MIT Open Access Articles

Capacity constrained accessibility of high-speed rail

CrossMark

Capacity constrained accessibility of high-speed rail

Yu Shen[1,2] · Jinhua Zhao[1]

Abstract This paper proposes an enhanced measure of accessibility that explicitly considers circumstances in which the capacity of the transport infrastructure is limited. Under these circumstances, passengers may suffer longer waiting times, resulting in the delay or cancellation of trips. Without considering capacity constraints, the standard measure overestimates the accessibility contribution of transport infrastructure. We estimate the expected waiting time and the probability of forgoing trips based on the $M/G^B/1$ type of queuing and discrete-event simulation, and formally incorporate the impacts of capacity constraints into a new measure: capacity constrained accessibility (CCA). To illustrate the differences between CCA and standard measures of accessibility, this paper estimates the accessibility change in the Beijing–Tianjin corridor due to the Beijing–Tianjin intercity high-speed railway (BTIHSR). We simulate and compare the CCA and standard measures in five queuing scenarios with varying demand patterns and service headway assumptions. The results show that (1) under low system loads condition, CCA is compatible with and absorbs the standard measure as a special case; (2) when demand increases and approaches capacity, CCA declines significantly; in two quasi-real scenarios, the standard measure overestimates the accessibility improvement by 14–30 % relative to the CCA; and (3) under the scenario with very high demand and an unreliable timetable, the CCA is almost reduced to the pre-BTIHSR level. Because the new CCA measure effectively incorporates the impact of capacity constraints, it is responsive to different arrival rules, service distributions, and system loads, and therefore provides a more realistic representation of accessibility change than the standard measure.

Keywords Accessibility · Capacity constraint · Discrete event simulation · High-speed rail · Queuing theory

✉ Jinhua Zhao
   jinhua@mit.edu

[1] Department of Urban Studies and Planning at Massachusetts Institute of Technology, Cambridge, USA

[2] DECIVIL/CERIS at Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

🍦 Springer

Introduction

Accessibility, as a vital nexus between transportation and land use systems, is widely used in both scientific studies and policy-making practices. Hansen's (1959) seminal work on accessibility first demonstrated two important elements in its measurement, the "spatial distribution of activities" and the "ability/desire to overcome spatial separation." Following Hansen's work, there have been numerous works studying the definitions, measures, and applications of accessibility. These works have been systematically reviewed by Bhat et al. (2000), Geurs and Ritsema Van Ecka (2001), Geurs and van Wee (2004) and Curtis and Scheurer (2010). In evaluating the relationship between land use and transportation strategies, Geurs and van Wee (2004) identify four basic perspectives on accessibility measures: (1) infrastructure-based measures, e.g. levels of congestion or average travel speed on the road network; (2) location-based measures, e.g. the sum of the product of jobs and travel time from origin to destination; (3) person-based measures, e.g. the activities in which an individual can participate within a given time budget; and (4) utility-based measures, e.g. the derived logsum of utilities from a combined mode-destination choice set. In studies of the accessibility impacts of high-speed rail (HSR), the location-based measures are the most commonly used measurements (Baptiste et al. 2003; Gutiérrez 2001; Gutiérrez et al. 1996; Gutiérrez and Urbano 1996; Shen et al. 2014; Spiekermann and Wegener 2006; Vickerman and Ulied 2009). Though given in various mathematical forms, location-based measures can be generically expressed as:

$$A_i = \sum_{j=1}^{J} g(O_j) f(I_{ij}), \tag{1}$$

where A_i location-based accessibility indicator at origin i; $g(O_j)$ activity function of opportunities O_j, e.g. population, number of jobs in destination j, representing the "spatial distribution of activities"; $f(I_{ij})$ impendence function of generalized travel cost or travel time from origin i to destination j, reflecting the element of "the ability/desire to overcome spatial separation."

The specific forms of Eq. (1) vary according to research purposes. Bhat et al. (2000) show that past studies have developed abundant mathematical forms of the activity function $g(O_j)$. Recent studies have further refined the measures of accessibility by integrating different theories into the calculation of activity functions, e.g. the competition of opportunities at destination, or the evaluation of equity (Geurs and van Wee 2004; Lucas et al. 2015; van Wee et al. 2001). As macro-level measures, accessibility models tend not to incorporate the details of transit operations in order to maintain the simplicity of the models in general cases. However, we argue that there is at least one important operational detail to model: capacity constraint, which could have significant impact on the accessibility measure in the planning and strategic phases of transport development. The discussion of capacity constraints is mainly found in research on transit network assignments (see Leurent (2011) and more details in "Capacity constraints in transit networks" section). A review from Kitamura (2009) reveals the importance of capacity effects on measuring accessibility and also shows that past studies fail to establish links between these two. As shown in Bhat et al. (2000), there is little discussion of the impendence function, $f(I_{ij})$, especially from an operational perspective. Thus, there is a clear gap between the accessibility literature and the transit network assignment literature. This paper aims to bridge the two by focusing on circumstances in which the capacity of transport infrastructure or

service is limited, and explicitly incorporating capacity constraints into the accessibility measurement.

In their review of accessibility measurements, Curtis and Scheurer (2010) point out that travel time is possibly the most commonly used measure to quantify travel impediment. In a transit network, waiting time (or transfer time) is an essential component of travel time calculation. Past studies convey the significance of travel time reliability (Bates 2001; Geurs and van Wee 2004) but do not take into account constraints on service capacity. This results in two common assumptions: (1) that waiting time can be estimated simply by taking the half of headway as expected waiting time (Curtis and Scheurer 2010), or by taking half of headway multiplied by a function of the coefficient of headway variation (Osuna and Newell 1972); (2) that all passengers are able to get onto the trains, i.e. no one is left behind. In densely populated areas, especially during peak hours, both assumptions could significantly diverge from reality. When capacity is constraining, passenger boarding becomes probabilistic. For those who are left behind, the service effectively becomes unavailable; for those who board the train, actual waiting times can be far longer than the expected waiting time estimated by the public transit timetable. By ignoring capacity constraints and making these unrealistic assumptions, the standard measure may substantially overestimate the accessibility impact of a new infrastructure or service. The problem can affect both urban public transit networks and intercity railway networks, including HSR, whenever demand approaches the service capacity.

Capacity constraints in transit networks

From the interactions between passengers, vehicles, stations, and lines emerge various capacity issues, which are classified by Leurent (2011) into seven groups: (1) the vehicle capacity of an infrastructure; (2) the operational capacity of vehicle fleet in a route; (3) the passenger capacity of a vehicle; (4) the capacity of vehicles covering a route during a certain period; (5) the passenger capacity of a station; (6) the vehicle storage and movement capacity of a station; and (7) the capacity of a station for interface with personal transport modes. Among these capacity phenomena, the literature review of Fu et al. (2012) points out that "the waiting time for boarding a transit vehicle at a stop/platform is an indispensable factor that needs special treatment." The estimation of passengers' waiting time becomes more complicated under scenarios with capacity constraints. In recent years, some multi-agent transit assignment models, e.g. BusMezzo, have been implemented to simulate the performance of public transit networks based on a set of capacity enhancement schemes (Cats 2013; Cats and Jenelius 2015).

Focusing on stations, previous studies propose various assumptions about boarding strategies to estimate the capacity constrained waiting time of passengers. The first assumption is that passengers follow a first-in-first-out (FIFO) discipline that establishes an ordered queue (Cominetti and Correa 2001; Hamdouch and Lawphongpanich 2008; Hamdouch et al. 2004a, b; Poon et al. 2004). In this case, the passengers' waiting time is modeled by means of bulk queuing. The other approach assumes that the passengers' wait at platform is mingled, meaning that if the capacity of the arriving vehicle is less than the number of candidate passengers, all these passengers have an equal probability of boarding (Kurauchi et al. 2003; Leurent and Chandakas 2012; Leurent et al. 2014; Schmöcker et al. 2011; Shimamoto et al. 2005). If the waiting passengers are mingled, the waiting time of a passenger is thus dependent on the probability of success-to-board (or failure-to-board).

Among the capacitated transit assignment models reviewed above, the main strategy is to minimize the waiting time and the main purpose is to find the network equilibrium,

whereas in this study, we aim mainly to find the distribution of waiting time under capacity constraints. To derive the waiting time distribution, one of the classical methods in the field of Queuing Theory is provided by Chaudhry and Templeton (1983). The study assumes that queues follow an M/G^B/1-FIFO system with the limiting behavior as t → ∞. After a series of mathematical transforms, the authors demonstrate that the Laplace-Stieltjes transform (L.-S.T.) of stationary waiting time distribution is dependent on (1) the limiting probability (as t → ∞) that there is no customer in queue; (2) the limiting probability (as t → ∞) that there are j customers remaining in the queue immediately after a departure epoch, where $j = 1, 2, 3, …,$ etc.; (3) the L.-S.T. of a given service time density; (4) Poisson arrival rate (λ); (5) mean service rate (μ); and, (6) maximum batch size, i.e. capacity (B); under the condition of $\lambda/B\mu < 1$. Based on their findings, the analytical solutions for two specific cases of M/G^B/1 queuing system, namely M/E$_k^B$/1 (service distribution being Erlang k) and M/D^B/1 (service distribution being deterministic), are also provided (Chaudhry 1991). As the conditions to obtain an analytical solution for a queuing system with bulk service are quite restrictive, the analytical solutions are not available for other queuing systems. Thus, some works study the M/G^B/1 systems based on numerical methods under more general scenarios (Dümmler and Vicari 1999; Glazer and Hassin 1987; Gold and Tran-Gia 1993).

None of these solutions can address the specific waiting time distribution of queuing in HSR stations. Analytical or numerical analyses of queuing system focus on a stationary condition (as t → ∞) with a system load less than 1, whereas the main interest of our work is to study queue accumulation at public transit stations when the travel demand may sometimes exceed service capacity. As discussed by Gross et al. (2008), due to the complexity of the transit systems, it becomes necessary to implement simulation techniques in order to study the performance of such a queuing system.

In "Methods" section, we introduce an approach to estimate the expected waiting time and the probability of forgoing trips, based on M/G^B/1 type of queuing scenario and discrete-event simulation. The impacts of capacity constraints with a specific focus on HSR station are then formally incorporated into the new accessibility measure: capacity constrained accessibility (CCA). To illustrate the differences between CCA and the standard measure, we use the Beijing–Tianjin corridor in China as the case study and describe the increasing demand and crowding in the Beijing–Tianjin Intercity high-speed railway (BTIHSR). "Results" section simulates the performance at HSR station under five queuing scenarios with various assumptions of passengers' arrival rules, train service distributions, and system loads. We compute the accessibility to Tianjin for a 10-km buffer area around Beijing South Station and compare the capacity-constrained and standard accessibility measures. "Discussion" section concludes with a discussion of the planning implications of the new CCA measure and future research directions.

Methods

Capacity constrained accessibility

The capacity of transit services is limited and passenger demand may exceed service capacity in certain conditions, entailing that some passengers suffer longer waiting time, or worse, may be unable to use the transit service at all. The consequences of longer waiting or inability to board vary depending on trip purpose, trip frequency, trip distance (or time),

service frequency, ticketing policies, etc. In the context of the Beijing–Tianjin HSR, a route for which the service frequency is high, if passengers cannot board the first train, they have to wait for subsequent train until they are able to board. The sole consequence at this stage is longer waiting time. However, if wait-to-board passengers accumulate and the waiting time keeps increasing, some passengers may forgo waiting and cancel their trips or use other alternative modes.

