

Reliable frequency determination

Incorporating information on service uncertainty when setting dispatching headways

Gkiotsalitis, K.; Cats, O.

DOI

[10.1016/j.trc.2018.01.026](https://doi.org/10.1016/j.trc.2018.01.026)

Publication date

2018

Document Version

Final published version

Published in

Transportation Research Part C: Emerging Technologies

Citation (APA)

Gkiotsalitis, K., & Cats, O. (2018). Reliable frequency determination: Incorporating information on service uncertainty when setting dispatching headways. *Transportation Research Part C: Emerging Technologies*, 88, 187-207. <https://doi.org/10.1016/j.trc.2018.01.026>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

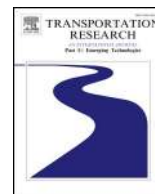
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' – Taverne project

<https://www.openaccess.nl/en/you-share-we-take-care>

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.



Reliable frequency determination: Incorporating information on service uncertainty when setting dispatching headways

K. Gkiotsalitis^{a,*}, O. Cats^b

^a NEC Laboratories Europe, Kurfürsten-Anlage 36, 69115 Heidelberg, Germany

^b Delft University of Technology, Postbus 5, 2600 AA Delft, The Netherlands

ARTICLE INFO

Keywords:

Reliability-based frequency setting
Tactical planning
Headway variability
Dispatching headway determination
Resource constrained optimization
Non-linear programming

ABSTRACT

Frequency setting requires the determination of the dispatching headways of all bus lines in a city network and constitutes the main activity in the tactical planning of public transport operations. Determining the dispatching headways of bus services in a city network is a multi-criteria problem that typically involves balancing between passenger demand coverage and operational costs. In this study, the problem of setting the optimal dispatching headways is formulated with the explicit consideration of operational variability issues for mitigating the adverse effects of passenger demand and travel time variations inherent to bus operations. The proposed model for setting the dispatching headways of bus lines considers the demand, headway and travel time variations along every section of each bus route for different times of the day, as well as operational costs, vehicle capacity and fleet size constraints.

We first formulate the problem while accounting for the consequences of variability in service operations. The resulting optimization problem is then solved by employing a Branch and Bound approach together with Sequential Quadratic Programming in order to find the optimal dispatching headway for each bus line. Experimental results demonstrate (a) the improvement potential of the base case dispatching headways when considering the service reliability; (b) the sensitivity of the determined dispatching headways to changes in different criteria, such as passenger demand and/or bus running costs, and (c) the convergence accuracy of the proposed solution method when compared to heuristic approaches.

1. Introduction

Public transport operators need to continuously update service frequencies to cater for changes in traffic conditions and passenger demand in both space and time. The service frequencies can be updated by modifying the dispatching headways of the respective bus services since the frequency of one bus line is inversely proportional to its dispatching headway. Bus line frequencies can be adjusted to the passenger travel needs subject to resource capacities and operational cost limitations by using information from passengers (i.e., smartcard logs (Pelletier et al., 2011; Ma et al., 2013; Munizaga and Palma, 2012; Luo et al., 2017), smartphones (Alexander et al., 2015; Gkiotsalitis and Stathopoulos, 2015; Calabrese et al., 2013; de Regt et al., 2017) and operating vehicles (Cortés et al., 2011).

In transit planning, frequency setting follows the network design and precedes the timetable design and vehicle and crew scheduling (refer to Kepaptsoglou and Karlaftis, 2009; Farahani et al., 2013; Ceder, 2007 for more details on public transport planning processes). Setting the frequencies by determining the dispatching headways of bus lines and network design are commonly

* Corresponding author.

E-mail addresses: konstantinos.gkiotsalitis@neclab.eu (K. Gkiotsalitis), o.cats@tudelft.nl (O. Cats).

considered as two consequent problems (with the exceptions of Silman et al., 1974; Ramos, 2014; Szeto and Wu, 2011). Frequency setting and network design are considered as two consequent problems due to the complexity of these problems, the inefficiency of solving approaches and the fact that frequency setting can be adjusted as part of tactical planning.

Methods to determine dispatching headways for setting the frequencies of the services are based on either passenger load profile rule-based techniques (Ceder, 2007, 1984; Hadas and Shnaiderman, 2012) or on minimizing passenger and operator costs Furth and Wilson (1981), Cipriani et al. (2012) and Gkiotsalitis and Cats (2017). For more details one can refer to the literature reviews by Ibarra-Rojas et al. (2015) and Guilhaire and Hao (2008). The unsatisfied demand is a main factor of the above-mentioned problem and is generally modeled by introducing a penalization weight in the objective function (Barra et al., 2007; Cipriani et al., 2012; Fan and Machemehl, 2008). Common practice in public transit planning is to determine the dispatching headways based on accumulated hourly passenger counts, average travel time, vehicle capacity and the minimum allowed frequency limit by time of day. One exception is the work of Hadas and Shnaiderman (2012) which presented a new approach for setting the dispatching headways by introducing the stochastic properties of Automatic Vehicle Location (AVL) and Automatic Passenger Counting (APC) data within a supply chain optimization model. The optimization elements of that approach were the: (a) empty-seat driven (unproductive cost) and (b) the overload and un-served demand (increased user cost).

Several studies have considered stochasticity in the tactical planning phase (Amberg et al., 2017). Li et al. (2013) considered stochastic parameters such as demand, arrival times, boarding/alighting times, and travel times via a stochastic optimization approach and a meta-heuristic solver that minimizes the sum of the expected value of the company costs and the waiting time costs for passengers. Bellei and Gkoumas (2010) modeled also the demand and dwell times stochastically while Barabino et al. (2017) proposed an offline framework that identifies the bus stops and the time periods in which the reliability of the bus operations is not sufficient using historical vehicle location data. An interesting extension of the models that determine the dispatching headways of bus lines which tries to minimize the passenger waiting times and operational costs while increasing ridership came from Gkiotsalitis et al. (2017), Verbas and Mahmassani (2013) and Verbas et al. (2015). Verbas et al. (2015) extended the model presented by Furth and Wilson (1981) considering demand variations along the route; thus, enabling the split of the route into sub-routes that enjoy homogeneous demand patterns in order to define dispatching headways for each sub-route independently. The variation of demand was modeled by assuming temporal and spatial heterogeneity of the ridership elasticity with respect to dispatching headways and the problem was formulated with a non-linear program which minimizes the weighted sum of ridership and wait time savings over all stops, lines, and time intervals subject to constraints such as budget, fleet size, headway bounds for each line pattern, and bounds for load factors.

Notwithstanding the above, to the best of the authors knowledge, none of the previous studies solved the problem of setting dispatching headways while considering the reliability of service operations and the consequences of travel time and demand variability during the day; even if the implications of the bus service reliability problem have been analyzed by several works such as the work of Chen et al. (2009). Neglecting service variability at the planning phase leads to the selection of sub-optimal solutions and the underestimation of both operational and passenger costs. Service reliability is mostly addressed at the operations control phase by re-adjusting planned schedules or applying other control measures such as bus holding or speed control in real-time for reacting to trip travel time and passenger demand changes (Gkiotsalitis and Maslekar, 2015; Moreira-Matias et al., 2016; Asgharzadeh and Shafahi, 2017). However, the consideration of service reliability already at the tactical planning phase can potentially generate solutions that tackle the inherent uncertainty of public transport operations which is particularly high at dense metropolitan areas with high-demand bus operations.

In the remainder of this paper, we develop and apply a reliability-based optimization framework for setting the dispatching headways of bus lines that considers historical operational data and is aware of the passenger waiting time variability at each stop and how it is affected when changing the planned dispatching headways. In the following section, the problem description is presented considering the demand variations and the travel time variability from bus stop to bus stop over time. In addition, the multi-objective problem of setting the optimal dispatching headways of several bus lines within a study area is formulated. An exact solution method for solving the resulting discrete non-linear programming problem is described. The method is applied by using General Transit Feed Specification (GTFS) data from 17 central bus lines in Stockholm and detailed AVL and APC data from central bus lines 1 and 3. After discussing the experimentation results, concluding remarks about practical implications and future work directions are presented in the closing section.

2. Reliability-based frequency setting problem

2.1. Problem formulation

In this work, we introduce the following notation for describing the main components of the frequency setting problem that requires the determination of the dispatching headways of all bus lines in a study area.

$\{L, S\}$	is a network with L bus lines and S bus stops.
t_l^{90th} (hour)	the total travel time value of a line l for which there is only a 10% chance for a bus trip to require more travel time than that. This travel time includes the boarding/alighting times at each bus stop and the layover times before starting a new trip
$L = \{L_1, L_2, \dots, L_{ L }\}$	the bus lines of the network

$S_l = \{S_{l,1}, S_{l,2}, \dots, S_{l, S_l }\}$	a vector denoting the stops of line $l \in L$ in a sequential order starting from the departure stop
q_l (Nr. of buses)	number of buses allocated to line l
γ (Nr. of buses)	maximum number of available buses at the network level
$b_{l,j,t}$ (passenger boardings/ hour/stop)	total number of observed boardings for line l at stop j for a time period t . A time period t of 1 h is commonly used in literature (Hadas et al. (2010)) for avoiding demand heterogeneity issues
k_l (monetary cost/trip)	operational cost per run for route l
h_l (minutes/trip departure)	the planned dispatching headway of bus line $l \in L$ over a period of the day in minutes (decision variable of the frequency setting problem)
$h_{l,j}(h_l)$ (minutes)	the expected operational headways at stops $j \in S_l$ when the planned dispatching headway is h_l
$Varh_{l,j}(h_l)$ (minutes ²)	the expected variance of the operational headways at stops $j \in S_l$ when the planned dispatching headway is h_l
$CVh_{l,j}$ (unitless)	the expected coefficient of variation of the operational headways at stops $j \in S_l$ when the planned dispatching headway is h_l
θ (Nr. of passengers)	the passenger capacity of the vehicle
$b_{l,j,t}(h_l)$ (passenger boardings/ hour/stop)	the expected boarding levels as a function of the planned dispatching headway h_l
$\Psi_{l,j,t}$ (passenger load/hour/ stop)	the observed hourly passenger load for line l at stops $j \in S_l$ for a time period t . It is computed as the difference of accumulated boarding passengers and alighting passengers
$\Psi_{l,j,t}(h_l)$ (passenger load/ hour/stop)	the expected passenger loading levels as a function of the planned dispatching headway h_l
$T = \{t_1, t_2, \dots, t_{ T }\}$	the hourly time periods for which the dispatching headway h_l is applied

The planned dispatching headway, h_l , of one bus line $l \in L$ can be selected from a pre-determined admissible set of values $h_l \in \{2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15\}$ min. in order to adhere to the cyclic bus timetable requirement. The upper bound is set to 15 min because we focus on high frequency services where the frequencies are sufficiently high so that passengers do not coordinate their arrival with vehicle arrivals (in this way, we allow at least $\frac{15 \text{ min headway per departure}}{60 \text{ minutes per hour}} = 4$ departures per hour). By doing so, we can assume random passenger arrivals at stops like most works in literature (see Furth and Wilson, 1981).

For setting the dispatching headways of bus lines, we form an objective function that considers four key components. First, we consider the passenger waiting cost at every stop $j \in S$. For this, we assume random passenger arrivals at stops resulting in a waiting time that is half the time of the headway at that stop. Ideally, the planned dispatching headway of a line l , which is the headway between successive trips at the departure stop, h_l , will be maintained at any other stop $j \in S_l - \{S_{l,1}\}$ of line l . However, this is not the case in real-world operations because of travel time variations that lead to bus bunching. For this reason, for each potential dispatching headway h_l we estimate the passenger waiting time at any stop j by using historical data of the observed headways at that stop from the real-world operations. Let assume that for different dispatching headways $\{h_l, h'_l, h''_l, \dots\}$ we observed headways $\{h_{l,j}, h'_{l,j}, h''_{l,j}, \dots\}$ at any bus stop $j \in S_l - \{S_{l,1}\}$. Then, for each bus stop $j \in S_l - \{S_{l,1}\}$ we can estimate its operational headway by using the dispatching headway at the departure stop according to the following expression:

$$h_{l,j}(h_l): \begin{cases} \text{expected operational headway at stop } j \in S_l - \{S_{l,1}\} \text{ for } h_l \\ h_l \text{ if } j = S_{l,1} \end{cases} \quad (1)$$

These headway observations from past operations refer to a specific value of the dispatching headway h_l and for each value of the dispatching headway h_l we use a different set of headway observations for calculating the average headway at that stop.

Using the expected value of the operational headways at any stop j for a planned dispatching headway h_l , the waiting times of passengers for that dispatching headway are:

$$\sum_{l \in L} \sum_{j \in S_l} \frac{h_{l,j}(h_l)}{2} \sum_{t=1}^{|T|} b_{l,j,t}(h_l)$$

, where the term $b_{l,j,t}(h_l)$ represents the hourly passenger boardings for line l and stop j as a function of the dispatching headway h_l .

In the above expression, we assumed random passenger arrivals where the waiting time at each stop is equal to half of the average headway at that stop. Many works in literature though consider also the headway variability for the estimation of the waiting time. Osuna and Newell (1972) for instance defined the waiting time for passenger random arrivals as half of the average headway plus the ration of the headway variance to twice the average headway. Adopting this approach, the expression of the waiting times of passengers can be expanded as follows:

$$O_1(h_l) = \sum_{l \in L} \sum_{j \in S_l} \left(\frac{h_{l,j}(h_l)}{2} + \frac{Varh_{l,j}(h_l)}{2h_{l,j}(h_l)} \right) \sum_{t=1}^{|T|} b_{l,j,t}(h_l) \quad (2)$$

where:

$$Varh_{l,j}(h_l): \begin{cases} \text{expected headway variance at stop } j \in S_l - \{S_{l,1}\} \text{ for } h_l \\ 0 \text{ if } j = S_{l,1} \end{cases} \quad (3)$$

The expected headway and headway variance for every bus stop j can be derived using the observed headways at each stop for different values of dispatching headways.

