Multi-Haul Quasi Network Flow Model for Vertical Alignment Optimization

by

Vahid Beiranvand

B.Sc. Hons., Islamic Azad University, Iran, 2008
M.Sc., National University of Malaysia, 2012

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE COLLEGE OF GRADUATE STUDIES

(Interdisciplinary Studies)

THE UNIVERSITY OF BRITISH COLUMBIA

(Okanagan)

January 2016

Abstract

The vertical alignment optimization problem for road design aims to generate a vertical alignment of a new road with a minimum cost, while satisfying safety and design constraints. We present a new model called multi-haul quasi network flow (MH-QNF) for vertical alignment optimization that improves the accuracy and reliability of previous mixed integer linear programming models. We evaluate the performance of the new model compared to two state-of-the-art models in the field: the complete transportation graph (CTG) and quasi network flow (QNF) models. The numerical results show that, within 1% relative error, the proposed model is robust and solves more than 93% of test problems. Whereas, the CTG only solves about 82% of test problems and QNF fails to solve any problem within 1% relative error. Moreover, in terms of computational time, on average the MH-QNF model solves the problems approximately 8 times faster than the CTG model.

Preface

This research is based on work conducted in the Centre for Optimization, Convex Analysis & Nonsmooth Analysis (COCANA) by Dr. W. Hare, Dr. Y. Lucet, and V. Beiranvand. I was responsible for developing the model, implementing it, and evaluating its performance by designing and performing experiments.

A version of this thesis has been submitted to a journal for publication as V. Beiranvand, W. Hare, Y. Lucet and S. Hossain (2015). Multi-Haul Quasi Network Flow Model for Vertical Alignment Optimization. I conducted all the testing and wrote the manuscript. Shahadat Hossain contributed by providing ideas.

Table of Contents

List of Tables

List of Figures

Acknowledgments

I would like to express my sincere gratitude to all the individuals who were involved directly or indirectly with this work. Specially, I would like to thank my supervisors Dr. Warren Hare and Dr. Yves Lucet for their continuous support during the research, their patience, motivation, and knowledge.

I am thankful to our industrial partner (Softree Technical Systems Inc.) for supporting the research. In particular, David Mills, Craig Speirs, and Alexis Guigue for providing technical information and valuable feedback throughout this thesis.

This research is funded by a Collaborative Research and Development (CRD) Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada, and Softree Technical Systems Inc.

Chapter 1

Introduction

Transportation infrastructures are an important sign of development and welfare in a given country [CCGC15, HL07]. While planners normally try to balance and coordinate among various transportation methods, roads and highways are considered the leading component of transportation infrastructures. Due to the fast increase of traffic volume, the high cost of road construction projects, and the environmental and safety impacts of these projects, road projects require continuing innovation to minimize the costs while satisfying environmental and safety constraints. Road design is one of the early and significant steps in road construction projects. Road design influences the construction cost and other contributing constraints remarkably. From the optimization point of view, the main objective in road design problems is to minimize the road construction costs (primarily excavation, embankment, and hauling costs), subject to road safety and quality constraints.

Road design starts by determining the corridor where the road is to be constructed. From here, a so called horizontal alignment is fixed. Next, vertical alignment and earthwork movement problems must be solved. In the literature, these last two steps are often combined for the sake of efficiency [MA04, HKL11]. To calculate horizontal alignment, the vertical alignment and earthwork movement problems should be solved accurately. If the vertical alignment and earthwork problems can be solved faster, then we can evaluate more possible horizontal alignments. Therefore by solving the vertical alignment and earth work movement problems accurately with a low computational cost, we can obtain a more precise horizontal alignment.

It should be noted that there are some studies that consider simultaneous optimization of horizontal alignment, vertical alignment, and earthwork movement [CGF89, HEAEH98, ESHS02, CL06]. However, by necessity, these approaches all use some heuristic(s), and the resulting roads generally still require a final detailed optimization of the vertical alignment and earthwork movement problems. In this research we focus on solving the combined vertical alignment and earthwork movement problem.

Considering the combined vertical alignment and earthwork movement

problem, we note that the final solution must satisfy various constraining factors, such as safety standards, economical limitations, environmental constraints, engineering requirements, and socioeconomic requirements [Fwa06, AAS11]. In addition to these considerations, inclusion of a set of design constraints and operational requirements makes the vertical alignment and earthwork movement problem complicated.

1.1 Horizontal alignment optimization

In a given corridor there are a lot of possible routes that can connect the source point to the destination point by building a road (see Figure 1.1). In horizontal alignment design, the objective is to identify the best route among all possible routes between the given points to build the road. The best route is the one with the lowest cost while satisfying all the design requirements. For the detailed understanding of the horizontal alignment optimization we refer to [MLH15].

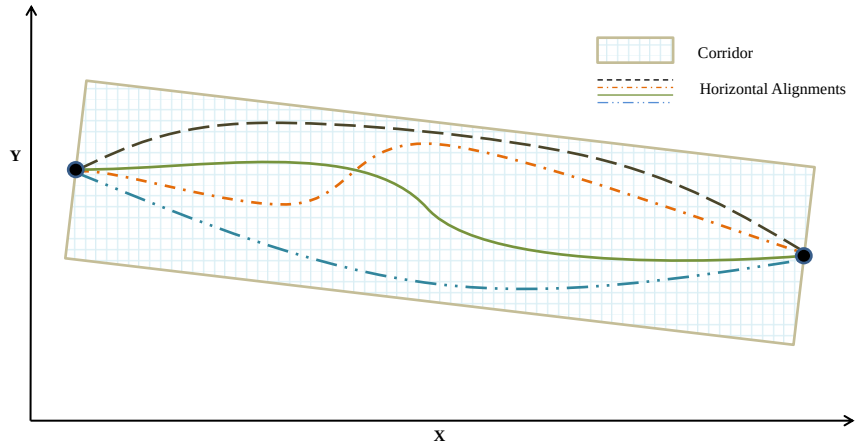

Figure 1.1: Some potential horizontal alignments in a corridor.

1.2 Vertical alignment optimization

In the vertical alignment optimization the horizontal alignment is fixed. The vertical alignment is a modification of the ground profile elevation required to build a road that minimizes the cost of road construction while satisfying safety and regulation constraints [HHLR14]. An example of a

vertical alignment is shown in Figure 1.2. In other words, the vertical alignment optimization seeks the minimum earthwork allocation cost satisfying the vertical alignment design constraints.

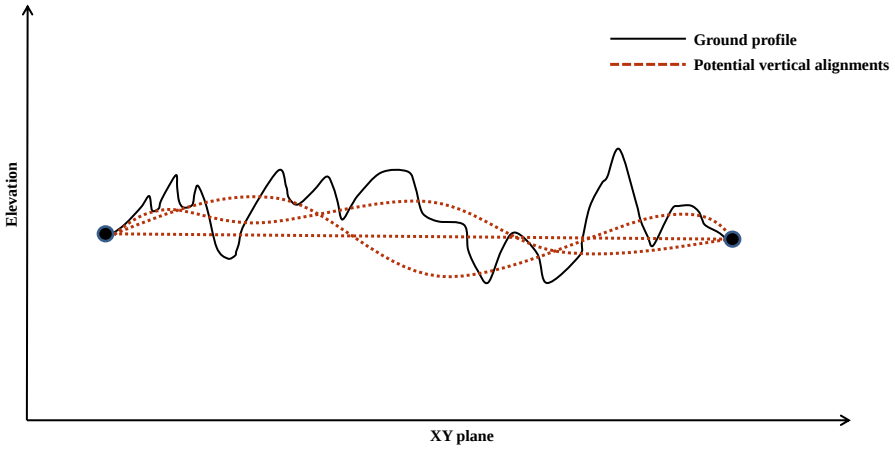

Figure 1.2: Some potential vertical alignments for a ground profile.

Past research on the vertical alignment problem has introduced a number of interesting techniques. Heuristic optimization methods, such as Genetic Algorithms (GAs), have been used by a number of authors. For example, Lee and Cheng [LC01] developed a mathematical model for vertical alignment and proposed a heuristic method to solve it. Jha et al. [JSJK06] present a comprehensive review of the road design problem and the related challenges. They also present solution algorithms based on GAs to optimize highway alignments by adapting operators and encoding schemes. Other heuristic based optimization methods for highway alignment optimization include [JS03, Aru05, KJS05, GLA09, KJS12]. Since the genetic algorithms and other heuristic based methods are non-deterministic they do not guarantee finding the optimum solution. Moreover, the results obtained by these methods are subject to change in different runs of the same algorithm on the same test problem. Therefore, we focus our research on mixed integer linear programming methods (discussed below).

Another early approach relied on dynamic programming. Dynamic programming was first employed by Goh et al. [GCF88] to develop two models for vertical alignment. They use piecewise-linear segments to represent the alignment. In general, compared to optimizing the horizontal alignment, the dynamic programming approach has a better performance in vertical alignment optimization; however, the resulting alignment is limited to a finite set

of points at each station. It means that only a portion of the problem search space is explored [Jon98]. Other models can be found in the literature which use dynamic programming [NEW75, Tri87, LPZL13].

The current state-of-the-art on the vertical alignment and earthwork movement problem relies on linear programming, or mixed integer linear programming. Easa [Eas88] proposed a two-steps linear programming model to optimize vertical alignment and earthwork movement. In the first step, by enumerating all technically feasible grades, a feasible road grade is chosen and for each grade the amount of cut and fill is calculated. In the second step, the earthwork transportation cost is minimized using linear programming. This trial and error approach does not guarantee finding the global solution. Moreb [Mor96] proposed a linear programming model for vertical alignment which incorporates the grade selection and earthwork allocation stages. In [MA04], Moreb and Aljohani extended the work of Easa [Eas88] to handle the sharp connectivity points problem in piecewise linear models by representing the road profile as a continuous quadratic curve. In [Mor09], Moreb proposed a model in which the road profile is represented as a one-dimensional spline. This model improves the computational efficiency and guarantees the global optimality by solving a single linear programming problem. Koch and Lucet [KL10] improved the Moreb model by introducing new gap and slope constraints that reduce errors in the model.

Hare et al. [HKL11] developed a mixed integer linear program (MILP) model for earthwork operations in which blocks are also taken into account. In [HLR15], the vertical alignment problem is dealt with and road side slopes are added to the previous MILP model proposed in [HKL11], and hence, a more accurate earthwork cost calculation is obtained. In another update [HHLR14], Hare et al. improved the performance of the MILP model (called complete transportation graph or CTG) and studied a novel quasi network flow (QNF) model for vertical alignment optimization considering the earthwork allocations. The QNF model presented significant improvements in computational time, but at a cost of decreased accuracy in the final solution. The CTG and QNF models play a key role in this thesis; further details on these models can be found in Section 2.3.

1.3 Three dimensional alignment optimization

In three dimensional alignment optimization, the above-mentioned road design phases (i.e., the vertical and horizontal alignments optimization) are optimized, *simultaneously*. For a detailed discussion on three dimensional

alignment optimization we refer to [MLH15].

1.4 Research overview

In this research, we improve the quasi network flow model of [HHLR14], by presenting a new model called the multi-haul quasi network flow (MH-QNF) model. The MH-QNF model balances the accuracy of the CTG model in approximating the earthwork cost with the low computational cost of the QNF model. The main idea behind the new model is to differentiate between hauls of different distances in earthwork allocation.

