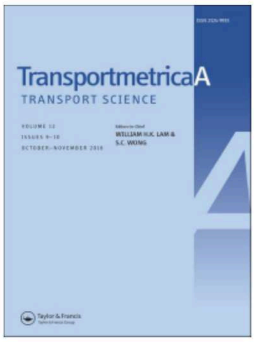

Transportmetrica A: Transport Science

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ttra21

Convergence behavior for traffic assignment characterization metrics

Priyadarshan N. Patil, Katherine C. Ross & Stephen D. Boyles

To cite this article: Priyadarshan N. Patil, Katherine C. Ross & Stephen D. Boyles (2021) Convergence behavior for traffic assignment characterization metrics, Transportmetrica A: Transport Science, 17:4, 1244-1271, DOI: 10.1080/23249935.2020.1857883

To link to this article: https://doi.org/10.1080/23249935.2020.1857883

Published online: 29 Dec 2020.

TRANSPORTMETRICA A: TRANSPORT SCIENCE
2021, VOL. 17, NO. 4, 1244–1271
https://doi.org/10.1080/23249935.2020.1857883

Taylor & Francis
Taylor & Francis Group

Convergence behavior for traffic assignment characterization metrics

Priyadarshan N. Patil[a], Katherine C. Ross[b] and Stephen D. Boyles[b]

[a]Graduate Program in Operations Research and Industrial Engineering, The University of Texas at Austin, Austin, TX, USA; [b]Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, Austin, TX, USA

ABSTRACT
Traffic assignment is used for infrastructure planning, based on metrics like total system travel time (TSTT), vehicle-miles traveled (VMT) and link or path flows. Algorithms for traffic assignment converge to an equilibrium solution over multiple iterations, but these metrics converge at different rates. Current guidance indicates that freeway link flows stabilize at a relative gap of roughly 10^{-4}. This study generalizes this guidance by testing additional networks and metrics, in more experimental settings. Our results reveal that aggregate metrics (VMT and TSTT) stabilize earlier (relative gap 10^{-4}) than link flows (relative gap 10^{-5}), which in turn stabilize slightly before the set of most likely used paths and flows on these paths (relative gap 10^{-6}). These results are stable across the TAPAS and Algorithm B methods for solving assignment. Our results also show strong linear correlations between alternative gap measures, allowing for the translation of stabilization results across other gap definitions as well.

ARTICLE HISTORY
Received 22 March 2020
Accepted 11 November 2020

KEYWORDS
Traffic assignment; convergence criteria; network optimization

1. Introduction

Traffic assignment is a common tool in transportation planning and predicts how travelers will choose routes accounting for congestion effects. Traffic assignment is used in long-term planning, as the final step of the traditional four-step model, to assist in decision-making based on link flows, select link analysis, or shortest path analysis. It also appears as a subproblem in network design, toll-setting, and other related bilevel optimization problems. Despite many advances in dynamic traffic modeling, static assignment remains common in current practice. And despite advances in technology and algorithmic efficiency, computation times are still a relevant issue as agencies move to more detailed, multi-class models, or when assignment is a subproblem in a larger iterative scheme (feedback models, trip table estimation, network design, and so forth). This article therefore focuses on the static traffic assignment problem (TAP) as it is traditionally formulated.

Algorithms for traffic assignment converge to an equilibrium solution in the limit, so a convergence criterion must be introduced to ensure output in finite time. Rose, Daskin,

CONTACT Priyadarshan N. Patil Priyadarshan@utexas.edu Graduate Program in Operations Research and Industrial Engineering, The University of Texas at Austin, 204 E. Dean Keeton St., C2200, Austin, TX 78712-1591, USA

and Koppelman (1988) considered the convergence behavior of the Frank–Wolfe algorithm on small networks (16 nodes). Their study used the relative gap metric (a common convergence criterion, defined below) based on the duality gap. They concluded that it was very difficult to obtain precise estimates of the equilibrium flows in networks which contain only a few O-D pairs with overlapping paths and called for more research on convergence behavior. Boyce, Ralevic-Dekic, and Bar-Gera (2004) found that link flows in the Philadelphia network stabilized once the relative gap (a different definition, also defined below) was below 10^{-4}. This article investigates the convergence behavior of other metrics – specifically total system travel time, vehicle-miles traveled, equilibrium path flows, and the set of used paths – on 12 standard networks. We thus aim to generalize the recommendations of Boyce, Ralevic-Dekic, and Bar-Gera (2004) based on other networks and metrics.

These metrics were chosen to represent a variety of applications. Aggregate measures, such as total system travel time or vehicle-miles traveled, are used to capture the overall state of a network (Harrison et al. 2006; Weisbrod 2008; Higgins 2013; Moudon and Stewart 2013). For instance, the North Carolina Department of Transportation strategic plan uses total system travel time to monitor network performance (Carolina Department of Transportation 2015), Litman (2016) uses vehicle-miles traveled as a sustainability indicator, Qian and Zhang (2012) use total system travel time and vehicle-miles traveled as factors to compare interstate closure scenarios in Sacramento, and the California Department of Transportation uses a reduction in vehicle-miles traveled as a strategic target (California Department of Transportation 2015). A few other examples include usage in the analysis of delivery vehicle impact (José, Cruz, and Ban 2013), credit- or permit-based demand management (Lessan and Fu 2019) and within network design problems as project selection criteria (Shayanfar and Schonfeld 2019). Disaggregate measures, such as link and path flows, more finely describe the impacts of projects on specific regions and populations. Such measures are commonly used by many practitioners and researchers (Cherlow 1981; Daniels, Ellis, and Stockton 1999; Bureau of Transportation Statistics 2015; Seattle Department of Transportation 2016; US Department of Transportation 2016, 2017; Astroza et al. 2017; Boyles et al. 2018; Maryland Department of Transportation 2018).

1.1. Why is static traffic assignment still relevant?

Static traffic assignment has been studied for over five decades now, starting with the convex optimization formulation by Beckmann, McGuire, and Winsten (1956) and described at length in Patriksson (1994) and Boyles, Lownes, and Unnikrishnan (2020). With significant advances in traffic flow theory and traffic assignment in the interim, including the development of dynamic traffic assignment and micro-simulation techniques, it is worth asking whether the traditional traffic assignment problem is still worth studying. Despite the important roles that these other methods play in transportation analysis, there are still several settings where static assignment remains a valuable tool.

Well-known advantages of static assignment include a standard formulation, efficient and provably correct solution algorithms, and guarantees of equilibrium existence and uniqueness. The latter concerns are not strictly mathematical, but have important implications for practice – it is unclear how projects should be evaluated or ranked if multiple, potentially very different solutions exist, or none at all.

A less-appreciated advantage is its greater robustness to errors in input data, such as origin-destination matrices or link and node parameters. Many dynamic traffic assignment models feature queue spillback, which is a significant contributor to traffic congestion in the field. However, spillback introduces discontinuities into the assignment process, potentially amplifying any error or noise in the model inputs. Boyles and Ruiz Juri (2019) showed that when the error in the trip table is sufficiently large, a model *without* spillback actually produces a smaller absolute error in delay estimations than a model *with* spillback. Relatedly, static models are easier to calibrate; despite advances in travel demand modeling, forecasting a time-dependent trip table in a large network remains highly challenging.

Static assignment can also be solved in a shorter amount of time. Even as computational resources expand and more efficient algorithms are developed, in the amount of time required to run a single dynamic assignment it is possible to run multiple static assignments. In applications requiring hundreds or thousands of assignment runs – examples include Monte Carlo simulation to simulate distributions over input parameters (Waller, Schofer, and Ziliaskopoulos 2001; Zhao and Kockelman 2002; Ukkusuri, Mathew, and Waller 2007; Duthie, Unnikrishnan, and Waller 2011), sensitivity analysis (Boyles 2012; Jafari, Pandey, and Boyles 2017), trip table estimation (Yang 1995; Lundgren and Peterson 2008), network design (Yang 1997; Yang and Bell 1998; Josefsson and Patriksson 2007), network pricing (Yang and Lam 1996), and other bilevel optimization problems (Yin 2000) – the computational advantages of static assignment are compelling, if for no other reason than a preliminary screening of alternatives to form a 'shortlist' for more detailed modeling.

We lastly point out a recent line of research showing how a variety of static and dynamic models can be generalized into a single common framework (Bliemer et al. 2017; Bliemer and Raadsen 2020), suggesting that research into one type of traffic assignment model may have relevance to the other as well.

For all of the reasons above, static assignment remains a commonly used tool in transportation planning practice. To be clear, none of this is to argue that static assignment should be universally applied. In applications where the input data are known with high precision, detailed congestion information is essential, and computation times are not constraining (e.g. present-day studies of work zone impacts), dynamic traffic assignment or even microsimulation are likely superior tools. Yet there remain applications where static assignment is preferred, as when inputs are poorly known, or when rapid assessment of a large number of alternatives is preferred to an in-depth assessment of a few (e.g. long-range regional planning, bilevel optimization), and it is such applications that this paper has in mind.

1.2. Why are runtimes still an issue?

With advances in computing and solution algorithms, it is worth asking whether run times are still relevant in static assignment, particularly given the time frame of long-term planning. As discussed in the previous section, there are applications requiring a large number of assignment runs, often with traffic assignment as a subproblem in an iterative scheme. For this reason, research continues in assessing and improving the computational performance of static assignment (Galligari and Sciandrone 2019; Schneck and Nökel 2020).

Furthermore, as computation power advances, network models have increased in scope and resolution. Regional planning models today commonly include tens of thousands of links and nodes, multiple user classes, and feedback to earlier modeling stages to ensure consistency. For large metropolitan areas, even using cutting-edge software and hardware, it is not uncommon for a single model run to take several hours. For a single scenario, this may be acceptable; as part of a bilevel trip table optimization requiring thousands of runs, it is clearly not.

The goal of this paper is to provide guidance on the level of convergence needed, depending on the metric of interest. This allows computational resources to be used as effectively as possible, and not wasted on unnecessary precision beyond the requirements of a particular application.

1.3. Contributions

The central question addressed in this paper is the level of precision needed in the solution. An insufficiently converged solution will not produce reliable estimates for planning. An overly-stringent convergence criterion, on the other hand, wastes computational resources that can be better spent on other model components or tasks. (For instance, examining additional alternatives in a Monte Carlo simulation or solutions in a bilevel program.) The appropriate convergence level depends on the application context: the specific network, the specific metrics of interest, and the decision being made. In particular, the appropriate convergence criterion when producing a single point prediction is likely different than that when selecting a preferred alternative among several. The latter problem introduces several complications relative to the former, so this article focuses primarily on the convergence level needed to stabilize a metric for a single modeling scenario.

The primary guidance to date is based on Boyce, Ralevic-Dekic, and Bar-Gera (2004), and a relative gap level of 10^{-4} or 10^{-5} is common in current software as a default convergence criterion. The manual for Caliper's TransCAD software further suggests that 'since traffic assignment problems vary in many dimensions, some experimentation is warranted to arrive at how much convergence is enough'. (Caliper Corporation 2018). While the study by Boyce, Ralevic-Dekic, and Bar-Gera (2004) played a critical role in determining the necessary level of precision, its experiments were conducted on a single network and considered a single metric (freeway link flows).

The main contribution of this article is to identify the convergence behavior of five metrics on twelve different networks, thus generalizing the analysis of Rose, Daskin, and Koppelman (1988) and Boyce, Ralevic-Dekic, and Bar-Gera (2004). We examine the rates of convergence of these metrics compared to that of relative gap (the most common convergence metric), and identify trends based on network size and congestion level. We additionally describe the heterogeneity in convergence rates between different links within the same network. These analyses primarily have implications for choosing a convergence level for analysis of a particular scenario, and also lay the groundwork for future studies on appropriate convergence criteria for multi-scenario analyses. Our experiments also include tests of different traffic assignment algorithms, and examining a scenario with heterogeneous user classes. We also compare alternative gap functions in current use, suggesting how our results for one gap function can be translated to these alternatives.

The rest of this article is structured as follows: Section 2 provides mathematical specifications of TAP and the metrics we study and reviews current solution algorithms. We next describe our experimental structure, the networks we use, and the design of particular scenarios in Section 3. We next provide the results of these experiments, and our interpretation of these results in Section 4. We finally conclude with a summary of our findings and topics for future study.

2. Background

Consider a directed network with a set of links A, and a set of zones Z. For each link $(i,j) \in A$, let l_{ij} denote its physical length, and t_{ij} its travel time, assumed to be a function of its flow x_{ij} alone. For each origin $r \in Z$ and destination $s \in Z$, let d_{rs} denote the demand for travel between these zones, and let Π^{rs} denote the set of network paths connecting these zones. Further, let Π be the set of all network paths. For a given path π, the number of travelers choosing that path is given by h_π.