In Eq. (2) we did not assume that the demand is inelastic with respect to passenger waiting times as many works in the literature do (Ibarra-Rojas et al., 2015). Instead, we used the function $b_{l,j,t}(h_l)$ to relate the number of boardings with the dispatching headway. Several studies, such as Paulley et al. (2006) and Preston and James (2000) have shown that there is an elasticity for bus demand with respect to passenger waiting times and the average value appears to be -0.64 for long-term operations. Let assume that a bus line l is operated for a long period with a dispatching headway $r_o = h'_l$ and for that dispatching headway the observed passenger demand was $b_{l,j,t}$ boardings and $\Psi_{l,j,t}$ loadings per line/stop and hour. Then, if for 1% waiting time increase we have a 0.64% decrease in passenger demand, the expected number of boardings for another dispatching headway, $h_l \neq h'_l$, can be approximated by the following iterative function:

$$b_{l,j,t}(h_l) = \begin{cases} \text{if } h_l > h'_l: & \begin{cases} r_o = h'_l, b_{l,j,t}(r_o) = b_{l,j,t} \\ r_n = r_{n-1} + 1\%r_{n-1} \\ b_{l,j,t}(r_n) = b_{l,j,t}(r_{n-1}) - 0.64b_{l,j,t}(r_{n-1}) \frac{r_n - r_{n-1}}{r_{n-1}} \end{cases} \\ \text{if } h_l < h'_l: & \begin{cases} r_o = h'_l, b_{l,j,t}(r_o) = b_{l,j,t} \\ r_n = r_{n-1} - 1\%r_{n-1} \\ b_{l,j,t}(r_n) = b_{l,j,t}(r_{n-1}) - 0.64b_{l,j,t}(r_{n-1}) \frac{r_n - r_{n-1}}{r_{n-1}} \end{cases} \end{cases} \quad (4)$$

where the initial value of $r_o = h'_l$ is updated in a recursive manner in order to investigate how much the expected passenger boardings per line/stop and hour are affected by the dispatching headway change according to the bus demand elasticity.

From Eq. (4) one can observe that if $h_l > h'_l$, then for every 1% increase of the dispatching headway ($r_n = r_{n-1} + 1\%r_{n-1}$) the number of boardings per line/stop and hour, $b_{l,j,t}(r_n)$, decrease by $-0.64b_{l,j,t}(r_{n-1}) \frac{r_n - r_{n-1}}{r_{n-1}}$ since $r_n - r_{n-1} = 1\%r_{n-1}$ which is always positive in this case. In contrast, if $h_l < h'_l$ for every 1% reduction of the dispatching headway ($r_n = r_{n-1} - 1\%r_{n-1}$) the boardings per line/stop and hour, $b_{l,j,t}(r_n)$, increase by $-0.64b_{l,j,t}(r_{n-1}) \frac{r_n - r_{n-1}}{r_{n-1}}$ since $r_n - r_{n-1} = -1\%r_{n-1}$ which is always negative in this case.

Eq. (4) shows that by reducing the planned dispatching headway the expected number of boarding passengers per line/stop for a given time window (i.e. hour), $b_{l,j,t}$, increases because passenger demand is sensitive to service improvements. However, this increase in the hourly passenger boarding rate may not necessarily increase the number of boarding passengers per vehicle trip because the increased demand is distributed over a larger number of trips.

After observing the form of the iterative function of Eq. (4), the expected number of boarding passengers for a planned dispatching headway can be further approximated using a closed-form, non-recursive expression. By plotting the output values of Eq. (4) for different dispatching headways (Fig. 1), one can observe that these values can be reproduced by a non-recursive function that has an exponential form. For this reason, we introduce a general exponential function for approximating Eq. (4) that has the following form:

$$b_{l,j,t}(h_l) = \psi_1 e^{-\psi_2 h_l} + \psi_3 \quad (5)$$

where ψ_1, ψ_2, ψ_3 are the parameters of the function which depend on the value of the dispatching headway under which the number of passenger boardings was observed. These parameters are computed by fitting the function to the data using non-linear least squares. To provide an example, a non-linear least squares fitting results to the parameter values of Fig. 1 for two hypothetical examples with (a) dispatching headway $h'_l = 10$ min. with 25 observed hourly passenger boardings and (b) dispatching headway of $h'_l = 5$ min. with

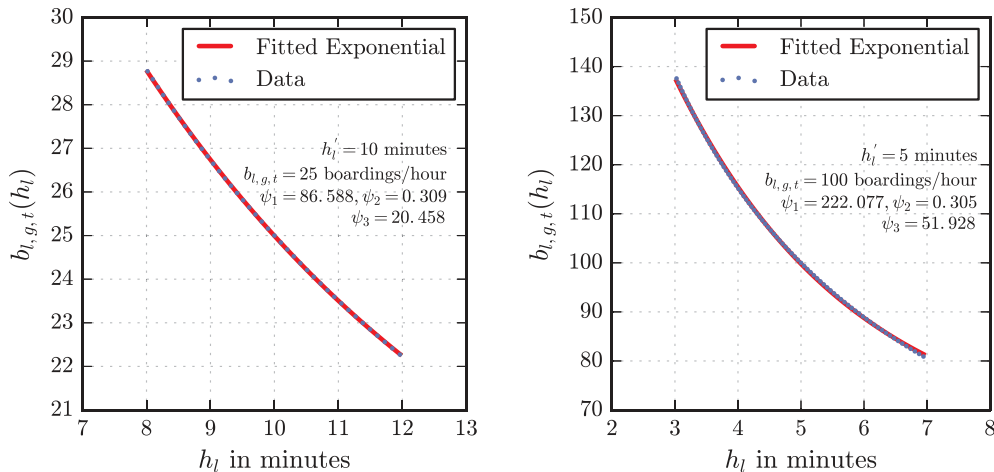


Fig. 1. Approximated Hourly Passenger Boardings, $b_{l,j,t}(h_l)$, for different values of dispatching headways [$h'_l = 10$ min and $b_{l,g,t} = 25$ boardings/h for the left figure; $h'_l = 5$ min and $b_{l,g,t} = 100$ boardings/h for the right figure], after approximating the iterative function of boardings with the exponential one of Eq. (5).

100 observed hourly passenger boardings.

The same logic of Eq. (4) applies for approximating the number of loadings for a dispatching headway change, $\Psi_{l,j,t}(h_l)$, after using the same incremental elasticity steps. Similarly, an exponential function can be used by following the logic of Eq. (5) for approximating the iterative function of $\Psi_{l,j,t}(h_l)$ with the use of a simpler exponential function.

Another objective is the improvement of service reliability. Service reliability is decoupled from the expected passenger waiting time because the cost of an unexpected waiting time is experienced as delay and therefore has a more negative impact to passengers than the anticipated waiting time as pointed out by Wardman (2004). In addition, transport authorities have increased the pressure to bus operators to improve operational service reliability. For this, they have defined specific KPIs such as the Excess Waiting Time of passengers (EWT) at stops for monitoring the service reliability and have developed incentive schemes for rewarding the performance of operators. For instance, the Land Transport Authority (LTA) in Singapore LTA (2017) offers a 6000 Singaporean dollar bonus per line for a 0.1 min reduction of the excess waiting time of passengers at stops.

For measuring the service reliability, Chen et al. (2009) tested a set of metrics that can be applied to the stop or the line level and proposed the use of the coefficient of variation, $CVh_{l,j}$, for a stop-level reliability assessment:

$$CVh_{l,j} = \frac{\sqrt{\text{Var}h_{l,j}(h_l)}}{h_{l,j}(h_l)} \quad (6)$$

Using this approach, the service reliability objective that considers also the level of boardings for giving different weights to different bus stops can be expressed as:

$$O_2(h_l) = \sum_{l \in L} \sum_{j \in S_l} CVh_{l,j} \sum_{t=t_1}^{t_{l,T}} b_{l,j,t}(h_l) \quad (7)$$

Finally, the objective function of setting the dispatching headways of bus lines should include the operational running costs which can be expressed as:

$$O_3(h_l) = \sum_{t=t_1}^{t_{l,T}} \sum_{l \in L} \frac{k_l}{h_l/60} \quad (8)$$

This cost component has been used by Furth and Wilson (1981) for calculating the running costs and is equal to the running costs of one bus allocated to a line l multiplied by the number of buses per hour. It represents the variable costs such as driver and technical staff, energy consumption and maintenance costs.

If we allocate a dispatching headway from the set $h_l = \{2,3,4,5,6,7,8,9,10,12,15\}$ min to a bus line, then we should ensure that the number of buses allocated to that line is an integer number. In the ideal case, the 90th percentile of the total travel time of a line, t_l^{90th} , should be such that for a planned dispatching headway h_l an integer number of vehicles is required. However, this is rarely the case and we might need to round upwards the number of buses according to the following equation:

$$q_l = \left\lceil \frac{t_l^{90th}}{h_l} \right\rceil \quad (9)$$

If there is a depreciation cost of operating an extra bus, the operator would be willing to deploy a solution that requires fewer buses. This can be translated to the following objective that penalizes the use of additional buses:

$$O_4(h_l) = \sum_{l \in L} \left\lceil \frac{t_l^{90th}}{h_l} \right\rceil \quad (10)$$

Apart from the objectives, there are also resource limitations. One first limitation is the availability of buses. The total number of buses assigned to every line should be at most equal to the total number of buses available at the network level:

$$\sum_{l \in L} \left\lceil \frac{t_l^{90th}}{h_l} \right\rceil \leq \gamma \quad (11)$$

where γ is the total number of available buses and is a positive integer.

Another limitation is the vehicle capacity. The maximum hourly loading point (MLP) method (Lee and Vuchic, 2005; Ceder and Israeli, 1998) has been extensively used as one of the two main methods for defining the optimal frequency per line that accomplishes the desired level of occupancy at the most crowded stop during the peak demand hour. A modification of this method can be used for defining the vehicle capacity constraint where the selected frequency should be such that the vehicle load at the peak hour and the peak stop is still lower than the vehicle capacity. If $\Psi_{l,j,t}(h_l)$ is the observed load for a planned dispatching headway h_l for each hourly time period for line l and stop j , then the maximum loading that occurs at stop j^* and hour t^* is expressed as:

$$\max \Psi_l(h_l) = \Psi_{l,j^*,t^*}(h_l) \quad (12)$$

Then, the dispatching headway of line l should satisfy the inequality:

$$h_l \leq \frac{60\theta}{\max \Psi_l(h_l)} \quad (13)$$

The importance of each one of these four objectives on the objective function of the problem depends on the operator's management preferences and the operational context. Weighting factors can be determined based on passenger and operator cost estimates (i.e. value of time, fixed and variable cost units). In the following, we form a single-objective function by introducing weight factors $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ that translate all costs to monetary terms, establishing thus trade-offs between compensatory objective function components:

$$\begin{aligned}
 \min_{\mathbf{h}=\{h_{L_1}, \dots, h_{L_{|L|}}\}} z(\mathbf{h}) &= \alpha_1 \sum_{l \in L} \sum_{j \in S_l} \left(\frac{h_{lj}(h_l)}{2} + \frac{\text{Var}h_{lj}(h_l)}{2h_{lj}(h_l)} \right) \sum_{t=t_1}^{t_{|T|}} b_{lj,t}(h_l) + \alpha_2 \sum_{l \in L} \sum_{j \in S_l} CVh_{lj}(h_l) \sum_{t=t_1}^{t_{|T|}} b_{lj,t}(h_l) + \alpha_3 \sum_{t=t_1}^{t_{|T|}} \sum_{l \in L} \frac{k_l}{h_l / 60} \\
 &+ \alpha_4 \sum_{l \in L} \left\lceil \frac{t_l^{90th}}{h_l} \right\rceil \\
 \text{subject to:} \quad \text{Available Buses:} \quad &\sum_{l \in L} \left\lceil \frac{t_l^{90th}}{h_l} \right\rceil \leq \gamma \\
 \text{Capacity: } h_l &\leq \frac{60\theta}{\max \Psi_l(h_l)} \text{ for all lines } l \in L \\
 \text{Allowed Dispatching Headways: } h_l &\in \Omega = \underbrace{\{2,3,4,5,6,7,8,9,10,12,15\}}_{\omega \text{ elements}} \text{ minutes}
 \end{aligned} \tag{14}$$

$\alpha_1, \alpha_2, \alpha_3, \alpha_4$ have positive values because the respective objective needs to be minimized and their units are: α_1 (monetary cost unit per waiting minute for each passenger); α_2 (monetary cost unit per passenger); α_3 (unitless); α_4 (monetary cost unit per utilized bus).

Finding the optimal dispatching headway for each bus line h_l results in a combinatorial problem since any changes in the dispatching headway of a bus line affects all other lines in the network because of changes in the allocation of a limited resource (i.e. vehicles). Hence, every choice regarding the planned dispatching headway of one line has implications to the dispatching headway choices of other lines; yielding an exponential number of $\omega^{|L|}$ combinations for calculating the optimal solution when using simple enumeration (brute-force). Due to the exponential time complexity, the problem is computationally intractable and allows an optimal solution search only on small networks.

One of the main problems of the objective function $z(\mathbf{h})$ is the absence of analytical expressions that relate the expected headway at bus stops, $h_{lj}(h_l)$ and the expected headway variance, $\text{Var}h_{lj}(h_l)$ with the planned dispatching headway, h_l . Because of this, the objective function can be either minimized with the use of simple enumeration -which is applicable only to very small networks- or, alternatively, with the use of heuristics that do not guarantee the convergence to the global optimum. For this reason, this work proposes the use of statistical analysis or empirical expressions for extracting the relationship between $h_{lj}(h_l)$, $\text{Var}h_{lj}(h_l)$ and the dispatching headway h_l from historical operational data.

Finally, the aim of the model described in Eq. (14) is to allocate buses to lines so that the objective function is minimized and, at the same time, the capacity of buses allocated to each line is sufficient for absorbing the respective passenger demand based on the max load point of the load profile. If this is not possible given the maximum allowable fleet size, γ , a feasible solution does not exist and the bus operator should decide which bus lines should exhibit oversaturation by relaxing the constraint $h_l \leq \frac{60\theta}{\max \Psi_l(h_l)}$ for some of the lines $l \in L$. Notwithstanding, our model will still attain a solution by distributing the set of available buses in such a way that the objective function is minimized.

