Time-Dependent Transportation Network Design that Considers Health Cost

Y. Jiang, W.Y. Szeto*

Department of Civil Engineering, The University of Hong Kong, Hong Kong, P. R. China

* *Phone: (852) 28578552, Fax: (852) 25595337, Email: ceszeto@hku.hk*

ABSTRACT

This paper proposes a bilevel optimization framework for time-dependent discrete road network design that considers health impacts. A general health cost function is proposed and captured in the framework. The function simultaneously considers the health impacts of road traffic emissions, noise, and accidents due to network expansion. To solve the problem, Artificial Bee Colony algorithm (ABC) is proposed to search the network design solutions of the upper level problem, while the method of successive averages (MSA) and the Frank-Wolfe algorithm is adopted to solve the lower-level time-dependent land-use transport problem. A repairing procedure is proposed to remedy infeasible solutions. A numerical study is set up to illustrate the conflict between maximizing consumer surplus and minimizing the health cost. This paper also reveals a paradox phenomenon that with an increasing amount of emissions, the health cost is decreasing. Moreover, the existence of a health inequity between different residential zones is demonstrated. The Sioux Falls network is adopted to show the performance of the solution algorithm as well as the effectiveness of the proposed repairing procedure.

Keywords: time-dependent transport network design; health cost; consumer surplus; artificial bee colony algorithm; discrete network design

1. Introduction

Currently, many road network improvement projects are still ongoing or under planning in many parts of the world. They should be carefully selected, as they influence travelers' route choices and travel costs, and are expensive. The design problem involved is called the transportation network design problem. This problem has been studied over 30 years. A recent review on the studies can be found in Farahani et al. (2013).

A classical modeling approach (i.e., LeBlanc, 1975; Boyce and Janson, 1980; Marcotte, 1986; Friesz et al., 1993; Meng et al., 2001; Ge et al., 2003; Chen and Yang, 2004; Chen et al., 2007; Ukkusuri et al., 2007; Ukkusuri and Waller, 2008; Chiou, 2009; Long et al., 2010; Wang and Lo, 2010) to this problem is based on the equilibrium traffic assignment framework. This approach allows analyzing the problem easily but cannot capture the realistic variations in demand during the planning horizon, the gradual changing network, and the land-use transport interaction.

To address the land-use transport interaction issue, some works (e.g., Smith and Liebman, 1978; Los, 1979; Meng et al., 2000; Yim et al., 2011; Li et al., 2014) were conducted, but still did not capture the changing demand over time and gradual network upgrades. On the other hand, in the last 10 years, time-dependent frameworks (e.g., Szeto and Lo, 2005, 2008; Kim et al., 2008; Lo and Szeto, 2009; Ukkusuri and Patil, 2009; Unnikrishnan, 2009; Xu et al., 2013) were proposed to deal with the changing demand and network upgrades over time. However, the land use component was still missing. Hence, Szeto et al. (2010, 2013) and Ma and Lo (2013) developed frameworks that can capture the land use transport interaction, the network upgrades, and the changing demand over time.

Another issue of most of the preceding frameworks is that they only consider single objective that minimizes the total system travel cost or maximize social surplus (e.g., Li et al. 2012a). They ignore the externalities associated with road traffic, such as traffic emissions, noise, and accidents, to the population in the residential areas in their objectives. They also fail to capture other important transportation planning or sustainability objectives, such as traffic emissions, noise, and accident reductions. As pointed out by May (2013), models of the urban transport system need to be able to reflect the full range of objectives associated with sustainability. To have a more sustainable network design, it is therefore important to develop multi-objective network design frameworks to incorporate different sustainability objectives.

In the literature, multi-objective network design models were proposed to handle the trade-off between different objectives, including those associated with sustainability (e.g., Lin and Feng, 2003; Chen et al., 2006, 2010; Ban et al., 2006; Kim and Kim, 2006; Qiu and Chen, 2007; Sharma and Mathew, 2011; Miandoabchi et al., 2012a, b, 2013). However, these objectives seldom directly consider health impacts that are increasingly concerned by the public. The public are concerned with road traffic emissions, such as carbon monoxides (CO), nitrogen oxides (NO), and particulate matter (PM), because of their negative health impacts and medical costs induced. Those emissions can hurt the respiratory system, induce asthma or even lung cancer, increase the risk of cardiovascular problems, or impede the psychomotor functions. The associated medical expense can be a burden to affected low-income families.

Road traffic emissions, one of sustainability indicators that has been included in the design objective (e.g., Lin and Feng, 2003; Yin and Lawphongpanich, 2006; Qiu and Chen, 2007; Szeto et al., 2014; Long et al., 2014), cannot be directly used to reflect health impacts and costs, because different types of emissions have various health impacts on different groups of people. To our best knowledge, only Sharma and Mathew (2011) captured health costs by adopting a health-damage cost function to convert traffic emissions to health costs in their network design problem. However, the characteristics (such as age, health conditions) and the size of affected population are not considered. In fact, they are critical factors to assess the health impacts (Fehr, 1999). Without considering them, the health impacts on residents may be wrongly estimated. For example, health damages can be underestimated (overestimated) if traffic emissions are low (high). It is because (i) traffic emissions can influence a large (small) population, (ii) the affected population can have a high (low) average age, and (iii) the affected population can be more (less) vulnerable to pollution damage.

More importantly, road traffic emissions are not the only sources that damage people's health. Traffic noise and accidents, the other two sustainability indicators that have been captured in the design objective (e.g., Szeto et al., 2014), also induce health problems, such as hearing problems and physical disabilities. Therefore, in this paper, a general health cost function, which simultaneously incorporates the impacts of traffic emissions, noise, and accidents, is proposed and captured in the discrete network design problem – designing the link capacity enhancement subject to the budgetary and spatial constraints. The problem is formulated as a bilevel problem where the upper level problem maximizes the sum of the increase in consumer surplus and the reduction in the health cost, while the lower level problem is the time-dependent land-use transport interaction problem. The bilevel problem is indeed bi-objective as it considers an economic objective as measured by the increase in consumer surplus and a health objective as measured by the health cost.

To solve this problem, Artificial Bee Colony algorithm (ABC) is proposed to search the network design solutions in the upper level problem, while the method of successive averages (MSA) and the Frank-Wolfe algorithm is adopted to solve the lower level time-dependent land-use transport problem. A repairing procedure is proposed to handle infeasible solutions. Numerical studies are set up to illustrate the performance of the solution algorithm, the repairing procedure, and the existence of the equity issue regarding consumer surplus and the health cost.

This paper differs from the study by Szeto et al. (2013) in twofold. First, this paper focuses on investigating the health impacts due to network expansion, while Szeto et al. (2013) focused on the measures of sustainability. Second, this paper uses a vehicle emission function that is more general than the one used by Szeto et al. (2013). This paper also has the following contributions:

- Proposing a framework for the bi-objective road network design problem with land-use transport interaction over time and health cost consideration;
- Proposing a general health cost function, which simultaneously captures the health impacts of traffic emissions, noise, and accidents, due to network improvement projects;
- Considering affected population when evaluating the health cost, and demonstrating its significance;
- Illustrating a paradox phenomenon that with the increasing amount of emissions, the emission induced health cost is decreasing;
- Developing an ABC algorithm to solve the bilevel problem; the algorithm and the model can be used as a decision making tool of transportation projects for many cities, and;
- Pointing out the existence of the health inequity issue;

The remainder of this paper is organized as follows: Section 2 describes the bi-level framework. The proposed algorithm is given in Section 3, followed by various numerical examples in Section 4. Section 5 discusses the need of conducting sensitivity analysis, model calibration and validation, and critical parameter identification. Finally, Section 6 provides some concluding remarks.

2. Framework

We consider a connected multi-modal transportation network with multiple Origin-Destination (OD) flows over the planning horizon [0, T]. The planning horizon is equally divided into N design periods. The mode here can be an individual mode or a combined mode. The following notations are used throughout the paper.

2.1 Notations

2.1.1 Indices

i, j: indices of zones;

m: index of travel mode;

h: index of the type of negative externalities; $h = 1,\ldots, n$ represents various types of traffic emissions, e.g., CO, PM or NO; $h = n + 1$ represents traffic noise; $h = n + 2$ represents traffic accidents, where n is the maximum number of the types of vehicle exhausts considered;

a: index of link;

τ: index of time period;

d: index of the coefficient associated with the emission function.

2.1.2 Set

N: set of zones;

P^{ij}: set of paths connecting OD pair ij;

L: set of links;

A_i^+ (A_i^-) set of links coming out (going into) zone i.

2.1.3 Parameters

N_a: number of candidate links;

α: economy-of-scale parameter to regulate the attractiveness of each zone on the allocation of residents;

$\tilde{\alpha}$: economy-of-scale parameter to regulate the attractiveness of each zone on the allocation of service employment;

$\alpha_{a,m}$, $\gamma_{a,m}$: parameters of the performance function of link a for mode m;

B: total budget;

β^r : parameter to regulate the effect of transport cost on the distribution of residents;

s: service employment-to-population ratio, which is less than 1;

β^s : parameter to regulate the effect of transport cost on the distribution of service employment;

$\overline{\beta}$: parameter in the logit model to regulate the effect of the mode travel cost;

$\overline{c}_a^0$: initial capacity of link a;

$\delta_a^{p,m}$: link-path incidence indicator for mode m, which equals one if link a is on route p, and zero otherwise;

$E_{i,\tau}^B$: basic employment in zone i during period τ;

$\tilde{g}_{E,i}$: growth rate of basic employment of zone i;

$\tilde{g}_{w,i}$, $\tilde{g}'_{w,i}$: growth rates of attractiveness of zone i ;

μ : population-to-employment ratio, which is intuitively not less than one and $\mu \le 1/s$;

ψ : cost of unit (travel) time;

λ^c, λ^h : conversion parameters that convert consumer surplus and the health cost measurements to period basis (i.e., if τ represents one year, traffic flow is hourly flow and the health impact study is on an annual basis, then $\lambda^c = 8760, \lambda^h = 1$);

$\rho_{a,\tau}^m$: toll for mode m or distance-based fare of passengers on mode m traversing link a during period τ;

$t_a^{0,m}$: free flow travel time on link a used by mode m;

θ^m : mode-specific cost of mode m;

$W_{j,\tau}$: attractiveness of zone j during period τ, which can be represented by the availability of floor space in zone j during that period for residential usage;

$\tilde{W}_{i,\tau}$: attractiveness in zone i during period τ, which can be represented by the availability of floor space in that zone during that period for commercial use;

$U_{a,\tau}$: unit construction cost of link a during period τ, which depends on the length of and the terrain around that link;

u_a: maximum allowable expansion capacity of link a;

l_a: length of link a;

$k^{m,h}, b_d^{m,h}$: parameters for computing vehicle emissions of type h generated by vehicles of mode m; $d = 0,1,\ldots,6$.

