

Energy-optimal Timetable Design for Sustainable Metro Railway Networks

Shuvomoy Das Gupta

Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA, USA.

ORCID: 0000-0001-9650-8746

sdgupta@mit.edu, <https://shuvomoy.github.io/>

Bart P.G. Van Parys

Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, USA.

vanparys@mit.edu, <https://www.vanparys.xyz/>

J. Kevin Tobin

Thales Canada Inc, 105 Moatfield Drive, Toronto, Ontario, Canada

kevin.tobin@urbanandmainlines.com, <https://ca.linkedin.com/in/kevin-tobin-45017b47>

Problem definition: We present our collaboration with Thales Canada Inc, the largest provider of communication-based train control (CBTC) systems worldwide. We study the problem of designing energy-optimal timetables in metro railway networks to minimize the *effective energy consumption* of the network, which corresponds to simultaneously minimizing total energy consumed by all the trains and maximizing the transfer of regenerative braking energy from suitable braking trains to accelerating trains.

Methodology/Results: We propose a novel data-driven linear programming model that minimizes the total effective energy consumption in a metro railway network, capable of computing the optimal timetable in real-time, even for some of the largest CBTC systems in the world. In contrast with existing works, which are either NP-hard or involve multiple stages requiring extensive simulation, our model is a single linear programming model capable of computing the energy-optimal timetable subject to the constraints present in the railway network. Furthermore, our model can predict the total energy consumption of the network without requiring time-consuming simulations, making it suitable for widespread use in managerial settings. We apply our model to Shanghai Railway Network’s Metro Line 8—one of the largest and busiest railway services in the world—and empirically demonstrate that our model computes energy-optimal timetables for thousands of active trains spanning an entire service period of one day in real-time (solution time less than one second on a standard desktop), achieving energy savings between approximately 20.93% and 28.68%.

Managerial implications: Our data-driven linear programming model’s robust, real-time capabilities offer transformative implications for the managerial decision-making process in metro railway networks, particularly those utilizing CBTC systems. The model’s computational efficiency negates the need for time-intensive simulations traditionally used for timetable optimization, thereby streamlining the planning process. It can also be generalized and deployed across any CBTC system, making it a broadly applicable tool for sustainable timetable planning. Given these compelling advantages, our model is in the process of being integrated into Thales Canada Inc’s industrial timetable compiler.

Key words: linear optimization; energy analytics; sustainable transportation

1. Introduction

Background. As projections indicate that the global urban population is poised to increase from 54% to 66% by the year 2050 (United Nations 2014), the urgency to enhance both the efficiency and sustainability of urban transit systems is escalating. Faced with an estimated influx of 2.5 billion urban inhabitants, there is an impending imperative to restructure existing transportation infrastructure. Rail transit, lauded for its substantial passenger-carrying capacity and safety metrics (Abril et al. 2008), is strategically positioned to play a crucial role in this infrastructural transformation. Notably, Communication-Based Train Control Systems (CBTC) have a significant impact on the effective functioning of these rail networks (Pascoe and Eichorn 2009). Utilized by some of the world’s busiest metros, CBTC systems offer high-resolution train location determination and continuous data communications. These capabilities significantly improve both safety and efficiency and allow for dynamic adjustments in operational schedules, offering solutions to traffic congestion and fluctuating demand. Thanks to technological advancements like built-in redundancy, modern CBTC systems are also more reliable and easier to maintain. CBTC systems are particularly conducive to developing sustainable and energy-efficient operations, owing to functionalities like the ability to implement automatic driving strategies and the precise implementation of a provided railway timetable (IRSE International Technical Committee 2009). In light of these multi-dimensional advantages, the integration of real-time, energy-efficient timetable optimization within CBTC-enabled railway networks emerges as a cornerstone strategy. This strategic integration not only amplifies operational efficiency and adaptability but also contributes to achieving broader environmental sustainability goals. Hence, it aligns perfectly with the imperatives of modern urban transport, which must be both highly efficient and demonstrably sustainable to meet the complex demands of an increasingly urbanized global landscape.

Motivation. Electricity is the energy source for trains in all modern metro railway networks that utilize CBTC systems. When trains make trips from origin platforms to destination platforms, their speed profile consists of four phases: acceleration, speed holding, coasting, and braking (Howlett and Pudney 1995), as shown qualitatively in Figure 1. Among these phases, trains consume most of their energy during acceleration. Conversely, when trains brake, the electric motors that make the trains move during the acceleration phase work in reverse, becoming generators. They convert the trains’ kinetic energy into electrical energy, referred to as *regenerative braking energy*. CBTC systems are capable of implementing a given timetable precisely, which allows for the transfer of this regenerative energy from braking trains to nearby accelerating trains. For empirical context, the New York City transit railway system, which consumes more than 1,600 GWh of electricity per year, has had all of its trains installed since 2018 capable of producing and transferring regenerative braking energy. Under favorable conditions, the regenerative energy produced can account for up to 50% of the energy consumed (Mohamed et al. 2018).

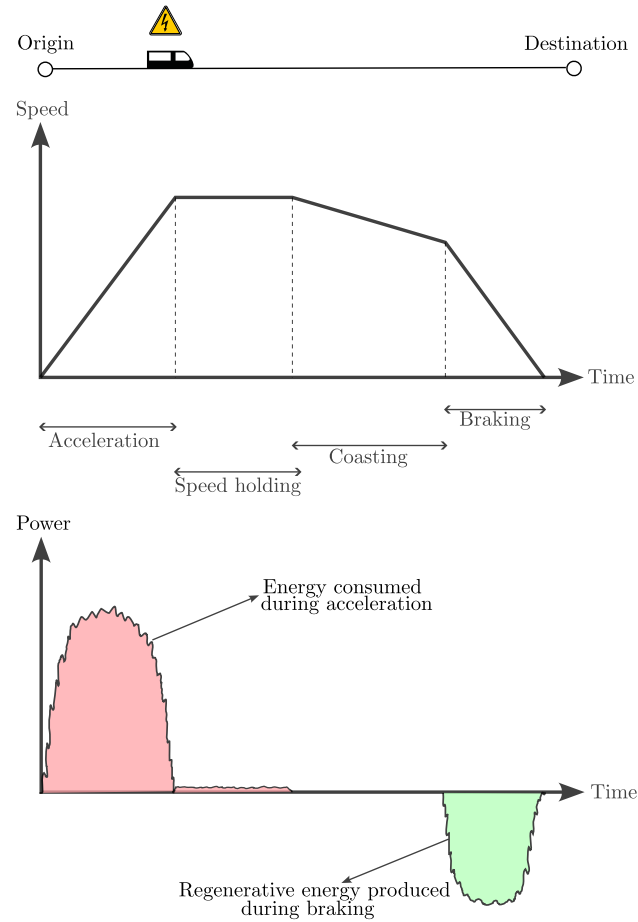


Figure 1 This figure graphically illustrates a train's energy consumption and regeneration cycle in CBTC systems. When a train makes a trip from the origin platform to the destination platform, it transitions through four phases - acceleration, speed holding, coasting, and braking, as shown in the speed vs. time graph at the top. The acceleration phase requires the most amount of energy energy, as denoted by the red-shaded area in the power vs. time graph at the bottom. In contrast, the speed holding and coasting phases entail minimal to zero energy usage. Notably, during the braking phase, the train produces regenerative braking energy, represented by the green-shaded area in the power vs. time graph. With appropriate scheduling, this energy can be strategically transferred to nearby accelerating trains.

Energy-optimal timetable. The conceptualization and implementation of an energy-optimal timetable offer avenues for substantial energy conservation, obviating the need for infrastructural alterations within CBTC-enabled railway systems. Formally, a railway timetable is a data structure, containing both the arrival and departure times of each train at every platform visited during a designated service period—typically spanning 18 to 24 hours in most operational networks. The core objective of an energy-optimal timetable is to strategically design these temporal decision variables to minimize the network's *effective energy consumption*, which equals the *total energy consumption for all trains during their accelerating phases* minus the *total regenerative braking energy successfully transferred to the accelerating trains from eligible braking counterparts*.

Related work. Recent years have seen noteworthy contributions to energy-efficient railway timetable computation. For instance, (Peña-Alcaraz et al. 2012) proposes a mixed integer programming (MIP) model limited to single train-lines with successful application on Madrid’s Line 3. The work in (Das Gupta et al. 2015) presents a more tractable MIP model for optimizing regenerative energy transfer between suitable train pairs, applied to the Dockland Light Railway, but does not directly address the actual energy savings. The work in (Li and Lo 2014) employs a genetic algorithm to calculate energy-efficient timetables. Their approach seeks to maximize the use of regenerative energy while minimizing the tractive energy of trains. Similarly, (Le et al. 2014) introduces a nonlinear integer programming model that utilizes simulated annealing. The model (Yang et al. 2020) proposes a non-dominated sorting genetic algorithm for a biobjective timetable optimization model, uniquely incorporating energy allocation and passenger assignment. Continuing in the same vein, Wang and Goverde have conducted studies on multi-train trajectory optimization: their method treats the problem as a multiple-phase optimal control problem, solved by a pseudospectral method (Wang and Goverde 2017). Furthermore, they introduce an innovative approach to energy-efficient timetabling that adjusts the running time allocation of given timetables using train trajectory optimization (Wang and Goverde 2019).

From an industry implementation standpoint, the two-stage linear optimization model (Das Gupta et al. 2016) is, to our knowledge, the sole optimization model incorporated into an industrial timetable compiler. This model presents a two-step linear optimization model for developing an energy-efficient railway timetable. The first stage minimizes total energy consumption for all trains, considering the railway network’s constraints. The second stage fixes the trains’ trip times to those computed in the first stage and maximizes the transfer of regenerative braking energy between suitable train pairs. However, this model has some limitations. First, although the optimization problems in the first and second stages are solved quickly (in minutes), transitioning between stages requires a time-consuming (in hours) intermediate simulation process, significantly extending the end-to-end runtime. Second, the model in (Das Gupta et al. 2016) does not directly model the network’s actual energy consumption, rather it employs proxy objective functions, the optimization of which might lead to an energy-efficient timetable indirectly. This roundabout approach necessitates extensive physics-based simulations a posteriori to predict the actual reduction in energy consumption. This could limit its applicability in managerial settings, where the model’s perceived utility is directly tied to its ability to predict energy consumption levels. Finally, the model requires the precise locations of energy consumption peaks, which are often unavailable, to formulate the optimization problem for the second stage.

Contribution. We develop a novel single-stage linear optimization model to construct energy-optimal timetables for all the trains in a CBTC-enabled metro railway network for an entire service period. We directly model the total energy consumption, total savings of regenerative energy, and the interaction between regenerative and consumed energy using a data-driven approach. Unlike existing works, our model can predict total energy consumption without requiring time-consuming simulations, making it suitable for widespread use in managerial settings. We also model the transfer of regenerative energy through a set of linear constraints, whereas the existing works use nonconvex constraints. We empirically demonstrate that our model performs extremely well when applied to Metro Line 8 of the Shanghai Railway network, which is one of the busiest railway services in the world in terms of ridership and number of trains. We deploy a warm-started parallel barrier algorithm to solve our linear optimization model, which computes energy-optimal timetables for a full service period of one day with thousands of active trains in real-time (solution time is less than a second on a standard desktop). The optimal timetables exhibit a significant reduction in effective energy consumption in comparison with existing real-world timetables, ranging between approximately 20.93% to 28.68%. Our proposed model offers transformative benefits for managerial decision-making within metro railway networks employing CBTC systems. It not only significantly speeds up the planning process by sidestepping traditional, time-intensive simulations but is also versatile enough for generalized application across any CBTC system. Recognizing these advantages, our model is in the process of being implemented into Thales Canada Inc’s industrial timetable compiler.

