European Journal of Operational Research 306 (2023) 1456–1472

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Innovative Applications of O.R.

Fuel-optimal truck path and speed profile in dynamic conditions: An exact algorithm

David P. Watling[a,*], Richard D. Connors[a,b], Haibo Chen[a]

[a] Institute for Transport Studies, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, UK
[b] Department of Engineering, University of Luxembourg, Esch-sur-Alzette, Luxembourg

ARTICLE INFO

Article history:
Received 18 January 2021
Accepted 16 July 2022
Available online 22 July 2022

Keywords:
Transportation
Space-time network
Path choice optimization
Dynamic information
Fuel consumption

ABSTRACT

We consider optimizing a truck's choice of path and speed profile to minimise fuel consumption, exploiting real-time predictive information on dynamically varying traffic conditions. Time-varying traffic conditions provide particular challenges, both from network-level interactions (e.g. slowing to consume more fuel locally may be beneficial to avoid congested periods downstream) and link-level phenomena (e.g. interaction between acceleration and gradient profiles). A multi-level, discrete-time decomposition of the problem is presented in which: (i) [sub-problems] speed profiles are optimized within each link, given boundary conditions of entry/exit times and speeds; (ii) [master problem] a space-time extended network representation is used to encode the dynamic interactions, within which the joint choice of path and speed profile is made. By instantiating the space-time network in (ii) with the optimal link profiles from (i), we are able to devise a tractable algorithm while optimizing speed profiles over a fine timescale. The solution approach is to pre-solve offline the computationally-intensive step (i), meaning that the representation in (ii) can be efficiently produced online in response to the real-time predictive information, whereby optimization of the path and speed profile is solved by a single shortest path search in the space-time network, for which many exact algorithms exist. The method is extended to additionally consider choice of discretionary stops and (pre-trip) departure time. Two representations are presented and investigated, depending on whether constraints are additionally imposed to ensure consistency of speed profiles across link boundaries. Numerical experiments are reported on a small illustrative example and a case-study network.

1. Introduction

There are several strong, real-life motivations for the problem considered in the present paper. Fuel combustion from transport is responsible for around a quarter of CO_2 emissions globally (IEA, 2018). Transport by road is by far the major contributor, and the trend has been increasing (Sims et al., 2014; SloCaT, 2018). While improvements in vehicle and fuel technology are making some impact, the trend is clearly an upward one, and one which technology alone will be unable to solve, given for example increases in vehicle-km (BEIS, 2017). Freight transport consumed nearly 45% of the total energy from transport in 2009, over half of this due to Heavy Goods Vehicles (HGVs) (Sims et al., 2014). With increasingly long journeys being made by HGVs, there are therefore potentially large gains to be made in the road freight industry in terms of reduced energy consumption and environmental impacts, with the potential interventions going beyond improvements to vehicle technology.

A first step towards realising this goal, then, is to understand the causal factors behind fuel consumption for HGVs (other than vehicle technology), and particularly to identify the complex chain of *interactions* between these factors. A natural place to begin describing this chain is the impact of road gradient, and the evidence of its strong interacting influence with the speed at which a truck is driven (Bandeira et al., 2018; Boriboonsomsin & Barth, 2009; Fröberg et al., 2006; He et al., 2016). Since we clearly cannot affect the gradient profile of a given path, but nevertheless may choose an alternative path when available, there is unsurprisingly evidence that path choice may also significantly affect overall fuel consumption for a journey (Scora et al., 2015). Alternatively, there is extensive evidence of the beneficial impacts on fuel consumed of modifying driving style (in terms of the speed/acceleration pro-

* Corresponding author.
E-mail address: d.p.watling@its.leeds.ac.uk (D.P. Watling).

https://doi.org/10.1016/j.ejor.2022.07.028

file followed), whether achieved through driver training (Ayyildiz, Cavallaro, Nocera & Willenbrock, 2017; Wu et al., 2018; Zavalko, 2018), on-board driver assistance/guidance systems (Huang et al., 2018; Vagg et al., 2013), or autonomously driven trucks (Nasri et al., 2018).

These factors all suggest general strategies that might be adopted to reduce fuel consumption, but the deployment of these strategies needs to be sensitive to the features of the particular network, road and traffic conditions within which a truck is driven, the characteristics of the vehicle and its load, and the availability of information for making decisions; i.e. a case-specific optimization approach is needed. This view is partially supported by the fact that general eco-driving training has seen to be widely varying in its effectiveness (Alam & McNabola, 2014). The need for a case-specific methodology is also highlighted in the contemporary debate over permitting longer and heavier trucks (Diaz-Ramirez et al., 2017; McKinnon, 2011; Rodrigues et al., 2015), which alters the trade-off between speed and choice of path due to the greater sensitivity of such trucks to burning large amounts of fuel in acceleraton events. Since such trucks may also be making longer journeys, there will clearly be more possibility for unexpected events to arise relative to the initially-planned itinerary. This in turn raises the positive possibility to influence the path choice and speed decision problem through real-time traffic information, where it is available (He et al., 2016; Jiménez et al., 2013).

In developing the approach in this paper, we wish to build on the findings and advances made in several previous relevant areas of research. A wide range of methods has now been developed for eco-driving, namely how a vehicle may be *optimally* driven along a given path to minimize fuel/environmental impacts (Dehkordi et al., 2019; Held et al., 2019; Hellström et al., 2009; Ozatay et al., 2017; Saboohi & Farzaneh, 2009; Saerens et al., 2013; Zeng & Wang, 2018). Such methods use detailed spatio-temporal models of speed and acceleration events in order to track fuel consumed during the journey. To some extent at the other extreme of the decision problem, several authors have considered the strategic planning of a vehicle's tour sequence and schedule around a number of pick-up/drop-off locations, as characterised by the Pollution Routing Problem and variants thereof (Bektas & Laporte, 2011; Franceschetti et al., 2013, 2017; Kramer et al., 2015; Kuo, 2010; Lin et al., 2014; Qian & Eglese, 2016). In these approaches there is no explicit representation of the paths available in the underlying road network, with each pair of locations joined by a single 'link'. While the travel time on each link may be time-dependent, there is no representation of acceleration events, so that (in these tour-planning approaches) fuel consumption is assumed to be a function of mean speed only.

At an intermediate level between the two classes of problem described above (detailed speed control, as in eco-driving, and complete tour re-planning), there are a handful of methods that consider the adaptation of a vehicle's *path*, such as between a sequential pair of pick-up/drop-off locations on a vehicle's tour; this focus is precisely the topic of the present paper. We review these in some detail, in order that the particular challenges of the problem may be highlighted, and the contributions of our own research may be exemplified. Nie and Li (2013) identify the important influence of acceleration events for fuel consumption, but note the difficulty in explicitly including such detailed factors in path-planning. They propose instead to approximate its effect by assuming a simple characteristic speed profile for each link, whereby a vehicle smoothly accelerates/decelerates following its entry to a link, to adjust to the assumed, pre-defined, time-independent 'cruising speed' of that link. Under such an assumption, vehicles have no *choice* of speed. The problem then considered is how to choose an optimal path for a vehicle to minimise the travel cost (a weighted sum of fuel and travel time), subject to an environmental constraint on CO_2 emissions being satisfied. The path choice problem must consider not only link-related factors, but 'movement' factors between links (as the vehicle changes speed). The resulting problem is a kind of constrained shortest path problem, which is known to be NP-complete. While illustrating the problem on a small example, they notice the difficulties in developing algorithms for general networks. Miao et al. (2018) considered a detailed powertrain-based model of fuel consumption, motivated by their interest in the specific effects of traffic signals in an urban setting. They discuss the difficulties in applying conventional shortest path methods to such problems due to the inter-dependent impacts of factors such as those discussed earlier (speed, gradient, etc.), and due to spatial inter-dependencies arising from the gear-shifting schedule or from the vehicle's management system. They go on to set out a problem of joint optimization of a vehicle's path and speed profile, subject to an upper bound on travel time. The optimization is constrained by instantaneous real-time estimates of point speeds at detector locations. A heuristic (genetic) algorithm is proposed for solving the problem, and a restricted form of the approach suggested for real-time applications, with the methods tested by implementing them in a traffic simulator.

This brings us to explaining a particular technical challenge that is addressed by the present paper, and which is central to understanding our own contribution, but which has either been avoided in previous work or has been addressed but at the cost of making other undesirable assumptions. Taking the work of Miao et al. (2018) as a reference point, it is noticeable that while this method may be implemented conditionally on time-dependent data, which would give paths that potentially vary by departure time, the method employs *instantaneous constraints* on vehicle speeds. That is to say, for a given path and departure time t, it is assumed that the speed profile along the path is constrained by point speeds along the path *at current time t*, rather than being constrained by the predicted (future) speeds at the time the vehicle would pass each downstream detector based on its speed profile to that point (a phenomenon that we shall refer to as *predictive constraints*). Time-dependent congestion means that instantaneous and predictive constraints may differ markedly, both in urban contexts (where congestion is more severe, causing a faster variation in time of congestion) and interurban contexts (where trips are longer, so that roads are traversed at very different times of day). Predictive constraints are, however, substantially more complex to incorporate than instantaneous constraints, since they introduce a kind of 'circularity' – in order to determine an optimal speed profile for a given path, we must define the speed constraints in the optimization, but to define the active time-dependent constraints at a given location, we must know the speed profile. This problem has been solved for a single path with an exact, quadratic-time algorithm (Hvattum et al., 2013). However, the problem greatly increases in complexity when speed profile optimization is combined with path choice, since the optimal link speeds are then in general *path-dependent*, and then the challenge is to avoid the computational burden of having to enumerate all paths.

The time-dependent path and speed profile optimization problem is thus highly challenging when combined with predictive constraints, and while previous authors have proposed methods that circumvent it, this is at the price of a significant loss in fidelity. In particular, they require the adoption of a *simplified representation* in which the detailed acceleration and gradient influences on fuel consumption noted earlier are neglected, with vehicles travelling at a speed that, while dependent on link entry time, is constant along a link's length. Such a simplified formulation has been used to develop several heuristic methods for determining fuel-optimal *tours* with speed optimization in a time-dependent setting. These methods include the tabu-search method proposed by Jabali et al. (2012) and the neighbourhood move heuristics pro-

posed by Qian and Eglese (2016). Huang et al. (2017) recently showed how such a formulation might be used to develop heuristics for the combined problem of tour, path, and speed optimization. In Watling et al. (2019), this simplified formulation was adopted to develop an algorithm for path and speed optimization, based on a time-discretisation of the underlying problem. Zeng et al. (2020) showed that under such an assumption of constant link speeds, a solution could be developed for finding minimum fuel paths under the constraint of a travel time budget and on-time arrival probability.

Our particular contribution, then, will be to address the problem of finding a fuel-optimal path and speed strategy for HGVs making long inter-urban journeys, while taking account of the detailed spatial-temporal interactions between acceleration, gradient and fuel (i.e. avoiding the simplified representation described above). Our work, therefore, has a similar starting point to that of Nie and Li (2013) and Miao et al. (2018). Similarly to Miao et al., we wish to represent the impact of detailed acceleration events on fuel consumption; however our different focus on inter-urban journeys and HGVs means our interest is more on the detailed interaction of acceleration events and gradient profiles, rather than their (urban-focused) study of gears and traffic signals. Compared to their urban context, the links we will use are thus potentially relatively long, reflecting the relative sparsity of path options for large heavy vehicles on inter-urban journeys, and thus we must include considerable detail within a link (e.g. a gradient *profile* rather than the assumed constant link gradient of Miao et al.). Unlike Nie and Li (2013) and Miao et al. (2018) we wish to handle time-dependent *predictive constraints* (as defined above), rather than assuming the simpler instantaneous constraints. However, we wish to do so without compromising on the fidelity of the speed recommendations and fuel consumption model, i.e. avoiding the constant speed assumptions of Huang et al. (2017) and Watling et al. (2019). A distinctive element of our method is that we wish to exploit real-time predictive information on time-dependent travel times, as is commonly available now in many map-based systems. Given our focus on long freight journeys, an additional important optimization dimension we wish to include is that of discretionary breaks, to avoid particularly serious congestion events. Unlike the previously cited works, we wish to develop an exact algorithm with guaranteed optimality, rather than a heuristic method. Nevertheless it must be sufficiently efficient to be implementable in a real-time context.

In summary, then, the contributions of our work will be:

- addressing the joint problem of optimizing path choice and speeds for an agent in a network in order to minimise fuel consumption;
- handling detailed speed/acceleration profiles in this framework, in order to realistically predict the impact of acceleration events and gradients on fuel consumption;
- dealing with predictive constraints on vehicle speeds in this time-dependent framework, rather than the simpler but less realistic instantaneous constraints;
- integrating the influence of real-time traffic information on time-dependent travel times;
- developing an exact algorithm (for its discretized form), as opposed to a heuristic method.

The paper is organised by explaining in a systematic order the key components of the approach. The formal definition of the problem is first given in Section 2, and a decomposition and then discretisation then derived. A fundamental element of the decomposition approach is its ability to solve, off-line, boundary optimization problems for each link independently. In Section 3 we investigate these sub-problems in detail. In Section 4, the solution approach is defined for both a relaxed form of the discretised problem and a non-relaxed form. A simple network example is employed in Section 5 to illustrate key features of the problem, approach and its solutions. A realistic application is presented in Section 6, based on the UK road network and using predictive travel time data harvested from a live travel information system. Finally, in Section 7, we summarise the main findings and conclusions of the research, and identify potential future research developments.

2. Formal problem definition and mathematical formulation

In the present section we formally describe the overall network-level problem to be considered, and then describe the specific way in which it is formulated for the purpose of solution.

2.1. Assumptions

The particular application context we have in mind is that of a large (e.g. 40T) truck making a long, inter-urban freight journey. While the vehicle's complete tour may typically involve several pickup/drop-off locations, we presume the sequence and planned schedule to visit these locations has been exogenously determined; the present paper considers only a given neighbouring pair of pickup/drop-off locations (i.e. a single 'leg') in the tour, and neglects any impacts on the rest of the tour. That is to say, our problem is not to optimize the *tour*, but to optimize the *path* for a single leg of the tour. The main motivation is to use real-time predictive traffic information to optimize fuel consumption for the tour leg. During the journey, we consider the possibility to dynamically adapt the speed, path and stops (discretionary breaks) for the truck; prior to the journey commencing, there is the additional possibility to modify the time of departure. These are therefore the decision variables for our optimization problem. At any given clock-time, the vehicle will either have not commenced its journey (the *pre-trip* case in which the 'origin' is set to be the start-point of the journey leg), or is *en route* (in which case the 'origin' is set to be the next downstream node on the vehicle's current path). With such a representation both cases can be handled with the same generic method, with the only proviso that choice of departure time is only an option in the pre-trip case.