2.1.4 Decision variables

$f_{p,ij,\tau}^m$: hourly flow of mode m on route p between OD pair ij during period τ;

f: $\left[f_{p,ij,\tau}^m\right]$;

$E_{ij,\tau}^S$: number of service employees who work in zone i and live in zone j or the number of service employment trips between OD pair ij during period τ;

$R_{ij,\tau}$: number of work-to-home trips between OD pair ij during period τ or the number of residents in zone j that work in zone i during period τ;

R: $\left[R_{ij,\tau}\right]$;

$y_{a,\tau}$: discrete variable, representing the capacity increment to link a at the beginning of period τ due to road widening (e.g., increasing the width of a lane or adding extra lanes);

y: $\left[y_{a,\tau}\right]$;

2.1.5 Functions of decision variables

$c_{ij,\tau}$: composite travel cost between OD pair *ij* during period τ, representing the inter-zonal impedance;

$c^m_{p,ij,\tau}$: travel cost of mode *m* on route *p* between OD pair *ij* during period τ;

$\overline{c}_{a,\tau}$: capacity of link *a* during period τ;

ΔCS: increase in consumer surplus;

ΔH: decrease in the health cost;

$E_{i,\tau}$: total employment, including service employment and basic employment, in zone *i* during period τ;

$E^S_{i,\tau}$: service employment or non-basic employment in zone *i* during period τ, or simply the number of service employees working there during that period;

$E^S_{ij,\tau}$: number of service employees who work in zone *i* living in zone *j* or the number of service employment trips between OD pair *ij* during period τ;

$g_{a,\tau}(y_{a,\tau})$: construction cost function of link *a* in period τ, which is generally assumed to be nonnegative, increasing, and differentiable;

$\pi^m_{ij,\tau}$: lowest travel cost between OD pair *ij* by mode *m* during period τ;

$q^m_{ij,\tau}$: demand for mode *m* between OD pair *ij* during period τ;

$R_{j,\tau}$: total number of residents in zone *j* during period τ;

$t^m_{a,\tau}$: link time (such as in-vehicle travel, waiting, and walking times) of mode *m* on link *a* during period τ;

$v^m_{a,\tau}$: hourly flow of mode *m* on link *a* during period τ;

$s^m_{a,\tau}$: average speed of mode *m* on link *a* during period τ;

$e^{h}_{a,\tau}$: negative externality *h* due to all modes traveling on link *a* during period τ;

$\Delta z^{h}_{i,\tau}$: change in the health impact due to externality *h* in zone *i* during period τ.

2.2 Bi-level Formulation Framework

Based on preceding notations, we can formulate our proposed network design problem as a bi-level one, which assumes that the decision maker intends to improve the traveler benefit or the network performance, measured by the increase in Consumer Surplus (CS), and the health condition of residential people, represented by the decrease in the health cost. This assumption is realistic when the decision maker is the government who aims to improve the benefits of both commuters and the residents by expanding transportation networks. The bi-level problem can be summarized as follows.

Upper level problem:

$$\max_{\mathbf{y},\mathbf{f},\mathbf{R}} \Delta CS + \Delta H$$

Subject to

design and financial constraints;

Lower level problem (expressed as a set of constraints):

time-dependent Lowry-type land-use constraints;

time-dependent modal-split/assignment constraints.

The first term in the upper level objective function represents the increase in consumer surplus ΔCS, while the second term is the reduction in the health cost ΔH. A larger value of ΔCS indicates a better performance of the transportation network, while a larger value of ΔH implies a better health condition of residents. The upper level objective value measures the impacts of transportation improvement projects on travelers and residential population. Although the upper level objective function implicitly assumes that consumer surplus and the health cost are equally important, there is no difficulty to add weighting parameters to cater the planner's different perception on these two objectives. The upper level problem is constrained by design requirements and resources. The lower level problem is represented by the constraints for land use transport interaction over time.

2.3 The Lower Level Problem

For the lower level problem, we consider two types of constraints: Time-dependent Lowry-type constraints and time-dependent modal-split/assignment constraints.

2.3.1 Time-dependent Lowry-type constraints

The Lowry model was developed to predict the population and employment distributions (Lowry, 1964). The Lowry model considers three sectors, namely *basic sector*, *non-basic sector*, and *household sector*. The basic sector represents industries that do not change their locations even if the transportation network changes over time. Their clients are mainly outside the study area. The basic sector provides jobs to local residents and the number of jobs offered by the basic sector is referred to as basic employment. The non-basic sector provides services to residential community. Their locations highly depend on the location and the size of population. The number of jobs offered by the non-basic sector is referred to as service employment. The household sector consists of residential population. Their residential location depends on the locations and numbers of basic and non-basic jobs available. Given model inputs, consisting of *composite travel cost* between zones and the spatial distribution of *basic employment* of base year (that can be obtained from travel characteristics surveys (Wong et al., 1999), the Lowry model determines model outputs such as the *service employment* and *residential population* of each zone.

The time-dependent Lowry-type constraints extend the Lowry model to a dynamic one in a way that during each design period, a Lowry-type equilibrium holds. The equilibrium conditions can be depicted by the following constraints.

$$R_{ij,\tau} = \frac{E_{i,\tau} W_{j,\tau}^{\alpha} \exp\left(-\beta^r c_{ij,\tau}\right)}{\sum_{j'} W_{j',\tau}^{\alpha} \exp\left(-\beta^r c_{ij',\tau}\right)}, \forall i, j, \tau , \tag{1}$$

$$E_{ij,\tau}^{S} = \frac{sR_{j,\tau} \tilde{W}_{i,\tau}^{\tilde{\alpha}} \exp\left(-\beta^s c_{ij,\tau}\right)}{\sum_{i'} \tilde{W}_{i',\tau}^{\tilde{\alpha}} \exp\left(-\beta^s c_{i'j,\tau}\right)}, \forall i, j, \tau , \tag{2}$$

$$E_{i,\tau} = E_{i,\tau}^{B} + E_{i,\tau}^{S}, \forall i, \tau , \tag{3}$$

$$E_{i,\tau}^{S} = \sum_{j} E_{ij,\tau}^{S}, \forall i, j, \tau , \tag{4}$$

$$R_{j,\tau} = \mu \sum_{i} R_{ij,\tau}, \forall j, \tau . \tag{5}$$

Equations (1) and (2) are gravity type constraints to determine home-to-work and service employment trip distributions for each period, respectively. Equations (3) and (4) are the conservation constraints of total and service employments during each time period. Equation (5) determines the zonal population distribution for each period. For the time-dependent Lowry-type constraints, more discussions can be found in Szeto et al. (2010, 2013).

The attractiveness of each zone and the basic employments for each period are model inputs. For simplicity, they are assumed to increase at a constant rate:

$$W_{i,\tau+1} = W_{i,\tau}\left(1 + \tilde{g}_{w,i}\right), \forall i, \tau , \tag{6}$$

$$\tilde{W}_{i,\tau+1} = \tilde{W}_{i,\tau}\left(1 + \tilde{g}'_{w,i}\right), \forall i, \tau , \tag{7}$$

$$E_{i,\tau+1}^{B} = E_{i,\tau}^{B}\left(1 + \tilde{g}_{E,i}\right), \forall i, \tau . \tag{8}$$

Other parameters in the Lowry model can be calibrated via the method described in Wong et al. (1999) using the data obtained from household surveys.

2.3.2 Time-dependent modal-split/assignment constraints

The time-dependent modal-split assignment constraints describe the route and mode choices over time and can be expressed as:

$$f^m_{p,ij,\tau}\left[c^m_{p,ij,\tau} - \pi^m_{ij,\tau}\right] = 0, \forall p \in P^{ij}, i, j, m, \tau\ , \tag{9}$$

$$c^m_{p,ij,\tau} - \pi^m_{ij,\tau} \geq 0, \forall p \in P^{ij}, i, j, m, \tau\ , \tag{10}$$

$$v^m_{a,\tau} = \sum_{ij}\sum_{p\in P^{ij}} f^m_{p,ij,\tau}\delta^{p,m}_a, \forall a, m, \tau\ , \tag{11}$$

$$t^m_{a,\tau} = t^{0,m}_a\left[1+\left(\frac{\sum_k \alpha_{a,k}v^k_{a,\tau}}{\bar{c}_{a,\tau}}\right)^{\gamma_{a,m}}\right], \forall a, m, \tau\ , \tag{12}$$

$$\bar{c}_{a,\tau} = \bar{c}^0_a + \sum_{\omega=1}^{\tau} y_{a,\omega}, \forall a, \tau\ , \tag{13}$$

$$c^m_{p,ij,\tau} = \sum_a\left(\psi t^m_{a,\tau} + \rho^m_{a,\tau}\right)\cdot\delta^{p,m}_a, \forall p \in P^{ij}, i, j, m, \tau\ , \tag{14}$$

$$c_{ij,\tau} = \frac{-\ln\left[\sum_m (\exp(-\bar{\beta}\left(\pi^m_{ij,\tau} + \theta^m\right))\right]}{\bar{\beta}}, \forall i, j, \tau, \tag{15}$$

$$q^m_{ij,\tau} = R_{ij,\tau}\left(\frac{\exp(-\bar{\beta}\left(\pi^m_{ij,\tau} + \theta^m_{ij}\right))}{\sum_k \exp(-\bar{\beta}\left(\pi^k_{ij,\tau} + \theta^k_{ij}\right))}\right), \forall i, j, m, \tau\ , \tag{16}$$

$$q^m_{ij,\tau} = \sum_{p\in P^{ij}} f^m_{p,ij,\tau}, \forall i, j, m, \tau\ , \tag{17}$$

$$f^m_{p,ij,\tau} \geq 0, \forall p, i, j, m, \tau\ . \tag{18}$$

Conditions (9) and (10) are the user-equilibrium conditions for each period. Equation (11) states that for each mode during each period, the link flow is obtained by summing the corresponding route flows on that link. Equation (12) computes link time by an asymmetric link performance function. Equation (13) expresses the capacity of link a used during period τ as the sum of its initial capacity $\bar{c}^0_a$ and the total improvements up to period τ. Equation (14) defines the route cost associated with a period to be the sum of the costs associated with the links on the route during that period. Equation (15) computes the composite cost associated with OD pair ij and period τ by aggregating the mode travel costs $\pi_{ij,\tau}^{m}+\theta^{m}$ over all the modes considered. The derivation of this composite cost can be found in Williams (1977). Equation (16) is the logit model for determining the OD demand for each mode and period. Equations (17) and (18) are the conservation and non-negativity conditions of route flows, respectively.

2.3.3 Lower level fixed point formulations

As shown by Szeto et al. (2013), the lower level problem can be formulated as

$$\begin{bmatrix}\mathbf{R}\\ \mathbf{f}\end{bmatrix}=\begin{bmatrix}\mathbf{F}\left(\mathbf{R},\mathbf{f}\right)\\ \mathbf{G}\left(\mathbf{R},\mathbf{f}\right)\end{bmatrix}, \tag{19}$$

where $\mathbf{F}\left(\mathbf{R},\mathbf{f}\right)$ and $\mathbf{G}\left(\mathbf{R},\mathbf{f}\right)$ are the vector functions defined by (1) - (8) and (9) - (17), respectively. According to the Brouwer's fixed-point theorem, if the solution set for (19) is compact and convex, and $\mathbf{F}\left(\mathbf{R},\mathbf{f}\right)$ and $\mathbf{G}\left(\mathbf{R},\mathbf{f}\right)$ are continuous, then there exists at least one solution to the lower level problem. For the proposed formulation, the continuity is guaranteed because each function in the constraints is continuous over the solution space. Moreover, the solution set is compact because the solution set is a closed, bounded ball (Szeto et al., 2013). Therefore, a solution must exist to the lower level problem. However, the uniqueness of solutions further requires that the mapping function is strictly monotone, which may not be satisfied as the cost function (12) is asymmetric with respect to path flows. This problem differs from other time-dependent assignment models (e.g., Nakayama and Connors, 2014) that the solution is unique.

2.4 The upper level problem

The upper level objective function equals the sum of two main components: the changes in consumer surplus and the changes in the health cost.

2.4.1 Increase in consumer surplus

Consumer surplus (CS) measures the difference between what consumers would be willing to pay for travel and what they actually pay. It internalizes the effect of network congestion and the public's propensity to travel. For the same network and demand characteristics, a higher CS implies a better performing system. Hence, when we plan for transportation network improvement, the network improvement implemented should increase CS; this implies that the change in CS should be positive, because the CS before improvement is fixed and larger than that before. Therefore, we can use the change in CS or the increase in CS as an alternative measure. Here, a common and simple approximation to this change, denoted as ΔCS, is employed (Williams, 1976) and is expressed in *its present value* as follows:

$$\Delta CS = \sum_{\tau} \sum_{ij} \sum_{m} \frac{\lambda^{c} \Delta CS_{ij,\tau}^{m}}{\left(1+\tilde{i}\right)^{\tau-1}}, \tag{20}$$

$$\Delta CS_{ij,\tau}^{m} \approx (1/2)(q_{ij,\tau}^{m,before} + q_{ij,\tau}^{m,after})(\pi_{ij,\tau}^{m,before} - \pi_{ij,\tau}^{m,after}), \forall i, j, m, \tau, \tag{21}$$

where the superscripts 'before' and 'after' denote before and after the improvement is finished, respectively. $\tilde{i}$ is the interest rate. $(1+\tilde{i})^{-(\tau-1)}$ is the discount factor for period τ. According to equation (20), ΔCS is the sum of the changes in CS for all modes and OD pairs over all periods, discounted to present value terms. Equation (21) is the rule-of-half definition for ΔCS and the derivation is presented in the appendix. Since the demand and flow are measured on an hourly basis, thus they are converted to the period basis by multiplying conversion parameter λ^{c} in equation (20).