1.1. Organization

The paper is organized as follows. In Section 2, we present the notation and notions used in this paper. Then in Section 3, we discuss the constraints that are necessary for proper functioning of a metro railway network. In Section 4, we present how we model the objective of effective energy consumption. The final optimization model is presented in Section 5. We present our numerical experiments applied to the Shanghai Railway Network in Section 6. We describe the architectural framework for industrial integration of our optimization model in Section 7.

2. Notation and notions

All the sets described in this paper are strictly ordered and finite unless otherwise specified. The cardinality and the i -th element of such a set S are denoted by $|S|$ and $S(i)$, respectively.

Table 1: List of symbols in the order they appear in the paper.

Symbol	Description
\mathcal{N}	The set of all platforms in a railway network

\mathcal{A}	The set of all tracks
\mathcal{T}	The set of all trains
\mathcal{N}^t	The set of all platforms visited by a train t in chronological order
\mathcal{A}^t	The set of all tracks visited by a train t in chronological order
a_i^t	The arrival time of train t at platform i (decision variable)
d_i^t	The departure time of train t from platform i (decision variable)
$[\tau_{ij}^t, \bar{\tau}_{ij}^t]$	The tack-based trip time window for train t from platform i to platform j
φ	The set of all crossing-overs in the network
\mathcal{B}_{ij}	The set of all train pairs involved in turn-around events on crossing-over (i, j)
$[\kappa_{ij}^{tt'}, \bar{\kappa}_{ij}^{tt'}]$	The trip time window for train t on the crossing-over (i, j)
$[\delta_i^t, \bar{\delta}_i^t]$	The dwell time window for train t at platform i
χ	The set of all platform pairs situated at the same interchange stations
\mathcal{C}_{ij}	The set of connecting train pairs for a platform pair $(i, j) \in \chi$
$[\chi_{ij}^{tt'}, \bar{\chi}_{ij}^{tt'}]$	The connection window between train t at platform i and train t' at platform j
\mathcal{H}_{ij}	The set of train-pairs who move along that track (i, j)
$h_i^{tt'}$	The headway time between train t and t' at or from platform i
$[\tau_{\mathcal{P}}^t, \bar{\tau}_{\mathcal{P}}^t]$	The total travel time window for train t to traverse its train path
$E_{i,j,t}^{\text{con,tr}}$	Energy consumed by train t during acceleration while going from platform i to platform j
$E_{i,j,t,t'}^{\text{con,cr}}$	Energy consumed by train t during acceleration while traversing the crossing-over $(i, j) \in \varphi$
Ω	The set of all platform pairs that are feasible for regenerative energy transfer
\mathcal{T}_i	The set of all trains which arrive at, dwell and then depart from platform i
\vec{t}	Temporally close train to the right of train t
\overleftarrow{t}	Temporally close train to the left of train t
\tilde{t}	Temporally closes train to train t
\mathcal{E}	The set of all synchronization processes between suitable train pairs
$\vec{\mathcal{E}}$	A subset of \mathcal{E} containing elements of the form (i, j, t, \vec{t})
$\overleftarrow{\mathcal{E}}$	A subset of \mathcal{E} containing elements of the form $(i, j, t, \overleftarrow{t})$
$E_{i,j,t,\vec{t}}^{\text{reg}}$	Regenerative energy transferred to accelerating train t on platform i from braking train \vec{t} on platform j
$\sigma_{ij}^{t,\vec{t}}$	Overlapping time between the effective accelerating phase of train t on platform i and the effective braking phase of train \vec{t} on platform j
$E_{i,j,t,\overleftarrow{t}}^{\text{reg}}$	Regenerative energy transferred to accelerating train \overleftarrow{t} on platform j from braking train t on platform i
$\sigma_{ij}^{t,\overleftarrow{t}}$	Overlapping time between the effective accelerating phase of train \overleftarrow{t} on platform j and the effective braking phase of train t on platform i
$[\beta_i^t, \bar{\beta}_i^t]$	The duration of the effective braking phase of train t around platform i
$[\alpha_i^t, \bar{\alpha}_i^t]$	The duration of the effective accelerating phase of train t around platform i

Consider a railway network where the set of all stations is denoted by \mathcal{S} . The set of all platforms in the railway network is indicated by \mathcal{N} . A directed arc between two distinct and non-opposite platforms is called a *track*. The set of all tracks is represented by \mathcal{A} .

The directed graph of the railway network is expressed by $\mathcal{G} = (\mathcal{N}, \mathcal{A})$. A *train-line* or *line* is a directed path, where the set of nodes represents non-opposite platforms and the set of arcs represents non-opposite tracks. A *crossing-over* is a special type of directed arc that connects two train-lines. If a train arrives at the terminal platform of a train-line, turns around by traversing the crossing-over, and starts traveling through another train-line, then the same physical train is

treated and labeled functionally as two different trains by the railway management. The set of all trains to be considered in our problem is denoted by \mathcal{T} . The sets of all platforms and all tracks visited by a train t in chronological order are denoted by $\mathcal{N}^t \subseteq \mathcal{N}$ and $\mathcal{A}^t \subseteq \mathcal{A}$ respectively. The *train-path* of a train is the directed path that contains all platforms and tracks visited by it in chronological order.

The decision variables to be determined are the *arrival times* and *departure times* of trains to and from the associated platforms respectively. Let a_i^t be the arrival time of train $t \in \mathcal{T}$ at platform $i \in \mathcal{N}^t$ and d_j^t be the departure time of train t from platform $j \in \mathcal{N}^t$.

3. Modelling the functional constraints

The functional constraints in the railway network show how the events are related and are necessary for the proper functioning of the railway network, which we present in Sections 3.1-3.5 below. We first discuss how we protect against uncertainty associated with the constraint data in the railway network as follows.

Incorporating robust constraints through box uncertainty. In practice, there can be some uncertainty in the data associated with the functional constraints, which can cause them to diverge from their intended values. To guard against such uncertainties, we integrate box uncertainty constraints into our model. Abstractly, all the functional constraints can be represented as $\ell \leq x - y \leq u$, where x and y stand for decision variables, and ℓ and u refer to the problem data. We then define the following uncertainty set: $\ell \in [\ell_{lb}, \ell_{ub}]$ and $u \in [u_{lb}, u_{ub}]$. The robust counterpart of the constraint $\ell \leq x - y \leq u$, for $\ell \in [\ell_{lb}, \ell_{ub}]$, $u \in [u_{lb}, u_{ub}]$ can be derived through straightforward computation as: $\ell_{ub} \leq x - y \leq u_{lb}$. In the interest of brevity, we shall present the robust versions of our constraints directly in the representations that follow. Now we are in a position to present the robustified railway constraints. While from an abstract mathematical point of view, all the constraints have a similar form, from a practical point of view, the constraints are very different. A pictorial representation of what the constraints physically represent is given by Figure 2.

3.1. Trip time constraint

The trip time constraints are crucial in determining train energy consumption and regenerative energy production. As discussed previously, while a train is making a trip from the origin platform to the destination platform, almost all the its required energy is consumed during the acceleration phase of its trip and all of the regenerative braking energy is produced during the braking phase. Trip time constraints can be classified into two types described as follows.

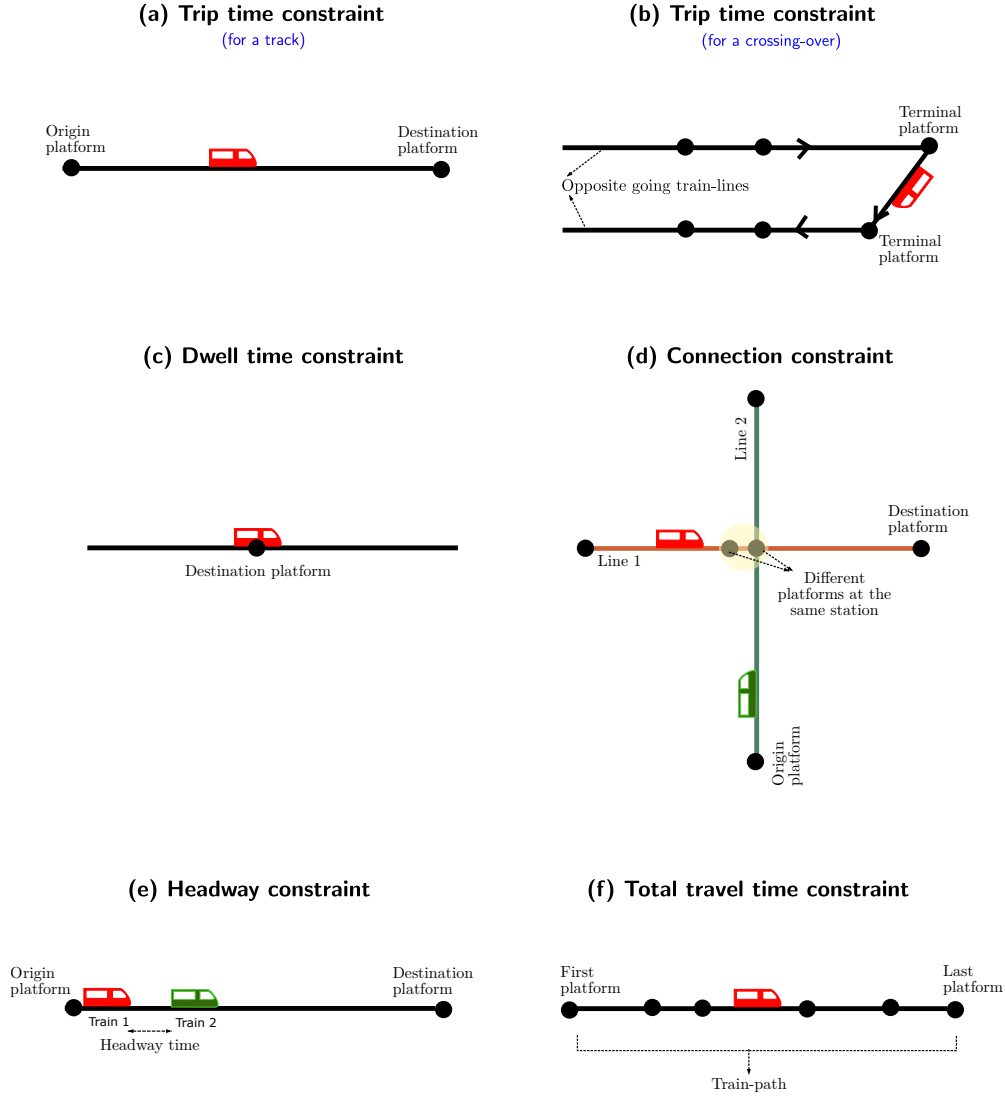


Figure 2 This figure graphically illustrates all the functional constraints present in a metro railway network that the timetable has to satisfy.

3.1.1. Track-based trip time constraint. Let us consider the trip of any train $t \in \mathcal{T}$ from platform i to platform j along the track $(i, j) \in \mathcal{A}^t$. The train departs from platform i at time d_i^t , arrives at platform j at time a_j^t , and its trip time can range from $\underline{\tau}_{ij}^t$ to $\bar{\tau}_{ij}^t$. We can express the trip time constraint as:

$$\underline{\tau}_{ij}^t \leq a_j^t - d_i^t \leq \bar{\tau}_{ij}^t, \quad \text{for } (i, j) \in \mathcal{A}^t, t \in \mathcal{T}. \quad (\text{TRACK})$$

This constraint is shown in Figure 2(a).

