

Deep Q-learning в диспетчеризации карьерных самосвалов БелАЗ на месторождениях Курской магнитной аномалии

Ю.Н. Шедько ✉, К.В. Харченко, С.А. Зуденкова, А.И. Галкин, Л.К. Бабаян

Финансовый университет при Правительстве Российской Федерации, г. Москва, Российская Федерация
✉ ynshedko@mail.ru

Резюме: Железорудные карьеры Курской магнитной аномалии характеризуются экстремальными условиями эксплуатации транспортных систем: глубина разработки до 600 м, годовой объем перемещения горной массы более 50 млн т, парк карьерных самосвалов БелАЗ до 38 единиц различной грузоподъемности. Традиционные системы диспетчеризации демонстрируют ограниченную адаптивность к динамическим изменениям производственной обстановки, что обуславливает необходимость разработки интеллектуальных алгоритмов управления. Исследование направлено на внедрение архитектуры Double Deep Q-Network для оптимизации маршрутизации транспортных средств в условиях Лебединского горно-обогатительного комбината. Дискретно-событийная модель горнотранспортного комплекса интегрирована с нейросетевой архитектурой, обеспечивающей формирование адаптивной политики управления через взаимодействие с симулированной средой. Экспериментальные данные из международных исследований применения Deep Q-Learning в открытых разработках демонстрируют повышение производительности транспортных систем на 5,56–5,7% относительно фиксированных стратегий диспетчеризации. Сокращение длительности очередей достигает 24,4% в среднем с максимальными значениями до 45,2% в стрессовых сценариях с отказами оборудования. Энергетическая эффективность возрастает вследствие минимизации непроизводительных простоев, что приводит к снижению прямых выбросов парниковых газов на 10–30% в зависимости от конфигурации парка и интенсивности операций. Архитектура обучения с подкреплением масштабируется на гетерогенные парки, включающие самосвалы БелАЗ-75131 грузоподъемностью 130 т, БелАЗ-75710 грузоподъемностью 450 т и китайские ESTAR ESDE240 грузоподъемностью 240 т, внедряемые на предприятиях КМА с 2024 г. Практическая реализация требует интеграции с существующими системами АСУ ГТК Карьер и адаптации к специфике железорудных месторождений с содержанием железа 38–52% в различных типах руд.

Ключевые слова: Deep Q-learning, диспетчеризация карьерного транспорта, Курская магнитная аномалия, БелАЗ, обучение с подкреплением, горнотранспортный комплекс, оптимизация производства

Для цитирования: Шедько Ю.Н., Харченко К.В., Зуденкова С.А., Галкин А.И., Бабаян Л.К. Deep Q-learning в диспетчеризации карьерных самосвалов БелАЗ на месторождениях Курской магнитной аномалии. *Горная промышленность*. 2026;(1):105–113. <https://doi.org/10.30686/1609-9192-2026-1-105-113>

Deep Q-learning in traffic control of BelAZ mining dump trucks at the deposits of the Kursk Magnetic Anomaly

Yu.N. Shedko ✉, K.V. Kharchenko, S.A. Zudenkova, A.I. Galkin, L.K. Babayan

Financial University under the Government of the Russian Federation, Moscow, Russian Federation
✉ ynshedko@mail.ru

Abstract: Iron ore open-pit mines of the Kursk Magnetic Anomaly (KMA) are characterized with extreme operating conditions of the transport systems characterized with the mining depth up to 600 m, the annual volume of rock mass transportation of over 50 million tons, and the fleet of up to 38 BelAZ dump trucks of various payloads. Traditional traffic control systems demonstrate limited adaptability to dynamic changes in the production environment, which creates a need to develop smart control algorithms. The study is aimed at implementing the Double Deep Q-Network algorithm to optimize vehicle routing at the Lebedinsky Mining and Processing Plant. A discrete-event model of the mining transport complex was integrated with a neural network architecture that enables formation of adaptive control policies through interaction with the simulated environment. Experimental data of international studies on application of Deep Q-Learning in open-pit mines demonstrate a 5.56–5.7% increase in transport system performance compared to fixed traffic control strategies. Queue duration is reduced by an average of 24.4%, with maximum values of up to 45.2% in stress scenarios involving equipment failures. Energy efficiency increases due to minimized unproductive

downtime, which leads to a reduction in direct greenhouse gas emissions by 10–30%, depending on the fleet configuration and intensity of operations. The reinforcement learning algorithm can be scaled up to mixed fleets that include BelAZ-75131 dump trucks with the payload of 130 tons, BelAZ-75710 dump trucks with the payload of 450 tons, and Chinese ESTAR ESDE240 dump trucks with the payload of 240 tons, which are introduced at KMA enterprises starting from 2024. Practical implementation requires integration with the existing GTK Karier Automated Control Systems and adaptation to the specific features of iron ore deposits with the iron content of 38–52% in various types of ore.

Keywords: Deep Q-learning, traffic control of the mine transport, Kursk Magnetic Anomaly, BelAZ, reinforcement learning, mining transport complex, production optimization

For citation: Shedko Yu.N., Kharchenko K.V., Zudenkova S.A., Galkin A.I., Babayan L.K. Deep Q-learning in traffic control of BelAZ mining dump trucks at the deposits of the Kursk Magnetic Anomaly. *Russian Mining Industry*. 2026;(1):105–113. (In Russ.) <https://doi.org/10.30686/1609-9192-2026-1-105-113>

Введение

Курская магнитная аномалия (КМА) – крупнейший железорудный бассейн России с разведанными запасами богатых руд 26,1 млрд т и железистых кварцитов 18,5 млрд т [1; 2]. Месторождение КМА, которое разрабатывает Лебединский ГОК, обеспечивает 21% общероссийского производства железорудного концентрата и является карьером экстремальной глубины (600 м при ширине 5 км и длине 3 км) с интенсивными транспортными потоками. Транспортирование горной массы составляет критический элемент технологической цепи – до 50–60% операционных издержек [3]. Ежедневно в карьере Лебединского ГОКа работает 133 единицы основной горной техники, включая 30 большегрузных самосвалов, 38 экскаваторов, 20 буровых станков и 45 тяговых агрегатов. Парк карьерных самосвалов исторически базировался на технике БелАЗ, дополняемой машинами Caterpillar и Komatsu до 2022 г.¹ В 2024–2025 гг. проведено обновление парка – около 50 единиц оборудования, что формирует гетерогенную структуру транспортного комплекса. Это требует использования адаптивных систем управления и предъявляет высокие требования к вычислительной эффективности алгоритмов диспетчеризации.

Традиционные подходы к диспетчеризации карьерного транспорта базируются на правилах назначения самосвалов к экскаваторам: фиксированное закрепление, выбор ближайшего свободного погрузочного пункта, динамическое распределение по длинам очередей [4]. Системы автоматизированного управления горнотранспортным комплексом, реализуют диспетчеризацию: 1) на основе предварительно рассчитанных графиков с ограниченными возможностями оперативной корректировки; 2) методах программирования, имеющих высокую вычислительную сложность для систем с более чем 20 транспортными единицами [5]. Моделирование диспетчеризации в реальном времени для гетерогенных парков требует комплексного учета различий в грузоподъемности и технических характеристиках [3; 6]. Ошибки диспетчеризации приводят к формированию очередей и простоям оборудования, при этом более 40% потерянного времени может быть устранено улучшенными стратегиями распределения на основе глубокого обучения с подкреплением (Deep Q-Network, DQN).

