Travel times in quasi-dynamic traffic assignment

Jeroen P.T. van der Gun*, Adam J. Pel, Bart van Arem

Delft University of Technology. P.O. Box 5048, 2600 GA Delft, The Netherlands.

*Corresponding author: j.p.t.vandergun@tudelft.nl

Abstract

By extending static traffic assignment with explicit capacity constraints, quasi-dynamic traffic assignment yields more realistic results while avoiding many disadvantages of dynamic assignment. We analyse the computation of travel times in quasi-dynamic assignment models. We formulate requirements for the correctness of resulting travel times, addressing both the calculation of travel times for individual routes and links itself, as well as the differences between travel times of different travel choices. We derive a new link travel time formula from vertical queuing theory that meets all requirements, unlike existing approaches in literature.

Keywords: quasi-dynamic traffic assignment, link travel time, vertical queuing theory, capacity constraint, link performance function.

1 Introduction

There is a long tradition of static traffic assignment (STA) in transportation research. Albeit dynamic traffic assignment (DTA) can yield more detailed and realistic results, it unfortunately also comes with greater model complexity, higher computational costs, higher data requirements, and poorer convergence. And hence, STA is still much used in practice and research. The most important limitation of traditional STA models is the lack of explicit capacity constraints, which results in errors in the modelled congestion patterns around bottlenecks, and consequently also errors in travel times and delays. To remedy this, Bakker et al. (1994), Bifulco and Crisalli (1998), Lam and Zhang (2000), Bundschuh et al. (2006), Gentile et al. (2014), and Brederode et al. (2018) formulated capacity-constrained STA models where excess traffic at bottlenecks can accumulate in residual queues. These queues are initially empty and absorb whatever traffic demand from the studied time period that exceeds capacity. Hence, while in traditional STA based on link performance functions (Bureau of Public Roads, 1964) traffic is omnipresent, instead in capacity-constrained STA traffic is instantaneously propagated over its entire route, but the fraction of a route's demand ending up in a residual queue is not propagated over the remainder of that route.

Bliemer et al. (2014) refer to this as quasi-dynamic traffic assignment (QDTA), and formulate the capacity-constrained traffic propagation and assignment as two fixed-point problems, including calculation of route travel times. Their formulation supports use of generic first-order node models (Tampère et al., 2011), and correctly places vertical queues in front of bottleneck nodes.

Within the context of quasi-dynamic traffic assignment, this paper focuses specifically on the computation of travel times, while we discuss the propagation of traffic through the network to the extent relevant. To this end, we formulate a list of requirements shown in Table 1 that can be divided into two categories:

- requirements for *absolute correctness*, that ensure a valid composition of the travel time calculation for individual routes and links (I-III);
- requirements for *relative correctness*, that ensure a valid comparison between travel times of alternative routes or alternative demand patterns (IV-VI).

The six requirements are listed here for reference and will be introduced and discussed throughout this paper. All requirements in Table 1 are realistic and satisfied by DTA, while QDTA formulations until now violate several requirements. This is solved in this paper where we present a new procedure to correctly compute travel times.

		STA	QDTA		DTA	
			BPR (1964)	Bliemer et al. (2014)	This paper	
Absolute correctness	I. Corridor Compatibility	(§2.2)	×	√	\checkmark	
	II. Route is Sum of Links	(§2.3)	√	×	\checkmark	\checkmark
	III. Steady State Consistency	(§2.3)	√	×	\checkmark	\checkmark
Relative correctness	IV. Correct Derivatives	(§3.1)	×	×	 Image: A second s	\checkmark
	V. First-In-First-Out	(§3.2)	√	×	 Image: A second s	\checkmark
	VI. Stops are Not Faster	(§3.2)	√	X	 Image: A second s	 Image: A second s

Table 1. Satisfaction of requirements for travel time calculation by various traffic assignment methods.

Existing QDTA travel time computation procedures fail to meet requirements for both absolute and relative correctness. For brevity, this paper will only show this for Bliemer et al. (2014), but the other QDTA models mentioned earlier also fail to meet multiple requirements. Our main contribution is a new QDTA travel time formula derived from queuing theory that does satisfy all requirements and can readily replace the Bliemer et al. (2014) formula. We select the Bliemer et al. (2014) formulation for our comparison because it both clearly illustrates all our requirements and its derivation is similar to our proposed formulation.

