Optimizing Truck Dispatching and Shovel Allocation Using Deep Reinforcement Learning in Open-Pit Mining: Literature Review

Khaleeq Ahmed[1] and Yashar Pourrahimian[2]

[1,2] School of Mining and Petroleum Engineering, University of Alberta, Edmonton, Canada

ABSTRACT

Material handling and transportation expenses can make up as much as 50% of operating costs in open-pit mining operations. The extracted material must be transported to different destinations like processing plants, dumps or stockpiles. Optimizing truck dispatching and shovel allocation is a major difficulty in open-pit mining operations, as it significantly affects both operating costs and production. In open-pit mining, dispatching efficiency is increased by simulation-based optimization and mixed integer programming models. Reinforcement learning (RL) has been applied in complex production control systems across several industrial processes in the last decade, yielding better results than previous approaches. Nevertheless, deep Reinforcement Learning (DRL) applications are relatively uncommon in the mining sector. This paper explores the potential of DRL-based approaches in mining operations for efficient fleet management systems. The research aims to improve understanding of current applications of these technologies in open-pit mine planning, as well as their potential in the future and any potential field constraints.

Keywords: *Optimization; Reinforcement learning (RL), Open-pit mines, fleet management*

1. Introduction

Over 60% of the output from surface mining is generated by open-pit operations [1]. In open-pit mining, the production process entails a number of tasks, including loading, moving while loaded, maneuvering at the dump, dumping, and returning to the loader [2]. The mining transport scheme's primary activities, loading and haulage, account for more than half of all operational expenditures [3]. A popular method of hauling, especially in large-scale open-pit mining operations, is the shovel-truck arrangement [4]. Drilling and blasting are common methods used in mining to recover material from hard rocks. After the material has been broken up, shovels are used to put it into trucks. These trucks subsequently move the debris to different locations, such as landfills for disposal, stockpiles for storage, or crushers for additional processing. This technique maximizes the total productivity of the mining operation by guaranteeing the effective handling and processing of mined materials [5].

Any mining operation's capacity to turn a profit depends on how well hauling and loading resources are allocated and used. The efficient management of these resources ensures prompt and economical transportation of materials from extraction regions to their intended locations, such as stockpiles, crushers, or waste dumps. In loading and hauling operations, the incorporation of optimization and strategic planning approaches can yield substantial cost savings, boost productivity, and improve

ISBN: 978-1-55195-520-9

overall profitability. Mining operations can attain peak efficiency and long-term financial gain by reducing idle time and optimizing the use of trucks and shovels [6].

Shovel allocation and truck dispatching are essential to open-pit mining's operational effectiveness. The productivity, cost-effectiveness, and environmental impact of mining operations are all directly impacted by these processes. Effective truck dispatching minimizes idle times and maximises the use of both trucks and shovels by ensuring that the appropriate number of trucks are available at the appropriate time and location. On the other hand, efficient shovel allocation guarantees a continuous and seamless mining process, eliminating bottlenecks and improving the mining workflow as a whole.

The optimization of truck dispatching and shovel allocation is a difficult task because of the dynamic and unpredictable nature of open-pit mining operations. Weather conditions, equipment failures, haul distances, and fluctuating ore grades increase the complexity of developing effective dispatching and allocation schedules. While somewhat efficient, traditional optimization techniques frequently find it difficult to adjust in real time to these quickly changing situations [3].

Creating mathematical models to optimize equipment allocation and pit sequencing has been a major area of research in open-pit mine short-term planning. Nonetheless, there are a lot of difficulties because pit production systems are inherently complicated and because the elements that make them up are random. These models are only useful for deterministic assessments since they frequently fail to capture all operational subtleties [7].

Machine learning (ML) and artificial intelligence (AI) advancements have created new opportunities to address these issues. Deep reinforcement learning (DRL) has become one of the most promising methods. DRL is an effective tool for managing the intricacies and uncertainties involved in open-pit mining operations because it combines the decision-making powers of reinforcement learning with the potent representation-learning capabilities of deep learning [8]. In order to optimize a goal, Reinforcement Learning (RL), a subfield of ML, uses a computational approach to learn from interactions with an environment [9]. DRL offers a highly adaptable data-driven production control framework and has been applied more frequently to optimize various engineering systems in the manufacturing, transportation, and heavy industries [10]. Mining firms may create strong and adaptable solutions for shovel allocation and truck dispatching by utilizing DRL, which will considerably increase operational efficiency and cost savings. Through better resource management and a decrease in the environmental impact of mining operations, the use of DRL can also support more environmentally friendly mining methods [11].

1.1. Shovel Allocation

An essential component of open-pit mining's operational effectiveness is shovel allocation. Proper shovel allocation guarantees a continuous and effective mining process by avoiding bottlenecks and maximizing workflow. To optimize output and reduce operational delays, the allocation of shovels entails selecting which shovels to use at particular pit locations. This work is complicated because mining operations are dynamic and unpredictable, with variations in ore grades, haul distances, equipment performance, and environmental conditions. Mathematical models have traditionally optimized shovel allocation, however these models are generally confined to deterministic evaluations and unable to capture all practical factors [12].

