

MODELLING THE PUBLIC TRANSPORT CAPACITY CONSTRAINTS' IMPACT ON PASSENGER PATH CHOICES IN TRANSIT ASSIGNMENT MODELS

Arkadiusz Drabicki¹, Rafal Kucharski², Andrzej Szarata³

^{1,2,3} Cracow University of Technology, Faculty of Civil Engineering, Department of Transportation Systems, Cracow, Poland

¹e-mail: adrabicki@pk.edu.pl

²e-mail: rkucharski@pk.edu.pl

³e-mail: aszarata@pk.edu.pl

Abstract: *The objective of this paper is to discuss the replication of passenger congestion (overcrowding) effects on output path choices in public transport assignment models. Based on a comprehensive literature review, the impact of passenger overcrowding effects was summarised in 3 main categories: the inclusion of physical capacity constraints (limits); the feedback effect between transport demand and supply performance; and the feedback effect on travel cost (discomfort penalty). Further on, sample case studies are presented, which prove that the inclusion of capacity constraints might significantly influence the assignment output and overall results in public transport projects' assessment – yet most state-of-the-practice assignment models would either miss or neglect these overcrowding-induced phenomena.*

In a classical 4-step demand model, their impact on passengers' travelling strategies is often limited to path (route) choice stage, while in reality they also have far-reaching implications for modal choices, temporal choices and long-term demand adaptation processes. This notion has been investigated in numerous research works, leading to different assignment approaches to account for impact of public transport capacity constraints – a simplified, implicit approach (implemented in macroscopic-based models, e.g. PTV VISUM), and a more complex, explicit approach (incorporated in mesoscopic-based models, e.g. BusMezzo). In the simulation part of this paper, sample tests performed on a small-scale network aim to provide a general comparison between these two approaches and arising differences in the assignment output. The implicit approach reveals some differences in assignment output once network capacity constraints are accounted for – though in a simplified manner, and producing somewhat ambiguous output (e.g. in higher congestion scenarios). The explicit approach provides a more accurate representation of overcrowding-induced phenomena - especially the evolving demand-supply interactions in the event of arising congestion in the public transport network. Further studies should involve tests on a city-scale, multimodal transport model, as well as empirical model validation, in order to fully assess the effectiveness of these distinct assignment approaches.

Highlights:

- *The paper discusses the inclusion of overcrowding effects on path choices in public transport assignment models*
- *These can be grouped into 3 main categories: physical constraints, demand-supply feedback and path discomfort cost*
- *Sample case studies show that their inclusion may substantially affect the assignment output*
- *Two general methods of modelling capacity constraints are: the implicit and explicit approach*
- *An illustrative example shows that both approaches produce different output with the explicit one being more specific and adequate*

Key words: *public transport assignment, passenger congestion, overcrowding, crowding discomfort, path choice, public transport capacity.*

1. Introduction

Path choice (or route choice) process comprises a crucial step within every single transport assignment model (Fig. 1). The path (route) choice algorithm is most commonly described by means of the probabilistic, discrete choice model and the random utility theory (Cascetta, 2001). The bottom line is that the probability of choosing a given O-D path is related to the cost-utility formula, which reflects the relative (dis)utility associated with travelling along that path, among all the alternative O-D paths (routes). The path cost formula comprises the following trip components: perceived travel times (i.e. in-vehicle, waiting, walking times), monetary costs (fares), transfer penalties and temporal utilities of earlier (or later) O-D connections – which are described in relative (weighted) terms, reflecting the user perceptions of disutility associated with particular trip stages (e.g. increased disutility associated with waiting and walking times). This path cost evaluation algorithm forms a key component within the classical 4-stage assignment model, where it is applicable at the modal split stage – i.e., used to evaluate the choice probability between the public and private transport modes (Szarata, 2014) – and eventually at the trip assignment stage – i.e., used to compute the choice probability of feasible network paths (routes, lines etc.).

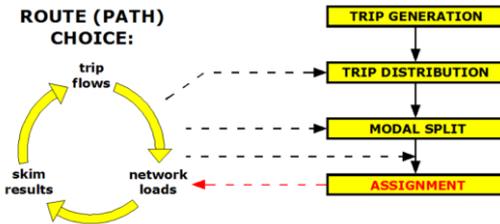


Fig. 1. Path (route) choice process in the 4-stage assignment model (source: Hartl, 2013)

In a summary, this means that the passenger choices in public transport networks are principally a function of journey times and service frequencies – i.e. main factors which are recurrently deemed most important according to the passenger surveys (Rudnicki, 1999). However, a major factor which is either missing or not properly exploited in most state-of-the-practice assignment models, concerns the inclusion of line (service) capacity, as well as the

associated (dis)comfort aspects – which might actually also have a notable effect on passengers' output choices, especially when considering various public transport modes with distinct capacity rates - i.e. mass transit (underground, urban rail) vs. feeder (light rail) systems, or conventional (bus, tram) vs. unconventional modes (monorail) (Drabicki et al., 2016).

The implications of network capacity constraints (limits) are more investigated in case of private transport assignment (Żochowska, 2014), where they are typically included in form of the volume-delay function (VDF). The VDF function describes the effects of increasing travel times as a result of rising traffic flows (volumes) – i.e. a non-linear travel time penalty which increases sharply once traffic volume approaches the saturation flow rate (i.e. the assumed road capacity limit). However, whereas the VDF functions are commonly available and widely applied to replicate the capacity limits in modern-day private transport (PrT) assignment models (Branston, 1976), the incorporation of capacity constraints effects in public transport (PuT) assignment models remains – to the best of our knowledge – much less examined and advanced, usually limited to individual case studies and modelling developments; in practical approach, the implications of capacity constraints on passenger path choices are often neglected in state-of-the-practice modelling algorithms.

The objective of this paper is to contribute to the ongoing research discussion on replicating the overcrowding effects in public transport assignment models. The literature review part of this paper will outline the main aspects of their impact on passenger path choices (which should be accounted for in simulation models) and present sample results from practical transportation studies. Further on, the simulation works on a small-scale network will reveal the arising differences in assignment output between the two common modelling approaches to public transport capacity constraints. Our aim is that the observations and conclusions from this study would illustrate the possibility of reproducing the overcrowding effects in these two main modelling algorithms, provide indications for their application on bigger-scale transport models – and together with a summary of the state-of-the-art in public transport congestion modelling, it would also point out fields for future improvement works.

The remainder of this paper is organised in a following way: section 2 focuses on literature review regarding the incorporation of public transport network capacity constraints and their impact on output passenger path choices. Section 3 highlights the importance of proper appraisal of public transport capacity constraints effects in assignment models, presenting results from sample case studies, where capacity constraints were taken into account and led to substantially different project assessment indicators. Section 4 presents two distinct modelling algorithms of public transport networks, where the influence of capacity constraints can be described in 2 various approaches. These are followed by practical simulations on sample networks in section 5, where both modelling approaches lead to different network performance, and consequently – distinct simulation output. Finally, section 6 provides the summary and conclusions for further research works, and indications for future applications on bigger, city-scale public transport assignment models.

2. Literature review

Substantial amount of research works in recent years has been devoted to the notion of public transport congestion, or more precisely – the passenger overcrowding: i.e., the way it affects the passengers' travelling strategies, user preferences, implications for the transport system performance, the issues of service optimisation etc. - with an ultimate goal of the inclusion of these (often mutually dependent) effects in assignment models. However, though these state-of-the-art assignment models aim to replicate the impact of passenger overcrowding on path choice decision models in a most plausible way

possible, they usually include only some of the overcrowding effects. Consequently, they often do not yield completely realistic results and are likely to underestimate the arising phenomena of passenger overcrowding.

In a general overview, the effects of passenger overcrowding on output path choice decisions – which should be accounted for in a model observing the public transport capacity constraints - can be summarised into three main categories, as listed below and elaborated in subsequent sections:

- physical capacity limits (constraints),
- feedback effect on service performance,
- feedback effect on passenger (dis)comfort.

2.1. Passenger congestion effects in assignment model – impact of physical capacity limits

The first category of public transport congestion effects concerns the direct impact of **physical capacity constraints** – i.e., the maximum permissible flow volume of passengers which can be carried by the components of public transport network within a specified time period. The major factor determining the physical capacity limits is the passenger load capacity of public transport vehicles (the max. no. of passengers able to “get on-board”) and the arising queuing phenomena at stops or platforms – while (Gentile and Noekel, 2015) also suggest that in some cases the finite capacity limits of stops (platforms) themselves – i.e. the space limitations – are also of relevant importance. In recent literature works, the impact of physical capacity limits in public transport assignment has been typically modelled in 2 following ways: by means of the (so-called) implicit or explicit approaches.