3.1.2. Crossing-over-based trip time constraint. A crossing-over is a directed arc that connects two train-lines, where a train-line is a directed path consisting of non-opposite platforms and non-opposite tracks. When a train turns around by traversing the crossing-over and starts

traveling through another train-line after arriving at the terminal platform of a train-line, it is considered as two different trains by the railway management.

Let φ be the set of all crossing-overs where turn-around events occur. Suppose we consider any crossing-over $(i, j) \in \varphi$ where the platforms i and j are located on different train-lines. Let \mathcal{B}_{ij} be the set of all train pairs involved in the corresponding turn-around events on the crossing-over (i, j) . Let $(t, t') \in \mathcal{B}_{ij}$, where train $t \in \mathcal{T}$ turns around at platform i by traveling through the crossing-over (i, j) , and beginning from platform j , starts traversing a different train-line as train $t' \in \mathcal{T} \setminus t$.

A time window $[\underline{\kappa}_{ij}^{tt'}, \bar{\kappa}_{ij}^{tt'}]$ must be maintained between the departure of the train from platform i (labeled as train t) and the arrival at platform j (labeled as train t'). The corresponding trip time constraint can be expressed as:

$$\underline{\kappa}_{ij}^{tt'} \leq a_j^{t'} - d_i^t \leq \bar{\kappa}_{ij}^{tt'}, \quad \text{for } (t, t') \in \mathcal{B}_{ij}, (i, j) \in \varphi. \quad (\text{CROSS})$$

This constraint is shown in Figure 2(b).

3.2. Dwell time constraint

When a train $t \in \mathcal{T}$ arrives at a platform $i \in \mathcal{N}^t$, it dwells there for a certain time interval denoted by $[\underline{\delta}_i^t, \bar{\delta}_i^t]$, during which passengers can embark or disembark. The train departs from the station once the dwell time has elapsed. The difference between the departure time d_i^t and arrival time a_i^t due to dwell time lies between $\underline{\delta}_i^t$ and $\bar{\delta}_i^t$. We can express the dwell time constraint as:

$$\underline{\delta}_i^t \leq d_i^t - a_i^t \leq \bar{\delta}_i^t, \quad \text{for } i \in \mathcal{N}^t, t \in \mathcal{T}. \quad (\text{DWELL})$$

This constraint is illustrated in Figure 2(c).

Note that the platform index i is varied over all elements of the set \mathcal{N}^t in Equation (DWELL). This is so because every train $t \in \mathcal{T}$ arrives at the first platform $\mathcal{N}^t(1)$ in its train-path either from the depot or by turning around from some other line, and departs from the final platform $\mathcal{N}^t(|\mathcal{N}^t|)$ to return to the depot or start as a new train on another line by turning around. Therefore, train t dwells at all the platforms in \mathcal{N}^t .

3.3. Connection constraint

In some cases, there might not be a direct train between the origin and destination of a passenger. To address this issue, the railway management employs connecting trains at interchange stations.

Let $\chi \subseteq \mathcal{N} \times \mathcal{N}$ be the set of platform pairs where passengers transfer between trains. If $(i, j) \in \chi$, then both platforms i and j are located at the same station, and there exists a train $t \in \mathcal{T}$ arriving

at platform i and another train $t' \in \mathcal{T}$ departing from platform j , such that a connection time window must be maintained between trains t and t' for passengers to transfer from the former to the latter. Note that order matters in this context.

Let \mathcal{C}_{ij} be the set of train pairs that enable passengers to make the corresponding connection or turn-around event for the platform pair $(i, j) \in \chi$. The connection constraint can be expressed as:

$$\underline{\chi}_{ij}^{tt'} \leq d_j^{t'} - a_i^t \leq \bar{\chi}_{ij}^{tt'}, \quad \text{for } (t, t') \in \mathcal{C}_{ij}, (i, j) \in \chi, \quad (\text{CONNECT})$$

where $\underline{\chi}_{ij}^{tt'}$ and $\bar{\chi}_{ij}^{tt'}$ are the lower and upper bounds, respectively, of the time window required to achieve the described connection between the associated trains. The connection constraint is shown in Figure 2(d).

3.4. Headway constraint

In any railway network, a minimum amount of time is always maintained between the departures of consecutive trains, known as the headway time. Let $(i, j) \in \mathcal{A}$ be the track between two platforms i and j , and let \mathcal{H}_{ij} be the set of train pairs that move along that track successively in the order of their departures.

Assume that train t and train t' move along this track in the same direction, where $(t, t') \in \mathcal{H}_{ij}$. Let $h_i^{tt'}$ and $h_j^{tt'}$ be the associated headway times at platforms i and j , respectively. The headway constraint can be expressed as follows:

$$h_i^{tt'} \leq d_i^{t'} - d_i^t \quad \text{and} \quad h_j^{tt'} \leq d_j^{t'} - d_j^t, \quad \text{for } (t, t') \in \mathcal{H}_{ij}, (i, j) \in \mathcal{A}. \quad (\text{HEADWAY})$$

The headway constraint is shown in Figure 2(e).

To ensure the safety of train operations, we must always maintain the headway constraints between two consecutive trains on the same track. Thus, when a train enters the braking phase and approaches a platform, the platform it enters cannot be immediately occupied by another train, as doing so would result in a temporal distance smaller than the headway time. It follows that selecting two consecutive trains, one accelerating and one braking at the same platform, is not feasible for synchronization from a safety standpoint. This underscores the importance of carefully considering the headway constraints when optimizing train schedules.

3.5. Total travel time constraint

In order to provide reliable service in the railway network, it is crucial to ensure that the total travel time for each train $t \in \mathcal{T}$ falls within a specified time window $[\underline{\tau}_{\mathcal{P}}^t, \bar{\tau}_{\mathcal{P}}^t]$, where $\underline{\tau}_{\mathcal{P}}^t$ and $\bar{\tau}_{\mathcal{P}}^t$ are the corresponding lower and upper bounds, respectively. This means that the time elapsed between

the train's departure from its first platform $\mathcal{N}^t(1)$ and its arrival at the last platform $\mathcal{N}^t(|\mathcal{N}^t|)$ must lie within the prescribed time window. Mathematically, we can express this constraint as:

$$\underline{\tau}_{\mathcal{P}}^t \leq a_{\mathcal{N}^t(|\mathcal{N}^t|)}^t - d_{\mathcal{N}^t(1)}^t \leq \bar{\tau}_{\mathcal{P}}^t, \quad \text{for } t \in \mathcal{T}. \quad (\text{TRAVEL})$$

It is worth noting that violating this constraint could result in delayed trains, missed connections, and reduced overall network capacity. This constraint is shown in Figure 2(f).

Domain of the event times

We can define the domain of the decision variables for the railway scheduling problem using the start and end times of the railway service period. To simplify the problem, we set the time of the first event of the service period to zero seconds, which corresponds to the departure of the first train of the day from some platform. We can then obtain an upper bound for the final event of the railway service period, which is the arrival of the last train of the day at some platform, by setting all trip times and dwell times to their maximum possible values. Let this upper bound be denoted by a positive number m . Then, the domain of the decision variables can be expressed as follows:

$$0 \leq a_i^t \leq m, \quad 0 \leq d_i^t \leq m, \quad \text{for } i \in \mathcal{N}^t. \quad (\text{DOMAIN})$$

4. Modeling the effective energy consumption

In this section, we discuss how we model the objective that minimizes the *effective energy consumption* of the railway network. The effective energy consumption is defined as the total energy consumed by all trains during their acceleration phases, denoted by E^{con} , minus the total regenerative braking energy transferred to such accelerating trains from eligible braking trains, denoted by E^{reg} . This is expressed as $E^{\text{con}} - E^{\text{reg}}$. We next discuss how we apply data-driven approaches to model E^{con} in Section 4.1 and E^{reg} in Section 4.2.

4.1. Modeling total consumed energy E^{con}

Relationship between energy consumption and the trip time constraint. Trains primarily consume electrical energy during their acceleration phase while going from an origin to a destination platform. As such, trip time constraints critically influence both energy consumption and the production of regenerative energy in trains. Once the trip time gets fixed, the energy-optimal speed profile can be efficiently computed in CBTC systems. A variety of software tools, e.g., **T**rain **K**inetics, **D**ynamics, and **C**ontrol (KDC) Simulator used in our research, can be employed for this purpose (Thales Canada Inc. 2023). The KDC simulator, based on the strategies of acceleration, speed holding, coasting, and braking, calculates the speed profile. This method aligns with the theoretical optimal speed profile as presented in the monograph (Howlett and Pudney 1995). For

a deeper dive into the calculation of optimal speed profiles, readers might consider the following papers (Jiaxin and Howlett 1993, Howlett 2000, Howlett et al. 2009, Khmelnitsky 2000, Liu and Golovitcher 2003), with a comprehensive review available in (Albrecht et al. 2016). The speed profile of a train on a track, e.g., the top subfigure of Figure 1, dictates its electrical power consumption and regeneration, leading to its power versus time graph or *power graph*, as shown in the bottom subfigure of Figure 1.

Energy consumption of all trains. We denote the total energy consumed by all trains by

$$\sum_{t \in \mathcal{T}} \sum_{(i,j) \in \mathcal{A}^t} \underbrace{E_{i,j,t}^{\text{con,tr}}}_{\text{depends on trip time } (a_j^t - d_i^t)} + \sum_{(i,j) \in \varphi, (t,t') \in \mathcal{B}_{i,j}} \underbrace{E_{i,j,t,t'}^{\text{con,cr}}}_{\text{depends on trip time } (a_j^{t'} - d_i^t)}$$

where

- $E_{i,j,t}^{\text{con,tr}} : \mathbb{R}_{++} \rightarrow \mathbb{R}_{++}$ is the energy consumed by a train $t \in \mathcal{T}$ during the acceleration phase of the trip from the origin platform i to the destination platform j with $(i,j) \in \mathcal{A}^t$. In a CBTC system, this function depends on the trip time $(a_j^t - d_i^t)$.
- $E_{i,j,t,t'}^{\text{con,cr}} : \mathbb{R}_{++} \rightarrow \mathbb{R}_{++}$ is the energy consumed by train t while traversing the crossing-over $(i,j) \in \varphi$; the energy is again consumed during the acceleration phase of t 's trip from the origin platform i to destination platform j , where it gets labeled as the train t' i.e., $(t,t') \in \mathcal{B}_{i,j}$. The trip time $a_j^{t'} - d_i^t$ associated with this crossing-over is the argument of $E_{i,j,t,t'}^{\text{con,cr}}$ in a CBTC system.

Data-driven approach to model the energy consumption. The exact analytical form of $E_{i,j,t}^{\text{con,tr}}$ or $E_{i,j,t,t'}^{\text{con,cr}}$ might be intractable as suggested by (Howlett et al. 2009, Das Gupta et al. 2016) and as shown qualitatively in Figure 1. Regardless, these functions exhibit a consistent characteristic: they are monotonically decreasing with trip time. That is to say, if an optimal speed profile is adhered to, the functions become *non-increasing* as trip time increases, as supported by (Milroy 1980). Notably, even in scenarios where trains are manually driven, potentially straying from optimal strategies, empirical evidence still supports the monotonic decrease in average energy consumption with increased trip time, as can be observed in (Pena Alcaraz et al. 2011, Figure 1).

