J Math Model Algor (2007) 6:151–173
DOI 10.1007/s10852-006-9054-1

A Queueing Framework for Routing Problems with Time-dependent Travel Times

Tom Van Woensel · Laoucine Kerbache · Herbert Peremans · Nico Vandaele

Received: 28 September 2004 / Accepted: 3 September 2005 /
Published online: 25 November 2006

Abstract Assigning and scheduling vehicle routes in a dynamic environment is a crucial management problem. Despite numerous publications dealing with efficient scheduling methods for vehicle routing, very few addressed the inherent stochastic and dynamic nature of travel times. In this paper, a vehicle routing problem with time-dependent travel times due to potential traffic congestion is considered. The approach developed introduces the traffic congestion component based on queueing theory. This is an innovative modelling scheme to capture the stochastic behavior of travel times as it generates an analytical expression for the expected travel times as well as for the variance of the travel times. Routing solutions that perform well in the face of the extra complications due to congestion are developed. These more realistic solutions have the potential to reduce real operating costs for a broad range of industries which daily face routing problems. A number of datasets are used to illustrate the appropriateness of the novel approach. Moreover it is shown that static (or time-independent) solutions are often infeasible within a congested traffic environment which is generally the case on European road networks. Finally, the effect of travel time variability (obtained via the queueing approach) is quantified for the different datasets.

T. Van Woensel (✉)
TU/e Eindhoven University of Technology, Department of Technology Management,
Den Dolech 2, Pav. F05, NL5600 MB Eindhoven, The Netherlands
e-mail: t.v.woensel@tm.tue.nl

L. Kerbache
HEC Paris, France

H. Peremans · N. Vandaele
University of Antwerp, Belgium

Key words time-dependent routing problems · queueing · dynamic travel times.

Mathematics Subject Classifications (2000) 60K25 · 60K30 · 90B06 · 90B10 · 90B15.

1 Introduction

Transportation is a main component of supply chain competitiveness since it plays a major role in the inbound, inter-facility, and outbound logistics. Transportation costs represent approximately 40%–50% of total logistics and 4%–10% of the product selling price for many companies [15]. Transportation decisions directly affect the total logistic costs. The passage of the transportation deregulation acts in the 1980s in the USA and in the 1990s in the EU drastically changed the business climate within which the transportation managers operate. Within the EU, the competition is becoming intense between transporters since they often operate at transnational levels and they must provide higher levels of service with lower costs to meet the various needs of customers. In this context, assigning and scheduling vehicle routes is a crucial management problem. Providing non-dominated vehicle routing planning schedules is a very hard combinatorial problem. Yet, a manager must rely on management techniques, using a proactive approach to identify and solve transportation problems and to provide the company with a competitive advantage in the marketplace. Despite numerous publications dealing with efficient scheduling methods for vehicle routing, very few address the inherent stochastic nature of this problem. Most research in this area has focused on routing and scheduling that incorporates variable customer demands. However, there has been very little research on routing and scheduling explicitly incorporating the congestion component directly. A selected literature overview will be presented in the next section.

The main contributions are twofold:

First, dynamic travel times are introduced in the vehicle routing problem. The approach developed here introduces the traffic congestion component modelled through a queueing approach. As such, the inherent stochastic nature of travel times is captured using a queueing approach to traffic flow theory. Speed is the result of the stochastic process of different vehicles interacting with each other under certain circumstances (e.g., bad weather). Combining speed with distance to be travelled leads to the expected travel time. Results show that total expected travel times can be improved significantly when explicitly taking into account congestion during the optimization. This is important not just because speed profiles can affect the objective(s) of the optimization, but also because the best solutions known for a static problem are in general infeasible when applied in a dynamic world.

Secondly, the analytical approach to congestion based on queueing models not only allows for the calculation of the expected travel times but also the variance of the travel times can be obtained. The latter allows for an evaluation of the routes based on the uncertainty involved. Combining the expected travel time with the variance of the travel time in the objective function generates new interesting insights on the impact of uncertainty on total travel times. Explicitly taking into account the variance of the travel times thus allows for evaluating routes on the risk involved. When doing so, depending on the risk profile of the manager/customer (risk averse, neutral or seeking), the planned route could be substantially different.

This paper is organized as follows: in Section 2 the routing models are formally defined, Section 3 gives a classification on routing problems with time-dependent travel times. In Section 4, the framework for modelling time-dependent travel times is discussed in detail, and in Section 5 computational results are presented. Section 6 deals with future research opportunities. The paper ends with general conclusions.

2 Defining Routing Models

The vehicle routing problem (VRP) can be described as a more general version of the well-known travelling salesman problem (TSP). The VRP aims to construct a set of shortest routes for a fleet of vehicles of fixed capacity. Each customer is visited exactly once by one vehicle which delivers the demanded amount of goods to the customer. Each route has to start and end at a depot, and the sum of the demands of the visited customers on a route must not exceed the capacity of the vehicle. Another constraint occurring in the real world is that the customer may specify time intervals in which he will be able to receive the deliveries. This additional restriction leads to the vehicle routing problem with time window constraints ($VRPTW$). The time windows can be either soft or hard indicating whether the time windows can be violated or not [37].

In this paper, it is assumed that there is only one depot from where the routes start and end for each vehicle. A homogeneous fleet consisting of several vehicles with fixed capacity, while each customer's demand is pre-determined i.e., static and deterministic demand (and not restricted by time windows). Formally, the vehicle routing problem can be represented by a complete weighted graph $G = (V, A, c)$ where $V = \{v_0, v_1, ..., v_n\}$ is a set of vertices and $A = \{(v_i, v_j) : i <> j\}$ is a set of arcs. The vertex 0 denotes the depot; the other vertices of V represent cities or customers. The non-negative weights $c_{(v_i,v_j)}$ which are associated with each arc (v_i, v_j) represent the cost (distance, travel time or travel cost) between v_i and v_j. For each customer, a non-negative demand qd_{v_i} and a non-negative service time d_{v_i} is given ($d_{v0} = 0$ and $qd_{v0} = 0$). The aim is then to find the minimum cost vehicle routes where the following conditions hold: every customer is visited exactly once by exactly one vehicle; all vehicle routes start and end at the single depot; every vehicle route has a total demand not exceeding the vehicle capacity Q; every vehicle route has a total route length not exceeding the maximum length L.

3 A Classification of Routing Problems

There has been limited research on routing and scheduling with congestion dependent travel times. Few researchers (see e.g., [19, 30, 31, 38, 39]) have dealt with dynamic travel times. Bertsimas and Simchi-Levi [7] provide an interesting survey of static (i.e., time-independent) vs. dynamic (i.e., time-dependent) routing problems (e.g., the travel time being a function of the time of the day due to traffic congestion or that the time in which demand occurs is a renewal process) as well as deterministic vs. stochastic routing problems (i.e., depending whether some of the characteristics follow a certain probability distribution or not; typical examples are probabilistic demand quantities or probabilistic travel times). If it seems reasonable to assume that the service time at each vertex (customer) is known in advance, it is definitely not the case for the travel time between two vertices. In fact, the travel times are the result of a stochastic process related to traffic congestion. Clearly, travel times depend greatly on the different number of vehicles occupying the road and on their speeds. It is assumed that service times at each node are known in advance. The travel times however are the result of a stochastic process due to the different number of vehicles occupying the road and their respective speeds. This section gives a comprehensive classification of vehicle routing problems based on the use of travel times following the logic outlined above. It should be noted that this classification only concerns the time dependency of travel times. For a more detailed review, we refer the reader to Kerbache and Van Woensel [34].

3.1 Routing Problems with Time-independent Travel Times

The time-independent routing problem can be referred to as the standard case. For this problem, most of the models and their solution approaches, assume that all characteristics are independent of the time of the day. Therefore, these models are far from real-life applications where, for instance congestion occurs on the used road network. Laporte [37] surveys the main research results on the time-independent vehicle routing problem. To solve these models, both exact and heuristic methods are available. Laporte [37] classifies exact algorithms into three categories: direct tree search methods, dynamic programming and integer programming. For larger and complex problems, heuristic algorithms are more appropriate. In the literature, the time-independent routing problem is modelled differently depending on the deterministic or stochastic characteristics of the parameters.

