment. All weights of the critical segments of $\Gamma_{\mathbb{R}^3}(U)$ have to be modified. The weight of $e_{ij}(k)$ is modified as follows

$$s_{ij}(k) = \begin{cases} \frac{\Delta h_{ij}(k)}{d'_{ij}(k)}, & \text{if } \Delta h_{ij}(k) \leq d'_{ij}(k) \cdot s_{\max} \text{ and } e_{ij}(k) \notin E_{\text{crit}}, \\ N_1, & \text{if } \Delta h_{ij}(k) > d'_{ij}(k) \cdot s_{\max} \text{ and } e_{ij}(k) \notin E_{\text{crit}}, \\ N_2, & \text{if } e_{ij}(k) \in E_{\text{crit}}, \end{cases} \quad (21)$$

where N_1 and N_2 are sufficiently large positive real values. A route is searched from s to t with minimal absolute gradients:

$$\sum_{e_{ij} \in E} x_{e_{ij}} \sum_{k=1}^{u_{ij}} s_{ij}(k) \longrightarrow \min! \quad (22)$$

Analogue to the most convenient routes with respect to the absolute altitude change, the modification of the weights from equation (21) for $N_1 = +\infty$ and/or $N_2 = +\infty$ can result in no solution in worst-case.

4 Pareto optimal routes

Multi-criteria optimization (Pareto optimization) is used to optimize an objective function with several criteria (such as minimizing the route length as well as minimizing the construction costs of this route) Ehrgott (2005). Often it is not possible to optimize all objectives simultaneously since in general, it is not possible to improve the first objective without worsening the other objectives. Since there is no unique

optimal solution defined. A set of solutions is determined in which the improvement of one objective function can be achieved only by the worsening of another, i.e., the set of optimal trade-offs (Pareto front). In our case, the Pareto front gives the optimum of the combination of the three criteria, the shortest route with respect to the distance, the cheapest route regarding the construction costs and the most convenient route in terms of the absolute elevation change or absolute gradient. In our investigation, we have defined a Pareto objective function $f_{\text{Pareto}}(E)$ on $\Gamma_{\mathbb{R}^3}(U)$ that consists of a distance function $f_D(E)$, a cost function $f_C(E)$, and an absolute elevation change function $f_H(E)$ as follows:

$$f_{\text{Pareto}}(E) = \alpha \cdot f_D(E) + \beta \cdot f_C(E) + \zeta \cdot f_H(E) \longrightarrow \min!, \quad (23)$$

with

$$f_D(E) = \sum_{e_{ij} \in E} x_{e_{ij}} d_{ij}, \quad (24)$$

$$f_C(E) = \sum_{e_{ij} \in E} x_{e_{ij}} c_{ij}, \quad (25)$$

and

$$f_H(E) = \sum_{e_{ij} \in E} x_{e_{ij}} w_{ij}, \quad (26)$$

such that for all $v \in V$:

$$\sum_{e_{ij} \in \delta^-(v)} x_{e_{ij}} - \sum_{e_{ij} \in \delta^+(v)} x_{e_{ij}} = \begin{cases} -1, & \text{if } v = s, \\ 1, & \text{if } v = t, \\ 0, & \text{else,} \end{cases}$$

$$\forall e_{ij} \in E : x_{e_{ij}} \geq 0,$$

where $\alpha + \beta + \zeta = 1$ for all $\alpha, \beta, \zeta \in [0, 1]$. Since $\Gamma_{\mathbb{R}^3}(U)$ is discrete, the minimum of the objective function $f_{\text{Pareto}}(E)$ generates a solution space S_{Pareto} of points in $\mathbb{R}^3$ for all $\alpha, \beta, \zeta \in [0, 1]$ with respect to the equation $\alpha + \beta + \zeta = 1$. This space lies on a hypersurface of $\mathbb{R}^3$. Each point of S_{Pareto} is a minimum of the function $f_{\text{Pareto}}(E)$. The Cartesian coordinate axes that depicted the solution set S_{Pareto} represent the three criteria distance, construction costs, and absolute elevation changes of a route. The values of each axes indicate the quantity or unit of the objective functions f_D in meters or kilometers, f_C in monetary units and f_H in meters, respectively. The values of S_{Pareto} indicate which values the respective objective functions have for each axis in $\mathbb{R}^3$. In case one is looking for Pareto optima with two criteria with respect to $\{(\alpha, \beta, \zeta) \in [0, 1]^3 \mid \alpha + \beta + \zeta = 1\}$ then one can orthogonally project the set S_{Pareto} from $\mathbb{R}^3$ to $\mathbb{R}^2$. The projection of S_{Pareto} is a curve in $\mathbb{R}^2$ in that each coordinate axis of this subspace indicates the values of the two remaining criteria. Let