2.4.2 Decrease in the health cost

The decrease in the health cost associated with transportation network expansion directly depends on health impacts, which in turn depend on at least three types of negative externalities of transportation networks: traffic emissions, noise, and accidents. Various models were proposed in the literature to predict these externalities (e.g., Hakim et al., 1991; Steele, 2001; Szeto et al., 2012). We use flow- or speed-dependent models because they are popularly used in the literature.

Traffic emissions. Vehicles generate various exhaust gases, such as CO, NO, and PM. The quantities produced depend on various factors, including the types of engines and vehicle speed. The paper adopts a polynomial model developed by Boulter et al. (2009) to compute quantities of each type of vehicle exhausts produced:

$$e_{a,\tau}^{h} = \sum_{m} v_{a,\tau}^{m} \left[k^{m,h} \sum_{d=0}^{6} b_{u}^{m,h} \left(s_{a,\tau}^{m} \right)^{d-1} \right] l_{a}, \forall h = 1,...,n, \forall a, \tau , \qquad (22)$$

where l_a is the length of link a. Equation (22) states that during the representative hour of period τ, the total emissions of type h generated by all modes on link a are determined by the sum of all of the products of the flow on that link and emissions generated by each vehicle traveling at the speed $s_{a,\tau}^{m}$. The above equation can also capture the emissions generated by various types of vehicles (i.e., private car or buses) by treating each type of vehicles as a separated mode.

Traffic noise. It makes people annoyance, brings stress, and causes hearing problems. To compute the traffic noise level and sound energy generated, a simplified FHWA (Menge et al., 1998) model is adopted, where the hourly A-weighted equivalent sound level generated by certain type of vehicles is used and obtained by adding different adjustments accounting for flow, distance, and shielding effects to the reference noise level. Mathematically, the sound level generated by mode m traveling on link a for period τ, $L_{a,\tau}^{m}$, can be calculated by

$$L_{a,\tau}^{m} = EL_{a,\tau}^{m} \left(s_{a,\tau}^{m} \right) + A_{a,traffic,\tau}^{m} \left(v_{a,\tau}^{m}, s_{a,\tau}^{m} \right) + A_{d} + A_{s}, \forall a, \tau, m , \qquad (23)$$

where $EL_{a,\tau}^{m}\left(s_{a,\tau}^{m} \right)$ represents the reference noise-emission level due to a mode-m vehicle traveling on link a during the represented hour of period τ, and is the maximum A-weighted sound level recorded by a microphone 15 meters (50 feet) to the side when that vehicle passes by over flat, generally absorptive terrain. $A^m_{a,traffic,\tau}\left(v^m_{a,\tau}, s^m_{a,\tau}\right)$ denotes the adjustment made due to the traffic flow of mode *m* on link *a* during the representative hour of period τ. A_d and A_s, respectively, represent the adjustment for the distance between the receiver and the roadway and the adjustment for all shielding and ground effects which are independent of vehicle type. In this study, they are assumed to be constant because the network improvement in this study only affects the capacities of candidate links which in turn slightly affect the corresponding distance and ground effects. The changes in A_d and A_s are too small compared with the first two terms in equation (23). The first term $EL^m_{a,\tau}\left(s^m_{a,\tau}\right)$ in equation (23) can be estimated by

$$EL^m_{a,\tau}\left(s^m_{a,\tau}\right) = 10\log_{10}\left((0.6214 s^m_{a,\tau})^{A/10}10^{B/10} + 10^{C/10}\right), \forall a, m, \tau \, , \tag{24}$$

where *A*, *B*, and *C* are calibrated parameters related to vehicle, pavement, and throttle types. For instance, when an automobile is under the full throttle condition and the pavement is made up of both DGAC (dense-graded asphaltic concrete) and PCC (Portland cement concrete), those parameters are set as follows: A = 41.740807, B = 1.148546, and C = 50.128316. $A^m_{a,traffic,\tau}\left(v^m_{a,\tau}, s^m_{a,\tau}\right)$ is calculated by

$$A^m_{a,traffic,\tau}\left(v^m_{a,\tau}, s^m_{a,\tau}\right) = 10\times\log_{10}\left(\frac{v^m_{a,\tau}}{s^m_{a,\tau}}\right) - 13.2dB, \forall a, \tau, m \, . \tag{25}$$

Equations (23) – (25) compute the noise level in unit of dB for each link. To aggregate noise level, the additive sound energy associated with link *a* during the representative hour of period τ is computed by

$$e^{n+1}_{a,\tau} = \sum_m 10^{L^m_{a,\tau}/10}, \forall a, \tau \, . \tag{26}$$

Traffic accidents. A higher speed causes a larger risk of drivers, as a longer stopping distance is required when emergence happens. A higher speed also induces more serious damages to pedestrians due to a larger momentum caused by a higher speed. Therefore, we adopt the following power model (Nilsson, 2004; Cameron and Elvik, 2010) to predict the resultant number of accidents as a result of the change in travel speed after road network expansion:

$$e_{a,\tau}^{n+2} = \sum_{m} Ac_{a,0} \left(\frac{s_{a,\tau}^{m}}{s_{a,0}^{m}} \right)^{P}, \forall a, \tau, \tag{27}$$

where $Ac_{a,0}$ denotes the number of accidents on link *a* before network expansion and is an input to the model. Similarly, $s_{a,0}^{m}$ is the average travel speed before network expansion. The power *P* depends on the level of injury severity or the type of accidents; in particular, when *P* = 4, it represents fatal crashes.

Decrease in health impacts in each zone. The above externalities are estimated for each link. To evaluate the health impacts in each zone, we assume that each negative externality of a link is equally divided into two zones connected to that link. As we are interested in the macro level impacts due to the change in the transportation network parameters such as link capacity, this assumption simplifies our calculation and caters the purpose of our study. Under the preceding assumption, the reduction in health impacts, $\Delta z_{i,\tau}^{h}$, due to the change in externality *h* in zone *i* during period τ is calculated by

$$\Delta z_{i,\tau}^{h} = \frac{1}{2} \sum_{a \in A_i^+ \cup A_i^-} \left(e_{a,\tau}^{h,before} - e_{a,\tau}^{h,after} \right), \forall i, \tau, h, \tag{28}$$

where $e_{a,\tau}^{h,before} \left(e_{a,\tau}^{h,after} \right)$ represents the negative externality *h* before (after) the network expansion due to all modes traveling on link a during period τ.

Decrease in health cost. To aggregate the different impacts on the health condition, all of the above externalities are converted and combined into one monetary value referred to as the health (impact) cost. In this paper, the method is similar to the one in BenMAP

(Environmental Benefit Mapping and Analysis Program) proposed and developed to estimate the health effects of a change in air quality via a "damage-function" approach. It was used by Lee et al. (2012) to evaluate the clean truck program and Davidson et al. (2007) to analyze the PM 2.5 control policy. This paper extends the BenMAP method to estimate a more general health cost incorporating the effects of traffic noise and accident injuries under the assumption that the health cost of traffic noise energy and accident injuries can be estimated via a similar "damage–function" as the one used in BenMAP for traffic emissions. Besides, we extend the BenMAP model to a time-dependent one, which can be represented as

$$\Delta H_{i.\tau} = \sum_{h=1}^{n+2} \Delta z_{i,\tau}^{h} \times \beta_{i,\tau}^{h} \times I_{i,\tau}^{h} \times VSL_{i,\tau}^{h} \times R_{i,\tau}, \forall i, \tau , \tag{29}$$

where $\Delta H_{i,\tau}$ is the reduction in the total health cost for residential zone i for time period τ. $\beta_{i,\tau}^{h}$ is the health impact multiplier of factor h in zone i during the representative hour of period τ, which estimates the percentage change in an adverse health effect for a unit change in the noise energy, each type of emissions, or the number of accidents. The change is defined as the difference between a baseline (i.e., starting) condition and a new condition. In practice, $\beta_{i,\tau}^{h}$ can be interpreted as a zone feature. For example, if more trees grow in one zone, then the impacts of traffic emissions and noise are fewer. If better hospital emergency services are provided, then the probability of losing life in traffic accidents is lower. $I_{i,\tau}^{h}$ is the health incidence due to factor h in zone i during period τ, which estimates the average number of adverse health effects (e.g., average number of people died or got aspiration illnesses due to factor h) per person in that zone in the concerned period. $I_{i,\tau}^{h}$ relates to the characteristics of the affected residential population. For example, if these people have a higher average age, then the health cost is higher, because they are more vulnerable to adverse health impacts; even if the total emissions is low. $VSL_{i,\tau}^{h}$ represents the value of statistical life. It is the monetary value that people are willing to pay to slightly reduce the risk of health impacts. It is also varied by time period and zone. $R_{i,\tau}$ defines the exposed population associated with zone *i* during period *τ.* Essentially, the method converts each health impact into the monetary value through the value of statistical life.

The first advantage of the above model is that it explicitly considers the affected population suffering those negative impacts directly. Second, the above framework is general and can be applied to various countries by calibrating different parameters, such as $\beta^h_{i,\tau}$ and $I^h_{i,\tau}$. To apply the above framework, it is necessary to synchronize unit with the project period considered, as it is noticed that most of the studies on "damage–function" are on an annual basis (Ballester, 2005). Thus, all above measurements adopted are converted to the period basis via multiplying a conversion parameter λ^h as shown in the following equation:

$$\Delta H = \sum_i \sum_\tau \frac{\lambda^h \Delta H_{i,\tau}}{(1+\tilde{i})^{(\tau-1)}}. \tag{30}$$

Equation (30) aggregates the health costs over all periods and computes the present value of the change in the health cost, *ΔH*. Similar to consumer surplus, if the change in the health cost is positive, then the health condition is improved as the health cost after network expansion is less than that before network expansion, and vice versa. Hence, this change can be interpreted as the reduction in the health cost and the larger the value is, the better the health condition is.

2.4.3 Design and financial constraints

The upper level problem also includes constraints. There are two types: capacity and budget. The capacity constraints are included to address the fact that a link (in road networks) cannot be built or expanded beyond an upper limit due to space limitation, and the budget constraint is included because the total expansion cost is bounded by the total budget available.

Mathematically, they can be, respectively, represented as

$$\bar{c}_{a,\tau} \le u_a, \forall a, \tau \text{, and} \tag{31}$$

$$\sum_{\tau}\sum_{a} g_{a,\tau}(y_{a,\tau}) \le B. \tag{32}$$

Equation (31) is the maximum allowable capacity constraint, which is to limit the total capacity of each link a after road expansion or highway construction at the beginning of period τ, $\bar{c}_{a,\tau}$, to be less than its maximum allowable capacity. Equation (32) states the total expenditure cannot exceed total budget B. $g_{a,\tau}(y_{a,\tau})$ is a construction cost function for link a during period τ, which is generally assumed to be nonnegative, increasing, and differentiable (Yang and Bell, 1998). Here, a linear relationship is adopted, which is

$$g_{a,\tau}(y_{a,\tau}) = U_{a,\tau} y_{a,\tau}, \forall a, \tau. \tag{33}$$

3. Solution method: hybrid artificial bee colony algorithm

The proposed problem involves integer decision variables, two levels, and nonlinear constraints. Hence, the proposed problem is *NP*-hard and extremely difficult to solve. Various meta-heuristic methods have been proposed in the literature to solve *NP*-hard problems. In particular, ABC has been recently used for solving transit network design problems and shown to be better than genetic algorithm (Szeto and Jiang, 2012, 2014). In general, ABC belongs to a class of evolutionary algorithms, which are inspired by the intelligent behavior of honeybees in finding nectar sources around their hives (Karaboga and Basturk, 2008). In this paper, a newly developed hybrid Artificial Bee Colony algorithm (ABC) is adopted to solve the problem, where ABC is used to solve the network design problem in the upper level, while the MSA and the Frank-Wolfe method are used to solve the lower level problem. The details of solving the lower level problem can be found in Szeto et al. (2014). For the proof of convergence of the MSA, the readers can refer to Daganzo (1983) or Theorem 3 in Cantarella (1997). In the following, we depict the method for solving the upper level problem in details.