3.1.1 *Routing Problems with Deterministic Time-independent Travel Times*

Many heuristic methods for these types of problems are derivations of methods developed for the *TSP*: e.g., the nearest neighbor algorithm, insertion algorithms, and tour improvement procedures. Nevertheless, there are some algorithms that are specifically developed for the vehicle routing problem. These heuristics are basically classified as follows [41]: constructive heuristics (e.g., the savings algorithm [13], and [25]), two-step heuristics (e.g., cluster-first-route-second methods, route-first-cluster-second methods, Christofides–Mingozzi–Toth two-phase algorithm [12]), incomplete optimization, local search heuristics (sweep algorithm [25]), metaheuristics (tabu search algorithm by [37]) and space filling curves.

3.1.2 *Routing Problems with Stochastic Time-independent Travel Times*

Stochastic Time-independent Routing Problems arise whenever some elements of the deterministic routing problems are assumed to be random. Common examples are stochastic demands and stochastic travel times. Sometimes, the set of customers to be visited is not known with certainty [24]. Since they combine the characteristics of stochastic and integer programs, these type of problems are often regarded as computationally intractable. Stochastic *VRPs* can be classified within the framework of stochastic programming. Stochastic programs are modelled in two stages. In a first stage, a planned or *a priori* solution is determined. The realizations of the random variables are then disclosed and, in a second stage, a recourse or corrective action is then applied to the first stage solution [24, 33].

3.2 Routing Problems with Time-dependent Travel Times

In this paper, the VRP problem considered deals with dynamic travel times. The motivation for using dynamic models is that the vehicles in the VRP operate in a traffic network which will be congested depending upon the time of the day. In the dynamic VRP, the non-negative weights c^p which are associated with each arc (i, j) represent the dynamic travel time between i and j starting in time bucket p. Whereas in the static VRP, these costs are not associated with a time dimension (i.e., the related costs do not change over time). The literature related to vehicle routing with dynamic travel times is rather scarce [5, 6, 19, 30, 31, 38, 39]. The reason is that the time-dependent VRP is much harder to model and to solve. The major shortcoming of the available models is the modelling of the travel time function. It is often discretized into a limited fixed number of time intervals (e.g., morning, midday and afternoon) with a distinct associated fixed mean speed. However, these speeds are modelled in an arbitrary way. For instance, Brown et al. [9] and Shen and Potvin [43] used a rough approximation of travel time by manually re-sequencing the route taking into account congestion. Ichoua et al. [31] used a model based on discrete travel speeds by adding correction factors to model the congestion. Models based on continuous travel times are very complex to solve (see [29]) and thus many simplifying assumptions had to be introduced to keep the model tractable. In general, the dynamic travel times can be modelled in two ways: deterministically or stochastically.

3.2.1 *Routing Problems with Deterministic Time-dependent Travel Times*

In the deterministic case, the travel times are known in advance and plugged in the solution heuristic depending upon the period of the day. The travel times are then a function of distance and mean speed. For instance, Ichoua et al. [31] consider three distinct time periods (where the first and third periods stand for the morning and evening rush hours, respectively, and the second period corresponds to the middle part of the day) and three different types of road links. This approach has been implemented within a parallel tabu search developed by Taillard et al. [44] for the fixed travel time version.

3.2.2 *Routing Problems with Stochastic Time-dependent Travel Times*

In the stochastic time-dependent models, the solution procedure takes into account the stochastic nature of the travel times. Travel times are in this case the result of taking into account not only mean travel time but ideally the travel time distribution itself. As the travel time distribution is derived from the speed distribution and the known distances, the approach requires realistic speed distributions. However, in practice neither travel time nor speed distributions are available in closed form [21]. For a general review on the stochastic vehicle routing, the reader is referred to Gendreau et al. [23].

4 Framework for Time-dependent Travel Times

In the time-dependent routing problem, the key issue is the computation of the travel times on arcs dependent upon the time period. The travel time T_{ij}^{p} during time period p is determined as the distance from i to j divided by the speed on the arc or:

$$T_{ij}^{p} = \frac{d_{ij}}{v_{ij}^{p}} \tag{1}$$

Hence, to determine the travel time on arc (i, j), one needs information on the distance between (i, j) and on the travel speed for that arc at time $p : v_{ij}^{p}$. The distance is readily available in the time-independent routing models, but the speed is a new variable that needs to be specified. In this paper, the time-dependent speeds are obtained using queueing models for traffic flows [28, 46]. This section is organized as follows: first, the queueing approach to traffic flow is presented, then the procedure to obtain the expected travel time is explained in detail, next the determination of the variance of the travel time is elaborated and finally, the added value of using queueing theory is discussed in detail.

4.1 A Queueing Approach

It is often observed that the speed for a certain time period tends to be reproduced whenever the same traffic flow is observed. Based on this observation, it seems reasonable to postulate that, if traffic conditions on a given road are stationary, there should be a relationship between flow, speed, and density. This relationship results in the concept of speed-flow-density diagrams. These diagrams describe the interdependence of traffic flow (q), density (k) and speed (v). The seminal work on speed-flow diagrams was the paper by [27]. Using well-known formulas of queueing models, speed-flow-density diagrams like the one shown in Figure 1 can be generated.

Figure 1 illustrates that, although every speed v corresponds with one traffic flow q, the reverse is not true. There are two speeds for every traffic flow: an upper branch (v_2) where speed decreases as flow increases and a lower branch (v_1) where speed increases. Intuitively it is clear that, as the flow moves from zero (at maximum speed v_f) to q_{max}, congestion increases but the flow rises because the decline in speed is over-compensated by the higher traffic density. If traffic tends to grow past q_{max}, flow falls again because the decline in speed more than offsets the additional vehicle numbers, further increasing congestion [16]. The flow-density diagram and the speed-density diagrams are an equivalent representation and can be interpreted in the same way.

Traditionally, these speed-flow-density diagrams are modelled empirically: speed and flow data are collected for a specific road and curves are fitted onto the data [16]. This traditional approach is limited in terms of predictive power and sensitivity analysis. Vandaele et al. [46] and Heidemann [28], showed that queueing models can also be used to explain uninterrupted traffic flows and thus offers a more practical approach, useful for sensitivity analysis, forecasts, etc. Jain and Smith [32] describe in their paper a state-dependent $M/G/C/C$ queueing model for traffic flows. Part of their logic is used here to extend our queueing models to state-dependent ones. Also a lot of research is done on a travel time-flow model originating from Davidson [17].

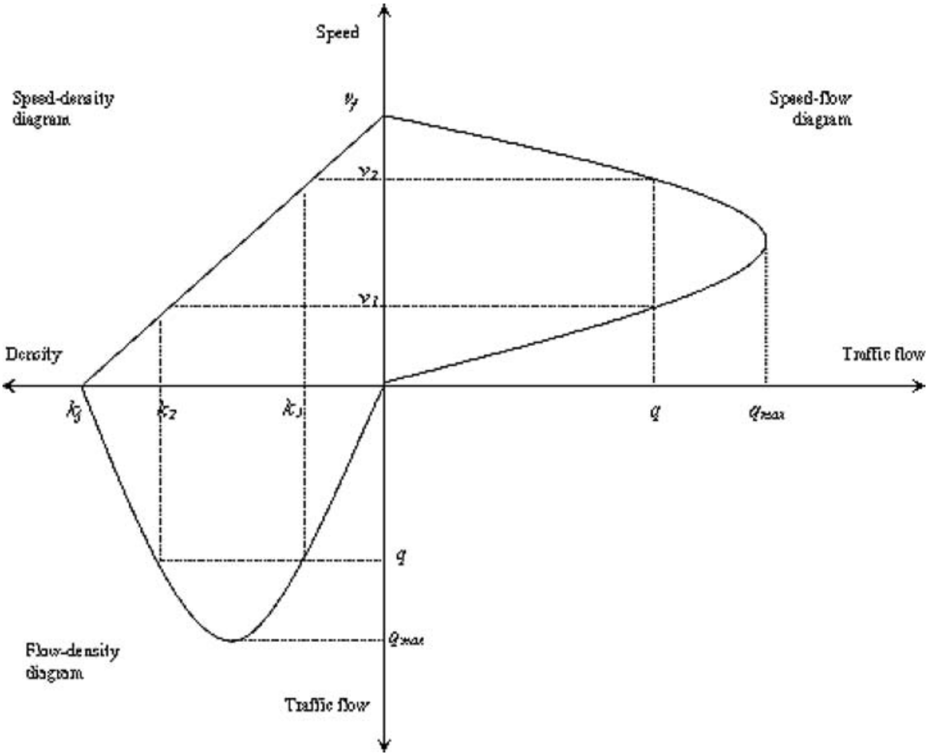


Figure 1 The relations between the speed-flow, the speed-density, and the flow-density diagrams.