1.2. Truck Dispatching

Another essential element of open-pit mining operations is truck dispatching. A general view of truck activities in a mining operation is shown in Figure 1. Effective truck dispatching minimizes idle times and maximizes the use of both trucks and shovels by ensuring that the appropriate number of trucks are available at the appropriate time and location. To ensure a smooth and continuous flow of materials, truck dispatching aims to construct an ideal plan that matches truck availability with shovel operations. The dynamic nature of mining conditions, truck travel periods, loading and unloading times, and other factors make this task difficult to balance. Although helpful, traditional optimization techniques frequently cannot adjust in real-time, which results in less-than-ideal performance [13].

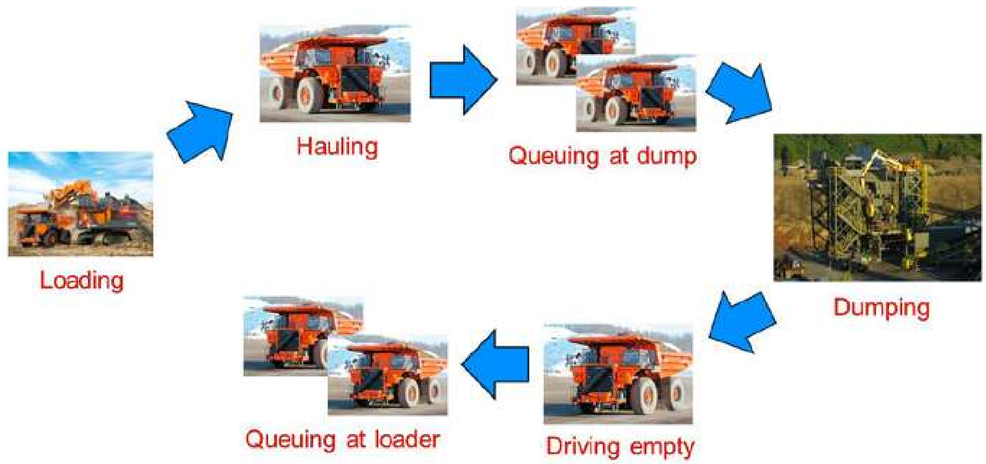

Figure 1. Truck activities in haul cycle during mining operation [14].

1.3. Reinforcement Learning

In RL, an agent picks up decision-making skills by interacting with its surroundings. RL uses a dynamic process of trial and error to maximize cumulative rewards, in contrast to supervised learning, where the model learns from a fixed dataset.

1.3.1. The Agent and Environment

The decision-maker in RL is the agent. It acts in accordance with a policy, which could be as basic as a rule or as sophisticated as a neural network. On the other hand, the environment is everything with which the agent comes into contact and everything that reacts to the agent's actions. A Markov Decision Process (MDP) is commonly used to simulate the interaction between the agent and the environment as shown in Figure 2.

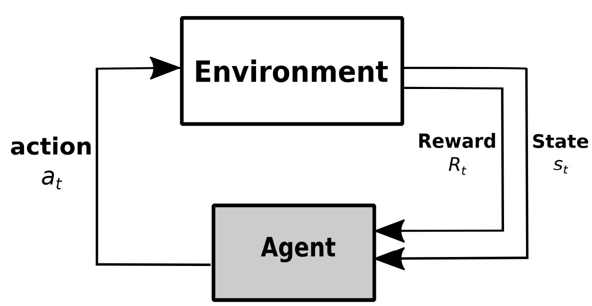

Figure 2. Reinforcement Learning, Agent and Environment [9].

a) State: The environment is in a particular state at any one time that offers all the data required to make decisions.
b) Action: Considering the existing circumstances, the agent chooses an action. The action space is the collection of all conceivable actions that the agent is capable of.
c) Reward: The agent receives a reward every time it interacts with the environment. This reward is a numerical value that indicates the immediate gain of the action.

d) Transition: After the interaction of the agent with the environment, the environment undergoes a new state. Stochastic components might potentially have an impact on this transition, representing the uncertainty found in real-world situations.

2. Literature Review

In open-pit mines, the Fleet Management System (FMS) is crucial to optimize the fleet management of mining vehicles, such as trucks, shovels, and other equipment, to boost output, save operating expenses, and increase safety. Because FMS offers real-time equipment tracking, mine operators can keep an eye on the whereabouts and condition of any piece of equipment. This guarantees effective resource allocation and usage[15]. Optimizing truck dispatching is one of the main purposes of FMS. In order to minimize idle time of loading and hauling equipment and shorten cycle times, the system assigns vehicles to loading and dumping locations using algorithms [16]. FMS monitors the health and performance of equipment by collecting data from sensors and diagnostic tools. This data can be used in machine conditioning to help in predictive maintenance and reduce unexpected downtimes [17]. The data-driven approach from FMS helps make informed decisions by generating reports on productivity, fuel consumption and operational efficiency [18]. Regarding safety of the equipment's operations and worker locations, FMS helps prevent accidents by enforcing safety protocols by restricting access to certain areas [19].