3. Solving the continuous problem of setting dispatching headways

The discrete Nonlinear Programming Problem (NLP) for determining the optimal dispatching headways for each bus line as described in Eq. (14) can be solved using discrete optimization algorithms. Several discrete optimization algorithms, such as the Branch and Bound (B&B) method which is used in this study and is described in the following section, deploy a strategy that requires solving a series of continuous NLP sub-problems for providing a direction towards the optimal discrete solution. The continuous NLP sub-problems are simpler than the original discrete NLP of Eq. (14) and their solutions constitute a state space search for finding the discrete optimal solution. In our case, one can transform the optimal dispatching headway determination problem (Eq. (14)) to the continuous problem of Eq. (15) by allowing the planned dispatching headways (decision variables) to take real values. The discrete set of allowed dispatching headway values for each bus line ($\{2,3,\dots,15\}$ min) is now used to set the boundary constraints:

$$\begin{aligned}
 \min_{\mathbf{h} \in \mathbb{R}^L} \quad & z(\mathbf{h}) \\
 \text{subject to} \quad & c_1(\mathbf{h}) = \gamma - \sum_{l \in L} \left\lceil \frac{t_l^{90th}}{h_l} \right\rceil \geq 0 \\
 & c_{2,\dots,|L|+2}(\mathbf{h}) = \frac{60\theta}{\max \Psi_l(h_l)} - h_l \geq 0 \text{ for all lines } l \in L \\
 & c_{|L|+2,\dots,2|L|+2}(\mathbf{h}) = h_l - 2 \min \geq 0 \text{ for all lines } l \in L \\
 & c_{2|L|+2,\dots,3|L|+2}(\mathbf{h}) = 15 \min - h_l \geq 0 \text{ for all lines } l \in L
 \end{aligned} \tag{15}$$

where $c_{|L|+2}, \dots, c_{3|L|+2}$ are the boundary constraints ensuring that all planned dispatching headway values, $\mathbf{h} = \{h_{L_1}, \dots, h_{L_{|L|}}\}$, are within the limits $\{2-15\}$ min. In addition, if E is the set of equality constraints, then $E = \emptyset$, whereas the set of inequality constraints is $I = \{1,2,3,\dots,3|L|+2\}$ and the total number of constraints is $m = I + E = 3|L| + 2$. For finding the optimal solution of the dispatching

headways with numerical optimization, we generate new iterates of an *initial guess* of dispatching headways denoted as \mathbf{h} by solving inequality constraint Quadratic subproblems (QP) at each iterate. The method of Sequential Quadratic Programming (SQP) is selected instead of others (such as Interior Point or the Augmented Lagrangian Method) because it is more suitable for tackling the nonlinear constraints of Eq. (15). Interior point methods (also referred to as barrier methods) and Lagrangian methods utilize multipliers to treat the constraints as penalties after introducing them to the main objective function. As expressed in many exact optimization books such as Nocedal and Wright (2006), these approaches, which are effective when constraints are linear, underperform compared to the SQP method when the optimization problem has significant nonlinearities in the constraints.

The SQP, which is developed in the 1970s, is well-described in the book of Nocedal and Wright (2006) and the article of Boggs and Tolle (1995) from which we adopted the main elements of the description of the SQP method that are presented in the appendix. SQP generates new iterates of the *initial guess* dispatching headways, $\mathbf{h}_{k=0}$, by solving inequality constraint Quadratic sub-problems (QP) at each iterate k . The idea behind the SQP solution method is to model the dispatching headways of the current iterate, \mathbf{h}_k , by a quadratic programming subproblem and then use the minimizer of this sub-problem to define a new iterate of dispatching headways, \mathbf{h}_{k+1} , until convergence.

For all those inequality QP sub-problems that should be solved at each iteration, the well-known active-set method can be utilized. In the active-set method, the equality constraints are always active and the active-set is updated at any iteration by solving an equality QP where different inequality constraints are considered as active (one can refer to Murty and Yu (1988) for a more detail description of the active set method).

4. State-space search for the discrete problem of setting dispatching headways with Branch and Bound (B&B)

After solving the continuous frequency setting problem assuming that dispatching headway variables for each bus line $\mathbf{h}^* = \{h_{L_1}, \dots, h_{L_{|L|}}\}$ can take any real value within the range of [2–15] min, one has to find the optimal solution of the discrete optimization problem where \mathbf{h} values belong to the discrete set $\Omega = \{2, 3, 4, \dots, 15\}$ min.

The goal of the branch-and-bound (B&B) method is to find the values of dispatching headways, \mathbf{h} , that minimize the value of the scalar objective function of Eq. (14), among some set Ω of candidate solutions. The search space contains $\omega^{L|}$ elements where ω is the length size of the discrete set $\Omega = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15\}$ min. from which the dispatching headways can take their values. It is evident that brute-force cannot be applied in our case because a mid-sized bus network with 35 bus lines requires the inspection of all $2.8102437e + 36$ elements of set Ω .

In B&B, our starting point is the enumeration tree which branches dynamically through newly developed nodes. Initially, our enumeration tree has only one node, the tree root. In our case, the root is the solution of the continuous frequency setting problem where the dispatching headways were considered continuous variables: \mathbf{h}^* . The objective function value of the continuous solution is $z(\mathbf{h}^*)$ and is the lowest bound (LB) of the INLP of Eq. (14). This is the best solution so far, known as the *incumbent solution*, and $z(\mathbf{h}^*)$ is the *incumbent objective*. However, this is not the solution of our discrete problem of setting the dispatching headways and, due to the absence of a feasible solution for the discrete NLP, we initialize the upper bound value of the problem as $b^{upper} = +\infty$.

A typical iteration has three main components: selection of the node to process, branching and bound calculation. If it can be established that the subspace of a node cannot contain the optimal solution, the whole subspace is discarded, else it is stored in the pool of live nodes $A = N - F$ where N is the set of all generated nodes and F is the set of discarded nodes. Its objective function value is also stored in the set of bounds B .

Initially, the selection of the node to process is trivial because we have only one node (the root). Then, new branches and nodes are dynamically generated. During branching, we assign all discrete dispatching headway values from set Ω to one variable $h_i \in \mathbf{h}$; thus, generating ω new branches starting from the root while all other variables remain continuous (real numbers). All new branches have new end-nodes and each end-node $i \in N$ has a set of assigned integer values for some dispatching headway variables \mathbf{h} while all other variables \mathbf{h} are continuous. The objective value (bound) of each node is calculated by solving the *restricted continuous NLP problem of setting the dispatching headways* with the algorithmic framework described in the previous section by assuming additional equality constraints that dictate a number of continuous dispatching headway variables \mathbf{h} to be equal to their already assigned integer values for this node. If at one node there are already g assigned integer values from set Ω to variables $\{h_1 = g_1, \dots, h_g = g_g\}$, then the bound of this node is computed from the restricted continuous NLP:

$$\begin{aligned}
 & \min_{\mathbf{h} \in \mathbb{R}^L} \quad z(\mathbf{h}; \{h_1, \dots, h_g\}) \\
 & \text{subject to} \quad c_1(\mathbf{h}) = \gamma - \sum_{l \in L} \left[\frac{t_l^{90th}}{h_l} \right] \geq 0 \\
 & \quad c_{2, \dots, |L|+2}(\mathbf{h}) = \frac{60\theta}{\max \Psi_l(h_l)} - h_l \geq 0 \text{ for all lines } l \in L \\
 & \quad c_{|L|+2, \dots, 2|L|+2}(\mathbf{h}) = h_l - 2 \min \geq 0 \text{ for all lines } l \in L \\
 & \quad c_{2|L|+2, \dots, 3|L|+2}(\mathbf{h}) = 15 \min - h_l \geq 0 \text{ for all lines } l \in L \\
 & \quad c_{3|L|+3}(\mathbf{h}) = h_1 - g_1 = 0, \text{ where } g_1 \in \Omega \\
 & \quad \dots \dots \dots \\
 & \quad c_{3|L|+g+2}(\mathbf{h}) = h_g - g_g = 0, \text{ where } g_g \in \Omega
 \end{aligned} \tag{16}$$

The solution of the restricted continuous NLP with $\{h_1, \dots, h_g\}$ already assigned dispatching headway values from set Ω is the bound of this node since if we continue branching from this node (i.e., assign integer values from Ω to more continuous variables of this

restricted continuous NLP) the newly generated subproblems would return *inferior* objective function values. After the third step of every B&B iteration (bound calculation of all newly generated nodes) we discard entire subspaces if it is established that some of the nodes cannot contain the optimal solution. If, for instance, one newly generated node cannot provide a feasible solution (i.e., given the already assigned values from set Ω to some of the variables $\{h_1, \dots, h_g\}$ of this node there is no solution of the restricted continuous NLP that satisfies all existed constraints), then this node and its entire subspace can be disregarded since there are no continuous values of the dispatching headways that can solve this restricted problem; thus, there would also not be any discrete dispatching headway values that provide a feasible solution. All discarded nodes are added to the set F whereas the remaining live set $A = N - F$ is left with fewer nodes. In addition, after each iteration the selected node that generated the new branches is also discarded from further consideration and is added to F .

The node for the new B&B iteration is the *incumbent solution* node, which is the node i at the live set $A = N - F$ that has the lowest bound value $B_i = \min B_j, \forall j \in A$. If after a number of B&B iterations we have a node at which all dispatching headways \mathbf{h} have been assigned with discrete values from the set Ω , then we have a first possible solution for the discrete NLP. Then, we can proceed with a more vigorous strategy for ruling out nodes from the set A and “pruning” the solution space. For doing this, we replace the upper bound of the problem $b^{upper} = +\infty$ with the objective function value of this candidate discrete NLP solution. Initially, nodes that belong to set A and their bound value is higher than the b^{upper} are omitted because if they currently have inferior objective function values, then their objective function values will remain inferior and would further degrade in next instances when some of their remaining continuous variables would be replaced with discrete ones. In addition, if later on we find another possible discrete solution of the NLP with lower objective function value than b^{upper} , then this becomes the currently chosen discrete NLP solution and we update also the value of b^{upper} while the procedure continuous until there are no remaining nodes at the live set $A = \emptyset$ (all branching possibilities have been exhausted).

To summarize the above, the algorithmic framework that is introduced for setting the optimal dispatching headways of bus lines while also considering operational data insights regarding the stop-level headway variability that is translated to waiting time reliability is composed of: (i) a method that estimates the relationship between the dispatching headways and the expected operational headways at each stop from historical data; (ii) a SQP algorithm that solves the continuous NLP problem of setting the dispatching headways; and (iii) a B&B algorithm that uses SQP at each iteration for solving restricted continuous NLP sub-problems and allocating buses to lines. This procedure is described in Fig. 2 for providing an overview of the proposed framework.

5. Experimental results

5.1. Case study description

The proposed method is applied to a case study network in Stockholm, Sweden. Our study area is the bus network of central Stockholm which contains 17 bus lines, $L = \{1, 2, 3, 4, 50, 53, 55, 56, 59, 61, 65, 66, 69, 72, 73, 74, 77\}$, for which GTFS data is available. Fig. 3 presents the case study network.

Stockholm bus operators design their schedules based on the 90th percentile of the vehicle running times during the respective time-of-day, day-of-the-week and season. In general, they derive the planned dispatching headways of the service lines based on the average load profile method while also considering a set of policy constraints.

In our experimentation, we select first two lines for detailed analysis in order to enable the enumeration of all solutions and benchmark the B&B/SQP approach against brute-force. Second, we apply our method to 17 lines operating in Stockholm inner-city to test the scalability of our method and its performance for a real-sized network.

5.2. Demonstration using two selected lines

Our small-scale demonstration of setting optimal dispatching headways uses data from bus lines 1 and 3, two high demand bus lines in the case study network. Detailed Automatic Vehicle Location (AVL) and Automatic Passenger Counting (APC) data are available for these lines for a five months period, from August to December 2011. Line 1 connects the main eastern harbor to a residential area in the western part of the city through the commercial centre. Line 3 serves as a north–south connection through Stockholm’s old city, connecting two large medical campuses. The datasets contain a total number of 8,241 trips and the vehicle travel time for each line (per direction) expressed as mean \pm standard deviation is presented in Table 1. Table 1 presents also the total number of boarding passengers per line trip for each direction and the 90th percentiles of the total round trip travel times.

For simplifying the notation, we denote the planned dispatching headway variables for each line as $h = \{h_1, h_2\}$. The bus stops of the bi-directional line 1 are $S_1 = \{1, 2, 3, 4, \dots, 65\}$ where 33 stops belong to direction 1 and 32 stops to direction 2. The bus stops of line 3 are $S_2 = \{1, 2, 3, 4, \dots, 51\}$ where 26 belong to the 1st direction and 25 to the 2nd. In this demonstration, we set the dispatching headways for the time period 8:00 am to 12:00 pm. The passenger boardings, $b_{l,j,t}$, alightings and the bus loads, $\Psi_{l,j,t}$, for that time period are presented in Fig. 4. Those values are the average hourly boardings/alightings and loadings from August to December 2011. For brevity reasons, Fig. 4 presents only the hourly demand levels of bus line 1.

As shown in Fig. 4, the most boardings at direction 1 occur at stops 5–9 and 11–15 while the highest loading levels are observed at stop 10 throughout the 4-h period. On the other hand, the boardings at direction 2 change significantly from hour to hour and do not follow a certain pattern. The same holds true for the passenger loading at stops where stops 7, 14–21 appear to be the most loaded ones but do not follow a common pattern over the 4-h period.