3.1 Artificial bee colony algorithm: overall algorithmic step

In ABC, a bee colony is divided into three types: employed bees, onlookers and scouts. Employed bees are responsible for exploiting available food sources (i.e., solutions) and gathering required information. They also share the information with the onlookers that select existing food sources to be further explored. When the food source is abandoned by its associated employed bee, the employed bee becomes a scout and starts to search for a new food source in the vicinity of the hive. The abandonment happens when the quality of the food source is not improved after performing a maximum allowable number of iterations called *limit*. The overall algorithmic step is given below.

Step 0. Generate initial solutions and assign each of them to one employed bee; Set iteration = 0.

Step 1. Iteration = iteration + 1.

Step 2. Employed bee phase.

For i = 1 to number of employed bees

Step 2.1. Employed bee i generates a neighbor solution.

Step 2.2. Repair the neighbor solution, if necessary.

Step 2.3. Solve the lower level problem and evaluate the fitness value of the neighbor solution.

Step 2.4. Replace the solution associated with employed bee i by its neighbor if the latter is better. Otherwise, increase the limit counter of the solution associated with employed bee i by 1.

Next

Step 3. Onlooker phase.

For i = 1 to number of onlookers

Step 3.1. Apply the roulette selection method to assign onlooker i to one of the solutions associated with employed bees and generates its neighbor solution.

Step 3.2. Repair the neighbor solution, if necessary.

Step 3.3. Solve the lower level problem and evaluate the fitness of the neighbor solution.

Step 3.4. Replace the solution associated with employed bee i by its neighbor if the latter is better. Otherwise, increase the limit counter of the solution associated with employed bee i by 1.

Next

Step 4. Scout phase.

Step 4.1. Identify the abandoned solutions, whose limit counters are greater than limit.

Step 4.2. Generate new solutions and replace the abandoned ones.

Step 5. If iteration is less than a predefined maximum value, return to Step 1, otherwise, stop and output best solutions.

3.2 Solution representation

Figure 1 illustrates the representation scheme. Each solution chromosome is divided into N_a parts, where N_a equals the number of candidate links, while the number of genes in each part equals the number of periods. $y_{a,\tau}, \tau = 1, ..., T$ represents the capacity expansion during period τ for link *a*.

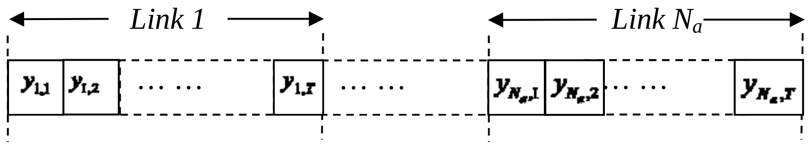


Figure 1 Solution Representation Scheme

3.3 Solution generation

To generate initial solutions in ABC, a link-based generating procedure is proposed. Initially, all the genes in the solution chromosome are set to be zero. Then, an expansion plan is determined link by link. For each candidate link, a random integer number between 1 and $T+1$ is produced to decide which year the expansion is initiated where $T+1$ means no expansion. Afterwards, the link expansion plan is obtained period by period based on the remaining allowable link capacity until the maximum allowable capacity or total budget is reached. Take link *a* for instance. If an integer number T' $(1 \le T' \le T)$ is generated, then the expansion plans from periods T' to T are determined; for each period τ $(T' \le \tau \le T)$, the capacity improvement is generated between 0 and $\min\left(u_a - \sum_{t=T'}^{\tau} y_{a,t}, \frac{B - \sum_{a'} \sum_{t=1}^{\tau} y_{a',t} U_{a',t}}{U_{a,\tau}} \right)$,

where $u_a - \sum_{t=T'}^{\tau} y_{a,t}$ is the remaining allowable capacity and $B - \sum_{a'} \sum_{t=1}^{\tau} y_{a',t}$ is the remaining budget. After designing one link, another candidate link is selected and expanded if there is still budget left. Following this procedure, the capacity constraints are satisfied automatically for the initial solutions.

The neighbor solutions are generated as follows: first, a candidate link is randomly selected; then its expansion plan is redesigned following the same procedure used in the initial solution generation except that the total budget constraint is ignored in order to allow more mutation.

3.4 Solution repairing

Since the neighborhood operation only incorporates the link capacity constraint, the total budget constraint is likely to be violated by the generated neighbor solutions. To remedy these infeasible solutions, a simple repair procedure is developed as follows:

Step 1. Calculate the total cost of each neighbor solution generated.

Step 2. Compare the total cost and the budget, if the cost is greater than the budget, proceed to next step; otherwise cease repairing;

Step 3. Select one expanded link randomly and set its capacity expansion to be zero one by one following the reverse chorological order until the total cost is not greater than the total budget or all of the expansions on the selected link is zero;

Step 4. Return to Step 2.

3.5 Fitness evaluation

The proposed fitness function is defined by

$$f = \Delta CS + \Delta H + C , \tag{34}$$

where C is a large constant to avoid generating negative fitness values. Note that there is no penalty term because the design and financial constraints are captured in the solution generation and repairing procedures.

3.6 Termination and convergence

For the convergence of (hybrid) ABC, it is widely acknowledged that metaheuristics including (hybrid) ABC may not guarantee to obtain a global optimum. In practice, the (hybrid) ABC is stopped after a predetermined number of iterations is reached or the best solution in the population has not been changed after a predetermined number of iterations.

3.7 Computation complexity

The computation complexity of the proposed algorithm in terms of the number of lower level problems solved is equal to $\left(N_e + N_o\right)C_{\max} + N_e\left\lfloor C_{\max} / \text{limit} \right\rfloor$, where N_e is the number of employed bees; N_o is the number of onlookers; $C_{\max}$ is the maximum number of iterations. It is because in each of the $C_{\max}$ iterations, the algorithm solves a total of $\left(N_e + N_o\right)$ lower level problems. Moreover, in the worst case, each solution does not change over limit iterations and is replaced by a new solution generated for every limit iterations. This occurs at most $\left\lfloor C_{\max} / \text{limit} \right\rfloor$ times.

4. Numerical studies

Two examples are set up to illustrate the properties of the model and the performance of the solution algorithm. The first example uses a simple network whereas the second example uses the Sioux Falls network. Both examples consider two modes: private vehicle and metro. In these numerical examples, congestion interaction between different modes is not our major focus. Therefore, we did not consider the interaction. Moreover, the health effects of exhausts, noise, and accidents due to metro are assumed to be negligible.

4.1 Example 1: A simple network

The simple network adopted is shown in Figure 2. There are 3 links in this network: links 1-3. Links 1 and 2 are links whose travel times are given by the BPR function. Link 3 is a separated metro link, as represented by a dashed line in the figure. There are 3 zones: E1, R2, and R3, in which ‘E’ stands for an employment zone whereas ‘R’ stands for a residential zone. The three zones form two OD pairs: E1-R2 and E1-R3. Both OD pairs are connected by highways but only OD pair E1-R2 has a segregated metro connection. In other words, there are two modes for OD pair E1-R2 but there is only one mode for OD pair E1-R3. It is proposed that link 2 is widened. For notation purposes, mode 1 represents the private vehicle mode and mode 2 represents the metro mode. For the private vehicle mode, two vehicle types are considered when computing the CO emissions. One is petrol private cars with their weight less than 3.5 tones, denoted as mode 1a, and the other is diesel private cars with their weight between 2.5-3.5 tones, denoted as mode 1b.The ratio of the two types of vehicles on each link is assumed to one to one..

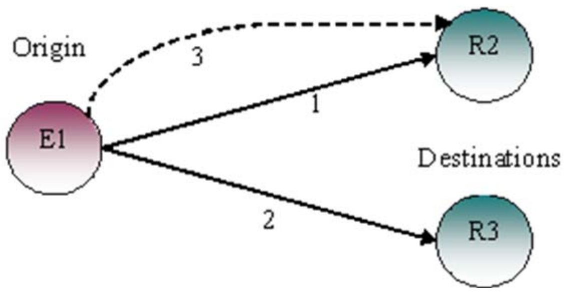


Figure 2 The small network

The parameters, unless otherwise specified, in this study are set to the following values:

1) Parameters in Lowry-type constraints:

$E_{1,1}^{B} = 5000$ jobs, $W_{1,1} =$ 3000 jobs, $W_{2,1} = W_{3,1} = 3000$ houses, $\tilde{W}_{2,1} = \tilde{W}_{3,1} = 3000$ houses,

$\mu = 5$, s = 0.1, $\beta^{r} = 0.04€^{-1}$, $\beta^{s} = 0.03€^{-1}$, $g_{E,1} = 0.04$, $\tilde{g}'_{w,1} = \tilde{g}'_{w,2} = \tilde{g}'_{w,3} = 0.05$

$\tilde{g}_{w,1} = \tilde{g}_{w,2} = \tilde{g}_{w,3} = 0.05$;

2) Parameters in modal-split/assignment constraints:

$\bar{c}_{1}^{0} = \bar{c}_{2}^{0} = 3000$ vph , $t_{1}^{0,1} = t_{2}^{0,1} = 1$ hour , $t_{3}^{0} = \dfrac{2}{3}$ hour , $\psi = €15/\text{h}$, $\theta^{1} = €16$, $\theta^{2} = €24$,

$\bar{\beta} = 0.05€^{-1}$, the transit fare on link 3 is €5;

3) Parameters in health cost calculation:

For computing the CO emissions, we set $k^{1a,1} = 1.0$, $b_{1}^{1a,1} = 2.2627\text{E}+1$, $b_{2}^{1a,1} = \text{-6.8548E-1}$, $b_{3}^{1a,1} = \text{-1.4443E-2}$, $b_{4}^{1a,1} = b_{5}^{1a,1} = b_{6}^{1a,1} = 0.0$, $k^{1b,1} = 0.633$, $b_{1}^{1b,1} = 1.4148\text{E+1}$, $b_{2}^{1b,1} = 1.1423$, $b_{3}^{1b,1} = \text{-4.3263E-3}$, and $b_{4}^{1b,1} = b_{5}^{1b,1} = b_{6}^{1b,1} = 0.0$.

For computing the number of accidents, it is assumed that the initial number of accidents on a link equals 3% of the link flow on that link, and P = 2 for the power model;

For computing the decrease in the health cost in each period τ, we set $\beta^1_{2,\tau} = 0.2\%$, $I^1_{2,\tau} = 0.2\%$, $\beta^2_{2,\tau} = 0.1\%$, $I^2_{2,\tau} = 0.1\%$, $\beta^3_{2,\tau} = 2.0\%$, and $I^3_{2,\tau} = 4.0\%$ for zone R2 and $\beta^1_{3,\tau} = 0.2\%$, $I^1_{3,\tau} = 0.1\%$, $\beta^2_{3,\tau} = 0.1\%$, $I^2_{3,\tau} = 0.2\%$, $\beta^3_{3,\tau} = 1.0\%$, and $I^3_{3,\tau} = 1.0\%$ for zone R3. The value of statistical life $VSL^h_{i,\tau} = €1.0e+6, \forall i, h, \tau$.

4) Interest rate: $\tilde{i} = 0.003$

Unless specified, we restrict that the capacity improvement is finished in the first period of link 2 and increase the capacity from 0 to 7500 vph. The lower level model was solved for each capacity setting to obtain results.

