



Received 28 October 2023, accepted 15 December 2023, date of publication 22 December 2023, date of current version 29 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3345879

RESEARCH ARTICLE

Optimization of Speed Profiles and Time Schedule of the Urban Rail Transit for Energy-Efficient Operation

YEUN SUB BYUN^(D) AND RAG GYO JEONG

Korea Railroad Research Institute, Uiwang, Gyeonggi 16105, Republic of Korea Corresponding author: Yeun Sub Byun (ysbyun@krri.re.kr)

This work was supported by the Research and Development Program through the Development of Autonomous Train Control Core Technology by the Korea Railroad Research Institute under Grant PK2301B1.

ABSTRACT Various optimal control strategies based on dynamic programming (DP) have been devised with the primary objective of energy conservation in the operation of Metro systems. Essentially, the DP algorithms strive to define a speed profile that reduces energy consumption while meeting specific journey distances and targeted travel times, thereby optimizing the operational speed of a train traversing between stations within permissible limits. However, there has been a relative paucity of research dedicated to ascertaining the ideal travel time for each station segment in an optimal manner. In this regard, this study devised a DP methodology with the goal of optimizing both the speed and the travel time for each station section, taking into account a predetermined total running time for a line composed of multiple stations under single train circumstances. To verify its effectiveness, the proposed optimization approach was compared with actual train operational data, encompassing recorded speeds and energy usage. Simulations were carried out mirroring real-world constraints, and the resultant solution adheres to the same speed limitations, distance, and travel time. Additionally, a methodology was introduced to account for any discrepancies between the actual train system and the simulation model. This was validated by comparing the actual energy consumption with the modeled energy consumption for each operational segment, demonstrating a modeling error of less than 1%.

INDEX TERMS Dynamic programming, energy efficient operation, optimal control, optimization of speed profile, urban rail transit, train.

I. INTRODUCTION

The significance of large-scale public transportation continues to grow, with urban rail systems earning increasing recognition owing to their operational efficiency, safety, and user-friendly features. As the urban population increases, the use of road transport in inner cities also increases, a trend that simultaneously increases energy consumption, pollution of various forms, and traffic congestion. Urban rail systems exhibit commendable efficiency in terms of energy consumption, consuming merely one-ninth of that used by passenger cars and half of that used by buses [1]. Despite this relative

The associate editor coordinating the review of this manuscript and approving it for publication was Hassan Omar^(D).

efficiency, railways do contribute to substantial energy use, prompting a multitude of studies aimed at enhancing energy efficiency in these systems. It has been reported that more than 80% of the total operational energy in railways is consumed by the train's internal system, with the train traction system alone constituting over 50% [2], [3]. Consequently, this study focuses on factors affecting the energy efficiency of urban rail vehicles within railway networks. Through the modification of individual train speed profiles while adhering to all operational constraints and without the need for costly infrastructure upgrades, it is suggested that operational energy efficiency can result in energy savings ranging from 5-20% [4], [5]. Pioneering research into the minimization of travel energy to reduce the operational costs of trains

was conducted by Ichikawa [6] in 1968. That seminal study used numerical methods to address the optimization issue of train speed trajectory. Following this, subsequent research has been wide-ranging and can be broadly categorized into analytical, numerical methods, and other methods. Analytical approaches are based on optimal control theory [7], integrating the maximum principle for matters relating to energy minimization. For instance, Howlett [8] employed the Pontryagin principle to establish optimal control for an ungraded track to minimize fuel consumption within a given run time, consisting of a power, hold, coast, and brake sequence.

Khmelnitsky [9] devised a rigorous strategy for the implementation of traction and braking to curtail energy consumption, accounting for varying grade profiles and adhering to arbitrary speed constraints. The optimal solution was procured through a maximum principle analysis. Meanwhile, Liu and Golovitcher [10] envisaged the operation of railway cars as an optimal control problem, utilizing a combination of five operational states to minimize energy consumption via a process rooted in the maximum principle.

Various numerical strategies addressing energy conservation challenges exploit diverse pathfinding algorithms anchored in discrete state space [11]. An investigation assessed the strengths and limitations of the three primary methodologies: dynamic programming (DP) [12], the gradient method [13], and sequential quadratic programming (SQP) [14]. A study proposed a DP method to tackle issues pertaining to constraints in three dimensions: time, distance, and speed [15]. Concurrently, another study introduced a weighted cost function in a two-dimensional space, encompassing distance and speed, to mitigate the computational time burden induced by multiple dimensions [16]. An additional study segmented a train's operational scenarios into four modes (acceleration, cruising, coasting, and braking) and optimized the sequence of these modes to compute the travel trajectory for energy optimization [17].

Noteworthy studies include [18], which utilized Bellman's dynamic programming to optimize train running speed, minimizing total energy consumption. Calderaro et al. [19] employed dynamic programming to identify a series of pseudo-optimal speed cycles with the objective of reducing the electrical energy required for traction. Bin et al. [20] optimized the speed profile of a high-speed train on a railway line with an unpowered neutral segment by deploying dynamic programming.

Various heuristic algorithms have also been examined. Kim and Chien developed a simulated annealing algorithm to identify optimal train operation for energy consumption minimization, taking into consideration factors like track alignment, speed limit, and timetable adherence [21]. Xiang and Hong [22] fashioned a combined model to optimize both the timetable and train speed profile utilizing a genetic algorithm, resulting in energy savings of up to 25%. Tang et al. [23] reduced the power consumption using a genetic algorithm on a traditional electrical model, which utilizes the regenerative energy between two trains [23].

VOLUME 11, 2023

Summarizing the findings of relevant previous studies yields the following observations.

A. SHORTCOMINGS OF PREVIOUS RESEARCH

1) Previous studies have primarily focused on optimizing the speed profiles for individual station-to-station segments while determining the segment travel times, approaching this aspect as a local optimization issue for each segment. Additionally, the method of setting the segment travel times during this process has not been clearly defined. Consequently, research on the methods of optimizing the travel times between stations across the entire railway route has been scarce.

2) Previous studies have not entailed a comprehensive analysis of the accuracy issues in system modeling. Furthermore, the research on mitigating the errors caused by discrepancies between the model and the real system has been limited.

