

Driver Behavior–aware Fuel Optimization Using a Digital Twin and Reinforcement Learning Approach for Open-pit Haul Trucks

Kusnawi Kusnawi*

Doctoral Program of Information Systems, Universitas Diponegoro, Semarang, Indonesia
Informatics, Universitas Amikom Yogyakarta, Yogyakarta, Indonesia
E-mail: kusnawi@students.undip.ac.id
ORCID iD: <https://orcid.org/0000-0002-2968-3744>

*Corresponding author

Mochamad Agung Wibowo

Postgraduate School Universitas Diponegoro, Semarang, Indonesia
E-mail: agung.wibowo@ft.undip.ac.id
ORCID iD: <https://orcid.org/0000-0002-5434-9107>

Ridwan Sanjaya

Department of Information Systems, Universitas Katolik Soegijapranata Semarang, Semarang, Indonesia
E-mail: ridwan@unika.ac.id
ORCID iD: <https://orcid.org/0000-0003-3865-5937>

Received: March 10, 2026; Revised: April 25, 2026; Accepted: June 15, 2026; Published: August 8, 2026

Abstract: Driver behavior, vehicle dynamics, and operating conditions strongly influence fuel consumption in open-pit mining operations. This study proposes a driver behavior-aware fuel optimization framework that integrates a digital twin architecture with reinforcement learning to improve fuel efficiency of heavy-duty haul trucks. The framework combines a data-driven vehicle dynamics surrogate, explicit modeling of driver behavior, and proximal policy optimization to enable safe and scalable policy learning within a realistic simulation environment. Historical telematics data were used to construct the digital twin and evaluate the learned policy under controlled operating conditions. Experimental results show that the reinforcement learning agent produces substantially smoother driving behavior, characterized by stable speed regulation and elimination of aggressive acceleration and braking events. Compared to historical operator driving, fuel consumption per kilometer, computed using rollout-level aggregation of cumulative fuel consumption and total traveled distance, was reduced from 4.45 L/km to 3.02 L/km, corresponding to a 32.05% improvement in fuel efficiency. The results demonstrate that explicitly modeling driver behavior within a digital twin-based reinforcement learning framework can yield significant fuel savings while maintaining realistic and interpretable driving strategies. The proposed approach provides a promising foundation for the development of decision-support and driver assistance systems aimed at improving energy efficiency in open-pit haulage operations.

Index Terms: Digital Twin, Reinforcement Learning, Fuel Efficiency, Driver Behavior Modeling, Open-pit Mining, Heavy-duty Haul Trucks, Proximal Policy Optimization.

1. Introduction

Fuel consumption in open-pit hauling operations is a major contributor to operational costs and environmental impacts. Large off-highway haul trucks operate under highly variable conditions, including load fluctuations, varying road grades, haul road characteristics, and diverse driving styles, which collectively lead to substantial variability and inefficiencies in energy use. Previous studies have reported that hauling activities alone account for a significant proportion of total energy consumption in surface mining operations, often exceeding one-third of total fuel use, making fuel optimization a critical objective for sustainable mining practices [1–3]. Conventional approaches to reducing fuel consumption in mining haulage systems primarily focus on improving vehicle design, increasing engine efficiency, and

optimizing haul roads. While these strategies have yielded measurable benefits, they typically require significant capital investment and infrastructure modifications. In contrast, driving behavior is increasingly recognized as a dominant and more flexible factor influencing fuel consumption, particularly in heavy-duty vehicles. Empirical studies have shown that speed management, acceleration smoothness, braking intensity, and anticipatory driving behavior can significantly impact fuel consumption even under similar operating conditions, without requiring changes to vehicle hardware [4–7].

Recent advances in onboard telematics systems have enabled the continuous collection of high-resolution operational data from mining fleets, including vehicle speed, payload, fuel consumption, road grade, and control inputs. These developments have created new opportunities for data-driven modeling and intelligent decision-support systems, aimed at enhancing operational efficiency. Machine learning-based fuel consumption models have been successfully applied to mining haul trucks and other heavy-duty vehicles, demonstrating improved prediction accuracy compared to traditional physics-based approaches under complex and nonlinear operating conditions [2, 8–11]. However, most existing studies still focus on predictive analytics and offline evaluation, providing limited support for adaptive or policy-based fuel optimization.

In parallel, the concept of the digital twin has emerged as a powerful paradigm for representing, analyzing, and optimizing complex industrial systems. Digital twins provide a virtual equivalent of a physical asset that is continuously informed by operational data, enabling simulation-based evaluation, performance analysis, and "what-if" scenario testing throughout the system's lifecycle [12–15]. In transportation and industrial applications, digital twins are increasingly being adopted for system monitoring and decision-making support. However, in the mining and transportation context, most digital twin implementations remain descriptive in nature and are rarely integrated with learning-based control or optimization frameworks [16–18].

Reinforcement learning (RL) has demonstrated strong potential for solving complex sequential decision-making problems in energy-efficient control and eco-driving applications. Policy gradient methods, such as Proximal Policy Optimization (PPO), have been shown to learn stable and fuel-efficient driving policies through iterative interactions with simulated environments [19–22].

Despite these advances, applications of reinforcement learning in mining transportation systems remain limited. Existing studies often rely on simplified or synthetic simulation environments and typically do not explicitly incorporate real-world driver behavior, even though driving style is a key determinant of fuel efficiency in real-world transportation operations [23–25]. This limitation has also been consistently identified in recent systematic literature reviews of adaptive machine learning applications for fuel optimization in mining haulage systems [26]. This gap highlights a key limitation of current research: the lack of an integrated framework that simultaneously combines real-world operational data, driver behavior modeling, and reinforcement learning within a digital twin environment for fuel optimization in open-pit mining haulage systems. Without such integration, reinforcement learning-based solutions risk generating control policies that are unrealistic, overly idealistic, or difficult to implement in a mining operational context.

To address this gap, this study proposes a driver behavior-aware fuel optimization framework that integrates a data-driven digital twin with a reinforcement learning agent. The proposed digital twin combines a surrogate vehicle dynamics model learned from real-world telematics data, an explicit representation of driver behavior, and an environmental model reflecting realistic operational constraints. Within this framework, a PPO-based reinforcement learning agent is trained to learn fuel-efficient driving policies through interaction with a virtual environment, rather than a physical system. By enabling a direct and fair comparison between historical operator driving behavior and reinforcement learning-derived policies under identical operating conditions, this study provides a practical and scalable approach to improving fuel efficiency in mining haulage operations. The proposed methodology can be applied to a wide range of heavy-duty vehicle systems and provides a foundation for future intelligent driver assistance and autonomous haulage applications.

2. Method

2.1. Study Area and Data Sources

A. Data Description

This study utilizes operational telematics data collected from an open-pit haulage system operating HD785 haul trucks. This dataset represents a real-life production haulage cycle performed under varying load and grade conditions and captures actual driver behavior during daily mining operations. The use of telematics data and on-board monitoring has become a standard and effective approach to analyzing fuel consumption and operational performance in mining haulage systems, as it provides direct access to high-resolution measurements that reflect real-world operating conditions [5,8–9]. Typical recorded variables include vehicle speed, instantaneous fuel consumption, payload mass, and road gradient. Together, these variables provide a comprehensive description of vehicle movement, energy consumption, and operating conditions throughout the transport cycle, forming a reliable basis for data-driven modeling and subsequent digital twin-based simulation and optimization [13].