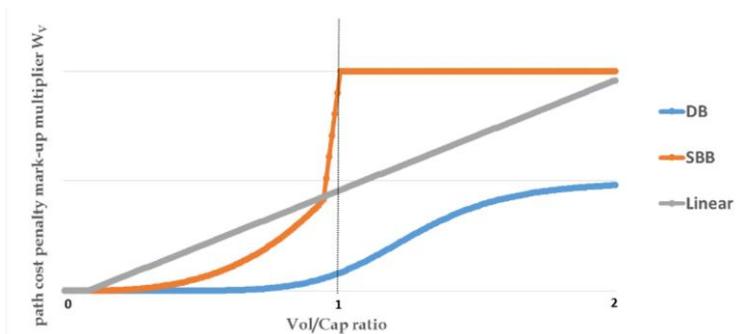


Fig. 2. Crowding (mark-up) cost discomfort functions, available in the implicit approach (based on (PTV VISUM Manual, 2016))

The **implicit approach** to capacity constraints follows the VDF-based method used as a default in most private transport assignment models to represent the congestion effects (described earlier). The passenger flow capacity of network links is not strictly bounded by a fixed limit value, but is instead defined with a non-decreasing, volume-dependent link cost function (Fig. 2). Typically, the function imposes an additional cost penalty above a certain threshold (e.g. the assumed seat capacity), which reflects the rising crowding discomfort. Further on, as passenger volume tends towards the capacity limit, the so-called *crush capacity*, the increase in cost penalty becomes non-linear and very sharp. Once the passenger flow exceeds the nominal line capacity (i.e. calculated with respect to the crush capacity of operating vehicles), travellers are not explicitly prohibited from using the service (vehicle run), but travel cost should have now risen so severely, that any additional passengers should be “discouraged” from boarding it – i.e., an *implicit* capacity limit is imposed upon that particular service (vehicle run). Analogous to the capacity-constrained traffic assignment model, the assignment is calculated in an iterative procedure: in each consecutive simulation run, the output demand flows (i.e. travellers' choices) depend concurrently on network parameters (i.e. travelling conditions) calculated in preceding simulation run. The assignment procedure gradually converges towards a stable solution, and the final output (passenger flows, line loads and travel costs) is obtained once an equilibrium state is achieved.

Typically, the path cost penalty in implicit approach (as e.g. in PTV VISUM model) is described either as a linear function of the volume-to-capacity ratio, or utilises a more nuanced, non-linear correlation as e.g. assumed by the DB and SBB functions (fig. 2) – with the latter solution being perhaps more appropriate, as it allows to account for the non-uniform increase rate in crowding discomfort. The two non-linear crowding cost functions used in the PTV VISUM model are analogous to the approaches used in rail demand modelling in the German railway system – the DB function (*Deutsche Bahn*), and the Swiss railway system - the SBB function (*Schweizerische Bundesbahnen*). In these two functions the path cost penalty due to passenger overcrowding is in general exponentially correlated with the rising volume-to-capacity ratio, with an

upper bound limit of the crowding cost penalty rate – beyond which it converges towards a fixed penalty rate (typically, this would occur once the *crush capacity* limit has been reached).

This forms a simplified method of representing the effects of passenger congestion in public transport assignment, which is usually applied within macroscopic models and available in common transport modelling tools (e.g. (PTV VISUM Manual, 2016)). As shown below on sample case studies, this assignment method enables to replicate some effects on passenger overcrowding on route choices and modal shifts, yet it comprises a rather simplified approach (e.g. by imposing a uniform cost both for travellers on-board and those waiting at the stops), missing the important, evolving congestion phenomena in public transport system.

The **explicit approach** to capacity constraints comprises a more specific (and thus more reliable) representation of public transport supply and its interactions with travel demand. Though it has not been applied yet on a wider scale – being developed mostly in individual algorithms (e.g. the BusMezzo algorithm (Cats, 2011)) and case study applications - its implementation has been hitherto possible in mesoscopic and microscopic assignment models. A more detailed modelling framework implies that the travel demand is represented by individual agents (passengers) progressing through the network, whereas travel supply is represented by individual vehicles (runs) defined with strict capacity limits, corresponding to the crush capacity values. Travellers arriving at the platform (stop) board the incoming vehicle runs according to their residual (available) capacity. If boarding volume exceeds the residual vehicle capacity, the remaining passengers are explicitly denied the boarding and have to wait for next vehicle departures – thus, important queuing phenomena arise at the platforms (stops). The queuing discipline at stops can be commonly reproduced in a number of ways, notably including the following two (Gentile and Noekel, 2016):

- the FIFO principle: “first in, first out” – an organised queuing process, consisting of the undersaturation queue (those who will board the nearest vehicle run) and oversaturation queue (those delayed and “forced” to wait yet for later vehicle runs),
- the so-called *mingling process*: no priority rules are in place, and passengers joining the residual

queue have roughly the same boarding probability as others waiting at the platform.

The resultant fail-to-board probability, which becomes significant as passenger congestion rises, has wider implications on the ensuing passenger path choices (Nuzzolo et al., 2012). Travellers who had to skip previous vehicle runs perceive additional disutility due to the boarding failure – i.e. the arising waiting cost is perceived as relatively more burdensome. Consequently, they may take a rerouting decision and consider other, less attractive O-D travel routes (paths).

2.2. Passenger congestion effects in assignment model – feedback on service performance

The second type of passenger overcrowding effects, resultant from the inclusion of public transport capacity limits, concerns the feedback interaction between the transport supply (service regularity and dwell times) and transport demand (passengers’ decisions and resultant volume flows) performance. A principal reason underlying this interaction is that the dwell-time of a public transport service trip is an increasing function of boarding and alighting passenger volumes. In a summary, the feedback effect demonstrates itself in the following manner: changes in passenger flows cause fluctuations in dwelling times at stops, which will conversely induce variations in service operating times and headway deviations. In turn, as vehicle arrivals (and departures) become irregular, passenger demand is now unevenly distributed among the individual runs – and further on, the feedback effect is amplified. This impedes the service regularity and reliability which is undesirable both for passengers (increasing travel times and crowding levels) as well as for operators (uneven utilisation of service supply).

This important phenomenon, “reinforced” by the arising passenger overcrowding, can only be replicated if the modelling framework allows to describe the impact of demand flows on vehicle dwell times – which is in practice often neglected especially in macroscopic assignment algorithms. The boarding and alighting processes are strictly related to passenger flow vector, which depends on vehicle exchange capacity, and the assumed “dwelling routine” (i.e. separate or mixed doors for boarding and alighting). Based on a wide range of literature sources (summarised by (Tirachini et al., 2013), (Gentile and Noekel, 2016)), it can be concluded that there is a roughly linear correlation between the dwell times and number of alighting (boarding) passengers – the values fall usually within the range of 2-4 secs/pass, though these are likely to increase even further (up to 6 secs/pass and beyond) in overcrowded conditions. (Gentile and Noekel, 2016) provide a detailed mathematical framework for describing the impact of dwelling flows on mean and, crucially, variance values of dwell times and service headways – i.e. the main “trigger” behind this feedback interaction process. Importantly, these time-dependent service variations may initially occur at individual stops or line sections, but will likely become amplified and propagate further downstream in the network.

The feedback loop between transport demand and transport supply performance is probably best manifested in a well-known phenomenon, which occurs in public transport networks during congested conditions – i.e. the so-called **bus bunching** effect (Fig. 3); its other denominations mentioned in literature sources are: bus platooning, clumping, pairing, the banana bus, the Bangkok effect (Moreira-Matias et al., 2012).

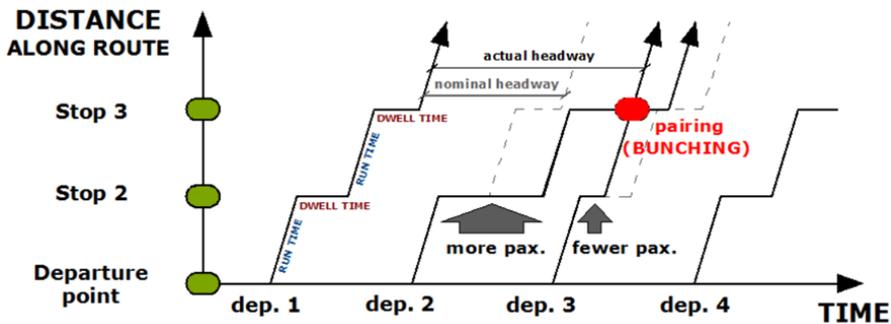


Fig. 3. Bus bunching effect, plotted on the space-time diagram (source: Attanucci, 2010)

The bus bunching effect can be explained intuitively on a space-time diagram (Attanucci, 2010), under the simple assumption of constant (Poisson-distribution based) passenger arrival process at stops, as follows: a certain vehicle run which arrives later than scheduled at the stop has to pick up a higher than average number of waiting passengers. Dwelling time takes longer than expected and once ready to depart, the vehicle is now delayed even further (relative to its nominal timetable). The same pattern will hold at the next downstream stop, where overcrowding conditions will likely become worse, the service delay will rise further, and so on. In contrast, the next (following) vehicle run has less waiting passengers to pick up, dwells shorter at the stop, and as a result, will run ahead of schedule. The relative headway between these 2 consecutive runs will likely decrease as they progress downstream in the network, and the second vehicle run may eventually catch up with the vehicle ahead of it – the stage where vehicle runs become “fully bunched” or paired together, at which the relative headway drops down to zero. The bunching phenomenon leads to substantial impairment in public transport service regularity, since journey times become longer, the waiting times are higher (due to uneven vehicle spacing), and recurrently – the average crowding levels increase due to uneven passenger loads' distribution among the individual vehicle runs.

One of the objectives of research works was to describe the main factors and critical conditions which induce the bus bunching effect. A common conclusion is that passenger demand (volume) has profound impact on service regularity, or more specifically, the resultant loading factor, defined as the ratio of pass. arrival rate (at stop) to pass. loading rate (on-board). (Newell and Potts, 1964) developed (possibly) a first mathematical framework for bunching effect, where they define this correlation by means of a critical bus bunching parameter. It describes transition from stable conditions to a self-reinforcing bunching phenomenon state, at which – if sustained over a longer time period - the buses will fall out of schedule even further. More advanced approaches emphasise the importance of passenger arrival pattern, which need not be always uniformly distributed in time. For example, (Fonzone et al., 2015) demonstrate on the proposed algorithm that various possible arrival patterns would require different critical conditions to trigger the bunching

effect (which could then develop in a substantially distinct degree). (Gentile and Noekel, 2015) propose a bus bunching coefficient variable, defined as a function of service headway between 2 consecutive vehicle runs. The coefficient can be used as a basic measure of arising bus bunching effect in the network, being calculated as the ratio of actual service headway (i.e. one resulting from fluctuations in dwell times) to the nominal scheduled headway – the higher the headway deviation rate, the bigger the on-going bunching effect. An analogous formula can be used to estimate the bunching coefficient at a downstream stop, resulting from current upstream service conditions and passenger dwelling flows. Additionally, research sources mention that the bus bunching effect is not only related to the demand-supply interactions at the stops, but may also be induced (or amplified) by other factors, such as general traffic characteristics, route design, road conditions etc. Numerous analytical models have been developed which allow to demonstrate their impact upon the output service regularity ((e.g. (Bağ, 2010), (Horbachov et al., 2015)) – however, in this paper we will focus primarily on the influence of passenger congestion and the consequent bunching phenomena.