B. ADVANTAGES AND CONTRIBUTIONS OF THE PROPOSED METHOD

1) This paper proposes a method to optimize the operation of the entire railway route, considering the travel times between each station, to achieve optimal operation across the entire route. This holistic optimization approach is more efficient than that presented in the previous studies and accounts for the interactions between multiple stations.

2) New methods are introduced to address the inaccuracies in system modeling. These methods identify and correct the differences between the actual energy consumption of trains, and the model-calculated energy usage by estimating the unknown external forces acting on the vehicle model. This improves the accuracy of the model and the performance of optimization solutions.

3) This paper presents methods for energy consumption and speed profile optimization for the entire operational line, which have not been addressed in previous studies. The results herein can help determine the range of the minimum and maximum travel times, along with the energy consumption for the entire route or each station segment, improving the efficiency of the overall operations and reducing the energy consumption and overall cost.

4) The results yielded by the proposed method are validated using real operational data, demonstrating a reduction of approximately 14% in energy consumption compared to existing operational methods. This demonstrates that the proposed method is both realistic and effective.

Therefore, this study can overcome the limitations of the previous studies and significantly contribute to the field by providing innovative methods of enhancing the energy efficiency of urban rail systems.

Reviewing prior research, it is evident that Pontryagin's minimum principle and dynamic programming optimization are frequently deployed to navigate driving optimization problems. Pontryagin's minimum principle tends to be used for simpler problems, while dynamic programming optimization is favored for more complex system models with multiple constraints [16]. In the current study, the optimal speed profile was determined using dynamic programming while accounting for several constraints. If the optimal travel speed between each inter-station segment is computed in parallel based on travel time, it can be documented as time-of-day energy data. This segmented information can subsequently be harnessed to optimize the operation of the entire section.

A two-phase procedure is herein proposed for the operational optimization design of the entire route. In the first phase, we employ a weighted parallel dynamic programming method [24]—applied to mixed cost functions of time and energy—to design the speed profile necessary for optimal travel for each station segment. This approach permits us to acquire the optimal speed profiles and corresponding energy consumption for various travel durations, which are segmented between minimum and maximum travel times.

In the second phase, we reapply the DP method, this time utilizing the travel time and energy data tables generated in the initial step to optimize the operation of the entire line. This optimization yields the hourly energy consumption for the entire line, and the speed profile for each station from the initial step is selected based on the total travel time.

The remainder of this paper is structured as follows. Section II describes the train model to be optimized and the power consumed during operation. This section also introduces a strategy to rectify potential inaccuracies that may arise during the modeling process. Section III provides a detailed description of the proposed two-stage DP algorithm for energy optimization across all operational segments. Section IV offers a comparison of pre- and post-optimization data to validate the efficacy of the suggested methodology. The final section outlines the conclusions obtained from the study.

II. RAILWAY SYSTEM

To develop a speed profile that minimizes the operational energy of a train, a dynamic model of the train is necessary, as is a definition and mathematical model of the external forces acting on the train during motion. Additionally, a power calculation equation is established to translate the force used during transit into energy. The data used in this paper are actual train journey measurements, incorporating factors such as each measurement's time, travel position, train speed, speed limit, line voltage, current, weight, and the distance between stations. The train operated in automated train operation (ATO) mode without energy optimization.

A. MATHEMATICAL FORMULATION OF TRAIN MOTION

The train model in this study is hypothesized to be a single point mass and is represented as follows:

$$ma = F_t - (F_r + F_g + F_x) \tag{1}$$

where *m* denotes the mass of the train, *a* refers to the acceleration of the train, and F_t symbolizes the traction and braking force implemented by the propulsion and braking system. F_r is the running resistance, which is influenced by the train's





travel speed. F_g is the gradient resistance, determined by the slope of the track. Moreover, the force obstructing the train's motion, which varies based on running position elements such as curve resistance and tunnel resistance, is defined as F_x . This includes model inaccuracies that may arise because of actual measurement errors, like unknown resistance and gradients that depend on the travel position, which are not defined in this mathematical model. The train's traction and braking forces are defined as (2), a function of the propulsion and braking forces approximated based on the parameters of the propulsion system. Fig. 1 presents a graph of the electric motor torque curve and running resistance.

$$F_{tMax} = \begin{cases} F_{acMax}, & v_h \le V_{ac} \cap a \ge 0\\ F_{acMax} \cdot V_{ac} / v_h, & v_h > V_{ac} \cap a \ge 0\\ F_{dcMax}, & v_h \le V_{dc} \cap a < 0\\ F_{dcMax} \cdot V_{dc} / v_h, & v_h > V_{dc} \cap a < 0 \end{cases}$$
(2)

where v_h represents the train velocity (km/h), V_{ac} is the speed at which the motor transitions into the constant power region, F_{acMax} is the maximum torque in the constant torque region, V_{dc} is the speed at which the motor enters the constant power region during braking, and F_{dcMax} is the maximum braking torque in the constant torque region.

The running resistance, encompassing both rolling resistance and air resistance, is calculated using the widely utilized Davis equation [25] as follows:

$$F_r = \left(a_r + b_r \cdot v_h + c_r \cdot v_h^2\right) \cdot mg/1000 \tag{3}$$

where a_r , b_r , and c_r represent the train's rolling characteristic coefficients and g represents the gravitational acceleration.

The gradient resistance is computed using (4), which combines the gradient of the train track and the weight of the train. The observed gradient data is applied based on the train's track position.

$$F_g = mg \cdot \sin(\theta(d)) \tag{4}$$

where $\theta(d)$ represents the track gradient at the train's current travel position. Table 1 lists the parameters utilized in the train model in this study.

B. TRAIN POWER AND ENERGY CONSUMPTION

A train dynamics model is deployed to calculate the commonly-used train power consumption model. Initially,

TABLE 1. Train model parameters.