The classical formulation of TAP identifies a network state which reflects traveler behavior (all travelers choose a shortest path between their origin and destination) and congestion effects (these shortest paths depend on the choices made by other travelers). Under mild regularity assumptions, such a network state can be identified by solving the following convex program (Beckmann, McGuire, and Winsten 1956):

$$\min_{\mathbf{x},\mathbf{h}} \quad \sum_{(i,j)\in A} \int_0^{x_{ij}} t_{ij}(x)\,\mathrm{d}x \tag{1}$$

subject to:

$$x_{ij} = \sum_{\pi\in\Pi:(i,j)\in\pi} h_\pi \quad \forall (i,j) \in A \tag{2}$$

$$\sum_{\pi\in\Pi^{rs}} h_\pi = d_{rs} \quad \forall (r,s) \in Z^2 \tag{3}$$

$$h_\pi \geq 0 \quad \forall \pi \in \Pi \tag{4}$$

If the link performance functions are strictly increasing, the objective function is strictly convex, and thus has a unique minimum solution in the link flows, which we denote by $\mathbf{x}^*$. This solution is called the user equilibrium (UE) state. We say that a path π is *used* in a solution to TAP if h_π is strictly positive, and define $\Pi_+(\mathbf{h})$ to be the set of used paths at a given solution.

In general, the UE path flow solution is not unique, since many path flow vectors $\mathbf{h}$ can generate the same link flow vector $\mathbf{x}$. The most likely path flows are the unique solution (denoted $\mathbf{h}^*$) to the following optimization problem, representing entropy maximization (Rossi, McNeil, and Hendrickson 1989):

$$\max_{\mathbf{h}} \quad -\sum_{(r,s)\in Z^2} \sum_{\pi\in\Pi^{rs}} h_\pi \log h_\pi \tag{5}$$

subject to:

$$\sum_{\pi\in\Pi:(i,j)\in\pi} h_\pi = x^*_{ij} \quad \forall (i,j) \in A \tag{6}$$

$$\sum_{\pi\in\Pi^{rs}} h_\pi = d_{rs} \quad \forall (r,s) \in Z^2 \tag{7}$$

$$h_\pi \geq 0 \quad \forall \pi \in \Pi \tag{8}$$

Note the constraint that the path flows must generate the UE link flows $\mathbf{x}^*$. The most likely path flows use as many paths as possible given the user equilibrium state (Bar-Gera 2006). Additionally, the formula for entropy calculation can be expressed in terms of link flows obtained from an origin-based assignment as follows:

$$E(\mathbf{x}) = -\sum_{p \in Z} \sum_{(i,j) \in A} x_{(i,j),p} \log \left(\frac{x_{(i,j),p}}{x_{j,p}} \right) \tag{9}$$

where, $x_{(i,j),p}$ is the flow on link (i,j) from origin p and $x_{j,p}$ is the flow through node j originating at p.

Link-based algorithms (Frank and Wolfe 1956; Daneva and Lindberg 2003; Mitradjieva and Lindberg 2013), path-based algorithms (Jayakrishnan et al. 1994; Florian, Constantin, and Florian 2009; Babazadeh et al. 2020), and bush-based algorithms (Bar-Gera 2002; Dial 2006; Nie 2010; Gentile 2014) have all been proposed to solve TAP. Link-based algorithms only track the link flows $\mathbf{x}$ at each iteration, and the path flows are implicit and must be calculated by post-processing. Such algorithms are economical on computer memory but are slow to obtain high-precision solutions. Path-based and bush-based algorithms produce precise solutions more quickly at the expense of additional memory requirements. Path-based algorithms explicitly track a path flow vector $\mathbf{h}$. Bush-based algorithms do not, but structure a disaggregated link flow solution in a way that a corresponding path flow vector can be easily calculated.

Most algorithms for TAP do not produce a most likely path flow solution, and a second algorithm is needed for that purpose. TAPAS (Bar-Gera 2010) is a recent algorithm which both solves TAP and provides a path flow solution satisfying *proportionality*, a slightly weaker condition than entropy maximization. TAPAS, and related algorithms derived from it, are highly efficient (Xie and Xie 2015). This study uses a combination of the TAPAS and Frank–Wolfe algorithms for its primary experiments, and also uses Algorithm B (a bush-based algorithm) to test transferability of the results.

In our analysis, we do not test any path-based algorithms. To economize on memory, such algorithms employ 'column dropping' rules to store as few paths as possible. However, such solutions have extremely low entropy (indeed, the most likely solution spreads flow over as many paths as possible), and thus the specific path flow solution is untrustworthy for further analysis; see the discussion and empirical results from Bar-Gera (2006) and Bar-Gera and Luzon (2007). Since we wish to examine convergence of the path flows in the solution, it is clearest to do so using an algorithm which converges to the (unique) entropy-maximizing path flow solution, rather than an arbitrary path flow equilibrium.

Given a feasible solution $(\mathbf{x}, \mathbf{h})$ to TAP, we select five metrics for analysis. (Since algorithms for TAP converge only in the limit, we do not demand optimal solutions to the above problems.) The *total system travel time* (TSTT) expresses the sum of each vehicle's travel time in the network:

$$\text{TSTT}(\mathbf{x}) = \sum_{(i,j) \in A} t_{ij} x_{ij} \tag{10}$$

Vehicle-miles traveled (VMT) expresses the total distance traveled by vehicles in the network:

$$\text{VMT}(\mathbf{x}) = \sum_{(i,j) \in A} l_{ij} x_{ij} \tag{11}$$

To measure convergence of these metrics, we calculate the relative difference between their values at the current solution $\mathbf{x}$ and the equilibrium solution $\mathbf{x}^*$:

$$\Delta\text{TSTT}(\mathbf{x}) = \frac{\text{TSTT}(\mathbf{x}) - \text{TSTT}(\mathbf{x}^*)}{\text{TSTT}(\mathbf{x}^*)} \tag{12}$$

$$\Delta\text{VMT}(\mathbf{x}) = \frac{\text{VMT}(\mathbf{x}) - \text{VMT}(\mathbf{x}^*)}{\text{VMT}(\mathbf{x}^*)}. \tag{13}$$

Both TSTT and VMT are aggregate metrics. To represent convergence of the specific link and path flows themselves, we measure the proportion of links (or paths) within a given relative threshold ϵ of their equilibrium values. Let $A^*_\epsilon(\mathbf{x})$ denote the set of links with flows within this threshold:

$$A^*_\epsilon(\mathbf{x}) = \{(i,j) \in A : |x_{ij} - x^*_{ij}| < \epsilon x^*_{ij}\}. \tag{14}$$

Let $\Pi^*_\epsilon(\mathbf{h})$ denote the set of paths whose flows are within this threshold:

$$\Pi^*_\epsilon(\mathbf{h}) = \{\pi \in \Pi_+(\mathbf{h}^*) : |h_\pi - h^*_\pi| < \epsilon h^*_\pi\}, \tag{15}$$

where $\mathbf{h}^*$ is the (entropy-maximizing) solution to the most likely path flows problem at equilibrium. Note that $\Pi^*_\epsilon(\mathbf{h})$ is a subset of the used path set at equilibrium. Using these sets, we define the *proportion of unconverged links (PUL)* as

$$\text{PUL}(\mathbf{x}, \epsilon) = 1 - \frac{|A^*_\epsilon(\mathbf{x})|}{|A|}, \tag{16}$$

and the *proportion of unconverged paths (PUP)* as

$$\text{PUP}(\mathbf{h}, \epsilon) = 1 - \frac{|\Pi^*_\epsilon(\mathbf{h})|}{|\Pi_+(\mathbf{h}^*)|}. \tag{17}$$

Finally, we define the *path set deviation (PSD)* to represent how the set of used paths converges to set of equilibrium paths by defining

$$\text{PSD}(\mathbf{h}) = 1 - \frac{|\Pi_+(\mathbf{h}) \cap \Pi_+(\mathbf{h}^*)|}{|\Pi_+(\mathbf{h}^*)|}. \tag{18}$$

We thus have PSD $= 1$ if the set of currently used paths and the set of equilibrium paths is disjoint, and PSD $= 0$ if every equilibrium path is in the current set of used paths. As with the other metrics, it should decrease to zero over successive iterations.

Both PUP and PSD are calculated with respect to the used paths at equilibrium. Some restriction of the path set is necessary, since the number of paths grows exponentially with network size and the vast majority of these are unused. Such paths should not be considered in our metrics, and we decided to measure PUP and PSD relative to the equilibrium path sets to be consistent with the other metrics (which are measured relative to the equilibrium link flows and most likely path flows). Defining PUP only based on the set of used paths at equilibrium is important because PUP is a relative error measure; any path for which $h^*_\pi = 0$ would thus appear 'unconverged' even with an infinitesimal flow value. This is not a serious deficiency, because any solution placing positive flow on a non-equilibrium path must also place the 'wrong' value on at least one path in $\Pi_+(\mathbf{h}^*)$ (by flow conservation), which would be detected by PUP with an appropriate ϵ value.

These five metrics – ΔTSTT, ΔVMT, PUL, PUP, and PSD – are directly related to practical applications of traffic assignment, and converge to zero as $\mathbf{x}$ and $\mathbf{h}$ approach $\mathbf{x}^*$ and $\mathbf{h}^*$, respectively. However, they are not suitable convergence criteria because they can only be evaluated if the equilibrium link flows or most likely path flows are already known, and there would be no need to solve TAP if this were true. Therefore, in practice convergence is measured using information available even at intermediate solutions.

The relative gap is one such measure. There are several definitions of relative gap in common practice; we use the following one for our experiments, and later discuss relationships with its alternatives. Let $\kappa_{rs}(\mathbf{x})$ denote the travel time on the shortest path from origin r to destination s using the link travel times corresponding to $\mathbf{x}$. The shortest path travel time (SPTT) can then give the total travel time we would expect if all vehicles were on shortest paths (as the UE condition requires):

$$\mathrm{SPTT}(\mathbf{x}) = \sum_{(r,s)\in Z^2} \kappa_{rs}(\mathbf{x}) d_{rs} \tag{19}$$

The gap and relative gap of a feasible solution, as defined in Rose, Daskin, and Koppelman (1988), are:

$$\mathrm{gap}(\mathbf{x}) = \mathrm{SPTT}(\mathbf{x}) - \mathrm{TSTT}(\mathbf{x}) \tag{20}$$

and

$$RG(\mathbf{x}) = -\frac{\mathrm{gap}(\mathbf{x})}{\mathrm{SPTT}(\mathbf{x})} = \frac{\mathrm{TSTT}(\mathbf{x})}{\mathrm{SPTT}(\mathbf{x})} - 1. \tag{21}$$

Relative gap is non-negative, and equal to zero only at equilibrium solutions, and thus is a valid gap function. Other gap metrics used for convergence include alternative definitions of relative gap, average excess cost (AEC) and average total reduced cost. We next define a variant of relative gap, the one used by Boyce, Ralevic-Dekic, and Bar-Gera (2004), and AEC, and discuss their relationship with the RG definition of equation (21). This section contains a brief mathematical discussion, and our results include an empirical comparison.

An alternate definition of relative gap (RG') normalizes the gap by a lower bound on the optimal value of the Beckmann function in equation (1). The lower bound calculated from a particular solution is given by

$$\mathrm{LB}(\mathbf{x}) = \sum_{(i,j)\in A} \int_0^{x_{ij}(k)} t_{ij}(x)\,\mathrm{d}x + \mathrm{gap}(\mathbf{x}), \tag{22}$$

and the best lower bound is the tightest bound over the flow vectors $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_k}$ seen over successive iterations thus far:

$$\mathrm{BLB} = \max_k \{\mathrm{LB}(\mathbf{x_k})\}. \tag{23}$$

The relative gap is then given by

$$RG'(\mathbf{x}) = -\frac{\mathrm{gap}(\mathbf{x})}{|\mathrm{BLB}|}. \tag{24}$$

The average excess cost is defined as

$$\mathrm{AEC}(\mathbf{x}) = -\frac{\mathrm{gap}(\mathbf{x})}{\sum_{(r,s)\in Z^2} d_{rs}}. \tag{25}$$

Observe that the numerator of all three gap functions is the same, and they differ only in how they are normalized. The ratio of AEC and RG equals the ratio of SPTT and $\sum_{(r,s)\in Z^2} d_{rs}$, which is the average travel time on the all-or-nothing assignment. As SPTT stabilizes close to convergence, the ratio between AEC and RG will approach a constant representing average travel time.