$\zeta = c_0 \in [0, 1)$ be a constant with respect to the equation (23). Then $\beta = 1 - c_0 - \alpha$ for all $\alpha \in [0, 1 - c_0]$. The set S_{Pareto} is projected for a fixed c_0 onto a curve in $\mathbb{R}^2$. This will give all optimal compromises between the two functions f_D and f_C for any c_0 part of the function f_H .

5 Computing optimal routes using CONTRA

For the construction of new routes, we developed a software called CONTRA, which is the acronym for **C**omputing an **O**ptimal **N**etwork of **T**ransit **R**outes through mathematical Algorithms. This tool is capable of computing and planning the construction of new routes connecting two locations (e.g., cities) in any country. These routes consist possibly of different infrastructure segments such as roads, railways, gas pipelines, etc. This software uses GIS-data, e.g., ground conditions and elevation data, and optimization techniques to compute the building of optimal routes. The output of the software is a summary of the data of the computed route, such as road coordinates, construction costs, and lengths. As an example, we compute routes between the city of *Khas Uruzgan* and the city of *Kabul*. In this part we computed routes that are below the elevation of 3500 m. In practice, it is unusual to build routes in elevated areas, because the oxygen decreases with increasing altitude.

5.1 Input data

CONTRA uses for the computation different GIS-data of a chosen area. To determine an optimal route, the following input data is required by this program:

1. Coordinates of the start and destination points (in latitude and longitude) are given in Table 1. These two points are colored in red in Fig. 7. The red dashed line is the line-of-sight connection between them.
2. For the national land cover of Afghanistan the shapefiles of the organization AIMS (1997), formerly of the Afghan Geodesy and Cartography Head Office, were used (AIMS 1997). The whole country is partitioned in polygonal shapes that describe the respective type of land, see Fig. 8.
3. For the topographic representation of Afghanistan, the SRTM-30m data of NASA are used Watkins (2019). The resolution of the SRTM-30m data for Afghanistan is 1-arcsecond (in total 36001×54001 pixels) in a latitude/longitude projection (EPSG:4326). As an example, in Fig. 9 the area of the province of Uruzgan in Afghanistan is depicted.

Table 1 Start and destination points as geographical coordinates

Coordinates/Points	Start-Khas Uruzgan	Destination-Kabul
Latitude	32.936415	34.513334
Longitude	66.681918	69.117687

Fig. 7 Start and destination points from the city of Khas Uruzgan to Kabul in Afghanistan

Fig. 8 A map of Afghanistan with different land cover. Each color represents a specific land cover (e.g., deserts or forests)

4. The maximum gradient angle of the routes is $\varphi_{\max} = 2.8624^\circ$, which corresponds to a maximum gradient of 5%. If this gradient is exceeded in steep areas, serpentine with a gradient of 5% are built.
5. According to reports of the Asian Development Bank and the Ministry of Public Works of the Islamic Republic of Afghanistan, the construction of a two-lane road in Afghanistan is on average approximately USD 1000 per meter (ADB 2007). The cost amount can depend on the land surface on which it is built (Operations

Fig. 9 The topography of the Uruzgan province in Afghanistan. The color spectrum from dark to light indicates the height of the terrain

- 2010). For our research we assumed an estimated cost depending on the different land covers. In general, it is possible to modify the estimated cost.
6. We assume the construction costs of roads along a mountain pass or on the mountainside on average USD 3000 per meter. The costs of quarrying and earthworks is higher than that of building an ordinary road.

5.2 Results

We compute optimal solutions for the construction of a route from the city of Khas Uruzgan to the city of Kabul in Afghanistan regarding three different single objectives: the shortest route in red, the cost-minimal or cheapest route in green, and the most convenient route w.r.t. the elevation change in blue, see Fig. 10. The columns in Table 2 describe (a) the construction costs of the routes in million USD, (b) the lengths of the routes in km and (c) the absolute elevation changes of the routes, i.e., the sum of all height changes from the starting point to the ending point along the route. Each of these three routes has a certain length and elevation profile.