Apart from the passenger demand, the AVL data from those five months is used for deriving the relationship between the

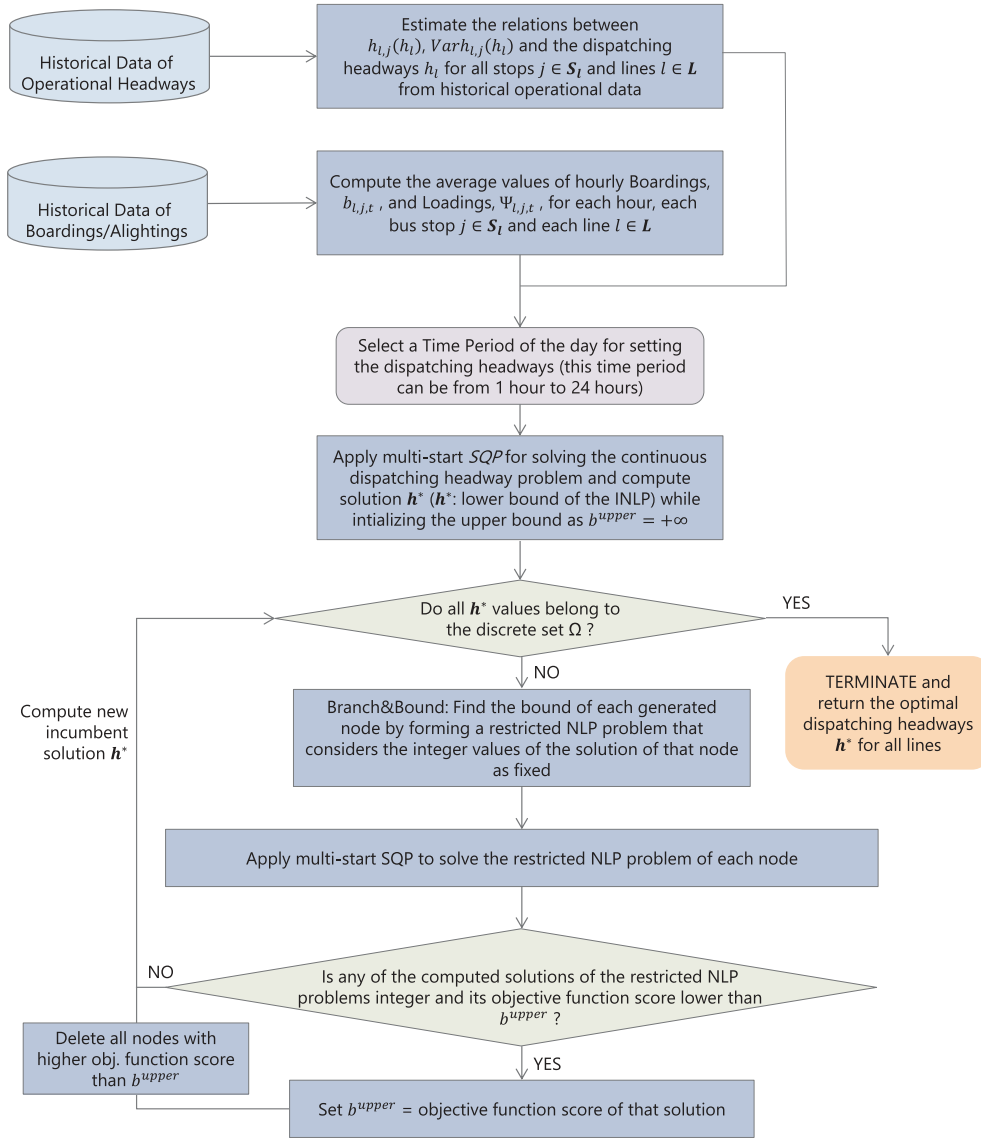


Fig. 2. Overview of the solution framework for exact optimization of dispatching headways.

dispatching headways and the headways at the bus stops. In this work, we utilize the Support Vector Regression (SVR) method using non-linear kernels. SVR is selected as the mean to establish the relationship between the dispatching headways and the resulting operational headways at stops but, in principle, it can be substituted by any other such method or empirical relationship. SVR is selected because it can be embedded within the SQP optimization framework and allows the SQP to compute objective function values for any possible dispatching headway solution. We should also note here that it is also a generic method and can be applied to datasets from different cities after an initial training.

The SVR model is trained with the use of historical data sets of observed headways and returns the expected values $h_{l,j}(h_l), Var h_{l,j}(h_l)$ for different dispatching headways h_l . As demonstrated in Fig. 5, the relationship between the headway at the first bus stop and the observed headways at the other stops may not be well-explained using linear fitting. For this reason, we use the Gaussian Radial Basis Function (RBF) kernel to compute the inner product in a higher dimensional feature space. For instance, if for some dispatching headway values $r_x = \{h_l, h'_l, h''_l, \dots\}$ we observed headways $r_y = \{h_{l,j}, h'_{l,j}, h''_{l,j}, \dots\}$ at stop j ; then, the Gaussian RBF is:

$$K(r_x, r_y) = \exp(-\gamma \|r_x - r_y\|^2) \quad (17)$$

For deriving the expected headway at every bus stop for different values of dispatching headways, RBF support vector regressions are performed. After applying the SVR RBF method to the historical data, we present the expected relation between the dispatching headways and the expected headway variability at any other stop in Fig. 6.

The key observations from Fig. 6 are: (i) the expected headway variation is close to zero at the first bus stops and exercises a

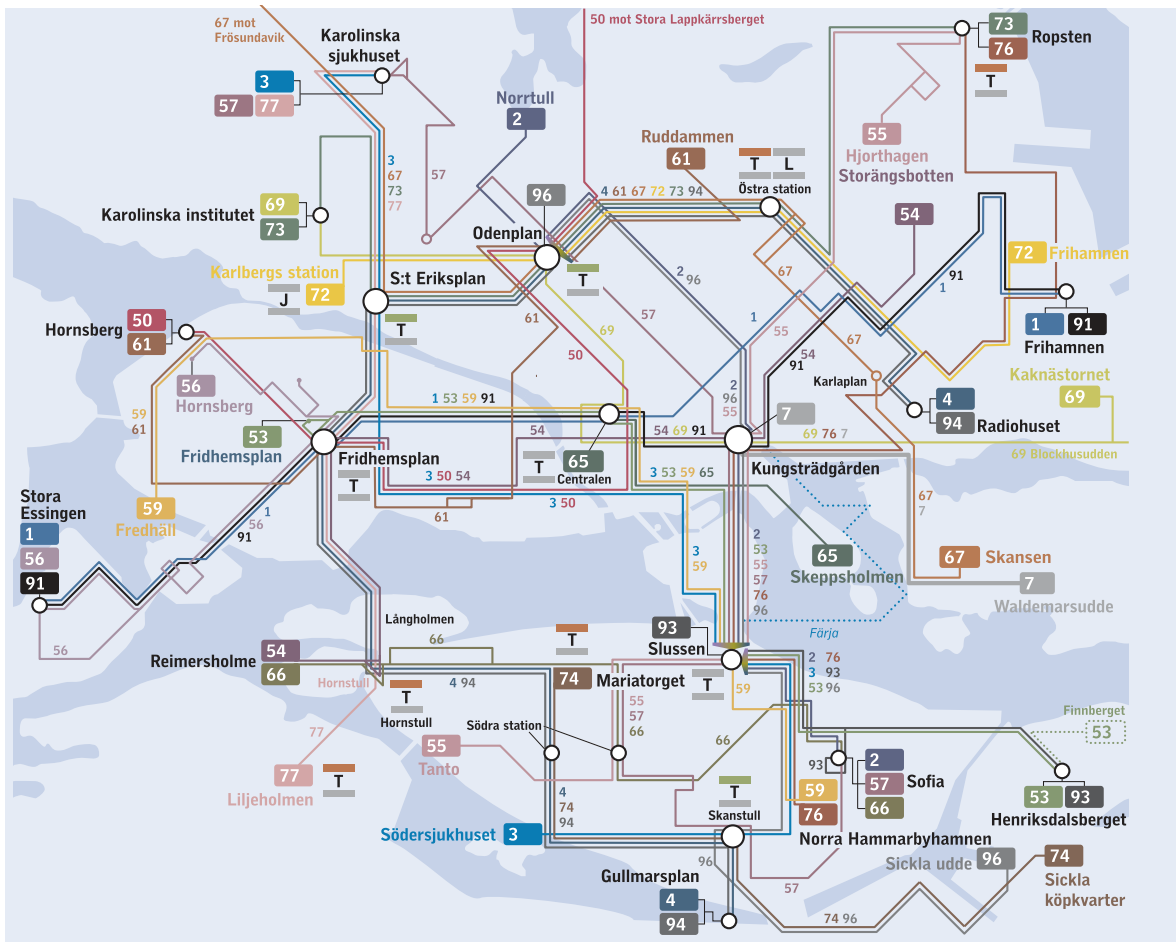


Fig. 3. A schematic network representation of Stockholm inner-city bus lines.

Source: <http://sl.se/ficktid/karta/vinter/SthlmInnerstad.pdf>.

Table 1

Statistics per line direction. The values are presented as mean \pm s.d.

	Trip travel times (s)	Passenger boardings (per trip)	Round trip t_l^{90th} (s)
Line 1: dir.1	3017 \pm 425	101 \pm 50	6807
dir.2	2755 \pm 480	98 \pm 51	
Line 3: dir.1	2607 \pm 465	70 \pm 37	6486
dir.2	2746 \pm 448	60 \pm 29	

proportional increase when moving to downstream stops; (ii) dispatching headway values of 4–12 min result in higher headway variations at the last stops of line 1 compared against dispatching headway values of 2–4 or 12–15 min; and (iii) line 3 does not show a clear pattern - dispatching headways of 2–6 min result in a higher downstream headway variability for direction 1 whereas the same variability for direction 2 is observed for dispatching headways of 6–15 min.

At this point, we should note that the headway variance at each stop measures the difference of the observed headways at that stop from the respective dispatching headways (in ideal operations where the travel times between bus stops and the dwell times remain stable, this variance should be equal to zero and the observed headway at any stop should be equal to the dispatching headway).

In this small-scale, two-line experiment, we can apply brute-force optimization for exploring the solution space and converging to the global optimum. The total number of buses that are available for bus lines 1 and 3 from 08:00 to 12:00 is $\gamma = 40$ and the vehicle capacity $\theta = 40$ passengers. In addition, the weight factors $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ that translate all objectives of Eq. (14) to a single compensatory term are selected so that the passenger waiting times $O_1(h_i)$ and the coefficient of variation $O_2(h_i)$ have almost equivalent importance as the running $O_3(h_i)$ and fixed $O_4(h_i)$ costs for the currently deployed solution (dispatching headways $h_1 = 5$ –8 min and $h_2 = 5$ –8 min

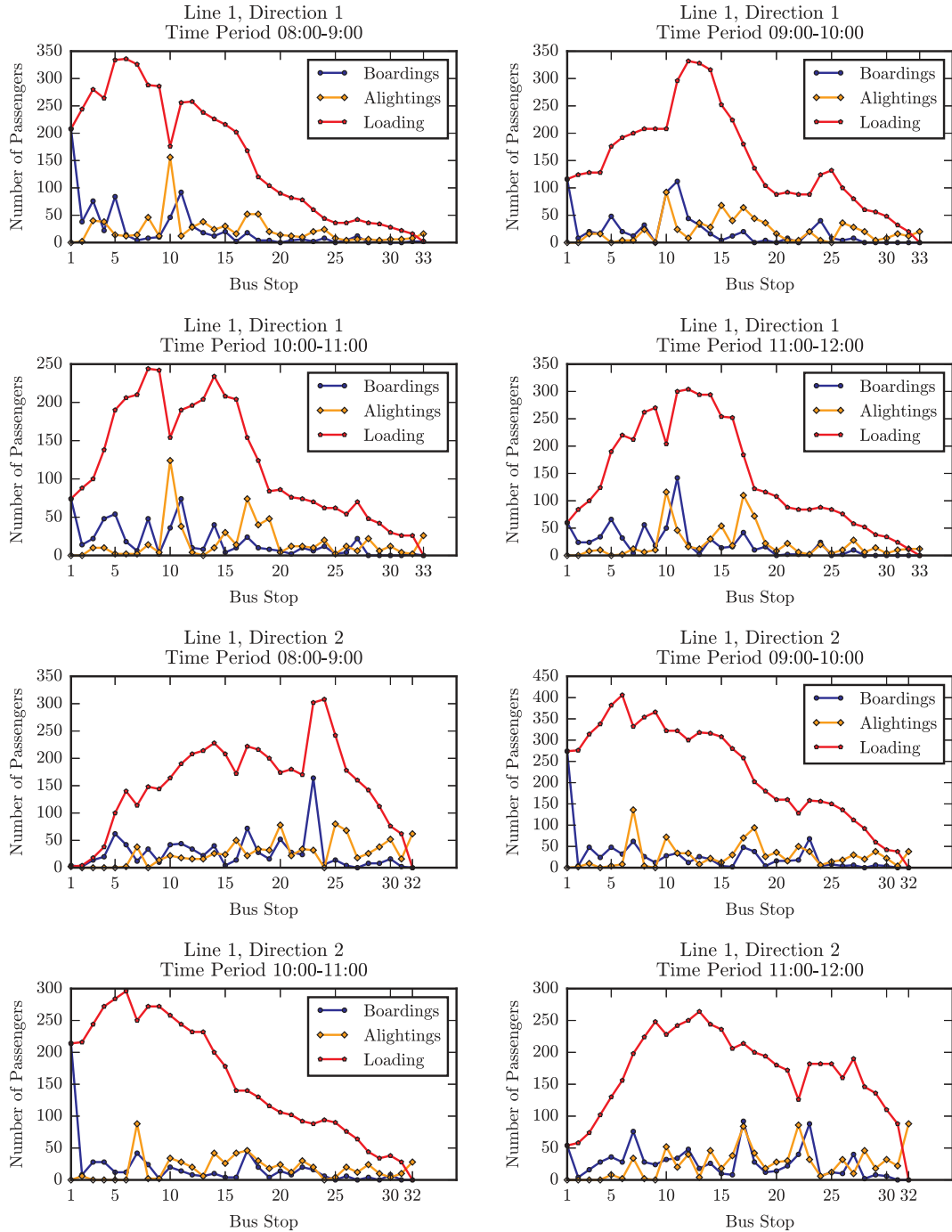


Fig. 4. Boardings, alightings and bus loadings for each hour from 08:00 to 12:00 for Line 1 (directions 1 and 2). At each hour, the trip frequency is ~ 8 –10 bus trips per line direction.

from 08:00 to 12:00). By adopting this approach, we derive the underlying weights that were implicitly used when setting the current dispatching headways which have balanced user and operator costs in the local context. Ratios between weights were set based on the local estimates of cost parameters Tra (2015) and this resulted in the following weight factor values: $\alpha_1 \approx 0.5$; $\alpha_2 \approx 1.5$; $\alpha_3 \approx 1.5$; $\alpha_4 \approx 1$.