To compare RG and RG', we compare LB and SPTT, and see that their ratio is

$$\frac{\mathrm{LB}(\mathbf{x})}{\mathrm{SPTT}(\mathbf{x})} = \frac{\sum_{(i,j)\in A}\int_0^{x_{ij}} t_{ij}(x)\,\mathrm{d}x}{\mathrm{SPTT}(\mathbf{x})} + \frac{\mathrm{SPTT}(\mathbf{x}) - \mathrm{TSTT}(\mathbf{x})}{\mathrm{SPTT}(\mathbf{x})} \tag{26}$$

The second term in this equation is $-RG(\mathbf{x})$, which approaches zero as equilibrium is reached. To analyze the first term, we note that $\mathrm{SPTT}(\mathbf{x})$ and $\mathrm{TSTT}(\mathbf{x})$ become asymptotically equal, and so

$$\frac{\sum_{(i,j)\in A}\int_0^{x_{ij}} t_{ij}(x)\,\mathrm{d}x}{\mathrm{SPTT}(\mathbf{x})} \approx \frac{\sum_{(i,j)\in A}\int_0^{x_{ij}} t_{ij}(x)\,\mathrm{d}x}{\sum_{(i,j)\in A} x_{ij} t_{ij}(x_{ij})}. \tag{27}$$

On the right-hand side of equation (27), both the numerator and denominator include a sum over links. For any specific link, the difference between its term in the numerator and its term in the denominator is illustrated in Figure 1, which shows a typical link performance function. The hatched area is the term in the numerator, whereas the area of the rectangle is its term in the denominator. We see visually that these two areas are approximately equal. To compare them numerically, consider a link for which $t_{ij} = t_{ij}^0(1 + 0.15(x_{ij}/u_{ij})^4)$, where t_{ij}^0 and u_{ij} are its free-flow time and 'practical capacity', respectively.[1] When the link is used relatively heavily ($x_{ij} = u_{ij}$), the ratio between the areas is 0.89. Even when the link is highly

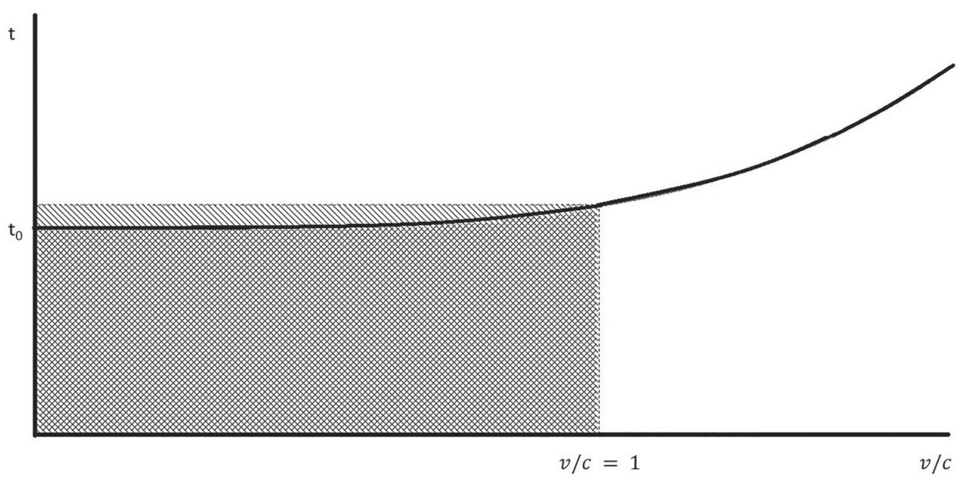

Figure 1. Visualization of RG and RG' ratio term.

congested ($x_{ij} = 1.5u_{ij}$), the ratio between the areas is 0.65, and the areas are of the same order of magnitude. We thus expect the ratio $\mathrm{LB}(\mathbf{x})/\mathrm{SPTT}(\mathbf{x})$ to be fairly close to one, and thus RG and RG' to have a similar order of magnitude.

To summarize, we expect our results for RG (which will be given in terms of order of magnitude) to translate more or less directly to RG', although the specific numerical value may differ by up to 30%. To translate them to AEC, one must multiply by the average travel time in the network (whose order of magnitude can be estimated *a priori*). If travel times are reported in minutes and the network represents a typical metropolitan region, we would expect the AEC for a given solution to be one to two orders of magnitude higher than the RG, and our results can be adapted accordingly. Our experiments below validate this analysis numerically.

3. Data and experiment design

The main objective of this paper is to determine the relationship between ΔTSTT, ΔVMT, PUL, PUP, and PSD (which carry more practical meaning) and the corresponding RG level (which can be calculated without knowing the equilibrium solution). We choose to index these results to relative gap, rather than iteration count or another measure of progress, because gap functions serve as an absolute measure of convergence that can be applied regardless of algorithm or parameter settings.

This section explains the procedures we used to determine the relationship between the five convergence metrics and relative gap. We first discuss the networks and algorithms used, and choices of specific parameters. We next discuss how we obtained solutions of a particular relative gap level for analysis.

The networks studied in this paper are shown in Table 1, all obtained from the Transportation Networks for Research repository (Stabler 2019). For ease of reference, we categorize the networks roughly by size: Sioux Falls through Anaheim are designated as small, Chicago Sketch through Terrassa are designated as medium, and the remaining networks are designated as large. The last column in this table shows the average equilibrium flow-to-capacity ratios, excluding centroid connectors. We consider networks with ratios of less than 0.5 to be uncongested, with ratios between 0.5 and 1.0 to be semi-congested, and networks with ratios greater than 1.0 to be congested. The Terrassa network is a clear outlier in this regard, assigning over 25 million trips in a region whose current population is around

Table 1. Description of networks used.

Network name	Zones	Links	Nodes	Trips	Average flow-to-capacity ratio
SiouxFalls	24	76	24	360,600	1.612
Eastern-Massachusetts	74	258	74	65,576	0.163
Anaheim	38	914	416	104,694	0.297
Chicago-sketch	387	2950	933	1,260,907	0.257
Berlin-Prenzlauerberg-Center	98	2184	975	23,648	0.121
Barcelona	110	2522	1020	184,679	1.137
Winnipeg	147	2836	1052	64,784	2.028
Terrassa	55	3264	1609	25,225,700	5.964
Austin	7388	18,961	7388	739,351	0.875
Berlin-Center	865	28,376	12,981	168,222	0.092
Chicago-Regional	1790	39,018	12,982	1,360,427	0.522
Philadelphia	1525	40,003	13,389	18,503,872	0.949

200,000, resulting in a flow-to-capacity ratio of almost 6. While such a demand level may not be realistic, we nevertheless include this network as a 'stress test' to see whether consistent trends can be seen even in extremely congested networks.

Calculating ΔTSTT, ΔVMT, PUL, PUP, and PSD requires the equilibrium link flow and most likely path flow solutions. Near-equilibrium link flows $\mathbf{x}^*$ and proportional path flows $\mathbf{h}^*$ were obtained using the TAPAS implementation by Perederieieva et al. (2015) (with default settings for TAPAS parameters used to determine cost-effective PAS and flow-effective PAS), setting a relative gap of 10^{-12} as the termination criterion. Our experiments will cover solutions over a range of gap levels between 10^{-3} and 10^{-8}. Over this range, we expect the distinction between using our reference solution with a gap of 10^{-12}, and an exact equilibrium, to be small. For calculating the proportions of unconverged links and unconverged paths, we chose a threshold of $\epsilon = 0.01$, to be consistent with Boyce, Ralevic-Dekic, and Bar-Gera (2004). (The results contain a sensitivity analysis with respect to this parameter.) All experiments were conducted on a machine with Ubuntu 18.04, 8 GB of memory and Intel i5 processor @ 3.30 GHz.

Our analysis involves solutions at six target gap levels: $10^{-3}, 10^{-4}, 10^{-5}, \ldots, 10^{-8}$. To facilitate comparison between different networks, we obtained solutions on each network whose relative gap was within 10% of these levels (e.g. between 0.0009 and 0.0011 for 10^{-3}). Obtaining solutions on each network with such specific gap levels is not trivial, since algorithms for TAP are designed to reach equilibrium as rapidly as possible, and not aim for a specific nonzero gap level. Therefore, we used the procedure described below to generate solutions at a specified target gap level γ. This procedure involves a hybrid of the TAPAS implementation described above, and an implementation of Frank–Wolfe (Boyles 2019) with ten bisection iterations per flow shift.

(1) Run TAPAS with a termination criterion of γ as relative gap. If the solution is in the acceptable range $[0.9\gamma, 1.1\gamma]$, return the link flows $\mathbf{x}$ and path flows $\mathbf{h}$, along with the values of ΔTSTT, ΔVMT, PUL, PUP, and PSD.
(2) If the returned solution has a gap level less than 0.9γ, examine the solution from the previous iteration to see if it is in the acceptable range $[0.9\gamma, 1.1\gamma]$. If so, return the link and path flows, and the five metrics, for that solution.
(3) If neither of the above solutions is in the acceptable range, initialize Frank–Wolfe with the TAPAS solution from the previous iteration. Perform iterations of Frank–Wolfe until the gap is in the acceptable range $[0.9\gamma, 1.1\gamma]$, and return that solution and the corresponding metrics. Frank–Wolfe is used due to its relatively slow solution improvement, which leads to flow values within the desired gap range without 'skipping over'.

This process is repeated for each network and target gap level. This procedure worked for all but seven scenario-*RG* value combinations, due to the Frank–Wolfe algorithm in the last step jumping over the acceptable gap range. In these remaining cases, we repeated the last step, restarting Frank–Wolfe algorithm with the prior flow pattern, but fewer bisection iterations.

Our experiments are divided into the following analyses:

(1) Identify the convergence rates of TSTT, VMT, and path/link flows to their equilibrium values using the procedure described above. We use both the base demand levels and

adjusted demand levels to study how congestion levels affect the convergence of these metrics. (This set of experiments is the most extensive, and is used as the basis for our core recommendations.)

(2) Repeating the analysis using Algorithm B (Dial 2006) to investigate transferability of results to algorithms besides TAPAS.
(3) Repeating the analysis in a setting with toll roads and two user classes with distinct values of time.
(4) Investigating the effect of different convergence levels when evaluating candidate solutions as a subproblem in network design, a bilevel program.
(5) Comparison of the alternative gap measures RG, RG', and AEC so results can be translated appropriately. The intent is to numerically validate the mathematical analysis in the previous section, which involved several approximations.

For the second set of experiments, we used our own implementation of Algorithm B; the source code is available at Boyles (2019). Aggregate metrics (TSTT, VMT) and PUL were calculated from these experiments. The path-based metrics PUP and PSD were not calculated, since Algorithm B does not aim to maximize entropy or provide a proportional path flow solution.

For the third set of experiments, we introduced two user classes distinguished by their value of time (\$15/h and \$30/h), and used the toll values given in the network instances, where present. For networks without toll roads, we randomly selected 10% of links to be tolled. These experiments used Algorithm B, as our implementation supports multi-class assignment and the available TAPAS implementation does not.

For the fourth set of experiments, we formulated a network design problem, in which a discrete set of links had to be chosen for "upgrade'. An upgrade increased a link's capacity by 50%, and had a cost proportional to its length. The budget allowed upgrading up to 5% of the total length of all links, and the objective is to minimize TSTT subject to equilibrium constraints. This is a classic bilevel problem which is intractable to solve exactly. We thus solved it heuristically, using the genetic algorithm implementation in the pyeasyga library (Remi-Omosowon 2020). Such an algorithm involves solving a number of TAP instances as subproblems to evaluate fitness of candidate solutions. We used the default values for parameters in this library. For each network, we varied the RG level used for evaluating the TAP fitness function. Each combination of objective function and RG level was solved five times, and average performance reported.

For the fifth set of experiments, we calculate RG, RG', and AEC for the solutions obtained in the previous analyses, and conduct a linear regression to investigate whether $RG \approx RG'$, and that RG and AEC differ by a nearly constant multiplicative factor, as was suggested by the approximate mathematical analysis in the previous section.

4. Results

This section reports the results from the experiments described above. Each set of results is presented in its own subsection. Experiment 1 forms the core of our analysis, and is described in the greatest detail. The remaining experiments are described more briefly, highlighting key differences from the core analysis results. The Appendix to the paper contains detailed results, separated for each of the 12 networks under consideration. The

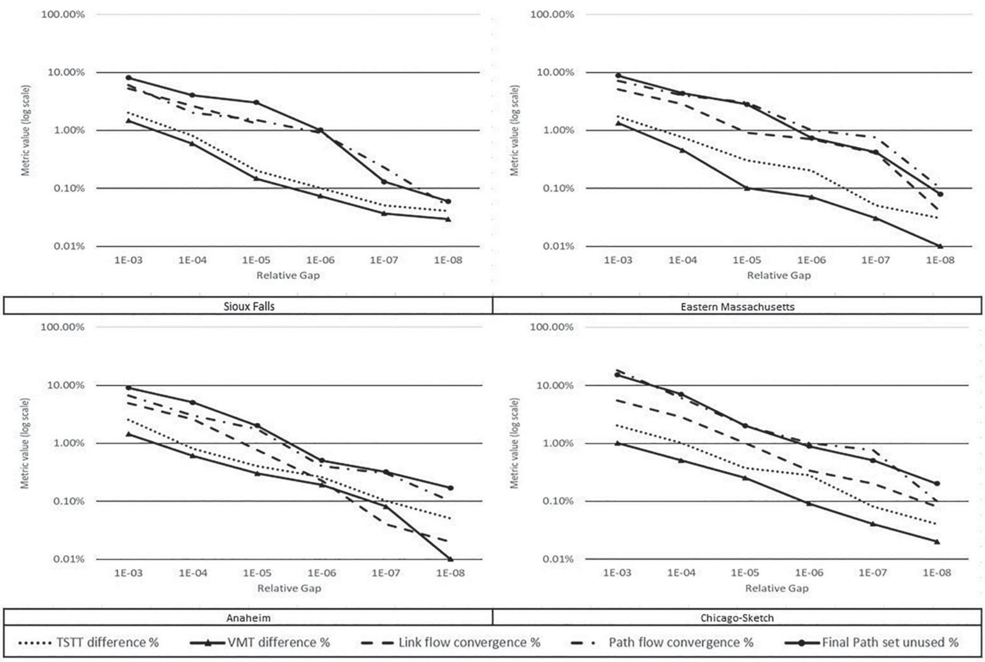

Figure 2. Stabilization behavior of metrics at default demand, small networks.

figures and tables in this section present summaries of this data, presenting the most important findings from each experiment.