2.1. Truck Fleet Dispatching

Determining where a haul truck will go when its load is dumped is known as truck dispatching. When managed well, truck dispatching, nested in relation to FMSs, logistics, and supply chain, can improve ore output and decrease truck waiting times by more than 15% [20]. In the past, heuristic algorithms and linear programming have been used to successfully handle the problem of truck fleet dispatching in open-pit mining. By resolving a series of linear equations, linear programming models maximize the distribution of trucks among loading and unloading locations while reducing overall operating expenses and cycle times. Heuristic algorithms like genetic algorithms, simulated annealing, and tabu search supplement these models by offering efficient and useful answers in complex and dynamic mining situations where decisions must be made quickly. Because of this hybrid strategy, dispatching decisions are guaranteed to be effective and flexible in response to the constantly shifting circumstances of open-pit mining operations [21].

Heuristic algorithms, however, are good at finding answers quickly, but they frequently depend on incomplete knowledge of the mining system and fail to take into consideration the major sources of uncertainty that are present in mining operations. These methods can ignore the unpredictability and variety of elements like ore quality, ambient circumstances, and equipment availability. Moreover, through Internet of Things (IoT) devices and sensors, contemporary mining operations produce enormous amounts of real-time data on equipment's location, status, and performance. Despite this abundance of data, the fleet dispatching systems in place today underuse these data streams, losing out on chances for more flexible and dynamic decision-making that might improve operational efficiency and flexibility. This gap could be filled by combining traditional dispatching methods with advanced data analytics and ML approaches, utilizing real-time data to manage uncertainties better and optimize fleet operations [22].

2.2. Short-Term Production Planning

Eivazy and Askari-Nasab [23] proposed a Mixed-Integer Programming (MIP) formulation for short-term production planning. It creates monthly extraction schedules for open-pit mines and allows for the mining of commodities to be taken to different destinations. This formulation considers the choice of ramps for removing ore from the pit and incorporates a horizontally directed mining method. The approach offers comprehensive and flexible schedules that maximize transportation logistics and resource allocation by concentrating on monthly intervals. L'Heureux et al. [24] presented a model that minimizes the operation's overall cost by taking into account decisions about drilling and blasting, shovel distribution, and movement between the mining zones. Another attempt was made by Kozan and Liu [25]. They considered the drilling and blasting operations while planning and optimizing the process to create a workable schedule for shift operations and aim to maximize throughput and reduce anticipated equipment idle time.

The main limitation of these methods is the complexity level necessary to depict the mining production environment adequately. These complex models are computationally demanding because they must consider many factors and limitations. Customized heuristics that introduce established rules and assumptions are frequently required to solve these models. The model is restricted to deterministic circumstances, and its versatility is limited by its dependence on heuristics. As a result, the inherent uncertainties of mining operations, such as variations in ore quality and equipment failures, are difficult for these models to simulate, which limits their usefulness and resilience in real-world situations [26].

To develop operating plans that specifically take equipment interactions inside the mine layout into account, more recent research efforts have combined discrete simulation with optimization engines. Discrete Event Simulation (DES) integration enables more robust and data-driven schedules, improving planning process accuracy and dependability. With the use of DES, mining operations' dynamic interactions and unpredictability may be modelled, resulting in more flexible schedules that are better equipped to adjust to changing conditions and uncertainties in the mine environment. Upadhyay and Askari-Nasab [27] presented a detailed discrete simulation of mining operations that uses CPLEX engine to obtain optimal shovel allocations to mining faces. In their extended approach [28], they extend their approach to optimize the truck and shovel allocation and optimal mining faces extraction sequences using multi-objective optimization approach within the simulation engine. In Germany, Shishvan and Benndorf [29] presented a comparable framework for simulation-optimizing operational choices in a continuous coal mining system. This framework records specific mining site excavation and disposal procedures using simulation. Instead of anticipating and learning from the uncertainties in the system, the optimization engine in this technique responds to the system's condition at each decision point. This makes it more difficult to optimize long-term results and deal with fluctuation in a proactive manner.

3. RL in Fleet Management Systems

Trucks and shovels act as agents when applying RL in open-pit mining systems. The environment is the pit itself, which includes places like dumps for waste, stockpiles, and crushers as well as different sources like mining faces as shown in Figure 3. The RL agent looks at the present state of the environment, including the parameters and conditions of the mining system, before making a choice on truck dispatch. The real-time dispatch solution generated by the RL agent is based on this observation and optimizes the movement and allocation of vehicles to improve productivity and operational efficiency. The system can react to shifting conditions and uncertainties in the mining environment with effectiveness due to its dynamic and adaptive approach.

