Transportation Research Part C 68 (2016) 194–214

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Prediction of vehicle CO_2 emission and its application to eco-routing navigation

Weiliang Zeng [a,*], Tomio Miwa [b], Takayuki Morikawa [c]

[a] Department of Civil Engineering, Nagoya University, Furo-cho, Chikusa, Nagoya 464-8603, Japan
[b] EcoTopia Science Institute, Nagoya University, Furo-cho, Chikusa, Nagoya 464-8603, Japan
[c] Institute of Innovation for Future Society, Nagoya University, Furo-cho, Chikusa, Nagoya 464-8603, Japan

ARTICLE INFO

Article history:
Received 10 December 2015
Received in revised form 6 April 2016
Accepted 6 April 2016
Available online 11 April 2016

Keywords:
CO_2 emission model
Fuel consumption
Eco-routing
Pareto-optimal
Probe vehicle
Navigation system

ABSTRACT

Transportation sector accounts for a large proportion of global greenhouse gas and toxic pollutant emissions. Even though alternative fuel vehicles such as all-electric vehicles will be the best solution in the future, mitigating emissions by existing gasoline vehicles is an alternative countermeasure in the near term. The aim of this study is to predict the vehicle CO_2 emission per kilometer and determine an eco-friendly path that results in minimum CO_2 emissions while satisfying travel time budget. The vehicle CO_2 emission model is derived based on the theory of vehicle dynamics. Particularly, the difficult-to-measure variables are substituted by parameters to be estimated. The model parameters can be estimated by using the current probe vehicle systems. An eco-routing approach combining the weighting method and *k*-shortest path algorithm is developed to find the optimal path along the Pareto frontier. The vehicle CO_2 emission model and eco-routing approach are validated in a large-scale transportation network in Toyota city, Japan. The relative importance analysis indicates that the average speed has the largest impact on vehicle CO_2 emission. Specifically, the benefit trade-off between CO_2 emission reduction and the travel time buffer is discussed by carrying out sensitivity analysis in a network-wide scale. It is found that the average reduction in CO_2 emissions achieved by the eco-friendly path reaches a maximum of around 11% when the travel time buffer is set to around 10%.

1. Introduction

With travel demand continuing to grow, fuel consumption and greenhouse gas (GHG) emissions are increasing unceasingly. Vehicle emissions contribute substantially to CO, CO_2, HC and NO_x. It has been noted that the transportation sector accounts for approximate 23% of total global CO_2 emissions, of which 73% are generated by road transport (JAMA, 2008; Birol, 2010). Urban traffic emission modeling and control have been attracted more and more attention (Tan and Gao, 2015; Csikos et al., 2015).

Even though alternative fuel vehicles such as all-electric and fuel cell vehicles will be the best solution, mitigating emissions by existing gasoline vehicles is an alternative countermeasure in the near term. Eco-driving, a term used for emerging driving assistance techniques that support the driver in optimizing route choice and driving behavior to reduce vehicle emissions, has been showing significant benefit in fuel saving and air quality improvement (Beusen et al., 2009; Mensing et al.,

* Corresponding author.
E-mail addresses: weiliangzeng49@gmail.com (W. Zeng), miwa@nagoya-u.jp (T. Miwa), morikawa@nagoya-u.jp (T. Morikawa).

http://dx.doi.org/10.1016/j.trc.2016.04.007
0968-090X/

2014). Eco-driving techniques can be classified into three decision-making levels (Sivak and Schoettle, 2012): strategic level (vehicle maintenance), tactical level (pre-trip eco-routing), and operational level (on-board driving assistance). In pre-trip eco-routing, a navigation system attempts to find the most eco-friendly path from origin to destination based on the estimate of average emission for all possible paths, while on-board driving assistance systems analyze drivers' behavior and instantaneous fuel consumption, and provide valuable feedback that helps them adjust their driving behavior to a more eco-friendly style. In this study, we mainly focus on the tactical level, i.e., eco-routing strategy.

Actually, establishing an efficient and preferable eco-routing navigation system is of great challenge because many aspects of this problem need to be considered. For example, what kind of tools can efficiently collect and estimate emissions in a large-scale transportation network? How should an eco-routing strategy be developed and how will such a strategy impact travel time and the environment? To fill these gaps, this study aims to find an eco-friendly path that produces minimum CO_2 emissions while satisfying travel time budget.

The rest of this paper is structured as follows. Section 2, following this introduction, offers a brief literature review related to vehicle emission estimation (or fuel consumption) and eco-routing problem. Section 3 describes the data collection method based on GPS and ODB devices. Section 4 describes the vehicle CO_2 emission model derived from the theory of vehicle dynamics. Section 5 states the eco-routing problem and the Pareto-optimal based routing approach. Section 6 validates the proposed CO_2 emission model and conducts a sensitivity analysis for the trade-off between CO_2 emission and travel time buffer. Finally, the achievements of this study and the direction for future research are outlined in Section 7.

2. Literature review

Routing approaches usually provide the optimal path based on distance, travel time and on-time arrival probability (Nazemi and Omidi, 2013; Zeng et al., 2015a, 2016). Few of them consider vehicle emissions. Intuitively, one may think that the shortest path or fastest path would also be the most eco-friendly path. However, a shortest path may take a driver through a heavy congested area, resulting in high vehicle emissions. On the other hand, there may be cases where a fastest path results in longer travel distance, albeit on less congested roadways. However, traveling on a path at a higher speed over a longer distance will also result in higher vehicle emission compared with a shorter path (Masikos et al., 2015). Many research found that the eco-routing navigation system is an application that promises reduced fuel consumption and emissions in the pre-trip stage (Boriboonsomsin et al., 2012, 2014; Yao and Song, 2013; Guo et al., 2013). For example, Yao and Song (2013) developed a dynamic eco-routing model utilizing standard shortest-path algorithm and consisting of a vehicle specific power (VSP) based model (Jimenez-Palacios, 1998) and a dynamic traffic information database. Similarly, an eco-routing navigation system based on shortest-path algorithm and integrating multisource historical and real-time traffic information was developed and validated in the Greater Los Angeles Metro area (Boriboonsomsin et al., 2012). Nie and Li (2013) estimated the CO_2 emission based on the Comprehensive Modal Emission Model (CMEM) and developed a multiple objective function to the constrained eco-routing problem where the price of travel time and fuel, and the effect of turning movement were considered. Though the solution approach was not discussed, they mentioned that the eco-routing problem can be regarded as a class of constrained shortest problem. Considering the impact of traffic signal, Sun and Liu (2015) formulated the eco-routing problem based on a Markov decision process. An investigation conducted in Sweden (Ericsson et al., 2006) found that 46% of trips based on the spontaneous route choice of the traveler were not the most eco-friendly. Vehicle emissions on these trips could be reduced by 8.2% if the most eco-friendly route were chosen. Similarly, Ahn and Rakha (2008) reported that a 4–20% reduction in vehicle emission can be achieved if an eco-routing strategy is adopted. On the other hand, eco-routing could result in significant reductions in emissions, but it naturally comes at the expense of increased travel time. A field study in Japan (Kono et al., 2008) found that the vehicle emission of the eco-friendly path is 9% lower than that of the least travel time path, while travel time is 9% longer. In such cases, an eco-routing navigation system might suggest the most eco-friendly path with lower vehicle emissions, but the travel time may exceed the travel time budget.

Besides navigation problems, the eco-routing concept was also applied to pollution routing problem, road pricing problem, and rail freight transports. For example, Bektas and Laporte (2011) addressed a pollution routing problem where emission was modeled as a function of vehicle speed and load. Koc et al. (2014) introduced a mix pollution routing problem by considering a heterogeneous vehicle fleet, which aimed to optimize the sum of vehicle fixed costs and routing cost. Franceschetti et al. (2013) proposed an integer linear programming formulation of the time-dependent pollution routing problem which taken into account traffic congestion at peak periods. Chen and Yang (2012) studied a Pareto-optimal pricing scheme that aims to take into account both vehicular congestion and CO emission. Kirschstein and Meisel (2015) developed mesoscopic greenhouse gas emission models for evaluating the eco-friendliness of rail freight transports.

The eco-routing problem discussed in this study can be regarded as the constraint shortest path problem. It is difficult to solve by using standard shortest-path algorithms (Dijkstra, 1959; Hart et al., 1968) due to the constrained term. Alternatively, path ranking techniques such as k-shortest path algorithm can be regarded as a solution by increasing k until a path satisfying the constraint first was found (Carlyle and Wood, 2005). However, it is intractable to a large-scale network due to the exponential increase in the computational effort if k is set to a large number. Another technique is called Lagrangian relaxation (Ahuja et al., 1993), by which the side constraints can be relaxed and supplemented into the objective function as Lagrangian terms. However, the optimal solution may not be found if feasible solutions fall into the duality gap. To fill this gap, Beasley and Christofides (1989) solved the constrained shortest path problem by using sub-gradient optimization

and developed a branch-and-bound approach to close the duality gap. Carlyle et al. (2008) proposed a depth first branch-and-bound approach, and a bisection searching technique was applied to determine the Lagrangian multipliers. Recently, Chen and Nie (2013) investigated a generalized bi-objective shortest path problem with non-additive cost and they used a *k*-shortest path ranking procedure to close the gap.

In summary, there is growing interest in routing problems with consideration of environmental aspects. However, few studies related eco-routing problems developed a practical link-based vehicle emission model and took travel time budget into account. Our study attempts to address this problem by seamlessly incorporating a vehicle dynamic based CO_2 emission model and a Pareto-optimal based routing approach, and discuss the benefit trade-off between travel time and emission in a simulation study.

3. Data collection for vehicle co_2 emission and travel time

As shown in Fig. 1, we collect data from probe vehicles with GPS and OBD (on-board diagnostics) device (Zeng et al., 2015b). The information collected includes fuel consumption, vehicle emission, speed, acceleration, brake use, accelerator position, and so on. The field data collection device in this study can be regarded as the portable emission measurement system, which is able to measure the exhaust concentrations of CO_2, CO, HC, and NO_x (Frey et al., 2012). Because CO_2 is a major greenhouse gas that contributes to global warming and it account for about 95% of the total exhaust gas (EPA, 2000), we choose CO_2 as the representative vehicle emission in this study. Through the OBD device, data on engine operation and fuel consumption are logged second-by-second. To calculate the CO_2 emission from fuel consumption, the carbon emission can be derived by multiplying the ratio of the molecular weight of CO_2 by the molecular weight of carbon (Coe, 2005). According to the previous study by the U.S. Environmental Protection Agency, CO_2 emission from a liter of gasoline is set to 2.32 kg. The Secure Digital (SD) card can be used as the data logger which records the GPS data and OBD data simultaneously. And then all the data can be uploaded to a central server. Considering the applicability and robustness of the routing model for a navigation system, we aggregate the emission data and travel time in a link-based level after map-matching (Miwa et al., 2012).

A real-world network with 4072 nodes and 12,877 links in Toyota city, Japan, is used to test the eco-routing procedure and evaluate the likely environment benefits of implementing it. The trip records are from 153 GPS probe vehicles over a period of 10 months in 2011. Fig. 2 gives an example of the link-based dataset for a single trip record. The link-based attributes include travel time, distance, CO_2 emission, etc., which makes them a convenient dataset for vehicle CO_2 emission modeling and its application to the eco-routing problem. The sample size in each link is shown in Fig. 3. About 80% of the whole network can be covered by the probe vehicles. The average number of probe vehicle is 11 for each link per week. The reliability of average link speed corresponds to the sample size in each link. Simulation test indicated that the network needs to have at least 10 probe vehicles passed through a link within the sampling period for an absolute error in estimated average link speeds to be less than 5 km/h (Long Cheu et al., 2002). The reliability of average link speed estimation can be defined as the probability that the relative error (ε) is less than the maximum acceptable relative error (ε_{max}). Assumed that the link speed follows normal distribution, the reliability of average link speed can be derived from the central limit theorem (Srinivasan and Jovanis, 1996).

$$r = \Phi\left(\frac{\varepsilon_{max}\mu}{\sigma/\sqrt{n}}\right) \tag{1}$$

where r is the reliability of average link speed, Φ is the cumulative function of standard normal distribution, μ is the expected link speed, σ is the expected standard deviation of link speed, n is the sample size.