<i>m</i> (kg)	168,000
FacMax (kN)	240.05
F _{dcMax} (kN)	-165
Vac (km/h)	35
V _{dc} (km/h)	40
ar	1.867
b _r	0.0359
<i>c</i> _r	0.000745

as shown in (5), the power supplied to the train from the catenary (P_{line}) is defined as the sum of the vehicle's propulsive power (P_t) and the power consumed by onboard auxiliary devices (P_{aux}). Alternatively, it can be defined as the product of the current (I_{line}) and voltage (V_{line}) in the catenary. Both the line voltage and current are measured while the vehicle is in motion at a sample time.

$$P_{line} = P_t + P_{aux} = V_{line}I_{line}$$
(5)

As stipulated in (6), the power required to propel the train (P_t) is defined as the product of the traction force (F_t) and the train speed (v) [26]. Here, v represents the speed of the train (m/s).

$$P_t = F_t v \tag{6}$$

Consequently, the total power consumed by a train can be defined as the sum of the power consumed for propulsion and the power consumed by the train's electrical auxiliaries, as outlined in the following equation:

$$P_{line} = F_t v + P_{aux} \tag{7}$$

where the total tractive force required to drive the train can be defined as the force necessary to accelerate the train and the external force applied from outside the train, as shown in (8).

$$F_t = ma + \left(F_r + F_g + F_x\right) = \frac{P_{line} - P_{aux}}{v} \tag{8}$$

Therefore, the connection between the total power consumed and train traction can be encapsulated as follows:

$$P_{line} = (ma + F_r + F_g + F_x) v + P_{aux}.$$
 (9)

In this study, the regenerative power produced during the braking phase was excluded from the total energy calculation. Our power computation concentrated solely on direct power consumption, as the usage of regenerative energy is subject to the allowable condition of the catenary line voltage, which is influenced by other train accelerations on the track and the state of regenerative braking.

C. ERROR CORRECTION FOR TRAIN SYSTEM MODELING

To evaluate the effectiveness of optimizing the operational speed profile, it is essential to measure the variation in operational energy before and after optimization. The most accurate method to accomplish this involves applying it to an actual train control system. However, during the optimization algorithm development phase, it is vital to base the approach on the most precise train model and line environment model feasible. This practice facilitates the prediction of improved optimization performance and augments the accuracy of the optimization solution in the actual system. The development of the optimal speed profile using the DP method leverages the train model of the line and the model of various resistances encountered by the train. The fidelity of these models has a direct bearing on the precision of the optimal solution design. Despite diligent efforts, there are inherent deficiencies in the mathematical model that cannot be accurately represented, such as train modeling errors and the slope, curvature, and tunnel resistance at each position along the line.

In this study, a range of errors not accounted for in the mathematical model were treated as a single disturbance. We obtained the model errors by calculating the travel position using the actual measured power and the applied model. These errors were then incorporated into the modeling errors to augment the optimization algorithm. Here, resistance elements comprise running resistance, which fluctuates with speed regardless of the train's travel position, and gradient resistance and curve resistance, which are fixed according to the characteristics of the travel position on the line. Assuming a well-modeled running resistance mathematical model for the train speed, the residual resistance is generated based on the travel position. These resistance elements can be utilized to compensate for modeling errors by estimating the resistance model error by travel position using instantaneously measured power on the train and the applied mathematical model.

The forces that inhibit train operation function as intricate, multifaceted disturbances contingent upon the train speed or the intrinsic attributes of the operating route. Among the key resistance elements, the running resistance (F_r) oscillates with speed due to the configuration or area of the train's frontage. Conversely, other resistances, such as curvature resistance and gradient resistance, maintain stable values owing to the characteristics of the line, depending upon the train's travel position. As a result, the disturbance can be approximated using (10), which deploys the instantaneous power consumption and resistance model equations at each position along the vehicle's line. In this case, instantaneous power consumption is assessed by employing line voltage and current data captured at each sample time and travel position during train operation.

The power measured on the line is quantified as the summation of propulsion power (P_t) and auxiliary power (P_{aux}). Hence, the power measured when the train is stationary, i.e., when no propulsion power is consumed, is considered the auxiliary power (P_{aux}) in the train. Subsequently, the auxiliary power (P_{aux}) is subtracted from the total line measurement power (P_{line}), and the resulting value is utilized as propulsion power (P_t). Here, the auxiliary power embodies the energy consumed by supplementary electrical devices within the vehicle and is presumed to be a constant value that remains unchanged over time. This facilitates the generation of unmodeled resistance elements as data by travel position, which can be integrated into the system model according to the train's travel position. In this context, d denotes the travel position.

$$F_x(d) = F_t (ma + F_r + F_g)$$

= $\frac{P_t}{v} - (ma + F_r + F_g)$
= $\frac{P_{line} (d) - P_{aux}}{v} - (ma + F_r + F_g)$ (10)

Figs. 1 and 2 present a comparison of the actual power measured during the journey between stations with the power consumed by the train model over the same stretch, based on the actual train speed. This comparison is intended to illustrate the aptness of the train model and the actual train. The patterns of power usage before and after the addition of disturbances to the train model validate the calibrated model's congruence with the actual vehicle.

The actual energy consumption between each station in the test was compared with the energy consumption derived from the implementation of the train's mathematical model using measured train speed and power data. Table 2 exhibits the test results for the seven sections, corroborating that the discrepancy in total energy consumption for each station segment is restricted within a 1% threshold.

III. OPTIMIZATION OF OPERATION ENERGY ACROSS ALL OPERATING SECTIONS

A. DYNAMIC PROGRAMMING METHOD

This section describes the utilization of DP in generating a speed profile optimized for kinetic energy in an electric rail vehicle. DP is a computational technique deployed to resolve multi-stage decision problems. It is predicated on Bellman's Optimality Principle [12] and can deliver optimal solutions to exceedingly intricate problems. In this study, the DP method is applied in two stages to optimize the operation of the entire line, encompassing multiple stations.

In the initial stage, a speed profile for each station section is constituted to optimize the operation energy. An optimization cost function is employed to balance the operation energy and travel time in a multi-tier ratio, thereby enabling the simultaneous generation of multiple speed profiles with minimal energy consumption per unit of travel time [24].

In the subsequent stage, given a line composed of multiple stations extending from the initial to the terminal station and a target travel time to traverse the entire line, the DP method is put to use. This is based on the optimal time and energy data table gathered in the first stage for each station segment. This process culminates in the optimal combination of traveling speed profiles for each station section, adhering to the target time while consuming the minimal possible amount of energy.



FIGURE 2. Comparison between measured power and modeled power.



FIGURE 3. Comparison between real-world and calibrated model energy consumptions.