Longer waiting time due to limited capacity leads to longer total travel time and therefore larger time–space distance between the origin and destination. The underestimation of travel impedance by the standard accessibility measure results in the overestimation of the accessibility benefits of the infrastructure or service, which will then lead to biases in any follow-up studies, such as changes in land value and the relocation of economic activities. This paper, focusing specifically on the nodal capacity constraints, introduces a new measure, CCA, to provide a more realistic representation of the accessibility contribution by a transportation infrastructure or service.

We assume that a new infrastructure (or service) m, signifying HSR in this study but also inclusive of metro or a bus rapid transit, reduces the travel time from origin i to destination j. A passenger arrives at station s to wait for service m based on the FIFO rule. When queue length exceeds the train capacity, passengers at the end of the queue will be unable to board the train and will have to remain in the queue for subsequent train arrivals. The accessibility offered by m starts to decline for left-behind passengers. If the waiting time is overly long, certain passengers will forgo trips or switch to other modes. For these passengers, the accessibility offered by the new infrastructure m is neutralized, equivalent to a period when the infrastructure m was not available.
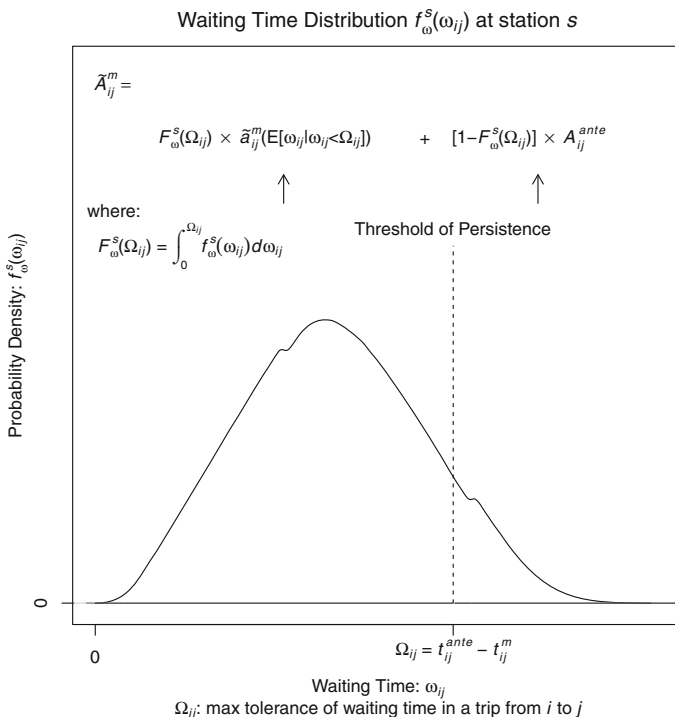

Fig. 1 Capacity constrained accessibility

To measure the accessibility from i to j via m with capacity constraints, we propose the following expression of CCA, as illustrated in Fig. 1:

$$\tilde{A}_{ij}^m = F_\omega^s(\Omega_{ij}) \times \tilde{a}_{ij}^m \left(E[\omega_{ij}|\omega_{ij} \le \Omega_{ij}] \right) + \left[1 - F_\omega^s(\Omega_{ij}) \right] \times A_{ij}^{ante}, \tag{2}$$

where $\tilde{A}_{ij}^m$ CCA from i to j via transit infrastructure or service m, ω_{ij} waiting time at station, denoted s, in a trip from i to j; F_ω^s the cumulative distribution function (CDF) of waiting time ω at station s; Ω_{ij} threshold of persistence, i.e. the maximum tolerance of waiting; $E[\omega_{ij}|\omega_{ij} \le \Omega_{ij}]$ expected waiting time of ω_{ij} given threshold of persistence Ω_{ij}; $\tilde{a}_{ij}^m(E[\bullet])$ accessibility via m as a function of expected waiting time $E[\bullet]$; A_{ij}^{ante} accessibility without m, e.g. accessibility before the inauguration of m.

As in Fig. 1, the probability of persistence when under threshold Ω_{ij}, $F_\omega^s(\Omega_{ij})$, is:

$$F_\omega^s(\Omega_{ij}) = \int_0^{\Omega_{ij}} f_\omega^s(\omega_{ij}) d\omega_{ij}, \tag{3}$$

where f_ω^s is the probability density function (PDF) of waiting time ω at station s. CCA is a weighted average of after-m and before-m accessibilities, with the probabilities of persistence and forgoing trips being the weights. By inserting Eq. (2) into the impedance function of Eq. (1), one gets a more general form of the location-based CCA measure:

$$\tilde{A}_i^m = \sum_{j=1}^J \left\{ g(O_j) \left[F_\omega^s(\Omega_{ij}) \times f\left(I_{ij}^m \left(E[\omega_{ij}|\omega_{ij} \le \Omega_{ij}] \right) \right) + \left(1 - F_\omega^s(\Omega_{ij}) \right) \times f\left(I_{ij}^{ante} \right) \right] \right\}, \tag{4}$$

where $\tilde{A}_i^m$ location-based CCA at origin i; $g(O_j)$ activity function of "opportunities" in destination j; $f\left(I_{ij}^{ante} \right)$ before-m impendence function from origin i to destination j $f\left(I_{ij}^m \left(E[\omega_{ij}|\omega_{ij} \le \Omega_{ij}] \right) \right)$ after-m impedance function incorporating $E[\omega_{ij}|\omega_{ij} \le \Omega_{ij}]$.

Assume that the travel times from i to j before and after the opening of m are t_{ij}^{ante} and t_{ij}^m, respectively, calculated based on standard infrastructure-based methods of calculating origin–destination (OD) travel time. Procedures for OD travel time calculation without capacity constraints are explicated in Appendix 1. With known t_{ij}^{ante} and t_{ij}^m, A_{ij}^{ante} and $\tilde{a}_{ij}^m$—without taking into account $E[\omega_{ij}|\omega_{ij} \le \Omega_{ij}]$—can be easily obtained. In this study, the threshold of persistence Ω_{ij} is defined as

$$\Omega_{ij} = t_{ij}^{ante} - t_{ij}^m. \tag{5}$$

In Eq. (2), the trickiest component to estimate is the waiting time distribution, F_ω^s. As in Fig. 1, the weights (i.e. probability of persistence) and the expected waiting time $E[\omega_{ij}|\omega_{ij} \le \Omega_{ij}]$ are both determined by the specific waiting time distribution.

Waiting time distribution and probability of persistence

The waiting time distribution of passengers is determined by passengers' arrival distribution and trains' service headway distribution. The queuing dynamics are shown in Fig. 2. As an example, the queue length starts to increase from t_0, and the trains arrive at time t_1, t_2, t_3, etc. If the lengths of queues at t_1, t_2, and t_3 are no greater than the capacity

B of the trains, all passengers are transported. If the length of the queue at t_4 is greater than the train's capacity, only B passengers are able to board, and the remaining passengers stay in the queue for the train arriving at t_5.

There is no closed form for the waiting time distribution in this application context. To simulate the queuing dynamics, this study applies a discrete-even stochastic simulation method suggested by Gross et al. (2008) with three major elements: (1) input modelling and generation, (2) bookkeeping of the queuing process, and (3) output analysis. The discrete-event simulation is programmed in Java based on AnyLogic simulation platform, and the outputs are recorded in SQL and analyzed in R.

The simulation system has two categories of inputs: passenger arrival headway and train service headway. Previous studies indicate that passenger's arrival incidence at transit stations appears to be random below a 10-min service headway (Abkowitz and Tozzi 1987; Fan and Machemehl 2002). Thus, if the service headway is less than or equal to 10 min, we assume that passengers arrive at station s randomly and singly, following the Poisson distribution. Therefore, the arrival headway in the simulation is randomly generated from an exponential distribution. We acknowledge the limitation of this idealization of the arrival process, since real-world passenger arrival (or the incidence behavior) at a transit station is proven to be a more complicated phenomenon than the Poisson process; it may be also influenced by the transit network, times of day, and service reliability, among other factors (Bates 2001; Frumin and Zhao 2012). For the other input, train service headway, we test two different distributions of service headways: uniform headway and real, schedule-based service headway with and without disturbances. We assume that the queuing system
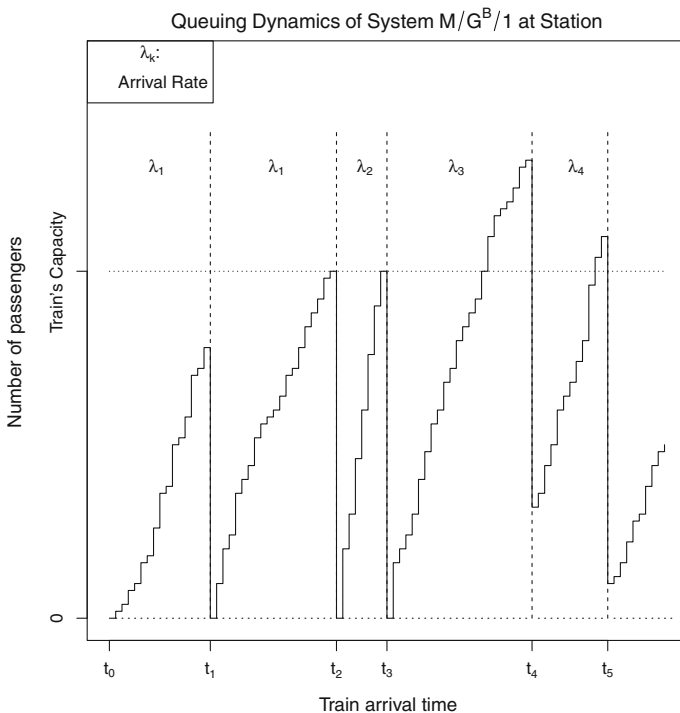

Fig. 2 Queuing dynamics with bulk service

has one server with an unlimited queue capacity (i.e. the station is large enough to accommodate all waiting passengers), and the service rule is FIFO.

With the predefined headway distribution t_H and service capacity B, the simulation tests the queuing performance with various system loads ρ from 70 to 115 %. For each system load ρ, the arrival rate λ is calculated by:

$$\lambda = \frac{\rho B}{t_H}. \tag{6}$$

During the simulation, when passenger P_n arrives at the station, the system prints the arrival time of P_n and simultaneously prints the length of the queue immediately after the arrival of P_n. When a train arrives and collects passengers, up to B passengers will leave the system. If P_n is located within the first B passengers, P_n leaves the queue and the model prints the waiting time of P_n as the difference between its arrival time and departure time. The simulation starts the bookkeeping after 30 min to allow a warm-up period, and keep recording the waiting time of each passenger and the queuing performance at the station for the next 240 min to simulate a 4-h peak period. The simulation stops at 400 min to make sure that all passengers P_n are able to leave the queue when the simulation ends. For each service headway and each system load, the simulation is replicated 1000 times.

Case: Beijing–Tianjin intercity high-speed railway

In addition to the large network coverage and high speed, what really differentiates China's HSR system is its frequency: along major corridors, such as the ones between Beijing and Tianji, and Shanghai and Nanijing, the peak hour headways are often below 10 min. The Beijing–Tianjin route operates with a 10-min average headway during peak hour now; the Shanghai–Nanjing route operates with a 7-min headway. These HSR lines operate as a high-frequency public transit service in general. Furthermore the wide adoption of "Transit IC card" and its integration between HSR and urban transit enable many passengers to use HSR just like a subway: they no longer time their arrivals according to the schedule, instead they simply show up at the train station and swipe their IC cards.

Beijing–Tianjin intercity high-speed railway (BTIHSR), inaugurated in August 2008, is one of the first passenger-only HSRs in China. The geographical location of this intercity rail service is shown in Fig. 3. It runs 117-km route between Beijing South Railway Station and Tianjin Railway Station, with a connection to Wuqing Station in Tianjin. Two additional intermediate stations in Beijing suburbs, Yizhuang and Yongle, are expected to be opened in future. The two termini are well integrated into the local public transit systems of each city. The Beijing South Station is one of the largest railway complexes in Asia, acting as an interchange for Beijing Subway Lines 4 and 14, national railway, and HSR service to Shanghai. The Tianjin Railway Station is also one of the major stations in China, with local connections to Tianjin Metro Line 2, 3 and 9, and national railway connections to Shanghai and Harbin.