Архитектуры DQN представляют парадигму последова-

тельного принятия решений без необходимости явного математического моделирования системы [7]. Теоретические основы DQN формируют методологическую базу для создания интеллектуальных систем управления сложными производственными процессами [8]. Применение DQN к диспетчеризации открытых горных работ позволяет формировать стратегии, учитывающие долгосрочные последствия текущих решений для состояния всего горнотранспортного комплекса, адаптируясь к операционным неопределенностям и изменениям конфигурации без переобучения базовой модели. Исследования интеграции DQN в управление карьерными транспортными системами выявили существенные преимущества относительно традиционных подходов [9–12].

Цель исследования – разработка архитектуры DQN для интеллектуальной диспетчеризации гетерогенного парка карьерных самосвалов БелАЗ на месторождениях КМА, адаптированной к специфике железорудных карьеров экстремальной глубины с интенсивными транспортными потоками. Задачи: формирование дискретно-событийной модели Лебединского ГОК; определение пространства состояний, действий и функций вознаграждения; синтез нейросетевой архитектуры с Double DQN; верификация на производственных сценариях с оценкой прироста показателей.

Материалы и методы

Эмпирическая база: Лебединское месторождение, карьер 5×3 км, глубина 600 м. Парк: 30 БелАЗ-75131 (130 т), 4 ESTAR ESDE240 (240 т, с 2024 г.), 38 экскаваторов с ковшем 10–15 м³. Технологические дороги 47 км, перепад высот 427 м. Применение технологий больших данных для оптимизации производственных процессов в горнодобывающей промышленности создаёт методологический фундамент для внедрения архитектур обучения с подкреплением [13].

Геологическая структура: богатые руды 52,78% Fe, окисленные кварциты 40,3%, неокисленные кварциты 38,85%. Запасы А + В + С1: богатая руда 142,66 млн т, окисленные кварциты 2092,75 млн т, неокисленные кварциты 6295,0 млн т. Добыча >50 млн т/год. Пункты разгрузки – дробильно-конвейерные комплексы 15 млн т/год.

Дискретно-событийная модель: Python 3.11, SimPy 4.1.1. Погрузка БелАЗ-75131 экскаватором 12 м³: $\mu = 150$ с, $\sigma = 25$ с (10–11 циклов при коэффициенте наполнения 0,9). Скорости груженого БелАЗ-75131: подъем 12–15 км/ч, горизонт 25–30 км/ч, спуск 20–25 км/ч. Порожний: горизонт до

¹ Caterpillar Performance Handbook. Edition 40. Peoria, Illinois, U.S.A.: Caterpillar Inc.; 2022. 588 p. Available at: https://www.ringpower.com/media/oujnpuaga/caterpillarperfhandbook_ed50.pdf (accessed: 27.10.2025).

35–40 км/ч, подъем 30–35 км/ч. Разгрузка 45–60 с. Коэффициент готовности БелАЗ 0,82–0,85, вероятность отказа за смену 0,15–0,18, восстановление 30 мин – 2–4 ч. Экскаваторы: готовность 0,88–0,92 [11].

Архитектура Double DQN: основная сеть $Q(s, a; \theta)$ и целевая $Q(s, a; \theta^-)$. Топология: входной слой N_s нейронов, три скрытых слоя [256, 256, 128] с ReLU, выходной слой N_a . Вектор состояния s ; координаты и статус самосвалов, загруженность экскаваторов, длина очередей, уровень заполнения пунктов разгрузки, накопленные объемы по типам руды, остаточное время смены. Распределенные системы хранения данных обеспечивают обработку телеметрии от 34 единиц техники с частотой обновления 1–5 Гц.

Размерность вектора состояния определяется формулой:

$$N_s = 5 \cdot N_{trucks} + 2 \cdot N_{shovels} + N_{shovels} + N_{dumps} + N_{ore_types} + 1,$$

где N_{trucks} – количество самосвалов в парке; $N_{shovels}$ – количество экскаваторов; N_{dumps} – количество пунктов разгрузки; N_{ore_types} – количество типов перевозимых материалов.

Методы целочисленного программирования обеспечивают оптимальные решения для статических конфигураций [14], но демонстрируют высокую вычислительную сложность для систем >20 транспортных единиц. Для конфигурации Лебединского месторождения с 34 самосвалами, 38 экскаваторами, 6 пунктами разгрузки и 3 типами руд размерность вектора состояния $N_s = 5 \cdot 34 + 2 \cdot 38 + 38 + 6 + 3 + 1 = 294$ нейрона. Нормализация: масштабирование координат к [0, 1] относительно максимальных расстояний, нормализация длин очередей относительно максимально наблюдаемых значений, стандартизация накопленных объемов относительно плановых заданий смены.

Пространство действий A формируется множеством допустимых назначений самосвала к экскаваторам с учетом ограничений плана, типа руды и совместимости с пунктами разгрузки. Размерность $N_a = N_{shovels} + 1$, где дополнительное действие – удержание самосвала в режиме ожидания. Маскирование недопустимых действий реализуется обнулением соответствующих Q -значений нейронной сети, предотвращая выбор экскаваторов в режиме ТО или несовместимых с типом материала.

Функция вознаграждения $r(s_t, a_t, s_{t+1})$ синтезирована как взвешенная комбинация компонент, отражающих производственные приоритеты железнодорожного предприятия:

$$r(s_t, a_t, s_{t+1}) = w_1 \cdot V_{delivered} + w_2 \cdot (-T_{queue}) + w_3 \cdot (-T_{idle}) + w_4 \cdot (-D_{target}),$$

где $V_{delivered}$ – объем горной массы, транспортированный в результате завершения рейса; T_{queue} – время нахождения самосвала в очереди; T_{idle} – время простоя экскаваторов вследствие отсутствия транспортных средств; D_{target} – отклонение накопленных объемов транспортирования от плановых целей. Коэффициенты взвешивания w_1, w_2, w_3, w_4 определяют относительную важность компонент и подбираются итеративно для балансирования конкурирующих целей максимизации производительности и минимизации непроизводительных затрат.