In case that some of the links are not covered by probe vehicles, the average speed and acceleration can be estimated by using the data of neighboring links or historical data. Various methodologies such as machine learning can be applied to the estimation. Because such topic is out of our research scope, we propose a simple weighting method in this study. As shown in Fig. 4, if there are no probe vehicles passing through the objective link, the neighboring links inside a specified sphere will be utilized for average speed estimation. The searching sphere can be defined by the center of the objective link and a specified radius (R). Considering that those neighboring links connected with the objective link may have higher correlation, a higher weight value can be set to the connected links.

$$v_e = \lambda v_c + (1-\lambda) v_n \tag{2}$$

where v_e is the estimated average speed of the objective link, v_c is the average speed of the neighboring links connected to the objective link inside a specified sphere, v_n is the average speed of the neighboring links without connection with the objective link, λ is the weight value. A suitable sphere radius and λ can be estimated from a large-scale data set by using data mining technology. For simplification of calculation, we set sphere radius as 100 m and λ as 0.7 in this study.

Fig. 5 shows the day-to-day link-based average speed at different times of the day. The off-peak and peak hours are significant on weekday. The average speed is about 9.4 m/s at peak hours (7:00–9:00 am, 18:00–19:00 pm) and about 10.1 m/s at off-peak hours. However, the peak hour at weekend is not significant. The average speed decreases from 13.3 m/s to 10.2 m/s as the time of day elapses from 6:00 am to 11:00 am. And then, it keeps about 10.5 m/s from 11:00 am to 22:00 pm. A possible reason is that people usually plan outdoor activities late in the morning at weekends, and the activity time is more flexible than commuting time.

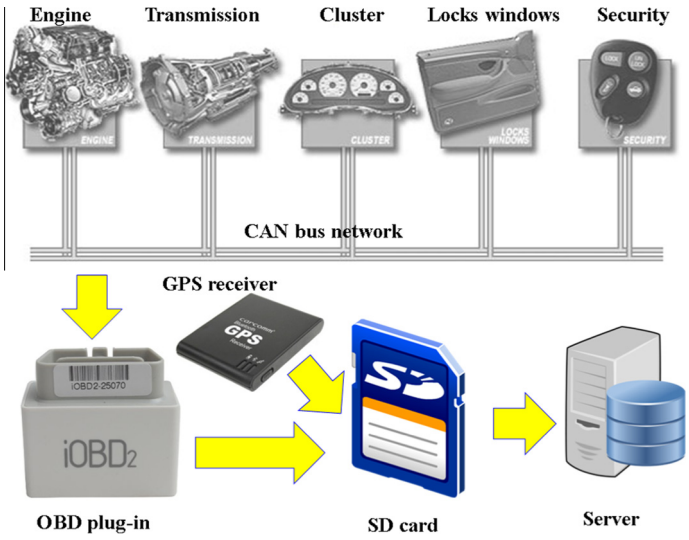

Fig. 1. Data collection for vehicle CO_2 emission.

Link ID	Vehicle Model	Displacement (cc)	Date	Travel time (s)	Distance (m)	Acceleration (m/s^2)	Road grade (degree)	CO_2 emission (kg)
7	Toyota Vitz	1496	2011-03-27 14:53:34	18	186	0.05	0.6	0.016
8	Toyota Vitz	1496	2011-03-27 14:53:45	11	230	0.12	0.7	0.014
9	Toyota Vitz	1496	2011-03-27 14:54:30	45	131	0.51	2.1	0.024
…	Toyota Vitz	…	…		…			…
15	Toyota Vitz	1496	2011-03-27 15:20:10	26	250	0.21	1.2	0.017

Fig. 2. Example of data collection.

The fluctuation of CO_2 emission will be influenced by the varying average speed in the road network. As shown in Fig. 6, the average CO_2 emission rate is 0.185 g/m at off-peak hours on weekday, while it increases to 0.214 g/m (15.7% increase) at peak hours. The CO_2 emission at morning peak hours of weekend is not significant because the average speed is relatively high at that time.

4. Fuel and CO_2 emission model based on vehicle dynamics

To develop an eco-routing navigation system, it is necessary to estimate the expected vehicle CO_2 emission for each link. A number of factors based on the characteristics of traffic, vehicle, and street configuration are found to affect vehicular fuel consumption or CO_2 emission (Park and Rakha, 2006; Ahn et al., 2002; Pandian et al., 2009; Guo et al., 2014). These variables can be roughly grouped into seven categories, i.e., travel-related (e.g., travel distance), weather-related (e.g., wind effects, snow, and rainfall), vehicle-related (e.g., acceleration, power demands, engine displacement, and catalyst for emissions),

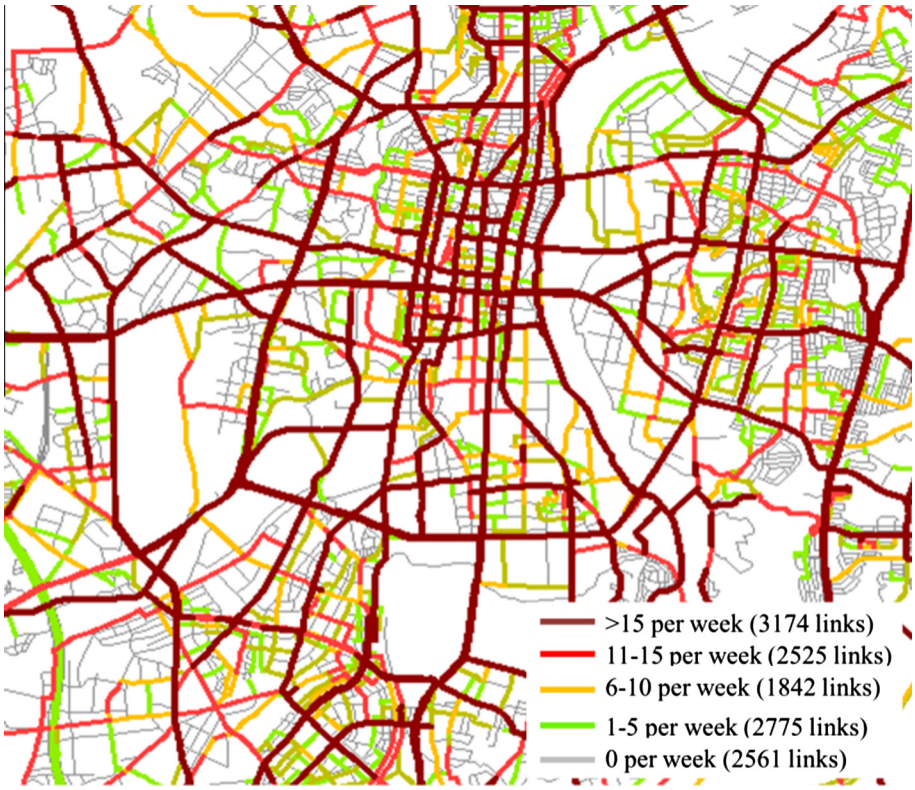

Fig. 3. Sample size in each link.

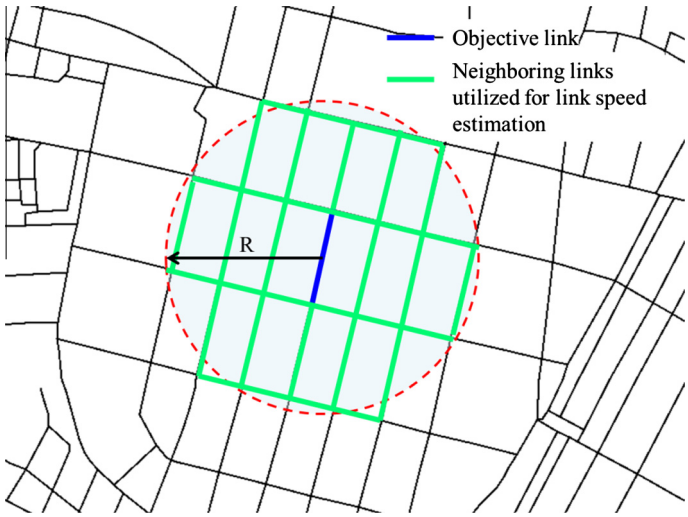

Fig. 4. Searching area for link speed estimation.

roadway related (e.g., road grade), traffic-related (e.g., congestion and average speed), driver-related factors (e.g., risk-taking and aggressive driving), and technology related (e.g., if communication technologies and sensors are enabled). In general, the selections of these factors to estimate fuel consumption or CO_2 emission usually depends on the context of the study, ease of measurement and ease of computation. Because most of these factors are related to the total engine power, our approach to estimate the fuel consumption or CO_2 emission is built on the theory of vehicle dynamics, which converts total engine power to fuel rate.

As shown in Fig. 7, the forces acting on a moving vehicle include driving force (F_d), rolling resistance force (F_r), aerodynamic drag force (F_a), and gravitational force (Rajamani, 2011). The vehicle dynamics function can be formulated as follows:

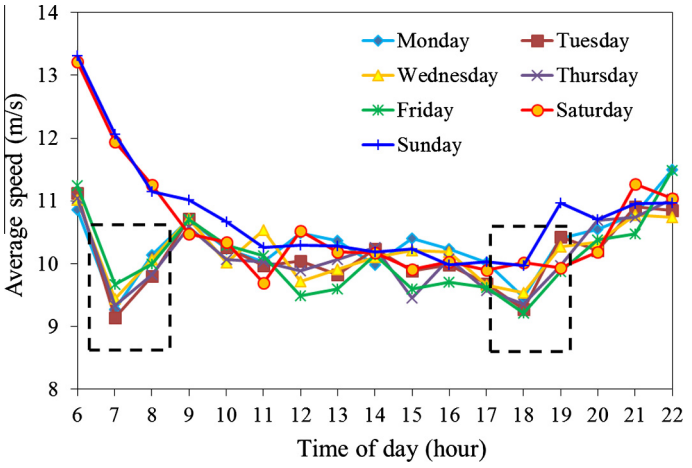

Fig. 5. Average link-based speed distribution in a time-varying road network.

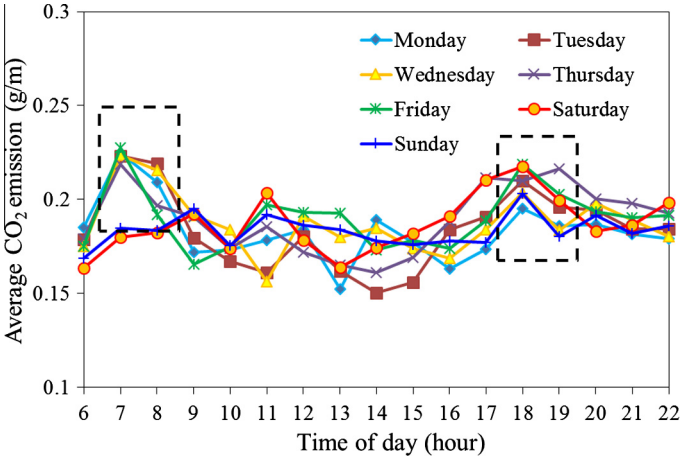

Fig. 6. Average link-based CO_2 emission distribution in a time-varying road network.

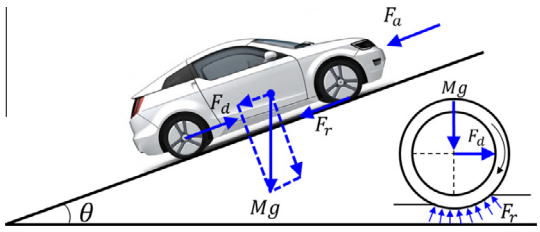

Fig. 7. Forces acting on a vehicle moving on an inclined road.

$$(M+m)a = F_d - F_r - F_a - Mgsin(\theta) \tag{3}$$

where

M: Mass of the vehicle;
m: Inertia weight when the vehicle accelerates;
a: Acceleration;
g: Acceleration of gravity;
θ: Angle of inclination of the road.