Furthermore, gathering practical measurements for the energy function is straightforward, either by analyzing historical data or employing physics-based simulations. A crucial observation into CBTC-enabled metro railway networks is the tight margin by which trip time is permitted to vary in Equations (TRACK) and (CROSS): such variations are often on the order of seconds. This leads us to the following assumption:

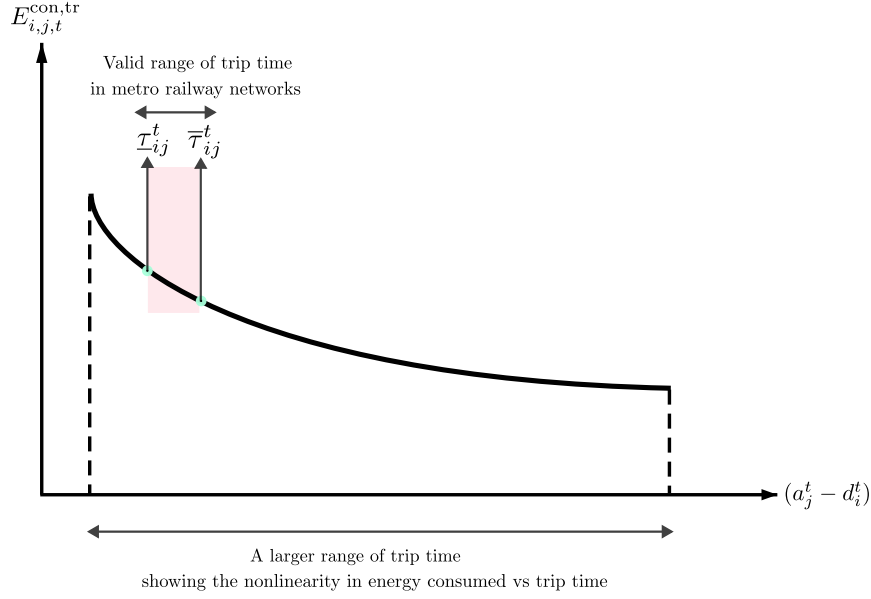


Figure 3 This figure graphically illustrates how the energy consumed by a train t going from platform i to j on track (i, j) varies as the trip time $a_j^t - d_i^t$ is varied. The longer range of trip time, while not practical for a metro railway network, is shown to illustrate the nonlinear nature of the consumed energy versus trip time in long inter-city travels that often span a few hours. On the other hand, the valid range of trip time in a metro network denoted by $[\tau_{ij}^t, \bar{\tau}_{ij}^t]$ is on the order of seconds, and in such a setup an affine approximation for the consumed energy can be reasonable.

ASSUMPTION 1. The amount by which the trip time is allowed to vary is on the order of seconds, i.e., $\bar{\tau}_{ij}^t - \tau_{ij}^t$ in (TRACK) and $\bar{\kappa}_{ij}^{tt'} - \underline{\kappa}_{ij}^{tt'}$ in (CROSS) are on the order of seconds.

Given the monotonically decreasing behavior of the energy function and Assumption 1, we can reasonably approximate the energy consumption functions as affine functions. This is qualitatively depicted in Figure 3 for track-based trip time constraints. Similar curves can be expected for crossing-over based trip time constraints. Our next objective is to derive the best affine approximation for energy consumed as a function of trip time. We'll achieve this by employing a least-squares method to fit a straight line to the energy versus trip time data, where the data can come from historical records or physics-based simulators, such as the KDC simulator, which we used in this work (Thales Canada Inc. 2023). In other words, we approximate consumed energy as:

$$\begin{aligned} E_{i,j,t}^{\text{con,tr}} &\triangleq c_{i,j,t}^{\text{con,tr}}(a_j^t - d_i^t) + b_{i,j,t}^{\text{con,tr}}, \quad \text{for } (i, j) \in \mathcal{A}^t, t \in \mathcal{T}, \\ E_{i,j,t,t'}^{\text{con,cr}} &\triangleq c_{i,j,t,t'}^{\text{con,cr}}(a_j^{t'} - d_i^t) + b_{i,j,t,t'}^{\text{con,cr}}, \quad \text{for } (i, j) \in \varphi, (t, t') \in \mathcal{B}_{i,j}, \end{aligned} \quad (1)$$

where

- $(c_{i,j,t}^{\text{con,tr}}, b_{i,j,t}^{\text{con,tr}})$ is computed by solving the following least-squares problem:

$$(c_{i,j,t}^{\text{con,tr}}, b_{i,j,t}^{\text{con,tr}}) = \underset{(\tilde{c}_{i,j,t}, \tilde{b}_{i,j,t})}{\operatorname{argmin}} \sum_{k=1}^p \left(\tilde{c}_{i,j,t} (a_j^t - d_i^t)^{(k)} + \tilde{b}_{i,j,t} - (\bar{E}_{i,j,t}^{\text{con,tr}})^{(k)} \right)^2,$$

where $(i, j) \in \mathcal{A}^t$, $t \in \mathcal{T}$, with $(\bar{E}_{i,j,t}^{\text{con,tr}})^{(k)}$ being the measured energy consumption associated with the track related trip time $(a_j^t - d_i^t)^{(k)}$ for observations $k = 1, 2, \dots, p$.

- $(c_{i,j,t,t'}^{\text{con,cr}}, b_{i,j,t,t'}^{\text{con,cr}})$ is computed by solving the following least-squares problem:

$$(c_{i,j,t,t'}^{\text{con,cr}}, b_{i,j,t,t'}^{\text{con,cr}}) = \underset{(\tilde{c}_{i,j,t,t'}, \tilde{b}_{i,j,t,t'})}{\operatorname{argmin}} \sum_{k=1}^q \left(\tilde{c}_{i,j,t,t'} (a_j^{t'} - d_i^t)^{(k)} + \tilde{b}_{i,j,t,t'} - (\bar{E}_{i,j,t,t'}^{\text{con,cr}})^{(k)} \right)^2,$$

where $(i, j) \in \varphi$, $(t, t') \in \mathcal{B}_{i,j}$ with $\bar{E}_{i,j,t,t'}^{\text{con,cr}})^{(k)}$ being the measured energy consumption related to crossing-over related trip time $(a_j^{t'} - d_i^t)^{(k)}$ for observations $k = 1, 2, \dots, q$.

4.2. Modeling transferred regenerative energy E^{reg}

Robust approximations of the power vs. time graph. Maximizing the transfer of regenerative braking energy between suitable train pairs is equivalent to maximizing the total overlapped area between the power vs. time graphs associated with power consumption and regeneration of those train pairs. The power vs. time graph for a train during its acceleration and braking phases depends on its trip time and is highly nonlinear in nature with no known analytical form in CBTC systems. However, the graph empirically exhibits a region where most of the power is concentrated, which allows us to apply the robust lumping method known as the FWHM (Full Width at Half Maximum) method, approximating the power vs. time graphs as rectangles. The FWHM method can be traced back to the early days of quantum mechanics (Gamow 1985), where it was instrumental in approximating the wave absorption vs. wavelength graph of molecules (Diem 2021, Chapter 5), and since then it has been used successfully in various fields such as numerical analysis, spectroscopy, signal processing, and statistics to approximate complicated graphs, and it's considered a useful, robust metric because it is not sensitive to the exact shape of the peak (Mahajan 2010, page 35-37). In the context of our problem, the FWHM method is applied by transforming the power vs. time graph into a rectangle (see Figure 4) whose height is the height of the peak and whose width is the full width at half maximum. These rectangles allow us to recognize the beginning and end of a train's effective accelerating or braking phases where most of the energy is concentrated, which we illustrate graphically in Figure 4.

Consider a train t arriving at (i.e., braking), then dwelling, and finally departing (i.e., accelerating) from platform i . As shown in Figure 4, the effective braking and accelerating phases of the train can be compactly contained in FWHM rectangle. We denote the beginning and end of the

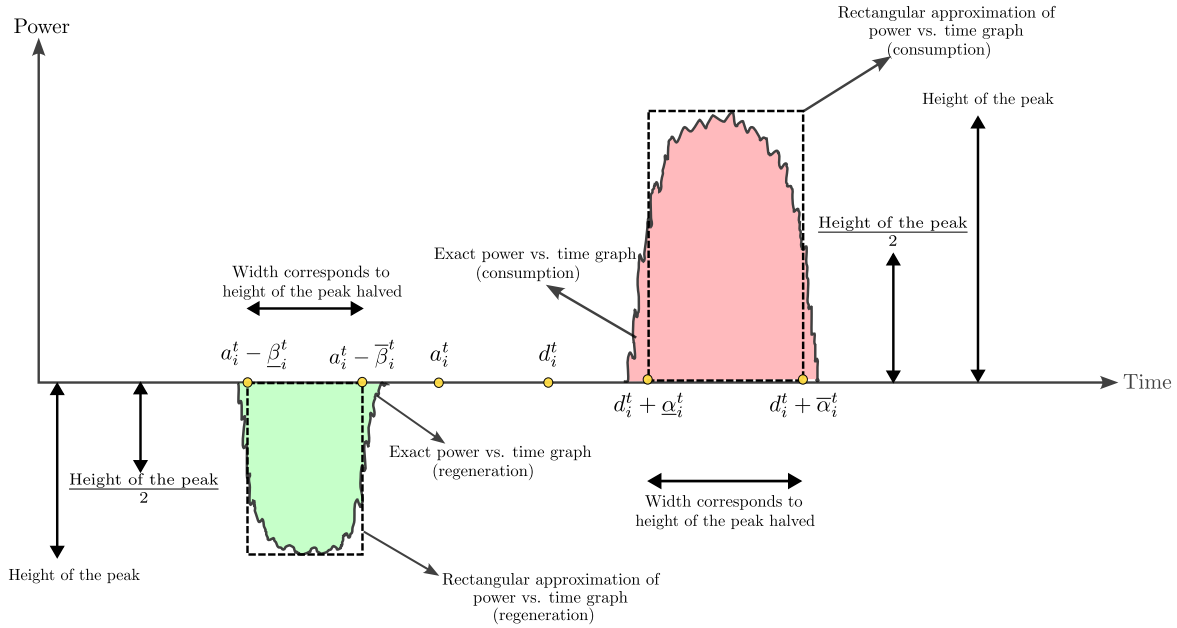


Figure 4 This figure illustrates how to compute the effective braking and acceleration phases of trains using the FWHM (Full Width at Half Maximum) method. Consider a train t arriving (i.e., braking) at, then dwelling, and finally departing (i.e., accelerating) from platform i . While the exact power vs. time graph of a train is challenging to model analytically, it is possible to compute a very robust approximation by applying the FWHM method. By doing so, the effective braking and accelerating phase of a train can be compactly represented by these rectangular approximations, which allows us to identify the beginning and end of the effective braking phase of train t , denoted by $a_i^t - \underline{\beta}_i^t$ and $a_i^t - \overline{\beta}_i^t$, respectively, and the beginning and end of the effective accelerating phase of the train t , represented by $d_i^t + \underline{\alpha}_i^t$ and $d_i^t + \overline{\alpha}_i^t$, respectively.

effective braking phase of t by $a_i^t - \underline{\beta}_i^t$ and $a_i^t - \overline{\beta}_i^t$, respectively, and the beginning and end of the effective accelerating phase of the train t by $d_i^t + \underline{\alpha}_i^t$ and $d_i^t + \overline{\alpha}_i^t$, respectively.