4.1.1 Tradeoff between ΔCS and ΔH

This scenario is created to illustrate the conflict between the change in consumer surplus, *ΔCS*, and the change in the health cost, *ΔH*. Figure 3 plots *ΔCS* and *ΔH* as a function of the improvement in the capacity of link 2. It is observed that *ΔCS* increases with the expansion of link 2, demonstrating that travelers are benefited from the network improvement, as expected. In contrast, the health condition is becoming worst.

The changes in the health cost can be explained by the following reasons. Once link 2 is expanded, the composited travel cost reduces; hence, more residents are attracted to zone R3. As a result, the flow on link 2 increases, and generates more traffic emissions and induces more traffic accidents. In contrast, these externalities reduce on link 1 (It should be reminded that those changes are only applicable to the proposed network and particular parameter setting.). After considering the affected population, the positive change in the health cost in zone R2 is less than the negative change in zone R3 Thus, the overall health condition deteriorates. Besides, with the drop in the composite travel cost, more people select to live in zone R3, magnifying the change in the health cost. The occurrence of such spatial redistribution is because of the ignorance of the health cost when people make decisions on living and working locations, travel mode, and route choices. Consequently, there is no reason to expect the resultant flow pattern and residential distribution can reduce the health cost given that people only consider their own generalized travel cost and accessibility.

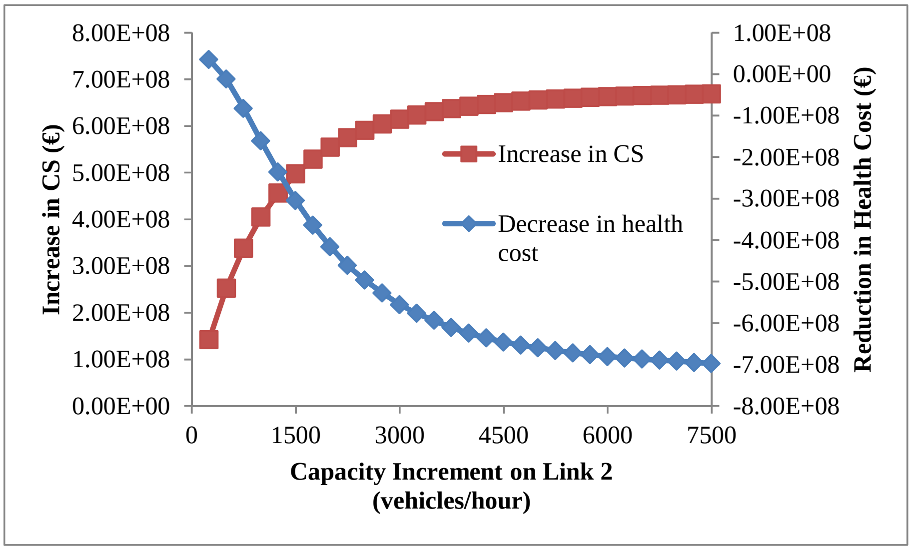


Figure 3 Conflict between ΔH and ΔCS

4.1.2 The effect of capacity expansion on the increase in social benefit

Figure 4 plots the objective values under different capacity expansion plans. When the capacity increment is 750 vph, the maximum objective value is achieved. Afterwards, the objective value reduces and falls below zero when the expansion is 4750 vph. This example not only shows the existence of the best strategy of the problem but also implies that if the expansion was wrongly designed (i.e., greater than 4750 vph), then the overall increase in social benefit, as a sum of the increase in consumer surplus and the reduction in the health cost, is worse than that of the do-nothing scenario.

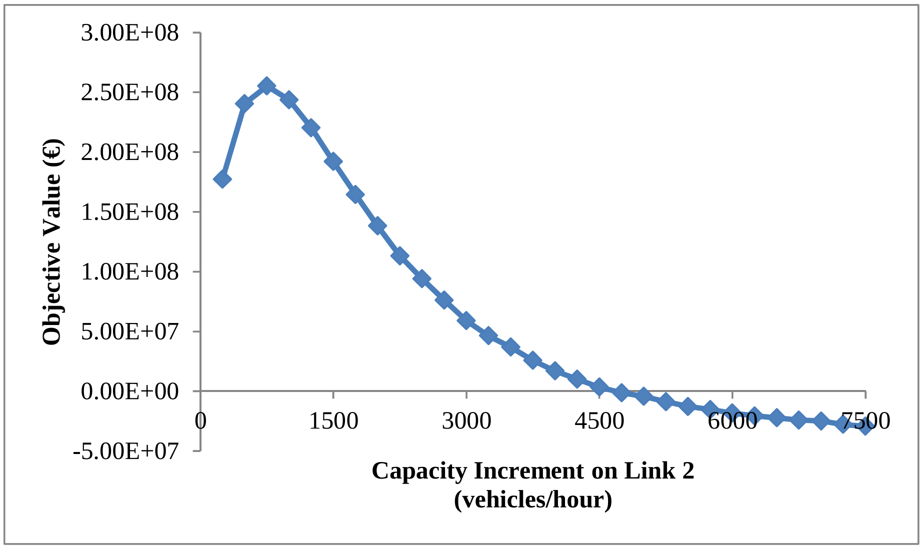


Figure 4 Objective values under different capacity expansion scenarios for link 2

4.1.3 A paradox on health cost with respect to emissions

This scenario aims to reveal a paradox that the health cost is reduced with the increment in the total amount of CO emissions. Figure 5 plots the reduction in the health cost and the reduction in the total amount of CO emissions as a result of capacity expansion. Clearly, the graph shows two different trends. The change in the health cost is positive and increasing; in other words, the health condition is better than the do-nothing scenario. In contrast, the reduction in the total amount of CO emissions is positive but decreasing.

This figure reveals a paradox that the health condition of the residents improves, as reflected by the changes in the health cost, with an increasing in the total amount of traffic emissions. Such paradox is because network expansion induces a change in not only the total traffic emissions, but also the spatial distributions of residents. More people move to a zone with less pollution and the zone with less pollution has more population than the zone with heavier pollution, because such movement roots in the consideration of the travel cost and accessibility. As a result, the total health cost decreases, despite the increment in the total CO emissions due to the change in speed. This case also explicitly demonstrates the importance of incorporating the affected population into the health cost calculation. Without considering the affected population, the health impacts can be wrongly estimated and the occurrence of the paradox may not be discovered. Although this example only considers CO for illustrating purposes, there is no conceptual difficulty to generalize the paradox conclusion to other health impacts because the root of the paradox is the neglect of affected population and the ignorance of the health impact when people make decisions on residential locations. Practically, it is suggested that the decision maker should consider the residents in influenced zones when evaluating the health impacts and costs of network expansion projects.

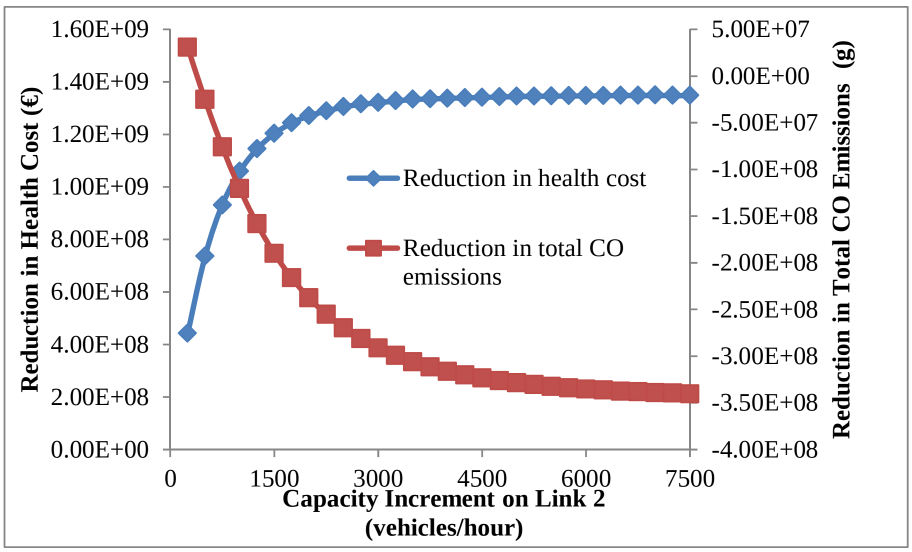


Figure 5 Comparing the health cost and CO emissions.

4.1.4 Health inequity issue

The proposed health cost function also helps demonstrate the issue of health inequity. Such issue is depicted in Figure 6, which separately plots the reductions in the health costs for zones R2 and R3. Obviously, it shows that after the capacity improvement, the health cost of residents in zone R3 deteriorates, while that in zone R2 ameliorates. The residents in zones R2 and R3 have different health conditions, which raises the health inequity issue between zones.

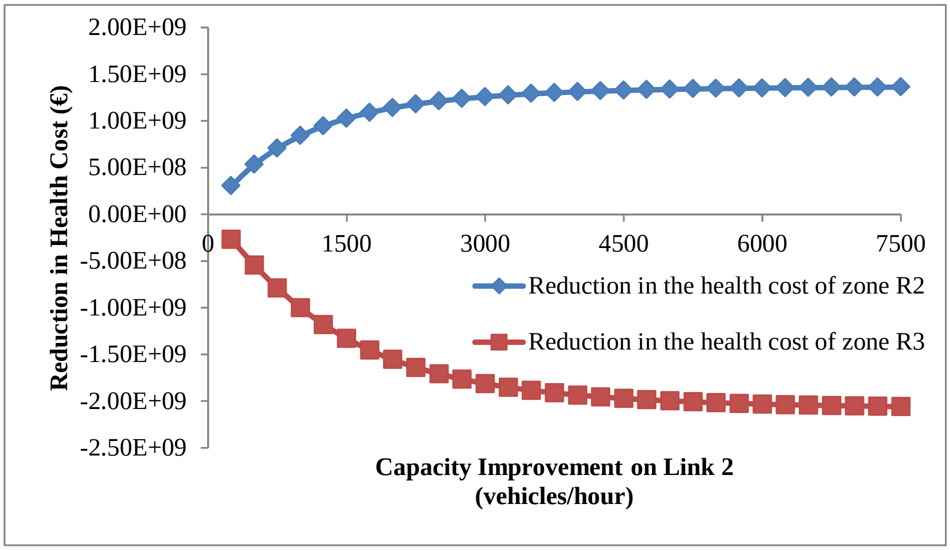


Figure 6 Reduction in the health cost for different zones

4.1.5 Effect of the health parameter on the objective value

To illustrate the effect of the health parameters on the optimal objective value and design, a sensitivity analysis on $\beta_{3,\tau}^{1}$ was conducted. $\beta_{3,\tau}^{1}$ was varied from 0.1% to 0.3%, representing the scenario that residents are more vulnerable to vehicle exhausts. The results are plotted in Figure 7. As expected, the larger the value of $\beta_{3,\tau}^{1}$ is, the more the health cost is, and the lower the optimal objective value (as reflected by the peak in the curve) is. Meanwhile, the value of $\beta_{3,\tau}^{1}$ affects the optimal expansion design. For example, the optimal expansion is 750 vph when $\beta_{3,\tau}^{1} = 0.2\%$. When $\beta_{3,\tau}^{1} = 0.1\%$, the optimal expansion is 7500 vph. This example clearly illustrates the importance of calibrating the values of parameters accurately.