The inter-city high-speed service travels at a maximum speed of 300 km/h, reducing the travel time between Beijing and Tianjin from ~ 70 min to ~ 35 min over conventional intercity rail. The bullet trains serving this line, China Railways CRH3 s, have about 550 seats, and passenger ticketing is restricted to the number of seats. The February 2014 morning timetable from the Ministry of Railways (MOR) is shown in Table 5 in Appendix 2. The schedule shows that, average peak-hour headway (from 06:35 to 08:31) from Beijing South to Tianjin is about 10 min and 35 s, with a minimum headway of 5 min. In

the future, the daily average service headway is expected to be reduced to 3 min. There are two major classes of tickets: first-class seats, the price of which is 66 Chinese Yuan (about US$10), and second-class seats, for 55 Chinese Yuan (about US$8).

The operation of HSR in China along corridors of rapid economic growth, such as the Beijing–Tianjin mega-region, attracts a large number of passengers (Wang et al. 2013). According to the MOR, since 2009, the average seat occupancy rate (i.e. loading rate) of the BTIHSR is over 70 %. During weekends and the "Golden Week" (a week-long national holiday), the seat occupancy rate often surges to 100 % (ChinaNews.com 2009).
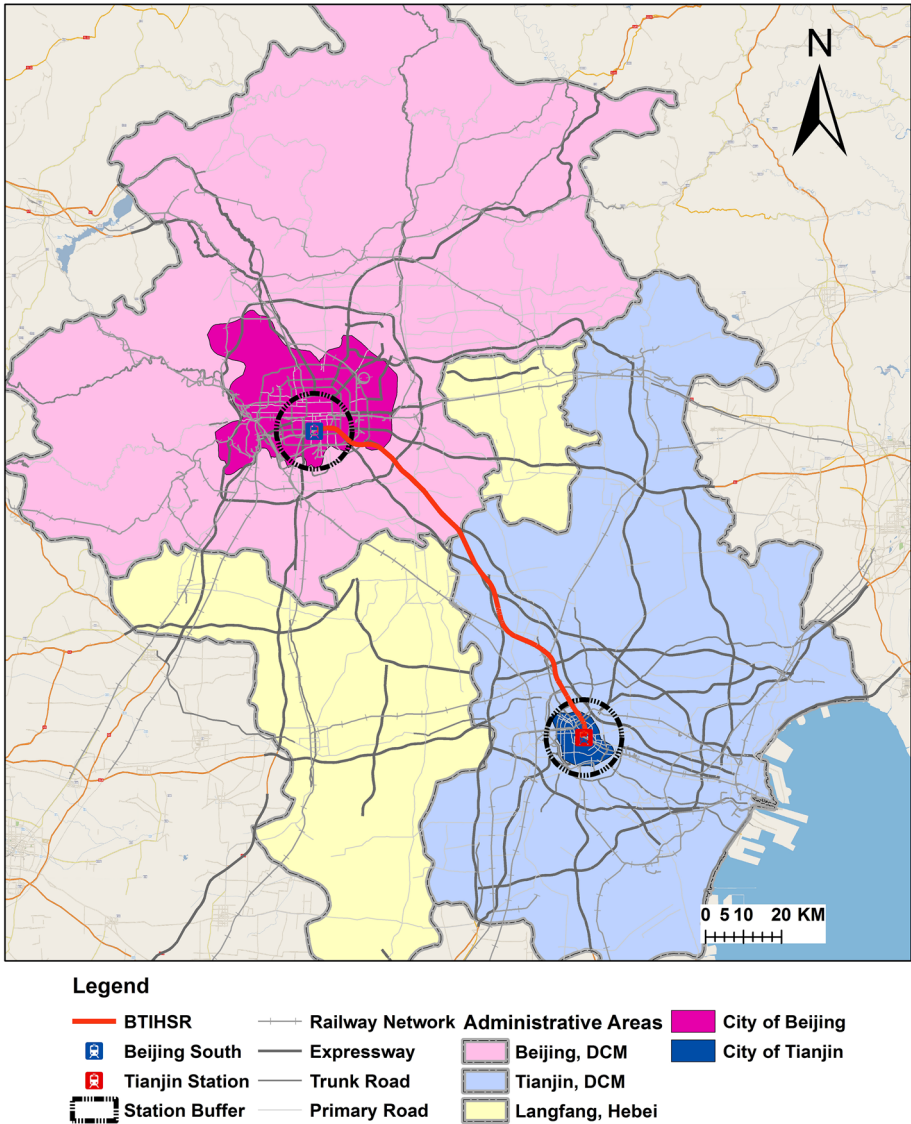

Fig. 3 Geography of Beijing–Tianjin Intercity Railway. (Created by the authors based on data from OpenStreetMap.org). *Note DCM* direct-controlled municipality

As of 2013, passenger use of the BTIHSR had grown at an annual rate of 20 % since its opening. Before the BTIHSR, annual ridership between these cities on conventional rail was about 8 million (Bullock et al. 2012). Twelve months after the inauguration of this line, the annual passenger volume increased to 18.70 million, and reached 22.26 million by August 2010, 24 months after the inauguration (Yin et al. 2014). The BTIHSR has become one of the most important travel modes for daily commuting between Beijing and Tianjin. Many people who live in Beijing choose to commute to Tianjin to work every day (Tan 2014). Although the MOR keeps updating the operation schedule of BTIHSR by shrinking its service headways, as travel demand on this line booms, service capacity is a critical constraint, especially during peak hours, weekends, and holidays.

Results

Five queuing scenarios in HSR station

The service of BTIHSR with queuing at Beijing South Railway Station is close to an ideal queuing system, fitting our assumptions. As one of the largest railway complexes in Asia, the waiting area in this station is very large, and the resulting capacity constraint on queue length is trivial. The single-line HSR network between Beijing and Tianjin is also quite simple. Thus, the prevailing bottleneck is the boarding of the train, and bottlenecks are unlikely to occur in other operational aspects. Since passenger ticketing is restricted to the number of seats and Beijing South Station is a terminal station for this HSR service, a fixed train capacity (B) of 550 can be set in our simulation. From each simulation, 10–15 million records of waiting time and queue length are produced, depending on the demand. For each service headway and each system load, the simulation is replicated 1000 times and an empirical probability density function (EPDF) is obtained. For each OD (from i to j), given a threshold of persistence Ω_{ij}, we calculate the probability of persistence and the expected waiting time $\mathrm{E}\left[\omega_{ij} | \omega_{ij} \leq \Omega_{ij}\right]$.

Five queuing scenarios with varying assumptions of arrival and service distributions are simulated for the analysis of the queuing performance at the Beijing South HSR station, where Scenario 1 is the least realistic scenario and Scenario 5 is the most realistic scenario based on our available data. The key characteristics of each scenario—service headway, disturbance of headway, passengers' arrival rate, and system loads—are summarized in Table 1.

In Scenario 1, we assume the queuing system follows an $M/D^B/1$ system, in which the service headway is exact 10 min without disturbance. Theoretically, if passengers arrive at the station independently and randomly without capacity constraints, the PDF of waiting time follows:

$$f_{\omega}^{s}\left(\omega_{ij}\right) = \begin{cases} \dfrac{1}{10 - 0} & \text{for } \omega_{ij} \in [0, 10] \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

As there is a known closed form of the unconstrained waiting time distribution, with this scenario we are able to validate that the simulated distribution is consistent with the analytical one.

In Scenario 2, holding all other assumptions constant, we change the train's headway to follow the February 2014 timetable of BTIHSR (Table 5; Appendix 2), forming an

Table 1 Summary of scenarios

Scenario	Service headway	Disturbance	Arrival rate	Load
1	10 min	No	Constant	Constant across the entire period
2	Based on the actual timetable	No	Subject to the service headway variation	Constant across the entire period
3	Based on the actual timetable	No	Ignore the service headway variation	Constant across the entire period
4	Based on the actual timetable	Triangular, max 2 min	Ignore the service headway variation	Constant across the entire period
5	Based on the actual timetable	No	Ignore the service headway variation	Surpasses 100 % at the first 2 h, then decreases to keep the average ρ less than 100 %.

M/Sch^B/1 systems. In this scenario, the passengers' arrival rate is assumed to accord with the service rate, i.e. passengers are aware of the headway variations and their demands change accordingly. For instance, assume $\rho = 100$ %. Between 9:01 and 9:06 the service headway is 5 min, meaning that an expected number of $\rho \times B = 550$ passengers will arrive at the station within this 5-min slot to take the 9:06 train. The next train departs at 9:19, meaning that the same expected number of passengers will arrive at the station within the 13-min slot.

In Scenario 3, holding all other assumptions of the second one constant, the arrival rate remains the same across the entire simulation period (i.e. passengers ignore the service headway variation), which is computed based on the 4-h average service headway from 6:35 to 10:29, about 10.65 min. Based on the findings of the previous studies mentioned in "Waiting time distribution and probability of persistence" section, the passengers' incidence behavior becomes random when the headway is 10 min or less. According to the MOR, the service headway of BTIHSR will be further reduced in the future. Therefore, we believe that this scenario is more realistic than Scenario 2 in that the passengers are more likely to arrive at the station regardless of the actual timetable.

In Scenario 4, holding all other assumptions of the third one constant, we introduce disturbance to the service headway. In this scenario, we allow the train's arrival times to deviate from the time table following a triangular distribution (with a maximum 2 min deviance from the timetable). The aim of this scenario is to test the waiting time distribution when the service is unreliable.

Scenario 5 is an extension of Scenario 3 but mimics a quasi-real scenario of passenger arrivals during morning hours. As the punctuality rate of BTIHSR is 98 % (ITourBeijing.com 2015) we assume that the service follows the timetable exactly. Instead of setting a uniform ρ, this scenario allows the build-up of demand at a higher ρ during a 2-h peak period from 30 to 150 min, then the system load declines to a lower percentage until the end of the simulation. As the average seat occupancy rate of BTIHSR is reported as 70 %, the simulation starts at 70 % for the warm-up period; 30 min later, ρ surpasses 100 % for 120 min, after which ρ decreases to keep the average ρ at 70 %. The scenario tests the peak hour demands at four levels, from 105 to 120 %.