It is assumed that the following characteristics are known:

(a) the features of the truck—in terms of its vehicle type, height, width and weight— as well as the load it will be transporting on the considered leg of the tour; and
(b) any road restrictions on routes relevant to the truck's journey, relating to a vehicle's height, width and axle load.

Taking (a) and (b) together, it is possible to define a permitted road network $G=\{V,E\}$ for this truck to use for the leg, where V and E respectively denote the node set and directed link set for the graph. These nodes represent locations at which potential diversion paths or stop locations occur. For a given {truck, tour-leg, load}, the graph $G=\{V,E\}$ thus defines a fixed set of information to allow the spatial feasibility of any path to be defined.

As such a truck makes its (typically long) journey leg, it is highly likely to encounter traffic conditions that regularly vary by time-of-day, e.g. as it passes near to urban centres at peak times where the inter-urban network is shared with urban traffic. In addition, it may be delayed by unexpected incidents, weather or special events, meaning that the time-varying delays may be routine or special to that journey. We suppose that at any given clock-time, real-time predictive information provides the predicted *minimum* travel time for the given vehicle to traverse a link, *for all future entry-times to a link* that are potentially relevant to the journey, and for all directed links in E. That is to say, the information is time-varying in two ways; different information will be available

at different clock-times, but also the predictive information available at any clock-time will differ by the predicted entry-time to a link. These predicted minimum travel times are exogenously determined, and can account for: (i) historical information including from previous days, giving a forecast of the typical time-of-day variation in travel times throughout the day, due to congestion and/or variations in demand; (ii) real-time information, e.g. on incidents, and the forecast impacts through a journey time prediction algorithm; (iii) fixed constraints due to vehicular/legal considerations, i.e. the minimum travel time is not less than the time it would take to traverse the link at minimum of the truck's instantaneous speed limit and the legal limit for the link for such a vehicle. When combined with the permitted road network $G=\{V,E\}$, the predicted minimum travel times in principle allow the space-time feasibility to be judged of our decision variables, namely start-time, path, link traversal speeds and discretionary stops.

Before continuing it is useful to clarify our assumptions regarding the information assumed available. Our method is developed on the premise that, at the time the path is planned, there is predicted information on the future travel times that will be experienced on each link of the path, at the (future) time that link is traversed. Several possibilities could be envisaged (with t_0 denoting the time that the path planning decision is made):

1. Information is based on the best real-time estimates of prevailing link travel times at time t_0. At time t_0, these would therefore not depend on the future time that the vehicle is predicted to enter each link, so in our sense we do not have time-dependent *predictive* information.
2. Information is based on historic data (from relevant previous days, combined in some way), to yield estimates of travel times for the whole of the present day, disaggregated by time of day. At time t_0, we are therefore able to capture the dependence on all future times that a vehicle may enter a link, so in our sense we have time-dependent predictive information.
3. Information is based on future forecasts of how today's traffic conditions will unfold in the future, based on current real-time estimates; thus, in the case of incidents for example, we would need to predict incident duration and impacts. At time t_0, we therefore again have time-dependent predictive information.

Other possibilities may be a combination of cases 1 and 3, or of cases 2 and 3. The method we shall describe is relevant to all such cases.

Thus far we have considered the network/spatial constraints and the space-time constraints for the journey. In addition, there are 'physical' within-link constraints which give upper and lower bounds on both the instantaneous speed and instantaneous acceleration of the vehicle. Clearly the constraints of type (iii) (as defined above) applied at the network-level also impose a kind of speed constraint, in that the minimum travel time to traverse a link gives a maximum traversal speed. Including the possibility of constraints on both instantaneous speed and the link traversal speed allows for the possibility to include a less strict constraint on instantaneous speed, e.g. setting the maximum instantaneous speed to be 10% higher than the legal speed limit for the link, and the maximum link traversal speed to be equal to the legal speed limit.

2.2. Formal problem definition

Here we give the mathematical definition of the problem to be addressed. Firstly, the basic *network-level* notation we require is as follows, with clock-time represented by the real variable $t \geq 0$:

$d \in \mathbb{R}$ $(d \geq 0)$ clock-time of the next decision-point, i.e. either the departure time from the origin, pre-trip, or the expected arrival time at the next downstream node, *en route*

$G=\{V,E\}$ = relevant permitted road network as defined in Section 2.1

i^* $(\in V)$ = starting node at which the current decisions are to be made, i.e. either the origin of the trip leg, pre-trip, or the next downstream node on the current path, if *en route*

j^* $(\in V)$ = destination node for the considered journey leg

$\tau_{ij,\min}(s,t)$ = at current clock-time $s \in \mathbb{R}$ $(s \geq 0)$, the predicted minimum travel time to traverse link $(i,j) \in E$ for a truck of the given type, if entering the link at (future) clock-time $t \in \mathbb{R}$ $(t > s)$, given real-time predictive information available at time s

For any distinct pair of nodes $i^* \neq j^* \in V$, denote the following feasible set of ordered node-sequences:

$$\mathcal{P}_{i^*j^*} = \{\mathbf{i} = (i_1, i_2, i_3, \ldots, i_n) : i_1 = i^*,\ i_n = j^* \ \&\ i_1 \to i_2 \to i_3 \to \ldots \to i_n \text{ is an acyclic path in } G = \{V, E\}\}.$$

A vehicle departing at time d from i^* and following such a path has some flexibility in choosing the travel time to traverse each link (by varing their speed), and thereby has some choice over the time at which each node on the path is passed. This therefore motivates the introduction of the notion of a *space-time path*, a sequence of nodes in $\mathcal{P}_{i^*j^*}$ each of which is additionally indexed by the (continuous) time it is passed on the path that begins at time d at node i^*:

$$\tilde{\mathcal{S}}_{i^*j^*d} = \left\{(\mathbf{i},\mathbf{t}) = \left((i_1,t_{i_1}),(i_2,t_{i_2}),\ldots,(i_n,t_{i_n})\right) : \mathbf{i} \in \mathcal{P}_{i^*j^*};\ \mathbf{t} \in \mathbb{R}^n;\ t_{i_n} > t_{i_{n-1}} > \ldots > t_{i_1} = d\right\}.$$

To be a *feasible* space-time path, it is required that any such t is chosen in a way to be consistent with the dynamic constraints imposed by real-time traffic information on link travel times, this information varying both by the time (s) at which the decision is made and the times (t) at which each node is passed (and therefore the time at which each subsequent link is entered). This gives rise to the following necessary conditions for feasibility (1), based on the minimum times to traverse the link, as predicted at current clock-time s:

$$t_{i_{k+1}} - t_{i_k} \geq \tau_{i_k i_{k+1},\min}(s, t_{i_k}) \ \ (k = 1,2,\ldots,n-1; s \in \mathbb{R},\ s \geq 0). \quad (1)$$

Link-level notation is now introduced in order to generate physical constraints, as well as leading to the definition of the objective function. Firstly, fixed link characteristics are denoted:

L_{ij} = length of link $(i,j) \in E$ in metres

$v_{ij,\text{MIN}}$ = minimum instantaneous permitted speed on link $(i,j) \in E$

$v_{ij,\text{MAX}}$ = maximum instantaneous permitted speed on link $(i,j) \in E$

$a_{ij,\text{MIN}}$ = maximum possible deceleration on link $(i,j) \in E$

$a_{ij,\text{MAX}}$ = maximum possible acceleration on link $(i,j) \in E$.

Secondly, time-dependent kinetic functions are defined; by later construction of the problem, it will be ensured that these will only be applied at relevant times (i.e. at times at which a link is being used) and will have relevant physical constraints applied (i.e. on link length and speed/acceleration limits):

$x_{ij}(t)$ = *location (continuous distance from entry point in metres) of vehicle on link* $(i,j) \in E$ *at clock-time* $t \geq 0$

$v_{ij}(t)$ = *speed (m/s) of vehicle on link* $(i,j) \in E$ *at clock-time* $t \geq 0$

$a_{ij}(t)$ = *acceleration (m/s^2) of vehicle on link* $(i,j) \in E$ *at clock-time* $t \geq 0$.

Thirdly, features are defined relevant to the objective function:

$\boldsymbol{\theta}_{ij}(x)$ = vector of fuel consumption parameters relevant to the location x metres from the entry to link $(i,j) \in E$, for $0 \leq x \leq L_{ij}$, $x \in \mathbb{R}$, containing both static features related to constant or typical features of the vehicle (e.g. drive-train efficiency, aerodynamic

drag) and location-dependent features related to the road environment (e.g. gradient, road surface)

$F_{ij}(v_{ij}, a_{ij}; \boldsymbol{\theta}_{ij}(x)) =$ instantaneous fuel rate at any location x $(0 \le x \le L_{ij};\ x \in \mathbb{R})$ on link $(i, j) \in E$, as a function of the instantaneous speed v_{ij}, instantaneous acceleration a_{ij}, and location-dependent parameter vector $\boldsymbol{\theta}_{ij}(x)$.

Now, constraints (1) are necessary but not sufficient for feasibility of a space-time path. In order to introduce sufficient conditions, the time-dependent kinetic functions defined above are applied successively along the links of a generic space time path $((i_1, t_{i_1}), (i_2, t_{i_2}), \ldots, (i_n, t_{i_n}))$ in $\tilde{\mathcal{S}}_{i^* j^* d}$ for a particular (i^*, j^*, d), which then generates the following physical constraints along any such space-time path:

$$v_{i_1 i_2}\left(t_{i_1}\right) = \tilde{v}_0 \tag{2}$$

$$v_{i_k i_{k+1}}\left(t_{i_k}\right) = v_{i_{k-1} i_k}\left(t_{i_k}\right)\ (k = 2, \ldots, n-1) \tag{3}$$

$$v_{i_k i_{k+1}}(t) = v_{i_k i_{k+1}}\left(t_{i_k}\right) + \int_{t_{i_k}}^{t_{i_{k+1}}} a_{i_k i_{k+1}}(w)dw \quad \left(t_{i_k} \le t \le t_{i_{k+1}};\ k = 1, 2, \ldots, n-1\right) \tag{4}$$

$$x_{i_k i_{k+1}}\left(t_{i_k}\right) = 0\ (k = 1, 2, \ldots, n-1) \tag{5}$$

$$x_{i_k i_{k+1}}\left(t_{i_{k+1}}\right) = L_{i_k i_{k+1}}\ (k = 1, 2, \ldots, n-1) \tag{6}$$

$$x_{i_k i_{k+1}}(t) = \int_{t_{i_k}}^{t_{i_{k+1}}} v_{i_k i_{k+1}}(w)dw \ \left(t_{i_k} \le t \le t_{i_{k+1}};\ k = 1, 2, \ldots, n-1\right) \tag{7}$$

Constraint (2) is used to denote the fact that the initial speed at the entry point to the path, $\tilde{v}_0$, is a decision variable, and so therefore must define the speed at the time t_{i_1} at which the first link (i_1, i_2) in the path is entered. Constraints (3) ensure that speeds are conserved at nodes, i.e. the speed entering a link must equal the speed at which the vehicle exited the previous link on the path. Constraints (4) consider the interval of time $t_{i_k} \le t \le t_{i_{k+1}}$ for which link (i_k, i_{k+1}) is used, and specify that during that time all relevant velocities can be generated from the acceleration profile. Thus, the acceleration profile can be used as the decision variable. The advantage over using speed as the decision variable is that well-defined releationships then exist between the kinetic variables without any constraints on the form of the permitted acceleration functions; discontinuous/step functions are permitted, for example. Constraints (5) and (6) ensure that the entry and exit times from each link are physically consistent with the link length, and constraints (7) specify that, again during the relevant time a link is used, the vehicle location can be generated from the velocity profile (which itself was generated from the acceleration profile in (4), hence the acceleration profile is again the 'seed') . Since no backtracking is permitted the inverse mappings of (5) and (6) effectively define the link entry and exit times $(t_{i_k}, t_{i_{k+1}})$ $(k = 1, 2, \ldots, n-1)$, and thereby the link travel times, consistent with the acceleration/speed profiles that generate $x_{i_k i_{k+1}}(.)$ through (7) and (4).

Finally, these variables are subject to constraints on instaneous speed (8) and acceleration (9):

$$v_{i_k i_{k+1},\mathrm{MIN}} \le v_{i_k i_{k+1}}(t) \le v_{i_k i_{k+1},\mathrm{MAX}} \ \left(t_{i_k} \le t \le t_{i_{k+1}};\ k = 1, 2, \ldots, n-1\right) \tag{8}$$

$$a_{i_k i_{k+1},\mathrm{MIN}} \le a_{i_k i_{k+1}}(t) \le a_{i_k i_{k+1},\mathrm{MAX}} \quad \left(t_{i_k} \le t \le t_{i_{k+1}};\ k = 1, 2, \ldots, n-1\right). \tag{9}$$

We are then in a position to define the overall optimization problem (10), for a given $i^* \ne j^* \in V$ and $d \in \mathbb{R}$ $(d \ge 0)$ and given the current clock-time of $s \in \mathbb{R}$ $(s \ge 0)$. Note that in the paper we use a modular approach to introducing the mathematical concepts. Thus having introduced and explained the individual numbered equations above, our approach will be to label/number the overall *problem* (such as we do in (10)), rather than the individual elements of the problem.