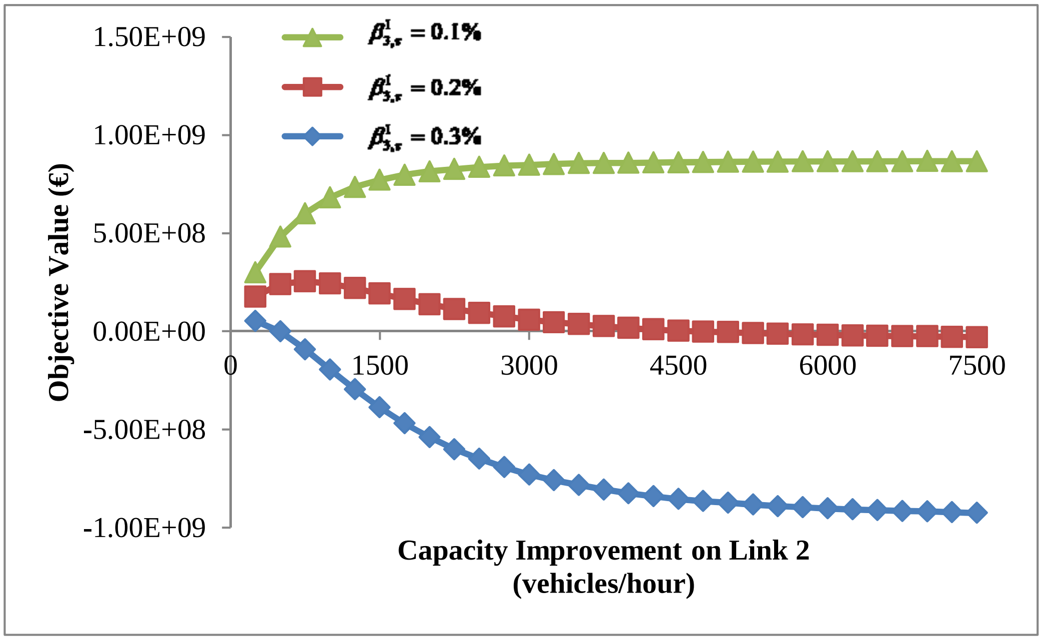


Figure 7 Effect of $\beta^{1}_{3,\tau}$

4.1.6 Pareto frontier

Figure 8 plots the Pareto frontier for the small network example. In this example, both links 2 and 3 were expanded from 0 to 7500 vph with a fixed increment of 250 vph. Thus, a total of 961 expansion combinations was evaluated. The Pareto frontier was obtained by the non-dominated sorting genetic algorithm proposed by Deb et al. (2002). In general, the graph shows the conflict between the two objectives when both links 2 and 3 are expanded. Meanwhile, it also shows that the Pareto frontier, under the parameter setting, is not convex due to the complexity of the health impact functions

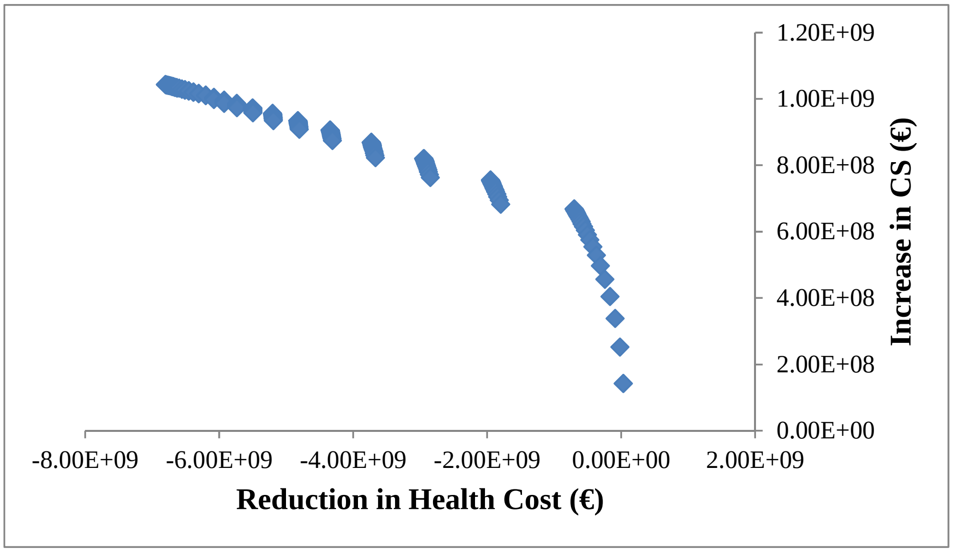


Figure 8 Pareto frontier

4.2 Example 2: Sioux Falls network

The Sioux Falls network is used in this example, but to cater the multimodal feature, a metro network is created on the Sioux Falls network and marked by the thick yellow lines as shown in Figure 9. The blue circle is the proposed developed area, where eight candidate links are included. For simplicity, the attractiveness of each zone are assumed to be equal for each zone and set as follows: $W_{j,1} = 1.0$, $\tilde{W}_{i,1} = 1.0$, $\beta^r = 0.02$, $\beta^s = 0.01$, for each period τ and zone i. The cost for constructing a lane with capacity 1500 vph is €5000 per kilometer, and the total budget is €300000. The other parameters are the same as those in example 1, except $\mu = 3.0$. For the link performance function, we set $\gamma_{a,m} = 4$ and the travel time and initial capacity setting were obtained from the website (Bar-Gera, 2014).

The parameters used in ABC are set as the following values: The colony size is set to 20, composing 10 employed bees and 10 onlookers. The stopping criterion of MSA is that the MSA stops when the error is less than 0.001. The maximum number of iterations in each MSA is 500. The program was coded with Fortran and run on a desktop with an Intel (R) Core (TM) i7-3770 CPU @ 3.4 GHz.

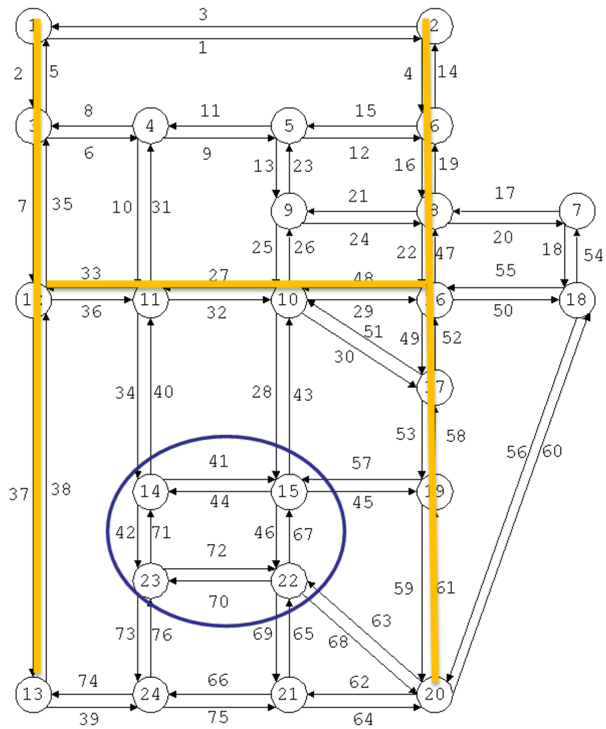


Figure 9 Modified Sioux-Falls Network with transit lines

4.2.1 Performance of hybrid ABC

The performance of the hybrid ABC algorithm is shown in Figure 10, from which, we can see that the objective value converges in a stepwise manner. Such convergence feature is also expected to occur in large network applications, because the objective value of the hybrid ABC algorithm remains stable until a better neighbor solution is found by the onlookers or the employed bees, or a new promising solution space is discovered by the scouts. For the Sioux Falls network scenario, the algorithm took 1262.9 seconds to conduct 300 major ABC iterations on average. Such a long computation time is mainly due to the computation time required to solve the lower level problem by the MSA and the Frank-Wolfe algorithm. To apply the algorithm in large networks, faster algorithms for traffic and transit assignment, such as origin-based algorithm (Bar-Gera, 2002) or approach-based algorithm (Szeto and Jiang, 2014), can be used and self-regulated averaging method can be incorporated in the future study. Moreover, the parallel computing technique can be introduced for large network applications.

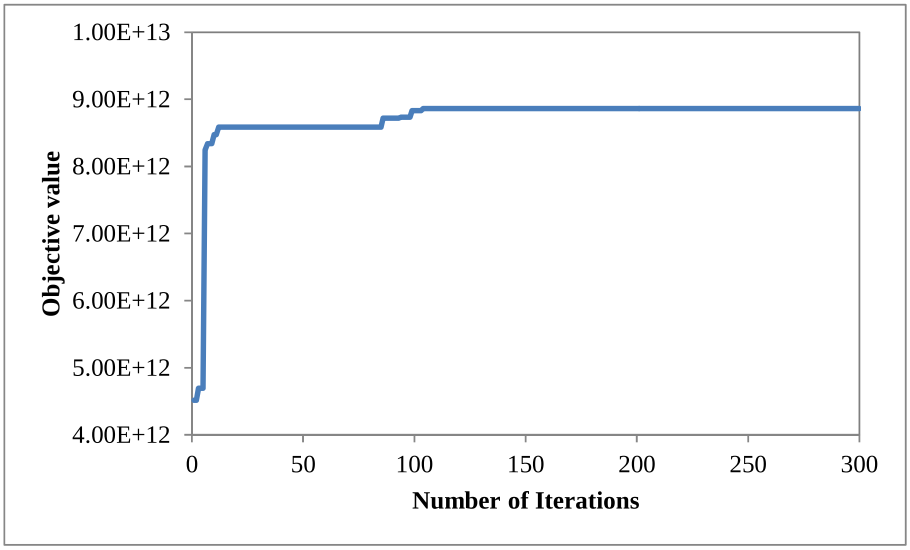


Figure 10 Convergence of hybrid ABC

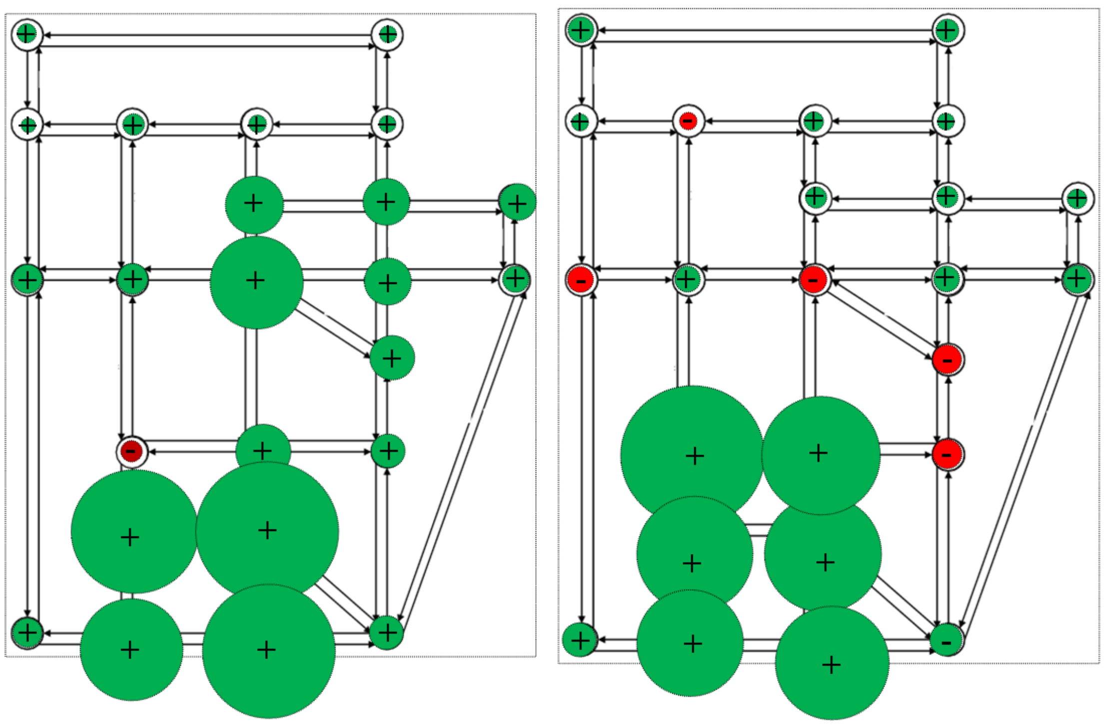
To test the effectiveness of the proposed repairing procedure, an ABC version without the repairing procedure was also coded. Other than removing the repairing procedure, the fitness function was adjusted by adding one penalty term to penalize the violation of the financial constraint. However, adding penalty value could also result in a large fitness value. To address such issue, in the employed bee and onlooker phases, an additional step was carried out to ensure that only a feasible neighbour solution was allowed to replace the associated solution. For each version of ABC, 20 runs with different initial seeds for generating random numbers were carried out. The computation results are presented in Table 1. The second and third columns are the results from the version without and with the repairing procedure, respectively. Two measures are compared. One is the average objective value of 20 runs, and the other is best objective value of 20 runs. The value in each pair of braces is the percentage increment with respect to the value of the same row. The table shows that the proposed repairing procedure improves the performance of the algorithm significantly. By adopting the repairing procedure, the average objective value is improved by 54.07%. Meanwhile, the best objective value of 20 runs increases by 13.08%. This is because the infeasible solutions are remedied and hence, more feasible solutions are explored and evaluated. A *t*-test was also conducted to examine whether the difference in the average objective value is statistically significant. The *t*-value obtained is 16.04 and hence we can conclude that the difference is significant at the 5% level and the average performance of the ABC algorithm with the repairing operator is better at this significance level