Determining the beginning and end of the effective braking and accelerating phases of the trains allows us to model the effective overlapping time in a very compact way. The *effective overlapping time* between two trains is defined as the duration of time during which the effective accelerating phase of the first train and the effective braking phase of the second train overlap, provided that both trains are powered by the same electrical substation, and the transmission loss associated with the transfer of regenerative energy of the braking train to the accelerating train is negligible. Thus, maximizing the transfer of regenerative energy positively correlates with the effective overlapping time between suitable train pairs, which we model next.

Constructing suitable train pairs for transfer of regenerative braking energy. The set that contains all platform pairs that are feasible for regenerative energy transfer is denoted by Ω . Consider any such platform pair $(i, j) \in \Omega$, and let $\mathcal{T}_i \subseteq \mathcal{T}$ be the set of all the trains that arrive at, dwell and then

depart from platform i . To avoid duplicates in Ω we construct the elements lexicographically with $i < j$. Suppose, $t \in \mathcal{T}_i$. Now, we are interested in finding another train \tilde{t} on platform j , i.e., $\tilde{t} \in \mathcal{T}_j$, which along with t would form a suitable pair for the transfer of regenerative braking energy. To achieve this, we start with an initial feasible timetable for the railway, which represents the desired service to be delivered. For most of the existing railway networks, the railway management has a feasible timetable. For every train t , this feasible timetable provides a feasible arrival time \bar{a}_i^t and a feasible departure time \bar{d}_i^t to and from every platform $i \in \mathcal{N}^t$ respectively. Intuitively, among all the trains that go through platform j , the one that is temporally close to t in the initial timetable would be a good candidate to form a pair with t . The temporal proximity can be of two types with respect to t , which results in the following definitions.

DEFINITION 1. Consider any $(i, j) \in \Omega$. For every train $t \in \mathcal{T}_i$, the train $\vec{t} \in \mathcal{T}_j$ is called *temporally close to the right of t* if

$$0 \leq \frac{\vec{a}_j^t + \vec{d}_j^t}{2} - \frac{\bar{a}_i^t + \bar{d}_i^t}{2} \leq r, \quad (2)$$

where r is an empirical parameter determined by the timetable designer and is much smaller than the time horizon of the entire timetable.

DEFINITION 2. Consider any $(i, j) \in \Omega$. For every train $t \in \mathcal{T}_i$, the train $\overleftarrow{t} \in \mathcal{T}_j$ is called *temporally close train to the left of t* if

$$0 < \frac{\bar{a}_i^t + \bar{d}_i^t}{2} - \frac{\overleftarrow{a}_j^t + \overleftarrow{d}_j^t}{2} \leq r. \quad (3)$$

DEFINITION 3. Consider any $(i, j) \in \Omega$. For every train $t \in \mathcal{T}_i$, the train $\tilde{t} \in \mathcal{T}_j$ is called **temporally close to t** if it is temporally close to the left or right of t .

So, any event where transfer of regenerative energy is possible can be described by specifying the corresponding $i, j, t \in \mathcal{T}_i$ and $\tilde{t} \in \mathcal{T}_j$ by using the definitions above. We construct a set of all such (i, j, t, \tilde{t}) with $i < j$ (to avoid duplicates), which we denote by \mathcal{E} . Naturally, we can partition \mathcal{E} into two sets denoted by $\vec{\mathcal{E}}$ and $\overleftarrow{\mathcal{E}}$, which contain elements of the form (i, j, t, \vec{t}) and $(i, j, t, \overleftarrow{t})$ respectively. For every $(i, j, t, \vec{t}) \in \vec{\mathcal{E}}$ (called *right event*), our strategy is to synchronize the effective accelerating phase of t with the effective braking phase of \vec{t} . On the other hand, for every $(i, j, t, \overleftarrow{t}) \in \overleftarrow{\mathcal{E}}$ (called *left event*), it is convenient to synchronize the effective accelerating phase of \overleftarrow{t} with the effective braking phase of t . For every $(i, j, t, \vec{t}) \in \vec{\mathcal{E}}$, the corresponding effective overlapping time is denoted by $\sigma_{ij}^{\vec{t}t}$, (called *right event overlapping time*) and for every $(i, j, t, \overleftarrow{t}) \in \overleftarrow{\mathcal{E}}$, the corresponding effective overlapping time is denoted by $\sigma_{ij}^{\overleftarrow{t}t}$ (called *left event overlapping time*). The overlapping time can be: *positive* when there is a positive temporal synchronization between effective acceleration phase and effective braking phase, *zero* where there is no synchronization and

the temporal distance between the phases is zero, or *negative* which corresponds to the case where their phases are apart by a certain temporal distance; we will illustrate these different overlapping times graphically later in this section.

The total regenerative energy transferred depends on the total effective overlapping time. Our objective is to maximize the transfer of regenerative energy, which we model by:

$$\sum_{(i,j,t,\vec{t}) \in \vec{\mathcal{E}}} \underbrace{E_{i,j,t,\vec{t}}^{\text{reg}}}_{\substack{\text{depends on} \\ \text{right event} \\ \text{overlapping time } \sigma_{ij}^{\vec{t} \vec{t}}}} + \sum_{(i,j,t,\overleftarrow{t}) \in \overleftarrow{\mathcal{E}}} \underbrace{E_{i,j,t,\overleftarrow{t}}^{\text{reg}}}_{\substack{\text{depends on} \\ \text{left event} \\ \text{overlapping time } \sigma_{ij}^{\overleftarrow{t} \overleftarrow{t}}}}, \quad (4)$$

where:

- $E_{i,j,t,\vec{t}}^{\text{reg}} : \mathbb{R} \rightarrow \mathbb{R}$ measures the transfer of regenerative energy from synchronizing the accelerating phase of t with the braking phase of \vec{t} , with the associated overlapping time $\sigma_{ij}^{\vec{t} \vec{t}}$ being the argument.
- $E_{i,j,t,\overleftarrow{t}}^{\text{reg}} : \mathbb{R} \rightarrow \mathbb{R}$ measures the transfer of regenerative energy from synchronizing the accelerating phase of \overleftarrow{t} with the braking phase of t , where $\sigma_{ij}^{\overleftarrow{t} \overleftarrow{t}}$ is the argument.

To model (4) in a tractable manner, we perform three steps as follows.

1. Computing data-driven approximations of $E_{i,j,t,\vec{t}}^{\text{reg}}$ and $E_{i,j,t,\overleftarrow{t}}^{\text{reg}}$.
2. Computing data-driven approximations of the effective accelerating and braking phases of trains.
3. Modeling the overlapping times.

Now we describe the three aforementioned steps in detail.

Step 1. Computing data-driven approximations of $E_{i,j,t,\vec{t}}^{\text{reg}}$ and $E_{i,j,t,\overleftarrow{t}}^{\text{reg}}$. Using a data-driven approach, we approximate $E_{i,j,t,\vec{t}}^{\text{reg}}$ and $E_{i,j,t,\overleftarrow{t}}^{\text{reg}}$ as affine functions of effective overlapping time $\sigma_{ij}^{\vec{t} \vec{t}}$ and $\sigma_{ij}^{\overleftarrow{t} \overleftarrow{t}}$, respectively. Due to Assumption 1, it is again reasonable to approximate the transferred regenerative energy functions as affine functions of the effective overlapping time. In other words, we denote:

$$\begin{aligned} E_{i,j,t,\vec{t}}^{\text{reg}} &\triangleq c_{i,j,t,\vec{t}}^{\text{reg}} \sigma_{ij}^{\vec{t} \vec{t}} + b_{i,j,t,\vec{t}}^{\text{reg}}, \\ E_{i,j,t,\overleftarrow{t}}^{\text{reg}} &\triangleq c_{i,j,t,\overleftarrow{t}}^{\text{reg}} \sigma_{ij}^{\overleftarrow{t} \overleftarrow{t}} + b_{i,j,t,\overleftarrow{t}}^{\text{reg}}, \end{aligned} \quad (5)$$

where $(c_{i,j,t,t}^{\text{reg } \rightarrow}, b_{i,j,t,t}^{\text{reg } \rightarrow})$ and $(c_{i,j,t,t}^{\text{reg } \leftarrow}, b_{i,j,t,t}^{\text{reg } \leftarrow})$ are positive coefficients computed via a constrained least squares problem where our input measurement data corresponds to effective overlapping time with output measurement data corresponds to the transferred regenerative braking energy. This measurement data can come from either historical data or physics-based simulation; in our numerical experiments, this data is generated by the KDC simulator. The positive values of $b_{i,j,t,t}^{\text{reg } \rightarrow}$ and $b_{i,j,t,t}^{\text{reg } \leftarrow}$ denote residual transfer of regenerative energy. This may occur due to some energy lying outside of our FWHM rectangles, even when the overlapping time is zero. Our affine modeling of $E_{i,j,t,t}^{\text{reg } \rightarrow}$ and $E_{i,j,t,t}^{\text{reg } \leftarrow}$ in (5) leads to the interpretation of the objective in (4), which we aim to maximize. Because the coefficients $(c_{i,j,t,t}^{\text{reg } \rightarrow}, b_{i,j,t,t}^{\text{reg } \rightarrow})$, $(c_{i,j,t,t}^{\text{reg } \leftarrow}, b_{i,j,t,t}^{\text{reg } \leftarrow})$ are nonnegative, positive overlappings would always lead to a positive transfer of regenerative energy. However, a negative or zero overlapping $\sigma_{ij}^{\rightarrow t}$ (or $\sigma_{ij}^{\leftarrow t}$) may have some small transfer of regenerative energy. For $\sigma_{ij}^{\rightarrow t} \leq -b_{i,j,t,t}^{\text{reg } \rightarrow}/c_{i,j,t,t}^{\text{reg } \rightarrow}$ (or $\sigma_{ij}^{\leftarrow t} \leq -b_{i,j,t,t}^{\text{reg } \leftarrow}/c_{i,j,t,t}^{\text{reg } \leftarrow}$), there will be little to no transfer of regenerative energy. In those cases, the negative values of $E_{i,j,t,t}^{\text{reg } \rightarrow}$ (or $E_{i,j,t,t}^{\text{reg } \leftarrow}$) introduce a *penalty* that reduces the value of the objective (4) being maximized. We make this modeling decision of introducing this penalty rather than setting the regenerative energy to zero directly because even when the overlapping time is negative, it is preferable to have the absolute value as small as possible. This approach aims to achieve some residual synchronization by reducing the penalty.