2.3. Passenger congestion effects in assignment model – feedback on passenger discomfort

The third major type of public transport congestion effects concerns the arising discomfort cost and its implications for passengers' travelling choices. Evidence from passenger surveys seems to reinforce the fact that crowding (dis)comfort is among the major factors relevant to users' travel experience – e.g. results from Transport for London's regular monitoring of customer satisfaction (Barry, 2015) indicate that travel comfort and crowding are rated as the (third and fourth) most important issues, right after the journey time and personal safety. Although journey times still form the baseline and most decisive factor in path choice process, the travel discomfort may also contribute its own mark-up “penalty” upon the travel cost. Overcrowding affects the travellers' comfort perception who become more reluctant if their public transport services are routinely congested. (Tirachini et al., 2013) mention a wide range of psychological, sensorial and social factors attributed to the overcrowding effects, such as: risk perception of personal safety, anxiety and

stress, possible ill-health, propensity to arrive late at work, possible loss of productive time.

Commonly, the **crowding discomfort factor** is included as an additional (mark-up) travel time multiplier in the general path cost formula. The relative (perceived) value of travel time components increases as rising passenger numbers (flows) produce a crowding externality (cost), relative to travelling in uncrowded conditions. The discomfort penalty is described as a non-linear, VDF-based function of volume-to-capacity ratio of a given travel alternative, which increases more sharply as crowding conditions deteriorate – the generalised crowding mark-up factor formula (Gentile and Noekel, 2015) is based on the same VDF function as used in the implicit approach to modelling the capacity constraints (described earlier).

A recurring question in literature sources is how to measure precisely the on-board crowding levels, with two basic approaches considered (Tirachini et al., 2013):

- discomfort cost as a function of *load factor* (percentage volume-to-capacity ratio): a simplified measure which can be related to the vehicle seat capacity, or in macroscopic approach – roughly to the generated line capacity – yet it says very little about the actual on-board crowding conditions themselves, which will vary depending on (among others) the vehicle interior arrangement; studies estimate that as such the crowding cost is “activated” from load factors between 60 – 90% onwards (Tirachini et al., 2013), (van Oort et al., 2015),

- discomfort cost as a function of *density of standees* (per square metre): perhaps a more relevant measure, since crowding discomfort becomes much more acute once passenger load surpasses the vehicle seat capacity, and the estimated available space per passenger provides a better picture of the degree of crowding “suffered” by standing travellers; here, the crowding mark-up penalty ranges between 1.0 – 1.6 (for those seated) and 1.5 – 2.4 (for those standing), and applies already if density of standees rises from zero (pax. per sq. m) (Whelan and Crockett, 2009).

The exact values of crowding discomfort factor differ among literature sources, being dependent on the methodology used, local context and user preferences, as well as individual public transport modes and trip characteristics (Tirachini et al., 2013). Literature review shows that crowding discomfort values are likely to be higher in case of rail systems and increase with trip length and duration. The majority of studies which aimed to provide an estimate of crowding discomfort costs on passengers’ choices focused mainly on long-distance urban trips (i.e. between the suburbs and city centre) made with suburban or metro railways (Tirachini et al., 2013), (Kroes et al., 2013), (Whelan and Crockett, 2009), as well as intercity rail trips (Lieberherr and Pritscher, 2012). A meta-study commissioned for the UK Department of Transport (Whelan and Crockett, 2009) provides a comprehensive valuation of overcrowding costs and the willingness-to-pay estimate for trips made in the British Rail system – which are often used as a guideline in transport practice (Fig. 4).

Table 4.2: Crowding Value of Time Multipliers

Load Factor	Sit	Stand	pass/m ²	Sit	Stand
80	1.00	1.50	0	1.00	1.53
100	1.08	1.50	1.0	1.11	1.62
120	1.23	1.67	2.0	1.21	1.70
140	1.38	1.85	3.0	1.32	1.79
160	1.53	2.02	4.0	1.42	1.87
180	1.68	2.20	5.0	1.53	1.96
200	1.83	2.37	6.0	1.63	2.04

Please note: The Load Factor and pass/m² (passengers per square meter) estimates vary by rolling stock type. The rows in this table are therefore **do not** match across different crowding metrics.

Fig. 4. Time cost multiplier factor due to crowding - acc. to the British Rail WTP meta-study (source: Whelan and Crockett, 2009)

The purpose of such modelling framework is to reflect the crowding discomfort impact on a certain share of travellers who would adjust their travel patterns, so as to avoid the worst overcrowding circumstances - and utilise other O-D travel alternatives. This should replicate the long-term adaptation process in travelling strategies, as repeated experience of overcrowding will impact the 3 important aspects:

- path (route) choice: travellers will be less likely to use notoriously overcrowded services and would seek other, perhaps less attractive, public transport connections; in the modelling approach, this could imply a demand shift towards services with higher spare capacity (e.g. mass transit systems), or less-popular public transport connections (e.g. a trade-off in longer journey times combined with less-crowded travel conditions),
- mode choice: as a consequence of routine overcrowding, public transport service would lose on their relative attractiveness, and travellers would likely revert to using private cars; for short-range trips, possibly an increase in walking (or cycling) trips could be observed,
- departure time choice: perhaps the most significant impact of passenger congestion on travel choices - in the day-to-day adaptation process travellers will seek to avoid the time periods of peak congestion, and would utilise the same O-D travel route but at less-popular travel times; the departure-time updating process would imply a higher passenger volume migration especially towards earlier departure runs.

This adaptability phenomenon of passengers' path choice strategies in response to crowding discomfort can be incorporated in the modelling framework by means of a conventional, iterative user equilibrium approach (in a simplified manner), or more reliably - by employing a day-to-day learning mechanism (Nuzzolo et al., 2012). In the latter case, travellers consider on day t the anticipated attribute values of path cost components, which are a weighted average of experienced and anticipated attribute values on day $t-1$ - thus, the path choice model is recurrently updated based on users' expectations and their prior experience:

The extent to which the overcrowding experience impacts the passengers' choices will differ profoundly, depending on the trip purpose (motivation). Literature sources (Tirachini et al.,

2013) and empirical surveys alike (London Assembly Report, 2009a) confirm that crowding discomfort is of limited significance for commuter (obligatory) journeys but might have substantial implications for leisure (non-obligatory) journeys. In former case, the necessity to arrive at destination on-time means that commuter travellers still assign much higher (relative) weight to travel times than on-board conditions, or as given by a cited London commuter (London Assembly Report, 2009b): "*You just have to use the Tube. There's just no choice, there is no option. Well, there is an option: just don't go to work but that's not really an option!*". The same report examines the ways in which commuters adapt to the frequently experienced travel conditions: around 66% of London rail commuters adjusted their departure times (e.g. chose earlier connections), and for one of the rail services ca. 20% of travellers would travel in the opposite direction first, just to have a higher chance of getting a seat at an upstream station. On the other hand, crowding seems to have much more suppressive impact on the non-obligatory trip motivations. The majority of leisure travellers do avoid travelling on London Underground during rush hours, and 25% of them change the time of travel during the day due to anticipated crowding. Additionally, sociodemographic factors themselves might be relevant as well: (Kim et al., 2009) indicate that specific user groups expose different "sensitivity" rates to crowding - e.g. elderly people are likely to sacrifice the extra travel time in favour of more comfortable trip conditions.

3. Appraisal of public transport capacity - sample case studies

Incorporation of passenger overcrowding effects on public transport system has been shown in a number of (both academic and practical) case studies to influence the overall projected network usage, performance results and assessment indicators - yielding distinct results when compared to the analysis "insensitive" to overcrowding phenomena. (Batarce et al., 2015) point out interestingly that (passenger) congestion in public transport plays an analogous role to (traffic) congestion in private transport (Fig. 5): investment in public transport systems increase both their transportation capacity and relative attractiveness, which spurs passenger demand growth. However, in longer run this induces

increase in travelling discomfort due to arising crowding, and the (finite) capacity of public transport supply itself may eventually become outstripped by the ever increasing passenger demand. In the end, this implies that further improvements in public transport systems are necessary - the ramifications of this feedback correlation may only be captured if public transport capacity constraints are taken into account; neglecting it would produce erroneous results in terms of public transport system effectiveness and capability.

Both implicit and explicit approaches to modelling the capacity constraints have been utilised in sample case studies to demonstrate the arising differences in public transport assignment output between congested vs. uncongested cases. In a recent case study for The Hague city (van Oort et al., 2015), an implicit, VDF-based approach revealed differences in passenger flows' distribution between the proposed tram line and the existing bus route along the same transport corridor (Fig. 6). A two-tier crowding mark-up penalty was assigned to the path

cost formula, which reflected first an increasing discomfort penalty due to rising on-board crowding (within the range of 1.0 – 1.7), and after reaching the assumed *crush capacity* it surged rapidly up to the constant value of 10.0. The method revealed a higher patronage rate of the tramway system – the passenger gains could be attributed both to its higher nominal capacity limit as well as better on-board comfort level, when compared to the existing bus system. Importantly, a reduction in service frequency need not necessarily imply a decline in passenger numbers, as envisaged by uncongested model. Inclusion of another principal factor – i.e. increasing service capacity (provided by tramway system) – mitigated these losses and even projected a slightly higher demand flow along the proposed tram line.