B. Data Preprocessing and Feature Engineering

Raw telematics signals are pre-processed and transformed into behavior-aware features for digital twin modeling and reinforcement learning. All variables are synchronized to a fixed three-second sampling interval to ensure temporal

alignment across vehicle conditions, control actions, and environmental conditions. Fuel consumption and distance traveled are computed at each time step by numerically integrating the recorded fuel rate and vehicle speed signals. Fuel efficiency is evaluated using fuel consumption per unit distance, which is defined as the ratio of cumulative fuel consumption to total distance traveled. This definition can be formally stated as:

$$\text{Fuel}_{\text{km}} = \frac{\sum_{t=1}^T F_t}{\sum_{t=1}^T D_t} \quad (1)$$

where F_t denotes the fuel consumption (L) at time step t , and D_t represents the traveled distance (km) at time step t , and T represents the total number of time steps within a haulage segment. These metrics enable consistent comparisons across driving profiles and operating conditions in heavy-duty haulage systems [2,8]. Load data are converted into discrete load states representing unloaded and loaded operations, while road gradient is described as both a continuous variable (in degrees) and a categorical class (downhill, flat, uphill) to capture physical effects and operational context [9,16]. Driver behavior is explicitly encoded through normalized control signals corresponding to acceleration, braking, and retarder use, scaled to the unit range (0, 1) for numerical stability during reinforcement learning. Additional features capturing short-term speed variations and stopping-related behavior are included to reflect driving smoothness and control intensity, which are known to impact fuel consumption in heavy-duty vehicles [23,27]. The resulting feature set integrates vehicle motion, energy use, road conditions, and driver behavior into a unified state representation that supports fuel-efficient and operationally realistic policy learning in a digital twin environment.

2.2. Digital Twin Architecture

The proposed digital twin represents a virtual counterpart of the physical haul truck system by integrating vehicle dynamics, driver behavior, and environmental conditions within a unified simulation framework. The digital twin is constructed in a modular manner to ensure interpretability, extensibility, and compatibility with reinforcement learning-based control. Such modular digital twin architectures are widely adopted to support data-driven modeling, simulation, and optimization in complex industrial systems, including mining operations [15–16].

A. Vehicle Dynamics Model

The Vehicle Dynamics Model (VDM) represents the relationship between vehicle operating conditions, driver control inputs, and fuel consumption. In this study, the VDM is implemented as a data-driven surrogate model trained on historical telematics data collected from open-pit mining haul truck operations. This surrogate model predicts instantaneous fuel consumption as a function of vehicle speed, road gradient, load, driver control intensity, and the status of the energy-saving mode. A machine learning-based surrogate model is adopted instead of a purely physics-based model to capture nonlinear interactions between operational variables that are difficult to model analytically, particularly under varying load and terrain conditions. Similar data-driven approaches have been shown to provide accurate fuel consumption predictions for mining haul trucks and other heavy-duty vehicles when sufficient telematics data is available [2,8–9].

In a digital twin, the VDM serves as a transition model for fuel consumption, enabling rapid and stable fuel usage simulations during reinforcement learning applications without the need for computationally expensive physical simulations.

B. Driver Behavior Model

Driver behavior is represented in the digital twin as an explicit control layer that links reinforcement learning policy actions to operationally meaningful driver inputs. In heavy haulage operations, fuel consumption is significantly influenced by driving style, particularly through the regulation of the speed and intensity of acceleration, braking, and retarder use. Empirical studies consistently demonstrate that smoother control inputs and moderated longitudinal actions can lead to substantial improvements in fuel efficiency for heavy-duty and off-road vehicles [4,27]. Therefore, driver behavior in the proposed framework is modeled as a structured and controllable component, rather than as stochastic variability.

In the digital twin environment, driver behavior is operationalized through a behavior-aware actuation mechanism that maps continuous RL actions to normalized driver control signals. At each simulation step, the RL policy generates continuous throttle commands. Positive actions are translated into accelerator intensity, while negative actions are converted into braking or retarder intensity using conditional logic that reflects realistic haul truck operation, such as increased braking emphasis at higher speeds and prioritized retarder use on downhill segments. The resulting behavioral signals, such as acceleration, braking, and retarder use, are normalized to the range (0, 1) to ensure numerical stability during learning and consistent interaction with vehicle dynamics surrogates.

This behavior-aware formulation enables the digital twin to generate realistic driver behavior trajectories that are directly governed by learned RL policies while remaining constrained by physically and operationally meaningful control patterns. As a result, the emerging driving behavior aligns with established principles of eco-driving and energy-efficient control, where subtle speed regulation and moderate control intensity play a central role in reducing fuel consumption [23,27].

C. Environment Model

The Environment Model defines the operational context and physical constraints within which haul trucks operate within the digital twin. Its primary role is to ensure that the state transitions generated during reinforcement learning remain realistic, safe, and representative of actual open-pit hauling conditions. The Environment integrates road gradients, load conditions, speed limits, and fuel consumption constraints into the simulation loop, thereby restricting the policy search space to a region that is physically and operationally feasible. Vehicle speeds are constrained within a predefined operational range that reflects typical haul truck operating practices. A lower speed limit is imposed to prevent unrealistic stop-and-go behavior, while an upper speed limit ensures compliance with safety considerations and mine operation policies. These constraints are enforced at each simulation step by clipping the predicted vehicle speed, preventing the reinforcement learning agent from exploiting unphysical or unsafe acceleration patterns.

Fuel consumption is regulated using a fuel-per-distance constraint within the simulation environment. The raw fuel consumption estimates generated by the vehicle dynamics surrogate model are converted to fuel consumption per distance values and further constrained within a plausible operating range. This mechanism prevents extreme or numerically unstable fuel estimates that might arise from surrogate model prediction errors, particularly under conditions of low speeds or near-zero distances. The corrected fuel consumption is then used for both reward calculations and performance evaluation, ensuring consistency between the learning objective and the evaluation metric.

Episode termination conditions are defined based on the maximum distance travelled and the number of simulation steps. Distance-based termination ensures that each episode represents a complete transport segment rather than an arbitrary time span. In contrast, step-based termination prevents excessively long episodes that could bias the learning process. Together, these termination criteria enable fair and repeatable evaluation of different driving policies under identical operating conditions. By explicitly embedding operational constraints and termination logic into the environment model, the digital twin provides a controlled yet realistic simulation setting for reinforcement learning. Environment design that takes these constraints into account reduces the likelihood of learning unsafe or impractical driving behaviors and aligns learned policies with real-world transport operations, a key requirement for digital twin-based optimization frameworks in mining and transportation systems [16,28].

2.3. Reinforcement Learning Framework

A. Problem Formulation and State Space

The fuel optimization task is formulated as a Markov Decision Process (MDP), where the goal is to learn a driving policy that minimizes fuel consumption per unit distance while maintaining smooth, safe, and operationally feasible vehicle behavior. At each discrete decision step, the agent observes the current operating state of the haul truck and selects a control action that determines the vehicle's motion and subsequent fuel consumption through interaction with the digital twin environment. This formulation is consistent with established reinforcement learning approaches for continuous control problems in transportation and energy-efficient driving applications [22,29].

The state vector is designed to capture key variables that characterize the haul truck's instantaneous operating context and directly affect fuel consumption. Specifically, these states include the vehicle's current speed, road gradient, load state, and eco-mode state. Vehicle speed reflects the motion state and direct control objectives, road gradient captures terrain-induced load variations, load represents mass-dependent energy demand, and eco-mode state encodes the vehicle's inherent efficiency constraints. Together, these variables provide a concise yet adequate representation of the operational context for fuel-related decision-making in transport operations.

Fuel consumption is not explicitly included in the state vector. Instead, it is implicitly determined by the interaction between the agent's control actions, state variables, and a data-driven surrogate for vehicle dynamics embedded in the environment. This design avoids oversimplifying the learning problem while ensuring that fuel-related outcomes remain causally linked to controllable driving decisions, a common practice in reinforcement learning-based energy optimization frameworks [30–31].

B. Action Space

The action space is defined as a continuous control variable representing throttle modulation, which directly regulates vehicle acceleration and deceleration. Actions are constrained within a fixed range and mapped to vehicle speed changes subject to operational and physical constraints imposed by the digital twin environment. A continuous action space is crucial for heavy-duty vehicle control, as discrete-action formulations tend to lead to abrupt speed changes, oscillatory behavior, and reduced ride smoothness, which are inconsistent with realistic haul truck operation.