Because the inertia weight (m) in Eq. (3) is roughly proportional to the inverse speed, it can be represented as follows:

$$m = m_0 + \frac{m_1}{v} \tag{4}$$

where m_0 and m_1 are the parameters to be estimated, v is the vehicle speed.

The driving force can be derived from the effective torque (Q_e) and the radius of tire (R).

$$F_d = \frac{Q_e}{R} \tag{5}$$

The torque can be derived from the effective engine power (P_e) as follows (Hendricks and Sorenson, 1990):

$$Q_e = \frac{r\eta P_e}{2\pi n} \tag{6}$$

$$n = \frac{vr}{2\pi R} \tag{7}$$

$$P_e = P - P_{idle} \tag{8}$$

$$P = \varepsilon H f_t \tag{9}$$

where

r: Overall gear reduction ratio;
η: Thermal efficiency multiplier;
v: Vehicle speed;
n: Engine revolution;
P_{idle}: Engine power in idling state;
P: Total engine power;
ε: Thermal efficiency;
H: Fuel energy constant;
f_t: Fuel rate.

When the tires are rotating, both the tires and the road surface are subject to deformation in the contact patch. The energy spent in deforming the tires cannot be recovered completely when the tires return to original shape (Rajamani, 2011). The energy consumption can be represented by the rolling resistance force that acts on the opposite moving direction of the vehicle. Typically, the rolling resistance force can be modeled as being roughly proportional to the offset load on the tires.

$$F_r = \mu Mg\cos(\theta) \tag{10}$$

where μ is the rolling resistance coefficient.

Aerodynamic drag force acts on any moving vehicle in the direction of the air free stream, which increases significantly with vehicle speed. The equivalent aerodynamic drag force on a vehicle can be represented as follows:

$$F_a = \frac{1}{2}\rho C_d A_f v^2 \tag{11}$$

where

ρ: Mass density of air;
C_d: Aerodynamic drag coefficient;
A_f: Front area of the vehicle.

According to Eqs. (3)–(11), the fuel rate (f_t) can be represented as follows:

$$f_t = \frac{(\mu\cos(\theta) + \sin(\theta))Mg}{\varepsilon\eta H} v + \frac{\rho C_d A_f}{\varepsilon\eta H} v^3 + \frac{M + m_0}{\varepsilon\eta H} a v + \frac{m_1}{\varepsilon\eta H} a + \eta P_{idle} \tag{12}$$

In practice, it is usually difficult to get all the variables such as η, P_{idle}, and μ. To simplify the fuel consumption model and make it applicable to the eco-routing problem, we only keep the easy-to-measure variables, i.e., v, a, and θ in the model. And the difficult-to-measure variables will be substituted by parameters to be estimated. Hence, the fuel rate (g/s) can be represented as follows:

$$f_t = \beta_1 \cos(\theta) v + \beta_2 \sin(\theta) v + \beta_3 v^3 + \beta_4 a v + \beta_5 a + \beta_6 \tag{13}$$

To make the fuel rate convenient to navigation problem, the fuel consumption per second can be transformed to fuel consumption per meter by dividing the speed,

$$f_l = \beta_1 \cos(\theta) + \beta_2 \sin(\theta) + \beta_3 v^2 + \beta_4 a + \beta_5 \frac{a}{v} + \beta_6 \frac{1}{v} + \beta_7 \tag{14}$$

where f_l is the fuel rate, g/m.

Because CO_2 emission is approximated linear to fuel consumption (Coe, 2005), CO_2 emission can be represented as follows:

$$E_{CO_2} = 2.32 * f_l \tag{15}$$

There were studies shown that instantaneous speed and acceleration play important roles in fuel consumption or vehicle emission (Ahn et al., 2002). However, obtaining instantaneous speed and acceleration for a large-scale transportation network is very difficult. Application of the instantaneous data to route choice is even more challenging. Considering that the urban transportation information system usually provides the average travel time or speed for each link in a certain interval, we use the average link speed and average acceleration as the surrogate inputs. The average acceleration is calculated as the change between the entering speed and the outgoing speed on a link. The angle of inclination can be obtained from the digital map. If the elevation data is not available in digital map, it can be extracted by using the Google Elevation API. The digital elevation dataset in Google Earth is mainly provided by SRTM (Shuttle Radar Topography Mission). And Google developers may improve it by adding data from other sources such as ASTER (Advanced Space-borne Thermal Emission and Reflection Radiometer). SRTM released a 1-arc second global digital elevation model in 2014 (http://www2.jpl.nasa.gov/srtm/) and most of the earth has been covered by this dataset ranging from 54° south to 60° north latitude except for the Middle East and North Arica. An experiment in Malaysia shown that the correlation of elevation data between Google Earth and SRTM for flat and hilly areas were 0.791 and 0.891, respectively (Rusli et al., 2014). And the correlation of elevation data from Google Earth and ASTER was 0.917 in mountainous areas. Therefore, it is believable that the accuracy of elevation derived from Google Elevation API is acceptable.

5. Eco-routing problem

5.1. Problem statement

Our objective is to find the minimum CO_2 emissions path between two nodes in a transportation network within the constraint of a certain travel time budget. The transportation network is modeled as a directed graph G(*N*,*A*), where $N = \{1, 2, \ldots, n\}$ represents the set of nodes and $A = \{a_{12}, a_{23}, \ldots, a_{mn}\}$ represents the set of links. Unlike a traditional network that only has a single attribute assigned to each link, such as travel time or length, the network considered in this study has two attributes: link CO_2 emission c_{ij} and link travel time t_{ij}. c_{ij} is estimated by using the proposed CO_2 emission model described in Section 4. Here, we wish to find the most eco-friendly path within the travel time budget (T). Consequently, the eco-routing problem from origin r to destination s can be described as the following integer programming problem:

$$\text{P1}: \quad \text{Min } Z(\boldsymbol{x}) = \sum_{ij \in A} c_{ij} x_{ij} \tag{16}$$

Subject to

$$\sum_{ij \in A} t_{ij} x_{ij} \leqslant T \tag{17}$$

$$\sum_{(i,j) \in A} x_{ij} - \sum_{(j,i) \in A} x_{ji} = g \tag{18}$$

$$g = \begin{cases} 1 & i = r \\ 0 & i \in N - \{r, s\} \\ -1 & i = s \end{cases} \tag{19}$$

where $x_{ij} \in \{0, 1\}$ indicates a link on the selected path and g represents the flow status for each node i in the network.

In this study, the concept of the Pareto frontier and the corresponding weighting method is extended to solve the eco-routing problem with a travel time constraint. To utilize a Pareto-optimal approach to the eco-routing problem, such as by implementing a weighting method, we transform P1 into a bi-objective like optimization problem as follows:

$$\text{P2}: \quad \text{Min } Z(\boldsymbol{x}) = w_t Z_t(\boldsymbol{x}) + w_c Z_c(\boldsymbol{x}) \tag{20}$$

Subject to constraints (17)–(19).

where

$$Z_t(\boldsymbol{x}) = \sum_{ij \in A} t_{ij} x_{ij} \tag{21}$$

$$Z_c(\boldsymbol{x}) = \sum_{ij \in A} c_{ij} x_{ij} \tag{22}$$

w_t: weighting parameter for the objective of travel time;
w_c: weighting parameter for the objective of CO_2 emissions.

Note that if constraint (17) is relaxed, solving P2 will generate a Pareto frontier (Pareto-optimal or non-dominated path set). Considering that the objective of this study is to find an optimal path but not the path set, we design a search algorithm to find the optimal solution along the Pareto frontier. The search procedure will stop when the travel time of a candidate Pareto-optimal solution reaches the specified travel time budget (T).

5.2. Routing approach

Prior to introducing the Pareto-optimal based approach, we first list some fundamental definitions relating to the concept of Pareto optimality.

Definition 1 (*Pareto-optimal solution and Pareto frontier*). Given a bi-objective shortest path problem, a path $p^* \in P$, where P is the path set from origin r to destination s, is called a Pareto-optimal solution if there is no other path $p \in P$ with $Z_t(\boldsymbol{x}) \leqslant Z_t(\boldsymbol{x}^*)$ and $Z_c(\boldsymbol{x}) \leqslant Z_c(\boldsymbol{x}^*)$ with at least one inequality being strict. An equivalent condition to the Pareto-optimal solution p^* is that there is no other feasible solution $p \in P$ that dominates p^*. The set of Pareto-optimal or non-dominated solutions is called the Pareto frontier.

Definition 2 (*Supported and unsupported Pareto-optimal solutions*). Supported Pareto-optimal solutions (or supported non-dominated solutions) are those feasible solutions that can be obtained as optimal solutions of a weighted sum optimization problem Min $(w_t Z_t + w_c Z_c)$ for $w_t, w_c > 0$. All other Pareto-optimal solutions or non-dominated solutions are called unsupported Pareto-optimal solutions (Sedeno-Noda and Raith, 2015).

As illustrated in Fig. 8(a), the supported Pareto-optimal solutions lie on the lower left boundary of the convex hull of the feasible solution set, whereas the unsupported Pareto-optimal solutions lie in the triangle areas determined by two adjacent supported Pareto-optimal solutions. The Pareto frontier comprises all the Pareto-optimal solutions including both supported and unsupported solutions. Assuming that the traveler prefers a fuel-efficient path (e.g., p^*) with the lowest CO_2 emissions within the travel time constraint, the optimal path can be found along the Pareto frontier. A weighting method (Coutinho-Rodrigues et al., 1999) can be applied to efficiently determine the Pareto frontier, using a weighting utility function combining all the objectives so as to convert a multi-objective problem into a single-objective problem with a range of parameter values. Unfortunately this method does not identify the unsupported Pareto-optimal solutions that lie in the interior of the convex hull. For example, if the travel time budget is set to T, the optimal solution should be p^* in Fig. 8(a). However, p^* is an unsupported Pareto-optimal solution in this case, so it cannot be found by the weighting method. To fill this gap, a three-step Pareto-optimal based algorithm combining the weighting method with k-shortest path algorithm is proposed in the following sections.

5.2.1. Step 1: Initialization

The algorithm gets started by locating two initial Pareto-optimal solutions using two weighting parameter sets, i.e., $(w_t, w_c) = (\varepsilon, 1-\varepsilon)$ and $(w_t, w_c) = (1-\varepsilon, \varepsilon)$, where ε is a sufficiently small number, i.e., $0 < \varepsilon \ll 1$. Note that a shortest path minimizing the CO_2 emissions ($Z_c(\boldsymbol{x})$) may be obtained when we solve the weighting program for $\varepsilon = 0$, but the corresponding shortest path could be a dominated path as the second objective is not taken into account. For example, as shown in Fig. 8(a), p' could be found if $w_t = 0$, but p' is not the Pareto-optimal solution because it is dominated by Z_2. Therefore, we use a small ε value in the initial weighting parameter sets to avoid the risk of choosing a dominated solution at the initial iteration. Suppose that the two Pareto-optimal solutions obtained by solving the two initial weighting problems are Z_1 and Z_2, three situations should be considered for different settings of travel time budgets. (1) If the travel time budget is less than the travel time of Z_1, i.e., $T < Z_{1,t}$, no paths can be found. (2) If the travel time budget is between the travel time of Z_1 and Z_2, i.e., $Z_{1,t} \leqslant T \leqslant Z_{2,t}$, the optimal path will be located on the Pareto frontier as shown in Fig. 8(a). (3) If the travel time budget is more than the travel time of Z_2, i.e., $T > Z_{2,t}$, the optimal path is Z_2 which can be found directly by using the shortest-path algorithm with the weighting parameter set $(w_t, w_c) = (\varepsilon, 1-\varepsilon)$. Here, we put the focus on how to find the optimal path for the second situation. The two resulting solutions, i.e. Z_1 and Z_2, are recorded into a list $\boldsymbol{L}$ sorted ascending by path travel time, i.e., $\boldsymbol{L} = \{Z_a^{(0)}, Z_b^{(0)}\}$, where $Z_a^{(0)} = Z_1$, $Z_b^{(0)} = Z_2$, and $Z_{a,t}^{(0)} \leqslant Z_{b,t}^{(0)}$. And then a weighting method (Coutinho-Rodrigues et al., 1999) approximates the true optimal solution in next step.