Another case study in the city of Stockholm (Cats et al., 2015) utilised a more specific, explicit approach with individually modelled vehicles and travellers (agents) to assess the projected performance of a new metro line proposed along an existing, busy bus corridor.

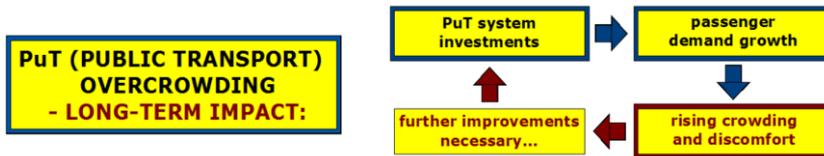


Fig. 5. Public transport (PuT) congestion (overcrowding) - a long-term feedback impact which may not be captured with conventional assignment models (source: Batarce et al., 2015)

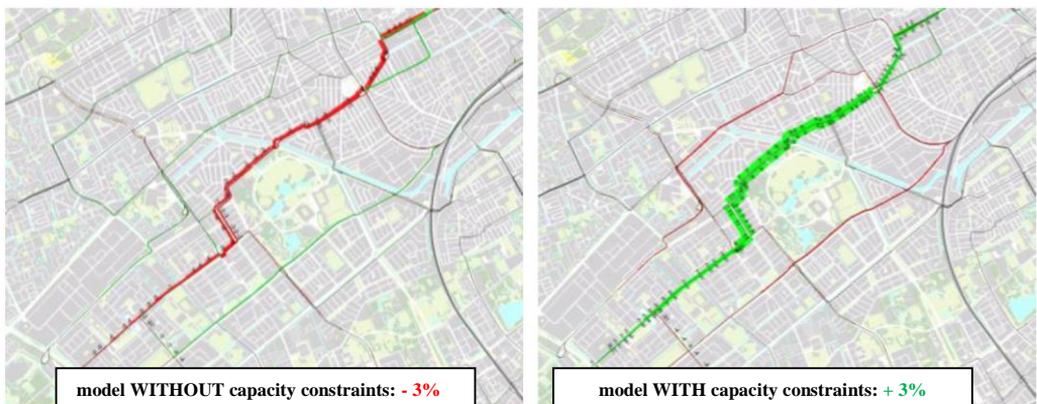


Fig. 6. Sample results of including the capacity constraints' and comfort effects in the implicit approach – estimated relative effect on daily ridership after conversion of bus line 25 to tram line in The Hague city (source: (van Oort et al., 2015))

In the existing scenario, when the busy corridor is served by 200 buses per hour (per direction), ca. 40% of buses are routinely overcrowded, and multiple denial-of-boarding events can be observed. In that case, the explicit approach captures the deteriorations in service quality and travel times, in the form of the bus bunching effect caused by (and correlated with) excessive demand flows – i.e. the very principal ramifications of the mutual demand-supply interactions. In contrast, a new metro line with much higher capacity would attract ca. 60% of bus users, and despite smaller service frequency it would still be less overcrowded and much more resilient to service disruptions. The absolute decrease in in-vehicle and waiting times is ca. 15%, but when weighted in relative (perceived) terms, the project would bring ca. 65% extra benefits, attributable to higher system capacity and reduced discomfort travel cost. In such case, a cost-benefit analysis based on uncongested static model would potentially miss a major share of gains coming from public transport system improvements.

A study for the Swiss railway system (Lieberherr and Pritscher, 2012) developed an implicit, VDF-based capacity restraint model, the so-called SBB crowding function (described in more detail in subsequent chapter), which has been now incorporated in the macroscopic PTV VISUM software. Application in pilot projects showed that the capacity-restraint assignment reduced the overestimation (overload) rate of railway system usage (measured in seat-km) by 30% - though the assignment model would still yield somewhat overestimated passenger flows, the obtained results would be more plausible. Additionally, the SBB crowding function revealed extra shifts from intercity to regional train services during overcrowded peak hours – a minor share of travellers (ca. 3% of total O-D flow) would switch towards slower but less-crowded trains. Furthermore, researchers reckon that a more far-reaching distinction between “seated” and “standing” crowding penalty itself might influence the assignment output. (Leurent, 2009) demonstrate that the predicted passenger load in Paris metro system is reduced by ca. 30%, when a congested model additionally distinguishes between the seated and standing crowding disutility.

Additionally, researchers (Small, 1999) indicate that the benefits of improving the public transport system

capacity may be not only quickly diminished (i.e. “eaten-up”) by passenger influx from alternative routes (modes), but furthermore – they might be actually partially (or even totally) undone by the phenomenon of latent (induced) demand (Szarata, 2013). The city of London provides a good example in terms of that narrative, illustrating how massive investment programmes in transportation systems can barely keep up with the ever growing demand pressure. A multi-billion improvement programme currently underway across the London Tube (underground rail) system is projected to increase the system capacity by approx. 30%, but analysis prepared for the busiest Tube line, the Northern Line (Fig. 7), shows already that by the time the works have been finished - the crowding levels will be even worse than before, virtually along each single section of the line (Transport for London, 2013). A flagship Crossrail project (ca. £17bn of total cost) is supposed to contribute 10% to the total urban transport network capacity – a substantial nominal gain in the city of 8m inhabitants - and become a core part of the public transport system. Though it is widely expected to relieve the existing Tube network, transport planners predict that once opened in 2018 the Crossrail “will be immediately full up with people” (Drabicki., 2015), and argue that a second Crossrail line is badly “needed” to counteract the anticipated passenger congestion. Numerous similar case studies can be found elsewhere in biggest urban metropolitan areas across the world, in case of which the public transport systems are particularly likely to become prone to massive passenger congestion and induced demand pressure. As mentioned earlier, impact of overcrowding on long-term passenger path choices also concerns the modal choices and departure time choices. (Tirachini et al., 2013) use stated-preference passenger survey data, and propose a range of MNL models to estimate demand choice models arising from inclusion of crowding discomfort in travel cost formula (Fig. 8). This Sydney-based study emphasises that models insensitive to crowding discomfort are likely to underestimate the value of in-vehicle travel times savings and overestimate the demand (model) share for high congestion levels (and vice versa for low congestion levels). An important observation is that for suburban railway trips, the inclusion of repeated overcrowding experience should produce a demand shift towards

private transport, with crowding “sensitivity” rate increasing as a function of trip duration: an uncongested model would yield a constant modal share of a sample rail line at ca. 5%, whereas for a congested model the modal share would range between 4 – 6% (travel time of 15 minutes) or ca. 3 – 8% (travel time of 40 minutes). In terms of departure time updating process, (Nuzzolo et al., 2012) incorporate a day-to-day learning mechanism in public transport assignment model, so as to emphasise the long-term implications of crowding

experience. The proposed framework shows that approx. 65% of commuters shift towards other (earlier or later) vehicle runs to mitigate the risk of on-board congestion, leaving on average 5 minutes earlier at the origin – a “spillback” effect can be observed in temporal demand distribution pattern: individual vehicle run loads might now substantially differ from their initial values once a congestion-induced adjustment takes place in passengers’ choice process.

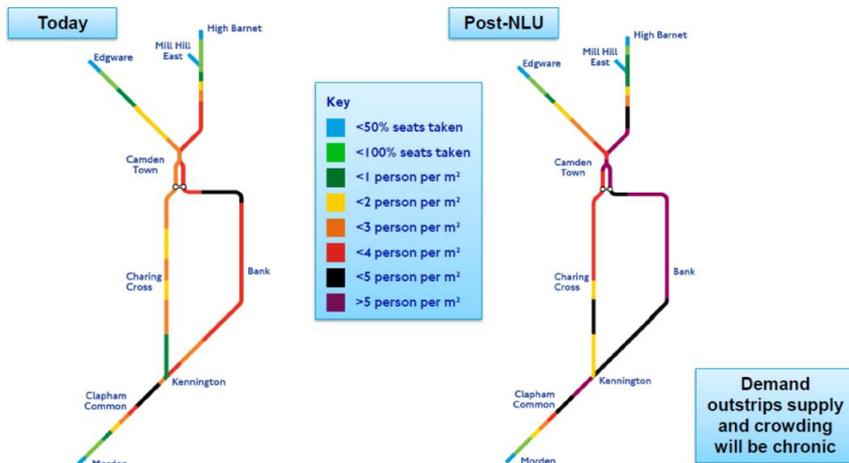


Fig. 7. Sample results for the London Underground case study: despite massive investment programme (NLU), capacity increases on the Northern Line will be quickly absorbed by induced passenger demand growth (source: Transport for London, 2013)