Using the above values, the result of setting the dispatching headways expressed in Eq. (14) is presented in Fig. 7 where the 2D plot enumerates all possible feasible and infeasible solutions after evaluating $\omega^{JL} = 112$ times the objective function. It can be observed by simple inspection that the dispatching headways $(h_1, h_2) = (8, 8)$ min with objective function score $z = 49280.554$ is the global optimum solution. We further analyze the result by investigating separately the results for the passenger-related and the operator-

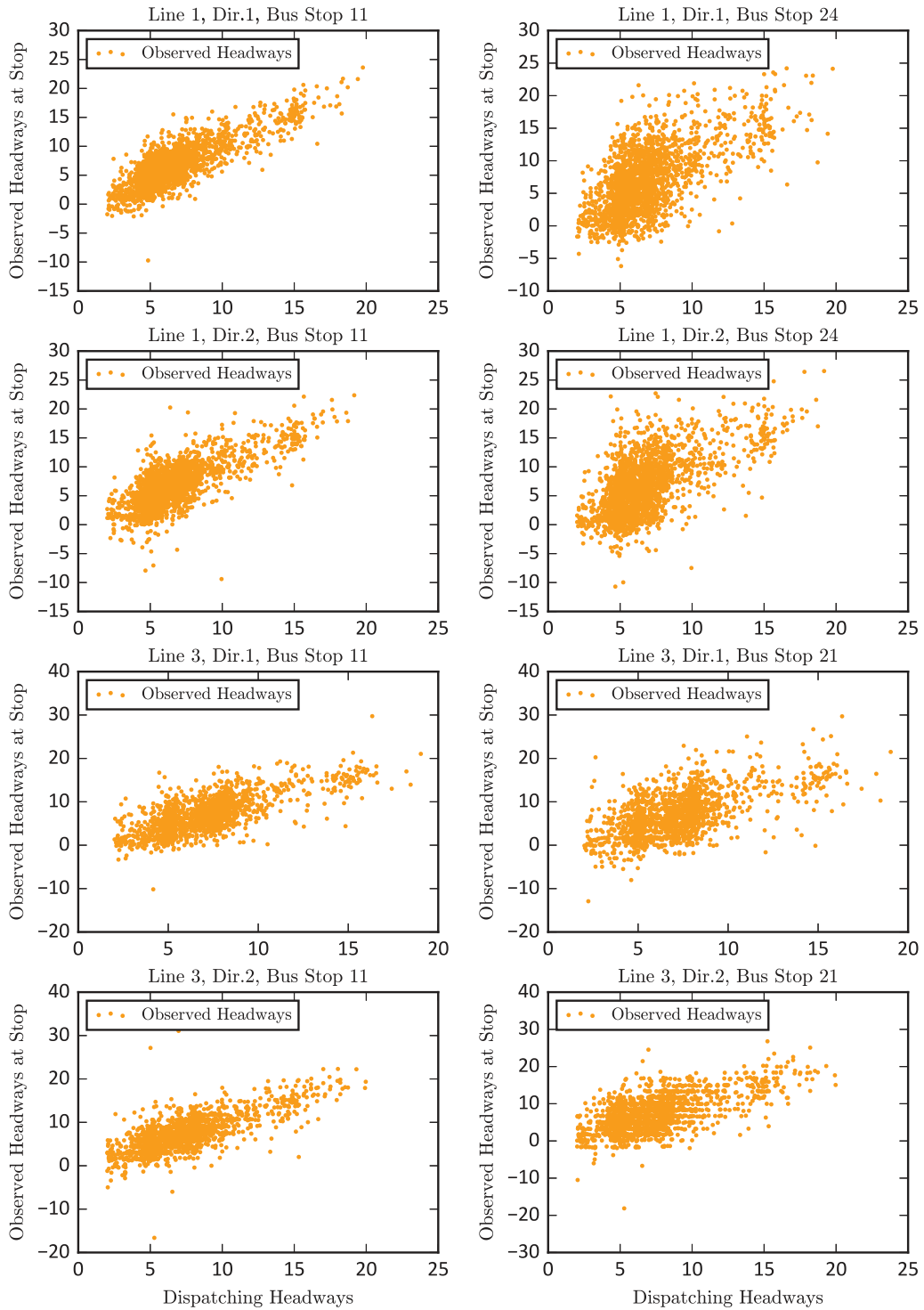


Fig. 5. Observed headways for different values of dispatching headways at selected stops according to the 5-month AVL data.

related objective function components. As it is clear from the problem formulation of Eq. (14), the first two terms of the objective function that are multiplied by weights α_1, α_2 reflect passenger waiting times and service reliability whereas the last two terms, that are multiplied by weights α_3, α_4 , correspond to operational costs. Therefore, in Fig. 8 we plot the first two and the last two terms of the objective function separately to explore the properties of passenger-related and operational-related costs for different values of the dispatching headways.

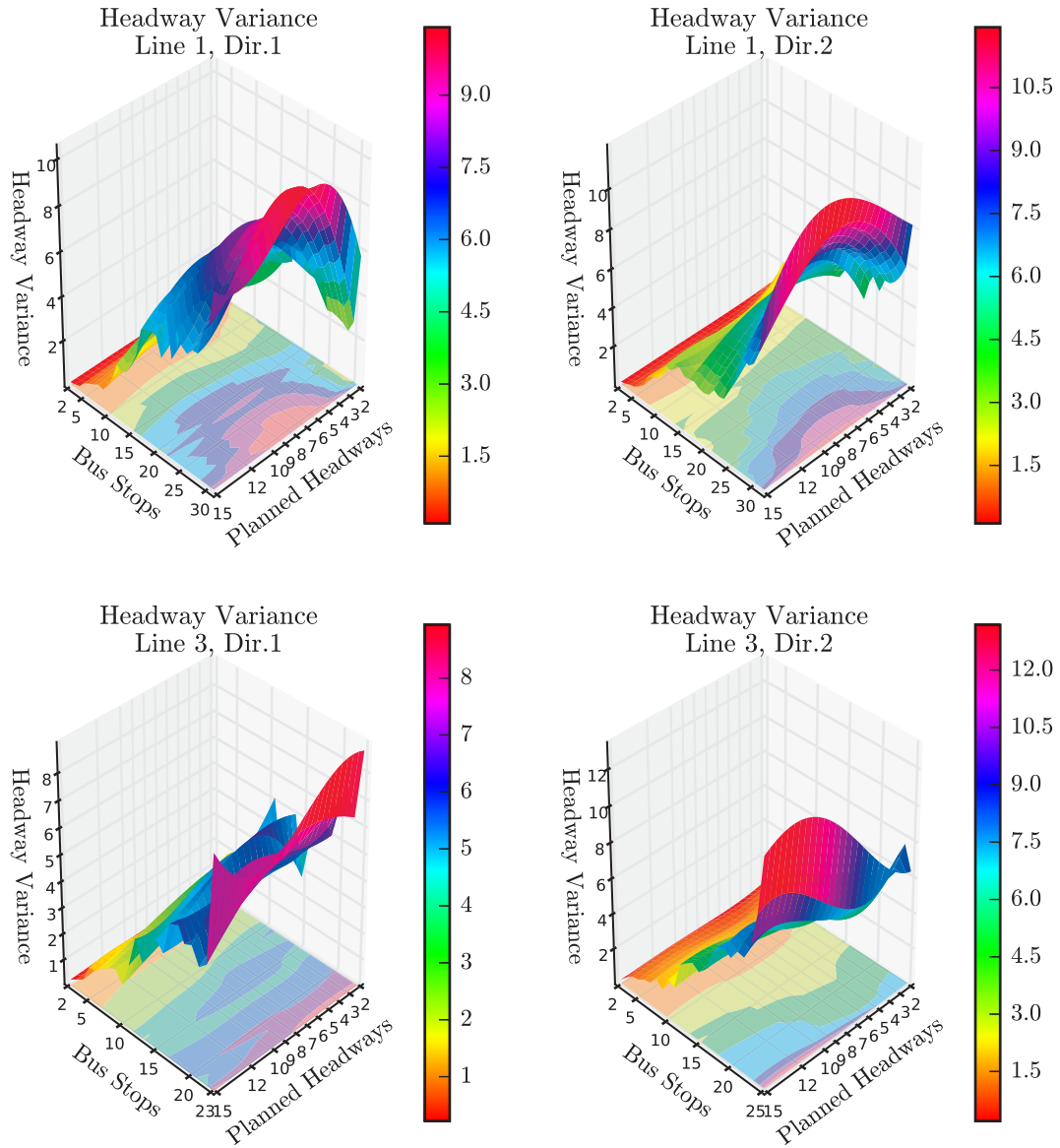


Fig. 6. Relation between the expected headway variance at each downstream stop and the planned dispatching headway as it is derived after applying RBF Support Vector Regression.

Fig. 8a depicts the form of the objective function if one considers only the first two terms of the objective function that are related to the passenger costs. From Fig. 8a one can observe that the dispatching headways that minimize the passenger-related costs are $(h_1, h_2) = (5, 6)$ min and their values are lower than the computed dispatching headways, $(h_1, h_2) = (8, 8)$ min, when considering also the operational costs. The reason behind this is that the function form of the operational costs which is presented in Fig. 8b exhibits a monotonically decreasing behavior that results in optimal dispatching headways of 15 min. This monotonic behavior is expected because if the dispatching headways increase, the number of bus trips decreases; thus, the operational costs decrease as well. Therefore, compromising the conflicting priorities of the passenger-related and the operational-related costs with the use of ratios between weights that are set based on local estimates of cost parameters Tra (2015) ($\alpha_1 \approx 0.5; \alpha_2 \approx 1.5; \alpha_3 \approx 1.5; \alpha_4 \approx 1$), results in the optimal solution of $(h_1, h_2) = (8, 8)$ minutes. This analysis can facilitate a discussion on the passenger-related costs and the operational-related costs with the transport planning authority for re-evaluating the ratios between different objective function terms. For instance, the transport authority might be willing to increase further the operational costs for satisfying the passenger-related costs by applying dispatching headways which are closer to the $(h_1, h_2) = (5, 6)$ min values.

Returning to the analysis of Fig. 7 that considers both the passenger-related and the operational-related costs, one can observe that very short dispatching headways yield infeasible solutions because they require many buses for performing the operations. For instance, for operating under dispatching headways $(h_1, h_2) = (2, 2)$ min a total number of $\left\lceil \frac{t_1^{90th}}{h_1} \right\rceil + \left\lceil \frac{t_2^{90th}}{h_2} \right\rceil = 112$ buses is required

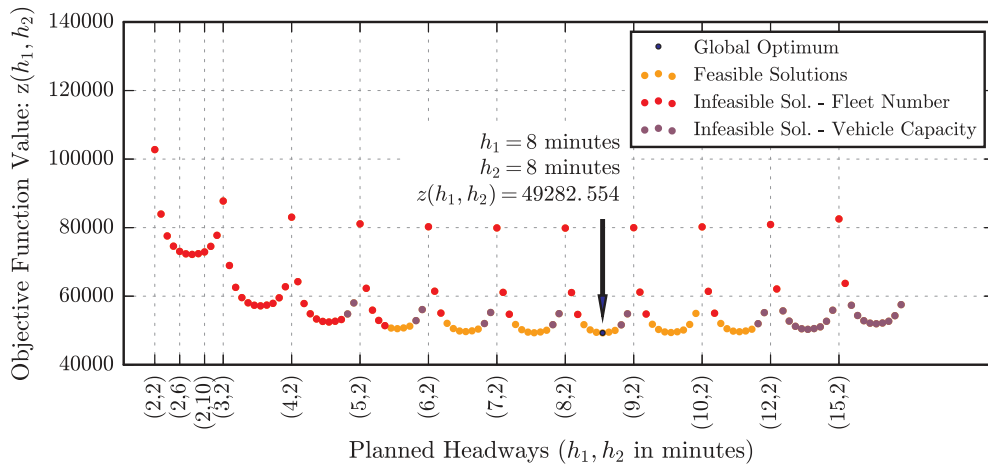


Fig. 7. Objective function score for different planned dispatching headways for line 1 and 3. Infeasible solutions violate either the line capacity constraint or require more buses than the available ones.

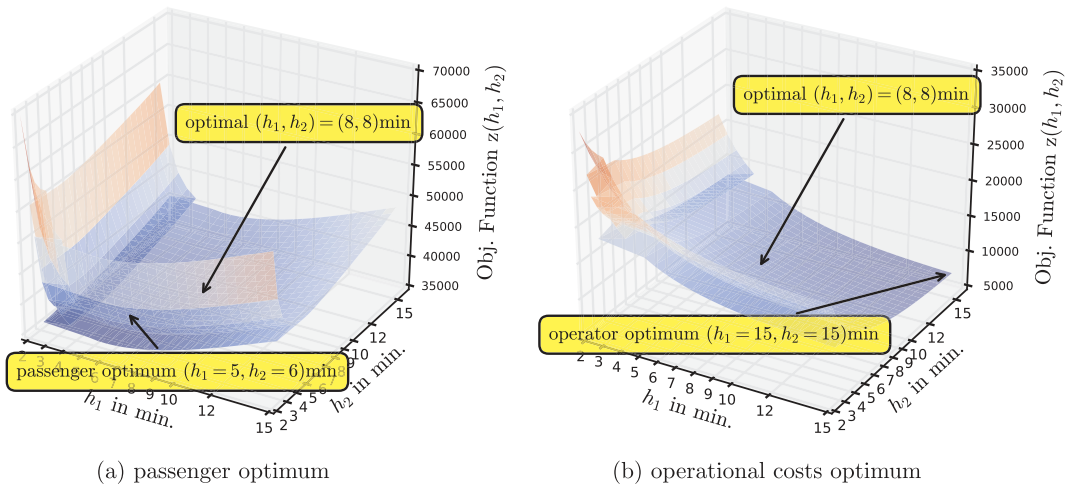


Fig. 8. Plot of the passenger-related optimum [left] and the optimum that minimizes the operational costs [right].

where the total number of available buses is $\gamma = 40$. Apart from the fleet number violation though, very short dispatching headways had also higher objective function scores because the operational costs are higher. In addition, dispatching headways of more than 12 min violated the vehicle capacity constraint for both lines 1 and 3 because the bus loading at the most congested stop is higher than the vehicle capacity of $\theta = 40$ passengers. Finally, the optimal solution, $(h_1, h_2) = (8, 8)$ min, requires 29 buses which are less than the total number of available buses.

