OPTIMISING AFRICA

HOW TO OPTIMISE INVESTMENTS IN THE NETWORK OF MEDICAL CENTRES ALONG THE AFRICAN HIGHWAYS

H. (Harwin) de Vries[1] - 311376

November 17th, 2011

ABSTRACT

This thesis describes the possibilities to model and solve the problem of investing optimally in a network of medical centres, called Roadside Wellness Centres (RWCs), along the African highways. By placing new RWCs at busy *truck-stops*, one intends to increase the number of truck drivers which have access to medical service. Next to that, the investments should create a network of RWCs, in which the truck drivers are provided with a *continuum of care*. This means that truck drivers, who suddenly need medical help, do not need to drive a long time along their routes before passing an RWC. In order to realise these goals, two MIP models are proposed in this thesis. The RWC Investment Model (RIM) models the problem of allocating locations to p new RWCs. The objective is to do this in such a way that the expected number of patients visiting the RWCs is maximised *and* that the *expected time* to the next RWC passed by an African truck driver is minimised. This model can be classified into the flow coverage models that also maximise the node demand covered. The RIM is extended to the RWC & Staff Investment Model (RSIM), which models the problem of optimally investing a budget increase by establishing new RWCs and hiring new employees. Both the RIM and the RSIM are extended so that de-investments can be optimised too. A case study shows that these models increase the benefits of investments a lot compared to the current investment strategy. The models can solve a large problem instance within an acceptable time. The sensitivity of the optimal solution to noise in the main input data and to changes in the user-defined parameters is analysed.

Keywords: Facility location; Optimisation; Integer program; Network; Flow refuelling; Flow coverage; Medical centres; Continuity of care; Truck driver; Corridor; Africa

ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics
Department of Econometrics
SUPERVISOR
Prof. Dr. Albert P.M. Wagelmans

ORTEC CONSULTING GROUP
Team North Star Alliance
SUPERVISORS
Drs. Frans van Helden
Drs. Timon van Dijk

[1] Master student at the Erasmus University Rotterdam, Erasmus School of Economics, Department of Econometrics, specialisation Quantitative Logistics and Operations Research and trainee at the ORTEC Consulting Group.

PREFACE

This preface marks the beginning of this thesis, but also the end of more than four years of study at the Erasmus University Rotterdam. Though I had a great time, I am also glad that my study Econometrics is finished. From now on, I can use my skills and knowledge to mean something for this world. Writing this thesis already showed me how great it is to make people happy and even to save lives with your knowledge.

Though writing a thesis is an individual assignment, many people contributed to the completion of this dissertation. I would like to use this preface to thank these people. First, I would like to thank Prof. Dr. Albert P.M. Wagelmans. Because of his support, insights, and valuable comments, he safe-guarded the academic value of this thesis. Furthermore, I thankfully acknowledge the support of my supervisors at the ORTEC Consulting Group, Frans van Helden and Timon van Dijk. Their theoretical background combined with their practical knowledge ensured that this thesis contains practical and well-founded solutions for North Star's problem. Last, I would like to thank all colleagues at ORTEC and at North Star Alliance. They provided a lot of support and a motivating working environment in the past months.

Special thank goes out to my family and friends who supported me during the last four years and the completion of this thesis. In particular, I would like to thank my wife Aline and my children Bertin and Henriëtte for the support, patience, and motivation provided in these years.

Finally, I hope that you like reading this thesis and that you share my enthusiasm for this topic.

Rotterdam, October 2011.

Harwin de Vries

NOTATION

a_{hk}	a coefficient equal to 1 if RWC location *k* is in combination *h*, 0 otherwise
AS	the set of all vectors of 4 binary variables, for which holds that only binary variable i is equal to 1 *and* binary variable i+1 take the value 1, for some $i \in \{1,2,3\}$.
b	current yearly budget (€/year)
bI	increase in the yearly budget (€/year)
bD	decrease in the yearly budget (€/year)
bR	maximum re-invested yearly budget (€/year)
bf_k	fixed yearly costs incurred when an RWC is established at location *k*.
bv_{ke}	variable yearly costs incurred when η_e employees occupy an RWC at location *k*.
$\bar{c}$	maximum possible continuity of care score of the network obtained after the investment
c_q	continuity of care score of path *q*
c_{hq}	continuity of care score of path *q*, when along this path combination of RWCs *h* is established
c_s	continuity of care score of a path, if it is assigned to scenario *s*
$\bar{d}$	maximum possible patient visits score of the network obtained after the investment
d_{ke}	estimated patient visits score at location *k* if η_e employees are employed there.
d_k	expected patient visits score for an RWC at location *k*.
D^q	destination of flow *q*. $D^q \subseteq L$
E	{0, 0.5, 1, 1.5,…, $\bar{\eta}$}: the set of possible numbers of FT employees (multiples of 0.5) occupying an RWC, indexed by *e*.
ec_k	element of E, corresponding to the *current* number of FT employees occupying location *k*
EI^k	$\{e \mid \eta_e > \eta_{ec_k}\}$ the set of numbers of FT employees that are larger than the current number of FT employees occupying location *k*
ED^k	$\{e \mid \eta_e < \eta_{ec_k}\}$ the set of numbers of FT employees that are smaller than the current number of FT employees occupying location *k*
η_e	parameter which is equal to the number of FT employees which element *e* corresponds to
ER_q	expected RWC time of path q
ER_{kl}	expected RWC time for truck drivers who are in segment (*k,l*)
ER_s	maximum value of ER_q for which path *q* can be assigned to scenario *s*
f_q	size of the flow along path *q* (nr. Of truck drivers starting this trip /day)
H^q	set of all 'relevant' potential *combinations of RWCs* at path *q*, indexed by *h*

i_{klq}	1 if locations *k* and *l* contain an RWC neighbour-pair (see definitions 4.1 and 5.1) at path q, 0 otherwise.
J^q	the set of all possible numbers of RWC neighbour-pairs defined along path *q*, indexed by *j*.
K	KC $\cup$ KP: the set of all current and potential RWC locations, indexed by *k*
KC	set of the locations of all RWCs which are *currently* in the network, indexed by *k*. KC $\subset$ K
KP	K-KC: the set of all *potential* RWC locations, indexed by *k*. KP $\subseteq$ K
K^{qk}	set of RWC locations along path *q* that are passed *after* passing location *k*
KA	K-KOD: the set of the locations of all RWCs that are *not at the origin or destination of a flow*, indexed by *k*. KA $\subset$ K
KOD	KO $\cup$ KD: the set of RWC locations situated at the origins of all paths, KO, united with the set of RWC locations situated at the destinations of all paths, KD, indexed by *k*
KEQ	set of locations at which RWC equivalents are situated. KEQ $\subseteq$ KC
L^q	$K^q \cup O^q \cup D^q$: the set of locations along path *q*
L^{qk}	set of locations along path *q* that are passed *after* location *k*. Note we pretend that $k \in KO^q$ is passed after O^q, and that $k \in KD^q$ is passed before D^q
M_q	an upper-bound on t_{klq}^{diff}, $\overline{t_q}$ and $t_q^{\max}$
n_q	number of (unordered) RWC neighbour-pairs along path *q*.
O^q	origin of flow *q*. $O^q \subseteq$ L
p	number of new RWCs to be located. $p \in \{0,...,\|kP\|\}$
pD	number of RWCs to be removed. $pD \in \{0,1,...,\|kC\| - \|KEQ\|\}$
pr_{klq}	probability that a truck driver is in segment (*k,l*) at a randomly chosen time in the time-line of path q
Q	set of non-zero flow paths in the network indexed by q
r, 1-r	relative importance of maximising $\sum_{k\in K} d_k x_k / \bar{d}$ and $\sum_{q\in Q} c_q f_q / \bar{c}$ respectively. $r \in [0,1]$
S	set of scenarios, indexed by *s*
t_{kl}	the driving time between location *k* and location *l* (see definition 4.1), if *k* ≠ l the time one has to spend at location *k* involuntarily, if *k* = *l*
tO_{qk}	driving time from location *k* to the origin of path q.
$\overline{t_q}$	average of the driving times between neighbour RWCs along path q
$t_q^{\max}$	maximum among the driving times between neighbour RWCs along path q
T_q	sum of the driving times between the RWC neighbours along path q
$\overline{\tau_s}$	maximum value of $\overline{t_q}$ for which path *q* can be assigned to scenario *s*

$\tau_s^{\max}$ maximum value of $t_q^{\max}$ for which path *q* can be assigned to scenario *s*

τ_s^{diff} maximum value of $|t_{kl} - \overline{t_q}|$ for all RWC neighbour-pairs $(k,l) \in N(\{x_k \mid k \in K^q\})$ for which path *q* can be assigned to scenario *s*

v_1 weigh factor for the patient visits score of the network. v_1=r/ $\overline{d}$

v_2 weigh factor for the continuity of care score of the network. v_2=(1-r)/ $\overline{c}$

w_1,w_2,w_3 weigh factors $\in$ [0,1]; $w_1+w_2+w_3=1$

x_k 1 if an RWC is placed at RWC location *k*, 0 otherwise

xe_{ke} 1 if the number of employees at location *k* is equal to η_e, 0 otherwise

y_{sq} 1 if path *q* is assigned to scenario *s*, 0 otherwise

y_{hq} 1 if along path *q* combination of RWCs *h* is established, 0 otherwise

Z objective value

CONTENTS

1. INTRODUCTION

Road transport is one of the main drivers of the Sub-Saharan African economies. By means of a network of Trans-African corridors, long distance movements of enormous amounts of goods flow to and from harbours, airports urban centres and other areas of high economic importance. Every day, thousands of long haul truck drivers and their assistants are on the move. Because of the huge distances that have to be travelled, they are on average 25 days per month separated from their homes and families (Ferguson & Morris, 2006).

Most of these truck drivers are men that are in the age of being sexually active, working under stressful conditions, and often carrying significant sums of cash to meet their travel needs (Caraël, 2005). These factors make them attractive customers to the sex industry that tends to be active in so-called 'hotspots' where trucks stop during night. This leads to enormous HIV prevalence rates among these truck drivers: exceeding 50% in some cases (Ramjee & Gouws, 2000). Next to that, they are very vulnerable to Sexually Transmitted Infections (STIs), TB, Hypertension, Malaria and other diseases. Due to their high mobility, they do not only contract lots of diseases, but also transfer them quickly. This makes the truck drivers one of the main dispersers of HIV in sub-Saharan Africa (Hudson, 1996).

These infections are not only a problem for those people that are infected themselves, but also for the Sub-Saharan countries as a whole. HIV infections among *employees* bring about increased absenteeism, increased staff-turnover, loss of skills, declining morale, and loss of tacit knowledge. These factors lead to increased costs and declining profits within the transport sector. Obviously, this brings about a competitive disadvantage for the Sub-Saharan economies (UNAIDS, 2000; Whiteside, 2006).

Since about ten years ago, decreasing the risk of truck drivers with respect to contracting and transferring infections has been getting more attention. Some studies came up with concrete suggestions about how to provide a sustainable solution for this problem. Truck drivers need to be provided with sexual health information, should be provided with condoms, and should be counselled, tested and treated for HIV and other infections (Ramjee & Gouws, 2000; Ramjee & Gouws, 2002; FHI, 1999). However, the traditional hospitals are often incapable to fulfil this role. They are generally situated at places that are hard to access by trucks and that have insufficient parking space. Moreover, the traditional hospitals are only accessible during daytime, whereas truck drivers generally only have time in the evening or at night (Gatignon & Van Wassenhove, 2008).

Several NGOs detected this problem and came up with the vision: 'If truck drivers won't come to the clinics, the clinics must go to them'. One of these NGOs is the North Star Alliance, which has been establishing 22 Roadside Wellness Centres (RWC) along the main corridors in East and Southern Africa since 2005. RWCs aim to provide truck drivers, sex workers and corridor communities with a collection of health services. This includes condom distribution, behaviour change communication (BCC), voluntary counselling and testing (VCT), and clinical services like treatment for STIs and other diseases.

North Star intends to build a whole network of health centres along the main transport corridors in Africa. They plan to expand this network up to 100-150 RWCs in 2015, because of multiple reasons. First, North Star hopes to enable truck drivers, which are travelling everywhere, to get medical help everywhere. Second, North Star intends to provide treatment for HIV and TB in the future. Patients being treated for these diseases should have very *quick access* to medical in case of complications or in case that they lose their pills. Third, expanding the network is important to make following up treatments or following up BCC lessons easier. In short, North Star aims to offer truck drivers a *continuum of care*.

RWCs are placed at so-called 'hotspots': truck stops and borders where large numbers of truck drivers stop and sex work and other informal trades flourish. However, till now these placements are done rather opportunistically. Instead of investing in RWCs in such a way that the *added value* to the target groups is maximised, the maximisation of the *number of patient visits* has been getting highest priority. This can be explained by the fact that North Star is a relatively young organisation, which has to prove that its concept works well in order to contract sponsors. However, this placement strategy has its drawbacks. The current network of RWCs is so fragmented, that only very few truck drivers have been visiting more than one RWC till now. Only 14 out of the 8929 registered patients have been visiting 2 or more different RWCs. Though not all visits and not all patients are registered, this statistic illustrates that you cannot speak about a *network* of RWCs yet. This shows that the possibility for offering truck drivers with continuity of care is a utopia as long as the current placement strategy in maintained.

This thesis proposes a new investment strategy in the form of a mathematical model. The main objective of this model is to maximise the added value of the entire network of RWCs for the truck drivers, the sex workers and the corridor communities. Two kinds of investments can be made in order to meet this objective: establishing new RWCs and hiring additional employees which occupy an RWC. The main issues that affect the 'fitness' of such investment or de-investment decision are: the costs, to what extent the number of patient visits changes, and to what extent the investment serves the objective of ensuring continuity of care to the truck drivers. We start with modelling the problem of optimising the locations of a batch of new RWCs as a Mixed Integer Programming (MIP) problem. Later, we expand the model by including the option to invest in additional staff. North Star also faces situations in which RWCs have to be closed and employees have to be fired because of a budget decrease. Therefore, both models are expanded for making de-investment decisions too. In short, the main objective of this thesis is to describe the possibilities to model and solve the problem of optimising the investments in the network of RWCs.

This thesis is organised as follows. Section 2 describes the problem in detail by means of a description of the relevant factors affecting the fitness of an investment in the network of RWCs. Section 3 reviews the literature dealing with problems that are similar to our problem. Section 4 frames our problem by means of assumptions about the problem structure. Based on these assumptions, section 5 describes mathematical models which optimise investments and de-investments in the network of RWCs. The performance of these models is tested in section 6. Last, in section 7 we draw some conclusions from the findings described in this paper, and make suggestions for future research.

2. PROBLEM DESCRIPTION

This section explains which issues affect the ‘fitness’ of an investment in new RWCs and/or new employees. These issues are split up in 4 categories: financial issues, patient visits, continuity of care, and location issues.

Some of these factors are related or even conflicting: optimising the way to invest money with respect to one issue has a negative effect in terms of another issue. In order to get insight in effects between them, the ways these issues are related are explained in subsection 2.2. Subsection 2.3 summarises the relation between all issues and the ‘fitness’ of an investment, and comes up with the formal problem statement.

2.1 Issues affecting the fitness of an investment

Financial issues

One of the most crucial steps that must to be taken before placing an RWC/ hiring employees is to make sure that sponsors are willing to fund such investment. The possibilities for *funding* are often dependent on the location where the money is invested. Namely, a great part of the money is provided by national/regional institutions, like ministries, road authorities, and transport operators.

Also the costs of investments in the network of RWCs may differ per country or per region. We can split up these costs in *fixed* yearly costs and *variable* yearly costs. Fixed yearly costs are the costs caused by the RWCs themselves, like costs of maintenance, rent, and amortisation (depends on the costs of establishing an RWC at a certain location). Costs are fixed costs if these are incurred when zero employees would occupy the RWCs. Variable yearly costs are the costs of *running* an RWC, like wages, costs of medicines, costs of electricity and water, and costs of replenishment. These costs largely depend on the number of employees at a certain location

In short, the fitness of an investment depends on the possibilities for funding at the location where that investment is made, and the variable and fixed yearly costs at that location.

Patient visits

One of North Star’s goals is to choose the location of an RWC such that the number of patient visits per day is maximised. A patient visit is defined as the event when one patient enters an RWC to get medical help (e.g. treatment or testing). Therefore it is important to choose a location where many potential patients come along or hang around.

Experience within North Star has uncovered a couple of variables that seem to be related to the number of patient visits per day: the *traffic volume* on the corridor next to the RWC, the *number of trucks stopping* at the truck stop or border the RWC is located at, the *average time trucks stop* at this place, and the *HIV&AIDS prevalence rates* among the truck drivers and sex workers visiting this place. In its turn, the amount of trucks stopping at the truck stop seems to be related to the amount of bars, sex shops and refuelling stations at this place.

However, some locations will ensure large numbers of patient visits, but are not attractive for North Star. This is simply because too many *local* people will visit the RWCs when placing them at these locations. Instead, North Star focuses mainly on providing health services to the *truck drivers*. This sounds rather discriminating, but as described in the introduction, there are multiple reasons why to take the truck drivers as target group. That is why North Star only wants to make investments at locations where truck drivers represent at least 50% of the potential patient visits.

Though it is also important to maximise the number of people which can be provided with condoms and with voluntary counselling services, we don't regard these objectives as determinants of the location of an RWC. Providing these services is more an organisational challenge than a matter of placing the RWC at the right location.

Summarising, the fitness of an investment depends on the resulting change in the number of patient visits and how many of these patients are truck drivers, sex workers, and local people.

Continuity of Care

As mentioned in the introduction, North Star wants to make investments in such a way that the network of RWCs provides a *continuum of care* to truck drivers. There are many reasons why this is important. First, North Star intends to enable truck drivers, which are travelling everywhere, to get medical help everywhere. Second, this ensures that there is always an RWC in the neighbourhood when a truck driver *suddenly* needs to get medical help. This is quite important, for example, when a truck driver gets Malaria, or when he loses his pills while being treated for Tuberculosis or HIV. Third, the barriers for undergoing follow-up treatments or follow-up BCC lessons are taken away. Truck drivers are highly mobile, so that it is difficult for them to undergo these at the same location. When ensuring a continuum of care, there is always an RWC in the neighbourhood when they *have to* undergo this follow-up treatment/ have to attend the lesson. The last reason is that placing multiple RWCs along a route results in *market expansion*. In section 2.2 we will come back to this.

Summarising, ensuring a continuum of care is primarily meant to ensure that truck drivers, who suddenly need to visit an RWC while travelling along their routes, do not need to drive a long time before passing an RWC. The extent to what the continuity of care ensured to *each* truck driver (travelling through the network) changes after an investment, determines the fitness of that investment.

Location issues

There are some properties of a location that determine the fitness of making investments at that location. First of all, it must be located in a safe region. Next, it must be supplied with water and electricity (which could eventually be provided by means of a generator). Third, having access to the internet is desirable, because this is required to run the COMETS program in the RWC. This program is an electronic health passport system that links centres together. This way, the patient records are accessible at every RWC, enabling truck drivers to continue treatment at every RWC. Last, possibilities to refer patients to hospitals that are close to the RWC make investing in that location more attractive.

2.2 Relations between issues

Continuity of care vs. patient visits

If North Star invests in new RWCs in order to improve the continuum of care to truck drivers, this often results in a negative side-effect. By placing multiple RWCs along the route of a flow of truck drivers, there will be *competition* among them. For example, it might be optimal (in terms of continuity of care) to place an RWC at a truck stop which is approximately one driving-day from a truck stop where another RWC is located. A number of truck drivers will probably visit both truck stops in two days, so that these RWCs have to 'share' part of the pool of potential patients. This results in cannibalisation. In contrast, when placing two RWCs very close to each others, they don't compete a lot. Trucks stopping at a certain truck stop will hardly ever go to a truck stop which is located a couple of miles away, because most of the truck drivers always attend the same truck stops. So, two RWCs which are situated close to each others are assumed to have a completely different pool of potential patients.

As mentioned in section 2.1, investments in new RWCs also bring about *market expansion*. That is: truck drivers tend to have more 'consumption' (i.e. patient visits) when they are provided with continuity of care. First, the barriers for attending an RWC diminish when truck drivers know that the quality of North Star's services is high (i.e. they are provided with a continuum of care). A second reason can be found in the so-called 'brand-effect'. In the field of consumer behaviour, this mechanism is used to explain how the brand of a certain product influences the consumer's willingness to buy it and the perceptions of its value and quality (Dodds, Monroe, & Grewal, 1991;

Rao & Monroe, 1989). Similarly, the willingness to visit an RWC might be larger when truck drivers know the brand 'North Star' from other RWCs they pass.

In short, investments in the continuity of care offered to truck drivers travelling along a certain route (i.e. how many RWCs are located along that route and where these are located) have an influence on the number of patient visits through the mechanisms of cannibalisation and market expansion.

Patient visits vs. financial Issues

Experience has shown that investing in additional nurses or outreach workers results in a significant increase in the number of patient visits, because of two reasons. First, the size and shape of the truck stop is often such that many truck drivers regard the RWC as being located too far away. An *outreach worker* can access these truck drivers and stress the importance of being counselled, tested and treated for HIV and other infections. Second, one nurse can handle about 25-30 patient visits per day. Investing in an additional *nurse* simply increases the number of patient visits, because the capacity of the RWC increases.

There is also an effect in the opposite direction: when making investments in order to increase the number of patient visits (by hiring nurses or outreach workers), this has a number of financial consequences. First, North Star provides its services for free, so that the costs of every patient visit (e.g. medicines, costs of testing, condoms) are covered by themselves. Also the opposite is true. When an investment results in an increase in the number of patient visits, sponsors might be more willingly to fund the corresponding RWC and North Star as a whole.

2.3 Summary

Figure 2.1 shows the 4 issues that determine the fitness of an investment, as well as the relations between these issues. From this figure it becomes clear that finding the optimal way to invest in RWCs and/or employees is not an easy task. First, the fact that there should be optimised along 4 dimensions makes the problem complex, because these have to be weighed in some way. Second, the fact that the relevant issues are influencing each other makes this problem complex: assumptions about the (mathematical) form of these influences have to be made. For example, how is the number of patient visits at a certain RWC influenced by the (distance to the) other RWCs? In section 4, some assumptions are made to simplify the optimisation.

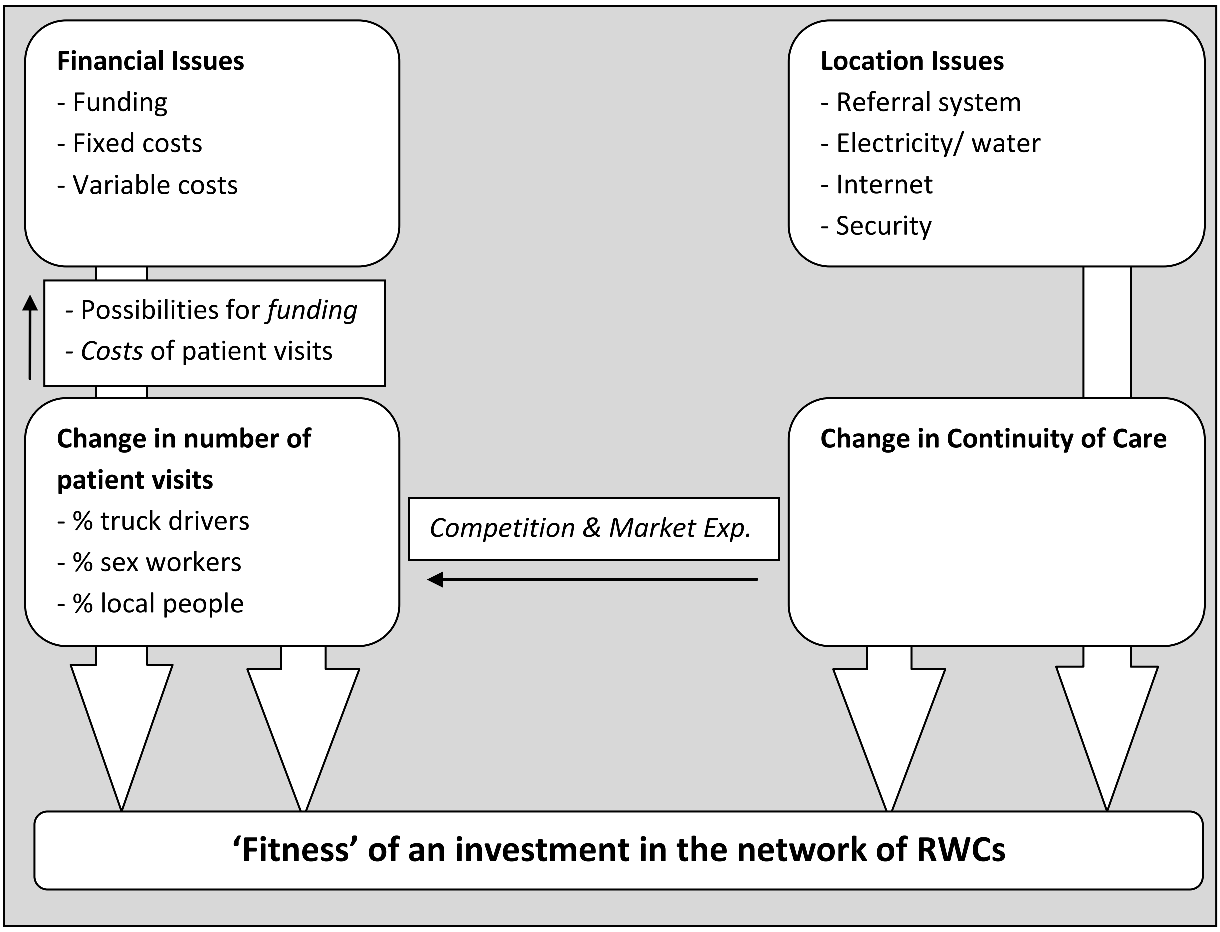


Figure 2.1 Issues affecting the fitness of an investment in the network of RWCs

The ultimate goal of this thesis is to model and solve the problem of taking the optimal decisions about where to invest or de-invest in RWCs and employees. The factors affecting the fitness of an investment in the network of RWCs as described in figure 2.1 are taken as a starting point.

Initially, we focus on investments in RWCs only, by proposing the RWC Investment Model (RIM). This model describes the problem of finding the *optimal locations* of a whole batch of RWCs that are going to be added to the network. After describing the RIM, we focus on extending this model to the RWC and Staff Investment model (RSIM), which models the problem of allocating an *investment budget* to investments in staff and to investments in RWCs. Last, these models are extended for cases in which de-investment decisions have to be taken.

3. LITERATURE REVIEW

From the previous sections it has become clear that this thesis deals with the problem of choosing the locations to invest in RWCs (and employees) among a set of nodes (hotspots), which are connected by means of arcs (roads). This makes our problem a *network location problem*. The network location literature is very broad and mainly devoted to the location of units like plants, facilities and services. For simplicity, we call these units 'facilities' from now on. These facilities generate or attract demand. In the network location literature the demand either comes from *static* demand units or from *moving* demand units, which make a trip from a specified origin (O) to a specified destination (D). The collection of static demand units at a certain location is called a 'demand point', whereas the collection of demand units moving along the same O-D path is called a 'flow'. When a demand point or flow is covered, intercepted, or captured by a certain facility, we just mean to say that this facility meets this demand.

We restrict ourselves to describing a couple of papers from three subclasses of network location problems that are more or less similar to our problem. This section starts with describing flow interception problems in subsection 3.1. Subsection 3.2 describes some location problems that take competition among different facilities into account. Subsection 3.3 reviews literature about the flow-refuelling problem and other multi-coverage problems. Last, subsection 3.4 summarises how the literature that is reviewed in this section can be used in order to model our problem.

3.1 Flow interception problems

Starting from about 1960 there has been a lot of research about network location problems. The list of references described by (Hale, 2006) contains over 3400 instances of these problems. Traditionally, in location theory the customers are seen as static points. Implicitly, this literature assumes that the customers always make a special purpose (dedicated) trip to obtain the service. That is why we call them 'static customers' from now on. To what degree such customer is covered depends on its distance to the facilities. Since the 1990s there has been interest in flow covering or flow interception problems, in which customers are seen as flows travelling along a path. As the name says, the facilities do not generate or attract flow, but intercept it. We call these customers 'flow-by customers' from now on. To what degree such customer is covered depends on the distance from its path to the facilities. Next to pure static customer coverage and pure flow-by customer coverage models, some papers combine them. The model proposed by (Zeng, 2007) maximises the convenience of the locations of the facilities to both types of customers and to the 'static and flow-by' customers. This last group refers to the consumers which choose a facility based on its greater convenience to either their *home or their path*.