1) SPEED PROFILE OPTIMIZATION FOR INTER-STATION

Trains travel a fixed distance between predetermined stations on the track. In this scenario, DP is used to determine an energy-efficient speed profile given a specific target travel time. To implement DP, a two-dimensional space is constructed, representing the train's distance traveled and its speed. This space is discretized at regular intervals, with the distance between stations and the maximum allowable speed serving as reference axes. As illustrated in Fig. 4, each intersection point between the distance and speed axes in this discrete two-dimensional space is referred to as a node, while the lines interconnecting these nodes are termed edges. For each edge connecting two nodes, the energy and time required to transition between them are estimated. Consequently, the optimal solution involves identifying node points at each location and minimizing the total cost of connecting these nodes by considering the energy and time associated with traversing each connecting node as dictated by the cost function. This process unfolds during the journey from the departure station to the destination station, with each corresponding node representing the target speed for the train at its specific travel position.

1. Establishing discrete spaces: Design and initialization.

Section Index	Operation energy consumption of actual data (kWh)	Operation energy consumption of corrected model (kWh)	Energy consumption error (%)	Running time (s)	Section length (m)
Section 1	4.6981	4.7413	0.92	50	303.05
Section 2	22.2121	22.2643	0.24	138	1,622.60
Section 3	16.2751	16.3759	0.62	123	1,306.50
Section 4	24.8971	25.0167	0.48	157	2,077.77
Section 5	21.4932	21.5772	0.39	127	1,654.84
Section 6	10.6188	10.7082	0.84	100	806.00
Section 7	21.5365	21.6474	0.51	101	1117.00

TABLE 2. Model error.

(Section 1): Hanam Geomdansan to Hanam City Hall, (Section 2): Hanam City Hall to Hanam Pungsan, (Section 3): Hanam Pungsan to Misa, (Section 4): Misa to Gangil, (Section 5): Gangil to Sangil-dong, (Section 6): Sangil-dong to Godeok, (Section 7): Godeok to Myeongil

As illustrated in Fig. 4, train operations are represented within a two-dimensional space defined by distance and velocity. This space is discretized at regular intervals along each axis. The x-axis represents the distance, with each discrete step identified as a stage. Correspondingly, the y-axis represents the velocity, with each stage representing a state within the discrete space. The intervals for speed or distance can be selected contingent on the requisite level of precision, and the speed of the train at the departure and arrival locations is stipulated as zero.

2. Designing a cost function with multiple weights

The time and energy necessary to traverse an edge connecting two proximate nodes in the discrete space are calculated and recorded by taking into consideration the change in speed between the nodes, which is presumed to be constant over the given distance. Furthermore, each corresponding edge index retains the weighted aggregate cost derived from these two values. In this context, the number of weight decompositions is multiplied by the number of weighted sum costs. The average acceleration (a_m) between two adjacent nodes is calculated using the speed at each node and the distance between the nodes, as shown in (11). The average acceleration obtained is applied to the acceleration (a) to calculate the traction force in the aforedescribed model equation.

To compute the cost, we multiply the time (T_i) and energy (E_h) required to travel between two nodes. As demonstrated in (11), the travel time is determined by the average speed between nodes (v_m) and the distance between nodes (d_s) . Meanwhile, the energy consumption can be estimated using the force and speed necessary for movement between the nodes, as depicted in (13).

$$a_m = \frac{v_{k+1}^2 - v_k^2}{2d_s} \tag{11}$$

$$T_i = \frac{d_s}{v_m} \tag{12}$$

$$E_h = (P_t + P_{aux}) \cdot T_i = (F_t v_m + P_{aux}) \cdot T_i$$
(13)

After acquiring the time and energy values for the relevant edges, a cost function is implemented to determine the optimal conditions that satisfy both the time and energy requirements, as shown in (14). The cost function incorporates time and energy with weighting factors. In this scenario, the weight comprises an array of values ranging from 0 to 1, divided into multiple steps. Each weight represents a ratio of time and energy allocation, with 0 representing the minimal travel time and 1 representing the energy conservation. Consequently, by altering these weights within the range of 0 to 1, the sum of the travel time and energy consumption required to reach the destination will fluctuate correspondingly. Alternative travel times and suitable speed profiles can be derived by appropriately adjusting the weights.

$$C_{k,k+1} = \frac{E_h}{E_{max} - E_{min}}\alpha + \frac{T_i}{T_{max} - T_{min}}(1 - \alpha) \quad (14)$$

where $C_{k,k+1}$ denotes the weighted sum cost of energy and time, representing the cost at that particular edge with multiple weights applied. If the weights are provided as an array, the cost is also treated as an array value. The weights are defined as follows:

$$\alpha = [\alpha_1 \alpha_2 \cdots \alpha_N]^T, \, \alpha_1 = 0, \, \alpha_N = 1,$$

$$\alpha_{n+1} > \alpha_n, \, n = 1, 2, \cdots, N.$$
(15)

 E_{max} represents the maximum energy consumed during node travel, E_{min} is the minimum interval energy, T_{max} is the maximum time taken to traverse a node, and T_{min} is the minimum node travel time.

3. Dynamic programming method for parallel processing of multi-weighted cost

Once the path cost between each node is established, the optimization problem across N_d stages aims to identify a connectivity path between the nodes that minimizes the



FIGURE 4. Speed profile according to dynamic program algorithm.

cumulative sum of the path costs C_k at each stage, as follows:

$$\min\sum_{1}^{N_d} C_k.$$
 (16)

In the established two-dimensional planar network, each node-to-node path is assigned a specific cost, factoring in elements such as distance and speed. To ascertain the lowest cost at each node, we utilize (17) to resolve the optimization problem employing the DP method. At stage k, the minimum cost, represented as J^*_k , is recognized as the smallest aggregate of the path cost $(C_{k,k+1})$ between stages k and k+1 and the minimum cost (J^*_{k+1}) from stage k+1 to the ultimate stage N_d . Significantly, this computation commences at the destination node (N_d) and progresses towards the initial node (k = 1) to pinpoint the minimum cost (J^*_k) .

$$[J_k^{*1 \times N}(x_k), I_k^{*1 \times N}] = \min_{col} [C_{k,k+1}^{s \times N}(x_{k+1}) + J_{k+1}^{*s \times N}(x_{k+1})]$$
(17)

where J_k^* represents the minimum cost from stage k to the final stage. It is an array comprising N minimum values, each calculated according to a particular split weight, as depicted in (18). While I_k^* represents an array consisting of N position indices, each indicating the minimum value for each weight, as demonstrated in (18). The term min_{col} stands for the process of determining the minimum value for each column in the matrix.