In the proposed framework, a reinforcement learning agent issues a single continuous throttle command at each decision-making step. Positive action values correspond to acceleration requests, while negative values represent deceleration commands. These commands are then translated into acceleration, braking, or retarder control signals through the driver behavior model, ensuring that speed transitions remain smooth and operationally feasible. This formulation enables the agent to learn more detailed speed regulation strategies that balance fuel efficiency and ride smoothness.

This type of continuous control formulation is widely adopted in studies of eco-driving and reinforcement learning-based vehicle energy management, where fine-grained longitudinal control is a key determinant of energy efficiency and

policy stability [22,30]. By restricting the action space to a single interpretable throttle dimension, the proposed approach maintains control realism while avoiding unnecessary policy complexity.

C. Reward Function

The reward function is designed to promote fuel-efficient and smooth driving behavior while respecting operational feasibility constraints inherent to open-pit haulage operations. Fuel consumption per unit distance is penalized to directly encourage energy-efficient driving behavior, whereas excessive speed fluctuations are discouraged to avoid aggressive driving patterns that may lead to unnecessary fuel usage, increased mechanical stress, or reduced drivability. At each decision step t , the reward is defined as:

$$R_t = -\alpha \text{Fuel}_{\text{km}} - \beta |\Delta v_t| \quad (2)$$

where Fuel_{km} denotes the instantaneous fuel consumption per kilometer estimated within the digital twin environment, Δv_t represents the change in vehicle speed between two consecutive time steps, and α and β are weighting coefficients controlling the trade-off between fuel efficiency and speed smoothness. The weighting coefficients are set to $\alpha = 1.0$ for the fuel consumption penalty and $\beta = 0.05$ for the speed smoothness penalty, thereby prioritizing fuel efficiency and discouraging abrupt speed variations during policy learning.

This reward formulation aligns with established reinforcement learning-based eco-driving and vehicle energy optimization studies, where penalizing energy usage and speed variability has been shown to yield stable and fuel-efficient control policies [22,32–33]. By explicitly incorporating speed variation penalties, the reward discourages oscillatory control behavior that is unrealistic for heavy-duty haul trucks operating under load. In addition to the primary reward components, operational penalties are enforced within the environment to discourage infeasible or unsafe behaviors, such as overspeeding under downhill conditions or violating predefined speed limits. These penalties are implemented as hard or soft constraints within the digital twin rather than as standalone reward terms, ensuring that the learned policy remains within realistic operational boundaries during training.

Beyond reward-based penalties, safety-related operational constraints are integrated into the digital twin environment to prevent infeasible driving behaviors. These constraints consist of predefined speed limits and feasibility checks that restrict the agent from generating unsafe actions, especially on downhill segments or under variable traction conditions. Enforcing these constraints during both training and evaluation ensures that the learned policy adheres to realistic operational safety boundaries.

D. PPO Configuration

Policy optimization is performed using PPO, a policy-gradient reinforcement learning algorithm designed to ensure stable learning through constrained policy updates [19]. By limiting the deviation between successive policy updates via a clipped surrogate objective, PPO effectively balances exploration and stability, making it particularly suitable for continuous control problems with noisy or complex dynamics. In the proposed framework, PPO is used to train a stochastic policy that maps the observed digital twin state variables to continuous throttle actions. Policy learning is carried out through repeated rollout simulations within the digital twin environment, where the agent iteratively updates its policy parameters based on cumulative rewards obtained from simulated haulage trajectories. This simulation-based training enables safe and efficient policy optimization without direct interaction with physical mining equipment.

PPO has been widely adopted in vehicle control and energy-efficient driving applications due to its robustness and ability to learn smooth control strategies under operational constraints. Prior studies have demonstrated the effectiveness of PPO-based approaches for fuel economy optimization, eco-driving, and energy management in both conventional and electrified vehicle systems [22,30–31]. These properties make PPO a suitable choice for learning stable and fuel-efficient driving policies within the proposed behavior-aware digital twin framework.

The reinforcement learning agent was trained using the Proximal Policy Optimization (PPO) algorithm as implemented in the Stable-Baselines3 library. Unless otherwise specified, the training configuration adhered to the default hyperparameter settings of Stable-Baselines3: a learning rate of 3×10^{-4} , a discount factor (γ) of 0.99, a clipping parameter of 0.2, a batch size of 64, and 2048 environment steps per policy update. The policy network employed a multilayer perceptron (MLP) architecture. Training was conducted for 100,000 interaction steps within the digital twin environment.

During training, the PPO agent maintains a stochastic policy that enables exploration of the action space through probabilistic action sampling. For evaluation purposes, however, the trained policy is executed with deterministic action selection to ensure consistent, reproducible performance comparisons across evaluation episodes.

Experiments were conducted using the Google Colab computational environment with GPU acceleration (NVIDIA A100). The surrogate vehicle dynamics model was trained on approximately 1.93 million telematics samples, and the PPO policy was trained for 100,000 interaction steps within the digital twin simulation environment. Although the dataset was relatively large, the training process required only moderate computational resources and was completed within practical runtime limits. These results indicate that the proposed digital twin-based reinforcement learning framework is computationally feasible for large-scale fleet optimization scenarios.

2.4. Training and Simulation

Reinforcement learning training is conducted entirely within the proposed digital twin environment, implemented as a custom Gymnasium-compatible simulation (HaulTruckEnv). This environment includes a data-driven vehicle dynamics surrogate model for fuel prediction, a behavior mapping component that converts policy actions into driver control signals such as acceleration, braking, or deceleration intensity, and operational constraints on speed and fuel consumption per unit distance to ensure realistic haul truck behavior.

Policy learning is performed using PPO, a stable policy gradient algorithm widely used for continuous control tasks [19]. The agent was trained using a multilayer perceptron policy (MlpPolicy) within a vectorized environment wrapper (DummyVecEnv) to ensure compatibility with the Stable-Baselines3 training pipeline [34]. Training is conducted over a fixed number of interactions totaling 100,000 time steps, during which the agent collects rollout trajectories, updates the policy using the PPO clipped surrogate objective, and iteratively improves the expected return through repeated policy updates.

Each simulation episode begins by sampling the initial operating context from historical telematics records, including initial speed, road gradient, payload, and energy-saving mode status. The agent then selects a continuous throttle action at each step, which translates into a speed change subject to operational speed constraints. Episode termination is triggered when a predetermined range limit is reached or the maximum number of steps is exceeded, ensuring that rollout trajectories are finite and comparable across experiments.

2.5. Evaluation and Validation

After training, the learned policy is evaluated using deterministic action selection to ensure repeatable performance assessment across multiple simulations. Policy evaluation is conducted over multiple episodes, each initialized with a randomly selected operating context. During each episode, total fuel consumption and total distance traveled are accumulated from the step-level simulation output, allowing for the calculation of fuel consumption per kilometer. Additionally, descriptive driving statistics, including average speed, maximum speed, and average slope, are calculated to characterize the operational regime and stability of the learned driving behavior.

The primary evaluation metric is fuel consumption per unit distance (L/km), calculated as the ratio of total fuel consumed to total distance traveled during each simulation. Aggregate results are reported as the average performance across all evaluation episodes. This simulation-based evaluation approach is consistent with standard reinforcement learning evaluation protocols in continuous control settings [19,29].

3. Results

3.1. Overview of Experimental Setup and Evaluation Scope

All experiments were conducted entirely within the proposed digital twin environment using a data-driven vehicle dynamics surrogate model. The evaluation compared historical operator driving behavior with driving policies generated by a reinforcement learning (RL) agent trained using Proximal Policy Optimization (PPO). To ensure a fair comparison, environmental conditions, including load and road gradient, were kept constant across both operator-based and RL-based simulations. The results presented in this section focus exclusively on the quantitative results obtained from the simulation runs and model outputs.

3.2. Digital Twin Surrogate Model Validation

The performance of the digital twin vehicle dynamics surrogate was evaluated by comparing predicted fuel consumption rates with actual telematics observations. The surrogate model was trained and tested on a large-scale dataset comprising approximately 1.93 million samples, using seven input features that represent vehicle motion, operating conditions, and driver behavior.