$$\text{Minimise} \sum_{k=1}^{n} \int_{t_{i_k}}^{t_{i_{k+1}}} F_{i_k i_{k+1}}\left(v_{i_k i_{k+1}}(t), a_{i_k i_{k+1}}(t); \theta_{i_k i_{k+1}}\left(x_{i_k i_{k+1}}(t)\right)\right)dt$$

with respect to:

$$\left(\left(i_1, t_{i_1}\right), \left(i_2, t_{i_2}\right), \ldots, \left(i_n, t_{i_n}\right)\right) \in \mathcal{S}_{i^* j^* d}$$

$$\tilde{v}_0\ \left(v_{i_1 i_2,\mathrm{MIN}} \le \tilde{v}_0 \le v_{i_1 i_2,\mathrm{MAX}}\right)$$

$$a_{i_k i_{k+1}}(t)\ \left(t_{i_k} \le t \le t_{i_{k+1}};\ k = 1, 2, \ldots, n-1\right)$$

subject to constraints:

$$t_{i_{k+1}} - t_{i_k} \ge \tau_{i_k i_{k+1},\min}\left(s, t_{i_k}\right)\ (k = 1, 2, \ldots, n-1; s \in \mathbb{R},\ s \ge 0)$$

$$v_{i_1 i_2}\left(t_{i_1}\right) = \tilde{v}_0$$

$$v_{i_k i_{k+1}}\left(t_{i_k}\right) = v_{i_{k-1} i_k}\left(t_{i_k}\right)\ (k = 2, \ldots, n-1)$$

$$v_{i_k i_{k+1}}(t) = v_{i_k i_{k+1}}\left(t_{i_k}\right) + \int_{t_{i_k}}^{t_{i_{k+1}}} a_{i_k i_{k+1}}(w)dw \tag{10}$$

$$\left(t_{i_k} \le t \le t_{i_{k+1}};\ k = 1, 2, \ldots, n-1\right)$$

$$x_{i_k i_{k+1}}\left(t_{i_k}\right) = 0\ (k = 1, 2, \ldots, n-1)$$

$$x_{i_k i_{k+1}}\left(t_{i_{k+1}}\right) = L_{i_k i_{k+1}}\ (k = 1, 2, \ldots, n-1)$$

$$x_{i_k i_{k+1}}(t) = \int_{t_{i_k}}^{t_{i_{k+1}}} v_{i_k i_{k+1}}(w)dw\ \left(t_{i_k} \le t \le t_{i_{k+1}};\ k = 1, 2, \ldots, n-1\right)$$

$$v_{i_k i_{k+1},\mathrm{MIN}} \le v_{i_k i_{k+1}}(t) \le v_{i_k i_{k+1},\mathrm{MAX}} \quad \left(t_{i_k} \le t \le t_{i_{k+1}};\ k = 1, 2, \ldots, n-1\right)$$

$$a_{i_k i_{k+1},\mathrm{MIN}} \le a_{i_k i_{k+1}}(t) \le a_{i_k i_{k+1},\mathrm{MAX}} \quad \left(t_{i_k} \le t \le t_{i_{k+1}};\ k = 1, 2, \ldots, n-1\right).$$

2.3. Problem decomposition

The primary source of difficulty in solving problem (10) arises from the non-convex constraints (1)–(9) that ensure feasibility of any space-time path, consistent with both the acceleration/speed profiles and the time-dependent constraints on minimum link travel times from predictive traffic information. This arises whether the problem is expressed in a link-based or path-based form, since for any given path the *active* feasibility constraints based on a given link's real-time predictive information depend on the values of the decision variables for links on the path both upstream and downstream of the current link. Thus, the major difficulty in directly solving problem (10) is that of ensuring *feasibility* of a combination of the decision variables. This is due to interdependencies between decision variables affecting the time-of-day at which

the vehicle enters a link, since this can affect which of the time-dependent constraints are active from the predictive travel time information. However, it is important to note that it is only the times at *node/link-entries* that are critical here; it is immaterial for network *feasibility* (in the sense of (1)) whether a vehicle traverses a 40 km link by travelling for an hour at a constant speed of 40 km/h or travels half an hour each at 20 km/h and 60 km/h, since in both cases the time from entry to exit will be one hour. While these different speed profiles will affect fuel optimality, they do not affect network feasibility, since our real-time travel information only constrains the entry-to-exit travel time. This property suggests a decomposition of (10), which decouples at least part of the problem of determining optimal link speed profiles:

$$\text{Minimise } \sum_{k=1}^{n} g_{i_k i_{k+1}}\left(t_{i_k}, t_{i_{k+1}}\right)$$

with respect to:

$$\left(\left(i_1, t_{i_1}\right), \left(i_2, t_{i_2}\right), \ldots, \left(i_n, t_{i_n}\right)\right) \in \mathcal{S}_{i^* j^* d}$$

subject to constraints:

$$t_{i_{k+1}} - t_{i_k} \geq \tau_{i_k i_{k+1},\min}\left(s, t_{i_k}\right) \ (k = 1, 2, \ldots, n-1; s \in \mathbb{R},\ s \geq 0)$$

where for each $k = 1, 2, \ldots, n-1$, $g_{i_k i_{k+1}}(t_{i_k}, t_{i_{k+1}})$ is an optimal objective value of the constrained sub-problem:

$$\text{Minimise } \int_{t_{i_k}}^{t_{i_{k+1}}} F_{i_k i_{k+1}}\left(v_{i_k i_{k+1}}(t), a_{i_k i_{k+1}}(t); \theta_{i_k i_{k+1}}\left(x_{i_k i_{k+1}}(t)\right)\right) dt$$

with respect to:

$$\tilde{v}_0 \ \left(v_{i_1 i_2,\text{MIN}} \leq \tilde{v}_0 \leq v_{i_1 i_2,\text{MAX}}\right) \text{ (in case } k = 1)$$

$$a_{i_k i_{k+1}}(t) \ \left(t_{i_k} \leq t \leq t_{i_{k+1}}\right)$$

subject to

$$\begin{aligned}
&v_{i_1 i_2}\left(t_{i_1}\right) = \tilde{v}_0 \text{ (in case } k = 1) \\
&v_{i_k i_{k+1}}\left(t_{i_k}\right) = v_{i_{k-1} i_k}\left(t_{i_k}\right) \text{ (in cases } k = 2, \ldots, n-1) \\
&v_{i_k i_{k+1}}(t) = v_{i_k i_{k+1}}\left(t_{i_k}\right) + \int_{t_{i_k}}^{t_{i_{k+1}}} a_{i_k i_{k+1}}(w) dw \ \left(t_{i_k} \leq t \leq t_{i_{k+1}}\right) \\
&x_{i_k i_{k+1}}\left(t_{i_k}\right) = 0 \\
&x_{i_k i_{k+1}}\left(t_{i_{k+1}}\right) = L_{i_k i_{k+1}} \\
&x_{i_k i_{k+1}}(t) = \int_{t_{i_k}}^{t_{i_{k+1}}} v_{i_k i_{k+1}}(w) dw \ \left(t_{i_k} \leq t \leq t_{i_{k+1}}\right) \\
&v_{i_k i_{k+1},\text{MIN}} \leq v_{i_k i_{k+1}}(t) \leq v_{i_k i_{k+1},\text{MAX}} \ \left(t_{i_k} \leq t \leq t_{i_{k+1}}\right) \\
&a_{i_k i_{k+1},\text{MIN}} \leq a_{i_k i_{k+1}}(t) \leq a_{i_k i_{k+1},\text{MAX}} \ \left(t_{i_k} \leq t \leq t_{i_{k+1}}\right).
\end{aligned} \tag{11}$$

Formulation (11) has several advantages over the original formulation (10). Firstly, it separates out the detailed determination of the acceleration/speed profiles for each link individually (the sub-problems), meaning that a high dimensional problem has been replaced by a number of low dimensional problems, thus leading to anticipated computational efficiency gains. Secondly, the master problem of (11) resembles a time-dependent shortest path problem, for which several existing algorithms exist as a starting point to develop a solution strategy (though the fact that constraints (1) are inequalities rather than equalities mean that it does not have the form of a regular shortest path problem). Thirdly, looking towards real-time applications, (11) has the advantage that clock time (s) only appears in the constraints of the master problem, not in the sub-problems. This suggests that, if some kind of offline "pre-processing" of the sub-problems is possible, then significant computational gains may be possible, which is important when a fast response is required as in *en route* re-planning.

2.4. Problem discretisation

The next step towards developing practical solution methods for (11) will be to develop a discretised version of the problem. The representation begins with the choice by the user of two time-discretisation levels, one for the master problem and one for the sub-problems. For the sub-problems the discretisation is used to approximate an integral, with the decision variables (acceleration profiles) defined as step functions. For the master problem, the discretisation is somewhat different, restricting the search to discretised version of the time variables, and requiring all time-related variables to be defined on that discretised level. Thus, the 'master-level' time increment (of size δ) governs the resolution at which all time-related factors are represented *at the network level*, be they clock-time, link entry/exit time, link traversal time or departure time (the predicted link travel times from real-time information could also be discretised, but in fact this is unnecessary so we will not do so). That is to say, this discretisation concerns all time-related factors other than those at a within-link level. A finer, 'minor' time increment (of size Δ) is used for discretising time at the within-link level, specifically for solving the link sub-problems. In order that the master problem and sub-problems knit together in time, Δ is chosen to be some integer sub-division of δ (i.e. $\delta = \kappa\Delta$ for some integer $\kappa \geq 1$).

The strategy to solve the sub-problems at a finer granularity than the master problem is justified by the fact that the fuel consumption objective of the sub-problems is sensitive to the number and duration of acceleration/deceleration events, these nonlinearites meaning that fuel estimates will likely be underestimated by temporal smoothing of the profiles. As noted in the previous section, the master problems are quite different in nature, with the routing decisions sensitive only to the link entry/exit times (given that the sub-problems embed optimal link fuel useage for any given link entry/exit times) and to the update frequency of the real-time link travel time information. As might be anticipated, a finer time resolution provides a better representation of the underlying physical system, but at a price of computational cost. This is particularly true for the master-level discretisation. Since a common master-level time unit must be used for all links of the network, a limiting factor on the coarseness of this discretisation is likely to be the fastest time to traverse the shortest link; it makes little sense to choose a 15-minute master-level discretisation if the shortest link could be traversed in 5 min, for example. However, it is still highly likely to be the case that a coarser discretisation will be sufficient for the master problem than the sub-problems, and so this is explicitly accounted for in the formulation.

In the formulation below we will use following notation to represent a discrete time-grid:

$$\delta\mathbb{N} = \{0, \delta, 2\delta, \ldots\}$$

and the discretised problem (12) is then:

$$\text{Minimise } \sum_{k=1}^{n} g_{i_k i_{k+1}}\left(t_{i_k}, t_{i_{k+1}}\right)$$

with respect to:

$$\left(\left(i_1, t_{i_1}\right), \left(i_2, t_{i_2}\right), \ldots, \left(i_n, t_{i_n}\right)\right) \in \mathcal{S}_{i^* j^* d}$$

subject to constraints:

$$t_{i_{k+1}} - t_{i_k} \geq \tau_{i_k i_{k+1},\min}\left(s, t_{i_k}\right) \ (k = 1, 2, \ldots, n-1; s \in \delta\mathbb{N})$$

$t_{i_k} \in \delta\mathbb{N} \ (k = 1, 2, \ldots, n)$

where for each $k = 1, 2, \ldots, n-1$, $g_{i_k i_{k+1}}(t_{i_k}, t_{i_{k+1}})$ is an optimal objective value of the constrained sub-problem:

$$\text{Minimise} \sum_{t \in \left[t_{i_k}, t_{i_{k+1}}\right] \cap \Delta\mathbb{N}} F_{i_k i_{k+1}}\left(v_{i_k i_{k+1}}(t), a_{i_k i_{k+1}}(t); \boldsymbol{\theta}_{i_k i_{k+1}}\left(x_{i_k i_{k+1}}(t)\right)\right)$$

with respect to:

$$\tilde{v}_0 \ \left(v_{i_1 i_2,\text{MIN}} \le \tilde{v}_0 \le v_{i_1 i_2,\text{MAX}}\right) \ \text{(in case } k = 1)$$

$$a_{i_k i_{k+1}}(t) \ \left(t \in \left[t_{i_k}, t_{i_{k+1}}\right] \cap \Delta\mathbb{N}\right)$$

subject to

$$\begin{aligned}
&v_{i_1 i_2}\left(t_{i_1}\right) = \tilde{v}_0 \ \text{(in case } k = 1)\\
&v_{i_k i_{k+1}}\left(t_{i_k}\right) = v_{i_{k-1} i_k}\left(t_{i_k}\right) \ \text{(in cases } k = 2, \ldots, n-1) \qquad (12)\\
&v_{i_k i_{k+1}}(t) = v_{i_k i_{k+1}}\left(t_{i_k}\right) + \sum_{w \in \left[t_{i_k}, t_{i_{k+1}}\right] \cap \Delta\mathbb{N}} a_{i_k i_{k+1}}(w)\\
&\qquad \left(t \in \left[t_{i_k}, t_{i_{k+1}}\right] \cap \Delta\mathbb{N}\right)\\
&x_{i_k i_{k+1}}\left(t_{i_k}\right) = 0\\
&x_{i_k i_{k+1}}\left(t_{i_{k+1}}\right) = L_{i_k i_{k+1}}\\
&x_{i_k i_{k+1}}(t) = \sum_{w \in \left[t_{i_k}, t_{i_{k+1}}\right] \cap \Delta\mathbb{N}} v_{i_k i_{k+1}}(w) \ \left(t \in \left[t_{i_k}, t_{i_{k+1}}\right] \cap \Delta\mathbb{N}\right)\\
&v_{i_k i_{k+1},\text{MIN}} \le v_{i_k i_{k+1}}(t) \le v_{i_k i_{k+1},\text{MAX}} \ \left(t \in \left[t_{i_k}, t_{i_{k+1}}\right] \cap \Delta\mathbb{N}\right)\\
&a_{i_k i_{k+1},\text{MIN}} \le a_{i_k i_{k+1}}(t) \le a_{i_k i_{k+1},\text{MAX}} \ \left(t \in \left[t_{i_k}, t_{i_{k+1}}\right] \cap \Delta\mathbb{N}\right).
\end{aligned}$$

In the following sections, we first investigate solutions to the sub-problems of (12) (Section 3) before moving on in Section 4 to the solution method for the master problem and the overall solution strategy.

3. Solution to the sub-problems

The sub-problems of (12) involve determing optimal acceleration profiles for each link in turn. We would particular highlight constraints (3) in (12), which mean that the solutions to the individual link problems are inter-related. In particular these constraints ensure consistent speeds at the times at which the vehicle crosses node boundaries, and imply that for a generic space-time path $((i_1, t_{i_1}), (i_2, t_{i_2}), \ldots, (i_n, t_{i_n}))$ in $\tilde{\mathcal{S}}_{i^* j^* d}$, the solution to the sub-problems for links (i_{k-1}, i_k) and (i_{k+1}, i_{k+2}) (the solutions to which yield $v_{i_{k-1} i_k}(t_{i_k})$ and $v_{i_{k+1} i_{k+2}}(t_{i_{k+1}})$) define input boundary conditions for link (i_k, i_{k+1}) . So now suppose that $v_{i_{k-1} i_k}(t_{i_k})$ and $v_{i_{k+1} i_{k+2}}(t_{i_{k+1}})$ are known, and consider the discretised sub-problem for link (i_k, i_{k+1}) from (12).

What particularly distinguishes this boundary problem from standard eco-routing problems is the additional constraints imposed. There are two variants of the problem that we shall consider:

P1. Given the *link entry/exit times*, what is the fuel-optimal speed profile for traversing the link? ("time-constrained problem")
P2. Given the *link entry/exit times and link entry/exit speeds*, what is the fuel-optimal speed profile for traversing the link? ("time- and speed-constrained problem")

Problem P1 is a relaxed version of problem P2, with constraint (3) neglected, so in this case there is no need to ensure consistent speeds at node boundaries. The reason to investigate P1 is that it affords some significant simplification (especially for the overall problem, as described later), and it might be hypothesised that any inconsistency is likely to be short-lived along the link, as the vehicle can quickly re-adjust to its speed anyway; so the effect of requiring speed consistency could be negligible, even though it makes the model more physically sensible. Even the relaxed problem P1 is a distinctive form of problem, due to the constraints (5) and (6), which implicitly constrain the travel time to some predefined level; so the question the sub-problem seeks to answer is: what is the fuel optimal acceleration profile for the link *given some link travel time*.