Step 2. Computing data-driven approximations of the effective accelerating and braking phases of trains. To compute the effective overlapping time, we need to compute the beginning and end of the effective accelerating and braking phases of the trains. Consider the trip of any train $t \in \mathcal{T}$ from platform i to platform j along the track $(i,j) \in \mathcal{A}^t$. The beginning and end of the effective accelerating phase of t while departing from platform i is denoted by $d_i^t + \underline{\alpha}_i^t$ and $d_i^t + \bar{\alpha}_i^t$ respectively. Similarly, the beginning and end of the effective braking phase of t while arriving at platform j by $a_j^t - \underline{\beta}_j^t$ and $a_j^t - \bar{\beta}_j^t$, respectively. The values for $\underline{\alpha}_i^t, \bar{\alpha}_i^t, \underline{\beta}_j^t, \bar{\beta}_j^t$ depend on the power graph of the train, which in turn depends on the trip time $a_j^t - d_i^t$. Analytical approximations of the aforementioned terms as a function of $a_j^t - d_i^t$ for a given train are not known, however, due to Assumption 1, it is reasonable to work with affine approximations of the trip times. So, for $t \in \mathcal{T}$, and $(i,j) \in \mathcal{A}^t$, we define $\underline{\alpha}_i^t, \bar{\alpha}_i^t, \underline{\beta}_j^t, \bar{\beta}_j^t$ as follows:

$$\begin{aligned} \underline{\alpha}_i^t &\triangleq c_{i,t}^{\alpha}(a_j^t - d_i^t) + b_{i,t}^{\alpha}, & \bar{\alpha}_i^t &\triangleq c_{i,t}^{\bar{\alpha}}(a_j^t - d_i^t) + b_{i,t}^{\bar{\alpha}}, \\ \underline{\beta}_j^t &\triangleq c_{i,t}^{\beta}(a_j^t - d_i^t) + b_{i,t}^{\beta}, & \bar{\beta}_j^t &\triangleq c_{i,t}^{\bar{\beta}}(a_j^t - d_i^t) + b_{i,t}^{\bar{\beta}}, \end{aligned} \quad (\text{EFF_ACCEL \& BRK_PHASES})$$

where the coefficients $(c_{i,t}^{\alpha}, b_{i,t}^{\alpha})$, $(c_{i,t}^{\bar{\alpha}}, b_{i,t}^{\bar{\alpha}})$, $(c_{i,t}^{\beta}, b_{i,t}^{\beta})$, and $(c_{i,t}^{\bar{\beta}}, b_{i,t}^{\bar{\beta}})$ are computed via least-squares with input measurement data corresponds to trip time with output measurement data corresponds to the beginning or end of the effective acceleration and braking phases.

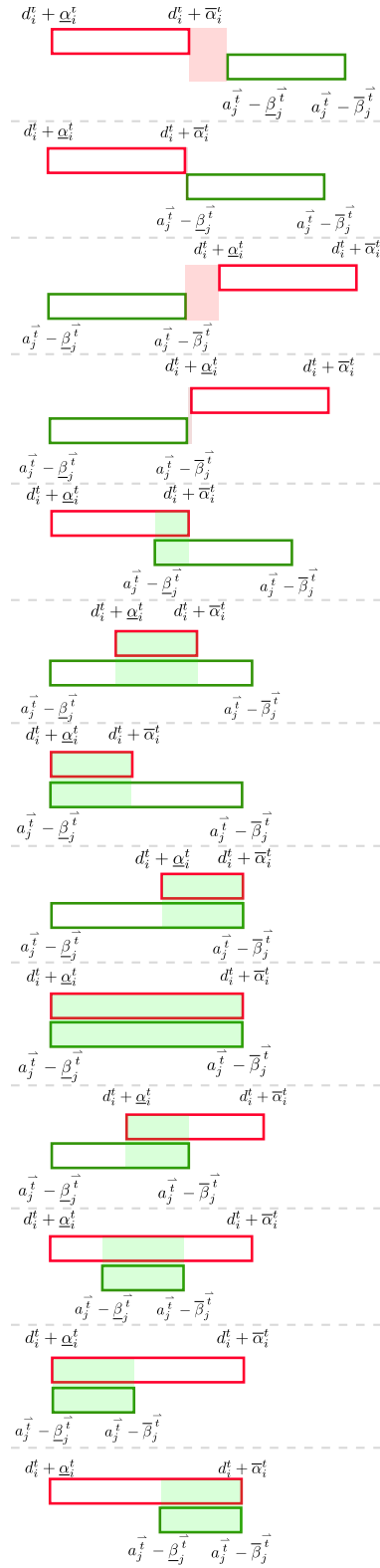


Figure 5 This figure illustrates all the possible overlapping times $\sigma_{ij}^{t \rightarrow t}$. For the first four cases, $\sigma_{ij}^{t \rightarrow t}$ shown using the red-shaded region is nonpositive, whereas for the next nine cases $\sigma_{ij}^{t \rightarrow t}$ is positive shown using the green-shaded region. The overlapping time $\sigma_{ij}^{t \rightarrow t}$ admits the closed form $\min \left\{ a_j^t - \bar{\beta}_j^t, d_i^t + \bar{\alpha}_i^t \right\} + \min \left\{ -d_i^t - \underline{\alpha}_i^t, -a_j^t + \underline{\beta}_j^t \right\}$.

Step 3. Modeling the overlapping times. Now, we describe how to model the effective overlapping time $\sigma_{ij}^{\overrightarrow{t} \overleftarrow{t}}$ and $\sigma_{ij}^{\overleftarrow{t} \overrightarrow{t}}$ in terms of the decision variables present in the system.

First, we discuss the modeling of right event overlapping time $\sigma_{ij}^{\overrightarrow{t} \overleftarrow{t}}$. Using Allen's interval algebra (Allen 1983), we know that there can be thirteen different kinds of overlapping possible between the accelerating phase of train t and the braking phase of train \overleftarrow{t} as shown in Figure 5. For the first four cases shown in Figure 5, the overlapping is nonpositive, i.e., the temporal blocks are apart from each other by a nonpositive distance (which is shown in the red-shaded region), whereas for the next nine cases, there is a positive overlapping in time (which is shown using the green-shaded region). Fortunately, we can write all these thirteen possible overlapping times in a closed form:

$$\begin{aligned} \sigma_{ij}^{\overrightarrow{t} \overleftarrow{t}} &= \min \left\{ a_j^{\overrightarrow{t}} - \overline{\beta}_j^{\overrightarrow{t}}, d_i^t + \overline{\alpha}_i^t \right\} - \max \left\{ d_i^t + \underline{\alpha}_i^t, a_j^{\overrightarrow{t}} - \underline{\beta}_j^{\overrightarrow{t}} \right\} \\ &= \min \left\{ a_j^{\overrightarrow{t}} - \overline{\beta}_j^{\overrightarrow{t}}, d_i^t + \overline{\alpha}_i^t \right\} + \min \left\{ -d_i^t - \underline{\alpha}_i^t, -a_j^{\overrightarrow{t}} + \underline{\beta}_j^{\overrightarrow{t}} \right\} \end{aligned} \quad (\text{RGHT_EVNT_OVLP_TIME})$$

which is a concave function in the decision variables $a_j^{\overrightarrow{t}}$, d_i^t , $\overline{\alpha}_i^t$, $\underline{\alpha}_i^t$, $\underline{\beta}_j^{\overrightarrow{t}}$, and $\overline{\beta}_j^{\overrightarrow{t}}$, as it is a point-wise minimum of affine functions in the decision variables, which preserves concavity (Boyd and Vandenberghe 2009, Section 3.2).

Next, we discuss modeling of left event overlapping time $\sigma_{ij}^{\overleftarrow{t} \overrightarrow{t}}$. Similar to the first case, we can write $\sigma_{ij}^{\overleftarrow{t} \overrightarrow{t}}$ as (easily proved by replacing i, j, t and \overleftarrow{t} in $\sigma_{ij}^{\overrightarrow{t} \overleftarrow{t}}$ with j, i, \overleftarrow{t} and t respectively)

$$\begin{aligned} \sigma_{ij}^{\overleftarrow{t} \overrightarrow{t}} &= \min \left\{ a_i^t - \overline{\beta}_i^t, d_j^{\overleftarrow{t}} + \overline{\alpha}_j^{\overleftarrow{t}} \right\} - \max \left\{ d_j^{\overleftarrow{t}} + \underline{\alpha}_j^{\overleftarrow{t}}, a_i^t - \underline{\beta}_i^t \right\} \\ &= \min \left\{ a_i^t - \overline{\beta}_i^t, d_j^{\overleftarrow{t}} + \overline{\alpha}_j^{\overleftarrow{t}} \right\} + \min \left\{ -d_j^{\overleftarrow{t}} - \underline{\alpha}_j^{\overleftarrow{t}}, -a_i^t + \underline{\beta}_i^t \right\}, \end{aligned} \quad (\text{LEFT_EVNT_OVLP_TIME})$$

which is again concave function in our decision variables.

5. Final optimization model

Using our developments in the previous sections, we arrive at the following final linear optimization model to minimize the effective energy consumption of the railway network:

minimize

$$\begin{aligned} & \sum_{t \in \mathcal{T}, (i,j) \in \mathcal{A}^t} (c_{i,j,t}^{\text{con,tr}} (a_j^t - d_i^t) + b_{i,j,t}^{\text{con,tr}}) + \sum_{(i,j) \in \varphi, (t,t') \in \mathcal{B}_{i,j}} (c_{i,j,t,t'}^{\text{con,cr}} (a_j^{t'} - d_i^t) + b_{i,j,t,t'}^{\text{con,cr}}) \quad \triangleright \text{consumption} \\ & - \sum_{(i,j,t,\vec{t}) \in \vec{\mathcal{E}}} \left(c_{i,j,t,\vec{t}}^{\text{reg}} \sigma_{ij}^{\vec{t}\vec{t}} + b_{i,j,t,\vec{t}}^{\text{reg}} \right) - \sum_{(i,j,t,\overleftarrow{t}) \in \overleftarrow{\mathcal{E}}} \left(c_{i,j,t,\overleftarrow{t}}^{\text{reg}} \sigma_{ij}^{\overleftarrow{t}\overleftarrow{t}} + b_{i,j,t,\overleftarrow{t}}^{\text{reg}} \right) \quad \triangleright \text{transferred regen. energy} \end{aligned}$$

subject to

$$\left. \begin{aligned} \sigma_{ij}^{\vec{t}\vec{t}} &\leq \varpi_{ij}^{\vec{t}\vec{t}} + \varphi_{ij}^{\vec{t}\vec{t}}, \quad \text{for } (i,j,t,\vec{t}) \in \vec{\mathcal{E}} \\ \varpi_{ij}^{\vec{t}\vec{t}} &\leq a_j^{\vec{t}} - \bar{\beta}_j^{\vec{t}}, \quad \text{for } (i,j,t,\vec{t}) \in \vec{\mathcal{E}} \\ \varpi_{ij}^{\vec{t}\vec{t}} &\leq d_i^{\vec{t}} + \bar{\alpha}_i^{\vec{t}}, \quad \text{for } (i,j,t,\vec{t}) \in \vec{\mathcal{E}} \\ \varphi_{ij}^{\vec{t}\vec{t}} &\leq -d_i^{\vec{t}} - \underline{\alpha}_i^{\vec{t}}, \quad \text{for } (i,j,t,\vec{t}) \in \vec{\mathcal{E}} \\ \varphi_{ij}^{\vec{t}\vec{t}} &\leq -a_j^{\vec{t}} + \underline{\beta}_j^{\vec{t}}, \quad \text{for } (i,j,t,\vec{t}) \in \vec{\mathcal{E}} \\ \sigma_{ij}^{\overleftarrow{t}\overleftarrow{t}} &\leq \varpi_{ij}^{\overleftarrow{t}\overleftarrow{t}} + \varphi_{ij}^{\overleftarrow{t}\overleftarrow{t}}, \quad \text{for } (i,j,t,\overleftarrow{t}) \in \overleftarrow{\mathcal{E}} \\ \varpi_{ij}^{\overleftarrow{t}\overleftarrow{t}} &\leq a_i^{\overleftarrow{t}} - \bar{\beta}_i^{\overleftarrow{t}}, \quad \text{for } (i,j,t,\overleftarrow{t}) \in \overleftarrow{\mathcal{E}} \\ \varpi_{ij}^{\overleftarrow{t}\overleftarrow{t}} &\leq d_j^{\overleftarrow{t}} + \bar{\alpha}_j^{\overleftarrow{t}}, \quad \text{for } (i,j,t,\overleftarrow{t}) \in \overleftarrow{\mathcal{E}} \\ \varphi_{ij}^{\overleftarrow{t}\overleftarrow{t}} &\leq -d_j^{\overleftarrow{t}} - \underline{\alpha}_j^{\overleftarrow{t}}, \quad \text{for } (i,j,t,\overleftarrow{t}) \in \overleftarrow{\mathcal{E}} \\ \varphi_{ij}^{\overleftarrow{t}\overleftarrow{t}} &\leq -a_i^{\overleftarrow{t}} + \underline{\beta}_i^{\overleftarrow{t}}, \quad \text{for } (i,j,t,\overleftarrow{t}) \in \overleftarrow{\mathcal{E}} \end{aligned} \right\} \begin{aligned} & \triangleright \text{hypograph} \\ & \text{constraints} \\ & \text{for(RGHT_EVNT_OVLP_TIME)} \end{aligned}$$