The structure of this paper is as follows. First, Section 2 revisits the derivation of quasidynamic travel time formulas from queuing theory, resulting in our new QDTA travel time formula. Requirements for absolute correctness are formulated and investigated in the process. Then, Section 3 formulates and investigates requirements for relative correctness, by considering derivatives of the link travel time and analysing how route travel times affect choice situations. Finally, Section 4 summarises our conclusions and discusses advantages of our approach.

2 Theoretical derivation of quasi-dynamic travel times

In this section, we will derive a formula for the travel time τ_a of link *a*. Like Bliemer et al. (2014, p. 371), we split the travel time into a free-flow travel time τ_a^{ff} and a queuing delay τ_a^{queue} :

$$\tau_a = \tau_a^{\rm ff} + \tau_a^{\rm queue} \,. \tag{1}$$

Here, τ_a^{ff} is either a constant or an increasing function of link inflow q_a . Nevertheless, our focus is on deriving a formula for τ_a^{queue} , that indicates the extra travel time due to the link's exit capacity being exceeded and the consequent queue being formed on the link. The formula for this queuing delay needs be consistent with the arrival and service rates of traffic, also in networks with more than one bottleneck.

Note that as we are dealing with QDTA, the (link and route) travel times in our computations are static, and hence there is no time index.

This section derives this link travel time formula for increasingly complex networks: a single link in Section 2.1, a diverges-only network in Section 2.2, and a general network in Section 2.3.

2.1 Single link

We first consider a network consisting of a single link *a*, subject to a constant traffic demand rate f_a departing during time period [0,T]. Since there are no upstream queues, the link inflow q_a is equal to the demand rate f_a . A queue may however build up at the exit of the link, consisting of traffic waiting to traverse the next node, due to insufficient capacity as determined by the node model. Let $\alpha_a \in (0,1]$ be the link flow reduction factor, such that the link outflow is $\alpha_a q_a$. Initially at time 0, there is no queue on the link. Because traffic propagation is instantaneous, a queue forms 'immediately'. At the end of the study time period, the number of vehicles in the queue is (Bliemer et al., 2014, p. 371)

$$Q_a = q_a T - \alpha_a q_a T = (1 - \alpha_a) q_a T .$$
⁽²⁾

Assuming the link outflow remains $\alpha_a q_a$ after *T* until all vehicles left, the last vehicle to enter the queue has to wait

$$\frac{Q_a}{\alpha_a q_a} = \frac{1 - \alpha_a}{\alpha_a} T = \left(\frac{1}{\alpha_a} - 1\right) T$$
(3)

time for the queue to dissolve before it can exit the link. Since the queue grew linearly from 0 to Q_a , the average delay for all $f_aT = q_aT$ vehicles equals half of this (Bliemer et al., 2014, p. 371; Gentile et al., 2014, p. 319):

$$\tau_a^{\text{queue}} = \left(\frac{1}{\alpha_a} - 1\right) \frac{T}{2} \,. \tag{4}$$

The build-up of the link's queue over time is depicted in Fig. 1 using cumulative curves. Because of the first-in-first-out property, the horizontal distance between the entrance and exit curves represents the delay of the N th vehicle. This shows visually how the above formulas are derived.



Fig. 1. Cumulative numbers of vehicles entering and exiting the queue over time, on a single link network.

2.2 Diverges-only network

Now consider a link *a* that is part of a corridor network consisting of multiple links, where all nodes are either one-to-one nodes or diverges. The link inflow q_a during time period [0,T] may now be constrained by queues in the set of links η_a upstream of link *a* (excluding itself, $a \notin \eta_a$). More precisely (Bifulco and Crisalli, 1998, p. 88; Bliemer et al., 2014, p. 369):

$$q_a = f_a \prod_{a' \in \eta_a} \alpha_{a'} \,. \tag{5}$$

Thus traffic will keep flowing into the queue at the exit of link *a* after time period [0,T] ended. We again assume the link outflow remains $\alpha_a q_a$ until all vehicles left. This implies inflows into following links will also remain constant. Because both the inflow and outflow of link *a* remain constant after [0,T] until the last vehicle enters/exits the link, therefore the last vehicle enters the queue on link *a* at time \overline{T}_a such that the full link demand is processed, i.e.