5.2.2. Step 2: Find the initial duality gap

To ensure a rapid approach to the optimal solution, a NISE-like algorithm is used to update the weighting parameters (Coutinho-Rodrigues et al., 1999). The weighting parameters are generated from the two Pareto-optimal solutions in list $\boldsymbol{L}$ iteratively. The most widely used method of generating parameters is the perpendicular method (Xie and Waller, 2012), which results in a parameter vector perpendicular to the line going through the two Pareto-optimal solution points.

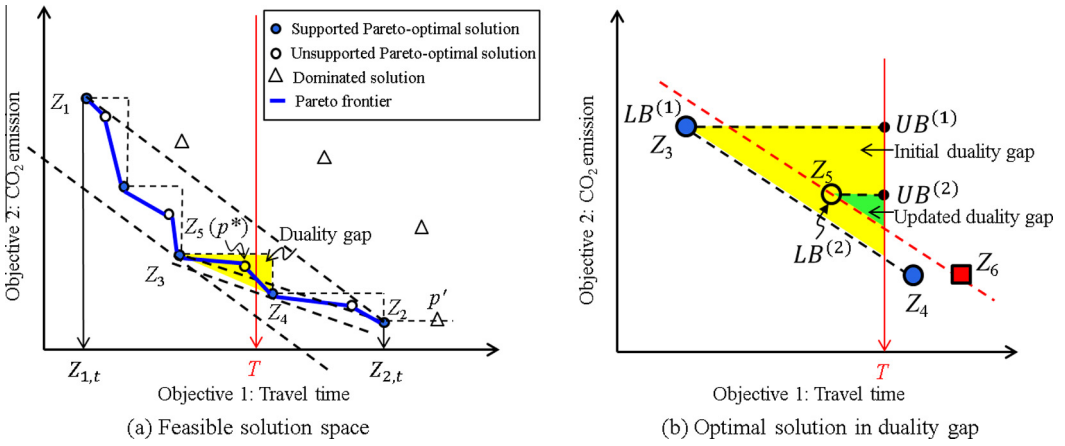

Fig. 8. Search procedure used by the Pareto-optimal based algorithm.

Specifically, given two Pareto-optimal solutions in list ***L***, the perpendicular method generates a new parameter set using the following linear equation system:

$$\begin{cases} w_t^{(n)}(Z_{a,t}^{(n-1)} - Z_{b,t}^{(n-1)}) + w_c^{(n)}(Z_{a,c}^{(n-1)} - Z_{b,c}^{(n-1)}) = 0 \\ w_t^{(n)} + w_c^{(n)} = 1 \end{cases} \tag{23}$$

Thus,

$$w_t^{(n)} = \frac{Z_{b,c}^{(n-1)} - Z_{a,c}^{(n-1)}}{(Z_{b,c}^{(n-1)} - Z_{a,c}^{(n-1)}) + (Z_{a,t}^{(n-1)} - Z_{b,t}^{(n-1)})} \tag{24}$$

$$w_c^{(n)} = \frac{Z_{a,t}^{(n-1)} - Z_{b,t}^{(n-1)}}{(Z_{b,c}^{(n-1)} - Z_{a,c}^{(n-1)}) + (Z_{a,t}^{(n-1)} - Z_{b,t}^{(n-1)})} \tag{25}$$

where

$Z_{a,t}^{(n-1)}$: Path travel time of $Z_a^{(n-1)}$ in list ***L***;
$Z_{b,t}^{(n-1)}$: Path travel time of $Z_b^{(n-1)}$ in list ***L***;
$Z_{a,c}^{(n-1)}$: Path CO_2 emission of $Z_a^{(n-1)}$ in list ***L***;
$Z_{b,c}^{(n-1)}$: Path CO_2 emission of $Z_b^{(n-1)}$ in list ***L***;
n: Iteration number.

The weighting single-objective shortest path problem specified by the updated $w_t^{(n)}$ and $w_c^{(n)}$ can be solved iteratively. A strategy is taken to identify the two candidate supported Pareto-optimal solutions that are closest to the true optimal solution. That is, the new supported Pareto-optimal solution is added to list ***L*** and meanwhile the previous solution, in which the travel time locates on the same side as the new solution relative to T, is removed from list ***L*** iteratively. For example, as shown in Fig. 8(a), Z_3 is obtained by solving the weighting shortest path problem with updated $w_t^{(1)}$ and $w_c^{(1)}$. We compare the travel time of Z_3 with T. Because Z_3 is on the left side of T ($Z_3 < T$), Z_2 should remain in ***L***. Thus, the items in ***L*** are updated by replacing Z_1, i.e., $Z_a^{(1)} = Z_3$, and $Z_b^{(1)} = Z_2$ at iteration 1. The iteration procedure is ended if no new supported Pareto-optimal solutions can be found. For example, since there are no other new supported Pareto-optimal solutions that can be found by updating the weighting parameters generated by Z_3 and Z_4, the parameter generation procedure is stopped. However, the actual optimal solution Z_5 should be located in the yellow triangle region, which is called the duality gap, though Z_3 is a potential optimal solution. The duality gap can be determined using the two adjacent supported Pareto-optimal solutions in list ***L***. It must be noted that Z_3 may not be the optimal solution because the NISE-like algorithm only identifies those supported Pareto-optimal solutions that lie on the boundary of the convex hull, but it cannot find unsupported Pareto-optimal solutions, e.g., Z_5. For this reason, we present an additional procedure to identify unsupported Pareto-optimal solutions in the duality gap and determine the optimal solution in the next step.

5.2.3. Step 3: Update the duality gap and determine the optimal solution

An efficient algorithm for closing the duality gap was introduced by Current et al. (1990) for the bi-objective routing problem, in which the upper bound is updated by checking the trade-off between two objective values according to the traveler's preference interactively. This inspires us to develop a heuristic approach to close the duality gap for the constrained eco-routing problem. In our algorithm, the upper bound and lower bound are updated according to the updated solution in the duality gap, but the weighting parameters do not need to update. And the duality gap is further limited to a smaller triangle region bounded by the travel time budget as shown in Fig. 8(b). Considering that the optimal solution may locate in the duality gap, a *k*-shortest path algorithm is applied to search for potential optimal solutions until there are no further new unsupported Pareto-optimal solutions in the duality gap. Specifically, the *k*-shortest path searching procedure stops when the upper bound (UB) of the duality gap has been reached. At each iteration, the upper bound is determined by the two solutions in list $\boldsymbol{L} = \{Z_a^{(n)}, Z_b^{(n)}\}$ and the travel time budget (T):

$$UB^{(n)} = w_t^{(n)} T + w_c^{(n)} Z_{a,c}^{(n)} \tag{26}$$

Accordingly, the objective function for the *k*-shortest path algorithm is as follows:

$$\text{Min } Z(\boldsymbol{x}) = \sum_{ij \in A} (w_t^{(n)} t_{ij} + w_c^{(n)} c_{ij}) x_{ij} \tag{27}$$

Subject to constraints (18) and (19).

Any *k*-shortest path algorithm for a directed network can be used in this study. Here, we use Yen's well-known algorithm (Yen, 1971) to find the unsupported Pareto-optimal solution. The UB is updated when a new unsupported Pareto-optimal solution is found in the duality gap by the *k*-shortest path algorithm. For example, at iteration 1, $Z_a^{(n)}$ in list $\boldsymbol{L}$ is updated to Z_5 and the UB is updated to $UB^{(1)}$ when Z_5 is found. Correspondingly, the lower bound (LB) can be defined as the weighted sum cost of the current unsupported Pareto-optimal solution, i.e., $Z_a^{(n)}$.

$$LB^{(n)} = w_t^{(n)} Z_{a,t}^{(n)} + w_c^{(n)} Z_{a,c}^{(n)} \tag{28}$$

Note that in using the *k*-shortest path searching procedure, it is possible that solutions outside the duality gap might be found, e.g., Z_6. These solutions will not be added to list $\boldsymbol{L}$ and the UB will not be updated; instead the next *k*th shortest path is calculated until a feasible solution is found in the duality gap.

5.3. An illustrative example for the proposed eco-routing approach

An example of the eco-routing problem with travel time constraint is illustrated here to demonstrate the effectiveness of the proposed search procedure. The tested network and the link attributes are shown in Fig. 9. The origin and destination are set to node 1 and node 12, respectively. The objective is to find the most eco-friendly path within the travel time budget (e.g., $T = 42$). The solutions in outcome space are shown in Fig. 10, and the iteration results are shown in Table 1.

In the first step, given the initial weighting parameter sets, ($w_t = 0.999$, $w_c = 0.001$) and ($w_t = 0.001$, $w_c = 0.999$), the first two supported Pareto-optimal solutions, i.e., Z_1 and Z_2 shown in Fig. 10, are identified by solving the weighted sum shortest path problem. Then, the travel times of the two solutions are checked against the specified travel time budget. Since the travel time budget is between the travel times of these two Pareto-solutions, i.e., $33 < T < 69$, Z_1 and Z_2 are added to list $\boldsymbol{L}$, and the calculation procedure goes to step 2.

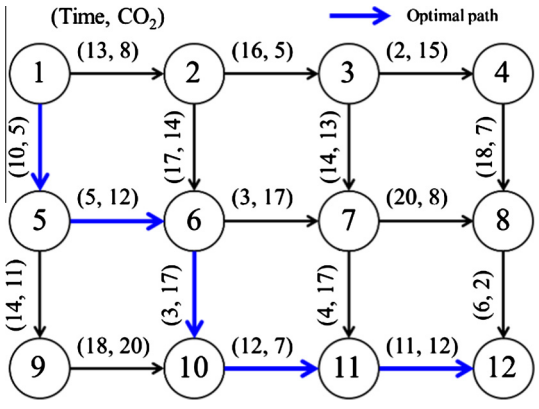

Fig. 9. Illustration of the tested network and the optimal path.

In step 2, the weighting parameters are updated using Eqs. (24) and (25), i.e., $w_t = 0.43$ and $w_c = 0.57$. Then, the next supported Pareto-optimal solution is found using the shortest-path algorithm, i.e., Z_3. After finding Z_3, we need to update $\boldsymbol{L}$. It is found that Z_3 is on the right side of T ($Z_3 > T$). Thus, Z_2 (on the right side of T) is removed and Z_3 is added to list $\boldsymbol{L}$. The weighting parameters are updated iteratively using the objective values of the two candidate solutions in list $\boldsymbol{L}$, i.e., Z_1 and Z_3. In iteration 2, because no other solutions can be found using the updated weighting parameters, i.e., $w_t = 0.63$ and $w_c = 0.37$, the duality gap can be identified by Z_1 and Z_3, and the calculation procedure goes to step 3.

In step 3, we find the optimal solution inside the duality gap. First, the UB value of the weighted sum cost of the k-shortest path is set using Eq. (26), i.e. $UB = 0.633 \times 42 + 0.367 \times 63 = 49.7$. The LB value is set as the weighted sum cost of one of the solutions in list $\boldsymbol{L}$, i.e., $LB = 0.633 \times 33 + 0.367 \times 63 = 44$. Then, the k-shortest path algorithm searches iteratively for the kth best path constrained by UB and LB. If the kth best path falls inside the duality gap, i.e., Z_4, the first solution in list $\boldsymbol{L}$ is replaced by the kth best path, i.e., $Z_a = Z_4$. And then UB and LB are updated by the new solutions in list $\boldsymbol{L}$, i.e., $UB = 0.633 \times 42 + 0.367 \times 53 = 46.03$, $LB = 0.633 \times 41 + 0.367 \times 53 = 45.4$. The k-shortest path algorithm is called again with the updated UB and LB. Because there are no more solutions to be found in the updated duality gap, the algorithm stops and the optimal solution, i.e., Z_4, is found.