Simulated probability density functions of waiting time from the 5 scenarios

Scenario 1   An M/D^B/1 queuing system with equal service headway
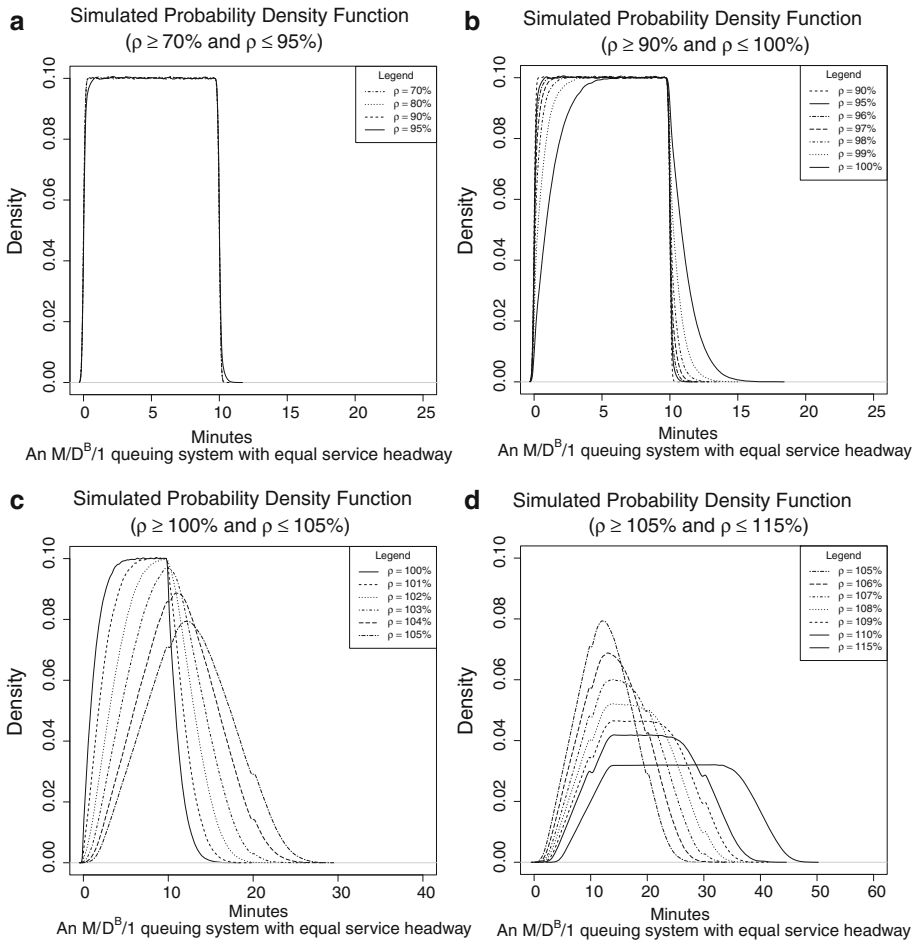
Fig. 4 Simulated probability density function of waiting time in scenario 1

The EPDFs of this scenario are shown Fig. 4. When system load ρ is less than 90 % (Fig. 4a), the EPDF is consistent with Eq. (7) and validates the simulation results in the low system load scenario. However, when ρ becomes more than 95 % (Fig. 4b), the EPDF starts to deviate from Eq. (7). When ρ is greater than 100 % (Fig. 4c, d), the EPDF becomes significantly different from (6). Table 2 reports the probability of persistence and the expected waiting time for any given waiting time budget based on the queuing simulation. With an unlimited threshold of persistence, when ρ is 100 %, the unconditional expected waiting time $E[\omega_{ij}]$ is 6.1 min. With only a 2 % increase of ρ ($\rho = 102$ %), $E[\omega_{ij}]$ becomes 8.3 min. When ρ reaches 110 %, $E[\omega_{ij}]$ grows to more than 20 min.

Scenario 2   An M/Sch^B/1 queuing system with a schedule-based arrival rate

The EPDFs are shown in Fig. 5. Under this scenario with a relatively lower ρ (as in Fig. 5a), passenger arrivals follow the service timetable, and within each headway the waiting time exhibits a uniform distribution. For instance, the minimum headway during

Table 2 Probability of persistence and expected waiting time for M/D^B/1 with 10 min headway

Probability of persistence

$\rho=$	70 %	90 %	95 %	100 %	102 %	104 %	106 %	108 %	110 %
Threshold of persistence (in min)									
1	0.100	0.100	0.097	0.036	0.006	0.001	0.000	0.000	0.000
2	0.200	0.200	0.196	0.106	0.028	0.006	0.001	0.000	0.000
3	0.300	0.300	0.297	0.194	0.067	0.020	0.006	0.002	0.000
4	0.400	0.400	0.397	0.289	0.125	0.044	0.018	0.008	0.003
9	0.900	0.900	0.897	0.788	0.569	0.320	0.188	0.121	0.080
10	1.000	1.000	0.997	0.888	0.668	0.404	0.243	0.160	0.109
11	1.000	1.000	1.000	0.952	0.762	0.491	0.302	0.201	0.139
12	1.000	1.000	1.000	0.982	0.841	0.579	0.368	0.247	0.174
13	1.000	1.000	1.000	0.994	0.901	0.661	0.436	0.297	0.212
14	1.000	1.000	1.000	0.998	0.943	0.735	0.504	0.349	0.253
15	1.000	1.000	1.000	1.000	0.971	0.800	0.571	0.401	0.295
16	1.000	1.000	1.000	1.000	0.987	0.855	0.636	0.453	0.337
19	1.000	1.000	1.000	1.000	0.999	0.958	0.804	0.607	0.462
20	1.000	1.000	1.000	1.000	1.000	0.974	0.848	0.657	0.504
24	1.000	1.000	1.000	1.000	1.000	0.999	0.969	0.840	0.670
25	1.000	1.000	1.000	1.000	1.000	1.000	0.983	0.876	0.711
26	1.000	1.000	1.000	1.000	1.000	1.000	0.991	0.907	0.751
27	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.933	0.789
28	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.954	0.825
29	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.969	0.858
30	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.979	0.886
34	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.973
35	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.984
36	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.991
40	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
41	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
42	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Expected waiting time

$\rho=$	70 %	90 %	95 %	100 %	102 %	104 %	106 %	108 %	110 %
Threshold of persistence (in min)									
1	0.50	0.50	0.51	0.60	0.65	0.70	0.80	0.82	0.96
2	1.00	1.00	1.01	1.22	1.37	1.49	1.63	1.67	1.73
3	1.50	1.50	1.52	1.80	2.06	2.24	2.43	2.53	2.58
4	2.00	2.00	2.02	2.37	2.73	2.96	3.16	3.31	3.45
9	4.50	4.50	4.52	4.98	5.79	6.35	6.59	6.78	6.99
10	5.00	5.00	5.02	5.49	6.35	7.00	7.26	7.45	7.66
11	5.00	5.00	5.03	5.83	6.86	7.62	7.90	8.07	8.27
12	5.00	5.00	5.03	6.00	7.29	8.21	8.54	8.72	8.91
13	5.00	5.00	5.03	6.08	7.63	8.75	9.16	9.36	9.57
14	5.00	5.00	5.03	6.11	7.90	9.22	9.75	9.97	10.21

Table 2 continued

Expected waiting time

$\rho=$	70 %	90 %	95 %	100 %	102 %	104 %	106 %	108 %	110 %
15	5.00	5.00	5.03	6.12	8.08	9.65	10.30	10.56	10.82
16	5.00	5.00	5.03	6.12	8.20	10.02	10.83	11.12	11.40
19	5.00	5.00	5.03	6.12	8.31	10.81	12.21	12.74	13.05
20	5.00	5.00	5.03	6.12	8.32	10.95	12.59	13.26	13.58
24	5.00	5.00	5.03	6.12	8.32	11.21	13.72	15.14	15.67
25	5.00	5.00	5.03	6.12	8.32	11.21	13.88	15.53	16.18
26	5.00	5.00	5.03	6.12	8.32	11.22	13.98	15.87	16.68
27	5.00	5.00	5.03	6.12	8.32	11.22	14.04	16.16	17.15
28	5.00	5.00	5.03	6.12	8.32	11.22	14.07	16.40	17.60
29	5.00	5.00	5.03	6.12	8.32	11.22	14.08	16.59	18.02
30	5.00	5.00	5.03	6.12	8.32	11.22	14.09	16.73	18.39
34	5.00	5.00	5.03	6.12	8.32	11.22	14.09	17.02	19.57
35	5.00	5.00	5.03	6.12	8.32	11.22	14.09	17.03	19.74
36	5.00	5.00	5.03	6.12	8.32	11.22	14.09	17.04	19.86
40	5.00	5.00	5.03	6.12	8.32	11.22	14.09	17.04	20.00
41	5.00	5.00	5.03	6.12	8.32	11.22	14.09	17.04	20.01
42	5.00	5.00	5.03	6.12	8.32	11.22	14.09	17.04	20.01

the simulated hours is 5 min, and below the capacity (i.e. the dashed lines in Fig. 5a) we can find a period of uniform distribution from 0 to 5 min. The longest headway is 19 min (from 8:42 to 9:01), and the distribution also shows the longest waiting time as 19 min. The aggregation of these uniform distributions becomes step-shaped.

Scenario 3   An M/Sch^B/1 queuing system with uniform arrival rate

The EPDFs can be found in Fig. 6. The passengers are assumed to neglect the timetable. When the service rate (e.g. with 5-min headway) is higher than the constant arrival rate ($\lambda = \rho B/10.65$), no passengers are likely to be left behind, but the train may depart with an actual seat occupancy rate lower than load ρ. When the service rate (e.g. with 15-min headway) is lower than arrival rate, a queue may still accumulate, even with ρ of less than 95 %. Therefore, in Fig. 6a, with full service capacity (the dashed lines), although the longest waiting time is still 19 min, the distribution is no longer step-shaped. When ρ increases to more than 90 %, even though the train is not fully loaded, the longest waiting may be up to 23 min.

Scenario 4   An extension of Scenario 3 with unreliable timetable

The EPDFs are shown in Fig. 7. The comparison between Scenario 3 and 4 shows that when the service timetable is unreliable, the maximum waiting time becomes much larger under a constant ρ. When ρ increases, the change of the EPDF with unreliable timetable is more significant than that with punctual train arrivals. The finding that reliability is an important factor influencing waiting time is consistent with past studies (Bates 2001). If capacity constraints and unreliable service occur together, the situation becomes much worse.

Scenario 5　　　　　　　A quasi-real scenario with variable system loads

Figure 8a shows that even when the average ρ is constant, if peak hour ρ is higher, the maximum and expected waiting times become much greater. Importantly, the results show that even an overall system load of 70 % can still be constrained by service capacity because of the demand concentration during the peak hours. Figure 8b, c, d show the (PDF) of waiting times when the average load of BTIHSR increases to ρ = 80, 90 and 100 %, which represent the cases during weekends, holidays, or other high demand periods of the year.