$$J_{k}^{*1 \times N} = \left[j_{\alpha_{1}}^{*} j_{\alpha_{2}}^{*} \dots j_{\alpha_{N}}^{*} \right]$$
(18)

$$I_k^{*1 \times N} = \left[i_{\alpha_1}^* i_{\alpha_2}^* \dots i_{\alpha_N}^* \right]$$
(19)

$$C_{k,k+1}^{s \times N} = \left[c_{\alpha_1}^{s \times 1} c_{\alpha_2}^{s \times 1} \dots c_{\alpha_N}^{s \times 1} \right]$$
(20)

$$J_{k+1}^{*s \times N} = \left[j_{\alpha_1}^{*s \times 1} j_{\alpha_2}^{*s \times 1} \dots j_{\alpha_N}^{*s \times 1} \right]$$
(21)

The variable $C_{k,k+1}$ corresponds to the weighted cost of each edge between stages k and k+1. This variable is expressed as an *s*- by -*N* matrix, as illustrated in (20). Conversely, J^*_{k+1} denotes the matrix of optimal costs calculated from node k+1 to the final destination. It is an *s*-by-*N* matrix, each element of which aligns with multiple weights, as exhibited in (21). In this instance, *s* represents the number of states in each stage, equivalent to the number of nodes along the speed axis.

To determine the most cost-effective path, the process initiates at the arrival point and advances towards the departure point, recording the smallest node number at each stage until the departure point is reached. The calculation for obtaining the minimal node index is described below. It is important to note that the starting and terminating nodes are excluded from this process as their speed is set to zero.

Stage k: K-1 (the stage immediately preceding the end of the distance axis)

$$P_{k,s} = I_k^{*T}, s = 0, 1, \dots, S : state index$$
 (22)

Here, k is the stage index, s is the state index, and $I *_k^T$ the transpose matrix of the number of least-cost nodes obtained from (19).

Stage k:
$$1 \le k < K-1$$
,
 $P_{k+1, I_k^{*T}} = \left[P_{k+1} \left(i_{\alpha_1}^* \right) P_{k+1} \left(i_{\alpha_2}^* \right) \cdots P_{k+1} \left(i_{\alpha_N}^* \right) \right]$ (23)

Within stage k, the optimal node index (I_k) traces the pathway among the highest-quality paths (P_{k+1}) stored in stage k+1. The total number of optimal nodes (P_n) , associated with the optimal node (I_k) from stage k to stage k+1, can be retrieved from paths corresponding to the weight order, as demonstrated in (24). In this case, '{}' denotes the cell structure data, while (N,:) signifies all column elements in the N_{th} row.

$$P_{n} = \begin{bmatrix} P_{k+1,I_{k}^{*T}}\{1\}(1,:) \\ P_{k+1,I_{k}^{*T}}\{2\}(2,:) \\ \vdots \\ P_{k+1,I_{k}^{*T}}\{N\}(N,:) \end{bmatrix}$$
(24)

Here, P_n represents the path with the minimum cost in stage k+1 for each weight. Therefore, the optimal path in stage k is generated by appending the optimal path (P_n) of stage k+1 to the optimal node (I_k) of stage k, as expressed in (25). This process is carried out in succession, progressing towards the starting node.

$$P_{k,s} = [I_k^{*T} P_n] \tag{25}$$

Ultimately, the minimum cost path extending from stage 1 to the final stage is extracted from the following equation:

$$P_{1,1}$$
: optimal path. (26)

The optimal path, inclusive of the starting node, for each weight is then completed using the following equation:

$$P_{0,0} = [1^{N \times 1} P_{1,1}^{N \times dK}].$$
(27)

4. Generating the speed profile using the optimal node

Equation (27) represents the node number of the optimal path at each stage. Therefore, when applying equal intervals of speed, the optimal speed profile using the node number of the optimal path can be expressed as follows:

$$V = [P_{0,0} \times d_v - d_v] = [v_0^{N \times 1} v_1^{N \times} \cdots v_{dK}^{N \times 1}]$$
(28)

where $d_v = v_{k+1} - v_k$ is the speed decomposition interval.

 Determining time and energy for each speed profile by weight

At this stage, the time (T_i) and energy (E_h) corresponding to each node can be obtained by tracing the node numbers from the optimal path acquired in (27), commencing from the departure node. With this information, the total time and energy required to traverse each station section for different weights can be calculated, as shown in Table 3. The weights between 0 and 1 were divided into 30 intervals. Table 3 presents the travel time and energy consumption for each weight. Fig. 5 displays the optimal speed profile designed for the 30 weights across the 7 station sections, taking into account factors such as speed limits and line altitude.

2) OPERATION TIME SCHEDULE OPTIMIZATION

In the previous step, the travel time and speed profiles for each station section were optimized. The ensuing step involves selecting a speed profile between each station that achieves the targeted travel time for the entire line, spanning from the departure station to the terminal station, while minimizing energy consumption. This is achieved by utilizing the data table of energy consumption predicated on travel time, as presented in Table 3.

1. Designing the discrete space: allocating costs among nodes

The optimal travel time and energy for each station section were determined based on the weight splits in the cost function, as exhibited in Table 3. To find the energy path that is optimal based on travel time for all operating sections, a network structure of nodes needs to be conceived, as illustrated in Fig. 6. The horizontal axis represents the stages, while the vertical axis represents the states. Stations are sequentially allocated in the odd stages of the horizontal axis, and the travel time (t) and energy consumption (e) for each segment are assigned to the states along the vertical axis, corresponding to each section from Table 3. This network structure allows for the optimization of the energy path based on travel time for the entire system.

To preserve the integrity of node connections, a dummy node is introduced between station nodes at even stages. The left edge of this dummy node is assigned zero time and zero energy cost to prevent any impact on the overall node connection cost. In Fig. 6, the right edge of the dummy node is associated with each state and signifies the designed duration and energy consumption for traveling between stations. With 30 split weights and 7 sections in this study, a total of 15 stages and 30 states are included, comprising 8 dummy nodes.