Table 1. Vehicle dynamics surrogate model performance

Metric	Training set	Test set
RMSE (L/h)	31.693	34.418
R ²	0.813	0.780

Table 1 reports the performance of the vehicle dynamics surrogate model using standard regression metrics. The model achieved a root mean square error (RMSE) of 31.693 L/h on the training dataset and 34.418 L/h on the test dataset, with corresponding coefficients of determination (R²) of 0.813 and 0.780, respectively. The close agreement between training and test performance indicates consistent predictive behavior across unseen operating conditions.

To further illustrate predictive accuracy, Fig. 1 presents a scatter plot comparing predicted and actual fuel consumption rates for the test dataset. The majority of data points are distributed around the diagonal reference line, indicating a strong correspondence between model predictions and observed fuel consumption values over a wide operating range.

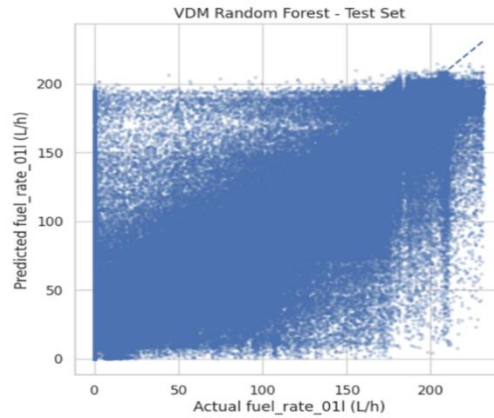


Fig.1. Random Forest VDM Performance: Actual vs Predicted Fuel Rate

Overall, these results indicate that the vehicle dynamics surrogate model accurately captures the nonlinear relationship between operating conditions and fuel consumption, supporting its use as a transition model for subsequent reinforcement learning experiments.

3.3. Reinforcement Learning Training Performance

The reinforcement learning agent was trained within the proposed digital twin environment using PPO for a fixed number of interaction steps. The training process remained numerically stable under the defined state, action, and reward formulation, with no evidence of divergence or collapse of the learned policy.

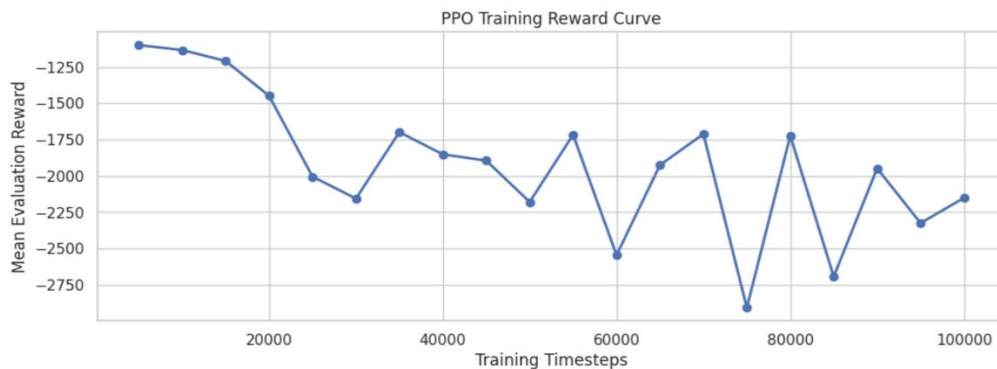


Fig.2. PPO training reward trajectory during learning

Fig. 2 presents the evolution of the mean training reward. Due to the penalty-based reward structure adopted in this study, reward values remain negative and exhibit non-monotonic fluctuations. Nevertheless, the reward trajectory remains bounded throughout training and does not show signs of instability or uncontrolled oscillations. These observations indicate that PPO is able to maintain stable policy updates within the constrained digital twin environment. The effectiveness of the learned policy is therefore assessed through rollout-based evaluations of driving behavior and fuel consumption performance, as presented in the subsequent sections.

3.4. Speed Profile Comparison: Operator vs RL

A. Time-Series Speed Profiles

Fig. 3 presents a comparison of time-series speed profiles between historical operator driving behavior and the speed trajectory generated by the RL agent under identical operating conditions. The operator speed profile exhibits frequent and abrupt fluctuations, including rapid accelerations and decelerations, as well as intermittent low-speed intervals throughout the haul cycle.

In contrast, the RL-generated speed profile maintains a comparatively smooth and stable velocity over time. The reduced oscillation in the RL-generated speed trajectory indicates more consistent speed regulation and fewer aggressive control actions. This difference highlights distinct driving patterns produced by the two strategies, with the RL policy favoring steady-state operation within the prescribed operational speed bounds.



Fig.3. Time-series speed profiles of historical operator driving and the RL agent under identical operating conditions

B. Speed Distribution Analysis

To further quantify the differences in speed regulation patterns, the vehicle speed distributions for historical operator driving and the RL-generated policy were analyzed. Fig. 4 presents the speed distributions for both driving strategies. The operator's speed distribution spans a wide operating range, reflecting substantial variability throughout the transport cycle. In contrast, the RL agent exhibits a highly concentrated speed distribution centered around a single operating speed, indicating consistent and tightly regulated speed control over time.

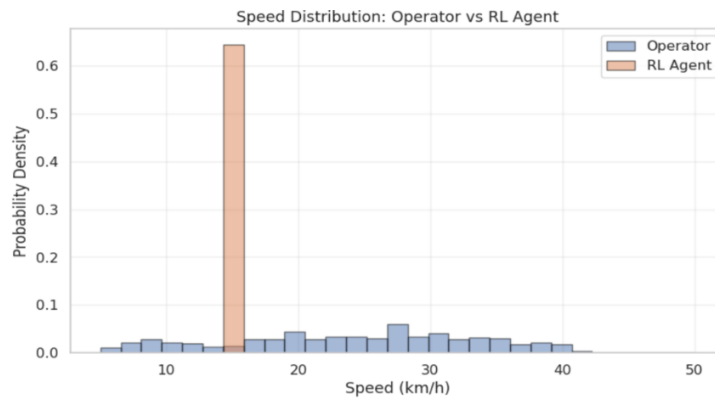


Fig.4. Speed distribution of historical operator driving and the RL agent

The cumulative characteristics of these distributions are illustrated in Fig. 5, which shows the cumulative distribution function (CDF) of vehicle speeds for the operator and the RL agent. The operator's CDF increases gradually, confirming significant dispersion in speed selection during operation. In comparison, the RL agent's CDF displays a nearly vertical transition, indicating that most operating time is spent within a narrow speed range with minimal variability.

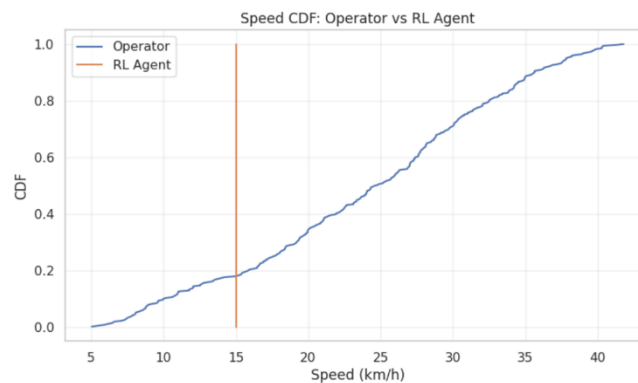


Fig.5. Cumulative distribution functions (CDF) of speed for historical operator driving and the RL agent

Quantitative statistics supporting these observations are summarized in Table 2, which reports key speed distribution metrics. Historical driving data showed a mean speed of 24.03 km/h with a standard deviation of 9.11 km/h, and speed values ranged from approximately 8.10 km/h (5th percentile) to 37.85 km/h (95th percentile). In contrast, the RL agent maintained a constant operating speed of 15.00 km/h throughout the evaluated episodes, resulting in zero variance and a fully concentrated interquartile range. These distribution-level statistics quantitatively confirm the reduction in speed variability and highly consistent speed regulation achieved by the RL-generated driving policy.