The model is based on some concepts of queueing theory but a direct derivation has not been clearly demonstrated [2, 3].

In a queueing approach to traffic flow analysis, roads are subdivided into segments, with length equal to the minimal space needed by one vehicle on that road (Figure 2). Define k_j as the maximum traffic density (i.e., average maximum number of cars on a road segment). This length is then equal to $1/k_j$ and matches the minimal space needed by one vehicle on that road. Each road segment is then considered as a service station, in which vehicles arrive at a certain rate λ and get served at another rate μ [28, 46, 51].

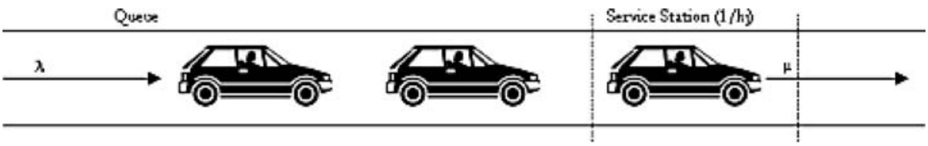


Figure 2 Queueing representation of traffic flows.

Table I The specific form of W_q for each queueing model

Queueing model	W_q
$GI/G/1KLB$	$\frac{\rho}{(1-\rho)}\frac{(c_a^2+c_s^2)}{2}\frac{1}{k_j v_f}\exp\left[\frac{-2(1-\rho)(1-c_a^2)^2}{3\rho(c_a^2+c_s^2)}\right]$
$GI/G/mK$	$\left(\frac{c_a^2+c_s^2}{2}\right)\left(\frac{\rho^{\left(\sqrt{2(m+1)}-1\right)}}{m(1-\rho)}\right)\left(\frac{1}{k_j v_f}\right)$
$GI/G/mW$	$\phi\left(\frac{c_a^2+c_s^2}{2}\right)\left(\frac{1}{k_j v_f}\right)W_{qM/M/k}$ With ϕ a correction factor defined in Whitt 1993([49]) and $W_{qM/M/m}$ the formula for the waiting time in an $M/M/m$ queue.

Vandaele et al. [46] developed different queueing models.[1] The $M/M/1$ queueing model (exponential arrival and service rates) is considered as a base case, but due to its specific assumptions regarding the arrival and service processes, it is not useful to describe real-life situations. Relaxing the specifications for the service process of the $M/M/1$ queueing model, leads to the $M/G/1$ queueing model (generally distributed service rates). Relaxing both assumptions for the arrival and service processes results in the $GI/G/m$ queueing model. Heidemann [28] and Vandaele et al. [46] showed that the speed v can be calculated by dividing the length of the road segment ($\frac{1}{k_j}$) by the total time in the system (W).

$$v = \frac{1/k_j}{W} \tag{2}$$

The total time in the system W is then different depending upon the queueing model used. The total time in the system W is then the sum of the waiting time W_q and the service time W_p, or $W = W_q + W_p$. Table I shows the specific form of W_q for the general queueing models. For the $GI/G/m$ queueing models, no exact solutions are available and one must rely on approximations. Here, three approximations are considered: the Kramer–Lagenbach-Belz approximation [36] is widely used but is limited to single servers only. To cope with multiple lanes, the heavy traffic or Kingman approximation [35] and the Whitt approximations [49] with multiple servers are used.

Results show that the developed queueing models can be adequately used to model traffic flows [52]. Moreover due to the analytical character of these models, they are very suitable to be incorporated in other models, e.g., the VRP.

In general, formula (2) can be rewritten in the following basic form with Ω a congestion pressure variable:

$$v = \frac{v_f}{1+\Omega} \tag{3}$$

Formula (3) shows that the speed is only equal to the maximum speed v_f if the factor Ω is zero. For positive values of Ω, v_f is divided by a number strictly larger

[1] In this paper, queueing models are referred to using the Kendall notation, consisting of several symbols – e.g., $M/G/1$. The first symbol is shorthand for the distribution of inter-arrival times, the second for the distribution of service times and the last one indicates the number of servers in the system.

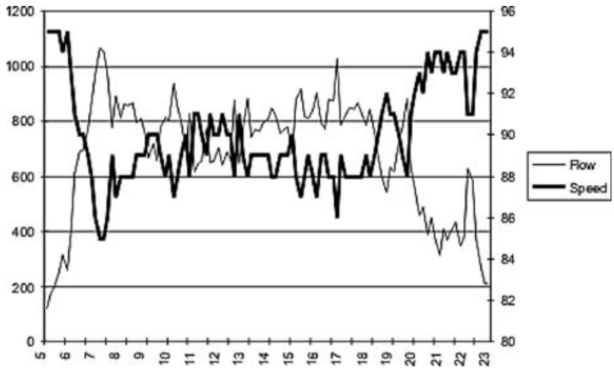


Figure 3 Example of a speed-flow diagram over time.

than one and speed is reduced. The factor Ω is thus the influence of congestion on speed. High congestion (reflected in a high Ω) leads to lower speeds than the maximum. The factor Ω is a function of a number of parameters depending upon the queueing model chosen: the traffic intensity, the coefficient of variation of service times and coefficient of variation of inter-arrival times. High coefficients of variation or a high traffic intensity will lead to a value of Ω strictly larger than zero. Actions to increase speed (or decrease travel time) should then be focussed on decreasing the variability or on influencing the traffic intensity, for example by manipulating the arrivals (arrival management and ramp metering). In the remainder of this paper, the $GI/G/m$ queueing models using the Whitt approximations will always be used. These have been widely used for their robustness and accuracy (see also [49]).

In Figure 3, an example is given for a speed and flow profile over time. The flow profile is based on observed data; the speeds are obtained using the queueing approach to traffic flows (only the upper speed v_2 of the speed-flow-density diagram is shown in Figure 1). As one can see, speed decreases as flow increases and vice versa.

For a more detailed discussion of the queueing models and their results, the interested reader is referred to Vandaele et al. [46], Van Woensel et al. [51], Van Woensel and Vandaele [52] and Van Woensel [50].

4.2 Mean Travel Time

To compute the travel time, one should note that in the time-dependent case, the travel speeds are no longer constant over the entire length of the arc. More specifically, one has to take into account the change of the travel speed when the vehicle crosses the boundary between two consecutive time periods. For example, the speed changes when going from time period p to time period $(p+1)$ from v_{ij}^{p} to $v_{ij}^{(p+1)}$.