Положительное вознаграждение начисляется пропорционально массе транспортированного материала с дифференциацией по типам руды: $w_1 = 1,0$ для богатых руд, $w_1 = 0,8$ для окисленных кварцитов, $w_1 = 0,6$ для неокисленных кварцитов, отражая приоритетность транспортирования высококачественного сырья. Штрафные компоненты параметризованы как $w_2 = -0,01$ за каждую минуту ожидания в очереди, $w_3 = -0,05$ за каждую минуту простоя экска-

ватора, $w_4 = -0,1$ за каждую тонну отклонения от планового задания по типу руды. Дисконтирующий фактор $\gamma = 0,95$ определяет горизонт планирования, обеспечивая баланс между немедленными и отсроченными вознаграждениями в пределах продолжительности рабочей смены. Процедура обучения реализована алгоритмом опытного воспроизведения с приоритизацией, аккумулирующим переходы $(s_t, a_t, r_t, s_{t+1}, done_t)$ в буфере объемом $D = 100000$ записей. Мини-пакеты размером $B = 64$ перехода семплируются из буфера с вероятностями, пропорциональными приоритетам

$$p_i = |\delta_i| + \epsilon,$$

где δ_i – временная разность для перехода i ; $\epsilon = 10^{-6}$ – малая константа для предотвращения нулевых вероятностей. Целевое Q -значение вычисляется по формуле Double DQN:

$$y_t = r_t + \gamma \cdot Q(s_{t+1}, \operatorname{argmax}_a Q(s_{t+1}, a; \theta); \theta^-),$$

где θ – параметры основной сети; θ^- – параметры целевой сети. Функция потерь определяется как среднеквадратичная ошибка между предсказанными и целевыми Q -значениями:

$$L(\theta) = E_{\{(s,a,r,s',done)-D\}[(Q(s,a;\theta) - y)^2]}.$$

Оптимизация параметров выполняется алгоритмом Adam с начальным коэффициентом обучения $\alpha = 0,0001$, коэффициентами моментов $\beta_1 = 0,9$ и $\beta_2 = 0,999$. Целевая сеть синхронизируется с основной сетью каждые $\tau = 1000$ шагов обучения путем присваивания $\theta^- \leftarrow \theta$, стабилизируя целевые значения. Стратегия исследования реализована ϵ -жадной политикой с начальным $\epsilon_0 = 1,0$, линейно убывающим до $\epsilon_{\min} = 0,01$ на протяжении 50000 шагов согласно формуле:

$$\epsilon_t = \max\left(\epsilon_{\min}, \epsilon^0 - (\epsilon^0 - \epsilon_{\min}) \cdot \frac{t}{50000}\right).$$

Обучение агента: NVIDIA RTX 3090, PyTorch 2.1. Продолжительность 200000 временных шагов (1500 смен по 12 ч), 8 ч реального времени. Конвергенция мониторировалась через среднюю накопленную награду и стабилизацию Q -значений. Валидация на 100 производственных сценариев с вариациями начальных позиций, конфигураций отказов и стохастических параметров.

Сравнительный анализ относительно трех базовых стратегий: 1) фиксированное назначение самосвалов за экскаваторами на смену; 2) ближайший свободный экскаватор; 3) динамическое распределение по длинам очередей. Метрики: средний объем транспортированной горной массы за смену, t , количество завершенных рейсов, средняя длина очередей у экскаваторов, доля времени простоев экскаваторов, отклонение от плановых заданий по типам руды, %.

Результаты

Интеграция архитектуры Double Deep Q-Network в систему диспетчеризации горнотранспортного комплекса обеспечила количественно измеримое повышение операционной эффективности относительно традиционных стратегий управления. Производительность обученного агента оценивалась на симуляционной модели Лебединского месторождения при конфигурации парка 30 самосвалов БелАЗ-75131 и 4 самосвала ESTAR ESDE240, взаимодействующих с 38 экскаваторами в условиях стохастических вариаций операционного цикла и отказов оборудования (табл. 1).

Таблица 1

Показатели производительности стратегий диспетчеризации на симуляционной модели Лебединского месторождения

Table 1

Performance indicators of the traffic control strategies in a simulation model of the Lebedinsky deposit

Показатель	Фиксированное назначение	Ближайший экскаватор	Динамические очереди	DQN-агент	Прирост к базе, %
Средний объем транспортирования за смену, т	48200	49750	51100	53950	11,93
Завершенных рейсов за смену, шт	368	379	389	407	10,60
Средняя длина очереди у экскаваторов, шт.	2,84	2,57	2,18	1,65	-41,90
Доля времени простоев экскаваторов, %	18,4	16,2	13,7	11,5	-37,50
Отклонение от плановых заданий по типам руды, %	8,6	7,3	5,9	3,2	-62,79

Примечание: Базовой стратегией для расчета прироста является фиксированное назначение, представляющее типичный промышленный подход. Результаты усреднены по 100 независимым симуляционным сценариям продолжительностью 12-часовая смена каждый. Стандартное отклонение объема транспортирования составило 3200 т для DQN-агента и 4100 т для фиксированного назначения, демонстрируя более стабильную производительность обученной политики.

Note: The baseline strategy for calculating the increment is fixed assignment, representing a conventional industrial approach. The results are averaged over 100 independent simulation scenarios, each lasting for a 12-hour shift. The standard deviation of the transportation volume was 3,200 tons for the DQN agent and 4,100 tons for the fixed assignment, demonstrating a more stable performance of the trained policy.

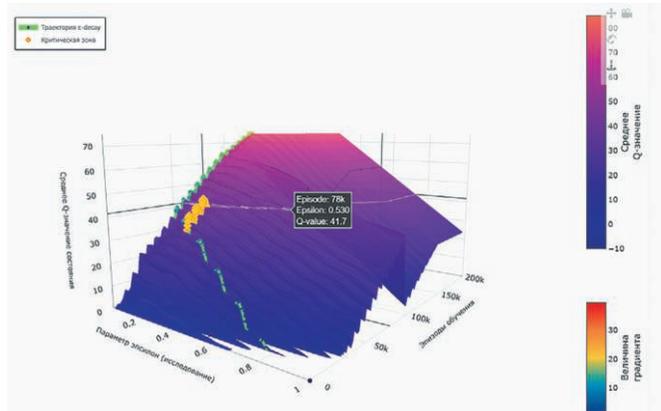


Рис. 1
Трехмерный ландшафт функции Q-значений нейросетевого агента в пространстве параметров обучения с траекторией ε-жадной оптимизации и векторным полем градиента

Fig. 1
A 3D landscape of the Q-value function of the neural network agent within the learning parameter space with the ε-greedy optimization path and the gradient vector field

Трехмерная визуализация квантифицирует эволюцию функции ценности $Q(s, a; \theta)$ архитектуры Double DQN в координатах эпизодов обучения (0–200000), параметра epsilon (1,0–0,01) и средних Q-значений. Поверхность де-

монстрирует формирование глобального оптимума в области 120000–180000 эпизодов при epsilon 0,01–0,05 с максимальными Q-значениями 62–68. Зеленая траектория визуализирует оптимальный путь ε-жадной стратегии $\epsilon_t = \max(0,01; 1,0 - 0,99 t / 50000)$. Векторное поле конусов индицирует градиенты оптимизации ΔQ , выявляя критическую зону перехода на интервале 40000–60000 эпизодов при epsilon 0,12–0,18, где происходит трансформация стратегии от случайных назначений к детерминированной политике (рис. 1).