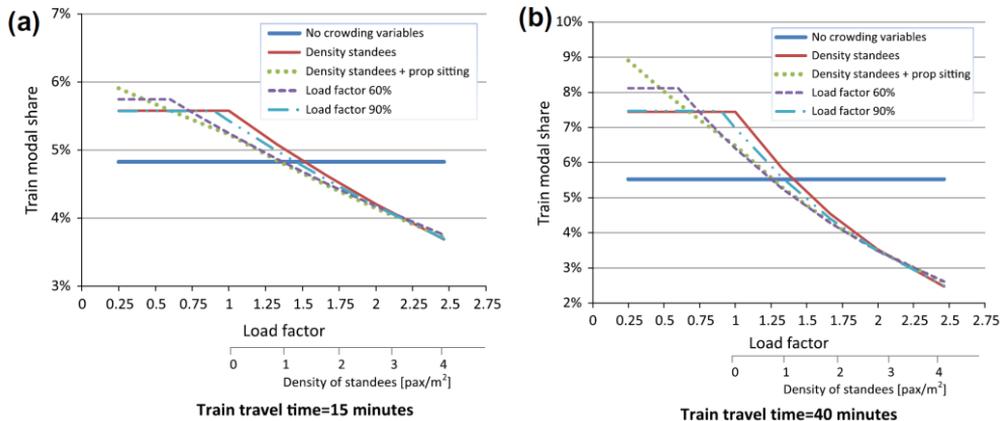


Fig. 8. Sample results obtained with different crowding cost functions – correlation between the overcrowding (discomfort) impact and the modal share of commuter rail system (source: Tirachini et al., 2013)

4. Assignment algorithms

In the practical part of this study, two public transport assignment algorithms will be tested on a sample transport network, to observe how they replicate the effects of passenger overcrowding on the evolving transport system performance and travelling experience (mainly in terms of journey times and service loads). Each algorithm utilises a distinct approach to modelling the capacity constraints of public transport systems, and assumes a different modelling aggregation level both on demand and supply sides, i.e.:

- **implicit approach:** timetable-based (i.e. schedule-based), macroscopic assignment model – as implemented in the commonly-used PTV VISUM software,
- **explicit approach:** simulation-based (i.e. agent-based), mesoscopic assignment model – as incorporated in the currently developed BusMezzo software.

4.1. Implicit capacity constraints' algorithm

The timetable-based assignment model operates on a macroscopic level, reproducing travel demand in form of aggregated link flows (within a certain time period). The path choice model is a one-off process triggered at the origin, when traveller chooses a complete O-D path (route), based on its (predetermined) utility value – and follows that single path all the way to his (her) destination. The baseline path utility formula is a sum of weighted (perceived) travel time components (in-vehicle, waiting, walking times), transfer penalties, and the temporal utility of that O-D connection. Additionally, once a capacity restraint model is introduced, a crowding mark-up penalty ($I + Av$) is assigned to the total path utility. The crowding penalty is recalculated in an iterative process, based

on the volume-to-capacity ratio of each link segment (importantly – not individual line segments), until a certain convergence (equilibrium) threshold is attained – i.e. a fixed-point problem solution after which a final path utility (impedance) rate is evaluated. The algorithm utilises a VDF-based procedure analogous to the private transport congested assignment model, with 3 crowding impedance functions available. For the purposes of this study, the SBB (Swiss Railway) function was assumed as it should allow us to replicate the two-tier effect of rising network overcrowding upon the path utility (impedance): for low volume-to-capacity rates, the effects of rising passenger discomfort (crowding mark-up penalty within range of 1.0 – 1.7), and a step-wise jump in path impedance (constant crowding mark-up penalty of 10.0) once passenger volume exceeds the assumed *crush capacity* (Fig. 9). This algorithm should reproduce, in a simplified – i.e. *implicit* – approach, the effects of capacity constraints on output passenger path choices: reductions in excessive (overestimated) passenger volumes and increasing attractiveness of less-crowded routes – though without considering the more specific, congestion-associated phenomena, particularly at the stops.

4.2. Explicit capacity constraints' algorithm

The simulation-based assignment model assumes a more disaggregate representation both of transport demand – individual agents (travellers), and transport supply – individual vehicles (trips) operating within the network. Here, the path utility is recurrently updated at each journey stage, when traveller may reconsider his (her) path (route) choice towards the destination – i.e. at each instance a boarding, alighting or connection decision process is triggered.

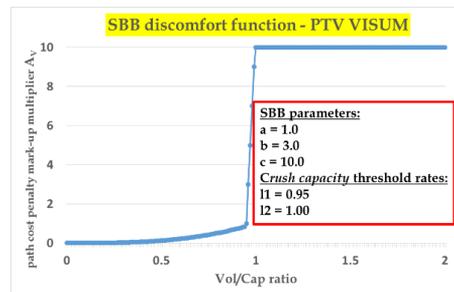
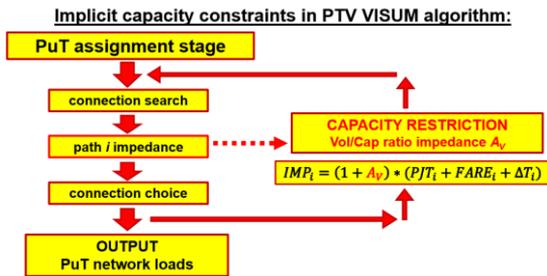


Fig. 9. Implicit capacity constraints' algorithm assumptions in simulations

Likewise, the path utility formula comprises the same set of travel time components plus transfer penalties, except for temporal utility of connection which was not yet included in the algorithm. Since the modelling algorithm operates on a more detailed, mesoscopic level, resultant passenger flows are an aggregate output of all the individual actions (path choice decisions) taken by agents (travellers) progressing through the network. Service supply is modelled as individual trips (runs) served by public transport vehicles, which are described with their distinguished properties - vehicle type, vehicle dynamics, and importantly - specified maximum passenger load capacity. Network performance is reproduced in a more stochastic manner – the actual travel times depend on real-time system conditions, and a dwell-time function is introduced to describe the direct impact of dwelling (i.e. boarding and alighting) flows onto dwell times – in our case, we will assume a linear dwell-time function of 2 secs/pass. The utilised modelling framework did not incorporate yet the impact of crowding discomfort upon the path cost-utility formula; nonetheless, it would allow us to observe the actual transport network performance and its implications for passengers’ travelling experience once network capacity constraints are modelled in an *explicit* approach – i.e. with strict denial-of-boarding and arising queuing phenomena occurring at stops if passenger flows exceed the system capacity, and the

very important demand-supply interactions (Fig. 10).

5. Results – implicit and explicit approaches

Simulations presented below were performed on a sample public transport network, i.e. the extended version (“SF ENet” (Fonzone and Schmoecker, 2014)) of the classical Spiess-Florian network (Spiess and Florian, 1989). The extended SF ENet layout is assumed on a network topology formulated by (Fonzone, Schmoecker 2015) and comprises a system of 7 bus stops (A to G) and 5 unidirectional bus lines (L1 to L5), situated along 2 parallel O-D routes (Fig. 11). A single origin-destination pair is assigned to the network. Travellers are allowed to transfer between bus lines at stops, and additionally a two-way, 3-minute walking connection is provided between the intermediate stops C and F. Vehicle runs are dispatched from origin stops at fixed intervals (headways), and line run times between consecutive stops remain constant. The crush capacity rate of each bus vehicle is assumed as 100 pax.; in explicit approach, a dwell-time function is introduced with a linear rate of 2 secs/pass.

To analyse the incorporation of passenger overcrowding effects in the sample network, 2 distinct modelling approaches were included, i.e. the implicit approach (PTV VISUM) and the explicit approach (BusMezzo) to modelling the capacity constraints.

Explicit capacity constraints in BusMezzo algorithm:

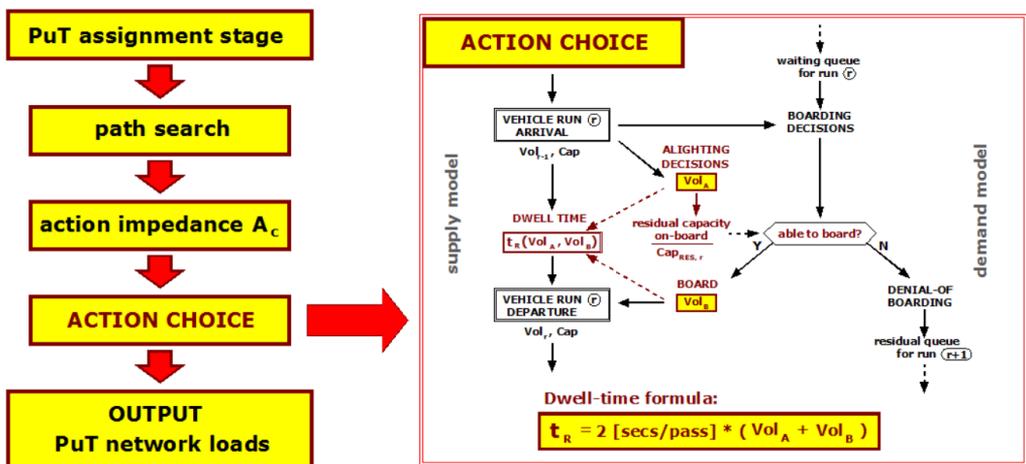


Fig. 10. Explicit capacity constraints' algorithm assumptions in simulations

For both of these, 4 individual O-D demand cases were assigned which should reflect the rising O-D demand conditions in the following stages:

- undersaturated conditions (1600 pax./hour – “LOW” congestion case),
- saturated network state (3200 pax./hour – “MID” congestion case),
- moderately and massively overcrowded conditions (6400 pax./hour – “HIGH” congestion case, and 16000 pax./hour – “V. HIGH” congestion case).

These respective O-D demand values correspond roughly to 50%, 100%, 200% and 500% of generated line capacity (per hour) combined for initial 3 line segments (L1, L2 and L5) departing from the origin. Total simulation run time is 120 minutes: service supply is generated during the whole 120 minutes, whereas passenger demand is assigned after initial 30 minutes and is generated within the next 60 minutes.