3.1. Markov Decision Process Components

The truck dispatching and shovel allocation problem in an open-pit mine can be represented as an MDP and solved using an RL approach. The components of this MDP are defined as follows:

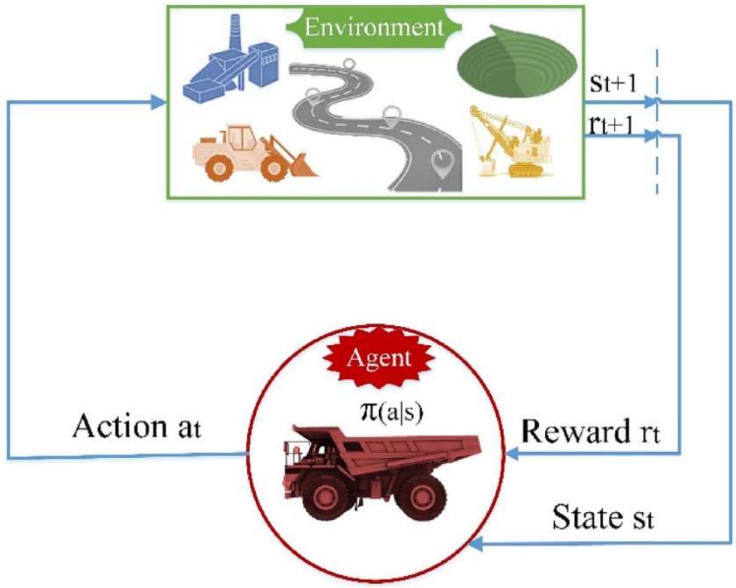

Figure 3. Agent interacting within a pit environment [30].

1. *State Space (S)*

   At any given time, the state space represents all possible system configurations. It includes:

   - Current locations of all trucks
   - The location and state of operation of every shovel at the moment
   - Current load status of each truck
   - Current stockpile levels at different destinations (for example, crushers, waste dumps)
   - Queue lengths at shovels and dumping locations
   - The state of the road network, including the traffic and maintenance status

2. *Action Space (A)*

   Action space consists of all possible actions that can be taken to dispatch trucks and allocate shovels. Actions include:

   - Assigning a specific truck to a specific shovel
   - Directing a truck that is loaded to a designated dump site.
   - Directing an empty vehicle to be loaded at a particular spot.
   - Decisions on shovel maintenance or modifications to operations.

3. *Transition Function (P)*

   The transition function defines the probability of transitioning from one state to another based on a specific action. It models the dynamics of the system, including:

   - Variations in the productivity of shovels

- Loading and Dumping times
- Breakdown and repair times for trucks and shovels
- Travel times between source and destinations

4. *Reward Function (R)*

   Based on the action taken by the RL agent, the reward function provides feedback to the RL agent. It typically includes:

   - Positive rewards for maximizing ore throughput to crushers
   - Penalty for truck idle times, shovel idle times, and excessive queuing
   - Penalties for breakdown of truck and shovels
   - Rewards for maintaining optimal stockpile levels and minimizing operational costs

5. *Agent*

   In the case of open-pit mines, entities like trucks and shovels will act as an agent that will interact with the environment by taking actions in the current state. The agent's objective is to learn an optimal policy by exploring and exploiting the state-action space. Different RL algorithms can be used to update the optimal policy and maximize the cumulative rewards.

6. *Policy (π)*

   The policy is a tactic the RL agent uses to choose what to do in a certain situation. The RL agent aims to learn an ideal policy that maximizes cumulative rewards over time.

The truck dispatching and shovel allocation problem can be modelled as an MDP, enabling the use of RL algorithms to discover effective policies that maximize the mining operation's overall performance. This formulation takes advantage of RL's capacity to manage open-pit mining situations' intricacy and dynamic character, leading to enhanced operational efficiency and decision-making.

3.2. Review of the Previous Work

A comprehensive review of the existing literature on FMS in open-pit mines was conducted to gather relevant studies on the advancement of this field of RL in open-pit mines. To ensure a thorough review, a set of precise keywords was employed to locate relevant publications and research papers. Figure 4 shows the network of keywords utilized in the literature search, emphasizing how several research themes linked to intelligent FMS in open-pit mining are interconnected. The visualization highlights the diverse methodology employed by scholars to address the complex nature of mining logistics and the crucial function of FMS in attaining operational excellence.

This research focuses primarily on RL-based FMSs after establishing the perspective of various intelligent approaches in FMSs for open-pit mines. This study aims to evaluate the existing level of understanding in this domain and pinpoint important avenues for future research pertaining to algorithmic and mining components of these intelligent systems. This study focuses on RL-based techniques to illustrate these methods' advantages and disadvantages in the mining industry.