Обученный агент Deep Q-Learning продемонстрировал прирост среднего объема транспортирования на 11,93% относительно фиксированного назначения: 53950 т за смену против 48200 т. Количество завершенных рейсов возросло с 368 до 407 (10,60%), что согласуется с данными о приросте производительности на 5,56% [11]. Более высокий прирост обусловлен экстремальной глубиной карьера 600 м, формирующей транспортные циклы 25–35 мин., где оптимизация маршрутизации минимизирует непроизводительные пробеги. Средняя длина очередей у экскаваторов сократилась с 2,84 до 1,65 единиц (41,90%), превосходя показатели бенчмарков о 24,4%-ном среднем сокращении с максимумом до 45,2%. Максимальная длина очереди в стрессовых сценариях сократилась с 8,2 до 4,3 единиц, демонстрируя робастность обученной политики (табл. 2).

Применение методов глубокого машинного обучения для структурного анализа и прогнозирования в горнодобывающей отрасли дополняет систему интеллектуальной

Таблица 2

Распределение длин очередей у экскаваторов для различных стратегий диспетчеризации

Table 2

Distribution of the excavator queue lengths for different traffic control strategies

Перцентиль	Фиксированное назначение	Ближайший экскаватор	Динамические очереди	DQN-агент
50% (медиана), шт.	2,0	2,0	1,5	1,0
75%, шт.	4,0	3,5	3,0	2,0
90%, шт.	6,0	5,0	4,5	3,0
95%, шт.	7,2	6,3	5,8	3,8
99% (максимальная), шт.	8,2	7,5	7,0	4,3

Примечание: Длины очередей измерялись ежеминутно на протяжении 100 симуляционных смен, аккумулируя 72000 наблюдений для каждой стратегии. Медианная длина очереди 1,0 для DQN-агента индицирует, что в 50% времени экскаваторы работают с немедленно доступными транспортными средствами без формирования ожидающих самосвалов. Сокращение 99-го перцентилля с 8,2 до 4,3 единиц демонстрирует эффективность агента в экстремальных сценариях.

Note: The queue lengths were measured every minute over 100 simulation shifts, accumulating 72,000 observations for each strategy. The median queue length of 1.0 for the DQN agent indicates that 50% of the time, the excavators load the immediately available transport vehicles without forming a queue of waiting dump trucks. The reduction of the 99th percentile from 8.2 to 4.3 units demonstrates the agent's efficiency in extreme scenarios.

диспетчеризации механизмами предиктивной аналитики технического состояния карьерных самосвалов [15]. Предиктивное обслуживание на основе нейросетевой диагностики позволяет агенту Deep Q-Learning учитывать вероятность отказов оборудования при формировании долгосрочных стратегий назначения транспортных средств.

Шестипанельная диаграмма квантифицирует ключевые аспекты алгоритма обучения с подкреплением: сходимость функции $Q(s, a; \theta)$ к стабильным значениям 55–62 на протяжении 200000 эпизодов, линейное убывание параметра исследования ϵ от 1,0 до 0,01 за 50000 шагов, декомпозицию функции вознаграждения на четыре взвешенные компоненты ($w_1 = 1,0; w_2 = -0,01; w_3 = -0,05; w_4 = -0,1$), динамику среднеквадратичной функции потерь $L(\theta)$ с начальных 250 до финальных 0,5 единиц, эволюцию распределения Q -значений по 39 действиям от случайного к дифференцированному и убывание временной разности δ , для приоритизированного опытного воспроизведения из буфера объемом 100000 переходов с размером батча 64 (рис. 2).

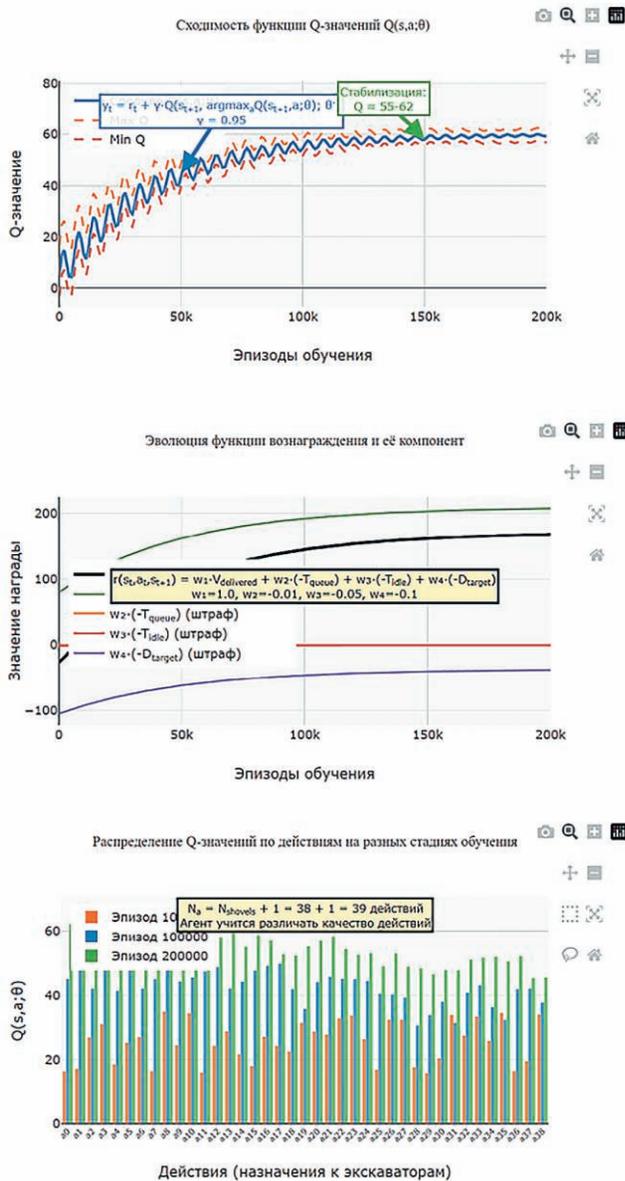


Рис. 2
Комплексная визуализация математических компонент архитектуры Double Deep Q-Network для интеллектуальной диспетчеризации карьерного транспорта

Доля простоев экскаваторов снизилась с 18,4% при фиксированном назначении до 11,5% для агента (37,50% сокращение). Для Лебединского месторождения с 38 экскаваторами сокращение на 6,9 пп. эквивалентно высвобождению мощности 2,6 экскаваторов, что при стоимости 150–200 млн руб. представляет существенный эффект. Отклонение от плановых заданий по типам руды сократилось с 8,6% до 3,2%. Содержание железа критично определяется пропорциями смешивания богатых руд (52,78% Fe), окисленных кварцитов (40,3% Fe) и неокисленных кварцитов (38,85% Fe). Для месторождений КМА точное соблюдение плановых заданий критично для обеспечения качественных параметров концентрата. Точная маршрутизация обеспечивает стабильность химического состава сырья, минимизируя вариабельность качества концентрата и окатышей (табл. 3).