Table 1
Optimal total fuel consumption (litres) from solving problem P2 for different boundary values of link exit time $t_{i_{k+1}}$ and link exit speed $v_{i_{k+1} i_{k+2}}(t_{i_{k+1}})$ (with $t_{i_k} = 0$ and $v_{i_{k-1} i_k}(t_{i_k}) = 40$ in all cases).

		$t_{i_{k+1}}$ (hours)			
		1.25	1.00	0.83	0.67
Constant speed		17.14	15.65	15.09	15.43
$v_{i_{k+1} i_{k+2}}(t_{i_{k+1}})$ (km/h)	0	17.36	15.97	15.59	16.31
	40	17.46	16.00	15.54	16.14
	80	17.93	16.47	15.88	16.23

Using the fuel rate model described in Appendix A, we consider an example of an undulating link of length $L_{i_k i_{k+1}} = 50$ km with the link elevation profile randomly generated at given points and a cubic spline fitted. The link profile is generated with zero net gradient (i.e. the ends of the link are at the same elevation), facilitating comparison with an entirely flat link. The profile is illustrated at the top of Fig. 1, with a maximum gradient of 3.11°. A discretisation of $\Delta = \frac{1}{120}$ hours (30 s) was used to solve several instances of problem P2. Optimality of the reported solutions was verified to a resolution of $\frac{1}{3600}$ hours (1 second), to ensure no artificial "flattening" of hills has occurred, as may arise if the discretisation is too coarse. Other assumed parameters were $a_{i_k i_{k+1},\text{MIN}} = -2$ m/s^2, $a_{i_k i_{k+1},\text{MAX}} = +2$ m/s^2, $v_{i_k i_{k+1},\text{MIN}} = 0$ km/h and $v_{i_k i_{k+1},\text{MAX}} = 110$ km/h.

The solution to a particular instance is illustrated in Fig. 1 for $t_{i_k} = 0$, $t_{i_{k+1}} = 0.67$ hours (40 min), and $v_{i_{k-1} i_k}(t_{i_k}) = v_{i_{k+1} i_{k+2}}(t_{i_{k+1}}) = 40$ km/h. It was solved with an interior-point algorithm[1] using a barrier function in a computation time of 4.8 s, given convergence tolerances on objective function value and constraints of 10^{-3}. The acceleration (and implied instantaneous fuel consumption) trajectory can be verified from the figure to track the uphill/downhill segments broadly as may be expected, with a total fuel consumption to traverse the link (at the given travel time and entry/exit speeds) of 16.14 litres. For comparison, if the link were flat and traversed at a constant speed ($= \frac{L_{i_k i_{k+1}}}{t_{i_{k+1}} - t_{i_k}}$) given the specified travel time, then the fuel consumption would have been 15.43 litres.

In Table 1 the optimal objective function values are reported for twelve instances of the example, all assuming $t_{i_k} = 0$ and $v_{i_{k-1} i_k}(t_{i_k}) = 40$ km/h, but with different combinations of input values for $t_{i_{k+1}}$ and $v_{i_{k+1} i_{k+2}}(t_{i_{k+1}})$. For comparison, if the vehicle were to drive at constant speed (neglecting $v_{i_{k-1} i_k}(t_{i_k})$ and $v_{i_{k+1} i_{k+2}}(t_{i_{k+1}})$, of course), then the four values for $t_{i_{k+1}}$ correspond to traversing the link at a speed of approximately 40, 50, 60 and 80 km/h, with respective fuel consumption levels of 17.14, 15.65, 15.09 and 15.43, characteristic of the classical u-shaped fuel consumption versus speed relationship. Considering the results across the columns of Table 1 for any given exit speed $v_{i_{k+1} i_{k+2}}(t_{i_{k+1}})$, qualitatively the

[1] https://bit.ly/2zpGAQq accessed 25/4/20

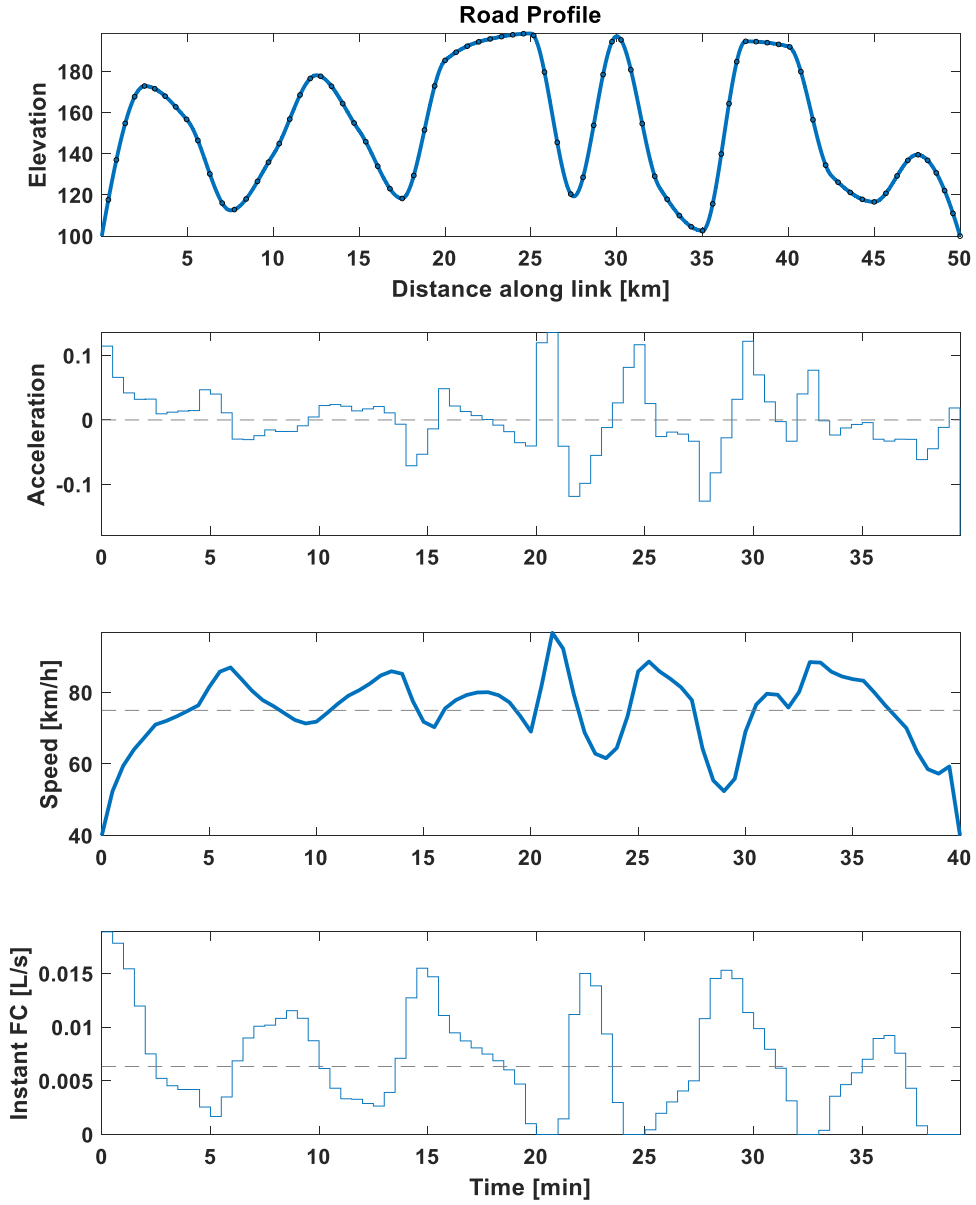

Fig. 1. Solution to Problem P2 for case of $(t_{i_k}, t_{i_{k+1}}, v_{i_{k-1}i_k}(t_{i_k}), v_{i_{k+1}i_{k+2}}(t_{i_{k+1}})) = (0, 0.67, 40, 40)$. Upper plot shows road elevation (m) vs distance along the road (km) with a black dot at each time step (1 min discretisation). Lower three plots show Acceleration, Speed and Instantaneous Fuel Consumption vs time (mins). Dashed horizontal lines show 75 km/h constant speed solution for flat link for relaxed problem P1.

same kind of u-shaped relationship is observed with the optimal values, but with the effect more pronounced due to the impact of gradient. The results are also seen to be sensitive, as expected, to the boundary value exit speed $v_{i_{k+1}i_{k+2}}(t_{i_{k+1}})$.

It is worth noting that this optimization problem admits other objectives such as vehicle emissions. Abdelmegeed and Rakha (2017) propose that instantaneous emissions of NOx, CO, and HC can be effectively modelled as polynomial functions of instantaneous fuel consumption and speed: $E(t) = E(F(t), v(t))$. Analytic emissions functions of this sort make it simple to swap the fuel consumption objective used here, for NOx for example, or for a weighted combination of fuel consumption and emissions objectives.[2] The same procedure may then be used to determine the optimal acceleration profile for an undulating road section with boundary conditions as described above.

4. Solution of the master problem and overall solution strategy

In the present secton we describe the solution method adopted for two variants of problem (12), in Sections 4.1 and 4.2, including the overall strategy for bringing together all the elements described so far.

4.1. Solution of the relaxed problem

The relaxed problem is defined as problem (12) with the speed consistency constraint (3) relaxed. This has implications not only for the sub-problems (as discussed in Section 3) but also for the master problem, since the relaxation allows the relevant space-time paths to be defined only in terms of their link sequences and node passage times. Although it is less realistic than assuming con-

[2] Simultaneously minimising fuel consumption, travel time and emissions in a multi-objective context is a more demanding problem outside the scope of this paper.

sistent speeds at node boundaries, it is considered as it is likely to have significant computational advantages. It is also useful for pedagogical reasons, as it is gives the most direct connection to existing methods; the methods defined in Section 4.2 for the non-relaxed problem are an extension of the methods described here.

Here, we particularly draw on our own previous work in using specially-designed Space-Time Extended Networks (STENs) which in this earlier work were used to solve highly simplified versions of the problems considered in the present paper (Watling et al., 2019). Our previous work demonstrated the computational attractiveness of a STEN-based approach, but without the realistic modelling of acceleration events and speed profiles (since we assumed a single constant speed to traverse each link, and neglected the impact of acceleration on fuel consumption). The drawbacks to such an assumption were discussed in Section 1.

In order to introduce such an approach some additional variables are needed:

$\tau_{ij,\max}(s,t)$ = at current clock-time $s \in \delta\mathbb{N}$, the predicted *maximum* travel time to traverse link $(i,j) \in E$ for a truck of the given type, if entering the link at (future) clock-time $t \in \delta\mathbb{N}$ $(t > s)$, given real-time predictive information available at time s

Note that the predicted *minimum* travel times (as defined earlier) vary by time of day to take account of phenomena such as congestion and traffic incidents, which may mean that the truck is not able to travel at its maximum permitted speed, and thus they will influence the decision variables. The predicted *maximum* travel times, on the other hand, are included only for implementing the STEN approach; they have no physical interpretation, and the intention is that they are sufficiently large so as not to be binding at optimality. The maximum times could be assumed to be independent of the time of day, or they could adapt based on the minimum times, and so to allow both possibilities they are defined as potentially depending on both t and s. Note that while they have no physical interpretation, the maximum times have a potentially significant impact on computation speed. Let $\boldsymbol{\tau}_{\min}(s)$ and $\boldsymbol{\tau}_{\max}(s)$ respectively denote vectors of minimum and maximum travel time across $s \in \delta\mathbb{N}$.

Following the procedure described in Watling et al (2019), a STEN is then created, which consists of a set $\tilde{V}$ of space-time nodes indexed as (i,t) for $i \in V$ and $t \in \delta\mathbb{N}$ and a set $\tilde{E}^{(s)}$ of space-time links indexed as $(i,t) \to (j,t+\theta)$ for $(i,j) \in E, t \in \delta\mathbb{N}, \theta \in \delta\mathbb{N}$ given by:

$$\tilde{V} = \{(i,t) \ : \ i \in V \setminus \{i^*\}; \ t \in \delta\mathbb{N}\} \cup (i^*,d) \cup (j_{\text{SINK}}, \text{free}) \tag{13}$$

$$\begin{aligned}\tilde{E}^{(s)} = \{&(i,t) \to (j,t+\theta) : (i,t) \in \tilde{V}; (i,j) \in E; \\ &\theta \in \left[\tau_{ij,\min}(s,t), \tau_{ij,\max}(s,t)\right] \cap \delta\mathbb{N}\} \\ &\cup \{(i^*,d) \to (j,d+\theta) : (i^*,j) \in E; \\ &\theta \in \left[\tau_{i^*j,\min}(s,d), \tau_{i^*j,\max}(s,d)\right] \cap \delta\mathbb{N}\} \\ &\cup \{(j^*,t) \to (j_{\text{SINK}}, \text{free}) : t \in \delta\mathbb{N}\}\end{aligned} \tag{14}$$

where in (13) the time-extended node set comprises time-indexed nodes in V in combination with $(j_{\text{SINK}}, \text{free})$, the latter denoting an ultimate super-destination to which all space-time flows are headed; and where the time-extended link set (14) consists of (i) multiple edges joining each pair of time-extended nodes, with an edge for each possible (discrete) travel time, (ii) similar multiple edges from the origin node i^* but at the pre-defined departure time d; and (iii) a single edge from the time-extended destination node j^* to the dummy sink j_{SINK}, which is reached regardless of the destination arrival time. The procedure above differs from those reported elsewhere for creating space-time networks, in its use of *multiple* space-time links exiting from each space-time node, which it does to embed the choice of link traversal time. (If we were to set $\tau_{ij,\min}(s,t) = \tau_{ij,\max}(s,t)$ then we would obtain a standard STEN, in which travel times are determined between pairs of space-time nodes.) This basic formulation can be readily extended to incorporate discretionary stops, by including waiting links by appending $\{(i,t) \to (i,t+1) : \ t \in \delta\mathbb{N}\}$ to $\tilde{E}^{(s)}$. The formulation can also be extended to incorporate departure time choice by removing the special treatment of i^* in $\tilde{V}$ and $\tilde{E}^{(s)}$ (i.e. treat the initial node just like any other node), and by adding a single super-origin $(i_{\text{SOURCE}}, \text{free})$ to $\tilde{V}$ with corresponding space-time links in $\tilde{E}^{(s)}$ given by $\{(i_{\text{SOURCE}}, \text{free}) \to (i^*,t) : \ t \in \delta\mathbb{N}\}$.