Accessibility with and without capacity constraints

To illustrate and to visualize the changes of CCA, the origin–destination (OD) matrix is defined as the travel time from the 10-km buffer area of Beijing South Station to the 10-km buffer area of Tianjin Station. For each buffer area, 1324 square cells of 500 × 500 meters
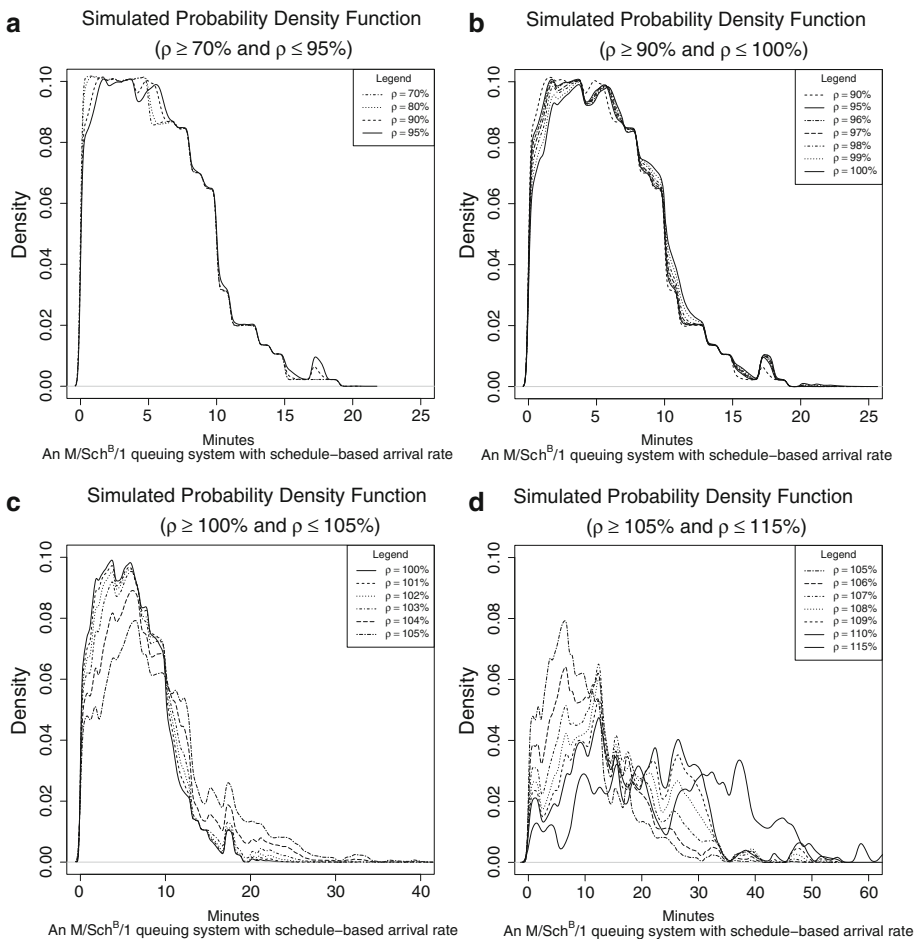
Fig. 5 Simulated probability density function of waiting time in scenario 2
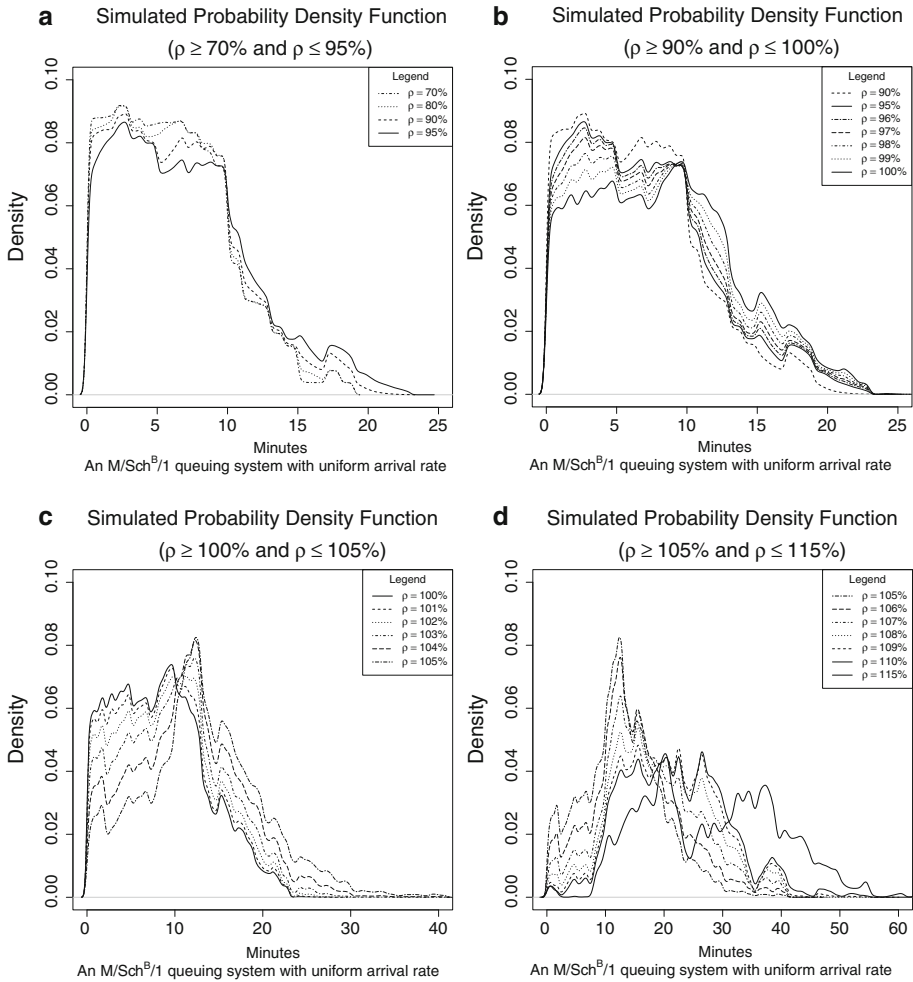
Fig. 6 Simulated probability density function of waiting time in scenario 3

are tessellated, forming an OD matrix of 1324×1324. The cell's size is consistent with the spatial resolution of MODIS C5 urban land-cover data (Schneider et al. 2009, 2010).

Based on our definition of CCA in Eq. (2), we can calculate and compare the CCA of the BTIHSR using the standard accessibility measure without capacity constraints. As there are no closed analytical solutions for $M/G^B/1$ queuing, the CCA is calculated based on the simulated PDF of waiting time, by looking up the probability of persistence and expected waiting time (PPEWT) in a table created under the five abovementioned scenarios, similar to Table 2. The CCA of origin i is expressed as:

$$\tilde{A}_i^{HSR} = \kappa \cdot \frac{1}{N_j} \cdot \sum_{j=1}^{N_j} \tilde{A}_{ij}^{HSR}, \tag{8}$$

where $\tilde{A}_i^{HSR}$ CCA at origin cell i, i.e. Beijing South Station's buffer area; N_j total number of land parcel cells in destination j, i.e. 1324; κ: empirical scale parameter to keep the value
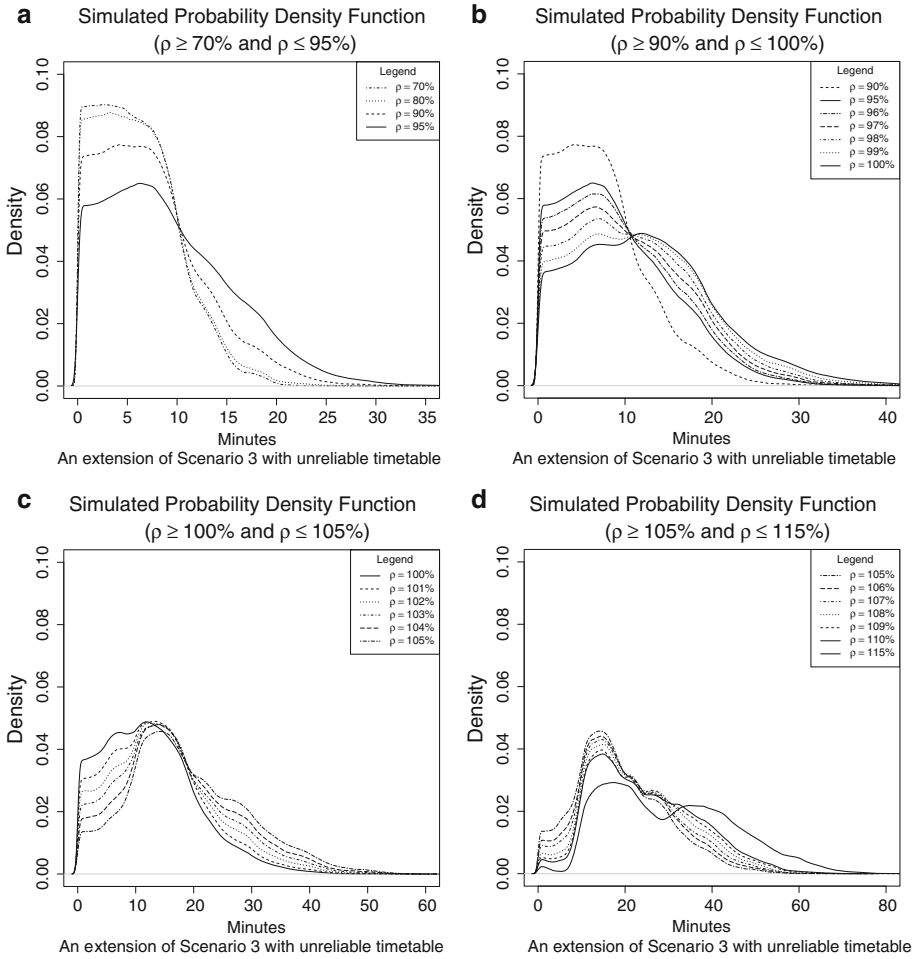
Fig. 7 Simulated probability density function of waiting time in scenario 4

of $\tilde{A}_i^{HSR}$ around 1, set to 100; $\tilde{A}_{ij}^{HSR}$ CCA from origin i to destination cell j, i.e. Tianjin Station's buffer area, which is calculated as:

$$\tilde{A}_{ij}^{HSR} = g(O_j)\left\{\text{Prob}(\omega_{ij}|\omega_{ij}\leq\Omega_{ij})\cdot\tilde{t}_{ij}\big(\text{E}[\omega_{ij}|\omega_{ij}\leq\Omega_{ij}]\big) + [1-\text{Prob}(\omega_{ij}|\omega_{ij}\leq\Omega_{ij})]\cdot t_{ij}^{ante}\right\}^{-1},$$
(9)

where $g(O_j)$ activity functions of opportunities at j. ω_{ij} waiting time of persistence of OD pair i to j; $\tilde{t}_{ij}(\bullet)$ travel time from i to j with queuing at HSR station, as a function of expected waiting time with queuing; t_{ij}^{ante} travel time from i to j before the opening of HSR; Prob($\bullet$) probability of persistence, which is obtained by looking up the table; $\text{E}[\omega_{ij}|\omega_{ij}\leq\Omega_{ij}]$ expected waiting time at HSR station in a trip from i to j, based on the threshold of persistence Ω_{ij}, obtained from the look-up table. The threshold of persistence Ω_{ij} is calculated based on Eq. (5).
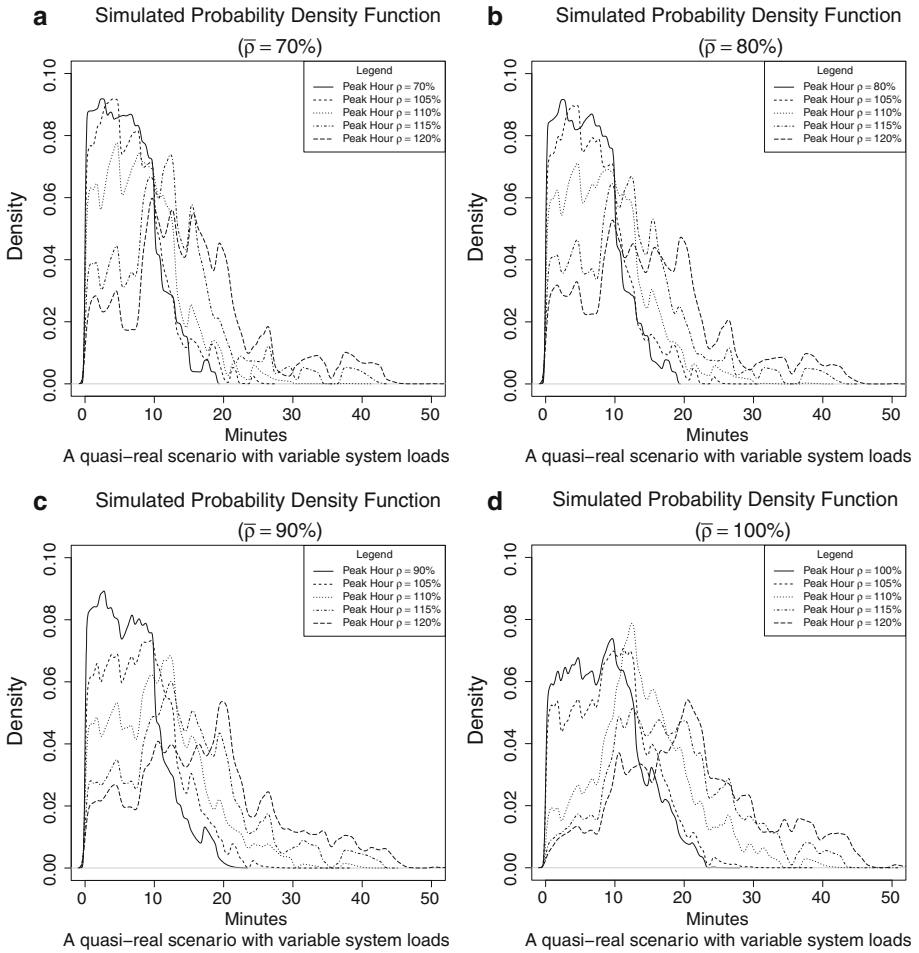
Fig. 8 Simulated probability density function of waiting time in scenario 5