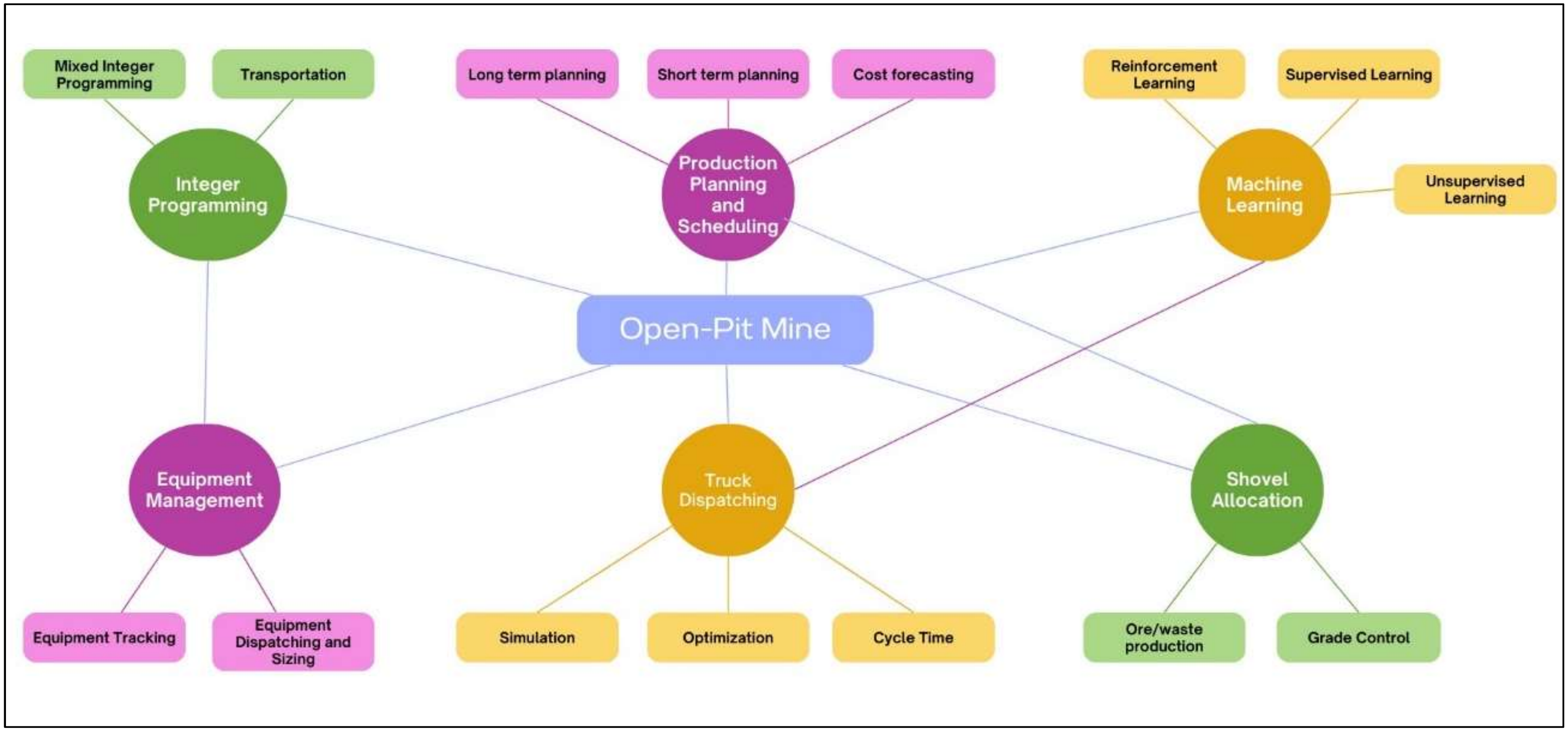

Figure 4. Keywords used to find the related articles in the literature regarding FMS in Open-pit Mines.

The development and use of RL and other intelligent approaches in FMS for open-pit mining have been greatly aided by the important publications included in the following Table 1. These studies demonstrate the history and various methodologies used to optimize mining operations, including various algorithms and approaches. The table shows the development of this field's research by being arranged chronologically. The authors, the algorithm, and a synopsis of their work are all included in each entry.

Table 1. Table of Key Articles on RL-based FMS in Open-Pit Mines.

Authors	Approach Used	Explanation
Siami-Irdemoosa & Dindarloo (2015) [31]	Neural Network (NN)	Predicted mining trucks' fuel and energy usage using artificial intelligence and data-driven methods. NN was tested with promising results to anticipate fuel usage per operational cycle based on truck payload, loading and unloading times, and idle while loading and emptying times.
Bastos et al. (2011) [32]	Time Dependent MDP	Presented a time-dependent MDP-based basic combinatorial formulation for truck dispatching. Developed a single-dependent agent technique to make multi-agent issues easier to understand. A two-phase system was used that combined real-time dispatching and offline MDP model solving. The simulation, conducted in SimEvents®, involved 15 trucks, one crusher, three shovels, and a 10-hour shift scenario. The results showed a slight improvement over baseline performance, on average.
Paduraru & Dimitrakopoulos (2017, 2019) [33; 34]	Reinforcement Learning (RL)	Trained an RL agent to discover the best destination choices for every mining block for a particular production schedule, introducing RL for mining operational decision-making
Lamghari & Dimitrakopoulos (2020) [35]	Hyper-Heuristic Framework	RL principles were reintroduced in a hyper-heuristic framework for long-term open-pit mining scheduling (OPS). This hyper-heuristic approach selects from a range of heuristic options, iteratively determining which heuristic yields the best solution.