4.1. Experiment 1: Network metric behavior results

The values of the main convergence metrics (ΔTSTT, ΔVMT, PUL, PUP, and PSD) are shown in Figures 2–4 (presented according to network and size), and in Figures 5–9 (presented according to each metric). In the latter set of figures, the thin lines represent the values of each metric in one of the twelve networks tested, and the thick line represents the average value. Both sets of figures use logarithmic axes both for the relative gap, and for each metric, to focus on the orders of magnitude in these values. Table 2 shows the numerical values of the metric means, as well as the highest and lowest values seen at a particular level across all networks. The raw data, containing the specific values for over all twelve networks, is found in the Appendix (Table A1).

All five metrics converged at roughly similar rates, despite significant differences in the size and congestion level of the networks tested. This is encouraging from the standpoint of providing transferable, practical advice on convergence thresholds.

In all of the networks, the aggregate metrics (TSTT and VMT) are already very near stabilization at a relative gap of 10^{-3}. For the small and medium networks, these values are within 1% of the equilibrium values when the relative gap is 10^{-4}, and for the large networks, they are within 2%. Both ΔTSTT and ΔVMT converge at roughly similar rates, but ΔVMT is usually slightly lower at a particular gap level. We believe this is because the link

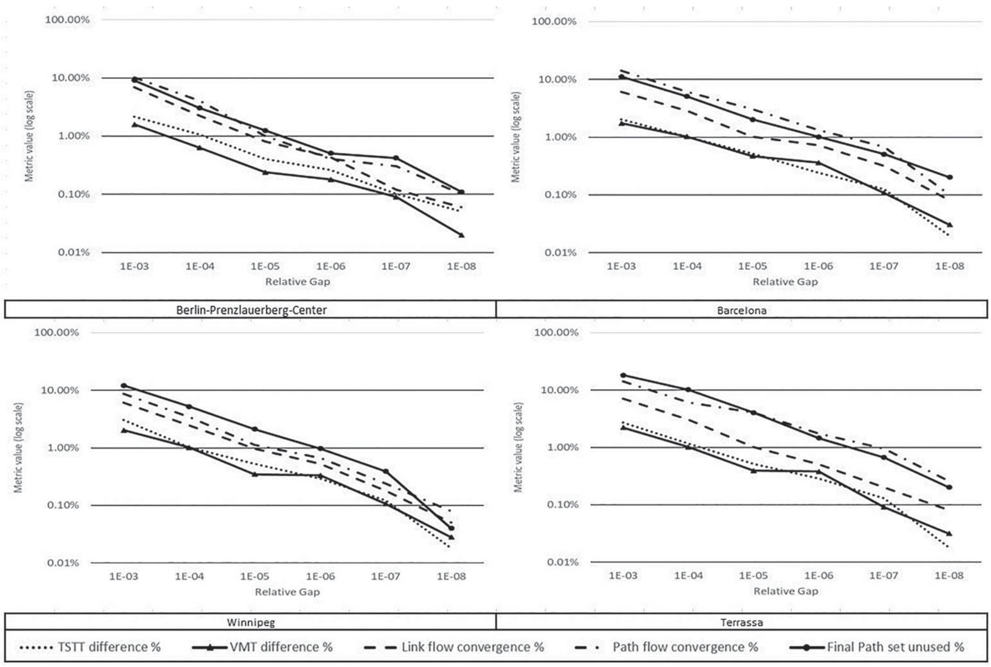

Figure 3. Stabilization behavior of metrics at default demand, medium networks.

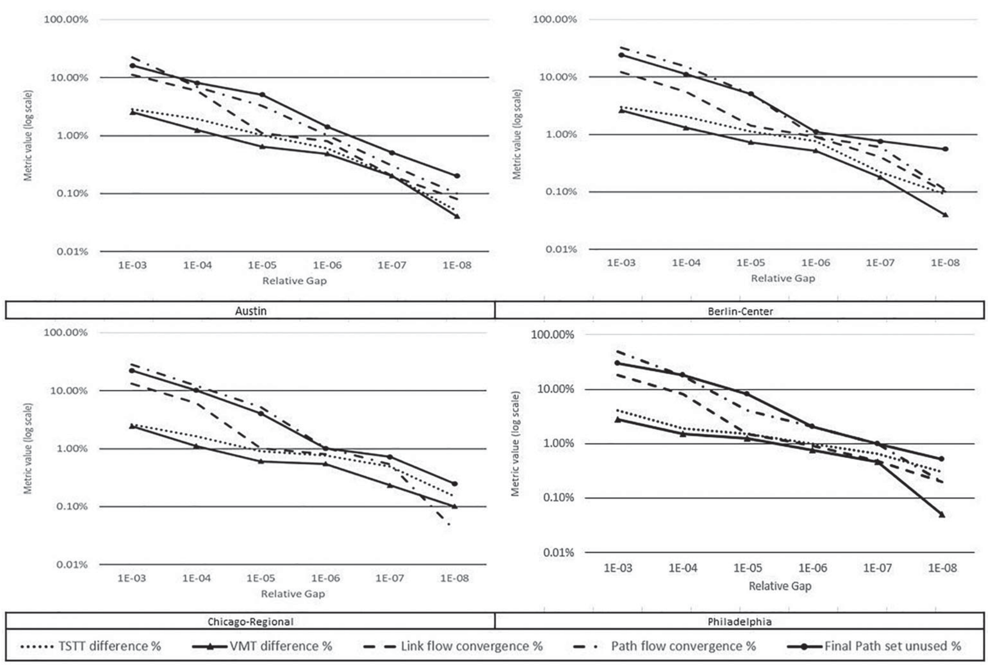

Figure 4. Stabilization behavior of metrics at default demand, large networks.

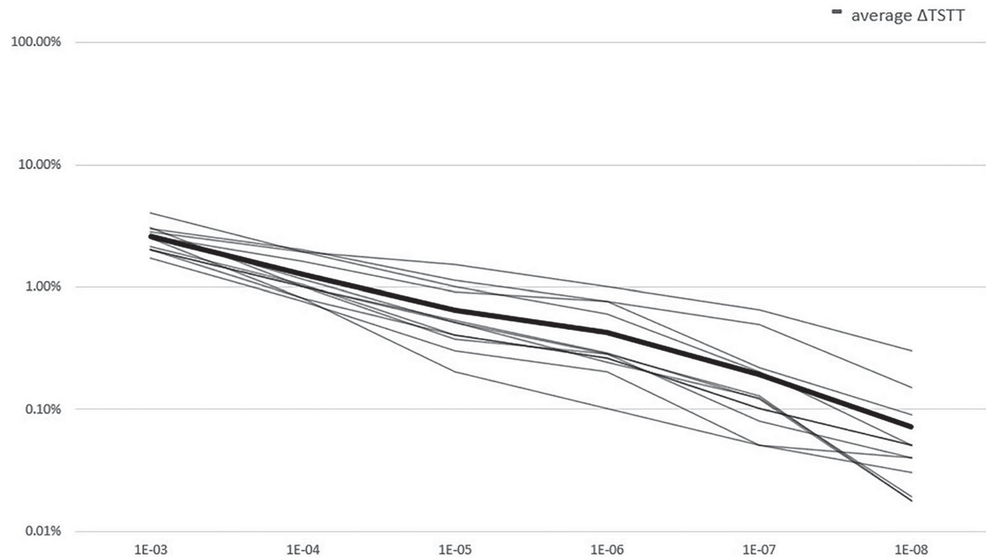

Figure 5. ΔTSTT trends for different gap levels.

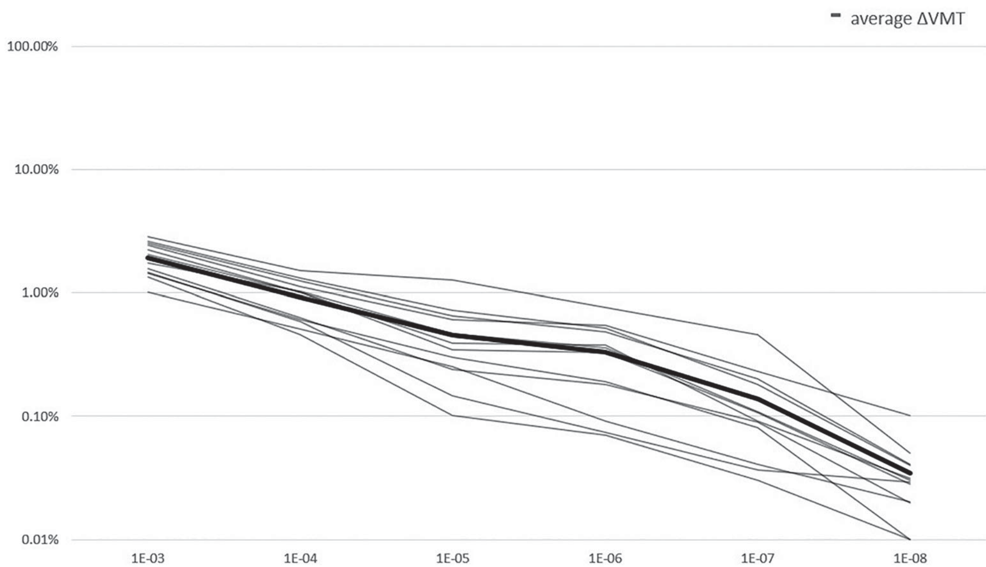

Figure 6. ΔVMT trends for different gap levels.

lengths are constant, and thus only the flows are changing between iterations when calculating VMT. By contrast, the calculation of TSTT involves flows and travel times, both of which are changing.

The proportion of unconverged links was the metric originally studied by Boyce, Ralevic-Dekic, and Bar-Gera (2004) for the Philadelphia regional network. They found that a gap of 10^{-4} was required to approach convergence for freeway links, defining convergence as a PUL of 1% or less. To achieve this level of convergence for arterial links as well as freeway

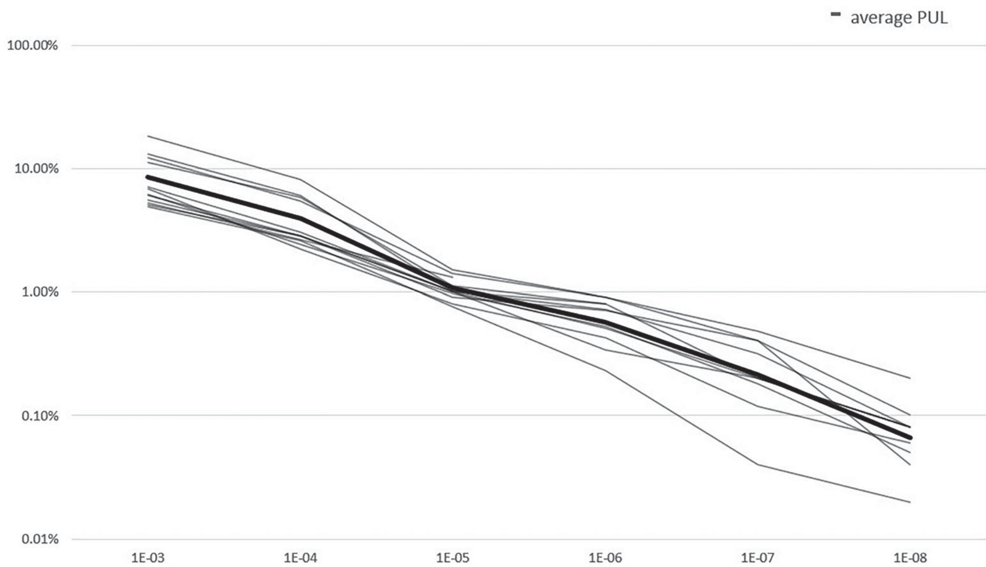

Figure 7. PUL trends for different gap levels.

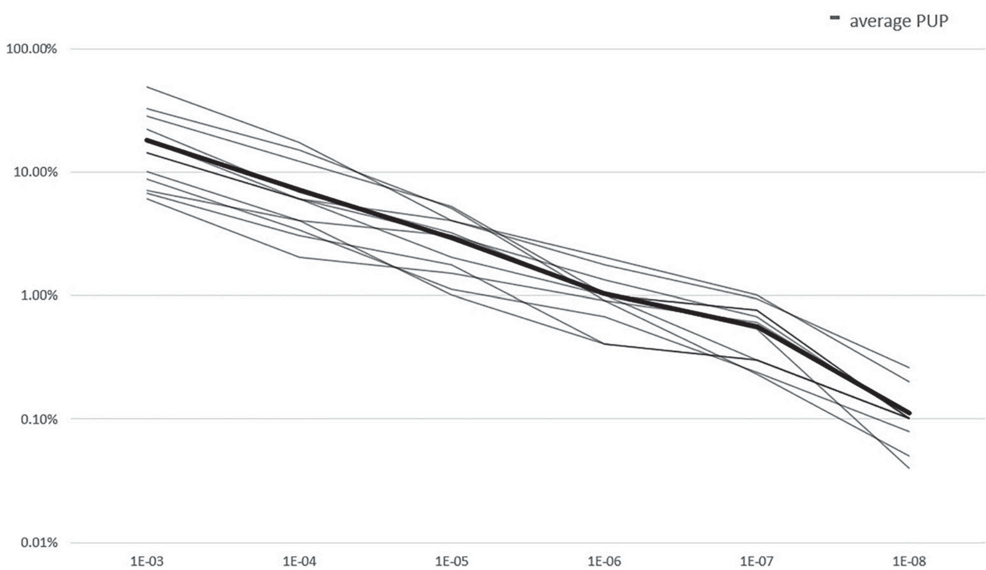

Figure 8. PUP trends for different gap levels.

links, a relative gap of 10^{-5} was needed. Our results show that this latter conclusion generally holds across the other networks tested, and that 99% of link flows are accurate to within 1% of equilibrium values at this gap level.