Each edge connecting the nodes records the energy consumption for each travel time, along with the combined cost that incorporates the weighted trade-off between travel time and energy consumption. To parallelize the cost function, the mixed cost is calculated using (29) at this stage. The mixed cost recorded at each edge corresponds to an array representing the split weights applied to the problem.

$$W_{k,k+1} = \frac{S_h}{S_{max} - S_{min}}\beta + \frac{T_{si}}{T_{smax} - T_{smin}} (1 - \beta) \quad (29)$$

where W_k denotes the cost at a particular edge, with multiple weights dictating the ratio between time and energy. If weights are employed as an array, the cost is also treated accordingly as an array value. The weights are defined as in (30).

$$\beta = [\beta_1 \beta_2 \cdots \beta_N]^T, \, \beta_1 = 0, \, \beta_N = 1, \, \beta_{n+1} > \beta_n, \\ n = 1, 2, \cdots, N$$
(30)

 S_{max} represents the maximum node travel energy, S_{min} denotes the minimum interval energy, T_{smax} signifies the maximum node travel time, and T_{smin} represents the minimum node travel time.

Generating the minimum cost and optimal path from each node in reverse: With the edge costs determined for the network connecting the nodes in the two-dimensional plane, the same DP method can be applied by following the steps from (17) to (27).

As shown in (31), the final path obtained from this process will include a dummy node $(1^{N_{x1}})$ for each odd stage.

$$P_{0,0} = [1^{N \times 1} P_2^{N \times 1} 1^{N \times 1} P_4^{N \times 1} 1^{N \times 1} \cdots P_{14}^{N \times 1} 1^{N \times 1}] \quad (31)$$

As a result, the final optimal path for each section, after removing the dummy nodes, can be expressed as follows:

$$P_{final} = [P_2^{N \times 1} P_4^{N \times 1} P_6^{N \times 1} \cdots P_{14}^{N \times 1}]$$
(32)

where N represents the number of split weights.

Table 4 presents the results of the paths identified in this study. The findings indicate the interval velocity profile contexts for each cost function weight.

IV. RESULTS AND DISCUSSION

In this study, a dynamic programming method was introduced to optimize energy consumption throughout all the operational segments. The operational optimization was conducted for seven segments of a station line, and the optimization performance was benchmarked against the actual energy consumption data. Simulations were performed using both real-world data and a modified model within a MATLAB environment. The objective was to determine the optimal travel time and speed profiles between each station while adhering to the same constraints and total travel time as that of the current line. During the actual operation, it took 794.5 s to traverse the seven station sections. A two-stage travel energy optimization method was employed to achieve operational optimization, utilizing 801 parallel split weights. This method yielded the total travel time, energy consumption, and selection profile indices for each station section, as shown in Table 4. The operational optimization helped in achieving a travel time range between 648 and 1739 s, while considering the station dwell time and operational buffer time.

Among the results, index 562 conformed to the actual travel time of 794 s, demonstrating a travel time condition of

	Secti	on 1	Secti	on 2	Secti	on 3	Secti	on 4	Secti	on 5	Secti	on 6	Secti	on 7
Weight index	t(s)	kWh	t(s)	kWh	t(s)	kWh	t(s)	kWh	t(s)	kWh	t(s)	kWh	t(s)	kWh
1	42.66	6.91	111.82	30.55	94.73	24.85	128.75	35.15	107.40	28.93	75.56	16.42	87.35	25.58
2	42.66	6.91	111.87	29.92	94.81	23.91	128.78	34.72	107.40	28.93	75.56	16.42	87.35	25.58
3	42.66	6.91	112.04	29.20	94.88	23.68	128.95	34.05	107.41	28.89	75.63	16.13	87.35	25.58
:	:		:	:	:	:	:	:	:	:	:	:	:	:
13	46.74	4.90	122.00	23.96	112.41	16.06	155.57	20.36	132.37	16.29	83.43	12.21	100.41	18.66
14	47.45	4.73	127.38	22.75	117.55	14.92	161.19	19.11	138.17	15.02	85.77	11.67	103.39	17.98
15	49.15	4.41	129.37	22.37	121.51	14.20	167.93	17.85	142.42	14.22	90.28	10.82	107.83	17.15
16	50.28	4.23	132.96	21.83	125.96	13.50	175.18	16.71	149.66	13.09	92.92	10.41	109.70	16.86
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
28	78.40	2.93	239.00	18.13	204.90	9.82	315.50	10.19	257.30	8.41	151.67	7.33	199.51	12.81
29	81.73	2.92	268.46	17.97	218.55	9.75	352.40	9.98	282.77	8.27	157.70	7.30	216.30	12.73
30	81.73	2.92	308.17	17.91	249.74	9.71	382.53	9.93	310.08	8.22	172.50	7.28	234.62	12.71

TABLE 3. Optimal travel time and energy by weight in each section.



FIGURE 5. Optimal speed profile by time-energy weight in each station section.

791.3 s. In the simulation results presented in Table 4, when the weight (β) is set to 0, it corresponds to the condition where the energy-related term in the cost function (Equation 28) becomes zero. Consequently, the cost function is reduced to a function that aims to minimize time, with only the time-related term on the right side remaining. Therefore, this condition represents a scenario where the entire operational segment is driven as fast as possible within the train's performance limits. The highest speed profile is selected among the generated profiles, utilizing the maximum acceleration and deceleration, resulting in the shortest travel time of 648.27 s and the maximum energy consumption. Conversely, when the weight (β) is set to 1, it represents a condition where only the energy-related term on the left side of the cost function (28) remains, making it a condition that seeks to minimize energy consumption. In this case, the travel time is the longest at 1739.38 s, and the speed profiles for each segment aim to minimize the energy consumption. In Table 4, under the condition of weight 1, profile 29 was selected out of 30 profiles in Section I. This was attributed to the fact that profiles 29 and 30 yielded the same results during the search process. The results ranging from 0 to 1 demonstrate that as the weight increases towards 1, the significance of the energy term increases, while the cost related to time decreases. This gradually shifts the emphasis towards energy minimization, resulting in longer travel times and reduced energy consumption. At this point, the profile number selected for each station segment can be easily identified. Notably, under the condition of weight 1, the selected travel condition minimizes energy consumption. However, beyond this point, increasing the vehicle's travel time does not reduce the energy consumption and instead causes it to increase again. This phenomenon is attributed to the energy consumption of the vehicle's onboard electrical devices, even when the vehicle is at a standstill.

Table 5 presents the results of optimizing the travel times between the station sections and the speed profiles

IEEEAccess



FIGURE 6. All station network for DP algorithm.