Having instantiated the STEN, the space-time links are embued with "costs". This is where we first depart from our earlier work, in that here we use the costs to incorporate the sub-problems of (12). Consider a particular space-time link $(i,t) \to (j,t+\theta)$ in $\tilde{E}^{(s)}$. Now, consider a generic sub-problem of the relaxed form of (12), namely Problem P1 as defined in Section 3, and set its boundary values as $(t_{i_k}, t_{i_{k+1}}) = (t\delta, (t+\theta)\delta)$. For this, $n = \theta\frac{\delta}{\Delta} = \theta\kappa$ increments will be needed in the discretisation of speeds/acceleration (and n is an integer since both θ and κ are). Then if $\{v_i^*, a_i^*, x_i^* \ : i = 1,2,\ldots,n\}$ is an optimal solution to the sub-problem then we set the cost of the space-time link $(i,t) \to (j,t+\theta)$ to be $\sum_{i=1}^{n} F(v_i^*, a_i^*; \theta(x_i^*))$.

Formally, the connection between problem (12) and the representation as a STEN is as follows. Denote the special form of STEN that arises from (13)/(14) by the mapping:

$$\left(\tilde{V}, \tilde{E}^{(s)}\right) = \Phi(V, E, i^*, j^*, d, \tau_{\min}(s), \ \tau_{\max}(s)) \ .$$

The relaxed form of (12) is then:

$$\text{Minimise} \ \sum_{k=1}^{n} g_{i_k i_{k+1}}\left(t_{i_k}, t_{i_{k+1}}\right)$$

with respect to:

$$\begin{aligned}&\left(\left(i_1, t_{i_1}\right), \left(i_2, t_{i_2}\right), \ldots, \left(i_n, t_{i_n}\right)\right) \in \left(\tilde{V}, \tilde{E}^{(s)}\right) \\ &\quad = \Phi(V, E, i^*, j^*, d, \tau_{\min}(s), \ \tau_{\max}(s))\end{aligned}$$

where for each $k = 1,2,\ldots,n-1$, $g_{i_k i_{k+1}}(t_{i_k}, t_{i_{k+1}})$ is an optimal objective value of the constrained sub-problem:

$$\text{Minimise} \sum_{t \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}} F_{i_k i_{k+1}}\left(v_{i_k i_{k+1}}(t), a_{i_k i_{k+1}}(t); \boldsymbol{\theta}_{i_k i_{k+1}}\left(x_{i_k i_{k+1}}(t)\right)\right)$$

with respect to:

$$\tilde{v}_0 \ \left(v_{i_1 i_2,\text{MIN}} \le \tilde{v}_0 \le v_{i_1 i_2,\text{MAX}}\right) \ (\text{in case } k = 1)$$

$$a_{i_k i_{k+1}}(t) \ \left(t \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}\right)$$

subject to

$$\begin{aligned}&v_{i_1 i_2}\left(t_{i_1}\right) = \tilde{v}_0 \ (\text{in case } k = 1) \\ &v_{i_k i_{k+1}}(t) = v_{i_k i_{k+1}}\left(t_{i_k}\right) + \sum_{w \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}} a_{i_k i_{k+1}}(w) \\ &\qquad \left(t \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}\right) \\ &x_{i_k i_{k+1}}\left(t_{i_k}\right) = 0 \\ &x_{i_k i_{k+1}}\left(t_{i_{k+1}}\right) = L_{i_k i_{k+1}} \\ &x_{i_k i_{k+1}}(t) = \sum_{w \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}} v_{i_k i_{k+1}}(w) \ \left(t \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}\right) \\ &v_{i_k i_{k+1},\text{MIN}} \le v_{i_k i_{k+1}}(t) \le v_{i_k i_{k+1},\text{MAX}} \ \left(t \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}\right) \\ &a_{i_k i_{k+1},\text{MIN}} \le a_{i_k i_{k+1}}(t) \le a_{i_k i_{k+1},\text{MAX}} \ \left(t \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}\right).\end{aligned} \tag{15}$$

An important point to note is that the optimal values of the within-link sub-problems of (15) are independent of the network

decision variables $(\mathbf{i}, \mathbf{t})$, since they presume an overall time to travel between the nodes as boundary values. They are also independent of the real-time information (i.e. they are independent of s). That is to say, even though $\tilde{E}^{(s)}$ in (14) does depend on s, the costs assigned (inside (15)) to the space-time links in $\tilde{E}^{(s)}$ do not. This is a key property as it suggests a very efficient approach to the real-time deployment of the solution algorithm:

1. Set the following:
   (a) time discretisation step $= \delta$
   (b) origin node $= i^*$
   (c) destination node $= j^*$
   (d) set of nodes where stops are allowed $= V_{STOP}$
   (e) overall time horizon $T \subseteq \delta\mathbb{N}$
   (f) departure/arrival time constraints $= T^D, T^A \subseteq T$
2. Instantiate a reference STEN according to (13)/(14) based on the bounds in step 1, denoting it by $(\tilde{V}, \tilde{E})$. For this reference STEN, embue the links with costs by solving the relevant fuel consumption sub-problems in (16) indepedently for each space-time link.
3. As real-time information emerges in the form of $\tau_{\min}(s)$ create the appropriate STEN $(\tilde{V}, \tilde{E}^{(s)})$ simply by deleting infeasible space-time links from the reference edge set $\tilde{E}$.
4. Solve a shortest path problem on STEN $(\tilde{V}, \tilde{E}^{(s)})$.

Importantly, Step 2 may be conducted once, off-line; thus even though there will typically be a large number of sub-problems to solve, they only need to be solved once and effectively 'stored' in the reference STEN for whenever they are needed. Steps 3 and 4 are then the only steps conducted in real-time. A shortest path on such a STEN then corresponds to a fuel-optimal path and acceleration/speed-profile along the path (and will also optimize departure time, if relevant, and stops, if included, according to the extensions described earlier in the section). That is to say, it not only gives an exact algorithm for the relaxed version of (12), but it does so in a way that allows the link-level problems (and hence the overall acceleration/speed profile for the path) to be solved at a fine temporal resolution, off-line, to ensure that the impact of acceleration events and gradients on fuel consumption are not lost, but in such a way that this fine resolution does not affect the on-line computational performance of the method. Perhaps most critically, the STEN-based approach presented ensures that all dynamic predictive constraints (1) are satisfied, since any feasible path in the STEN automatically does so.

4.2. Solution of the non-relaxed problem

A disadvantage of the approach described in secton 4.1 is that it neglects a potentially important physical characteristic in (12), by relaxing constraint (3). In that method, while speed profiles are solved in detail at the link-level, there is no 'matching' of the speed profiles across links in a path, meaning that a vehicle may exit one link at a very different speed to that at which it enters the next. In the present section, we describe an extension of the methods previously described to address this issue.

To apply a STEN approach, we introduce the idea of node transfer speeds, and define U to be a discrete set of permitted transfer speeds. In the new STEN, multiple copies of each space-time node are created, one for each permitted transfer speed, with duplicate copies of the relevant space-time links created for each such transfer speed. Thus the STEN now consists of a set $\hat{V}$of space-time-speed nodes indexed as (i,t,u) for $i \in V$, $t \in \delta\mathbb{N}$ and $u \in \mathcal{U}$; and a set $\hat{E}^{(s)}$ of space-time-speed links indexed as $(i,t,u) \rightarrow (j, t+\theta, w)$ for $(i,j) \in E$, $t \in \delta\mathbb{N}$, $\theta \in \delta\mathbb{N}$, $u \in \mathcal{U}$, $w \in \mathcal{U}$, given by (with u^* denoting the given initial speed at node i^*):

$$\hat{V} = \{(i,t,u) \ : \ i \in V \setminus \{i^*\}; \ t \in \delta\mathbb{N}; \ u \in \mathcal{U}\} \cup (i^*, d, u^*) \cup (j_{\text{SINK}}, \text{free}, \ \text{free}) \tag{17}$$

$$\begin{aligned}\hat{E}^{(s)} = \{&(i,t,u) \rightarrow (j, t+\theta, w) : (i,t,u) \in \hat{V}; (i,j) \in E; \\ &\theta \in \left[\tau_{ij,\min}(s,t), \tau_{ij,\max}(s,t)\right] \cap \delta\mathbb{N}; w \in \mathcal{U}\} \\ \cup\{&(i^*, d, u^*) \rightarrow (j, d+\theta, w) : (i^*, j) \in E; \\ &\left[\tau_{i^*j,\min}(s,d), \tau_{i^*j,\max}(s,d)\right] \cap \delta\mathbb{N}; w \in \mathcal{U}\} \\ \cup\{&j^*, t, u) \rightarrow (j_{\text{SINK}}, \text{free}, \text{free}) : t \in \delta\mathbb{N}; u \in \mathcal{U}\}\end{aligned} \tag{18}$$

To associate costs with the space-time-speed links, we now consider Problem P2 defined in Section 4.2. For a particular space-time-speed link $(i,t,u) \rightarrow (j, t+\theta, w)$ in $\hat{E}^{(s)}$, the boundary values of the relevant sub-problem are $(t_{i_k}, t_{i_{k+1}}, v_{i_{k-1}i_k}(t_{i_k}), v_{i_{k+1}i_{k+2}}(t_{i_{k+1}})) = (t, t+\theta, u, w)$, with again $n = \theta \frac{\delta}{\Delta} = \theta\kappa$. Then if $\{v_i^*, a_i^*, x_i^* \ : i = 1, 2, \ldots, n\}$ is an optimal solution to the sub-problem, we set the cost of the space-time-speed link $(i,t,u) \rightarrow (j, t+\theta, w)$ to be $\sum_{i=1}^{n} F(v_i^*, a_i^*; \theta(x_i^*))$.

Formally, feasibility with respect to the STEN defined by (17) and (18) is written as

$$(\mathbf{i}, \mathbf{t}, \mathbf{u}) \in \left(\hat{V}, \hat{E}^{(s)}\right) = \Omega(V, E, i^*, j^*, d, u^*, \tau_{\min}(s), \ \tau_{\max}(s)) \tag{19}$$

with the resulting problem then:

$$\text{Minimise} \ \sum_{k=1}^{n} h_{i_k i_{k+1}}\left(t_{i_k}, t_{i_{k+1}}, u_{i_k}, u_{i_{k+1}}\right)$$

with respect to:

$$(\mathbf{i}, \mathbf{t}, \mathbf{u}) \in \left(\hat{V}, \hat{E}^{(s)}\right) = \Omega(V, E, i^*, j^*, d, u^*, \tau_{\min}(s), \ \tau_{\max}(s))$$

where for each $k = 1, 2, \ldots, n-1$, $h_{i_k i_{k+1}}(t_{i_k}, t_{i_{k+1}}, u_{i_k}, u_{i_{k+1}})$ is an optimal objective value of the constrained sub-problem:

$$\text{Minimise} \sum_{t \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}} F_{i_k i_{k+1}}\left(v_{i_k i_{k+1}}(t), a_{i_k i_{k+1}}(t); \boldsymbol{\theta}_{i_k i_{k+1}}\left(x_{i_k i_{k+1}}(t)\right)\right)$$

with respect to:

$$a_{i_k i_{k+1}}(t) \ \left(t \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}\right)$$

subject to

$$v_{i_k i_{k+1}}(t) = v_{i_k i_{k+1}}\left(t_{i_k}\right) + \sum_{w \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}} a_{i_k i_{k+1}}(w) \qquad \left(t \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}\right) \tag{20}$$

$$v_{i_k i_{k+1}}\left(t_{i_k}\right) = u_{i_k}$$

$$v_{i_k i_{k+1}}\left(t_{i_{k+1}}\right) = u_{i_{k+1}}$$

$$x_{i_k i_{k+1}}\left(t_{i_k}\right) = 0$$

$$x_{i_k i_{k+1}}\left(t_{i_{k+1}}\right) = L_{i_k i_{k+1}}$$

$$x_{i_k i_{k+1}}(t) = \sum_{w \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}} v_{i_k i_{k+1}}(w) \ \left(t \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}\right)$$

$$v_{i_k i_{k+1},\text{MIN}} \leq v_{i_k i_{k+1}}(t) \leq v_{i_k i_{k+1},\text{MAX}} \ \left(t \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}\right)$$

$$a_{i_k i_{k+1},\text{MIN}} \leq a_{i_k i_{k+1}}(t) \leq a_{i_k i_{k+1},\text{MAX}} \ \left(t \ \in \ \left[t_{i_k}, t_{i_{k+1}}\right] \ \cap \ \Delta\mathbb{N}\right).$$

Note that in formulation (20), no longer do we optimize the initial speed in the sub-problems (since it is defined by the choice of $(\mathbf{i}, \mathbf{t}, \mathbf{u})$ from the STEN), and the sub-problems now included addition constraints defining the speeds at the entry and exit times (again defined by the choice of $(\mathbf{i}, \mathbf{t}, \mathbf{u})$ from the STEN). Many of the features that were observed in Section 4.1 for the relaxed problem transfer over. Thus, the formulation is readily extended to include the choice of discretionary stops and/or departure time. The

Table 2
Overview of Examples.

Example	Within-link Speed	Gradient of Link 2	Speed at Node 1 and Node 4	Speed at mid-route nodes (2,3)	Feasible Link Travel Times	En-route Stop Allowed
1A	Constant	Constant	<free>	<free>	Static	No
1B	Optimised	Constant	0	Matched	Static	No
1C	Optimised	Varying	0	Matched	Static	No
1D	Optimised	Varying	0	Matched	Dynamic	No
1E	Optimised	Varying	0	Matched	Dynamic	Yes

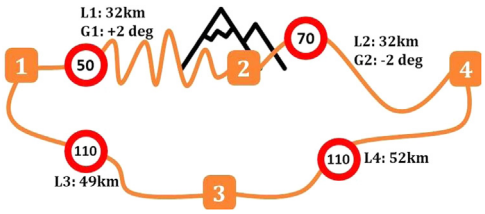

Fig. 2. Simple example network with link gradients, lengths and speed limits shown. See also Table 3.

optimal values of the sub-problems for determining costs on the extended network are independent of s, and so the same strategy may be adopted of instantiating the network off-line and only solving a shortest path problem on-line. A shortest path in the new form of STEN now automatically optimizes fuel consumed with respect to the additional decision variable of transfer speed. In terms of encoding speed into the STEN, two clear options are to have a constant discretisation (e.g. in 10 km/h increments), or to have a variable-sized mesh dependent on the sensitivity of fuel to speed at that speed level. Whichever is chosen, it is anticipated that a rather coarse discretisation will be used, since the impact of any discretisation 'error' should be relatively small (the speeds are only instantaneous speeds at the boundaries, and so the vehicle can very quickly adjust), while the impact of a fine discretisation on computation time can be very significant. An overview of the steps needed to fully implement the procedure is given in Appendix B.