In this two-line experiment, we can also demonstrate the use of our algorithmic framework. At first, the continuous NLP problem of setting the dispatching headways is solved with the SQP algorithm returning dispatching headways $\{\mathbf{h}^* = 7.73, 7.89\}$ min which is the lowest bound of the discrete NLP with $z(\mathbf{h}^*) = 49273.971$. After three branching iterations presented in Fig. 9, the B&B attains a discrete solution $(h_1, h_2) = (8, 8)$ min with $z(\mathbf{h}^*) = 49282.554$. The B&B search terminates at the third iteration because an integer solution $(h_1, h_2) = (8, 8)$ min is calculated and the upper bound becomes equal to $b^{upper} = 49282.554$ which is lower than any other value from the remaining nodes that belong to the live set $A = \{2, 4, 5\}$.

In Fig. 9 we obtained the objective function score of each one of the five (5) generated nodes of the B&B method by applying the SQP optimization method for solving the respective nonlinear optimization problem. In order to demonstrate how the proposed method performs in terms of iterations until convergence, we present the iterations of the SQP optimization algorithm for solving each one of the 5 nonlinear optimization problems in Table 2. Each nonlinear problem was solved in 4–13 iterations with the use of SQP and the total number of evaluations of the objective function with our method is $13 + 10 + 13 + 4 + 6 = 46$ compared to the 121 evaluations required by the simple enumeration method that was presented in Fig. 7. As presented in Table 2, the objective function values oscillate significantly at the first iterations of the SQP method but stabilize rapidly (in most cases, after 2–5 iterations).

The optimal dispatching headways are expected to vary when the passenger demand or the operational costs change. This is examined in this study by performing a sensitivity analysis for deriving the stability region of each dispatching headway solution. Fig. 10 presents the optimal dispatching headways for different levels of passenger demand and bus running costs. From Fig. 10a one

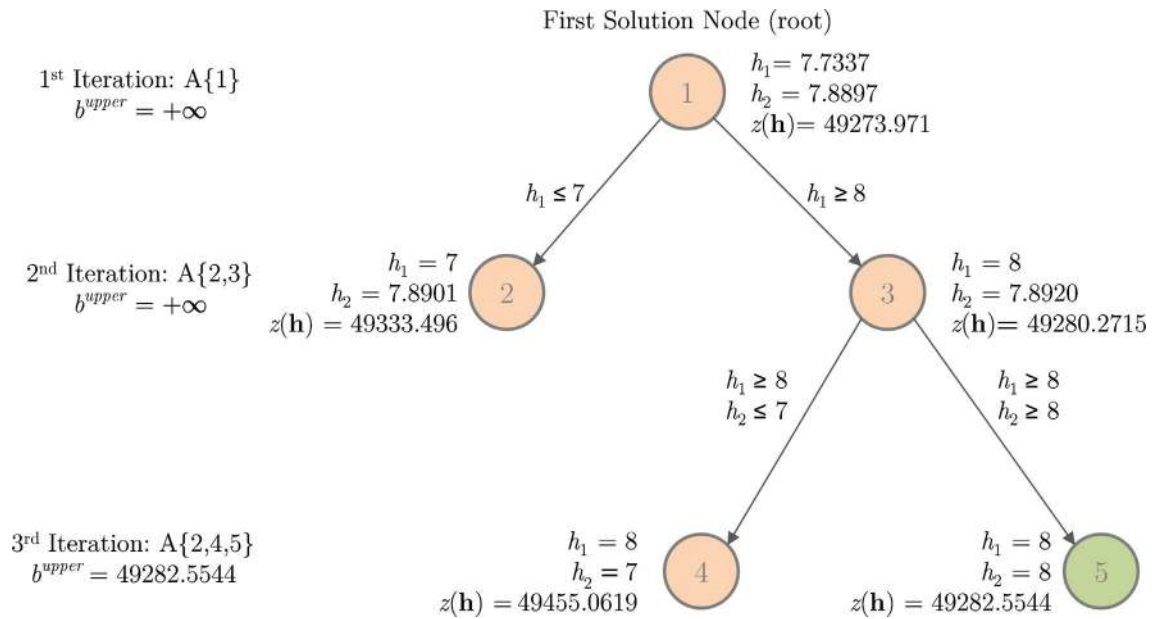


Fig. 9. Solution of the discrete NLP problem of setting the optimal dispatching headways with branch and bound.

Table 2

SQP iterations required until convergence along with the respective objective function score changes for each one of the branches generated by the B&B method.

Iter.	$2 \leq h_1 \leq 15$ $2 \leq h_2 \leq 15$			$2 \leq h_1 \leq 7$ $2 \leq h_2 \leq 15$			$8 \leq h_1 \leq 15$ $2 \leq h_2 \leq 15$		
	h_1	h_2	$z(\mathbf{h})$	h_1	h_2	$z(\mathbf{h})$	h_1	h_2	$z(\mathbf{h})$
1	14.986	15.000	57516.91	6.988	15.000	54908.04	14.986	15.000	57516.91
2	9.250	2.000	80031.52	7.000	2.000	79918.44	9.250	2.000	80031.52
3	8.924	6.960	49575.33	7.000	2.000	79918.44	8.924	6.960	49575.33
4	8.831	8.172	49383.41	7.000	4.346	53452.67	8.831	8.172	49383.41
5	8.803	7.960	49364.70	7.000	7.426	49379.58	8.803	7.960	49364.70
6	8.731	7.836	49353.13	7.000	7.957	49334.32	8.731	7.836	49353.13
7	8.528	7.716	49332.77	7.000	7.896	49333.50	8.528	7.716	49332.77
8	8.015	7.617	49297.42	7.000	7.890	49333.50	8.015	7.617	49297.42
9	7.776	7.722	49279.91	7.000	7.890	49333.50	8.000	7.745	49284.64
10	7.717	7.877	49274.05	7.000	7.890	49333.50	8.000	7.886	49280.28
11	7.733	7.889	49273.97				8.000	7.893	49280.27
12	7.734	7.890	49273.97				8.000	7.891	49280.27
13	7.733	7.890	49273.97				8.000	7.888	49280.27
Iter.	$8 \leq h_1 \leq 15$ $2 \leq h_2 \leq 7$			$8 \leq h_1 \leq 15$ $8 \leq h_2 \leq 15$			$8 \leq h_1 \leq 15$ $8 \leq h_2 \leq 15$		
	h_1	h_2	$z(\mathbf{h})$	h_1	h_2	$z(\mathbf{h})$	h_1	h_2	$z(\mathbf{h})$
1	14.991	7.000	52122.34	14.986	15.000	57516.91			
2	8.823	7.000	49541.87	8.886	8.000	49379.71			
3	8.000	7.000	49455.06	8.802	8.000	49365.85			
4	8.000	7.000	49455.06	8.404	8.000	49313.31			
5				8.000	8.000	49282.55			
6				8.000	8.000	49282.55			

can observe that the dispatching headway solution $(h_1, h_2) = (8, 8)$ min is valid if the passenger demand is reduced by less than 40% or does not increase by more than 33%. If it decreases by more than 40%, the optimal dispatching headway of bus line 1 becomes 9 min. Even if the optimal solution remains stable for up to 40% passenger demand reduction; further reductions affect significantly the optimal solution because the running costs for serving low-demand bus services become too high. Consequently, for another 40% passenger demand reduction the optimal dispatching headway of bus line 1 increases significantly to 12 min and of line 3–10 min. Then, the dispatching headways take their higher possible values of $(h_1, h_2) = (15, 15)$ min for a significant demand reduction of more than 94%.

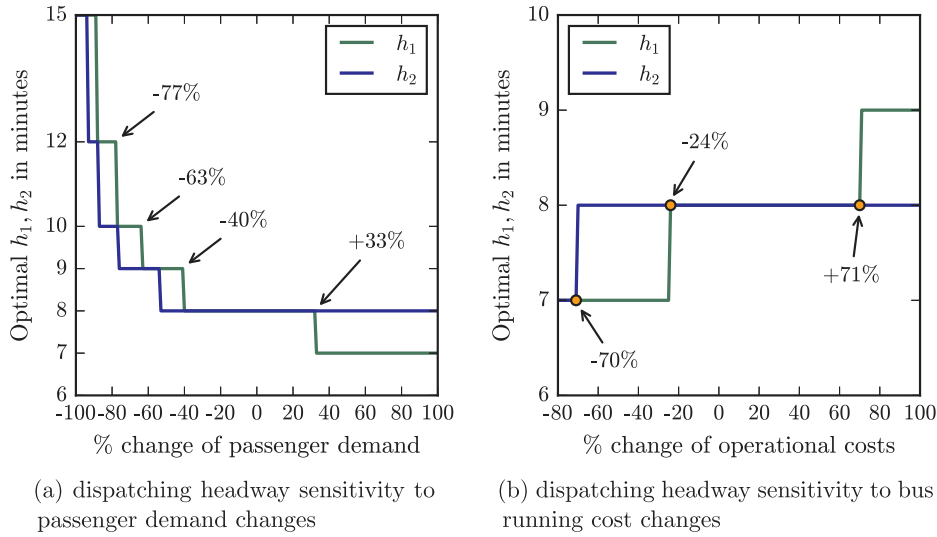


Fig. 10. Sensitivity of optimal frequency setting solutions.

The dispatching headways demonstrate a different sensitivity pattern when the bus running costs change as presented in Fig. 10b. For instance, a 24% reduction of running costs requires a change of optimal dispatching headways into $(h_1, h_2) = (7, 8)$ min whereas a reduction of more than 70% change the optimal dispatching headways values to $(h_1, h_2) = (7, 7)$ min. In addition, if the running costs rise by more than 71%, bus line 1 should increase its planned dispatching headway to 9 min. The information of the stability regions of the optimal headways is very useful for bus operators because it gives them a degree of confidence that the deployed dispatching headways remain optimal when slight passenger demand and running cost variations are observed. Furthermore, those regions provide information to bus operators regarding the threshold values of passenger demand and running costs upon which they should act by changing the values of the deployed dispatching headways.

5.3. Impacts of considering service reliability in setting the dispatching headways

The optimal dispatching headways of this small-scale, 2-line scenario were computed by considering also the service reliability aspect expressed in Eq. (14). In Eq. (14) the service reliability recognizes the passenger waiting time variability at downstream stops during real-world operations and is also explicitly expressed in the objective function as the coefficient of variation of passenger waiting times at each stop as proposed by Chen et al. (2009). If one does not consider the service reliability in the analysis, the waiting time variations of passengers at stops are excluded from Eq. (14) and the problem of setting the dispatching headways becomes merely a trade-off between passenger demand coverage and running costs reduction as expressed in the modified Eq. (18) that does not consider operational headway variabilities.

$$\begin{aligned}
 \min_{\mathbf{h}=\{h_{L1}, \dots, h_{Ll}, \dots\}} \quad & z(\mathbf{h}) = \alpha_1 \sum_{l \in L} \sum_{j \in S_l} \left(\frac{h_l}{2} \right) \sum_{t=t_1}^{t_{lT}} b_{l,j,t}(h_l) + \alpha_3 \sum_{t=t_1}^{t_{lT}} \sum_{l \in L} \frac{k_l}{h_l / 60} + \alpha_4 \sum_{l \in L} \left[\frac{t_l^{90th}}{h_l} \right] \\
 \text{subject to:} \quad & \text{Available Buses: } \sum_{l \in L} \left[\frac{t_l^{90th}}{h_l} \right] \leq \gamma \\
 & \text{Capacity: } h_l \leq \frac{60\theta}{\max \psi_l(h_l)} \text{ for all lines } l \in L \\
 & \text{Allowed Dispatching Headways: } h_l \in \Omega = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15\} \text{ minutes} \\
 & \quad \quad \quad \omega \text{ elements}
 \end{aligned} \tag{18}$$

Optimizing the base-case scenario of Eq. (18) that does not consider operational headway variations results in different optimal dispatching headway values ($h_1 = 8, h_2 = 7$) min as expressed in Fig. 11.

Fig. 11 shows the objective function form for different values of dispatching headways. As it is evident from Fig. 11, the objective function has significantly higher scores for dispatching headways of more than 10 min and less than 4 min. This occurs because (i) the line-level unsatisfied demand increases rapidly for a dispatching headway of more than 10 min and penalizes disproportionately the objective function; and (ii) the line-level operational costs increase significantly for dispatching headways of 4 min or less due to the need of many more running buses.

In contrast, our reliability-based optimization of the dispatching headways, which was demonstrated at the previous sub-section, yields a profoundly different objective function as can be observed in Fig. 12.

The form of the objective function of Fig. 12 for different dispatching headway values is less smooth due to the consideration of headway variability at downstream stops and the explicit inclusion of the reliability-dependent expression for passenger waiting

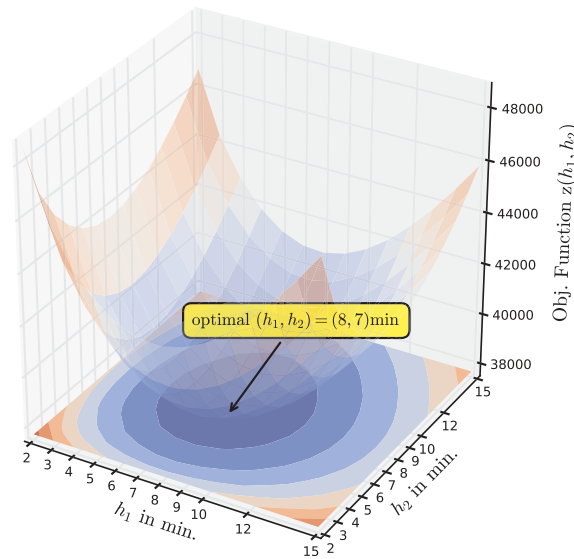


Fig. 11. Objective function values for different dispatching headways without considering the passenger waiting time variability.