$$q_a \overline{T}_a = f_a T , \qquad (6)$$

so

$$\overline{T}_a = \frac{f_a}{q_a} T = \frac{1}{\prod_{a' \in \eta_a} \alpha_{a'}} T .$$
⁽⁷⁾

At time \bar{T}_a , the number of vehicles in the queue equals

$$\overline{Q}_a = q_a \overline{T}_a - \alpha_a q_a \overline{T}_a = (1 - \alpha_a) q_a \overline{T}_a = \frac{1 - \alpha_a}{\prod_{a' \in \eta_a} \alpha_{a'}} q_a T \,. \tag{8}$$

The last vehicle thus experiences a delay of

$$\frac{\overline{Q}_a}{\alpha_a q_a} = \frac{1 - \alpha_a}{\alpha_a \prod_{a' \in \eta_a} \alpha_{a'}} T = \frac{1}{\prod_{a' \in \eta_a} \alpha_{a'}} \left(\frac{1}{\alpha_a} - 1\right) T .$$
⁽⁹⁾

The average delay on the link for all $f_a T$ vehicles therefore equals half of this (Bliemer et al., 2014, p. 372):

$$\tau_a^{\text{queue}} = \frac{1}{\prod_{a'\in\eta_a} \alpha_{a'}} \left(\frac{1}{\alpha_a} - 1\right) \frac{T}{2} \,. \tag{10}$$

Fig. 2 depicts the queue build-up in terms of cumulative numbers of vehicles over time.



Fig. 2. Cumulative numbers of vehicles entering and exiting a link's queue over time, in a network with multiple links.

The total average delay of link *a* and all links before it equals (Bliemer et al., 2014, p. 372)

$$\sum_{a'\in\eta_a\cup\{a\}}\tau_{a'}^{\text{queue}} = \sum_{a'\in\eta_a\cup\{a\}}\frac{1}{\prod_{a'\in\eta_a'}\alpha_{a'}}\left(\frac{1}{\alpha_{a'}}-1\right)\frac{T}{2} = \sum_{a'\in\eta_a\cup\{a\}}\left(\frac{1}{\prod_{a''\in\eta_a\cup\{a'\}}\alpha_{a''}}-\frac{1}{\prod_{a''\in\eta_a'\cup\{a'\}}\alpha_{a''}}\right)\frac{T}{2} = \left(\frac{1}{\prod_{a'\in\eta_a\cup\{a\}}\alpha_{a'}}-1\right)\frac{T}{2}.$$
 (11)

As discovered by Bliemer et al. (2014, pp. 371-372), this is consistent with vertical queuing theory on a corridor: the queuing delay of a series of consecutive bottlenecks is equal to the queuing delay of a single bottleneck with severities multiplied. It thereby satisfies our first requirement in Table 1:

Requirement I. (Corridor Compatibility):

Travel times on a corridor network are compatible with queuing theory calculations with the corridor inflow as arrival rate and the final link outflow as service rate. Depending on the capabilities of the traffic propagation model, either vertical queuing or horizontal queuing may be used.

2.3 General network

We now proceed to calculating travel times in general networks. In a network with merging traffic, the traffic flow rate and the composition of traffic after a merge may vary. Hence the calculation of delays in general networks is a bit more involved. While Bliemer et al. (2014, p. 372) immediately infer the corridor result to also hold for general networks, we follow a more rigorous derivation.

We can convert a general network into a diverge-only network by replacing links after merges with multiple parallel links, and assigning flows to the parallel links in such a way that traffic flows never merge. Let P_a be the set of parallel links that replace link *a*:

$$\sum_{p \in P_a} f_p = f_a$$

$$\sum_{p \in P_a} q_p = q_a$$
(12)

Each parallel link $p \in P_a$ now has a well-defined set of predecessor links η_p . Using the same reasoning as for the diverge-only network, we can compute the average delay on any link $p \in P_a$:

$$\forall p \in P_a: \tau_p^{\text{queue}} = \frac{1}{\prod_{p' \in \eta_p} \alpha_{p'}} \left(\frac{1}{\alpha_p} - 1\right) \frac{T}{2} = \frac{f_p}{q_p} \left(\frac{1}{\alpha_p} - 1\right) \frac{T}{2} \,. \tag{13}$$