Chapter 2

Multi-haul quasi network flow model

In this chapter we describe the variables, parameters, and sets required to model the vertical alignment of roads. We first provide extensive details on the multi-haul quasi network flow (MH-QNF) model. Then we describe the quasi network flow (QNF) and complete transportation graph (CTG) models. All models are designed to optimize vertical alignment for a known horizontal alignment.

2.1 Model variables and parameters

The road is approximated as a quadratic spline. It is split into m segments indexed by $\mathcal{G} = \{1, 2, ..., m\}$. For all $g \in \mathcal{G}$, each segment is represented using the following equation

$$P_g(s) = a_{g,1} + a_{g,2}s + a_{g,3}s^2 \tag{2.1}$$

in which s is the distance from the beginning of the segment.

Each segment is further split into sections so that the g^{th} spline segment is made of n_g sections indexed by the set $\mathcal{S}_g = \{1, 2, ..., n_g\}$. The total number of sections in a road is $n = \sum_{g \in \mathcal{G}} n_g$ and these sections are indexed by the set $\mathcal{S} = \{1, 2, ..., n\}$. The section indexes of a specific spline segment g are mapped to the actual section index set $\mathcal{S}$ using the function $\varphi : (\mathcal{G} \times \mathcal{S}_g) \to \mathcal{S}$. For example, if $\varphi(g, j) = i$ then $s_i = s_{\varphi(g,j)}$ for all $i \in \mathcal{S}, g \in \mathcal{G}, j \in \mathcal{S}_g$. So the spline function representing the road profile is as follows

$$P(s) = \begin{cases} P_1(s) & \textbf{if } s_{\varphi(1,1)} \leq s \leq s_{\varphi(1,n_1)}, \\ P_2(s) & \textbf{if } s_{\varphi(2,1)} \leq s \leq s_{\varphi(2,n_2)}, \\ \vdots & \\ P_{n_g}(s) & \textbf{if } s_{\varphi(n_g,1)} \leq s \leq s_{\varphi(2,n_{n_g})}. \end{cases}$$

In the vertical alignment problem, one goal is to obtain the optimal offset between the ground profile and the road profile. In addition, the required cut and fill volumes should be calculated for each offset. The offset is denoted by $u_i, i \in \mathcal{S}$, and the cut and fill volumes of a section i are denoted by V_i^+ and $V_i^-, i \in \mathcal{S}$ respectively. When constructing a road we need to consider borrow pits to bring in material and waste pits to dump extra materials. They are modeled as external sections and indexed by the sets $\mathcal{B} = \{1, 2, ..., n_\beta\}$ and $\mathcal{W} = \{1, 2, ..., n_w\}$ in which n_β is the number of borrow pit sections and n_w is the number of waste pit sections. A borrow pit index is mapped to the section index to which it is attached by the function $\vartheta : \mathcal{B} \to \mathcal{S}$. Similarly, the function $\delta : \mathcal{W} \to \mathcal{S}$ maps the waste pit index to the section index to which it is attached.

Access roads are required in road construction and they are used as gateways to the road being constructed. Access roads are linked to a section of the road and are indexed by the set $\mathcal{R} = \{1, 2, ..., n_r\}$ in which n_r is the number of access roads in the construction site. The function $\varrho : \mathcal{R} \to \mathcal{S}$ maps an access road index to the section index to which it is linked. The capacity of the i^{th} borrow pit (respectively waste pit) is denoted by C_i^b (respectively C_i^w). The set $\mathcal{N} = \mathcal{S} \cup \mathcal{B} \cup \mathcal{W} = \{1, 2, ..., n+n_\beta+n_w\}$ represents all the indexes for sections, borrow pits, and waste pits. The *dead haul distance* is the distance between a borrow/waste pit and the section to which it is linked. The dead haul distance of the i^{th} borrow or waste pit is denoted by $\tilde{d}_i$.

The side slope is defined as the steady decrease/increase of height when moving orthogonally from the road profile to the ground profile in a cut/fill section. Side slopes are represented as trapezoid shaped cross-sections and are approximated using stacked rectangles to preserve the linearity of the model, see [HHLR14].

Another important consideration in vertical alignment optimization are blocks [HKL11]. Blocks are obstacles that need to be removed to access some parts of the corridor. For example, they could indicate a river or mountain, and require building a bridge or a tunnel. We define n_b as the number of blocks in the corridor and the blocks are indexed by the set $\mathcal{I} = \{1, 2, ..., n_b\}$. The function $\gamma : \mathcal{I} \to \mathcal{S}$ maps a block index to the section index to which it is linked. To model the block removal process we use a time step t which specifies the time at which a block is removed. In [HKL11], it is shown that to remove n_b blocks we need at most $n_b + 1$ time steps. So we define the set $\mathcal{T} = \{0, 1, 2, ..., n_b\}$ to model the required time steps and we use the binary variables $y_{k,t}$ for each block $k \in \mathcal{I}$ and time step $t \in \mathcal{T}$ to determine whether a block is removed or not. After a block is removed we can move material

over its section.

2.2 Multi-haul quasi network flow model

In the MH-QNF model, similar to the QNF model [HHLR14], we draw sections and pits as nodes while arcs show feasible movements between them. In the QNF model the authors use a single haul to move materials. You can think of this single haul as a single type of earthmoving equipment such as a bulldozer. In real roadway construction sites, different equipments are required for earthmoving tasks. For example, for the short distances using a bulldozer may be more economical, while for a long distance movement a truck may be preferred. Therefore, to obtain a more realistic model, we extend the previous QNF model [HHLR14] by using multiple hauling paths.

Figure 2.1 shows the typical flows for a section in the proposed MH-QNF model. In Figure 2.1, there are three hauling paths at each side. At each side, you can think of each hauling path as a specific type of construction machinery with known hauling and loading costs. Depending on the hauling distance we can determine which machinery is more efficient to be used. In this model, the material can be moved from the current section i to the next section $(i+1)$ or the previous one $(i-1)$. Depending on the distance of a hauling task, the material movement can be performed through one of a number of hauling paths indexed by $\mathcal{H} = \{1, 2, ..., n_h\}$. (In our figures and numerical tests we shall apply $n_h = 3$, which is based on consultation with our industry partner, whose standard software includes *free-haul*, *over-haul*, and *end-haul*.)

When the flow of material reaches a section, materials can be unloaded to fill the section, or a cut from the section can be performed with the new materials added to the flow, or the section remains unchanged and the material is transferred to the next section. To transfer material to the left or right we use virtual transit nodes for both left and right directions. Since we have n_h different hauls, n_h groups of transit nodes are denoted in the model, see Figure 2.2 (which represents three hauls for consistency with numerical experiments). If material is moved to the next section, the left-side transit nodes are employed, otherwise if the material is moved to the previous section the right-hand side transit nodes are used. In Figure 2.2, we define $f^{r,s}_{i,i+1,t}$ and $f^{r,s}_{i,i-1,t}$ (for all $i \in \mathcal{S}, t \in \mathcal{T}$) as the flow of material from the left (right) short haul transit node i to the left (right) short haul transit node $i+1$ (or $i-1$) at time step t, respectively. Similarly, we use the variables $f^{r,m}_{i,i+1,t}$ and $f^{r,m}_{i,i-1,t}$ (for all $i \in \mathcal{S}, t \in \mathcal{T}$) for the middle hauling

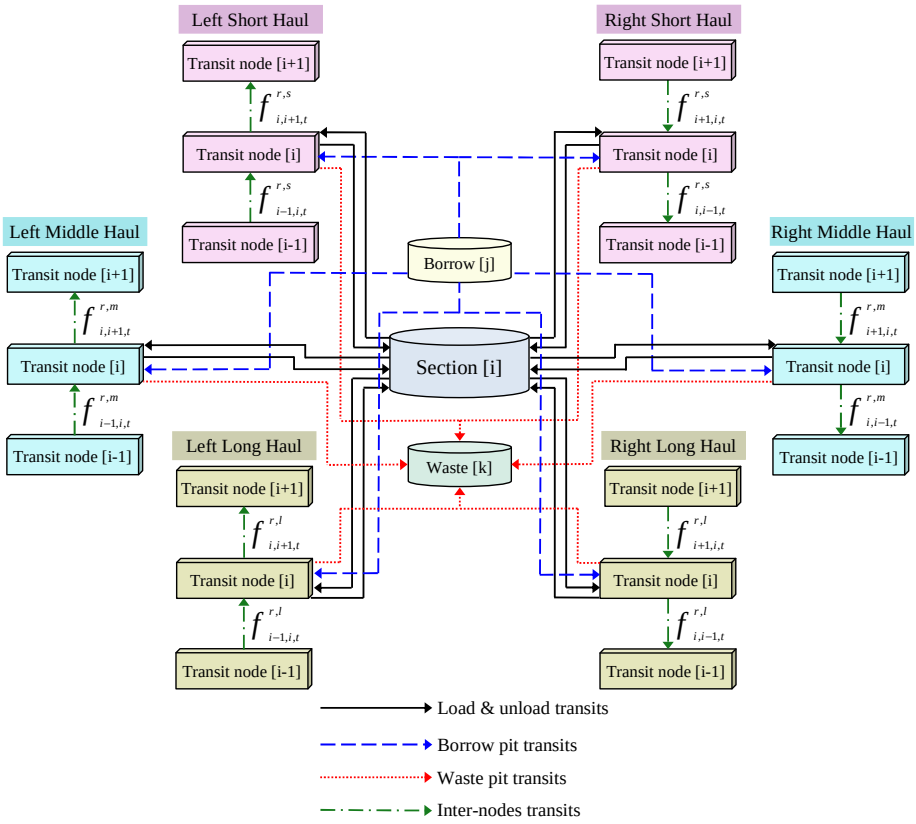

Figure 2.1: A typical section in MH-QNF model

path, and the variables $f^{r,l}_{i,i+1,t}$ and $f^{r,l}_{i,i-1,t}$ (for all $i \in \mathcal{S}, t \in \mathcal{T}$) for the long hauling path.