Flow interception location models (FILM) have been used to solve lots of problems. We mention some of them that are related to transport or traffic. For a more comprehensive overview we refer to (Hodgson, 1998). FILMs can be split up in *flow oriented* and *gain oriented* problems. Whereas flow oriented problems focus on intercepting flow, gain oriented problems focus on the gain that can be obtained at a node by intercepting a flow there. Examples of these problems are maximising the amount of flow (gain) intercepted with a number of facilities, or minimizing the amount of facilities required to intercept a certain amount of flow (gain). Because these problems are NP-hard, the paper of (Sterle, 2010) proposes greedy, ascent, and tabu search heuristics to solve these problems.

The paper of (Zeng, 2007) describes another way to classify the objectives of flow interception models. Four types of customers are used to explain the differences among them: the *video* buyers, the *coffee* consumers, the *pizza* consumers, and the *hamburger* consumers. These four types of optimisation problems can be summarised in the Generalized Flow-Interception Location Allocation Model (GFILAM). We explain the four types of problems by means of an example.

Consider truck driver Mamadou who starts off his trip very early in the morning. Of course, he is in desperate need for *coffee*. A network of coffee-selling facilities is placed optimally with respect to the preferences of people like Mamadou, when it can intercept as many coffee buyers *as early as possible* on their paths. An example of a network location problem of this type is the placement of stations which inspect a network in order to detect dangerous vehicles (Hodgson et al., 1996). The objective is to detect them as soon as possible. This problem is modelled by means of an MIP and solved with a greedy heuristic. Another example is the detection for failures in a network. The paper of (Rosenkrantz, 2000) describes algorithms to locate failure inspection stations, based on the probability that a failure occurs at each path segment and the corresponding costs.

During the day, Mamadou looks around for a video store, so that he can watch a movie when he is back at home. Obviously, he does not care *where* to buy the video, but just wants to have it at the end of its trip. If you have to decide where to locate facilities which intercept consumers like video buyers, you do not take the location where the flows are intercepted into account. Instead, the objective is just to capture as much flow as possible.

Mamadou suddenly gets hungry around 12 o'clock and looks around for a hamburger restaurant. He does not have an absolute preference for the location of the restaurant, but just wants to have it close at the moment he gets hungry. The flow refuelling location problem is an example of a problem in which you try to meet demand at locations where many consumers are ‘hungry’. This problem deals with locating the refuelling or gasoline stations such that the vehicles can be refuelled at the moment they need to.

At the end of the day, Mamadou is tired and hungry, so that he decides to buy a pizza and to eat it when he is home. Mamadou prefers to eat the pizza when it is still warm, so that he wants to meet a pizza shop *as late as possible* on its trip. The location problems dealing with this kind of preferences aim to intercept as much flow *as late as possible* on their trips.

These examples illustrate that flow interception models differ in terms of the definition of the preferences (values) of the flows that are to be intercepted. These models also differ in the *definition of the interception* itself. The standard flow interception model assumes that customers like Mamadou do not deviate from their predefined paths in order to visit the facility. This means that they are only intercepted in case that a facility is located on their paths. In reality, Mamadou and his colleagues are willing to deviate in some cases. Therefore the paper of (Berman et al., 1995) describes three different variants of the way the interception is defined, and proposes (greedy) heuristics and algorithms to solve the resulting flow interception problems. The first assumes that Mamadou is willing to deviate at most Δ miles from his path in order to visit the facility. The second generalization assumes that Mamadou’s desire to visit the facility is so large, that he *always* deviates from its path to the nearest facility if there is no facility on its preferred path. Two examples of problems making this assumption are the ‘median problem with deviation distances’ and the ‘generalized maximum market size problem’ (Berman, 1997). In these problems, the facilities should be placed such that (a function of) the total deviation distance of all flows is minimized. The third generalization of the flow interception model assumes that the probability that Mamadou deviates from its preferred path in order to visit a facility depends on the deviation distance (larger distance means smaller probability). This idea of demand-points or flows being partially covered is also used in set-covering problems. For example, the paper of (Berman et al., 2002) defines the coverage of a customer as a step-wise function of its distance to the facility. The model presented in (Drezner et al., 2010) assumes that the coverage declines linearly from ‘fully covered’ to ‘not covered’ between two critical distances, and that these distances are random variables. Next to these three definitions of the interception of a flow, some papers assume that this also depends on the amount of competition among the facilities. We come back to this issue in section 3.2.

Other extensions of the flow interception problem are to assume limited capacities of the facilities, or to assume multi-type of flows. An example of the latter is the problem of locating both gas stations and convenience stores. These types of facilities have positive influence on each others.

When Mamadou stops for refuelling his vehicle, he always gets a cup of coffee when a convenience store is co-located with this refuelling station. Similarly, when Mamadou takes a break in order to eat a hamburger, he also refuels his vehicle if a gas station is co-located. So, co-locating facilities with other types of facilities could stimulate customers to have more consumption. The MIP model presented in (Dandan, 2009) takes this interaction into account when optimising the locations of 2 types of facilities. Because this problem is NP-hard, two greedy heuristics are proposed to solve it quickly.

The problems described above show some similarity to our problem. One of the objectives stated in section 2 is to optimise way to invest in the network of RWCs with respect to the number of patient visits. This number is to a large extent determined by the amount of traffic along the corridor the RWC is placed at. More traffic tends to result in larger numbers of patient visits. So, maximising the number of patient visits could be done by maximising the amount of truck drivers (flow) that is intercepted by the RWCs (facilities). The problem of locating an RWC can be modelled as a flow interception problem. Second, our problem deals with patients represented by flows (truck-drivers), but also with patients who can be regarded as static (sex workers and local people). A third feature of our problem is that the truck drivers can be seen as 'hamburger' consumers. Optimally, the RWCs are placed such that a truck driver passes one of them at the moment he 'gets hungry' for North Star's services. Fourth, truck drivers might deviate from their pre-planned paths in order to visit an RWC. Last, truck-drivers who stop along the road may have multiple objectives. For example, they might want to make use of the sex industry, might want to take some rest, might want to refuel their truck, or want to visit a health centre. Like described in (Dandan, 2009), this fact can be exploited by co-locating an RWC with other flow-attracting facilities. Namely, part of the flow attracted by other facilities may be 'captured' by the RWC, while these truck drivers would not have visited the RWC if it was not co-located.

3.2 Competitive facilities

When locating a set of retail facilities, two key effects need to be considered. The first is *cannibalisation*, which occurs when these facilities capture part of the demand from pre-existing facilities. The second is *market expansion*, which occurs when the total consumer demand increases as a result of opening new facilities. The paper of (Berman et al., 2002) describes a location optimisation model which captures these effects. Each demand point assigns a utility value to each facility, depending on the *distance* to such facility and based on the facility's *attractiveness*. The market expansion effect is taken into account by modelling the total expenditures of each market (i.e. a demand point) as a function of the utility values of all facilities. The cannibalisation effect is included by modelling the share a market that is captured by a certain facility as a function of its own utility and the utilities of the other facilities. Based on these two functions, the paper describes an IP model which maximises the total revenues of locating a number of facilities, and proposes algorithms to solve it.

The paper of (Berman et al., 2002) regards customers as static. As we said before, this implies the assumption that customers, who wish to visit a facility, always do this by making a special purpose dedicated trip. This assumption makes the model that is presented in their paper unsuitable for facilities that (partly) rely on flow-by customers (like gas stations and convenience stores). For these cases, (Berman et al., 1998) defines the utility value such flow-by customer assigns to a facility. Similar to the paper described above, this utility is defined as a function of the minimum deviation distance of the path to such facility and of the attractiveness of that facility. In order to also include the static customers in the model, these are also seen as flow-by customers who travel from and to the same point, and deviate to a facility while they make this trip. Based on the utility values an (imaginary) flow assigns to all facilities, the paper describes the problem of finding the optimal locations for a number of facilities. The objective is to maximise the market share, which is defined as the total utility the customers assign to your *own* network of facilities divided by the total utility the customers assign to *all* facilities (also those of the competitors). This problem is modelled as an IP model. A branch-and-bound procedure is proposed to solve this problem.

The model described in (Wu et al., 2003) also deals with maximising the share of flow-by customers that is intercepted by a number of facilities. However, the authors use the so-called gravity model (Huff, 1964; Huff, 1966) to define the utility a flow-by customer assigns to a facility. This definition is slightly different from the one presented in (Berman et al., 1998). Again, the problem is modelled as an IP model, and solved with a greedy heuristic.

The papers described above are similar to our problem in two aspects. First, the utility a truck driver assigns to an RWC also depends on the attractiveness of the location and on the distance he has to deviate to reach that facility. Second, when placing an RWC, we also face the effects of cannibalisation and market expansion. Cannibalisation occurs when two RWCs share part of the pool of potential patients. For example, this effect may occur when two RWCs at two different truck stops are attended consecutively by the same truck drivers. Market expansion also occurs when placing new RWCs. More truck drivers can be intercepted, and the 'consumption' of the truck drivers (i.e. the number of patient visits) increases because of the aforementioned brand effect.

3.3 Multi-coverage problems

The models described in the previous two subsections implicitly assume that a certain demand is covered if there is at least *one* ‘conveniently located facility’ (e.g. a facility along the path of a flow-by customer, or within a certain distance from a demand point). This assumption makes these models unsuitable when it is beneficial or even required to have *multiple* facilities on a flow.

One instance of a location allocation problem that takes the benefits of multi-coverage of flows into account is the location allocation of billboards. From the field of marketing research it becomes clear that seeing an advertisement multiple times may increase its effects on the consumer. The papers of (Averbakh et al., 1996; Hodgson et al., 1997) describe two IP models which do not only determine whether to place a billboard along a path, but also the number of billboards to place along it.

Another example is the *flow refuelling problem*. The paper of (Kuby et al., 2005) describes an algorithm to generate all combinations of nodes (refuelling stations) along a flow that can refuel this flow. A combination of facilities can *refuel a flow* when vehicles can repeat a trip from O to D and back to O multiple times. This is only the case if a vehicle does not face the situation in which the *driving distance* (note: different from absolute distance) to the next refuelling station it will pass is such that it runs out of fuel before the vehicle reaches it. Based on these combinations, the paper proposes an MIP model to determine the locations of a number of refuelling stations. We call this the Flow Refuelling Location Model (FRLM) from now on. The FRLM’s objective is to maximise the total amount of flow that can be refuelled. The authors apply this model in order to design a network of hydrogen stations in Florida (Kuby et al., 2009). Because of the computation times, solving the FRLM becomes very impractical when considering large networks. Three heuristic algorithms are proposed to solve this problem efficiently, including greedy-adding, greedy-adding with substitution, and genetic algorithms (Lim et al., 2010).

One feature of the FRLM is that it only selects locations among a pre-defined set of candidates. However, in reality it is sometimes more important to *generate* promising candidate locations. The authors tackle this issue by proposing three methods to add new *promising candidate locations* for refuelling stations (Kuby et al., 2007).

The FRLM is extended by also taking into account the customers’ willingness to deviate from their pre-planned path in order to get their vehicle refuelled. The paper of (Kim, 2010) proposes the Deviation-Flow Refuelling Location Model (DFRLM), which locates facilities to maximise the total amount of flow refuelled on deviation paths. The willingness to deviate from the pre-planned path is assumed to be decreasing when the deviation distance increases. The paper describes two heuristics to solve the DFRLM, including greedy-adding and greedy-adding with substitution algorithms.

Whereas the (D)FRLM focuses on maximising the amount of flow that can be refuelled with a certain amount of facilities, the paper of (Yang et al., 2010) approaches the flow refuelling problem in a different way. The authors describe an MIP model which optimises the *amount* of facilities and their locations, under the restriction that all vehicle flows (inter-city traffic) can be refuelled. This model balances the conflicting objectives of minimizing the costs of running the facilities and of maximising the node demand (intra-city traffic) that is covered by the facilities.

The FLRM requires O–D flow data, which are generally not easy to obtain. Instead, (Bapna et al., 2002) approach the problem of refuelling traffic flows as a Maximum Covering/Shortest Spanning Subgraph Problem (MC3SP). The problem is to decide which arcs (road between two cities) to provide with a combination of refuelling stations, enabling traffic along that arc. This is done by balancing their costs (the fixed costs associated with adding and/or upgrading enough stations on that arc, plus the variable cost of traffic along that arc) and their benefits (populations on or near the arc). The main restriction is that all significant cities are connected to the network, so that traffic between all these cities is possible. The paper models this problem as an IP model and proposes a spanning tree heuristic to solve it.

As was said before, the problem of refuelling vehicles on a flow looks a lot like the problem of ensuring continuity of care for a certain flow of truck drivers. First, the (D)FRLM aims to avoid cases in which refuelling stations are so far away from each others that vehicles get out of fuel before they pass the next station. Similarly, we want to avoid cases in which truck drivers do not pass an RWC for a long time. This means that a flow is not simply covered when *one* RWC is placed at this flow. Namely, sometimes multiple RWCs should be placed to ensure continuity of care. Second, also in terms of the objective of the location problem, our problem looks like the FRLM. Both problems want to place a fixed number of facilities optimally. Instead, the location models proposed in (Bapna et al., 2002) and (Yang et al., 2010) leave the amount of facilities open and determine this amount by balancing costs and benefits.

There are also differences between our problem and the (D)FRLM. First, the (D)FRLM regards a flow as being *covered or not*. In our problem, truck drivers are covered (i.e. provided with a continuum of care) *to some extent*, depending on the driving times between the RWCs passed along a route (if any). Second, the (D)FRLM maintains only has the objective to maximise the amount of flow covered. Instead, our problem deals with optimising investments with respect to multiple objectives. It might be beneficial to deviate from the 'optimal' locations in terms of continuity of care, in order to 'catch' more patient visits. Third, whereas the (D)FRLM places refuelling stations based on the *distance* between them, we place RWCs based on the *driving-time* between them. It is hard to transform driving-time to distance. Due to truck failures (which occur quite often) and delays at border posts, it is very hard to say how long it takes to cover a certain distance.

A difference between our problem and the problem modelled in the DFRLM is that truck drivers generally don't deviate from their *paths* in order to get medical help. Therefore, it is still open for discussion whether the DFRLM is applicable to our problem. We will come back to this in section 4.

3.4 Conclusions

In section 2 we described the problem of making investments in a network of RWCs in such a way that the total fitness value of the resulting network is maximised. In this subsection we summarise how the literature described above can be used for modelling this problem. Consecutively, we describe this for the problem of maximising the number of patient visits, for optimising the network with respect to offering continuity of care, and for the relation between these problems.

Patient visits

The problem of maximising the number of patient visits can be seen as a flow interception problem. The number of truck drivers which are 'intercepted' by the RWC depends on the number of trucks stopping at the truck stop the RWC is located at. This amount, in its turn, depends on the amount of traffic along the corridor next to the truck stop and on the attractiveness of the truck stop itself. Also the distance *other flows* have to deviate to reach this location may have an influence on the amount of trucks stopping at a certain location. In short, the decision about the truck stop that is attended by a truck driver can be modelled by means of the distance he has to deviate in order to reach it and the truck stop's attractiveness.

Another way to model the decisions of truck drivers about which truck stops to attend, is to regard them as multi-objective customers. Possible objectives are to visit a truck stop which is close at the moment they intend to take a break, to visit a truck stop with bars, a truck stop with sex workers, a truck stop with an RWC, or a truck stop with a refuelling station.

The number of patient visits is also determined by the size of the local community and the amount of sex workers hanging around. As described in section 3.2, these static visitors can easily be added to the flow interception model by seeing them as flow-by visitors, travelling from and to the same point, and deviating to an RWC.

Continuity of care

As described in section 2.1, a network has a large fitness value if it ensures truck drivers a continuum of care. Suppose that we can define all combinations of locations that ensure this on a flow. In that case, we can model the problem of building a network of RWCs can be modelled as a (deviation-) Flow Refuelling Location Model. However, reality is slightly different from the assumption underlying this approach. Whether a flow of trucks is ensured continuity of care is not a matter of *true or false*. Instead, this is ensured this to some *extent*. The FRLM needs to be adapted to include this idea of flows being *partially covered*.

Continuity of care vs. patient visits

The number of patient visits at a certain RWC is to some extent affected by the locations of all other RWCs because of competition. As described in section 3.2, the competition among facilities can be modelled by assuming that every demand unit assigns a utility value to a facility, based on the deviation distance and the attractiveness of that facility. The market share of each facility is defined as a function of the utility of that facility *and* the utilities of all facilities. Similarly, competition among RWCs can be modelled by assuming that each truck driver assigns a utility to each RWC. This utility is based on the attractiveness of the location the RWC is located at, and on the distance a truck driver has to deviate from his preferred truck stop in order to stop at the truck stop where the RWC is located. The 'market share' of each RWC can be calculated afterwards.

The network design does not only have an effect on the amount of competition among the RWCs, it also determines the number of patient visits for all RWCs. Therefore, the idea about how to take into account the market expansion effect of placing an RWC may be very useful for modelling our problem. As described in section 3.2, this can be done by modelling the total expenditures of each customer as a function of the utility values of all facilities in the network. Similarly, the number of patient visits 'performed' by a truck driver can be modelled as a function of the total utility he assigns to all RWCs in the network.

4. MODEL

In this section we make assumptions about the problem structure, so that we can model the problem of optimising the investments in the network of RWCs. Some of these assumptions are relaxed later.

Subsection 4.1 contains assumptions about the issues and relations between these issues that determine the fitness of a (de-)investment in the network of RWCs. The data required to describe and score such network of RWCs are described in subsection 4.2. Subsection 4.3 works out the first issue determining the fitness of a (de-)investment: the *patient visits score* of a network. The second issue determining this fitness, the *continuity of care score* of a network, is described in subsection 4.4. Last, subsection 4.5 summarises section 4.

For sake of simplicity, we regard de-investments as (a special kind of) investments: de-investing in a certain budget or in a number of RWCs is the same as investing a negative valued budget or in a negative number of RWCs. So, we will mainly talk about investments from now on.

4.1 Assumptions about the problem structure

A1.1: The fitness score of an investment in the network of RWCs *only* depends on (this assumption is changed later):

1. The resulting change in the total number of truck drivers, sex workers and locals visiting the RWCs in the network per day
2. The resulting change in the extent to what a continuum of care is ensured to the truck drivers who are travelling through the network

A2: There are no market expansion and competition effects in the network

These assumptions simplify the problem summarised in figure 2.1. Initially, we do not take the factors ‘financial issues’ and ‘location issues’ and the relationship between the factors ‘continuity of care’ and ‘patient visits’ into account.

By assuming A1 and A2, we implicitly make the following assumptions. First, we do not care about the fact that the costs of running and establishing an RWC and the possibilities for funding an investment may differ per location. Second, we do not care about the fact that the possibilities for referring patients to hospitals may differ per location. Last, we assume that placing or removing an RWC does not affect the numbers of patient visits at other RWCs (assumption A2).

The other location issues (whether the region of a potential RWC location is safe and whether electricity, water, and internet are available) can be taken into account by deleting the potential locations that do not meet the fixed constraints about these issues from the initial set of potential locations where could be invested in a new RWC.

4.2 Assumptions about the available data

We assume that we have access to the following data:

A3: The set of current and potential RWC locations in the network

A4.1: (Predicted) number of patient visits per day for each (potential) RWC location, and what number of these patient visits is performed by truck drivers, sex workers, and locals, respectively (this assumption is changed later).

A5: The expected driving times between all relevant locations. These driving times include the durations of involuntary stops

A6: 1. The set of non-zero flow paths through the network

2. The size of each flow (average number of trucks which start travelling along this route per day)

3. The set of RWC locations (current and potential locations) that are passed when travelling along a path

With respect to the flows of truck drivers, we assume the following:

A7: 1. All flows are O–D flows: not circular paths

2. The entire flow for a given O–D pair follows the same path through the network (no deviation)

3. The O–D flow matrix is symmetrical.

Assumptions A3 – A6 provide the information needed to score a network of RWCs. Obviously, *scoring a network* of RWCs enables *scoring investments* in the network of RWCs, by looking at the difference between the score of the network before the investment and the score of the network after the investment.

A4.1 provides the information needed to score a network with respect to the number of patient visits. A5 and A6 provide the information needed to score to what extent a continuum of care is ensured to the truck drivers (see assumption A11.1). This score is path-based: the route of a truck driver has to be known in order to calculate the driving times between RWCs passed along its route, i.e. to define to what extent a continuum of care is ensured to this truck driver. From now on, we call this the *continuity of care score of a path*. The ‘relevant locations’ mentioned in A5 are: potential

RWC locations, locations of RWCs that are already in the network, locations of RWC equivalents (see assumption A10.1), locations of origins and destinations of flows.

Because of A4.1, we can easily identify the potential RWC locations that do not meet the constraint that truck drivers represent at least 50 percent of the patient visits at an RWC. These potential RWC locations can be erased from the initial set of locations.

Assuming A7 brings about the advantage that the number of paths in the network is decreased a lot. This automatically decreases the number of variables in the models presented in section 5. If part 2 of A7 is not assumed, the number of potential paths through the network is infinite. Part 2 decreases this number to n^2, where n is the number of origin or destination nodes. By assuming part 1, this number is decreased further to n(n-1), because the path from location i to location i is not an option any more. By stating that the O–D matrix is also symmetrical, the ordered pairs i,j and j,i are identical so that these can be treated as an unordered pair i,j. This decreases the maximum possible number of O–D paths to n(n-1)/2.

4.3 Definition of the patient visits score

A8: Let dt_k, ds_k, and dl_k be the (predicted) number of patient visits at (potential) RWC location *k* per day that are performed by truck drivers, sex workers and locals, respectively. Let wdt, wds, and wdl be weigh parameters. Then the patient visits score of *location* k*,* d_k, is defined as: $wdt * dt_k + wds * ds_k + wdl * dl_k$

A9: The patient visits score of a *network* of RWCs is defined as the sum of the patient visits scores of all RWCs in that network.

These definitions enable that the importance of optimising the investments with respect to the truck drivers, sex workers, and locals is differentiated. North Star uses the values 3, 2, and 1 for wdt, wds, and wdl, respectively.

4.4 Definition of the continuity of care score

A10.1: RWC equivalents are situated at the origin and at the destination of each flow (this assumption is changed later)

A11.1: The continuity of care score of a path depends on the following factors (this assumption is changed later):

1. The average of the driving times between RWC neighbours (see definition 4.1) along this path
2. The variance among these driving times
3. The maximum among these driving times

A12: Let c_q be the continuity of care score of *path* $q \in Q$, and f_q the flow size of path q, then the continuity of care score of a *network* of RWCs is $\sum_{q \in Q} f_q c_q$

The origins and destinations of flows are often places where medical help is available. So, when determining the continuity of care score, these places can be regarded as RWC equivalents: places which can be seen *as if an RWC is located there*. This explains assumption A10.1. In the model presented in section 5, we represent these RWC equivalents by means of *dummy-RWCs*. For sake of simplicity, we generally regard these as normal RWCs from now on.

Assumption A12 defines the continuity of care score of a network as a weighed sum of the continuity of care scores of all paths in the network. This way, we prefer ensuring a continuum of care to *large* flows of truck drivers.

Assumption A11.1 requires some additional explanation. Why do we choose these three variables to determine to what extent a continuity of care is ensured to truck drivers along a path? Before we answer this question, we give a definition of the driving time between two neighbour RWCs along a path.

Definition of the driving time between neighbour RWCs along a path

Such driving time corresponds to the *maximum time to medical help* when travelling between two RWC neighbours (see definition 4.1) along a path. The durations of involuntary stops while travelling between two RWCs are also included in such driving time. Namely, these times definitely increase the maximum time to medical help. Examples of places where involuntary stops have to be made are: border posts, origins (loading time), and destinations (unloading time). In contrast, we do not regard stopping in order to take a break as an involuntary stop. First, this makes it very hard to estimate the driving time between two RWCs. Second, the maximum time to medical help refers to emergency cases in which these voluntary stops can be skipped. Summarising, we come to the following definitions:

Definition 4.1 Under assumption A10.1, two RWCs are *neighbours along a path* if and only if a truck driver travelling along that path will consecutively pass these two (dummy-) RWCs, without passing a third (dummy-) RWC. (A slightly different version of this definition is proposed in section 5 for the case that A10.1 does not hold).

Definition 4.2 The *driving time between (dummy-) RWC neighbours* along a path is the time needed to drive from the one to the other in case that the truck would not stop during this trip, plus the durations of involuntary stops which are made along the trip. The duration of voluntary stops, like stops for taking a break, is not included.

Definition 4.2 already highlights the importance of locating an RWC at places where involuntary stops are made. For example, consider case 1 as described in figure 4.1. Here, we visualise the trip from an origin to a destination by means of a time-line with events. The events are the moments at which an RWC is passed. Because RWC *6* is located at a place where a long-lasting involuntary stop is made, only one RWC is needed in the second part of the time-line of this trip to get the driving times to the next RWC very small. In contrast, the RWCs passed during the first part of the time-line of this path are not placed at long-lasting involuntary stops. Many RWCs are needed to get the driving times between RWCs as small as in the second part of this trip.

Figure 4.1: Time-line case 1. Black (black/white) boxes: moments at which RWCs (dummy-RWCs) are passed. Black rectangle: duration of an involuntary stop at a place where an RWC is located. Lines: driving times between (dummy-) RWC neighbours.

Calculation of continuity of care score 1

In order to calculate the continuity of care score of a path, we score three factors affecting the continuity of care ensured to truck drivers travelling along that path (assumption A11.1). Specifically, we assume that factor 1 gets score 0 if the average of the driving times between RWC neighbours along that path is larger than some threshold value $\hat{t}_2$, and gets the maximum score (>0) when this average is smaller than some threshold value $\hat{t}_1$. In order to score factor 2, we assign a score to the deviation of each individual driving time from this average. We assume that the score 0 is assigned to such driving time when the deviation is larger than some threshold value $\hat{t}_3$, and the maximum score when this deviation is equal to 0. Factor 3 gets score 0 if the maximum of all driving times between RWC neighbours along the path is larger than $\hat{t}_2$, and gets the maximum score when this maximum is smaller than $\hat{t}_1$. In section 5 we describe the functions determining the continuity of care score in detail.

Ideas behind continuity of care score 1

In section 3.1 we explained that ensuring a continuum of care is primarily meant to ensure that truck drivers, who suddenly need to visit an RWC while travelling along their routes, do not need to drive a long time before passing an RWC. Because of the following reasons, the three factors listed in A11.1 are used to score to what extent this is the case along a path.

First, a large average of the driving times between the RWC neighbours along a path indicates that truck drivers, who need medical help, often have to drive a long time before they pass an RWC. In contrast, a very small average indicates that there is often an RWC nearby when they need medical help. So, the average of the driving times between RWC neighbours along a path says something about the continuity of care along a path. This explains part 1 in assumption A11.1.

Second, it would not be correct to use only the average of the driving times between RWC neighbours to determine the continuity of care score of a path. To illustrate this, consider case 2 as described in picture 4.2. The average of the driving-times between RWC neighbours is very small in this case. But the RWCs are situated in such a way that the driving times between some of the RWC neighbours along this path are very large (e.g. the driving time between dummy-RWC *1* and RWC *2*). When truck drivers who are travelling along this path need medical help, they often need a very long time before passing an RWC. So, the continuity of care score of a path should not only be determined by the average of the driving times, but also by *the variance among these driving times*. This explains part 2 in assumption A11.

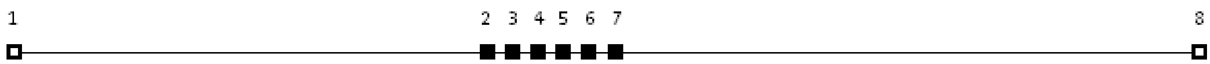


Figure 4.2: Time-line case 2. Black (black/white) boxes: moments at which RWCs (dummy-RWCs) are passed

Last, it is not entirely correct to use only the factors stated in parts 1 and 2 of assumption A11.1 to determine the continuity of care score of a path. Consider case 3, as described in figure 4.3. The average of the driving times between (dummy-) RWC neighbours and the variance among these driving times are such that these result in a reasonable continuity of care score. However, the driving time between RWC 7 and dummy-RWC *8* is so large that it becomes quite risky to undergo treatment for HIV and TB while travelling along this path. This stresses the importance of avoiding *large* gaps in the continuum of care of a path. Therefore, we also take the *maximum* among the driving times into account when determining the continuity of care score of a path. This explains part 3 in assumption A9.