Table 2. Statistical summary of speed distribution metrics for historical operator driving and the RL agent

Metric	Operator	RL Agent
Mean speed (km/h)	24.03	15.00
Standard deviation (km/h)	9.11	0.00
5th–95th percentile (km/h)	8.10 – 37.85	15 – 15
Interquartile range (IQR)	13.14	0.00

As indicated by the statistics in Table 2, the RL agent maintains a nearly constant operating speed across the evaluated episodes. This constant-speed behavior reflects the relatively stable operating conditions represented in the digital twin simulation. In the evaluated haul segment, road gradient, payload state, and operational constraints remained largely consistent, leading the RL agent to adopt a steady-state speed regulation strategy that minimizes unnecessary acceleration and braking. Such behavior aligns with established eco-driving principles for heavy-duty vehicles, where maintaining a stable speed reduces transient engine load fluctuations and improves fuel efficiency. In real-world open-pit mining operations, additional factors, such as traffic interactions, queuing at loading or dumping points, and abrupt terrain variations, may introduce greater speed variability. Incorporating these dynamics into the digital twin environment represents an important direction for future research to enhance policy realism further.

3.5. Fuel Consumption Performance

Fuel consumption performance is evaluated using aggregated fuel-per-distance metrics calculated over the full simulation, complemented by time-series visualizations to assess the temporal consistency of fuel use under identical operating conditions. Performance comparisons are made between the operator's past driving behavior and the policy generated by the reinforcement learning agent in the same digital twin simulation environment. Fuel efficiency is quantified as fuel consumption per unit distance, calculated from cumulative fuel use and total distance traveled during each evaluation episode.

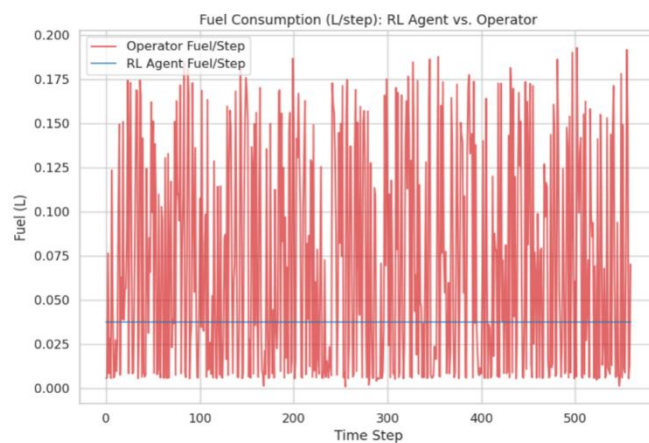


Fig.6. Time-series comparison of fuel consumption per step (L/step) between historical operator driving and the RL agent under identical operating conditions

Fig. 6 presents the time-series comparison of fuel consumption per step between the operator and the RL agent. The operator's fuel consumption profile exhibits pronounced fluctuations across simulation time steps, reflecting frequent changes in speed and throttle input during the haul cycle. In contrast, the RL agent maintains a nearly constant fuel consumption per step, consistent with its stable speed regulation behavior observed in the preceding analyses.

Quantitatively, the historical operator recorded an average fuel consumption rate of 4.45 L/km. In comparison, the RL agent consumed 3.02 L/km under the same operating conditions across multiple evaluation episodes, corresponding to an average reduction of approximately 32.05% in fuel consumption per kilometer relative to the operator benchmark. These values are obtained from aggregated fuel and distance metrics over the evaluated haul cycle. Figure 6 further illustrates the temporal evolution of fuel consumption per step, highlighting the substantially reduced variability achieved by the RL-generated driving policy.

3.6. Driving Behavior Metrics Comparison

Driving behavior metrics were calculated to quantitatively assess the differences in smoothness and driving control between the historical operator's driving and the RL-generated policy. The evaluated metrics included speed variability, acceleration stability, stopping behavior, and the occurrence of harsh acceleration and braking events.

Table 3. Driving behavior metrics comparison between historical operator driving and the RL agent

Metric	Operator	RL Agent
Number of samples	600	560
Mean speed (km/h)	13.71	15.00
Speed standard deviation (km/h)	13.59	0.00
Mean absolute speed change (km/h per step)	15.30	0.00
Standard deviation of speed change (km/h per step)	19.43	0.00
Mean absolute acceleration (m/s ²)	1.42	0.00
Stop ratio (%)	43.67	0.00
Harsh acceleration ratio (%)	41.40	0.00
Harsh braking ratio (%)	39.23	0.00

As summarized in Table 3, the operator's past driving behavior exhibited substantial variability in speed and acceleration. The operator exhibited a standard deviation of 13.59 km/h in speed and an average absolute speed change of 15.30 km/h per step, indicating frequent fluctuations in speed. In contrast, the RL agent maintained a constant operating speed of 15.00 km/h, resulting in zero speed variance and no changes in speed between steps.

Acceleration-related metrics further highlight this difference. The operator's driving profile exhibited an average absolute acceleration of 1.42 m/s², while the RL agent produced no acceleration or deceleration events during the evaluated episodes. Furthermore, the operator exhibited a high proportion of stopping events (43.67% of time steps), as well as frequent sudden accelerations (41.40%) and sudden braking events (39.23%). These events were largely absent from the RL-generated driving policy. This comparison confirms that the RL agent learned a very stable and controlled driving behavior in the digital twin environment. The observed reduction in variability in speed, acceleration magnitude, and sudden driving events is consistent with the smoother speed profiles and lower fuel consumption reported in the previous section.

The zero values observed for several driving behavior metrics in Table 3 reflect the steady-state operating regime learned by the RL agent. Because the learned policy maintains a stable speed trajectory throughout the simulated haul segment, unnecessary acceleration, braking, and stop-and-go behavior are largely avoided. Such smooth longitudinal control is consistent with eco-driving principles, where minimizing abrupt speed changes contributes to improved fuel efficiency. In real-world haulage operations, however, additional operational disturbances such as traffic interactions, loading delays, and terrain variations may introduce greater variability in these metrics.

Overall, the experimental evaluations demonstrate that the proposed digital twin-based reinforcement learning framework yields consistent and measurable performance improvements. The vehicle dynamics surrogate model achieves sufficient prediction accuracy to support reinforcement learning-based simulations, and the PPO agent demonstrates stable training convergence in the digital twin environment. Compared to historical operator driving, the RL-generated policy produces a smoother speed profile with significantly lower variability and lower fuel consumption. Under identical operating conditions, fuel consumption per kilometer is reduced from 4.45 L/km for operator driving to 3.02 L/km for the RL agent, which equates to a 32.05% improvement. Driving behavior metrics further confirm the improved smoothness and stability of the RL-generated driving policy.

3.7. Interpretation of Fuel Efficiency Improvement

The experimental results demonstrate that the proposed behavior-aware digital twin and reinforcement learning framework can achieve substantial reductions in fuel consumption compared to historical operator driving under identical operating conditions. The observed improvement is primarily attributable to the learned policy's ability to maintain a stable and near-constant operating speed, thereby reducing unnecessary acceleration and deceleration events. The results suggest that digital twin-based, behavior-aware reinforcement learning constitutes a promising approach in heavy-duty vehicles, where longitudinal control smoothness directly affects engine load and energy losses [4,35].

These findings are consistent with prior empirical studies, which indicate that smoother speed regulation and moderated control intensity can yield significant fuel savings without requiring changes to vehicle hardware or infrastructure [4,27,35]. In the context of open-pit haulage operations, where vehicles operate under repetitive cycles and relatively predictable routes, the influence of driving behavior on fuel consumption becomes even more pronounced [1]. The results therefore reinforce the notion that driver behavior optimization represents a high-impact and operationally feasible lever for improving energy efficiency in mining transportation systems.