In this paper, the time horizon is discretized into P time periods of equal length $\triangle p$ with a different travel speed associated with each time period p $(1 \leq p \leq P)$. The travel speeds are obtained using the above discussed queueing models for traffic flows. Formally, the travel time T_{ij} going from customer i to customer j, starting at some time p_0, is defined by the following expression:

$$\int_{p_0}^{p_0+T_{ij}^{p_0}} \left[v_{ij}^{p}\right] dp = d_{ij}$$

With v_{ij}^{p} denoting the speed in time period p on the arc (i, j) and d_{ij} the distance travelled. Solving this integral for T_{ij} and making use of the discrete time horizon, results in:

$$d_{ij} = \triangle p \left(\varphi v_{ij}^{p_0} + v_{ij}^{p_0+1} + ...v_{ij}^{p_0+(k-2)} + \phi v_{ij}^{last}\right)$$

Rewriting as a function of the time slices, gives:

$$T_{ij}^{p_0} = \varphi \triangle p_{first} + (k-2)\triangle p + \phi \triangle p_{last}$$

With $\triangle p_{first}$ the first time zone which contains p_0 and $\triangle p_{last}$ the last time zone used to cover the distance d_{ij} and k the total number of time slices needed (which is a function of the different speeds). The expected travel time is thus the sum of the following components:

1. The fraction of travel time still available in the first time zone, given by $\left(\varphi \triangle p_{first}\right)$, with $\triangle p_{first}$ the first time zone which contains p_0 and φ the fraction parameter $(0 \leq \varphi \leq 1)$.
2. The travel times of the $(k-2)$ intermediate time zones passed: $(k-2)\triangle p$.
3. The fraction of the travel time in the last time zone, given by $(\phi \triangle p_{last})$, with ϕ the fraction parameter $(0 \leq \phi \leq 1)$.

Using this procedure, the travel time $T_{ij}^{p_0}$ from customer i to customer j, starting at time p_0 can easily be determined based on the distance d_{ij} and the speed v_{ij}^{p} for the different time periods p obtained from the queueing models given a certain flow q in that time period.

4.3 Variance of the Travel Time

In this section, an expression for the variance of the travel time is derived. Again, using the queueing approach to traffic flow, the variance can be obtained in a closed form. For each time period p, the variance of the travel time can be determined as follows (using formulae 1 and 2):

$$\begin{aligned} Var\left(T_{ij}^{p}\right) &= Var\left(\frac{d_{ij}}{v_{ij}^{p}}\right) \\ &= d_{ij}^2 * Var\left(k_j * W^p\right) \\ &= d_{ij}^2 * k_j^2 * Var\left(W^p\right) \end{aligned}$$

The distance from i to j, d_{ij} and the jam density of the road k_j are assumed to be known. The variance of the total time in the system in period p, W^p is the sum of the waiting time and the service time. As there is no exact form for the variance of the waiting time, one needs to rely on approximations to obtain the variance of the waiting time. The waiting time can be obtained using the two moment approximations from Whitt [48]. These approximations have already proven their value and usability in production management ([47, 48], and others). Whitt [48] describes an approximation for the variance of the waiting time for the case when the expected waiting time is known (either exact or approximated). Only the necessary results for the analysis of the variance of the waiting time are presented here. For a detailed discussion, the reader is referred to Whitt [48]. The approximation has the following general form:

$$Var\left(W^p\right) = Var_{WaitingTime} + Var_{ServiceTime} \tag{4}$$

$$= \left[W_q\right]^2 c^2_{W_q} + \left[W_p\right]^2 c^2_s \tag{5}$$

with: $c^2_{W_q}$ the squared coefficient of variation of the waiting times defined in Whitt [48].

4.4 Added Value of the Queueing Approach

The assumption that everything in transportation goes according to a schedule is unrealistic resulting in a planning gap, i.e., the performance difference between the planned route and the actual route. The main consequence of considering congestion effects is that this route-planning gap will decrease as the planned route is now much more realistic compared to the classic VRP which is optimized in terms of distances. In order to realize this goal, one needs a good representation of the traffic process itself. The key issue is thus the characterization of the speed distribution and the resulting travel times on each link over different time slices.

Ichoua et al. [31] used a model based on a discrete speed distribution by adding correction factors to model the congestion but for a limited number of time slices. The proposed framework in the current paper has some similarities to this model proposed by Ichoua et al. [31]. In their paper, the travel time estimation is done using an arbitrarily step function. A recourse procedure is used as the travel time calculation takes into account travel speed changes when crossing the boundaries between two consecutive time periods. The major difference with the current paper lies in the estimation of the time dependent travel speed on a give link at a given time period when the departure is taken place: in the current paper a queueing model is used to estimate this travel speed. Ichoua et al. [31] recognize that travel speeds change continuously over time and therefore correctly claim that using step functions to compute travel speeds is a more reasonable assumption. Unfortunately, the authors do simulations with only three time slices, which is a too wide granularity to be a good approximation for a real-life observed speed-flow pattern (e.g., with two peaks during a normal weekday). Moreover, in the Ichoua et al. [31] paper, the speeds are modelled in an arbitrary way within each time slice: first, the arcs are partitioned in different subsets based on their physical characteristics (e.g., width, one/two ways, etc...), and their geographical location, then for each arc in the subset different weights are provided. Any information on how to link the weights of the arcs in the model with the crucial real-life physical characteristics is missing.

The major strength of using the queueing models is that, given the physical characteristics of the road network, it can immediately be mapped onto the parameters of the queueing model. The flow q is a parameter that is determined empirically over time, allowing the determination of realistic velocity profiles as a function of time. In the queueing formulas, one has four parameters that allows one to model any possible situation: the coefficient of variation of the inter-arrival times c_a, the coefficient of variation of the service times c_s, the jam density k_j and the free flow speed v_f. In practice, the jam density and the free flow speed are fixed for a given arc (i, j), leaving the coefficients of variation to represent the specific traffic conditions (e.g., bad weather, etc.). Analytical queueing models based on traffic counts thus model the behavior of traffic flow as a function of the most relevant determinants (e.g., free flow speeds, jam density, variability due to weather, etc.). An empirical validation of the queueing approach is provided in Van Woensel and Vandaele [52]; validation based on simulation results is provided in Van Woensel et al. [53]. Consequently, the travel times can be modelled much more realistically using these speeds (i.e., expressed in kilometer per hour) and are directly related to the physical characteristics and the geographical location on the arc.

There is also support in the literature in favor of the proposed combined routing-queueing approach: Bertsimas and Simchi-Levi [7] note that although queueing theory and vehicle routing are two well-studied disciplines individually, the effort to combine both has been very limited (see also [20] for a similar reasoning). The current paper is as such the first incursion into this combined routing–queueing research area. Bertsimas and Simchi-Levi [7] argue that analytical analysis of the vehicle routing problem offers new insights into the algorithmic structure and it makes performance analysis of classical algorithms possible. Moreover, it leads to a better understanding of models when integrating vehicle routing with other issues like inventory control. Furthermore, they point out that dealing with stochasticity in the VRP provides insights that can be useful when constructing practical algorithms for the VRP within a dynamic and stochastic environment.

5 Solution Approach

In this section, a solution strategy based on local search is proposed. According to Aarts and Lenstra [1], local search is a solution process that tries to improve a given initial solution by making relatively small changes in several steps in the solution space. The quality of the solutions is determined with the cost function of the problem. Local search techniques will result in a good but not necessarily optimal solution within reasonable computing time. In this section the local search heuristic used for obtaining solutions for the deterministic dynamic VRP is presented. More specifically, Ant Colony Optimization (ACO) is considered and implemented to compare and verify the results as an illustration. This heuristic is often used in the current literature to 'solve' combinatorial optimization problems. The success of the method is due to several factors: general applicability of the approach, flexibility for taking into account specific constraints in real cases and ease of implementation [42].

ACO is a stochastic optimization algorithm specifically intended to solve discrete optimization problems. The inspiration of the ACO algorithm comes from the observation of the trail laying and the trail following behavior of a real ant species (Linepithaeme humile). As the ants move in search for food, they deposit an aromatic essence called pheromone on the ground. The amount deposited generally depends upon the quality of food sources found. Other ants, observing the pheromone are more likely to follow the pheromone trail, with a bias towards stronger trails. As such, the pheromone trails reflect the memory of the ant population and over time, trails leading to good food sources will be reinforced while paths leading to remote sources will be abandoned [14]. The current implementation for ACO is based on the ones described in [10, 11, 18]. The major change made to this basic algorithm is to replace distance by dynamic travel time. To do so, the pheromone information is extended with an extra dimension representing the time zones. Consequently, the evaporation rate had to be adjusted to retain intermediate high-quality solutions long enough to direct ACO towards the most promising areas in the higher dimensional solution space.