6. Numerical experiment and discussion

6.1. Estimation and evaluation for the proposed CO_2 emission model

The parameters of the proposed CO_2 emission model are estimated by using the maximum likelihood estimation (MLE). It is assumed that the error of f_l in Eq. (14) follows normal distribution with zero mean and standard deviation (σ). The likelihood L_k of one sample can be represented as follows:

$$L_k(\boldsymbol{\beta}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(f - f_l(v,a,\theta|\boldsymbol{\beta}))^2}{2\sigma^2}} \tag{29}$$

where f and $f_l(v, a, \theta|\boldsymbol{\beta})$ are the observed and estimated fuel rate, respectively. Considering the entire samples, the likelihood of the samples given for model parameters can be represented as follows:

$$L(\boldsymbol{\beta}) = \prod_{k=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(f - f_l(v,a,\theta|\boldsymbol{\beta}))^2}{2\sigma^2}} \tag{30}$$

where n is the number of sample. To facilitate the computation, the logarithm is taken on both sides of Eq. (30). Thus, the maximum log-likelihood estimates of model parameters ($\boldsymbol{\beta}$) are obtained such that the following equation is maximized.

$$\ln L(\boldsymbol{\beta}) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{k=1}^{n}(f - f_l(v,a,\theta|\boldsymbol{\beta}))^2 \tag{31}$$

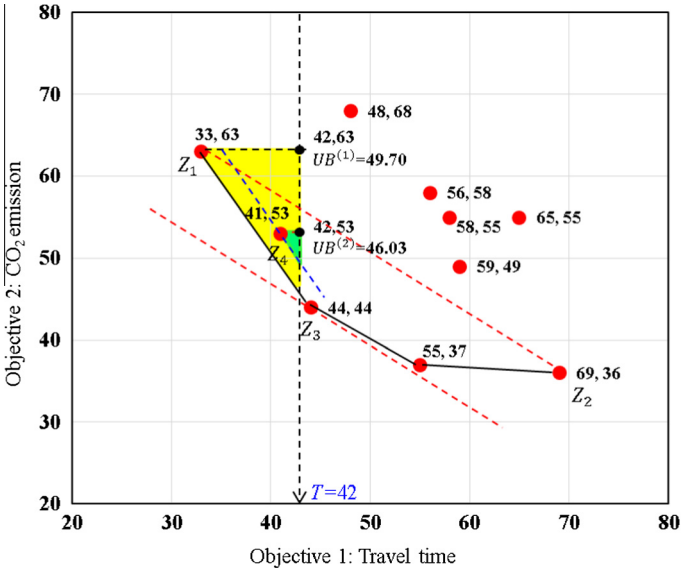

Fig. 10. Solutions in outcome space.

Table 1
Solution search process and iteration results for the illustrative eco-routing example.

Step	Iteration	w_t	w_c	*LB*	*UB*	List $\boldsymbol{L}$		Objective value			
						Z_a	Z_b	$Z_{a,t}$	$Z_{a,c}$	$Z_{b,t}$	$Z_{b,c}$
1	1	0.999	0.001	Null	Null	1–5–6–7–11–12	Null	33	63	Null	Null
1	2	0.001	0.999	Null	Null	1–5–6–7–11–12	1–2–3–7–8–12	33	63	69	36
2	1	0.429	0.571	Null	Null	1–5–6–7–11–12	1–5–6–7–8–12	33	63	44	44
2	2	0.633	0.367	Null	Null	1–5–6–7–11–12	1–5–6–7–8–12	33	63	44	44
3	1	0.633	0.367	44.00	49.70	1–5–6–10–11–12	1–5–6–7–8–12	41	53	44	44
3	2	0.633	0.367	45.40	46.03	1–5–6–10–11–12	1–5–6–7–8–12	41	53	44	44

The estimation result is shown in Table 2. Since the sample size is sufficient for a statistical evaluation, a parameter has a statistical significance with 95% confidence level if the *t*-value is larger than 1.96. We can see that all of the parameters are statistical significant.

The estimation performance is evaluated by using the following statistical metrics, namely, the mean absolute percentage error (*MAPE*) and squared correlation coefficient (r^2). The smaller the values of *MAPE*, the closer are the predicted values to the observed values. r^2 provides an indication of the correlation between the predicted values and the observed values. A larger value closing to one suggests a better predictor.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{32}$$

$$r^2 = \frac{\left(n\sum_{i=1}^{n}\hat{y}_i y_i - \sum_{i=1}^{n}\hat{y}_i\sum_{i=1}^{n}y_i\right)^2}{\left(l\sum_{i=1}^{n}(\hat{y}_i)^2 - \left(\sum_{i=1}^{n}\hat{y}_i\right)^2\right)\left(l\sum_{i=1}^{n}(y_i)^2 - \left(\sum_{i=1}^{n}y_i\right)^2\right)} \tag{33}$$

where n is the total number of samples, y_i and $\hat{y}_i$ represent the actual and the estimated outputs, respectively.

To demonstrate the advantage of the proposed CO_2 emission model, the Virginia Tech Microscopic Energy and Emission Model (VT-Micro), Support Vector Machine (SVM) model and artificial neural network (ANN) model, are used for performance comparison.

VT-Micro was developed as a third-order regression model that estimates emission rates as a function of the speed and acceleration (Rakha et al., 2004).

$$MOE_e = \sum_{i=0}^{3}\sum_{j=0}^{3}(K_{i,j}^{e} v^i a^j) \tag{34}$$

where MOE_e is the fuel consumption rate, $K_{i,j}^{e}$ is the coefficient to be estimated, v is the instantaneous speed, and a is the instantaneous acceleration.

SVM is a novel supervised learning method used for regression (Cristianini and Shawe-Taylor, 2000). It can be adjusted to map the complex input–output relationship for the nonlinear system without dependent on the specific functions. Given a set of data points $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_l, y_l)\}$, SVM approximates the function using the following form:

$$f(\boldsymbol{x}) = \boldsymbol{\omega}\cdot\Phi(\boldsymbol{x}) + b \tag{35}$$

where $\Phi(\boldsymbol{x})$ denotes the high-dimensional feature spaces which are nonlinearly mapped from the input space $\boldsymbol{x}$, i.e., average speed, average acceleration, and angle of inclination of the road. The coefficients $\boldsymbol{\omega}$ and b are estimated by minimizing the regularized risk function:

$$\text{Minimize } \frac{1}{2}\|\boldsymbol{\omega}\|^2 + C\frac{1}{l}\sum_{i=1}^{l}L_{\varepsilon}(y_i, f(\boldsymbol{x}_i)) \tag{36}$$

$$L_{\varepsilon}(y_i, f(\boldsymbol{x}_i)) = \begin{cases} |y_i - f(\boldsymbol{x}_i)| - \varepsilon, & |y_i - f(\boldsymbol{x}_i)| \geqslant \varepsilon \\ 0 & otherwise \end{cases} \tag{37}$$

where $\|\boldsymbol{\omega}\|^2$ is the regularized term. Minimizing $\|\boldsymbol{\omega}\|^2$ will make a function as flat as possible, thus playing role of controlling the function capacity. l is the number of training samples. The second term $C\frac{1}{l}\sum_{i=1}^{l}L_{\varepsilon}(y_i, f(\boldsymbol{x}_i))$ is the empirical error measured by the ε-insensitive (ε-SV) loss function, which is defined by Vapnik (2000). The constant $C > 0$ determines the trade-off between the flatness of $f(\boldsymbol{x})$ and the amount up to which deviations larger than ε are tolerated.

The classical backpropagation neural network with one hidden layer of H hidden nodes, and logistic activation functions and one output node with a linear function will be used to establish the ANN model (Cortez et al., 2009). The structure of ANN model can be written as follows:

Table 2
Estimation result for the proposed CO_2 emission model.

Parameters	Coefficient	*t*-value	*P*-value
β_1	−2.68	−15.11	0.00
β_2	0.450	8.98	0.00
β_3	0.0000650	22.71	0.00
β_4	0.00411	20.12	0.00
β_5	0.266	9.55	0.00
β_6	0.533	74.83	0.00
β_7	2.77	12.69	0.00
Sample size	70,056		
R-square	0.82		

$$\hat{y}_{ANN} = w_{o,0} + \sum_{j=m+1}^{o-1} f\left(\sum_{i=1}^{m} x_i w_{j,i} + w_{j,0}\right) w_{o,i} \tag{38}$$

$$f(x) = \frac{1}{e^{-x}+1} \tag{39}$$

where $w_{j,i}$ denotes the weight of the connection from node j to i, and o is the output node. The number of hidden nodes (H) is determined by a regularization method (Hastie et al., 2001).

Fig. 11 shows the goodness-of-fit for the vehicle CO_2 emission predictions by the four models. It can be seen from the r^2 that the prediction of the proposed model is much closer to the observed target than the other two models. The MAPEs are 13.2%, 14.1%, 16.9% and 23.9%, respectively. The superiority of the proposed model is probably because we derive its structure from the theory of vehicle dynamic. Even though SVM model and ANN model also have an acceptable accuracy, it is difficult to get any insights on the structure of the function being approximated. The performance of VT-Micro model is not so satisfactory because its structure is not so consistent with the actual CO_2 emission model and the angle of inclination is neglected.

6.2. Sensitivity analysis for CO_2 emission model

To reveal how the selected factors affect the CO_2 emission, we propose the following simple methodology for sensitivity analysis. Once the parameters of the proposed model have been estimated on a large set of input variables, we calculate an average value for each input variable. Then, holding all variables at their average values but one each time, vary the one input over its entire range and analyze the variability produced in the outputs.

Here, sensitivities of the three explanatory variables are examined. Fig. 12(a) portrays a non-linear variation in the CO_2 emission. The minimal CO_2 emission appears to occur at the average speed of approximate 16.9 m/s (61 km/h). This finding indicates that an increase in average travel speed from 1 m/s to 16.9 m/s could result in a decrease of 39% in the CO_2 emission, while an increase in average travel speed from 16.9 m/s to 30 m/s could also result in an increase of 20% in the CO_2 emission. It indicates that driving too slowly or too fast will lead to high CO_2 emission. Fig. 12(b) illustrates the linear relationship between the average acceleration and CO_2 emission. It indicates that the CO_2 emission increases by 0.0307 g/m as a result of 0.1 m/s^2 increase in average acceleration. Fig. 12(c) shows how the angle of inclination of the road affects the CO_2 emission. Interestingly, the CO_2 emission increases slowly when the angle of inclination is negative, while it increases dramatically when the angle of inclination is between 0° and 15°. It indicates that driving in hilly area will result in significant fuel consumption and CO_2 emission.

To better understand the relative importance of each explanatory variable to the CO_2 emission, we conduct the elasticity analysis after estimating the parameters. Elasticity measures the percentage reaction of a dependent variable to a percentage change in an independent variable. The elasticity can be written as follows:

$$E_a = \frac{df(\boldsymbol{x})}{f(\boldsymbol{x})} \Big/ \frac{dx_a}{x_a} \tag{40}$$

where E_a denotes the elasticity for variable x_a, a denotes the subscript for each explanatory variable, $f(\boldsymbol{x})$ is the regression function. The output of Eq. (40) can be obtained by numerical solution after estimating the parameters by MLE.