Figure 11 shows the elevation profile of a shortest road with respect to the distance between the cities Khas Uruzgan (left) and Kabul (right). Note that there is

Fig. 10 Three single-objective optimal roads between the cities Khas Uruzgan and Kabul in Afghanistan computed with CONTRA: shortest (red), cheapest (green), and most convenient (blue). Image rendering: Google Earth

Table 2 The result of the three routes regarding the cost, length, and absolute elevation changes

Optimal route type	Cost in million USD	Length in km	Absolute elevation changes in m
Shortest route	463	368	16,278
Cheapest route	360	381	18,216
Most convenient route	910	479	11,597

Fig. 11 The elevation of a shortest road on terrain from left the city of Khas Uruzgan to right the city of Kabul in Afghanistan. Image rendering: Google Earth

Fig. 12 The elevation of a cheapest road on terrain from left Khas Uruzgan to right Kabul in Afghanistan. Image rendering: Google Earth

Fig. 13 The elevation of a most convenient road with minimal absolute elevation changes from left Khas Uruzgan to right Kabul. Image rendering: Google Earth

a slight difference in the lengths of that route computed by our algorithm (337 km, see Table 2) and the re-rendering of that very path in Google Earth (299 km). We believe that this is due to the fact that Google Earth does not take the elevation changes into account when computing a path length. Figure 12 shows the elevation profile of a cheapest road with respect to the construction costs from Khas Uruzgan to Kabul. Analogously, Fig. 13 visualizes the elevation profile of the most convenient road above sea level between these two cities. One can see that this route is a longer distance, but most parts of the route have low absolute elevation changes.

For the Pareto optimum, the following objective function of cost and elevation change is minimized for different $0 \leq \lambda \leq 1$:

$$\sum_{e_{ij} \in E} x_{e_{ij}} (\lambda c_{i,j} + (1 - \lambda) w_{i,j}) \longrightarrow \min! \quad (27)$$

Figure 14 shows the Pareto front of the routes from Khas Uruzgan to Kabul regarding the two objectives “construction costs” and “elevation changes”. This chart shows on the horizontal axis the construction costs of the routes in million USD and on the vertical axis the absolute elevation changes in meters. The small red circles represent specific routes between these two cities. The leftmost circle represents the cheapest route and the right most circle represents the most convenient route with minimal altitude variation. The circles that lie between these two extremal circles, are other optimal routes from Khas Uruzgan to Kabul which are trade-offs of these two criteria. The gaps between the circles in this figure indicate that there is an unsuitable area (obstacle) between two routes so that the construction of a route across this area is sub-optimal. Thus, the routes run around such “obstacle areas” and generate gaps in the Pareto front.

Fig. 14 The Pareto front of the most convenient and cheapest routes regarding absolute elevation change and construction costs from Khas Uruzgan to Kabul. Each red circle represents an optimal trade-off routes between the cheapest and the most convenient routes where the most left circle represents the cheapest route and the most right circle represents the most convenient route

Figure 15 shows the optimal routes picked from the red circles in Fig. 14 for six different λ values. For the extremal value $\lambda = 1$ the Pareto optimal route is the cheapest route between Khas Uruzgan and Kabul (see top left depiction in Fig. 15). The second extremal value $\lambda = 0$ presents the most convenient route between these two cities (see below right depiction in Fig. 15). All the other depictions are optimal compromises between these two extremes.

6 Conclusions

The search for a suitable trajectory for the construction of a route in the transport or energy sector can be very time-consuming considering the environmental information. This paper presented a new method of computing the shape of a new route using discrete optimization and mathematical algorithms. The relevant information was taken from the GIS data such as the geographical environment, structure of the soil, topographic elevation data of terrain, and average construction costs of an infrastructure segment. In the first step, U was discretely approximated by a weighted grid Γ . In the second step, the weights of this grid were computed regarding the three criteria: (a) distance, (b) construction costs and (c) absolute elevation changes or absolute gradients. In the third step, routes with minimal edge weights were determined between two locations whereby these routes are globally optimal. Furthermore, with Pareto optimization, we have