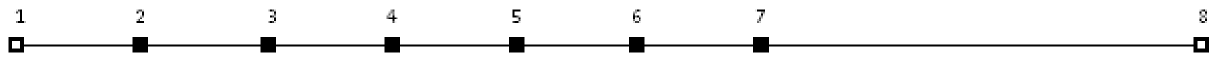


Figure 4.3: Time-line case 3. Black (black/white) boxes: moments at which RWCs (dummy-RWCs) are passed

Though the factor stated in part 2 of assumption A11.1 also provides a motive to avoid these gaps, there are two reasons why also part 3 should be used to define the continuum of care score of a path. First, part 2 does not increase the penalties for large driving times if the deviation from the average of the driving times becomes larger than $\hat{t}_3$. Second, part 2 does not look at the absolute size of a gap, but only at the size of the gap *compared to* the average of all driving times between the RWCs. Therefore, the penalties for large gaps are too small when they are determined by part 2 only.

Continuity of care score 2

As we show in section 5, using continuity of care score 1 (i.e. assuming A11.1) in the RIM and the RSIM results in very large and complex models. Some preliminary tests have been shown that it is very hard to solve large problem instances when using this definition. A good alternative is to replace assumption A11.1 by the following:

A11.2: The continuity of care score of a path depends on the expected driving time to the next RWC when travelling along the path.

This variable is defined as follows:

Definition 4.3 Consider the time-line of the trip of a truck driver travelling along path q. Suppose that *tO* $\in [0, T]$ is the time at which a truck driver is currently driving or making an involuntary stop while travelling along his path, i.e. the *driving time from the origin* to his current location. T is the driving time from the origin to the destination of the path he is travelling along. Let now *er(tO)* be the expected driving time to the next RWC that is passed from this moment. Then the expected driving time to the next RWC when travelling along the path is equal to $\int_0^T \frac{er(tO)}{T} dtO$. For sake of simplicity, we call this variable the *RWC time of a path.*

Specifically, we assume that the continuity of care score 2 of a path is equal to 0 if the expected RWC time is larger than some threshold value $\hat{t}_5$, and gets the maximum score (>0) when this variable is smaller than some threshold value $\hat{t}_4$.

This variable scores the continuity of care of a path in a similar way as the variables described in assumption A11.1. The RWC time of a path is large (small) when the average of the driving times between RWC neighbours along a path is small (large), when the variance among these driving times is small (large) and when there are no gaps (lots of gaps) in the continuum of care.

Assuming A11.2 instead of A11.1 brings about one disadvantage. The *RWC time* of a path will always decrease when placing an additional RWC along this path. This may stimulate placements of RWCs at roads that are already quite densely occupied with RWCs. Using A11.1 to determine the continuity of care score would avoid this effect, because factor 2 'penalises' placing RWCs close to each other (assumed that there exist parts of the path that are less densely occupied with RWCs).

4.4 Summary

Assumptions A1 – A12 frame the problem of optimising investments in the network of RWCs. This problem is summarised in figure 4.3. The objective is to choose the investment among all possible investments that has the maximum fitness value. This means: invest in the network of RWCs in such a way that the fitness value of the resulting network is maximised. This fitness value is a function of the patient visits score of the network and the continuity of care score of the network. This continuity of care score is either determined based on the driving times between neighbours RWCs along the paths in the network or based on the so-called expected RWC time of a path. These scores are called the continuity of care score 1 and 2, respectively.

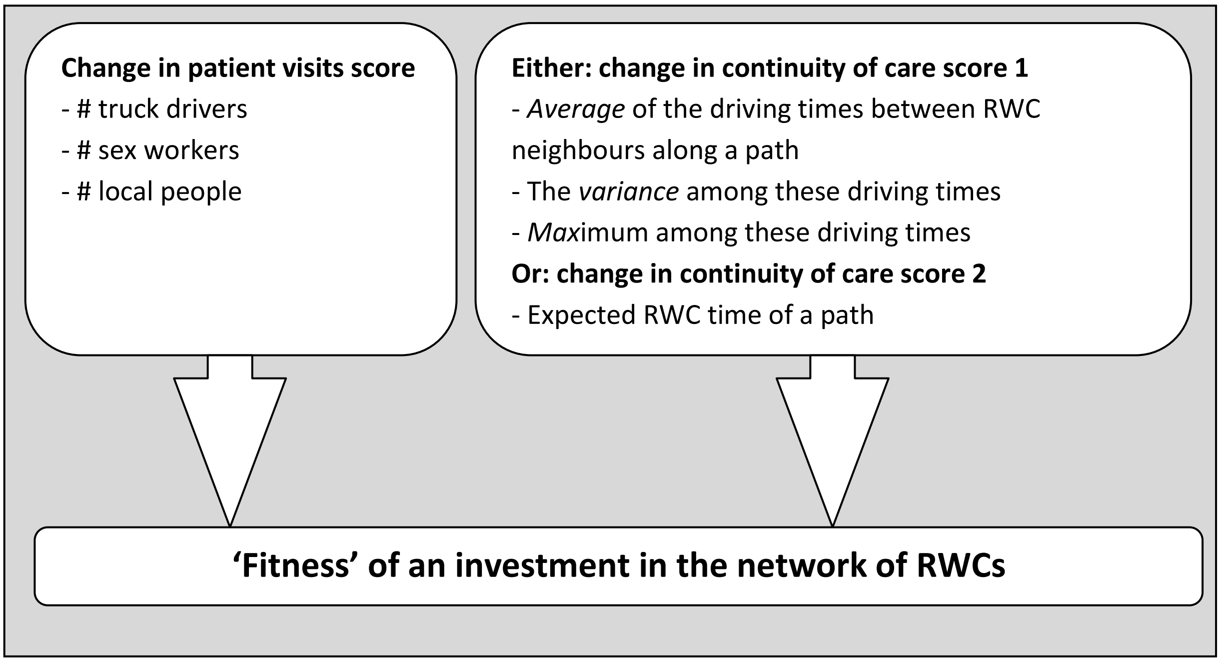


Figure 4.3: Issues determining the 'fitness' of an investment in the network of RWCs, based on assumptions A1.1, A2-A3, A4.1, A5-A10.1, A11.1 or A11.2, and A12

5. SOLUTION METHODS

In this section we model the problem of optimising the investments in a network of RWCs. We start with the problem of optimising the locations of *p* RWCs that are to be added to the network. Subsection 5.1 describes this problem by means of an IP model, which we call the RWC Investment Model (RIM) from now on. Here, we do not define the continuity of care score of a path yet. In subsection 5.2 we propose three approaches to include the calculation of the continuity of care score 1 into the RIM. As explained in section 4, this score is based on the driving times between RWC neighbours along a path. In subsection 5.3 we adapt these three approaches to include the continuity of care score 2. This score is based on the expected RWC time of a path. Subsection 5.4 describes three relevant extensions to the RIM. Last, section 5 is summarised in subsection 5.5.

5.1 RWC Investment Model (RIM)

This subsection models the problem of allocating locations, selected from a set of *potential* locations *KP*, to *p* RWCs that are going to be added to the network. The *current* RWCs and RWC equivalents (which we also regard as if they are current RWCs) are located at the set of locations *KC*. Together, KC and KP make up the set of *RWC locations,* denoted by *K*. The set *KOD* contains the set of RWC locations at the origins and the set of RWC locations at the destinations of all paths, denoted by *KO* and *KD* respectively. All of these sets are indexed by *k*. The decision variable x_k indicates whether an RWC is placed at location *k* (x_k=1), or not (x_k=0).

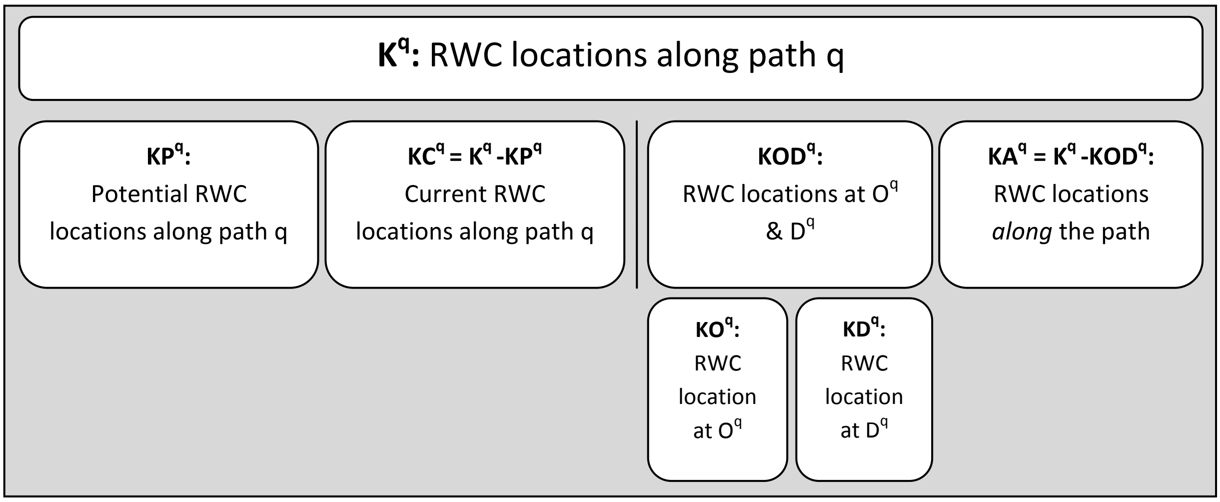


Figure 5.1: Overview of sub-sets of the set of RWC locations *K*

The objective function consists of two parts, corresponding to the objectives of maximising the patient visits score and of maximising the continuity of care score of the network of RWCs, respectively. The relative importance of both objectives is incorporated by means of the weigh factors v_1 and v_2. As we assumed in A4.1, we have estimates of the potential patient visits score at potential RWC location *k* and the current patient visits scores of the current RWC locations. These estimates are represented by d_k.

The set of paths *q* is denoted by *Q*. Every path has a unique combination of an origin and a destination. The *variable* c_q represents the continuity of care score of path *q*.

The resulting IP model is the following:

$$\max Z = v_1 \sum_{k \in K} d_k x_k + v_2 \sum_{q \in Q} f_q c_q \tag{5.1}$$

$$c_q = g(\{x_k \mid k \in K^q\}) \quad \forall q \in Q \tag{5.2}$$

$$\sum_{k \in KP} x_k = p \tag{5.3}$$

$$x_k = 1 \quad \forall k \in KC \tag{5.4}$$

$$x_k \in \{0,1\} \quad \forall k \in K \tag{5.5}$$

Where[2]:

KC: set of the locations of all RWCs which are *currently* in the network, indexed by *k*. KC $\subset$ K

KP: K-KC: the set of all *potential* RWC locations, indexed by *k*. KP $\subseteq$ K

K: KC $\cup$ KP: the set of all current and potential RWC locations, indexed by *k*

v_1: weigh factor for the patient visits score of the network. $v_1 = \frac{r}{\bar{d}}$

v_2: weigh factor for the continuity of care score of the network. $v_2 = \frac{1-r}{\bar{c}}$

r, 1-r: the relative importance of maximising $\sum_{k \in K} d_k x_k / \bar{d}$ and $\sum_{q \in Q} c_q f_q / \bar{c}$, respectively. $r \in [0,1]$

$\bar{d}$: maximum possible patient visits score of the network obtained after the investment

$\bar{c}$: maximum possible continuity of care score of the network obtained after the investment

Q: set of non-zero flow paths in the network indexed by *q*

x_k: 1 if an RWC is placed at RWC location *k*, 0 otherwise

d_k: expected patient visits score for an RWC at location *k*, defined as wdt*dt_k+wds*ds_k+wdl*dl_k (see assumption A8)

---

[2] If a set of locations K, KC, KP, or another subset of locations contains the index *q*, we refer to the subset of these locations which are along path q.

p: number of new RWCs to be located. $p \in \{0,...,|kP|\}$

f_q: size of the flow along path *q*

c_q: continuity of care score of path *q*

The objective function (5.1) maximises a function of the patient visits score and the continuity of care score of the network obtained after adding *p* RWCs to the initial network. Constraint (5.2) defines the continuity of care score of each path *q* based on the values of x_k, i.e. the locations of the RWCs that are located along path *q* (if any). Constraint (5.3) ensures that exactly *p* new RWCs are allocated to a potential RWC location. The RWCs (and RWC equivalents) that are already in the network are incorporated by means of constraint (5.4). Here, the value of x_k is forced to be equal to 1 if an RWC is already situated at location *k*. Finally, constraint (5.5) stipulates that an RWC is either allocated to location *k* or not.

Because our problem has to deal with two conflicting objectives, one has to answer the question: how important do I regard the maximisation of the patient visits score vs. the maximisation of the continuity of care score? This question can be answered by means of the parameter *r*. The smaller (larger) this parameter, the larger the importance of maximising the patient visits score (the continuity of care score). This parameter is defined as follows. Suppose that the maximum attainable patient visits score after adding *p* RWCs is $\bar{d}$, and that the maximum attainable continuity of care score after adding *p* RWCs is $\bar{c}$. Then, getting the patient visits score 1 percent-point closer to $\bar{d}$ is (1-r)/r times as important as getting the continuity of care score 1 percent-point closer to $\bar{c}$. For example, suppose that *r* is equal to 0.2. Then one regards a 4 percent-point (from $\bar{d}$) increase in the patient visits score as being equal to a 1 percent-point (from $\bar{c}$) increase in continuity of care score.

In order to solve the RIM, problem (5.1) – (5.5) has to be run *three times*. The first is meant to determine the value of $\bar{d}$, which is equal to Z after solving (5.1) – (5.5) with v_1=1 and v_2=0. The second is meant to determine the value of $\bar{c}$, which is equal to Z after solving (5.1) – (5.5) with v_1=0 and v_2=1. Given these values, (5.1) – (5.5) has to be solved again to obtain the optimal way to invest in the network of RWCs.

The function $g(\{x_k \mid k \in K^q\})$ calculates the continuity of care score of path *q*, based on the criteria mentioned in assumption A11.1 or A11.2. In the sections 5.2 and 5.3 we present how this score could be calculated and how it could be integrated in the model.

5.2 Calculation of continuity of care score 1

We propose three approaches to define the value of $g(\{x_k \mid k \in K^q\})$, the continuity of care score of path *q*, based on assumption A11.1. In section 5.2.1 we integrate this score within the basic problem itself. We call this the *integrated approach*. The main advantage of this approach is that it determines the continuity of care score of a path exactly as proposed in assumption A11.1. The main weakness of this approach is its complexity: many additional variables and constraints are needed to include the calculation of the continuity of care score 1 in (5.1) – (5.5).

Therefore, we present a second way to model the value of $g(\{x_k \mid k \in K^q\})$ in section 5.2.2. Here, the set of driving times between RWC neighbours along path *q* is classified into a couple of scenarios. Every scenario corresponds to a pre-defined continuity of care score. So, if the set of driving times along path *q* is assigned to a certain scenario, this means that the continuity of care score of that path is equal to the score of this scenario. We call this the *scenario approach* from now on. Also this approach has the disadvantage that many additional variables and constraints are needed to determine which scenario should be assigned to a set of driving times along path q. Moreover, this approach uses a step-wise linear function between the continuity of care score of path *q* and the driving times between RWC neighbours. As we explain later, a piece-wise linear function, which is used in the scenario approach, would be more appropriate.

The third approach we use to model the value of $g(\{x_k \mid k \in K^q\})$ is totally different from the others. We use the idea behind the Flow-Refuelling Location Model (FRLM) to calculate this value. Therefore, we call this the FRLM approach. The advantage is that this model allows us to input the score of $g(\{x_k \mid k \in K^q\})$ for each vector of x_k-values *as a parameter*. The main disadvantage of using the FRLM to model this value is that the number of 'relevant' combinations of x_k-values may explode when the number of RWC locations along a path grows (in section 5.2.3 we explain what is meant with 'relevant'). Because the number of variables and constraints depends on the number of relevant combinations, using the FLRM approach may result in very large models.

5.2.1 Integrated approach

The idea of the integrated approach is to identify the set of driving times between RWC neighbours by means of an IP model, and to score this set of driving times afterwards. In A5 we assume that we know the driving times between all locations. So, we only have to find out which RWCs are passed along a path and in what sequence these are passed. We make use of the definition of RWC neighbours described in definition 4.1 in order to do this.

The number of RWCs that are situated along path q is equal to $\sum_{k \in K^q} x_k$. Because we assume RWC equivalents at the origin and destination of a flow, the number of (unordered) RWC neighbour-pairs, n_q, is equal to $\sum_{k \in K^q} x_k - 1$. For example, consider again case 3 described in figure 4.3. Eight RWCs are situated there, whereas seven RWC neighbour-pairs can be defined: (1,2); (2,3); (3,4); (4,5); (5,6); (6,7); (7,8)

In order to integrate the continuum of care 1 of path q in model (5.1) - (5.5), the integrated approach adds the following constraints:

$$g(\{x_k \mid k \in K^q\}) = w_1 h_1(\overline{t_q}) + w_2 \sum_{k \in K^q} \sum_{l \in K^{qk}} i_{klq} h_2(t_{kl} - \overline{t_q}) + w_3 h_3(t_q^{\max}) \quad \forall q \in Q \tag{5.6}$$

$$i_{klq} = 1_{(k,l) \in N(\{x_k \mid k \in K^q\})} \quad \forall q \in Q, k \in K^q, l \in K^{qk} \tag{5.7}$$

$$\overline{t_q} = \sum_{k \in K^q} \sum_{l \in K^{qk}} i_{klq} t_{kl} \Big/ \left(\sum_{k \in K^q} x_k - 1 \right) \quad \forall q \in Q \tag{5.8}$$

$$t_q^{\max} = \max_{k \in K^q, l \in K^{qk}} \{ i_{klq} t_{kl} \} \quad \forall q \in Q, k \in K^q, l \in K^{qk} \tag{5.9}$$

Here, the new terms are:

$N(\{x_k \mid k \in K^q\})$: the set of RWC neighbour-pairs at path q, given the values of x_k

K^{qk}: the set of RWC locations along path q that are passed *after* passing location k

i_{klq}: 1 if locations k and l contain an RWC neighbour-pair (see definition 4.1) at path q, 0 otherwise. These variables are only defined for set of pairs $\{(k,l\} \in K^q \times K^q \mid l \in K^{qk}\})$

t_{kl}: the driving time between location k and location l (see definition 4.1), if $k \neq$ l
the time one has to spend at location k involuntarily, if $k = l$

$\overline{t_q}$: the average of all driving times between neighbour RWCs along path q

$t_q^{\max}$: the maximum among all driving times between neighbour RWCs along path q

w_1, w_2, w_3: weigh factors $\in [0,1]$; $w_1+w_2+w_3 = 1$

Constraint (5.6) defines the continuity of care score along path q. By means of constraint (5.7) we define whether locations k and l can be regarded as RWC neighbours at path q ($i_{klq} = 1$), or not ($i_{klq} = 0$). In constraint (5.8) and (5.9) we define the average of the driving times between RWC neighbours along path q and the maximum among these driving times, respectively. Next, we explain these constraints and the way we linearized them in detail.

Constraint (5.6)

Constraint (5.6) consists of three parts, analogous to the three sub-objectives stated in assumption A11.1. The first, $w_1 h_1(\overline{t_q})$, provides the score for the sub-objective of minimizing the average of all driving times between neighbour RWCs along path *q*: $\overline{t_q}$. The second, $w_2 \sum_{k \in K^q} \sum_{l \in K^{qk}} i_{klq} h_2(t_{kl} - \overline{t_q})$, provides the score for the sub-objective of minimizing the difference between such driving time and this average: $t_{kl} - \overline{t_q}$. That is, the objective of minimising the variance among these driving times. The third, $w_3 h_3(t_q^{\max})$, provides the score for the sub-objective of minimizing the maximum among these driving times, $t_q^{\max}$.

We define the functions $h_1(\overline{t_q}), h_2(t_{kl} - \overline{t_q})$ and $h_3(t_q^{\max})$, based on four *desired properties*. First, these functions should be piecewise linear in order to make solving the model easier. Second, the resulting value for $g(\{x_k \mid k \in K^q\})$ should be between 0 and 1, corresponding to 'no continuum of care' and 'maximum continuum of care', respectively. We ensure this by choosing the values of w_1, w_2, and w_3 in such a way that the sum of these values is equal to 1, *and* by ensuring that the factors these weights are multiplied with in constraint (5.6) take values between 0 and 1.

Third, $h_1(\overline{t_q})$ and $h_3(t_q^{\max})$ must be non-increasing functions in $\overline{t_q}$ and $t_q^{\max}$, respectively. The reason is that a larger average (or maximum) of the driving times between neighbour RWCs does not improve the continuity of care ensured to truck drivers travelling along that path. As explained in section 4, $h_1(\overline{t_q})$ should take the value 1 when $\overline{t_q}$ is smaller than $\hat{t}_1$, and 0 when $\overline{t_q}$ is larger than $\hat{t}_2$. The same holds for the relation between $h_3(t_q^{\max})$ and $t_q^{\max}$.

Last, the function $h_2(|t_{kl} - \overline{t_q}|)$ must be concave. Here, the reason is that we would like to have the variance among these driving times as small as possible. A larger deviation from the average of the driving times should not increase the score for continuity of care along a path. When this deviation is larger than $\hat{t}_3$, $h_2(t_{kl} - \overline{t_q})$ should take the value 0. In order to ensure that the factor w_2 is multiplied with is between 0 and 1, this function should take its maximum value $1/n_q$ at 0, where $n_q = \sum_{k \in K^q} x_k - 1$.

Based on these properties, we propose the following piecewise linear functions:

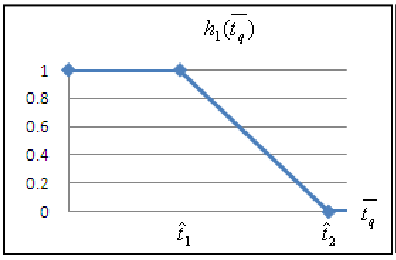


Figure 5.1: function h_1

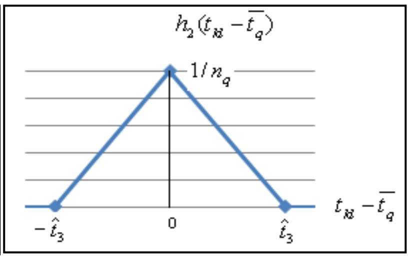


Figure 5.2: function h_2

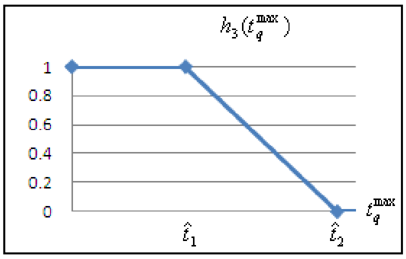


Figure 5.3: function h_3

In order to integrate this function in the model, we replace constraint (5.6) by the following linear constraints:

$$g(\{x_k \mid k \in K^q\}) = w_1(\lambda 1_{1q} + \lambda 1_{2q}) + w_2 \sum_{k \in K^q} \sum_{l \in K^{qk}} ih_{klq} + w_3(\lambda 3_{1q} + \lambda 3_{2q}) \quad \text{(5.6A)}$$

$$\lambda 1_{1q} + \lambda 1_{2q} + \lambda 1_{3q} + \lambda 1_{4q} = 1 \quad \forall q \in Q \quad \text{(5.6B)}$$

$$\lambda 1_{1q} 0 + \lambda 1_{2q} \hat{t}_1 + \lambda 1_{3q} \hat{t}_2 + \lambda 1_{4q} M_q = \overline{t_q} \quad \forall q \in Q \quad \text{(5.6C)}$$

$$\lambda 1_{iq} \le z1_{iq} \quad \forall q \in Q, i \in \{1,2,3,4\} \quad \text{(5.6D)}$$

$$\{z1_{1q}, z1_{2q}, z1_{3q}, z1_{4q}\} \in AS \quad \forall q \in Q \quad \text{(5.6E)}$$

$$\lambda 3_{1q} + \lambda 3_{2q} + \lambda 3_{3q} + \lambda 3_{4q} = 1 \quad \forall q \in Q \quad \text{(5.6F)}$$

$$\lambda 3_{1q} 0 + \lambda 3_{2q} \hat{t}_1 + \lambda 3_{3q} \hat{t}_2 + \lambda 3_{4q} M_q = t_q^{\max} \quad \forall q \in Q \quad \text{(5.6G)}$$

$$\lambda 3_{iq} \le z3_{iq} \quad \forall q \in Q, i \in \{1,2,3,4\} \quad \text{(5.6H)}$$

$$\{z3_{1q}, z3_{2q}, z3_{3q}, z3_{4q}\} \in AS \quad \forall q \in Q \quad \text{(5.6I)}$$

$$\lambda 1_{iq}, \lambda 3_{iq} \ge 0 \quad \forall q \in Q, i \in \{1,2,3,4\} \quad \text{(5.6J)}$$

$$z1_{iq}, z3_{iq} \in \{0,1\} \quad \forall q \in Q, i \in \{1,2,3,4\} \quad \text{(5.6K)}$$

Where:

ih_{klq} : linear representation of $i_{klq} h_2(t_{kl} - \overline{t_q})$ (see appendix A)

$\lambda 1_{iq}, \lambda 3_{iq}$: multiplier variables

$z1_{iq}$: variables indicating whether $\overline{t_q}$ is adjacent to point i

$z3_{iq}$: variables indicating whether $t_q^{\max}$ is adjacent to point i

AS : the set of all vectors of 4 binary variables, for which holds that only binary variable i is equal to 1 *and* binary variable i+1 take the value 1, for some $i \in \{1,2,3\}$. E.g. {0,1,1,0} is an element of AS.

M_q: an upper bound on $\overline{t_q}$ and $t_q^{\max}$ (e.g. the driving time from the origin to the destination of path q)

Suppose that the value $\overline{t_q}$ lies between the points $t_{(1)}$ and $t_{(2)}$, taking the values $h(t_{(1)})$ and $h(t_{(2)})$. There are unique values of $\lambda_{(1)}$ and $\lambda_{(2)}$, such that $\overline{t_q} = \lambda_{(1)} t_{(1)} + \lambda_{(2)} t_{(2)}$. If the function $h(\overline{t_q})$ is linear between these points, it is also true that $h(\overline{t_q}) = h(\lambda_{(1)} t_{(1)}) + h(\lambda_{(2)} t_{(2)})$.

This is also the idea behind the way we modelled constraint (5.6). Constraints (5.6B) – (5.6E) define $\overline{t_q}$ as a linear combination of two *adjacent* points among 0, $\hat{t}_1$, $\hat{t}_2$, and M_q. Similarly, constraints (5.6G) – (5.6I) define $t_q^{\max}$ as a linear combination of two adjacent points among 0, $\hat{t}_1$, $\hat{t}_2$, and M_q. Because the functions h_1 and h_3 are linear between all of these adjacent points, the values of $h_1(\overline{t_q})$ and $h_3(t_q^{\max})$ can be found by taking the same linear combination of the *function values* of these adjacent points. This is done in constraint (5.6A).

Only the way to model $i_{klq} h_2(t_{kl} - \overline{t_q})$ is not described yet. However, the linearization of this function results in many additional variables and constraints. Because this makes computation times increase a lot, it may be undesirable to include this function in the model. In appendix A the way this is modelled and linearized can be found.