Figure 2.2 shows the load and unload flows that model the fill and cut volumes of materials for each section. Whenever we want to move material from a section i to the section j we need to cut material from section i. Then, the cut material should be loaded to the suitable machinery (i.e., the material is unloaded from section) depending on the hauling distance. This loading task is modeled by the loading flows $f^{l,h}_{i-1,i,t}$ and $f^{l,h}_{i+1,i,t}$ (for all $i \in \mathcal{S}, t \in \mathcal{T}, h \in \mathcal{H}$) depicted in Figure 2.2. Similarly, at the destination section j the unloading of material into the section is modeled using the unloading flows $f^{u,h}_{i,i+1,t}$ and $f^{u,h}_{i,i-1,t}$ (for all $i \in \mathcal{S}, t \in \mathcal{T}, h \in \mathcal{H}$). After cutting the materials we can transfer it to the left or right using any of the three hauling options (short, middle, or long). Similarly, a section can

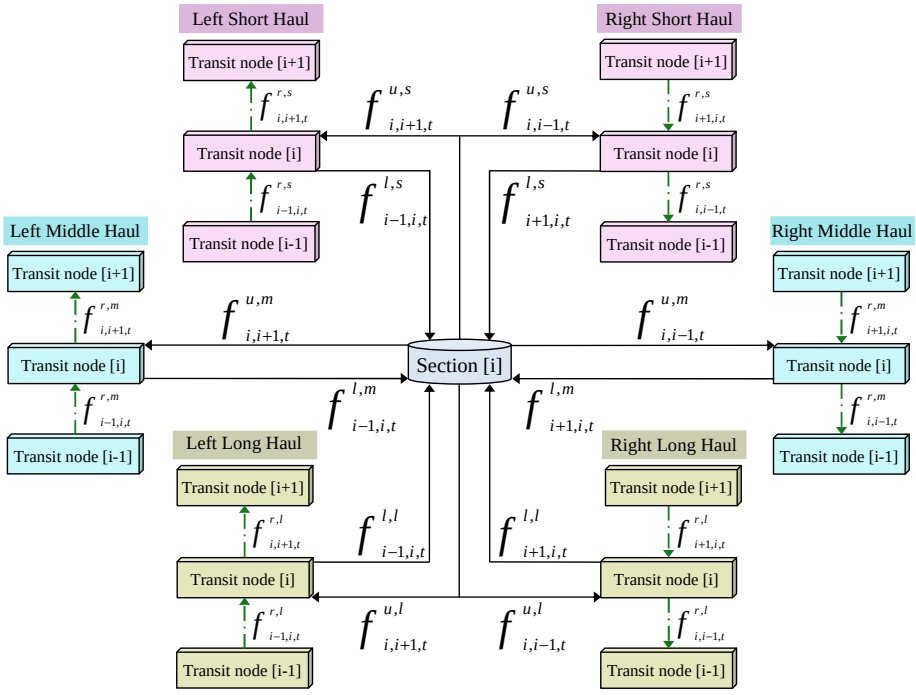

Figure 2.2: Load and unload flows in the MH-QNF model

be filled using the materials taken from left or right transit nodes. To model these possibilities, we introduce the variables $f^{l,s}_{i-1,i,t}$ and $f^{l,s}_{i+1,i,t}$ (for all $i \in \mathcal{S}, t \in \mathcal{T}$) as the load (fill) flows of materials from the left and right transit nodes for the short hauling path. The variables $f^{l,m}_{i-1,i,t}$, $f^{l,m}_{i+1,i,t}$; and $f^{l,l}_{i-1,i,t}$ and $f^{l,l}_{i+1,i,t}$ (for all $i \in \mathcal{S}, t \in \mathcal{T}$) are used as the load flows of materials for the middle and long hauling paths, respectively. Similarly, the variables $f^{u,s}_{i,i+1,t}$ and $f^{u,s}_{i,i-1,t}$ (for all $i \in \mathcal{S}, t \in \mathcal{T}$) are the unload (cut) flows of materials to the left and right transit nodes for the short hauling path. In the same way, the variables $f^{u,m}_{i,i+1,t}$, $f^{u,m}_{i,i-1,t}$; and $f^{u,l}_{i,i+1,t}$, $f^{u,l}_{i,i-1,t}$ (for all $i \in \mathcal{S}, t \in \mathcal{T}$) are the unload flows for the middle and long hauling paths, respectively.

Figure 2.3 shows the borrow flows which are the flow of materials from a borrow pit to a section. In some cases, some extra material is required for the construction project which is borrowed from a borrow pit. This means that the material should be cut from the borrow pit. Then, depending on the distance between the borrow pit and destination section, the material is loaded to the appropriate machinery (or hauling path) to be transfered to the destination section. Finally, at the destination section, the material is

unloaded into the section. The borrow flows $f^{b,h}_{j,\vartheta(j)+1,t}$ and $f^{b,h}_{j,\vartheta(j)-1,t}$ (for all $j \in \mathcal{B}, t \in \mathcal{T}, h \in \mathcal{H}$) in Figure 2.3, model this process.

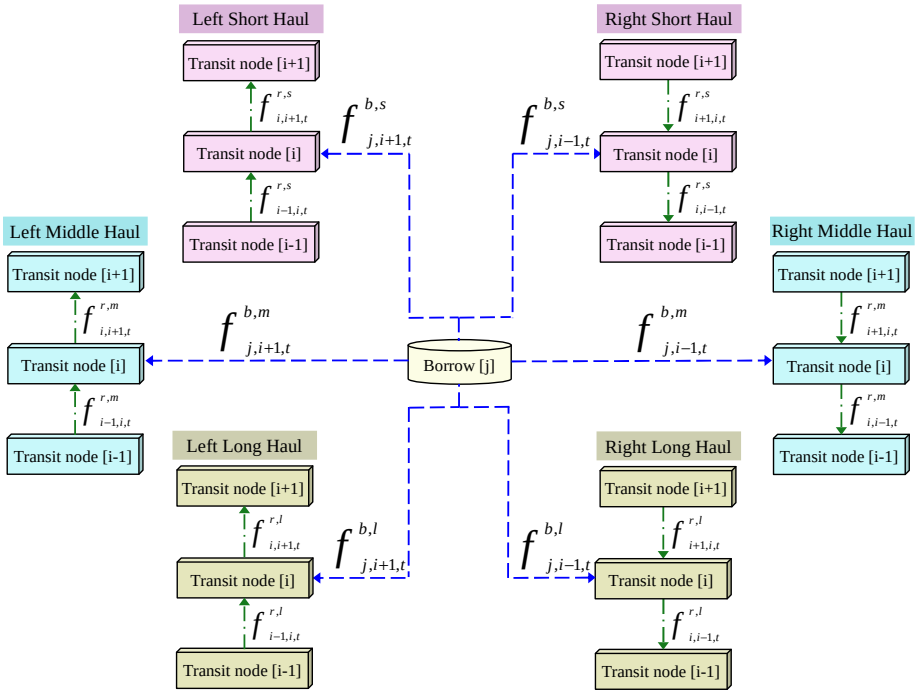

Figure 2.3: Borrow flows in the MH-QNF model

Each borrow pit is attached to a section. In fact, a borrow flow denotes a cut or unload flow from a borrow pit. The variables $f^{b,s}_{j,\vartheta(j)+1,t}$ and $f^{b,s}_{j,\vartheta(j)-1,t}$ (for all $j \in \mathcal{B}, t \in \mathcal{T}$) can be used to transfer materials from a borrow pit to the next or the previous section via the short haul path. Similarly, to represent the borrow flows for the middle and long hauling paths we define the variables $f^{b,m}_{j,\vartheta(j)+1,t}$, $f^{b,m}_{j,\vartheta(j)-1,t}$; and $f^{b,l}_{j,\vartheta(j)+1,t}$ and $f^{b,l}_{j,\vartheta(j)-1,t}$ (for all $j \in \mathcal{B}, t \in \mathcal{T}$), respectively. Waste flows are handled similarly, see Figure 2.4.

In Figure 2.4 the waste flows are shown. In some cases, in the construction site we have some unused material at a section that cannot be used in other sections of the road. In such conditions we can dump the unused material to a waste pit. To do so, we need to move material to the waste pit using the appropriate hauling path (or construction machinery) and then unload and fill the waste pit by the transferred material. Similar to bor-

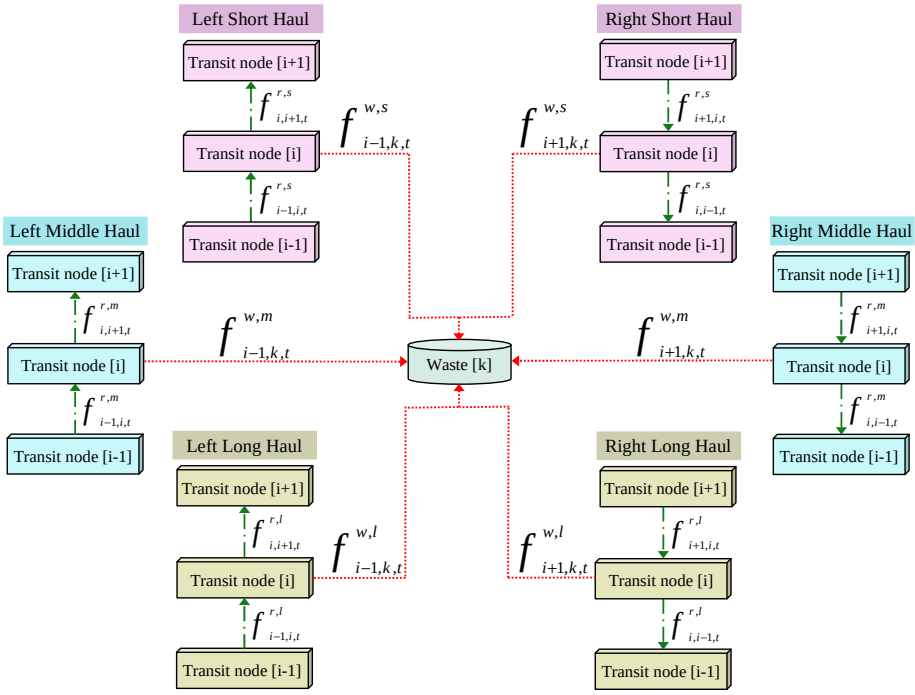

Figure 2.4: Waste flows in the MH-QNF model

row flows, waste flows show the flow of material from a section to a waste pit. Waste pits are attached to specific sections. Material can be moved to a waste pit in both left and right directions via different hauling paths. The variables $f^{w,s}_{\delta(k)-1,k,t}$ and $f^{w,s}_{\delta(k)+1,k,t}$ (for all $k \in \mathcal{W}, t \in \mathcal{T}$) represent the waste flows to the waste pit k from both left and right directions through the short hauling path. Similarly, for the middle and long hauling paths the waste flow variables $f^{w,m}_{\delta(k)-1,k,t}$, $f^{w,m}_{\delta(k)+1,k,t}$; and $f^{w,l}_{\delta(k)-1,k,t}$, $f^{w,l}_{\delta(k)+1,k,t}$ (for all $k \in \mathcal{W}, t \in \mathcal{T}$) are defined, respectively.

A typical real-world road construction project involves several types of materials with different excavation, embankment, and hauling costs. The proposed model is able to handle more than one material. We define the index set $\mathcal{M} = \{1, 2, ..., n_m\}$ corresponding to the indexes of the different types of materials. Different types of material can be excavated using different types of construction machinery. On the other hand, the excavation and embankment costs depend on the material types. Therefore, by adding more material types or changing the excavation costs associated with a specific material type, in our model we are able to model different construction

machineries used to cut the earth. In this research, we chose $n_m = 4$ based on consultation with our industry partner.

2.2.1 Cost components

The cost components of a vertical alignment problem consist of the embankment, excavation, and hauling costs of materials in the construction site. The per unit of volume excavation (cut or unload from a section) cost of material is represented by p_1, p_2, ... p_{n_m} for different types of materials, respectively. Similarly, q_1, q_2, ... q_{n_m} denote the per unit volume of embankment (fill or load to a section) cost of material for the different types of materials. Finally, for the short hauling path, the hauling cost of materials from the section i to the section $i-1$ (respectively $i+1$) is defined as $c^{r,s}_{i,i-1} = c_s d_{i,i-1}$ (respectively, $c^{r,s}_{i,i+1} = c_s d_{i,i+1}$) in which c_s is the cost of moving one unit volume of materials per unit distance for the short hauling path, and $d_{i,j}$ is the distance between sections i and j. Similarly, for the middle hauling path we define the hauling cost components $c^{r,m}_{i,i-1} = c_m d_{i,i-1}$ and $c^{r,m}_{i,i+1} = c_m d_{i,i+1}$; and for the long hauling path we have $c^{r,l}_{i,i-1} = c_l d_{i,i-1}$ and $c^{r,l}_{i,i+1} = c_l d_{i,i+1}$. There is another cost when cutting materials from a section or a borrow pit and it is the loading cost. In the real world the loading cost of different construction equipments is different. Therefore, we define the loading cost components y_s, y_m, and y_l as the cost of loading different types of construction equipments which use short, middle, and long hauling paths, respectively. We do not consider the unloading cost of material from construction equipments because based on consultation with our industrial partner, it is cheap enough to be ignored.