Eq. (EFF_ACCEL_&_BRK_PHASES), \triangleright beginning and end of acceleration and braking

Eq. (TRACK), (CROSS), (DWELL), (CONNECT), (HEADWAY), (TRAVEL), \triangleright operational constraints

Eq. (DOMAIN), \triangleright domain of decision variable

(OPT_MODEL)

where the decision variables are a_i^t , d_i^t , $\sigma_{ij}^{\vec{t}\vec{t}}$, $\varpi_{ij}^{\vec{t}\vec{t}}$, $\varphi_{ij}^{\vec{t}\vec{t}}$, $\sigma_{ij}^{\overleftarrow{t}\overleftarrow{t}}$, $\varpi_{ij}^{\overleftarrow{t}\overleftarrow{t}}$, and $\varphi_{ij}^{\overleftarrow{t}\overleftarrow{t}}$. Note that the first ten constraints recast (RGHT_EVNT_OVLP_TIME) and (LEFT_EVNT_OVLP_TIME) as linear constraints using the hypograph approach (Boyd and Vandenberghe 2009, Section 4.1.3).

Predicting effective energy consumption from the solution. Once we have computed an optimal solution to (OPT_MODEL), it provides us with the energy-optimal timetable. An estimate of the effective energy consumption of the final timetable is given by:

$$\begin{aligned} & \sum_{t \in \mathcal{T}, (i,j) \in \mathcal{A}^t} (c_{i,j,t}^{\text{con,tr}} (a_j^t - d_i^t) + b_{i,j,t}^{\text{con,tr}}) + \sum_{(i,j) \in \varphi, (t,t') \in \mathcal{B}_{i,j}} (c_{i,j,t,t'}^{\text{con,cr}} (a_j^{t'} - d_i^t) + b_{i,j,t,t'}^{\text{con,cr}}) \\ & - \sum_{(i,j,t,\vec{t}) \in \vec{\mathcal{E}}} \max \left\{ c_{i,j,t,\vec{t}}^{\text{reg}} \sigma_{ij}^{\vec{t}\vec{t}} + b_{i,j,t,\vec{t}}^{\text{reg}}, 0 \right\} - \sum_{(i,j,t,\overleftarrow{t}) \in \overleftarrow{\mathcal{E}}} \max \left\{ c_{i,j,t,\overleftarrow{t}}^{\text{reg}} \sigma_{ij}^{\overleftarrow{t}\overleftarrow{t}} + b_{i,j,t,\overleftarrow{t}}^{\text{reg}}, 0 \right\}, \end{aligned}$$

(PRED_EFF_ENG_CNISM)

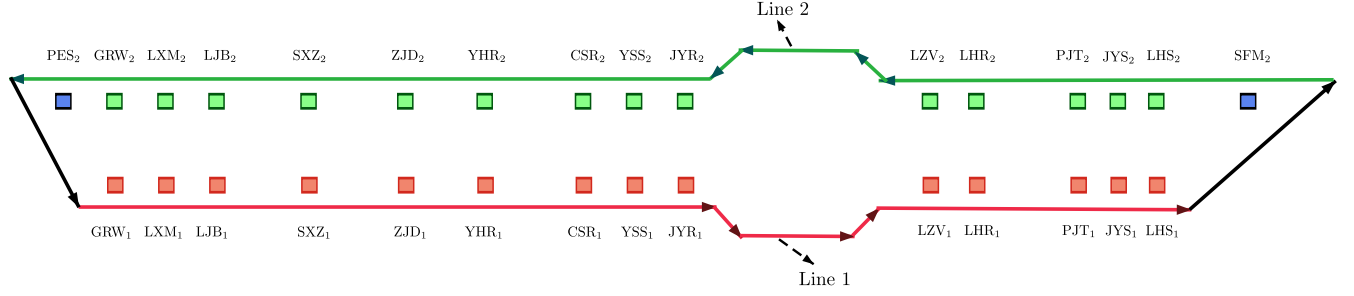


Figure 6 Railway network considered for the numerical study.

where the first line computes the energy consumption during the acceleration, whereas the second line is the negative of the actual transfer of regenerative energy, thus the final expression corresponds to the predicted effective energy consumption. Recall that when $\sigma_{ij}^{\vec{t}} \leq 0$ and $\sigma_{ij}^{\leftarrow t} \leq 0$, the summands $c_{i,j,t,t}^{\text{reg}} \rightarrow \sigma_{ij}^{\vec{t}} + b_{i,j,t,t}^{\text{reg}} \rightarrow$ and $c_{i,j,t,t}^{\text{reg}} \leftarrow \sigma_{ij}^{\leftarrow t} + b_{i,j,t,t}^{\text{reg}} \leftarrow$ act as penalty terms in the objective function of (OPT_MODEL), so $\max\{c_{i,j,t,t}^{\text{reg}} \rightarrow \sigma_{ij}^{\vec{t}} + b_{i,j,t,t}^{\text{reg}} \rightarrow, 0\}$ and $\max\{c_{i,j,t,t}^{\text{reg}} \leftarrow \sigma_{ij}^{\leftarrow t} + b_{i,j,t,t}^{\text{reg}} \leftarrow, 0\}$ correspond to the actual transferred regenerative energy for the individual events, respectively.

6. Numerical experiment

This section has the following organization. In Section 6.1, we describe the railway network where we apply our model. In Section 6.2, we present the results of our numerical study.

6.1. Railway network in consideration

Our empirical analysis focuses on Metro Line 8 of the Shanghai Railway network, a critical component of the world's most extensive metro system by route length. The Shanghai Railway Network also stands second in terms of station count while leading in global metro ridership with approximately 3.88 billion passengers served in 2019 (Wikipedia 2023a). The Metro Line 8 infrastructure, in particular, traverses some of Shanghai's highest-density residential areas and records an average daily ridership of about 1.08 million. The line spans 37.4 km and encompasses 28 functional stations (Wikipedia 2023b).

Metro Line 8 features three distinct services, each governed by individually optimized timetables generated through communication-based train control systems. In this study, we formulate energy-efficient timetables specifically for the PES₂-SFM₂ service route within Metro Line 8, as depicted in Figure 6. The said network is composed of two lines—Line 1 and Line 2—with 14 stations represented in all uppercase in the figure. Notably, each station comprises dual platforms catering to trains from both lines; for instance, LXM Station includes platforms LXM₁ and LXM₂, which serve Line 1 and Line 2, respectively.

The line under examination has a total length of 37 km. The stations have an average inter-station distance of 1.4 km, ranging from a minimum of 738 m (between YHR and ZJD) to a maximum of 2.6 km (between PJT and LHR). Operational constraints are subject to well-defined physical limits, including the track slope, which lies within the range of $[-2.00453^\circ, 2.00453^\circ]$, and the maximum permissible acceleration and deceleration of trains, specified as 1.04 m/s^2 and -0.8 m/s^2 , respectively. The conversion efficiency from electrical to kinetic energy is 0.9, and from kinetic to regenerative braking energy is 0.76, as well as the transmission loss factor of regenerative electricity is 0.1.

6.2. Results of the numerical study

To compute the energy-efficient timetables for the PES₂-SFM₂ service on Shanghai Railway Network's Metro Line 8, we consider 11 distinct operational instances. Each instance encompasses unique parameters, including train count, headway durations, train velocities, track gradients, and energy consumption profiles during acceleration and braking phases. Thales Canada Inc. has provided us with existing timetables corresponding to these 11 scenarios. Upon the computation of our energy-optimized timetables, we quantify the improvements by comparing the energy usage against these baseline timetables.

Table 2 Results of the numerical study applied Metro Line 8 of Shanghai Railway Network.

#Trains	#Variables	#Constraints	Solution time (s)	Initial effective energy consumption (kWh)	Predicted final effective energy consumption (kWh)	Reduction predicted by our model (%)	Reduction predicted by SPSIM (%)
1000	47581	151259	0.669	353721.62	252282.87	28.68	30.27
1032	49104	156102	0.674	364832.93	263288.46	27.83	31.34
1066	50726	161254	0.731	376017.98	271346.61	27.84	33.04
1100	52339	166391	0.683	386087.42	284325.21	26.36	29.10
1132	53865	171239	0.760	389844.65	291664.27	25.18	26.56
1166	55487	176391	0.753	407717.20	301714.07	26.00	26.55
1198	57010	181234	0.763	417767.32	330319.76	20.93	19.27
1232	58626	186376	0.855	447025.23	334638.29	25.14	25.86
1266	60242	191518	0.867	438657.08	332747.39	24.14	26.94
1298	61765	196361	0.901	449379.24	333858.16	25.71	24.83
1332	63387	201513	0.906	475354.59	350642.67	26.24	25.97

The computational experiments were conducted on a high-performance computing infrastructure comprising an AMD Ryzen 9 7950X CPU with 16 cores and 32 threads, accompanied by 32 GB of RAM, and operating on a Windows 11 Pro platform. We employed the algebraic modeling language JuMP, a state-of-the-art, open-source framework, to articulate our optimization problem (Lubin et al. 2023). Within this framework, we implemented the parallel interior-point algorithm of the Gurobi Optimizer 10.0 (Gurobi Optimization, LLC 2023). To accelerate the optimization process,

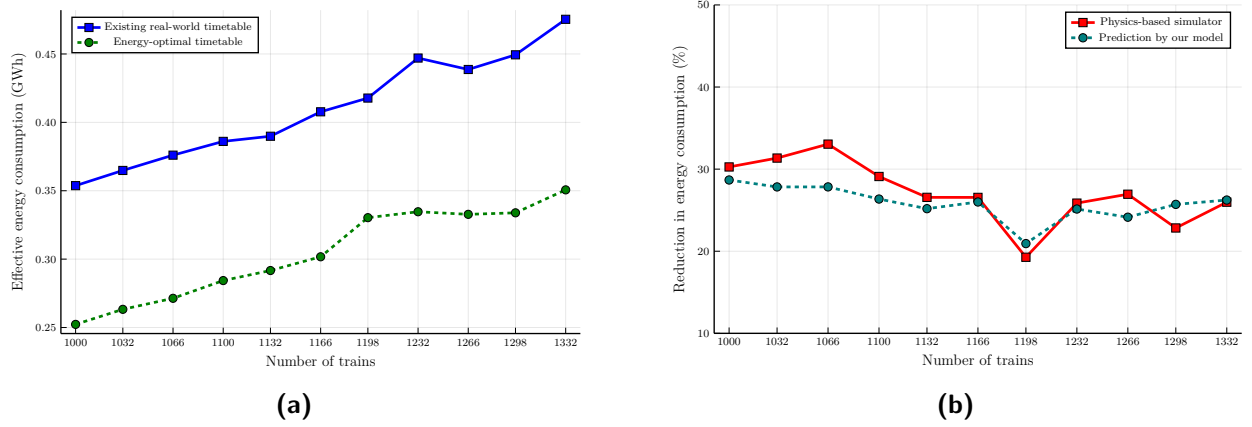


Figure 7 Figure 7a compares the predicted effective energy consumption of the energy-optimal timetables (prediction by (PRED_EFF_ENG_CNEM)) to that of the existing real-world timetables. Figure 7b shows that the reduction predicted by our model is quite close to proprietary physics-based simulator SPSIM.

we initialized the algorithm using existing timetables generously supplied by Thales Canada Inc, significantly reducing the computational time required to reach an optimal solution.