5. Example 1: simple network

We construct a simple example to illustrate key elements of the problem formulation and solution, again using the fuel consumption model of Appendix A. The baseline case assumes constant link speed bounded by the speed limits, constant link gradient and is relaxed so there is no requirement for consistent node transfer speeds. We then introduce, in turn, elements of the model proposed in Section 3 and related constraints, and illustrate the impact on both fuel consumption and path choice. When incorporating node transfer speeds, we consider two cases of: (1) node transfer speeds all set to be 0 km/h; and (2) a discrete feasible set of potential node transfer speeds: [0, 30, 50, 90] km/h.

We consider a two-path network where each path from node 1 → node 4 comprises two links with lengths and speed limits as shown in Fig. 2 and Table 3. Nodes 2 and 3 provide potential en-route stopping points. The elevation of node 2 is approximately 1 km above nodes 1,3,4, these latter all being at the same elevation. This results in Link 1 having an overall uphill gradient of $+2°$, with link 2 (of almost the same length) having an average gradient of $-2°$ i.e. downhill.

While link 1 is of constant uphill gradient, and links 3 & 4 are flat, link 2 is assumed to be undulating with local fluctuations in gradient (see Fig. 3 below). Within the collection of examples presented below, the non-uniform gradient of link 2 is deliberately ignored in some cases and included in others to highlight the impact of assuming a constant gradient.

5.1. Example 1A: baseline

The baseline case assumes a constant gradient, whether the gradient is zero or non-zero i.e. no undulations. As a reference point, we first examine what would happen if links were analysed independently to find optimal fuel-minimising link speed, i.e. by solving problem P1 as defined by (7). The optimal speed on link 1 is 54.64 km/h (giving fuel consumption of 26.77 litres) but this exceeds the link speed limit, and hence within the speed constraints the optimal speed is 50 km/h. There is no constraint on link-end speeds: departure and arrival occur at high speed, and speed jumps from 50 km/h to 70 km/h at node 2. If these independently-optimized link speeds were adopted, then overall path 1 would be the more fuel-efficient, as shown in Table 3.

5.2. Building the link travel time-optimal fuel consumption lookup table

For each link, we now consider a (discretized) range of travel times and a set of node transfer speeds. Initially, we suppose that we have no dynamic travel time information available, so consider a case where the minimum and maximum travel times are constant over the day. Link 1 has a length of 31.92 km with a minimum speed of 25 km/h and a maximum speed of 50 km/h, giving the range of link travel times [38.3, 76.61] minutes. Discretizing the STEN at a 1-minute interval, we therefore consider possible travel times to be [39, 40, …, 77] minutes. The upper speed-limit is considered a hard constraint and hence disallows link travel time of 38 min (corresponding speed would be 50.4 km/h) or less. We consider a discrete set of node transfer speeds of [0, 30, 50, 90] km/h. For each link travel time, and each possible combination of node transfer speeds, we solve to find the acceleration profile which results in minimum fuel consumption. We record the fuel consumption in a lookup table (see Table 4), having coordinates (link number, link travel time, A-node Speed, B-node Speed, fuel consumed). The time-discretisation for solving the link sub-problems was 30 s. In practice, we solved the sub-problems in such a way that the numerical solution was not guaranteed to exactly satisfy all constraints. We therefore capture diagnostic information such as whether the speed limit was instantaneously exceeded (we allow 5% of timesteps) and if the link length constraint was met (we allow 5% error in distance travelled along the link).

The STEN is constructed using the link-specific discretized link travel times and node speeds, and the relevant link costs are taken from the lookup table (Table 4). Links of the STEN are constructed (or not) to capture various constraints. For example, the origin and destination node speeds are fixed at 0 km/h in all cases.

5.3. Example 1B: use optimal link speed profiles

For this and all subsequent examples we compute the fuel minimising path using the relevant lookup table (as in Table 4) and the STEN constructed from it. In this example Link 2 is assumed to have constant downhill gradient.

Controlling and matching node transfer speeds is desirable for consistency of the model and to represent network features e.g. if

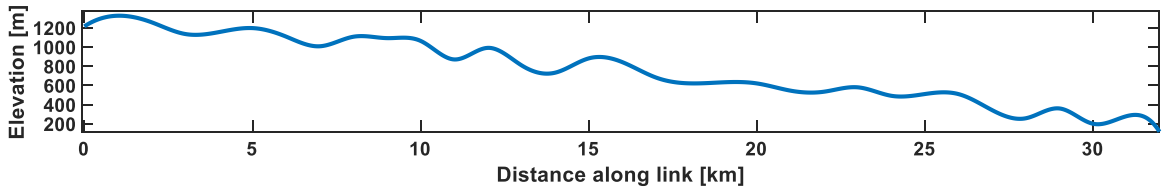

Fig. 3. Undulations on link 2. Overall gradient = −2°.

Table 3
[Ex 1A] Independent link-based fuel consumption minimisation (constant speed, no undulations).

Link	Length	Speed Limit min/max	Link Travel Time (min)	Link Speed (km/h)	Fuel Consumption (L)	Path	Fuel Consumption (L)	Travel Time (mins)
1	31.92	25/50	38.30	50.00	26.83	1	26.83	65.77
2	32.05	25/70	27.47	70.00	0			
3	48.96	40/110	44.70	65.72	14.70	2	31.38	92.36
4	52.20	40/110	47.66	65.72	15.68			

Note that zero fuel is consumed on link 2 due to the constant downhill gradient.

Table 4
Extract from Optimal Fuel Consumption lookup table. Note minimum travel time is equivalent to travelling constantly at the maximum speed.

Link	A-node Transfer Speed (km/h)	B-node Transfer Speed (km/h)	Link Travel Time (min)	Fuel (L)	A-node Transfer Speed (km/h)	B-node Transfer Speed (km/h)	Link Travel Time (min)	Fuel (L)
1	0	0	39	27.15	50	50	39	26.85
1	0	0	40	27.13	50	50	40	26.89
1	0	0	41	27.14	50	50	41	26.94
1	0	0	42	27.16	50	50	42	26.99
1	0	0	43	27.17	50	50	43	27.05

Table 5
Real-time predictive information (LTTT).

Time	Link01	Link02	Link03	Link04
12-Jul-2019 08:00:00	38.30	21.37	26.71	28.47
12-Jul-2019 08:01:00	38.30	21.37	26.7128.47	
12-Jul-2019 08:28:00	38.30	21.37	26.71	28.47
12-Jul-2019 08:29:00	38.30	21.37	26.71	28.47
12-Jul-2019 08:30:00	38.30	21.37	26.71	90.00
12-Jul-2019 08:31:00	38.30	21.37	26.71	90.00
12-Jul-2019 08:32:00	38.30	21.37	26.71	90.00

Table 6
Real-time predictive information (LTTT).

Time	Link01	Link02	Link03	Link04
12-Jul-2019 08:00:00	38.30	21.37	26.71	90.00
12-Jul-2019 08:01:00	38.30	21.37	26.71	90.00
12-Jul-2019 08:02:00	38.30	21.37	26.71	90.00
12-Jul-2019 08:03:00	38.30	21.37	26.71	90.00
12-Jul-2019 08:04:00	38.30	21.37	26.71	90.00
12-Jul-2019 08:05:00	38.30	21.37	73.00	90.00
12-Jul-2019 08:57:00	38.30	21.37	73.00	90.00
12-Jul-2019 08:58:00	38.30	21.37	73.00	89.00
12-Jul-2019 08:59:00	38.30	21.37	73.00	88.00
12-Jul-2019 09:58:00	38.30	21.37	73.00	29.00
12-Jul-2019 09:59:00	38.30	21.37	73.00	28.47
12-Jul-2019 10:00:00	38.30	21.37	73.00	28.47

a node is a starting/stopping node such as a junction. To illustrate the impact of specifying node transfer speeds in the formulation we initially consider a problem with all node transfer speeds set at 0 km/h for all nodes. Fuel consumption is minimised on path 1 as shown in Table 7 - 1B(i). Subsequently we allow optimisation of node transfer speeds, selected from $U = [0,30,50,90]$ km/h, while maintaining 0 km/h at origin and destination nodes. The optimal path does not change, though fuel consumption is improved (see Table 7 - 1B(ii)). Note that the speeds reported in Table 7 are averages of the optimal link speed profile, computed from link travel time and link length; they are not the maximum speed attained within the link.

5.4. Example 1C: include link undulations

In the calculations so far, link 2 has been assumed to have constant gradient.

Including the gradient profile (shown in Fig. 3) requires recomputing the optimal, fuel minimising speed profiles, and updating the lookup table i.e. fuel consumption data in Table 4 is updated. The minimum fuel consumed, under any feasible travel time, when traversing link 2 is now 17.05 litres The resulting increase in fuel consumed therefore changes the optimal path, as shown in Table 7 – 1C.

5.5. Example 1D: include dynamic link travel times

To illustrate how the method adapts to capture real-time predictive information of dynamic network conditions we consider the onset of severe congestion on link 4 (e.g. as might be caused by an incident). Adopting similar terminology to that suggested by Eglese et al. (2006), we store the link travel times in a Link Travel Time Timetable (LTTT): using a one-minute discretisation, then at each minute during the period of analysis we specify the minimum travel time to traverse the link when entering the link at that time. As real-time information emerges, it is used to override the default travel time levels, e.g. obtained by assuming a vehicle travels at the speed limit for the whole link length.

The real-time information assumed is given in Table 5, where the travel time on link 4 has a step increase at 08:30. Entering link 4 later than this will impose a minimum link travel time of 90 min; as an approximate guide, this corresponds to an average link speed of 34.8 km/h (compared with 65.72 km/h in previous examples), which due to the U-shaped fuel consumption curve (see Appendix A) would significantly increase fuel consumption if it were the only information used (it is not, the calculations below are based on the speed profile, not just average speed). The

Table 7
Examples 1A-1E Results.

Example	Path	Link	Link Travel Time (min)	Path Travel Time (min)	Link End Speeds (km/h)	Link Entry Time	Average Speed (km/h)	Fuel (L)	Fuel (L)
1A	1	1	38.30	65.77	50	N/A	50.00	26.83	26.83
		2	27.47		70	N/A	70.00	0	
1B (i)	1	1	40	113	0,0	08:00	47.88	27.13	27.21
		2	73		0,0	08:40	26.34	0.08	
1B (ii)	1	1	40	102	0,30	08:00	47.88	27.12	27.12
		2	62		30,0	08:40	31.02	0.00	
1C	2	3	46	95	0,50	08:00	63.87	15.05	30.81
		4	49		50,0	08:46	63.92	15.76	
1D	2	3	29	78	0,90	08:00	101.3	18.30	33.64
		4	49		90,0	08:29	63.92	15.34	
1E	2	3	47	147	0,0	08:00	62.51	15.08	31.13
		–	50			08:47	0	0	
		4	50		0,0	09:37	62.64	16.05	

Table 8
Extract from dynamic Link Travel Time Timetable (LTTT).

Day	Time Of Day	Link1 (mins)	Link2 (mins)	Link3 (mins)	Link4 (mins)	Link5 (mins)
12-Jul-19	06:40:00	84.98	15.69	18.95	27.44	13.43
12-Jul-19	06:41:00	84.92	15.75	18.97	27.39	13.46
12-Jul-19	06:42:00	84.86	15.81	18.99	27.34	13.50
12-Jul-19	06:43:00	84.80	15.87	19.01	27.30	13.53
12-Jul-19	06:44:00	84.75	15.94	19.03	27.25	13.56
12-Jul-19	06:45:00	84.69	16.00	19.06	27.20	13.59

Table 9
STEN graph properties for different time spans.

Earliest Start	Latest End	Time Span [hrs]	STEN Edges	STEN Nodes	RAM [MB]	Build STEN [s]	Shortest Path [s]
06/07/2019 00:00	10/07/2019 23:59	120	275,995,555	1513,261	4124	599.0	5.59
06/07/2019 00:00	06/07/2019 23:59	24	47,667,068	338,221	730	95.7	0.80
06/07/2019 06:00	06/07/2019 18:00	12	19,195,920	191,341	294	40.2	0.37

Table 10
Results of tests on the UK motorway network.

Row	Start Time	End Time	Path	Fuel [litres]	Travel Time [hh:mm]	Path Length [km]	Average Speed [km/h]	Notes
1	06:00	18:00	1	223.98	11:23	742.68	65.24	Node speeds 50 km/h
2	06:00	14:00	1	246.39	08:00	742.68	92.84	Median node speed 90 km/h
3	06:00	18:00	2	249.41	12:00	823.01	68.58	42 and 62 Closed
4	06:00	18:00	3	243.30	11:39	803.13	68.94	42 and 62 re-open *en route*
5	06:00	18:00	4	275.50	12:00	882.74	73.56	42 and 62 close *en route*

Table 11
Fuel consumption model parameters.

	Parameter Value
β_1	0.000344636826390
β_2	0.000000543265083
β_3	0.042822544388554
β_4	0.006708663250830
β_5	0.002327916266460
β_6	0.319097080735411

time at which the truck departs is also now relevant in the case of dynamic traffic information – in all cases, we assume a fixed departure time of 08:00. The optimal solution is shown in Table 7 – 1D.

5.6. Example 1E: allow stops

We construct the LTTT (see Table 6) to illustrate the benefit of including optional stops in the formulation. Here Link 3 becomes congested from 08:05 onwards, whereas link 4 is congested until 09:00, after which it clears (NB: we make sure that the link travel time decay over time does not contravene FIFO). With departure at 08:00 and no arrival time constraint, fuel consumption is minimised by stopping at node 3, as shown in Table 7 – 1E.

5.7. Simple network results

The base case (Ex 1A) ignores any node speed constraints and assumes constant gradient on all links. This results in the minimum fuel consumption of all the examples. The simplest specification ensuring origin/destination node speeds are zero, and that all node transfer speeds match, is to force all link-end speeds to be zero.

Example 1B(i) shows an increase in fuel consumption compared with Example 1A due to the need to accelerate from rest at the start of each link; in particular link 2 now has non-zero fuel consumption despite being assumed to have constant downhill gradient in this example. By allowing multiple node transfer speeds, [0,30,50,90] km/h, but maintaining 0 km/h at origin and destination nodes, (Table 7 - 1B(ii)), fuel consumption is lowered slightly compared with Ex 1B(i), due to passing node 2 at 30 km/h and hence avoiding the need to accelerate from a stationary start to link 2. Path 1 continues to be the most fuel efficient.