2.2.2 Objective function and constraints

In the MH-QNF model the objective is to minimize the total cost of excavation, embankment, and hauling tasks. Therefore, the objective function is defined as follows.

$$\begin{aligned}\min \quad & \Bigg[\sum_{\substack{i\in\mathcal{S}\\ m\in\mathcal{M}\\ h\in\mathcal{H}}} (p_m + y_h)V_i^+ + \sum_{\substack{i\in\mathcal{S}\\ m\in\mathcal{M}}} q_m V_i^- \\ & + \sum_{\substack{i\in\mathcal{S}\\ h\in\mathcal{H}\\ t\in\mathcal{T}}} (c_{i,i-1}^{r,h} f_{i,i-1,t}^{r,h} + c_{i,i+1}^{r,h} f_{i,i+1,t}^{r,h}) \\ & + \sum_{\substack{j\in\mathcal{B}\\ m\in\mathcal{M}\\ h\in\mathcal{H}\\ t\in\mathcal{T}}} (p_m + y_h + c_h\tilde{d}_j)(f_{j,\vartheta(j)-1,t}^{b,h} + f_{j,\vartheta(j)+1,t}^{b,h}) \\ & + \sum_{\substack{k\in\mathcal{W}\\ m\in\mathcal{M}\\ h\in\mathcal{H}\\ t\in\mathcal{T}}} (q_m + c_h\tilde{d}_k)(f_{\delta(k)-1,k,t}^{w,h} + f_{\delta(k)+1,k,t}^{w,h}) \Bigg]. \end{aligned} \tag{2.2}$$

Since the transit nodes are only intermediate virtual points that provide transits for the flows of materials, the sum of the input flows to a transit node must be equal to the sum of the output flows from the node. Therefore, for all $i \in \mathcal{S}, h \in \mathcal{H}, t \in \mathcal{T}$ the *flow constraints* are defined as

$$f_{i-1,i,t}^{r,h} + f_{i,i+1,t}^{u,h} + \sum_{\substack{j\in\mathcal{B}\\ \vartheta(j)=i}} f_{j,i+1,t}^{b,h} = f_{i,i+1,t}^{r,h} + f_{i-1,i,t}^{l,h} + \sum_{\substack{k\in\mathcal{W}\\ \delta(k)=i}} f_{i-1,k,t}^{w,h}, \tag{2.3}$$

$$f_{i+1,i,t}^{r,h} + f_{i,i-1,t}^{u,h} + \sum_{\substack{j\in\mathcal{B}\\ \vartheta(j)=i}} f_{j,i-1,t}^{b,h} = f_{i,i-1,t}^{r,h} + f_{i+1,i,t}^{l,h} + \sum_{\substack{k\in\mathcal{W}\\ \delta(k)=i}} f_{i+1,k,t}^{w,h}. \tag{2.4}$$

The section nodes are treated as the source and destination of the flows. Therefore, the total sum of load flows to a section or a waste pit should be equal to the total fill volume of materials from that section or waste pit. Similarly, the total sum of unload flows from a section or a borrow pit should be equal to the total cut volume of materials from that section or borrow pit. These constraints are called *balance constraints* and defined as

$$\sum_{t\in\mathcal{T}} (f_{i,i+1,t}^{u,h} + f_{i,i-1,t}^{u,h}) = V_i^+, \quad i \in \mathcal{S}, h \in \mathcal{H}, \tag{2.5}$$

$$\sum_{t\in\mathcal{T}} (f_{j,\vartheta(j)+1,t}^{b,h} + f_{j,\vartheta(j)-1,t}^{b,h}) = V_j^+, \quad j \in \mathcal{B}, h \in \mathcal{H}, \tag{2.6}$$

$$\sum_{t\in\mathcal{T}}(f^{l,h}_{i-1,i,t}+f^{l,h}_{i+1,i,t})=V^{-}_{i},\quad i\in\mathcal{S},h\in\mathcal{H},\tag{2.7}$$

$$\sum_{t\in\mathcal{T}}(f^{w,h}_{\delta(k)-1,k,t}+f^{w,h}_{\delta(k)+1,k,t})=V^{-}_{k},\quad k\in\mathcal{W},h\in\mathcal{H}.\tag{2.8}$$

The borrow and waste pit constraints ensure that the total sum of borrow/waste flows from/to a borrow/waste pit must not exceed the capacity of that borrow/waste, i.e.,

$$\sum_{t\in\mathcal{T}}(f^{b,h}_{j,\vartheta(j)+1,t}+f^{b,h}_{j,\vartheta(j)-1,t})\leq C^{b}_{j},\quad j\in\mathcal{B},h\in\mathcal{H},\tag{2.9}$$

$$\sum_{t\in\mathcal{T}}(f^{w,h}_{\delta(k)-1,k,t}+f^{w,h}_{\delta(k)+1,k,t})\leq C^{w}_{k},\quad k\in\mathcal{W},h\in\mathcal{H}.\tag{2.10}$$

Before introducing block constraints we require some additional notations. Let $\bar{\mathcal{I}}^2=\{(k_1,k_2)\in\mathcal{I}\times\mathcal{I}: k_1<k_2,$ and k_1,k_2 are two consecutive blocks with no access road in between them$\}$, $\bar{\mathcal{I}}_{\leftarrow}=\{k\in\mathcal{I}:$ the set of blocks k before which there is no access road$\}$, and $\bar{\mathcal{I}}_{\rightarrow}=\{k\in\mathcal{I}:$ the set of blocks k after which there is no access road$\}$. To keep track of when a block is removed, we introduce binary variables $y_{k,t}$ (i.e., it can be either 0 or 1) for each block $k\in\mathcal{I}$ and $t\in\mathcal{T}$. We also use the parameter M_i as the largest possible cut or fill volume at a section i. For the sake of simplicity, we use M instead of M_i.

Earth movement is not allowed between two sections when there is a block between them. This means that for each hauling path $h\in\mathcal{H}$, if a section i is a block then material cannot be moved over that block or section. In other words, the whole material moved to the section i from the section $i-1$ must be loaded to section i; i.e., for the left transit nodes $\forall h\in\mathcal{H}$, $f^{r,h}_{i-1,i,t}=f^{l,h}_{i-1,i,t}$. Moreover, if the section i is a block the only material that can move from this section to the next section $(i+1)$ is the material that has been excavated from section i, i.e., for the left transit nodes $\forall h\in\mathcal{H}$, $f^{r,h}_{i,i+1,t}=f^{u,h}_{i,i+1,t}$. In the case of the right transit nodes, $f^{r,h}_{i+1,i,t}=f^{l,h}_{i+1,i,t}$ and $f^{r,h}_{i,i-1,t}=f^{u,h}_{i,i-1,t}$ ensure that no material can move across section i to the previous section. Therefore, for the left transit nodes, the constraints

$$-(f^{r,h}_{\gamma(i)-1,\gamma(i),t}-f^{l,h}_{\gamma(i)-1,\gamma(i),t})\leq My_{i,t-1},\quad\forall i\in\mathcal{I},h\in\mathcal{H},t\in\mathcal{T},\tag{2.11}$$

$$f^{r,h}_{\gamma(i)-1,\gamma(i),t}-f^{l,h}_{\gamma(i)-1,\gamma(i),t}\leq My_{i,t-1},\quad\forall i\in\mathcal{I},h\in\mathcal{H},t\in\mathcal{T},\tag{2.12}$$

$$-(f^{r,h}_{\gamma(i),\gamma(i)+1,t} - f^{u,h}_{\gamma(i),\gamma(i)+1,t}) \leq My_{i,t-1}, \quad \forall i \in \mathcal{I}, h \in \mathcal{H}, t \in \mathcal{T}, \tag{2.13}$$

$$f^{r,h}_{\gamma(i),\gamma(i)+1,t} - f^{u,h}_{\gamma(i),\gamma(i)+1,t}) \leq My_{i,t-1}, \quad \forall i \in \mathcal{I}, h \in \mathcal{H}, t \in \mathcal{T}, \tag{2.14}$$

ensure that no material movement is allowed across a block. In the same way, for the right transit nodes the constraints

$$-(f^{r,h}_{\gamma(i)+1,\gamma(i),t} - f^{l,h}_{\gamma(i)+1,\gamma(i),t}) \leq My_{i,t-1}, \quad \forall i \in \mathcal{I}, h \in \mathcal{H}, t \in \mathcal{T}, \tag{2.15}$$

$$f^{r,h}_{\gamma(i)+1,\gamma(i),t} - f^{l,h}_{\gamma(i)+1,\gamma(i),t} \leq My_{i,t-1}, \quad \forall i \in \mathcal{I}, h \in \mathcal{H}, t \in \mathcal{T}, \tag{2.16}$$

$$-(f^{r,h}_{\gamma(i),\gamma(i)-1,t} - f^{u,h}_{\gamma(i),\gamma(i)-1,t}) \leq My_{i,t-1}, \quad \forall i \in \mathcal{I}, h \in \mathcal{H}, t \in \mathcal{T}, \tag{2.17}$$

$$f^{r,h}_{\gamma(i),\gamma(i)-1,t} - f^{u,h}_{\gamma(i),\gamma(i)-1,t} \leq My_{i,t-1}, \quad \forall i \in \mathcal{I}, h \in \mathcal{H}, t \in \mathcal{T}, \tag{2.18}$$

ensure no material movement across a block.