Constraint (5.7)

By means of constraint (5.7) we define whether locations *k* and *l* can be regarded as an RWC neighbour-pair along path *q* ($i_{klq} = 1$), or not ($i_{klq} = 0$). We call the RWCs at locations *k* and *l* from such RWC neighbour-pair the *predecessor* and the *successor*, respectively. The following constraints ensure that the variables i_{klq} take the true values, given the x_k-values.

$$\sum_{l \in K^q} i_{klq} = x_k \quad \forall q \in Q, k \in KA^q \tag{5.7A}$$

$$\sum_{l \in K^q} i_{klq} = 1 \quad \forall q \in Q, k \in KO^q \tag{5.7B}$$

$$\sum_{k \in K^q} i_{klq} = x_l \quad \forall q \in Q, l \in KA^q \tag{5.7C}$$

$$\sum_{k \in K^q} i_{klq} = 1 \quad \forall q \in Q, l \in KD^q \tag{5.7D}$$

$$i_{klq} \in \{0,1\} \quad \forall q \in Q, k \in K^q, l \in K^{qk} \tag{5.7E}$$

Here, the new terms are:

KA: K-KOD: the set of the locations of all RWCs that are *not at the origin or destination of a flow*, indexed by *k*. KA $\subset$ K

KOD: KO$\cup$KD: the set of RWC locations situated at the origins KO, united with the set of RWC locations situated at the destinations KD of all paths, indexed by *k*

Theorem 5.1 *Constraints (5.7A) – (5.7E) ensure that* $i_{klq} = 1$ *if and only if locations k and l are RWC neighbours along path q*

Proof: First, observe that constraints (5.7A) – (5.7D) force that two locations can only be regarded as RWC neighbours if there are (dummy-) RWCs placed at both of them. Second, constraints (5.7A) and (5.7C) force that all RWCs that are located along the path have exactly one predecessor and one successor at a path. Taking the origin (which has only a successor: constraint (5.7B)) and the destination (which has only a predecessor: constraint (5.7D)) of this path into account, we come to a number of $n_q = \sum_{k \in K^q} x_k - 1$ predecessors and the same number of successors on a path. This proves that for a given path q, constraints (5.7A) – (5.7E) ensure that the number of predecessor – successor pairs is exactly n_q. As explained above, this is the true number. Last, because $i_{klq} \in \{0,1\}$, these n_q predecessor – successor pairs refer to n_q *different* predecessor – successor pairs. I.e. n_q *different* i_{klq} variables get the value 1

The only thing that is left to be proven is that constraints (5.7A) - (5.7E) automatically select the *true* n_q *predecessor-successor pairs* from the RWCs that are placed along flow q. We do this by means of *induction*. For simplicity, when we talk about *RWC k being a neighbour of RWC l at flow q*, we mean to say that $i_{klq} = 1$.

Consider the first RWC 'k_1' a truck driver at path *q* will pass after 'passing' the dummy-RWC 'k_0' at the origin of path *q* (O^q). Constraint (5.7C) ensures that exactly 1 other RWC is the neighbour of 'k_1'. This neighbour RWC is forced to be passed *before* 'k_1', because we only defined i_{klq} if location *l* is passed after location k. The only RWC that can be regarded as a neighbour of 'k_1' is the dummy-RWC at O^q. So, $i_{k_0 k_1 q}$ is forced to be 1, which is true. Next, consider the RWC 'k_2' a truck driver at path *q* will pass after passing RWC 'k_1'. Again, constraint (5.7C) ensures that exactly 1 other RWC is the neighbour of 'k_2'. The definition of i_{klq} forces that this RWC should be passed *before* 'k_2'. Because constraint (5.7A) prevents dummy-RWC 'k_0' from being the predecessor of more than 1 RWC, RWC 'k_1' is the only one that can possibly be regarded as the predecessor of ‘k_2’. So, $i_{k_1k_2q}$ is forced to be 1, which is true. Now consider RWC ‘k_i’, where i > 1. Because constraint (5.7A) prevents RWCs $k_0,\ldots,$ k_{i-2} from being the predecessor of more than 1 RWC, RWC ‘k_{i-1}’ is the only one that can be regarded as a predecessor of ‘k_i’. So, $i_{k_ik_{i-1}q}$ is forced to be 1 for all i, which is true. This completes the proof that constraints (5.7A) - (5.7E) automatically assigns $i_{klq}=1$ for the *true* predecessor-successor pairs (k,l) from the RWCs that are placed along flow q.

So, we have proven that constraints (5.7A) – (5.7E) select exactly n_q *different* i_{klq} variables to get the value 1 *and* that these variables refer to n_q *true* RWC neighbour-pairs. Obviously, this implicates that $i_{klq}=1$ *if and only if* locations *k* and *l* contain an RWC neighbour-pair ☐

Constraint (5.8)

In constraint (5.8) we define the average of the driving times between RWC neighbours along path q $\overline{t_q}$. This can be obtained by T_q / n_q, where T_q is the sum of the driving times between the RWC neighbours along this path. This definition is highly non-linear, because both T_q and n_q are variables. We can linearize constraint (5.7) by replacing it by the following constraints:

$$\overline{t_q} = \sum_{j\in J^q} nT_{jq} / \theta_j \tag{5.8A}$$

$$T_q = \sum_{k\in K^q}\sum_{l\in K^{qk}} i_{klq}t_{kl} \quad \forall q\in Q \tag{5.8B}$$

$$\theta_j n_{jq} = \sum_{k\in K^q} x_k - 1 \quad \forall q\in Q, j\in J^q \tag{5.8C}$$

$$nT_{jq} \le n_{jq} \quad \forall q\in Q, j\in J^q \tag{5.8D}$$

$$nT_{jq} \le T_q \quad \forall q\in Q, j\in J^q \tag{5.8E}$$

$$nT_{jq} \ge T_q - \hat{T}_q(1-n_{jq}) \quad \forall q\in Q, j\in J^q \tag{5.8F}$$

$$n_{jq} \in \{0,1\} \quad \forall q\in Q, j\in J^q \tag{5.8G}$$

Where:

J^q: $\{0,1,\ldots,|K^q|-1\}$ the set of all possible numbers of RWC neighbour-pairs defined along path *q*, indexed by *j*

n_{jq}: 1 if θ_j neighbour-pairs can be defined along path *q* (i.e. if the number of RWCs along path *q* is θ_j -1), 0 otherwise

θ_j : parameter which is equal to the number of RWC neighbour-pairs element *j* corresponds to

T_q : sum of the driving times between the RWC neighbours along path *q*

$\hat{T}_q$: an upper bound on the value of T_q (e.g. the driving time from the origin to the destination of path *q*)

nT_{jq} : T_q if the variable n_{jq} is equal to 1, 0 otherwise

Constraint (5.8C) stipulates that the variable n_{jq} is only equal to 1 if the number of neighbour-pairs along path *q* is equal to θ_j. Constraint (5.8A) defines the average of the driving times between RWC neighbours along path *q*, by dividing the total driving time T_q by the number of neighbour-pairs. This division is equivalent to the multiplication $T_q n_{jq} / \theta_j$ if n_{jq} is equal to 1. In order to linearize the multiplication $T_q n_{jq}$, we replace this with the variable nT_{jq}, which is equal to T_q if n_{jq} is equal to 1, and equal to 0 otherwise. This is ensured by means of constraints (5.8D) – (5.8F). Furthermore, constraint (5.8B) defines the value of T_q, and constraint (5.8.G) defines n_{jq} as binary variables.

Constraint (5.9)

This constraint defines the maximum among the driving times between RWC neighbours along path q, and can be replaced by:

$$t_q^{\max} \geq i_{klq} t_{kl} \quad \forall k \in K^q, l \in K^{qk}, q \in Q \tag{5.9A}$$

Constraint (5.9A) forces that the value of $t_q^{\max}$ is not smaller than all driving times between RWC neighbours along path q. In order to maximise the value of c_q, the value of $t_q^{\max}$ will also be forced to be as small as possible. So, we do not need to restrict $t_q^{\max}$ to be *equal to* the largest among the driving times.

5.2.2 Scenario approach

The idea behind the scenario approach is to assign each path to a scenario, based on the driving times between the set of RWC neighbours along the path. Each scenario corresponds to a pre-defined score for $g(\{x_k \mid k \in K^q\})$. The continuity of care of a path is equal to the pre-defined continuity of care score of the scenario this path is assigned to.

Each scenario is characterised by unique ranges for the three variables explained in assumption A11.1: the *average* of the driving times between neighbour RWCs along path q, the *difference* between such driving time and this average, and the *maximum* among these driving times. If all of these variables take a value that lies within a given range defined in such scenario, this path is allowed to be classified into that scenario.

Next to the constraints (5.7A) – (5.7E), (5.8A) – (5.8G), and (5.9A), the scenario approach adds the following constraints to model (5.1) – (5.5):

$$g(\{x_k \mid k \in K^q\}) = \sum_{s \in S} c_s y_{sq} \quad \forall q \in Q \tag{5.10}$$

$$\sum_{s \in S} y_{sq} \leq 1 \quad \forall q \in Q \tag{5.11}$$

$$\overline{t_q} \leq y_{sq}\overline{\tau_s} + (1 - y_{sq})M_q \quad \forall q \in Q, s \in S \tag{5.12}$$

$$t_q^{\max} \leq y_{sq}\tau_s^{\max} + (1 - y_{sq})M_q \quad \forall q \in Q, s \in S \tag{5.13}$$

$$t_{klq}^{diff} \leq y_{sq}\tau_s^{diff} + (1 - y_{sq})M_q \quad \forall q \in Q, k \in K^q, l \in K^{qk}, s \in S \tag{5.14}$$

$$t_{klq}^{diff} \geq i_{klq} \mid t_{kl} - \overline{t_q} \mid \quad \forall q \in Q, k \in K^q, l \in K^{qk} \tag{5.15}$$

$$y_{sq} \in \{0,1\} \quad \forall q \in Q, s \in S \tag{5.16}$$

Where:

S : the set of scenarios, indexed by *s*

y_{sq} : 1 if path *q* is assigned to scenario *s*, 0 otherwise

c_s : the continuity of care score of a path, if it is assigned to scenario *s*

$\overline{\tau_s}$: the maximum value of $\overline{t_q}$ for which path *q* can be assigned to scenario *s*

$\tau_s^{\max}$: the maximum value of $t_q^{\max}$ for which path *q* can be assigned to scenario *s*

τ_s^{diff} : the maximum value of $\mid t_{kl} - \overline{t_q} \mid$ for all RWC neighbour-pairs $(k,l) \in N(\{x_k \mid k \in K^q\})$ for which path *q* can be assigned to scenario *s*

Constraint (5.10) defines the score for continuity of care along path *q*, depending on the scenario this path is assigned to. Based on the values of $\overline{t_q}$, $t_q^{\max}$, t_{klq}^{diff}, the scenarios this path can be assigned to are determined in constraints (5.12) – (5.15). Constraints (5.11) and (5.16) ensure that at most one scenario is assigned to a path. If multiple scenarios can be assigned, the one with the largest value of c_s is chosen automatically, because we want to maximise the value of $g(\{x_k \mid k \in K^q\})$ (see objective function (5.1) and constraint (5.2)).

In order to linearize constraint (5.15), it has to be replaced by the following constraints:

$$t_{klq}^{diff} \geq it_{klq}^{|diff|} \quad \forall k \in K^{q}, l \in K^{qk}, q \in Q \tag{5.15A}$$

$$it_{klq}^{|diff|} \leq i_{klq} M_{q} \quad \forall k \in K^{q}, l \in K^{qk}, q \in Q \tag{5.15B}$$

$$it_{klq}^{|diff|} \leq t_{klq}^{|diff|} \quad \forall k \in K^{q}, l \in K^{qk}, q \in Q \tag{5.15C}$$

$$it_{klq}^{|diff|} \geq t_{klq}^{|diff|} - M_{q}(1 - i_{klq}) \quad \forall k \in K^{q}, l \in K^{qk}, q \in Q \tag{5.15D}$$

$$t_{klq}^{|diff|} \geq t_{kl} - \overline{t_{q}} \quad \forall k \in K^{q}, l \in K^{qk}, q \in Q \tag{5.15E}$$

$$t_{klq}^{|diff|} \geq \overline{t_{q}} - t_{kl} \quad \forall k \in K^{q}, l \in K^{qk}, q \in Q \tag{5.15F}$$

Where:

$it_{klq}^{|diff|}$: linear representation of $i_{klq} \, | t_{kl} - \overline{t_{q}} \, |$

$t_{klq}^{|diff|}$: linear representation of $| t_{kl} - \overline{t_{q}} \, |$

The value of $| t_{kl} - \overline{t_{q}} \, |$ is linearized in constraints (5.15E) and (5.15F). Because a larger value of this absolute difference will never result in a larger objective value, these constraints force that $t_{klq}^{|diff|}$ is chosen as small as possible. So, we do not need to restrict $t_{klq}^{|diff|}$ to be *equal to* the absolute difference. Constraints (5.15A) – (5.15D) ensure that t_{klq}^{diff} is equal to $t_{klq}^{|diff|}$ if i_{klq} is equal to 1, and that t_{klq}^{diff} is equal to 0 if i_{klq} is equal to 0.

5.2.3 Flow Refuelling Location Model (FRLM) approach

As explained in section 3.3, the Flow Refuelling Location Model exogenously determines which combinations of facilities along path *q* can refuel the flow along this path. Similarly, we exogenously determine the set of combinations of RWCs located along path *q* that result in a strictly positive score for ensuring a continuum of care along this path. This set is denoted by H^{q} and indexed by *h*.

Each combination of RWCs *h* is scored in exactly the same way as described in section 5.2.1. We determine the driving-times between RWC neighbours in such combination, the average of these times, $\overline{t_{q}}$, and the maximum among these times, $t_{q}^{\max}$. These values are evaluated in $h_{1}(\overline{t_{q}})$, $h_{2}(t_{kl} - \overline{t_{q}})$, and $h_{3}(t_{q}^{\max})$, respectively (see figures 5.1, 5.2 and 5.3). Next, we use constraint (5.6) to determine the continuity of care score along path *q*, when along this path combination of RWCs *h* is established. This score is represented by the parameter c_{hq}.

In order to include the continuity of care score 1 of path q in model (5.1) - (5.5), the FLRM approach adds the following constraints:

$$g(\{x_k \mid k \in K^q\}) = \sum_{h \in H^q} c_{hq} y_{hq} \quad \forall q \in Q \tag{5.17}$$

$$\sum_{h \in H^q} y_{hq} \leq 1 \quad \forall q \in Q \tag{5.18}$$

$$a_{hk} x_k \geq y_{hq} \quad \forall q \in Q, h \in H^q, k \in K^q \mid a_{hk} = 1 \tag{5.19}$$

$$0 \leq y_{hq} \quad \forall q \in Q, h \in H^q \tag{5.20}$$

Here, the new terms are:

H^q: set of all 'relevant' potential *combinations of RWCs* at path q (consisting of all current RWC locations along path q and *a set of* potential RWC locations along path q), indexed by h

a_{hk}: a coefficient equal to 1 if RWC location k is in combination h, 0 otherwise

y_{hq}: 1 if along path q combination of RWCs h is established, 0 otherwise

c_{hq}: continuity of care score of path q, when along this path combination of RWCs h is established

The continuity of care score of path q is determined in constraint (5.17), depending on the combination of RWCs h that is established along this path. Whether or not combination h can be regarded as established at path q is determined in constraint (5.19). Constraints (5.18) and (5.20) ensure that at most one 'relevant' combination of RWCs h is established at path q (i.e. y_{hq} is equal to 1 for at most one h in H^q). Though we do not restrict y_{hq} to be binary, this is automatically ensured by the model.

Theorem 5.2 *Constraints (5.17) – (5.20) force that the variables y_{hq} take binary values*

Proof: constraint (5.19) ensures that combination h can only be regarded as established, if RWCs are allocated to *all* locations defined within this combination. If this is not the case for a certain combination h, this constraint forces y_{hq} to take the value 0. The variables y_{hq} for the combinations h that *can* be regarded as established, are all allowed to take values between 0 and 1 because of constraints (5.18) and (5.20). Nevertheless, the model ends up with exactly 1 combination h for which y_{hq} is 1, while the others are zero. Namely, we want the term in (5.17) to be maximised (see objective function (5.1) and constraint (5.2)), while the sum of the y_{hq} variables must not exceed 1. This implicates that only the variable y_{hq} with *the largest value for c_{hq}* takes the value 1, while y_{hq} is 0 for all other combinations h. □

The number of combinations of RWCs that could possibly be established along path *q* grows exponentially when the number of current and potential RWC locations along this path ($|K^q|$) grows. This number can be decreased a lot by only taking the 'relevant' combinations into account. First, we do not include the combinations of RWCs *h* that do not contain all current RWC locations. Second, we decrease this number further by leaving all combinations *h* for which holds that $c_{hq} \approx 0$ out of consideration. The remaining combinations make up the set H^q.

5.2.4 Summary

The following sets of constraints have to be added to model (5.1) – (5.5) in order to include the calculation of the continuity of care score 1:

- The integrated approach: add constraints (5.6A) – (5.6K), (A1) – (A13) (see appendix A), (5.7A) – (5.7E), (5.8A) – (5.8E), and (5.9A)
- The scenario approach: add constraints (5.7A) – (5.7E), (5.8A) – (5.8E), (5.9A), (5.10) – (5.14), (5.15A) – (5.15F), and (5.16)
- The FLRM approach: add constraints (5.17) – (5.20)

The integrated approach results in a very large model (in terms of variables and constraints). In case that the scenario approach is used, this model is considerably smaller, but still very large. Another weakness is that this approach simplifies the definition of the continuity of care score. When the FRLM approach is applied to include this score, the size of the model grows exponentially in the number of RWC locations along a path.

Based on these observations, it is not surprising that some preliminary tests show that it is hard to solve the RIM with these three approaches to include continuity of care score 1. A good alternative is to change the definition of the continuity of care score a little. The impact of this change on the way to include the calculation of the continuity of care score in the RIM is described in the next subsection.

5.3 Calculation of continuity of care score 2

In this section we present three approaches to calculate the continuity of care score of a path and to implement it in model (5.1) – (5.5). The difference with section 5.2 is that the continuity of care score of a path is now based on the *expected RWC time* of this path, denoted by ER_q.

Section 5.3.1 explains how to integrate this score in model (5.1) – (5.5). The scenario approach is explained in section 5.3.2. Last, section 5.3.3 explains how to apply the FLRM in order to calculate the continuity of care score 2 of a path.

5.3.1 Integrated approach

Next to (5.7A) – (5.7E), the following constraints have to be added to model (5.1) – (5.5) in order to integrate the continuity of care score 2:

$$g(\{x_k \mid k \in K^q\}) = h_4(ER_q) \quad \forall q \in Q \tag{5.21}$$

$$ER_q = \frac{1}{2t_{O^qD^q}} \sum_{k \in K^q} \sum_{l \in K^{qk}} i_{klq} t_{kl}^{\;2} \quad \forall q \in Q \tag{5.22}$$

Where:

ER_q : expected RWC time of path q

Constraint (5.21) defines the continuity of care score 2 of path *q* as a function of ER_q. This expected RWC time of path *q* is defined in constraint (5.22). Next, we explain these constraints and the way we linearized them in detail.

Constraint (5.21)

We define the function $h_4(ER_q)$ based on the same *desired properties* as we based the functions $h_1(\overline{t_q})$ and $h_3(t_q^{\max})$ on (see section 5.2.1). First, this functions should be piecewise linear. Second, the resulting value for $g(\{x_k \mid k \in K^q\})$ should be between 0 and 1, corresponding to 'no continuum of care' and 'maximum continuum of care', respectively. Third, $h_4(ER_q)$ must be a non-increasing function of ER_q. The reason is that a larger expected RWC time does not improve the continuity of care ensured to truck drivers travelling along that path. Last, as explained in section 4, $h_4(ER_q)$ should take the value 1 when ER_q is smaller than $\hat{t}_4$, and equal to 0 when ER_q is larger than $\hat{t}_5$.

Based on these properties, we propose the following piecewise linear function for $h_4(ER_q)$:

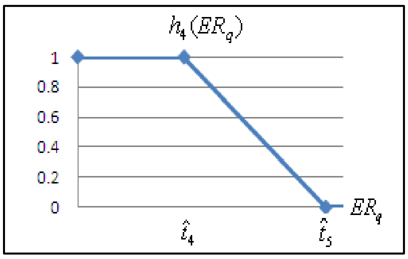


Figure 5.4: function h_4

In order to integrate this function in the model, we replace constraint (5.21) by the following linear constraints:

$$g(\{x_k \mid k \in K^q\}) = \lambda 4_{1q} + \lambda 4_{2q} \quad \forall q \in Q \tag{5.21A}$$

$$\lambda 4_{1q} + \lambda 4_{2q} + \lambda 4_{3q} + \lambda 4_{4q} = 1 \quad \forall q \in Q \tag{5.21B}$$

$$\lambda 4_{2q} \hat{t}_4 + \lambda 4_{3q} \hat{t}_5 + \lambda 4_{4q} M_q = ER_q \quad \forall q \in Q \tag{5.21C}$$

$$\lambda 4_{iq} \leq z4_{iq} \quad \forall q \in Q, i \in \{1,2,3,4\} \tag{5.21D}$$

$$\{z1_{1q}, z1_{2q}, z1_{3q}, z1_{4q}\} \in AS \quad \forall q \in Q \tag{5.21E}$$

$$\lambda 4_{iq} \geq 0 \quad \forall q \in Q, i \in \{1,2,3,4\} \tag{5.21F}$$

$$z4_{iq} \in \{0,1\} \quad \forall q \in Q, i \in \{1,2,3,4\} \tag{5.21G}$$

Here, the new terms are:

$\lambda 4_{iq}$: multiplier variables

$z4_{iq}$: variables indicating whether ER_q is adjacent to point i

Constraints (5.21B) – (5.21G) define ER_q as a linear combination of two *adjacent* points among 0, $\hat{t}_4$, $\hat{t}_5$, and M_q. Because the function h_4 is linear between all of these adjacent points, the value of $h_4(ER_q)$ can be found by taking the same linear combination of the *function values* of these adjacent points. This is done in constraint (5.21A).

Constraint (5.22)

Consider the trip of a truck driver along path *q*. We could visualise this trip by means of a time-line. Every time that an RWC is passed could be seen as an *event* on this time-line. Let us now define the part of the time line between two of these events as a *segment*. Such segment always refers to the trip from one RWC to another.

Next, consider a truck driver who is travelling along path *q*. Currently, he is in segment (*k*,*l*) within the time-line of this path. Suppose that we would draw a randomly chosen time within this segment. What is the expected time *from* this randomly chosen time *to* the end of the segment? This question is equivalent to: what is the expected RWC time for truck drivers in a certain segment of the time-line?

As we prove in theorem 5.3, the expected RWC time for truck drivers in segment (*k*,*l*) is equal to $\frac{1}{2} t_{kl}{}^2$. The expected RWC time for truck drivers along *path q* (i.e. for truck drivers in all segments) can be found by conditioning this variable on the segments defined in the path. Here, the probability that a truck driver is in segment (*k,l*) is equal to $\frac{t_{kl}}{t_{O^qD^q}}$. The resulting expression for the RWC time of a path is the one described in constraint (5.22). Next, we give a formal prove that this formula is correct.

Theorem 5.3 *The expected RWC time of a path, ER_q, can be calculated by:* $\frac{1}{2t_{O^qD^q}}\sum_{k\in K^q}\sum_{l\in K^{qk}} i_{klq}t_{kl}{}^2$

Proof: The timeline of a trip from O^q to D^q can be split up in two types of segments. First, a segment can refer to a trip from one RWC to another. Second, it can refer to the 'imaginary' trip from an RWC to the same RWC. This is the case when an involuntary stop (e.g. a stop made because one has to wait to cross a border or a bridge) is made at the location where the RWC is situated.

The expression for ER_q can be obtained by conditioning this variable on these segments: we treat each segment as a sub-path, and calculate the expected RWC time of that path. Let ER_{kl} be the expected RWC time for truck drivers who are in segment (*k*,*l*). I.e. they are travelling from location *k* to location *l*. Next, let ER_{kk} be the expected RWC time for truck drivers who are in segment (*k,k*). I.e. these are making an involuntary stop at this location. Last, let pr_{klq} be the probability that a truck driver is in segment (*k,l*) at a randomly chosen time in the time-line of path q.

The expression for ER_q can be derived as follows:

$$ER_q = \sum_{k\in K^q}\sum_{l\in K^{qk}} i_{klq}\, pr_{klq} ER_{kl} + \sum_{k\in K^q | t_{kk}>0} x_k\, pr_{kkq} ER_{kk} \tag{5.22A}$$

$$= \sum_{k\in K^q}\sum_{l\in K^{qk}} i_{klq} \frac{t_{kl}}{t_{O^qD^q}} ER_{kl} + \sum_{k\in K^q | t_{kk}>0} x_k \frac{t_{kk}}{t_{O^qD^q}} * 0 \tag{5.22B}$$

$$= \sum_{k\in K^q}\sum_{l\in K^{qk}} i_{klq} \frac{t_{kl}}{t_{O^qD^q}} \int_{tO_{qk}}^{tO_{ql}} \frac{er(tO)}{tO_{ql}-tO_{qk}}\, dtO \tag{5.22C}$$

$$= \sum_{k\in K^q}\sum_{l\in K^{qk}} i_{klq} \frac{t_{kl}}{t_{O^qD^q}} \int_{tO_{qk}}^{tO_{ql}} \frac{(tO_{ql}-tO)}{t_{kl}}\, dtO \tag{5.22D}$$

$$= \sum_{k\in K^q}\sum_{l\in K^{qk}} i_{klq} \frac{1}{t_{O^qD^q}} \left[tO_{ql}tO - \frac{1}{2}tO^2 \right]_{tO_{qk}}^{tO_{ql}} \tag{5.22E}$$

$$= \sum_{k\in K^q}\sum_{l\in K^{qk}} i_{klq} \frac{1}{2t_{O^qD^q}} (tO_{ql}-tO_{qk})^2 \tag{5.22F}$$

$$= \frac{1}{2t_{O^qD^q}} \sum_{k\in K^q}\sum_{l\in K^{qk}} i_{klq}t_{kl}{}^2 \tag{5.22G}$$

Here, the new terms are:

tO_{qk}: driving time from location k to the origin of path q.

ER_{kl}: expected RWC time for truck drivers who are in segment (k,l)

pr_{klq}: probability that a truck driver is in segment (k,l) at a randomly chosen time in the time-line of path q

Formula (5.22A) conditions ER_q on the two sets segments described above. The first set of segments is identified by looking at all variables i_{klq} that are equal to 1. The second set of segments is identified by defining all locations k where an RWC is located (i.e. x_k=1) and where an involuntary stop is made (i.e. t_{kk}>0). Obviously, the expected RWC time of such activity is 0. Therefore, ER_{kk} is 0 in (5.22B). The probability p_{klq} can be calculated by dividing the duration of the activity, t_{kl}, by the total duration of all activities, $t_{O^q D^q}$ (the driving time from the origin to the destination of path q).

The only thing that is left to be derived is the expression for ER_{kl}, which is done in (5.22C)-(5.22F). In (5.22C) we apply the definition of the expected RWC time stated in definition 4.3. From a given moment tO, the expected time to the next RWC l, er(tO), is equal to the difference between the driving times to the origin of these moments: tO_{ql}-tO. This explains formula (5.22D). Formulas (5.22E) and (5.22F) derive an expression for the integral stated in formula (5.22D). (5.22F) contains the final expression for ER_q. □

5.3.2 Scenario approach

As explained in section 5.2.2, the scenario approach assigns each path to a scenario. Each scenario corresponds to a pre-defined score for $g(\{x_k \mid k \in K^q\})$. Here, the assignment is based on the expected RWC time of the path.

Next to the constraints (5.7A) – (5.7E) and (5.22), the scenario approach adds the following constraints to model (5.1) – (5.5):

$$g(\{x_k \mid k \in K^q\}) = \sum_{s \in S} c_s y_{sq} \quad \forall q \in Q \tag{5.23}$$

$$\sum_{s \in S} y_{sq} \leq 1 \quad \forall q \in Q \tag{5.24}$$

$$ER_q > (1 - y_{sq}) ER_s \quad \forall q \in Q, s \in S \tag{5.25}$$

$$y_{sq} \in \{0,1\} \quad \forall q \in Q, s \in S \tag{5.26}$$

Here, the new term is:

ER_s: maximum value of ER_q for which path *q* can be assigned to scenario *s*

Constraint (5.23) defines the continuity of care score 2 of path *q*. This score depends on the scenario this path is assigned to. Based on the value of ER_q, the set of scenarios this path can be assigned to is determined in constraint (5.25). Constraints (5.24) and (5.26) ensure that at most 1 scenario is assigned to a path. Again, a path is automatically assigned to the feasible scenario (i.e. constraint (5.25) is met for the scenario) with the largest value of c_s, because we want to maximise the value of $g(\{x_k \mid k \in K^q\})$.

5.3.3 FLRM approach

Solving the RIM, using continuity of care score 2, can also be done by means of the FLRM approach described in section 5.2.3. As explained, the RIM has to be provided with the parameters c_{hq}: the continuity of care score of path *q* in case that combination of RWCs *h* is established. The only thing that needs to be changed in order to adapt this approach to continuity of care score 2 is to calculate the parameters c_{hq} based on this score. This can be done by calculating $h_4(ER_q)$ for path *q* for the case that combination of RWCs *h* is established.