Keeping all input variables at their average values except x_a which varies through its entire range with $i \in \{1, 2, \ldots, n\}$ intervals, the point elasticity (E_a^i) for each input x_a^i can be written as follows:

$$E_a^i = \frac{y_a^{i+1} - y_a^i}{x_a^{i+1} - x_a^i} \frac{x_a^i}{y_a^i} \tag{41}$$

If a given input variable ($x_a \in \{x_a^1, \ldots, x_a^n\}$) is relevant then it should produce a larger sum of absolute point elasticity ($|E_a^i|$). Thus, its relative importance (R_a) can be given by Cortez et al. (2009):

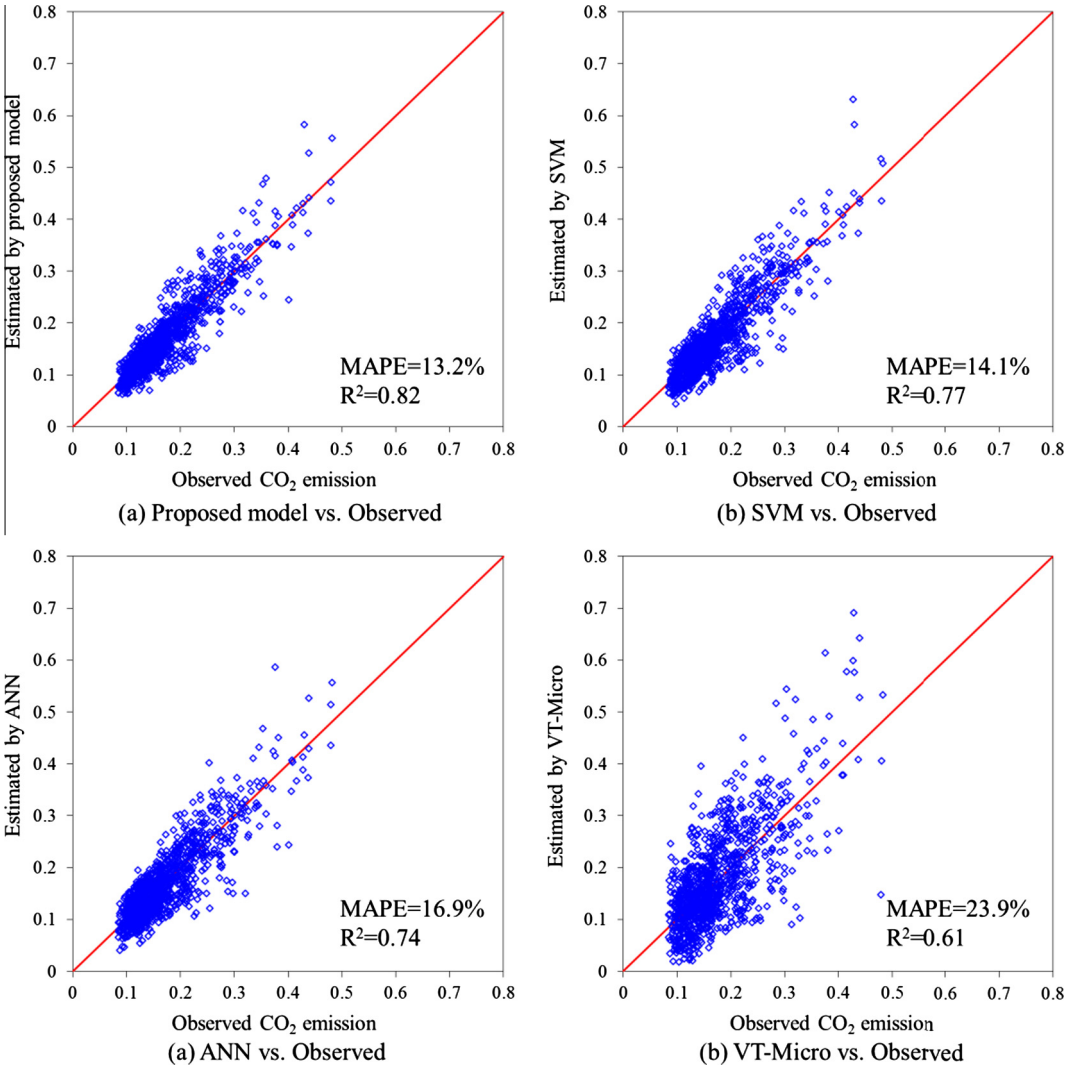

Fig. 11. Model performance comparison.

$$R_a = \frac{\sum_{i=1}^{n} |E_a^i|}{\sum_{a=1}^{m} \sum_{i=1}^{n} |E_a^i|} \tag{42}$$

where n is the number of samples, m is the number of explanatory variables.

The relative importance for explanatory variables in the proposed CO_2 emission model is shown in Fig. 13. It demonstrates that the average speed plays the most important role in vehicle CO_2 emission, while the angle of inclination of the road has less influence. The average speed and average acceleration occupy 85.6% relative importance to the model. It indicates that a vehicle with lower average acceleration running in around 61 km/h is likely to save fuel consumption and reduce the CO_2 emission. This ranking helps travelers determine an economic and eco-friendly path.

6.3. Performance analysis of eco-routing in a real-world network

In Fig. 14, a randomly selected real-world OD (origin–destination) pair in the central area of Toyota city is extracted to demonstrate the performance of the proposed eco-routing approach. For this demonstration, the travel time budget is set as 1.2 times the observed travel time. That means we allow the eco-routing model to find an eco-friendly path with 20% extra travel time relative to the observed path. As shown in Fig. 14 and Table 3, different routing strategies and their associated path performances are compared with the proposed eco-routing strategy. The shortest distance path takes a direct route through the central urban area with a high density of intersections, which may lead to great fluctuations in speed. The observed path selected by the driver has similar features to the shortest distance path. Even though the shortest path and

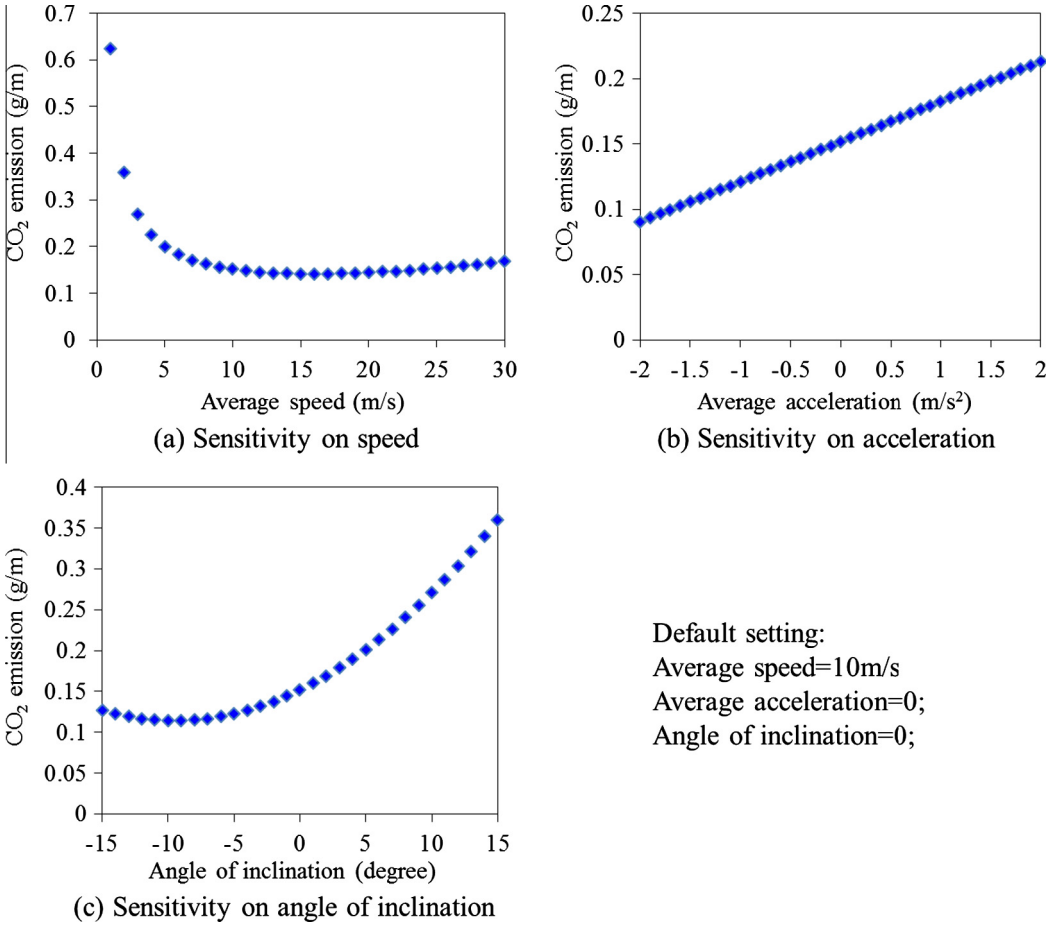

Fig. 12. Sensitivity analysis on explanatory variables.

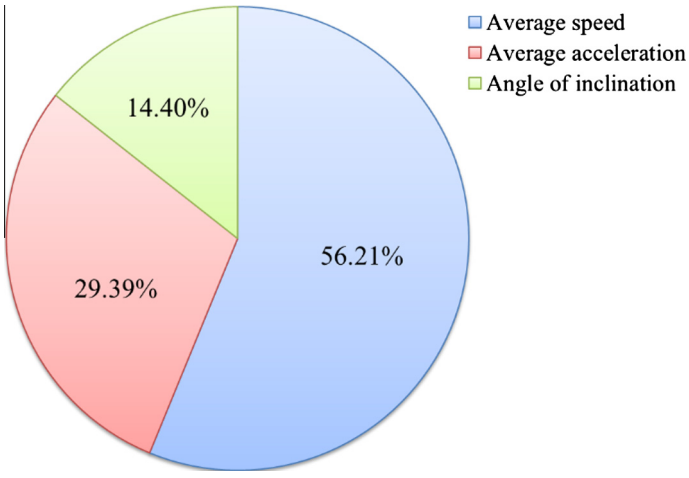

Fig. 13. Relative importance for each explanatory variable.

the observed path are shorter, the high density of intersections and high value of coefficient of variance (COV) of link average speed will result in higher CO_2 emissions as compared to the proposed eco-friendly path. On the other hand, the least travel time path detours onto expressways to profit from their higher average travel speed. As compared with the shortest distance path, the travel time saving is 10.26% even though the distance traveled is 1.16 times longer. However, this path generates the most CO_2 emissions because it is the longest among the four paths.

Fig. 14. Case study for eco-routing.

Table 3
Path performance comparison.

Performance indices	OP	DP	TP	EP	Relative difference (%)		
					EP vs. OP	EP vs. DP	EP vs. TP
Trip distance (km)	8.71	8.00	9.33	9.01	3.42	12.58	−3.43
Trip average speed (km/h)	17.23	20.52	26.66	21.45	24.53	4.54	−19.52
Intersection density (km^{-1})	8.15	8.75	7.50	5.88	−27.82	−32.75	−21.60
COV of link average speed	0.63	0.74	0.58	0.43	−31.62	−42.16	−26.21
Travel time (h)	0.51	0.39	0.35	0.42	−16.95	7.69	20.00
CO_2 emission (kg)	1.38	1.36	1.41	1.29	−6.98	−5.15	−10.17

Note: OP = Observed path; DP = Shortest distance path; TP = Least travel time path.
EP = Eco-friendly path; COV = Coefficient of variance.

As shown in Table 3, the eco-friendly path, determined as proposed in this study, offers significantly better performance in regard to CO_2 emissions at the cost of a very small increase in travel time and a minor detour. The eco-friendly path reduces CO_2 emissions by 6.98%, 5.15%, and 10.17% relative to the observed path, the shortest distance path, and the least travel time path, respectively. In particular, in comparison with the observed path as selected empirically by the driver, the eco-friendly path offers a significant advantage in both travel time (reduced by 16.95%) as well as CO_2 emissions (reduced by 6.98%), though the distance traveled is a little longer (by 0.3 km). We note that the saving in CO_2 emissions compared with the least travel time path naturally comes at the expense of increased travel time, because the objective of the routing algorithm is no longer to minimize travel time.