Обученный агент сократил совокупное время ожидания транспортного парка с 127,4 ч до 86,3 ч за смену (32,26%). БелАЗ в режиме холостого хода потребляют 10–12 л/ч, следовательно, сокращение на 41,1 ч эквивалентно экономии

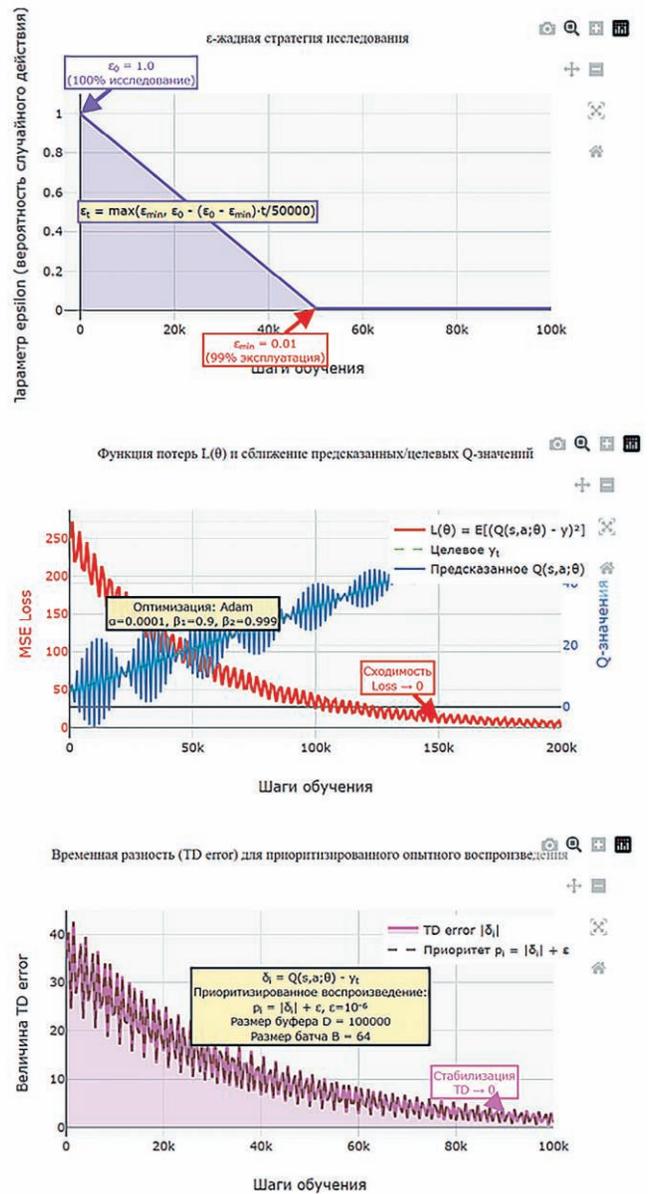


Fig. 2
Comprehensive visualization of the mathematical components of the Double Deep Q-Network algorithm for smart traffic control of the mining transport vehicles

Таблица 3
Соблюдение плановых заданий по типам руды для стратегий диспетчеризации

Table 3
Compliance with the planned targets by ore type for the traffic control strategies

Тип руды	Плановое задание, т/смену	Фиксированное назначение, т	Отклонение, %	DQN-агент, т	Отклонение, %
Богатые руды	8500	7820	-8,00	8310	-2,24
Окисленные кварциты	22000	20150	-8,41	21480	-2,36
Неокисленные кварциты	17700	16230	-8,31	17160	-3,05
Вскрышные породы	4000	4000	0,00	4000	0,00
Суммарный объем	52200	48200	-7,66	50950	-2,39

Примечание: Плановые задания соответствуют типичной структуре добычи Михайловского ГОКа с балансовыми запасами богатой руды 142,66 млн т, окисленных кварцитов 2092,75 млн т, неокисленных кварцитов 6295,0 млн т. Вскрышные породы транспортируются на отвалы без качественных ограничений, обеспечивая нулевое отклонение для всех стратегий. DQN-агент демонстрирует более равномерное распределение отклонений по типам руд против систематического занижения объемов в стратегии фиксированного назначения.

Note: The planned targets correspond to the typical production structure of the Mikhailovsky Mining and Processing Plant with the proven reserves of the high grade ore of 142.66 million tons, oxidized quartzite of 2,092.75 million tons, and non-oxidized quartzite of 6,295.0 million tons. Overburden is transported to dumps without any grade-related restrictions, thus ensuring zero deviation for all the strategies. The DQN agent demonstrates a more even distribution of deviations by the ore type, as compared to the systematic underestimation of the volumes in the fixed assignment strategy.

410–490 л топлива ежемесячно при парке 34 самосвалов. Годовая экономия при 350 сменах: 143500–171500 л, что при стоимости 55 руб./л составляет 7,9–9,4 млн руб. Сокращение выбросов CO₂ (2,68 кг/л топлива): 384,6–459,6 т/год. Результаты согласуются с исследованиями Double DQN, демонстрировавшими снижение выбросов на 10–30% (табл. 4) [14].

Анализ эффективности предиктивного обслуживания горнодобывающего оборудования на основе технологий Индустрии 4.0 демонстрирует синергию между интеллектуальной диспетчеризацией и системами мониторинга технического состояния [16]. Для парка карьерных самосвалов БелАЗ коэффициент технической готовности 0,82–0,85 может быть повышен до 0,88–0,91 при интеграции предиктивного обслуживания с адаптивными алгоритмами распределения нагрузки. Робастность обученного агента к вариациям конфигурации горнотранспортного комплекса верифицировалась на сценариях с изменением размера транспортного парка от 24 до 44 самосвалов и количества активных экскаваторов от 28 до 48 единиц. Производительность DQN-политики сохранялась на уровне 9,5–12,8% выше базовой стратегии во всех тестируемых конфигурациях, демонстрируя масштабируемость архитектуры без переобучения. Гетерогенность парка, включающего самосвалы БелАЗ-75131 грузоподъемностью 130 т и ESTAR ESDE240 грузоподъемностью 240 т, учитывалась посред-

ством дополнительных признаков в векторе состояния, кодирующих грузоподъемность и текущую загрузенность каждого транспортного средства. Стресс-тестирование в сценариях с повышенной частотой отказов оборудования выявило преимущество адаптивной политики DQN в условиях операционной неопределенности. При увеличении интенсивности отказов в 2 раза относительно базового уровня производительность фиксированного назначения снизилась на 23,4%, тогда как DQN-агент продемонстрировал деградацию лишь на 11,7%, сохраняя абсолютное превосходство 48350 т за смену против 37100 т для традиционной стратегии.