Simulations performed on a sample network reveal that both modelling approaches produce different assignment output as a consequence of rising O-D passenger volumes, with respect to each individual (described above) category of passenger overcrowding effects. Starting with the inclusion of physical capacity limits, the implicit approach has

relatively more limited impact on output network performance: in aggregate terms, average journey times increase from 21.7 mins (“LOW” congestion case) to just 24.3 mins (“V. HIGH” congestion case) (Fig. 12). These changes in journey times are pretty much minimal and can be merely attributed to the relative shifts in O-D path choices (i.e. paths with longer in-vehicle travel times become somewhat more attractive), but they do not reflect any changes in waiting times - which remain virtually constant (or even decrease slightly) in the event of massive passenger congestion. This stands in stark contrast to the explicit constraints’ algorithm which reveals much more significant changes in travel times: as a consequence of rising passenger congestion, average journey times increase from 31.4 mins (“LOW” congestion case) to 63.7 mins (“V. HIGH” congestion case). Here, the average in-vehicle travel times remain constant, but a significant surge in waiting times takes place now due to congestion-induced queuing phenomena at stops, which are evidently captured by the explicit algorithm: as O-D demand volume exceeds the system capacity, a rising share of passengers is denied the boarding and becomes increasingly delayed as they try to reach the destination.

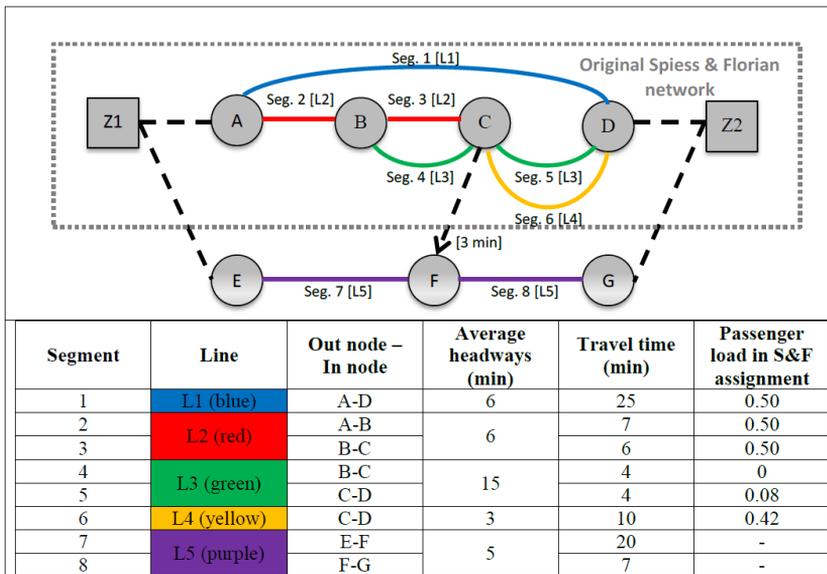


Fig. 11. Spiess-Florian extended network (SF ENet) -topology of sample bus transport network used in simulation works (source: Fonzone and Schmoecker, 2015).

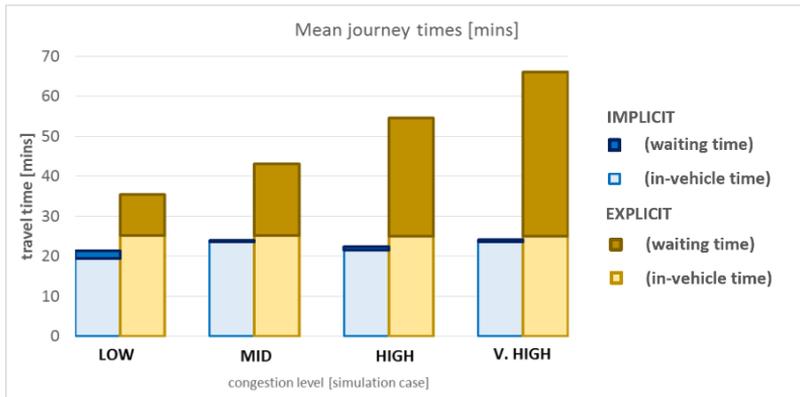


Fig. 12. Simulation results – substantial differences in mean journey times between the explicit and implicit algorithms

In the implicit approach, the effects of arising congestion are described by the travel cost penalty imposed by the SBB function: it reflects the travellers’ willingness to shift towards less-crowded connections, but does not account for strict denial-of-boarding: in the end, 100% of travellers will

reach the destination successfully and the whole O-D demand volume would be redistributed within the whole 2-hour simulation period to earlier or later departures, even if it implies volume-to-capacity ratio values reaching up to 500% on individual line segments (Fig. 13).

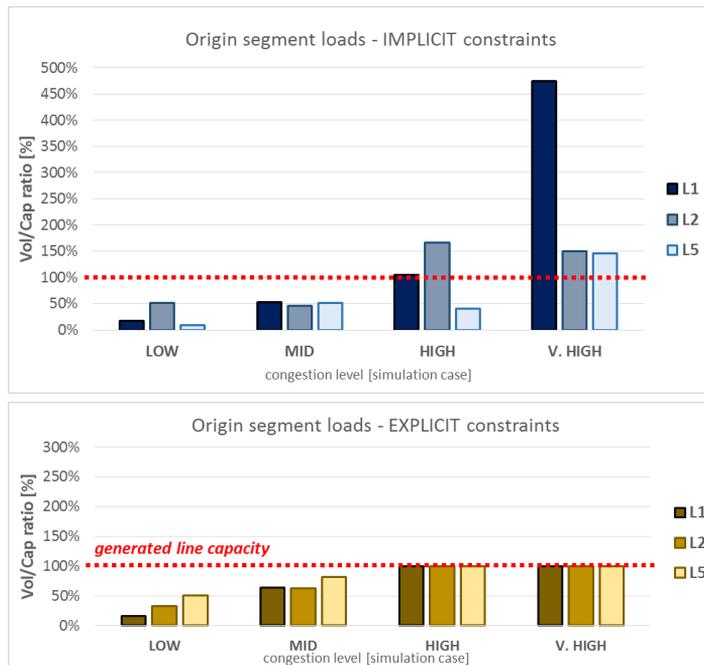


Fig. 13. Results – differences in origin segments’ (L1, L2 and L5) Vol/Cap ratios - plotted against the generated line capacity threshold

In contrast, in the explicit approach a strict denial-of-boarding principle is observed for every additional passenger beyond the capacity limit: volume-to-capacity ratio will never exceed 100%, and travellers would have to wait for the ensuing service runs which will have spare on-board capacity. Consequently, the probability-of-arrival at the destination decreases sharply as overcrowding develops in the SF ENet: for “HIGH” and “V. HIGH” congestion cases, 46% and 71% of travellers respectively will not make it to the destination after 120 minutes of simulation run time, and will still remain stranded somewhere in the network (Fig. 14).

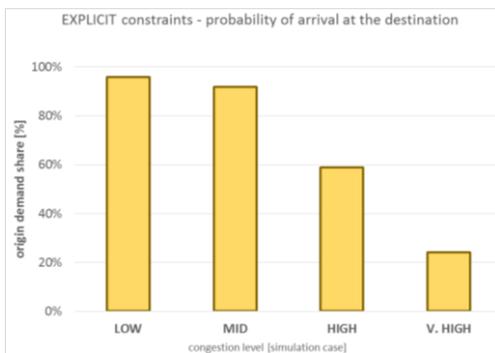


Fig. 14. Results - rising failure-to-arrive probability at the destination in the explicit approach, as a consequence of increasing network congestion

An important remark regarding the mesoscopic-based (explicit) algorithm performance should be made here, which is related to distinct assumptions utilised in the probabilistic discrete choice algorithm. Each time (i.e. at each instance) the traveller makes a travel decision, each alternative he (she) considers in the O-D choice set is described with a non-zero probability – thus, he (she) will *most likely – but not necessarily* – choose the O-D alternative with the highest utility value. Simulation works assumed a default MNL *theta* parameter value of 0.50 – which should be in practice properly calibrated (i.e. most likely, increased) to match the expected probability rate of rational choice behaviour. This comprises a significantly distinct feature of mesoscopic-based algorithm assumptions

– and therefore, the exact travel time values should not be interpreted in absolute terms (e.g. in comparison to macroscopic-based algorithm) but rather used to observe relative changes as a consequence of system overcrowding. This is also the reason behind a non-zero failure-to-arrive probability rate (at the destination) even in low congestion scenarios; in higher congestion levels, the additional rises in this probability rate can be directly attributed to the implications of overcrowding-induced phenomena.

A major difference in the assignment output concerns the replication of demand-supply interactions, i.e. the feedback effect between passenger congestion and service performance. This cannot be captured within the implicit approach, where both service run times and dwell times remain unaffected despite the passenger overcrowding – regardless of all the simulation cases. However, it is of utmost importance in case of explicit approach, where mutual dynamic developments on-going in the congested network have profound implications both on the demand (passengers’) and the supply (services’) side. A significant growth in dwell times can be evidently observed for individual vehicle runs as passenger boarding and alighting flows increase at the stops, which result in up to 50% longer total run times of bus trips in the SF ENet. The demand-supply feedback loop is perhaps best demonstrated when plotting dwelling flows against service headways for consecutive vehicle runs (Fig. 15): it shows that service headways are likely to deviate from their nominal values when fluctuations in dwelling flows grow higher. Importantly, the biggest headway deviation values are correlated not with the extreme demand magnitude - but principally with the *extreme demand variance*: the biggest “bumps” in line headways tend to overlap with the highest “bumps” in dwelling flows. This is a characteristic feature of the on-going bus bunching effect (described above), which in highly overcrowded simulation cases (“HIGH” and “V. HIGH” cases) becomes a self-sustaining phenomenon, reflecting that the network performance falls out of stability state - and will only diminish in the final 30 minutes of simulation period once O-D demand generation ceases and the SF ENet finally “recovers” from massive congestion.