The other scenario which should be ensured by the block constraints is that material movement is not allowed between two blocks, before the first block, and after the last block with no access roads, until the blocks are removed. Thus, for all $h \in \mathcal{H}$, the transit, borrow, and waste flows are forbidden between two blocks with no access roads, until one of the blocks is removed, i.e.,

$$f^{r,h}_{i,i+1,t} \leq M(y_{k_1,t-1} + y_{k_2,t-1}),$$
$$\forall i \in \mathcal{S}, (k_1,k_2) \in \bar{\mathcal{I}}^2, h \in \mathcal{H}, t \in \mathcal{T}, \gamma(k_1) \leq i, i+1 \leq \gamma(k_2), \tag{2.19}$$

$$f^{r,h}_{i+1,i,t} \leq M(y_{k_1,t-1} + y_{k_2,t-1}),$$
$$\forall i \in \mathcal{S}, (k_1,k_2) \in \bar{\mathcal{I}}^2, h \in \mathcal{H}, t \in \mathcal{T}, \gamma(k_1) \leq i, i+1 \leq \gamma(k_2), \tag{2.20}$$

$$f^{b,h}_{j,\vartheta(j)+1,t} \leq M(y_{k_1,t-1} + y_{k_2,t-1}),$$
$$\forall j \in \mathcal{B}, (k_1,k_2) \in \bar{\mathcal{I}}^2, h \in \mathcal{H}, t \in \mathcal{T}, \gamma(k_1) \leq \vartheta(j)-1, \vartheta(j)+1 \leq \gamma(k_2), \tag{2.21}$$

$$f^{b,h}_{j,\vartheta(j)-1,t} \leq M(y_{k_1,t-1} + y_{k_2,t-1}),$$
$$\forall j \in \mathcal{B}, (k_1,k_2) \in \bar{\mathcal{I}}^2, h \in \mathcal{H}, t \in \mathcal{T}, \gamma(k_1) \leq \vartheta(j)-1, \vartheta(j)+1 \leq \gamma(k_2), \tag{2.22}$$

$$f^{w,h}_{\delta(j)-1,j,t} \leq M(y_{k_1,t-1} + y_{k_2,t-1}),$$
$$\forall j \in \mathcal{W}, (k_1,k_2) \in \bar{\mathcal{I}}^2, h \in \mathcal{H}, t \in \mathcal{T}, \gamma(k_1) \leq \delta(j)-1, \delta(j)+1 \leq \gamma(k_2), \tag{2.23}$$

$$f^{w,h}_{\delta(j)+1,j,t} \leq M(y_{k_1,t-1} + y_{k_2,t-1}),$$
$$\forall j \in \mathcal{W}, (k_1,k_2) \in \bar{\mathcal{I}}^2, h \in \mathcal{H}, t \in \mathcal{T}, \gamma(k_1) \leq \delta(j)-1, \delta(j)+1 \leq \gamma(k_2). \tag{2.24}$$

After the last block with no access road, the transit, borrow, and waste flows are not allowed until the block is removed:

$$f^{r,h}_{i,i+1,t} \leq M y_{k,t-1}, \quad \forall i \in \mathcal{S}, k \in \bar{\mathcal{I}}_{\rightarrow}, h \in \mathcal{H}, t \in \mathcal{T}, \gamma(k) \leq i, i+1 \leq n, \tag{2.25}$$

$$f^{r,h}_{i+1,i,t} \leq M y_{k,t-1}, \quad \forall i \in \mathcal{S}, k \in \bar{\mathcal{I}}_{\rightarrow}, h \in \mathcal{H}, t \in \mathcal{T}, \gamma(k) \leq i, i+1 \leq n, \tag{2.26}$$

$$f^{b,h}_{j,\vartheta(j)+1,t} \leq M y_{k,t-1},$$
$$\forall j \in \mathcal{B}, k \in \bar{\mathcal{I}}_{\rightarrow}, h \in \mathcal{H}, t \in \mathcal{T}, \gamma(k) \leq \vartheta(j)-1, \vartheta(j)+1 \leq n, \tag{2.27}$$

$$f^{b,h}_{j,\vartheta(j)-1,t} \leq M y_{k,t-1},$$
$$\forall j \in \mathcal{B}, k \in \bar{\mathcal{I}}_{\rightarrow}, h \in \mathcal{H}, t \in \mathcal{T}, \gamma(k) \leq \vartheta(j)-1, \vartheta(j)+1 \leq n, \tag{2.28}$$

$$f^{w,h}_{\delta(j)-1,j,t} \leq M y_{k,t-1},$$
$$\forall j \in \mathcal{W}, k \in \bar{\mathcal{I}}_{\rightarrow}, h \in \mathcal{H}, t \in \mathcal{T}, \gamma(k) \leq \delta(j)-1, \delta(j)+1 \leq n, \tag{2.29}$$

$$f^{w,h}_{\delta(j)+1,j,t} \leq M y_{k,t-1},$$
$$\forall j \in \mathcal{W}, k \in \bar{\mathcal{I}}_{\rightarrow}, h \in \mathcal{H}, t \in \mathcal{T}, \gamma(k) \leq \delta(j)-1, \delta(j)+1 \leq n. \tag{2.30}$$

In the same way, before the first block with no access road, the transit, borrow, and waste flows are not allowed until the block is removed:

$$f^{r,h}_{i,i+1,t} \leq M y_{k,t-1}, \quad \forall i \in \mathcal{S}, k \in \bar{\mathcal{I}}_{\leftarrow}, h \in \mathcal{H}, t \in \mathcal{T}, 1 \leq i, i+1 \leq \gamma(k), \tag{2.31}$$

$$f^{r,h}_{i+1,i,t} \leq M y_{k,t-1}, \quad \forall i \in \mathcal{S}, k \in \bar{\mathcal{I}}_{\leftarrow}, h \in \mathcal{H}, t \in \mathcal{T}, 1 \leq i, i+1 \leq \gamma(k), \tag{2.32}$$

$$f^{b,h}_{j,\vartheta(j)+1,t} \leq M y_{k,t-1}, \quad \forall j \in \mathcal{B}, k \in \bar{\mathcal{I}}_{\leftarrow}, h \in \mathcal{H}, t \in \mathcal{T}, 1 \leq \vartheta(j)-1, \vartheta(j)+1 \leq \gamma(k), \tag{2.33}$$

$$f^{b,h}_{j,\vartheta(j)-1,t} \leq M y_{k,t-1}, \quad \forall j \in \mathcal{B}, k \in \bar{\mathcal{I}}_{\leftarrow}, h \in \mathcal{H}, t \in \mathcal{T}, 1 \leq \vartheta(j)-1, \vartheta(j)+1 \leq \gamma(k), \tag{2.34}$$

$$f^{w,h}_{\delta(j)-1,j,t} \leq M y_{k,t-1}, \quad \forall j \in \mathcal{W}, k \in \bar{\mathcal{I}}_{\leftarrow}, h \in \mathcal{H}, t \in \mathcal{T}, 1 \leq \delta(j)-1, \delta(j)+1 \leq \gamma(k), \tag{2.35}$$

$$f^{w,h}_{\delta(j)+1,j,t} \leq M y_{k,t-1}, \quad \forall j \in \mathcal{W}, k \in \bar{\mathcal{I}}_{\leftarrow}, h \in \mathcal{H}, t \in \mathcal{T}, 1 \leq \delta(j)-1, \delta(j)+1 \leq \gamma(k). \tag{2.36}$$

After excavating/embanking the required amount of earth from/to a section with a block, we should consider that block as a removed block which does not exist anymore. The *block removal indicator constraints* are used to satisfy this expectation as follows

$$\sum_{t=0}^{n_b} (f^{u,h}_{\gamma(k),\gamma(k)+1,t} + f^{u,h}_{\gamma(k),\gamma(k)-1,t}) + M(1-y_{k,t}) \geq V^{+}_{\gamma(k)}, \quad \forall k \in \mathcal{I}, h \in \mathcal{H}, t \in \mathcal{T}, \tag{2.37}$$

$$\sum_{t=0}^{n_b} (f^{l,h}_{\gamma(k)-1,\gamma(k),t} + f^{l,h}_{\gamma(k)+1,\gamma(k),t}) + M(1-y_{k,t}) \geq V^{-}_{\gamma(k)}, \quad \forall k \in \mathcal{I}, h \in \mathcal{H}, t \in \mathcal{T}. \tag{2.38}$$

At the end of each time step $t \in \mathcal{T}$, at least one block should be removed so that removing all the blocks does not take more than $n_b + 1$ time steps. The *block removal enforcement* constraint guarantees this,

$$\sum_{t=0}^{u} \sum_{k \in \mathcal{I}} y_{k,t} \geq t+1 \quad \forall t \in \mathcal{T} \setminus \{n_b\}. \tag{2.39}$$

In addition, when a block is removed it should remain removed. This is done by the *monotonicity constraint* as follows

$$y_{k,t} \geq y_{k,t-1}, \quad \forall k \in \mathcal{I}, t \in \mathcal{T} \setminus \{0\}. \tag{2.40}$$

The continuity constraints,

$$P_{g-1}(s_{\varphi(g,1)}) = P_g(s_{\varphi(g,1)}) \quad \forall g \in \mathcal{G} \setminus \{1\}, \tag{2.41}$$

$$P'_{g-1}(s_{\varphi(g,1)}) = P'_g(s_{\varphi(g,1)}) \quad \forall g \in \mathcal{G} \setminus \{1\}, \tag{2.42}$$

are used to ensure that the height and the slope of the first section of a segment is equal to the height and the slope of the last section of the previous segment.

The *volume constraints*

$$V_i^+ - V_i^- = A_i u_i, \tag{2.43}$$

in which A_i is the area of section i, guarantee that the total excavated/ embanked volume of material from/to a section equals the volume difference between the road profile and the ground profile for that section.

There are restrictions on the grade of the road profile to satisfy safety considerations. *Slope constraints*

$$G_L \leq P'_g(s_{\varphi(g,1)}) \leq G_U \quad \forall g \in \mathcal{G} \setminus \{1\}, \tag{2.44}$$

are used to constrain the spline segments within a closed interval $[G_L, G_U]$, in which G_L and G_U are the minimum and the maximum valid grades, respectively.

Finally, We need *bound constraints* to restrict the domain of variables as

$$0 \leq V_i^+ \leq M, \quad \forall i \in \mathcal{S} \cup \mathcal{B}, \tag{2.45}$$

$$0 \leq V_i^- \leq M, \quad \forall i \in \mathcal{S} \cup \mathcal{W}, \tag{2.46}$$

$$f^{r,h}_{i,i+1,t} \geq 0, \quad f^{r,h}_{i,i-1,t} \geq 0 \quad \forall i \in \mathcal{S}, h \in \mathcal{H}, t \in \mathcal{T}, \tag{2.47}$$

$$f^{l,h}_{i-1,i,t} \geq 0, \quad f^{l,h}_{i+1,i,t} \geq 0 \quad \forall i \in \mathcal{S}, h \in \mathcal{H}, t \in \mathcal{T}, \tag{2.48}$$

$$f^{u,h}_{i,i+1,t} \geq 0, \quad f^{u,h}_{i,i-1,t} \geq 0 \quad \forall i \in \mathcal{S}, h \in \mathcal{H}, t \in \mathcal{T}, \tag{2.49}$$

$$f^{b,h}_{j,\vartheta(j)+1,t} \geq 0, \quad f^{b,h}_{j,\vartheta(j)-1,t} \geq 0 \quad \forall j \in \mathcal{B}, h \in \mathcal{H}, t \in \mathcal{T}, \tag{2.50}$$

$$f^{w,h}_{\delta(j)-1,j,t} \geq 0, \quad f^{w,h}_{\delta(j)+1,j,t} \geq 0 \quad \forall j \in \mathcal{W}, h \in \mathcal{H}, t \in \mathcal{T}. \tag{2.51}$$

2.3 The CTG and QNF Models

In this section we describe two models that solve the vertical alignment problem for a known horizontal alignment. In the next chapter, we will use these models to evaluate the performance of MH-QNF model.

The *Complete Transportation Graph* model (henceforth CTG model) is a mixed integer linear programming model which was first developed by Hare et al. in [HLR15]. This model is based on the spline technique proposed by Moreb [Mor09] in which the road profile is modeled using piecewise polynomials and the vertical alignment problem is solved as a single linear program. The CTG model was improved in [HHLR14] by adding some speedup techniques to improve the computational time.

We use the CTG model as a benchmark model in our experiments. We do not refer to the details of this model in our research. Full details of the CTG model, which appear in [HHLR14, HLR15], are technical. Therefore, in this discussion, we aim to provide only a broad stroke overview. In essence, instead of hauling paths, for every pair of nodes (i, j), the CTG model places an arc moving from i to j. The flow through the arc is a continuous variable and the arc cost is computed based on i and j. Figure 2.5 demonstrates an example of CTG model.