To preserve the composition of inflow into the node downstream of any original link a', including conservation of turning fractions (Daganzo, 1995, p. 88; Tampère et al., 2011, p. 295), we require

$$\forall p' \in P_{a'} : \alpha_{p'} = \alpha_{a'}, \tag{14}$$

resulting in

$$\forall p \in P_a : \tau_p^{\text{queue}} = \frac{1}{\prod_{p' \in \eta_n} \alpha_{p'}} \left(\frac{1}{\alpha_a} - 1\right) \frac{T}{2} = \frac{f_p}{q_p} \left(\frac{1}{\alpha_a} - 1\right) \frac{T}{2}.$$
(15)

Eq. (15) matches the final result reported by Bliemer et al. (2014, p. 372). Even though the links $p \in P_a$ now have equal outflow-to-inflow ratios α_a , they still have separate queues with potentially different delays. Original link *a* does not yet have a unique travel time: instead, the travel time one experiences still varies depending on what predecessor links they came from, i.e. traffic following different routes experience different travel times on the same physical link. Because this difference cannot be attributed to differences in departure time or total prior travel time, the following requirement in Table 1 is not met:

Requirement II. (Route is Sum of Links):

The travel time of a route equals the sum of travel times of the links along the route. The travel time traffic experiences on a link may vary with the time the traffic enters the link, but not with where the traffic originates from.

In order to satisfy Requirement II, we continue our derivation by merging the parallel links $p \in P_a$ back into a single link *a* with a single travel time τ_a . Naïve summation of the timedependent inflows and outflows of links $p \in P_a$ may lead to non-constant inflow and outflow for link *a*, because according to Eq. (7), the summed inflows q_p can have different durations \overline{T}_p and the summed outflows $\alpha_a q_p$ can have different durations $\overline{T}_p / \alpha_a$. This is also problematic, because it violates the stationarity of the traffic composition implied by the instantaneous propagation of flows. Because the instantaneous flow propagation does not take travel times into account, it is impossible to conclude from the flow profiles that traffic from one origin arrives on average earlier at some point than the other traffic. Within quasidynamic modelling, the only way to avoid implying such conclusions is to assume a constant inflow for link *a* in which the separate contributions of different origins can no longer be identified.¹ This leads to the next requirement:

Requirement III. (Steady State Consistency):

Link travel times are based on both instantaneous traffic propagation and homogeneous traffic composition, or on neither. Put differently, either traffic propagation and demand both are time-dependent, or both are in steady state.

Below we continue by seeking to construct a homogenous constant inflow profile for link *a* that is consistent with the inflow profiles of links $p \in P_a$. The inflow rate q_a is already given by Eq. (12). To preserve the total inflow, the duration \overline{T}_a is then the average of durations \overline{T}_p weighted with inflow rates q_p . Substituting Eq. (7), this weighted average simplifies to

$$\bar{T}_{a} = \sum_{p \in P_{a}} \frac{q_{p}}{q_{a}} \bar{T}_{p} = \sum_{p \in P_{a}} \frac{q_{p}}{q_{a}} \frac{f_{p}}{q_{p}} T = \sum_{p \in P_{a}} \frac{f_{p}}{q_{a}} T = \frac{f_{a}}{q_{a}} T .$$
(16)

¹ An alternative solution is to instead get rid of instantaneous flow propagation, but then the quasi-dynamic model turns into a traditional dynamic model (with rectangular demand profiles and without spillback).

Now that we made the link inflow constant, the link outflow is also constant. The number of vehicles in the combined queue on link *a* at time \overline{T}_a therefore equals

$$\overline{Q}_a = q_a \overline{T}_a - \alpha_a q_a \overline{T}_a = (1 - \alpha_a) q_a \overline{T}_a = (1 - \alpha_a) q_a \frac{f_a}{q_a} T = (1 - \alpha_a) f_a T.$$
(17)

The delay of the last vehicle is therefore

$$\frac{\overline{Q}_a}{\alpha_a q_a} = \frac{(1 - \alpha_a) f_a T}{\alpha_a q_a} = \frac{f_a}{q_a} \left(\frac{1}{\alpha_a} - 1\right) T$$
(18)

such that the average delay is

$$\tau_a^{\text{queue}} = \frac{f_a}{q_a} \left(\frac{1}{\alpha_a} - 1 \right) \frac{T}{2} \,. \tag{19}$$

This result is fully consistent with the previous result for a diverges-only network. The evolution of the cumulative numbers of vehicles entering and exiting the combined queue is the same as Fig. 2.