Undulations increase fuel consumption compared with a constant downhill slope. Including the within-link undulations forces

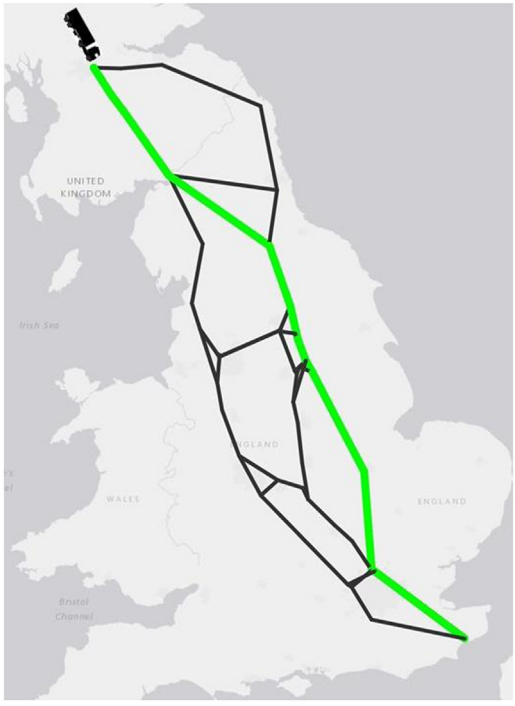

Fig. 4. Black links show the network. The optimal path is highlighted in green (Path 1). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

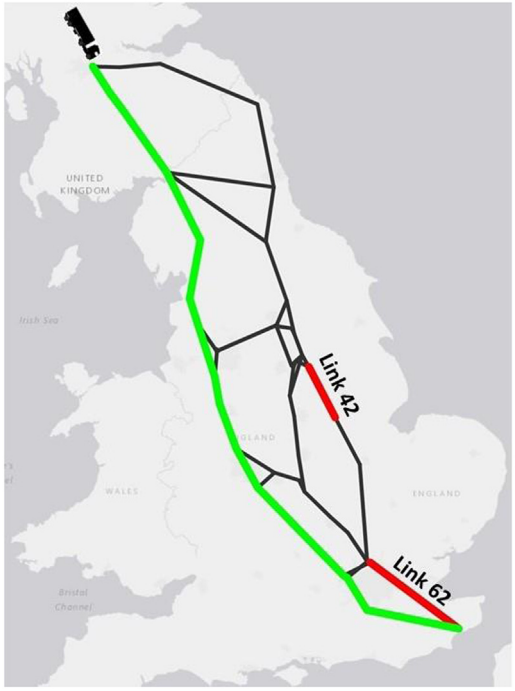

Fig. 5. Before departure, red links close. The new optimal path (Path 2) follows green links. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

recomputation of the optimal link speed profile for each combination of link-end speeds. The result is a significant increase in fuel consumption on link 2, sufficient to change the optimal path to path 2 (Ex 1C).

In Example 1D severe congestion occurs on link 4 from 08:30, leading to dynamic traffic conditions. The minimum fuel consumption path is path 2 (as in Ex 1C), departing at 08:00 and travelling fast on link 3 to enter link 4 before the link travel time increases. Naturally, this consumes additional fuel on link 3, but the total fuel consumption is still less than going via path 1 (with link 2 undulations). Having entered link 4 in time, and with no arrival time constraint, the truck can return to optimal cruising speed (since link 4 is flat). Note that the fuel consumed on link 4 is slightly less than in Ex 1C due to the faster link entry speed (90 km/h here) allowing an initial fuel-conserving deceleration.

In Example 1E en-route stops are possible. The traffic conditions captured in the LTTT (Table 6) result in the following optimal path plan: having departed at 08:00 when link 3 is uncongested, stop for 50 min at node 3, then continue on the now cleared link 4. There is a slight fuel penalty compared with Ex 1C since the node speed when stopping is set to be 0 km/h. Hence stopping incurs an additional acceleration event when departing.

6. Example 2: implementation on a larger network

As a case study for applying our methods, a simplified representation of the UK motorway network was developed, sufficient to consider routing a 40T truck from the edge of the city of Glasgow to the Eurotunnel. This 'primal' network of motorway-standard links suitable for a truck of such a size/weight comprises 51 nodes and 66 directional links. Online mapping resources were used to determine link lengths ($\min = 2.26$ km, $\max = 145.13$ km) and link gradients ($\min = -0.60^{o}$, $\max = 0.59$ deg).

Feasible node transfer speeds were set to be [0, 30, 50, 90] km/h. The range of link speeds [40 km/h, 110 km/h] was used to generate a range of possible link travel times for each link, discretised at 1 min intervals. In total 42,320 within-link speed profile optimization problems were solved. As in the preceding simple example, we used the fuel consumption model defined in Appendix A. Link gradients were used, but constant gradient was assumed on each link. The resulting lookup table records the optimal fuel consumption for each link, for each travel time and each combination of node transfer speeds.

To illustrate how traffic data can be incorporated, we use predicted link travel time data obtained by polling Bing Maps live traffic API every 15 min (for each link) for several days. For each link, these "raw" travel time data are resampled on a regular discrete grid (of 1 min) for link entry times. This gives us a dynamic link travel time timetable (LTTT) with entries such as those in Table 8.

As an illustrative example of optimal routing, we consider a journey made by a truck travelling from Glasgow to the Eurotunnel. The earliest start time is fixed to be 00:00 on 6th July 2019. A range of latest arrival times is considered. These limits are encoded into the STEN by only creating space-time links from the origin and to the destination that are within these bounds.

For each link entry time, the minimum link travel time is the maximum of (i) the LTTT travel time and (ii) the time to traverse the link at the upper speed-limit (110 km/h). The maximum link travel time is the maximum of (i) the dynamic LTTT travel time and (ii) the time to traverse the link at the lower speed limit (40 km/h). In sufficiently-congested conditions, when the LTTT corresponds to a link speed below 40 km/h, the only available link travel time is this single value.

The minimum fuel consumption path, considering departure time, link speed, node speed, stops and path is then achieved by Dijkstra shortest path search in the STEN. The performance statistics for the method are given in Table 9; the experiments were run on a standard desktop PC (Intel i7–4790 CPU 3.60 GHz with 32GB RAM), written in Matlab 2020a, and run under Windows 10. The main focus of our study is the space-time network created for the single day of 06.07.2019. As an experiment, a space-time network

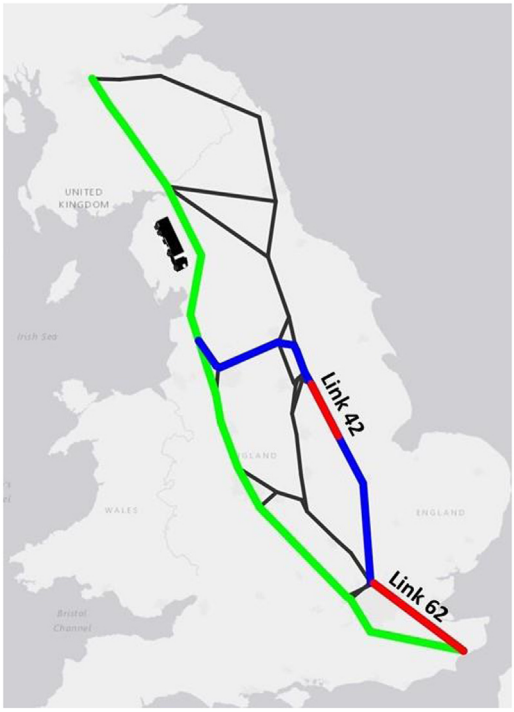

Fig. 6. At origin, red links are closed so the truck departs on Path 2 (green). While en route (position shown) red links re-open, truck diverts to blue/red path. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

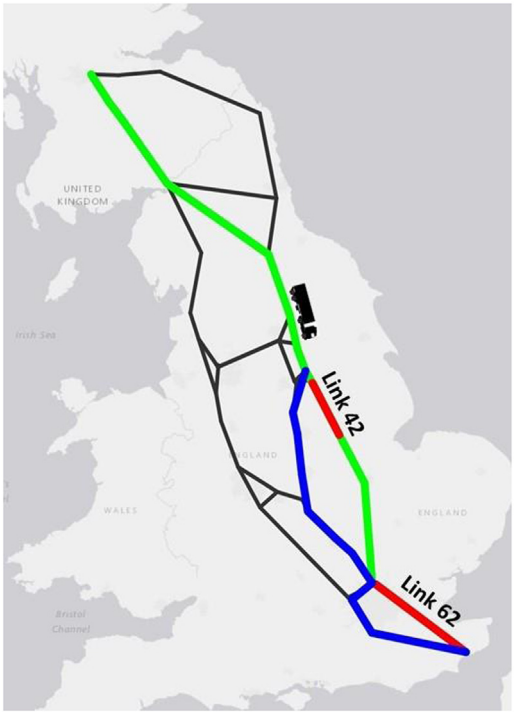

Fig. 7. At origin all links are open, truck departs on Path 1 (green). While en route (position shown) red links close. The truck diverts to the blue path. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

was also created for a multi-day period (first row of Table 9), at the same one-minute temporal resolution; it is not unusual for international truck journeys to take several days, so such an example is practically relevant. The space-time network in this case may have 276 million links, yet again the optimal path can be computed in a few seconds, providing evidence of the scalability of the method.

We return to the single day of 06.07.2019 to illustrate some of the phenomena demonstrated in Section 5, the results being in Table 10. With the earliest departure time 06:00, row 1 is unconstrained by arrival time. The optimal path is Path 1 (Fig. 4). Bringing forward the latest arrival time (row 2) we see an increase in fuel consumption, but without changing path.

If links 42 and 62 are closed (or equivalently have very long travel times) and this is known before departure, the optimal path switches to the west side of the network (Path 2 in Fig. 5). The latest arrival time of 18:00 becomes an active constraint since this path is longer and hence the average speed must be increased, consuming more fuel (row 3, Table 10)

If links 42 and 62 re-open while following Path 2, there may be the chance to divert and re-join Path 1. Path 3 is an example of an optimal re-routing. Fig. 6 shows the truck originally following the green path (since red links are closed). The truck is notified that the red links due to re-open before the truck's expected link entry time allowing it to follow the blue path. The resulting path length and fuel consumption are shown in row 4 of Table 10.

If links close while *en route*, the diversion options are likely to be worse than if this information is available before departure. Consider the truck departing on Path 1, following both node sequence and optimal link speed profiles (green path in Fig. 7). While *en route*, links 42 and 62 close. The truck optimally re-routes to follow the blue path. The latest arrival time constraint becomes active and the truck needs to speed up significantly on the much longer diversion path. Fuel consumption increases accordingly (Row 5 in Table 10).

7. Conclusions and further research

The dynamic and uncertain nature of the traffic conditions to be faced on a long truck delivery leg provide considerable potential for the use of real-time predictive information to improve the journey experience, and thereby reduce costs to operators, delays to those shipping/receiving goods and/or the environmental impacts of the journey. In the present paper, we have focused on reducing fuel consumption in a time-varying network. Such a focus sets particular challenges, given on the one hand the need to represent the influence of detailed micro-behaviour such as acceleration events on fuel use, and on the other hand the requirement to represent network effects, where time-varying congestion constraints may have significant influences on consumption far downstream from the location they occur, generating complex predictive dynamic constraints. The method developed in the present paper addresses this by embedding a link-level optimization with boundary values within an overall master problem that addresses the network level. This has the advantage that a bank of link-level optimization problems can be solved off-line, and a space-time network instantiated with the solutions. Then fast real-time optimizations at the network level may be performed based on real-time traffic information, with information on incidents or unexpected congestion (which imply lower shortest travel times) handled simply by deleting links of the space-time network.

In the paper, we have reported results in which the link level is represented by the VT-CPFM model, as described for example by Park et al. (2013). However, one advantage of the method presented is its flexibility, with the possibility to incorporate and test alternative models, such as those reviewed for eco-routing by Zhou et al. (2016). Miao et al. (2018) considered some additional complexity not captured in the model we have tested, such as the possibility for negative fuel consumption. This could also be incorporated in the method we have presented, since we are imagining

the links to be sufficiently long that any locally-negative consumption will be balanced out such that the overall consumption on a link will be positive, thus avoiding the problems with shortest path algorithms identified by Miao et al. One issue not explicitly addressed in the present paper is the issue of 'turning costs' identified by Nie and Li (2013), and incorporating such complexity could be an interesting direction for further development of the model. A final important question is how to make trade-offs between differing objectives facing freight operators, including various kinds of emissions, travel time and energy useage. For such problems the development of multi-objecitve optimization methods would be especially relevant (Ferreira et al., 2020).

The focus in the present paper has been on developing individual recommendations that might be provided to drivers. A future direction could be to also consider how drivers may actually respond to the recommendations (as in Javanmardi et al. (2017)) and potentially include the anticipated response in the optimization, to give a kind of behavioural routing policy. Alternatively, if we imagine routing information is provided to a much wider group of drivers, then the question arises as to the system impacts of many vehicles following such advice. In this spirit, Ahn and Rakha (2013) and Bandeira et al. (2018) considered a kind of dynamic eco-routing problem at the network level, defining a user equilibrium model of the resulting problem. Thus it is possible to analyse the effect of optimization as it scales up from the individual level to the wider level of a fleet or class of vehicles. Finally, we may consider combining the optimization with additional dimensions of control, such as that of the infrastructure. For example, Staubach et al. (2014) considered cooperative systems for eco-driving that communicate with traffic lights Table 2.

Acknowledgements

The research reported in this paper was conducted within the projects Optitruck (optitruck.eu) and Modales (modales-project.eu), which were financially supported by the European Union's Horizon 2020 Research and Innovation Program, under grant agreement numbers 713788 and 815189.

Appendix A. Fuel Consumption Function

Many methods could be used to compute instantaneous fuel consumption. To be consistent with the proposed methodology, the minimum requirement is that it should be a function of the vehicle speed and acceleration, and the road gradient. Fast evaluation of the fuel consumption function is not strictly necessary (since it is only explicitly used offline), and so many candidates could be included. However, since it must be optimized repeatedly under many different boundary conditions, then there is some advantage to using an analytic model that is both fast to evaluate and repeatable. Here we adopt a closed-form analytic function, inspired by the functional form of the VT-CPFM (Park et al., 2013), based on the instantaneous slope θ, speed v, and acceleration a:

$$F(v,a;\theta,\beta) = \Big[\big(\beta_1+\beta_2 v^2+\beta_3\sin\theta+\beta_4 a\big)^2 v^2+\beta_6\big(\beta_1+\beta_2 v^2 + \beta_3\sin\theta+\beta_4 a\big)v+\beta_5\Big]_+$$

Fitting this model to instantaneous fuel consumption data results in estimates for the coefficients, β_i, which implicitly capture the dependence of $F(t)$ on a vehicle's characteristics such as its mass, load, rolling reistance, frontal area, etc.