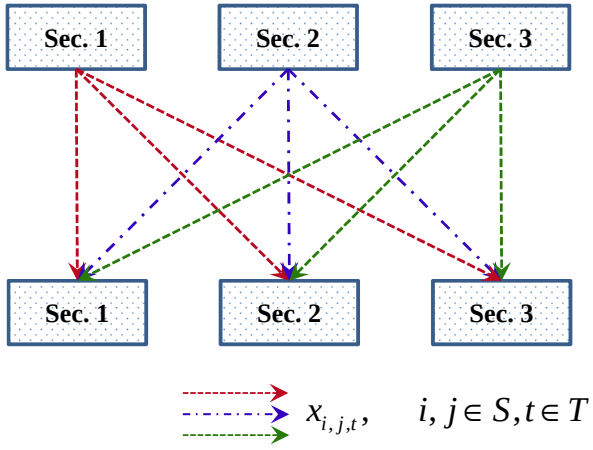

Figure 2.5: CTG model for a three-sections-long road.

A complete transportation graph for a road with n sections, is a graph that includes n source nodes, n destination nodes, and an edge between each pair of source/destination nodes. The CTG model allocates a variable for each edge so the required number of material movement variables is $O(n_s^2)$ in

which n_s is the number of sections. This model was extended in [HHLR14] by adding blocks and side slopes to the mixed integer linear programming model proposed by Hare et al. [HLR15].

For ease of explanation, we introduce variables for material movement without considering blocks, which can be introduced via time steps, similar to their use in MH-QNF. The full description of the model can be found in [HHLR14, HLR15]. To track material movement, we use the variable x_{ij}, which shows the amount of material moved from section i to section j. For each $i \in \mathcal{N}$, we use the index set $\mathcal{N}^i_{\rightarrow}$ which contains all indices j so that x_{ij} is a valid move:

$$\mathcal{N}^i_{\rightarrow} = \begin{cases} j \in \mathcal{S} \cup \mathcal{W} & \text{if } i \in \mathcal{S} \\ j : \quad j \in \mathcal{S} & \text{if } i \in \mathcal{B} \\ \emptyset & \text{if } i \in \mathcal{W} \end{cases} \tag{2.52}$$

Similarly, for each $i \in \mathcal{N}$, we use the index set $\mathcal{N}^i_{\leftarrow}$ which contains all indices j so that x_{ij} is a valid move:

$$\mathcal{N}^i_{\leftarrow} = \begin{cases} j \in \mathcal{S} \cup \mathcal{W} & \text{if } i \in \mathcal{S} \\ j : \quad \emptyset & \text{if } i \in \mathcal{B} \\ j \in \mathcal{S} & \text{if } i \in \mathcal{W} \end{cases} \tag{2.53}$$

Note that for any $i, j \in \mathcal{N}$, $j \in \mathcal{N}^i_{\rightarrow}$ if and only if $i \in \mathcal{N}^j_{\leftarrow}$. Finally, $\mathcal{N}^2$ consists of all index pairs (i, j) such that x_{ij} is a valid move:

$$\mathcal{N}^2 = \{(i, j) : j \in \mathcal{N}^i_{\rightarrow}\}. \tag{2.54}$$

Note that $\mathcal{N}^2 \neq \mathcal{N} \times \mathcal{N}$, as $\mathcal{N}^2$ omits index pairs (i, j) that correspond to illogical moves, such as movement from a road section into a borrow pit.

Similar to the other two models, the cost of vertical alignment depends on the embankment, excavation and hauling of materials from section to section. Let p_i represents the excavation cost per unit of volume and q_i represents the embankment cost per unit of volume for section $i \in \mathcal{N}$. Let r_{ij} be the hauling cost per unit of volume from section i to section j.

The objective function of the CTG model is to minimize the total excavation, embankment, and hauling costs when moving materials between sections, borrow pits, and waste pits over all time steps. The objective function can be represented as follows:

$$\min \Big(\sum_{i \in \mathcal{S} \cup \mathcal{B}} p_i V_i^+ + \sum_{i \in \mathcal{S} \cup \mathcal{W}} q_i V_i^- + \sum_{\substack{(i,j) \in \mathcal{N}^2 \\ t \in \mathcal{T}}} r_{ij} x_{ijt} \Big), \tag{2.55}$$

where V_i^+ (for $i \in \mathcal{S} \cup \mathcal{B}$), and V_i^- (for $i \in \mathcal{S} \cup \mathcal{W}$) represent the total cut and fill volumes of material respectively.

The constraints are similar to the constraints explained for MH-QNF model, which are adapted based on CTG model idea.

The *Quasi Network Flow* model (henceforth QNF model) is studied in [HHLR14]. It reduces the number of variable in the CTG model from quadratic to linear, but at the cost of a more restrictive objective function. The QNF model is quite easy to describe at this stage – the QNF model is a special case of the MH-QNF model, where the number of haul types is one: $n_h = 1$. Figure 2.6 illustrates a sample QNF model.

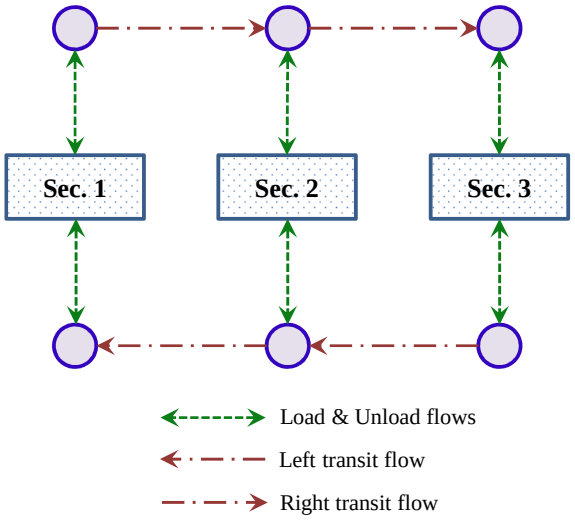

Figure 2.6: QNF model for a three-sections-long road.

For comparison, in Figure 2.7, the MH-QNF version of the road shown in Figures 2.5 and 2.6 is presented, where multiple hauling paths are employed to move materials. This figure seems more reasonable for longer roads where we can differentiate short, middle, and long distances. The examples presented here are only to show differences between three types of models introduced earlier.

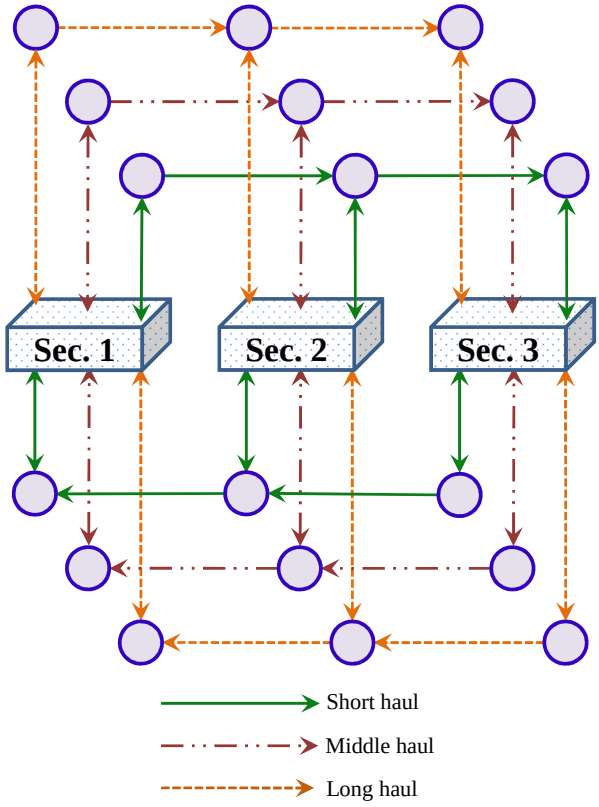

Figure 2.7: MH-QNF model for a three-sections-long road.

Chapter 3

Numerical results

In this chapter, we present our experimental data and discuss the model performance on real-world problems. To evaluate the performance of the MH-QNF model, we compare the efficiency, robustness and accuracy of the proposed model with both CTG and QNF models in terms of timing and optimal cost. The efficiency refers to the computational effort required to obtain a solution. We compare the efficiency of models using solving time. Robustness of an optimization method is defined as the ability of the method to perform well over a wide range of problems. We compare the robustness of the models using the success rate measure as will be discussed later. Finally, the accuracy is the quality of the solution obtained by solving a model. We evaluate the accuracy by comparing the relative errors in cost values obtained by all the models.

3.1 Experimental setup

In a fair comparison of optimization methods, the first step is to assemble or choose a proper problem test set [Bei15]. The selected problem test collection consists of 60 problems. These problems are generated by changing various parameters (i.e., number of sections, blocks, access roads, offset levels and section lengths) in 7 distinct road samples, provided by our industrial partner (see Table 3.1). All tests represent realistic road design problems. In the test collection, the section length and the number of sections in a road remain unchanged in all the problems.

We consider 5 hours timeout for to solve each individual test problem. Because of limitations on the computational power, 5 hours was the highest computational time we could spend on solving each problem to finish all the experiments in a reasonable time. If a test cannot be solved within this time limitation, then it is reported as unsuccessful. Linear programming feasibility tolerance and relative mixed integer programming gap tolerance parameters are set to 10^{-06} and 1%, respectively. The upper bound and lower bound for slope constraints are 0.1 and -0.1. We use the default settings for the other parameters for solver and models.

Table 3.1: Basic roads used to assemble the test collection.

Road	Length (km)	Section Length (m)	Number of sections
A	1	20	50
B	5	100	50
C	2	20	100
D	3	20	150
E	15	100	150
F	20	100	200
G	9	20	450

After selection of the test collection, we solved the problems by different models using the same computational environment. The workstation used for experiments has an Intel(R)Xeon(R) CPU W3565 3.20GHz (4 cores and 8 threads) processor and 24.0 GB of RAM. To solve the developed mathematical models we use the academic version of the IBM ILOG CPLEX optimizer 12.51 edition `http://www.cplex.com`.

Table 3.2: Cost components in $\$/cu.m.$

Cost component	Cost Type	Value	Avg Value
Excavation cost	M1	4.000	
	M2	4.000	
	M3	20.000	
	M4	4.000	
Embankment cost	M1	2.000	
	M2	2.000	
	M3	1.800	
	M4	2.000	
Loading cost	Short	0.000	1.067
	Middle	0.600	
	Long	2.600	
Hauling cost	Short	0.008	0.005
	Middle	0.004	
	Long	0.002	

One of the influencing factors on our experiments is the cost components used in the MH-QNF model. Table 3.2 shows the cost components

involved in a vertical alignment problem and the associated values selected in consultation with our industry partner. Excavation and embankment cost components depend on the different types of materials. Loading and hauling cost components depend on the type of hauling path. The Avg Value column is the average of the cost values for a cost component. All the three models are adapted and parametrized using the same cost components. Both MH-QNF model and the CTG model use multiple hauling paths with different hauling and loading costs. Therefore, in the case of the QNF model which has a single hauling path, for the loading and hauling cost components we need to test the model with different pairs of hauling and loading cost components.