The remainder of this section is restricted to the definition of the neighborhood around the current solution found and the definition of the objective function evaluated. Each solution is checked for 2-optimality and is improved if possible. A route is 2-optimal if it is not possible anymore to improve the route by exchanging two arcs. As opposed to the static deterministic VRP where the gain is calculated based on distances, in the static dynamic VRP, the gain is calculated in terms of travel time. As the evaluation is done in terms of travel times, the triangle inequality does not hold anymore. Therefore, the gain of the complete solution has to be re-evaluated. In addition to the 2-opt improvement, extra improvement heuristics are performed taking into account explicitly the time-dependent nature of the problem. First, all the different tours that make up a complete VRP solution, are checked to see if it would be advantageous in terms of travel time to break up the tour in two parts by adding the depot. This procedure is repeated until no more improvements can be realized or if the maximum number of trucks is exceeded. Secondly, all starting times of the different tours that make up a complete VRP solution, are shifted in time to evaluate the effect of the start time on the total travel time. In case of improvement, the starting time of the associated tour is updated. The rationale behind this optimization is that in a dynamic reality, a truck can decide to leave earlier or later to avoid periods of (anticipated) high congestion.

Define a solution as a set S with m routes $R_1, R_2, ..., R_m$ where $R_r = (v_0, v_{r_1}, v_{r_2}, ..., v_0)$ and each vertex v_i $(i \geq 1)$ belongs to exactly one route. The routes in S might be feasible with respect to their tour length restrictions or vehicle capacity constraints but do not have to be. For the ease of notation, write $v_i \in R_r$ if the vertex is part of the route R_r and write $(v_i, v_j) \in R_r$ if v_i and v_j are two consecutive vertices of R_r. Also define $T^p_{(v_i,v_j)}$ as the travel time needed to cover the distance between (v_i, v_j) leaving vertex v_i at time p. The basic objective function which needs to be minimized is similar to Gendreau et al. [22] but now expressed in terms of travel times:

$$F_1(S) = \sum_r \sum_{(v_i,v_j)\in R_r} T^p_{(v_i,v_j)} + \beta \sum_r \left[\sum_{v_i \in R_r} qd_{vi} - Q \right]^+ + \gamma \sum_r \left[\left(\sum_{(v_i,v_j)\in R_r} T^p_{(v_i,v_j)} + \sum_{v_i\in R_r} \delta_{vi} \right) - L \right]^+$$

where $[x]^+ = max(0, x)$ and γ, δ positive parameters. If the solution is feasible the second and third part of the equation are cancelled; on the other hand, if the solution is infeasible with respect to capacity (tour length) a penalty proportional to γ (δ,

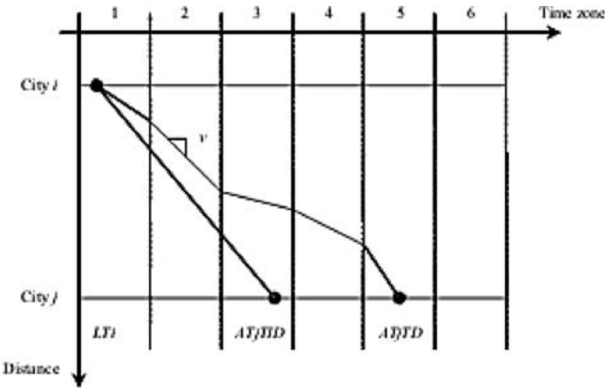


Figure 4 Example of the travel time functions.

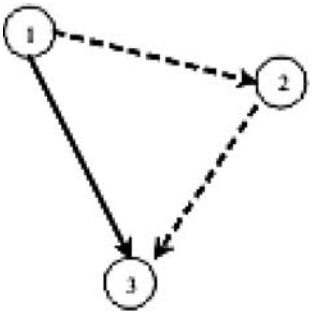
respectively,) is added. An illustration of the travel time results is shown in Figure 4. Figure 4 shows the different travel times over the link from city i to city j.

As one can see, the unbroken line gives the speed–space profile for the static deterministic case. The broken line gives the expected travel time assuming time buckets of 10 min. Leaving at time LT_i (leaving time city i), the vehicle can either arrive at time $AT_{jStatic}$ (arrival time at city j static speeds) or $AT_{jDynamic}$ (arrival time at city j dynamic speeds). The difference in both arrival times is the result of not taking into account congestion.

Most decision makers are risk-averse when taking decisions, as such only taking into account the mean ignores the risk attribute of the planners. It is argued that risk can be associated with the variance factor [40]. Note that the proposed approach is similar to mean-variance analysis used in financial planning of portfolios [8, 26]. The objective function $F_1(S)$ is thus extended by adding the variance of the travel times, leading to $F_2(S)$:

$$F_2(S) = F_1(S) + \alpha \sum_r \sum_{(v_i,v_j)\in R_r} Var\left(T^p_{(v_i,v_j)}\right)$$

where α, β, γ are positive parameters. Higher risk averseness will be reflected in an increase of the parameter α resulting in more weight for the variance in the objective function. An illustration of this objective function is presented in Figure 5. In this highly simplified example, the shortest link in terms of travel time is going directly from 1 to 3 (travel time is 10 compared to $6 + 7$ when going via node 2). If one takes into account the extended objective function $F_2(S)$ the actual routing decision will change depending upon the choice of α. If the weight α increases the longer route via node 2 will be preferred over the direct route as the longer route has less variability then the direct route (as reflected in the variance information). With objective function $F_2(S)$ together with the queueing approach (which allows for finding an analytical expression for $Var\left(T^p_{(v_i,v_j)}\right)$) the effect of incorporating travel time variability (as a measure for uncertainty) can be evaluated.



Node	(1,3)	(1,2)	(2,3)
T	10	6	7
Var(T)	15	3	4

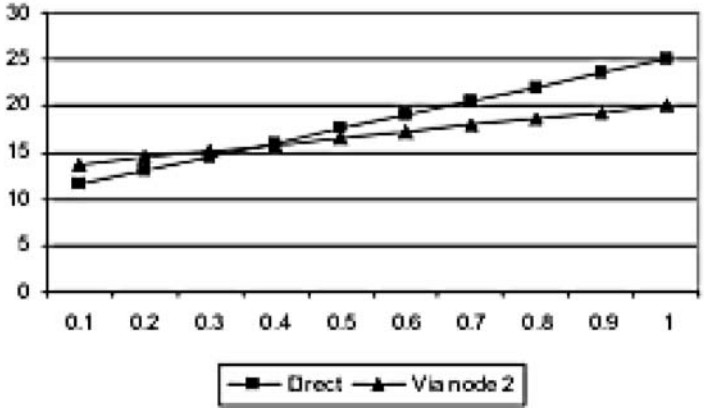


Figure 5 Illustration of objective function $F_2(S)$.

6 Computational Results

In this section, computational results will be presented. First, the problem instances are used to obtain the best solution for the time-independent routing models, i.e., minimizing total distance. Then, using the speeds from the queueing models, the obtained time-independent route can be recalculated in terms of time-dependent travel times. In a last step, the routing model which immediately takes into account time-dependent travel times is solved and compared with the latter. To validate this approach, explicit enumeration is used on small problem instances. As such, all possible solutions are obtained and evaluated both in terms of distance and in terms of travel times. As the explicit enumeration is limited to small problem instances only (the complexity is at least $n!$), a metaheuristic Ant Colony Optimization is introduced in the second section and applied on the benchmark problems described in Augerat et al. [4] for both objective functions considered.

6.1 Explicit Enumeration

As a first indication of the validity of the modelling approach, the analysis is done for small datasets (10 cities). This dataset is solved by explicit enumeration, i.e., all possible solutions are generated and evaluated both in distance and in travel times.