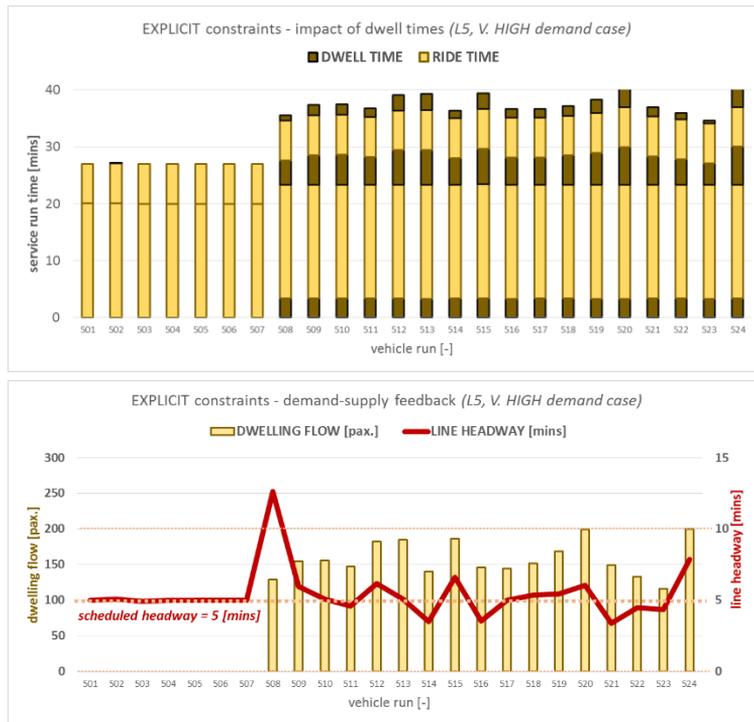


Fig. 15. Results - mutual demand-supply interactions captured in the explicit approach: sample effects on service run times (top) – up to 50% longer service times, and headway deviations (bottom), induced by passenger flows

The differences in the observable assignment output can be attributed to the assumed aggregation level within the simulation algorithm. The implicit approach operates on a macroscopic level, where transport demand and transport supply systems can only be traced in terms of aggregate flows and link segments for the whole assignment period: a more exact examination of service run times' or journey times' distribution is not possible within the scope of this algorithm, and output network performance is principally measurable with average (aggregate) indicator rates. The explicit approach assumes a more disaggregate representation both on the demand as well as the supply side, and thus enables to observe much more detailed output for each individual component of the transport system – i.e. journey times of individual travellers, and service run times of individual vehicle runs. This allows us to reproduce an interesting passenger arrival pattern at the destination, which also mimics the demand-

supply feedback interaction: for higher congestion cases, the rising bus bunching effect eventually induces a **“passenger bunching” pattern**, with O-D demand arrivals becoming more concentrated (“bunched”) due to system capacity bottlenecks (Fig. 16).

Finally, distribution of path choice patterns also exposes substantial differences between the two assignment algorithms, as seen on the example of path choice shares between 3 line segments at the origin (L1, L2 and L5) (Fig. 17). In the implicit approach, the path choice formula reflects the discomfort cost penalty already for low and moderate crowding conditions. Thus, a substantial shift can be observed when congestion rises from the “LOW” to the “MID” case: the O-D demand becomes pretty much equally distributed between the 3 segments, and for each of them the volume-to-capacity ratio stabilises between 46% to 52%.

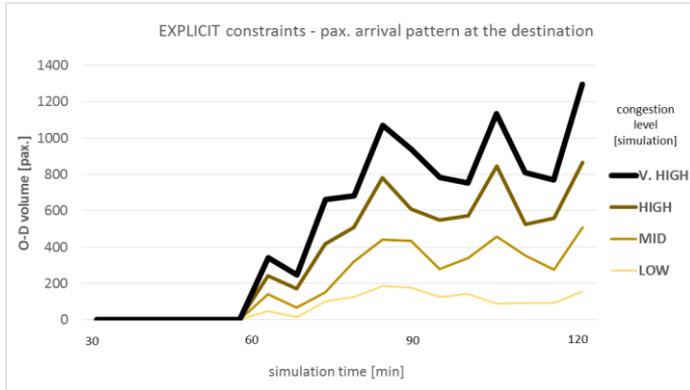


Fig. 16. Results - implications of demand-supply feedback in the explicit approach: fluctuations in service performance (bus bunching) eventually influence the overall pass. arrival ("pass. bunching") pattern at the destination.

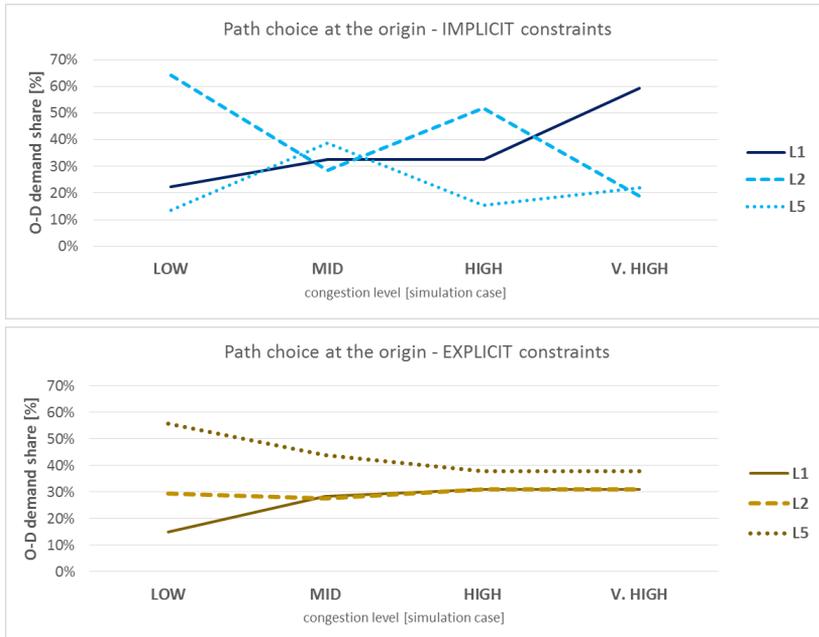


Fig. 17 Results - arising differences in path choice at the origin between the implicit (top) and explicit (bottom) approaches

However, for further (“HIGH” and “V. HIGH”) congestion cases no consistent path choice pattern can be derived or explained: the O-D demand shares alternately jump up or drop down, suggesting that the network output could not reach a stable

(equilibrium) solution - the small-scale SF ENet with its simple topology becomes simply a few times more overloaded than its generated capacity rate. In the explicit approach, no discomfort cost penalty was included in the path cost formula yet, and the

output path shares reflect merely the *explicit* impact of line segments' capacity limits. This actually produces a pretty much consistent picture of evolving demand distribution pattern: in the relatively less crowded simulation cases, a bulk of O-D demand share is concentrated along the most attractive L5 line segment (56% in the "LOW" congestion case), whereas increasingly congested conditions in the SF ENet result in a more even utilisation (distribution) of the available system capacity and hitherto less attractive O-D routes: in the "HIGH" congestion case, the L5 patronage rate drops to 38% whereas both the L1 and L2 patronage rates reach 31%.

6. Evaluation and conclusions

The objective of this paper was to discuss the incorporation of passenger congestion (overcrowding) effects in public transport assignment models. The first part of this work, being based on a comprehensive literature review, aimed to outline main categories of passenger overcrowding effects, which should be accounted for in the assignment output – i.e. the inclusion of physical capacity constraints (limits); the feedback effect between passenger demand and service supply performance (e.g. the well-known bus bunching effect); and the feedback effect on passenger discomfort (travel cost). Then, a brief presentation of sample case studies followed up, which demonstrated that the inclusion of public transport capacity constraints might significantly affect the actual assignment output and final analysis results – with implications not only for the path (route) choice stage, but also going back to the mode choice and long-term demand adaptation process (e.g. departure time choice). Further on, the simulation part of this

paper aimed to examine the replication of passenger overcrowding effects on a small-scale, sample transport network for 2 distinct modelling approaches to public transport capacity constraints – i.e. macroscopic and mesoscopic assignment models, implemented respectively in the PTV VISUM and BusMezzo algorithms.

Both assignment algorithms can reproduce passenger overcrowding effects in a substantially different manner, and will consequently yield quite distinct assignment output (Table 1). The macroscopic algorithm assumes an implicit, simplified approach to modelling capacity constraints of public transport system – i.e. in the form of VDF-based procedure. The increasing travel cost penalty aims to reflect the two-tier effect of arising congestion on passenger path choices – i.e. shifts due to travel discomfort (in low-congested conditions) and demand outflow towards alternative connections (once volume exceeds capacity). As congestion arises in the network, the implicit approach tends to redistribute the O-D demand towards available system capacity and hitherto less attractive O-D paths (routes) – though a certain drawback of this algorithm is that for massive congestion levels, the assignment procedure might not actually converge to a stable (equilibrium) solution and would produce erroneous path choices. However, the implicit-based algorithm does not observe the exact capacity limits of travel supply – eventually 100% of O-D demand will be assigned to the network - nor does it capture the resultant queuing phenomena. Perhaps more importantly, the mutual interaction between transport demand vs. transport supply performance, induced by passenger congestion phenomena, is also missing.