Since the MH-QNF model is an extension of the QNF model, to get a better understanding of the importance of loading and hauling cost components, in addition to the average values of loading and hauling costs (1.067, 0.005, respectively), we test the QNF model with the other three pairs of cost values for loading and hauling cost components reported for different hauls (Table 3.2); i.e., we run the QNF model four times with different loading and hauling costs: average case (1.067, 0.005), short haul (0.000, 0.008), middle haul (0.600, 0.004), and long haul (2.600, 0.002). Thus, we have 4 different setups for QNF model as follows.

- **QNF-S**: the model QNF configured with short hauling and loading cost values.
- **QNF-M**: the model QNF configured with middle hauling and loading cost values.
- **QNF-L**: the model QNF configured with long hauling and loading cost values.
- **QNF-A**: the model QNF configured with average hauling and loading cost values.

As shown in Figure 3.1, to evaluate the efficiency of using a variety of hauling paths over a single path in a road construction site, the loading and hauling cost components are selected such that

- if **hauling distance** $< 150m$ then short hauling path will be picked as the optimum method of transferring material;
- if $150m \leq$ **hauling distance** $< 1000m$ then middle hauling path will be picked;

– and finally, if **hauling distance** $\geq 1000m$ then the long hauling path will be selected.

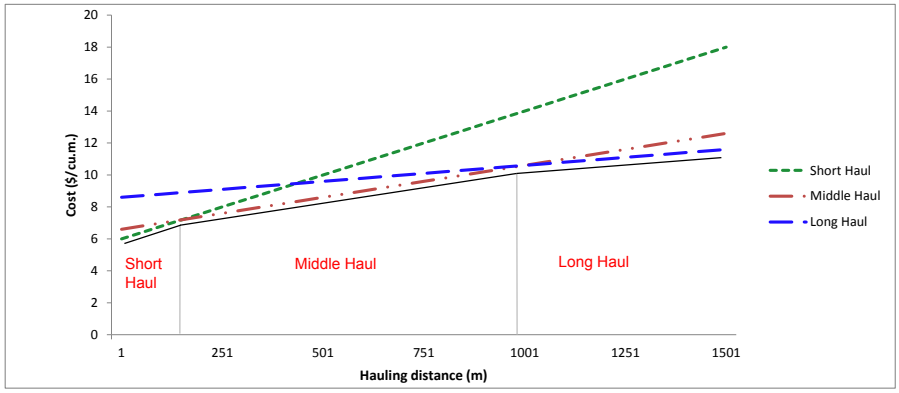

Figure 3.1: Cost configuration.

In [HHLR14] the authors proposed six potential techniques to improve the speed of solution search. They extensively tested all the techniques with two models and identified the setup which works the best for both models. Their results highlighted two techniques as the most efficient methods. Therefore, we test all models using two different configurations. In the first configuration, we use SOS2 variables for the volume constraints and SOS1 variables for modeling the block constraints in all the models. In the second configuration, again we use SOS2 variables for the volume constraints but for the block constraints the basic technique is employed.

In general, in our experiments we use 12 model configurations. To increase the readability of the report on the results we use the following naming convention.

1. Models that use the *basic* technique for blocks:
   - **MH-QNF-B**: The MH-QNF model.
   - **CTG-B**: The CTG model with multi haul cost components.
   - **QNF-SB**: The QNF-S model.
   - **QNF-MB**: The QNF-M model.
   - **QNF-LB**: The QNF-L model.
   - **QNF-AB**: The QNF-A model.
2. Models that use the *SOS1* techniques for blocks:

- **MH-QNF-S**: The MH-QNF model.
- **CTG-S**: The CTG model with multi haul cost components.
- **QNF-SS**: The QNF-S model.
- **QNF-MS**: The QNF-M model.
- **QNF-LS**: The QNF-L model.
- **QNF-AS**: The QNF-A model.

3.2 Reporting the results

As a first pass at examining model accuracy, we summarize the relative errors in cost values obtained by all models in Table 3.3. The column "opt. found" is the number of problems where the model optimum value was found before timeout occurred (regardless of relative error). The columns "min./mean/max. error" are the minimum/mean/maximum relative error over all problems solved when compared to CTG configurations (CTG-B or CTG-S), based on only problems solved by both CTG configuration and the model examined.

Table 3.3: Summary of model accuracy for each model.

#	Model	opt. found	min. error (%)	mean error (%)	max. error (%)
1	CTG-B	51	–	–	–
2	MH-QNF-B	56	0.00	0.47	2.20
3	QNF-SB	56	27.15	57.10	100.92
4	QNF-MB	56	12.33	25.65	40.74
5	QNF-LB	56	18.93	25.81	37.95
6	QNF-AB	56	22.19	40.22	62.98
7	CTG-S	45	–	–	–
8	MH-QNF-S	57	0.00	0.37	1.43
9	QNF-SS	56	27.15	57.56	100.92
10	QNF-MS	56	12.45	25.78	40.74
11	QNF-LS	56	18.93	25.81	37.39
12	QNF-AS	56	22.36	40.44	62.98

As we expected, among the above 12 model configurations, none of the QNF model configurations are able to solve a single problem satisfactorily. The reason is that the QNF model considers a linear relation between the

hauling (and loading) cost and the hauling distance. In other words, QNF assumes that only one type of machinery is used to transfer materials in the construction site, regardless of the hauling distance. We therefore discard these methods as they are inaccurate.

Table 3.4 shows the solution times for the remaining 4 configurations applied to 60 test problems when only 1% relative error is allowed. "NaN" values indicate that the solver exceeded the 5 hours timeout.

The reliability and robustness of an optimization method is defined as the ability of the method to perform well over a wide range of optimization problems. To evaluate the reliability of the model we use two measures: success rate and computational accuracy. The success rate is defined as the number of problems in the given problem set that are successfully solved to optimality by the optimization method. We use performance profiles [DM02] to evaluate the reliability and robustness of the models. While performance profiles are now prevalent in optimization benchmarking, we provide a brief review for completeness. The performance profile of a solver s is defined as follows

$$\rho_s(\alpha) = \frac{1}{|\mathcal{P}|}\text{size}\{p \in \mathcal{P} : r_{p,s} \leq \alpha\}, \tag{3.1}$$

where $|\mathcal{P}|$ represents the cardinality of the problem set $\mathcal{P}$ and $r_{p,s}$ is the performance ratio defined as

$$r_{p,s} = \begin{cases} \frac{t_{p,s}}{\min\{t_{p,s}:s\in\mathcal{S}\}} & \text{if the convergence test passed} \\ \infty & \text{if the convergence test failed} \end{cases} \tag{3.2}$$

in which $t_{p,s}$ is the performance measure which is the computational time in our case. In this equation, for a specific problem and the best solver in terms of $t_{p,s}$ we will have $r_{p,s} = 1$.

One of the advantages of performance profiles, is that they implicitly includes the performance ratio *success rate* as a reliability factor. The value of $\rho_s(\alpha)$ gives a sense of how promising the solutions found by an optimization algorithm are relative to the best solution found by all the optimization algorithms that are compared together.

Table 3.4: Solution times (in second) for 1% relative error.

Test #	MH-QNF-B	CTG-B	MH-QNF-S	CTG-S	Test #	MH-QNF-B	CTG-B	MH-QNF-S	CTG-S
1	23.84	19.02	15.21	33.71	31	9.89	179.33	9.83	35.04
2	11.00	27.18	15.28	35.41	32	18.15	619.50	78.73	241.43
3	26.51	190.64	20.51	237.69	33	11.03	1361.16	25.27	675.14
4	168.62	7214.67	172.48	4848.31	34	30.08	2385.43	91.46	NaN
5	42.07	434.39	52.39	452.03	35	22.71	2463.64	35.74	2543.56
6	13.05	232.84	28.26	1014.80	36	120.08	NaN	273.13	9350.41
7	874.97	603.49	581.46	1928.96	37	57.23	1702.74	90.71	NaN
8	NaN	12605.20	2409.50	NaN	38	3.12	4.96	4.95	4.98
9	19.93	169.51	78.17	309.46	39	17.09	230.50	36.15	623.73
10	13.40	297.75	17.13	141.41	40	18020.00	NaN	NaN	NaN
11	66.34	2348.93	60.28	NaN	41	3361.52	NaN	9224.50	NaN
12	146.36	3929.18	124.01	9853.74	42	392.59	NaN	455.71	NaN
13	4.96	86.70	11.38	33.63	43	308.46	NaN	313.39	NaN
14	2.15	39.36	4.97	17.24	44	NaN	NaN	NaN	NaN
15	18.75	392.67	35.29	451.63	45	70.47	2014.47	168.04	NaN
16	319.90	NaN	366.57	NaN	46	50.75	1289.76	122.84	3528.95
17	13.81	81.61	23.76	174.21	47	8.85	7.44	17.43	7.78
18	238.37	1369.81	383.95	4559.33	48	5.70	9.02	10.61	9.76
19	30.17	799.55	68.35	1657.61	49	3.22	1.87	6.72	1.89
20	81.99	632.69	1175.49	4146.94	50	1.36	0.94	2.46	0.99
21	455.83	NaN	2743.00	NaN	51	4.83	6.11	7.72	6.14
22	223.27	8901.20	1426.83	NaN	52	8.97	28.74	14.42	28.87
23	65.00	1306.09	553.65	2437.76	53	14.37	40.89	23.76	41.09
24	3.65	91.19	7.87	55.34	54	13.58	27.02	21.90	27.86
25	9.64	750.86	25.49	314.50	55	13.68	30.44	23.01	31.26
26	76.04	5486.19	158.04	NaN	56	3.78	4.74	6.34	4.68
27	6.87	353.50	14.64	57.91	57	73.50	1373.68	94.70	1361.67
28	6.16	493.60	14.19	248.43	58	NaN	NaN	NaN	NaN
29	5.52	67.06	9.20	19.16	59	15.24	122.75	15.90	128.76
30	30.44	2454.93	51.14	2432.61	60	6.99	14.09	8.58	16.29

For each solution obtained for problem s using the model m the percent relative error is calculated using the following formula:

$$E_{s,m} = \frac{Obj_s - Obj_b}{Obj_b}, \tag{3.3}$$

in which Obj_s is the optimal objective value obtained for the problem s by the model m and Obj_b is the objective value obtained by the benchmark model b. In our experiments, a solution found by model m is considered *successful* if $|E_{s,m}| \leq 1.0\%$.

When available, we consider the CTG model as the benchmark and we calculate the relative error for the solutions obtained using the MH-QNF model. When the CTG model does not provide a solution but the MH-QNF does, we consider the MH-QNF relative error to be 0.0%. This is done since only 5 problems out of 60 have such condition (solved by MH-QNF but not CTG). Adding or removing these 5 problems does not change the final conclusions of the experiments. Moreover, as discussed previously, the accuracy of MH-QNF model is pretty close to CTG in all the problems solved by both models. Therefore, we decided to trust MH-QNF model in the case of these 5 test problems. Figure 3.2 provides the resulting performance profile.