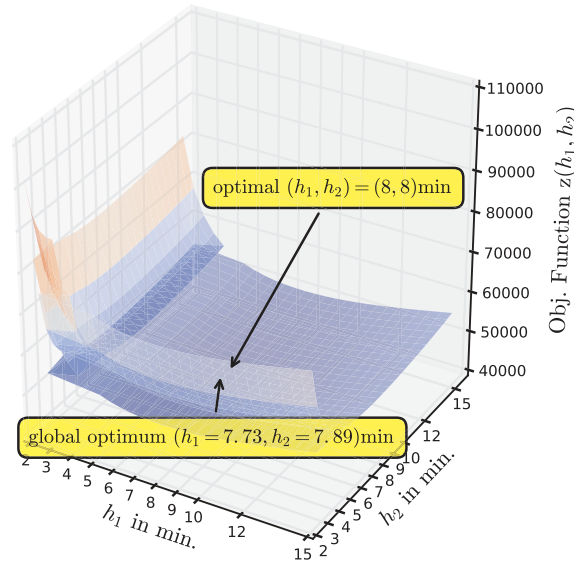


Fig. 12. Objective function values for different dispatching headways considering the passenger waiting time variability.

times in the objective function.

From Fig. 12 one can observe that the optimal dispatching headways ($h_1 = 8; h_2 = 8$ min) differ from the case of not considering operational headway variations. In addition, the objective function value increases significantly only for dispatching headways of 4 min or less. In contrast, this pattern is not observed for dispatching headways longer than 10 min because the increased unsatisfied demand is penalizing the objecting function; but, at the same time, the lower headway variability that is observed at the downstream bus stops of line 1 for such dispatching headway values compensates for this reduction in performance as can be seen in Fig. 6.

In general, the analysis of the historical AVL data in Stockholm resulted in a headway variation pattern that supports higher dispatching headways and this is why the optimal dispatching headway solution that considers the service reliability is not the same as the base-case one. However, this can change from city to city and time period to time period. For this reason, the utilization of the proposed method for setting the dispatching headways of bus lines should be repeated on a regular basis (i.e., as part of seasonal timetable design routines) to adjust to the new patterns of observed headways.

5.4. Network-wide application

For deriving the planned schedules of the $L = \{1, 2, \dots, 17\}$ central bus lines of Fig. 3 we developed a data processing module for

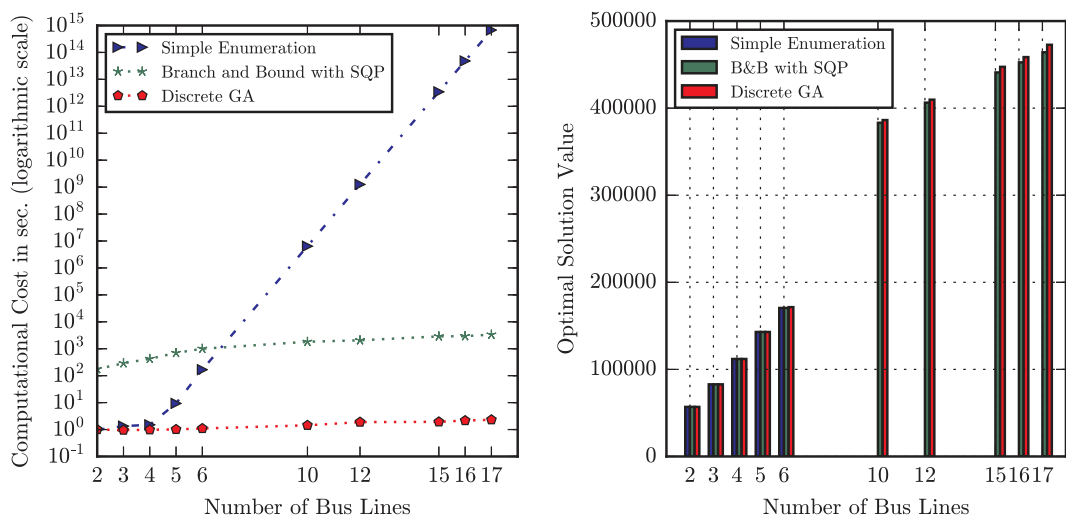
converting GTFS data from.txt formal to sql databases. This facilitates data queries and enables the development of web-based applications providing a front-end to the operational control team.

For the scalability and algorithmic convergence tests, we compared the simple enumeration results against (i) the B&B technique with continuous sub-problem optimization with SQP and (ii) a heuristic Genetic Algorithm (GA) solution method. The B&B and the heuristic GA techniques are programmed in Python 2.7 following the respective algorithmic steps described in this paper. For implementing the SQP algorithm, we use pyOpt which is a Python-based package for formulating and solving nonlinear constrained optimization problems. More specifically, we deploy a multi-start SQP strategy where we generate initial solution guesses and we apply the SQP steps for each one of the solution guesses by using the Feasible Sequential Quadratic Programming (FSQP) method from the pyOpt package which is capable of handling nonlinear inequality constraints and is based on the paper of [Lawrence and Tits \(1996\)](#). The computational performance tests of the above-mentioned algorithms are implemented on a 2556 MHz processor machine with 1024 MB RAM.

The GA solution method is tailored to the characteristics of the dispatching headway setting problem and is comprised of three population members (two “parent” candidate solutions and one “offspring” candidate solution). Initially, the first parent candidate solution is a vector of dispatching headways denoted as $\mathbf{h}^A = \{h_{L_1}^A, \dots, h_{L_{|L|}}^A\}$ where each bus line dispatching headway $\{h_{L_1}^A, \dots, h_{L_{|L|}}^A\}$ is randomly selected from the discrete set $\Omega = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15\} \text{min}$. In a similar manner, the second parent candidate solution is denoted as $\mathbf{h}^B = \{h_{L_1}^B, \dots, h_{L_{|L|}}^B\}$ and the offspring candidate solution is denoted as $\mathbf{h}^O = \{h_{L_1}^O, \dots, h_{L_{|L|}}^O\}$ where initially all dispatching headways take random values from the set Ω .

At each GA iteration it is checked whether the offspring satisfies the constraints of the optimization problem and performs better than its parents by checking if $z(\mathbf{h}^O) < z(\mathbf{h}^A)$ or $z(\mathbf{h}^O) < z(\mathbf{h}^B)$. If this is the case, then the parent solution with the worst performance is replaced with the offspring solution (elitism) by setting $\mathbf{h}^A = \mathbf{h}^O$ if $z(\mathbf{h}^A) > z(\mathbf{h}^B)$ or $\mathbf{h}^B = \mathbf{h}^O$ if $z(\mathbf{h}^B) \geq z(\mathbf{h}^A)$. If this is not the case, the offspring candidate solution is updated during a crossover/mutation phase. At the crossover phase, the dispatching headway of each bus line $h_i^O \in \mathbf{h}^O$ has 50% chance to be replaced by the corresponding dispatching headway of the first parent, h_i^A , and 50% chance to be replaced by the dispatching headway of the second parent, h_i^B . Mutation is also allowed by introducing a 10% probability for each bus line headway $h_i^O \in \mathbf{h}^O$ to take a random value from the set Ω since the dispatching headways of the bus line that do not belong to the two parent candidate solutions might yield a further minimization of the objective function and need to be explored. Finally, the GA terminates if for many iterations the crossover/mutation phases cannot generate a new offspring candidate solution that performs better than its parents. In such case, we assume that we have reached a minimum point and declare the better performing parent solution as the approximate global optimum.

The scalability and algorithmic convergence tests demonstrate the computational complexity of each solution method and their accuracy level (convergence rate to the global optimum). The scalability/convergence tests include larger parts of the central bus network of Stockholm starting from 2 bus lines and moving up to the 17 bus lines shown in [Fig. 3](#). For the simple enumeration method, only results from 6 bus lines were able to be computed due to the computational complexity and memory exhaustion. For instance, optimizing the entire central bus network of Stockholm requires several years if we use simple enumeration while the proposed B&B multi-start strategy returns a solution in 55 min. This computational time attests for its applicability as part of the tactical planning routine. In [Fig. 13a](#) the detailed computational costs of (i) Simple Enumeration, (ii) B&B with a multi-start SQP strategy that generates a large number of initial guesses scattered across the solution space from which each convergence starts and (iii) GA are presented. The figure is constructed based on the results of 10 test scenarios. Each one of these scenarios contains a



(a) Computational Cost with Simple Enumeration and B&B with multi-start SQP strategy

(b) Objective function improvement with Simple Enum., B&B with SQP and discrete GA

Fig. 13. Computational and convergence results from setting the dispatching headways of the central Stockholm bus lines.

different number of bus lines in central Stockholm drawn from the set: {2,3,4,5,6,10,12,15,16,17}. The final scenario with 17 bus lines allocates the desired dispatching headways to all $L = \{1,2,3,4,50,53,55,56,59,61,65,66,69,72,73,74,77\}$ bus lines in central Stockholm. The dispatching headways test cases of {10,12,15,16,17} bus lines or more are computed only with the Branch and Bound and the Discrete GA solution method due to the prohibitive computational cost of Simple Enumeration. Therefore, the computational cost of Simple Enumeration for 10, 12, 15, 16 and 17 bus lines in Fig. 13a is approximated.

From Fig. 13a it is evident that the computational cost of Simple Enumeration increases exponentially with the number of bus lines whereas the GA and the B&B with a multi-start SQP strategy enjoy a polynomial computational cost increase. In addition, Fig. 13b displays the objective function scores and the convergence rates of the optimal dispatching headway solutions attained by Simple Enumeration (for up to 6 bus lines), the proposed B&B method and the discrete GA respectively. It is evident that for up to 5 bus lines all solution methods converge to the global optimum which is also the solution with Simple Enumeration. In the case of 6 bus lines, the GA solution is inferior to the global optimum (convergence rate of 97.89%) while our B&B solution method reaches still a 100% convergence.

For the remaining test-case scenarios of {10,12,15,16,17} bus lines we cannot use simple enumeration to validate whether the B&B solutions or the discrete GA solutions are the global minimizers. Our B&B solution method managed though to compute dispatching headway solutions that improved the objective function score by up to 18% compared with the discrete GA solutions yielding a strong benefit for bus operators that will be able to allocate buses to lines more efficiently.

6. Discussion and concluding remarks

In this work we modeled the network-wide problem of setting optimal dispatching headways that determine the frequencies of bus services with the explicit consideration of the inherent variability to bus operations (demand, travel time and passenger waiting times variations). In addition, we introduced a B&B method with a multi-start strategy for generating initial guesses and using SQP in order to converge to the global optimum solution of the discrete, nonlinear problem of setting the dispatching headways. Results from a real-size network demonstrate that our solution method converged to the global optimum computed with brute-force on small-sized bus networks. Moreover, the proposed method returns superior solutions by up to 18% when compared to heuristic solution methods such as GA for larger-scale scenarios.

Using historical AVL and APC data from two bi-directional bus lines in central Stockholm, the optimal bus dispatching headways based on several aspects (passenger demand coverage, waiting time variability at stop level, operational costs, cost of utilizing extra buses) were examined by investigating the impact of the corresponding aspect to the objective function.

The problem formulation and solution method presented in this paper can be used for tactical planning of dispatching headways that considers explicitly the variabilities during bus operations. The sensitivity analysis of the determined dispatching headways that considers reliability aspects allowed also to identify the range of values under which the specification of weight values does not affect the solution obtained, allowing the service operator to select solutions that are robust to differences in priority setting. The method proposed in this paper assumes that AVL data from all bus trips at the network level is available for deriving the travel time and, subsequently, the headway patterns at downstream stops for different values of dispatching headways. Our problem formulation assumes that there is no interlining, i.e. the dispatching headway for each line is determined separately, assuming that vehicles run back and forth on the same route. By providing information on deadheading possibilities, future research can enhance the fleet allocation flexibility with interlining; something that is especially advantageous in case of a strongly directional (i.e. asymmetric) demand (Ceder and Stern, 1981). The implications of the planned dispatching headways on operations and possible interactions with real-time control strategies can be potentially studied using a transit simulation model that is capable of accounting for day to day variations in waiting time uncertainty (Cats and Gkioulou, 2017).

Regarding the practicality of the proposed method that determines the dispatching headways, we note that our approach utilizes the existing resources (available vehicles) without (i) requesting extra resources for improving the passenger-related costs or (ii) deploying service or rolling stock variants such as short-turning and interlining services. Given the computational efficiency of the proposed approach, it is considered feasible to apply the proposed solution approach as part of the transit planning. The results suggest that the solution obtained can lead to substantial savings compared with existing approaches, although the extent to which these savings will materialize depends on the exact vehicle and crew scheduling.

The model presented in this paper can be further extended in future research by considering additional constraints such as bus drivers' rostering and by including the associated labor in the objective function. Determining the optimal daily times when the dispatching headways should change values as function of temporal variations in travel time and demand patterns can be also a promising subject for future research.

Appendix A. Detailed description of the SQP method

SQP generates new iterates of the *initial guess* dispatching headways, $\mathbf{h}_{k=0}$, by solving inequality constraint Quadratic sub-problems (QP) at each iterate k . The idea behind the SQP solution method is to model the dispatching headways of the current iterate, \mathbf{h}_k , by a quadratic programming QP subproblem and then use the minimizer of this sub-problem to define a new iterate of dispatching headways, \mathbf{h}_{k+1} , until convergence.

Given that the boundary constraints are inequality constraints and both the objective function and the constraints, z, c_i , are continuously differentiable at a point \mathbf{h}_k , then if \mathbf{h}_k is a local optimum and the *regularity conditions* are satisfied at this point there is a Lagrange multiplier vector λ_k with m elements such that the *first order necessary Karush-Kuhn-Tucker* (KKT) conditions are satisfied:

$$\begin{aligned}
&\text{Stationarity } \nabla_{\mathbf{h}} \mathcal{L}(\mathbf{h}_k, \lambda_k) = 0 \\
&\text{Primal Feasibility: } c_i(\mathbf{h}_k) = 0, \forall i \in E = \emptyset \text{ and } c_i(\mathbf{h}_k) \geq 0, \forall i \in I = \{1, 2, 3, \dots, 3|L| + 2\} \\
&\text{Dual Feasibility: } \lambda_{k,i} \geq 0, \forall i \in I \\
&\text{Complementarity: } \lambda_{k,i} c_i(\mathbf{h}_k) = 0, \forall i \in E \cup I
\end{aligned} \tag{A.1}$$

where

$$\mathcal{L}(\mathbf{h}, \lambda) = z(\mathbf{h}) - \sum_{i \in E \cup I} \lambda_i c_i(\mathbf{h}) \tag{A.2}$$

is the Lagrangian function $\mathcal{L}: \mathbb{R}^{L+m} \rightarrow \mathbb{R}$ of the constrained NLP of setting the dispatching headways and, at the initial iteration, an *initial guess* of the Lagrange multipliers $\lambda_{k=0}$ should be also provided.