Table II The results of the explicit enumeration analysis

Problem	Distance	Time(*SolTID*)	Time (*SolTD*)	Difference
a	3, 121.81	1, 826.88	1, 795.41	−31.47(−1.72%)
b	2, 114.29	1, 184.37	1, 150.49	−33.88(−2.86%)
c	2, 399.08	1, 346.67	1, 332.91	−13.76(−1.02%)
d	2, 057.44	1, 170.95	1, 137.47	−33.48(−2.85%)
e	2, 122.06	1, 174.40	1, 166.87	−7.53(−0.06%)

The datasets are all random subsets of different problems from Augerat et al. [4]. All coordinates are multiplied with a constant factor of five to ensure that multiple time zones are covered over longer distances. Of course, this operation does not effect in any way the results and comparison presented here between the time-independent and the time-dependent routing models. In the time-dependent routing models, the starting time is allowed to be different (comparable to the real-life decisions of leaving earlier or later due to congestion). Trucks are in this case allowed to start between 7 A.M. and 9 A.M. It should be clear that these decisions only make sense for the time-dependent routing models as in the time-independent models no time dimension is involved in the analysis. For the explicit enumeration, the truck capacity is set to 45.

Table II shows the results for the different small datasets considered using objective function $F_1(S)$.[2] Focussing on problem *a*, reveals that the best solution in terms of distance (time-independent routing models), is 3,121.81 km. Using this solution $(0-1-7-0-2-3-6-0-9-8-4-0-5-10-0)$ and applying the speeds obtained from the queueing approach, results in a total travel time of 1,826.88 hr. Solving the same problem instances in terms time-dependent travel times reveals that the minimum travel time is 1,795.41 h. Moreover, the specific solution is also significantly different : $(0-1-4-8-0-6-9-10-5-0-7-2-3-0)$. For the problem set *a*, the total gain was 31.47 h or 1.72%.

Overall, for the five small datasets analyzed, the conclusion is similar. As expected, the results show that the incorporation of the congestion component in the routing models yields inferior results in terms of travel times compared to the time-independent model. Moreover, for the above five random subsets, the best solution in the time-dependent case differs from the best one in the time-independent case showing that, recalculating the time-independent solution by taking the speed information into account always yields worse results than solving the time-dependent case directly.

6.2 Heuristic Solutions

Unfortunately, the explicit enumeration approach only works for relatively small problems due to the computation times. Therefore if one wants to process realistic sizes of datasets, one needs to rely on approximation methods or heuristics. The VRP

[2] Objective function $F_2(S)$ could not be evaluated with explicit enumeration as the state space became to large even for these small problem sets due to the addition of the variance component in the objective function.

with time-dependent travel times was tested on the benchmark problems described in Augerat et al. [4]. These problems contain between 32 and 100 customers in addition to the depot. All coordinates are multiplied with a constant factor of five to ensure that multiple time zones are covered over longer distances. The length of the time period for this experiment is set equal to half an hour. Note that the choice of the time settings is purely arbitrarily, i.e., in the extreme case, time periods of 1 min can be considered. All capacities of the trucks are set to 100.

6.2.1 *Objective Function* $F_1(S)$

Table III shows the results for the time-independent case in terms of travel time, compared with the best, average and worst solution obtained for the time-dependent VRP over 10 runs. The starting time was shifted over a maximum of 10 time

Table III Comparing the deterministic time-independent VRP and the deterministic time-dependent VRP

		Time	Time dependent (% difference with time independent)					
	n	indep	Best	(%)	Average	(%)	Worst	(%)
$A1$	32	83.6741	77.2388	(−7.6)	77.3920	(−7.5)	77.5912	(−7.3)
$A2a$	33	69.5829	67.3420	(−3.2)	67.7153	(−2.7)	68.4364	(−1.6)
$A2b$	33	75.1773	73.2944	(−2.5)	73.7363	(−1.9)	74.4467	(−1.0)
$A3$	34	82.4392	79.2325	(−3.9)	80.5169	(−2.3)	81.4240	(−1.2)
$A4$	36	86.5221	79.9325	(−7.6)	81.0086	(−6.4)	81.8755	(−5.3)
$A5a$	37	69.3512	67.0384	(−3.3)	68.2538	(−1.6)	69.3507	(−0.0)
$A5b$	37	100.8685	91.0713	(−9.0)	94.9525	(−5.8)	97.2003	(−3.6)
$A6$	38	77.0037	73.4557	(−4.6)	74.7673	(−2.9)	76.7076	(−0.3)
$A7a$	39	85.9634	80.1364	(−6.8)	80.6482	(−6.4)	81.6714	(−5.0)
$A7b$	39	85.9416	80.5680	(−6.3)	82.0977	(−4.5)	84.2558	(−2.0)
$A8$	44	96.8239	88.7271	(−8.4)	90.7408	(−6.3)	92.8712	(−4.1)
$A9a$	45	105.1757	91.1591	(−13.3)	92.3050	(−12.3)	94.5917	(−10.1)
$A9b$	45	119.9593	111.1500	(−7.3)	113.9039	(−5.0)	116.2822	(−3.1)
$A11$	46	99.8330	87.3667	(−12.5)	88.5097	(−11.3)	90.3458	(−9.5)
$A12$	48	120.0614	103.6655	(−13.6)	106.7327	(−11.1)	109.7273	(−8.6)
$A13$	53	112.7714	96.6962	(−14.3)	101.4383	(−10.1)	106.3747	(−5.7)
$A14$	54	128.2866	115.2570	(−10.2)	116.4982	(−9.2)	118.7220	(−7.5)
$A15$	55	113.1931	103.0351	(−9.0)	105.6013	(−6.7)	108.1423	(−4.5)
$A16$	60	164.2826	136.6314	(−16.8)	139.2443	(−15.2)	141.6384	(−13.8)
$A17$	61	112.5518	96.9277	(−13.9)	101.2580	(−10.0)	105.6470	(−6.2)
$A18$	62	155.5044	126.1025	(−18.9)	132.0772	(−15.1)	135.9084	(−12.6)
$A19a$	63	176.9465	149.7263	(−15.4)	157.3567	(−11.1)	163.9178	(−7.4)
$A19b$	63	160.2347	128.5738	(−19.7)	132.6417	(−17.2)	136.9360	(−14.5)
$A20$	64	168.5061	138.8681	(−17.6)	142.7777	(−15.3)	149.7541	(−11.1)
$A21$	65	140.9648	117.8127	(−16.4)	120.6168	(−14.4)	123.3941	(−12.5)
$A22$	69	145.1921	121.7341	(−16.2)	126.6075	(−12.8)	130.9841	(−9.8)
$A23$	80	223.0433	179.3210	(−19.6)	186.0885	(−16.6)	195.0871	(−12.5)
$A24$	100	122.0448	92.5627	(−24.2)	96.6729	(−20.8)	99.7973	(−18.2)

n = number of customers
All numbers are expressed in hours of travel time

periods, which is equivalent with a decision: start a new tour in the morning or in the afternoon.

Table III gives the results of the analysis. In this table, all solutions are expressed in travel time. In other words, the obtained static solution (by minimizing distance) is recalculated in times. The average reduction in travel time comparing the recalculated static solution versus the dynamic solution over all sets is 22.2%. The results show that the routes improve significantly when one explicitly takes into account the dynamic character of the problem. Using the dynamic congestion information results in routes that are (considerably) shorter in terms of travel time. The larger the dataset, the larger the potential improvements become. Moreover, the spread between the best and the worst solution for the time-dependent VRP is (in most cases) shown to be small.

6.2.2 *Objective Function $F_2(S)$*

The second objective function $F_2(S)$ is evaluated for the Augerat datasets: the links are randomly selected as being highly variable versus low variable with equal weights, i.e., 50% of the roads have high coefficients of variations as input in the queueing formulas (c_a and c_s both equal to 0.9) and 50% of the roads have low coefficients of variations as input in the queueing formulas (c_a and c_s both equal to 0.1). The parameter α (the weight for the variance in the objective function) is varied between 0 and 0.1 with steps of 0.01. After extensive testing, it was found that $\alpha > 0.1$ did not result in any further gains. This is mainly due to the scale of the different factors in the objective function. Moreover, after an extra evaluation of the objective function only in the variance of the travel time (i.e., dropping the expected travel time) revealed that the variance was very close to this solution (i.e., minimum variance). Figure 6 gives the coefficient of variation (standard deviation of travel times divided by the mean travel times) of the solutions obtained for a selected number of datasets.