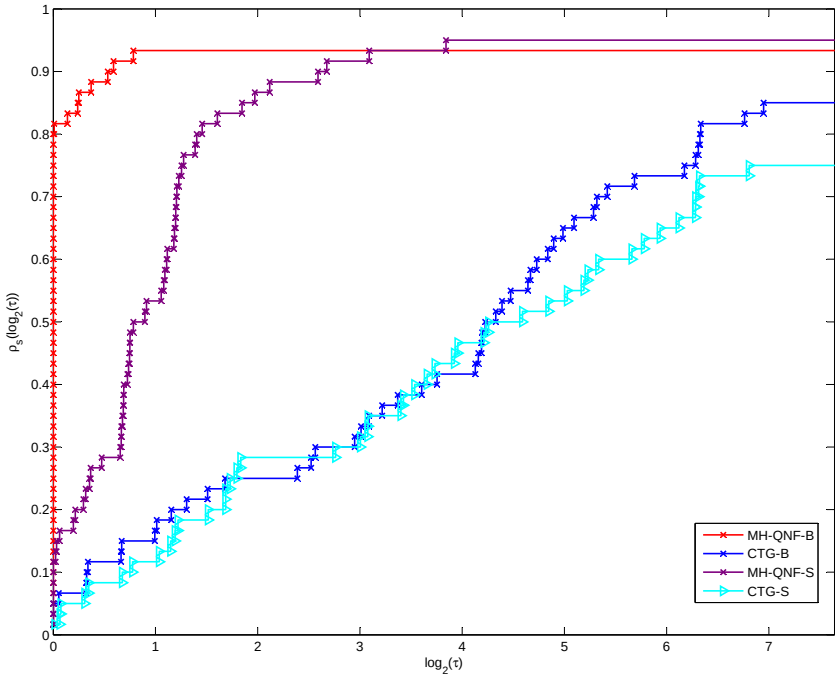

Figure 3.2: Performance profiles with up to 1% relative error.

Examining Figure 3.2, it is clear that MH-QNF greatly outperforms CTG. In terms of the block technique, the CTG model has a relatively better performance when it uses the basic technique for blocks by solving almost 82% of the problems compared to the SOS1 version (CTG-S) which ends up with solving 75%. Similarly, in the case of MH-QNF model, the model using the basic technique (MH-QNF-B) is considerably faster than when it uses the SOS1 technique (MH-QNF-S). Except on one test, MH-QNF-B solves most of the problems faster than MH-QNF-S. In general, among all the configurations, MH-QNF-B is the winner and has a better performance than the CTG configurations.

Chapter 4

Conclusions

In this research, we extended a model based on a previous quasi network flow model by Hare et al. [HHLR14] for the vertical alignment problem considering the earthwork cost. In particular, we added multiple hauling flows to the model in order to improve the accuracy of the model and provide more modeling flexibility for the users. The model considers features such as blocks and side-slopes. We compared the model with two state-of-the-art models in the field: CTG and QNF models.

The CTG model allows a great deal of flexibility in arc-costs, but the resulting number of variables is very large. The total number of variables grows quadratically with the number of sections: $O((n_s)^2 n_h n_m)$, where n_s, n_h and n_m are the number of sections, hauling paths, and material types, respectively [HHLR14]. As the number of sections is typically the dominating factor in these models, the CTG model is typically much larger than the QNF and MH-QNF models.

In particular, in the MH-QNF model (and hence the QNF model) the total number of variables grows linearly with the number of sections: $O(n_s)$ [HLR15]. The number of variables in MH-QNF model is $O(n_h * n_s)$, which is still linear in n_s. (Note that n_h is typically 3, so have little impact on problem size.) However, in QNF model there is a linear relation between the hauling cost and the distance traveled by the volume of materials (e.g., the cost of moving 1 unit of earth for $100m$ equals to 100 times the cost of moving 1 unit for $1m$), which is not true in real world. Therefore, we proposed the MH-QNF model to integrate the low computational cost of QNF model with the accuracy of CTG in solving the vertical alignment problem. The experimental results confirm that the new model is both highly accurate and considerably faster than the CTG model.

4.1 Future research

4.1.1 Three dimensional road design optimization

In three dimensional alignment optimization, the vertical and horizontal alignments are optimized, simultaneously. The horizontal alignment optimization problem is related to the vertical alignment optimization problem. Often, three dimensional optimization involves repeatedly solving the vertical alignment problem. As discussed in previous chapters, the proposed model in this thesis (i.e., MH-QNF model) accurately solves the vertical alignment optimization problem and is computationally cheaper than CTG model. Therefore, as a future work MH-QNF model can be used in three dimensional optimization process to get a precise final solution.

4.1.2 Model uncertainty

In the proposed model there are some parameter values (such as the length of segments, the number of sections in a segment, and the number of blocks in a road) that we have not studied with respect to their importance in the final solution provided by the model. These parameters are subject to change and error. As a future work we can use *model uncertainty*, to investigate these potential changes and errors and their impacts on the conclusions to be drawn from the model.

4.1.3 Human-based experiment

The models developed for road design optimization problems are practically used by civil engineers. The sophisticated nature of road design problem imposes a lot of constraints that need to be dealt with in order to provide a practical solution to engineers who use the model. In this regard, for the sake of providing a quality solution, it is required to design new experiments in order to reveal potential unknown constraints in the real-world road design problem which is not handled by the model proposed in this research. As a future experiment it is suggested to design a test and solve the test set by both automated optimization tool and an experienced engineer. Then we can evaluate the results obtained in both cases, which might provide more insight on the strength and quality of the developed model.

Bibliography

[AAS11] AASHTO, editor. *A Policy on Geometric Design of Highways and Streets*. American Association of State Highway and Transportation Officials (AASHTO), 6th edition, 2011. → pages 2

[Aru05] K. Aruga. Tabu search optimization of horizontal and vertical alignments of forest roads. *Journal of Forest Research*, 10(4):275 – 284, 2005. → pages 3

[Bei15] V. Beiranvand. A survey on benchmarking of optimization algorithms. Technical report, University of British Columbia, 2015. → pages 24

[CCGC15] C. Celauro, F. Corriere, M. Guerrieri, and B. L. Casto. Environmentally appraising different pavement and construction scenarios: A comparative analysis for a typical local road. *Transportation Research Part D: Transport and Environment*, 34(0):41 – 51, 2015. → pages 1

[CGF89] E. P. Chew, C. J. Goh, and T. F. Fwa. Simultaneous optimization of horizontal and vertical alignments for highways. *Transportation Research Part B: Methodological*, 23(5):315 – 329, 1989. → pages 1

[CL06] J. Cheng and Y. Lee. Model for three-dimensional highway alignment. *Journal of Transportation Engineering*, 132(12):913–920, 2006. → pages 1

[DM02] E. D. Dolan and J. J. Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91:201–213, 2002. → pages 29

[Eas88] S. M. Easa. Selection of roadway grades that minimize earthwork cost using linear programming. *Transportation Research Part A: General*, 22(2):121 – 136, 1988. → pages 4

[ESHS02] S. M. Easa, T. R. Strauss, Y. Hassan, and R. R. Souleyrette. Three-dimensional transportation analysis: Planning and design. *Journal of Transportation Engineering*, 128(3):250–258, 2002. → pages 1

[Fwa06] T. F. Fwa. *The Handbook of Highway Engineering*. CRC Press, 1st edition, 2006. → pages 2

[GCF88] C. J. Goh, E. P. Chew, and T. F. Fwa. Discrete and continuous models for computation of optimal vertical highway alignment. *Transportation Research Part B: Methodological*, 22(6):399 – 409, 1988. → pages 3

[GLA09] A. B. Goktepe, A. H. Lav, and S. Altun. Method for optimal vertical alignment of highways. *Proceedings of the ICE - Transport*, 162(4):177–188, 2009. → pages 3

[HEAEH98] Y. Hassan, S. Easa, and A. Abd El Halim. Highway alignment: Three-dimensional problem and three-dimensional solution. *Transportation Research Record: Journal of the Transportation Research Board*, 1612(1):17–25, 1998. → pages 1

[HHLR14] W. Hare, S. Hossain, Y. Lucet, and F. Rahman. Models and strategies for efficiently determining an optimal vertical alignment of roads. *Computers & Operations Research*, 44:161 – 173, 2014. → pages 2, 4, 5, 7, 8, 20, 21, 22, 27, 33

[HKL11] W. L. Hare, V. R. Koch, and Y. Lucet. Models and algorithms to improve earthwork operations in road design using mixed integer linear programming. *European Journal of Operational Research*, 215(2):470 – 480, 2011. → pages 1, 4, 7

[HL07] A. Herranz-Loncán. Infrastructure investment and spanish economic growth, 1850-1935. *Explorations in Economic History*, 44(3):452 – 468, 2007. → pages 1

[HLR15] W. Hare, Y. Lucet, and F. Rahman. A mixed-integer linear programming model to optimize the vertical alignment considering blocks and side-slopes in road construction. *European Journal of Operational Research*, 241(3):631 – 641, 2015. → pages 4, 20, 21, 33

[Jon98] J. C. Jong. *Optimizing highway alignments with genetic algorithms*. PhD thesis, University of Maryland College Park, United States - Maryland, 1998. → pages 4

[JS03] J. C. Jong and P. Schonfeld. An evolutionary model for simultaneously optimizing three-dimensional highway alignments. *Transportation Research Part B: Methodological*, 37(2):107 – 128, 2003. → pages 3

[JSJK06] M.K. Jha, P. Schonfeld, J.-C Jong, and E. Kim. *Intelligent road design*, volume 19. WIT Press, 2006. → pages 3

[KJS05] E. Kim, M. K. Jha, and B. Son. Improving the computational efficiency of highway alignment optimization models through a stepwise genetic algorithms approach. *Transportation Research Part B: Methodological*, 39(4):339 – 360, 2005. → pages 3

[KJS12] M. W. Kang, M. K. Jha, and P. Schonfeld. Applicability of highway alignment optimization models. *Transportation Research Part C: Emerging Technologies*, 21(1):257 – 286, 2012. → pages 3

[KL10] V. R. Koch and Y. Lucet. A note on: Spline technique for modeling roadway profile to minimize earthwork cost. *Journal of Industrial and Management Optimization*, 6(2):393 – 400, 2010. → pages 4

[LC01] Y. Lee and J. Cheng. Optimizing highway grades to minimize cost and maintain traffic speed. *Journal of Transportation Engineering*, 127(4):303 – 310, 2001. → pages 3

[LPZL13] W. Li, H. Pu, H. Zhao, and W. Liu. *Bidirectional Dynamic Programming Approach for Highway Alignment Optimization*, volume 779-780 of *Advanced Materials Research*, chapter Chapter 2: Traffic and Transportation Engineering, Civil and Mechanical Engineering, pages 700–704. 2013. → pages 4

[MA04] A. A. Moreb and M. S. Aljohani. Quadratic representation for roadway profile that minimizes earthwork cost. *Journal of Systems Science and Systems Engineering*, 13(2):245 – 252, 2004. → pages 1, 4

[MLH15] S. Mondal, Y. Lucet, and W. Hare. Optimizing horizontal alignment of roads in a specified corridor. *Computers & Operations Research*, 64:130 – 138, 2015. → pages 2, 5

[Mor96] A. A. Moreb. Linear programming model for finding optimal roadway grades that minimize earthwork cost. *European Journal of Operational Research*, 93(1):148 – 154, 1996. → pages 4

[Mor09] A. A. Moreb. Spline technique for modeling roadway profile to minimize earthwork cost. *Journal of Industrial and Management Optimization*, 5(2):275 – 283, 2009. → pages 4, 20

[NEW75] A. J. Nicholson, D. G. Elms, and A. Williman. Optimal highway route location. *Computer-Aided Design*, 7(4):255 – 261, 1975. → pages 4

[Tri87] D. Trietsch. Comprehensive design of highway networks. *Transportation Science*, 21(1):26–35, 1987. → pages 4