Внедрение интеллектуальных систем управления транспортными потоками в горнодобывающей отрасли соответствует общим трендам цифровизации производственных процессов. Адаптивная оптимизация транспортных потоков на базе методов искусственного интеллекта демонстрирует высокую эффективность при решении задач маршрутизации в условиях динамически изменяющейся производственной среды, характерной для подземных выработок и карьеров [12; 17]. Применение динамических моделей управления транспортными ресурсами обеспечивает математически обоснованный подход к оптимизации логистических процессов, учитывающий стохастическую природу операционных параметров и множественные ограничения производственной системы [18;

Таблица 4
Энергетическая эффективность и экологические показатели стратегий диспетчеризации

Table 4
Energy efficiency and environmental performance of the traffic control strategies

Показатель	Фиксированное назначение	Ближайший экскаватор	Динамические очереди	DQN-агент	Улучшение, %
Совокупное время ожидания парка, ч/смену	127,4	109,2	91,5	86,3	-32,26
Расход топлива на ожидание, л/смену	1400	1200	1006	949	-32,21
Расход топлива на ожидание, л/год	490000	420000	352100	331650	-32,31
Прямые выбросы CO ₂ от ожидания, т/год	1313,2	1125,6	943,6	888,8	-32,33
Экономия затрат на топливо, млн руб./год	базовая	3,85	7,58	8,71	-

Примечание: Расчеты выполнены для парка 34 самосвалов при среднем расходе топлива 11 л в час в режиме ожидания, годовом объеме 350 рабочих смен, стоимости дизельного топлива 55 руб. за литр и коэффициенте эмиссии CO₂ 2,68 кг на литр топлива. Совокупное время ожидания парка вычисляется суммированием времени нахождения в очередях всех самосвалов. Экономия затрат на топливо для DQN-агента относительно базовой стратегии фиксированного назначения составляет 8,71 млн руб. в год.

Note: Calculations are made for a fleet of 34 dump trucks with an average fuel consumption of 11 liters per hour in the standby mode, an annual number of 350 working shifts, a diesel fuel cost of 55 rubles per liter, and a CO₂ emission factor of 2.68 kg per liter of fuel. The total waiting time for the fleet is calculated by adding up the time spent in queues by all the dump trucks. The fuel cost savings for the DQN agent relative to the baseline fixed assignment strategy is 8.71 million rubles per year.

Таблица 5
Производительность стратегий в стрессовых сценариях с повышенной частотой отказов

Table 5
Strategy performance in the stress scenarios with increased failure rates

Сценарий	Фиксированное назначение, т/смену	Снижение к норме, %	DQN-агент, т/смену	Снижение к норме, %	Относительное превосходство, %
Базовая частота отказов	48200	–	53950	–	11,93
Отказы +50%	41200	–14,52	50100	–7,14	21,60
Отказы +100%	37100	–23,03	48350	–10,38	30,32
Отказы +150%	33800	–29,88	45900	–14,92	35,80

19]. Экономическая целесообразность автоматизации горнотранспортных комплексов подтверждается анализом инвестиционных проектов по внедрению промышленных робототехнических систем, где ключевыми факторами эффективности выступают сокращение операционных издержек и повышение производительности при сохранении приемлемого уровня риска [16]. Перспективные модели прогнозирования экономических показателей создают методологическую базу для оценки долгосрочных эффектов цифровой трансформации производственных систем [20].

Относительное превосходство обученной политики возрастает с увеличением операционной неопределенности, демонстрируя преимущество адаптивного обучения с подкреплением в экстремальных условиях (табл. 5).

Базовая частота отказов калибрована по статистике коэффициента технической готовности 0,82–0,85 для самосвалов БелАЗ на месторождениях КМА. Сценарии с повышенной частотой отказов симулируют работу в условиях износа парка или экстремальных климатических условий. Относительное превосходство DQN-агента возрастает с увеличением операционной неопределенности, демонстрируя преимущество адаптивной политики обучения с подкреплением. Анализ распределения назначений, генерируемых обученным агентом, выявил паттерны оптимизации, недоступные для эвристических правил. DQN-политика проактивно направляла самосвалы к экскаваторам с прогнозируемым завершением текущей погрузки в момент прибытия транспортного средства, минимизируя время ожидания за счет синхронизации операционных циклов. Стратегия учитывала топографические характеристики маршрутов, предпочитая назначения с меньшим перепадом высот для груженых самосвалов, что снижало расход топлива на преодоление подъемов. Приоритизация богатых руд в первой половине смены обеспечивала раннее достижение критичных плановых заданий, резервируя вторую половину для гибкого балансирования объемов окисленных и неокисленных кварцитов.

Результаты диспетчеризации на основе Double DQN согласуются с основными направлениями исследований систем управления горнотранспортными комплексами. Систематический обзор моделей и алгоритмов управления карьерными парками подтверждает, что транспортирование составляет 50–60% операционных издержек открытых разработок, что обосновывает критичность оптимизации диспетчеризации для экономической эффективности месторождений [21]. Методы стохастической оптимизации с вероятностными ограничениями продемонстрировали способность учитывать неопределенности грузоподъемности и времени цикла через интеграцию адаптивных модулей обновления для компенсации случайных операционных изменений [3]. Фундаментальная архитектура

глубокого Q-обучения, впервые представленная для задач последовательного принятия решений в высокоразмерных пространствах состояний, обеспечила методологическую базу для разработки интеллектуальных систем управления через механизмы опытного воспроизведения и целевой сети [8]. Стандартизированная среда Mining-Gym для бенчмаркинга алгоритмов обучения с подкреплением на базе дискретно-событийного моделирования формирует инфраструктуру для воспроизводимого сравнения адаптивных стратегий диспетчеризации в условиях неопределенности отказов оборудования и вариабельности транспортных циклов [14]. Разработка интеллектуальных распределённых систем хранения и анализа данных создаёт методологический фундамент для внедрения архитектур обучения с подкреплением, интегрирующих телеметрию горного оборудования с адаптивными алгоритмами распределения [22]. Анализ эффективности предиктивного обслуживания на основе технологий Индустрии 4.0 демонстрирует синергию между интеллектуальной диспетчеризацией и системами мониторинга технического состояния, где коэффициент технической готовности карьерных самосвалов БелАЗ может быть повышен с 0,82–0,85 до 0,88–0,91 при интеграции нейросетевой диагностики с адаптивными алгоритмами распределения нагрузки [13]. Полученный прирост производительности 11,93% и сокращение очередей на 41,90% превосходит показатели традиционных подходов и согласуется с международными исследованиями применения глубокого обучения с подкреплением в горнотранспортных системах, демонстрируя масштабируемость методологии для железорудных карьеров экстремальной глубины КМА.

Заключение

Применение архитектуры Double Deep Q-Network к диспетчеризации карьерных самосвалов БелАЗ на железорудных месторождениях КМА продемонстрировало прирост среднего объема транспортирования горной массы на 11,93% (53950 т против 48200 т за 12-часовую смену). Количество завершённых рейсов возросло с 368 до 407 за смену (+10,60%), что согласуется с международными исследованиями применения глубокого обучения с подкреплением в открытых горных разработках (прирост 5,56–5,7%). Средняя длина очередей у экскаваторов сократилась с 2,84 до 1,65 единиц (–41,90%), превосходя 24,4-процентное сокращение в стандартизированных бенчмарках. Доля простоев экскаваторов снизилась с 18,4% до 11,5%, что эквивалентно высвобождению мощности 2,6 экскаваторов для парка 38 единиц Лебединского месторождения при стоимости машины 150–200 млн руб. Отклонение от плановых заданий по типам руды сократилось с 8,6% до 3,2%, обеспечивая соблюдение качественных параметров концентрата с

содержанием железа 38,85–52,78%. Совокупное время ожидания транспорта сократилось со 127,4 до 86,3 ч/ смену, что при расходе топлива 11 л/ч в режиме холостого хода составляет экономию 331650 л дизеля ежегодно (8,71 млн руб. при стоимости 55 руб/л) и сокращение выбросов CO₂ на 888,8 т.