Based on the figure, it is clear that taking into account variance of the travel time reduces the coefficient of variation of the solution, i.e., the variability of a solution decreases. In general, the coefficient of variation decreased with on average with 54.30% with $\alpha \in [0, 0.1]$. The price for this reduction in variability is an increase in the average travel times: this increased on average with 27.87%. On the other hand, the variance of the travel times reduced with 66.67%. Moreover, when examining the

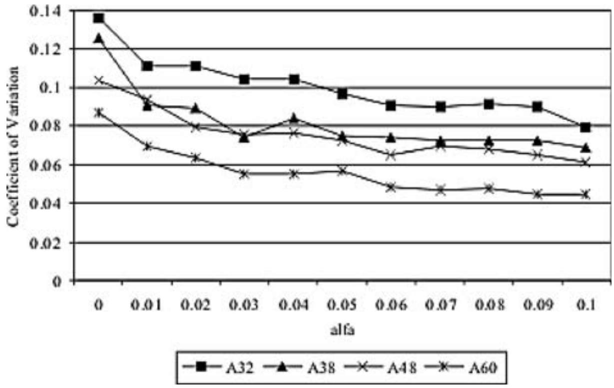


Figure 6 Evaluation of objective function $F_2(S)$ in terms of the coefficient of variation.

routes themselves, one can observe that the higher α, the more likely it becomes that links with less variability will be preferred at the expense of a higher average travel time, which is in line with the previous observation.

6.2.3 *Discussion of the Results*

The following observations can be made based on the above experiments:

1. Results show that the total travel times are improved when explicitly taking into account congestion. On average a reduction of 22.2% is achieved in travel times. This is in line with the findings reported in Ichoua et al. [31], meaning that the proposed advanced framework is at least as good as one of the best proposed methodology available in the literature. Unfortunately, a direct comparison with Ichoua et al. [31] proofed to be hard due to the different approach in the modelling of the travel time distributions. For this reason, setting up the experimental design such that the queueing approach could be adequately mapped with the settings of Ichoua et al. [31] turned out to be a difficult exercise which is left for future research.
2. The analysis revealed another added value of the queueing approach: it allowed for finding an analytical expression for the variance of the travel times. The results showed that the variability significantly decreased when taking into account the variance in the optimization, depending upon the weight given to the variance component in the objective function. The gain in terms of variance was 66.66% compared to a more limited increase of expected travel times of 27.87%. The approach of Ichoua et al. [31] does not make it possible to generate estimates of the variance of the travel times.
3. Not taking into account travel time variability (i.e., $\alpha = 0$) gives highly variable routes in all sets analyzed (see also Figure 6). However, giving a minimal weight (i.e., $\alpha = 0.1$) to the variance component in the objective function reduces the variance already with an average of 26.92% compared to an increase of only 5.19% in expected travel times. This suggests that a relatively small weight yields already significant gains in the planning process.
4. Computational results suggest that in systems where the constant speed approximation is no longer valid, it is crucially important to explicitly consider the variability due to traffic congestion in the model. The impact of dynamic components will be even more important when relating the approach to urban contexts (see also [45] for more insights on the impact of city logistics on routing decisions).

7 Future Research Opportunities

Based on the proposed framework for time-dependent routing problems, different future research opportunities exist: one could extend the existing models to cope with time windows, new and better heuristics could be developed, etc. In this paper, the Ant Colony Optimization approach is followed to find solutions for the described problem. Of course, any other solution methodology could be used (e.g., Tabu Search).

Using the queueing approach presented, one can also incorporate time windows in the time-dependent routing models in a realistic way. The issue is that a decision maker needs an indication to what degree the (hard or soft) time window will be met. In the case of hard time windows, this probability is either equal to one (if the hard time window at customer j is met) or equal to zero (if the hard time window at customer j is not met). This results always in a probability zero for customer j if the time window cannot be met. In the extreme case, this will result in as many routes as there are customers, i.e., each customer is incorporated in a single route. For soft time windows, such a probability can have all values between 0 and 1, with a value close to zero a strong probability that the time window will not be met and a value of one a strong probability that the time window will be met. Starting from the predetermined time window $[t_l^j, t_u^j]$ of customer j $(t_l^j < t_u^j)$, and defining AT_j as the expected arrival time of the vehicle at customer j, then the probability ψ_{ij} can be defined as:

$$\psi_{ij} = P\left(t_l^j \leq AT_j \leq t_u^j\right)$$

Of course, the expected arrival time of the vehicle at customer j is the sum of the time the vehicle left customer i (LT_i) and the travel time from i to j, given leaving time $(T_{ij}^{LT_i})$ or:

$$AT_j = LT_i + T_{ij}^{LT_i}$$

The travel time from i to j, given leaving time $(T_{ij}^{LT_i})$ will then depend upon the distribution assumed for the travel times.

In the current paper the fleet size is assumed to be unlimited. However, the benefits of a more realistic estimate of the true travel times will become even more important when the fleet size is limited (i.e., only a number of trucks are available). Indeed, many more opportunities and improvement possibilities exist for the limited fleet size problem when taking into account time. For example, a manager might take advantage of a travel time reduction of 25% by letting the same truck do two tours on a day (e.g., morning and afternoon).

8 Conclusions

This paper aimed at obtaining vehicle routing solutions that perform well in the face of the extra complications due to congestion, which eventually leads to a better solution in practice. These more realistic solutions have the potential to reduce real operating costs for a broad range of industries which daily face routing problems. Recently, the problem considered has received increasing attention due to its relevance to real-life problems. A framework for routing problems with time-dependent travel times due to potential traffic congestion was presented. The approach developed introduces the traffic congestion component in the standard VRP models. The traffic congestion component was modelled using a queueing approach to traffic flows. By making use of this analytical approach to traffic flows, the limited necessary data to model congestion is easily obtained which opens the door for real-life applications. Both the time-independent as the time-dependent VRP were solved using Ant Colony Optimization. Results showed that the total travel times can be improved significantly when explicitly taking into account congestion during the optimization. Moreover, the framework allows for finding an expression for the variance of the travel time which can be easily integrated in the objective function. Results showed that depending upon the weight chosen for the variance component in the objective function, the obtained results (in terms of chosen routes, mean travel times and variance of the travel times) can be significantly different. The capability of analytically taking into account dynamic travel times is extremely valuable, not only because speeds profiles do affect the objective function of the optimization, but also as demonstrated the best solutions for the static problem applied in a dynamic context, are in general suboptimal. Compared to the actual realization in real-life, the planning gap (difference between plan and actual) can be reduced substantially when taking into account a manifestation of the traffic on the roads. Consequently, due to the resulting reduced planning gap, less time and effort needs to be invested in replanning (in real-time) during the day.

Acknowledgements The authors would like to thank the anonymous referees and the editor for their valuable input during the process of (re-)writing the paper. Also many thanks to Christophe Lecluyse (University of Antwerp) for helping with implementing the heuristics.