Table 1. Summary - inclusion of 3 distinguished categories of passenger overcrowding effects in the 2 analysed modelling algorithms

Modelling approach:	IMPACT OF PUBLIC TRANSPORT CONGESTION (PASSENGER OVERCROWDING) EFFECTS – SUMMARY:		
	Physical capacity constraints (limits)	Feedback on demand-supply performance	Feedback on passenger discomfort
MACROSCOPIC, timetable-based (PTV VISUM)	IMPLICIT	no*	yes
MESOSCOPIC, simulation-based (BusMezzo)	EXPLICIT	yes	yes**

* limited functionality (i.e. simplified elongation of service run times) only

** not available yet at the time of this research work

The mesoscopic algorithm utilises a more detailed representation of transport system both on demand and supply sides, with a more explicit, detailed approach to modelling capacity constraints – i.e. by observing exact capacity limits of public transport vehicle runs and the ensuing passengers' sequential travel choices. Passengers experience strict denial-of-boarding once a vehicle becomes overcrowded, and as a result the queuing phenomena occur at stops. The passengers' arrival probability is thus heavily influenced by the capacity restraint regime of transport system: the excessive O-D demand share will not arrive yet at the destination if volume exceeds (generated) system capacity. Importantly, the explicit approach enables to reproduce much more dynamic demand-supply interactions, in the form of the on-going feedback between passenger flows and service performance. Since dwell times depend mutually on dwelling flows, a clear-cut, developing bus bunching effect can be traced for individual (consecutive) vehicle runs which is characteristic for congested public transport networks. Resultant path choices reflect shifts both due to available network capacity and current service performance: no feedback effect on path cost (discomfort) was tested yet at the time of this research work, but this can now also be incorporated in the mesoscopic-based algorithm.

Based on the summarised state-of-the-art research works, as well as own simulation works, it seems conceivable that the implicit (macroscopic) modelling approach can be used to model the impact of passenger congestion on path choices in a simplified manner - though a more accurate and reliable representation of the congestion-induced effects is feasible only with an explicit (mesoscopic) modelling approach. The implicit approach is more commonly implemented in the state-of-the-practice assignment models and can be already utilised to replicate a certain (limited) overcrowding impact on route choices and modal shifts in city-scale transport models; the explicit approach remains less developed and often constrained to individual application studies, but comprises a more promising and comprehensive method of representing the whole complexity of public transport congestion effects. Future research works should involve a more specific investigation of model calibration and validation, as well as comparison of assignment output on a city-scale, multimodal transport model –

not only just in terms of induced shifts in path choices, but also the far-reaching influence on modal choices, temporal choices and long-term demand adaptation processes (Drabicki et al., 2016) - so as to assess the overall effectiveness of both implicit and explicit algorithms in modelling the public transport network capacity constraints.

References

- [1] ATTANUCCI, J., 2010. Public Transportation Systems, Lecture 14. Lecture presented at the Massachusetts Institute of Technology, Cambridge, MA. Available from: https://ocw.mit.edu/courses/civil-and-environmental-engineering/1-258j-public-transportation-systems-spring-2010/lecture-notes/MIT1_258JS10_lec14.pdf [Accessed 21st May 2017]
- [2] BARRY, J., 2015. London's Bus Service – monitoring satisfaction. In *Smart Public Transport Conference*, Warsaw, Poland.
- [3] BATARCE, M., MUÑOZ, J. C., ORTUZAR, J. D., RAVEAU, S., MOJICA, C., & RÍOS, R. A., 2015. Valuing crowding in public transport systems using mixed stated/revealed preferences data: the case of Santiago. In *TRB 94th Annual Meeting Compendium of Papers*, Washington DC.
- [4] BAĞ, R., 2010. Simulation model of the bus stop. *Archives of Transport*, 22(1), 5-25.
- [5] BRANSTON, D., 1976. Link capacity functions: A review. *Transportation Research*, 10(4), 223-236.
- [6] CASCETTA, E., 2013. Transportation systems engineering: theory and methods (Vol. 49). *Springer Science & Business Media*
- [7] CATS, O., 2011. Dynamic Modelling of Transit Operations and Passenger Decisions. Doctoral thesis in Transport Science with specialisation in Transport Systems. KTH – Royal Institute of Technology, Stockholm, Sweden.
- [8] CATS, O., WEST, J., & ELIASSON, J., 2014. Appraisal of increased public transport capacity. In *hEART Conference 2014* in Leeds, UK.
- [9] DRABICKI, A., 2015. Incorporation of public transport overcrowding effects on passenger path choices in transit assignment. Master

- Thesis. Cracow University of Technology, Krakow, Poland.
- [10] DRABICKI, A., KUCHARSKI, R., & SZARATA, A., 2016. Zastosowanie ograniczeń przepustowości sieci transportu publicznego w makroskopowym rozkładzie ruchu. (ENG.: Incorporation of public transport network capacity constraints in macroscopic trip assignment model) *Transport Miejski i Regionalny*, 08/2016., Krakow, Poland.
- [11] FONZONE, A., & SCHMÖCKER, J. D., 2014. Effects of transit real-time information usage strategies. *Transportation Research Record: Journal of the Transportation Research Board*, (2417), 121-129
- [12] FONZONE, A., SCHMÖCKER, J. D., & LIU, R., 2015. A model of bus bunching under reliability-based passenger arrival patterns. *Transportation Research Part C: Emerging Technologies*, 59, 164-182.
- [13] GENTILE, G., & NOEKEL, K. (eds.), 2016. Modelling public transport passenger flows in the era of intelligent transport systems. *Gewerbestrasse: Springer International Publishing*.
- [14] HARTL, M., 2013. Route choice in macroscopic and microscopic assignment models for public transport. Master thesis. Universitaet Stuttgart, Germany.
- [15] HORBACHOV, P., NAUMOV, V., KOLII, O., 2015. Estimation of the bus delay at the stopping point on the base of traffic parameters. *Archives of Transport*, 35(3), 15-25.
- [16] KIM, J. K., LEE, B., & OH, S., 2009. Passenger choice models for analysis of impacts of real-time bus information on crowdedness. *Transportation Research Record: Journal of the Transportation Research Board*, (2112), 119-126.
- [17] LEURENT, F., 2009. On seat congestion, passenger comfort and route choice in urban transit: a network equilibrium assignment model with application to Paris. In *Annual Meeting of the Transportation Research Board Session Transit Capacity and Quality of Service* (pp. TRB-09). TRB.
- [18] LIEBERHERR, J., & PRITSCHER, E., 2012. Capacity-restraint railway transport assignment at SBB-Passenger. In *Proceedings of the 12th Swiss Transport Research Conference*.
- [19] LONDON ASSEMBLY TRANSPORT COMMITTEE REPORT, 2009. The Big Squeeze. Rail overcrowding in London. Report commissioned by the Greater London Authority, UK
- [20] LONDON ASSEMBLY TRANSPORT COMMITTEE REPORT, 2009. Too close for comfort. Passengers' experiences of the London Underground. Report commissioned by the Greater London Authority, UK.
- [21] KROES, E., KOUWENHOVEN, M., DEBRINCAT, L., & PAUGET, N., 2013. On the Value of Crowding in Public Transport for Ile-de-France.
- [22] MOREIRA-MATIAS, L., FERREIRA, C., GAMA, J., MENDES-MOREIRA, J., & DE SOUSA, J. F., 2012, July. Bus bunching detection by mining sequences of headway deviations. In *Industrial Conference on Data Mining* (pp. 77-91). Springer Berlin Heidelberg.
- [23] NEWELL, G. F., & POTTS, R. B., 1964. Maintaining a bus schedule. In *Australian Road Research Board (ARRB) Conference*, 2nd, 1964, Melbourne (Vol. 2, No. 1).
- [24] NUZZOLO, A., CRISALLI, U., & ROSATI, L., 2012. A schedule-based assignment model with explicit capacity constraints for congested transit networks. *Transportation Research Part C: Emerging Technologies*, 20(1), 16-33.
- [25] PTV AG, 2015. VISUM 15 User Manual. Karlsruhe, Germany.
- [26] RUDNICKI, A., 1999. Jakość komunikacji miejskiej. (ENG.: Quality of urban public transport.) In *Zeszyty Naukowo-Techniczne Oddziału Stowarzyszenia Inżynierów i Techników Komunikacji w Krakowie*, (71). Krakow, Poland.
- [27] SMALL, K. A., & GOMEZ-IBANEZ, J. A., 1999. Urban transportation. Handbook of regional and urban economics, 3, 1937-1999.
- [28] SPIESS, H., & FLORIAN, M., 1989. Optimal strategies: a new assignment model for transit networks. *Transportation Research Part B: Methodological*, 23(2), 83-102.
- [29] SZARATA, A., 2013. Modelowanie podróży wzbudzonych oraz tłumionych zmianą stanu infrastruktury transportowej. (ENG.: Modelling of induced and suppressed trips resulting from changes in transport

- infrastructure) Cracow University of Technology, Krakow, Poland.
- [30] SZARATA, A., 2014. The multimodal approach to the modelling of modal split. *Archives of Transport*, 29(1), 55-63.
- [31] TIRACHINI, A., HENSHER, D. A., & ROSE, J. M., 2013. Crowding in public transport systems: effects on users, operation and implications for the estimation of demand. *Transportation research part A: policy and practice*, 53, 36-52.
- [32] TRANSPORT FOR LONDON, 2013. Capacity for growth at Camden stations. Presented in the London Borough of Camden, 18th September 2013, Greater London, UK.
- [33] VAN OORT, N., DROST, M., BRANDS, T., & YAP, M., 2015. Data-driven public transport ridership prediction approach including comfort aspects. In *13th CASPT Conference*, Rotterdam, The Netherlands.
- [34] WHELAN, G. A., & CROCKETT, J., 2009, March.. An investigation of the willingness to pay to reduce rail overcrowding. In *International Choice Modelling Conference 2009*.
- [35] ŻOCHOWSKA, R., 2014. Selected issues in modelling of transport flows in congested urban networks. *Archives of Transport*, 29(1), 77-89.