Table 1 continued

Authors	Approach Used	Explanation
Choi et al. (2020) [36]	Support Vector Machine (SVM)	Suggested an alternative method of production forecasting based on supervised learning strategies. A large dataset, collected by IoT devices deployed in an open-pit mine, was analyzed to predict ore production. SVM demonstrated the best performance among the approaches examined. The study emphasized the potential for making the most of the data produced by mine monitoring equipment.
Kumar et al. (2020) and Kumar & Dimitrakopoulos (2021) [37; 38]	Reinforcement Learning (RL)	Expanded the study to consider fresh data gathered in real time using sensors and other monitoring tools. The focus was on integrating fresh data related to mineral grades and properties, emphasizing RL's potential for mining systems that are adaptable and self-learning.
Icarte et al. (2020)	Contract Net Protocol (CNP)	Suggested an innovative method of truck dispatching that makes use of a multi-agent system in which separate intelligent agents represent trucks, shovels, and unloading locations. A CNP was used for the truck-shovel interaction. The method was benchmarked using a DES from a copper mine in Chile, comparing it against heuristic and quantitative optimization models. The results showed that production targets were met while reducing operational expenses by 18%. The study was further extended to optimally reschedule trucks in response to machine failures.
Zhang et al. (2020) [39]	Deep Q-Network (DQN)	Presented a multi-agent reinforcement learning system for the dynamic management of diverse mining fleets. Used a typical DQN with experience sharing and memory tailoring, utilizing a single network with shared parameters for both decentralized execution and centralized learning. Fifty trucks of three different sorts, three shovels, and three dump trucks were used to simulate a 12-hour shift in SimPy ®. Compared to a simple heuristic, productivity increased by almost 5%.

Table 1 continued

Authors	Approach Used	Explanation
De Carvalho and Dimitrakopoulos (2021) [40]	Double Deep Q-Network (DDQN)	Integrated geological uncertainty and processing plant aims to dispatch based on reinforcement learning. used a DDQN model in conjunction with a DES to make decisions. The elaborately modelled copper-gold mining complex included two pits, four shovels, twelve trucks, a sizable waste dump, a mill, and a leach pad. Exceeded baselines by 12–16% and 20%–23% in terms of copper and gold production, respectively.
Huo et al. (2023) [41]	Q-Learning	Assessed emission reduction capabilities of RL-based dispatching. Applied the standard Q-learning algorithm with a limited state space. Defined a reward function addressing six aspects, including correct delivery of materials and timely maintenance. Simulated a hypothetical open-pit mine with three shovels, varying truck numbers, a mill, a garage, and a waste dump in OpenAI®. Achieved up to 43% higher production and up to 37% less greenhouse gas (GHG) emissions compared to baselines.
Avalos & Ortiz (2023) [42]	Deep Q-Learning (DQL)	Suggested using a DQL algorithm to optimize the long-term open-pit mine design while considering metal grade uncertainty and metallurgical factors. The DQL algorithm selects extraction blocks at daily time intervals to maximize the net present value (NPV) of the mining project, while adhering to geotechnical constraints. Demonstrated how DRL techniques can be used to address uncertainties and incorporate a variety of variable types while producing useful outcomes.
Levinson et al. (2023) [43]	Actor-Critic RL Algorithm	An Actor-Critic RL system was proposed for the stochastic optimization of planning decisions, such as the extraction sequence, destination policy, stockpiling, and preconcentration of a copper open-pit mining complex. Under the unpredictability in metal quality, the RL agent effectively learns a method to optimize short-term planning decisions.

These papers thoroughly summarize the many techniques and developments in RL-based FMS in open-pit mines, presenting diverse approaches, simulation settings, and accomplished results. When taken as a whole, they demonstrate the great potential of RL and other clever approaches to maximize operational decision-making, boost productivity, and cut expenses in mining operations.

In this specific area, the open-pit production planning and scheduling problem has been the focus of the great bulk of the study. To address the challenges of large-scale production scheduling for the life-of-mine strategic plan, a great deal of effort has gone into creating metaheuristics and intelligent computation methods. Short- and long-term planning has received the greatest focus among the many difficulties.

In context of open-pit strategic planning, ML has experienced exponential growth with successful applications across various areas of interest. The use of these intelligent algorithms can handle the large and complex problems of mine planning and scheduling. Also, DES, which uses vast amounts of past mining data to build digital twins of mining operations and aid in decision-making, is gaining popularity. The rise in supervised learning and RL approaches further demonstrates the potential for ML adoption in operational management activities. These developments highlight how AI and data-driven approaches are revolutionizing the efficacy and efficiency of strategic planning in surface mining operations.

4. Way Forward

Noriega and Pourrahimian [26; 44] thoroughly analyzed the RL framework for shovel allocation and truck fleet dispatching. An AI agent was able to communicate with and get feedback from the system using a digital twin or simulation model of the system. The agent acquired and trained an ideal stochastic decision-making policy with the aid of this technique. The research suggested a virtual environment for the open-pit production system utilizing a DES model. To enable the DRL agent to engage with the model and learn the best decision-making techniques to accomplish certain objectives, uncertainties must be included throughout the loading and hauling operational cycles. The authors presented the DRL framework for dynamic truck dispatching and shovel allocation planning. A case study replicating a month of production in a DES was used to test this approach. The outcomes demonstrated that the agent successfully picked up on the dynamics of the open-pit mining environment, consistently accomplishing the predetermined objectives in various scenarios.

An extensive data set from the industry—more precisely, from an oil-sand mining operation—has been gathered. This raw data must be carefully filtered and cleaned at the project's current phase to guarantee its quality and appropriateness for using machine learning techniques. The allocation of shovels and truck fleet dispatching, two crucial tasks for raising mining operations' efficiency, will be optimized with the help of these algorithms.