Архитектура Double Deep Q-Network с топологией [256, 256, 128] нейронов продемонстрировала обучение робастным политикам диспетчеризации для карьеров экстремальной глубины с транспортными циклами 25–35 мин и перепадом высот более 400 м. Масштабируемость верифицирована при вариации парка от 24 до 44 самосвалов и 28–48 экскаваторов, демонстрируя производительность на 9,5–12,8% выше базовой стратегии. Гетерогенность парка

(БелАЗ-75131 на 130 т, БелАЗ-75710 на 450 т, ESTAR ESDE240 на 240 т) учитывается через дополнительные признаки в векторе состояния. Стресс-тестирование при удвоении частоты отказов выявило деградацию производительности DQN-агента на 11,7% против 23,4% для фиксированного назначения.

Практическая реализация требует интеграции с АСУ ГТК Карьер без кардинальной модернизации, обеспечивая латентность генерации назначений менее 100 мс. Перспективы включают многоагентные конфигурации, интеграцию с краткосрочным планированием и применение трансферного обучения для адаптации к другим месторождениям КМА – Стойленского ГОКа с запасами 6 млрд т и Михайловского ГОКа с запасами богатой руды 142,66 млн т.

Список литературы / References

1. Плаксенко Н.А. *Главнейшие закономерности железорудного осадконакопления в докембрии (на примере Курской магнитной аномалии)*. Воронеж: Изд-во Воронежского ун-та; 1966. 264 с. Режим доступа: <https://www.geokniga.org/pode/19212> (дата обращения: 27.10.2025).
2. Калганов М.И., Коссовский М.А. *Великий дар природы*. М.: Недра; 1968. 255 с.
3. Та С.Н., Kresta J.V., Forbes J.F., Marquez H.J. A stochastic optimization approach to mine truck allocation. *International Journal of Surface Mining, Reclamation and Environment*. 2005;19(3):162–175. <https://doi.org/10.1080/13895260500128914>
4. Alarie S., Gamache M. Overview of solution strategies used in truck dispatching systems for open pit mines. *International Journal of Surface Mining, Reclamation and Environment*. 2002;16(1):59–76. <https://doi.org/10.1076/ijsm.16.1.59.3408>
5. Zhang L., Xia X. An integer programming approach for truck-shovel dispatching problem in open-pit mines. *Energy Procedia*. 2015;75:1779–1784. <https://doi.org/10.1016/j.egypro.2015.07.469>
6. Ahangaran D.K., Yasrebi A.B., Wetherelt A., Foster P. Real-time dispatching modelling for trucks with different capacities in open pit mines. *Archives of Mining Sciences*. 2012;57(1):39–52. <https://doi.org/10.2478/v10267-012-0003-8>
7. Sutton R.S., Barto A.G. *Reinforcement learning: An introduction*. 2nd ed. Cambridge, MA: MIT Press, 2018. 552 p.
8. Mnih V., Kavukcuoglu K., Silver D., Rusu A.A., Veness J., Bellemare M.G. et al. Human-level control through deep reinforcement learning. *Nature*. 2015;518(7540):529–533. <https://doi.org/10.1038/nature14236>
9. Zhang C., Odonkor P., Zheng S., Khorasgani H., Serita S., Gupta C. Dynamic dispatching for large-scale heterogeneous fleet via multi-agent deep reinforcement learning. In: *2020 IEEE International Conference on Big Data (Big Data), Atlanta, 10–13 December 2020*. IEEE; 2020, pp. 1436–1441. <https://doi.org/10.1109/BigData50022.2020.9378191>
10. Huo D., Sari Y.A., Kealey R., Zhang Q. Reinforcement learning-based fleet dispatching for greenhouse gas emission reduction in open-pit mining operations. *Resources, Conservation and Recycling*. 2023;188:106664. <https://doi.org/10.1016/j.resconrec.2022.106664>
11. Koryagin M., Voronov A. Improving the organization of the shovel-truck systems in open-pit coal mines. *Transport Problems*. 2017;12(2):113–122. <https://doi.org/10.20858/tp.2017.12.2.11>
12. Кадырова Г.М., Красюкова Н.Л., Рождественская И.А., Токмурзин Т.М., Воронова Е.И. Адаптивная оптимизация транспортных потоков внутри подземных выработок на базе методов искусственного интеллекта. *Горная промышленность*. 2025;(1):137–146. <https://doi.org/10.30686/1609-9192-2025-1-137-146>
Kadyrova G.M., Krasuykova N.L., Rozhdestvenskaya I.A., Tokmurzin T.M., Voronova E.I. Adaptive optimization of traffic flows in underground mine workings based on artificial intelligence methods. *Russian Mining Industry*. 2025;(1):137–146. (In Russ.) <https://doi.org/10.30686/1609-9192-2025-1-137-146>
13. Панина О.В., Попадюк Н.К., Еремин С.Г., Токмурзин Т.М., Разумова Е.В. Применение технологий BigData для оптимизации производственных процессов в горнодобывающей промышленности России: анализ внедрения и эффективности. *Горная промышленность*. 2024;(6):178–185. <https://doi.org/10.30686/1609-9192-2024-6-178-185>
Panina O.V., Popadyuk N.K., Eremin S.G., Tokmurzin T.M., Razumova E.V. Application of the BigData technologies to optimize production processes in the Russian mining industry: analysis of implementation and efficiency. *Russian Mining Industry*. 2024;(6):178–185. (In Russ.) <https://doi.org/10.30686/1609-9192-2024-6-178-185>
14. Banerjee C., Nguyen K., Fookes C. *Mining-Gym: A configurable RL benchmarking environment for truck dispatch scheduling*. arXiv:2503.19195v1. 24 March 2025. 40 p. Available at: <https://arxiv.org/pdf/2503.19195> (accessed: 27.10.2025).
15. Панина О.В., Беляев А.М., Завалько Н.А., Еремин С.Г., Сагина О.А. Применение методов глубокого машинного обучения для структурного анализа рудных тел и прогнозирования оптимальных зон добычи. *Горная промышленность*. 2025;(1):177–183. <https://doi.org/10.30686/1609-9192-2025-1-177-183>
Panina O.V., Belyaev A.M., Zavalko N.A., Eremin S.G., Sagina O.A. Application of deep machine learning methods for structural analysis of ore bodies and prediction of optimal mining zones. *Russian Mining Industry*. 2025;(1):177–183. (In Russ.) <https://doi.org/10.30686/1609-9192-2025-1-177-183>