5.3.4 Summary

The following sets of constraints have to be added to model (5.1) – (5.5) in order to include the calculation of the continuity of care score 2:

- The integrated approach: add constraints (5.7A) – (5.7E), (5.21A) – (5.21G), and (5.22)
- The scenario approach: add constraints (5.7A) – (5.7E), and (5.22) – (5.26)
- The FLRM approach: add constraints (5.17) – (5.20)

In terms of numbers of variables and constraints, including the continuity of care score 2 instead of continuity of care score 1 is better in case that the integrated approach or the scenario approach is used. For the FLRM approach, the model size is not influenced by the definition of this score. This implies that this approach still brings about the disadvantage that the size of the model grows exponentially in the number of RWC locations along a path.

The advantage that the scenario approach results in much smaller models than the integrated approach does not hold for the continuity of care score 2. Now the sizes of these models are almost equal. If both models can be solved equally fast too (which is not necessarily the case), we prefer the integrated approach over the scenario approach. Namely, the scenario approach simplifies the definition of the continuity of care score.

5.4 Extensions

In this subsection we propose three extensions to model (5.1) – (5.5). First, subsection 5.4.1 describes how to adapt the RIM such that it can determine the optimal *pD* RWCs which should be removed. Second, we extend the RIM to the RWC & Staff Investment Model (RSIM), which optimises the way to invest a certain budget in RWCs (determining their locations) and in additional employees. This extension is described in subsection 5.4.2. Last, in subsection 5.4.3 we adapt the RIM and the RSIM for the case that assumption A10.1 (RWC equivalents at the origin and at the destination of each flow) does not hold.

5.4.1 De-investments

In some cases, North Star has a budget decrease. Therefore, they have to perform a de-investment in the network of RWCs, by closing RWCs. In other cases, they want to re-invest part of the network, by moving RWCs to a different location. This section explains how to adapt model (5.1)-(5.5) in order to deal with these cases.

Let *pD* be the number of RWCs to be removed from the network. In order to generalise model (5.1)-(5.5) such that it can handle investments, de-investments and re-investments, it should be replaced with the following model:

$$\max\ Z = v_1 \sum_{k \in K} d_k x_k + v_2 \sum_{q \in Q} f_q c_q \tag{5.27}$$

$$c_q = g(\{x_k \mid k \in K^q\}) \quad \forall q \in Q \tag{5.28}$$

$$\sum_{k \in K} x_k = |KC| + p - pD \tag{5.29}$$

$$\sum_{k \in KC} x_k \geq |KC| - pD \tag{5.30}$$

$$\sum_{k \in KP} x_k \leq p \tag{5.31}$$

$$x_k = 1 \quad \forall k \in KEQ \tag{5.32}$$

$$x_k \in \{0,1\} \quad \forall k \in K \tag{5.33}$$

Where:

pD: number of RWCs to be removed. $pD \in \{0,1,\ldots,|kC| - |KEQ|\}$

KEQ = set of locations at which RWC equivalents are situated. KEQ $\subseteq$ KC

Objective function (5.27) maximises a function of the patient visits score and the continuity of care score of the network obtained after making an investment, a de-investment or a re-investment. Constraint (5.28) defines the continuity of care score of path *q*. One of the approaches described in section 5.2 or 5.3 can be used to implement this score in model (5.27) – (5.32). Next, (5.29) stipulates that the number of established RWCs in the resulting network is equal to the current number of established RWCs minus the number of removed RWCs, plus the number of new RWCs. In the resulting network, *at most pD* RWCs are removed and *at most* p new RWCs are placed. This is ensured by constraints (5.30) and (5.31), respectively. Together, constraints (5.29) – (5.31) ensure that:

1. p RWCs are added to the network in case of an investment (i.e. p>0, pD=0)
2. pD RWCs are removed in case of a de-investment (i.e. p=0, pD>0)
3. at most min{p,pD} RWCs are moved to another place, exactly p-min{p,pD} new RWCs are established, and exactly pD-min{p,pD} RWCs are removed in case of a re-investment (i.e. p>0, pD>0). Note that it is not necessarily beneficial to move all RWCs to another place.

Last, constraint (5.32) stipulates that RWC equivalents cannot be removed and (5.33) defines that either an RWC is established at location *k* or not.

Summary

The RIM, as described in (5.1) – (5.5), optimises the investment in *p* new RWCs. However, North Star sometimes wants to optimise the de-investment in pD RWCs or to optimise the re-investment of at most min{p,pD} RWCs. Model (5.27) – (5.33) extends the RIM such that it can solve these problems.

5.4.2 RWC & Staff Investment Model (RSIM)

In the RWC Investment Model, described in (5.1)-(5.5) or (5.27)-(5.33), we implicitly assume that the only way to invest in the network is to add a fixed number of RWCs to the network. As we explained in section 2, another way to invest in the network of RWCs is to hire additional nurses or outreach workers. This increases the number of patient visits because of increased capacity and because these employees can actively access truck drivers to promote North Star's activities. In this section, we include the decision to invest in additional RWC staff members in the RIM. The resulting model is called the RWC & Staff Investment Model (RSIM) from now on.

Problem Definition

The RIM restricts that the number of RWCs to be placed cannot exceed a given number. Instead, the RSIM restricts that the total costs of the investments in the network cannot exceed a given budget. Therefore, we replace assumption A1.1 by the following:

A1.2: The fitness score of an investment in the network of RWCs *only* depends on:

1. The resulting change in the total number of truck drivers, sex workers and locals visiting the RWCs in the network per day
2. The resulting change in the extent to what a continuum of care is ensured to the truck drivers who are travelling through the network
3. Whether or not the fixed yearly costs + the variable yearly costs (which partly depend on the numbers of patient visits) exceed the budget

Replacing assumption A1.1 by A1.2 has its implications for the set of issues affecting the fitness of an investment (for the RIM: see figure 4.3). For the RSIM, these are summarised in figure 5.5. In short, the problem is to choose how to invest, de-invest or re-invest a certain budget in RWCs and staff, such that the fitness of the resulting network is maximised. The 'fitness value' is defined as a function of the resulting change in the patient visits score of the network and the resulting change in the continuity of care scores ensured to truck drivers travelling through the network. Either continuity of care score 1 or continuity of care score 2 can be applied to determine these. Though not included in the objective function, the financial issues implicitly determine the fitness of an investment. This is ensured by a budget constraint, which determines whether an investment is allowed or not.

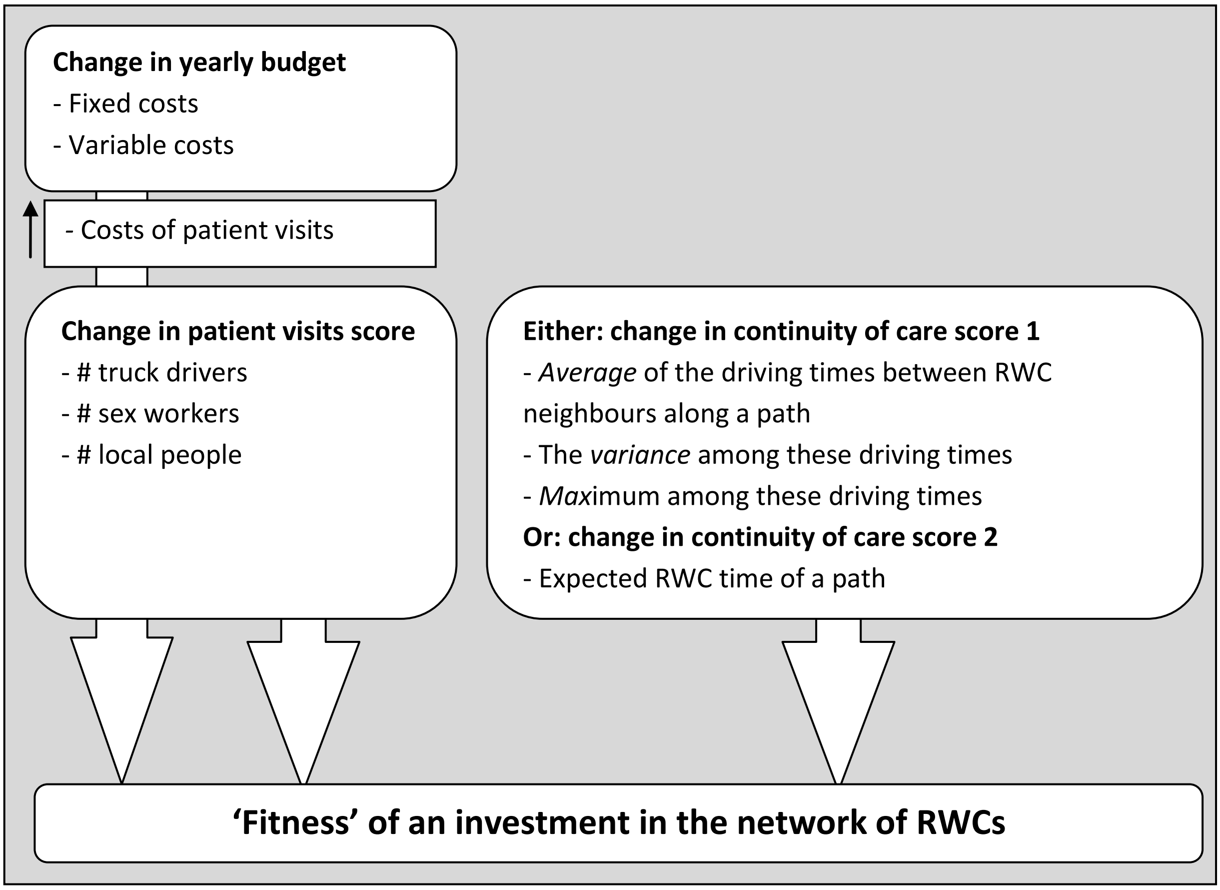


Figure 5.5: Issues determining the fitness the RSIM assigns to an investment in the network of RWCs

The model

Let *b* be the yearly budget which is spent on the current network of RWCs. Three types of decisions can be taken about this budget. First, one could decide to increase this budget with *bI*. Then the problem is to optimise the way to *invest* at most this budget increase by establishing new RWCs and by hiring additional employees. Second, the budget could be decreased with *bD*. In that case, the problem becomes to optimise the way to *de-invest* at least this budget decrease by removing RWCs and by hiring employees. Last, one could decide to re-invest *bR*. This means that at most bR of the costs incurred by the current RWCs and employees are cut. Next, at most these cut costs are spent by establishing new RWCs and by hiring new employees. Eventually, one could decide to combine a re-investment with an investment *or* a de-investment. Again, the objective is to invest, de-invest, or re-invest in such a way that a function of the patient visits score and the continuity of care score of the resulting network is maximised.

In order to perform investments, de-investments, or re-investments in RWCs and employees, we have to know something about their costs. As described in section 2.1, costs can be split up in fixed yearly costs (e.g. amortisation, rent, maintenance) and variable yearly costs (e.g. wages, medicines, electricity). The fixed yearly costs, which are incurred when an RWC is established at location k, are represented by bf_k. As explained, the variable costs depend on the number of employees occupying a location (for simplicity, we do not split up the employees in nurses and outreach workers). Let E be the set of possible numbers of employees occupying an RWC, indexed by *e*. The parameter η_e is equal to the number of FT employees which element *e* corresponds to. Given that η_e FT employees occupy location *k*, the variable yearly costs incurred at this location are represented by bv_{ke}.

When changing the number of employees at a location *k*, also the patient visits score of that location changes. Therefore it does not suffice to provide the model with *one* patient visits score any more. Instead, one has to provide the parameters d_{ke}, the patient visits score of location *k* in case that η_e FT employees occupy an RWC at that location.

The decision whether to employ η_e FT employees at location *k*, can be made by means of the binary decision variables xe_{ke}. Let $ec_k \in$ E be the element corresponding to the current number of employees at location *k*, η_{ec_k}. An investment in employees is made when this number is increased to some $\eta_e > \eta_{ec_k}$. That is, xe_{ke} is 1 for some $e \in EI^k$. A de-investment in employees is made by decreasing the number of FT employees at location *k* to some $\eta_e < \eta_{ec_k}$. That is, xe_{ke} is 1 for some $e \in ED^k$. The sets EI^k and ED^k are finite sets. First, North Star only hires Full-Time (FT) employees or 0.5 FT employees (i.e. Part-Time employees). Second, the total number of FT employees occupying an RWC can be at most $\overline{\eta}$, because of the limited capacity of an RWC.

As listed above, the RIM and the RSIM require different data. Therefore, we replace assumption A4.1 by assumptions A4.2 – A4.4.

A4.2: Estimates of d_{ke}, the patient visits score of location *k*, if η_e FT employees occupy an RWC at that location

A4.3: Estimates of bf_k, the fixed yearly costs incurred when an RWC is established at location *k*

A4.4: Estimates of bv_{ke}, the variable yearly costs incurred when η_e FT employees occupy an RWC at location *k*

Based on assumptions A1.2, A2-A3, A4.2-A4.4, A5-A10.1, A11.1 or A11.2, and A12, the problem of investing, de-investing or re-investing in the network of RWCs can be modelled as follows:

$$\max Z = v_1 \sum_{k\in\{K-KEQ\}} \sum_{e\in E} d_{ke} xe_{ke} + v_2 \sum_{q\in Q} f_q c_q \tag{5.34}$$

$$c_q = g(\{x_k \mid k \in K^q\}) \quad \forall q \in Q \tag{5.35}$$

$$\sum_{k\in K} bf_k x_k + \sum_{k\in\{K-KEQ\}} \sum_{e\in E} bv_{ke} xe_{ke} \le b + bI - bD \tag{5.36}$$

$$\sum_{k\in KP} bf_k x_k + \sum_{k\in\{K-KEQ\}} \sum_{e\in EI^k} (bv_{ke} - bv_{kec_k}) xe_{ke} \le bR + bI \tag{5.37}$$

$$\sum_{k\in KC} bf_k (1 - x_k) + \sum_{k\in\{K-KEQ\}} \sum_{e\in ED^k} (bv_{kec_k} - bv_{ke}) xe_{ke} \ge bD \tag{5.38}$$

$$x_k = 1 - xe_{k0} \quad \forall k \in \{K - KEQ\} \tag{5.39}$$

$$x_k = 1 \quad \forall k \in KEQ \tag{5.40}$$

$$\sum_{e\in E} xe_{ke} = 1 \quad \forall k \in \{K - KEQ\} \tag{5.41}$$

$$xe_{ke} \in \{0,1\} \quad \forall k \in \{K - KEQ\}, e \in E \tag{5.42}$$

$$x_k \in \{0,1\} \quad \forall k \in K \tag{5.43}$$

Here, the new terms are:

b: current yearly budget (€/year): $\sum_{k\in KC} (bf_k + bv_{kec_k})$

bI: increase in the yearly budget (€/year): bI $\in [0, \sum_{k\in K} (bf_k + \max_{e\in E} \{bv_{ke}\}) - b]$. Note: either bI>0 or bD>0 bD: decrease in the yearly budget (€/year): bD $\in [0, b]$. Note: either bI>0 or bD>0

bR: maximum re-invested yearly budget (€/year): bR $\in [0, b]$

E: {0, 0.5, 1, 1.5, …, $\overline{\eta}$ }: the set of possible numbers of FT employees (multiples of 0.5) occupying an RWC, indexed by *e*. The parameter $\overline{\eta}$ is equal to 3 in reality, so that |E|=7.

η_e : parameter which is equal to the number of FT employees which element *e* corresponds to

ec_k: element of E, corresponding to the *current* number of FT employees occupying location *k*

EI^k: {e| $\eta_e > \eta_{ec_k}$ }: the set of numbers of FT employees that are larger than the current number of FT employees occupying location *k*

ED^k: {e| $\eta_e < \eta_{ec_k}$ }: the set of numbers of FT employees that are smaller than the current number of FT employees occupying location *k*

xe_{ke} : 1 if the number of employees at location *k* is equal to η_e , 0 otherwise

bf_k: fixed yearly costs incurred when an RWC is established at location *k*. $bf_k = 0 \ \forall k \in KEQ$

bv_{ke}: variable yearly costs incurred when η_e employees occupy an RWC at location *k*. $bv_{ke} = 0 \ \forall k \in KEQ$

d_{ke} : estimated patient visits score at location *k* if η_e employees are employed there. $d_{ke} = 0 \ \forall k \in KEQ$

The objective function (5.34) maximises a function of the patient visits score and the continuity of care score of the resulting network. Again, the continuity of care score of the network is based on the continuity of care scores of all paths. These are determined in constraint (5.35). One of the approaches described in sections 5.2 and 5.3 can be used to include continuity of care score 1 or 2 in the RSIM.

Constraint (5.39) defines that an RWC can be regarded as established at an RWC location if and only if at least 0.5 FT employee occupies this location. By means of constraint (5.41) we stipulate that for a given RWC location *k*, xe_{ke} has the value 1 for exactly one $e \in$ E. This way, we choose whether to employ 0 employees (i.e. no RWC), in 0.5 FT employee, in 1FT employee, etc. Constraints (5.42) and (5.43) define xe_{ke}, and x_k as binary variables. Of course, RWC equivalents cannot be removed. Therefore, (5.40) ensures that x_k is equal to 1 for these locations.

Constraints (5.36) – (3.38) ensure that the budget constraints in case of an investment, a de-investment, or a re-investment (eventually combined with an investment or a de-investment) are met. Next, we explain why.

Of course, one cannot exceed the investment budget in case of an investment (i.e. bI>0, bD=0, bR=0), whereas one has to cut at least bD from the current budget in case of a de-investment (i.e. bI=0, bD>0, bR=0). For these cases it suffices to restrict that the costs of new RWCs and employees in the network do not exceed bI (5.37) and that the costs saved by removing current RWCs and employees are at least bD (5.38).

Suppose that one wants to re-invest at most bR (i.e. bI=0, bD=0, bR>0). Then (5.37) stipulates that the yearly costs of new RWCs and new employees cannot exceed bR, whereas (5.36) restricts that the costs of the network after the re-investment cannot exceed the current budget. This implies that at least the costs of these new RWCs and employees have to be cut somewhere else by removing RWCs and employees.

Next, consider the case that one wants to re-invest at most bR and to invest at most bI (i.e. bI>0, bD=0, bR>0). Then (5.36) restricts that the increase in the yearly costs of the network of RWCs is at most bI, whereas (5.37) restricts the yearly costs of new RWCs and new employees cannot exceed bR + bI. So, in case that these costs exceed bI, these constraints ensure that at least the difference between these costs and bI has to be cut somewhere else by removing current RWCs and employees.

Last, suppose that one wants to re-invest at most bR and to de-invest at most bD (i.e. bI=0, bD>0, bR>0). Then (5.37) stipulates that the yearly costs of new RWCs and new employees cannot exceed bR. Constraint (5.36) restricts that the decrease in the yearly costs of the network of RWCs is at least bD. Therefore, these constraints ensure that at least the costs of the additional RWCs and employees plus bD should cut somewhere else by removing RWCs and employees.

Summary
This subsection describes the RWC & Staff Investment Model (RSIM). The RSIM models the problem of optimising the investment of a budget increase in new RWCs and new employees. Furthermore, this model can be used to solve the problem of optimising the de-investment of a budget decrease by removing RWCs and employees, and to optimise re-investments in the network. The objective of these investments, de-investments and re-investments is to maximise the patient visits score and the continuity of care score of the resulting network.

5.4.3 Generalized RWC equivalents

In assumption A10.1 we state that the origin and destination of a flow can be regarded *as if an RWC is placed there*. The idea behind this assumption is that medical help is often available at these places.

In this subsection, we generalize this assumption in two ways. First, we assume that RWCs equivalents are situated at *a subset of* the origins and destinations of all flows (regard: the empty set is also a subset). Second, we take the possibility to refer truck drivers to other healthcare providers *along a path* into account. There are multiple other NGOs providing medical services along the road. In some cases, these health centres could be regarded as if they are RWCs. For example, North Star could make an agreement to cooperate with this centre if such centre offers treatment for HIV or TBC. So, next to a subset of the origins and destinations of all flows, also a set of other locations can be regarded as if an RWC is placed there. That is why we replace assumption A10.1 by the following assumption:

A10.2: RWC equivalents are situated at a set of locations

All changes which have to be made in order to adapt the RIM or RSIM to assumption A10.2 are related to the way the continuity of care score of a path is calculated. Note that both continuity of care score 1 and 2 are based on the driving times between the set of RWC neighbours *k* and *l*, i.e based on the driving time between the next RWC passed and the last RWC passed. By means of the example in figure 5.6, we explain why these driving times have to be calculated differently if A10.2 is assumed instead of A10.1.

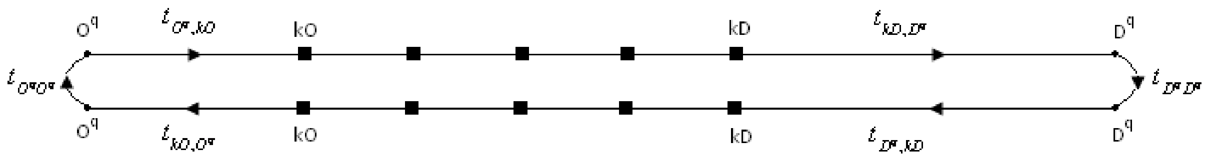


Figure 5.6: Time-line of a circular tour, e.g. O^q -> D^q -> O^q, explaining the driving time to the next RWC if no RWC is situated at O^q or D^q

In this example, no (dummy-) RWC is located at the destination of a path. Suppose that the RWC at location 'kD' is the last RWC which is passed before reaching the destination. A truck driver who just passed this RWC will reach the destination after t_{kD,D^q} time units, where he stays $t_{D^qD^q}$ time units. Afterwards, he starts off the trip back to the origin, passing the RWC at location 'kD' again after $t_{D^q,kD}$ time units. So, because A10.1 does not hold in our example (i.e. A10.2 holds), the driving time to the next RWC, after passing the RWC at location 'kD' for the first time is equal to $t_{kD,D^q} + t_{D^qD^q} + t_{D^q,kD}$. If assumption A10.1 would be made, this driving time is equal to t_{kD,D^q}. Similarly, the driving time to the next RWC, after passing the RWC at location 'kO' (the last RWC passed before reaching the origin O^q) for the first time, is equal to $t_{kO,O^q} + t_{O^qO^q} + t_{O^q,kO}$ time units. So, in order to identify the driving times between RWC neighbours along path q, it does not suffice to look only at the driving times between RWC neighbours *between O^q and D^q*. Instead, we look at the driving times between RWC neighbours for *a whole circular tour along path q* (e.g. from O^q to D^q, and back to O^q).

The following subsections explain how the integrated approach, the scenario approach and the FLRM approach to include continuity of care score 1 or 2 have to be adapted to the case that assumption A10.2 holds.

5.4.3.1 Changes in the *integrated* approach to include the continuity of care score *1*

Next, we explain what changes need to be made in the constraints added by the integrated approach to include the continuity of care score 1.

Constraint (5.7)

In order to calculate the continuity of care score (constraint (5.6)) of a path, we change the interpretation of the variables i_{klq} . Therefore we start with explaining the changed definition of these variables, and describe how to change constraint (5.6) afterwards.

Let us define the set of *locations* along path q, L^q, as $K^q \cup O^q \cup D^q$. I.e. the set of RWC locations united with the locations of the origin and the destination of path q. Figure 5.7 describes the definitions of all subsets of the set of locations L. Observe that the set L^q may also contain RWC locations at O^q and D^q, next to the locations of O^q and D^q themselves. Though the driving time between O^q and an RWC at O^q is equal to 0, we pretend that: a truck driver sequentially passes O^q, an eventual (dummy-) RWC at O^q, eventual (dummy-)RWCs along the path, an eventual (dummy-) RWC at D^q , and D^q.

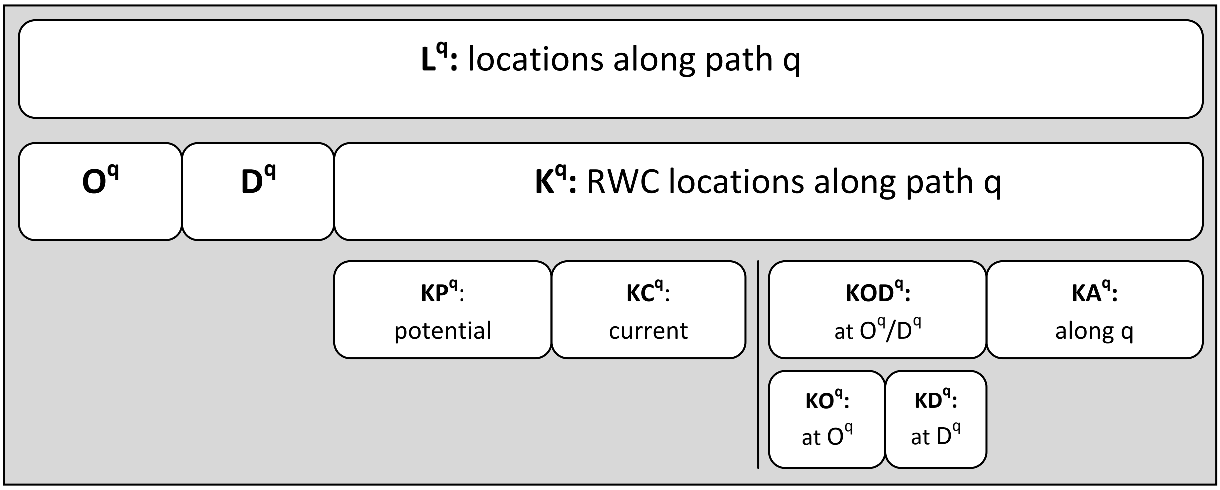


Figure 5.7: Overview of sub-sets of the set of locations L

According to definition 4.1, an RWC can only be a neighbour of another RWC. However, now we change the definition as follows:

Definition 5.1 Under assumption A10.2, a pair of locations is an RWC neighbour-pair at path q if and only if:

- Both locations contain a (dummy-) RWC, or are the origin or destination of a the path
- A truck driver travelling along path q will consecutively pass the two locations in such pair, without passing another RWC/origin/destination.

This means that an RWC can be a neighbour of another RWC or a neighbour of an origin/destination. The variables i_{klq} are redefined accordingly, by replacing constraints (5.7A) – (5.7E) by constraints (5.7A.1) – (5.7E.1).

$$\sum_{l \in L^q} i_{klq} = x_k \quad \forall q \in Q, k \in K^q \tag{5.7A.1}$$

$$\sum_{l \in L^q} i_{klq} = 1 \quad \forall q \in Q, k \in O^q \tag{5.7B.1}$$

$$\sum_{k \in L^q} i_{klq} = x_l \quad \forall q \in Q, l \in K^q \tag{5.7C.1}$$

$$\sum_{k \in L^q} i_{klq} = 1 \quad \forall q \in Q, l \in D^q \tag{5.7D.1}$$

$$i_{klq} \in \{0,1\} \quad \forall q \in Q, k \in L^q, l \in L^{qk} \tag{5.7E.1}$$

Where:

i_{klq} : 1 if locations *k* and *l* contain an RWC neighbour-pair at path *q*, 0 otherwise. These variables are only defined for set of pairs $\{(k,l\} \in L^q \times L^q \mid l \in L^{qk}\})$

L^q: $K^q \cup O^q \cup D^q$: the set of locations along path *q*

L^{qk}: the set of locations along path *q* that are passed *after* location *k*. Note: we pretend that $k \in KO^q$ is passed after O^q, and that $k \in KD^q$ is passed before D^q

O^q : the origin of flow *q*. $O^q \subseteq L$

D^q : the destination of flow *q*. $D^q \subseteq L$

One can prove that constraints (5.7A.1) – (5.7E.1) ensure that $i_{klq} = 1$ if and only if locations *k* and *l* contain an RWC neighbour-pair, in exactly the same way as described in theorem 5.1.

Constraint (5.6)

The parts $w_1 h_1(\overline{t_q})$ and $w_3 h_3(t_q^{\max})$, as well as the way these are linearized, remain the same. The second part of the objective function has to be adapted. This part compares the driving times between RWC neighbours along a path with the average of these driving times. As described in figure

5.6, the driving time to the next RWC is sometimes equal to the time to return to the RWC you just passed, via the origin or the destination. In order to take this possibility into account, we replace constraint (5.6) by the following:

$$g(\{x_k \mid k \in K^q\}) = w_1 h_1(\overline{t_q}) +$$
$$w_2\left(2\sum_{k\in K^q}\sum_{l\in K^{qk}} i_{klq} h_2(t_{kl} - \overline{t_q}) + \sum_{k\in O^q}\sum_{l\in KA^q} i_{klq} h_2(2t_{kl} + t_{kk} - \overline{t_q}) + \sum_{k\in KA^q}\sum_{l\in D^q} i_{klq} h_2(2t_{kl} + t_{ll} - \overline{t_q})\right) +$$
$$w_3 h_3(t_q^{\max}) \quad \forall q \in Q \tag{5.6.1}$$

The second part of this constraint consists of three summations. A truck driver making a circular tour at path q, will have to 'cover' the driving times between every *pair of RWCs* that are neighbours along this path twice. The deviation of these driving times from the average of the driving times is scored in the first summation. The second summation compares the time to return to an RWC via the origin of a path with the average of the driving times. It only does so if there is no RWC located between this RWC and the origin and if this RWC is not located at the origin. This is ensured by checking whether an RWC that is not located at the origin or destination (i.e. an element from the set KA) is the neighbour of the origin of a flow. The same is done for the time to return to an RWC via the destination of a flow.