TABLE 4. Travel time and energy optimization by 801 split weights.

Index	Weights	Travel time (s)	Energy (kWh)	Sec 1 (Profile)	Sec 2 (Profile)	Sec 3 (Profile)	Sec 4 (Profile)	Sec 5 (Profile)	Sec 6 (Profile)	Sec 7 (Profile)
1	0	648.27	168.39	1	1	1	1	1	1	1
:	:	:	:	:	:	:	:	:	:	:
560	0.3311	784.59	105.49	15	15	14	14	14	14	14
561	0.3327	784.59	105.49	15	15	14	14	14	14	14
562	0.3343	791.33	104.22	15	15	14	15	14	14	14
563	0.3359	800.08	102.58	15	15	14	15	15	15	14
564	0.3375	804.52	101.75	15	15	14	15	15	15	15
:	:	:	:	:	:	•	:	:	:	:
801	1	1739.38	68.68	29	30	30	30	30	30	30

for each segment while satisfying the actual train travel time of 794 s. This optimization was implemented to compare the energy consumption and travel times for each station section. In sections II and 6, the energy consumption increased. However, based on the overall optimization of the station segments, it is evident that when completing the final journey, there is a reduction in the total energy consumption by approximately 14% when compared to the previous mode of operation. The results of this study were obtained through a comparison of the power usage between the actual operation on the real line, where the power consumption was measured, and travel simulations were conducted through the optimization method. Therefore, the amount of energy saved varies based on factors such as the number of stations or the conditions of the infrastructure

(grades, curves, etc.), as well as the characteristics of the train (propulsion, braking, regenerative capabilities, rolling resistance, etc.).

Fig. 7 is based on the recorded empirical data during the operation and route design. Here, a speed of 0 corresponds to the location of each of the eight stations. The figure depicts the operational records of vehicles operated by ATO and the speed profile results generated for each station, based on the applied weight conditions. The thick blue line represents the actual operating speed of the train, while the bold red line presents the optimized speed profiles obtained for each station under the condition of equal travel time to the actual operation. Since there are various gradient conditions between the station sections, this enables us to determine the locations at which deceleration and acceleration must be

Section (Profile)	Actually used energy (kWh)	Optimized energy (kWh)	Energy saving (%)	Measured travel time (s)	Optimized travel time (s)	Travel distance (m)
Sec 1(15)	4.6981	4.4101	-6.13	50.5	49.15	303.05
Sec 2(15)	22.2121	22.3715	0.72	137.75	129.37	1,622.60
Sec 3(14)	16.2751	14.9221	-8.31	122.25	117.55	1,306.50
Sec 4(15)	24.8971	17.8451	-28.32	155.5	167.93	2,077.77
Sec 5(14)	21.4932	15.0157	-30.14	127.25	138.17	1,654.84
Sec 6(14)	10.6188	11.6721	9.92	100	85.77	806.60
Sec 7(14)	21.5365	17.9828	-16.50	101.25	103.39	1117.0
Total	121.7309	104.2194	-14.39	794.5	791.33	8,888.36

TABLE 5. Performance comparison by optimization (energy reduction: 14.39%, time reduction: 3 s).



FIGURE 7. Optimal speed profile and actual train speed.

performed to obtain the optimal speed profile for each weight condition during the journey.

V. CONCLUSION AND SUMMARY

This study devised a dynamic programming method for optimizing the travel energy of the entire line of operating sections, along with individual station sections. It entailed the meticulous design of optimal travel speed profiles to cater to the minimum and maximum travel times of the operating sections, considering the operational prerequisites of several stations within each section. A two-stage dynamic programming method was conceived to maximize the energy efficiency of the entirety of the operating sections by concurrently processing multiple weighted cost functions. The first stage entailed designing an optimal speed profile for each station section, predicated on travel time. The second stage involved architecting a node network for the dynamic programming method to optimize the overall operation, utilizing the data table from the first stage's designed station region, complemented by an optimization procedure. Furthermore, the train system model required for the optimization design process was supplied, along with a strategy to compensate for model errors.

To offset model errors, the discrepancy between the measured power and the estimated power was identified as a component originating from modeling errors or external forces not encompassed by the model. This model error was addressed by creating data predicated on the travel position and interpolating it accordingly. The proposed method for correcting model errors can be uniformly applied to urban rail systems based on catenary lines on different routes. However, the level of model errors may vary depending on the estimation performance of external disturbances. To minimize modeling errors, it is advisable to accurately model and apply elements such as the running resistance based on varying driving speeds, tailored to the trains on the specific route.

To corroborate the proposed technique, operational data gathered during the actual operational process was employed. This data encompassed information such as the train's position on the line, line voltage, current, speed limit, vehicle weight, and grade at distinct time intervals. Initially, the model interpolation approach was leveraged to evaluate the discrepancy between the design model and the implemented system, substantiating that the total energy consumed by each segment remained within a 1% margin under identical operational conditions. Moreover, to verify the optimal travel speed profile design, a comparison was made of the energy consumption when utilizing the optimized speed profile under the same operational objectives and constraints as the real data. Ultimately, through operational optimization, a speed profile was designed that achieved approximately 14% energy savings compared to the prevailing operational energy usage.

This study was not aimed at generating real-time optimized speed profiles that consider the changing operational conditions of the train. Instead, it aims to provide the target travel times and target speed profiles during the initial timetable design phase. The design of real-time optimal speed profiles will be performed in future studies.