For a fixed route, e.g., the observed route in Fig. 14, we can adjust the driving speed to reduce the CO_2 emission if the traffic condition allows. As shown in Fig. 15, the CO_2 emission of the observed route will be the same as the eco-route if the average speed increases to 26.5 km/h. The CO_2 emission can further decrease to 1.01 kg if the average speed increases to 61 km/h. However, it is difficult to adjust the driving speed to a bigger value that greatly exceeds the average link speed in a non-free traffic flow condition. Therefore, it is necessary to plan an eco-route before departure.

6.4. Sensitivity analysis of potential CO_2 emission reduction

Because the benefit trade-off between CO_2 emissions and travel time budget is important both to individual travelers and to the provision of routing guidance, the network-wide reduction in CO_2 emissions that can potentially be realized with

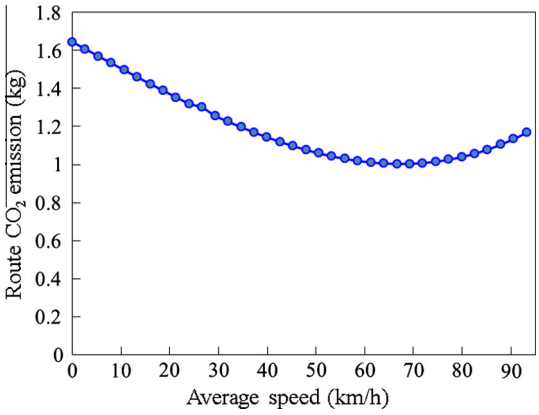

Fig. 15. CO_2 emission trend as average speed increases for a specified route.

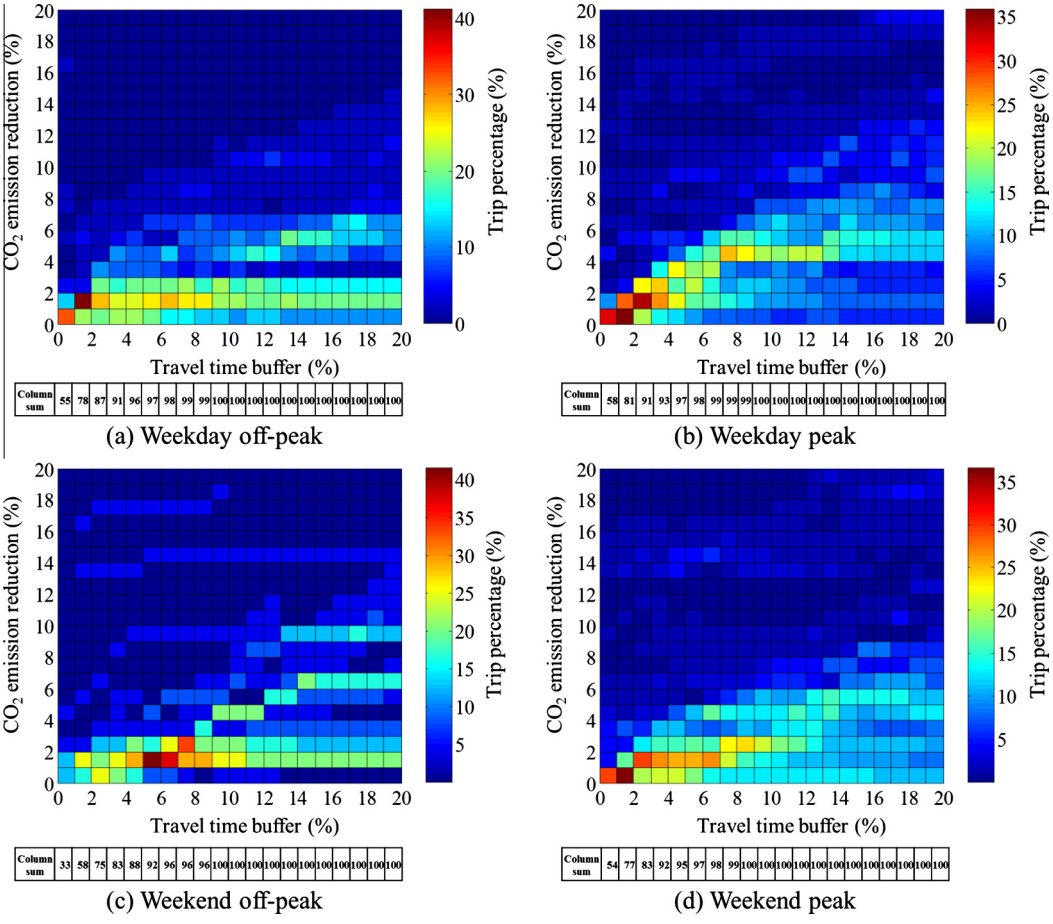

Fig. 16. Impact of travel time buffer on the percentage of trips with CO_2 emission reduction.

various settings of travel time budget is analyzed in this section. The concept of a travel time buffer is used, defined as the percentage increment over the least travel time. OD pairs collected from 7989 real-world trip records are used to conduct this sensitivity analysis. For each trip, the eco-routing model calculates eco-friendly paths under various travel time buffers, and then the potential reduction in CO_2 emissions is calculated by comparing these with the least travel time path.

Fig. 16 illustrates how many percentage of trips for which CO_2 emissions are reduced by a certain amount under different conditions. The *x*-axis denotes the percentage increment in travel time budget (the travel time buffer). The *y*-axis denotes the percentage reduction in CO_2 emissions compared to the least travel time path. The color of each cell denotes the percentage

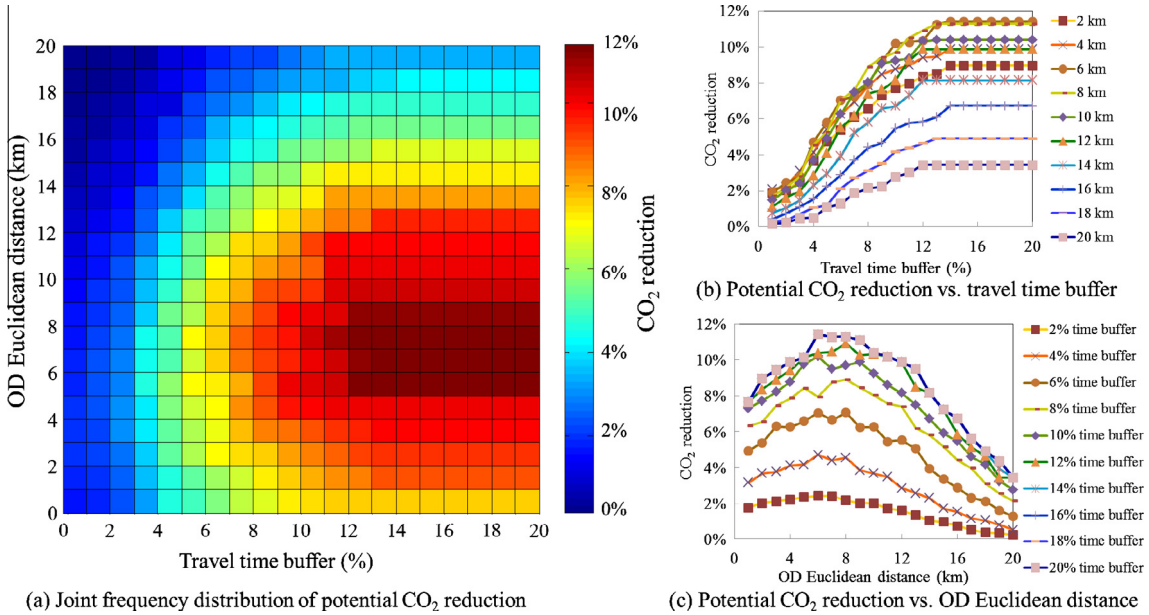

Fig. 17. Impact of travel time buffer and OD distance on CO_2 reduction.

of trips. The single row table below each sub-graph denotes the cumulative percentage of trips with that buffer value for which CO_2 emissions are reduced. This shows that the percentage of trips with reduced CO_2 emissions at the cost of a very small increase in travel time. However, the degree of the CO_2 reduction varies by trips. For example, some trips can achieve a reduction of 15% while others only reach 5%, even when the travel time buffer is increased to 20% or more. This indicates that eco-routing does not always provide a significant CO_2 reduction as compared to traditional routing based on travel time. Interestingly, if the travel time buffer is increased to 10%, almost all trips see some degree of CO_2 emissions reduction. This suggests that a relatively small increase in travel time enables network-wide CO_2 emissions to be effectively reduced by the eco-routing strategy. On the other hand, the cumulative percentage of the trips for which CO_2 emissions are reduced on weekdays increases faster with increasing travel time buffer than on weekends. This indicates that the overall reduction in CO_2 emissions on weekdays is more sensitive. Further, the cumulative percentage of trips for which CO_2 emissions are reduced in peak hours increases faster with increasing buffer than in off-peak hours. This suggests that the eco-routing strategy has a significant potential to reduce CO_2 emissions during peak hours.

Fig. 17(a) shows the joint percentage distribution of potential CO_2 reduction for different OD distances and travel time buffers. The cell color[1] denotes the percentage CO_2 emissions reduction compared to the least travel time path. In general, it is found that the eco-friendly path reduces CO_2 emissions by an average of 11% for OD distances between 6 km and 9 km when the travel time buffer is greater than 10%. This indicates that a travel time buffer of 10% is appropriate for the eco-routing strategy, since this provides the greatest CO_2 emissions reduction for the least cost in travel time. Fig. 17(b) shows the curves of potential CO_2 reduction as the travel time buffer increases. There is a significant rise in CO_2 emissions reduction as the travel time buffer increases from 1% to 12%. However, with further rises in travel time buffer, the CO_2 reduction remains relatively stable. That is, few paths with lower CO_2 emissions can be found once the travel time buffer increases beyond a certain threshold (e.g., approximately 12% in the studied network). On the other hand, compared to trips with longer OD distance, shorter trips have the potential for a larger CO_2 reduction percentage. For example, for trips with 6 km OD distance the potential reduction in average CO_2 emissions is from 2% to 11%, while for trips with 20 km OD distance the potential reduction is only from 0.2% to 3%. Fig. 17(c) shows the trend of potential CO_2 reduction percentage as OD Euclidean increases. The curves are shaped like a mountain for all buffer values. The CO_2 reduction rises to a peak for OD distances up to 8 km, then falls for OD distances greater than 8 km. Trips with larger OD distances are not as sensitive to CO_2 reduction as shorter trips; this is because a larger percentage of least travel time paths and eco-friendly paths may overlap due to the same choice of expressway or major road.

7. Conclusions and future work

This study proposes a vehicle dynamics based CO_2 emission model and an eco-routing approach to address the problem of finding the most eco-friendly path in terms of minimum CO_2 emissions constrained by a travel time budget. The method of Pareto-optimal optimization is introduced to solve this routing problem. The benefits of the proposed method are mainly

[1] For interpretation of color in Fig. 17, the reader is referred to the web version of this article.

from two aspects. First, comparing to microscopic CO_2 emission models such as CMEM and VSP, the proposed are more applicable to eco-routing problem because the input variables are average speed, average acceleration and angle of inclination. These variables are easy to measure from current transportation information systems. Second, comparing to traditional eco-routing method, the proposed method not only optimizes the CO_2 emission, but also take the travel time budget into account, which guarantees both on-time arrival and environmental friendliness or fuel economy.