3 Relative correctness of quasi-dynamic travel times

Now that we have derived a travel time formula for links in a general network, we proceed to assess the relative correctness of travel times in QDTA, i.e. the ability to compare route travel times. This section is split into two parts. First, we look at the derivatives of the link travel time formula while keeping the link exit capacity fixed. Second, we check the behavioural incentives for travel choices.

3.1 Link with fixed exit capacity

We now study a link with a fixed exit capacity. Because of the invariance principle (Lebacque and Khoshyaran, 2005, pp. 370-371), this occurs when the turning fractions and competing flows at its downstream node remain constant, even in advanced first-order node models (Tampère et al., 2011, p. 295). Assuming link *a* has fixed exit capacity C_a , we have (Bifulco and Crisalli, 1998, p. 88)

$$\alpha_a = \min\left(1, \frac{C_a}{q_a}\right). \tag{20}$$

Combining this with Eqs. (1) and (19) yields the following link travel time:

$$\tau_a = \begin{cases} \tau_a^{\rm ff} & \text{if } q_a \le C_a \\ \tau_a^{\rm ff} + \left(\frac{f_a}{C_a} - \frac{f_a}{q_a}\right) \frac{T}{2} & \text{if } q_a \ge C_a \end{cases}.$$
(21)

This structure of this formulation is comparable to link travel time functions in static assignment, for example the well-known Bureau of Public Roads (1964, p. V20) function:

$$\tau_a = \tau_a^{\rm ff} \left(1 + 0.15 \left(\frac{f_a}{C_a} \right)^4 \right). \tag{22}$$

The comparison reveals the fundamental difference that in (non-capacitated) static assignment the travel time τ_a only depends on the total demand f_a for the link, and not also on the actual flow q_a that is able to get into the link during the studied time period. (Static models don't compute q_a .) Both static models and our quasi-dynamic model satisfy

$$\begin{cases} \forall q_a \in (0, f_a]: \frac{\partial \tau_a}{\partial f_a} \ge 0\\ \forall q_a \in (C_a, f_a]: \frac{\partial \tau_a}{\partial f_a} > 0 \end{cases}, \tag{23}$$

but our quasi-dynamic model additionally satisfies

$$\begin{cases} \forall q_a \in (0, f_a] : \frac{\partial \tau_a}{\partial q_a} \ge 0 \\ \forall q_a \in (C_a, f_a] : \frac{\partial \tau_a}{\partial q_a} > 0 \end{cases},$$
(24)

whereas static models always have $\partial \tau_a / \partial q_a = 0$. Note that this also holds if τ_a^{ff} in Eq. (21) is not constant but itself increases with q_a , as suggested at the beginning of Section 2. In conclusion, we have the following requirement:

Requirement IV. (Correct Derivatives):

In case of a link with a fixed exit capacity, the partial derivatives of the link's travel time with respect to both the link's demand and the link's flow have the correct sign, in accordance with Eqs. (23) and (24).

If one were to derive a representative link travel time directly from Eq. (15) (Bliemer et al., 2014, p. 372), it would be based on the total travel time of all traffic demand of all routes on the link, divided by the total traffic demand. This corresponds to the previously mentioned idea of summing the cumulative inflow and outflow curves. One can prove that this results in a possibility of $\partial \tau_a / \partial f_a < 0$, violating Eq. (23). For brevity, we omit this proof here.

3.2 Behaviour in choice situations

We will now study how differences between route travel times affect choice situations, which we use to formulate two more requirements for relative correctness. Note that satisfaction of Requirement II (Route is Sum of Links) automatically leads to satisfaction of the new requirements.

In Eq. (15) (Bliemer et al., 2014, p. 372), the travel time one experiences on a link depends on the severity of congestion earlier in your route. However, the severity of congestion may not coincide with the time spent travelling earlier in your route. Consequently, you may be able to speed up your total trip by replacing part of your route with a slower detour. This results in Requirement V:

Requirement V. (First-In-First-Out):

Route travel times respect the first-in-first-out property of links: it is not possible to leave a link earlier by entering it later. In non-dynamic contexts, the words 'earlier' and 'later' are to be interpreted in terms of the travel time spent since the beginning of the route.