Link flow behavior for multiple ϵ thresholds can be seen in Figure 10. Trends are similar within network size grouping, and variations therein are caused by differing congestion levels. For instance, Austin and Philadelphia show a similar proportion of links in various

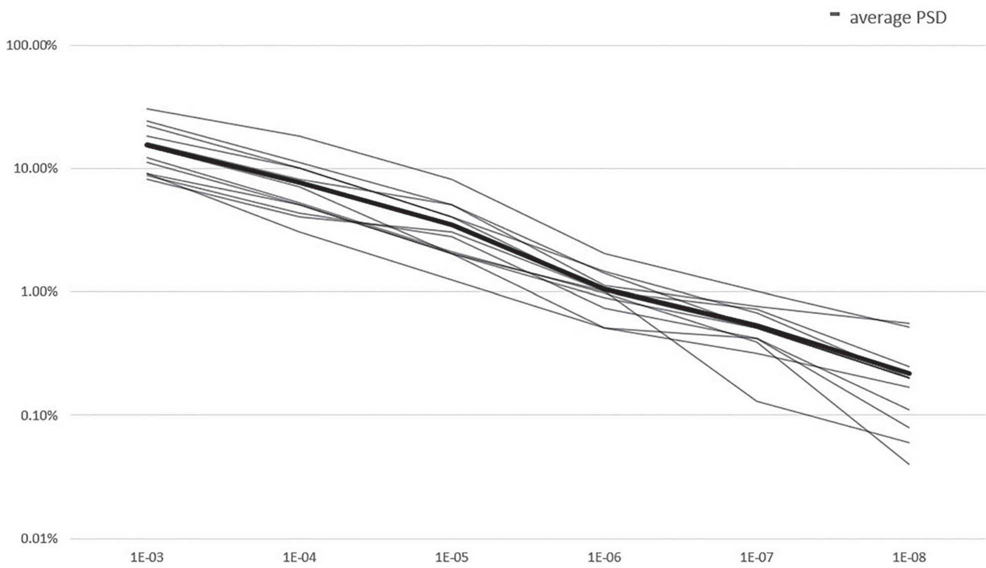

Figure 9. PSD trends for different gap levels.

Table 2. Metric stabilization behavior data (average) using TAPAS.

Gap Level	ΔTSTT			ΔVMT			PUL		
	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max
1E-03	1.72%	2.56%	4.05%	1.01%	1.93%	2.84%	4.89%	8.46%	18.23%
1E-04	0.76%	1.26%	2.02%	0.46%	0.92%	1.52%	2.22%	3.92%	8.10%
1E-05	0.20%	0.65%	1.52%	0.10%	0.46%	1.27%	0.77%	1.07%	1.52%
1E-06	0.10%	0.42%	1.01%	0.07%	0.33%	0.76%	0.00%	0.58%	0.91%
1E-07	0.05%	0.19%	0.66%	0.03%	0.14%	0.46%	0.00%	0.21%	0.49%
1E-08	0.02%	0.07%	0.30%	0.01%	0.03%	0.10%	0.00%	0.07%	0.20%

Gap Level	PUP			PSD		
	Min	Mean	Max	Min	Mean	Max
1E-03	6.08%	18.07%	48.60%	8.10%	15.41%	30.38%
1E-04	2.03%	7.19%	17.21%	3.04%	7.63%	18.23%
1E-05	1.01%	2.93%	5.26%	1.25%	3.47%	8.10%
1E-06	0.41%	1.04%	2.03%	0.51%	1.06%	2.03%
1E-07	0.23%	0.56%	1.01%	0.13%	0.53%	1.01%
1E-08	0.04%	0.11%	0.26%	0.04%	0.22%	0.56%

relative error regimes, but Winnipeg and Chicago Sketch differ due to higher congestion on the Winnipeg network. This also relates to the relationship between congestion level and stabilization, explored later.

The remaining two metrics (proportion of unconverged paths and path set deviation) are the last to stabilize. Relative gap levels of 10^{-6} were needed before these metrics decreased to 1% or less. We believe this occurs because the number of used paths grows quickly with network size. For instance, in the Philadelphia network, the equilibrium solution uses over 300 million paths. Most of these paths necessarily have small flow, and changes in even a single link will change the flows across many paths.

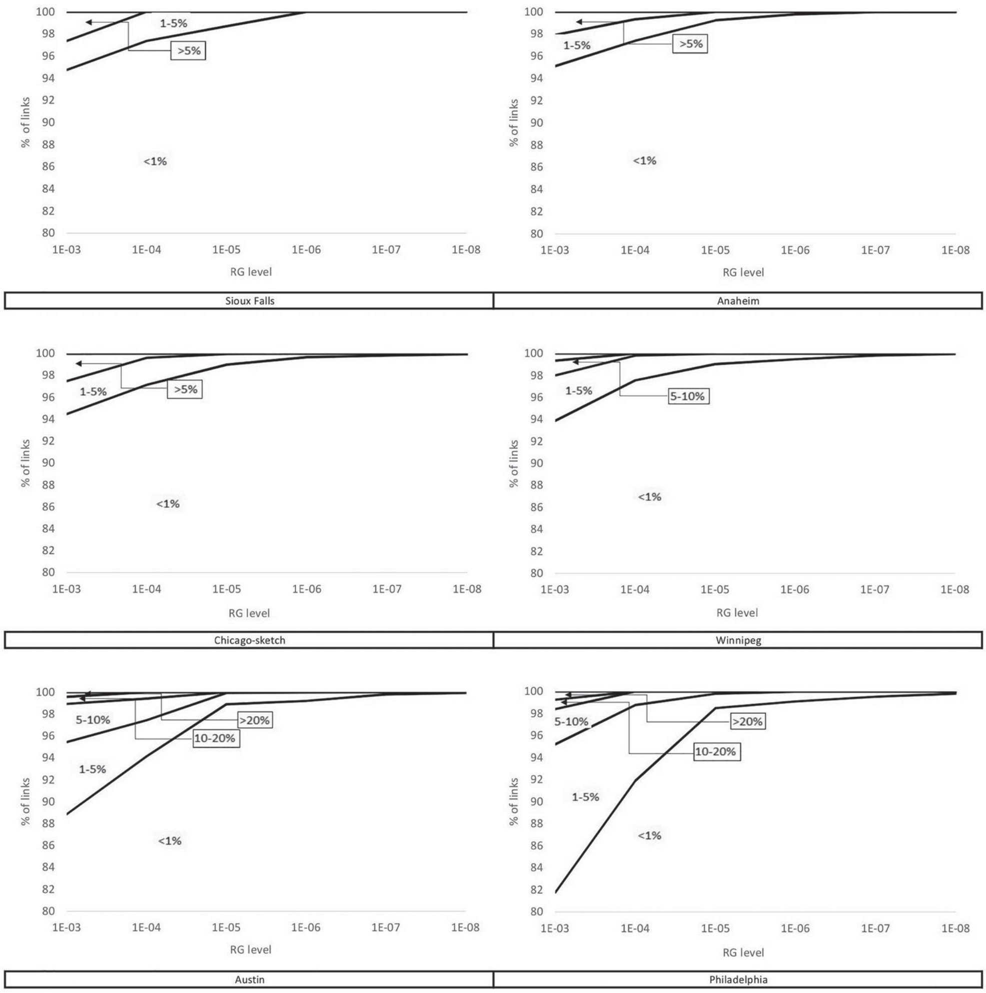

Figure 10. Link flow trends for various ϵ thresholds.

Table 3 shows how the values of PUP vary across networks for different choices of ϵ, for a fairly converged solution of $RG = 10^{-6}$. This allows us to see the distribution of path convergence, similar to Figure 10 for links. We see that virtually all paths ($\approx$98%) are within 1% of their equilibrium values; almost all ($\approx$95% or more) are within 0.1% of their equilibrium values; and the significant majority ($>$ 85%) are within 0.01%. A negligible number of paths (roughly one in a thousand) remain more than 10% from their equilibrium values.

Table 4 provides the entropy values for Chicago-Regional and the Philadelphia networks at various RG levels. The entropy values show a clear increasing and convergent trend towards the final entropy value for each network-algorithm pair. Thus, as the network flow stabilizes, it tends to increase entropy, regardless of the algorithm used. As path flow patterns are intricately linked to entropy values, it stabilizes to within 1% of the convergence value at a RG level of 10^{-6} and below, in line with the observed behavior of PUP and PSD metrics.

Table 3. PUP sensitivity analysis w.r.t. ϵ.

Sensitivity analysis for PUP(ε) (RG level of 10^{-6})				
Net \ ε	0.0001	0.001	0.01	0.1
Sioux Falls	7.02%	2.24%	0.91%	0.05%
Anaheim	4.53%	1.38%	0.41%	0.02%
Chicago	8.92%	3.30%	1.01%	0.03%
Winnipeg	5.42%	1.85%	0.67%	0.01%
Austin	9.18%	3.79%	1.01%	0.06%
Philadelphia	14.23%	5.41%	2.03%	0.17%

Table 4. Entropy values for varying *RG* values.

	Chicago-regional		Philadelphia	
Relative gap	AlgB	TAPAS	AlgB	TAPAS
10^{-3}	370,682.16	850,027.30	2,609,915.27	4,797,572.35
10^{-4}	380,578.59	863,914.76	2,654,382.74	4,958,732.03
10^{-5}	389,116.51	885,997.08	2,724,551.75	5,054,152.48
10^{-6}	395,337.48	898,332.39	2,783,409.60	5,137,311.26
10^{-7}	397,353.30	908,074.29	2,793,543.55	5,189,488.01
10^{-8}	397,600.29	919,463.21	2,794,690.11	5,218,702.51
10^{-12}	397,654.25	920,159.60	2,794,750.93	5,226,184.24

Table 5. Used paths (in millions) for various relative gap values.

Gap Level	Chicago Regional	Philadelphia
1E-03	86.877	352.014
1E-04	88.819	358.008
1E-05	91.837	371.205
1E-06	91.482	368.607
1E-07	91.866	369.873
1E-08	91.895	369.998
1E-12	92.265	370.108

We further investigate how path flows converge on the largest networks, by showing how the number of used paths stabilizes on the Chicago Regional and Philadelphia networks at convergence levels up to 10^{-12} relative gap. This is shown in Table 5. We see that the number of used paths increases with the solution precision, but that this number appears to converge, perhaps to the number of used paths at the exact equilibrium.

Table 6. Metric stabilization behavior data (average) for Algorithm B.

Gap Level	ΔTSTT		ΔVMT		PUL	
	TAPAS	AlgB	TAPAS	AlgB	TAPAS	AlgB
1E-03	2.56%	2.38%	1.93%	1.73%	8.46%	7.57%
1E-04	1.26%	1.13%	0.92%	0.83%	3.92%	3.57%
1E-05	0.65%	0.56%	0.46%	0.38%	1.07%	0.98%
1E-06	0.42%	0.34%	0.33%	0.26%	0.58%	0.54%
1E-07	0.19%	0.14%	0.14%	0.09%	0.21%	0.21%
1E-08	0.07%	0.04%	0.03%	0.02%	0.07%	0.06%

4.2. Experiment 2: Algorithm B comparison

The second experiment was performed using Algorithm B, to test transferability of the results to other traffic assignment algorithms. The results are summarized in Table 6, which compares the average values of each metric between TAPAS and Algorithm B.

The full data from these results are shown in Tables A2 (raw data for Algorithm B) and A3 (for a side-by-side comparison). The path-based metrics are not computed or compared in this experiment, since Algorithm B does not attempt to provide a most likely path flow solution, and therefore its path flow results cannot be fairly compared to those of TAPAS (and indeed should not be used in practice, as with any other path flow solution which does not have high entropy).

The trends are very similar between the two algorithms, and the values of each metric are always of the same order of magnitude, and almost always nearly identical numerically. This finding is encouraging, suggesting that the conclusions of Experiment 1 are applicable to other algorithms, and that the relative gap is a good universal measure of convergence, regardless of the specific assignment algorithm.

4.3. Experiment 3: Heterogeneous driver results

The third experiment divided the travel demand into two groups with different values of time, introducing tolls on 10% of the network links. Table 7 compares the values of each metric between the base case (Experiment 1) and this two-class setting. Raw data is shown in Table A4. Since these experiments were performed using our implementation of Algorithm B, path-based metrics are not computed or compared for the same reasons as in Experiment 2.

All three metrics behave extremely similar to single-class Algorithm B experiment metric behavior, indicating that the presence of multiple user classes does not significantly affect the convergence rates of these metrics.

4.4. Experiment 4: Network design application results

Our fourth experiment investigated the effects of subproblem precision in the network design problem, a bilevel program. In this experiment set, we varied the *RG* threshold used

Table 7. Metric stabilization behavior data (average) for single-class and multi-class Algorithm B.