A earlier study that was cited in [26] examined data from the mining operation for a month. Although this yielded insightful information, it was found that more datasets are necessary to improve DRL algorithm performance. The algorithms will be able to learn more efficiently with a larger dataset since it will capture a wider variety of operational circumstances.

Expanding the dataset to include a broader range of operational situations and a longer time period is the next step in the future studies This will boost not only the DRL algorithms' training process but also their generalization and dependability in real-world applications. Further studiesendeavors to enhance operational efficiency and productivity in the mining industry by developing more robust and efficient optimization methods for truck fleet dispatching and shovel allocation, through the utilization of a more comprehensive dataset.

5. References

[1] Hartman, H.L. and J.M., Mutmansky, Introductory Mining Engineering. 2002.

[2] Fisonga, M., and V., Mutambo, (2017). *Optimization of the fleet per shovel productivity in surface mining: Case study of Chilanga Cement, Lusaka Zambia.* Cogent Engineering, **4**.

[3] Gamache, M., and S., Alarie, *Overview of solution strategies used in truck dispatching systems for open-pit mines.* International Journal of Surface Mining, Reclamation and Environment, 2002. **16**, pp. 59-76.

[4] Czaplicki, J.M., Shovel-Truck Systems: Modelling, Analysis and Calculations. 2008.

[5] Whittle, D., Open-Pit Planning and Design. 2011. pp. 25.

[6] Elijah, K., J., Githiria, M., Samuel, and D., Mauti, *Optimisation of shovel-truck haulage system in an open-pit using queuing approach.* Arabian Journal of Geosciences, 2021. **14.**

[7] Blom, M., A., Pearce, and P., Stuckey, *A decomposition-based algorithm for the scheduling of open-pit networks over multiple time periods.* Management Science, 2016. **62.**

[8] Mnih, V., K., Kavukcuoglu, D., Silver, A.A., Rusu, J., Veness, M.G., Bellemare, A., Graves, M., Riedmiller, A.K., Fidjeland, G., Ostrovski, S., Petersen, C., Beattie, A.,Sadik, I., Antonoglou, H., King, D., Kumaran, D., Wierstra, S., Legg, and D., Hassabis, *Human-level control through deep reinforcement learning.* Nature, 2015. **518**(7540), pp. 529-533.

[9] Sutton, R.S., and A.G., Barto, Reinforcement Learning: An Introduction. 2018.

[10] Panzer, M., and B., Bender, B., *Deep reinforcement learning in production systems: a systematic literature review.* International Journal of Production Research, 2021. **60**, pp. 1-26.

[11] Fu, Y., and C., Aldrich, *Deep learning in mining and mineral processing operations: A review.* IFAC-PapersOnLine, 2020. **53**(2), pp. 11920-11925.

[12] Osanloo, M., J., Gholamnejad, and B., Karimi, *Long-term open-pit mine production planning: A review of models and algorithms.* International Journal of Mining, Reclamation and Environment, 2008. **22**, pp. 3-35.

[13] Newman, A.M., E., Rubio, R., Caro, A., Weintraub, and K., Eurek, *A review of operations research in mine planning.* Interfaces, 2010. **40**, pp. 222-245.

[14] Zhang, C., (2019). *Integrating machine learning, optimization and simulation to increase equipment utilization: Use case study on open-pit mines.* Hitachi: Industrial AI blog, 2019.

[15] Chanda, E., and S., Gardiner, *A comparative study of truck cycle time prediction methods in open-pit mining.* Engineering, Construction and Architectural Management, 2010. **17**, pp. 446-460.

[16] Yao, J., Z., Wang, H., Chen, W., Hou, X., Zhang, X., Li, and W., Yuan, *Open-pit mine truck dispatching system based on dynamic ore blending decisions.* Sustainability, 2023. **15**(4), p. 3399.

[17] Dayo-Olupona, O., B., Genc, T., Celik, and S., Bada, *Adoptable approaches to predictive maintenance in mining industry: An overview.* Resources Policy, 2023. **86**, p. 104291.

[18] Burt, C.N., and L., Caccetta, *Match factor for heterogeneous truck and loader fleets.* International Journal of Mining, Reclamation and Environment, 2007. **21**(4), pp. 262-270.

[19] Song, Z., X., Li, R., Huo, and L., Liu, (2024). *Intelligent early-warning platform for open-pit mining: Current status and prospects.* Rock Mechanics Bulletin, 2024. **3**, p. 100098.

[20] Moradi, A., S., Upadhyay, and H., Askari Nasab, *Truck dispatching in surface mines - Application of fuzzy linear programming.* Journal of the Southern African Institute of Mining and Metallurgy, 2021. **121**, pp. 1-8.

[21] Moradi Afrapoli, A., and H., Askari-Nasab, *Mining fleet management systems: A review of models and algorithms.* International Journal of Mining, Reclamation and Environment, 2019. **33**, pp. 42-60.

[22] Chaowasakoo, P., H., Seppälä, H., Koivo, and Q., Zhou, *Digitalization of mine operations: Scenarios to benefit in real-time truck dispatching.* International Journal of Mining Science and Technology, 2017. **27**(2), pp. 229-236.