The linearization of constraint 5.6.1 changes a bit with respect to the linearization of constraint 5.6. We describes these changes in appendix B,

Constraint (5.8)

Because of the changed definitions of n_q and T_q, we replace the constraints (5.8A) - (5.8C) by the constraints (5.8A.1) - (5.8C.1), and change the definitions of J^q, n_{jq}, T_q, and $\hat{T}_q$. This way we adapt constraints (5.8A) – (5.8G) to the generalization presented in this section:

$$\overline{t_q} = \sum_{j\in J^q} nT_{jq} / \theta_j + i_{O^q D^q q}\hat{t}_2 \tag{5.8A.1}$$

$$T_q = 2\sum_{k\in L^q}\sum_{l\in L^{qk}} i_{klq} t_{kl} + \sum_{k\in KOD^q} (1 - x_k) t_{kk} \quad \forall q \in Q \tag{5.8B.1}$$

$$\theta_j n_{jq} \le \sum_{k\in K^q} 2x_k + \sum_{k\in KOD^q} x_k \quad \forall j \in J^q, q \in Q \tag{5.8C.1}$$

Where:

J^q: $\{1,...,2|K^q|+2\}$: the set of all possible numbers of RWC neighbour-pairs (see definition 5.1) defined when making a whole circular tour along path *q*, indexed by *j*

n_{jq}: 1 if the numbers of neighbour-pairs defined for path *q* (i.e. RWCs passed when making a whole circular tour along this path) is equal to *j*, 0 otherwise

T_q: sum of the driving times between the n_q RWC neighbours defined for a circular tour along path *q*

$\hat{T}_q$: an upper bound on the value of T_q (e.g. $2t_{O^qD^q} + t_{D^qD^q} + t_{O^qO^q}$)

Consider a truck driver who makes whole circular tour from O^q to D^q, and back to O^q. If no RWCs are placed at O^q and D^q, i.e. the set KOD^q is empty, the number of times an RWC is passed during its tour is equal to $\sum_{k \in KA^q} 2x_k$. If an RWC is placed at O^q or at D^q, the amount of times an RWC is passed during its tour is equal to $\sum_{k \in KA^q} 2x_k + 1$. This number is equal to $\sum_{k \in KA^q} 2x_k + 2$ if RWCs are placed both at O^q and D^q. Because we make a circular tour, the number of RWCs passed during such circular tour is equal to the number of RWC neighbour-pairs, n_q . Constraint (5.8C.1) ensures that n_q takes these values, depending on the presence of RWCs at the origin and/or destination of a path.

Let T_q represent the sum of the driving times between the n_q RWC neighbours. In figure 5.5, this corresponds to the sum of the lines connecting the RWCs. We use the redefined variables i_{klq} to do this. If RWCs are placed both at O^q and D^q, this sum is equal to $2\sum_{k \in L^q} \sum_{l \in L^{qk}} i_{klq} t_{kl}$. If no RWC is placed at O^q and/or D^q, this sum is increased with $t_{O^qO^q}$ and $t_{D^qD^q}$, respectively (see figure 5.5). Constraint (5.8B.1) ensures that T_q takes the correct value.

Because of assumption A10.2, it is possible that *no single (dummy-) RWC* is passed when travelling along path *q*. Whether this is the case is identified by the variable $i_{O^qD^qq}$. Namely, if the origin of path *q* is the neighbour of the destination, this means that no RWC is located in between. For this case, we ensure that the resulting continuity of care score of that path is 0, by making the scores for all three sub-objectives (see assumption A11.1) equal to 0. For the part $h_1(\overline{t_q})$, this can be done by setting the value for $\overline{t_q}$ at $\hat{t}_2$. That is exactly what the term $i_{O^qD^qq}\hat{t}_2$ in constraint (5.8A.1) does.

Constraint (5.9)

The way the value of $t_q^{\max}$ is obtained also changes for our generalization. The following constraints should be added to constraint (5.9A):

$$t_q^{\max} \geq \sum_{k \in KA^q} i_{klq}(2t_{kl} + t_{ll}) \quad \forall q \in Q, l \in D^q \tag{5.9B}$$

$$t_q^{\max} \geq \sum_{l \in KA^q} i_{klq}(2t_{kl} + t_{kk}) \quad \forall q \in Q, q \in O^q \tag{5.9C}$$

$$t_q^{\max} \geq i_{O^q D^q q} \hat{t}_2 \quad \forall q \in Q \tag{5.9D}$$

Constraint (5.9B) defines that the maximum among the driving times between RWC neighbours along path *q* is not smaller than $t_{kD,D^q} + t_{D^q D^q} + t_{D^q,kD}$ if no (dummy-) RWC is situated at the destination of that flow. As described in figure 5.5, this is the driving time from the last RWC along the path 'kD' to the destination of the path, and back to 'kD'. Similarly, constraint (5.9C) defines that the maximum driving time is not smaller than $t_{kO,O^q} + t_{O^q O^q} + t_{O^q,kO}$ if no (dummy-) RWC is situated at the origin of the flow. Constraint (5.9D) ensures that $h_3(t_q^{\max})$ is equal to 0 if no single (dummy-) RWC is located along path q.

Summary

In order to adapt the calculation of $g(\{x_k \mid k \in K^q\})$ in the RIM or the RSIM to the RWC equivalents generalization, the following changes have to be made to constraints (5.6) – (5.9):

- Replace constraint (5.6) by constraint (5.6.1)
- Replace constraints (5.7A) – (5.7E) by constraints (5.7A.1) – (5.7E.1)
- Replace the constraints (5.8A) - (5.8C) by constraints (5.8A.1) - (5.8C.1)
- Add constraints (5.9B) - (5.9D)

5.4.3.2 Changes in the *scenario* approach to include the continuity of care score *1*

This subsection describes how to adapt the scenario approach to include the continuity of care score 1 in case that assumption A10.1 does not hold.

As explained in section 5.2.2, this approach adds constraints (5.7) – (5.9) and (5.10) – (5.15) to the RIM or the RSIM in order to include the calculation of the continuity of care score 1. Section 5.4.2.1 already describes how to adapt constraints (5.7) - (5.9) to our generalisation. Among the other constraints only (5.14) and (5.15) need to be changed. Constraint (5.14) has to be replaced with (5.14.1), and constraints (5.15.1) - (5.15.3) have to be added.

$$t_{klq}^{diff} \leq y_{sq}\tau_s^{diff} + (1-y_{sq})M_q \quad \forall q \in Q, k \in L^q, l \in L^{qk}, s \in S \tag{5.14.1}$$

$$t_{klq}^{diff} \geq \sum_{l \in KA^q} i_{klq} \,|\, 2t_{kl} + t_{kk} - \overline{t_q} \,| \quad \forall q \in Q, k \in O^q \tag{5.15.1}$$

$$t_{klq}^{diff} \geq \sum_{k \in KA^q} i_{klq} \,|\, 2t_{kl} + t_{ll} - \overline{t_q} \,| \quad \forall q \in Q, l \in D^q \tag{5.15.2}$$

$$t_{klq}^{diff} > i_{O^q D^q q} \max_{s \in S} \{\tau_s^{diff}\} \quad \forall q \in Q \tag{5.15.3}$$

Where:

M_q: an upper-bound on t_{klq}^{diff} (e.g. $t_{O^q D^q} + t_{O^q O^q} + t_{D^q D^q}$)

A truck driver making a circular tour at path *q* has to cover the driving times between every RWC neighbour-pair along this path. In case that a two RWCs are neighbours, constraints (5.15) defines the absolute difference between the driving time between these RWCs and the average of all driving times. Similarly, (5.15.1) and (5.15.2) define these absolute differences in case that the driving time to the next RWC is the time to return to an RWC via the origin or the destination, respectively. Based on the entire set of absolute differences, constraint (5.14.1) defines the set of scenarios path *q* can be assigned to. Because of assumption A10.2, it is possible that no single (dummy-) RWC is passed when travelling along path *q*. Constraint (5.15.3) ensures that this path cannot be assigned to any scenario, so that the continuity of care score of such path is equal to 0.

Constraints (5.15.1) and (5.15.2) can be linearized in exactly the same way constraint (5.15) is linearized. This linearization is explained in detail in appendix C.

Summary

In order to adapt the calculation of $g(\{x_k \mid k \in K^q\})$ in the RIM or the RSIM to the RWC equivalents generalization, the following changes have to be made to constraints (5.7A) – (5.7E), (5.8A) – (5.8G), (5.9A), and (5.10) – (5.15):

- Replace constraints (5.7A) – (5.7E) by constraints (5.7A.1) – (5.7E.1)
- Replace the constraints (5.8A) - (5.8C) by constraints (5.8A.1) - (5.8C.1)
- Add constraints (5.9B) - (5.9D)
- Replace constraint (5.14) by constraint (5.14.1)
- Add constraints (5.15.1) – (5.15.2) (see appendix C), and (5.15.3)

5.4.3.3 Changes in the *FLRM* approach to include the continuity of care score *1*

The only thing that has to be changed in order to adapt the FRLM approach to the RWC equivalent generalization is to define the set of driving times between RWC neighbours, given that combination of RWCs h is established, based on *a circular tour along path q* (see figure 5.5). Next, combination of RWCs h can be scored by means of constraint (5.6.1). This score is represented by the parameter c_{hq}. Given these parameters, the RIM and the RSIM can be solved, using constraints (5.17) – (5.20) to define the continuity of cares score of a path.

5.4.3.4 Changes in the *integrated* approach to include the continuity of care score *2*

In case that assumption A10.1 holds, constraints (5.7A)-(5.7E), (5.21A)-(5.21F) and (5.22) have to be added to the RIM or the RSIM in order to integrate the continuity of care score 2. To adapt this approach for the case that assumption A10.1 does not hold, constraints (5.7A)-(5.7E) have to be replaced by (5.7A.1)-(5.7E.1) and (5.22) has to be replaced by (5.22.1). The reasons why these have to be replaced are explained in section 5.4.3.1.

$$ER_q = \frac{1}{2t_{O^qD^q}}\left(2\sum_{k\in K^q}\sum_{l\in K^{qk}} i_{klq}t_{kl}^{\ 2} + \sum_{k\in O^q}\sum_{l\in KA^q} i_{klq}(2t_{kl}+t_{kk})^2 + \sum_{k\in KA^q}\sum_{l\in D^q} i_{klq}(2t_{kl}+t_{ll})^2\right) +$$

$$i_{O^qD^qq}\hat{t}_5 \quad \forall q\in Q \qquad (5.22.1)$$

In (5.22.1) the definition of the expected RWC time of path q is applied to a circular tour along that path (see figure 5.5). Because of assumption A10.2, it is possible that no single (dummy-) RWC is passed when travelling along path q. The term $i_{O^qD^qq}\hat{t}_5$ ensures that the continuity of care score of such path is equal to 0.

5.4.3.5 Changes in the *scenario* approach to include the continuity of care score *2*

As explained in section 5.3.2, the scenario approach adds constraints (5.7A) – (5.7E) and (5.22) – (5.26) to the RIM or the RSIM, to integrate the continuity of care score 2. In order to adapt this approach for the case that assumption A10.1 does not hold, constraints (5.7A)-(5.7E) have to be replaced by (5.7A.1)-(5.7E.1) and (5.22) has to be replaced by (5.22.1). Again, the reasons why these have to be replaced are explained in section 5.4.3.1.

5.4.3.6 Changes in the *FLRM* approach to include the continuity of care score *2*

Adapting the FLRM to the case that assumption A10.1 does not hold, is a matter of providing the right values for the parameters c_{hq}: the continuity of care score of path q in case that combination of RWCs h is established. To this end, we have to define the set of driving times between RWC neighbours, given that combination of RWCs h is established, based on *a circular tour along path q* (see figure 5.5). Next, the value of c_{hq} can be obtained by calculating $h_4(ER_q)$ for path *q*, where the value of ER_q can be calculated by means of constraint (5.22.1). Given these parameters, Given these parameters, the RIM and the RSIM can be solved, using constraints (5.17) – (5.20) to define the continuity of cares score of a path.

5.5 Summary

In section 5, the RWC Investment Model (RIM) and the RWC & Staff Investment Model (RSIM) are proposed. The RIM models the problem of optimising the investment, de-investment, or re-investment of a given number of RWCs. The RSIM models the problem of optimising the investment of a budget increase, the de-investment of a budget decrease, or the re-investment of part of the current budget in RWCs *and* employees. The objective of the RIM and the RSIM is to maximise the patient visits score and the continuity of care score of the network obtained after the investment, de-investment or re-investment.

Initially, the continuity of care score of a network is based on the continuity of care scores 1 of all paths. Three different approaches are proposed to include this score in the RIM and the RSIM, which all have their advantages and disadvantages in terms of complexity and accuracy. The integrated approach defines the variables which determine the continuity of care score by means of a set of linear constraints, and defines this score as a piece-wise linear function of these variables. The scenario approach does almost the same, but uses a piece-wise constant function between the continuity of care score and these variables. The FLRM approach 'tells' the RIM and the RSIM what the continuity of care score of a path would be in case that a given set of RWCs would be established along a path. This way, this score is not a variable any more, but a parameter.

Because some preliminary tests showed that all these approaches result in models which are very hard to solve, these approaches are adapted for the case that continuity of care score 2 is used. In terms of complexity and accuracy, the scenario approach to integrate this score shows most potential.

The continuity of care score 1 and 2 are determined based on the time required to drive from an RWC to the next RWC passed. The next RWC passed may be the same RWC again, which occurs when no (dummy-)RWC is located at the origin or the destination of a path. If this is the case, the continuity of care score of a path has to be calculated differently. Therefore, the way to calculate the continuity of care score 1 and 2 are generalised.

6. Performance Analysis

This section deals with the performance of the RWC Investment Model (RIM) and the RWC & Staff Investment Model (RSIM). Subsection 6.1 contains some general conclusions about the complexity of these models and the three approaches to include the continuity of care score of a path. By means of a case study, we analyse the models delivered to North Star. This is done in subsection 6.2. Subsection 6.3 summarises section 6.

6.1 Model analysis

When the continuity of care score of a path would not be included in the RIM and the RSIM, these would be relatively simple models. Both models are MIP models, and the numbers of variables and constraints are quite small. The RIM, as described in (5.27) – (5.33), consists of |K| binary decision-variables, 4+|K| +|KEQ| 'normal' constraints and |Q| constraints which have to be replaced by a set of *shadow-constraints* from one of the approaches described in section 5.2 or 5.3. For the RSIM, as described in (5.34) – (5.43), these numbers are: |K|+|K|*|E| binary decision-variables, 4+4|K|-2|KEQ| 'normal' constraints, and |Q| constraints which have to be replaced by a set of shadow constraints. The sets K, Q, and E, which determine the size of the model, tend to be quite small in reality. As we explain in section 6.2, North Star expects that the cardinalities of these sets do not exceed 100, 100, and 7, respectively. So, if the continuity of care score of a path would not be included into the RIM or the RSIM, solving these models would be relatively easy.

However, constraint (5.28) from the RIM and constraint (5.35) from the RSIM have to be replaced by |Q| sets of shadow constraints. The additional numbers of shadow constraints and variables differ a lot, depending on the definition of the continuity of care score used (continuity of care score 1 or 2) and on the approach used to implement this score (integrated, scenario or FLRM approach). Tables 6.1 and 6.2 summarise these numbers under assumption A10.2 (i.e. for the case of generalised RWC equivalents).

Though these numbers do not necessarily say something about the possibilities to solve a problem within an acceptable time, they often give some information about this. In general, solving problems with many constraints and (integer) variables are more difficult to solve than problems with few constraints and (integer) variables.

	Continuity of care score 1[3]		
Approach:	Integrated	Scenario	FLRM
#decision-variables	$\sum_{q\in Q} \lvert K^q\rvert^3 + r1$	$\frac{1}{2}\sum_{q\in Q}(\lvert K^q\rvert^2 + 4\lvert K^q\rvert + 4) + \lvert Q\rvert\lvert S\rvert$	$\sum_{q\in Q} \lvert H^q\rvert$
#binary	$\frac{1}{2}\sum_{q\in Q} \lvert K^q\rvert^2 + r2$	$\frac{1}{2}\sum_{q\in Q}(\lvert K^q\rvert^2 + 4\lvert K^q\rvert + 4) + \lvert Q\rvert\lvert S\rvert$	0
#constraints	$2\sum_{q\in Q} \lvert K^q\rvert^3 + r1$	$(3 + \frac{1}{2}\lvert S\rvert)\sum_{q\in Q} \lvert K^q\rvert^2 + 2\lvert Q\rvert\lvert S\rvert + r2$	$\sum_{q\in Q} \lvert H^q\rvert(\lvert K^q\rvert + 1) + 1$

Table 6.1: Additional number of decision-variables, decision-variables that are binary variables, and constraints when including continuity of care score 1 in the RIM or the RSIM

	Continuity of care score 2		
Approach:	Integrated	Scenario	FLRM
#decision-variables	$8\lvert Q\rvert + \frac{1}{2}*$ $\sum_{q\in Q}(\lvert K^q\rvert^2 + 2\lvert K^q\rvert + 2)$	$\lvert S\rvert\lvert Q\rvert + \frac{1}{2}*$ $\sum_{q\in Q}(\lvert K^q\rvert^2 + 2\lvert K^q\rvert + 2)$	$\sum_{q\in Q} \lvert H^q\rvert$
#binary	$4\lvert Q\rvert + \frac{1}{2}*$ $\sum_{q\in Q}(\lvert K^q\rvert^2 + 2\lvert K^q\rvert + 2)$	$\lvert S\rvert\lvert Q\rvert + \frac{1}{2}*$ $\sum_{q\in Q}(\lvert K^q\rvert^2 + 2\lvert K^q\rvert + 2)$	0
#constraints	$3 + 2\sum_{q\in Q} \lvert K^q\rvert + 11\lvert Q\rvert$	$1 + 2\sum_{q\in Q} \lvert K^q\rvert + (4 + \lvert S\rvert)\lvert Q\rvert$	$\sum_{q\in Q} \lvert H^q\rvert(\lvert K^q\rvert + 1) + 1$

Table 6.2: Additional number of decision-variables, decision-variables that are binary variables, and constraints when including continuity of care score 2 in the RIM or the RSIM

Looking at the numbers in table 6.1, the integrated approach to include the *continuity of care score 1* does not show much potential. For example, if $\lvert K^q\rvert$ is equal to 10 for some *q* (which is quite realistic), this path brings about more than 2000 variables and over 3000 constraints. The scenario approach shows a bit more potential, because this approach does not include the set of constraints needed to linearize constraint 5.6.1 (see appendices A and B). The FLRM approach has the advantage that it does not bring about any integer variables. However, the size of the set H^q, as well as the number of variables and constraints, grows exponentially in $\lvert K^q\rvert$. Suppose that $\lvert K^q\rvert$ RWC locations are passed

[3] r1 : a term with the form $a1\sum_{q\in Q} \lvert K^q\rvert^2 + a2\sum_{q\in Q} \lvert K^q\rvert + a3\lvert Q\rvert + a4$

r2 : a term with the form $a5\sum_{q\in Q} \lvert K^q\rvert + a6\lvert Q\rvert + a7$

along path q. Then the total number of combinations which have to be checked in order to identify whether such combination is 'relevant' (see section 5.2.3) is equal to $\sum_{q \in Q} 2^{|K^q|}$. The FLRM approach for including the continuity of care score 1 only seems to be promising in case that $|K^q|$ is small for all q.

In terms of numbers of (integer) variables and constraints, the *continuity of care score 2* is always at least as good as the continuity of care score 1. Only for the FLRM approach, these numbers do not depend on the definition of the continuity of care score. In case that the integrated approach is applied, the model size is decreased a lot when using continuity of care score 2. This decrease is somewhat smaller for the scenario approach. The benefits of using the scenario approach instead of the integrated approach are very small for this score. Depending on the number of scenarios defined, using the scenario approach may even result in a model which is larger than the model in case that the integrated approach is used.

6.2 Quantitative analysis: case study

The objective of the case study is to discover whether the RIM and the RSIM can be solved within an *acceptable amount of time*. Specifically, this should be the case for a problem instance which North Star regards as a large one. The size of the problem instance is an upper bound on the size of the instances they intend to solve in the future. Next to that, the case study is meant to discover whether the solutions are logical, and to find out how *sensitive* the optimal solution is for changes in the user-defined parameters and for *noise* in the data.

In order to perform the case study, we make use of the program POLARIS (Program for Optimising the Long-term Achievements of the RWC Investment Strategy), which we built for North Star to solve the RIM and the RSIM. A screenshot of POLARIS can be found in appendix D. This program is built by means of AIMMS 3.11, and makes use of the GUROBI 4.0 solver to solve the RIM and the RSIM. A Dell Latitude D630 with a 2.2GHz Intel Core 2 Duo processor and 1.99GB of RAM is used to run the tests.

Initially, the FLRM approach was used to include the continuity of care score 1 the RIM and the RSIM. However, it turned out that it took a very long time to determine the continuity of care scores of all $\sum_{q \in Q} 2^{|K^q|}$ combinations of RWCs *h* (see section 6.1). This makes this approach highly impractical. Also the integrated approach and the scenario to include continuity of care score 1 result in large models. Some preliminary tests showed that solving a very small problem instance with these approached is already very complex.

That is why we started focussing on the continuity of care score 2. As explained in section 6.1, the advantage that the scenario approach results in much smaller models than the integrated approach does not hold for the continuity of care score 2. Moreover, the scenario approach brings about the disadvantage that the piecewise linear function $h_4(ER_q)$ is replaced by a piecewise constant function. Because of that, a little noise in the driving times data could have a large effect on the continuity of care score of a path. Next to that, the continuity of care score may increase disproportionally much after a small improvement in the continuity of care score of a path when using this approach. That is why we decided to use the *integrated approach* to include *continuity of care score 2* in the RIM and the RSIM.

6.2.1 Case description

We test the RIM and the RSIM on a large realistic sample network in Southern and Eastern Africa. Part of the data required to describe his network is based on the current network. The remaining data are generated as realistically as possible.

The nodes in this network are 71 O-D nodes, 25 current RWCs and 75 potential RWC locations. North Star provided 28 of these potential locations. Strategic locations in Sub-Saharan Africa, like mines, ports, and large cities, are chosen to complete this set of potential locations. The arcs in the network are the roads connecting these 171 locations. The length of an arc, i.e. the smallest driving time between two locations, is determined by means of the OMR (ORTEC Map & Route) module, which determines these by means of a shortest-path algorithm. These times also used to determine K^q: the set of RWC locations passed along the shortest (i.e. minimum driving time) route from O^q to D^q.

Because information about traffic flows in Africa is very scarce, we had to generate a set of routes ourselves. This set consists of 100 routes from and to the main ports, cities, and mines in Southern and Eastern Africa. The size of route q, f_q, is generated randomly.

The parameters bf_k (fixed yearly costs at location k) and bv_{ke} (variable yearly costs at location *k*, in case that η_e employees occupy an RWC at that location), are based on costs data from the 7 RWCs in Kenya. For simplicity, we assume that the costs are equal for each location: bf_k = €15.000/year, bv_{ke}= η_e*€8.000/year. Next to that, we assume that the number of FT employees occupying the current RWC locations is 1, which is true for most of them.

Last, the parameters d_{ke} are generated based on the patient visits scores of the current RWCs. Similar to the current RWCs, the expected patient visits scores among the potential RWCs differ a lot in reality. In order to include this variation in our case, these scores are more or less randomly drawn. The minimum and the maximum of the patient visits scores of the current RWCs are taken as bounds on these randomly chosen patient visits scores. A convergent relation between the patient visits score, d_{ke}, and the number of employees at location *k*, η_e, is used. So, in terms of patient visits the marginal benefits on an additional employee decreases when the number of employees already hired at a certain location increases. This relation is a realistic one too. Namely, the size of the pool of potential patients on a certain day is different each day. This size does not change when adding an employee. Therefore, compared to the other employees, the new employee will more often face a day in which the capacity he brings about in terms of patient visits is not (fully) needed.

Based on the case described above, the RIM has 3771 binary variables and 601 fractional variables. The number of constraints in this model is 2290. For the RSIM these numbers are: 4471 binary variables, 601 fractional variables, and 2515 constraints.

There are multiple reasons why the case described above can be regarded as a *relatively large* one. First, North Star has some global information about the number of main long distance truck routes in Sub-Saharan Africa. This number turns out to be relatively small. Namely, the bad condition of the roads in Africa restricts the set of roads which these long-haul vehicles drive on. Because of this, the goods are often transported from or to a relatively small number of 'distribution points'. Together with the main harbours, airports, and mines, these make up the set of origins and destinations of the routes. Because the set of O-D nodes is relatively small, the size of the set Q is relatively small too.

Second, North Star intends to apply the model for parts of the network only. For example, when they receive a budget for investments in a certain country, they do not need to include potential RWC locations in other countries. Next to that, they do only need to take the flows into account which (partly) go through that country or region. So, this restricts the size of the sets K and Q (and their subsets).

Last, as explained, the number of employees occupying an RWC can be at most 3 because of capacity constraints. Since North Star only hires multiples of 0.5FT employees, the size of the set of possible numbers of employees occupying an RWC at location *k* is *at most* 6, i.e. $|E|=7$.

Based on these observations, North Star expects that $|Q|$ is smaller than 100, that $|K|$ is smaller than 100, and $|E|$ is at most 7 in the problem instances they are going to run in the future. This supports the statement that the case described above is a relatively large one. If the RIM and the RSIM can be solved within an acceptable time for our case study, it is likely that most of North Star's future problem instances can be solved within an acceptable time too.

6.2.2 Results

This subsection describes a solution of the RIM and one of the RSIM, as well as why these solutions make sense. Both models perform the investments and/or de-investments on the network described in subsection 6.2.1. For the parameters *r*, $\hat{t}_4$, and $\hat{t}_5$ we use the values which North Star regards as default values: 0.5, 4 (hours), and 16 (hours), respectively.

Figure 6.1 describes the optimal network after moving 5 RWCs to a different place in the network. This is the network obtained after running the RIM with p=5 and pD=5. Because of this re-investment, the patient visits score of the network increases from 732 to 831, whereas the continuity of care score of the network increases from 704 to 1959. As expected, RWCs are removed from locations with a relatively small patient visits score (10, 17, 13, 17, and 20) and added at locations with a relatively large patient visits score (20, 53, 40, 33, and 30). Next to that, RWCs are removed from locations that hardly contribute to the continuity of care of the network, and added at locations that contribute a lot. When closing one of the five removed RWCs in the initial network, the continuity of care score of that network decreases with 3, 0, 0, 0, and 0, respectively. In contrast, when closing one of the added RWCs in the optimised network, the continuity of care score of that network decreases with 227, 411, 220, 469, and 380, respectively.