Gap Level	ΔTSTT		ΔVMT		PUL	
	AlgB	Multiclass	AlgB	Multiclass	AlgB	Multiclass
1E-03	2.38%	2.37%	1.73%	1.68%	7.57%	7.27%
1E-04	1.13%	1.01%	0.83%	0.77%	3.57%	3.37%
1E-05	0.56%	0.52%	0.38%	0.37%	0.98%	0.96%
1E-06	0.34%	0.32%	0.26%	0.26%	0.54%	0.52%
1E-07	0.14%	0.12%	0.09%	0.09%	0.21%	0.19%
1E-08	0.04%	0.04%	0.02%	0.03%	0.06%	0.06%

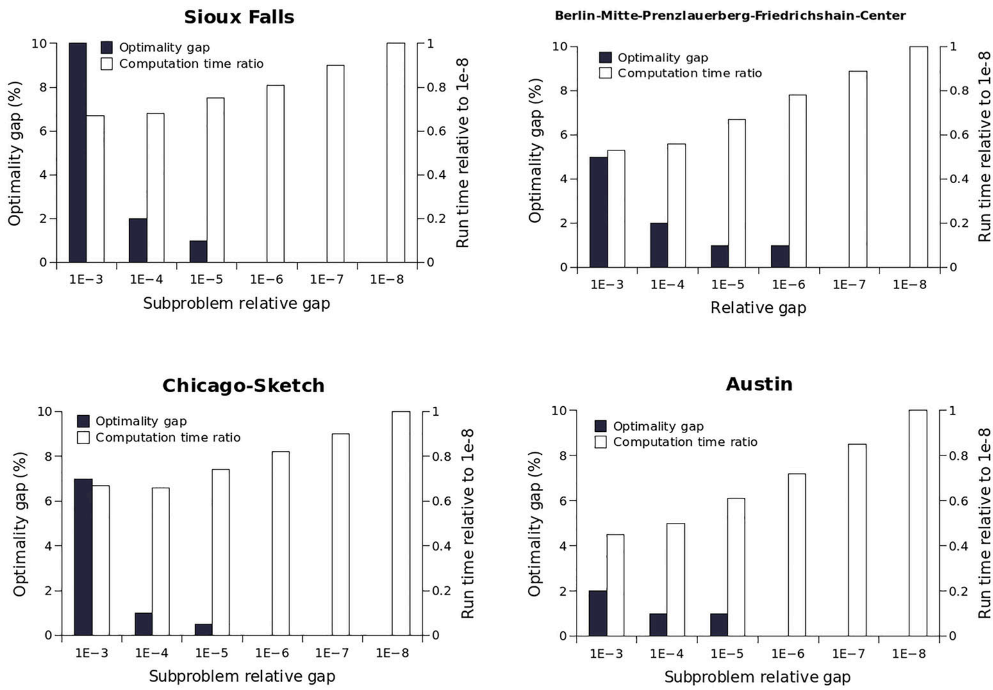

Figure 11. Network design performance with varying RG levels.

in the TAP solutions used in the lower level of this optimization problem, ranging from 10^{-3} to 10^{-8}), for minimizing TSTT. The resulting solutions at the end of the heuristic were then evaluated to a gap of 10^{-8} to compare their performance with a 'benchmark' solution to the network design problem with solved all of its subproblems to a gap of 10^{-8}.

Figure 11 shows the gap between the objective function values with the subproblems solved at a looser gap to those with 10^{-8} (measured by the percentage difference), and the computation times (reported as the fraction of time taken when solving all subproblems to 10^{-8}). The plotted values are averaged over five solutions of the genetic algorithm, which operates randomly.

Table 8. Regression of RG' and AEC with RG.

	RG'			AEC		
Network	Coefficient	Intercept	R^2	Coefficient	Intercept	R^2
Sioux Falls	0.716	0.000	0.999	22.301	0.000	0.999
Anaheim	1.023	0.000	0.999	6.923	0.000	0.999
Chicago Sketch	0.968	0.000	0.999	27.565	0.000	0.999
Barcelona	0.913	0.000	0.999	8.935	0.000	0.995
Austin	1.008	0.000	0.999	7.819	0.000	0.998
Philadelphia	0.974	0.000	0.999	18.276	0.000	0.999

For the higher convergence levels, there was no objective function gap, because the best-found solutions involved expanding the same set of links as in the solution for a gap of 10^{-8}. At these gap levels (around 10^{-6} or 10^{-7}), there was no advantage in solving the subproblems further. When the subproblems are solved to a relative gap of 10^{-4} or tighter, the objective function was within 2% of the benchmark value, and run times were decreased by 40–60%. This may be acceptable in certain applications, given the uncertainty in other components of the planning process (model specification, demand forecasting, etc.).

4.5. *Experiment 5: Gap function comparison*

The fifth experiment set compared the values of three gap functions for the solutions obtained in the previous experiments: RG, RG', and AEC. Linear regressions were performed on RG' vs. RG, and AEC vs. RG, with the results shown in Table 8.

The two definitions of relative gap (RG' and RG) are nearly identical, as shown by R^2 values greater than 0.999, an intercept of essentially zero, and a coefficient close to one. This confirms the analysis at the end of Section 2, and suggests that our conclusions can be equally applied regardless of which relative gap definition is being used.

We also observe that the ratio between AEC and RG is essentially constant within each network ($R^2 > 0.99$ and essentially zero intercept). As expected, this constant differs by network, as it reflects the average travel time on the shortest path available to travelers. For the sizes of networks used in common practice, and for the common choice of minutes as the unit for travel time, we see that AEC is roughly an order of magnitude larger than RG. This suggests that our conclusions can be readily transferred to the AEC gap measure by translating them accordingly.

5. Conclusions

We studied the convergence rate of five metrics as the relative gap reduces over successive iterations of traffic assignment, in twelve networks of varying size and congestion levels. Across these networks, we observed trends for network metric behavior which are summarized below:

- The aggregate metrics (total system travel time and vehicle-miles traveled) were within 1% of their equilibrium values once the relative gap was below 10^{-4} (earlier for smaller networks)

- Link flows achieved stability (less than 1% of the links more than 1% away from equilibrium values) at a relative gap of 10^{-5}.
- Path flows and the sets of used paths stabilized later, at a relative gap of 10^{-6}.
- The above conclusions were seen whether TAPAS or Algorithm B was used to solve for equilibrium, and for both single and two-class assignments.
- In the network design problem, solving the subproblems to a gap level of 10^{-4} instead of 10^{-8} increased the objective function value by less than 2%, but decreased computation time by 40–60%.
- There are strong linear relationship between RG, RG' and AEC ($R^2 > 0.99$). This indicates the transferability of results between different gap metrics: RG and RG' can essentially be used interchangeably, whereas AEC differs from RG by a constant multiple representing average travel time.

The main limitations of the work are (1) that we propose no underlying theory to explain these findings, but present the analysis empirically; and (2) that we restrict our investigation to absolute levels of accuracy, as if a point prediction were sought for a single scenario in isolation. It is also unclear whether this guidance can be generalized to other traffic models, such as dynamic traffic assignment.

Future research should address all of these issues. In particular, regarding (1), the consistent convergence trends across very different networks (spanning several orders of magnitude in both size and congestion level) suggest that there may be a more fundamental relationship between relative gap and these metrics. It may be possible to derive analytical relationships describing such a relationship, at least in stylized settings that roughly approximate practical traffic networks. While the current empirical results span a variety of network sizes and congestion levels to provide meaningful trends and useful guidelines to practitioners and researchers, theoretical bounds shall help generalize the findings of this study.

Regarding (2), another common application involves comparison of multiple alternatives or scenarios, where it is important to determine a stable ranking (or at least a preferred alternative). It would be valuable to see what gap levels are needed before project rankings become stable, although such a gap level would depend critically on how distinct the project impacts are, and a careful investigation is needed to account for this factor.

Note

1. This is the commonly used Bureau of Public Roads function with standard values for its shape parameters Bureau of Public Roads (1964).

Acknowledgments

This research was partly supported by the National Science Foundation (Grant Nos. 1254921 and 1562291) and the Data-Supported Transportation Operations and Planning Center. The authors would like to thank Rishabh Thakkar for his help with multi-class TAP code.

Data availability statement

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was partly supported by the National Science Foundation [grant number 1254921] and [grant number 1562291] and the Data-Supported Transportation Operations and Planning Center.

References

Astroza, Sebastian, Priyadarshan N. Patil, Katherine I. Smith, and Chandra R. Bhat. 2017. "Transportation Planning to Accommodate Needs of Wind Energy Projects." *Transportation Research Record: Journal of the Transportation Research Board* 2669 (1): 10–18.

Babazadeh, Abbas, Babak Javani, Guido Gentile, and Michael Florian. 2020. "Reduced Gradient Algorithm for User Equilibrium Traffic Assignment Problem." *Transportmetrica A: Transport Science* 16 (3): 1111–1135.

Bar-Gera, Hillel. 2002. "Origin-based Algorithm for the Traffic Assignment Problem." *Transportation Science* 36 (4): 398–417.

Bar-Gera, Hillel. 2006. "Primal Method for Determining the Most Likely Route Flows in Large Road Networks." *Transportation Science* 40 (3): 269–286.

Bar-Gera, Hillel. 2010. "Traffic Assignment by Paired Alternative Segments." *Transportation Research Part B: Methodological* 44 (8–9): 1022–1046.

Bar-Gera, Hillel, and Amos Luzon. 2007. "Differences Among Route Flow Solutions for the User-equilibrium Traffic Assignment Problem." *Journal of Transportation Engineering* 133 (4): 232–239.

Beckmann, Martin, Charles B. McGuire, and Christopher B. Winsten. 1956. *Studies in the Economics of Transportation*. Technical Report RM-1488-PR. RAND Corporation.

Bliemer, Michiel C. J., and Mark P. H. Raadsen. 2020. "Static Traffic Assignment with Residual Queues and Spillback." *Transportation Research Part B: Methodological* 132: 303–319.

Bliemer, Michiel C. J., Mark P. H. Raadsen, Luuk J. N. Brederode, Michael G. H. Bell, Luc J. J. Wismans, and Mike J. Smith. 2017. "Genetics of Traffic Assignment Models for Strategic Transport Planning." *Transport Reviews* 37 (1): 56–78.

Boyce, David, Biljana Ralevic-Dekic, and Hillel Bar-Gera. 2004. "Convergence of Traffic Assignments: How Much is Enough?" *Journal of Transportation Engineering* 130 (1): 49–55.

Boyles, Stephen D. 2012. "Bush-based Sensitivity Analysis for Approximating Subnetwork Diversion." *Transportation Research Part B: Methodological* 46 (1): 139–155.

Boyles, Stephen D. 2019. "tap-b Implementation." Accessed May 25, 2019. https://github.com/spartalab/tap-b

Boyles, Stephen D., Nicholas E. Lownes, and A. Unnikrishnan. 2020. *Transportation Network Analysis*. Vol. 1.

Boyles, Stephen, Priyadarshan Patil, Venktesh Pandey, and Cesar Yahia. 2018. *Beyond Political Boundaries: Constructing Network Models for Megaregion Planning*. Technical Report CM2-11. The consortium of Cooperative Mobility for Competitive Megaregions (CM2).

Boyles, Stephen D., and Natalia Ruiz Juri. 2019. "Queue Spillback and Demand Uncertainty in Dynamic Network Loading." *Transportation Research Record: Journal of the Transportation Research Board*2673 (2): 38–48.

Bureau of Public Roads. 1964. *Traffic Assignment Manual for Application with a Large, High Speed Computer*. Vol. 37. US Department of Commerce, Bureau of Public Roads, Office of Planning.

Bureau of Transportation Statistics. 2015. "Travel Time Index." Accessed June 25, 2019. https://www.bts.gov/content/travel-time-index

California Department of Transportation. 2015. "CalTrans Strategic Management Plan 2015–2020." Accessed May 25, 2019. https://www.dot.ca.gov/perf/library/pdf/Caltrans_Strategic_Mgmt_Plan_033015.pdf

Caliper Corporation. 2018. *TransCAD Version 8 User Guide*. Newton, MA: Caliper Corporation.

Cherlow, Jay R. 1981. "Measuring Values of Travel Time Savings." *Journal of Consumer Research* 7 (4): 360–371.

Daneva, Maria, and Per Olov Lindberg. 2003. "A Conjugate Direction Frank-Wolfe Method with Applications to the Traffic Assignment Problem." In *Operations Research Proceedings 2002*, 133–138. Springer.

Daniels, Ginger, David R. Ellis, and Win R. Stockton. 1999. *Techniques for Manually Estimating Road User Costs Associated with Construction Projects*. Vol. 3. College Station, TX: Texas Transportation Institute.

Dial, Robert B. 2006. "A Path-based User-equilibrium Traffic Assignment Algorithm that Obviates Path Storage and Enumeration." *Transportation Research Part B: Methodological* 40 (10): 917–936.

Duthie, Jennifer C., Avinash Unnikrishnan, and S. Travis Waller. 2011. "Influence of Demand Uncertainty and Correlations on Traffic Predictions and Decisions." *Computer-Aided Civil and Infrastructure Engineering* 26 (1): 16–29.

Florian, Michael, Isabelle Constantin, and Dan Florian. 2009. "A New Look At Projected Gradient Method for Equilibrium Assignment." *Transportation Research Record: Journal of the Transportation Research Board* 2090 (1): 10–16.