REFERENCES

- X. S. Yu, "Energy-saving optimization of train and energy evaluation in urban rail transit," M.S. thesis, Beijing Jiaotong Univ., Beijing, China, 2012.
- [2] Y. Ding, "Study on train movement calculation and operation optimization simulation system," Ph.D. thesis, Beijing Jiaotong Univ., Beijing, China, 2005.
- [3] Y. M. Wang, "Quantification analysis on the energy factors of the urban rail transit system," M.S. thesis, Beijing Jiaotong Univ., Beijing, China, 2011.
- [4] Q. Gu, T. Tang, and Y.-D. Song, "A survey on energy-saving operation of railway transportation systems," *Meas. Control*, vol. 43, no. 7, pp. 209–211, Sep. 2010, doi: 10.1177/002029401004300704.
- [5] I. A. Hansen and J. Pachl, *Railway Timetabling & Operations*. Hamburg, Germany: Eurailpress, 2014.
- [6] K. Ichikawa, "Application of optimization theory of bounded state variable problems to the optimization of train," *Bull. JSME*, vol. 11, no. 47, pp. 857–865, 1968.
- [7] L. S. Pontryagin, *The Mathematical Theory of Optimal Processes*. Hoboken, NJ, USA: Wiley, 1962.
- [8] P. Howlett, "Optimal strategies for the control of a train," *Automatica*, vol. 32, no. 4, pp. 519–532, Apr. 1996, doi: 10.1016/0005-1098(95) 00184-0.
- [9] E. Khmelnitsky, "On an optimal control problem of train operation," *IEEE Trans. Autom. Control*, vol. 45, no. 7, pp. 1257–1266, Jul. 2000, doi: 10.1109/9.867018.
- [10] R. F. Liu and I. M. Golovitcher, "Energy-efficient operation of rail vehicles," *Transp. Res. A, Policy Pract.*, vol. 37, no. 10, pp. 917–932, Dec. 2003, doi: 10.1016/j.tra.2003.07.001.
- [11] M. Miyatake and H. Ko, "Optimization of train speed profile for minimum energy consumption," *IEEJ Trans. Electr. Electron. Eng.*, vol. 5, no. 3, pp. 263–269, May 2010, doi: 10.1002/tee.20528.
- [12] R. Bellman and R. Kalaba, Dynamic Programming and Modern Control Theory. New York, NY, USA: Academic, 1964.
- [13] H. Ko and M. Miyatake, "A numerical method for optimal operating problem of a train considering DC power feeding system," *IEEJ Trans. Ind. Appl.*, vol. 126, no. 8, pp. 1104–1112, 2006, doi: 10.1541/ieejias.126.1104.
- [14] M. Miyatake and K. Matsuda, "Energy saving speed and charge/discharge control of a railway vehicle with on-board energy storage by means of an optimization model," *IEEJ Trans. Electr. Electron. Eng.*, vol. 4, no. 6, pp. 771–778, Nov. 2009, doi: 10.1002/tee.20479.

- s posivehicle
 If F. Mensing, "Optimal energy utilization in conventional, electric and hybrid vehicles and its application to eco-driving," Ph.D. dissertation, INSA Lyon, Villeurbanne, France, 2013.
 - [17] L. Pröhl, H. Aschemann, and R. Palacin, "The influence of operating strategies regarding an energy optimized driving style for electrically driven railway vehicles," *Energies*, vol. 14, no. 3, p. 583, Jan. 2021, doi: 10.3390/en14030583.

[15] F. Mensing, R. Trigui, and E. Bideaux, "Vehicle trajectory optimization for application in ECO-driving," in *Proc. IEEE Vehicle Power Propuls. Conf.*,

Sep. 2011, pp. 1-6, doi: 10.1109/VPPC.2011.6042993.

- [18] H. Ko, T. Koseki, and M. Miyatake, "Application of dynamic programming to the optimization of the running profile of a train," *WIT Trans. Built Environ.*, vol. 74, 2004.
- [19] V. Calderaro, V. Galdi, G. Graber, A. Piccolo, and D. Cogliano, "An algorithm to optimize speed profiles of the metro vehicles for minimizing energy consumption," in *Proc. Int. Symp. Power Electron., Electr. Drives, Autom. Motion.* Ischia, Italy: SPEEDAM, Jun. 2014, pp. 813–819, doi: 10.1109/SPEEDAM.2014.6872030.
- [20] Z. Bin, Y. Shijun, Z. Lanfang, L. Daming, and C. Yalan, "Energyefficient speed profile optimization for high-speed railway considering neutral sections," *IEEE Access*, vol. 9, pp. 25090–25100, 2021, doi: 10.1109/ACCESS.2021.3056387.
- [21] K. Kim and S. I.-J. Chien, "Optimal train operation for minimum energy consumption considering track alignment, speed limit, and schedule adherence," J. Transp. Eng., vol. 137, no. 9, pp. 665–674, 2011, doi: 10.1061/(ASCE)TE.1943-5436.0000246.
- [22] X. Li and H. K. Lo, "An energy-efficient scheduling and speed control approach for metro rail operations," *Transp. Res. B, Methodol.*, vol. 64, pp. 73–89, Jun. 2014, doi: 10.1016/j.trb.2014.03.006.
- [23] H. Tang, C. T. Dick, and X. Feng, "Improving regenerative energy receptivity in metro transit systems," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2534, no. 1, pp. 48–56, Jan. 2015, doi: 10.3141/2534-07.
- [24] Y. S. Byun and R. G. Jeong, "Optimization of driving speed of electric train using dynamic programming based on multi-weighted cost function," *Appl. Sci.*, vol. 12, no. 24, p. 12857, Dec. 2022, doi: 10.3390/app122412857.
- [25] W. J. Davis, "Tractive resistance of electric locomotives and cars," *Gen. Electr. Rev.*, vol. 29, no. 10, pp. 685–708, Oct. 1926.
- [26] H. Partab, Modern Electric Traction. Delhi, India: Natraj Offset Printer, Gagan Kapur for Dhanpat Rai& Co.(P) LTD Educational and Technical Publisher, 2010.



YEUN SUB BYUN received the Ph.D. degree in electronic engineering from Chungbuk National University, Cheongju, South Korea, in 2012. Currently, he is a Principal Researcher with the Korea Railroad Research Institute (KRRI), Uiwang, South Korea. He also holds the position of an Adjunct Professor of transportation system engineering with the University of Science and Technology. In the past, he was a Visiting Researcher with Tennessee State University. His

research interests include power savings, navigation, guidance, control, and localization in the context of intelligent vehicles.



RAG GYO JEONG received the B.S., M.S., and Ph.D. degrees in electrical engineering from Inha University, Incheon, South Korea, in 1991, 1999, and 2005, respectively. In 1995, he joined the Korea Railroad Research Institute (KRRI), Uiwang, South Korea, as a Senior Researcher. He currently holds the positions of a Chief Researcher and the General Director of the Smart Electrical and Signaling Division, KRRI. From 1990 to 1994, he was a Staff Engineer with

Hanjin Heavy Industries Company Ltd. His research areas of interests include autonomous train control systems, electric-powered transportation systems, personal rapid transit (PRT), and system engineering.