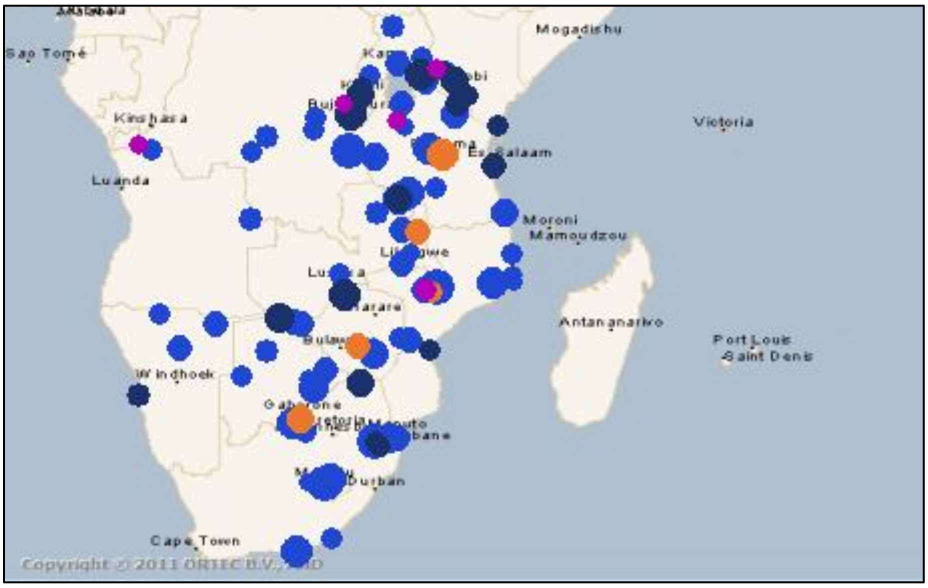


Figure 6.1: optimal network after moving 5 RWCs to another location. Orange: added RWC, purple: removed RWC, dark blue: current RWC, light blue: potential RWC. The size of a node reflects the patient visits score

The time required to solve this problem is 648 seconds. Determining the parameters $\bar{d}$ and $\bar{c}$ took 1 second and 235 seconds, respectively. The remaining 412 seconds were needed to perform the final optimisation.

Figures 6.2 and 6.3 describe the optimal network after increasing the yearly budget with €150.000 (the current yearly budget is €575.000). At most 7 new RWCs with at least 0.5FT employee can be established with this money. In terms of continuity of care, it is optimal to place as many new RWCs as possible. In terms of patient visits score, it is optimal to invest in 2 new RWCs only. Because investing in additional employees is a relatively cheap way of increasing the patient visits score, it is optimal to spend the remainder of the budget on this. Since *r* has the value 0.5, the optimal way to invest is a mix of these extremes. In total, 6 new RWCs are established and 7.5 FT employees are hired.

Because of this investment, the patient visits score of the network increases from 732 to 1059, whereas the continuity of care score of the network increases from 704 to 2168. Again, RWCs are added at locations that contribute a lot to the continuity of care in the network. If one of the six added RWCs would be closed, the continuity of care score of the network decreases with 209, 411, 220, 491, 150, and 409.

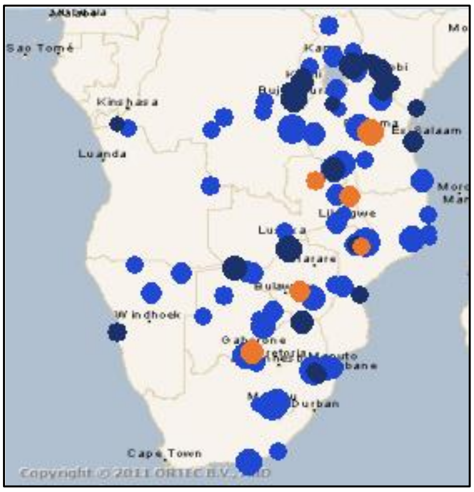


Figure 6.2: optimal network of RWCs after increasing the yearly budget with €150.000

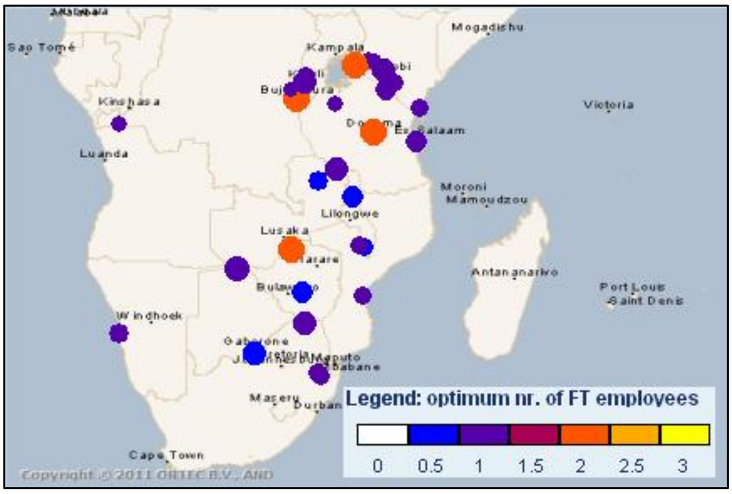


Figure 6.3 optimal numbers of FT employees after increasing the yearly with €150.000

The time required to run the RSIM to solve this problem is 684 seconds. Determining the parameters $\bar{d}$ and $\bar{c}$ took 1 second and 243 seconds, respectively. The remaining 440 seconds were needed to perform the final optimisation.

6.2.3 Effects of *r*, *p* and *bI*

The solutions of the RIM and the RSIM depend largely on the importance of the conflicting objectives of maximising the patient visits score and maximising the continuity of care score, represented by the parameter *r*. As explained in section 5.1, this parameter can be interpreted as follows. Suppose that the maximum attainable patient visits score after the investment is $\bar{d}$, and that the maximum attainable continuity of care score after the investment is is $\bar{c}$. Then, getting the patient visits score 1 percent-point closer to $\bar{d}$ is (1-r)/r times as important as getting the continuity of care score 1 percent-point closer to $\bar{c}$. This subsection describes how sensitive the solutions of the RIM and the RSIM are for changes in r and to what extend these effects depend on the size of the investment. These are very relevant issues. A large sensitivity implies that decision makers should not simply fix the parameters, but analyse what happens to the optimal network and its fitness value if they change one of these parameters.

Figures 6.4, 6.5 describe the patient visits score and the continuity of care score of the network obtained after optimising the investment in a certain number of RWCs. Next, figures 6.6 and 6.7 describe these scores of the networks obtained after increasing the yearly budget with a certain amount. In order to see the effect of the parameter *r* on these scores, these investments are done for *r*=0, 0.2, 0.5, 0.8, and 1. The exact scores can be found in appendix E. Note, the trend-line is only added to get a better overview over the relation between the variables. This may be misleading. *Only the dots in the graphs refer to solved problem instances.*

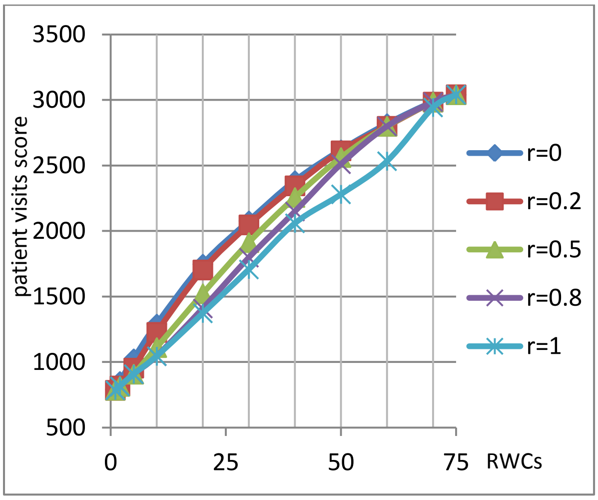


Figure 6.4: trade-off curve between patient visits score of a network and the number of RWCs added to that network.

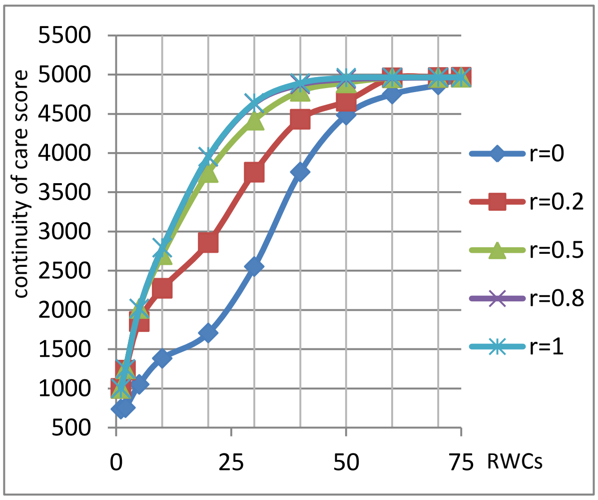


Figure 6.5: trade-off curve between continuity of care score of a network and number of RWCs added to that network.

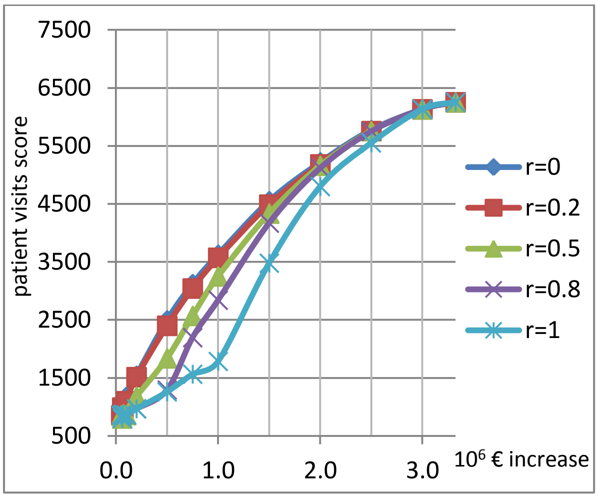


Figure 6.6: trade-off curve between patient visits score of a network and the budget increase in that network.

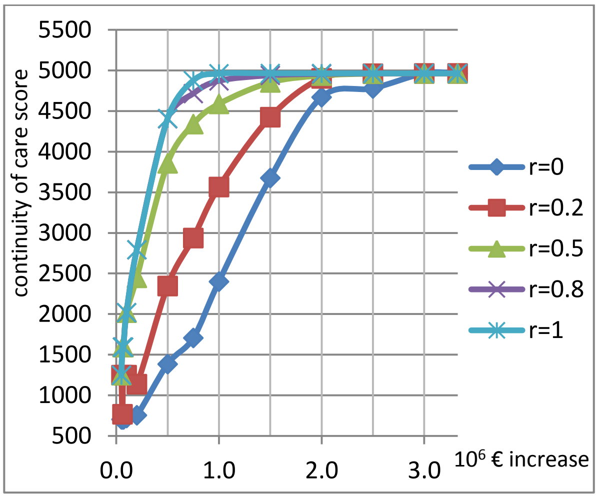


Figure 6.7: trade-off curve between continuity of care score of a network and the budget increase in that network.

A number of important conclusions can be drawn from these solutions. First, the following relation counts for a given investment size. If the value of *r* increases, the patient visits score decreases (or remains the same) and the continuity of care score increases (or remains the same). The optimal network is very sensitive for changes in *r*. A different value for *r* generally results in a quite different network. Because not all aspects the fitness of an investment can be included in the model, it might be beneficial to analyse the optimal networks obtained for multiple values of *r*.

The investment strategy which North Star used till now could be seen as a strategy with *r* = 0. Our case shows that this approach may perform badly in terms of continuity of care. For instance, when investing in 5 additional RWCs, the continuity of care score increases with 348 (50% increase) in case that *r* = 0 is used, whereas this increase is 1315 (187% increase) when *r*=0.5 is used. The nice thing about this change in *r* is that the effects on the patient visits score of the resulting network are relatively small. The patient visits score increases with 320 (44% increase) in case that *r* = 0 is used, whereas this increase is 176 (24% increase) when *r*=0.5 is used. So, the second conclusion is that the current investments strategy can be improved a lot in terms of continuity of care, and that changing the strategy (i.e. changing *r*) may have relatively small effects on the increase in the patient visits score after an investment.

Third, with the set of potential RWC locations in our test network, it is not possible to provide each route with the maximum possible continuity of care. In case that RWCs would be established at all 75 potential RWC locations, the continuity of care score of the network is 4964. If all routes would be provided with a complete continuum of care, this score would be 8273. This highlights the fact that optimising an investment is also a matter of providing the RIM or the RSIM with *strategically located potential RWC locations*.

Fourth, for *r*=1, the trade-off curve between the continuity of care score and the size of the investment is monotonically non-decreasing, but not convex (e.g. see figure 6.8). The concavities in this function can be explained by the fact that sometimes a couple of new RWCs along a path are required to cause a significant increase in the continuity of care score of that path. For values of *r* that are smaller than 1, such concavity is 'filled up' by making investments which improve the patient visits score of the network. This explains why the trade-off function of the patient visits score vs. the size of the investment is rather 'impulsive' for *r*>0 and why the function of the continuity of care score vs. the size of the investment is rather 'impulsive' for *r*<1 (e.g. see figure 6.9). Therefore, it might be beneficial not to fix the investment budget, but to see what networks are obtained when decreasing or increasing this budget a little. Maybe, a little increase may fill up a gap in the continuum of care, so that it results in a much better network. Or maybe a large decrease results in a network which fitness is almost equal.

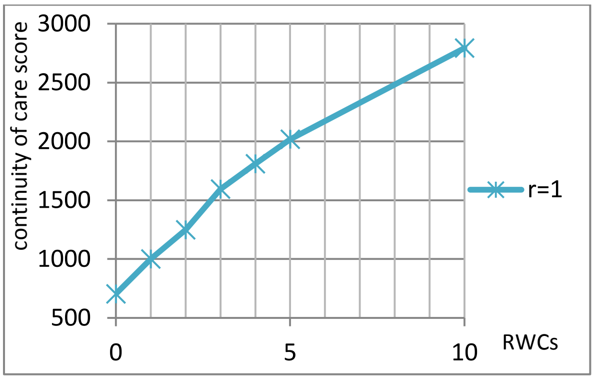


Figure 6.8: example of a concavity in the trade-off curve between continuity of care score of a network and the number of RWCs added, in case that *r*=1.

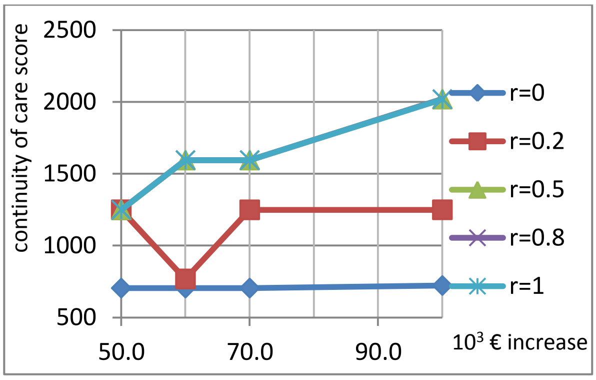


Figure 6.9: example of 'impulsive' behaviour in the trade-off curve between the continuity of care score of a network and the budget increase in that network.

Last, even though some parts of the curves are concave, figures 6.4 – 6.7 show that the marginal benefits of additional investments steadily decrease. In the long term North Star has to make a decision to stop investing in additional RWCs and employees. Namely, at some point the marginal benefits of an investment in terms of patient visits and continuity of care do not exceed the marginal costs of the investment any more. This stresses the importance of thinking about the question: what should be the minimal benefits in terms of patient visits and continuity of care obtained after investing a certain budget in the network of RWCs? In fact, one has to express the value of an additional continuity of care score unit and an additional patient visits score unit in Euros. This way, both the benefits and the costs of an investment can be measured with the same unit.

6.2.4 Effects of noise in f_q, d_k and d_{ke}

Data about the flow volumes are very scarce. Currently, these have to be estimated based on a small collection of traffic counts, which are often quite outdated. Probably, this results in noise in the parameters f_q, the flow volume at path q. Furthermore, almost all values of the parameters d_k (for the RIM) and d_{ke} (for the RSIM) have to be estimated. Because predictions are almost always wrong, these parameters tend to be noisy too.

Noise in these data may cause that the optimal solution would not be optimal in case that the correct estimates were used. Next, we explain how sensitive the optimal solution is for noisy data.

Theorem 6.1 *If the noise in the parameters f_q , d_k, and d_{ke} is at most pNoise %, then the optimal solution based on the noisy parameters is at most* $dZ = 100 * \frac{2 * pNoise}{100 + pNoise}$ *% worse than the optimal solution based on the correct parameter values.*

Proof: Consider the values of the optimal solution and the second best solution, based on the noisy parameters. These are represented by $Z1$ and $Z2$, respectively. Let us denote the set of locations *k* for which x_k=1 in the best solution and the second best solution by K^{best1} and K^{best2}. Furthermore, let us represent the set of flows *q* for which c_q>0 in the best solution and the second best solution by Q^{best1} and Q^{best1}, respectively.

Observe that the objective function of the RIM is linear in f_q and d_k, whereas the objective function of the RSIM is linear in f_q and d_{ke}. Let $\hat{Z}1$ and $\hat{Z}2$ be the values of $Z1$ and $Z2$ when the correct parameter-values were used. If the parameters d_k, d_{ke}, and f_q for *all* $k \in K^{best1}$ and all $q \in Q^{best1}$ are *pNoise* % smaller in reality, the value $\hat{Z}1$ is *pNoise* % smaller than $Z1$ *too*. Similarly, if the parameters d_k, d_{ke}, and f_q for *all* $k \in K^{best2}$ and all $q \in Q^{best2}$ are *pNoise* % larger in reality, the value $\hat{Z}2$ is *pNoise* % larger too.

So, the value of $Z1$ could be at most *pNoise* % smaller in reality, whereas the value of $Z2$ could be at most *pNoise* % larger in reality. Of course a parameter cannot increase and decrease with *pNoise* % at the same time. Therefore, a condition for $Z1$ to decrease with this percentage *and* $Z2$ to increase with this percentage is that $K^{best1} \cap K^{best2} = \emptyset$ and that $Q^{best1} \cap Q^{best2} = \emptyset$.

If these conditions are met, the value of $\hat{Z}1$ is at most $\frac{Z1 * (100 - pNoise)}{100}$ and the value of $\hat{Z}2$ is at most $\frac{Z2 * (100 + pNoise)}{100}$. This means that the upper bound on dZ, the difference between $\hat{Z}2$ and $\hat{Z}1$ (in % of $\hat{Z}2$), can be calculated as follows:

$$dZ = 100 * \frac{\hat{Z}2 - \hat{Z}1}{\hat{Z}2} \tag{6.1}$$

$$\leq 100 * \frac{Z2*(100 + pNoise) - Z1*(100 - pNoise)}{Z2*(100 + pNoise)} \tag{6.2}$$

$$\leq 100 * \frac{Z1*(100 + pNoise) - Z1*(100 - pNoise)}{Z1*(100 + pNoise)} \tag{6.3}$$

$$= 100 * \frac{2*pNoise}{100 + pNoise} \tag{6.4}$$

In (6.2) the upper-bounds on $\hat{Z}2$ and $\hat{Z}1$ are filled in. Obviously, dZ is maximised when $Z2 = Z1$, because $Z2$ cannot be larger than $Z1$ ($Z1$ is the value of the optimal solution). This explains formula (6.3). Last, formula (6.4) reduces formula (6.3), and completes this prove. ☐

Observe that this expression for dZ is an upper-bound. First, dZ can only be equal to this upper-bound if *all* parameters f_q , d_k, and d_{ke} change with *exactly pNoise* %. Obviously, this is not likely to occur. Second, the objective values of the best solutions are generally not equal (see formula 6.2 for the effects). Third, the sets K^{best1} and K^{best2} (and K^{best3}, K^{best4},...) do overlap a lot in reality. Namely, the current RWC locations are always in these sets (except for the case that all these are removed by a de-investment or a re-investment). Moreover, it is very likely that potential RWC locations that have a relatively large patient visits score and contribute a lot to the continuity of care in the network all show up in the best solutions in case of an investment. In case of a de-investment or a re-investment, this is the case for current RWC locations that have a relatively large patient visits score. The more the sets K^{best1} and K^{best2} (and K^{best3}, K^{best4},...) overlap, the smaller the upper-bound on dZ. Namely, noise in the patient visits scores of the RWCs that are in the best solutions does not change the absolute difference in the objective values of these solutions at all.

Last, the sets Q^{best1} and Q^{best2} (and Q^{best3}, Q^{best4},...) do overlap a lot in reality. Because of the reasons explained above, the sets K^{best1} and K^{best2} (and K^{best3}, K^{best4},...) tend to overlap a lot. The locations that are in all these sets generally contribute to the continuity of care of some routes, so that these routes show up in Q^{best1} and Q^{best2} (and Q^{best3}, Q^{best4},...) too. Again, it holds that the value of upper-bound on dZ gets smaller when the sets overlap more.

6.2.5 Computation times

Tables 6.3 and 6.4 show the computation times of solving the RIM and the RSIM for the problem instances described in section 6.2.3.

Budget Increase	r=0	r=0.2	r=0.5	r=0.8	r=1
€ 50,000	2	24	81	48	113
€ 60,000	2	25	70	60	39
€ 70,000	2	48	60	94	74
€ 100,000	2	33	332	148	254
€ 200,000	2	40	399	394	351
€ 500,000	2	131	821	103	90
€ 750,000	2	576	268	450	35
€ 1,000,000	60	385	892	950	24
€ 1,500,000	2	144	75	181	901
€ 2,000,000	2	456	22	805	41
€ 2,500,000	1	13	99	33	77
€ 3,000,000	2	10	15	850	25
€ 3,325,000	1	4	15	19	17

Table 6.3: computation times (seconds) of solving the RSIM with different parameters.

p	r=0	r=0.2	r=0.5	r=0.8	r=1
1	1	14	10	12	11
2	1	45	63	72	44
5	3	88	115	191	169
10	3	119	921	421	773
20	2	61	101	126	439
30	2	32	19	49	41
40	2	11	12	12	19
50	2	5	12	9	17
60	3	4	11	9	5
70	1	3	4	5	7
75	2	3	3	7	10

Table 6.4: computation times (seconds) of solving the RIM with different parameters.

All problem instances were solved within 950 seconds. Though there is no one-to-one relation between the computation time and the values of *r*, *p*, or *bI*, these parameters seem to have some influence on the computation time. First, problem instances with a small value for *r* seem to be solved quicker. A possible explanation can be found in the fact that the problem of searching for the investment that optimises the patient visits score can be solved relatively easy (this problem is a knapsack problem). In contrast, optimising an investment in terms of continuity of care is a complex combinatorial problem. Second, optimising the way to invest a very small budget increase (i.e. close to 0) or a very large budget increase (i.e. close to the maximum possible investment: € 3,325,000) seems to be easier than a 'moderate' budget increase. The number of ways to invest *small* budget increase is relatively small, which might explain the small computation times for these cases. One could also regard a choice where to invest in additional RWCs and employees as a choice *where not to invest* in these. In case of a *large* budget increase the set of options where not to invest is relatively small, which could explain the small computation times for these cases.

6.3 Summary

In terms of the numbers of (integer) variables and constraints, including the continuity of care score 1 in the RIM or the RSIM results in large models. Some initial tests indicate that these models are indeed hard to solve, even for small test cases. Therefore we chose to use the continuity of care score 2 in these models. Specifically, the integrated approach to include this score in the RIM and the RSIM was used, because the FLRM approach still results in large models, and because the scenario approach simplifies the definition of this score.

The solutions provided by the RIM and the RSIM are logical. Investments are made at locations that contribute a lot to the continuity of care in the network and have a large patient visits score. Moreover, the investments balance the conflicting objectives of maximising the patient visits score and maximising the continuity of care score.

Currently, North Star mainly focuses on maximising the patient visits score. Test cases show that a little more emphasis on the maximisation of the continuity of care score results in a large increase in this score and a relatively small decrease in the patient visits score. Therefore, it is important not to fix the importance of both objectives, but to consider the effects of a small change.

The test cases also show that it is important not to fix the investment budget. A little increase could make it possible to provide a flow with a significant continuity of care score, which would not be possible with the original budget.

Furthermore, the results show that the marginal benefits of additional investments steadily decrease. This highlights the importance of defining the required benefits of an invested euro in terms of patient visits and continuity of care.

The presence of noise in the estimates of the patient visits scores and the flow volumes is a matter-of-fact. This may imply that the optimal solution of the RIM or the RSIM would not be optimal in case that the correct estimates were used. An upper-bound on the sensitivity of the optimal solution to this noise is derived.

Last, the RIM and the RSIM are tested on a large sample network. Because all problem instances were solved within an acceptable time, we expect that North Star will not face large computation times when solving future problem instances either.

7. Conclusions and Discussion

7.1 Conclusions

As explained in the introduction, the main objective of this thesis is to describe the possibilities to model and to solve the problem of how to invest optimally in the network of RWCs. In short, the following conclusions can be drawn.

There are two types of investments which can be made: establishing new RWCs and hiring new employees. The fitness of such investment depends on the costs of the investment, the characteristics of the locations where the investments are made, the resulting change in the patient visits score of the network, and the resulting change in the continuity of care score of the network. Assumptions are made to simplify the definition of this fitness value of an investment. This value is a function of the resulting change in the patient visits score and the resulting change in the continuity of care score of the network.

This thesis proposes two models to solve the investment optimisation problem. The RWC Investment Model (RIM) models the problem of locating *p* new RWCs in such a way that the fitness value of the resulting network is maximised. In order to include investments in additional employees, the RIM is extended to the RWC & Staff Investment Model (RSIM). The RSIM models the problem of how to invest a budget increase by establishing new RWCs and hiring new employees in such a way that the fitness value of the resulting network is maximised. Both the RIM and the RSIM are extended so that de-investments can be optimised too.

The continuity of care score of a network is defined as a weighed sum of the continuity of care scores of all paths in the network. This score is large if truck drivers, who suddenly need to visit an RWC while travelling along their path, do not need to drive a long time before passing an RWC. Instead, this score is small if truck drivers have to drive a long time before passing an RWC. Including such score into the RIM and the RSIM brings about computational problems. Therefore, this thesis proposes two different definitions of this score, which both define this score based on the extent to what a truck driver is always close to an RWC. Continuity of care score 1 does so by looking at the maximum, the average, and the variation among the driving times between the RWCs passed when travelling along a path. The continuity of care score 2 depends on the expected driving time to the next RWC when travelling along a path. Moreover, we describe three approaches (integrated, scenario and FLRM approach) to include these scores into the RIM and the RSIM.

Some initial tests indicate that it is hard to solve the models obtained when including the continuity of care score 1. Therefore we chose to focus on continuity of care score 2. Specifically, we regard the integrated approach to include this score in the RIM and the RSIM as the best option, because the FLRM approach often results in large models which are hard to solve, and because the scenario approach simplifies the definition of this score.

Test problems show that the RIM and the RSIM, using the integrated approach to include continuity of care score 2, provide logical answers and can be solved within an acceptable time. All 120 problem instances are solved within 950 seconds, whereas the average computation time is 129 seconds.

7.2 Contribution to literature

As described in the literature review, our problem can be classified into the network location problems. Specifically, it belongs to the subclass of flow interception problems, because locations of new facilities are optimised with respect to moving demand units: the truck drivers. Since a flow is not simply provided with a continuum of care when one RWC is placed along the corresponding route, our problem can be further classified into the subclass of multi-coverage problems.

There are only two known problems within this field. The first is the billboard location problem. Because it is beneficial to place multiple billboards along the route of a flow, the degree of coverage is defined as a non-decreasing function of the number of billboards along the route. Our problem is quite different, because this degree of coverage is not only determined by the number of RWCs along a route, but also by the driving times between these RWCs. The second problem in this field is the flow refuelling location problem. Here, a flow is regarded as being covered if a vehicle, driving along the corresponding path, will always meet a refuelling station before it runs out of fuel. Though this problem does take the driving times between the facilities into account, our problem is quite different again. The flow refuelling location problem regards a flow as being covered or not. In contrast, flows are covered (i.e. provided with a continuum of care) to some extent in our problem. The RIM and the RSIM are the only multi-coverage flow covering models known in literature that define the coverage of a flow as a fractional variable *and* determine this degree of coverage based on the driving times between the facilities located along a flow.

Two other innovations in the models proposed in this thesis are the so-called integrated approach and the scenario approach to include the degree of coverage (i.e. the continuity of care score) of a path in the RIM and the RSIM. These approaches use a set of linear constraints to determine the degree of coverage of a path as a function of the driving times between the established RWCs along that path. In the other multi-coverage flow covering models known, the degree of coverage is a matter of providing the models with a set of parameters.

7.3 Discussion and future research directions

In terms of usefulness, the RIM and the RSIM have their strengths and weaknesses. This subsection describes these, as well as the corresponding opportunities for future research

Weaknesses

One of the main disadvantages of the models proposed in this thesis is that these require O-D flow data. Particularly for Africa, these data are very scarce and may be incorrect. Furthermore, these models require estimates of the patient visits scores, which tend to be noisy too. This may imply that the optimal solution of these models may not be optimal when the correct data are provided.