Since this study focuses on the impacts of time–space separation due to capacity constraints from public transit, we do not go into details of the activity function. Thus, $g(O_j)$ is normalized to 1 and is removed from the equation. The calculation of CCA is based on the PPEWT table (e.g. Table 2) generated in each scenario. Given a particular threshold of persistence for each OD pair from i to j, obtained from Eq. (5), the CCA can be calculated by looking up the PPEWT table. Take Scenario 1 as an example, with $\rho = 90\ \%$. Assume Ω_{ij} for OD pair i to j is 15 min. According to the table, all passengers will wait for the train and the expected waiting time $E[\omega_{ij}|\omega_{ij} \leq 15]$ is 5 min. Equation (2) can be rewritten as:

$$\tilde{A}_{ij}^{HSR} = 1.0 \times \tilde{a}_{ij}^{HSR}\left(E\left[\omega_{ij}\big|\omega_{ij} \leq 15\right]\right) + 0 \times A_{ij}^{ante} = A_{ij}^{HSR}, \text{ where } E\left[\omega_{ij}\big|\omega_{ij} \leq 15\right] = 5,$$
(10)

which is equivalent to the standard accessibility measures. When ρ increases to 104 %,

only 80 % passengers would persist in waiting based on the 15-min Ω_{ij}. And the expected waiting time increases to 9.65 min. For the other 20 %, accessibility declines to the before-BTIHSR situation. For the 80 % of passengers we need to use the expected waiting time $\mathrm{E}\left[\omega_{ij}\middle|\omega_{ij}\le 15\right] = 9.65$ min in their accessibility calculation. In this case, the CCA is thus calculated as:

$$\tilde{A}_{ij}^{HSR} = 0.8 \times \tilde{a}_{ij}^{HSR}\left(\mathrm{E}\left[\omega_{ij}\middle|\omega_{ij}\le 15\right]\right) + 0.2 \times A_{ij}^{ante}, \text{ where } \mathrm{E}\left[\omega_{ij}\middle|\omega_{ij}\le 15\right] = 9.65. \quad (11)$$

Figure 9a shows the accessibility patterns before the opening of BTIHSR. The calculation of OD travel time follows the discussions in Appendix 1. Since Tianjin is located in the southeast of Beijing, travel via road network was generally fastest before the inauguration of the BTIHSR, and the travel time from southeastern cells, especially those close to ring roads and motorways, is shorter than that from other places. After the opening of the BTIHSR, the travel time from central areas becomes the shortest without considering capacity constraint, as shown in Fig. 9b. Figure 9c through 9h show the accessibility patterns in Scenarios 2–5 based on the CCA measures.

The aggregate value of CCA is calculated as:

$$\tilde{A}^{HSR} = \frac{1}{N_i} \sum_{i=1}^{N_i} \tilde{A}_i^{HSR} \quad (12)$$

The calculations of the aggregate standard accessibility after the opening of HSR A^{HSR} and the aggregate before-HSR accessibility A^{ante} follow the same idea. To quantify the difference between the new CCA measure and the standard measure in terms of the accessibility improvement thanks to the BTIHSR, we calculate the overestimation ε by the standard measure relative to the CCA as

$$\varepsilon = \frac{\Delta A}{\Delta \tilde{A}} - 1 = \frac{A^{HSR} - A^{ante}}{\tilde{A}^{HSR} - A^{ante}} - 1 \quad (13)$$

Table 3 reports the aggregate accessibility from Beijing South Station's buffer area to Tianjin and the overestimation error of the standard measure in each scenario.

Figure 9c is almost identical as Fig. 9b and the value of $\tilde{A}^{HSR}$ (1.215) is also very close to unconstrained accessibility (1.214). With a low system load, CCA is equivalent to the standard measure. Thus, CCA is compatible with and absorbs the standard accessibility measure as a special case where capacity constraints are not binding.
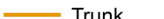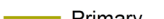
Figure 9d shows the results in Scenario 3 with $\rho = 100$ % of with lighter colors of many cells than in Fig. 9b. The aggregate $\tilde{A}^{HSR}$ decreases from 1.214 to 1.178, and in this case the standard measure overestimates the accessibility improvement by 16.7 %. When the timetable becomes less reliable in scenario 4 (Fig. 9e), the CCA decreases even further to 1.133. Under an extreme situation in scenario 4 with $\rho = 115$ % and unreliable timetable (Fig. 9f), the aggregate accessibility declines greatly to 1.015, which almost triples the overestimation error ε. With unreliable timetable and high loads ρ together, the value accessibility becomes much lower. As the results, Fig. 9f looks more similar to Fig. 9a than Fig. 9b, i.e. for much of the study area, the accessibility nearly declines to the situation before the opening of BTIHSR.

Figure 9g, h depict the CCA patterns of two quasi-real situations in Scenario 5. In Fig. 9g, the average load ρ is set at 70 % and the 2-h peak load ρ_{peak} is set at 110 % to represent a typical morning. A comparison of Fig. 9g, b, in which $\tilde{A}^{HSR}$ equal 1.181 and

Fig. 9 Accessibility to Tianjin from Beijing South Railway Station's Buffer Area
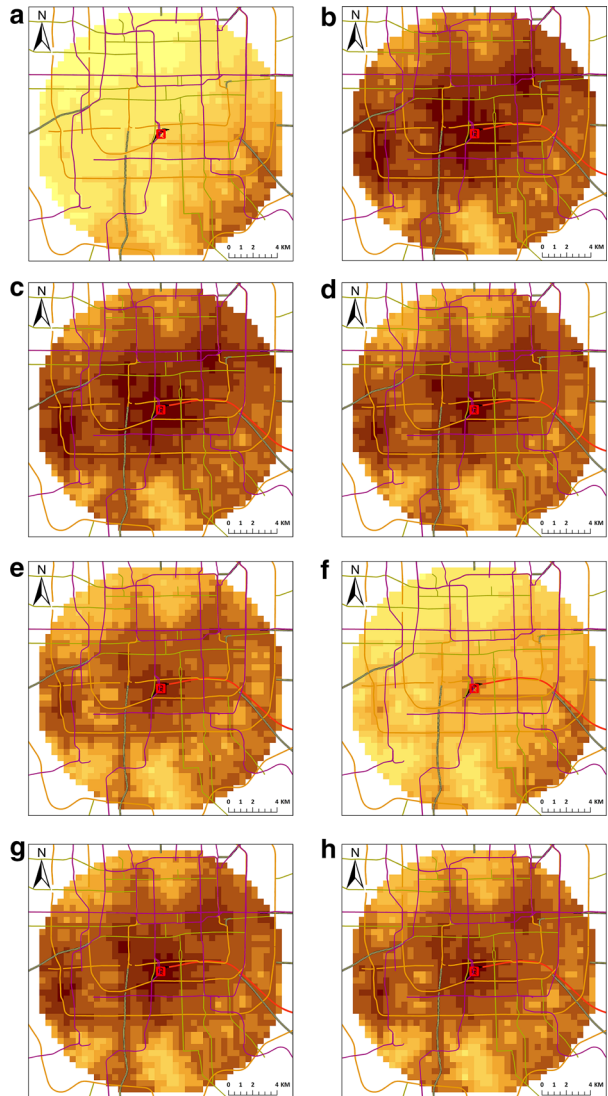
Legend

Beijing South Station
Beijing-Tianjin HSR
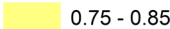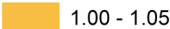Beijing Subway
Motorway
Trunk
Primary

Accessibility
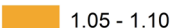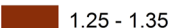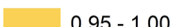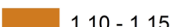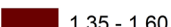
0.75 - 0.85
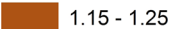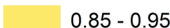0.85 - 0.95
0.95 - 1.00
1.00 - 1.05
1.05 - 1.10
1.10 - 1.15
1.15 - 1.25
1.25 - 1.35
1.35 - 1.60

Table 3 Aggregate accessibility and overestimation in each scenario

Figure 9a–h	Accessibility $\tilde{A}^{HSR}$	Overestimation ε
9a: Accessibility before HSR	0.948	N.a.
9b: Unconstrained accessibility after HSR	1.214	N.a.
9c: Scenario 2, $\rho = 70\ \%$	1.215	−0.4 %
9d: Scenario 3, $\rho = 100\ \%$	1.178	16.7 %
9e: Scenario 4, $\rho = 100\ \%$	1.133	43.8 %
9f: Scenario 4, $\rho = 115\ \%$	1.015	297.0 %
9 g: Scenario 5, average $\rho = 70\ \%$	1.181	14.2 %
9 h: Scenario 5, average $\rho = 90\ \%$	1.153	29.8 %

N.a. not applicable

1.214 respectively, shows that even under a moderate average load of 70 % accessibility can suffer if the system is saturated during peak hours. If the average load increases to 90 %, such as on weekends and holidays, $\tilde{A}^{HSR}$ decreases to 1.153, as shown in Fig. 9h. The visual difference between Fig. 9h, b becomes obvious. The overestimations of accessibility by the standard measure are respectively 14.2 and 29.8 %, which are substantial errors for any practical purposes in the accessibility applications.

Applying Eq. (10) at the individual level, we can calculate the overestimation of the accessibility changes for each cell and examine the variation in the errors. Figure 10 plots the distributions of the overestimation errors for the two quasi-real situations in Scenario 5. The solid line shows the overestimation ranges between 10 and 30 %. When ρ increases to 90 %, show as a dotted line, the overestimation can be as high as 70 %. By ignoring the capacity constraint, the standard measure can result in huge errors of estimation beyond any reasonably acceptable level.

Discussion

Prior accessibility measures rarely take into account the transit network's operational details. In this paper, we present one such detail, capacity constraint, which is too important to ignore even at the planning and strategic stages of the accessibility discussion. In densely-populated areas where passenger demand is high, travel time is not only determined by the availability of infrastructure, but also constrained by service capacity. The longer waiting times and forgone trips resulting from capacity constraints can significantly change the accessibility patterns.

We bridge the gap between the accessibility literature and the transit network assignment literature to propose a new accessibility measure that incorporates the capacity constraints of a transit network. We formalize the measure of CCA as the weighted average of accessibilities before and after the new infrastructure. The weights, the probability of forgoing trips and the probability of persistence, are determined by the threshold of persistence and distribution of waiting time as simulated by the $M/G^B/1$ queuing model.