References

1. Aarts, E.H.L., Lenstra, J.K.: Local Search in Combinatorial Optimization. Wiley, New York (1997)
2. Akçelik, R.: Travel time functions for transport planning purposes: Davidson's function, its time-dependent form and an alternative travel time function. Aust. Road Res. **21**(3), 49–59 (1991)
3. Akçelik, R.: Relating flow, density, speed and travel time models for uninterrupted and interrupted traffic. Traffic Eng. Control **37**(9), 511–516 (1996)
4. Augerat, P., Belenguer, J.M., Benavent, E., Corber, A., Naddef, D.: Separating capacity constraints in the CVRP using tabu search. Eur. J. Oper. Res. **106**, 546–557 (1998)
5. Bertsimas, D., Van Ryzin, G.: A stochastic and dynamic vehicle routing problem in the Euclidian plane. Oper. Res. **39**, 601–615 (1991)
6. Bertsimas, D., Van Ryzin, G.: Stochastic and dynamic vehicle routing problems in the Euclidean plane with multiple capcitated vehicles. Oper. Res., **41**, 60–76 (1993)
7. Bertsimas, D.J., Simchi-Levi, D.: A new generation of vehicle routing research: robust algorithms, addressing uncertainty. Oper. Res., **44**(2), 286–304, (1996)
8. Best, M.J., Grauer, R.R.: Sensitivity analysis for mean-variance portfolio problems. Manage. Sci. **37**(8), 980–989 (1991)
9. Brown, G.G., Ellis, C.J., Lenn, G., Graves, W., Ronen, D.: Real time, wide area dispatch of mobile tank trucks. Interfaces **17**, 107–120 (1987)
10. Bullnheimer, B., Hartl, R.F., Strauss, Ch.: Applying the ant system to the vehicle routing problem. In: Voss, S., Martello, I., Osman, H., Roucairol, C. (eds.) Meta-Heuristics: Advances and Trends in Local Search Paradigms for Optimization. Kluwer, Boston (1999)
11. Bullnheimer, B., Hartl, R.F., Strauss, Ch.: An improved ants system algorithm for the vehicle routing problem. Ann. Oper. Res. **89**, 319–328 (1999)
12. Christofides, N., Mingozzi, A.: Vehicle Routing: Practical and Algorithmic Aspects. Pergamon Press, Oxford, UK (1989)
13. Clarke, G., Wright, J.W.: Scheduling of vehicles from a central depot to a number of delivery points. Oper. Res. **12**(4), 568–581 (1964)
14. Corne, D., Dorigo, M., Glover, F.: New Ideas in Optimization. Advanced Topics in Computer Science Series. McGraw-Hill, London, UK (1999)
15. Coyle, J.J., Bardi, E.J., Langley Jr. J.J.: The Management of Business Logistics. West, St. Paul, Minnesota (1996)

16. Daganzo, C.F.: Fundamentals of Transportation and Traffic Operations. Elsevier, New York (1997)
17. Davidson, K.B.: The theoretical basis of a flow-travel time relationship for use in transportation planning. Aust. Road Res. **8**(1), 32–35 (1978)
18. Dorigo, M., Stützle, T.: The ant colony optimization metaheuristic: algorithms, applications and advances. In: Handbook of Metaheuristics, International Series in Operations Research and Management Science, vol. 57, pp. 251–285. Kluwer, Massachusetts (2002)
19. Fleischmann, B., Gietz, M., Gnutzmann, S.: Time-varying travel times in vehicle routing. Transp. Sci. **38**(2), 160–173 (2004)
20. Gans, N., Van Ryzin, G.: Dynamic vehicle dispatching: optimal heavy traffic and practical insights. Oper. Res. **47**(5), 675–692 (1999)
21. Gao, S., Chabini, I.: The best routing policy problem in stochastic time-dependent networks. Transp. Res. Rec. **1783**, 188–196 (2002)
22. Gendreau, M., Hertz, A., Laporte, G.: A tabu search heuristic for the vehicle routing problem. Manage. Sci. **40**(10), 1276–1290 (1994)
23. Gendreau, M., Laporte, G., Seguin, R.: Stochastic vehicle routing. Eur. J. Oper. Res. **88**(1), 3–12 (1996)
24. Gendreau, M., Laporte, G., Séguin, R.: A tabu search heuristic for the vehicle routing problem with stochastic demands and customers. Oper. Res. **44**(3), 469–477 (1996)
25. Gillett, B.E., Miller, L.R.: A heuristic algorithm for the vehicle-dispatch problem. Oper. Res. **22**(2), 340–349 (1974)
26. Grauer, R.R., Hakansson, N.H.: On the use of mean-variance and quadratic approximations in implementing dynamic investment strategies: a comparison of returns and investment policies. Manage. Sci. **39**(7), 856–871 (1993)
27. Greenshields, B.D.: A study of traffic capacity. In: Highway Research Board Proceedings, Washington, District of Columbia, vol. 14, pp. 448–477, 1935
28. Heidemann, D.: A queueing theory approach to speed-flow-density relationships. In: Transporation and Traffic Theory. Proceedings of the 13th International Symposium on Transportation and Traffic Theory, Lyon, France, 1996
29. Hickman, M.D., Wilson, N.H.M.: Passenger travel time and path choice implications of real-time transit information. Trans. Res. Circ. **3**(4), 211–226 (1995)
30. Hill, A.V., Benton, W.C.: Modeling intra-city time-dependent travel speeds for vehicle scheduling problems. Eur. J. Oper. Res. **43**(4), 343–351 (1992)
31. Ichoua, S., Gendreau, M., Potvin, J-Y.: Vehicle dispatching with time-dependent travel times. Eur. J. Oper. Res. **144**, 379–396 (2003)
32. Jain, R., MacGregor Smith, J.: Modeling vehicular traffic flow using M/G/C/C state dependent queueing models. Transp. Sci. **31**, 324–336 (1997)
33. Kenyon, A.S., Morton, D.P.: Stochastic vehicle routing with random travel times. Transp. Sci. **37**(1), 69–82 (2003)
34. Kerbache, L., Van Woensel, T.: Planning and Scheduling Transportation Vehicle Fleet in a Congested Traffic Environment. Copenhagen Business School, Denmark (2005)
35. Kingman, J.F.C.: The single server queue in heavy traffic. Proc. Camb. Philos. Soc. **57**, 902–904 (1964)
36. Kraemer, W., Lagenbach-Belz, M.: Approximate formulae for the delay in the queueing system GI/GI/1. In: Congressbook of the Eight International Teletraffic Congress, Melbourne, Australia, pp. 235–1/8, 1976
37. Laporte, G.: The vehicle routing problem: an overview of exact and approximate algorithms. Eur. J. Oper. Res. **59**(3), 345–358 (1992)
38. Malandraki, C., Daskin, M.S.: Time dependent vehicle routing problems: formulations, properties and heuristic algorithms. Transp. Sci. **26**(3), 185–200 (1992)
39. Malandraki, C., Dial, R.B.: A restricted dynamic programming heuristic algorithm for the time dependent traveling salesman problem. Eur. J. Oper. Res. **90**, 45–55 (1996)
40. Mulvey, J.M., Vanderbei, R.J., Zenios, S.A.: Robust optimization of large-scale systems. Oper. Res. **43**(2), 264–281 (1995)
41. Osman, I.: Vehicle routing and scheduling: applications, algorithms and developments. In: Proceedings of the International Conference on Industrial Logistics, Rennes, France, 1993
42. Pirlot, M.: General local search methods. Eur. J. Oper. Res. **92**, 493–511 (1996)
43. Shen, Y., Potvin, J.-Y.: A computer assistant for vehicle dispatching with learning capabilities. Ann. Oper. Res. **61**, 189–211 (1995)

44. Taillard, É., Badeau, P., Gendreau, M., Guertin, F., Potvin, J.Y.: A tabu search heuristic for the vehicle routing problem with soft time windows. Transp. Sci. **31**, 170–186 (1995)
45. Taniguchi, E., Thompson, R.G., Yamada, T., Van Duin, R.: City Logistics: Network Modelling and Intelligent Transport Systems. Pergamon, New York (2001)
46. Vandaele, N., Van Woensel, T., Verbruggen, A.: A queueing based traffic flow model. Trans. Res. D. **5**(2), 121–135 (2000)
47. Vandaele, N.J.: The impact of lot sizing on queueing delays: multi product, multi machine models. PhD thesis, Katholieke Universiteit Leuven, Department of Applied Economics (1996)
48. Whitt, W.: The queueing network analyzer. Bell Syst. Tech. J. **62**(9), 2779–2815 (1983)
49. Whitt, W.: Approximations for the GI/G/m queue. Prod. Oper. Manag. **2**(2), 114–161 (1993)
50. Van Woensel, T.: Modelling uninterrupted traffic flows, a queueing approach. PhD Dissertation, University of Antwerp, Belgium (2003)
51. Van Woensel, T., Creten, R., Vandaele, N.: Managing the environmental externalities of traffic logistics: the issue of emissions. In: POMS journal, Special Issue on Environmental Management and Operations, vol. 10 (2001)
52. Van Woensel, T., Vandaele, N.: Empirical validation of a queueing approach to uninterrupted traffic flows. J. Oper. Res. **4**(1), 59–72 (2006)
53. Van Woensel, T., Wuyts, B., Vandaele, N.: Validating state-dependent queueing models for uninterrupted traffic flows using simulation. J. Oper. Res. **4**(2), 159–174 (2006)