Table 1 Effectiveness of the proposed repairing procedure

	Without repairing procedure	With repairing procedure
Average objective value of 20 runs	5.656E+12 -	8.714E+12 (54.07%)
Best objective value of 20 runs	7.837E+12 -	8.862E+12 (13.08%)

4.2.2 Spatial inequity

Figure 10 plots ΔCS and ΔH at each node at the final period between the scenario with expansion and that without. The increment in CS of each node *i* was calculated by summing up the change in CS of each OD pair over destination, time, and mode, discounted to present value terms ($\sum_{\tau}\sum_{j}\sum_{k}\frac{\Delta CS_{ij,\tau}^{k}}{\left(1+\tilde{i}\right)^{\tau-1}}$). Similar to Figure 9, the green circle with "+" sign indicates an improvement in *CS* or health cost, while the red one with "-" represents the deterioration; the radius of the circle represents the change in magnitude. Figure 10 shows that the health cost and consumer surplus can be improved simultaneously for some nodes, such as nodes 21, 22, 23, and 24. Meanwhile, the conflict between CS and health cost is also observed. For example, at nodes 4 and 10, their own CS increases while the corresponding health cost decreases. By comparing the size of the circles, it is concluded that the residents live within the improvement area face more significant changes in CS and the health cost. The further the zone is away from the improvement area, the less the influence of the improvement is. In addition, this figure demonstrates the existence of spatial inequity at the node level. Despite the fact that most of the residents are benefitted from network improvement, some residents are suffered either in terms of health or consumer surplus. To handle these spatial inequity issues, we can add suitable inequity indicators or constraints in our formulation without conceptual difficulty.



(a) ΔCS at each node **(b) ΔH at each node**

Figure 11 ΔCS and ΔH at each node

4.2.3. Best solution obtained

The best solution obtained from the hybrid ABC is presented in Table 2. This table shows that not all the links are expanded due to the budget constraint. For the four links selected to improve, two of them reach the maximum capacity of 7500 vph, implying that constraint (31) is binding for these three links. Interestingly, it is observed that not all of the expansions are carried out in the first period. In particular, link 46 is even expanded in the last period. Such phenomenon is attributed to the conflict between the objectives of travelers and residents. Commuters' prefer to increase the capacity in the early period to gain more consumer surplus. In contrast, residents intend to oppose the expansion or delay the expansion due the reduction in the health condition. Consequently, the expansion could be implemented in the middle or even at end of the planning horizon. These results reveal that in a time-dependent transportation network design framework, the conflict between the different objectives affects the timing of link expansion, and the transportation network improvements may not be performed in each design period. The results also illustrate the importance of incorporating the time dimension into the transportation network design framework.

Table 2 Best solution obtained from hybrid ABC

	Link Expansion (vph)							
	Link 41	Link 42	Link 44	Link 46	Link 67	link 70	Link 71	Link 72
$\tau = 1$	0	0	1500	3000	1500	6000	0	0
$\tau = 2$	0	0	0	3000	0	1500	0	0
$\tau = 3$	0	0	0	0	0	0	0	0
$\tau = 4$	0	0	0	0	0	0	0	0
$\tau = 5$	0	0	0	1500	0	0	0	0

5. Discussion

Our proposed model has many parameters. Their values may highly affect the ultimate design as illustrated in one of the examples in Section 4.1.5 in this paper. Hence, their values should be carefully calibrated using real data and the maximum likelihood method, and the resultant model should be validated by another set of data before the model is applied for practical applications.

A sensitivity analysis for each parameter should also be performed to determine the effects of parameter values on the design and the objective value and identify critical parameters to the design. This step can guide us to determine the allocation of limited resource to collect, very often time consuming and expensive, data for calibrating the values of parameters. For critical parameters, more resources should be allocated to collect the corresponding data to improve the accuracy of the model prediction. If necessary, we need to calibrate and validate the model again after performing the sensitivity analysis and collecting more data for calibrating critical parameter values.

Since the critical parameters and the sensitivity of the parameters depend on the scenario setting, we cannot conclude which parameters are critical in general. However, they should be found by conducting sensitivity analyses in practical applications, because the consequence of wrong estimation on those parameters may result in a very suboptimal design as illustrated in Section 4.1.5 and an overestimation of the benefit of the design.

6. Concluding remarks

This paper develops a bilevel time-dependent road network design model to consider the health impacts and consumer surplus. A general health cost function is proposed and used in the model to capture the health impacts of traffic emissions, noise and accidents due to network expansion projects. Those impacts are formulated as flow or speed dependent variables. Hence, the effects of the changes in human health due to network improvements can be captured. In addition, the affected population is considered in the health cost. To solve the bilevel problem, a hybrid artificial bee colony algorithm is proposed to search the upper-level network design solutions, while the MSA and the Frank-Wolfe algorithm is adopted to handle the lower-level transportation land-use interaction problem. Numerical studies are also set up to illustrate the performance of the hybrid ABC, the effectiveness of the proposed repairing procedure, the conflict between the health cost objective and consumer surplus objective, a paradox on the health cost with respect to traffic emissions, the importance of incorporating the affected population into the health cost calculation, and the importance of incorporating time dimension into the transportation network design framework. More importantly, we found that although capacity expansion is expected to increase the overall consumer surplus, the health costs among different zones varies. This raises the issue of health inequity.

In the future, the proposed framework can be extended to consider uncertainties (e.g., Ng et al., 2011; Szeto et al., 2011a,b), other types of dynamics (e.g., Lo and Szeto, 2002, 2005; Watling and Cantarella, 2013; Szeto, 2013; Wong et al. 2014), and transit network design decisions (e.g., Szeto and Wu, 2011). The solution algorithm can also be improved for large network applications, based upon origin-based algorithm (Bar-Gera, 2002), approach-based algorithm (Szeto and Jiang, 2014), and self-regulated averaging method, and the parallel computing technique (Wong et al., 2000).

Acknowledgments

The research was jointly supported by grants (201011159026 and 201211159009) from the University Research Committee and Faculty of Engineering Top-up Grant of the University of Hong Kong, and a grant from National Natural Science Foundation of China (71271183). The authors are grateful to the four reviewers for their constructive comments.

References

Ballester, F., 2005. Air pollution and health: an overview with some case studies, Environmental Health Impacts of Transport and Mobility, edited by Nicolopoulou-Stamati, P., Hens, L., and Howard, C.V., Springer, Dordrecht.

Ban, J.X., Liu, H.X., Ferris, M.C., and Ran, B., 2006. A general MPCC model and its solution algorithm for continuous network design problem. *Mathematical and Computer Modelling*, 43(5-6), 493-505.

Bar-Gera, H. 2014. Transportation test problems. http://www.bgu.ac.il/~bargera/tntp/, accessed March 3, 2014.

Bar-Gera, H., 2002. Origin-based algorithm for the traffic assignment problem. *Transportation Science*, 36(4), 398-417.

Boulter, P.G., Barlow, T.J., and McCrae, I.S., 2009. Emission factors 2009: Report 3-exhaust emission factors for road vehicles in the United Kingdom. Version 6, Published Project Report PPR 356, TRL Limited.

Boyce, D.E. and Janson, B.N., 1980. A discrete transportation network design problem with combined trip distribution and assignment. *Transportation Research Part B*, 14(1-2), 147-154.

Cameron, M.H. and Elvik, R., 2010. Nilsson's Power Model connecting speed and road trauma: Applicability by road type and alternative models for urban roads. *Accident Analysis & Prevention*, 42(6), 1908-1915.

Cantarella, G.E., 1997. A general fixed-point approach to multimode multi-user equilibrium assignment with elastic demand. *Transportation Science*, 31 (2), 107–128.

Chen, A. and Yang, C., 2004. Stochastic transportation network design problem with spatial equity constraint. *Transportation Research Record: Journal of the Transportation Research Board*, 1882, 97-104.

Chen, A., Kim, J., Lee, S., and Kim, Y., 2010. Stochastic multi-objective models for network design problem. *Expert Systems with Applications: An International Journal*, 37(2), 1608-1619.

Chen, A., Kim, J., Zhou, Z., and Chootinan, P., 2007. Alpha reliable network design problem. *Transportation Research Record: Journal of the Transportation Research Board*, 2029, 49-57.

Chen, A., Subprasom, K., and Ji, Z., 2006. A Simulation-based Multi-Objective Genetic Algorithm (SMOGA) for build-operate-transfer network design problem. *Optimization and Engineering Journal*, 7(3), 225-247.

Chiou, S.W., 2009. A subgradient optimization model for continuous road network design problem. *Applied Mathematical Modelling*, 33(3), 1386-1396.

Daganzo, C.F., 1983. Stochastic network equilibrium with multiple vehicle types and asymmetric, indefinite link cost Jacobians. *Transportation Science*, 17 (3), 282–300.

Davidson, K., Hallberg, A., McCubbin, D., and Hubbell, B., 2007. Analysis of PM2.5 using the environmental benefits mapping and analysis program (BenMAP). *Journal of Toxicology and Environmental Health Part A*, 70(3-4), 332-346.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.A.M.T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evolutionary Computation, IEEE Transactions on*, 6(2), 182-197.

Farahani, R.Z., Miandoabchi, E., Szeto, W.Y., and Rashidi, H. 2013. A review of urban transportation network design problems. *European Journal of Operational Research*, 229(2), 281-302.

Fehr, R, 1999. Environmental health impact assessment: Evaluation of a ten-step model. *Epidemiology*, 10(5), 618-625.

Friesz, T.L., Anandalingam, G., Mehta, N.J., Nam, K., Shah, S.J., and Tobin, R.L., 1993. The multiobjective equilibrium network design problem revisited: A simulated annealing approach. *European Journal of Operational Research*, 65(1), 44-57.

Ge, Y.E., Zhang, H.M., and Lam, W.H.K. 2003. Network reserve capacity under influence of traveler information. *Journal of Transportation Engineering-ASCE*, 129(3), 262-270.

Hakim, S., Shefer, D., Hakkert, A.S., and Hocherman, I., 1991. A critical review of macro models for road accidents. *Accident Analysis & Prevention*, *23*(5), 379-400.

Karaboga, D. and Basturk, B., 2008. On the Performance of Artificial Bee Colony (ABC) Algorithm. *Applied Soft Computing*, 8(1), 687-697.

Kim, B.J., and Kim, W., 2006. An equilibrium network design model with a social cost function for multimodal networks. *The Annals of Regional Science*, 40(3), 473-491.

Kim, B.J., Kim, W.K., and Song, B.H., 2008. Sequencing and scheduling highway network expansion using a discrete network design model. *The Annals of Regional Science*, 42(3), 621-642.

LeBlanc, L.J., 1975. An algorithm for discrete network design problem. *Transportation Science*, 9(3), 183-199.

Lee, G., You, S., Ritchie, S.G., Saphores, J.D., Jayakrishnan, R., and Ogunseitan, O., 2012. Assessing air quality and health benefits of the Clean Truck Program in the Alameda corridor, CA. *Transportation Research Part A*, 46(8), 1177-1193.