Frank, Marguerite, and Philip Wolfe. 1956. "An Algorithm for Quadratic Programming." *Naval Research Logistics Quarterly* 3 (1–2): 95–110.

Galligari, Alessandro, and Marco Sciandrone. 2019. "A Computational Study of Path-based Methods for Optimal Traffic Assignment with Both Inelastic and Elastic Demand." *Computers & Operations Research* 103: 158–166.

Gentile, Guido. 2014. "Local User Cost Equilibrium: A Bush-based Algorithm for Traffic Assignment." *Transportmetrica A: Transport Science* 10 (1): 15–54.

Harrison, R., M. Schofield, L. Loftus-Otway, D. Middleton, and J. West. 2006. *Freight Performance Measures Guide*. Project Progress Report no. 0-5410-P3. Texas Department of Transportation.

Higgins, S. F. 2013. "Estimating Economic Impacts from Transportation Investments Using the Texas Statewide Analysis Model and TREDIS." BS thesis, The University of Texas at Austin, Austin, TX.

Holguín-Veras, José, Coral A. Torres Cruz, and Xuegang Ban. 2013. "On the Comparative Performance of Urban Delivery Vehicle Classes." *Transportmetrica A: Transport Science* 9 (1): 50–73.

Jafari, Ehsan, Venktesh Pandey, and Stephen D. Boyles. 2017. "A Decomposition Approach to the Static Traffic Assignment Problem." *Transportation Research Part B: Methodological* 105: 270–296.

Jayakrishnan, R., Wei T. Tsai, Joseph N. Prashker, and Subodh Rajadhyaksha. 1994. *A Faster Path-Based Algorithm for Traffic Assignment*. Technical Report. University of California Transportation Center.

Josefsson, Magnus, and Michael Patriksson. 2007. "Sensitivity Analysis of Separable Traffic Equilibrium Equilibria with Application to Bilevel Optimization in Network Design." *Transportation Research Part B: Methodological* 41 (1): 4–31.

Lessan, Javad, and Liping Fu. 2019. "Credit-and Permit-based Travel Demand Management State-of-the-Art Methodological Advances." *Transportmetrica A: Transport Science* 1–24. https://doi.org/10.1080/23249935.2019.1692963.

Litman, Todd. 2016. *Well Measured*. Victoria, Boritish Columbia: Victoria Transport Policy Institute.

Lundgren, Jan T., and Anders Peterson. 2008. "A Heuristic for the Bilevel Origin–Destination-matrix Estimation Problem." *Transportation Research Part B: Methodological* 42 (4): 339–354.

Maryland Department of Transportation. 2018. "Maryland Department of Transportation Strategic Plan." Accessed May 25, 2019. https://www.roads.maryland.gov/OPPEN/2018_MDOT_TSMO_Strategic_Plan.pdf

Mitradjieva, Maria, and Per Olov Lindberg. 2013. "The Stiff is Moving – Conjugate Direction Frank-Wolfe Methods with Applications to Traffic Assignment." *Transportation Science* 47 (2): 280–293.

Moudon, Anne Vernez, and Orion Stewart. 2013. "Tools for Estimating VMT Reductions from Built Environment Changes." Washington State Department of Transportation, Olympia, Washington.

Nie, Yu Marco. 2010. "A Class of Bush-based Algorithms for the Traffic Assignment Problem." *Transportation Research Part B: Methodological* 44 (1): 73–89.

North Carolina Department of Transportation. 2015. "North Carolina Department of Transportation Strategic Plan F.Y. 2018–2022." Accessed May 25, 2019. https://www.ncdot.gov/about-us/our-mission/Performance/Documents/StrategicPlan_2015_2017.pdf

Patriksson, Michael. 1994. *The Traffic Assignment Problem: Models and Methods*. Mineola, New York: VSP.
Perederieieva, Olga, Matthias Ehrgott, Andrea Raith, and Judith Y. T. Wang. 2015. "A Framework for and Empirical Study of Algorithms for Traffic Assignment." *Computers & Operations Research* 54: 90–107.
Qian, Z., and M. H. Zhang. 2012. "Full Closure Or Partial Closure? Evaluation of Construction Plans for the I-5 Closure in Downtown Sacramento." *Journal of Transportation Engineering* 139 (3): 273–286.
Remi-Omosowon, Ayodeji. 2020. "Pyeasyga Repository." Accessed May 25, 2020. https://github.com/remiomosowon/pyeasyga
Rose, Geoffrey, Mark S. Daskin, and Frank S. Koppelman. 1988. "An Examination of Convergence Error in Equilibrium Traffic Assignment Models." *Transportation Research Part B: Methodological* 22 (4): 261–274.
Rossi, Thomas F., Sue McNeil, and Chris Hendrickson. 1989. "Entropy Model for Consistent Impact-Fee Assessment." *Journal of Urban Planning and Development* 115 (2): 51–63.
Schneck, Arne, and Klaus Nökel. 2020. "Accelerating Traffic Assignment with Customizable Contraction Hierarchies." *Transportation Research Record: Journal of the Transportation Research Board* 2674 (2): 0361198119898455.
Seattle Department of Transportation. 2016. "Seattle Department of Transportation Strategic Plan F.Y. 2018–2022." Accessed May 25, 2019. https://www.seattle.gov/Documents/Departments/SDOT/TechnologyProgram/ITSStrategicPlan20102020.pdf
Shayanfar, Elham, and Paul Schonfeld. 2019. "Selecting and Scheduling Interrelated Road Projects with Uncertain Demand." *Transportmetrica A: Transport Science* 15 (2): 1712–1733.
Stabler, Ben. 2019. "Transportation Networks." Accessed May 25, 2019. https://github.com/bstabler/TransportationNetworks
Ukkusuri, Satish V., Tom V. Mathew, and S. Travis Waller. 2007. "Robust Transportation Network Design Under Demand Uncertainty." *Computer-Aided Civil and Infrastructure Engineering* 22 (1): 6–18.
US Department of Transportation. 2016. "US Department of Transportation Value of Time Memorandum." Accessed May 25, 2019. https://www.transportation.gov/sites/dot.gov/files/docs/Value%20of%20Travel%20Time%20Memorandum.pdf
US Department of Transportation. 2017. "US Department of Transportation Strategic Plan F.Y. 2018–2022." Accessed May 25, 2019. https://www.transportation.gov/sites/dot.gov/files/docs/mission/administrations/office-policy/304866/dot-strategic-planfy2018-2022508.pdf
Waller, S. Travis, Joseph L. Schofer, and Athanasios K. Ziliaskopoulos. 2001. "Evaluation with Traffic Assignment Under Demand Uncertainty." *Transportation Research Record: Journal of the Transportation Research Board* 1771 (1): 69–74.
Weisbrod, Glen. 2008. "Models to Predict the Economic Development Impact of Transportation Projects: Historical Experience and New Applications." *The Annals of Regional Science* 42 (3): 519–543.
Xie, Jun, and Chi Xie. 2015. "Origin-based Algorithms for Traffic Assignment: Algorithmic Structure, Complexity Analysis, and Convergence Performance." *Transportation Research Record: Journal of the Transportation Research Board* 2498 (1): 46–55.
Yang, Hai. 1995. "Heuristic Algorithms for the Bilevel Origin-destination Matrix Estimation Problem." *Transportation Research Part B: Methodological* 29 (4): 231–242.
Yang, Hai. 1997. "Sensitivity Analysis for the Elastic-demand Network Equilibrium Problem with Applications." *Transportation Research Part B: Methodological* 31 (1): 55–70.
Yang, Hai, and Michael G. H. Bell. 1998. "Models and Algorithms for Road Network Design: A Review and Some New Developments." *Transport Reviews* 18 (3): 257–278.
Yang, Hai, and William H. K. Lam. 1996. "Optimal Road Tolls Under Conditions of Queueing and Congestion." *Transportation Research Part A: Policy and Practice* 30 (5): 319–332.
Yin, Yafeng. 2000. "Genetic-Algorithms-Based Approach for Bilevel Programming Models." *Journal of Transportation Engineering* 126 (2): 115–120.
Zhao, Yong, and Kara Maria Kockelman. 2002. "The Propagation of Uncertainty Through Travel Demand Models: An Exploratory Analysis." *The Annals of Regional Science* 36 (1): 145–163.

Appendix. Raw data

Table A1. Metric stabilization behavior using TAPAS.

Sioux Falls					
Gap Level	ΔTSTT	ΔVMT	PUL	PUP	PSD
1E-03	2.02%	1.47%	5.26%	6.08%	8.10%
1E-04	0.81%	0.59%	2.63%	2.03%	4.05%
1E-05	0.20%	0.15%	1.31%	1.52%	3.04%
1E-06	0.10%	0.07%	0.00%	0.91%	1.01%
1E-07	0.05%	0.04%	0.00%	0.23%	0.13%
1E-08	0.04%	0.03%	0.00%	0.05%	0.06%

Berlin-Mitte-Prenzlauerberg-Friedrichshain-Center					
Gap Level	ΔTSTT	ΔVMT	PUL	PUP	PSD
1E-03	2.15%	1.58%	6.87%	10.13%	9.11%
1E-04	1.06%	0.63%	2.22%	4.05%	3.04%
1E-05	0.41%	0.24%	0.81%	1.01%	1.25%
1E-06	0.26%	0.18%	0.43%	0.41%	0.51%
1E-07	0.10%	0.09%	0.12%	0.30%	0.42%
1E-08	0.05%	0.02%	0.06%	0.10%	0.11%

Austin					
Gap Level	ΔTSTT	ΔVMT	PUL	PUP	PSD
1E-03	2.84%	2.51%	11.14%	22.28%	16.20%
1E-04	1.92%	1.25%	5.87%	6.89%	8.10%
1E-05	1.01%	0.65%	1.11%	3.24%	5.06%
1E-06	0.61%	0.49%	0.81%	1.01%	1.42%
1E-07	0.20%	0.20%	0.20%	0.30%	0.51%
1E-08	0.05%	0.04%	0.08%	0.10%	0.20%

Eastern Massachusetts					
Gap Level	ΔTSTT	ΔVMT	PUL	PUP	PSD
1E-03	1.72%	1.34%	5.06%	7.09%	8.75%
1E-04	0.76%	0.46%	2.84%	4.05%	4.36%
1E-05	0.30%	0.10%	0.91%	3.04%	2.82%
1E-06	0.20%	0.07%	0.71%	1.01%	0.74%
1E-07	0.05%	0.03%	0.41%	0.76%	0.42%
1E-08	0.03%	0.01%	0.04%	0.10%	0.08%

Barcelona					
Gap Level	ΔTSTT	ΔVMT	PUL	PUP	PSD
1E-03	2.03%	1.74%	6.08%	14.18%	11.14%
1E-04	1.01%	1.02%	2.85%	6.08%	5.06%
1E-05	0.52%	0.47%	1.01%	3.04%	2.03%
1E-06	0.24%	0.36%	0.73%	1.34%	1.01%
1E-07	0.13%	0.11%	0.32%	0.68%	0.51%
1E-08	0.02%	0.03%	0.08%	0.10%	0.20%

Berlin Center					
Gap Level	ΔTSTT	ΔVMT	PUL	PUP	PSD
1E-03	2.98%	2.60%	12.15%	32.40%	24.30%
1E-04	2.02%	1.31%	5.47%	15.19%	11.14%
1E-05	1.13%	0.73%	1.42%	5.06%	5.06%
1E-06	0.77%	0.52%	0.91%	0.91%	1.11%
1E-07	0.22%	0.18%	0.41%	0.61%	0.76%
1E-08	0.09%	0.04%	0.10%	0.11%	0.56%

Anaheim					
Gap Level	ΔTSTT	ΔVMT	PUL	PUP	PSD
1E-03	2.53%	1.44%	4.89%	6.69%	9.11%
1E-04	0.81%	0.61%	2.61%	3.04%	5.06%
1E-05	0.41%	0.30%	0.77%	1.77%	2.03%
1E-06	0.26%	0.19%	0.23%	0.41%	0.51%
1E-07	0.10%	0.08%	0.04%	0.30%	0.32%
1E-08	0.05%	0.01%	0.02%	0.10%	0.17%

Winnipeg					
Gap Level	ΔTSTT	ΔVMT	PUL	PUP	PSD
1E-03	3.04%	2.03%	6.12%	8.67%	12.15%
1E-04	1.01%	1.01%	2.45%	3.42%	5.23%
1E-05	0.53%	0.35%	0.97%	1.12%	2.11%
1E-06	0.29%	0.33%	0.53%	0.67%	0.97%
1E-07	0.12%	0.11%	0.18%	0.24%	0.39%
1E-08	0.02%	0.03%	0.05%	0.08%	0.04%

Chicago-Regional					
Gap Level	ΔTSTT	ΔVMT	PUL	PUP	PSD
1E-03	2.63%	2.43%	13.16%	28.35%	22.28%
1E-04	1.62%	1.11%	6.08%	12.15%	10.13%
1E-05	0.91%	0.61%	1.01%	5.26%	4.05%
1E-06	0.76%	0.55%	0.81%	1.03%	1.01%
1E-07	0.50%	0.23%	0.00%	0.54%	0.72%
1E-08	0.15%	0.10%	0.00%	0.04%	0.25%

Chicago Sketch					
Gap Level	ΔTSTT	ΔVMT	PUL	PUP	PSD
1E-03	2.03%	1.01%	5.53%	18.23%	15.19%
1E-04	1.01%	0.51%	2.86%	6.08%	7.09%
1E-05	0.37%	0.25%	1.01%	2.03%	2.03%
1E-06	0.28%	0.09%	0.34%	1.01%	0.89%
1E-07	0.08%	0.04%	0.20%	0.77%	0.51%
1E-08	0.04%	0.02%	0.08%	0.10%	0.20%

Terrassa					
Gap Level	ΔTSTT	ΔVMT	PUL	PUP	PSD
1E-03	2.70%	2.23%	7.09%	14.18%	18.23%
1E-04	1.16%	1.01%	3.04%	6.08%	10.13%
1E-05	0.52%	0.39%	1.01%	4.05%	4.05%
1E-06	0.29%	0.38%	0.51%	1.76%	1.46%
1E-07	0.13%	0.09%	0.20%	0.95%	0.67%
1E-08	0.02%	0.03%	0.08%	0.26%	0.20%

Philadelphia					
Gap Level	ΔTSTT	ΔVMT	PUL	PUP	PSD
1E-03	4.05%	2.84%	18.23%	48.60%	30.38%
1E-04	1.92%	1.52%	8.10%	17.21%	18.23%
1E-05	1.52%	1.27%	1.52%	4.05%	8.10%
1E-06	1.01%	0.76%	0.91%	2.03%	2.03%
1E-07	0.66%	0.46%	0.49%	1.01%	1.01%
1E-08	0.30%	0.05%	0.20%	0.20%	0.52%

Table A2. Metric stabilization behavior data using Algorithm B.