An intrinsic property of a model is that it simplifies reality, which is also the case for the RIM and the RSIM. First, the competition effects and the market expansion effects of placing multiple RWCs along a path are not included. Further research is required to see whether these effects can be included in the models. As explained in sections 3.2 and 3.3, there has been written a lot about modelling these effects, for example in (Kim, 2010). Second, the RIM and the RSIM do not model the relation between the possibilities for funding an investment and the region where the investment is made. Third, the RSIM regards all employees as equal, while the costs (salary) and benefits (increase in the patient visits score) differ between nurses and outreach workers. Relatively small adaptations in this model need to be made to overcome this shortcoming.

Furthermore, the RIM and the RSIM implicitly assume that truck drivers do not deviate from their planned paths in order to get medical help. This is not always true. Sometimes, truck drivers are willing to deviate a little, for example to visit a renowned hospital in a city they pass. However, this willingness to deviate decreases with the distance they have to deviate. In fact, this means that medical centres contribute to the continuum of care ensured to a route, depending on the distance to that route. A further investigation is required to discover the possibilities to change the continuity of care score of a path accordingly.

Future Research

The fitness of the possible investments partly depends on the set of locations where can be invested in new RWCs. Namely, these determine to what extent the gaps in the continuum of care of a path can be filled up, and how many RWCs are required to do this. The fitness of future investments might be improved a lot if strategically located potential RWC locations are selected. This highlights the importance of investigating what the properties of promising potential RWC locations are.

The Flow Refuelling Location Model (FLRM), described in (Kuby et al., 2005), is a special case of the RIM. The FLRM faces the disadvantage that the number of possible combinations of facilities, i.e. the number of variables, grows exponentially in the number of facilities along a route. This computational disadvantage does not show up in case that the integrated approach or the scenario approach is used to determine whether a flow can be 'refuelled' or not. For this reason, it would be interesting to investigate whether using the RIM makes solving the FLRM easier.

The RIM and the RSIM can be used to optimise investment plans. However, making these plans brings about another optimisation problem: what is the optimal order in which the investments in such plan should be executed? Further research is required to answer this question.

Strengths

One of the strengths of the models proposed in this thesis is the fact that the RIM and the RSIM provide the optimal investments, de-investments, and re-investments, *given* the preferences of the decision maker. Does the decision maker want to choose those investments that maximise the patient visits score of the resulting network, does he want only to maximise the continuity of care score, or does he want it to make a trade-off between these two objectives? These questions can be answered by setting the value of one parameter.

The RIM and the RSIM provide decision makers with the opportunity to plan ahead all future investments within a certain region. That is, to optimise the network obtained after making all future investments in that region. Currently, one does not take the future into account when deciding how to invest a given budget increase. This can be identified as a greedy approach. Because solving the RIM or the RSIM gives the optimal way to invest all future investments, the solutions provided by these models are always at least as good as the solutions provided by a greedy approach.

Therefore, the current investment strategy is like a person who is building a house without a plan. Every time he gets some money, he makes a little plan for how to spend that money on the house. In contrast, the RIM and the RSIM are like architects. They make a well-thought construction plan for a new house, which can be executed afterwards.

REFERENCES

Averbakh, I., & Berman, O. (1996). Locating flow-capturing units on a network with multi-counting and diminishing returns to scale. *European Journal of Operational Research 91* , 495-506.

Bapna, R., Thakur, L.S., & Nair, S.K. (2002). Infrastructure development for conversion to environmentally friendly fuel. *European Journal of Operational Research 142* , 480–496.

Berman, O., Bertsimas, D., & Larson, R.C. (1995). locating discretionary service Facilities, II: Maximising market size, minimizing inconvenience. *Operations Research* , 623-632.

Berman, O. (1997). Deterministic Flow-Demand Location Problems. *The Journal of the Operational Research Society 48* , 75-81.

Berman, O., &Krass, D. (1998). Flow intercepting spatial interaction model: a new approach to optimal location of competitive facilities. *Location Science 6* , 41-65.

Berman, O., &Krass, D. (2002a). Locating Multiple Competitive Facilities: Spatial Interaction Models with Variable Expenditures. *Annals of Operations Research 111* , 197–225.

Berman, O., &Krass, D. (2002b). The generalized maximal covering location problem. *Computers & Operations Research 29* , 563–581.

Caraël, E. (2005). *HIV and Mobile Workers: a review of risks and programmes among truckers in West Africa.* Geneva: International Organisation for Migration.

Dandan, H. (2009). Model and algorithm of flow interception problem with multi-type of flows. *Journal of Systems Engineering and Electronics 20* , 755–761.

Dodds, W. B., Monroe, K. B., & Grewal, D. (1991). Effects of Price, Brand, and Store Information on Buyers' Product Evaluations. *Journal of Marketing Research* , 307-319.

Drezner, T., Drezner, Z., & Goldstein, Z. (2010). A Stochastic Gradual Cover Location Problem. *Naval Research Logistics 57* , 367-372.

Ferguson, A., & Morris, C. (2006). Mapping transactional sex on the Northern Corridor highway in Kenya. *Health & Place* , 504-519.

FHI. (1999). *Corridors of Hope in Southern Africa: HIV Prevention Needs and Opportunities in Four Border Towns.*

Gatignon, A., & Van Wassenhove, L. (2008). *Paving the Road to Healthy Highways: a Partnership to Scale Up HIV&AIDS clinics in Africa.* Fontainebleau: INSEAD.

Hale, T. (2006). Accessed at May 20, 2011, from Trevor Hale’s location science references: http://gator.dt.uh.edu/~halet/

Hodgson, M. J., & Berman, O. (1997). A billboard location model. *Geographical & Environmental Modeling 1* , 25-45.

Hodgson, M.J., Rosing, K.E., & Zhang, J. (1996). Locating Vehicle Inspection Stations to Protect a Transportation Network. *Geographical Analysis* , 299-314.

Hodgson, M.J. (1998). Developments in flow-based location-allocation models. In D. Griffith, C. Amrhein, & J.-M. Huriot, *Economic Advances in Spatial Modelling and Methodology* (pp. 119-132). Kluwer Academic Publishers.

Hudson, C. (1996). AIDS in rural Africa: a paradigm for HIV-1 infection. *Royal Society of Medicine Press* , 236-243.

Huff, D. (1966). A programmed solution for approximating an optimum retail location. *Land Economics 42* , 293-303.

Huff, D. (1964). Defining and estimating a trade area. *Journal of Marketing 28* , 34-38.

Kim, J-G. (2010). *Location of Refuelling Stations for Alternative Fuel Vehicles Considering Driver Deviation Behavior and Uneven Consumer Demand: Model, Heuristics, and GIS.* ProQuest.

Kuby, M., & Lim, S. (2005). The flow-refuelling location problem for alternative-fuel vehicles. *Socio-Economic Planning Sciences 39* , 125–145.

Kuby, M., & Lim, S. (2007). Location of Alternative-Fuel Stations Using the Flow-Refuelling Location Model and Dispersion of Candidate Sites on Arcs. *Networks and Spatial Economics 7* , 129-152.

Kuby, M., Lines, L., Schultz, R., Xie, Z., Kim, J.-G., & Lim, S. (2009). Optimisation of hydrogen stations in Florida using the Flow-Refuelling Location Model. *international journal of hydrogen energy 34* , 6045-6064.

Lim, S., & Kuby, M (2010). Heuristic algorithms for siting alternative-fuel stations using the Flow-Refuelling Location Model. *European Journal of Operational Research 204* , 51–61.

Ramjee, G., & Gouws, E. (2002). Prevalence of HIV among truck drivers visiting sex workers in KwaZulu-Natal, South Africa. *Sexually Transmitted Diseases* , 44-49.

Ramjee, G., & Gouws, E. (2000). *Targeting HIV-prevention efforts on truck drivers and sex workers: implications for a decline in the spread of HIV in Southern Africa.* Durban: South Africa Medical Research Council.

Rao, A. R., & Monroe, K. B. (1989). The Effect of Price, Brand Name and StoreName on Buyers' Perception of Product Quality: an Integrative Review. *Journal of Marketing Research* , 351-357.

Rosenkrantz, D.J., Tayi, G.K., & Ravi, S.S. (2000). Algorithms for path-based placement of inspection stations on networks. *Informs Jounal on Computing* , 136-149.

Sterle, C. (2010). *Location Routing Models and Methods for Freight Distribution and Infomobility in City Logistics.* Montreal: CIRRELT.

UNAIDS. (2000). *The Business Response to AIDS: Impact and Lessons Learned.* Geneva.

Whiteside, A. (2006). *The Socio-Economic Impact of HIV/AIDS in Swaziland.* Mbabane: National Emergency Response Committee on HIV/AIDS.

Wu, T.-H., & Lin, J.-N. (2003). Solving the competitive discretionary service facility location problem. *European Journal of Operational Research 144* , 366–378.

Yang, Y.-W., & Wang, C.-R. (2010). Locating passenger vehicle refuelling stations. *Transportation Research Part e 46* , 791–801.

Zeng, W. (2007). GIS Based Facility Location Planning with Different Types of Consumers. *Journal of Dissertation* , 1-126.

APPENDICES

A: Linearization of $\sum_{k \in K^q} \sum_{l \in K^q} i_{klq} h_2(t_{kl} - \overline{t_q})$ from constraints (5.6) & (5.6.1)

$$\sum_{k \in K^q} \sum_{l \in K^{qk}} i_{klq} h_2(t_{kl} - \overline{t_q}) = \sum_{k \in K^q} \sum_{l \in K^{qk}} ih_{klq} \quad \text{(A1)}$$

$$ih_{klq} \leq i_{klq} \quad \forall q \in Q, k \in K^q, l \in K^{qk} \quad \text{(A2)}$$

$$ih_{klq} \leq \sum_{j \in J^q} h_{jklq} \quad \forall q \in Q, k \in K^q, l \in K^{qk} \quad \text{(A3)}$$

$$ih_{klq} \geq \sum_{j \in J^q} h_{jklq} - (1 - i_{klq}) \quad \forall q \in Q, k \in K^q, l \in K^{qk} \quad \text{(A4)}$$

$$h_{jklq} \leq n_{jq} \quad \forall q \in Q, j \in J^q, k \in K^q, l \in K^{qk} \quad \text{(A5)}$$

$$h_{jklq} \leq \lambda 2_{3klq} / \theta_j \quad \forall q \in Q, j \in J^q, k \in K^q, l \in K^{qk} \quad \text{(A6)}$$

$$h_{jklq} \leq \lambda 2_{3klq} / \theta_j - (1 - n_{jq}) \quad \forall q \in Q, j \in J^q, k \in K^q, l \in K^{qk} \quad \text{(A7)}$$

$$\lambda 2_{1klq} + \lambda 2_{2klq} + \lambda 2_{3klq} + \lambda 2_{4klq} + \lambda 2_{5klq} = 1 \quad \forall q \in Q, k \in K^q, l \in K^{qk} \quad \text{(A8)}$$

$$\lambda 2_{1klq}(-M) + \lambda 2_{2klq}(-\hat{t}_3) + \lambda 2_{3klq} 0 + \lambda 2_{4klq} \hat{t}_3 + \lambda 2_{5klq} M_q = t_{kl} - \overline{t_q} \; \forall q \in Q, k \in K^q, l \in K^{qk} \quad \text{(A9)}$$

$$\lambda 2_{iklq} \leq z2_{iklq} \quad \forall q \in Q, k \in K^q, l \in K^{qk}, i \in \{1,2,3,4\} \quad \text{(A10)}$$

$$\{z2_{1klq}, z2_{2klq}, z2_{3klq}, z2_{4klq}\} \in AS \quad \forall q \in Q, k \in K^q, l \in K^{qk} \quad \text{(A11)}$$

$$ih_{klq} \geq 0 \quad \forall q \in Q, k \in K^q, l \in K^{qk} \quad \text{(A12)}$$

$$h_{jklq} \geq 0 \quad \forall q \in Q, j \in J^q, k \in K^q, l \in K^{qk} \quad \text{(A13)}$$

Where,

ih_{klq} : linear representation of $i_{klq} h_2(t_{kl} - \overline{t_q})$

J^q : $\{0,1,\ldots,|K^q|-1\}$ the set of all possible numbers of RWC neighbour-pairs defined along path q, indexed by j

n_{jq} : 1 if j neighbour-pairs can be defined along path *q* (i.e. if the number of (dummy-) RWCs along path *q* is j-1), 0 otherwise. This is ensured by means of constraint (5.8C) in case that assumption A10.1 holds, and by means of constraint (5.8C.1) in case that assumption A10.1 does not hold.

θ_j : parameter which is equal to the number of RWC neighbour-pairs element j corresponds to

M_q : an upper bound on $|t_{kl} - \overline{t_q}|$

$\lambda 2_{iklq}$: multiplier variables

$z2_{iklq}$: variables indicating whether $t_{kl} - \overline{t_q}$ is adjacent to point i

As we say in section 5, $h_2(t_{kl} - \overline{t_q})$ ranges between 0 and $1/n_q$, where $n_q = \sum_{k \in K^q} x_k - 1$ under assumption A10.1, and $\sum_{k \in K^q} 2x_k + \sum_{k \in KOD^q} x_k$ under assumption A10.2. Alternatively, one could rescale $h_2(t_{kl} - \overline{t_q})$ to a function which takes values between 0 and 1, and pre-multiply this function with $1/n_q$, which is represented by the summation $\sum_{j \in J^q} n_{jq} / \theta_j$.

In constraints (A8) – (A11), we define $t_{kl} - \overline{t_q}$ as a linear combination of two adjacent points among –M, $-\hat{t}_3$, 0, $\hat{t}_3$, and M (see figure 5.2). Because function h_2 is linear between all of these adjacent points, the values of (the rescaled function of) $h_2(t_{kl} - \overline{t_q})$ can be found by taking the same linear combination of the *function values* of these adjacent points. This is done in constraint (A9). The resulting value of the rescaled function, given the value of $t_{kl} - \overline{t_q}$ is equal to $\lambda 2_{3klq}$.

Multiplying this value with the summation $\sum_{j \in J^q} n_{jq} / \theta_j$ results in the value of $h_2(t_{kl} - \overline{t_q})$. However, models in which two variables are multiplied are very hard to solve. Therefore, we introduce the variable h_{jklq}, which is equal to $\lambda 2_{3klq} / \theta_j$ if the number of neighbour-pairs defined at path *q* is equal to j, and 0 otherwise. This is ensured by means of constraints (A5) – (A7).

Last, constraints (A1) – (A4) linearize the multiplication between i_{klq} and $h_2(t_{kl} - \overline{t_q})$, (the latter represented by $\sum_{j \in J^q} h_{jklq}$). These ensure that ih_{klq} is equal to $\sum_{j \in J^q} h_{jklq}$ if i_{klq} is equal to 1, and equal to 0 if i_{klq} is equal to 0.

B: Linearization of constraint (5.6.1)

Linearization of $\sum_{k\in O^q}\sum_{l\in KA^q} i_{klq} h_2(2t_{kl}+t_{kk}-\overline{t_q})$ *from constraint (5.6.1):*

$$\sum_{k\in O^q}\sum_{l\in KA^q} i_{klq} h_2(2t_{kl}+t_{kk}-\overline{t_q}) = \sum_{k\in O^q}\sum_{l\in KA^q} ih_{klq} \quad \text{(B1)}$$

$$ih_{klq} \le i_{klq} \quad \forall q\in Q, k\in O^q, l\in KA^q \quad \text{(B2)}$$

$$ih_{klq} \le \sum_{j\in J^q} h_{jklq} \quad \forall q\in Q, k\in O^q, l\in KA^q \quad \text{(B3)}$$

$$ih_{klq} \ge \sum_{j\in J^q} h_{jklq} - (1-i_{klq}) \quad \forall q\in Q, k\in O^q, l\in KA^q \quad \text{(B4)}$$

$$h_{jklq} \le n_{jq} \quad \forall q\in Q, j\in J^q, k\in O^q, l\in KA^q \quad \text{(B5)}$$

$$h_{jklq} \le \lambda 2_{3klq}/\theta_j \quad \forall q\in Q, j\in J^q, k\in O^q, l\in KA^q \quad \text{(B6)}$$

$$h_{jklq} \le \lambda 2_{3klq}/\theta_j - (1-n_{jq}) \quad \forall q\in Q, j\in J^q, k\in O^q, l\in KA^q \quad \text{(B7)}$$

$$\lambda 2_{1klq} + \lambda 2_{2klq} + \lambda 2_{3klq} + \lambda 2_{4klq} + \lambda 2_{4klq} = 1 \quad \forall q\in Q, k\in O^q, l\in KA^q \quad \text{(B8)}$$

$$\lambda 2_{1klq}(-M) + \lambda 2_{2klq}(-\hat{t}_3) + \lambda 2_{3klq} 0 + \lambda 2_{4klq}\hat{t} + \lambda 2_{5klq} M_q = t_{kl} - \overline{t_q} \quad \forall q\in Q, k\in O^q, l\in KA^q \quad \text{(B9)}$$

$$\lambda 2_{iklq} \le z2_{iklq} \quad \forall q\in Q, k\in O^q, l\in KA^q, i\in\{1,2,3,4\} \quad \text{(B10)}$$

$$\{z2_{1klq}, z2_{2klq}, z2_{3klq}, z2_{4klq}\} \in AS \quad \forall q\in Q, k\in O^q, l\in KA^q \quad \text{(B11)}$$

$$ih_{klq} \ge 0 \quad \forall q\in Q, k\in O^q, l\in KA^q \quad \text{(B13)}$$

$$h_{jklq} \ge 0 \quad \forall q\in Q, j\in J^q, k\in O^q, l\in KA^q \quad \text{(B14)}$$

Linearization of $\sum_{k\in KA^q}\sum_{l\in D^q} i_{klq} h_2(2t_{kl}+t_{ll}-\overline{t_q})$ *from constraint (5.6.1):*

$$\sum_{k\in KA^q}\sum_{l\in D^q} i_{klq} h_2(2t_{kl}+t_{ll}-\overline{t_q}) = \sum_{k\in KA^q}\sum_{l\in D^q} ih_{klq} \quad \text{(B15)}$$

$$ih_{klq} \le i_{klq} \quad \forall q\in Q, k\in KA^q, l\in D^q \quad \text{(B16)}$$

$$ih_{klq} \le \sum_{j\in J^q} h_{jklq} \quad \forall q\in Q, k\in KA^q, l\in D^q \quad \text{(B17)}$$

$$ih_{klq} \ge \sum_{j\in J^q} h_{jklq} - (1-i_{klq}) \quad \forall q\in Q, k\in KA^q, l\in D^q \quad \text{(B18)}$$

$$h_{jklq} \le n_{jq} \quad \forall q\in Q, j\in J^q, k\in KA^q, l\in D^q \quad \text{(B19)}$$

$$h_{jklq} \le \lambda 2_{3klq}/\theta_j \quad \forall q\in Q, j\in J^q, k\in KA^q, l\in D^q \quad \text{(B20)}$$

$$h_{jklq} \le \lambda 2_{3klq}/\theta_j - (1-n_{jq}) \quad \forall q\in Q, j\in J^q, k\in KA^q, l\in D^q \quad \text{(B21)}$$

$$\lambda 2_{1klq} + \lambda 2_{2klq} + \lambda 2_{3klq} + \lambda 2_{4klq} + \lambda 2_{4klq} = 1 \quad \forall q \in Q, k \in KA^q, l \in D^q \tag{B22}$$

$$\lambda 2_{1klq}(-M) + \lambda 2_{2klq}(-\hat{t}_3) + \lambda 2_{3klq}0 + \lambda 2_{4klq}\hat{t} + \lambda 2_{5klq}M_q = t_{kl} - \overline{t_q} \ \forall q \in Q, k \in KA^q, l \in D^q \tag{B23}$$

$$\lambda 2_{iklq} \leq z2_{iklq} \quad \forall q \in Q, i \in \{1,2,3,4\} \tag{B24}$$

$$\{z2_{1klq}, z2_{2klq}, z2_{3klq}, z2_{4klq}\} \in AS \quad \forall q \in Q \tag{B25}$$

$$ih_{klq} \geq 0 \quad \forall q \in Q, k \in KA^q, l \in D^{q^q} \tag{B26}$$

$$h_{jklq} \geq 0 \quad \forall q \in Q, j \in J^q, k \in KA^q, l \in D^q \tag{B27}$$

See appendix A for an explanation of the terms used here, and of the way the parts of constraint (5.6.1) are linearized.

C: Linearization of constraints (5.15.1) and (5.15.2)

Constraint (5.15.1)

$$t_{klq}^{diff} \geq \sum_{l \in KA^q} it_{klq}^{|diff|} \quad \forall q \in Q, k \in O^q \tag{5.15.1A}$$

$$it_{klq}^{|diff|} \leq i_{klq} M_q \quad \forall q \in Q, k \in O^q, l \in KA^q \tag{5.15.1B}$$

$$it_{klq}^{|diff|} \leq t_{klq}^{|diff|} \quad \forall q \in Q, k \in O^q, l \in KA^q \tag{5.15.1C}$$

$$it_{klq}^{|diff|} \geq t_{klq}^{|diff|} - M_q (1 - i_{klq}) \quad \forall q \in Q, k \in O^q, l \in KA^q \tag{5.15.1D}$$

$$t_{klq}^{|diff|} \geq 2t_{kl} + t_{kk} - \overline{t_q} \quad \forall q \in Q, k \in O^q, l \in KA^q \tag{5.15.1E}$$

$$t_{klq}^{|diff|} \geq \overline{t_q} - 2t_{kl} - t_{kk} \quad \forall q \in Q, k \in O^q, l \in KA^q \tag{5.15.1F}$$

Constraint (5.15.2)

$$t_{klq}^{diff} \geq \sum_{k \in KA^q} it_{klq}^{|diff|} \quad \forall q \in Q, l \in D^q \tag{5.15.2A}$$

$$it_{klq}^{|diff|} \leq i_{klq} M_q \quad \forall q \in Q, k \in KA^q, l \in D^q \tag{5.15.2B}$$

$$it_{klq}^{|diff|} \leq t_{klq}^{|diff|} \quad \forall q \in Q, k \in KA^q, l \in D^q \tag{5.15.2C}$$

$$it_{klq}^{|diff|} \geq t_{klq}^{|diff|} - M_q (1 - i_{klq}) \quad \forall q \in Q, k \in KA^q, l \in D^q \tag{5.15.2D}$$

$$t_{klq}^{|diff|} \geq 2t_{kl} + t_{kk} - \overline{t_q} \quad \forall q \in Q, k \in KA^q, l \in D^q \tag{5.15.2E}$$

$$t_{klq}^{|diff|} \geq \overline{t_q} - 2t_{kl} - t_{kk} \quad \forall q \in Q, k \in KA^q, l \in D^q \tag{5.15.2F}$$

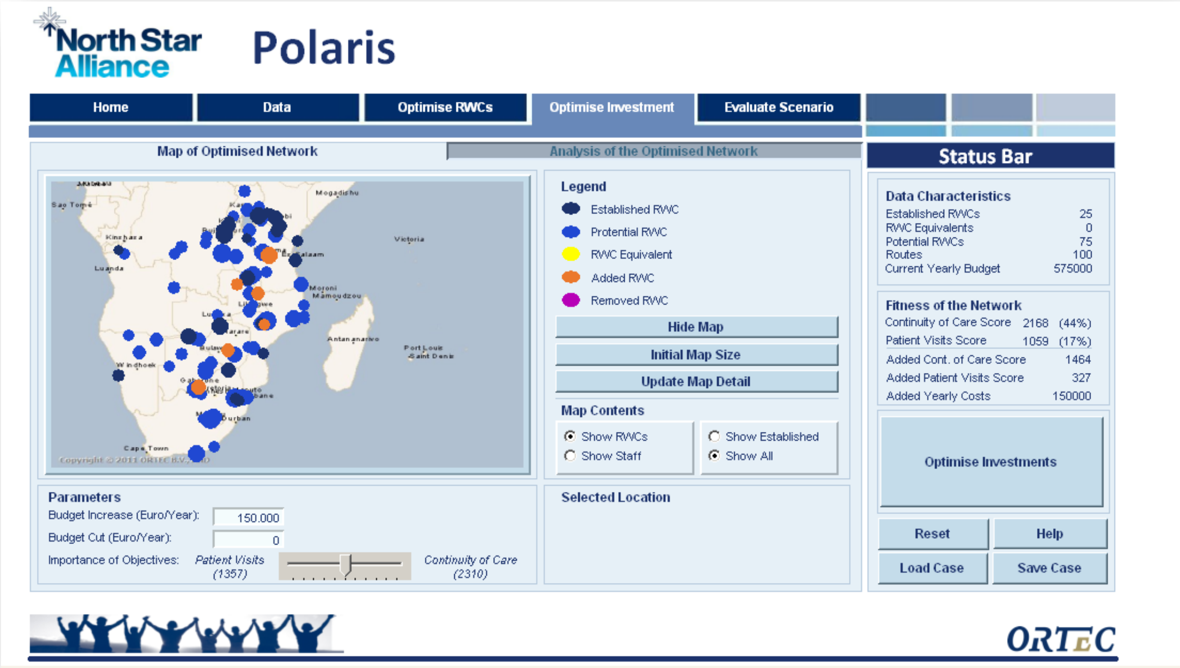
Where:

$it_{klq}^{|diff|}$: linear representation of $i_{klq} \,| t_{kl} - \overline{t_q} \,|$

$t_{klq}^{|diff|}$: linear representation of $| t_{kl} - \overline{t_q} \,|$

The value of $| 2t_{kl} + t_{kk} - \overline{t_q} \,|$ is linearized in constraints (5.15.1E) and (5.15.1F). Because a larger value of this absolute difference never results in a larger objective value, these constraints force that $t_{klq}^{|diff|}$ is chosen as small as possible. So, we do not need to restrict $t_{klq}^{|diff|}$ to be *equal to* the absolute difference. Constraints (5.15.1A) – (5.15.1D) ensure that t_{klq}^{diff} is equal to $t_{klq}^{|diff|}$ if i_{klq} is equal to 1, and that t_{klq}^{diff} is equal to 0 if i_{klq} is equal to 0. The linearization of constraint (5.15.2) can be explained in the same way.

D: Screenshot of POLARIS



E: Test results

RIM

p	r=0	r=0.2	r=0.5	r=0.8	r=1
1	792	785	785	785	785
2	852	818	815	815	815
5	1025	954	908	908	908
10	1290	1227	1111	1045	1045
20	1747	1703	1522	1409	1372
30	2077	2047	1912	1797	1707
40	2384	2345	2256	2149	2059
50	2619	2612	2561	2510	2280
60	2819	2799	2799	2799	2533
70	2987	2984	2984	2984	2943
75	3040	3040	3040	3040	3040

Table E.1: patient visits scores of optimal networks obtained after solving the RIM with different parameters.

p	r=0	r=0.2	r=0.5	r=0.8	r=1
1	737	1001	1001	1001	1001
2	755	1233	1249	1249	1249
5	1052	1852	2019	2019	2019
10	1384	2276	2704	2794	2794
20	1707	2858	3751	3950	3955
30	2553	3755	4418	4632	4639
40	3758	4430	4786	4873	4892
50	4483	4660	4893	4941	4960
60	4751	4958	4958	4958	4965
70	4868	4962	4962	4962	4965
75	4965	4965	4965	4965	4965

Table E.2: continuity of care scores of optimal networks obtained after solving the RIM with different parameters

RSIM

bl	r=0	r=0.2	r=0.5	r=0.8	r=1
50000	982	854	854	854	854
60000	1030	990	802	802	802
70000	1062	965	874	874	874
100000	1173	1099	862	862	862
200000	1538	1514	1176	968	968
500000	2493	2400	1839	1292	1260
750000	3118	3043	2570	2195	1565
1000000	3624	3572	3251	2840	1785
1500000	4548	4488	4335	4170	3479
2000000	5217	5182	5159	5117	4800
2500000	5774	5756	5756	5750	5552
3000000	6132	6132	6132	6132	6132
3325000	6253	6253	6253	6253	6253

Table E.3: patient visits scores of optimal networks obtained after solving the RSIM with different parameters.

bl	r=0	r=0.2	r=0.5	r=0.8	r=1
50000	704	1249	1249	1249	1249
60000	704	768	1595	1595	1595
70000	704	1249	1595	1595	1595
100000	722	1249	2019	2019	2019
200000	755	1138	2448	2794	2794
500000	1385	2347	3858	4406	4410
750000	1707	2938	4341	4716	4877
1000000	2401	3566	4590	4873	4962
1500000	3677	4426	4854	4941	4965
2000000	4672	4904	4939	4958	4965
2500000	4780	4962	4962	4963	4965
3000000	4965	4965	4965	4965	4965
3325000	4965	4965	4965	4965	4965

Table E.4: continuity of care scores of optimal networks obtained after solving the RSIM with different parameters