To illustrate the difference between the standard measure and CCA measure, we examine the accessibility change in the Beijing–Tianjin mega-region resulting from the introduction of the BTIHSR by using five queuing scenarios with varying assumptions about arrival rules, service distributions, and system loads. The case study shows that,
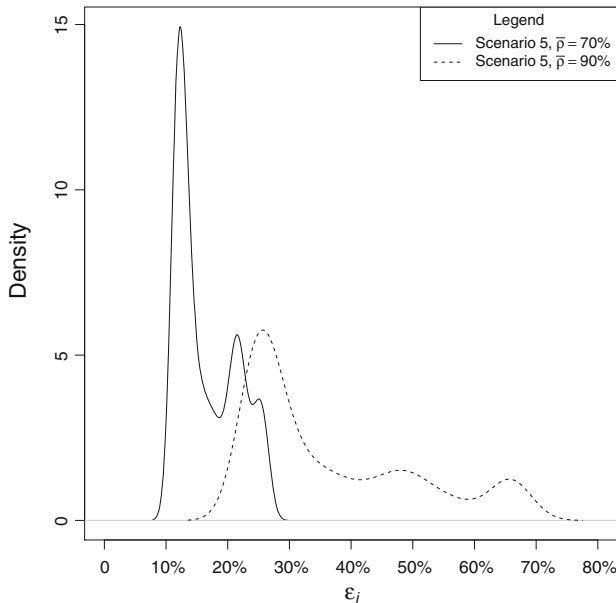
Fig. 10 Distribution of overestimation errors by the standard measure

under low system load, CCA is compatible with and absorbs the standard accessibility measure as a special case. If the system is saturated, the standard measure overestimate the accessibility contribution of new infrastructure, with overestimation errors ranging from 14 to 30 % in two quasi-real scenarios tested in our study. The CCA measure can effectively incorporate the impact of capacity constraints: it is responsive to different arrival rules, service headway distributions, and system loads, and provides a more realistic representation of the accessibility changes than the standard measure. The use of CCA makes the measurement of accessibility more complex, but our findings demonstrate that avoiding these significant overestimation errors requires this more complex approach.

HSR infrastructure in China influences population distribution, investment patterns, and the functional integration of Chinese cities by dramatically reducing transit times across and between urban regions (Fang 2013). At station-level, a study in Europe concluded that the usage of railway depends on the quality of integration into local transport infrastructures (Brons et al. 2009). Adding to the conclusions from past studies, this paper highlights that capacity constraints also influence the level of access to rail service. By taking into account capacity constraints, the value of accessibility becomes dynamic instead of static because capacity constraints vary with many operational factors including travel demand, service frequency, etc. Planners and policy-makers should not neglect the effect of operational conditions on accessibility dynamics. In the worse-case scenario we test—unreliable service and 115 % passenger load during peak hour—the accessibility improvements of HSR are almost nullified by capacity constraints. With an overestimation of accessibility, post hoc analysis of a transit system's wider economic and spatial impacts is likely to be misjudged. Policy adaptations or timely interventions, such as increasing service

frequency or offering alternative travel modes can mitigate the accessibility effects of capacity constraints.

The CCA is a generic method that can be applied in the other transit systems in which capacity constraints are critical, even though the queuing details have to be modified to reflect the specific system characteristics. Many urban transit systems in large cities are close to capacity, including those of London and New York in the West and Seoul, Shanghai, Hong Kong and Singapore in Asia. These metro systems often operate at saturation level throughout the peak hours. To apply CCA in the study of these systems, it may be possible to simplify the calculation by creating parametric functions for the probability of persistence and expected travel time as functions of threshold of persistence, system loads, peak-hour periods, service headway, etc., based on a more general and comprehensive PPEWT table.

In addition to the route-level analysis presented in this paper, CCA can also be used for network-level analyses of accessibility. Our paper shows that the CCA offers a lower (and more realistic) estimate of the accessibility contribution by a new transportation service. If the creation of a new service reduces crowding in other service lines by diverting demand, CCA can be extended to detect the accessibility improvement in those lines. Therefore the CCA not only better represents the aggregate value of the accessibility improvement (as in Table 3), but can also capture the spatial distribution of its benefits.

CCA can be further extended to examine the interchange stations between inter-city rail system and the urban transit system, which are often the bottleneck for both systems. Local demand for rail usage and intermodal accessibility are both essential for urban rail planning (Horner and Grubesic 2001). CCA is able to provide an additional tool in the planning process which also takes into account intermodal demand for accessibility measurement. For instance, the High Speed 2 project in the U.K. is planned to connect Northern England to London's Euston Railway Station, where passengers currently transfer to Victoria Station or Northern lines, both of which are heavily congested during peak hours. It is critical to understand the capacity constraints in the London Underground in order to estimate the accessibility benefit of the High Speed 2 project.

Another application is to more accurately calculate the accessibility gain in intermediary cities along the high speed rail lines. High passenger demand at upstream stations can restrict accessibility gains at intermediary stations. A city may be seemingly connected to a high speed rail line but not experience substantial accessibility improvements if the actual seat capacity is constrained. The standard accessibility measure is not sensitive to such details. Since the actual service capacity can be a function of the seat allocation policy set by the operating companies or government authorities, it is important to use the CCA measure to evaluate or design such policies.

One limitation in this paper is that we assume the passenger capacity is the dominant constraint on accessibility. Although the studies reviewed in "Capacity constraints in transit networks" section demonstrate that train boarding is perhaps more important than other capacity issues in the context of BTIHSR network, we acknowledge that capacity constraints may also occur elsewhere in a given transit network. To implement CCA in transit networks where other capacity constraints prevail, a more system-wide analysis is necessary. Another limitation of the paper is the lack of treatment of the pre-trip seat reservation: some passengers reserve the tickets in advance and therefore they do not need to join the queue. In our future work, we will try to acquire the data of the percentage of pre-ordered tickets in order to enhance the accuracy of the model. We anticipate a complex dynamics of ticketing choice behavior. The percentage may not be constant. Rather, people are more likely to reserve the tickets ahead of time when the crowding increases, and the

government may also adjust its ticket allocation policy out of fairness concerns and/or to encourage the IC card adoption.

Appendix 1: calculation of OD travel time

In this study, travel time from origin i to destination j without capacity constraints is calculated by using the actual road and railway network data for the municipalities of Beijing and Tianjin as well as the intermediary city of Langfang in Hebei Province. The transportation network in the study area is obtained from OpenStreetMap.org and contains more than 150,000 links, including all types of roads from residential street to motorway, local metro systems in Beijing and Tianjin, and national railway network including BTIHSR. The main transportation infrastructures are shown in Fig. 3. The travel time before and after the opening of BTIHSR are calculated by the same procedures.

To illustrate the general calculation procedures, in the following discussions, we denote $TT(i,j)$ as the travel time from origin i in the Beijing South Station area to destination j in the Tianjin Station area. In other words, before the opening of BTIHSR, $TT(i,j)$ is given as t_{ij}^{ante}; after BTIHSR, $TT(i,j)$ is given as t_{ij}^m for standard accessibility measures, or given as $\tilde{t}_{ij}^m$ for CCA measures. The travel time $TT(i,j)$ is the minimum between public transit travel time $TT^R(i,j)$ and car travel time $TT^C(i,j)$:

$$TT_{ij} = \min\left[TT^R(i,j), TT^C(i,j)\right]. \tag{14}$$

Without taking waiting time into account, the in-vehicle public transit travel time is calculated as:

$$
\begin{aligned}
TT^R(i,j) = &\ TT^{walk}(i,N_i) + TT^{bus}(N_i,D_i) + TT^{subway}(D_i,S_{BJ}) + TT^{icr}(S_{BJ},S_{TJ}) \\
&+ TT^{subway}(S_{TJ},D_j) + TT^{bus}(D_j,N_j) + TT^{walk}(N_j,j) + \mathrm{E}[\omega]
\end{aligned}, \tag{15}
$$

which is the sum of (1) walking time from i to the closet road node N_i, $TT^{walk}(i,N_i)$; (2) bus travel time from N_i to the closest subway station D_i, $TT^{bus}(N_i,D_i)$; (3) subway travel time from D_i to Beijing or Beijing South railway station S_{BJ}, $TT^{subway}(D_i,S_{BJ})$; (4) intercity railway travel time from Beijing to Tianjin station, $TT^{icr}(S_{BJ},S_{TJ})$; (5) subway travel time in Tianjin to station D_j, the closest one to destination j, $TT^{subway}(S_{TJ},D_j)$; (6) bus travel time from D_j to N_j, the closest road node to destination j, $TT^{bus}(D_j,N_j)$; (7) walking time from N_j to j, $TT^{walk}(N_j,j)$. The walking speed is set as 4.4 km/h; and the bus speed in the city area of Beijing and Tianjin is assumed to be the same as the car speed, which will be discussed later. Without capacity constraint, the classical expression of waiting time $\mathrm{E}[\omega]$ (indifferent to OD pairs) is used in Eq. (15), based on Osuna and Newell (1972):

$$\mathrm{E}[\omega] = \frac{1}{2}\mathrm{E}[t_H] \cdot \left(1 + CV(t_H)^2\right), \text{ where } CV(t_H) = \sqrt{\frac{\mathrm{var}(t_H)}{\mathrm{E}[t_H]^2}}. \tag{16}$$

$\mathrm{E}[t_H]$ is the expected headway distribution t_H; and $CV(t_H)$ is the coefficient of variation of t_H; and $\mathrm{var}(t_H)$ is the variation of t_H. As the subway and bus services in Beijing and Tianjin are highly dense and frequent (1- or 2-min headway during peak hours), in this

study, the expected waiting time is calculated only as the waiting time for BTIHSR. Taking into account the capacity constraints, the expected waiting time $E[\omega]$ is replaced by $E\left[\omega_{ij} \middle| \omega_{ij} \leq \Omega_{ij}\right]$ for each OD pair. In a mega-city like Beijing, capacity constraints should not only occur in inter-city trips, but also in all nodes of the intra-city transit network. To measure the CCA of whole transit network, a more realistic—but much more complex—approach is to incorporate capacity constraints into all transit network nodes. In this study, we only illustrate the CCA of HSR, but we hope to model network-wide capacity constraints in future work.

The car travel time from i to j is calculated as:

$$
\begin{aligned}
TT^C(i,j) = {}& TT^{walk}(i, N_i) + TT^{car}_{loc}(N_i, G_i) + TT^{car}_{reg}(G_i, G_j) \\
& + TT^{car}_{loc}(G_j, N_j) + TT^{walk}(N_j, j)
\end{aligned}, \tag{17}
$$

which is the sum of (1) walking time from i to the closet road node N_i, $TT^{walk}(i, N_i)$; (2) intra-city car travel time in Beijing from N_i to the closest inter-city highway entry G_i, $TT^{car}_{loc}(N_i, G_i)$; (3) inter-city highway travel time from Beijing to Tianjin, $TT^{car}_{reg}(G_i, G_j)$; (4) intra-city car travel time in Tianjin from highway exit G_j to closest road node N_j to destination j, $TT^{car}_{loc}(G_j, N_j)$; and (5) walking time from N_j to final destination j, $TT^{walk}(N_j, j)$.

The assignment of road speed is a critical task to calculate roadway travel time. In the cities of Beijing and Tianjin, the roadway traffic is highly congested, especially during peak hours. The actual road speed in these cities is much lower than the designed speed. As the actual travel speed of in each road segment of both cities is unavailable, we use the following estimation to get the proximate travel speed in each type of road. In 2013, the average vehicle speed $\bar{V}$ in Beijing during peak hours was 20 to 25 km/h (World Resources Institute 2015). Here, we let $\bar{V}$ equal to 25 km/h in both Beijing and Tianjin. The estimation of road speed V_k in each road type k is solved by the following formula:

$$
\sum_k V^{city}_k \cdot L^{city}_k = \bar{V} \cdot \sum_k L^{city}_k, \tag{18}
$$

where L^{city}_k is the total length of each road type k in the city of Beijing or Tianjin; and V^{city}_k is the estimated travel speed of road type k in Beijing or Tianjin. As the design speed of each type of road is known, by keeping the ratio between each road's design speed constant, this formula is thus solvable. The estimated road speeds of each type of road in

Table 4 Estimated road speed in the City of Beijing and Tianjin

	Design speed	Speed in Beijing	Speed in Tianjin
Expressways	100	40.60	46.39
Trunk roads	80	32.48	37.11
Primary roads	60	24.36	27.83
Secondary roads	40	16.24	18.56
Tertiary roads	30	12.18	13.92

The unit of speed is km/h

the cities of Beijing and Tianjin are listed in Table 4, and these are used to calculate the

intra-city roadway travel time. The city areas of Beijing and Tianjin are shown in Fig. 3.