Sioux Falls			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.12%	1.52%	5.26%
1E-04	0.90%	0.66%	2.63%
1E-05	0.35%	0.18%	1.31%
1E-06	0.07%	0.03%	0.00%
1E-07	0.02%	0.01%	0.00%
1E-08	0.00%	0.00%	0.00%

Berlin-Mitte-Prenzlauerberg-Friedrichshain-Center			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.03%	1.45%	6.39%
1E-04	0.98%	0.58%	1.93%
1E-05	0.37%	0.23%	0.77%
1E-06	0.23%	0.17%	0.43%
1E-07	0.09%	0.08%	0.10%
1E-08	0.05%	0.02%	0.05%

Austin			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.39%	2.18%	9.42%
1E-04	1.60%	1.09%	5.50%
1E-05	0.85%	0.58%	0.92%
1E-06	0.58%	0.40%	0.71%
1E-07	0.20%	0.19%	0.19%
1E-08	0.05%	0.04%	0.08%

Eastern Massachusetts			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	1.80%	1.28%	4.59%
1E-04	0.77%	0.43%	2.64%
1E-05	0.29%	0.10%	0.86%
1E-06	0.20%	0.06%	0.73%
1E-07	0.05%	0.03%	0.42%
1E-08	0.03%	0.01%	0.04%

Barcelona			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	1.79%	1.64%	5.63%
1E-04	0.94%	0.94%	2.55%
1E-05	0.44%	0.43%	0.91%
1E-06	0.23%	0.31%	0.66%
1E-07	0.12%	0.10%	0.28%
1E-08	0.02%	0.03%	0.08%

Berlin Center			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.57%	2.29%	10.77%
1E-04	1.90%	1.12%	4.71%
1E-05	0.99%	0.65%	1.21%
1E-06	0.63%	0.45%	0.79%
1E-07	0.19%	0.16%	0.33%
1E-08	0.08%	0.03%	0.09%

Anaheim			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.49%	1.40%	4.85%
1E-04	0.78%	0.56%	2.63%
1E-05	0.40%	0.30%	0.71%
1E-06	0.25%	0.19%	0.24%
1E-07	0.10%	0.08%	0.04%
1E-08	0.05%	0.01%	0.02%

Winnipeg			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.83%	1.84%	5.47%
1E-04	0.95%	1.01%	2.19%
1E-05	0.47%	0.31%	0.92%
1E-06	0.26%	0.30%	0.52%
1E-07	0.11%	0.10%	0.16%
1E-08	0.02%	0.03%	0.05%

Chicago-Regional			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.21%	1.54%	11.51%
1E-04	1.24%	0.86%	5.12%
1E-05	0.62%	0.43%	0.95%
1E-06	0.47%	0.35%	0.76%
1E-07	0.18%	0.08%	0.28%
1E-08	0.04%	0.02%	0.04%

Chicago Sketch			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	1.91%	0.94%	4.84%
1E-04	0.91%	0.45%	2.53%
1E-05	0.37%	0.25%	0.97%
1E-06	0.29%	0.08%	0.32%
1E-07	0.08%	0.04%	0.19%
1E-08	0.04%	0.02%	0.07%

Terrassa			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.56%	2.16%	6.49%
1E-04	0.99%	0.98%	2.82%
1E-05	0.52%	0.37%	1.00%
1E-06	0.29%	0.37%	0.48%
1E-07	0.13%	0.08%	0.20%
1E-08	0.02%	0.03%	0.08%

Philadelphia			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	3.85%	2.49%	15.61%
1E-04	1.63%	1.22%	7.59%
1E-05	1.05%	0.78%	1.27%
1E-06	0.63%	0.35%	0.82%
1E-07	0.37%	0.14%	0.37%
1E-08	0.10%	0.03%	0.18%

Table A3. Metric stabilization behavior comparison between TAPAS and Algorithm B.

Sioux Falls						
Gap Level	ΔTSTT		ΔVMT		PUL	
	TAPAS	Alg-B	TAPAS	Alg-B	TAPAS	Alg-B
1E-03	2.02%	2.12%	1.47%	1.52%	5.26%	5.26%
1E-04	0.81%	0.90%	0.59%	0.66%	2.63%	2.63%
1E-05	0.20%	0.35%	0.15%	0.18%	1.31%	1.31%
1E-06	0.10%	0.07%	0.07%	0.03%	0.00%	0.00%
1E-07	0.05%	0.02%	0.04%	0.01%	0.00%	0.00%
1E-08	0.04%	0.00%	0.03%	0.00%	0.00%	0.00%

Winnipeg						
Gap Level	ΔTSTT		ΔVMT		PUL	
	TAPAS	Alg-B	TAPAS	Alg-B	TAPAS	Alg-B
1E-03	3.04%	2.83%	2.03%	1.84%	6.12%	5.47%
1E-04	1.01%	0.95%	1.01%	1.01%	2.45%	2.19%
1E-05	0.53%	0.47%	0.35%	0.31%	0.97%	0.92%
1E-06	0.29%	0.26%	0.33%	0.30%	0.53%	0.52%
1E-07	0.12%	0.11%	0.11%	0.10%	0.18%	0.16%
1E-08	0.02%	0.02%	0.03%	0.03%	0.05%	0.05%

Anaheim						
Gap Level	ΔTSTT		ΔVMT		PUL	
	TAPAS	Alg-B	TAPAS	Alg-B	TAPAS	Alg-B
1E-03	2.53%	2.49%	1.44%	1.40%	4.89%	4.85%
1E-04	0.81%	0.78%	0.61%	0.56%	2.61%	2.63%
1E-05	0.41%	0.40%	0.30%	0.30%	0.77%	0.71%
1E-06	0.26%	0.25%	0.19%	0.19%	0.23%	0.24%
1E-07	0.10%	0.10%	0.08%	0.08%	0.04%	0.04%
1E-08	0.05%	0.05%	0.01%	0.01%	0.02%	0.02%

Austin						
Gap Level	ΔTSTT		ΔVMT		PUL	
	TAPAS	Alg-B	TAPAS	Alg-B	TAPAS	Alg-B
1E-03	2.84%	2.39%	2.51%	2.18%	11.14%	9.42%
1E-04	1.92%	1.60%	1.25%	1.09%	5.87%	5.50%
1E-05	1.01%	0.85%	0.65%	0.58%	1.11%	0.92%
1E-06	0.61%	0.58%	0.49%	0.40%	0.81%	0.71%
1E-07	0.20%	0.20%	0.20%	0.19%	0.20%	0.19%
1E-08	0.05%	0.05%	0.04%	0.04%	0.08%	0.08%

Chicago-Sketch						
Gap Level	ΔTSTT		ΔVMT		PUL	
	TAPAS	Alg-B	TAPAS	Alg-B	TAPAS	Alg-B
1E-03	2.03%	1.91%	1.01%	0.94%	5.53%	4.84%
1E-04	1.01%	0.91%	0.51%	0.45%	2.86%	2.53%
1E-05	0.37%	0.37%	0.25%	0.25%	1.01%	0.97%
1E-06	0.28%	0.29%	0.09%	0.08%	0.34%	0.32%
1E-07	0.08%	0.08%	0.04%	0.04%	0.20%	0.19%
1E-08	0.04%	0.04%	0.02%	0.02%	0.08%	0.07%

Philadelphia						
Gap Level	ΔTSTT		ΔVMT		PUL	
	TAPAS	Alg-B	TAPAS	Alg-B	TAPAS	Alg-B
1E-03	4.05%	3.85%	2.84%	2.49%	18.23%	15.61%
1E-04	1.92%	1.63%	1.52%	1.22%	8.10%	7.59%
1E-05	1.52%	1.05%	1.27%	0.78%	1.52%	1.27%
1E-06	1.01%	0.63%	0.76%	0.35%	0.91%	0.82%
1E-07	0.66%	0.37%	0.46%	0.14%	0.49%	0.37%
1E-08	0.30%	0.10%	0.05%	0.03%	0.20%	0.18%

Table A4. Metric stabilization behavior data for multi-class assignment.

Sioux Falls			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.05%	1.36%	3.94%
1E-04	0.87%	0.56%	1.97%
1E-05	0.24%	0.12%	0.98%
1E-06	0.07%	0.03%	0.00%
1E-07	0.03%	0.00%	0.00%
1E-08	0.00%	0.00%	0.00%

Berlin-Mitte-Prenzlauerberg-Friedrichshain-Center			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.08%	1.54%	6.96%
1E-04	1.02%	0.61%	2.17%
1E-05	0.38%	0.24%	0.81%
1E-06	0.25%	0.17%	0.41%
1E-07	0.10%	0.08%	0.12%
1E-08	0.05%	0.02%	0.06%

Austin			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.43%	1.90%	9.32%
1E-04	1.13%	0.86%	4.53%
1E-05	0.76%	0.60%	1.03%
1E-06	0.47%	0.36%	0.62%
1E-07	0.16%	0.12%	0.15%
1E-08	0.04%	0.03%	0.06%

Eastern Massachusetts			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	1.68%	1.24%	4.92%
1E-04	0.76%	0.46%	2.93%
1E-05	0.30%	0.10%	0.90%
1E-06	0.18%	0.05%	0.67%
1E-07	0.04%	0.02%	0.35%
1E-08	0.02%	0.00%	0.03%

Barcelona			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.00%	1.73%	5.78%
1E-04	0.96%	1.04%	2.73%
1E-05	0.51%	0.46%	0.96%
1E-06	0.23%	0.35%	0.71%
1E-07	0.12%	0.11%	0.31%
1E-08	0.02%	0.03%	0.08%

Berlin Center			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.69%	1.93%	9.28%
1E-04	1.22%	0.94%	4.21%
1E-05	0.84%	0.68%	1.19%
1E-06	0.52%	0.41%	0.77%
1E-07	0.19%	0.14%	0.37%
1E-08	0.08%	0.06%	0.09%

Anaheim			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.73%	1.55%	4.83%
1E-04	0.81%	0.64%	2.48%
1E-05	0.41%	0.31%	0.75%
1E-06	0.26%	0.20%	0.23%
1E-07	0.11%	0.08%	0.04%
1E-08	0.05%	0.01%	0.02%

Winnipeg			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	3.05%	1.93%	5.81%
1E-04	1.03%	0.98%	2.41%
1E-05	0.51%	0.33%	0.95%
1E-06	0.29%	0.34%	0.53%
1E-07	0.12%	0.10%	0.17%
1E-08	0.02%	0.03%	0.05%

Chicago-Regional			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.20%	1.53%	10.54%
1E-04	1.15%	0.86%	5.13%
1E-05	0.69%	0.48%	0.77%
1E-06	0.42%	0.31%	0.74%
1E-07	0.13%	0.10%	0.00%
1E-08	0.04%	0.03%	0.00%

Chicago Sketch			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.14%	0.96%	5.68%
1E-04	1.03%	0.52%	2.97%
1E-05	0.39%	0.25%	1.05%
1E-06	0.28%	0.09%	0.35%
1E-07	0.08%	0.04%	0.20%
1E-08	0.04%	0.02%	0.09%

Terrassa			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	2.60%	2.17%	6.86%
1E-04	1.14%	0.97%	2.89%
1E-05	0.52%	0.40%	1.01%
1E-06	0.29%	0.37%	0.51%
1E-07	0.13%	0.09%	0.20%
1E-08	0.02%	0.03%	0.08%

Philadelphia			
Gap Level	ΔTSTT	ΔVMT	PUL
1E-03	3.46%	2.63%	14.76%
1E-04	1.24%	0.98%	6.77%
1E-05	0.80%	0.60%	1.25%
1E-06	0.58%	0.42%	0.84%
1E-07	0.20%	0.16%	0.36%
1E-08	0.11%	0.08%	0.15%