We fitted this fuel consumption model to simulation data of a 40T diesel truck (using the modelling approach described in Gao et al., 2020), which resulted in the parameter values in Table 11. The resulting relationship between fuel consumption and speed is illustrated in Fig. 8, for the case of zero gradient and acceleration; the optimal cruising speed is 65 km/h.

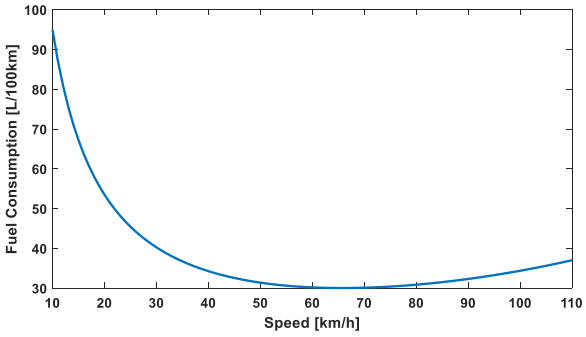

Fig. 8. Marginal relationship between fuel consumption and speed for assumed fuel consumption model (for the case of zero gradient, zero acceleration).

Appendix B. Practical application of the procedure

Link optimisation

1. Set the following:
   (a) time discretisation step $= \delta$
   (b) origin node $= i^*$
   (c) destination node $= j^*$
   (d) set of nodes where stops are allowed $= V_{STOP}$
   (e) set of node transfer speeds $= \mathcal{U}$
   (f) overall time horizon $T \subseteq \delta\mathbb{N}$
   (g) departure/arrival time constraints $= T^D, T^A \subseteq T$
2. Load/define link characteristics and connectivity i.e. link-node incidence matrix
3. Load link travel time timetable (LTTT):
   (a) Data records the minimum travel time for each link at each time of day
   (b) If needed, subsample to have LTTT on discrete time basis
4. For each road link connecting node i to node j
   (a) set the min/max travel time to consider for this link (consider link travel time timetable data and practical minimum for average link speed).
   (b) for every combination of (i) link entry speed, $u_i \in \mathcal{U}$ (ii) link exit speed, $u_j \in \mathcal{U}$ (iii) link entry time, t_i, (iv) link exit time, t_j that could arise from the link entry time, compute the fuel-optimal objective function value $h_{ij}(t_i, t_j, u_i, u_j)$

STEN Construction

5. For each node $i \in V$, create space-time-speed nodes (i,t,u) for every discrete time point $t \in T$ and each node transfer speed $u \in \mathcal{U}$
6. Create a dummy source node, i_{SOURCE},and a dummy sink node, j_{SINK}.
7. If there is a physical link connecting node i to node j then we connect the associated space-time nodes (i,t,u) to $(j,t+\theta,w)$ for all $u,w \in \mathcal{U}$, whenever θ is a feasible travel time for this link.
   Travel time θ must not violate the LTTT minimum travel time (when leaving node i at time t) and should not exceed the maximum link travel time considered (see step 4a).
8. The cost of the link connecting (i,t,u) to $(j,t+\theta,w)$ is $h_{ij}(t,t+\theta,u,w)$ (as computed in step 4b above).
9. Stop nodes: for nodes $i \in V_{STOP}$ where stops are allowed
   (a) connect $(i,t,0)$ to $(i,t+\theta,0)$ for all allowable stop durations θ.
   (b) Assign a zero costs to these links (as stop does not in itself consume fuel).

10. Departure/Arrival links (which have zero cost):
    (a) connect, i_{SOURCE} to $(i^*, t, 0)$ for each permitted departure time $t \in T^D$
    (b) connect, $(j^*, t, 0)$ to j_{SINK} for each permitted arrival time $t \in T^A$

Note that the STEN can be trimmed as follows:

- Not all space-time nodes are needed in step 5. Consider the earliest allowed departure time from the origin, t^*, and the shortest travel time from the origin to each node in the network, $\{\theta_i^*\}$. Space-time nodes, (i, t, u), are only required for $t \geq t^* + \theta_i^*$.
- Prior to the real time optimal path calculation, for any changes to accessible nodes or feasible link travel times, remove elements as appropriate from the STEN.

The shortest path can now be computed using standard algorithms from i_{SOURCE} to j_{SINK} with link costs as stated above.

References

Abdelmegeed, M. A. E., & Rakha, H. (2017). Heavy-duty diesel truck emissions modeling. *Transportation Research Record, 2627*, 26–35.

Ahn, K., & Rakha, H. A. (2013). Network-wide impacts of eco-routing strategies: A large-scale case study. *Transportation Research Part D, 25*, 119–130.

Alam, M. S., & McNabola, A. (2014). A critical review and assessment of Eco-Driving policy & technology: Benefits & limitations. *Transport Policy, 35*, 42–49.

Ayyildiz, K., Cavallaro, F., Nocera, S., & Willenbrock, R. (2017). Reducing fuel consumption and carbon emissions through eco-drive training. *Transportation Research Part F, 46*, 96–110.

Bandeira, J. M., Fernandes, P., Fontes, T., Pereira, S. R., Khattak, A. J., & Coelho, M. C. (2018). Exploring multiple eco-routing guidance strategies in a commuting corridor. *International Journal of Sustainable Transportation, 12*(1), 53–65.

BEIS. (2017). 2015 UK greenhouse gas emissions, final figures. *2016 UK greenhouse gas emissions, provisional figures. statistical release: National statistics. February 7th 2017*. Department for Business, Energy and Industrial Strategy.

Bektas, T., & Laporte, G. (2011). The pollution-routing problem. *Transportation Research Part B, 45*, 1232–1250.

Boriboonsomsin, K., & Barth, M. (2009). Impacts of road grade on fuel consumption and carbon dioxide emissions evidenced by use of advanced navigation systems. *Transportation Research Record, 2139*, 21–30.

Dehkordi, S. G., Larue, G. S., Cholette, M. E., Rakotonirainy, A., & Rakha, H. A. (2019). Ecological and safe driving: A model predictive control approach considering spatial and temporal constraints. *Transportation Research Part D: Transport and Environment, 67*, 208–222.

Díaz-Ramirez, J., Giraldo-Peralta, N., Flórez-Ceron, D., Rangel, V., Mejía-Argueta, C., Huertas, J. I., et al., (2017). Eco-driving key factors that influence fuel consumption in heavy truck fleets: A Colombian case. *Transportation Research Part D, 56*, 258–270.

Eglese, R., Maden, W., & Slater, A. (2006). A Road Timetable to aid vehicle routing and scheduling. *Computers & Operations Research, 33*(12), 3508–3519.

Ferreira, J. C., Teresinha, M., Steiner, A., & Canciglieri, O., Jr (2020). Multi-objective optimization for the green vehicle routing problem: A systematic literature review and future directions. *Cogent Engineering, 7*, 1.

Franceschetti, A., Demir, E., Honhon, D., Van Woensel, T., Laporte, G., & Stobbe, M. (2017). A metaheuristic for the time-dependent pollution-routing problem. *European Journal of Operational Research, 259*, 972–991.

Franceschetti, A., Honhon, D., Van Woensel, T., Bektaş, T., & Laporte, G. (2013). The time-dependent pollution-routing problem. *Transportation Research Part B: Methodological, 56*, 265–293.

Fröberg, A., Hellström, E., & Nielsen, L. (2006). Explicit fuel optimal speed profiles for heavy trucks on a set of topographic road profiles. *SAE Technical Paper 2006-01-1071*.

Gao, J., Chen, H., Dave, K., Chen, J., & Jia, D. (2020). Fuel economy and exhaust emissions of a diesel vehicle under real traffic conditions. *Energy Science & Engineering, 8*, 1781–1792.

He, C. R., Maurer, H., & Orosz, G. (2016). Fuel consumption optimization of heavy–duty vehicles with grade, wind, and traffic information. *ASME Journal of Computational and Nonlinear Dynamics, 11*(6).

Held, M., Flärdh, O., & Mårtensson, J. (2019). Optimal speed control of a heavy-duty vehicle in urban driving. *IEEE Transactions on Intelligent Transportation Systems, 20*(4), 1562–1573.

Hellström, E., Ivarsson, M., Åslund, J., & Nielsen, L. (2009). Look-ahead control for heavy trucks to minimize trip time and fuel consumption. *Control Engineering Practice, 17*, 245–254.

Huang, Y., Ng, E. C. Y., Zhou, J. L., Surawski, N. C., Chan, E. F. C., & Hong, G. (2018). Eco-driving technology for sustainable road transport: A review. *Renewable and Sustainable Energy Reviews, 93*, 596–609.

Huang, Y., Zhao, L., Van Woensel, T., & Gross, J.-. P. (2017). Time-dependent vehicle routing problem with path flexibility. *Transportation Research Part B, 95*, 169–195.

Hvattum, L. M., Norstad, I., Fagerholt, K., & Laporte, G. (2013). Analysis of an exact algorithm for the vessel speed optimization problem. *Networks, 62*, 132–135.

IEA. (2018). *CO2 emissions from fuel combustion*. International Energy Agency.

Jabali, O., van Woensel, T., & deKok, A. G. (2012). Analysis of travel times and CO2 emissions in time-dependent vehicle routing. *Production and Operations Management, 21*, 1060–1074.

Javanmardi, S., Bideaux, E., Trégouët, J. F., Trigui, R., Tattegrain, H., & Nicouleau Bourles, E. (2017). Driving style modelling for eco-driving applications. *IFAC PapersOnLine, 50*(1), 13866–13871.

Jiménez, F., López-Covarrubias, J. L., Cabrera, W., & Aparicio, F. (2013). Real-time speed profile calculation for fuel saving considering unforeseen situations and travel time. *IET Intelligent Transport Systems, 7*(1), 10–19.

Kramer, R., Maculan, N., Subramanian, A., & Vidal, T. (2015). A speed and departure time optimization algorithm for the pollution-routing problem. *European Journal of Operational Research, 247*, 782–787.

Kuo, Y. (2010). Using simulated annealing to minimize fuel consumption for the time-dependent vehicle routing problem. *Computers & Industrial Engineering, 59*, 157–165.

Lin, C., Choy, K. L., Ho, G. T. S., Chung, S. H., & Lam, H. Y. (2014). Survey of green vehicle routing problem: Past and future trends. *Expert Systems with Applications, 41*, 1118–1138.

McKinnon, A. (2011). Improving the sustainability of road freight transport by relaxing truck size and weight restrictions. In P Evangelista, A McKinnon, & Sweeney (Eds.), *Supply chain innovation for competing in highly dynamic markets: Challenges and solutions*. London: Premier Reference Source.

Miao, C., Liu, H., Zhu, G. G., & Chen, H. (2018). Connectivity-based optimization of vehicle route and speed for improved fuel economy. *Transportation Research Part C, 91*, 353–368.

Nasri, M. I., Bektas, T., & Laporte, G. (2018). Route and speed optimization for autonomous trucks. *Computers and Operations Research, 100*, 89–101.

Nie, Y., & Li, Q. (2013). An eco-routing model considering microscopic vehicle operating conditions. *Transportation Research Part B: Methodological, 55*, 154–170.

Ozatay, E., Ozguner, U., & Filev, D. (2017). Velocity profile optimization of on road vehicles: Pontryagin's Maximum Principle based approach. *Control Engineering Practice, 61*, 244–254.

Park, S., Rakha, H. A., Ahn, K., & Moran, K. (2013). Virginia tech comprehensive power-based fuel consumption model (VT-CPFM): Model validation and calibration considerations. *International Journal of Transportation Science and Technology, 2*(4), 317–336.

Qian, J., & Eglese, R. (2016). Fuel emissions optimization in vehicle routing problems with time-varying speeds. *European Journal of Operational Research, 248*, 840–848.

Rodrigues, V. S., Piecyk, M., Mason, R., & Boenders, T. (2015). The longer and heavier vehicle debate: A review of empirical evidence from Germany. *Transportation Research Part D, 40*, 114–131.

Saboohi, Y., & Farzaneh, H. (2009). Model for developing an eco-driving strategy of a passenger vehicle based on the least fuel consumption. *Applied Energy, 86*, 1925–1932.

Saerens, B., & Van den Bulck, E. (2013). Calculation of the minimum-fuel driving control based on Pontryagin's maximum principle. *Transportation Research Part D, 24*, 89–97.

Scora, G., Boriboonsomsin, K., & Barth, M. (2015). Value of eco-friendly route choice for heavy-duty trucks. *Research in Transportation Economics, 52*, 3–14.

Sims, R., Schaeffer, R., Creutzig, F., Cruz-Núñez, X., D'Agosto, M., Dimitriu, D., et al., (2014). Transport. In O. Edenhofer, R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, & K. Seyboth (Eds.), Climate change 2014: Mitigation of climate change. *contribution of working group III to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press [(eds.)]..

SloCaT. (2018). *Transport and climate change global status report 2018* Available at http://slocat.net/tcc-gsr .

Staubach, M., Schebitz, N., Köster, F., & Kuck, D. (2014). Evaluation of an eco-driving support system. *Transportation Research Part F, 27*, 11–21.

Vagg, C., Brace, C. J., Hari, D., Akehurst, S., Poxon, J., & Ash, L. (2013). Development and field trial of a driver assistance system to encourage eco-driving in light commercial vehicle fleets. *IEEE Transactions on Intelligent Transportation Systems, 14*(2), 796–805.

Watling, D. P., Connors, R. D., Milne, D. S., & Chen, H. (2019). Optimization of route choice, speeds and stops in time-varying networks for fuel-efficient truck journeys. *European Journal of Transport and Infrastructure Research, 19*(4), 265–290.

Wu, Y., Zhao, X., Rong, J., & Zhang, Y. (2018). The effectiveness of eco-driving training for male professional and non-professional drivers. *Transportation Research Part D, 59*, 121–133.

Zavalko, A. (2018). Applying energy approach in the evaluation of eco-driving skill and eco-driving training of truck drivers. *Transportation Research Part D, 62*, 672–684.

Zeng, W., Miwa, T., & Morikawa, T. (2020). Eco-routing problem considering fuel consumption and probabilistic travel time budget. *Transportation Research Part D, 78*, Article 102219.

Zeng, X., & Wang, J. (2018). Globally energy-optimal speed planning for road vehicles on a given route. *Transportation Research Part C, 93*, 148–160.

Zhou, M., Jin, H., & Wang, W. (2016). A review of vehicle fuel consumption models to evaluate eco-driving and eco-routing. *Transportation Research Part D, 49*, 203–218.