Table 5 BTIHSR morning timetable

Departure	Headway	Departure	Headway	Departure	Headway	Departure	Headway
6:13	0:08	7:50	0:10	9:29	0:11	11:10	0:08
6:21	0:14	8:00	0:08	9:40	0:10	11:18	0:12
6:35	0:15	8:08	0:08	9:50	0:10	11:30	0:10
6:50	0:10	8:16	0:15	10:00	0:10	11:40	0:25
7:00	0:13	8:31	0:11	10:10	0:05	12:05	0:11
7:13	0:10	8:42	0:19	10:15	0:14	12:16	0:14
7:23	0:08	9:01	0:05	10:29	0:11	12:30	0:08
7:31	0:10	9:06	0:13	10:40	0:15	12:38	
7:41	0:09	9:19	0:10	10:55	0:15		

Times given in hour:minute format

Appendix 2: BTIHSR morning timetable (Version Feb. 2014)

See Table 5.

References

Abkowitz, M., Tozzi, J.: Research contributions to managing transit service reliability. J. Adv. Transp. **21**, 47–65 (1987)

Baptiste, H., et al. (2003) Transport services and networks: territorial trends and basic supply of infrastructure for territorial cohesion. ESPON Project 1.2.1 Second interim report, Luxembourg

Bates, J.: Reliability—the missing model variable. In: Hensher, D.A. (ed.) Travel Behaviour Research: the Leading Edge, pp. 527–546. Pergamon, Oxford (2001)

Bhat, C., Handy, S.L., Kockelman, K., Mahmassani, H., Chen, Q., Weston, L.: Urban Accessibility Index: literature Review. Center of Transportation Research, University of Texas at Austin, Springfield (2000)

Brons, M., Givoni, M., Rietveld, P.: Access to railway stations and its potential in increasing rail use. Transp. Res. Part A **43**, 136–149 (2009)

Bullock, R.H., Salzberg, A., Jin, Y.: High-speed rail—the first three years: taking the pulse of China's Emerging Program. World Bank (2012)

Cats, O.: Multi-agent transit operations and assignment model. Proced. Comput. Sci. **19**, 809–814 (2013)

Cats, O., Jenelius, E.: Planning for the unexpected: the value of reserve capacity for public transport network robustness. Transp. Res. Part A (2015). doi:10.1016/j.tra.2015.02.013

Chaudhry, M.L.: Numerical issues in computing steady-state queueing-time distributions of single-server bulk-service queues: M/Gb/1 and M/Gd/1. INFORMS J. Comput. **4**, 300–310 (1991)

Chaudhry, M.L., Templeton, J.G.C.: A First Course in Bulk Queues. Wiley, New York (1983)

ChinaNews.com.: Ministry of rail: High-speed rail leads to new development of green transportation (2009). http://www.chinanews.com/cj/cj-cyzh/news/2009/08-01/1800203.shtml. Accessed 6–7 2015

Cominetti, R., Correa, J.: Common-lines and passenger assignment in congested transit networks. Transp. Sci. **35**, 250–267 (2001)

Curtis, C., Scheurer, J.: Planning for sustainable accessibility: developing tools to aid discussion and decision-making. Prog. Plan. **74**, 53–106 (2010)

Dümmler, M., Vicari, N.: A numerical analysis of the M/Db/N queueing system. Institute of computer science research report series, Univeristy of Würzburg (1999)

Fan, W., Machemehl, R.B.: Characterizing bus transit passenger waiting times. Paper presented at the 2nd Material Specialty Conference of the Canadian Society for Civil Engineering Montréal, Canada (2002)

Fang, W.: Dispersion of agglomeration through transport infrastructure. Doctoral dissertation, Massachusetts Institute of Technology (2013)

Frumin, M., Zhao, J.: Analyzing passenger incidence behavior in heterogeneous transit services using smartcard data and schedule-based assignment. Transp. Res. Rec. **2274**, 52–60 (2012)

Fu, Q., Liu, R., Hess, S.: A review on transit assignment modelling approaches to congested networks: a new perspective. Proced.—Soc. Behav. Sci. **54**, 1145–1155 (2012)

Geurs, K.T., Ritsema Van Ecka, J.R.: Accessibility measures: Review and applications. Rijksinstituut voor Volksgezondheid en Milieu (National Institute of Public Health and the Environment), Utrecht University, Bilthoven/Utrecht, the Netherlands (2001)

Geurs, K.T., van Wee, B.: Accessibility evaluation of land-use and transport strategies: review and research directions. J. Transp. Geogr. **12**, 127–140 (2004)

Glazer, A., Hassin, R.: Equilibrium arrivals in queues with bulk service at scheduled times. Transp. Sci. **21**, 273–278 (1987)

Gold, H., Tran-Gia, P.: Performance analysis of a batch service queue arising out of manufacturing system modelling. Queueing Syst. **14**, 413–426 (1993)

Gross, D., Shortle, J.F., Thompson, J.M., Harris, C.M.: Fundamentals of Queuing Theory, 4th edn. Wiley, New York (2008)

Gutiérrez, J.: Location, economic potential and daily accessibility: an analysis of accessbility impact of high-speed line Madrid-Barcelona-French border. J. Transp. Geogr. **9**, 229–242 (2001)

Gutiérrez, J., González, R., Gómez, G.: The European high-speed train network: predicted effects on accessibility patterns. J. Transp. Geogr. **4**, 227–238 (1996)

Gutiérrez, J., Urbano, P.: Accessibility in the European Union: the impact of the trans-European road network. J. Transp. Geogr. **4**, 15–25 (1996)

Hamdouch, Y., Lawphongpanich, S.: Schedule-based transit assignment model with travel strategies and capacity constraints. Transp. Res. Part B **42**, 663–684 (2008)

Hamdouch, Y., Marcotte, P., Nguyen, S.: Capacitated transit assignment with loading priorities. Math. Program. **101**, 205–230 (2004a)

Hamdouch, Y., Marcotte, P., Nguyen, S.: A strategic model for dynamic traffic assignment. Netw. Spat. Econ. **4**, 291–315 (2004b)

Hansen, W.G.: How accessibility shapes land use. J. Am. Inst. Plan. **25**, 73–76 (1959)

Horner, M.W., Grubesic, T.H.: A GIS-based planning approach to locating urban rail terminals. Transportation **28**, 55–77 (2001)

ITourBeijing.com.: Beijing Travel Guide: Beijing–Tianjin Intercity Railway (2015). http://www.itourbeijing.com/beijing-tour/beijing-tianjin-intercity-railway.htm. Accessed Jun 7th 2015

Kitamura, R.: The effects of added transportation capacity on travel: a review of theoretical and empirical results. Transportation **36**, 745–762 (2009)

Kurauchi, F., Bell, M.G.H., Schmöcker, J.-D.: Capacity constrained transit assignment with common lines. J. Math. Model. Algorithms **2**, 309–327 (2003)

Leurent, F.: Transport capacity constraints on the mass transit system: a systemic analysis. Eur. Transp. Res. Rev. **3**, 11–21 (2011)

Leurent, F., Chandakas, E.: The transit bottleneck model. Proced. Soc. Behav. Sci. **54**, 822–833 (2012)

Leurent, F., Chandakas, E., Poulhès, A.: A traffic assignment model for passenger transit on a capacitated network: bi-layer framework, line sub-models and large-scale application. Transp. Res. Part C **47**, 3–27 (2014)

Lucas, K., van Wee, B., Maat, K.: A method to evaluate equitable accessibility: Combining ethical theories and accessibility-based approaches. Transportation In Press (2015) doi:10.1007/s11116-015-9585-2

Osuna, E.E., Newell, G.F.: Control strategies for an idealized public transportation system. Transp. Sci. **6**, 52–72 (1972)

Poon, M.H., Wong, S.C., Tong, C.O.: A dynamic schedule-based model for congested transit networks. Transp. Res. Part B **38**, 343–368 (2004)

Schmöcker, J.-D., Fonzone, A., Shimamoto, H., Kurauchi, F., Bell, M.G.H.: Frequency-based transit assignment considering seat capacities. Transp. Res. Part B **45**, 392–408 (2011)

Schneider, A., Friedl, M.A., Potere, D.: A new map of global urban extent from MODIS satellite data. Environ. Res. Lett. **4**, 1–11 (2009)

Schneider, A., Friedl, M.A., Potere, D.: Mapping global urban areas using MODIS 500-m data: new methods and datasets based on 'urban ecoregions'. Remote Sens. Environ. **114**, 1733–1746 (2010)

Shen, Y., de Abreu e Silva, J., Martínez, L.M.: Assessing High-speed Rail's impacts on land cover change in large urban areas based on spatial mixed logit methods: a case study of Madrid Atocha railway station from 1990 to 2006. J. Transp. Geogr. **41**, 184–196 (2014)

Shimamoto, H., Kurauchi, F., Iida, Y., Bell, M.G.H., Schmöcker, J.-D.: Evaluating public transit congestion mitigation measures using a passenger assignment model. J. East. Asia Soc. Transp. Stud. **6**, 2076–2091 (2005)

Spiekermann, K., Wegener, M.: Accessibility and spatial development in Europe. Sci. Reg. **5**, 15–46 (2006)

Tan, X.: Commuters between Beijing and Tianjin, Trains fully occupied every day: Three years of Beijing–Tianjin Intercity HSR. Bohai Morning Post (2014). http://roll.sohu.com/20140321/n396990204.shtml. Accessed June 7th 2015

van Wee, B., Hagoort, M., Annema, J.A.: Accessibility measures with competition. J. Transp. Geogr. **9**, 199–208 (2001)

Vickerman, R., Ulied, A.: Indirect and wider economic impacts of high speed rail. In: de Rus, G. (ed.) Economic Analysis of High Speed Rail in Europe, pp. 89–118. Fundación BBVA, Madrid (2009)

Wang, J.J., Xu, J., He, J.: Spatial impacts of high-speed railways in China: a total-travel-time approach. Environ. Plan. A **45**, 2261–2280 (2013)

World Resources Institute.: Beijing says "no" to vehicle ownership growth (2015). http://www.wri.org.cn/en/news/beijing-says-%E2%80%9Cno%E2%80%9D-vehicle-ownership-growth. Accessed June 7th 2015

Yin, M., Bertolini, L., Duan, J.: The effects of the high-speed railway on urban development: International experience and potential implications for China. Prog. Plan., 1–52 (2014). doi:10.1016/j.progress.2013.11.001

Yu Shen is a Doctoral Candidate in the DECIVIL/CERIS at Instituto Superior Técnico, Universidade de Lisboa, and a Visiting Researcher in the Department of Urban Studies and Planning at Massachusetts Institute of Technology. His research interests cover Transportation Land-Use Interaction Modelling, Discrete Choice Methods, Agent-based Modelling, GIS, Transport Policy and Behavior, and ITS.

Jinhua Zhao is the Edward H. and Joyce Linde Assistant Professor in the Department of Urban Studies and Planning at Massachusetts Institute of Technology. He studies (1) the behavioral foundation for transport policies; (2) public transit management using information technology; and (3) China's urbanization and urban mobility. Prof. Zhao has been working with London, Chicago, Boston, Vancouver, Shanghai, Beijing and Singapore to help improve their urban transportation systems and policies. Prof. Zhao directs the urban mobility research lab JTL and is the co-PI for the Transit Lab at MIT.