Because the travel time of a route is not the sum of travel times of links in the route, an additional problem arises when a route is split up in two. Since the combined travel time of the two new routes can be less than the travel time of the original route. This results in Requirement VI:

Requirement VI. (Stops are Not Faster):

Insertion of intermediate stops in a route cannot reduce the route's travel time. In other words, when a route is split in two, the sum of the travel times of the route parts cannot be less than the travel time of the original route.

4 Discussion

In this paper, we derived a new formula for link travel times in quasi-dynamic assignment, theoretically underpinned with vertical queuing and instantaneous propagation of traffic flows. We showed that the existing travel time computation procedure of Bliemer et al. (2014) violate requirements for both the absolute correctness and relative correctness of the resulting travel times. We demonstrated that our own proposed definition of link travel times, Eq. (19), does satisfy all requirements and can be used to substitute the Bliemer et al. (2014) formulation.

In addition to satisfying our requirements, our new link travel time definition has several practical advantages. The existence of link travel times simplifies pathfinding in traffic demand modelling. Because link travel times are directly related to the traffic situation on the link, extensions of the traffic flow theory are also easier to incorporate, for example accounting for multiple vehicle types (e.g. trucks, public transport, automated vehicles).

Acknowledgement

This research effort was undertaken for the project *Spatial and Transport impacts of Automated Driving (STAD)*, part of the research programme *Smart Urban Regions of the Future (SURF)* of the Netherlands Organisation for Scientific Research (NWO).

References

- Bakker, D., Mijjer, P., Hofman, F. (1994). QBLOK: een toedelingsmethodiek voor het modelleren van de afhankelijkheid tussen knelpunten en de voorspelling van blokkades. In: Jager, J.M. (ed.), *Colloquium Vervoersplanologisch Speurwerk* 7994. Colloquium Vervoersplanologisch Speurwerk, Amsterdam, The Netherlands, pp. 313-332.
- Bifulco, G.N., Crisalli, U. (1998). Stochastic user equilibrium and link capacity constraints: formulation and theoretical evidencies. Proceedings of the European Transport Conference, pp. 85-96.

- Bliemer, M.C.J., Raadsen, M.P.H., Smits, E., Zhou, B., Bell, M.G.H. (2014). Quasi-dynamic traffic assignment with residual point queues incorporating a first order node model. *Transportation Research Part B*, 68, 363-384.
- Brederode, L., Pel, A., Wismans, L., De Romph, E., Hoogendoorn, S. (2018). Static Traffic Assignment with Queuing: model properties and applications. *Transportmetrica A*.
- Bundschuh, M., Vortisch, P., Van Vuuren, T. (2006). Modelling queues in static traffic assignment. European Transport Conference. Strasbourg, France.
- Bureau of Public Roads. (1964). Theory. In *Traffic assignment manual for application with a large, high speed computer*. U.S. Government Printing Office, Washington, D.C., pp. V1-V24.
- Daganzo, C.F. (1995). The cell transmission model, part II: Network traffic. *Transportation Research Part B*, 29(2), 79-93.
- Gentile, G., Velonà, P., Cantarella, G.E. (2014). Uniqueness of stochastic user equilibrium with asymmetric volume-delay functions for merging and diversion. *EURO Journal* on Transportation and Logistics, 3(3-4), pp. 309-331.
- Lam, W.H.K., Zhang, Y. (2000). Capacity-Constrained Traffic Assignment in Networks with Residual Queues. *Journal of Transportation Engineering*, 126(2), 121-128.
- Lebacque, J.P., Khoshyaran, M.M. (2005). First-Order Macroscopic Traffic Flow Models: Intersection Modeling, Network Modeling. In: Mahmassani, H.S. (ed.), *Transportation and Traffic Theory: Flow, Dynamics and Human Interaction -Proceedings of the 16th International Symposium on Transportation and Traffic Theory*. Elsevier, Oxford, United Kingdom, pp. 365-386.
- Tampère, C.M.J., Corthout, R., Cattrysse, D., Immers, L.H. (2011). A generic class of first order node models for dynamic macroscopic simulation of traffic flows. *Transportation Research Part B*, 45(1), 289-309.