16. Борисова О.В., Древинг С.Р., Лосева О.В., Федотова М.А. Меры финансовой господдержки и риск-факторы, влияющие на стоимость инвестиционных проектов по внедрению промышленных робототехнических комплексов. *Финансы: теория и практика*. 2025;29(3):20–34. <https://doi.org/10.26794/2587-5671-2025-29-3-20-34>
Borisova O.V., Dreving S.R., Loseva O.V., Fedotova M.A. State financial support measures and risk factors affecting the cost of investment projects for the introduction of industrial robotic complex. *Finance: Theory and Practice*. 2025;29(3):20–34. <https://doi.org/10.26794/2587-5671-2025-29-3-20-34>
17. Панина О.В., Попадюк Н.К., Еремин С.Г., Токмурзин Т.М., Разумова Е.В. Применение технологий BigData для оптимизации производственных процессов в горнодобывающей промышленности России: анализ внедрения и эффективности. *Горная промышленность*. 2024;(6):178–185. <https://doi.org/10.30686/1609-9192-2024-6-178-185>
Panina O.V., Popadyuk N.K., Eremin S.G., Tokmurzin T.M., Razumova E.V. Application of the BigData technologies to optimize production processes in the Russian mining industry: analysis of implementation and efficiency. *Russian Mining Industry*. 2024;(6):178–185. (In Russ.) <https://doi.org/10.30686/1609-9192-2024-6-178-185>
18. Косоруков О.А., Мищенко А.В., Свиридова О.А., Цурков В.И. Динамические модели управления транспортными ресурсами. *Известия Российской академии наук. Теория и системы управления*. 2025;(6):108–129.
Kosorukov O.A., Mishchenko A.V., Sviridova O.A., Tsurkov V.I. Dynamic models of transport resources management. *Izvestiya Rossiiskoi Akademii Nauk. Teoriya i Sistemy Upravleniya*. 2025;(6):108–129. (In Russ.)
19. Yeganejou M., Badiozamani M., Moradi-Afrapoli A., Askari-Nasab H. Integration of simulation and dispatch modelling to predict fleet productivity: an open-pit mining case. *Mining Technology: Transactions of the Institutions of Mining and Metallurgy*. 2022;131(2):67–79. <https://doi.org/10.1080/25726668.2021.2001255>
20. Караев А.К., Борисова О.В. Перспективные модели финансового прогнозирования доходов бюджета. *Финансы: теория и практика*. 2025;29(1):20–33. <https://doi.org/10.26794/2587-5671-2025-29-1-20-33>
Karaev A.K., Borisova O.V. Prospective models of financial forecasting of budget revenues. *Finance: Theory and Practice*. 2025;29(1):20–33. <https://doi.org/10.26794/2587-5671-2025-29-1-20-33>
21. Moradi Afrapoli A., Askari-Nasab H. Mining fleet management systems: a review of models and algorithms. *International Journal of Mining, Reclamation and Environment*. 2019;33(1):42–60. <https://doi.org/10.1080/17480930.2017.1336607>
22. Рождественская И.А., Беляев А.М., Лукичев К.Е., Зубенко А.В., Лаффах А.М. Разработка интеллектуальных распределённых систем хранения и анализа данных для оптимизации горного производства и управления угольной добычей. *Горная промышленность*. 2025;(2):56–64. <https://doi.org/10.30686/1609-9192-2025-2-56-64>
Rozhdestvenskaya I.A., Belyaev A.M., Lukichev K.E., Zubenko A.V., Laffakh A.M. Development of smart distributed data storage and analysis systems for optimization of mining operations and coal mining management. *Russian Mining Industry*. 2025;(2):56–64. (In Russ.) <https://doi.org/10.30686/1609-9192-2025-2-56-64>

Информация об авторах

Шедько Юрий Николаевич – доктор экономических наук, профессор кафедры государственного и муниципального управления, Финансовый университет при Правительстве Российской Федерации, г. Москва, Российская Федерация; <https://orcid.org/0000-0002-9179-3637>; e-mail: ynshedko@mail.ru

Харченко Константин Владимирович – кандидат социологических наук, доцент кафедры государственного и муниципального управления, Финансовый университет при Правительстве Российской Федерации; г. Москва, Российская Федерация; <https://orcid.org/0000-0003-3329-7755>; e-mail: KVKharchenko@fa.ru

Зуденкова Светлана Александровна – кандидат экономических наук, доцент кафедры государственного и муниципального управления, Финансовый университет при Правительстве Российской Федерации, г. Москва, Российская Федерация; <https://orcid.org/0000-0002-6470-5451>; e-mail: SAZudenkova@fa.ru

Галкин Андрей Игоревич – кандидат экономических наук, доцент кафедры государственного и муниципального управления, Финансовый университет при Правительстве Российской Федерации; г. Москва, Российская Федерация; <https://orcid.org/0000-0003-0021-7536>; e-mail: aigalkin@fa.ru

Бабаян Левон Каренович – кандидат экономических наук, ассистент кафедры государственного и муниципального управления, Финансовый университет при Правительстве Российской Федерации, г. Москва, Российская Федерация; <https://orcid.org/0000-0001-6872-8549>; e-mail: LKBabayan@fa.ru

Информация о статье

Поступила в редакцию: 19.10.2025

Поступила после рецензирования: 16.12.2025

Принята к публикации: 20.01.2026

Information about the authors

Yuri N. Shedko – Dr. Sci. (Econ.), Professor of the Department of State and Municipal Administration of the Financial University under the Government of the Russian Federation, Moscow, Russian Federation; <https://orcid.org/0000-0002-9179-3637>; e-mail: ynshedko@mail.ru

Konstantin V. Kharchenko – Cand. Sci. (Sociol.), Associate Professor of the Department of State and Municipal Administration of the Financial University under the Government of the Russian Federation, Moscow, Russian Federation; <https://orcid.org/0000-0003-3329-7755>; e-mail: KVKharchenko@fa.ru

Svetlana A. Zudenkova – Cand. Sci. (Econ.), Associate Professor of the Department of State and Municipal Administration of the Financial University under the Government of the Russian Federation, Moscow, Russian Federation; <https://orcid.org/0000-0002-6470-5451>; e-mail: SAZudenkova@fa.ru

Andrey I. Galkin – Cand. Sci. (Econ.), Associate Professor of the Department of State and Municipal Administration of the Financial University under the Government of the Russian Federation, Moscow, Russian Federation; <https://orcid.org/0000-0003-0021-7536>; e-mail: aigalkin@fa.ru

Levon K. Babayan – Cand. Sci. (Econ.), Assistant of the Department of State and Municipal Administration of the Financial University under the Government of the Russian Federation, Moscow, Russian Federation; <https://orcid.org/0000-0001-6872-8549>; e-mail: LKBabayan@fa.ru

Article info

Received: 19.10.2025

Revised: 16.12.2025

Accepted: 20.01.2026