Based on the results of a numerical experiment in Toyota city, Japan, the contribution and key findings are summarized as follows:

(1) The relative importance analysis indicates that the average speed and average acceleration occupy 85.6% relative importance to the CO_2 emission model.
(2) Eco-friendly path offers significantly reduced CO_2 emissions at little cost in terms of increased travel time and detours. On average, an eco-friendly path can reduce CO_2 emissions by 6.98%, 5.15%, and 10.17% relative to the observed path, the shortest distance path, and the least travel time path, respectively.
(3) Compared to the observed path as selected empirically by the driver, the eco-friendly path offers significant advantage in terms of travel time (reduced by 16.95%) and CO_2 emissions (reduced by 6.98%) though the travel distance is slightly longer (by 0.3 km).
(4) In an eco-routing experiment using all the observed OD pairs, it is found that the percentage of trips in which CO_2 emissions are reduced increases as the travel time buffer increases. Interestingly, when the travel time buffer reaches 10%, a certain degree of CO_2 emissions reduction is achieved for almost all trips.
(5) The average reduction in CO_2 emissions achieved by the eco-friendly path reaches a maximum of around 11% for trip OD distances between 6 km and 9 km and when the travel time buffer is around 10%. This indicates that setting a travel time buffer of 10% is appropriate for this eco-routing model, because this results in the greatest reduction in CO_2 emissions for the least cost in terms of travel time.

Potential directions for future research include extending the proposed CO_2 emission model by considering other related factors and estimating the network-wide vehicle emission. It appears to be significant benefit to a green transportation system and human health by applying the eco-routing navigation proposed in this study.

Acknowledgment

The first author would like to thank the China Scholarship Council (CSC) for financial support.

References

Ahn, K., Rakha, H., 2008. The effects of route choice decisions on vehicle energy consumption and emissions. Transp. Res. Part D: Transp. Environ. 13 (3), 151–167.
Ahn, K., Rakha, H., Trani, A., Van Aerde, M., 2002. Estimating vehicle fuel consumption and emissions based on instantaneous speed and acceleration levels. J. Transp. Eng. 128 (2), 182–190.
Ahuja, R.K., Magnanti, T.L., Orlin, J.B., 1993. Network Flows: Theory, Algorithms, and Applications. Prentice-Hall, Upper Saddle River, NJ, pp. 598–648.
Beasley, J.E., Christofides, N., 1989. An algorithm for the resource constrained shortest path problem. Networks 19 (4), 379–394.
Bektas, T., Laporte, G., 2011. The pollution-routing problem. Transp. Res. Part B: Methodol. 45 (8), 1232–1250.
Beusen, B., Broekx, S., Denys, T., Beckx, C., 2009. Using on-board logging devices to study the longer-term impact of an eco-driving course. Transp. Res. Part D: Transp. Environ. 14 (7), 514–520.
Birol, F., 2010. World Energy Outlook 2010. International Energy Agency.
Boriboonsomsin, K., Barth, M.J., Zhu, W., Vu, A., 2012. Eco-routing navigation system based on multisource historical and real-time traffic information. IEEE Trans. Intell. Transp. Syst. 13 (4), 1694–1704.
Boriboonsomsin, K., Dean, J., Barth, M., 2014. Examination of attributes and value of ecologically friendly route choices. Transp. Res. Rec.: J. Transp. Res. Board 2427 (1), 13–25.
Carlyle, W.M., Wood, R.K., 2005. Near-shortest and k-shortest simple paths. Networks 46 (2), 98–109.
Carlyle, W.M., Royset, J.O., Kevin Wood, R., 2008. Lagrangian relaxation and enumeration for solving constrained shortest-path problems. Networks 52 (4), 256–270.
Chen, P., Nie, Y.M., 2013. Bicriterion shortest path problem with a general nonadditive cost. Transp. Res. Part B: Methodol. 57, 419–435.
Chen, L., Yang, H., 2012. Managing congestion and emissions in road networks with tolls and rebates. Transp. Res. Part B: Methodol. 46 (8), 933–948.
Coe, E., 2005. Average Carbon Dioxide Emissions Resulting from Gasoline and Diesel Fuel Tech. rep. United States Environmental Protection Agency.
Cortez, P., Cerdeira, A., Almeida, F., Matos, T., Reis, J., 2009. Modeling wine preferences by data mining from physicochemical properties. Dec. Supp. Syst. 47 (4), 547–553.
Coutinho-Rodrigues, J.M., Clímaco, J.C.N., Current, J.R., 1999. An interactive bi-objective shortest path approach: searching for unsupported nondominated solutions. Comput. Oper. Res. 26 (8), 789–798.
Cristianini, N., Shawe-Taylor, J., 2000. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press.
Csikos, A., Varga, I., Hangos, K.M., 2015. Modeling of the dispersion of motorway traffic emission for control purposes. Transp. Res. Part C: Emerg. Technol.
Current, J.R., Revelle, C.S., Cohon, J.L., 1990. An interactive approach to identify the best compromise solution for two objective shortest path problems. Comput. Oper. Res. 17 (2), 187–198.
Dijkstra, E.W., 1959. A note on two problems in connexion with graphs. Numer. Math. 1 (1), 269–271.
EPA, 2000. Average Annual Emissions and Fuel Consumption for Passenger Cars and Light Trucks. Transportation and Air Quality.
Ericsson, E., Larsson, H., Brundell-Freij, K., 2006. Optimizing route choice for lowest fuel consumption–potential effects of a new driver support tool. Transp. Res. Part C: Emerg. Technol. 14 (6), 369–383.
Franceschetti, A., Honhon, D., Van Woensel, T., Bektaş, T., Laporte, G., 2013. The time-dependent pollution-routing problem. Transp. Res. Part B: Methodol. 56, 265–293.

Frey, H.C., Yazdani-Boroujeni, B., Hu, J., Sandhu, G., Liu, B., Jiao, W., Graver, B., 2012. Field measurements of 1996 to 2013 model year light duty gasoline vehicles. In: Proceedings of the Air and Waste Management Association's Annual Conference and Exhibition, Chicago, June 24–28, 2013.
Guo, L., Huang, S., Sadek, A.W., 2013. An evaluation of environmental benefits of time-dependent green routing in the greater Buffalo–Niagara region. J. Intell. Transp. Syst. 17 (1), 18–30.
Guo, C., Yang, B., Andersen, O., Jensen, C.S., Torp, K., 2014. EcoMark 2.0: empowering eco-routing with vehicular environmental models and actual vehicle fuel consumption data. GeoInformatica, 1–33.
Hart, P.E., Nilsson, N.J., Raphael, B., 1968. A formal basis for the heuristic determination of minimum cost paths. IEEE Trans. Syst. Sci. Cybernet. 4 (2), 100–107.
Hastie, T., Tibshirani, R., Friedman, J., 2001. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer-Verlag, NY.
Hendricks, E., Sorenson, S.C., 1990. Mean Value Modeling of Spark Ignition Engines. SAE Technical Paper Series, Paper No. 900616.
JAMA, 2008. Reducing CO_2 Emissions in the Global Road Transport Sector <http://www.jama-english.jp/>.
Jimenez-Palacios, J.L., 1998. Understanding and Quantifying Motor Vehicle Emissions with Vehicle Specific Power and TILDAS Remote Sensing Doctoral Dissertation. Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA.
Kirschstein, T., Meisel, F., 2015. GHG-emission models for assessing the eco-friendliness of road and rail freight transports. Transp. Res. Part B: Methodol. 73, 13–33.
Koc, C., Bektaş, T., Jabali, O., Laporte, G., 2014. The fleet size and mix pollution-routing problem. Transp. Res. Part B: Methodol. 70, 239–254.
Kono, T., Fushiki, T., Asada, K., Nakano, K., 2008. Fuel consumption analysis and prediction model for "Eco" route search. In: 15th World Congress on Intelligent Transport Systems and ITS America's 2008 Annual Meeting.
Long Cheu, R., Xie, C., Lee, D.H., 2002. Probe vehicle population and sample size for arterial speed estimation. Comp.-Aided Civil Infrastruct. Eng. 17 (1), 53–60.
Masikos, M., Demestichas, K., Adamopoulou, E., Theologou, M., 2015. Energy-efficient routing based on vehicular consumption predictions of a mesoscopic learning model. Appl. Soft Comput. 28, 114–124.
Mensing, F., Bideaux, E., Trigui, R., Ribet, J., Jeanneret, B., 2014. Eco-driving: an economic or ecologic driving style? Transp. Res. Part C: Emerg. Technol. 38, 110–121.
Miwa, T., Kiuchi, D., Yamamoto, T., Morikawa, T., 2012. Development of map matching algorithm for low frequency probe data. Transp. Res. Part C: Emerg. Technol. 22, 132–145.
Nazemi, A., Omidi, F., 2013. An efficient dynamic model for solving the shortest path problem. Transp. Res. Part C: Emerg. Technol. 26, 1–19.
Nie, Y.M., Li, Q., 2013. An eco-routing model considering microscopic vehicle operating conditions. Transp. Res. Part B: Methodol. 55, 154–170.
Pandian, S., Gokhale, S., Ghoshal, A.K., 2009. Evaluating effects of traffic and vehicle characteristics on vehicular emissions near traffic intersections. Transp. Res. Part D: Transp. Environ. 14 (3), 180–196.
Park, S., Rakha, H., 2006. Energy and environmental impacts of roadway grades. Transp. Res. Rec.: J. Transp. Res. Board 1987, 148–160.
Rajamani, R., 2011. Vehicle Dynamics and Control. Springer, Berlin.
Rakha, H., Ahn, K., Trani, A., 2004. Development of VT-Micro model for estimating hot stabilized light duty vehicle and truck emissions. Transp. Res. Part D: Transp. Environ. 9 (1), 49–74.
Rusli, N., Majid, M.R., Din, A.H.M., 2014. Google Earth's derived digital elevation model: a comparative assessment with Aster and SRTM data. In: IOP Conference Series: Earth and Environmental Science, vol. 18(1), pp. 1–6.
Sedeno-Noda, A., Raith, A., 2015. A Dijkstra-like method computing all extreme supported non-dominated solutions of the biobjective shortest path problem. Comput. Oper. Res. 57, 83–94.
Sivak, M., Schoettle, B., 2012. Eco-driving: strategic, tactical, and operational decisions of the driver that influence vehicle fuel economy. Transp. Policy 22, 96–99.
Srinivasan, K., Jovanis, P., 1996. Determination of number of probe vehicles required for reliable travel time measurement in urban network. Transp. Res. Rec.: J. Transp. Res. Board 1537, 15–22.
Sun, J., Liu, H.X., 2015. Stochastic eco-routing in a signalized traffic network. Transp. Res. Part C: Emerg. Technol. 59, 32–47.
Tan, Z., Gao, H.O., 2015. Traffic control for air quality management and congestion mitigation in complex urban vehicular tunnels. Transp. Res. Part C: Emerg. Technol. 58, 13–28.
Vapnik, V., 2000. The Nature of Statistical Learning Theory. Springer.
Xie, C., Waller, S.T., 2012. Optimal routing with multiple objectives: efficient algorithm and application to the hazardous materials transportation problem. Comp.-Aided Civil Infrastruct. Eng. 27 (2), 77–94.
Yao, E., Song, Y., 2013. Study on eco-route planning algorithm and environmental impact assessment. J. Intell. Transp. Syst. 17 (1), 42–53.
Yen, J.Y., 1971. Finding the k shortest loopless paths in a network. Manage. Sci. 17 (11), 712–716.
Zeng, W., Miwa, T., Wakita, Y., Morikawa, T., 2015a. Application of Lagrangian relaxation approach to α-reliable path finding in stochastic networks with correlated link travel times. Transp. Res. Part C: Emerg. Technol. 56, 309–334.
Zeng, W., Miwa, T., Wakita, Y., Morikawa, T., 2015b. Exploring trip fuel consumption by machine learning from GPS and CAN bus data. J. East. Asia Soc. Transp. Stud. 11, 906–921.
Zeng, W., Miwa, T., Morikawa, T., 2016. Application of hyperpath strategy and driving experience to risk-averse navigation. IET Intell. Transp. Syst. http://dx.doi.org/10.1049/iet-its.2015.0065.