The findings of this computational study are encapsulated in Table 2. Remarkably, in every scenario, the solution time was confined to less than one second, effectively rendering the model amenable to real-time applications. The baseline effective energy consumption metrics, associated with the initial timetables, were ascertained through SPSIM, a physics-based simulation tool utilized by Thales Canada Inc. Upon obtaining the optimal solution, as formulated in equation (OPT_MODEL), we proceeded to evaluate the effective energy consumption of the optimized timetables utilizing equation (PRED_EFF_ENG_CNEM). The optimized timetables demonstrated a substantial improvement over their real-world counterparts, with the reductions in effective energy consumption spanning from approximately 20.93% in the least favorable instances to as much as 28.68% in the most advantageous cases. This is also graphically illustrated in Figure 7a.

To corroborate the validity of our model's predictions, we conducted a comparative analysis with the physics-based simulator SPSIM. The latter's estimates were generated by inputting our optimized timetables into the simulator, calculating the resultant effective energy consumption and subsequently the degree of reduction. The comparative metrics, enumerated in the last two columns of Table 2 and Figure 7b, reveal a strong alignment between the two prediction methodologies, although our model tends to yield slightly more conservative estimates on average. The variance can be attributed to the methodology employed by SPSIM for computing the transfer of regenerative energy, which is based on the precise area of overlap between power-versus-time graphs for pairs of accelerating and decelerating trains. This may occasionally exceed the energy transfer estimates

provided by our model, which employs rectangular approximations of these graphs, as delineated in Section 4.2.

7. Architectural framework for industrial integration of the proposed model

This section describes the architectural framework that facilitates the incorporation of the optimization model proposed in this paper into Thales Canada Inc's industrial timetable compiler. As shown in Figure 8, the core architecture has two primary constituents: (i) the existing timetable compiler, showcased on the left, and (ii) the energy-optimal timetable generator, represented on the right.

For the purpose of this integration framework, the existing timetable compiler is partitioned into three discrete modules:

- *Timetable data*: This module houses the necessary data to solve the optimization problem. Additionally, it includes other information specific to the railway network under consideration, such as gradient attributes and speed limitations in various track segments; all presented in their raw form.
- *Optimization class*: This module transforms the raw data into a structure that is compatible with the optimization algorithm we employ.
- *CSCPP class*: This class establishes a bi-directional communication link between the industrial timetable compiler and the energy-optimal timetable generator. It injects the input data into the optimization algorithm, retrieves the optimized timetable, and converts it into a format amenable for deployment in a communication-based train control system.

The energy-optimal timetable generator is similarly sub-divided into three modules:

- *Type definitions*: This module contains the data structures essential for describing a railway network, as previously elaborated in Sections 3 and 5.
- *Utility functions*: This module contains the critical functions required for the optimization algorithm's operation. Included are the description of the optimization model (OPT_MODEL), a function to predict effective energy consumption via (PRED_EFF_ENG_CNSM), and another to transform the output timetable in a CSCPP-compatible format.
- *Optimization algorithm*: This segment contains the parallel interior-point algorithm used for solving (OPT_MODEL), supplemented by a warm-start mechanism. Once the final timetable is computed, the algorithm forwards the results to the CSCPP class through the utility functions.

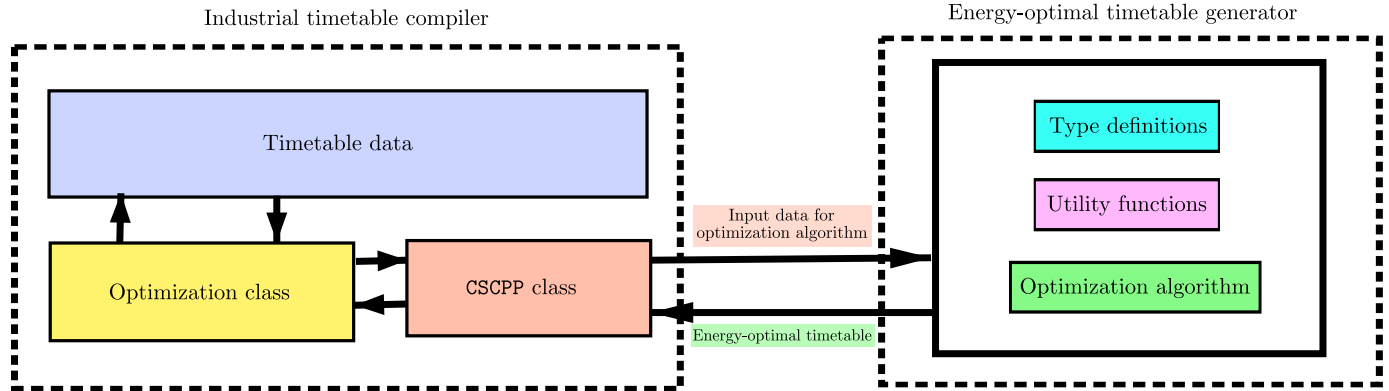


Figure 8 The integration architecture for integrating our optimization model with an industrial timetable compiler

8. Conclusion

In this paper, we have proposed a novel single-stage linear optimization model to compute energy-optimal timetables for sustainable communication-based train control systems. Our model simultaneously minimizes the total energy consumption of all trains and maximizes the transfer of regenerative braking energy. Distinct from existing models, that are either NP-hard or require multi-stage simulations, our model facilitates real-time decision-making by producing energy-optimal timetables subject to the inherent functional constraints of metro railway networks. We demonstrated the model's impact via its application to Shanghai Railway Network's Metro Line 8, achieving energy savings between 20.93% and 28.68% in comparison with existing real-world timetables, with sub-second computational times on a standard desktop. Given its managerial implications and computational robustness, our model is poised for global implementation as an integrated component of Thales Canada Inc's timetable compiler.

References

- Abril M, Barber F, Ingolotti L, Salido M, Tormos P, Lova A (2008) An assessment of railway capacity. *Transportation Research Part E: Logistics and Transportation Review* 44(5):774–806.
- Albrecht A, Howlett P, Pudney P, Vu X, Zhou P (2016) The key principles of optimal train control - part 1: Formulation of the model, strategies of optimal type, evolutionary lines, location of optimal switching points. *Transportation Research Part B: Methodological* 94:482–508.
- Allen JF (1983) Maintaining knowledge about temporal intervals. *Communications of the ACM* 26(11):832–843.
- Boyd S, Vandenberghe L (2009) *Convex optimization* (Cambridge university press).
- Das Gupta S, Pavel L, Kevin Tobin J (2015) An optimization model to utilize regenerative braking energy in a railway network. *American Control Conference (ACC)*.

- Das Gupta S, Tobin JK, Pavel L (2016) A two-step linear programming model for energy-efficient timetables in metro railway networks. *Transportation Research Part B: Methodological* 93:57–74.
- Diem M (2021) *Quantum Mechanical Foundations of Molecular Spectroscopy* (John Wiley & Sons).
- Gamow G (1985) *Thirty years that shook physics: The story of quantum theory* (Courier Corporation).
- Gurobi Optimization, LLC (2023) Gurobi Optimizer Reference Manual. URL <https://www.gurobi.com>.
- Howlett P (2000) The optimal control of a train. *Annals of Operations Research* 98:65–87.
- Howlett PG, Pudney PJ (1995) *Energy-efficient train control* (Springer Science & Business Media).
- Howlett PG, Pudney PJ, Vu X (2009) Local energy minimization in optimal train control. *Automatica* 45(11):2692–2698.
- IRSE International Technical Committee (2009) Semi-automatic, driverless, and unattended operation of trains. Available at: <http://tinyurl.com/2k2u3ktv>.
- Jiaxin C, Howlett P (1993) A note on the calculation of optimal strategies for the minimization of fuel consumption in the control of trains. *IEEE Transactions on Automatic Control* 38(11):1730–1734.
- Khmelnitsky E (2000) On an optimal control problem of train operation. *IEEE transactions on automatic control* 45(7):1257–1266.
- Le Z, Li K, Ye J, Xu X (2014) Optimizing the train timetable for a subway system. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* URL <http://dx.doi.org/10.1177/0954409714524377>.
- Li X, Lo HK (2014) An energy-efficient scheduling and speed control approach for metro rail operations. *Transportation Research Part B: Methodological* 64:73–89.
- Liu RR, Golovitcher IM (2003) Energy-efficient operation of rail vehicles. *Transportation Research Part A: Policy and Practice* 37(10):917–932.
- Lubin M, Dowson O, Dias Garcia J, Huchette J, Legat B, Vielma JP (2023) JuMP 1.0: Recent improvements to a modeling language for mathematical optimization. *Mathematical Programming Computation* URL <http://dx.doi.org/10.1007/s12532-023-00239-3>.
- Mahajan S (2010) *Street-fighting mathematics: the art of educated guessing and opportunistic problem solving* (The MIT Press).
- Milroy IP (1980) *Aspects of automatic train control*. Ph.D. thesis, Loughborough University.
- Mohamed A, Reid A, Lamb T (2018) White paper on wayside energy storage for regenerative braking energy recuperation in the electric rail system. *Con Edison: New York, NY, USA* 1–20.
- Pascoe RD, Eichorn TN (2009) What is communication-based train control? *IEEE Vehicular Technology Magazine* 4(4):16–21.

- Peña-Alcaraz M, Fernández A, Cucala AP, Ramos A, Pecharromán RR (2012) Optimal underground timetable design based on power flow for maximizing the use of regenerative-braking energy. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 226(4):397–408.
- Pena Alcaraz M, Webster M, Ramos A (2011) An approximate dynamic programming approach for designing train timetables .
- Thales Canada Inc (2023) SelTrac communications-based train control for urban rail. <https://www.thalesgroup.com/en/communications-based-train-control-cbtc-signalling>.
- United Nations (2014) World urbanization prospects. <https://esa.un.org/unpd/wup/Publications/Files/WUP2014-Report.pdf>.
- Wang P, Goverde RM (2017) Multi-train trajectory optimization for energy efficiency and delay recovery on single-track railway lines. *Transportation Research Part B: Methodological* 105:340–361.
- Wang P, Goverde RM (2019) Multi-train trajectory optimization for energy-efficient timetabling. *European Journal of Operational Research* 272(2):621–635.
- Wikipedia (2023a) Description of shanghai metro network. https://en.wikipedia.org/wiki/Shanghai_Metro, online; Wikipedia, The Free Encyclopedia.
- Wikipedia (2023b) Line 8 (shanghai metro). [https://en.wikipedia.org/wiki/Line_8_\(Shanghai_Metro\)](https://en.wikipedia.org/wiki/Line_8_(Shanghai_Metro)), online; Wikipedia, The Free Encyclopedia.
- Yang S, Liao F, Wu J, Timmermans HJ, Sun H, Gao Z (2020) A bi-objective timetable optimization model incorporating energy allocation and passenger assignment in an energy-regenerative metro system. *Transportation Research Part B: Methodological* 133:85–113.