While the RL-generated policy converges toward a near-constant operating speed under the evaluated conditions, the reinforcement learning framework does not enforce a predefined speed heuristic. Rather, the RL agent develops a

control policy through interactions with the digital twin environment, utilizing observed operational states such as vehicle speed, road gradient, and load conditions. Consequently, the observed steady-state behavior emerges as an optimal strategy under the relatively stable operating conditions modeled in the simulation, rather than being explicitly defined as a deterministic rule.

3.8. Role of Explicit Driver Behavior Modeling

The main contribution of this study lies in the explicit representation of driver behavior within a digital twin framework. Unlike conventional approaches that treat driver behavior as stochastic noise or implicitly embed it in vehicle dynamics models, the proposed framework models driver behavior as a controllable and learnable component that directly mediates the relationship between control actions and fuel consumption.

By mapping reinforcement learning actions to interpretable driver control signals such as acceleration, braking, and retarder usage, the digital twin enables the learning process to internalize eco-driving principles in a structured manner. This behavior-aware formulation allows the agent to discover fuel-efficient driving strategies that remain operationally realistic and consistent with established driving practices, in which longitudinal control patterns play a dominant role in determining energy consumption for heavy-duty vehicles [7]. Similar observations have been reported in recent studies on driving behavior recognition and fuel economy evaluation, highlighting the dominant role of longitudinal control patterns in determining energy consumption for heavy-duty vehicles [23–24].

From a methodological perspective, treating driver behavior as an explicit control layer improves interpretability and practical relevance. Learned policies can be more easily analyzed, validated, and potentially translated into driver assistance or training systems, bridging the gap between data-driven optimization and real-world implementation [36,37].

3.9. Contribution of the Digital Twin Architecture

The proposed digital twin architecture provides an effective platform for integrating real-world telematics data, data-driven vehicle dynamics modeling, and reinforcement learning-based control within a unified simulation environment. By embedding a surrogate vehicle dynamics model trained on historical operational data, the Digital Twin enables fast and stable simulation of fuel consumption while preserving the essential nonlinear relationships between vehicle state, driver behavior, and environmental conditions, which are critical for realistic heavy-duty vehicle modeling [38].

Compared to existing digital twin applications in mining, which are often limited to monitoring, visualization, or descriptive analytics, the proposed framework extends the digital twin paradigm toward prescriptive optimization through learning-based control. This aligns with recent reviews emphasizing the need for digital twins to evolve from passive representations into active decision-support and optimization tools [16,39].

Moreover, the modular separation of vehicle dynamics, driver behavior, and environment models enhances extensibility and scalability. Such modular digital twin architectures have been widely recognized as a key enabler for adapting models across different assets, operating conditions, and optimization objectives without fundamental redesign [14–15,40]. The proposed framework can therefore accommodate alternative surrogate models, additional behavioral features, or more complex environmental representations, supporting its applicability to a broader range of haulage systems and operating scenarios.

While the surrogate vehicle dynamics model achieves an R^2 of 0.780 on the test dataset, its primary purpose is to serve as a computationally efficient transition model for reinforcement learning simulations in the digital twin environment. Both the reinforcement learning-generated policy and the historical operator driving profiles are evaluated using identical surrogate model assumptions, ensuring that the reported fuel efficiency improvements reflect relative performance differences rather than absolute fuel consumption estimates.

3.10. Comparison with Existing Studies

The results of this study are broadly consistent with prior reinforcement learning-based eco-driving and energy management research, which has demonstrated meaningful reductions in fuel or energy consumption under controlled simulation conditions. Studies in on-road and off-road domains report that reinforcement learning can effectively learn smooth and energy-efficient control policies when supported by appropriate state representations and reward formulations [22,31]. More generally, recent surveys and reviews of reinforcement learning applications in control and transportation systems have highlighted the effectiveness of policy-gradient methods for learning energy-efficient control strategies in complex environments [41].

Direct quantitative comparison with existing work remains challenging due to differences in vehicle types, operating environments, evaluation protocols, and performance metrics. In particular, many prior studies rely on simplified vehicle models or synthetic driving cycles to facilitate training and evaluation [22]. In contrast, the present work leverages real-world telematics data and explicitly models driver behavior within a digital twin environment. As such, the reported fuel savings should be interpreted as indicative of the framework's effectiveness within its defined scope rather than as a universal benchmark across mining operations, a limitation commonly acknowledged in recent reviews of digital twin-enabled mining and transportation systems [16,18].

3.11. Limitations and Future Research Directions

Despite the promising results obtained in this study, several limitations exist that contextualize the scope and applicability of the proposed framework. First, the experimental evaluation was conducted entirely within a digital twin environment using a data-driven vehicle dynamics surrogate model. Although the surrogate model demonstrated adequate predictive accuracy and supported stable reinforcement learning, discrepancies between simulated and real-world behavior may still emerge, particularly under operating conditions that are rarely represented in historical data or during extreme events. Second, the reinforcement learning policy evaluated in this study was optimized under controlled operating assumptions, resulting in a highly stable driving pattern with limited speed variability during the evaluation episodes. While this behavior is desirable for isolating fuel efficiency effects and ensuring fair comparisons with historical operator data, it does not reflect the full range of adaptability required in real-world haulage operations involving traffic interactions, stochastic disturbances, or dynamically changing routes.

Third, the evaluation focused on a limited set of operators, routes, and operating conditions, with load and road gradients held constant to ensure methodological fairness. While this design choice allows for robust performance comparisons, it limits the generalizability of the reported fuel savings across a broader and more heterogeneous range of mining scenarios.

Future research will address these limitations by extending the proposed framework in several directions. Multi-episode and multi-route evaluations will be conducted to assess policy robustness across different operational contexts and operator profiles. Environmental uncertainty and stochastic disturbances will be incorporated into the digital twin to increase realism and test the policy's adaptability. Furthermore, hybrid control strategies that combine reinforcement learning with rule-based safety constraints or human involvement mechanisms will be explored to enhance operational acceptability. Finally, integrating with real-time fleet management and driver assistance systems is a promising path to translating the proposed approach into a practical decision-support tool for surface mining operations.

4. Conclusions

This study presented a driver behavior-aware fuel optimization framework for open-pit haul trucks by integrating a digital twin architecture with reinforcement learning. The proposed framework combines a data-driven vehicle dynamics surrogate, explicit driver behavior modeling, and PPO to enable safe and scalable learning of fuel-efficient driving strategies within a realistic simulation environment. Experimental results demonstrate that the digital twin surrogate provides sufficient predictive fidelity to support stable reinforcement learning. Compared to historical operator driving, the reinforcement learning policy produced substantially smoother speed regulation, eliminated aggressive acceleration and braking events, and reduced speed variability. Under identical operating conditions, fuel consumption per kilometer was decreased from 4.45 L/km to 3.02 L/km, corresponding to a 32.05% improvement in fuel efficiency. A key contribution of this work lies in explicitly modeling driver behavior as an integral component of the optimization process. Rather than optimizing fuel consumption solely through abstract control objectives, the proposed approach establishes a clear behavioral pathway linking speed stability and smooth driving patterns to measurable fuel savings. This behavior-aware formulation enhances interpretability and aligns learning-based optimization with practical operational considerations in heavy-duty haulage. While the learned policy exhibits highly stable speed regulation under controlled evaluation conditions, this behavior reflects the objective of isolating fuel efficiency effects and provides a foundation for future extensions incorporating greater operational variability. Overall, the results indicate that digital twin-based, behavior-aware reinforcement learning represents a promising approach for fuel efficiency improvement in open-pit mining operations. By enabling controlled experimentation and interpretable policy learning within a realistic simulation environment, the proposed framework provides a foundation for future development of decision support and driver assistance systems aimed at reducing fuel consumption while maintaining realistic and operationally feasible driving behavior.

All the Declarations and Statements

Author Contributions Statement

Kusnawi Kusnawi – Conceptualization, Methodology, Data Analysis, and Writing – Original Draft Preparation.