To model the current dispatching headway solution \mathbf{h}_k by a quadratic programming (QP) subproblem and then use the minimizer of this subproblem to define a new iterate \mathbf{h}_{k+1} until convergence, a linearization of the constraints is required since QP problems tackle only linear constraints. This can be modeled by using the current iteration values of the dispatching headways vector \mathbf{h}_k and the Lagrange multipliers λ_k for finding the minimizer p which is a vector of $|L|$ elements by solving the following QP subproblem:

$$\begin{aligned}
&\min_{p \in \mathbb{R}^L} \quad z(\mathbf{h}_k) + \nabla z(\mathbf{h}_k)^T p + \frac{1}{2} p^T \nabla_{hh}^2 \mathcal{L}(\mathbf{h}_k, \lambda_k) p \\
&\text{subject to} \quad \nabla c_i(\mathbf{h}_k)^T p + c_i(\mathbf{h}_k) \geq 0, i \in I \\
&\quad \quad \quad \nabla c_i(\mathbf{h}_k)^T p + c_i(\mathbf{h}_k) = 0, i \in E = \emptyset
\end{aligned} \tag{A.3}$$

where $J(\mathbf{h})^T = [\nabla c_1(\mathbf{h}), \nabla c_2(\mathbf{h}), \dots, \nabla c_m(\mathbf{h})]$ is the *Jacobian* matrix of the constraints vector with $|L| \times m$ elements and $\nabla_{hh}^2 \mathcal{L}(\mathbf{h}_k, \lambda_k)$ is the *Hessian* matrix of the Lagrange function with $|L| \times |L|$ elements. After solving the above inequality QP problem the iterate values are updated $(\mathbf{h}_{k+1}, \lambda_{k+1}) = (\mathbf{h}_k + p_k, \lambda_{k+1})$ where p_k and λ_{k+1} are the solution and the corresponding Lagrange multiplier of the inequality QP. Iterations then continue until convergence according to the algorithm:

```

1: function SQP( $z(\mathbf{h}), c_1(\mathbf{h}), \dots, c_m(\mathbf{h}), E, I$ )
2:   With an initial guess of dispatching headways and Lagrange multipliers choose an initial iteration pair  $(\mathbf{h}_k, \lambda_k)$ ;
3:   Set  $k \leftarrow 0$ ;
4:   while a convergence test is satisfied
5:     Compute  $z(\mathbf{h}_k), \nabla z(\mathbf{h}_k)^T, \nabla_{hh}^2 \mathcal{L}(\mathbf{h}_k, \lambda_k), J(\mathbf{h}_k)^T, c(\mathbf{h}_k)$ , where  $c$  is vector of  $m$  elements;
6:     Solve the Inequality QP( $z(\mathbf{h}), \nabla_{hh}^2 \mathcal{L}(\mathbf{h}_k, \lambda_k), \nabla z(\mathbf{h}_k)^T, E, I, J(\mathbf{h}_k)^T, -c(\mathbf{h}_k)$ ) sub-problem:
7:        $\min_{p \in \mathbb{R}^L} z(\mathbf{h}_k) + \nabla z(\mathbf{h}_k)^T p + \frac{1}{2} p^T \nabla_{hh}^2 \mathcal{L}(\mathbf{h}_k, \lambda_k) p$ 
8:       subject to  $J(\mathbf{h}_k)p + c(\mathbf{h}_k) \geq 0$ 
9:       Inequality QP function returns solution pair  $(p_k, \lambda_{k+1})$ 
10:      Set  $(\mathbf{h}_{k+1}, \lambda_{k+1}) \leftarrow (\mathbf{h}_k + p_k, \lambda_{k+1})$ 
11:    end while
12:    return optimal solution  $\mathbf{h}^*$  and the objective function value  $z(\mathbf{h}^*)$ ;
13: end function

```

The Hessian, Jacobian and other first and second order derivatives can be numerically approximated with the use of finite differences in an adaptive manner, coupled with Richardson's extrapolation methodology (Richardson, 1911) to provide a maximally accurate result. The SQP algorithm convergence criterion can be the step direction stagnation (i.e., reach at an inequality QP subproblem where its solution returns $p_k = \{0, \dots, 0\}$ which indicates that there is no better direction than the current one).

References

- Alexander, L., Jiang, S., Murga, M., González, M.C., 2015. Origin-destination trips by purpose and time of day inferred from mobile phone data. *Transport. Res. Part C: Emerg. Technol.* 58, 240–250.
- Amberg, B., Ionescu, L., Kliever, N., 2017. Robust efficiency in public bus transport and airline resource scheduling. In: *Operations Research Proceedings 2015*. Springer, pp. 259–264.
- Asgharzadeh, M., Shafahi, Y., 2017. Real-time bus-holding control strategy to reduce passenger waiting time. *Transport. Res. Rec.: J. Trans. Res. Board* (2647), 9–16.
- Barabino, B., Di Francesco, M., Mozzoni, S., 2017. An offline framework for the diagnosis of time reliability by automatic vehicle location data. *IEEE Trans. Intell. Transport. Syst.* 18(3), 583–594.
- Barra, A., Carvalho, L., Teypaz, N., Cung, V.-D., Balassiano, R., 2007. Solving the transit network design problem with constraint programming. In: *11th World Conference in Transport Research-WCTR 2007*.
- Bellei, G., Gkoulas, K., 2010. Transit vehicles' headway distribution and service irregularity. *Public Transport* 2 (4), 269–289.
- Boggs, P.T., Tolle, J.W., 1995. Sequential quadratic programming. *Acta Numer.* 4, 1–51.
- Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira, J., Ratti, C., 2013. Understanding individual mobility patterns from urban sensing data: a mobile phone trace example. *Transport. Res. Part C: Emerg. Technol.* 26, 301–313.
- Cats, O., Gkioulou, Z., 2017. Modeling the impacts of public transport reliability and travel information on passengers waiting-time uncertainty. *EURO J. Transport. Logist.* 6 (3), 247–270.
- Ceder, A., 1984. Bus frequency determination using passenger count data. *Transport. Res. Part A: Gen.* 18 (5–6), 439–453.
- Ceder, A., 2007. *Public Transit Planning and Operation: Theory, Modeling and Practice*. Elsevier, Oxford.

- Ceder, A., Israeli, Y., 1998. User and operator perspectives in transit network design. *Transport. Res. Rec.: J. Transport. Res. Board* (1623), 3–7.
- Ceder, A., Stern, H.L., 1981. Deficit function bus scheduling with deadheading trip insertions for fleet size reduction. *Transport. Sci.* 15 (4), 338–363.
- Chen, X., Yu, L., Zhang, Y., Guo, J., 2009. Analyzing urban bus service reliability at the stop, route, and network levels. *Transport. Res. Part A: Policy Pract.* 43 (8), 722–734.
- Cipriani, E., Gori, S., Petrelli, M., 2012. Transit network design: a procedure and an application to a large urban area. *Transport. Res. Part C: Emerg. Technol.* 20 (1), 3–14.
- Cortés, C.E., Gibson, J., Gschwendner, A., Munizaga, M., Zúñiga, M., 2011. Commercial bus speed diagnosis based on gps-monitored data. *Transport. Res. Part C: Emerg. Technol.* 19 (4), 695–707.
- de Regt, K., Cats, O., van Oort, N., van Lint, H., 2017. Investigating potential transit ridership by fusing smartcard data and gsm data. *Transport. Res. Rec.* 2652 (2017), 50–58.
- Fan, W., Machemehl, R.B., 2008. Tabu search strategies for the public transportation network optimizations with variable transit demand. *Comput.-Aid. Civil Infrastruct. Eng.* 23 (7), 502–520.
- Farahani, R.Z., Miandoabchi, E., Szeto, W.Y., Rashidi, H., 2013. A review of urban transportation network design problems. *Euro. J. Oper. Res.* 229 (2), 281–302.
- Furth, P.G., Wilson, N.H., 1981. Setting frequencies on bus routes: theory and practice. *Transport. Res. Rec.* 818 (1981), 1–7.
- Gkiotsalitis, K., Cats, O., 2017. Optimization of bus frequency settings considering demand and trip time variations. In: 96th Annual Meeting of the Transportation Research Board.
- Gkiotsalitis, K., Maslekar, N., 2015. Improving bus service reliability with stochastic optimization. In: 2015 IEEE 18th International Conference on Intelligent Transportation Systems (ITSC). IEEE, pp. 2794–2799.
- Gkiotsalitis, K., Stathopoulos, A., 2015. A utility-maximization model for retrieving users willingness to travel for participating in activities from big-data. *Transport. Res. Part C: Emerg. Technol.* 58, 265–277.
- Gkiotsalitis, K., Wu, Z., Cats, O., 2017. Tactical service design and vehicle allocation optimization. In: hEART 2017: 6th Symposium of the European Association for Research in Transportation.
- Guihaire, V., Hao, J.-K., 2008. Transit network design and scheduling: a global review. *Transport. Res. Part A: Policy Pract.* 42 (10), 1251–1273.
- Hadas, Y., Shnaiderman, M., 2012. Public-transit frequency setting using minimum-cost approach with stochastic demand and travel time. *Transport. Res. Part B: Methodol.* 46 (8), 1068–1084.
- Hadas, Y., Shnaiderman, M., Cedar, A., 2010. Public transit frequency setting using minimum cost approach with stochastic demand. In: Proceedings of the 45th Annual Conference of the ORSNZ, Auckland, New Zealand. pp. 353–362.
- Ibarra-Rojas, O., Delgado, F., Giesen, R., Muñoz, J., 2015. Planning, operation, and control of bus transport systems: a literature review. *Transport. Res. Part B: Methodol.* 77, 38–75.
- Kepaptsoglou, K., Karlaftis, M., 2009. Transit route network design problem: review. *J. Transport. Eng.* 135 (8), 491–505.
- Lawrence, C.T., Tits, A.L., 1996. Nonlinear equality constraints in feasible sequential quadratic programming. *Optim. Meth. Software* 6 (4), 265–282.
- Lee, Y.-J., Vuchic, V.R., 2005. Transit network design with variable demand. *J. Transport. Eng.* 131 (1), 1–10.
- Li, Y., Xu, W., He, S., 2013. Expected value model for optimizing the multiple bus headways. *Appl. Math. Comput.* 219 (11), 5849–5861.
- Lta: Ewt Incentive Structure, 2017. < [https://www.lta.gov.sg/data/apps/news/press/2014/20140124_BSRF\(final2\)-Annex.pdf](https://www.lta.gov.sg/data/apps/news/press/2014/20140124_BSRF(final2)-Annex.pdf) > (Accessed: 2017-07-27).
- López Ramos, F., 2014. Conjoint Design of Railway Lines and Frequency Setting Under Semi-Congested Scenarios. PhD Thesis — Universitat Politècnica de Catalunya.
- Luo, D., Cats, O., van Lint, H., 2017. Constructing transit origin-destination matrices using spatial clustering. *Transport. Res. Rec.* 2652 (2017), 39–49.
- Ma, X., Wu, Y.-J., Wang, Y., Chen, F., Liu, J., 2013. Mining smart card data for transit riders travel patterns. *Transport. Res. Part C: Emerg. Technol.* 36, 1–12.
- Moreira-Matias, L., Cats, O., Gama, J., Mendes-Moreira, J., de Sousa, J.F., 2016. An online learning approach to eliminate bus bunching in real-time. *Appl. Soft Comput.* 47, 460–482.
- Munizaga, M.A., Palma, C., 2012. Estimation of a disaggregate multimodal public transport origin-destination matrix from passive smartcard data from Santiago, Chile. *Transport. Res. Part C: Emerg. Technol.* 24, 9–18.
- Murty, K.G., Yu, F.-T., 1988. *Linear Complementarity, Linear and Nonlinear Programming*, vol. 3. Heldermann Berlin.
- Nocedal, J., Wright, S.J., 2006. *Numerical Optimization*. Springer.
- Osuna, E., Newell, G., 1972. Control strategies for an idealized public transportation system. *Transport. Sci.* 6 (1), 52–72.
- Paulley, N., Balcombe, R., Mackett, R., Titheridge, H., Preston, J., Wardman, M., Shires, J., White, P., 2006. The demand for public transport: the effects of fares, quality of service, income and car ownership. *Transport Policy* 13 (4), 295–306.
- Pelletier, M.-P., Trépanier, M., Morency, C., 2011. Smart card data use in public transit: a literature review. *Transport. Res. Part C: Emerg. Technol.* 19 (4), 557–568.
- Preston, J., James, T., 2000. Analysis of Demand for Bus Services. Final report. Transport Studies Unit, University of Oxford.
- Richardson, L.F., 1911. The approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a masonry dam. *Philos. Trans. Roy. Soc. Lond. Ser. A, Contain. Papers Math. Phys. Character* 210, 307–357.
- Silman, L.A., Barzily, Z., Passy, U., 1974. Planning the route system for urban buses. *Comput. Oper. Res.* 1 (2), 201–211.
- Szeto, W.Y., Wu, Y., 2011. A simultaneous bus route design and frequency setting problem for tin shui wai, hong kong. *Euro. J. Oper. Res.* 209 (2), 141–155.
- Trafikverket, 2015. samhällsekonomiska principer och kalkylvrden fr transport sek- torn: Asek5.2, version2015-04-01, borlnge.
- Verbas, I., Mahmassani, H., 2013. Optimal allocation of service frequencies over transit network routes and time periods: formulation, solution, and implementation using bus route patterns. *Transport. Res. Rec.: J. Transport. Res. Board* (2334), 50–59.
- Verbas, I.Ö., Frei, C., Mahmassani, H.S., Chan, R., 2015. Stretching resources: sensitivity of optimal bus frequency allocation to stop-level demand elasticities. *Public Transport* 7 (1), 1–20.
- Wardman, M., 2004. Public transport values of time. *Transport Policy* 11 (4), 363–377.