Li, Z.C., Lam, W.H.K., and Wong, S.C. 2012a. Modeling intermodal equilibrium for bimodal transportation system design problems in a linear monocentric city. *Transportation Research Part B*, 46(1), 30-49.

Li, Z.C., Lam, W.H.K., Wong, S.C., and Sumalee, A., 2012b. Environmentally sustainable toll design for congested road networks with uncertain demand. *International Journal of Sustainable Transportation*, 6(3), 127-155.

Li, Z.C., Li, Z.K., and Lam, W.H.K., 2014. An integrated design of sustainable land use and transportation system with uncertainty in future population. *Transportmetrica A*, 10(2), 160-185.

Lin, J.J. and Feng, C.M., 2003. A bi-level programming model for the land use – network design problem. *The Annals of Regional Science*, 37(1), 93-105.

Lo, H.K. and Szeto, W.Y., 2002. A cell-based dynamic traffic assignment model: formulation and properties. *Mathematical and Computer Modelling*, 35(7-8), 849-865.

Lo, H.K. and Szeto, W.Y., 2009. Time-dependent transport network design under cost-recovery. *Transportation Research Part B*, 43(1), 142-158.

Lo, H.K., and Szeto, W.Y., 2005. Road pricing modeling for hyper-congestion. *Transportation Research Part A*, 39(7-9), 705-722.

Long, J.C., Gao, Z.Y., Zhang, H.Z., and Szeto, W.Y., 2010. A turning restriction design problem in urban road networks. *European Journal of Operational Research*, 206(3), 569-578.

Long, J.C., Szeto, W.Y., and Huang, H.J., 2014. A bi-objective turning restriction design problem in urban road networks. *European Journal of Operational Research*, 237(2), 426-439.

Los, M., 1979. A discrete-convex programming approach to the simultaneous optimization of land use and transportation. *Transportation Research Part B*, 13(1), 33-48.

Lowry, I.S, 1964. *A Model of Metropolis* Santa Monica: Rand Corporation.

Ma, X. and Lo, H.K, 2012. Modeling transport management and land use over time. *Transportation Research Part B*, 46, 687-709.

Marcotte, P., 1986. Network design problem with congestion effects: A case of bilevel programming. *Mathematical Programming*, 34(2), 142-162.

May, A.D., 2013. Urban transport and sustainability: The key challenges. International *Journal of Sustainable Transportation*, 7(3), 170-185.

Meng, Q., Yang, H., and Bell, M.G.H., 2001. An equivalent continuously differentiable model and a locally convergent algorithm for the continuous network design problem. *Transportation Research Part B*, 35(1), 83-105.

Meng, Q., Yang, H., and Wong, S.C., 2000. A combined land-use and transportation model for work trips. *Environment and Planning B*, 27(1), 93-103.

Menge, C.W., Rossano, C.F., Anderson, G.S., and Bajdek, C.J., 1998. *FHWA Traffic Noise Model Technical Manual, Report No FHWA-PD-96-010*. Federal Highway Administration, Washington, DC.

Miandoabchi, E., Daneshzand, F., Szeto, W.Y., and Farahani, R.F., 2013. Multi-objective discrete urban road network design. *Computers and Operations Research*, 40(10), 2429-2449.

Miandoabchi, E., Farahani, and R.Z, Szeto, W.Y., 2012b. Bi-objective bimodal urban road network design using hybrid metaheuristics. *Central European Journal of Operations Research*, 20(4), 583-621.

Miandoabchi, E., Farahani, R.Z., Dullaert, W. and Szeto, W.Y., 2012a. Hybrid evolutionary metaheuristics for concurrent multi-objective design of urban road and public transit networks. *Networks and Spatial Economics*, 12(3), 441-480.

Nakayama, S. and Connors, R., 2014. A quasi-dynamic assignment model that guarantees unique network equilibrium. *Transportmetrica A*, 10(7), 669-692.

Ng, M.W., Szeto, W.Y., and Waller, S.T., 2011. Distribution-free travel time reliability assessment with probability inequalities. *Transportation Research Part B*, 45(6), 852-866.

Nilsson, G., 2004. *Traffic safety dimensions and the Power Model to describe the effect of speed on safety*. Bulletin 221, Lund Institute of Technology, Department of Technology and Society, Traffic Engineering, Lund, Sweden. http://www.dissertations.se/dissertation/a9952d343f/.

Qiu, Y. and Chen, S., 2007. Bi-level programming for continuous network design of comprehensive transportation system based on external optimization. *Proceedings of 2007 IEEE International Conference on Grey Systems and Intelligent Services*, November 18-20, Nanjing, China.

Sharma, S. and Mathew, T.V., 2011. Multi-objective network design for emission and travel time tradeoff for sustainable large urban transportation network. *Environment and Planning B*, 38(3), 520-538.

Smith, J.M. and Liebman, J.S., 1978. A zero-one integer-programming formulation of the problem of land-use assignment and transportation-network design. *Environment and Planning B*, 5(1), 101-115.

Steele, C., 2001. A critical review of some traffic noise prediction models. *Applied acoustics*, 62(3), 271-287.

Szeto, W.Y., Wang, Y., and Wong, S.C., 2014. The chemical reaction optimization approach to solving the environmentally sustainable network design problem. *Computer-Aided Civil and Infrastructure Engineering*, 29(2), 140-158.

Szeto, W.Y., 2013. Routing and scheduling hazardous material shipments: Nash game approach. *Transportmetrica B*, 1(3), 237-260 .

Szeto, W.Y. and Jiang, Y., 2012. A hybrid artificial bee colony for transit network design problem. *Transportation Research Record: Journal of the Transportation Research Board*, 2284, 47-56.

Szeto, W.Y. and Jiang, Y., 2014. Transit assignment: Approach-based formulation, extragradient method, and paradox. *Transportation Research Part B*, 62, 51-76.

Szeto, W.Y. and Lo, H.K., 2005. Strategies for road network design over time: Robustness under uncertainty. *Transportmetrica*, 1(1), 47-63.

Szeto, W.Y. and Lo, H.K., 2008. Time-dependent transport network improvement and tolling strategies. *Transportation Research Part A*, 42(2), 376-391.

Szeto, W.Y. and Wu, Y.Z., 2011. A simultaneous bus route design and frequency setting problem for Tin Shui Wai, Hong Kong. *European Journal of Operational Research*, 209 (2), 141-155.

Szeto, W.Y., Jaber, X.Q., and O'Mahony, M., 2010. Time-dependent discrete network design frameworks considering land use. *Computer-Aided Civil and Infrastructure Engineering*, 25(6), 411-426.

Szeto, W.Y., Jaber, X.Q., and Wong, S.C. 2012. Road network equilibrium approaches to environmental sustainability. *Transport Reviews*, 32, 491-518.

Szeto, W.Y., Jiang, Y., and Sumalee, A., 2011a. A cell-based model for multi-class doubly stochastic dynamic traffic assignment. *Computer-Aided Civil and Infrastructure Engineering*, 26, 595-611.

Szeto, W.Y., Jiang, Y., Wang, D.Z.W., and Sumalee, A., 2013. A sustainable road network design problem with land use transportation interaction over time. *Networks and Spatial Economics*, DOI: 10.1007/s11067-013-9191-9.

Szeto, W.Y. and Jiang, Y., 2014. Transit route and frequency design: Bi-level modeling and hybrid artificial bee colony algorithm approach. *Transportation Research Part B,* accepted.

Szeto, W.Y., Solayappan, M., and Jiang, Y., 2011b. Reliability-based transit assignment for congested stochastic transit networks. *Computer-Aided Civil and Infrastructure Engineering*, 26(4), 311-326.

Ukkusuri, S.V. and Patil, G., 2009. Multi-period transportation network design under demand uncertainty. *Transportation Research Part B*, 43(6), 625-642.

Ukkusuri, S.V. and Waller, S.T., 2008. Linear programming models for the user and system optimal dynamic network design problem: Formulations, comparisons and extensions. *Networks and Spatial Economics*, 8(4), 383-406.

Ukkusuri, S.V., Mathew, T.V., and Waller, S.T., 2007. Robust transportation network design under demand uncertainty. *Computer-Aided Civil and Infrastructure Engineering*, 22(1), 6-18.

Unnikrishnan, A., Valsaraj, V., Damnjanovic, I., and Waller, S.T., 2009. Design and management strategies for mixed public private transportation networks: A meta-heuristic approach. *Computer-Aided Civil and Infrastructure Engineering*, 24(4), 266-279.

Watling, D.P. and Cantarella, G.E.. 2013. Modelling sources of variation in transportation systems: theoretical foundations of day-to-day dynamic models. *Transportmetrica B*, 1 (1), 3-32.

Wang, D.Z. and Lo, H.K., 2010. Global optimum of the linearized network design problem with equilibrium flows. *Transportation Research Part B*, 44(4), 482-492.

Williams, H., 1976. Travel demand models, duality relations and user benefit analysis. *Journal of Regional Science*, 16(2), 147-166.

Williams, H., 1977. On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning A*, 9(3), 285-344.

Wong, C.K., Tong, C.O., and Wong, S.C., 1999. The development and calibration of a Lowry model with multiple market segments. *Environment and Planning A*, 31(11), 1905-1918.

Wong, S.C., Wong, C.K., and Tong, C.O., 2000. A parallelized genetic algorithm for the calibration of Lowry model. *Parallel Computing*, 27(12), 1523-1536.

Wong, R.C.P., Szeto, W.Y., Wong, S.C., and Yang, H., 2014. Modeling multi-period customer-searching behavior of taxi drivers. *Transportmetrica B*, 2(1), 40-59.

Xu, W., Yang, H., and Han, D., 2013. Sequential experimental approach for congestion pricing with multiple vehicle types and multiple time periods. *Transportmetrica B*, 1(2), 136-152.

Yang, H., and Bell, M.G.H., 1998. Models and algorithms for road network design: A review and some new developments. *Transport Reviews*, 18, 257-278.

Yim, K.K.W., Wong, S.C., Chen, A., Wong, C.K., and Lam, W.H.K., 2011. A reliability-based land use and transportation optimization model. *Transportation Research Part C*, 19(2), 351-362.

Yin, Y. and Lawphongpanich, S., 2006. Internalizing emission externality on road networks. *Transportation Research Part D*, 11(4), 292-301.

Appendix

This appendix shows the derivation of the rule-of-half following Williams (1976). In Figure 12, x-axis represents the travel demand for model *m*, $q_{ij,\tau}^{m}$, and y-axis represents the composite cost for model $m, \pi_{ij,\tau}^{m}$. For the assumed curve of the demand function $q\left(\pi_{ij,\tau}^{m}\right)$, the change in consumer surplus due to a fall in the composite cost from $\pi_{ij,\tau}^{m,before}$ to $\pi_{ij,\tau}^{m,after}$ equals the shaded area. Such area can be approximated to the area of the trapezium. Mathematically, the change in consumer surplus can be expressed as

$$\begin{aligned}\Delta CS_{ij,\tau}^{m} &= -\int_{\pi_{ij,\tau}^{m,before}}^{\pi_{ij,\tau}^{m,after}} q\left(\pi_{ij,\tau}^{m}\right) d\left(\pi_{ij,\tau}^{m}\right) \\ &\approx (1/2)(q_{ij,\tau}^{m,before} + q_{ij,\tau}^{m,after})(\pi_{ij,\tau}^{m,before} - \pi_{ij,\tau}^{m,after})\end{aligned}$$

which is equation (21).

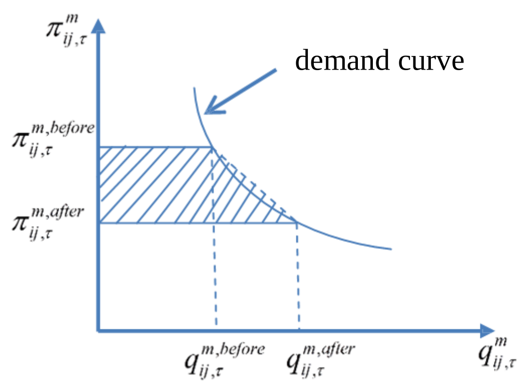


Figure 12