Mochamad Agung Wibowo – Supervision, Validation, and Writing – Review & Editing.

Ridwan Sanjaya – Formal Analysis, Visualization, and Writing – Review & Editing.

All authors have read and agreed to the published version of the manuscript.

Conflict of Interest Statement

The authors declare no conflict of interest.

Funding Declaration

This research received no external funding.

Data Availability Statement

The data used in this study are available upon reasonable request.

Ethical Declarations

This study did not involve human participants or animal experiments. Therefore, no ethical approval was required.

Acknowledgments

The authors would like to express their gratitude to Amikom University Yogyakarta and the Doctoral Program in Information Systems at Diponegoro University, Semarang, Indonesia, for their academic support and provision of a research environment that enabled the completion of this research.

Declaration of Generative AI in Scholarly Writing

The use of AI was limited to grammar checking and language editing. All generated manuscripts were reviewed and verified by the authors, who are solely responsible for the final submission.

Abbreviations

The following abbreviations are used in this manuscript:

CDF – Cumulative Distribution Function
IQR – Interquartile range
RL – Reinforcement Learning
PPO – Proximal Policy Optimization
VDM – Vehicle Dynamics Model
MDP – Markov Decision Process
MLP – Multilayer Perceptron
RMSE – Root Mean Square Error

References

- [1] P. Bodziony and M. Patyk, “The Influence of the Mining Operation Environment on the Energy Consumption and Technical Availability of Truck Haulage Operations in Surface Mines,” *Energies*, vol. 17, no. 11, 2024, doi: 10.3390/en17112654.
- [2] Q. Wang, R. Zhang, S. Lv, and Y. Wang, “Open-pit mine truck fuel consumption pattern and application based on multi-dimensional features and XGBoost,” *Sustainable Energy Technologies and Assessments*, vol. 43, p. 100977, Feb. 2021, doi: 10.1016/j.seta.2020.100977.
- [3] L. Chen *et al.*, “Autonomous mining through cooperative driving and operations enabled by parallel intelligence,” *Communications Engineering*, vol. 3, no. 1, p. 75, May 2024, doi: 10.1038/s44172-024-00220-5.
- [4] F. Mensing, R. Trigui, H. Tategrain, and E. Bideaux, “Ecodriving. From processing the ideal speed profile to its use during driving activity,” *European Journal of Electrical Engineering*, vol. 17, no. 5–6, pp. 397–418, Dec. 2014, doi: 10.3166/ejee.17.397-418.
- [5] Q. Wang, R. Zhang, Y. Wang, and S. Lv, “Machine learning-based driving style identification of truck drivers in open-pit mines,” *Electronics (Switzerland)*, vol. 9, no. 1, pp. 1–24, 2020, doi: 10.3390/electronics9010019.
- [6] X. Zhu, X. Shen, K. Chen, and Z. Zhang, “Research on the prediction and influencing factors of heavy duty truck fuel consumption based on LightGBM,” *Energy*, vol. 296, 2024, doi: 10.1016/j.energy.2024.131221.
- [7] Z. Ma, B. N. Jørgensen, and Z. Ma, “A Scoping Review of Energy-Efficient Driving Behaviors and Applied State-of-the-Art AI Methods,” *Energies (Basel)*, vol. 17, no. 2, p. 500, Jan. 2024, doi: 10.3390/en17020500.
- [8] T. Bousonville, D. Cheubou Kamga, T. Krüger, and M. Dirichs, “Data driven analysis and forecasting of medium and heavy truck fuel consumption,” *Enterp. Inf. Syst.*, vol. 16, no. 6, 2022, doi: 10.1080/17517575.2020.1856417.
- [9] S. Alamdari, M. Hossein Basiri, A. Mousavi, and A. Soofastaei, “Application of Machine Learning Techniques to Predict Haul Truck Fuel Consumption in Open-Pit Mines,” *Journal of Mining and Environment*, vol. 13, no. 1, pp. 69–85, 2022, doi: 10.22044/jme.2022.11577.2145.
- [10] P. Fan, G. Song, Z. Zhai, Y. Wu, and L. Yu, “Fuel consumption estimation in heavy-duty trucks: Integrating vehicle weight into deep-learning frameworks,” *Transp. Res. D Transp. Environ.*, vol. 130, p. 104157, May 2024, doi: 10.1016/j.trd.2024.104157.
- [11] L. Liu, W. Li, D. Wang, Y. Wu, R. Yang, and W. Shi, “Fuel Rate Prediction for Heavy-Duty Trucks,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 8, pp. 8222–8235, Aug. 2023, doi: 10.1109/TITS.2023.3265007.
- [12] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, and F. Sui, “Digital twin-driven product design, manufacturing and service with big data,” *The International Journal of Advanced Manufacturing Technology*, vol. 94, no. 9–12, pp. 3563–3576, Feb. 2018, doi: 10.1007/s00170-017-0233-1.
- [13] F. Tao, M. Zhang, and A. Y. C. Nee, *Digital twin driven smart manufacturing*. Academic Press, 2019.
- [14] Q. Qi and F. Tao, “Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison,” *IEEE*