[23] Eivazy, H., and H., Askari-Nasab, *A mixed integer linear programming model for short-term open-pit mine production scheduling.* Mining Technology, 2012. **121**(2), pp. 97-108.

[24] L'Heureux, G., M., Gamache, and F., Soumis, *Mixed integer programming model for short term planning in open-pit mines.* Mining Technology, 2013. **122**(2), pp. 101-109.

[25] Kozan, E., and S.Q., Liu, *A new open-pit multi-stage mine production timetabling model for drilling, blasting and excavating operations.* Mining Technology, 2016. **125**, pp. 47-53.

[26] Noriega, R., and Y., Pourrahimian, *Shovel allocation and scheduling for open-pit mining using deep reinforcement learning.* International Journal of Mining, Reclamation and Environment, 2024. **38,** pp. 1-18.

[27] Upadhyay, S.P., and H., Askari-Nasab, *Dynamic shovel allocation approach to short-term production planning in open-pit mines.* International Journal of Mining, Reclamation and Environment, 2017. **33**, pp. 1-20.

[28] Upadhyay, S.P., and H., Askari-Nasab, *Simulation and optimization approach for uncertainty-based short-term planning in open-pit mines.* International Journal of Mining Science and Technology, 2018. **28**(2), pp. 153-166.

[29] Shishvan, M.S., and J., Benndorf, *Simulation-based optimization approach for material dispatching in continuous mining systems.* European Journal of Operational Research, 2019. **275**(3), 1pp. 108-1125.

[30] Hazrathosseini, A., and A., Moradi Afrapoli, *Transition to intelligent fleet management systems in open-pit mines: A critical review on application of reinforcement-learning-based systems.* Mining Technology, 2024. **133**, pp. 50-73.

[31] Siami-Irdemoosa, E., and S.R., Dindarloo, *Prediction of fuel consumption of mining dump trucks: A neural networks approach.* Applied Energy, 2015. **151**, pp. 77-84.

[32] Bastos, G.S., L.E., Souza, F.T., Ramos, and C.H., Ribeiro, *A single-dependent agent approach for stochastic time-dependent truck dispatching in open-pit mining*. in *Proceedings of 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), IEEE.* 2011. pp. 1057-1062.

[33] Paduraru, C., and R., Dimitrakopoulos, *Adaptive policies for short-term material flow optimization in a mining complex.* Mining Technology, 2017. **127**, pp. 56-63.

[34] Paduraru, C., and R., Dimitrakopoulos, *Responding to new information in a mining complex: Fast mechanisms using machine learning.* Mining Technology, 2019. **128**(3), pp. 129-142.

[35] Lamghari, A., and R., Dimitrakopoulos, *Hyper-heuristic approaches for strategic mine planning under uncertainty.* Computers & Operations Research, 2020. **115**, 104590.

[36] Choi, Y., H., Nguyen, X.N., Bui, T., Nguyen-Thoi, and D., Park, *Estimating ore production in open-pit mines using various machine learning algorithms based on a truck-haulage system and support of internet of things.* Natural Resources Research, 2021. **30**(2), pp. 1141-1173.

[37] Kumar, A., R., Dimitrakopoulos, and M., Maulen, *Adaptive self-learning mechanisms for updating short-term production decisions in an industrial mining complex.* Journal of Intelligent Manufacturing, 2020. **31**, pp. 1795-1811.

[38] Kumar, A., and R., Dimitrakopoulos, *Production scheduling in industrial mining complexes with incoming new information using tree search and deep reinforcement learning.* Applied Soft Computing, 2021. **110**, pp. 107644.

[39] Zhang, C., P., Odonkor, S., Zheng, H., Khorasgani, S., Serita, C., Gupta, and H., Wang, *Dynamic dispatching for large-scale heterogeneous fleet via multi-agent deep reinforcement learning*, in *Proceedings of 2020 IEEE International Conference on Big Data (Big Data), IEEE*. 2020. pp. 1436-1441.

[40] de Carvalho, J.P., and R., Dimitrakopoulos, *Integrating production planning with truck-dispatching decisions through reinforcement learning while managing uncertainty.* Minerals, 2011. **11**(6), p. 587.

[41] Huo, D., Y.A., Sari, R., Kealey, and Q., Zhang, *Reinforcement learning-based fleet dispatching for greenhouse gas emission reduction in open-pit mining operations.* Resources, Conservation and Recycling, 2023. **188**, p. 106664.

[42] Avalos, S., and J.M., Ortiz, *Multivariate Geostatistical Simulation and Deep Q-Learning to optimize mining decisions.* Mathematical Geosciences, 2023. **55**(5), pp. 673-692.

[43] Levinson, Z., R., Dimitrakopoulos, and J., Keutchayan, *Simultaneous stochastic optimization of an open-pit mining complex with preconcentration using reinforcement learning.* Applied Soft Computing, 2023. **138**, p. 110180.

[44] Noriega, R., and Y., Pourrahimian, *Truck Fleet Dispatching Control in Open-Pit Mining Based on Reinforcement Learning and Discrete Event Simulation.* Society for Mining, Metallurgy and Exploration, 2023.