Fig. 15 The variants of optimal trade-offs for $\lambda = 1, 0.5737, 0.3413, 0.0925, 0.0325, 0$: transition from the cheapest route (top left) to the most convenient route (bottom right)

combined these optimization criteria to determine optimal compromises between them that correspond to the real conditions. Finally, we modeled the area between Khas Uruzgan and Kabul in Afghanistan with our computer program CONTRA and with the corresponding GIS data and determined the shortest, cheapest, and most convenient roads between these two cities. With the Pareto optimization we determined several compromises between the cheapest and the most convenient route and presented the change between both extremes as trajectories. In further work, we have the goal to compute the construction of optimal routes with tunnels and bridges. The calculation of tunnels or bridges along a route is more complicated than routes on the surface. When planning tunnels and bridges, the

computation must be carried out in a three-dimensional space. This makes the calculation more difficult and time-consuming.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- ADB (2007) Proposed Asian development fund grant islamic republic of afghanistan: road network development project 1. Report and Recommendation of the President to the Board of Directors 40333, Asia Development Bank
- Afghan Geodesy and Cartography Head Office and AIMS. Land Cover of Afghanistan (1997). Technical report, Afghan Geodesy and Cartography Head Office, 2014
- Basu S, Snoeyink J (2007) Terrain Representation using Right-Triangulated Irregular Networks. In: CCCG, pp 133–136. Citeseer
- Belaidi H, Bentarzi H, Belaidi A, Hentout A (2014) Terrain traversability and optimal path planning in 3D uneven environment for an autonomous mobile robot. *Arab J Sci Eng* 39(11):8371–8381
- Dijkstra EW (1959) A note on two problems in connexion with graphs. *Numerische Mathematik* 1(1):269–271
- Ehrgott M (2005) Multicriteria optimization, 2nd edn. Springer, Berlin
- Fay M, Yepes T (2003) Investing in infrastructure: What is Needed from 2000 to 2010? World Bank Policy Research Working Paper Series, 3102
- Fowler IL, Little J (1979) Automatic extraction of irregular network digital terrain models. *Comput Graph* 13(2):199–207
- Gaal Hassan Osman (2017) Lack of Infrastructure: The Impact on Economic Development as a case of Benadir region and Hir-shabelle, Somalia. *Developing Country Studies* 7(1)
- Gurara D, Klyuev V, Mwase N, Presbitero A, Xu XC, Bannister G (2017) Trends and challenges in infrastructure investment in low-income developing countries. IMF working papers
- Hanisch F (2004) Marching square. In: CGEMS: computer graphics educational materials. The Eurographics Association
- Ho C-C, Wu F-C, Chen B-Y, Chuang Y-Y, Ouhyoung M (2005) Cubical marching squares: adaptive feature preserving surface extraction from volume data. *Comput Graph Forum* 24(3):537–545
- Mitchell JSB, Sharir M (2004) New results on shortest paths in three dimensions. In: Proceedings of the twentieth annual symposium on Computational geometry: SCG'04. ACM Press
- Muñoz P, R-Moreno MD, Castaño B (2017) 3Dana: a path planning algorithm for surface robotics. *Eng Appl Artif Intell* 60:175–192
- Operations (2010) Northern Land Use Guidelines - Access: Roads and Trails. Report Volume 5, Indigenous and Northern Affairs Canada
- Oyedele Olufemi Adedamola (2016) Infrastructure Problems of Developing Nations and Sustainable Development. In: *International Public Procurement Conference, IPPC7, Bali, Indonesia*
- Saranya C, Unnikrishnan M, Akbar Ali S, Sheela DS, Lalithambika VR (2016) Terrain Based D* Algorithm for Path Planning. *IFAC-PapersOnLine* 49(1):178–182
- Sunday Dan (2001) Inclusion of a Point in a Polygon. Online, Internet
- Turau V (2004) Algorithmische Graphentheorie. Oldenbourg, Deutschland

- Vincenty T (1975) Direct and Inverse Solutions of Geodesics on the Ellipsoid with Application of Nested Equations. *Surv Rev* 23(176):88–93
- Watkins Derek (2019) SRTM 30m Digital Elevation Data, SRTMHGT files at 1-arcsecond resolution in a latitude/longitude projection (EPSG:4326). downloaded from NASA servers. Technical report, NASA
- Zazai MF (2020) Berechnung der Trajektorien für die Entwicklung von optimalen Routen. PhD thesis, Brandenburg University of Technology (BTU) Cottbus-Senftenberg
- Zazai MF, Fügenschuh A (2018) Computing pareto-optimal transit routes through mathematical algorithms. In *Operations research proceedings*, pp 467–473. Springer International Publishing

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.