- Access*, vol. 6, pp. 3585–3593, Jan. 2018, doi: 10.1109/ACCESS.2018.2793265.
- [15] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, “Characterising the Digital Twin: A systematic literature review,” *CIRP J Manuf Sci Technol*, vol. 29, pp. 36–52, May 2020, doi: 10.1016/j.cirpj.2020.02.002.
 - [16] P. Nobahar, C. Xu, P. Dowd, and R. Shirani Faradonbeh, “Exploring digital twin systems in mining operations: A review,” *Green and Smart Mining Engineering*, vol. 1, no. 4, pp. 474–492, Dec. 2024, doi: 10.1016/j.gsme.2024.09.003.
 - [17] N. Elbazi, A. Tigami, O. Laayati, A. El Maghraoui, A. Chebak, and M. Mabrouki, “Digital Twin-Enabled Monitoring of Mining Haul Trucks with Expert System Integration: A Case Study in an Experimental Open-Pit Mine,” in *2023 5th Global Power, Energy and Communication Conference (GPECOM)*, IEEE, Jun. 2023, pp. 168–174, doi: 10.1109/GPECOM58364.2023.10175789.
 - [18] H. Wu, P. Ji, H. Ma, and L. Xing, “A Comprehensive Review of Digital Twin from the Perspective of Total Process: Data, Models, Networks and Applications,” *Sensors*, vol. 23, no. 19, p. 8306, Oct. 2023, doi: 10.3390/s23198306.
 - [19] B. Li, Y. Cui, Y. Xiao, S. Fu, J. Choi, and C. Zheng, “An improved energy management strategy of fuel cell hybrid vehicles based on proximal policy optimization algorithm,” *Energy*, vol. 317, p. 134585, Feb. 2025, doi: 10.1016/j.energy.2025.134585.
 - [20] H. Chen, G. Guo, B. Tang, G. Hu, X. Tang, and T. Liu, “Data-driven transferred energy management strategy for hybrid electric vehicles via deep reinforcement learning,” *Energy Reports*, vol. 10, pp. 2680–2692, Nov. 2023, doi: 10.1016/j.egy.2023.09.087.
 - [21] Z. Yang, Z. Zheng, J. Kim, and H. Rakha, “Eco-driving strategies using reinforcement learning for mixed traffic in the vicinity of signalized intersections,” *Transp Res Part C Emerg Technol*, vol. 165, p. 104683, Aug. 2024, doi: 10.1016/j.trc.2024.104683.
 - [22] H. Kim, H. Pyeon, J. S. Park, J. Y. Hwang, and S. Lim, “Autonomous vehicle fuel economy optimization with deep reinforcement learning,” *Electronics (Switzerland)*, vol. 9, no. 11, pp. 1–19, 2020, doi: 10.3390/electronics9111911.
 - [23] Z. Zhang, R. Zhang, and J. Sun, “Research on the Comprehensive Evaluation Method of Driving Behavior of Mining Truck Drivers in an Open-Pit Mine,” *Applied Sciences*, vol. 13, no. 20, p. 11597, Oct. 2023, doi: 10.3390/app132011597.
 - [24] L. Fang, H. Li, Y. Zheng, and X. Luo, “Driving behavior recognition and fuel economy evaluation for heavy-duty vehicles,” *Research in Transportation Business & Management*, vol. 60, p. 101371, Jun. 2025, doi: 10.1016/j.rtbm.2025.101371.
 - [25] Z. Zhang, E. Demir, R. Mason, and C. Di Cairano-Gilfedder, “Understanding freight drivers’ behavior and the impact on vehicles’ fuel consumption and CO₂e emissions,” *Operational Research*, vol. 23, no. 4, p. 59, Dec. 2023, doi: 10.1007/s12351-023-00798-2.
 - [26] Kusnawi, M. Agung Wibowo, and R. Sanjaya, “A Systematic Literature Review of Adaptive Machine Learning Approaches for Real-Time Fuel Efficiency Optimization in Open-Pit Mining Trucks,” *Sistem Informasi dan Komputer*, vol. 15, pp. 40–46, 2025, doi: 10.32736/sisfokom.v15i1.2527.
 - [27] H. I. Ashqar, M. Obaid, A. Jaber, R. Ashqar, N. O. Khanfar, and M. Elhenawy, “Incorporating driving behavior into vehicle fuel consumption prediction: methodology development and testing,” *Discover Sustainability*, vol. 5, no. 1, Dec. 2024, doi: 10.1007/s43621-024-00511-z.
 - [28] A. Fuller, Z. Fan, C. Day, and C. Barlow, “Digital Twin: Enabling Technologies, Challenges and Open Research,” *IEEE Access*, vol. 8, pp. 108952–108971, 2020, doi: 10.1109/ACCESS.2020.2998358.
 - [29] R. S. Sutton and A. G. Barto, *Reinforcement learning: an introduction*. The MIT Press, 2020.
 - [30] H. Hailemichael, B. Ayalew, L. Kerbel, A. Ivanco, and K. Loiseau, “Safe Reinforcement Learning for an Energy-Efficient Driver Assistance System,” *IFAC-PapersOnLine*, vol. 55, no. 37, pp. 615–620, 2022, doi: 10.1016/j.ifacol.2022.11.250.
 - [31] Y. Ye, B. Xu, H. Wang, J. Zhang, B. Lawler, and B. Ayalew, “Deep reinforcement learning-based energy management system enhancement using digital twin for electric vehicles,” *Energy*, vol. 312, p. 133384, Dec. 2024, doi: 10.1016/j.energy.2024.133384.
 - [32] H. Ding, W. Li, N. Xu, and J. Zhang, “An enhanced eco-driving strategy based on reinforcement learning for connected electric vehicles: cooperative velocity and lane-changing control,” *Journal of Intelligent and Connected Vehicles*, vol. 5, no. 3, pp. 316–332, Oct. 2022, doi: 10.1108/JICV-07-2022-0030.
 - [33] X. Tang *et al.*, “Reinforcement Learning-Based Energy Management for Hybrid Power Systems: State-of-the-Art Survey, Review, and Perspectives,” *Chinese Journal of Mechanical Engineering*, vol. 37, no. 1, p. 43, May 2024, doi: 10.1186/s10033-024-01026-4.
 - [34] A. Raffin *et al.*, “Stable-Baselines3: Reliable Reinforcement Learning Implementations,” 2021. [Online]. Available: <https://github.com/DLR-RM/stable-baselines3>.
 - [35] M. Miotti, Z. A. Needell, S. Ramakrishnan, J. Heywood, and J. E. Trancik, “Quantifying the impact of driving style changes on light-duty vehicle fuel consumption,” *Transp Res D Transp Environ*, vol. 98, p. 102918, Sep. 2021, doi: 10.1016/j.trd.2021.102918.
 - [36] E. S. Sumner *et al.*, “Personalizing driver safety interfaces via driver cognitive factors inference,” *Sci. Rep.*, vol. 14, no. 1, p. 18058, Aug. 2024, doi: 10.1038/s41598-024-65144-8.
 - [37] D. L. Fisher, W. J. Horrey, J. D. Lee, and M. A. Regan, Eds., *Handbook of Human Factors for Automated, Connected, and Intelligent Vehicles*. Boca Raton, FL: CRC Press, 2020, doi: 10.1201/b21974.
 - [38] E. Tomanik, A. J. Jimenez-Reyes, V. Tomanik, and B. Tormos, “Machine-Learning-Based Digital Twins for Transient Vehicle Cycles and Their Potential for Predicting Fuel Consumption,” *Vehicles*, vol. 5, no. 2, pp. 583–604, May 2023, doi: 10.3390/vehicles5020032.
 - [39] D. M. Botín-Sanabria, A.-S. Mihaita, R. E. Peimbert-García, M. A. Ramírez-Moreno, R. A. Ramírez-Mendoza, and J. de J. Lozoya-Santos, “Digital Twin Technology Challenges and Applications: A Comprehensive Review,” *Remote Sens (Basel)*, vol. 14, no. 6, p. 1335, Mar. 2022, doi: 10.3390/rs14061335.
 - [40] C. K. Lo, C. H. Chen, and R. Y. Zhong, “A review of digital twin in product design and development,” *Advanced Engineering Informatics*, vol. 48, Apr. 2021, doi: 10.1016/j.aei.2021.101297.
 - [41] D. J. Joshi, I. Kale, S. Gandewar, O. Korate, D. Patwari, and S. Patil, “Reinforcement Learning: A Survey,” in *Advances in Intelligent Systems and Computing*, vol. 1311, 2021, pp. 297–308, doi: 10.1007/978-981-33-4859-2_29.

Authors' Profiles



Kusnawi Kusnawi received his Bachelor of Computer Science (S.Kom) in Informatics from STMIK AMIKOM Yogyakarta, Indonesia, and his Master of Engineering (M.Eng.) in Information Technology from the Faculty of Electrical Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia. He is currently a doctoral candidate in Information Systems at the Graduate School of Diponegoro University, Semarang, Indonesia. His research interests include decision support systems, information systems, data science, data warehousing, big data, and machine learning.



Mochamad Agung Wibowo is the Dean of the Graduate School of Diponegoro University. Previously, he served as Dean of the Faculty of Engineering of Diponegoro University for the 2015–2019 and 2019–2024 periods. He earned a bachelor's degree in civil engineering from Diponegoro University, Semarang, Indonesia, in 1992, followed by a master's degree in management from the same university in 1996. He later continued his studies at Nottingham University, UK, and obtained a Master of Science degree in 1999 and a doctorate (Ph.D.) in 2004. He has published extensively in reputable national and international journals. His areas of research expertise include green supply chain management in construction, risk management, integration of design and procurement strategies in government infrastructure projects, implementation of lean construction to support the Sustainable Development Goals (SDGs), and project safety audits utilizing digital technology. His contributions span a wide range of academic and higher education policy areas in Indonesia, particularly in advancing sustainable research and innovation.



Ridwan Sanjaya is a full professor in information systems at Soegijapranata Catholic University, Semarang, Indonesia. He joined the SCU faculty in 2002, initially serving as a lecturer. He obtained his M.Sc. in Internet and E-Commerce Technology and his Ph.D. in Computer Information Systems from the Graduate School of Information Technology, Assumption University, Bangkok, Thailand (in 2006 and 2011, respectively). He has expertise in utilizing technology to transform educational content for students, to align with the business, and to empower people.

How to cite this paper

Kusnawi Kusnawi, Mochamad Agung Wibowo, Ridwan Sanjaya, "Driver Behavior-aware Fuel Optimization Using a Digital Twin and Reinforcement Learning Approach for Open-pit Haul Trucks", *International Journal of Information Technology and Computer Science(IJITCS)*, Vol.18, No.4, pp.155-169, 2026. DOI:10.5815/ijitcs.2026.04.10